Reliable LC-MS Multi-Attribute Method for Biotherapeutics by Run-Time Response Calibration

Zhongqi Zhang*, Bhavana Shah and Xiaoyan Guan

Process Development, Amgen Inc., One Amgen Center Drive, Thousand Oaks, California 91320

*E-mail: zzhang@amgen.com

Supporting Information

Table of Content

I. Principle of the Method

Notations. Notations used in this article are summarized as follows: *A* represents the abundance of an isoform as a fractional value (or percentage). *I* represents the measured ion intensity (area under the peak in the selected-ion chromatogram) of an isoform. *k* represents the response factor, and *a* represents the correction factor for the response factor. Superscript 0 represents the reference standard (RS). For example, A^0 represents the known abundance of an isoform in RS, and I^0 represents the measured ion intensity of an isoform in RS. Subscripts 0, 1, 2, … *i*, … *n* represent the *n*+1 isoforms associated with a specified residue (including the unmodified form). Subscript 0 represents the most abundant isoform (usually the unmodified form). For example, A_i represents the abundance of isoform *i*, I_i represents the ion intensity of isoform *i*, k_i represents the response factor of isoform *i*, and a_i represents the response factor correction factor for isoform *i*. The value of a_0 is defined as 1. The subscript 1 may be omitted in certain equations when there is only one isoform present other than the major isoform (*n* = 1). Note when isoforms *i* = 0, 1, … *n* are mentioned in this report, they are usually associated with a single residue, regardless of whether this is explicitly stated or not. For example, there can be three isoforms associated with an Asn residue in a peptide, including the unmodified Asn, its deamidated form, and its succinimide form. An oxidized Met residue on the same peptide, however, belong to a different set of isoforms (assuming the two modifications are stochastically distributed and therefore independent of each other).

Conventional MAM. Assume an amino acid residue has *n*+1 modification states (0, 1, …, *n*). The most abundant state, usually the unmodified, is denoted as *i* = 0. There are a total of *n*+1 peptide isoforms related to the residue of interest. The abundance of each isoform is calculated based on eq 1.

$$\begin{cases} kA_i = I_i, i = 0,...,n \\ \sum_{j=0}^{n} A_j = 1 \end{cases} \quad (1)$$

where *k* is the response factor for each of the isoforms of the peptide (a constant value based on the conventional assumption), A_i and I_i are the abundance and MS intensity of isoform *i*, respectively.

Solving eq 1 yields,

$$\begin{cases} kA_i = I_i \\ \sum_{j=0}^{n} A_j = 1 \end{cases} \rightarrow \begin{cases} A_i = \frac{I_i}{k} \\ \frac{1}{k}\sum_{j=0}^{n} I_j = 1 \end{cases} \rightarrow \begin{cases} A_i = \frac{I_i}{k} \\ k = \sum_{j=0}^{n} I_j \end{cases} \rightarrow \begin{cases} k = \sum_{j=0}^{n} I_j \\ A_i = \frac{I_i}{\sum_{j=0}^{n} I_j}, i = 0,...,n \end{cases}$$

The response factor *k*, which is the sum of intensities of all isoforms, represents a combination of the both peptide recovery and the MS response.

Calibrating Response Factors Using the Reference Standard. Let's say the response factor for each isoform *i* is modified by a factor a_i, then we express the response factor for isoform *i* in the RS as $a_i k^0$ (superscript of 0 stands for RS), and the response factor for isoform *i* in the sample as $a_i k$. After considering both RS and the sample, eq 1 becomes

$$\begin{cases} a_i k^0 A_i^0 = I_i^0, i = 0,...,n \\ a_i k A_i = I_i, i = 0,...,n \\ a_0 = 1 \\ \sum_{j=0}^{n} A_j = 1 \end{cases} \quad (2)$$

Note that the response factor for the RS and the sample can be different because of the slight difference in sample preparation, instrument sensitivity, etc. Additionally, for the equations to be solvable, factor *a* for one of the isoforms must be set to a constant value. It is a good idea to set *a* as 1 for the most abundant isoform, which is usually the unmodified form ($a_0 = 1$).

Solving eq 2 yields,

$$\begin{cases} a_i k^0 A_i^0 = I_i^0 \\ a_i k A_i = I_i \\ a_0 = 1 \\ \sum_{j=0}^{n} A_j = 1 \end{cases} \rightarrow \begin{cases} \frac{a_i k^0 A_i^0}{a_0 k^0 A_0^0} = \frac{I_i^0}{I_0^0} \\ a_i k A_i = I_i \\ a_0 = 1 \\ \sum_{j=0}^{n} \frac{I_j}{a_j k} = 1 \end{cases} \rightarrow \begin{cases} \frac{a_i A_i^0}{A_0^0} = \frac{I_i^0}{I_0^0} \\ a_i k A_i = I_i \\ \\ \frac{1}{k} \sum_{j=0}^{n} \frac{I_j}{a_j} = 1 \end{cases} \rightarrow \begin{cases} a_i = \frac{I_i^0}{I_0^0} \frac{A_0^0}{A_i^0} \\ a_i k A_i = I_i \\ \\ k = \sum_{j=0}^{n} \frac{I_j}{a_j} \end{cases} \rightarrow \begin{cases} a_i = \frac{I_i^0}{I_0^0} \frac{A_0^0}{A_i^0} \\ A_i = \frac{I_i}{a_i k} \\ \\ k = \sum_{j=0}^{n} \frac{I_j}{a_j} \end{cases}$$

$$\rightarrow \begin{cases} a_i = \frac{I_i^0}{I_0^0} \frac{A_0^0}{A_i^0}, i = 0,...,n \\ A_i = \frac{I_i / a_i}{\sum_{j=0}^{n} \frac{I_j}{a_j}}, i = 0,...,n \end{cases}$$

Eq 2 can be used to perform response factor calibration (a-calibration) based on known attribute abundance in the RS (A_i^0), and to calculate the abundance of each isoform in the sample (A_i).

Calibrating Artificial Modifications Using the Reference Standard. Let's use a factor *b* to denote the extent of artificial change for each attribute. After considering both RS and the sample, eq 1 becomes

$$\begin{cases} k^0[A_i^0 + b_i S_i^0] = I_i^0, i = 0,...,n \\ k[A_i + b_i S_i] = I_i, i = 0,...,n \\ \sum_{j=0}^{n} A_j = 1 \end{cases} \quad (3)$$

In eq 3, S_i represents the substrate of the artificial change of isoform *i*. For example, for oxidation and deamidation (type-3 attributes), the substrate is the unmodified peptide. For succinimide (a type-2 attribute), the substrate is the modified peptide.

Depending on the nature of the substrate, eq 3 may become complex. For example, using eq 3 for multiple modifications on the same residue, such as N-glycosylation on asparagine residues, may create a complex variant of eq 3. It is therefore not recommended to use eq 3 for multiple modifications on the same residue. For residues with single modification, eq 3 is reduced to eqs 4 and 5, depending on whether the substrate is the modified peptide (type-2 attributes) or unmodified peptide (type-3 attributes).

For type-2 attributes

$$\begin{cases} k^0(1 - A^0 - bA^0) = I_0^0 \\ k^0(A^0 + bA^0) = I^0 \\ k(1 - A - bA) = I_0 \\ k(A + bA) = I \end{cases} \quad (4)$$

Solving eq 4 yields

$$\begin{cases} k^0(1 - A^0 - bA^0) = I_0^0 \\ k^0(A^0 + bA^0) = I^0 \\ k(1 - A - bA) = I_0 \\ k(A + bA) = I \end{cases} \rightarrow \begin{cases} k^0[1 - (1+b)A^0] = I_0^0 \\ k^0(1+b)A^0 = I^0 \\ k[1 - (1+b)A] = I_0 \\ k(1+b)A = I \end{cases} \rightarrow \begin{cases} k^0 - I^0 = I_0^0 \\ k^0(1+b)A^0 = I^0 \\ k - I = I_0 \\ k(1+b)A = I \end{cases} \rightarrow \begin{cases} k^0 = I_0^0 + I^0 \\ k^0(1+b)A^0 = I^0 \\ k = I_0 + I \\ k(1+b)A = I \end{cases}$$

$$\rightarrow \begin{cases} k^0 = I_0^0 + I^0 \\ k = I_0 + I \\ \dfrac{k(1+b)A}{k^0(1+b)A^0} = \dfrac{I}{I^0} \end{cases} \rightarrow \begin{cases} k^0 = I_0^0 + I^0 \\ k = I_0 + I \\ A = \dfrac{Ik^0}{I^0 k} A^0 \end{cases} \rightarrow A = \frac{I(I_0^0 + I^0)}{I^0(I_0 + I)} A^0$$

For type-3 attributes:

$$\begin{cases} k^0[1-A^0-b(1-A^0)] = I_0^0 \\ k^0[A^0+b(1-A^0)] = I^0 \\ k[1-A-b(1-A)] = I_0 \\ k[A+b(1-A)] = I \end{cases} \quad (5)$$

Solving eq 5 yields

$$\begin{cases} k^0[1-A^0-b(1-A^0)] = I_0^0 \\ k^0[A^0+b(1-A^0)] = I^0 \\ k[1-A-b(1-A)] = I_0 \\ k[A+b(1-A)] = I \end{cases} \rightarrow \begin{cases} k^0[A^0+b(1-A^0)] = I^0 \\ k^0 = I_0^0 + I^0 \\ k[A+b(1-A)] = I \\ k = I_0 + I \end{cases}$$

$$\rightarrow \begin{cases} (I_0^0+I^0)[A^0+b(1-A^0)] = I^0 \\ (I_0+I)[A+b(1-A)] = I \end{cases} \rightarrow \begin{cases} b = \dfrac{1}{1-A^0}\left(\dfrac{I^0}{I_0^0+I^0} - A^0\right) \\ A = \dfrac{1}{1-b}\left(\dfrac{I}{I_0+I} - b\right) \end{cases}$$

Calibrating Both Response Factors and Artificial Modifications Using Two Different Standards. The RS can be used to correct either the response factor (*a*) or artificial modification (*b*); it cannot be used to correct for both. To correct for both *a* and *b*, an additional standard is needed. To get a different standard, the RS or another sample can be stressed to create a second standard containing higher level of the attributes of interest, and then both standards are analyzed together with the samples. Known attribute abundance of the two standards and their determined MS responses will be used to correct for both *a* and *b*.

