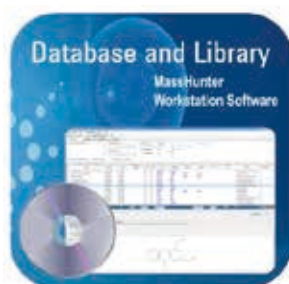


アジレントの精査された化合物データベース およびライブラリによる未知化合物の同定の 信頼性向上



著者

Emma E. Rennie,
Robert H. Williams,
Ruben Garnica, and
Maria VanDamme

概要

ターゲット化合物および未知化合物の同定の信頼性は、同定に用いられるデータベースおよびスペクトルライブラリに登録されているデータの質、すなわち真正性と完全性によって決まります。データベースおよびライブラリの情報は、インターネットで公開されている他、多くのベンダーからも提供されていますが、その大半がクラウドソーシングによって収集されたもので、情報の確認はほとんど、またはまったく行われていないのが現状です。ライブラリの品質が低いと、ラボの結果の品質にも悪影響がおよぶ可能性があります。

博物館の館長や美術品のコレクターが美術品や芸術品を収集するときには、広範な調査によってその作品一つひとつの真正性と完全性を評価します。同じように、アジレントの化合物データベースおよびライブラリに含まれるデータや精密質量スペクトルも、それぞれの品質を確保するために、きわめて詳細な精査プロセスで作成されています。

このホワイトペーパーでは、アジレントの精密質量ライブラリおよびデータベースがどのように開発されているのか、その精査プロセスはどのようなものなのか、またこのプロセスによってラボの結果の品質がなぜ向上するのかを紹介します。

Agilent 精密質量パーソナル化合物データベースライブラリとは

Agilent 精密質量パーソナル化合物データベースライブラリ (PCDL: Personal Compound Database and Library) は、詳細な化合物情報のデータベースと、四重極飛行時間型質量分析計 (Q-TOF) で採取された高分解能精密質量スペクトルのライブラリで構成されており、ガスクロマトグラフィーおよび液体クロマトグラフィーのアプリケーションに利用できます。PCDL と Agilent MassHunter データ解析ソフトウェア、TOF または Q-TOF を組み合わせて使用することで、化合物を精密質量、同位体比率、精密質量プロダクトイオンスペクトル、および保持時間 (RT:Retention Time) の最大 4 つの情報確認によって同定できます。Agilent イオンモビリティ四重極 TOF (IM-Q-TOF) では、さらに衝突断面積 (CCS) という追加の情報からの確認もできます。PCDL があれば、特定の対象化合物をターゲットにする場合も、数千種類のターゲットをスクリーニングする場合も、各化合物の標準物質は必要ありません。また、各 PCDL に付属する Agilent MassHunter PCDL Manager ソフトウェアパッケージを使用して、RT、新規化合物、および新規質量スペクトルの他、IM-Q-TOF MS の場合はイオンモビリティの CCS の情報を追加するなど、特定の分析またはシステム構成に合わせて PCDL をカスタマイズすることも簡単にできます。

精密質量スペクトルに同位体比率の理論値を組み合わせると、非常に信頼性の高い化合物の同定が可能になります。使用するクロマトグラフィーメソッドに固有の RT を追加すれば、信頼性がさらに高まります。RT の追加は、異性体化合物を分析する場合に特に有効です。また、ターゲット PCDL に化合物を追加して、過去のサンプルに遡ってデータ解析を簡単に行えます。さらに、Agilent MassHunter Qualitative Analysis Workflows ソフトウェアに搭載されている自動精査ワークフローを実行するときは、アジレントのケミストが使用するものと同じインポート設定で、精査済みの新しい LC/MS/MS スペクトルを簡単に追加できます。

アジレントの精密質量データベースおよびライブラリは、Agilent MassHunter ソフトウェアの定性分析および定量分析の両方で使用できます。MassHunter PCDL Manager は、Agilent MassHunter ソフトウェアスイートに含まれる Qualitative Analysis (定性分析) および Quantitative Analysis (定量分析) データ解析ソフトウェアパッケージや、Mass Profiler、Mass Profiler Professional、Profiler などの統計解析および視覚化ソフトウェアパッケージと直接連携します。

すべての Agilent MassHunter PCDL 製品には、データサンプルおよび取り扱い方法を参照できるユーザーガイドが付属しています。また、アップグレードによって、初回購入後も常に最新の情報を利用できます。GC/Q-TOF 用 MassHunter 農業 PCDL およびワークフローなど一部の PCDL には、e メソッド、推奨構成、およびワークフローガイドも含まれ、PCDL および化合物同定ワークフローを分析ラボにすばやく導入できます。

