

PepNovo と PEAKS を用いた *de novo* ペプチド配列解析

磯野 直人^{1*}・柳原 和典¹・小林 裕子^{2,3}・小林 一成^{2,3}・久松 眞¹

¹ 三重大学大学院生物資源学研究科,

² 三重大学生命科学研究支援センター,

³ 三重大学大学院地域イノベーション学研究科

De novo peptide sequencing by PepNovo and PEAKS

Naoto ISONO^{1*}, Kazunori YANAGIHARA¹, Yuhko KOBAYASHI^{2,3},
Issei KOBAYASHI^{2,3} and Makoto HISAMATSU¹

¹ Graduate School of Bioresources, Mie University, Tsu, Mie 514-8507, Japan

² Life Research Center, Mie University, Tsu, Mie 514-8507, Japan

³ Graduate School of Regional Innovation Studies, Mie University, Tsu, Mie 514-8507, Japan

Abstract

We present a technique of *de novo* peptide sequencing from MALDI TOF/TOF tandem mass spectra by web-based programs, PepNovo and PEAKS. Furthermore, we show the result of *de novo* peptide sequencing of kidney bean granule-bound starch synthase I (GBSSI).

Key Words: *de novo* peptide sequencing, PepNovo, PEAKS, MALDI TOF/TOF tandem mass spectrometry

はじめに

近年、質量分析法の発展やゲノム解析の進歩により、ペプチドマスフィンガープリンティング (PMF) 法や MS/MS イオンサーチ法を用いて、タンパク質を簡単に同定できるようになった。しかし、これらはゲノムデータベースや EST データベースなどが整備された生物種に限り有効な実験手法である。データベースを利用せずにアミノ酸配列を決定する方法として、タンデム質量分析装置を用いた *de novo* ペプチド配列解析法 (*de novo* peptide sequencing) がある¹⁾。これはペプチドイオンの開裂によって生じた生成イオンのスペクトル (MS/MS スペクトル) から数学的演算によりアミノ酸配列を算出する技術である。通常、MS/MS スペクトル中には多くのピークが観察さ

れる。例えば、10 残基程度のペプチドの場合、明瞭なピークだけでも数十本あることが多く、さらに多数のノイズが見られる。そのため、マニュアル操作で構造未知のペプチドのアミノ酸配列を正確に決定するのはほとんど不可能であると考えられる。最近、AUDENS, Lutefisk, NovoHMM, PepNovo, PEAKS をはじめとする、多くの優れた *de novo* ペプチド配列解析用のアルゴリズム (プログラム) が考案された²⁻⁶⁾。現在では、これらのプログラムを利用したアミノ酸配列決定が、比較的容易に実行できるようになりつつある。

三重大学では 2009 年に 4800 Plus MALDI TOF/TOF Analyzer (エービー・サイエックス社製) が設置された。本機種の標準機能では ProteinPilot ソフトウェア (エービー・サイエックス社製) や Mascot (マトリックスサイエンス

社製)を用いたタンパク質の解析が可能である。これらはデータベースや既知のアミノ酸配列情報を利用し、既知タンパク質の同定や修飾解析を行うことを目的とするプログラムである。構造が未知でデータベース未登録のタンパク質のアミノ酸配列は決定できない。そこで、本解説では4800 Plus MALDI TOF/TOF Analyzerで得られたMS/MSスペクトルから、ウェブベースのソフトウェアであるPepNovoとPEAKS Onlineを用いて、*de novo* ペプチド配列解析する方法について紹介する。以下では、インゲンマメのGranule-bound starch synthase I (GBSSI, EC 2.4.1.242)を試料として用いた場合の解析方法と結果を示す。

方 法

1. SDS-ポリアクリルアミドゲル電気泳動 (SDS-PAGE) およびゲル内消化

既報⁷⁾にしたがって精製した組換え型GBSSI (5 μ g) をSDS-PAGEで分離した後、ゲルをBio-Safe CBB G-250 ステイン (バイオラッド社製) で染色し、目的のバンドを含むゲル片を切り出した。続いて、東京都老人総合研究所産学公連携プロテオーム共同研究センターの標準操作法にしたがって、還元アルキル化とトリプシンによるゲル内消化を行った⁸⁻⁹⁾。得られた消化液を遠心エバポレーターで2 μ L程度まで濃縮した後、40 μ Lの溶媒A (2% アセトニトリル, 0.1% トリフルオロ酢酸) と混合した。

