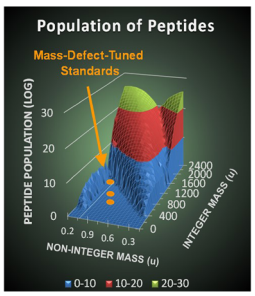
J. Am. Soc. Mass Spectrom. (2018) 29:490–500
DOI: 10.1007/s13361-017-1849-y

RESEARCH ARTICLE

Iodine-Containing Mass-Defect-Tuned Dendrimers for Use as Internal Mass Spectrometry Calibrants

Joseph A. Giesen,[1] Benjamin J. Diament,[2] Scott M. Grayson[1]

[1]Department of Chemistry, 2015 Stern Hall, Tulane University, New Orleans, LA 70118-5636, USA
[2]Department of Computer Science and Engineering, University of Washington, Box 352350, Seattle, WA 98195-2350, USA



Abstract. Calibrants based on synthetic dendrimers have been recently proposed as a versatile alternative to peptides and proteins for both MALDI and ESI mass spectrometry calibration. Because of their modular synthetic platform, dendrimer calibrants are particularly amenable to tailoring for specific applications. Utilizing this versatility, a set of dendrimers has been designed as an internal calibrant with a tailored mass defect to differentiate them from the majority of natural peptide analytes. This was achieved by incorporating a tris-iodinated aromatic core as an initiator for the dendrimer synthesis, thereby affording multiple calibration points (*m/z* range 600–2300) with an optimized mass-defect offset relative to all peptides composed of the 20 most common proteinogenic amino acids.

Keywords: Calibrant, Mass defect, Dendrimer, MALDI-TOF MS, ESI MS, Peptides

Received: 14 April 2017/Revised: 18 October 2017/Accepted: 13 November 2017/Published Online: 19 December 2017

Introduction

Matrix-assisted laser desorption ionization (MALDI) [1] and electrospray ionization (ESI) [2, 3] MS were developed as soft ionization techniques suitable for very large molecular weight ions. These two ionization techniques in combination with time-of-flight (TOF) mass analyzers have provided a means of characterizing analytes with a molecular weight as high as one million [4]. Because TOF instruments use the flight time of a given analyte from ion extraction at the source to ion impact at the detector to determine *m/z*, any alteration to the ionization conditions or acquisition parameters can affect the observed time of flight, and therefore the determined *m/z*. As a result, precise and regular calibration is required to provide highly accurate mass values.

Current calibrants for TOF MS systems include peptides, proteins, ion clusters [5–10], and polydisperse synthetic polymers [11, 12]. Although peptides and proteins have been frequently used for MS calibration, they are susceptible to modifications including deamidation [13], methionine oxidation [14], disulfide bridge formation/scission [15], and post-translational modifications, which can cause signal broadening or misidentification, in addition to ambiguity as to their true mass [16]. Furthermore, their inherent instability toward these reactions, as well as proteases and chemical degradation, requires expensive purification and results in limited shelf-lives even when refrigerated [17]. Many synthetic polymers and nanocluster formulations have been proposed that overcome the concerns with cost and shelf-life; however, when used for internal calibration, their multiplicity of signals increases the likelihood of signal overlap with a potential analyte. Furthermore, if the spacing within the peak distribution is relatively narrow, specific calibration points are more likely to be misidentified (as one repeating unit greater or less), skewing the calibration curve [18]. Recently, polyester-based dendrimers have been proposed as an attractive alternative for both MALDI and ESI MS calibration because of their efficient synthesis, high purity, broad molecular mass range, and true monodispersity [19–21]. Additional practical advantages, such as broad compatibility with MALDI matrices and extended shelf-lives, have led to their commercialization.

With the growing potential of mass spectrometry for the rapid screening of peptides and proteins, the use of internal calibrants is particularly appealing for maximizing mass accuracy and thereby improving the success of peptide identification. However, most biological and synthetic macromolecular calibrants consist predominately of H, C, N, and O, which all exhibit a mass defect that is either positive or negligibly negative. As a result, these calibrants exhibit a similar positive

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s13361-017-1849-y) contains supplementary material, which is available to authorized users.

Correspondence to: Scott Grayson; *e-mail:* sgrayson@tulane.edu

mass defect to those expected for biological analytes, increasing the likelihood that the analyte signal might be obscured, shifted, or misidentified because of a nearby or overlapping calibrant signal. An ideal internal calibrant set would have a mass defect [22] signature that clearly differentiates it from the majority of peptide analytes across the entire mass range of interest. While mass-defect labeling has been explored to tag peptide analytes, [23, 24] the concept of incorporating a mass defect label into a calibrant has only been demonstrated to date by the inclusion of multiple fluorine atoms into the mass standards [25]. However, while the negative mass defect of multiple fluorine atoms can provide contrast relative to the positive mass defect observed in most synthetic or biological polymers, a significant number of fluorine atoms (nearly 40) would be required to maximize the mass-defect offset relative to common analytes. An attractive alternative is the incorporation of iodine, which exhibits a much larger negative mass defect, [26] nearly 60 times greater per atom than F. In order to design calibrants with an optimized mass defect, the mass-defect distributions among all peptides (composed of the 20 most common proteinogenic amino acids) were first calculated, and this data set was used to identify tris-iodinated cores as the ideal initiating groups for the synthesis of dendrimer-based mass-defect calibrants.

Three specific goals are addressed in this study: (1) a computational analysis of the statistical distribution of peptides to identify the most effective calibrant targets, (2) the design and synthesis of a series of mass defect-tuned dendrimer calibrants, and (3) the assessment of the tris-iodinated dendrimers as internal calibrants for MALDI-TOF and ESI MS analysis.

Experimental and Methods

Materials

All reagents including 2,4,6-triiodophenol, 2,2-bis(hydroxymethyl)propionic acid (bis-MPA), 2,2-dimethoxypropane, polyethylene glycol (PEG) 750 M_n, propargyl bromide, *p*-toluenesulfonic acid monohydrate (*p*-TsOH), 4-(dimethylamino)pyridine (DMAP), palladium on carbon 10% wt. loading (Pd/C), and Celite as well as matrices and salts: *trans*-2-[3-(4-*tert*-butylphenyl)-2-methyl-2-propenylidene] malononitrile (DCTB), 1,8-dihydroxyanthracen-9(10H)-one (dithranol, DIT), α-cyano-4-hydroxycinnamic acid (CHCA), sodium trifluoroacetate, trifluoroacetic acid, sodium bicarbonate ($NaHCO_3$), sodium hydrogen sulfate ($NaHSO_4$), and sodium chloride (NaCl) were obtained from Sigma Aldrich (St. Louis, MO, USA) and used without further purification. Sodium hydride (NaH) was also obtained from Sigma Aldrich as a suspension in oil. This reagent was rinsed with hexanes and the solvent was decanted off. This washing was repeated six times and the NaH was dried prior to use. Solvents, including acetone, acetonitrile, dichloromethane (DCM), ethanol (EtOH), methanol (MeOH), tetrahydrofuran (THF), hexanes, and glacial sulfuric acid (H_2SO_4), were obtained as reagent grade from Fisher Scientific (Fair Lawn, NJ, USA) and used as received. Dowex acid resin (50Wx4, 200-400 mesh, Acros Organics) was reactivated with 1M H_2SO_4 (250 mL) and then washed with 60 mL each of EtOH, THF, acetone, DCM, and hexanes to remove small molecular weight impurities, and the resulting light yellow colored solid was collected and dried on high vacuum overnight. Deionized water was purified in-house with an Elga PureLab Prima system (ELGA, Wycombe, UK). Graphite matrix was used as is from a Ticonderoga # 2 pencil. A sample of trypsin-digested BSA (CAM-modified) was acquired from New England Biolabs (Ipswich, MA, USA).