Let's use a superscript of 0 to represent the RS and a superscript of 1 to represent the stressed standard. After considering both standards as well as the sample, the corresponding mathematics are shown in eqs 7 and 8, depending on the substrate of the artificial modification. Note that eqs 7 and 8 apply only to residues with a single modification ($n = 1$).

For type-2 attributes

$$\begin{cases} k^0[1-A^0-bA^0]=I_0^0 \\ ak^0[A^0+bA^0]=I^0 \\ k^1[1-A^1-bA^1]=I_0^1 \\ ak^1[A^1+bA^1]=I^1 \\ k[1-A-bA]=I_0 \\ ak[A+bA]=I \end{cases} \qquad \textbf{(7)}$$

Solving eq 7 yields

$$\begin{cases} k^0[1-A^0-bA^0]=I_0^0 \\ ak^0[A^0+bA^0]=I^0 \\ k^1[1-A^1-bA^1]=I_0^1 \\ ak^1[A^1+bA^1]=I^1 \\ k[1-A-bA]=I_0 \\ ak[A+bA]=I \end{cases} \rightarrow \begin{cases} ak^0[1-A^0-bA^0]=aI_0^0 \\ ak^0[A^0+bA^0]=I^0 \\ ak^1[1-A^1-bA^1]=aI_0^1 \\ ak^1[A^1+bA^1]=I^1 \\ ak[1-A-bA]=aI_0 \\ ak[A+bA]=I \end{cases} \rightarrow \begin{cases} ak^0=aI_0^0+I^0 \\ ak^0[A^0+bA^0]=I^0 \\ ak^1=aI_0^1+I^1 \\ ak^1[A^1+bA^1]=I^1 \\ ak=aI_0+I \\ ak[A+bA]=I \end{cases}$$

$$\rightarrow \begin{cases} (aI_0^0+I^0)[A^0+bA^0]=I^0 \\ (aI_0^1+I^1)[A^1+bA^1]=I^1 \\ (aI_0+I)[A+bA]=I \end{cases} \rightarrow \begin{cases} (aI_0^0+I^0)[A^0+bA^0]=I^0 \\ (aI_0^1+I^1)[A^1+bA^1]=I^1 \\ (aI_0+I)[A+bA]=I \end{cases} \rightarrow \begin{cases} \dfrac{I^0}{(aI_0^0+I^0)A^0}=\dfrac{I^1}{(aI_0^1+I^1)A^1} \\ (aI_0^0+I^0)[A^0+bA^0]=I^0 \\ (aI_0+I)[A+bA]=I \end{cases}$$

$$\rightarrow \begin{cases} (aI_0^0+I^0)A^0I^1=(aI_0^1+I^1)A^1I^0 \\ (aI_0^0+I^0)[A^0+bA^0]=I^0 \\ (aI_0+I)[A+bA]=I \end{cases} \rightarrow \begin{cases} aI_0^0A^0I^1+I^0A^0I^1=aI_0^1A^1I^0+I^1A^1I^0 \\ (aI_0^0+I^0)[A^0+bA^0]=I^0 \\ (aI_0+I)[A+bA]=I \end{cases}$$

$$\rightarrow \begin{cases} a=\dfrac{I^1A^1I^0-I^0A^0I^1}{I_0^0A^0I^1-I_0^1A^1I^0} \\ (aI_0^0+I^0)[A^0+bA^0]=I^0 \\ (aI_0+I)[A+bA]=I \end{cases} \rightarrow \begin{cases} a=\dfrac{I^0I^1(A^1-A^0)}{I_0^0I^1A^0-I^0I_0^1A^1} \\ b=\dfrac{I^0}{A^0(aI_0^0+I^0)}-1 \\ A=\dfrac{I}{(aI_0+I)(b+1)} \end{cases} \rightarrow \begin{cases} a=\dfrac{I^0I^1(A^1-A^0)}{I_0^0I^1A^0-I^0I_0^1A^1} \\ b=\dfrac{I^0}{A^0(aI_0^0+I^0)}-1 \\ A=\dfrac{I(aI_0^0+I^0)}{I^0(aI_0+I)}A^0 \end{cases}$$

For type-3 attributes:

$$\begin{cases} k^0[1-A^0-b(1-A^0)]=I_0^0 \\ ak^0[A^0+b(1-A^0)]=I^0 \\ k^1[1-A^1-b(1-A^1)]=I_0^1 \\ ak^1[A^1+b(1-A^1)]=I^1 \\ k[1-A-b(1-A)]=I_0 \\ ak[A+b(1-A)]=I \end{cases} \quad \textbf{(8)}$$

Solving eq 8 yields

$$\begin{cases} k^0[1-A^0-b(1-A^0)]=I_0^0 \\ ak^0[A^0+b(1-A^0)]=I^0 \\ k^1[1-A^1-b(1-A^1)]=I_0^1 \\ ak^1[A^1+b(1-A^1)]=I^1 \\ k[1-A-b(1-A)]=I_0 \\ ak[A+b(1-A)]=I \end{cases} \rightarrow \begin{cases} ak^0[1-A^0-b(1-A^0)]=aI_0^0 \\ ak^0[A^0+b(1-A^0)]=I^0 \\ ak^1[1-A^1-b(1-A^1)]=aI_0^1 \\ ak^1[A^1+b(1-A^1)]=I^1 \\ ak[1-A-b(1-A)]=aI_0 \\ ak[A+b(1-A)]=I \end{cases} \rightarrow \begin{cases} ak^0=aI_0^0+I^0 \\ ak^0[A^0+b(1-A^0)]=I^0 \\ ak^1=aI_0^1+I^1 \\ ak^1[A^1+b(1-A^1)]=I^1 \\ ak=aI_0+I \\ ak[A+b(1-A)]=I \end{cases}$$

$$\rightarrow \begin{cases} (aI_0^0+I^0)[A^0+b(1-A^0)]=I^0 \\ (aI_0^1+I^1)[A^1+b(1-A^1)]=I^1 \\ (aI_0+I)[A+b(1-A)]=I \end{cases} \rightarrow \begin{cases} \dfrac{\dfrac{I^0}{aI_0^0+I^0}-A^0}{1-A^0}=\dfrac{\dfrac{I^1}{aI_0^1+I^1}-A^1}{1-A^1} \\ (aI_0^0+I^0)[A^0+b(1-A^0)]=I^0 \\ (aI_0+I)[A+b(1-A)]=I \end{cases}$$

$$\rightarrow \begin{cases} (\dfrac{I^0}{aI_0^0+I^0}-A^0)(1-A^1)=(\dfrac{I^1}{aI_0^1+I^1}-A^1)(1-A^0) \\ (aI_0^0+I^0)[A^0+b(1-A^0)]=I^0 \\ (aI_0+I)[A+b(1-A)]=I \end{cases}$$

$$\rightarrow \begin{cases} [I^0-A^0(aI_0^0+I^0)](aI_0^1+I^1)(1-A^1)=[I^1-A^1(aI_0^1+I^1)](aI_0^0+I^0)(1-A^0) \\ (aI_0^0+I^0)[A^0+b(1-A^0)]=I^0 \\ (aI_0+I)[A+b(1-A)]=I \end{cases}$$

$$\rightarrow \begin{cases} [I^0-A^0aI_0^0-A^0I^0](aI_0^1+I^1)(1-A^1)=[I^1-A^1aI_0^1-A^1I^1)](aI_0^0+I^0)(1-A^0) \\ (aI_0^0+I^0)[A^0+b(1-A^0)]=I^0 \\ (aI_0+I)[A+b(1-A)]=I \end{cases}$$

$$\rightarrow \begin{cases} [aI_0^1 I^0 - aI_0^1 A^0 aI_0^0 - aI_0^1 A^0 I^0 + I^1 I^0 - I^1 A^0 aI_0^0 - I^1 A^0 I^0](1-A^1) \\ = [aI_0^0 I^1 - aI_0^0 A^1 aI_0^1 - aI_0^0 A^1 I^1 + I^0 I^1 - I^0 A^1 aI_0^1 - I^0 A^1 I^1](1-A^0) \\ (aI_0^0 + I^0)[A^0 + b(1-A^0)] = I^0 \\ \\ (aI_0 + I)[A + b(1-A)] = I \end{cases}$$

$$\rightarrow \begin{cases} aI_0^1 I^0 - aI_0^1 A^0 aI_0^0 - aI_0^1 A^0 I^0 + I^1 I^0 - I^1 A^0 aI_0^0 - I^1 A^0 I^0 - A^1 aI_0^1 I^0 + A^1 aI_0^1 A^0 aI_0^0 + A^1 aI_0^1 A^0 I^0 - A^1 I^1 I^0 + A^1 I^1 A^0 aI_0^0 + A^1 I^1 A^0 I^0 \\ = aI_0^0 I^1 - aI_0^0 A^1 aI_0^1 - aI_0^0 A^1 I^1 + I^0 I^1 - I^0 A^1 aI_0^1 - I^0 A^1 I^1 - A^0 aI_0^0 I^1 + A^0 aI_0^0 A^1 aI_0^1 + A^0 aI_0^0 A^1 I^1 - A^0 I^0 I^1 + A^0 I^0 A^1 aI_0^1 + A^0 I^0 A^1 I^1 \\ (aI_0^0 + I^0)[A^0 + b(1-A^0)] = I^0 \\ \\ (aI_0 + I)[A + b(1-A)] = I \end{cases}$$

$$\rightarrow \begin{cases} aI_0^1 I^0 - aI_0^1 A^0 aI_0^0 - aI_0^1 A^0 I^0 = aI_0^0 I^1 - aI_0^0 A^1 aI_0^1 - aI_0^0 A^1 I^1 \\ (aI_0^0 + I^0)[A^0 + b(1-A^0)] = I^0 \\ \\ (aI_0 + I)[A + b(1-A)] = I \end{cases}$$