2. ナノLCによるペプチドの分離

ゲル内消化した試料はトラップカラムHiQ sil C 18 W (内径500 μ m×長さ1 mm, 粒子径3 μ mのC 18 粒子を充填, ケーワイエーテクノロジー社製) を用いて脱塩した後、逆相カラムHiQ sil C 18 W (内径50 μ m×長さ50 mm, 粒子径3 μ mのC 18 粒子を充填, ケーワイエーテクノロジー社製) を接続したDiNAダイレクトナノHPLCシステム (ケーワイエーテクノロジー社製) に供した。溶媒Aで10分間カラムを洗浄した後、0-50%溶媒B (70% アセトニトリル, 0.1% トリフルオロ酢酸) の直線濃度勾配 (45分間), 50-100% 溶媒Bの直線濃度勾配 (10分間), 100-0% 溶媒Bの直線濃度勾配 (10分間)

でペプチドを分画した。操作はすべて流速300 nL/分で行った。カラムから溶出されたペプチド試料をマトリックス溶液 (0.4% α -シアノ-4-ヒドロキシケイ皮酸, 2% クエン酸二アンモニウム, 0.1% トリフルオロ酢酸, 70% アセトニトリル) と混合し、MALDIプレート (エービー・サイエックス社製) に30秒/スポットの速度で合計119スポットに分画した。

3. MALDI TOF/TOF タンデム質量分析

マトリックス支援レーザー脱離イオン化飛行時間型タンデム質量分析計 (MALDI TOF/TOF MS) である4800 Plus MALDI TOF/TOF Analyzer (エービー・サイエックス社製) を用いてペプチドの質量分析を行った。前駆イオンを検出するために、MS Reflector Positive モードで検出範囲を800~4000 Da, Focus Massを2400 Daに設定した。イオンカウントが1万カウント程度になるようにレーザー強度を設定し、1スポットあたり25箇所について50回レーザーをランダムに照射して全てのスポット (119スポット) に含まれる分子イオンのTOF質量スペクトルを測定した。さらに、各スポットから検出されるスペクトルのうち、イオン強度の高い上位10本のピークを前駆イオンとして、MS-MS 1 kV Positive モードでMS/MS自動測定を実施した。衝突誘起開裂 (CID) コントロールを“CID Off”, レーザー強度は前駆イオン検出時の1.2倍に設定し、1スポットあたり50箇所について50回レーザーをランダムに照射した。

4. データ変換

4800 Plus MALDI TOF/TOF Analyzerによって得られたSpot Setデータをソフトウェア4000 Series Explorerで開き、Peaks to Mascotの機能を用いて単一のMascot形式ファイルに変換した。Experiment Typeとして“Protein-Peptide by Spot Set”を選択した。MS/MS Peak Filterでは、Min S/Nを20, Min Areaを1000に設定した (MS/MSスペクトルのピークが小さい場合は、これらの設定値を小さくする)。Peaks to Mascotによって出力されるファイルの拡張子はtxtであるが、PepNovoで解析する時にエラーとなるため、拡張子をmgfに変更した。

5. PepNovo

PepNovo は University of California, San Diego の Center for Computational Mass Spectrometry の Web サイト (<http://proteomics.ucsd.edu/LiveSearch/>) から無料で利用できる。メインページの Tool selection では Tool を “PepNovo” とし、Spectrum file で Mascot 形式ファイル (拡張子 mgf) を選択した。また, “Use spectrum Charge” と “Use spectrum precursor m/z” の両方にチェックを入れた。More options では PepNovo search type を “De Novo” とし, Number of desired solutions を “1” とした (各ペプチドについて多数の予想配列を得たい場合はこの数値を増やす)。

