

アミノ酸組成によるタンパク質同定手法の 精度向上に関する研究

古田 舜*・熊本沙也加*・荒木朋洋**

Study on improvement for protein identification method by amino acid composition

Shun FURUTA, Sayaka KUMAMOTO and Tomohiro ARAKI

(Received 29 November 2018; accepted 18 January 2019)

ABSTRACT

Protein identification using amino acid composition analysis identifies proteins by amino acid composition delivered from hydrolyzed protein. In hydrolysis, Cys and Trp can not be recovered because the decomposition of their structures. In addition, Ile-Ile bonding, Ile-Val bonding and Ile-Leu bonding cause the low recovery rate of amino acid. These amino acids cause the decrease of the quality of protein identification. Therefore, improvement of identification was required by using Cys, Trp and Ile factors for protein identification. As a result, we established the improvement method for protein identification using by corrected value of Cys, Trp and Ile.

緒言

プロテオーム解析とは生体内で発現されるタンパク質を網羅的に解析することである (1)。生体内のタンパク質は遺伝情報から作り出されている。この全遺伝情報をゲノムといい、ゲノムから発現されるタンパク質の総称をプロテオームと呼ぶ。ゲノムは同じ生物種であれば同じ遺伝子情報を持つが、プロテオームは時期や生育環境などの違いによって発現するタンパク質が変化する。このことからプロテオーム解析はそれぞれの条件下で発現しているタンパク質を網羅的に同定する (2)。プロテオーム解析におけるタンパク質の同定は一般的に質量分析計を用いて行われる。生物の組織または細胞のタンパク質を 2D-PAGE により分離し (3)、ディファレンシャルディスプレイによって検出された特異的なタンパク質を同定する (4, 5)。

タンパク質はゲル中で酵素消化後、得られたペプチド断片を質量分析計で検出し、ペプチドマスフィンガープリント法 (PMF 解析) やデノボシーケンス法などの解析法を用いて同定する (6-9)。PMF 解析は、酵素消

化後のタンパク質から得られたペプチド断片を質量分析計で検出した際の質量スペクトルをデータベース化された理論的な質量スペクトルの質量パターンと照合、比較する方法である (10)。この方法はハイスループットなタンパク質同定法として汎用されるが、ゲノム未知生物由来タンパク質のようなデータベース上に存在しないタンパク質は同定が困難である。そのような場合には、検出したペプチドは二次マス (MS/MS) 解析を行い、得られたフラグメントイオンの質量差からアミノ酸配列情報を得るデノボシーケンス法で、得られたアミノ酸配列情報を用いて配列相同性からタンパク質を同定する。これらの解析法はプロテオミクスにおいて汎用されているが、前処理が煩雑であることや、機器やその維持費が高額である。一方、アミノ酸組成分析計はタンパク質を加水分解したのち、アミノ酸組成やアミノ酸含有量を正確に定量することが可能である (11)。本手法の特徴は質量分析とは同定の際に用いる情報が大きく異なっているという点である。質量分析計を用いた同定法は部分アミノ酸配列情報を用いて同定を行うが、アミノ酸組成を用いた同定手法はアミノ酸組成比情報を用いる。その情報の性質の違いから新しい視点からのタンパク質同定が期待される。この同定手法はプロテオーム解析に一般的に活用されていないが、プロテオーム解析での有意性が見出されている (12)。また、Swiss Institute of

*東海大学大学院農学研究科

** 東海大学農学部バイオサイエンス学科

連絡先 荒木朋洋 e-mail:araki@agri.u-tokai.ac.jp

Bioinformatics のポータルサイト ExPASy には AACCompIdent (13) というアミノ酸組成比を用いたタンパク質同定ツールが公開されており、このツールを用いた研究もいくつか報告されている (14, 15)。本研究室でもアミノ酸組成分析によるタンパク質同定システム (Ver 1.2; 特開2010-96501) を開発した。以上のようにプロテオーム解析におけるアミノ酸組成分析の利用はこれまでいくつか報告されている (16)。

本研究はより簡便なアミノ酸組成分析を用いたタンパク質同定手法の確立を目的とし、精度評価と本手法の精度向上を検討した。現在の手法は加水分解の影響により脱アミド化されそれぞれ Asp と Glu になる Asn や Gln, その構造が破壊される Cys や Trp を除く、全20種類中16種類のアミノ酸で行っている。また、タンパク質配列中の Ile-Ile 結合, Ile-Val 結合及び Ile-Leu 結合は加水分解を受けにくく、アミノ酸組成分析において Ile 回収率が低下すると報告されており、タンパク質同定において同定精度を低下させている。そこでタンパク質同定に Cys と Trp の情報を加えると共に、Ile の回収率を補正することができればタンパク質同定精度が向上すると考えられる。そこで、Cys, Trp を含む18種のアミノ酸及び回収率を補正した Ile の情報を用いてタンパク質同定精度の向上を検討した。