Agilent 精密質量パーソナル化合物データベースライブラリの開発

各 PCDL は、次の 4 つの主要ステップで構成される詳細な標準手順に従って作成されます。

1. 第一線の分析エキスパートとの協力により精選されたターゲット化合物リストが作成されます。
2. 化合物データベースを作成し、各ターゲット項目の情報を検証します。
3. 高純度の標準物質を使用してプロダクトイオンスペクトルデータが収集されます。
4. アジレントの厳格な品質管理精査プロセスに従って精査されたプロダクトイオンスペクトルのライブラリが作成されます。

ステップ 1: 外部エキスパートとの協力による化合物リストの作成

データベースおよびライブラリ作成プロセスの第一ステップでは、PCDL に含める、または追加する化合物が決定されます。アジレントは、研究分野ごとに関連性の高い包括的な最新の製品を作り上げるために、幅広い分野の第一線の分析エキスパートと提携し、これらのターゲットリストの優先順位を決定しています。

各 PCDL 製品の内容を精選した後 (内容リストの例については、図 1 を参照)、プロダクトイオンスペクトルを収集する化合物の優先順位が決定されます。

ステップ 2: 各化合物のデータベース項目の作成

PCDL に含める各ターゲット化合物のデータベース項目が作成されます。各項目には、化合物に関するさまざまな情報が含まれ、各情報が手作業で検証されます。経験を積んだアジレントのケミストにより、分子式、化合物名、構造、CAS 登録番号、および化合物類タグが評価の高い複数のリソース (SciFinder、ChemSpider、PubChem、ChemIDplus、Chemistry Dashboard など) に照らして検討されます。また、個々の化学的な情報がその情報元サイトで手作業により検証されます。この作業を通して、各項目に含まれる化合物情報で高度な正確さが確保されます。

データベースに登録される化学的な情報の具体的な種類は、PCDL のアプリケーション分野によって異なります。これには次のような項目があります。

- 一般名
- IUPAC 名
- 別名
- 分子式
- CAS 登録番号
- ChemSpider ID
- PubChem ID
- InChI
- SMILES 表記
- ベンダー製品コード

代謝物については、KEGG、METLIN ID、HMDB ID、LMP ID、ChEBI ID、BioCyc ID などの代謝物系の ID が含まれます。

Compound Name	CAS #	Compound Name	CAS #
(1R)-cis-Permethrin	54774-41-9	Acetamidic acid	105419-29-7
(1R)-trans-Permethrin	61989-77-7	Acetaminophen	14256-82-1
1,2,3,5-tetrachlorobenzene	834-90-3	Acibenzolar-S-methyl (BTH)	135158-54-2
1,2,3-Trichlorobenzene	87-41-6	Acifluorfen-methyl	50594-67-7
1,2,4,5-tetrachlorobenzene	95-54-3	Aciclovir	74073-46-6
1,2,4-Trichlorobenzene	120-82-1	Acinathrin (Rufact)	103833-18-7
1,2-Dibromo-3-chloropropane	96-12-9	Akton	1757-18-2
1,2-Dichlorobenzene (o-Dichlorobenzene)	95-50-1	Alachlor	15672-90-9
1,3-Dichlorobenzene (m-Dichlorobenzene)	541-70-1	Aldrin	309-80-2
1,4-Dichlorobenzene (p-Dichlorobenzene)	106-46-7	Alfodrin	584-79-2
1-Naphthaleneacetic acid	85-47-3	Alidochlor	83-71-0
1-Naphthal	90-15-3	Ametostatin	855318-97-4

図 1. 各 PCDL には、登録されているすべての化合物とそのケミカルアブストラクツサービス (CAS) 登録番号の内容リストが含まれています。

Compounds view, containing accurate and up-to-date compound information:
Common name, IUPAC name, accurate mass of neutral molecule, molecular formula, ChemSpider, CAS, PubChem, KEGG, METLIN ID, HMDB ID, LMP ID, InChI, SMILES, vendor product codes, and so forth

Common name:
Verified from multiple different sources for commonality

Structure and Notes Area:
• Structure
• Compound class tags
• Regulation tags and references
• Chinese, Japanese and English synonyms
• Toxicology research references
• Outdated and alternative CAS identifiers
• Compound descriptions