なお, 翻訳後修飾が予想される場合はメインページの Allowed Post-Translational Modifications で, 考慮する修飾の種類を設定することができる。ただし, 翻訳後修飾を考慮すると解析に必要な時間が長くなる。また, 実際は未修飾のペプチドに対して, 修飾を考慮した解析を行うと, 正しいアミノ酸配列が得られないことが多い。

6. PEAKS Online

PEAKS Online は Bioinformatics Solutions Inc. の Web サイト

(<http://www.bioinfor.com:8080/peaksonline/>) から無料で利用できる。ただし, はじめて使用する前にアカウントを作成する必要があり, 一つのアカウントからの解析は 1 日 5 回に制限されている。多数の試料について, 短時間で解析する必要がある場合は Bioinformatics Solutions Inc. から購入できるスタンドアローン版の PEAKS Studio を利用しても良い。

Search ページの Enzyme は “Semi Trypsin”, Function は “Find de novo sequences” とした。Preprocess で “Do preprocess” を選択し, Instrument は “TOF-TOF” とした。Merge tol., Parent tol., Fragment tol. はすべて 0.5 Da とし, Precursor mass は “Monoisotopic” とした。Data file として Mascot 形式ファイル (拡張子 mgf) を選択した。データベースに関する項目は PEAKS Online では無効であるため, 初期設定のままとした。なお, PepNovo と同様, 翻訳後修飾について考慮した解析をするように設定することも可能である。

解析結果

インゲンマメ GBSSI はデンプンのアミロース合成に関与する酵素であり, すでに cDNA の全塩基配列が決定され, 国際塩基配列データベースに登録されている (アクセッション番号 AB029546)⁷⁾。トリプシン消化した GBSSI を MALDI TOF/TOF 質量分析計で分析したところ合計 86 個のペプチドイオンの MS/MS スペクトルが得られた。このデータを用いて *de novo* 配列解析を行った。ただし, *de novo* 配列解析には, アミノ酸配列や塩基配列のデータベース情報は使用しなかった。

PepNovo による解析では, 図 1 のような予想アミノ酸配列と PepNovo スコアの一覧表が得られる。PepNovo スコアが大きな値の場合, 予想配列が実際のアミノ酸配列と一致あるいは類似している可能性が高い。GBSSI の分析結果では 78 個のペプチドの予想配列が示された (PepNovo スコア: -135~112)。PepNovo スコアが 60 以上のペプチドは 22 個あり, このうち 11 個は予想配列が実際の配列と連続 7 残基以上一致していた (表 1)。つまり, PepNovo はこれらのペプチドの配列をかなり正しく決定できたことがわかった。一方, スコア 60 以下の場合, PepNovo 予想配列の信頼性は低く, 特にスコア 50 以下の予想配列は実際の配列とほとんど一致しなかった (データは示さない)。また, PepNovo はペプチドの N 末端側もしくは C 末端側の分析結果の信頼性が低い場合, この部分の予想配列を返さずにペプチド末端からの質量 (N-Gap および C-Gap) を用いた結果を示すことがある。一例を示すと, GBSSI のトリプシン分解物である $m/z = 1750.722$ のペプチドイオン (前駆イオン) の予想配列は LFVGA EVA であり, N-Gap と C-Gap がそれぞれ 431.341 と 533.328 であった。実際の配列は GMNLI FVGA EVAPWSK であり, 下線部の配列と予想配列は完全に一致していた (後述するように, I と L は質量が同一であり識別できない)。

PEAKS では, 図 2 のような予想アミノ酸配列候補とそれぞれのスコアの一覧表が得られる。一つのペプチドについて, 20 種類の予想配列候補が提示される。スコアが大きな値の場合, 予想配