Experimental Methods

Mass spectra were obtained on a Bruker Autoflex III MALDI-TOF MS (Bruker Daltonics, Billerica, MA, USA) using 1 kHz smartbeam II Nd:YAG laser. Data were acquired in reflector-positive ion mode, with pulsed ion extraction (20 ns) from a stainless steel target plate (MTB 384) purchased from Bruker (Billerica, MA, USA). Bruker Daltonics FlexControl 3.0 software was used for data acquisition, and data analysis was carried out utilizing Bruker Daltonics FlexAnalysis 3.0 software. Samples were prepared from stock solutions of analytes, namely, the tris-iodinated calibrants, dipropargyl PEG, and SpheriCal calibrants (2 mg mL^{-1}). Additionally, stock solutions of matrices, DCTB, DIT, CHCA (10 mg mL^{-1}), and the cation source, sodium trifluoroacetate (1 mg mL^{-1}), were prepared in THF. The stock solution of endomorphin I (2 mg mL^{-1}) was prepared in 50/50 deionized water/acetonitrile. Sample solutions were then prepared by mixing 10 μL of matrix solution, 5 μL of analyte, and 5 μL of the cation stock solutions, and 3 μL of the resulting solution was plated via dried-droplet method. Alternatively, the two-layer method was utilized for graphite matrix with graphite first being scribed onto the target spot followed by addition of 3 μL of a 1:1 analyte:salt solution prepared by mixing 5 μL of both analyte and cation stock solutions. BSA digest samples were prepared as recommended by NEB protocol, and the calibrant sample (1 μL) added directly to this for internal calibration, with no additional sodium. Data were collected using approximately 5000–10,000 laser shots under the following acquisition parameters: ion source 1: 19.0 kV; ion source 2: 16.55 kV; lens: 8.50 kV, reflector 1: 21.0 kV, reflector 2: 9.64 kV; detector: 2.0 kV, and a low mass gauge at 300 *m/z*. The laser power was set to the minimum value that would yield high resolution spectra. The initial mass scale for the MALDI-TOF MS analysis was calibrated using SpheriCal mass standards (Polymer Factory, Sweden) prepared via the dried-droplet method using identical concentrations and ratios as above.

ESI analysis was performed on a Bruker MicrOTOF II MS, utilizing micrOTOFControl 3.0 and DataAnalysis 4.0 software for data acquisition and analysis, respectively. Samples were prepared in micromolar solutions of acetonitrile for the acetonide protected dendrimers and methanol for the polyhydroxyl dendrimers. The BSA digest sample was prepared according to the supplied protocol and then doped with an

equal volume of the 0.1–1.0 µM G1-G3 calibrant mixture. Spectra were obtained via direct injection at 0.01 mL min^{-1}, end plate offset: –500 V; capillary: 2800 V; nebulizer: 0.3 Bar; dry gas: 4.0 L min^{-1}; dry gas temp: 180 °C; capillary exit: 90.0 V; skimmer 1: 30.0 V; hexapole 1: 23.0 V; hexapole rfRF: 400.0 Vp-p; skimmer 2: 22.0 V; lens 1 transfer: 72.0 µs; and lens 1 pre-pulse storage: 5.0 µs.

To further confirm the structure of the synthetic calibrants, NMR data were acquired on a 400 MHz Varian Mercury spectrometer (Palo Alto, CA, USA) or a 300 MHz Bruker Avance III spectrometer (Billerica, MA, USA). The residual solvent signals were used as the reference for both. NMR spectra were obtained in chloroform-d ($CDCl_3$) and methanol-d_4 (CD_3OD), purchased from Cambridge Isotope Laboratories (Andover, MA, USA), for the acetonide-protected dendrimers and the deprotected hydroxylated dendrimers, respectively. Gel permeation chromotography (GPC) was carried out on a Waters model 1515 series pump, equipped with a Waters model 2707 autosampler and fitted with a Waters model 2487 differential refractometer detector (Milford, MA, USA). Sample analysis was performed with a THF mobile phase (1 mL min^{-1} flow rate) utilizing a series of two columns from Polymer Standard Services, (Mainz, Germany) (1) guard column (50 mm), (1) analytical linear M (3 µm, 8 × 300 mm) and (1) analytical 100 Å (3 µm, 8 × 300 mm) calibrated with polystyrene standards for all generations of the protected dendrimers. Deprotected dendrimers were not analyzed by GPC because of limited solubility in the THF mobile phase resulting from their high hydroxyl content.

Synthesis of the acetonide protected monomer was carried out as previously reported by Ihre, et al [27, 28]. Dendronization was carried out with the acetonide anhydride **1**, with DMAP catalyst using triiodophenol as the dendrimer core, per literature [29, 30]. The subsequent acid-catalyzed removal of the acetonide protecting groups was carried out using Dowex acidic resin in methanol at 40 °C under reduced pressure (556 mbar). Full synthetic details can be found in the Supporting Information.

Computational Investigations

In order to optimize the design for a set of mass-defect calibrants, the population of all possible peptides (MW 0-2400) was computed and the population graphed with respect to the integer mass and non-integer mass. For this discussion, it is important to first clarify the definitions of "integer mass" and "non-integer mass". Integer mass corresponds to the whole number portion of a compound's molecular mass, regardless of the contribution of mass defect, whereas non-integer mass corresponds to the remainder of the mass (always <1.00 u). For example, the monoisotopic mass of the peptide $(Leu)_8(Lys)_5$ ($C_{78}H_{150}N_{18}O_{14}$) is 1563.1579 where 1563 is the integer mass and 0.1579 is the non-integer mass (Figure 1). These mass definitions are in contrast with the more frequently used terms "nominal mass" and "mass defect." Nominal mass of a molecule corresponds to the sum of the nominal masses of each of a

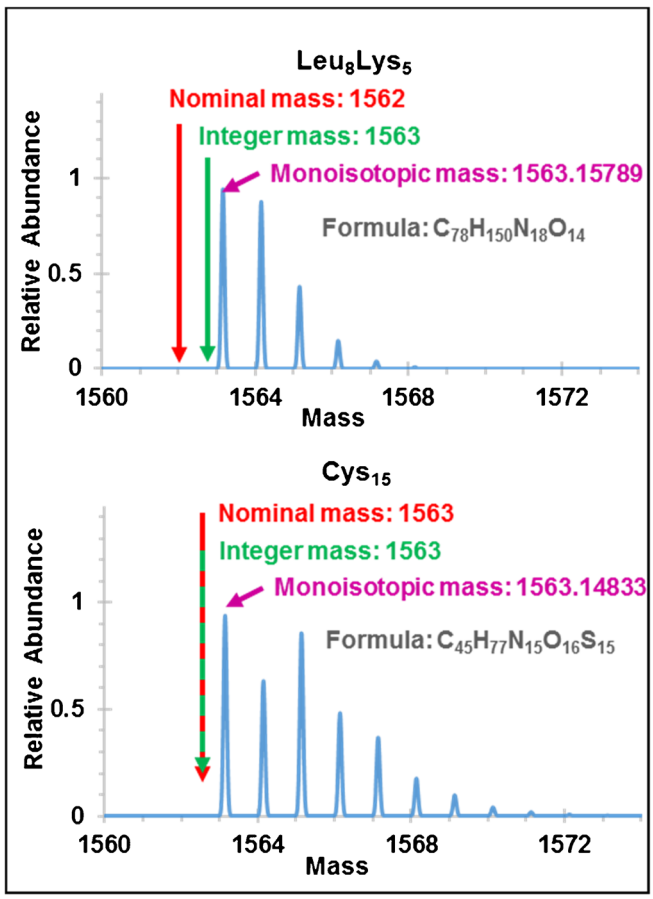


Figure 1. Comparison of nominal mass and integer mass for two representative peptides

compound's constituent atoms, whereas mass defect of a compound represents the difference between its monoisotopic mass and its nominal mass, which is also the sum of the mass defects of each of the compound's constituent atoms. Using the same example as above, the nominal mass of the peptide $(Leu)_8(Lys)_5$ is 1562, which corresponds to the sum of the nominal atomic masses (78 × 12) + (150 × 1) + (18 × 14) + (14 × 16) = 1562. Likewise, the mass defect of the 13-mer peptide is the sum of the atomic mass defects (78 × 0.00000) + (150 × 0.007825) + (18 × 0.003074) + (14 × –0.005085) = 1.1579. In this case, the 1 u difference between nominal mass and integer mass is the result of a mass defect greater than 1. This is in contrast to the cysteine 15-mer ($C_{45}H_{77}N_{15}O_{16}S_{15}$) (Figure 1), the integer mass and nominal mass of which are both 1563, owing to a mass defect less than 1 (45 × 0.00000) + (77 × 0.007825) + (15 × 0.003074) + (16 × –0.005085) + (15 × –0.027929) = 0.14833). It is important to distinguish this difference when designing a mass defect-tuned calibrant because the ideal mass targets should be designed based on the distribution of non-integer masses relative to integer masses, rather than the distribution of mass defects relative to nominal masses. The former are important when analyzing unknowns because they are the masses that are observed empirically, whereas the latter can only be determined if the atomic composition of the compound in question is known. The use of

integer and non-integer mass becomes more important at higher mass ranges where molecular mass defects greater than 1 are increasingly common, and the ramifications for mass defect-tuned calibrants will be explored in more detail below.

