PROTEIN METRICS

Project Metrics Application Note

Release 1.0

Monika Bhatt

Feb 08, 2018

CONTENTS

SEQUENCE VARIANT ANALYSIS (SVA) PART I: BYONIC™ SEARCHES

Marshall Bern, Ph.D., K. Ilker Sen, Ph.D., Eric Carlson, Ph.D., and Chris Becker, Ph.D. Protein Metrics Inc. www.proteinmetrics.com June 2016

1.1 Benefits summary

- Best practices for running a sequence variant search

- How-to for using Byonic™ software to effectively search for sequence variants, for subsequent inspection and quantification by Byologic® software.

1.2 Method

Almost all therapeutic protein data sets include mass spectra from peptides that differ from the corresponding database peptide by one amino acid substitution at some concentration. Sequence variants may arise from mutations or mistranslations in the host cell line or problems with the bioreactor feed. Searching for sequence variants is especially challenging, because there are a large number of possible amino acid substitutions, and many of the mass deltas are either exactly or approximately the same as the mass deltas from posttranslational modifications and sample preparation artifacts. In this application note, we show how to search for sequence variants in biotherapeutics. Sequence variant analysis aims to identify and quantify all the variant peptides down to very small percentages of their "wild type" (native or expected sequence) counterparts. Because the protein database is very small for a therapeutic protein, typically only the product plus trypsin or other digestive enzyme and decoys, Byonic running time is not a concern. It is feasible to search for all possible one-AA-substitution peptides in a few minutes. The challenge is distinguishing the relatively rare true sequence variants from false positives. This Application Note focuses on the first step, the Byonic search for SVA.

Step	Detail
Choice of alkylating agent	Many amino acid substitutions have the exact same mass difference as the common alkylating agents. We recommend choosing one based on the desired analysis: • Ideally, use C13 labeled Iodoacetic acid. • Iodoacetic acid tends to generate fewer over-alkylated peptides than iodoacetamide, and has fewer delta mass overlaps with all possible amino acid substitutions list. • If, however, only single base substitutions that lead to sequence variants will be considered, rather than the entire list, carbamidomethyl may be the desired agent as there are no conflicting masses within this shorter list.
Use a high resolution mass spectrometer	Use high resolution MS1. Ideally, also use high resolution MS2. Make sure the instrument is well calibrated and operating near peak performance.
MSMS Search: Modification rules settings.	Identify artifacts first with Protein Metrics' Preview™ or Byonic Wildcard Search™. • Do not include deamidation in the list of variants (Asn->Asp). Instead include them as named rare1 modifications. Include all the types of non-variant modifications known to be present in your sample, including all the sample preparation artifacts particular to your laboratory. Use rare1 for all modifications and sequence variants except common1 for pyro-glu modifications. • Use common max 1 and rare max 1. • See example in Table 2 for a typical modification list. Have the sequence variants in the search as rare1 modifications (see Figure 1 below for the full list, and an exemplary short list in Table 3). • A list of modifications for sequence variants is available from Protein Metrics. This list can be copied from a text editor and pasted into the Modification pop-up box in Byonic. • In addition, for mAbs with N-glycosylation, include the table of 50 N-glycans on the Glycan tab, and check Show all N-glycopeptides on the Advanced tab. Use common1 setting.

1.2.1 Discussion regarding the use of isotopically labeled iodoacetic acid or iodoacetimide as alkylating agents.

We suggest using iodoacetic acid for alkylation of Cysteines because this typically produces fewer sample preparation artifacts compared to iodoacetimide. In either case however, there are a significant number of confusable masses of sequence variants with alkylation artifacts. Table 4 below lists confusable masses for unlabeled and C13-labeled alkylations and sequence variants. Examination of the table shows that for commercially available labeled iodoacetic acid and iodoacetimide, that the single-C13-labeled iodoacetic acid will give the best separation and hence the fewest (or none) confusion with sequence variants. Note that if a three-C13-labeled iodoacetic acid is available, the +61 modification has no confusable masses with sequence variants.

For example, the `Settings >> Configuration >> Device Groups` allows the user to add new device groups, read the available device groups, update the device groups and delete device groups, on a regular LogPoint.

Fig. 1.1: Device Groups on a regular LogPoint

However, when the LogPoint becomes Fabric-enabled, the user can only view the available device groups.

Fig. 1.2: Device Groups on a Fabric-enabled LogPoint

For add operation, the user needs to use the **DeviceGroups - Create API**. For update operation, the **DeviceGroups - Edit API** is used while the delete operation is carried out via the use of **DeviceGroups - Trash API**.

Hence, the create, update, and delete operations of certain functionalities of a Fabric-enabled LogPoint need to be carried out via their corresponding API exposed applications.

Note: A limited number of features exposed by the API server are available in the user interface of the Director Console. Please refer to the **Director Console Manual** for details on the list of features that can be carried out via its user interface.

The list of applications and features that are implemented in Director Console API are listed below:

1. Configuration

   - Devices
   - Device Groups
   - Distributed Collectors
   - Normalization Policy
   - Repos
   - Distributed LogPoints
   - Parsers
   - Enrichment Policy
   - Enrichment Source
   - Processing Policy
   - Routing Policies
   - SNMP Policy

2. Knowledge Base

   - Normalization Package

3. System

   - System Settings Enrichment

   - System Settings General

   - System Settings NTP

   - System Settings SMTP

   - System Settings SNMP

   - System Settings SSH

   - System Settings Support Connection

   - Upload

   - Open Door

   - Backup and Restore

   - Snapshots

   - Plugins

4. Built-in Collectors and Fetchers

   - File System Collector

   - FTP Collector

   - FTP Fetcher

   - SCP Fetcher

   - SFlow Collector

   - Snare Collector

   - SNMP Fetcher

   - SNMP Trap Collector

   - Syslog Collector

   - WMI Fetcher

Refer to the Director Console API Documentation in the LogPoint Help Center for further details.

Note:

- The Devices - Attach/Detach feature is used in the Distributed LogPoint setup. It is used to attach/detach devices on behalf of the collector LogPoint, from the main LogPoint.

- The Backup and Restore - CreateConfigBackup feature is used to backup the configuration details of Fabric-enabled LogPoint and all of its collectors/fetchers.

- The Backup and Restore - CreateLogsChecksumBackup feature is used to back up the logs, index of the logs and the checksum of the logs of the Fabric-enabled LogPoint.

- The Upload - Upload feature is used to upload applications and patches to the Fabric Storage.

- The Upload - Install feature is used to install the applications and patches already uploaded in the Fabric Storage, to the Fabric-enabled LogPoint.

- The Snapshots feature is supported only in the LogPoint v6.0.0 ISO.

1.3 Different Types of Audit Logs

1.3.1 Audit Logs for Fabric Server

LogPoint generates and stores audit logs of all actions carried out in the Fabric Server. With these audit logs, every action in Fabric Server, such as what was changed, who changed it and when those changes were made etc., can be traced.