表 1. PepNovo と PEAKS による GBSSI の予想アミノ酸配列¹⁾

No.	ペプチド イオン m/z	実際のアミノ酸配列	PepNovo 予想配列 ^{2,3)}	PepNovo スコア	PEAKS 予想配列 ^{2,4)}	PEAKS スコア	PepNovo-PEAKS 共通配列 ^{2,5)}
1	1786. 928	FIDQNVQIITLGTGK	LFDDVLLLVQV+490. 352	112	EMDNLNQLLLLDNNK	10	--
2	1734. 7841	GMNLFVGAEVAPWSK	TSNLLFVKEVA+533. 315	105	GMNLLFVQEVANDER	15	NLLFVXEVA ⁶⁾
3	1240. 5389	ALGTYGTVAMEK	ALGTYGTV+407. 325	95	ALGTYGTVAMEK	82	ALGTYGTV
4	904. 377	ELVSSGDR	ELVSSVDR	85	ELVSSVDR	56	ELVSSVDR ⁷⁾
5	1750. 722	GMNLFVGAEVAPWSK	431. 341+LFVGAEVA+533. 328	83	QFRFLVGAEVAVGSDR	10	LFVGAEVA
6	1896. 944	FLLKEALQAEVGLPVR	501. 438+EALKEV+655. 367	83	SSAKKEALQAEVGLFYR	10	EALKEV ⁶⁾
7	1766. 728	TDKFDLHFDTTSVK	491. 346+LDLHFDTTSVK	82	KLCFLDLHFDTTSVK	14	LDLHFDTTSVK
8	1745. 729	QIEQLEEIYPEK	763. 345+LQEKLEEL+1745. 763	81	RDVQLEELYLLQAR	10	XLEEL ⁶⁾
9	1796. 729	LYGPSAGVYEDNQLR	276. 425+GALAGVYED+530. 242	81	EFGPSAGVYEDNQLR	10	AGVYED
10	1589. 705	SAFDFTDGHLPVR	783. 362+GHLQVPR	80	SSMDFDGHLPVRSK	10	GHLQ ⁶⁾
11	1738. 764	GMNLFVGAEVAPWSK	431. 354+LFVGAEV+592. 357	79	SDNFLVGAEVADCQR	10	LFVGAEV
12	884. 3989	FDGPLAHK	FDGPLAHK	78	FDGPLAHK	49	FDGPLAHK
13	1256. 537	ALGTYGTVAMEK	ALGTYGTV+423. 237	75	ALGTYGTVAFEK	38	ALGTYGTV
14	1564. 675	VAFCHNISVQGR	727. 291+NLSVAGAR	74	SSGMDLHNSVGNAR	10	NLSV
15	2094. 917	?	217. 397+EKVSTAHFD+863. 508	68	FAQQVSTAHFDEEEAEK	10	VSTAHFD
16	2000. 833	HPPEPPLLNLNPEVR	510. 401+DFPLLNL+678. 323	68	HPPEPPLLNLNLPAGVDR	10	DFPLLNL
17	2013. 918	TGGLGDVGLGPSALAEHGHR	385. 412+DVLGGLPSAL+706. 268	67	--	--	--
18	1395. 678	EALQAEVGLPVR	271. 075000+AVAEVGL+485. 266	65	ELAQAEEVGLVPR	10	AEVGL
19	2105. 804	YDQYKDAWDTNVTVEVK	1101. 36+DTNVTVEVK	64	FVQVSHDACDNTVPCPMK	13	DTNV
20	1952. 924	AALESDLVLTSPYYAK	497. 358+SDVLTW+641. 431	63	QCHESDLVLTVSACCCNK	23	SDVLTW ⁷⁾
21	2073. 834	YDQYKDAWDTNVTVEVK	1069. 449+DTNVTVEVK	63	FMGMCQDAWDTNVPCEAR	10	DTNV
22	1518. 639	QIEQLEEIYPEK	611. 278+EEYLPTR	61	KLMYLEELQHSK	10	EE
23	1672. 646	IIQNCMAQDFSWK	LLQNCFAQD+583. 374	58	EPQNCFAQDQDSMCK	10	CFAQD
24	2089. 8149	YDQYKDAWDTNVTVEVK	ENYKDAWDT+788. 368	55	FMQYKDAWDTNVTVPCR	10	YKDA ⁶⁾
25	1952. 9449	AALESDLVLTSPYYAK	497. 501+MALVLTW+641. 304	54	GVVLESDLVLTVSAHTNK	14	LVLTX ⁷⁾
26	1456. 576	FFHCYKQVDR	754. 236+QAGVDR	52	RDGCGCYKQDR	10	XXXDR ^{6,8)}
27	1404. 6121	VFVDHPCFLEK	597. 364+PSSFLEK	51	AGFLFACPHDAAGK	10	--
28	1419. 676	ILAGSDFIMIPSR	174. 08+PSLMLFD+442. 413	51	LLAGSDFLMLPSR	14	LML