In Figure 2a and b, the integer mass of each possible peptide is measured on the x-axis, its corresponding non-integer mass is measured on the y-axis, and the z-axis represents the population of peptides with that specific mass (the population values calculated for 0.01 u widths in the non-integer mass for each integer mass). Figure 2c–f represent cross-sections of the 3-D graph at given integer masses. It should be noted that these initial population calculations were determined assuming an unbiased statistical incorporation of the 20 most common proteinogenic amino acids residues, rather than the actual frequency of occurrence, and without taking into consideration the effect of post-translational modifications.

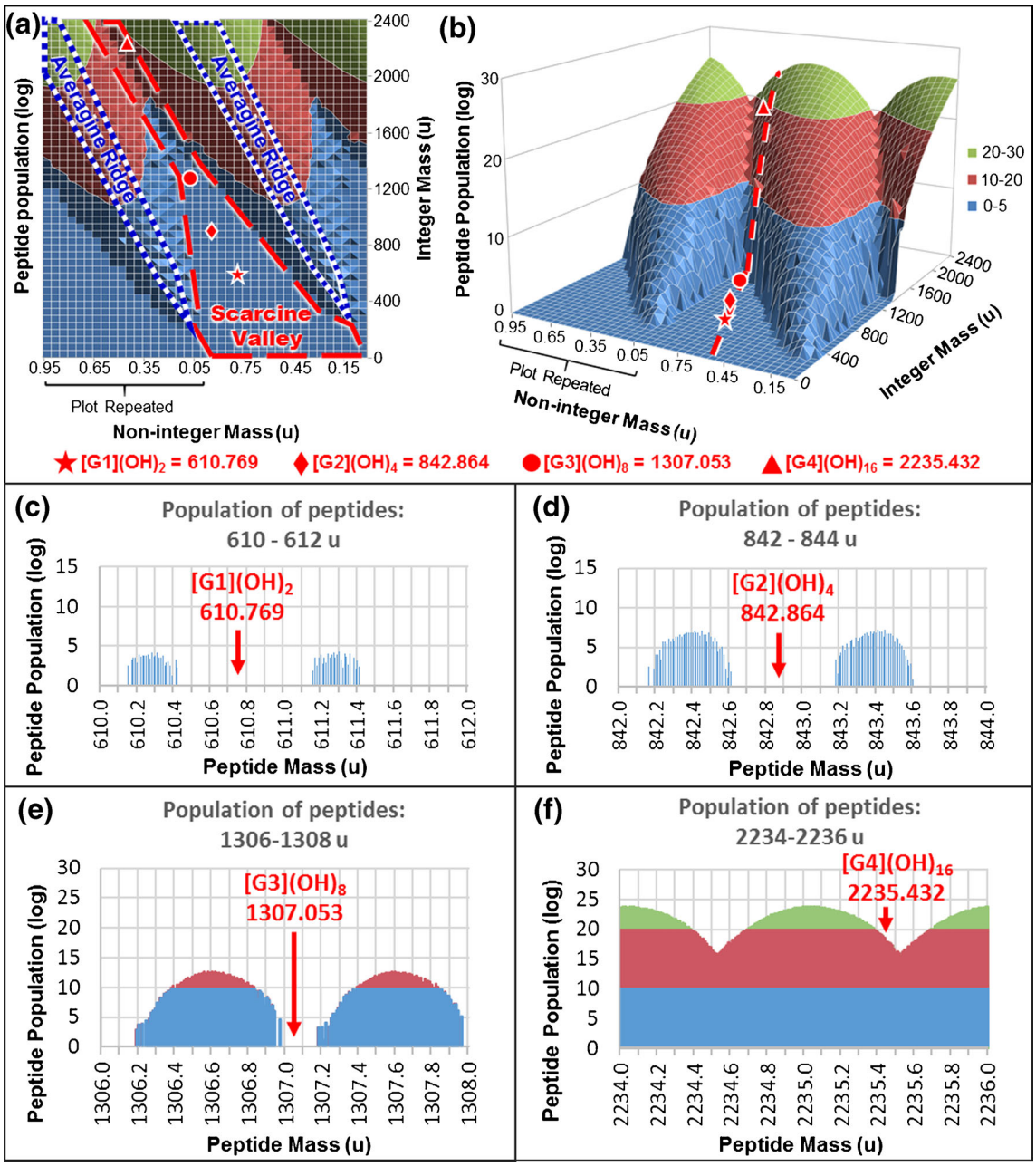


Figure 2. **(a)** 2-D overview of peptide population highlighting the most populous non-integer masses with respect to each integer mass (averagine ridge) and the least populous non-integer masses (scarcine valley); **(b)** 3-D oblique view of the same data set, with an overlay of the mass-defect-tuned calibrant compounds and their trend line along the scarcine valley (mass-defect-tuned dendrimer monoisotopic masses given as sodium adducts). Note: the x-axis (non-integer mass) has been duplicated to yield a width of 2.0 u in order to aid in visualization of the topology of the data set; **(c)**-**(f)** cross-sections of the 3-D graph at the integer masses of each of the dendrimer calibrants. Note: each cross-section's non-integer mass scale is inverted with respect to the 3-D graph

Table 1. Mass defect for common non-metals

Element	Isotope	Isotopic abundance (%)	Monoisotopic mass (u)	Mass defect (u)	Mass defect per u ($\times 10^{-6}$)
Hydrogen	^{1}H	99.9885	1.00783	0.00783	7770
	^{2}H	0.0115	2.0141	0.0141	700
Carbon	^{12}C	98.93	12	0	0
	^{13}C	1.07	13.00335	0.00335	258
Nitrogen	^{14}N	99.632	14.00307	0.00307	219
	^{15}N	0.368	15.00011	0.00011	7.3
Oxygen	^{16}O	99.757	15.99491	–0.00509	–318
	^{17}O	0.038	16.99913	–0.00087	–51.2
	^{18}O	0.205	17.99916	–0.00084	–467
Fluorine	^{19}F	100	18.9984	–0.0016	–84.2
Chlorine	^{35}Cl	75.787	34.96885	–0.03115	–891
	^{37}Cl	24.22	36.96885	–0.03419	–925
Bromine	^{79}Br	50.69	78.91834	–0.08166	–1035
	^{81}Br	49.31	80.90585	–0.08371	–1035
Iodine	^{127}I	100	126.90447	–0.0955	–753
$Averagine_{ud}$ $C_{5.35}H_{7.85}N_{1.45}O_{1.45}S_{0.10}$			118.8057	0.0557	469

When examining the shape of the peptide population graph (Figure 2a and b), it is observed that no peptides are identified at masses below 132 u (the integer mass of diglycine), as expected. For an integer mass of 610 u the data set shows a relatively narrow 0.27 u range of possible non-integer masses (from 0.15 to 0.42 u) (Figure 2c). This leaves a large "unoccupied" window of non-integer masses for which there are no peptides with this particular integer mass (i.e., 610). With increasing integer masses, the non-integer mass range that is populated with peptides broadens because of the larger number of amino acids per peptide and the variation of the relative mass defect from one amino acid to another. To help visualize the origin of this non-integer mass broadening, the homopolymers of each amino acid have been graphed in a 2-D plot of integer mass versus non-integer mass (Supplementary Figure S1). For example, leucine and isoleucine (131.09463) and lysine (146.10553) exhibit the largest mass defect per nominal mass (7.224×10^{-4} and 7.228×10^{-4} mass defect/nominal mass, respectively), whereas cysteine, due to its sulfur content, exhibits the lowest mass defect per nominal mass (1.632×10^{-4} mass defect/nominal mass). As a result, at integer masses of 1307 u, the populated non-integer masses range increases to be 0.78 u wide, from 0.18 to 0.96 u (Figure 2e), leaving a much narrower unoccupied window where no peptides exhibit that combination of integer and non-integer masses. Finally, above 1500 u, the width of the populated non-integer masses occupies the entire range from 0.00 to 0.99 u because the signals of the peptides in this range with the lowest possible mass defect (e.g., the 15-mer of polycysteine, $(Cys)_{15}$, with and integer mass of 1563 and a mass defect of 0.148) begin to overlap with the peptides with the highest mass defect from one integer mass below (e.g., $(Leu)_8(Lys)_5$ with an integer mass of 1562, but a mass defect of 1.1579. If the peptide population is graphed on a continuous 2-D plot, relative to the monoisotopic mass, the regions populated with peptides will generate a sawtooth pattern with each tooth spaced ~1 u apart. Segments of this sawtooth graph are depicted in Figure 2c–f. At low masses (Figure 2c) there is an appreciable unpopulated region between teeth. However, at higher masses (above integer masses of 1500) these unpopulated regions narrow until the teeth merge because there are no longer non-integer masses that are unoccupied by peptides. Figure 2a and b are the full data set for integer masses 0–2400 in a 3-D graph where Figure 2c–f represent the cross-sections of this 3-D graph at particular integer masses. Although the range of the mass defect should only be 0.00–0.99, this 3-D data set (Figure 2a and b) is graphically repeated once, in order to better visualize the topographical features that cross over graph's boundary.