$$\rightarrow \begin{cases} I_0^1 I^0 - I_0^1 A^0 aI_0^0 - I_0^1 A^0 I^0 = I_0^0 I^1 - I_0^0 A^1 aI_0^1 - I_0^0 A^1 I^1 \\ (aI_0^0 + I^0)[A^0 + b(1-A^0)] = I^0 \\ \\ (aI_0 + I)[A + b(1-A)] = I \end{cases}$$

$$\rightarrow \begin{cases} I_0^0 A^1 aI_0^1 - I_0^1 A^0 aI_0^0 = I_0^0 I^1 - I_0^0 A^1 I^1 + I_0^1 A^0 I^0 - I_0^1 I^0 \\ (aI_0^0 + I^0)[A^0 + b(1-A^0)] = I^0 \\ \\ (aI_0 + I)[A + b(1-A)] = I \end{cases}$$

$$\rightarrow \begin{cases} a = \dfrac{I_0^0 I^1 - I_0^0 A^1 I^1 + I_0^1 A^0 I^0 - I_0^1 I^0}{I_0^0 A^1 I_0^1 - I_0^1 A^0 I_0^0} \\ (aI_0^0 + I^0)[A^0 + b(1-A^0)] = I^0 \\ \\ (aI_0 + I)[b + (1-b)A)] = I \end{cases}$$

$$\rightarrow \begin{cases} a = \dfrac{I_0^0 I^1 (1-A^1) - I^0 I_0^1 (1-A^0)}{I_0^0 I_0^1 (A^1 - A^0)} \\ b = \dfrac{1}{1-A^0}\left(\dfrac{I^0}{aI_0^0 + I^0} - A^0\right) \\ A = \dfrac{1}{1-b}\left(\dfrac{I}{aI_0 + I} - b\right) \end{cases}$$

Limit of Quantitation (LOQ). LOQ is usually defined as the minimum analyte concentration when the corresponding response has at least a signal-to-noise ratio of 10, where noise is defined as the standard deviation of signal variation, or 10-folds of the standard deviation in the determined analyte concentration (σ/S) when the concentration is approaching zero (ICH guideline Q2(R1)). Both definitions specify the LOQ as the minimum analyte concentration that is at least 10-times the standard deviation of measured concentration, corresponding to a 10% relative standard deviation (RSD). We therefore use the minimum attribute abundance to reach 10% RSD as a means to estimate the LOQ.

II. Experimental Section

Tryptic Digestion of Antistreptavidin IgG2. The IgG2 RS and test samples (~120 µg each) were digested with trypsin using the following procedure. First, each sample was treated with 8 mM dithiothreitol (DTT, Sigma-Aldrich, St. Louise, MO) at 25°C for 30 min under a denaturing solution containing 6.5 M guanidine hydrochloride (Macron Fine Chemicals, Stroudsburg, PA) and 0.2 M Tris (TEKnova, Hollister, CA) at pH 7.5 to reduce the disulfide bonds. The reduced IgG2 were then alkylated with 14 mM iodoacetic acid (Sigma-Aldrich) at 25°C for 20 minutes in dark. Alkylation was quenched with 6 mM DTT.

To intentionally create some difference in the sample preparation procedure, each reduce/alkylated sample was digested with trypsin using two different methods. In the first method, the reduced/alkylated sample (~ 1.2 mg/mL IgG2 concentration) was exchanged into the digestion buffer containing 0.1 M tris and 50 mM methionine (pH 7.5) using a Bio-Rad (Hercules, CA) Bio-Spin 6 column according to the manufacturer's recommended procedure. After buffer exchange, appropriate amount of trypsin (Sigma-Aldrich) was added to achieve an enzyme:substrate ratio of 1:12, followed by incubation at 37°C for 60 min. Digestion was quenched using equal volume of 0.25 M acetate buffer (pH 4.8) in 8 M guanidine hydrochloride. Final IgG2 concentration in the digest was ~0.5 mg/mL.

In the second method, each reduced/alkylated sample was exchanged into the same digestion buffer using a Microcon-30kDa filter (Millipore Sigma, Burlington, MA). First, each reduced/alkylated sample was spun down at 14000 g, and the flow-through was discarded. Then, the process was repeated three more times after adding 250 µL of digestion buffer each time to the filter. Trypsin digestion was carried out on the same filter by adding 140 µL digestion buffer and 10 µg trypsin (at 1 mg/mL), followed by incubating at 37°C for 60 min. After digestion equal volume of quench solution was added to the filter and was spun down at 14000 g to collect peptides in a new receiving tube. Final IgG2 concentration in the digest was ~0.4 mg/mL.

Tryptic Digestion of the Fc-Fusion Protein. The Fc-fusion protein was first denatured in a solution containing 7.5 M guanidine HCl, 250 mM tris (pH 7.5), and 2 mM EDTA, at a protein concentration of 1 mg/mL. Prior to proteolytic digestion, 2 µl of 500 mM DTT solution was added to 100 µL of denatured protein solution, followed by incubation at 25°C for 30 min to reduce the disulfide bonds. Then, 4 µL of 500 mM sodium iodoacetate was added, followed by incubation at 25°C for 20 min, to alkylate the cysteine side chains. After buffer-exchanging into 100 mM tris, 50 mM methionine, pH 7.5 solution by a Bio-Rad Bio-Spin desalting column, every 50 µg of desalted sample was digested with 5 µg of trypsin at 37ºC for 30 min. To quench the digestion, 2% formic acid was added to each digest for a final acid concentration of 0.2%.

Details of the LC-MS Methods. For analyzing tryptic digests of antistreptavidin IgG2, two different LC methods were used in the three LC-MS systems (Table S1). For the first LC method (systems A and B), proteolytic peptides were eluted on a Waters Acquity peptide CSH column (150 × 2.1 mm, 1.7 µm particle, 170 Å pore size) at a flow rate of 0.3 mL/min with the column temperature maintained at 60 ºC. Mobile phase A was 0.1% formic acid in water, and mobile phase B was 0.1% formic acid in acetonitrile (ACN). After initial hold at 0.5% B for 5 minutes, mobile phase B linearly increased to 35% in 40 minutes. Column wash is achieved by increasing mobile phase B to 99% in 4 minutes with hold for 1 minute. Column was equilibrated with 0.5% B for 15 min. For the second LC method (system C), peptides were eluted on a Waters Acquity BEH C18 column (2.1 × 150 mm, 1.7 µm particle) at a flow rate of 0.3 mL/min with the column temperature maintained at 60 ºC. Mobile phase A was 0.1% formic acid and 0.02% trifluoroacetic acid (TFA) in water, and mobile phase B was 0.1% formic acid and 0.02% TFA in acetonitrile. After initial hold at 0.5% B for 5 minutes, phase B linearly increased to 40% in 40 minutes. Column wash is achieved by increasing phase B to 99% in 4 minutes with hold for 1 minute. Column was equilibrated with 0.5% B for 15 min. About 3 to 4 µg of each tryptic digest was injected for analysis.

The Q Exactive Plus Biopharma (System A) was set up to perform full-scan MS at a resolution of 70,000 and AGC (automatic gain control) target of 1×10^6, followed by five data-dependent higher-energy collision dissociation (HCD) MS/MS (normalized collision energy = 27%) of the most abundant ions. For the Fusion Lumos (Systems B and C), full-scan MS data was collected with a resolution of 60,000 and AGC = 4×10^5, followed by top speed data dependent collision-induced dissociation (CID) MS/MS in the ion-trap (normalized collision energy = 30%). Instrument control and data collection were accomplished by Thermo Scientific Xcalibur software.

For analyzing tryptic digests of the Fc-fusion protein, proteolytic peptides (3 µg) were eluted on an Agilent Zorbax C18 RR HD column (2.1 × 150 mm, 1.8 µm particle, 300 Å pore size) at a flow rate

of 0.25 mL/min with the column temperature maintained at 50ºC. Mobile phase A was 0.1% formic acid in water, and mobile phase B was 0.1% formic acid in acetonitrile. Beginning with 1.0%, mobile phase B linearly increased to 40% after 70 min and to 90% at 76 min. After washing at 90% for 5 min, the column was equilibrated with 1% B for 11 min.

The Q Exactive Biopharma (system D) and Exactive Plus (systems E and F) were set up to perform full-scan MS at a resolution of 140,000 and AGC target of 1 x 10^6. No MS/MS spectra were collected. Instrument control and data collection were accomplished by Thermo Scientific Chromeleon software.

Table S1. Details of the LC-MS systems used in this work.

System	*MS*	*HPLC*	*HPLC column*	*HPLC method**
Antistreptavidin IgG2				
A	Q Exactive Biopharma	Agilent 1290	Waters peptide CSH at 60°C	Additive: 0.1% formic acid Gradient: 0.5% to 35% of ACN in 40 min Flowrate: 0.3 mL/min
B	Fusion Lumos	Agilent 1290	Waters peptide CSH at 60°C	Additive: 0.1% formic acid Gradient: 0.5% to 35% of ACN in 40 min Flowrate: 0.3 mL/min
C	Fusion Lumos	Agilent 1260	Waters BEH C18 at 60°C	Additive: 0.1% formic acid + 0.02% TFA Gradient: 0.5% to 40% of ACN in 40 min Flowrate: 0.3 mL/min
Fc-fusion protein				
D	Q Exactive Biopharma	Agilent 1290	Agilent Zorbax C18 RR HD at 50°C	Additive: 0.1% formic acid Gradient: 1% to 40% of ACN in 70 min Flowrate: 0.25 mL/min
E	Exactive Plus (unit 1)	Agilent 1290		
F	Exactive Plus (unit 2)	Agilent 1290		

*ACN: acetonitrile; TFA: trifluoroacetic acid

III. Results

Antistreptavidin by Systems A and B. Using the RS for response factor calibration, the abundances of each of the 177 attributes in the three samples were calculated. For demonstration purpose, measured abundances of two attributes in sample 2 on two different instrument setups and two different sample preparation procedures are shown in Figure S1. After response factor calibration, the variations caused by difference in sample preparation and instrument setup are greatly reduced, as

indicated by the relative standard deviation (RSD) of the 12 measurements decreasing from ~50% to ~6%.

