OECD QSAR Toolbox v.4.1

のアプリケーション　マニュアル

(F1 help)

本翻訳は、OECD により作成されたものではなく、OECD の公式な翻訳ではありません。翻訳の内容と品質について OECD は責任を負いません。

原文と本翻訳に相違がある場合は、原文を優先してください。

本翻訳物は、OECD より公開された OECD/ECHA (2017), OECD QSAR Toolbox Application Manual（http://www.oecd.org/chemicalsafety/risk-assessment/TB4_Application_manual_F1.compressed.pdf）を NITE が仮訳したものです。
日本語訳
© 2018 National Institute of Technology and Evaluation (NITE) for this Japanese edition

目次

A. 導入 6

A.1. 目的 7
A.2. このガイダンス文書の使い方 8

B. ようこそ 9

B.1. 謝辞 10
B.2. OECD QSAR Toolbox とは 11
B.3. 略語集 12
B.4. アイコン 13

C. ユーザ インターフェース（システム表示と入力方法） 14
C.1. ステージツールバー 16
C.2. アクションツールバー 17
C.3. ステージオプションパネル 18
C.4. データ・マトリックス 23

C.4.1. エンドポイントツリーのエリア 24
C.4.2. 選択された化学物質のエリア 39
C.4.3. データのエリア 44

C.5. 情報と活用できるアクセサリ 47

D. ワークフロー 49
D.1. 入力 50
D.1.1. 単一の化学物質 55

D.1.1.1. CAS 番号による化学物質の入力 56
D.1.1.2. 化学名で化学物質を入力する 58
D.1.1.3. 化学構造式による化学物質の入力 59

D.1.1.3.1. 単一構造の入力 60
D.1.1.3.2. 混合物の入力 62

D.1.1.4. 化学組成による化学物質の入力 63
D.1.1.5. 最後に使用した化学物質の中から化学物質を入力する 68
D.1.1.6. データベースから化学物質を入力する 70
D.1.1.7. インベントリ・リストから化学物質を入力する 73
D.1.1.8. ファイルから化学物質を入力する 76

D.1.2. 化学物質のリスト 78

D.1.2.1. データベースを読み込む 79
D.1.2.2. インベントリ・リストを読み込む 81
D.1.2.3. カスタムファイルを読み込む 83

D.1.3. 検索のオプション機能 85

D.1.3.1. Substructure（サブ構造） (SMARTs) 86
D.1.3.2. Query tool（クエリツール） 90

D.1.4. 標的エンドポイントを決める 107

D.1.4.1. in vitro（イン・ビトロ；試験管内試験）での遺伝子突然変異 114
D.1.4.2. in vivo（イン・ビボ；生体内試験）での遺伝子突然変異 116
D.1.4.3. in vitro（イン・ビトロ；試験管内試験）試験での染色体異常 117
D.1.4.4. in vivo（イン・ビボ；生体内試験）試験での染色体異常 118
D.1.4.5. Skin sensitization（皮膚感作性） 119
D.1.4.6. 他のエンドポイント 120

D.2. プロファイリング 121

D.2.1. Profiling methods（プロファイリング手法）.....122

D.2.1.1. Predefined（既定のカテゴリ）.....123
D.2.1.2. General mechanistic（一般的な反応メカニズム）.....125
D.2.1.3. Endpoint specific（エンドポイント特化型）.....131
D.2.1.4. Empiric（経験則的）.....138
D.2.1.5. Toxicological（毒物学的）.....141
D.2.1.6. Custom（カスタム）.....142

D.2.2. Metabolism（代謝）.....143

D.2.2.1. Documented metabolism（実測記録をもつ代謝）.....144
D.2.2.2. Simulated metabolism（代謝のシミュレーション予測）.....146

D.2.3. 補助機能.....149
D.2.4. プロファイリング方法の説明.....152
D.2.5. プロファイリング結果の説明.....155
D.2.6. カスタマイズしたプロファイリングを作る.....160
D.2.7. プロファイルを削除する.....163

D.3. データ 164

D.3.1. データベース.....165
D.3.2. インベントリ.....181
D.3.3. データを集積する.....183
D.3.4. データをインポートする.....186

D.3.4.1. Vertical import（縦断的なインポート）：データベース.....187
D.3.4.2. Horizontal Import（横断的インポート）：データベース.....195
D.3.4.3. インベントリをインポートする.....199
D.3.4.4. サポート情報.....201
D.3.4.5. IUCLID6 からデータをインポートする.....202

D.3.5. データをエクスポートする.....206

D.3.5.1. csv ファイルにデータをエクスポートする.....207
D.3.5.2. IUCLID6 にデータをエクスポートする.....209

D.4. Category definition（カテゴリ決定） 215

D.4.1. Grouping methods（グルーピング法）.....216
D.4.2. カテゴリ構築の基本.....223
D.4.3. カテゴリを決定する.....224

D.4.3.1. 単一の化学物質のカテゴリ化.....229
D.4.3.2. 混合物のカテゴリ化.....230
D.4.3.3. 階層型のプロファイリングスキームのプロファイリング結果を使用したカテゴリ化.....231
D.4.3.4. 代謝によるカテゴリ化.....235
D.4.3.5. アラート性能.....238

D.4.4. サブカテゴリ化.....241
D.4.5. カテゴリを結合させる.....243

D.5. Data Gap Filling（データギャップ補完） 244

D.5.1. データギャップ補完方法.....245
D.5.2. データギャップ補完の手順.....246

D.5.3. データギャップ補完の画面.....248
D.5.4. 3 つのギャップ補完アプローチに共通の機能 250

D.5.4.1. 左サイドパネル 251

D.5.4.1.1. Descriptors（予測グラフの X 軸と Y 軸）.....252
D.5.4.1.2. Prediction（予測）.....253
D.5.4.1.3. Adequacy（妥当性）.....254
D.5.4.1.4. Cumulative frequency（累積頻度）.....255
D.5.4.1.5. Residuals（残差）.....256
D.5.4.1.6. Statistics（統計情報）.....257

D.5.4.2. データギャップ補完のメニュー 258

D.5.4.2.1. Select/filter data(選択/データのフィルタリング）.....259

D.5.4.2.1.1. Subcategorize（データギャップ補完の後で行う
サブカテゴリ化）.....260
D.5.4.2.1.2. Mark chemicals by WS（水溶解度で印を付ける）.....273
D.5.4.2.1.3. Mark chemicals by descriptor values（軸上の
値により印を付ける）.....275
D.5.4.2.1.4. Filter points by test conditions（試験条件による
フィルタリングポイント）.....276
D.5.4.2.1.5. Mark focused chemical（化学物質にフォーカス
して印をつける）.....277
D.5.4.2.1.6. Mark focused points（データ点にフォーカスして
印をつける）.....278
D.5.4.2.1.7. Remove marked data（マークしたデータを削除）....280
D.5.4.2.1.8. Clear existing marks（既存のマークを消去）.....281
D.5.4.2.1.9. Mark outliers（外れ値のマーク）.....282

D.5.4.2.2. Gap filling approach（ギャップ補完のアプローチ）.....283
D.5.4.2.3. Descriptors/Data（グラフ軸の記述子/データ）.....284

D.5.4.2.3.1. Change descriptor units（グラフ軸記述子の単位
変更）.....285
D.5.4.2.3.2. Edit descriptor options（記述子のオプション編集）..287
D.5.4.2.3.3. Select endpoint tree descriptor （エンドポイント
ツリー記述子の選択）.....288

D.5.4.2.4. Model/(Q)SAR（モデル/QSAR）.....290

D.5.4.2.4.1. Shown domain（ドメインの表示）.....291
D.5.4.2.4.2. Save model （モデルの保存）.....292
D.5.4.2.4.3. Save domain as category （カテゴリとして
ドメインを保存）.....294
D.5.4.2.4.4. Calculate Q2（Q2 を計算する）.....295

D.5.4.2.5. Calculation options（計算オプション）.....296

D.5.4.2.5.1. Data usage options（データの使用）.....297
D.5.4.2.5.2. Prediction approach options（予測手法の
オプション）.....300
D.5.4.2.5.3. Use Target Data for prediction（予測に
標的物質のデータを使用する）.....305
D.5.4.2.5.4. Set level of significance（有意水準を設定する）.....306

D.5.4.2.6. Visual options（グラフ表示のオプション）.....307

D.5.4.2.6.1. Set units in figure title label（図のタイトルに単位を設定する) 308
D.5.4.2.6.2. Axis range tool（軸の範囲ツール）.....309
D.5.4.2.6.3. Confidence range tool（信頼範囲ツール）.....310
D.5.4.2.6.4. Intercorrelation tool（相互相関ツール）.....311

D.5.4.2.7. Information（情報）.....312

D.5.4.2.7.1. Details of focused chemical（注目する化学物質の詳細）.....313
D.5.4.2.7.2. Details of target chemical（標的化学物質の詳細）...314
D.5.4.2.7.3. Details of focused points（注目するデータ点の詳細）315
D.5.4.2.7.4. Differences to target（標的物質との違い）.....316
D.5.4.2.7.5. Region selection tool（グラフ内の領域の選択ツール）318
D.5.4.2.7.6. Legend tool（凡例ツール）.....319

D.5.4.2.8. Miscellaneous（その他）.....320
D.5.4.3. Right click functionality（右クリック機能）.....321

D.5.5. （Q）SAR の特色 322

D.5.6. Automated and Standardized workflow（自動化・標準化ワークフロー）.....328

D.5.6.1. 自動化・標準化ワークフローのアルゴリズム.....329
D.5.6.2. Automated workflow（自動ワークフロー）.....336
D.5.6.3. Standardized workflow（標準ワークフロー）.....340

D.5.7. 一群の化学物質（混合物）の毒性推定の方法論.....346

D.5.7.1. 例：定量的な混合物の毒性予測 347

D.5.7.1.1. 混合物モデル：方法論 353

D.5.7.1.1.1. Independent MOA（独立した作用機序）.....354

D.5.7.1.1.1.1. 実例 355

D.5.7.1.1.2. Similar MOA（類似した作用機序）.....356

D.5.7.1.1.2.1. 実例 357
D.5.7.1.1.2.2. 例 358

D.6. Report（レポート）.....360

D.6.1. 予測レポート 361
D.6.2. データ・マトリックス　レポート 372
D.6.3. カテゴリ　レポート.....373
D.6.4. QMRF レポート.....375

E. Options（オプション）.....378

E.1. General （一般）.....379
E.2. Modules （モジュール）.....383
E.3. Profiling （プロファイリング）.....385
E.4. Data Gap Filling （データギャップ補完）.....386
E.5. Interface （インターフェイス）.....392
E.6. Data （データ）.....394

A 章. 導入

QSAR Toolbox End - user Manual/

現在のオンラインヘルプの目的は、OECD QSAR Toolbox を化学物質をカテゴリにグループ化するための使用にあたり必要な情報を提供することです。ツールボックスの主な目的は、ユーザが（Q）SAR 手法を使用して化学物質をカテゴリに分類したり、リード・アクロス、トレンドアナリシス、および (Q) SAR によってデータギャップを補完できるようにしたりすることです。化学カテゴリの概念に関する詳細な背景情報については、OECD 環境健康安全刊行物のテストおよび評価シリーズに掲載されている、化学物質のグルーピングに関するガイダンス文書を参考にしてください。

［OECD (2007);ENV/JM/MONO(2007)28: http://www.oecd.org/officialdocuments/displaydocument/?doclanguage=en&cote=env/jm/mono(2007)28].

QSAR Toolbox の専用インターネットサイト[http://www.qsartoolbox.org]、OECD（Q）SAR プロジェクトのインターネットサイト[http://www.oecd.org/env/existingchemicals/qsar]、および QSAR Toolbox の開発者のインターネットサイト[http://toolbox.oasis-lmc.org/]で、他にもガイダンスとトレーニング資料を入手することができます。ユーザは、これらのインターネットサイトを定期的にご確認下さい。

QSAR Toolbox は、経済協力開発機構（OECD）が欧州化学品庁（ECHA）と協力して行うプロジェクトであって、Laboratory of Mathematical Chemistry によって開発されています。

新しい Toolbox 4 は、Toolbox 3.4 ソースコードを書き換えました。このシステムは全く新しいものであり、ボトムアップから設計され、新しい基盤技術とデータベース管理サブシステムを使用して書き直されています。これは、従来と全く異なる展開を体験できるということです。

A.1. 目的

この QSAR Toolbox エンドユーザマニュアルは、QSAR Toolbox ソフトウェアの使用を支援するものです。ヘルプデスクに確認したり研修に参加したりせずとも、すべてのユーザが QSAR Toolbox の使用方法を理解できるようになることを目的としています。また、QSAR Toolbox のすべての機能、オプション、およびパスの完全な参照マニュアルとしても使用できます。

A.2. このガイダンス文書の使い方

QSAR Toolbox End - ユーザ マニュアル/

マニュアルは、主に次のパートで構成されています。

A. 導入 - この章では、このマニュアルの目的、使用方法のガイダンス、高度なサポートオプションについてのヒントなどを紹介します。

B. ようこそ- QSAR Toolbox の背景と歴史の概要、その目的、スキルとツールを寄付する化学的なチームと関係者のための情報の概要を提供します。

C. ユーザインターフェース - この章は、ソフトウェアのインターフェース、ステージオプションパネル、ソフトウェアのデータ・マトリックスをユーザに提示することを目的としています。

D. ワークフロー - この章では、特定の化学物質のエンドポイントを予測する典型的なワークフローの概要を提供します。この章には、Toolbox のワークフローの主要な 6 つのセクション（Input, Profiling, Data, Category definition, Data Gap filling, and Reporting）の詳細な説明が含まれています。

E. オプション - この章では、Toolbox の一般的なオプションに関する情報を提供します。この章では、プログラムの以下のセクションの設定を変更する方法について説明します。 Main, Modules, Units, Gap filling, Reports です。

B 章　ようこそ

ようこそ

データの欠如は、有害性およびリスク評価の大きな問題のひとつとして認識されることがよくあります。利用可能なデータは通常、断片的で異なる形式で保存され、時には一貫性にかけることが多いものです。試験のレポートや出版物のクオリティは非常に重要な考慮事項であるため、信頼性の低い研究を除外し、最も信頼性が高いと思われるものにさらなるリソースを集中すべきです。データとそのクオリティに関連する情報は、後で再利用できるようにデータベース内のデータを効率よく構成するための基礎的要素となります。そのような最適なデータ構造は、特定の情報を検索し、データギャップを見つけて補完し、有害性およびリスク評価を実行するための効果的なツールを構築するための大前提となります。リソースと動物愛護の理由から、科学的な根拠付けが可能であるならば、動物実験の回数を減らすことが重要です。

ケモメトリクス手法は、環境と健康安全性評価のために動物実験を削減できる可能性を持っています。カテゴリアプローチ法と（Q）SAR は、資源を節約し、有害性およびリスク評価を加速するために使用できる代替手法です。あらゆる種類のエンドポイントに関する多数の QSAR 式を文献で見つけることができます。これらの研究のほとんどすべてが何かしらの観点から関連性があり、興味を持たれています。反応メカニズムについて間接的な情報を提供する QSAR もあれば、提示されているモデルがリスク評価プロセスに適用可能であるため関心を持たれている QSAR もあります。環境運命と健康影響のための既存の推定モデルを使用し批判的評価をするには、（Q）SAR のライブラリとその特性に関する関連情報が必要です。この点で、コンピュータシステムにおいてしっかり構成されたレビュー済みデータのデータベース、（Q）SAR ライブラリ、検索エンジンをうまく構築することは、証拠の重み付けと化学物質の規制にとって」実用的価値があります。

B.1. 謝辞

ようこそ

OECD QSAR Toolbox の開発は大規模な共同作業であり、多くの科学チームと関係者がツールボックスに統合するための技術とツールを寄進しています。

[http//www.oecd.org/env/chemicalsafetyandbiosafety/assessmentofchemicals/donorstotheqsartoolbox.htm]を参照のこと。

B.2. OECD QSAR Toolbox とは

ようこそ

OECD QSAR Toolbox は、動物実験の使用を減らし、試験のコストを削減し、より多くの化学物質についてヒトの健康と環境への影響評価ができるよう設計されたソフトウェアです。OECD QSAR Toolbox は、有害性およびリスク評価に必要な実験データのギャップを補完するためのカテゴリアプローチを適用する科学的計算方法と情報技術を提供します。このシステムを利用すると、有害性とリスクの評価者は以下のことを実施できます。

- 既定のカテゴリを使用・改良したり、新しいカテゴリを作成したりする。

- ユーザが選択した特性に基づいて類似化学物質（またはカテゴリ）を特定する。消失率、安定した代謝物の生成、高反応性中間体の生成、不活性化経路など、代謝を考慮し化学物質をカテゴリ化する。

- 信頼性に関する情報（実験誤差、使用した分析手法や計算方法、再現性など）とともに、ローカルおよびリモート（Web 上の）データベースから入手可能なすべての実測データまたは予測データを抽出する。

- リードアクロス、トレンドアナリシス、(Q) SAR モデルなどのケモメトリクス手法を利用して、カテゴリ内で欠落している情報のギャップを補完する。

QSAR の予測結果には、その反応機序的背景、トレーニングセットに入っている化学物質、統計値、適用範囲、妥当性に関する情報が含まれます。

OECD QSAR Toolbox は、インストールされている計算ツール、データベースマネージャ、（Q）SAR ライブラリ、カテゴリ化モデルなどのすべての構成要素（モジュール）間の情報フローをナビゲートする、拡張可能なアプリケーションです。

B.3. 略語集

ようこそ

AP	アラート（警告）表示
AW	自動ワークフロー
BCF	生物濃縮係数
EC50	半数影響濃度
IGC50	半数成長阻害濃度（統計学的にまたはグラフ上で推定された試験物質の特定の濃度であって、50%に成長阻害を引き起こすことが予想されるときの濃度）
InChi	IUPAC 国際化学物質識別法
KOW	オクタノール／水の分配係数
LC50	半数致死濃度（試験した生物群の 50%が死に至ると予想される、統計的に推定された濃度）死亡は、死亡率、毒性および集団影響を受けた群によって決定される。
LOEC	試験した生体に統計的に有意な悪影響を及ぼす最小影響濃度 (LOEC) または最小影響量 (LOEL)
MATC	最大許容毒性濃度; NOEC と LOEC の濃度の間の幾何平均である仮想の閾値濃度
MW	分子量
NOEC	無影響濃度(NOEC) /無影響量 (NOEL);報告された統計的検定試験によりコントロール群の反応と有意には異ならない効果を生じる濃度または用量
QSAR	定量的構造活性相関
SAR	構造活性相関
SMILES	SMILES 記法；簡略化した分子文字列入力システム
SW	水溶解度
SW	標準ワークフロー
UVCBs	組成が不明または不定の化学物質、複雑な反応生成物または生物材料

B.4. アイコン

ようこそ

Document ツリーの各ノードに新しいアイコンが実装されています。ユーザが実行したアクティビティを示します。Documents パネルに表示されるアイコンの意味は次のとおりです。

 - 現在のドキュメントは 2D モードです（立体構造を加味しません）

 - 現在のドキュメントは 2.5D モード（立体構造を加味します）です

 - CAS 番号で化学物質を入力しました

 - 化学物質名で化学物質を入力しました

 - 化学構造を描画して化学物質を入力しました

 - 組成を持つ（持たない）化学物質を入力しました

 - リストから化学物質を入力しました

 - データベースを読み込みました

 - インベントリを読み込みました

 - リストを読み込みました

 - SMARTS で化学物質を検索しました

 - クエリツールで化学物質を検索しました

 - カテゴリを決定しました

 - データギャップを補完しました

 - プロファイリングスキームを使用してサブカテゴリを作成しました

 - 代謝とプロファイリングスキームを組み合わせて使用し、サブカテゴリを作成しました

 - ギャップ補完方法を変更しました

 - Target multiplication（標的物質が代謝などで数種類の物質に増え多元化すること）を適用しました

 - Multiplication（代謝などで多元化すること）の生成物を示しています

 - Combine（結合させる）ボタンを使用しました

 - Tautomerism（互変異性）を適用しました

 - Metabolism/transformations（代謝/転換）のいずれかを適用しました

 - 手動で化学物質が追加されました

C 章. ユーザインターフェース（システム表示と入力方法）

QSAR Toolbox End - ユーザ マニュアル

ツールボックスのインターフェースは、所与の化学物質（標的化学物質といいます）のエンドポイントを予測するために、典型的なワークフローに従うように設計されています。これは、アプリケーションのウィンドウの最上部にあるツールバー(1)上のワークフロー(Input、Profiling、Data、Category definition、 Data Gap Filling、 Report）の主な 6 つのステージのことです（図 1）。ステージツールバーの下にはアクションツールバー（2）があります。このアクションツールバーは、現在のステージに関連する非常に重要なアクションを提供します。メインウィンドウの左側には、Documents パネル（3）とステージオプションパネル（4）があります。Documents パネル（3）はツールバー（1）から独立しています。このパネルはツールバーの各セクションを通して存在しており、ユーザは Document ツリーを用いて、サブカテゴリ化のステップを進んだり戻ったりすることが可能になります。この機能は、Toolbox Ver.3.4 以前のデータギャップ補完ウィンドウの Return to matrix ボタンにとってかわるものです。ステージオプションパネル (4) は、現在のステージの特定のコンテンツと、このコンテンツに関連するアクションを提供します。メインフォームの一番広い部分は**データ・マトリックス**（5）が占めています。これはすべてのステージで利用でき、システムに読み込まれた化学物質の実験値と予測値の両方を照会したデータを表示します。

右上隅のボタン（6）は次のとおりです：現在のアクティビティを停止するボタン、 オプションボタン、進む/戻るボタン、情報ボタン（訳注：E 章を参照）。最後の情報ボタンには、ユーザがシステムをダウンロードしたり、トレーニング資料を探したり、質問をするためのリンクが含まれています（図 2）。管理、請負業者、資金調達、寄付者、QSAR Toolbox 開発の歴史に関する情報もあります。情報ウィンドウの最後のタブで新しいバージョンを確認することもできます。

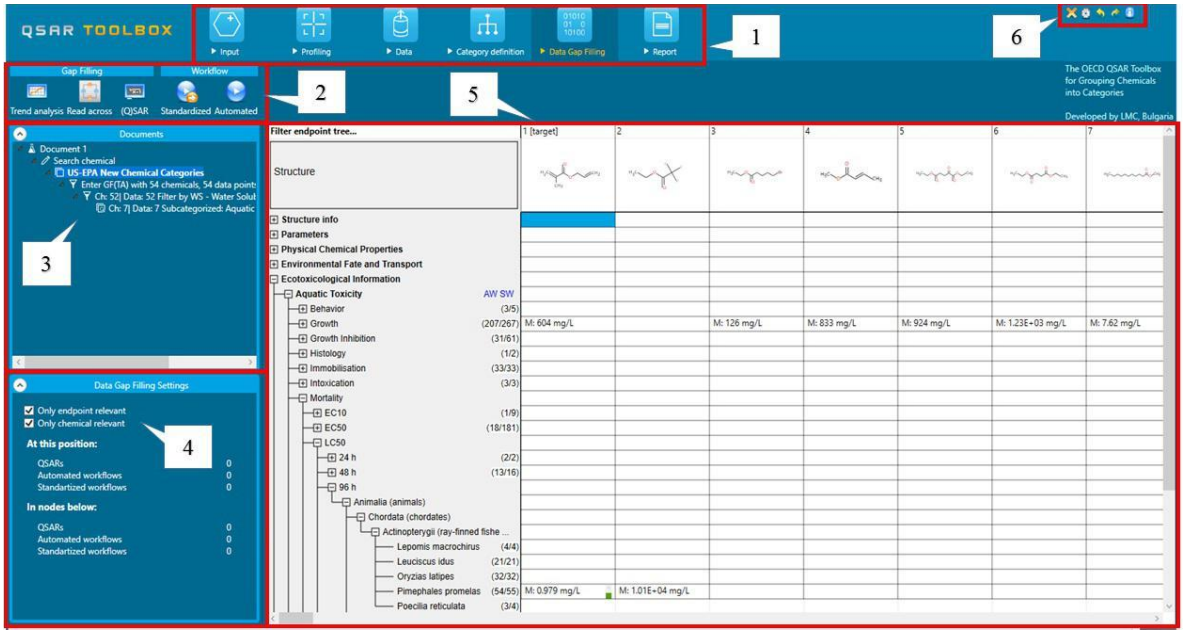


図 1

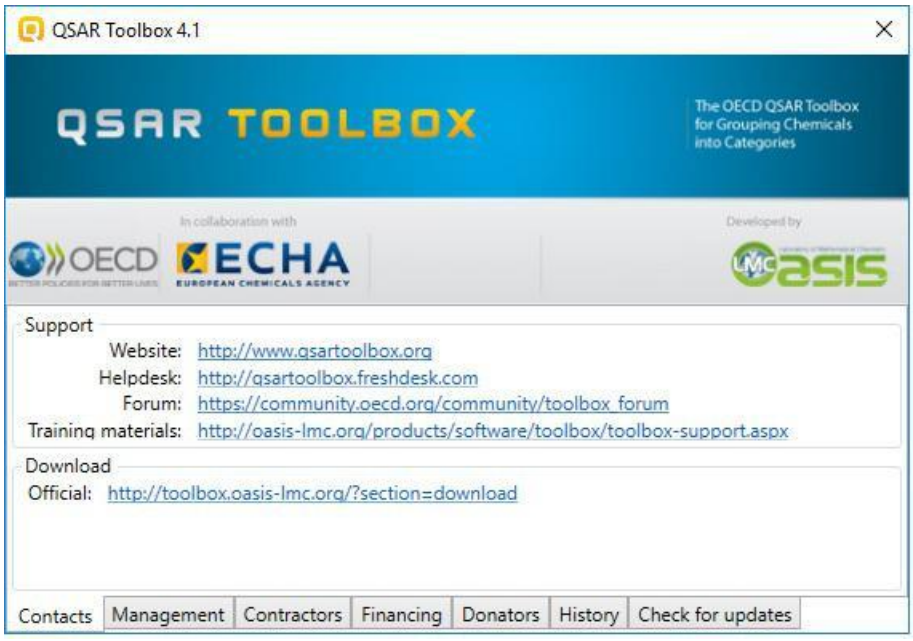

QSAR Toolbox 4.1
QSAR TOOLBOX
The OECD QSAR Toolbox for Grouping Chemicals into Categories
In collaboration with
OECD
ECHA
EUROPEAN CHEMICALS AGENCY
Developed by
Support
Website: http://www.qsartoolbox.org
Helpdesk: http://qsartoolbox.freshdesk.com
Forum: https://community.oecd.org/community/toolbox_forum
Training materials: http://oasis-lmc.org/products/software/toolbox/toolbox-support.aspx
Download
Official: http://toolbox.oasis-lmc.org/?section=download
Contacts
Management
Contractors
Financing
Donators
History
Check for updates

図 2

C.1. ステージツールバー

ステージツールバーは、ツールボックスのインターフェース上に常に表示されます。これにより、プログラムのワークフローのメインステージ間を簡単にナビゲートできます。各ステージがツールバーボタンで表示され、選択されたステージに関連するインターフェースを呼び出します。以下に例を示します（図 1-2）：

「入力」ステージのツールバーボタン（1）とインターフェース（2） (図 1) 。

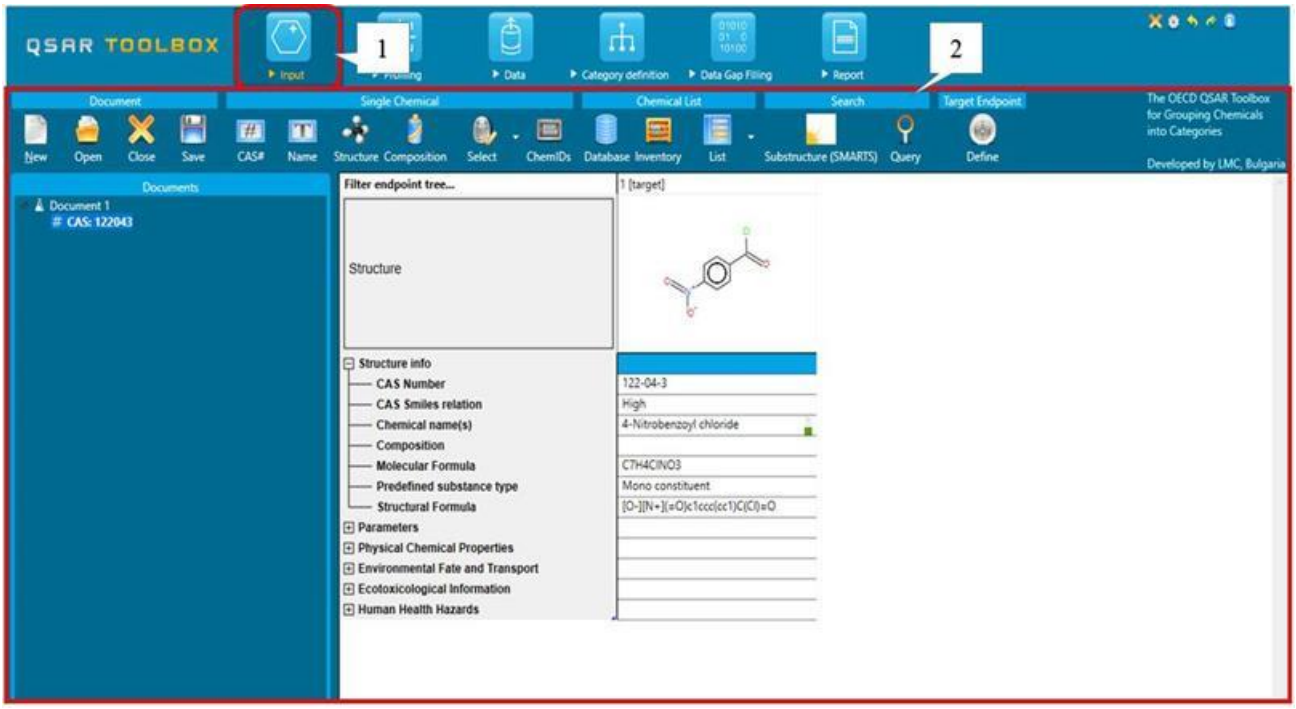


図 1

「データギャップ補完」ステージのツールバーボタン（1）とインターフェース（2）。(図 2)

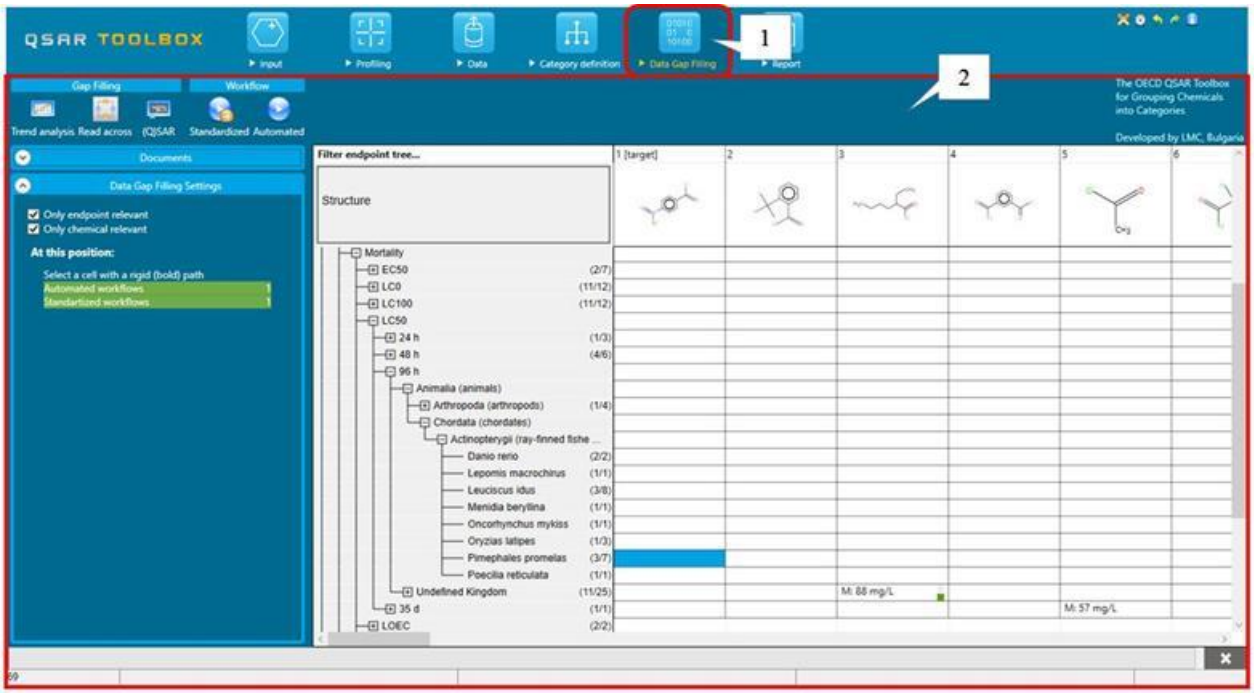


図 2

C.2. アクションツールバー

アクション（作業）ツールバーは、プログラムのワークフローのステージの基本的なアクションを提供します。各ステージに特定のアクションがあるため、ツールバーの内容はステージによって異なります。ユーザの便宜のために、アクションをグループに分けることもあります。以下に例を示します（図 1-2）：

「Input（入力）」ステージの基本的なアクション。（図 1）

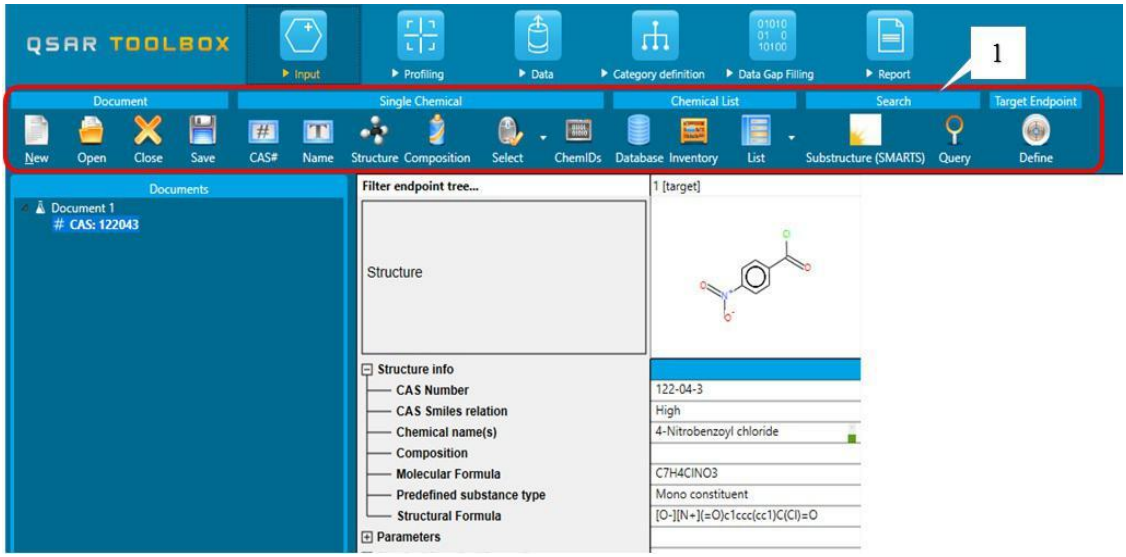


図 1

「Data Gap Filling（データギャップ補完）」ステージの基本的なアクション。(図 2)

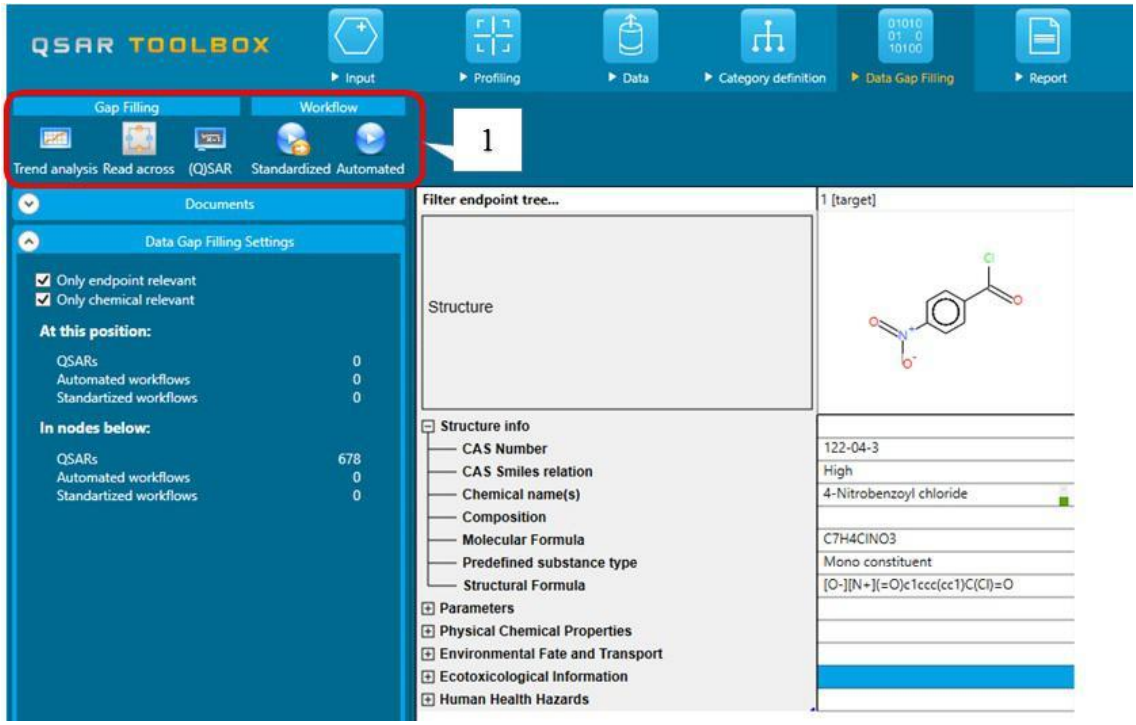


図 2

C.3. ステージオプションパネル

ユーザ インターフェース

ステージオプションパネルは、現在のステージの特定のコンテンツと、このコンテンツに関連するアクションを提供します。各ステージにはそれぞれの機能があり、そのためステージオプションパネルの内容が異なります。以下に例示します:

I. Input（入力）： 入力ステージのステージオプションパネルには、作業中の文書、文書の内容のリストが表示されますまた、標的物質構造の関連物質をいくつか抽出する 3 つのアプローチを提供します。互変異性による構造の多元化、代謝による多元化、そして標的物質が 2 つ以上の成分による多元化です。代謝は、5 つの実測値シミュレータおよび 10 個の代謝シミュレーション値によるシミュレータによって適用することができます。:

- 実測値 :
  - Observed Mammalian metabolism（哺乳類の代謝の実測）
  - Observed Microbial metabolism（微生物代謝の実測）
  - Observed Rat In vivo metabolism（ラット in vivo（生体内）代謝）
  - Observed rat liver metabolism with quantitative data（定量データを含むラットの肝臓代謝の実測）
  - Observed Rat Liver S9 metabolism（ラット肝 S9 代謝の実測）

- 予測値（シミュレーション値）:
  - Autoxidation simulator（自動酸化シミュレータ）
  - Autoxidation simulator (alkaline medium)　（自動酸化シミュレータ、アルカリ媒体）
  - Dissociation simulator（解離シミュレータ）
  - Hydrolysis (acidic)｛加水分解（酸性）｝
  - Hydrolysis (basic)　｛加水分解（塩基性）｝
  - Hydrolysis (neutral)　｛加水分解（中性）｝
  - In vivo Rat metabolism simulator（In vivo ラット代謝シミュレータ）
  - Microbial metabolism simulator（微生物代謝シミュレータ）
  - Rat liver S9 metabolism simulator（ラット肝臓 S9 代謝シミュレータ）
  - Skin metabolism simulator（皮膚代謝シミュレータ）

1. Multiplication by Metabolism（代謝による多元化）

読み込まれた標的物質の化学構造の多元化を行うには、ステージオプションパネル（1）で標的物質を右クリックし、ポップアップメニュー（2）から Multiplication を選択してから、Metabolism / Transformations（3）をクリックし、 最後に Skin metabolism simulator（4）などのシミュレータを選択します。（図 1）

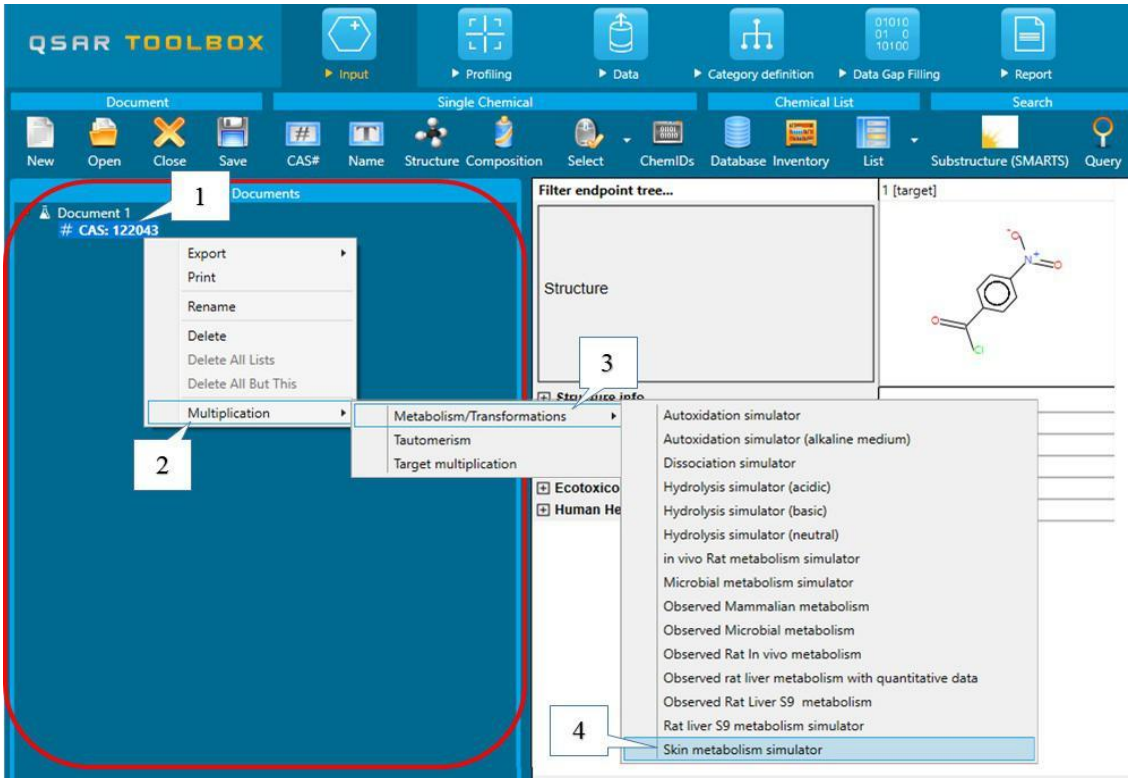


図 1

実測値による代謝シミュレータ 5 つから 1 つを選んで同じ手順を踏むと、構造を Multiplication できます。

2. Tautomerism by Multiplication （多元化による互変異性体化）

構造体（1）を右クリックし、ポップアップメニュー（2）から Multiplication を選択してから、 Tautomerism（3）を押します。(図 2)

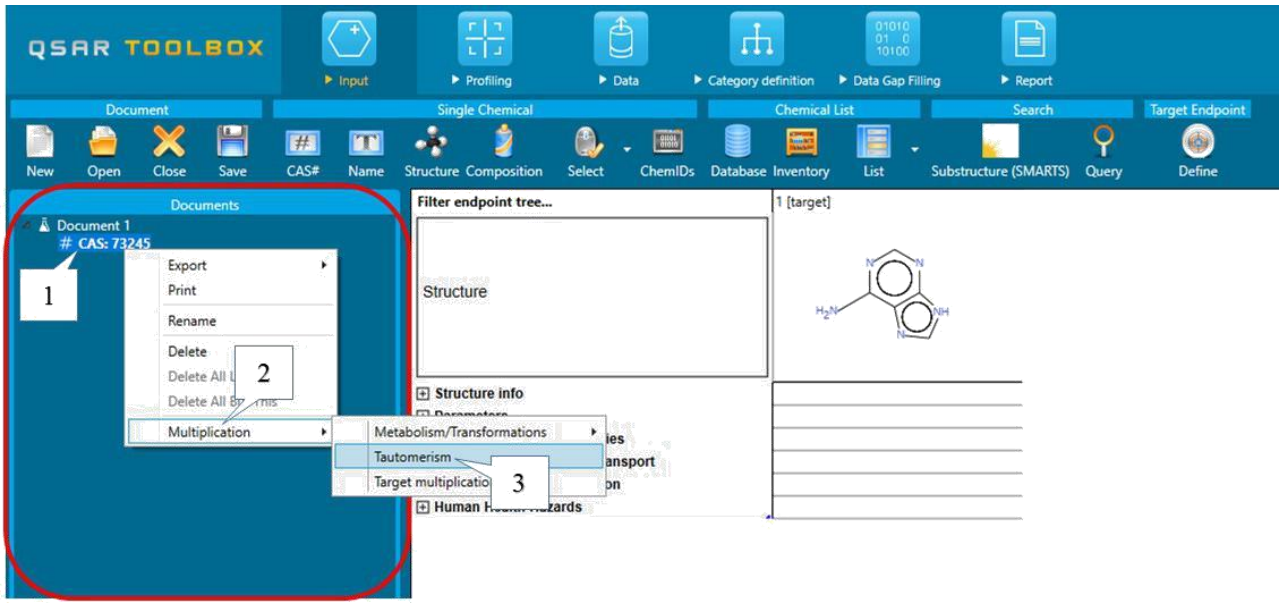


図 2

親物質リスト（1）は、互変異性体セット全体から構成され、標的物質と各セットの成分の子物質リスト（2）が作成されます。(図 3)

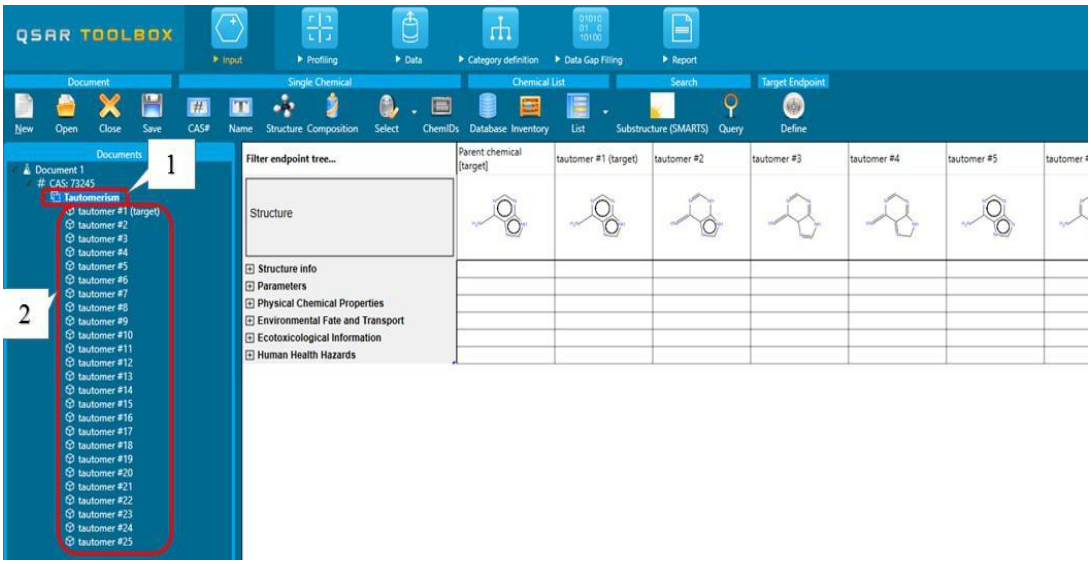


図 3

3. Multiplication by Target multiplication（標的物質多元化による多元化）

物質によっては、複数の成分（混合物、添加物の存在、不純物などを含む）からなるものもあります。こういった種類の物質で個々の成分をすべて見るには、ステージオプションパネル（1）で標的物質を右クリックし、ポップアップメニュー（2）から Multiplication を選択し、Target multiplication (3)を選択します。(図 4) 物質組成中の成分の子物質リスト（4）が作成されます。

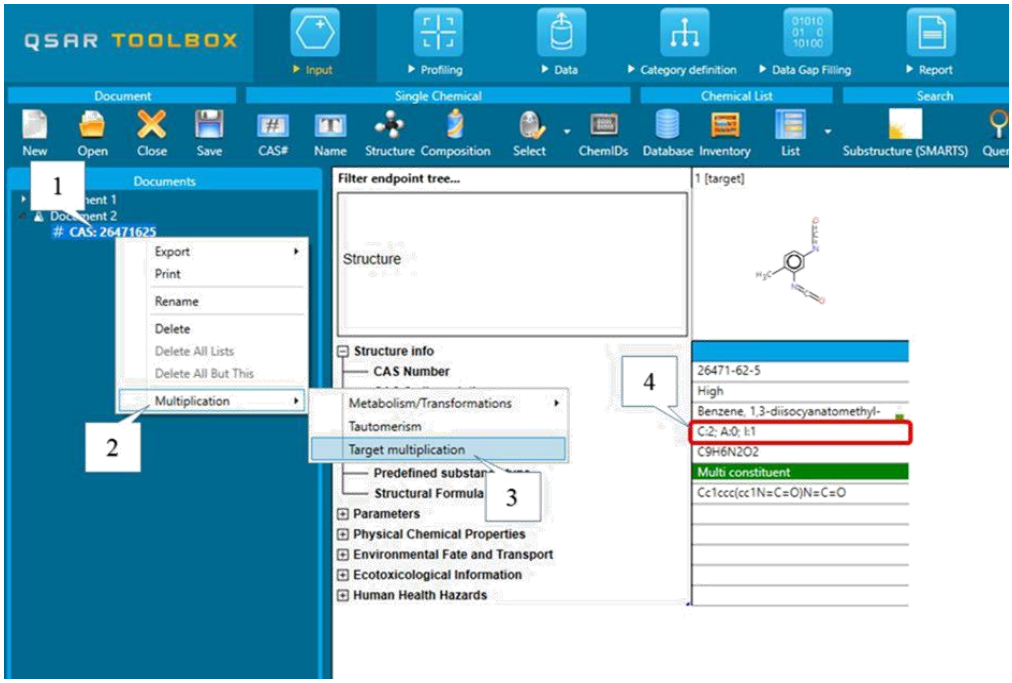


図 4

II. プロファイリング

新しい機能として、エンドポイントに応じたプロファイラの関連性に関する情報を提供します。エンドポイントが**データ・マトリックス**（1）で選択されると、Options（2）内で、**データ・マトリックス**で選択されたエンドポイント（またはエンドポイントが事前定義されている場合は標的エンドポイント）ごとに Profiling モジュール内のプロファイラを Color（3）や Group（4）で色付け・グルーピングできます（図 5）。凡例（5）ボタンをクリックすると、どの色が何を意味しているか見ることができます。サブカテゴライゼーションの過程で類似物質を検索したり、データギャップを補完したりするときにも、同じ機能を category definition（カテゴリ決定）に使用できます。

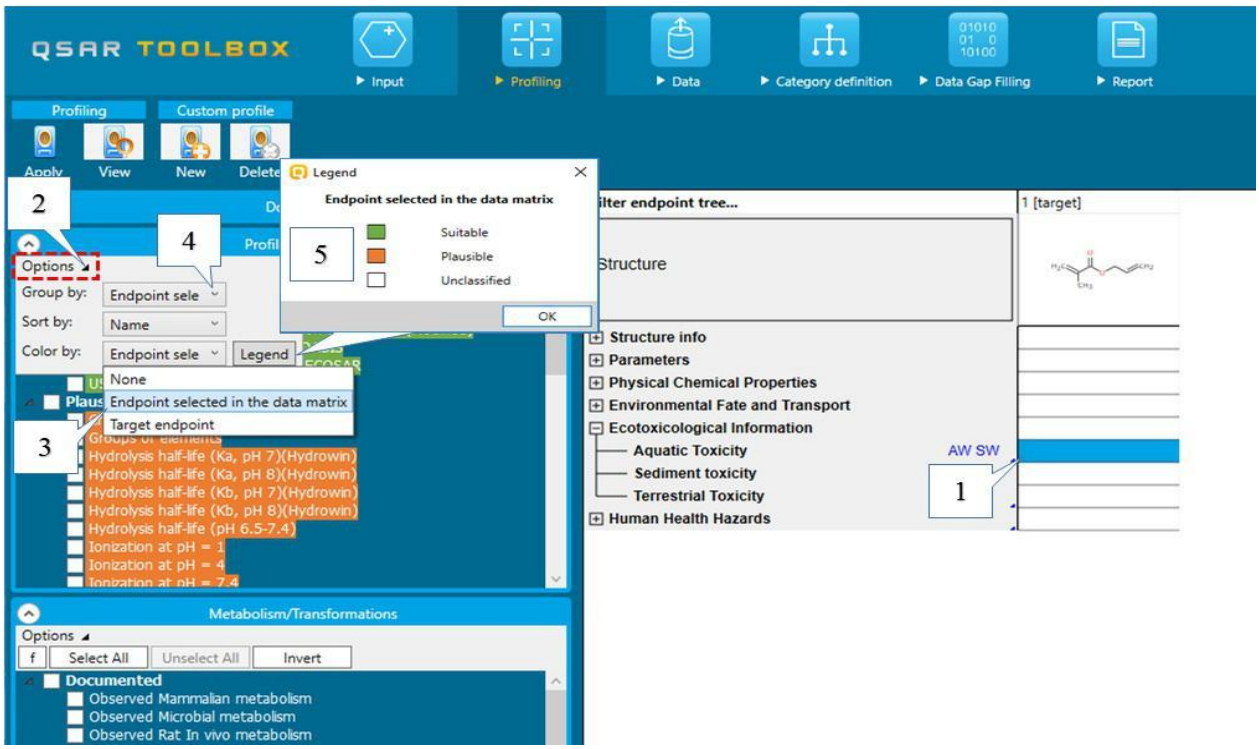


図 5

III. データ

Toolbox v.4.0（およびそれ以降のバージョン）では、オプション（3）において、**データ・マトリックス**で選択されたエンドポイント（またはエンドポイントが事前定義されている場合は標的エンドポイント）（2）の色付けを選択することで、どのデータベースに選択したエンドポイント(1)に関連するデータが含まれているかを確認できるようになっています。データが存在するデータベースは、緑色で強調表示されています（4）。（図 6）

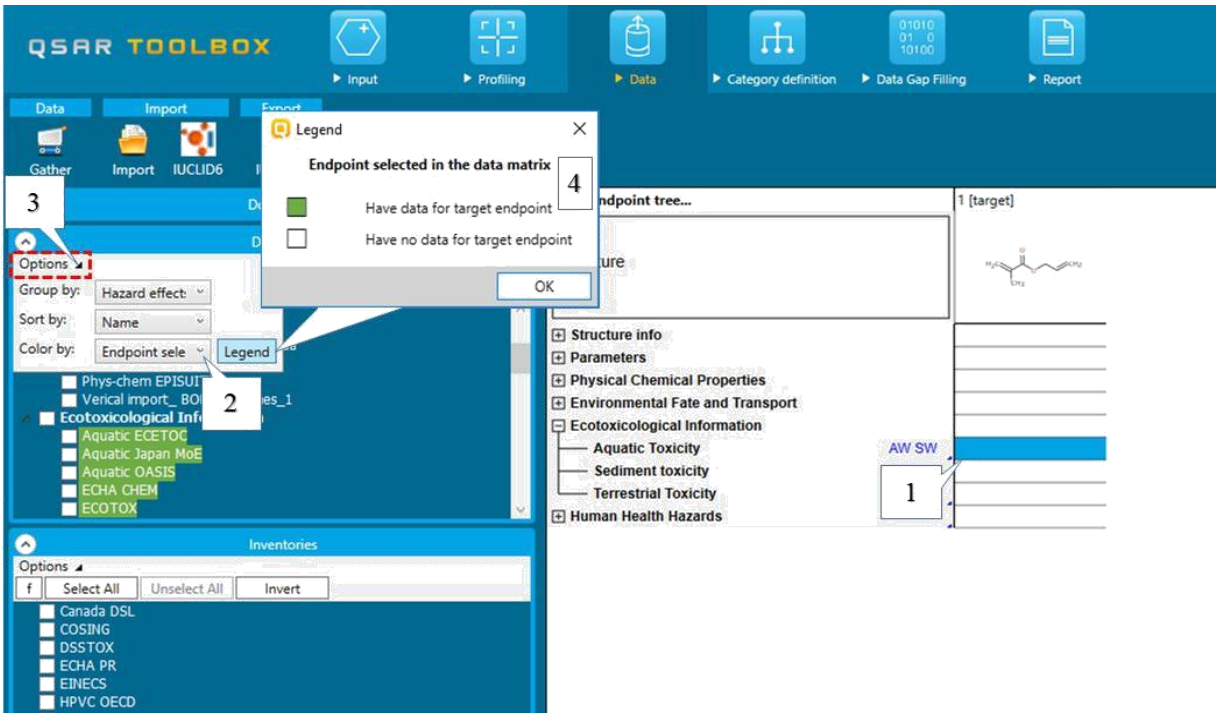

QSAR TOOLBOX
Input
Profiling
Data
Category definition
Data Gap Filling
Report
Data
Import
Gather
Import
IUCLID6
Legend
Endpoint selected in the data matrix
Have data for target endpoint
Have no data for target endpoint
OK
4
3
Options
Group by:
Hazard effect
Sort by:
Name
Color by:
Endpoint sele
Legend
2
Phys-chem EPISUI
Verical import_ BO
Ecotoxicological Inf
Aquatic ECETOC
Aquatic Japan MoE
Aquatic OASIS
ECHA CHEM
ECOTOX
Inventories
Options
Select All
Unselect All
Invert
Canada DSL
COSING
DSSTOX
ECHA PR
EINECS
HPVC OECD
ndpoint tree...
1 [target]
Structure info
Parameters
Physical Chemical Properties
Environmental Fate and Transport
Ecotoxicological Information
Aquatic Toxicity
Sediment toxicity
Terrestrial Toxicity
Human Health Hazards
AW SW
1

図 6

C.4. データ・マトリックス

ユーザ インターフェース

下のスナップショットは**データ・マトリックス**ィンドウ（1）を表わしています。 (図 1).

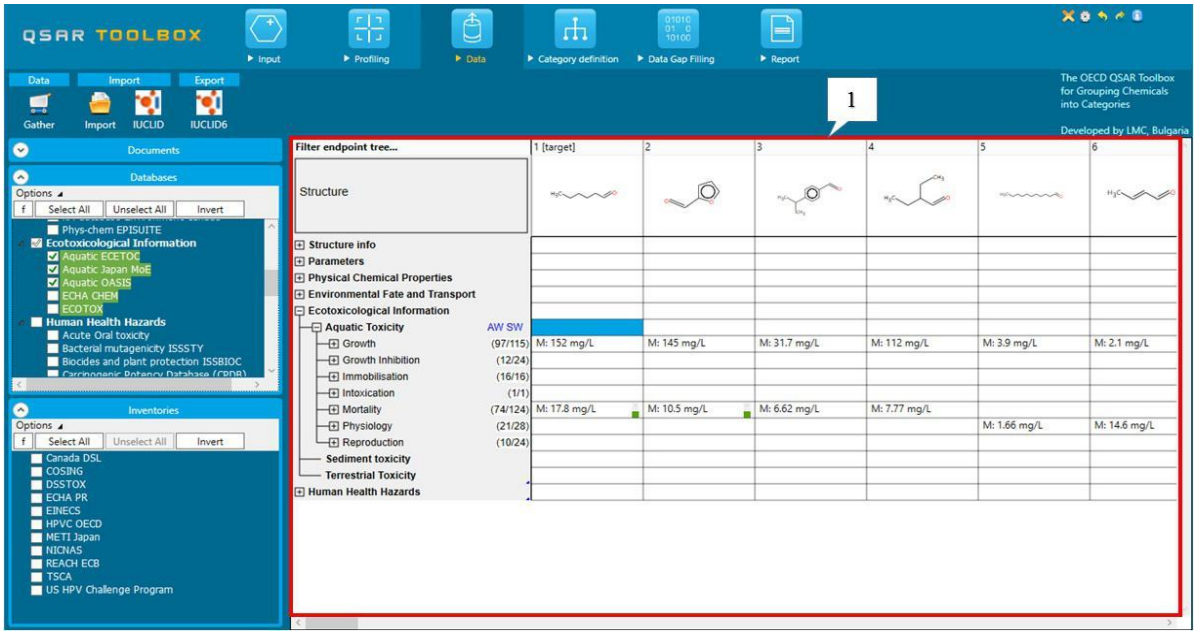


図 1

データ・マトリックスィンドウには主要な部分が 3 つあります（図 2）。

エンドポイントツリーのエリア（1）
選択された化学物質のエリア（2）
データのエリア（実測値、予測値）（3）

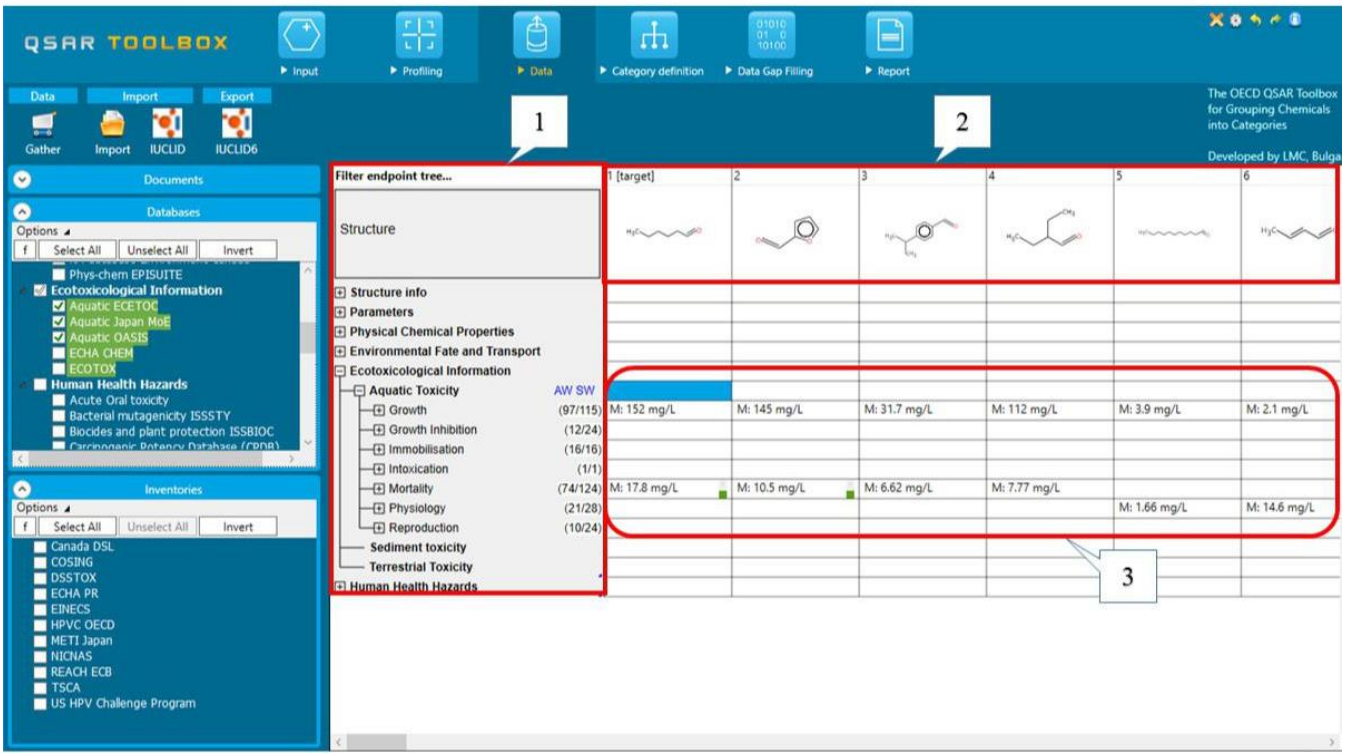


図 2

C.4.1. エンドポイントツリーのエリア

エンドポイントツリーのエリアの内容:

- エンドポイントツリーのノード（＋印で表されるツリー構造の分岐点）
- 構造
- ツリー階層の設定
- エンドポイントツリーのフィルタリング
- ヒント

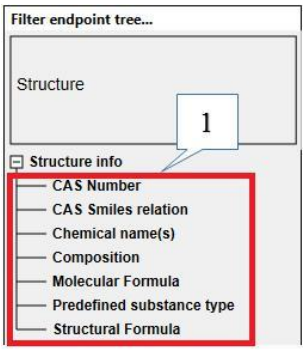
4.1.1. エンドポイントツリーのノード:

エンドポイントツリーには 6 つの汎用的ノードがあります:

- Structure info（構造情報）
- Parameters（パラメーター）
- Physical Chemical Properties（物理化学的性質）
- Environmental Fate and Transport（環境運命と移送）
- Ecotoxicological Information（生態毒性情報）
- Human Health Hazard（ヒト健康有害性）

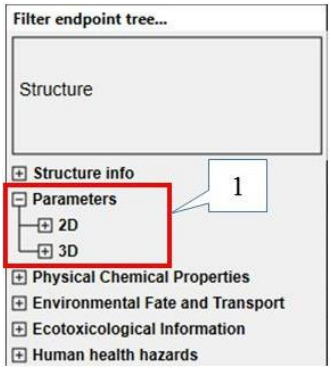
- Structure info（構造情報）には CAS number（CAS 番号）、 CAS Smiles relation（CAS 番号と SMILES 表記法との関係）; Chemical Name(s)（化学物質名）、Composition（組成）、Molecular formula（分子式）、Predefined substance type（既定の物質のタイプ）、Structural formula（構造式）など、選択された化学物質の物質情報を表示するサブノードが含まれています。（図 1）：

図 1



- Parameters は、2D と 3D（1）の 2 つのサブノードに分かれています。 (図 2).

図 2



対象のパラメータ（1）上でマウスを右クリックし、オプションの 1 つを選択すると、そのパラメータを計算することができます。オプションには、to calculate all parameters（すべてのパラメータを計算）、to calculate only 2D (3D) parameters（2D または 3D パラメータのみを計算） 、 to calculate only the current parameter（現在のパラメータのみを計算）があります（2） 。ここで選択したオプションは、**データ・マトリックス**に読み込まれたすべての化学物質のパラメータの計算に適用されます。（図 3）

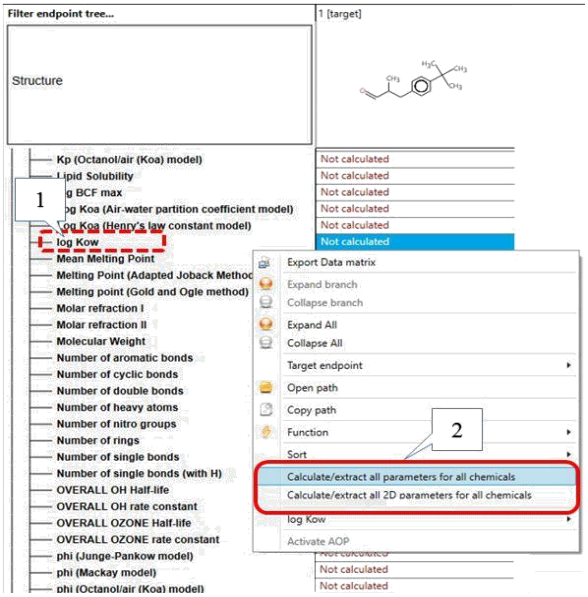


図 3

ある特定の化学物質またはすべての化学物質のパラメータの計算をすることができます。必要なパラメータの計算をしたり、化学物質のセル上（1）にマウスを持ってきて、右マウスクリックで現れるポップアップメニューからオプションの一つを選んでください（2）。オプションには Calculate/extract …（計算／抽出…） ... や、 Calculate/Extract all 2D parameters（すべての 2D パラメータを計算/抽出）、などがあります。 (図 4)

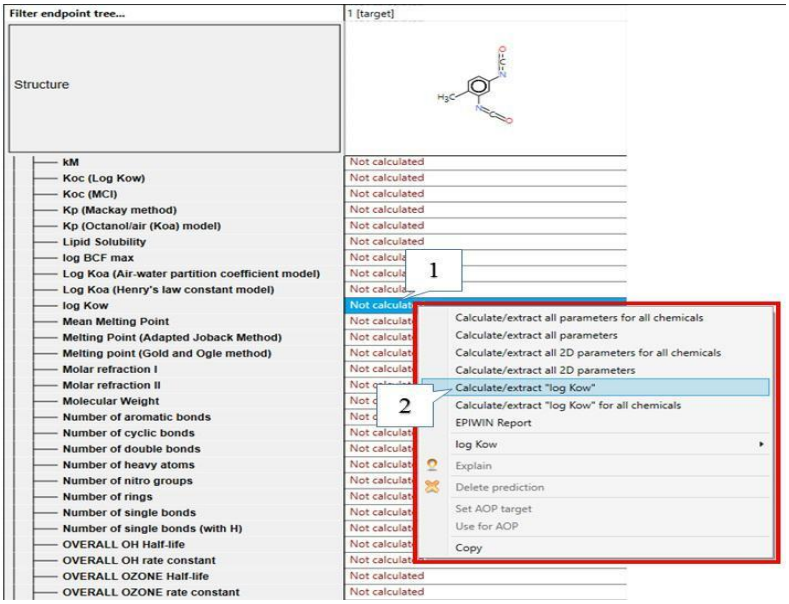


図 4

Toolbox データベースで利用できる実測データは、他の 4 つの大きいノードとそのサブノードに振り分けされています。これらの 4 つのノードは、割り当てられた実測データのタイプに応じて 4 つの基本セクションに分けられたものです。
例えば、融点や分配係数（図 5）の結果は、Physical Chemical Properties（物理化学的性状）ノード（1）のサブノードである Melting/freezing point（融点/凝固点）や Partition Coefficient（分配係数）に振り分けされています。

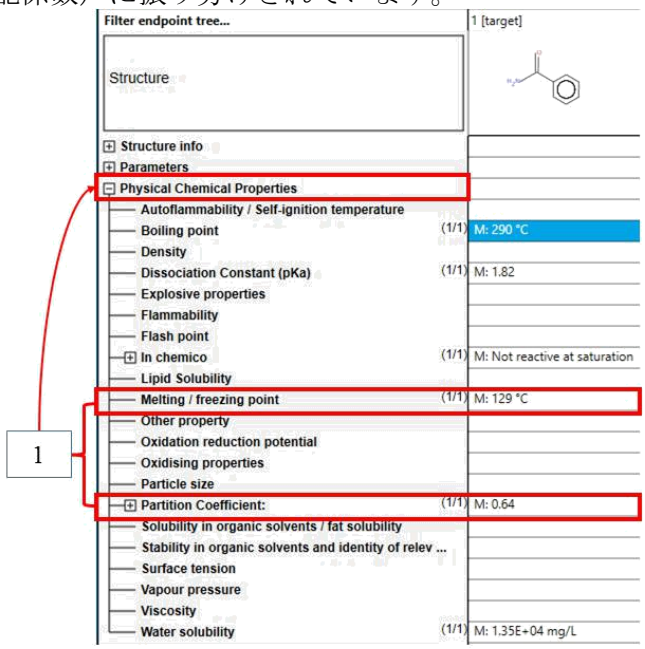


図 5

また、Ames 試験や染色体異常に関連するデータは、Human Health Hazards（ヒト健康有害性）ノードのサブノードである Bacterial reverse mutation assay (e.g. Ames Test)（細菌の逆突然変異アッセイ、例えば Ames 試験）および In vitro mammalian chromosome aberration test（In vitro 哺乳類染色体異常試験）（2）に割り当てられています。（図 6）

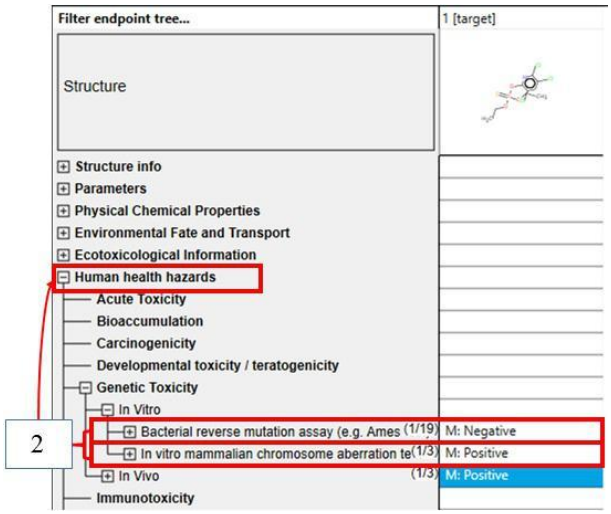


図 6

4.1.2. エンドポイントツリーの構造

エンドポイントツリーは、既定部分と可動部分の 2 つの部分で構成されています。既定部分は固定されていて並べ替えできませんが、可動部分は柔軟に並べ替えられます。様々なデー

タベースから得られる実測データが多様であるためこの機能が実装されました。既定部分はデフォルトで太字になっており（図 7）、可動部分は太字ではありません（図 8）。可動部分は、利用可能なメタ（付随）データによって変わります。したがって、ユーザは、**データ・マトリックス**上の化学物質のデータを収集した後に、エンドポイントツリーのこの部分を表示することができます（Collecting data のセクションを参照）。

- **Predefined part（既定部分、図 7）**

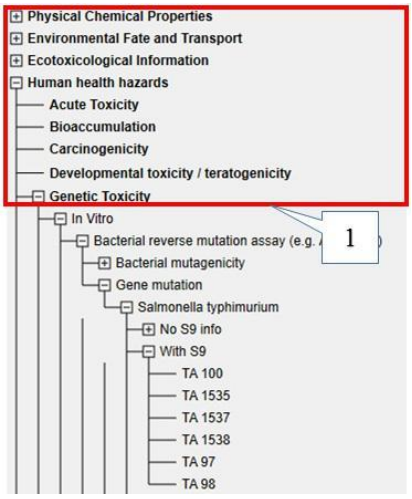


図 7

- **Dynamic part (可動部分、図 8)**

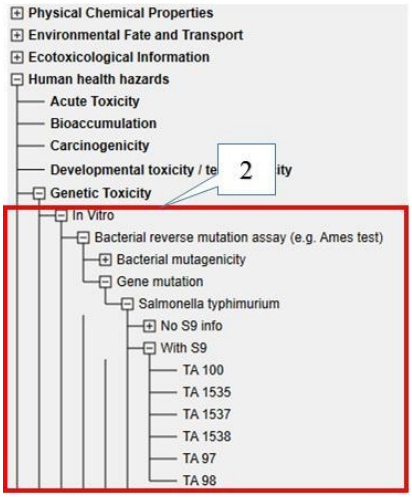


図 8

エンドポイントツリーの可動部分は、実測データに関するメタデータ欄を使って構築されます。図８，９の事例において、 in vitro および in vivo は、Type of method（手法の種類）（1）と呼ばれるメタデータ欄の要素です。（図 9）

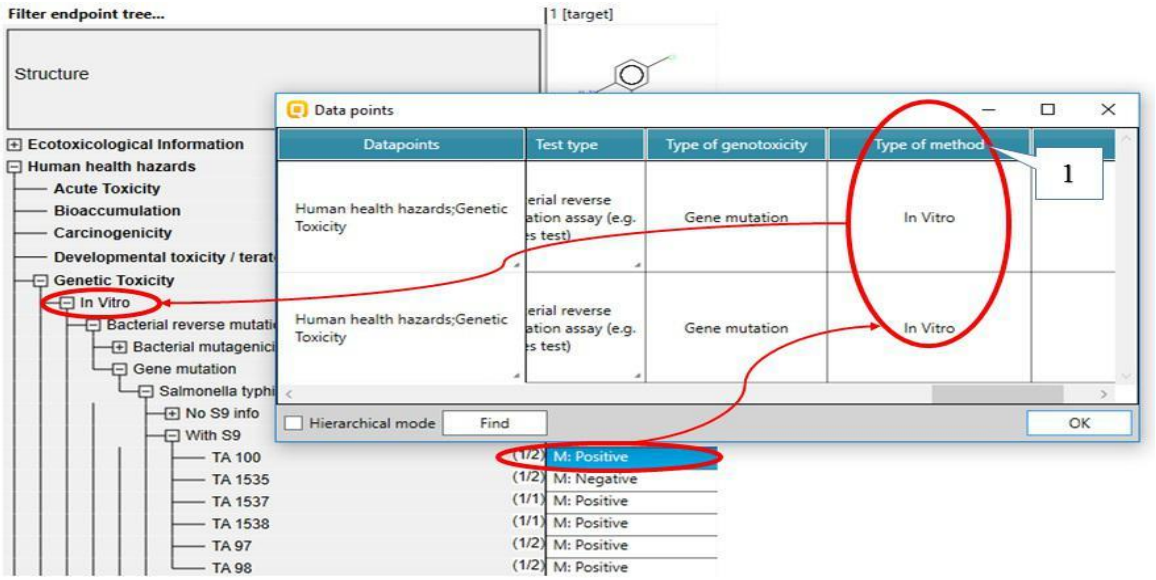


図 9

次の Bacterial reverse mutation assay (e.g Ames test)（細菌の逆突然変異アッセイ、例えばエイムス試験）ノードは Test type （試験の種類）（2）です。（図 10）

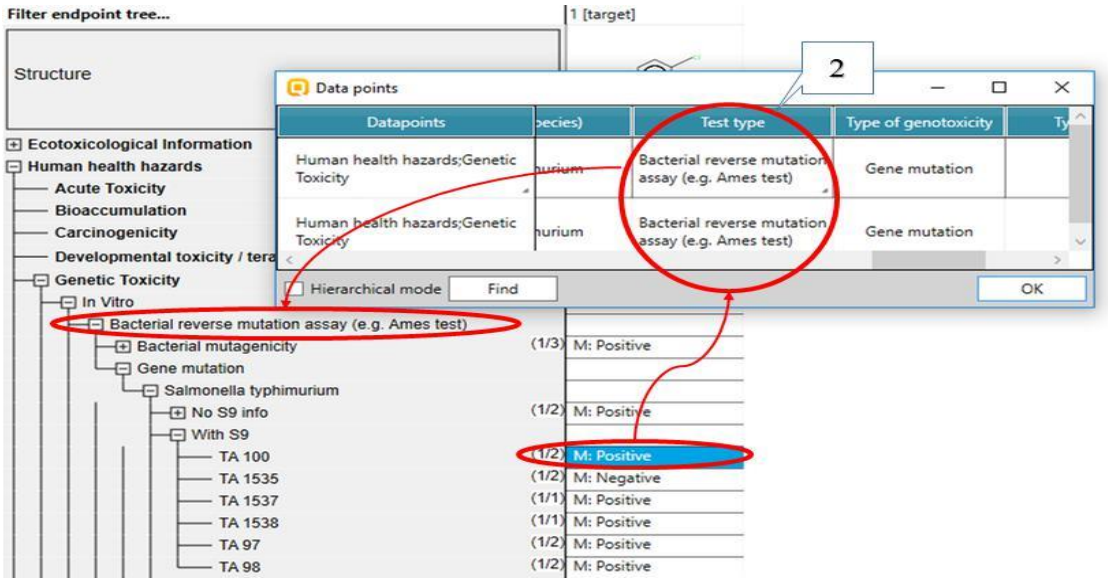


図 10

その下の 2 つのノード、Gene mutation（遺伝子突然変異）と Salmonella typhimurium（ネズミチフス菌）は、データベースの遺伝毒性の種類および試験生物（種）の欄と連動しています。（3）（図 11）

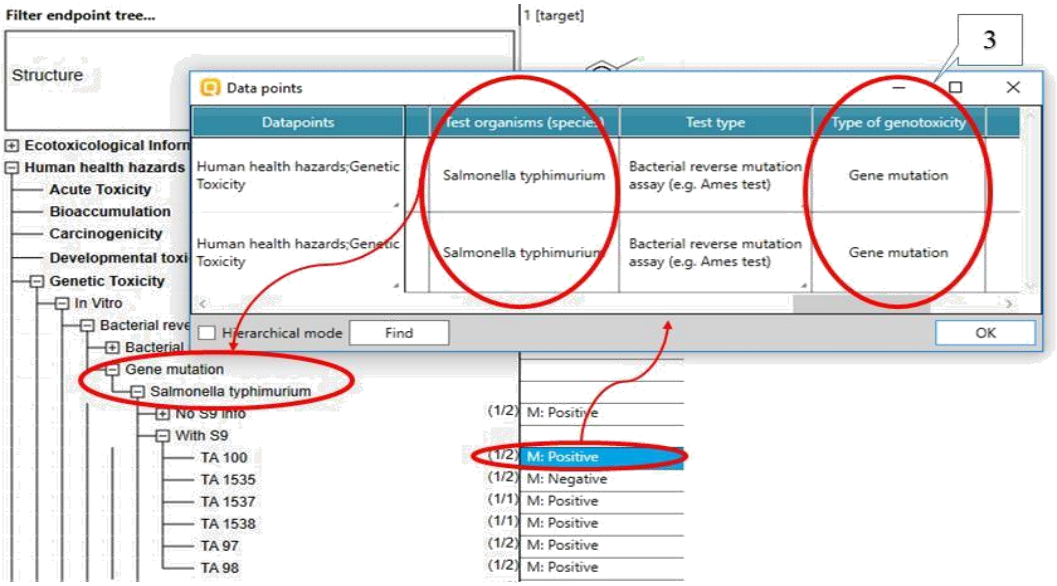


図 11

最末端の 2 つのノードは、次の 2 つのフィールド、Metabolic activation（代謝活性化）および　Strain　（生物系統）（4）と連動しています。（図 12）

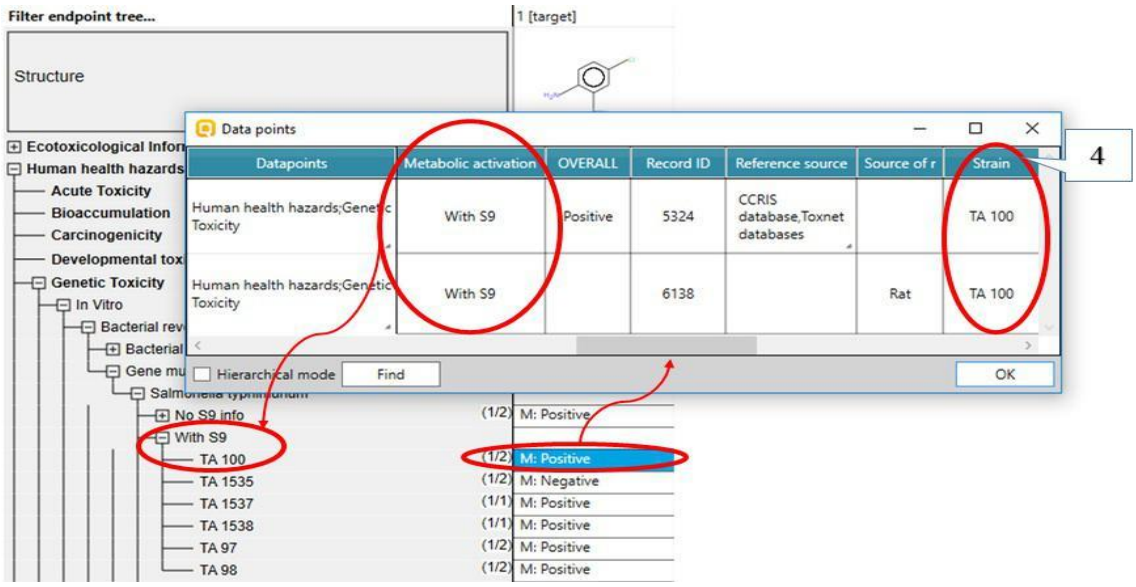


図 12

これらの各フィールドは、Set tree hierarchy（階層構造の設定）機能を使用して並べ替えることができます。このオプションを利用するには、階層構造を並べ替えたいノード（固定されている部分に相当、(1)）を右クリックし、メニューから Set tree hierarchy（2）をクリックします。小さな青い三角は、階層構造を設定できるレベルのノードであることを示します（3）。（図 13）

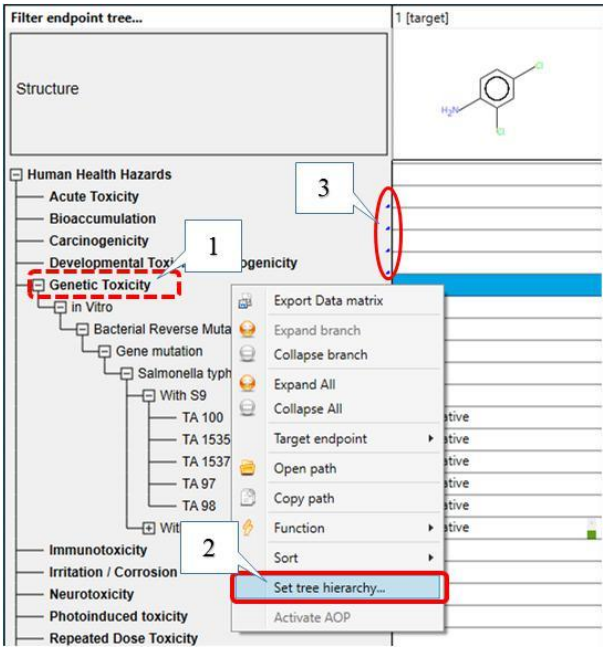


図 13

Set tree hierarchy ウィンドウが表示されます。

4.1.3. ツリー階層機能の設定

Set tree hierarchy ウィンドウには、Metadata labels（1）と Sub-nodes（2）の 2 つのパネルがあります。（図 14）

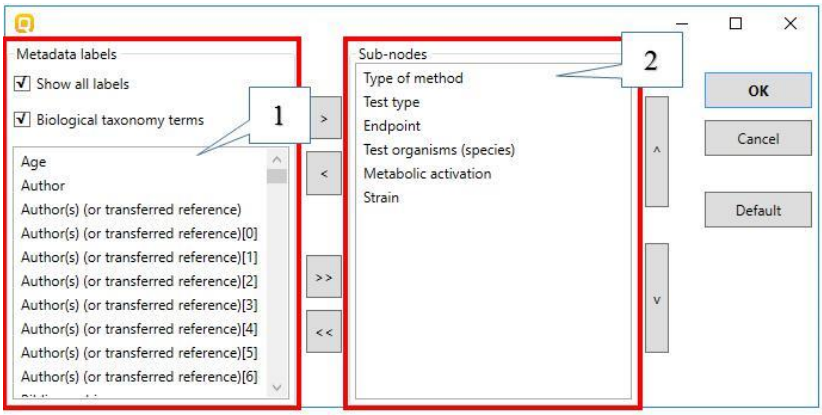


図 14

ツールボックスでは、デフォルトの階層構造が設定されています。Metadata labels パネルには、最も使われそうな欄のリストが含まれています。ユーザが別のフィールドをサブノードとして設定したい場合は、Show all labels ボックス（1）をチェックすると、様々なデータベースで使用可能なすべてのラベルのリストが表示されます（2）。（図 15）

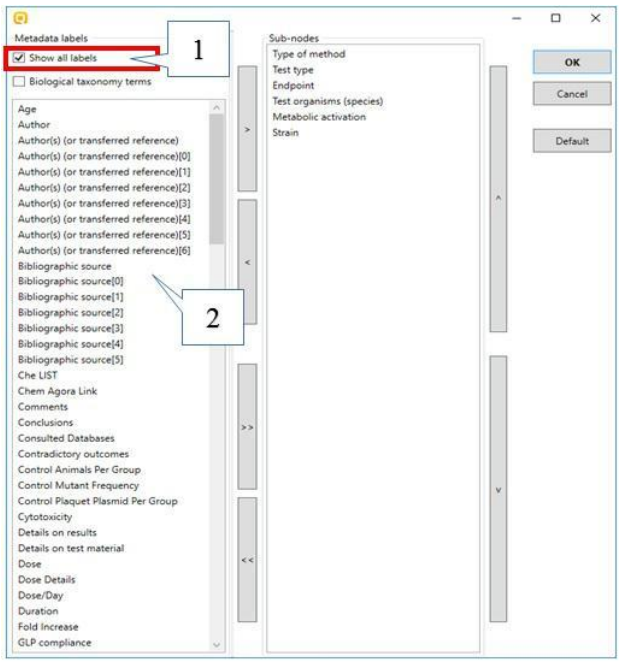


図 15

さらに Biological taxonomy terms （生物分類学用語）ボックス（1）をチェックすると、試験生物の生物分類学に関するすべての用語を列挙することができます。利用可能なメタデータラベルの後に用語が表示されます（2）。（図 16）

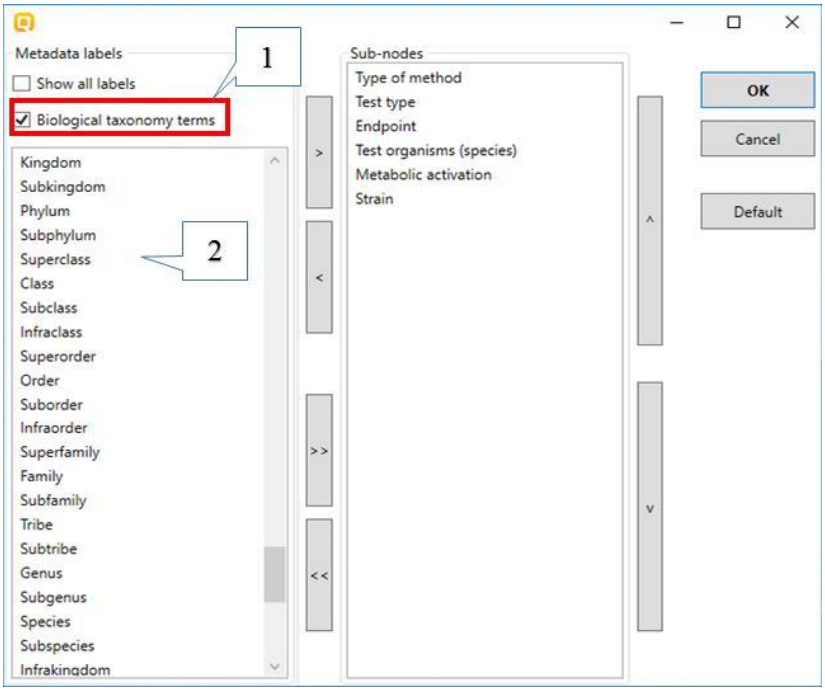


図 16

Sub-nodes パネルに表示される一連のフィールド（1）は、エンドポイントツリーのノードの体系を示します（2）。（図 17）

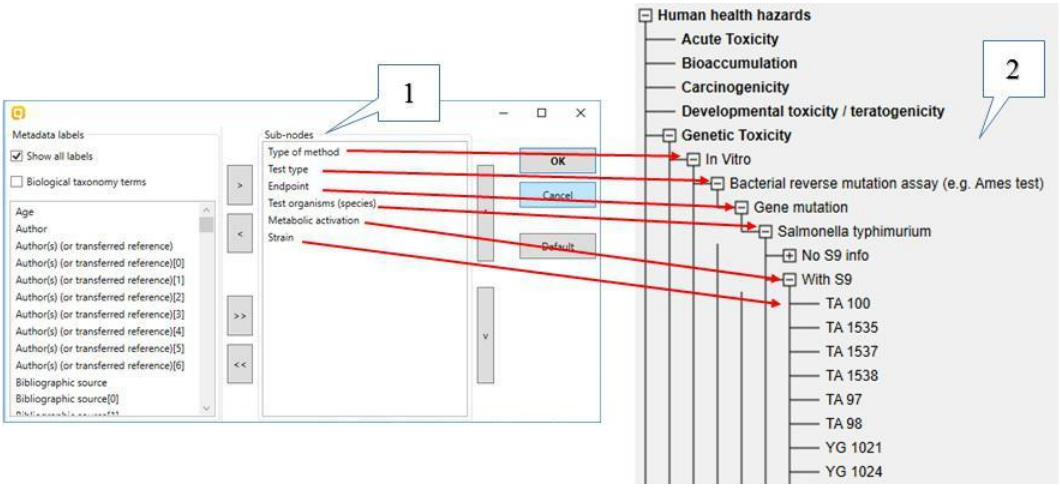


図 17

補助ボタン（1）を使用すると、フィールドを追加したり、サブノードとして既に指定されているフィールドを削除したりすることができます。サブノードは、上方向および下方向ボタン（2）で並べ替えることができます。もしエンドポイントツリーのデフォルト設定をリセットしたい場合、デフォルトボタン（3）をクリックしてください。(図 18)

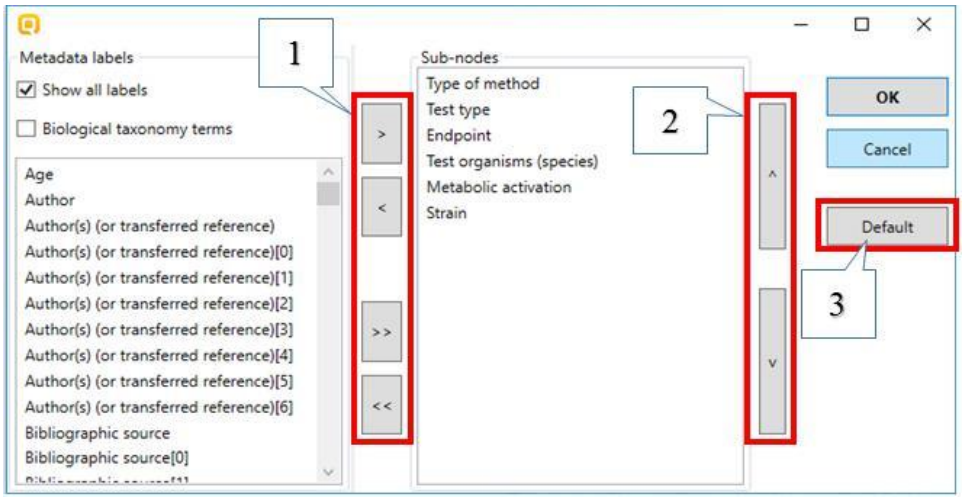


図 18

OK ボタンを押すとエンドポイント階層の変更が確定します。

4.1.4. エンドポイントツリーのノードのフィルタリング

エンドポイントツリーのノードは、 Filter endpoint tree…（エンドポイントツリーをフィルタリングする）でフィルタリングできます。エンドポイントツリーをフィルタリングするには、Filter endpoint tree ...（1）という名前の空白のフィールドに目的のクエリ（訳注：D.1.3.2 章 Query tool 参照）を書き込んでください。この時エンドポイントツリーの白い部分が緑色になるので（2）、Enter ボタン（3）を押してください。（たとえば、 "skin"と入力して Enter を押すと、エンドポイントツリーがフィルタリングされ、skin に関連するノードのみが表示されます）（図 19-20）

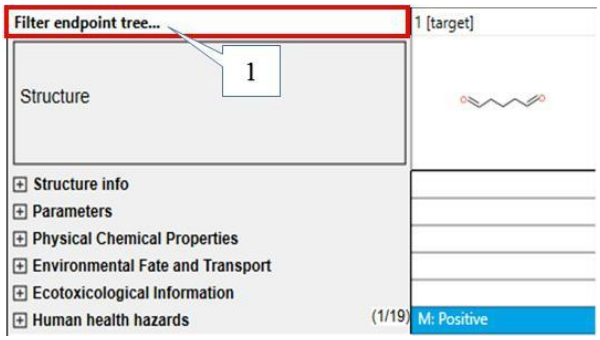


図 19

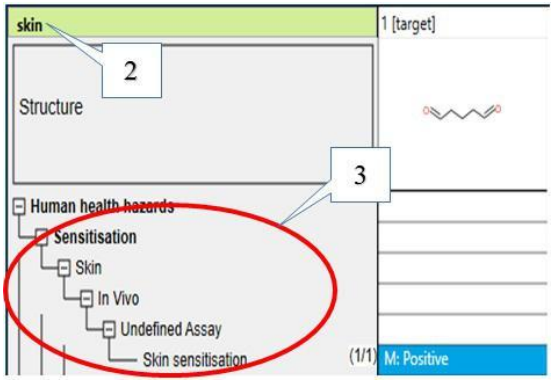


図 20

ユーザが明示したクエリを削除し、キーボードの Enter ボタンを押すと、システムはエンドポイントツリーをデフォルト設定に復元します。

4.1.5. ヒント

エンドポイントツリーの領域をマウスで右クリックすることで他の追加機能を利用できます。

- **Expand（展開） -** Expand branch（ツリーの枝部を展開）（1）および Expand all（すべて展開）（2）のオプション機能により、現在のノードまたはエンドポイントツリー上の折りたたまれたすべてのノードを展開できます。Expand branch オプション機能は、展開できる枝部に対してのみ有効です（3）(図 21).

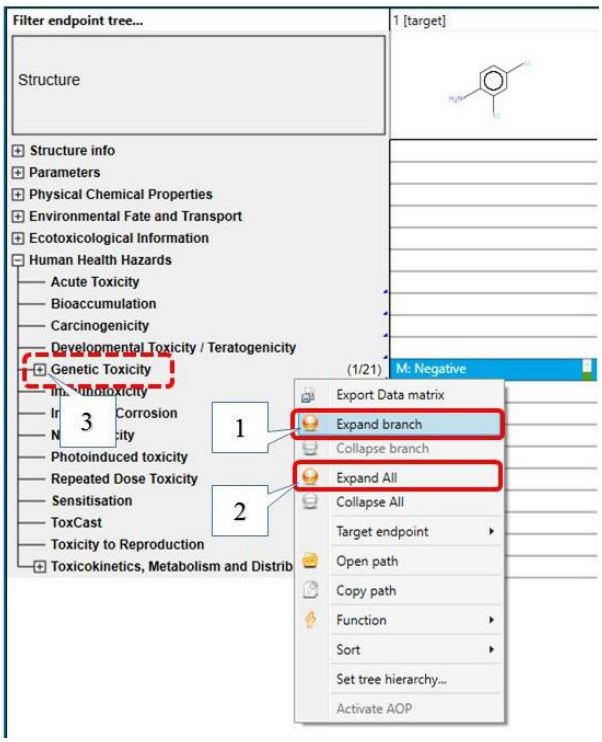


図 21

- **Collapse （折りたたみ）** - Collapse branch（枝部の折りたたみ）（1）と Collapse all（すべての折りたたみ）（2）オプション機能により、現在のノードまたはエンドポイントツリー上のすべての展開されたノードを折りたたむことができます。Collapse branch オプション機能は、一度展開された枝部に対してのみ有効です（3）。（図 22）

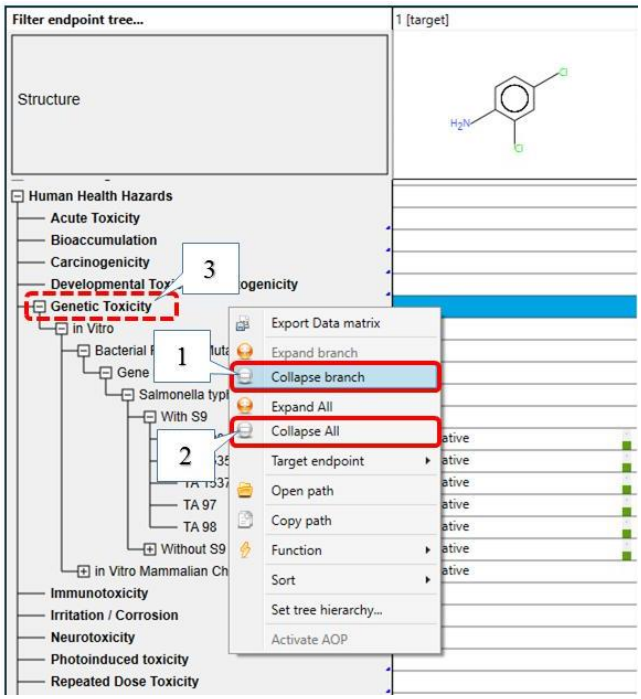


図 22

●**Export Data matrix（データ・マトリックスエクスポート）** - このオプションによりデータ・マトリックスから利用可能なデータをエクスポートできます（1）（図 23）。

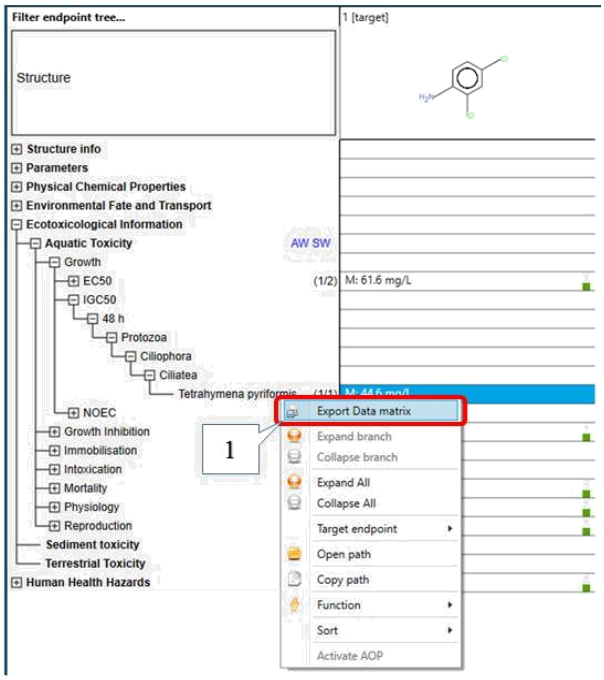


図 23

- **Target endpoint（標的エンドポイント）** - このオプションによりエンドポイントのツリーから標的エンドポイントを直接決定できます。そのためには、データが事前に集められていなければなりません。目的のエンドポイントの最末端（1）を右クリックしてください。Target endpoint を選び、Define（2）を選択すると、Define endpoint のウィンドウが表示され、ここで選択したエンドポイント（1）に従ってすべての入力欄を埋められます（図 24）。最後に、Finish（4）をクリックして確定します。

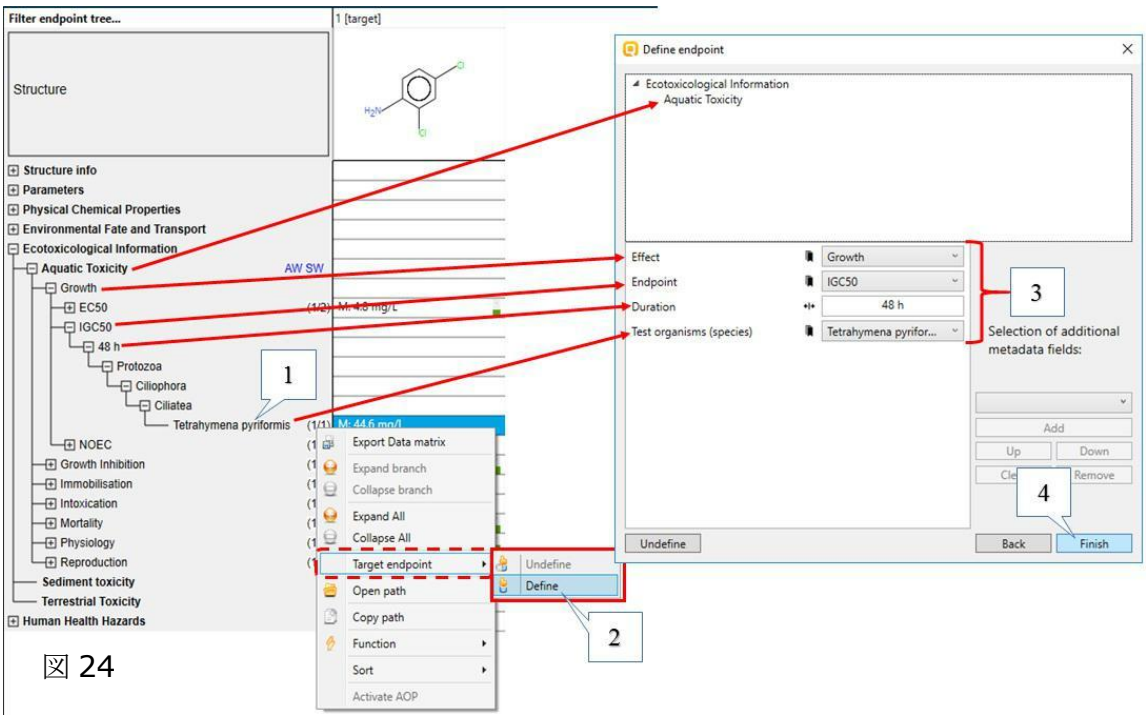


図 24

- **Copy path（エンドポイントツリーにおける位置のパスをコピー）** - このオプションにより、エンドポイントのファイル場所（1）を現在の位置（2）にコピーできます（図 25）。

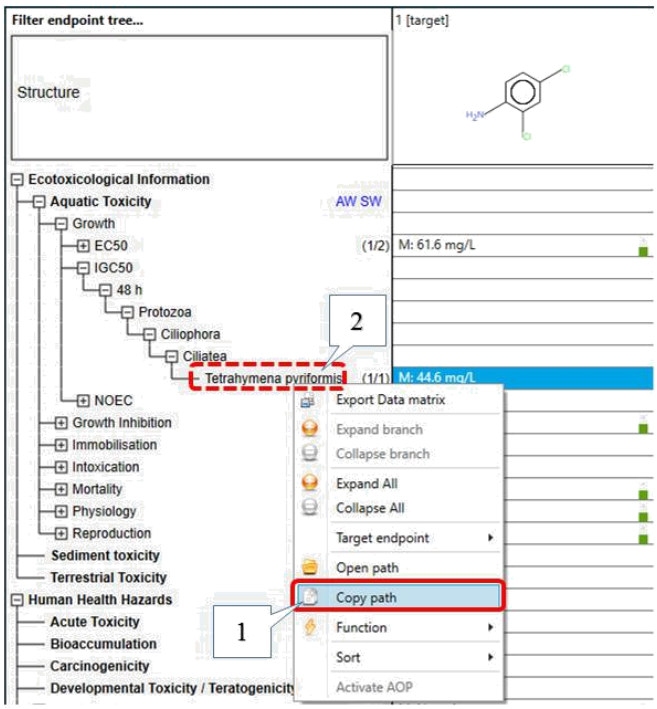
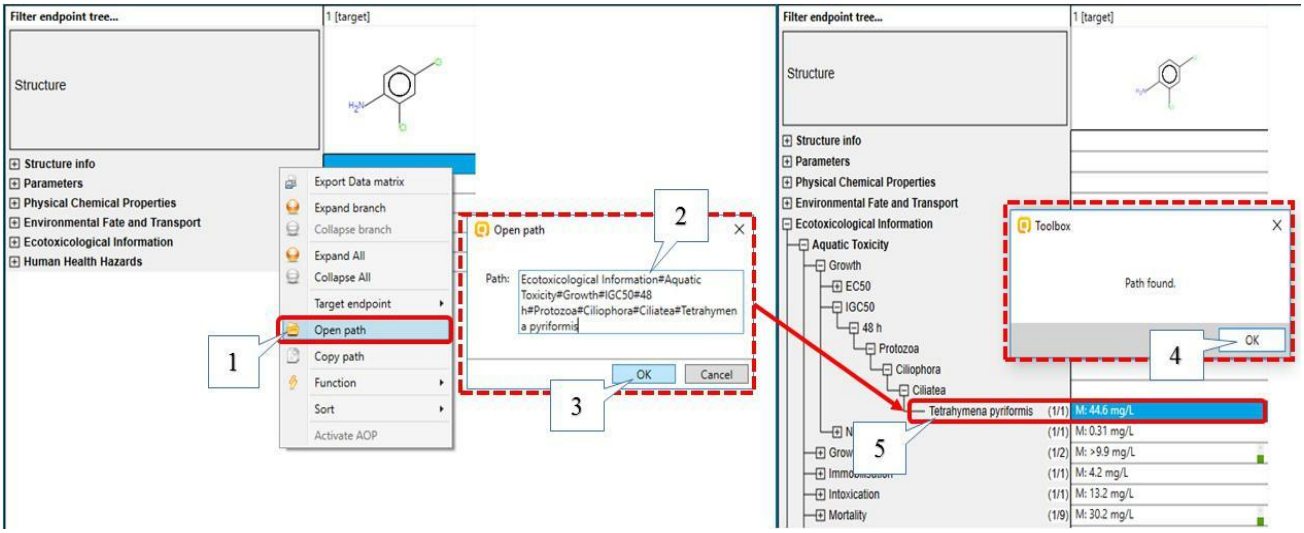


図 25

- **Open path（パス指定によるエンドポイントツリーを表示）** - エンドポイントツリーにおける位置のパスをコピーしておくと（図 25）、エンドポイントツリーが折りたたまれている場合でも、右クリックオプション（図 26）内で Open path（1）を選択することによってパス指定した部分を表示させることができます。ファイル場所を空欄に貼り付け（2）、OK（3）をクリックしてください。Path found と表示されたら、OK（4）で確定します。パスで指定したエンドポイント（5）が表示されます。



. 図 26

ただし、同じパス（1）を貼り付けても、そのエンドポイントに対応する利用可能なデータが**データ・マトリックス**にない場合、エンドポイントツリーの既定部分のみが表示されます（2）（図 27）。

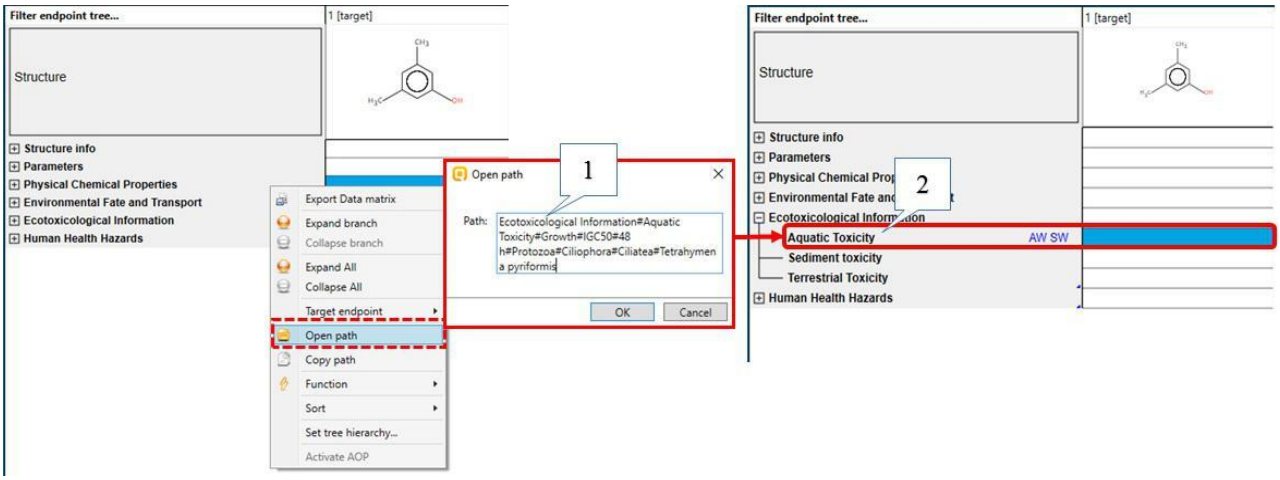


図 27

● **Function（関数） -** １つの化学物質に複数の実験データがある場合、この機能を使うと最小値、最大値または平均値を表示します。この機能は、指定された行のデータすべてに対して機能します。デフォルトですべての値が表示されています（1）（図 28）。

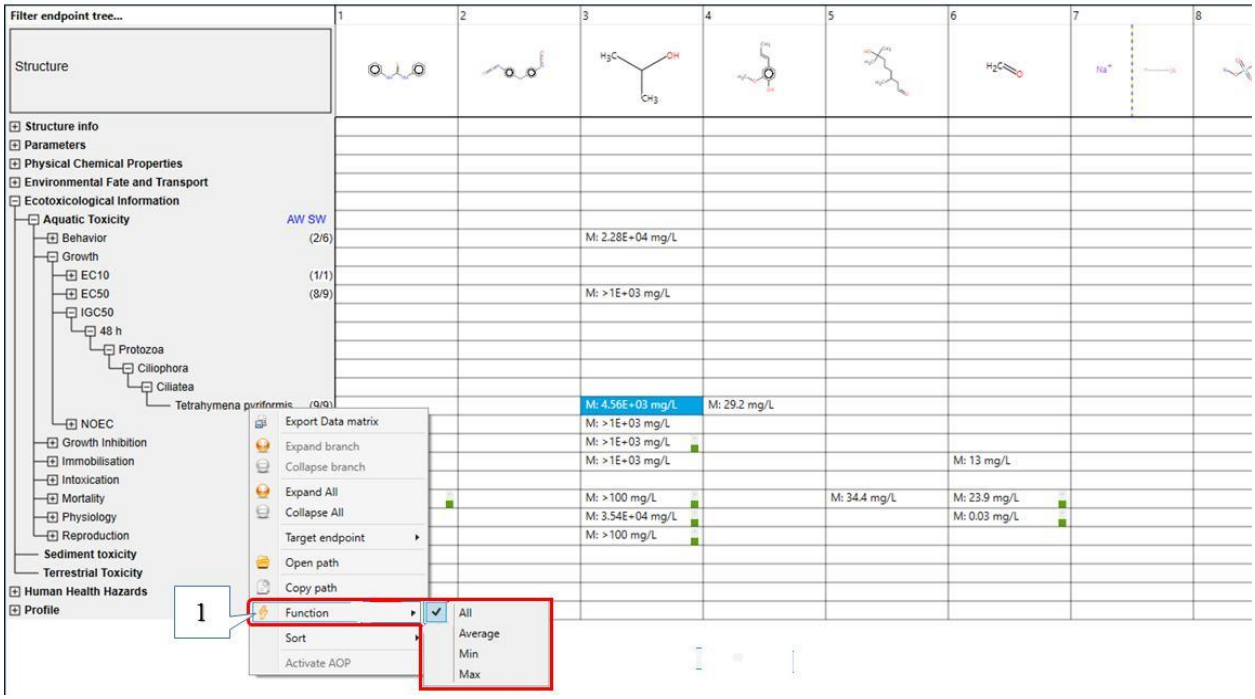


図 28

- **Sort（並べ替え）- データ・マトリックス**に表示された指定の行の CAS 登録番号、計算されたパラメータ、および/または実験データの並べ替えができる機能があります。フィルタリングの対象となるデータを持つノードを右クリックして、 Sort（1）を選択してください（図 29）。このオプションにより化学物質は、標的化学物質の優先順位（もしあれば）を考慮して、実験データにより降順または昇順に並べ替えられます。優先順位とは、**データ・マトリックス**に標的化学物質がある場合、それは最初（一番左）の列に残し、他の化学物質は標的物質の後に降順または昇順で配置されるという意味です。

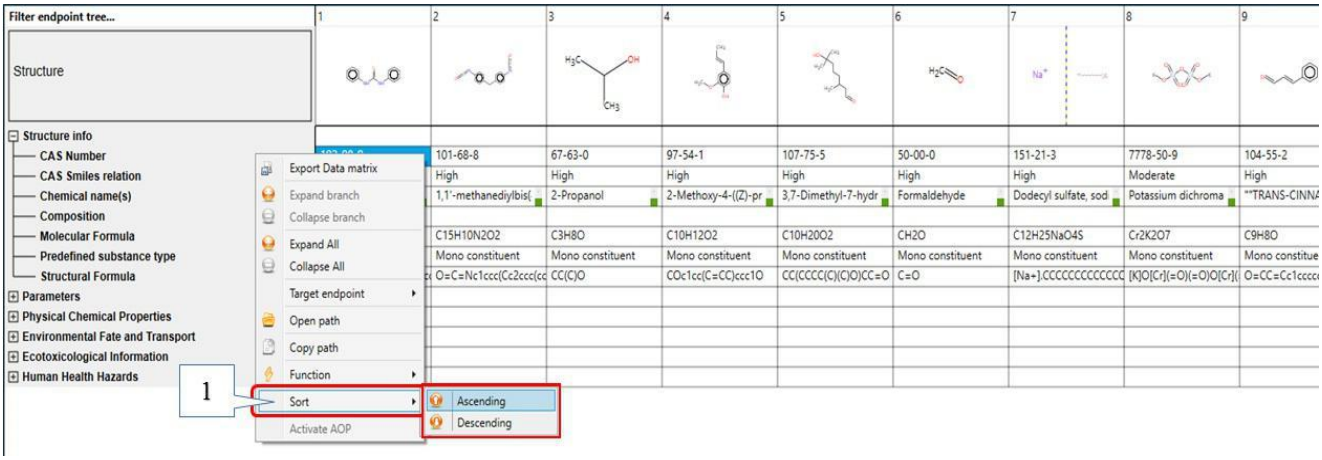


図 29

- **Activate AOP（AOP のアクティベート） -** 皮膚感作性の概念実証 AOP が QSAR Toolbox に導入されています。AOP は、Sensitization の行（1）に移動し、Activate AOP（2）（図 30）を選択することによってアクティブにすることができます。より詳細な AOP の使い方については、ツールボックスの「AOP ワークフローの導入に関するチュートリアル：皮膚感作」へのリンク（http://oasis-lmc.org/products/software/toolbox/toolbox-support.aspx）を参照してください。

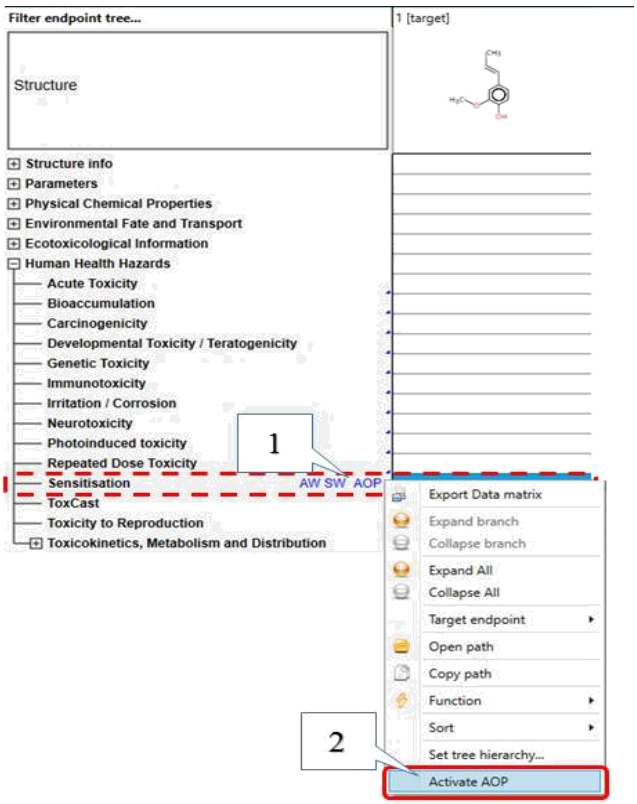


図 30

C.4.2. 選択された化学物質のエリア

選択された化学物質のエリア （図１-２）

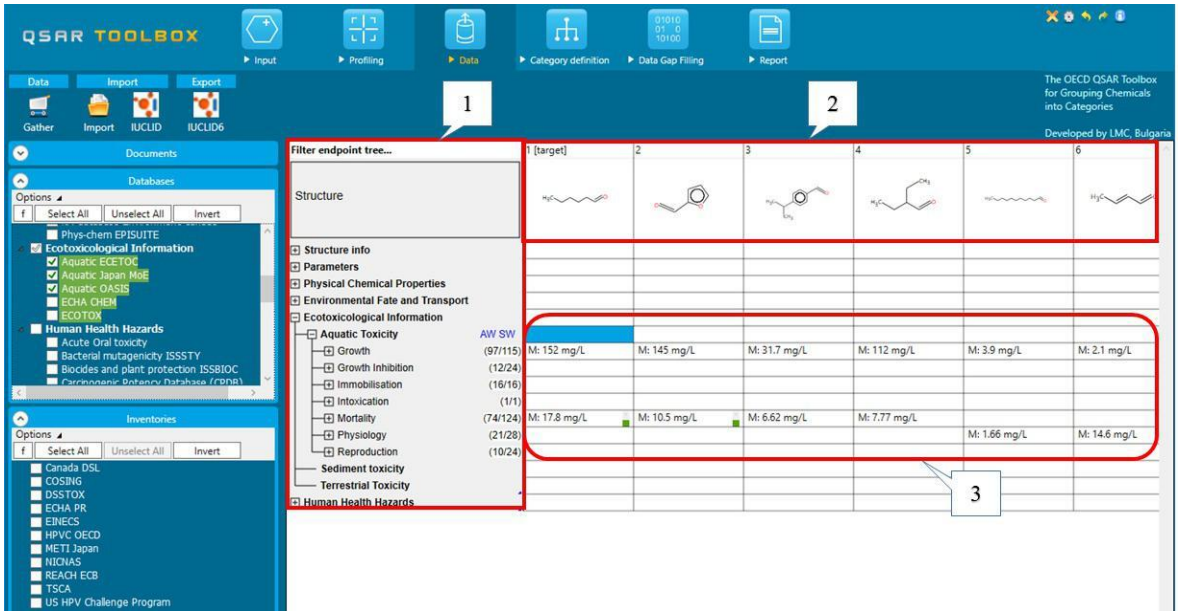


図 1

4.2.1. Chemical information（化学的情報）

この機能は成分、添加物および不純物に関する利用可能な情報を提供します（1）。標的構造を右クリックして Chemical information（化学的情報）（2）を選択すると、それらを見ることができます。新しいポップアップウィンドウが表示されます。この画面は、主成分（3）と各構成要素の詳細（4）の 2 つの部分で構成されています。物質の成分、添加物および不純物がもしあるならば、詳細が書かれた画面のほうの分離されたタブ（5）に配置されています（図 2）。

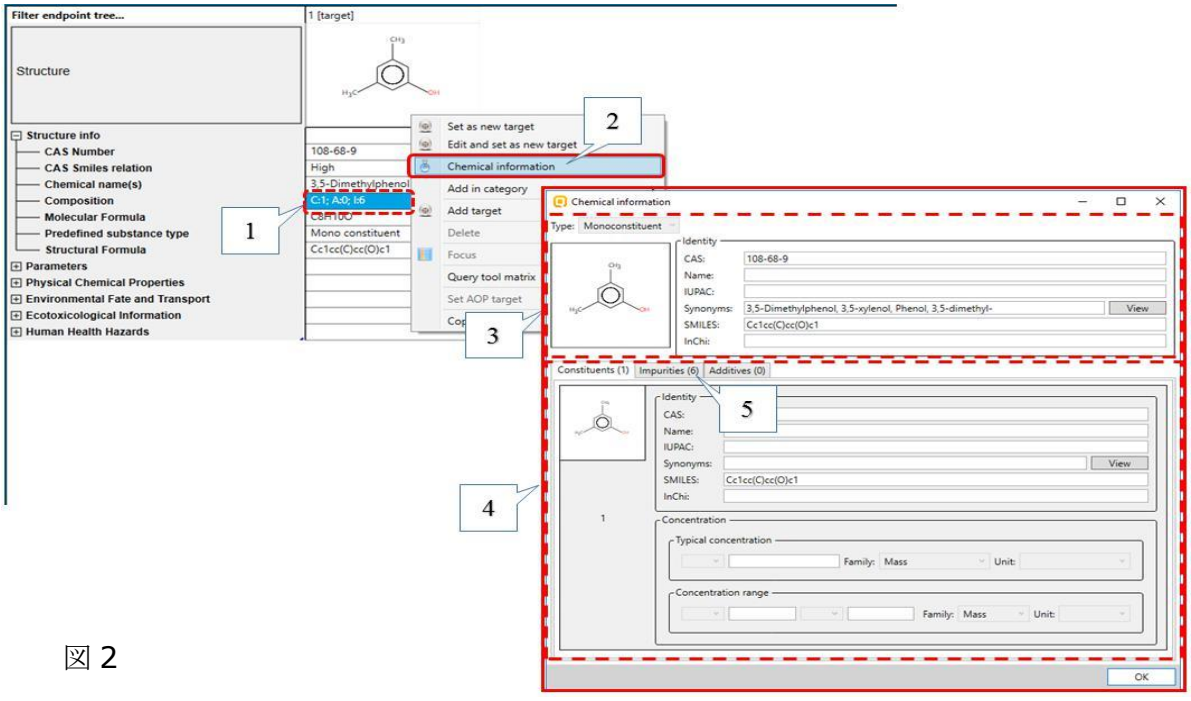


図 2

4.2.2. Query tool matrix（クエリツールマトリックス）

ユーザは、**データ・マトリックス**上から化学物質の中を検索することができます。右クリックオプション機能で Query tool matrix（クエリツールマトリックス）（1）を選択すると実行できます。（訳注：標的化学物質の上で右クリックすること） (図 3).

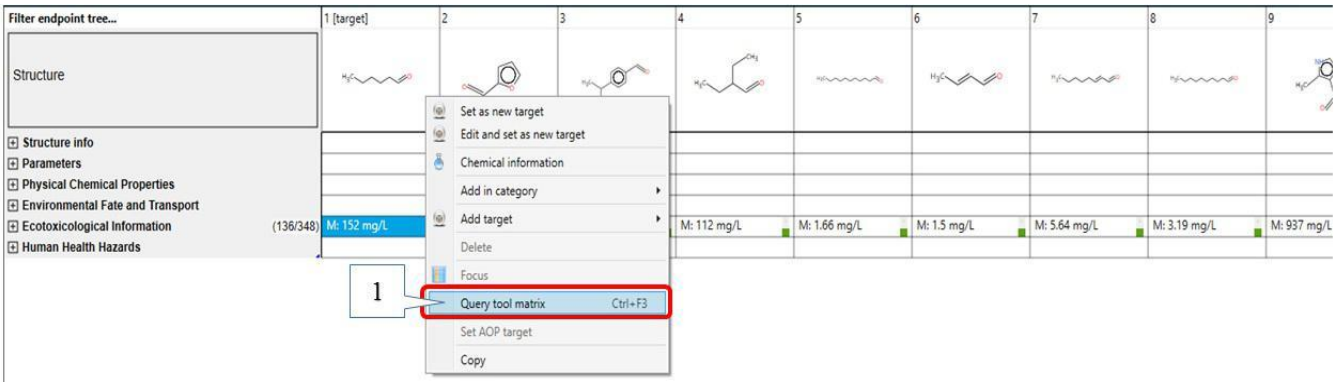


図 3

4.2.3. 新しい標的化学物質の追加

標的化学物質を別の化学物質に変更できる右クリックオプションがいくつかあります：

- **.Set as new target**（新規の標的物質をセットする）（図 4）- **データ・マトリックス**で利用可能な別の化学物質を新しい標的物質として選択することができます。

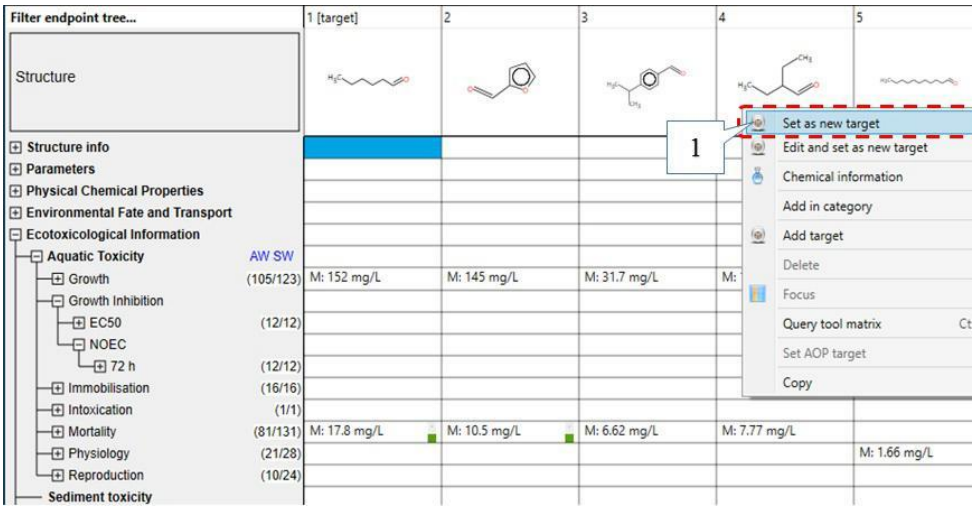


図 4

- **Edit and set as new target**（編集して新規の標的物質としてセットする）（図 5－1） - 2D エディタ（2）で、**データ・マトリックス**で利用可能な化学物質を編集し、新しい標的物質（3）として設定します。

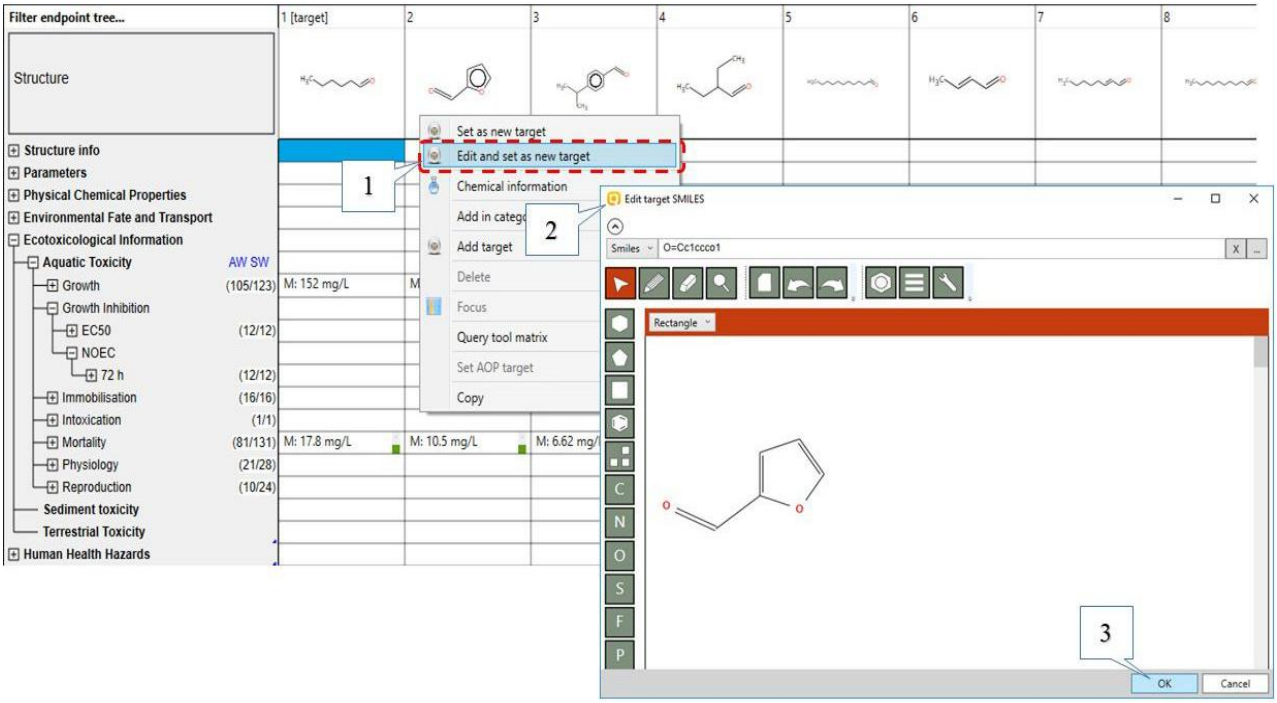


図 5

- **Add target**（標的物質の追加）（図 6－１） - ユーザが Input モジュール上で作業しているかどうかにかかわらず、すべての Input モジュールと同様に CAS 番号、化学物質名、化学構造などによって新しい標的物質を追加することができます。

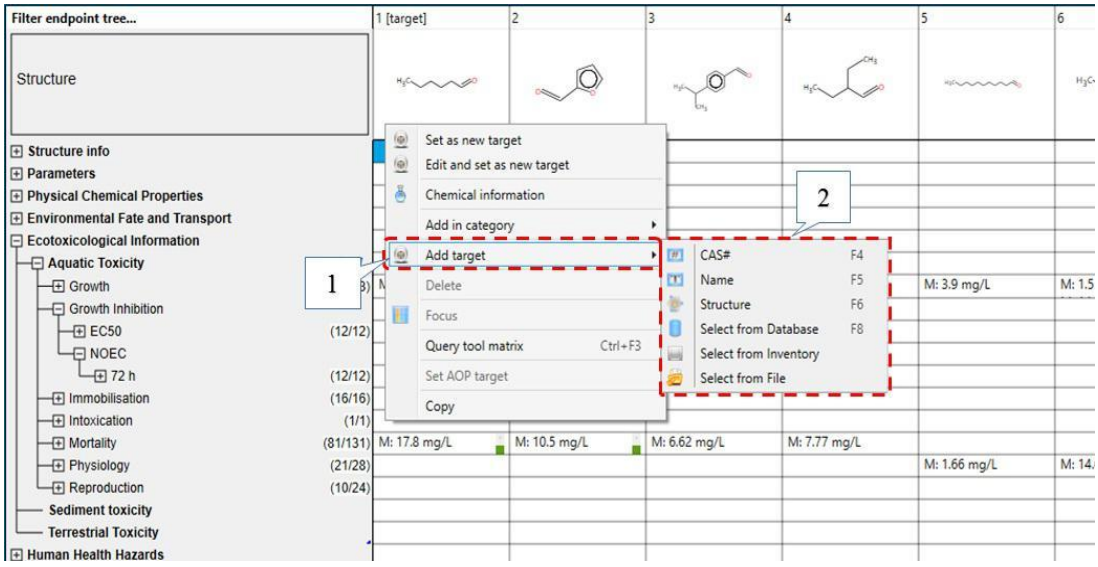


図 6

4.2.4. 付属の右クリックオプション

- **Add in category** (カテゴリ内に追加) （図 7－１） - CAS 番号、化学物質名、化学構造など（2）で**データ・マトリックス**に化学物質を手動で追加することができます。

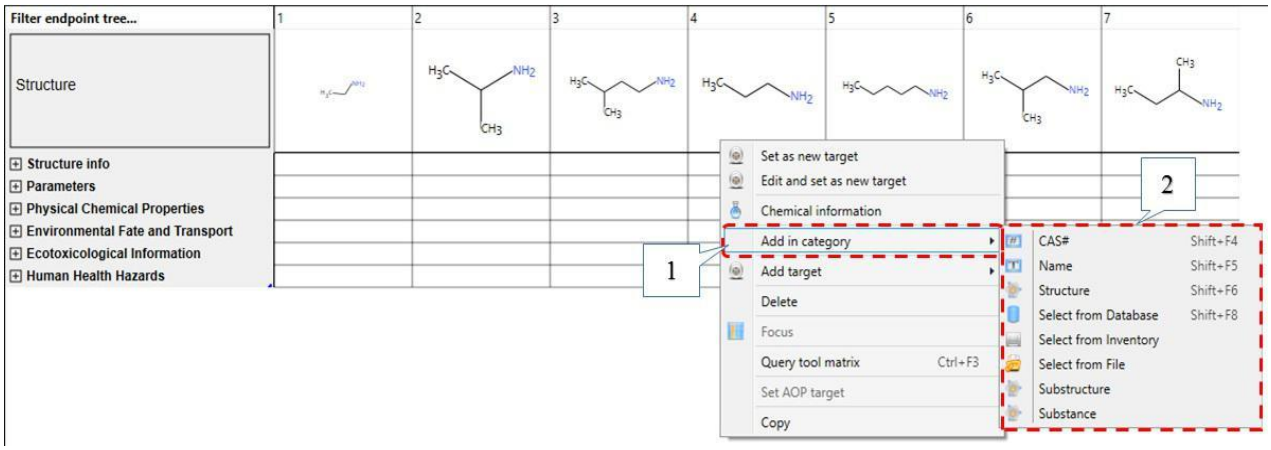


図 7

- **Delete（消去）** （図 8－１） **-** この機能により、**データ・マトリックス**内の化学物質を消去することができます。

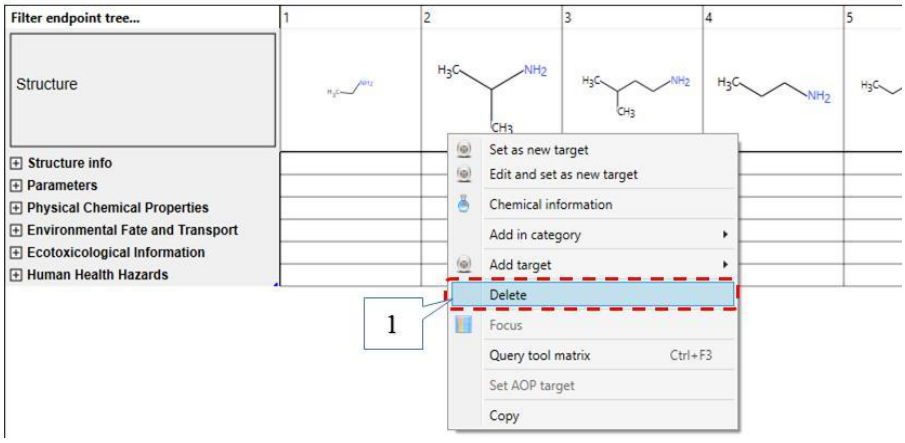


図 8

- **Focus （フォーカスする）- 標的物質が（代謝などで）多元化したとき（1）、選択された代謝物の右クリックオプションで Focus（2）を選択することにより、特定の代謝物についての作業を続けることができます。

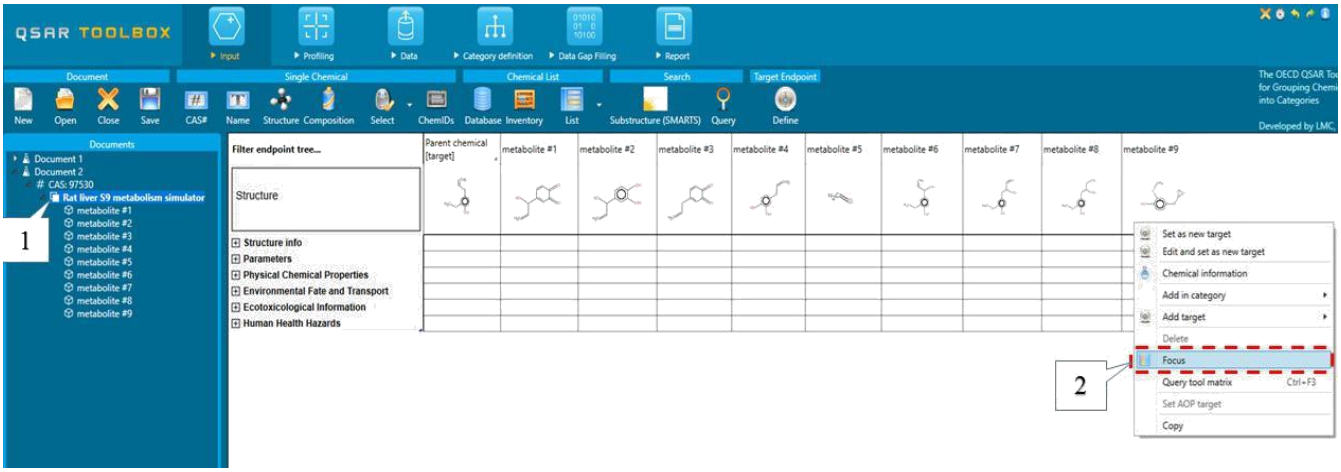


図 9

- **Set AOP target （AOP の標的を設定する） -** AOP がアクティブになったら（1）、分析される化学物質を AOP の標的（2）に設定する必要があります（図 10）。

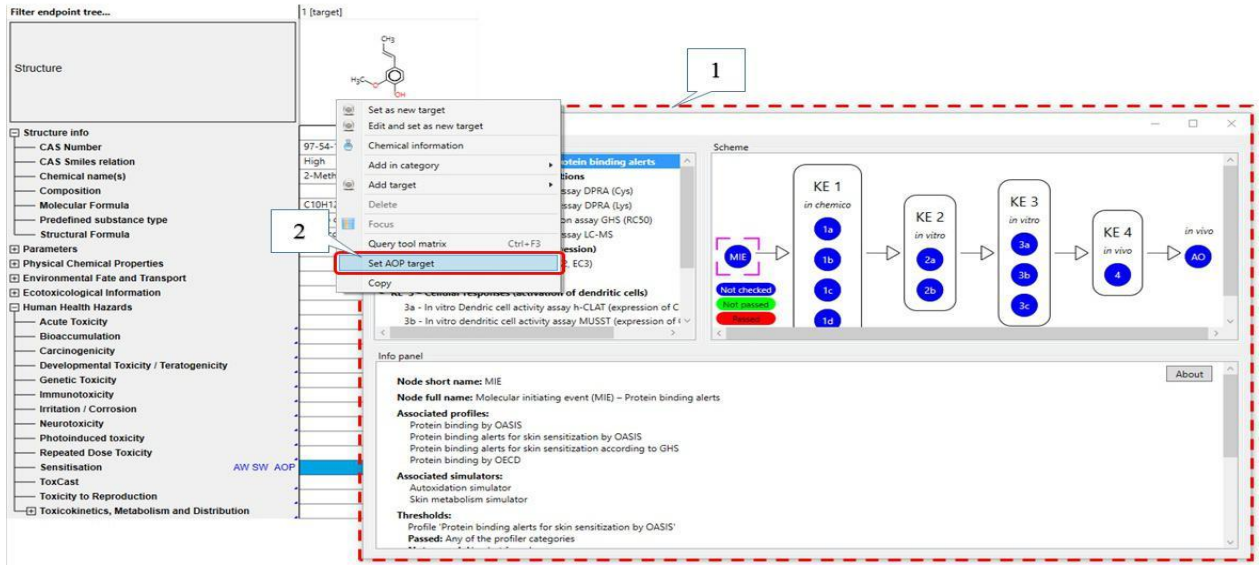


図 10

- **Copy(コピーする) -** 現在作業中の化学構造の SMILES をコピーできます（図 11－1）。

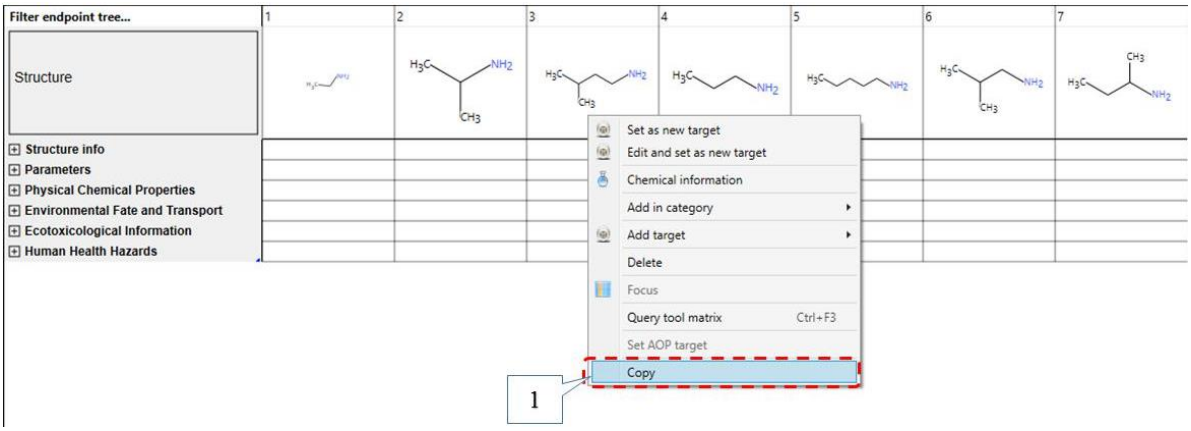


図 11

C.4.3. データのエリア

Area with data （図 1-３）

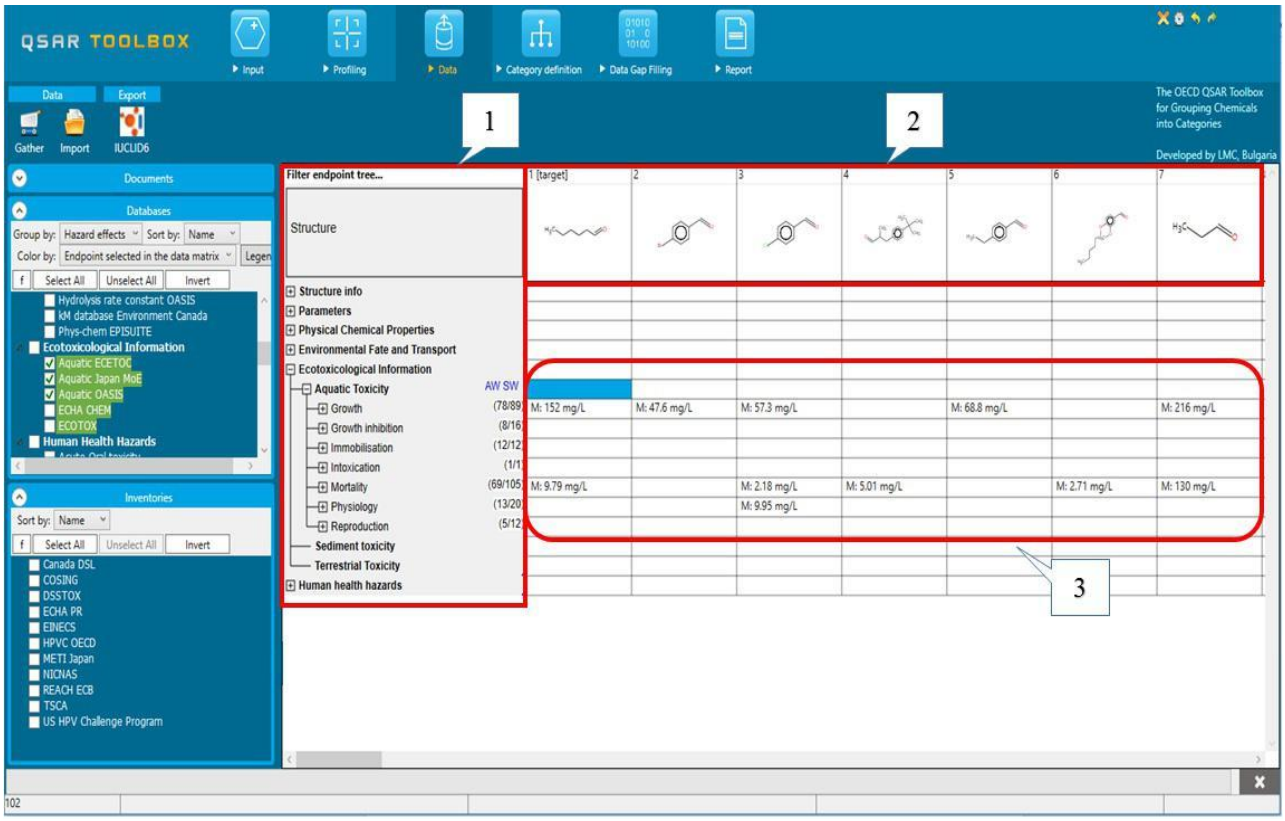


図 1

4.3.1. Search（検索）

Search（検索）オプション機能を使用すると、外部ソースで CAS 登録番号または化学物質名を検索できます。

- **Search of CAS Number（CAS 登録番号の検索） -** CAS 番号（1）のセルを右クリックすると、自動的にポップアップメニューが表示され、**ユーザ**は eChemPortal または Google のウェブサイト（2）内で選択した CAS 番号を検索できます。 （図 2）

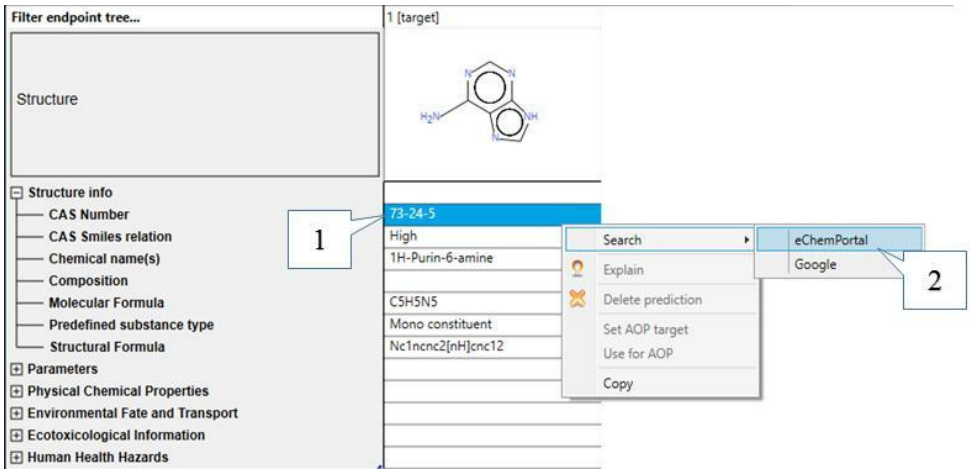


図 2

- **Search of Chemical name(s)（化学物質名の検索） -** Chemical name(s)

（1）のセルを右クリックすると、自動的にポップアップメニューが表示され、**ユーザ**は eChemPortal、PubChem、Wikipedia、Google ウェブサイト内の使用可能な化学名（2）を検索できます（3）。 （図 3）

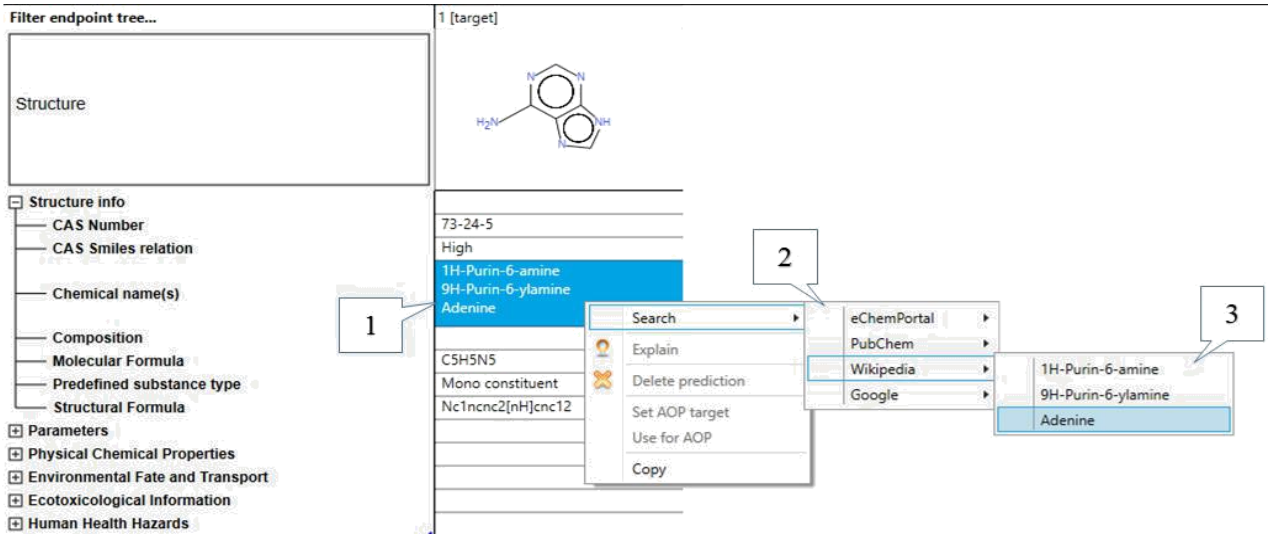


図 3

4.3.2. Quality assessment (QA)（品質評価）- CAS 番号と SMILES の対応関係がどれくらい信頼できるか、Toolbox で使用可能なすべての化学物質について示しています。QA（1）をダブルクリックするか、右クリックして Explain（2）を選択すると、CAS-SMILES 対応関係のウィンドウが表示されます。現在見ている化学物質の CAS 登録番号と SMILES が上部（3）に表示されます。このウィンドウには、以下の情報も含まれています。化学物質が存在するすべての情報源名と情報形式（4）、 情報源の品質（5）および 情報源内の SMILES が他のデータベースで決められたものなのかの情報（6）（図 4）。

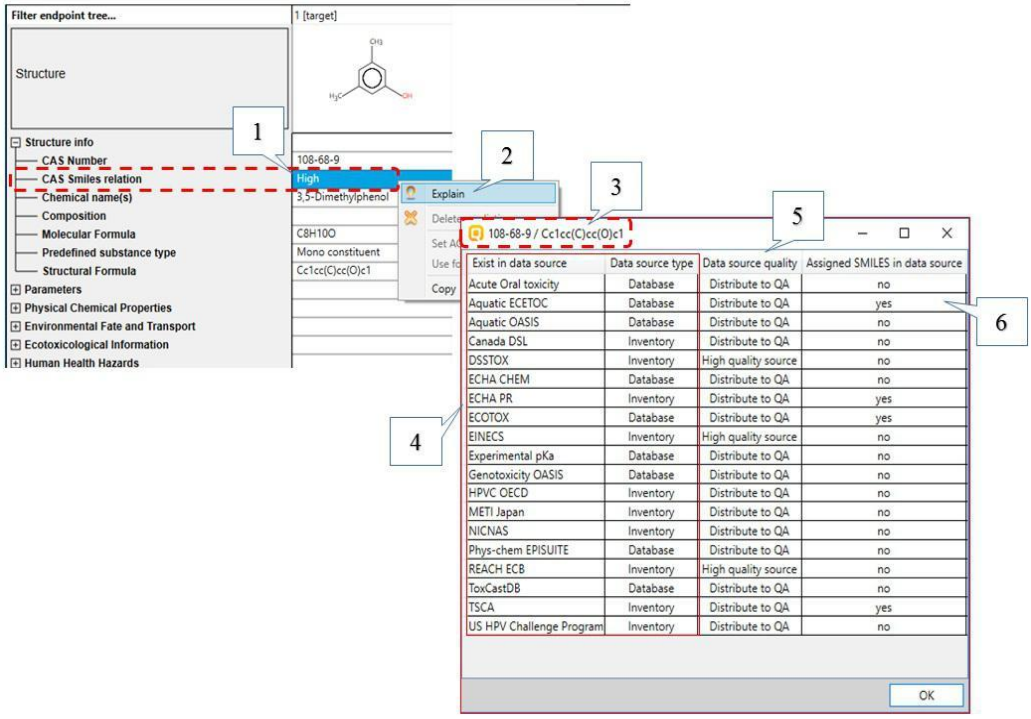


図 4

4.3.3. データに伴うラベル

データ・マトリックス上で見えるデータのラベルがあります：

1. **“M”** - データベースから抽出した測定データを意味する
2. **“R”** - リード・アクロス分析により得られた予測結果
3. **“T”** - トレンドアナリシスにより得られた予測結果
4. **“Q”** - QSAR 予測により得られた予測結果
5. **“IMOA”** - 独立した作用機序に基づき、構成成分から得られた予測結果
6. **“SMOA” -** 類似した作用機序に基づき、構成成分から得られた予測結果

C.5. 情報と活用できるアクセサリ

ユーザ インターフェース

1. 補助機能

ギャップ補完アプローチで使用する類似物質に関連した特定の情報を、ユーザに提供するための新しい機能が実装されました。この補助機能は、さまざまな種類の情報を提供します。例を以下に示します。:

- 類似物質に関する特定の情報を提供する（例えば、類似物質によってはその LC50 値が標的化学物質の対水溶解度より大きいことがある）。システムは、（理論上存在するが）除去すべき不適切なデータ点を検出するために、体積濃度の単位群で表されるデータ値を、最大対水溶解度の計算値と比較する。



「対水溶解度より大きな値があります。対水溶解度より大きな値を除いてください。」

- 警告メッセージ、ユーザへの注意喚起:
  不等号等（例えば、> 5、=100 など）を含むデータ値をもつ類似物質がある



「現在のギャップ補完の状態では不等号等を含むデータ値が含まれます」

組成物を持つ類似物質がある（例えば、混合物、UVCB など）



「現在のギャップ補完の状態では組成物を持つ類似物質が含まれます」

- 通知メッセージの例
  現在の状況で得られたトレンドアナリシス予測結果は、統計によって許容できる



「この予測は統計上許容できるものです。（内挿かつ R2≧0.7 かつ類似物質数≧10）」

- リード・アクロス予測は、5 つ（デフォルト）よりも多くの類似物質によって得られる。記述子の値が等しい化学物質による可能性もある（logKow 値）



「現在のリードアクロスでは予測値に 6 つの類似物質を使用しています」

補助機能がギャップ補完のステージに表示されたら、一度クリックするとメッセージが表示されます。ヘルパーをクリックしてスライドさせると削除できます。

2. 稼働中のアクションおよびバックグラウンドで実行中のアクションのステータスを表示する

稼働中のアクションの進行状況を追跡したり、自動ワークフローの実行時に稼働中のアクションをキャンセルしたりするために使用する機能。(図 1)

稼働中のアクションの進行状況を見るには、時計のアイコン（1）をクリックしてください。

稼働中のアクションをキャンセルするには、停止ボタン（2）をクリックします。

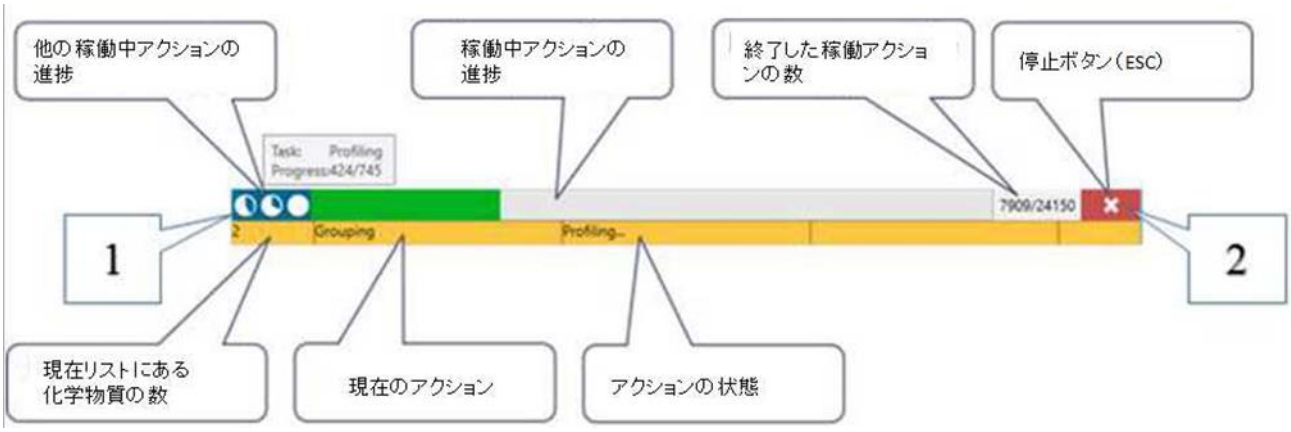


図 1

D 章. ワークフロー

ツールボックスは、ユーザを化学物質の入力からナビゲートするワークフローを提供し、実験データと類似物質を収集してデータギャップを補完したり、予測を報告したりします。

6 つのワークフローセクションがあり、プログラムのインターフェース内に明確に表示されています。:

- Input（入力）
- Profiling（プロファイリング）
- Data（データ）
- Category definition（カテゴリ決定）
- Data Gap Filling（データギャップ補完）
- Report（レポート）

D.1. 入力

ワークフロー

化学物質は、この OECD (Q) SAR のアプリケーション Toolbox において、通常は二次元 (2D) 構造として定義できる原子結合によって識別されます。原子結合は、「原子結合の平面」を一続きの文字列に変換する一連のルールに基づいて、いろいろな識別子でコード化することができます。化学物質はその物質名称、CAS 登録番号、2D 構造の描画などを通してシステムに入力することができ、本システムはそれに該当する原子結合を識別でます。類似物質の検索、アクティビティの予測など、次に示すアクティビティの多くは、この原子結合に基づいています。しかし化学物質に固有な識別子として原子結合を使用すると、いくつか不都合が生じることがあります。

例えば、別々の CAS 登録番号を有する化学物質であっても、それらの準分子状態のために同じ原子結合を持つことがあり得ます。化学的に同一かどうか解釈があいまいな場合、システムは化学物質を異なる標的物質として扱うので、ユーザが**データ・マトリックス**から化学物質の一部を削除することができます。

Input モジュールは、ツールボックスの機能を適用させる標的化学物質の ID をユーザが決定するところです。ここには 2 つの選択肢があります：

- 単一の化学物質を読み込む、
- 化学物質のリストを読み込む。

- **New document（新規ドキュメント作成）**

Input （1）セクションで New （2）をクリックすると、新規ドキュメントが開きます（図 1）。ドキュメントの名前を入力することができ（3）、立体異性を扱うかどうかを決定し（4）、OK（5）をクリックします。（訳注：立体異性体を特に識別したい場合のみ、（4）で Account stereo をクリックすること）

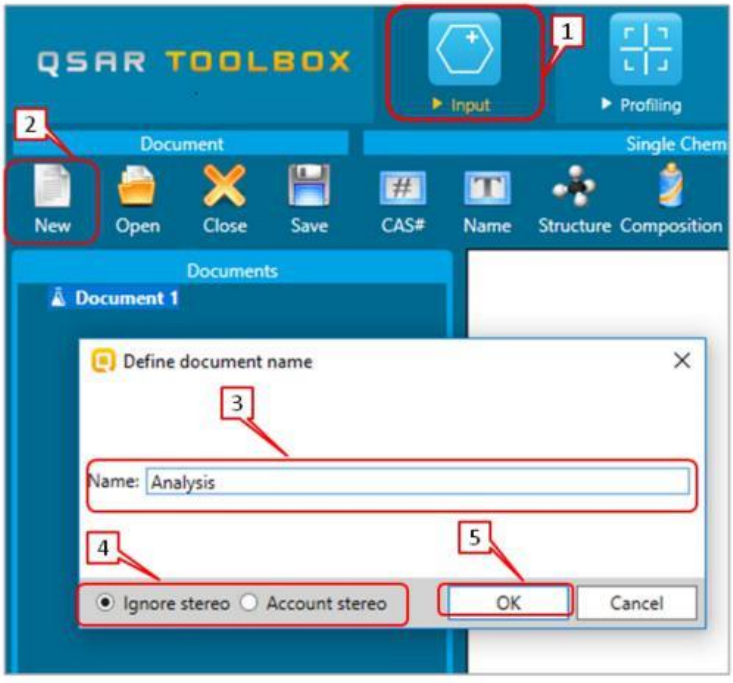


図 1

新しいドキュメントは Documents パネル（1）に表示され、その名前もソフトのウィンドウの上部に表示されます（2）。（図 2）

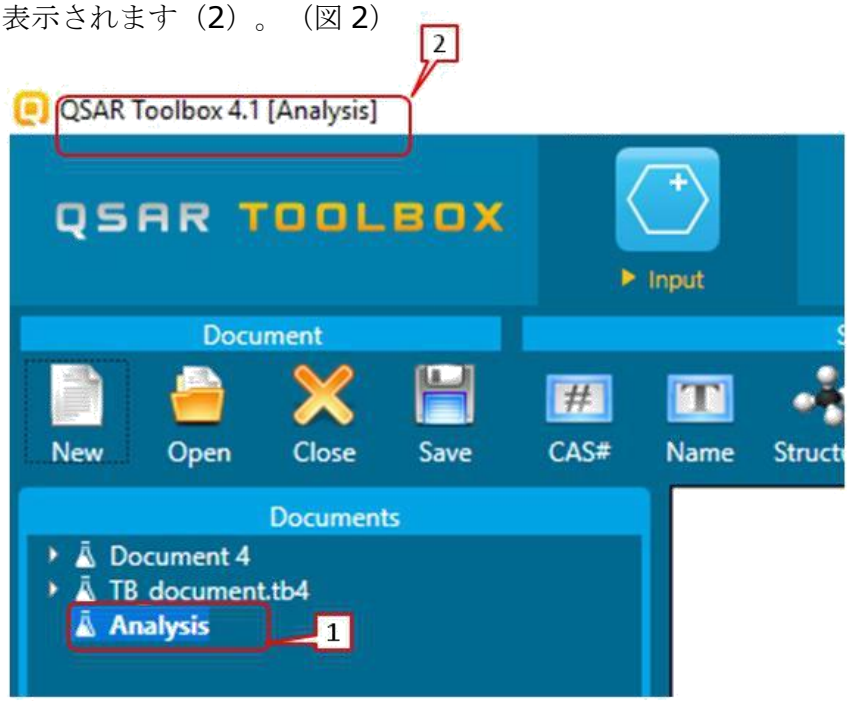


図 2

- **Close a document**（ドキュメントを閉じる）Close (1) をクリックすると、選択したドキュメントが閉じます。(図 3)

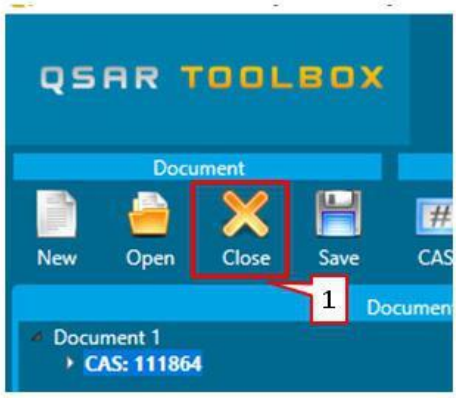


図 3

確認メッセージが表示され Yes （1）をクリックするとドキュメントが閉じられます（図 4）。

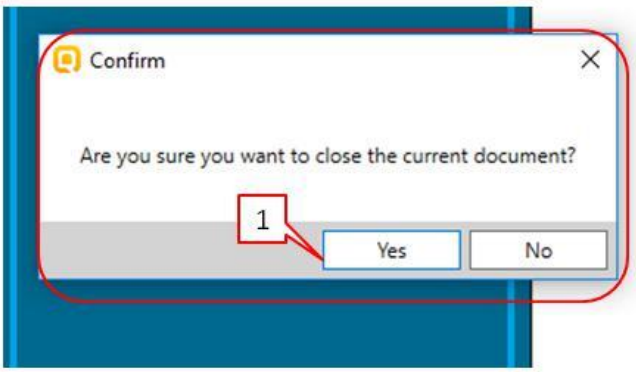


図 4

- **Open a document**（ドキュメントを開く）

* .tb4 ファイルを開くには、 Open （1）をクリックします（図 5）

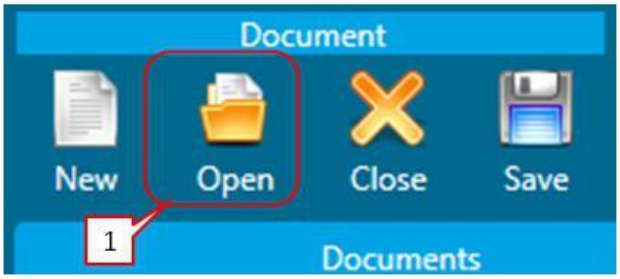


図 5

選択したドキュメント（1）が tb4 拡張子で表示されます（2）。次に、 Open （4）をクリックします（図 6）

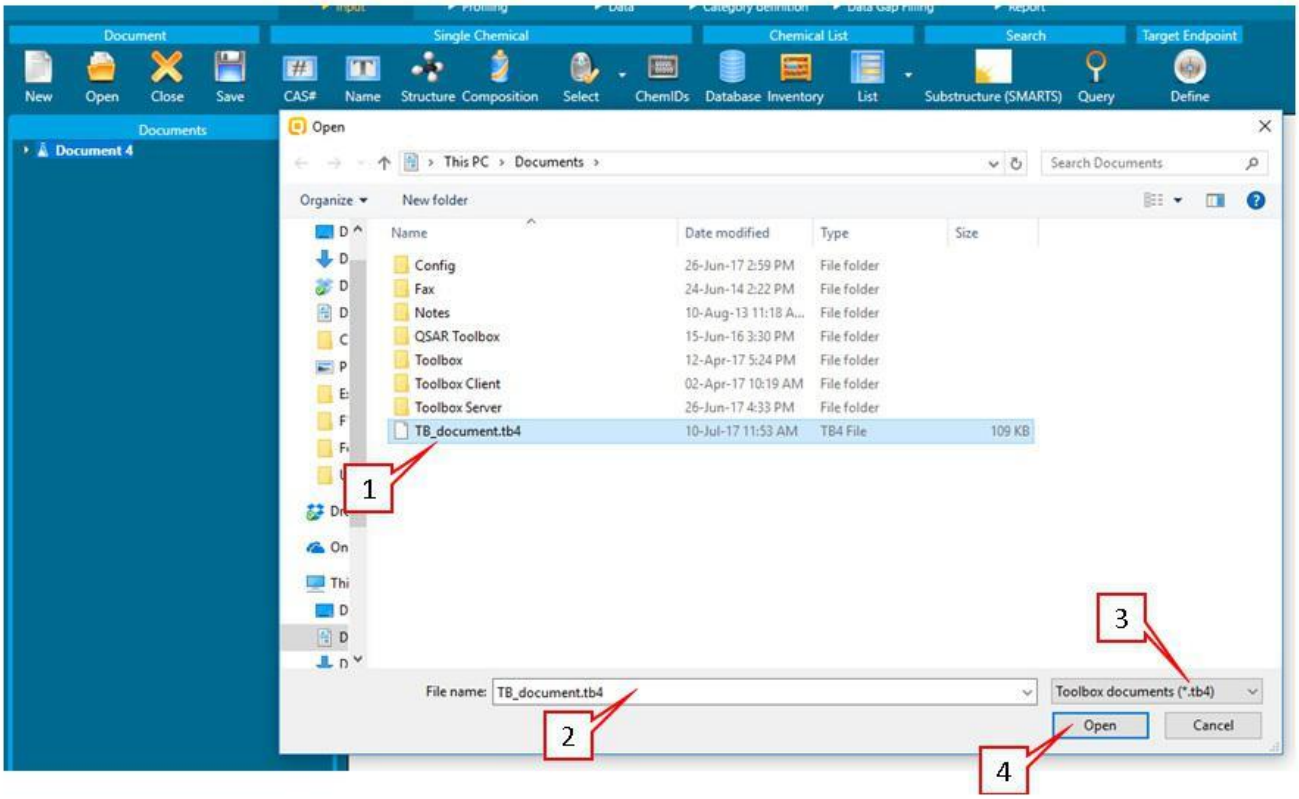


図 6

確認メッセージが表示され（1）、ファイルの内容が Document セクション（2）に表示されます。（図 7）

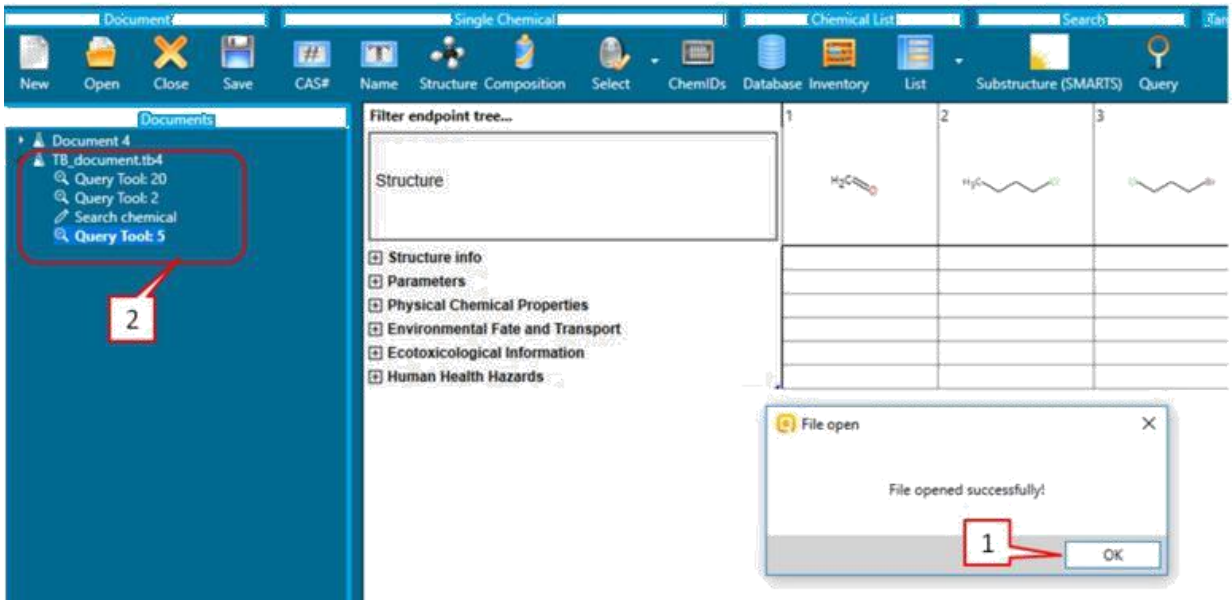


図 7

整合性チェックが実行されます。すべてのコンテンツが適切に開いている場合 Info:OK というメッセージがポップアップウィンドウに表示されます。そうでない場合、問題は同じポップアップウィンドウに表示されます。（図 8）

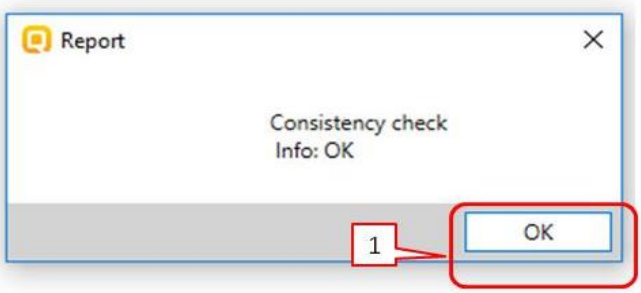


図 8

- **Save a document**（ドキュメントを保存する）

  Save (1) をクリックします（図 9）。

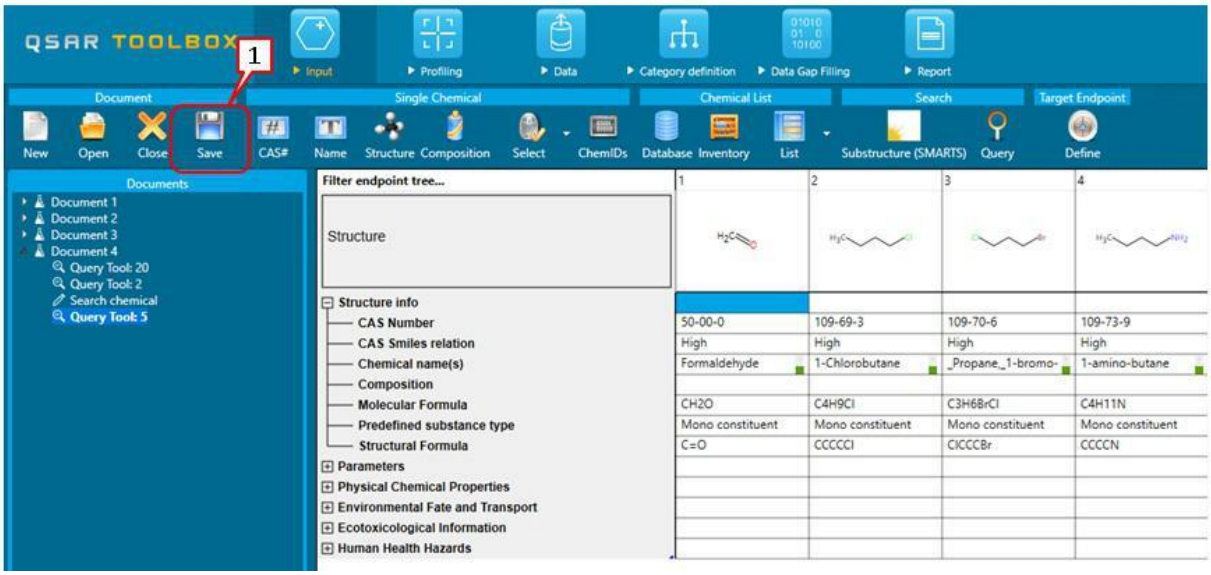


図 9

ファイル名を入力し（1）、 Save（2）をクリックします（図 10）。

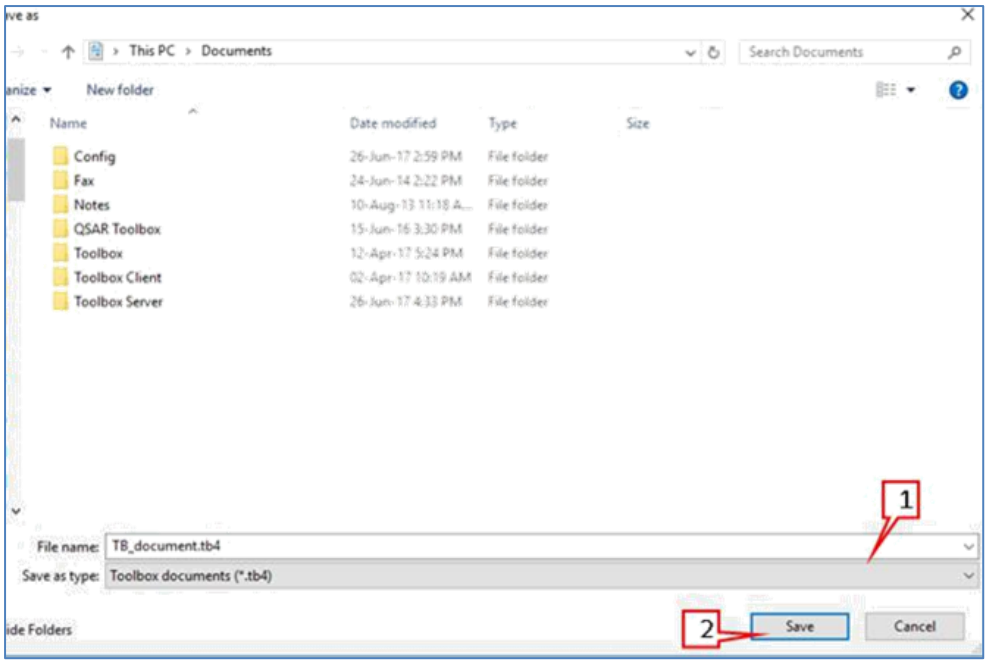


図 10

表示された確認メッセージで OK をクリックします（図 11）。

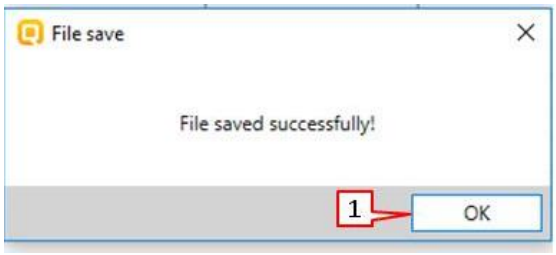


図 11

D.1.1. 単一の化学物質

単一の化学物質はこのシステムでは次のものを使って入力できます:

- CAS 登録番号
- 化学物質名（ワイルドカード可）
- 2D 構造（SMILES を入力するか手動で化学構造を描く）
- Composition のセクション（訳注：D.1.1.4.を指す）では、単一および多成分物質、ポリマーおよび UVCBs、例えば質量%、不純物、添加物の組成特性を入力する方法をご説明します。
- 既存のデータベースからの選択
- 既存のインベントリからの選択
- カスタムファイルからの選択

D.1.1.1. CAS 番号による化学物質の入力

CAS （ケミカル・アブストラクト・サービス） 番号で化学物質を入力するには、CAS＃ボタン（1）を押し、ハイフンを入れずに化学物質の CAS 番号を入力し（2）、Search（3）ボタンを押します（図 1）。CAS＃が複数の化学物質にかかっている場合には、複数の化学物質の ID を検索することができます。適切な化学構造（4）にチェックを入れてください。選択した選択肢は物質リスト（5）の上に表示され、スクロールバーを使用してその構造に起因するすべての名前を見ることができます（6）。

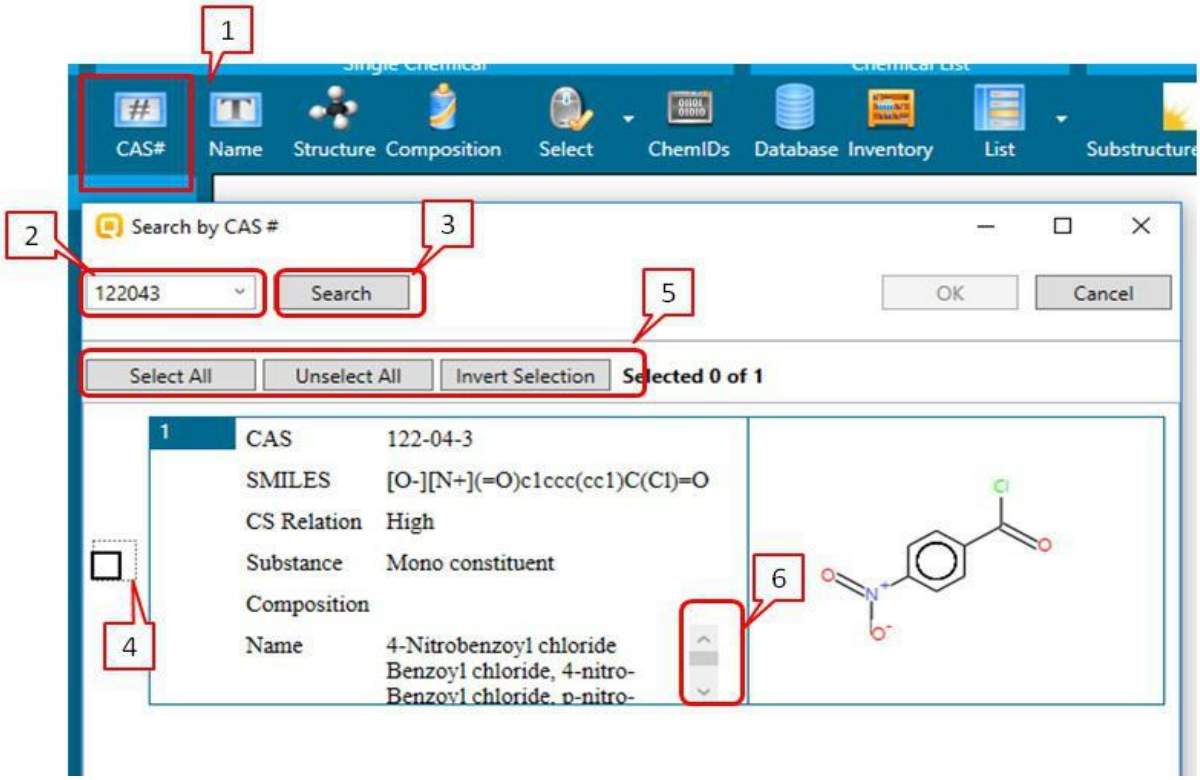


図 1

適切な化学構造を選択したら（1）、OK（2）をクリックします。(図 2)

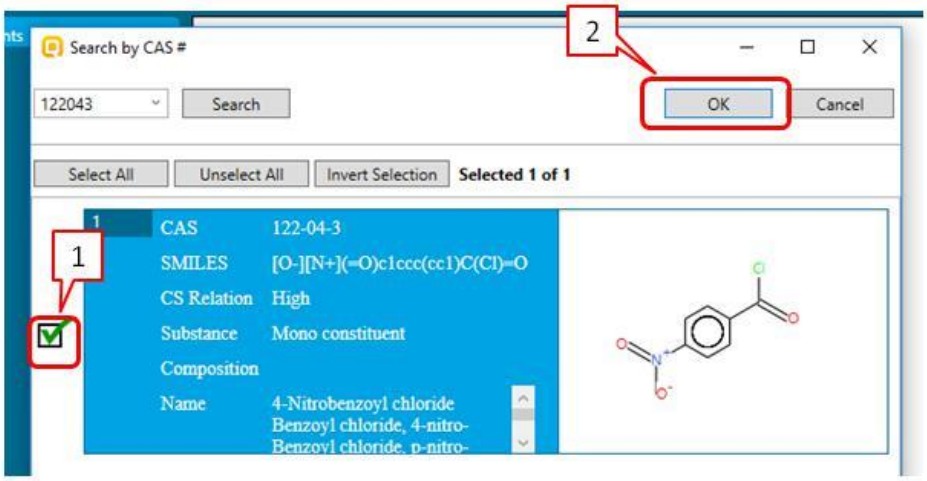


図 2

物質の種類も考慮されます。すなわち、一つの化学物質が別々の物質の種類に所属する場合（該当するのは ECHA Chem データベースのみ）、各物質の種類によって、独立した別の化学物質とみなされ、表示されます。

D.1.1.2. 化学名で化学物質を入力する

化学物質名を入力するには、 Name ボタン（1）をクリックし、検索オプション（2）を選択し、目的の化学物質（またはその一部）の名前を書き込んでから（3）、Search（4）をクリックします 。検索条件に一致するすべての化学物質がリストされます。その隣にあるボックス(5)にチェックを入れるか、または選択ボタンを使用して、関心のあるものを選択してください（6）。

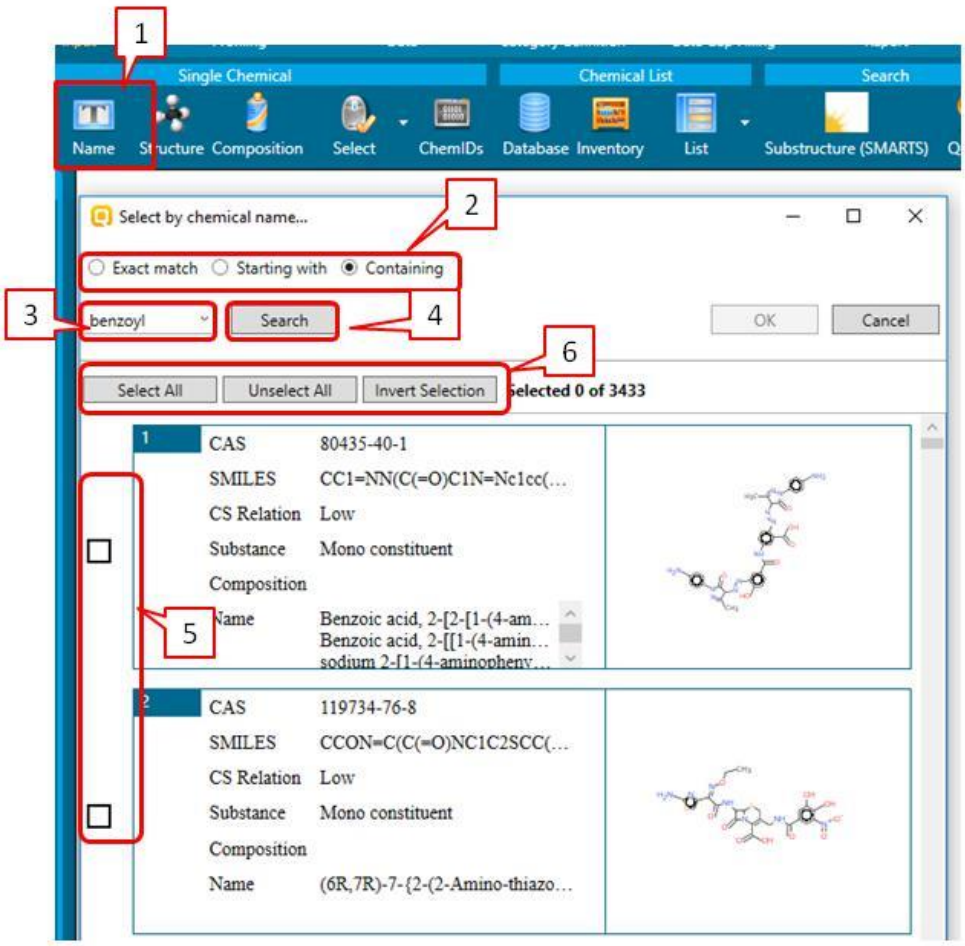


図 1

物質名検索の入力欄（3）には最大 192 文字を入力することができます。（図 1）

D.1.1.3. 化学構造式による化学物質の入力

2 つの種類の化学構造が描画によって入力可能です：

- 単一の物質構造
- 混合物

D.1.1.3.1. 単一構造の入力

原子同士の結合は、化学物質の 2D 構造を描いて制定することができます。Structure（化学構造）（1）を選択し、Structure Drawing（化学構造描画） ウィンドウ（3）内で対象化学物質の構造を描きます。OK（4）（図 1）をクリックして完全な図面を確認してください。

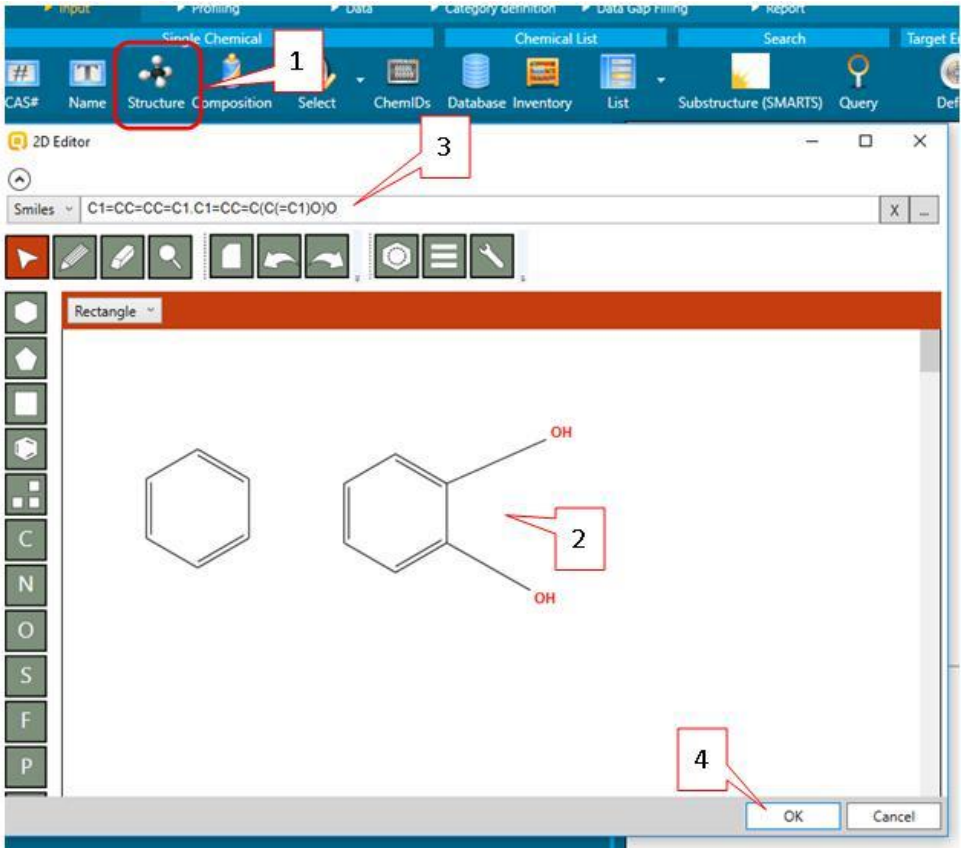


図 1

化学結合を入力するもう一つの方法として、化学物質の SMILES（Simplified Molecular Input Line Entry System、スマイルス記法）があります。対応するコードを SMILES 入力欄（2）に入力（書込みまたは貼付け）してください（図 1）。原子の結合が正しくつながっていれば、対応する 2D 構造、CAS 登録番号、化学名が表示されます。 (Toolbox で情報が入手できれば).