材料および方法

材料

本研究に用いたニワトリ卵白由来 C-type lysozyme (hen egg lysozyme Code No. 100940) は生化学工業株式会社から購入し、本研究室で CM イオン交換クロマトグラフィーにてリクロマトし、凍結乾燥したものを使用した。ガチョウ卵白由来 G-type lysozyme (goose egg lysozyme) は、本研究室で遺伝子を合成して発現、精製を行ったリコンビナント G-type lysozyme を使用した。大豆由来 trypsin inhibitor (Code No. 106K7055), ウシ由来 serum albumin (Code No. 114K7545) はシグマ アルドリッチ ジャパンから購入したものを使用した。

リコンビナント G-type lysozyme の作製と精製

リコンビナント G-type lysozyme の遺伝子合成はアミノ酸配列をもとに PCR 法で行い、リコンビナント G-type lysozyme の全体遺伝子を作製した。その後、BMGY 培地を用いてタンパク質発現を行った。発現したタンパク質は CM- トヨパールカラムクロマトグラフィーで精製し凍結乾燥して使用した。

試薬

還元剤 β -メルカプトエタノール (β -ME) は GE ヘルスケアから購入した。その他の試薬は市販の特級または生化学用試薬を用いた。

試料の調製

各タンパク質試料は精製標品を 1 mg/ml になるように蒸留水に溶解し、調製した。

SDS- ポリアクリルアミドゲル電気泳動 (SDS-PAGE)

Laemmli らの方法にしたがって SDS-PAGE を行った (17)。すなわち、分離ゲル用緩衝液 (A 液) は、トリス (ヒドロキシメチル) アミノメタン 18.2 g, 1 M 塩酸 24 ml, ドデシル硫酸ナトリウム (SDS) 0.4 g を蒸留水に溶解し、pH を 8.8 に調整後、全容を 100 ml にフィルアップした。濃縮ゲル用緩衝液 (B 液) は、トリス (ヒドロキシメチル) アミノメタン 6.0 g, 1 M 塩酸 48 ml, SDS 0.4 g を蒸留水に溶解し、pH 6.8 に調整後、全容を 100 ml にフィルアップした。アクリルアミド/ビス混合液 (C 液) はアクリルアミド (モノマー) 30 g, N,N'-メチレンビスアクリルアミド 0.8 g を蒸留水に溶解後、全容を 100 ml にフィルアップした。10% (w/v) 過硫酸アンモニウム (APS) 溶液 (D 液) は過硫酸アンモニウム 0.1 g を蒸留水 1 ml で溶解した。SDS 試料調製溶液はグリセリン 2.0 ml, 濃縮ゲル用緩衝液 2.5 ml, β -ME 1 ml, SDS 0.4 g を加え、さらに 0.2% プロモフェノールブルー (BPB) を含むメタノール溶液を 0.1 ml 加え、蒸留水で 10 ml にフィルアップした。電気泳動用緩衝液はトリス (ヒドロキシメチル) アミノメタン 6.0 g, SDS 1 g, Gly 28.8 g を蒸留水に溶解し、0.2% BPB を含むメタノール溶液 0.1 ml を加え蒸留水で 1000 ml にフィルアップした。染色はコマジープリリアントブルー R-250 (CBB) を用いた。泳動はミニスラブ電気泳動装置 (AE-6440, ATTO 製) を使用した。

トランスブロッティング

トランスブロッティングはセミドライ式で行った。ブロッティング緩衝液は 3-シクロヘキシル-1-プロパンスルホン酸 (CAPS) 緩衝液を用いた。膜は、Poly Vinylidene Difluoride (PVDF 膜, Bio-Rad 製) を使用しセミドライ式ブロッティング装置 (ホライズプロット 4 M-R WSE-4040・M 型, ATTO 製) 装置を用い 120 mA, 2 時間で行った。

加水分解

タンパク質バンドを転写した PVDF 膜は 12 mm 試験管内で十分に乾燥させた後、0.05% β -ME を含む混合酸 (6 M HCl : トリフルオロ酢酸 (TFA) = 2 : 1) 300 μ l を加え、減圧下で封管後、110 $^{\circ}$ C で 24 時間加水分解を行った (18)。加水分解物は減圧乾固した後、200 μ l の 0.02 M HCl に溶解し、0.45 μ m のメンブレンフィルター (日本ミリポア製 Cosmonice Filter ϕ 2 mm) で濾過した。

アミノ酸組成分析

試料は高速アミノ分析計 L-8900 (日立製作所製) を用いてアミノ酸組成分析を行った。アミノ酸は波長 570 nm で検出を行い、Pro は波長 440 nm で検出を行った。

