Supporting Information for

Quantitative Proteomics: Measuring Protein Synthesis Using ^{15}N Amino Acids Labeling in Pancreas Cancer Cells

Yingchun Zhao, † Wai-Nang Paul Lee, ‡ Shu Lim, ‡ Vay Liang Go, ‡ Jing Xiao, † Rui Cao, † Hengwei Zhang, † Robert Recker, † and Gary Guishan Xiao*,†

Genomics & Functional Proteomics Laboratories, Osteoporosis Research Center, Creighton University Medical Center, 601 N 30th Street, Suite 6730, Omaha, NE 68131, Metabolomics Core, UCLA Center of Excellence in Pancreatic Diseases, Harbor-UCLA Medical Center, Torrance, CA 90502

*To whom correspondence should be addressed: E-mail: gxiao@creighton.edu, Fax: 402-280-4284.

† Creighton University.

‡ University of California at Los Angeles.

BRIEF SUMMARY

In this supplemental information section we present briefly the theory of the concatenation model with an illustration. The data deduction method for the determination of fractional protein synthesis is demonstrated using experimental data from which results shown in Figure 3 are based.

Quantitative analysis of mass spectrum would be simple if there were no naturally existing isotopes. The existence of isotopes of carbon (^{13}C), hydrogen (^{2}H), nitrogen (^{15}N), oxygen (^{18}O) and sulfur (^{33}S, ^{34}S) presents challenges to the interpretation of mass spectrum and opportunities for using these isotopes for biological investigations. The presence of different amount of isotopes in molecules of a protein generates molecules with different masses. These molecules of the same compound with different molecular weights form the family of mass isotopomers. The distribution of mass isotopomers gives rise to the appearance of the isotope envelop observed in a peptide spectrum.

Isotopomer distribution of a peptide is predominantly influenced by the number of carbon, nitrogen and hydrogen atoms in the molecule and the isotopic abundance of each species. In a peptide where the natural abundances of ^{15}N, ^{18}O and ^{2}H are relatively low compared to that of ^{13}C, the isotopomer distribution can be approximated by a binomial distribution governed by N_C and p, the number of carbon atoms and its natural abundance. The individual relative intensity of isotopomer (n) is given by

$$M(n)= (N_C!/n!(N_C-n)!)p^n(1-p)^{Nc-n}. \qquad \text{Equation 1}$$

In this formula, n stands for the number of ^{13}C atoms and N_C stands for the total number of carbon atoms in the peptide. The individual intensity is given by $M_0 = (1-p)^N$, $M_1= Np\times(1-p)^{N-1}$, etc. ; and the sum of these individual intensities is equal to 1. This distribution can be determined experimentally by normalizing individual intensity by the total intensity such that the sum of normalized intensities equals to 1. The normalized intensity is given the name of RIA (relative isotopologue abundance) in the paper by Vogt et al. [1;2] A consequence of a normalized distribution is that the sum of product of $M(n)\times n$, which we defined as ΣMn,[3] has the meaning of average mass. When the

contribution of natural abundance is removed, ΣMn also has the same meaning as #isotope atom/molecule, the equivalent of specific activity. For a binomial distribution,

$\Sigma Mn = N \times p$ Equation 2

When a peptide has two populations of isotopes ^{13}C and ^{15}N as in the case of the present study, the distribution of isotopomer is not given by a binomial distribution. When the presence or absence of ^{13}C is independent of the presence or absence of ^{15}N and vice versa, the distribution is the concatenation of isotopomers from these stable isotope species.[4] That is, the observed isotopomer distribution of the peptide is the concatenation of a binomial distribution of ^{13}C isotopomers governed by (N_C, p) and a binomial distribution of ^{15}N isotopomers governed by (N_N, p'). The mathematical operation of the concatenation function is as follows: If C-distribution is (a_0, a_1, a_2, a_3,…, a_n), N-distribution is (b_0, b_1, b_2, b_3,…, b_n), and combined distribution is (c_0, c_1, c_2, c_3,…, c_n), then newly synthesized distribution using concatenation operation is $c_0 = a_0b_0$; $c_1 = a_0b_1 + a_1b_0$; $c_2 = a_0b_2 + a_1b_1 + a_2b_0$; $c_3 = a_0b_3 + a_1b_2 + a_2b_1 + a_3b_0$ [4]. As a result, the average mass of the peptide is the sum of average mass of these ^{13}C and ^{15}N distributions:

$\Sigma Mn = N_C \times p + N_N \times p'$ Equation 3

The combined distribution of a peptide containing ^{13}C and ^{15}N isotopomers and its implication on changes in average mass (mass shift) is illustrated in Figure S-1. The ^{13}C isotopomer distribution of a hypothetical natural peptide and ^{15}N isotopomer distribution of a labeled peptide are typical binomial distribution (Fig. S-1). The mean and variance of the ^{13}C isotopomer distribution around the monoisotopic peak (m_0) are N_Cp and $N_Cp(1-p)$ respectively, where N_c is the number of carbon atoms in the peptide, and p is the natural abundance of ^{13}C. Similarly, the mean and variance of the ^{15}N isotopomer distribution are

$N_N p'$ and $N_N p'(1-p')$ respectively and p' is the average ^{15}N enrichment. The isotopomer distribution after concatenation is also shown in Figure S-1. The mean and variance of the new distribution is given by $(N_C p + N_N p')$ and $[N_C p(1-p) + N_N p'(1-p')]$ respectively. The mass shift of the new distribution is $[(N_C p + N_N p') - N_C p] = N_N p'$. Therefore, the mass shift is a function of the ^{15}N isotopomer distribution.

The observed spectrum of a peptide (both preexisting and newly synthesized) is then linear combination of isotopomers of natural (unlabeled) and the expected labeled peptides, and the ratio of labeled to the total isotopomers provides a molar fraction of the newly synthesized protein.

$$\Sigma Mn = N_C \times p + FSR \times (N_N \times p') \qquad \text{Equation 4}$$

$$\Delta \text{average mass} = \Sigma Mn - N_C \times p = FSR \times (N_N \times p') \qquad \text{Equation 5}$$

Since $N_C \times p$ is given by the unlabeled peptide, FSR can be determined when $N_N \times p'$ is known. $N_N \times p'$ can be determined in peptide whose FSR is 1 (100%). In the method of Vogt et al using $[U^{13}C_6]$-glucose as the source of isotope [2], the theoretical change in average mass is $\Sigma N_i \Delta k_i$, where N_i is the number of amino acid (i) and Δk_i is the change in average mass of that amino acid (i) after labeling. Δk_i is determined from peptides with FSR close to 100% and their amino acid sequences.

Because N_N can be determined from the known peptide sequence, the ^{15}N enrichment (p') can be determined by curve fitting (mass shift/N_N). The mass shift is first obtained by inspection. Its value is changed by small increment or decrement to arrive at a new p' such that the ratio of the theoretical M(i) based on the binomial distribution (N_N, p') to the observed M(i) in regions with no overlap from the unlabeled spectrum is constant. Once the natural isotopomers and the ^{15}N isotopomers are known, the isotopomers of the