誤った結合が入力された場合、その誤った入力は SMILES の入力欄に赤色で表示され、構造は表示されません。SMILES 入力欄の下に簡単な説明テキストが表示されます（図 2）。

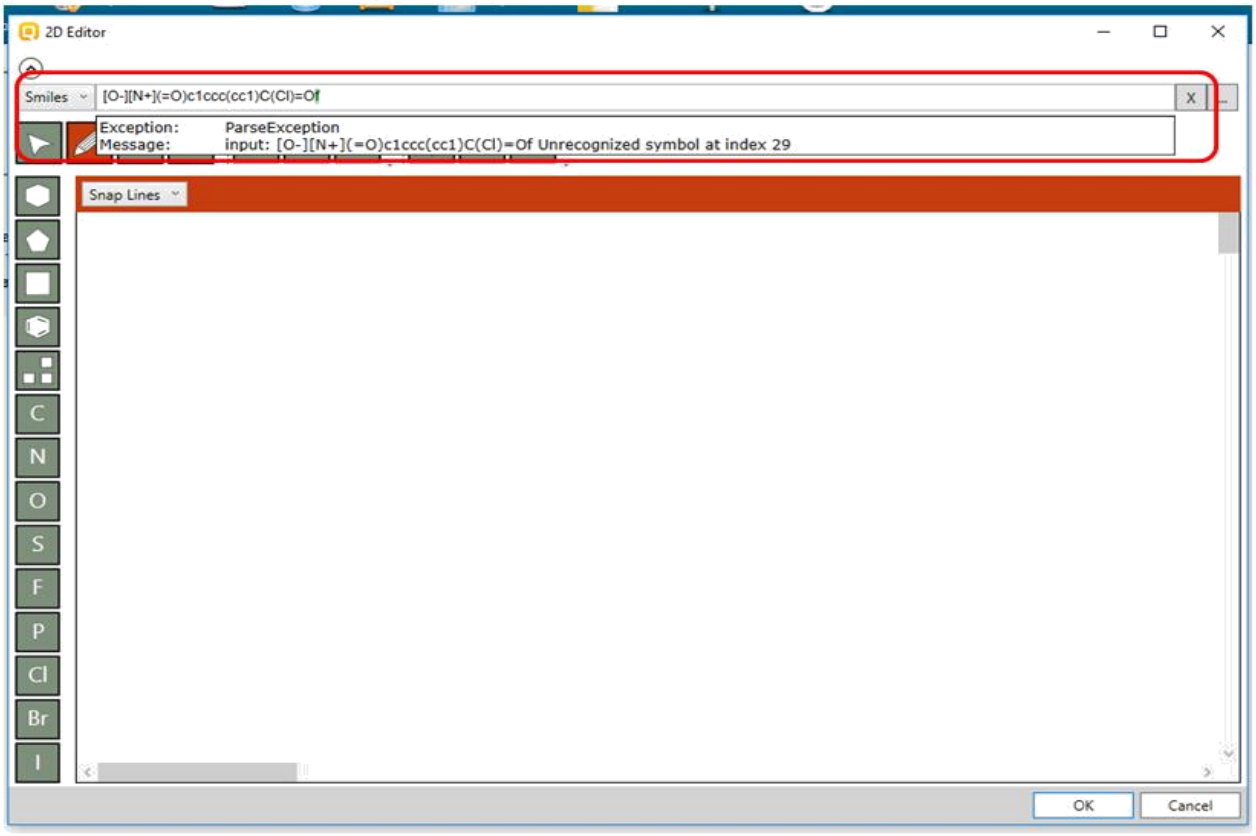


図 2

描画すると、システムは入力された SMILES のデータベースとインベントリを検索します。特定できた SMILES コードが表示されますが、複数の SMILES が特定された場合にはベンゼンの場合のように適切な化学物質を選択してください （図 3 の 1,2）。OK をクリックします（図 3）。

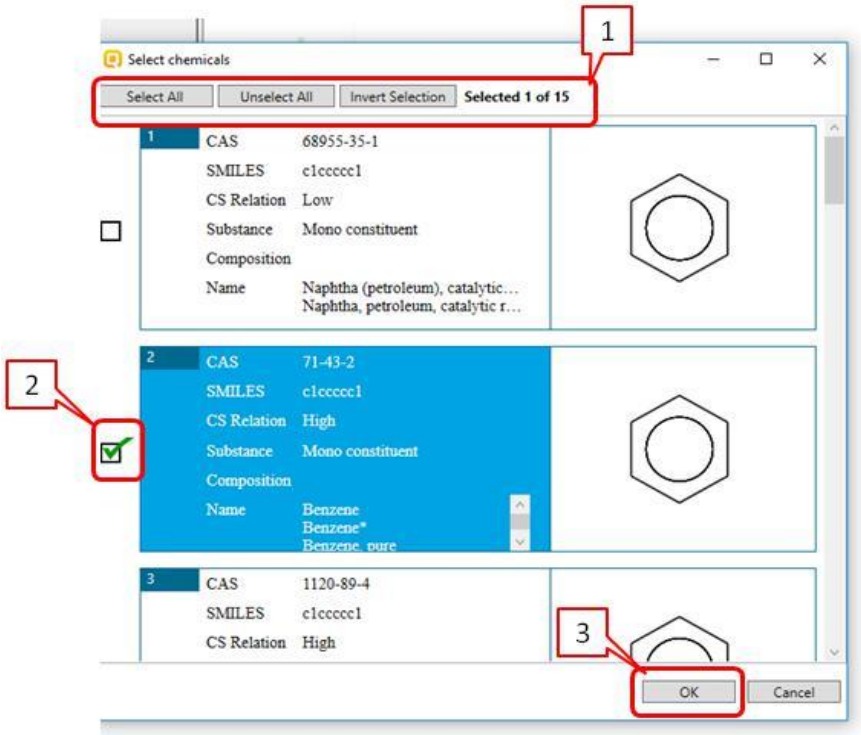


図 3

D.1.1.3.2. 混合物の入力

原子同士の結合は、化学物質の 2D 構造を描いて制定することができます。Structure（1）をクリックし、混合物の成分（2）を描くか、SMILES 入力欄（3）に SMILES を貼り付けます。混合物に別々の成分があれば、SMILES 表記ではドット文字で分割されています。手描き構造が完成したら OK（4）ボタンで確定してください。（図 1）

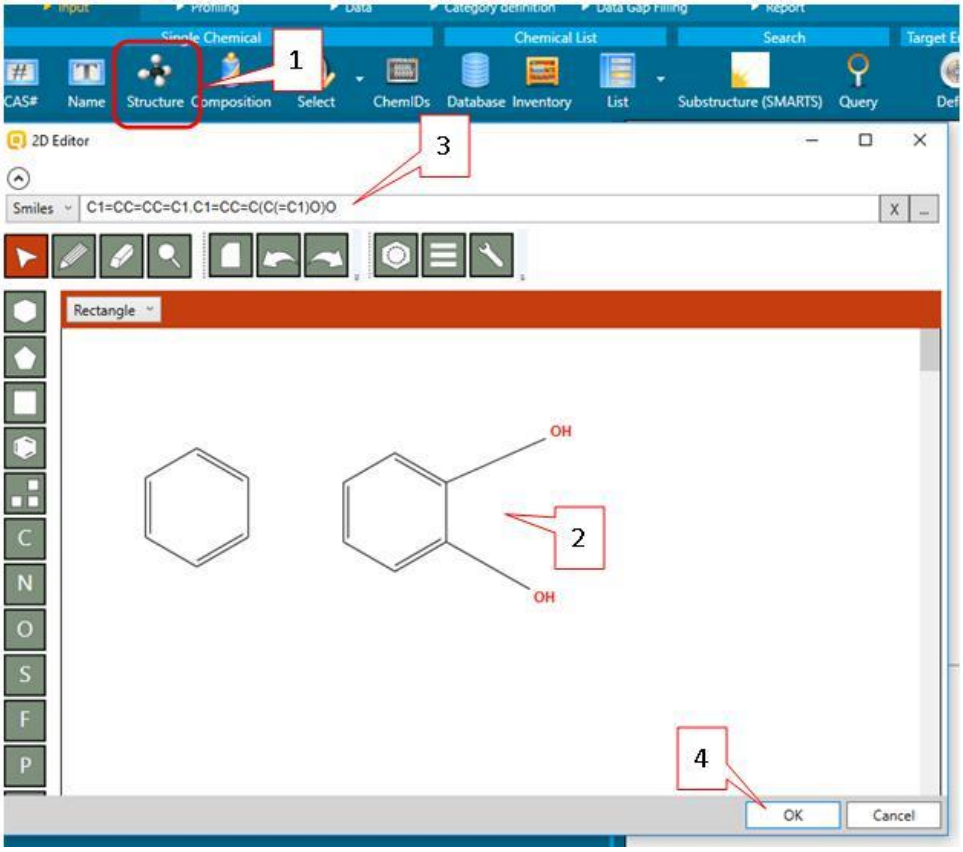


図 1

D.1.1.4. 化学組成による化学物質の入力

化学物質の組成式情報を含むオプションが提供されます。それを使うには、最初に Composition（化学組成式）（1）をクリックし、ドロップダウンメニュー（2）から化学物質の Type（タイプ）を選択します。（図 1）

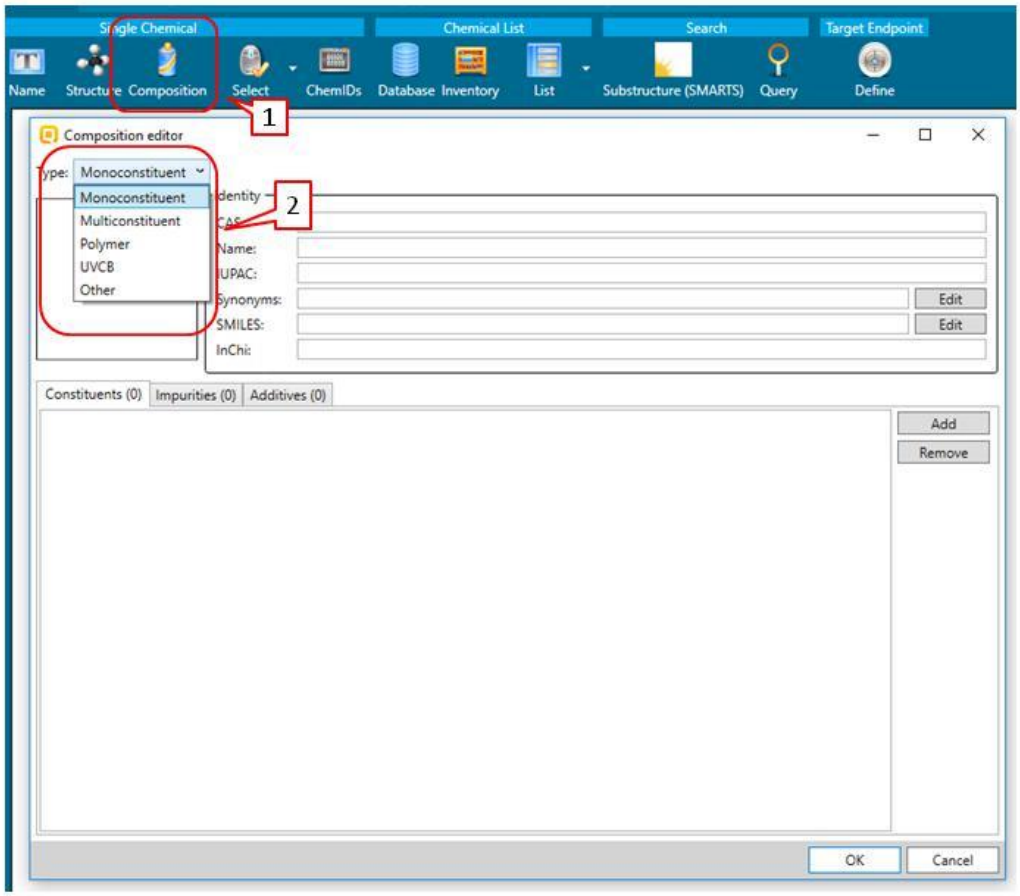


図 1

たとえば混合物の場合は、Multiconstituent（複数成分）を選択し、Constituent タブ（2）、Add（3）の順に選択します（図 2）。Edit（編集）（4）を押すと 2D エディタが開き、構造を描いたり SMILES(5)を貼り付けたりできます 2D エディタの詳細については、D 章 1.3.1. サブ構造（SMARTS）を参照してください。構成要素ごとに、Add を押して新しいダイアログを開く必要があります。

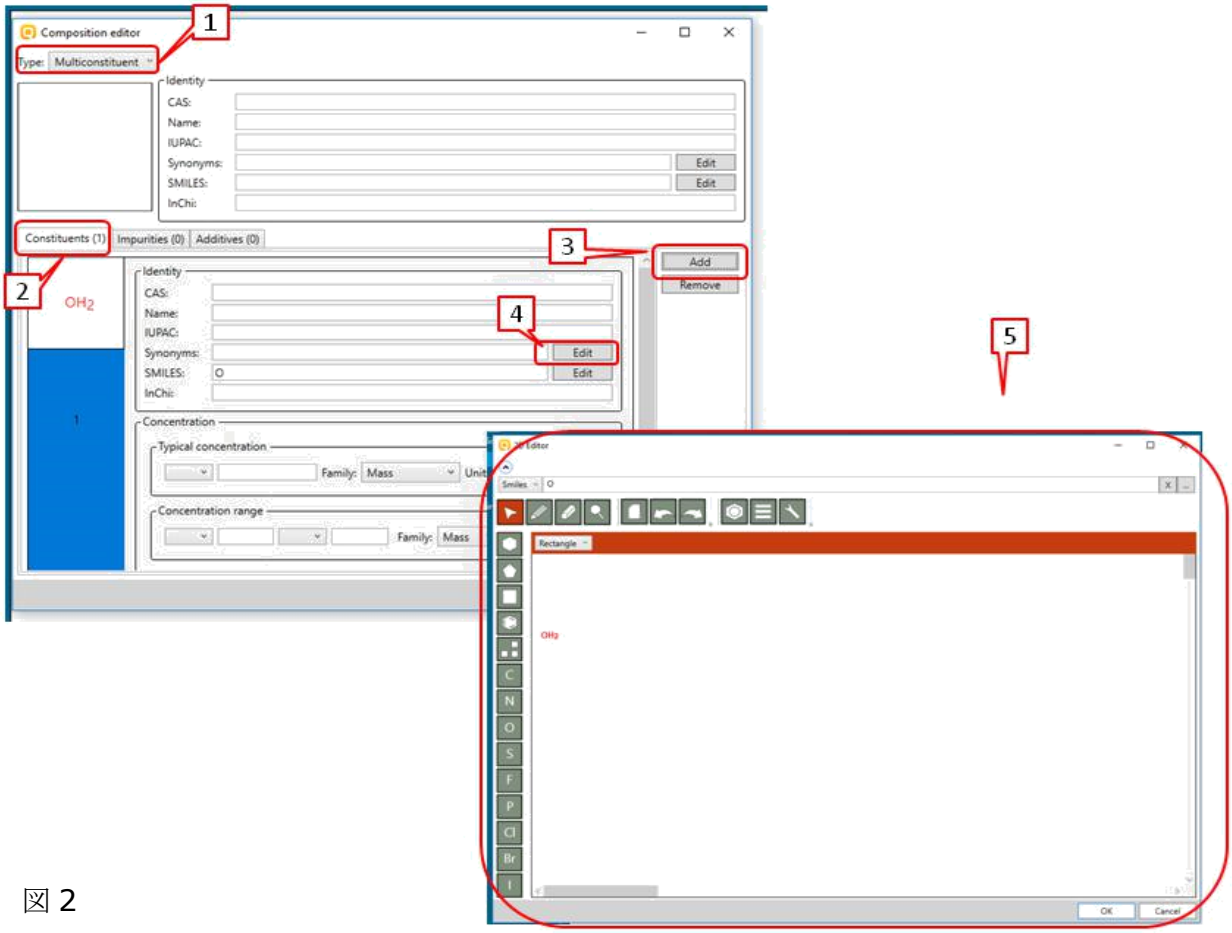


図 2

成分の数は（1）に示されています（図 3）。Additives（添加物）と Impurities（不純物）に関する定量・定性的情報を加えることができます（2）。書込みを削除するには、Remove（削除）を押します。Identity（識別情報）セクション（4）には、CAS 登録番号、Name（物質名）などの入力欄があり、ユーザが入力することができます。

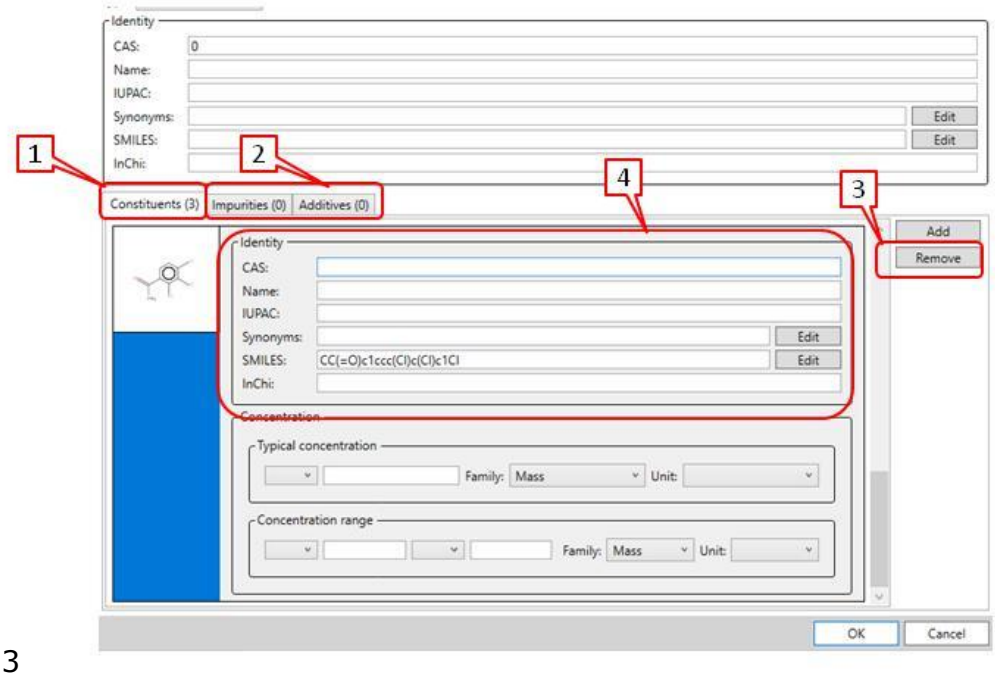


図 3

Concentration（濃度）セクションでは、各成分（1,2）の量を表すオプションがいくつかあります（図 4）。

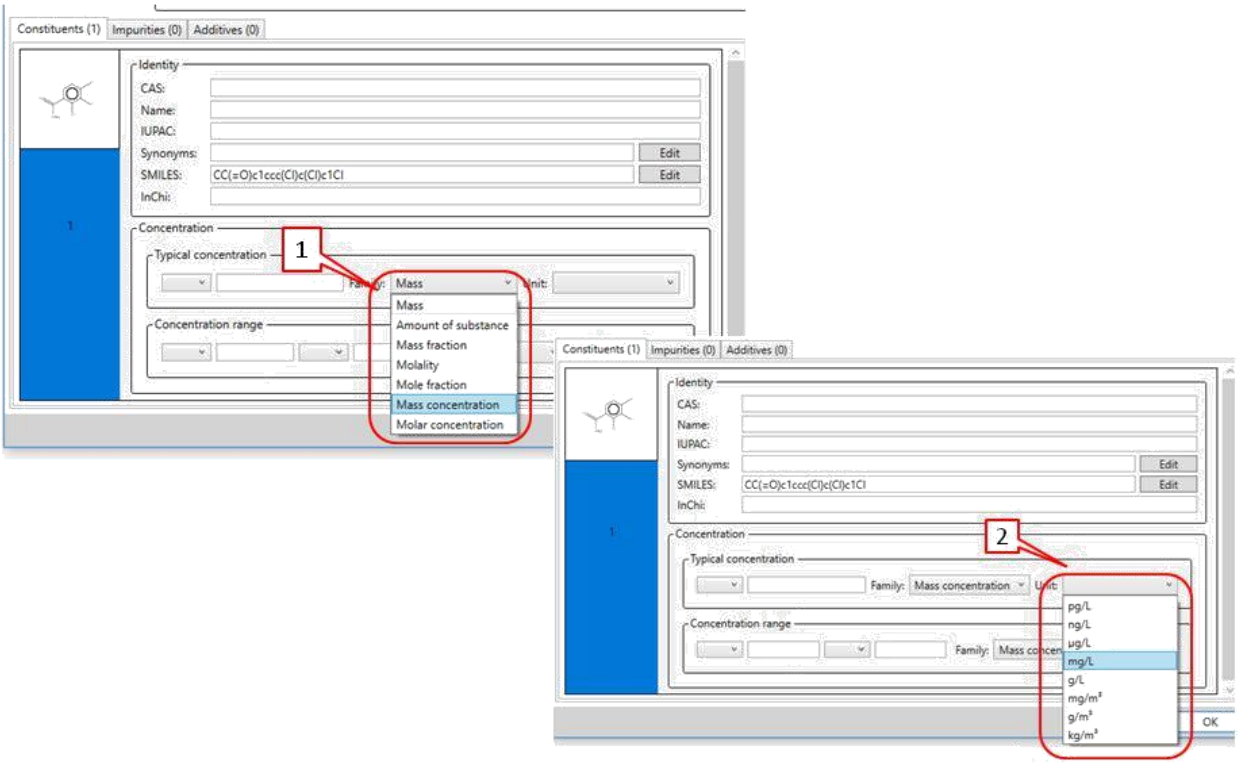


図 4

ドロップダウンメニューから符号修飾子を選択し、数字（1）を入力し、family（系統）と unit（単位）（2）を選択します（図 5）

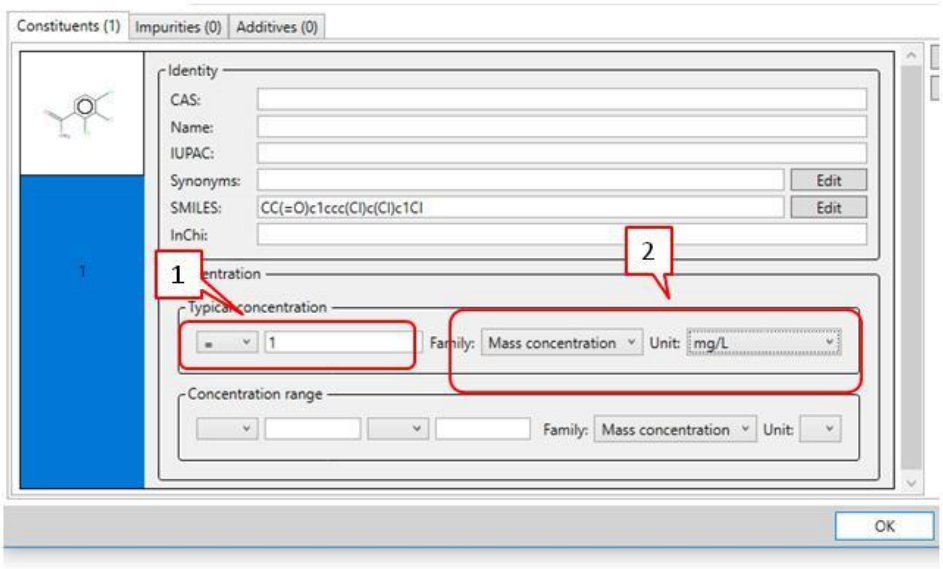


図 5

すべての数量が入力されたら OK を押します。(図 6)

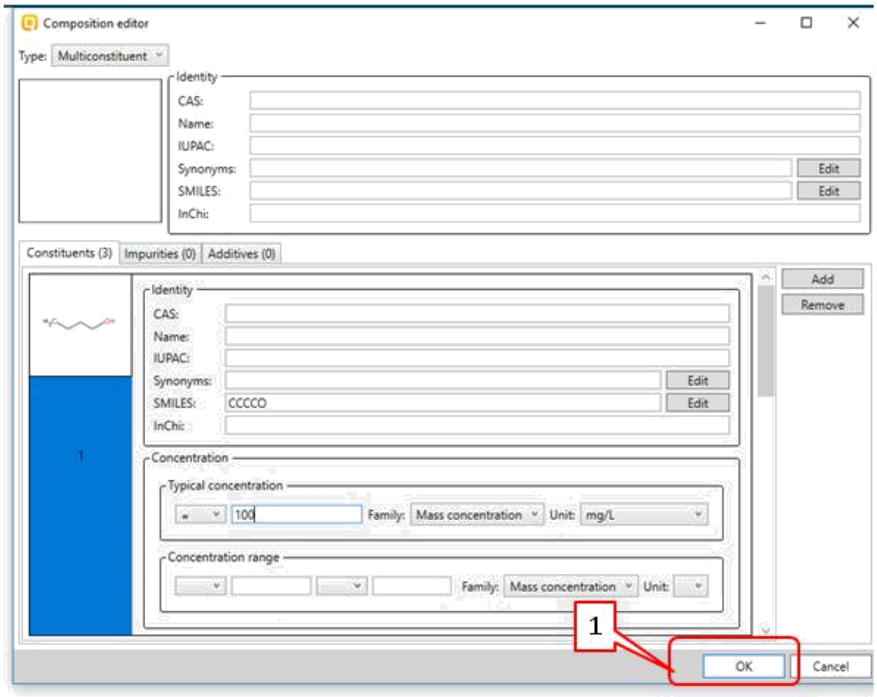


図 6

多成分の化学物質が**データ・マトリックス**（1）に表示されます（図 7）。

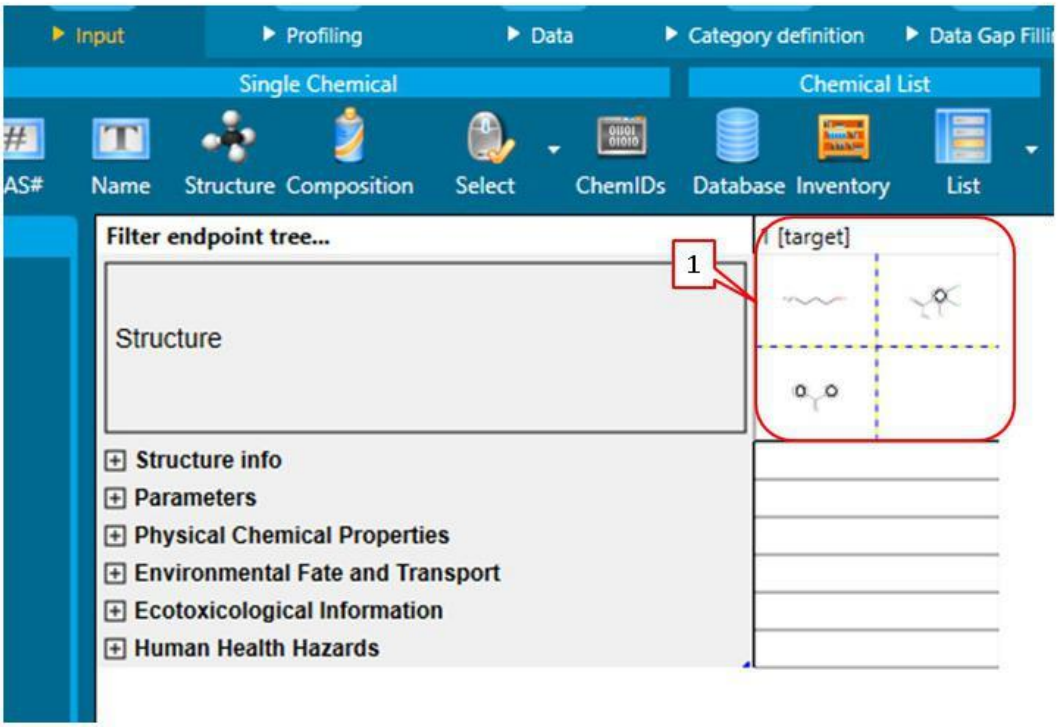


図 7

Structure info（1）の横にある+記号を展開し、セル（2）を展開すると、入力されたすべての化学情報が表示されます（3）（図 8）

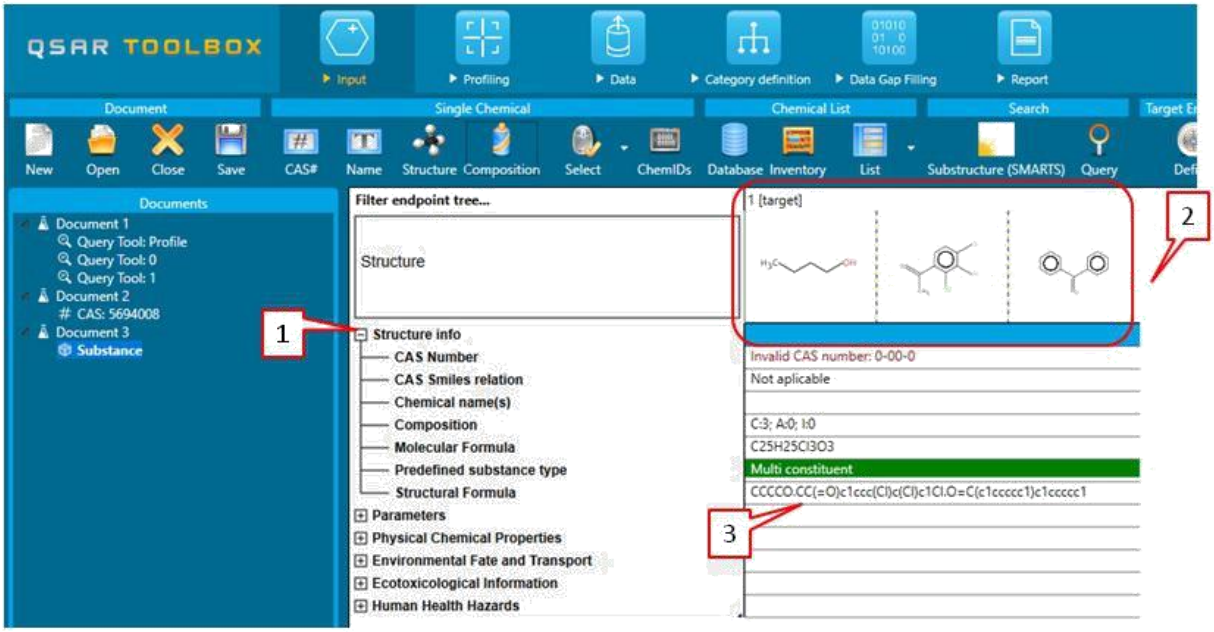


図 8

D.1.1.5. 最後に使用した化学物質の中から化学物質を入力する

Select アイコンのドロップダウンの▽印を押すと、化学物質を選択するオプションが表示されます（図 1）：

- last used chemicals（最後に使用した化学物質）
- database（データベース）
- inventory（インベントリ）
- file（ファイル）

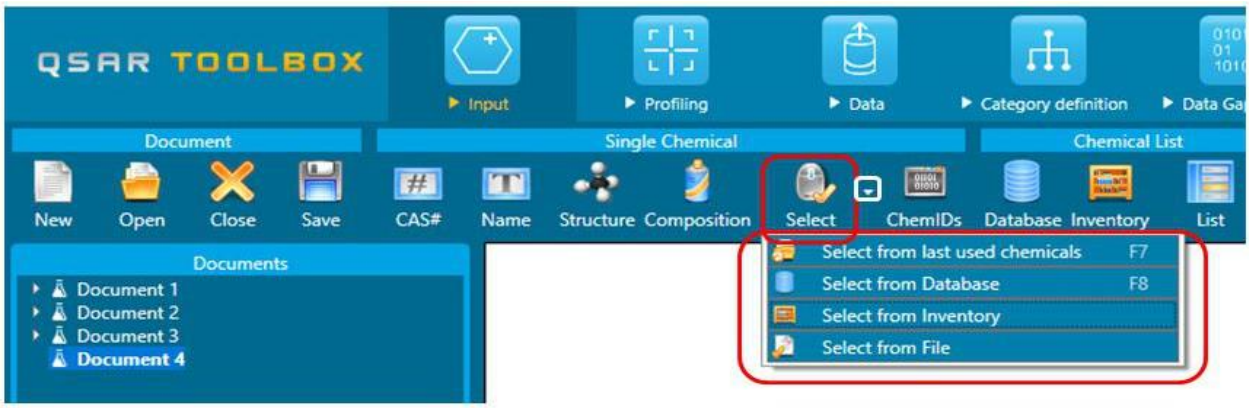


図 1

- ***last used chemicals***（最後に使用した化学物質から選択する）

最後に使用した化学物質から化学物質を選択する場合（F7 ショートカットでも可）、最後に使用した 100 種類以下の化学物質が一覧表示されます（図 2）。

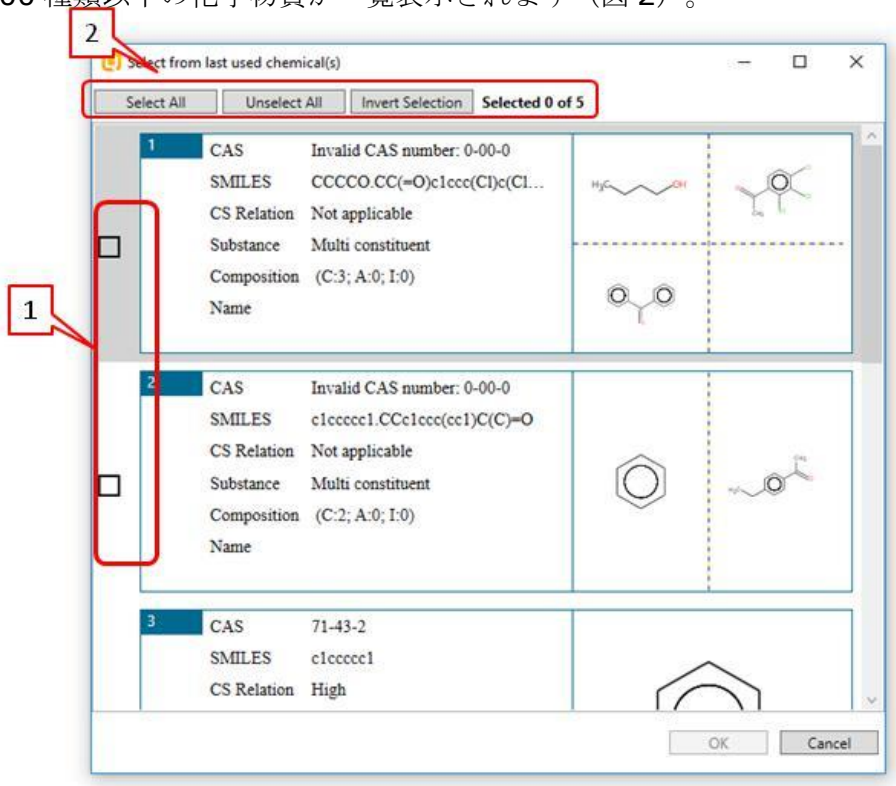


図 2

ユーザは以下のことを行うことができます：

- 目的の化学物質をチェックする/チェックを外す（1);
- select all ですべての化学物質を選択する（2);
- unselect all（すべての選択を解除）し、必要な化学物質だけを選択する（2)
- Invert selection で選択を反転する（2).

標的物質として使う化学物質にチェックを入れ（1）、OK（2）をクリックします。(図 2)

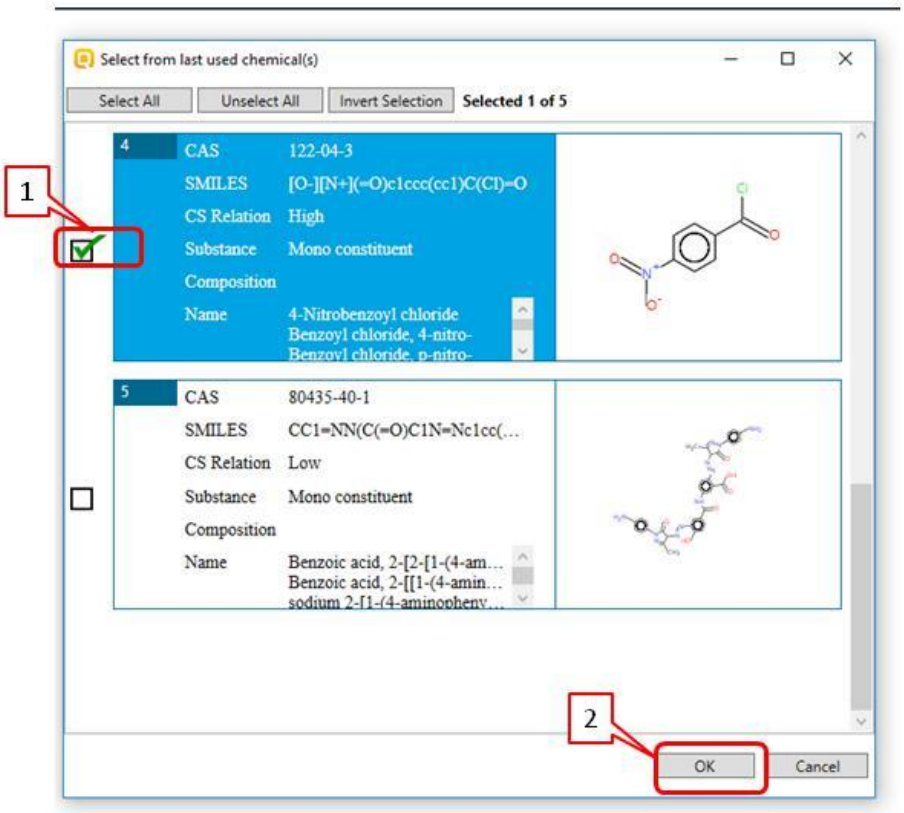


図 2

D.1.1.6. データベースから化学物質を入力する

Select アイコンのドロップダウンの▽印を押すと、化学物質を選択するオプションが表示されます（図 1）：

- last used chemicals（最後に使用した化学物質）
- database（データベース）
- inventory（インベントリ）
- file（ファイル）

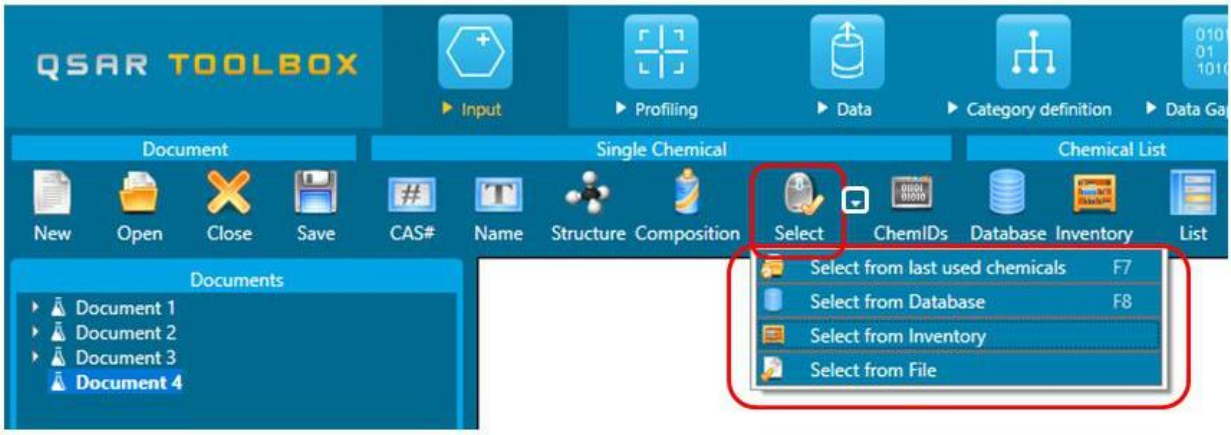


図 1

- ***Select from database*（データベースから選択する）**

Select from database（F8 ショートカット）オプションは、ツールボックス 4.1 に含まれるすべてのデータベースから化学物質を選択できます。ある特定のデータベースが選択されると（例えば Experimental pKa）（1）、About ボタンが表示され（2）、About を選択するとデータベースの簡単な説明が与えられます。OK ボタンをクリックすると、データベースからすべての化学物質が表示されます（3）。

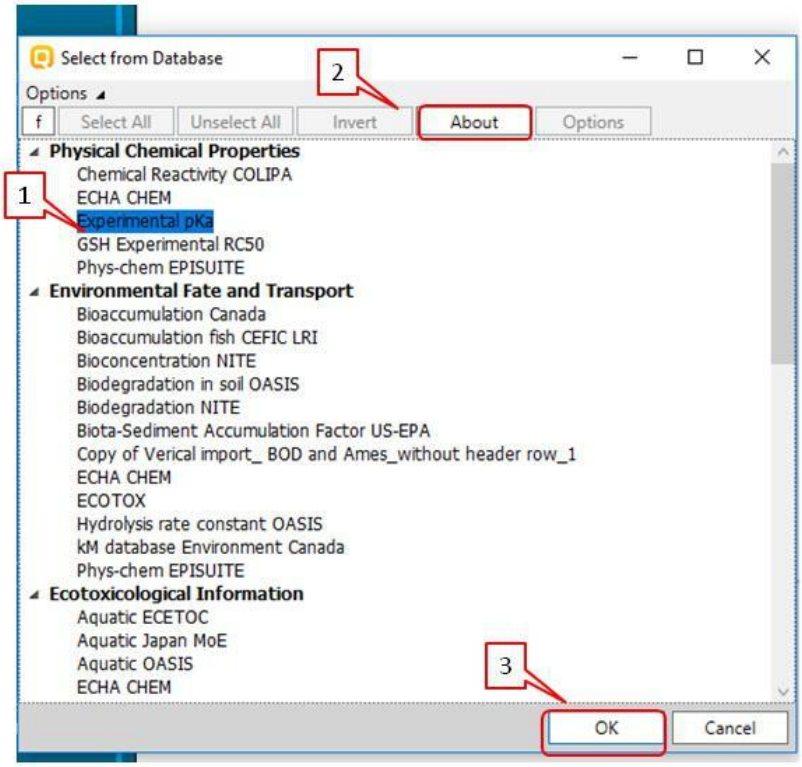


図 2

f ボタンをクリックしてフィルタ入力欄（2）を開きます（図 3）。bio など、キーワードを入力して、表示されたデータベースを絞り込みます。Close をクリックしてフィルタを閉じます。

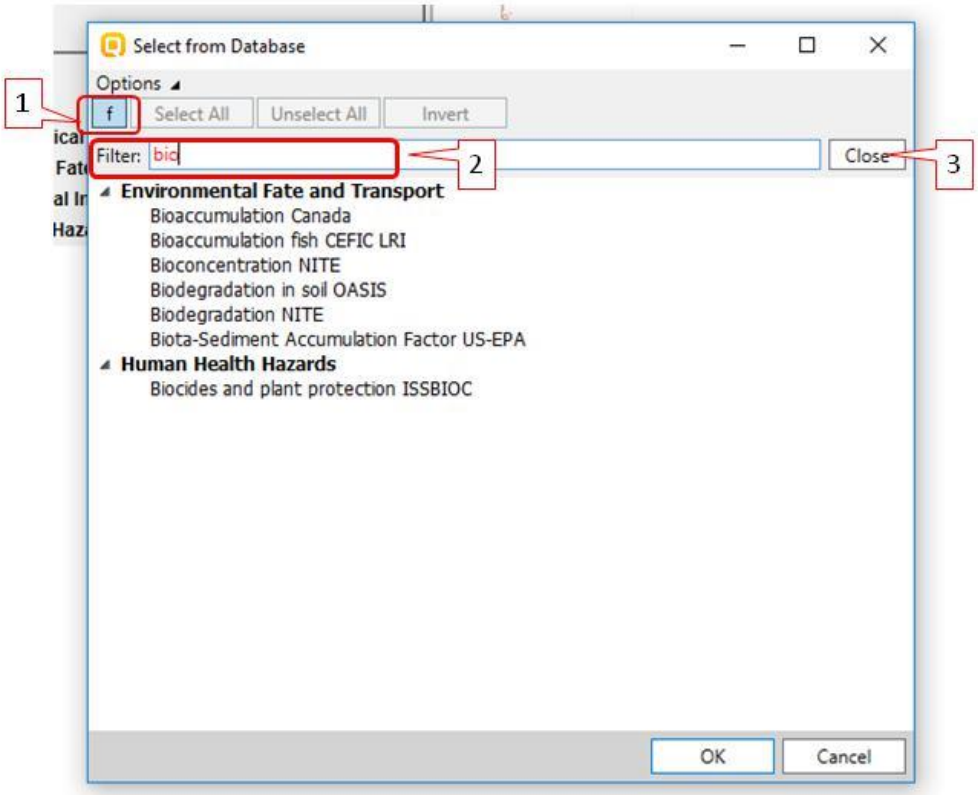


図 3

Options (1) のドロップダウンメニューをクリックして、データベースを Group （グループ化）または Sort （振り分け）します（2）（図 4）。このデータベースは有害性効果（3）に基づいて分類されています。グループ内では、Sorted by で Name を選ぶと名前で区分できます（4）。

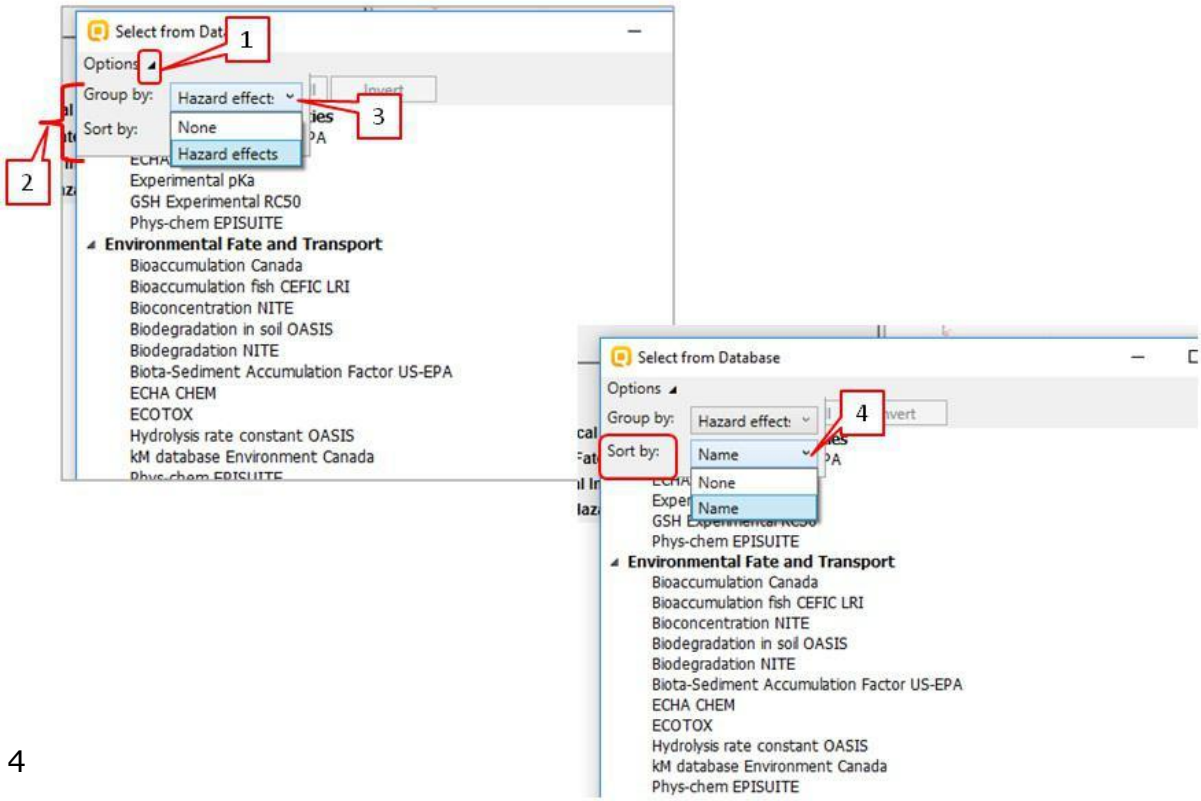


図 4

続いて、選択したデータベースのすべての化学物質を表すウィンドウが表示されます。標的物質にしようとする化学物質を選択し（1）、OK ボタン（2）で選択してください。（図 5）

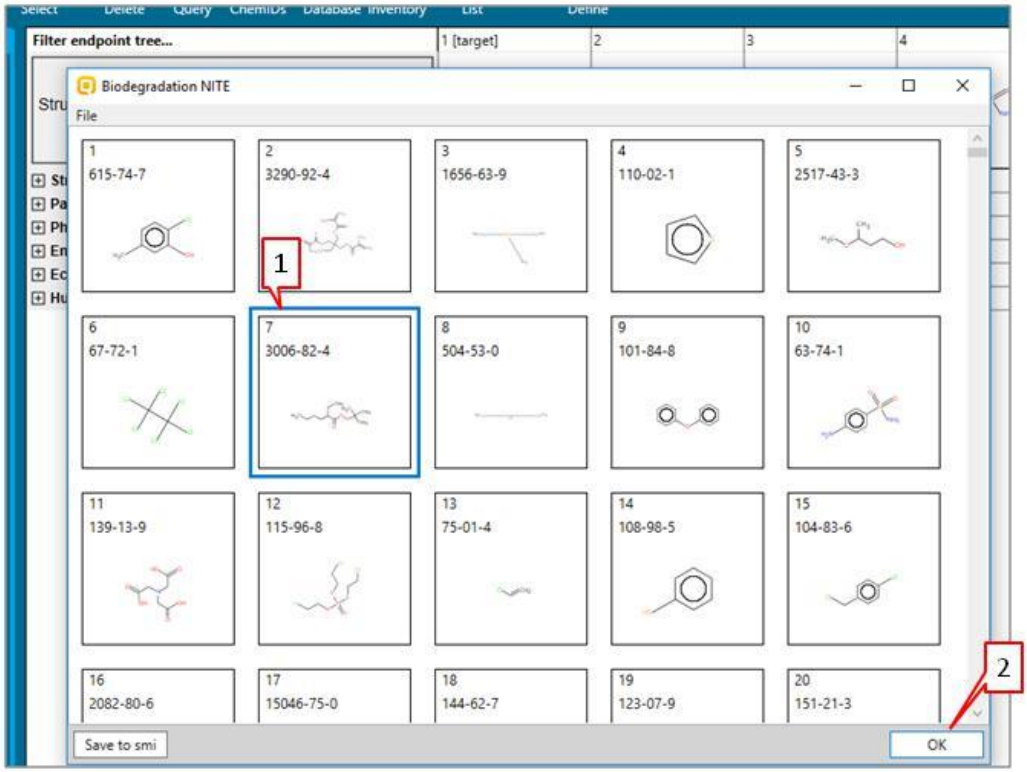


図 5

D.1.1.7. インベントリ・リストから化学物質を入力する

Select アイコンのドロップダウンの▽印を押すと、化学物質を選択するオプションが表示されます（図 1）：

- last used chemicals（最後に使用した化学物質）
- database（データベース）
- inventory（インベントリ）
- file（ファイル）

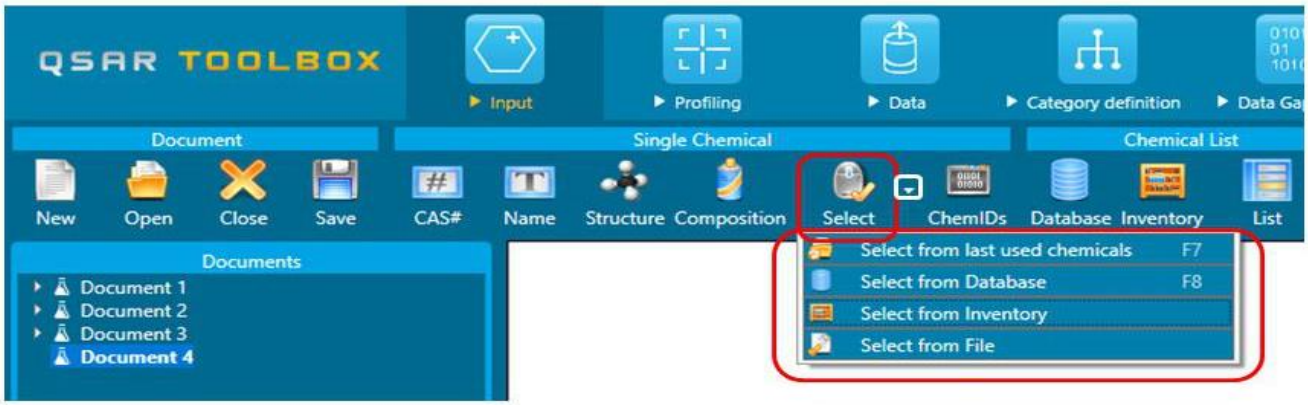


図 1

- ***Select from inventory*（インベントリから選択する）**
  インベントリ・リストから標的化学物質を選択します。利用可能なインベントリを含むリストが表示されます（図 2）。オプションボタン（1）が含まれています。

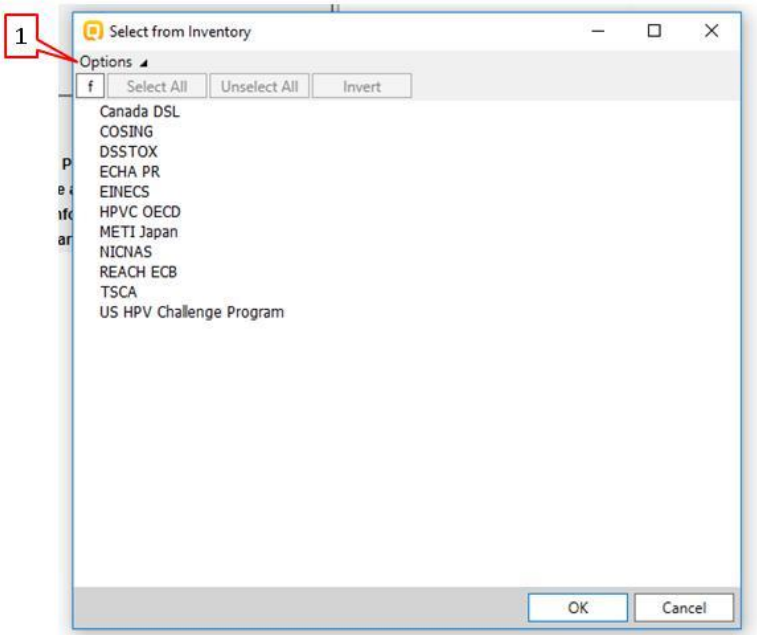


図 2

インベントリは Name（物質名）で区分できます。（図 3-１）

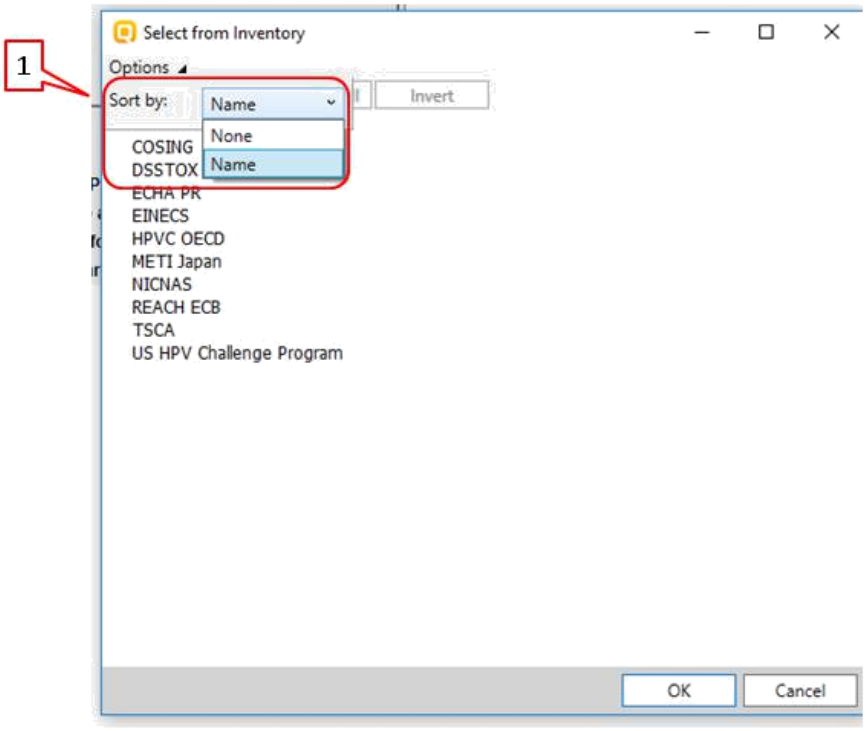


図 3

インベントリが選択されると（1)、 About ボタンが表示され、クリックされるとインベントリに関する情報が表示されます（図４）。About ウィンドウを閉じ、OK をクリックしてインベントリを表示します。

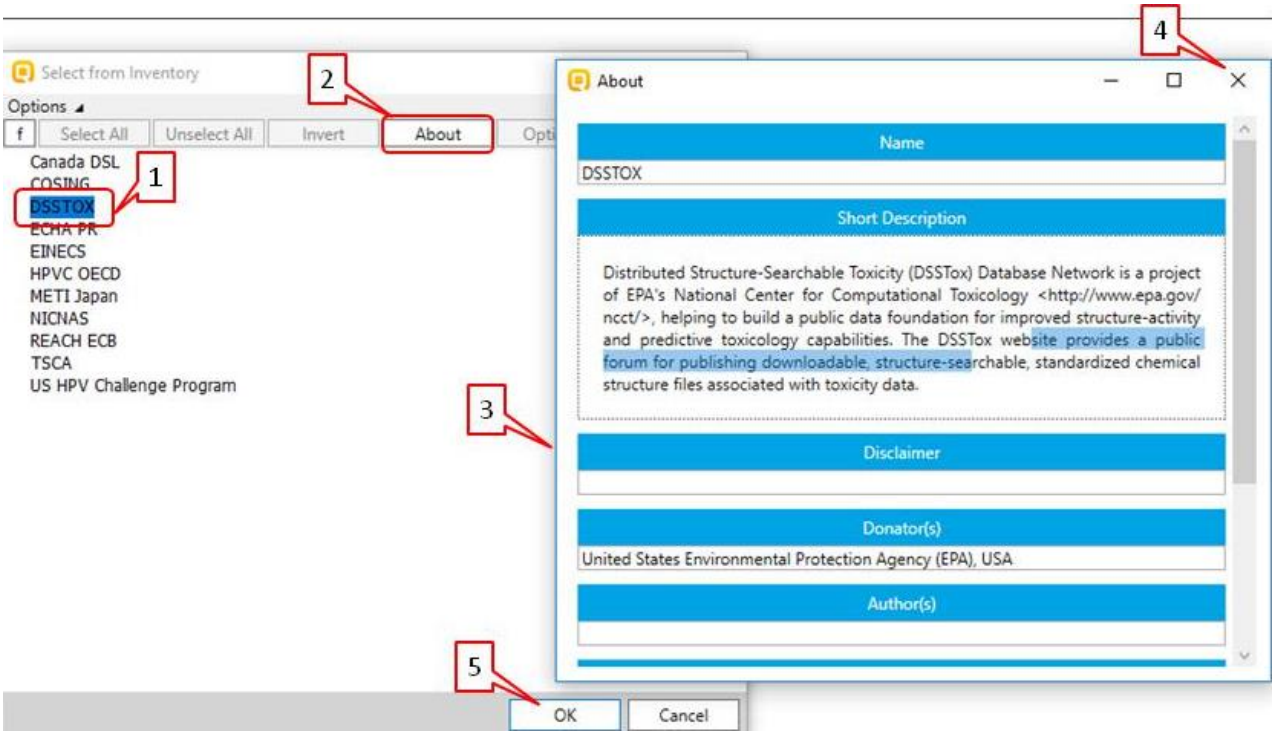


図 4

続いて、選択したインベントリからすべての化学物質を含むウィンドウが表示されます。標的物質にしようとする化学物質を選択し（1）、OK ボタン（2）を介して選択してください。（図 4）

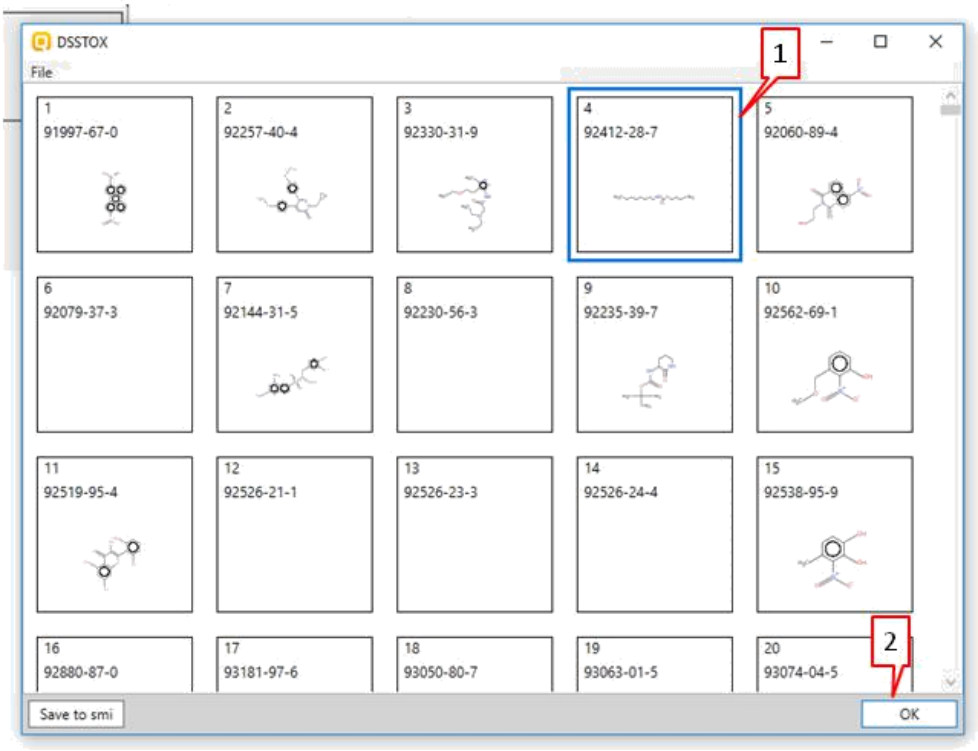


図 4

D.1.1.8. ファイルから化学物質を入力する

Select アイコンのドロップダウンの▽印を押すと、化学物質を選択するオプションが表示されます（図 1）：

- last used chemicals（最後に使用した化学物質）
- database（データベース）
- inventory（インベントリ）
- file（ファイル）

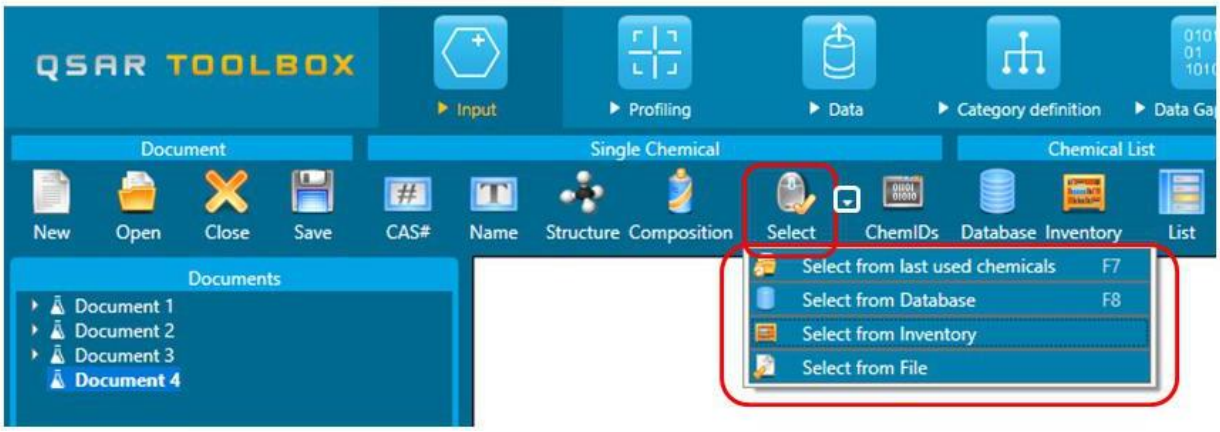


図 1

Select from file（ファイルから選択する）

標的化学物質は、ファイルに保存されている、すでに特定された原子結合のリストから選択できます。現在、原子結合のシステムのファイル形式でサポートされているのは、SDF と SMI です。SMILES のファイルには、TAB で区切られた CAS 番号と化学名の後に、対応する SMILES 文字列が続きます。これらのファイルの拡張子は* .smi です。　パスウェイ（経路）（1）、ファイルフォーマット（2）、ファイル名（3）を選択します（図 2）。

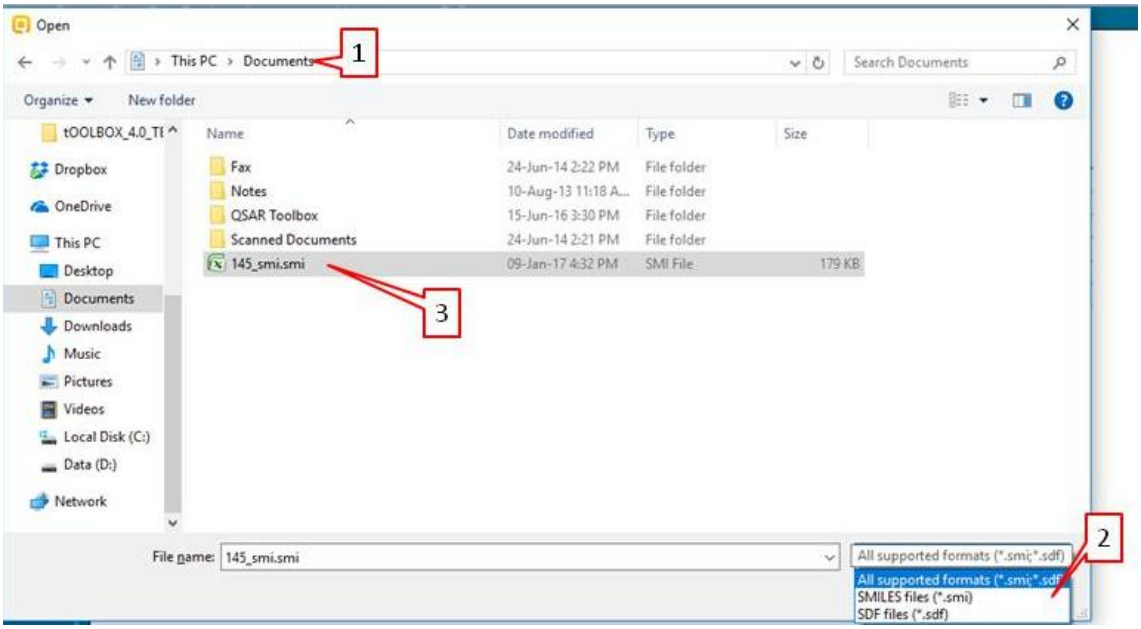


図 2

続いて、選択したファイルからすべての化学物質を表すウィンドウが表示されます（図 3）。標的物質にしようとする化学物質を選択し（1）、OK ボタン（2）で選択を確認してください。

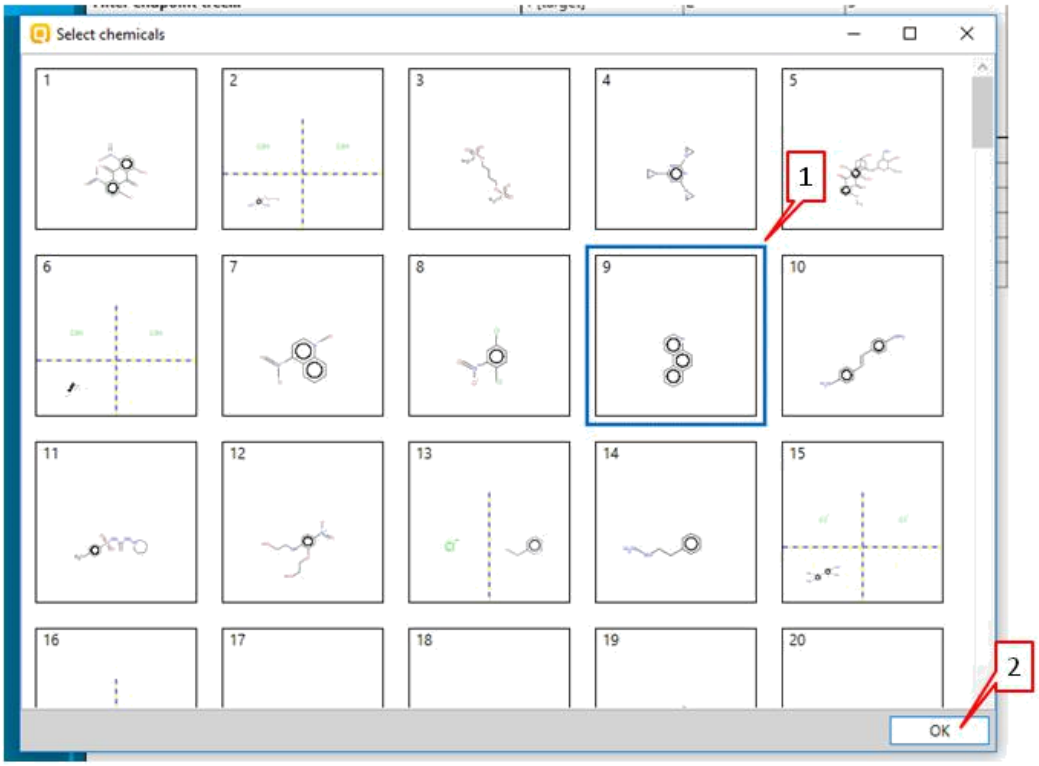


図 3

D.1.2. 化学物質のリスト

単一の化学物質を入力する代わりに、化学物質リストを**データ・マトリックス**に読み込ませ、その化学的・毒物学的特性、様々なエンドポイントの反応性、またはデータギャップ補完を実行することができます。

化学物質リストは以下の場所から読み込みができます：

- データベース
- インベントリ・リスト
- カスタムファイル

D.1.2.1. データベースを読み込む

既存のデータベースを読み込むには、Database（1）ボタンを押します。使用可能なデータベースを含むリストが表示されます。データベース（2）を選択し、About ボタン（データベースに関する詳細情報が必要な場合）（3）をクリックし、OK ボタン（4）で選択を確定してください（図１）。

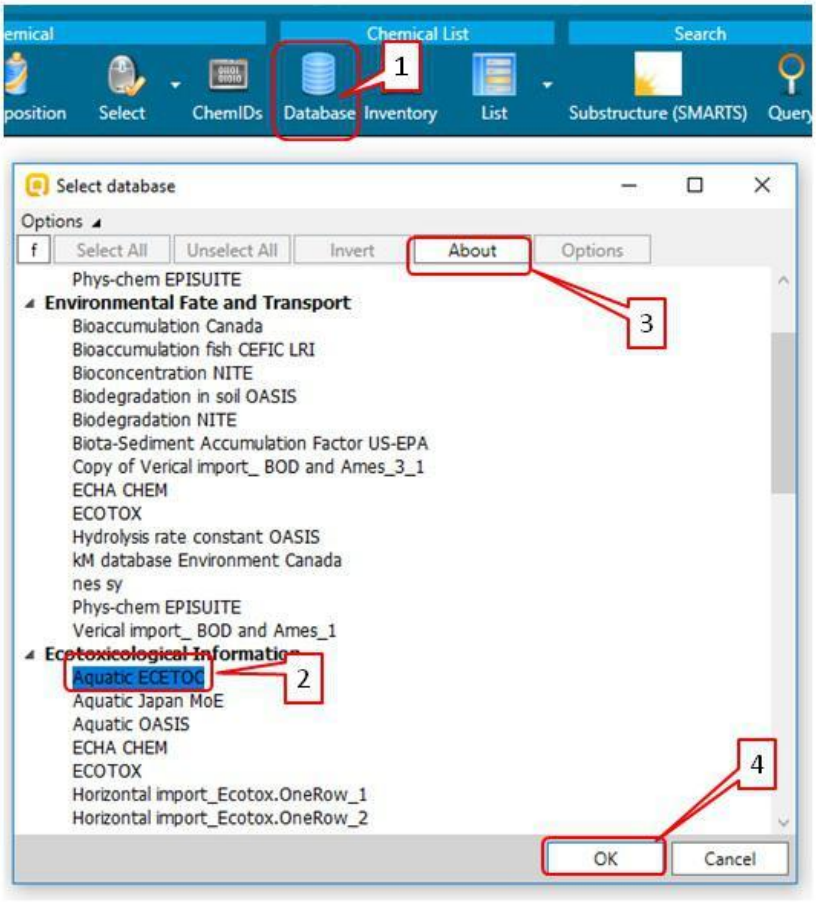


図 1

ユーザがデータベースを選択する際に役立つオプションは次の通りです（図 2）：

- group by hazard effect（有害性の効果によるグループ分け） (1);
- sort by name 物質名による区分 (2);

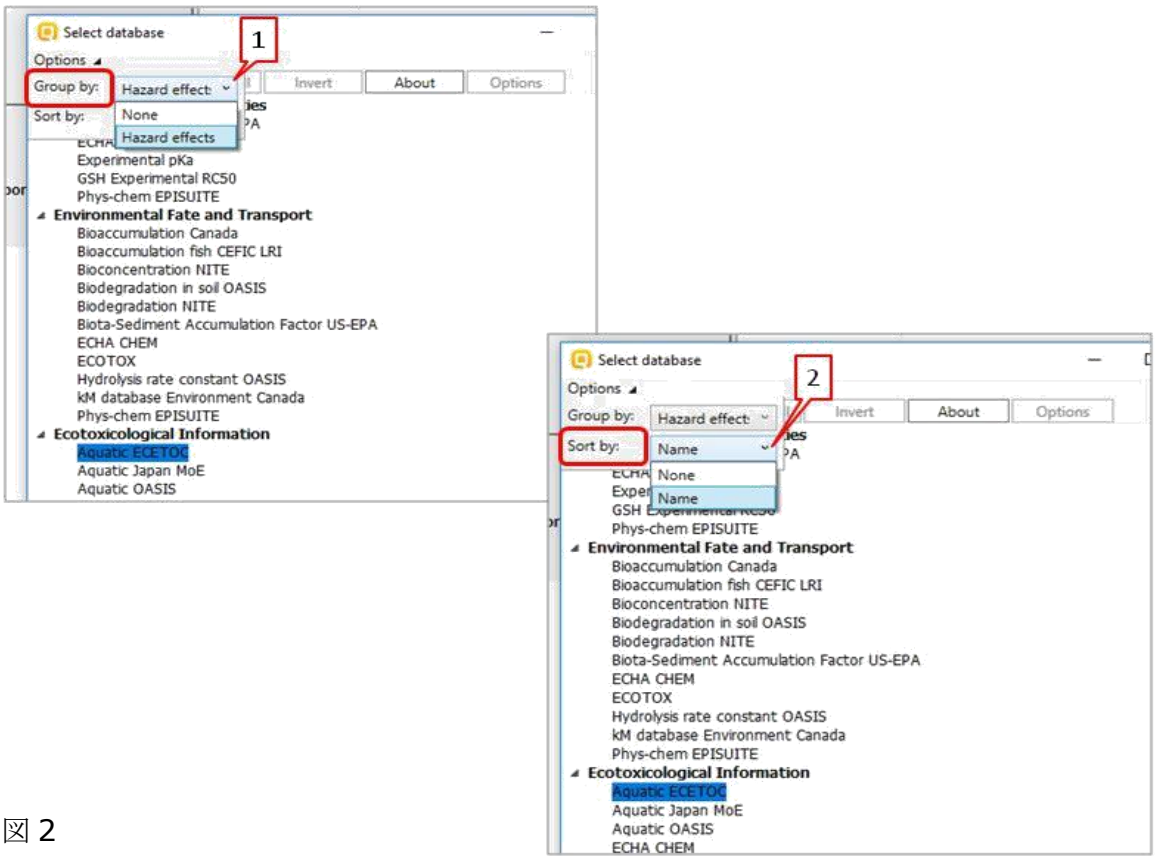


図 2

f（フィルタ）ボタン（1）をクリックすると新しい入力欄が開き（2）、目的のデータベースの名前の一部または全体を書き込むことができます（図 3）。データベース（3）を選択し、OK（4）をクリックします。

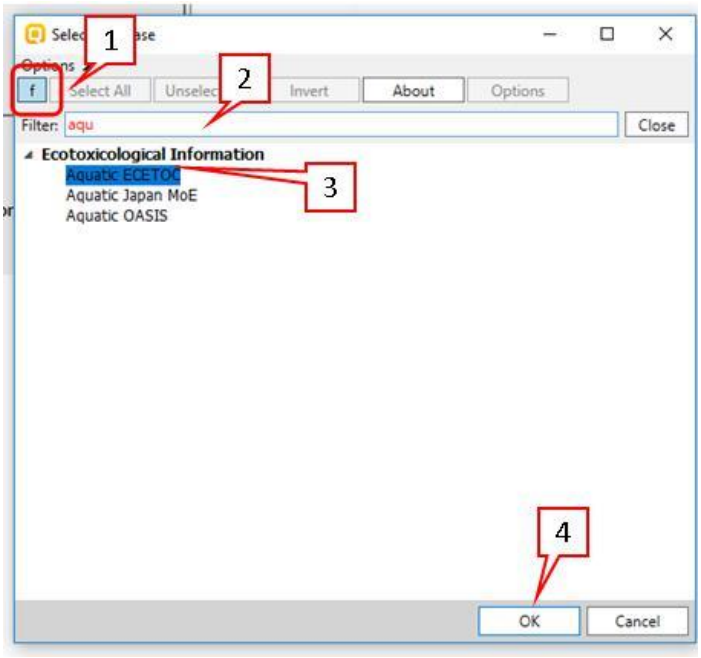


図 3

D.1.2.2. インベントリ・リストを読み込む

既存のインベントリを読み込むには、Inventory（1）ボタンを押します。利用可能なインベントリを含むリストが表示されます。標的物質を選択するインベントリ（2）を選択し、データベースに関する詳細情報が必要な場合は About ボタン（3）をクリックし、OK ボタン（4）で選択を確定してください。

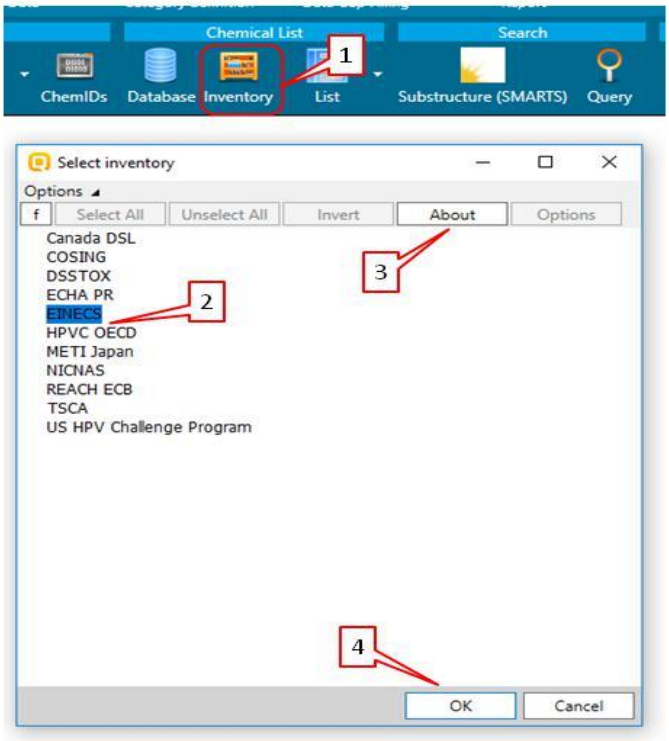


図 1

Sort by name（物質名による区分） (1,2)も導入されています (図 2)。

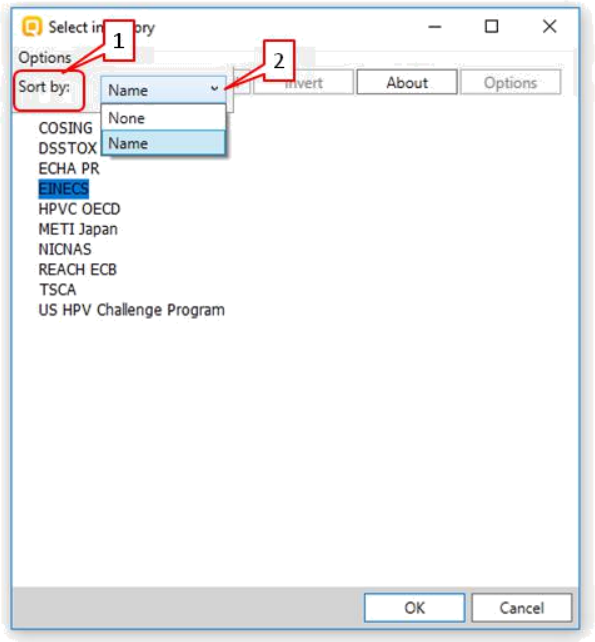


図 2

f（フィルタ）ボタン（1）をクリックすると新しい入力欄が開き（2）、目的のインベントリの名前の一部または全体を書き込むことができます（図 3）。名前にそのキーワードを持つインベントリだけがウィンドウに残ります（3）。

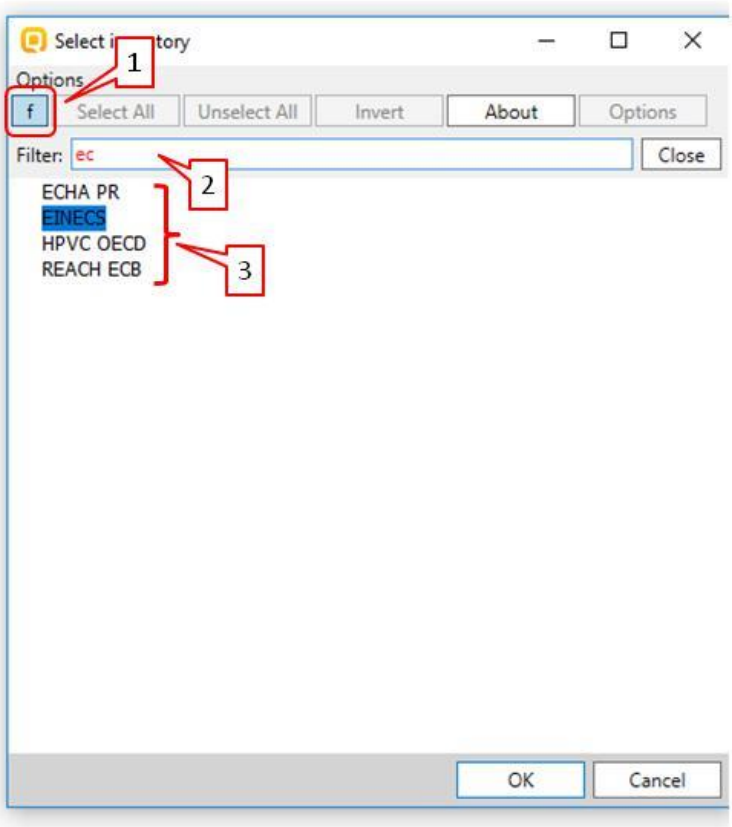


図 3

D.1.2.3. カスタムファイルを読み込む

ファイルに格納されている特定の原子結合を有するリストを介して、化学物質のセットを読み込むことができます。現在、原子結合システムファイルフォーマットでサポートされているのは、SDF、SMI、MOL です。SMILES のファイルには、タブで区切られた CAS 番号と化学名の後に、対応する SMILES 文字列が続きます。通常これらのファイルの拡張子は* .smi です。
List ボタン（1）をクリックしてファイル（2）を選択し、最後に Open（3）をクリックします。

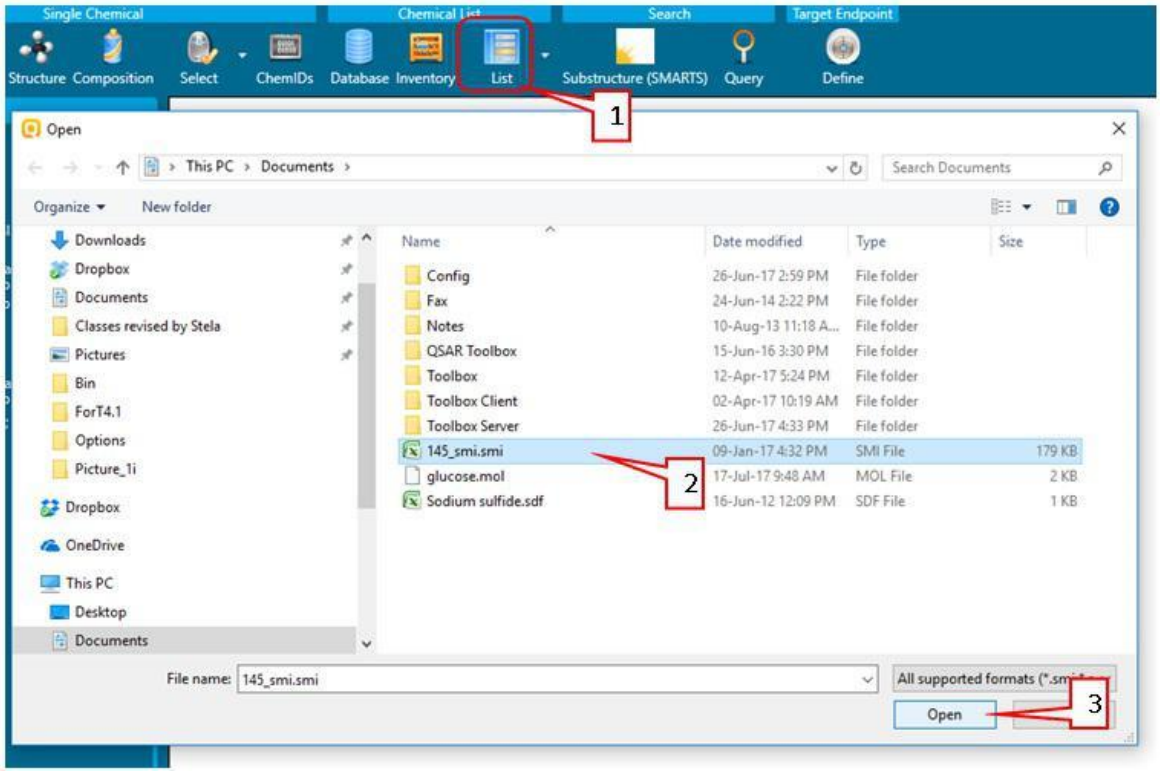


図 1

さらに、最後に使用したフォルダまたは Toolbox のサンプル例のファイルからファイルを読み込むオプションがあります（図 2）。

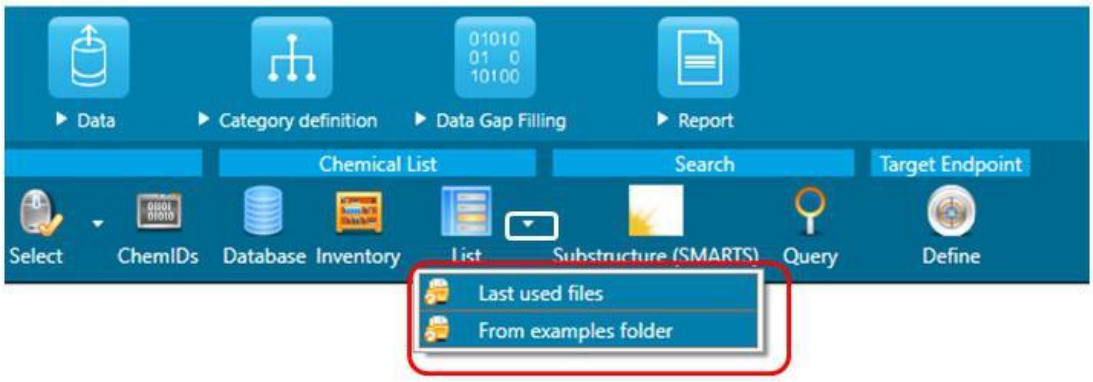


図 2
Last used files（最後に使用したファイル）（1）を選択し、開くファイルを選択して（2）、[OK]（3）をクリックします（図 3）。

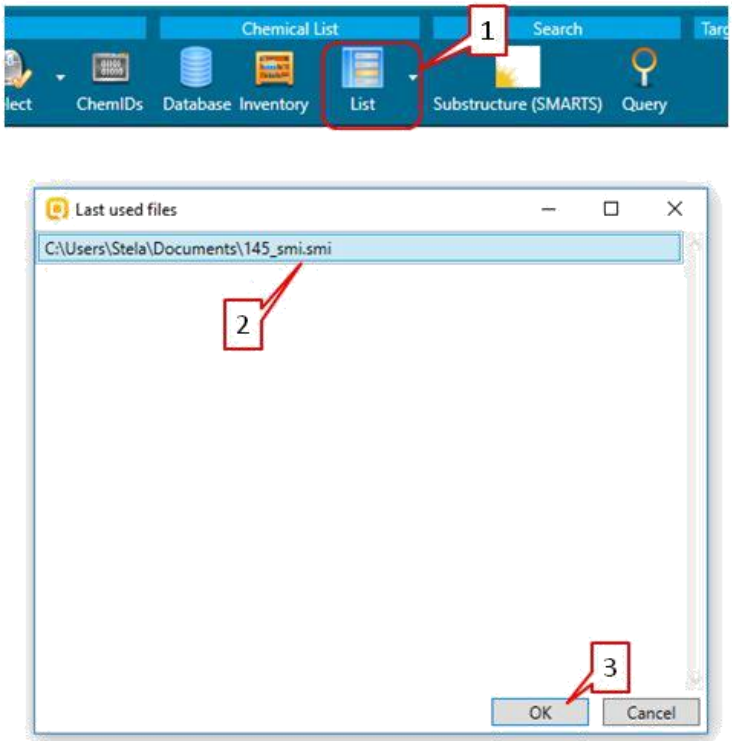


図 3

From Example folders（使用例フォルダから選ぶ）（図 2）から選択し、Toolbox の使用例フォルダを開きます（図 3）。ファイルを選択し（1）、Open をクリックします（2）（図 4）。

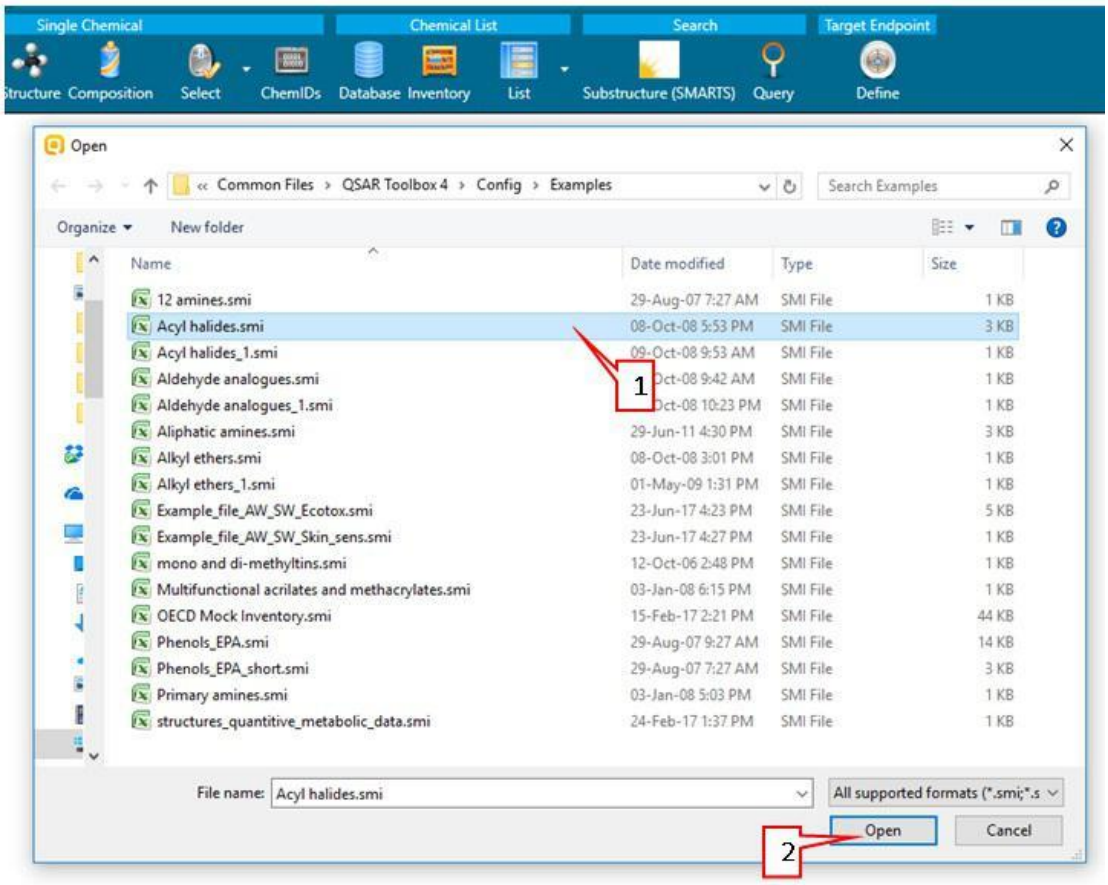


図 4

D.1.3. 検索のオプション機能

検索オプションには 2 つのタイプがあります：

- Substructure（部分構造） (SMARTs 記法による入力)
- クエリツール

D.1.3.1.Substructure（サブ構造） (SMARTS)

特定の化学物質を念頭に置いているわけではなく、むしろ同じ構造的特徴を含む化学物質のグループを探したいときは、Substructure（SMARTS）オプション（図 1）を選択することができます。

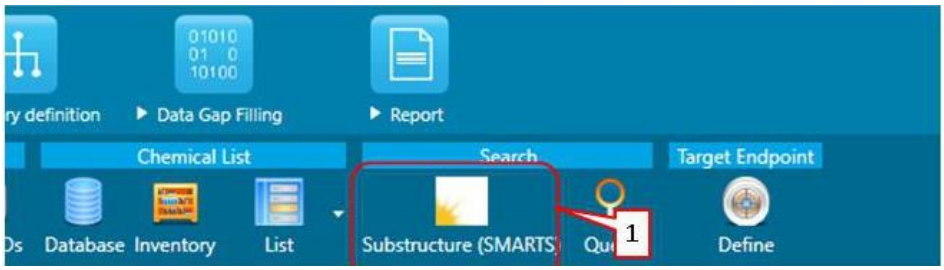


図 1

検索を開始する前に、少なくとも 1 つのデータベースかインベントリを選択する必要があります。選択していないとリマインダメッセージが表示されます（図 2）。

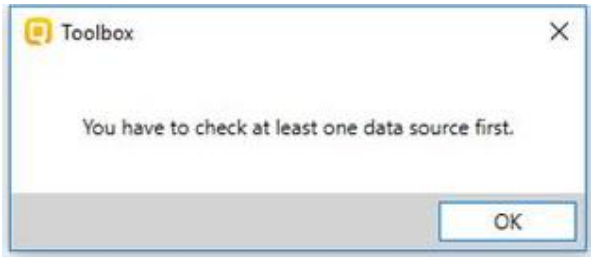


図 2

SMART 言語を組み込んだ 2D エディタを図 3 に示します。主なオプションは次のとおりです。

- 編集用ボタン (1): 鉛筆ボタン（描画）、消しゴムボタン（削除）、拡大鏡ボタン（ズーム）。空白ページ（すべて削除）、矢印ボタン（元に戻す/やり直し）
- SMART タイプのフラグメントボタン（2）（ユーザがボタン上にカーソルを置くとヒントが表示されます）：
- " [ ]" ボタン（原子の列挙）：分子内の特定の原子に結合する可能性のある原子を列挙できます。（訳注：図 7，8 参照）
- "$"ボタン（再帰 SMART 表記を可能にする）：反復表記は、ある原子の周囲を記述することを可能にします。例えば、それはニトロ基の一部ではない N（窒素）原子を表すのに使用することができます。（訳注：例えば入力欄に[$(N--)]と書くとその N 原子の価数を指定して検索可能。）
- "Rpt"ボタン（反復フラグメント）：原子またはフラグメントが分子内で何回繰り返されるかを指定できます。
- "Rpc" ボタン（置換フラグメント）：分子内の指定された原子を別の化学元素に置き換えるオプションです。
- "Exh" ボタン（限定フラグメント）：ここで列挙された原子またはフラグメントのみが所定の原子に結合することができます。
- 周期表 (3)
- 頻繁に使用する原子 (4)
- 選択ツール (5) - ユーザがその特性を定義すべき原子をポイントアウトしたいときはいつでも使用してください。

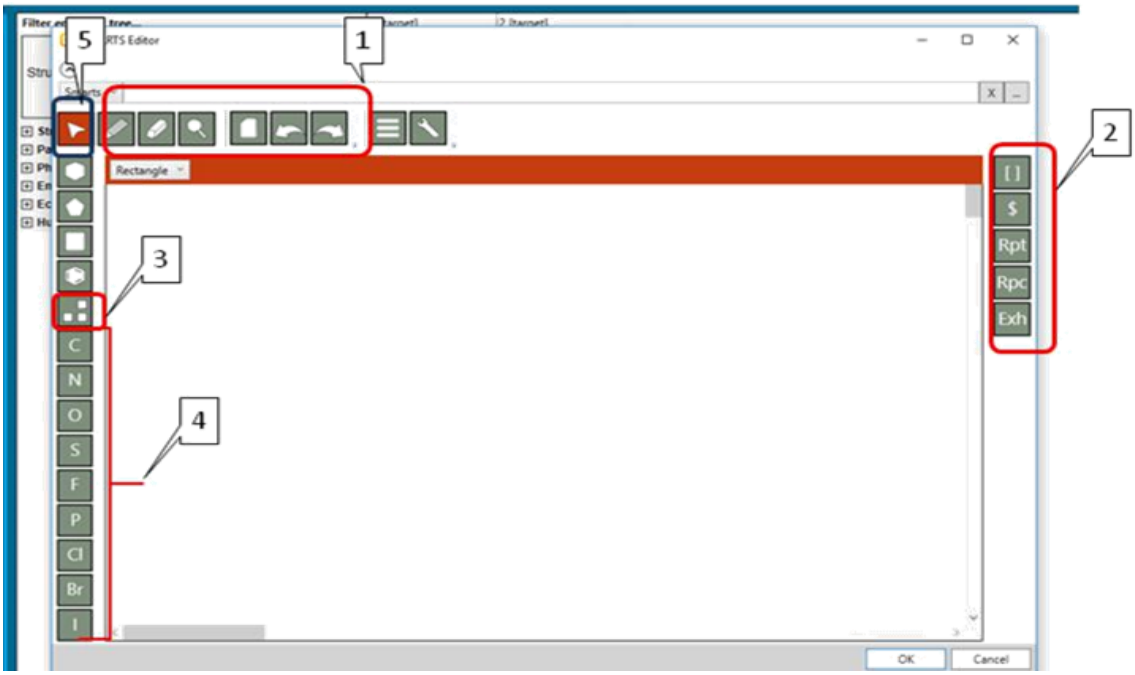


図 3

（訳注：前回の Toolbox Ver4.0 マニュアルには以下の説明が追記されていた。
「その他のオプションを図 4 に示します:

- 4 つの最も一般的な環状構造（1）のテンプレート。
- ボタン（2）を押すと、環状構造のテンプレートのリスト（3）とフラグメントのリスト（4）を含むフラグメントブラウザが開きます。フラグメントは 2D エディタの右側のフラグメントと同じです。」）

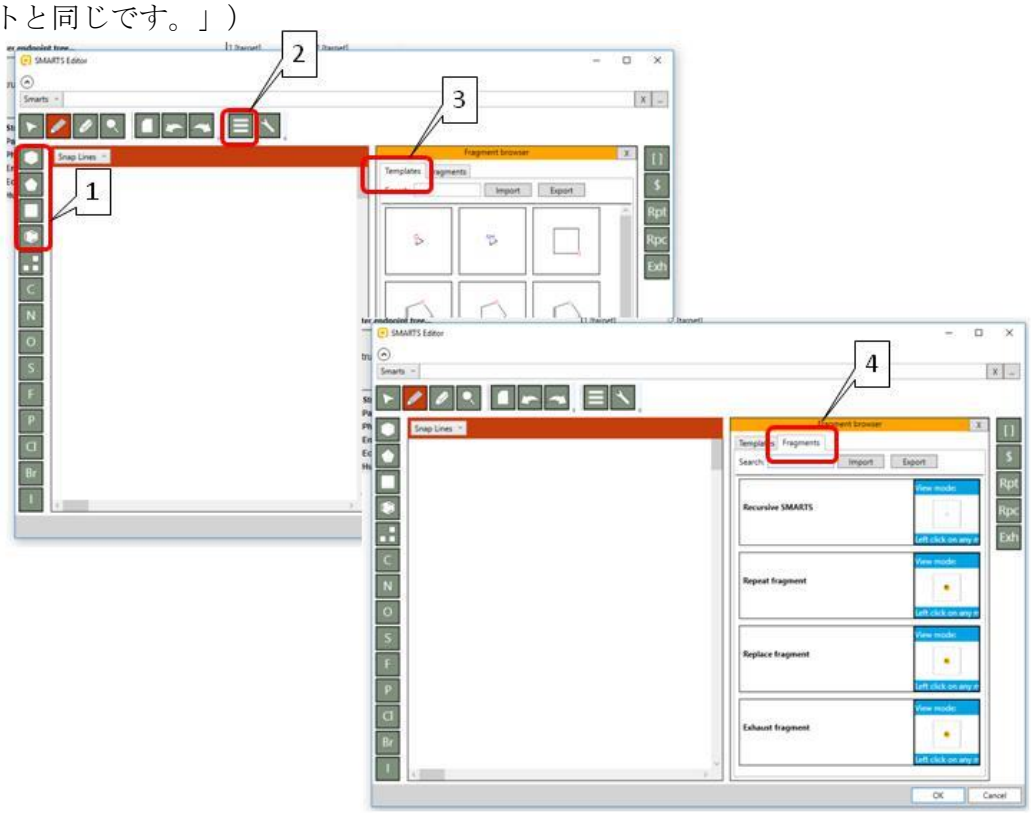


図 4

ユーザはフラグメントのライブラリに構造を追加することもできます（図 5）。化学構造を描いたら、右クリックし、Add to library（ライブラリに追加）を選択します。フラグメントのライブラリをエクスポートして sfl ファイルとして保存し、必要に応じて後でインポートすることもできます。

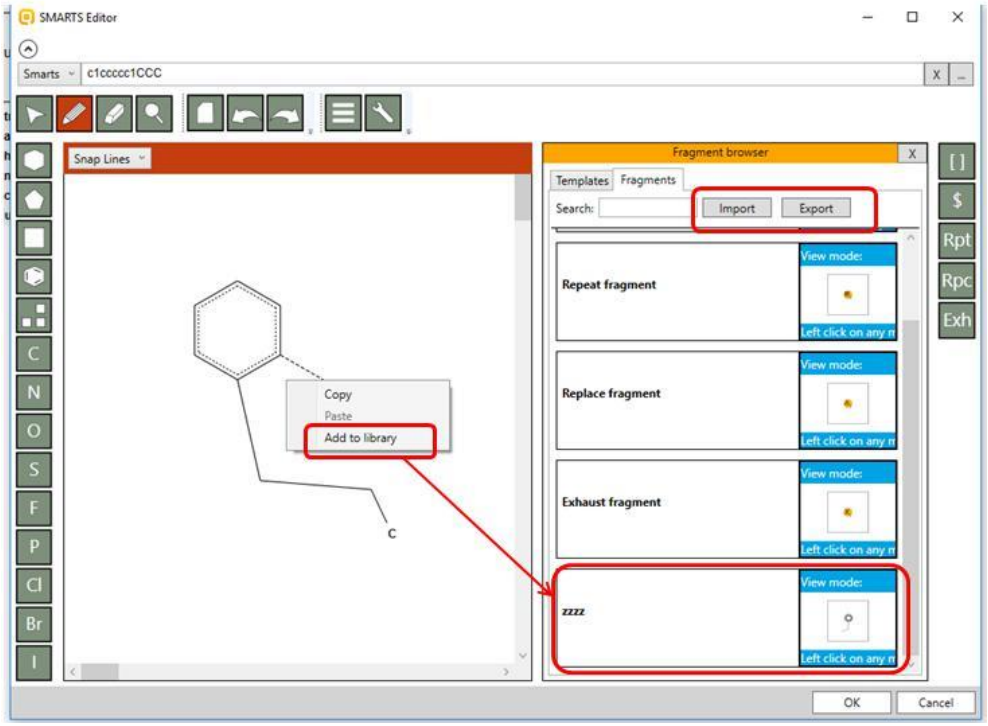


図 5

SMARTS フラグメントのライブラリは、Toolbox に備えられています。Import （インポートする）（1）をクリックするとファイルの場所が表示され（2）、Custom_Library_SMART.sfl（3）を選択して Open（4）をクリックします。

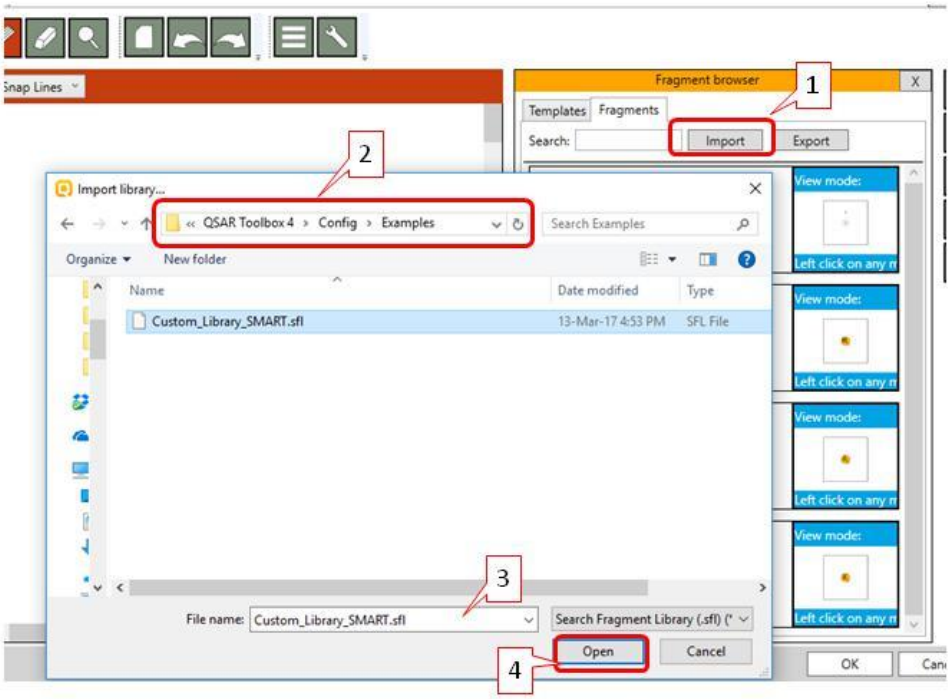


図 6

ライブラリが Fragment browser（フラグメントブラウザ）（1）に表示されます（図 7）。フラグメント（2）を選択して描画パネル（3）に表示します。フラグメントを見やすく表示するには、selection tool（選択ツール）（4）をクリックします。

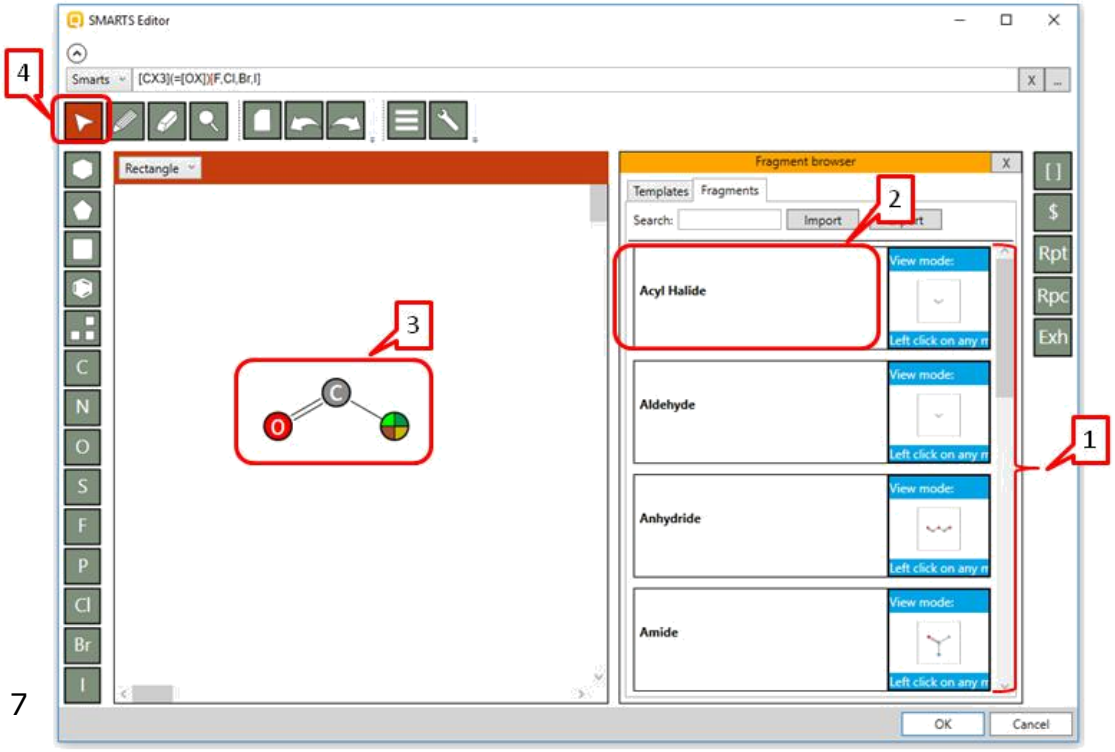


図 7

サブ構造の検索例を図 8 に示します。ハロゲン化ベンゼンを部分構造に持つ化学物質が選択したデータベースおよびインベントリから選択されるところです。

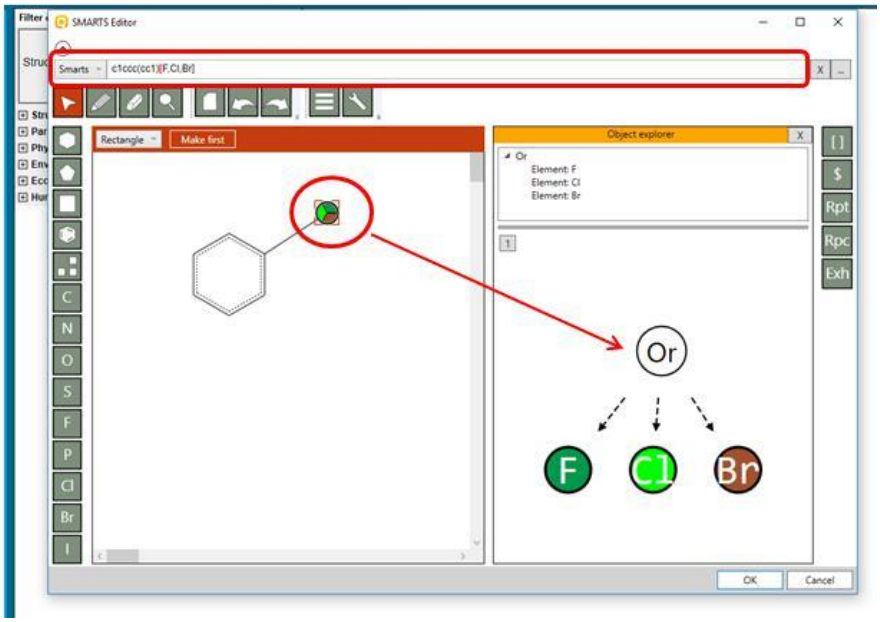


図 8

D.1.3.2. Query tool（クエリツール）

Query tool（クエリツール）は、ツールボックスの入力パネル（1）からアクセスできます（図 1）。

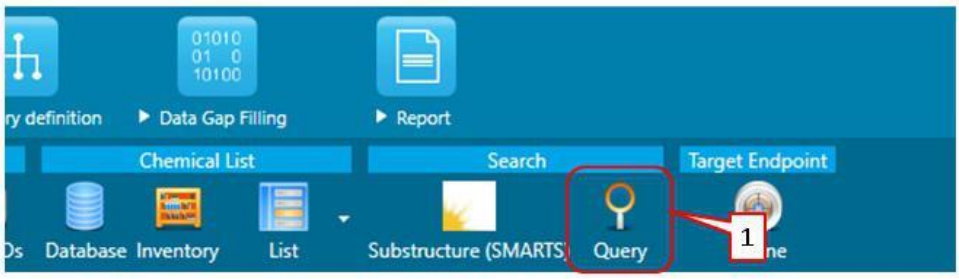


図 1

クエリ検索は、選択したデータベースとインベントリ上で実行されます。検索を続行する前に、データベースやインベントリを選択するようメッセージが表示されます（図 2）。

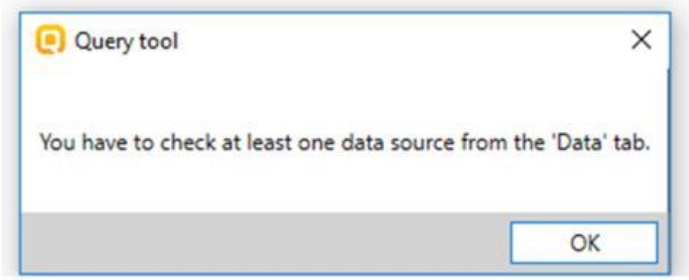


図 2

クエリが事前に作成され、新しいクエリのデータソースが変更された場合、情報メッセージが表示されます（図 3）。

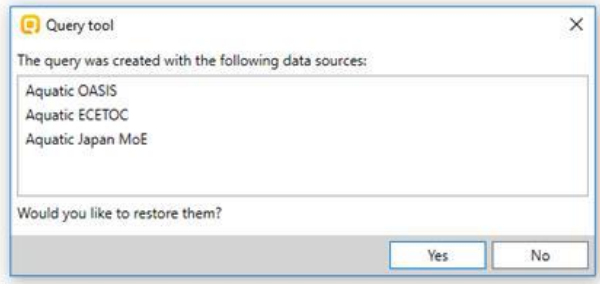


図 3

クエリツールの主要な構成要素は次のとおりです（図 4）：

- クエリ編集パネル（1）
- クエリ ツリーロジックパネル（2）

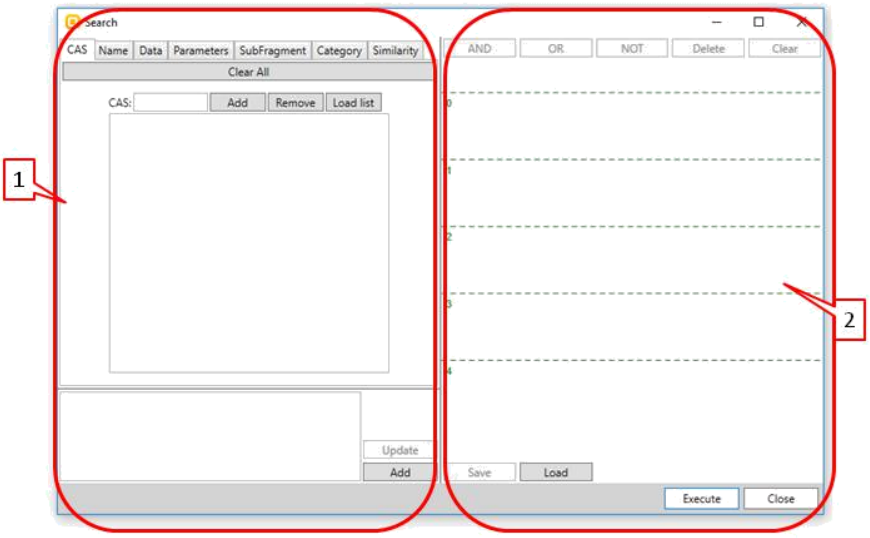


図 4

- <u>クエリツールの編集パネルの一般的な機能（図 **5**）：</u>

- メイン検索タブ（1): CAS、Name（物質名）、 Data（データ）、Parameters（パラメータ）、SubFragment（サブフラグメント）、Category（カテゴリ）、Similarity（類似性）

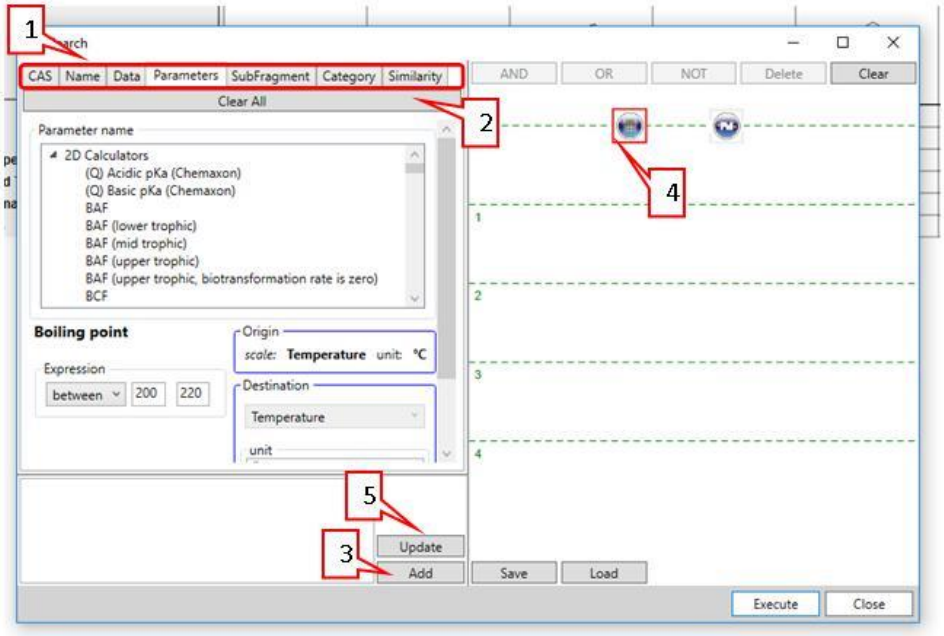


図 5

- **Clear All（全削除）**ボタン（2）は検索グループパネルのすべての書込みを削除します。
- **Add （追加）**ボタン：1 つのタブから選択したすべての検索条件を右側の実行パネルのクエリ（4）として追加します。
- クエリを選択すると、赤枠で囲まれます（4）。書き込んだものをいくつかを変更してから Update（5）を押すことで、アクティブなクエリを変更するというオプションもあります。
- クエリを右クリックすると黄色に変わり（1）、論理演算子の Not（2）が追加されます（図 5）。 また、クエリは Delete（消去）（2）することができます。Clear（クリ

ア）ボタンは、選択したクエリだけでなく、パネル内のすべてのクエリと論理演算を削除します。Save（保存）（3）を押すと、すべてのクエリと論理演算を保存できます。Load（読み込み）（4）を押すと、保存したクエリを読み込みできます。Execute（実行）（5）をクリックしてクエリを実行します。

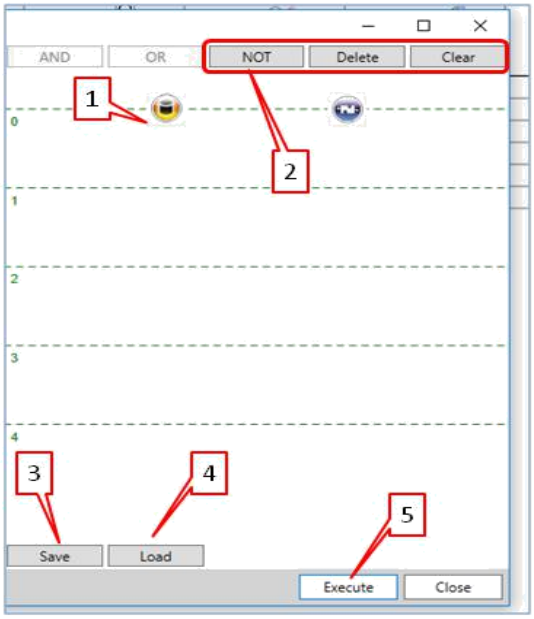


図 6

- 論理演算子 AND や OR を挿入するには、選択したクエリ（黄色で表示）（1）を右クリックし、次に論理演算（2）を選択する必要があります（図 7）。論理演算（赤色で表示）を右クリックすると（4）NOT を追加したり、Delete をしたりすることができます（5）。

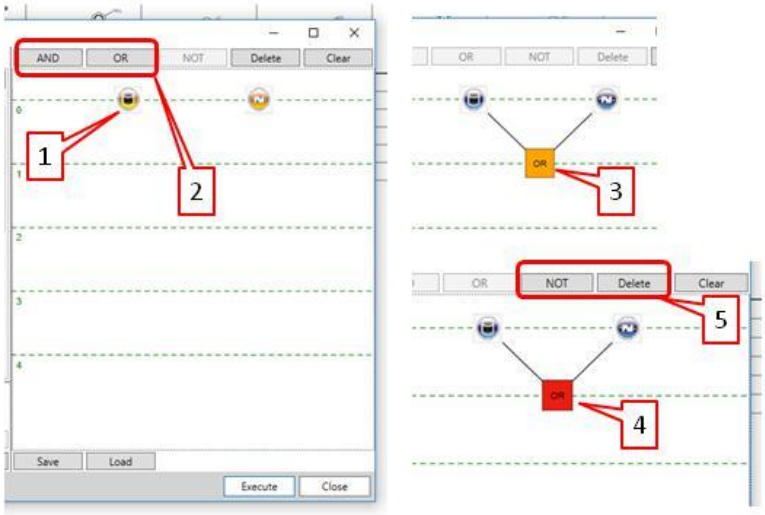


図 7

- **CAS 番号**

CAS 番号パネルでは、CAS 登録番号にもとづいて、選択したデータベースまたはインベントリを検索することができます。 (2)のボタンですが、セル内に書き込まれた CAS 登録番号（図 8 の 1）を削除し、Add をクリックして 番号を入れなおすオプションも備えられています。

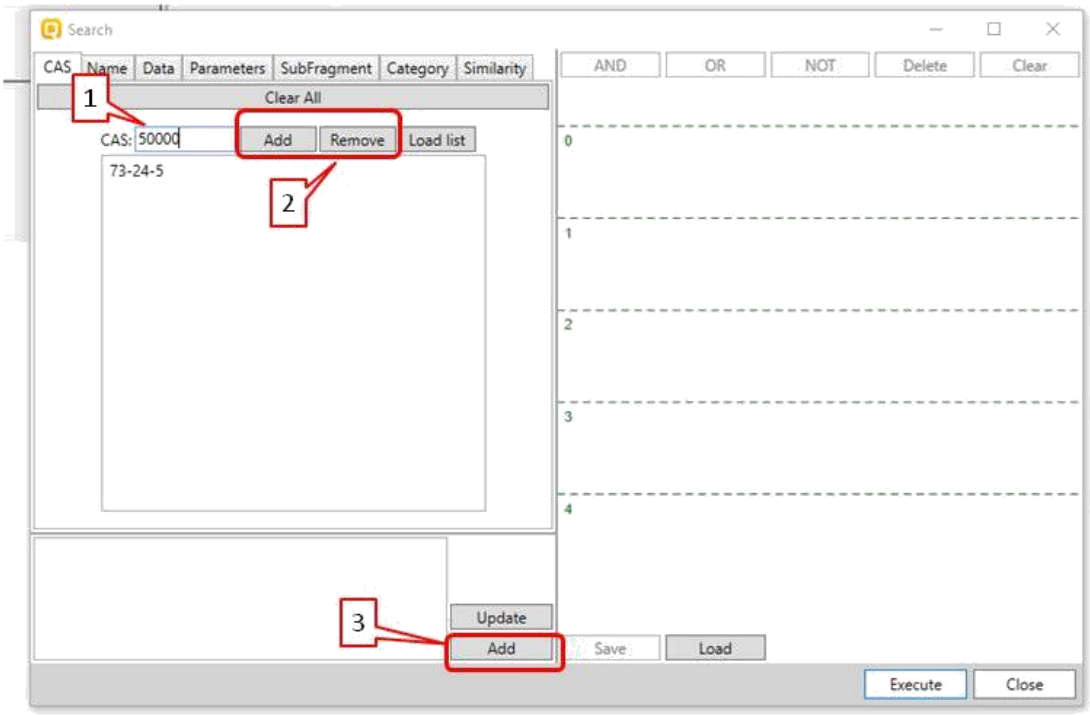


図 8

対象となる CAS 登録番号が入力されたら、Add（1）を押すと新しいクエリが表示されますので（2）、次に Execute（実行）（3）を押します（図 9）。

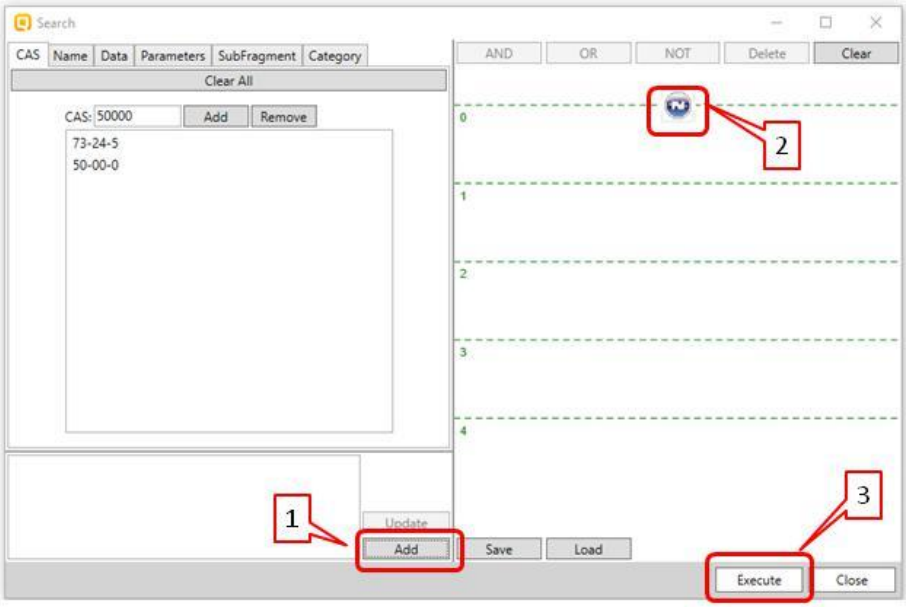


図 9

選択された情報源の化学物質であって、その CAS 登録番号がクエリ内の CAS 登録番号と一致するすべてのものが、マトリックスに表示されます（図 10 参照）。見つかった化学物質の数が Documents のツリー（1）に表示され、Structure info（化学構造情報）（2）が展開されて化学物質の識別情報（3）が表示されます。

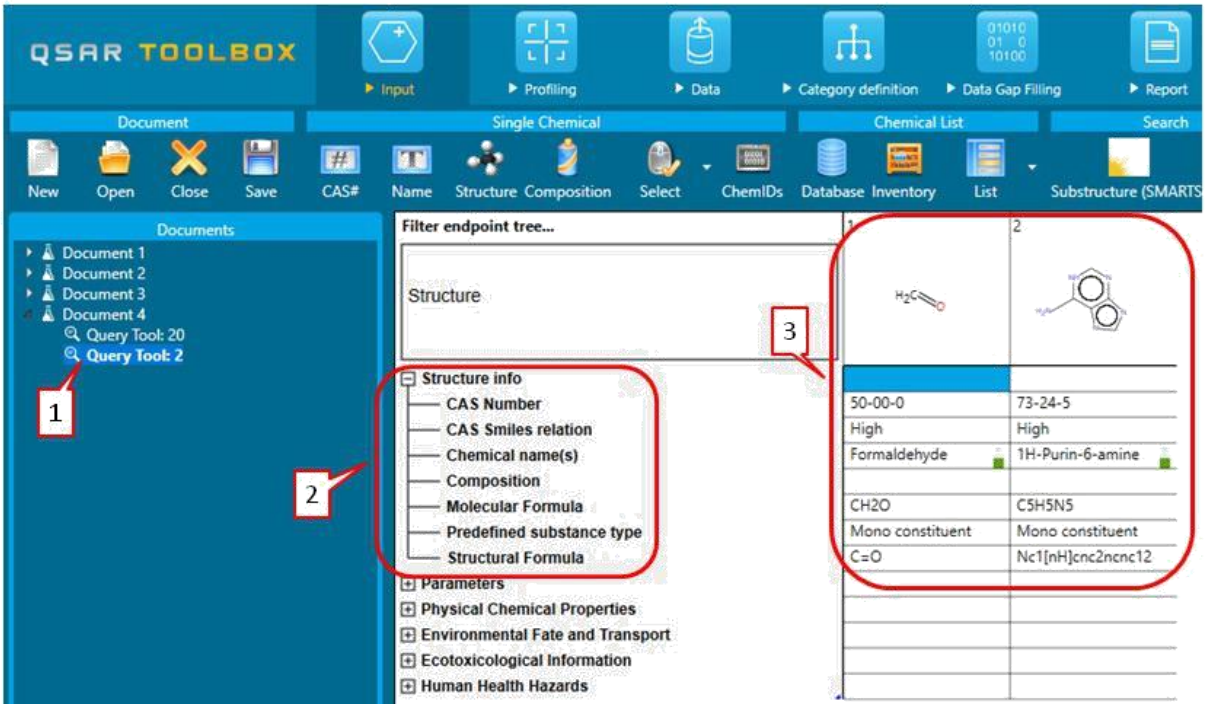


図 10

また、txt 形式（2,3,4）でなければならない CAS 番号リスト（1）を読み込むこともできます。（図 11）

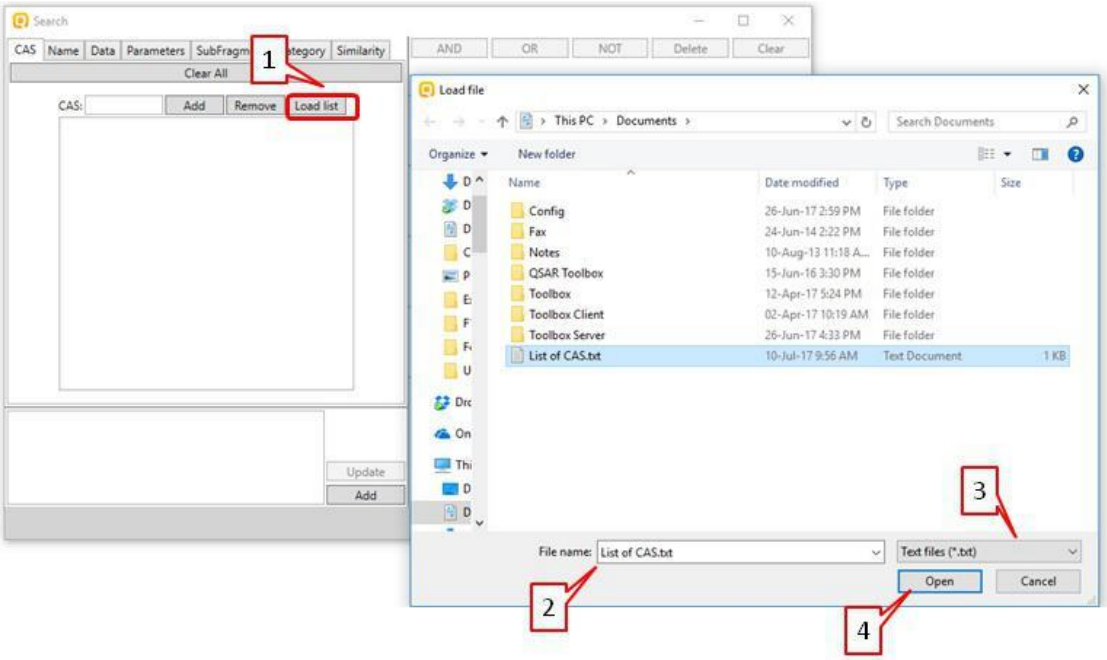


図 11

CAS 登録番号が表示されたら（1）、Add（2）をクリックしてから Execute（3）をクリックします（図 12）。

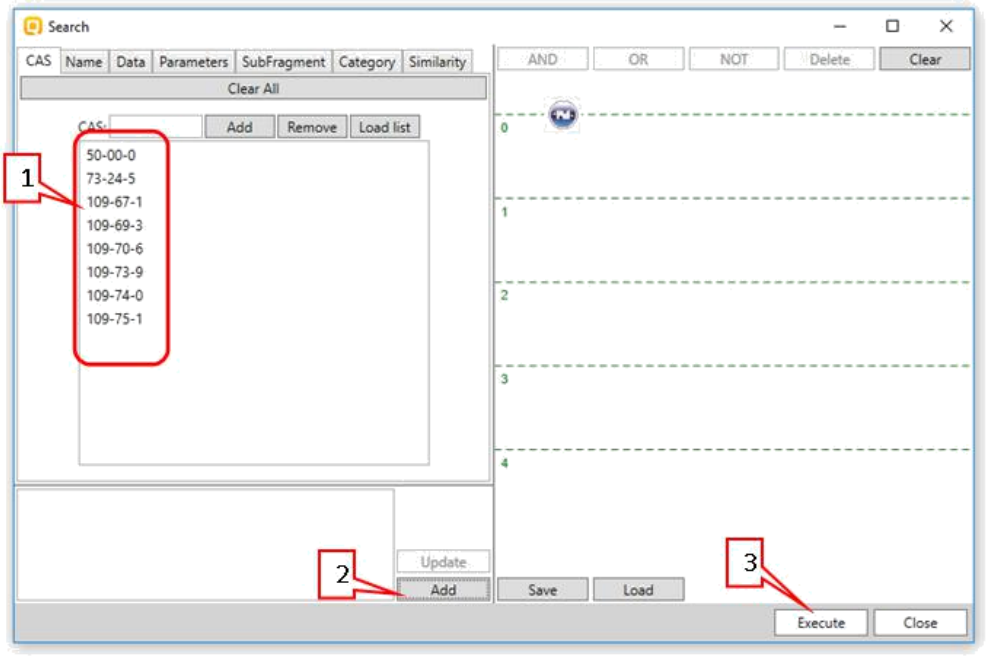


図 12

クエリと一致するすべての化学物質が表示されます（図 13）。

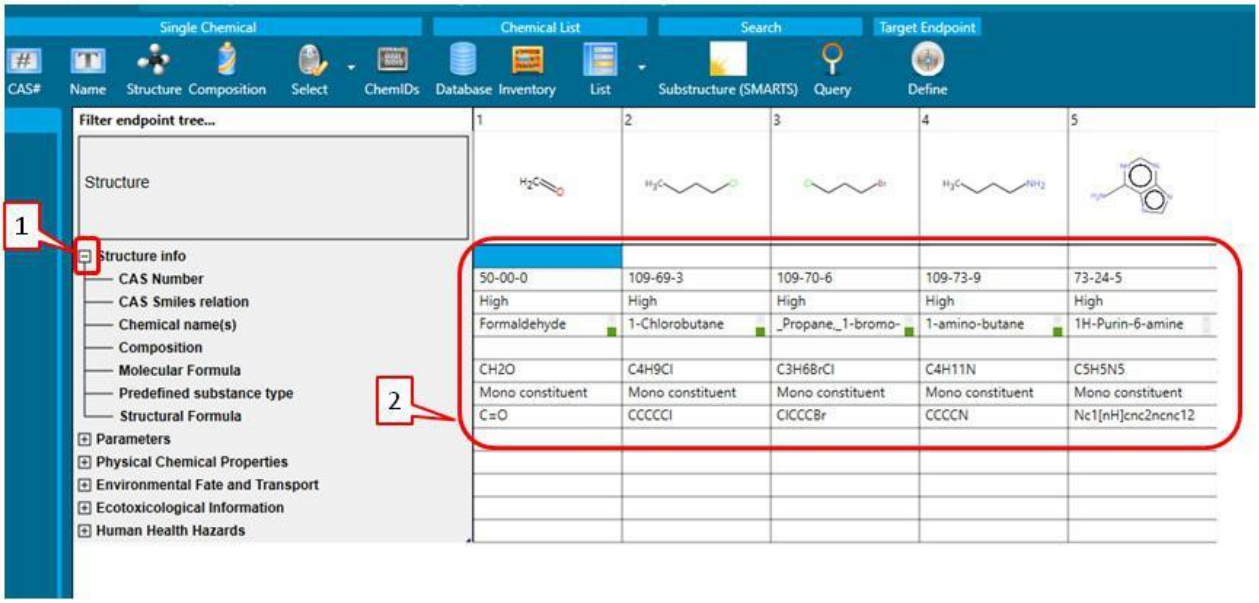


図 13

- **物質名**

Name（物質名）パネルでは、物質名に基づいて、選択したデータベースまたはインベントリ（Data の項を参照）内で検索を実行できます。まず、セルに名前を書き込んでください（図 14 の 1）。Name（物質名）に関する追加のオプションを選択し（2）、Add（追加）ボタン（3）を押します。入力された名前を削除するオプションも導入されました（3）。

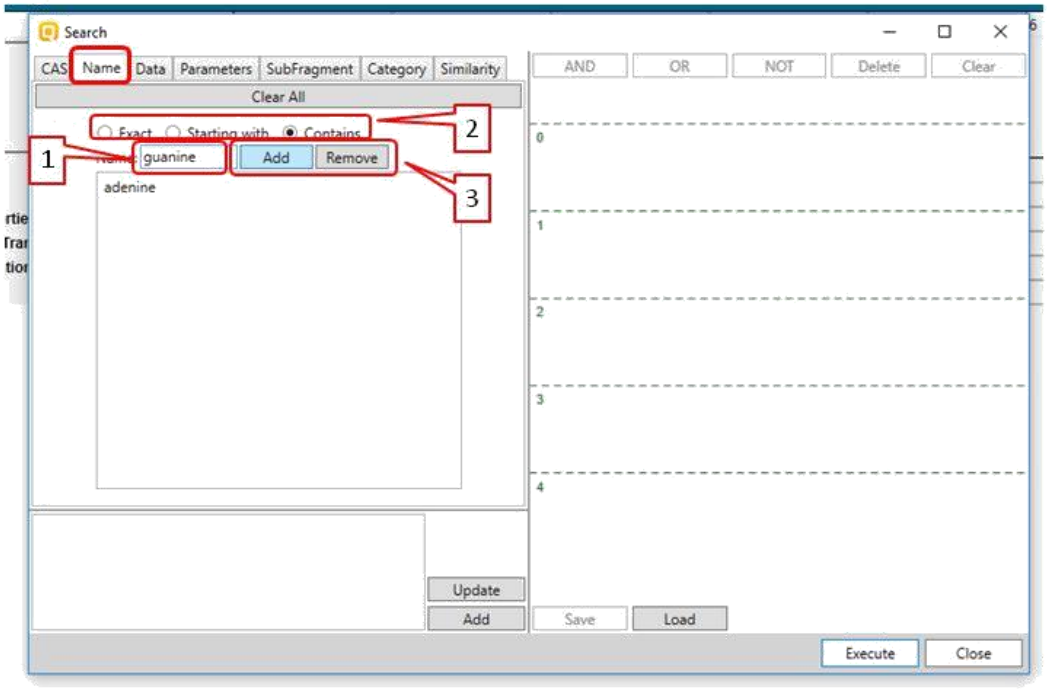


図 14

対象となる CAS 登録番号が入力されたら、Add（1）を押すと新しいクエリが表示されますので（2）、次に Execute（実行）（3）を押します。(図 15)

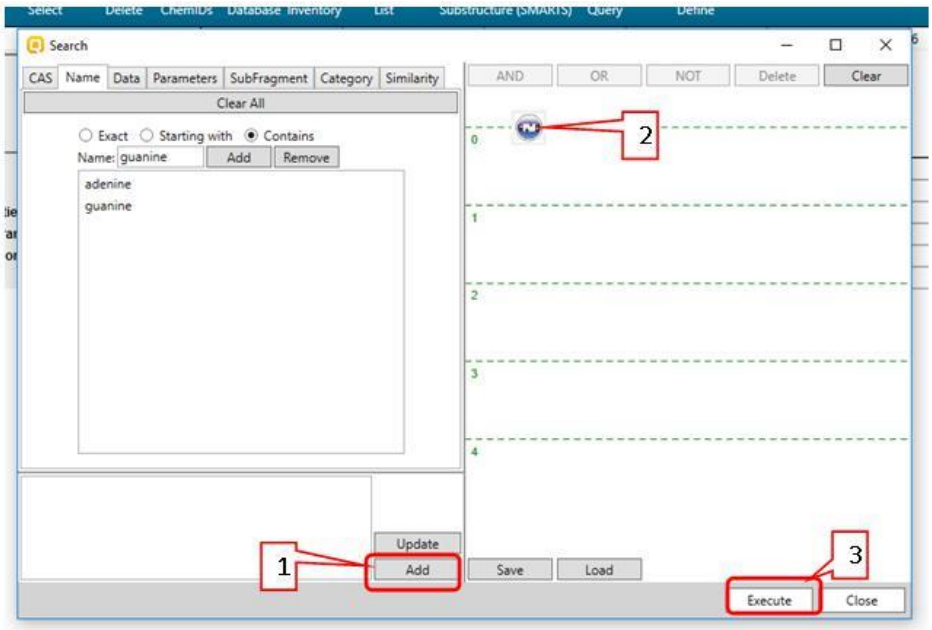


図 15

図 15 の例では、名前にアデニンまたはグアニンが含まれている化学物質が、選択したデータベースやインベントリ内で検索されます。結果は**データ・マトリックス**に表示されます。基準に合致する化学物質の数（1）が Document ツリーに書き込まれます（図 16）。化学情報（3）を見たいときは、Structure info（2）を展開してください。

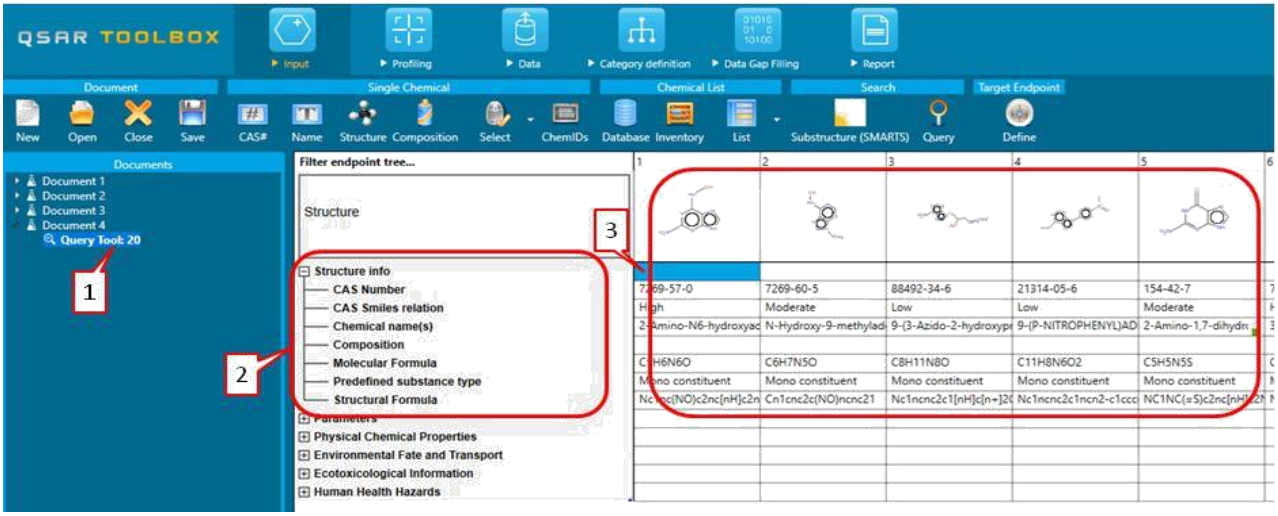


図 16

- <u>データ</u>

Data パネル（1）で表示されたエンドポイント（2）は、あらかじめ選択されたデータベースに基づいています（図 17）。例えば、Ecotoxicological（生態毒性）の情報を含むデータベースのみを選択した場合、そのノードだけ表示されます。

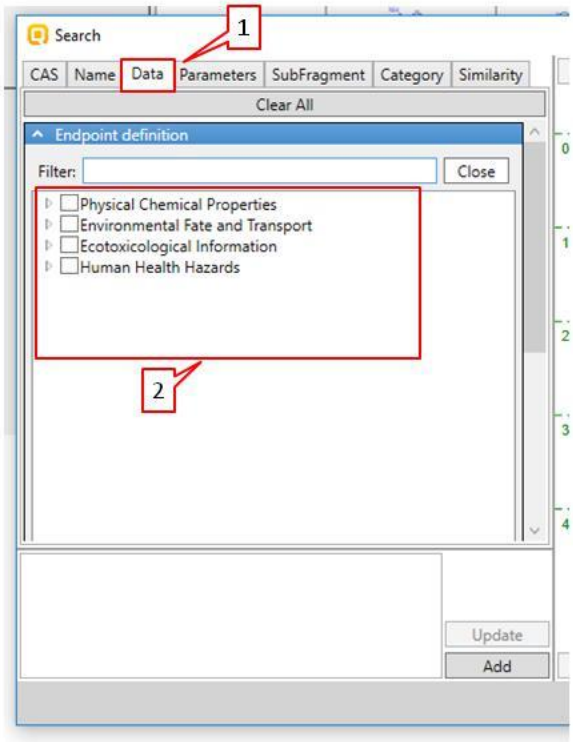


図 17

選択手順には、以下のステップがあります。

- endpoint（エンドポイント） (1) :

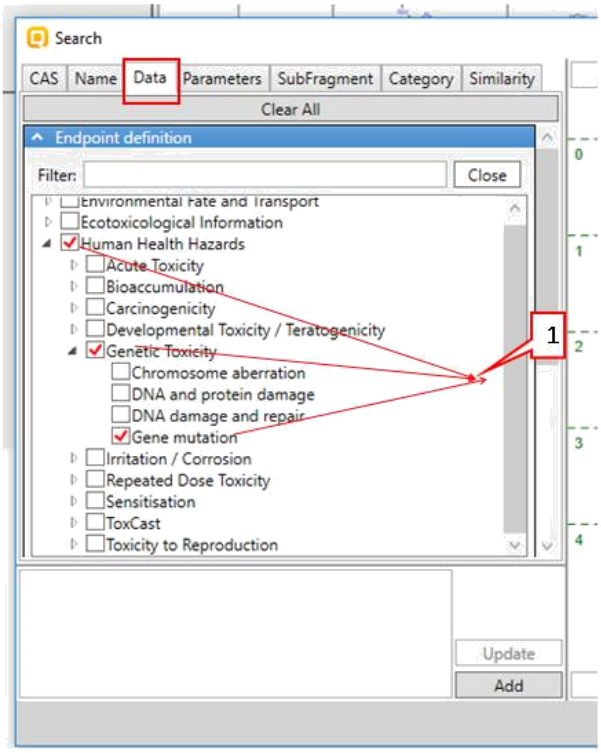


図 18

- Meta data（付随データ） : ドロップダウン矢印（1）から付随データのタイプを選択し、次に Add（追加）（2）を押します（図 19）。（訳注 : Endpoint Definition 画面を下にスクロールすると図 19 の通り Metadata を指定する部分が表示される）

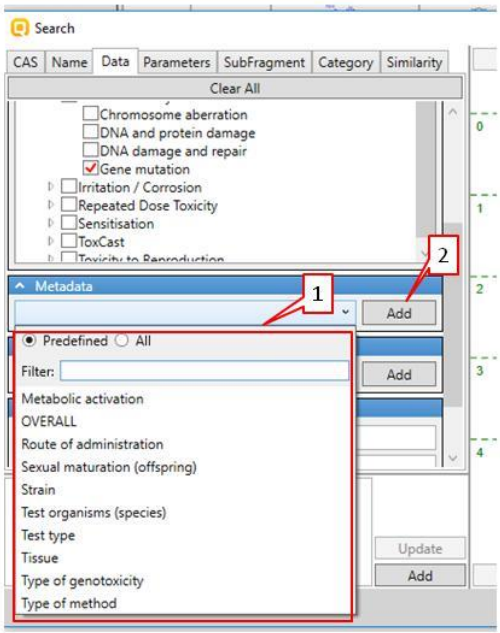


図 19

例えば、 Test organism species（試験生物種）を選択した場合は（1）、種のタイプを選択し（2）、Add（3）を押します（図 20）。

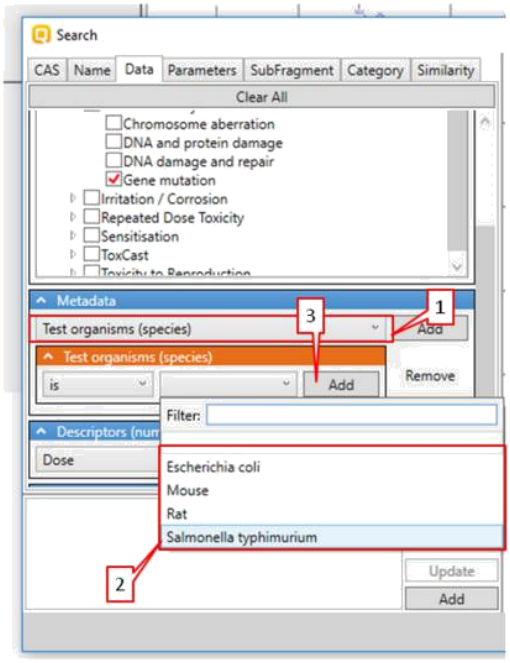


図 20

スクロールダウンし（1）、もし Descriptor (numerical data)（数値データの記述子）、例えば、持続時間、溶解度などが選択したエンドポイントと関連する場合、それは（2）に表示されます（図 21）。

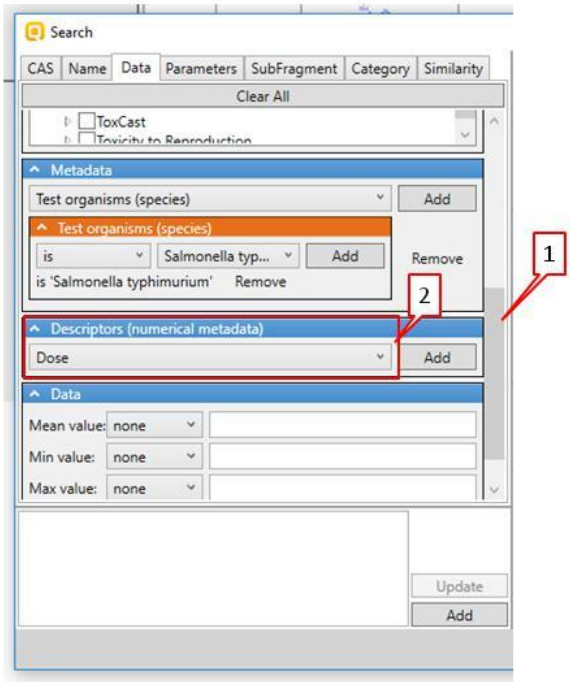


図 21

単位（尺度）（1）とカテゴリデータ（2）の選択も含まれています（図 22）。

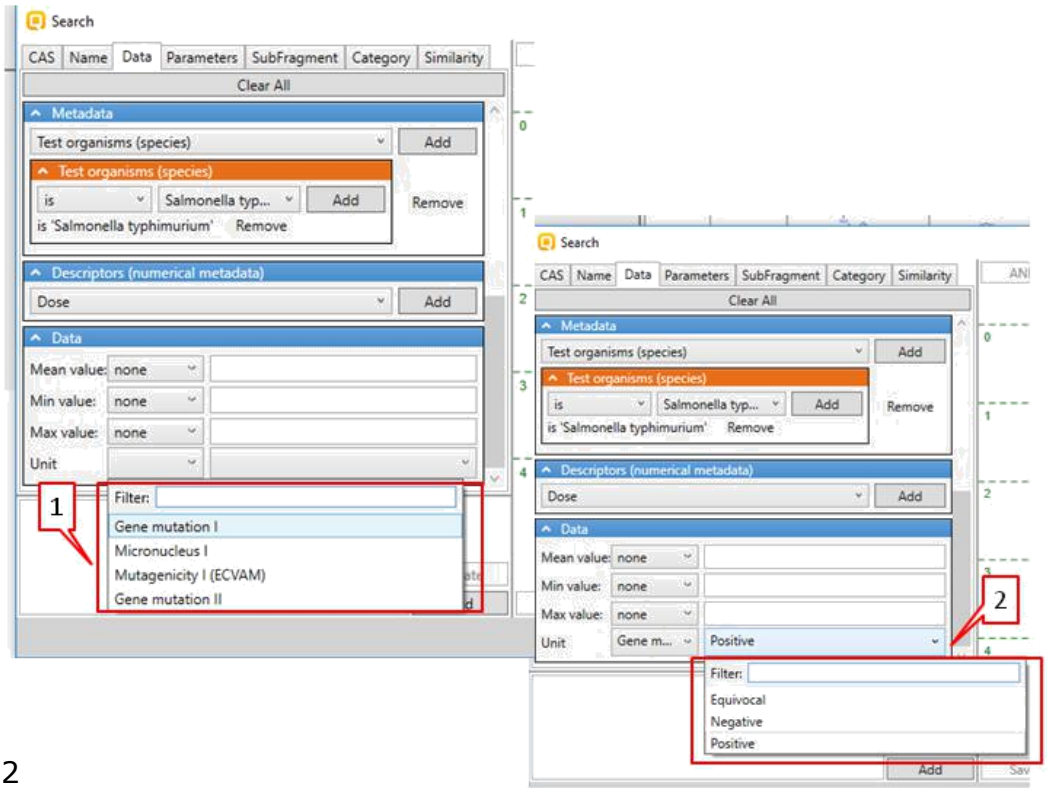


図 22

入力欄に全て入力したら、Add（1）を押すと新しいクエリが表示されますので（2）、次に Execute（実行）（3）を押すか、クエリをダブルクリックします。 (図 23).