タンパク質同定システムを用いたタンパク質同定

アミノ酸組成情報を用いたタンパク質の同定はアミノ酸組成分析による同定システム (特開 2010-96501) を使用した。

同定用データベースは Uni-Prot 上に登録されている全タンパク質アミノ酸配列情報を利用し、シグナル配列が決定されているタンパク質はシグナル配列を除いた成熟タンパク質部分の配列を、不明なものはシグナル配列を含む全配列を使用し、20 種類のアミノ酸のアミノ酸組成比 (mol%) として作成した組成比データベースを用いた。Asn および Gln は加水分解により脱アミド化されそれぞれ Asp と Glu になることから、同定に用いるデータベースでは Asn と Gln の量をそれぞれ Asp と Glu に残基数を加算して設定した。スコアはデータベース上の理論アミノ酸組成比 (mol%) と実験によって得られたアミノ酸組成比 (mol%) とのマッチングにより算出した。理論値—実験値 = 0 (完全一致) である場合、スコア 100 とした。

アミノ酸濃度に対する補正

同定スコアの変動のシミュレーションは、アミノ酸組成分析で得られた各アミノ酸濃度に対して回収ができな

いアミノ酸及び回収率の低いアミノ酸に対してアミノ酸濃度を増加させ、スコアの変動をシミュレーションした。Cys と Trp はアミノ酸組成分析にて定量されないため Cys と Trp の回収された濃度を 0 として仮想濃度にて Cys 濃度と Trp 濃度を 0 から増加させた。補正はアミノ酸濃度を増加させた際の同定スコアの変動をシミュレーションし、同定スコアが最大となった点の Cys と Trp の濃度を用いる方法で行った。また、モデルタンパク質を用いた補正シミュレーションにおいて補正值に用いた Cys と Trp の濃度と実際にアミノ酸組成分析で得られたタンパク質量から計算される Cys と Trp の濃度の理論値を比較することで補正が適切に行えているか確認を行った。Ile はアミノ酸組成分析にて得られた Ile 濃度を増加させる補正を行い、スコアの変動をシミュレーションした。Ile に関しても、補正值として用いた Ile 濃度と実際に分析したタンパク質量における Ile の理論値と比較することで補正が適切に行えているか確認を行った。

結果及び考察

アミノ酸組成分析を用いたタンパク質同定

アミノ酸に対しての補正シミュレーションを行うために、ニワトリ (*Gallus gallus*) 由来 C-type lysozyme, ダイズ (*Glycine max*) 由来 Trypsin inhibitor, ダチョウ (*Anser anser*) 由来 G-type lysozyme, ウシ (*Bos Taurus*) 由来 Serum albumin の 4 種のタンパク質についてアミノ酸組成分析を行い、各アミノ酸実測値に対して補正をせずにタンパク質同定を行った (Table 1)。このアミノ酸組成分析の結果を使用して、Cys, Trp に対する補正とタンパク質配列中の特定の結合により回収率が低下する Ile に対する補正の検討を行った。

アミノ酸組成分析結果の Cys 実測値に対する補正

Cys は加水分解によって 2 分子結合を起こしてシスチンとして回収される (19)。しかし、シスチンの一部は破壊されることや、Cys に還元されることがあるため

Table 1 Identification score of model proteins (Without correction)

Protein	Species	Score (Without correction)	Score With Cys ^{a)}
C-type lysozyme	<i>Gallus gallus</i>	78.92	55.63
Trypsin inhibitor	<i>Glycine max</i>	82.37	75.91
G-type lysozyme	<i>Anser anser</i>	73.29	68.39
Serum albumin	<i>Bos taurus</i>	77.20	57.47