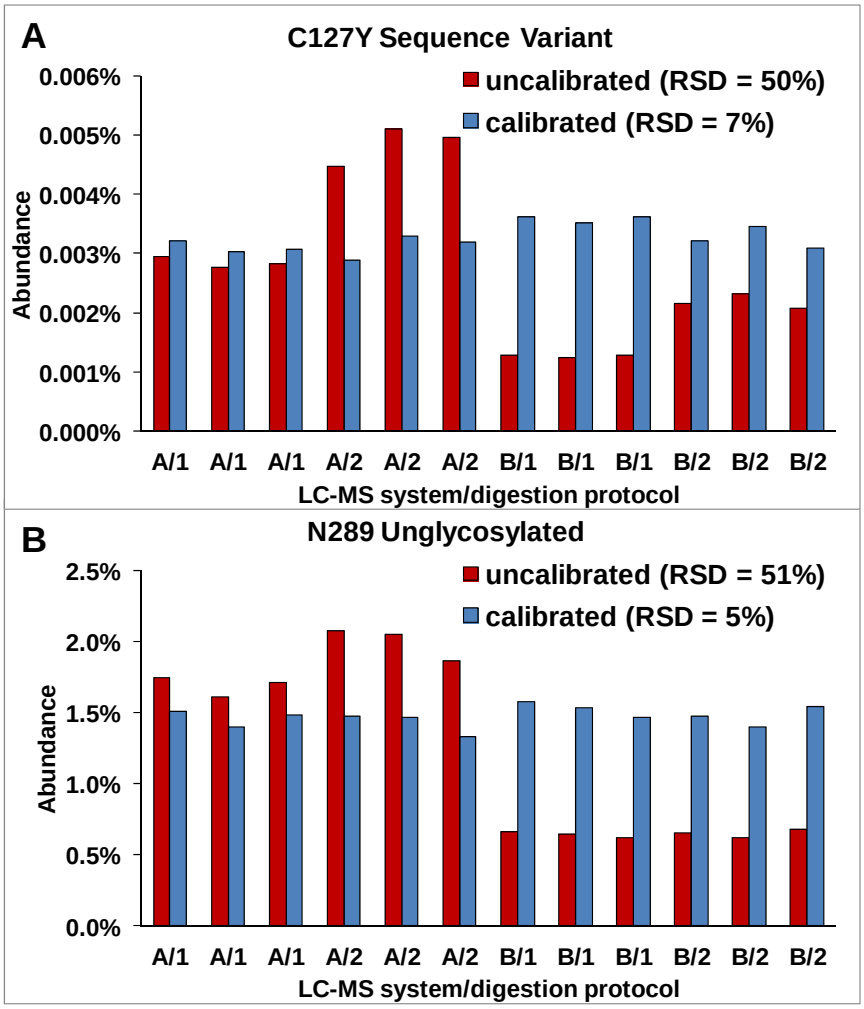

Figure S1. Measured abundance in the three replicates of sample-2, with and without response factor (a) calibration, of heavy chain Cys127Tyr sequence variant (A) and unglycosylated Asn-289 (B) by two LC-MS systems (systems A and B on the horizontal axes) and two digestion protocols (/1 and /2 on the horizontal axes). After response factor calibration, the variations between instruments and sample preparation procedures are greatly reduced, with RSD decreased by > 7 folds for the two attributes.

For type-2 attributes, variations in artificial modifications between different sample preparation conditions can be corrected using eq 4. Type-2 b-calibration was applied to the two N-terminal free glutamines on the heavy and light chains. Improvement in intermediate precision of the two type-2 attributes (2 attributes in 3 samples generated 6 measurements) after b-calibration is shown in Figure S2,

and is compared to the precision after response factor calibration (i.e., a-calibration). The two calibration methods generated very similar RSD.

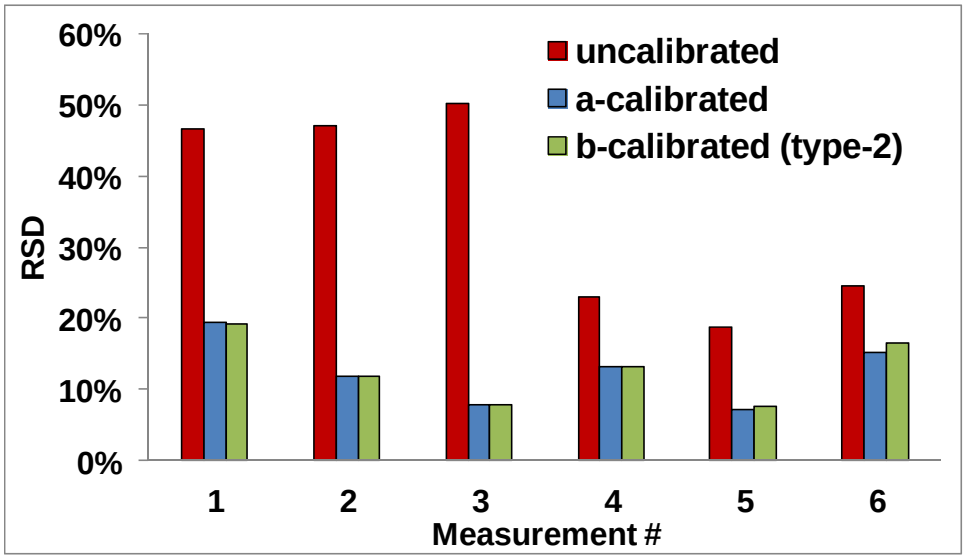

Figure S2. Calibration of artificial modification for type-2 attributes generated similar result as response-factor (a-) calibration. Measurements 1-3 are heavy-chain N-terminal free glutamine in samples 1-3 respectively, and measurements 4-6 are light-chain N-terminal free glutamine in samples 1-3, respectively.

The 177 attributes monitored in the study are listed in Table S2. All calibration calculations were performed on Microsoft Excel. Four Excel files for antistreptavidin IgG2 calibration are also included in the Supporting Information, including a-calibration, type-2 b-calibration, type-3 b-calibration, and type-3 a&b-calibration.

Antistreptavidin by Systems A and C. The 117 attributes monitored by systems A and C are listed in Table S3. The Excel file for the comparison of systems A and C is also included in the Supporting Information.

Table S2. The 177 attributes monitored by systems A and B (HC: heavy chain; LC: light chain).

Chain	Modification	Chain	Modification	Chain	Modification
HC	Q1+NH3	HC	N289+A1G1F	HC	M389I
HC	K13+Glycation	HC	N289+A1G1M4	HC	D393+H2O loss
HC	S30+C-term clip	HC	N289+A1G1M4F	HC	S400N
HC	S31+N-term clip	HC	N289+A1G1M5	HC	M420+Oxidation
HC	W34+Double Oxidation	HC	N289+A1G1M5F	HC	M420I
HC	W34+Oxidation to kynurenine	HC	N289+A1S1	HC	N426+Deamidation
HC	W37+Double Oxidation	HC	N289+A1S1F	HC	N426+NH3 loss
HC	K44+Glycation	HC	N289+A1S1M4	HC	S436N
HC	W48+Double Oxidation	HC	N289+A1S1M4F	HC	G438-58.0054 (Pro-amide)
HC	W48+Oxidation	HC	N289+A1S1M5	HC	G438+Lys
HC	S55+C-term clip	HC	N289+A1S1M5F	LC	Q1+NH3
HC	G56R	HC	N289+A2G0	LC	S2+N-term clip
HC	G56+N-term clip	HC	N289+A2G0F	LC	T5+GalNAc
HC	N59K	HC	N289+A2G1	LC	P14+Hydroxylation
HC	K65+Glycation	HC	N289+A2G1F	LC	R17K
HC	R67Q	HC	N289+A2G2	LC	G24R
HC	V68I	HC	N289+A2G2F	LC	N28K
HC	T107+GalNAc-3SG	HC	N289+A2S1G0F	LC	G30R
HC	T107+N-term clip	HC	N289+A2S1G1F	LC	G32+N-term clip
HC	K117+Hydroxylation	HC	N289+A2S2F	LC	H36Q
HC	S120+SAHexHexNAc	HC	N289+A3G1F	LC	W37+Double Oxidation
HC	V121I	HC	N289+A3G2F	LC	W37+Oxidation
HC	F122Y	HC	N289+Deamidation	LC	L48F
HC	C127+C-term clip	HC	N289+Gn	LC	G52D
HC	C127Y	HC	N289+M5	LC	N53K
HC	R129K	HC	N289+M6	LC	N53+Deamidation
HC	T131+N-term clip	HC	N289+M7	LC	S54+SAHexHexNAc
HC	S132+N-term clip	HC	N289+M8	LC	G59E
HC	S134N	HC	N289+M9	LC	E83K
HC	S134+N-term clip	HC	N289+Unglycosylated	LC	S98+N-term clip
HC	T135+N-term clip	HC	V294M	LC	L100+N-term clip
HC	A136T	HC	S296N	LC	K108+Glycation
HC	C140Y	HC	V297A	LC	V109I
HC	V142I	HC	V297I	LC	V111A
HC	W154+Double Oxidation	HC	V300I	LC	V111I
HC	W154+Oxidation	HC	H302+Double Oxidation	LC	Q114+Deamidation
HC	N155K	HC	H302Q	LC	D144+H2O loss
HC	A158+N-term clip	HC	D304+H2O loss	LC	D144G
HC	T160+C-term clip	HC	W305+Double Oxidation	LC	W154+Double Oxidation
HC	S161+N-term clip	HC	W305+Oxidation	LC	W154+Oxidation
HC	T191+N-term clip	HC	N307+Deamidation	LC	W154+Oxidation to kynurenine
HC	D199+H2O loss	HC	N307+NH3 loss	LC	K155+Glycation
HC	K238+Glycation	HC	K309+Glycation	LC	K162+Glycation
HC	K240+Glycation	HC	K318+Glycation	LC	G164E
HC	D241+H2O loss	HC	G319S	LC	V165M
HC	M244+Double Oxidation	HC	P321+Hydroxylation	LC	K177+Glycation
HC	M244+Oxidation	HC	K352+Glycation	LC	S185R
HC	M244I	HC	N353+NH3 loss	LC	S185N
HC	S259R	HC	N353K	LC	W191+Double Oxidation
HC	D262+C-term clip	HC	C359Y	LC	W191+Oxidation
HC	N268K	HC	K362+Hydroxylation	LC	K192+Glycation
HC	W269+Double Oxidation	HC	K362+Glycation	LC	S196N
HC	K280+Glycation	HC	D368N	LC	S198N
HC	N289+A1G0	HC	A370T	LC	V201I
HC	N289+A1G0F	HC	S375N	LC	V201M
HC	N289+A1G0M4	HC	N376+Deamidation	LC	H203Q
HC	N289+A1G0M5	HC	N376+NH3 loss	LC	G205D
HC	N289+A1G0M5F	HC	N382+NH3 loss	LC	S206N
HC	N289+A1G1	HC	K384+Glycation	LC	K210+Glycation

Table S3. The 117 attributes monitored by systems A and C.

