

# CHEMISTRY COMPONENT COLLECTION

Chemistry Collectionは、分子のプロパティの計算、フィルタリング、解析等、化学構造式に関連する処理を行うのに最適で包括的なスイートを提供します。このコレクションは、PipelinePilotの標準機能に、化合物情報処理やケムインフォマティクス解析などの機能を追加するコレクション群になります。この拡張機能を使用すると、以下のような化学構造式に関連した幅広い分野のプロトコルを作成することができます。

## CHEMISTRY COMPONENT COLLECTIONを使用すると、次のことを実現できます。

化合物ライブラリの収集:

- ・ ライブラリのクリーンアップと標準化
- ・ 複数ライブラリの比較
- ・ 部分構造と類似構造の検索
- ・ 幅広い特性プロファイリングとサブセットの選択

コンビケムライブラリの設計:

- ・ 骨格および反応に基づく列挙

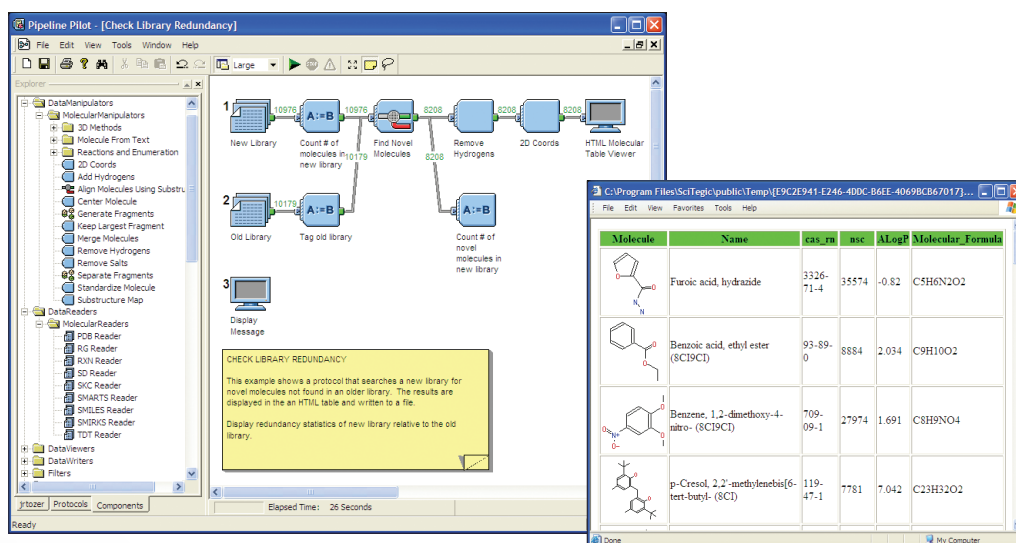
## Modeling Collectionとの

組み合わせによって以下の事が実現:

- ・ 構造活性モデルの作成
- ・ 化合物のクラスタリング
- ・ MCS(Maximal common substructure)の抽出





## READERS AND WRITERS

Chemistry Collection には、多くの化学ファイルフォーマットに対応したリーダー/ライター機能が搭載されています。SD、RG、RXN、SMILES、SMIRKS、SMARTS、TDT、MOL2、およびPDBなど



The screenshot displays the Pipeline Pilot interface with a workflow for checking library redundancy. The workflow includes steps for reading a new library, counting novel molecules, finding novel molecules, removing hydrogens, and displaying the results in an HTML table. A yellow box highlights the 'CHECK LIBRARY REDUNDANCY' step, which searches for novel molecules not found in an older library.

The resulting HTML table shows the following data:

Molecule	Name	cas_rn	usc	ALogP	Molecular_Formula
	Furoic acid, hydrazide	3326-71-4	35474	-0.82	CSH6N2O2
	Benzoic acid, ethyl ester (SC19CT)	93-89-0	8884	2.034	C9H10O2
	Benzene, 1,2-dimethoxy-4-nitro- (SC19CT)	709-09-1	27974	1.691	CSH9NO4
	p-Cresol, 2,2-methylenebis(6-tert-butyl- (SC1)	119-47-1	7781	7.042	C23H32O2

の形式です。ISIS™、MDL Direct、ActivityBase、Accord などのデータベースから構造データを読み込むこともできます。

## VIEWERS

Chemistry Collectionでは、Pipeline Pilotの標準ビューアに加えて、DS VisualizerやISIS for Excelなどのビューアともシームレスに統合することができます。さらに、分子データをグラフ表示できる、さまざまなWebビューアが利用できます。

## MANIPULATORS

Chemistry Collectionは、塩、Tautomers、立体化学、電荷に基づいて構造を修正するマニピュレータを備えています。このため、一連の分子を正規化してから、分子の比較、化学反応の適用、化合物ライブラリのマージ、およびコンビケムエニュミレーションを行うことができます。構造をアライメントするマニピュレータと、高速で高品質な2Dおよび3Dレイアウト用のマニピュレータを使用できます。

## FILTERS

化合物のプロパティ構造または計算結果に基づいてフィルタをかけるコンポーネントが提供されています。これらのフィルタは、簡単な特性値の閾値 (Lipinski Filter の閾値など) から、より高度な部分構造検索や重複チェック等のフィルタを行うことが可能です。ダイバーシティまたは類似性に基づくサブセットの抽出にもフィルタを使用できます。個別に必要なフィルタを作成する事もできます。

## PROPERTY CALCULATORS

Chemistry Collection では、さまざまな分子特性計算を高速で行うことができます。1秒あたり数百、数千の分子を解析できます。分子特性は、AlogP、logD、pKaと溶解度、分子量などがあります。他にも、トポロジー インデックス、さまざまな分子特性数などがあります。

## STRUCTURAL FINGERPRINTS

構造上のフィンガープリントを計算する Pipeline Pilot 独自の方法が Chemistry Collection に含まれています。この方式は Extended Connectivity Fingerprints (ECFP) 法として知られています。最大40億のさまざまな構造的特徴を利用して分子の各原子の環境にインデックスを付けることができます。検索、クラスタリング、およびモデリングの適用分野に有効な、非常に高速な手法です。ECFP法は、別コレクションであるModeling Collection の Bayesian 学習法と併せて使用すると、大規模データセットに対する説予測を構築できます。

## MOLECULAR TOOLKIT

Molecular ToolkitのJavaおよびPerl APIをIntegration Collection と組み合わせて使用すると、分子データモデルの検索をプログラムから利用できます。

Pipeline Pilotの詳細については、次のURLを参照して下さい。  
<http://accelrys.co.jp/products/pipeline-pilot>