

Byonic™: Modification Fine Control™

Marshall Bern, Ph.D., Protein Metrics Inc. (www.proteinmetrics.com)

October 2014

Benefits summary

- Understand how to use Byonic's Modification Fine Control™ to structure efficient, sensitive, comprehensive searches.

Method

Byonic gives the user detailed control of the types and numbers of posttranslational modifications allowed on candidate peptides. Modification Fine Control™ enables Byonic to focus its search on the candidate peptides most likely to match mass spectra, so that with the right parameter settings, Byonic can find more true matches more quickly than any other proteomics search engine. This application note explains how to choose the right parameter settings. (This application note explains ordinary modifications. Glycan and wildcard modifications are special cases, explained in other [application notes](#).)

We illustrate Modification Fine Control with the example shown in Figure 1. In this example, **Total common max** is set to 2 and **Total rare max** (just out of view) is set to 1. Oxidized methionine is common2 so Byonic will consider a candidate peptide with two M[+15.99]'s but not one with three M[+15.99]'s. Q[+0.98] is common1, so Byonic will not consider a peptide with two Q[+0.98]'s. Byonic will not consider a peptide with one M[+15.99], one N[+0.98], and one Q[+0.98], because **Total common max** sets a limit of two common modifications per peptide. Byonic will, however, consider a peptide with N-terminal Q[-17.03], one M[+15.99], and one N[+0.98], because pyro-Glu Q is charged against **Total rare max** rather than **Total common max**.

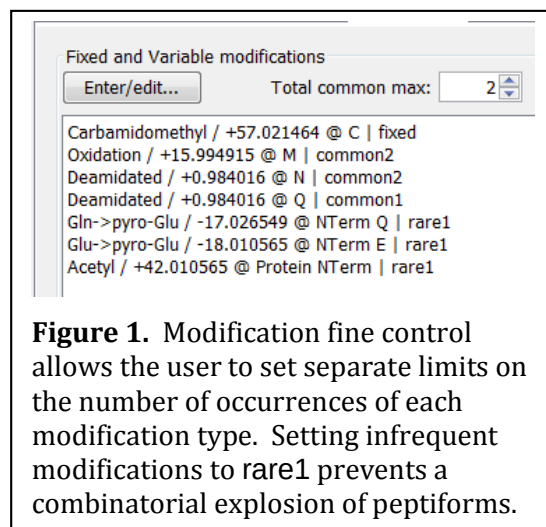


Figure 1. Modification fine control allows the user to set separate limits on the number of occurrences of each modification type. Setting infrequent modifications to rare1 prevents a combinatorial explosion of peptidforms.

The standard strategy for Modification Fine Control is to make the frequent modifications **common** and the infrequent modifications **rare**:

1. Set the most frequent modification types (for example, oxidized methionine and deamidated asparagine) to **common2** or **common3**. Set the modification types that cluster together (for example, phosphorylated serine) to **common2**, **common3**, or **common4**. Set **Total common max** to 2, 3, or 4, depending upon the overall modification level and search time budget. The common modifications are now