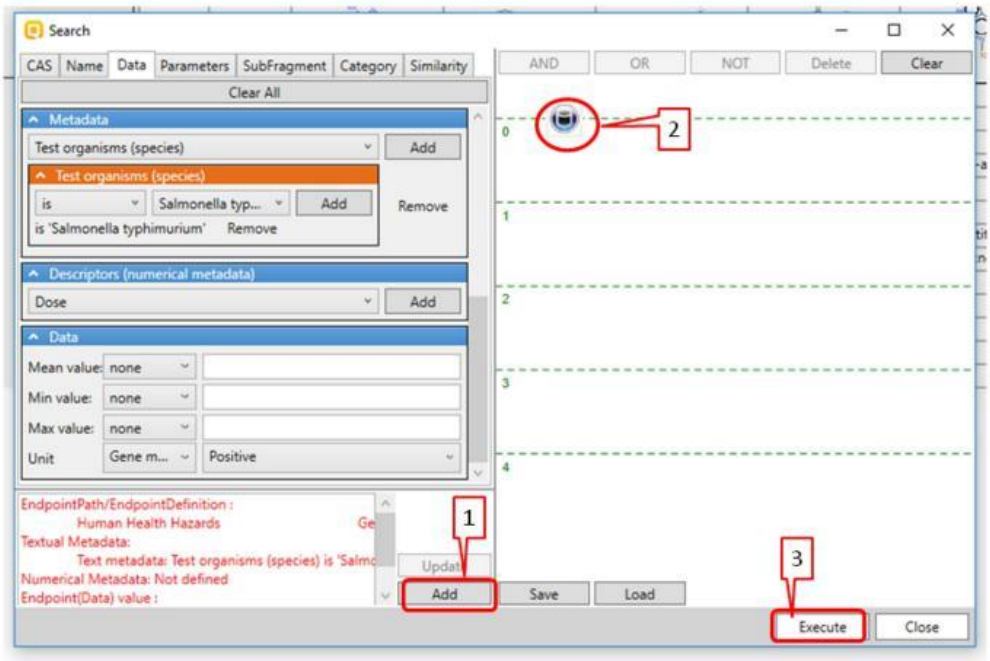


図 23

- **Parameters（パラメータ）パネル**

4 つ目の Parameters パネル（1）では、ユーザが 2D または 3D パラメータ（2）でデータベースを検索することができます（図 24）。さまざまな検索オプションを設定するために不等号などが使用できます。（3）

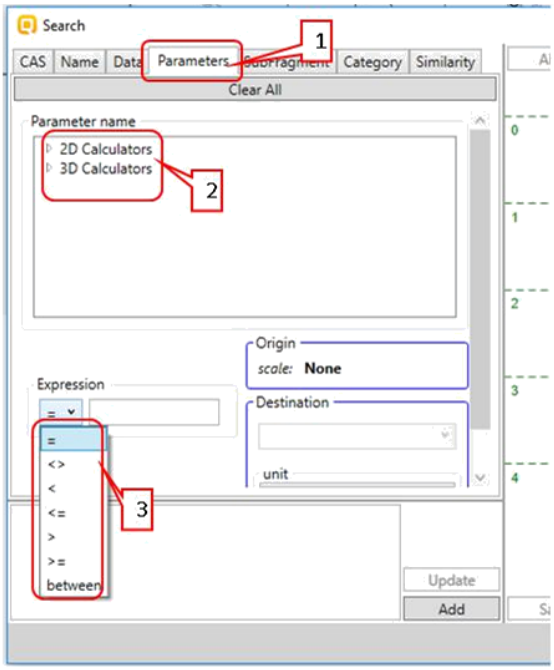


図 24

選択過程は以下の手順を含みます：パラメータ（1）を選択し、不等号などを選択して値（2）を入力し、expression（表記）のタイプ（3）を選択し、単位（4）と表示のタイプ（5）を選択します（図 25）。

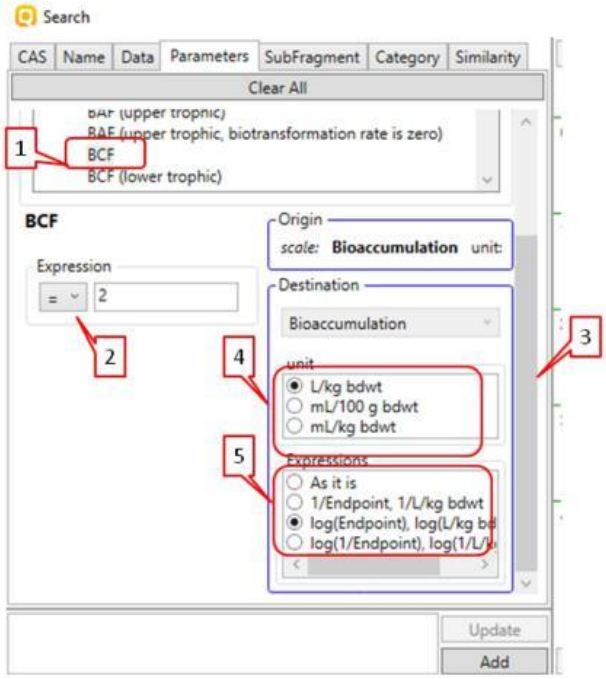


図 25

- **SubFragment（サブフラグメント）パネル**

5 番目の SubFragment パネル（1）では、Add（2）ボタンを押すと描画（4）（または SMARTS（3）をペースト）で化学構造を検索できます（図 26）。SMARTS エディタが表示されるので、そこにサブフラグメントの SMARTS（3）を貼り付けたり、描画したりできます（4）。追加の検索オプションも備えられています（5）。SMARTS エディタの使い方について詳細を知りたいときはサブストラクチャ（SMARTS）を参照してください。

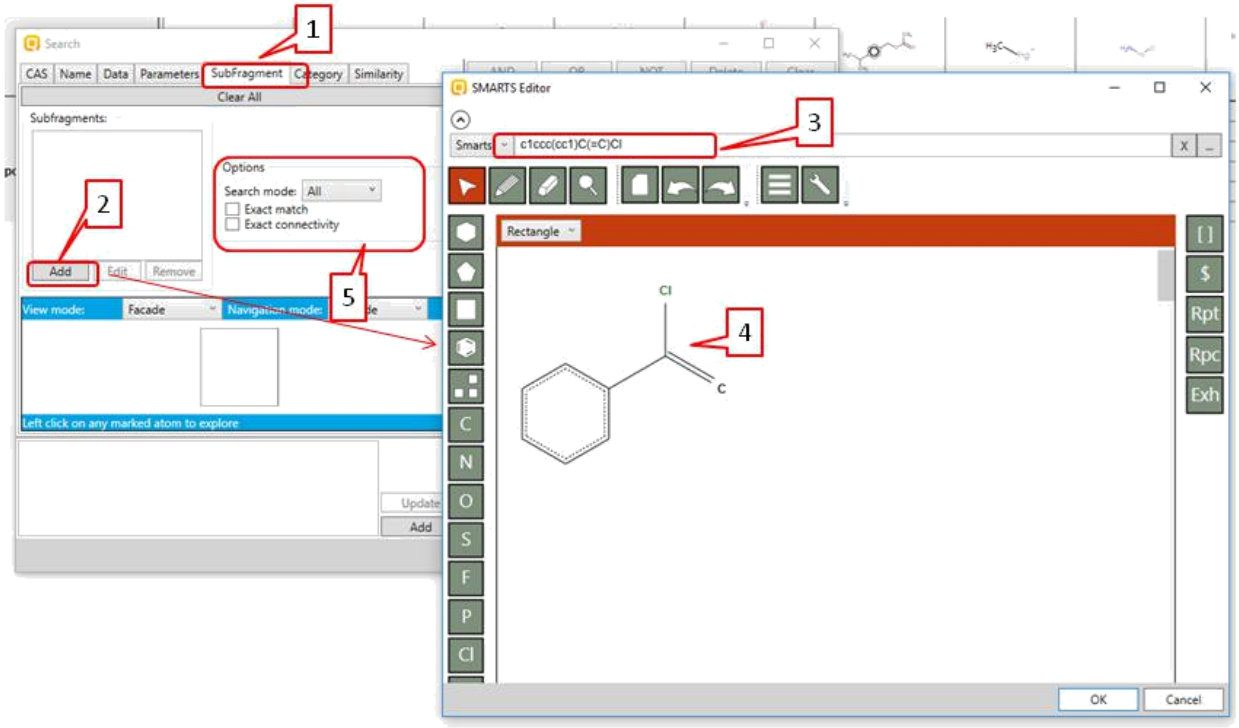


図 26

- **Category（カテゴリ）パネル**

Category （カテゴリ）パネルでは、使用可能な profiles（プロファイル）（1）のいずれかによる検索が可能です。（図 27）フィルタ入力欄を使用して、プロファイラの名前にあるキーワードを挿入します。

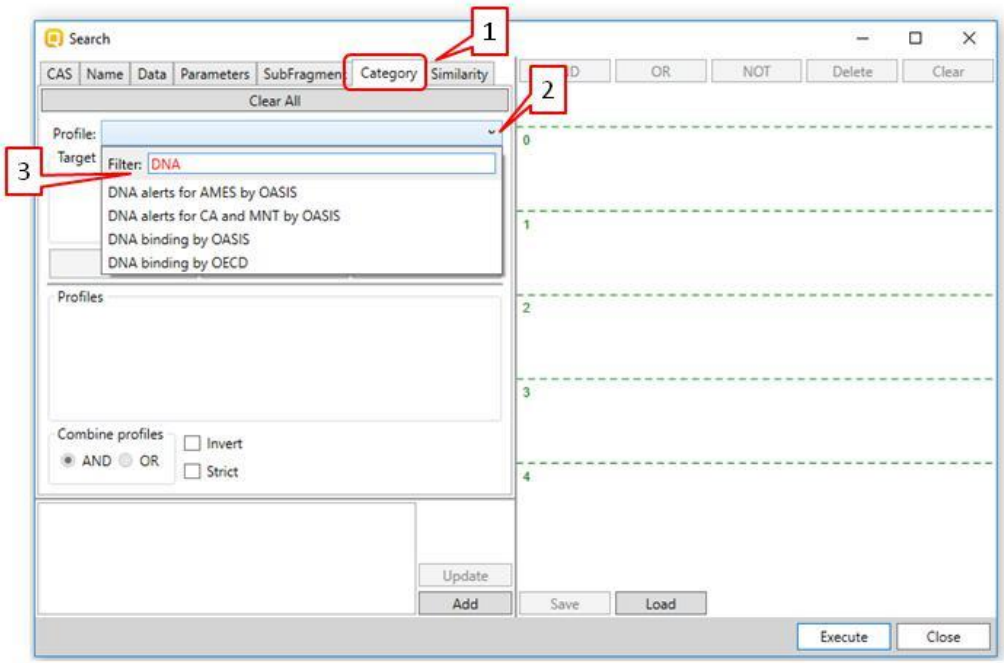


図 27

プロファイルを選択したら（1）、パネル（2）から目的のカテゴリを選択し、（3）のボタンでパネル（4）に移動します（図 28）。他にも検索オプション（5）が使用できます(図 28)。

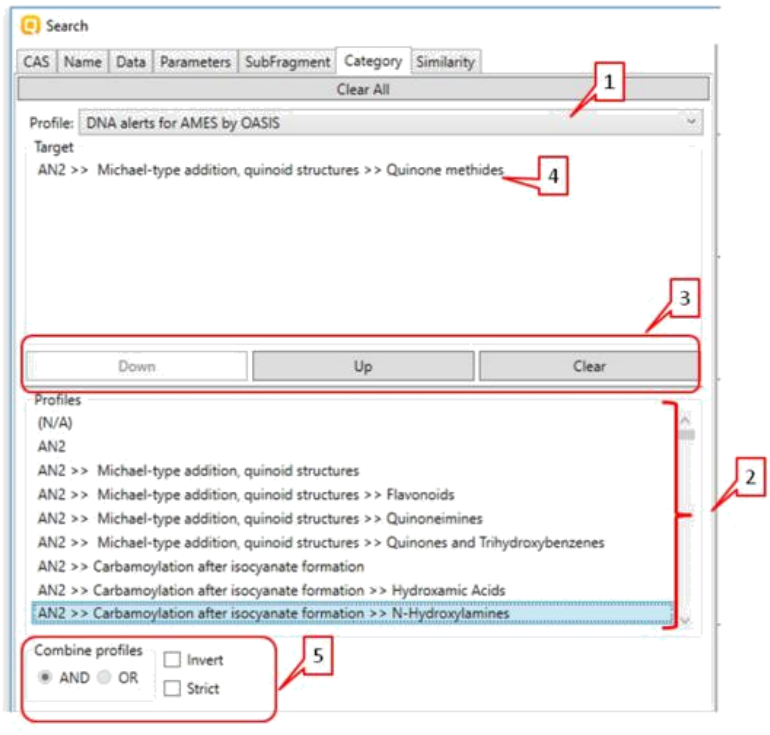


図 28

- **Similarity（類似性）panel**

標的化学物質の類似性基準に基づいて化学物質を検索することができます。
Similarity パネル（1）で、Options ボタン（図 29 の 2）をクリックして、さまざまな類似性オプションを表示/調整します（図 30）。

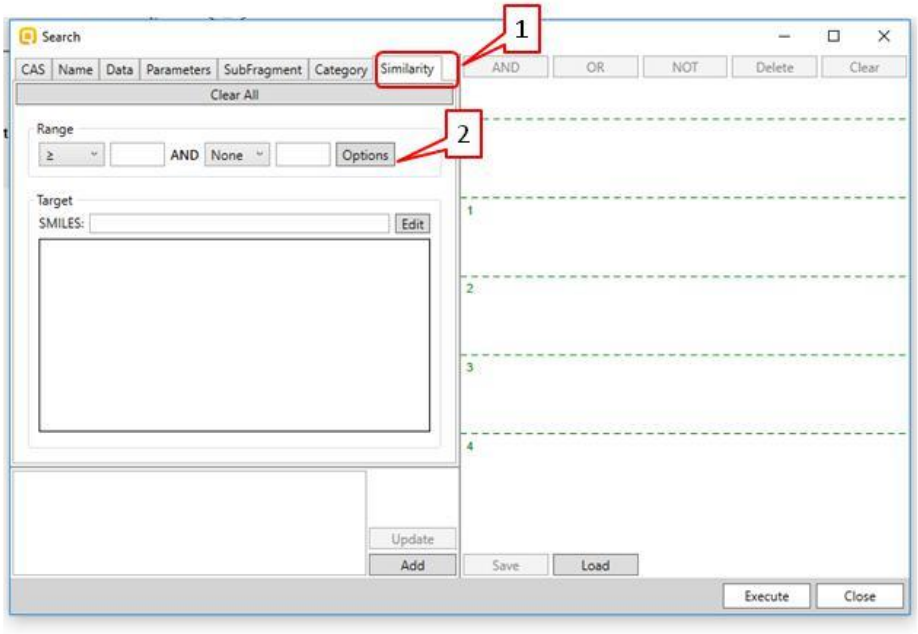


図 29

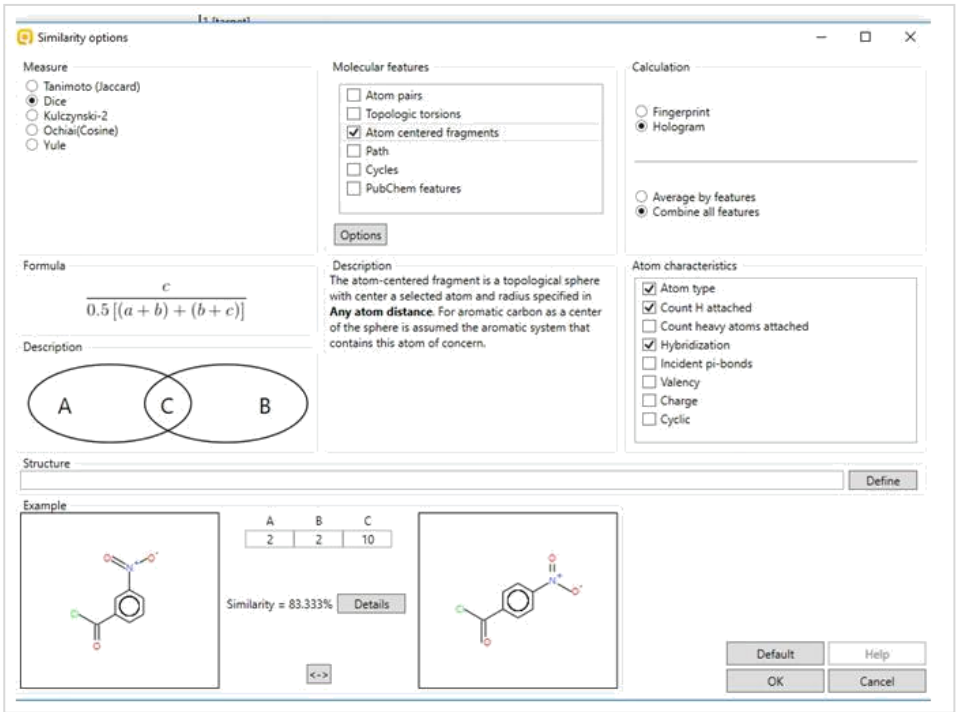


図 30

標的物質を 2D エディタ（図 31 の 1,2,3）に描くか、または SMILES 入力欄（4）に SMILES を貼り付けることができます。

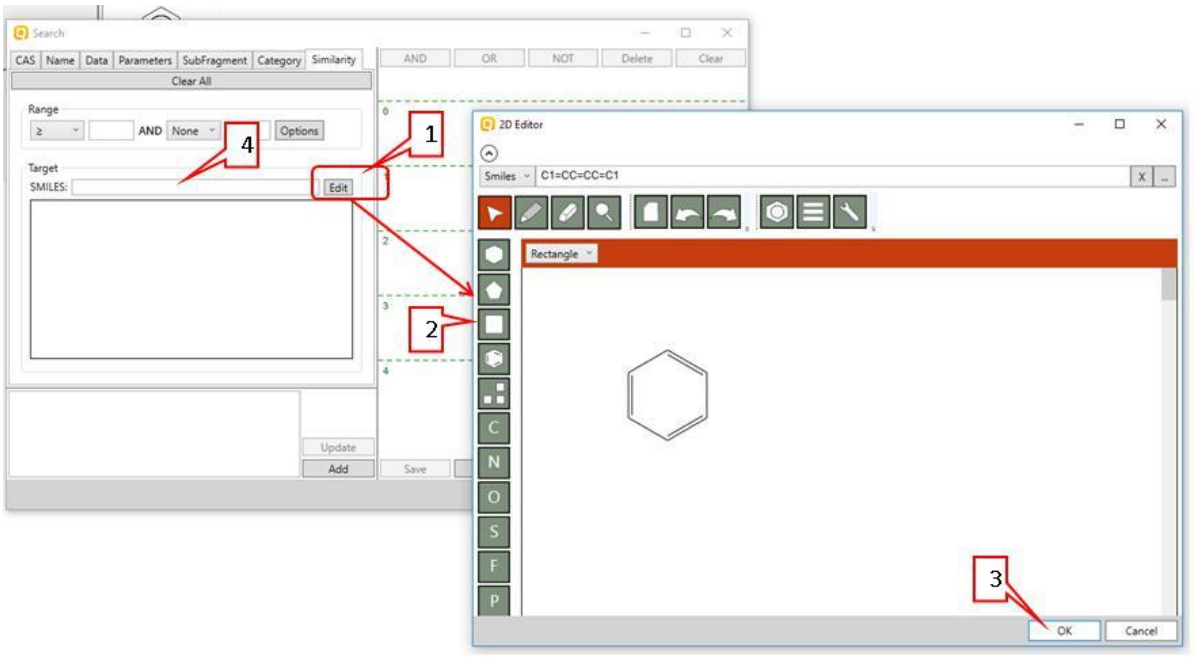


図 31

標的物質が表示されたら（1）、類似度の範囲（%）（2）を選択してください（図 32）。

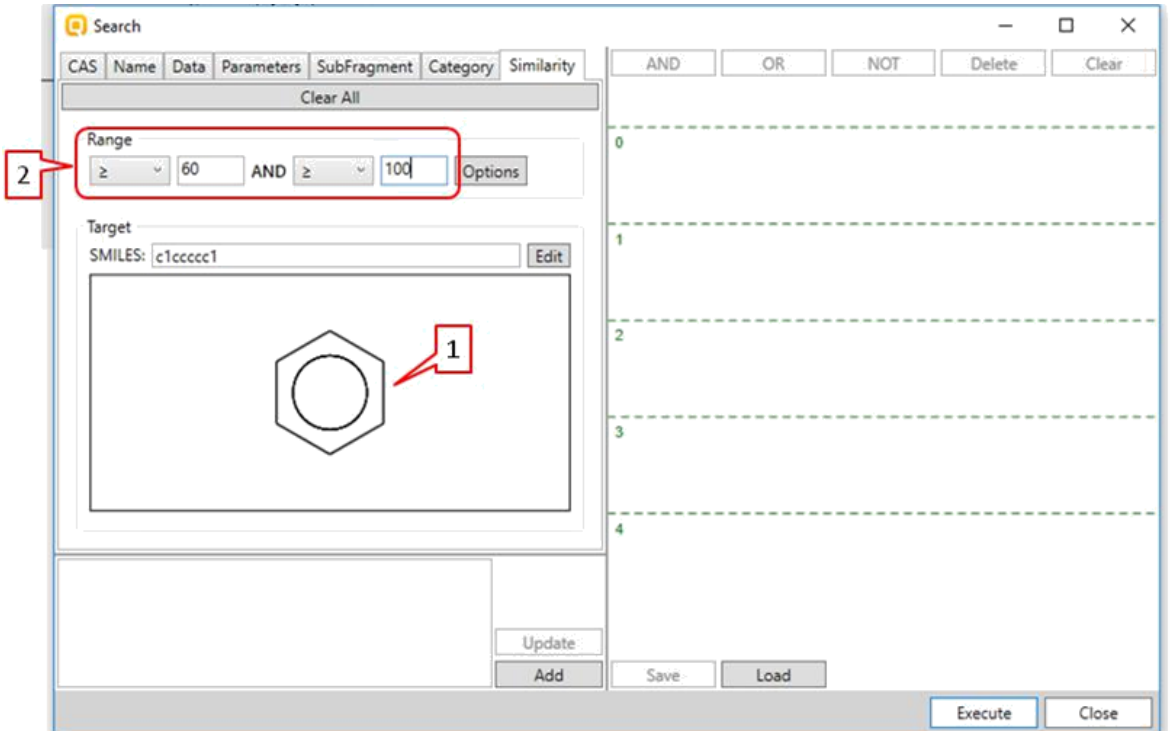


図 32

Add（1）をクリックして新しいクエリを作成し（2）、Execute （3）をクリックします。（図 33）

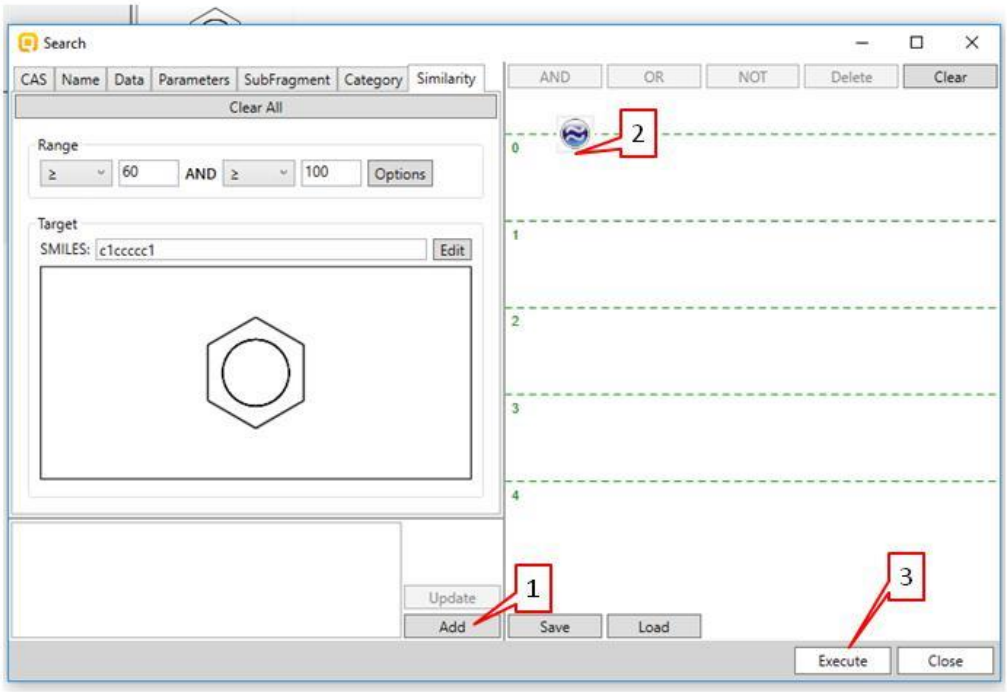


図 33

D.1.4. 標的エンドポイントを決める

Toolbox 4.1 では、標的エンドポイントを事前に決定できます。この機能により、EC3, LC50, gene mutation（遺伝子突然変異）など、調査しようとするエンドポイントを入力できるようになりました。選択したエンドポイントに関連するプロファイルとデータベースのための情報をさらに含めることができます。

特定の付随データ情報に基づいて、同じエンドポイントに対してもさまざまな関連性スコアを提供できます。アラートパフォーマンス（AP）の計算は、標的エンドポイントが事前に選択されている場合にのみ可能です。

標的エンドポイントを選択する主な手順は次のとおりです（図 1）。

- CAS＃ボタン（1）を押す
- （2）に CAS 番号を書く
- Search ボタン（3）を押す
- 化学物質の横にあるチェックボックスをオンにする（4）
- OK をクリック　（5）

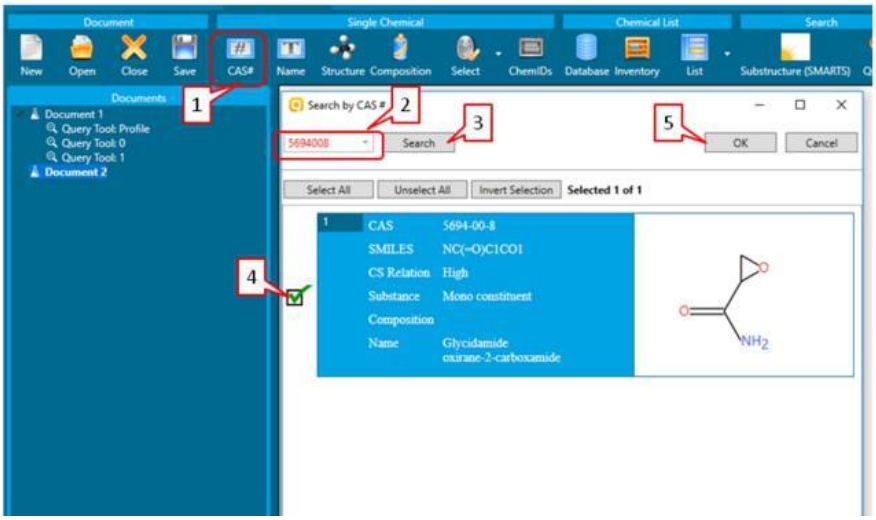


図 1

化学構造式が**データ・マトリックス**（1）に追加されたら、Define Target Endpoint（標的エンドポイントを決める）を選択します。（図 2 の 2）

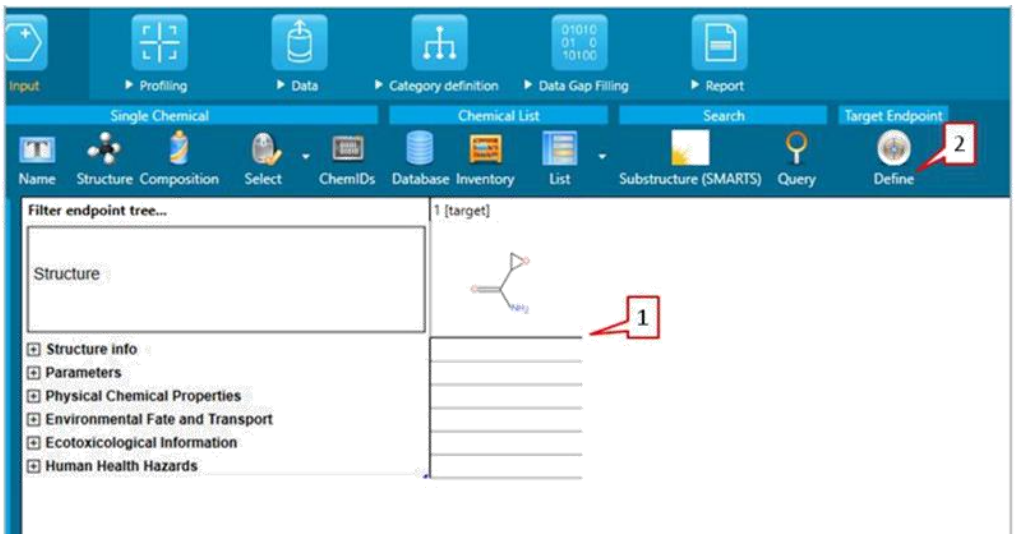


図 2

Define endpoint （エンドポイントを決める）ウィンドウが表示されます (図 3)。

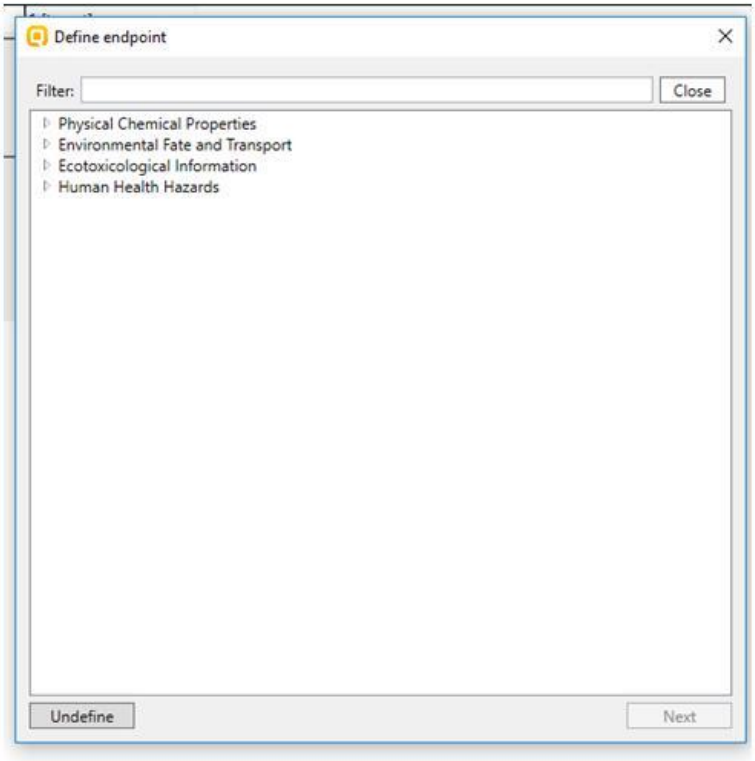


図 3

ツリーを展開してエンドポイントのメインレベルを選択するか（1）、Filter の入力欄に入力して（2）、Next（3）を押します。（図 4）

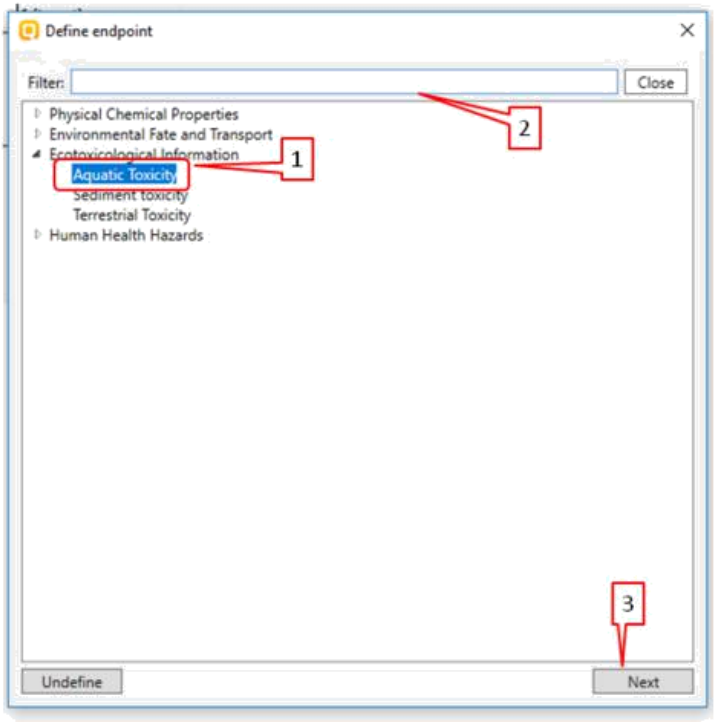


図 4

フィルタ入力欄（3）を使用して、ドロップダウンメニュー（2）からエンドポイント（1）を選択します（図 5）。

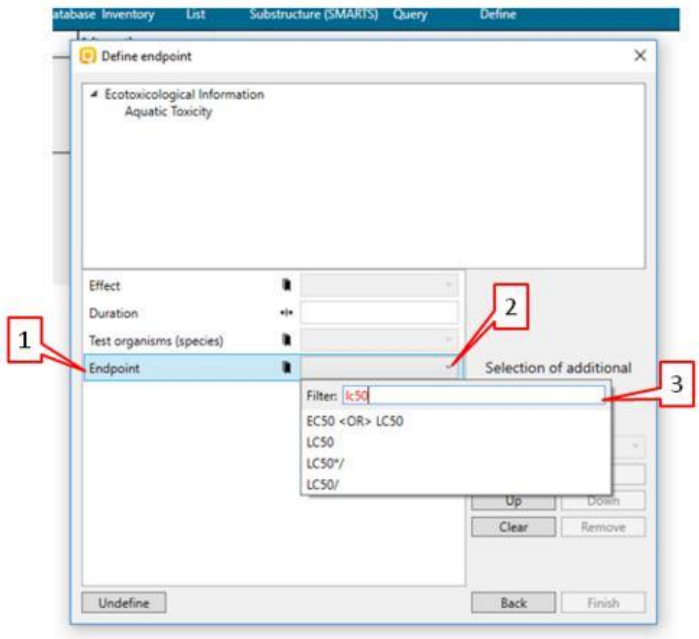


図 5

フィルタ入力欄（3）を使用して、Test organism (species)（生物種を試験する） (1)をドロップダウンメニュー （2）から選択します （3） (図 6)。
次に、持続時間を入力します（例：96 h）（4）。

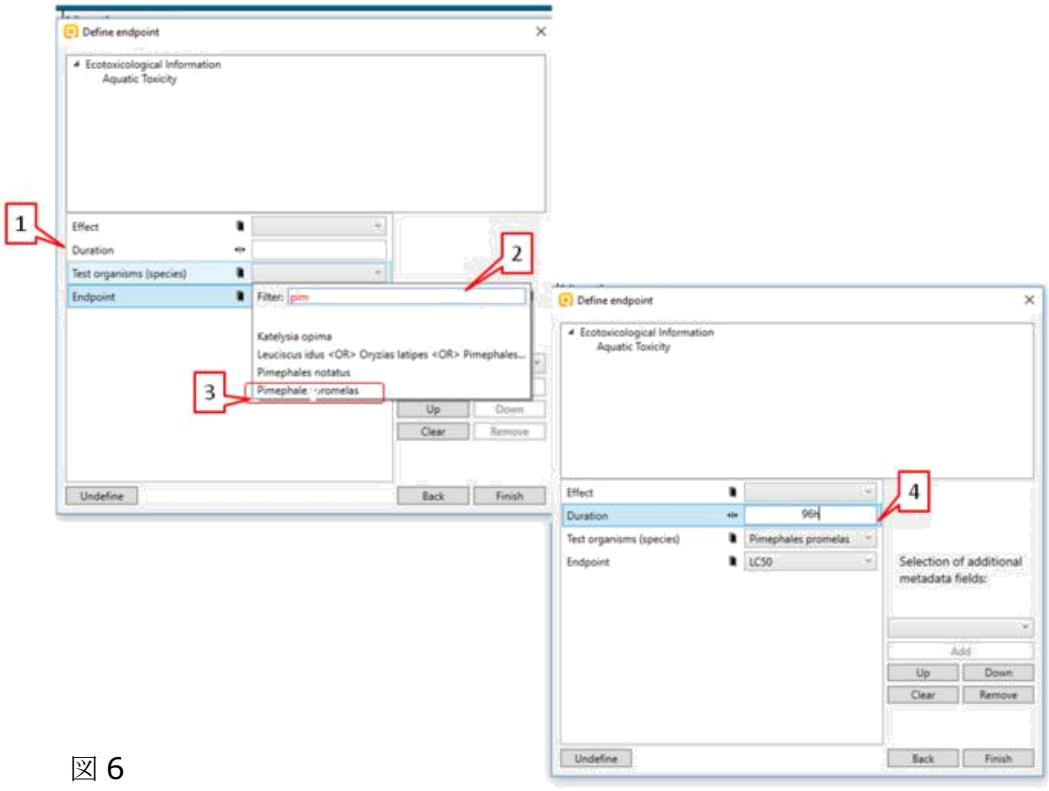


図 6

フィルタ入力欄（2）にキーワードを入力してドロップダウンメニューから Effect（1）を選択し、次に例えば mortality（致死）を選択します。 （図 7）。

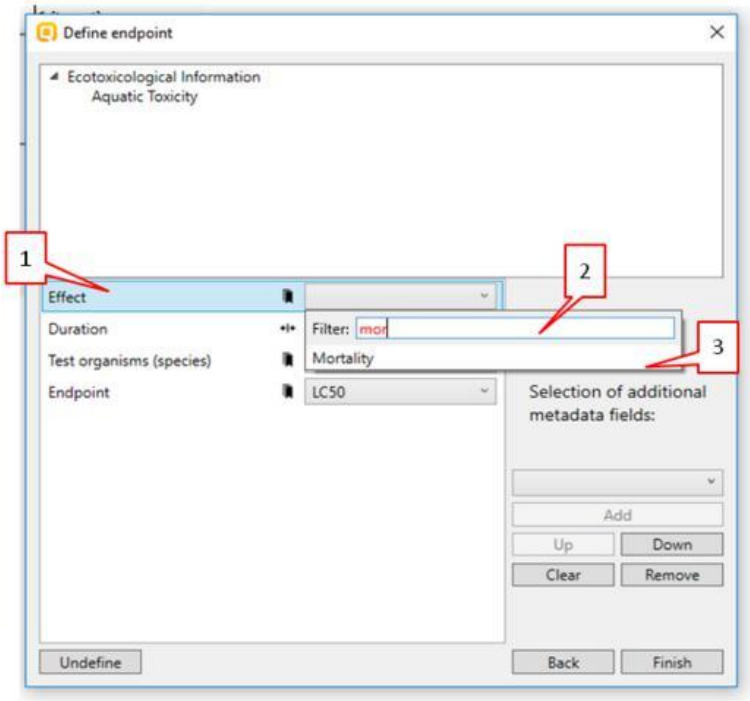


図 7

追加の付随データ入力欄は、ドロップダウンメニュー（1）から追加できます（図 7）。フィルタ入力欄（2）を使用してキーワード、例えばガイドラインなどを入力します。次に Add (3)をクリックします。(図 8－A)

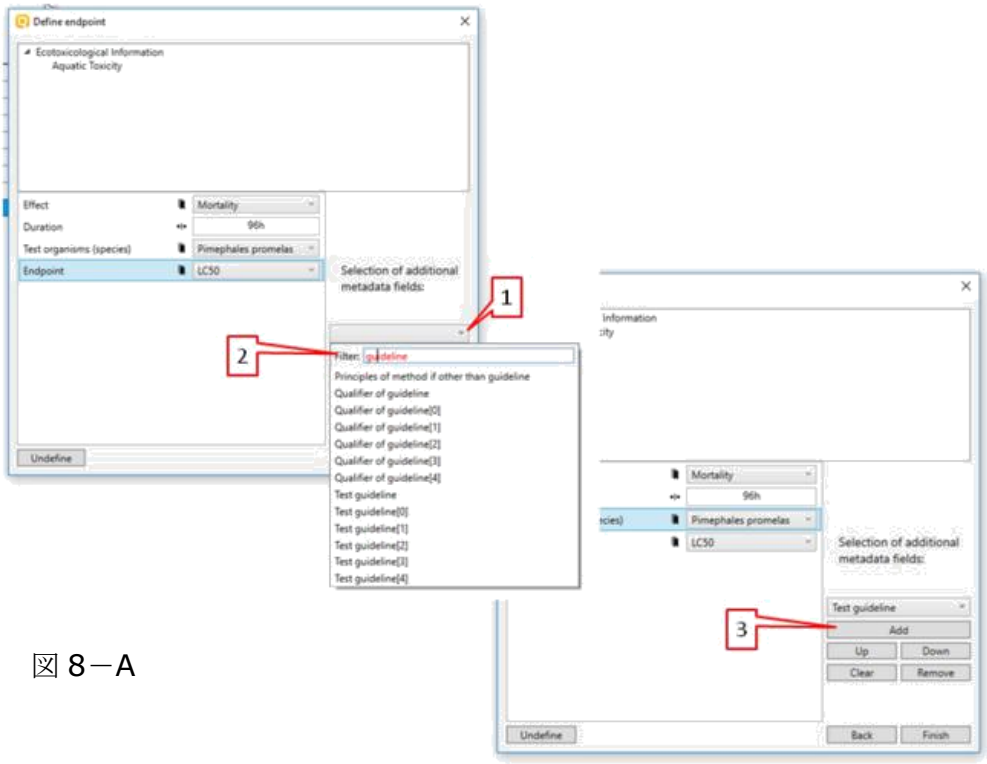


図 8－A

エンドポイントツリーが展開され、図８－B に示すように選択したエンドポイントが**データ・マトリックス**内で黄色く強調表示されます。エンドポイントツリーのリーフは、上記のエンドポイントの定義セクションで指定された階層に従います。

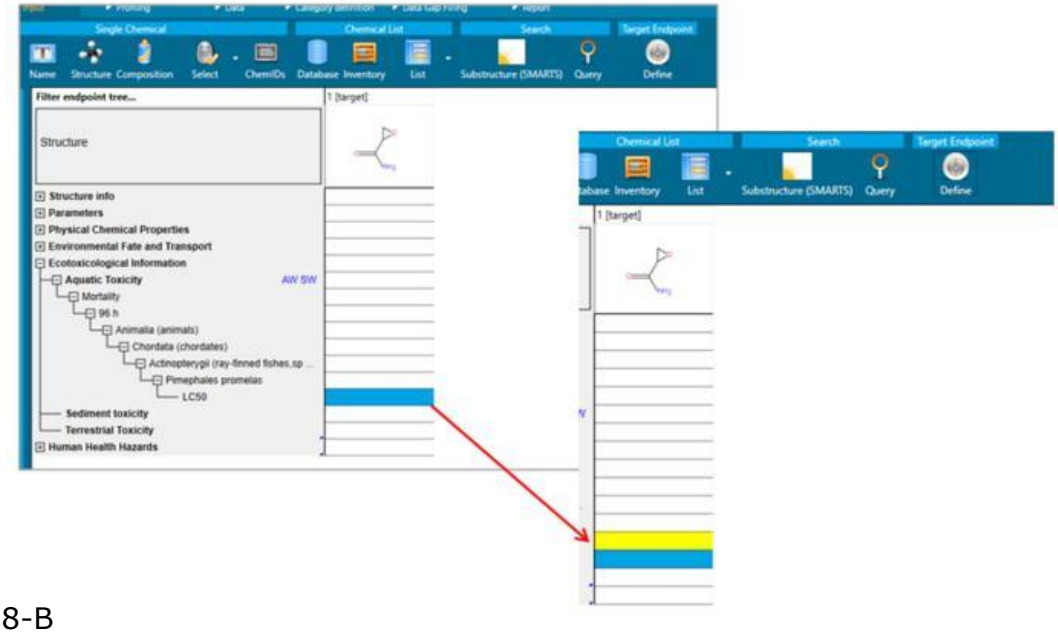


図 8-B

標的エンドポイントを決定する別の方法は、エンドポイントツリーから選択することです。まず、カテゴリを作成し、データを収集する必要があります（データの収集の章を参照）。たとえば、Skin Sensitization エンドポイントです。次に、通常であれば最も実験値を多く持つエンドポイントを選択します。図 8 では、EC3（1）のデータが収集されています。リーフの横で右クリックし、Target endpoint（2）を選択し、 Define（3）をクリックします。

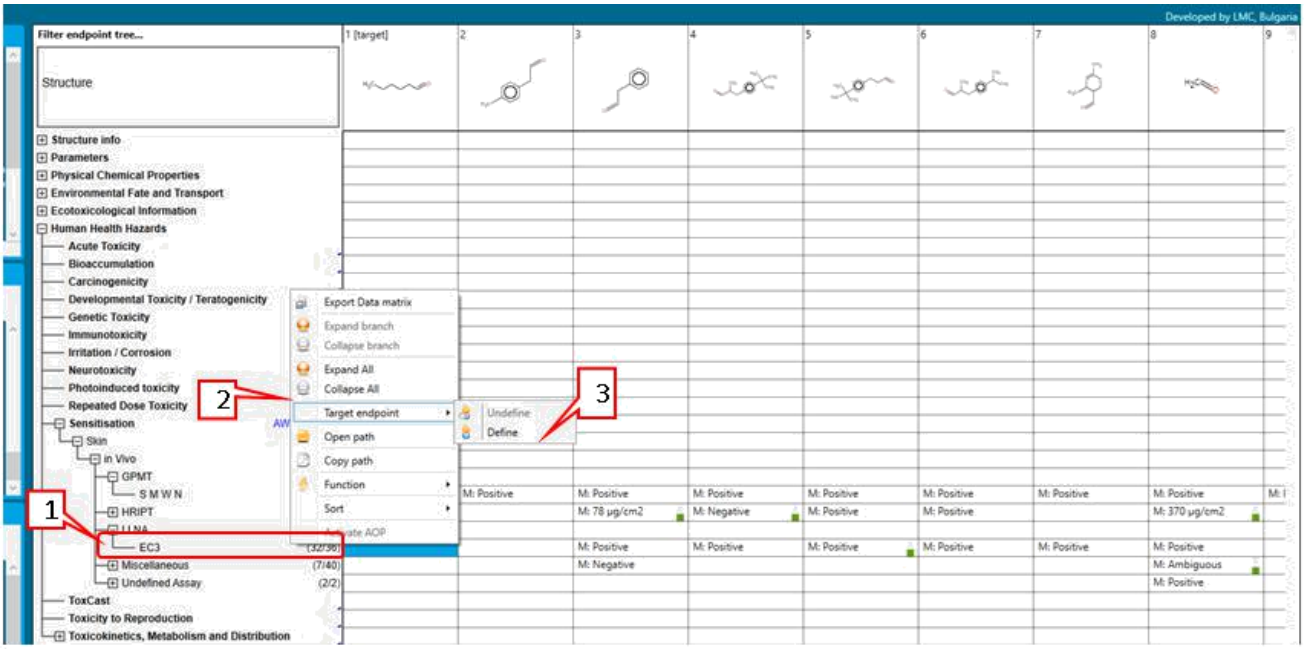


図 9

Define endpoint ウィンドウでは、すべての必須フィールドがすでに選択されています。(図 10)

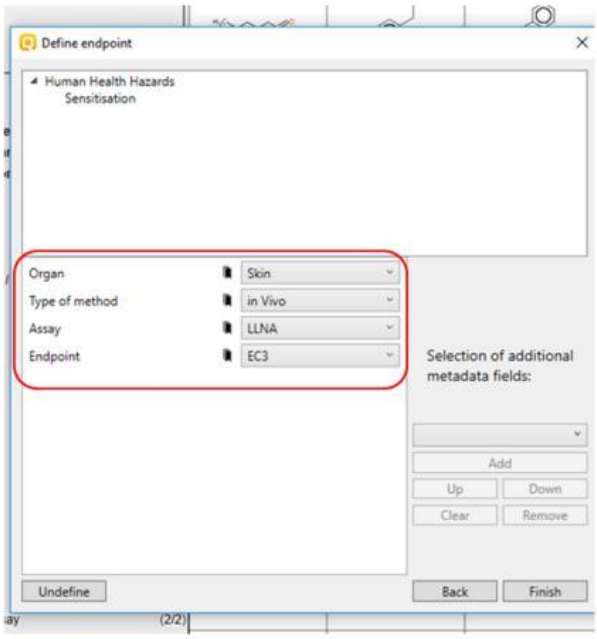


図 10

標的エンドポイントの付随データレベルの階層を変更することができます。たとえば、Endpoint（1）をクリックし、Add（2）をクリックすると、Endpoint が上部（3）に移動します（図 11）。最後に OK をクリックします（4）。

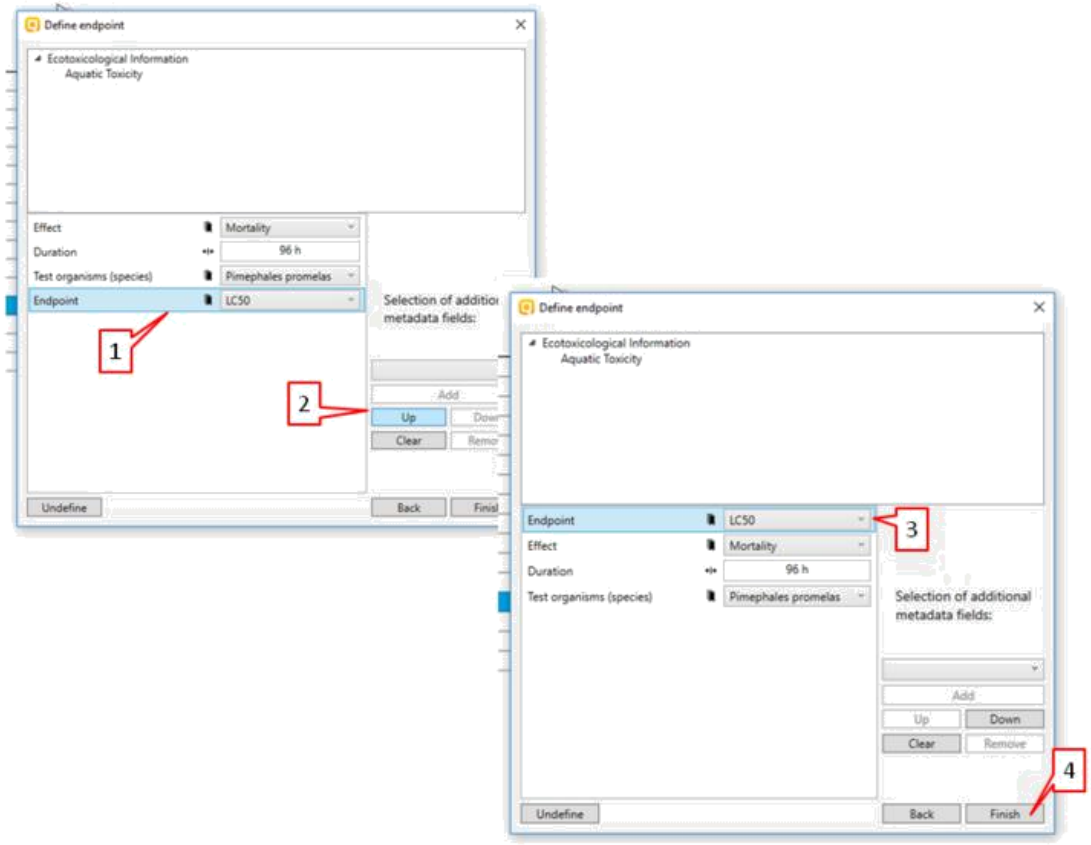


図 11

階層を変更することによって、試験生物種（ここでは Pimephalis promelas）がツリーの中で下のリーフになり、それが標的エンドポイントです（図 12）。

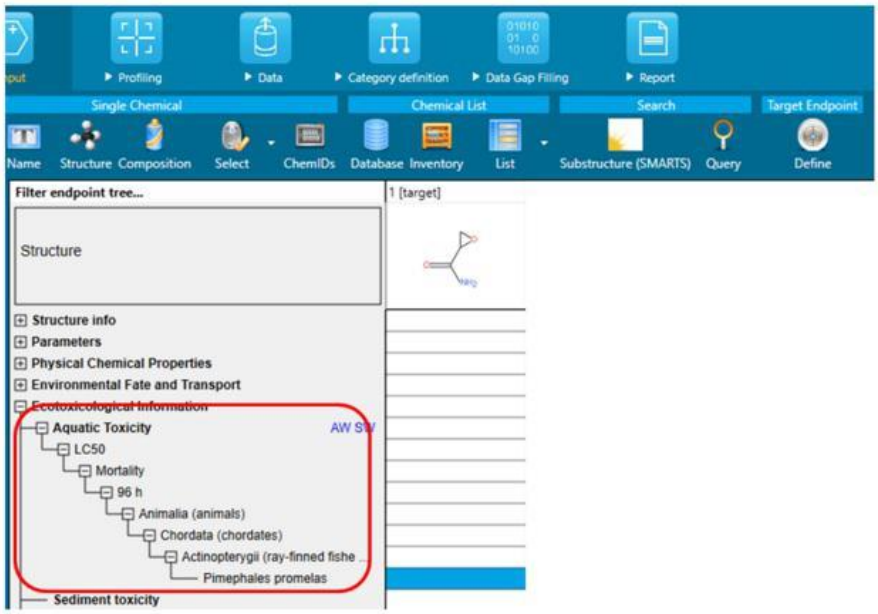


図 12

標的エンドポイントを定義解除するには、Define（1）をクリックした後 Undefine ボタン（2）をクリックするか、または灰色のフィールド（3）のエンドポイントの隣をクリックして、Target endpoint（4）をクリックし、Undefine（5）をクリックします（図 13）。

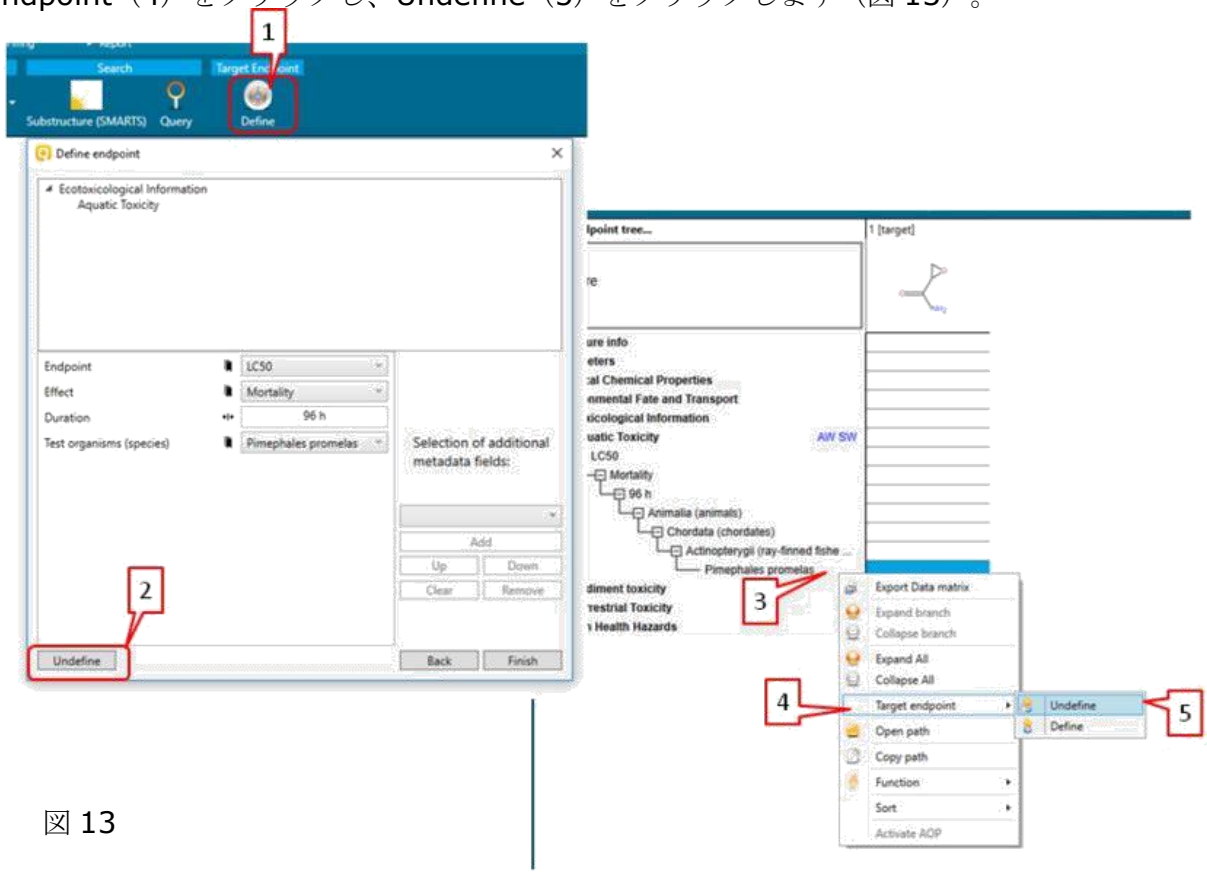


図 13

D.1.4.1. in vitro（イン・ビトロ；試験管内試験）での遺伝子突然変異

組み合わせが可能な in vitro 遺伝子突然変異のパスを以下に示します。組み合わせが可能なデータベースを色付け表示しています（Databases の項を参照）。
（訳注：表の左側青の部分で調査したい事柄の組み合わせを選んだとき、そのデータを入手可能なデータベースを表の右側緑色の部分で示しています）

Predefined （既定のもの）	Endpoint （エンドポイント）	Type of method （手法の種類）	Test type （試験の種類）	Test organisms (species) （試験生物種）	Metabolic activation （代謝活性）	Strain （種）	Bacterial mutageni -city ISSSTY	Genotoxi -city OASIS	ECVAM Genotoxicty and Carcinogenicity	Toxicity Japan MHLW	ECHA CHEM
Genetic Toxicity			Bacterial reverse mutation assay (e.g. Ames test)				relevant DB	relevant DB	relevant DB	relevant DB	relevant DB
Genetic Toxicity	Gene mutation	In Vitro					relevant DB	relevant DB	relevant DB	relevant DB	
Genetic Toxicity	Gene mutation	In Vitro	Bacterial reverse mutation assay (e.g. Ames test)				relevant DB	relevant DB	relevant DB	relevant DB	
Genetic Toxicity	Gene mutation	In Vitro	Bacterial reverse mutation assay (e.g. Ames test)	Salmonella typhimu -rioum			relevant DB	relevant DB		relevant DB	
Genetic Toxicity	Gene mutation	In Vitro	Bacterial reverse mutation assay (e.g. Ames test)	Salmonella typhimu -rioum	Without S9		relevant DB	relevant DB		relevant DB	
Genetic Toxicity	Gene mutation	In Vitro	Bacterial reverse mutation assay (e.g. Ames test)	Salmonella typhimu -rioum	With S9		relevant DB	relevant DB		relevant DB	
Genetic Toxicity	Gene mutation	In Vitro	Bacterial reverse mutation assay (e.g. Ames test)	Salmonella typhimu -rioum	No S9 info			relevant DB			
Genetic Toxicity	Gene mutation	In Vitro	Bacterial reverse mutation assay (e.g. Ames test)	Escherichia coli						relevant DB	
Genetic Toxicity	Gene mutation	In Vitro	Bacterial reverse mutation assay (e.g. Ames test)		Without S9		relevant DB	relevant DB	relevant DB	relevant DB	
Genetic Toxicity	Gene mutation	In Vitro	Bacterial reverse mutation assay (e.g. Ames test)		With S9		relevant DB	relevant DB	relevant DB	relevant DB	
Genetic Toxicity	Gene mutation	In Vitro	Bacterial reverse mutation assay (e.g. Ames test)		No S9 info			relevant DB	relevant DB		
Genetic Toxicity	Gene mutation	In Vitro	Bacterial reverse mutation assay (e.g. Ames test)	Undefined Test organisms (species)					relevant DB		
Genetic Toxicity	Gene mutation	In Vitro	Bacterial reverse mutation assay (e.g. Ames test)	Salmonella typhimu -rioum		TA 100	relevant DB	relevant DB		relevant DB	

Genetic Toxicity	Gene mutation	In Vitro	Bacterial reverse mutation assay (e.g. Ames test)		TA 98	relevant DB	relevant DB		
Genetic Toxicity	Gene mutation	In Vitro	Bacterial reverse mutation assay (e.g. Ames test)		TA 98, TA 100		relevant DB		
Genetic Toxicity	Gene mutation	In Vitro	In vitro mammalian cell micronucleus test				relevant DB		
Genetic Toxicity	Gene mutation	In Vitro	Mammalian cell gene mutation assay				relevant DB	relevant DB	
Genetic Toxicity	Gene mutation	In Vitro	Mammalian cell gene mutation assay	With S9			relevant DB	relevant DB	
Genetic Toxicity	Gene mutation	In Vitro	Mammalian cell gene mutation assay		T-lymphoma cell lines		relevant DB		

D.1.4.2. in vivo（イン・ビボ；生体内試験）での遺伝子突然変異

組み合わせが可能な in vivo 遺伝子突然変異のパスを以下に示します。組み合わせが可能なデータベースを色付け表示しています（Databases の項を参照）。

Predefined （既定のもの）	Endpoint （エンドポイント）	Type of method （手法の種類）	Test type （試験の種類）	Test organisms (species) （試験生物種）	ECVAM Genotoxicty and Carcinogenicity	Transgenic Rodent Database
Genetic Toxicity	Gene mutation	In Vivo			relevant DB	relevant DB
Genetic Toxicity	Gene mutation	In Vivo	Transgenic Rodent Mutation			relevant DB
Genetic Toxicity	Gene mutation	In Vivo	Transgenic Rodent Mutations		relevant DB	
Genetic Toxicity	Gene mutation	In Vivo		Rat		relevant DB
Genetic Toxicity	Gene mutation	In Vivo		Mouse		relevant DB

D.1.4.3. in vitro（イン・ビトロ）での染色体異常

組み合わせが可能な in vitro 染色体異常のパスを以下に示します。組み合わせが可能なデータベースを色付け表示しています（Databases の項を参照）。

Predefined（既定のもの）	Endpoint（エンドポイント）	Type of method（手法の種類）	Test type（試験の種類）	Test organisms (species)（試験生物種）	Metabolic activation（代謝活性）	Strain（種）	Genotoxicity OASIS	ECVAM Genotoxicty and Carcinogenicity	Toxicity Japan MHLW	ECHA CHEM
Genetic Toxicity	Chromosome aberration						relevant DB	relevant DB	relevant DB	relevant DB
Genetic Toxicity	Chromosome aberration	In Vitro					relevant DB	relevant DB	relevant DB	
Genetic Toxicity	Chromosome aberration	In Vitro	In Vitro Mammalian Chromosome Aberration Test				relevant DB	relevant DB	relevant DB	
Genetic Toxicity	Chromosome aberration	In Vitro	In Vitro Mammalian Chromosome Aberration Test	Chinese hamster			relevant DB		relevant DB	
Genetic Toxicity	Chromosome aberration	In Vitro	In Vitro Mammalian Chromosome Aberration Test	Chinese hamster	Without S9		relevant DB			
Genetic Toxicity	Chromosome aberration	In Vitro	In Vitro Mammalian Chromosome Aberration Test	Chinese hamster	With S9		relevant DB			
Genetic Toxicity	Chromosome aberration	In Vitro	In Vitro Mammalian Chromosome Aberration Test	Chinese hamster	No S9 info		relevant DB			
Genetic Toxicity	Chromosome aberration	In Vitro			No S9 info		relevant DB	relevant DB		
Genetic Toxicity	Chromosome aberration	In Vitro	Micronucleus assay					relevant DB		

D.1.4.4. in vivo（イン・ビボ）での染色体異常

組み合わせが可能な in vivo 染色体異常のパスを以下に示します。組み合わせが可能なデータベースを色付け表示しています（Databases の項を参照）。

Predefined (既定のもの)	Endpoint (エンドポイント)	Type of method (手法の種類)	Test type (試験の種類)	Test organisms (species) (試験生物種)	Predefined (既定のもの)	Genotoxicity OASIS	ECVAM Genotoxicty and Carcinogenicity	Toxicity Japan MHLW	Micronucleus OASIS	Micronucleus ISSMIC
Genetic Toxicity	Chromosome aberration	In Vivo				relevant DB	relevant DB	relevant DB	relevant DB	relevant DB
Genetic Toxicity	Chromosome aberration	In Vivo	Micronucleus assay			relevant DB	relevant DB	relevant DB	relevant DB	relevant DB
Genetic Toxicity	Chromosome aberration	In Vivo	Micronucleus assay	Mouse			relevant DB	relevant DB	relevant DB	relevant DB
Genetic Toxicity	Chromosome aberration	In Vivo	Micronucleus assay	Rat			relevant DB		relevant DB	relevant DB
Genetic Toxicity	Chromosome aberration	In Vivo	Micronucleus assay	Rat, mouse			relevant DB			
Genetic Toxicity	Chromosome aberration	In Vivo	Micronucleus assay	Rat, hamster			relevant DB			
Genetic Toxicity	Chromosome aberration	In Vivo	Micronucleus assay	Rabit					relevant DB	
Genetic Toxicity	Chromosome aberration	In Vivo	Micronucleus assay	Human					relevant DB	
Genetic Toxicity	Chromosome aberration	In Vivo	Micronucleus assay	Mammalia					relevant DB	
Genetic Toxicity	Chromosome aberration	In Vivo	Micronucleus assay	Hamster			relevant DB			
Genetic Toxicity	Chromosome aberration	In Vivo	Micronucleus assay	Mouse	Bone marrow cells			relevant DB		
Genetic Toxicity	Chromosome aberration	In Vivo	Micronucleus assay		Bone marrow cells			relevant DB		
Genetic Toxicity	Chromosome aberration	In Vivo	In Vivo Mammalian Chromosome Aberration Test				relevant DB			

D.1.4.5. Skin sensitization（皮膚感作性）

組み合わせが可能な in vivo 皮膚感作性のパスを以下に示します。組み合わせが可能なデータベースを色付け表示しています（Databases の項を参照）。

Predefined (既定のもの)	Endpoint (エンドポイント)	Assay (試験の種類)	Type of method (手法の種類)	Organ (器官)	Skin Sensiti-zation	REACH Skin sensitisation database (normalised)	Dendritic cells COLIPA	Keratinocyte gene expression Givaudan	Keratinocyte gene expression LuSens	Skin Sensiti-zation ECETOC	ECHA CHEM
Sensitisation	ABC	Miscellaneous	in Vivo	Skin	relevant DB						
Sensitisation	CD54	Dendritic Cell Activity (h-CLAT)	in Vitro	Skin			relevant DB				
Sensitisation	CD86	Dendritic Cell Activity (h-CLAT)	in Vitro	Skin			relevant DB				
Sensitisation	CD86	Dendritic Cell Activity (mMUSST)	in Vitro	Skin			relevant DB				
Sensitisation	CD86	Dendritic Cell Activity (MUSST)	in Vitro	Skin			relevant DB				
Sensitisation	EC3	LLNA	in Vivo	Skin	relevant DB	relevant DB					
Sensitisation	EC1.5	KeratinoSens	in Vitro	Skin				relevant DB			
Sensitisation	EC1.5	LuSens	in Vitro	Skin					relevant DB		
Sensitisation	EC2	KeratinoSens	in Vitro	Skin				relevant DB			
Sensitisation	EC2	LuSens	in Vitro	Skin					relevant DB		
Sensitisation	EC3	KeratinoSens	in Vitro	Skin				relevant DB			
Sensitisation	LOEL	HRIPT	In Vivo	Skin	relevant DB						
Sensitisation	NOEL	HRIPT	In Vivo	Skin	relevant DB						
Sensitisation	Respiratory sensitisation		In Vivo	Respiratpry Tract						relevant DB	
Sensitisation	SMWN	GPMT	In Vivo	Skin	relevant DB						
Sensitisation	SWAN	Miscellaneous	In Vivo	Skin	relevant DB						
Sensitisation	Skin sensitisation		In Vivo	Skin	relevant DB					relevant DB	relevant DB
Sensitisation	Skin Sensitisation	Buehler Test	In Vivo	Skin							relevant DB
Sensitisation	Skin Sensitisation	Guinea Pig Maximisation Test	In Vivo	Skin							relevant DB
Sensitisation	Skin Sensitisation	Mouse Local Lymphnode Assay (LLNA)	In Vivo	Skin							relevant DB

D.1.4.6. 他のエンドポイント

組み合わせが可能な他のパスを以下に示します。組み合わせが可能なデータベースを色付け表示しています（Databases の項を参照）。

Predefined (既定のもの)	Endpoint (エンドポイント)	Type of method (手法の種類)	Test type (試験の種類)	Test organisms (species) (試験生物種)	Genotoxicity -city OASIS	ECVAM Genotoxicty and Carcinogenicity	Carcinogenic Potency DB (CPDB)	Carcinogenicity and mutagenicity ISSCAN	Biocides amd Plant protection ISSBIOC
Genetic Toxicity	DNA and protein damage	In Vivo	Liver micronucleus		relevant DB				
Genetic Toxicity	DNA damage and repair	In Vivo	Single cell gel electrophoresis (comet) assay		relevant DB	relevant DB			
Genetic Toxicity	DNA damage and repair	In Vivo	Unscheduled DNA Synthesis			relevant DB			
Carcinogenicity	Carcinogenicity					relevant DB			
Carcinogenicity	Summary carcinogenicity						relevant DB	relevant DB	relevant DB
Carcinogenicity	Summary carcinogenicity			Mouse			relevant DB	relevant DB	relevant DB
Carcinogenicity	Summary carcinogenicity			Rat			relevant DB	relevant DB	relevant DB
Carcinogenicity	Summary carcinogenicity			Hamster			relevant DB		
Carcinogenicity	TD50						relevant DB	relevant DB	
Carcinogenicity	TD50			Mouse			relevant DB	relevant DB	
Carcinogenicity	TD50			Rat			relevant DB	relevant DB	

D.2. プロファイリング

"Profiling"モジュールを使用すると、化学物質の識別情報やその化学構造に基づいて情報を検索できます。

Toolbox には現在、既定のカテゴリへの属性を認識するプロファイリング手法（プロファイラ）の大規模なリストがあります。代謝物の実測値や予測値を考慮に入れるものが含まれています。

それぞれのプロファイラの説明は、目的のプロファイラ（1）を選択し、About ボタン（2）をクリックすることで確認できます（図 1）。大半のプロファイラの詳細については、pdf 形式（3）で提供されています。

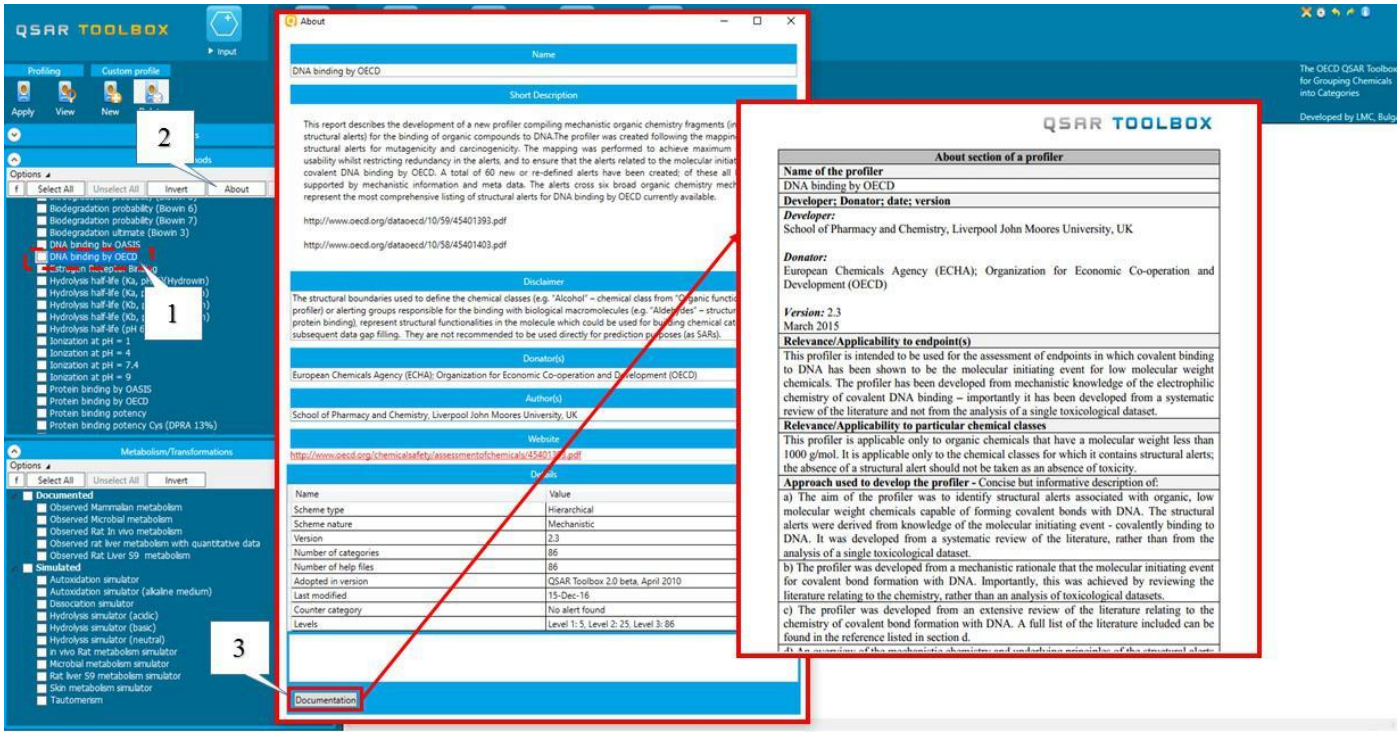


図 1

プロファイラ内のカテゴリの詳細な説明もあります。

D.2.1. Profiling methods（プロファイリング手法）

この章では、以下の項目を記載しています。

- プロファイラの分類
- 各プロファイラの概要について
- 各プロファイラの説明

Profiling methods（プロファイリング手法、プロファイラ）は、既定のカテゴリ（例えば、新しい化学物質の評価用に米国環境保護庁が規定するカテゴリや OECD HPV 化学物質プログラムで評価しているカテゴリ）、反応メカニズムや作用機序との関係を特定します。また、プロファイリングの結果により類似物質を検索する最適な方法が決まります。

実際、すべてのプロファイラがすべての規制エンドポイントに関連するわけではありません。タンパク質結合などの一般的な反応メカニズムプロファイラは多くのエンドポイント（皮膚感作性、急性毒性、水生毒性）に関連していますが、一部のプロファイラは特定のエンドポイントのみのカテゴリを構築するようできています。

プロファイリング手法は、化学物質評価に使用される定性的情報の種類に基づいて 5 つのグループに分かれています。

- Predefined（既定のカテゴリ）、
- General mechanistic（一般的な反応メカニズム）、
- Endpoint specific （エンドポイント特化型）、
- Empiric（経験則的）、
- Toxicological（毒性学的）、
- Custom（カスタム）

ユーザは自分用のカスタムプロファイラを作成することもできます。カスタムプロファイラの作成の詳細については、Creating a custom profiler のセクションを参照してください。

D.2.1.1. Predefined（既定のカテゴリ）

Predefined(既定の)カテゴリとは規制当局が定めた､化学物質をグループ化する規則の群であり、「スタンダードな」カテゴリ化の手法とみなすことができます。
ツールボックスは以下の既定のプロファイラを備えています:

- Database affiliation（データベース属性）
- Inventory affiliation（インベントリ属性）
- OECD HPV Chemicals categories（OECD 高生産量化学物質カテゴリ）
- Substance type（物質の種類）
- US-EPA New Chemicals categories（米国環境保護庁の新規化学物質カテゴリ）

上に挙げた各プロファイラの簡単な説明を以下に示します。

Database affiliation（データベース属性）

このプロファイラは、化学物質（データベースに属しているもの）のデータベース属性に関する情報を提供します。化学物質があるデータベースに属しているからと言って、そのデータベースから実験データが入手できるとは限りません。あるデータベースに属しているということは、その物質の ID（SMILES 表記を利用できれば）が当該データベースに属していることを意味します。

Inventory affiliation（インベントリ属性）

このプロファイラは、化学物質（インベントリに属しているもの）のインベントリ属性に関する情報を提供します。

OECD HPV Chemicals categories（OECD 高生産量化学物質カテゴリ）

このプロファイラは、2010 年 4 月までに評価された OECD 高生産量化学物質のリストに基づいて開発されました。それは少なくとも 1 つの加盟国/地域で、年間 1000 トン以上の水準で生産または輸入される化学物質です。このプロファイラは、OECD 高生産量化学物質プログラムに対応して、国が人の健康と環境に対する有害性の評価を行う化学物質を選択することをサポートします。
これらの有害性評価は、OECD 既存化学物質プログラムの基本的な部分であり、スクリーニング情報データセット（SIDS）の比較的制限されたデータ要素に基づいています。その目的は潜在的な有害性について化学物質をスクリーニングすることであり、これにより懸念される化学物質のさらなる研究にリソースを注ぐことができます。

Substance type（物質の種類）

システムは以下のどの物質の種類かを特定します: # - 個々の化学物質 # - 解離性化学物質 # - 混合物 # - ポリマー # - 組成不明の化学物質です。特性または組成物質によって、物質はこれらの 5 つの物質タイプの 1 つに分類されます。

US-EPA New Chemicals categories（米国環境保護庁の新規化学物質カテゴリ）

米国環境保護庁（EPA）の新規化学物質プログラムのプロファイラでコード化された規則によって、米国環境保護庁公害防止・毒物局の公式文書である「TSCA 新化学物質プログラム（NCP）/化学物質のカテゴリ」に引用されているオリジナルのカテゴリが再現されています。

しかし、カテゴリは非常に多様でその境界は主に物理的考察に基づいており、化学構造を基本としたルールは含まれていないため、すべてのカテゴリがコード化されているわけではありません。これらのカテゴリは次のとおりです:

- Category: Acid Dyes and Amphoteric Dyes（酸性染料と両性染料）
- Category: Cationic Dyes（陽イオン性染料）
- Category: Polyanionic Polymers (& Monomers) {ポリ陰イオン性ポリマー（&モ
- ノマー）}
- Category: Polycationic Polymers（ポリ陽イオン性ポリマー）
- Category: Respirable, Poorly soluble Particulates（吸入性、難溶性微粒子）

D.2.1.2. General mechanistic（一般的な反応メカニズム）

反応メカニズムの分類スキームは、公開情報に基づいて重要な化学特性を識別するためのルールで構成されています。実質的に、これらは研究された現象に関する仮説に依存しています。この点において、これらの適応は、開発目的の現象に傾倒しています。

ツールボックスは以下の一般的な反応メカニズムの分類スキームを備えています:

- DNA binding by OASIS（DNA 結合、OASIS）
- DNA binding by OECD（DNA 結合、OECD）
- Estrogen receptor binding（エストロゲン受容体結合）
- Protein binding by OASIS（タンパク質結合、OASIS）
- Protein binding by OECD（タンパク質結合、OECD）
- Protein binding potency（タンパク質結合能）
- Protein binding potency Cys (DPRA 13%)（タンパク質結合能 システイン DPRA13%）
- Protein binding potency Lys (DPRA 13%)（タンパク質結合能 リジン DPRA 13%）
- Biodeg BioHC half-life (Biowin)（生分解 BioHC 半減期 Biowin）
- Biodegradation probability (Biowin 1)　（生分解確率 Biowin 1）
- Biodegradation probability (Biowin 2)（生分解確率 Biowin 2）
- Biodegradation probability (Biowin 5)（生分解確率 Biowin 5）
- Biodegradation probability (Biowin 6)（生分解確率 Biowin 6）
- Biodegradation probability (Biowin 7)（生分解確率 Biowin 7）
- Biodegradability ultimate (Biowin 3)（生分解性 最終生成物 Biowin 3）
- Biodegradability primary (Biowin 4)（生分解性 一次反応生成物 Biowin 4）
- Hydrolysis half-life (Ka, pH7)（加水分解半減期　Ka、pH7）
- Hydrolysis half-life (Ka, pH8)（加水分解半減期　Ka、pH8）
- Hydrolysis half-life (Kb, pH7)（加水分解半減期　Kb、pH7）
- Hydrolysis half-life (Kb, pH8)（加水分解半減期　Kb、pH8）
- Hydrolysis half-life (pH 6.5-7.4)（加水分解半減期　pH 6.5-7.4）
- Ionization at pH=1（pH=1 におけるイオン化）
- Ionization at pH=4（pH=4 におけるイオン化）
- Ionization at pH=7.4（pH=7.4 におけるイオン化）
- Ionization at pH=9（pH=9 におけるイオン化）
- Toxic hazard classification by Cramer (original)（Cramer の毒性有害性分類、オリジナル）
- Toxic hazard classification by Cramer (with extension)（Cramer の毒性有害性分類、拡張版）

上に挙げた各プロファイラの簡単な説明を以下に示します。

DNA binding by OASIS（DNA 結合、OASIS）

このプロファイラは、OASIS TIMES システムの AMES 変異原性モデル部分に基づいています。このプロファイラは、AMES 変異原性モデルで解析された DNA との相互作用の原因の 85 の構造アラートで構成されています。このプロファイラの調査範囲は、DNA と相互作用しうる標的分子内の構造アラートがあるかどうか調べることです。85 個の構造アラートのリストは、8 つの**反応メカニズム**のドメインに分かれています。**反応メカニズム**の各ドメインは、**反応メカニズム**的アラートに分かれています。アラートのうち 31 件は更新されています。プロファイリングを行うと、標的物質はそれに対応する構造アラート、**反応メカニズム**アラートと**反応メカニズム**のドメインに割り当てます。

DNA binding by OECD（DNA 結合、OECD）

この説明文は、有機化学物質フラグメント反応メカニズム（構造的アラートの形で）をコンパイルした、DNA に結合する有機化合物に関する新しいプロファイラの開発について記載しています。
このプロファイラは、変異原性と発がん性に関する既存の構造アラートのマッピングにより作成されました。このマッピングにより、アラートの重複を制限しつつも最大限の重なりとユーザビリティが達成され、DNA 共有結合の分子開始事象に関連するアラートを保証しています。合計 60 件の新しいアラートや再定義されたアラートが作成されました。 そのうち 2 つを除いた全アラートが反応メカニズム的情報とメタデータに支持されています。このアラートは、6 つの広範な有機化学反応メカニズムにまたがり、現在入手可能な OECD の DNA 結合の構造アラートの最も包括的なリストを提供しています。

Estrogen receptor binding（エストロゲン受容体結合）

エストロゲン受容体（ER）結合は、タンパク質結合と非常によく似た分子開始事象です。これはいくつかの包括的なデータベースが存在するエンドポイントです。これにより ER 結合とそれに伴う可能性のある内分泌かく乱を予測するための、（Q）SAR を用いた各アプローチが開発されています。Toolbox に組み込んだ ER 結合プロファイリングスキームは、文献情報源から抽出され、実験データに支持された構造的・パラメータ的なルールに基づいています。この ER 結合プロファイラは、分子
量（MW）と構造的特徴によって、その化学物質を非結合性物質・結合性物質に分類します：

- 非常に強力な結合性物質：200～500Da の分子量であって、ヒドロキシル基が 2 つの環状構造のどちらかに結合している化学物質。
- 強力な結合性物質：遮蔽されていないヒドロキシル基またはアミノ基のある 5 員または 6 員炭素環を少なくとも 1 つ有し、200～500Da の分子量の化学物質。
- 中程度の結合性物質：遮蔽されていないヒドロキシル基 またはアミノ基のある 5 員または 6 員炭素環を少なくとも 1 つ有し 170～200Da の分子量の化学物質。
- 弱い結合性物質：遮蔽されていないヒドロキシル基またはアミノ基のある 5 員または 6 員炭素環を少なくとも 1 つ有し、170Da 未満の分子量の化学物質。

もし対象化学物質が上記の構造的・パラメータ的要件のどれかを満たさない場合、それは非結合性物質に分類されます：

- ヒドロキシル基またはアミノ基が機能していない非結合性物質
- 500 Da 以上の分子量である非結合性物質
- ヒドロキシル基またはアミノ基のない非結合性物質
- 環状構造を持たない非結合性物質

Protein binding by OASIS （タンパク質結合、OASIS）

このタンパク質結合アラートは、ExxonMobil、P&G、Unilever、香料材料研究所（RIFM）、Dow とデンマーク国立食品研究所を含む産業コンソーシアムがブルガス数理科学研究所と共に D.Roberts 博士と共同で、皮膚感作性を予測する TIMES モデルの一部として開発しました。L'Oreal は、Mekenyan 教授（OASIS-LMC）と 2007 年に締結した研究契約において、QSAR Toolbox に提供する化学物質カテゴリの評価と改善に貢献しました。このプロファイラの調査範囲は、タンパク質との相互作用に関与する標的分子内に構造アラートがあるかどうか調べることです。101 個の構造アラートのリストがあり、11 の反応メカニズムのドメインに分かれています。反応メカニズムの各ドメインは、2 つ以上の反応メカニズム的なアラートに分かれています。プロファイリングを行うと、標的物質は該当する構造アラート、反応メカニズムのドメイン、反応メカニズムアラートに割り振られます。

Protein binding by OECD（タンパク質結合、OECD）

Protein binding by OECD プロファイラは、理論有機化学に基づいた直接作用の構造アラートの分析によって開発されました（このプロファイラには代謝的/非生物的に活性化した構造アラートは含まれません）。アラートを集積して分析し、文献に含まれる情報を化学反応メカニズムの枠組みに入れました。この Protein binding byOECD プロファイラを利用する際、この化学反応メカニズムを化学物質のカテゴリ形成の基礎として使用することができます。5 つの各反応メカニズムドメインにおいて、共通した反応

性部位を持っている関連構造アラートは、いわゆる反応メカニズムアラートに分類されています。化学物質カテゴリを形成するには、このプロファイラを使用して**反応メカニズム**アラートまたは構造アラートレベルのいずれかで実行できます。この Protein binding by OECD プロファイラには、52 個の構造アラートを含む 16 個の反応メカニズムアラートが含まれています。これらのデータは、化学反応メカニズムと文献情報をで支持されています。

Protein binding potency（タンパク質結合能）

このプロファイラは in chemico RC50 値で表されるチオール反応性の実測値に基づき開発されています。データは標的化学物質とグルタチオン（GSH）のチオール基との共有結合を測定することによって得られています。タンパク質結合の構造アラートは、GSH　RC50 実測値が判明している約 400 種の化学物質から引用されます。 化学物質は全て、SN2 反応を介した GSH 相互作用と、Michael 付加反応メカニズム（MA）を介した相互作用の 2 つの共通する求電子性反応メカニズムを有しています。

プロファイラには、49 個の MA と 46 個の SN2 カテゴリが含まれています。合わせて 95 種類になる構造アラートは 5 段階の結合能カテゴリに分類されています : Extremely（極端に強い）、Highly（強い）、Moderately（中程度）、Slightly reactive（わずかに反応）、Suspect（結合能を持つか疑わしい）です。結合能カテゴリの分類は以下の通りです : Extremely reactive（極めて反応性が高い ; RC50 < 0.099 mmol / L）、Highly reactive（高反応性 ; RC50 = 0.100 ~ 0.990 mmol / L）、Moderately reactive（中程度の反応性 ; RC50 = 1.000 ~ 15.000 mmol / L）、Slightly reactive（わずかに反応性 ; RC50 = 16.000 ~ 70.000 mmol / L）、Suspect（結合能は疑わしい ; RC50 = 71.000 ~ 135.000 mmol / L）。プロファイリングを行うと、標的物質は化学構造的な基準に基づいて、該当する結合能カテゴリに振分けられます。

Biodeg BioHC half-life（生分解 BioHC 半減期、 Biowin）

BioHC half-life (Biowin)プロファイラは、BioHCwin プログラムで実施される個々の石油系炭化水素の定量的一次生分解半減期を決定するためのモデルに基づいています。このモデルは、フラグメントベースの手法を使用し、複数の線形回帰を使用して半減期（日）を推定します。

Biodegradation probability (Biowin 1) （生分解確率 Biowin 1）

化学物質の生分解確率は、線形回帰式による各フラグメントの値（フラグメント係数）を合計し、定数を加算することによって計算されます。計算された確率により以下のように生分解性を判定します:

- 0.5 以上の確率→ Biodegrades Fast（迅速に生分解する）
- 0.5 未満の確率→ Does NOT Biodegrades（生分解しない）

Biodegradation probability (Biowin 2)（生分解確率 Biowin 2）

この Biodegradation probability （Biowin 2）プロファイラは、非線形回帰式による各フラグメントの値（フラグメント係数）を合計し、定数を加算することによって、化学物質の生分解確率を計算します。計算された確率により以下のように生分解性を判定します:

- 0.5 以上の確率→ Biodegrades Fast（迅速に生分解する）
- 0.5 未満の確率→ Does NOT Biodegrades（生分解しない）

Biodegradability ultimate (Biowin 3)（生分解性最終生成物 Biowin 3）

このモデルは、「完全に」生分解を完了させるのに必要な時間を予測します。生分解の最終形態とは、親物質が二酸化炭素と水、試験化合物中に存在する他のあらゆる元素の無機酸化物、および新しい細胞材料に変換したことです。Biowin 3 は典型的な水生環境や「評価用」の水生環境で完全な最終生分解を達成するために必要な時間を予測します

Biodegradability primary (Biowin 4)（生分解性 一次反応生成物 Biowin 4）

このモデルは、一次生分解が完了するのに必要な時間を予測します。一次生分解とは、親物質が一次代謝物に変換することです。Biowin 4 は典型的な水生環境や「評価用」の水生環境で一次生分解を達成するために必要な時間を予測します。

Biodegradation probability (Biowin 5)（生分解確率 Biowin 5）

Biodegradation probability (Biowin 5) プロファイラは、EPISUIT の Biowin 5 のモデルに基づいています。Biowin5 は、日本の MITI （旧：通商産業省、現：経済産業省）の良性生分解試験、すなわち OECD 301C における化合物の生分解性を評価するための予測モデルのことです。生分解性の予測は、重線形回帰分析を使用して開発されたフラグメント定数に基づいています。

- 0.5 以上の確率→ Readily Degradable（良分解性である）
- 0.5 未満の確率→ Does NOT Biodegrades（良分解性でない）

Biodegradation probability (Biowin 6)（生分解確率 Biowin 6）

Biodegradation probability (Biowin 6) プロファイラは、EPISUIT の Biowin 6 のモデルに基づいています。Biowin 6 は、日本の MITI （旧：通商産業省、現：経済産業省）の良性生分解試験、すなわち OECD 301C における化合物の生分解性を評価するための予測モデル のことです 。生分解性の予測は、そのモデルに応じて重非線形回帰分析を使用して開発されたフラグメント定数に基づいています。

- 0.5 以上の確率→ Readily Degradable（良分解性である）
- 0.5 未満の確率→ Does NOT Biodegrades（良分解性でない）

Biodegradation probability (Biowin 7)（生分解確率 Biowin 7）

Biodegradation probability (Biowin 7) プロファイラは、EPISUIT の Biowin 7 のモデルに基づいています。Biowin 7 はメタン生成の嫌気性条件下での迅速な生分解の確率を予測します。具体的には、「血清ボトル」嫌気性生分解スクリーニング試験の条件下です。このエンドポイントは、典型的な嫌気性消化装置における分解を予測するものと仮定されています。

- 0.5 以上の確率→ Readily Degradable（良分解性である）
- 0.5 未満の確率→ Does NOT Biodegrades（良分解性でない）

Hydrolysis half-life (Ka, pH7)（加水分解半減期　Ka、pH7）

Hydrolysis half-life (Ka, pH7)プロファイラは Hydrowin に基づき、全酸触媒加水分解速度定数で加水分解の半減期を推定します。酸触媒加水分解速度定数の半減期は pH7 条件下で計算されます。

Hydrolysis half-life (Ka, pH8)（加水分解半減期　Ka、pH8）

Hydrolysis half-life (Ka, pH8)プロファイラは Hydrowin に基づき、全酸触媒加水分解速度定数で加水分解の半減期を推定します。酸触媒加水分解速度定数の半減期は pH8 条件下で計算されます。

Hydrolysis half-life (Kb, pH7)（加水分解半減期　Kb、pH7）

Hydrolysis half-life (Kb, pH7)プロファイラは Hydrowin に基づき、全塩基触媒加水分解速度定数で加水分解の半減期を推定します。塩基触媒加水分解速度定数の半減期は pH7 条件下で計算されます。

Hydrolysis half-life (Kb, pH8)（加水分解半減期　Kb、pH8）

Hydrolysis half-life (Kb, pH8)プロファイラは Hydrowin に基づき、全塩基触媒加水分解速度定数で加水分解の半減期を推定します。塩基触媒加水分解速度定数の半減期は pH8 条件下で計算されます。

Hydrolysis half-life（加水分解半減期　pH 6.5-7.4）

このモデルは次の実験条件で、個々の有機化学物質の加水分解の半減期を予測します。中性・ほぼ中性の pH、大気温または室温、大気圧下です。加水分解プロファイラは、半減期に応じて化学物質をカテゴリに分類します。半減期の計算値により加水分解は次のようにランク付けされます。 very slow（非常に

遅い、> 100 日）、slow（遅い、30-100 日）、moderate（中程度、10-30 日）、fast （速い、5-10 日）、very fast（非常に速い、1-5 日）、extremely fast （非常に速い、1 日未満）。 (US EPA, REACH, D. Aronson, R. Boethling, P. Howard, W. Stiteler, Estimating biodegradation half-lives for use in chemical screening, Chemosphere, 2006, 63(11), 1953-1960)

Ionization at pH=1（pH=1 におけるイオン化）

この Ionization at pH=1 プロファイルは、pH=1 での分子のイオン化成分の部分濃度（イオン化、%）を計算します。任意の化学物質について、このプロファイルは予め計算された pKa 値に基づいて、最も強い酸性と最も強い塩基性サイトのイオン化のみから生じるイオン化成分の部分濃度を計算します。計算された部分濃度を使い、化学物質をイオン化 10 区分の範囲、すなわち[0,10]、[10,20] ... [90,100]に分類します。この範囲は、イオン化部位の特性とその pKa 値に基づいて、酸・塩基性の水準を示しています。

注意：このプロファイルは、pKa 計算値のデータベースに基づいています。したがって、このプロファイルは Toolbox データベースに含まれる化学物質にのみ適用できます。

Ionization at pH=4（pH=4 におけるイオン化）

この Ionization at pH=4 プロファイルは、pH=4 での分子のイオン化成分の部分濃度（イオン化、%）を計算します。任意の化学物質について、このプロファイルは予め計算された pKa 値に基づいて、最も強い酸性と最も強い塩基性サイトのイオン化のみから生じるイオン化成分の部分濃度を計算します。計算された部分濃度を使い、化学物質をイオン化 10 区分の範囲、すなわち[0,10]、[10,20] ... [90,100]に分類します。この範囲は、イオン化部位の特性とその pKa 値に基づいて、酸・塩基性の水準を示しています。

Ionization at pH=7.4（pH=7.4 におけるイオン化）

この Ionization at pH=7.4 プロファイルは、pH=7.4 での分子のイオン化成分の部分濃度（イオン化、%）を計算します。任意の化学物質について、このプロファイルは予め計算された pKa 値に基づいて、最も強い酸性と最も強い塩基性サイトのイオン化のみから生じるイオン化成分の部分濃度を計算します。計算された部分濃度を使い、化学物質をイオン化 10 区分の範囲、すなわち[0,10]、[10,20] ... [90,100]に分類します。この範囲は、イオン化部位の特性とその pKa 値に基づいて、酸・塩基性の水準を示しています。

Ionization at pH=9（pH=9 におけるイオン化）

この Ionization at pH=9 プロファイルは、pH=9 での分子のイオン化成分の部分濃度（イオン化、%）を計算します。任意の化学物質について、このプロファイルは予め計算された pKa 値に基づいて、最も強い酸性と最も強い塩基性サイトのイオン化のみから生じるイオン化成分の部分濃度を計算します。計算された部分濃度を使い、化学物質をイオン化 10 区分の範囲、すなわち[0,10]、[10,20] ... [90,100]に分類します。この範囲は、イオン化部位の特性とその pKa 値に基づいて、酸・塩基性の水準を示しています。

Toxic hazard classification by Cramer (original)（Cramer の毒性有害性分類、オリジナル）

この Toxic hazard classification by Cramer プロファイラは G.M. Cramer and R.A.Ford の発案した論文、有害性評価 - ディシジョンツリーに基づいています(Food and Cosmetics Toxicology, Volume 16, Issue 6, December 1978, Page 255-276.)。化学物質をさまざまなレベルの毒物学的懸念（経口投与された場合）に分類するカテゴリ化ルールは、ツリー状の構造となっています。ディシジョンツリーは 33 の化学構造規則を備えており、次の 3 つの分類のいずれかに化合物を割振ります:

- Low (Class I)（低：分類Ｉ）
- Intermediate(Class II) （中程度：分類ＩＩ）

- High (Class III)（高：分類ＩＩＩ）

このカテゴリのうち 2 つは ToxTree v.2.1.0 から借用した「一般的な食物の成分」の 440 種類の化合物と「体の通常の成分」の 107 種類の化合物を含む外部ファイルを使用しています。

Toxic hazard classification by Cramer (with extension)(Cramer の毒性有害性分類、拡張版)

この Toxic hazard classification by Cramer (with extension)プロファイラは G.M. Cramer and R.A.Ford の発案した論文、有害性評価 - ディシジョンツリーに基づいています。(Food and Cosmetics Toxicology, Volume 16, Issue 6, December 1978, Page 255-276.)化学物質をさまざまなレベルの毒物学的懸念(経口投与された場合)に分類するカテゴリ化ルールは、ツリー状の構造となっています。ディシジョンツリーは 33 の化学構造規則を備えており、次の 3 つの分類のいずれかに化合物を割振ります:

- Low (Class I)（低：分類Ｉ）
- Intermediate(Class II)　（中程度：分類ＩＩ）
- High (Class III)（高：分類ＩＩＩ）

このカテゴリのうち 2 つは ToxTree v.2.1.0 から借用した「一般的な食物の成分」の 440 種類の化合物と「体の通常の成分」の 107 種類の化合物を含む外部ファイルを使用しています。

Protein binding potency Cys (DPRA 13%)（タンパク質結合能 システイン DPRA 13%)

このプロファイルは、皮膚感作性の有害性発現経路（AOP）を導入して構築されています。ペプチド結合性試験（DPRA）から得られたデータの基礎に基づいて開発されました。DPRA は、化学物質がタンパク質と反応する能力を評価する結合性試験です。還元型グルタチオンと 2 種類の合成ペプチド、すなわちリジンとシステインをモデルペプチドとして使用しています。リシンとシステインとの反応時間は両方とも 24 時間です。ペプチド結合性は、ペプチド減少分がパーセントで報告されます。このプロファイルは、システイン減少の実測値を有する約 229 種の化学物質から抽出された 77 個の構造アラートを備えています。この 77 個の構造アラートのセットは、次の 3 つのカテゴリに分類されます。DPRA が 21%以上（DPRA 13%）、DPRA は 9%未満（DPRA 13%）、グレーゾーン 9-21%（DPRA 13%）です。結合能カテゴリの分類は、L'Oreal（Dimitrov et al., 2016）との共同研究で発表された分析法に基づいています。

Protein binding potency Lys (DPRA 13%)（タンパク質結合能 リジン DPRA 13%)

このプロファイルは、皮膚感作性の有害性発現経路（AOP）を導入して構築されています。ペプチド結合性試験（DPRA）から得られたデータの基礎に基づいて開発されました。DPRA は、化学物質がタンパク質と反応する能力を評価する結合性試験です。還元型グルタチオンと 2 種類の合成ペプチド、すなわちリジンとシステインをモデルペプチドとして使用しています。リジンとシステインとの反応時間は両方とも 24 時間です。ペプチド結合性は、ペプチド減少分がパーセントで報告されます。このプロファイルは、リジン減少の実測値を有する約 228 種の化学物質から抽出された 73 個の構造アラートを備えています。この 73 個の構造アラートのセットは、次の 3 つのカテゴリに分類されます。DPRA が 21%以上（DPRA 13%）、DPRA は 9%未満（DPRA 13%）、グレーゾーン 9-21%（DPRA 13%）です。結合能カテゴリの分類は、L'Oreal（Dimitrov et al., 2016）との共同研究で発表された分析法に基づいています。

D.2.1.3. Endpoint specific　（エンドポイント特化型）

エンドポイントに特化した以下のカテゴリ化のスキームが Toolbox に備えられています:

- Acute aquatic toxicity classification by Verhaar (Modified)（Verhaar 急性水生生物毒性分類修正版）
- Acute aquatic toxicity MOA by OASIS（OASIS の急性水生生物毒性作用機序）
- Aquatic toxicity classification by ECOSAR　（ECOSAR の水生生物毒性の分類）有害性発現
- Bioaccumulation - metabolism alerts（生物蓄積 - 代謝アラート）
- Bioaccumulation - metabolism half-lives（生物蓄積 - 代謝半減期）
- Biodegradation fragments (BioWIN MITI)　（生分解性フラグメント BioWIN MITI）
- Carcinogenicity (genotox and nongenotox) alerts by ISS（ISS、発がん性アラート – 遺伝毒性・非遺伝毒性）
- DART scheme（発生毒性物質スキーム）
- DNA alerts for AMES by OASIS（OASIS、AMES 試験の DNA アラート）
- DNA alerts for CA and MNT by OASIS（OASIS、CA（染色体異状）と MNT（小核試験）の DNA アラート）
- Eye irritation/corrosion Exclusion rules by BfR（BfR、目の刺激/腐食の除外ルール）
- Eye irritation/corrosion Inclusion rules by BfR（BfR、目の刺激/腐食の包含ルール）
- in vitro mutagenicity (Ames test) alerts by ISS（ISS、in vitro 突然変異誘発 AMES 試験アラート）
- in vivo mutagenicity (Micronucleus) alerts by ISS（ISS、in vivo 突然変異誘発 微小核アラート）
- Keratinocyte gene expression（ケラチン生成細胞遺伝子発現）
- Oncologic Primary Classification（Oncologic、一次分類）
- Protein binding alerts for Chromosomal aberration by OASIS（OASIS 染色体異常のタンパク質結合アラート）
- Protein binding alerts for skin sensitization according to GHS（GHS 皮膚感作のタンパク質結合アラート）
- Protein binding alerts for skin sensitization by OASIS（OASIS 皮膚感作のタンパク質結合アラート）
- Protein Binding Potency h-CLAT（h-CLAT によるタンパク質結合能）
- Respiratory sensitization（呼吸器感作）
- Retinoic Acid Receptor Binding（レチノイン酸受容体結合）
- rtER Expert System ver.1 - US EPA（rtER エキスパートシステム ver.1 - 米国環境保護庁）
- Skin irritation/corrosion Exclusion rules by BfR（BfR の皮膚の刺激/腐食の除外ルール）
- Skin irritation/corrosion Inclusion rules by BfR（BfR の皮膚の刺激/腐食の包含ルール）

上に挙げた各プロファイラの簡単な説明を以下に示します（訳注：以下は原文の並び順に従っているため ABC 順ではない）。

Acute aquatic toxicity classification by Verhaar (Modified)（Verhaar の急性水生生物毒性分類　改定版）

この Acute aquatic toxicity classification by Verhaar (Modified) は、Toxtree ソフトウェアが開発した Verhaar 則を適応するため LMC が開発したパラメータ的・構造的なルールを備えています。この分類システムは Toxtree バージョン 2.5 改訂版 Verhaar スキーム部分に基づいています(Verhaar ら、1992, Enoch ら、2008)。 このシステムは、カテゴリ化導入部分や、後続の試験を実施する化学物質の優先順位付けに使用することができます。さらにこれらの予測は、危険性および有害性評価において大きな価値を持つ可能性があります。多数の有機化学物質を作用機序に基づいて 4 つに分類することにより、化学物質の構造とその急性水生生物毒性との間の定量的関係を特定したりすることができます。4 つの分類とは、**(1)**

不活性な化学物質（ベースライン毒性において）、**（2）**（1）より不活性ではない化学物質、**（3）**反応性化学物質、 **（4）**特に反応性の強い化学物質です。

Acute aquatic toxicity MOA by OASIS（OASIS の急性水生生物毒性作用機序）

このプロファイラは、化学物質をその急性毒性作用機序（MOA）に従って様々なカテゴリに分類します。2D 化学構造の情報のみに基づき、化学物質の作用機序（MOA）が特定されます。理論的・経験的知識に基づき、MOA は以下の 7 つの階層に順序づけられ区別されています：アルデヒド、α、β-不飽和アルコール、フェノールおよびアニリン、エステル、ナルコティックアミン、ベースラインナルコティクス。

Aquatic toxicity classification by ECOSAR （ECOSAR の水生生物毒性の分類）

この Aquatic Toxicity Classification by ECOSAR プロファイラは、LMC と OECD が米国環境保護庁の生態系構造 - 活動プログラム (ECOSAR™) 内におけるケミカルクラスの構造的定義を適用するために開発した分子決定の機能を備えています。ECOSAR™には、化合物の評価に適したケミカルクラスを選択するための専門家のルールに基づくディシジョンツリーを備えた、水生生物毒性を予測するためのケミカルクラスベースの SAR ライブラリが含まれています。ECOSAR™は、現在、118 のケミカルクラスを識別するようにプログラムされています。このプロファイラは定量的な毒性値を予測するためではなく、ECOSAR™の分類定義を使用した化学物質カテゴリ化のために導入
されています。

Bioaccumulation - metabolism alerts（生物蓄積 - 代謝アラート）

生物蓄積とは生体における取り込みと排泄の速度が競合した正味の蓄積結果であり、典型的には BCF・BAF 測定値とモデル予測値を用いて魚類で評価されます。BCFBAF モデルに含まれる魚の生体内変換モデルの詳細情報は、EPI Suite バージョン 4.0 のモデルのヘルプファイルからも入手できます。各化合物中の各フラグメントの出現数に各化合物の logKow と分子量を加えたマトリックスに対して最終的な多重線形回帰を行いました。マトリックスの解の列は、各化合物の生体内変換半減期の実測値の対数であり、付録では個々のフラグメントをリストにしています。多重線形回帰は、CoStatTM 統計ソフトウェアを用いて行いました。このモデルは、スクリーニングレベルで全身の一次生体内変換半減期 HL（日）と、魚体内の各有機化学物質の生体内変換速度定数 kM（1 /日）を見積もります。このモデルには、さまざまな化学構造に広く適用できるように、独自の構造フラグメントの大きなセットが含まれています。しかしながら、これらのフラグメントは有機化学物質で存在しうる構造フラグメントの全ドメインを反映してはいません。魚類における 632 の kM 生体内変換率の実測値セット（対数生体内半減期の単位で集計）は、モデル導出用の 421 化合物のトレーニングセットと、モデル試験のための 211 化合物の検証セットに分けられています。

Bioaccumulation - metabolism half-lives（生物蓄積 - 代謝半減期）

生物蓄積とは生体における取り込みと排泄の速度が競合した正味の蓄積結果であり、典型的には BCF・BAF 測定値とモデル予測値を用いて魚類で評価されます。BCFBAF モデルに含まれる魚類の生体内変換のモデルに関する叙述が Arnot. JA ら（2009）の公表文献に記載されています。BCFBAF モデルに含まれる魚類の生体内変化モデルの詳細情報は、EPI Suite バージョン 4.0 のモデルのヘルプファイルからも入手できます。このモデルは、スクリーニングレベルで全身の一次生体内変換半減期 HL（日）と、魚体内の各有機化学物質の生体内変換速度定数 kM（1 /日）を見積もります。プロファイラは、化学物質を以下の生体内変換速度のカテゴリに分類します。very slow（非常に遅い）、slow（遅い）、moderate（中程度の）、 fast（速い）、very fast（非常に速い）。これにより、同様の生体内変換速度を持つ化学物質を同定することができます。5 つのカテゴリの上限・下限値は、Arnot. JA ら（2009）によって提案されたものです。

Biodegradation fragments (BioWIN MITI) （生分解性フラグメント BioWIN MITI）

BIOWIN は、環境微生物の混合集団の存在下で有機化合物が迅速に好気性・嫌気性生分解をする比率を見積もります。BIOWIN には 7 つの各生分解モデルが含まれています。生分解性の予測は、モデルに依存す

る線形回帰分析または非線形回帰分析から得られたフラグメント係数に基づいています。BIOWIN 生分解性分類手法は、MITI 生分解性確率モデルで使われる構造フラグメントに基づいています。

DNA alerts for AMES by OASIS（OASIS の AMES 試験の DNA アラート）

このプロファイラは、OASIS TIMES システムの AMES 変異原性モデル部分に基づいています。このプロファイラは、AMES 変異原性モデルから抽出された DNA との相互作用を担う 85 の構造アラートに基づいています。このプロファイラの調査範囲は、AMES 変異原性に関する、DNA との相互作用を起こす標的分子内の構造アラートがあるかどうか調べることです。このプロファイラは、DNA と相互作用するアラートをもつ化学物質が電子的・立体的な要因のために相互作用しなくなることが考慮されています。それは、いくつかのアラートに付与されているマスキングによって明確に定義されています。85 個の構造アラートのリストは、8 つの反応メカニズムドメインに分かれています。各反応メカニズムドメインは反応メカニズムアラートに分かれています。アラートのうち 31 件は更新されています。プロファイリングを行うと、標的物質はそれに対応する構造アラート、反応メカニズムアラートと反応メカニズムドメインに割り振られます。

DNA alerts for CA and MNT by OASIS（OASIS の染色体異状と小核試験の DNA アラート）

このプロファイラは、染色体異常モデルから抽出された DNA との相互作用を担う化学物質の 85 の構造アラートに基づいています。ローカルトレーニングセットの化学物質の差異により in vitro AMES モデルと、CA モデルにおける DNA アラートにはわずかな差異があります。このプロファイラの調査範囲は、染色体異常および小核試験と関連する DNA との相互作用を起こす標的分子内の構造アラートがあるかどうか調べることです。このプロファイラは、DNA と相互作用するアラートをもつ化学物質が電子的・立体的な要因のために相互作用しなくなることが考慮されています。それは、いくつかのアラート付与されているマスキングによって明確に定義されています。85 個の構造アラートのリストは、8 つの反応メカニズムのドメインに分かれています。反応メカニズムの各ドメインは反応メカニズム的なアラートに分かれています。プロファイリングを行うと、標的物質はそれに対応する構造アラート、反応メカニズムアラートと反応メカニズムのドメインに割り振られます。

Protein binding alerts for skin sensitization by OASIS（OASIS の皮膚感作のタンパク質結合アラート）

このタンパク質結合アラートは、ExxonMobil、P&G、Unilever、香料品香料原料安全性研究所（RIFM）、Dow とデンマーク国立食品研究所を含む産業コンソーシアムがブルガス数理化学研究所と共に D.Roberts 博士の協力のもと、皮膚感作性を予測する TIMES モデルの一部として開発しました。このプロファイラの調査範囲は、タンパク質、特に皮膚のタンパク質との相互作用を起こす標的分子内の構造アラートがあるかどうか調べることです。このプロファイラは、タンパク質と相互作用するアラートをもつ化学物質が電子的・立体的な要因のために相互作用しなくなることが考慮されています。それは、いくつかのアラートに付与されているマスキングがあるのかどうかによって明確に定義されています。100 個の構造アラートのリストは、11 の反応メカニズムのドメインに分かれています。反応メカニズムの各ドメインは、2 つ以上の反応メカニズム的なアラートに分かれています。プロファイリングを行うと、標的物質はそれに対応する構造アラート、反応メカニズムアラートと反応メカニズムのドメインに割り振られます。

Protein binding alerts for skin sensitization according to GHS（GHS の皮膚感作のタンパク質結合アラート）

このプロファイラは、GHS 基準によって皮膚感作性を予測するために最近開発された OASIS TIMES モデルに基づいています。タンパク質結合アラートは、モデリング用のトレーニングセットとして使用される 517 化学物質から抽出されます。現在のバージョンのプロファイラでは、化学物質を GHS カテゴリ 1A と 1B に分類するアラートが 131 件あります。カテゴリ 1B の領域は、モデルとそれぞれのプロファイラでわずかに修正されています。その結果、分類の上限・下限値は次のようになります。 - カテゴリ 1A [EC3

（LLNA）= 2%。 NOEL（HRIPT）=500 μg/ cm2] - カテゴリ 1B [2% ＜EC3（LLNA）＜50%; 500 μg/ cm2 ＜ EL（HRIPT）＜ 12,500 μg/ cm2] - No alert found（警告なし）[EC3（LLNA）= 50%; 陰性（LLNA）; NOEL（HRIPT）＞12,500 μg/cm2]。 各アラートは反応メカニズムドメイン、例えば Schiff 塩基形成、マイケル付加、アシル化などと関連しており、アラートとなる官能基と皮膚タンパク質との相互作用のメカニズムについてサポートするヘルプファイルに記載されています。このプロファイラは、化学構造的な要求およびパラメータ的（log Kow）な要求で構成されています。

DART scheme（発生毒性物質スキーム）

QSAR Toolbox のこの部分は、以下のジャーナル投稿論文で概説されている、生殖毒性や発生毒性物質として作用しうる化学構造的特徴を持つ、化学物質を同定するためのフレームワークの改作です：Wu S.,Fisher J.,Naciff J.,Laufersweiler M., Lester C., Daston G., Blackburn K. 著「発生毒性物質または生殖毒性物質として作用する可能性に関連する構造的特徴を有する化学物質を同定するためのフレームワーク」Chem Res Toxicol. 2013 Dec 16;26(12):1840-61. 既知の作用機序（MOA）とそれに関連する構造的特徴と MOA 情報が不足している場合の生殖または発生毒性物質（DART）化学物質分子内の構造フラグメントの経験的な関連付けとの組み合わせに基づいて開発されました。

このツールは、DART である可能性について評価された 716 の化学物質（陽性 664、陰性 16、データ不十分 36）の詳細なレビューに基づいて設計されています。これらの化学物質は、既定の受容体結合と化学的特性に基づいて、そして既知の場合にはその MOA に基づいて、25 のさまざまなカテゴリと 129 のサブカテゴリに分類されています。ディシジョンツリーを基に化学物質を評価すると、その化学物質が DART を有すると知られている化学構造に関連しているか関連していないか、 あるいは DART ディシジョンツリーの化学物質の範囲外の化学構造を持っているかがわかります。ディシジョンツリーのは、トレーニングセットに含まれていない化学物質群を対象にして試験され、既知の DART 効果を有する化学物質を高い割合で特定することが確認されています。このディシジョンツリーは独立型ツールとして使用することを意図したものではなく、生殖および/または発育毒性効果の事例の化学物質に類似した特徴を持つ化学物質を広く捕捉することを意図して設計されています。

このツールは、懸念される化学物質を特定するためのスクリーニングシステムの一部分として、また構造活性相関（SAR）に基づく決定の「Weight of evidence（証拠の重み）」の一部として、加で実験をしてデータを取得することなくデータギャップを補完するために使用できます。このツールに記載されている化学構造グループ、ならびにこれらのグループに合致する新しい化学物質は、MOA を明らかにする in vitro 試験の仮説の開発の起点として使用されて最終的には MOA を組み込んだ DART 用 SAR の、原理の改良開発に使用することが期待されます。我々は MOA の見識を強化することで、その結果 DART 陽性と DART 陰性化学物質の分類を促進するディシジョンツリーが今後改善され、このツールにより確実に予測が可能な化学物質ドメインを拡大することを予期しています。また、化学ドメインの適応範囲を拡大するためにさらなるグループがツリーに追加されることを予期しています。

Keratinocyte gene expression（ケラチン生成細胞遺伝子発現）

このプロファイラは、皮膚感作性の有害性発現経路（AOP）の実施に関連して構築されています。これは、ヒトの不死化ケラチノサイトに安定に挿入されたヒト AKR1C2 遺伝子の ARE（訳注：Adenylate - uridylate - rich elements、AU リッチの意）の単一コピーの制御下で、化学物質がルシフェラーゼレポーター遺伝子の発現を誘発する可能性を調べた、KeratinoSens 試験データを基に開発されています。皮膚感作との関連性は、Keap1-Nrf2-ARE 調節経路と、感作に関する求電子性化学物質の検出との関係から推測されています。試験から得られたデータに基づいて、3 つのエンドポイント EC1.5、EC2、EC3（ルシフェラーゼ誘導において 1.5 倍、2 倍、3 倍の増加を誘発する濃度）が報告されています。EC1.5、EC2 および/または EC3 が > 2000 である場合、遺伝子誘導は観察されません。このプロファイラには、EC1.5、EC2、

EC3 データに含まれる約 100 種類の化学物質から抽出された 22 種類の構造アラートが含まれています。22 種類の構造アラートのセットは、4 つのカテゴリに分かれています：非常に高い遺伝子発現、高い遺伝子発現、中程度の遺伝子発現、低い遺伝子発現です。

カテゴリの分類は EC3 値に依存し、以下の通りです：非常に高い遺伝子発現を有する化学物質（EC3 ≦ 15 μM）、高遺伝子発現（EC3 = 15 ~ 50 μM）、中程度の遺伝子発現（EC3 = 50 ~ 100 μM）、低遺伝子発現（EC3 = 100 ~ 1999 μM）です。

プロファイリングを行うと、化学構造基準に適合するかどうかに基づいて、標的物質は該当する結合能カテゴリに割り振られます。

Eye irritation/corrosion Exclusion rules by BfR（BfR の目の刺激/腐食の除外ルール）

眼刺激性/腐食性の除外基準は、眼刺激性や腐食性を示さない化学物質を特定するための物理化学的上限・下限値に基づいています。目の刺激性のルールを規定するために使うパラメータは、脂質溶解度、オクタノール・水分配係数、水溶性、融点、分子量です。

Eye irritation/corrosion Inclusion rules by BfR（BfR の目の刺激/腐食の包含ルール）

このプロファイラは 17 個の構造アラートで構成されており、経験的に導出した化学構造的包含ルールに基づいています。眼の刺激や腐食の可能性がある化学物質を特定します。このプロファイラは既知の作用機序（眼および/または結膜組織内の生化学反応）にも基づいています。このモデルは、少なくとも 95%の純度であり、O、N、S、P、Si、ハロゲンをヘテロ原子として含むがそれ以外の原子は含まない有機物質に適用可能です。

Carcinogenicity (genotox and nongenotox)alerts by ISS（ISS の遺伝毒性および非遺伝毒性発がん性）

このプロファイラは、ソフトウェア Toxtree の該当モジュールの拡張バージョンです。55 の構造アラート（SAs）リストに基づいて、発がん性を推定するためのディシジョンツリーとして機能します。そのうち、35 は Toxtree モジュールに由来し、20 は新たに導出されたものです。新しい SAs のほとんどは非遺伝毒性の発がん性に関連していますが、リストの最初の SA は主に遺伝毒性の発がん性をコードしています。発がん性の SAs とは、発がん性化学物質に関係することが知られている分子官能基または部分構造です。分子構造に含まれている 1 つまたは複数の SAs が認識されると、このシステムは化学物質の発がん性の可能性があると警告します。

in vitro mutagenicity (Ames test) alerts by ISS（ISS の in vitro 突然変異誘発（AMES 試験）アラート）

このプロファイラは、ソフトウェア Toxtree の 変異原性 / 発がん性モジュールに基づいています。これは 30 の構造的アラート（SAs）のリストに基づいて、in vitro（AMES 試験）変異原性を推定するためのディシジョンツリーとして機能します。変異原性の SA は、化学物質の突然変異誘発活性に関係することが知られている分子官能基や部分構造です。分子構造に含まれている 1 つまたは複数の SAs が認識されると、このシステムは化学物質に突然変異を起こす可能性があると警告します。現在の SAs のリストは、非遺伝毒性の発癌性に関する SAs を除外して得られたオリジナルの Toxtree リストのサブセットです。

in vivo mutagenicity (Micronucleus) alerts by ISS（ISS の in vivo 突然変異誘発（小核試験）アラート）

このプロファイラは、ソフトウェア Toxtree の ToxMic ルールベースに基づいています。このルールベースは、潜在的な in vivo 突然変異原を事前にスクリーニングするための 35 の構造アラート（SAs）のリストを提供しています。この SAs とは、in vivo 小核試験における効果の誘導に関係することが知られている分子官能基または部分構造です。ここで提供されているげっ歯類における in vivo 小核試験用の SAs の

集積は、毒性作用機構に関する既存の知識と、この分析法で試験される化学物質の構造解析の両方に基づいたものです。

Retinoic acid receptor（レチノイン酸受容体）

Retinoic acid receptor (RAR) binding とは、Molecular initiating event（分子開始事象）の 1 つです。レチノイドは、胚発生と成人のホメオスタシスにとって重要な内在性化学物質であるビタミン A（レチノール）と機能的に類似した化学物質です。In vivo ではレチノールはレチノイン酸（RA）に酸化され、細胞増殖と分化の制御に関与します。RA はレチノイン酸受容体（RAR）とレチノイド X 受容体（RXR）との相互作用を介して遺伝子転写を変調し、それぞれ 3 つのアイソタイプ（α、β、γ）として発現します。胚において RA は、胚葉形成と、心臓、膵臓、指、耳、眼、肺の器官形成を統制します。レチノイドの影響で RAR シグナル伝達が過度に活性化すると、心臓血管、CNS、四肢、耳、下顎、泌尿生殖器の欠損を含むさまざまな発達上の欠陥をもたらす可能性があります。Wnt への影響は、状況（RAR 配位子を含む）に特異的で RAR アイソフォーム間で異なるように思われます [Yasuhara et al., 2010]。催奇形性の正確なメカニズムは不明であり、RAR シグナル伝達は、多くの遺伝子標的を制御する早期胚発生において複雑な役割を有しています[Kam et al., 2012]。これらのカテゴリは、（開発中の）RAR 受容体活性化作用の有害性発現経路を開発するために公開された文献の調査と、各カテゴリに引用された文献からの in vitro 試験データに基づいて開発されました。カテゴリに入らなかったが活性である類似物質は、別の結合機構を介して作用すると考えられています。

Respiratory sensitization（呼吸器感作）

このプロファイラは、低分子量化学物質の呼吸器感作性の評価に使用します。このプロファイラは、呼吸器感作を誘発する段階の**反応メカニズム**の知識から開発され、それにより肺の中のタンパク質に共有結合する化学物質を同定します。この**反応メカニズム**の仮定によって、このプロファイラは皮膚と肺の両方を介して呼吸器感作を誘発しうる化学物質を同定するのに適したものとなっています。（特定の構造アラートを持つ化学物質は感作の誘導期と誘発期の両方で同一でなければならない）。

Protein binding alerts for Chromosomal aberration（染色体異常のタンパク質結合アラート）

このプロファイラは、トポイソメラーゼ、細胞タンパク質付加物などの特定のタンパク質と、化学物質の相互作用の原因となる 33 の構造的アラートに基づいています。

明確な反応メカニズムの根拠と関連して、これらのアラートは第 2 の反応性要素（DNA 反応性を相補する）として、 in vitro 染色体異常 OASIS TIMES 変異原性モデルに含まれています。このプロファイラの調査範囲は、標的分子が染色体異常を誘発する能力を調べることです。分子内で立体的（または電子的）障害を起こすためにタンパク質との相互作用を妨げる性質は明確に定義されており、「阻害」マスクとしていくつかのアラートに関連付けられています。

Oncologic Primary Classification（腫瘍の一次分類）

この OncoLogic Primary Classifier プロファイラは、LMC と OECD が開発した分子定義で構成されており、米国環境保護庁の発がん性予測用の OncoLogic™がんエキスパートシステムが対象としている、発癌性物質の化学的分類の構造基準を適用しています。QSAR Toolbox では、OncoLogic Primary Classifier は OncoLogic™分類の規定に基づき、分類する目的でのみ使用されます。このプロファイラは、発がん性を予測するためではなく、分類目的のために導入されています。

rtER Expert System ver.1 - USEPA（rtER エキスパートシステム ver.1 - 米国環境省）

この rtER Expert System ver.1 - USEPA プロファイラは、米国 EPA エストロゲン受容体エキスパートシステム（ERES）の適用範囲である、エストロゲン受容体結合剤の可能性を持つ化学物質分類の構造基準に従った分子決定からなります。この ERES プロファイラは、エストロゲン受容体結合親和性の予測に使用される、影響ベースの自動化システムです。この ERES はもともと既定の規制目的に対応するために開発されたもので、特にステロイド型化学構造を持つ化学物質を含まず、（したがってより高い親和性の ER 相互作用をすることができない）2 つの特別なインベントリ、すなわち食品用農薬不活性物（FI）および抗菌性農薬（AM）に含まれる化学物質をの優先順位付けのために開発されました。このシステムは、既定の

規制インベントリをカバーする化学物質トレーニングセットに基づいて構築されており、陰性だと予測された化学物質は ER に結合する可能性が低いという確証を高めるために、分析中に化学物質溶解度や細胞毒性まで試験することによって、結合のあらゆる兆候を拾い上げるように特別に最適化された in vitro 分析を使用しています。

化学物質分類ベースのアプローチは、影響ベースの化学物質カテゴリとリード・アクロスの概念を採用することによって、特徴がよく知られている限られた数のトレーニングセット化学物質から、より広範囲の化学物質のインベントリに外挿することができるように設計されました。この ERES は、明確に定義されたエンドポイントまたは ER バイオ分析ドメインの化学的・生物学的側面に関して専門家の反応メカニズムの合意をコード化する、論理的ルールベースのディシジョンツリーです。ERES を開発するアプローチにおいて、意図している目的に使うシステムの透明性（予測される化学物質と試験された化学物質との関係）や、有用性（FI および AM 化学物質に提供される予測）が強調されました。例えば ERES 化学物質群に見出されている、相対結合親和性（RBA）と LogKow との関係は、予測している化学物質がトレーニングセット化学物質の領域内であることを確実にするために各群内で使用されました。活性であるか不活性であるか予測可能な既知の範囲の外に分類される化学物質は、「未知の結合力」（UnkBP）を有するとみなされていました。自動化バージョンの ERES を使うと、予測される化学物質を各化学群（すなわち、ディシジョンツリーのノード）内のトレーニングセット化学物質と比較することができます。ツールボックスでは、rtER Expert System ver.1 - USEPA プロファイラは、オリジナルの ERES 化学群の構造定義に基づいてカテゴリ化するために使用されます。この rtER Expert System ver.1 - US-EPA プロファイラは、相対的結合親和性（RBA）を予測するためではなく、カテゴリ化目的のために導入されています。

Skin irritation/corrosion Exclusion rules by BfR（BfR の皮膚の刺激/腐食の除外ルール）

皮膚刺激性/腐食性の除外基準は、皮膚刺激性や腐食性を示さない化学物質を特定するための物理化学的上限・下限値に基づいています。BfR 新化学物質データベースにある 1,358 の化学物質を用いて、物理化学的限界ルールを導出します。これらのルールは、すべての物質に適用できるルールと特定の化学物質サブ分類に適用できるルールに分割されます（詳細はプロファイラの基本モードで利用可能です）。皮膚の刺激性のルールを規定するために使うパラメータは、脂質溶解度、表面張力、オクタノール・水分配係数、蒸気圧、水溶性、融点、分子量です。

Skin irritation/corrosion Inclusion rules by BfR（BfR の皮膚の刺激/腐食の包含ルール）

プロファイラには、その**反応メカニズム**によって刺激、腐食、またはその両方を引き起こす化学物質の積極的な分類に使用できる構造アラートが含まれています。これら 40 種類の有機構造フラグメントは、今のところ皮膚刺激と腐食の包絡ルールをカバーしています。

D.2.1.4. Empiric（経験則的）

経験則的カテゴリ化手法には、標的化学物質を構成する化学元素や、その化学官能基性、化学構造類似性を決定するために使用される規則が含まれます。

ツールボックスは以下の既定の経験則的カテゴリ化手法を備えています:

- Chemical elements（化学元素），
- Groups of elements（元素周期表の「族」），
- Lipinski Rule Oasis（Lipinski 則 、Oasis），
- Organic functional groups（有機官能基），
- Organic functional groups (nested)（有機官能基・入れ子状），
- Organic functional groups (US-EPA)（有機官能基・米国環境保護庁），
- Organic functional groups, Norbert Haider (checkmol)（有機官能基、Norbert Haider・Checkmol）
- Structure similarity（化学構造の類似性）
- Tautomers unstable（不安定な互変異性体）

上に挙げた各プロファイラの簡単な説明を以下に示します。

Chemical elements（化学元素）

このプロファイラには、18 の族で構成された周期表のすべての化学元素が含まれています。1 族と 2 族には、アルカリとアルカリ土類化学元素があります。アクチノイドとランタノイドは、周期表におけるそれらの位置により 2 分割した 3 族として示されています。周期表内で 3〜12 族に格納されている遷移金属は、周期表の族番号と同じ様に、族に編成されています。
13 族には半金属と金属が示されています。Sn、Pb と半金属の Si、Ge は炭素とともに第 14 族に位置しています。周期表の第 15 族の金属と半金属は、3 つの別々のカテゴリに含まれます。酸素、硫黄、セレンは、3 つの別々のカテゴリに分類されて 16 族と呼ばれています。ハロゲン原子（F、Cl、Br、I、At）は 17 族に位置しています。18 族には希ガス（He、Ne、Ar、Kr、Xn、Rn）が含まれています。

Groups of elements（元素周期表の「族」）

このプロファイラには、9 つのカテゴリに分類された周期表のすべての化学元素が含まれています：
アルカリとアルカリ土類化学物質は、最初の 2 つのカテゴリに分類されています。
ハロゲン原子（F、Cl、Br、I、At）は、「ハロゲン」カテゴリに表示されます。
半金属元素のホウ素（B）、シリコン（Si）、ゲルマニウム（Ge）、ヒ素（As）、アンチモン（Sb）、テルル（Te）、ポロニウム（Po）はカテゴリ「Metalloids（半金属）」に位置しています。
金属および非金属元素は、2 つの別々のカテゴリに分類されます。
アクチノイドおよびランタノイドは「Rare Earth（希土類）」カテゴリに格納されています。
希ガス（He、Ne、Ar、Kr、Xn、Rn）は「Noble gases（希ガス）」カテゴリに含まれています。
すべての遷移金属は 1 つの別々のカテゴリに分類されています。

Lipinski Rule Oasis（Lipinski 則 Oasis）

Lipinski's Rule of Five（リピンスキーのルール・オブ ・ファイブ）とは、薬物の類似性を評価したり、特定の薬理学的または生物学的活性を有する化合物が、ヒトにおいて経口活性の高い薬物になりうる特性かどうかを判断する経験則のことです。この法則は、1997 年に Christopher A. Lipinski によって、ほとんどの医薬品が比較的小さな親油性の分子であるという観測結果に基づいて案出されました。この法則により、人体における薬物の薬物動態にとって、吸収、分布、代謝、排泄（4 つ

の頭文字をとって「ADME」）を含む重要な分子特性が説明されます。しかしながら、この法則では化合物が薬理学的に活性であるかどうかを予測できません。この法則は、活性と選択性、ならびにリピンスキーの法則が説明する薬物様特性を増加するために、薬理学的に活性なリード構造が一歩一歩最適化されていく薬物開発にとって重要なものです。分子構造を改変すると、より高分子量、より多くの環状構造、より多くの回転可能な結合、より高い親油性を有する薬物になることがよくあります。

Organic functional groups（有機官能基）

この Organic Functional Groups (OFG) システムは、大規模データベースから有機化学物質の様々な機能性・特徴的構造フラグメントの分類化と体系化を導入し、構造的類似化学物質を同定できるよう設計されています。Organic functional groups（有機官能基）とは、その分子の特徴的な化学反応を決める、分子内の原子および/または結合の特定の群です。官能基は、有機分子の特徴的な構造フラグメントの大事な部分です。1 つの官能基が別の官能基に変化することもありますし、官能基が別の分子と反応してより大きな構造を構築することもあります。所与の官能基の相対的反応性は、近くに位置する他の官能基によって改変されることがあります。したがって、官能基は、有機化学における主要な「反応単位」であり、構造と化学反応性との間の基本的な関係を決定します。この OFG プロファイルは現在、同じ官能基を含む OFG と OFG（入れ子状）に細分されています。その相違点とは、OFG は標的化合物中に存在する全ての官能基を表示するのに対し、OFG（入れ子状）はより大きな官能基に入れ子にされている官能基を表示しないことです（すなわち、アルデヒド基はカルボキシル基に入れ子になっており、OFG（入れ子状）プロファイラでは示されていない）。したがって、OFG（入れ子状）システム内で「重なっている」基は、実際には別の官能基を持っています。

Organic functional groups (nested)（有機官能基・入れ子状）

この Organic Functional Groups (OFG) システムは、大規模データベースから有機化学物質の様々な機能性・特徴的構造フラグメントの分類化と体系化を導入し、構造的類似化学物質を同定できるよう設計されています。Organic functional groups 有機官能基）とは、その分子の特徴的な化学反応を決める、分子内の原子および/または結合の特定の群です。官能基は、有機分子の特徴的な構造フラグメントの大事な部分です。1 つの官能基が別の官能基に変化することもありますし、官能基が別の分子と反応してより大きな構造を構築することもあります。所与の官能基の相対的反応性は、近くに位置する他の官能基によって改変されることがあります。したがって、官能基は、有機化学における主要な「反応単位」であり、構造と化学反応性との間の基本的な関係を決定します。この OFG プロファイルは現在、同じ官能基を含む OFG と OFG（入れ子状）に細分されています。その相違点とは、OFG は標的化合物中に存在する全ての官能基を表示するのに対し、OFG（入れ子状）はより大きな官能基に入れ子にされている官能基を表示しないことです（すなわち、アルデヒド基はカルボキシル基に入れ子になっており、OFG（入れ子状）プロファイラでは示されていない）。したがって、OFG（入れ子状）システム内で「重なっている」基は、実際には別の官能基を持っています。

Organic functional groups (US-EPA)（有機官能基・US-EPA）

645 の構造フラグメントと補正係数がこの[Organic Functional Groups（US-EPA）]拡張版で使用されています。プロファイラは米国環境保護庁の EPI Suite プログラムの KOWWIN フラグメント ライブラリ由来のものです。塩基のライブラリに加えて、特定のもっと大きい環状構造系を特徴付けるために、原子フラグメントセットに追加のフラグメントを加えました。. 一般に、化学構造中の水素でない各原子（例えば、炭素、窒素、酸素、硫黄など）は、フラグメントの「骨格」であって、正確なフラグメントは、その原子に何が結合しているかによって決定されます。いくつかの官能基は骨格の「原子」として扱われる。カルボニル基（C=O）、チオカルボニル基（C=S）、ニトロ基（-NO_2）、硝酸塩基（ONO_2）、シアノ基（-C / N）、イソチオシアン酸（-N=C=S）が含まれています。各骨格の「原子」への結合は、一般的な結合または特定の結合のいずれかになります。特定の結合は一般的な結合よりも優先されます。例えば、芳香族炭素原子、芳香族酸素原子および芳香族硫黄

原子は、一般的な結合のみを有します。すなわち、そのフラグメントはその原子に何が結合していても同じということです。対照的に、5 つの芳香族窒素フラグメントが存在します。：
(b)　（a）5 員環の場合、（b）6 員環の場合、（c）窒素が酸化物タイプの場合、すなわちピリジンオキシド、（d）窒素が縮合環位置、すなわちインドリジン、（e）窒素が+5 価、すなわち N-メチルピリジニウムヨウ化物の場合、酸化物タイプが最も特定の結合であるため、他の 4 つより優先されます。別の例として脂肪族炭素原子があります。 -CH3、-CH2-、または -CH< に何が結合していても関係なく、フラグメントは同じです。しかしながら、水素が付いていない脂肪族炭素は、（a）3 つかそれ以上の炭素結合を有する 4 つの単結合が存在する場合、および（b）前記の基準を満たさない、のいずれかの 2 つのフラグメントを有する可能性があります。米国環境保護庁の KOWWIN フラグメント ライブラリの開発に関する詳細や、フラグメントのリストについては EPI Suite のヘルプメニューを参照してください。このプログラムは、米国環境保護庁のウェブサイト（http://www.epa.gov/oppt/exposure/pubs/episuite.htm）からダウンロードすることができます。Meylan, W.M. and P.H.Howard.1995.Atom/fragment contribution method for estimating octanol-water partition coefficients. J. Pharm.Sci.84: 83-92.をご覧ください。
KOWWIN 方法論のより完全な解説についてはジャーナルの記事を参照してください。KOWWIN ライブラリは、初期の原子 - フラグメント構造解析活性の基礎として、米国環境保護庁の類似物質識別方法（AIM）でも使用されています。類似物質識別方法（AIM）は Web ベースのツールで、http://aim.epa.gov からアクセスできます。

Organic functional groups, Norbert Haider (checkmol)（有機官能基、Norbert Haider ・Checkmol）

このプロファイルは、 "Checkmol"というプログラムによって認識される 204 の有機官能基を基盤に作成されました。分子がさまざまなカテゴリに分類されることにより、入力された分子が官能基の存在について分析されます。

Structure similarity（化学構造の類似性）

分子の類似性とは状況に応じて定義される概念であり、すなわちその分子を比較するためのエンドポイントによります。例えば、2 つの分子がある場合、AMES 変異原性（DNA に損傷を与えうる同じ官能基を有する）に関しては類似しているかもしれないが、急性水生生物毒性に関しては類似していない（タンパク質/脂質損傷、または同じ作用機序を起こす共通の機能を有さない）ということがあり得ます。このような類似性・非類似性は反応メカニズム的といわれていますが、巨大分子との相互作用を支配する官能基（最終的には分子のエンドポイントに関わる）と直接には関係のない、抽象的な分子特徴に基づく類似性とは異なります。後者は構造的類似性と呼ばれ、反応メカニズム的類似性と比較してより一般的な性質を有しています。構造的類似性は、QSAR 研究においてよく用いられます。また構造類似性は、分子・原子の特性に基づいて構造的に類似の化学物質を同定するために使う分類方法として使用することができます。

Tautomers unstable（不安定な互変異性体）

この Tautomers unstable プロファイラは、水相・気相中の互変異性体形態に関する利用可能なデータと理論計算の基礎に基づいて開発されています。この不安定な互変異性形態は、独立したカテゴリとして表示されます。このプロファイラはリード・アクロスに有効です。

D.2.1.5. Toxicological（毒物学的）

化学物質の毒性プロファイラを特定するために使用される反復投与プロファイラです。プロファイラには、有害性評価支援システム（HESS）のデータベースから抽出された反復投与毒性試験データに基づく境界が含まれています。このプロファイラは経済産業省（日本）のプロジェクトで LMC の協力のもと、NITE（（独）製品評価技術基盤機構）によって開発されました。

- Repeated dose toxity (HESS)（反復投与量、HESS）

この RD プロファイラの簡単な説明を以下に示します。

Repeated dose (HESS)（反復投与量、HESS）

このプロファイラは、反復投与経口毒性において同様の毒物学的影響を誘発すると予想される、カテゴリ領域からなります。これらのカテゴリ領域は、有害性評価支援システム（HESS）のデータベースにおける反復投与毒性試験データに基づいて開発されました。各カテゴリ（反応メカニズム的または経験的情報）の根拠付けについて説明されています。

D.2.1.6. Custom（カスタム）

これらはユーザが作成したプロファイリングスキームです。いずれの場合でも、新しいプロファイラを作成する際に、それをユーザは上記のいずれかのセクションに配置することを選択できます。

優先順位付けスキーム(PBT) の例が Custom プロファイラに入っています。この手法は、実測値および/または予測値に基づいて、標的化学物質の残留性、生物蓄積性および毒性を特定します。

- Example Prioritization Scheme (PBT)（優先順位付けスキームの例、PBT）

D.2.2. Metabolism（代謝）

Toolbox では、以下の方法で生成される 2 種類の代謝物を提供することができます。:

- Documented metabolism（実測された代謝物）
- Simulated metabolism（シミュレーション予測による代謝物）

すべての代謝物が実測値も予測値も検索可能であり、同じ手法でプロファイリングされています。それは標的化学物質がユーザに、標的化学物質の有毒な代謝物を形成する可能性を通知するからです。研究された種や実験条件に関する追加情報は、データベースによっても提供されています。

D.2.2.1. Documented metabolism（実測記録をもつ代謝）

微生物と哺乳動物における 400 種類の実測済み代謝経路を集めた 3 つのデータベースがツールボックスに備えられています。

Observed microbial catabolism（微生物の異化作用の実測）
551 の化学物質から炭素とエネルギーを得るために微生物が使用する分解経路が、コンピュータで異化作用の情報へアクセスしやすくする特別なファイル形式で保存されています。このコレクションには、C1 化合物、脂肪族炭化水素、脂環式環状構造、フラン、ハロゲン化炭化水素、芳香族炭化水素、ハロゲン化芳香族化合物、アミン、スルホン酸塩、硝酸塩、ニトロ誘導体、ニトリル、複数の官能基を含む化合物の異化作用が含まれています。経路の大部分は好気性環境に関連しています。研究論文、科学記事、UM-BBD（L.B.M.Ellis, D. Roe、L. P. Wackett. Nucleic Acids Res., 34, D517（2006）; http://umbbd.msi.umn.edu/）といった公開ウェブサイトなど、さまざまな情報源を利用してデータベースをコンパイルしました。

Observed mammalian metabolism（哺乳類の代謝の実測）
さまざまな哺乳類について 630 もの研究が行われた、100 の化学物質について実測記録をもつ代謝経路があり、コンピュータで代謝情報へアクセスしやすいデータベース形式で保存されています。このコレクションには、様々な官能基をもつ化学物質すなわち脂肪族アミン、ハロゲン化アルキル・アリール、エーテル、エステル、カルバミン酸塩、カルボン酸エステル、多官能基化合物が含まれています。In vivo と in vitro 研究を用いて化学物質の代謝運命を分析しました。In vivo 研究を備えた代謝マップは、圧倒的に研究のコレクションといえます（49 のマップに 347 の研究が含まれる）。ミクロソームの研究については in vitro 研究に含まれる他の実験系よりもたくさんあります（研究の約 50 %）。in vivo 研究に含まれる投与経路の約 50 % は、経口投与経路を指しています。研究論文、科学記事、公開ウェブサイトを含むさまざまな情報源を利用してデータベースをコンパイルしました。

Observed Rat In vivo metabolism（ラット in vivo（生体内）代謝）
科学文献から抽出されており、げっ歯類（主にラット）における生体異物の化学物質の in vivo 生物変換に関連した、647 の化学物質について実測記録をもつ代謝経路が、コンピュータで代謝情報へアクセスしやすくするデータベース形式で保存されています。このデータベースには、脂肪族と芳香族炭化水素、ハロゲン化炭化水素、アルコールとフェノール、カルボニル化合物、カルボン酸とエステル、ニトロ化合物、アミン、有機硫化物、複素環状構造、そして主に多官能基化学物質などの様々な官能基を持ち構造的に異なる化学物質が含まれています。適応される分野には、工業用化学物質、溶剤、モノマー、医薬品、農薬、いくつかの植物化学物質、アゾ化学物質などがあります。主として研究論文、科学記事、公開ウェブサイトを含むさまざまな情報源を利用して代謝情報を抽出し、データベースをコンパイルしました。

Observed Rat Liver S9 metabolism（ラット肝 S9 代謝の実測）
げっ歯類（主にラット）肝ミクロソームや S9 断片のような in vitro 実験系を用いて観測した 261 種の化学物質のための実測代謝経路が、コンピュータで代謝情報へアクセスしやすくするデータベース形式で保存されています。このデータベースには、脂肪族と芳香族炭化水素、ハロ

ゲン化炭化水素、カルボン酸とエステル、ニトロ化合物、アミン、有機硫化物、複素環状構造、多官能基化学物質などの様々な官能基を持ち構造的に異なる化学物質が含まれています。主として研究論文、科学記事、公開ウェブサイトを含むさまざまな情報源を利用して代謝情報を抽出し、データベースをコンパイルしました。

Observed rat liver metabolism with quantitative data（定量データを含むラットの肝臓代謝の実測）

「定量的代謝情報および関連機能の実測値」は、定量的代謝データを考慮に入れることで、Toolbox の代謝情報の使用を拡大し、容易にすることを目的としています。この課題に関して、定量的代謝データを含む新しいデータベースが作成され、標準の Toolbox の装備に含められました。ツールボックスのユーザインターフェースでは、こうした情報の詳細を表示できます。また、定量的な代謝データを、類似物質のグループ分けとフィルタリングに使用することもできます。

プロファイリングの後、右クリックオプション（1）（図 1）で実測値代謝マップを見ることができます。ラベル「QTY」（2）は、標的物質/代謝物にいくつか定量的データがあることを示しています。文献から抽出された研究の情報は、右側のパネル（3）に位置しています。ラベルは数字で構成され、この場合は「1.14.14.1」（4）で酵素情報を示しており、METAPATH（訳注：計算化学用のソフト名）のプラットフォームで見ることができます。

親物質および/または代謝物の量を知りたい場合、利用可能な試験の 1 つ上のドロップダウンボックスをクリックしてください。自動的にそれぞれの代謝物の量がウィンドウ（5）の下部に表形式で表示され、時間に対して対象物質の量が減少するのがわかります（例えば、対象物質の消滅）。

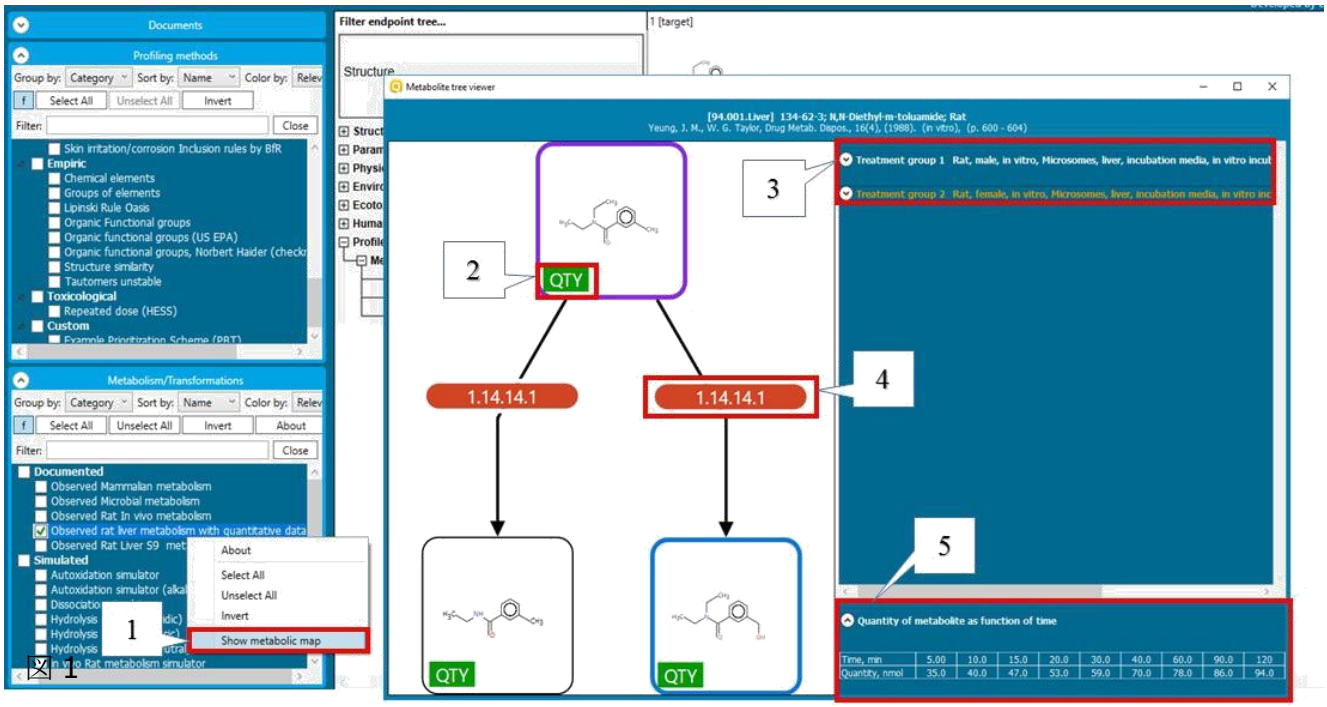


図 1

D.2.2.2. Simulated metabolism（代謝のシミュレーション予測）

微生物、肝臓と皮膚の代謝、ならびに非生物的ヒドロキシル化を適応する分子変換のシミュレータもシステムに備えられています。以下に各シミュレータの簡単な説明を示します。

Autoxidation simulator（自動酸化シミュレータ）

自動酸化（Autoxidation : AU）とは有機分子の自発的な空気誘発酸化です。これは化学物質と酸素分子とのフリーラジカル連鎖反応であり、その結果、酸化物が形成されます。この酸化物のうち、有機ヒドロペルオキシドは接触アレルギーのような有害作用を誘発することに関して最も重大であると考えられています。それゆえ AU 経路実測値をシミュレーションするために、AU モデルが開発されました。この目的のために、133 の化学物質（テルペン、単純脂肪族とポリエチレングリコールエーテル、アルデヒド、アミノフェノール）のトレーニングセットを、AU 経路に関連して公表されたデータと共に用いました。次の条件下でデータ収集をすることで、データの一貫性を維持しました : 空気または酸素曝露、室温、大気圧、バルクまたはほぼ中性（pH 7～7,5）またはわずかにアルカリ性（pH 8～9）のさまざまな溶媒の存在下での AU、中間体、AU の持続時間は数時間から数ヶ月までです。

Autoxidation simulator (alkaline medium)（自動酸化シミュレータ、アルカリ媒体）

133 の化学物質（テルペン、単純脂肪族とポリエチレングリコールエーテル、アルデヒド、アミノフェノール）のトレーニングセットを、AU 経路に関連して公表されたデータと共に用いました。次の条件下でデータ収集をすることで、データの一貫性を維持しました : 空気または酸素曝露、室温、大気圧、バルクまたはさまざまなアルカリ溶媒(pH 10.2～11.5)の存在下での AU、中間体、AU の持続時間は数時間から数ヶ月までである。AU シミュレータは、AU 経路の実測値から抽出された一連の分子変換を備えています。

Dissociation simulator（解離シミュレータ）

Hydrolysis (Acidic) simulator（加水分解・酸性　シミュレータ）

Hydrolysis (Acidic) simulator は、以下の実験条件で、別個の有機化学物質の加水分解生成物を予測します。酸性 pH、大気の温度または室温、および大気圧下です。このシミュレータは、論文や公開ウェブサイトを含むさまざまな情報源から収集されたデータに基づいて開発されました。このモデルには、次の化学物質の分類が含まれています。エポキシド、アジリジン、エステル、カルバメート、ハロゲン化メタン、選択されたハロゲン化アルキル、無水物、ジチオカルバメート、イソシアネート、イソチオシアネート、塩化スルホニル、ラクトン、ニトリル、アミド、N-ハロゲン化アミン、ジケテン、有機過酸化物などです。

Hydrolysis (Basic) simulator（加水分解・塩基性　シミュレータ）

Hydrolysis (Basic) simulator は、以下の実験条件で、別個の有機化学物質の加水分解生成物を予測します。塩基性 pH、大気の温度または室温、および大気圧下です。このシミュレータは、論文や公開ウェブサイトを含むさまざまな情報源から収集されたデータに基づいて開発されました。このモデルには、次の化学物質の分類が含まれています。スルホニルハロゲン、有機リン化合物、エポキシド、アジリジン、エステル、カルバメート、ハロゲン化メタン、選択されたハロゲン化アルキル、無水物、ジチオカルバメート、イソシアネート、イソチオシアネート、塩化スルホニル、ラクトン、ニトリル、アミド、N-ハロゲン化アミン、ジケテン、有機過酸化物などです。

Hydrolysis (Neutral) simulator（加水分解・中性　シミュレータ）

加水分解シミュレータは非生物モデルであり、中性またはほぼ中性の pH 下において、試験 28 日目のときの化学物質の加水分解生成物とその量を予測します。このモデルは以下の実験条件で、個々の有機化学物質の加水分解生成物を予測します。中性またはほぼ中性の pH、大気の温度または室温、および大気圧下です。このシミュレータは、記事や公開 Web サイトを含むさまざまなソースから収集されたデータに基づいて開発されました。このモデルには、次の化学物質の分類が含まれています。個々の有機化合物、エポキシド、アジリジン、エステル、カルバメート、ハロゲン化メタン、選択されたハロゲン化アルキル、無水物、ジチオカルバメート、イソシアネート、イソチオシアネー

ト、塩化スルホニル、ラクトン、ニトリル、アミド、N-ハロゲン化アミン、ジケテン、有機過酸化物などです。

In vivo Rat metabolism simulator（In vivo ラット代謝シミュレータ）

現在の in vivo ラット肝代謝シミュレータ（形質転換表）とは、構造的に汎用化し、階層的に配置した 609 の非生物的および酵素反応の電子的に設計されたセットのことであり、げっ歯類（ほとんどがラット）のような in vivo 試験系の代謝の特徴を表しています。このシミュレータの主な適用性は、生殖、ならびに細菌の突然変異誘発や染色体異常などの in vitro 遺伝毒性効果を誘発し得る、代謝活性化反応と生体異物の経路の予測に関連しています。シミュレータ内の形質転換はそれぞれ、親物質と生成物の構造フラグメントと、阻害するマスキングを備えています。発生確率はそれぞれの主な形質転換によるもので、主な形質転換によって形質転換リスト内でのその階層が決まります。肝臓代謝に関する公表データを用いて、代謝反応実測値と経路を有する、幅広い構造的多様性を持った外因性化学物質のトレーニングセットが構築されました。その代謝に関するデータは、主に選ばれた科学ジャーナル、研究論文、ウェブサイトの領域の研究出版物から収集され、さまざまな構造を有する化学物質でよく観測される生体内肝代謝反応に関連しています。分子転換のセットは、試験の持続時間と比較して非常に高い速度で生じると考えられる、26 の非生物的反応と 583 の酵素制御反応を一部分として含み、これらの反応が最優先にされます。またこの反応のサブセットは高反応性官能基と中間体の変換、すなわち互変異性化、フェノールへのアレーンエポキシド転位なども含んでいます。全体として、シミュレータには、脂肪族の C の酸化、芳香族の C のヒドロキシル化、酸化した N と O の脱アルキル化、エポキシ化、エステルとアミド加水分解、カルボニル基還元、ニトロとアゾ基還元、酸化的脱アミノ化、β-酸化、環開裂、加水分解開裂、芳香族化、脱カルボキシル化、脱ハロゲン化などの 479 の酵素的な第 I 相転移も含みます。さらに、 in vitro 系と異なり、グルクロニド化、硫酸化、グルタチオン抱合、N-アセチル化などの 104 の酵素的な第 II 相転移が起こると考えられており、 in vivo で高い優先度で起こると考えられます。

Microbial metabolism simulator（微生物代謝シミュレータ）

微生物代謝のオリジナルの CATABOL シミュレータがシステム内に備えられています。(J.Jaworska, S. Dimitrov, N. Nikolova, O. Mekenyan. SAR QSAR Environ. Res., 13, 307 (2002); S. Dimitrov, R Breton, D. Mackdonald, J. Walker, O. Mekenyan. SAR QSAR Environ. Res., 13, 445 (2002); S. Dimitrov, V. Kamenska, J.D.Walker, W. Windle, R. Purdy, M. Lewis, O. Mekenyan. SAR QSAR Environ. Res., 15, 69 (2004).) 多経路異化作用は、公開されている代謝経路データベースから抽出された、階層順序に並べられた主な分子変換を介する非生物的反応と酵素媒介反応を用いて、シミュレーションされます。形質変換の階層は、化学物質の異化作用マップの増殖を制御するために使用されています。シミュレーションは、最上位の階層にある形質変換に関連する親分子とソースフラグメントとの一致を検索することから始まります。一致が見つからない場合、次の階層の形質変換などで検索が行われます。一致が見つかると、形質変換生成物が生成されます。新たに形成された生成物についてこの手順が繰り返されます。Observed microbial catabolism データベースに保存された 200 の化学物質の公開された異化作用に基づいて評価される予測性（代謝物が予測できると仮定したとき、代謝物が実測される確率）は 83%です。

Rat liver S9 metabolism simulator（ラット肝臓 S9 代謝シミュレータ）

現在の in vitro ラット肝代謝シミュレータ（形質転換表）とは、構造的に汎用化し、階層的に配置した 509 の生体内変換反応の電子的に設計されたセットのことであり、げっ歯類（主にラット）肝ミクロソームや S9 画分などの in vivo 試験系の代謝の特徴を表しています。このシミュレータの主な適用性は、生殖、ならびに細菌の突然変異誘発や染色体異常などの in vitro 遺伝毒性効果を誘発し得る、代謝活性化反応と生体異物の経路の予測に関連しています。

シミュレータ内の形質転換はそれぞれ、親物質と生成物の構造フラグメントと、阻害するマスキングを備えています。発生確率はそれぞれの主な形質転換によるもので、主な形質転換によって形質

転換リスト内でのその階層が決まります。げっ歯類肝臓代謝や S9 画分に関する公表データを用いて、代謝反応実測値と経路を有する、幅広い構造的多様性を持った外因性化学物質のトレーニングセットが構築されました。トレーニングセット中の有機化合物は、さまざまな種類の工業用化学物質、すなわち単環と縮合環のアレーン、フェノール、ハロゲン化アルカンとハロゲン化アレン、芳香族・脂肪族アミン、ニトロアレーン、アルカンとシクロアルカン、アルケン、エーテル、カルボン酸およびその誘導体、ハロゲン化炭化水素、アルコール、エポキシド、N-ニトロソアミン、アゾ化学物質、などに属しています。その代謝に関するデータは、主に選ばれた科学ジャーナル、研究論文、ウェブサイトの領域の研究出版物から収集され、さまざまな構造を有する化学物質でよく観測される in vitro 肝代謝反応に関連しています。

分子転換のセットの一部分に、試験の持続時間と比較して非常に高い速度で生じると考えられる、25～30 の非生物的反応と数個の酵素制御反応が含まれ、これらの反応が最優先にされます。またこの反応のサブセットには高反応性官能基と中間体の変換、すなわち互変異性化、フェノールへのアレーンエポキシド転位なども入っています。全体としてシミュレータは、脂肪族の C の酸化、芳香族の C のヒドロキシル化、酸化した N と O の脱アルキル化、エポキシ化、エステルとアミド加水分解、カルボニル基還元、脱アミノ化、β-酸化、環開裂、加水分解開裂、芳香族化、脱カルボキシル化、脱ハロゲン化などの 450～470 の酵素的な第 I 相転移からなります。さらに、グルクロニド化、硫酸化、グルタチオン抱合、N-アセチル化などの 15～20 の酵素的な第 II 相転移が第 I 相よりもかなり低い優先度で含まれています。

Skin metabolism simulator（皮膚代謝シミュレータ）

Skin metabolism simulator は、皮膚部分の化学物質の代謝に従っています。皮膚代謝データの報告が無く、肝臓で測定された反応に類似する反応を介して吸収された生体異物を皮膚酵素が代謝することができるという広く普及した仮定のもとに、このシミュレータは簡略化された哺乳動物肝代謝シミュレータとして開発されました。皮膚代謝シミュレータには、階層的に順序付けられた主な形質変換のリストが含まれており、その変換は反応速度で決まる変換と反応速度では決まらない変換の 2 つの主要なタイプに分けられます。反応速度で決まる変換とは、C-ヒドロキシル化、エステル加水分解、酸化、グルタチオン抱合、グルクロニド化、スルホン化など第 I 相および第 II 相です。反応速度では決まらない変換には、反応性の高い中間体の分子変換が含まれます。シミュレータは、親分子を、発生確率が最も高い変換に関連する反応フラグメントと突き合わせることによって開始します。これにより一次代謝物のセットが作られます。次に、この導出された各代謝物は、第 2 段階の代謝物を生成するために、階層的に順序付けられた形質変換の同じリストに提出されます。この手順は代謝増殖の制約が満たされるまで繰り返される（例えば、代謝物を得る確率が低いことや、もっぱら第 II 相反応になるなど）。

シミュレータ（1）を選択して所定の化学物質に適用できます。生成された代謝物は、**データ・マトリックス**（2）上に読み込まれます。代謝物の入ったセル上でダブルクリックすると、代謝物が表示されたウィンドウが表示されます（3）。（図 1）

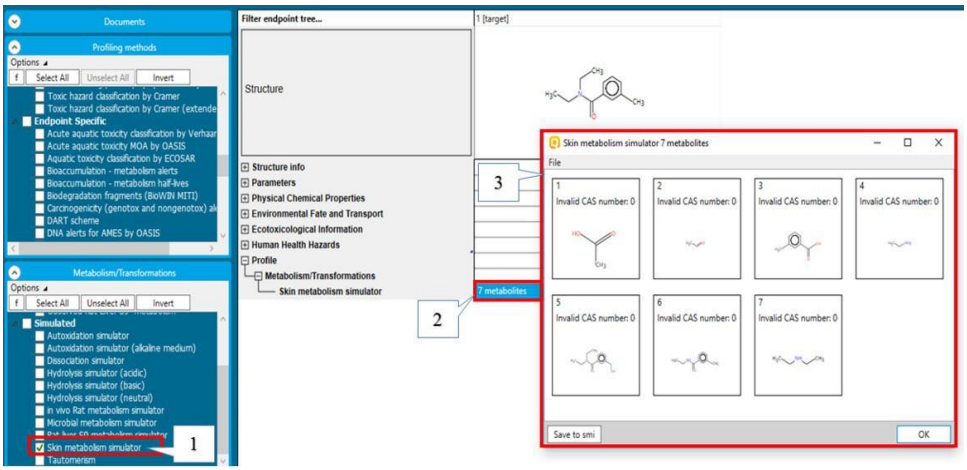


図 1

D.2.3. 補助機能

2.3.1. プロファイラ手法の関連性

プロファイラの適切性は、選択したエンドポイントに関してカテゴリを構築するプロファイラがどれほどふさわしいかという状態によります。エンドポイントが選択されると（1）、Options（2）内の関連するプロファイラが強調表示（3）され、Endpoint selected in the data matrix を選べばグループ化（4）されます。Legend（凡例）ボタン（5）をクリックすると、それぞれの色の意味を見ることができます。（図 1）

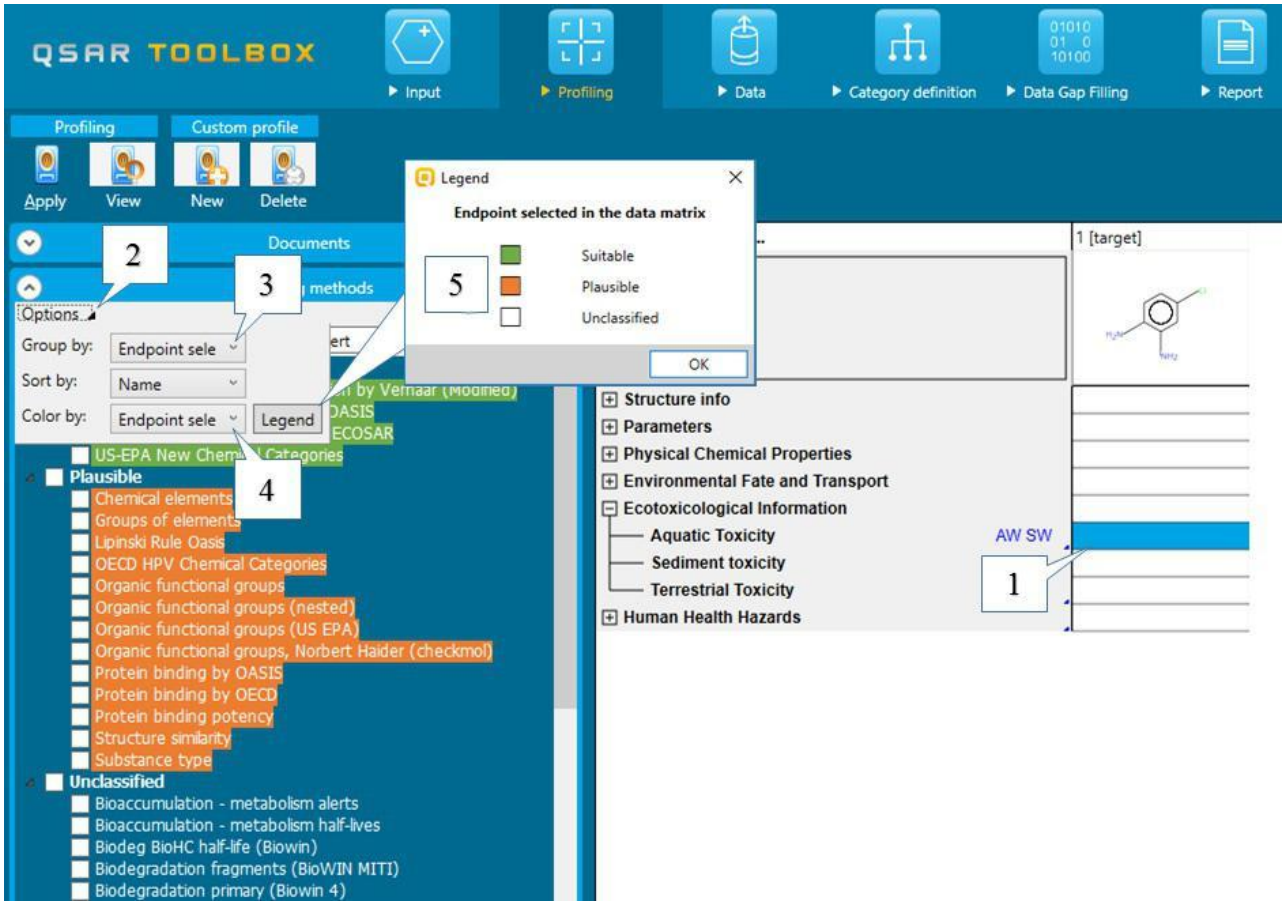


図 1

この評価では、プロファイラの開発に使用された実験データと専門知識との関係性が考慮されます。関連性は 3 種類のスコアをつけることができます：

- **関連性有り**（緑色で強調表示） - プロファイラは、標的物質のエンドポイントのデータ/知識を使用して開発されている
- **関連性の可能性有り**（橙色で強調表示） - プロファイラを構築するために使用したデータ/知識は、何らかの形で標的物質のエンドポイントに関連することが知られる
- **関連性は未知**（強調表示されていない） - 使用されたデータ/知識と標的物質のエンドポイントとの間の関係に関する証拠はない

ユーザがエンドポイントツリーの他の位置（1）に移動すると、それに関連するプロファイラが変更されます（2）（図 2）。このプロファイラで使われる、同じ関連性のスコア（訳注：緑・橙色の強調表示のこと）は、代謝マップの予測値と実測値（3）に対しても適用されています。

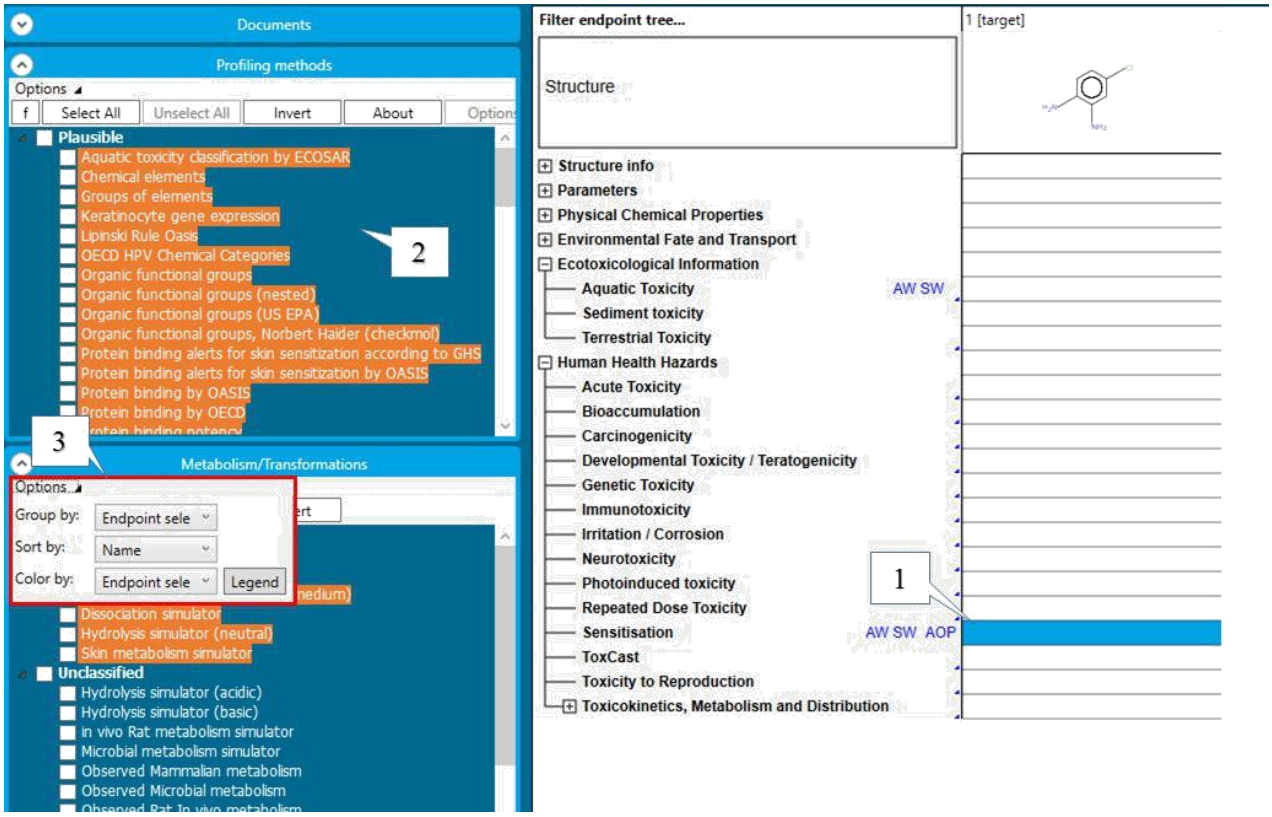


図 2

標的物質のエンドポイント（2）に従って色分けし、グループ化すると、ユーザからは既定のエンドポイント（1）に関連するプロファイラだけが見えます。このオプションを選択すると、エンドポイントツリー（3）の位置に対応するのではなく、選択されたエンドポイント（1）にのみ対応します。(図 3)

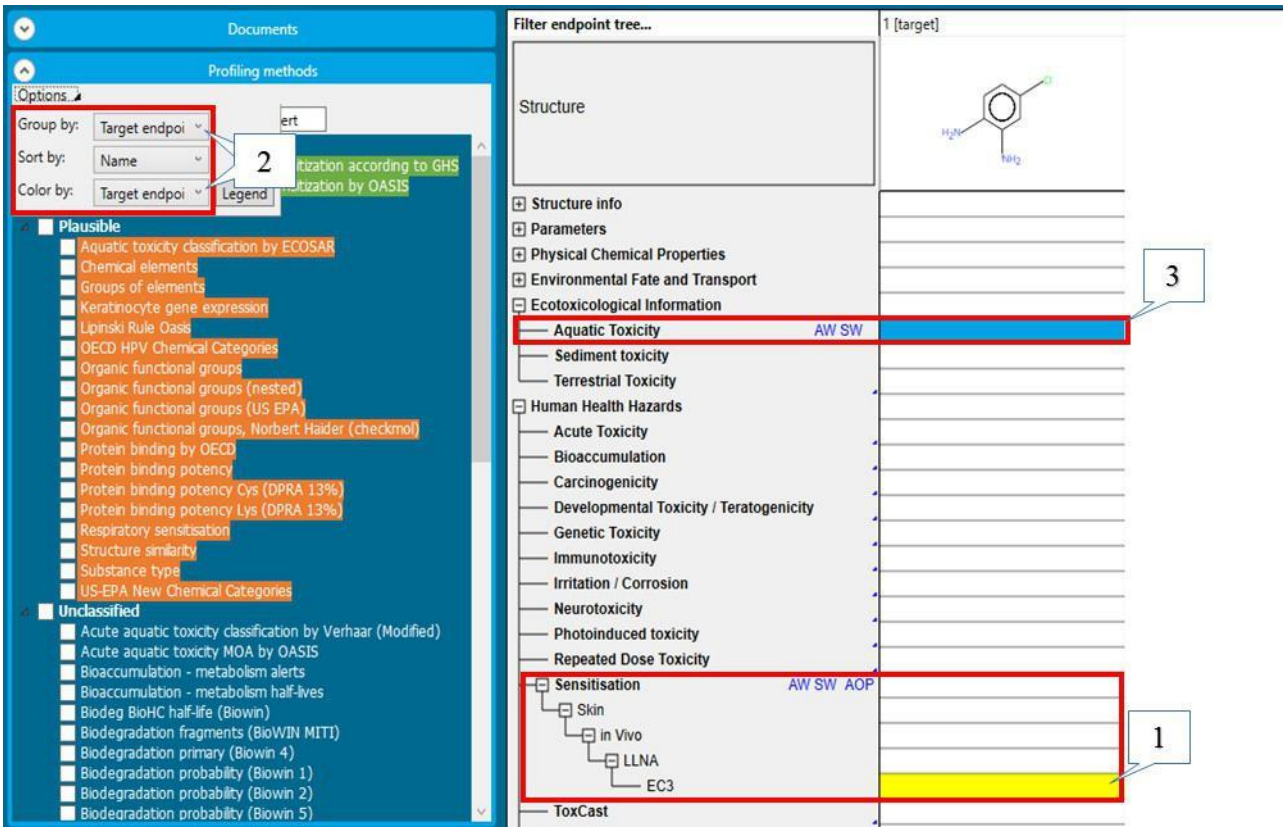


図 3

2.3.2. Filtering（フィルタリング）

フィルタリングにより、ユーザは興味のあるプロファイラおよび/またはシミュレータをより迅速に見つけることができます。F ボタン（1）をクリックし、キーワード（2）を入力してください。これで、要件を満たすプロファイラがフィルタリングで残ります（3）。(図 4)

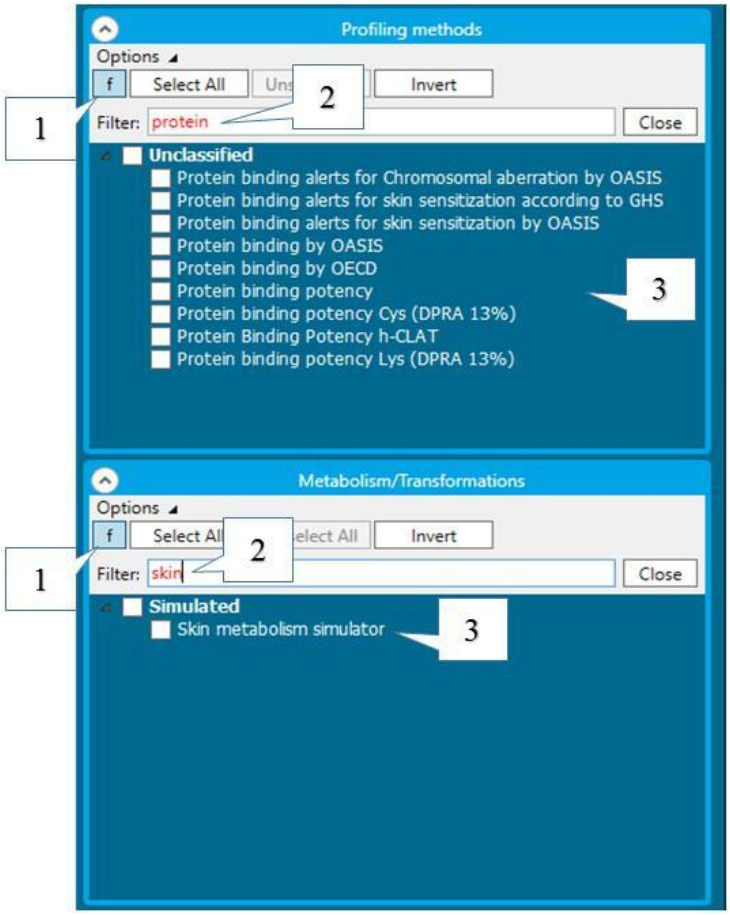


図 4

D.2.4. プロファイリング方法の説明

構造に基づくプロファイリング方法には 3 つのタイプがあります:

- Standard type profiling scheme（標準型プロファイリングスキーム）
- Hierarchical type profiling scheme（階層型プロファイリングスキーム）
- Dendroidal type profiling scheme（ツリー構造型プロファイリングスキーム）

Toolbox 4.0 では、化学構造類似性がプロファイリングセクションのプロファイラとして追加されました。

Standard type profiling scheme（標準型プロファイリングスキーム）

プロファイリングスキームからカテゴリの説明を見るには、
スキーム（1）を選択して、View（2）ボタンをクリックしてください。これにより、Profiling Scheme Browser（プロファイリング・スキーム・ブラウザ）（3）が実行されます。(図 1)

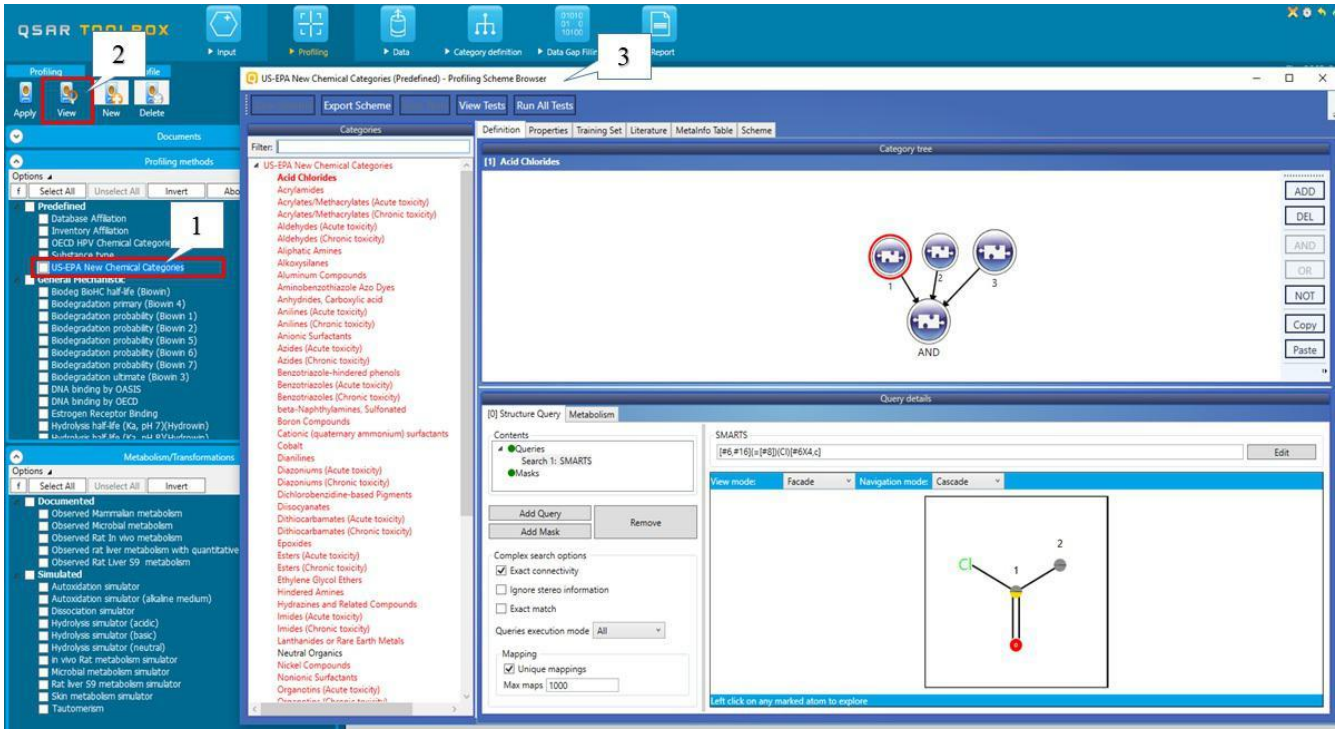


図 1

クエリツリーと選択された領域の内容を見ることができます。新しいカテゴリを定義したり、既存のカテゴリの名前変更または削除をしたり、ルールを追加または削除したり、新しい境界を設定したりすることができます。Toolbox のインストールによって提供されるプロファイラの変更は保存できません。

Hierarchical type profiling scheme（階層型プロファイリングスキーム）

プロファイリングスキーム（図 2）のカテゴリの階層構造は Toolbox が 3.0 のときに開発されました。 このようなタイプの構造を持つプロファイリングスキームを以下にいくつか示します。

- DNA binding by OASIS（DNA 結合；OASIS に基づく）
- DNA binding by OECD（DNA 結合；OECD に基づく）
- DNA alerts for AMES by OASIS（OASIS の AMES 試験の DNA アラート）
- DNA alerts for CA and MNT by OASIS（OASIS の CA と MNT の DNA アラート）

- Protein binding alerts for Chromosomal aberration by OASIS（OASIS の染色体異常のタンパク質結合アラート）
- Protein binding alerts for skin sensitization by OASIS（OASIS の皮膚感作のタンパク質結合アラート）
- Protein binding alerts for skin sensitization according to GHS（GHS による皮膚感作のタンパク質結合アラート）
- Protein binding by OASIS v1.4（タンパク質結合；OASIS に基づく）
- Protein binding by OECD（タンパク質結合；OASIS に基づく） Protein binding potency（タンパク質結合能）
- Protein binding potency Cys (DPRA 13%)（タンパク質結合能 システイン DPRA 13%）
- Protein binding potency Lys (DPRA 13%)（タンパク質結合能 リジン DPRA 13%）
- Keratinocyte gene expression（ケラチン生成細胞遺伝子発現）

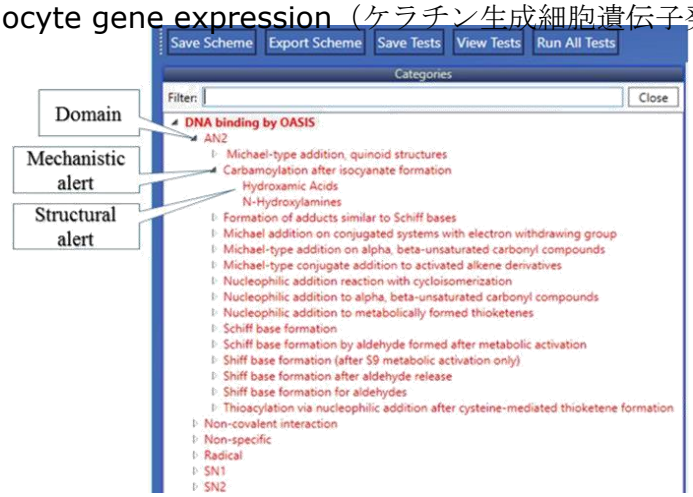


図 2

Dendroidal type profiling scheme（ツリー構造型プロファイリングスキーム）

図 3 でツリー構造型スキームを図示しました。

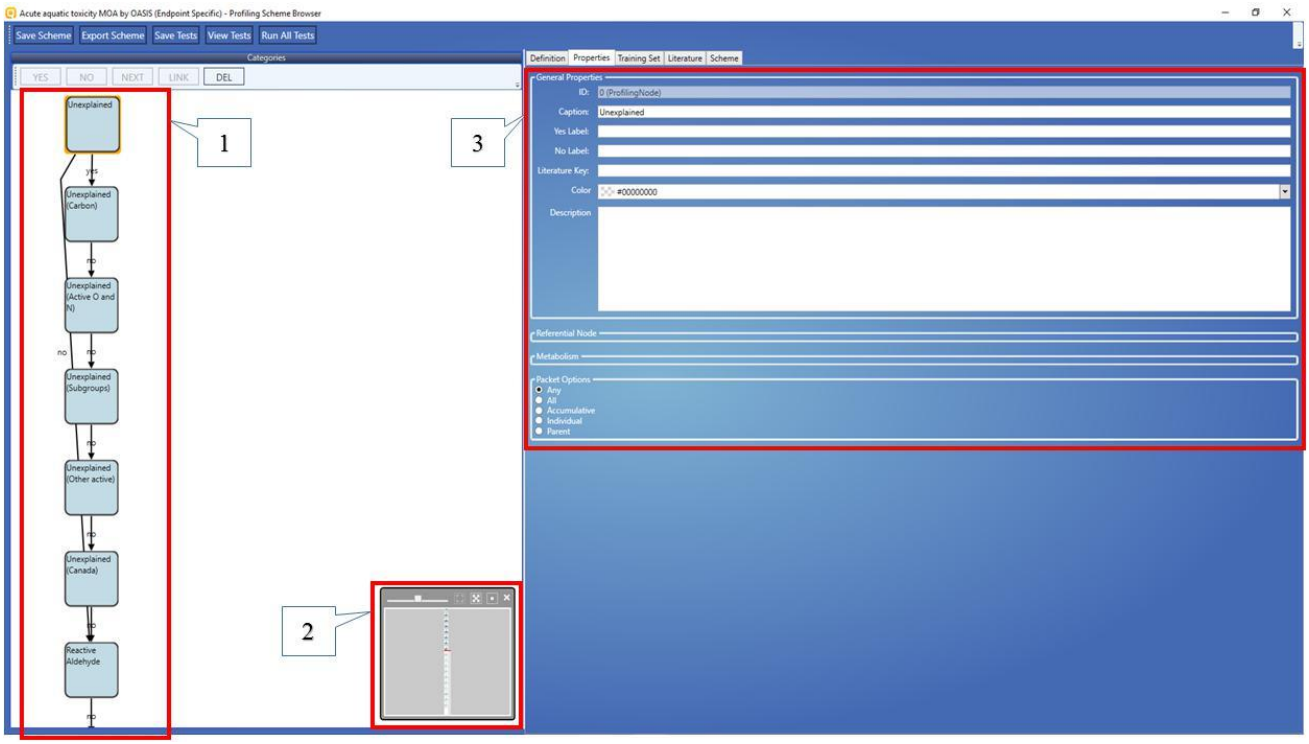


図 3

1 - **logic panel**（論理パネル）:ここに、論理ノードとして構成されたプロファイリングスキーム論理が表示される
2 - **navigation panel**（ナビゲーションパネル）
3 - **a panel with properties of nodes**（ノードのプロパティを示すパネル）

Structure similarity（化学構造の類似性）

構造類似性プロファイラは、経験的プロファイラ（1）に含まれています（図 4）。Toolbox 4.0 の構造類似性オプションは Toolbox 4.1 で拡張されており、PubChem の下部構造の特徴（2）を備えています。

PubChem は、各化学構造について二次サブ構造のフィンガープリント（特徴を特定するもの）を生成する。サブ構造とは化学分子構造の断片です。フィンガープリントは、バイナリ（1/0）ビットの順序付きリストです。各ビットは、特定の原子における Boolean 判定や、さらなる類似性境界と類似性の検索のために使用される試験的特色を表しています。PubChem の機能は以下の 7 つのグループに区別され使用されています。：

- 階層要素数
- 環状構造
- 単原子対
- 最近傍の単原子
- 原子近傍の詳細
- 単純な SMARTS のパターン
- 複雑な SMARTS のパターン

一度 PubChem 機能が選択されると両方の例示構造の間の構造的類似性のパーセンテージが変更されました（3）。2 つの構造間の類似性については、Details をクリックすると詳細を表示します（4）。新しいウィンドウが表示され、どちらのフラグメントが共通（緑色）であるか、両方の化学物質（赤色）に固有のフラグメントを確認できます（5）。

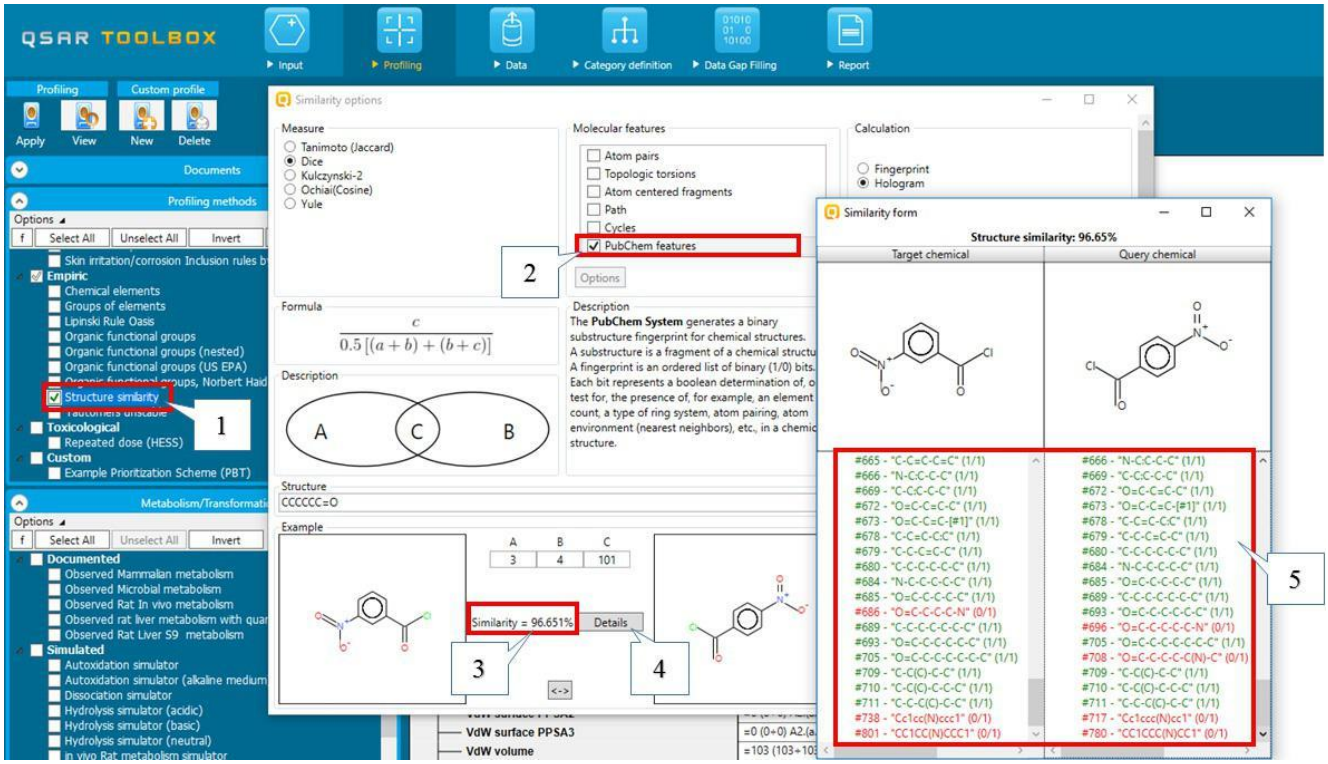


図 4

D.2.5. プロファイリング結果の説明

データ・マトリックスに表示されるプロファイリング結果のさまざまな構成は、3 つのタイプのプロファイリング方法に従って利用できます。

- 標準的な説明
- 階層的な説明

標準的な説明

標準型プロファイリングスキームで得られたプロファイリング結果を図 1 に示します。

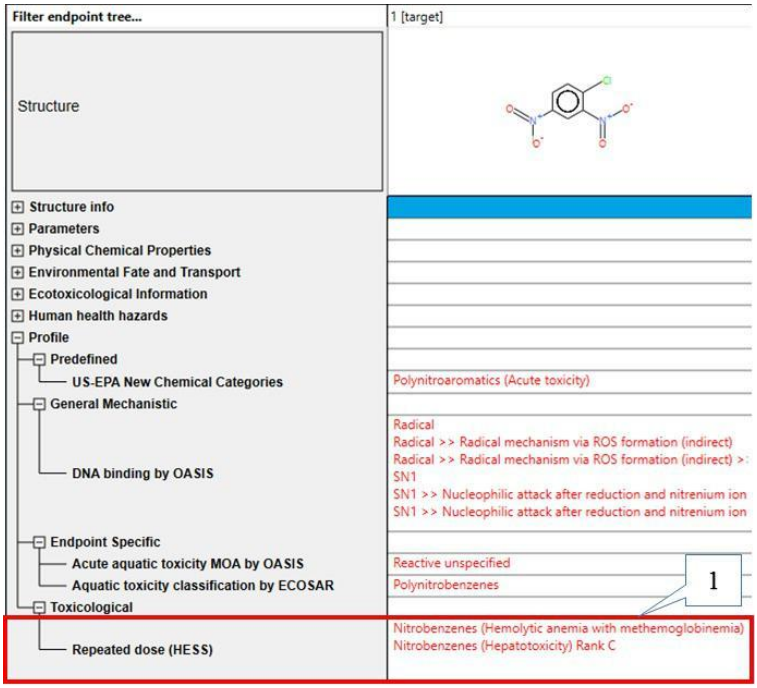


図 1

プロファイリング結果（1）でセルを右クリックし、Explain（2）を選択すると、プロファイリング結果の詳細な説明が表示されます。(図 2)

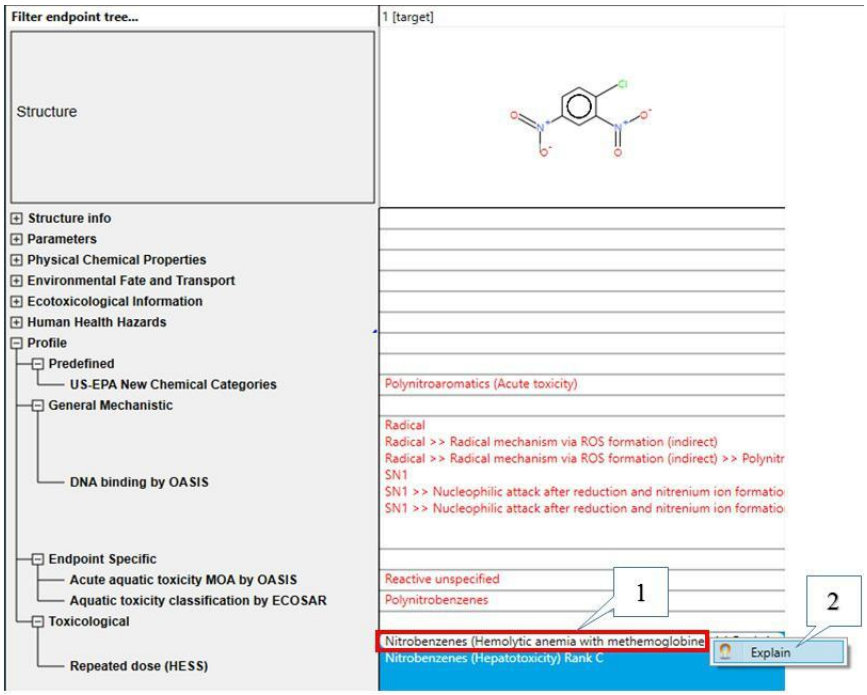


図 2

この動作をすることで、プロファイリング方法の領域と反応メカニズム的な根拠を表示するウィンドウが現れます（1）。（図 3）

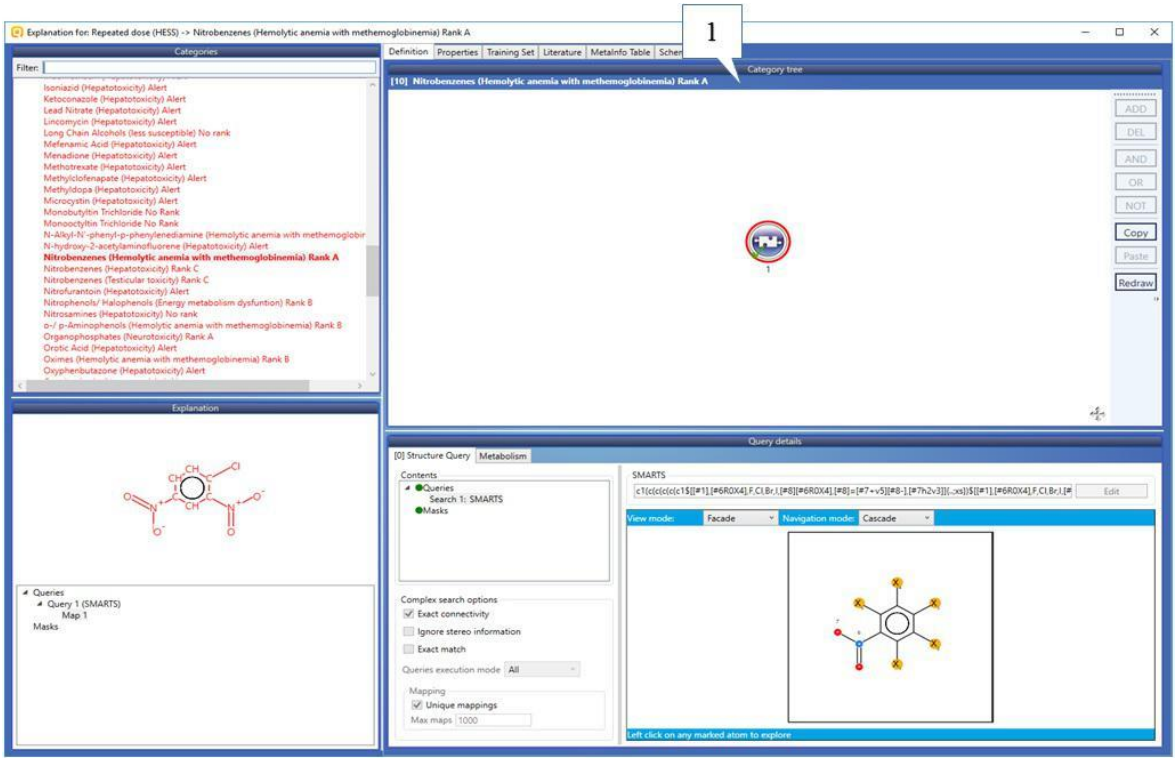


図 3

プロファイリング結果のウィンドウのフィールド（画面内の領域）を以下に説明します（図 4）。

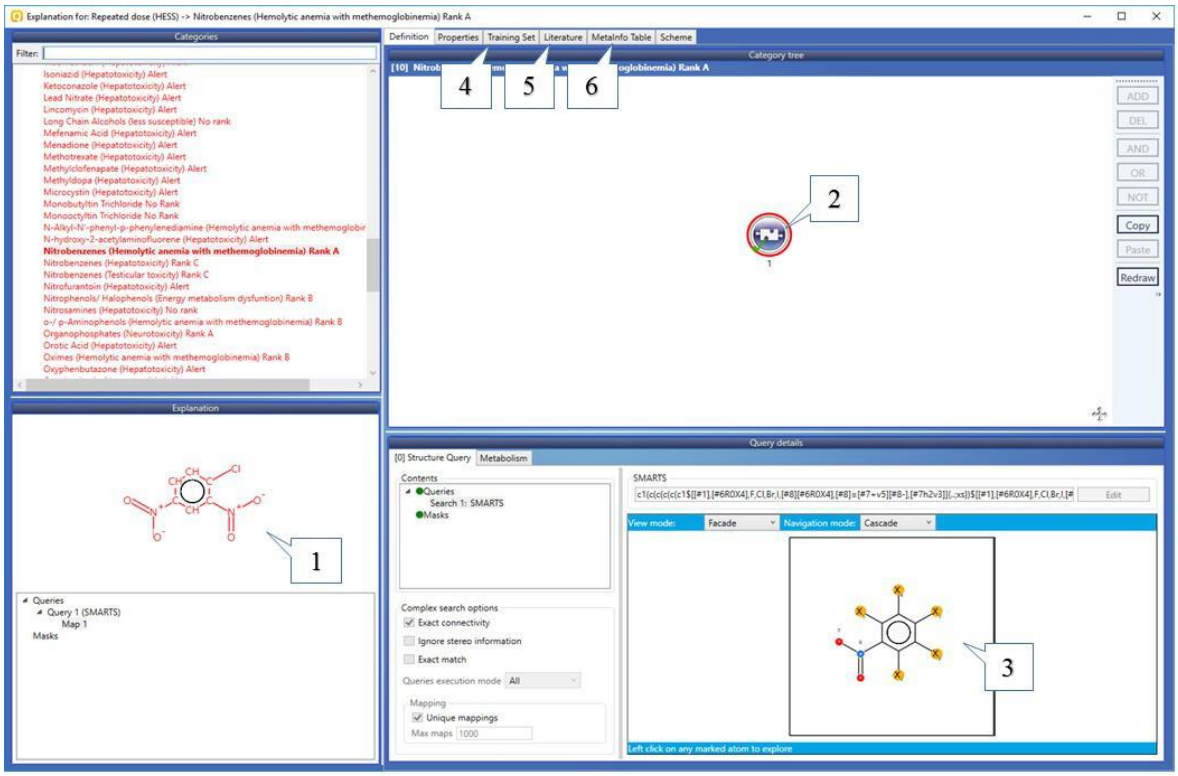


図 4

1 - 標的化学物質を含むパネル。標的物質のいずれかの部分が、表示されているカテゴリの基準を満たす場合、それは赤で強調表示されます。
2 - 表示されているカテゴリの構造的領域。標的物質が領域の基準を満たしている場合は、緑の印でマークされます。
3 - 表示されたカテゴリの構造領域を定義するために使用される構造フラグメント。
4 - 表示されているカテゴリに関連するトレーニングセットの化学物質。
5 - 表示されている各カテゴリに関連する**反応メカニズム**的根拠。
6 -トレーニングセット化学物質の試験のメタデータ情報。