a) Cys concentration was set to 0

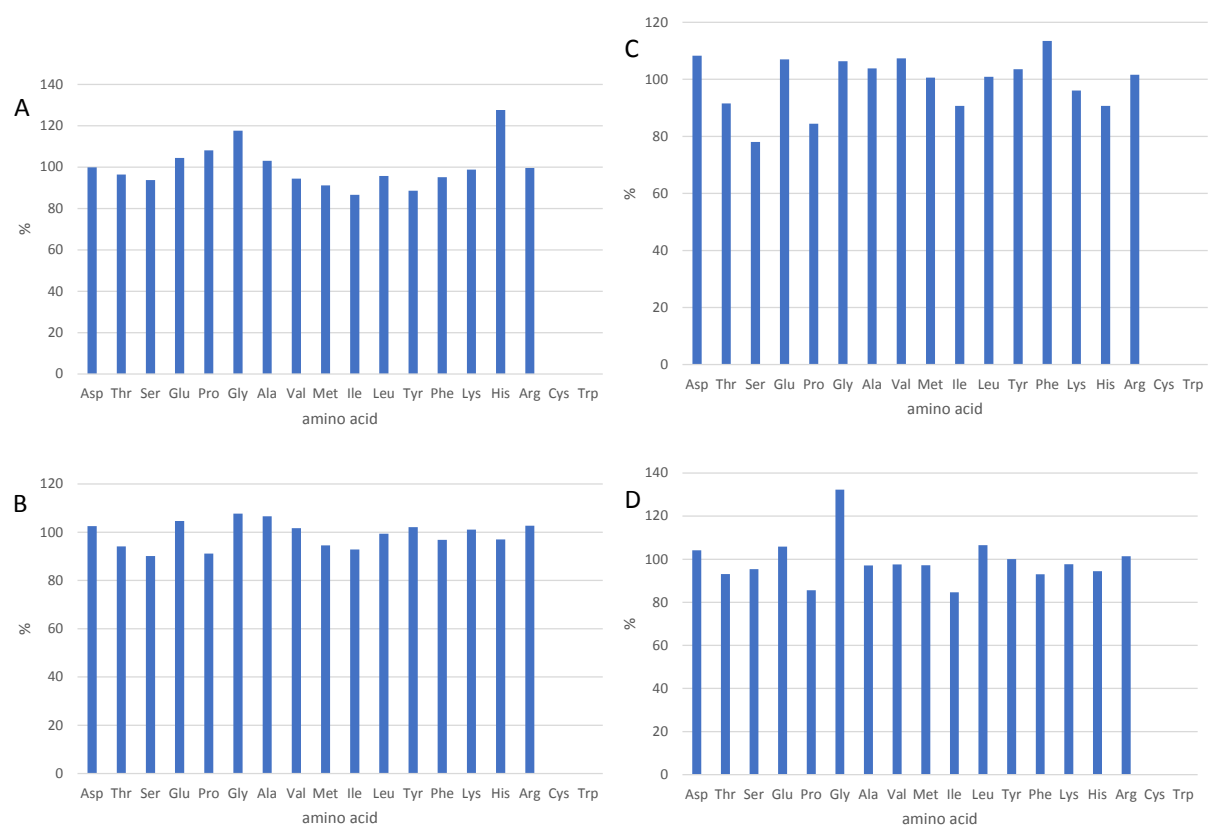


Fig.1 Amino acid recovery compared to the theoretical value of model protein.

A ; C-type lysozyme (*Gallus gallus*), B ; Trypsin inhibitor (*Glycine max*), C ; G-type lysozyme (*Anser anser*), D ; Serum albumin (*Bos taurus*). The percentage on the vertical axis shows the yield based on the theoretical value when the theoretical value was taken as 100%.

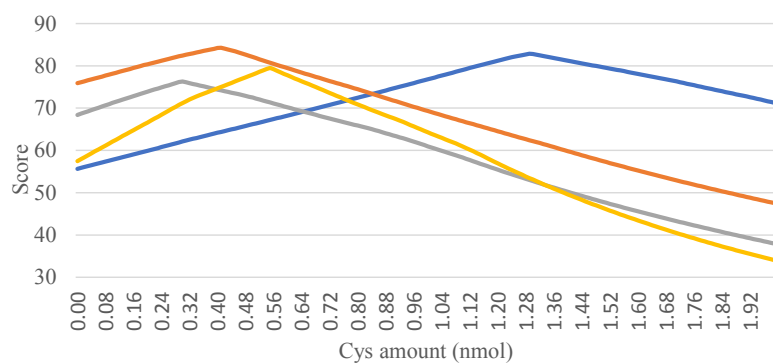


Fig.2 Simulation of relationship between corrected concentration of Cys and identification score.

Blue ; C-type lysozyme (*Gallus gallus*), Orange ; Trypsin inhibitor (*Glycine max*), Gray ; G-type lysozyme (*Anser anser*), Yellow ; Serum albumin (*Bos taurus*).

Table 2 Identification score and Cys concentration

	Score	Corrected score	Corrected Cys concentration ^{a)}	Theoretical Cys concentration ^{b)}
C-type lysozyme	55.63	82.31	1.33 nmol	1.37 nmol
Trypsin inhibitor	75.91	84.28	0.43 nmol	0.43 nmol
G-type lysozyme	68.39	76.30	0.32 nmol	0.32 nmol
Serum albumin	57.47	78.91	0.57 nmol	0.60 nmol

a) Cys concentration for highest identification score

b) The theoretical value of Cys concentration calculated from the amount of protein recovered as a result of amino acid composition analysis

Cys を安定して定量することは難しい。そのため Cys の定量は還元アルキル化を行い、カルボキシメチルシステインとして定量する方法が用いられている (20)。アルキル化を用いる方法は、従来のアミノ酸組成分析に加えて化学修飾反応操作を行う必要がある。このことから現在の同定手法は Cys の情報を用いていない。そこで補正により Cys の値をシミュレートすることでタンパク質の同定精度向上を検討した。各モデルタンパク質のアミノ酸組成の理論値を 100 % として実際のアミノ酸組成分析において回収されたアミノ酸ごとの回収量を示した (Fig.1)。Cys はアミノ酸組成分析で回収されていないため 0 % とした。Cys 補正が 0 時点の各タンパク質の同定結果は Cys の濃度を 0 として同定を行っているため、通常の Cys を含まずにを行った同定結果よりも低い値となっている (Table 1)。補正は Cys 濃度を 0 から増加させていき、スコアが最大となった Cys 濃度を用いた。スコアが最大となった Cys 濃度は濃度の増加量と同定スコアとの関係をシミュレーションして求めた (Fig.2)。また、補正後のスコアが最大となった Cys の濃度が理論値と合っているかの判断基準としてアミノ酸組成分析結果で回収されたタンパク質量から計算される Cys 濃度と比較した (Table 2) その結果、C-type lysozyme は 1.33 nmol