1) PepNovo スコアが 50 以上のペプチド (28 個) の解析結果を示した。

2) 実際のアミノ酸配列と一致する連続した 2 残基以上の配列を赤字で示した。また記号 L は「(イソロイシン) あるいは L (ロイシン)」を意味している。

3) N-Gap と C-Gap (図 1) がある場合は予想配列の前後に数値を記載した。2 残基以上の配列を下線で示した。

4) 20 種類の予想配列のうち、最も高いスコアを示す配列を記載した。最高スコアが 10 のペプチドについては seq 1 (図 2) の配列を示した。また、PepNovo 予想配列と一致する連続した 2 残基以上の配列を下線で示した。

5) PepNovo と PEAKS の両方のプログラムで共通して予想された配列を示した。

6) K≡Q=AG=GA (質量)

7) W≡AD=DA=EG=GE=SV=VS (質量)

8) R≡GV=VG (質量)

Filter	Title	Scan	Rank Score	PepNovo Score	N-Gap	C-Gap	[M+H]	Charge	Sequence
<input type="checkbox"/>	Label: B13, Spot_Id: 126779, Peak_List_Id: 175504, MSMS Job_Run_Id: 13735, Comment: (SQS 0)	1	0.14800	34.16200	0.000	0.000	805.376	1	THHASPR
<input type="checkbox"/>	Label: C4, Spot_Id: 126795, Peak_List_Id: 175520, MSMS Job_Run_Id: 13735, Comment: (SQS 0.000)	2	-0.57700	-40.54900	0.000	0.000	810.340	1	FFHFGR
<input type="checkbox"/>	Label: B12, Spot_Id: 126780, Peak_List_Id: 175505, MSMS Job_Run_Id: 13735, Comment: (SQS 0.000)	3	0.48500	-17.60900	0.000	0.000	816.465	1	LSWVGVR
<input type="checkbox"/>	Label: B6, Spot_Id: 126786, Peak_List_Id: 175508, MSMS Job_Run_Id: 13735, Comment: (SQS 0.000)	4	1.49700	-13.79200	0.000	0.000	839.401	1	ELQLYSS
<input type="checkbox"/>	Label: C4, Spot_Id: 126795, Peak_List_Id: 175524, MSMS Job_Run_Id: 13735, Comment: (SQS 0.000)	5	0.19200	-37.88000	0.000	0.000	844.321	1	EHFQRK
<input type="checkbox"/>	Label: B3, Spot_Id: 126789, Peak_List_Id: 175512, MSMS Job_Run_Id: 13735, Comment: (SQS 0.000)	6	-1.10600	-57.07000	0.000	0.000	847.833	1	QEHCEN
<input type="checkbox"/>	Label: B7, Spot_Id: 126785, Peak_List_Id: 175507, MSMS Job_Run_Id: 13735, Comment: (SQS 0.000)	7	1.71500	46.46000	0.000	0.000	851.401	1	AHFDFK
<input type="checkbox"/>	Label: E5, Spot_Id: 126844, Peak_List_Id: 175584, MSMS Job_Run_Id: 13735, Comment: (SQS 0.000)	8	-0.59500	-30.43100	0.000	0.000	864.428	1	LPSQYEK
<input type="checkbox"/>	Label: B3, Spot_Id: 126789, Peak_List_Id: 175514, MSMS Job_Run_Id: 13735, Comment: (SQS 0.000)	9	0.78900	77.65400	0.000	0.000	884.399	1	FDGPLAHK
<input type="checkbox"/>	Label: C1, Spot_Id: 126792, Peak_List_Id: 175517, MSMS Job_Run_Id: 13735, Comment: (SQS 0.000)	11	-1.11400	-27.38700	0.000	0.000	901.325	1	FFASSWR