Chain	Modification	Chain	Modification	Chain	Modification
HC	Q1+NH3	HC	N289+A1G1M5F	HC	N376+Deamidation
HC	K13+Glycation	HC	N289+A1S1	HC	N376+NH3 loss
HC	S30+C-term clip	HC	N289+A1S1F	HC	N382+NH3 loss
HC	S31+N-term clip	HC	N289+A1S1M4	HC	K384+Glycation
HC	W34+Double Oxidation	HC	N289+A1S1M4F	HC	M389I
HC	W34+Oxidation to kynurenine	HC	N289+A1S1M5	HC	M420+Oxidation
HC	W37+Double Oxidation	HC	N289+A1S1M5F	HC	G438-58.0054 (Pro-amide)
HC	W48+Double Oxidation	HC	N289+A2G0	HC	G438+Lys
HC	W48+Oxidation	HC	N289+A2G0F	LC	Q1+NH3
HC	S55+C-term clip	HC	N289+A2G1	LC	P14+Hydroxylation
HC	G56+N-term clip	HC	N289+A2G1F	LC	R17K
HC	K65+Glycation	HC	N289+A2G2	LC	G24R
HC	K117+Hydroxylation	HC	N289+A2G2F	LC	G30R
HC	S120+SAHexHexNAc	HC	N289+A2S1G0F	LC	W37+Double Oxidation
HC	F122Y	HC	N289+A2S1G1F	LC	W37+Oxidation
HC	C127+C-term clip	HC	N289+A2S2F	LC	N53+Deamidation
HC	S134N	HC	N289+A3G1F	LC	S54+SAHexHexNAc
HC	S134+N-term clip	HC	N289+A3G2F	LC	S98+N-term clip
HC	T135+N-term clip	HC	N289+Deamidation	LC	K108+Glycation
HC	C140Y	HC	N289+Gn	LC	V109I
HC	W154+Double Oxidation	HC	N289+M5	LC	V111I
HC	W154+Oxidation	HC	N289+M6	LC	Q114+Deamidation
HC	N155K	HC	N289+M7	LC	D144+H2O loss
HC	T160+C-term clip	HC	N289+M8	LC	W154+Double Oxidation
HC	T191+N-term clip	HC	N289+M9	LC	W154+Oxidation
HC	K238+Glycation	HC	N289+Unglycosylated	LC	W154+Oxidation to kynurenine
HC	K240+Glycation	HC	V297I	LC	K155+Glycation
HC	M244+Oxidation	HC	V300I	LC	K162+Glycation
HC	S259R	HC	D304+H2O loss	LC	G164E
HC	D262+C-term clip	HC	W305+Double Oxidation	LC	K177+Glycation
HC	K280+Glycation	HC	W305+Oxidation	LC	S185R
HC	N289+A1G0	HC	N307+Deamidation	LC	S185N
HC	N289+A1G0F	HC	N307+NH3 loss	LC	W191+Double Oxidation
HC	N289+A1G0M5	HC	K309+Glycation	LC	W191+Oxidation
HC	N289+A1G1	HC	G319S	LC	K192+Glycation
HC	N289+A1G1F	HC	P321+Hydroxylation	LC	S198N
HC	N289+A1G1M4	HC	K352+Glycation	LC	H203Q
HC	N289+A1G1M4F	HC	K362+Hydroxylation	LC	S206N
HC	N289+A1G1M5	HC	K362+Glycation	LC	K210+Glycation

IV. Discussion

Error Propagation Analysis. Response calibration, as any calibration methods, requires additional measurements of isoform intensities in the RS. These additional measurements introduce additional errors in the attribute abundance calculation. For example, when a residue has only two modification states (n = 1), in the conventional method described in eq 1, only two measurements (I_0 and I_1) are required to derive the attribute abundance A_1. During a-calibration with eq 2, however, two additional measurements (I_0^0 and I_1^0) are required to calculate A_1. Improvement in intermediate precision is only achieved if the errors caused by these additional measurements are smaller than the variation between laboratories and instruments. To simplify error analysis, we only consider a common case of two modification states (modified and unmodified, n = 1). For conventional MAM, eq 1 is written as

$$A_1 = \frac{I_1}{I_0 + I_1} \qquad (9)$$

Assuming the variance of measuring each signal intensity is independent of each other, the relative variance in determining A_1 can be estimated by the following,

$$\begin{aligned} \frac{s_1^2}{A_1^2} &= \left(\frac{1}{A_1}\frac{\partial A_1}{\partial I_0}s_{I_0}\right)^2 + \left(\frac{1}{A_1}\frac{\partial A_1}{\partial I_1}s_{I_1}\right)^2 = \left[\frac{-I_1}{A_1(I_0+I_1)^2}s_{I_0}\right]^2 + \left[\frac{I_0}{A_1(I_0+I_1)^2}s_{I_1}\right]^2 \\ &= \left[\frac{-1}{(I_0+I_1)}s_{I_0}\right]^2 + \left[\frac{I_0}{I_1(I_0+I_1)}s_{I_1}\right]^2 \\ &= \left[\frac{s_{I_0}}{I_0}\right]^2 + \left[\frac{s_{I_1}}{I_1}\right]^2, \text{ when } I_1 << I_0 \end{aligned} \qquad (10)$$

For response factor calibration (when n = 1), eq 2 is written as

$$\begin{cases} a_1 = \dfrac{I_1^0}{I_0^0}\dfrac{1-A_1^0}{A_1^0} \\ A_1 = \dfrac{I_1/a_1}{I_0 + \dfrac{I_1}{a_1}} = \dfrac{I_1}{I_0 a_1 + I_1} \end{cases} \qquad (11)$$

The relative variance in determining A_1 can be estimated by the following,

$$
\begin{aligned}
\frac{s_1^2}{A_1^2} &= \left(\frac{1}{A_1}\frac{\partial A_1}{\partial a_1}\frac{\partial a_1}{\partial I_0^0}s_{I_0^0}\right)^2 + \left(\frac{1}{A_1}\frac{\partial A_1}{\partial a_1}\frac{\partial a_1}{\partial I_1^0}s_{I_1^0}\right)^2 + \left(\frac{1}{A_1}\frac{\partial A_1}{\partial I_0}s_{I_0}\right)^2 + \left(\frac{1}{A_1}\frac{\partial A_1}{\partial I_1}s_{I_1}\right)^2 \\
&= \left[\frac{(1-A_1^0)I_1^0 I_0 I_1}{A_1 A_1^0 (I_0^0)^2 (I_0 a_1 + I_1)^2}s_{I_0^0}\right]^2 + \left[\frac{(1-A_1^0)I_0 I_1}{A_1 A_1^0 I_0^0 (I_0 a_1 + I_1)^2}s_{I_1^0}\right]^2 + \left[\frac{I_1 a_1}{A_1 (I_0 a_1 + I_1)^2}s_{I_0}\right]^2 + \left[\frac{I_0 a_1}{A_1 (I_0 a_1 + I_1)^2}s_{I_1}\right]^2 \\
&= \left[\frac{(1-A_1^0)I_1^0 I_0}{A_1^0 (I_0^0)^2 (I_0 a_1 + I_1)}s_{I_0^0}\right]^2 + \left[\frac{(1-A_1^0)I_0}{A_1^0 I_0^0 (I_0 a_1 + I_1)}s_{I_1^0}\right]^2 + \left[\frac{a_1}{(I_0 a_1 + I_1)}s_{I_0}\right]^2 + \left[\frac{I_0 a_1}{I_1 (I_0 a_1 + I_1)}s_{I_1}\right]^2 \\
&= \left[\frac{(1-A_1^0)A_1 I_1^0 I_0}{A_1^0 (I_0^0)^2 I_1}s_{I_0^0}\right]^2 + \left[\frac{(1-A_1^0)A_1 I_0}{A_1^0 I_0^0 I_1}s_{I_1^0}\right]^2 + \left[\frac{a_1}{(I_0 a_1 + I_1)}s_{I_0}\right]^2 + \left[\frac{I_0 a_1}{I_1 (I_0 a_1 + I_1)}s_{I_1}\right]^2 \\
&\approx \left[\frac{(1-A_1^0)A_1 I_1^0 I_0}{A_1^0 I_0^0 I_1}\right]^2 \left[(\frac{s_{I_0^0}}{I_0^0})^2 + (\frac{s_{I_1^0}}{I_1^0})^2\right] + \left[\frac{s_{I_0}}{I_0}\right]^2 + \left[\frac{s_{I_1}}{I_1}\right]^2, \text{ when } I_1 << I_0 a_1 \\
&= \left[\frac{s_{I_0^0}}{I_0^0}\right]^2 + \left[\frac{s_{I_1^0}}{I_1^0}\right]^2 + \left[\frac{s_{I_0}}{I_0}\right]^2 + \left[\frac{s_{I_1}}{I_1}\right]^2, \text{ because } I \text{ is proportional to } A \text{ and } A_1^0 << 1
\end{aligned}
\tag{12}
$$

Note the latter two terms in eq 12 equal the variance of conventional MAM described in eq 10, and therefore, the increase of variance caused by a-calibration can be estimated by the first two terms, which is simply the sum of relative variance of the two signal measurements in the RS. For a-calibration to be effective, this increase in variance must be smaller than the variance caused by laboratory-to-laboratory variations. This requires that the attribute in the RS having high enough abundance to be quantified with good precision (e.g., above LOQ).

