

平成 21 年度戦略的基盤技術高度化支援事業
「血液診断バイオマーカーのための定量比較 LC-MS ロボットにおける
組込みソフトウェアの開発」

研究開発成果等報告書

平成 22 年 3 月

委託者 関東経済産業局

委託先 株式会社 MCB I

第1章 研究開発の概要	
1-1 研究開発の背景・研究目的及び目標	2
1-1-1 研究開発の背景	2
1-1-2 研究目的及び目標	3
1-1-3 実施結果等	3
用語および略号表	3
1-2 研究体制	4
1-2-1 研究組織及び管理体制	4
1-2-2 管理員及び研究員	4
1-2-3 経理担当者及び業務管理者の所属、氏名	5
1-2-4 他からの指導・協力者	5
1-3 成果概要	6
1-4 当該研究開発の連絡窓口	7
第2章 本論	
2-1 LCにおけるデータの再現性の確保	8
2-1-1 対象としたLCデータ	8
2-1-2 組み込みソフトウェアによるデータ処理	9
2-2 MSにおける質量分析データの正確な取扱いの保証	12
2-2-1 対象としたLC-MSデータ	12
2-2-2 分析結果	13
2-2-3 UFLC-MSデータの保証	15
2-3 標準サンプルを用いたLC-MSロボットのPERFORMANCEの検証	15
2-3-1 基本的な考え方	15
2-3-2 クラスタリングの方法	16
2-3-3 対象としたLC-MSデータ	16
2-3-4 LS-MSデータとその処理	16
2-3-5 ランドマークの応用	18
最終章 全体総括	
1. 研究開発成果	20
2. 研究開発後の課題・事業展開	20
2-1 研究開発後の課題	20
2-2 事業展開	20

第1章 研究開発の概要

1-1 研究開発の背景・研究目的及び目標

1-1-1 研究開発の背景

本研究開発は、特定ものづくり基盤技術において達成しようとする高度化の目標、特定ものづくり基盤技術高度化指針のうち、以下の項目に対応している。

(一) 組込みソフトウェアに係る技術に関する事項

1 組込みソフトウェアに係る技術において達成すべき高度化目標

1) 川下製造業者の抱える課題及び要請（ニーズ）

これについて、対応する事項ごとに研究開発の背景を記述する。

(1) 産業機械及び産業用ロボットに関する事項

エ. 新たな適合分野への対応

産業用ロボットは、製造業において主に溶接・塗装・ワーク脱着等の作業工程で使用されていたが、近年医療分野においてもいわゆる「医療用ロボット」が実用化されつつある。大手の病院や検査センターでは、血液から血清・血漿を分離し、通常の生化学検査装置でほぼオートメーションで処理と計測が行うことができるようになってきているが、これは医療用ロボットの一つである。さらに最近では、外科手術をサポートする手術支援ロボットが実用化され、これも産業用ロボットの新たな適合領域となっている。一方、これらの装置を製造する川下製造業者（OEM 供給業者を含む）では、医学研究の高度化・高速化に伴って次々に生み出される新技術への対応も必要とされるようになってきている。その中でも、生体試料の成分を分離する液体クロマトグラフィー（LC）と田中耕一氏がノーベル賞を受賞した質量分析（MS）技術の組合せ（LC-MS）は生体分子の検出、とくに血液中のタンパク質・ペプチドの検出に多大の成果を挙げた。最近ではこの計測技術が診断分野に応用され、従来にない高精度の診断を可能にした。しかしながら、現状ではマニュアル操作に頼る研究室レベルの応用に留まり、臨床診断の現場で広く活用されるに至っていない。その大きな理由の一つがこの技術の広範囲な活用を可能にする組込みソフトウェアの不足である。すなわち、LC-MS 装置を製造する川下業者は、組込みソフトウェアの高度化による当該分野への適応が求められている。具体的には、本研究開発で取り組む課題として、計測装置である高速液体クロマトグラフィー装置と質量分析装置が正しく稼働しているかをテスト・検証する技術、質量分析データから計測対象物を定量する技術、ヒトの健常と病気の状態における検査データを比較・評価する技術があげられる。これら技術を充足するソフトが組み込まれた LC-MS 装置が求められる血液検査ロボットである。

(5) 川下分野横断的な共通の課題・ニーズに関する事項

ウ. 品質の向上

当該新規適応分野である医療用ロボットでは特に高い品質・信頼性が求められる。そのため、前述のように血液検査ロボットが正しく稼働しているかを随時テスト・検証することが可能でなければならない。プロセスを可視化し、管理者がそれを利用して直ちに対応できるような組込みソフトウェアが求められている。

エ. 国際規格への対応

国際規格に基づく品質保証が必要である。研究分野ではタンパク質・ペプチドに関する質量分析データについての国際規格を国際学会（Human Proteome Organisation, HUPO）が定めている。産業レベルではデータ処理についての国際規格がないため、研究分野に準じた国際規格とこれに基づいたデータのみを扱うようなソフトウェアが求められている。

1-1-2 研究目的及び目標

本研究開発の目的は、現状ではマニュアル操作に頼っている液体クロマトグラフィー（LC）と質量分析（MS）技術の組合せ（LC-MS）からなる生体試料中のタンパク質・ペプチドの定量と結果の客観的比較を、組み込みソフトウェアの開発により広範囲の疾病の診断に活用できるようにすることである。本研究開発で取り組む課題として、計測装置である高速液体クロマトグラフィー装置と質量分析装置が正しく稼働しているかをテスト・検証する技術、質量分析データから計測対象物を定量する技術、ヒトの健常と病気の状態における検査データを比較・評価する技術があげられる。これらの技術を充足するソフトが組み込まれた LC-MS 装置が求められる血液検査ロボットであり、これを「定量比較 LC-MS ロボット」と呼ぶことにする。

研究の目標は、本ロボットから得られる結果についての「信頼性の向上」である。そのためには、LC で同じ試料を計測する場合の再現性が担保されていることが必要であり、MS で質量分析データのノイズピークを取り除き、正しくピークを認識して、データ処理に供することが必要である。本研究開発の組み込みソフトウェアにおいて、これら 2 つの装置が正しく稼働しているかどうかをテスト・検証する技術および質量分析データから計測対象物を正確に定量する技術の開発を目標とする。

1-1-3 実施結果等

LC-MS による生体試料中のタンパク質・ペプチドの定量と結果の客観的比較を可能にする定量比較 LC-MS ロボットのための組み込みソフトウェアの開発を目的とし、本ロボットから得られる結果についての「信頼性の向上」を目標としたが、これらの目的・目標は本研究開発において達成された。

用語および略号表

LC : liquid chromatography、液体クロマトグラフィー

MS : mass spectroscopy、質量分析

TIC : total ion chromatogram。吸光度のモニターの代わりに、LC における溶出分画を横軸とし、MS から得られる全てのイオンの強度の和を縦軸としたクロマトグラムを total ion chromatogram (TIC) と呼ぶ。

BPC : TIC からバックグラウンド信号を除き、各溶出分画の最大信号を溶出分画に対してプロットしたものを base peak chromatogram (BPC) と呼ぶ。