^{13}C and ^{15}N labeled peptide can be constructed using concatenation operation. The resultant distribution is the expected distribution of the new peptide and can be used to determine newly synthesized fraction. The data processing algorithm is illustrated in Figure 3. First, the intensity distribution of individual peaks in the peptide without labeling (control) was normalized (Fig 3A). After normalization, formula $M_1/M_0=N\times p/(1-p)$ ($p = 0.0111$) was used to calculate the carbon atoms by assuming ^{13}C natural abundance to be 1.11%. Once the N_C and p are known, a binomial distribution was set up following the formula: intensity of isotopomer (n) = $(N_C!/n!(N_C-n)!)p^n(1-p)^{Nc-n}$. In this formula, n stands for the number of ^{13}C atoms and N_C stands for the total number of carbon atoms in the peptide. Example of such computation is shown in Table S-1. The near perfect fit between theoretical and experimental values are shown in Figure 3A. The theoretical carbon number is somewhat larger than the number calculated from peptide sequence due to other minor natural enrichments, like ^{2}H, ^{18}O, ^{15}N, and ^{33}S. In fact, the natural distribution is a concatenation of all these distributions. The approximation of the observed distribution by the theoretical distribution of ^{13}C suggests that the mass isotopomer distribution is predominantly influenced by natural abundance of ^{13}C (Fig. 3A) and validate the use of binomial model for interpretation of peptide spectrum.

Both 50% and 33% of artificial enrichment of ^{15}N in the medium caused obvious mass shift in spectrum distribution. The mass shift of the labeled spectrum can be determined from the change in average mass after subtracting the natural spectrum (Fig. 3B). Based on the mass shift and the number of nitrogen atoms in the specific fragment (from sequence information), the average ^{15}N enrichment can be deduced. After the N_N and p'

are known, the theoretical ^{15}N isotopomer distribution can be generated based on a binomial distribution function (Tables S-2 and S-3). The concatenation of ^{13}C (Table S-1) and ^{15}N distributions (Tables S-2 and S-3) represents isotopomer distribution of the newly synthesized peptide. The isotopomer distribution after ^{15}N labeling which is represented by the concatenated distribution is shown in Figures 3C (50%) and 3D (33%). It is important to note that the sum of all isotopomers in any distribution, C-isotopomer, N-isotopomer or the concatenated isotopomer, is equal to 1.

The observed isotopomer distribution of a peptide is the weighted sum of isotopomers of the natural (preexisting) and the labeled (new) peptide (Fig. 3C&D). By multiple linear regression analysis using the observed distribution as the dependent variable and the preexisting and newly synthesized parts as the independent variables, the contribution of each of pre-existing and new peptides can be determined.

Table S-1. The unlabeled spectrum of 1699 m/z in spot 6 can be simulated by a binomial distribution.

	peak intensity	Normalized	M_1/M_0	carbon # [a)]	Theoretical distribution [b)]
M_0	57603.5	0.347	1.044	94.1	0.350
M_1	60153.5	0.363			0.370
M_2	31848.5	0.192			0.193
M_3	11834.5	0.071			0.066
M_4	3415.5	0.021			0.017
M_5	749.5	4.5E-03			3.4E-03
M_6	198.5	1.2E-03			5.7E-04
M_7	82.5	5.0E-04			8.0E-05

a) The carbon number calculation was based on the formula $M_1/M_0=N\times p/(1-p)$, p=0.0111.

b) Binomial distribution was based on N=94, p=0.0111.

Table S-2. Concatenation analysis of newly synthesized protein based on peptide spectrum from 50% ^{15}N labeling.

