

# UNITY<sup>®</sup>

LOCATE COMPOUNDS IN DATABASES THAT MATCH  
A PHARMACOPHORE OR FIT A RECEPTOR SITE



Three-dimensional database searching is an effective means of accelerating the discovery of lead compounds.<sup>1</sup> UNITY combines database searching with the molecular design and analysis tools in SYBYL<sup>®</sup> to provide an integrated environment for new compound discovery. With UNITY, researchers perform structural searches, explore relational data, and retrieve, manage, and analyze the resulting hits. UNITY creates databases for enterprise-wide access.

### Three-dimensional and Two-dimensional Searching of Databases

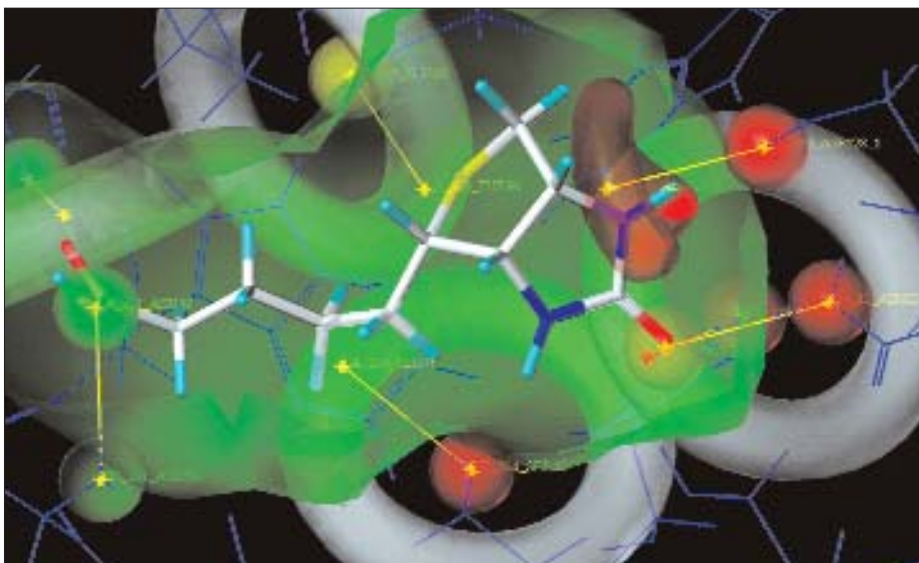
Conventional 3D database searching finds only molecules whose stored conformations match the constraints of the query. UNITY's conformationally flexible 3D searching finds molecules that can achieve a matching conformation regardless of the stored conformation. The Directed Tweak algorithm<sup>2</sup> in UNITY quickly finds molecules that could match the constraints of the query. Important molecules in compound classes never before considered are often identified.

UNITY allows you to build structural queries based on molecules, molecular fragments, pharmacophore models, or receptor sites.

In addition to atoms and bonds, 3D queries can include features such as lines, planes, centroids, extension points, hydrogen bond sites, and hydrophobic sites. Distance, angle, excluded volume, surface volume, and spatial constraints define the geometric relationships between features. In receptor-based queries, UNITY additionally has the ability to specify multiple excluded volumes that represent the binding site structure, or to define the containing volume of the receptor cavity.

Partial match directives (for example, match at least four but no more than seven of 10 possible H-bond donor sites) allow a search for compounds containing only some of the features specified in the query.

Two-dimensional database searching



A UNITY query constructed at the active site of the streptavidin/biotin complex (1STP). Yellow lines originate at hydrogen bonding sites of the protein (shown as spheres) and terminate within the spatial constraint for complementary ligand sites. The rotameric positions of hydroxyl and amine H-bond donors are shown as toroidal constraints. A surface constraint at the protein/ligand interface is shown in green. The spatial cap in red accounts for a bifurcated interaction with an Asp carboxyl. Partial match groups are shown in different colors: red, yellow, or green.

looks for connectivity patterns in molecules as specified by the query. UNITY's 2D search options include exact and substructure searching. A unique option of UNITY provides substructure searching of combinatorial libraries. Queries may also specify variable elements such as Markushes. Precomputed fingerprints that encode molecular features optimize search times by quickly eliminating compounds that cannot match a query. Similarity searching in UNITY is based on a comparison of these fingerprints.

#### Features

- Directed Tweak provides rapid, conformationally-flexible 3D searching

- An extensive set of features and constraints are available to construct 3D queries
- Partial matching of query features is allowed for any 3D search
- Full Markush capabilities enable variable query specification
- Fingerprints are used as filters to speed 2D and 3D searching
- Queries contain realistic hydrogen bond donor/acceptor site representations
- Fingerprints can be customized and edited

## Applications

- Conformationally flexible searching of compound databases for ligands that fit a receptor site
- Exploration of databases for compounds consistent with a pharmacophore hypothesis
- Lead explosion by retrieving similar compounds
- Virtual screening of compound databases to discover lead compounds
- Determining reagents in commercial databases that support combinatorial chemistry synthesis

## Storage of Compound Structures and Associated Data

UNITY creates Tripos databases that store compound structural data such as name, RegID, fingerprints, and 3D coordinates. Standard or custom fingerprints are calculated during database creation.

UNITY uses SYBYL Line Notation<sup>3,4</sup> (SLN), a flexible and concise language for storing molecular structures and for creating search queries. An extension of SLN, combinatorial SLN (cSLN), can specify a library of compounds as a single searchable entry.