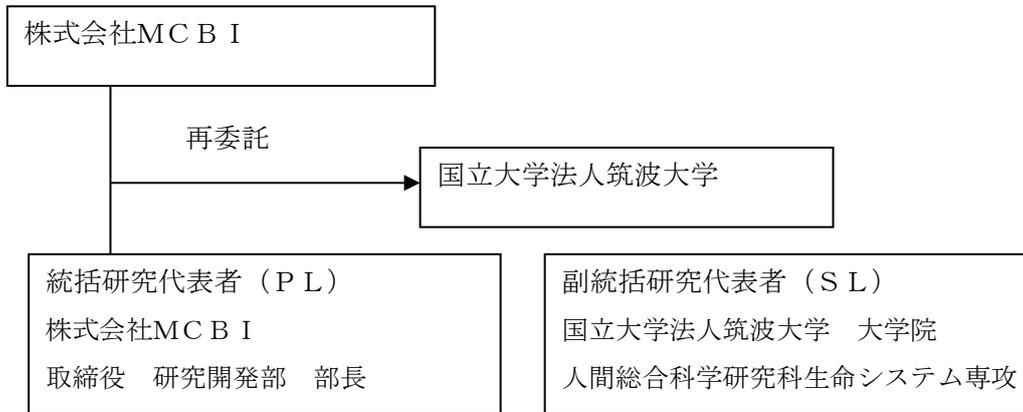
RT : LC における溶出時間 (retention time の略)。LC-MS では LC の溶出分画と 1:1 に対応する。

UFLC : 超高速 LC、UF は ultrafast の略。

1-2 研究体制

1-2-1 研究組織及び管理体制

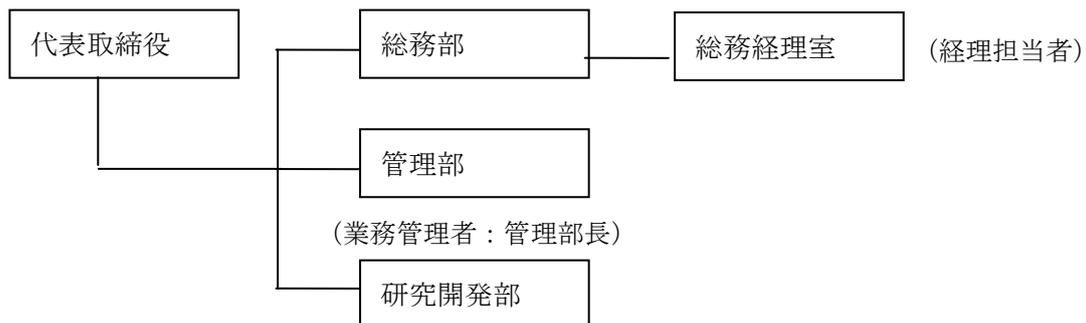
1) 研究組織（全体）



2) 管理体制

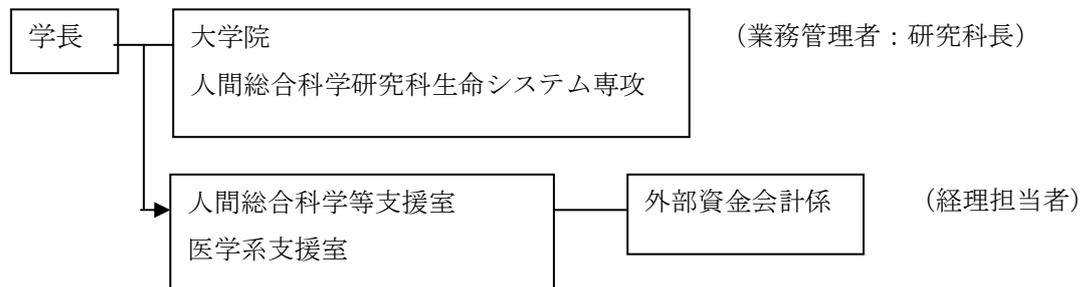
①事業管理者

【株式会社MCBI】



② 再委託先

【国立大学法人筑波大学】



1-2-2 管理員及び研究員（以下の表の実施内容（番号）は1-3 成果概要を参照）

【事業管理者】株式会社MCBI

①管理員

氏名	所属・役職	実施内容（番号）
永島 廉平	管理部長兼研究開発部長	④（=プロジェクト管理運営）
酒井 弘枝	総務部総務経理室長	④（=プロジェクト管理運営）
内山 典子	管理補助員	④（=プロジェクト管理運営）

② 研究員

氏名	所属・役職	実施内容 (番号)
目野 浩二	研究開発部・室長	②③
石井 俊	研究開発部・室長	①②③
鈴木 秀昭	研究開発部・室長	①②③
山口 英司	研究開発部・室長	②③
是永 龍巳	研究開発部・研究員	①②
S.K.Dissanayake	研究開発部・研究員	①②
皆川 昇	研究開発部・研究員	①②③

【再委託先】

(研究員)

氏名	所属・役職	実施内容 (番号)
内田 和彦	国立大学法人筑波大学 大学院 人間総合科学研究科生命システム専攻 准教授	①②③

1-2-3 経理担当者及び業務管理者の所属、氏名

(事業管理者)

株式会社MCBI

(経理担当者) 総務部総務経理室長 酒井 弘枝

(業務管理者) 管理部長 永島 廉平

(再委託先)

国立大学法人筑波大学

(経理担当者) 人間総合科学等支援室 医学系支援室 外部資金会計係長 菅谷 哲宏

(業務管理者) 大学院 人間総合科学研究科 研究科長 五十殿 利治

1-2-4 他からの指導・協力者

氏名	所属・役職	備考
永島 廉平 内田 和彦	株式会社MCBI 管理部長兼研究開発部長 国立大学法人筑波大学 大学院 人間総合科学研究科生命システム専攻 准教授	委 PL SL
朝田 隆	国立大学法人筑波大学 大学院 人間総合科学研究科疾患制御医学専攻 教授	アドバイザー
高野 純	株式会社島津製作所 経営戦略室兼基盤技術研究所	アドバイザー

1-3 成果概要

① LCにおけるデータの再現性の確保（実施：株式会社MCBI、国立大学法人筑波大学）

実施内容

LCでは、使用されるカラムの劣化や試薬の品質の違い、サンプルの特質などにより、毎回の試行（ランと称する）において溶出データにバラツキが生ずる。現状では特定波長の吸光度を連続してモニターし、「通常得られる」クロマトグラム（画像）かどうかを目視によって判断している。そこで、生体試料由来の適当な標準サンプル（例えば、よく保存された健常人の血清サンプル）を用いてデータを蓄積しておき、一定の基準を設けることにより「通常得られる」クロマトグラムかどうかを判断できるとよい。本研究開発の組込みソフトウェアはこの作業を実験者に代わって行う。すなわち、実サンプルについての検査の前または後に標準サンプルのテストランを行ったとき、組込みソフトウェアはLCのperformanceについての合否を判断し、LCデータの再現性についての客観的な情報を出力する。通常、生体試料由来サンプルをLCに負荷する前に純水をサンプルとして用い（これをblank runと称する）、カラムの劣化や試薬の品質の良・不良を調べる。ついで、生体試料由来の標準および被検サンプルをLCに負荷し、被検サンプルのデータを標準サンプルの場合と比較することになる。

成果概要

装置としてのLCを初めて使用するとき、何らかの修理を施したとき、カラム、チュービング、ポンプなどの部品の交換を行ったとき、新しいロットの溶出溶媒を使用するときなどには、LCが所定のperformanceを示すかどうかを確認する必要がある。そのために、①純水および②測定対象とする生体試料（健常人血清および被検血清などに所定の前処理を加えたもの）をサンプルとしてLC装置に負荷し、期待外の異常が何ら生じないこと、および繰り返し試行でのバラツキが少なく一定範囲内であることを確認する。

「定量比較LC-MSロボット」では、これを数値的データに基づいて組込ソフトウェアが行う。そのプログラムを構築し、実データの処理を行って期待通りの確認を行うことができた。

② MSにおける質量分析データの正確な取扱いの保証（実施：株式会社MCBI、国立大学法人筑波大学）

実施内容

MSによって得られる質量分析データも横軸に m/z を目盛り、縦軸にシグナル強度を目盛った一種の画像データである。実際のデータでは m/z を異にする多数のイオンのピークが見られる。