図 2. Agilent MassHunter PCDL Manager: 一般的な化合物データベース項目が表示されています。

各データベース項目には、各化合物の化学構造およびアノテーションも含まれています。注記は化合物によって異なり、次のようなものがあります。

- 化合物、規制、および国固有の化合物類タグ
- 中国語名、日本語名、および英語名
- 毒物学研究参考文献

- 削除された (または古い) CAS 登録番号と代替 CAS 登録番号
- 化合物の説明

各化合物項目についてデータベースに登録される上記の情報は、PCDL Manager で簡易検索または詳細検索機能を使用してすべて検索することができます (図 2 を参照)。

これらの項目を基準にデータベースをフィルタリングし、各 PCDL 製品からより小規模で編集可能な PCDL のサブセットを作成することも可能です。例えば、特定の化合物グループまたは存在が疑われる化合物リストのスクリーニングワークフローを作成するために、独自のカスタム PCDL を作成することができます。

最後に、各化合物項目が Agilent Master PCDL と照合され、重複する化合物を排除することができます。これにより、データ解析での結果の確認時間が大幅に短縮されます。

一般名、CAS、構造、注記などの化合物識別情報は、MassHunter データ解析ソフトウェアおよび MassHunter PCDL Manager から利用できます (図 3 を参照)。

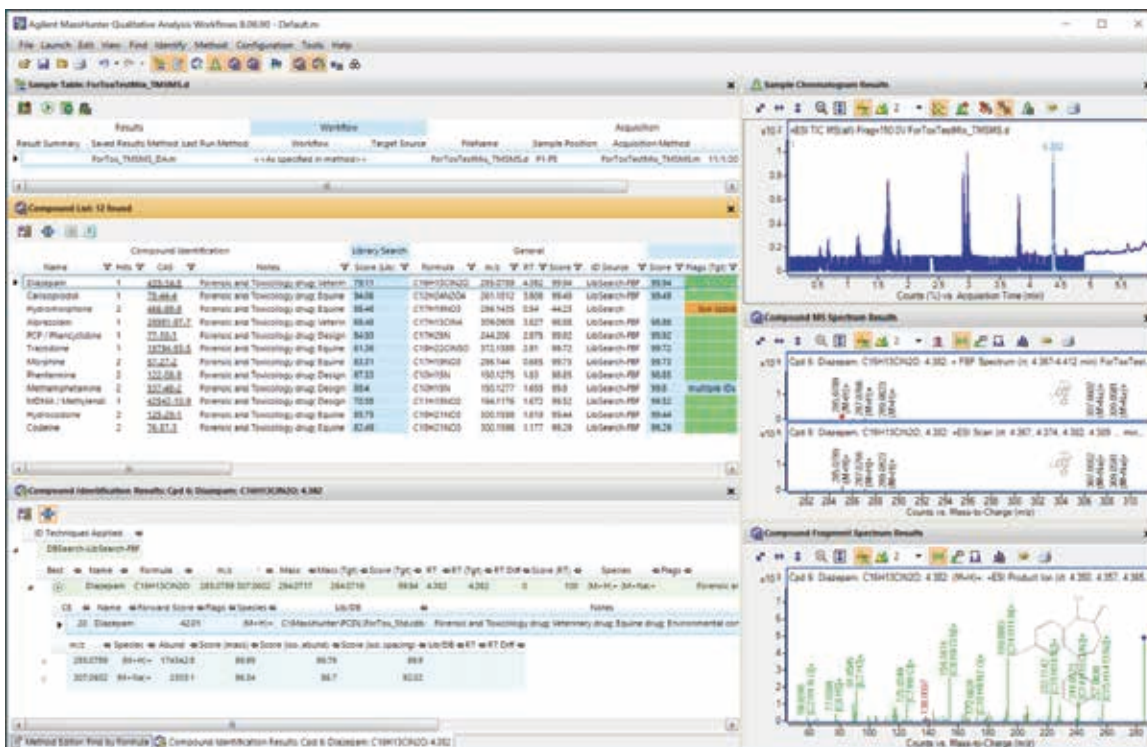


図 3. Agilent MassHunter Qualitative Analysis Workflows: 構造、CAS、注記など、LC/MS/MS で分析された化合物の同定結果が表示されています。

ステップ 3: 高純度標準物質による質量スペクトルデータの収集

高純度の単一化合物標準物質を使用して精密質量スペクトルが採取されます。データの採取は、常に厳格な標準操作手順書 (SOP) に従って行われます。SOP には、機器のチューニングおよびキャリブレーションの他、品質管理サンプルのデータ採取方法が規定されているため、すべてのデータを同一の実験条件下で採取することができます。真菌または細菌代謝物など単一の標準物質として入手できない化合物は、アジレントと提携している第一線の科学協力エキスパートによって合成されます。