When further examining the shape of the peptide population topology, a number of important trends are observed. First, while O and S exhibit slightly negative mass defects, their contribution is minor relative to the mass defect of H, C, and N; and therefore, the peptide population exhibits a trend of an increasing non-integer mass with respect to increasing integer mass. The quantification of this slightly positive average mass defect in a statistical population of peptides has been described previously [31]. This concept was refined by Senko et al., and the term "averagine" was coined to describe the "average" amino acid residue based on the observed ratio of amino acids, yielding the molecular formula of $C_{4.9384}H_{7.7583}N_{1.3577}O_{1.4773}S_{0.0417}$ and a monoisotopic mass of 111.0543 u [32]. It is important to note that this more commonly used averagine value was calculated based on the frequency of occurrence for each amino acid residue by the Protein Identification Resource Database. Therefore, this definition of averagine should be considered a variable, dependent upon the database used.

For our initial computational investigation, a "database independent" analysis was carried out based on an equal likelihood for each amino acid residue to occur within the population of peptides. Thus the older averagine calculation, based on an unbiased "uniform distribution" of amino acids [31] is more appropriate for this data set and for clarity will be named $averagine_{ud}$. The formula for $averagine_{ud}$ is $C_{5.35}H_{7.85}N_{1.45}O_{1.45}S_{0.10}$, which corresponds to a monoisotopic

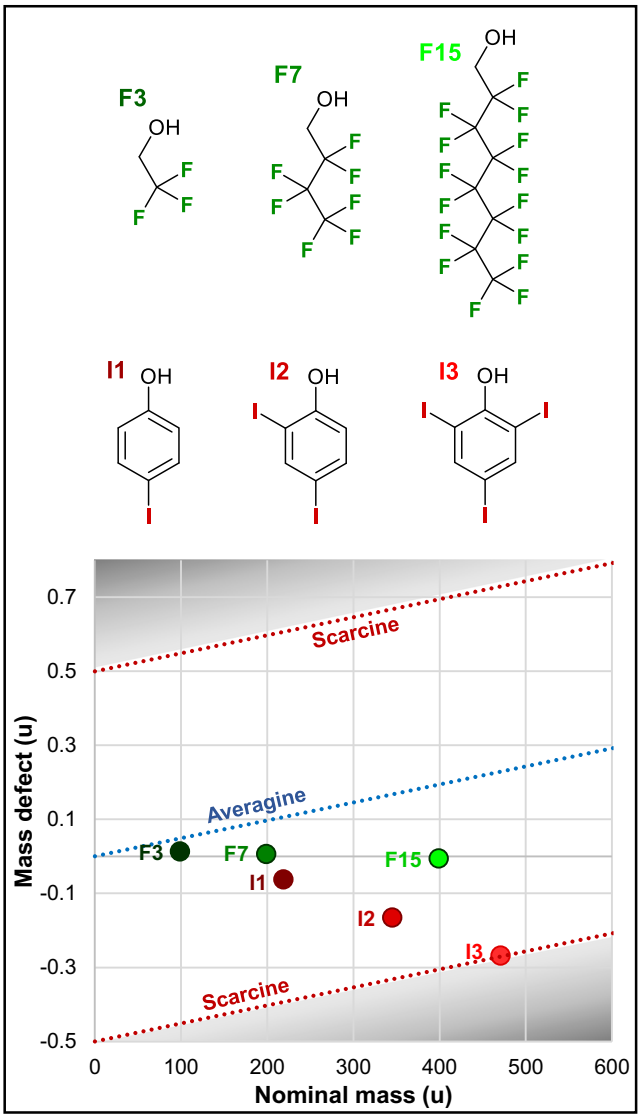


Figure 3. The mass defect of halogenated alcohols was graphed with respect to their nominal mass for a range of potential cores: F3) trifluoroethanol, F7) heptafluorobutanol, F15) pentadecafluorooctanol, I1) 4-iodophenol, I2) 2,4-diiodophenol, I3) 2,4,6-triiodophenol

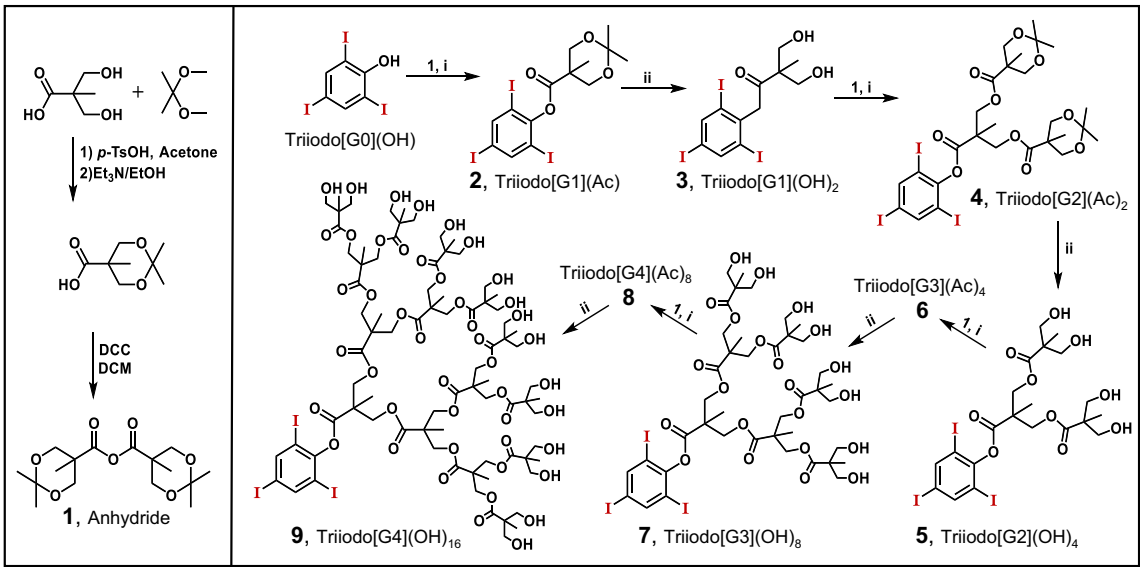
mass of 118.80572. When defining the mass defect with respect to the nominal mass, it is also important to note that the monoisotopic mass of averagine$_{ud}$ has a nominal mass component of 118.75000, which is not a whole number (a consequence of the non-whole number molecular formula), and a mass-defect component of 0.05572. While the ratio of these two values defines the averagine$_{ud}$ trend in a 2-D graph of mass defect versus nominal mass (with a slope of 4.691×10^{-4}), in the 3-D topology this same concept is observed as the "averagine ridge": the maximum population of peptides for each non-integer mass, relative to its integer mass.

In contrast to this topological maxima represented by the averagine ridge, for each integer mass there are either completely unpopulated non-integer mass regions (in the mass range below ~1500 u), or a minimally populated "valley" (for masses above ~1500 u). For the sake of discussion, this region will be termed the "scarcine valley" and represents the ideal target region for mass defect-labeled calibrants in order to minimize the potential of signal overlap with natural peptide analytes. It is also useful to note that the least populated "scarcine$_{ud}$" trend line should exhibit a mass-defect offset of ~0.5 u, with respect to the averagine$_{ud}$ trend line, but exhibiting the same slope of 4.691×10^{-4} (non-integer mass versus integer mass). Finally, when examining the peptide population data (Figure 2) it is important to reiterate that for the purpose of designing mass standards with monoisotopic masses that are not too close to those of any peptide analyte, integer and non-integer masses are the only appropriate means to analyze the peptide population data sets because once the mass defect for a given compound has increased to exactly 1 u, it is indistinguishable experimentally from a compound with no mass defect but 1 additional unit of nominal mass.