LCの溶出分画ごとに得られるMSにおける質量分析データの正確な取扱いを保証する組込みソフトウェアは、研究分野でタンパク質・ペプチドに関して国際学会（HUPO）が定めている国際規格を組込みソフトウェアに導入し、質量分析データの正確な取扱いを保証できるようにするものである。

成果概要

本研究開発の組込みソフトウェアに関係する国際規格は、①タンパク質・ペプチド特定のためのretention time (RT)を限定できること、②質量許容誤差範囲内にMSデータが収まることである。

LC-MS装置から得られる全てのイオンの強度の和からなるTICからバックグラウンド信号を除き、各溶出分画（RTと1:1に対応）の最大信号を溶出分画に対してプロットしたものをbase peak chromatogram (BPC)と呼ぶが、BPCの分析によって、MSデータの正確な取扱いの保証が可能で

ある。標準合成ペプチドを用いて LC-MS 装置の RT および MS に関する performance を検証するプログラムを構築し、実データについて質量分析データの正確な取扱いの保証を行うことができた。

③ 標準サンプルを用いた LC-MS ロボットの performance の検証（実施：株式会社MCBI、国立大学法人筑波大学）

実施内容

LC-MS ロボットでは、LC 単独あるいは MS 単独での出力に関する信頼性の保証のみでは不十分で、統合的かつシームレスに LC-MS としての performance を検証できるようにすることが求められる。本研究項目ではこのような機能をもつ統合的ソフトウェアを開発する。

成果概要

2D-HPLC および Ultraflex TOF/TOF を使用した場合の LC-MS ロボットの performance の検証は以下のステップによって行われる。

- 1) 複数の健康人血清についての LC-MS 測定データ (mzXML 形式ファイル) を入力する。
- 2) それぞれのデータを内部標準によって標準化し、ピーク強度が一定の内部標準比以上かつ連続する一定数分画以上で出現するピークからなる「標準化 m/z-RT データ」を得る。
- 3) 標準化 m/z-RT データについてユークリッド平方距離の閾値 D_2 の下でクラスタリングを行う。
- 4) 複数の健康人血清についての標準化 m/z-RT データに共通するクラスターをランドマークとして、重心の m/z 値および RT 値を登録する。これを「登録標準ランドマーク」と呼ぶ。
- 5) 被検患者血清の LC-MS 測定データを入力してステップ 2)、3) を行い、得られた各クラスターの重心が登録標準ランドマークに一定の m/z 値および RT 値の範囲内で一致するかどうか調べる。
- 6) 上記 5) における一致率が一定の閾値以上である場合に合格とする。
- 7) ステップ 5) および 6) を、被検サンプルがなくなるまで繰り返す。

上記の手順を実行するためのプログラムを構築し、このプログラムが実データについて期待通りの動作を行うことを確認した。

1-4 当該研究開発の連絡窓口

株式会社MCBI

研究開発部長 永島 廉平

〒305-8562 茨城県つくば市東 1-1-1

独立行政法人産業技術総合研究所第 4 事業所中央第 4-1 4F

臨床インフォマティクス研究施設内

TEL 029-861-3000

FAX 029-855-5271

E-mail: nagashimarenpei@cbiri.org

第2章 本論

2-1 LCにおけるデータの再現性の確保

LCでは、使用されるカラムの劣化や試薬の品質の違い、サンプルの特質などにより、毎回の試行（ランと称する）において溶出データにバラツキが生ずる。現状では分光光度計による特定波長の吸光度を連続してモニターし、「通常得られる」クロマトグラム（画像の一種）かどうかを目視によって判断している。「通常得られる」ものと同じかどうかの判断は実験者の経験による判断に依存し、主観的なものにならざるを得ない。そこで、適当な生体試料由来の標準サンプル（例えば、よく保存された患者の血清サンプル）を用いてデータを蓄積しておき、一定の基準を設けることにより「通常得られる」クロマトグラムかどうかを判断できるとよい。正しい基準を設定できるまではトライアル・アンド・エラーでの基準の設定と目視による判断を繰り返すことになるが、一旦正しい基準が設定されれば、組込みソフトウェアがこの作業を実験者に代わって行うことができるようになる。すなわち、実サンプルについての検査の前後に標準サンプルのテストランを行って、組込みソフトウェアが出力する「良・不良」の判断からデータの再現性についての客観的な情報が得られることになる。通常、生体試料由来サンプルをLCに負荷する前に純水をサンプルとして用い（これを blank run と称する）、カラムの劣化や試薬の品質の良・不良を調べる。ついで、生体試料由来の標準および被検サンプルをLCに負荷し、被検サンプルのデータを標準サンプルの場合と比較することになる。blank run の場合に問題がないにもかかわらず、生体由来被検サンプルでは標準との差が大きいときは、その生体サンプル固有の異常があるとの判断を行い、得られたデータの解析を保留する。

2-1-1 対象としたLCデータ

1) 使用したLC装置

高速液体クロマトグラフ 2次元マイクロ自動分析システム(2D-HPLC)（島津製作所）

図1に流路図を示す。

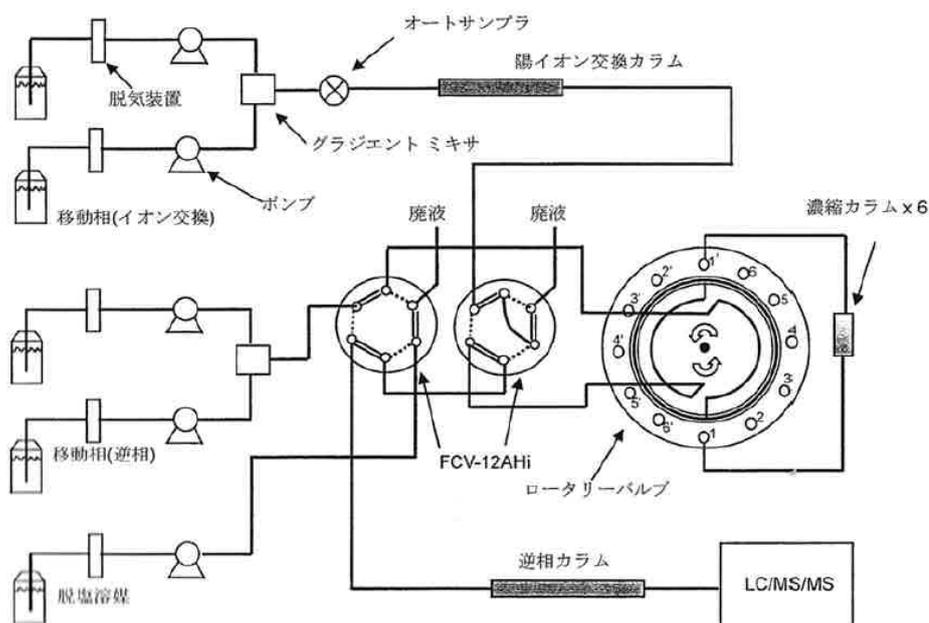


図1 2D-HPLC（島津製作所）の流路図

2) 使用サンプル

blank run : 純水

生体試料由来サンプル

標準サンプル : 健常人 5 例の血清

被検サンプル : アルツハイマー病患者 5 例の血清

4) 2D-LC

1次元目は6分画あり、本報告書ではSCX1の場合のみ例示する。SCX1は、2次元目のクロマトグラムが変化に富んでいるので、例として取り上げるのに適しているからである。(ここに記載する方法はむしろ他の分画にも問題なく適用できる。)

