

Distill

DETERMINE AND VISUALIZE STRUCTURE-ACTIVITY RELATIONSHIPS



The heart of any drug discovery effort lies in elucidating structure-activity relationships (SARs), whether the endeavor is based on testing a handful of compounds or assaying thousands of compounds by high-throughput screening. Distill™ is a hierarchical clustering tool that classifies compounds according to their common substructures and organizes the results in a display that enables visualization of SARs.

Applications

- Extract structure-activity relationships from screening data
- Formulate and evaluate hypotheses about structural features important to activity
- Compare new compounds with previously tested compounds to assess whether or not they may be active
- Prioritize compounds for synthesis from a virtual compound library
- Construct queries to search databases for other compounds containing substructures correlated with activity
- Facilitate 3D QSAR analysis by identifying common core structures

Distill's clustering algorithm creates a hierarchy in which each node contains compounds that share a common structural core. The results are presented graphically in an interactive dendrogram, where each node represents a substructure and the set of compounds containing that substructure. By averaging a property such as activity for the compounds at a node, and then coloring each node according to this average, Distill can visually relate substructures to activity. Navigation tools facilitate movement within the hierarchy.

By clicking on any node in the dendrogram, you can display the substructure associated with that node, as well as those of the parent and child nodes. Statistics about the

properties of compounds in a node are displayed. By clicking on a substructure, you can see all the compounds containing that substructure. This colored roadmap shows not only how incremental change in structure adds or detracts from activity, but how combinations of changes may provide a pathway to enhanced activity.

Distill classifies compounds from scores based on the number of atoms, bonds, ring bonds, heteroatoms, and branched atoms in the common substructure. Several heuristics improve on the computational time required for hierarchical clustering, which depends on the number and diversity of the compounds clustered. The property color mapping is independent of the clustering, which is strictly a function of structure.

Advantages

■ Because Distill orders and sorts compounds the way a medicinal or synthetic chemist would, researchers can objectively evaluate their own conclusions about structure-activity relationships for a given receptor.

■ Distill exploits the power of the human eye for pattern recognition of chemical

structures, making it possible to easily formulate SAR hypotheses from overwhelming volumes of data.

- Unlike other clustering tools that relate descriptors to properties, Distill relates the components of structure (atoms, bonds, and connectivity) to a specific property, facilitating the next generation design by chemists.
- Users can select nodes in the dendrogram from which Distill constructs UNITY® 2D search queries in order to identify similar interesting compounds from another database.

Features

- Construction of structural hierarchies independent of property or activity data.
- Flexible, tunable scoring function for focusing on important substructures.



Distill's interface showing a dendrogram color-coded by activity for a validation test set, and the linked display of the common substructures.

- Fuzzy atom and bond capabilities based on a Markush representation that allow identification of larger skeletal backbones.
- Statistics of activities or properties calculated for all compounds at each node in the dendrogram.
- Automatic search query generation for compounds containing exclusively the substructure in one node and not others.
- Instant traversal of the dendrogram up, down or sideways with interactive links to the corresponding common substructures and compounds.
- Alignment of molecules within a node based on their common core for subsequent use in 3D QSAR tools such as QSAR with CoMFA®.

Validation

Distill has been tested against more than twenty pharmaceutical data sets, from a number of disparate biological receptor classes. The results show that, in every case, the most active compounds cluster tightly into one or more branches of the dendrogram. When compounds from several different chemical series are included in a single data set, the highest level of branching in the dendrogram reflects the different chemical series.

Hardware and Software Requirements

Distill requires a separate license and is accessible through the SYBYL® expert molecular modeling environment. UNITY searching options spawned from Distill require a UNITY 2D searching license. Distill, SYBYL, and UNITY run on workstations operating under IRIX® (SGI®) or Linux® (x86).

Complementary Software

- **Advanced CoMFA®** for refining and enhancing 3D QSAR models.
- **Almond™** for calculating and utilizing alignment independent molecular descriptors.
- **HQSAR™** for performing automated QSAR analyses.
- **QSAR with CoMFA** for building predictive structure-activity and structure-property models.
- **VolSurf™** for predicting ADME properties.
- **UNITY** for locating compounds in databases that match a pharmacophore or fit a receptor site.



WWW.TRIPOS.COM				CONTACT_US@TRIPPOS.COM		
AUSTRALIA +61 (7) 5439 9775	CANADA +1 450 4334500	FRANCE +33 1 69 59 29 49	GERMANY +49 89 45 10 300	JAPAN +81 3 5166 1721	UNITED KINGDOM +44 1 908 650000	UNITED STATES 800 323 2960 +1 314 647 1099
©Tripos, Inc. 2005 0075 Printed in USA All trademarks are the property of their respective owners.						