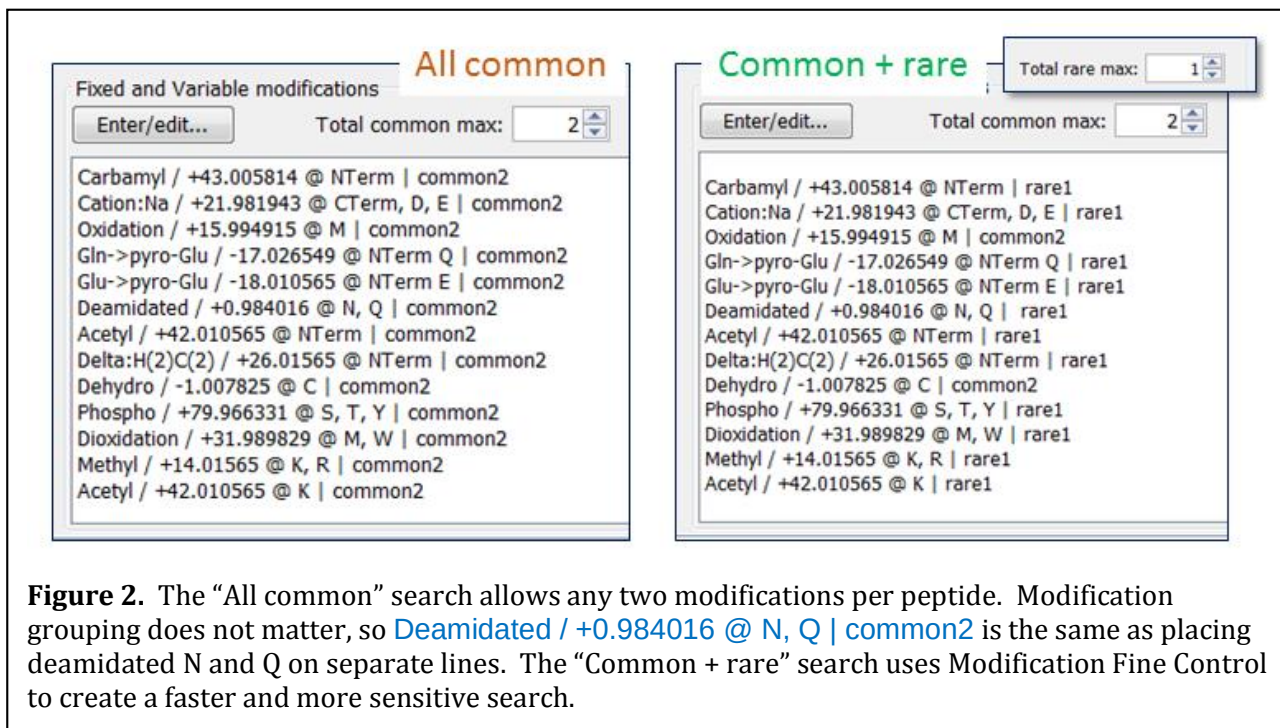
“fully enabled” much as they would be with a conventional search engine. We rarely use more than 7 or 8 fully enabled modification types. We never use **Total common max** larger than 4 when searching against a large database. Search time grows exponentially with **Total common max**, so a search with **Total common max** 4 may take 100s of times longer than a search with **Total common max** 2.

2. Set the infrequent modification types (for example, doubly oxidized tryptophan) to **rare1**. So long as **Total rare max** is kept to 1, the number of rare1 modifications can grow quite large (to 10s or even 100s) without the search time exploding. We usually make N-terminal modifications (pyro-Glu, N-terminal acetylation, over-alkylation) **rare1** as well, even though they may be quite frequent; this way they do not compete with common modifications.

One subtlety: The standard strategy only makes sense when **Total common max** exceeds **Total rare max**. If you set both **Total common max** and **Total rare max** to 1 in order to simplify the search and make it run faster, then you should divide the most frequent modifications evenly between the common and rare categories. You might even set the most frequent modification to be both common and rare in order to allow two per peptide.

Example

Figure 2 shows two different modification settings for a tryptic digest of Jurkat cell lysate run with Orbitrap MS1, low-energy CID fragmentation, and ion-trap MS2. Cysteines were not reduced or alkylated, so we enabled **Dehydro, common2** to find peptides with a single internal disulfide bond. “All common” shows a search without Modification Fine Control: all modifications are treated equally. “Common + rare” shows a search using the standard strategy described above. Table 1 gives results of the two searches. These results are typical for routine proteomics, for which Modification Fine Control offers small improvements in sensitivity and specificity but significant improvement in running time. However, for searches with 10s or 100s of modification types (for example, oxidative footprinting), Modification Fine Control is essential.



Search	# PSMs	PSM FDR	# Unique Peptides	# Proteins	Running Time
All common	7278	0.1%	6098	589	17 min 35 sec
Common + rare	7363	0.1%	6158	598	6 min 1 sec

Table 1. Byonic results for the two searches.