GC/MS 精密質量スペクトルは、豊富なフラグメンテーションパターンと同位体クラスタ情報が得られる EI フルスキャンモードで収集されます (図 4 を参照)。

LC/MS/MS 精密質量スペクトルは、 m/z が最小の正および負のモノアイソトピックイオンについて、複数の付加イオン種とコリジョンエネルギーで収集されます (図 5 を参照)。多様な安定性をもつ化合物について豊富なプロダクトイオンスペクトルを提供するために、各化合物は 10、20、および 40 eV のコリジョンエネルギーで採取されます。低い電圧である 10 V のスペクトルでもフラグメンテーションが過剰な場合 (非常に不安定な化合物) や、高いエネルギーの 40 V のスペクトルでも解裂が不十分な場合は (非常に安定性の高い化合物)、必要に応じて追加のコリジョンエネルギーでスペクトルが採取されます。

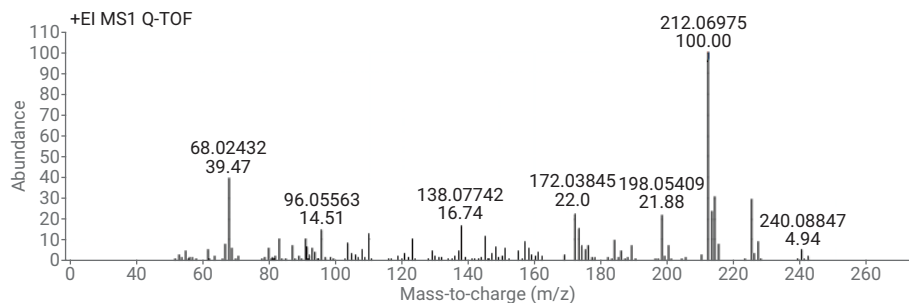


図 4. 農薬シアナジンの GC/MS EI 精密質量スペクトル

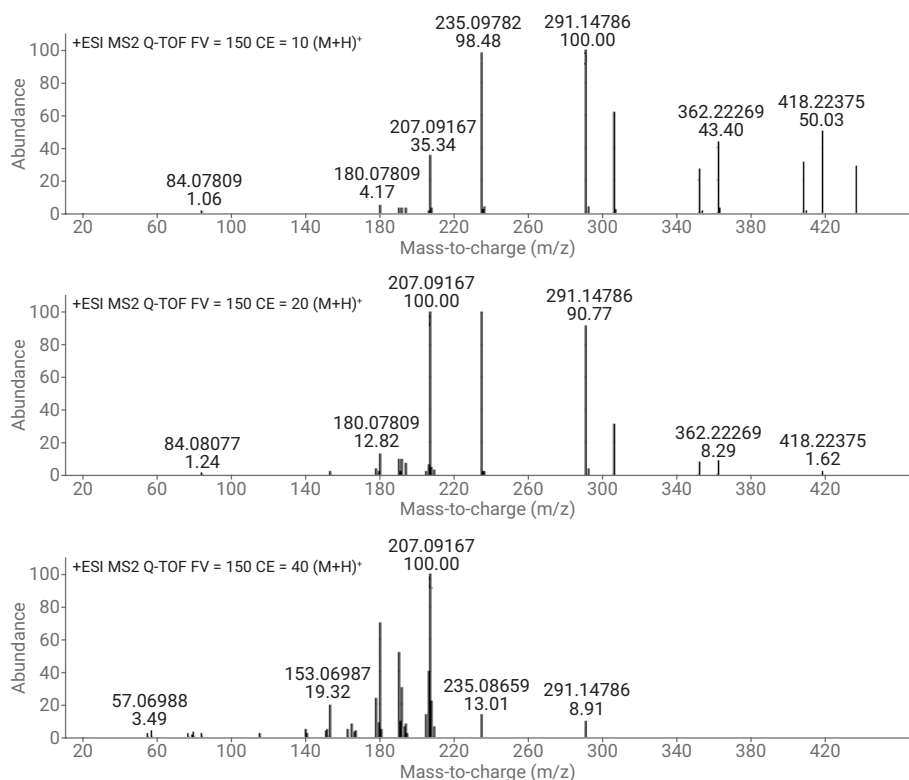


図 5. 薬物バルサルタンの LC/MS/MS ESI 精密質量スペクトル

ステップ 4: 優先化合物リスト のライブラリスペクトル項目の 作成

アジレントのアプリケーションケミストまたはアジレントと提携している協力エキスパートにより、適切な SOP に従って質量スペクトル情報が採取されます。このとき選択される SOP は、使用するクロマトグラフィーメソッドによって異なります。