m/z	Peak intensity	Normalize	^{15}N binomial [a]	Concatenation process [b]				Expected Newly synthesized [c]
M_0	31808.5	0.096	2.0E-04	7.0E-05				7E-05
M_1	36121.5	0.109	0.002	0.001	7.3E-05			8.0E-04
M_2	19490.5	0.059	0.011	0.004	0.001	3.8E-05		0.004
M_3	9941.5	0.030	0.033	0.012	0.004	4.0E-04	1.4E-05	0.016
M_4	11476.5	0.035	0.075	0.026	0.012	0.002	1.5E-04	4.1E-06	0.040
M_5	20352.5	0.062	0.127	0.044	0.027	0.006	0.001	4.3E-05	0.079
M_6	30234.5	0.092	0.170	0.059	0.046	0.014	0.002	2.2E-04	0.122
M_7	34092.5	0.103	0.182	0.063	0.062	0.024	0.005	0.001	0.155
M_8	33282.5	0.101	0.159	0.055	0.066	0.033	0.009	0.002	0.165
M_9	32893.5	0.100	0.115	0.040	0.058	0.035	0.012	0.003	0.148
M_{10}	27913.5	0.084	0.069	0.024	0.042	0.031	0.013	0.003	0.113
M_{11}	20675.5	0.063	0.035	0.012	0.025	0.022	0.011	0.004	0.075
M_{12}	13205.5	0.040	0.014	0.005	0.013	0.013	0.008	0.003	0.043
M_{13}	6398.5	0.019	0.005	0.002	0.005	0.007	0.005	0.002	0.022
M_{14}	2392.5	0.007	0.001	4.9E-04	0.002	0.003	0.002	0.001	0.010
M_{15}	70.5	2.1E-04	3.3E-04	1.2E-04	0.001	0.001	0.001	0.001	0.004

a) ^{15}N binomial distribution was based on N=21 (from peptide chemical structure), p =0.333 (calculated from equation (mass shift=Np). For mass shift of 7, p= 7/21=0.333).

b) Concatenation was done based on ^{13}C (Table S-1) and ^{15}N binomial distributions.

c) The calculation of newly synthesized protein followed the instruction in Footnote 2. The spectrum of this newly synthesized peptide is shown in Figure 3 Panel C.

Table S-3. Concatenation analysis of newly synthesized part based on peptide spectrum from 33% ^{15}N labeling[a)].

m/z	Peak intensity	Normalize	^{15}N binomial	Concatenation process				Expected Newly synthesized
M_0	29743	0.089	0.006	0.002				0.002
M_1	34353	0.102	0.036	0.013	0.002			0.015
M_2	28029	0.083	0.099	0.034	0.013	0.001		0.049
M_3	29876	0.089	0.170	0.059	0.036	0.007	4.5E-04	0.102
M_4	41305	0.123	0.209	0.073	0.062	0.019	0.003	1.3E-04	0.156
M_5	46585	0.139	0.194	0.067	0.076	0.033	0.007	7.4E-04	0.184
M_6	42364	0.126	0.141	0.049	0.070	0.040	0.012	0.002	0.174
M_7	31863	0.095	0.082	0.029	0.051	0.037	0.015	0.004	0.136
M_8	23328	0.069	0.039	0.014	0.030	0.027	0.014	0.004	0.090
M_9	15770	0.047	0.016	0.005	0.014	0.016	0.010	0.004	0.051
M_{10}	9021	0.027	0.005	0.002	0.006	0.008	0.006	0.003	0.025
M_{11}	3545	0.011	0.001	4.8E-04	0.002	0.003	0.003	0.002	0.011
M_{12}	108	3.2E-04	3.1E-04	1.1E-04	0.001	0.001	0.001	0.001	0.004
M_{13}	59	1.8E-04	5.9E-05	2.1E-05	1.1E-04	2.7E-04	3.6E-04	3.2E-04	0.001

a) The calculation is the same as in Table S-2 except mass shift being 4.5. The result of the newly synthesized peptide is shown in Figure 3 Panel D.