図1 PepNovoの解析結果画面

列が実際のアミノ酸配列と一致あるいは類似している可能性が高い。GBSSIの分析結果では80個のペプチドの予想配列が示された（PEAKSスコア：10～82）。PEAKSスコアが25以上のペプチドは合計12個あり、このうち5個は予想配列が実際の配列と連続7残基以上一致していた（データは示さない）。ただし、PEAKSスコアが最低値の10であるにもかかわらず、予想配列の一部が実際の配列と一致するケースも多かった。例えば、表1ではPepNovoスコアが50以上のペプチド（28個）が示されているが、PEAKSスコア10のペプチドが17個あり、このうち7個は予想配列が実際の配列と連続7残基以上一致していた。また、PEAKSはPepNovoと異なり、ペプチドの末端についても必ず予想配列を提示する。しかし、PEAKSスコアが極めて高いケースを除き、ペプチドの末端側の配列の信頼性は低かった（表1）。

このように、PepNovoとPEAKSは優れた*de novo* ペプチド配列解析プログラムであり、部分的ではあるが、正確にアミノ酸配列を決定するこ

とができた。ただし、*de novo* ペプチド配列解析では気をつけなければならない点がいくつかある。(1) アミノ酸残基LとIは同一の質量であるため識別できない。そのため、PepNovoとPEAKSの予想配列中の記号Lは「IまたはL」を意味している。(2) アミノ酸残基KとQの質量差は0.036しかない。また、QはAGあるいはGAと同一の質量である。スペクトルによってはこれらのアミノ酸残基（配列）の判別ができないことがある。同様に、(3) NとGGの識別、(4) W, AD, DA, EG, GE, SV, VSの識別、(5) R, GV, VGの識別も難しいので、これらのアミノ酸残基（配列）が予想配列中にある場合は、注意する必要がある。

タンパク質をコードする遺伝子やcDNAをクローニングするために、アミノ酸配列情報を基にして、(RT-)PCRに用いるオリゴヌクレオチドプライマー（縮重プライマー）を設計することがある。しかし、PepNovoあるいはPEAKSのいずれか一方のプログラムの結果のみを用いて効果的なプライマーを設計するのは難しい。例えば、

Search Parameter Summary

User name : Naoto Isono
 Email : isono@bio.mie-u.ac.jp
 Submit time : 2010-09-22 00:00:18
 Search title : PvGBSSI
 MS data file : pps_20100628isono_1278740092.mgf
 Search function : Find de novo sequences
 Do preprocess : Yes
 Use filter : No
 Database : N/A
 Taxonomy :
 Enzyme : Semi Trypsin
 Fixed modifications :
 Variable modifications :
 Instrument : TOF-TOF
 Precursor mass : Monoisotopic
 Merge tolerance : 0.5 Da
 Parent tolerance : 0.5 Da
 Fragment tolerance : 0.5 Da
 Max missed cleavages : 3
 Max variable PTM : 3

Peptide Candidates for each Spectrum

index	m/z	z	seq1	score1	seq2	score2
1	801.2928	1	YWNYR	10	VYKVHR	10
2	805.376161	1	DECVSPR	34	DCEVSPR	34
3	810.3398	1	YMHECK	49	YMHECK	49
4	816.4654	1	RETVRR	10	GVETVRR	10
5	839.4005	1	QLNCCMK	10	QLNCCMK	10
6	844.320561	1	CCCCGPR	10	CCCCGPR	10
7	847.8333	1	QETRVSK	10	EQRTSVK	10
8	851.4006	1	KLNTTFK	10	KLNAMFK	10
9	864.4277	1	GGDPVEYK	10	GGSSH TYR	10
10	884.3989	1	FDGLPAHK	49	FDGLPAHK	49