トレーニングセットの化学物質は Training set パネルに記載されています。通常、構造的境界は、カテゴリに関連するこれらのトレーニングセットの化学物質に基づいて抽出されます（図5）。

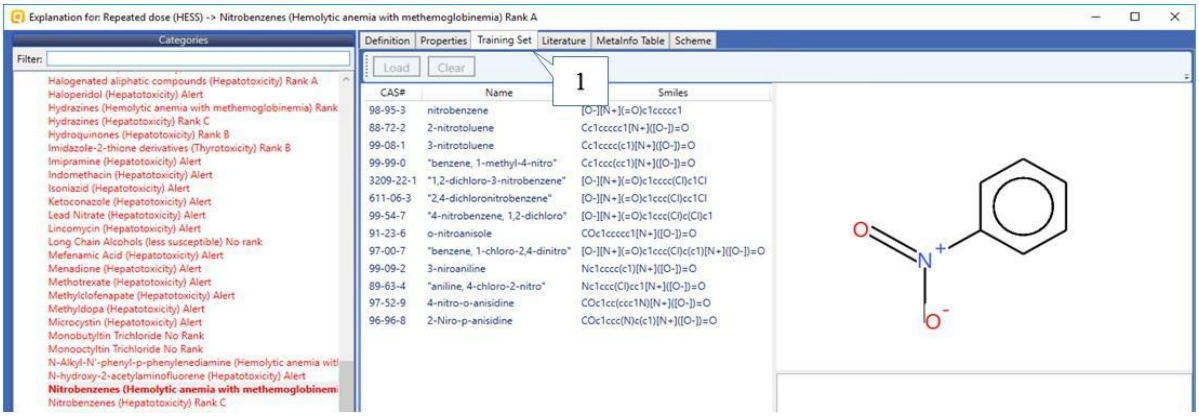


図 5

表示されているカテゴリの反応メカニズム的根拠は、Literature パネルにあります。（図 6）

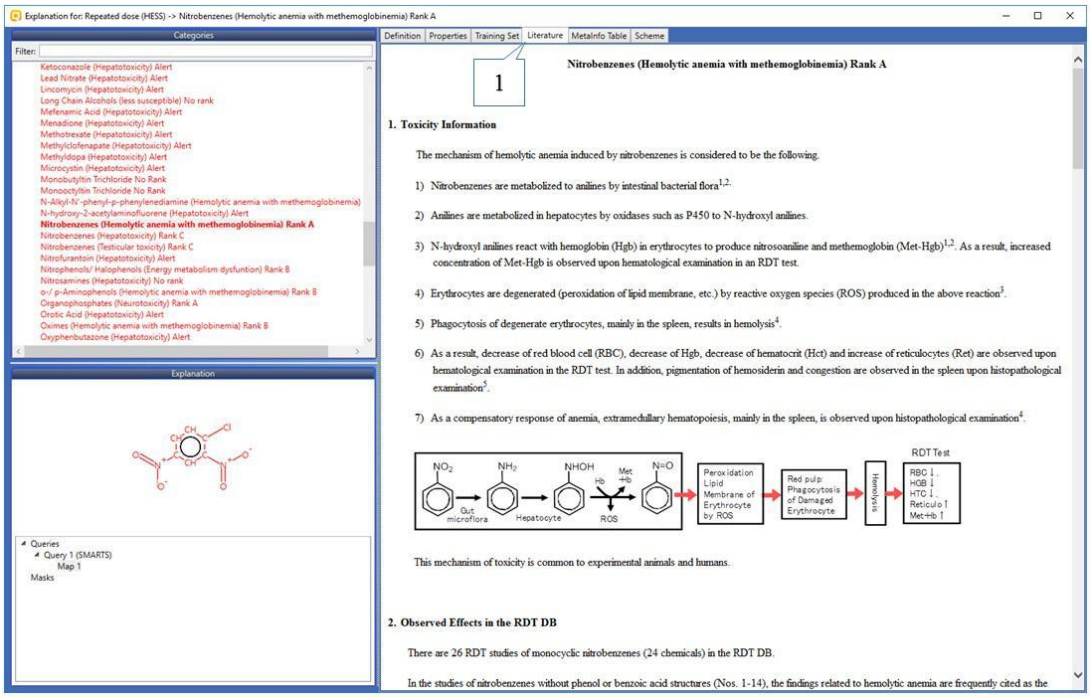


図 6

検査項目、器官、組織および効果に関する追加データは、MetaInto Table（1）（図 7）に示されています。表に記載されている情報は、トレーニングセットの化学物質と対応するエンドポイント（すなわち、この例ではメトヘモグロビン血症を伴う溶血性貧血）に関連しています。

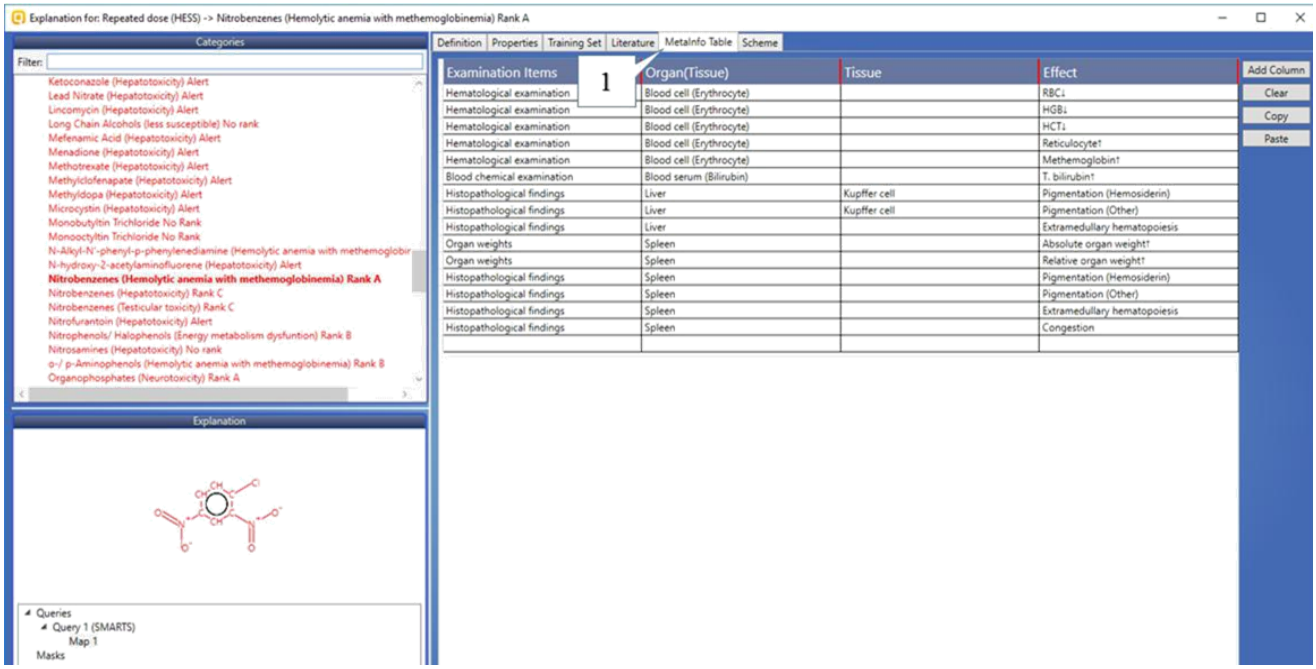


図 7

階層的な説明

この階層型構造に基づいて、これらのタイプのスキームから得られたプロファイリング結果を図 8 に示します。

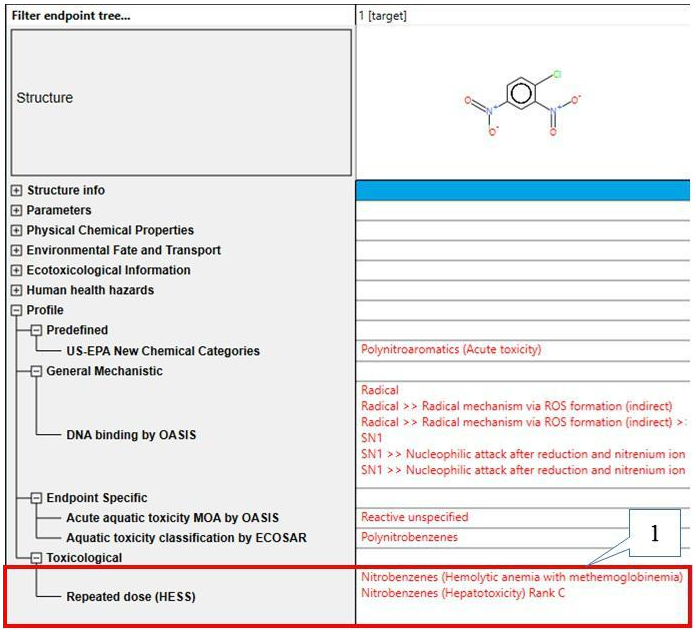


図 8

構造アラートの結果（1）を右クリックし、Explain（2）を選択して詳細を表示します。（図 9）

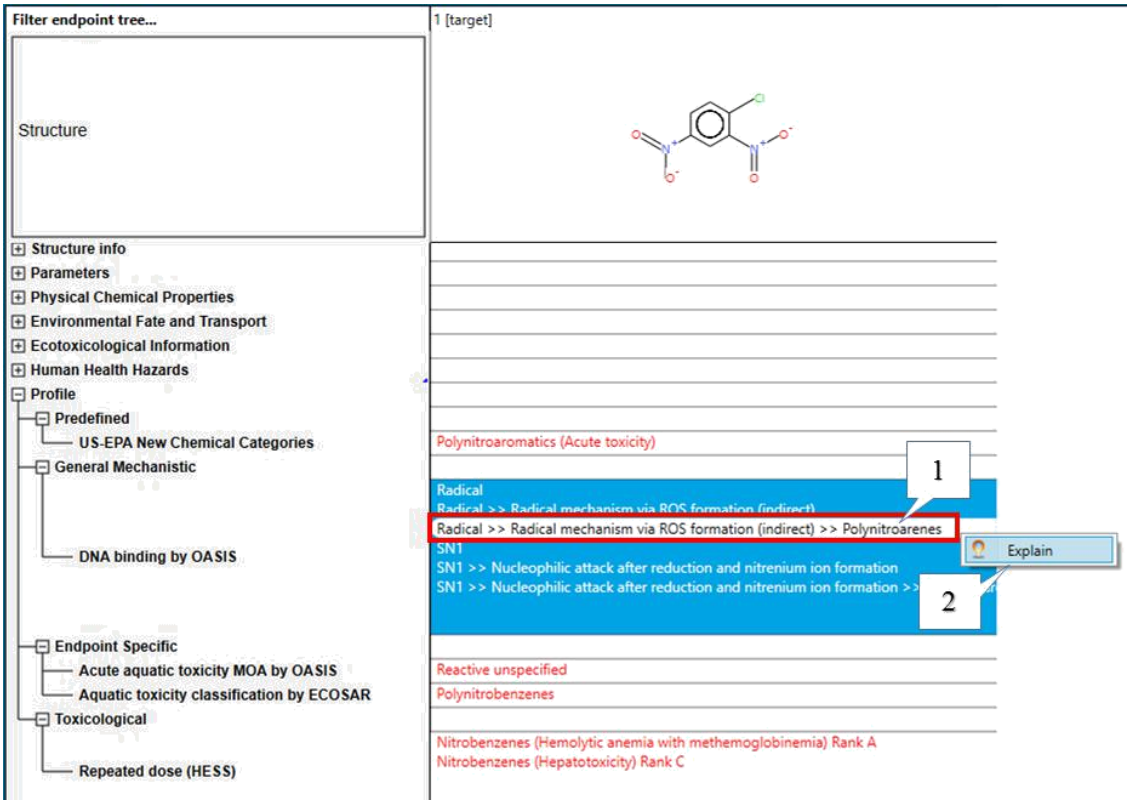


図 9

D.2.6. カスタマイズしたプロファイリングを作る

あらかじめインストールされている分類スキームに追加して、ユーザの理論に基づく新しいプロファイリングスキームを作成できます。
これを行うには、ボタン New（1）をクリックしてください。ダイアログが表示されるので、そこに新しいカテゴリ化スキーム（2）の名前を入力してください（図１）。

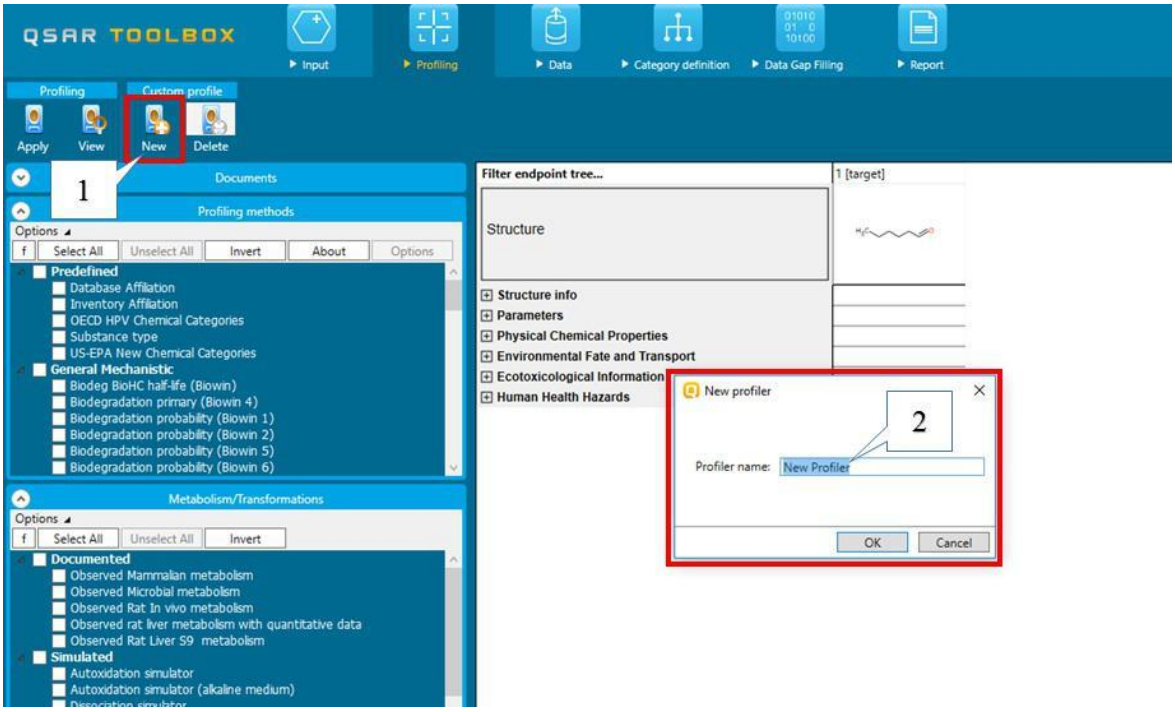


図 1

名前を入力すると、スキームのタイプに関連する新しいウィンドウが表示されます。そのプロファイリングスキームを線形（または階層型）かツリー型（または優先順位付けスキーム）に選択できます（図 2）。

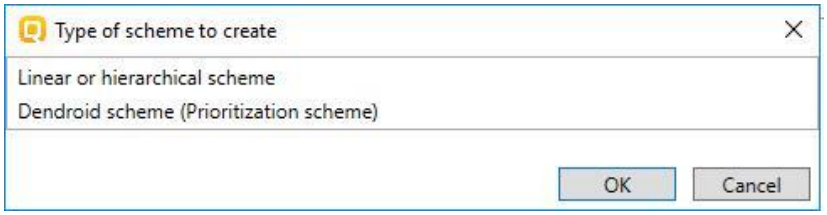


図 2

その後、Custom セクションの下に新しい空のスキームが作成され、プロファイリングスキーミングのブラウザが表示されます

2.6.1. 新しいカテゴリを追加する

新しいカテゴリを追加するには、右クリックオプションで Add Category（1）を選択してください（図 3）。カテゴリ名（2）を入力して OK をクリックすると、新しいカテゴリが作成されます。次に、領域を追加してください。

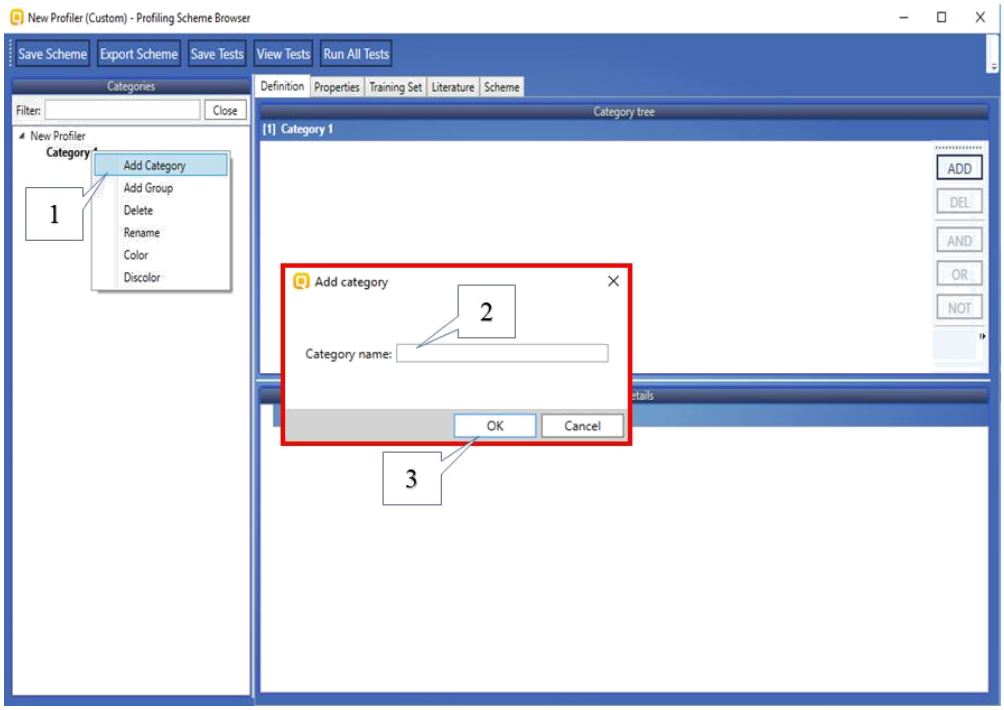


図 3

2.6.2. カテゴリにクエリを追加する

クエリを追加するには ADD（1）ボタンをクリックしてください（図 4）。カテゴリを表す 7 タイプのクエリを表示するリストが現われます（2）。

- data（データ）
- label（ラベル）
- parametric（パラメータ）
- QSAR
- referential（参照）
- similarity（類似性）
- structural（化学構造）です。

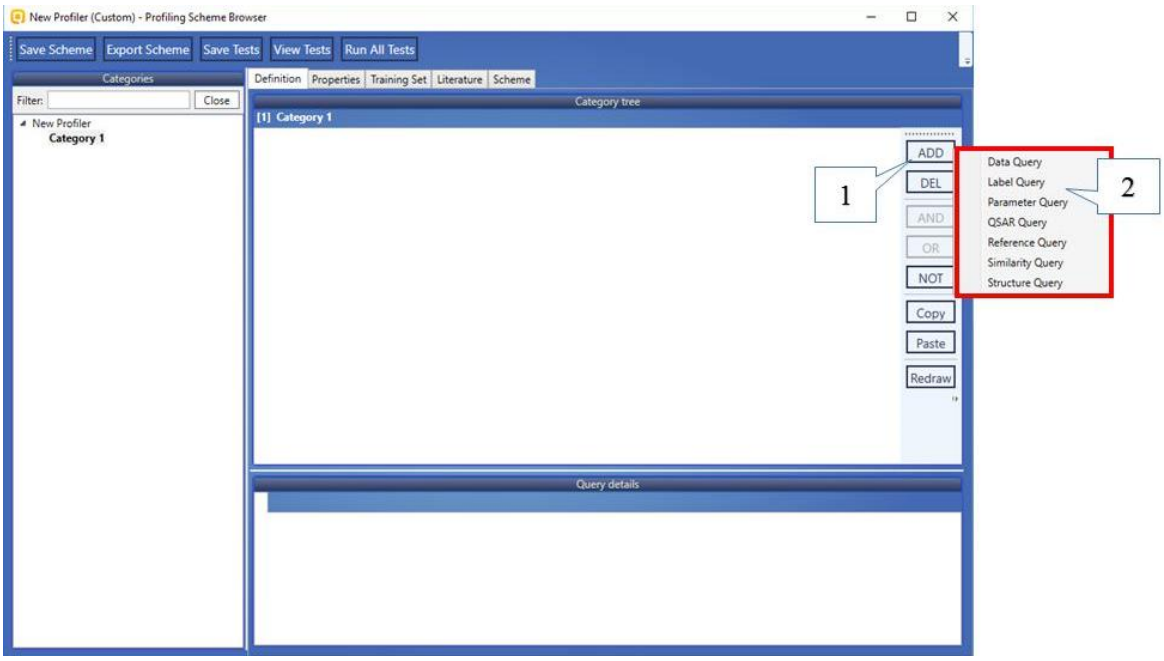


図 4

選択したカテゴリのタイプに応じて、クエリを設定するためにさまざまなオプションが用意されています。例えば、パラメータ・クエリには限界値（最小、最大、またはその両方）が必要であり、化学構造クエリのためには構造フラグメントが必要です。

2.6.3. カテゴリ・クエリの組合せ

論理上 2 つ以上のクエリを結合するためには、カテゴリスキームの論理を記述するクエリを作成する必要があります。クエリの各セットは、クエリ項（1）で表示されます（図 5）。 NOT ボタン（2）を使用すると、クエリ項を反転できます。さらに、元の項やそれから派生した項は、AND（3）または OR（4）の論理ボタンを使用して結合できます。結合させたいクエリを選択するには、Ctrl キーを押しながらマウスをクリックしてください。
下の図は、パラメータ的なものと構造的なものの 2 つのクエリセットを含むカテゴリを示しています。
さらに、これらのクエリは、次の様に組み合わされて基準を明確にします。

- 100Da を超える MW（分子量）を有すること、および
- それらの構造に C（C）= O のフラグメントを持たないこと。

カテゴリが構築された後、Save scheme スキーム（5）でスキームを保存してください（図 5）。

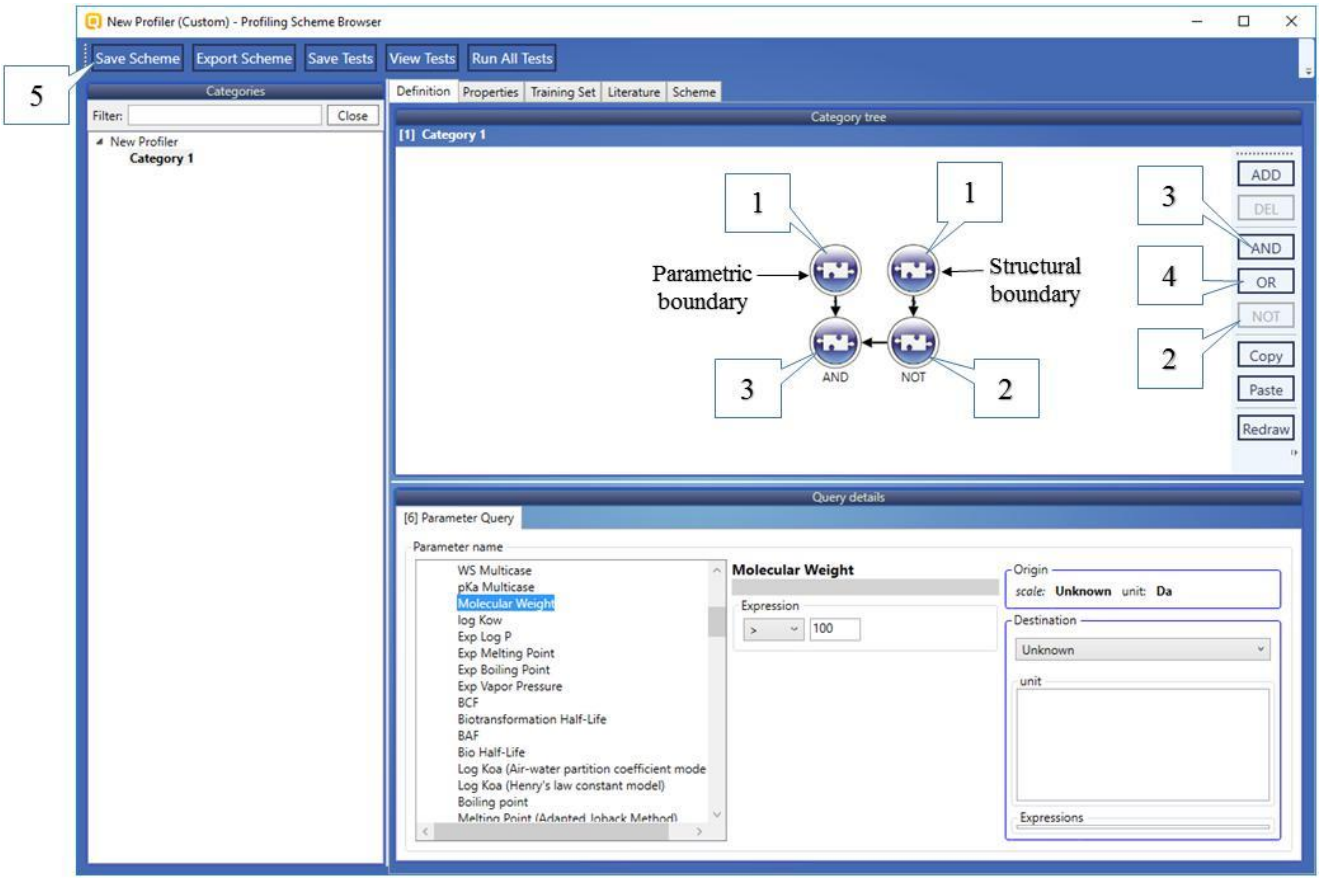


図 5

D.2.7. プロファイルを削除する

ユーザはカスタムプロファイラしか削除することができません。デフォルトのプロファイリングスキーム、観測値データベースとシミュレータは削除できません。

カスタムプロファイリングスキームを削除するには、削除するスキーム（1）を選択し、スキームを削除する Delete ボタン（2）をクリックします。（図 1)

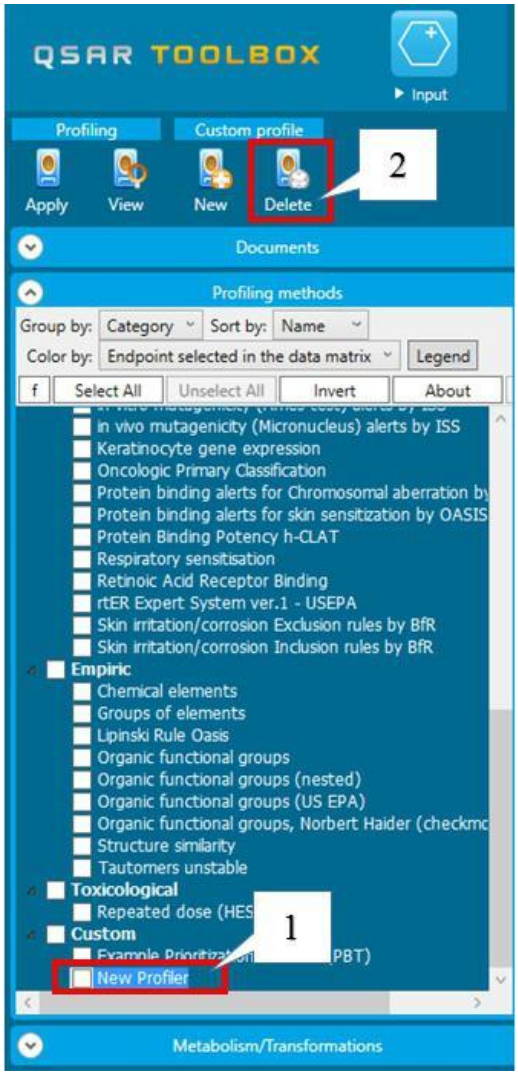


図 1

D.3. データ

Data モジュールでは、ツールボックスに導入されたデータベースから実験結果を取得できます。

エンドポイントツリーパネルで主に表示される部分を図 1 に示します。すべてのデータベースがパネル（1）にリストされ、インベントリはパネル（2）に表示されます。(図 1)

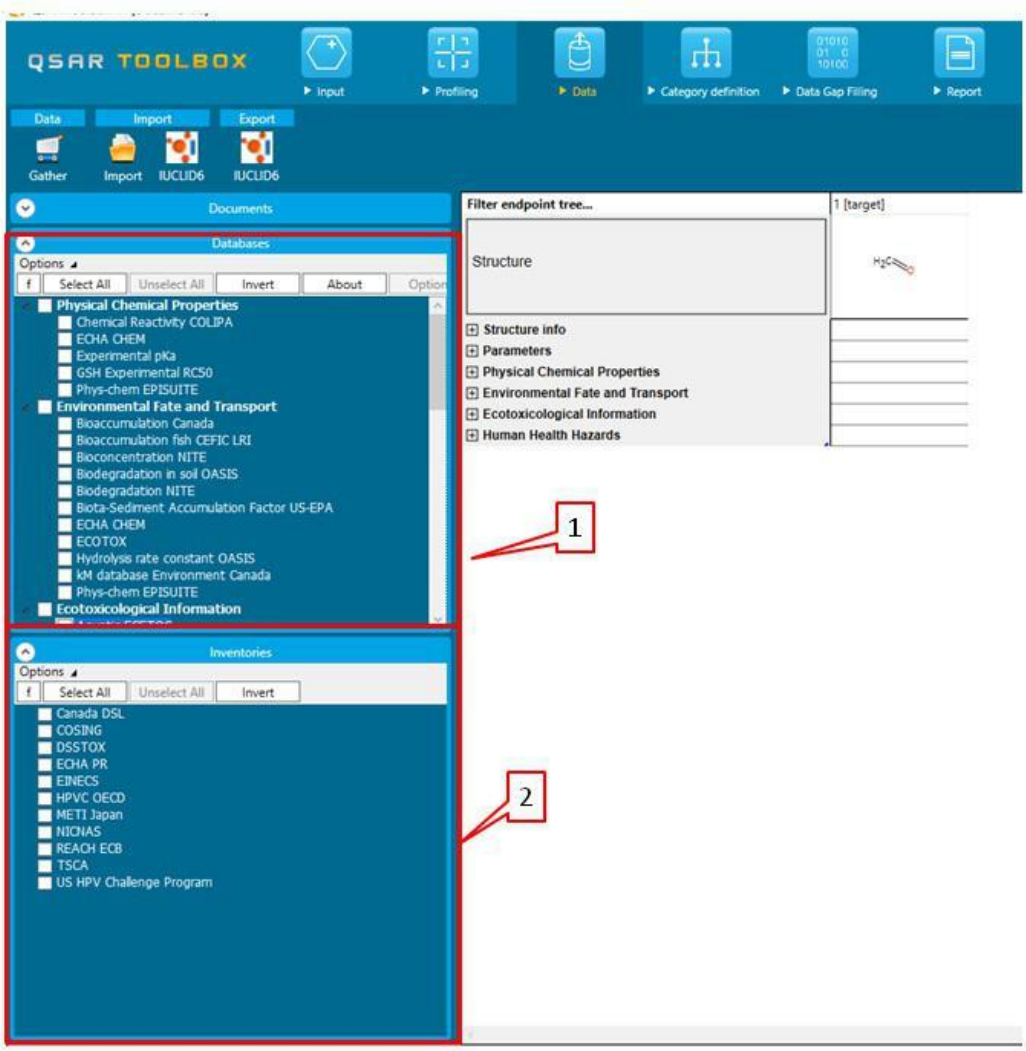


図 1

この Data セクションには、以下のサブセクションがあります。

- Databases（**データベース**）
- Inventories（**インベントリ**）

Caution

The databases include chemicals and their experimental data, while the Inventories include chemicals **ONLY**.

「注:データベースには化学物質と実験データが含まれ、インベントリは化学物質**のみ**が含まれます」

D.3.1. データベース

Toolbox の各データベースには、実測値を持つ化学物質が含まれています。化学物質は、判明している限りの CAS RN（登録番号）、化学物質名、SMILES などの化学物質 ID 情報とともに表示されます。実測値（**データ・マトリックス**で「M」と表示）は、「定量的な」値や「定性的な」値が含まれます。この点において、データの変換ができるように開発され（判定基準の変換）、利用可能な観測値をすべて確保して使用できるようになっています。判定基準と、判定基準の変換の詳細については、Options/General のセクション （訳注：E.1 章）を参照してください 。

Toolbox 4.1 で利用可能な 50 のデータベースを表 1 に示します。

表 1. Toolbox 4.1 で利用可能なデータベースのリスト

DATABASE（データベース）	
Acute Oral toxicity (急性経口毒性)	Human Half-life (ヒトの体内での半減期)
Aquatic ECETOC (水生生物の生態毒性；ECETOC)	Hydrolysis rate constant (加水分解速度定数) OASIS
Aquatic Japan MoE (水生生物の生態毒性、国立環境研)	Keratinocyte gene expression (角化細胞遺伝子発現) LuSens
Aquatic OASIS (水生生物の生態毒性；OASIS)	
Bacterial mutagenicity ISSSTY (遺伝子突然変異;伊 国衛研)	Keratinocyte gene expression (角化細胞遺伝子発現) Givaudan
Bioaccumulation Canada （水生生物の生物蓄積性；カナダ環境省)	kM database Environment Canada (代謝変換速度定数;カナダ環境省)
Bioaccumulation fish CEFIC LRI （魚の生物蓄積性；欧州化学工業連盟 LRI)	Micronucleus ISSMIC (小核試験；ISSMIC)
	Micronucleus OASIS (小核試験；OASIS)
Biocides and plant protection ISSBIOC （殺虫剤および植物保護；伊国衛研)	MUNRO non-cancer EFSA（反復投与毒性、非発がん性；欧州食品安全機関)
Bioconcentration NITE(生物濃縮;製品評価技術基盤機構)	Phys-chem EPISUITE (物理化学的性状)
Biodegradation in soil OASIS (土壌生分解性；OASIS)	Phys-chem EPISUITE (物理化学的性状)
Biodegradation NITE (生分解性;製品評価技術基盤機構)	REACH Skin sensitisation database (normalised) (標準化 皮膚感作性データベース;欧州 REACH 規則)
Biota-Sediment Accumulation Factor US-EPA (生物圏堆積物蓄積係数)	Receptor mediator effect（受容体伝達物質効果)
Carcinogenic Potency Database (CPDB) （発がん性データベース；CPDB)	Rep Dose Tox Fraunhofer ITEM (反復投与毒性；独 Fraunhofer 研)
Carcinogenicity&Mutagenicity（発がん性および変異原性) ISSCAN	Repeated Dose Toxicity (反復投与毒性) HESS
Cell Transformation Assay (細胞形質転換分析) ISSCTA	Rodent Inhalation Toxicity（齧歯類の吸入毒性) Database
Chemical Reactivity COLIPA(化学反応性；欧州化粧品工業会)	Skin irritation (皮膚刺激性)
Dendritic cells COLIPA (樹状細胞；欧州化粧品工業会)	Skin sensitization (皮膚感作性)
Developmental toxicity ILSI (発生毒性；ILSI)	Skin sensitization (皮膚感作性) ECETOC
ECHA CHEM	Toxcast DB
ECOTOX	Toxicity to reproduction (生殖毒性)
ECVAM genotoxicity and Carcinogenicity(遺伝子毒性および発がん性)	ToxRefDB US-EPA (毒性参照データベース;米国環境保護庁)
Experimental pKa (解離定数 pKa 実測値)	Toxicity Japan MHLW (毒性；国衛研)
Eye irritation ECETOC （眼刺激性;ECETOC)	Transgenic rodent (遺伝子形質転換齧歯類) database
Genotoxicity OASIS （遺伝毒性;OASIS)	Yeast estrogen assay database University of Tennessee-Knoxville (USA)
GSH Experimental RC50(グルタチオン反応性)	ZEBET database (ZEBET データベース)

さまざまなタイプの実験データに基づいて、データベースの構成が開発されました。データベースの編成は、エンドポイントツリーの既定ノードの構成に従います。
データベースは次の 4 つの基本セクションに分類されています。

- Physical Chemical Properties （物理化学的性質）
- Environmental Fate and Transport（環境運命と輸送）
- Ecotoxicological Information （生態毒性情報）
- Human Health Hazards（ヒトの健康への有害性）

注：1 つのデータベースに複数のセクションのエンドポイントが含まれている場合は、各セクションに渡ってそのデータベース名があります。このようなデータベースの例として、すべてのセクションに含まれている ECHA CHEM があります。
いくつかのオプションが導入されています（1）。これは、データベースを選択する際の補助目的です（図 1）。

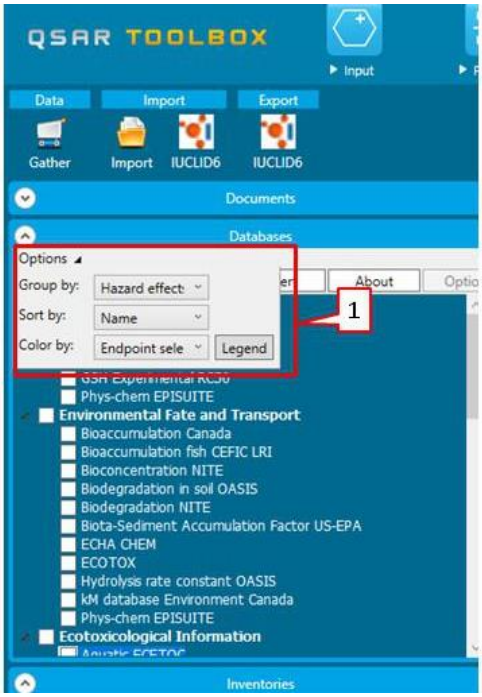


図 1

- ***Group by オプション***（図 2-1）：有害性の影響に基づいてデータベースをグループ化できます。

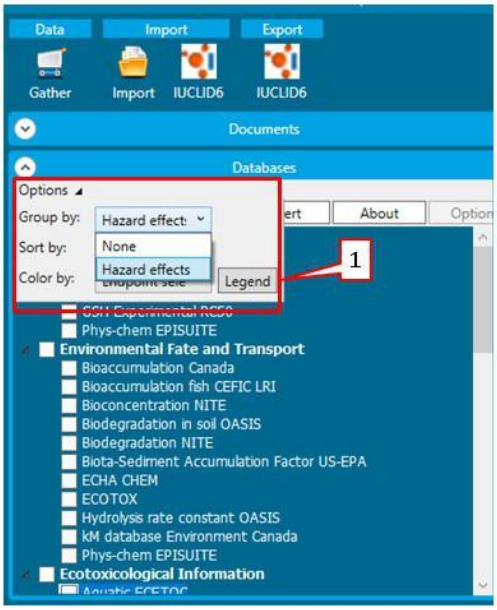


図 2

- ***Sort by オプション*** (図 3-1): データベース名で並べられます。

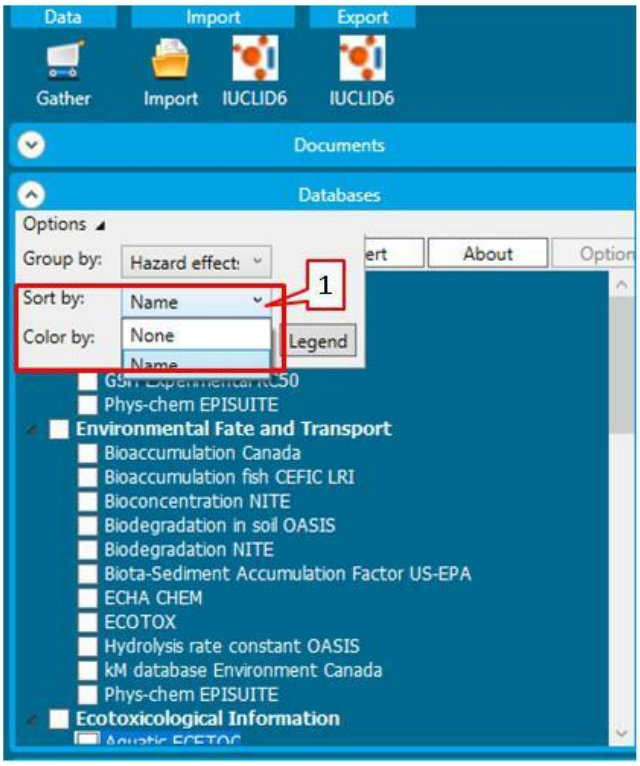


図 3

- ***Color by オプション*** (図 4-1) in Figure 4 : **データ・マトリックス**内の選択されたエンドポイントまたは標的エンドポイントの実測データを持つデータベースを色付けできます。

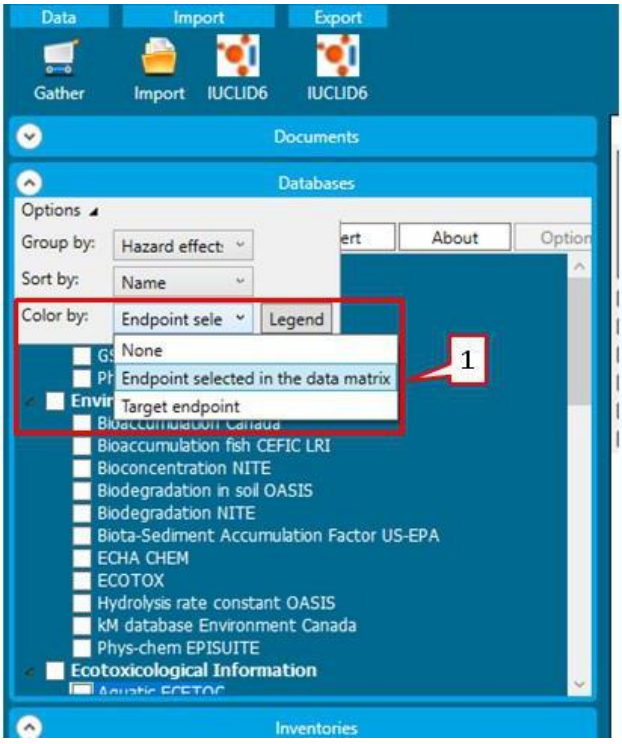


図 4

Color by endpoint selected in the data matrix を選択すると、エンドポイントが**データ・マトリックス**で選択されます。例えば Aquatic toxicity（水生毒性）(2) の場合、エンドポイントの測定データを含む関連データベースは緑色（4）に着色されます（図 5)。また、配色（3）について説明する legend（凡例）もあります。

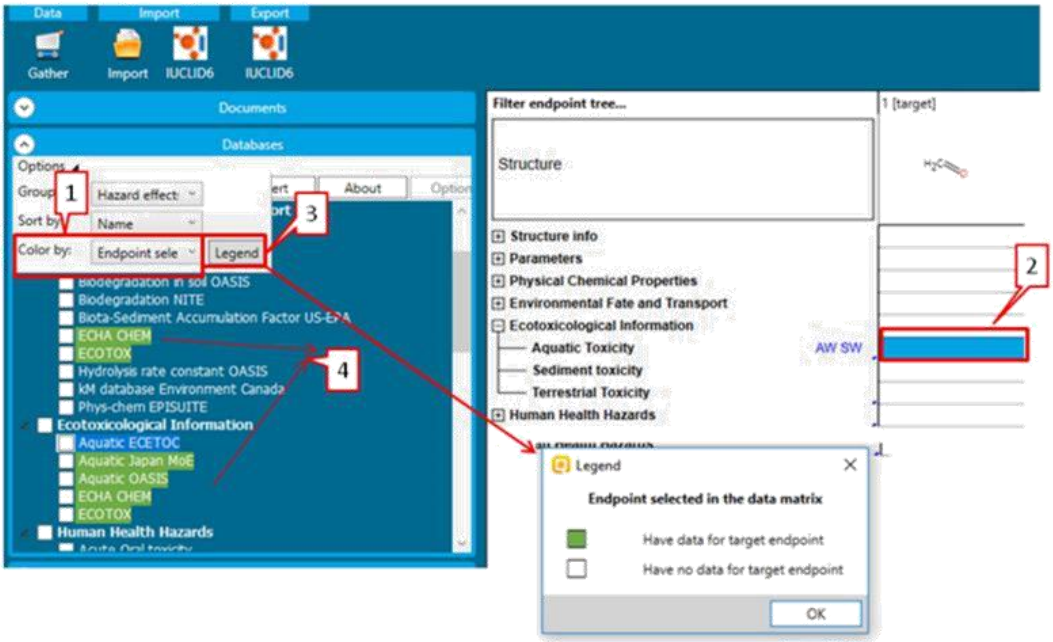


図 5

データベースが強調表示されたら、データベースの一部または全部を選択してから（1）、Gather ボタン（2）をクリックします（図 6）。ポップアップウィンドウ（3）が表示され、すべての・または選択されたデータポイントのデータを収集するオプションが表示されます。Choose オプション（4）をクリックすると、エンドポイントツリーが現れ、ドロップダウンメニュー矢印から関心のあるエンドポイント、例えば Aquatic toxicity（水生毒性）を選択することができます。OK（6）をクリックします。

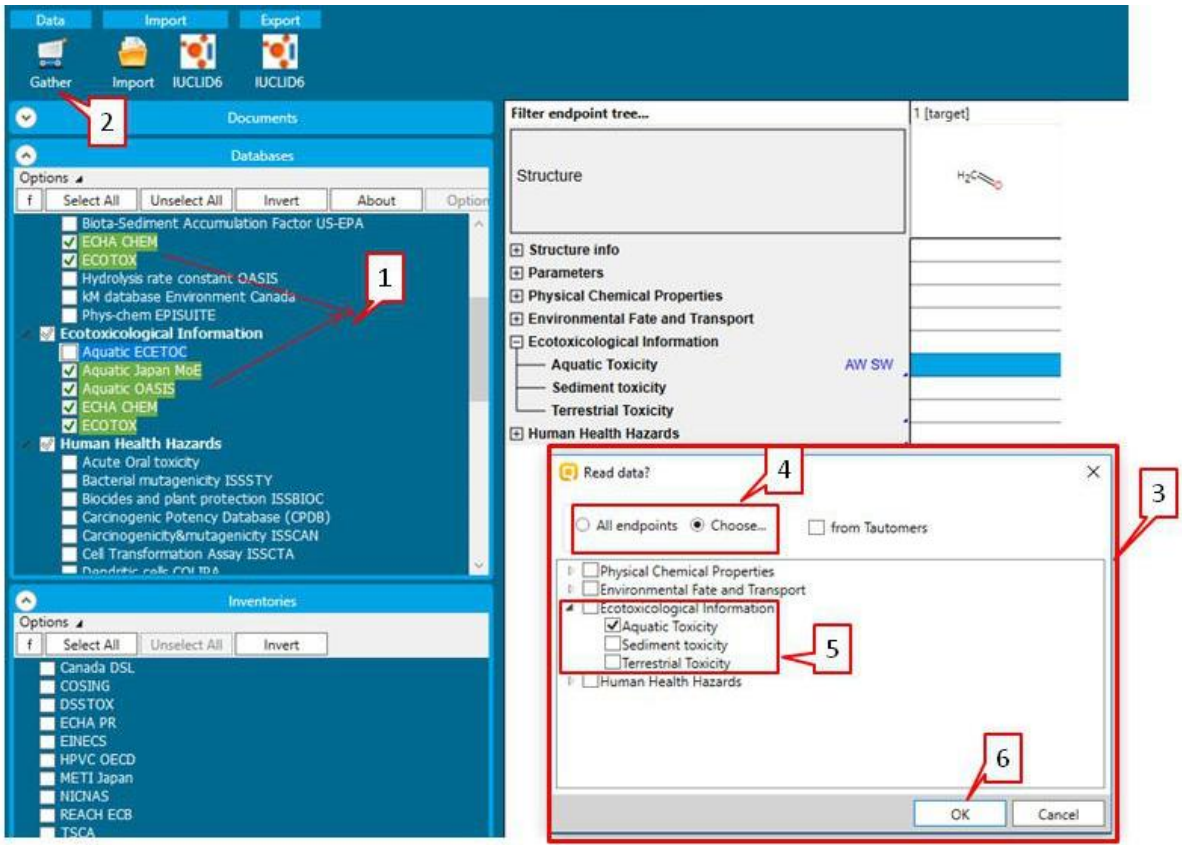


図 6

情報ウィンドウ（図 7-1）は、選択された化学物質について収集された測定データの数に関する詳細を示します。収集されたすべてのデータが**データ・マトリックス**に表示されます（図 7−2）。エンドポイントツリーを展開して、測定データを持つノードを表示します（図 8 の 1〜4）。

図 7

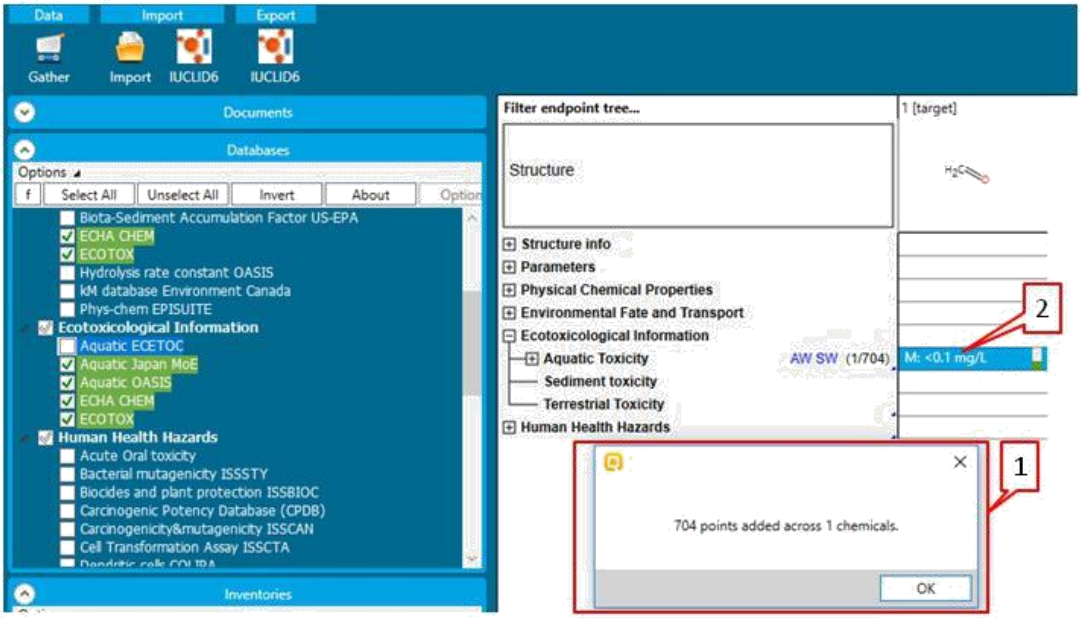
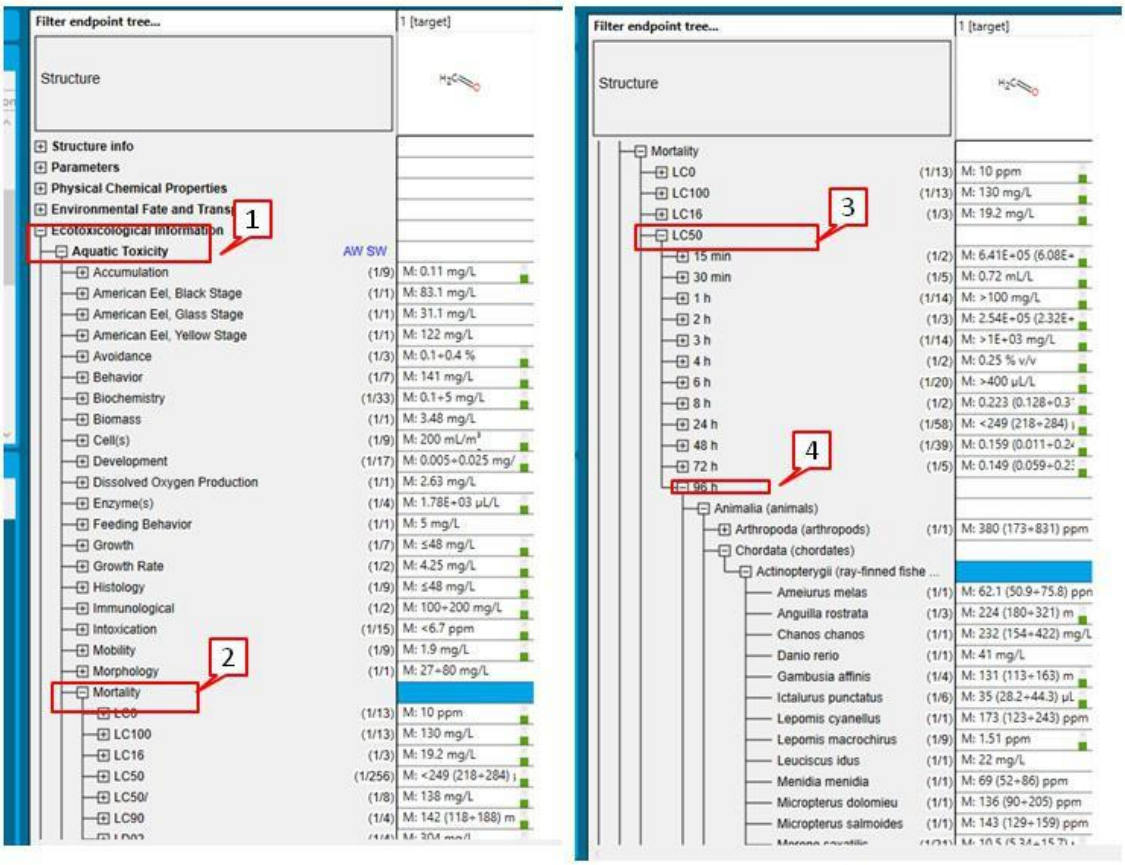


図 8



選択オプションは、データベースの上にあるボタンで、またはデータベースの 1 つを右クリックしてアクティブにすることができます（図 9）。

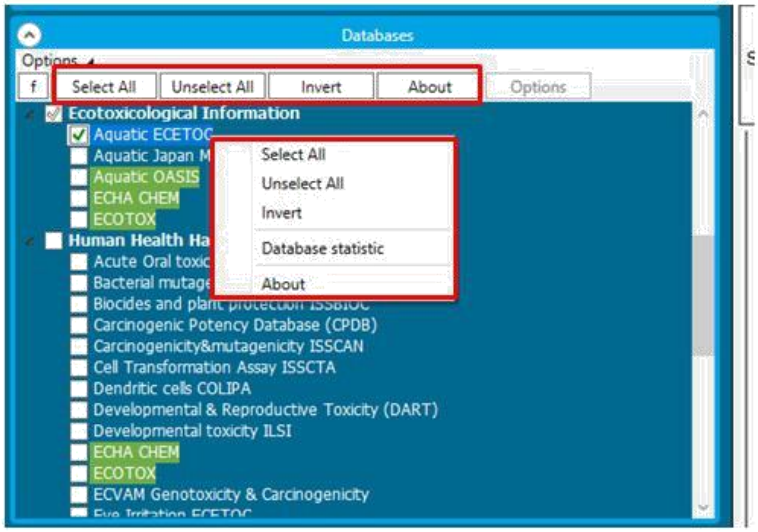


図 9

1 - Select All - すべてのデータベースを選択。Select All（1）をクリックすると、すべてのデータベースが選択されます（2）。 （図 10）

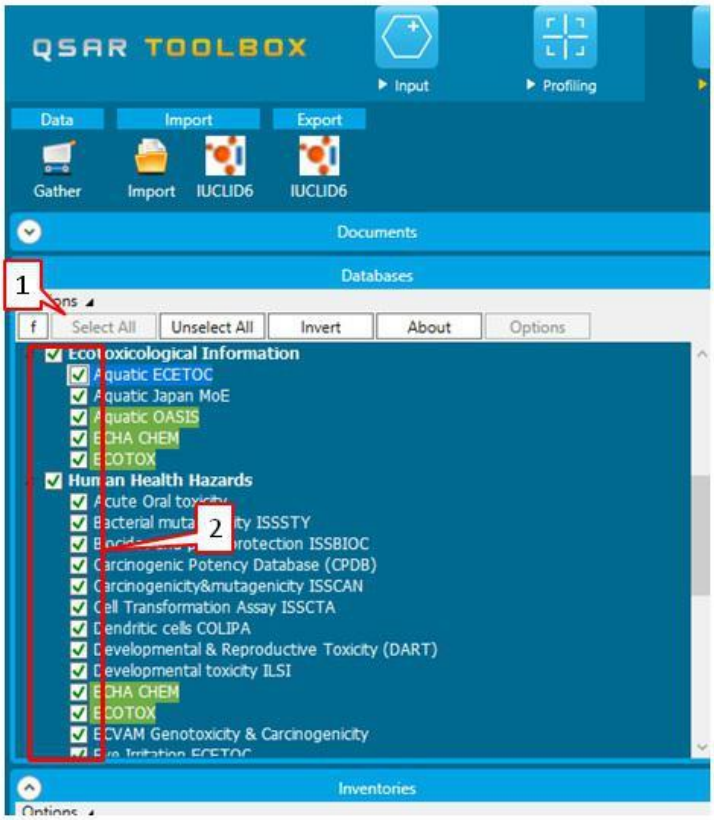


図 10

2 - Unselect All - すべてのデータベースの選択を解除。 Unselect All（1）をクリックすると、すべてのデータベースの選択が解除されます（2）。

(図 11)

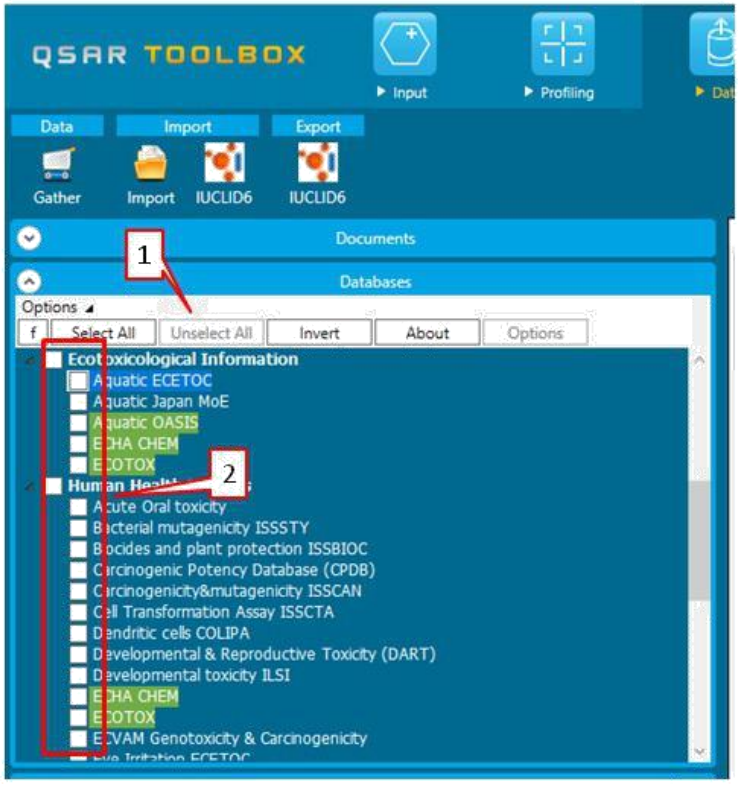


図 11

3 - Invert – 直前に実行されたデータベースの選択を反転します。例えば、ユーザが Bioconcentration NITE（1）を選択し、Invert（2）をクリックすると、ソフトウェアは、Bioconcentration NITE（3）を選択解除しつつ他のすべてのデータベースを選択します（図 12）。

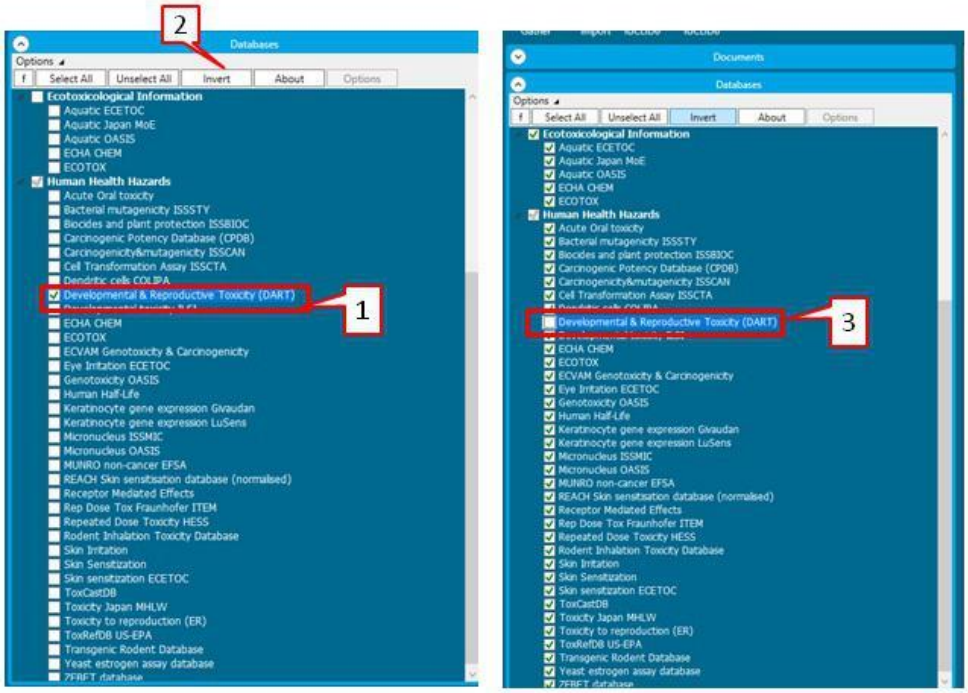


図 12

4 - About - 強調表示されたデータベース（1）の簡単な説明を表示します（図 13）。 About（2）をクリックすると、簡単な説明が表示されます。図 13 のように Documentation ボタン（3）をクリックすると、データベースの内容（4）の詳細を表示します。

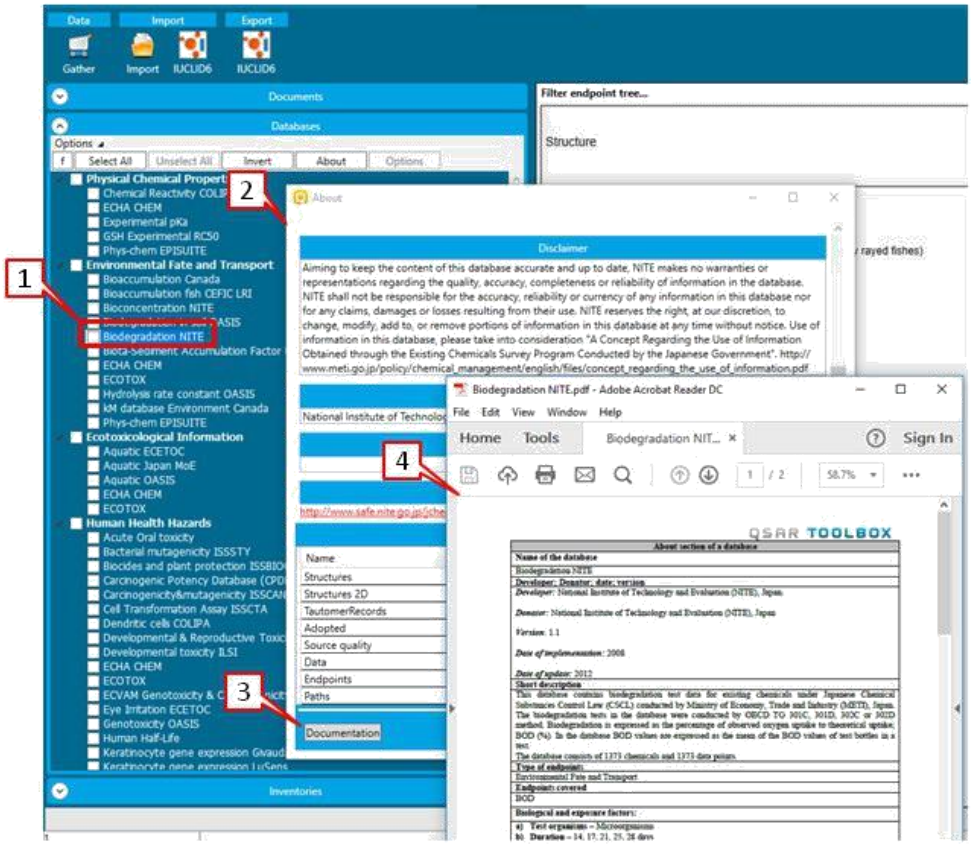


図 13

フィルタオプション（図 14－1）は、データベース内でのクイック検索をするために導入されました。選択すると、新しい入力欄（2）が表示されるので、調査したいデータベース名の一部または全部を入力して検索します。

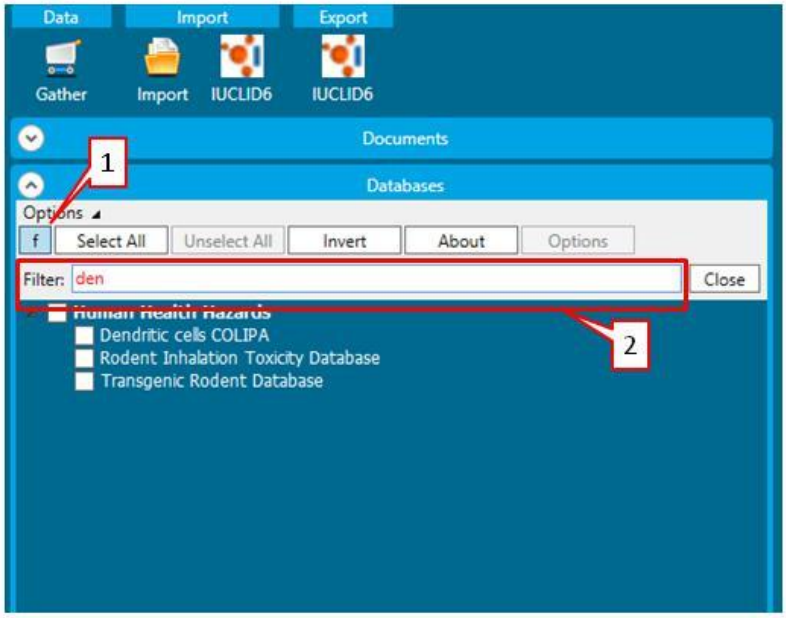
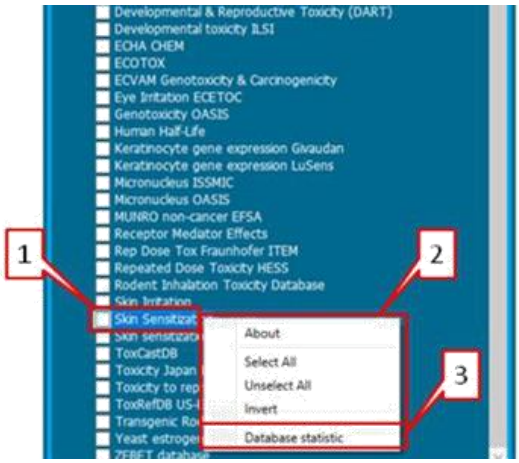


図 14

- **データベースに関連する品質属性：*Database statistics(データベース統計)***

各データベースを右クリックします。 例えば Skin Sensitization（1）をクリックすると、ウィンドウ（2）が開くので、Database statistics（3）をクリックしてください（図 15）。

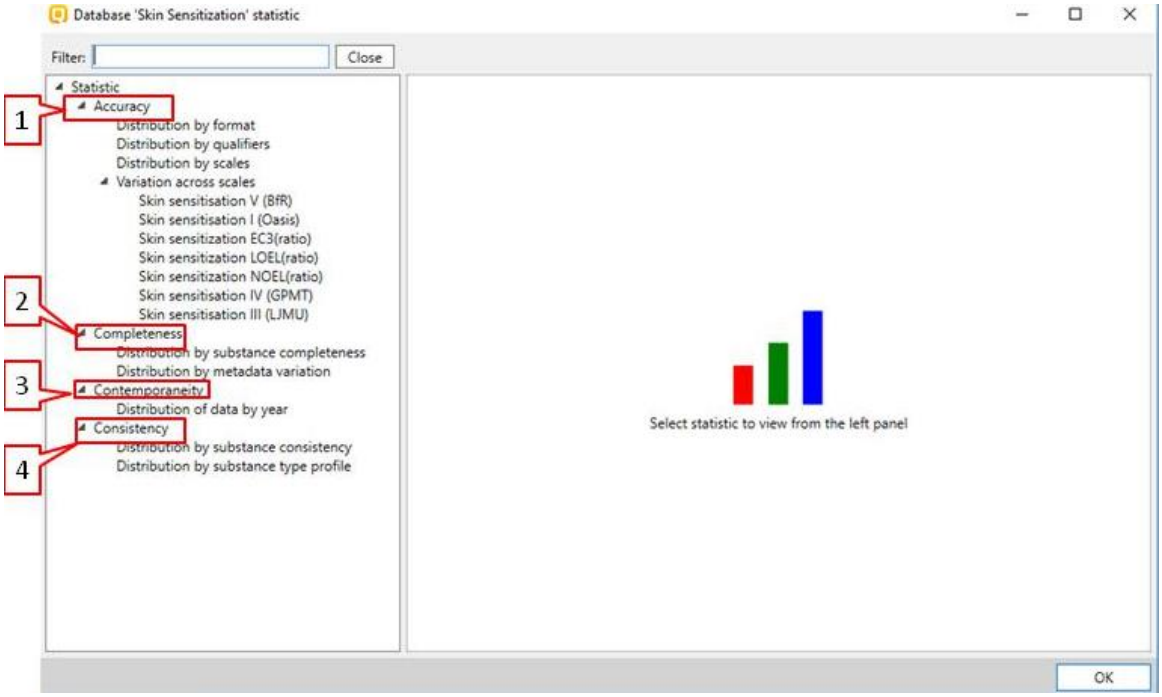
図 15



Skin Sensitization（皮膚感作）データベースのデータベース統計を図 16 に示します。

（訳注：図 16～26 までデータベースの中の皮膚感作データベースを例として取り上げ、どの程度信頼できるかを統計的に表示させるデモンストレーションを行っている）

図 16



この統計は、1）Accuracy（正確性）、2）Completeness（完備性）、3）Contemporaneity（同時性）および 4）Consistency（一貫性）の 4 つの属性からなっています。

1)　Accuracy（正確性）

精度とは、真の値または真であると受容される値に対し、実測値がどれくらい近いかです。これは、現象の実測値とは通常、真の値の推定値であるだけだと考えられることを示しています。「真の値」は実際にはわからないため、正確性を評価することは非常に困難です。 以下の「正確性の指標」が導入されています。：

- Format（書式：カテゴリまたは数値データの値：最小、最大、平均形式）
- Qualifiers（符号：　<、>、 ~）
- Scales（判定基準、判定基準のタイプ）
- Variation across the scales (data categories)（判定のバラツキ（データカテゴリ））

Accuracy の下にある各項目の詳細

- **Format（書式）** - データが数値なのかカテゴリ値（訳注：判定結果が＋＋やーなどにカテゴリ分けされるようなデータ）なのかという分布を提供します。数値データは最小値、最大値および平均値を備えていることがあります。手順に従ってください（図 17）。：

- Distribution by format をクリックする（1）
- バー（訳注：図 17 の緑のグラフ柱）を左クリックして保持すると、そのタイプの形式に該当する化学物質の数が表示されます（2）。

図 17 に示すように、731 個の数値データがあり（バー上で左クリック）、それらはすべて平均値です。もうひとつのバーはカテゴリ値をもつものの総数です。

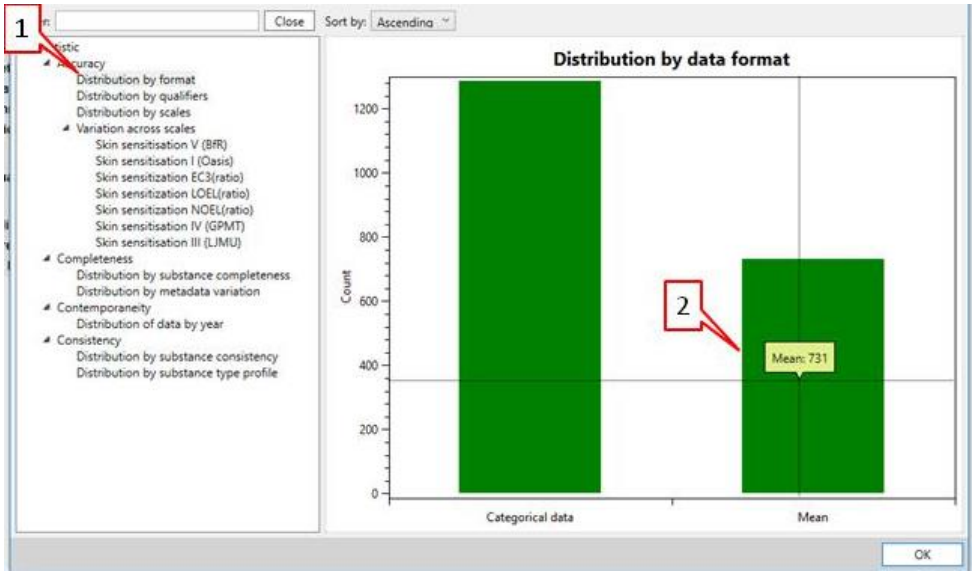


図 17

- ***Qualifiers*（符号）** - 数値データには符号が付くことがありますが、カテゴリ別データに修飾子はありません。符号を持つデータ値は現在、絶対値として扱われています。手順に従ってください（図 18）。：

- Distribution by qualifiers（1）をクリック – SS（皮膚感作）データベースのカテゴリデータエントリは 1288 件 （2）です。

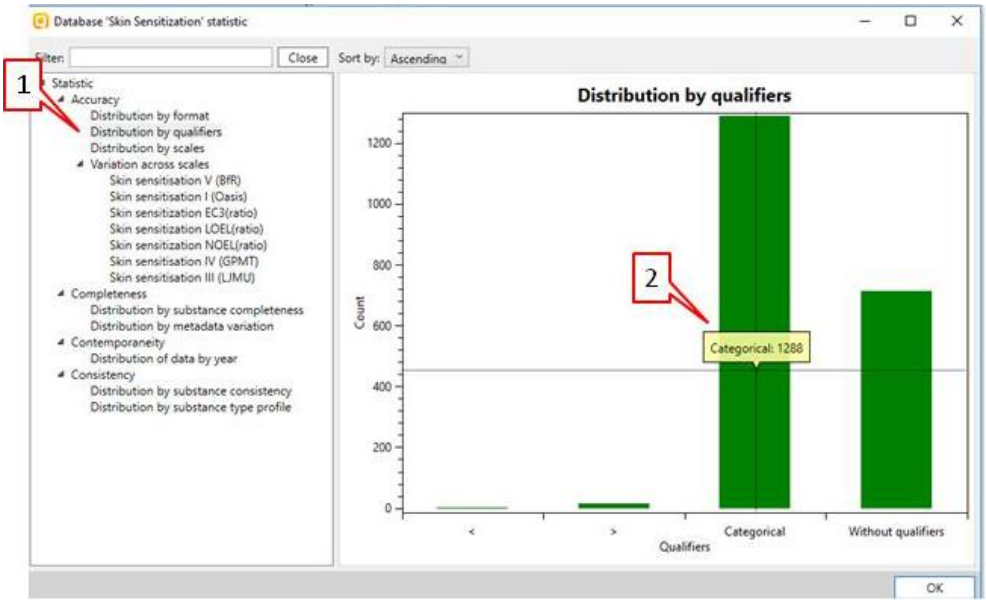


図 18

- **Scales（判定基準　訳注：試験結果が陽性かどうかなど）** - オリジナルの判定結果に従ったデータ分布を提供します（図 19）。:

Distribution by scales（1）をクリック – 濃度による判定基準（ratio scale）は最も情報量のある判定基準です。皮膚感作 IV（GPMT）が ratio scale の例です。

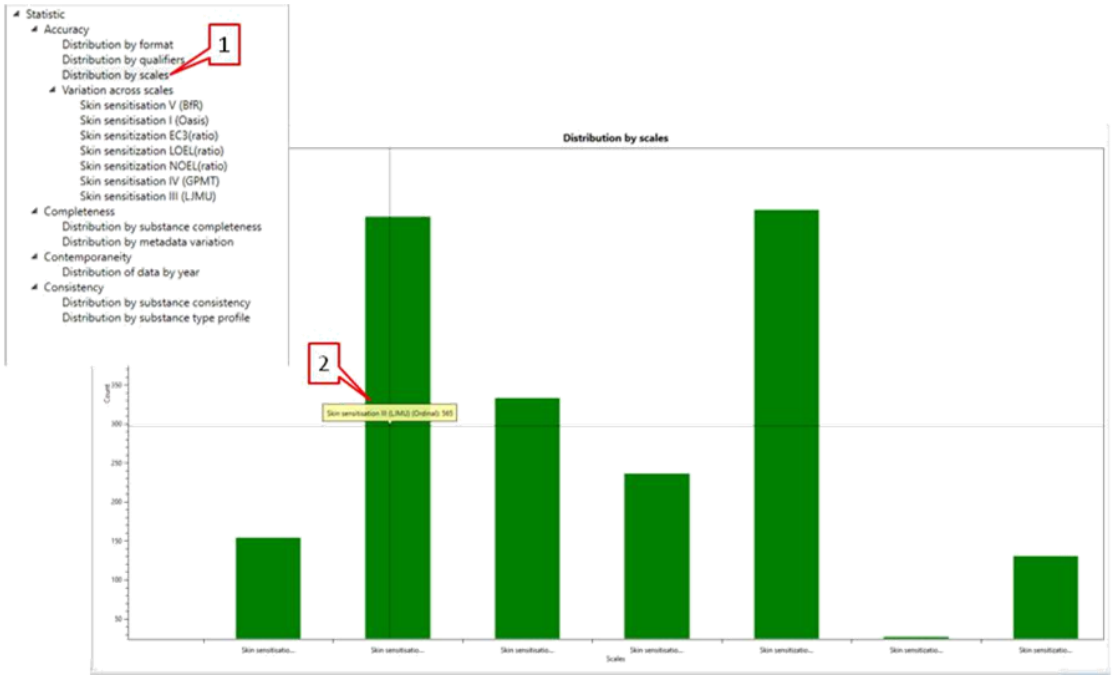


図 19

- **Variation of data across the scales（データの判定基準のばらつき）** - この精度の評価は、元のデータベースファイルで使用可能な判定基準に応じていくつかのサブレベルに分けられます。現在、カテゴリ的データ（訳注：陽性陰性などのデータ）を持つ判定基準にのみ適用されます。 1 つのデータベースに存在する化学物質は、同一の判定基準について例えば陽性、陰性など違う判定結果のデータを有することもあります。これらの化学物質は、オリジナルのデータに従ってバーに分配されます。 2 つ以上のデータを持つ化学物質は、One category、Neighboring categories、Distant categories、All categories のグループに分かれています。

- *One category*（判定結果が同一カテゴリ値の中にある）
  - いずれの試験結果判定も同一であるデータを持った化学物質の数を表します。
- *Neighboring categories*（複数の判定結果が、近いカテゴリ値の中にある）
  - 近い判定結果（例えば、皮膚感作性 が弱い陽性の化学物質と、皮膚感作性が強い陽性の化学物質）の間でデータが小さくばらつくような化学物質の数を表します。
- *Distant categories*（判定結果が大きく違うカテゴリ値）
  - 試験判定結果に大きな差があるデータを持つ化学物質の数を表します（例えば、陽性と陰性の判定結果データを持つ化学物質）。
- *All categories*（全カテゴリ値に結果がちらばったもの）
  - オリジナルの判定基準において、すべてのカテゴリ値にデータがばらつく化学物質の数を表します。

次の手順に従ってください（図 20）。

- Variation across scales をクリックします（1）
- Skins sensitization III（LJMU）**判定基準**をクリックします（2）
- ご覧の通り、ある化学物質について大部分のデータは、One category に分類されています（3）。

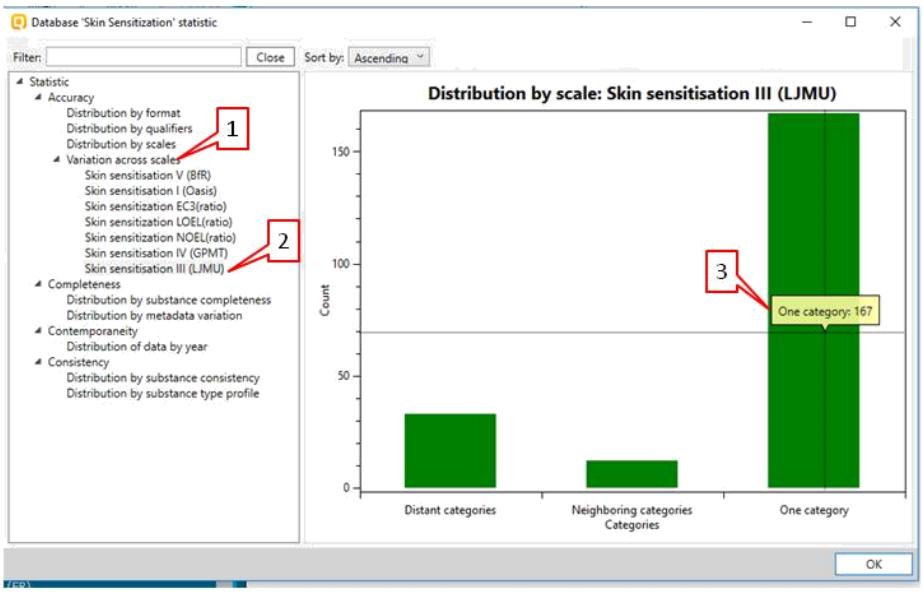


図 20

2) Completeness（完備性）

- ***Substance completeness***（物質の完備性）- 選択されたデータベースの CAS 登録番号および SMILES といった一般物質 ID の可用性と関連性をユーザに通知します。次の手順に従ってください（図 21）。
    - Completeness を展開する（1）
    - Distribution by substance completeness をクリックする （2)
    - CAS 登録番号や SMILES を有する各化学物質の数を見るには、バー（例えば、 Yes: CAS / SMILES）を左クリックする（3）



図 21

- ***Metadata completeness***（付随データの完備性　訳注：試験生物種等の情報のこと） - いくつのデータが追加情報を持っているかを示します。1 つまたは複数の標的エンドポイントを選択後、どのタイプの付随データを見たいかを選択することができます。データベースに必要な情報がない場合は、Undefined のバーが現われます。この場合、皮膚感作性のデータ中の試験方法の分布が示されます。次の手順に従ってください（図 22）。

- Completeness を展開する（1）
- Distribution by metadata completeness をクリックする(2)
- Select endpoint ウィンドウが表示される（3）
- Human health hazard を拡張して skin sensitization の階層を表示させる（4）
- 使用可能なすべてのエンドポイントを選択する（5）
- Metadeta を開き、Assay（6）を選択してから、Add ボタン（7）をクリックする
- OK（8）をクリック

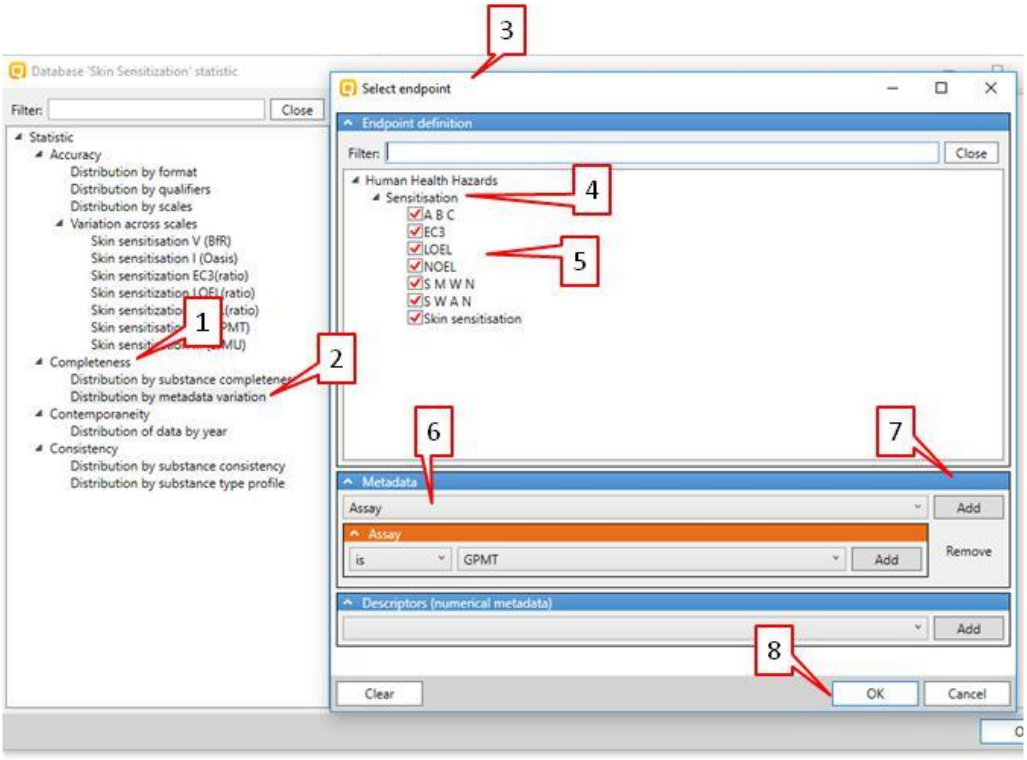


図 22

その結果、分類は次のようになります（図 23 参照）。

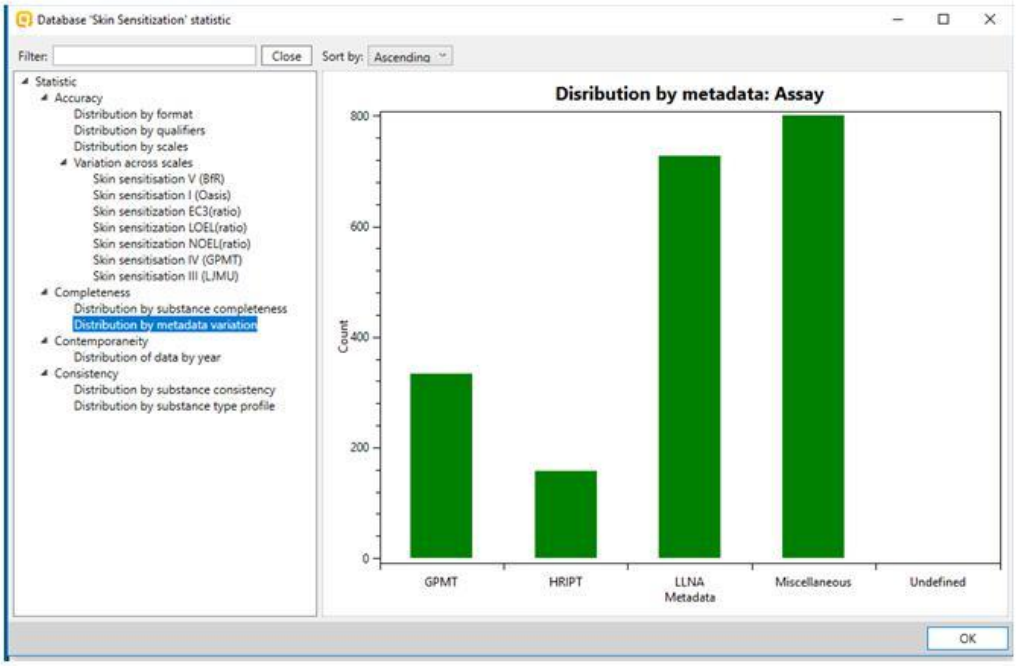


図 23

3) Contemporaneity（年代）

選択したデータベースのすべてのデータを公開年次に従って分類することができます。
次の手順に従います（図 24）。:

- Contemporaneity のレベル（1）を展開します。
- Distribution by year (2)をクリックします。
- バーを左クリックすると、所定の年に試験された化学物質の数が表示されます（3）。

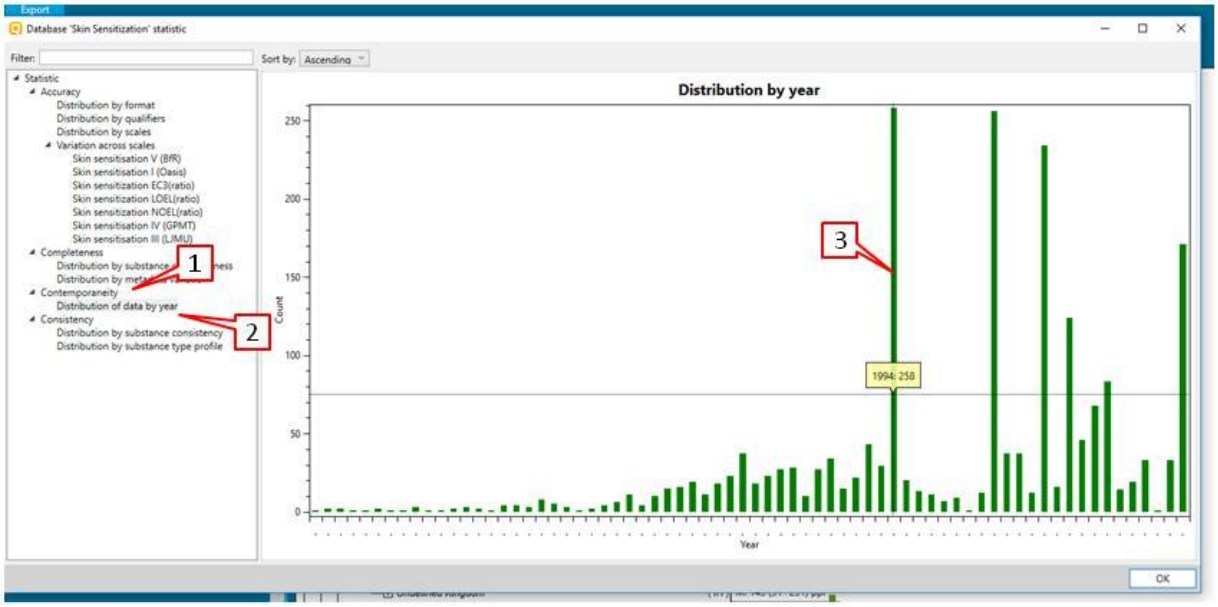


図 24

4)　　Consistency（整合性）

-　　***Substance consistency***（物質の整合性）- 現在のデータベースで CAS 登録番号を割り振られた SMILES 表記の数に関する情報を提供します。複数の構造に対応する CAS 登録番号を持つ化学物質がある場合には、表示されます。図 25 に示す手順に従います。:

- Consistency （1）をクリックします。
- Distribution by substance consistency （2）をクリックします。
- 1236 の CAS 登録番号に SMILES（3）が付いています。

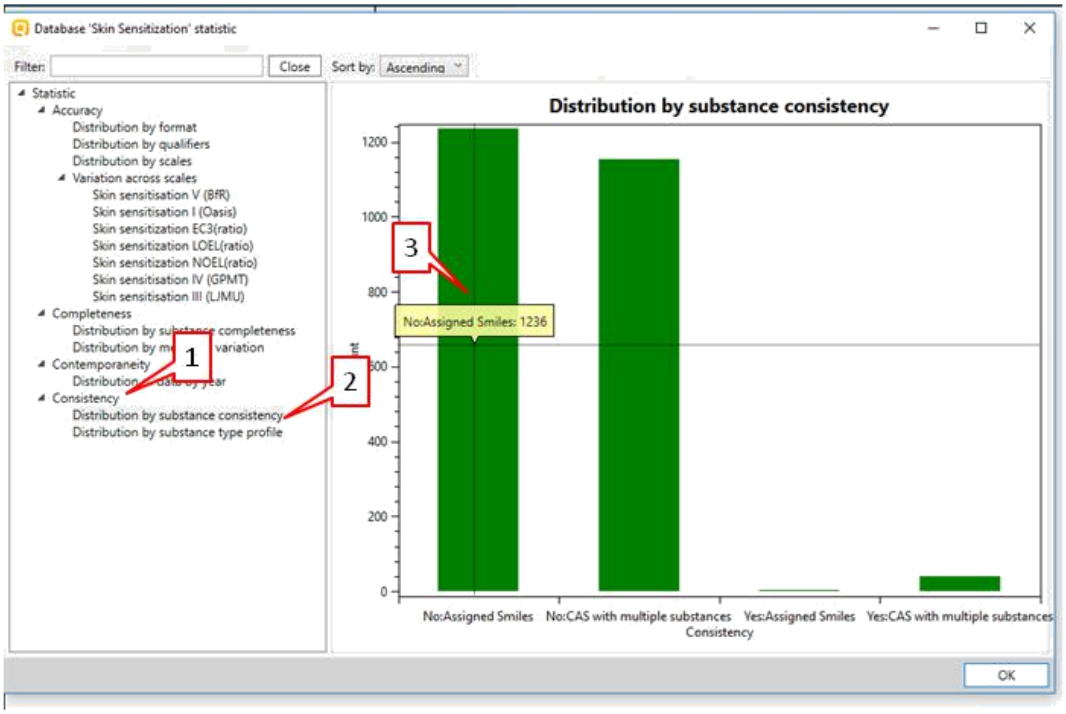


図 25

- ***Substance type profiler***（物質タイプ・プロファイラ） - 物質のタイプごとの統計をとり、現在のデータベースの化学物質に複数の構成物質や UVCB が存在するかどうかをユーザに通知します。図 26 に示される手順に従います。（訳注：物質が混合物かどうか、イオンかどうかなど）

  - Consistency （1）をクリックします。
  - Distribution by substance type profiler（2）をクリックします。
  - descrete chemicals（混合物ではなく単一である化学物質）が皮膚感作データベース（バー（3）を左クリック）で 1181 個あると識別されています。

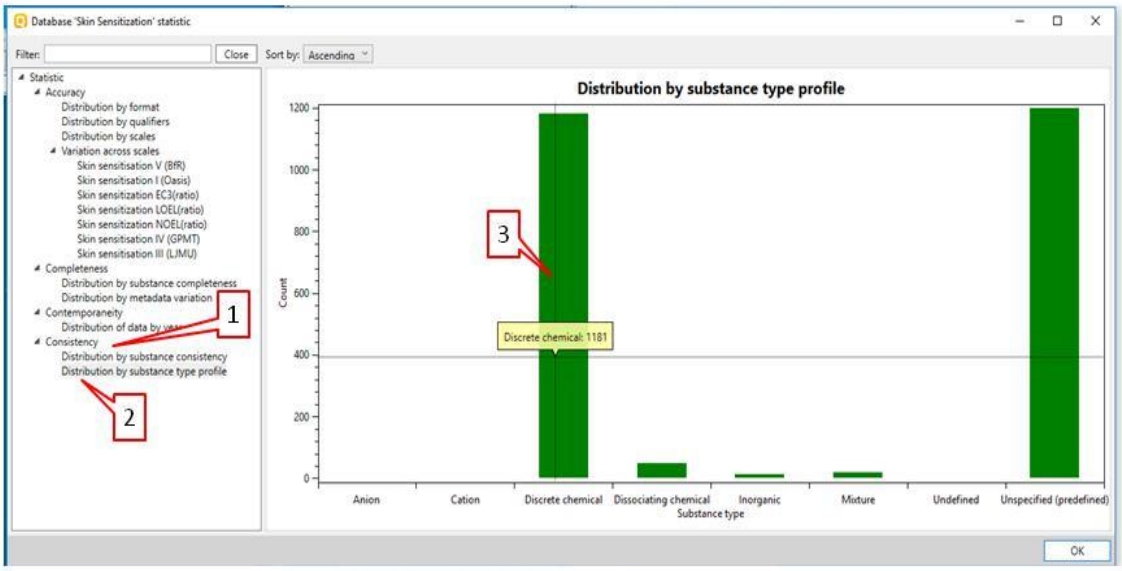

Database 'Skin Sensitization' statistic
Filter:
Close
Sort by:
Ascending
Statistic
Accuracy
Distribution by format
Distribution by qualifiers
Distribution by scales
Variation across scales
Skin sensitisation V (BfR)
Skin sensitisation I (Oasis)
Skin sensitization EC3(ratio)
Skin sensitization LOEL(ratio)
Skin sensitization NOEL(ratio)
Skin sensitisation IV (GPMT)
Skin sensitisation III (LJMU)
Completeness
Distribution by substance completeness
Distribution by metadata variation
Contemporaneity
Distribution of data by year
Consistency
Distribution by substance consistency
Distribution by substance type profile
1
2
3
Distribution by substance type profile
Count
1200
1000
800
600
400
200
0
Discrete chemical: 1181
Anion
Cation
Discrete chemical
Dissociating chemical
Inorganic
Mixture
Undefined
Unspecified (predefined)
Substance type
OK

図 26

D.3.2. インベントリ

Toolbox 4.0 の利用可能なインベントリを表 2 に示します。データベースには利用可能な実験データを持つ化学物質が含まれ、インベントリには化学物質のみが含まれています。インベントリの化学物質は、CAS、化学名、SMILES とともに表示されます。

表 2. Toolbox 4.0 で利用可能なデータベースのリスト

Inventries （インベントリ）

NICNAS （豪；国家工業化学品届出審査機構）	EINECS （欧州既存商業化学物質リスト）
Canada DSL （カナダ；国内物質リスト）	HPVC OECD （OECD 高生産量化学物質リスト）
COSING （欧州規制下の化粧品材料・添加物データベース	MITI Japan 通産省（現：経産省）
DSSTox （米国環境保護庁；検索可能な毒物のデータベース）	REACH_ECB （欧州化学物質局）
ECHA PR （欧州化学品庁）	TSCA （米国化学物質法規制）
US HPV Challenge Program （米；高生産量化学物質チャレンジプログラム）	

データベースと同じ選択肢を利用できます。 - select （選択する）(1)、unselect （選択を解除）(2)、 invert select（選択を反転） (3)、 About（詳細） (4) (図 1)

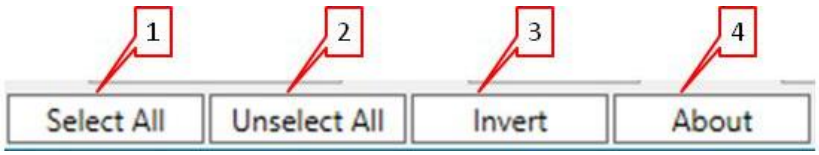


図 1

1- Select All – すべてのインベントリを選択。

2- Unselect All – すべてのインベントリの選択を解除。

3- Invert –最後に実行されたインベントリの選択を反転します。

4- About – 強調表示されたインベントリの簡単な説明を示します。インベントリ（1）を強調表示し、About（2）をクリックすると、短い説明用ウィンドウ（3）が表示されます（図 2）。

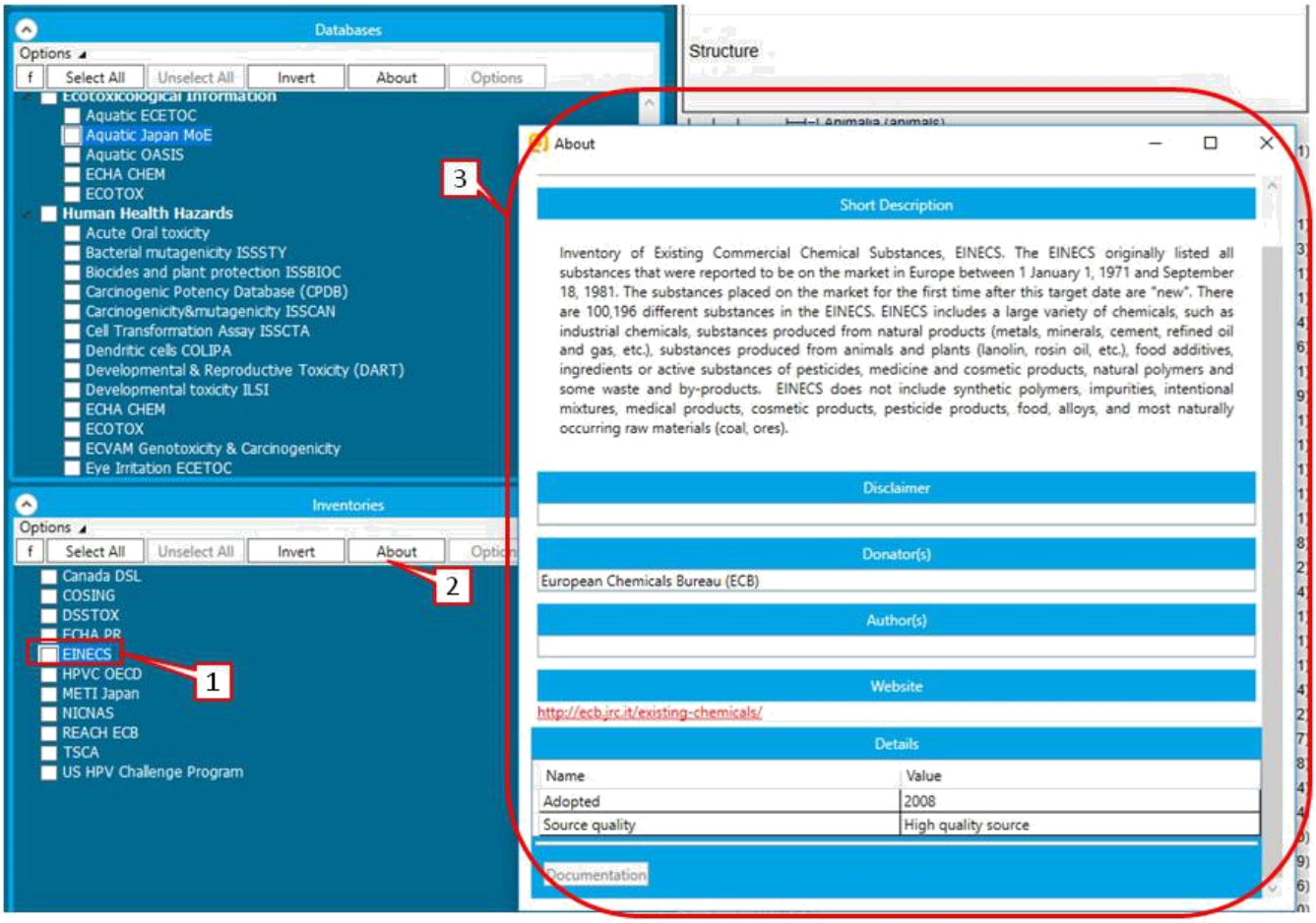

Databases
Options
Select All
Unselect All
Invert
About
Options
Ecotoxicological Information
Aquatic ECETOC
Aquatic Japan MoE
Aquatic OASIS
ECHA CHEM
ECOTOX
Human Health Hazards
Acute Oral toxicity
Bacterial mutagenicity ISSSTY
Biocides and plant protection ISSBIOC
Carcinogenic Potency Database (CPDB)
Carcinogenicity&mutagenicity ISSCAN
Cell Transformation Assay ISSCTA
Dendritic cells COLIPA
Developmental & Reproductive Toxicity (DART)
Developmental toxicity ILSI
ECHA CHEM
ECOTOX
ECVAM Genotoxicity & Carcinogenicity
Eye Irritation ECETOC
Inventories
Options
Select All
Unselect All
Invert
About
Canada DSL
COSING
DSSTOX
ECHA PR
EINECS
HPVC OECD
METI Japan
NICNAS
REACH ECB
TSCA
US HPV Challenge Program
1
2
3
Structure
About
Short Description
Inventory of Existing Commercial Chemical Substances, EINECS. The EINECS originally listed all substances that were reported to be on the market in Europe between 1 January 1, 1971 and September 18, 1981. The substances placed on the market for the first time after this target date are "new". There are 100,196 different substances in the EINECS. EINECS includes a large variety of chemicals, such as industrial chemicals, substances produced from natural products (metals, minerals, cement, refined oil and gas, etc.), substances produced from animals and plants (lanolin, rosin oil, etc.), food additives, ingredients or active substances of pesticides, medicine and cosmetic products, natural polymers and some waste and by-products. EINECS does not include synthetic polymers, impurities, intentional mixtures, medical products, cosmetic products, pesticide products, food, alloys, and most naturally occurring raw materials (coal, ores).
Disclaimer
Donator(s)
European Chemicals Bureau (ECB)
Author(s)
Website
http://ecb.jrc.it/existing-chemicals/
Details
Name
Value
Adopted
2008
Source quality
High quality source
Documentation

図 2

D.3.3. データを集積する

データ収集は、次の 2 つの基本的な方法のいずれかで実行できます。

- ***全エンドポイントの全データの収集：*** Select all （1）をクリックしてすべてのデータベースを選択し、 Gather （2）をクリックして、すべてのエンドポイントのデータを抽出するよう確定します（3）（図 1）。（訳注：先に Input モジュールで調べたい物質を入力していないと Data モジュールは使えません）

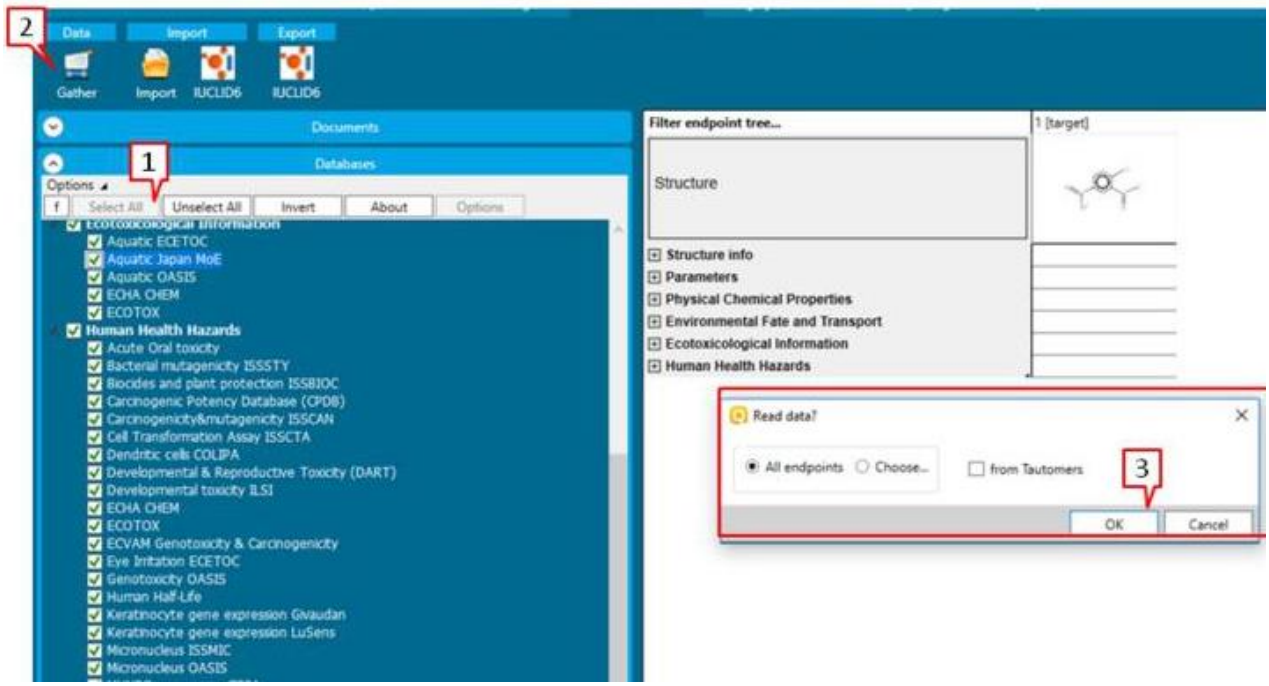


図 1

表示された情報メッセージで OK をクリックします（図 2）。

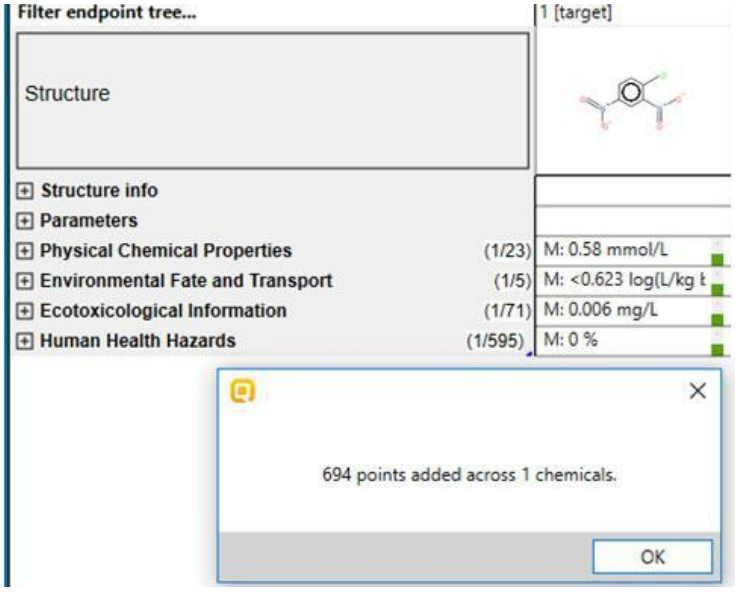


図 2

すべての利用可能な測定データは、(a) **データ・マトリックス**（2）に読み込まれた化学物質について抽出されます。実験データを持つ化学物質のデータポイント数も示されています（1）。

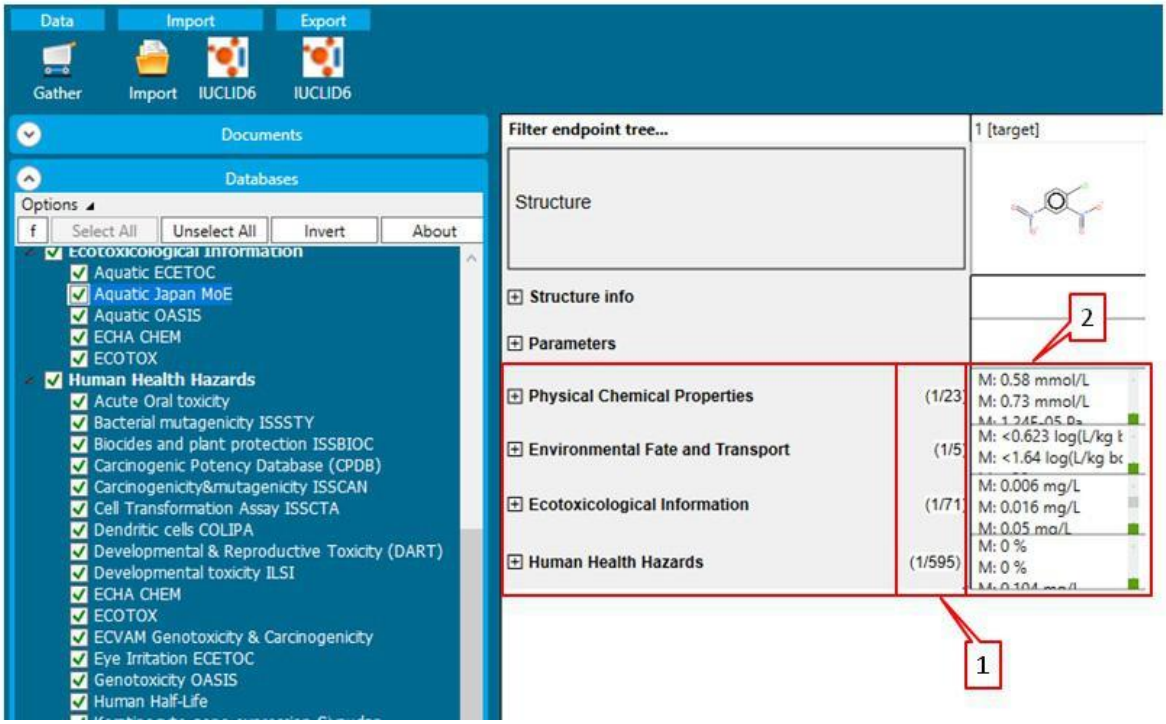


図 3

- **一つのエンドポイントについてのデータ収集**（たとえば、一つあるいは限定数のエンドポイントのデータ収集）：研究されたエンドポイント（1）に関するデータを持つデータベースを選択します。例えば Aquatic ECETOC を選び、Gather（2）をクリックします（図 4）。**データ・マトリックス**に化学物質の測定データを読み込ませる前に、Read data? のダイアログウィンドウが表示され（3）、 all endpoints (4) または Choose an endpoint (4)の選択を促します。（訳注：ここで Choose を選択すると図 4-5 の部分が開く）該当するエンドポイントが選択されたら（5）、OK（6）をクリックします。

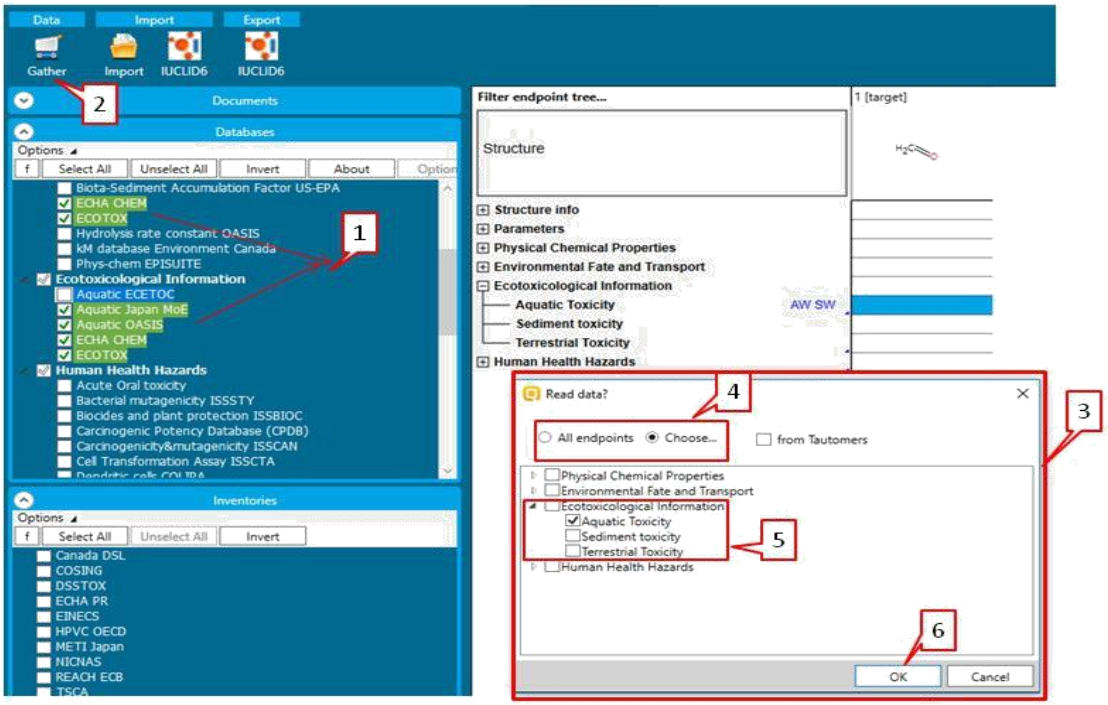


図 4

測定データの数に関する情報メッセージが表示されます（1）（図 5）。標的化学物質について、選択したエンドポイントのすべての測定データが**データ・マトリックス**（2）に表示されます。

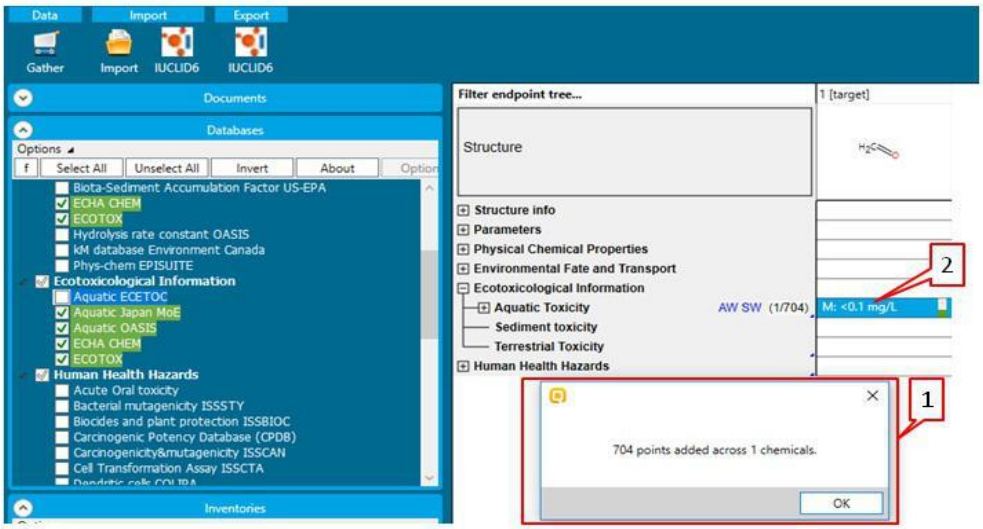


図 5

エンドポイントツリー（1-4）を展開すると、収集された付随データが表示されます（図 6）。

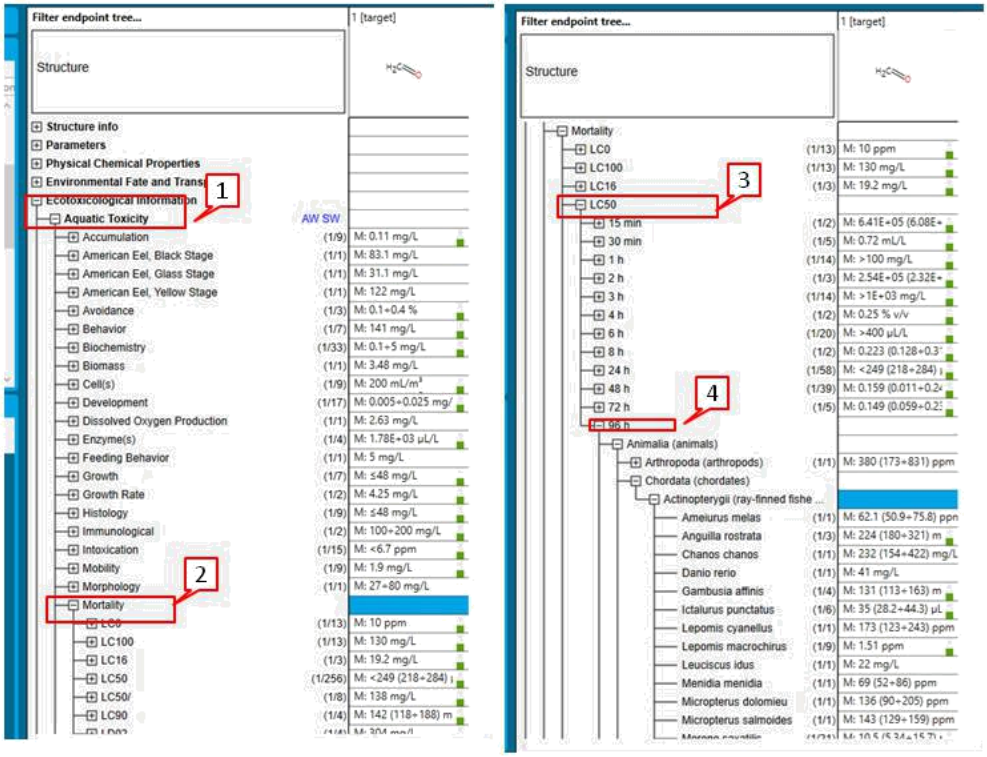


図 6

注意:

1. ツールボックス・データベースには実験データのある化学物質が含まれていますが、インベントリには化学物質のみが含まれており、実験データは含まれていません。

2. データベースは先にプロファイリング・計算をされ、結果はキャッシュに保存されます。これは、実験データ（カテゴリーを決める）で類似物質を検索するプロセスをさらに早めるために行われます。

D.3.4. データをインポートする

2 種類のインポートがあり、Excel ファイルとタブ区切りファイル経由で使用することができます。両方の場合の手順は同じです。

Toolbox 4.0（以降）は、次のデータ・タイプで動作します（図 1）。

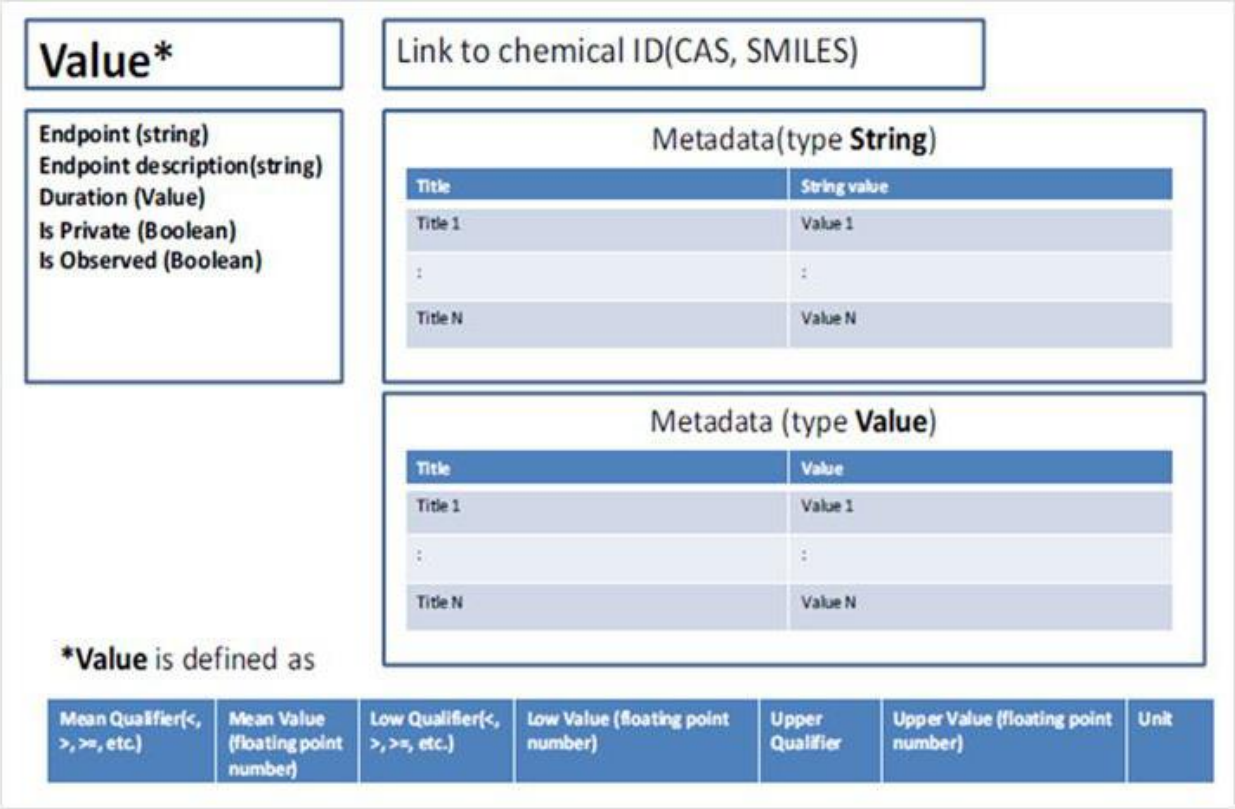


図 1

インポート機能は、ファイルから情報を翻訳し、これを別々のまとまりになるよう区切り（上の図を参照）、それらをデータベースに組み込むことです。情報は、関連する化学物質、数値、付随データからなります。言い換えれば、インポートの要所は、データポイントのリスト（ユーザが**データ・マトリックス**内で見てギャップ補完に使用する数）を、それに対応する付随データ、すなわち試験期間、試験生物、エンドポイントなどに関する追加情報とともに決定することです。

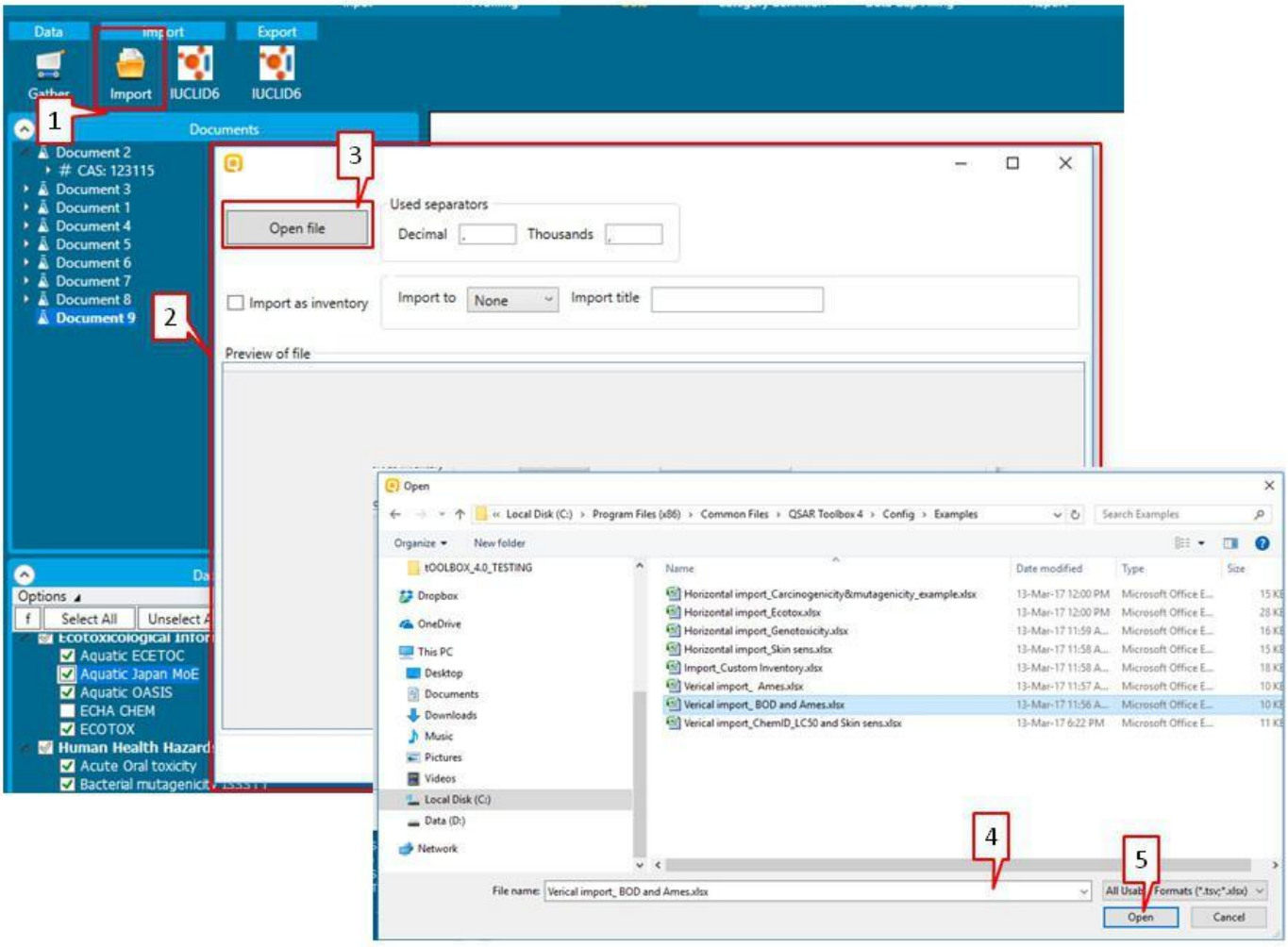
D.3.4.1. Vertical import（縦断的なインポート）: データベース

（訳注：Toolbox をダウンロードするとサンプルのデータベース Excel ファイルがついており、この章でインポートの練習ができる）

バーティカル（縦断的な）インポートは、一貫した実験データと、同一のサポート情報（エンドポイント、試験生物、試験条件、作成者など）を持つ一連の化学物質に適しています。

カスタムデータベースをインポートするには、次の手順を実行してください。:

- Import button (1)をクリックします（図 1)。インポートウィンドウが表示されます（2)。Open file (3) をクリックし、（訳注：ここでローカルディスク C:¥Program Files (x86)を開ける）Common files/QSAR Toolbox 4.1/Config/Examples のサンプルファイルから、データベースの Vertical import _BOD and Ames.xlsx (4)を選択します。Open(5)をクリックします。



- データベースのデフォルト名は、ファイルの名前+追加番号と同じになります (図 2-1)。ファイルからの情報はタイトルの列（2）に示されています。Next (3)をクリックします。

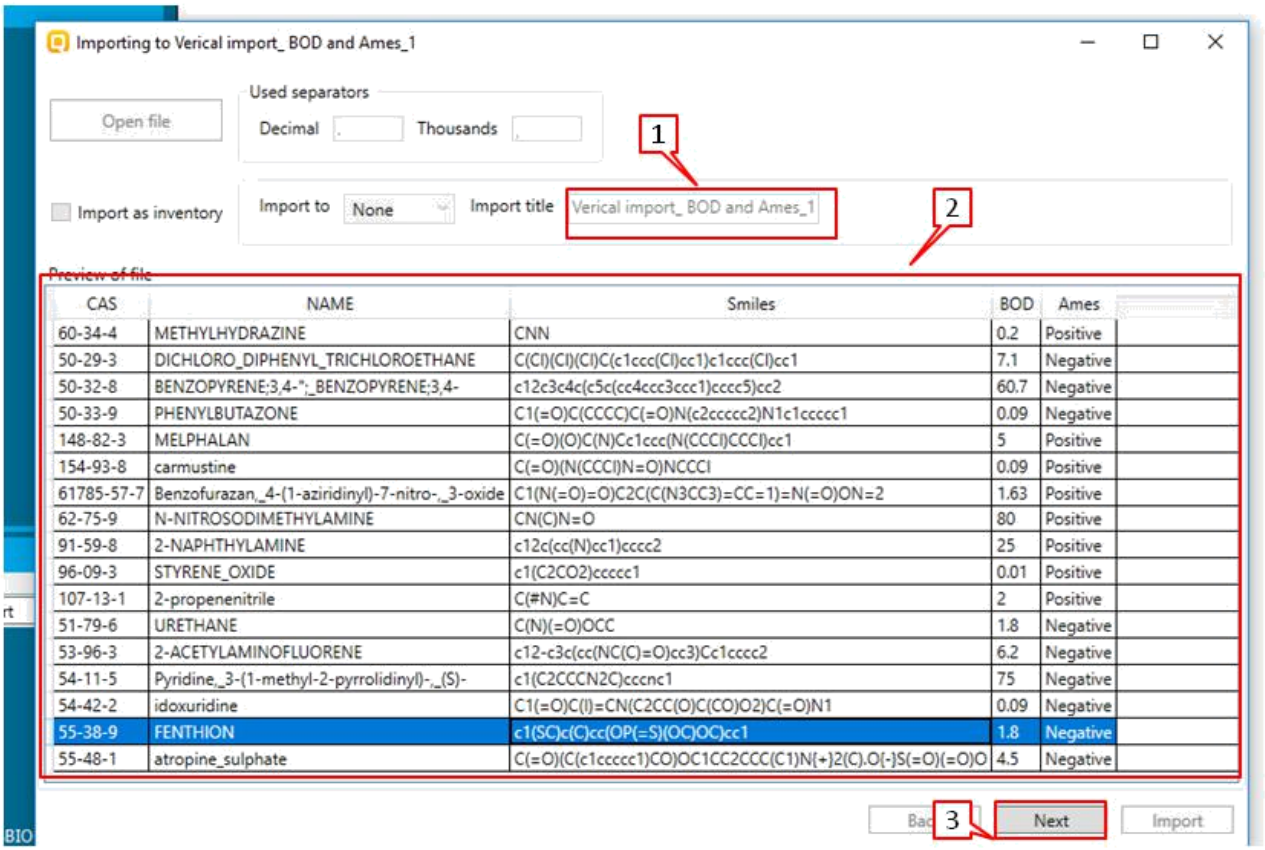


図 2

- 図 3 の Vertical radio ボタン（1）を押します。エンドポイント名（3）の下の各 No endpoint selected ボタンを個別にクリックして、BOD（生物学的酸素要求量）および Ames (Bacterial reverse mutation assay) （2）に関わる関連エンドポイントを決定します。

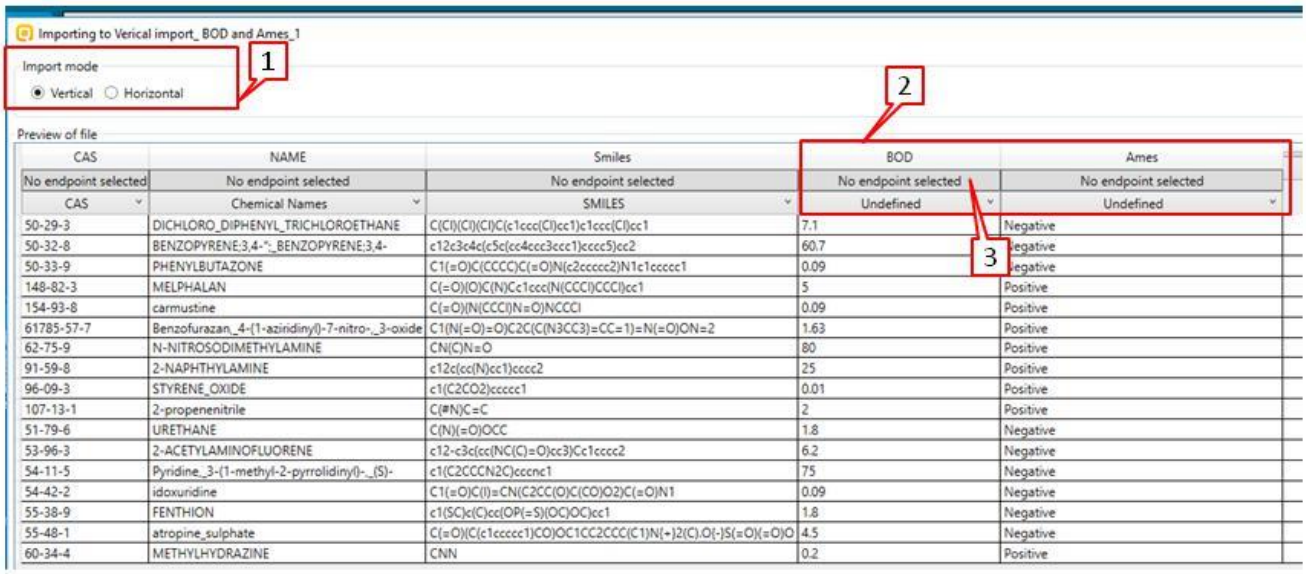


図 3

- **<u>BOD</u>**: エンドポイント選択ウィンドウが表示されます。No endpoint selected (1)をクリックし、フィルターに bio を入力し（2)、エンドポイントツリーから Ready Biodegradability（3）を選択します。Family / Scale のドロップダウンメニューで、フィルタ入力欄（5）に Bio と入力し、Biodegradability（%）（6）を選択します（図 4)。

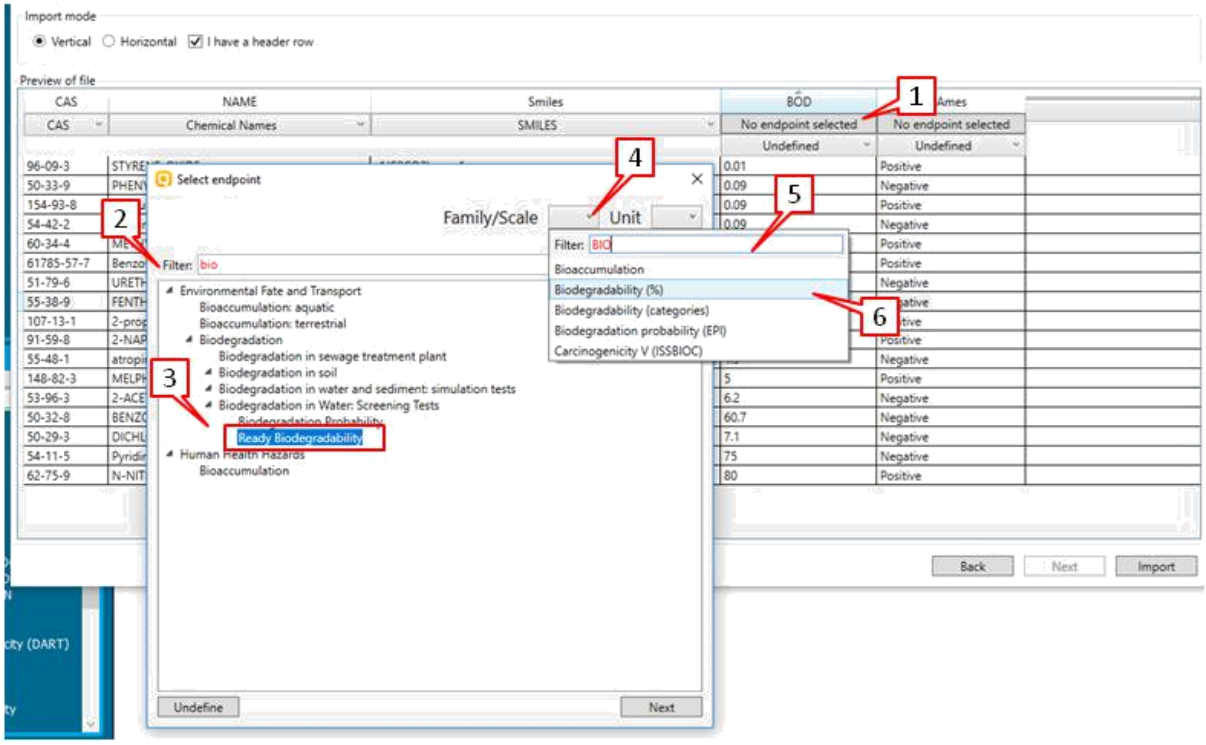


図 4

Unit ドロップダウンメニュー（1）から　%（Biodegradability（%））を選択し、Next（2）をクリックします（図 5）。

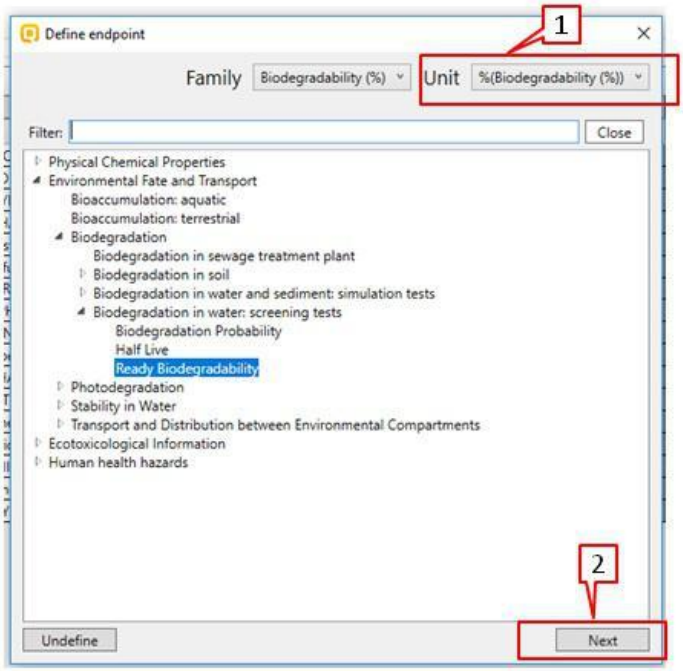


図 5

次のウィンドウで、ドロップダウンメニューから BOD エンドポイント（1,2）を選択します（図 6）。

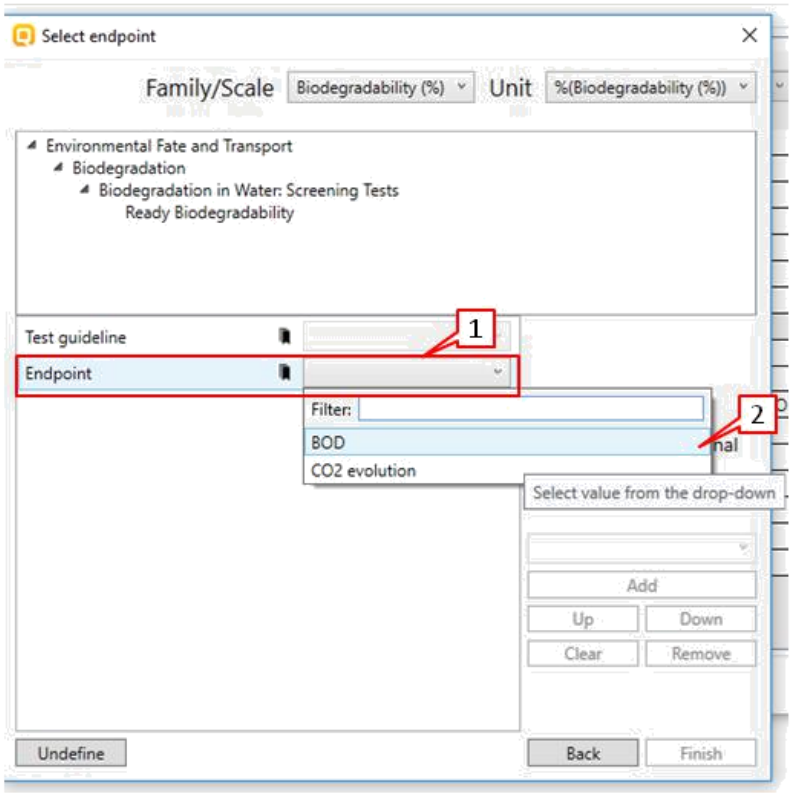


図 6

test guideline（テストガイドライン）(1) のドロップダウンメニューから、OECD 301C (2) を選択します。(図 7)

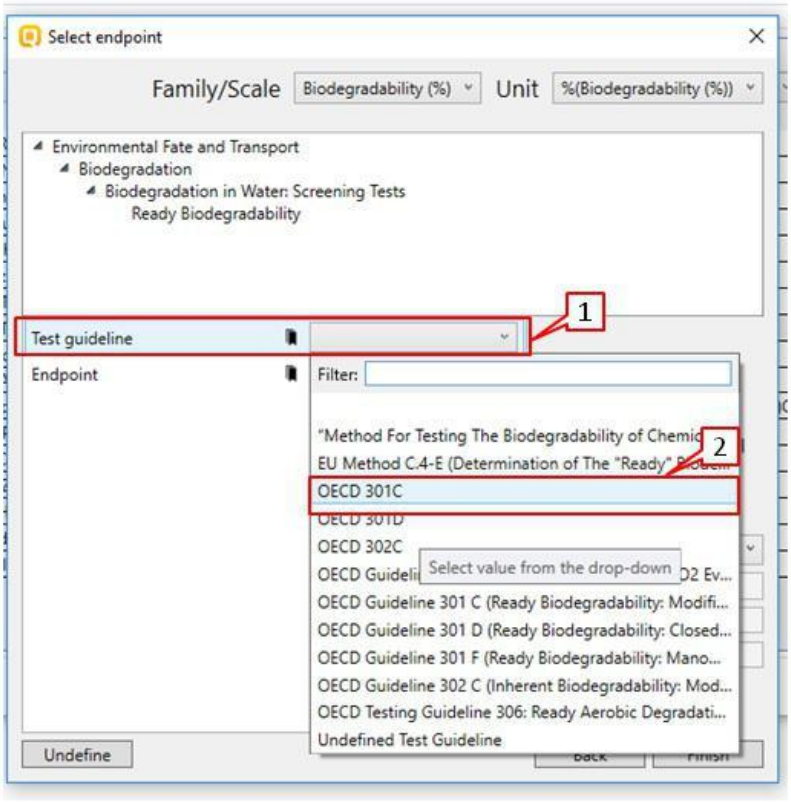


図 7

Finish（完了）(1) をクリックします。(図 8)

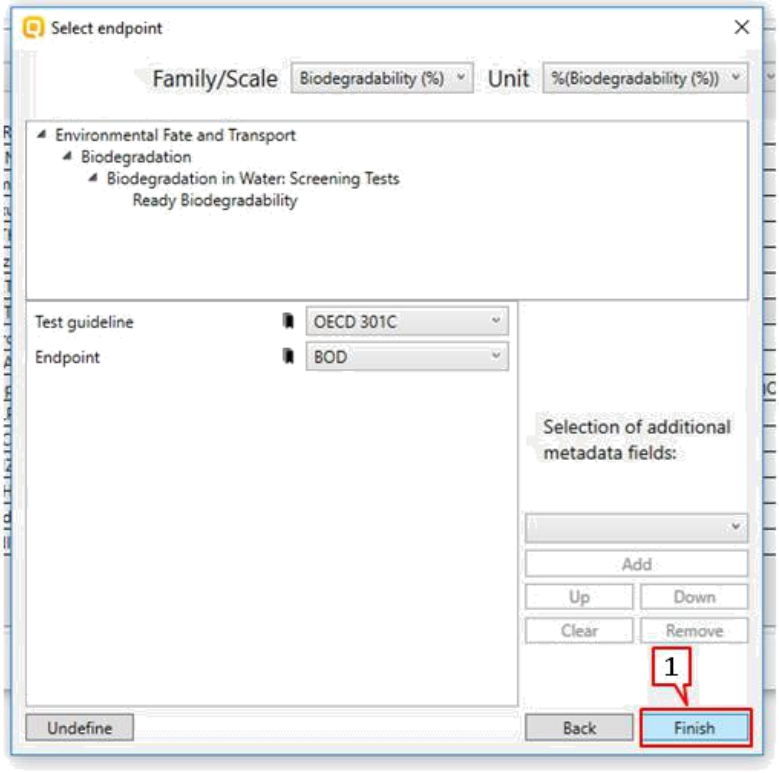


図 8

編集した入力欄がメインテーブル（1）に表示されます（図 9）。

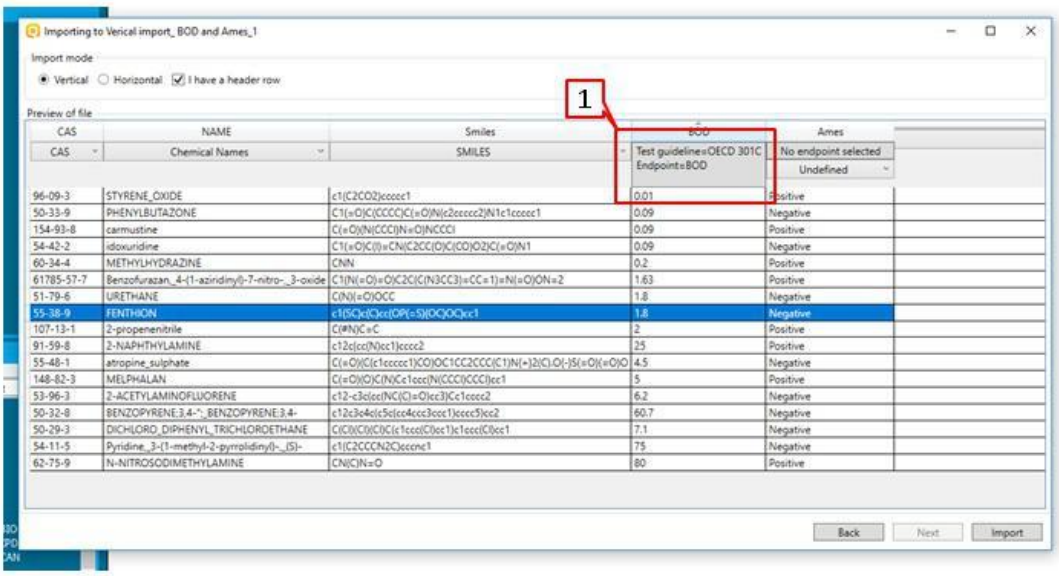


図 9

- **Ames**：（訳注：まず図 9 の Ames の下の No endpoint selected を押す）Select endpoint window が表示されます。フィルタ（1）を使用してエンドポイントツリーから Genetic Toxicity（2）を選択し、フィルタオプションを使用して family (Gene mutation I) （3）を選択し、Next（4）をクリックします。

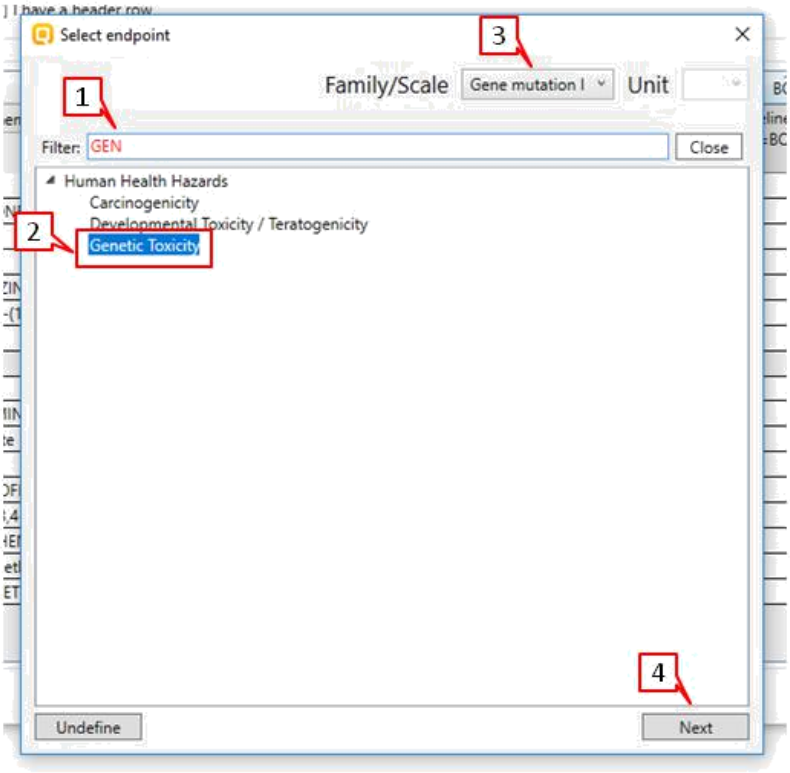


図 10

ドロップダウンメニュー（2）から Gene mutation エンドポイント（1）を選択します（図 11）。

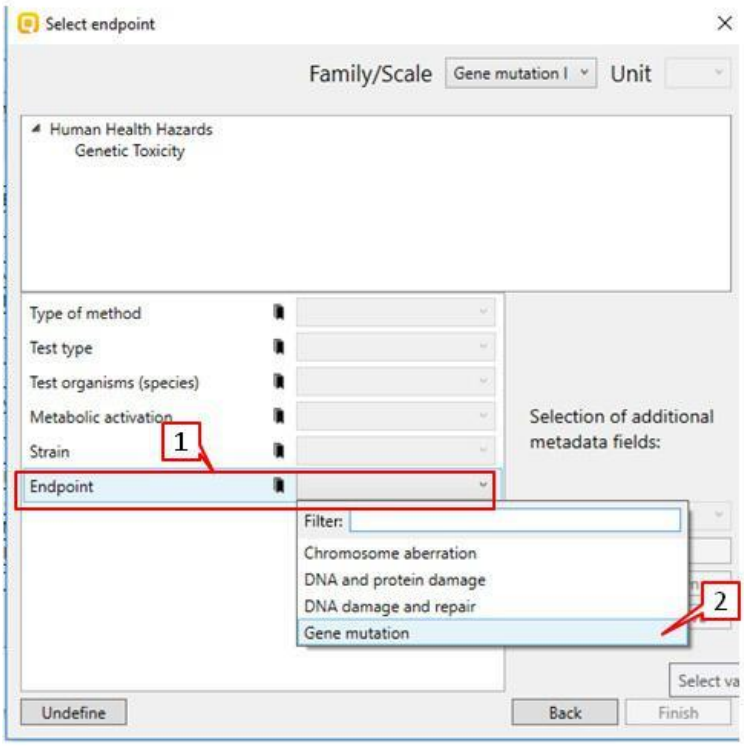


図 11

インポートされたデータのフォーマットに関連する場合は、追加の付随データ（1）を選択できます（図 12）。Finish（2）をクリックします。

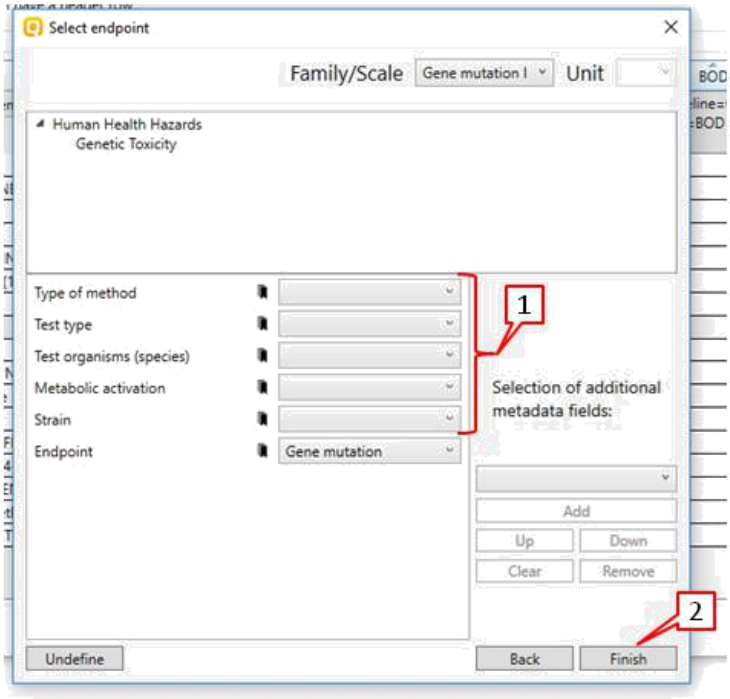


図 12

編集した入力欄（1）はメインテーブル（1）に表示されます（図 13）。Import (2)をクリックします。インポートが完了すると情報メッセージが表示されます。OK（3）をクリックします。

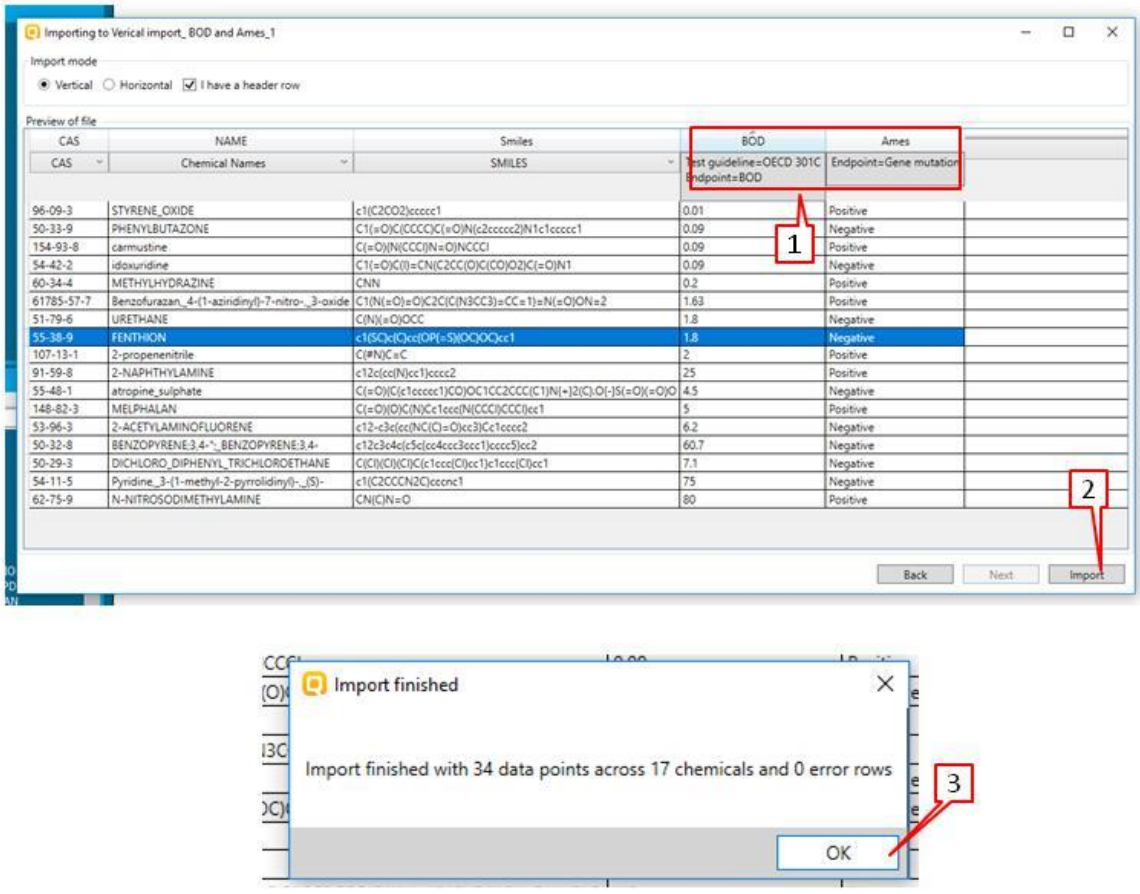


図 13

新しいデータベースが database list（1）に表示されます（図 14）。右クリックメニュー（2）は、データベース統計を表示したり、データベースを削除したりすることができるように導入されました。

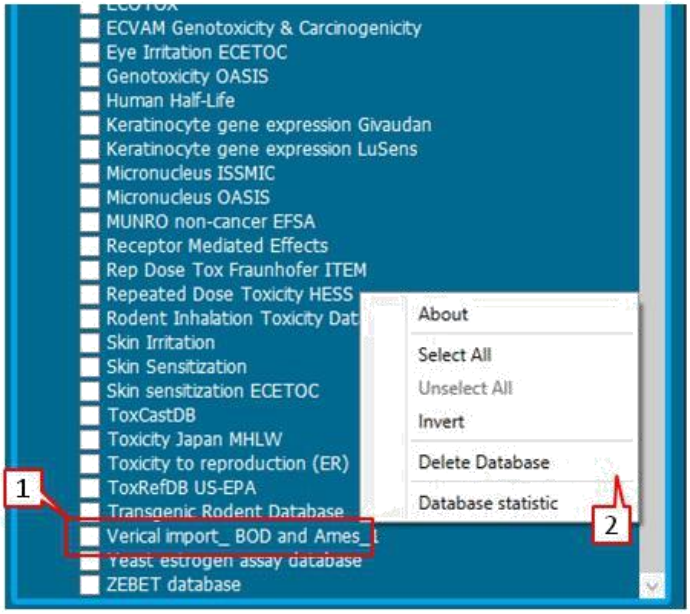

ECVAM Genotoxicity & Carcinogenicity
Eye Irritation ECETOC
Genotoxicity OASIS
Human Half-Life
Keratinocyte gene expression Givaudan
Keratinocyte gene expression LuSens
Micronucleus ISSMIC
Micronucleus OASIS
MUNRO non-cancer EFSA
Receptor Mediated Effects
Rep Dose Tox Fraunhofer ITEM
Repeated Dose Toxicity HESS
Rodent Inhalation Toxicity Dat
Skin Irritation
Skin Sensitization
Skin sensitization ECETOC
ToxCastDB
Toxicity Japan MHLW
Toxicity to reproduction (ER)
ToxRefDB US-EPA
Transgenic Rodent Database
Verical import_ BOD and Ames_t
Yeast estrogen assay database
ZEBET database
About
Select All
Unselect All
Invert
Delete Database
Database statistic
1
2

図 14

D.3.4.2. Horizontal import（横断的インポート）：データベース

ホリゾンタル（横断的な）インポートは、サポート情報（エンドポイント、試験条件、試験生物、著者など）など、さまざまなタイプの実験データを持つ一連の化学物質に適しています。

Import (1)をクリックします（図 1）。インポートウィンドウが表示されます。インポートされたファイルとは別の名前をそのデータベースに付ける場合は、Import title 入力欄（2）に入力します。すでに導入されたデータベース（3）にインポートすることもできます。Open file （4）をクリックします。

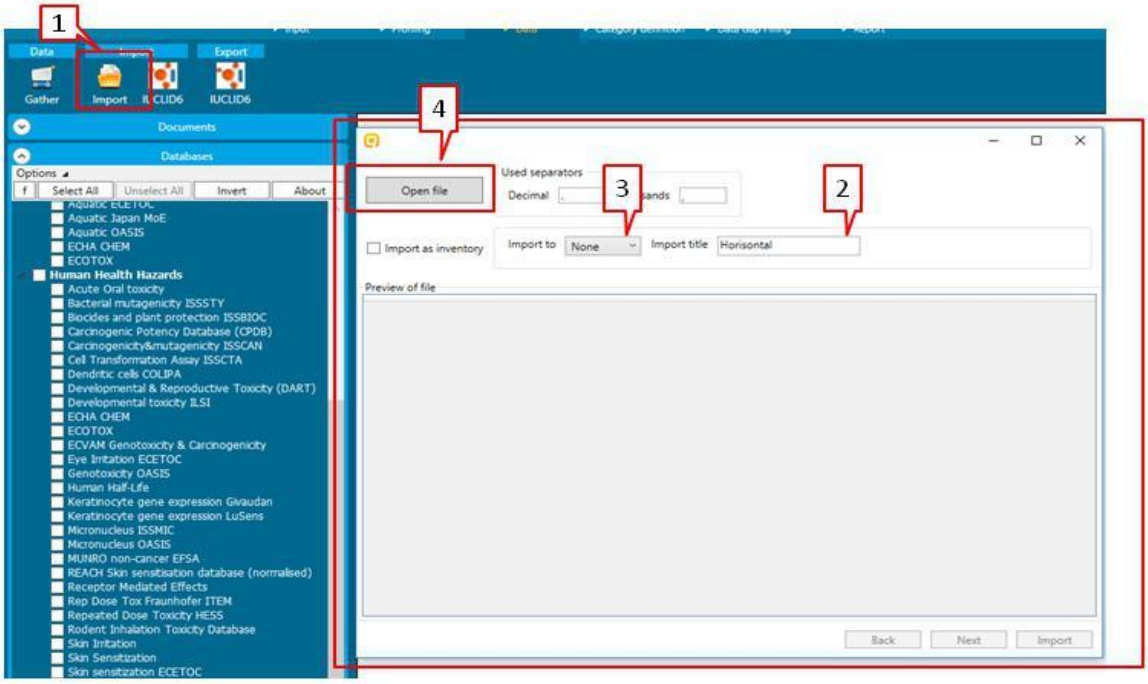


図 1

たとえば、Common files/QSAR Toolbox 4/Config/Examples（1）のサンプルファイルから、Horizontal import_Genotoxicity.xlsx データベース(2)を選択します。Open をクリックします。(図 2)

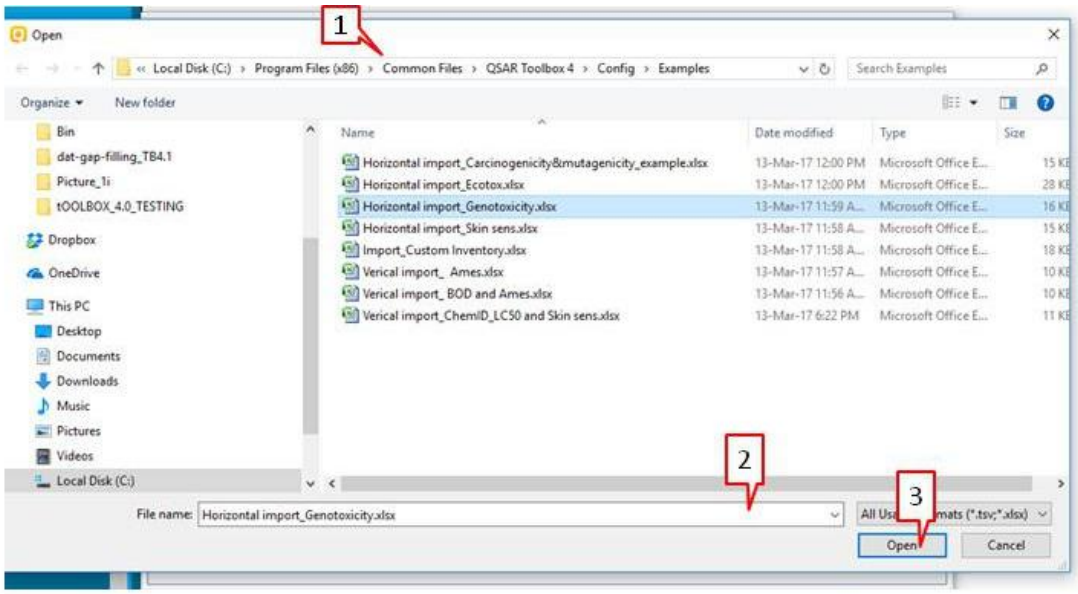


図 2

ファイルからの情報はタイトルの列（1）に示されています。Next (2)をクリックします。(図 3)

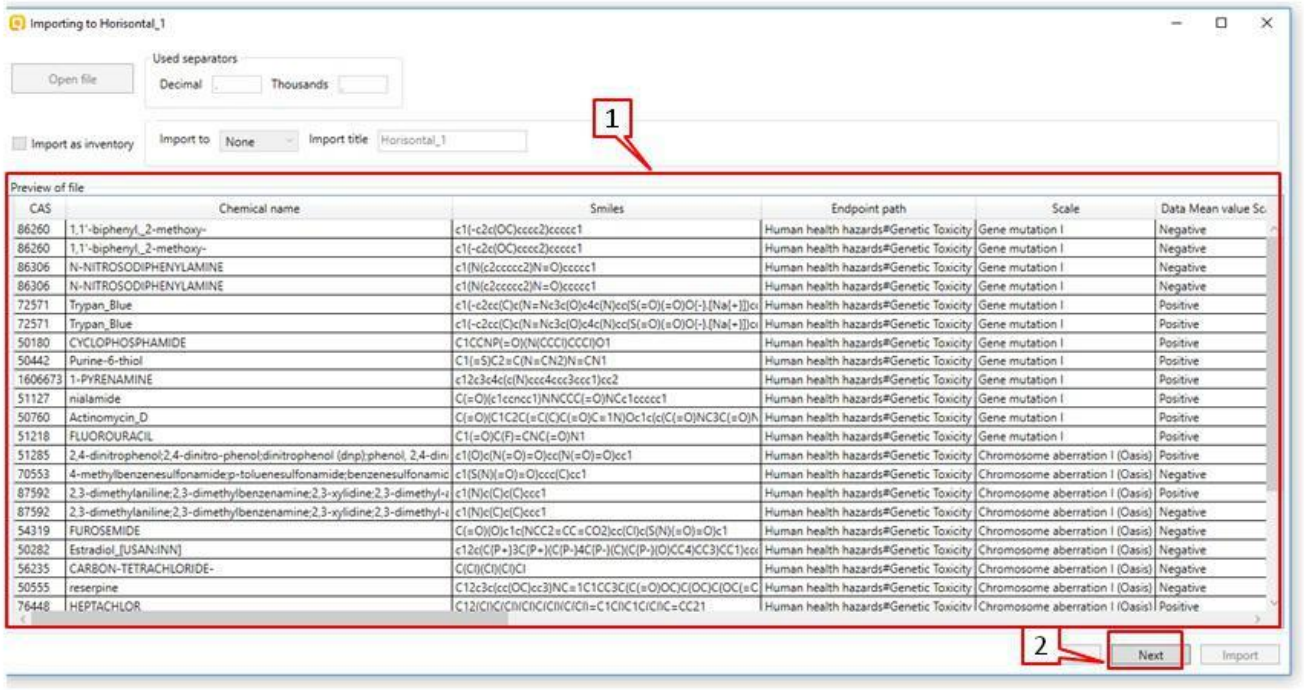


図 3

Horizontal ラジオボタン（1）をクリックします（図 4)。インポートされたファイルのある付随データが、Toolbox（2）の付随データフィールドにマップされていないと、表の上に赤色のメッセージが表示されます。付随データ名（3）の下のドロップダウンメニューから選択するという、付随データフィールドの手動マッピングが必要になります。

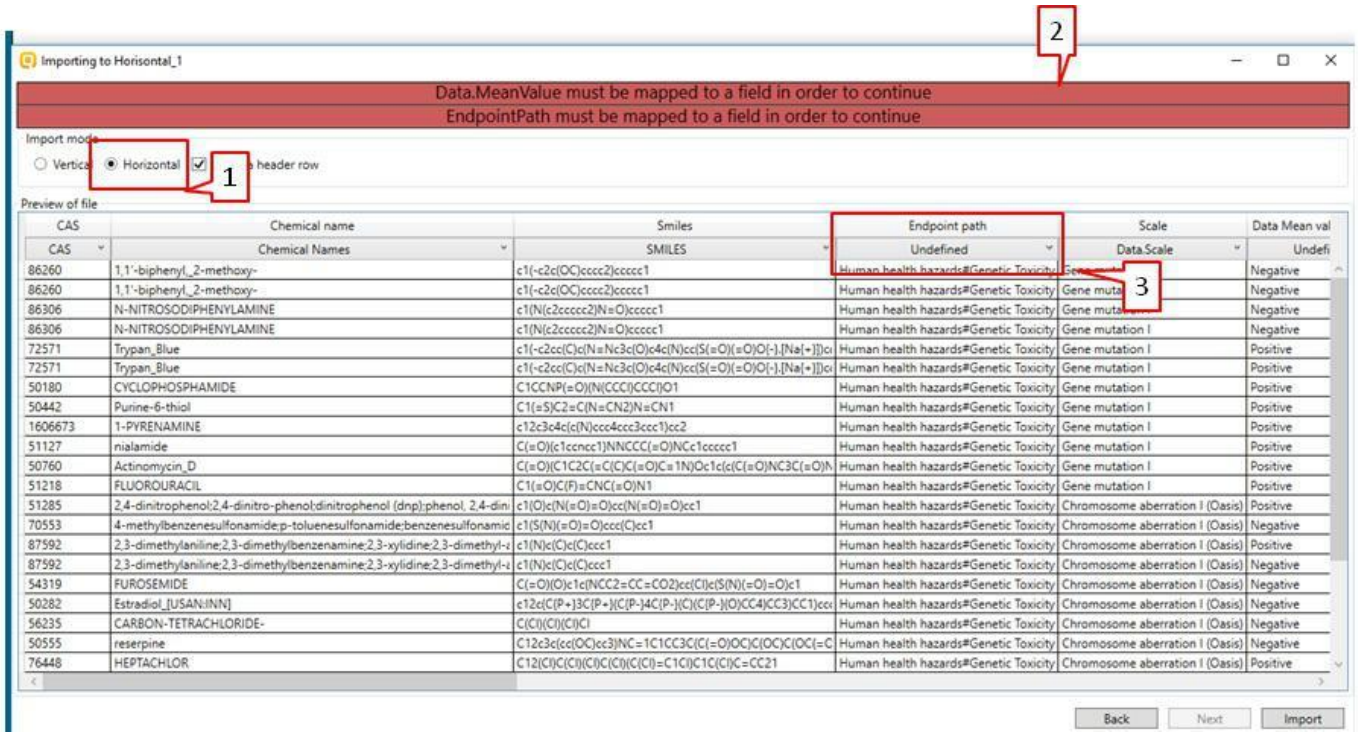


図 4

ドロップダウンメニュー（1）から、装備されているフィルタ（2）を使用してドロップダウンメニューで EndpointPath（3）を検索して選択します（図 5)。

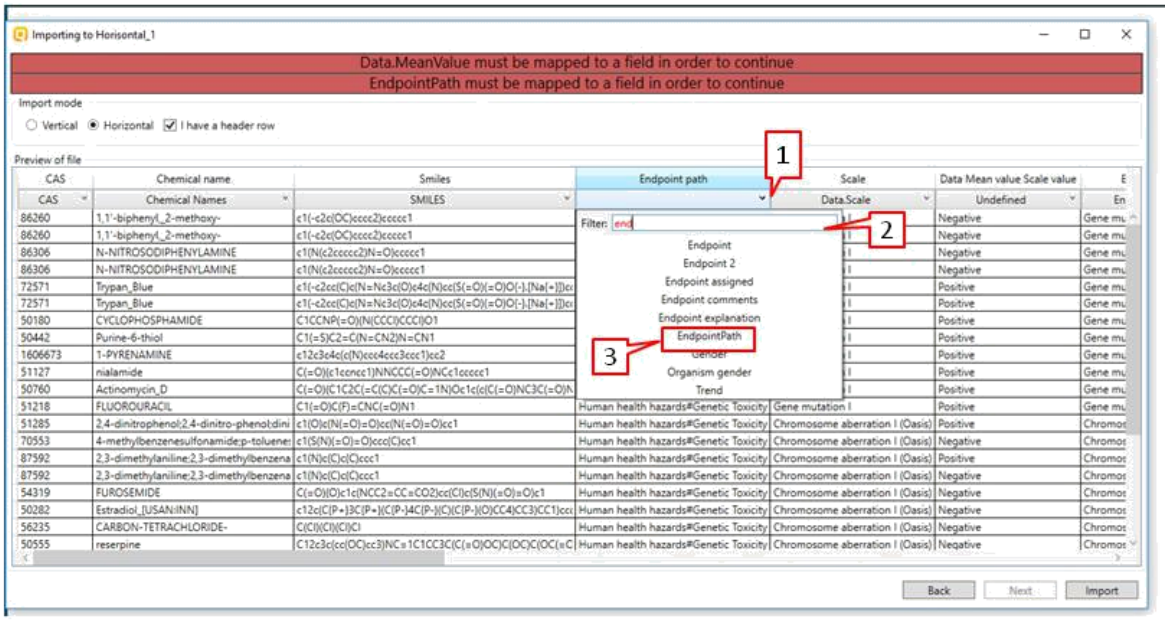


図 5

同様の方法で、 Data. Meanvalue （データ平均値）（図 6 の 1）をマップしてください。Import (2)をクリックします。

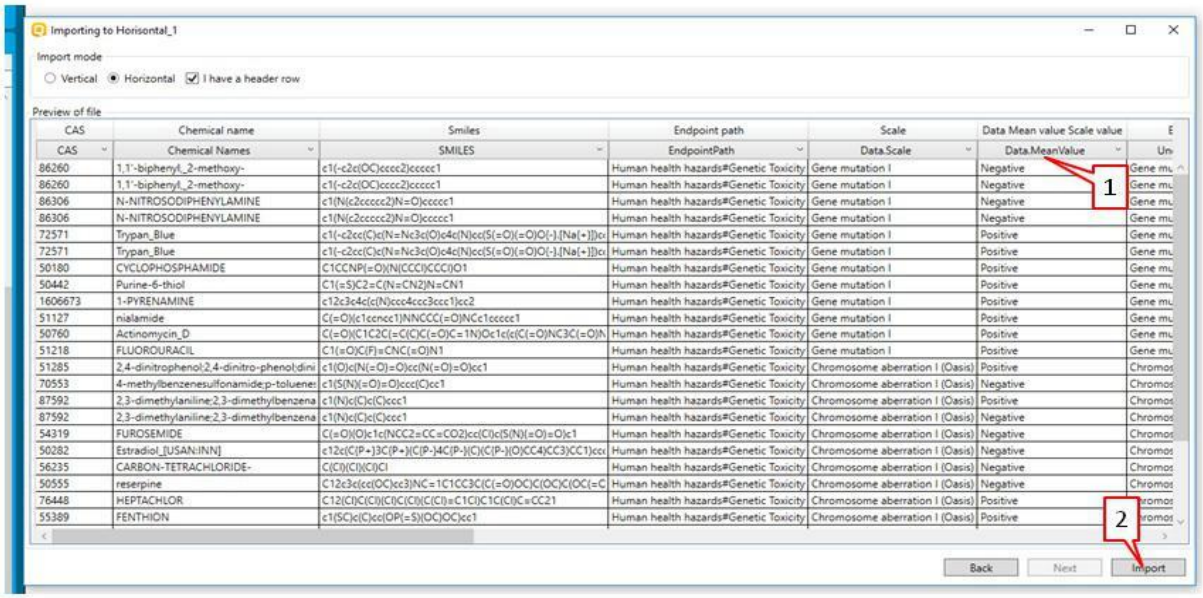


図 6

インポートが成功したことを示すメッセージが表示されます（図 7）。OK (1)をクリックします。

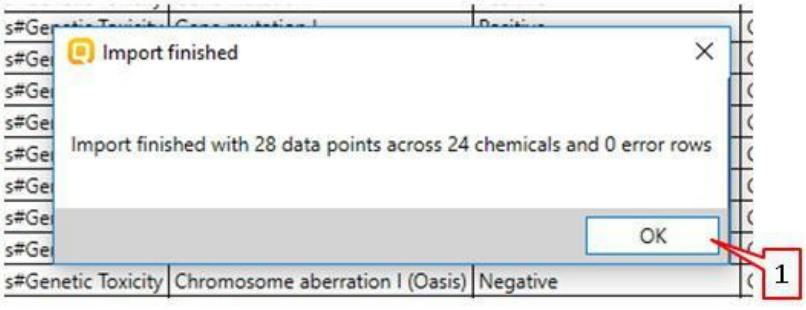


図 7

新しくインポートされたデータベースが、データベースのリストに追加されました（図 8）。右クリックメニュー（2）で、データベースの統計情報やデータベースの削除など、いくつかのオプションが提供されます。

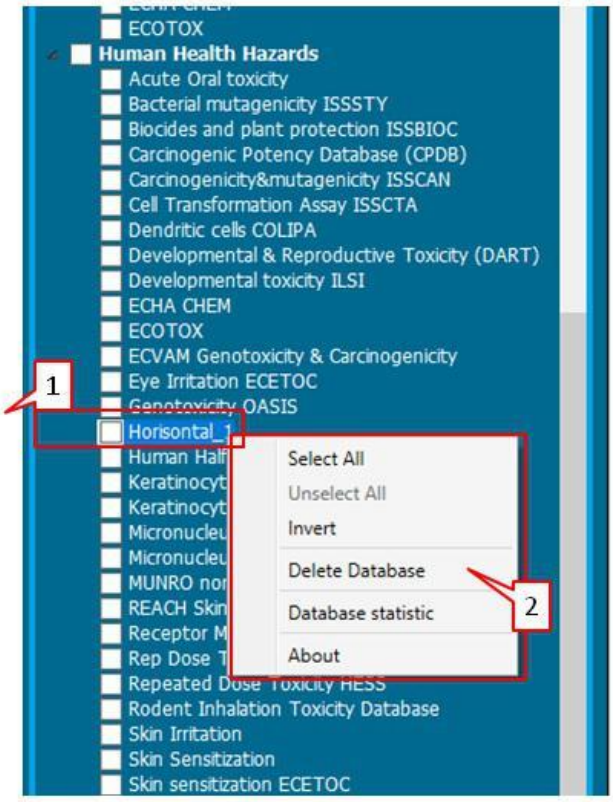


図 8

D.3.4.3. インベントリをインポートする

インベントリをインポートするには、以下の手順を実行してください。:
- Import （1）をクリックし、 Import as inventory チェックボックス（2）をオンにし、Open file（3）をクリックして、インベントリを選択します
（たとえば、 Common files/QSAR Toolbox 4/Config/Examples）（4）。そして Open（5）をクリックします（図 1）。