LC/MS/MS スペクトルの精査

MassHunter Qualitative Analysis データ解析ソフトウェアに搭載されているアルゴリズムを使用して、LC/MS/MS 精密質量スペクトルが抽出され、解析されます。まず、Find by Formula アルゴリズムによってスペクトルが抽出されます。次に、分子式生成 (MFG) 機能により、質量スペクトルの各 m/z プロダクトイオンピークに組成式が割り当てられます。その後、各ピークについて、その組成式がプリカーサイオンの有効な部分組成式かどうかを検証されます。ほとんどのプロダクトイオンピークでは、MFG の質量許容範囲内に、割り当てられる組成式候補が複数存在します。その場合、イオンピークは、質量の最も近い組成式に割り当てられます。このプロセスで使用される MFG の質量許容範囲は質量によって異なり、ライブラリ検索のデフォルト質量許容範囲の約 1/3 です。スペクトルを PCDL にインポートするときに、すべての m/z イオンピークが理論的な精密質量に合わせて補正されます (図 6A および 6B を参照)。プリカーサイオンの部分組成式に割り当てることのできないプロダクトイオンピークは、化学的ノイズまたは不純物として排除されます。また、ベースピークの 1% 以下のプロダクトイオンピークは、常にノイズとして排除されます。

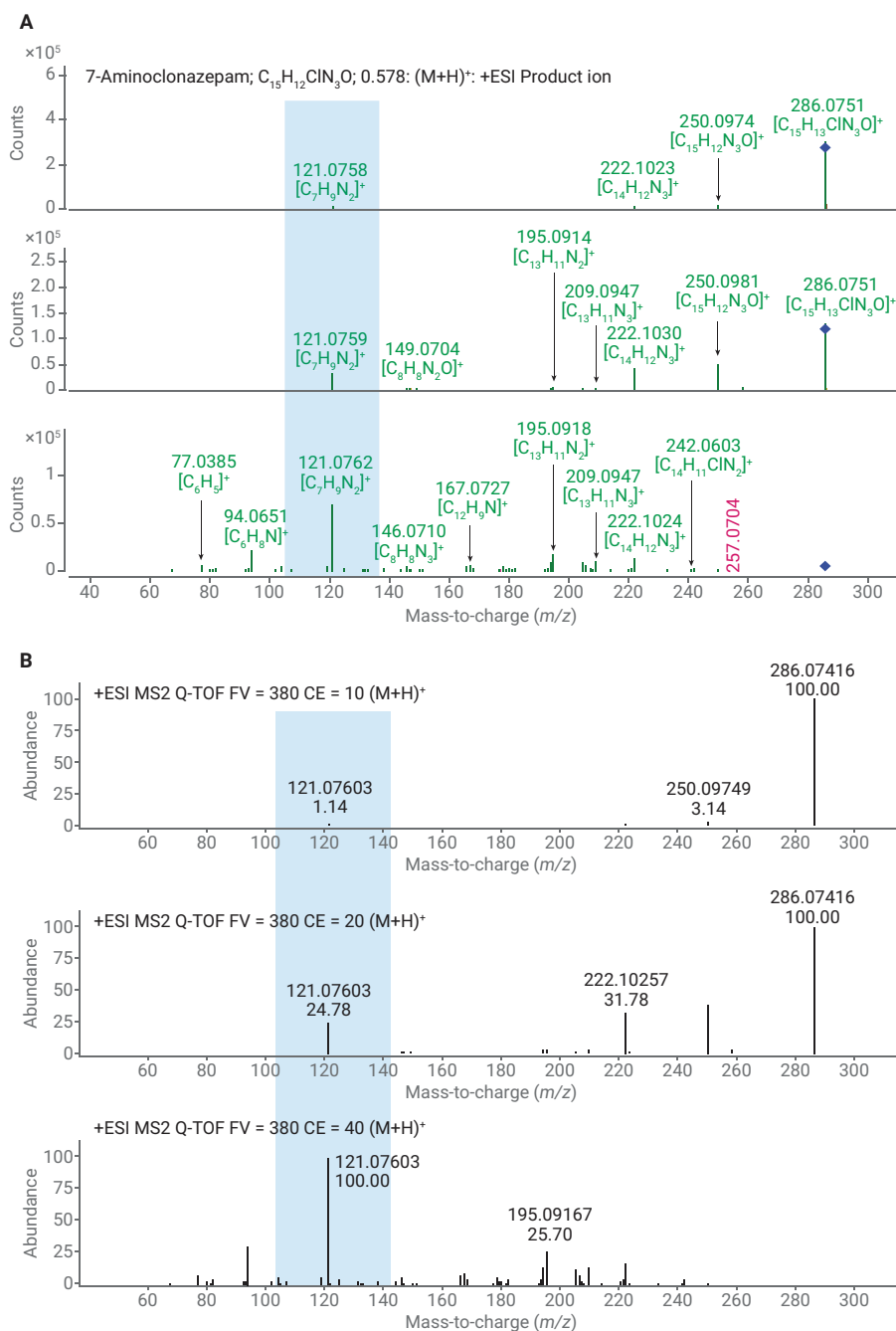


図 6. A) 実験により得られた 7-アミノクロナゼパムの LC/MS/MS ESI 精密質量スペクトルと B) 理論的な精密質量に合わせて補正された 7-アミノクロナゼパムの LC/MS/MS ESI スペクトル。青でハイライトされた領域は質量が補正されています。

GC/MS スペクトルの精査

MassHunter Qualitative Analysis データ解析ソフトウェアに搭載されているアルゴリズムを使用して、精密質量スペクトルが抽出され、解析されます。まず、特徴的なイオンの抽出イオンクロマトグラム (EIC) をもとに、スペクトルが手作業で抽出されます。次に、MFG により、化合物の既知の分子式にもとづいて、同位体比や同位体質量差を考慮しながら、質量スペクトルの各 m/z フラグメントイオンピークに組成式が割り当てられます。ほとんどのフラグメントイオンピークでは、MFG の質量許容範囲内に、割り当てられる組成式候補が複数存在します。その場合、イオンピークは、質量が最も近く、同位体比および同位体質量差が適切なフラグメント組成式に割り当てられます。このプロセスで使用される MFG の質量許容範囲は質量によって異なり、ライブラリ検索のデフォルト質量許容範囲の約 1/3 です。スペクトルを PCDL にインポートするときに、すべての m/z イオンピークが理論的な精密質量に合わせて補正されます。分子イオンの