Finally, this data set also clarifies an important limitation of mass defect-tuned calibrants. A comparison of the maximum population of peptides (averagine) and the minimum population of peptides (scarcine) for various integer masses shows that this ratio decreases rapidly with increasing molecular weight. As detailed in Supplementary Table S1, while the peptide population for the integer mass of 2000 u varies by 9 orders of magnitude depending upon the non-integer mass, above 8000 u this variation is reduced to less than 2 orders of magnitude. Therefore, mass defect-tuned calibrants are most useful for the identification of single peptides or the peptide fragments of proteins, and increasingly relevant in the lower mass ranges (600–2300), which corresponds to the typical mass range of protein digests.

Mass-Defect Calibrant Design

The versatility of the polyester dendrimer synthesis makes it particularly amenable to preparing mass defect-tuned calibrants. Because the selected dendritic repeat units consist of C, H, and O in similar ratios ($C_5H_8O_3$) as averagine, their ratio of mass defect with respect to nominal mass is sufficiently close to that of averagine to minimize deviation in the slope (mass defect/nominal mass) across the mass range of interest (<2400). However, a core molecule must be identified that provides an appropriate non-integer mass offset relative to the peptide population, i.e., falls near the scarcine trend line. Therefore, such a core must include multiple atoms that exhibit a significant negative mass defect. Although many high mass elements exhibit a desirably negative mass defect, it is advantageous to focus on monoisotopic elements to maintain a narrow isotopic distribution. The halogens are particularly useful to consider because they can also be easily incorporated into organic compounds via replacement of a C–H bond with a C–X bond (where X = halogen). Although halogens have been explored as mass-defect labels, bromine, [33, 34] which has been utilized most frequently, [35, 36] is complicated by two stable isotopes, ^{79}Br and ^{81}Br, and therefore can cause



Scheme 1. Synthesis of dendrimer by repetitive dendritic growth (i) and deprotection (ii) steps. Conditions: (i) **1** (>1.2 eq per OH), DMAP (10% wt), DCM, and (ii) Dowex acid resin, MeOH, 556 mbar, 40 °C

undesirable isotopic broadening when multiple bromines are introduced. Fluorine, on the other hand, is monoisotopic, but exhibits a relatively small mass defect (–0.00160 u) even when considered in proportion to its nominal mass of 19 ($–8.42 \times 10^{-5}$ u mass defect per u nominal mass). Iodine, however, is monoisotopic and possesses both a much larger molecular weight (127) and a much larger mass defect per atomic mass unit ($–7.52 \times 10^{-4}$ u mass defect per u nominal mass). This combination yields a negative mass defect 59.7 times greater per atom than fluorine (relative mass defects tabulated in Table 1). However, because organic iodides are very reactive towards nucleophilic substitution, to ensure stability of the generated calibrants, iodine is best incorporated using the relatively stable aryl–iodine bonds. Figure 3 depicts a graph of the mass defect versus nominal mass for a few common fluorinated and iodinated mono-ol compounds that could be considered as core candidates from which dendrimers could be grafted. The incorporation of only three iodine atoms provides a mass defect of –0.490 u with respect to the averagine_{ud} trend line, near the ideal maximum value of –0.500 u. Conversely, the incorporation of as many as 15 fluorine atoms, in the case of perfluorooctanol, falls significantly short of the desired mass defect (only –0.193 u relative to averigine_{ud}). In fact, it would take 37 fluorine atoms to achieve the optimal mass defect of –0.5, but would therefore yield calibrants with an integer mass of 950 or greater, (e.g., perfluorononadecanol) limiting accurate calibration below this mass. In addition to the limited synthetic accessibility of such highly fluorinated compounds, their poor solubility in water and polar solvents limits their use for the internal calibration of peptides, and other highly fluorinated compounds (e.g., Ultramark) are known to be "sticky," exhibiting in-source persistence that results in carryover from one MS sample run to the next [37]. Therefore, 2,4,6-triiodophenol was selected as the core from which a polyester dendrimer would be grafted to generate a series of mass-defect-tuned calibrants.

Mass-Defect Calibrant Synthesis

The tris-iodinated core polyester dendrimers were prepared using an analogous synthetic procedure as described by Ihre et al. [28], and Gillies and Fréchet [30]. Reaction of triiodophenol with the acid anhydride monomer, **1**, resulted in the target "first generation" ester product, **2** (Scheme 1), the term "generation" referring to the number of synthetic iterations and therefore the number of layers of dendritic monomers. Acid-catalyzed hydrolysis of the acetonide protecting group was carried out selectively to yield the first generation diol, **3**, without any evidence for the hydrolysis of the phenolic ester. Repetition of the esterification reaction then yielded the second generation protected dendrimer, **4**, which was again deprotected in a nearly quantitative fashion to afford the second generation tetraol, **5**. These same steps were repeated to generate the third generation protected dendrimer, **6**, and its deprotected octaol, **7**. An additional iteration yielded the fourth generation protected dendrimer, **8**, and its deprotected hexadecanol, **9**.

Evaluation of Mass-Defect Calibrant

The purity of the dendrimers was initially assessed by MALDI-TOF MS using non-iodinated SpheriCal dendrimers as a calibrant. The resultant spectra exhibited only a single, well-resolved signal (Figure 4) corresponding closely to the expected sodium adducts, and all data were in close agreement with theoretical values (Table 2). Furthermore, the observed isotopic distribution for each calibrant was consistent with having a significant mass fraction of monoisotopic iodine. Most importantly for use as internal calibrants, the non-integer mass for each of the tris-iodinated dendrimers was offset by more than 0.40 u relative to the averagine trend line.

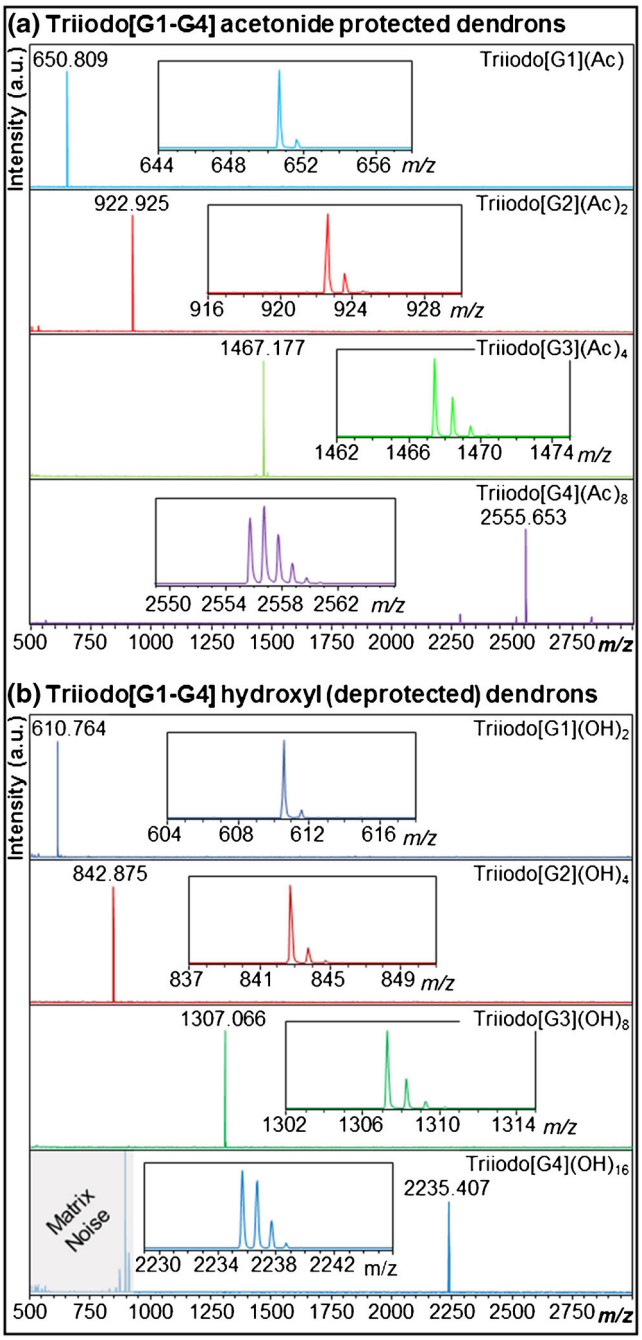


Figure 4. MALDI-TOF spectra for the isolated G1, G2, G3, and G4 tris-iodinated calibrants with inset to highlight the isotopic distribution for the **(a)** acetonide protected and **(b)** deprotected hydroxylated dendrimers