2-1-2 組込みソフトウェアによるデータ処理

データ処理の基本的な考え方は以下のとおりである。

図2Aの左側は、ある日に得たblank runの1次元目で得たSCX1分画を2次元目のC-18カラムで溶出したもので、横軸は溶出時間(秒)、縦軸は210nmにおける吸光度を示す。組込みソフトウェアによるデータ処理のために、これに対応する数値データを高速フーリエ変換(Fast Fourier Transform, FFT)して、図2Aの右側の図(「スペクトル」と呼ぶ)を得る。このように変換することにより、複雑なクロマトグラムが単純化され、異なるクロマトグラム間の比較を行うことが容易になる。図2Bは同じことを別の日に得たblank runについて行った結果である。

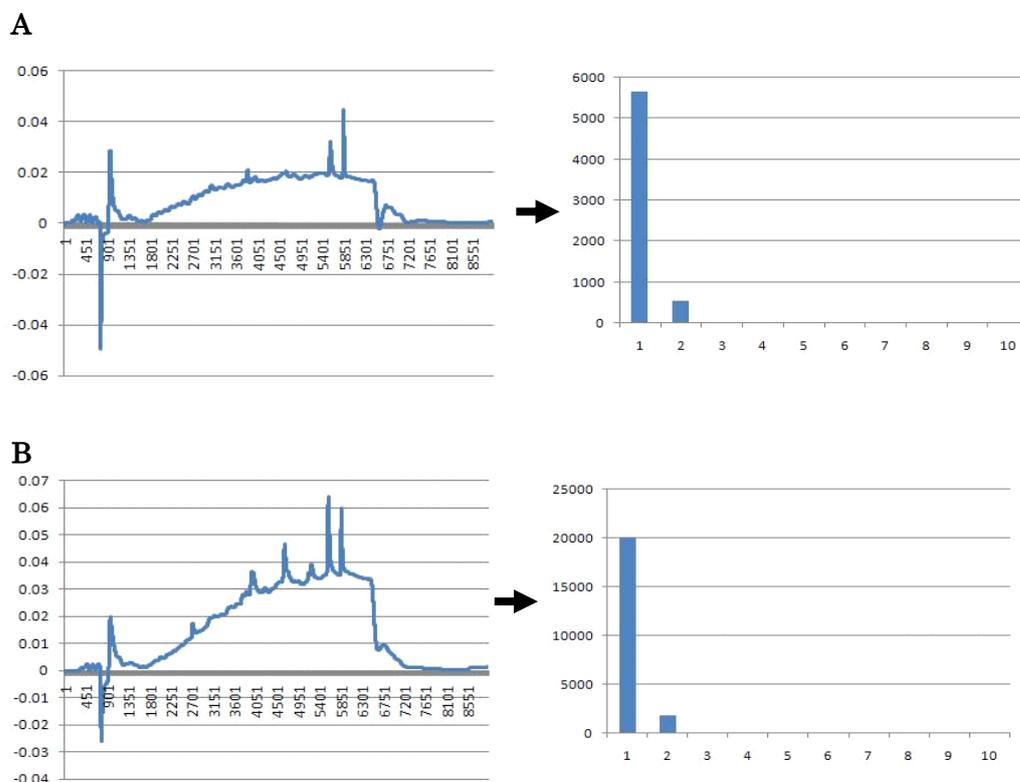


図2 クロマトグラム(左)のFFT変換(右)

つぎの手順は、スペクトル同士の比較である。図 2A のスペクトルの周波数成分を異なる次元と考えると、座標は (5600, 500) である。同様に図 2B のスペクトルの座標は (20000, 1500) となる。スペクトル 2A とスペクトル 2B の類似・非類似は、上述の 2 座標の距離が小さいか大きいかによって判断できる。

これまでおよび今後の記述で用いられるクロマトグラムデータの数値データ形式は netCDF¹であり、これを元にしたFFT変換のプログラムは、CZTアルゴリズム²によった。

2-1-2-1 blank run の場合

今、blank run のクロマトグラムとスペクトルが 6 組あるとする。通常得られる blank run の標準として、この 6 個のスペクトルの平均を求める。図 3 はそのような標準 blank run スペクトルである。

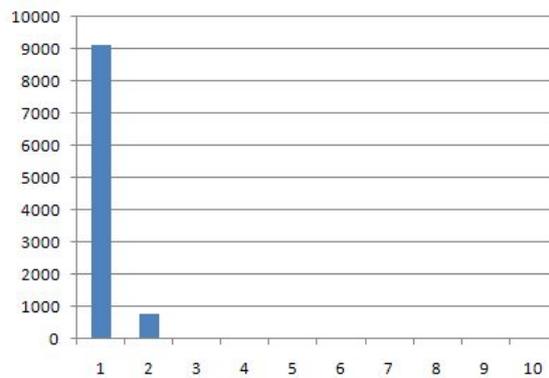


図 3 標準 blank run スペクトル (6 個の blank run の平均)

新たに blank run を行ったときは、この標準スペクトルと比較して、合否判定を行う。

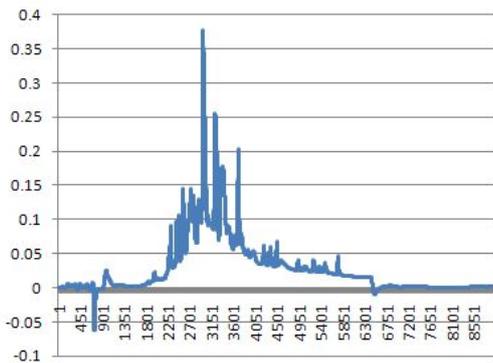
2-1-2-2 血清サンプルの場合

血清サンプルの場合も基本的に blank run の場合と同様に LC データの合否を判断することができる。ただし、この場合の標準スペクトルは健常人複数例からなるパネルの血清サンプルからの平均と標準偏差を用い、被検血清サンプルのスペクトルの合否を判定する。健常人パネル 5 例、被検サンプル 5 例について、この手続きを行った。図 4 は健常人血清の場合の 1 例を示し、図 5 は健常人 5 例の平均スペクトル (標準血清スペクトルと呼ぶ) を示す。

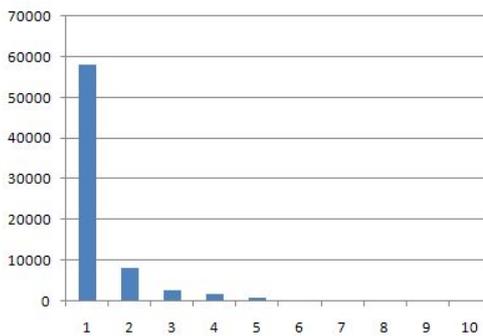
なお、健常人血清と差のある患者固有のスペクトル成分が存在する可能性はあるが、これまでのデータから、その差は波長が短く強度が非常に小さい成分に止まり、スペクトル間の距離にほとんど影響しないことが判明している。

¹ <http://www.unidata.ucar.edu/software/netcdf/>

² <http://cetusa.sakura.ne.jp/softlab/srcpatch/#czt>



091216-FHVNR166-SCX1



標準血清スペクトルからの距離 : 1243

図4 健常人血清1例の個別データ。

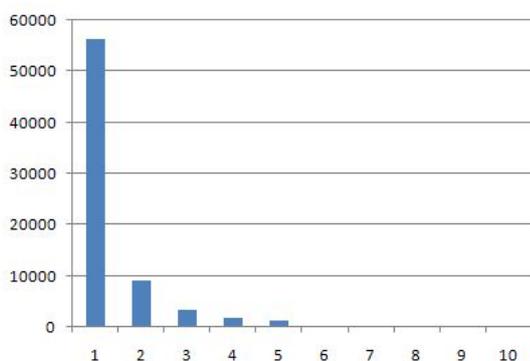
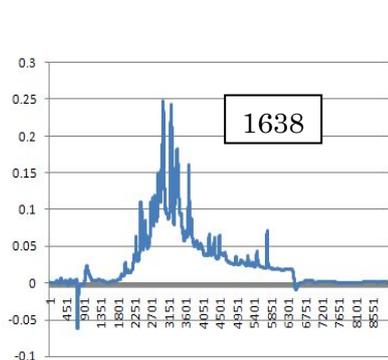