部分分子式に割り当てることのできないフラグメントイオンピークは、化学的ノイズまたは不純物として排除されます。また、イオンカウントが 100 以下のフラグメントイオンピークは、常にノイズとして排除されます。

高品質で検索可能なスペクトルを得るために、アジレントの科学者によってデータセットが綿密に調査され、次のようなスペクトルは自動的に排除されます。

- 機器パラメータの設定が正しくない
- ベースピークのイオンカウントが 1,000 未満
- ノイズまたは不純物としてスペクトルから排除されたフラグメントイオンピークや、ライブラリ検索スコアを低下させるフラグメントイオンピーク
- 単一化合物標準物質の LC/MS/MS クロマトグラムに現れる複数の化合物や成分など、不純物、分解物、または変換生成物と考えられるピーク

すべての PCDL スペクトルの精査プロセスの最終ステップとして、高度な技能を持つ専門の質量分析スペシャリストにより、PCDL にインポートされた各スペクトルの適格性が MassHunter Qualitative Analysis データ解析で目視で判断されます。

データ品質を確保するために、各項目に対して複数レベルの確認が行われます。この精査プロセスでは、すべてのデータが同じように収集、精査されていることが確認され、採取条件が一貫していないデータによる誤差が排除されます。採取条件に一貫性のないデータは、オープンアクセスのライブラリやユーザーの貢献により成り立っているライブラリの大きな欠点の 1 つです。これが、このようなデータベースやライブラリによる検索結果の品質や結果の信頼性を低下させる原因となっています。

法医中毒学分野の多くの薬剤化合物については、トレーサビリティを提供するために、各スペクトルに使用された単一の化合物標準物質に関する内容が LC/MS/MS スペクトル化合物情報に含まれています。

Agilent 精密質量パーソナル化合物データベースライブラリの使用法

Agilent PCDL では、Agilent MassHunter ソフトウェアサイトとの直接的な連携により、分子式および参照ライブラリスpekトルが化合物情報とともに提供されます。これはアジレントならではの化合物同定機能であり、幅広い化合物データマイニングおよび同定ワークフローに活用できます。

代謝パスウェイ、化合物類、規制、国固有のリストなどを基準に、特定の解析プロセス用により小規模なターゲット PCDL を作成することも可能です。PCDL と化合物データマイニングおよび同定パイプラインから対象外の化合物を除外することにより、データ解析の効率と精度を大幅に高めることができます。

必要な化合物だけを収載したターゲット PCDL は次の手順で簡単に作成できます。

1. PCDL Manager で、関連する化合物類、規制、または国固有のタグを検索します (例については、図 7 を参照)。
2. 検索結果で目的の化合物をハイライト表示します。
3. 右クリックして「**Create Subset PCDL**」または「**Append to PCDL**」を選択します。