For type-3 b-calibration (eq. 5), the variance in determining A is estimated by

$$
\begin{aligned}
s^2 &= \left(\frac{\partial A}{\partial b}\frac{\partial b}{\partial I_0^0}s_{I_0^0}\right)^2 + \left(\frac{\partial A}{\partial b}\frac{\partial b}{\partial I^0}s_{I^0}\right)^2 + \left(\frac{\partial A}{\partial I_0}s_{I_0}\right)^2 + \left(\frac{\partial A}{\partial I}s_I\right)^2 \\
&= \left[\frac{I^0 I_0}{(1-b)^2(1-A^0)(I_0^0 + I^0)^2(I_0 + I)}\right]^2 + \left[\frac{-I_0^0 I_0}{(1-b)^2(1-A^0)(I_0^0 + I^0)^2(I_0 + I)}s_{I^0}\right]^2 + \left[\frac{-I}{(1-b)(I_0 + I)^2}s_{I_0}\right]^2 + \left[\frac{I_0}{(1-b)(I_0 + I)^2}s_I\right]^2 \\
&= \left[\frac{I_0^0 I^0 I_0}{(1-b)^2(1-A^0)(I_0^0 + I^0)^2(I_0 + I)}\right]^2 \left[(\frac{s_{I_0^0}}{I_0^0})^2 + (\frac{s_{I^0}}{I^0})^2\right] + \left[\frac{-I}{(1-b)(I_0 + I)^2}s_{I_0}\right]^2 + \left[\frac{I_0}{(1-b)(I_0 + I)^2}s_I\right]^2 \\
&= \left[\frac{I_0^0 I^0 I_0}{(1-b)^2(1-A^0)(I_0^0 + I^0)^2(I_0 + I)}\right]^2 \left[(\frac{s_{I_0^0}}{I_0^0})^2 + (\frac{s_{I^0}}{I^0})^2\right] + \left[\frac{I_0 I}{(1-b)(I_0 + I)^2}\right]^2 \left[(\frac{s_{I_0}}{I_0})^2 + (\frac{s_I}{I})^2\right] \\
&\approx \left[\frac{I_0^0 I^0}{(1-A^0)(I_0^0 + I^0)^2}\right]^2 \left[(\frac{s_{I_0^0}}{I_0^0})^2 + (\frac{s_{I^0}}{I^0})^2\right] + \left[\frac{I}{I_0 + I}\right]^2 \left[(\frac{s_{I_0}}{I_0})^2 + (\frac{s_I}{I})^2\right], \text{when } b << 1, I << I_0 \\
&\approx \left[\frac{I^0}{I_0^0 + I^0}\right]^2 \left[(\frac{s_{I_0^0}}{I_0^0})^2 + (\frac{s_{I^0}}{I^0})^2\right] + \left[\frac{I}{I_0 + I}\right]^2 \left[(\frac{s_{I_0}}{I_0})^2 + (\frac{s_I}{I})^2\right], \text{when } A^0 << 1, I^0 << I_0^0
\end{aligned}
\tag{13}
$$

Eq. 13 suggests that b-calibration for type-3 attributes is more tolerant to low intensity of the attribute in the RS (I^0), because lower I^0 reduces the contribution of errors from the RS measurements.

Requirements for Effective Calibrations. To understand the requirements for an effective

calibration by each method, let's examine the equations describing each method.

The a-calibration described by eq. 2 requires 1) MS signal is proportional to and only to the attribute abundance, and 2) the same a-correction factor can be applied to both RS and samples. Requirement 1 assumes good method specificity and linearity, and requirement 2 assumes good overall repeatability within the sequence. The a-calibration holds effective as long as the two criteria are met.

The b-calibration described by eq. 4 and 5 requires 1) MS signal is proportional to and only to the attribute abundance, and 2) artificial modification is consistent within the same LC-MS sequence. Requirement 1 assumes good method specificity and linearity, and requirement 2 assumes good digestion repeatability within the sequence. The b-calibration holds effective as long as the two criteria are met.

The a&b-calibration described by eq. 7 and 8 requires 1) MS signal is proportional to and only to the attribute abundance, 2) the same a-correction factor can be applied to both RS and samples, and 3) artificial modification is consistent within the same LC-MS sequence. Requirement 1 assumes good method specificity and linearity, requirement 2 assumes good overall repeatability within the sequence, and requirement 3 assumes good digestion repeatability within the sequence. The b-calibration holds effective as long as the three criteria are met.

More In-Depth Discussion. The new method is significantly advantageous over the conventional MAM because it greatly reduces the lab-to-lab variability. It effectively eliminates the requirement of MAM to use a consistent equipment, which is a major difficulty in the current MAM workflow. The new approach solves the well-recognized instrument-to-instrument variability problem of MAM, which is often the reason for skepticism, therefore makes MAM potentially more acceptable by the biopharmaceutical industry and regulatory agency. With conventional MAM, validation and transfer of the method are challenging because of the requirement of the same instrument condition across different analytical laboratories. With the calibration methods described here, method validation and transfer become more straight-forward, because instrument models and settings that relate to response factors become less important. As described earlier, the calibration methods require acceptable specificity, linearity, and repeatability, as any other conventional analytical methods. Compared to conventional MAM, the only additional requirement is that specificity, linearity, and repeatability apply to the RS as well, meaning the attributes of interest in the RS must be above the LOQ. This additional requirement is not believed to be a major problem, considering an estimated LOQ of 0.002% for attributes that do not change during sample preparation (see below).

One drawback of conventional MAM methodology is that the measured attribute abundance is not absolute, because of the assumption that all isoforms have the same response factor, which may not

be true for modifications that involve change in charge, hydrophobicity or peptide length. In the example shown in this work, the reference attribute abundances in the RS were determined from six replicates by conventional MAM. These reference abundances, however, can also be determined from other methods. For example, glycan abundance can be determined by hydrophilic interaction liquid chromatography (HILIC) analysis of released glycans. Oxidation, deamidation, isomerization etc. can be quantified by peptide mapping with UV detection, if these peptide isoforms are chromatographically resolved from other peptides. Once the reference abundance of an attribute in the RS is determined, it will be defined as the true abundance and used for the life of the standard. More importantly, if the attribute abundance in the RS is determined by a technique with accurate absolute quantitation, attribute abundance determined from a-calibration also becomes an absolute value. The attribute abundance in the secondary or working standards can be determined using the method described here with the primary RS as calibrant.

Note that the methodology described in this report applies only to targeted attribute quantitation. Another important part of MAM is nontargeted new-peak detection, which is not affected in anyway by the methodology described here.

Additional Antistreptavidin Data with MS Settings Optimized for Quantitation. Response calibration effectively converts the requirement of long-term method repeatability to a short-term intra-sequence repeatability. Therefore, intra-sequence repeatability is the most critical performance parameter in the methodology. For the antistreptavidin dataset reported, the mass spectrometer spent significant amount of time collecting MS/MS. When analyzing samples in a routine setting, however, better data quality will be achieved by spending all available instrument time collecting full-scan MS data for quantitation purpose. To demonstrate, six replicate digests of antistreptavidin IgG2 were analyzed with LC-MS systems G and H with full MS scan only (Table S4), and with other instrument parameters optimized for maximal ion detection, including 2 micro scans to increase S/N of each scan, a low resolution of 70k to maximize scan speed, and a high AGC target of 3e6 to maximize ion injection.

Table S4. The two additional LC-MS systems used to analyze antistreptavidin digest in full-scan only mode.

System	*MS*	*HPLC*	*HPLC column*	*HPLC method*
G	Q Exactive Biopharma (unit 2)	Agilent 1290	Agilent Zorbax RRHD Stable Bond C18 (2.1 × 150 mm, 1.8 μm particle, 300 Å pore size) at 50°C	Additive: 0.1% formic acid Gradient: 1.0% acetonitrile for 5 min, then to 10% in 1 min, then to 35% in 64 min Flowrate: 0.25 mL/min
H	Q Exactive Biopharma (unit 3)	Waters Acquity		

Before analysis of the six replicates, the antistreptavidin digest were analyzed on System G four times by LC-MS/MS for peptide identification purpose. Processing the four LC-MS/MS files by MassAnalyzer identified a different set of attributes compared to the datafiles from systems A and B, among which 126 are in common. Abundance of these 126 attributes were determined by MassAnalyzer from each of the 6 files obtained from system G and 6 files from system H, and their relative standard deviation (RSD) determined. Table S5 lists the 126 attributes as well as their average abundances and RSD as determined from System G. It is apparent from Table S5 that a reproducible measurement of tryptophan oxidation is difficult, presumably because of artificial oxidations occurring during sample preparation. Excluding tryptophan oxidations, the average RSD for all other attributes determined by system G is 2.2%, including attributes with abundance less than 0.01%. Figure S3 shows the comparison of RSD distribution at different attribute abundance determined by systems A, G and H. It is apparent that the RSD decreased greatly with system G and H in full-scan only mode. In fact, the RSD was reduced by an average of 5-folds in systems G and H compared to system A. The median RSD decreased from 9.0% in system A to 2.5% in systems G and H for the 126 attributes. The intra-sequence precision for type-1 attributes is even better on systems G and H (Figure S4), with 97% of all type-1 attributes (abundance 0.002% to 40%) above quantitation limit (i.e., RSD < 10%).

Table S5. The 126 attributes monitored by system A, G and H. The average abundance and RSD of each attribute (n = 6) determined by system G are also shown.