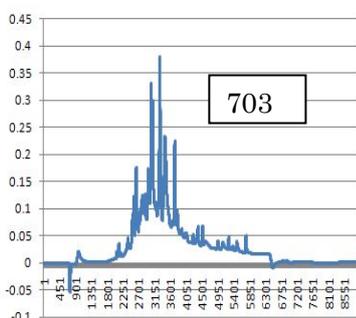
図5 標準血清スペクトル (健常人血清5例の平均)

この標準血清スペクトルからの距離を基準として被検サンプルの LC データの合否を判定することになる。

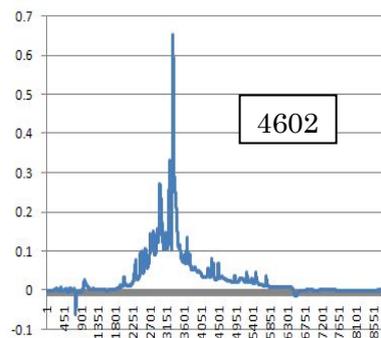
図6は被検サンプル (患者血清) 5例のクロマトグラムである。図中の数値は標準血清スペクトルからの距離を示す。



091215-FHAD001



091217-FHAD002



100106-FHAD003

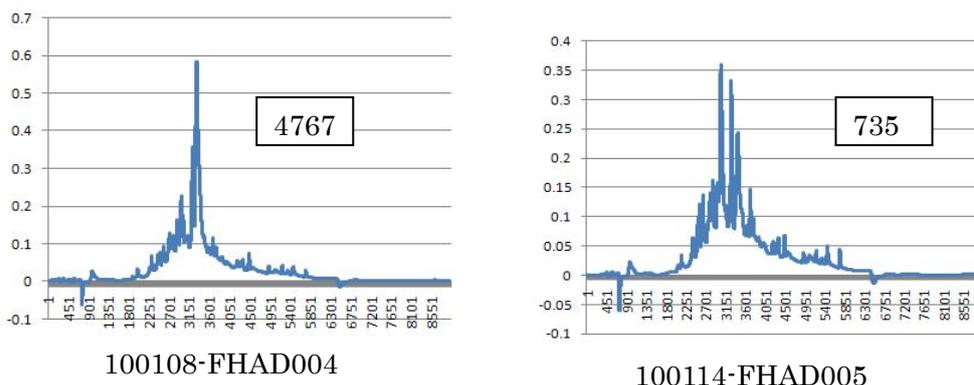


図6 被検サンプル (AD) 5例のクロマトグラムと標準血清スペクトルからの距離 (図中の数値)

2-1-2-4 特定の周波数範囲のみを抽出して比較する場合

これまでのクロマトグラムの比較は高強度の周波数が小さいもの (波長の長いもの) を主たる対象としたが、場合によっては、特定の周波数範囲のスペクトルを抽出して比較したほうが目的に合うことがある。例えば、目的の溶出成分 (複数成分の場合でもよい) を特定して、そのピークの検出の可否を調べたい場合である。調べたい周波数成分の範囲を定めることは可能なので、これまでに記載した手法を用いて、その範囲のスペクトルだけを異なるサンプル間で比較すればよい。

以上まとめると、「定量比較 LC-MS ロボット」の組込ソフトウェアとしてプログラムを構築し、実データの処理を行って期待通りの確認を行うことができた。

2-2 MSにおける質量分析データの正確な取扱いの保証

本研究開発の組込みソフトウェアに関係する国際規格は、①タンパク質・ペプチド特定のための retention time (RT) を限定できること、②質量許容誤差範囲内に MS データが収まることである。

base peak chromatogram (BPC) の分析によって、MS データの正確な取扱いの保証が可能である。ここでは、分析対象としてペプチドを想定し、標準合成ペプチドを用いた LC-MS 装置の performance の保証について記述する。

2-2-1 対象としたLC-MSデータ

1) 使用した LC-MS 装置

以下からなる解析システム 1 式

- ・ UFLC : 超高速液体クロマトグラフ UFLXR 前処理自動化システム (島津製作所)
- ・ MS : Q-STAR およびその TurboIonSpray Source (ABI)
- ・ PC サーバー・クライアント (OS)

図7 は使用した UFLC-MS 装置の流路図である。

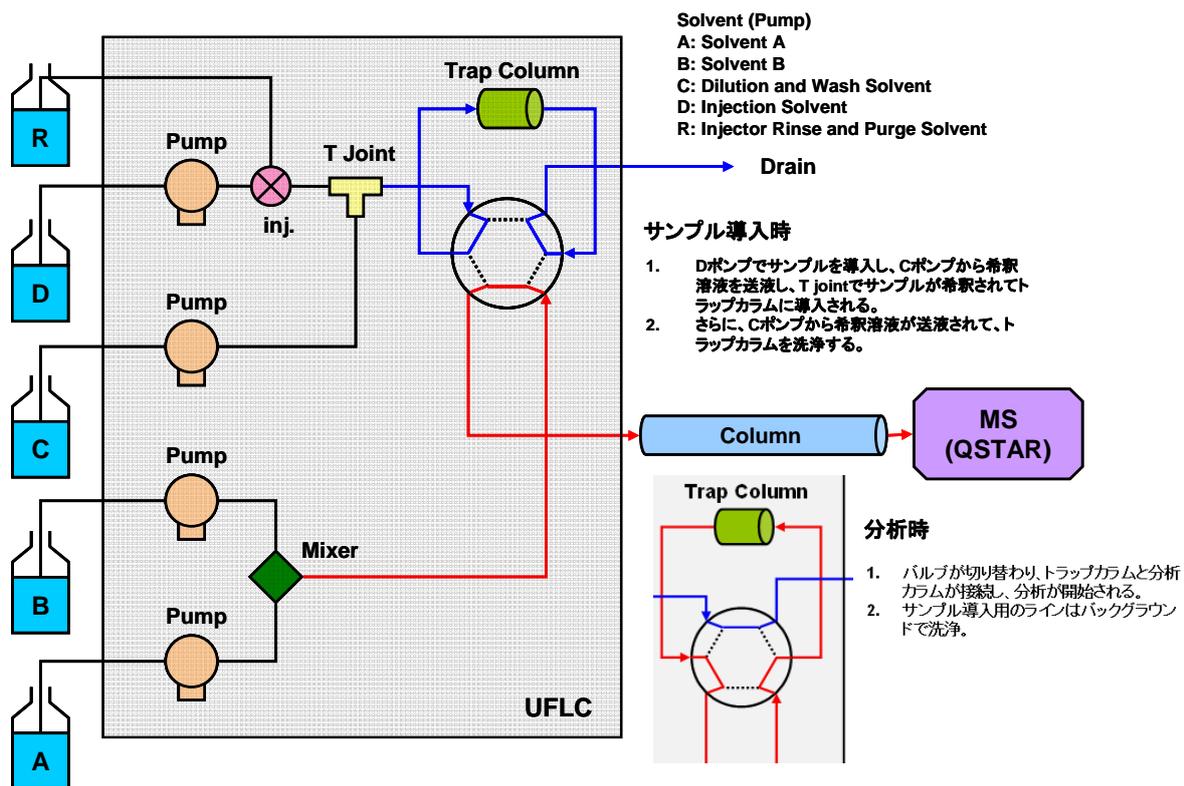


図 7 UFLC-MS 装置の流路図

2) 使用サンプル

以下のペプチド (表 1) を各々 1 pmol 含む水溶液 (標準テスト溶液)