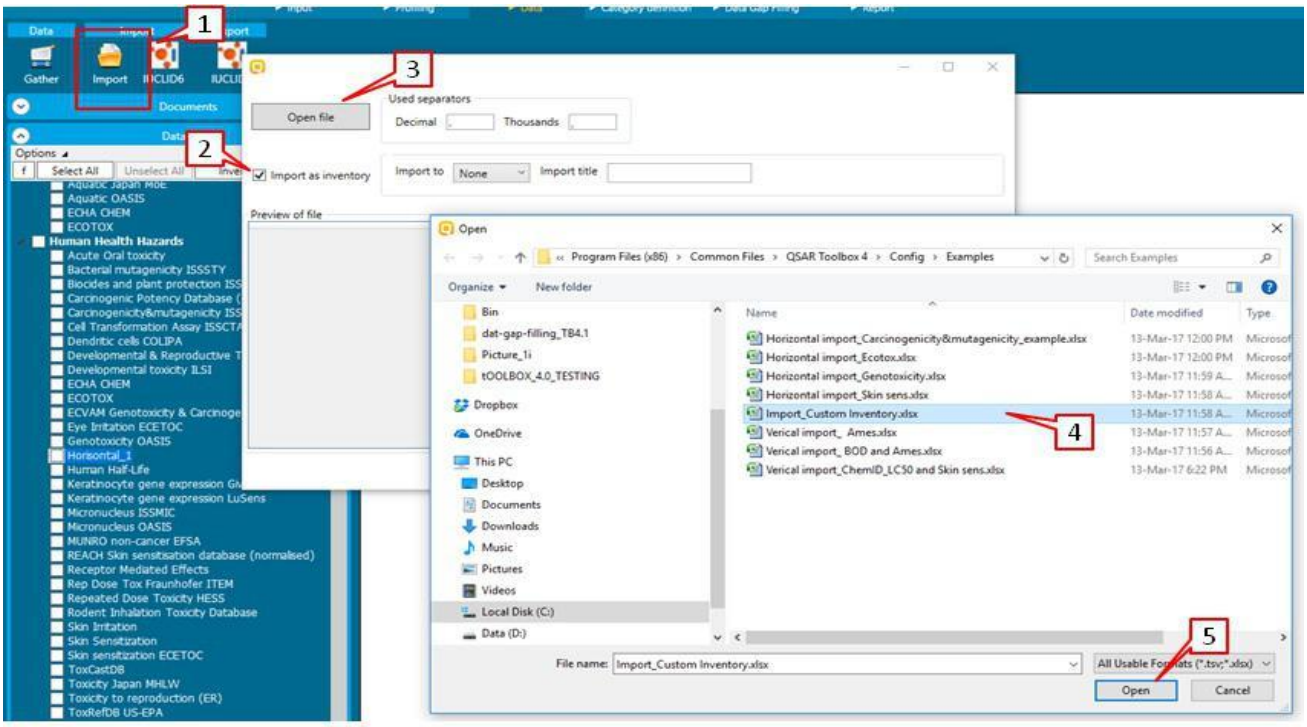


図 1

- インポート・ファイルの内容が（2）に示されているように、インポート・タイトルはインポート・ファイルと同名に数字（1）を追加した名前と同じです。Next (3)をクリックします。(図 2)



図 2

- インポートが成功したことを確認するメッセージで、 Import（1）をクリックし、次に OK（2）をクリックします。(図 3)

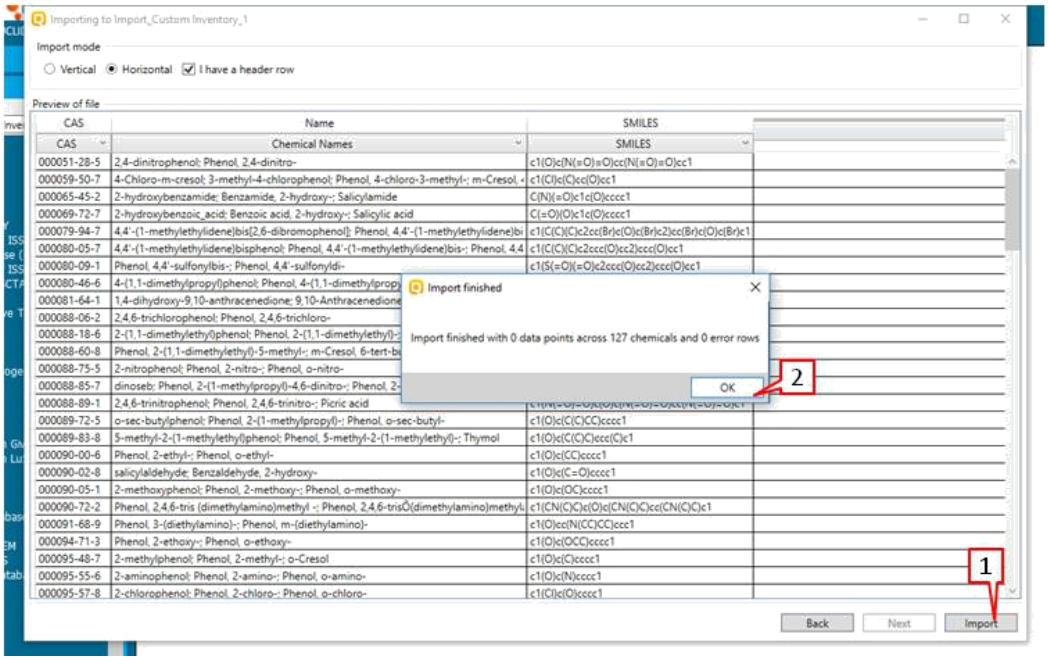


図 3

- 新しくインポートされたインベントリがインベントリ・リスト（1）に追加されました（図 4）。右クリックメニューには、削除オプションと inventory statistic（インベントリ統計情報）（2）が表示されます。

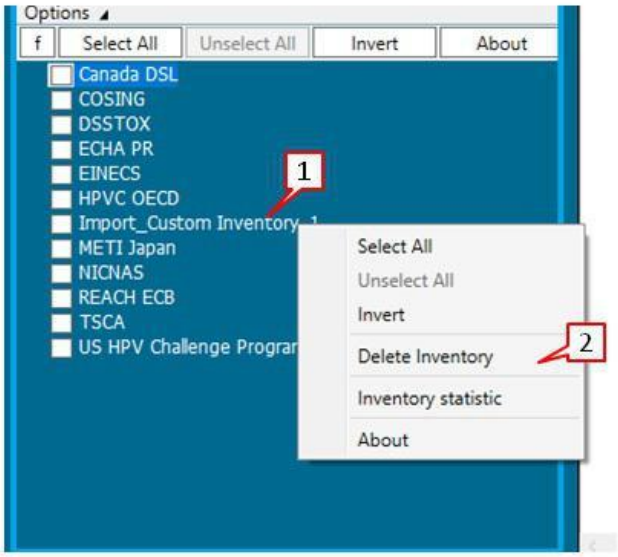


図 4

D.3.4.4. サポート情報 ワークフロー / データ / データをインポートする / D.3.4.4. サポート情報

最も重要なエンドポイントを以下の表に示します。この表の情報は、新しいデータベースのインポートをしやすくすることを目的としています。

Hazard effects	Endpoint(エンドポイント)	Unit family	Unit member (e.g mg/l)
Physical chemical properties	Boiling point（沸点）	Temperature	°C;°F; K
	Dissociation constant(pKa)(解離定数)	None	
	Melting/feezing point	Temperature	°C;°F; K
	Partition coefficient: logKow	None	
	Surface tension	None	mN/m
	Vapour pressure	Pressure	Pa; mmHg; Torr
	Water solubility	Mass concentration	mg/L
Environmental Fate and Transport	Bioaccumulation (e.g. BCF,BAF)	Bioaccumulation	L/kg bdwt
	Biodegradation (BOD)	Biodegradability (%)	%
	Biodegradation (BOD)	Biodegradability (categories)	Ready; Not ready
	Biodegradation (BOD)	Biodegradability(EPI)	Readily Degradable; Not Readily Degradable
Ecotoxicological information	Acute aquatic toxicity (LC50, EC50)	Mass concentration	mg/L; ppm
	Terrestrial Toxicity (EC50, NOEC)	Mass concentration	mg/L; ppm
		Volume concentration	gal/100gal
Human health hazards	Acute toxicity (LD50)	Mass fraction	mg/kg
	Gene mutation	Gene mutation I	positive
			equivocal
			negative
	DNA and protein damage		
	Chromosome aberration	Chromosome aberra on (Oasis)	positive
			negative
	Summary carcinogenicity	Carcinogenicity I (ISSCAN)	positive
			equivocal
			negative
	Repeated dose toxicity (LOEL, NOEL)	Administrated dose (mass)	mg/kg bdwt/d
	Skinsensitization(all type endpoints)	Skin sensitization II (ECETOC)	Positive,negative
	Skin sensitization (EC3)	Skin sensitization EC3(ratio)	%
		Skin sensitisation I (Oasis)	Strongly positive
			Weakly positive
			Negative
	Skin sensitization (SMWN)	Skin sensitization IV (GPMT)	Strong sensitizer
			Moderate sensitizer
			Weak sensitizer
			Non sensitizer
	Skin sensitization (LOEL, NOEL) Skin sensitization (ABC)	Skin sensization LOEL (ratio) Skin sensitization V (BfR)	µg/cm2 Category A, B, C
	Skin sensitization (SWAN)	Skin sensitization III (LJMU)	Strongly sensitizing
			Moderately sensitizing
			Ambiguous
			Not sensitizing

D.3.4.5. IUCLID6 からデータをインポートする

以下の手順で、IUCLID6 から新しく作成したデータベースにデータをインポートする方法を示します。
WebServices 経由で直接転送を開始するには、Data（1）を押し、 Import のアイコンの下にある IUCLID6（2）を選択する必要があります。(図 1)

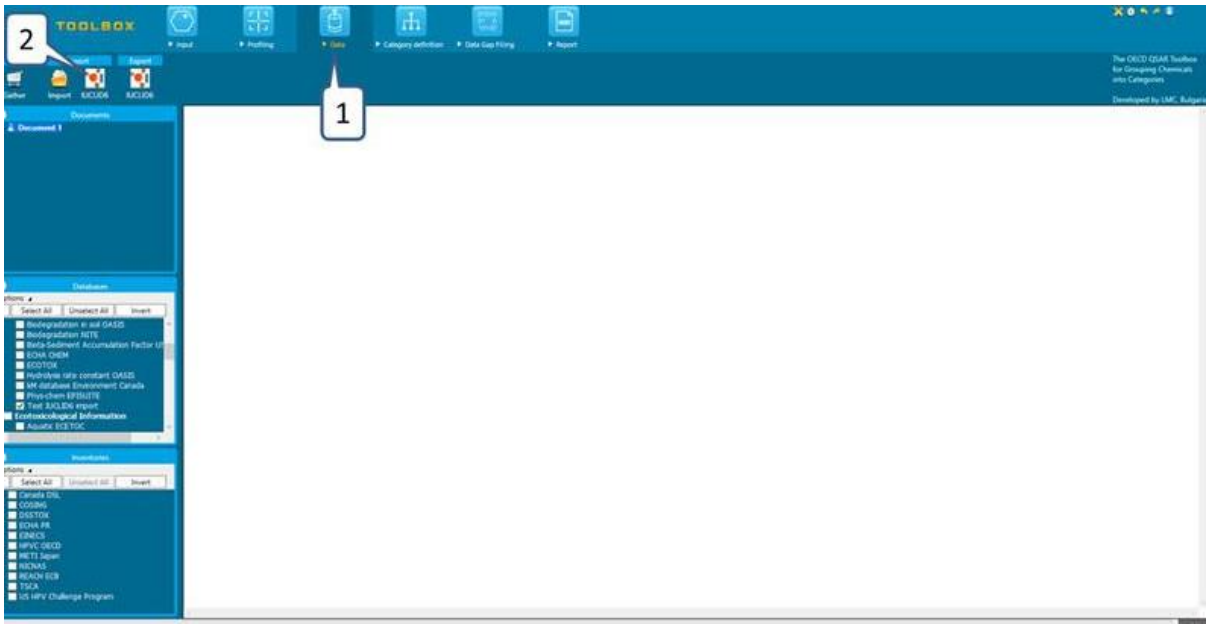


図 1

Toolbox の既存のデータベースにデータをインポートすることはできません。カスタムデータベースへのインポートだけが可能です。Choose destination ウィンドウで New database（1）を選択し、名前を付けます（2)。ソフトウェアが何らかの問題を示している場合、それらはログファイルに保存されます。インポート手順（3）中に起きたエラーまたはメッセージを調べるには、"Show log after import"（インポート後のログを表示する）(3）をチェックすることをお勧めします。Next (4)をクリックします。(図 2)

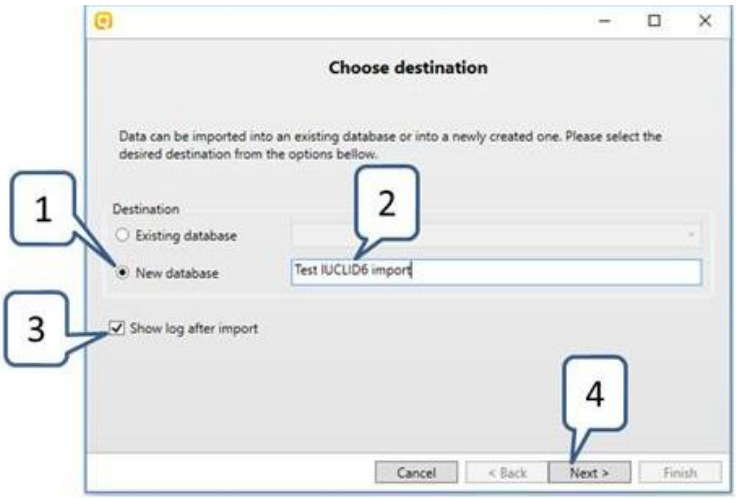


図 2

接続パラメータが Connect to an IUCLID 6 server ウィンドウ（図 3）に入力されます。

- WebServices Server：実行中の IUCLID 6 サーバの IP アドレスまたは DNS 名（1）。
- Port: the TCP ポート（通常 8080） (2)。
- I6 Username: サーバに有効な IUCLID 6 ユーザ名（3）。
- I6 Password: 該当するパスワード（4）。

上記のすべてのパラメータを入力したら、次へ（5）をクリックします。

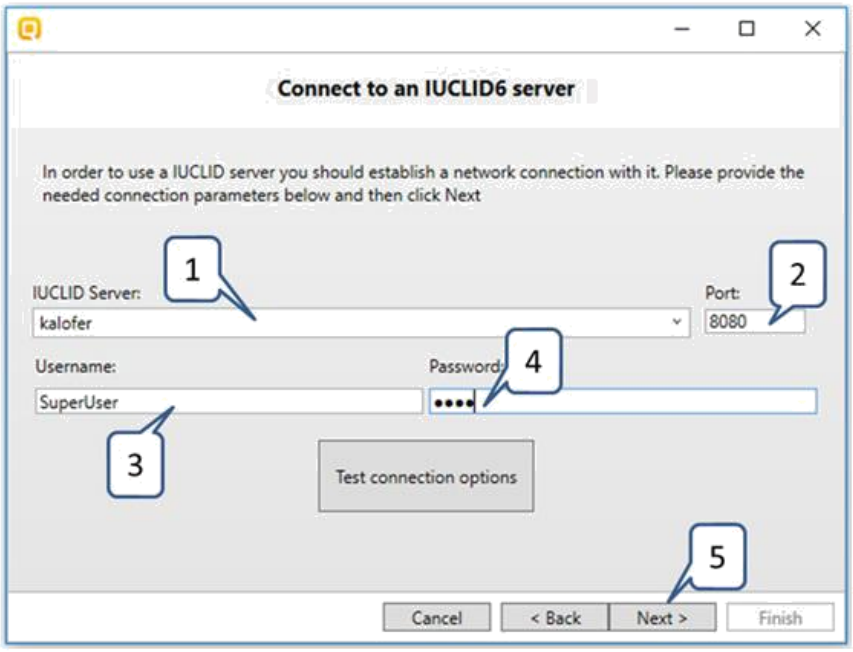


図 3

接続に成功すると、サーバ、ポート、およびユーザ名が保存され、後でリストからサーバを選択することで取得できます。

正しいパラメータが入力され、接続が確立されていると、次の画面が表示されます（図 4）。この段階では Get All Substances ボタンをクリックして、IUCLID 6 データベースの該当するすべての物質を含むリストを取得できます。

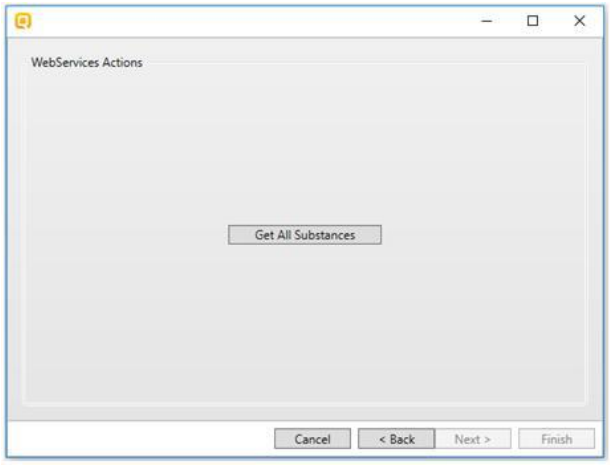


図 4

ウィザードの次のページにすべての物質のリストが表示されます。それらから 1 つを選択してください（1)。そのあと Finish（2）をクリックします。(図 5)

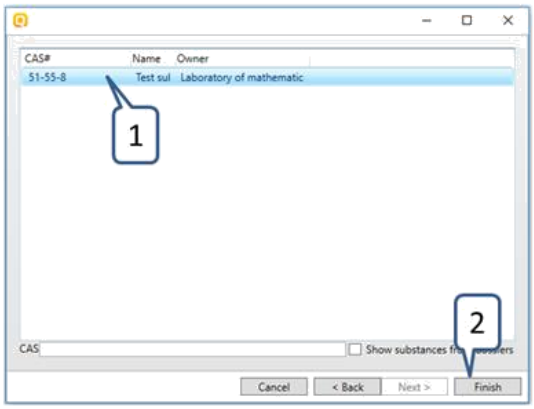


図 5

実際にインポートプロセスが開始されます。転送を実行するのに必要な時間は、接続速度とコンピュータのパフォーマンスによって異なる場合があります。進捗状況は緑色のバーで示されます。(図 6)
選択した物質が転送されると、“Import completed!”（インポートが完了！）というメッセージが表示され、OK をクリックするとウィザードが終了します。

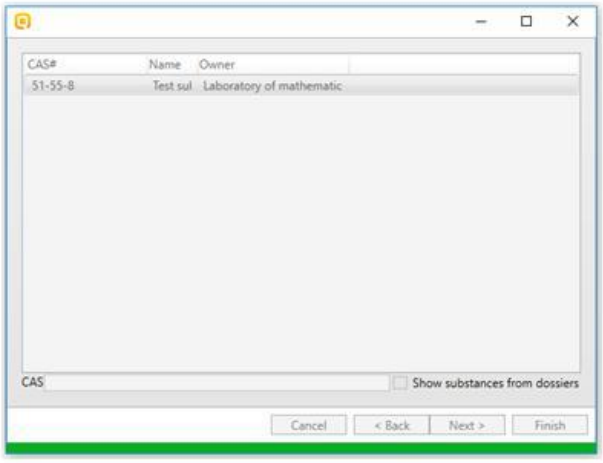


図 6

インポートウィザードを終了すると、インポートされた物質を含む新しく作成されたデータベースを読み込んで表示することができます。Input モジュールに行き、 Database をクリックして下さい（図 7)。
“Select database” ウィンドウが表示されます。インポートされたデータベースを選択してください。
データベース（1）を選択して OK（2）をクリックすると、**データ・マトリックス**上のデータベースから化学構造が読み込まれます。(図 7)

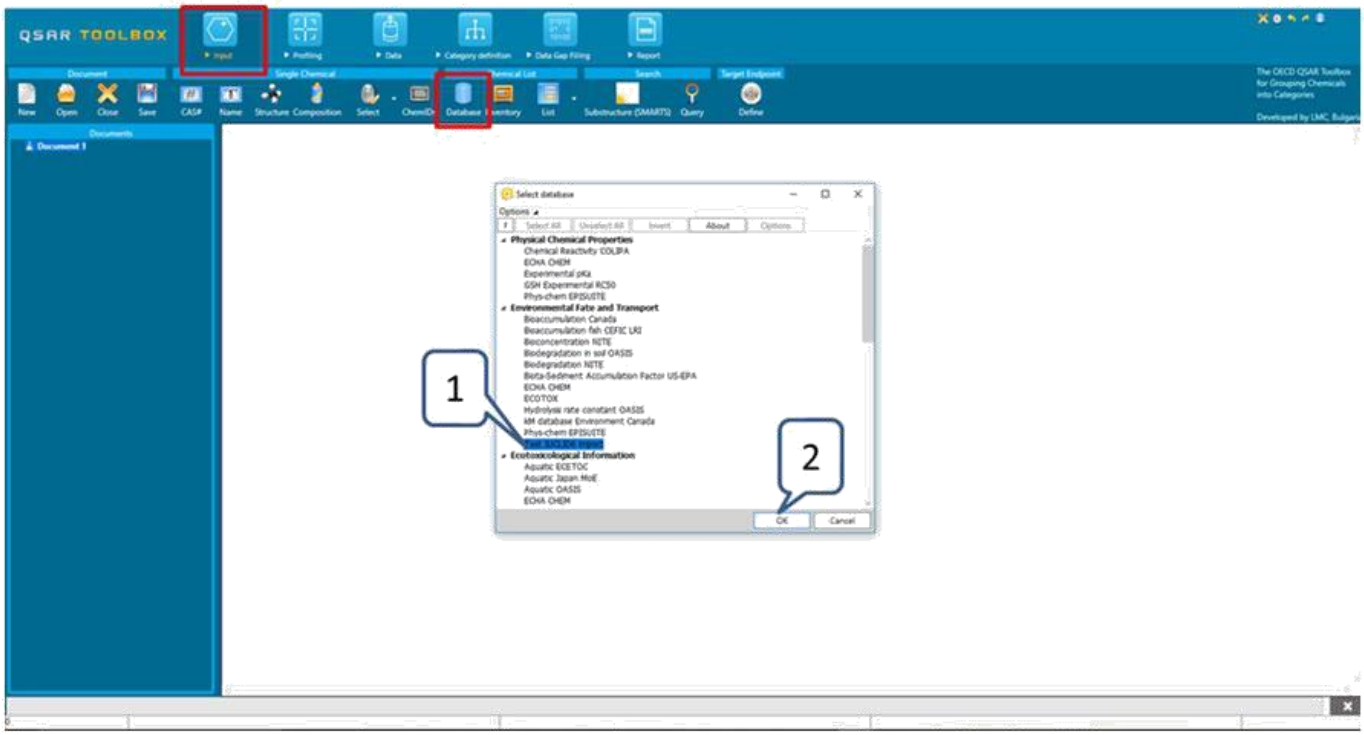


図 7

エンドポイントデータを読み込むには、 Data タブ（1）をクリックし、データベース（2）を選択してから Gather （3）ボタンをクリックしてデータ収集を開始してください。(図 8)

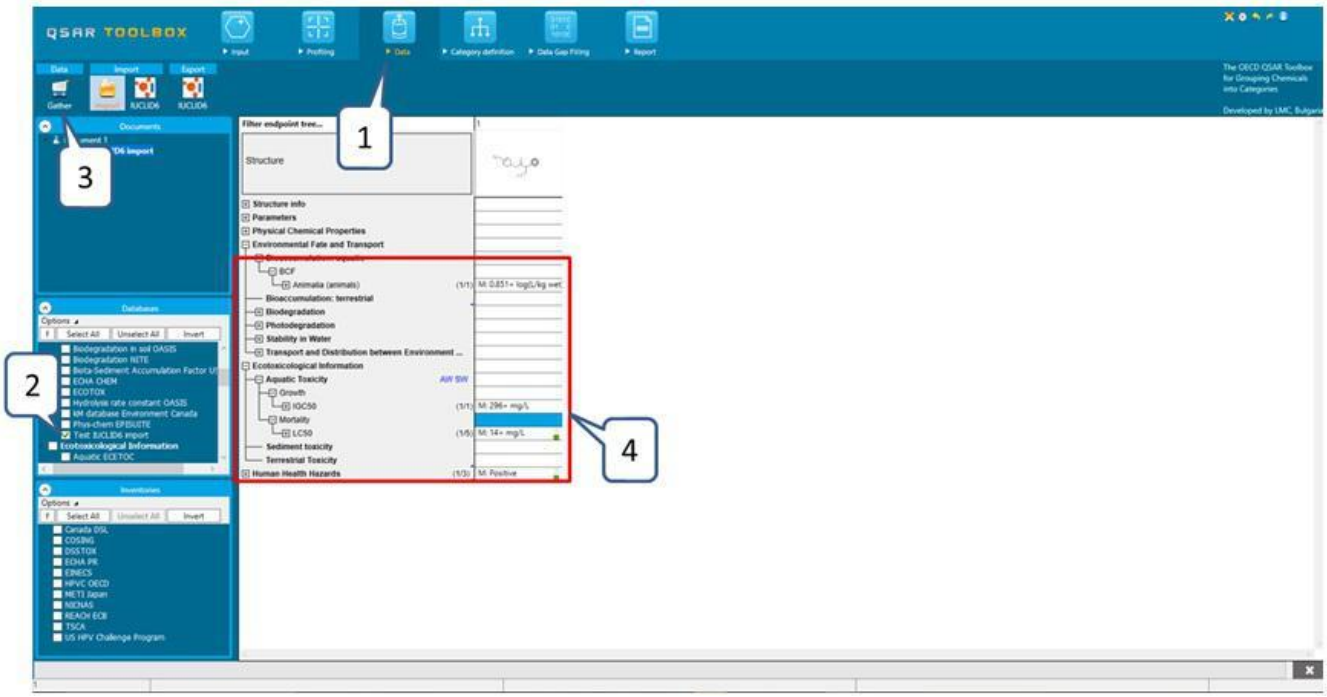


図 8

D.3.5. データをエクスポートする

2 種類のエクスポートがあります。:

- csv file を使う
- IUCLID 6

D.3.5.1. csv ファイルにデータをエクスポートする

実験データを集めたら（Collecting data のセクションを参照）、csv ファイルとしてエクスポートできます。

次の手順に従ってください：
（訳注：Input で適当な物質を入力し、Data を押して左の Databases から項目を選択し、左上の Gather アイコンを押してデータを集めておく）
関連するデータを含むセルを選択します。たとえば Ecotoxicological information（生態毒性情報）です（図 1）。灰色の部分（1）のタイトル脇をクリックし（右クリック）、Export data matrix（2）をクリックします。関連するノードが選択されている所の Matrix report が開きます（3）。ただし、他の情報を含めたい場合は、選択を変更することができます（前もってデータを集めていた場合）。選択ボタンもあります（4）。最後に Export（5）をクリックします。

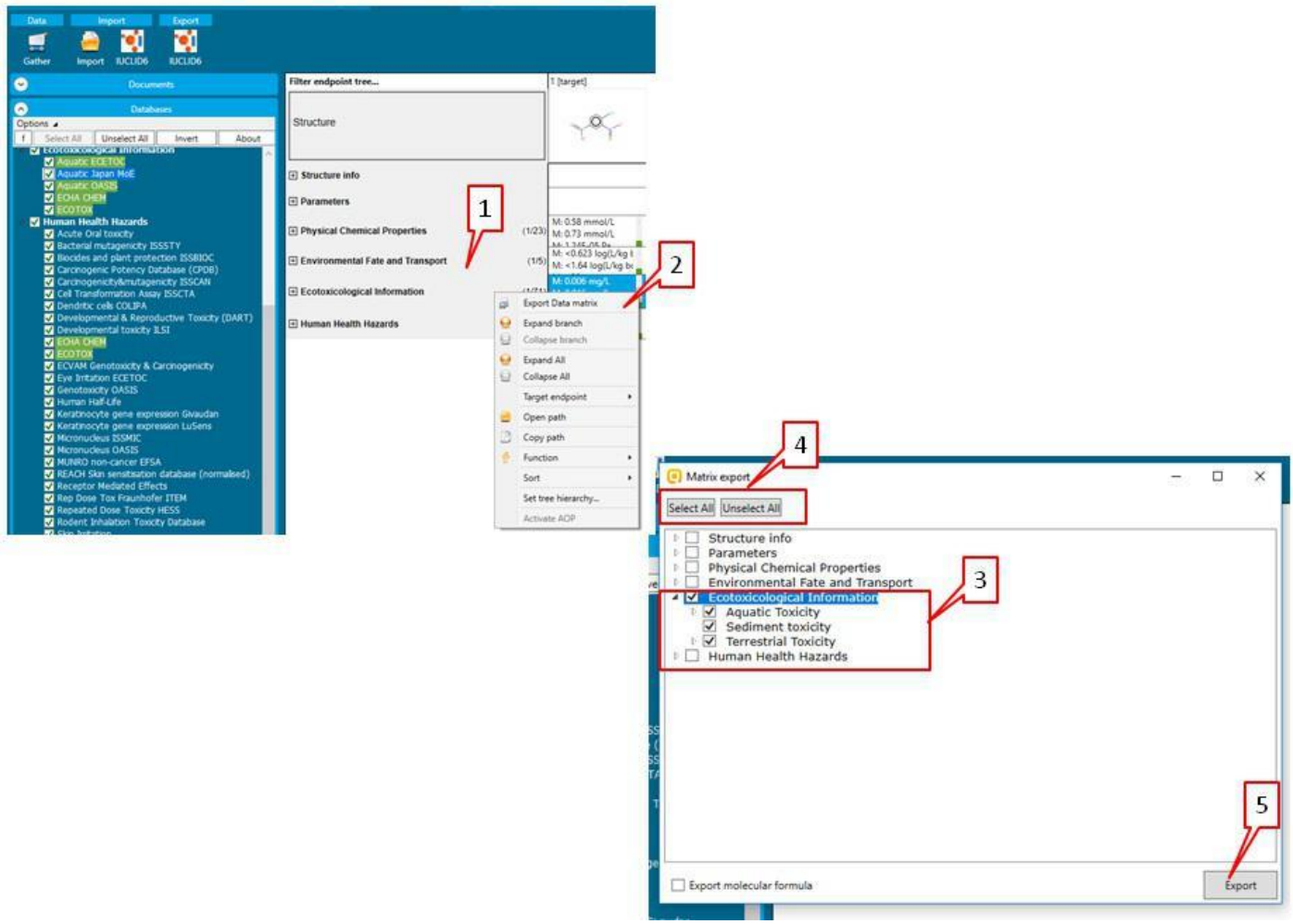


図 1

ファイルの場所（1）を選択し、ファイルの名前を入力して Save（3）をクリックします（図 2）。csv ファイルは Excel（4）で開くことができます。

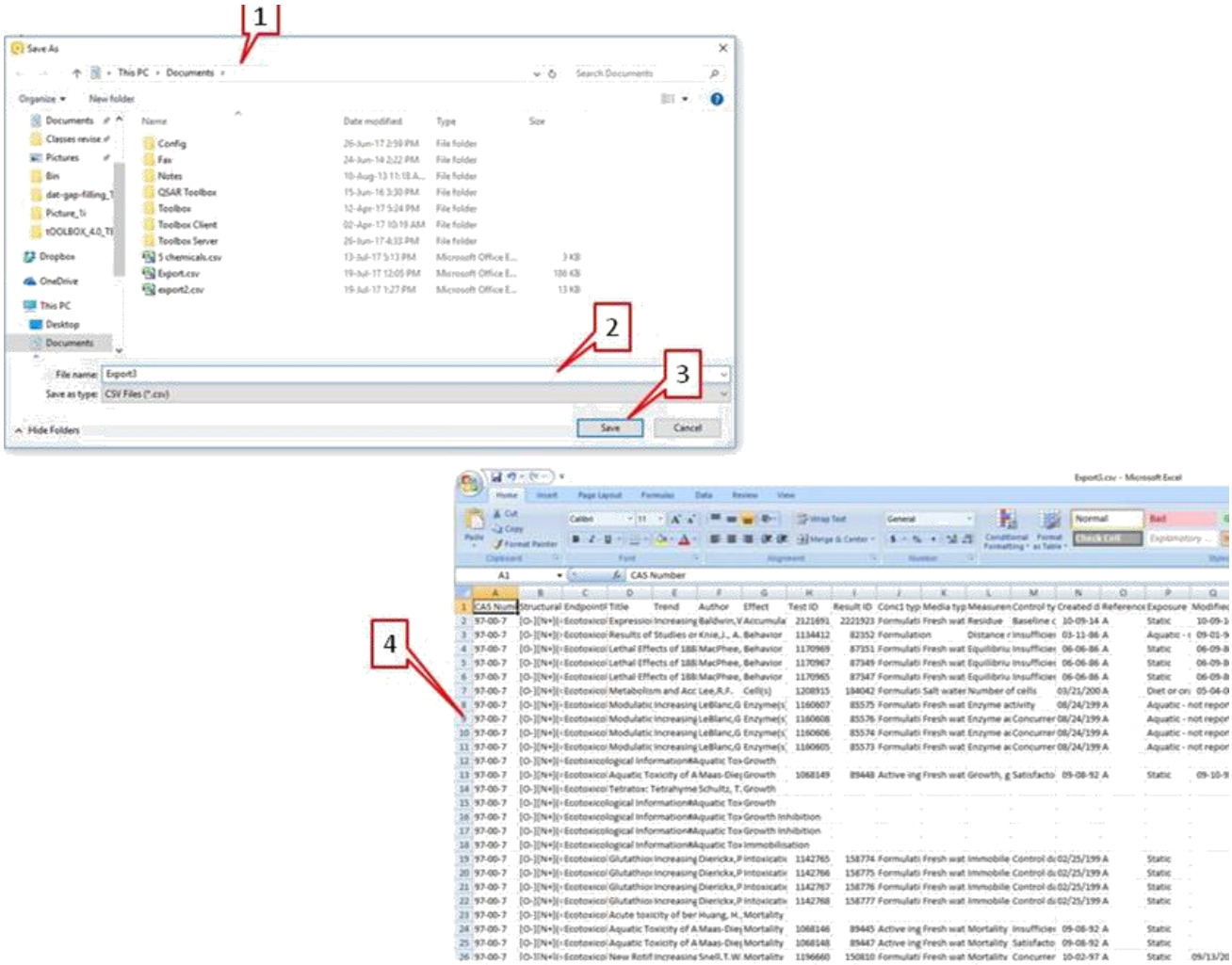

1
2
3
4
Save As
This PC
Documents
Search Documents
Organize
New folder
Name
Date modified
Type
Size
Config
Fas
Notes
QSAR Toolbox
Toolbox
Toolbox Client
Toolbox Server
5 chemicals.csv
Export.csv
export2.csv
File name: Export3
Save as type: CSV Files (*.csv)
Hide Folders
Save
Cancel
Export3.csv - Microsoft Excel
CAS Number

図 2

D.3.5.2. IUCLID6 にデータをエクスポートする

ツールボックスの予測値を選択して、IUCLID 6 データベースの物質に直接割り振ることができます。

QSAR Toolbox を使用すると、(Data Gap Filling ツールを使用して) 予測データを IUCLID 6 にエクスポートすることができます。現時点ではオンラインで、IUCLID 6 サーバーに直接接続し (WebServices 経由)、予測されたエンドポイントデータを選択した物質に割り振ることができます。

データを IUCLID にエクスポートするには、Data (1) をクリックし、予測値のセル (2) を選択し、エクスポート用の IUCLID6 ボタン (3) (図 1) をクリックします。

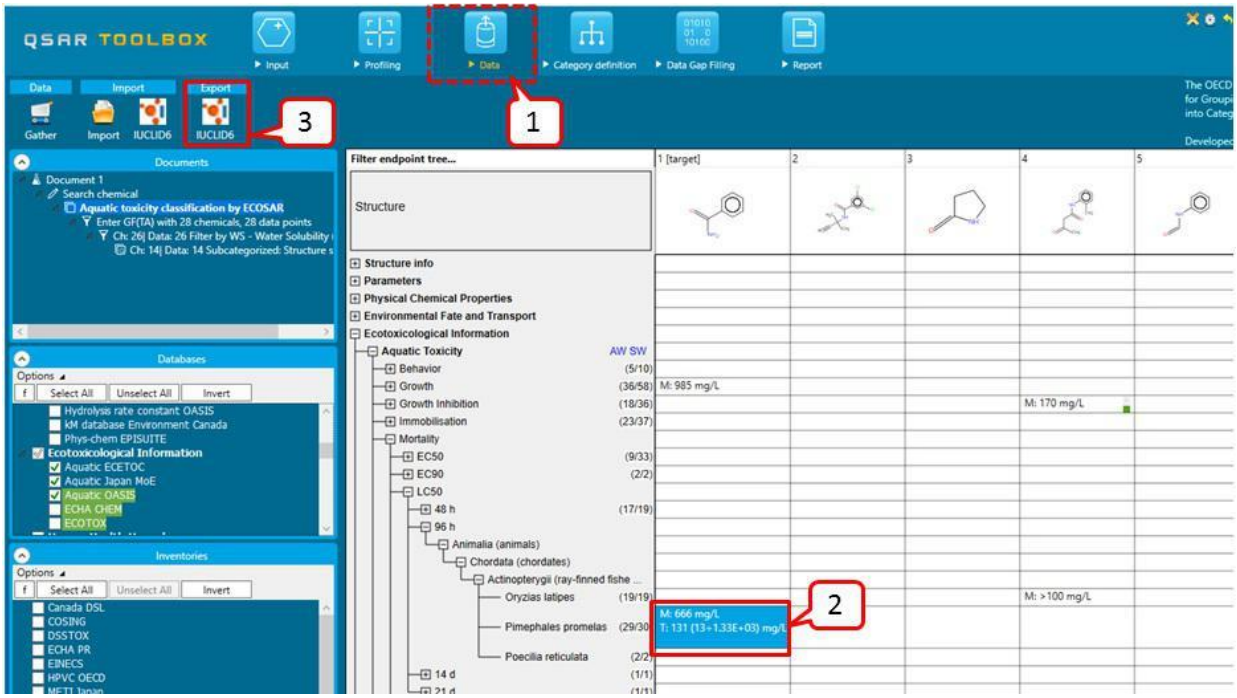


図 1

データを準備する

ウィザードの最初のページで、エクスポートする必要がある prediction (予測) を選択してください (1)。このページでは、レポート情報の編集もできます (2)。(図 2)

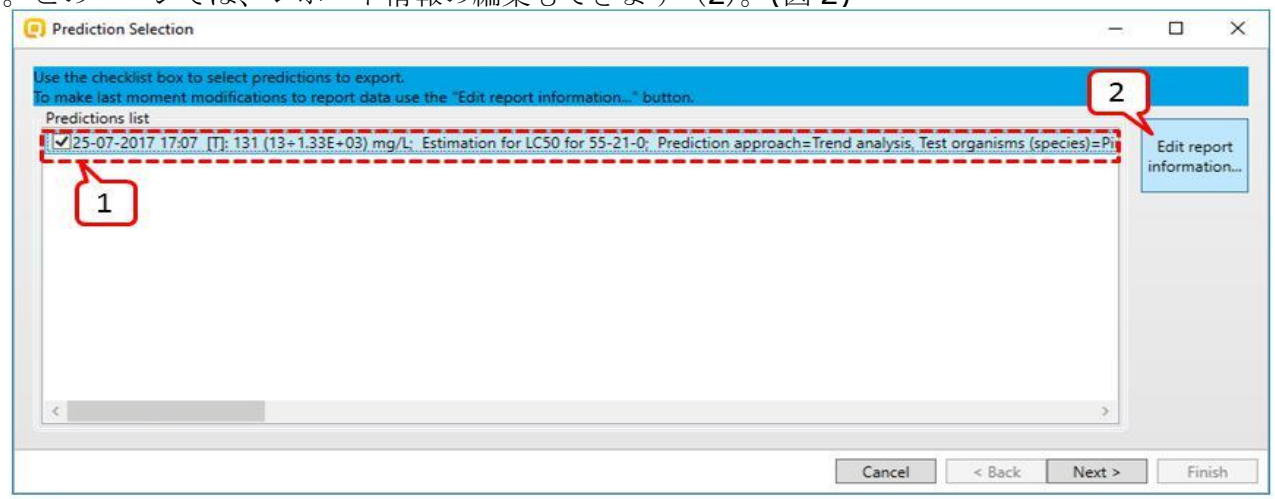


図 2

Edit report information （レポート情報の編集）ボタン（図 2-2）をクリックすると、レポートをカスタマイズするための新しいウィンドウが開きます（1）（図 3）。次に、レポートの利用可能な入力欄に追加情報を編集および/または追加することができます。

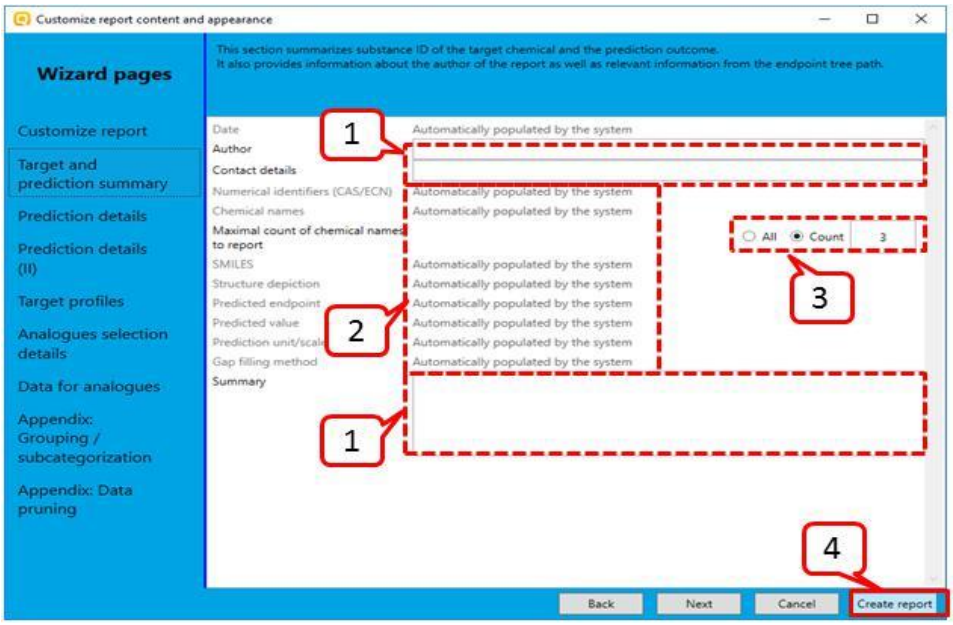


図 3

手動でいくつかの入力欄に情報を追加できます（1）。また、自動的に入力されて変更することができない情報があります（2）。標的物質は多くの別名を持つ可能性があるため、レポートする化学物質名の最大数に関する追加オプションも作成されます。All ラジオボタンをクリックするか、または Count ラジオボタン（3）を選択して、化学物質名の正確な数を指し示すように、レポート内の標的物質の利用可能な化学物質名をすべて表示するかどうかを選択できます。各入力欄に情報を追加した後、Create report（レポート作成）ボタン（4）をクリックしてください。変更は生成されたレポートに保存されます。

レポート入力欄の additions/correction（追加/修正）を指定したら、 Next （1）をクリックしてください。(図 4)

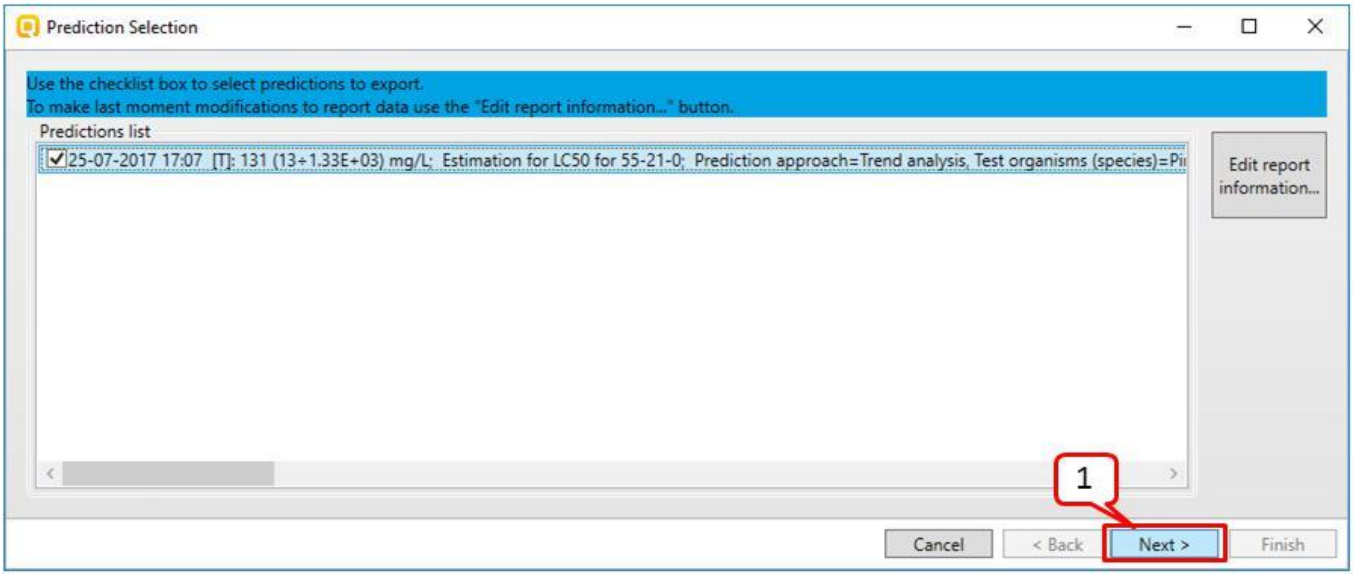


図 4

次のウィンドウでは、選択された各予測（1）に、リストから（2）該当する Harmonized Template（ふさわしいテンプレート）（図 5）を合わせてください。テンプレートを選択すると、予測が緑色で強調表示されます。すべての予測に対応するテンプレートが割り当てられるとすぐに、 Next ボタン（3）が有効になります。インポートウィザードのこのステップで、エクスポート準備が整ったかどうか入力欄を確認できます（4）。

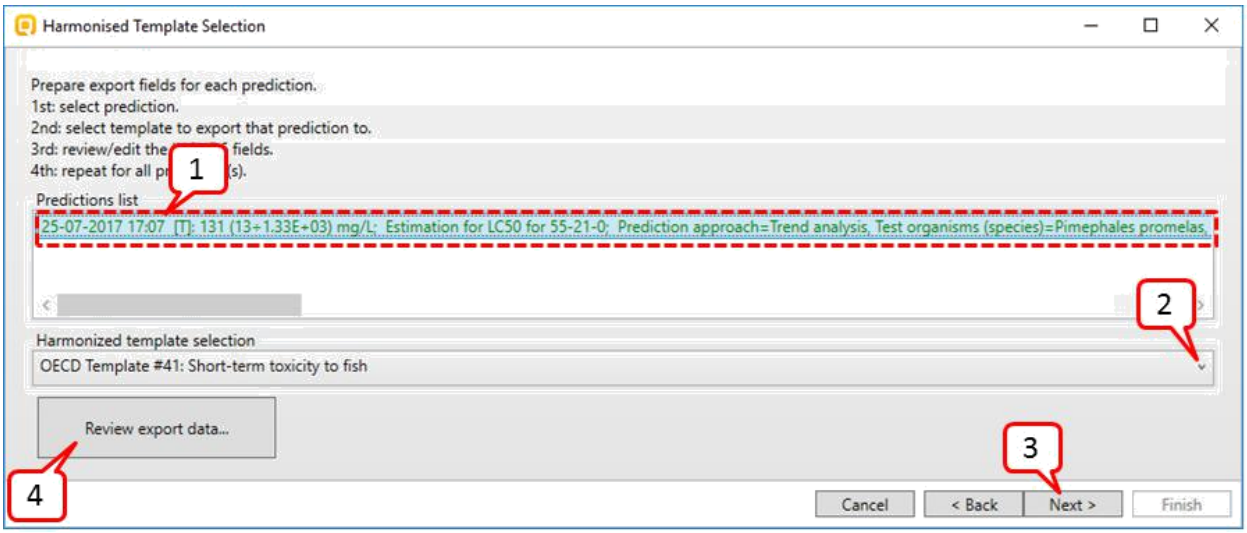


図 5

"Review export data"（エクスポートするデータを確認する）ボタンをクリックすると、IUCLID6 export ...ウィンドウが表示され、評価者（1）に必須と印をつけられている入力欄をレビュー/編集し、他のすべての入力欄もレビュー/編集できます。編集が終了したら OK（2）をクリックします（図 6）。

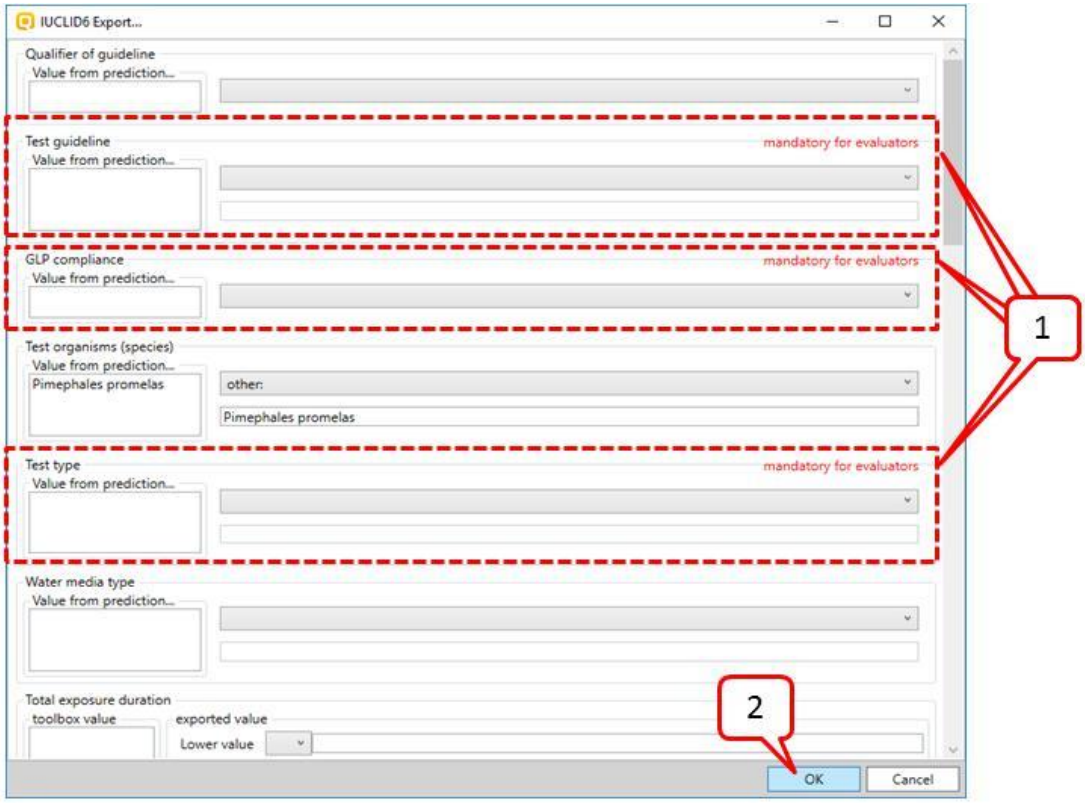


図 6

インポートを続行するには、 Next をクリックします（1）（図 7）。

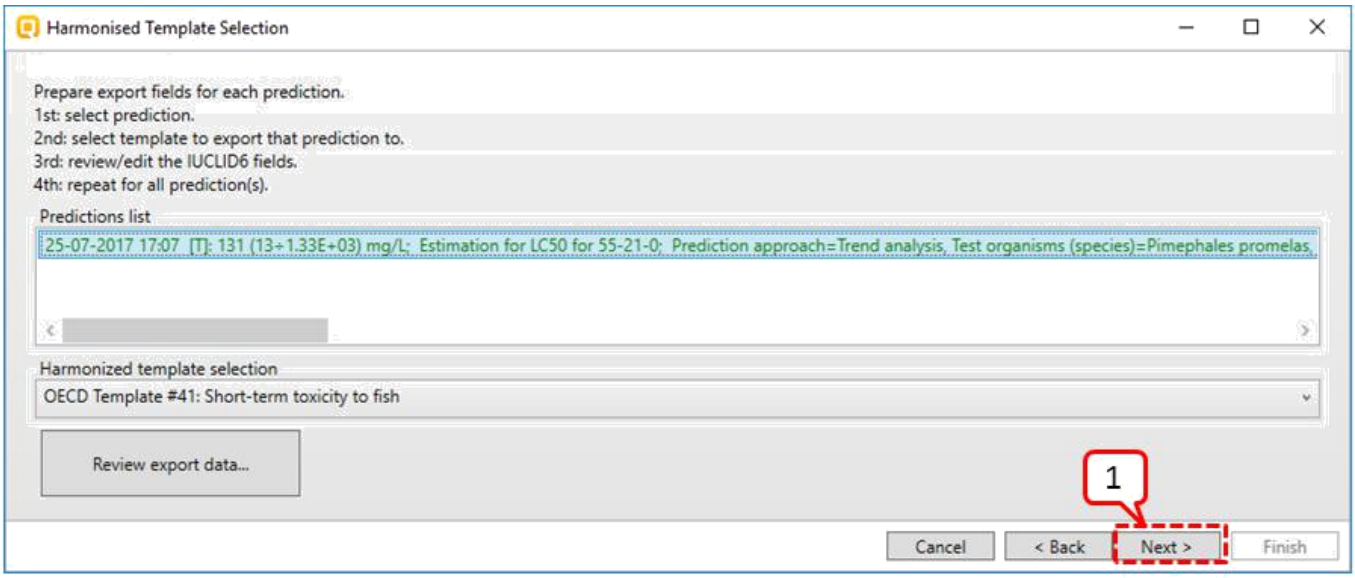


図 7

接続パラメータ（1）を提供してください（図 8）。これらには、以下のものがあります：

- WebServices Server：実行中の IUCLID 6 サーバの IP アドレスまたは DNS 名。
- Port: サーバが感知している TCP ポート（通常 8080）。
- I6 Username：サーバに有効な IUCLID 6 ユーザ名。
- I6 Password: ：対応するパスワード。

上記のすべてのパラメータを入力した後、 Next （2）をクリックしてサーバへの接続を試みてください。接続に成功すると、サーバ、ポート、およびユーザ名が保存され、後でリスト（3）からサーバを選択することで取得できます。

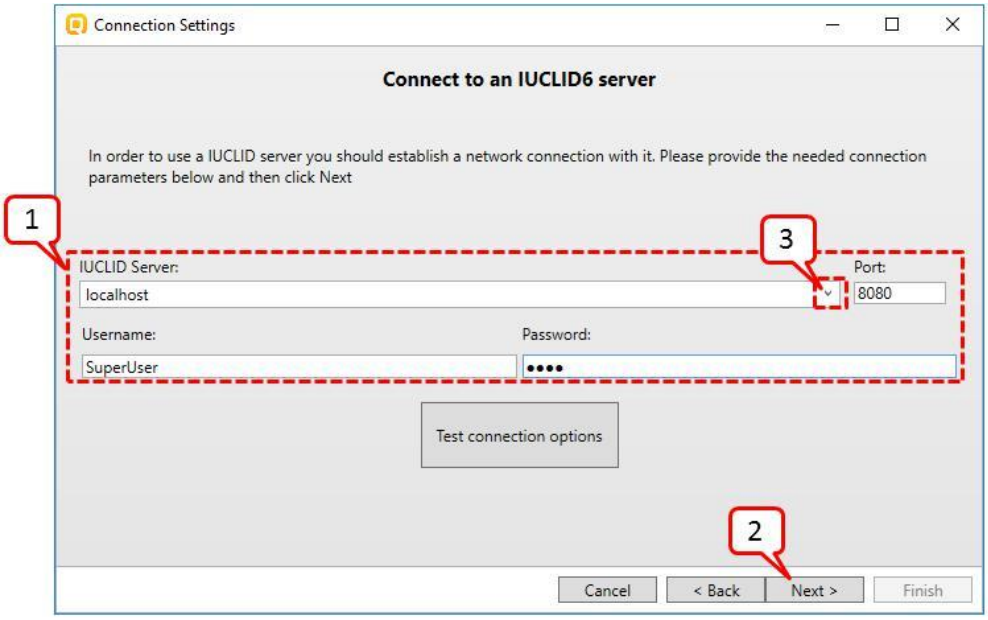


図 8

正しいパラメータが入力され、接続が確立されていると、次の画面が表示されます。この段階では Get All Substances ボタンをクリックして、IUCLID 6 データベースの該当するすべての物質を含むリストを取得できます。(図 9)

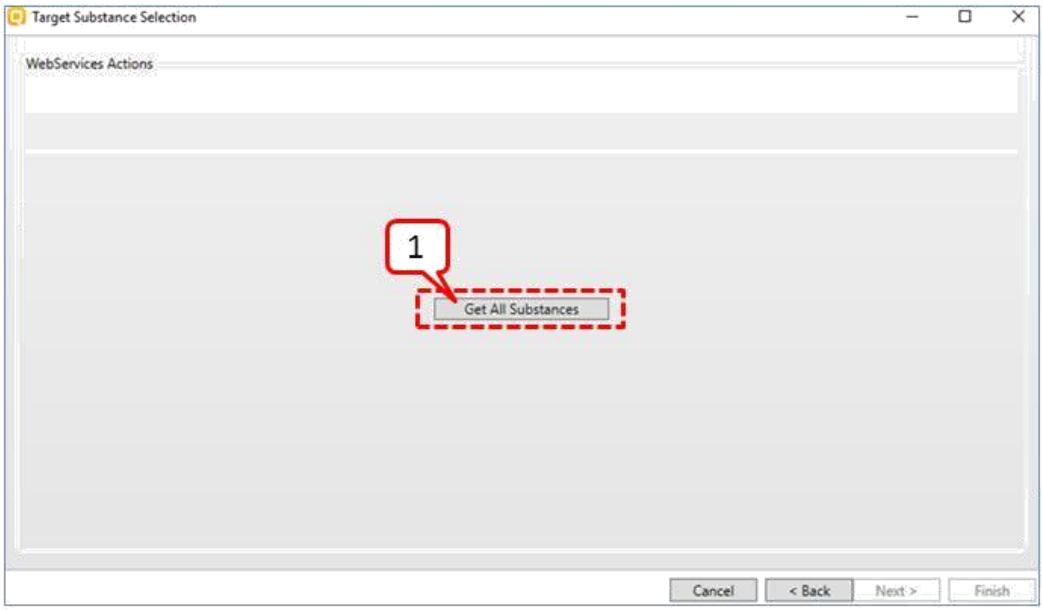


図 9

この段階で、IUCLID 6 データベース内で見つかったすべての物質のリストが提示されます。エンドポイント予測は、これらの物質のいずれかに割り振られます。そのどれかを選択して（1）Next（2）をクリックするとエクスポート処理が開始され、1 分ほどかかる場合があります（図 10）。

注：現在のリストは 1 つの物質のみで構成されています。物質を含むリストが利用可能であれば、CAS 登録番号（3）によって特定の物質を検索することができます（図 10）。

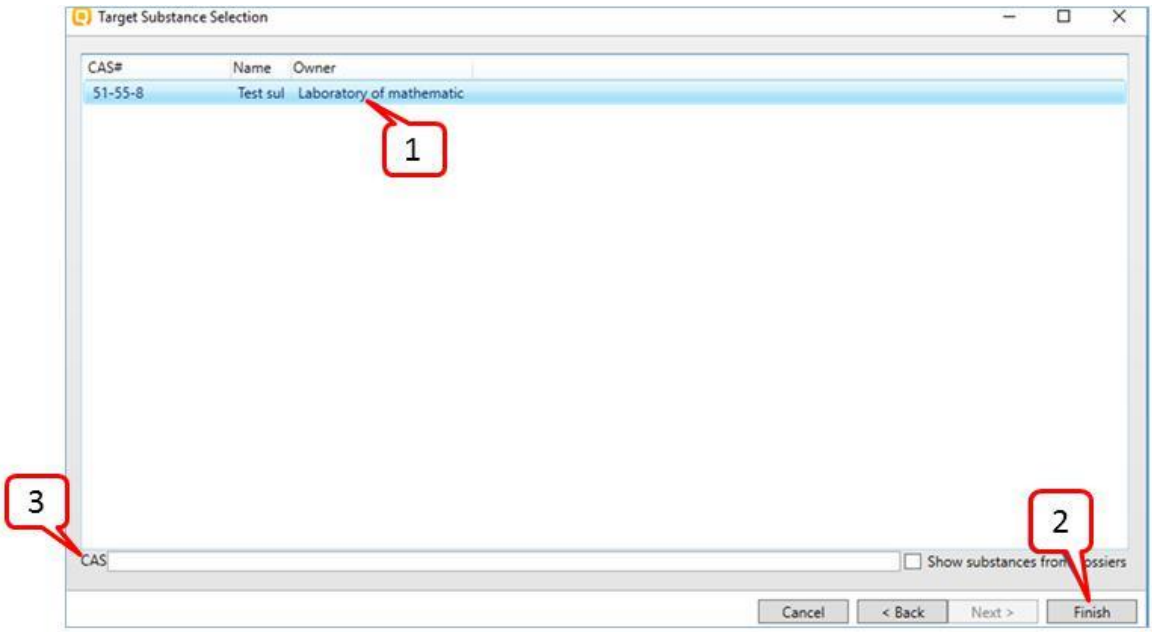
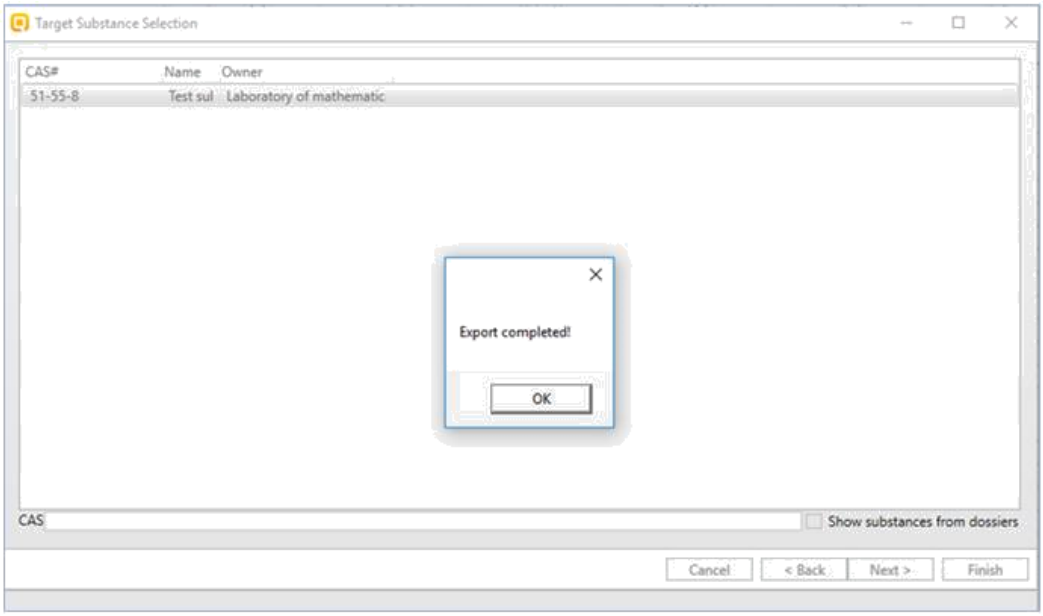


図 10

エクスポートプロセスが終了するとすぐに、データが IUCLID 6 に正常に転送されたかどうかが通知されます（図 11）。問題が発生した場合、詳細なログが表示され、管理者/開発者が問題を識別するのに役立ちます。

図 11


Target Substance Selection
CAS#
Name
Owner
51-55-8
Test sul
Laboratory of mathematic
Export completed!
OK
CAS
Show substances from dossiers
Cancel
< Back
Next >
Finish

D.4. Category definition（カテゴリ決定）

（訳注：この章の Category definition（カテゴリ決定）や Categorization（カテゴリ化）という言葉は、標的物質と何らかの類似性を持つ化学物質を集めてきて、ひとつの「物質のグループ」を作ることを指す。詳しい手順は Input モジュールで標的物質を入力した後、D.4.3 章以降の手順を参照のこと。D.5 章でデータギャップ補完を行うためには事前にこの D.4 章の通りに「物質のグループ」を作成しておくことが必要となる。）

基本情報

このモジュールを使うと、いくつかの化学物質グルーピング方法が使えるようになり、標的分子が含まれていて毒物学的な意味をもつカテゴリに振り分けることができるようになります。このモジュールはグルーピングメソッドに基づいており、様々な「類似性」の測定によって、化学物質カテゴリに化学物質群を振り分けることができます。これによりカテゴリ内のデータギャップをリード・アクロスやトレンドアナリシスで補完できるようになります。ここはワークフローにおける重要なステップです。サブカテゴリ化をすることでよりよいカテゴリ決定ができるよう、ツールボックス内にいくつかの支援オプションがあります。

たとえば、ある大きなインベントリ内では、水生毒性の作用機序に応じて化学物質を分類することができます。または、標的化学物質に特定の作用機構があると判別したところから（カテゴリ決定を）始めると、それと同じ反応メカニズムによって化学結合し、実験結果が入手できる類似物質を見つけることができます。もし調査されるエンドポイントに関連した標的化学物質について、特定の作用機構や作用機序がないと判別されたときには、標的化学物質と構造的に類似した化学物質を検索してみてください。そこで、特定の作用機構や作用機序がある化学物質を除外すれば、よりよい検索結果が得られます。

D.4.1.Grouping methods（グルーピング法）

グルーピング法のリストにプロファイリング方法のリストも含めました（表 1）。

表 1.プロファイリングとグルーピング法のリスト

Profiling method（プロファイリング方法）	Grouping method（グルーピング法）
Predefined（既定のもの）	
Database affiliation（データベースの所属）	Database affiliation（データベースの所属）
Inventory affiliation（インベントリの所属）	Inventory affiliation（インベントリの所属）
OECD HPV Chemicals categories（OECD 高生産量化学物質カテゴリ）	OECD HPV Chemicals categories（OECD 高生産量化学物質カテゴリ）
Substance type（物質のタイプ）	Substance type（物質のタイプ）
US-EPA New Chemicals categories（米国環境保護庁の新化学物質カテゴリ）	US-EPA New Chemicals categories（米国環境保護庁の新化学物質カテゴリ）
General mechanistic（一般的な反応メカニズム）	
Biodeg probability（生分解率）(Biowin1)	Biodeg probability（生分解率）(Biowin1)
Biodeg probability（生分解率）(Biowin2)	Biodeg probability（生分解率）(Biowin2)
Biodegradation ultimate(最終生分解) (Biowin3)	Biodegradation ultimate（最終生分解）(Biowin3)
Biodegradation Primary（一次生分解） (Biowin4)	Biodegradation Primary（一次生分解） (Biowin4)
Biodeg probability（生分解率） (Biowin5)	Biodeg probability（生分解率） (Biowin5)
Biodeg probability（生分解率） (Biowin6)	Biodeg probability（生分解率） (Biowin6)
Biodeg probability（生分解率） (Biowin7)	Biodeg probability（生分解率） (Biowin7)
BioHC half-life（生分解 BioHC 半減期） (Biowin)	BioHC half-life（生分解 BioHC 半減期） (Biowin)
DNA binding by OECD（DNA 結合；OECD）	DNA binding by OECD（DNA 結合；OECD）
DNA binding by OASIS v1.4（DNA 結合；OASIS）	DNA binding by OASIS v1.4（DNA 結合；OASIS）
Hydrolysis half-life(加水分解半減期）(Ka, pH7,Hydrowin)	Hydrolysis half-life(加水分解半減期)(Ka, pH7,Hydrowin)
Hydrolysis half-life(加水分解半減期）(Ka, pH8,Hydrowin)	Hydrolysis half-life(加水分解半減期)(Ka, pH8,Hydrowin)
Hydrolysis half-life(加水分解半減期）(Kb, pH7,Hydrowin)	Hydrolysis half-life(加水分解半減期)(Kb, pH7,Hydrowin)
Hydrolysis half-life(加水分解半減期）(Kb, pH8,Hydrowin)	Hydrolysis half-life (加水分解半減期)(Kb, pH8,Hydrowin)
Hydrolysis half-life（加水分解半減期）(pH 6.5-7.4)	Hydrolysis half-life（加水分解半減期）(pH 6.5-7.4)
Ionization at pH=1（pH=1 におけるイオン化）	Ionization at pH=1（pH=1 におけるイオン化）
Ionization at pH=4（pH=4 におけるイオン化）	Ionization at pH=4（pH=4 におけるイオン化）
Ionization at pH=7.4（pH=7.4 におけるイオン化）	Ionization at pH=7.4（pH=7.4 におけるイオン化）
Ionization at pH=9（pH=9 におけるイオン化）	Ionization at pH=9（pH=9 におけるイオン化）

Protein binding by OASIS v1.4（タンパク質結合；OASIS）	Protein binding by OASIS v1.4（タンパク質結合；OASIS）
Protein binding by OECD（タンパク質結合；OECD）	Protein binding by OECD（タンパク質結合；OECD）
Protein binding potency（タンパク質結合能）	Protein binding potency（タンパク質結合能）
Toxic hazard classification by Cramer (original)（Cramer による毒性有害性分類、オリジナル）	Toxic hazard classification by Cramer (original)（Cramer による毒性有害性分類、オリジナル）
Toxic hazard classification by Cramer (with extension)（Cramer による毒性有害性分類、拡張版）	Toxic hazard classification by Cramer (with extension)（Cramer による毒性有害性分類、拡張版）
Endpoint specific （特定のエンドポイント）	
Acute aquatic toxicity classification by Verhaar (Modified)（Verhaar による急性水生生物毒性分類）	Acute aquatic toxicity classification by Verhaar (Modified)（Verhaar による急性水生生物毒性分類）
Acute aquatic toxicity MOA by OASIS（OASIS, 急性水生生物毒性作用機序）	Acute aquatic toxicity MOA by OASIS（OASIS, 急性水生生物毒性作用機序）
Aquatic toxicity classification by ECOSAR（ECOSAR, 水生生物毒性の分類）	Aquatic toxicity classification by ECOSAR（ECOSAR,水生生物毒性の分類）
Bioaccumulation – metabolism alerts（生物蓄積 – 代謝アラート）	Bioaccumulation – metabolism alerts（生物蓄積 – 代謝アラート）
Bioaccumulation – metabolism half-lives（生物蓄積 – 代謝半減期）	Bioaccumulation – metabolism half-lives（生物蓄積 – 代謝半減期）
Biodegradation fragments (BioWIN MITI)（生分解性フラグメント BioWIN MITI）	Biodegradation fragments (BioWIN MITI)（生分解性フラグメント BioWIN MITI）
Carcinogenicity (genotox and nongenotox)alerts by ISS（ISS, 遺伝毒性および非遺伝毒性発がん性）	Carcinogenicity (genotox and nongenotox)alerts by ISS（ISS, 遺伝毒性および非遺伝毒性発がん性）
DART scheme（DART スキーム）	DART scheme（DART スキーム）
DNA alerts for AMES by OASIS（AMES 試験の DNA アラート、OASIS）	DNA alerts for AMES by OASIS（AMES 試験の DNA アラート、OASIS）
DNA alerts for CA and MNT by OASIS（CA と MNT の DNA アラート、OASIS）	DNA alerts for CA and MNT by OASIS（OASIS、CA と MNT の DNA アラート）
Eye irritation/corrosion Exclusion rules by BfR（BfR,目の刺激/腐食の除外ルール）	Eye irritation/corrosion Exclusion rules by BfR（BfR,目の刺激/腐食の除外ルール）
Eye irritation/corrosion Inclusion rules by BfR（BfR,目の刺激/腐食の包含ルール）	Eye irritation/corrosion Inclusion rules by BfR（BfR,目の刺激/腐食の包含ルール）
in vitro mutagenicity (Ames test) alerts by ISS（ISS, in vitro 突然変異誘発（AMES 試験）アラート）	in vitro mutagenicity (Ames test) alerts by ISS（ISS, in vitro 突然変異誘発（AMES 試験）アラート）
in vivo mutagenicity (Micronucleus) alerts by ISS（ISS にもとづく in vivo 突然変異誘発（微小核）アラート）	in vivo mutagenicity (Micronucleus) alerts by ISS（ISS, in vivo 突然変異誘発（微小核）アラート）
Keratinocyte gene expression（ケラチン生成細胞遺伝子発現）	Keratinocyte gene expression（ケラチン生成細胞遺伝子発現）

Oncologic Primary Classification（Oncologic 一次分類）	Oncologic Primary Classification（Oncologic 一次分類）
Protein binding alerts for Chromosomal aberration by OASIS（OASIS, 染色体異常のタンパク質結合アラート）	Protein binding alerts for Chromosomal aberration by OASIS(OASIS, 染色体異常のタンパク質結合アラート)
Protein binding alerts for skin sensitization according to GHS（GHS, 皮膚感作のタンパク質結合アラート）	Protein binding alerts for skin sensitization according to GHS（GHS, 皮膚感作のタンパク質結合アラート）
Protein binding alerts for skin sensitization by OASIS（OASIS, 皮膚感作のタンパク質結合アラート）	Protein binding alerts for skin sensitization by OASIS（OASIS, 皮膚感作のタンパク質結合アラート）
Respiratory sensitization（呼吸器感作）	Respiratory sensitization（呼吸器感作）
Retinoic Acid Receptor Binding（レチノイン酸受容体結合）	Retinoic Acid Receptor Binding(レチノイン酸受容体結合)
rtER Expert System - USEPA（rtER エキスパートシステム - 米国環境省）	rtER Expert System - USEPA（rtER エキスパートシステム - 米国環境省）
Skin irritation/corrosion Exclusion rules by BfR（BfR, 皮膚の刺激/腐食の除外ルール）	Skin irritation/corrosion Exclusion rules by BfR（BfR, 皮膚の刺激/腐食の除外ルール）
Skin irritation/corrosion Inclusion rules by BfR（BfR, 皮膚の刺激/腐食の包含ルール）	Skin irritation/corrosion Inclusion rules by BfR（BfR, 皮膚の刺激/腐食の包含ルール）
Empiric（経験則的）	
Chemical elements（化学元素）	Chemical elements（化学元素）
Groups of elements（元素周期表の「族」）	Groups of elements（元素周期表の「族」）
Lipinski Rule Oasis（Lipinski 則, Oasis）	Lipinski Rule Oasis（Lipinski 則, Oasis）
Organic functional groups（有機官能基）	Organic functional groups（有機官能基）
Organic functional groups (nested)（有機官能基・入れ子状）	Organic functional groups (nested)(有機官能基・入れ子状)
Organic functional groups （有機官能基）(US-EPA)	Organic functional groups （有機官能基）(US-EPA)
Organic functional groups, Norbert Haider (checkmol)（有機官能基・ Norbert Haider ・ Checkmol）	Organic functional groups, Norbert Haider (checkmol)（有機官能基・Norbert Haider ・ Checkmol）
Structure similarity（化学構造の類似性）	Structure similarity（化学構造の類似性）
Tautomers unstable（不安定な互変異性体）	Tautomers unstable（不安定な互変異性体）
Toxicological（毒物学）	
Repeated dose Toxity (HESS)（反復投与毒性、HESS）	Repeated dose Toxity (HESS)（反復投与毒性、HESS）
Custom（カスタマイズする）	
Example Prioritization Scheme (PBT)（実例優先順位付けスキームの例 、PBT）	Example Prioritization Scheme (PBT)（実例優先順位付けスキームの例、PBT）

グルーピング法の背景情報の概要については、表 2 に記載されています。

表 2.グルーピング法の背景情報の概要：

グルーピング法	背景情報の概要	グルーピング法	背景情報の概要
General mechanistic（一般的な反応メカニズム）		Biodegradation fragments（生分解性フラグメント）(BIOWIN MITI)	BIOWIN は、環境微生物の混合集団の存在下で有機化合物が迅速に好気性・嫌気性生分解される確率を見積もります。
DNA binding by OECD （DNA 結合；OECD）	このグルーピング法には DNA 結合に関するカテゴリや化学的手法が含まれています。 このグルーピング法は、特に genotoxicity（遺伝毒性） エンドポイントに関連しています。	Eye irritation / corrosion exclusion（眼刺激性/腐食の除外） BfR の規則	このグルーピング法は眼刺激性/腐食に関連する物理化学的な除外基準をベースにしています。
DNA binding by OASIS v1.4 （DNA 結合；OASIS）	このグルーピング法には DNA 結合に関するカテゴリや化学的手法が含まれています。このグルーピング法は、 特に genotoxicity（遺伝毒性） エンドポイントに関連しています。	Eye irritation / corrosion inclusion（眼刺激性/腐食の包含） BfR の規則	このグルーピング法は、目の炎症や腐食の構造アラートに基づいています。
Estrogen receptor binding （エストロゲン受容体結合）	このグルーピング法には、エストロゲン受容体（ER）結合に関するシンプルなカテゴリが含まれています。 その方法は、魚と哺乳類の生殖毒性エンドポイントに関連しています。	in vivo mutagenicity (Micronucleus) alerts by ISS （ISS, in vivo 突然変異誘発（微小核）アラート）	このプロファイラは、 in vivo 小核試験における影響の誘導に関連することが知られている構造アラートに基づいています。
Protein binding by OASIS v1.4 （タンパク質結合；OASIS）	このグルーピング法にはタンパク質結合のカテゴリや化学的反応メカニズムが含まれています。 その方法は、特に皮膚および呼吸器感作性と、急性水生毒性に関連しますが、また染色体異常、急性吸入毒性にも関連します。	in vitro mutagenicity (Ames test) alerts by ISS （ISS, in vitro 突然変異誘発（AMES 試験）アラート）	このプロファイラは、ソフトウェア Toxtree の 変異原性 / 発がん性モジュールに基づいています。これは 30 の構造的アラート（SAs）のリストに基づいて、in vitro（AMES 試験）変異原性を推定するための決定ツリーとして機能します。

		Carcinogenicity (genotox and nongenotox)alerts by ISS （ISS, 遺伝毒性および非毒性発がん性）	このプロファイラは、ソフトウェア Toxtree の該当モジュールの拡張バージョンです。55 の構造アラート（SAs）リストに基づいて、発がん性を推定するための決定ツリーとして機能します。　構造アラートは主に遺伝毒性発がん物質の作用機序の知識を参照しています（したがって、細菌の変異原性の活性も参照しています）が、潜在的な非遺伝毒性発がん物質アラートに目印をつける構造アラートも多数含まれています。
Toxic hazard classification by Cramer (original)（Cramer による毒性有害性分類、オリジナル）	これはヒトの健康評価のために、化学物質を 3 つの区分けのうち 1 つに分類する、歴史的なグルーピング法です。	Oncologic Primary Classification（腫瘍の一次分類）	この OncoLogic Primary Classifier プロファイラは、LMC と OECD が開発した分子決定で構成されており、米国環境保護庁の発がん性予測用の OncoLogic™がんエキスパートシステムが対象としている、発がん性物質の化学的分類の構造基準にならっています。
Endpoint specific　（特定のエンドポイント）		Skin irritation/corrosion Exclusion rules by BfR（BfR にもとづく皮膚の刺激/腐食の除外ルール）	この皮膚刺激性/腐食性の除外基準は、皮膚刺激性や腐食性を示さない化学物質を特定するための物理化学的カットオフ値に基づいています。
Acute aquatic toxicity MOA by OASIS （OASIS, 急性水生生物毒性作用機序）	このグルーピング法は、1997 年以前に ファットヘッドミノーを使用した化学物質の試験に基づき、急性水生毒性（行動）作用機序によって化学物質を分類しています。 この方法は、特に急性水生毒性のエンドポイントに関連しています。この方法ではファットヘッドミノーのデータからの作用機序しか説明できないため、急性水生毒性にリンクされている他のグルーピング法との組み合わ	Skin irritation/corrosion Inclusion rules by BfR（BfR, 皮膚の刺激/腐食の包含ルール）	その反応メカニズムによる刺激、腐食、または刺激と腐食の両方を引き起こすことについて、陽性判定である化学物質を分類できる構造アラートが含まれています。今では、この 40 の有機構造フラグメントが、皮膚の炎症や腐食の包含規則をカバーしています。

	せで使用されることをお勧めします。		
Aquatic toxicity classification by ECOSAR （ECOSAR、水生生物毒性の分類）	ECOSAR は化学物質の分類を判別するためのグルーピング法です。これは特に水生毒性のエンドポイントに関連しています。この方法は最も堅牢な反応メカニズム的なグルーピング法であるため、選定手段になることが多いです。	Empiric（経験則的）	
Acute aquatic toxicity classification by Verhaar （Verhaar による急性水生生物毒性分類）	これは、1992 年までの魚の急性毒性の経験則に基づく反応性による歴史的なグルーピング法です。この方法は、特に急性水生毒性のエンドポイントに関連しています。 この方法では、反応性の大まかな分類（不活性、不活性、反応、および特定の作用をする化学物質）しか表せないので、急性水生毒性に関わる他のグルーピング法との組み合わせで使用されることをお勧めします。	Organic functional groups （有機官能基）	この Organic Functional Groups (OFG) システムは、大規模データベースから有機化学物質の様々な機能性・特徴的構造フラグメントの分類化と体系化を導入し、構造的類似化学物質を同定できるよう設計されています。 Organic functional groups （有機官能基）とは、その分子の特徴的な化学反応を決める、分子内の原子および/または結合の特定の群です。
Bioaccumulation – metabolism alerts （生物蓄積 – 代謝アラート）	これら 2 つのグルーピング法は、米国環境保護庁が開発した生物濃縮の BCFBAF モデルに使用されている構造アラートに基づいています。 それは特に生物濃縮のリード・アクロスで化学物質を分類するときに、分類不可能なものを識別することに関連しています。	Organic functional groups (nested) （有機官能基・入れ子状）	この OFG プロファイルは現在、同じ官能基を含む OFG (一般) と OFG（入れ子状）に細分されています。その違いは、OFG (一般)は 標的化合物に存在する全ての官能基を表示するのに対し、OFG（入れ子状）では、大きな官能基の一部分である官能基については表示しません。したがって、OFG（入れ子状）システム内で「重なっている」基は、実際には新しい別の官能基を持っています。
Bioaccumulation – metabolism half-lives （生物蓄積 – 代謝半減期）			

標的化学物質の類似物質を検索するときは、プロファイリングの出力によって最も適切な方法が決定されます。次のような推奨事項があります。

Recommendations

If specific mechanisms or modes of action are identified for a target chemical, which are relevant for the investigated endpoint, then it is recommended to search for chemicals which have the same mechanisms or modes of action. The search results can then be refined by eliminating those chemicals which are structurally most dissimilar.

If no specific mechanisms or modes of action are identified for a target chemical, which are relevant for the investigated endpoint, then it is recommended to search for chemicals which are structurally similar to the target chemical. The search results can then be refined by eliminating those chemicals which have specific mechanisms or modes of action.

[推奨事項：もし、調査されたエンドポイントに関連する標的化学物質について特定の反応メカニズムや作用機序があると判明した場合、同じメカニズムまたは作用機序をもつ化学物質を探すことが推奨されます。そして、構造的に最も類似していない化学物質を排除することにより、検索結果を絞り込むことができます。

もし、調査されたエンドポイントに関連する標的化学物質について特定の反応メカニズムや作用様式が見つからなかった場合、標的化学物質と構造的に類似する化学物質を検索することが推奨されます。そして、特定の反応メカニズムや作用様式をもつ化学物質を排除することにより、検索結果を洗練することができます。]

注意していただきたいのは、類似物質の検索は、Inventories の下の selected Databases or inventories（選択されたデータベースまたはインベントリ）内の化学物質に対し行われるということです。たとえば、「Skin Sensitation（皮膚感作性）ECETOC」と「Skin Sensitation」データベースのみを選択した場合、類似物質の検索は、これらのデータベースで入手可能になっている皮膚感作性の実験データをもつ化学物質の中だけで実行されます。同様に、化学物質のインベントリの検索を拡張するよう決定することができます。たとえば、データベースの「Carcinogenicity &Mutagenicity ISSCAN（発がん性と変異原性、ISSCAN）」と「Genotoxicity OASIS（遺伝毒性、OASIS）」と、ならびにインベントリの「Canada DSL（カナダ DSL）」を選択すると、ツールボックスではこれらのデータベースにも、この個別のインベントリにも類似物質があるかどうか調査してくれます。

Caution

The inventories contain between 5 000 and 100 000 substances. In order to accelerate the process of identifying the similar analogues only the databases are preliminary profiled and indexed (2D and 3D calculations). In this respect if one is to include one of them in the query this will produce longer calculation times.

[注意：インベントリには 5000～100000 もの物質が含まれています。似ている類似物質を同定するプロセスをさらに速くおこなうため、データベースだけが事前にプロファイリングとインデックス付け（2D・3D 計算）されています。
このため、クエリで（インベントリを）1 つ含める場合、計算時間が長くなります。]

D.4.2. カテゴリ構築の基本

データベース/インベントリを選択してカテゴリ決定に適用させるときは、次の事項が推奨されます。
ユーザの目的が、特定のエンドポイントにおいて実験データが入手可能な類似物質だけを検索したいという場合、このエンドポイントについての（試験）結果を持つデータベースのみを選択する必要があります。その際はインベントリを選択しないでください。

カテゴリ化する手順の最初のステップ は、エンドポイントに特化しないグルーピング法（オレンジ色に着色）を使用して、より構造的に似ている類似物質を探すことです。この点において、より広範なグループを選択する、化学構造に基づいたグルーピング法を使用されることをお勧めします。例えば：

- US EPA Categorization（米国環境保護庁　カテゴリ化）
- OECD Categorization（経済協力開発機構　カテゴリ化）
- Organic functional group（有機官能基）
- Structural similarity（化学構造の類似性）
- ECOSAR

次のステップ はサブカテゴリ化の手順を使用して、より広範なグループを絞り込むことです。このステップは反応メカニズム的なベースであり、エンドポイントに特化したグルーピング法（緑色で強調表示されています ）を使用することができます:

- DNA binding mechanism（DNA 結合反応メカニズム）
- Protein binding mechanism（タンパク質結合反応メカニズム）
- Genotoxicity/carcinogenicity（遺伝毒性/発がん性）
- Cramer rules（Cramer 則）
- Verhaar rule（Verhaar 則）
- Skin/eye irritation corrosion rules（皮膚/眼刺激性、腐食の法則）

サブカテゴリ化の最終ステップーカテゴリ化の最初のステップを行います。

D.4.3. カテゴリを決定する

カテゴリ決定の手順

カテゴリの作成は容易です。 類似物質の検索に使うプロファイル（1）を選び、**Define** をクリックします（2）。（図 1）

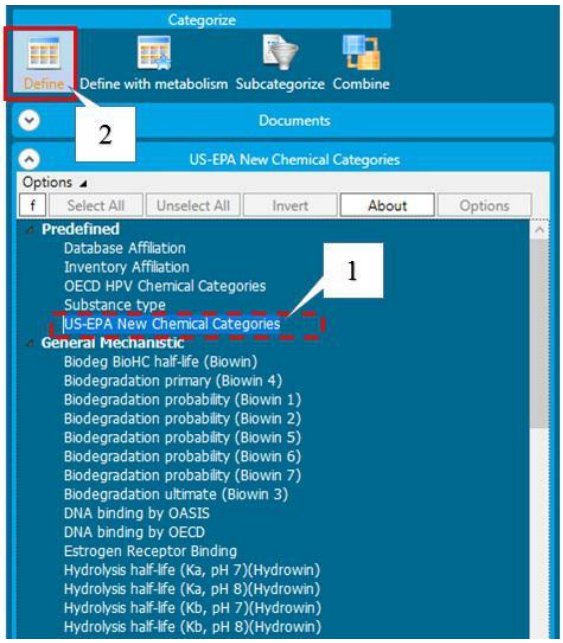


図 1

グルーピング法を実行する際には、選択したグルーピング法のすべてのカテゴリが提供され、標的化学物質（1）のカテゴリ（もしあれば）には、**Target** と **Profiles** のリストボックスが配置されます。（図 2）

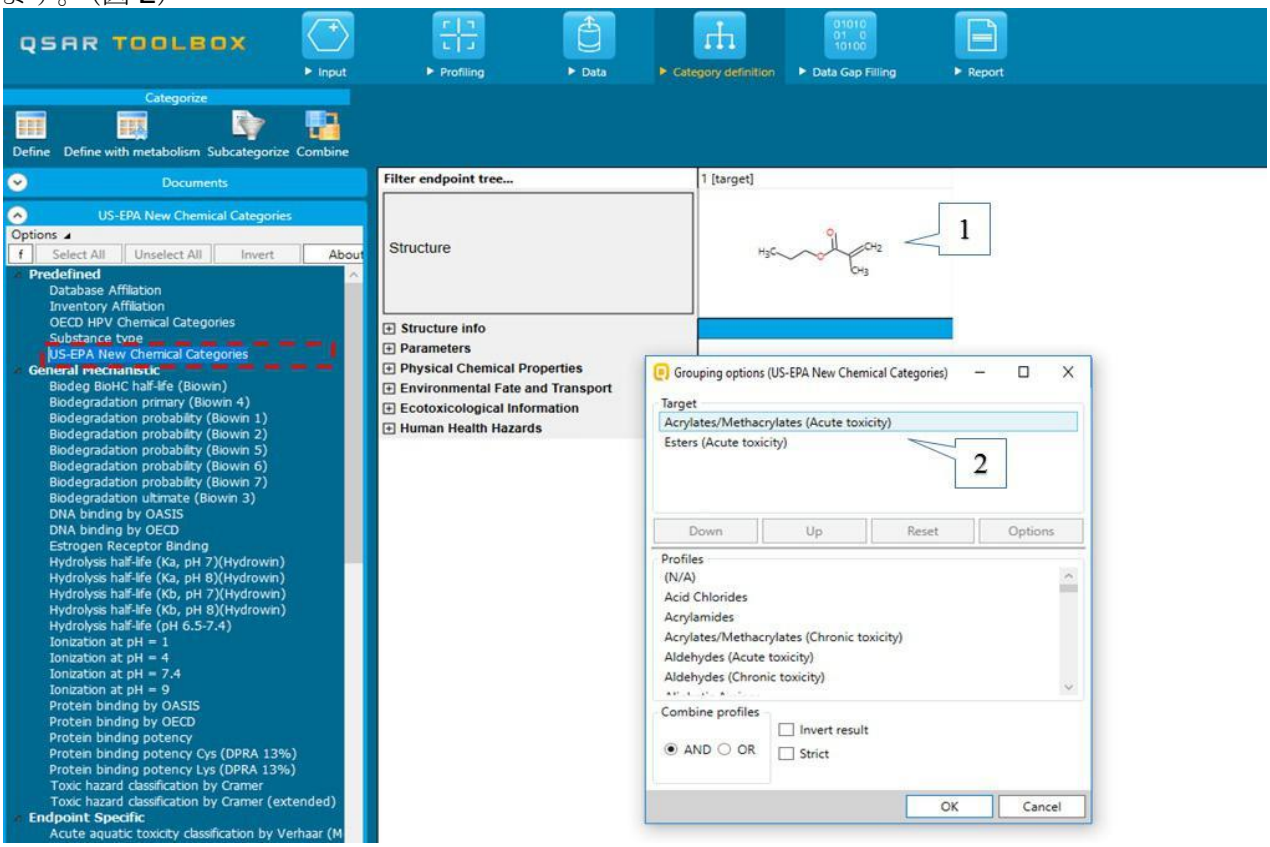


図 2

この段階でユーザは以下のことができます。

- （1）のいずれか 1 つを選択して**標的のカテゴリを削除**したり、**Down** ボタン（3）またはダブルクリックによりそれをすべて **Profiles** のパネル（2）に移動したりできます。（図３）

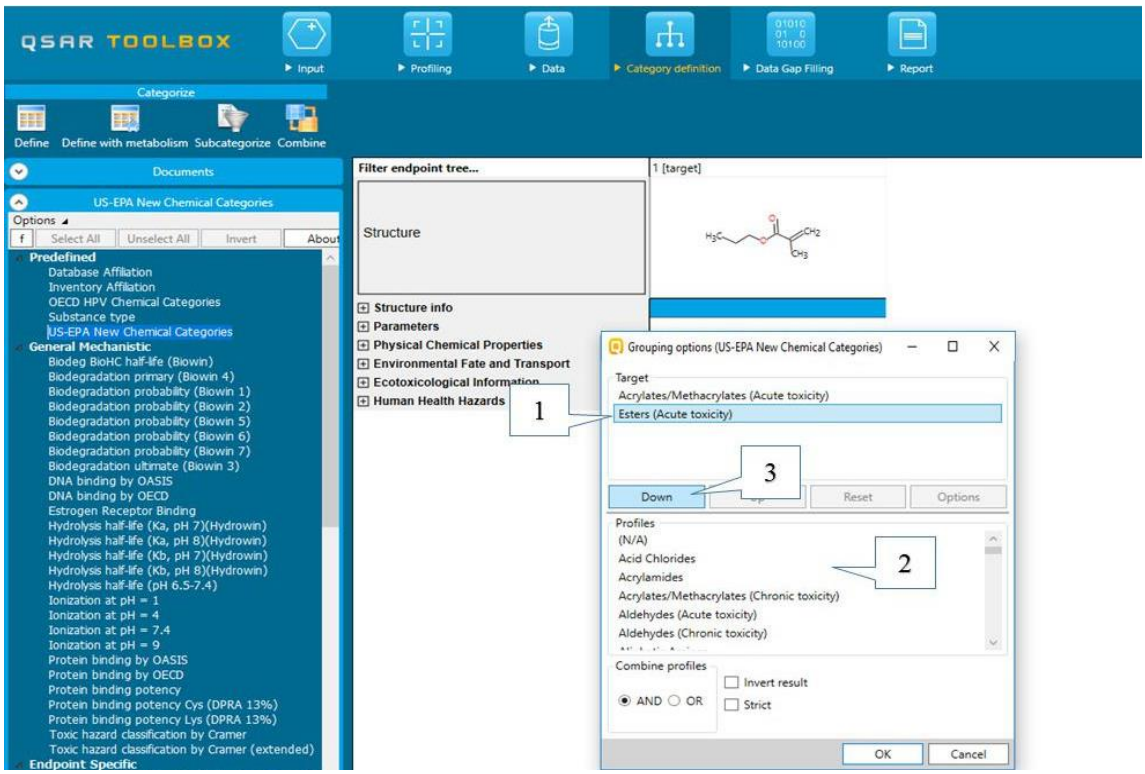


図 3

- すべてのプロファイルパネル（2）から標的プロファイル（1）のいずれか 1 つを選択することで、それを選択してカテゴリに追加することができます。また、**Up ボタン**（4）またはダブルクリックを押すことで、**Targets** と **Profiles** パネル（３）にそれを移動させることができます。（図 4）
-

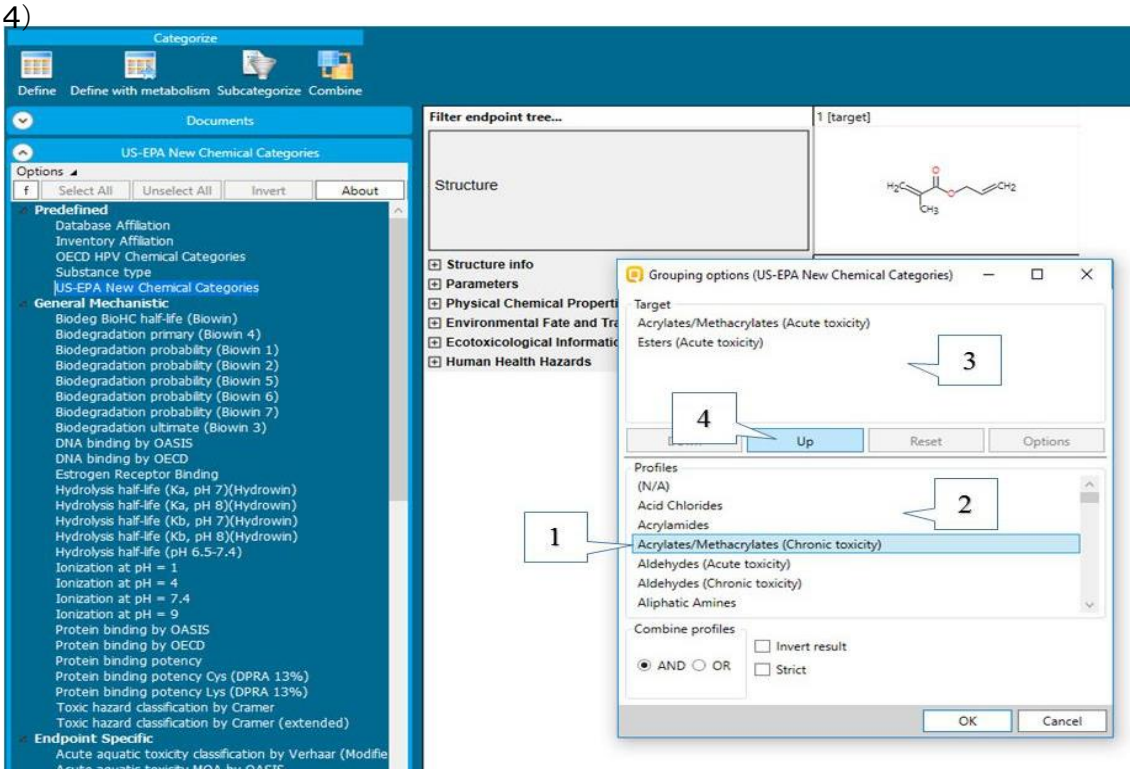
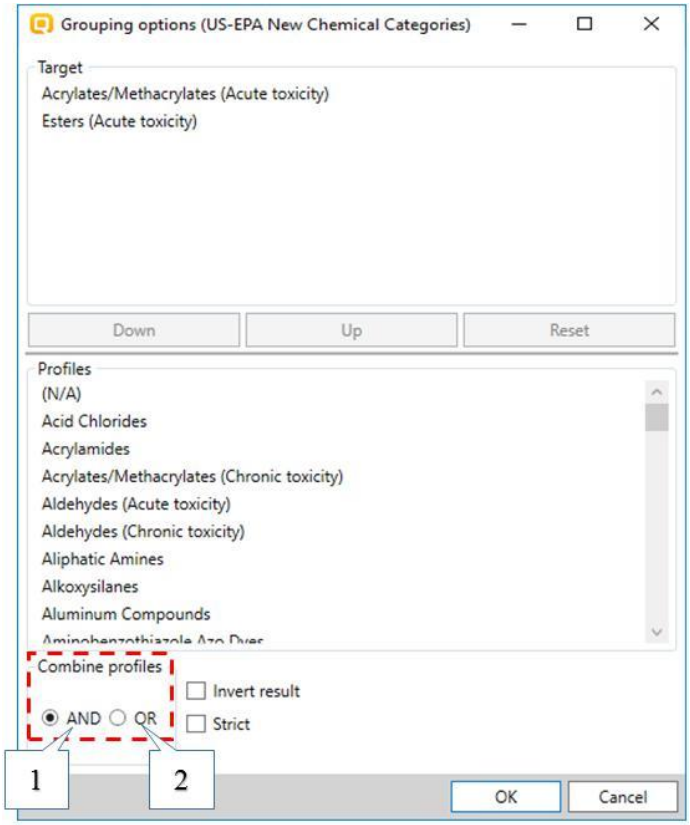


図 4

- **Combine profiles（プロファイルを組み合わせる）** - ここでは、And/Or 操作を使用して、標的プロファイルの組み合わせを選択することができます。**And** (1) を選ぶと選択されたすべてのカテゴリに、属する化学物質が検索されます。**Or** (2) を選ぶと、選択したいずれかのカテゴリに属する化学物質が検索されます。デフォルトでは **AND** が選択されています（図 5)。

図 5



- **Invert results** (1) - この機能は、標的プロファイルとは異なるプロファイルを持つ化学物質を検索します。下のスクリーンショットでは(図 6) 、もし標的が **Acrylates/Metacrylates (Acute toxicity)**（アクリレート／メタクリレート、急性毒性）や **Esters (Acute toxicity)**（エステル、急性毒性）に分類されていているときに Inverts results が選択された場合は、このソフトウェアは上記以外のプロファイルを持っている化学物質を検索します。
- **Strict** オプション（図 6-2） - Strict のチェックボックスがオンになっていると、選択されたカテゴリ以外に属する化学物質は検索されません。

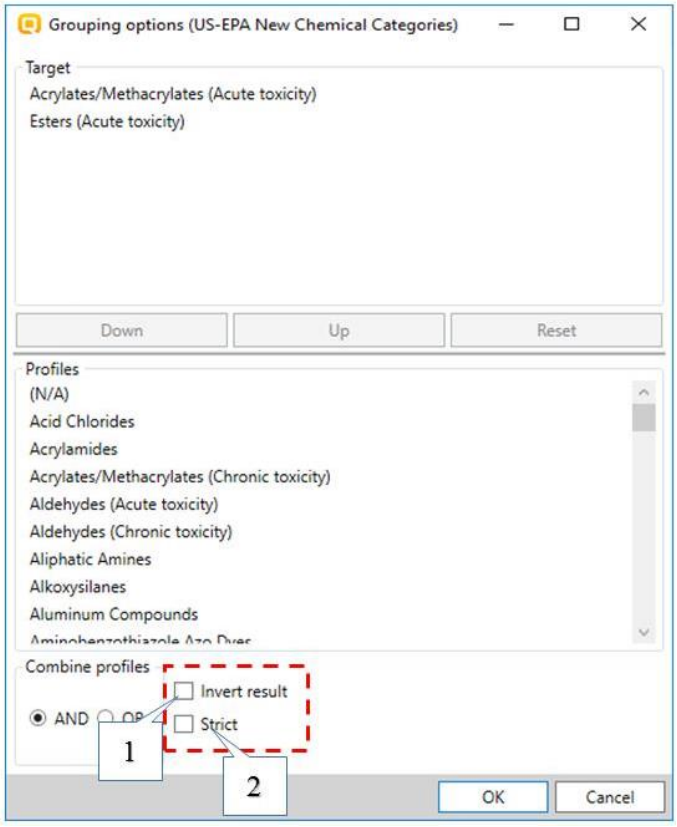


図 6

カテゴリを決定するウィンドウを閉じた後、カテゴリが作成され、**Grouping results** ダイアログボックスが表示されます（図 7）。

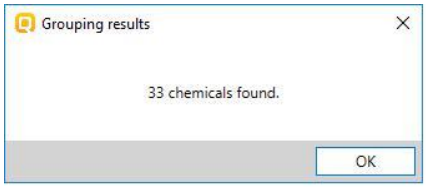


図 7

このソフトウェアが、ある標的化学物質をもとにして、同一のプロファイルにおいて選択されたデータベースから 33 の化学物質を特定したところです。

Grouping results ウィンドウを閉じた後、ソフトウェアは自動的にデータを収集するアクションを開始します。特定のエンドポイント（**Choose...**）またはすべてのエンドポイントを集める（**All endpoints**）デフォルト選択ができます（以下を参照してください）（図 8 ）。調査したエンドポイントに関連するデータベースをユーザが以前選択していた場合は 、どちらの選択肢も同じ結果を返します。



図 8

Recommendations

If the user has selected all databases under the Endpoint section, and selects **All endpoint** the gather data operation could be very time consuming due to the diversity of endpoints and size of databases. In this respect the user is recommended to always select only those databases, and endpoint paths, which are related to the investigated endpoints.

[推奨事項：
もし、そのエンドポイントのセクションについて全てのデータベースを選択し、All endpoint を選択したら、エンドポイントやデータのサイズが多岐にわたるため、データを集める操作に大変時間がかかります。そのため、調査したいエンドポイントにかかるデータベースとエンドポイントパスだけを選択することをおすすめします]。

データを読み込むと最後に、決定したカテゴリがドキュメントツリーに現れます。 (図 9-1)

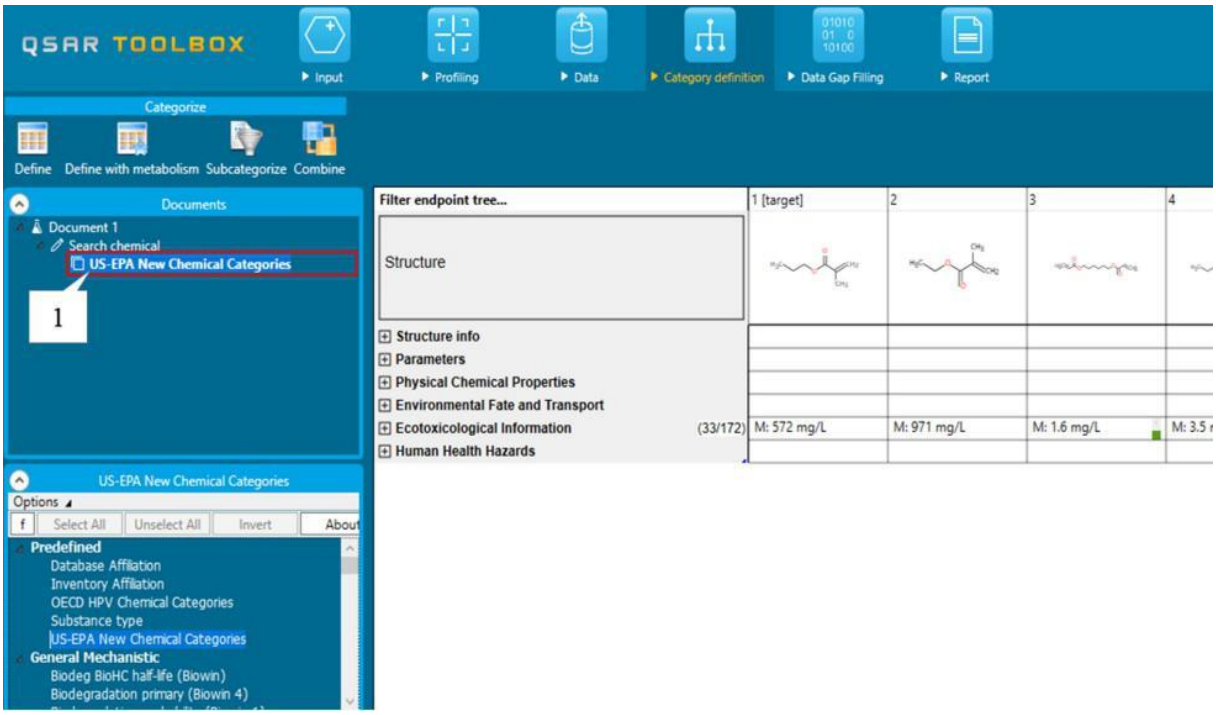


図 9

D.4.3.1. 単一の化学物質のカテゴリ化

単一の化学物質のカテゴリ化は、カテゴリを決定する手順のセクション（前ページまで、D.4.3）で説明しました　（カテゴリを決定するには「カテゴリ決定」「手順」のセクションを参照してください ）。

D.4.3.2. 混合物のカテゴリ化

あるひとまとまりの混合物をカテゴリ化すると、その混合物の成分の全てのプロファイルも考慮されるようになります。(図 1-1)

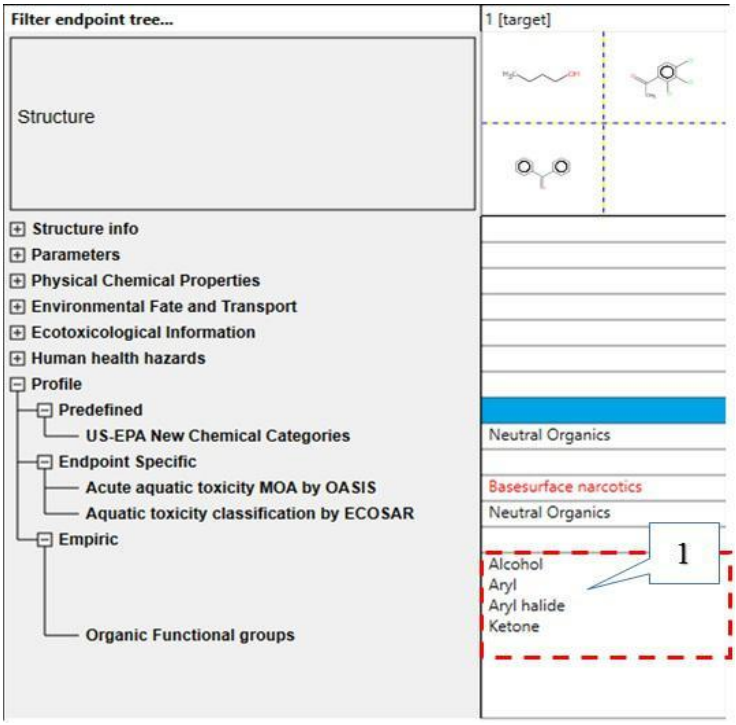


図 1

また、カテゴリ化パネルでは、 その混合物の成分の全てのプロファイルが標的物質のプロファイルとして考慮に入れられます（図 2-1)。

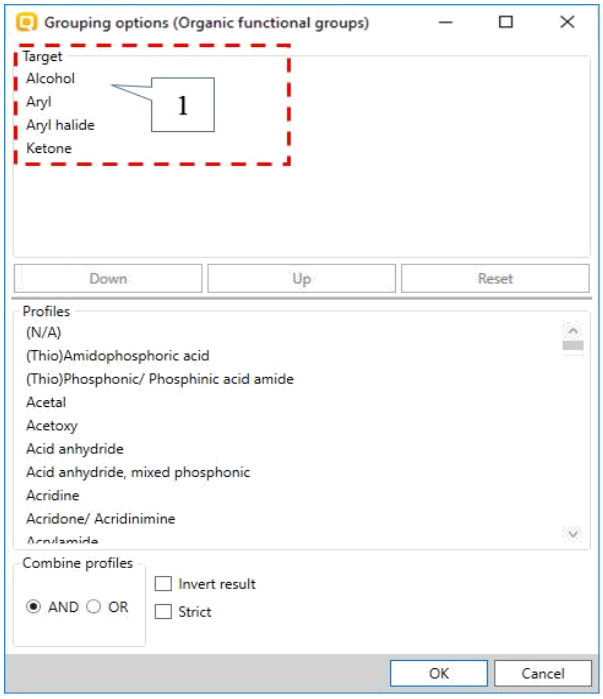


図 2

D.4.3.3. 階層型のプロファイリングスキームのプロファイリング結果を使用したカテゴリ化

例えばタンパク質、 DNA 結合などの階層型スキームからのプロファイリング結果は、Domain（ドメイン、範囲の意味）(1)、Mechanistic Alert（反応メカニズム的アラート）(2) と Structual Alert（構造アラート）(3) についての情報を提供します（図 1)。プロファイリング結果が階層的に表示されます。

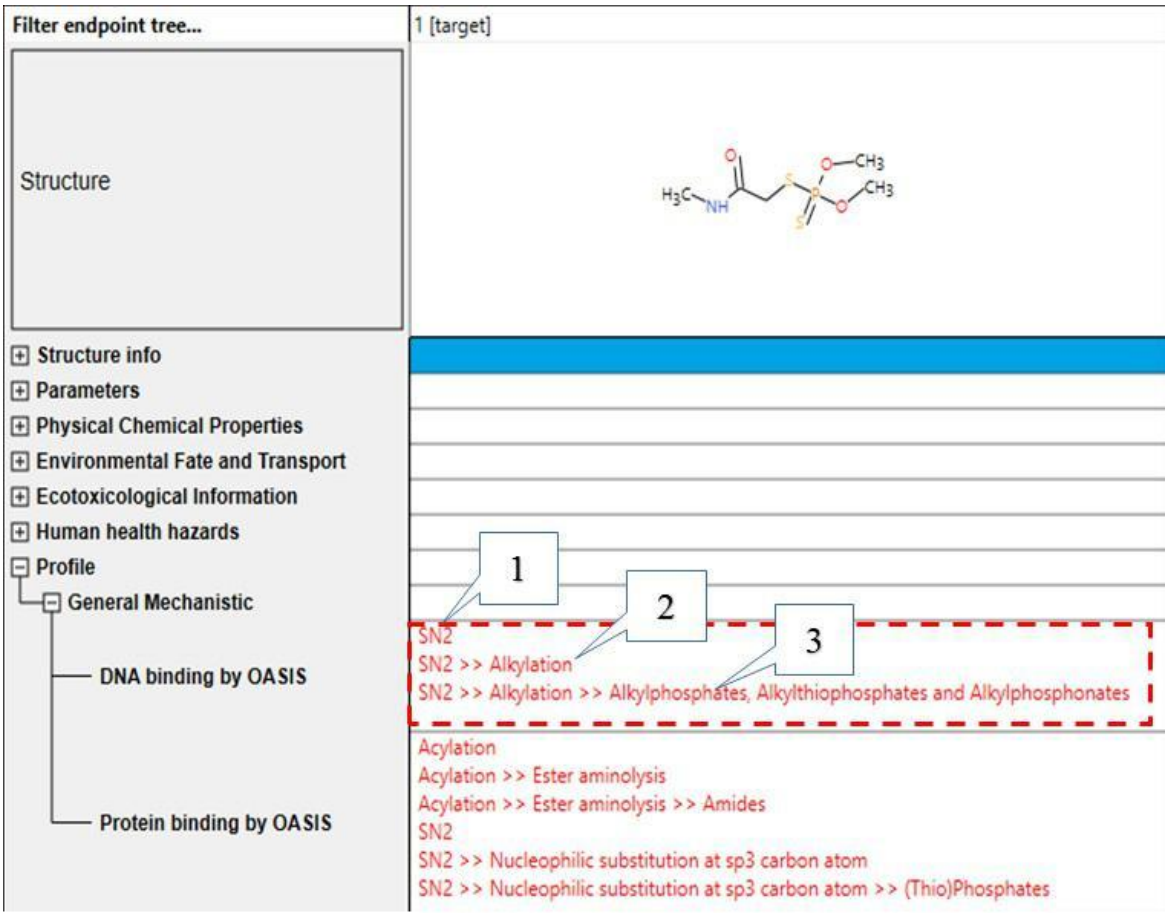


図 1

1. Domain （訳注：ここでは化学反応を大まかに区分した大範囲）
2. Domain >> Mechanistic alert
（訳注：Domain に入っている中サイズの範囲）
3. Domain >> Mechanistic alert >> Structural alert
（訳注：Mechanistic alert に入っている小範囲）

ツールボックスは、これらの各プロファイル結果を使用してカテゴリを決定する機会を提供しています。

カテゴリ化をしようとして"Domain（図 2）"カテゴリを使用する場合、ソフトウェアは "Domain"カテゴリの基準を反映させた化学物質を検索します（例えば SN2)。SN2 というドメインには 次のカテゴリ（訳注：Mechanistic alert のカテゴリ）が含まれています。

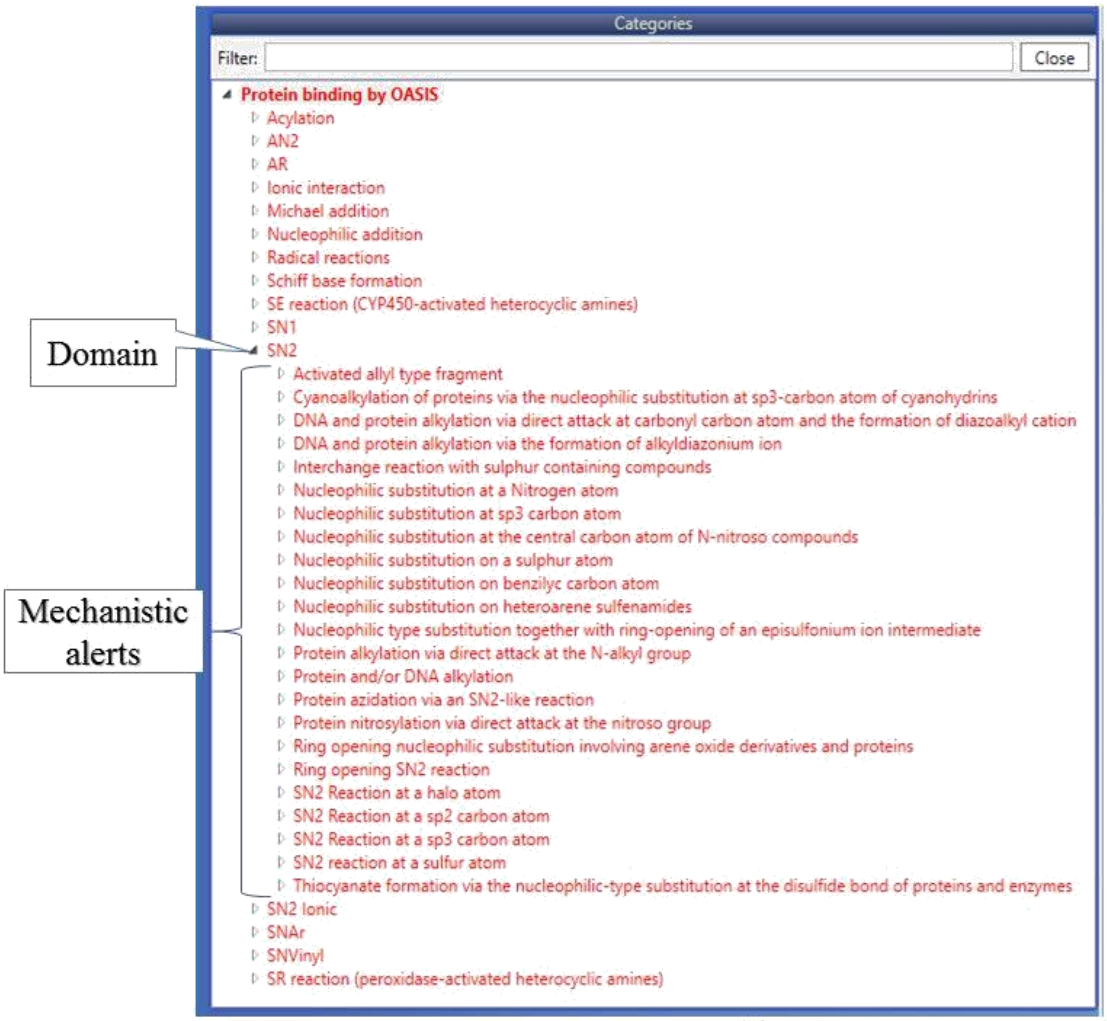


図 2

既定のカテゴリには、これら Mechanistic alert の 1 つ 、または全部一緒に分類されている化学物質が含まれており、それはカテゴリ決定で使われた上記のロジカルな操作（訳注：And/Or 操作など）で決められています。

既定のカテゴリ「SN2」は、7 つの Structural alert（図 3）を含んでいる”Nucleophilic substitution at sp3 carbon atom（sp3 炭素原子の求核置換反応）”（Mechanistic alert）より広範囲です。

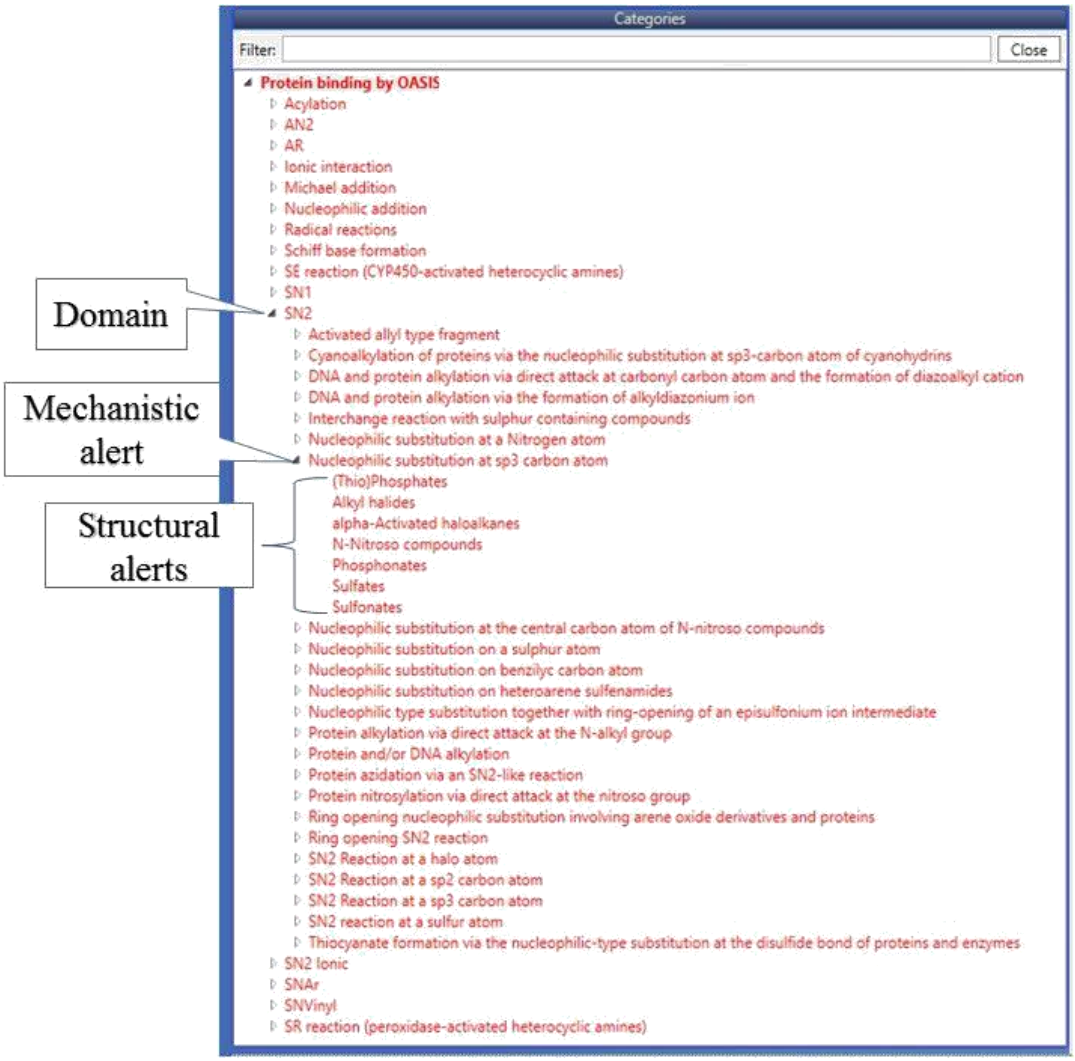


図 3

カテゴリ化の手順

もし階層型グルーピング法を使用してカテゴリを決定している場合は、カテゴリ化の手順中に別々にまたは同時に、Domain（ドメイン）、Mechanistic alert（反応メカニズム的なアラート）、Structural alert（構造アラート）を使用できます（図 4）。

カテゴリ（1）を選択し、下へ移動させて、ある構造アラートに関するカテゴリを除くこともできますし（図 4）、あるいはデフォルトで残したままにすることもできます。

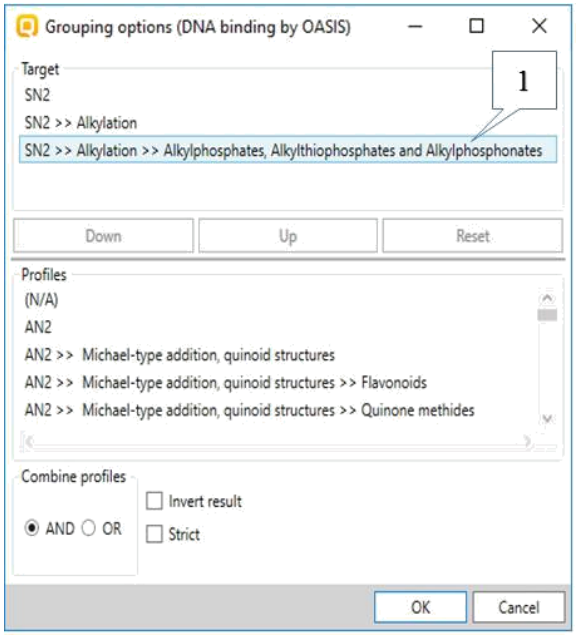


図 4

4. AND で組み合わせた Domain、 Mechanistic alert、Structural alert の基準を満たすような類似物質をこのソフトウェアが見つけられなかった場合は、次のメッセージが表示されます（図 5 ）。

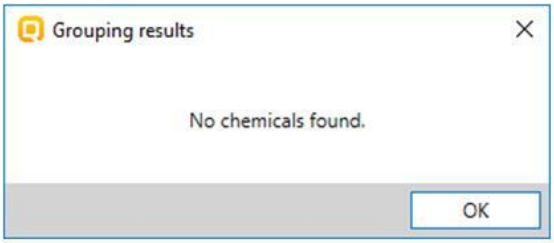


図 5

その際、より細かく特定したカテゴリ（Structural alert）（図 6）を削除することでカテゴリを広げることができるので、Domain と Mechanistic alert カテゴリを引き続き使用してください。

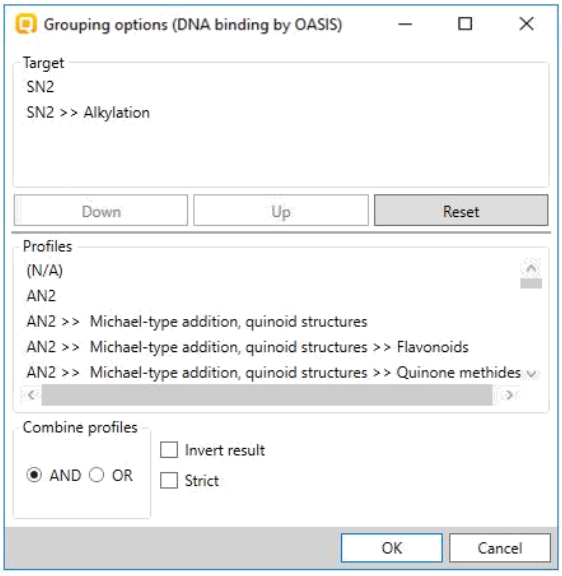


図 6

D.4.3.4. 代謝によるカテゴリ化

代謝変換を考慮したグループ化とは、化学物質の代謝活性化を考慮した類似物質を探すための手順です。ツールボックス 3 でも実装されていましたが、改良され機能拡張がなされたツールボックス 4 では、それぞれの代謝物のために様々な条件を提供しています。

代謝を考慮したカテゴリを作るには、**Define with metabolism**（代謝を含めて決定する）(1) をクリックします。(図 1) 新しいウィンドウが表示され、代謝（Documented＝実験値または Simulated＝予測値）を選択するよう尋ねられます (2)。標的のエンドポイント (3) が preliminary defined（既定）の場合 、関連する代謝変換が強調表示されます (4)。

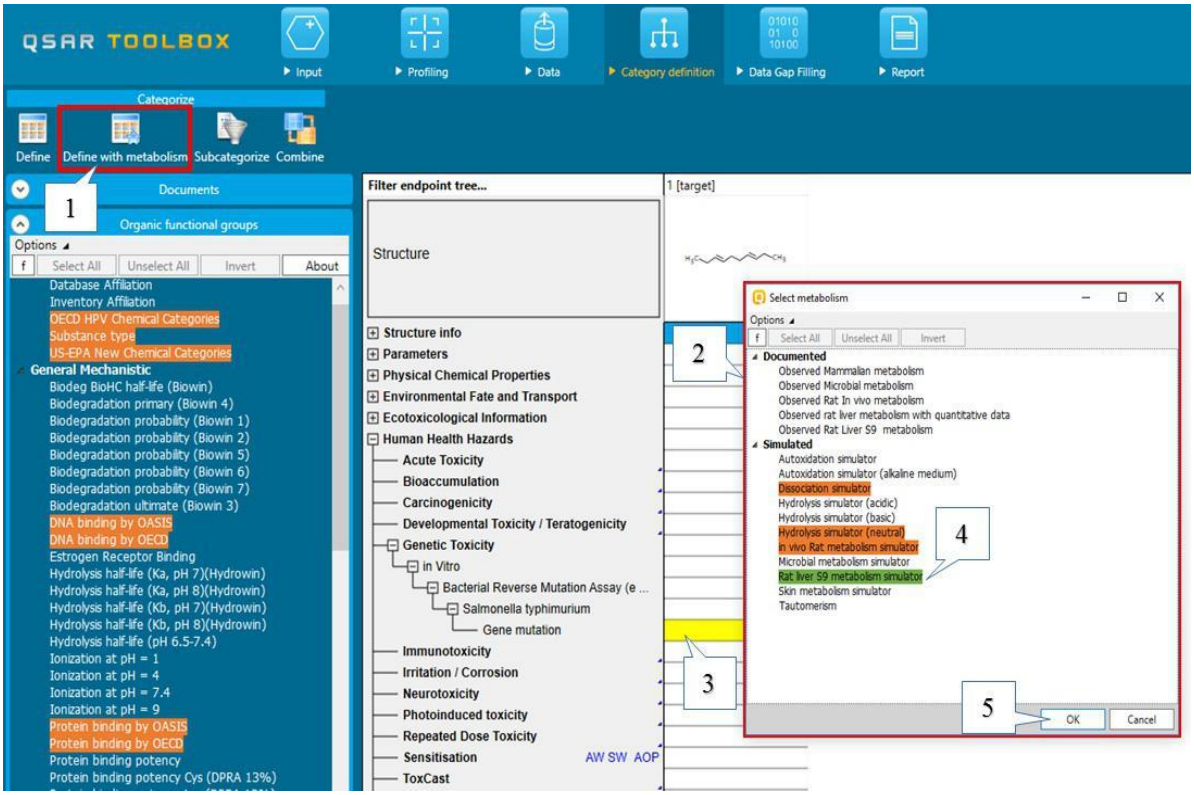


図 1

Map similarity options（代謝マップの類似性オプション）ダイアログが現れ、事前に選択されたシミュレータで生成した、標的化学物質のすべての代謝物が表示されます。このダイアログには 2 つのサブセクションがあり、以下図 2 に示して説明します。

- 1 つめのサブセクション（1）の親物質と、生成された各代謝物（事前に選択された代謝シミュレータに生成されたもの）は別々の行で表示されます。これにより、類似物質を見つけるために、それぞれの化学構造に様々な基準を決定できるようになりました。
- 2 つめのサブセクション（2）では、「親物質 + 代謝物」全体のパッケージとして作業します。すなわちこの基準は、全体のパッケージのために提供されているものであって、個々の代謝物のためのものではありません。

各々の化学構造式 ("Query"の列にある）のドロップダウン・メニュー（3）を利用すると、類似物質のさらなる検索基準の種類を設定できます。

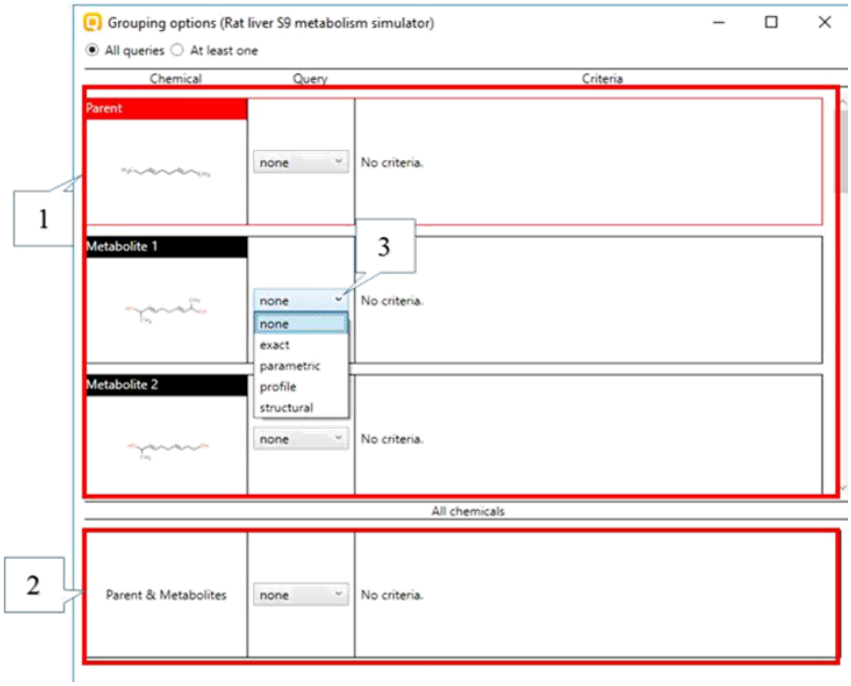


図 2

ドロップダウン・メニューの各オプションについて説明します。
None（なし） - デフォルトのオプション ; 基準が設定されていません。
Exact（完全一致） まったく同一の代謝物を持つ類似物質を検索する機会を提供しています。**Exact** オプションは、代謝物と「親物質+代謝物」パッケージのみで利用できますが、親物質については利用できません。
Parametric（パラメータ的基準） - 特定の値または特定の可動範囲（ツールボックスで現在利用可能なすべてのパラメータを含むリスト）を持つパラメータを与えます。パラメータが選択されると、目下の化学構造が自動的に計算されます。
Profile（プロファイル） - 選択されたプロファイルの特定のカテゴリ（ツールボックスで現在利用可能なすべてのカテゴリを含むリスト）を与えます。
Structural（化学構造的） – フラグメント中心の原子に基づいた特定の類似性を与えます。

簡単に言えば、ツールボックス 4.1 の代謝のグルーピングで以下のものをもつ類似物質を見つけることができます（図 3）。

- 既定のプロファイルを持つ親物質とその代謝物（1）
- 既定のプロファイルをもつ代謝物（2）
- 完全一致する代謝物（3）
- 既定のパラメータ値をもつ代謝物（4）
- 既定のものに類似した代謝物（5）
- 既定のプロファイル、パラメータ値や構造類似性を持つ親物質（6）
- 上記の組み合わせ

さらに、すべて、または少なくとも 1 つの既定の基準を満たしているかどうかを選択することができます。（7）
Alert performance（アラート性能）（8）のセクションでは、 私たちが標的エンドポイント（2）を決定しているかどうかが表示されます。

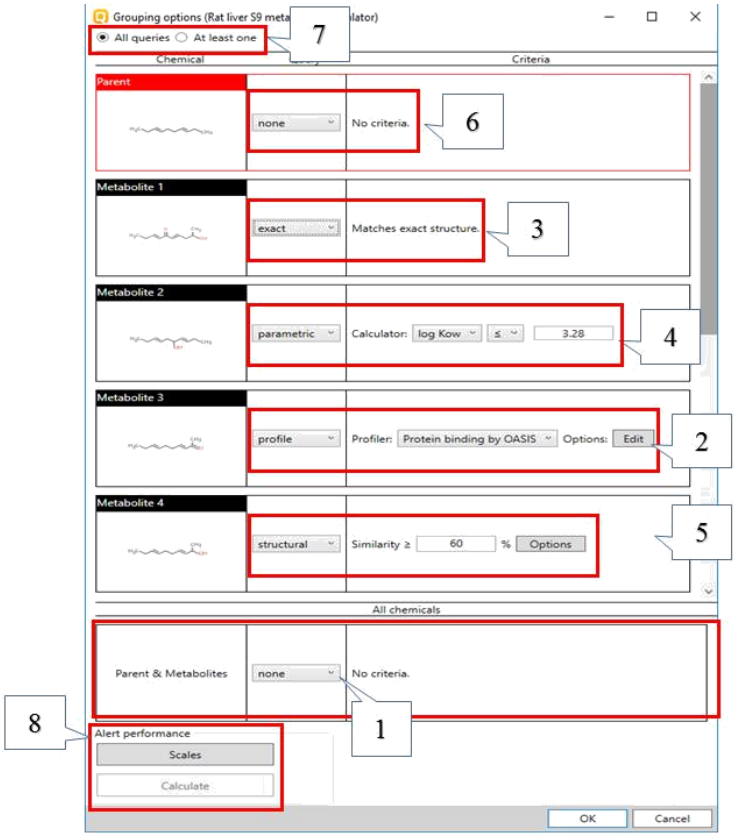


図 3

アラート性能は、他の既定クエリと同様、選択されたプロファイリングスキームに基づいて計算することができます。たとえば、parametric クエリで（MW < 168 Da）（1）が選択されている場合、**Calculate** (2)をクリックすると、その表示（3）は、この要件を満たす代謝物を作り出す化学物質に基づいたものになります。（図 4）

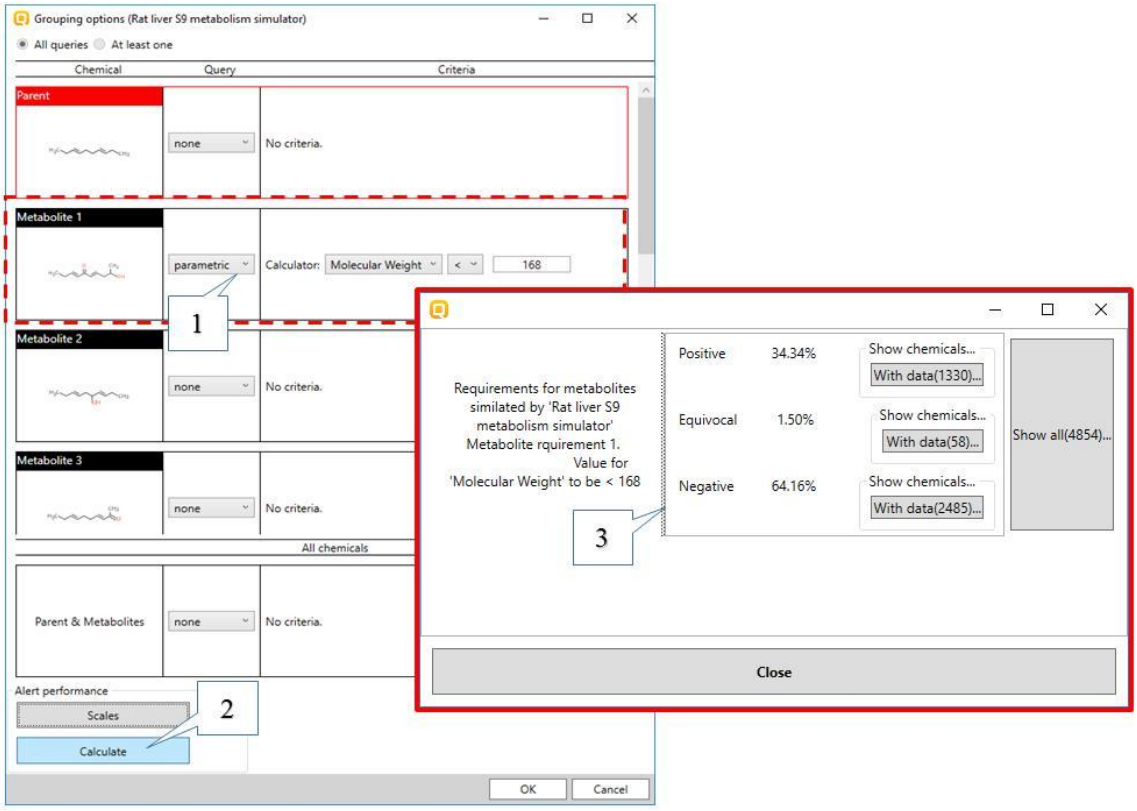


図 4

D.4.3.5. アラート性能

（訳注：このセクションの手順で同じ有害性を持つ化学物質を抽出し、ひとつの集団として D.5 章以降に用いることが可能であるということ）

アラートは、多くのプロファイリングスキーム（プロファイラ）のメインのカテゴリ構築ユニットであり、理論的知識と経験的な観測に基づいてアラート付けを行います。選択されているアラートがデータギャップ補完に適しているかどうかユーザに通知するために、アラート性能を評価するためのスコアが開発されています。この点において、アラート性能の計算とは、選択されたエンドポイントに対する所定のアラートの信頼性を示しています。

アラート性能（AP）は、アラートが標的エンドポイントにどれほど関連しているかはっきりさせるために使用されています。これはカテゴリを形成するためのアラートの有効性を反映しており、すべてのプロファイラとすべてのエンドポイントに適用できます。

信頼性とは、カテゴリ形成をする際に、選択されたアラートが適しているかどうかを伝えるものです。ツールボックス内のアラートの主な用途の 1 つは、特定のエンドポイントに関する、既定の化学物質のカテゴリを絞り込むことです。アラートの信頼性は環境に依存しており、標的エンドポイント、選択したデータベース、選択したモードとスケール（判定基準）が変化すると大幅に影響を受けます。

この機能を使用するためには、標的エンドポイントが事前に設定されている必要があります。類似物質の検索に使用されている、対応するデータベースも選択してください。

profiling scheme（プロファイリングスキーム）（1）を選び、***Define***（決定）（2） をクリックすると、target structure（標的化学構造）（3）内で見つかったすべてのプロファイラのカテゴリが含まれるカテゴリ化のウィンドウが表示されます。(図 1) そして、エンドポイントが決定されている場合 (4) は、***Alert performance*** (アラート性能) を計算する追加のセクションが使用できるように設計されています。（5）

AP のセクションは、 2 つのボタンからなります。*Scales* (判定) - 別のオプションを設定できます。*Calculate* （計算）-選択された *Scale* によって AP を計算します。

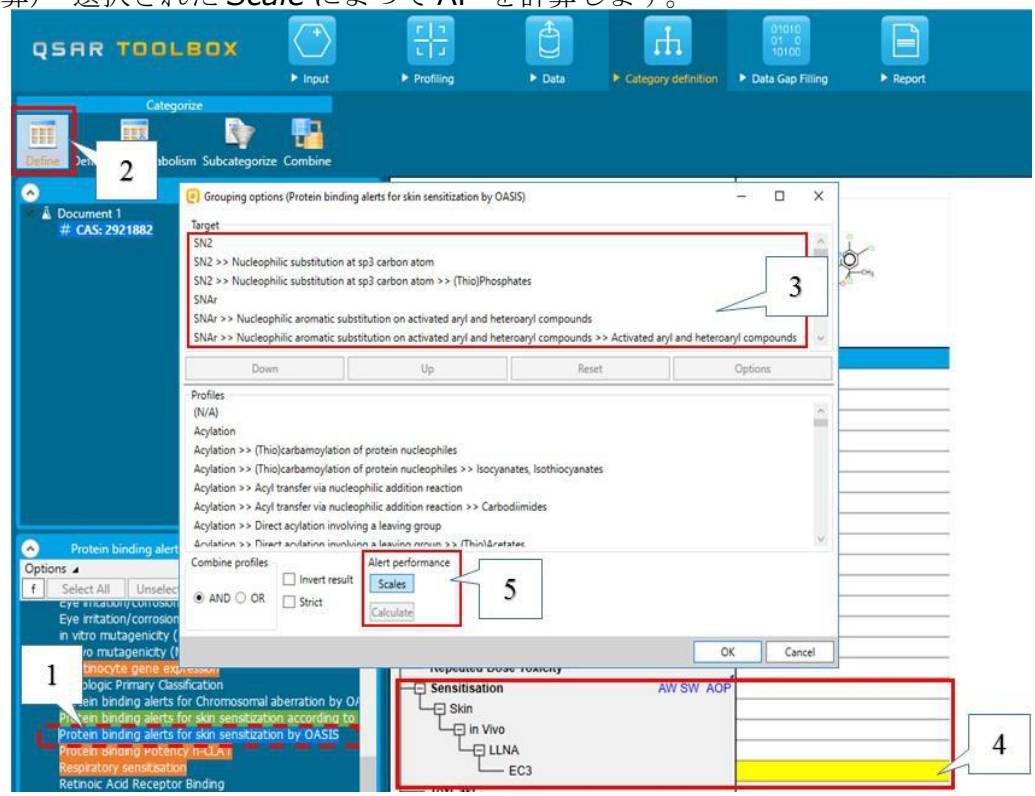


図 1

アラート性能を計算するためのさまざまなオプションは、***Aggregation options*** （集計オプション　訳注：図 1 で *Scale* ボタンを押す）内で使用可能です。アラート性能を計算するためには、特有の判定基準(1)に適用させる可能性があります(皮膚感作性 II、すなわち ECETOC を、陰性と陽性のデータに変換します)。利用可能なデータにさまざまな重み付けを適用できる追加オプションも、ドロップダウン・メニュー（2）で提供されています。ワーストケースシナリオがデフォルトとして設定されています（つまりこの場合では Maximal）（図 2)。

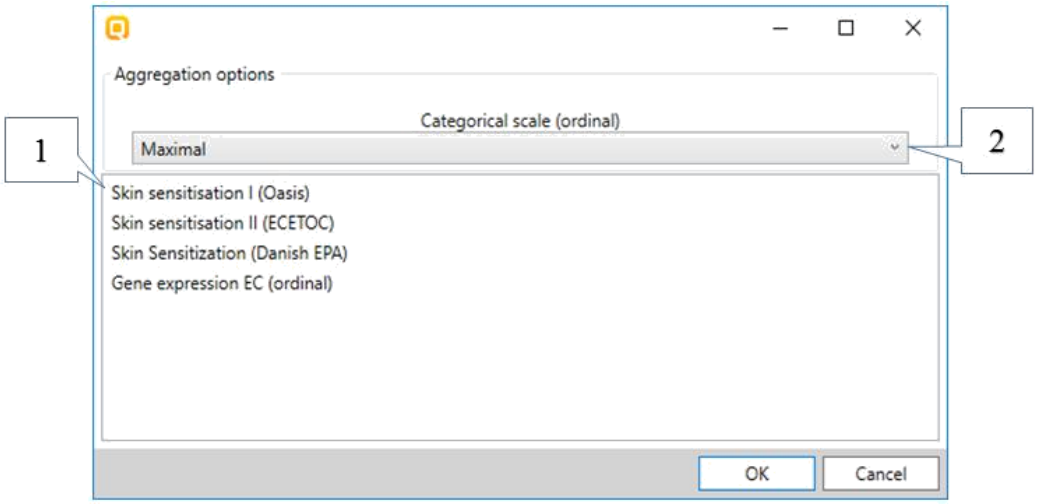


図 2

このオプションが設定されている場合、アラート性能は標的化学構造で見つかったすべてのアラート（1）も、各アラートも、別々に計算します。（(1）の AP のうち、前者の AP は（2)、 後者の AP は（3))（図 3）

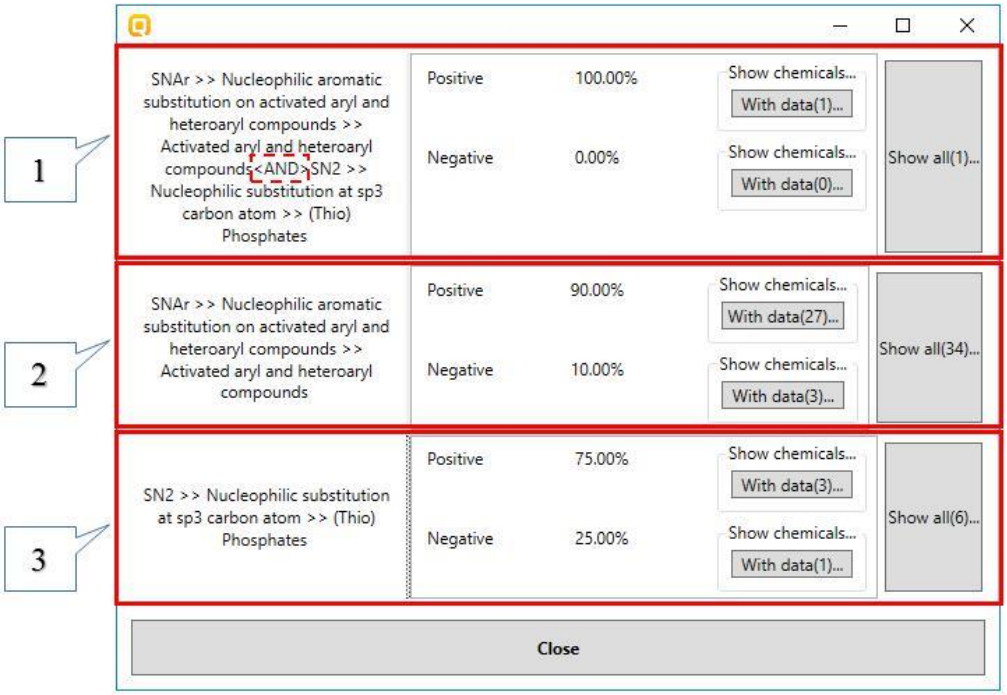


図 3

Show all（すべてを表示）ボタン を押すと、対応するアラートを持つ、選択されたデータベースにあるすべての化学物質の数が表示されます。化学物質はこのボタンをクリックすると見られます。（訳注：実測値がないものも含まれる）

Show chemicals...（化学物質を表示） ボタンを押すと、既定の判定基準に関連した特定のデータを持つ化学物質の数が表示されます。（訳注：アラートに関する試験結果が陽性や陰性といった、数値ではない実測値を持つものの数が表示される）

アラート性能は、特定のデータを持つ化学物質と、既定の判定基準に関連するデータを持つすべての化学物質との関係を表しています。

例：1 つめのアラート（SNAr >> 活性化アリールおよびヘテロアリール化合物上の求核芳香族置換>> 活性化アリールおよびヘテロアリール化合物）の アラート性能情報は、次の情報（図 4）を提供しています。

- 同じアラートをもつ 34 の化学物質が、選択したデータベースにあります。(1)
- 34 のうち、30 (27+3)の化学物質は、事前に既定の判定基準（2）に関連する実験データを備えています。
- 30 のうち 27 の化学物質（90 %）は陽性判定のデータを持ちます（3)。
- 30 のうち 3 の化学物質（10 %）は陰性判定のデータを持ちます（4)。

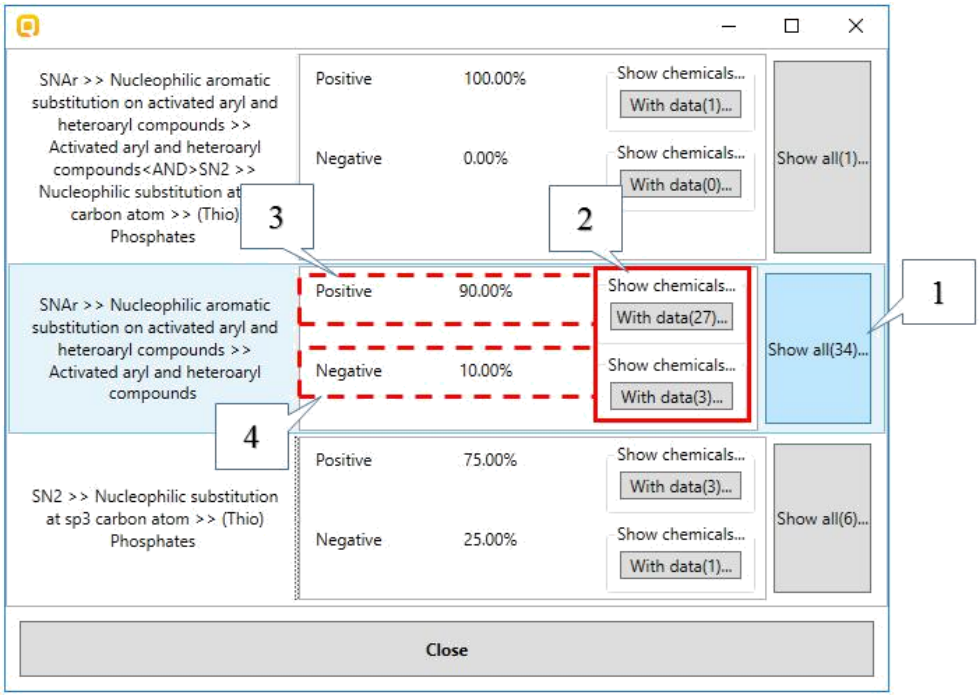


図 4

代謝を考慮したアラート性能も可能です。

D.4.4. サブカテゴリ化

広範なカテゴリを絞り込み、構造的・反応メカニズム的に類似した類似物質のカテゴリを決定するという第 2 のステップとして、サブカテゴリ化の手順を示します。反応メカニズム的にも確実に類似していると言える物質の中だけで、類似物質アプローチを行えます。特定の反応メカニズムや作用機序によって類似物質を特定した場合、標的化学物質と類似物質は既に、同一の関連した反応メカニズムと作用機序を有しているかもしれません。しかしながら、類似物質は、その分子中に付加官能基を持つため、さらなる反応メカニズムや作用機序を持っていることがあります。このため、サブカテゴリ化の手順を適用してカテゴリを絞り込みます（類似しない化学構造を排除する）。

Category definition モジュールでサブカテゴリ化をすると、Data Gap Filling モジュールに進む前に、より広いカテゴリを絞り込むことができます。たとえば 242 個のエステルが標的化学物質として US-EPA カテゴリで識別されています (図 1-1)。この点で、サブカテゴリ化手順を使用すると、これらの 242 化学物質すべてが同じタンパク質結合反応メカニズムを有するかどうか確認できます。これを行うには、**Subcategorize** ボタン（2）をクリックし、Subcategorization パネルでタンパク質結合プロファイラー（3）の 1 つを選択してください。Subcategorization パネルには、プロファイリングのセクションにリストされているのと同じ、グルーピング/プロファイリング法を持つリストが含まれています。サブカテゴリ化の手順では、代謝を考慮することができます。

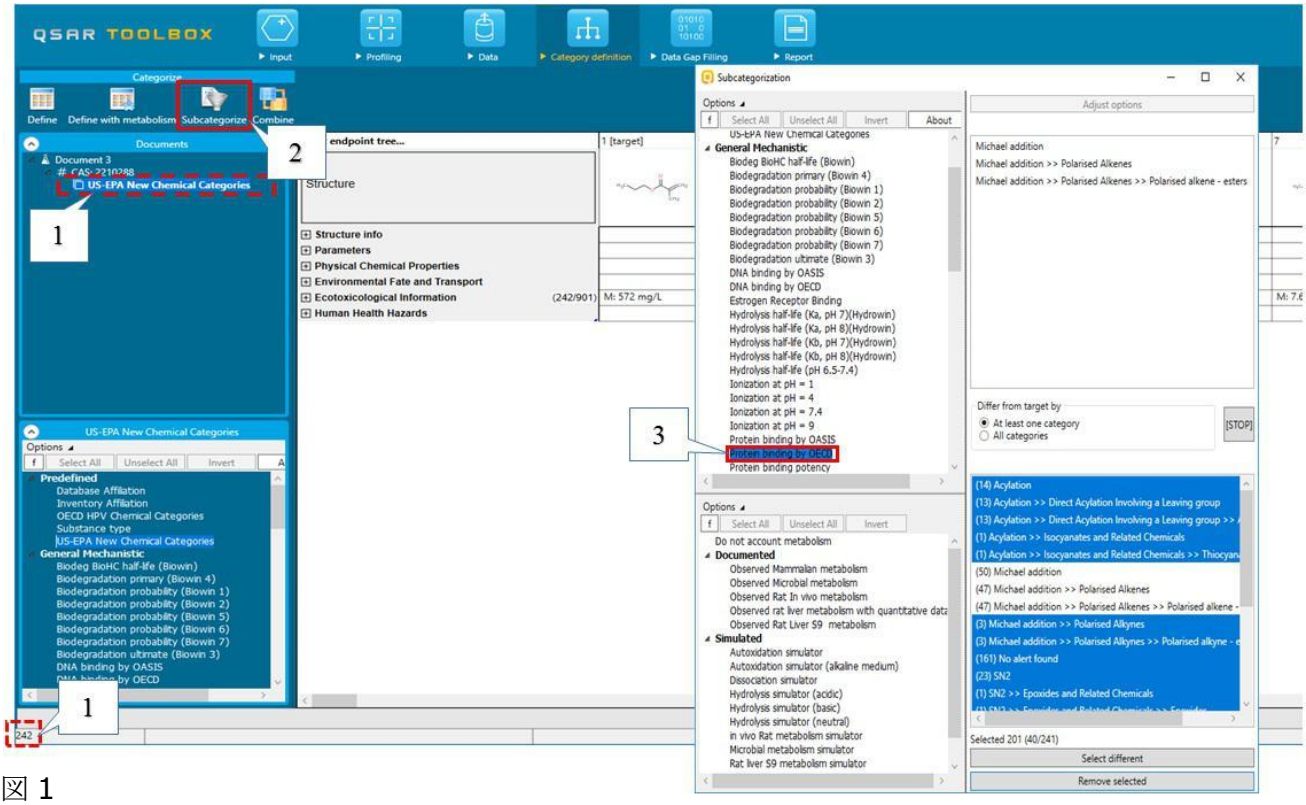


図 1

標的化学物質とは異なる相互作用のメカニズムを有する化学物質 / 類似物質は、青色の背景で強調表示されています。強調表示されていない他のものは、標的物質と同じカテゴリを少なくとも一カ所持っています。

all categories ラジオボタンが選択されている場合（図 2-1）、絞り込まれたカテゴリに含まれるすべての類似物質が、論理結合（AND）(2) で結合されているすべてのカテゴリを持っているということです。**Remove selected** ボタン（3）をクリックし、類似しない類似物質を除いた後、このソフトウェアは親カテゴリのサブカテゴリである新しいカテゴリを決定します。

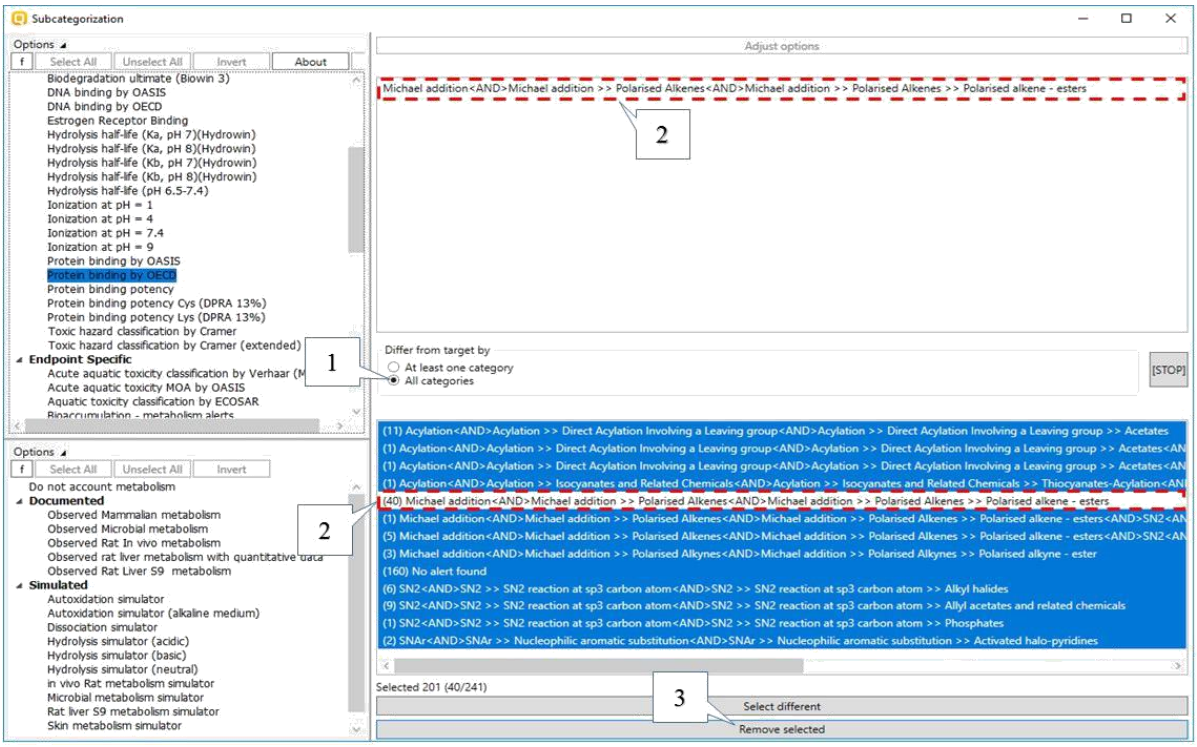


図 2

サブカテゴリは、第 1 のカテゴリのサブノード（1）として表示されます。新しいデータ・マトリックスは、サブカテゴリが作成された後に残る化学物質から構成されます（2）。(図 3)

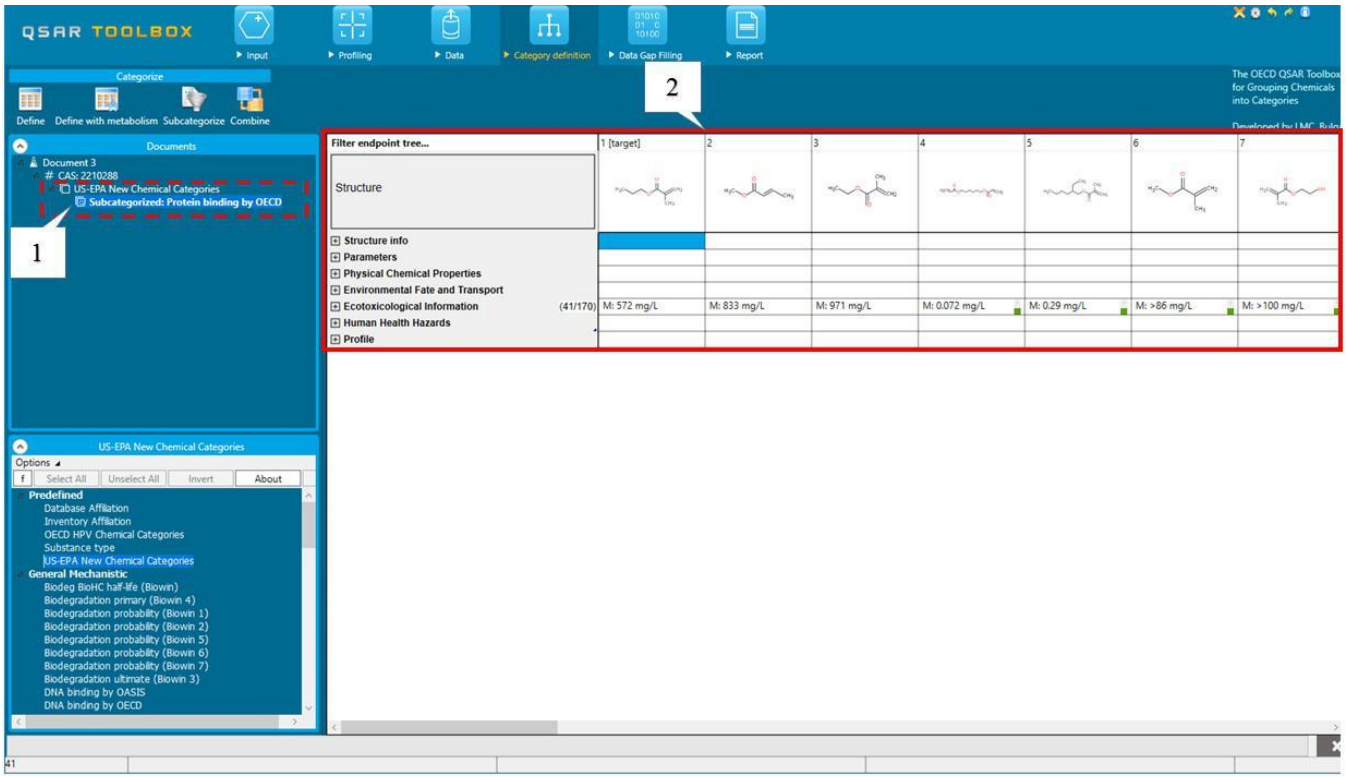


図 3

D.4.5. カテゴリを結合させる

カテゴリ結合とは、既存の化学物質リストから論理的に構成要素を結合することによって新しいカテゴリを決定することです。Document ツリー（1）に複数の化学物質リストがある場合、ユーザは **Combine** ボタン（2）をクリックしてそれらを合体させることができます。使用可能なすべてのリストを含む Combine categories ウィンドウが表示されます。一部またはすべての化学物質のリストを選択し、化学物質を論理操作 AND / OR（4）で組み合わせることができます。OK（5）をクリックすると、新しいカテゴリの作成が完了します（図 1）。

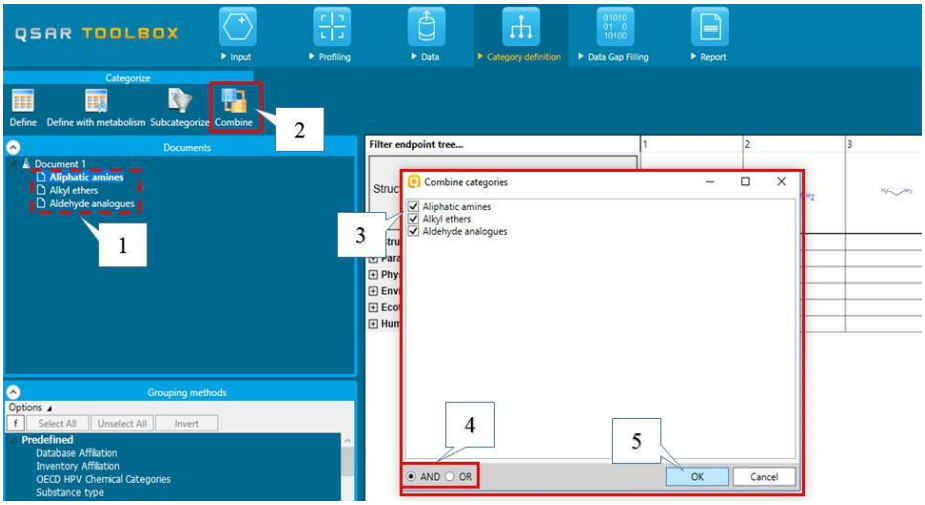


図 1

組合せの論理操作（AND / OR 操作）の選択後、ソフトウェアは、既定のカテゴリ（1）の基準に合う化学物質を含む新しいカテゴリを決定します。（図 2）

- AND を選択すると、新しい化学物質リストには、選択したリストに共通する化学物質のみが含まれます。 この場合（図 2）、選択されたすべてのリストに存在する化学物質は存在しません（図 1）。
- OR を選択すると、新しい化学物質リストには、選択したリストに表示されているすべての固有の化学物質が含まれます。

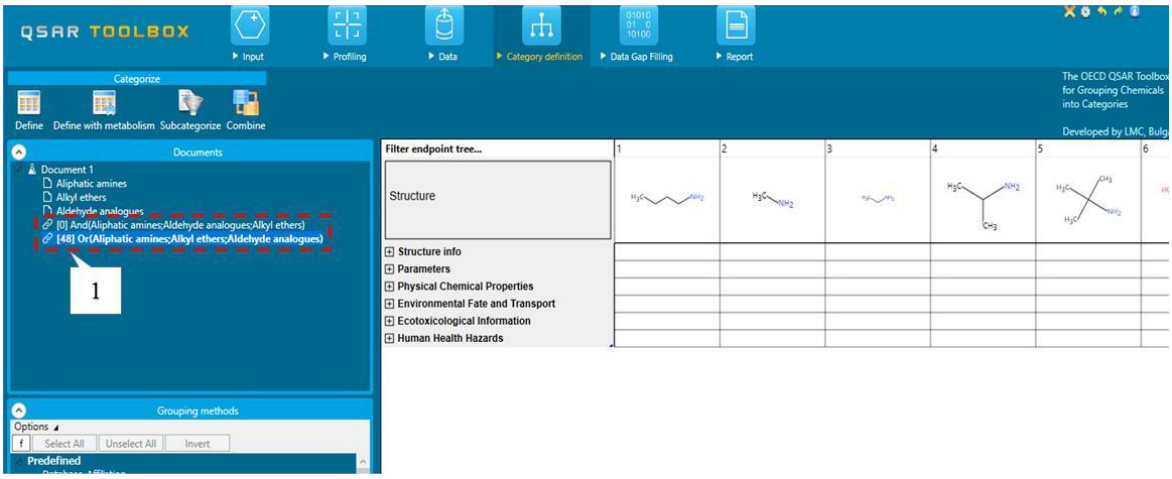


図 2

D.5.　Data Gap Filling（データギャップ補完）

モジュールの目的

Data gap filling モジュールの目的とは 3 種類の様々なデータギャップ補完ツールへのアクセスを提供することです。

- リード・アクロス
- トレンドアナリシス
- （Q）SAR モデル

これらに加えて 2 つのワークフローが開発されました。
- Standardized workflow（標準ワークフロー）
- Automated workflow（自動ワークフロー）

ワークフローの目的は、迅速かつ信頼性の高い結果を得るために、最良の順序でステップを選択することを通じてユーザをナビゲートすることです。

- リード・アクロス
- トレンドアナリシス
- （Q）SAR モデル

リード・アクロスとトレンドアナリシスでは、データ・マトリックス内の利用可能な実験データを使用してデータギャップを補完します。「(Q) SAR モデル」は、ツールボックスに統合された外部（Q）SAR モデルのライブラリへのアクセスを提供します。

最も関連性の高いデータギャップ反応機構を選択することについては、次の考慮事項に依存します。

- リード・アクロスは、きまった数の結果（例えば、陽性、陰性、不明確）をとりうる皮膚感作性または変異原性のような「定性」エンドポイントのための、適切なデータギャップ補完法です。それに加えて、実験結果を有する類似物質の数が少ない場合の「定量的エンドポイント」（例えば、魚類の 96h-LC50）にはリード・アクロスが推奨されます。
- トレンドアナリシスは、実験結果を有する多数の類似物質が同定された場合、「定量的エンドポイント」（例えば、魚類の 96h -LC50）のための適切なデータギャップ充填方法です。
- 標的化学物質に適切な類似物質が見つからない場合、「(Q) SAR モデル」を使用すると、データギャップを補完することができます。

リード・アクロスとトレンドアナリシスを選択すると、データ・マトリックス内の利用可能なデータがデータギャップを補完するために使用されます。ユーザは、プロファイラを使用して標的化学物質と比較し、異なるプロファイルを有する化学物質を排除することによって、データセットをさらに減らすことができます。

2 つのシナリオが区別されます。

- 標的化学物質について、そのエンドポイントに関連する特定の反応機構または作用機序が特定される場合、考慮される類似物質もすべて同じ反応機構または作用機序を有していなくてはなりません。
- 標的化学物質について、そのエンドポイントに関連する特定の反応機構または作用機序が特定されない場合、考慮される類似物質は同じ反応機構または作用機序を有していません。

Caution

Categorical type data such as data for skin sensitization or Ames mutageicity endpoints are calculated using read across or QSAR methods.

注意：Skin sentization や Ames mutagenicity といったエンドポイントはカテゴリ的データ（訳注：数値ではなく陽性陰性などのデータ）であり、リード・アクロスか QSAR 法で計算される。

（訳注：先にデータ・マトリックスに類似物質が並んでいる必要があるため、次の作業を参考に事前準備をされたい。Input モジュールの Cas#アイコンを押して、122-04-3 と入力して塩化 4-ニトロベンゾイルを標的化学物質とする。そして Data モジュールで全ての Database を選択して左上の Gather アイコンを押す。次に、Category definition モジュールで、例えば US-EPA New Chemical Categories などある程度大きなカテゴリをクリックし、続けて左上の Define アイコンを押し、現れた Option の小画面で OK を押すとデータ・マトリックスに類似物質が並び、物質数を小画面で報告される。これに OK を押すと「Read Data?」という小画面が現れるので OK を押すと、エンドポイントのデータがマトリックスに入る。この状態が図 1 に相当しており、以下の本文はこの状態からスタートしていると考えられる）

Category definition（カテゴリ決定）セクションで説明したように　、ツールボックスは、標的化学物質と同様の構造的機能を有し、実験結果が得られる化学物質を特定します。 塩化 4-ニトロベンゾイル（CAS No 122-04-3）を例として図 1 に示します。皮膚感作性は「定性的（訳注：具体的数値ではない、陽性や陰性といったデータのこと）」なエンドポイントであるため、データギャップはリード・アクロスを使って補完することができます。

補完すべきデータギャップに対応するマトリックス内のセル（1）を強調表示した後、Gap Filling 方法（2）（図 1）を選択してください。また、 2 つのワークフローオプションが含まれています。Standardized（SW、標準化）と Automated（AW、自動化）です（3）。これらは自動化されたデータギャップ補完のためのアルゴリズムを表しています。これらのアルゴリズムは、皮膚感作性（ LLNA と GPMT データ）および魚に対する急性水生毒性（Pimephales promelas、Mortality、LC50、96 時間）のために開発されています。いったん開始されると、AW（自動化されたワークフロー）は装備されたロジックに従い、ユーザとやりとりすることなく予測が完了します。データギャップ補完の標準化を実施するにあたって、同じアルゴリズムが実装されています。AW と比較した主な違いは、アプリケーションのドメインが SW（他の生物種、試験実施期間などを含む）内では拡張されていることと、SW はユーザとの相互作用を可能にし、したがって AW で行えるものとは異なる選択が可能であることです。

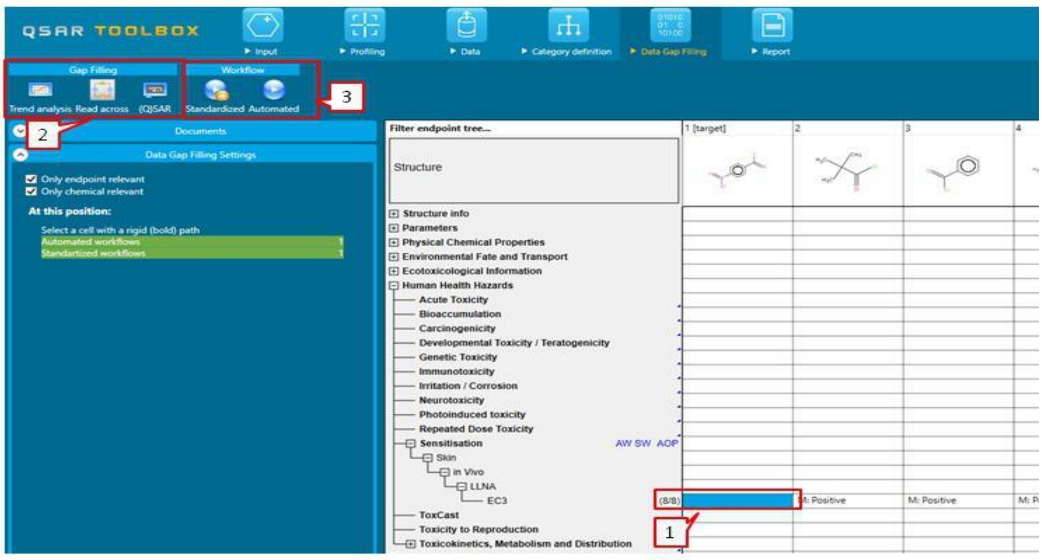
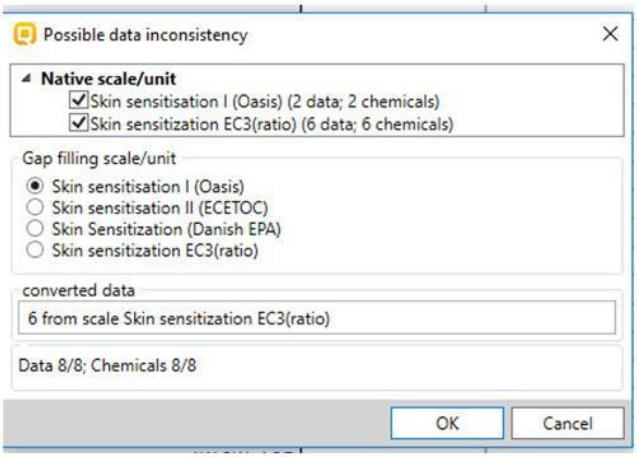


図 1

データギャップ補完ウィンドウに入る前に **possible data inconsistency** ウィンドウが現れます（図 2-1）（訳注：図 1-1 のセルをクリックで青くしてから左上の Trend analysis や Read-across のアイコンをクリックしたときに出る小画面。なお、1 つでもデータ・マトリックス内にデータが無い試験の場合は Gap Filling はできない。）

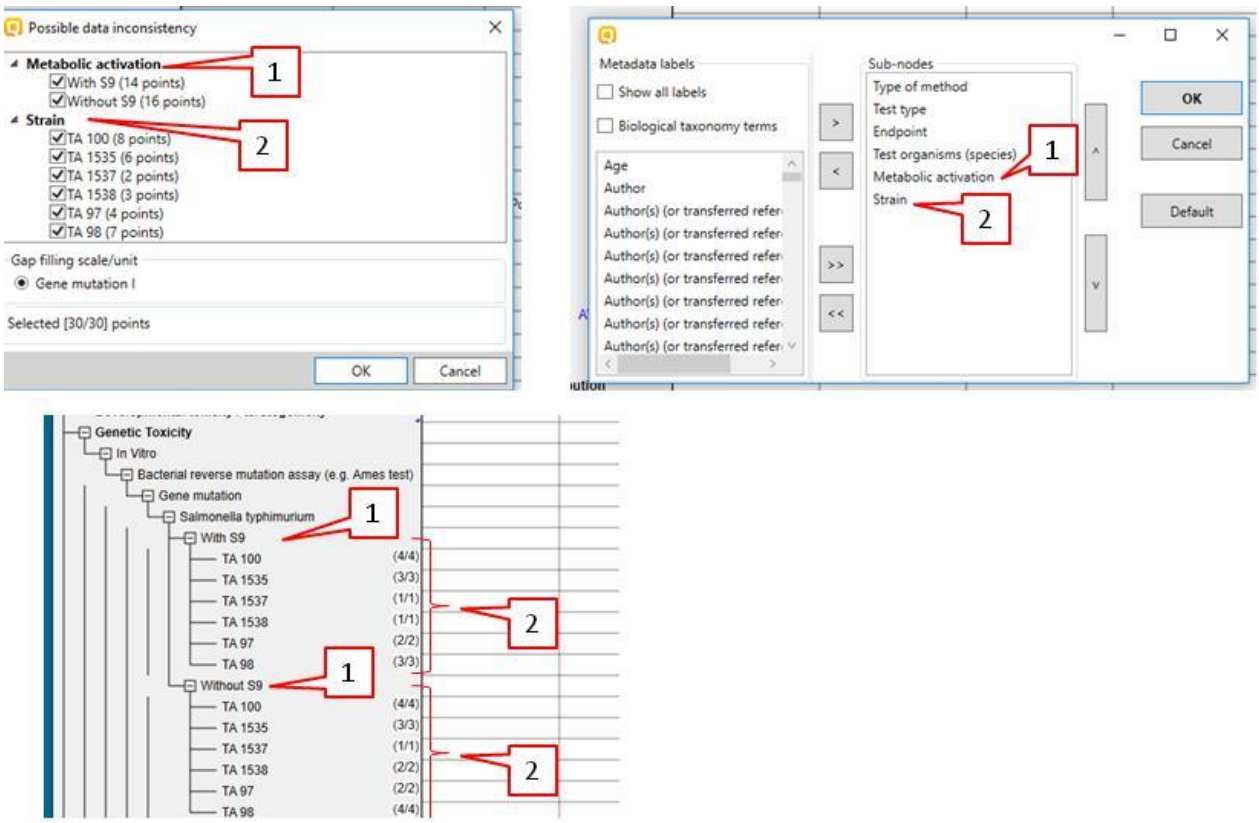
図 2



この機能はデータの不一致をユーザに通知します。

図 3 に示す例では、Salmonella typhimurium（サルモネラ・ティフィムリウム）のデータに関連し、不一致である可能性のあるデータは、**Metabolic activation（代謝活性化）** (1) と **Strain**（系統種） (2)。これらの 2 つの付随データ欄に含まれるデータは、データギャップ補完において混在しています。

図 3



スケール（判定基準）の詳細については、About の項目に記載されています。

Note

Only one scale/unit is allowed in data gap filling.