の Cys 濃度でスコアが 55.63 から 82.86 まで増加し、Cys 濃度の理論値は 1.37 nmol で補正に用いた Cys 濃度とほぼ同じとなった。Trypsin inhibitor は 0.43 nmol の Cys 濃度でスコアが 75.91 から 84.28 まで増加し、Cys 濃度の理論値は 0.43 nmol で補正に用いた Cys 濃度と同じとなった。G-type lysozyme は 0.32 nmol の Cys 濃度でスコアが 68.39 から 76.30 まで増加し、Cys 濃度の理論値は 0.32 nmol で補正に用いた Cys 濃度と同じとなった。同様に Serum albumin は 0.57 nmol の Cys 濃度として用いたところスコアが 57.47 から 78.91 まで増加し、Cys 濃度の理論値は 0.60 nmol で補正に用いた Cys 濃度と同じとなった。以上の結果より Cys の濃度を増加させていき、スコア最大となった濃度を用いる補正を行うことで、より精度の高いタンパク質同定が行えることが明らかになった。

アミノ酸組成分析結果の Trp 実測値に対する補正

Trp は側鎖にインドール環を持ち加水分解条件で酸化分解されるため現在の同定手法では Trp の情報を用いていない (19)。したがって補正により Trp の値をシミュレートすることでタンパク質の同定精度向上を検討した。各タンパク質は Trp 実測値の初期値を 0 として補正を行った (Fig.1)。補正は Trp 濃度を 0 から増加させていき、

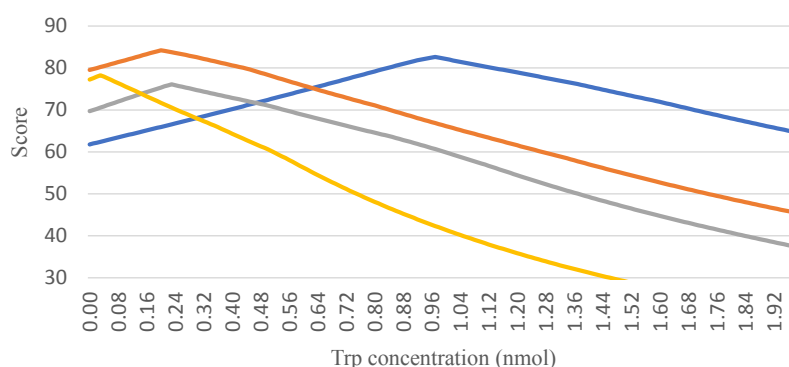


Fig.3 Simulation of relationship between corrected concentration of Trp and identification score.
Blue ; C-type lysozyme (*Gallus gallus*), Orange ; Trypsin inhibitor (*Glycine max*), Gray ; G-type lysozyme (*Anser anser*), Yellow ; Serum albumin (*Bos taurus*).

Table 3 Identification score and Trp concentration

Protein	score	corrected score	Trp corrected concentration ^{a)}	Trp theoretical concentration ^{b)}
C-type lysozyme	62.48	82.63	0.97 nmol	0.99 nmol
Trypsin inhibitor	79.52	84.22	0.20 nmol	0.20 nmol
G-type lysozyme	69.69	76.11	0.25 nmol	0.25 nmol
serum albumin	77.23	78.25	0.03 nmol	0.02 nmol

a) Trp concentration for highest identification score

b) Theoretical value of Trp concentration calculated from the amount of protein recovered as a result of amino acid composition analysis

スコアが最大となった Trp 濃度を補正值として用いた。スコアが最大となった Trp 濃度は濃度の増加量と同定スコアとの関係をシミュレーションして求めた (Fig.3)。また, Cys と同様に Trp 濃度の理論値と比較した (Table 3)。その結果, C-type lysozyme は0.97 nmol の Trp 濃度でスコアが62.48から82.63まで増加し, 理論値とはほぼ同じとなった。Trypsin inhibitor は0.20 nmol の Trp 濃度でスコアが79.52から84.22まで増加し, 理論値と同じとなった。G-type lysozyme は0.25 nmol の Trp 濃度でスコアが69.69から76.11まで増加し, 理論値と同じとなった。同様に Serum albumin は0.03 nmol の Trp 濃度として用いたところスコアが77.23から78.25まで増加し, 理論値とはほぼ同じとなった。以上の結果より Trp の濃度を増加させていき, スコア最大となった濃度を用いる補正を行うことで, より精度の高いタンパク質同定が可能となった。

アミノ酸組成分析結果の Ile 実測値に対する補正

Ile, Val 及び Leu はアミノ酸側鎖の α 鎖の近傍に枝分かれ構造を持つため, タンパク質配列中の Ile-Ile 結合, Ile-Val 結合及び Ile-Leu 結合は加水分解を受けにくく, アミノ酸組成分析において Ile 回収率が低下する (19)。

