BIOVIA

3DEXPERIENCE®

CHEMISTRY COMPONENT COLLECTION

Datasheet

The Chemistry Collection offers a comprehensive suite of molecular property calculators, filters, and manipulators. This collection of modular components extends the standard capabilities of BIOVIA Pipeline Pilot to include compound processing and cheminformatics analyses. With these enhancements, you can create protocols for a broad array of chemistry applications.

WITH THE CHEMISTRY COMPONENT COLLECTION YOU CAN:

Conduct compound library acquisition:

- Library cleanup and standardization
- Multiple library comparisons
- Substructure and similarity searching
- Extensive property profiling and subset selection

Perform combinational library design:

- Scaffold and reaction driven enumeration

Combined with the separately available Modeling Collection, you can perform:

- Structure activity modeling
- Compound clustering
- Maximal common substructure searching

READERS AND WRITERS

The Chemistry Collection includes file readers and writers for most industry-standard chemistry formats. It includes SD, RG, and RXN (from MDL), SMILES, SMIRKS, SMARTS, and TDT (from Daylight), MOL2 (from Tripos) and public formats such as PDB. You can also read molecular data from databases such as ISIS™ and BIOVIA Direct, ActivityBase (from IDBS), and Accord (from BIOVIA).

VIEWERS

In addition to BIOVIA Pipeline Pilot's standard viewers, the Chemistry Collection seamlessly integrates with several popular structural viewers including Discovery Studio Viewer and Insight for Excel. Furthermore, a variety of Web viewers are available providing a way to graphically display your molecular data.

MANIPULATORS

The Chemistry Collection includes manipulators that modify structures according to their salts, tautomers, stereochemistry, charges, to name but a few. You can normalize sets of molecules before comparing them, apply chemical reactions, merge collections of molecules and even enumerate combinatorial libraries. For optimal display characteristics, you can employ manipulators for structural alignment and for fast, high quality 2D and 3D layout.

DASSAULT SYSTEMES

FILTERS

Numerous filter components are provided to separate molecules into groups based on chemical properties or calculations. These filters range from simple property value thresholds, like those in the Lipinski Filter, to more sophisticated criteria, such as occurrences of substructures or duplicates. Filters are available to select representive subsets based on diversity or similarity. An interactive filter allows you to select a subset of molecules for further processing. What's more, you can create custom chemistry filters for any property values that you require.

PROPERTY CALCULATORS

The Chemistry Collection includes a variety of fast molecular property calculators. You can characterize hundred or thousands of molecules per second. Whole molecule properties include AlogP, logD, pKa and solubility, and molecular weight. Other properties include electrotopological indices, solvent accessible surface area and volume descriptors and a large variety of molecular property counts.

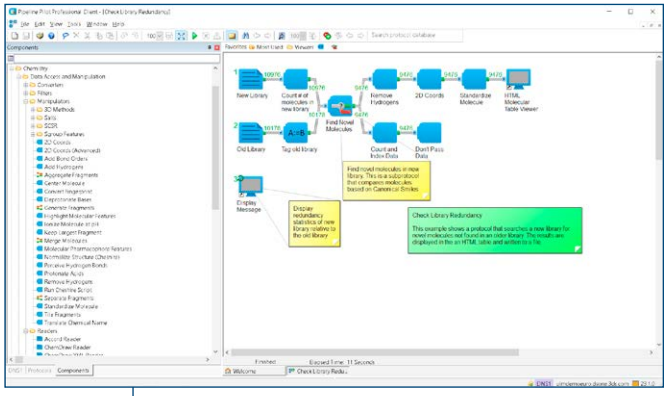


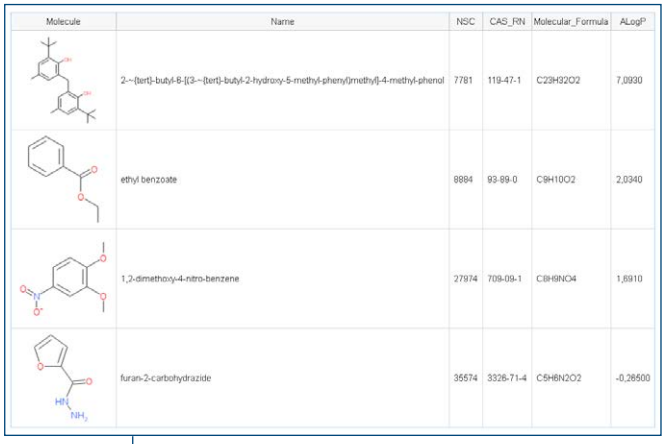
Figure 1. An example protocol showing cleanup by removing duplicates from a library

STRUCTURAL FINGERPRINTS

A proprietary method for calculating structural fingerprints is included in the Chemistry Collection. Known as Extended Connectivity Fingerprints (ECFP), this method offers excellent characterization of molecules, indexing the environments of every atom in a molecule using up to four billion different structural features. It is an extremely rapid method useful in searching, clustering, and modeling applications. When used with Bayesian learning technologies in the separately available Modeling Collection, ECFP provides interpretable predictive models for large data sets. You can also use Daylight,MDL or user defined keys within the system.

MOLECULAR TOOLKIT

When combined with the Integration Collection, the Molecular Toolkit's Python, Java and Perl API's provide programmatic access to the molecular data model's methods.



Molecule	Name	NSC	CAS_RN	Molecular_Formula	ALogP
	2-~{tert}-butyl-6-[(3-~{tert}-butyl-2-hydroxy-5-methyl-phenyl)methyl]-4-methyl-phenol	7781	119-47-1	C23H32O2	7,0930
	ethyl benzoate	8884	93-89-0	C9H10O2	2,0340
	1,2-dimethoxy-4-nitro-benzene	27974	709-09-1	C8H9NO4	1,6910
	furan-2-carbohydrazide	35574	3326-71-4	C5H6N2O2	-0,26500

Figure 2. Simple browser-based viewing of structures and calculated properties

LEARN MORE

Our **3D**EXPERIENCE® platform powers our brand applications, serving 11 industries, and provides a rich portfolio of industry solution experiences.

Dassault Systèmes, the **3DEXPERIENCE** Company, is a catalyst for human progress. We provide business and people with collaborative virtual environments to imagine sustainable innovations. By creating 'virtual experience twins' of the real world with our **3DEXPERIENCE** platform and applications, our customers push the boundaries of innovation, learning and production.

Dassault Systèmes' 20,000 employees are bringing value to more than 270,000 customers of all sizes, in all industries, in more than 140 countries. For more information, visit **www.3ds.com**.

3DEXPERIENCE®

DASSAULT SYSTEMES | The **3DEXPERIENCE**® Company

Americas
Dassault Systèmes
175 Wyman Street
Waltham, Massachusetts
02451-1223
USA

Europe/Middle East/Africa
Dassault Systèmes
10, rue Marcel Dassault
CS 40501
78946 Vélizy-Villacoublay Cedex
France

Asia-Pacific
Dassault Systèmes K.K.
ThinkPark Tower
2-1-1 Osaki, Shinagawa-ku,
Tokyo 141-6020
Japan