［注意：Scale（判定基準）と Unit（単位）はデータギャップ補完では一種類しか使えません］

なお図 1 のウインドウ下部にある数字（例えば、図中選択された 8/8）は、合計 8 個のデータポイントから 8 個のデータポイントがデータギャップ充填モジュールに入ることを意味します。

D.5.3. データギャップ補完の画面

次のスナップショットはさまざまなタイプのデータギャップ補完を示しています。

- Read across(リード・アクロス) ウィンドウ

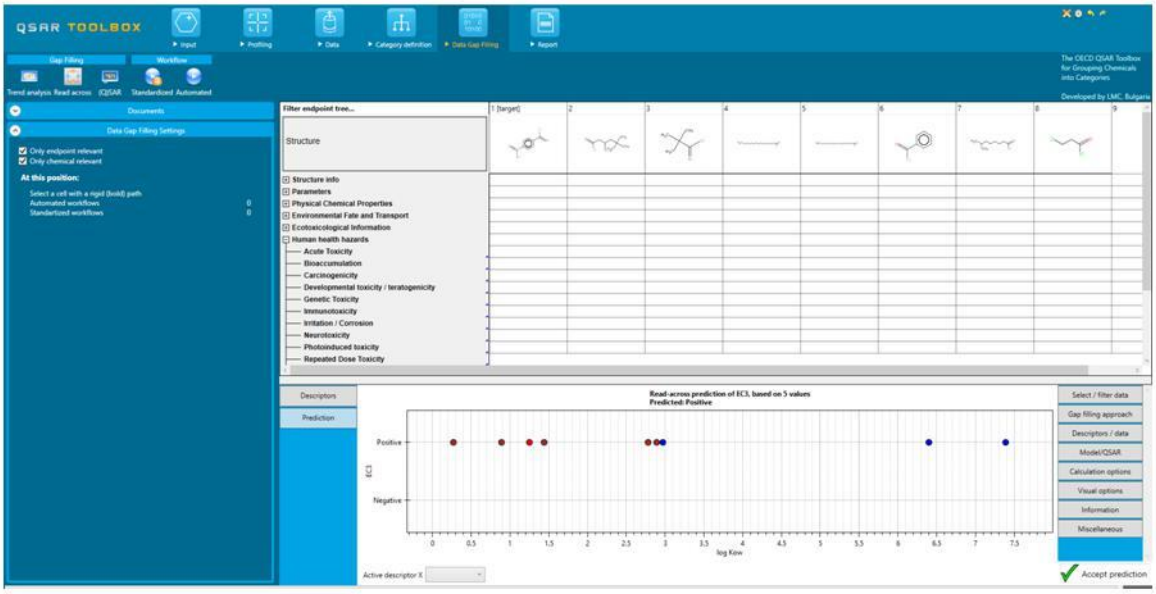


図 1

- Trend analysis(トレンドアナリシス) ウィンドウ ：

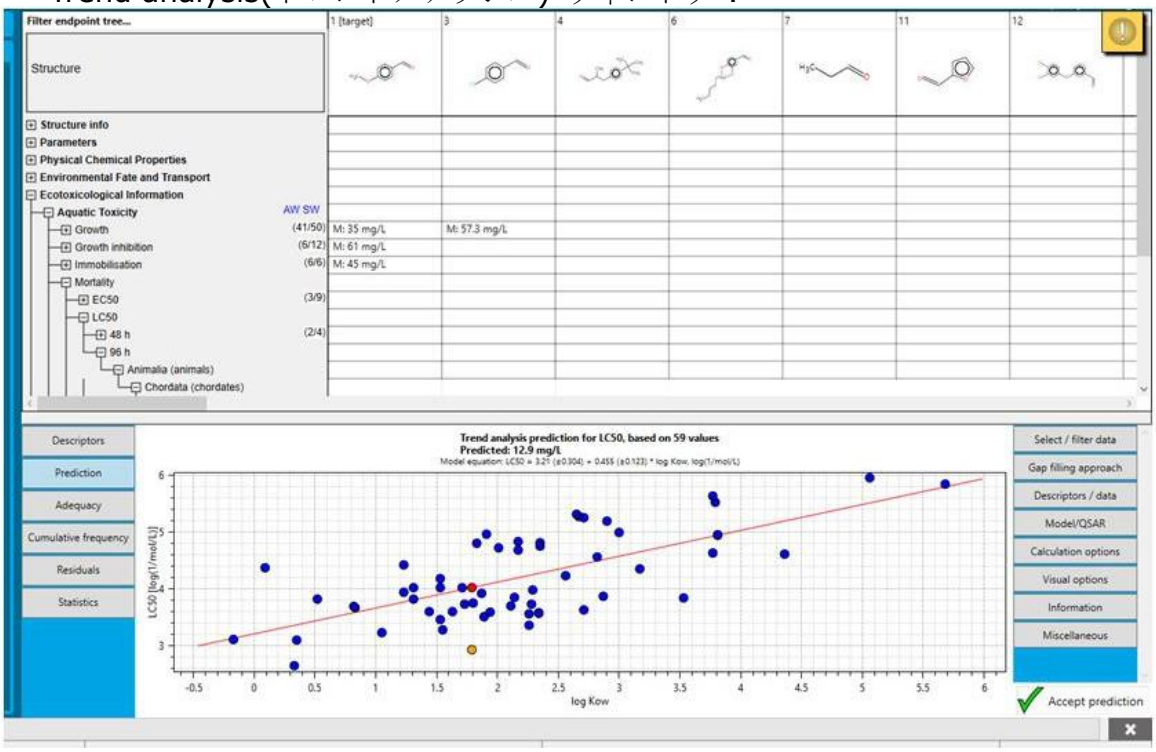


図 2

- （Q）SAR models ウィンドウ

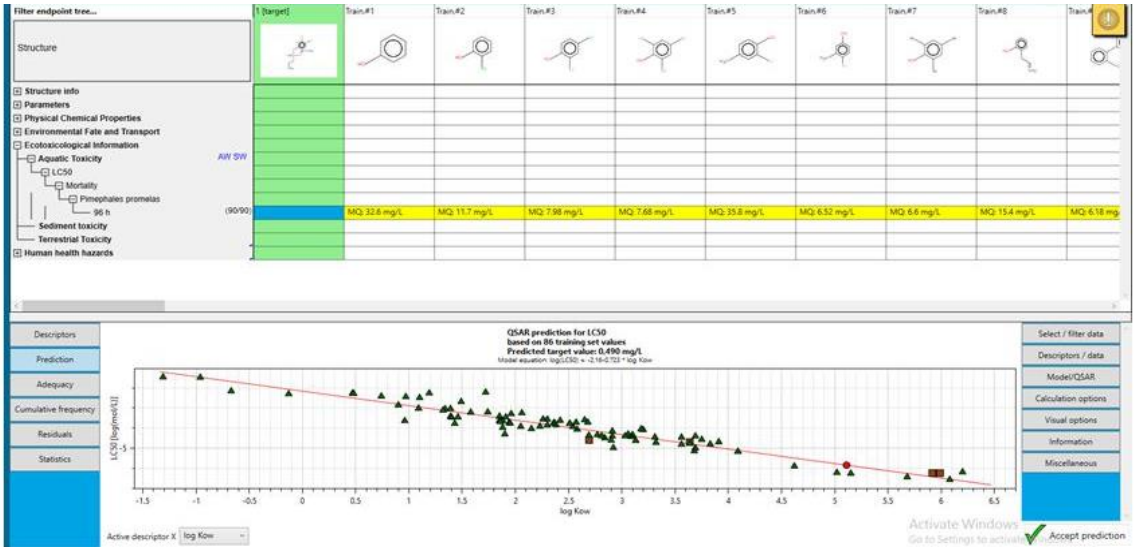


図 3

- Automated workflow（AW,自動ワークフロー）

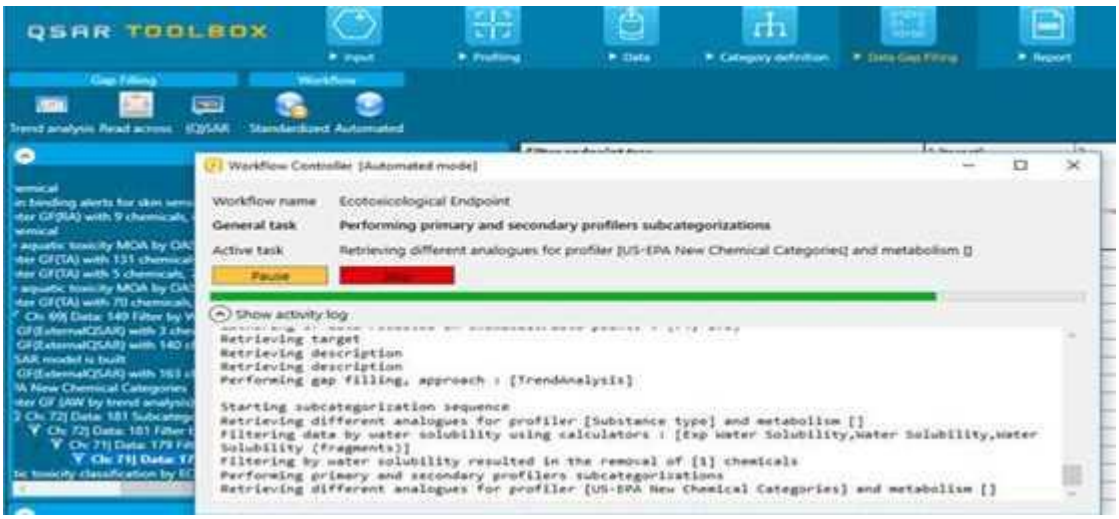


図 4

- Standard workflow(SW, 標準ワークフロー)

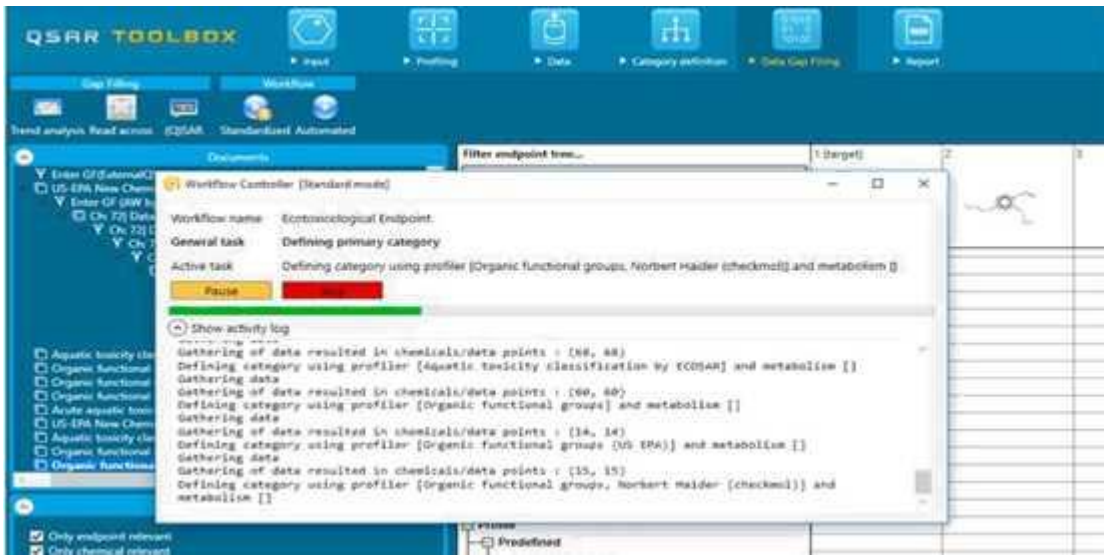


図 5

D.5.4. 3 つのギャップ補完アプローチに共通の機能

3 つの各ギャップ補完アプローチは、それぞれ異なっているものの、共通した機能を持っておりすべて共有されています。

（訳注：以下の共通機能は、データギャップ補完が済み、データ・マトリックスの下にグラフが出現したときに、グラフの中に表示される）

- 左サイドパネル
- データギャップ補完のメニュー
- 右クリックの機能

D.5.4.1. 左サイドパネル

左サイドパネルが（図 1）に示されています。

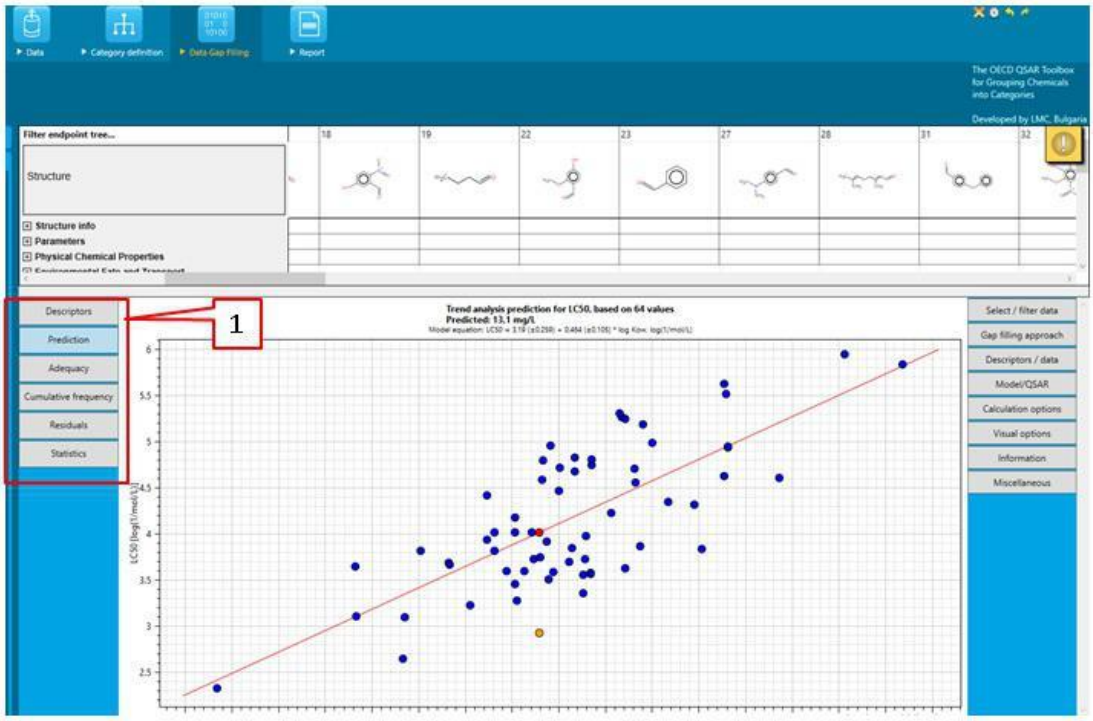


図 1

それには 6 つのセクションが含まれています。

- Descriptors（訳注：ここでは予測グラフの横軸である X 軸を選択する機能）
- Prediction（訳注：Y 軸をエンドポイント、X 軸を Descriptors で選択した指標で作ったグラフ）
- Adequacy（妥当性）
- Cumulative frequency（累積頻度）
- Residuals（残差）
- Statistics（統計情報）

D.5.4.1.1.　Descriptors（予測グラフの X 軸と Y 軸）

Descriptors パネルが図 1 に示されています。

（訳注：D.5.2 章冒頭の訳注で述べた状態、すなわち D.5.2 章の図 1 の状態で、Trend analysis アイコンを押してデータギャップを補完したときにデータ・マトリックスの下に現れる画面がこの下にある図 1。Trend analysis は数値のデータを 1 つ以上もつ試験のデータギャップ補完に使う。Read-across は陽性・陰性など数値ではない結果に使える）

ここでは、グラフの X 軸に使用する記述子を選択できます。Y 軸の点は、データ・マトリックスのデータ点の値であり、最も一般的には対数スケールに設定されます。Y 軸の単位は、Options（訳注：ツールボックス全体画面右上の小さな歯車のアイコン）の Data gap filling で変更できます。

. Descriptors パネルは Read-across, Trend analysis および (Q)SAR models のデータギャップ補完で利用できます。通常、Log Kow がデフォルトのグラフ軸記述子です。Descriptor を変更するには、アクティブなグラフ軸記述子（1）をダブルクリックして削除し、下のリストから選択したグラフ軸記述子（2）をダブルクリックします。

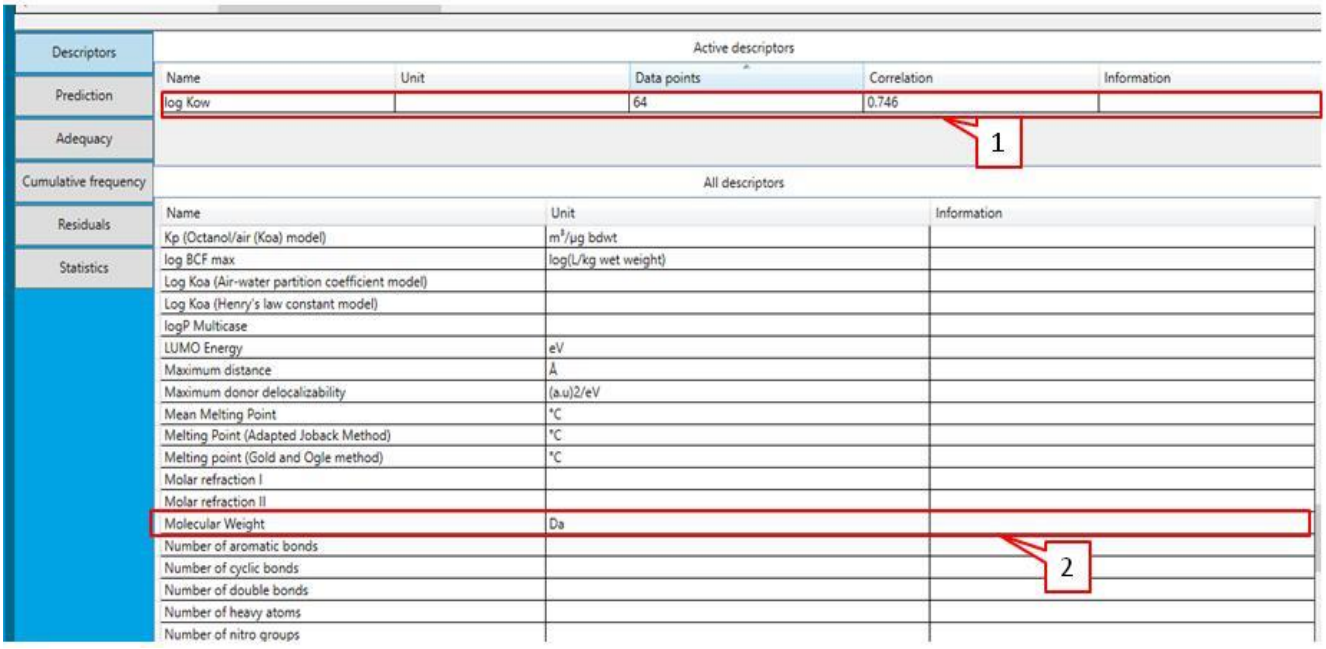


図 1

D.5.4.1.2. Prediction（予測）

図 1 に示す ***Prediction*** パネルは、3 つのデータギャップ補完アプローチで若干の違いがあるもののだいたい同じです。リード・アクロスではグラフ軸記述子に関する隣接データポイントを強調表示し（デフォルトでは 5 つの最隣接点ですが、これは Prediction アプローチオプションで変更することができます)、トレンドアナリシスでは傾向線が表示されます。どのアプローチでも、このパネルはデータ点の分布、およびグラフ上の標的化学物質の予測値を示しています。

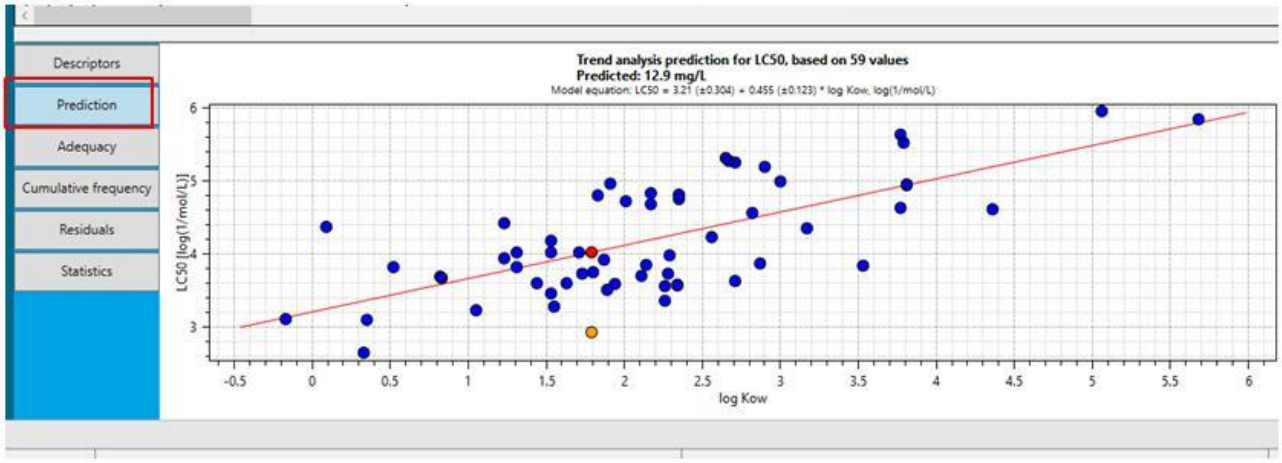


図 1

D.5.4.1.3.　Adequacy（妥当性）

図 1 の Adequacy パネルには、妥当性のグラフが入っています。これは、x 軸上の値はデータ・マトリックス上のすべての物質の実測値であり、y 軸上の値はそれらの予測値というグラフです。また、グラフの上には、決定係数（R2）と調整係数（R2 adj）（1）が表示されます。

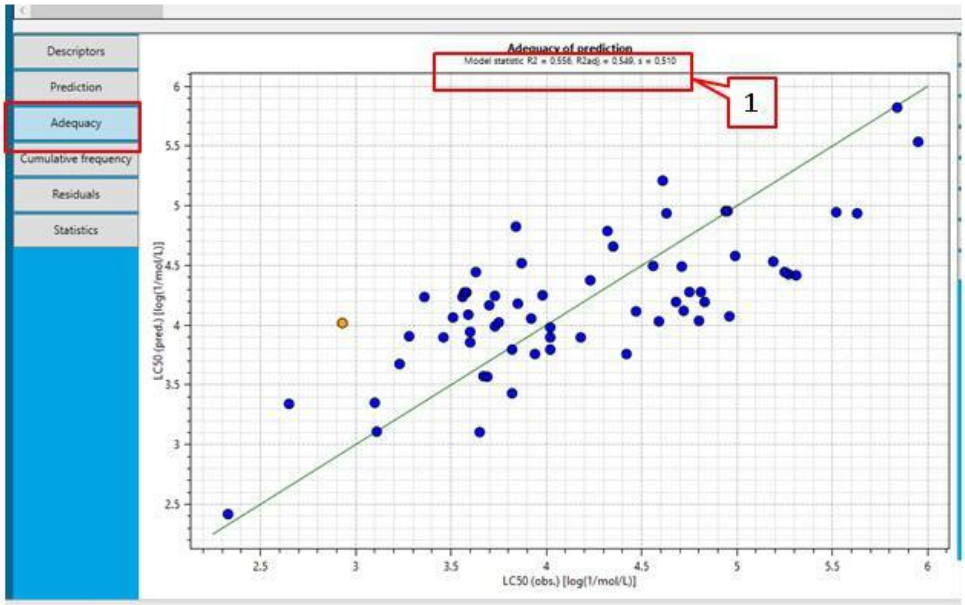


図 1

D.5.4.1.4.　Cumulative frequency（累積頻度）

累積頻度は、残差の値（実測値－計算予測値）が基準残差値以下であるという頻度です。

例えば、累積頻度が残存値 0.2 で 70%である場合、トレーニングセットの構成要素の 70%は 0.2 以下の残差を有します。以下は、累積頻度のスナップショットです。(図 5)

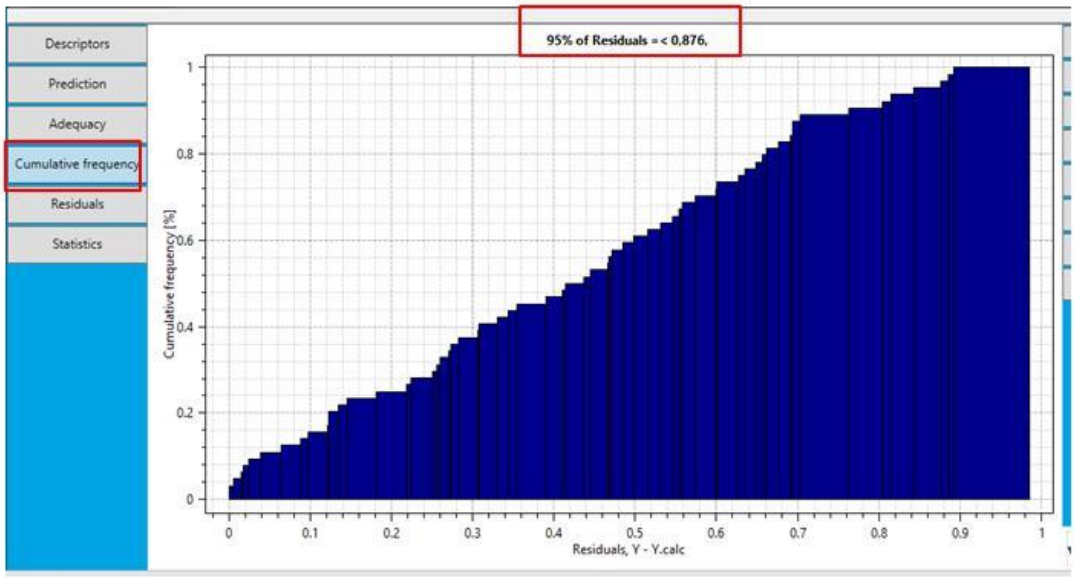


図 5

D.5.4.1.5　Residuals（残差）

このパネルには、所定のエンドポイントと所定のグラフ軸記述子の残差すべての分布が示されているグラフが入っています。(図 1)

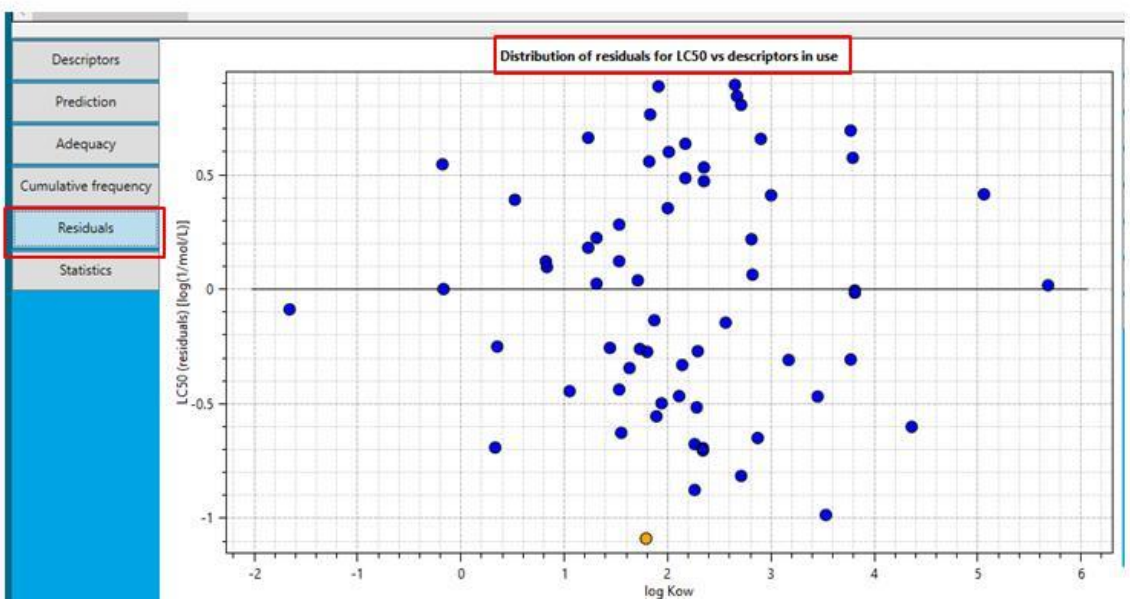
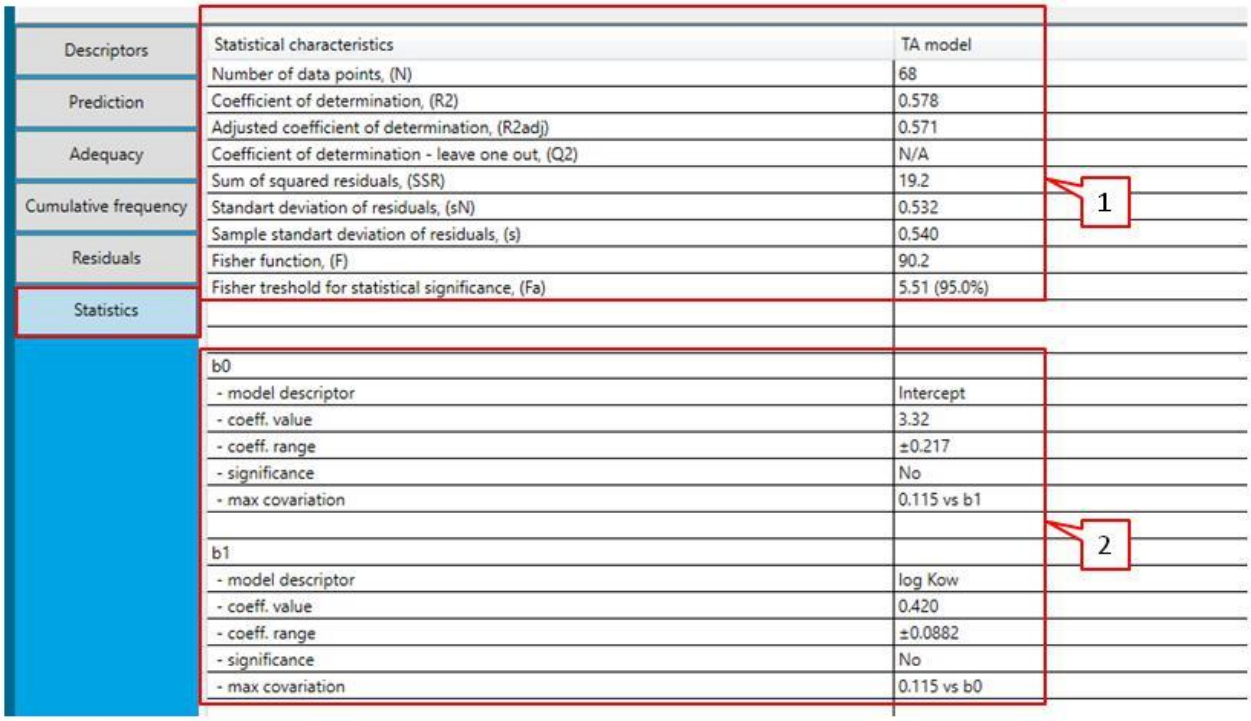


図 1

D.5.4.1.6. Statistics（統計情報）

このパネルは、 Prediction パネル（トレンドアナリシスと QSAR アプローチで利用可能）における回帰式の統計的特性を示しています（図 1)。パネルの上部には回帰式（1）の統計が含まれ、続く部分では回帰式に含まれる係数が示されています（2)。



Statistical characteristics	TA model
Number of data points, (N)	68
Coefficient of determination, (R2)	0.578
Adjusted coefficient of determination, (R2adj)	0.571
Coefficient of determination - leave one out, (Q2)	N/A
Sum of squared residuals, (SSR)	19.2
Standart deviation of residuals, (sN)	0.532
Sample standart deviation of residuals, (s)	0.540
Fisher function, (F)	90.2
Fisher treshold for statistical significance, (Fa)	5.51 (95.0%)
b0	
- model descriptor	Intercept
- coeff. value	3.32
- coeff. range	±0.217
- significance	No
- max covariation	0.115 vs b1
b1	
- model descriptor	log Kow
- coeff. value	0.420
- coeff. range	±0.0882
- significance	No
- max covariation	0.115 vs b0

図 1

D.5.4.2. データギャップ補完のメニュー

データギャップ補完のメニューは、図 1 に示されています。

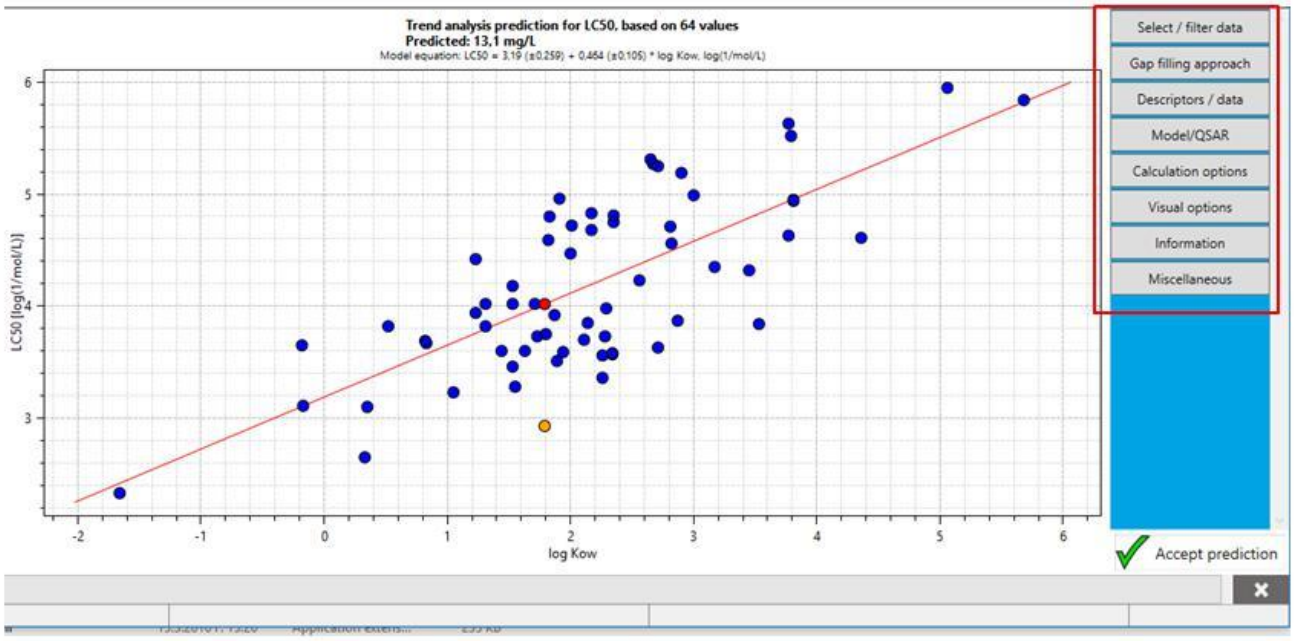


図 1 - データギャップ補完のメニュー

各オプションをクリックすると、より多くの機能を持つメニューが開きます。それらは、D.5.4.2 の中のサブセクションで説明します。

D.5.4.2.1.　Select /filter data　(選択/データのフィルタリング)

この機能には、データのフィルタリングと選択のさまざまなオプションが含まれます。(図 1)
（訳注：Select /filter data ボタンをクリックすると細かい 9 項目が並び、図 1 のようになる）

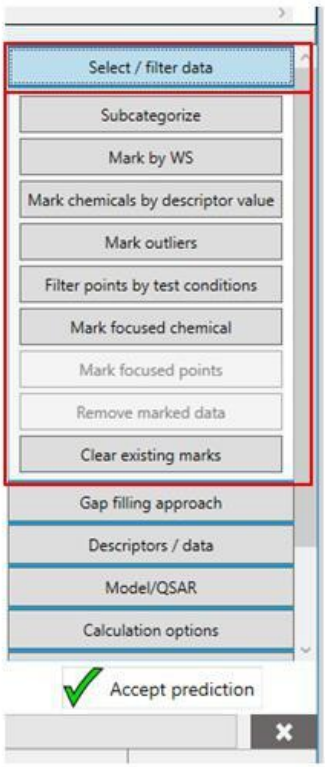


図 1

- Subcategorize（サブカテゴリ化）
- Mark by WS (water solubility)（水溶解度によるマーク）
- Mark chemical by descriptor value（軸記述子による化学物質のマーク）
- Mark outliers（外れ値のマーク）
- Filter points by test conditions（試験条件によるフィルタリングポイント）
- Mark focused chemical（化学物質にフォーカスしたマーク）
- Mark focused points（ポイントにフォーカスしたマーク）
- Remove mark data（マークしたデータを削除）
- Clear existing marks（既存のマークを消去）

D.5.4.2.1.1. Subcategorize（データギャップ補完の後で行うサブカテゴリ化）

Subcategorize （図 1）は、ユーザが利用できる最も強力なツールの 1 つです。これにより、類似物質のカテゴリが広すぎれば、より厳選し、その化学的特性と特性が標的化学物質により近いような化学物質のグループに絞り込む機能が提供されます。

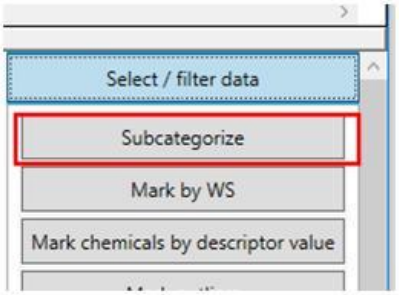


図 1

Acute toxicity（急性毒性：LC50, 96h、Pimephalis Promelas）を予測するトレンドアナリシス アプローチを例として説明します：

Subcategorize（図 1）をクリックすると、Subcategorisation ウィンドウ（図 2）が開きます。Options（1）を展開し、*Colour by* で *Target endpoint*（標的エンドポイントによる色付け）（2,3）を選択して、どのプロファイラがサブカテゴリ化に最も適しているかを見ます。*Legend*（4）をクリックすると、色の凡例（5）が表示されます。まず、これらのプロファイラを使用して類似物質のリストを絞り込むことをお勧めします（6）。

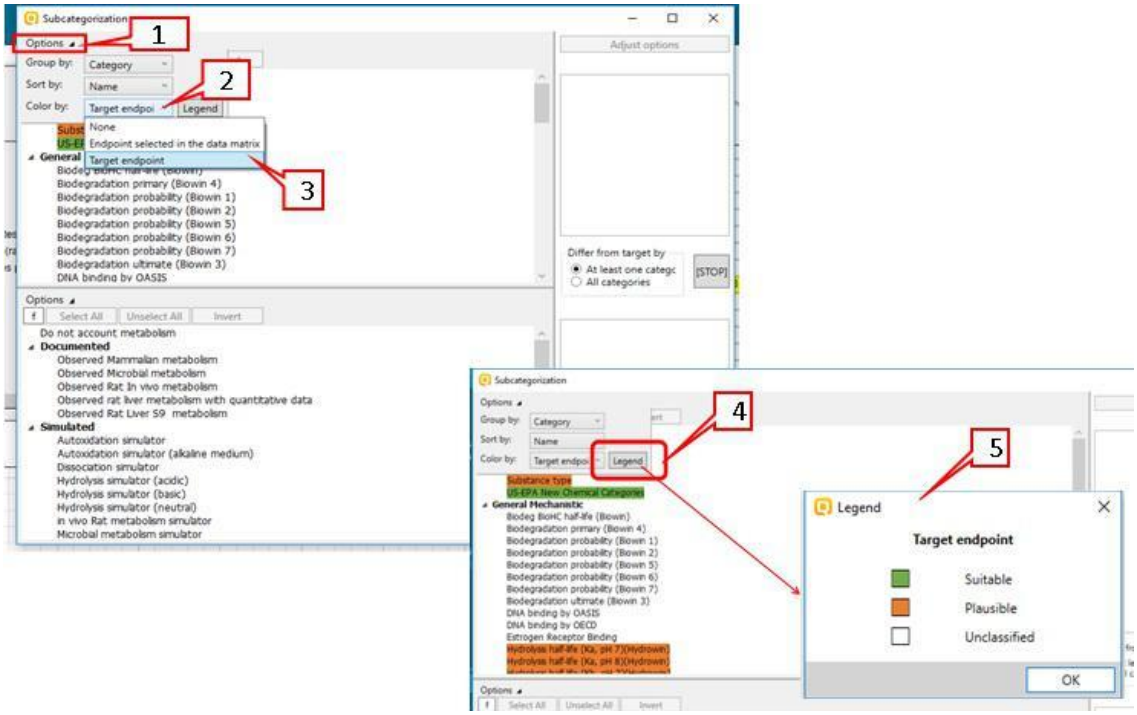


図 2

適切なすべてのプロファイラをグループ化するには、*Options*（1）を展開し、*Group by*（2）を選択してから Target endpoint （3）を選択します（図 3）。

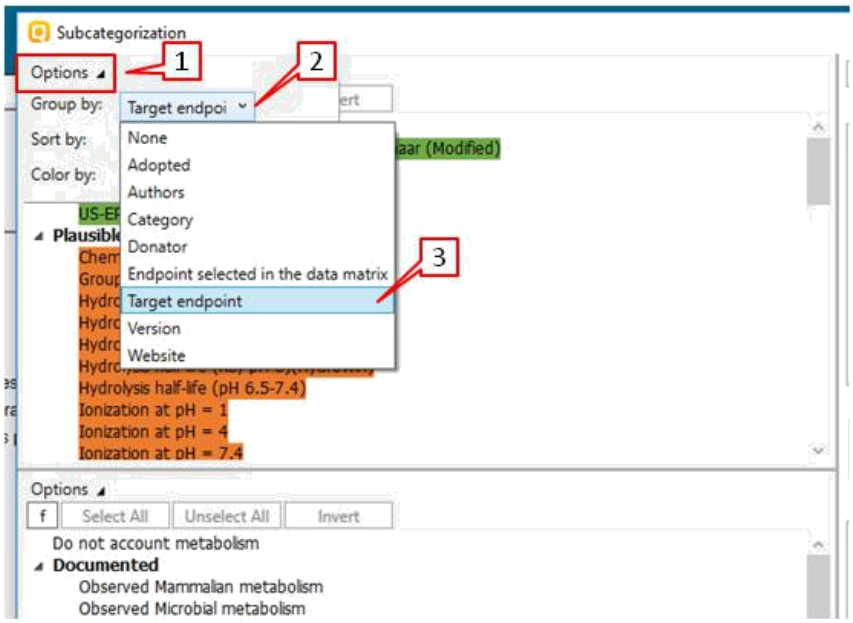


図 3

Acute aquatic classification by Verhaar (Modified) プロファイラを選択すると、標的物質とすべての類似物質を再プロファイリングします（図 4）。次に、類似物質の特性を標的化学物質の特性と比較します（3）。この例では、4 つの類似物質が Class 5 に分類され、標的（Class 3 に分類される）と区別され、異なる作用機序を示すことが予想されています（このグループは青色で着色される）。選択された類似物質の数と、比較された類似物質の総数も（3）に示されています。また、グラフ上（淡い青色の点）およびデータ・マトリックス上（緑色）（4）で色分けされています。*Remove selected*（選択対象を削除）（5） をクリックすると、これらの類似物質を削除します。

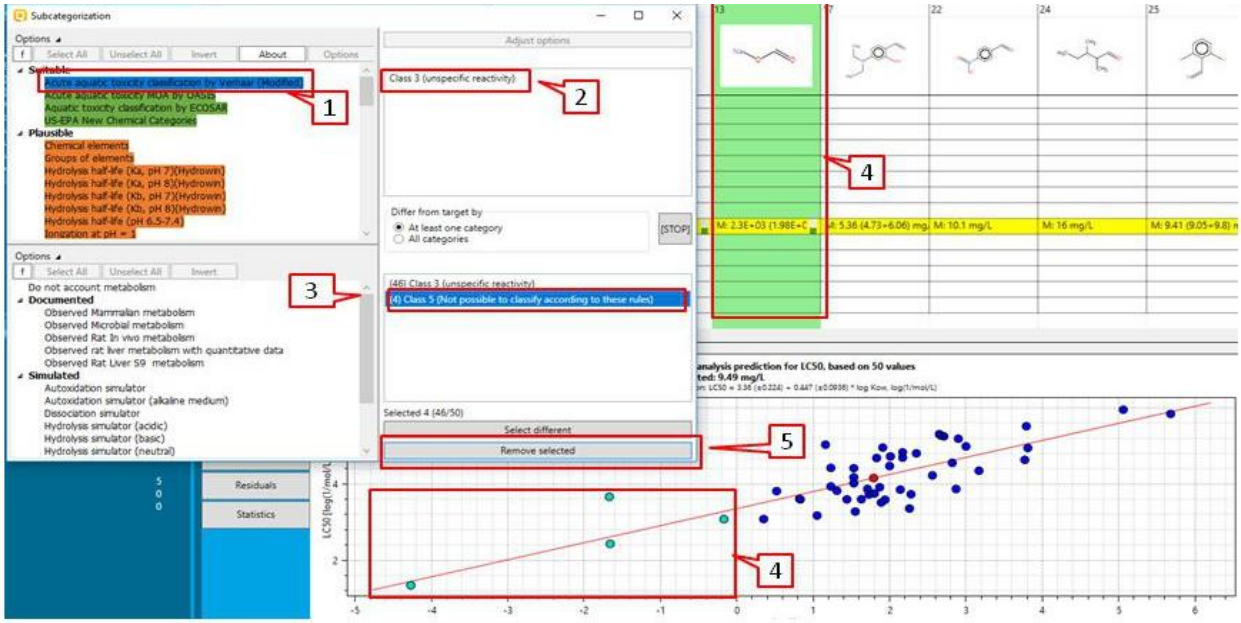


図 4

Subcategorization パネルの左側には、カテゴリの評価と絞り込みに使用できるすべてのプロファイリング/カテゴリ化方法が含まれています。エンドポイントおよび関連するサブカテゴリ化方法の詳細については、Grouping methods の項目を参照してください。プロファイリング方法（1）とともに、すべての代謝の実測値と、転換の予測値の表（2）がサブカテゴリ化パネルに含まれています。(図 5)

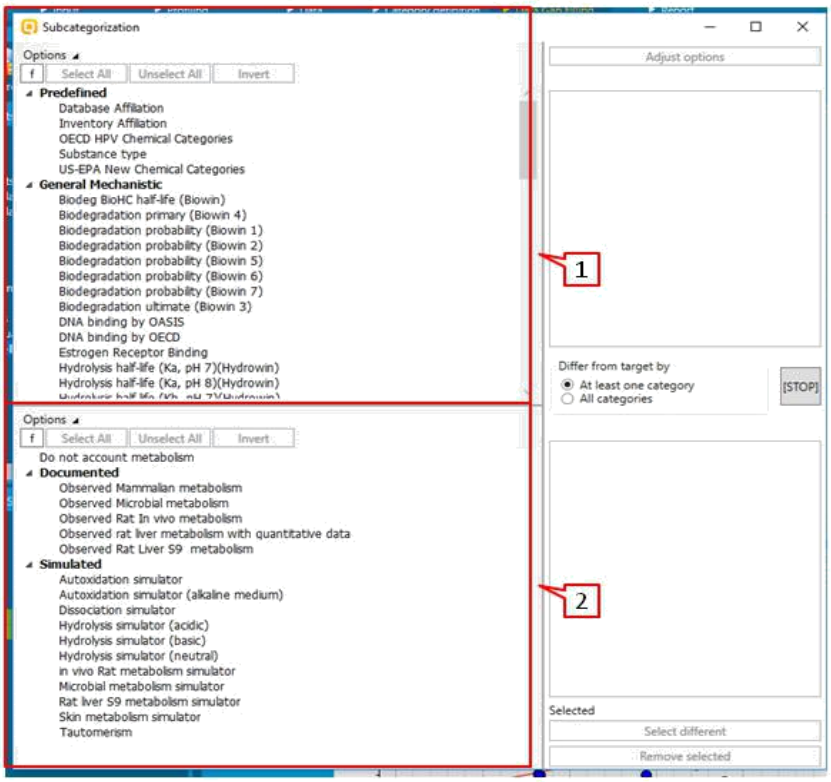


図 5

いくつかの利用可能なシミュレータ（非生物/生物）を適用するためには、プロファイリング方法（1）と関連するシミュレータ（2）を選択します（図 6）。皮膚感作性の予測をするために、次のシミュレータを使用することもできます：Skin sensitization and Autoxidation simulators（皮膚感作性および自動酸化 シミュレータ）。

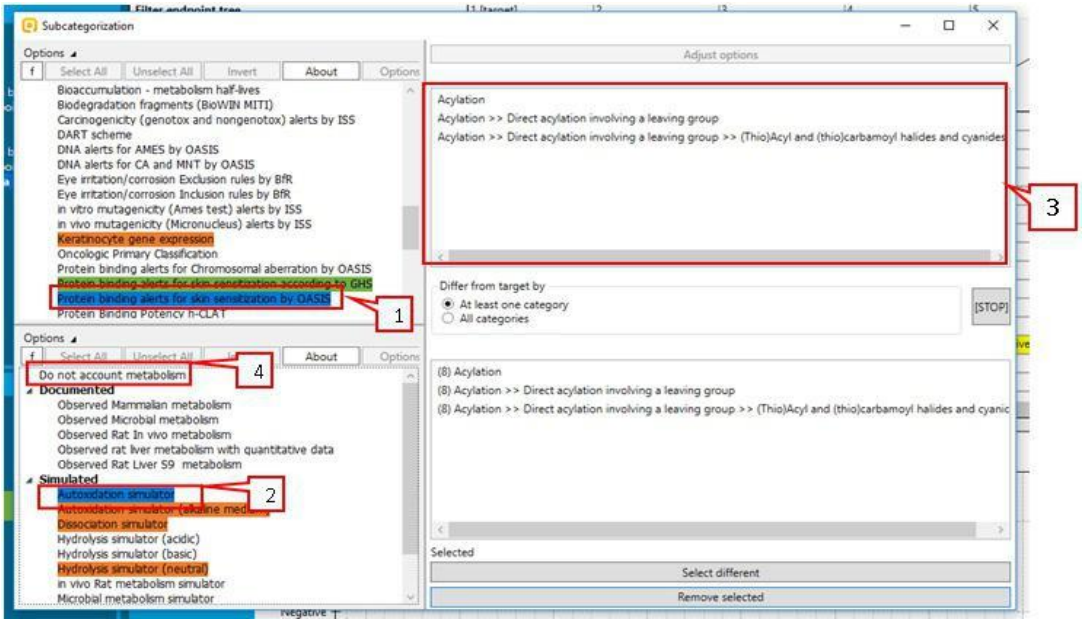


図 6

次のシミュレータは、 遺伝毒性予測に関連しています。Rat liver metabolism（ラット肝臓代謝）シミュレータや Rat liver S9 metabolism（ラット肝臓 S9 代謝）シミュレータです。

プロファイリングするプロセスにシミュレータが含まれていると、ソフトウェアはそのカテゴリの化学物質の代謝（の予測）を開始し、選択したプロファイルを標的とその代謝物のパッケージに適用します。このプロセスは、データベースが以前に代謝をおこなっていないと時間がかかることがあります。パッケージ、標的および代謝物のプロファイリング結果は、Subcategorization ウィンドウ（3）の標的パネルに表示されます（図 6）。

代謝物の代謝とプロファイリングを作動させないためには *Do not account for metabolism*（図 6-4）をクリックしてください。

注意: すべてのデータベースとインベントリにあるものが代謝されるわけではないことに注意してください。

プロファイラ（1）と simulated/observed metabolism simulators（2）（図 7）で選択オプションをご利用できます。

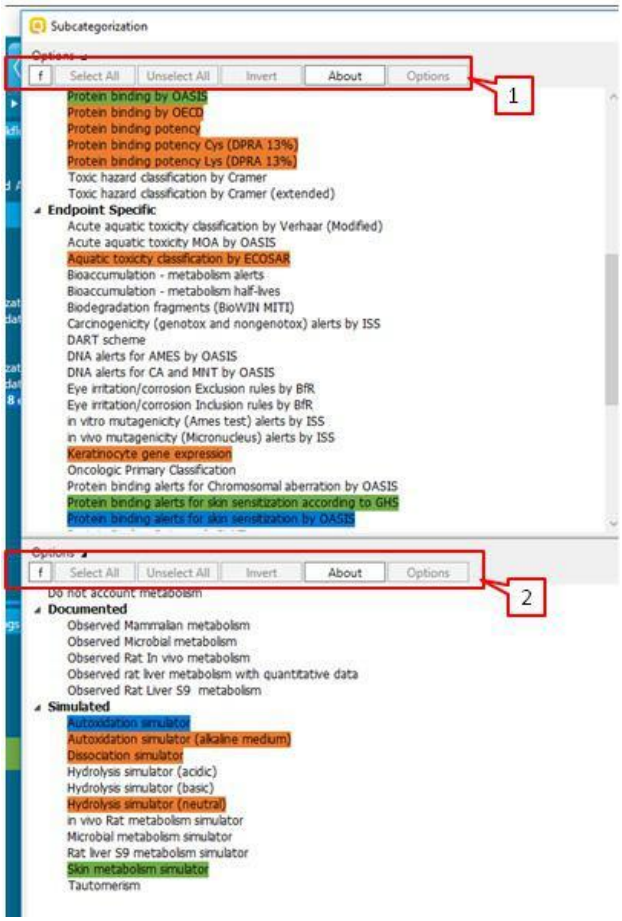


図 7

- ***Group by***： さまざまな基準に基づいてグループ化できます（図 8 を参照）。Options（1）を展開し、Group by （2）を選択して展開し、ドロップダウンリスト（3）からグループ化する基準を選択します。

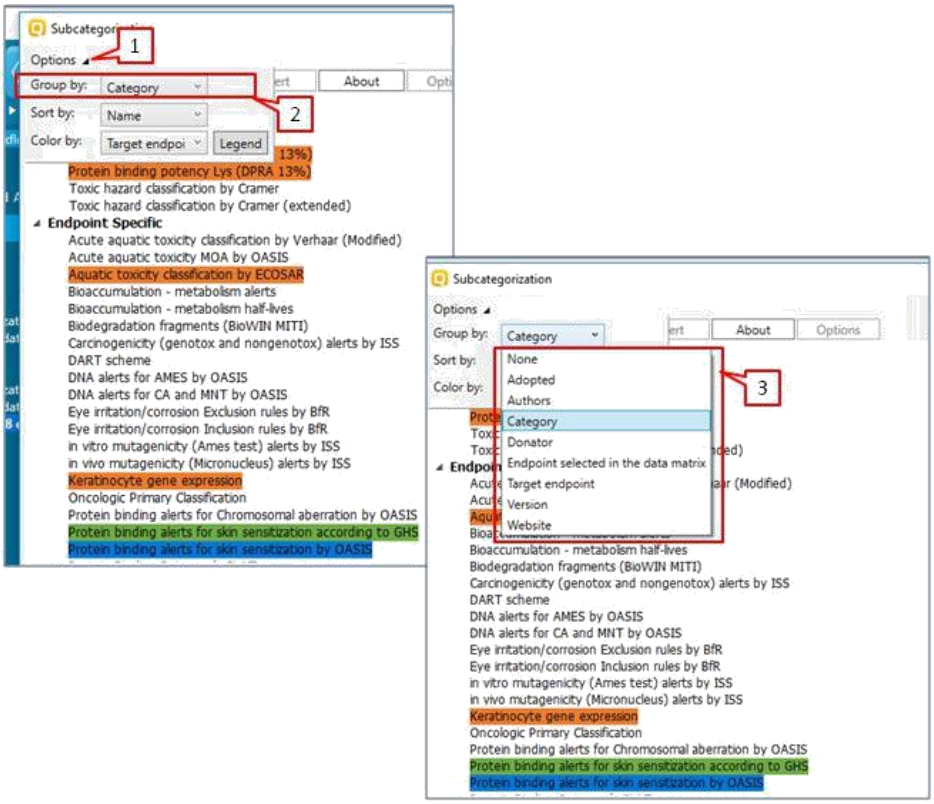


図 8

○ ***Sort by***： さまざまな基準に基づいて並べ替えを行うことができます（図 9 を参照）。Options（1）を展開し、 Sort by （2）を選択して展開し、ドロップダウンリスト（3）から並べ替えをおこなう基準を選択します。

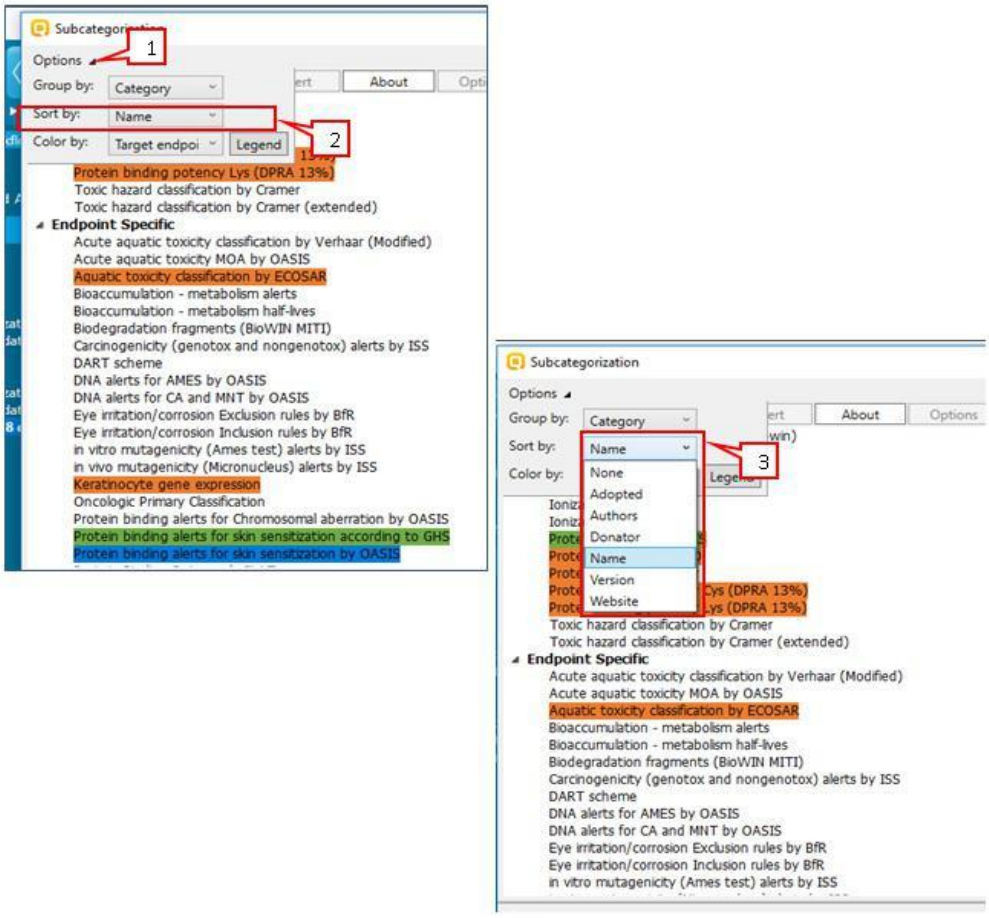


図 9

- ***Color by*** : データ・マトリックスまたは標的エンドポイントで選択されたエンドポイントに基づいてプロファイラー/シミュレーターに色付けします（図 10）: Options（1）を展開し、 *Color by* （2）を選択して展開し、ドロップダウンリスト（3）から並べ替えをおこなう基準を選択します。

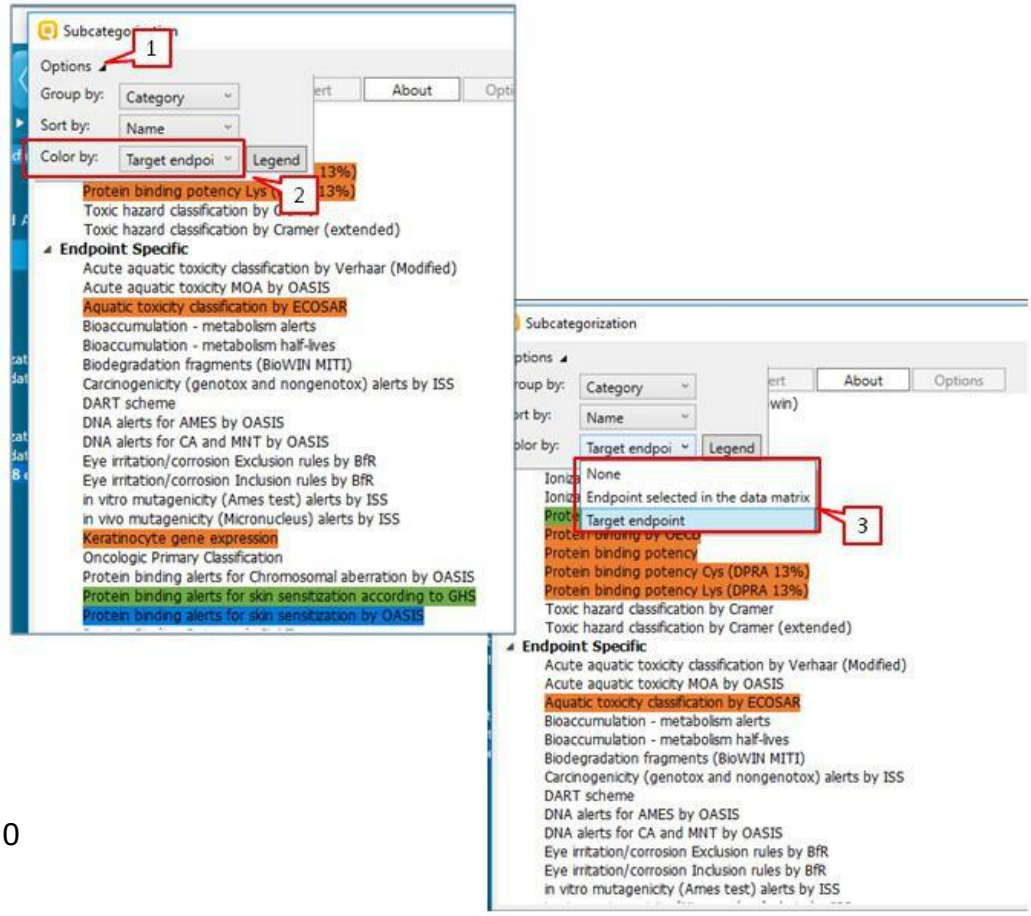


図 10

- **About** - 各プロファイラの背後にある情報とドキュメントを表示することができます（図 11）。

- ｆ （フィルタ）ボタン（1）をクリックし、キーワード（2）を入力します。
- プロファイラ（3）を選択します。
- About（2）をクリックして About（情報）ウィンドウを表示します。
- *Documentation* （5）をクリックすると、プロファイラ（6）に関するすべての情報が表示されます。

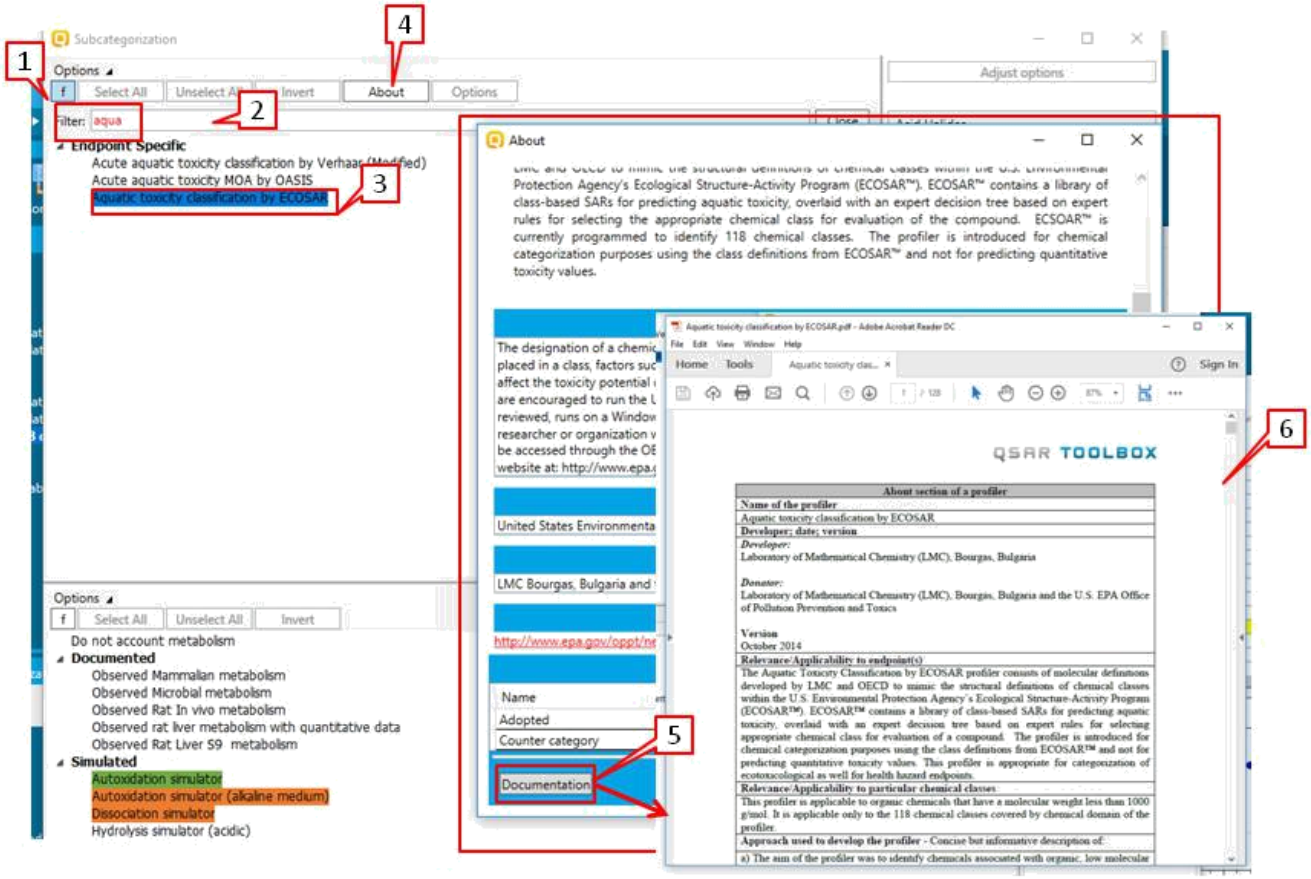


図 11

Subcategorization ウィンドウ内の右側のパネルは、2 つの部分で構成されています。標的物質と類似物質です。両方のパネルには標的物質と類似物質のプロファイル結果が含まれています。

標的物質パネルには、利用可能な場合は、その標的物質と代謝物（3）の、選択されたプロファイリング方法（2）にまたがるプロファイリング結果（1）が含まれています。(図 12)

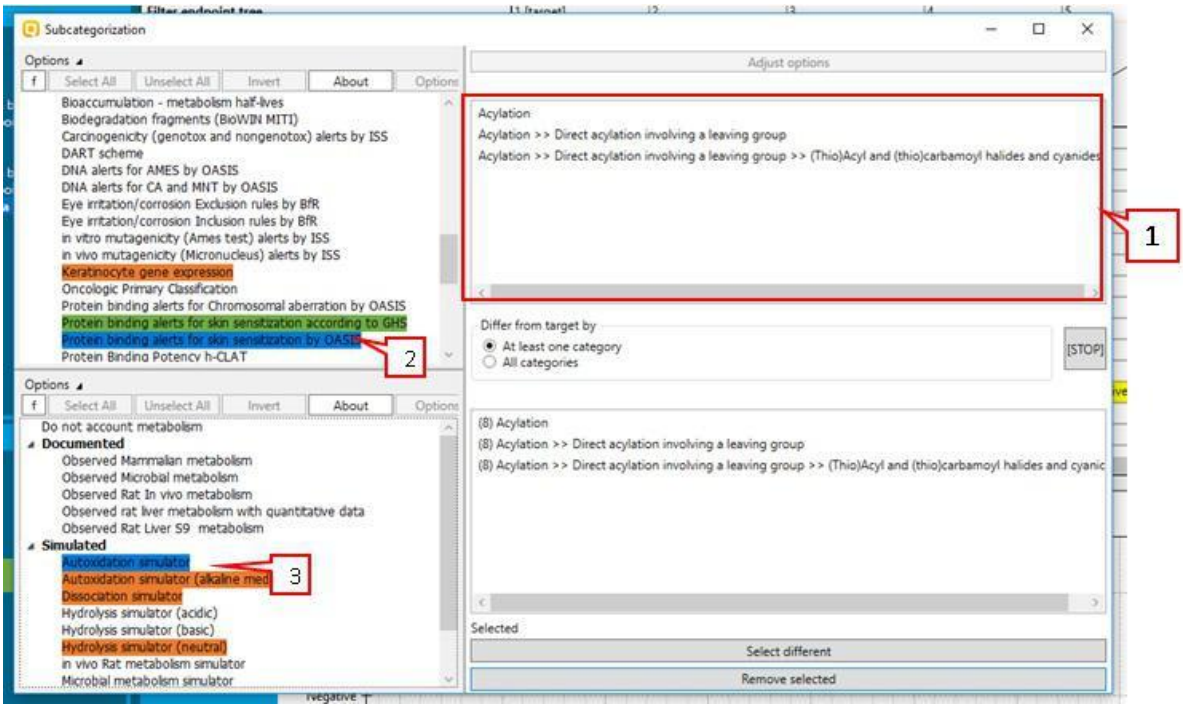


図 12

Analogues パネル（1）には、適用させたプロファイラ（2）とシミュレータ（3）に基づく類似化学物質（1）のプロファイリング結果が含まれています。(図 13) 標的化学物質と異なるプロファイリング結果を有する類似物質は、青色の背景に着色されます（4)。括弧内の数字は、それぞれのカテゴリ （ハロゲン化アルキル）の類似物質の数を示します。類似していない物質を削除するには、削除（5）をクリックします。(図 13)

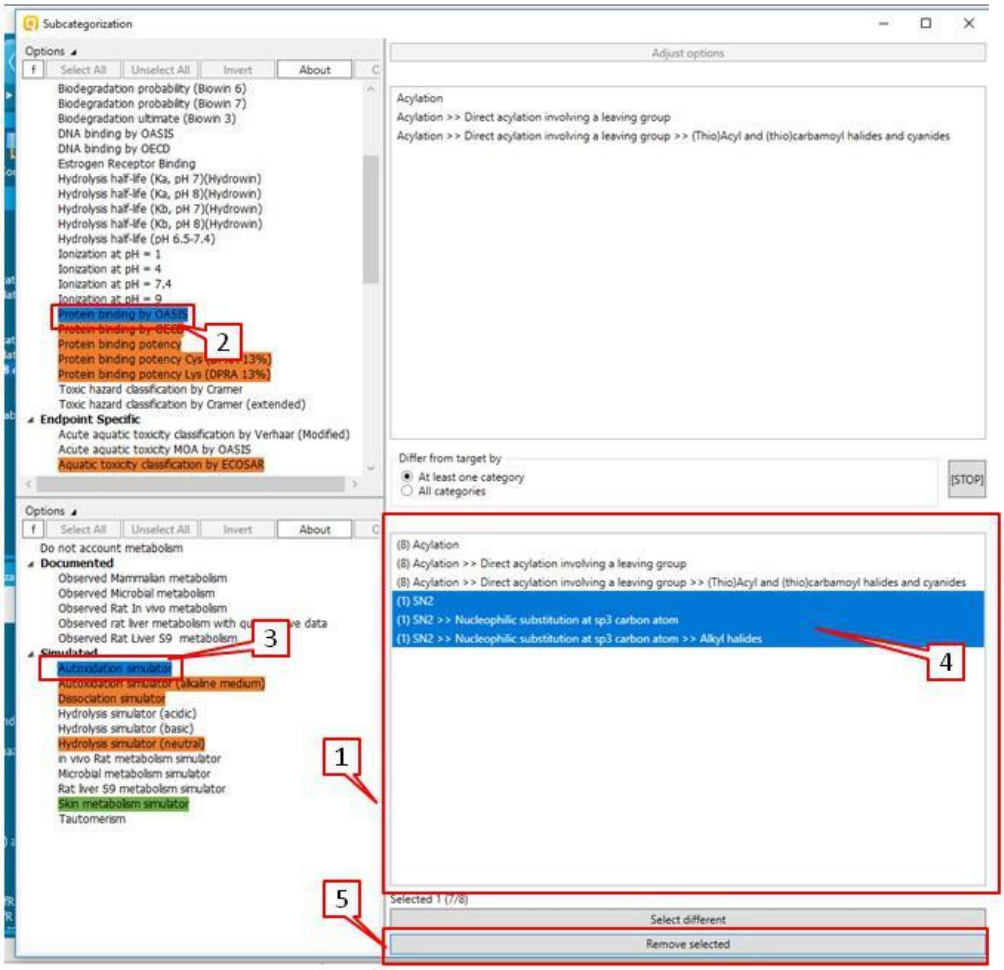


図 13

カテゴリを右クリックすると、二次機能のメニューが開きます（1）(図 14)

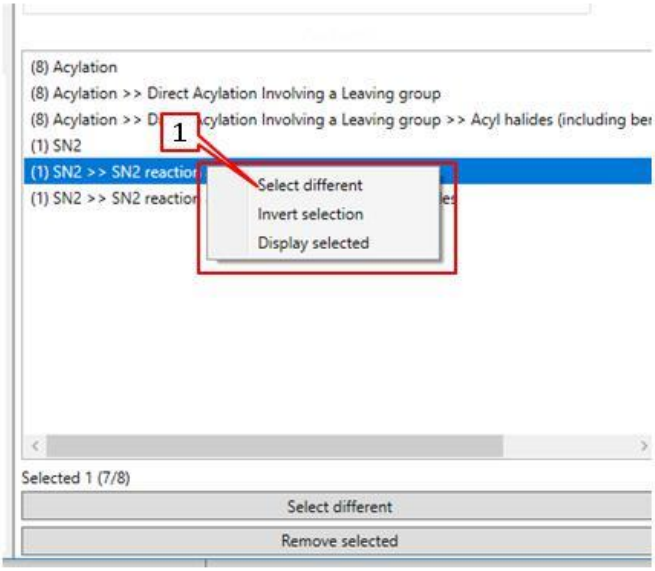


図 14

Display/selected をクリックすると、選択したすべての類似物質を含む新しいウィンドウが開きます（図 15）。



図 15

- *Select different* を選択した場合：プロファイリングの結果（図 16）について、類似物質が標的化学物質と区別するため 2 つのオプションがあり ます。
  - 標的物質の有するものとは異なっている、少なくとも 1 つのプロファイリングカテゴリを有する類似物質（1）
  - 標的物質の有するものとは、すべてが同じではないプロファイリングカテゴリを有する類似物質（2）

第 1 のオプションは、このソフトウェアが、標的化学物質に関連するカテゴリ（プロファイリング結果）の少なくとも 1 つを有する類似物質を検索するという意味です。

第 2 のオプションは、このソフトウェアが、標的化学物質に関連するカテゴリ（プロファイリング結果）すべてを同時に有する類似物質を検索するという意味です。

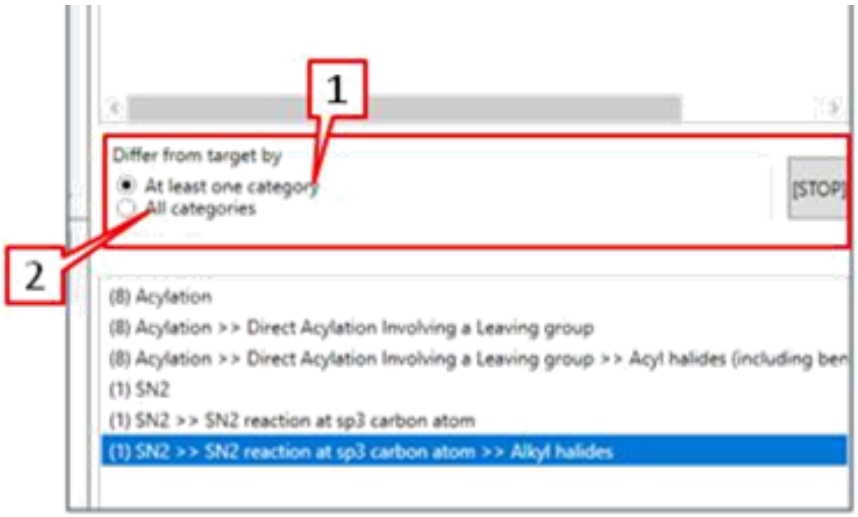


図 16

- ***Adjust options（調整オプション）***

Adjust options ボタンは、 *Structure similarity* （構造類似性）プロファイラ（1）で使用されます。（訳注：Subcategorization 内の Options で Group by Category にすると、図 17 の状態になるので図 17-1 の Structure similarity を押すと右上の Adjust options ボタンが作動可能になる。*Adjust options* ボタンの使い方は図 21 以降を参照）Structure similarity プロファイラは、標的化学物質と類似物質との間のさまざまなレベルの構造類似性に基づいて、化学物質（類似物質）を同定するために開発されています。Structure similarity プロファイラ（1）を選択すると、ソフトウェアはすべての類似物質を類似性の区分（2）に振り分けします（図 17）。

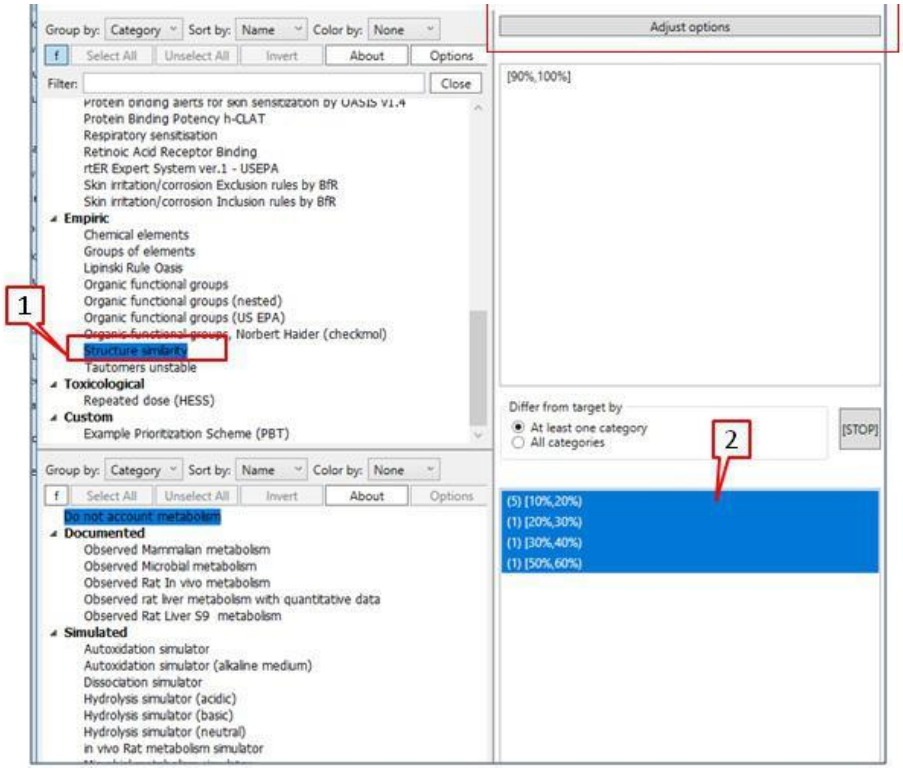


図 17

各区分に入る化学物質を見るには、区分（1）をクリックして右クリックし、ポップアップウィンドウ（2）を表示します（図 18）。*Display selected* をクリックすると、化学物質の構造を含むウィンドウが開きます。化学構造（3）を選択し、右クリックしてポップアップウィンドウの Explain オプション（4）を開き、クリックすると新しいウィンドウのプロファイリング結果が開きます。

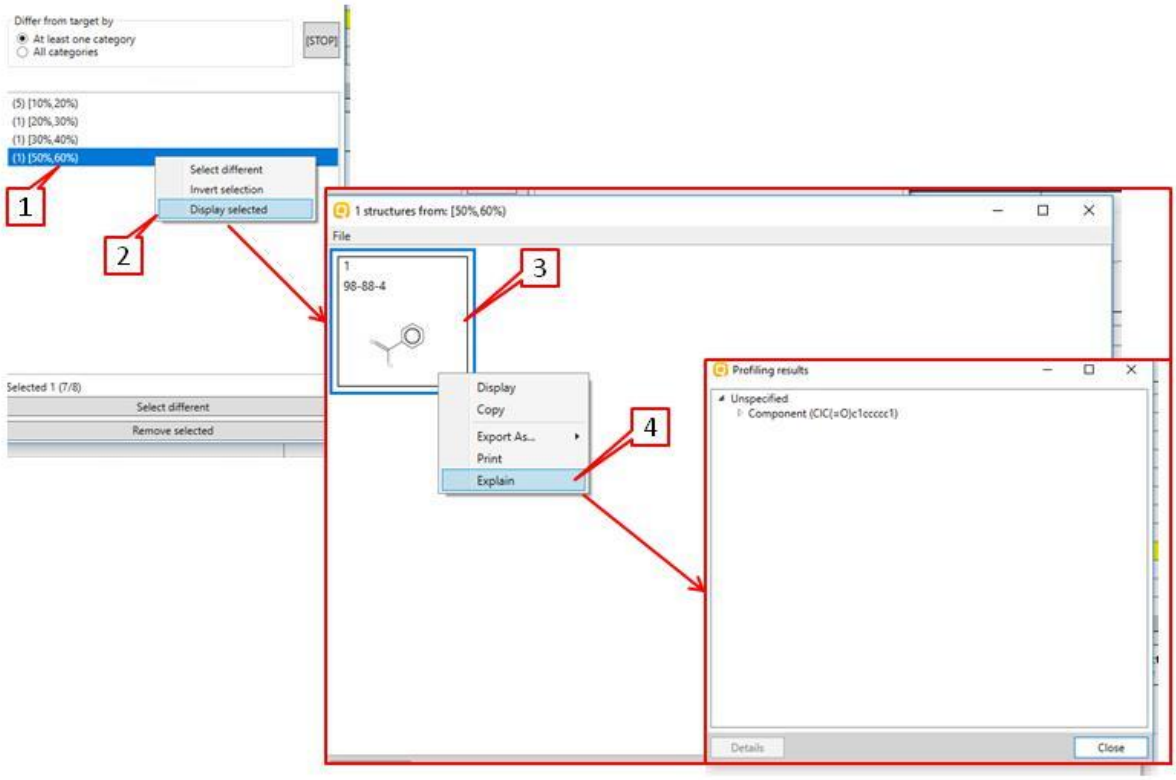


図 18

Component（1）の部分を展開し、similarity の範囲（2）を選択してから、*Details*（3）をクリックして Similarity explain ウィンドウ（4）（訳注：旧版では Similarity form ウィンドウ）を開きます（図 19）。

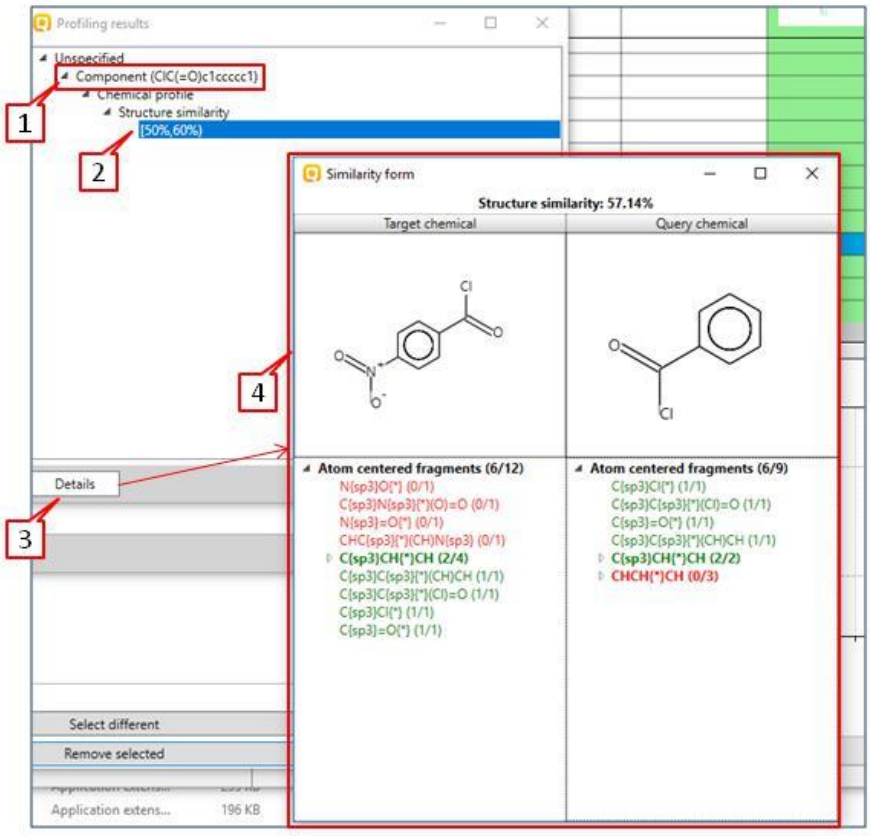


図 19

ウィンドウの上部には、構造間の類似度のパーセンテージが表示されます（1）。標的化学物質の分子構造と比較構造が表示されます（2）。2 つの化学物質間の類似性を決定する際に使用される特徴（フラグメント）（3）も表示されています。緑色（4）は、2 つの比較分子に共通する特徴を示すものであり、赤色は、差異（5）の指標です。(図 20)

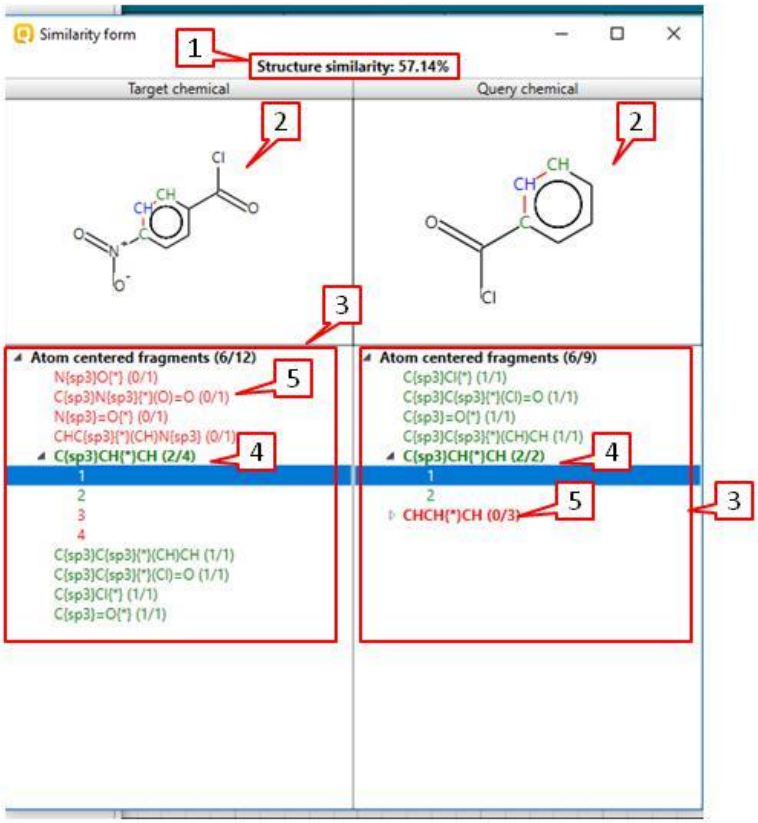


図 20

- ***類似性の設定***

類似度計算の設定リストは、***Adjust option*** ボタン（1）（図 21）をクリックすると利用できます。次の ***Similarity options*** のウィンドウ が含まれています。

- 2 - 類似度の計算方法
- 3 - 類似度の計算式を表示するパネル
- 4 - 類似性の特徴の重なりの図示
- 5 - 計算に使用される分子の特徴
- 6 - 対応する分子の特性に関連するオプション
- 7 - 選択された分子の特性の短い説明
- 8 - 計算オプション
- 9 - の類似度の計算に使用される原子の特性
- 10 - 選択された方法と類似性の特性の図例
- 11 - SMILES 入力欄
- 12 - デフォルトオプションを設定する

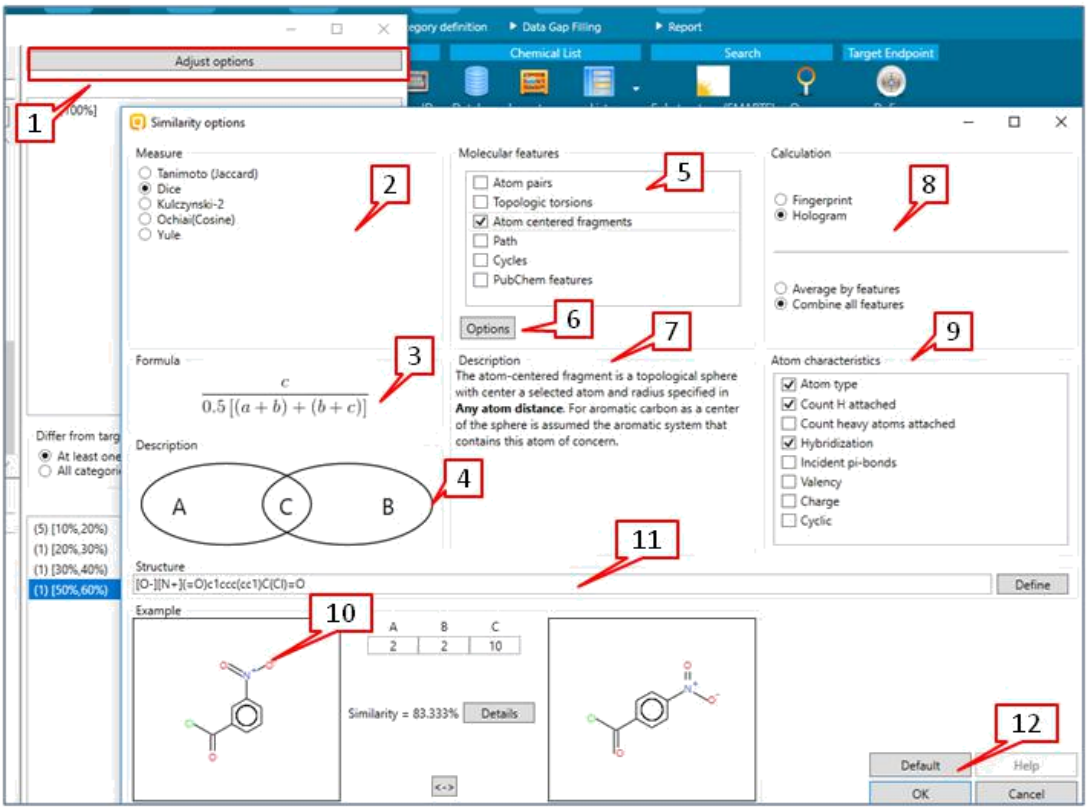

Adjust options
1
Similarity options
Measure
Tanimoto (Jaccard)
Dice
Kulczynski-2
Ochiai(Cosine)
Yule
2
Molecular features
Atom pairs
Topologic torsions
Atom centered fragments
Path
Cycles
PubChem features
5
Options
6
Calculation
Fingerprint
Hologram
8
Average by features
Combine all features
Formula
$\frac{c}{0.5\,[(a+b)+(b+c)]}$
3
Description
7
The atom-centered fragment is a topological sphere with center a selected atom and radius specified in **Any atom distance**. For aromatic carbon as a center of the sphere is assumed the aromatic system that contains this atom of concern.
9
Atom characteristics
Atom type
Count H attached
Count heavy atoms attached
Hybridization
Incident pi-bonds
Valency
Charge
Cyclic
Description
A
C
B
4
11
Structure
[O-][N+](=O)c1ccc(cc1)C(Cl)=O
Define
Example
10
A
B
C
2
2
10
Similarity = 83.333%
Details
<->
12
Default
Help
OK
Cancel
(5) [10%,20%)
(1) [20%,30%)
(1) [30%,40%)
(1) [50%,60%)

図 21

D.5.4.2.1.2. Mark chemicals by WS（水溶解度で印を付ける）

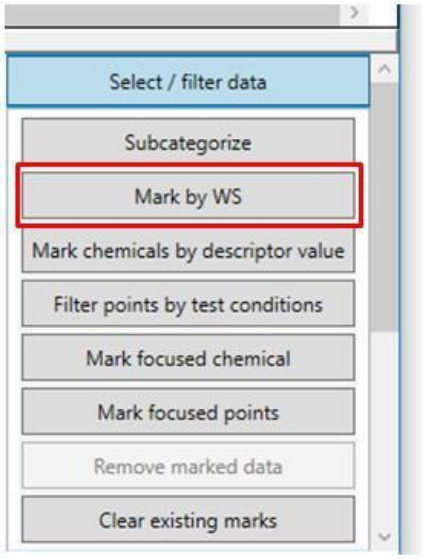


図 1

この機能は、容積濃度単位の群にあるデータ値を、最大水溶解度の計算値と比較して、実験誤差を検出します。このオプションが選択されている場合には 、ポップアップウィンドウが表示されます（図 2)。

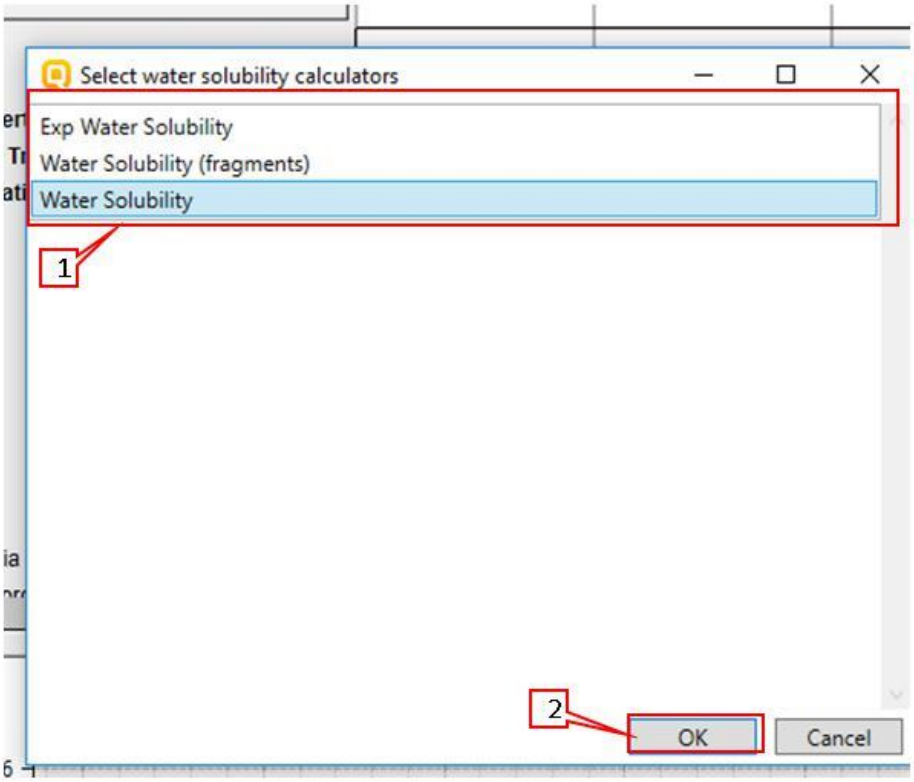


図 2

使用可能な方法（1）のいずれかを選択し、**OK** をクリックしてください。（図 2-2）
情報メッセージが表示されます（図 3)。

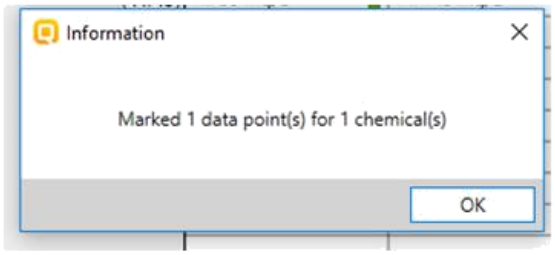


図 3

マークされた化学物質は淡い青色で着色されています（図 4）。*Mark focused point* をクリックします（1）。

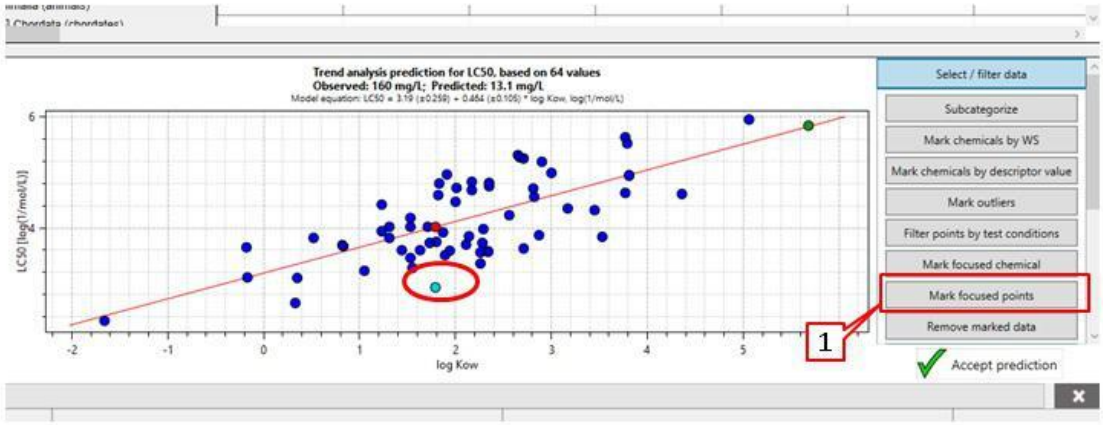


図 4

ポイントがマークされ（つまり緑色に変わり）、削除することもできます（図 5）

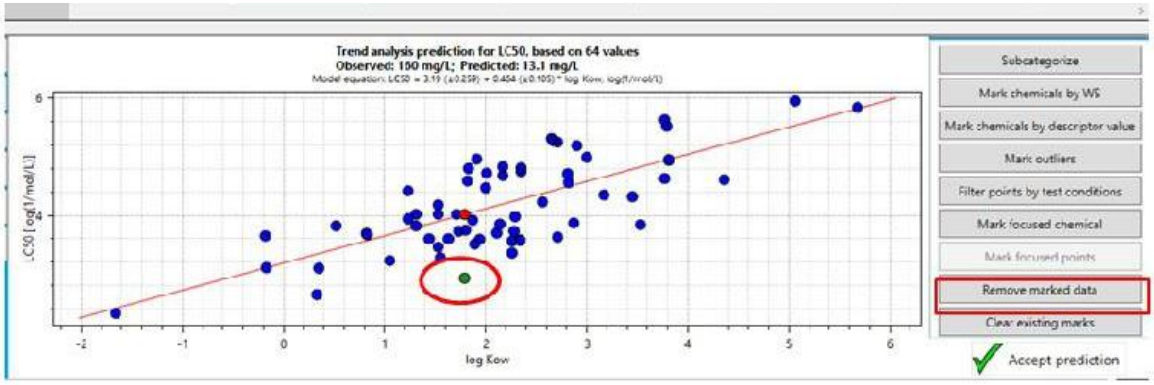


図 5

図 6 では、マークされた点が除去されています。

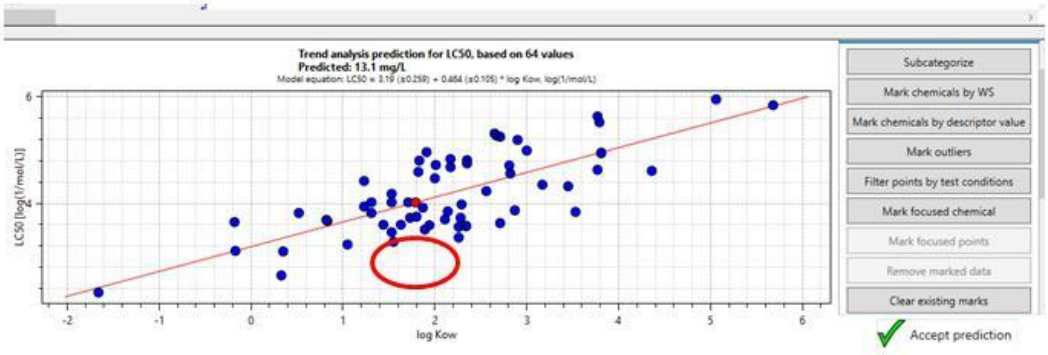


図 6

D.5.4.2.1.3.　Mark chemical by descriptor value（軸上の値により印を付ける）

この機能は、特定のパラメータの範囲内に含まれる化学物質を選択するときに使用されています。*Mark chemicals by descriptor value*　（1）をクリックして使用可能なディスクリプタ（2）の 1 つを選択し、必要な範囲を決定し（3）、OK（4）をクリックします。(図 1)

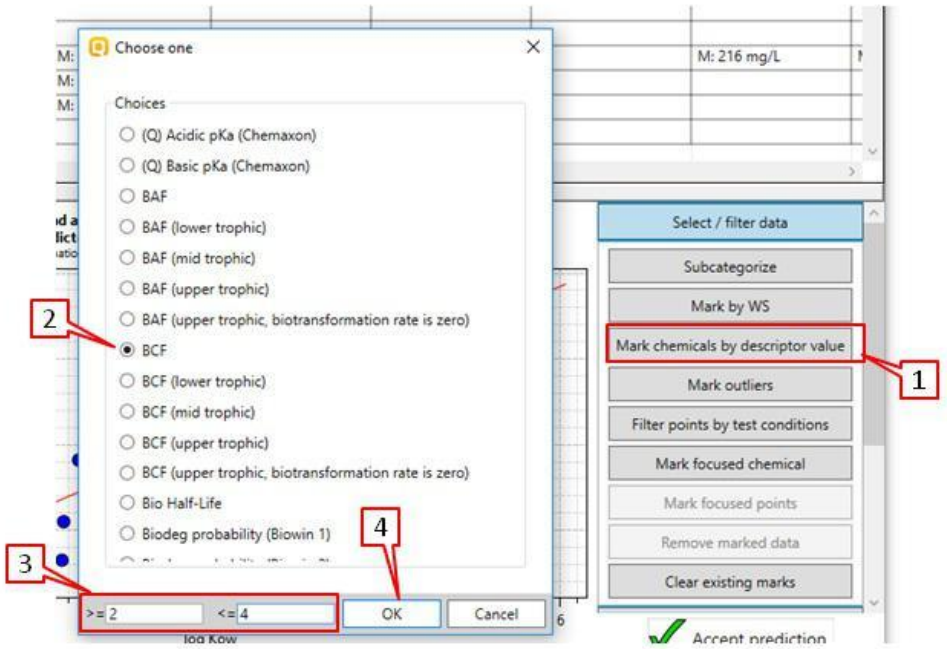


図 1

決定した範囲内にある化学物質が選択されます（緑色で着色されています）（1）（図 2）。

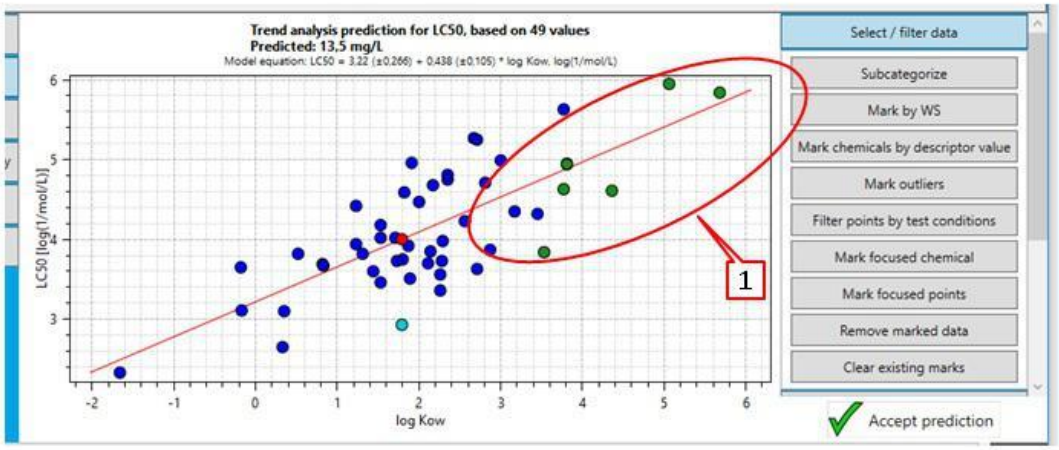


図 2

D.5.4.2.1.4.　Filter points by test conditions（試験条件によるフィルタリングポイント）

この機能は、付随データに基づいてデータポイントの一部を削除する機能を提供します。

Filter points by test conditions ボタン（1）をクリックして、*Data Filter* ウィンドウを開きます。付随データフィールド（2）で、例えば test organisms (species)をクリックし 、標的物質と選択された類似物質との間の試験生物種の違いを見ます（3)。標的と異なる系統種に基づく付随データは、薄い灰色になっています。また、 select different （4）で、灰色でないものを選択することもできます。*Remove* をクリックして、選択した化学物質の実験データを削除します。

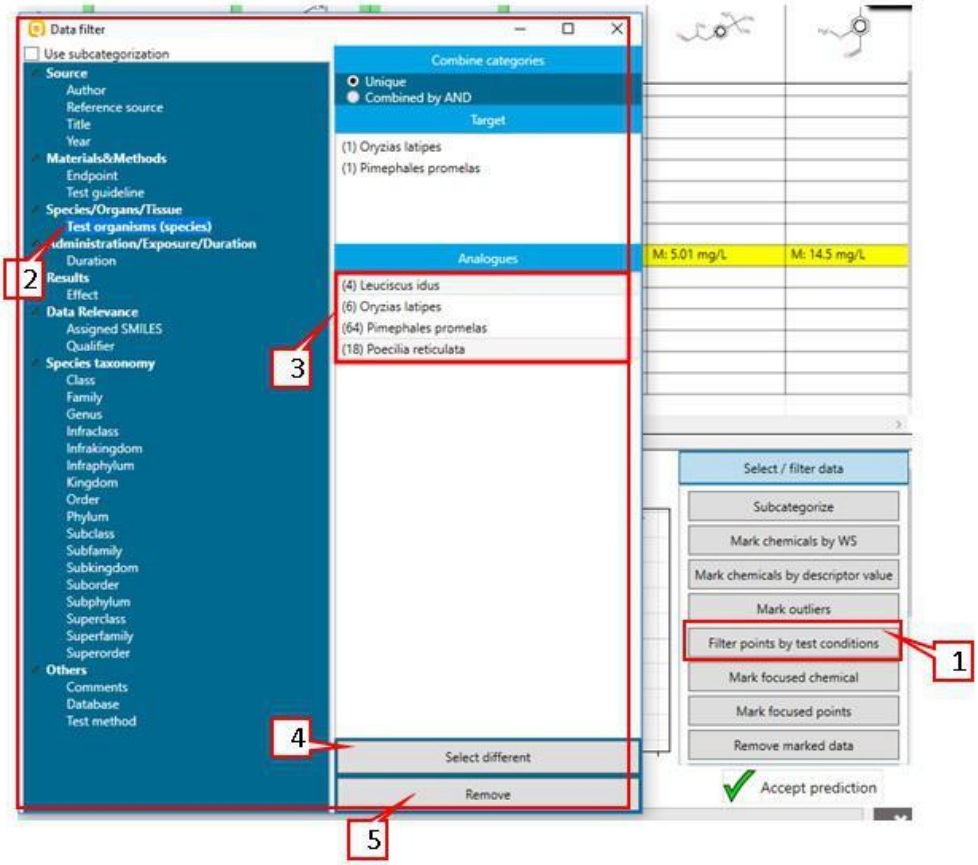


図 1

Note

Filter by test condition functionality removes experimental data for a given chemical.

［注意：Filter by test condition の機能は所定の化学物質の実験値を除去します。］

D.5.4.2.1.5. Mark focused chemical（化学物質にフォーカスして印をつける）

グラフからひとつの化学物質をクリックすると、薄い青色で着色されるので（1）、*Mark focused chemical* (2)をクリックすると、その化学物質を緑色で着色します（3）（図 1）。これは、ユーザが特定のデータを削除したい場合に実行されます（次々章の Remove marked chemicals / points を参照してください）。

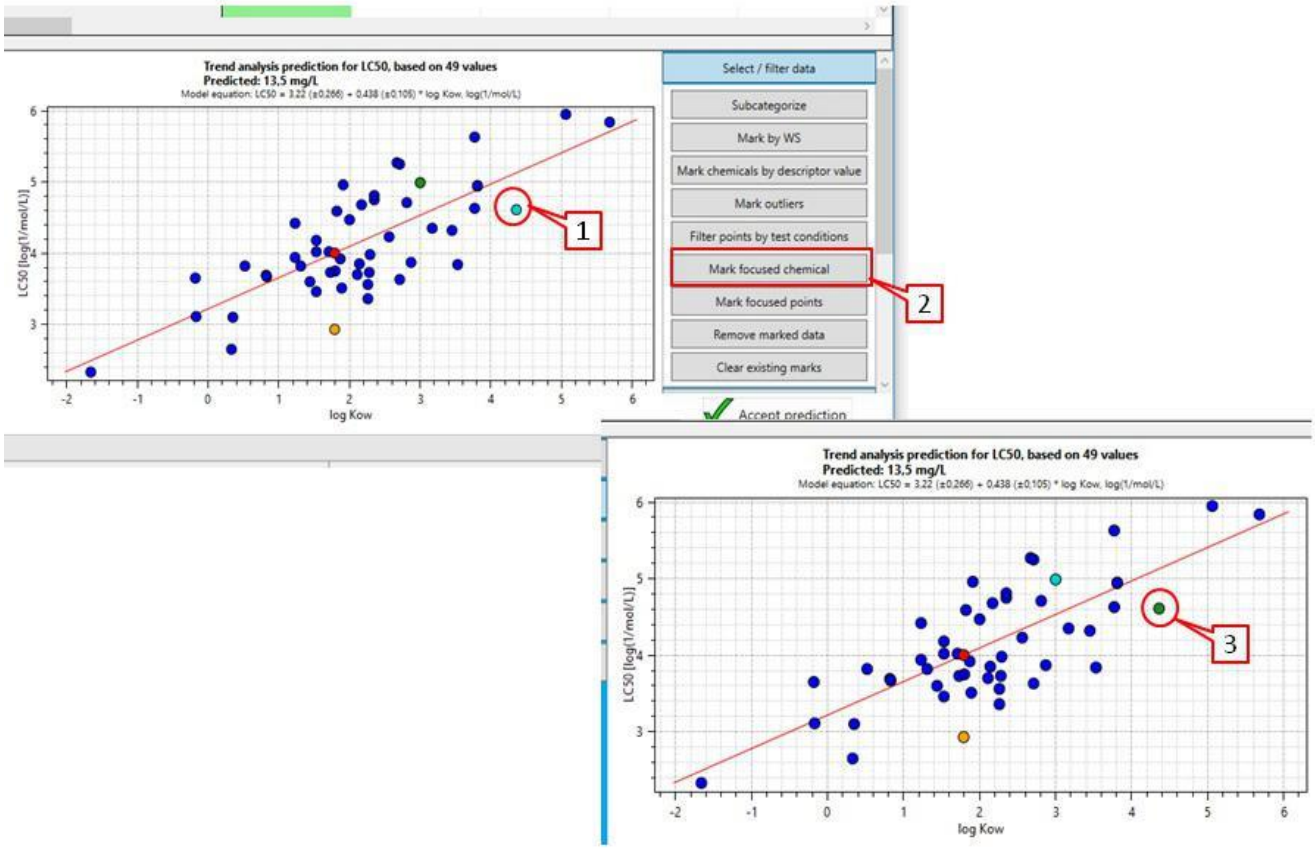


図 1

D.5.4.2.1.6. Mark focused points（データ点にフォーカスして印をつける）

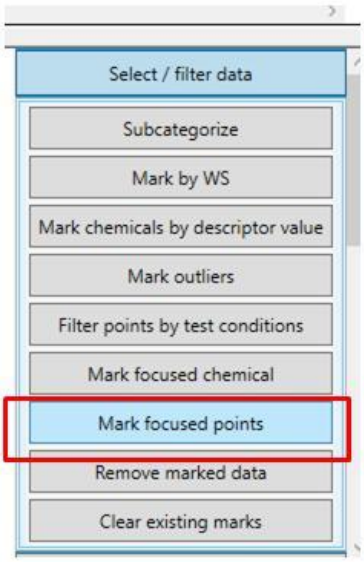


図 1

この機能を使用するには、まずグラフ上の 1 つまたはいくつかの点を選択する必要があります。これはたとえば ***Region selection*** ツール（2）（***Information*** ボタンを押すと下位に表示されるサブボタン）（図 2-1）を使用して行います。これにより、グラフ（3）内のデータ点の範囲を選択できます。

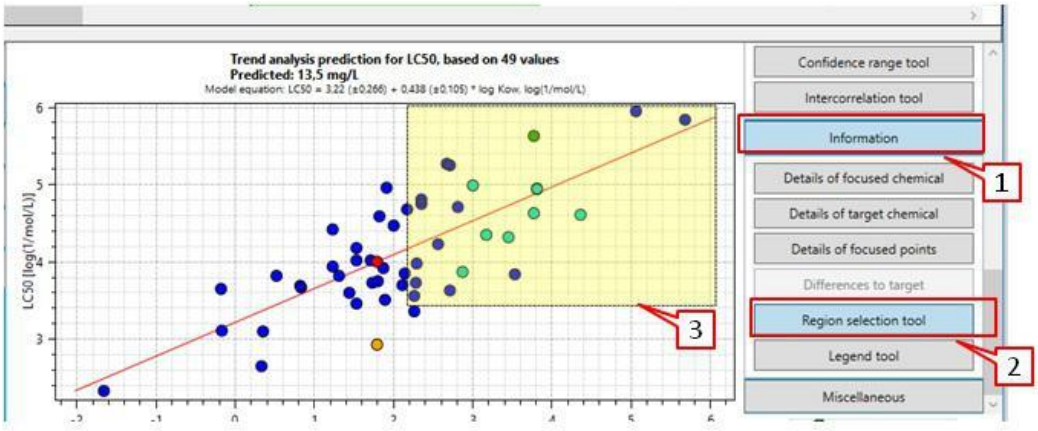


図 2

選択したデータ点は、淡いブルーに塗りつぶされます（1）。(図 3)

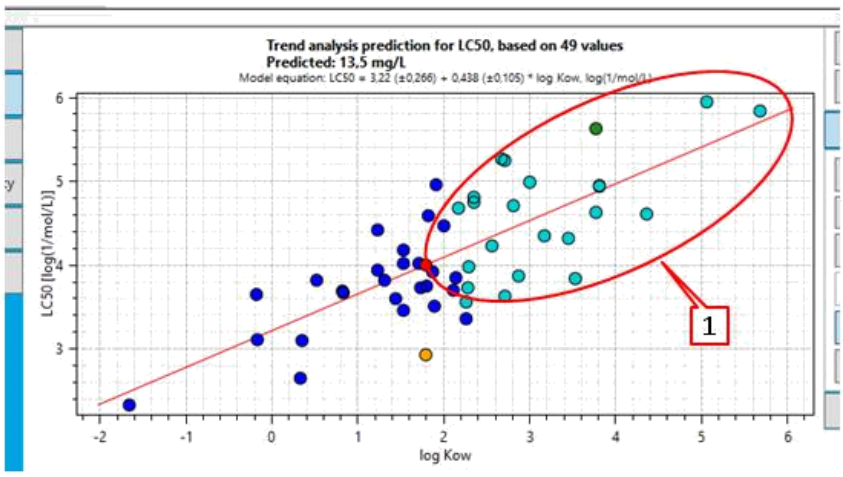


図 3

データ点が強調表示されたら、 *Mark focused points (1)*をクリックします。この機能は、現在フォーカスされているデータポイント（図 4-2）を緑色でマークします。

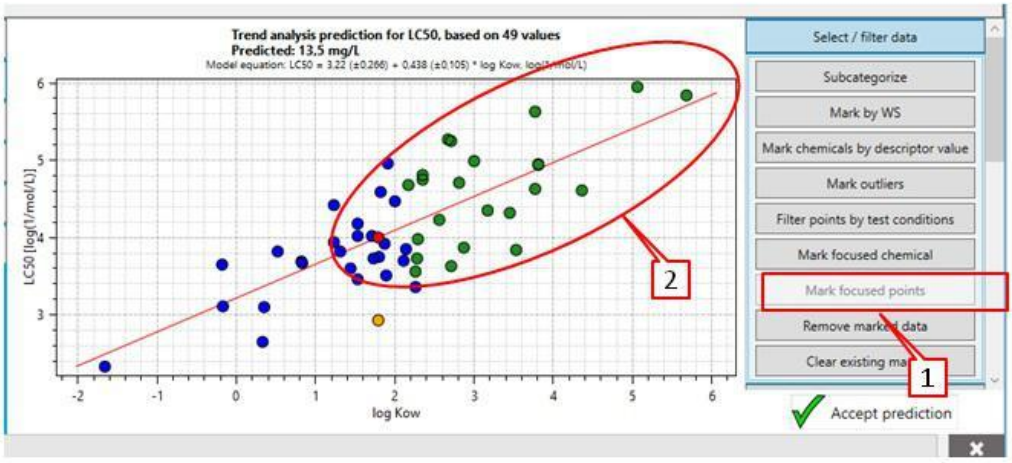
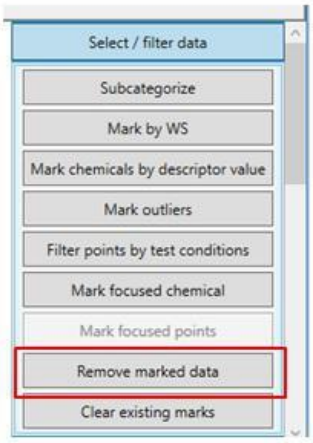


図 4

D.5.4.2.1.7. Remove marked data（マークしたデータを削除）



この機能（1）は、マークされた（緑色で着色された）化学物質/データ点を除去します（図 1）。

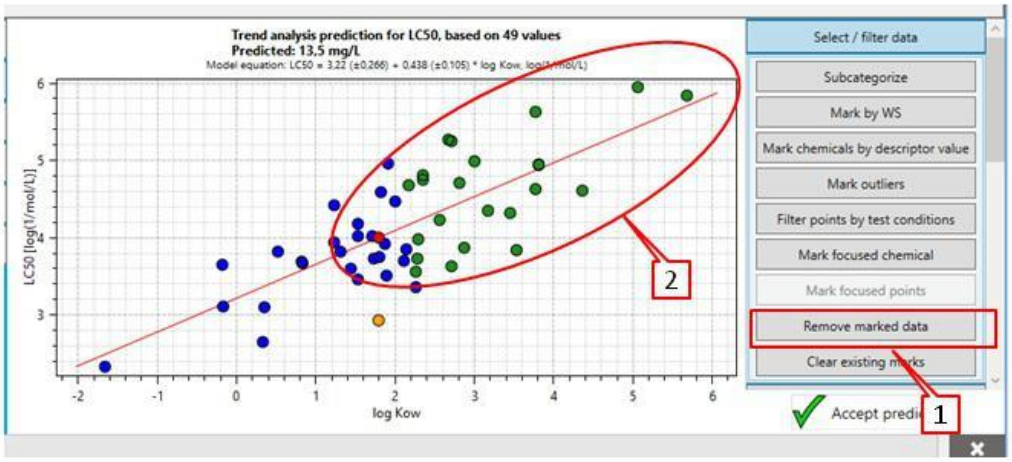
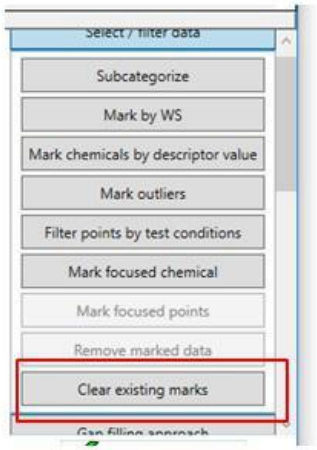


図 1

D.5.4.2.1.8.　Clear existing marks（既存のマークを消去）



この機能は、Mark focused chemical または Mark focused points の実行後に化学物質/データ点のマークを外します。

D.5.4.2.1.9. Mark outliers（外れ値のマーク）

この機能は、外れ値である化学物質を選択するために使用されます。*Mark outlier* (1)をクリックします（図 1）。外れ値がない場合、メッセージが表示され（2）、そうでない場合、外れ値はグラフ上で緑色で表示されます。

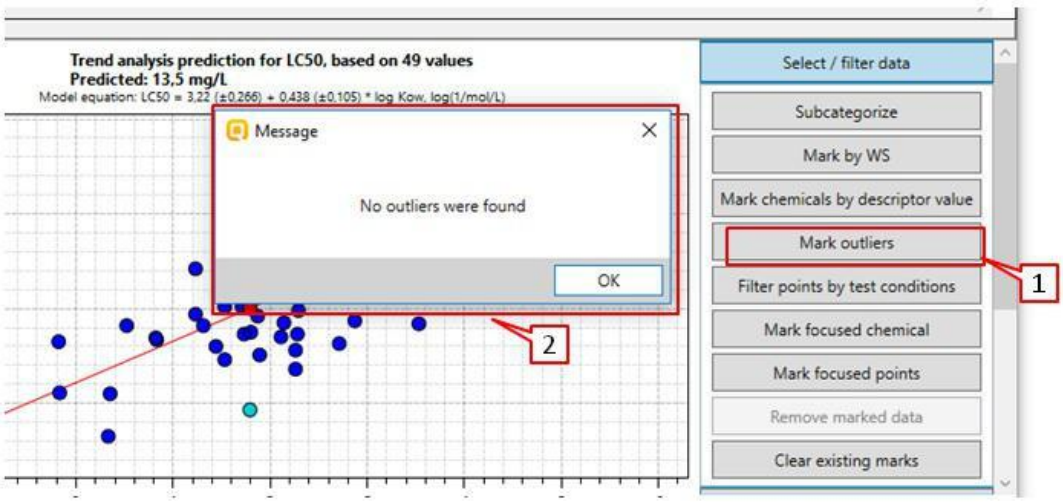


図 1

D.5.4.2.2.　Gap filling approach(ギャップ補完のアプローチ)

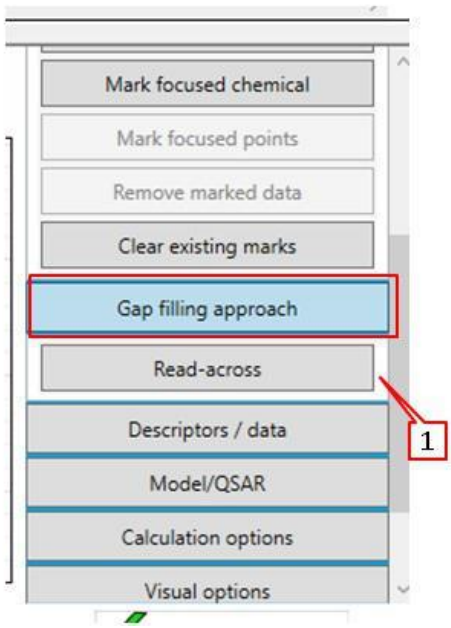


図 1

この機能により、2 つのデータギャップ補完モジュール（1）を切り替えることができます（図 1）

- ***Read-across*** をクリックすると、トレンドアナリシスからリード・アクロスに切り替わります。
- ***Trend analysis*** をクリックすると、リード・アクロスからトレンドアナリシスに切り替わります。

D.5.4.2.3. Descriptors/data（グラフ軸の記述子/データ）

このパネル（図 1）には、3 つの機能が含まれています。

- Change descriptor units（グラフ軸記述子の単位変更）
- Edit descriptor options（グラフ軸記述子のオプション編集）
- Select endpoint tree descriptor（エンドポイントツリー記述子の選択）

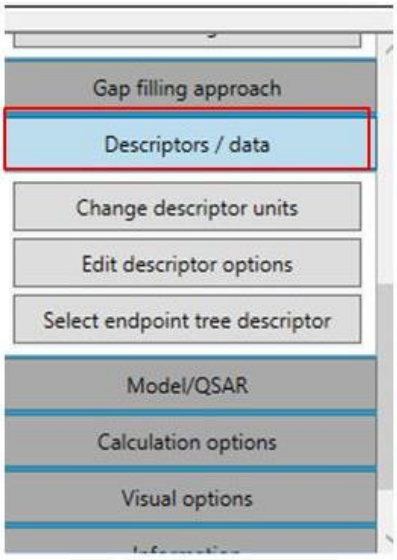


図 1

D.5.4.2.3.1. Change descriptor units（グラフ軸記述子の単位変更）

Change descriptor units （2）は、x 軸パラメータが単位で表現される場合に適用できます。

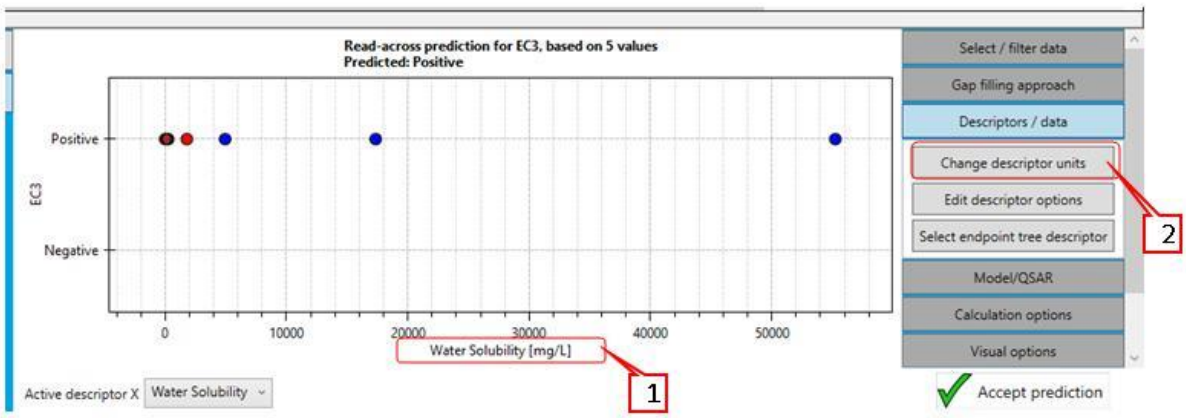


図 1

Mass concentration （質量濃度）（1）を選択すると、現在の単位が mg / L（2）と表示され、データギャップ補完のグラフ（図 2）にそのように表されます（図 3）。式を log（1/endpoint）（4）に変更して、Accept（5）をクリックします。

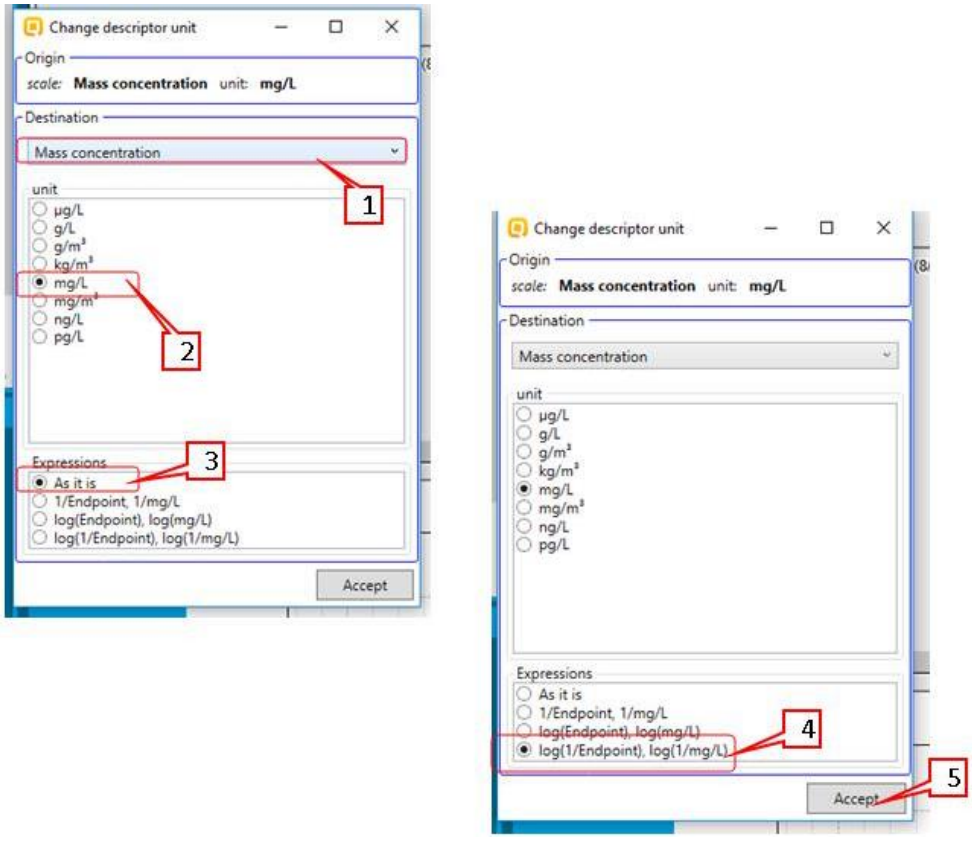


図 2

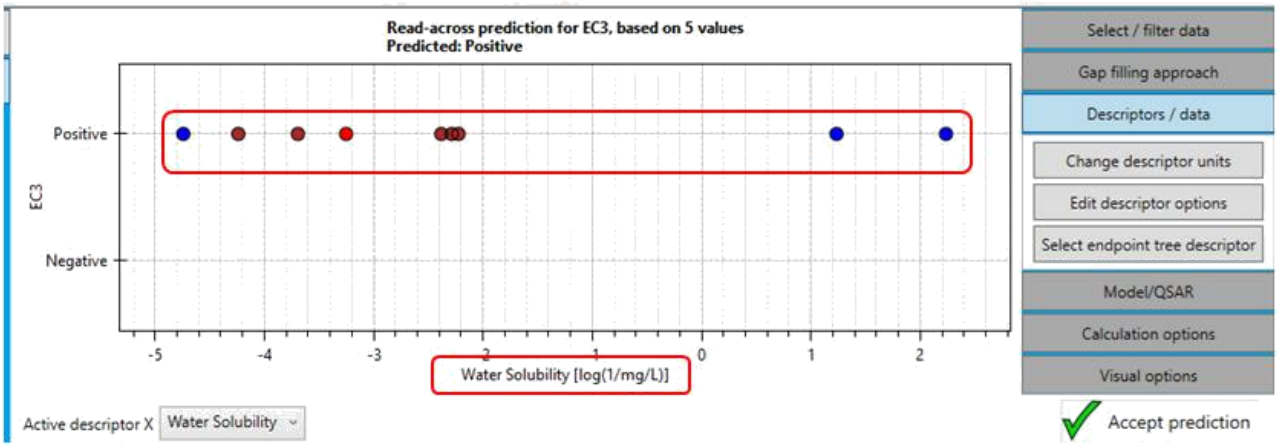


図 3

単位変更の結果、グラフ上の点の分布が変更されます（図 3）。

D.5.4.2.3.2.　Edit descriptor options（記述子のオプション編集）

Edit descriptor options (1) は、SMILES（2）のセットから代表的なパラメータを計算するためのさまざまなオプションを提供します（図 1）。x 軸パラメータが類似度[%]の場合は、*Additional options* (3) をクリックして similarity options ウィンドウを表示し、類似性オプションを変更します。OK（5）をクリックします。

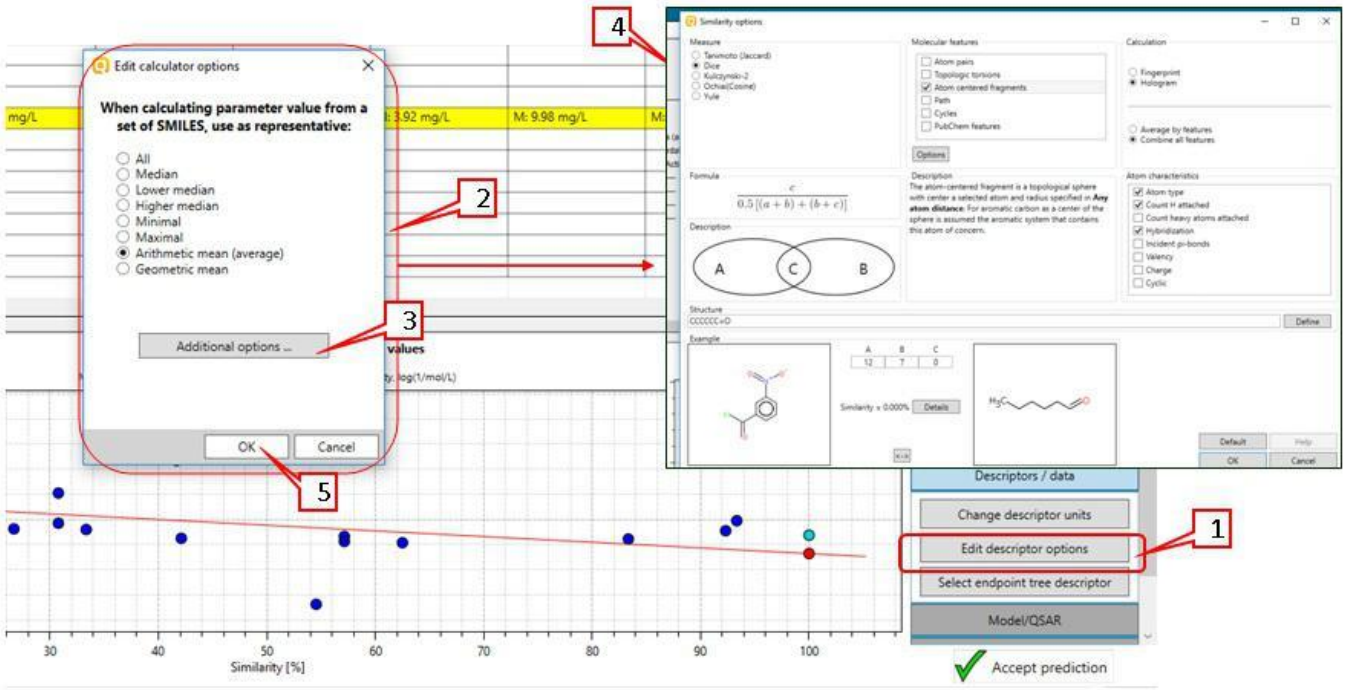


図 1

D.5.4.2.3.3. Select endpoint tree descriptor（エンドポイントツリー記述子の選択）

データギャップ補完の典型的な方法は、目的とするエンドポイントの実験データと 2D や 3D パラメータとの間の相関/依存性を調べることです。モデリングの観点からは、目的のエンドポイントは従属変数であり、2D / 3D パラメータは独立変数です。しかし、異なるエンドポイント間の相関を調査して使用するためには、他のエンドポイントの実験データが独立変数として使用される可能性があります。これは、***Select endpoint tree descriptor*** オプションを選択して行います。エンドポイント変数を記述子として選択できる新しいウィンドウ（2）を開くには、Select endpoint tree descriptor （1）をクリックします（図 1）。

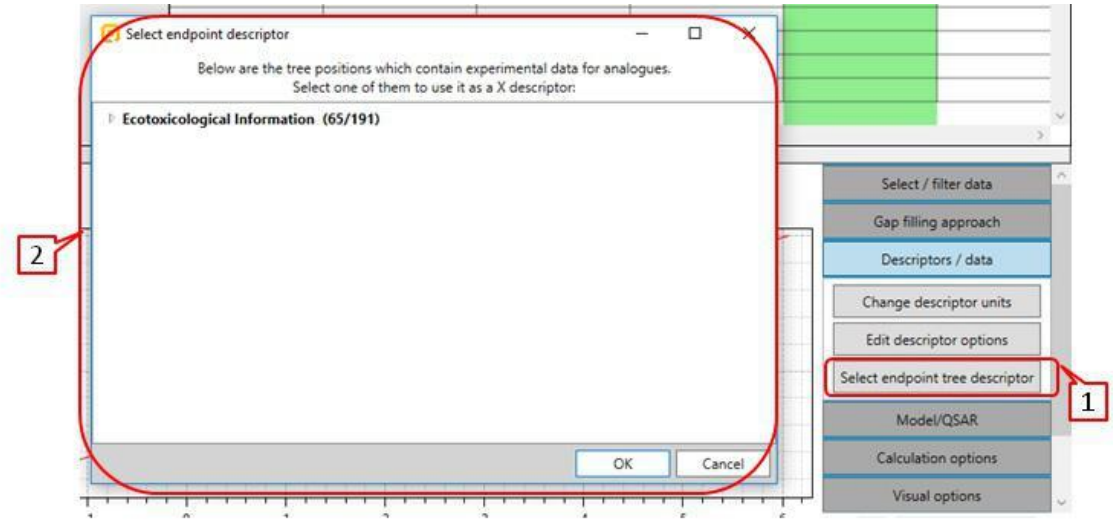


図 1

Select endpoint tree descriptor は、データギャップ補完法で使用される類似物質のさまざまな実験データの比較を可能にします。この機能を選択すると、システムはデータマトリクスで収集されたすべてのデータを繰り返し処理し、実験データを持つノードのみを含むエンドポイントツリー（元のエンドポイントツリーのサブセット）を構築します（図 2）。ツリーを展開し、目的のエンドポイント（1）を選択します。OK (2)をクリックします。

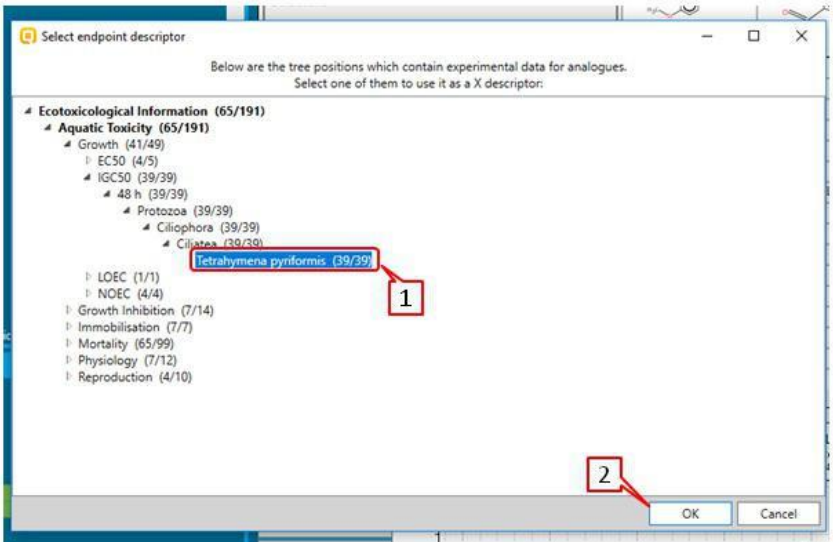


図 2

Possible data inconsistency ウィンドウが表示される可能性があります（図 3）。scale/ unit を選択し、OK（1）をクリックします。

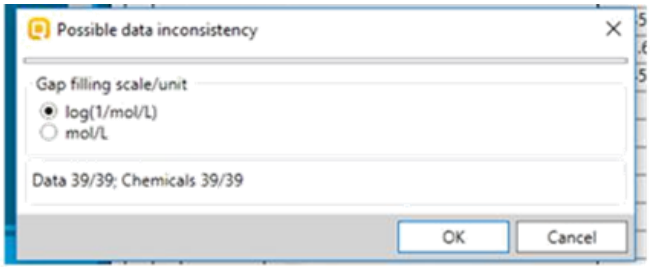


図 3

2 つのエンドポイント間の相関関係がグラフ上に表示されます（図 4 ）。

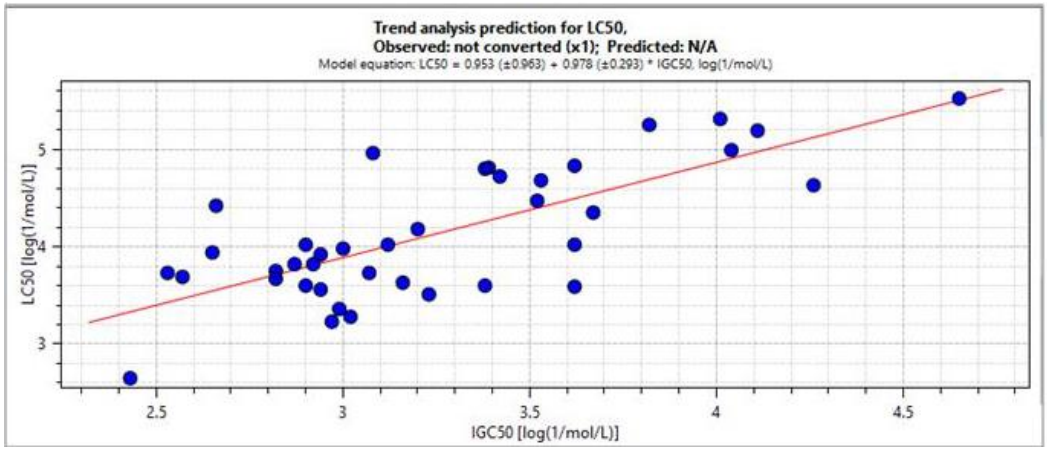


図 4

NOTE

The endpoint tree node which is selected in above manner will appear as parameter in the parameters list of the gap filling module and main options form. There is a specific distinction in the way the endpoint tree parameters work compared to the standard (2D/3D) parameters. While the standard parameters return the same value for a given chemical, the endpoint tree parameters return a value based on the experimental data currently available on the data matrix. As at any given moment different databases might be selected, the available experimental data for chemicals might also vary. For example, if the user applies a QSAR based on endpoint tree parameters, the prediction for a given chemical might vary depending on the current selection of databases at the moment of prediction.

［注意：上記の方法で選択されたエンドポイントツリーノードは、Gap filling モジュールとメインオプションフォームのパラメータリストに、パラメータとして表示されます。標準（2D / 3D）パラメータと比較して、エンドポイントツリーパラメータの作用は、明確な違いがあります。 標準的なパラメータは、所与の化学物質について同じ値を返しますが、エンドポイントツリーパラメータは、現在データ・マトリックス上で利用可能な実験データに基づいて値を返します。どのようなデータベースが選択されても、利用可能な実験データや化学物質は異なる可能性があります。 例えば、ユーザがエンドポイントツリーパラメータに基づいて QSAR を適用する場合、所与の化学物質の予測は、予測のその時点における現在のデータベースの選択に応じて変化する可能性があります。

D.5.4.2.4. Model/(Q)SAR （モデル/QSAR)

4 つのオプション（図 1）が含まれています。

- Show domain （ドメインの表示）
- Save model （モデルの保存）
- Save domain as category （カテゴリとしてドメインを保存）
- Calculate Q2（Q2 を計算する）

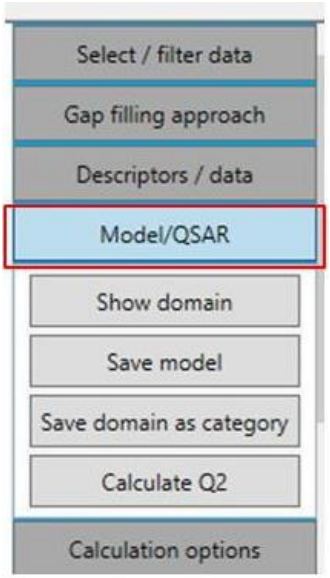


図 1

D.5.4.2.4.1.　Show domain　（ドメインの表示）

Show domain　(1)をクリックします（図 1）。

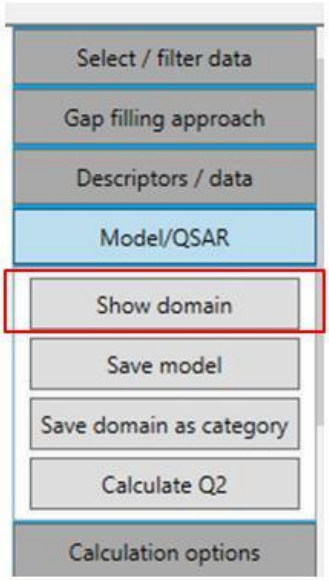


図 1

モデルのドメインが表示されます（図 2）。

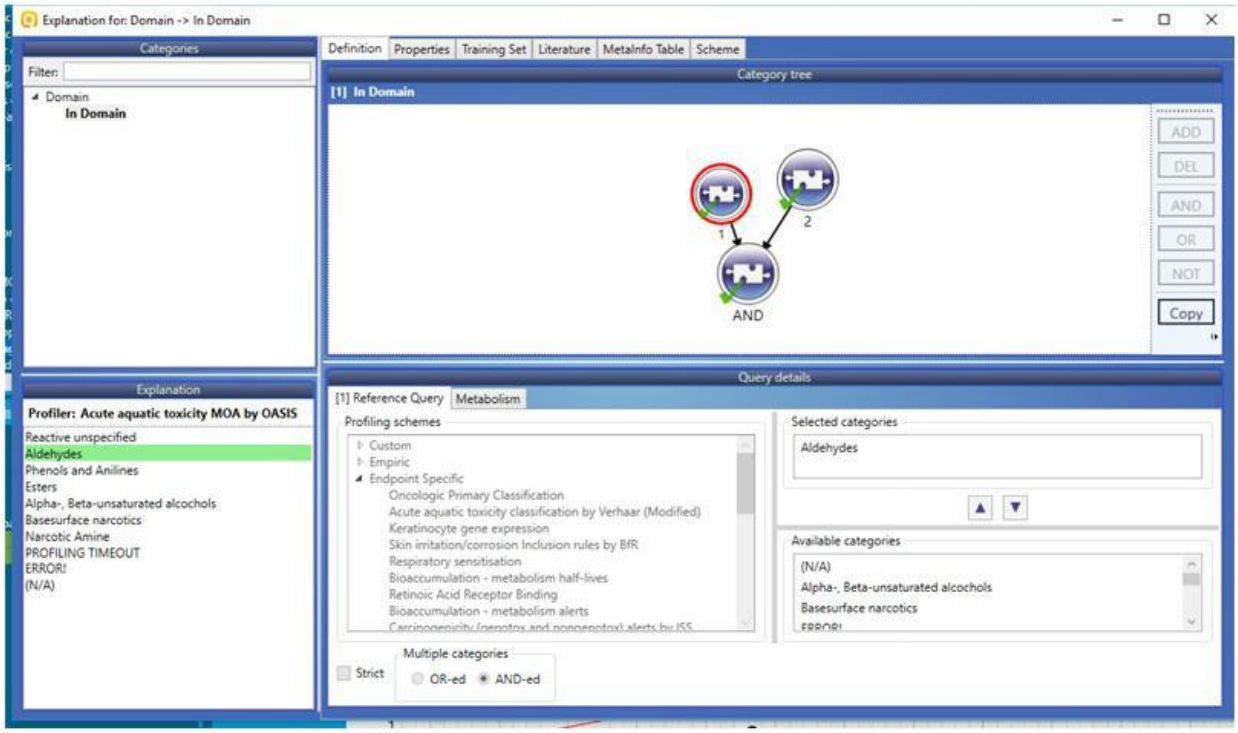


図 2

D.5.4.2.4.2. Save model （モデルの保存）

Save model をクリックします（図 1）。

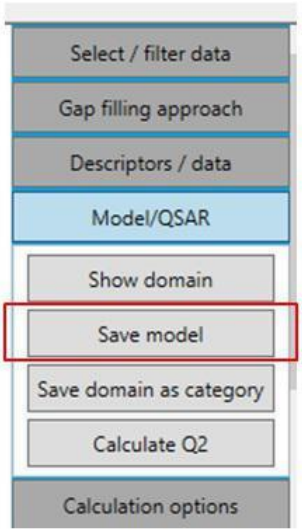


図 1

モデルに沿ったすべての情報を入力するフォームが表示されます（図 2）。
モデルとバージョンのタイトルは必須入力欄となっています（1）。Next (2)をクリックします。

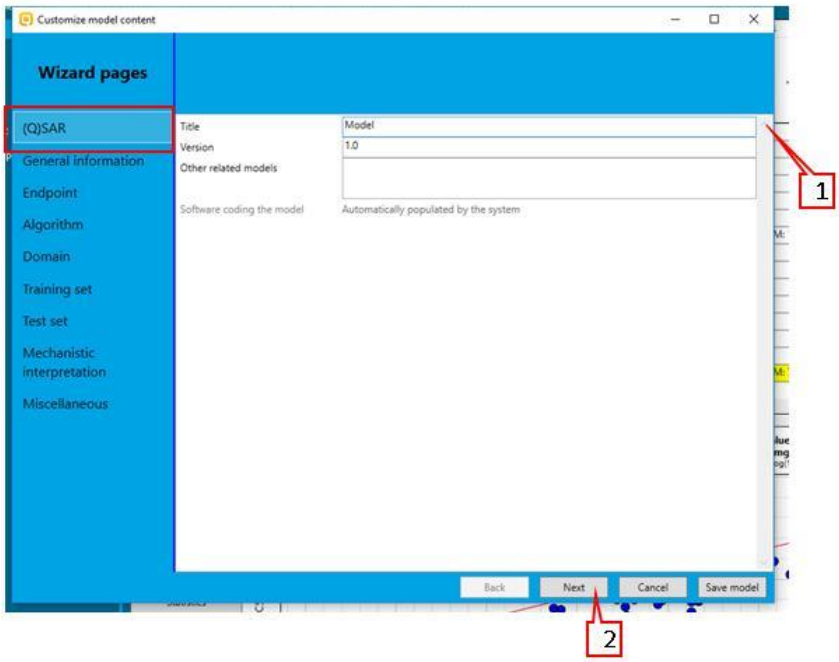


図 2

他のセクションの中には、ユーザ（1）が編集可能な一般情報（1）セクションもあります。すべての情報が入力されたら、*Save* (2)をクリックします（図 3）。確認メッセージが表示されます（ 3 ）。OK (2)をクリックします。

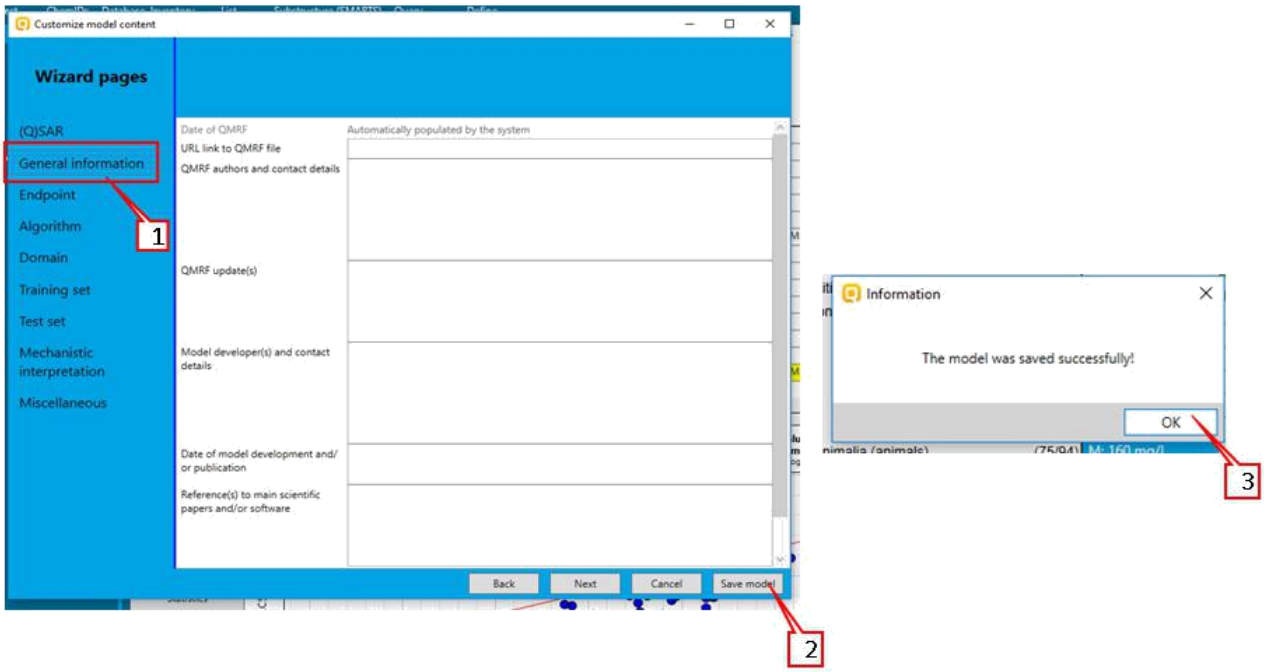

Customize model content
Wizard pages
(Q)SAR
General information
Endpoint
Algorithm
Domain
Training set
Test set
Mechanistic interpretation
Miscellaneous
Date of QMRF
Automatically populated by the system
URL link to QMRF file
QMRF authors and contact details
QMRF update(s)
Model developer(s) and contact details
Date of model development and/ or publication
Reference(s) to main scientific papers and/or software
Back
Next
Cancel
Save model
1
2
Information
The model was saved successfully!
OK
3

図 3

D.5.4.2.4.3. Save domain as category （カテゴリとしてドメインを保存）

Save domain as category (1) をクリックします（図 1)。カスタムのプロファイラが生成されます（2)。OK (3)をクリックします。

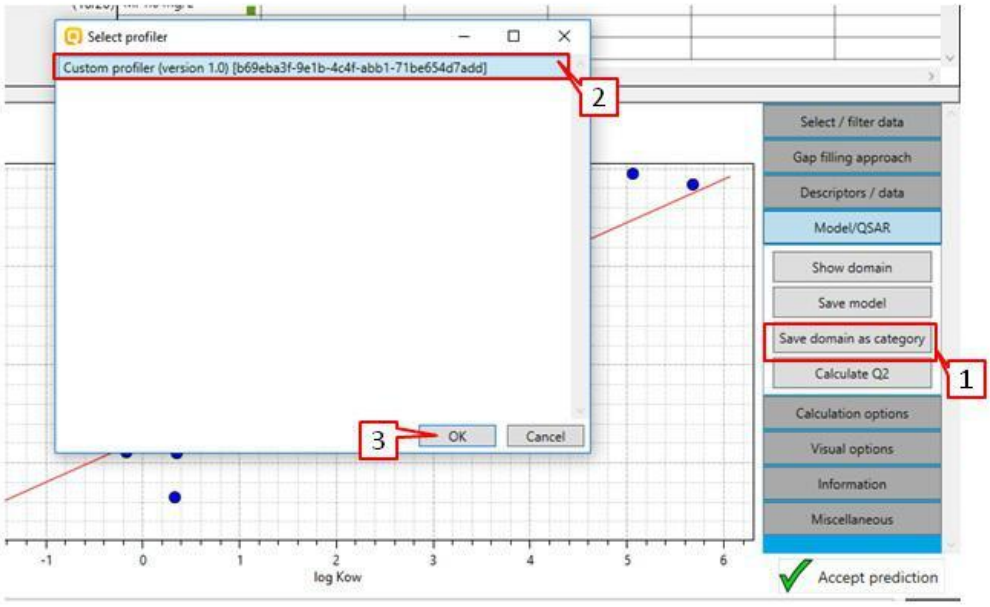


図 1

カテゴリ名（1）をタイプして、 OK （2）をクリックします。プロセスが成功したことを示すメッセージが表示されます（図 2)。OK (3)をクリックします。(図 2)

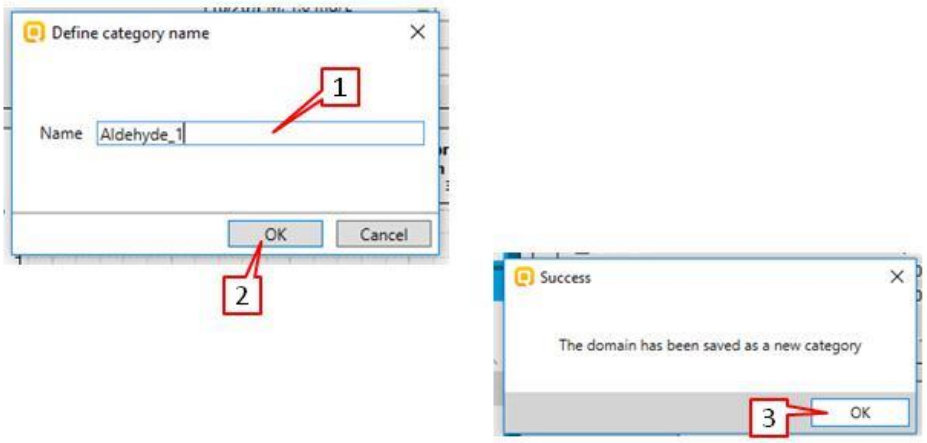


図 2

D.5.4.2.4.4.　Calculate Q2（Q2 を計算する）

（訳注：統計において Q2 はモデルの予測性能、R2 はモデルの直線性を表す）

このプログラムは 50 以上の類似物質を含むカテゴリの Q2 パラメータを手動（で追加して）計算することができます。これは、50 以上の類似物質からなるカテゴリの Q2 計算にかなりの時間を要するため、自動的には実行されないからです。

Calculate Q2 をクリックします（図 1）。

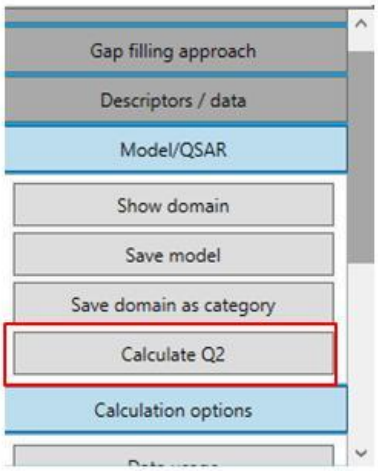
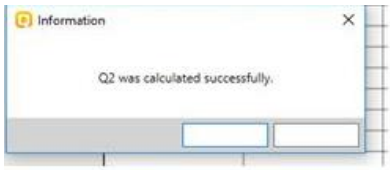


図 1

情報メッセージが表示されます（図 2）。*Statistics* (1)　をクリックし、Q2 値（2）を参照してください。



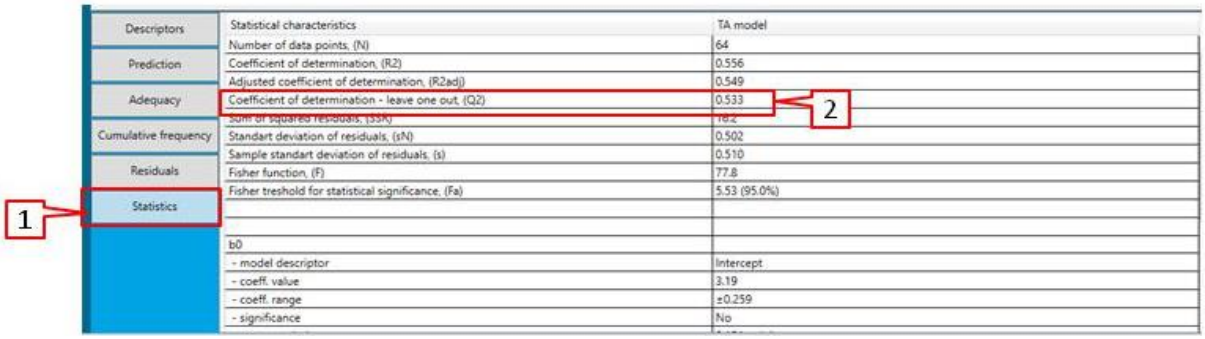


図 2

D.5.4.2.5.　Calculation options (計算オプション)

5 つの機能（図 1）が含まれています。

- Data usage（データの使用状況）
- Prediction approach options（予測手法のオプション）
- Use target data for prediction / Do not use target for prediction（予測に標的データを使う / 予測に標的データを使わない）
- Set level of significance（有意性のレベルを設定する）

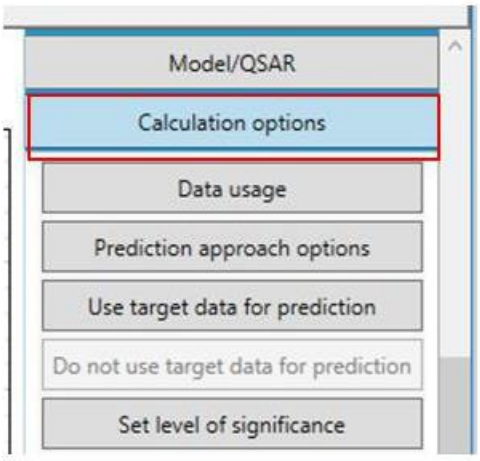


図 1

D.5.4.2.5.1. Data usage options（データの使用）

Data usage (1) をクリックすると 1 つの化学物質に割り当てられた複数の値(2)を処理する方法を選択します。(図 1)

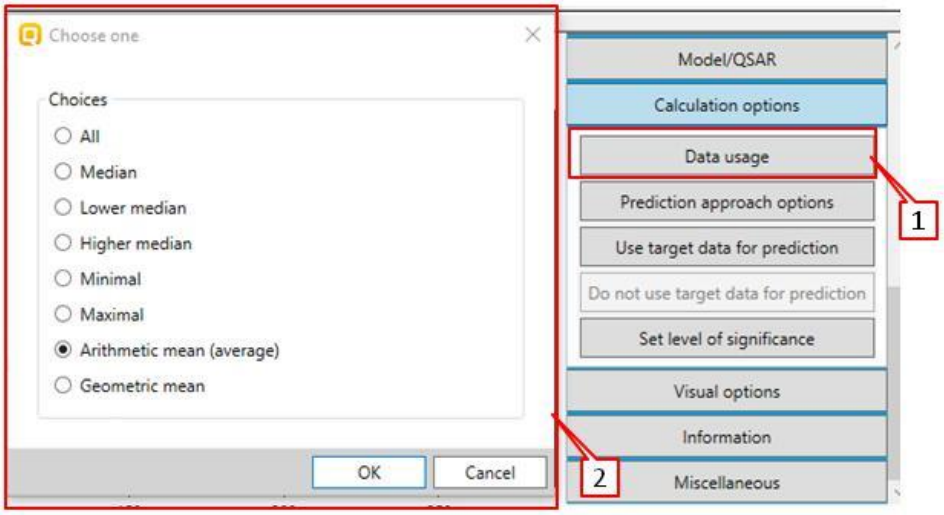


図 1

次の例では、データの使用方法とカテゴリのスケール（判定基準）での予測方法の各アプローチを示しています。ここでは選択されたエンドポイントについて、4 つの判定値をとりうる仮想判定基準と、2 つ以上の実測値を有する 3 つの化学物質を用意しています。

最初の図は、3 つの化学物質の実測値を表しています。

1) **DataUsage = All**、 すべてのデータ点が計算に使用される

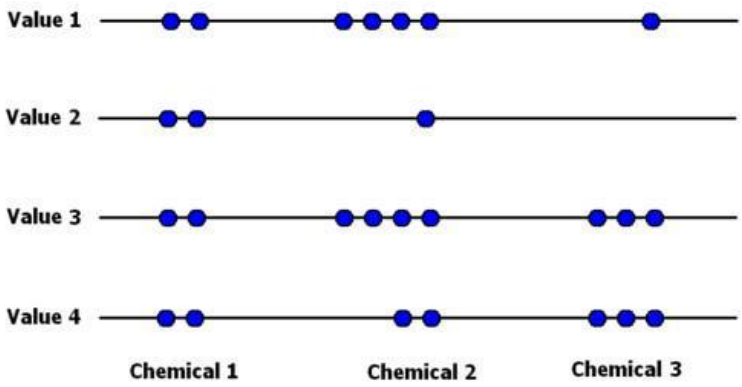


図 2 化学物質の実測値（すべてのデータ点を使用した場合の同じ画像）

次の図は、3 つの化学物質の再計算された値を表しています。実測値は空白とし、再計算された値は青で表示しています。

2) **DataUsage = Minimal**, 1 化学物質につき 1 点が与えられる

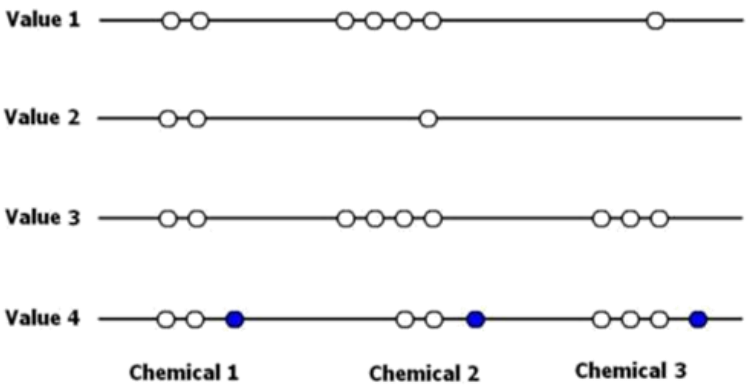


図 3　化学物質の再計算値（各化学物質の最小値を使用している場合）

3) **DataUsage** = **Maximal**,　1 化学物質につき 1 点が与えられる

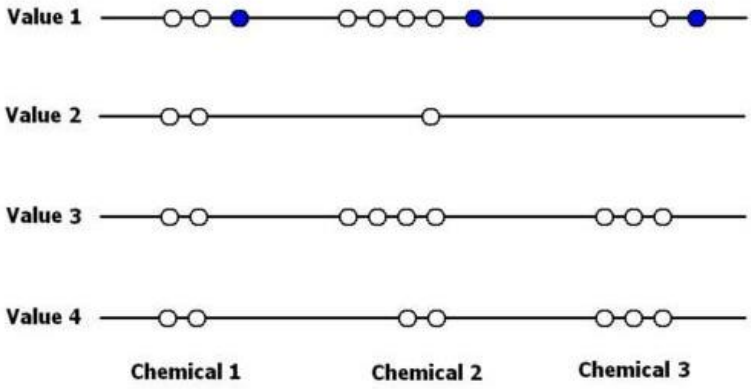


図 4　化学物質の再計算値（各化学物質の最大値を使用している場合）

4) **DataUsage** = **Median**（中央値）, Chemical 1 と Chemical 2 には 2 つのメジアン（中央値）があります。Chemical 3 は 1 つの中央値を持ちます（ここでは value 2 は考慮にいれません)。

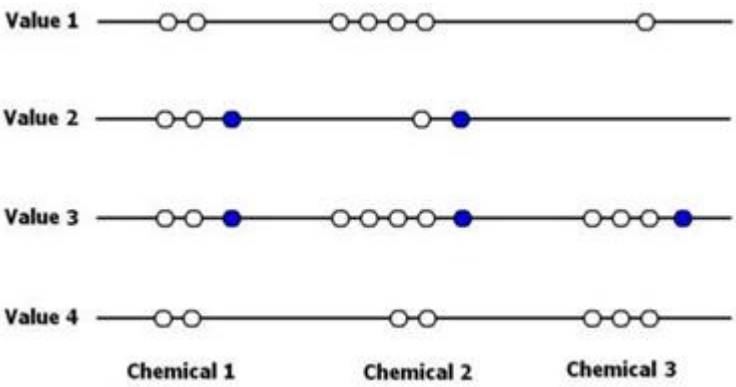


図 5　化学物質の再計算値（各化学物質の中央値を使用している場合）

5) **DataUsage = Lower median（低い方の中央値）**， １化学物質につき１点が与えられる

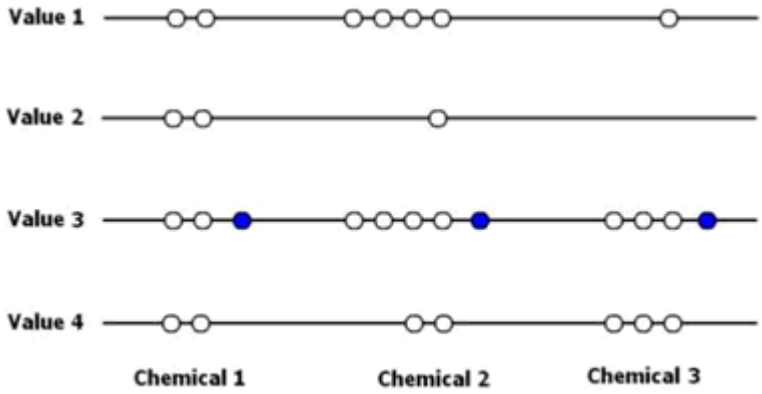


図 6　化学物質の再計算値（各化学物質の低い方の中央値を使用している場合）

6) **DataUsage = Higher median（高い方の中央値）**， １化学物質につき１点が与えられる

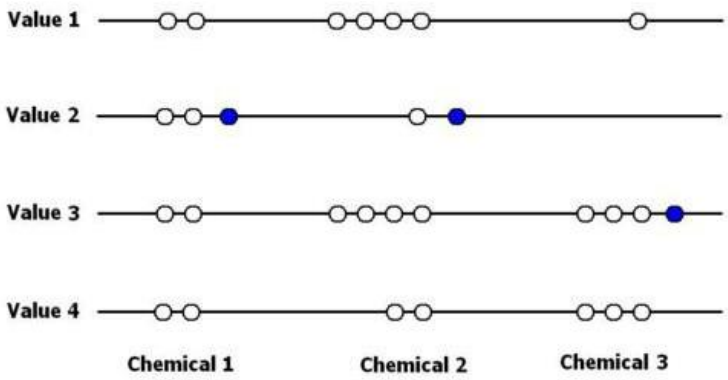


図 6　化学物質の再計算値（各化学物質の低い方の中央値を使用している場合）

D.5.4.2.5.2. Prediction approach options（予測手法のオプション）

Prediction approach options（1）をクリックすると、予測近似方法が設定されます。トレンドアナリシス（2）については linear （１次曲線近似）と quadratic（２次曲線近似)、リード・アクロス（3）については minimal（最小）, maximal（最大）, average（平均）, median（中央値）, lower median（低い方の中央値）, higher median（高い方の中央値）, mode（最頻値）, lowest mode（低い方の最頻値）そして highest mode（高い方の最頻値）があります（図1)。

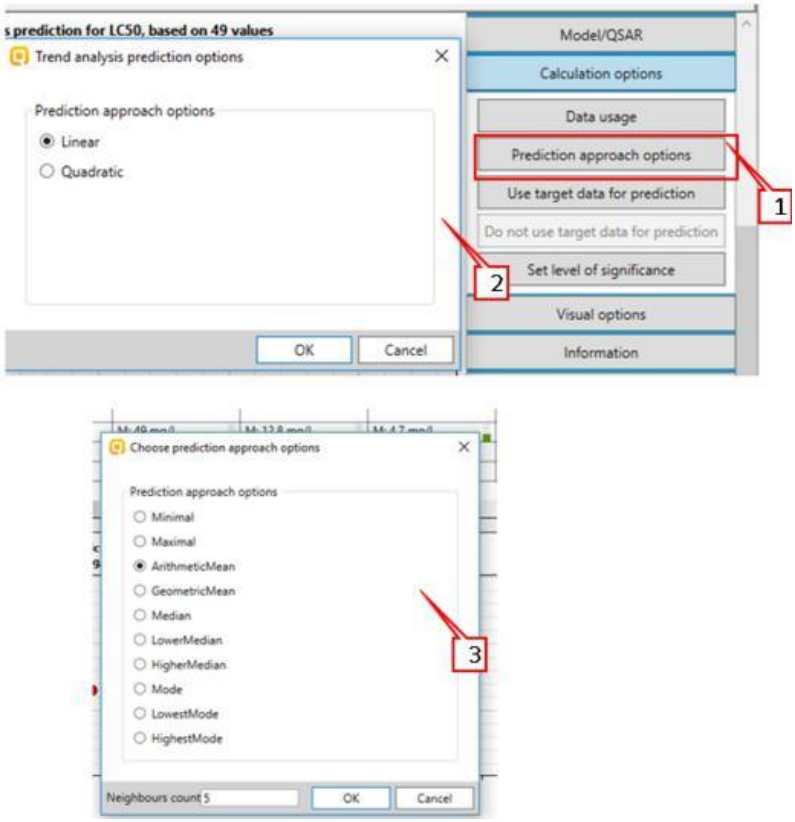


図 1

予測の例

Chemical 2 と Chemical 3 が、予測を決定する近傍（類似物質）であるとしましょう。（訳注：前ページの図の Chemical1～3 を参照のこと）上記で示した prediction approach options の様々なケースを以下に例示します。

1) **DataUsage = All** では、予測値は次のとおりです。

- **Value 4**, 近似タイプが “Minimal” の場合
- **Value 1**, 近似タイプが “Maximal” の場合
- **No value**, 近似タイプが “Median” の場合は値がありません - Value2 と Value 3 は両方とも中央値であるため、システムは自動的に決定を行うことはできません
- **Value 3**, 近似タイプが “Lower median” の場合
- **Value 2**, 近似タイプが “Higher median” の場合

- **Value 3**, 近似タイプが “Mode”, “Lowest mode” または “Highest mode” の場合 -この値には 7 つの近傍点が使用できます。この場合は 1 つの最頻値しか使用できないため、最後の 3 つの近似タイプでは同じ予測値を与えます。

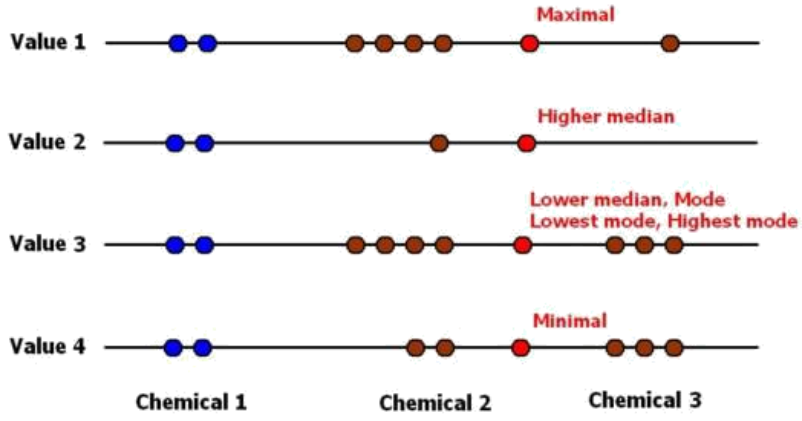


図 1　　すべてのデータ点を使用した場合の予測値

2)　　**DataUsage = Minimal**　では、予測値は次のとおりです。

- **Value 4**,すべての近似タイプに対して

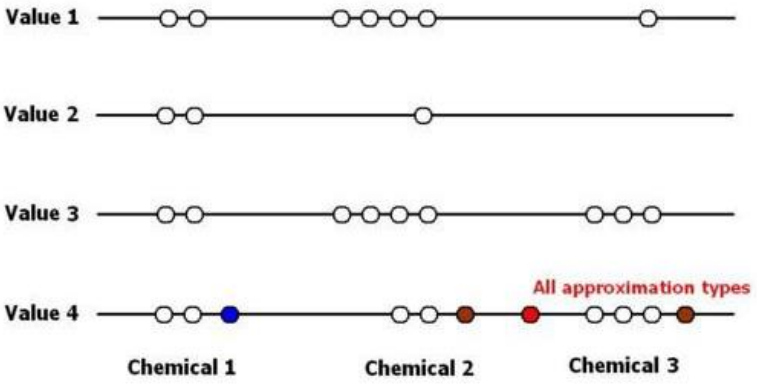


図 2　化学物質の再計算値（各化学物質の最小値を使用している場合）

3) **DataUsage = Maximal**　では、予測値は次のとおりです。

- **Value 1**,すべての近似タイプに対して

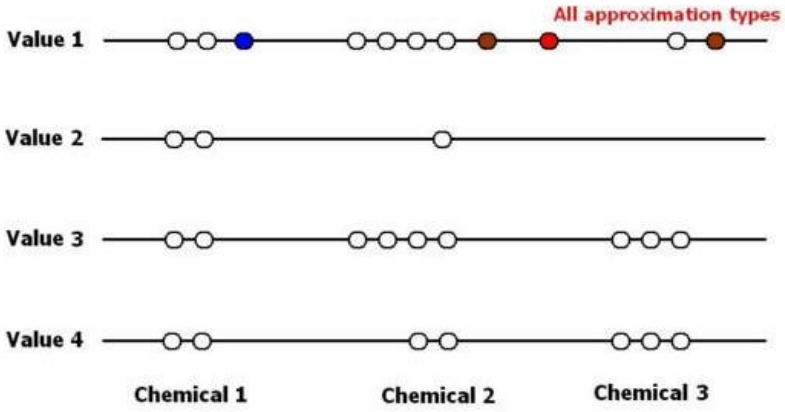


図 3　化学物質の再計算値（各化学物質の最大値を使用している場合）

4) **DataUsage = Median** では、予測値は次のとおりです。

- **Value 3**, 近似タイプが「最小」の場合
- **Value 2**, 近似タイプが「最大」の場合
- **No value**, 近似タイプが“Median”の場合は値がありません - Value 2 と Value 3 は両方とも中央値であるため、システムは自動的に決定を行うことはできません
- **Value 3**, 近似タイプが“Lower median”の場合
- **Value 2**, 近似タイプが “Higher median”の場合
- **Value 3**, 近似タイプが “Mode”, “Lowest mode” または “Highest mode” の場合 -この値には 2 つの近傍点が使用できます。この場合は 1 つの最頻値しか使用できないため、最後の 3 つの近似タイプでは同じ予測値を与えます。

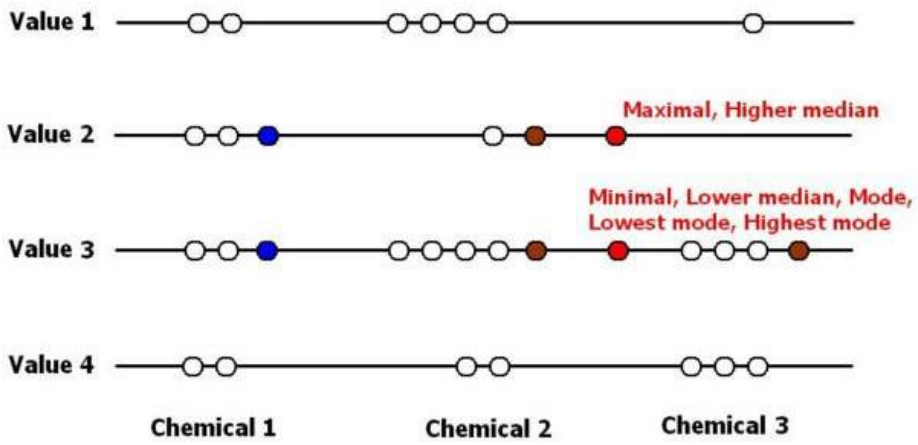


図 4 化学物質の再計算値（各化学物質の中央値を使用している場合）

5) **DataUsage = Lower Median** では、予測値は次のとおりです。**Value 3**,すべての近似タイプに対して

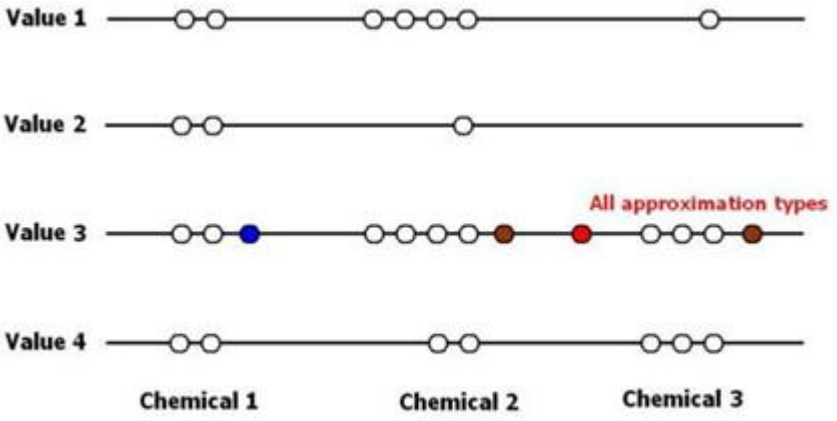


図 5 化学物質の再計算値（各化学物質の低い方の中央値を使用している場合）

6) **DataUsage = Higher Median** では、予測値は次のとおりです。

- **Value 3**, 近似タイプが「最小」の場合
- **Value 2**, 近似タイプが「最大」の場合
- **No value**, 近似タイプが“Median”の場合は値がありません - Value 2 と Value 3 は両方とも中央値であるため、システムは自動的に決定を行うことはできません
- **Value 3**, 近似タイプが“Lower median”の場合

- **Value 2**, 近似タイプが “Higher median”の場合
- **No value**, 近似タイプが“Mode”の場合は値がありません - Value2 と Value 3 は両方とも最頻値であるため、システムは自動的に決定を行うことはできません
- **Value 3**, 近似タイプが“Lowest mode”の場合
- **Value 2**, 近似タイプが “Highest mode”の場合

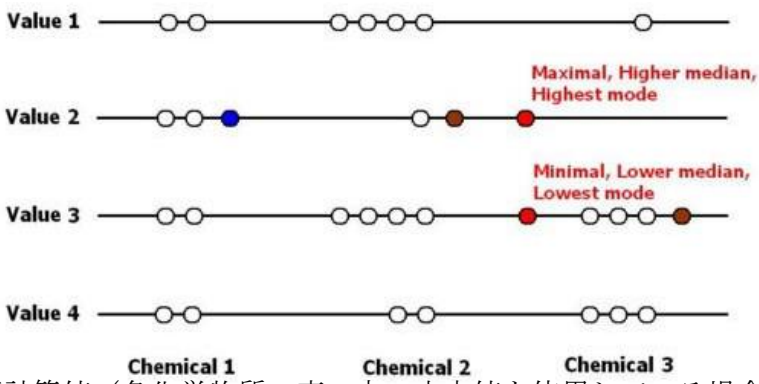


図 6　化学物質の再計算値（各化学物質の高い方の中央値を使用している場合）

7) **DataUsage = Mode**　では、予測値は次のとおりです。

- **Value 4**, 近似タイプが「最小」の場合
- **Value 1**, 近似タイプが「最大」の場合
- **Value 3**, 近似タイプが　“Median”, “Lowest median” または “Highest median” の場合 - この場合は 1 つの中央値しか使用できないため、この 3 つの近似タイプでは同じ予測値を与えます。
- **Value 3**, 近似タイプが　“Mode”, “Lowest mode” または “Highest mode” の場合 -この値には 2 つの近傍点が使用できます。この場合は 1 つの最頻値しか使用できないため、最後の 3 つの近似タイプでは同じ予測値を与えます。

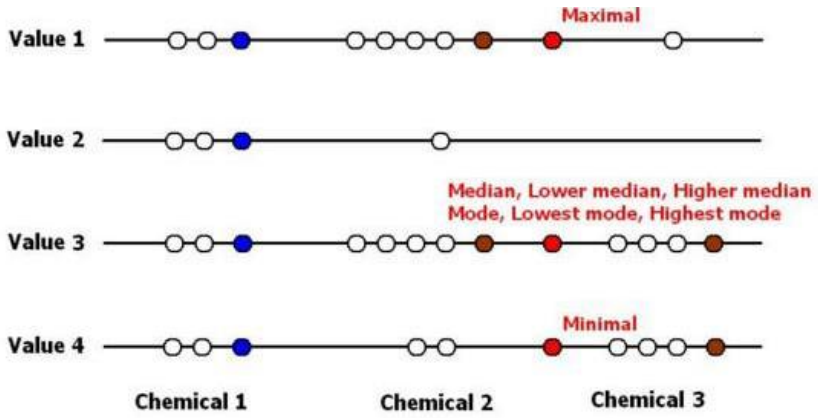


図 7　化学物質の再計算値（各化学物質の最頻値を使用している場合）

8) **DataUsage = Lowest mode**　では、予測値は次のとおりです。

- **Value 4**, 近似タイプが「最小」の場合
- **Value 3**, 近似タイプが「最大」の場合
- **No value**, 近似タイプが“Median”の場合は値がありません - Value 3 と Value 4 は両方とも中央値であるため、システムは自動的に決定を行うことはできません

- **Value 4**, 近似タイプが“Lower median”の場合
- **Value 3**, 近似タイプが “Higher median”の場合
- **No value**, 近似タイプが“Mode”の場合は値がありません - Value 3 と Value 4 は両方とも最頻値であるため、システムは自動的に決定を行うことはできません
- **Value 4**, 近似タイプが“Lowest mode”の場合
- **Value 3**, 近似タイプが“Highest mode”の場合

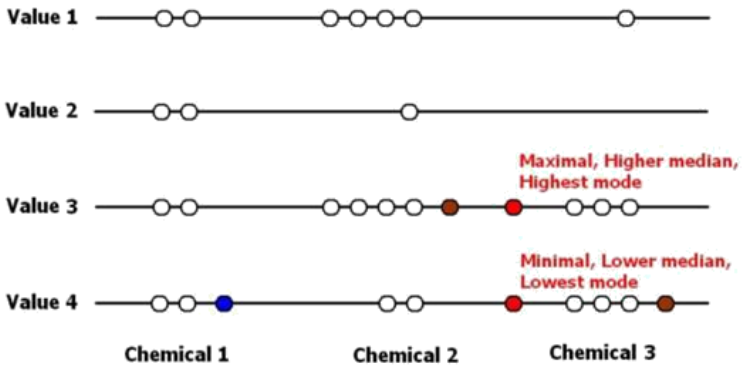


図 8　化学物質の再計算値（各化学物質の低い方の最頻値を使用している場合）

9) **DataUsage** = **Highest mode**　では、予測値は次のとおりです。
- **Value 3**, 近似タイプが「最小」の場合
- **Value 1**, 近似タイプが「最大」の場合
- **No value**, 近似タイプが“Median”の場合は値がありません - Value 1 と Value 3 は両方とも中央値であるため（Value 2 はここでは考慮に入れません）、システムは自動的に決定を行うことはできません

- **Value 3**, 近似タイプが“Lower median”の場合
- **Value 1**, 近似タイプが “Higher median”の場合
- **No value**, 近似タイプが“Mode”の場合は値がありません - Value 1 と Value 3 は両方とも最頻値であるため、システムは自動的に決定を行うことはできません
- **Value 3**, 近似タイプが“Lowest mode”の場合
- **Value 1**, 近似タイプが“Highest mode”の場合

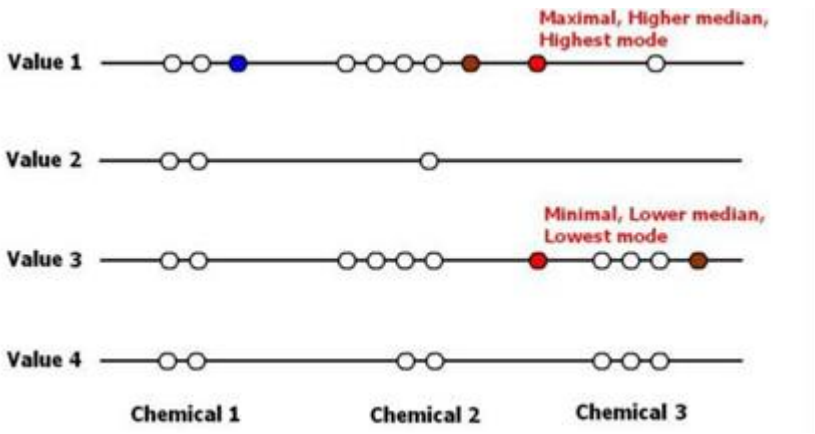


図 9　化学物質の再計算された値（各化学物質の高い方の最頻値を使用している場合）

D.5.4.2.5.3. Use Target Data for prediction（予測に標的物質のデータを使用する）

This functionality allows using (or not using) the experimental data of target, if available, in the data gap filling calculation (1 in Figure 1).