そこでアミノ酸組成分析を用いたタンパク質同定において回収率が低い Ile の実測値を補正することでタンパク質の同定精度向上の可能性が考えられる。そこで各タンパク質のアミノ酸回収量の理論値を100 %として実際に得られたアミノ酸回収量と比較した結果, すべてのタンパク質で Ile のアミノ酸回収量が100 %を下回っていた (Fig.1)。各タンパク質の実測値に補正せずに同定した結果, C-type lysozyme の同定スコアは78.92, Trypsin inhibitor は82.37, G-type lysozyme は73.29, Serum albumin は77.20となった (Table 1)。そこで, Ile の値を + 方向へ補正することによりどのように同定のスコアが変動するかについて, Ile の補正量と同定スコアの関係性をシミュレーションした (Fig.4)。また, 各タンパク質の Ile の実測値と補正後の Ile の補正值とその補正が理論値と合っているかの判断基準としてアミノ酸組成分析結果で回収されたタンパク質量から計算される Ile 濃度の理論値と比較した (Table 4)。その結果, Ile の補正により C-type lysozyme の同定スコアは78.92から80.44まで上昇した。Ile 実測値は0.89 nmol で補正によって1.05 nmol でスコア最大となった。Ile 濃度の理論値は1.12 nmol で補正に用いた Ile 濃度とはほぼ同じとなった。Trypsin inhibitor 同定スコアは82.37から85.04まで上昇し,

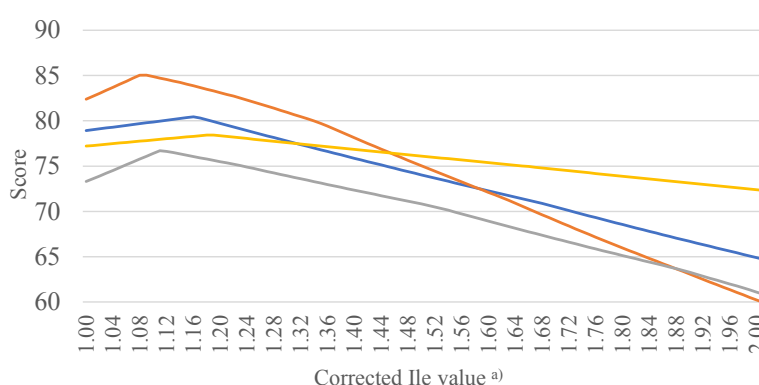


Fig.4 Simulation of relationship between corrected Ile value and identification score.

Blue ; C-type lysozyme (*Gallus gallus*), orange ; Trypsin inhibitor (*Glycine max*), gray ; G-type lysozyme (*Anser anser*), yellow ; Serum albumin (*Bos taurus*).

a) Corrected Ile value was indicated by simulated value of increased rate for measured value of Ile (1.00).

Table 4 Ile concentration and simulated values

Protein	Measured concentration	Corrected concentration ^{a)}	Theoretical concentration ^{b)}
C-type lysozyme	0.89 nmol	1.05 nmol	1.12 nmol
Trypsin inhibitor	1.29 nmol	1.38 nmol	1.44 nmol
G-type lysozyme	0.95 nmol	1.06 nmol	1.06 nmol
Serum albumin	0.18 nmol	0.22 nmol	0.22 nmol

a) Ile concentration when the score becomes maximum by Ile correction

b) Ile theoretical concentration calculated from the amount of protein recovered in the amino acid analysis

実測値の1.29 nmol から1.38 nmol でスコア最大となった。Ile 濃度の理論値は1.44 nmol で補正に用いた Ile 濃度とほぼ同じとなった。同様に G-type lysozyme の同定スコアは73.29から76.70まで上昇し、補正値は Ile 濃度の理論値と同じとなった。さらに Serum albumin のタンパク質の同定スコアは77.20から78.42まで上昇し、同様に補正値は理論値と同じとなった。これを各タンパク質配列中の Ile 数、Ile-Ile 結合数、Ile-Val 結合数、Ile-Leu 結合数と、同定スコアの関係性を調べた結果、Ile-Ile 結合を2つもつ G-type lysozyme の Ile の実測値を補正した際のスコア上昇が他のタンパク質を補正した際のスコア上昇と比較して一番大きかった。このことから Ile-Ile 結合が Ile-Val 結合、Ile-Leu 結合よりも同定スコアへの影響が大きいと考えられた。また、Ile-Val 結合と Ile-Leu 結合によりアミノ酸回収率に対して影響があると考えられる Val と Leu は、各タンパク質のアミノ酸回収率で理論値に対してほぼ100 %に近い回収率となっており、Val と Leu に対しては補正の必要性は小さいと考えられる。以上の結果より Ile の実測値を補正することで、より精度の高いタンパク質同定が行えることが明らかになった。ここまでの Cys, Trp 及び Ile への補正はいずれも同定スコアの上昇による同定精度向上が見込めるため、3種のアミノ酸全てを用いて補正を行うとより精度の高い同定