Chain	Modification	Avg Abundance	RSD	Chain	Modification	Avg Abundance	RSD
HC	Q1+NH3	1.5799%	1.1%	HC	N289+M6	1.5391%	0.4%
HC	K13+Glycation	0.0255%	1.1%	HC	N289+M7	0.8134%	0.4%
HC	S30+C-term clip	0.0238%	4.4%	HC	N289+M8	0.8744%	1.8%
HC	W34+Double Oxidation	0.0974%	18.8%	HC	N289+Unglycosylated	0.8176%	0.7%
HC	W34+Oxidation to kynurenine	0.0413%	1.6%	HC	S296N	0.0165%	2.4%
HC	W37+Double Oxidation	0.1171%	29.0%	HC	V297I	0.0330%	1.1%
HC	K44+Glycation	0.0458%	1.3%	HC	V300I	0.0155%	2.4%
HC	W48+Double Oxidation	0.0305%	26.4%	HC	H302+Double Oxidation	0.0050%	9.1%
HC	W48+Oxidation	0.0083%	2.1%	HC	D304+H2O loss	0.0160%	3.7%
HC	G56+N-term clip	0.0042%	4.8%	HC	W305+Double Oxidation	0.0637%	37.5%
HC	N59K	0.0054%	1.1%	HC	W305+Oxidation	0.0315%	10.5%
HC	K65+Glycation	0.0357%	2.4%	HC	N307+Deamidation	0.6276%	1.4%
HC	R67Q	0.0068%	3.6%	HC	N307+NH3 loss	2.8816%	1.7%
HC	K117+Hydroxylation	4.8148%	0.7%	HC	K309+Glycation	0.0699%	2.5%
HC	S120+SAHexHexNAc	0.0205%	3.0%	HC	K318+Glycation	0.3106%	1.2%
HC	V121I	0.0080%	1.3%	HC	P321+Hydroxylation	0.1814%	2.0%
HC	F122Y	0.0030%	2.8%	HC	K352+Glycation	0.0470%	0.7%
HC	R129K	0.0061%	2.2%	HC	N353+NH3 loss	0.0774%	1.1%
HC	T131+N-term clip	0.0067%	2.6%	HC	K362+Hydroxylation	0.0051%	4.0%
HC	S132+N-term clip	0.0309%	1.5%	HC	N376+Deamidation	0.9491%	1.2%
HC	S134N	0.0241%	1.2%	HC	N376+NH3 loss	1.7793%	0.7%
HC	S134+N-term clip	0.0281%	1.4%	HC	N382+NH3 loss	0.4244%	0.5%
HC	T135+N-term clip	0.0051%	3.0%	HC	D393+H2O loss	0.5742%	3.4%
HC	C140Y	0.0215%	1.3%	HC	S400N	0.0205%	1.4%
HC	V142I	0.0039%	2.5%	HC	M420+Oxidation	0.6178%	0.8%
HC	W154+Double Oxidation	0.1188%	41.2%	HC	M420I	0.0304%	1.2%
HC	W154+Oxidation	1.6758%	0.8%	HC	N426+Deamidation	0.0817%	4.7%
HC	N155K	0.0054%	3.3%	HC	N426+NH3 loss	0.5183%	0.9%
HC	T191+N-term clip	0.0068%	13.2%	HC	G438+Lys	1.6757%	0.2%
HC	K238+Glycation	0.1361%	1.0%	LC	Q1+NH3	1.2864%	0.5%
HC	K240+Glycation	0.0686%	5.5%	LC	S2+N-term clip	0.0903%	0.7%
HC	M244+Oxidation	1.9956%	0.6%	LC	P14+Hydroxylation	0.2951%	0.9%
HC	M244I	0.0117%	1.5%	LC	R17K	0.0240%	2.1%
HC	S259R	0.0045%	3.6%	LC	N28K	0.0031%	1.9%
HC	D262+C-term clip	0.2140%	3.5%	LC	G30R	0.0045%	2.8%
HC	N268K	0.0375%	4.8%	LC	G32+N-term clip	0.0055%	4.6%
HC	W269+Double Oxidation	0.0884%	39.2%	LC	W37+Double Oxidation	0.0782%	35.6%
HC	K280+Glycation	0.3554%	2.8%	LC	W37+Oxidation	0.0123%	18.6%
HC	N289+A1G0	0.3501%	0.6%	LC	L48F	0.0066%	3.0%
HC	N289+A1G0F	0.9472%	0.9%	LC	G52D	0.0049%	3.1%
HC	N289+A1G1	0.2540%	0.6%	LC	N53K	0.0022%	2.3%
HC	N289+A1G1F	5.0031%	0.8%	LC	N53+Deamidation	0.4672%	3.0%
HC	N289+A1G1M4F	1.1888%	1.1%	LC	S54+SAHexHexNAc	0.0273%	3.1%
HC	N289+A1G1M5	0.7722%	0.9%	LC	G59E	0.0046%	2.5%
HC	N289+A1G1M5F	1.1327%	1.2%	LC	E83K	0.0145%	2.3%
HC	N289+A1S1	0.0310%	2.6%	LC	S98+N-term clip	0.0787%	3.2%
HC	N289+A1S1F	0.8934%	1.0%	LC	K108+Glycation	0.3554%	3.8%
HC	N289+A1S1M4	0.1247%	0.9%	LC	V109I	0.0108%	0.6%
HC	N289+A1S1M4F	0.5591%	0.4%	LC	V111A	0.0033%	2.0%
HC	N289+A1S1M5	0.1899%	0.6%	LC	V111I	0.0254%	9.2%
HC	N289+A1S1M5F	0.4202%	0.5%	LC	Q114+Deamidation	0.0048%	2.5%
HC	N289+A2G0	1.1835%	0.2%	LC	D144+H2O loss	0.0108%	6.8%
HC	N289+A2G0F	35.5540%	0.6%	LC	W154+Double Oxidation	0.2351%	29.6%
HC	N289+A2G1	0.2412%	1.8%	LC	W154+Oxidation	0.0377%	19.4%
HC	N289+A2G1F	34.1080%	0.7%	LC	K155+Glycation	1.4753%	1.8%
HC	N289+A2G2F	6.1746%	0.7%	LC	K177+Glycation	0.1195%	2.4%
HC	N289+A2S1G0F	0.1850%	1.3%	LC	S185R	0.0137%	1.2%
HC	N289+A2S1G1F	0.3235%	1.0%	LC	S185N	0.0444%	0.8%
HC	N289+A2S2F	0.0746%	1.9%	LC	W191+Double Oxidation	0.0544%	34.0%
HC	N289+A3G2F	0.0350%	11.0%	LC	W191+Oxidation	0.0348%	14.9%
HC	N289+Deamidation	0.0209%	2.7%	LC	K192+Glycation	0.0894%	1.7%
HC	N289+Gn	0.0162%	2.8%	LC	S198N	0.0600%	1.6%
HC	N289+M5	5.1831%	0.6%	LC	K210+Glycation	0.0621%	1.9%

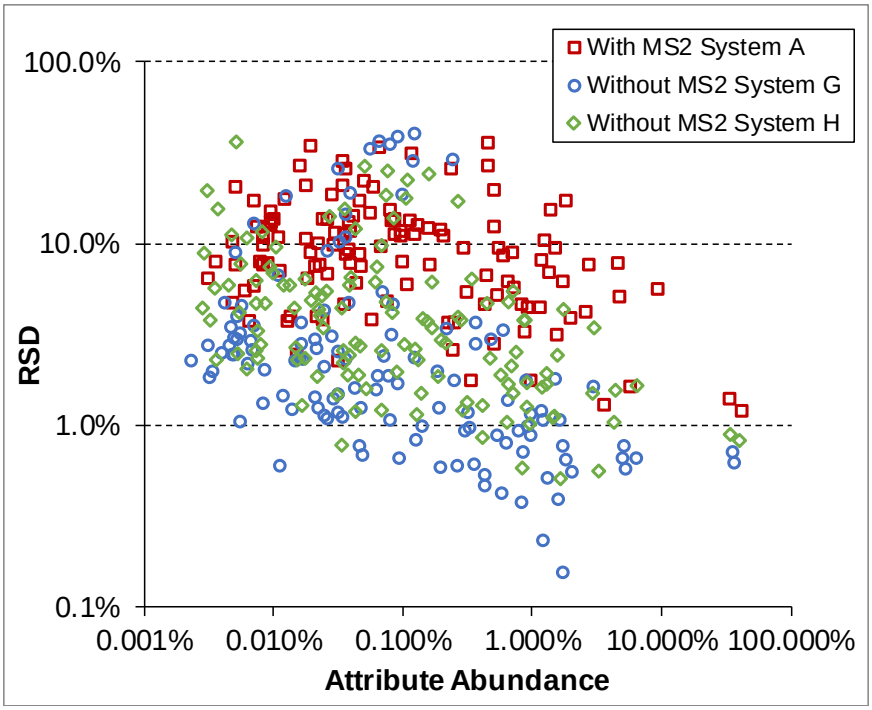

Figure S3. Reduction of RSD in full-scan MS only mode (systems G and H) compared with MS+MS/MS mode (system A).

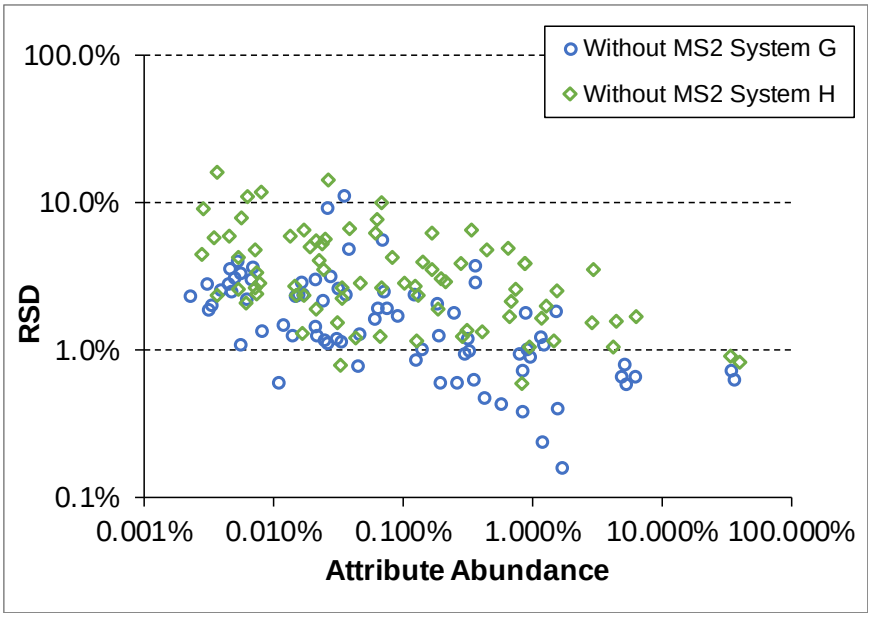

Figure S4. Distribution of intra-sequence RSD for type-1 attributes in antistreptavidin determined by systems G and H. Among the total 158 measurements ranging from 0.002% to 40%, only 5 have RSD above 10%, suggesting an LOQ approaching 0.002% for type-1 attributes.