表 1 使用したペプチド

ペプチド	m/z	
	[M+H] ⁺ (mono)	2 価 ion
ADP10xy	1276.56	638.78
ADP13pq	1487.72	744.36

2-2-2 分析結果

数値データは mzXML 形式で与えられる。このデータから各分画 (RT) の最大信号のみを RT に対してプロットし BPC を得た。Run 1~Run 3 までの 3 回実験して得た BPC を図 8 に示した。UFLC の 1 分画は 6 秒 (s) である。

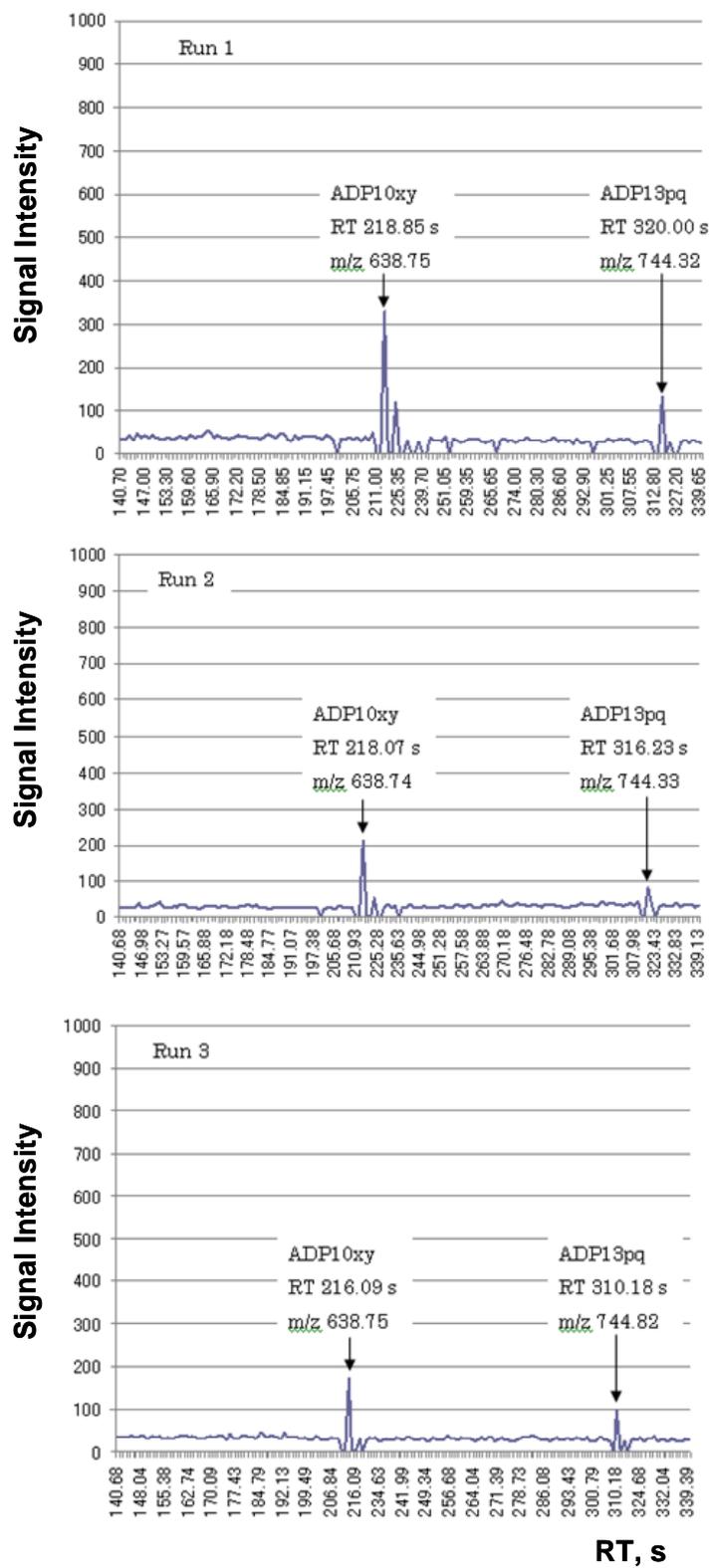


図8 2種のペプチドADP10xyとADP13pqを含む標準テスト溶液のUFLC-MSによるBPC(3回の繰り返し実験により再現性を検討した。)

2-2-3 UFLC-MSデータの保証

標準ペプチドを用いた UFLC-MS データの保証は、装置から出力される BPC に関する数値データについて、標準ペプチドの期待される m/z 値 ± 0.5 以内に入る信号が、期待される RT 値（例えば ± 30 s）に検出されるかどうかの合否判定によってなされる。そのためのプログラムは、BPC 出力のプログラムに、 m/z 値と RT を survey して期待される信号の有無と合否の結果を出力するプログラムを追加すればよい。

以上まとめると、標準合成ペプチドを用いて LC-MS 装置の RT および MS に関する performance を検証するプログラムを構築し、実データについて質量分析データの正確な取扱いの保証を行うことができた。

2-3 標準サンプルを用いたLC-MSロボットのperformanceの検証

LC と MS の両者を組み合わせた LC-MS ロボットについては、総合的かつシームレスにその performance を検証できるようにすることが求められる。本研究項目では LC および MS の両者の performance を統合的に検証できる機能をもつ組み込みソフトウェアを開発する。

以下、2-3-1 から 2-3-5 まで、LC 装置として高速液体クロマトグラフ 2次元マイクロ自動分析システム(2D-HPLC)（島津製作所）、MS 装置として Ultraflex TOF/TOF (BRUKER DALTONICS) を使用した場合について記述した。UFLC（超高速液体クロマトグラフ UFLXR 前処理自動化システム、島津製作所）および MS 装置として Q-STAR とその TurboIonSpray Source (ABI) を使用した場合については、2-2における記述を参照されたい。

2-3-1 基本的な考え方

MS によって得られる質量分析データは基本的には個々のサンプル・ウエル[2次元目の LC における溶出分画に相当し、溶出時間 (retention time, RT) と 1:1 に対応する]に関する出力であって、横軸に m/z （質量/荷電）を目盛り、縦軸にシグナル強度を目盛った一種の画像データである。実際のデータでは m/z を異にする多数のイオンのピークが見られる。複数の連続した溶出分画についてシグナル強度を可視化したものが図 9 の heatmap 表示である。この図では縦軸に連続した分画が目盛られ、横軸に m/z が目盛られている。



図 9 連続した LC 分画ごとの MS 強度の heatmap 表示

図 9 の左と右は異なる健常人の血清に関するデータであるが、 m/z 値を異にし、連続した 4~6 分画に現れた 4 個の正方形で囲まれた左右で共通したシグナルを見てとることができる。質量分析データの正確な取扱いの保証は、健常人血清に見られるこのようなシグナル（以後「ランドマーク」と呼ぶ）が各データについて認められる場合に与えられる。実際のデータ処理は複数の血清からのデータについて heatmap のもととなった数値データを用いてクラスターとしてランドマークを抽出し、その各クラスターが標準血清サンプルの場合の LC 分画および m/z ($0.5 m/z$ 以内の誤差を許す) と一致する場合に合格とし、そうでないデータを排除することによって行う。