図2 PEAKS の解析結果画面

GBSSI の場合、PepNovo ではスコア 60 以上のペプチド (22 個) のうち、提示された予想配列が完全に正しいペプチドは 10 個であった (表 1)。また、PEAKS ではスコア 25 以上のペプチド (12 個) のうち、予想配列が完全に正しいペプチドは 3 個だけであった。つまり、これらのペプチド予想配列を用いて作成したプライマーの多くは、(RT-) PCR で機能しないことが予想された。そこで、私たちは PepNovo と PEAKS の両方のプログラムで、共通して予想された配列を調べた。その結果、連続した 7 残基以上の配列が共通しているペプチドは 10 個あることが明らかとなった (表 1)。この共通予想配列のうち 8 個については実際のアミノ酸配列と完全に一致していた。このように、PepNovo と PEAKS を併用し、情報を絞り込めば、より確実に正しいアミノ酸配列を取得できることがわかった。これらの共通予想配列を用いれば、(RT-) PCR に有効なプライマーを、効率的に作成することができると推定した。

最近、私たちは本解説で紹介した手法を用いて、微細藻類 *Ochromonas danica* 由来 β -1,3-glucan phosphorylase (EC 2.4.1.97) のアミノ酸配列を決定した (データ未発表)。本酵素は *Ochromonas* 属以外では発見されておらず、これまで構造が不明であり、*Ochromonas danica* や類縁生物のゲノム配列も決定されていなかった。また、アミノ酸配列情報から作成した縮重プライマーを用いて RT-PCR を行い、本酵素をコードする cDNA のクローニングにも成功した。このように *de novo* ペプチド配列解析は、構造未知のタンパク質の研究を進める際に非常に役立つ技術であり、今後さらに利用が進むであろう。

要 約

マトリックス支援レーザー脱離イオン化飛行時間型タンデム質量分析計 (MALDI TOF/TOF MS) によって得られた MS/MS スペクトルを用

いて、*de novo* ペプチド配列決定する手法を解説する。ウェブベースのソフトウェアである PepNovo と PEAKS online を使用した解析方法を紹介する。また、インゲンマメ Granule-bound starch synthase I (GBSSI) を試料とした解析例を示す。

引用文献

- 1) MEDZIHRADSKY, K. F. (2005) Peptide sequence analysis. *Methods Enzymol.* **402**: 209-244.
- 2) GROSSMANN, J., ROOS, F. F., CIELIEBAK, M., LIPTÁK, Z., MATHIS, L. K., MÜLLER, M., GRUISSEM, W. and BAGINSKY, S. (2005) AUDENS: a tool for automated peptide de novo sequencing. *J. Proteome Res.* **4**: 1768-1774.
- 3) TAYLOR, J. A. and JOHNSON, R. S. (2001) Implementation and uses of automated de novo peptide sequencing by tandem mass spectrometry. *Anal. Chem.* **73**: 2594-2604.
- 4) FISCHER, B., ROTH, V., ROOS, F., GROSSMANN, J., BAGINSKY, S., WIDMAYER, P., GRUISSEM, W. and BUHMANN, J. M. (2005) NovoHMM: a hidden Markov model for de novo peptide sequencing. *Anal. Chem.* **77**: 7265-7273.
- 5) FRANK, A. and PEVZNER, P. (2005) PepNovo: de novo peptide sequencing via probabilistic network modeling. *Anal. Chem.* **77**: 964-973.
- 6) MA, B., ZHANG, K., HENDRIE, C., LIANG, C., LI, M., DOHERTY-KIRBY, A. and LAJOIE, G. (2003) PEAKS: powerful software for peptide de novo sequencing by tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* **17**: 2337-2342.
- 7) ISONO, N., ITO, H., SENOURA, T., YOSHIKAWA, M., NOZAKI, K. and MATSUI, H. (2003) Molecular cloning of a cDNA encoding granule-bound starch synthase I from kidney bean seeds and expression in *Escherichia coli*. *J. Appl. Glycosci.* **50**: 355-360.
- 8) 戸田年総 (1997) 固定化 pH 勾配ゲルストリップを用いた二次元電気泳動. *生物物理化学* **41**: 13-20.
- 9) http://www.proteome.jp/2D/J_2DEmethod.html