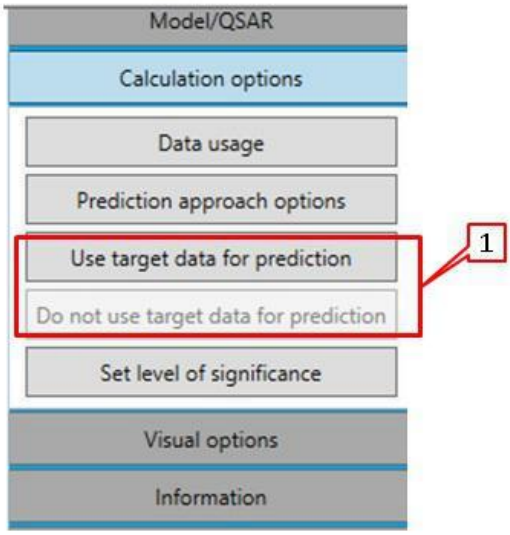


図 1

D.5.4.2.5.4. Set level of significance（有意水準を設定する）

Set level of significance - confidence level （信頼水準）と standard deviation（標準偏差）を設定します。

Set level of significance （1）をクリックすると新しいウィンドウが開き、デフォルトの有意水準（0.9500）を変更することができます（2）（図 1）。OK (3)をクリックします。

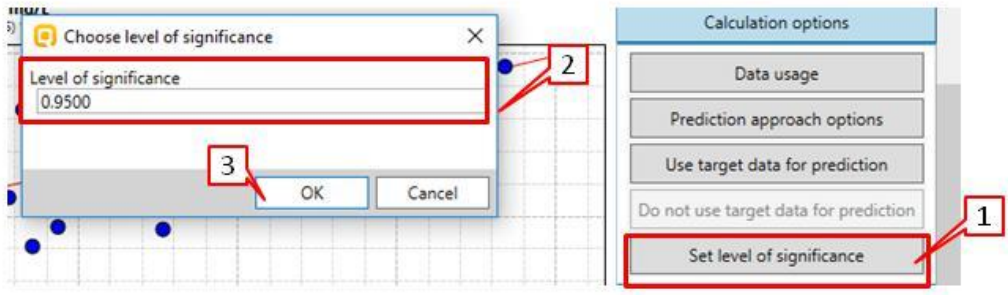


図 1

D.5.4.2.6. Visual options （グラフ表示のオプション）

4 つの機能（図 1）が含まれています。

- Set units in figure title（図のタイトルに単位を設定する）
- Axes range tool（軸の範囲ツール）
- Confidence range tool（信頼範囲ツール）
- Intercorrelation tool（自己相関ツール）

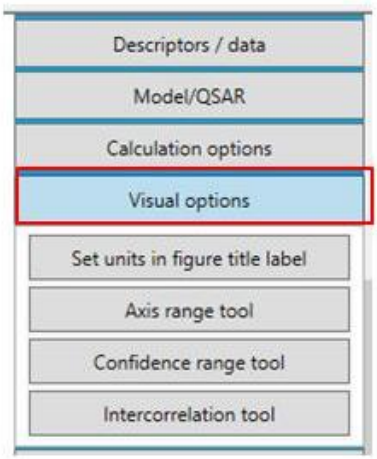


図 1

D.5.4.2.6.1. Set units in figure title(図のタイトルに単位を設定する)

Set units in figure title - グラフ表示のオプションを設定します。(図 1)

Set units in figure title をクリックします。（1）単位（2）と式（3）が、図のタイトル（4）の単位を決定します。*Accept* (5)をクリックします。

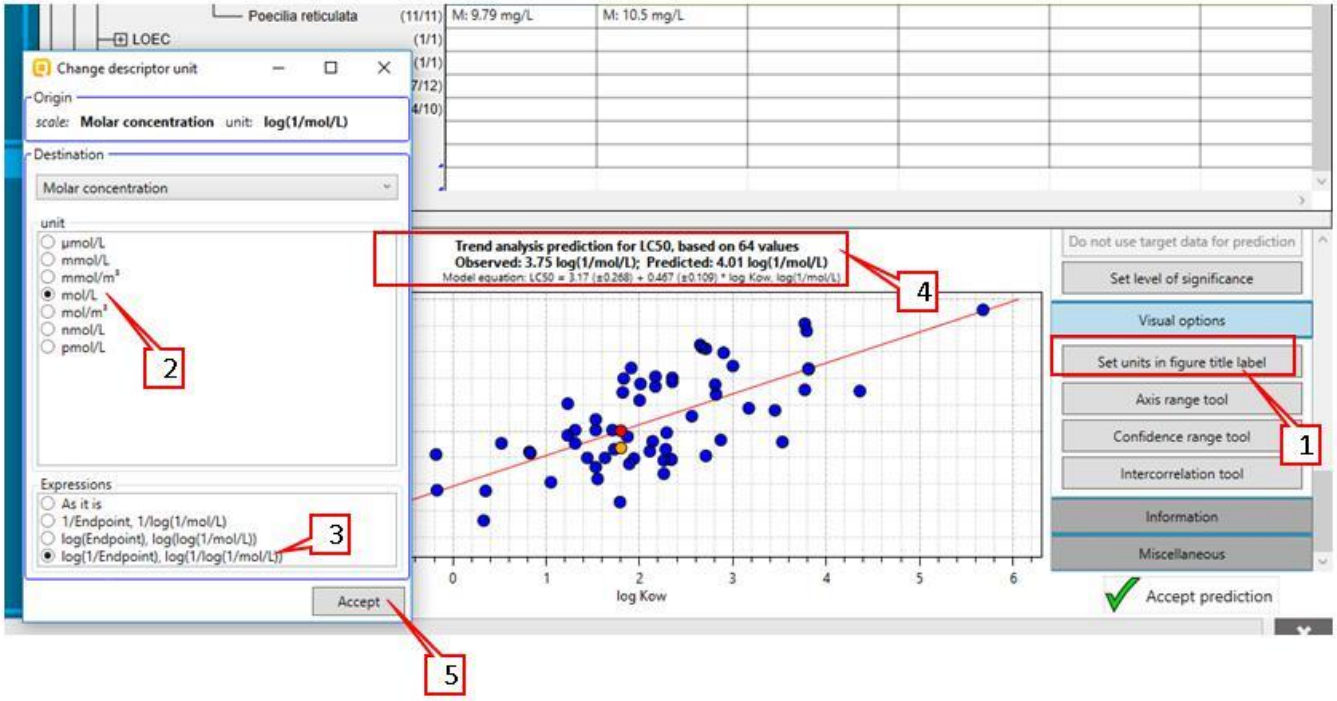


図 1

D.5.4.2.6.2.　Axis range tool（軸の範囲ツール）

Axis range tool - X 軸 と Y 軸の範囲を手動で設定することができます。(図 1)

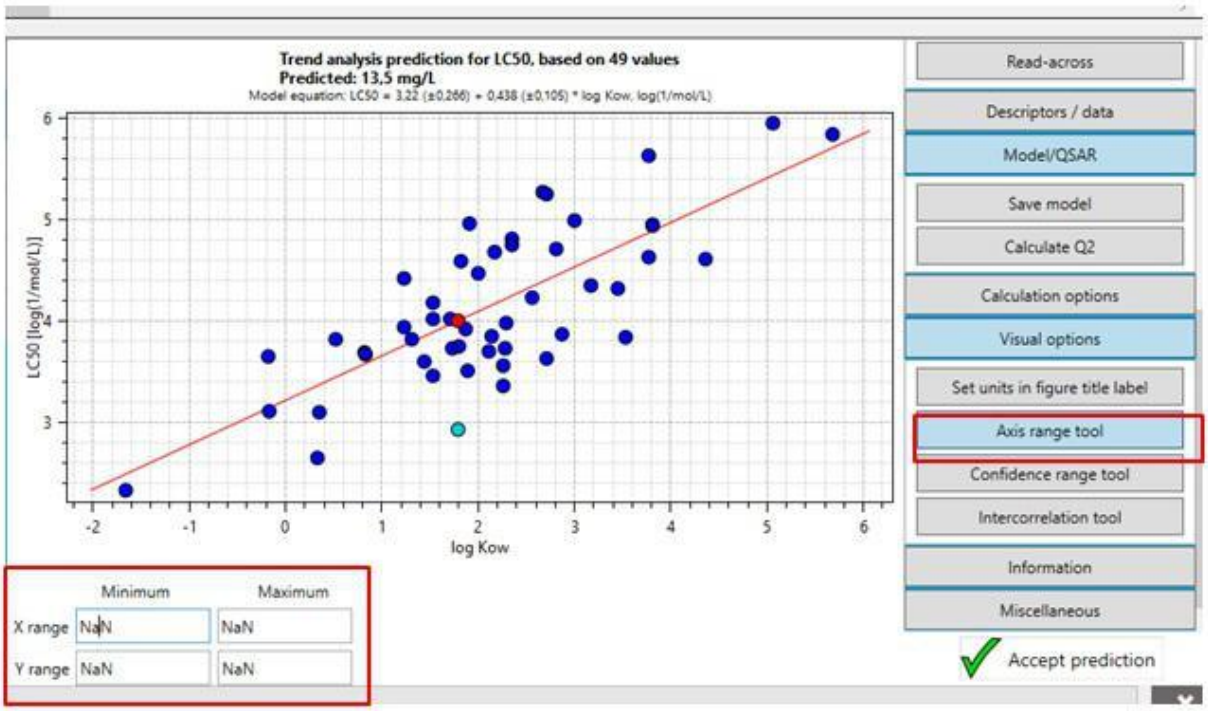


図 1

D.5.4.2.6.3. Confidence range tool（信頼範囲ツール）

Confidence range tool - 予測パネル（1）の信頼範囲を表示/非表示します。内側の範囲は回帰式（2）の信頼範囲を示し、外側の範囲は一度の予測（3）の信頼範囲を示します（図 1）。

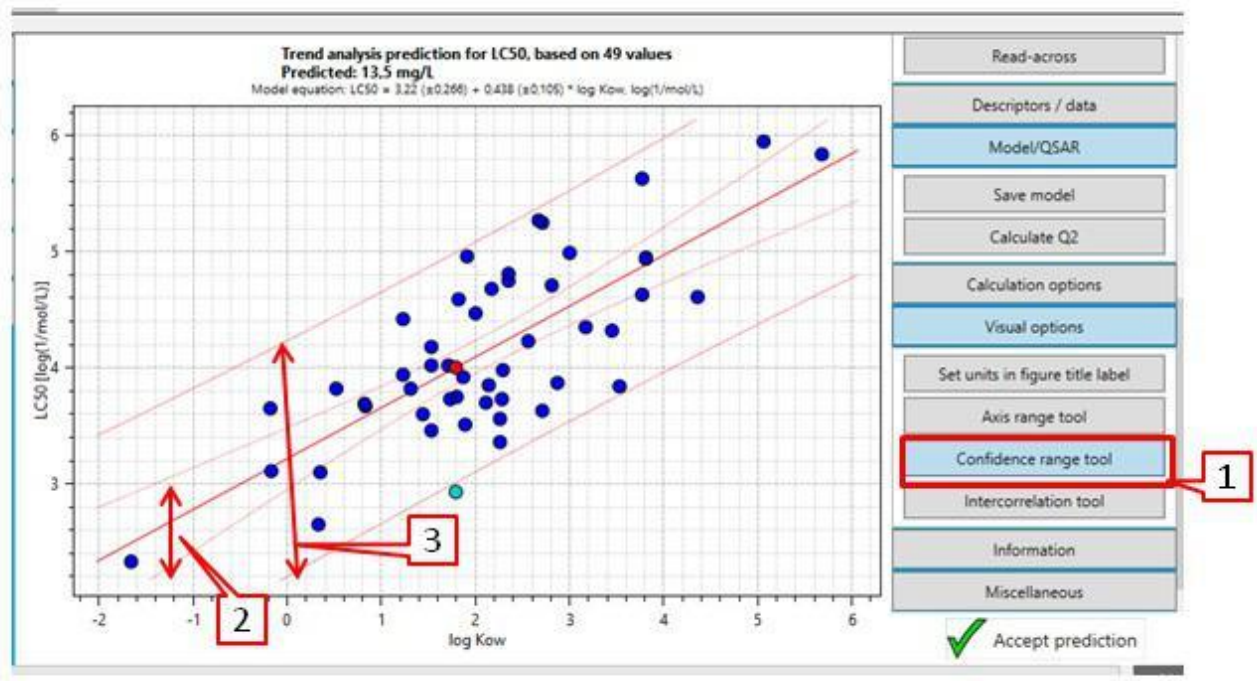


図 1

D.5.4.2.6.4. Intercorrelation tool（相互相関ツール）

Intercorrelation tool（1） - 相互相関パネルを示します。このボタンをクリックすると、グラフ上の点が消え、ユーザは推定したいと思う相関について 2 つのグラフ軸記述子（2）を選択することができます。グラフ軸記述子を選択したら、Intercorrelation tool（3）をクリックしてください。この相関関係は、グラフ（4）上に表示されます（図 1）。

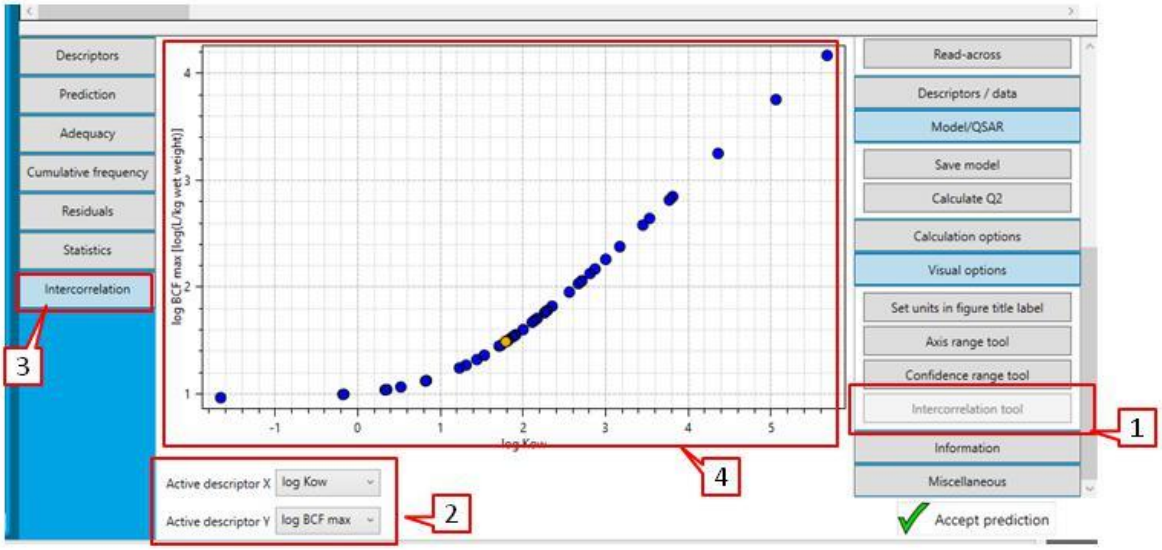


図 1

D.5.4.2.7. Information（情報）

Information （情報）

6 つの機能が含まれています。

- Details of focused chemical（注目する化学物質の詳細）
- Details of target chemical（標的化学物質の詳細）
- Details of focused points（注目するデータ点の詳細）
- Differences to target（標的物質との違い）
- Region selection tool（領域の選択ツール）
- Legend tool（凡例ツール）

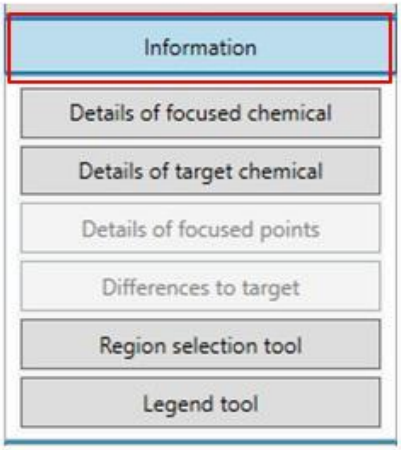


図 1

D.5.4.2.7.1. Details of focused chemical（注目する化学物質の詳細）

Details of focused chemical (1) - 選択されたデータポイントの化学物質に関する詳細（2）を示します。このウィンドウ（3）には、Chem ID 情報、化学構造、計算された記述子を含むパネル、付随データを含むパネルが含まれています。(図 1)

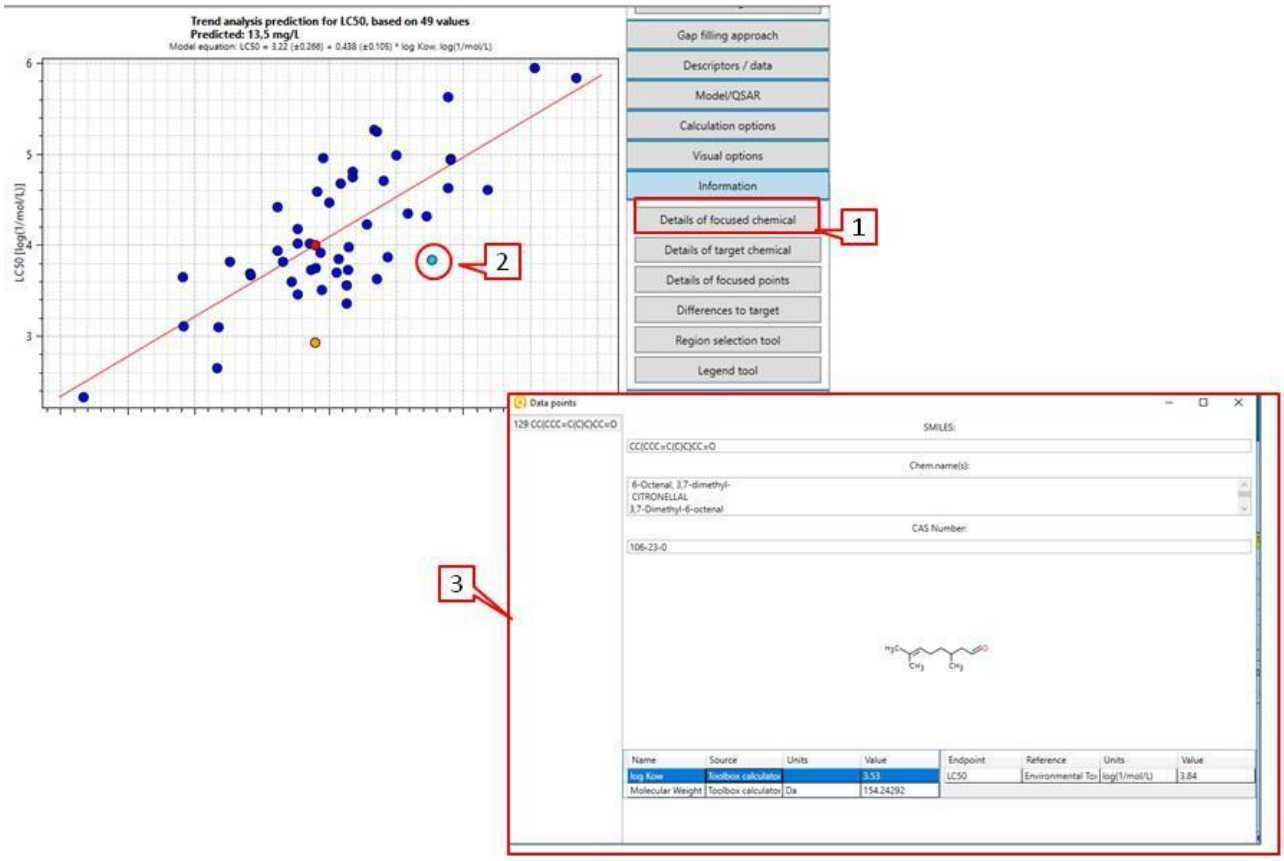


図 1

D.5.4.2.7.2. Details of target chemical（標的化学物質の詳細）

Details of target chemical (1) - 標的化学物質（2）に関する追加情報を示します（図 1）。

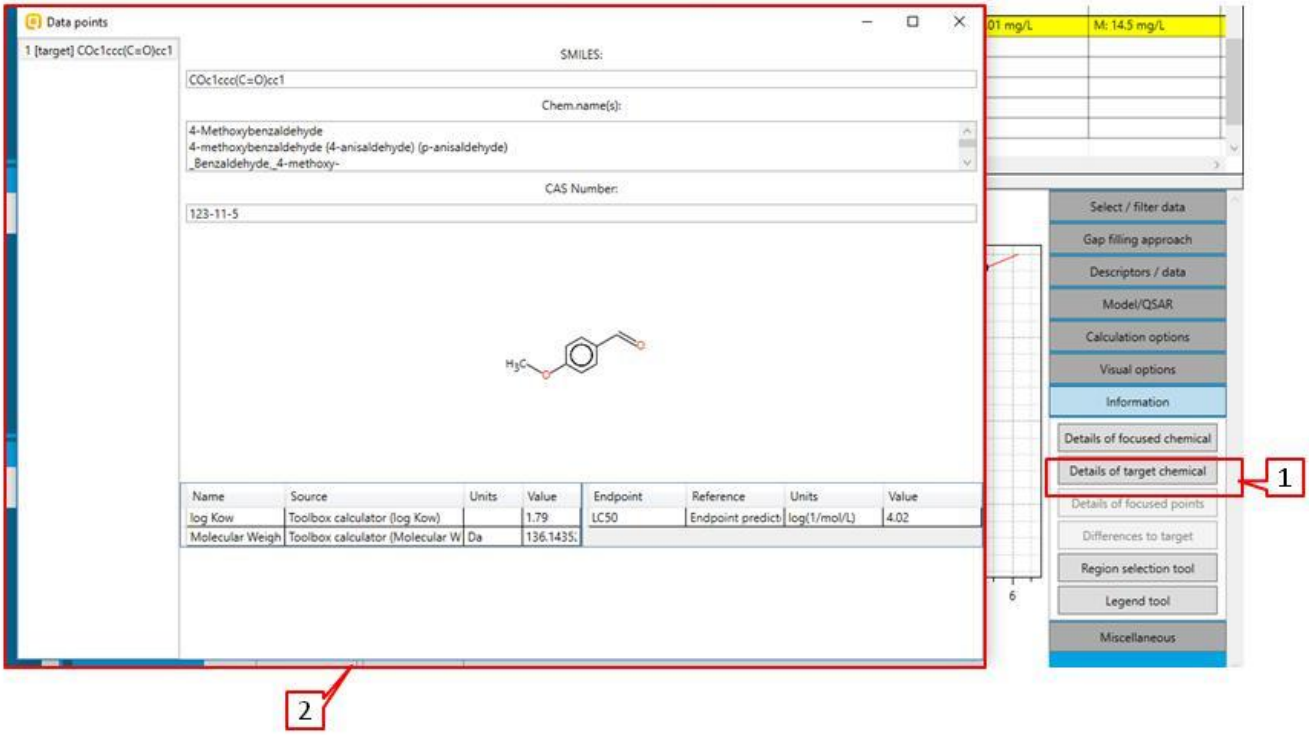


図 1

D.5.4.2.7.3. Details of focused points（注目するデータ点の詳細）

Details of focused points（注目するデータ点の詳細）

最初に、 *Region selection tool* （1）（訳注：この使い方は D.5.4.2.7.5 章を参照）を使用してグラフ上のいくつかのポイントを選択します（2）（図 1）。

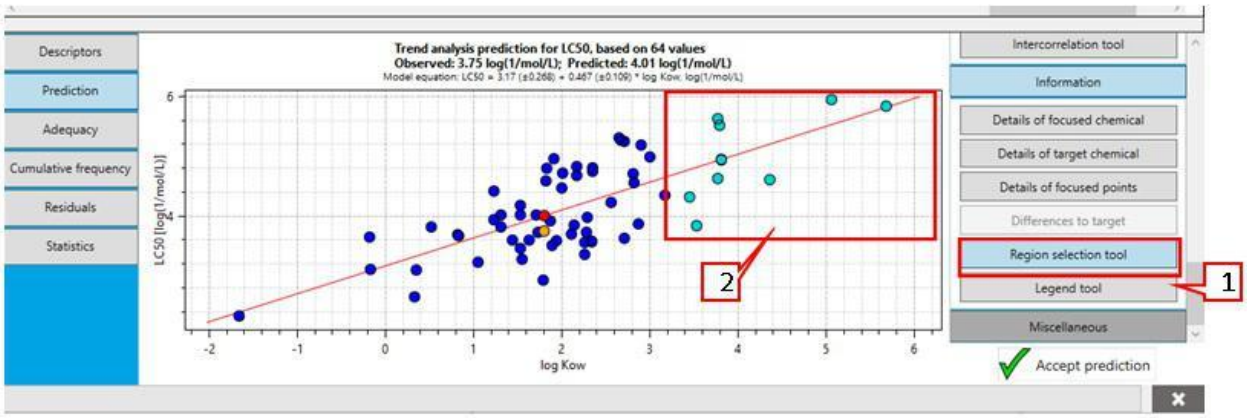


図 1

Details of focused points （1）をクリックして Data points ウィンドウ（2）を表示し、SMILES のリストから化学物質を選択してその付随データを表示することができます。(図 2)

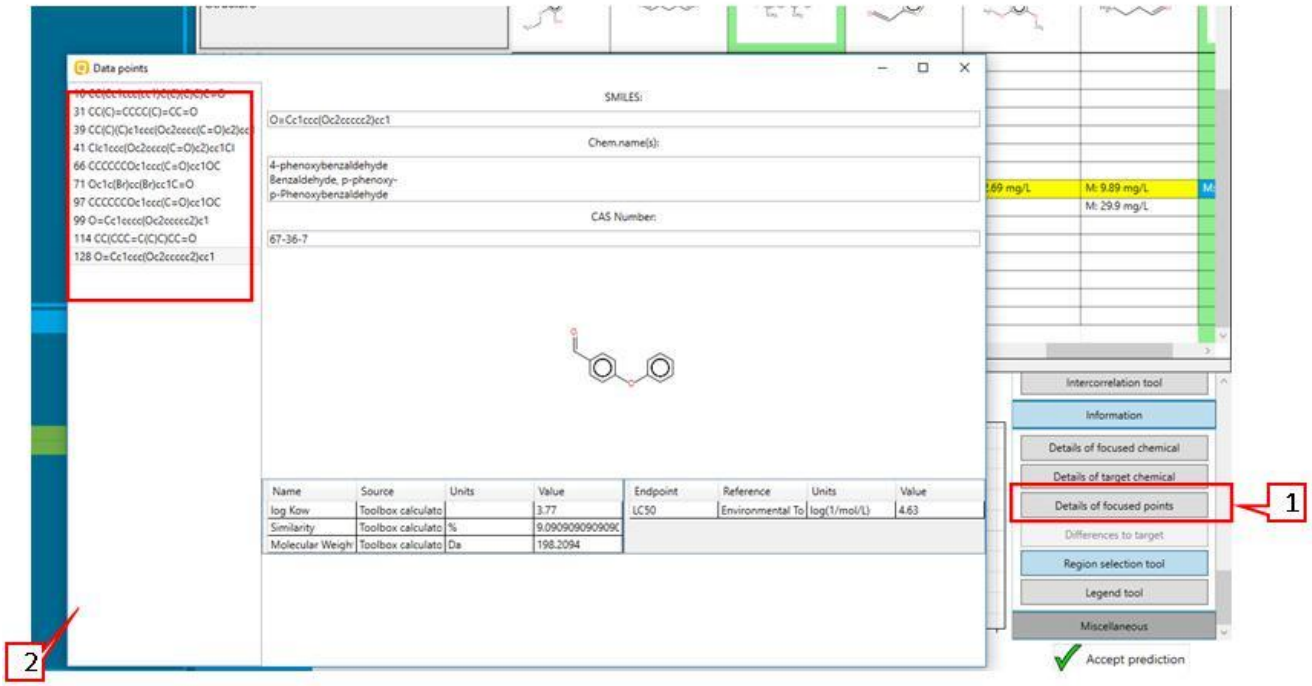


図 2

D.5.4.2.7.4. Differences to target（標的物質との違い）

Differences to target（標的物質との違い）

このオプションは、使用可能なすべてのプロファイリング方法において、選択したデータポイント(2)と標的化学物質との相違点を示します。(訳注：この作業前に、全体画面左側の Document ツリーで調査しようとする Gap Filling の行、“Enter GF(TA) with..."をクリックしておくことに留意。途中で別の作業をすると Document ツリーで別の行をクリックした状態になり、調査が進まなくなる。)

Differences to target を押すと自動で Subcategorization の小画面が開き、図 1 はその状態である)

Options を展開します（3)。

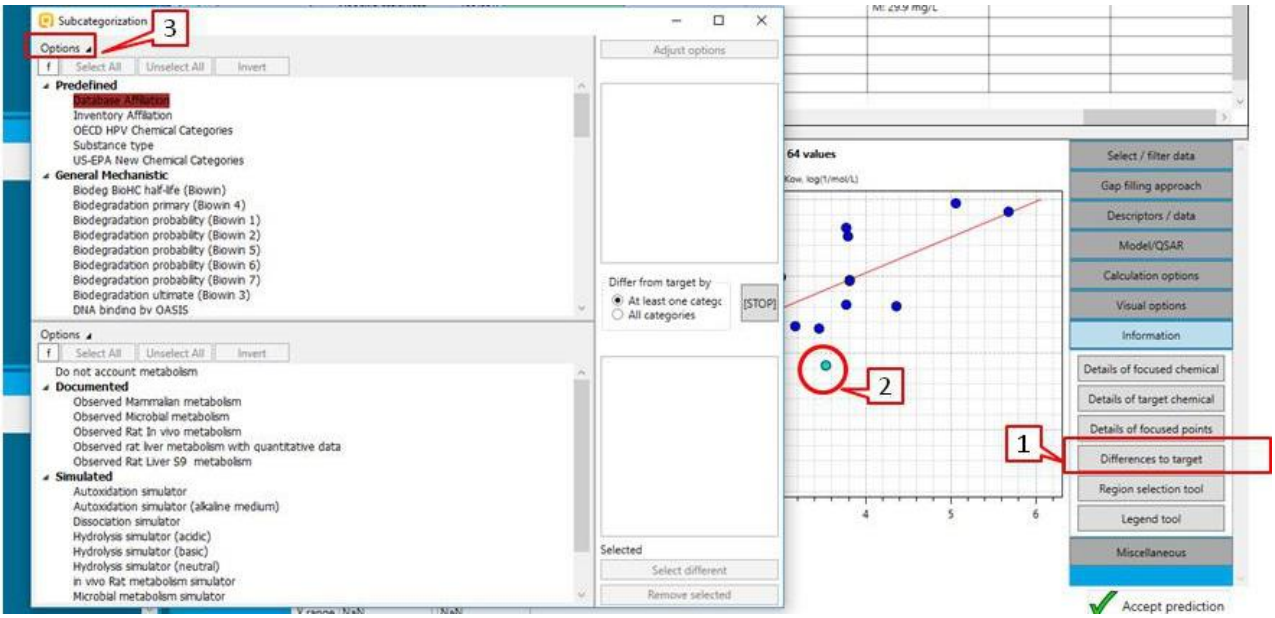


図 1

Color by Different (1), *Group by* Different (2) を選択します。(図 2) 多少の違いがあるプロファイリング方法がグループ化され、濃赤に色付けされます。

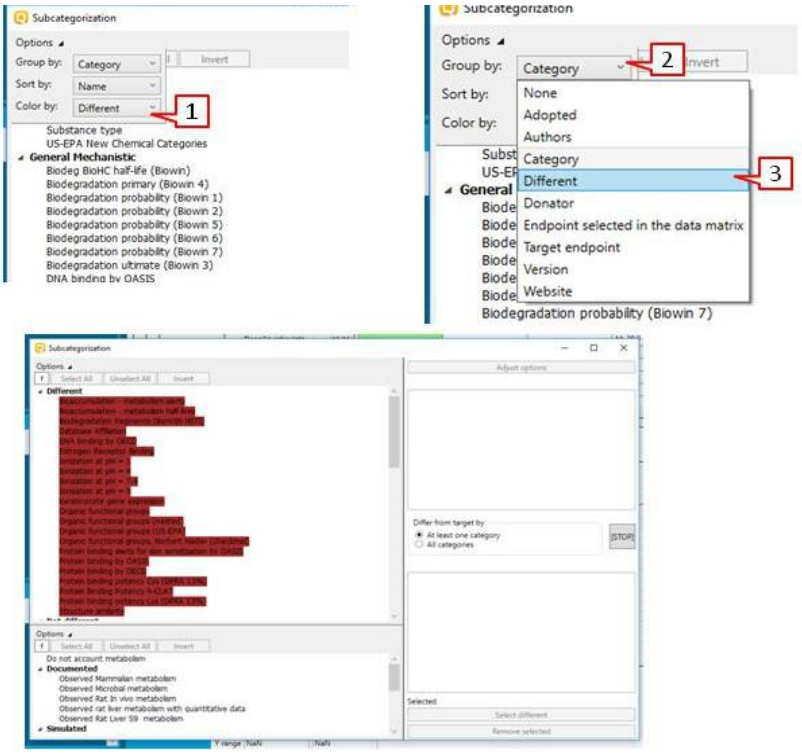


図 2

プロファイル（1）を選択すると、標的物質（2）と類似物質（3）がプロファイルされます（図3）。標的物質と異なる類似物質のプロファイリング結果は青色で表示されます（3）。

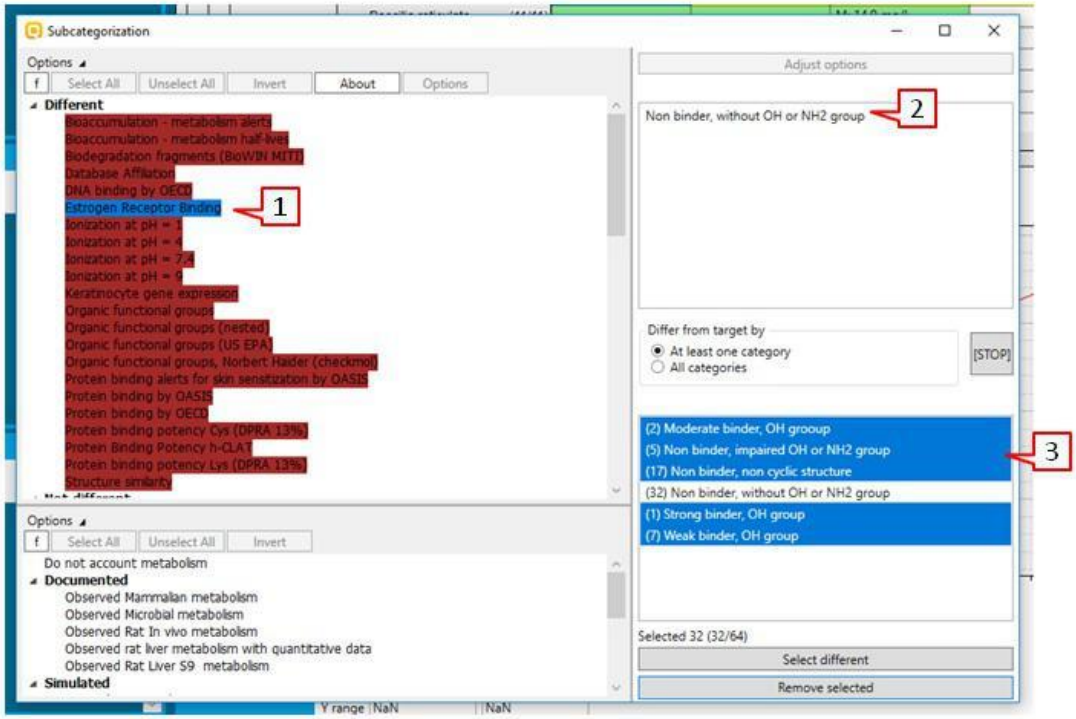


図 3

D.5.4.2.7.5. Region selection tool（グラフ内の領域の選択ツール）

Region selection tool（グラフ内の領域の選択ツール）

このオプションは、カテゴリ内の化学物質が同じ logKow を持つ場合、つまり、データポイントがグラフで二重に重なる場合に使用されます。この機能により、ユーザは 1 つの領域内のすべての化学物質（データ点）を見ることができます。*Region Selection tool* (1)をクリックしてから、マウス（左マウスボタン）をドラッグして長方形の領域（2）を指定してください（図 1)。その後、注目されている各データ点に右クリックオプションを使用します（Right-click options を参照)。

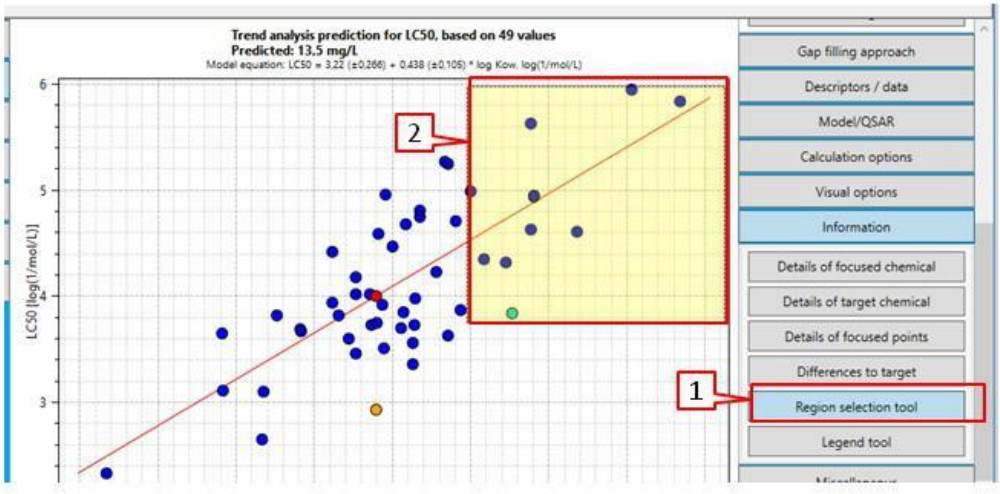


図 1

D.5.4.2.7.6. Legend tool（凡例ツール）

Legend tool (1) - チャートの凡例（2）を表示/非表示します。(図 1)

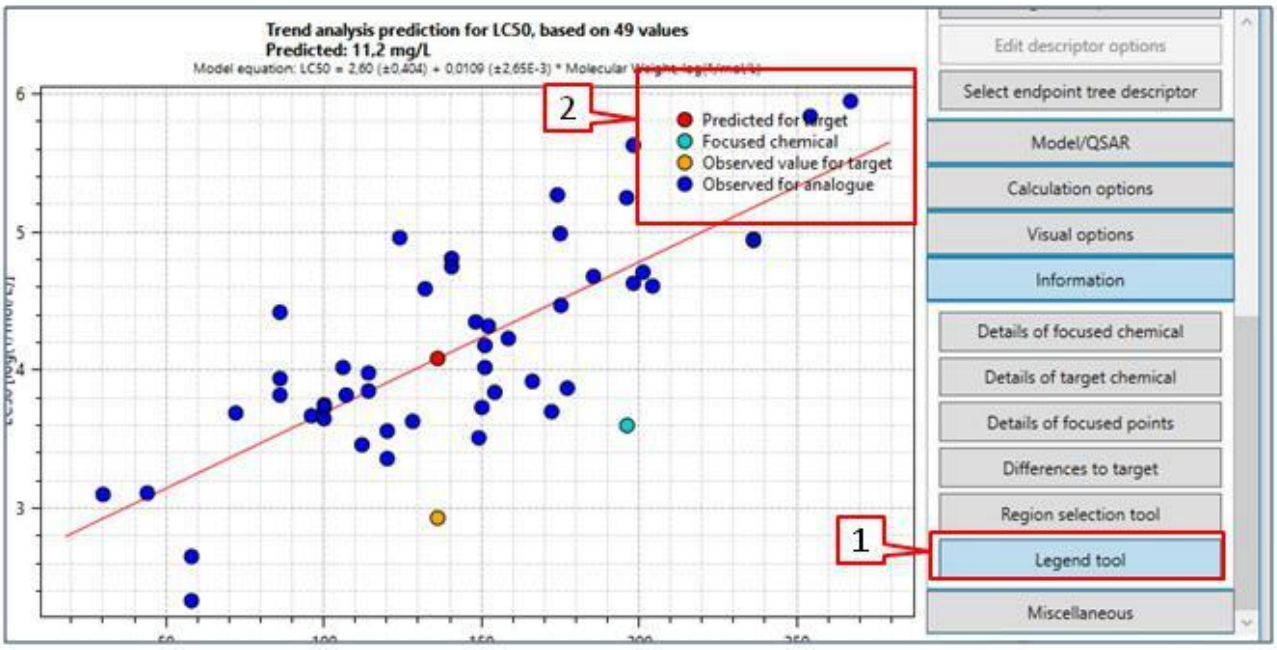


図 1

D.5.4.2.8. Miscellaneous （その他）

Miscellaneous （その他）

Miscellaneous (1) (Figure 1) : *Save picture* (2) オプションが実装されています。ファイル名（3）を選択し、保存（4）してください。

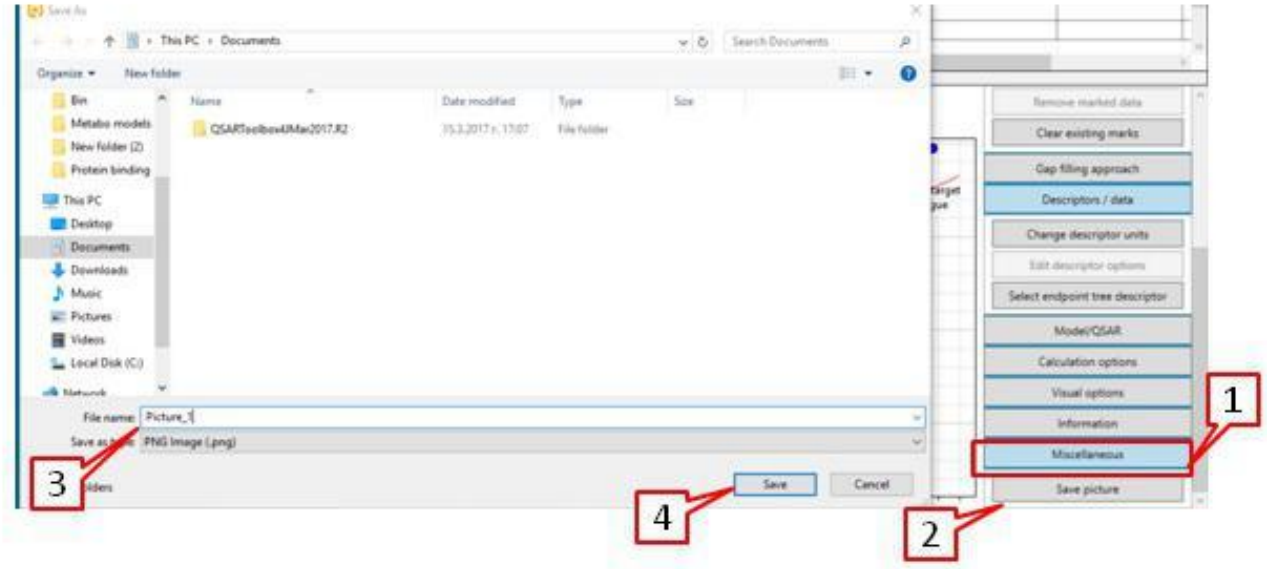


図 1

D.5.4.3. Right click functionality （右クリック機能）

3 つのギャップ補完方法でもう 1 つ共通する機能は、図 1 の右クリックメニュー（1）です。選択したデータ点を右クリックすると、データギャップ補完のメニューにいくつかのオプションが付きます。

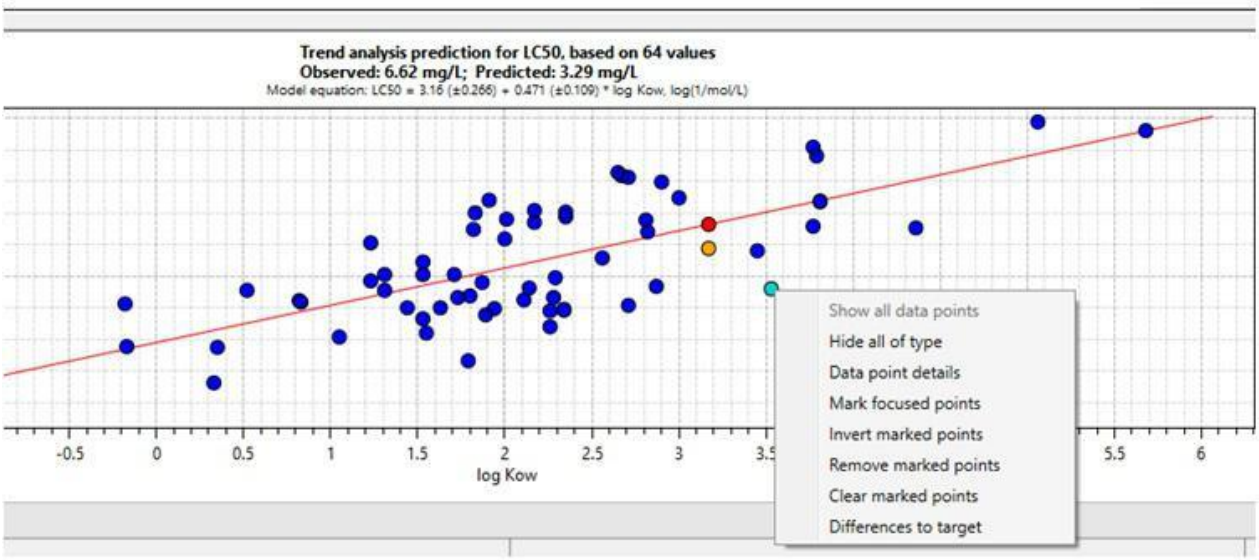


図 1

D.5.5.（Q）SAR の特色

（Q）SAR の特色

（訳注:以下の作業の前にに Input モジュールで標的化学物質を入力しておくこと。下の例では CAS 番号 123-11-5 が入力されている） エンドポイントツリーを展開し、サブノード、たとえば Aquatic Toxicity (1)を選択し、QSAR アイコン（2）をクリックします。（図 1）

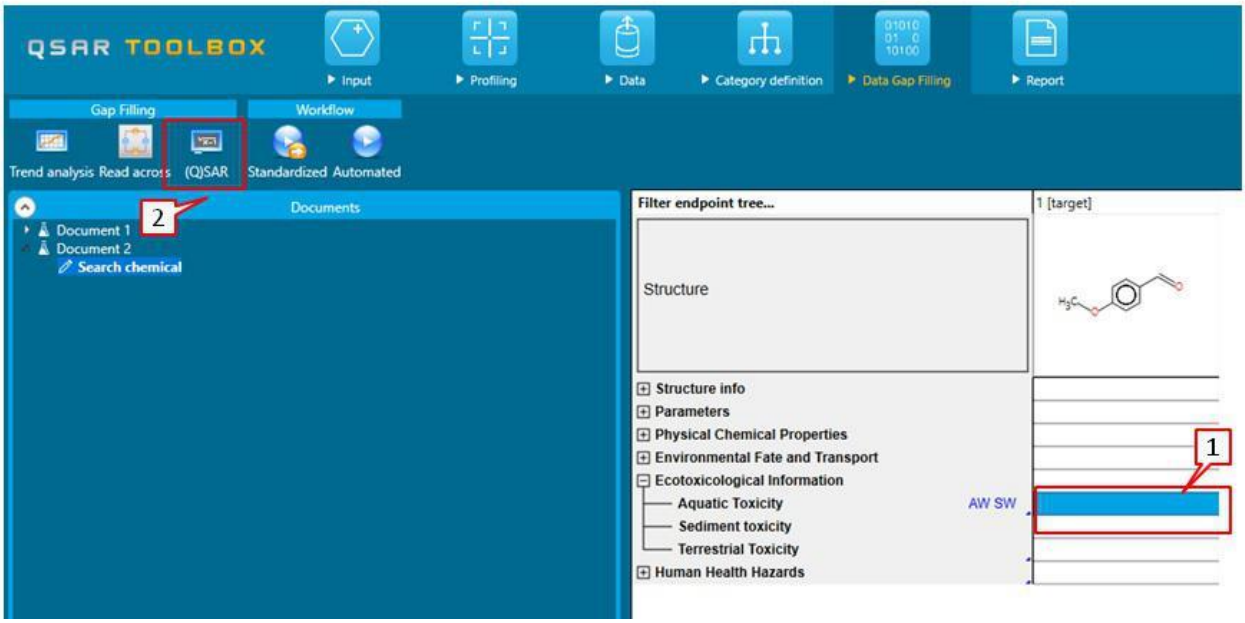


図 1

このレベル（1）で利用可能な QSAR は 674 あります（2）（図 2）。

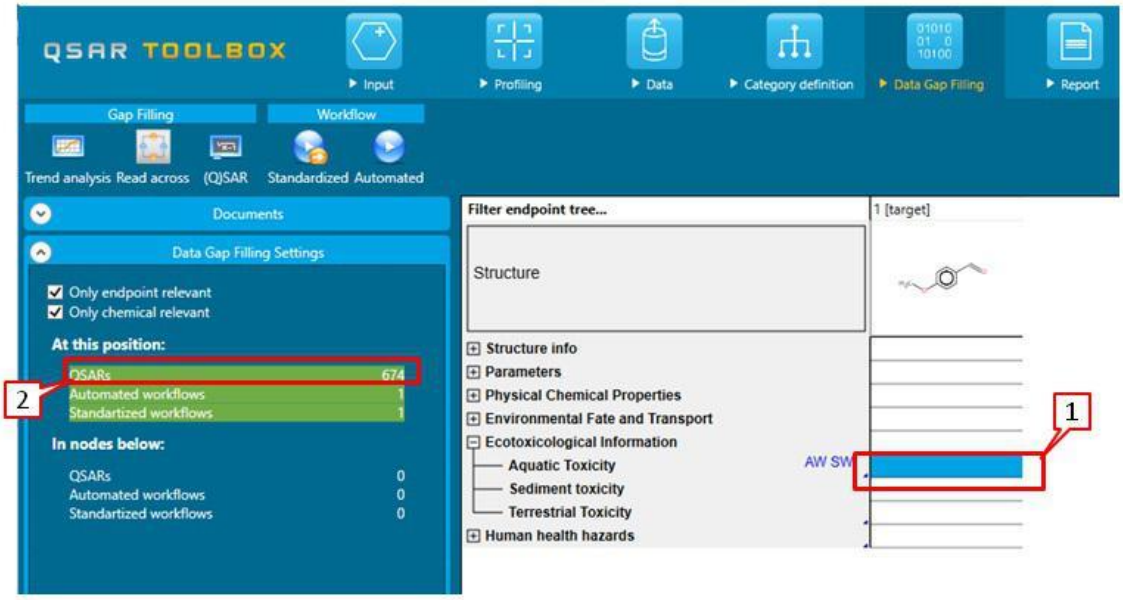


図 2

Only endpoint relevant（1）をオフにすると、エンドポイントに関係なく、ソフトウェア内の利用可能な QSAR がすべて数字（2）で表されます。（図 3）

（訳注：ツールボックスのバージョン 4.1 ではこのボックスの黒いチェックが図 2 のように表示されない不具合がある）

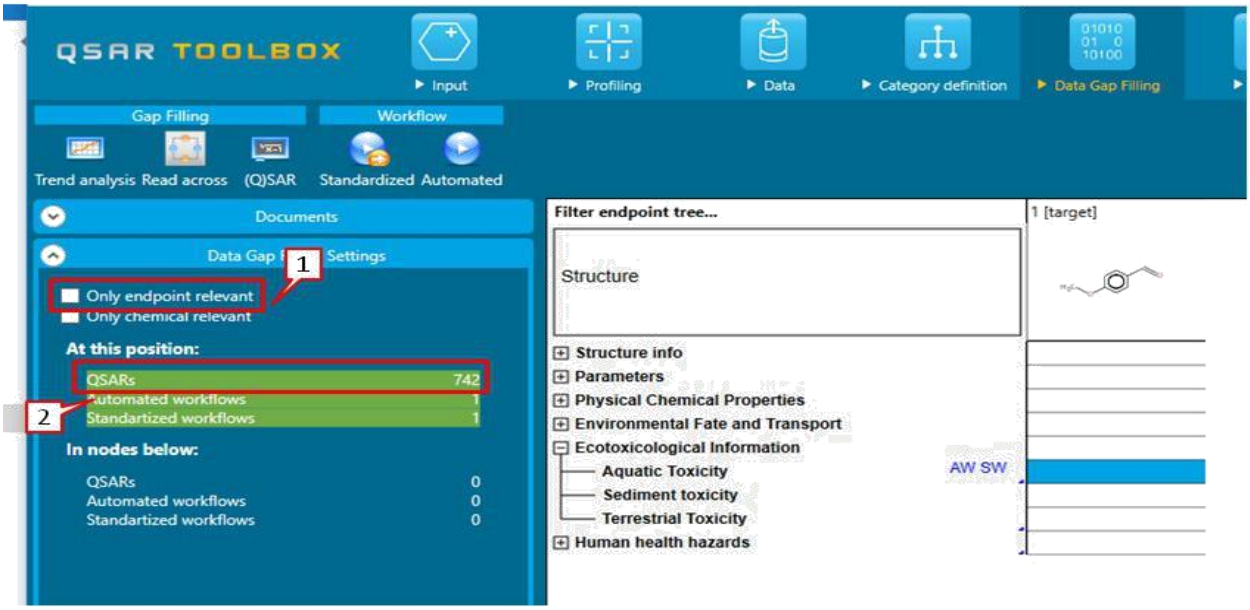


図 3

QSAR を実行するには、次の手順を実行してください。（図 4）：

- 関連するエンドポイント（1）を選択します。

- **Only chemical relevant** (2)をチェックします。これにより、標的化学物質の官能性に基づいて、適用される QSAR（5）の数が削減されます。このオプションをオフにすると、標的化学物質に存在する官能基に関係なく、エンドポイントに関連するすべての QSAR が適用されます。

- (Q)SAR (3)をクリックします。

- ポップアップ・ウィンドウが表示され、実行されたすべての QSAR のリスト（5）とその予測（6）が表示されます。

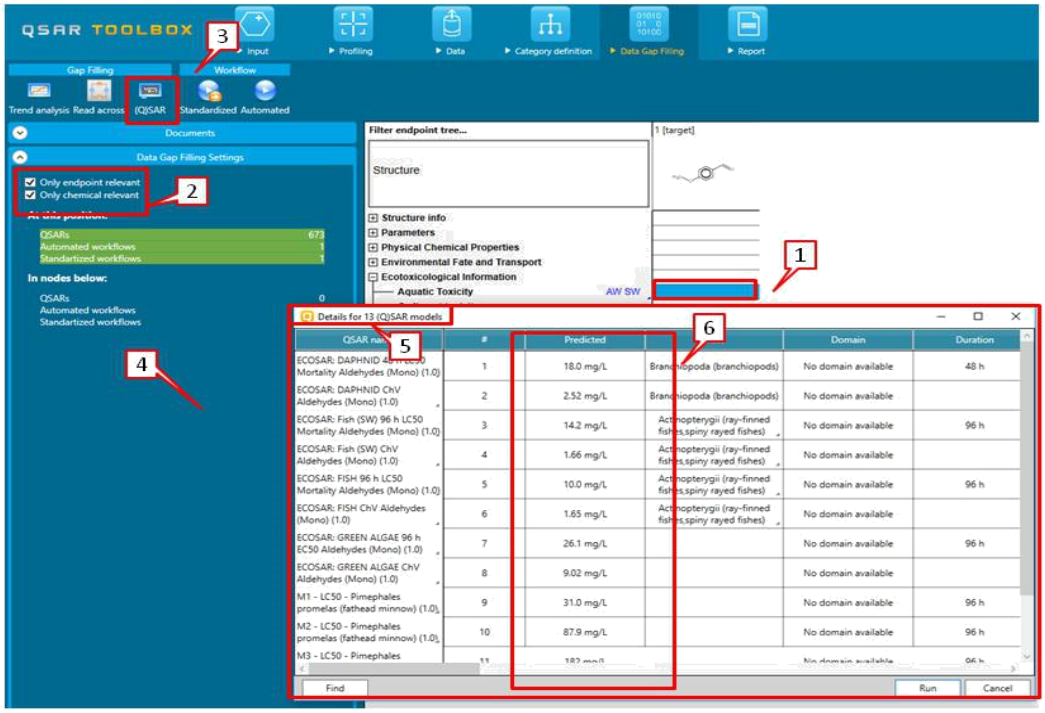


図 4

（Q）SAR の名前（1）を右クリックし、ポップアップメニューを表示し、 Show training set (2) をクリックするとトレーニングセット（3）の化学物質を表示します。(図 5)

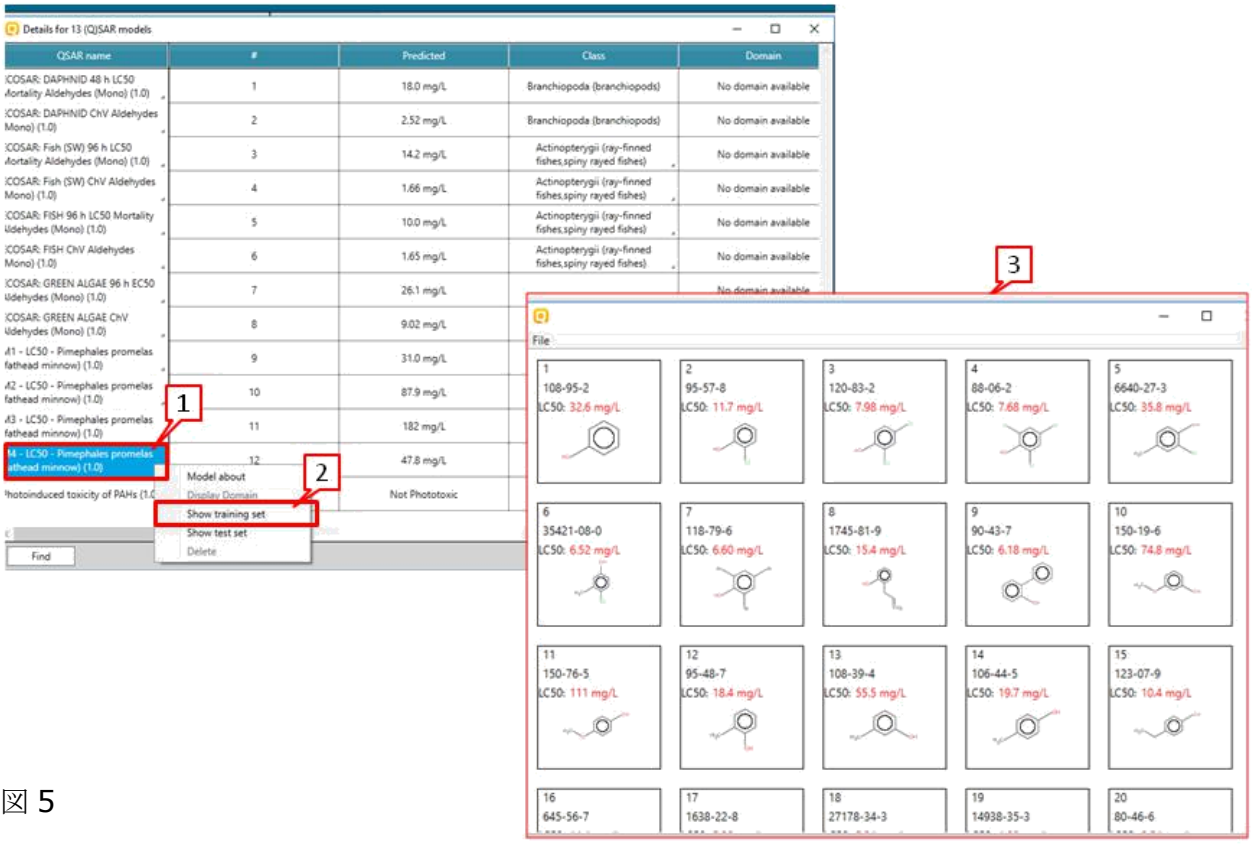


図 5

予測手順を続行するには、Run（1）をクリックして Select QSAR method（2）から選択します。（図 6）

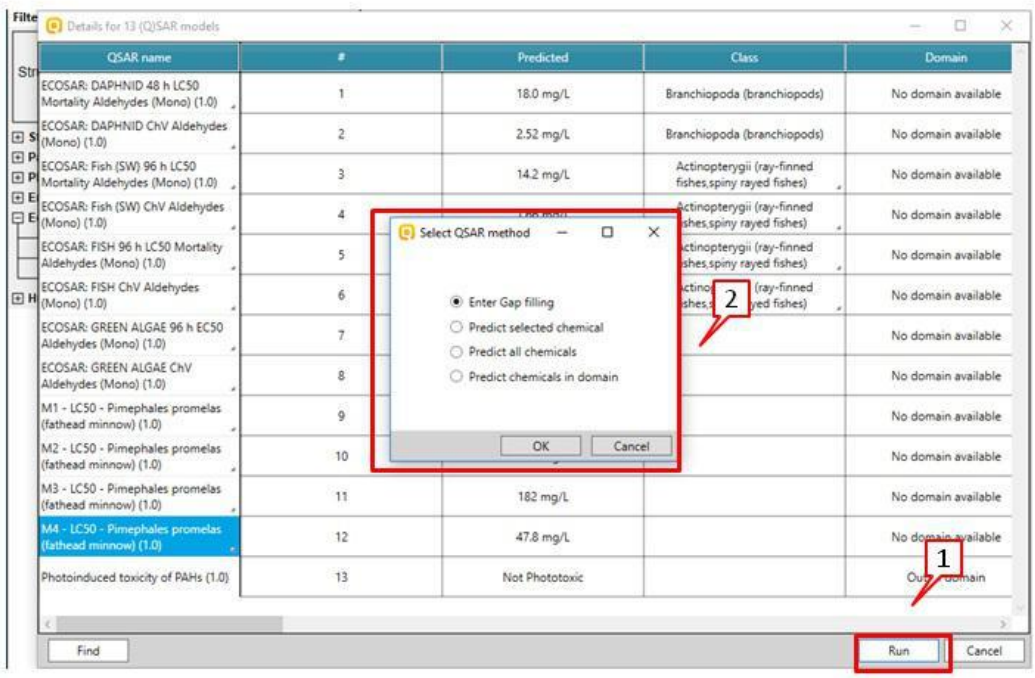


図 6

- Enter Gap filling: QSAR のトレーニングセットの化学物質（2）と予測（赤い点）と回帰式（3）に基づいて回帰グラフ（1）を示します（図 7）。この予測が受容されたら(4,5)（訳注：Accept prediction を押すことで確認メッセージが表示され、Yes をクリックすると、Success というメッセージが表示され、受容が完了する）、データ・マトリックスに表示されます。

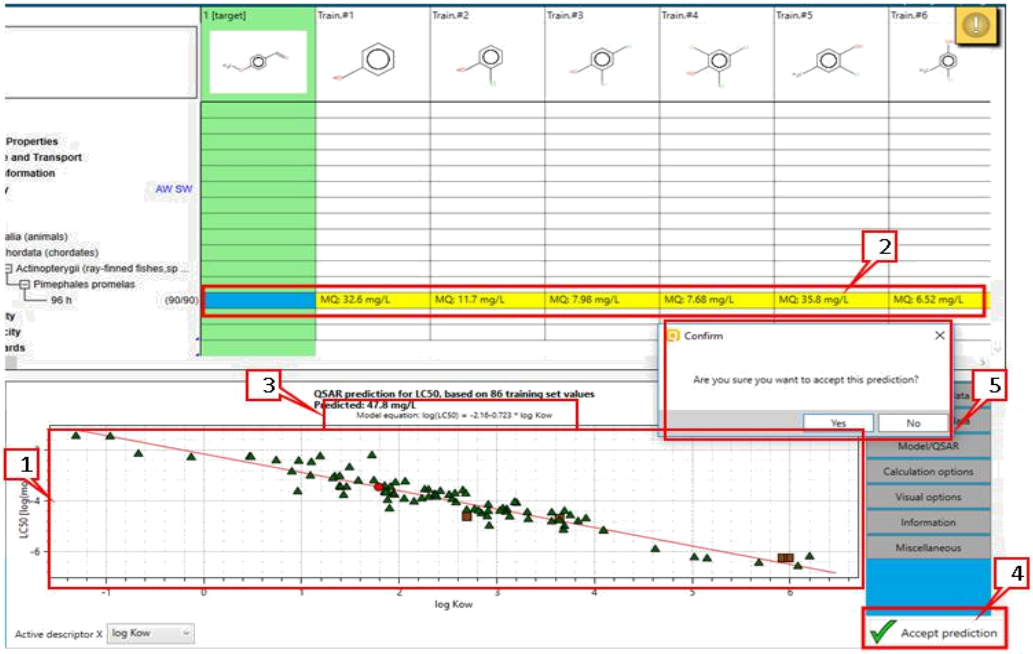


図 7

- Predict Selected chemical: （Q）SAR（1）を選択し、 Run （2）をクリックし、 *Predict selected chemical* (3) をクリックして OK（4）をクリックします。この予測は、データ・マトリックス（5）上に表示されます。

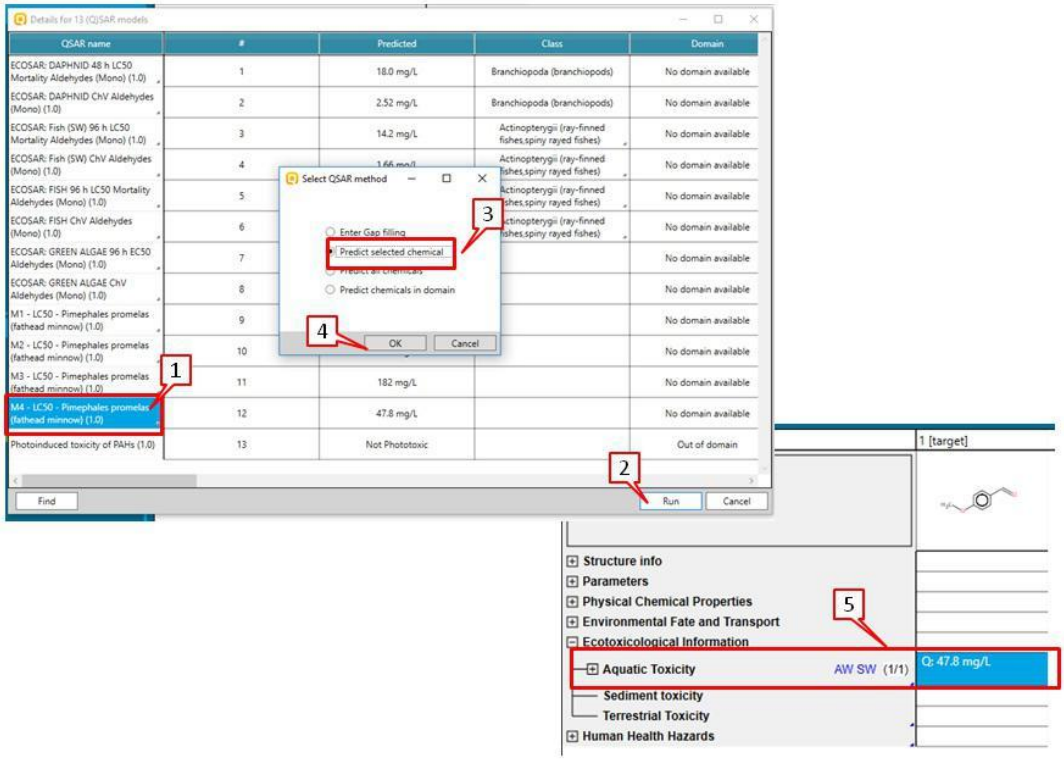


図 8

- Predict all chemicals (図 9)は一群の化学物質がデータ・マトリックスに読み込まれているときに使用できます。Predict all chemicals （1）をクリックし、OK（2）をクリックすると、予測がデータマトリックス（3）に表示されます。

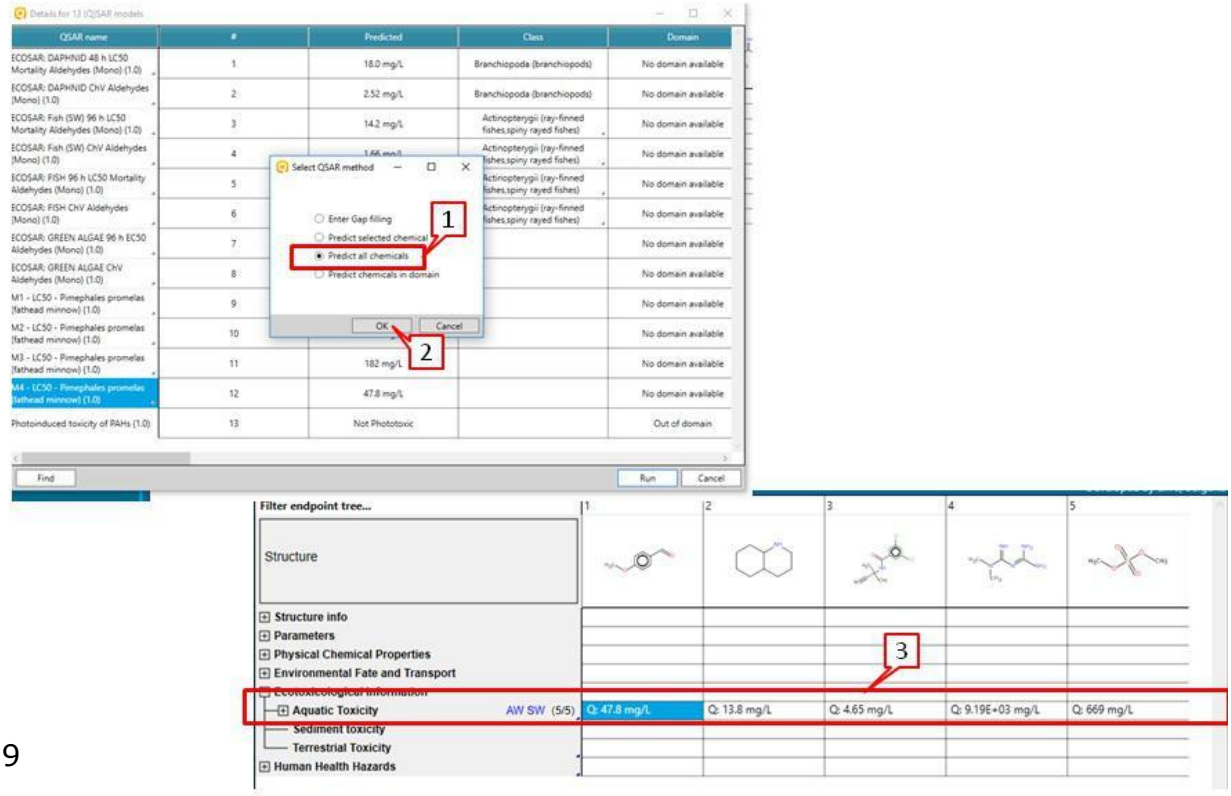


図 9

- Predict chemicals in domain はドメイン内に標的化学物質が入っている場合に適用可能です。

（訳注：ここでの「ドメイン内」とは、予測に使われるトレーニングセットの物質群の「予測適応範囲内」に、標的物質が含まれているという意味である。また現在、図 10 の予測モデルのうち Photoinduced toxicity（光誘発毒性）等、一部のモデルのみに対応している。）

373 NRU データ（図 10）に基づく光誘発毒性の例：予測モデル（1）を右クリックし、Display domain (2) をクリックすると、化学物質がドメイン内かドメイン外かということを含め、ドメインパラメータが表示されます。

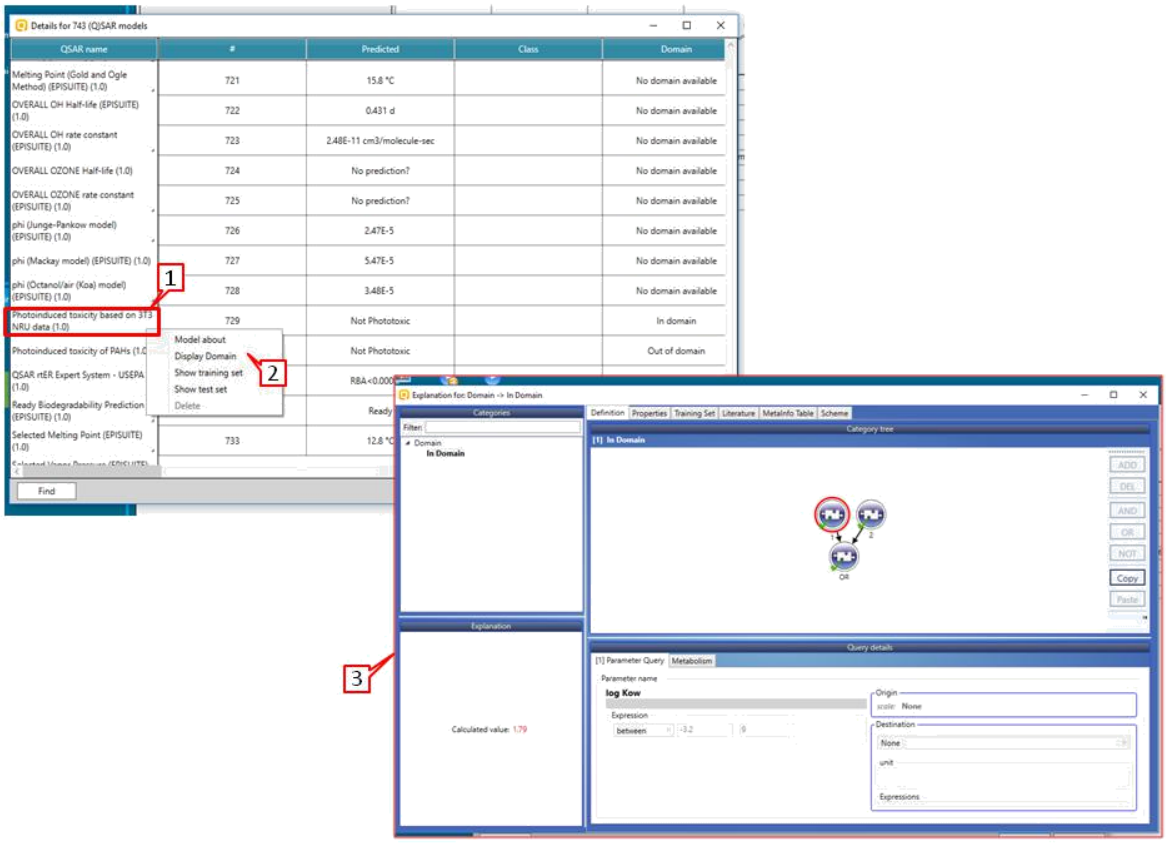


図 10

Predict chemicals in domain (1)をクリックし、 OK (2) をクリックします(図 11)

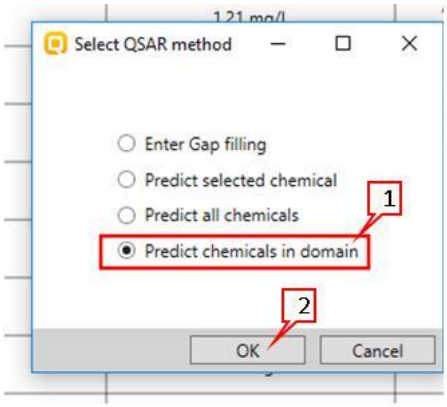


図 11

するとドメイン内の化学物質だけが予測されます。ドメイン外のものは空白のままになります。

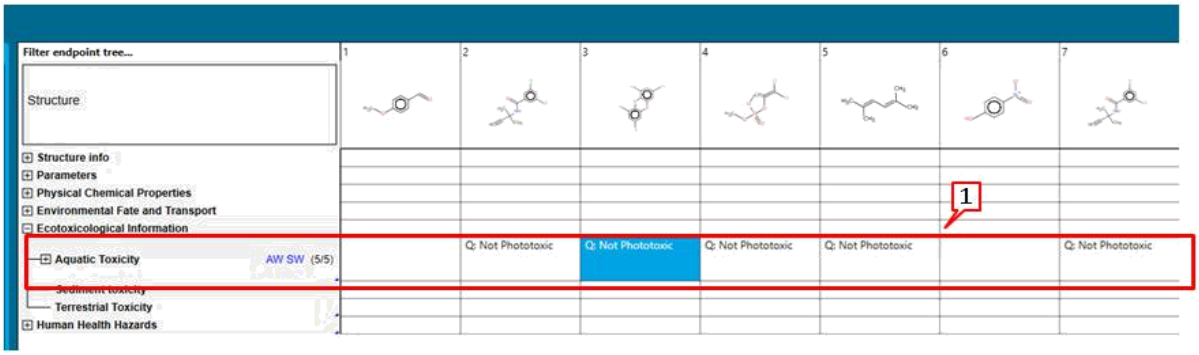


図 12

D.5.6. Automated and Standardized workflow（自動化・標準化ワークフロー）

自動化・標準化ワークフロー

これらのアルゴリズムは、皮膚感作性（LLNA と GPMT データ）および魚に対する急性水生毒性（Pimephales promelas、Mortality、LC50、96 時間）のデータギャップ補完のために開発されました。いったん開始されると、AW（自動化されたワークフロー）は装備されたロジックに従い、ユーザとやりとりすることなく予測が完了します。標準化データギャップ補完（SW）を実施するときも同じアルゴリズムが実装されています。AW と比較したときの主な違いは、アプリケーションのドメインが SW（他の生物種、試験実施期間などを含む）内では拡張されていることと、SW はユーザからの干渉を可能にすることであり、したがって AW では実行できないものも選択できるということです。

D.5.6.1.自動化・標準化ワークフローのアルゴリズム

1. ECOTOX エンドポイント用の SW/AW のアルゴリズム

1. AW の標的エンドポイント：

- Sub-hazard（有害性） - Aquatic toxicity（水生毒性）
- Effect（影響） - Mortality（死亡率）
- エンドポイント - Lc50
- 試験期間 - 96 h
- 試験生物種 - *P. promelas*

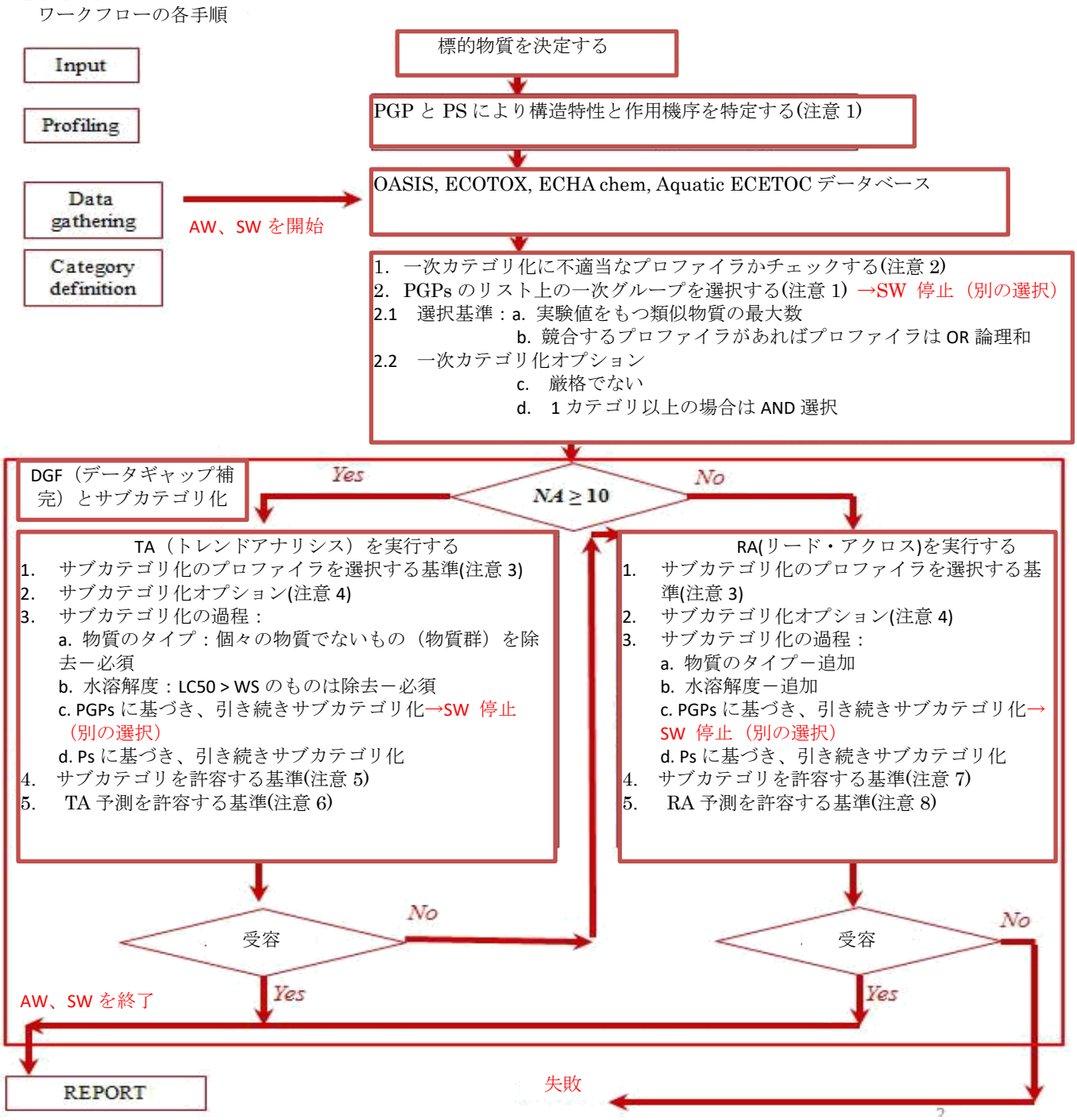
***2.* SW の標的エンドポイント：**

SW は、以下のエンドポイントの組み合わせに対して提案されます。効果、生物種、および期間（表 1 参照）。

表 1.エンドポイントの SW のワークフローの組み合わせ

Endpoint（エンドポイント）	**Effect（影響）**	**Duration（期間）**	**Phylum（動物分類上の門）**	**Class（分類）**
LC50	Immobilization（運動抑制）	24h- 96h	Arthropoda (arthropods、節足動物)	Branchiopoda（ branchiopods、鰓脚綱 ）
LC50	Intoxication（中毒）	24h- 96h	Arthropoda (arthropods、節足動物)	Branchiopoda（ branchiopods、鰓脚綱 ）
LC50	Mortality（死亡）	24h- 96h	Chordata (chordates、脊索動物)	Actinopterygii (条鰭類、棘条魚)
LC50	Mortality（死亡）	24h- 96h	Arthropoda (arthropods、節足動物)	Branchiopoda（ branchiopods、鰓脚綱 ）
EC50	Growth（成長）	24h - 96h	Chlorophyta (緑藻類)	Chlorophyceae（緑藻類）
EC50	Immobilization（運動抑制）	24h- 96h	Arthropoda (arthropods、節足動物)	Branchiopoda（ branchiopods、鰓脚綱 ）
EC50	Intoxication（中毒）	24h- 96h	Arthropoda (arthropods、節足動物)	Branchiopoda（ branchiopods、鰓脚綱 ）
EC50	Mortality（死亡）	24h- 96h	Chordata (chordates、脊索動物)	Actinopterygii (条鰭類、棘条魚)
EC50	Mortality（死亡）	24h- 96h	Arthropoda (arthropods、節足動物)	Branchiopoda（ branchiopods、鰓脚綱 ）
EC50	Population（個体群）	24h - 96h	Chlorophyta (緑藻類)	Chlorophyceae（緑藻類）

3. ワークフロー アルゴリズム



注意 1：Primary （一次）と secondary （二次）のグルーピングプロファイラ

Primary grouping profilers (PGPs、一次グルーピングのプロファイラ):

- US EPA,
- Verhaar
- MOA
- ECOSAR
- OFG （Organic Functional Groups)
- OFG （ USEPA ）
- OFG, Norbert Haider

Profilers for subcategorization (PS、サブカテゴリ化のプロファイラ):

- Substance type
- Protein binding (OASIS v1.4/OECD),
- Chemical elements（化学元素）

注意 **2:** 一次分類には適さないプロファイラ

プロファイラ	プロファイラの結果
US-EPA	N/A Not categorized（分類されていません） ERROR!（エラーが発生しました !） PROFILING TIMEOUT（プロファイリングのタイムアウト）
Verhaar	Class 2 (less inert compounds)（不活性な化合物が少ない) Class 3 (unspecific reactivity)（非特異的反応性） Class 4 (Compounds and groups of compounds acting by a specific mechanism)（特定の反応機構によって作用する化合物や化合物群） Class 5 (Not possible to classify according to these rules) should be excluded（これらの規則に従って分類することは不可能。除外すべき） ERROR!（エラーが発生しました !） PROFILING TIMEOUT（プロファイリングのタイムアウト） (N/A)
MOA（作用機序）	Reactive unspecified（反応しない） (N/A) ERROR!（エラーが発生しました !） PROFILING TIMEOUT（プロファイリングのタイムアウト）
ECOSAR	(N/A) ERROR!（エラーが発生しました !） No alert found（アラートは見つかりませんでした） Not classified（分類されていません） Not Related to an Existing ECOSAR Class Definition（既存の ECOSAR クラスの定義に関連していません） PROFILING TIMEOUT（プロファイリングのタイムアウト） SHOULD NOT BE PROFILED（プロファイルできません） Salt（塩）
Protein binding OASIS（タンパク質結合、OASIS）	N/A No alert found（アラートは見つかりませんでした） ERROR!（エラーが発生しました !） PROFILING TIMEOUT（プロファイリングのタイムアウト）
Protein binding OECD（タンパク質結合、OECD）	N/A No alert found（アラートは見つかりませんでした） ERROR!（エラーが発生しました !） PROFILING TIMEOUT（プロファイリングのタイムアウト）
Organic functional groups（有機官能基） Organic functional groups (USEPA) ; OFG (Norbert Haider);	ERROR!（エラーが発生しました !） N/A No functional group found（官能基が見つかりません） PROFILING TIMEOUT（プロファイリングのタイムアウト）

注意 3: サブカテゴリ化用のプロファイラの選択基準

- 標的物質と（とても）異なる類似物質が最小数であること
- 競合するプロファイラがある場合には、“acceptance（受容）”統計を提供するプロファイラが適用されます。

注意 4: サブカテゴリ化のオプション：少なくとも 1 つのカテゴリが選択されていることです。

注意 5: TA（トレンドアナリシス） のサブカテゴリ化の受容条件は次のとおりです。

1. PGP によるサブカテゴリ化

IF 内挿 **AND**（R^2↑ **OR** 残差の 95%↓）**AND** NA ≧10
THEN サブカテゴリ化を受容し、次の PGP に進む
ELSE サブカテゴリ化を拒絶し、次の PGP に進む

2. PSs によるサブカテゴリ化

IF 内挿 **AND**（R^2↑OR 残差の 95%↓）**AND** NA ≧10
THEN サブカテゴリ化を受容し、次の PS に進む
ELSE 次の PS に進む

Difinition（上記の言葉の規定）：
内挿： 標的物質の log Kow は類似物質の log Kow の範囲内でなければならない
相関係数 R^2↑： R^2 が増加する **OR** R^2 ≧0.85
残差の 95%↓： 残差の 95%減少 **OR** 残差の 95%≦1.0
NA： 類似物質の数

注意 6: TA（トレンドアナリシス） の受容条件は次のとおりです。

すべての PGP および PS によるサブカテゴリ化の後
IF R^2≧0.7 **And** NA≧10
THEN 予測を受容してレポートを生成する
ELSE リード・アクロスに切り替える

注意 7: リード・アクロスを実行する際のサブカテゴリ化ステップの受容基準

1. PGP によるサブカテゴリ化

IF 内挿 **AND**（LC50↓ **OR** log Kow↓）**And** NA ≧5
THEN サブカテゴリ化を受容し、次のプロファイラに進む
ELSE サブカテゴリ化を拒絶し、次のプロファイラに進む

2. PSs によるサブカテゴリ化

IF 内挿 **AND**（LC50↓ **OR** log Kow↓）**And** NA ≧5
THEN サブカテゴリ化を受容し、次のプロファイラに進む
ELSE サブカテゴリ化を拒絶し、次のプロファイラに進む

Difinition（上記の言葉の規定）：
内挿： 標的物質の log Kow は類似物質の log Kow の範囲内でなければならない
LC50↓： 最も類似した 5 つの類似物質について、LC50 の変動範囲が減少する **OR** 変動範囲が≦2 log 単位
log Kow↓： 最も類似した 5 つの類似物質について、log Kow の変動範囲が減少する **OR** 変動範囲が≦2 log 単位
NA： 類似物質の数

注意 **8: RA**（リード・アクロス） の受容条件は次のとおりです。

すべての PGP および PS によるサブカテゴリ化の後

IF 内挿 **AND** LC50≦2 log 単位 **OR** log Kow≦2 log 単位 **AND** NA≧5

THEN 予測を受容し、レポートに進む

Difinition（上記の言葉の規定）：

内挿： 標的物質の log Kow は類似物質の log Kow の範囲内でなければならない

LC50=2： 最も類似した 5 つの類似物質について、LC50 の変動範囲が減少する **OR** 変動範囲=2 log 単位

log Kow=2： 最も類似した 5 つの類似物質について、log Kow の変動範囲が減少する **OR** 変動範囲=2 log 単位

NA： サブカテゴリ化を適用した後の類似物質の数

II. 皮膚感作性の AW/SW のワークフローのアルゴリズム

1. SS（皮膚感作性）の AW で使用されるプロファイルとシミュレータのリスト

ケース I：標的構造にアラートがある場合

一次グルーピングのプロファイル一覧：

- Protein binding alerts for skin sensitization by OASIS（OASIS、皮膚感作のタンパク質結合アラート）

サブカテゴリのみに使用する補助となるプロファイル一覧：

- Substance type（物質のタイプ）
- Protein binding alerts for skin sensitization by OASIS（OASIS、皮膚感作のタンパク質結合アラート）

ケース II ：AU または SM シミュレーションの後にアラートが見つかった場合

一次グルーピングのプロファイル一覧：

- Protein binding alerts for skin sensitization by OASIS v 1.4（OASIS v 1.4、皮膚感作のタンパク質結合アラート）
  ＋ Autoxidation (or Skin Metabolism) simulators（自動酸化または皮膚代謝シミュレータ）

サブカテゴリのみに使用する補助となるプロファイル一覧：

- Substance type（物質のタイプ）
- Protein binding alerts for skin sensitization by OASIS（OASIS、皮膚感作のタンパク質結合アラート）
- Structure similarity（化学構造の類似性）

ケース III ：標的（親）物質構造にも、生成された代謝物にもアラートが検出されない場合

一次グルーピングのためのプロファイル一覧：

- US-EPA New Chemicals categories（US-EPA の新化学物質カテゴリ）
- Aquatic toxicity classification by ECOSAR （ECOSAR、水生生物毒性の分類）
- Protein binding alerts for skin sensitization by OASIS（OASIS、皮膚感作のタンパク質結合アラート）
- Organic functional groups（有機官能基）
- Organic functional groups (nested)（有機官能基・入れ子状）
- Organic functional groups (US-EPA)（有機官能基・US-EPA）
- Organic functional groups, Norbert Haider (checkmol)（有機官能基、Norbert Haider、Checkmol）

サブカテゴリのみに使用する補助となるプロファイル一覧：

- Substance type（物質のタイプ）
- Protein binding potency（タンパク質結合能）
- Keratinocyte gene expression（ケラチン生成細胞遺伝子発現）
- Structure similarity（化学構造の類似性）

Protein binding alerts for skin sensitizaiton by OASIS と組み合わせて使用するシミュレータ

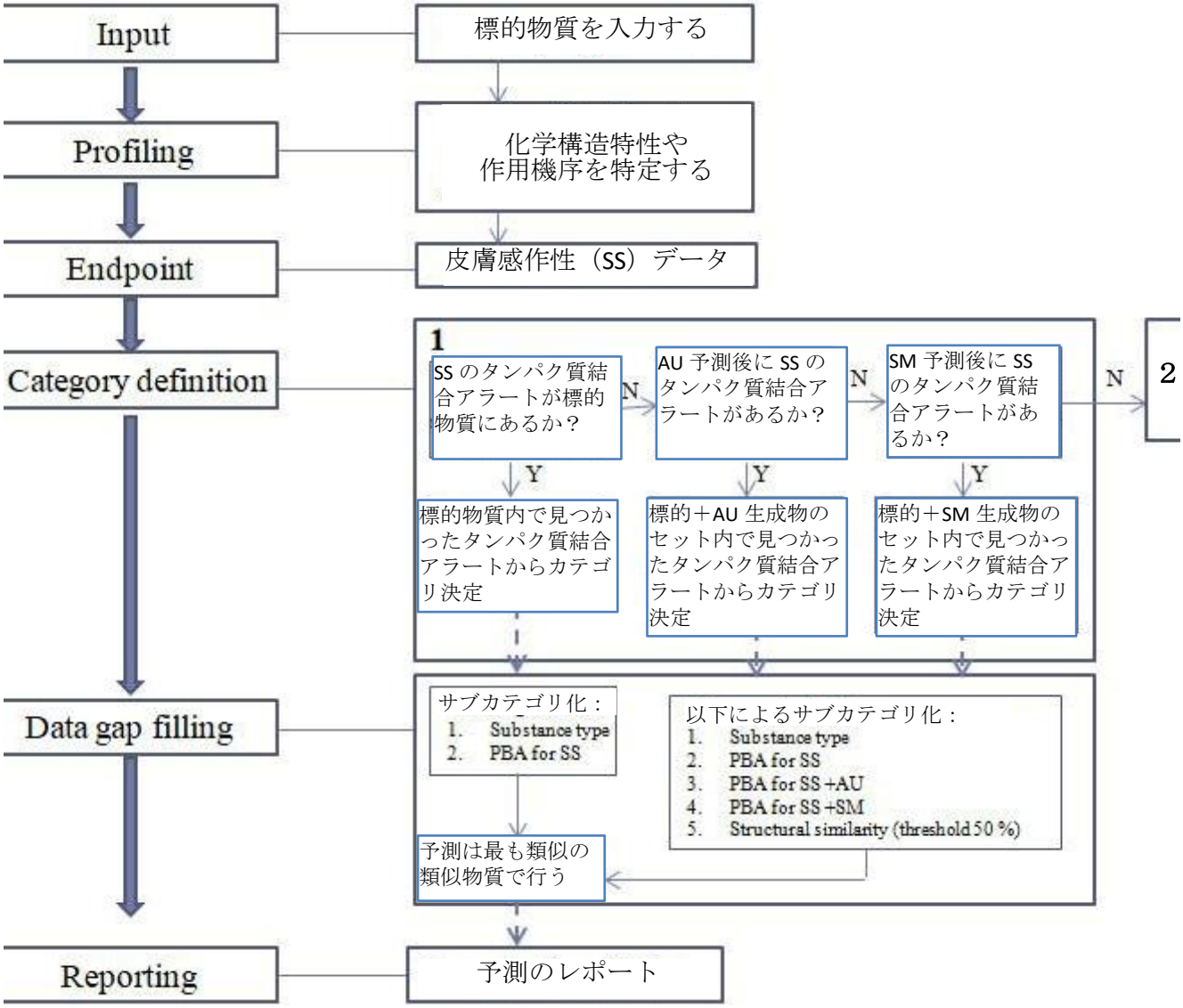
- Autoxidation simulator（AU、自動酸化シミュレータ）
- Skin metabolism simulator（SM、皮膚代謝シミュレータ）

Databases（データベース）

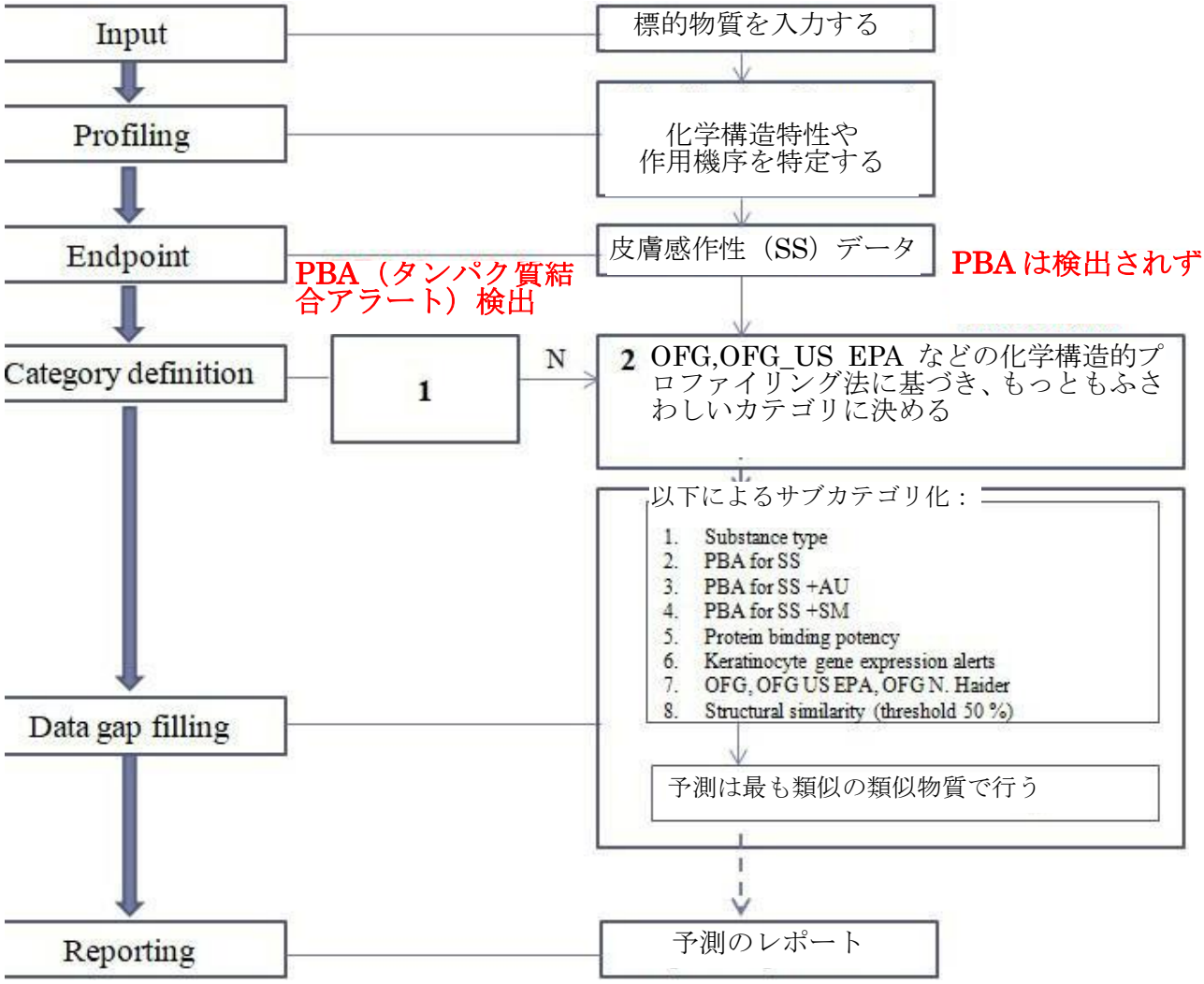
- Skin sensitization（皮膚感作性）
- Skin sensitization ECETOC（皮膚感作性；ECETOC）

2. アルゴリズム

SS（皮膚感作性）評価のための AW /SW - Part 1



SS（皮膚感作性）評価のための AW /SW　- Part 2



PBA = Protein binding alert（タンパク質結合アラート）

D.5.6.2. Automated workflow（自動ワークフロー）

I. 単一の化学物質の生態毒性エンドポイントに関する自動ワークフロー

生態毒性を予測するための自動ワークフロー（AW）は、 1つのエンドポイントだけに行うよう開発されています。

- エンドポイント：試験動物の半分が死ぬ毒性物質の濃度 - LC50
- Effect（効果） - Mortality（死亡率）
- Species（試験生物種） - 魚（P. promelas）
- 試験期間 - 96 h

自動化ワークフローを実行するには、以下の説明の手順（図 1）に従ってください。

（訳注：Input モジュールを選び適当な化学物質を入力し、Data モジュールで全 Database を選択し Gather アイコンを押してから以下の操作を開始できる）

- Data Gap Filling（データギャップ補完）モジュールを選択します。
- Automated（自動化、AW） (2)を選択し ます。
- ポップアップウィンドウ（4）の Ecotoxicological Endpoint（3）ラジオボタンをクリックしてください。
- 最後に OK（5）を押します。

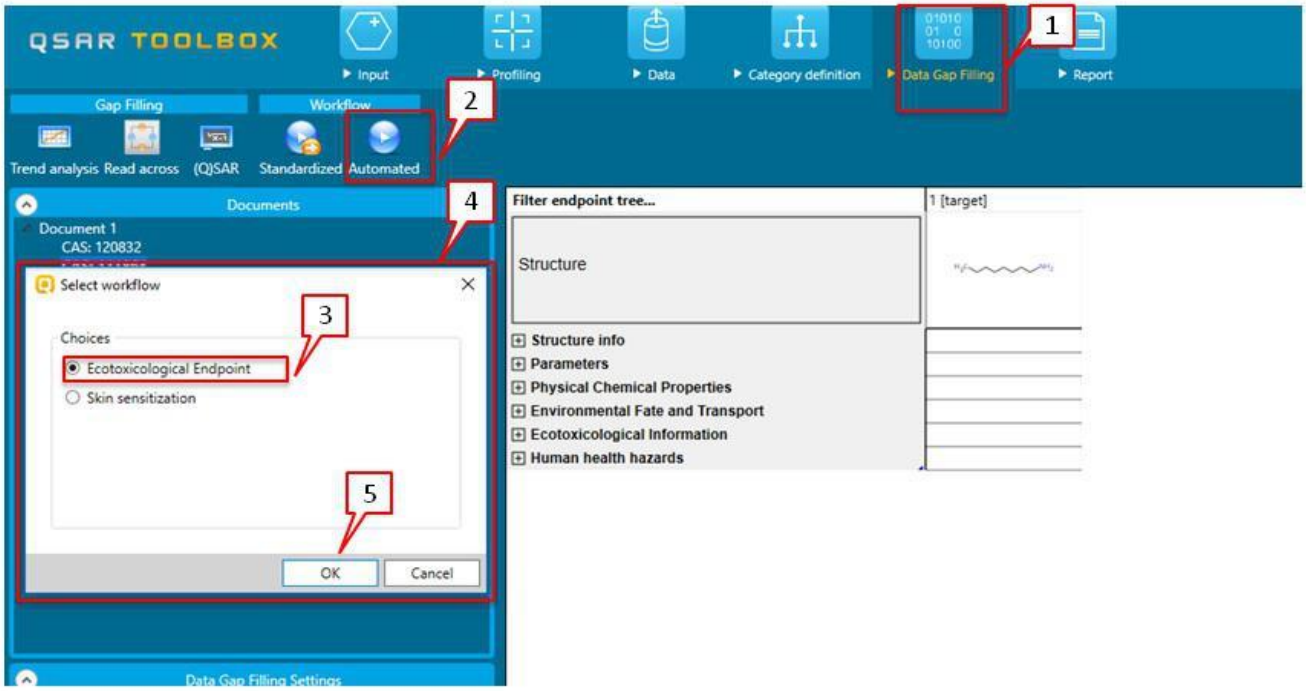


図 1 AW で Ecotoxicological エンドポイントをアクティベーションする

ダイアログウィンドウでエンドポイントパス（1,2）を選択することができます（図 2）

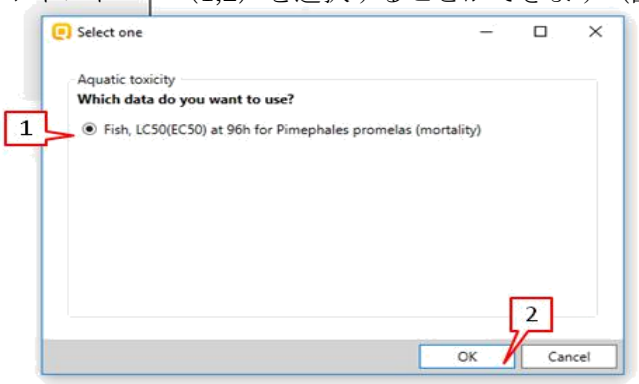


図 2

次に表示されるウィンドウは、ワークフローコントローラといわれるものです（図 3）。2 つの主要ナビゲーションボタン Pause と Stop（1）がついています。ワークフロー名（2）のキャプションの下には、現在実行されているアクションについてのジェネラルタスク（3）と、ジェネラルタスク内で実行されているサブタスクに関する情報を提供する、アクティブタスク（4）が含まれます。ワークフロー実行中に適用される一連の手順は、アクティビティログに要約されています。Show activity log（アクティビティログを表示）（5）をクリックして、一連の手順を確認します。

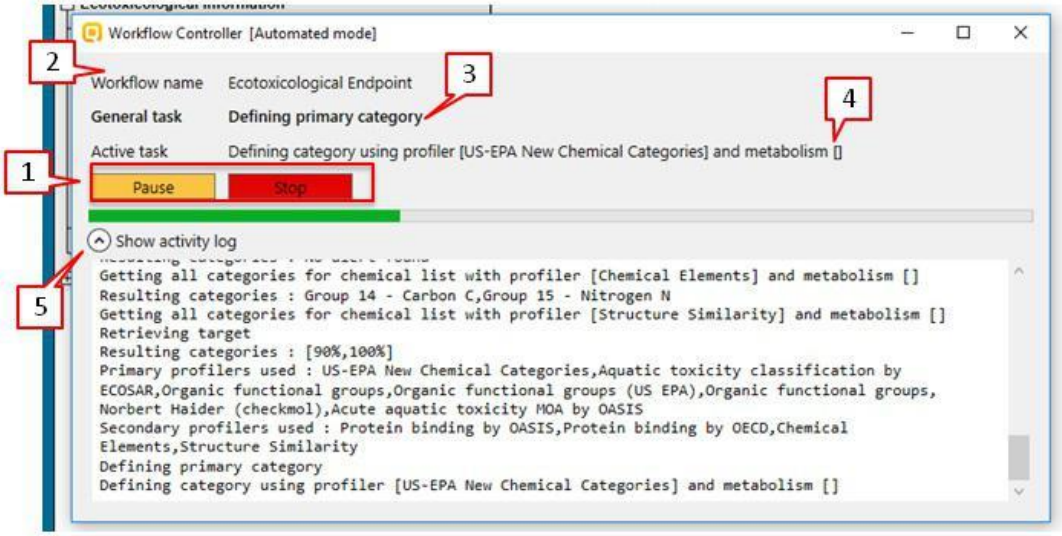


図 3

一連のステップは、AW の実行中に自動的に適用されます。ここではユーザの操作は必要ありません。

AW が終了すると、“Prediction accepted successfully（予測が正常に受容されました）”というメッセージが表示され（1）、プログレスバーは完全に色が変わります（2）（図 4）。終了したワークフローは、ワークフローコントローラ（3）のタイトルに書き込まれます。OK（4）をクリックし、X 印（5）をクリックしてワークフローコントローラウィンドウを閉じます。

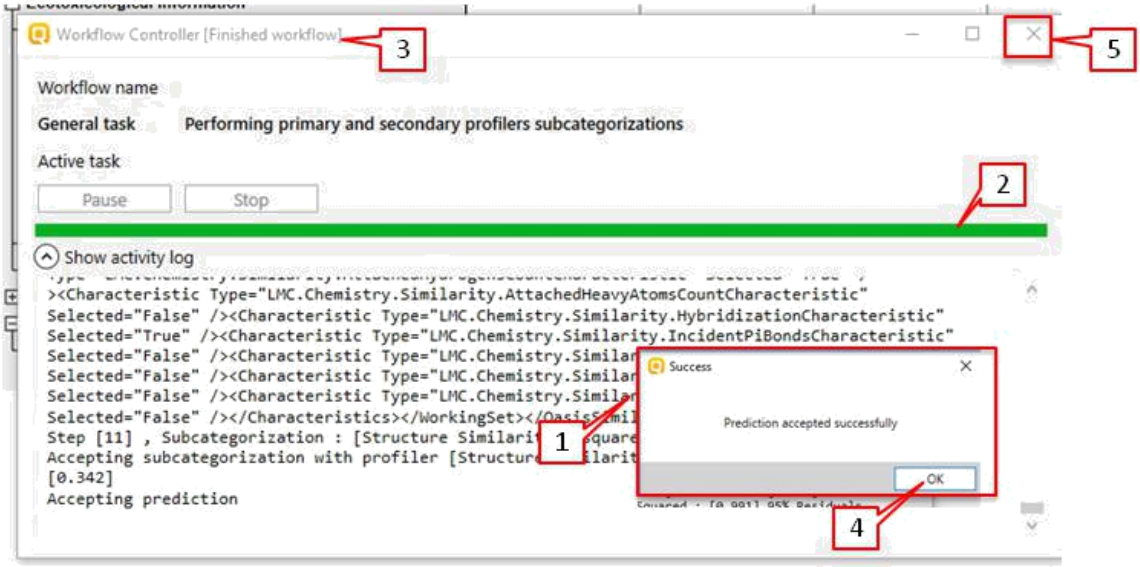


図 4

予測が予測受容基準を満たせない場合、または 1 次グループ分けのために十分なデータが収集されない場合、次のメッセージが表示されます（図 5）。

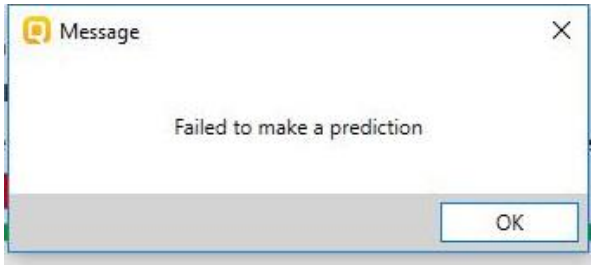


図 5

II.　一群の化学物質の生態毒性エンドポイントに関する自動ワークフロー

Input の章で説明した一群の化学物質の入力方法を参照してください。（訳注： Input モジュールで Substracture のアイコンを押し、ある化学構造を入力すると、それを構造内に含む化学物質が全部表示されるので、全選択したのち Data モジュールで全 Database を選び Gather アイコンを押してから以下の作業を開始する）

Data Gap filling （1）を選択し、Automated (2)をクリックし、ポップアップウィンドウ（3）から Ecotoxicological Endpoint（4）をクリックして、OK（5）をクリックします（図 6）

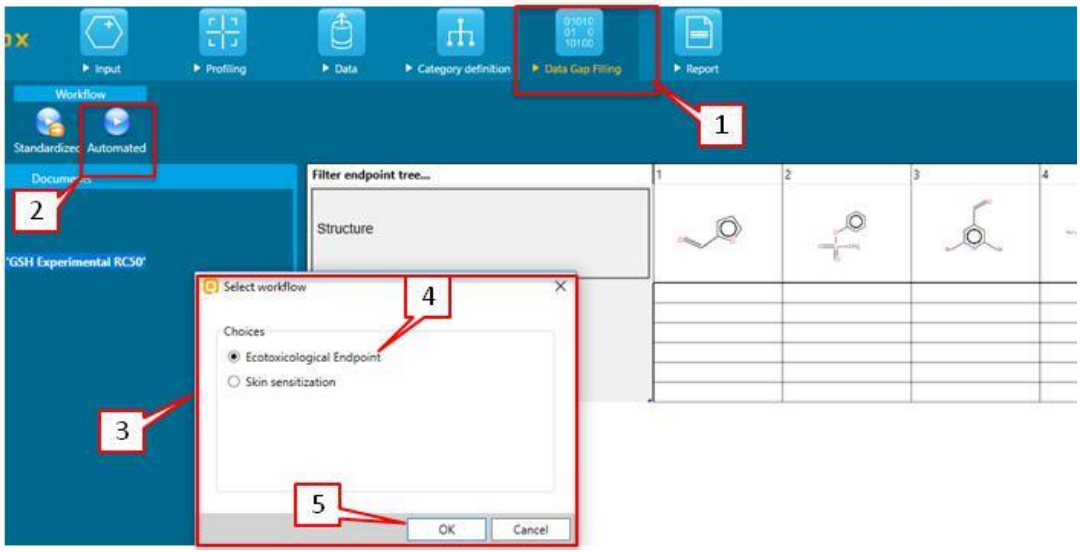


図 6

ポップアップウィンドウ（1）が出るので、エンドポイントパスを選択します（2）。図 7

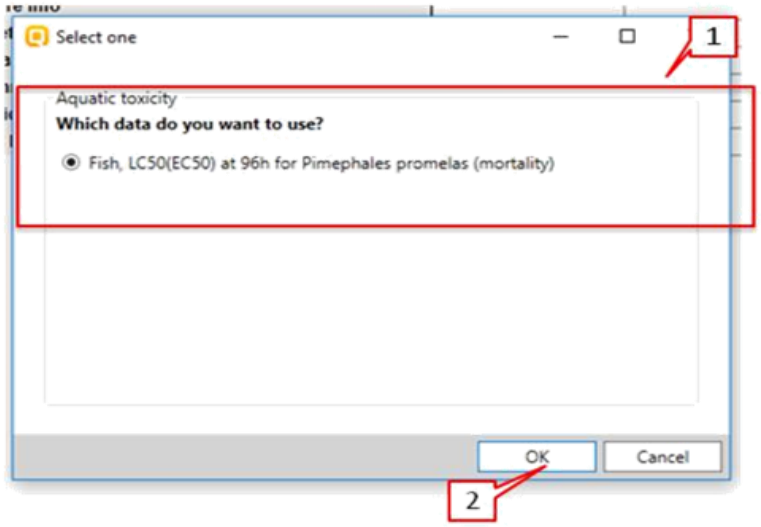


図 7

ワークフローコントローラウインドウが表示されますがアクティブではありません（図 8-1）。ポップウィンドウの Select range（2）が表示されます。ここで一群から予測すべき化学物質の範囲を選択してください。（訳注：選択された物質はデータ・マトリックスで左から番号が付けられている。どの番号の物質をギャップ補完に使うか選べる）OK (3)をクリックします。または、リストから化学物質の特定の範囲を予測する場合、Add range（範囲の追加）（4）をクリックします（5）。OK (6)をクリックします。

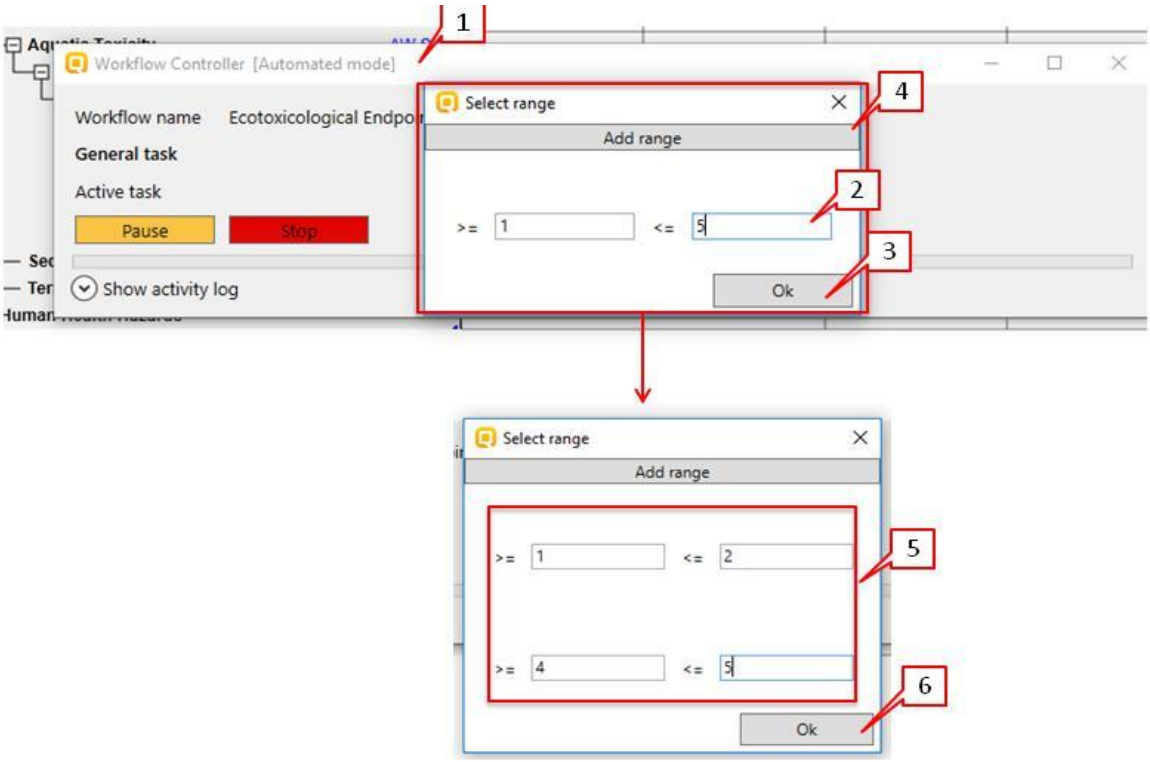


図 8

ワークフローが終了すると、ワークフローコントローラ（1）に表示されます（図 9）。また、進行状況バーが完全に塗りつぶされています（2）。この予測は、データ・マトリックス（3）上に表示されます。さらに、4 つの化学物質のうち 4 つが予測されることが示されています。
最後に、X 印ボタン（4）をクリックすると、ワークフローのウィンドウが閉じます。

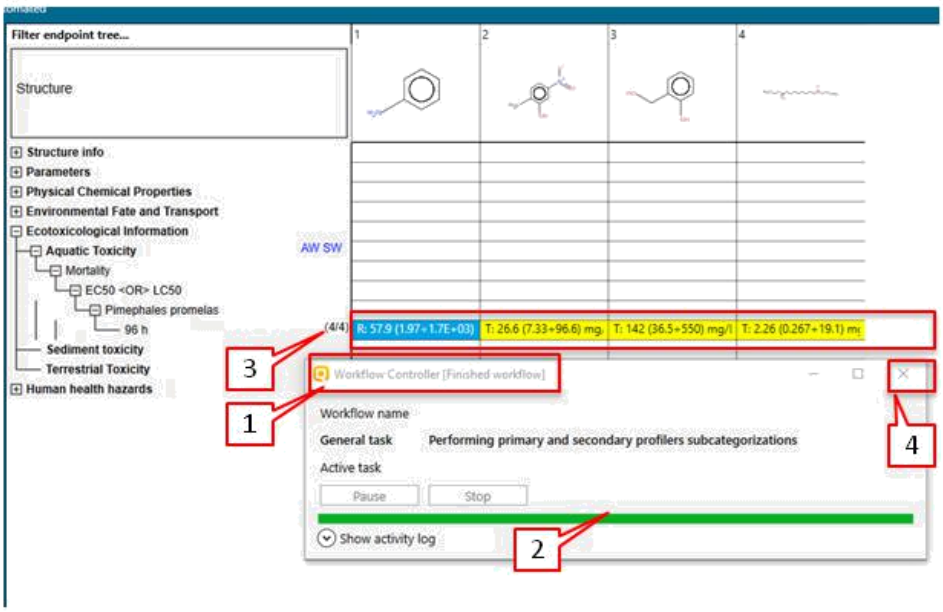


図 9

D.5.6.3. Standardized workflow（標準ワークフロー）

Standardized workflow（SW, 標準ワークフロー）
トレンドアナリシスを行う SW の使用例：
標的エンドポイント：LC50, 96h all fish (CAS 番号 111-86-4)

（訳注：事前に Input モジュールの CAS#アイコンを押し CAS 番号の 111-86-4 を入力後、Data モジュールで全 Database をクリックし、Gather アイコンを押してから以下の作業を開始するとよい）
次のステップを実行してください：Data Gap filling モジュール（1）に行き、 Standardized （2）をクリックし、ポップアップウィンドウ (4) から Ecotoxicological エンドポイント (3) を選択し、OK (5) をクリックします。(図 1)

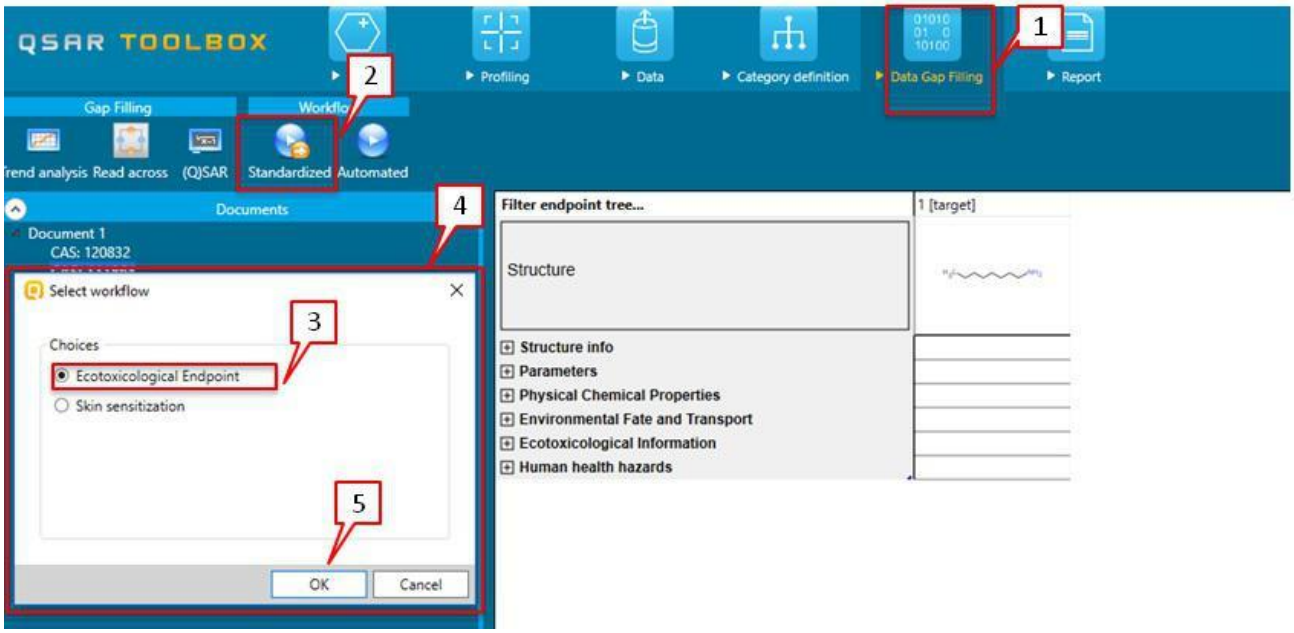


図 1

プレビューされたエンドポイントを選択するようダイアログボックスが表示されます (図 2)。エンドポイント（1）を選択し、OK（2）をクリックします。

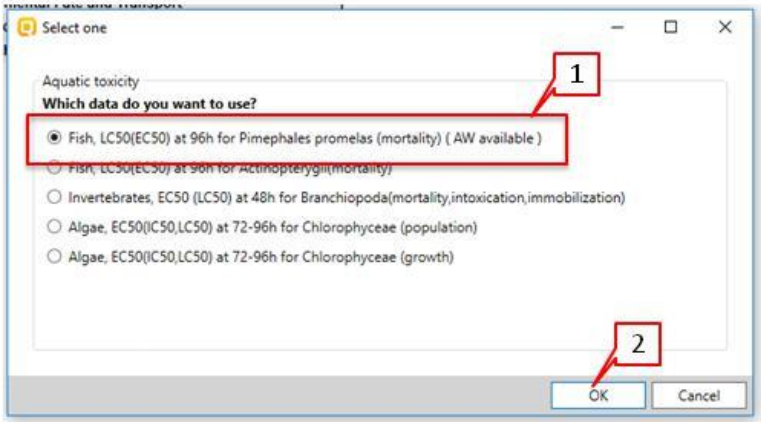


図 2

標的エンドポイントが選択されると、ワークフローコントローラウインドウが表示されます (1) (図 3)。数秒後にワークフローコントローラが表示され、選択するデータベースのリストを含む追加のウィンドウ

が表示されます。AW で使用されているデータベースが表示されています。調べたいデータベース（2）を選択し、OK（3）をクリックします。

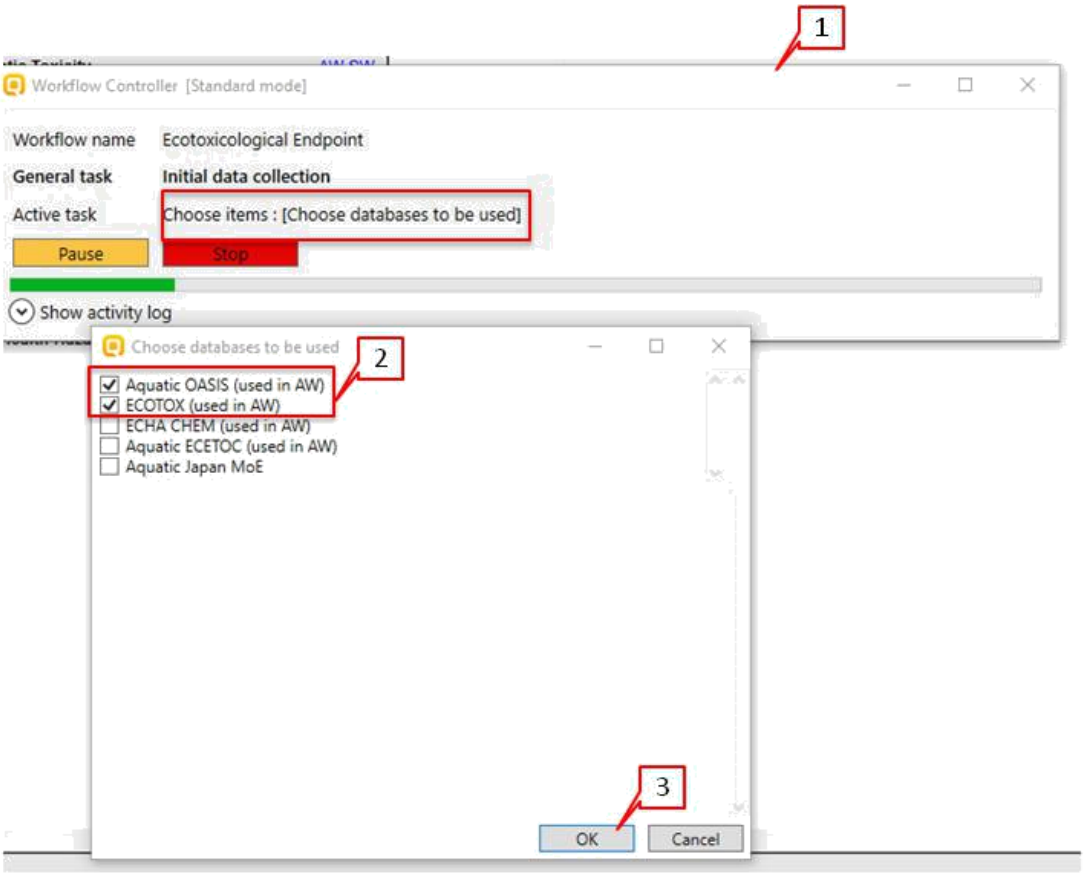


図 3

データベースの選択後、ワークフローはそれぞれの一次グループ分けプロファイラを標的化学物質に適用します。結果として、適用されたすべてのプロファイルを含むリストが、対応する化学物質/データ統計情報とともに提供されます（図 4）。プロファイル（1）を選択し、OK（2）をクリックします。

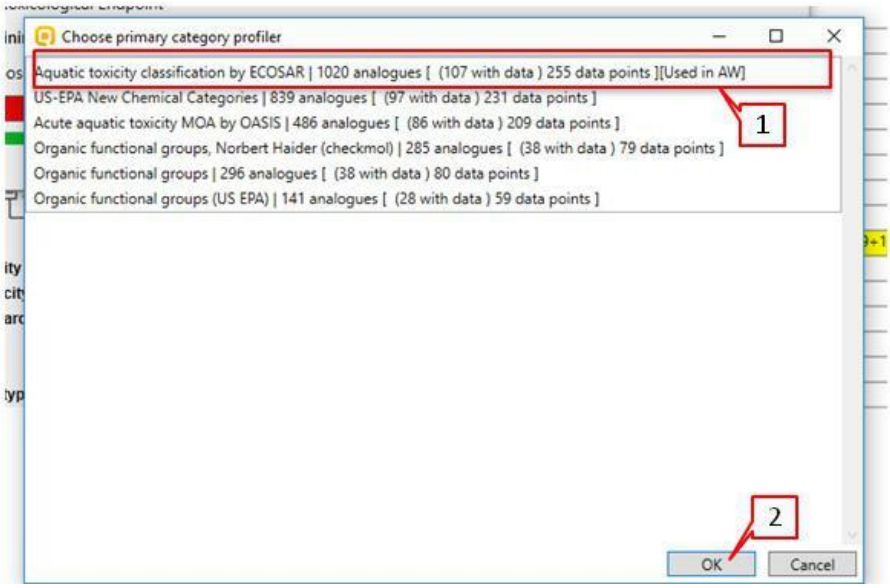


図 4

サブカテゴリ化が開始されると、一次グループと二次グループ（2）の 2 つのセクションに分かれている Subcategorisation ウィンドウ（1）が表示されます（図 5）。

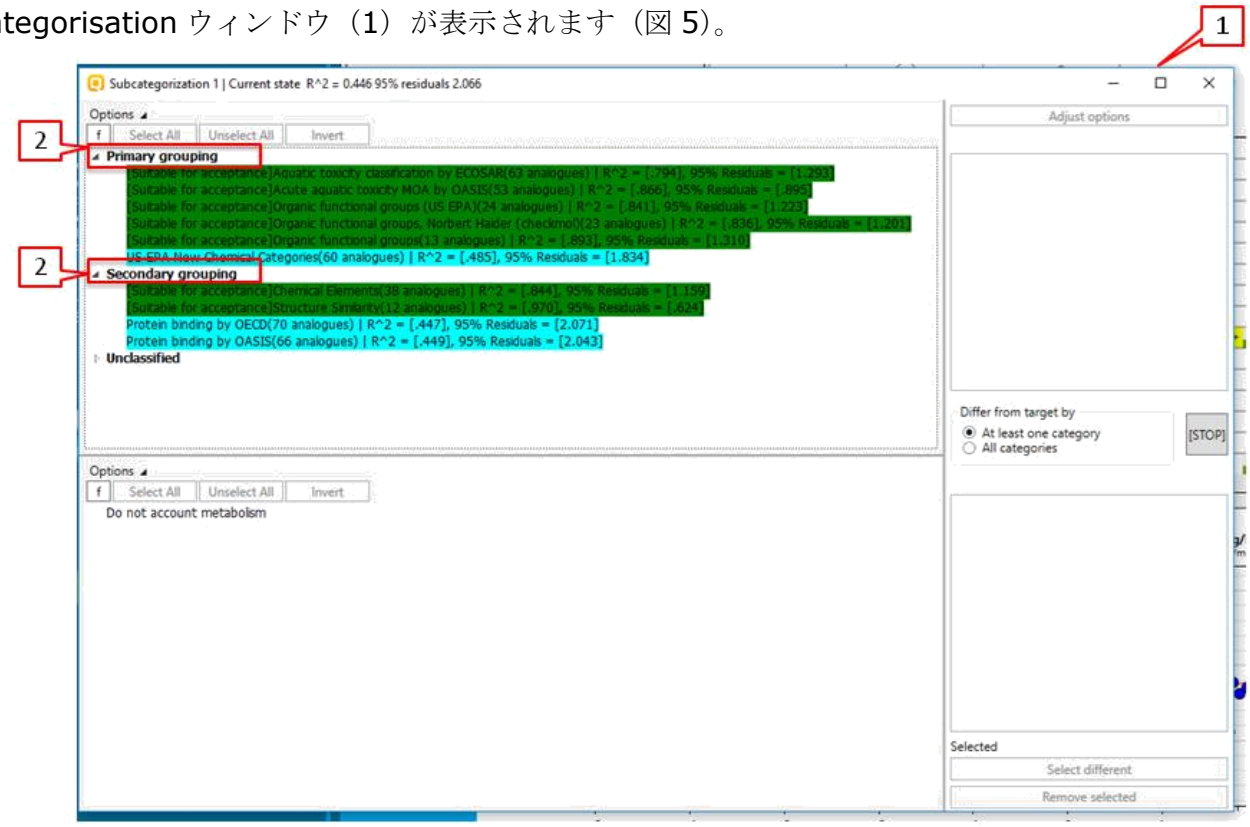


図 5

Options(1) を展開して、プロファイルを Group（グループ化）/Sort(並べ替え)/Color(色分け)または Filter（フィルタリング）します。*Legend* (2) をクリックしてカラーチャートを開きます。

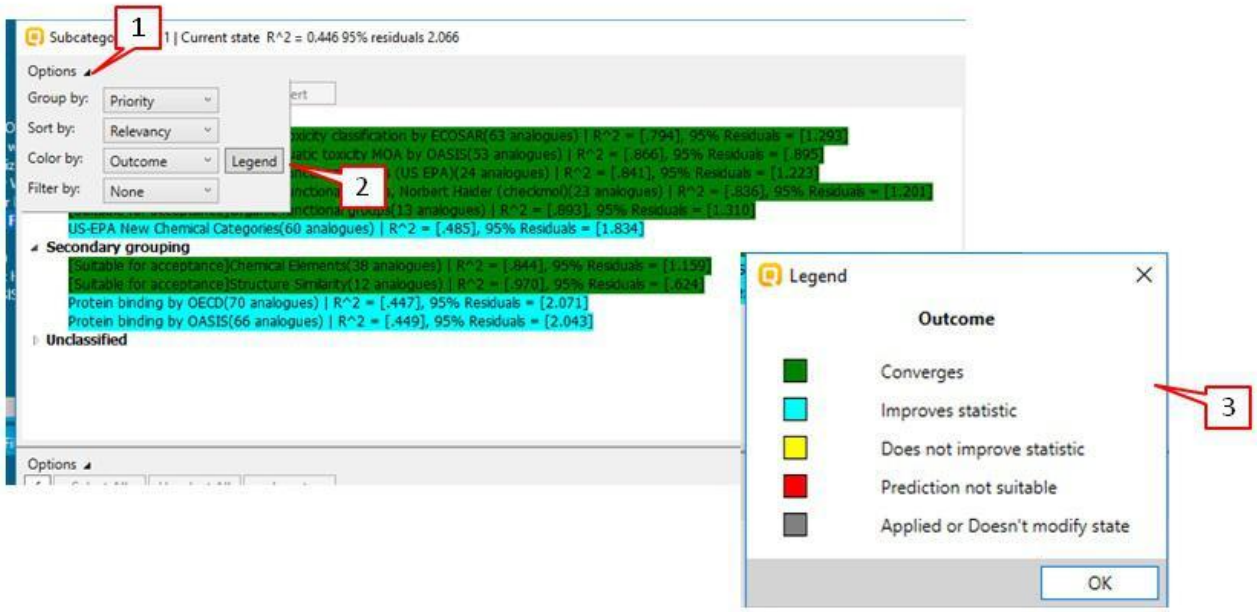


図 6

色の凡例：

グリーン　- プロファイラの適用は、予測受容基準に収まる

ブルー　- プロファイラを適用すると、統計のみが改善されます（たとえば、相関係数 R2 を上げる、または残差の 95%を減らす）

黄色 のプロファイラ適応では、統計情報は変更されず

赤 - サブカテゴリ化の受容基準に到達せず

グレー - プロファイラがすでに適用された

Subcategorisation ウィンドウのその他の機能は次のとおりです（図 7）。

- 現在の状態の統計的基準（1）
- 予測の受容条件（2）
- プロファイラの適合性の表示（3）
- 選択された基準に基づいてプロファイラの様々な順序を可能にするナビゲーションバー（4）

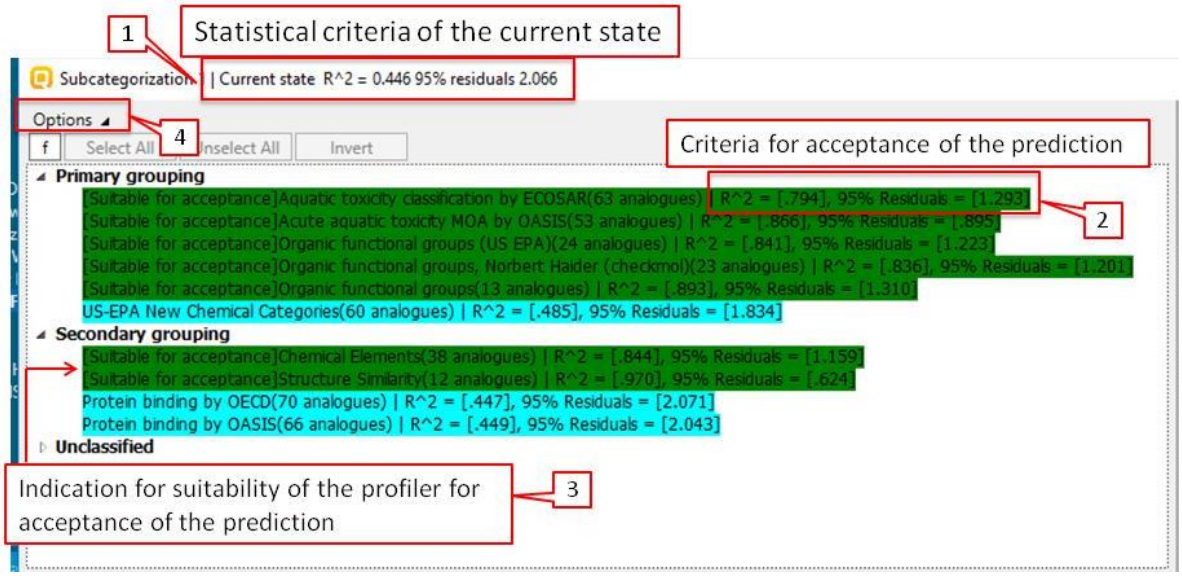


図 7

一次グループの最初のプロファイラをクリックすると、プロファイリング結果が標的化学物質と異なるすべての化学物質（2）が強調表示されます（図 8）。*Remove selected* (3)をクリックします。

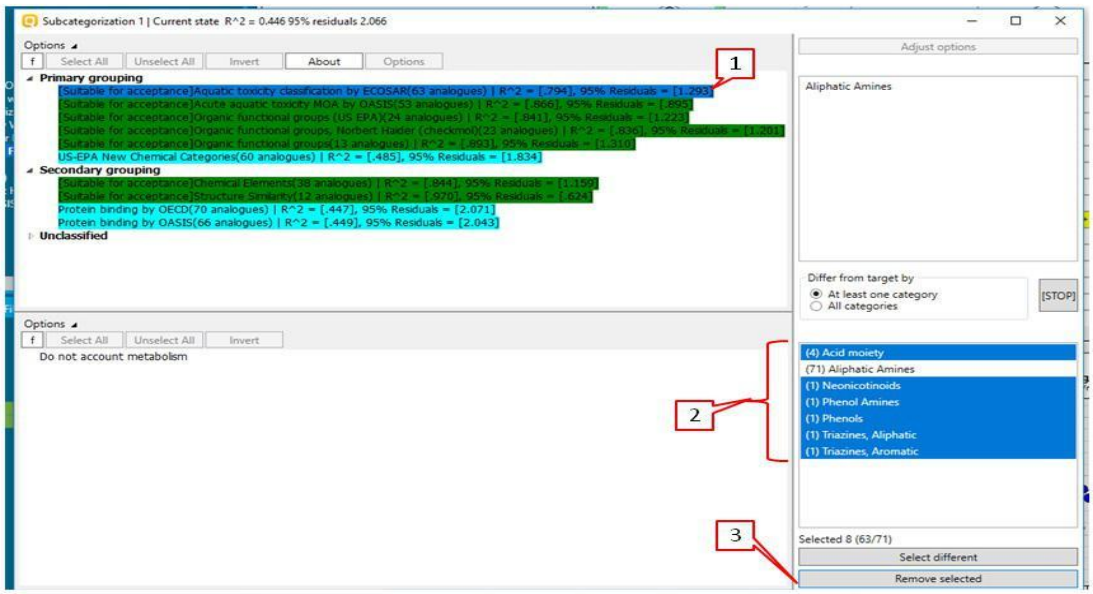


図 8

標的と（プロファイリング結果が）異なる類似物質が消去されます。結果が予測の受容基準を満たす場合、次のメッセージが表示されます（図 9-1）。

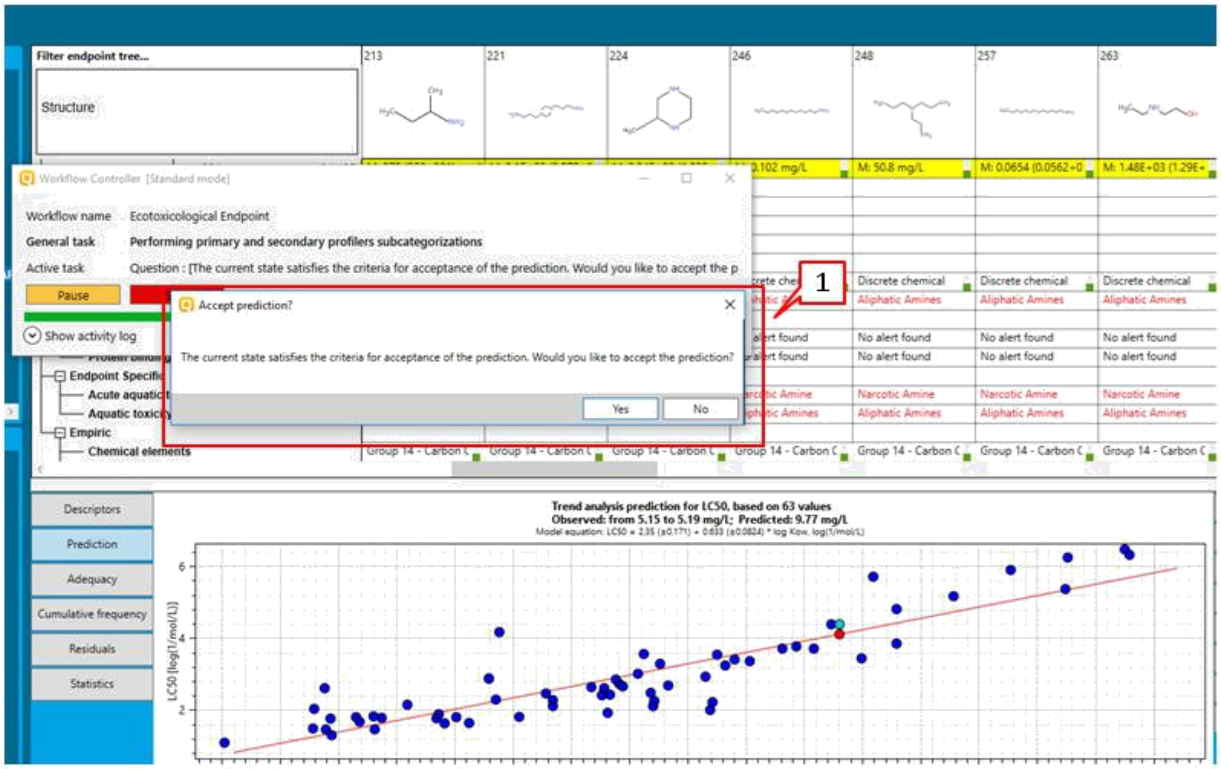


図 9

Yes をクリックして予測を受容するか、予測の基準に満足するまでは No をクリックして、サブカテゴリの他のプロファイラを続行できます。

Yes をクリックすると、次のメッセージが表示されます（図 10）。OK (1)をクリックします。

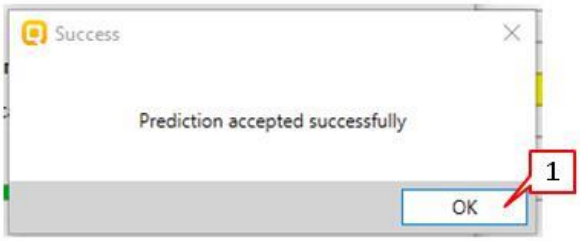


図 10

ワークフローコントローラで Stop（1）をクリックし、ワークフローの終了を確定します（2）（図 11）。

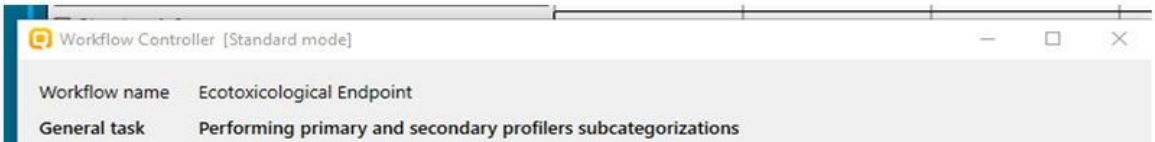
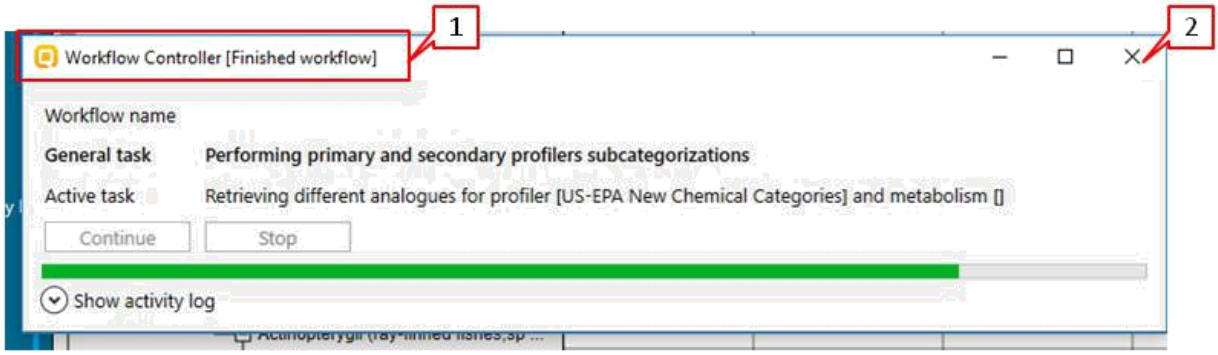


図 11

Finished workflow (1)メッセージがワークフローコントローラに書き込まれます。X 印（2）をクリックしてコントローラを閉じます（図 12)。

図 12



この結果は、データ・マトリックス（13）上に表示されます。ワークフローの過程は、ドキュメントツリー（2）に記載されます。

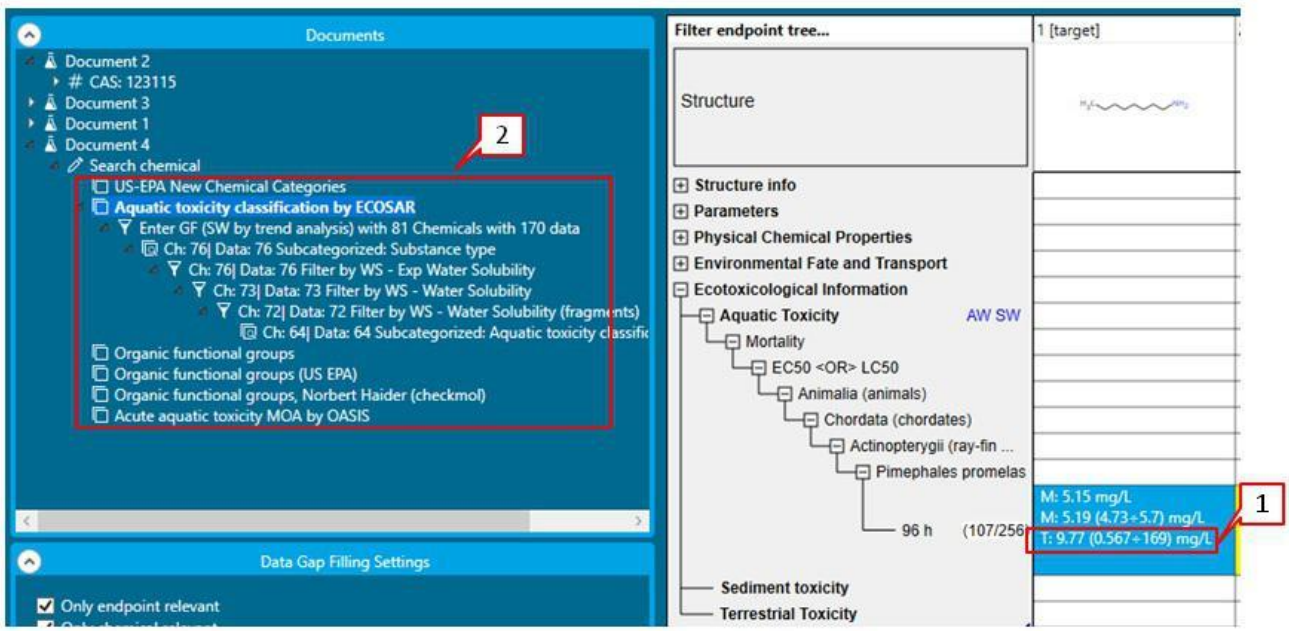


図 13

D.5.7. 一群の化学物質（混合物）の毒性推定の方法論

化学物質の毒性推定をするための 3 つの手法が開発されました。

- 独立したモード：**Independent action** （異なった作用）
- 類似したモード ：**Dose/concentration addition**（用量／加算濃度）
- 特定のモデル：**Specific models**

Independent action および Dose/concentration addition の両方の概念は、混合物中の化学物質が互いの毒性に影響しない、すなわち生物学的な標的部位で互いに相互作用しないという仮定に基づいています。そのような化学物質は、共通、または類似の作用機序によって同様の反応を誘発するか、またはそれらは独立して作用し、異なるエンドポイントを持ったり、かつ/または異なる標的器官をもつかもしれません。
両方の概念は、化学物質混合物の規制リスク評価におけるデフォルトアプローチとして提案されています。

Independent action (response addition, effects addition) （独立した作用：付加の反応、付加の効果） とは、化学物質がお互いに影響を与えず、互いに影響しあわないような異なった作用機序によって互いに独立して作用する場合に、生じます。

Dose/concentration addition (similar action, similar joint action)（用量／加算濃度：類似した作用、類似した合同の作用） とは、混合物内の化学物質が同じ反応機構 / 作用機序で反応し、その効果だけが変化するという場合に生じます。(訳注：他に混合物があろうと影響が無視できて、その注目する化合物は、濃度に応じて効果を強めるという状況に使えるモード）原理的には、単一成分の用量や濃度は、個々の物質の効力を表す倍率によって乗算して加味されます。

同じような作用をする化合物の混合物の効果は、各化合物の効力補正（調整）用量/濃度の合計の効果に相当します。

Specific models （特定のモデル） - この方法論は、混合物の毒性予測をする目的で、一群の化学物質（混合物）に基づいて開発された QSAR モデルを使用することを目指しています。このセクションでは、現在開発中です。

一群の化学物質を扱うこれらの手法をベースに、互変異性体、混合物、代謝物を扱うための様々な方法が開発されています。

Caution

In case the gap filling is entered with set of chemicals with undefined quantities of the components equimolar quantities for all components are assumed for the gap filling calculations.

注意：混合量が不明の成分が混合物に含まれているときにデータギャップ補完を行う場合、全ての混合成分が等モル量で含まれていると仮定してギャップ補完を行います。

D.5.7.1. 例：定量的な混合物の毒性予測

混合量が明らかになっている混合物を、親物質から増えた化学物質として扱います。（訳注：親物質が代謝されるなどして親物質の他にも物質が増えたケースをこの章の例として扱っている）
構成成分の作用機序に基づいた、混合物の予測のための 3 つの新しいオプションが利用できます。

- ***Acting Independently*** (with different mode of action)　（独立した作用：異なる作用機序で）
- ***Acting Similarly*** (with same mode of action)　（類似した作用：同じ作用機序で）

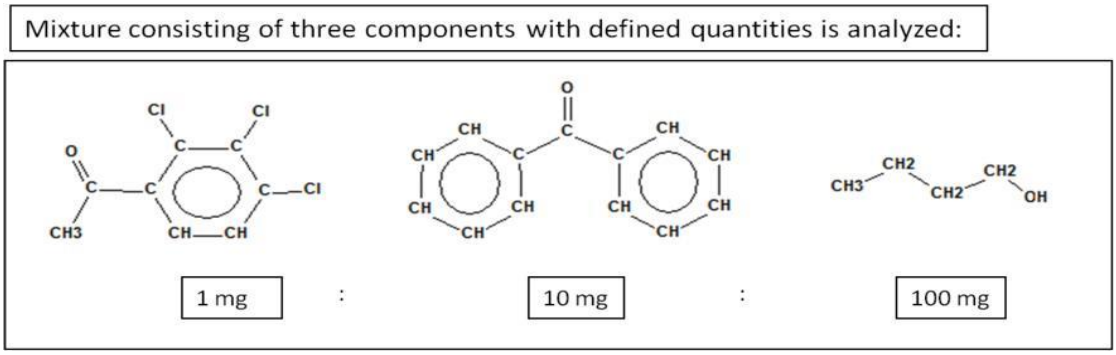


図 1　（量の分かっている 3 つの成分からなる混合物を分析していきます。）

混合物の水生毒性を予測します。　手順は以下の通りです。

- **混合物を入力します。**ツールボックス 4.1 Input の章で物質成分の入力の方法に関する部分を参照してください（訳注：Input モジュールで Composition のアイコンを押し、エディタ図 1 の 3 物質を入力する）　。
  混合物が入力されると、データ・マトリックス（1）に表示されます（図 2)。混合物を予測するには、まず分解する必要があります：Document ツリーの物質を右クリックして（2)、Multiplication　（3）をクリックし、Target multiplication　（4）をクリックします。

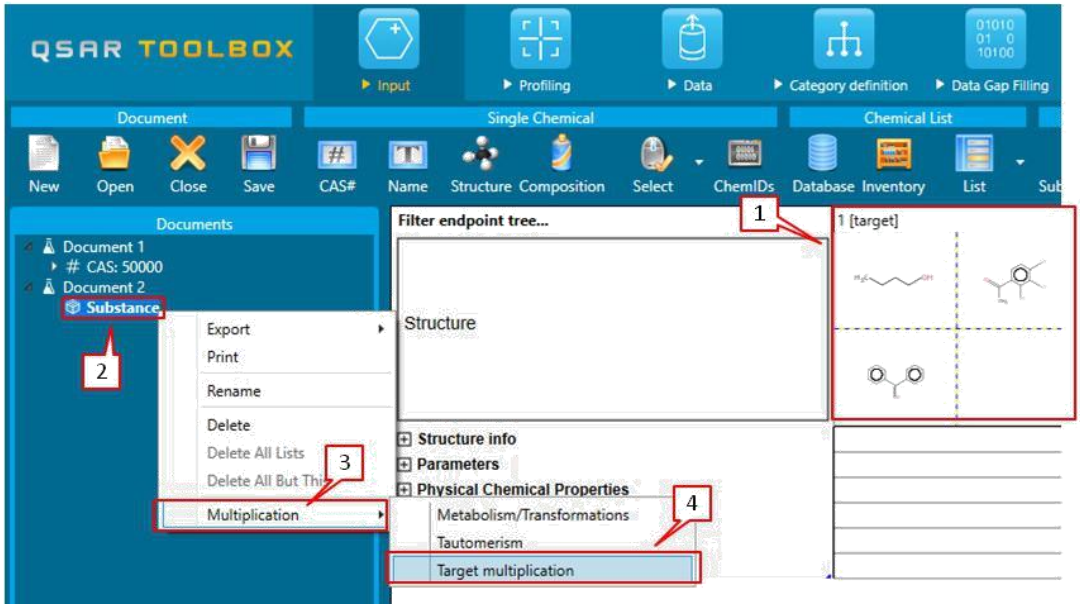


図 2

生成されているコンポーネントの数をユーザに通知するメッセージが表示されます。OK (2)をクリックします。(図 3)

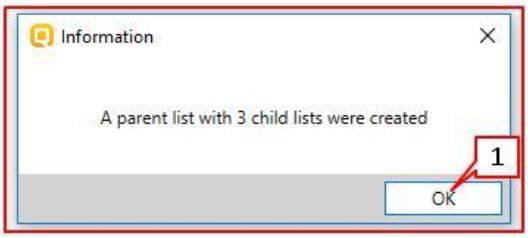


図 3

混合物とその 3 つの成分がデータ・マトリックス上に表示されます（図 4 ）。

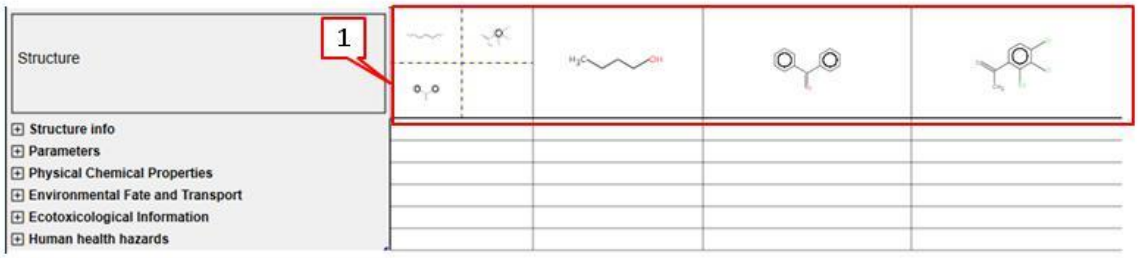


図 4

- 混合物の成分のプロファイリング：プロファイリングパネル（1）に移動し、以下のプロファイルを選択します。US-EPA New chemical category, Acute aquativ toxicity MOA by OASIS, Acute toxicity classification by ECOSAR (2)を選択して Apply(3)を押します。 ご覧の通り、すべての成分は同じ作用機序を持ちます（4)。(図 5)

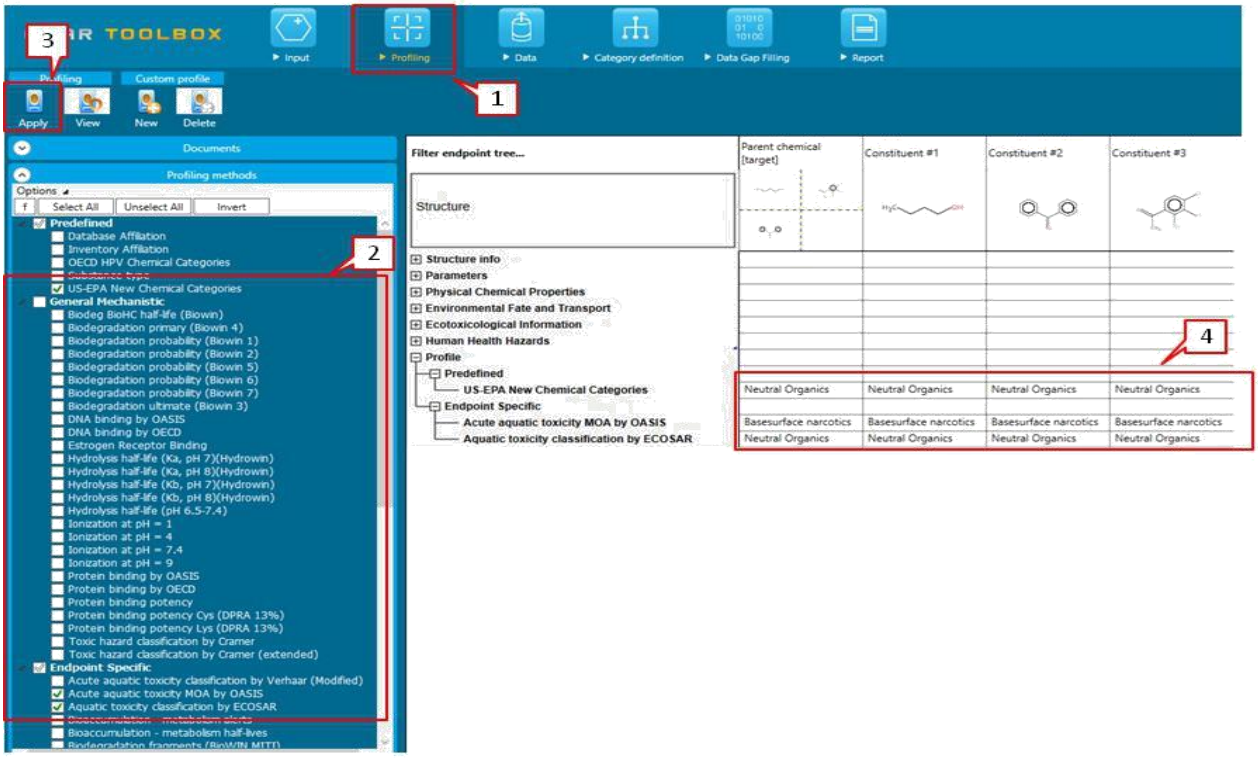


図 5

実験データを収集 します。Data（1）に移動し、Ecotoxicological　データベース（2）を選択し、Gather（3）をクリックします。Ecotoxicological データの一部のみを選択できます（4）。OK（5）をクリックします。(図 6)

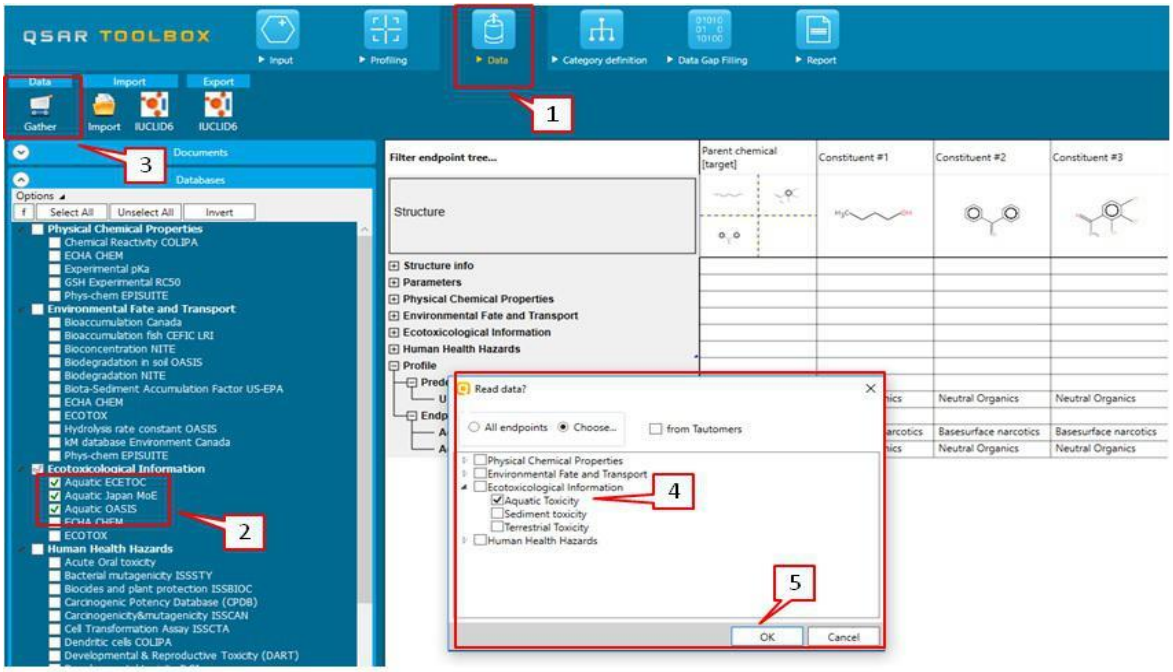


図 6

- ***Similar mode of action***（類似の作用機序）を用いたデータギャップ補完のアプローチ：この特定のケーススタディでは、調査中の混合物の量が決まっており、すべての成分が同じ作用機序を有するということに基づいて、計算をするために類似の作用機序が適用されます。Data gap Filling モジュール（1）に行き、混合物の LC 50,96h、Pimephales promelas（2）についてのセルをクリックして、同じ（皆が共通して持っている）作用機序（3）をクリックしてください。（図 7）

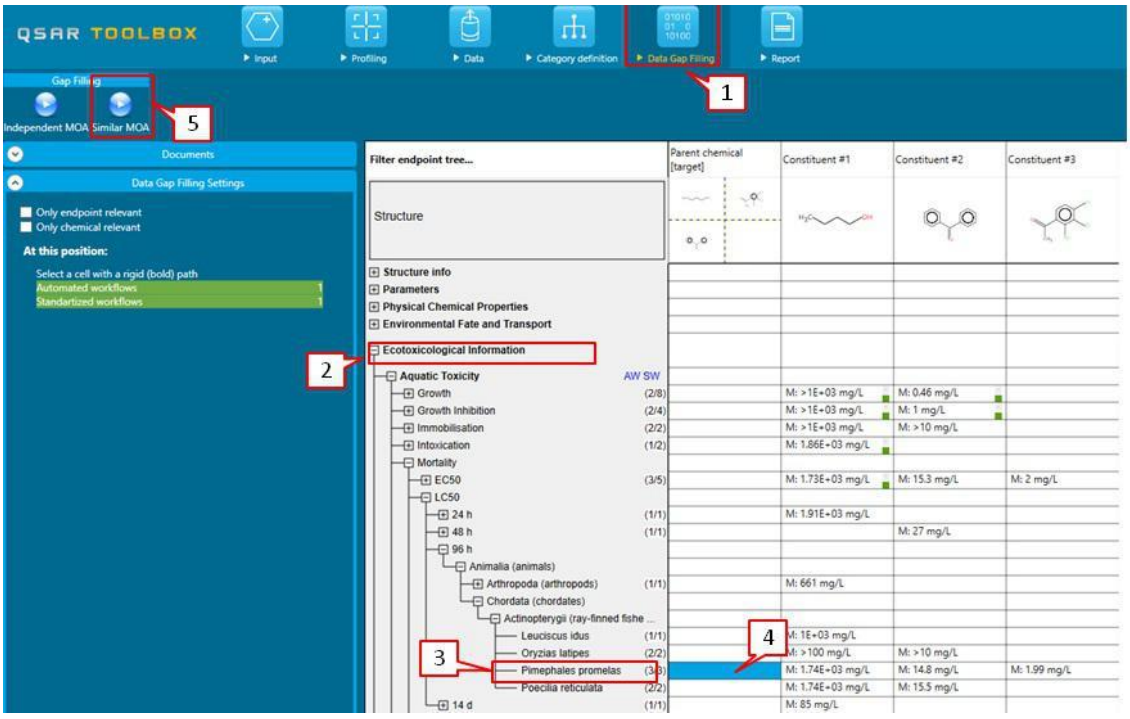


図 7

Possible data inconsistency ウィンドウが表示されますので、そこでデータギャップ補完に使用する単位を選択する必要があれば、ユニットを選択して OK（1）をクリックします。（図 8）

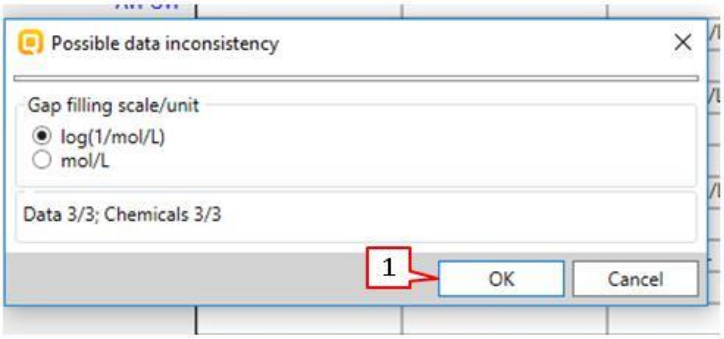


図 8

予測結果は各成分の量を考慮し、LC 50（1）の予測に用量濃度計算を使用します（図 9）。 *Accept the prediction* (2)を押して下さい。

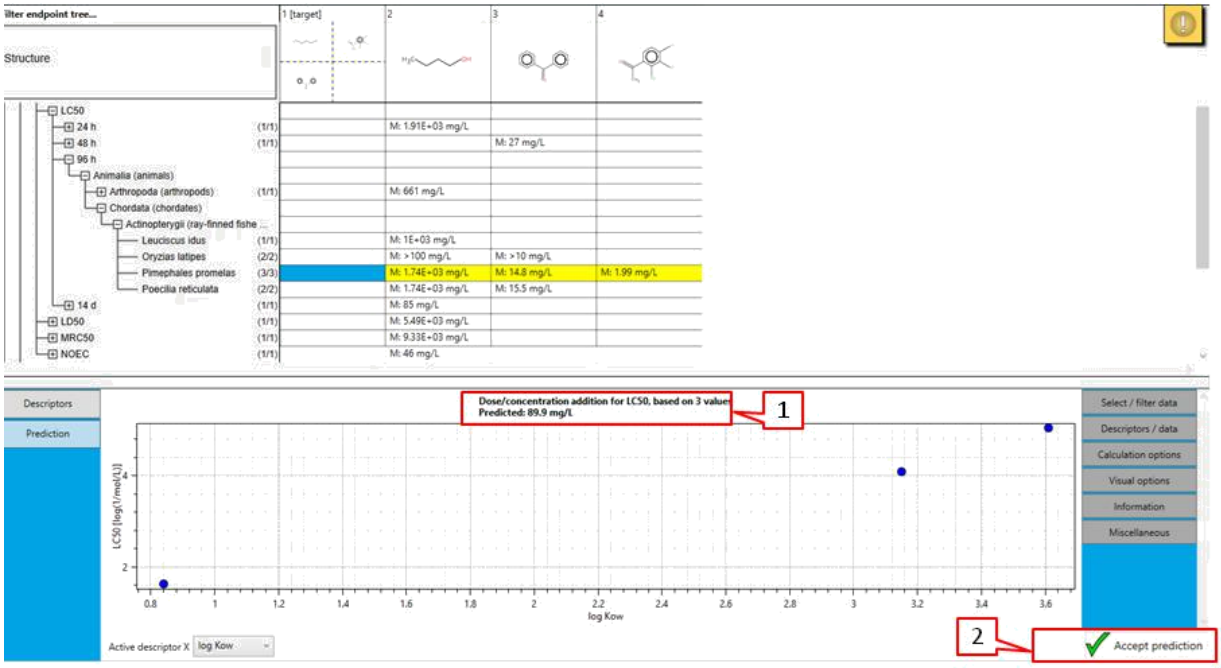


図 9

この予測は、データ・マトリックス上に表示されます。（図 10）

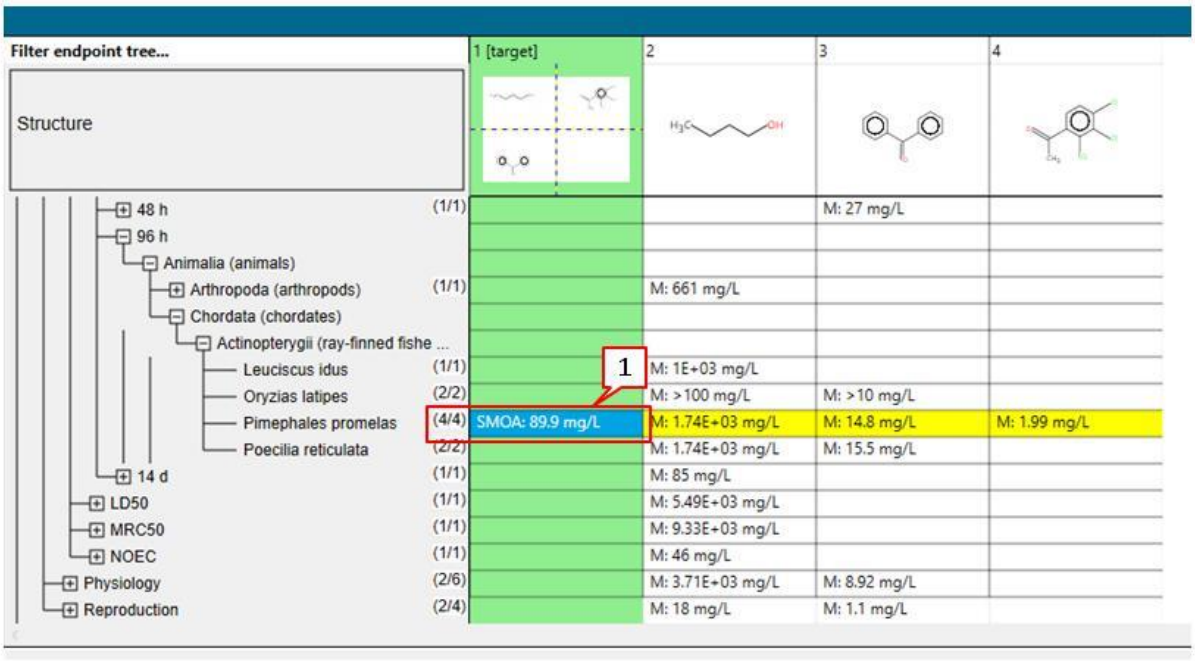


図 10

- 独立した作用機序の場合は次のステップに従って実行して下さい。Data gap Filling モジュール（1）に行き、混合物の LC 50,96h、Pimephales promelas（2）についてのセルをクリックして、同様の作用機序（3）をクリックしてください。（図 11）

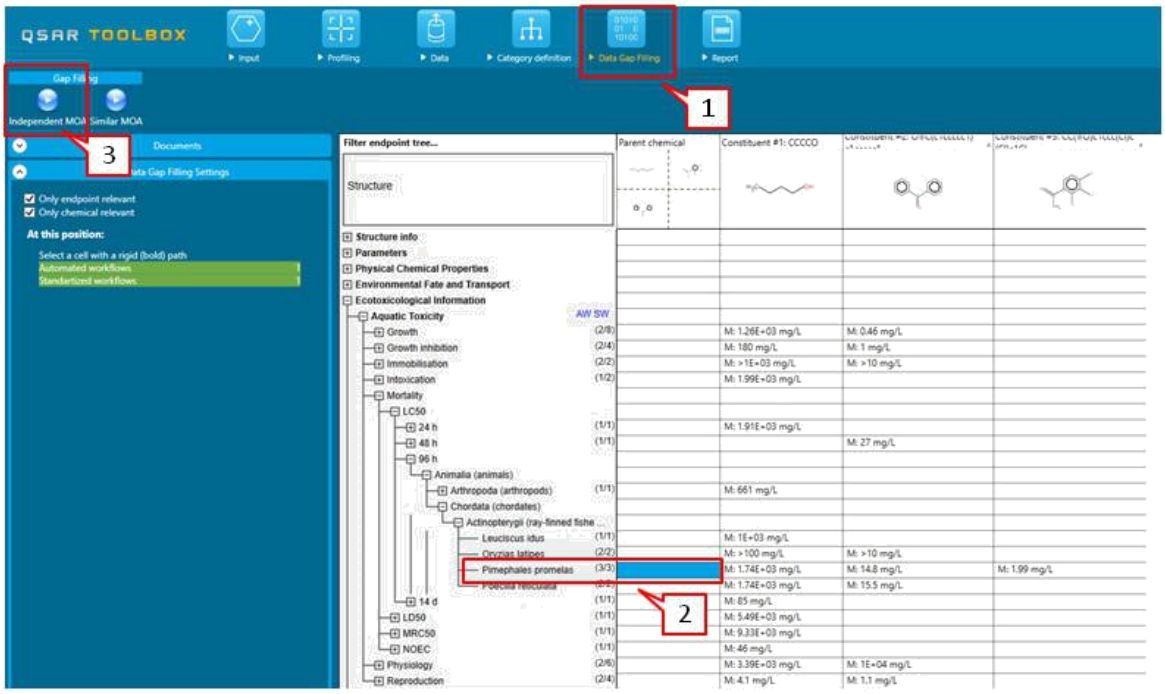
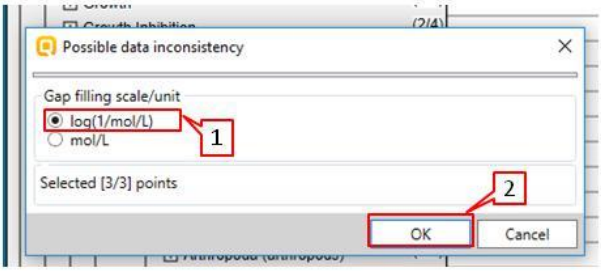


図 11

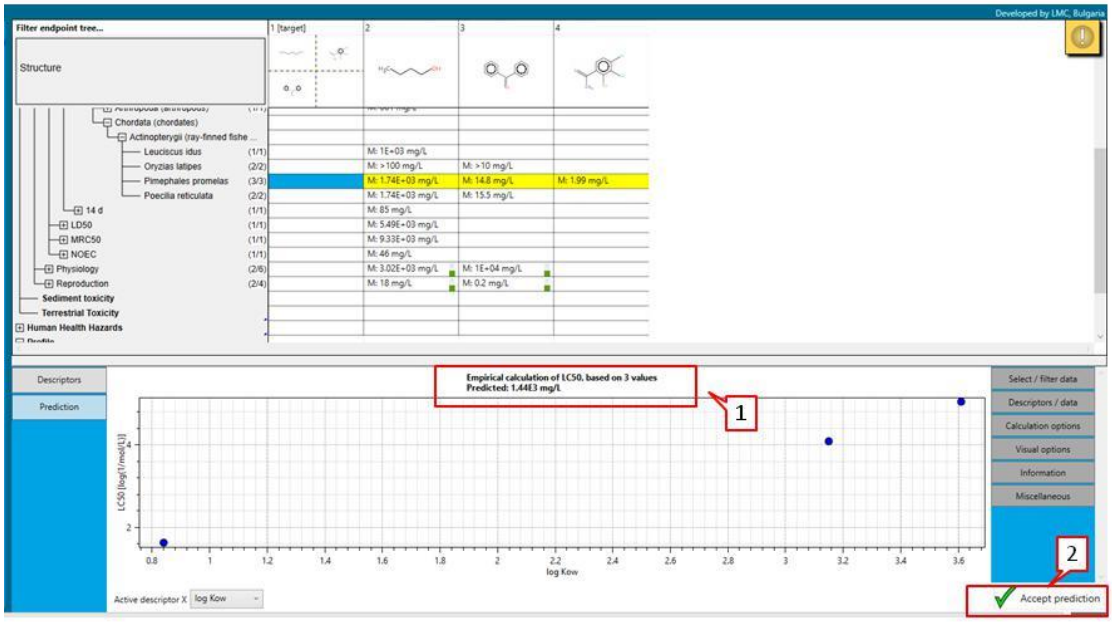
Possible data inconsistency ウィンドウが表示されますので、そこでデータギャップ補完に使用する単位を選択する必要があれば、ユニットを選択して OK（2）をクリックします。（図 12）

図 12



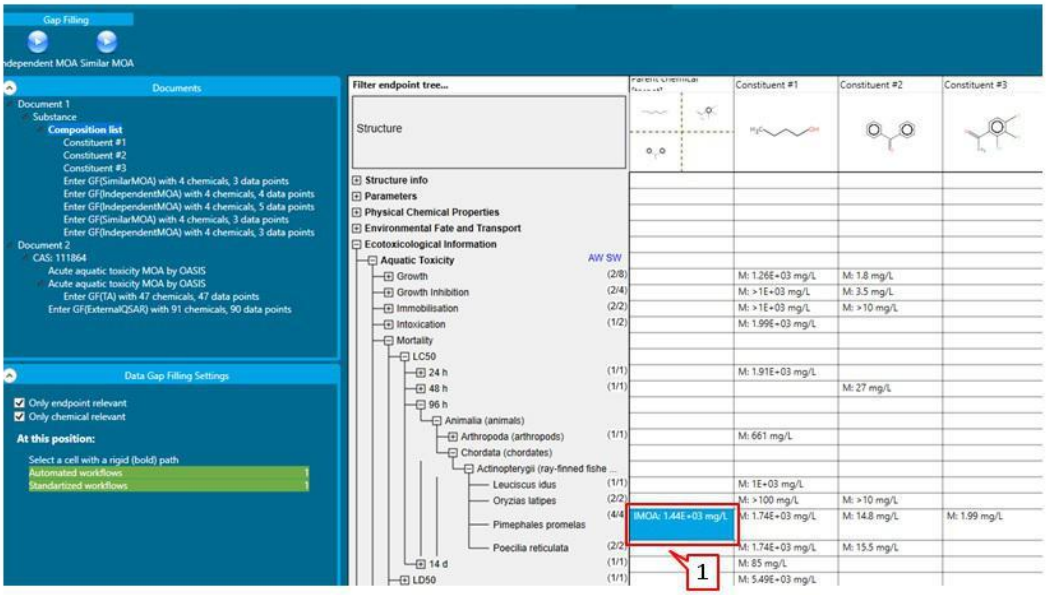
この予測結果（1）は、グラフの上に表示されます（図 13）。*Accept the prediction* (2)を押して下さい。

図 13



この予測は、データ・マトリックス上(1) に表示されます。（図 14）

図 14



D.5.7.1.1. 混合物モデル：方法論

混合物モデル： 2 つの作用機序を使用して混合物を計算します。

Independent MOA（独立した作用機序）

Similar MOA （類似した作用機序）

D.5.7.1.1.1. Independent MOA（独立した作用機序）

独立した作用機序によって作用する成分の混合モデル（増加したときの作用、増加した時の効果）とは、化学物質が互いに影響しあわないような異なった作用機序によって、互いに独立して作用する場合に使います。

仮定 - 組合せ効果は、独立したランダム事象の統計的概念に従い、個々の混合成分に引き起こされる影響から計算することができる

混合物の反応：

$$E(C_{Mix}) = 1 - \prod_{i=1}^{N} [1 - E(C_i)]$$

$E(C_{Mix})$ - 混合物全体によって引き起こされる効果

$E(C_i)$ - 混合物中に存在する濃度で単独で適用された場合、個々の成分が引き起こす影響

問題 - 用量－反応の関係は実際には分かっていない。

近似の種類 -用量－反応が不明であるとき、以下の近似の種類を使用しています（図 1）。

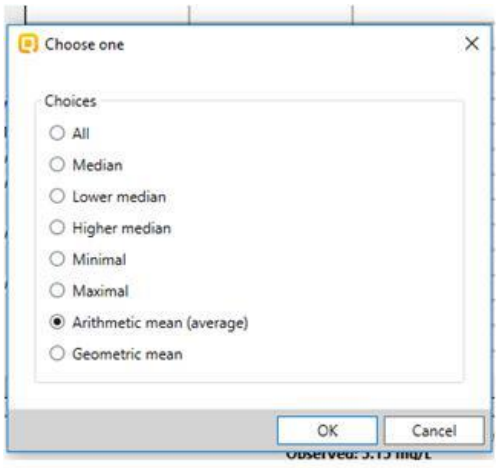


図 1

D.5.7.1.1.1.1.　実例

独立した作用機序によって作用する成分の混合モデル

実例

$$y_{Mixture} = f\left(y_j, j = 1,2,3,\ldots\right)$$

f - Mean, Mode, Median, Max, Min, etc.