以下の記載では、LC において溶出分画の代わりに RT (retention time) を用いる。すなわち、図 9 の縦軸は RT (秒、s) として表現される。今回のデータでは、1 分画は 6 s に相当する。

2-3-2 クラスタリングの方法

本組込みソフトウェアでは、ランドマークの抽出は Ward 法によるクラスタリング³によって行う。青木によると、Ward 法は分類感度が高く、最も明確なクラスターを作るとされている。

手順を以下に記す。すべての点同士の平方距離を計算し、最短のペアから順に Cluster 1、Cluster 2 のようにクラスターを作り、番号を付ける。予め、閾値 (D_t) を定め、その閾値を越えるペアはクラスターとしないことにする。クラスター自身は 1 個の点（その位置はクラスター内の点の重心とする）とみなされ、それに近い点を併合する。新たなクラスターの番号は最初のもので同一とする。新たなクラスターが作られるたびに、その重心がいずれも点とみなされる。次に行おうとする併合が平方距離 D_t を越えるとき、それ以上クラスターを作らないと定める。点の数が多く、 D_t が十分小さいときは、最初に作られたペアのクラスター番号をそれぞれ持つ多数のクラスターができる。

2-3-3 対象とした LC-MS データ

1) 使用した LC-MS 装置

- ・ LC : 高速液体クロマトグラフ 2次元マイクロ自動分析システム(2D-HPLC) (島津製作所)
- ・ MS : Ultraflex TOF/TOF (BRUKER DALTONICS)

2) 使用サンプル

健常人および患者血清

3) 数値データの形式

一般的なデータフォーマットの mzXML 形式。

2-3-4 LS-MS データとその処理

1) 方針

健常人血清サンプルから得た LS-MS データを標準とし、これを患者血清から得たデータと比較する。

2) 前処理

091216-FHVN166-SCX1 (健常人) および 091215-FHAD001-SCX1 (患者) の LC-MS データ

³ 神島敏弘 : 人工知能学会誌 18 : 59, 2003; 青木繁伸 : <http://aoki2.si.gunma-u.ac.jp/lecture/stats-by-excel/vba/html/clustan.html>

のうちピーク強度 2000 以上のピークを横軸 m/z 、縦軸 RT の平面上にプロットして図 10 を得た。

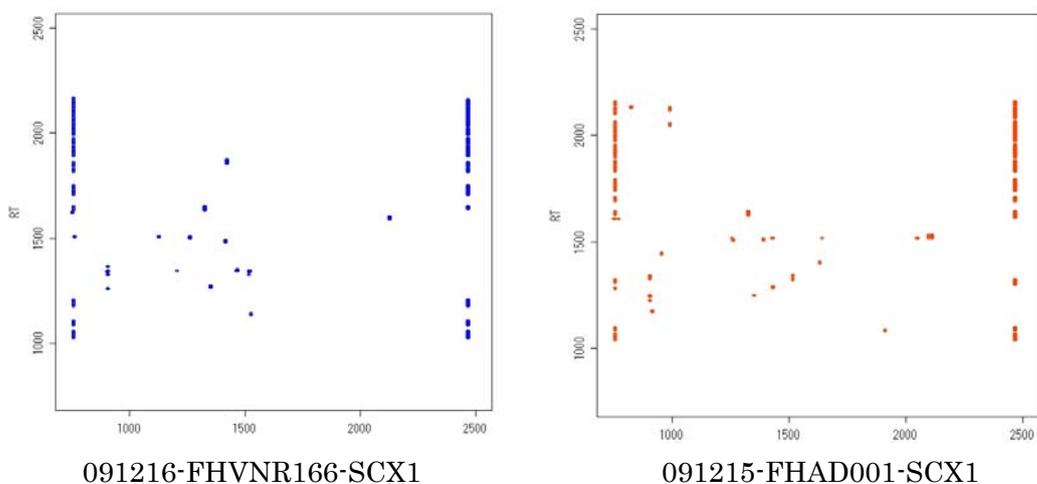


図 10 ピーク強度 2000 以上のピークに限定した LC-MS データ

多数のサンプルについての比較を可能にするために、全サンプルに加えられた一定量 (250 fmol/well) の内部標準によるピーク (図 10 の左端の連続するピーク) の強度で血清由来のピークの強度を除して標準化し、標準化後のピーク強度が 3.0 以上のピークのみを選んだ。また、連続する分画数が少ないピークは、再現性が乏しいと判断されるので、少なくとも 3 分画以上で出現するピークのみを選んだ。これら 2 つの限定により、図 11 が得られた。

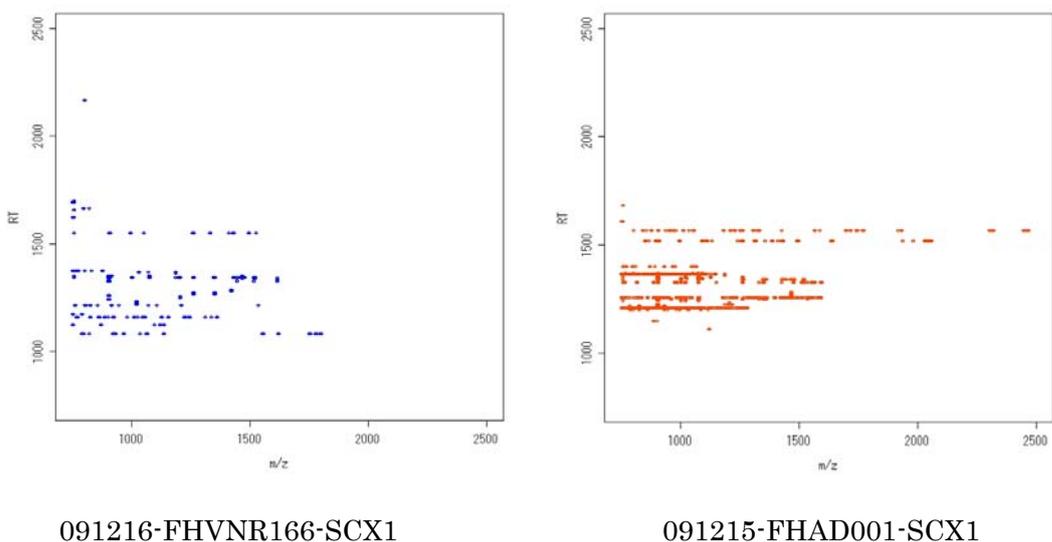


図 11 ピーク強度を内部標準のピーク強度で除して標準化し、標準化後のピーク強度が 3.0 以上、RT が 3 分画以上連続するピークの m/z -RT プロット

3) クラスタリング

次に図 11 の m/z -RT データについて、クラスタリングを行った。ユークリッド平方距離の閾値

D_2 を定め、健常人血清と患者血清で対応するクラスター（すなわち、ランドマーク）の間のユークリッド平方距離が D_2 より小さい場合は、同一のランドマークとみなした。