メタボロミクスワークフローでは、MassHunter Pathways to PCDL ソフトウェアを使用して、パスウェイにもとづくデータ解析用の PCDL を作成できます (図 8 を参照)。

新規成分が出現した場合も、MassHunter PCDL Manager のインポート機能により次の操作を簡単に実行できます。

- 新規化合物を化学的識別子とともにインポート
- 化合物またはそのサブセットの RT を追加または更新

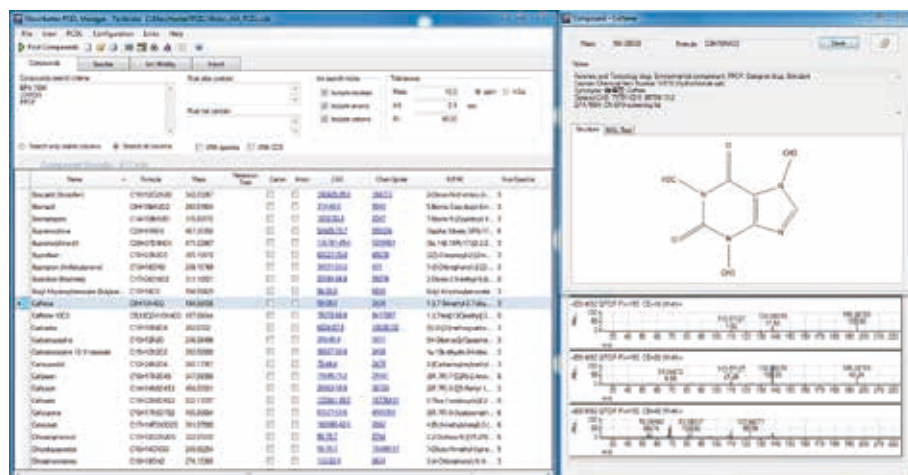


図 7. Agilent MassHunter PCDL Manager B.08.00: 米国および日本の政府規制タグ (EPA 1694 および JDWQS) とパーソナルケア製品 (PPCP) の化合物類タグを指定して化合物を検索した結果が表示されています。

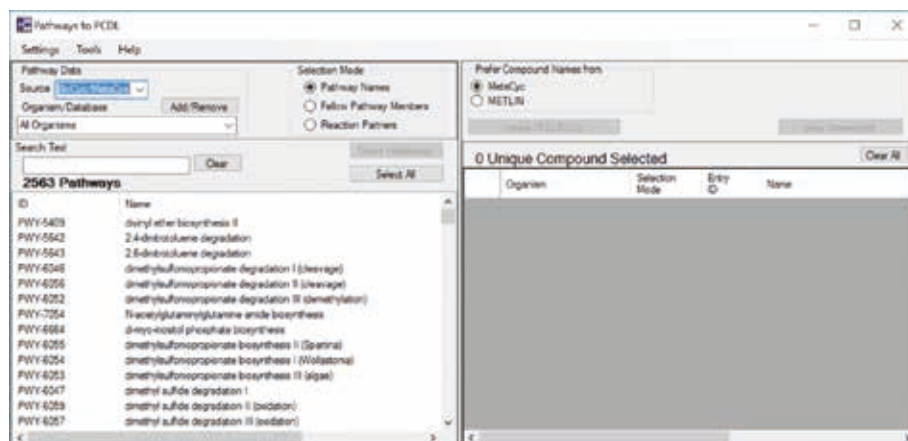


図 8. Agilent MassHunter Pathways to PCDL ソフトウェアによるパスウェイにもとづく解析

- CCS 値とその他の詳細なイオンモビリティ情報をインポート (IM-Q-TOF LC/MS)

新しいスペクトルは、MassHunter Qualitative Analysis から追加、他の PCDL からコピー、またはジャーナル参考文献から追加できます。

精査済みの LC/MS/MS スペクトルは、MassHunter Qualitative Analysis Workflows に搭載されている「**Send Spectra to PCDL**」自動精査ワークフローにより、アジレントの R&D 部門で用いられているものと同じインポート設定を使用して

簡単に追加できます (図 6 を参照)。

新しいスペクトルを追加するときと同じインポートパラメータを使用することで、スペクトルライブラリに含まれるすべての化合物に対して一貫したスコアランキングが適用され、化合物の同定結果の正確さがスペクトルライブラリ全体で維持されます。

MassHunter Qualitative Analysis では、複数の PCDL をデジチェーンに連結して連続的に照会することにより、より短いデータ解析時間で包括的な検索を実行できます。最も一般的な DB/ライブラリの連続検索ワークフローの 1 つが、小規模な化合物のターゲット