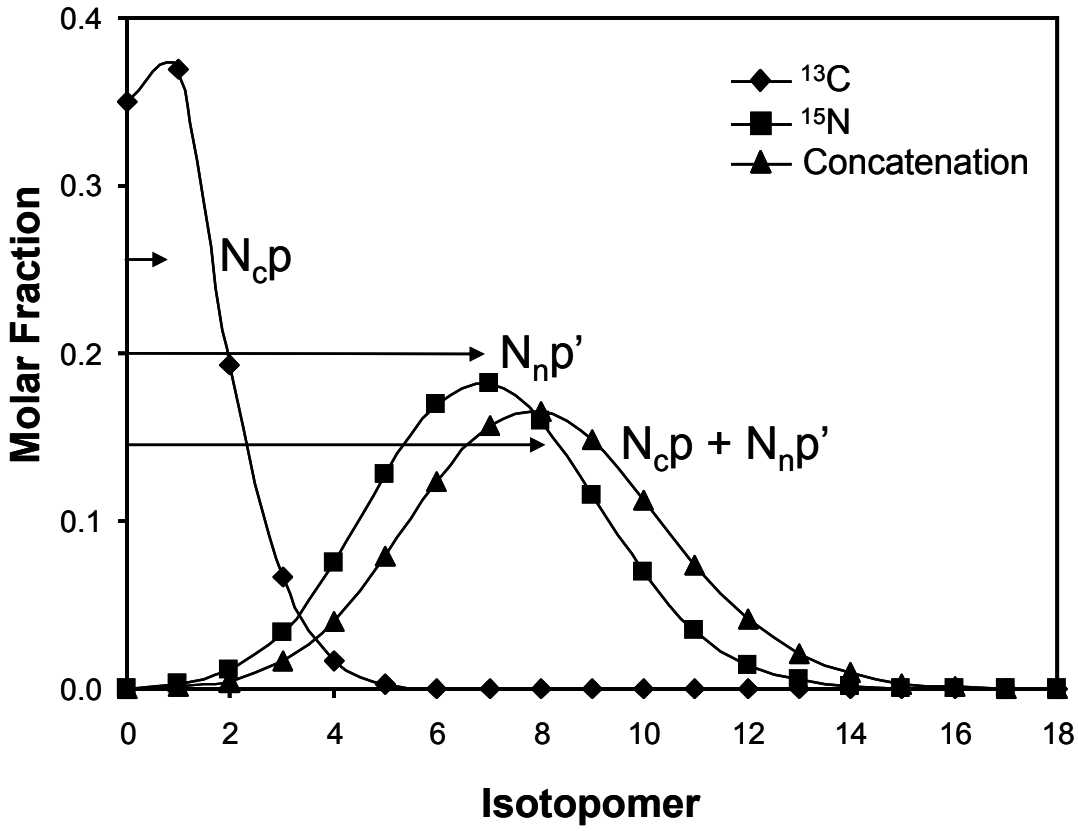

Figure S-1. The isotope envelops of three theoretical mass isotopomer distributions showing effect of concatenation. The first curve is that of the distribution of ^{13}C isotopomers calculated from a binomial distribution assuming carbon number to be 94 and natural enrichment of ^{13}C to be 0.0111. The second curve represents the distribution of ^{15}N isotopomers with nitrogen number of 21 and enrichment of 0.333. The third curve is that of concatenation of these two distributions representing the newly synthesized protein after incorporation of ^{15}N. The sum of molar fractions for each curve is 1.

References

(1) Vogt, J. A.; Schroer, K.; Holzer, K.; Hunzinger, C.; Klemm, M.; Biefang-Arndt, K.; Schillo, S.; Cahill, M. A.; Schrattenholz, A.; Matthies, H.; Stegmann, W. *Rapid Commun. Mass Spectrom.* **2003,** *17*, 1273-82.

(2) Vogt, J. A.; Hunzinger, C.; Schroer, K.; Holzer, K.; Bauer, A.; Schrattenholz, A.; Cahill, M. A.; Schillo, S.; Schwall, G.; Stegmann, W.; Albuszies, G. *Anal. Chem.* **2005,** *77*, 2034-42.

(3) Lee, W. N.; Bassilian, S.; Ajie, H. O.; Schoeller, D. A.; Edmond, J.; Bergner, E. A.; Byerley, L. O. *Am. J. Physiol*. **1994,** *266*, E699-E708.

(4) Xiao, G. G.; Garg, M.; Lim, S.; Wong, D.; Go, V. L.; Lee, W. N. *J. Appl. Physiol.* **2008,** *104*, 828-36.