図 12 は、健常人と患者のランドマークの比較である。両者で一致する RT-m/z 平面上の位置にランドマークがあることが理解できる。

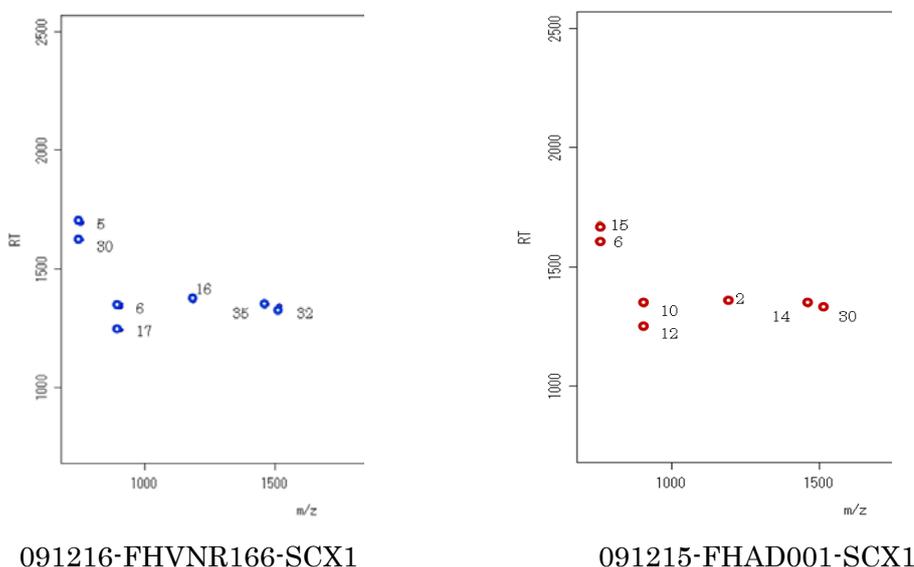


図 12 健常人（左）と患者（右）のランドマークの比較

2-3-5 ランドマークの応用

ランドマークとしては標準サンプルを用いて同定されたものが適当である。標準サンプルとしては通常、健常人由来の生体試料に規定の前処理を加えたものがよい。ここでは、生体試料を血清として、図 17 の健常人血清のクラスターをランドマークとして用いる場合を例示する（表 2）。

表 2 図 17 のランドマークの測定された m/z と RT（複数の健常人血清に共通するクラスターから得られたランドマークを登録標準ランドマークという）

ランドマーク	クラスター番号	m/z	RT (s)
LM1	5	755.0	1692
LM2	30	755.0	1620
LM3	6	905.0	1338
LM4	17	905.0	1242
LM5	16	1187.5	1362
LM6	35	1465.0	1344
LM7	32	1517.5	1338

次のステップとして、被検血清（疾患サンプルなど）から得た数値データについて、同じ手法でクラスタリングを行う。ついで、このステップで得たクラスターについて、標準サンプルのランドマークに対応するクラスターがあるかどうかを調べる。合否判定の基準となる m/z 値ならびに RT 値（両者のセットを登録標準ランドマークと呼ぶ）およびその範囲については、複数の健常人血清サンプルからの十

分なデータ蓄積を得て、決定しなければならない。

登録標準ランドマークは複数の健常人サンプルに共通のクラスターとして定義され、ランドマーク数に関する限定はない。被検サンプルについても、サンプルごとにクラスター数は異なるはずである。したがって、登録標準ランドマークの総数のうち何個が被検クラスターと m/z 、 RT の点で一致するかの一一致率により合否判定することになる。合否判定のためのプログラムには一致率についての閾値を予め入力できるようにしておく。

以上まとめると、LC-MS ロボットの performance を検証するためのプログラムを構築し、このプログラムが実データについて期待通りの動作を行うことを確認した。

参考文献

1. <http://www.unidata.ucar.edu/software/netcdf/>
2. <http://cetus.sakura.ne.jp/softlab/srcpatch/#czt>
3. 神島敏弘：人工知能学会誌 18：59, 2003; 青木繁伸：<http://aoki2.si.gunma-u.ac.jp/lecture/stats-by-excel/vba/html/clustan.html>

最終章 全体総括

1. 研究開発成果

以下の課題について、研究開発を行った。成果を課題ごとに記載する。

① LCにおけるデータの再現性の確保

①純水および②測定対象とする生体試料をサンプルとして LC 装置に負荷して得られる数値的データを基に、本研究開発の組込みソフトウェアが期待外の異常が何ら生じないこと、および繰り返し試行でのバラツキが少なく一定範囲内であることを確認する。プログラムを構築し、実データの処理を行い、期待通りの確認を行うことができた。

② MSにおける質量分析データの正確な取扱いの保証

標準合成ペプチドを用いて LC-MS 装置の RT および MS に関する performance を検証するプログラムを構築し、実データについて国際規格に基づく質量分析データの正確な取扱いの保証を行うことができた。

③ 標準サンプルを用いた LC-MS ロボットの performance の検証

2D-HPLC および Ultraflex TOF/TOF を使用した LC-MS ロボットの performance を検証するためのプログラムを構築し、このプログラムが実データについて期待通りの動作を行うことを確認した。

2. 研究開発後の課題・事業展開

2-1 研究開発後の課題

以下の課題があるので、それぞれの技術的目標値を記載する。

2-1-1 組込みソフトウェアによるデータ処理時間の短縮

処理時間の短縮では、一日あたり 500-1,000 サンプルの処理ができることを目標とする。

2-1-2 サンプルから得られる情報の可視化とこれによる比較解析機能の構築

これまでに開発したプロトタイプより可視化、比較解析機能とも格段に見やすく、使い勝手の良いものに改善する。完成すれば新規かつ革新的な組込みソフトウェアとなる。

2-1-3 複数のサンプルに関する統計解析機能の付与

これまでに開発したプロトタイプでは、統計解析機能が不十分であるので、これを拡充する。前項と同様、完成すれば新規かつ革新的な組込みソフトウェアとなる。

2-2 事業展開

2-2-1 事業化目標

組込みソフトウェアを組み込んだ定量比較 LC-MS ロボットの研究開発終了 (H24 年) から 5 年後の売上目標は以下のとおりである。

国内 50 億円 (シェア 60%、売り上げ 30 億)

国外 500 億円 (シェア 18%、売り上げ 90 億)

総売上 550 億円 (組込みソフトウェア 27.5 億)

組込みソフトウェアは定量比較 LC-MS ロボット売価の約 5% とする。

2-2-2 事業化の実施体制

株式会社 MCBI、株式会社島津製作所、医療機器販売会社（未定）によって実施体制を形成する。株式会社 MCBI は組込みソフトウェアを川下製造業者の株式会社島津製作所や医療機器製造会社に販売する。またそのバージョンアップなどのメンテナンスを有償で行う。

2-2-3 事業化計画

定量比較 LC-MS ロボットの対象は臨床検査市場である。想定する市場としては中間業者ないしエンドユーザーで考えた場合、以下の複数の市場が対象となる。

- ・医療機器市場
- ・体外診断薬市場
- ・予防医療・健康支援市場

製品・サービスと事業化手段は以下のとおりである。

①血液診断バイオマーカーのための定量比較 LC-MS ロボット

臨床検査システムとして1システムあたり 30,000,000 円

（当社はのうち組込みソフトウェア：1,500,000 円）

当該ロボットを使用した検査として、これまでの当社ならびに共同研究機関の研究成果としてある肝疾患（肝炎、肝硬変、肝がん）の高精度診断、高血圧・動脈硬化などメタボリック症候群から発症する疾患、現在定量的検査が困難な精神神経疾患が対象となる。これらの疾患における本検査の臨床有効性を前述の医療機関との共同研究、ならびに別途共同研究を結んでいる体外診断薬メーカーと臨床有効性試験を行い、厚生労働省への申請を行うことを同時に進める。その承認前に、これらの共同研究を実施した医療機関における評価ならびに成果を活用して販売促進を行う。具体的には、学会での発表、展示会での周知、さらに、すでに既存検査システムの販売や検査受託で実績のある共同事業先との連携によって、本検査システムの有用性を顧客に理解してもらうようにするなどの販売促進の工夫を行い、事業目標を達成する（事業化の実施体制を参照）。

②体外診断薬

保険適応を得た場合、1検査あたり 5,000 円（ライセンス料 10%）

③予防医療・健康支援

自由診療・自主健康診断である。受託検査の形で1検査あたり 10,000 円（当社マージン 30%）