ト PCDL から始まり、次に存在が疑われる化合物の PCDL に移り、最後に大規模な汎用 PCDL ですべての化合物候補を検索するものです。このとき、すべての PCDL の同定結果を化合物ごとに報告することも、複数の化合物の同定を連続して実行し、データ処理時間を節約することもできます。

結論

化合物の同定を迅速に、高い信頼性で行うためには、質量分析データベースおよびライブラリの情報が高品質であることが不可欠です。すべての Agilent MassHunter PCDL は、アジレントのデータ収集および精査プロセスによって各化合物項目および精密質量スペクトルの品質が確保されています。

PCDL は、Agilent MassHunter ソフトウェアスイートの一部としてデータ解析、統計解析、および視覚化ソフトウェアパッケージと直接連携し、多様な化合物データマイニングおよび同定ワークフローにおいて、一貫したスコアランキングによる信頼性の高い化合物の同定を可能にします。

また、各 PCDL に付属する MassHunter PCDL Manager ソフトウェアパッケージを使用して、RT、化合物、およびスペクトルの他、IM-Q-TOF MS の場合はイオンモビリティ情報を追加または削除して、ワークフローに合わせて PCDL をカスタマイズすることもできます。

付録: Agilent 精密質量パーソナル化合物データベースライブラリ製品

食品安全性

LC/TOF および Q-TOF 用精密質量 LC/MS/MS PCDL	化合物数	精密質量 LC/MS/MS スペクトルのある化合物数	スペクトルの総数	RT のある化合物数
農業	1,750	> 825	> 2,700	0
動物用医薬品	> 2,150	> 1,525	> 5,200	104
マイコトキシン	> 450	> 300	> 1,350	0
GC/Q-TOF 用精密質量 GC/MS PCDL	化合物数	精密質量 GC/MS スペクトルのある化合物数	スペクトルの総数	RT のある化合物数
農業	> 850	> 850	> 850	> 840 (15 × 15.20 分および 40 分のメソッド) 最大 750 (5 × 15.20 分のメソッド)

環境分析

LC/TOF および LC/Q-TOF 用精密質量 LC/MS/MS PCDL	化合物数	精密質量 LC/MS/MS スペクトルのある化合物数	スペクトルの総数	RT のある化合物数
農業	1,750	> 825	> 2,700	0
水質汚染物質	> 1,400	> 1,050	最大 3,900	268
GC/Q-TOF 用精密質量 GC/MS PCDL	化合物数	精密質量 GC/MS スペクトルのある化合物数	スペクトルの総数	RT のある化合物数
農業	> 850	> 850	> 850	> 840 (15 × 15.20 分および 40 分のメソッド) 最大 750 (5 × 15.20 分のメソッド)

法中毒学

LC/TOF および LC/Q-TOF 用精密質量 LC/MS/MS PCDL	化合物数	精密質量 LC/MS/MS スペクトルのある化合物数	スペクトルの総数	RT のある化合物数
Broecker、Herre、Pragst による法中毒学化合物	> 9,200	> 3,900	> 13,500	0

本製品は法医学分野の実験用途での使用を想定しており、医薬品医療機器等法に基づく登録を行っておりません。

メタボロミクス

LC/TOF および LC/Q-TOF 用精密質量 LC/MS/MS PCDL	化合物数	精密質量 LC/MS/MS スペクトルのある化合物数	スペクトルの総数	RT のある化合物数
代謝物*	> 249,450	> 11,000	> 37,260	> 680
NIST 2014 MS/MS#	9,345	9,345	> 234,000	0

本製品は一般的な実験用途での使用を想定しており、医薬品医療機器等法に基づく登録は行っておりません。

NIST 2014 MS/MS PCDL は精密質量および単位質量 LC/MS/MS スペクトルで構成されています。

医薬品

LC/TOF および LC/Q-TOF 用精密質量 LC/MS/MS PCDL	化合物数	精密質量 LC/MS/MS スペクトルのある化合物数	スペクトルの総数	RT のある化合物数
抽出物と浸出物	> 1,000	> 350	> 1,300	129

ホームページ

www.agilent.com/chem/jp

カスタムコンタクトセンター

0120-477-111

email_japan@agilent.com

本製品は一般的な実験用途での使用を想定しており、医薬品医療機器等法に基づく登録を行っておりません。本文書に記載の情報、説明、製品仕様等は予告なしに変更されることがあります。

アジレント・テクノロジー株式会社
© Agilent Technologies, Inc. 2017
Printed in Japan, October 25, 2017
5991-8580JAJP