が可能であると考えられる。

Cys と Trp の補正濃度を使用し、Ile 実測値に対する補正を行った3重補正

モデルタンパク質を用いた通常のアミノ酸組成分析において、タンパク質同定にて得られていない Cys, Trp に対する補正とタンパク質配列中の特定の結合により回収率が低下する Ile に対する補正を検討した結果、3種のアミノ酸全ての補正で同定精度向上の効果が確認できた。そこで3種のアミノ酸全てを用いて3重補正を行うことにより、補正手法の確立を目指した。すなわち各モデルタンパク質同定は、Cys と Trp の濃度に対する補正によってスコア最大となった Cys 濃度と Trp 濃度を加えた18種のアミノ酸組成を使用し、Ile 実測値に対する補正を行うことで3重補正についてのスコア変動を調べた (Fig.5)。3重補正後の同定スコアを補正なしでの同定スコア、Cys, Trp 及び Ile に対して個々で補正を行った際のスコア、Ile 補正を行う前の Cys, Trp を用いた2重補正時のスコアと比較した (Table 5)。その結果、全てのモデルタンパク質で Cys, Trp 及び Ile に対する3重補正は同定スコアの上昇を示した。以上の結果より Cys, Trp 及び Ile に対する補正を個々で行うよりもスコア最大となった Cys, Trp 濃度を使用し、Ile 実測値に対する

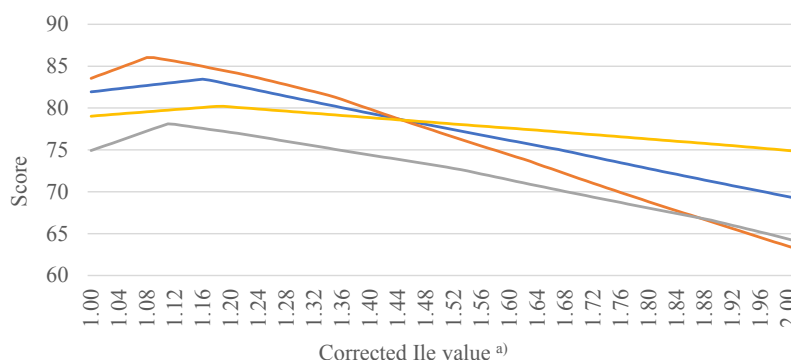


Fig.5 Simulation of Ile concentration for Cys and Trp corrected values.

Blue ; C-type lysozyme (*Gallus gallus*), Orange ; Trypsin inhibitor (*Glycine max*), Gray ; G-type lysozyme (*Anser anser*), Yellow ; Serum albumin (*Bos taurus*).

a) Corrected Ile value was indicated by simulated value of increased rate for measured value of Ile (1.00).

Table 5 Identification score of model protein (without correction), Cys corrected score, Trp corrected score, Ile corrected score, Cys,Trp Corrected Score, and triple correction score

Protein	No corrected Score	Cys Corrected score	Trp Corrected score	Ile Corrected score	Cys,Trp Corrected Score	Cys,Trp,Ile Corrected score
C-type lysozyme	78.92	82.63	80.91	80.44	81.94	83.45
Trypsin inhibitor	82.37	84.28	84.22	85.04	83.54	86.02
G-type lysozyme	73.29	76.30	76.11	73.70	74.93	78.10
Serum albumin	77.20	78.91	78.25	78.42	79.02	80.19

補正を行うことにより精度の高いタンパク質同定が可能であることが明らかになった。

Cys と Trp は通常加水分解では回収できないが、モデルタンパク質のアミノ酸組成分析結果に Cys と Trp を仮想濃度にて追加する補正方法は、各モデルタンパク質の同定スコアを上昇させた。また、Cys と Trp の仮想濃度は各モデルタンパク質のアミノ酸組成分析結果のタンパク質回収量から計算される Cys と Trp の理論値に近い濃度となった。このことから Cys と Trp の仮想濃度として用いた濃度は理論的にも正しい濃度であり、この補正方法は従来のアミノ酸組成分析を用いたタンパク質の同定よりも精度が高いと考えられる。また特定のアミノ酸配列により回収率が低下する Ile には、Ile 濃度を増加させる補正を行ったところ、各モデルタンパク質の同定スコアが上昇した。補正にてスコアが最大となった Ile 濃度は、各モデルタンパク質のアミノ酸組成分析結果のタンパク質回収量から計算される Ile の理論値に近い濃度となった。このことから回収率が低下した Ile 濃度を増加する補正は、従来のタンパク質同定と比較して精度が高くなったと考えられる。さらに、Cys と Trp の仮想濃度を使用し、Ile の濃度を増加させる 3 重補正は、Cys, Trp 及び Ile の補正を個々で行うよりもさらに精度が上昇し、本補正手法はアミノ酸組成を用いたタンパク質同定手法の補助手法として実用可能であると考えられる。