Peptide Characterization with Internal Calibrant

In order to initially evaluate these dendrimers as suitable candidates for internal calibration, peptides or synthetic macromolecules were selected that exhibit an ion of nearly identical integer mass in order to demonstrate the ability to resolve nearly overlapping signals from the analyte and the mass defect-tuned calibrant. Because the dendrimers did not incorporate amino groups or other easily protonated functional groups but did readily ionize via complexation with sodium cations, target analytes were selected that had nearly identical integer *m/z* when ionized as the sodiated dendrimer complexes. Endomorphin I was selected as the first demonstration because its protonated form (*m/z* = 611.298) exhibits the same nominal mass (and an adjacent integer mass) as the sodium adduct of first generation tris-iodinated dendrimer (*m/z* = 610.769). When mixed together with 2,3-dihydroxybenzoic acid (DHBA) as matrix, the resultant spectrum clearly shows distinct resolution between the two monoisotopic signals that are separated by approximately 0.635 u (Figure 5a). The M+1 signal for the calibrant differed by only 0.402 u with respect to the monoisotopic signal of the analyte, yet, as long as the sample conditions and instrument design enable resolution greater than 1000, this ~0.4 u non-integer mass offset should enable unambiguous identification of both the analyte and the calibrant. Furthermore, because the calibrant exhibits a substantially reduced M+1 isotopic signal (11.9%), relative to its monoisotopic signal, (compared with endomorphin with a 36.8% signal intensity for M+1), the identities of the two isotopic distributions can be easily differentiated as protein (orange) and tris-iodinated calibrant (blue).

Polymer Characterization with Internal Calibrant

Although the mass defect-tuned calibrants were designed specifically for internal calibration of peptides, most common synthetic polymers exhibit mass-defect trends similar to peptides, and therefore this calibrant can be equally useful for the characterization of many synthetic polymers. A dipropargyl poly(ethylene glycol) (PEG) molecule was prepared because the M+1 isotopic signal for the Na^+ adduct of its 27-mer (*m/z* = 1306.743) exhibits a mass within 1 u of the Na^+ adduct of the third generation tri-iodo calibrant (*m/z* = 1307.053). Although the mass defect difference in this case is only 0.311 u, again the two signals can be clearly distinguished from each other, and the characteristic isotopic distribution of the tris-iodinated calibrant (green) enables simple confirmation of its identity (Figure 5b). The application of these tris-iodinated calibrants has also been shown to have equal utility for ESI analysis, again yielding an unambiguously identification of the G3 calibrant and the analyte, despite having interspersed isotopic distributions (see Supplementary data Figure S3).

Protein Digest Characterization with Internal Calibrant

Finally, the set of mass defect-tuned standards was used as an internal calibrant for a trypsin digest of bovine serum albumin (BSA) using both MALDI and ESI instruments in order to highlight this calibrant's utility when dealing with a mixture of numerous peptides. As depicted in Figure 6a, the digest exhibited over 40 identifiable peptide fragments between *m/z* 500 and 2000, the structures of which were previously reported [38]. As demonstrated for the PEG example above, the calibrant signal at 1307.056 (Figure 6b), is clearly resolved from the peptide fragment at 1305.710 despite the fact that their isotopic distributions overlap. Again, the mass defect-tuned calibrant exhibits a much higher M/M+1 signal ratio, a consequence of three heavy monoisotopic iodine atoms. This enables

Table 2. Observed monoisotopic masses and calculated monoisotopic masses of all compounds and the observed errors (*all samples ionized with Na^+, except Endomorphin which was ionized via protonation)

	Chemical comp. (salt)	Calc. (u)	MALDI (u)	Δ (u)	ESI (u)	Δ (u)
Triiodo[G1](Ac)	$C_{14}H_{15}I_3O_4(Na^+)$	650.800	650.809	0.009	650.798	0.002
Triiodo[G2](Ac)$_2$	$C_{27}H_{35}I_3O_{10}(Na^+)$	922.926	922.925	0.001	922.929	0.003
Triiodo[G3](Ac)$_4$	$C_{53}H_{75}I_3O_{22}(Na^+)$	1467.178	1467.177	0.001	1467.182	0.004
Triiodo[G4](Ac)$_8$	$C_{105}H_{155}I_3O_{46}(Na^+)$	2555.682	2555.653	0.029	—	—
Triiodo[G1](OH)$_2$	$C_{11}H_{11}I_3O_4(Na^+)$	610.769	610.764	0.005	610.768	0.001
Triiodo[G2](OH)$_4$	$C_{21}H_{27}I_3O_{10}(Na^+)$	842.864	842.875	0.011	842.852	0.012
Triiodo[G3](OH)$_8$	$C_{41}H_{59}I_3O_{22}(Na^+)$	1307.053	1307.066	0.013	1307.055	0.002
Triiodo[G4](OH)$_{16}$	$C_{81}H_{123}I_3O_{46}(Na^+)$	2235.432	2235.407	0.025	—	—
Endomorphin I*	$C_{34}H_{39}N_6O_5(H^+)$	611.298	611.255	0.043	—	—
Dipropargyl PEG (27-mer)	$C_{60}H_{114}O_{28}(Na^+)$	1305.739	1305.752	0.013	1305.755	0.016

unambiguous identification of the calibrant signal relative to the analyte signal, in the case their overlapping signals might otherwise complicate their correct assignments.

Furthermore, when all of the peptide fragments are plotted with respect to their integer and non-integer masses, the contrast between traditional calibrants (Figure 6c) and the mass defect-tuned calibrant (Figure 6d) is clear. In the case of the typical peptide-based calibrants (gray), these compound masses directly overlap with the trend of peptide fragments, nearly tracing the averagine trend line, confirming that these peptides would be poor internal calibrants because of the likelihood of signal overlapping with one or more analyte. On the other hand, Ultramark 1621 [39] (yellow), Agilent tune mix [40] (green), and CsI ion clusters [5] (pink) all exhibit slopes that differ significantly from the averagine trend line (Figure 6c). Therefore, a number of these mass standards, particularly in the *m/z* = 1500–2100 range, will exhibit masses that are likely to overlap with peptide analytes in this same mass range. In stark contrast to these traditional calibrants, the mass defect-tuned calibration standards exhibit a trend that is nearly parallel to the averagine trend line, and very close to the ideal scarcine target, where the minimal number of peptides would fall (Figure 6d). Furthermore, across the 500–2000 mass range, each of the calibration compounds exhibits at least a 0.25 u non-integer mass offset relative to peptide fragments of a similar mass range, such that a mass spectrometer with resolving power >10,000 should easily differentiate the calibrants (separated at FWHM) from the analytes. In addition to the unique mass distributions these dendrimers exhibit, they demonstrate many of the same features of the previously reported dendrimer mass calibrants including broad compatibility with a range of solvents and matrices, as well as multi-year shelf-lives [19].

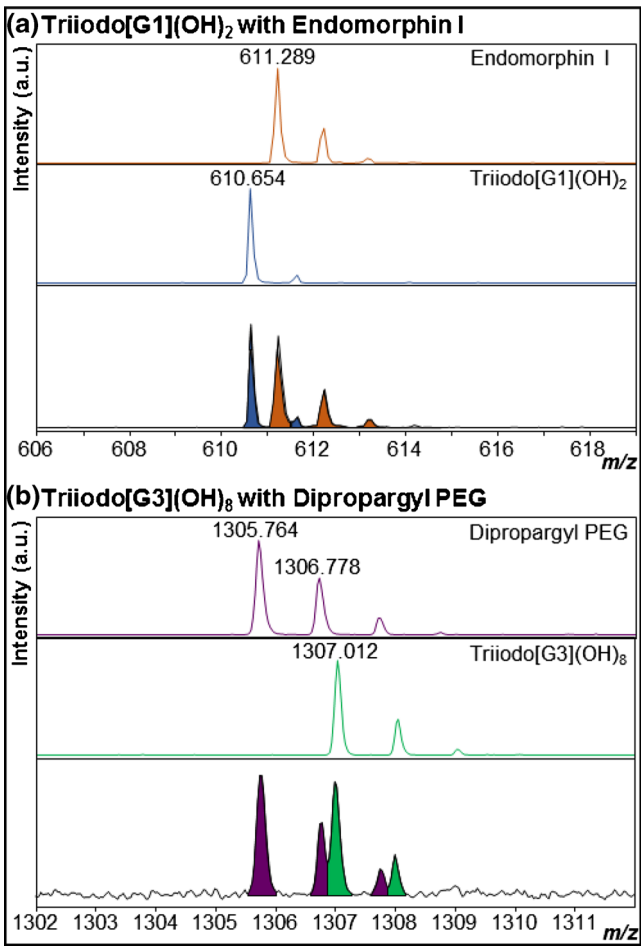


Figure 5. **(a)** MALDI-TOF mass spectra for the endomorphin I as $[M + H]^+$ (top, in orange), tris-iodo [G1](OH)$_2$ as $[M + Na]^+$ (middle, in blue), and their combined spectrum for internal calibration (bottom). **(b)** Spectra of dipropargyl PEG (27-mer) as $[M + Na]^+$ (top, in purple), tris-iodo [G3](OH)$_8$ as $[M + Na]^+$ (middle, in green), and their combined spectrum for internal calibration (bottom)