In a routine setting, sample analyses are bracketed by RS analyses. To simulate this process, the first and sixth replicates were treated as standard and used to calibrate response factors of other four samples. Intermediate precision RSDs for the eight injections on two different systems were then calculated and compared in Figure S5. The percentages of attributes with intermediate precision RSD < 10% increased from 55% to 87% (out of a total of 126 attributes). Among the 126 attributes, only 12 type-3 attributes had RSD > 15% while all type-1 and type-2 attributes had RSD below 15%. Among the 12 attributes with RSD > 15%, 10 of them were tryptophan oxidations. This suggests that artificial oxidations of tryptophan were occurring during sample preparation and/or analysis, and more work is needed to reduce these artificial tryptophan oxidations. Note that tryptophan oxidation usually generates several peaks with different retention times. Integration of these peaks were checked manually to ensure the same peaks were integrated and quantified.

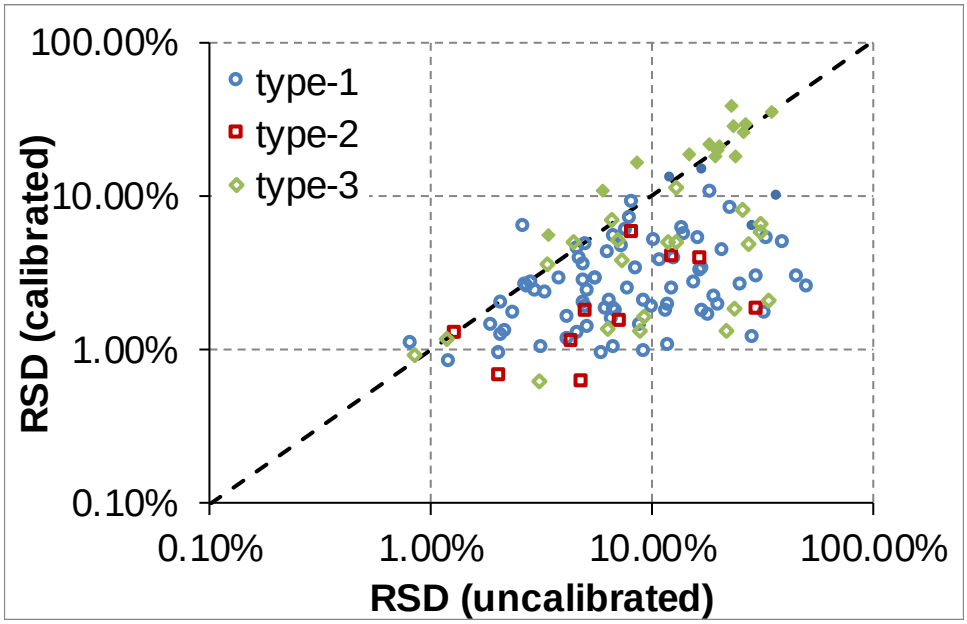

Figure S5. Comparison of intermediate precision (indicated by RSD of the 4 replicate measurements on each of the two LC-MS systems) with and without response factor calibration. The shape and color represent type of attributes, and the open shapes represent attributes with high enough abundance in the RS to be precisely quantified (intra-sequence RSD < 10%).

In the MAM workflow describe here, the reference abundance is defined as the true attribute abundance in the RS. When the reference abundance is used to calibrate the system, as long as the basic requirements of specificity, linearity and repeatability are met, the accuracy of the system is guaranteed. This is a basic principle of virtually all measurement systems, in which a system with acceptable

specificity, linearity and repeatability is calibrated with one or more reference standards to ensure accuracy. To demonstrate, in the same data shown above, when the first and sixth replicates were used as standard to calibrate response factors of other four samples, the median relative error of the 126 attributes determined by the two LC-MS systems decreased from 24.5% before calibration to 1.67% after calibration (Figure S6).

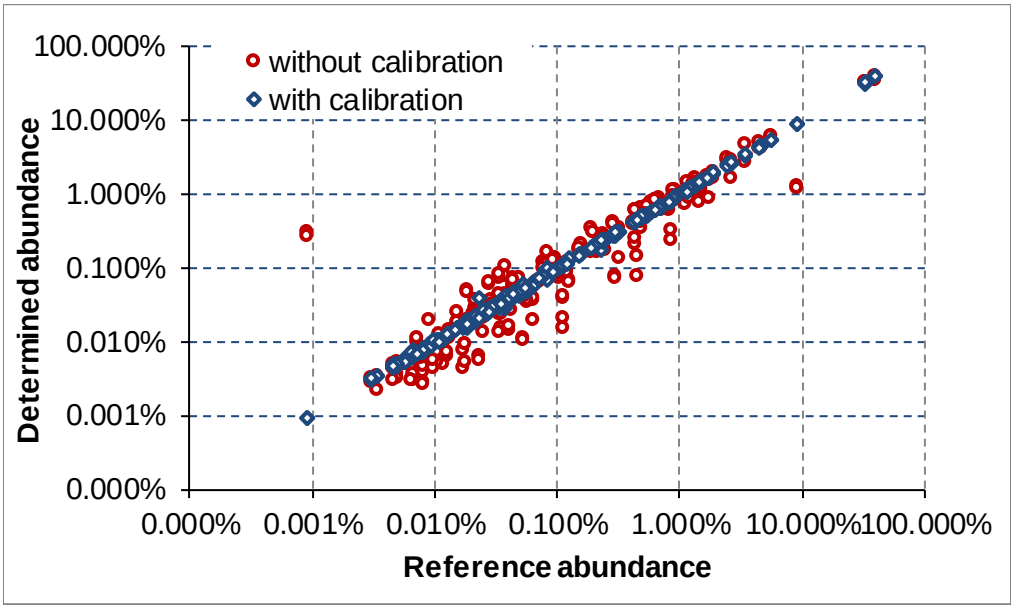

Figure S6. Comparison of determined attribute abundance (average of 4 measurements) in RS to the reference abundance with and without a-calibration. The accuracy improved greatly, with median error reduced from 24.5% to 1.67%.

LOQ. LOQ can be estimated from the minimum concentration of an attribute with intra-sequence RSD below 10%. Figure S7 shows the relationship of attribute abundance and intra-sequence RSD for the three attribute types in antistreptavidin IgG2 (on LC-MS System A). For type-1 attributes, the RSD for most attributes are below 10% with abundances as low as 0.003% (LOQ = 0.003%). For type-3 attributes, however, most attributes below 0.1% have RSD > 10% and most attributes above 1% have RSD < 10%, suggesting that the LOQs are typically between 0.1% and 1%, depending on the amount of variations introduced during sample preparation. Type-2 attributes appear to be between type-1 and type-3 attributes.

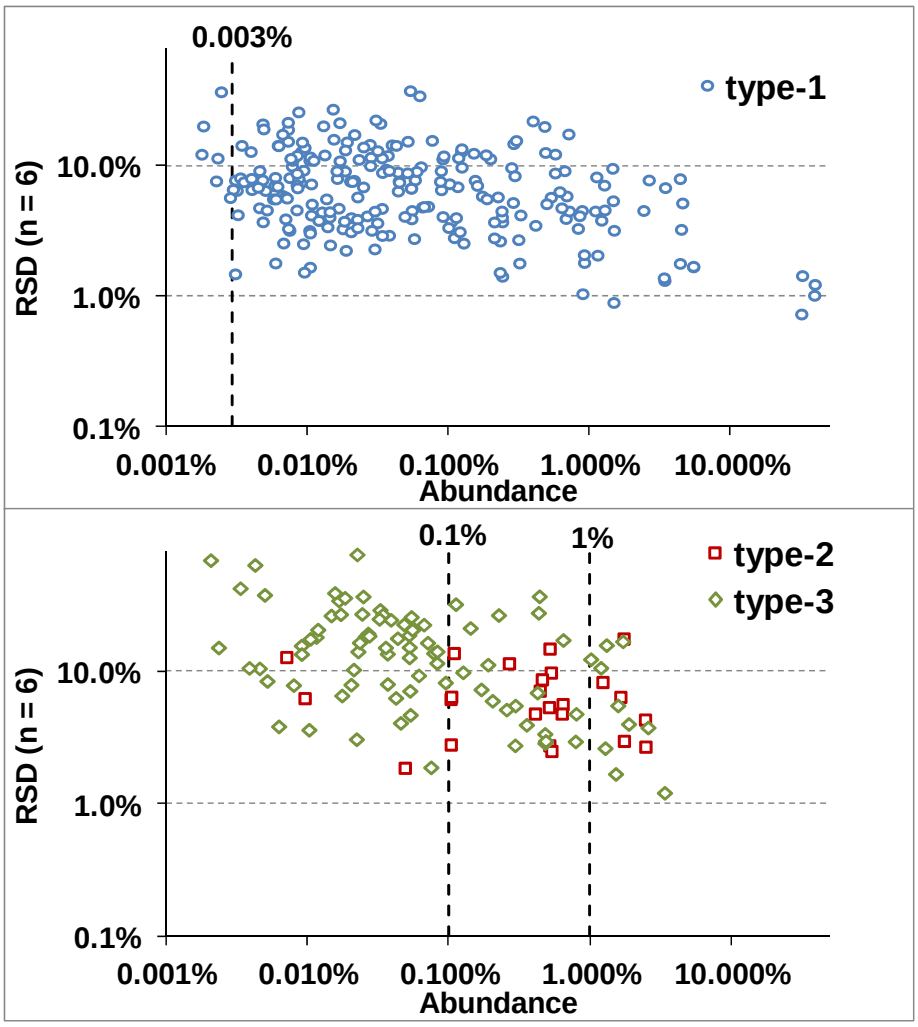

Figure S7. Relationship of attribute abundance and intra-sequence RSD (n=6) for different attribute types. For majority of type-1 attributes (top), intra-sequence RSDs are below 10%, with abundance down to 0.003%, indicating the LOQ of the LC-MS system is as low as 0.003% when the attribute does not change during sample preparation. Quantitation limits are much higher (0.1 – 1%) for most type-3 attributes (bottom).