平均、モード、中央値、最大値、最小値、など

Arithmetic	Geometric	Harmonic
算術的	幾何学的	調和的
$y_{Mixture} = \sum_{n=1}^{N} x_n y_n$	$y_{Mixture} = \prod_{n=1}^{N} y_n^{x_n}$	$y_{Mixture} = \frac{\sum_{n=1}^{N} x_n}{\sum_{n=1}^{N} \frac{x_n}{y_n}}$

D.5.7.1.1.2. Similar MOA （類似した作用機序）

類似した作用機序によって作用する成分の混合モデル

仮定 - 混合物中の成分は、それらの割合と個々の効力に比例して、共同効果に寄与します。

- **同じ標的サイトの成分が起こす作用**
- **同じ反応機構により成分が起こす作用**
- **成分が類似した効果（反応機構ではない）を持つ**

$$RPF_j^{(i)} = \frac{ED_{resp}^{(i)}}{ED_{resp}^{(j)}}$$

相対的な効能係数 （RPF）

i - インデックス（参照）化学物質

ED_{resp} - 特定の反応（ある割合の動物が反応した、測定された生理学的な値に関してある割合で変化が起きたなど）を引き起こす化学物質投与量（濃度）

$$CED_j^{(i)} = RPF_j^{(i)} d_j$$

化学的等価線量（濃度）（CED）

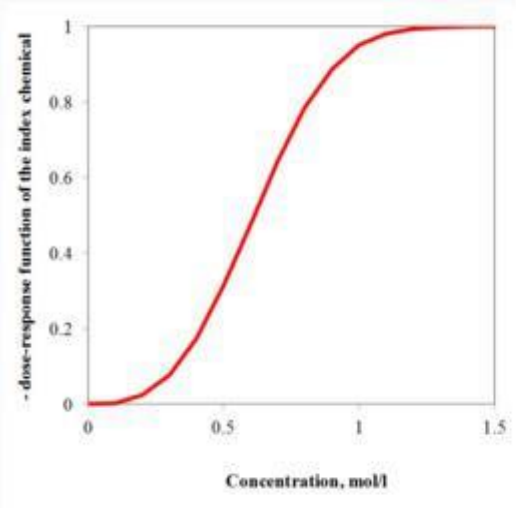
投与量（濃度）d_j における化学物質 j と同じ効果をもたらす基準化学物質 i の投与量（濃度）

$$ICED = \sum_{j=1}^{J} CED_j^{(i)} = \sum_{j=1}^{J} RPF_j^{(i)} d_j$$

化学的等価線量（濃度）の指数 （ICED）

混合物として同じ効果をもたらす基準化学物質 i の等価投与量（濃度）

混合物の毒性効果 - 混合物を曝露した結果としての反応（動物が反応した割合、測定された生理学的な値に関して変化が起きた割合など）



$$Effect^{Mixture} = f_i(ICED)$$

f_i - dose-response function of the index chemical

f_i は基準化学物質における用量―反応の関数

D.5.7.1.1.2.1. 実例

類似した作用機序によって作用する成分の混合モデル

実例

参照物質：成分 1（i = 1）

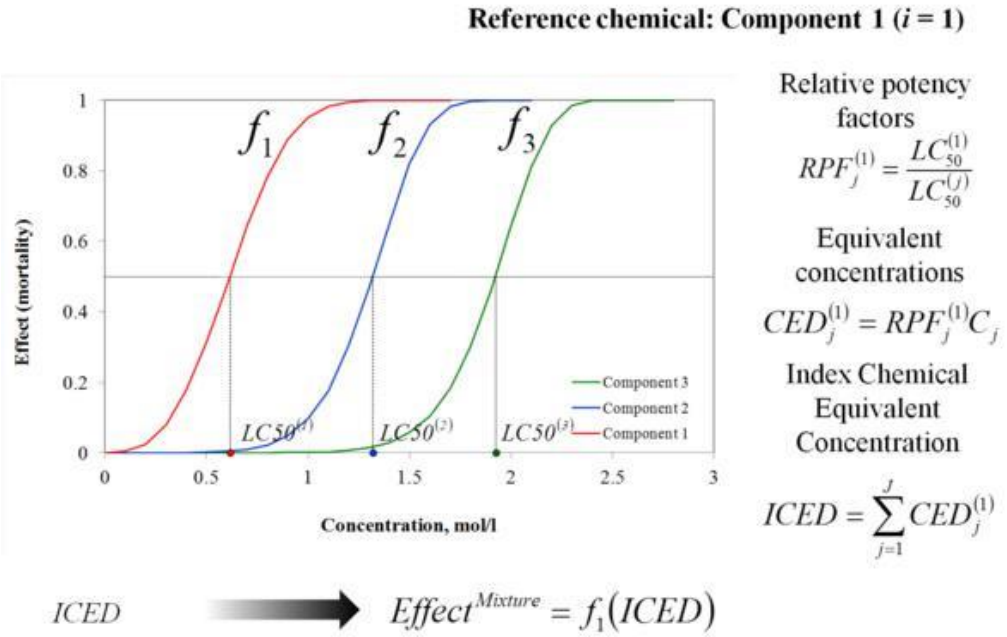


図 1

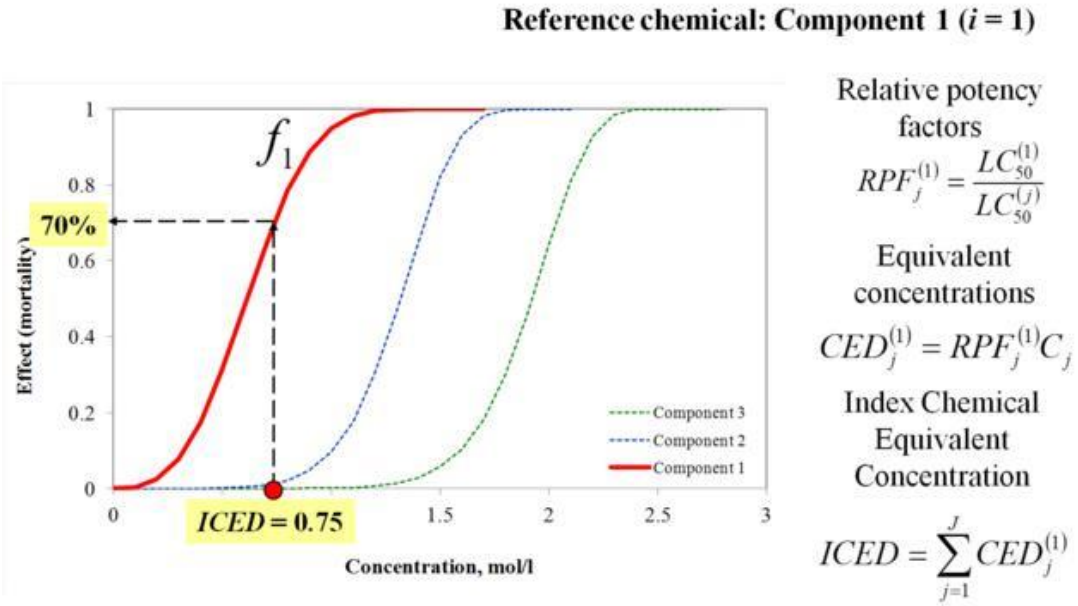


図 2

D.5.7.1.1.2.2.例：

類似した作用機序によって作用する成分の混合モデル

例：

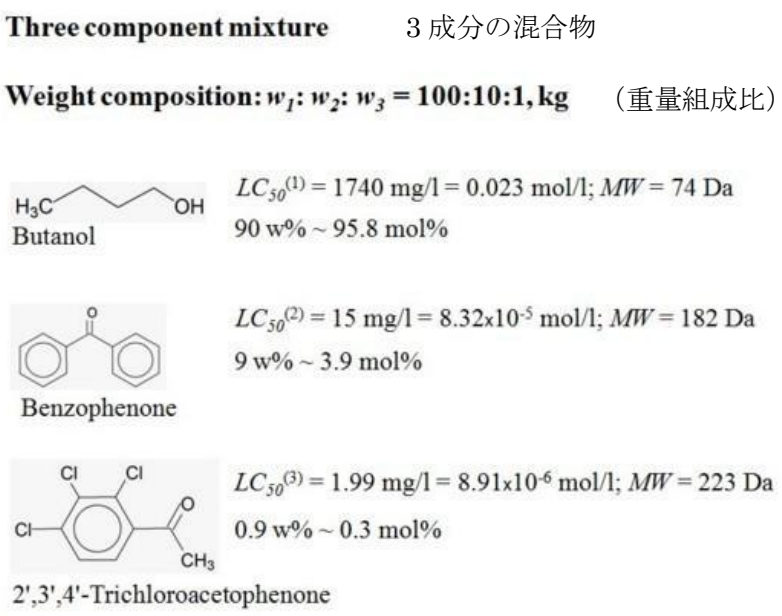


図 1 混合物の成分

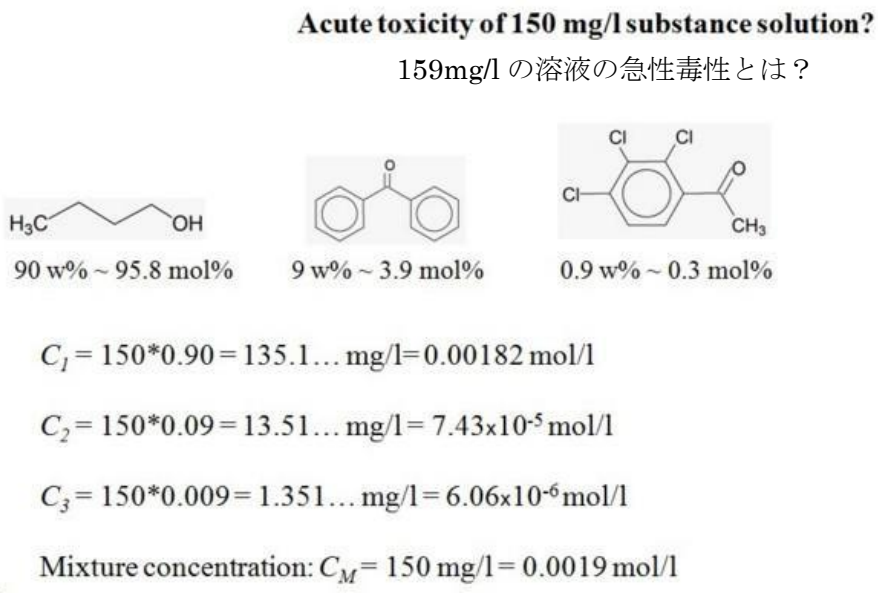
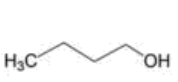
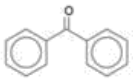


図 2　成分の濃度を決定する

Acute toxicity of 150 mg/l substance solution?



90 w% ~ 95.8 mol%



9 w% ~ 3.9 mol%

0.9 w% ~ 0.3 mol%

$C_1 = 0.00182\ mol/l$ $C_2 = 7.43\text{x}10^{-5}\ mol/l$ $C_3 = 6.06\text{x}10^{-6}\ mol/l$

$LC_{50}^{(1)} = 0.023\ mol/l$ $LC_{50}^{(2)} = 8.32\text{x}10^{-5}\ mol/l$ $LC_{50}^{(3)} = 8.91\text{x}10^{-6}\ mol/l$

$RPF_1^{(1)} = 1$ $RPF_2^{(1)} = 282$ $RPF_3^{(1)} = 2630$

Index Chemical Equivalent Concentration $ICED = \sum_{j=1}^{3} RPF_j^{(1)} C_j = 0.039\ mol/l$

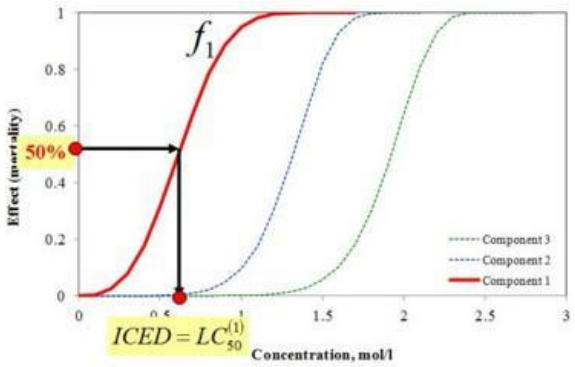
$ICED = 0.039 \rightarrow Effect^{Mixture} = f_1(ICED)$ **Unfortunately f_1 is unknown**

$ICED = 0.039\ mol/l > LC_{50}^{(1)} = 0.023\ mol/l$

the mortality caused by mixture will be > 50%

図 3 各コンポーネントの RPF （相対的効能係数係数）の計算

Quantity of mixture causing 50% mortality?



$Effect^{Mixture} = 50\%$ mortality $\rightarrow ICED = LC_{50}^{(1)} \rightarrow C_M = LC_{50}^{M}$

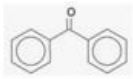
図 4　　50 %致死率を引き起こす、混合物の量の計算を示しています。

Concentration of the substance that will cause 50% mortality: $LC_{50}^{M} = ?$

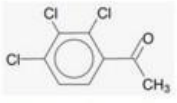
$ICED = C_1 + RPF_2^{(1)} C_2 + RPF_3^{(1)} C_3 = LC_{50}^{(1)} = 0.023\ mol/l$



90 w% ~ 95.8 mol%



9 w% ~ 3.9 mol%



0.9 w% ~ 0.3 mol%

Mole composition:
$x_1 : x_2 : x_3 = 95.8 : 3.9 : 0.3$

$C_2 = \frac{3.9}{95.8} C_1$ $C_3 = \frac{0.3}{95.8} C_1$

$$ICED = C_1 + \frac{3.9}{95.8} RPF_2^{(1)} C_1 + \frac{0.3}{95.8} RPF_3^{(1)} C_1 = 0.023\ mol/l$$

$C_1 = 0.0011\ mol/l$ $C_2 = 4.48 \times 10^{-5}\ mol/l$ $C_3 = 3.66 \times 10^{-6}\ mol/l$

$$LC_{50}^{M} = 0.00115\ mol/l = 89.9\ mg/l$$

図 5　50 %致死率を引き起こす、混合物の濃度の計算を示しています。

D.6. Report（レポート）

モジュールの目的

レポートモジュールを使用すると、ツールボックスで実行した予測のレポートを生成できます。ツールボックスには、既定のレポートテンプレートと、ユーザーが独自のカスタムテンプレートを決められるテンプレートエディタが含まれています。

QSAR Toolbox v.4.1（1）では、4 つのレポートテンプレートが開発され、実装されています（図 1）。

- 予測レポート
- データ・マトリックス　レポート
- カテゴリ　レポート
- QMRF モデル　レポート

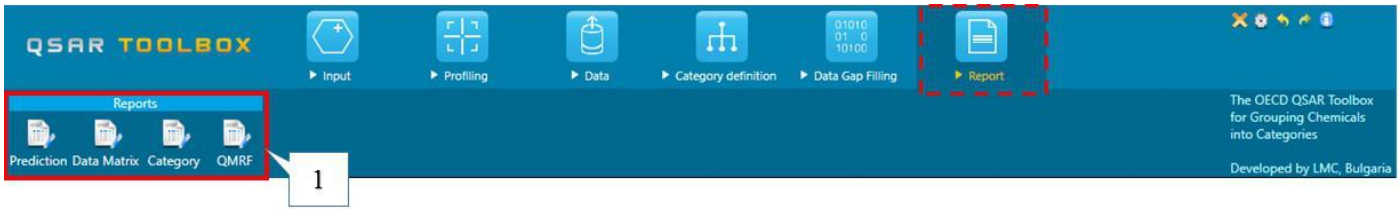


図 1

D.6.1. 予測レポート

6.1.1.単一の化学物質の予測

Prediction （予測）ボタンは 2 つのファイルを生成します。1 つは予測がどのように得られたかを示し、もう 1 つは標的化学物質と、使用された類似物質の情報を示します。

レポートを作成するには、予測データ（1）のあるセルを選択し、次に **Prediction**（2）をクリックする必要があります。選択したセル内で、複数の予測結果が利用可能な場合は、ポップアップウィンドウが表示されます。**Prediction** （3）を選択し、OK（4）（図 1）をクリックしてください。

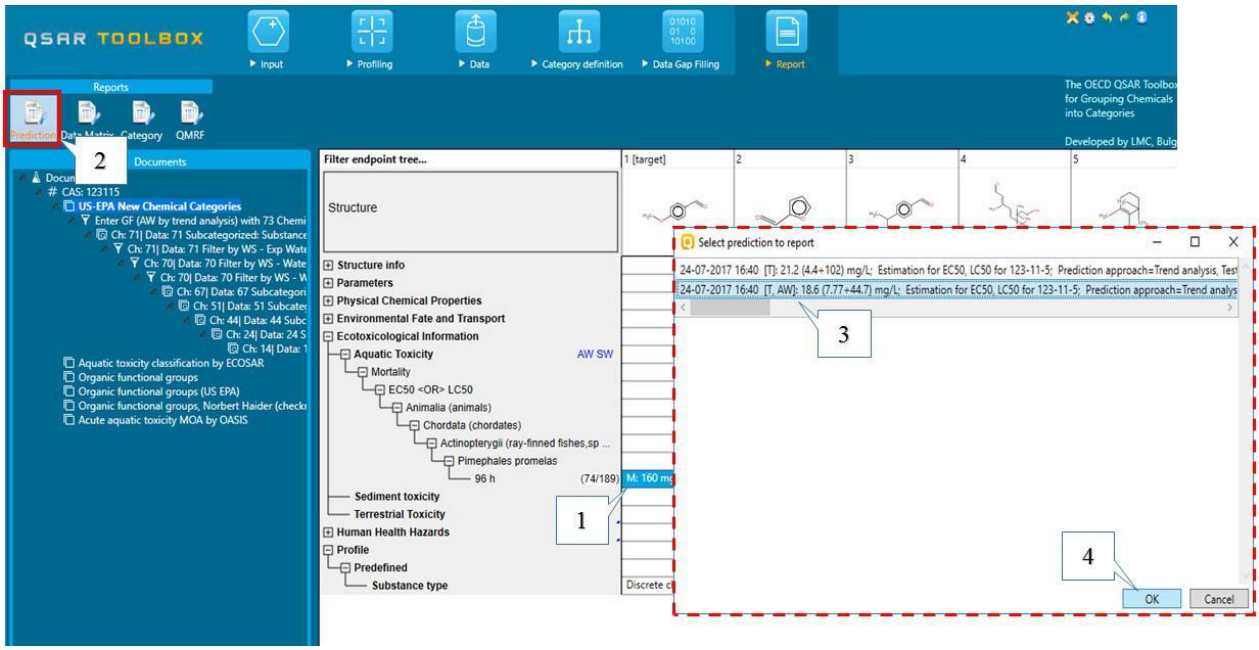


図 1

予測を選択した後は、 **Report wizard**（**レポートウィザード**） が表示されます。レポート（1）に含まれるセクションはカスタマイズすることができます。*Back* ボタンと *Next* ボタン（2）を使用してセクションを行き来できます（図 2）。

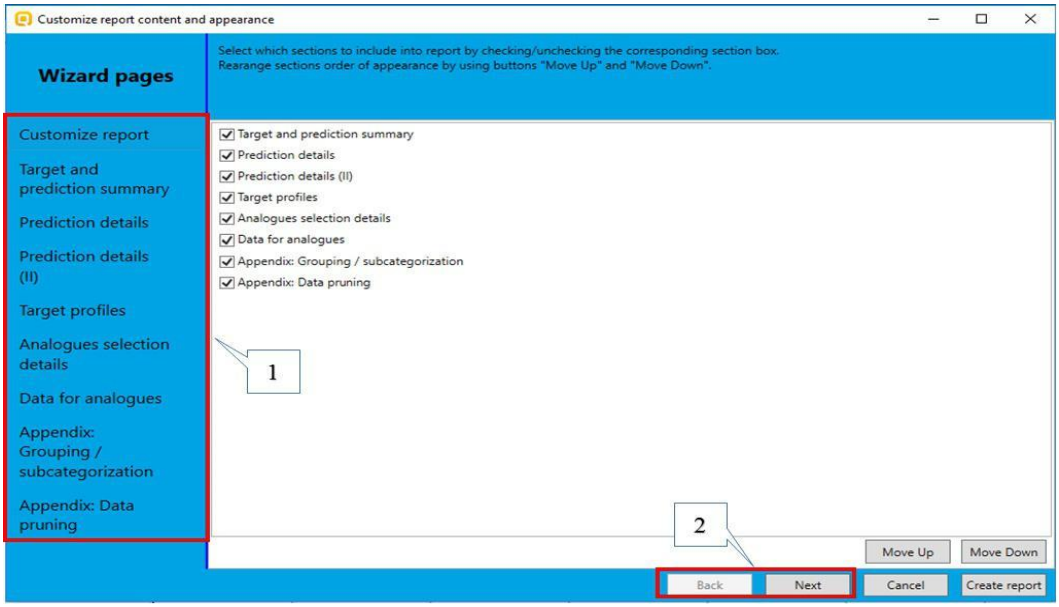


図 2

- **Customize report**（レポートのカスタマイズ） - このセクションでは、対応するチェックボックスをオン/オフ（1）することで、レポートにセクションを含めたり除外したりすることができます。セクションの表示順序は、**Move Up/Down** ボタン（2）（図 3）で変更することもできます。

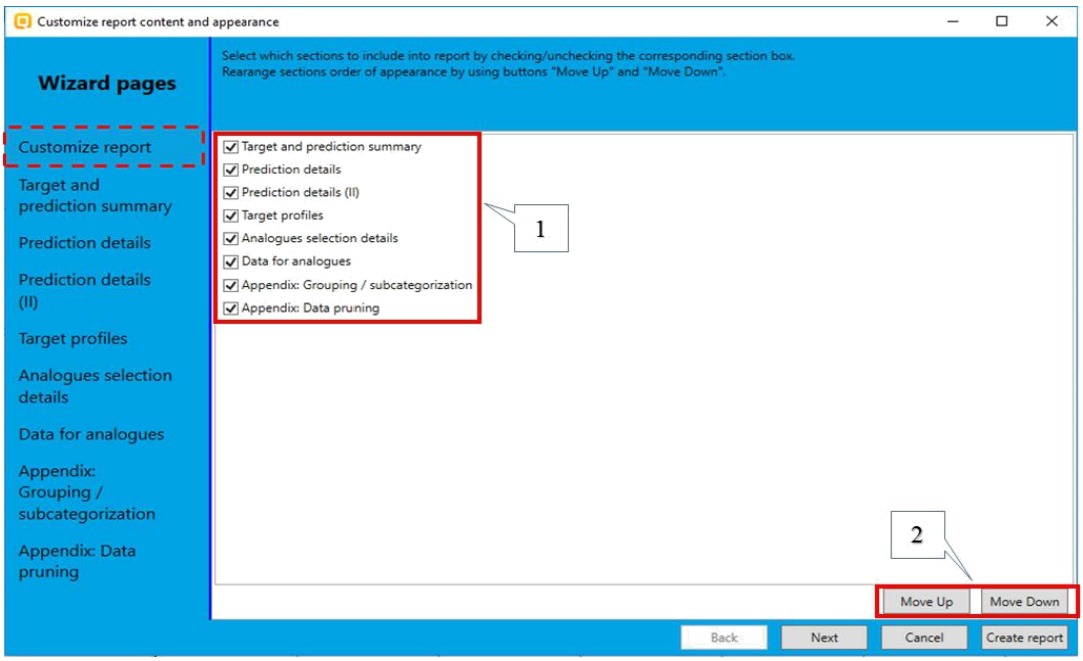


図 3

- **Target and prediction summary**（標的物質と予測の要約） - このセクションには、標的化学物質の物質 ID と予測結果が含まれます。システムによって自動的に設定された入力欄が表示されます。ここでは、著者、連絡先の詳細および要約情報（1）に関する情報を追加することができます。標的物質は多くの別名を持つ可能性があるため、レポートする化学物質名の最大数に関する追加オプションも作成されます。All (2)ラジオボタンを選択するか、または Count（3）ラジオボタンを選択して、化学物質名の正確な数を指し示すように、レポート内の標的物質の利用可能な化学物質名をすべて表示するかどうかを選択できます。（図 4）

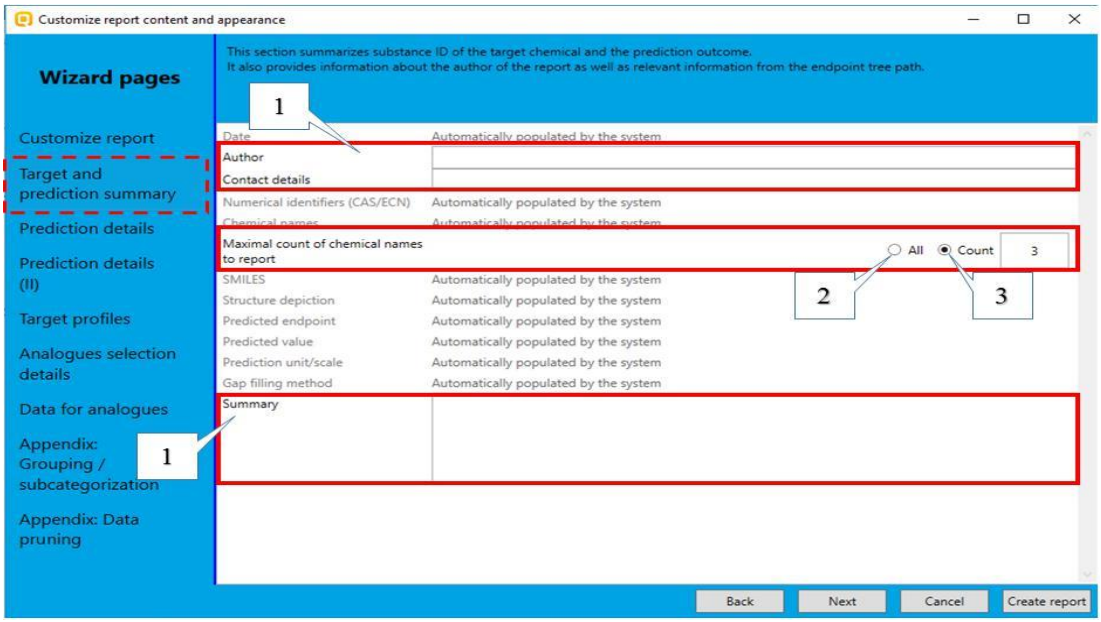
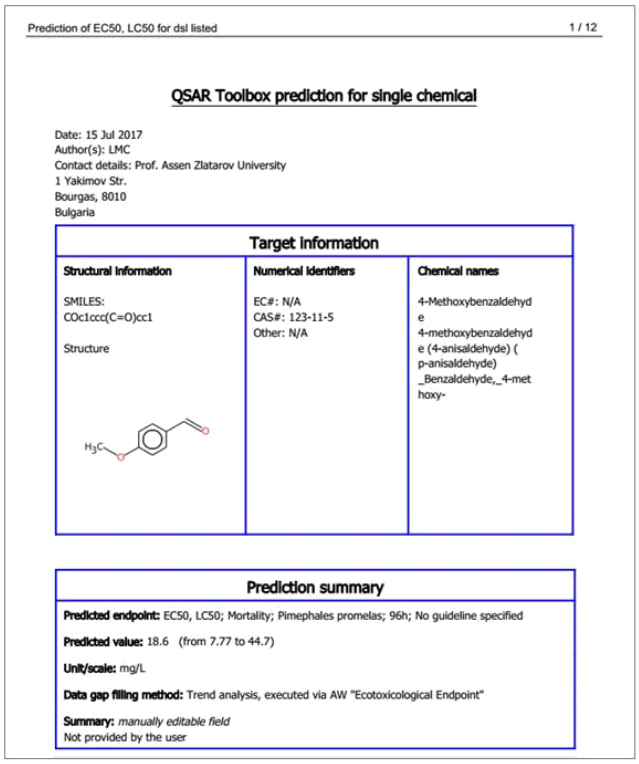


図 4

生成されたレポートでこのセクションがどのように表示されるかを図 5 に例示します。



Prediction of EC50, LC50 for dsl listed 1 / 12

QSAR Toolbox prediction for single chemical

Date: 15 Jul 2017
Author(s): LMC
Contact details: Prof. Assen Zlatarov University
1 Yakimov Str.
Bourgas, 8010
Bulgaria

Target information

Structural information	Numerical identifiers	Chemical names
SMILES: COc1ccc(C=O)cc1 Structure	EC#: N/A CAS#: 123-11-5 Other: N/A	4-Methoxybenzaldehyde 4-methoxybenzaldehyde (4-anisaldehyde) (p-anisaldehyde) _Benzaldehyde,_4-methoxy-

Prediction summary

Predicted endpoint: EC50, LC50; Mortality; Pimephales promelas; 96h; No guideline specified

Predicted value: 18.6 (from 7.77 to 44.7)

Unit/scale: mg/L

Data gap filling method: Trend analysis, executed via AW "Ecotoxicological Endpoint"

Summary: *manually editable field*
Not provided by the user

図 5

- **Prediction details** （予測の詳細）と **Prediction details (II)** － Prediction details セクションは、予測とその信頼性に関する詳細を提供します。**Prediction details (II)**は任意であり、ギャップ補完方法に応じて予測に関する特定の情報を提供しています。予測詳細情報はシステムによって自動的に入力され（1)、Prediction details（II）セクションは自分で入力する欄で構成されます（図 6)。

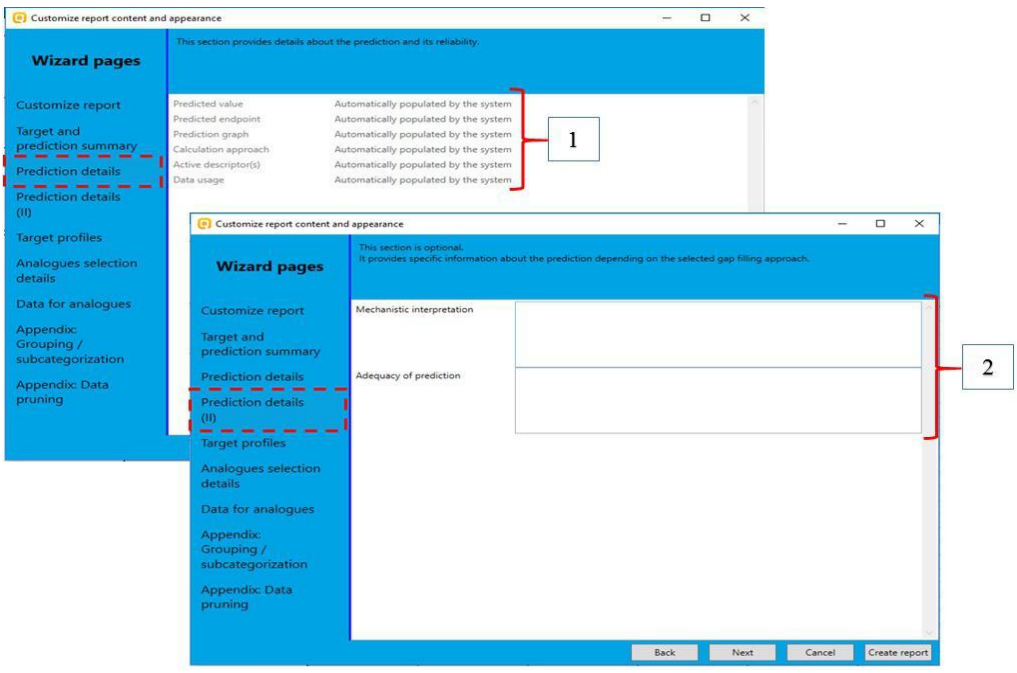


図 6

レポート内に出力された *Prediction details* 欄の例を図 7 に示します。

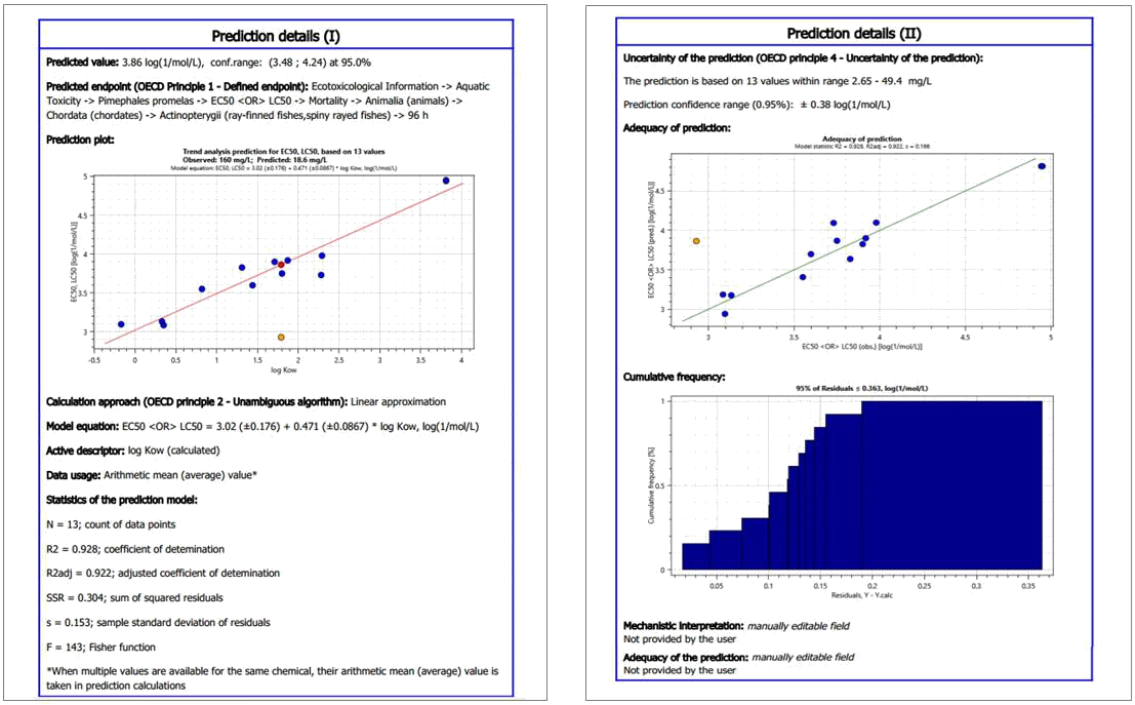


図 7

- **Target profiles** （標的プロファイル） - このセクションでは、予測に使用されるプロファイルを要約します。最初のサブセクション Profiles のシステムが、一次カテゴリ化とサブカテゴリ化に使用したプロファイラを、適用した通りに提供します（1）（図 8）。プロファイル付きのリストは、（Add を押して）新規追加、または一部削除などで拡張できます（2）。デフォルトで *Show in report* オプションがチェックされています（3）。これはオプションの設定です。*Exported in an Excel format* (4)でプロファイラのリストを Excel 形式（4）でエクスポートできます。

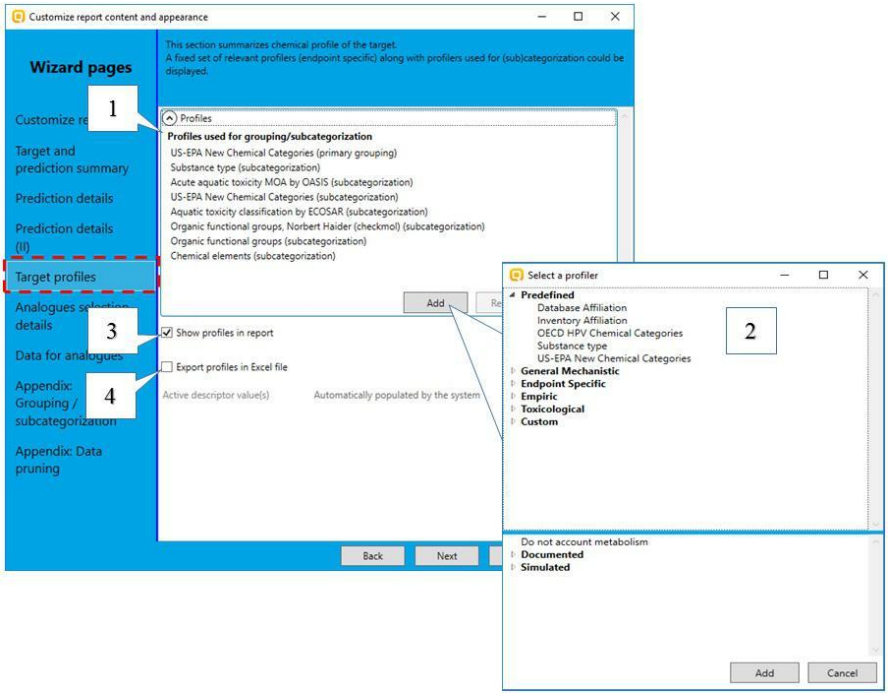
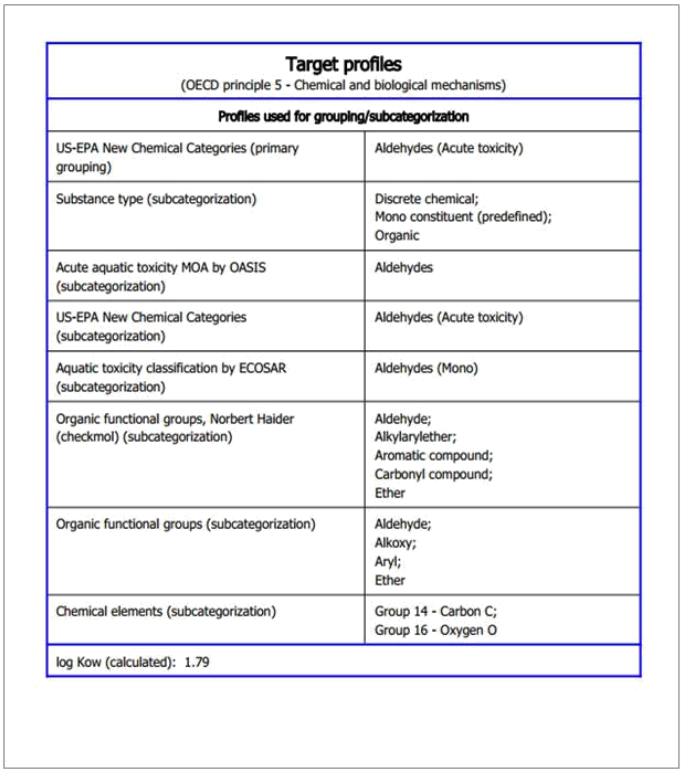


図 8

現在のセクションが出力されたレポートを図 9 に示します。



Target profiles
(OECD principle 5 - Chemical and biological mechanisms)

Profiles used for grouping/subcategorization	
US-EPA New Chemical Categories (primary grouping)	Aldehydes (Acute toxicity)
Substance type (subcategorization)	Discrete chemical; Mono constituent (predefined); Organic
Acute aquatic toxicity MOA by OASIS (subcategorization)	Aldehydes
US-EPA New Chemical Categories (subcategorization)	Aldehydes (Acute toxicity)
Aquatic toxicity classification by ECOSAR (subcategorization)	Aldehydes (Mono)
Organic functional groups, Norbert Haider (checkmol) (subcategorization)	Aldehyde; Alkylarylether; Aromatic compound; Carbonyl compound; Ether
Organic functional groups (subcategorization)	Aldehyde; Alkoxy; Aryl; Ether
Chemical elements (subcategorization)	Group 14 - Carbon C; Group 16 - Oxygen O
log Kow (calculated): 1.79	

図 9

- **Analogues selection details** （類似物質の選択の詳細） - このセクションでは、類似物質がどのように選択されたかを示します。選択されたデータベース、カテゴリの境界、適用範囲のドメイン、手動で削除され追加されたデータの数（もしあれば）が表示されます（図 10）。この情報は自動的に与えられます。

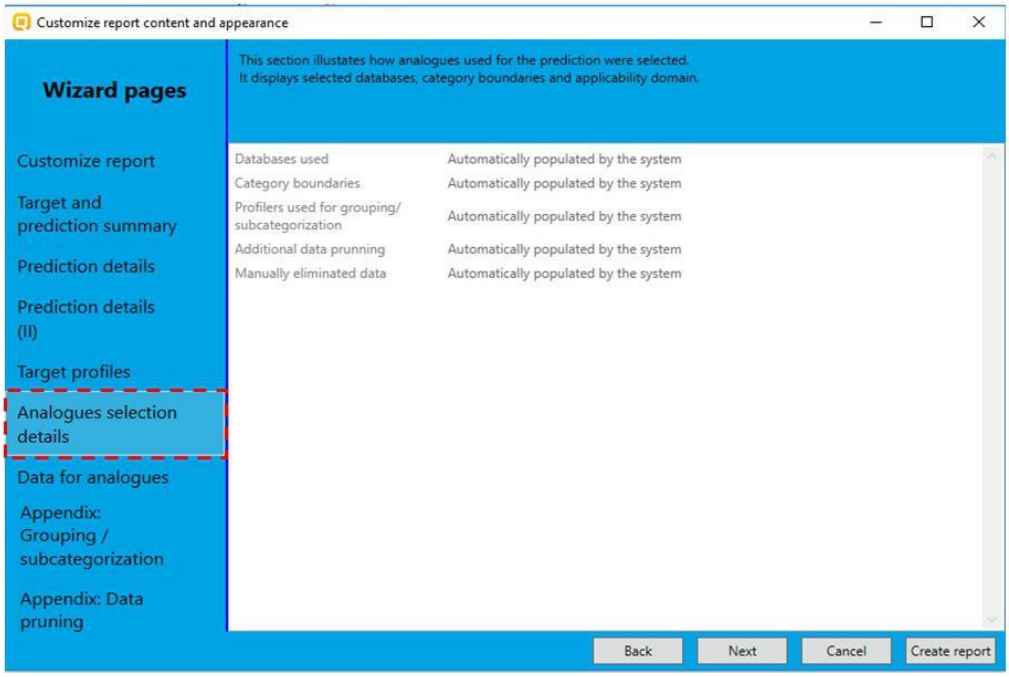
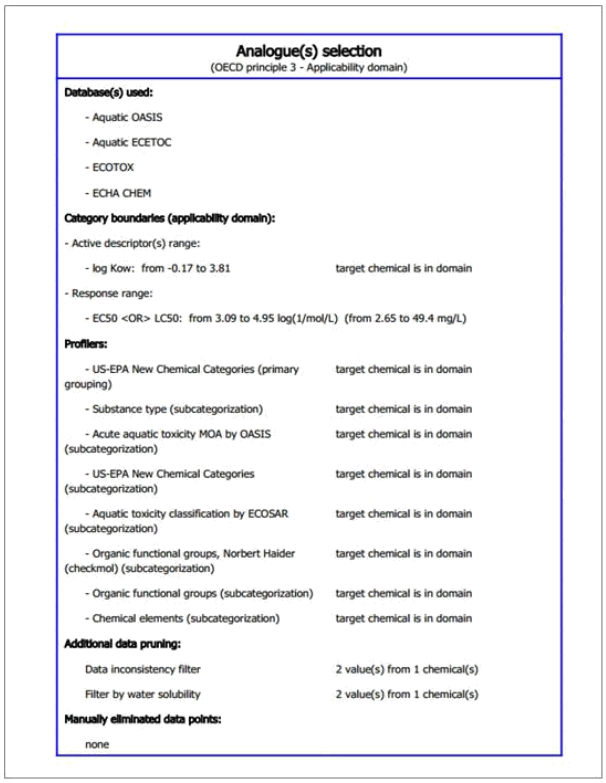


図 10

生成されたレポートで現在のセクションがどのように表示されるかを図 11 に例示します。



Analogue(s) selection
(OECD principle 3 - Applicability domain)

Database(s) used:

- Aquatic OASIS
- Aquatic ECETOC
- ECOTOX
- ECHA CHEM

Category boundaries (applicability domain):

- Active descriptor(s) range:
  - log Kow: from -0.17 to 3.81 — target chemical is in domain
- Response range:
  - EC50 <OR> LC50: from 3.09 to 4.95 log(1/mol/L) (from 2.65 to 49.4 mg/L)

Profilers:

- US-EPA New Chemical Categories (primary grouping)	target chemical is in domain
- Substance type (subcategorization)	target chemical is in domain
- Acute aquatic toxicity MOA by OASIS (subcategorization)	target chemical is in domain
- US-EPA New Chemical Categories (subcategorization)	target chemical is in domain
- Aquatic toxicity classification by ECOSAR (subcategorization)	target chemical is in domain
- Organic functional groups, Norbert Haider (checkmol) (subcategorization)	target chemical is in domain
- Organic functional groups (subcategorization)	target chemical is in domain
- Chemical elements (subcategorization)	target chemical is in domain

Additional data pruning:

Data inconsistency filter	2 value(s) from 1 chemical(s)
Filter by water solubility	2 value(s) from 1 chemical(s)

Manually eliminated data points:

none

図 11

- **Data for analogues** （類似物質に関するデータ） - このセクションでは、パラメータ、プロファイラ、実験データを含み、予測を得るために使用される、類似物質の関連詳細情報を提供します。それぞれのメニューを開くと、利用可能なパラメータを含むリストが表示されます（1）（図 12）。予測の導出に使用されるプロファイラのリストは、Profiles サブセクション（2）に示され、リストは追加したり、削除したりできます（3）。類似物質の実験データは、Add（5）ボタンをクリックして Include experimental data（実験データを含める）サブセクション（4）の中に追加することができます。エンドポイントと、（フィルターメニューを使用するか、ツリーを展開して目的のエンドポイントを見つけることができます）追加の付随データを選択できます。Finish ボタン（6）をクリックすると、データの選択が完了します。

注意:
このセクション(で説明したこと)はデータ・マトリックスファイルを生成してレポートにされます。予測を得るために使用される類似物質の実験データは、システムによって自動的に提供されることに留意してください。

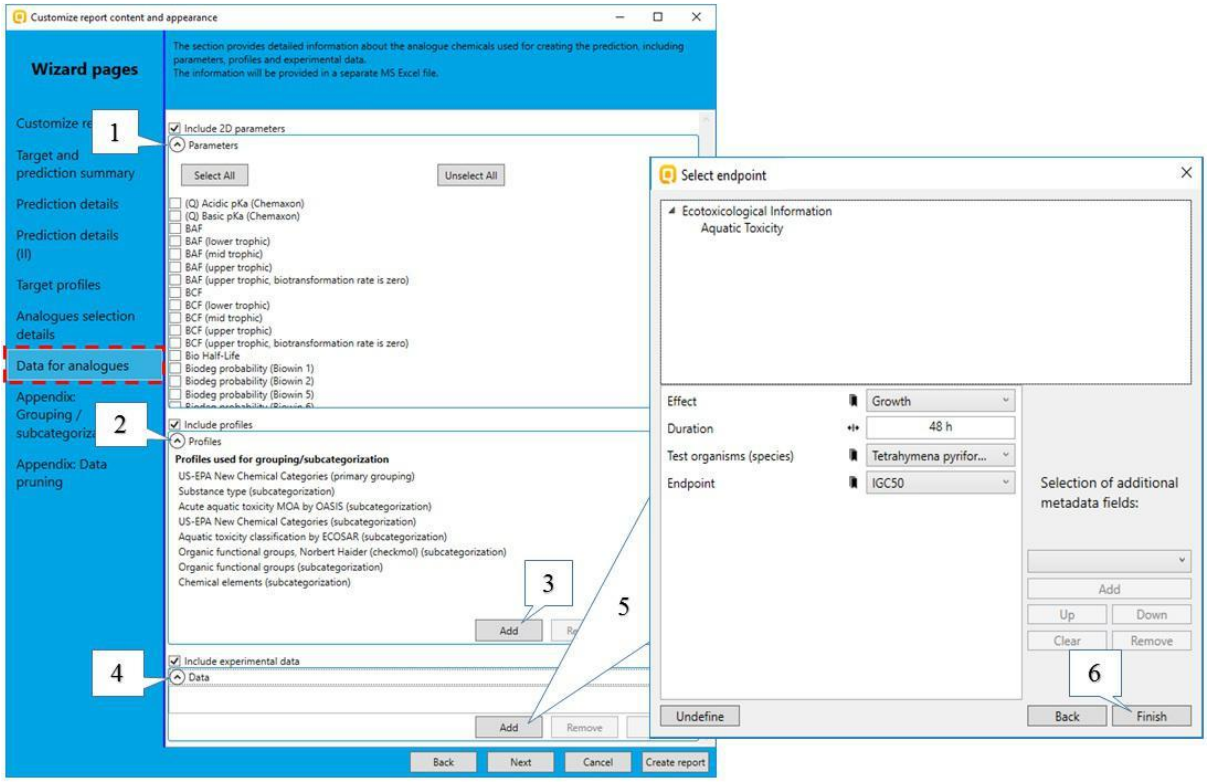


図 12

- **.Appendix: Grouping / subcategorization** （付録：グルーピング／サブカテゴリ化）と、 **Appendix: Data pruning** （付録：データの切り落とされた部分）-これらのセクションには、一次カテゴリ化に使用されるカテゴリ、およびシステムによって自動的に設定されるサブカテゴリ化中に削除されるカテゴリとメタデータ（1-2）が含まれます。手動で除去されたデータ（3）について、その正当性を提供することができます（図 13）。**Appendix: Data pruning** は、手動ワークフローが実行されたときにのみ表示されます。

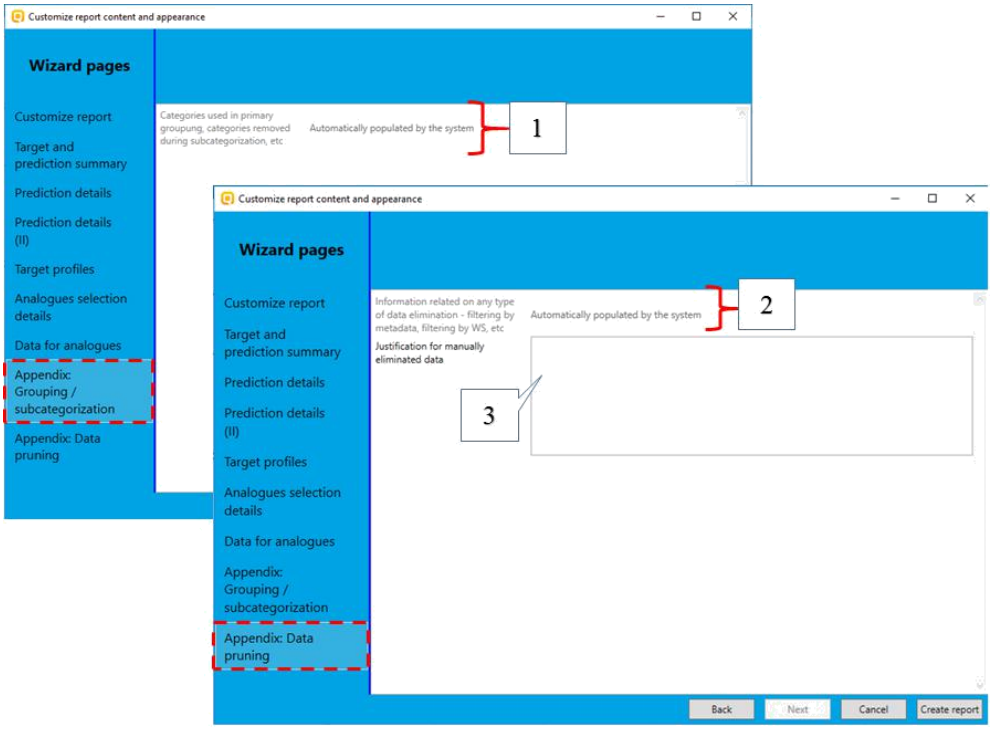
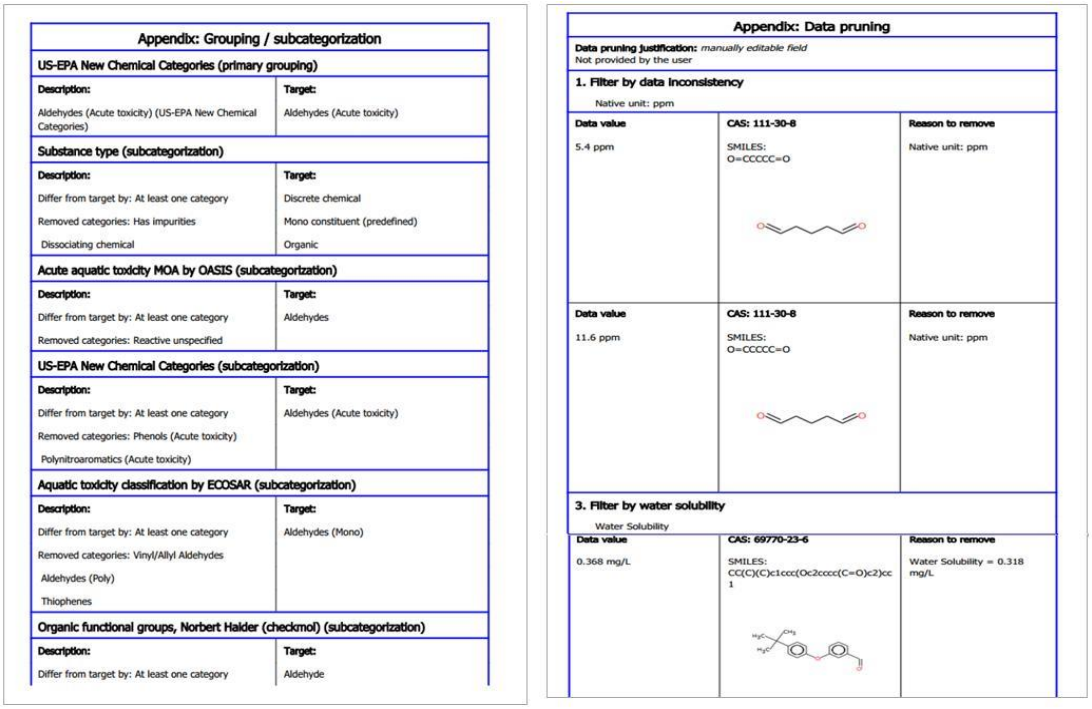


図 13

生成されたレポート内でこのセクションがどのように表示されるかを図 14 に示します。



Appendix: Grouping / subcategorization

US-EPA New Chemical Categories (primary grouping)	
Description:	**Target:**
Aldehydes (Acute toxicity) (US-EPA New Chemical Categories)	Aldehydes (Acute toxicity)
Substance type (subcategorization)	
Description:	**Target:**
Differ from target by: At least one category	Discrete chemical
Removed categories: Has impurities	Mono constituent (predefined)
Dissociating chemical	Organic
Acute aquatic toxicity MOA by OASIS (subcategorization)	
Description:	**Target:**
Differ from target by: At least one category	Aldehydes
Removed categories: Reactive unspecified	
US-EPA New Chemical Categories (subcategorization)	
Description:	**Target:**
Differ from target by: At least one category	Aldehydes (Acute toxicity)
Removed categories: Phenols (Acute toxicity)	
Polynitroaromatics (Acute toxicity)	
Aquatic toxicity classification by ECOSAR (subcategorization)	
Description:	**Target:**
Differ from target by: At least one category	Aldehydes (Mono)
Removed categories: Vinyl/Allyl Aldehydes	
Aldehydes (Poly)	
Thiophenes	
Organic functional groups, Norbert Haider (checkmol) (subcategorization)	
Description:	**Target:**
Differ from target by: At least one category	Aldehyde

Appendix: Data pruning

Data pruning justification: *manually editable field*
Not provided by the user

1. Filter by data inconsistency

Native unit: ppm

Data value	CAS: 111-30-8	Reason to remove
5.4 ppm	SMILES: O=CCCCC=O	Native unit: ppm
11.6 ppm	SMILES: O=CCCCC=O	Native unit: ppm

3. Filter by water solubility

Water Solubility

Data value	CAS: 69770-23-6	Reason to remove
0.368 mg/L	SMILES: CC(C)(C)c1ccc(Oc2cccc(C=O)c2)cc1	Water Solubility = 0.318 mg/L

図 14

6.1.2. 混合物の予測

混合物のモデルから得られた、予測のための特別なレポートテンプレートがあります。
SMOA（IMOA）予測をしたら、それをクリックして（1）、**Prediction** （2）をクリックします（図 15）。Wizard ページが表示されます。左側のツリー（3）を展開して、レポート内のセクションの外観をカスタマイズしたり、レポートを直接作成すること（Create report）ができます（4）。

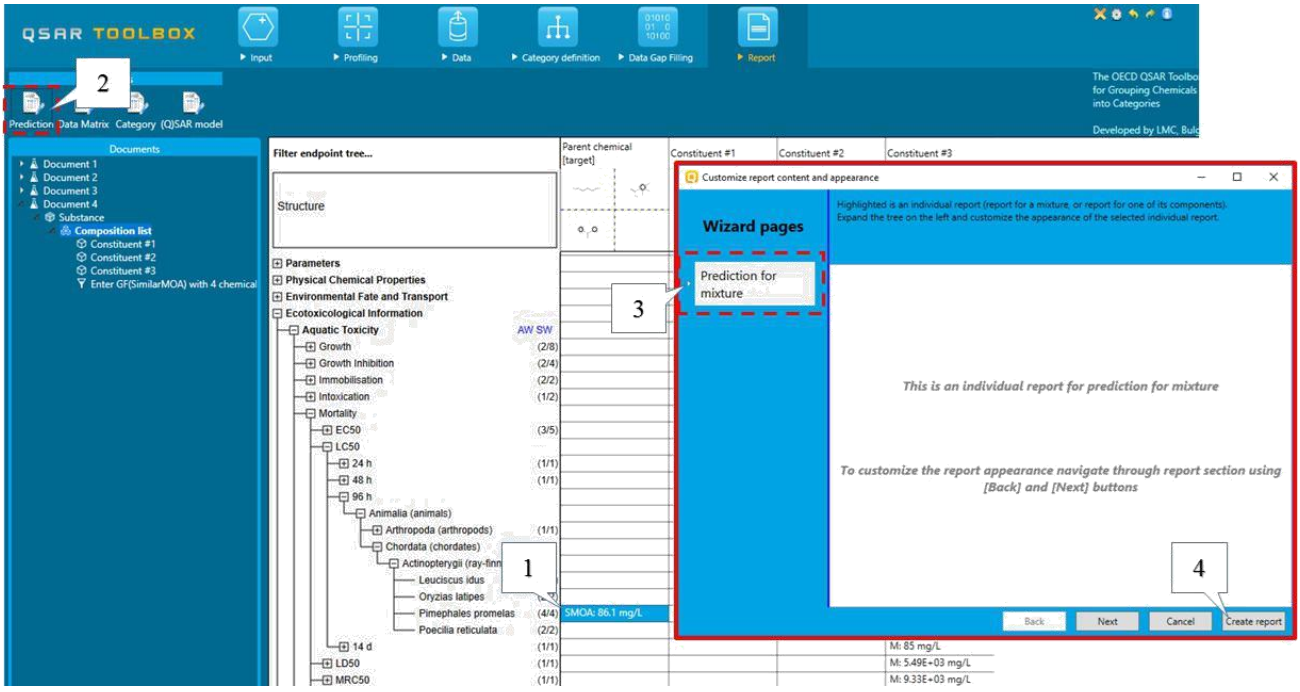


図 15

混合物についてのレポート　テンプレート（図 16-1）の部分は、単一の化学物質の部分とだいたい同じです。ここで違うのは、混合物の成分が、収集された類似物質の代わりに予測に使用されることです。このレポートでは、対応するチェックボックスのチェック（2）を外してセクションの一部を除外するか、または Move up/Move down ボタン（3）でセクションを再配置することによって、レポートをカスタマイズできます。

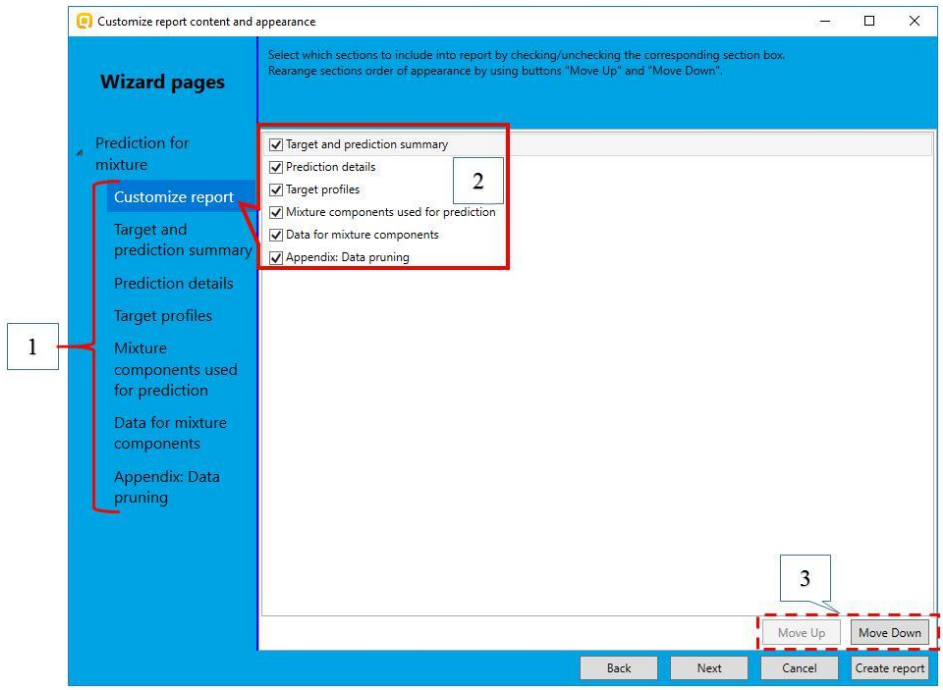


図 16

6.1.3. レポートの作成

Create report (1)をクリックしてレポートを生成 します。2 つのオプションがついた追加のウィンドウが表示されます。**Prediction report**（予測レポート） と **Data matrix** (2)です。ファイル（3）を開いたり保存したりすることができます（図 17）。

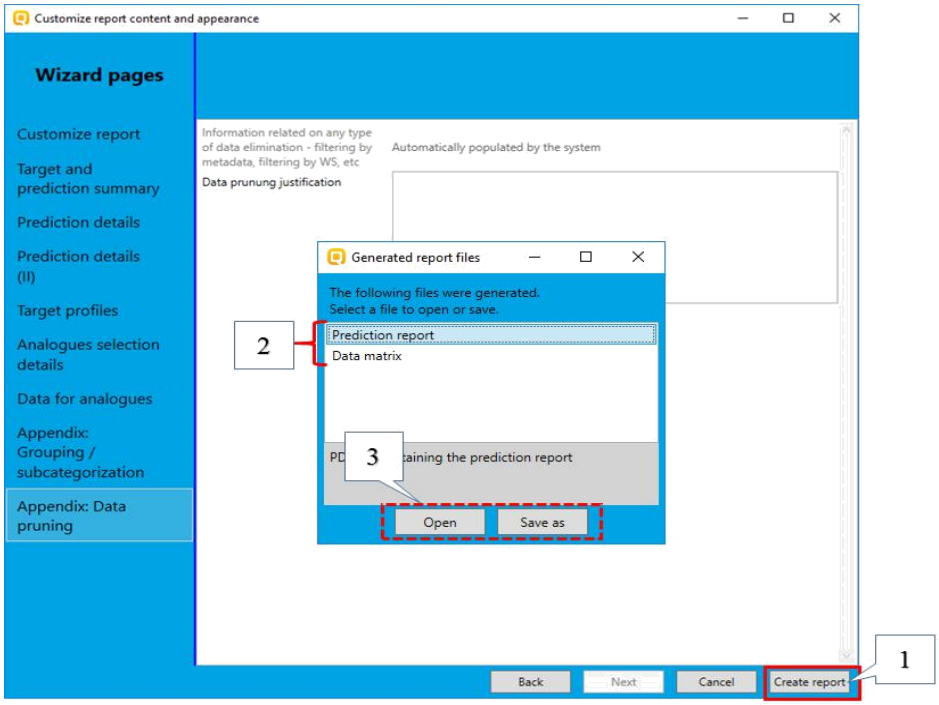
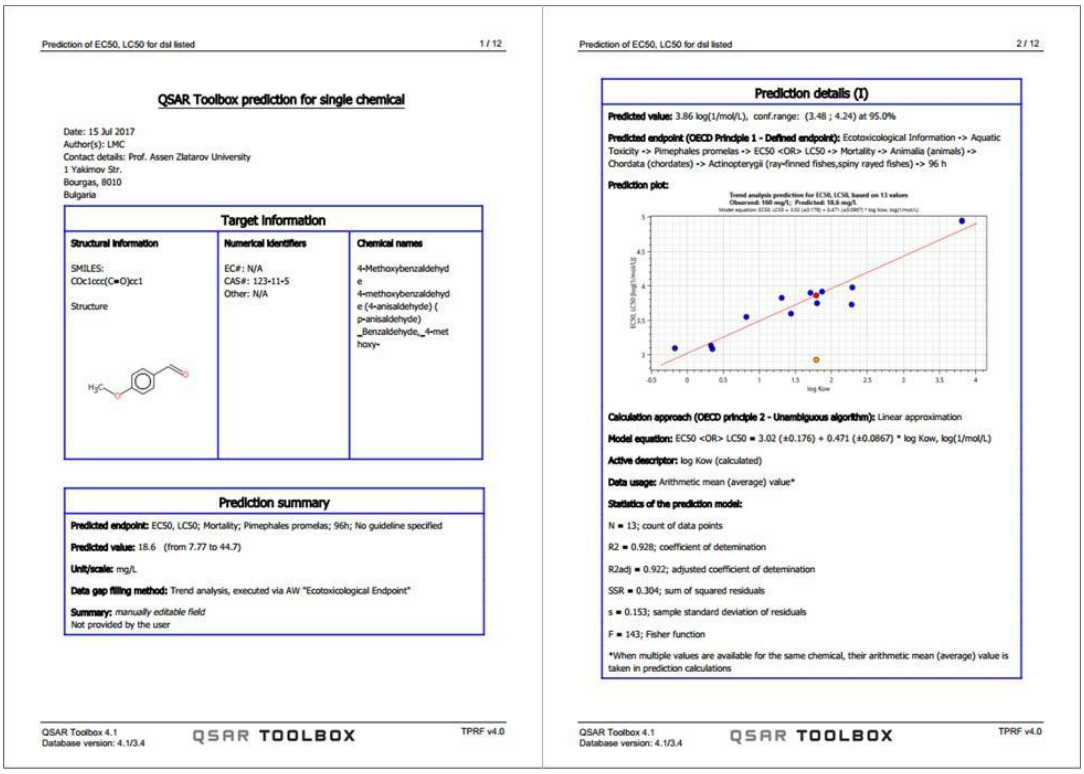


図 17

Prediction report（予測レポート）は、標的に関連する予測情報を含む PDF ファイルです。

- 単一の化学物質の予測（図 18）



Prediction of EC50, LC50 for dsl listed 1 / 12

QSAR Toolbox prediction for single chemical

Date: 15 Jul 2017
Author(s): LMC
Contact details: Prof. Assen Zlatarov University
1 Yakimov Str.
Bourgas, 8010
Bulgaria

Target information

Structural information	Numerical identifiers	Chemical names
SMILES: COc1ccc(C=O)cc1 Structure	EC#: N/A CAS#: 123-11-5 Other: N/A	4-Methoxybenzaldehyde 4-methoxybenzaldehyde (4-anisaldehyde) (p-anisaldehyde) _Benzaldehyde,_4-methoxy-

Prediction summary

Predicted endpoint: EC50, LC50; Mortality; Pimephales promelas; 96h; No guideline specified

Predicted value: 18.6 (from 7.77 to 44.7)

Unit/scale: mg/L

Data gap filling method: Trend analysis, executed via AW "Ecotoxicological Endpoint"

Summary: *manually editable field*
Not provided by the user

QSAR Toolbox 4.1
Database version: 4.1/3.4
QSAR TOOLBOX
TPRF v4.0

Prediction of EC50, LC50 for dsl listed 2 / 12

Prediction details (I)

Predicted value: 3.86 log(1/mol/L), conf.range: (3.48 ; 4.24) at 95.0%

Predicted endpoint (OECD Principle 1 - Defined endpoint): Ecotoxicological Information -> Aquatic Toxicity -> Pimephales promelas -> EC50 <OR> LC50 -> Mortality -> Animalia (animals) -> Chordata (chordates) -> Actinopterygii (ray-finned fishes,spiny rayed fishes) -> 96 h

Prediction plot:

Trend analysis prediction for EC50, LC50, based on 13 values
Observed: 160 mg/L; Predicted: 18.6 mg/L

EC50, LC50 [log(1/mol/L)]
log Kow

Calculation approach (OECD principle 2 - Unambiguous algorithm): Linear approximation

Model equation: EC50 <OR> LC50 = 3.02 (±0.176) + 0.471 (±0.0867) * log Kow, log(1/mol/L)

Active descriptor: log Kow (calculated)

Data usage: Arithmetic mean (average) value*

Statistics of the prediction model:

N = 13; count of data points

R2 = 0.928; coefficient of determination

R2adj = 0.922; adjusted coefficient of determination

SSR = 0.304; sum of squared residuals

s = 0.153; sample standard deviation of residuals

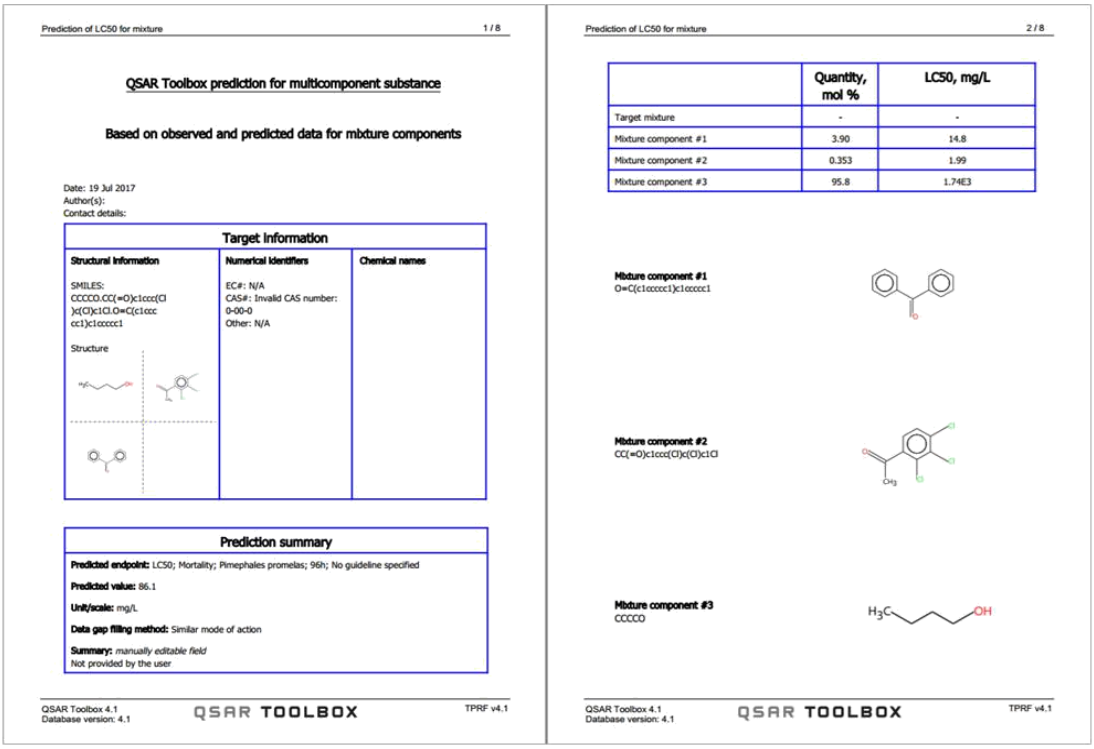
F = 143; Fisher function

*When multiple values are available for the same chemical, their arithmetic mean (average) value is taken in prediction calculations

QSAR Toolbox 4.1
Database version: 4.1/3.4
QSAR TOOLBOX
TPRF v4.0

図 18

- 混合物の化学物質の予測（図 19）



Prediction of LC50 for mixture 1 / 8

QSAR Toolbox prediction for multicomponent substance

Based on observed and predicted data for mixture components

Date: 19 Jul 2017
Author(s):
Contact details:

Target information

Structural information	Numerical identifiers	Chemical names
SMILES: CCCCO.CC(=O)c1ccc(Cl)c(Cl)c1Cl.O=C(c1ccccc1)c1ccccc1 Structure	EC#: N/A CAS#: Invalid CAS number: 0-00-0 Other: N/A	

Prediction summary

Predicted endpoint: LC50; Mortality; Pimephales promelas; 96h; No guideline specified

Predicted value: 86.1

Unit/scale: mg/L

Data gap filling method: Similar mode of action

Summary: *manually editable field*
Not provided by the user

QSAR Toolbox 4.1
Database version: 4.1
QSAR TOOLBOX
TPRF v4.1

Prediction of LC50 for mixture 2 / 8

	Quantity, mol %	LC50, mg/L
Target mixture	-	-
Mixture component #1	3.90	14.8
Mixture component #2	0.353	1.99
Mixture component #3	95.8	1.74E3

Mixture component #1
O=C(c1ccccc1)c1ccccc1

Mixture component #2
CC(=O)c1ccc(Cl)c(Cl)c1Cl

Mixture component #3
CCCCO

QSAR Toolbox 4.1
Database version: 4.1
QSAR TOOLBOX
TPRF v4.1

図 19

Data matrix（訳注：図 17-2 下段）で出力されるものは、選択されたパラメータ、プロファイル、およびエンドポイントツリーに関するデータとともに予測に使用される化学物質の入った MS Excel ファイルです。

- 単一の化学物質のデータ・マトリックス（図 20）

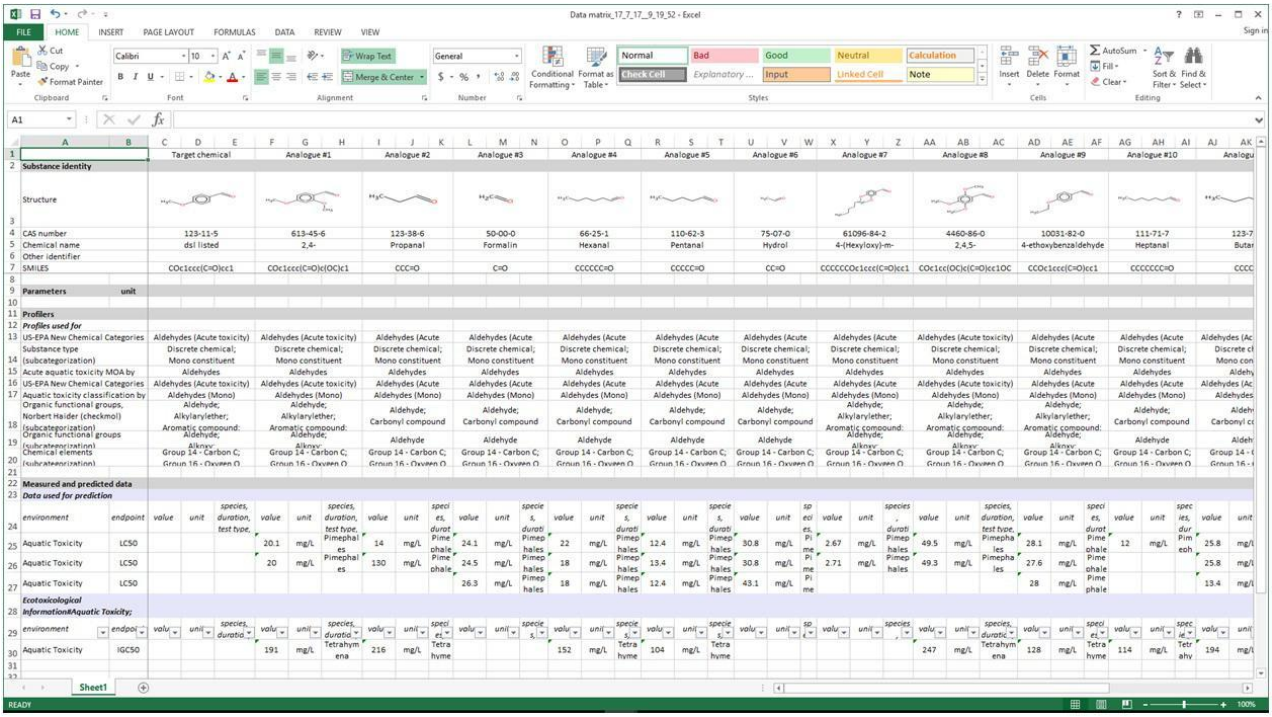


図 20

- 混合物の化学物質のデータ・マトリックス（図 21）

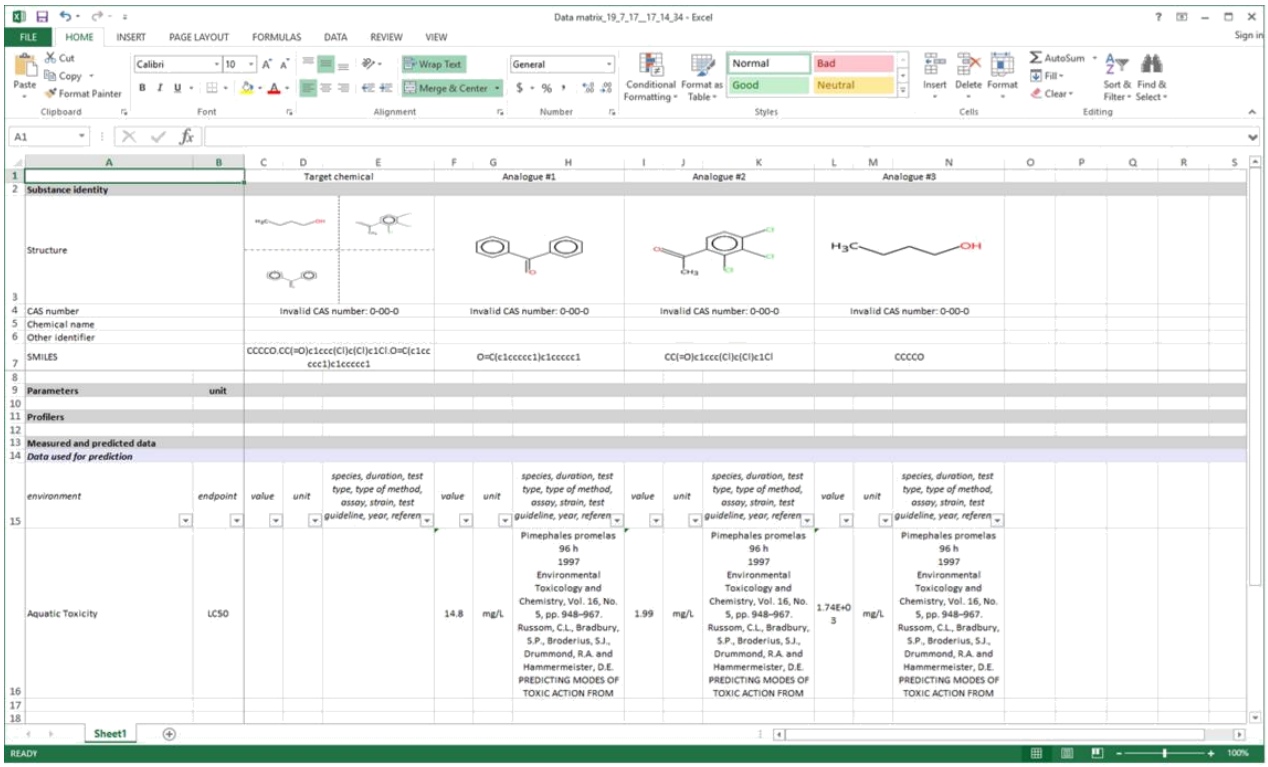


図 21

D.6.2. データ・マトリックス　レポート

データ・マトリックスレポートでは、予測レポート（D.6.1.6　Data for analogues を参照）（訳注：現在この章は無い）と同様の Excel ファイルを生成することができます。ただし、この場合、利用可能な予測は必要ありません。

このタイプのレポートを作成するには、レポートモジュールで **Data matrix** （1）をクリックし、表示する化学物質データ（パラメータおよび/またはプロファイル結果および/または実験データ）を選択します（図 1）。プロファイルを作成したり、パラメータを計算したり、データを集める必要はありません。カテゴリを作成するためにプロファイリングスキームが使用されている場合、それは自動的に選択されます（2）。最後に Create report（3）をクリックしてファイルを生成します。

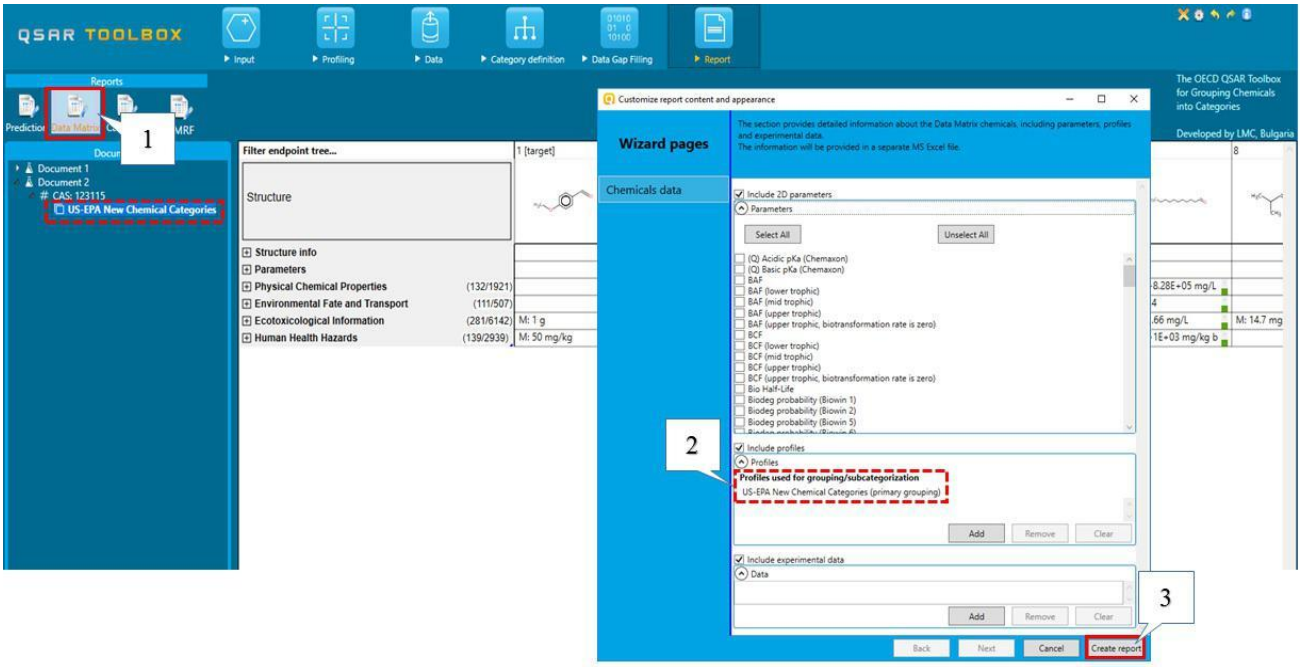


図 1

D.6.3. カテゴリ　レポート

Category の 機能は、現在の化学物質リストのレポートを作成します。Category report（1）は、ユーザがカテゴリを検証できるようにするため開発されました（図 1）。Category report のウィザードページでは、レポートの内容を次のようにカスタマイズすることができます（2）。手動で情報（カテゴリ仮説、カテゴリのカバーするエンドポイント、化学物質の純度や既知の不純物）（3）を含めたり、 Toolbox が提供する任意の情報（プロファイリング結果、2D（3D）パラメータ、実験データ）を含めたりできます。

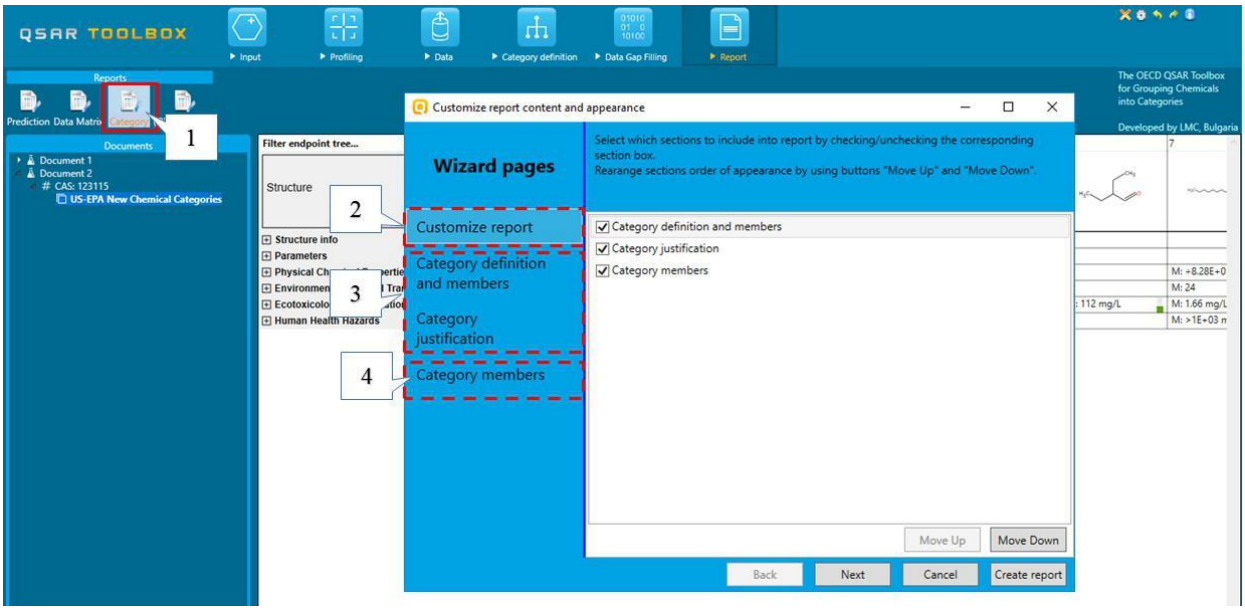


図 1

レポートをカスタマイズしたら、**Create report**（1）をクリックしてファイルを生成します（図 2）。Category report は Chemicals category report （2）と Data matrix（3）の 2 つの部分で構成されています。これらは開くか保存することができます（4）。Chemicals category report は、化学物質カテゴリの定義と妥当性を含む PDF レポートです。

Data matrix は、選択されたパラメータ、プロファイル、およびエンドポイントツリーに関するデータとともに予測に使用される化学物質の入った MS Excel ファイルです。

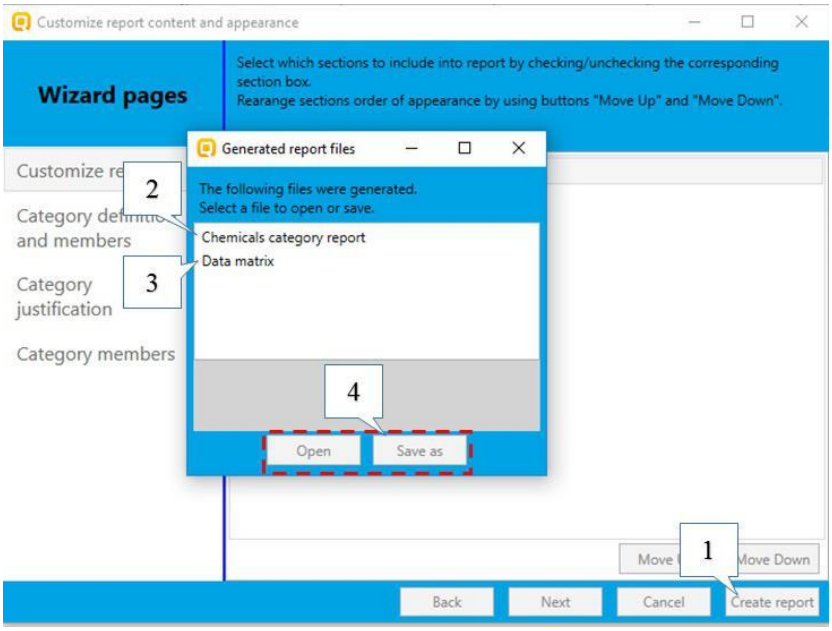


図 2

作成されたカテゴリレポートのタイトルは、ドキュメントツリー内のカテゴリの名前に対応します（図 3）。したがって、最初にカテゴリの名前を変更してからレポートを生成すると、タイトルを変更できます。

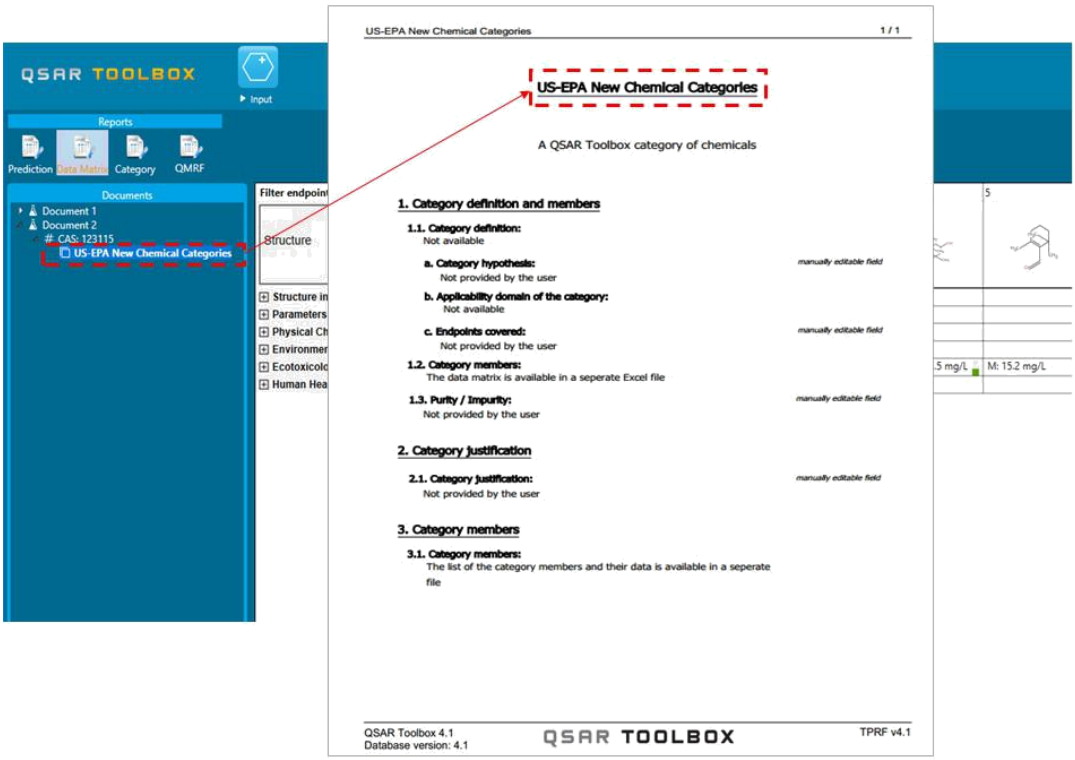


図 3

D.6.4.QMRF レポート

QMRF オプションは、選択した（Q）SAR モデルの QMRF レポートを作成します。
特定のモデルの QMRF を生成するには、(Q) SAR モデルの対応する行をクリックする必要があります。ただし、正確な行がわからない場合は、エンドポイントツリーのより広範囲な上位部分を選択することができます。例えば Human Health Hazard (1)です（図 1)。行が選択されたら、QMRF（2）をクリックします。選択した行に関連する、すべての利用可能な（Q）SAR モデルがリストされます。関心のあるモデルを見つけ、（例： Photoinduced toxicity based on 3T3 NRU data (3) ）OK をクリックします。

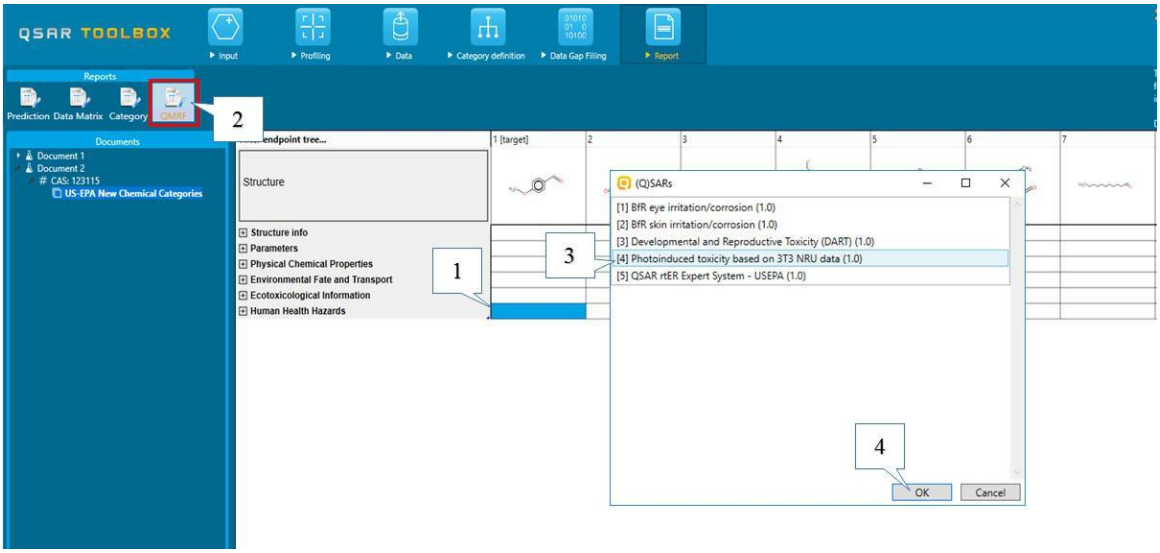


図 1

Wizard ページが表示されます。モデルのトレーニングセットとテストセットデータのチェックボックスをオフにすることだけが可能です（図 2)。レポートの残りのセクションはシステムによって自動的に入力され（2)、各セクションの内容のみが表示されます。トレーニングセットとデータセットセクション（4）は、現在のモデルのトレーニングおよびデータセットのすべての化学物質について、2D / 3D パラメータ、プロファイリング結果および/または実験データ（5）を含めることができます。これらのセットは別の Excel ファイルとして生成されます。Create report をクリックしてファイルを生成します（6)。

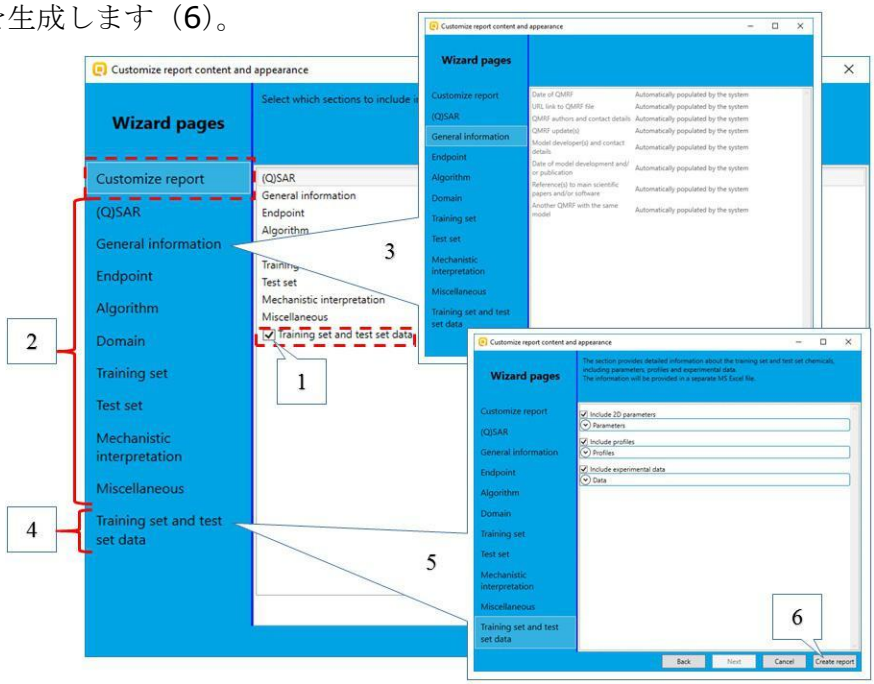


図 2

3 つのファイルが生成され（1)、開いたり保存したりすることができます（2）（図 3)。

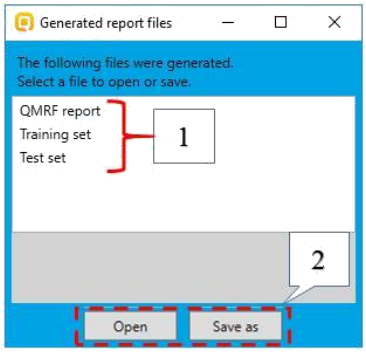


図 3

1 ） **QMRF レポート-** 選択された（Q）SAR の QMRF レポートを含む PDF ファイル（図 4)。
2 ） **トレーニングセット -** 選択されたパラメータ、プロファイル、エンドポイントツリーの位置 (1) に関する QSAR のトレーニングセット化学物質を含む、MS Excel ファイル。（図 5）
3 ） **テストセット -** QSAR のテストセット化学物質と、選択されたパラメータ、プロファイル、エンドポイントツリー位置（2）に関するデータを含む MS Excel ファイル。（図 5）

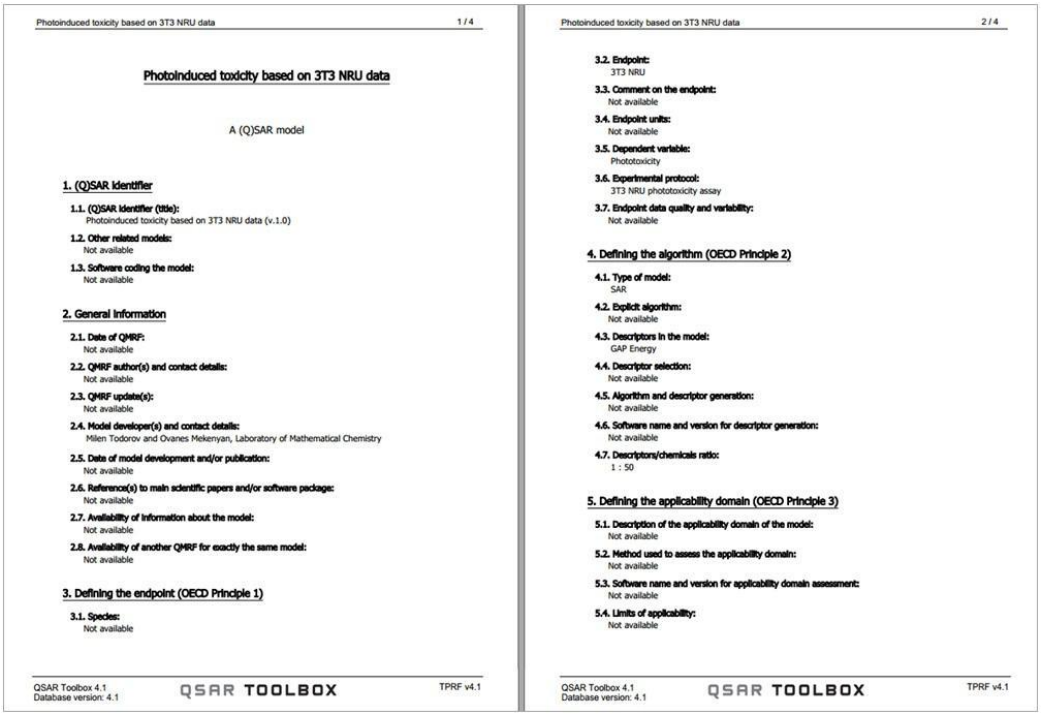


図 4

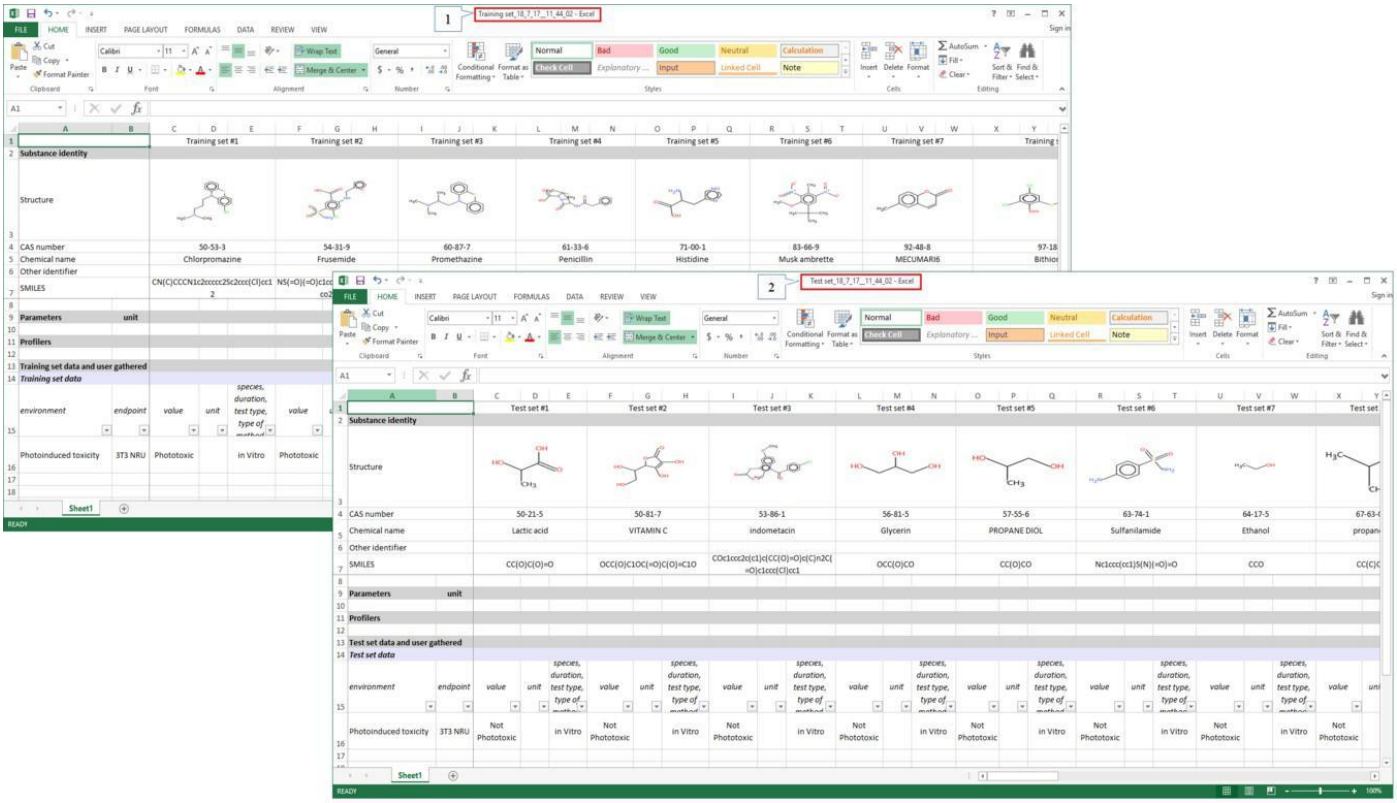

1
Training set_18_7_17_11_44_02 - Excel
Training set #1
Training set #2
Training set #3
Training set #4
Training set #5
Training set #6
Training set #7
Substance identity
Structure
CAS number
50-53-3
54-31-9
60-87-7
61-33-6
71-00-1
83-66-9
92-48-8
Chemical name
Chlorpromazine
Frusemide
Promethazine
Penicillin
Histidine
Musk ambrette
MECUMARI6
Other identifier
SMILES
Parameters
unit
Profilers
Training set data and user gathered
Training set data
environment
endpoint
value
unit
species, duration, test type, type of
Photoinduced toxicity
3T3 NRU
Phototoxic
in Vitro
2
Test set_18_7_17_11_44_02 - Excel
Test set #1
Test set #2
Test set #3
Test set #4
Test set #5
Test set #6
Test set #7
50-21-5
50-81-7
53-86-1
56-81-5
57-55-6
63-74-1
64-17-5
Lactic acid
VITAMIN C
indometacin
Glycerin
PROPANE DIOL
Sulfanilamide
Ethanol
CC(O)C(O)=O
OCC(O)C1OC(=O)C(O)=C1O
COc1ccc2c(c1)c(CC(O)=O)c(C)n2C(=O)c1ccc(Cl)cc1
OCC(O)CO
CC(O)CO
Nc1ccc(cc1)S(N)(=O)=O
CCO
Not Phototoxic
Sheet1
READY

図 5

E 章 . Options（オプション）

E 章 . オプション

Toolbox の一般オプションは、以下のプログラムの設定を変更することができます。

1. General
2. Modules
3. Profiling
4. Data Gap Filling
5. Interface
6. Data

Options ボタン（図 1）は、ツールボックスのメインウィンドウの右上隅に配置されています。

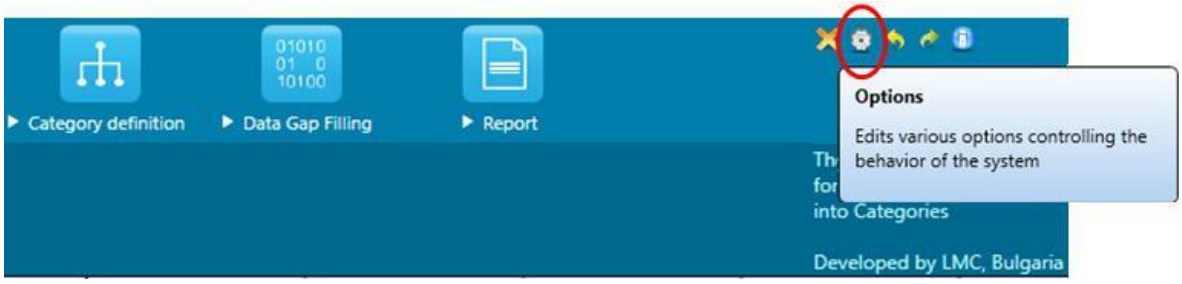


図 1

次のウィンドウ（図 2）が開き、すべての機能がウィンドウの左側に表示されます（1）。設定を変更した場合は、Apply を選択してから OK（2）を選択してください。Set Default（デフォルト設定に戻す）オプションもあります。

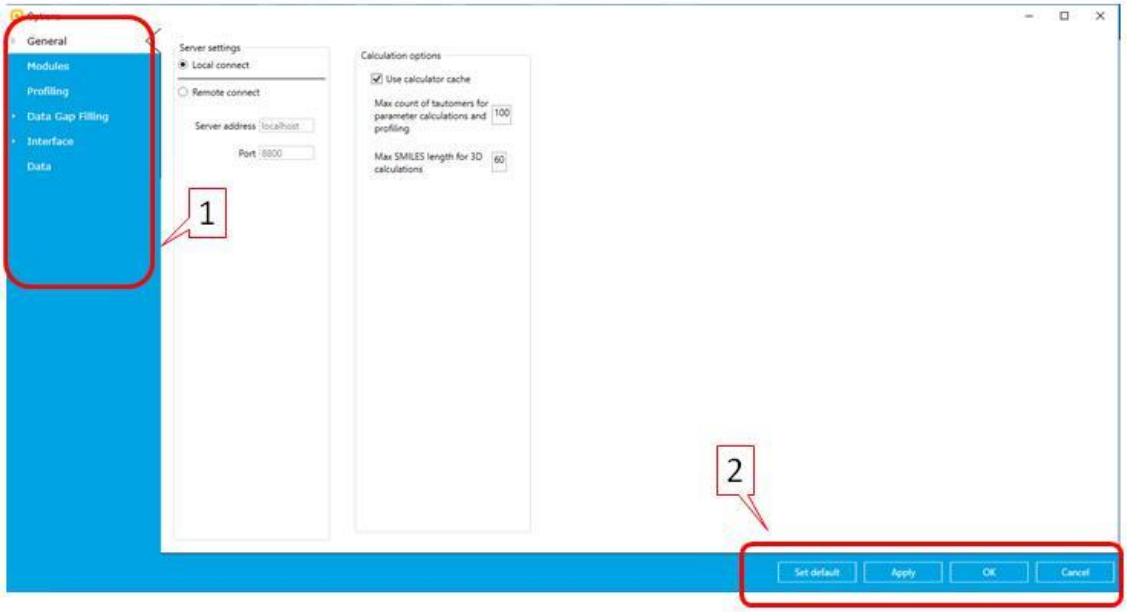


図 2

E.1. General（一般）

General（一般）設定では、サーバ設定（ローカルまたはリモート）(1）と計算設定（2）（図 1）を変更できます。

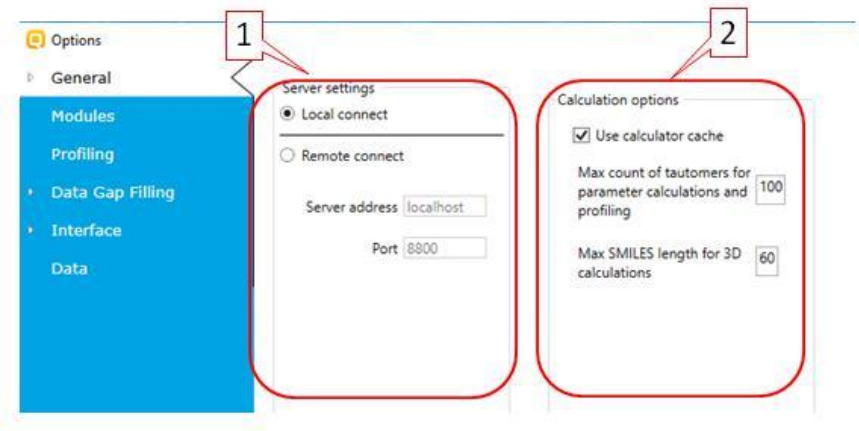


図 1

（ 1 ）サーバの設定：

1. ローカル接続オプションをオンにすると、ツールボックスはツールボックスがインストールされている同じ PC に保存されているデータベース（FDB ファイル）を使用します。データベースのインストールは、Toolbox のインストール中に実行されます。ここでは、作業 FDB ファイルを変更するためのオプションが用意されています。
2. Remote connect（リモート接続）オプションが選択されている場合には、ツールボックスには、データベースが格納されて使用でき、サーバに接続して使用できます。
3. サーバのアドレスと Port（ポート番号）を入力します。

(2) Calculation options:（計算オプション）

1. Use calculator cache オプションを使用すると、事前に計算されてツールボックスキャッシュに保存された値を使用/不使用して、2D および 3D パラメータの計算を実行できます。このオプションをオンにすると、計算プロセスが非常に遅くなることがあるため、ユーザは注意が必要です。
2. Max count of tautomers...（互変異性体の最大数…）オプションは、プロファイリングおよび 2D・3D 計算の対象となる特定の構造に対して、生成される互変異性体の数を明らかにします。計算およびプロファイリングの対象となる、互変異性形態の数のしきい値を変更することができます。
3. Max SMILES length for 3D calculations（3D 計算の SMILES の最大長さ）は、そのしきい値を変更することができます。

•Scale editor（スケールエディタ）

Scale editor（2）は、General（1）オプション（図 2）を拡張すると表示されます。Scale editor パネル（3）では、ツールボックス 4.0 に入っているすべてのスケールを表示します。

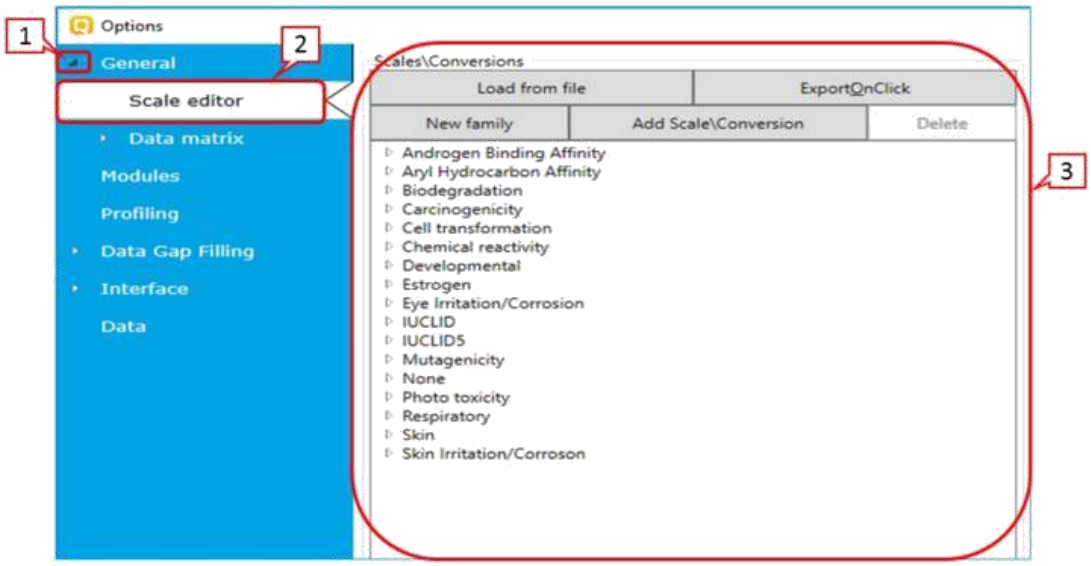


図 2

図 3 にあるとおり、New family スケール（1,2）を追加するオプションがあります。

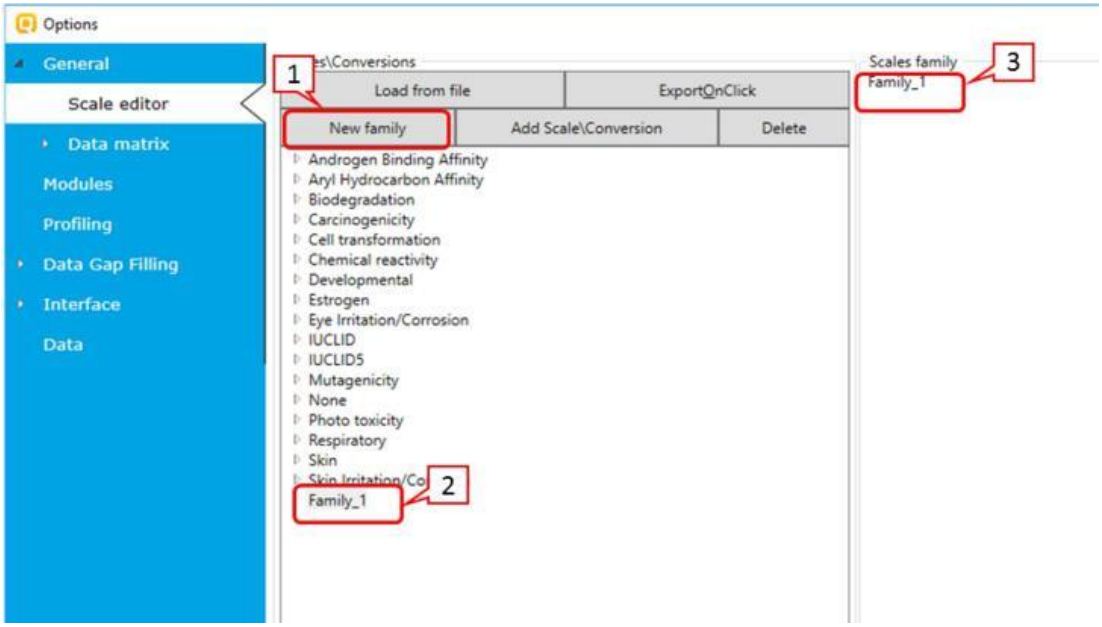


図 3

次に、図 4 ではファミリ（1,2）に新しいスケールを追加できます。

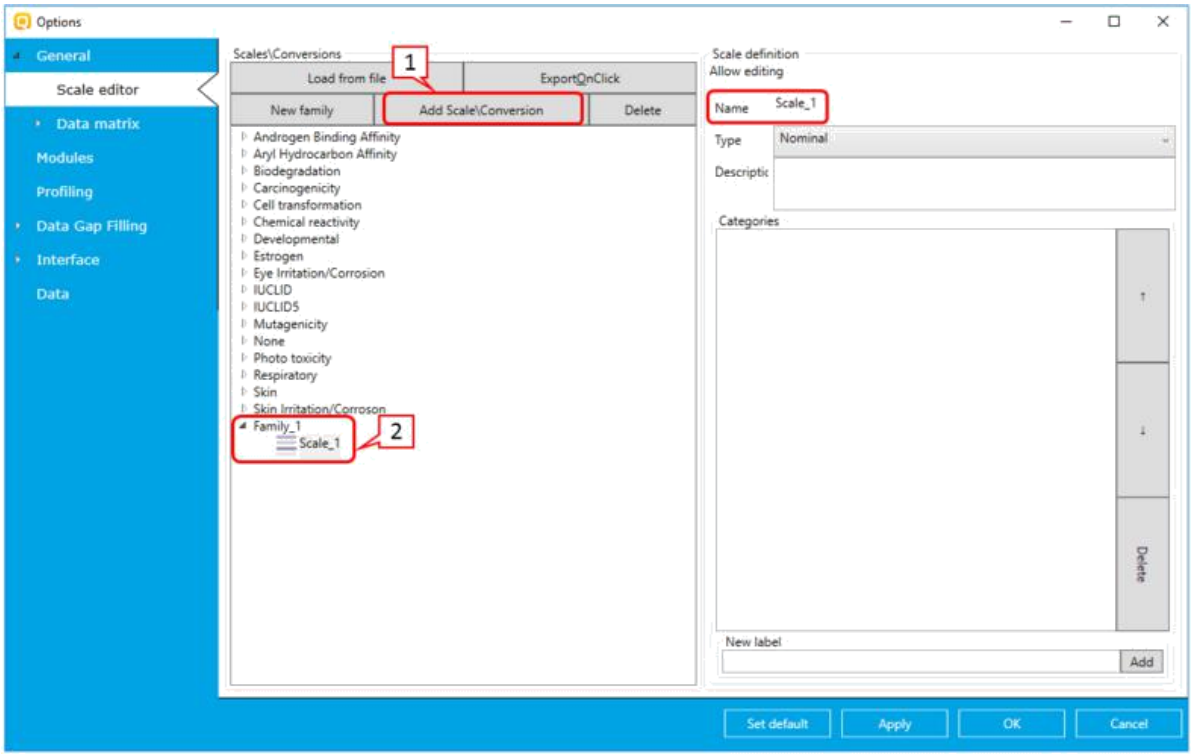


図 4

既存のスケールのファミリを拡張表示すると、このファミリに属するすべてのスケールとその変換も（もしあれば）表示されています（1）。（図 5）また、ユーザはスケールを Load （読み込み）（2）したり、スケールを Export（エクスポート）（3）することができます。

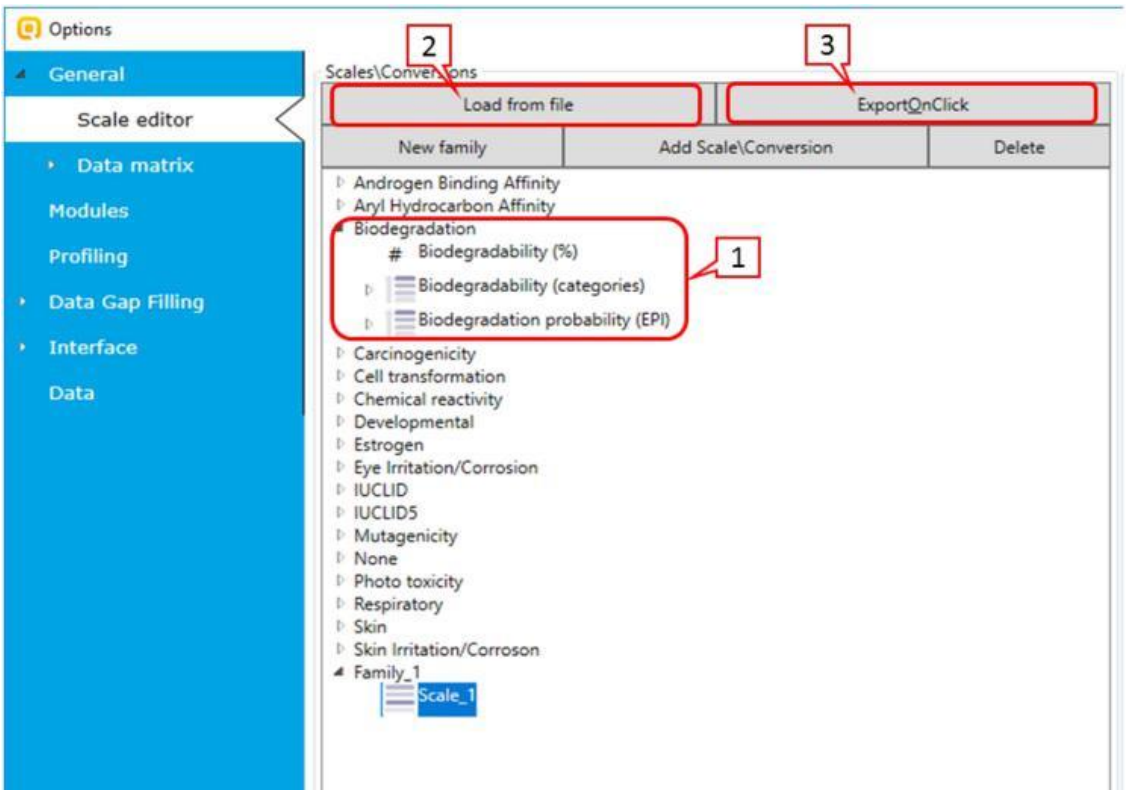


図 5

Data matrix　データ・マトリックス

Data matrix（データ・マトリックス）オプション（1）は、一般的なオプションのサブレベルにあります。メインの右パネルは、マトリックス（2）の寸法パラメータを与えます。（図 6）

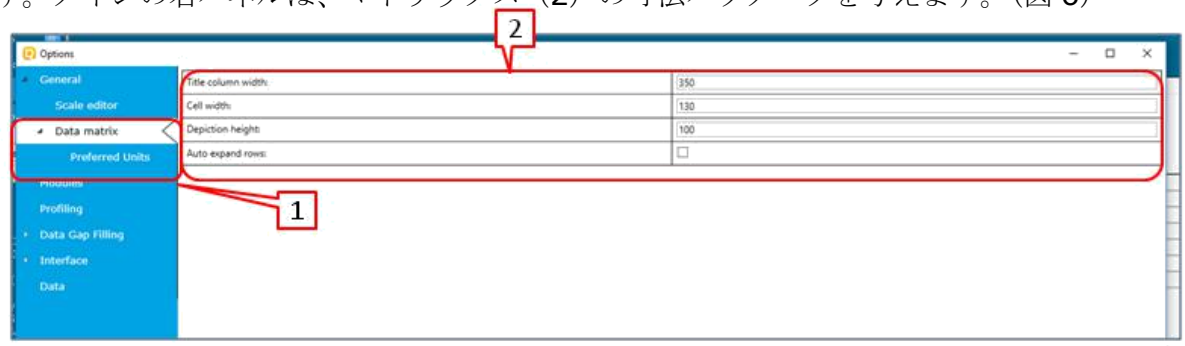


図 6

データマトリックスオプションの下位レベルに **Preferred units**　（優先される単位）（1）があります（図 7）。

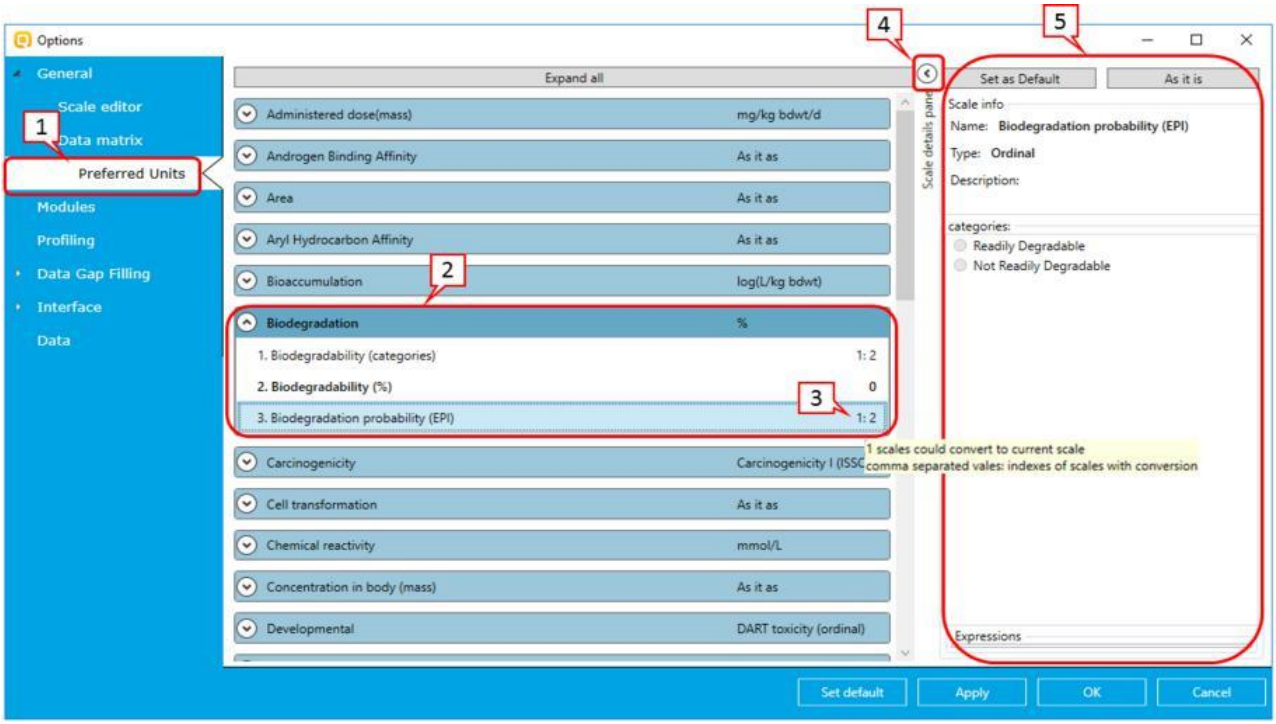


図 7

ここでは、Biodegradation（生分解）の単位を例として説明します。

- 使用可能なすべてのスケールを表示するには、名前の隣にある文字列を展開します（2）。
- 太字のスケールである生分解性（%）がデフォルトのスケールです。
- 最後の数字（3）はスケール変換を示しています。例えば、生分解確率（EPI）スケールは、生分解性%であるスケール番号 2 に変換されます。
- Scale details パネル（５）は、矢印（4）をクリックすると開くことができます。パネル（5）で、異なる単位および表現の選択ができます。

E.2.Modules（モジュール）

Modules パネルのツールボックスにすべての利用可能なモジュール、計算機能、データベース、在庫、プロファイラ、(Q)SAR モデル、代謝などのリストが含まれます（図 1）。

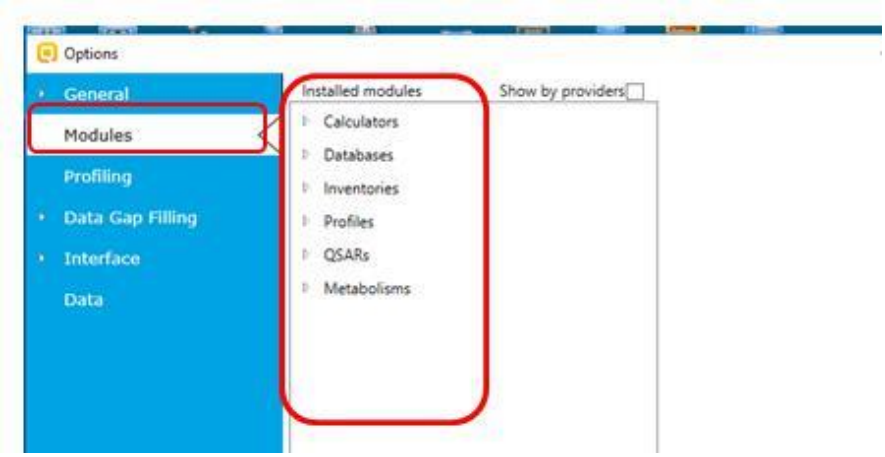


図 1

拡張の三角形アイコンをダブルクリックして、モジュールの内容を表示します。たとえば、calculators をダブルクリックすると、Toolbox（図 2）で使用可能なすべての 2D および 3D 計算機能が 2 つのリストに表示されます。

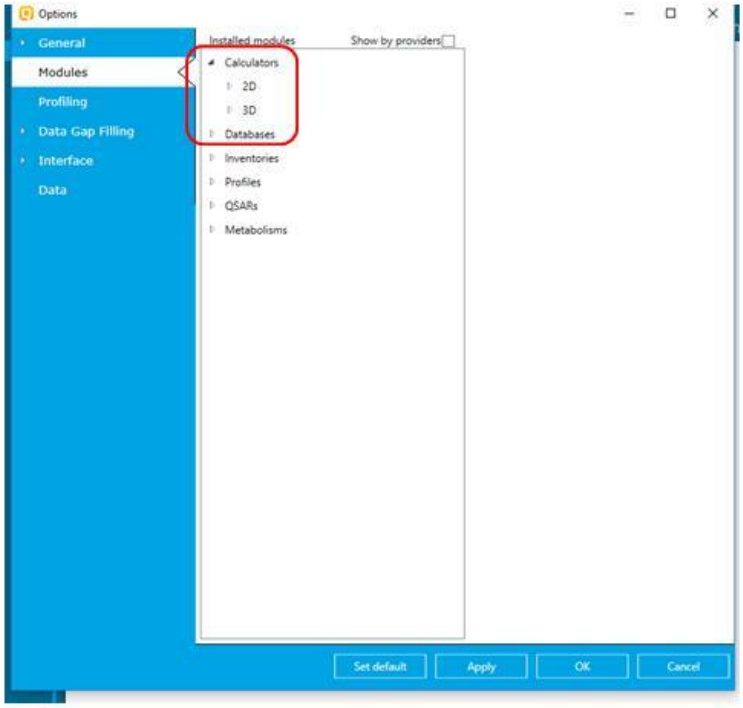


図 2

次の図（図 3）は、ノードになっている Calculators（計算機能）をより詳細に表しています。

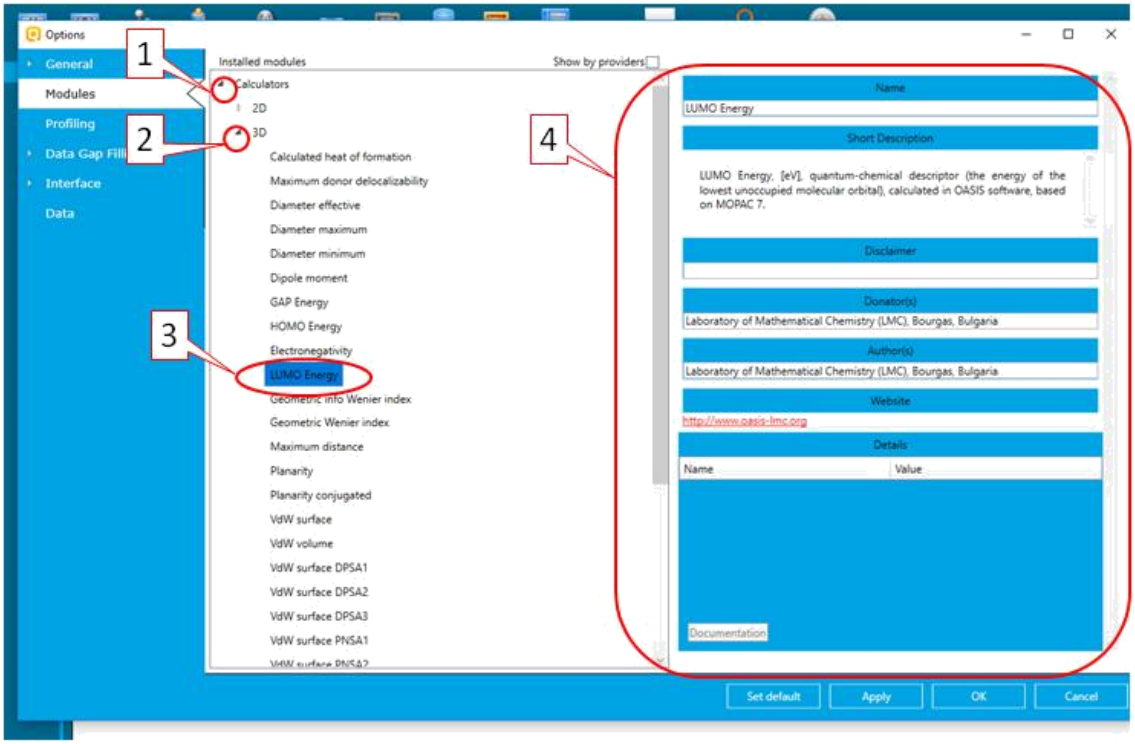


図 3

1. Calculators ノード（1）を展開し、3D ノード（2）を展開します。

2. 例えば LUMO Energy（3）などの各計算機でクリックすると、簡単な説明、著者、貢献者など、選択された計算機能（モジュール）に関するより詳細な情報を提供します。

E.3.Profiling（プロファイリング）

Profiling モジュールに関連する一般的なオプションを図 1 に示します。

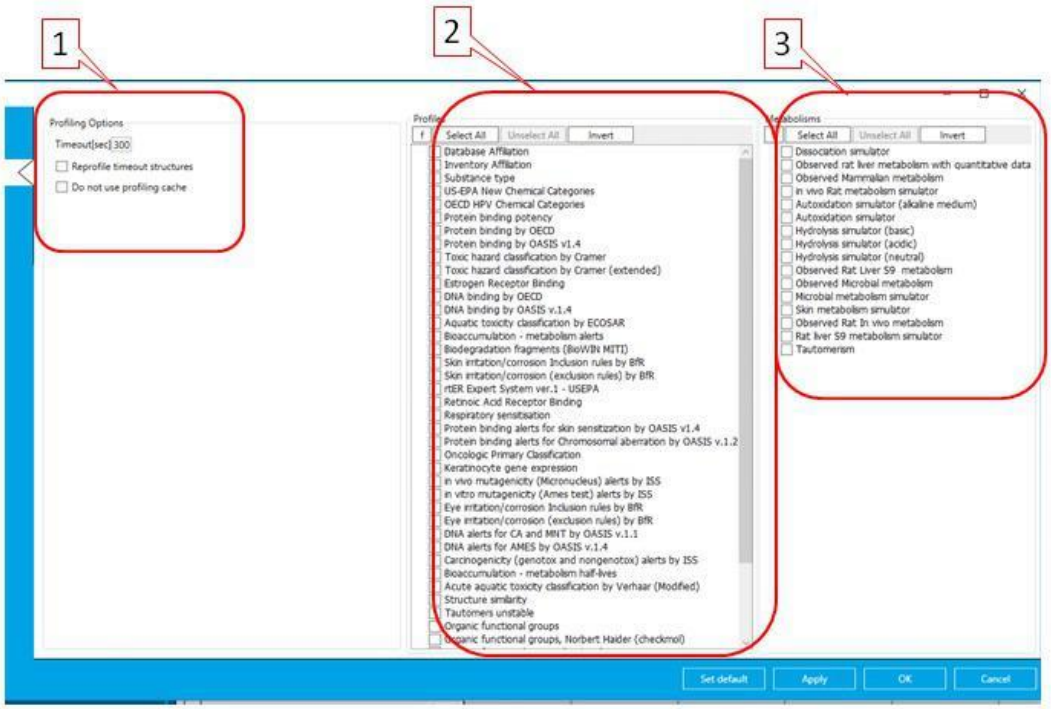


図 1

(1) - Timeout- ここでは、ユーザの増加は、化学構造のプロファイリングをするためのデフォルトの最大時間を増減できます。

注意: 大きくて複雑な構造のプロファイリングには時間がかかることがあるため、タイムアウトオプションが付けられました。そういう種類の構造をプロファイリングする時間が、設定された時間を超える場合、これらの構造体のプロファイリング結果はタイムアウトとなります。

Reprofile timeout structures - このオプションは上記に関連しています。Timeout のしきい値を上げると、タイムアウト構造のプロファイリングが可能になります。

Do not use profiling cache － プロファイリング計算に関連します。ここで、プロファイリングキャッシュを使用しないことを選択できます。大量の化学物質のプロファイリングのプロセスは非常に遅くなるため、キャッシュを使用しないよう選択するときは注意が必要です。(作動時間は) 格納されたプロファイリング結果のキャッシュに依存するため、カテゴリ化プロセスにも影響します。

注意: このオプションの選択は、新しく作成されたプロファイリングスキームの結果をチェックしたり、既存のスキームの変更をチェックしたりする場合に推奨されます。

(2) - ソフトウェアが起動するたびに選択されるプロファイラをユーザが選択できるようにする。

(3) - ソフトウェアが起動するたびに選択される代謝シミュレータをユーザが選択できるようにする。

E.4.Data Gap Filling（データギャップ補完）

データギャップ補完オプションには、4 つの主なサブレベル（図 1-1）が含まれています : 優先ユニット、計算機能、エンドポイントツリー、記述子です。

図 1

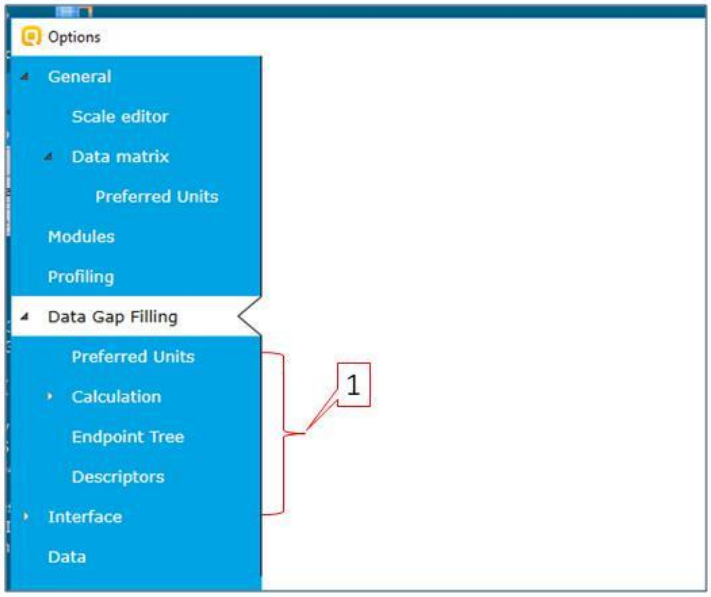


図 2 の Preferred Units (1)

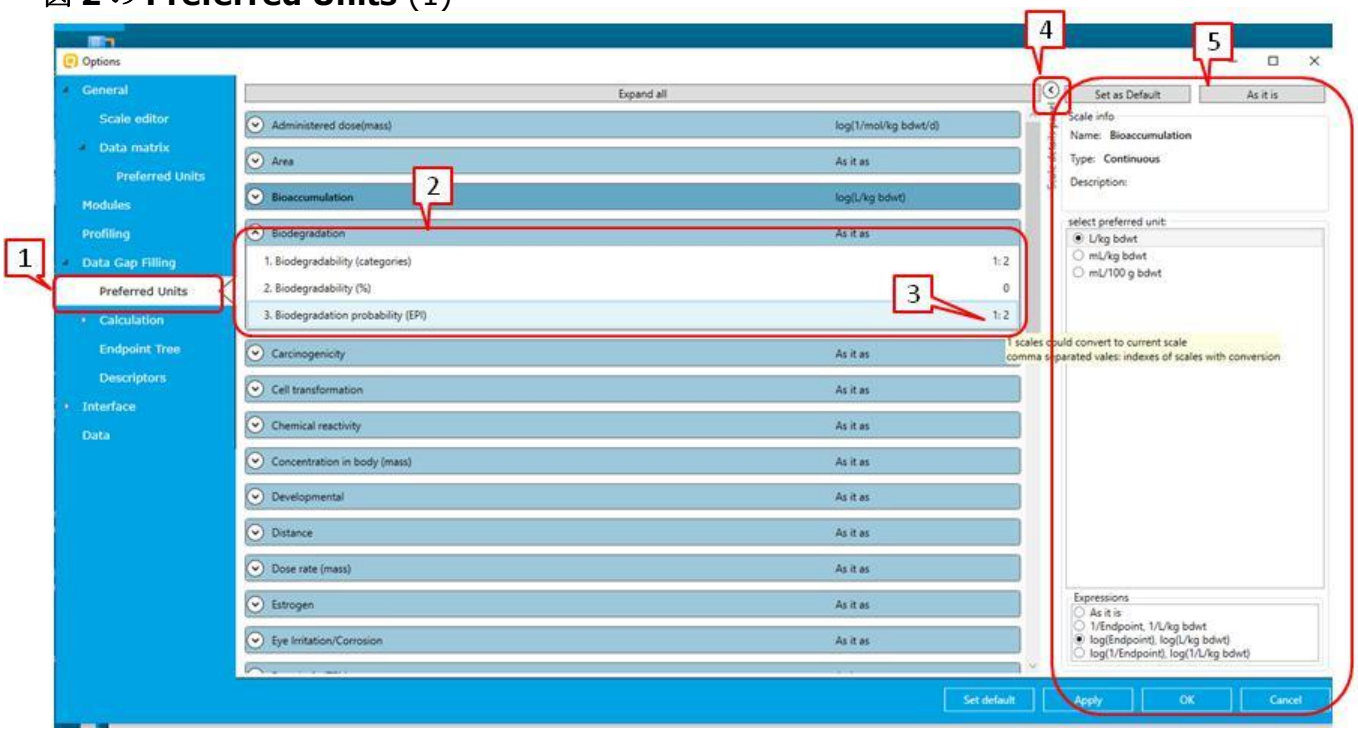


図 2

生分解単位が例として示されています（図 2）。

- 名前の横にある三角形を展開して、使用可能なすべてのスケールを表示します（2）。
- 太字のスケールである生分解性（%）がデフォルトのスケールです。

- 最後の数字はスケール変換を示しています。例えば、生分解確率スケールは、生分解性%であるスケール番号 2 に変換されます。

- Scale details パネル（５）は、クリック(4)すると開くことができます。このパネルにより、ユーザは様々なユニットと表現を選択することができます。

- **Calculation（計算機能）**

Calculation セクション（1）では、データギャップ補完方法（2）ごとに様々な計算オプションを選択できます（図 3）。

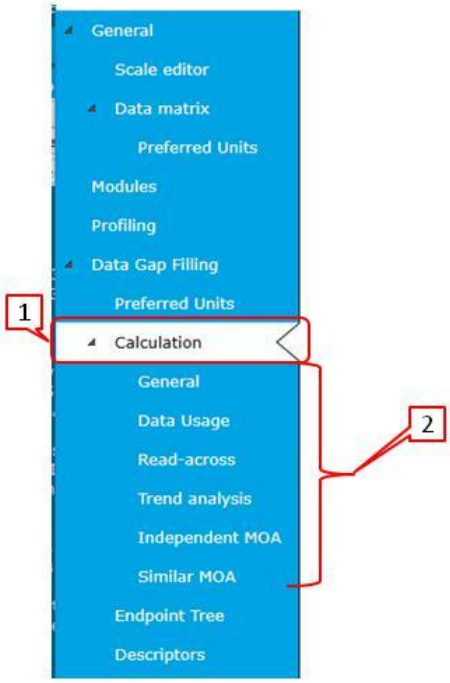


図 3

- General（１）：すべての方法に関連する一般的なオプション（2）を含みます。（図 4）

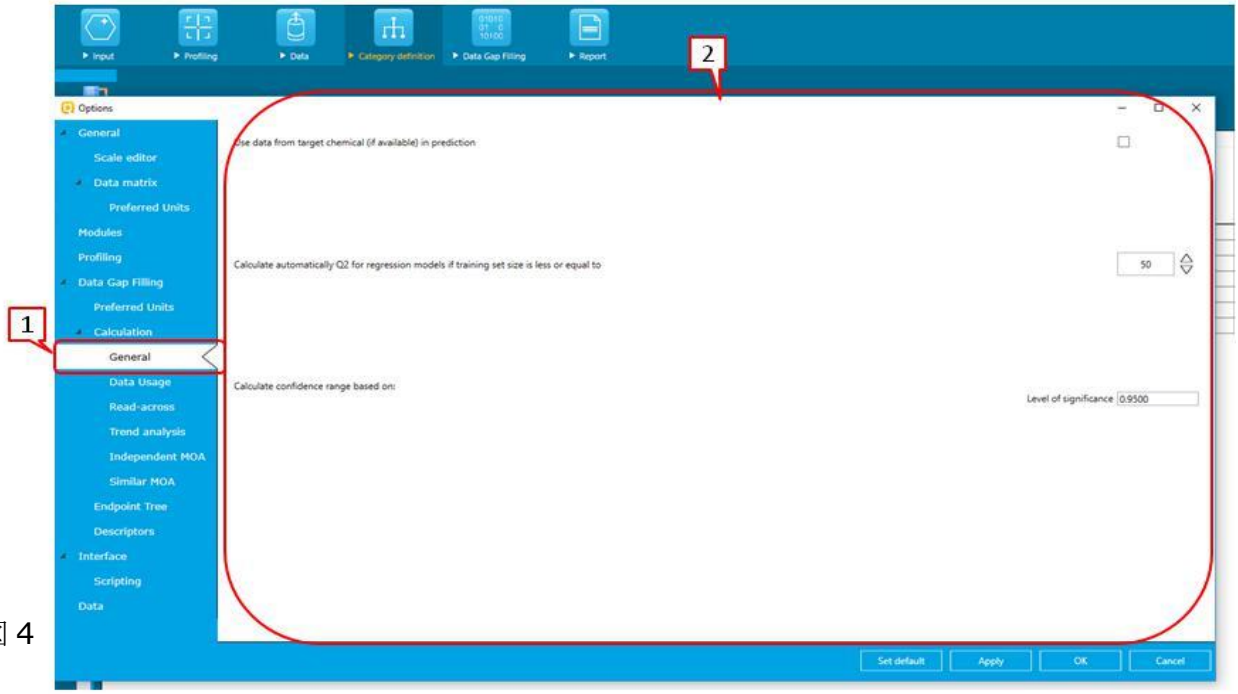


図 4

- Data usage (1): すべてのデータギャップ補完方法で使用されるスケールのタイプとそのデフォルト設定（2）を示します（図 5）。

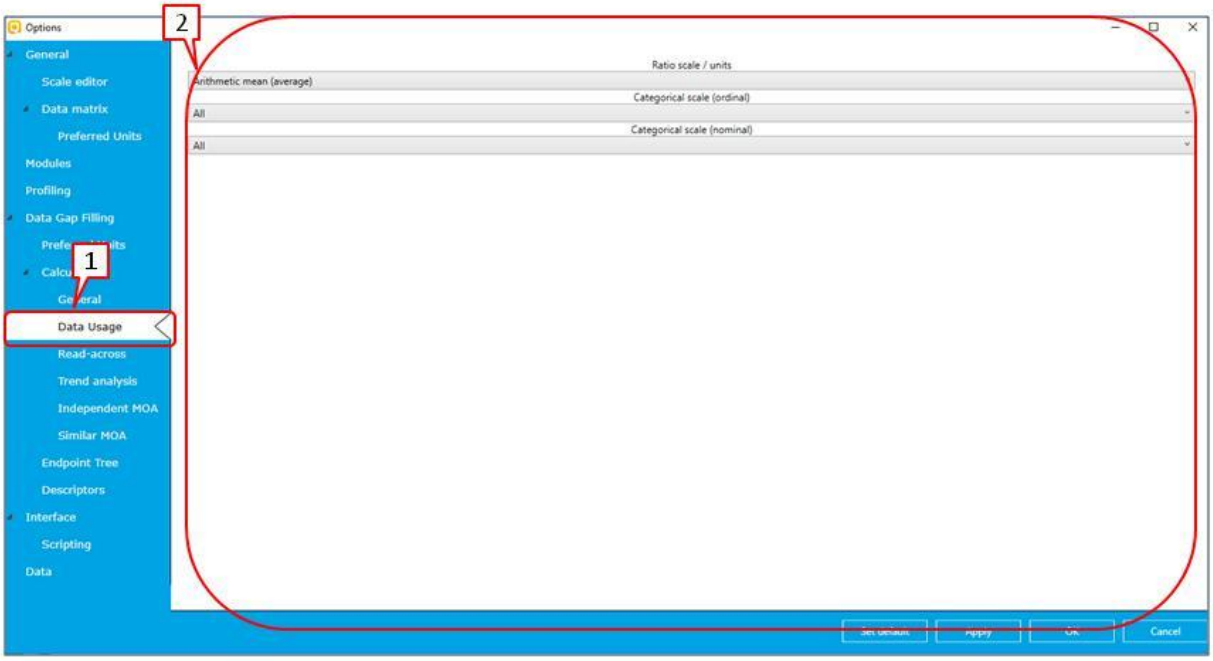


図 5

各スケールのドロップダウンメニュー（図 6-1）を選択することによってそれらを変更することができます。



図 6

- Read-across (1) : リード・アクロス予測（2）で使用される最も近い類似物質のデータ使用のデフォルト設定を示します。リード・アクロスで使用される最も近い類似物質の設定と数は、ユーザによって変更できます（図 7）。

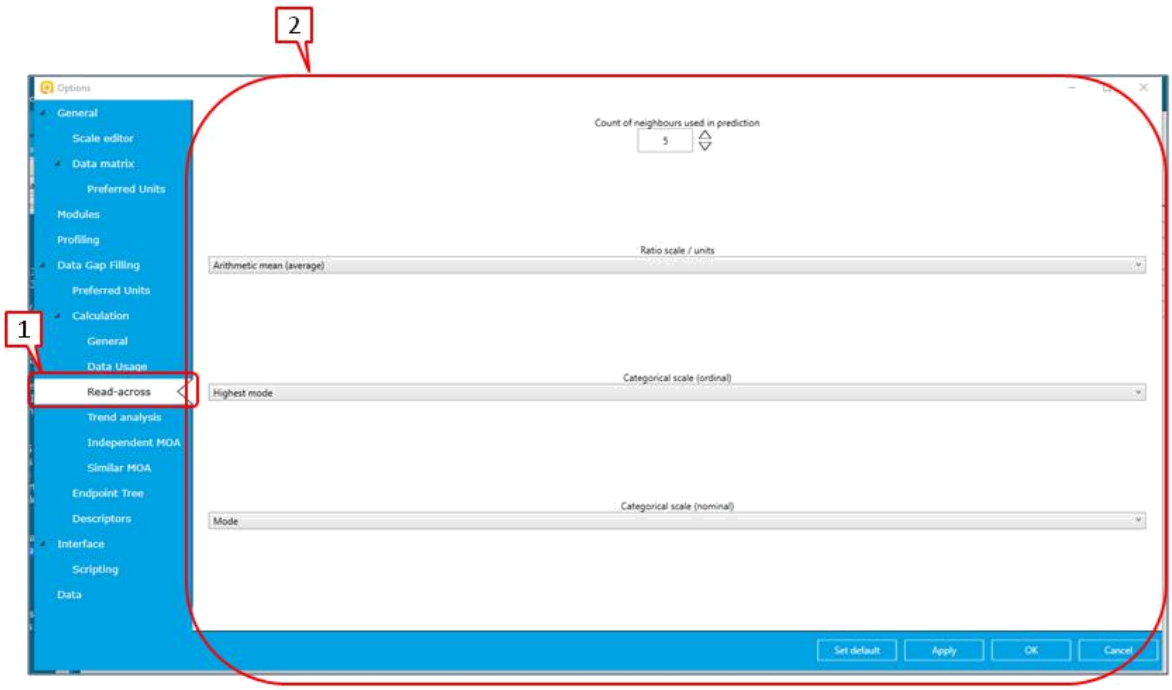


図 7

-Trend-analysis (1): トレンドアナリシス手法で使用される近似のタイプを線形または二次式(2)で与えます。(図 8)

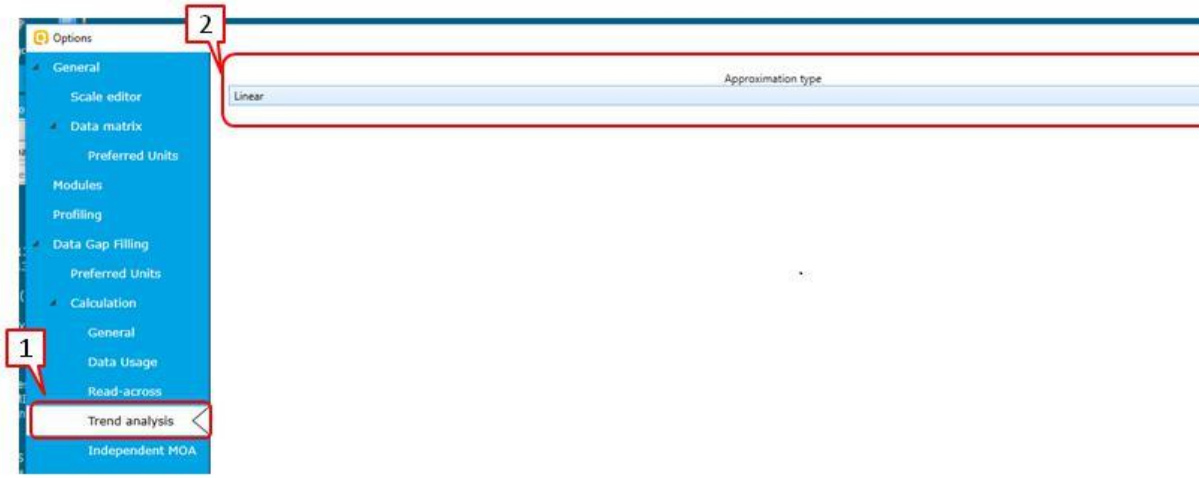


図 8

Independent MOA（独立した作用機序）（ 1 ）: このデータギャップ補完方法は混合物に使用されます。そのオプションは、最も好ましくないシナリオ（2）を表す設定になっています（図 9)。

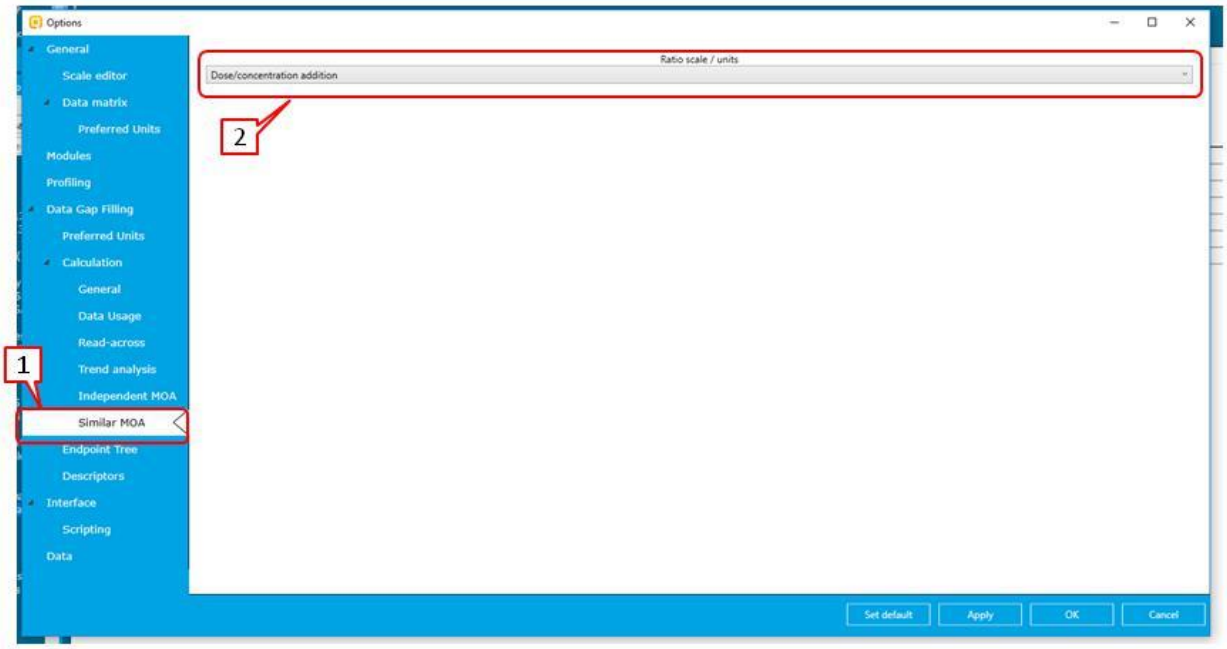


図 9

- Similar MOA（類似した作用機序）（1）（図 10）：このデータギャップ補完方法は混合物に使用されます。1 つだけ設定することができます（2）。個々の成分の数量を考慮するために選択します。

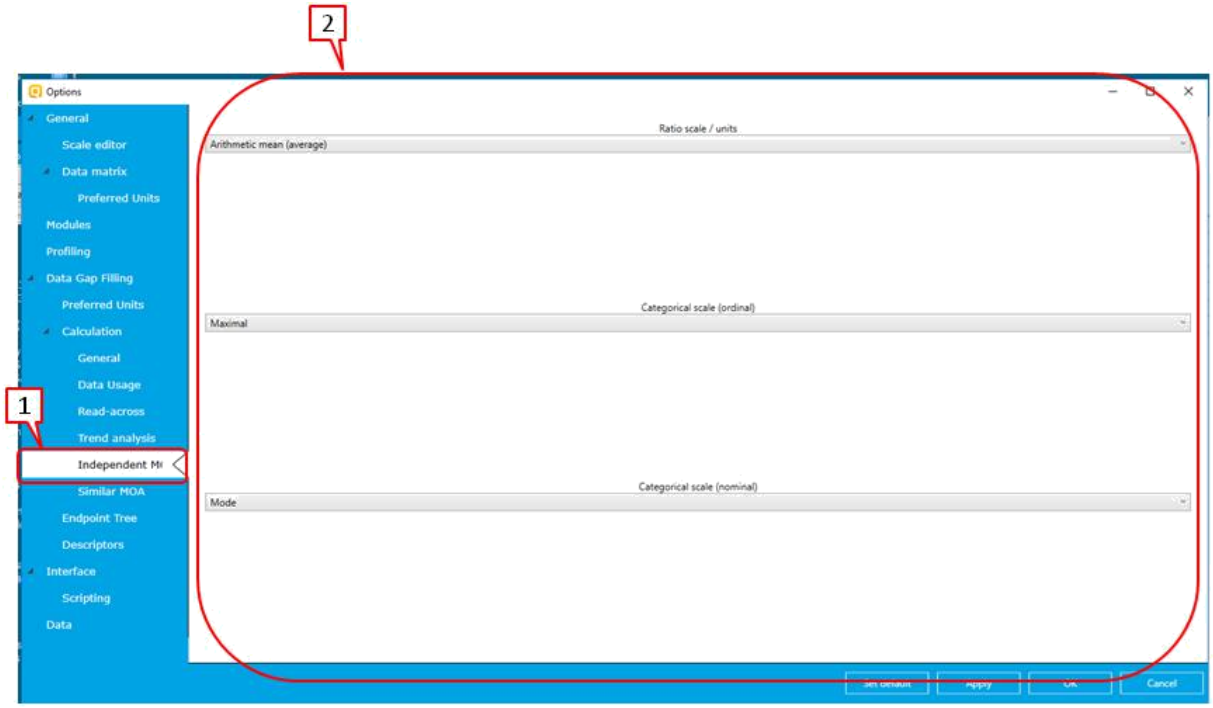


図 10

- **Endpoint tree（エンドポイントのツリー）** (1): このセクションでは、各エンドポイント予測の x 軸について適切な記述子を選択できます。(図 11)

  Default (2)を展開します。

  エンドポイントを選択し（3）、

  - 記述子（4）を設定します。デフォルトでは、x 軸の記述子は log Kow に設定されています。

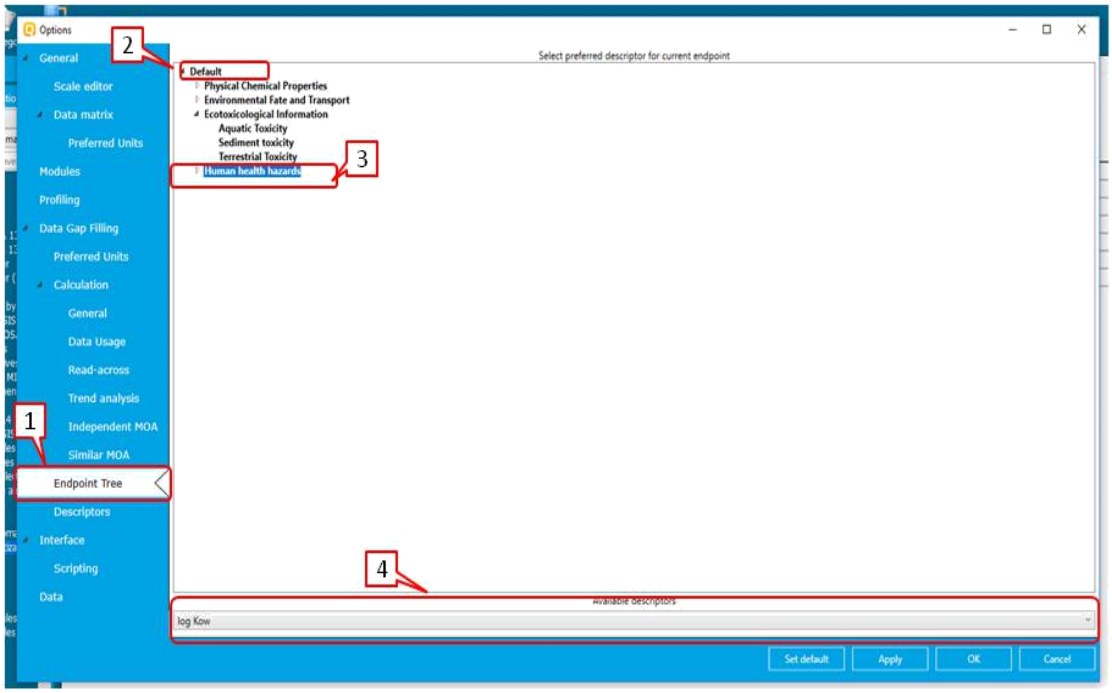


図 11

- **Descriptors** （グラフ軸の記述子）（図 12-1）: このセクションでは、予測の x 軸で使う記述子の優先単位と式を選択できます。
  - 記述子（2）を選択します。
- Scale details パネル（3）を開きます。
- 適切な単位と式（4）を選択します。

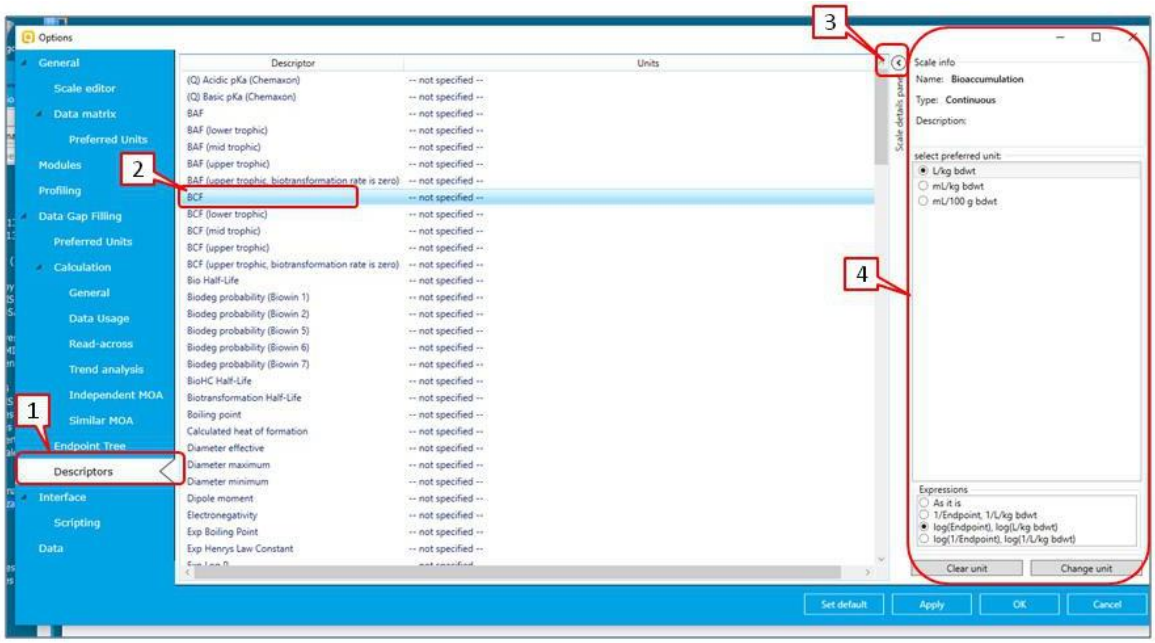
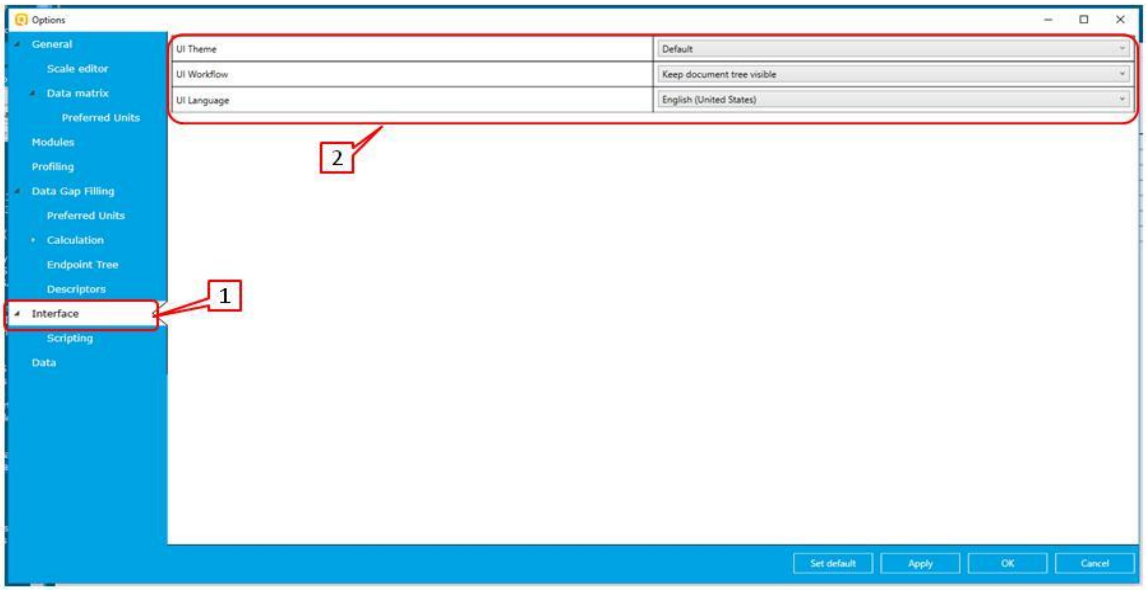


図 12

E.5.Interface（インターフェース）

このオプション（1）を使用すると、インターフェース（2）の一部の機能を変更できます（図 1）。

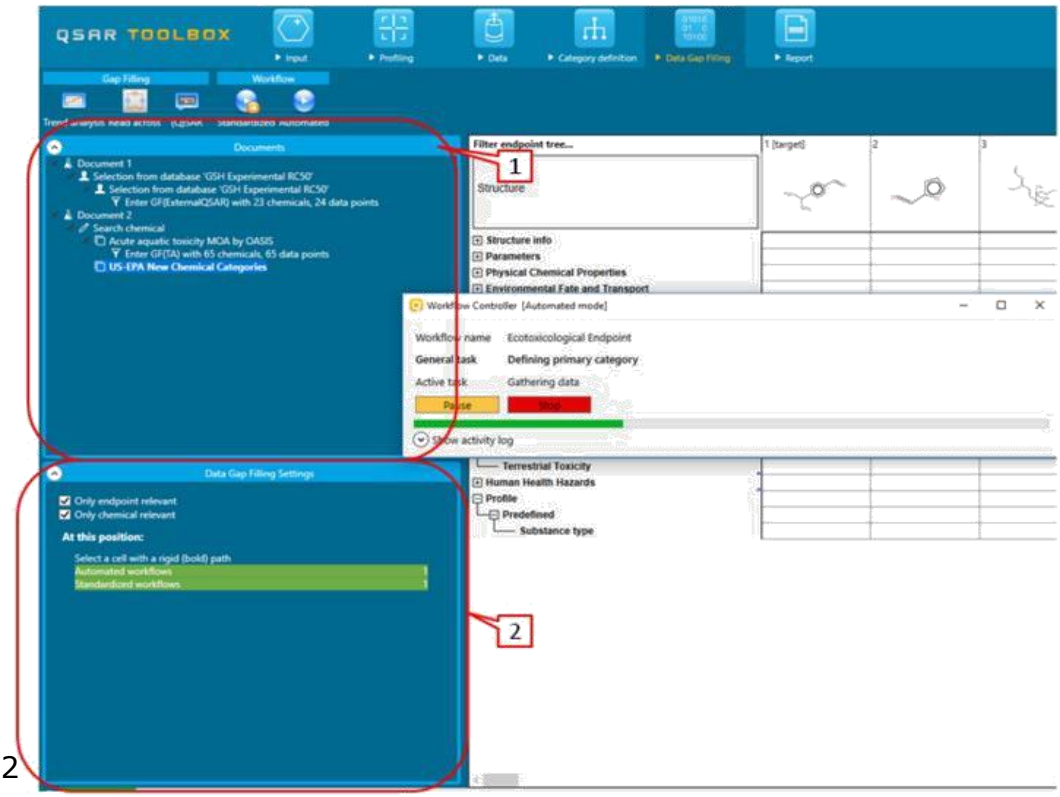
図 1



次の設定を使用できます。

- UI (User Interface) Theme:いくつかのテーマをご用意しています。
- UI （ユーザインターフェース） Workflow ：ワークフローの実行中に、左右のサイドパネルを制御します。ドキュメントツリーの view/hide（表示/非表示）（1）、データギャップ補完設定パネル（2）の view/hide のオプションを提供します（図 2）。

図 2



- Scripting（スクリプトの作成）

Scripting（図 3-1）には、Data gap filling で利用可能な、自動ワークフローと標準ワークフローのインターフェースに関連する設定（2）があります。Enable logging オプションは、実行中のワークフローの進行状況を確認できるようにするため、デフォルトではオンになっています（図 4-1）。

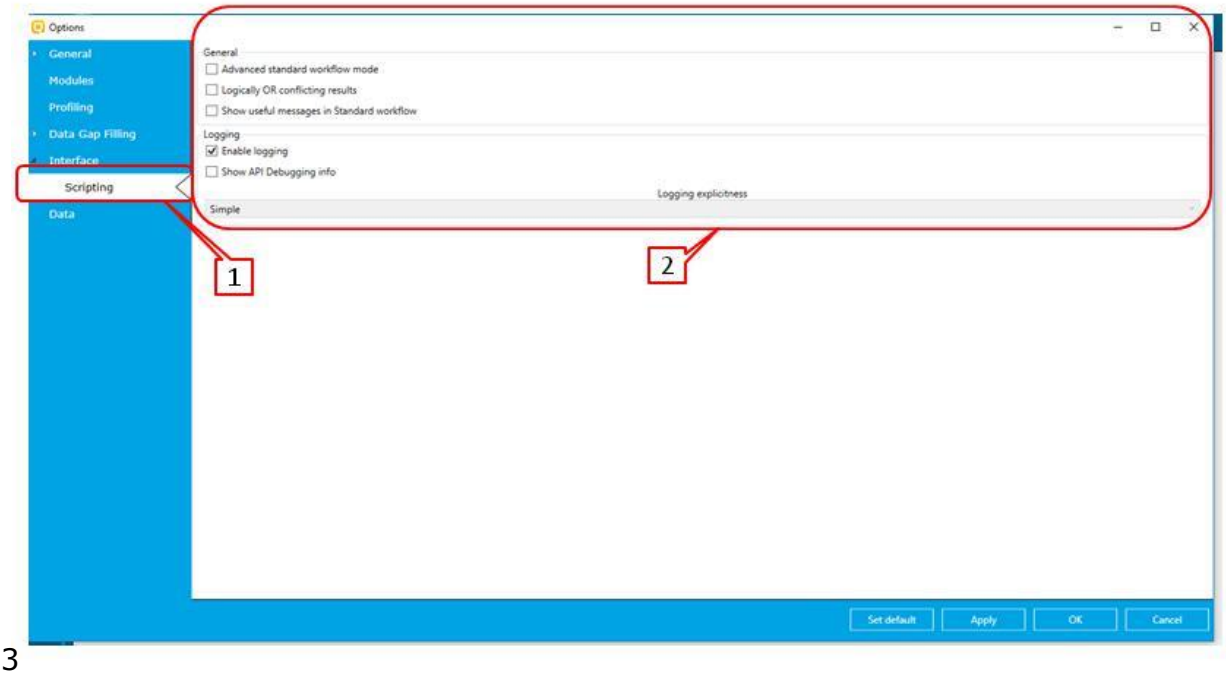


図 3

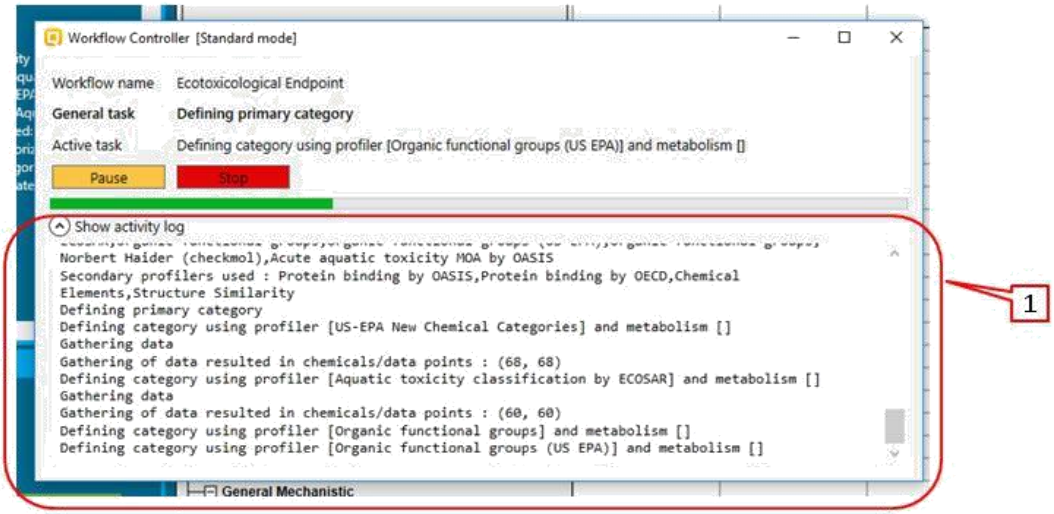


図 4

E.6.Data（データ）

Data（1）では、ソフトウェアが起動するたびにデフォルトで選択される、データベース（2）またはインベントリ（2）を選択できます。(図 1)

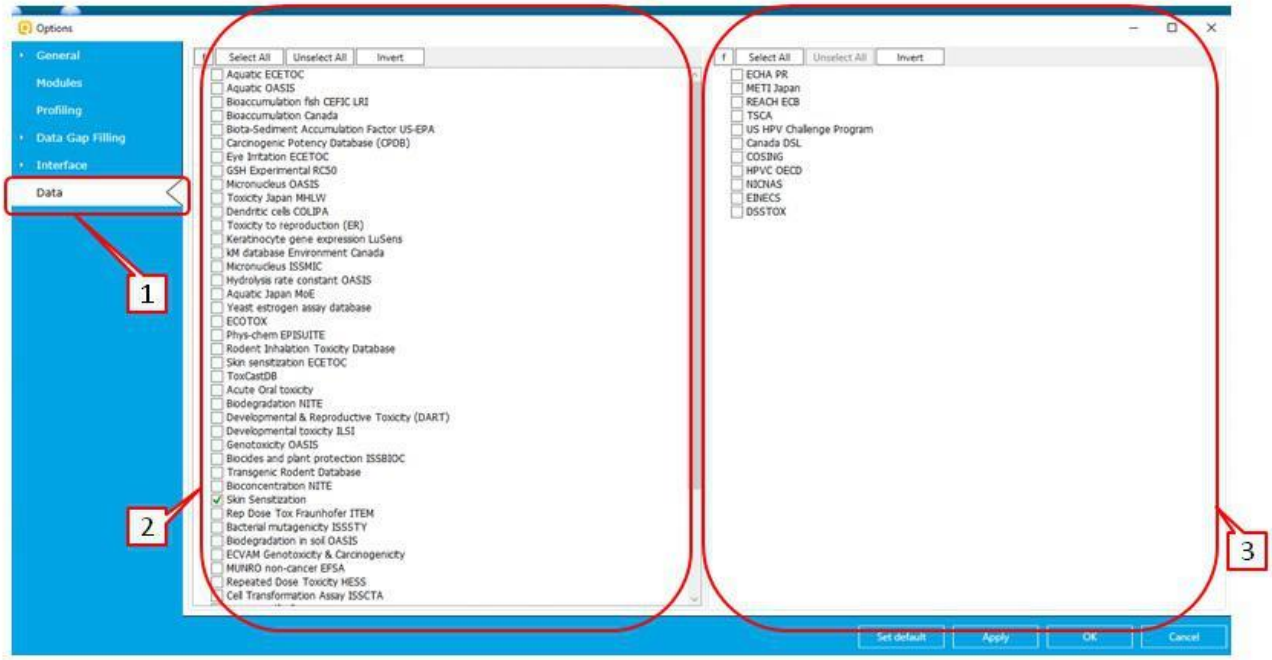


図 1