Conclusion

With the rapid development of mass-spectrometry-based profiling of biomedical samples, internal calibrants that can expedite the accurate identification of peptide disease-biomarkers are becoming increasingly valuable. By using an internal calibrant, the mass accuracy for peptide identification can be optimized, but problems of signal overlap between calibrants and the range of possible analytes arise. To address the issue of calibrant/analyte overlap, the population of possible peptide analytes has been modeled and the least populated non-integer mass regions across the range of integer masses have been determined and termed "scarcine." This computationally determined scarcine trend has been used to identify tris-iodinated dendrimers as an attractive target for mass defect-

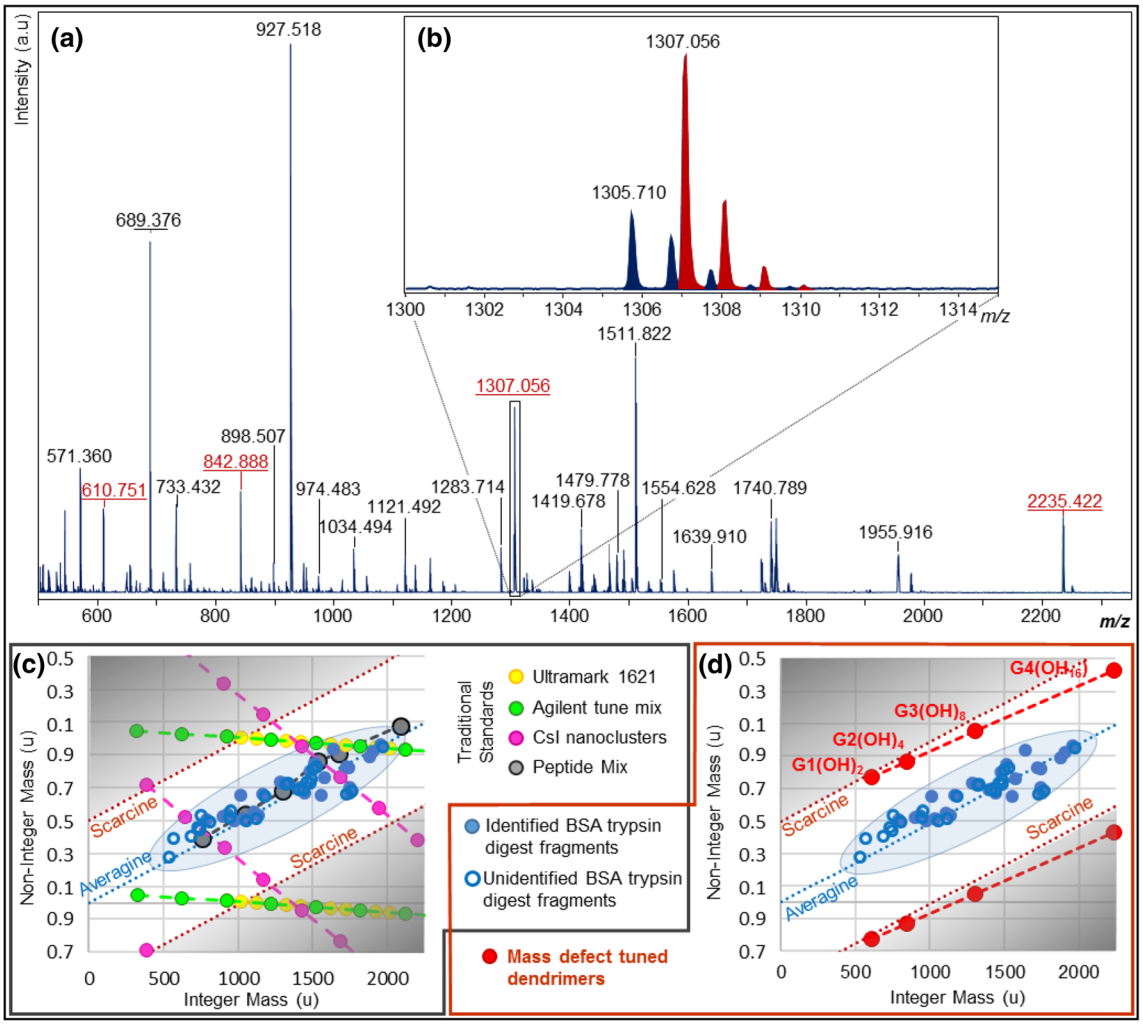


Figure 6. **(a)** MALDI-TOF mass spectrum of a BSA peptide digest with the entire calibrant set G1-G4 (red), and **(b)** inset of the triiodo [G3]-$(OH)_8$ calibrant (red) and BSA (frag 402-412) (blue). **(c)** The plot of the BSA digest mixture (blue) relative to peptide mix calibrants (gray), Ultramark 1621 (yellow), and CsI nanoclusters (pink) in positive mode. **(d)** The plot of BSA digest mixture relative to the mass defect-tuned dendrimers exhibiting no signal overlap. The peptide mix calibrant includes: bradykinin (frag 1-7), angiotensin II (human), angiotensin I, P14R, neurotensin, and TCTH (frag 1-17). The BSA digest included peptides that were identified in Giansanti P. et al. [38]

tuned calibrants. Furthermore, this peptide population model enables a quantitative visualization of the need for internal calibrants. With as many as 10^{20} possible peptides with an integer mass of 2000 and a variability of 0.01 u in non-integer mass, the improved mass accuracy made possible with internal calibration can substantially reduce the likelihood of analyte misidentification.

A set of tris-iodinated dendrimer calibrants were synthesized that exhibit an optimal mass defect offset to minimize calibrant/analyte signal overlap using both MALDI-TOF MS and ESI MS. Furthermore, this set of calibrants was evaluated against a trypsin protein digest, confirming this advantage of mass defect-tuned calibrants even when analyzing complex mixtures of peptides. In addition, the dendrimers' iodine-rich composition yields a unique isotopic distribution that can easily differentiate the calibrant signals from those of most biological or synthetic analytes. Because the number of analytes near the scarcine trend line increases with increasing integer mass, the utility of mass defect-tuned calibrants is most powerful for lower molecular weight peptides and protein fragments. However, in this lower mass range (600–2300), evaluations of the tris-iodinated dendrimers confirm their unique advantages as internal calibrants especially when an analyte exhibits an integer mass close to that of one of the calibrants.

Acknowledgments

Acknowledgment is made to the donors of the American Chemical Society Petroleum Research Fund, (53890-ND7) for partial support of this research, as well as the Louisiana Board of Regents for a graduate fellowship (J.A.G.), and Tulane University for a Carol Lavin Bernick Faculty Grant. S.M.G. thanks Drs. Alfred Yergey, P. Jane Gale, and Michael A. Grayson for helpful discussions.