謝辞

タンパク質同定システムの開発にご協力いただいた、株式会社日立ハイテックフォールディングに感謝いたします。

要約

プロテオーム解析における同定手法として質量分析など部分アミノ酸配列情報を用いた解析が一般的である。しかし、質量分析はハイスループットであるが分析コストが高額であり、デメリットがある。そこで本研究では、低コストで操作も簡便なアミノ酸組成分析を用いたタンパク質同定手法の確立を目的とした。現在の手法では加水分解の際に影響を受けるアミノ酸が数種類存在し、そのアミノ酸の組成情報が欠落あるいは定量的な回収が困難となっている。これらのアミノ酸組成情報を活用できれば同定の精度を向上できると考えられる。そこで加水分解で回収が難しい Cys, Trp および Ile の収量に対して補正方法を見出すことで同定精度の向上を検討した。す

なわちそれぞれのアミノ酸の濃度を増加させ、スコアが最大値となるよう補正值の検討を行った。その結果 Cys, Trp および Ile はいずれも補正により同定の精度が向上し、これらのアミノ酸濃度の 3 重補正により高精度のタンパク質同定が可能であることを明らかにした。

参考文献

1. Shevchenko A., Jensen O.N., Podtelejnikov A.V., Sagliocco F., Wilm M., Vorm O., Mortensen P., Shevchenko A., Boucherie H., Mann M., 1996, Linking genome and proteome by mass spectrometry: Large-scale identification of yeast proteins from two dimensional gels, *Proc. Natl. Acad. Sci. USA*, 93, 14440-14445.
2. Malcolm G.P., Bogdanova A., Lopez M., Gutierrez S., Pitt A.M., 2002, Multiwell In-gel protein digestion and microscale sample preparation for protein identification by mass spectrometry, *Proteomics*, 2, 145-150.
3. Fenyö D., 2000, Identifying the proteome: software tools, *Current Opinion in Biotechnology*, 11, 391-395.
4. Shen Y.N., Kim S.H., Yoon D.H., Lee H.G., Kang H.S., Seo K.S., 2012, Proteome analysis of bovine *Longissimus dorsi* muscle associated with the marbling score, *Asian-Aust. J. Anim. Sci.*, 25, 1083-1088.
5. Ghabooli M., Khatabi B., Ahmadi F.S., Sepehri M., Mirzaei M., Amirkhani A., Jorrín-Novo J.V., Salekdeh G.H., 2013, Proteomics study reveals the molecular mechanisms underlying water stress tolerance induced by *Piriformospora indica* in barley, *J. Proteomics*, 94, 289-301.
6. Perkins D.N., Pappin D.J.C., Creasy D.M., Cottrell J.S., 1999, Probability-based protein identification by searching sequence databases using mass spectrometry data, *Electrophoresis*, 20, 3551-3567.
7. Steven G., Binz P.-A., Hochstrasser D.F., Appel R. D., 2002, Peptide mass fingerprinting peak intensity prediction: Extracting knowledge from spectra, *Proteomics*, 2, 1374-1391.
8. Yergey A. L., Coorsen J. R., Backlund P. S., Blank P. S., Jr., Humphrey G.A., Zimmerberg J., 2002, De novo sequencing of peptides using MALDI/TOF-TOF, *J. Am. Soc. Mass Spectrom.*, 13, 784-791.
9. Standing K.G., 2003, Peptide and protein de novo sequencing by mass spectrometry, *Curr. Opin. Struct. Biol.*, 13, 595-601.

10. Aebersold R., Mann M., 2003, Mass spectrometry-based proteomics, *Nature*, 422, 198-207.
11. Masuda A., Docmae N., 2013, Examination of an absolute quantity of less than a hundred nanograms of proteins by amino acid analysis, *Anal. Bioanal. Chem.*, 405, 8073-8081.
12. Masuda A., Dohmae N., 2011, Amino acid analysis of sub-picomolar amounts of proteins by precolumn fluorescence derivatization with 6-aminoquinolyl-N-hydroxysuccinimidyl carbamate, *BioSci. Trends.*, 5 (6) , 231-238.
13. SIB, AACompIdent (protein identification by aa Composition), < <http://web.expasy.org/aacomident/> >
14. Fountoulakis M., Lahm H.W., 1998, Hydrolysis and amino acid composition analysis of proteins, *J. Chromatogr. A*, 826, 109-113.
15. Weiss M., Manneberg M., Juranville J. F., Lahm H. W., Fountoulakis M., 1998, Effect of the hydrolysis method on the determination of the amino acid composition of proteins, *J. Chromatogr. A*, 795, 263-275.
16. 細澤知香, 2015, ゲノム未知生物種の特異的発現タンパク質の同定に関する研究, 東海大学大学院平成27年度修士論文
17. Laemmli U.K., 1970, Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature*, 227, 680-685.
18. 植木 厚, 1987, 続生化学実験講座2 タンパク質の化学上, 東京化学同人, 192 - 193.
19. 高橋礼子, 1976, 生化学実験講座1 タンパク質の化学Ⅱ－一次構造決定法－, 東京化学同人, 30-36.
20. Crestfield A.M., Moore S., Stein W.H., 1963, The preparation and enzymatic hydrolysis of reduced and S-carboxymethylated proteins, *J. Biol. Chem.*, 238, 622-627.