UNITY utilities convert a file or database in one format to another. UNITY can transfer structures between Tripos databases and MOL2, MOL, and SD files, as well as Daylight SMILES strings. Commercial databases are easily converted to Tripos databases, and associated property data can be stored in Oracle®. Large 2D databases can be converted to 3D Tripos databases with Concord® (licensed separately), the industry standard for rapid conversion of 2D input to accurate, geometry-optimized 3D structures.

## Features

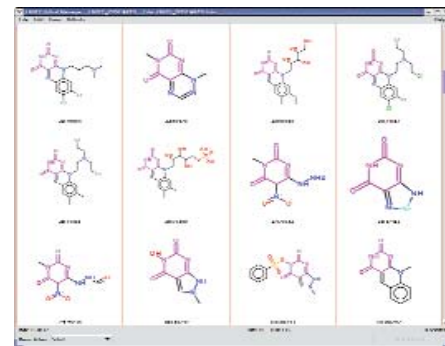
- Structures can be stored with 3D coordinates, fingerprints, and associated data
- Concise combinatorial library storage is made possible by combinatorial SLN
- UNITY 3D databases can be created from 2D databases using Concord and StereoPlex®
- UNITY is compatible with other vendors' databases

## Enterprise-wide Access to Structures and Data

UNITY can access databases anywhere on a network. Security locks may be implemented at the level of workgroups or individuals, and can be applied to specific sets of compounds.

## Features

- Multiple databases residing anywhere on a network can be searched simultaneously
- Network-wide accessibility eliminates the need to keep and synchronize multiple copies of databases



The UNITY Hitlist Manager, showing the results of a sub-structure search of the NCI database.

### Management, Post-processing, and Analysis of Search Results

UNITY and the SYBYL Molecular Spreadsheet™ simplify information management of hits returned by searches. Two-dimensional displays of structures, color-coded to show the portions that match the query, make it possible to quickly step through the search results. UNITY can generate derivative hitlists by performing Boolean operations on searches, by making refinements to the original query and rerunning the search, or by directly searching the hitlists themselves.

The structures returned by a database search can be refined with post-processing tools available in UNITY and SYBYL.

A torsional minimizer relaxes structures by decreasing internal energy while continuing to satisfy query constraints. Alternatively, the torsional minimizer can be used to tighten UNITY hits so that they better match the query constraints. UNITY ranks hits based on strain energy, RMSD of the features in hits to those in the query, and the number of rotatable bonds.

UNITY and SYBYL are tightly integrated. Candidate structures returned by UNITY from 3D searches are automatically aligned with the query to facilitate analysis by QSAR with CoMFA®. Structures returned from receptor-site queries are fit into the receptor cavity. FlexX™ and CScore™ enable virtual screening of ligands returned by UNITY.

### Features

- Results may be viewed in a Molecular Spreadsheet or the UNITY Hitlist Manager
- Boolean operations can be used to combine different hitlists
- Hitlists or Molecular Spreadsheets can also be searched
- Hits returned by UNITY are aligned to the query, facilitating analysis by QSAR with CoMFA
- Hits returned from receptor-site queries are docked to the receptor
- Post-processing of UNITY hits includes torsional minimization constrained by the query

### Hardware and Software Requirements

UNITY is accessible through the SYBYL expert molecular modeling environment. SYBYL requires a separate license. Licensing options for UNITY include: UNITY/Base for 2D searching, UNITY 3D for 3D searching, UNITY extra search engine for running multiple searches concurrently, and UNITY RDBMS for searching relational data in Oracle. UNITY is also available standalone. UNITY and SYBYL run on workstations operating under Mac OS X or Linux® (x86).

### Complementary Software

- **Advanced Computation** for exploring the conformational properties of compounds.
- **AMPAC™** for calculating transition states and spectral properties using semiempirical quantum mechanical methods.

- **Auspyx®** for storing and searching chemical structures and relational data directly inside Oracle databases.
- **Concord** for generating accurate 3D coordinates.
- **Confort™** for generating sets of diverse, low energy conformers.
- **CScore** for ranking the affinity of compounds bound to a target with consensus scoring.
- **Distill™** for determining and visualizing SARs.
- **Surflex-Dock™** for flexibly docking ligands into a binding site
- **QSAR with CoMFA** for building predictive structure-activity and structure-property models.
- **StereoPlex** for expanding the stereochemical diversity of a database.
- **Tuplets™** for pharmacophore-based virtual screening without a 3D model.

### References

1. Martin, Y.C. "3D Database Searching in Drug Design." *J. Med. Chem.* **1992**, 35, 2145-2154.
2. Hurst, T. "Flexible 3D Searching: The Directed Tweak Technique." *J. Chem. Inf. Comput. Sci.* **1994**, 34, 190-196.
3. Ash, S.; Cline, M.A.; Homer, R.W.; Hurst, T.; Smith, G.B. "SYBYL line notation (SLN): A versatile language for chemical structure representation." *J. Chem. Inf. Comput. Sci.* **1997**, 37, 71-79
4. Clark, R.D.; Homer, W.R.; Hurst, T.; Jilek, R.J.; Swanson, J. "SYBYL line notation (SLN): A single notation to represent chemical structures, queries, reactions, and virtual libraries." *J. Chem. Inf. Model* **2008**, Article ASAP



WWW.TRIPOS.COM			CONTACT_US@TRIPPOS.COM	
AUSTRALIA +61 (7) 5439 9775	FRANCE +33 1 69 59 29 43	GERMANY +49 89 45 10 300	JAPAN +81 3 3662 0521	UNITED STATES 800 323 2960 +1 314 647 1099
<small>© Tripos L.P., 2008 0075 Printed in USA All trademarks are the property of their respective owners.</small>				