References

1. Karas, M., Bachmann, D., Hillenkamp, F.: Influence of the wavelength in high-irradiance ultraviolet laser desorption mass spectrometry of organic molecules. Anal. Chem. **57**, 2935–2939 (1985)
2. Dole, M., Mack, L.L., Hines, R.L., Mobley, R.C., Ferguson, L.D., Alice, M.B.: Molecular beams of macroions. J. Chem. Phys. **49**, 2240–2249 (1968)
3. Yamashita, M., Fenn, J.B.: Electrospray ion source. Another variation on the free-jet theme. J. Phys. Chem. **88**, 4451–4459 (1984)
4. Schreimer, D.C., Li, L.: Detection of high molecular weight narrow polydisperse polymers up to 1.5 million Daltons by MALDI mass spectrometry. Anal. Chem. **68**, 2721–2725 (1996)
5. Pleasance, S., Thibault, P., Sim, P.G., Boyd, R.K.: Cesium iodide clusters as mass calibrants in ionspray mass spectrometry. Rapid Commun. Mass Spectrom. **5**, 307–308 (1991)
6. Anacleto, J.F., Pleasance, S., Boyd, R.K.: Calibration of ion spray mass spectra using cluster ions. Org. Mass Spectrom. **27**, 660–666 (1992)
7. Hop, C.E.C.A.: Generation of high molecular weight cluster ions by electrospray ionization: implications for mass calibration. J. Mass Spectrom. **31**, 1314–1316 (1996)
8. Moini, M., Jones, B.L., Rogers, R.M., Jiang, L.: Sodium trifluoroacetate as a tune/calibration compound for positive- and negative-ion electrospray ionization mass spectrometry in the mass range of 100–4000 Da. J. Am. Soc. Mass Spectrom. **9**, 977–980 (1998)
9. Konig, S., Fales, H.M.: Calibration of mass ranges up to *m/z* 10,000 in electrospray mass spectrometers. J. Am. Soc. Mass Spectrom. **10**, 273–276 (1999)
10. Kolářová, L., Prokeš, L., Kučera, L., Hampl, A., Peňa-Méndez, E., Vaňhara, P., Havel, J.: Clusters of monoisotopic elements for calibration of (TOF) mass spectrometry. J. Am. Soc. Mass Spectrom. **28**, 419–427 (2017)
11. McEwen, C.N., Larsen, B.S.: Accurate mass measurement of proteins using electrospray ionization on a magnetic sector instrument. Rapid Commun. Mass Spectrom. **6**, 173–178 (1992)
12. Cody, R.B., Tamura, J., Musselman, B.D.: Electrospray ionization/magnetic sector mass spectrometry: Calibration, resolution, and accurate mass measurements. Anal. Chem. **64**, 1561–1570 (1992)
13. Hao, P., Adav, S.S., Gallart-Palau, X., Sze, S.K.: Recent advances in mass spectrometric analysis of protein deamidation. Mass Spectrom. Rev. **36**, 677–692 (2017)
14. Chen, M., Cook, K.D.: Oxidation artifacts in the electrospray mass spectrometry of Aß peptide. Anal. Chem. **79**, 2031–2036 (2007)
15. Rombouts, I., Lagrain, B., Scherf, K.A., Lambrecht, M.A., Koehler, P., Delcour, J.A.: Formation and reshuffling of disulfide bonds in bovine serum albumin demonstrated using tandem mass spectrometry with collision-induced and electron-transfer dissociation. Sci. Rep. **5**, 12210 (2015)
16. Feng, R., Konishi, Y., Bell, A.W.: High accuracy molecular weight determination and variation characterization of proteins up to 80 ku by ion mass spectrometry. J. Am. Soc. Mass Spectrom. **2**, 387–401 (1991)
17. Lai, M.C., Topp, E.M.: Solid state chemical stability of proteins and peptides. J. Pharm. Sci. **88**, 489–500 (1999)
18. Garofolo, F.: LC-MS instrument calibration. In: Chan, C.C., Lee, Y.C., Lam, H., Zhang, X.-M. (eds.) Analytical method validation and instrument performance verification. John Wiley Inc., Hoboken, NJ (2004)
19. Grayson, S.M., Myers, B.K., Bengtsson, J., Malkoch, M.: Advantages of monodisperse and chemically robust "SpheriCal" polyester dendrimers as a "universal" MS calibrant. J. Am. Soc. Mass Spectrom. **25**, 303–309 (2014)
20. Casey, B.K., Grayson, S.M.: The potential of amine-containing dendrimer mass standards for internal calibration of peptides. Eur. J. Mass Spectrom. **21**, 747–752 (2015)
21. Gross, J.H.: Improved procedure for dendrimer-based mass calibration in matrix-assisted laser desorption/ionization-time-of-flight-mass spectrometry. Anal. Bioanal. Chem. **408**, 5945–5951 (2016)
22. Pourshahian, S.: Mass defect from nuclear physics to mass spectral analysis. J. Am. Soc. Mass Spectrom. **28**, 1836–1843 (2017)
23. Yao, X., Diego, P., Ramos, A.A., Shi, Y.: Averagine-scaling analysis and fragment ion mass defect labeling in peptide mass spectrometry. Anal. Chem. **80**, 7383–7391 (2008)
24. Bajrami, B., Shi, Y., Lapierre, P., Yao, X.: Shifting unoccupied spectral space in mass spectrum of peptide fragment ions. J. Am. Soc. Mass Spectrom. **20**, 2124–2134 (2009)
25. Fishman, V.N., Linclau, B., Curran, D.P., Somayajula, K.V.: Tris(perfluoroalkylethyl)silyl alkyl amines as calibration standards for electron ionization mass spectrometry in the mass range of 100–3000 Da. J. Am. Soc. Mass Spectrom. **12**, 1050–1054 (2001)
26. Shi, Y., Bajrami, B., Yao, X.: Passive and active fragment ion mass defect labeling: distinct proteomic potential of iodine-based reagents. Anal. Chem. **81**, 6438–6448 (2009)
27. Ihre, H., Hult, A., Fréchet, J.M.J., Gitsov, I.: Double-stage convergent approach for the synthesis of functionalized dendritic aliphatic polyesters based on 2,2-bis(hydroxymethyl)propionic acid. Macromolecules. **31**, 4061–4068 (1998)
28. Ihre, H., Padilla de Jesus, O.L., Fréchet, J.M.J.: Fast and convenient divergent synthesis of aliphatic ester dendrimers by anhydride coupling. J. Am. Chem. Soc. **123**, 5908–5917 (2001)
29. Malkoch, M., Malmström, E., Hult, A.: Rapid and efficient synthesis of aliphatic ester dendrons and dendrimers. Macromolecules. **35**, 8307–8317 (2002)
30. Gillies, E.R., Fréchet, J.M.J.: Designing macromolecules for therapeutic applications: polyester dendrimers-poly(ethylene oxide) "bow-tie" hybrids with tunable molecular weight and architecture. J. Am. Chem. Soc. **124**, 14137–14146 (2002)
31. Zubarev, R.A., Bondarenko, P.V.: An a priori relationship between the average and monoisotopic masses of peptides and oligonucleotides. Rapid Commun. Mass Spectrom. **5**, 276–277 (1991)
32. Senko, M.W., Beu, S.C., McLafferty, F.W.: Determination of monoisotopic masses and ion populations for large biomolecules from resolved isotopic distributions. J. Am. Soc. Mass Spectrom. **6**, 229–233 (1995)
33. Hall, M.P., Ashrafi, S., Obegi, I., Petesch, R., Peterson, J.N., Schneider, L.V.: "Mass defect" tags for biomolecular mass spectrometry. J. Mass Spectrom. **38**, 809–816 (2003)
34. Sleno, L.: The use of mass defect in modern mass spectrometry. J. Mass Spectrom. **47**, 226–236 (2012)
35. Hernandez, H., Niehauser, S., Boltz, S.A., Gawandi, V., Phillips, R.S., Amster, I.J.: Mass defect labeling of cysteine for improving peptide assignment in shotgun proteomic analyses. Anal. Chem. **78**, 3417–3423 (2006)
36. Mirzaei, H., Brusniak, M.-Y., Mueller, L.N., Letarte, S., Watts, J.D., Aebersold, R.: Halogenated peptides as internal standards (H-PINS): an introduction of an MS-based internal standard set for liquid chromatography-mass spectrometry. Mol. Cell. Proteomics. **8**, 1934–1946 (2009)
37. Cody, M.: Electrospray ionization mass spectrometry: history, theory, and instrumentation. In: Pramanik, B.N., Ganguly, A.K., Gross, M.L. (eds.) Applied Electrospray mass spectrometry. Marcel Dekker, New York, NY (2002)
38. Giansanti, P., Tsiatsiani, L., Low, T.Y., Heck, A.J.R.: Six alternative proteases for mass spectrometry-based proteomics beyond trypsin. Nat. Protoc. **11**, 993–1006 (2016)
39. Moini, M.: Ultramark-1621 as a calibrant reference compound for mass spectrometry 2. Positive-ion and negative-ion electrospray-ionization. Rapid Commun. Mass Spectrom. **9**, 711–714 (1994)
40. Deng, L., Ibrahim, Y.M., Hamid, A.M., Garimella, S.V.B., Webb, I.K., Zheng, X., Prost, S.A., Sandoval, J.A., Norheim, R.V., Anderon, G.A., Tolmachev, A.V., Baker, E.S., Smith, R.D.: Ultra-high resolution ion mobility separations utilizing traveling waves in a 13 m serpentine path length structure for lossless ion manipulations module. Anal. Chem. **88**, 8957–8964 (2016)