

StereoPlex[®]

EXPAND THE STEREOCHEMICAL DIVERSITY OF A DATABASE



Most compound collections (corporate databases, screening libraries, compound-provider catalogs) are comprised of 2D structures with incomplete and/or incorrect indications of stereochemistry. This inevitably leads to arbitrary and/or incorrect stereochemistry when these structures are converted to 3D. Although a compound may be registered with complete stereochemistry, often the "wrong" stereoisomer of an active compound is used in a virtual HTS application such as docking, 3D-searching, or 3D-QSAR. This may result in a missed hit of potentially immense importance. StereoPlex addresses this issue.

With the increasing emphasis on vHTS applications and increased compute-power to support those applications, StereoPlex offers a robust, industry-tested solution to a problem which simply can not be ignored. StereoPlex generates multiple stereoisomers of each input structure according to a user-specified limit on the number of stereoisomers and a user-specified priority rule which tells the program which stereoisomers to generate if the complete set would exceed the user's limit. Significantly, subject to the user-specified priority rule, StereoPlex "multiplexes" both atom-centered (R/S) and bond-centered (E/Z) chirality. Also significantly, StereoPlex does not simply output all allowed permutations of chiral specifications. Rather, StereoPlex employs a novel and unique algorithm to exclude topologically impossible stereoisomers.

Features and Benefits

- Enables user-controlled "multiplexing" of stereocenters (to avoid missing "hits" in docking, 3D-search, etc. due to using the incorrect stereoisomer for the given target)
- Addresses both atom-centered (R/S) and bond-centered (E/Z) chirality
- User specifies maximum number of stereoisomers to output per compound
- User specifies priority rule for determining which stereocenters to address if maximum would be exceeded
- Automatically eliminates topologically impossible stereoisomers of ring-systems
- Supports input and output in SMILES and SDfile format
- Easily incorporated into other programs, shell scripts, and "pipelined" protocols

The Problem

Bioreceptors are stereoselective, therefore 3D-search queries should be stereoselective. Concord[®] generates 3D structures which are 100% consistent with the stereochemistry specified in the 2D input. However, 2D databases frequently contain incorrect or incomplete stereochemical specification and, sometimes, contain no stereochemical specifications whatsoever. Even if the 2D stereochemical specifications are complete and correct, the particular stereoisomer stored in the database might not match the current 3D-search query.

Extending a database by adding additional stereoisomers of each compound addresses this problem. However, combinatorial "multiplexing" of all n stereocenters contained in each compound is obviously not a practical solution. Thus, several problems need to be addressed:

- Which of the 2^n possible stereoisomers should be generated?
- Which of the 2^n possible stereoisomers are topologically or sterically impossible?
- How should we treat those stereocenters for which stereochemical specifications were given in the 2D input?

- How should the new stereoisomers be "named"?

The Solution

StereoPlex is designed specifically to address the need for additional stereoisomers in 3D databases. StereoPlex functions as a 3D *database-extender* and can be used either by itself or as a Concord pre-processor to provide control over an otherwise unmanageable task.

Applications

Input to StereoPlex consists of 2D connection tables (CTs) in either SMILES or MDL[®] molfile format. These 2D CTs may contain either complete or incomplete stereochemical specifications. The maximum number of stereocenters to be multiplexed for any given compound (n_{\max}) may be specified. The user decides whether or not to multiplex centers which were stereo-specified on input. The user specifies the *priority rule* by which up to n_{\max} stereocenters will be selected by StereoPlex from the prioritization of the following classes of stereocenters:

- Acyclic double bonds
- Ring fusion atoms
- Ring bridgehead atoms

- Other ring atoms
- Acyclic atoms

The priority rule also specifies how to prioritize centers within the same class. StereoPlex computes a symmetry rank for each stereocenter which indicates whether the center is towards the "middle" of the molecule or closer to the "periphery." The user decides which gets higher priority. The priority rule could also specify that centers within one or more classes should not be multiplexed at all.

StereoPlex provides researchers with complete control while freeing them from a very important and difficult task. A novel algorithm identifies and discards stereoisomer-candidates which are topologically impossible (e.g., 254 out of the 28 = 256 possible candidates derived from a chirally substituted cubane).

StereoPlex even provides control of options for constructing unique stereoisomer names or database registration IDs.

Users can request StereoPlex output in either SMILES or MDL SDfile formats. On UNIX® systems, the use of "pipes" avoids the overhead of file-I/O, an especially convenient feature of StereoPlex.

Concord and StereoPlex

StereoPlex was designed to be used as a Concord pre-processor. Concord never generates 3D structures which are not 100% consistent with the stereochemistry specified in the 2D input. However, not all stereoisomers which are topologically possible are also sterically possible. In the process of converting StereoPlex output into

an extended 3D database, Concord provides a key advantage by eliminating sterically strained or impossible stereoisomers, thus maximizing "stereodiversity" while discarding "stereo-junk."

Molecular Diversity

In addition to extending databases, StereoPlex also acts as a *database-diversifier*. StereoPlex 2D fragment or fingerprint-based measures of compound similarity are useful descriptors of molecular diversity. However, they perceive two stereoisomers as being identical.

Bioreceptors, on the other hand, typically show strong discrimination between stereochemical isomers. Clearly, "stereochemical" diversity is just as important as "functional group" diversity.

StereoPlex can be used synergistically with other software to fully address molecular diversity requirements. Adding StereoPlex's "stereochemical" compounds to a database is as useful as adding compounds with diverse functional groups derived from other sources. Moreover, adding StereoPlex-generated compounds offers the following advantages:

- The StereoPlex-generated stereoisomer may already be covered under a patent for the "parent" compound.
- A sample of the stereoisomer may already be in the storeroom — either as a component of a racemic mixture or as a pure compound for which the original 2D (parent) CT was registered with incorrect or incomplete stereochemical specification.
- Even if the compound is not readily available, a researcher will have

experience in synthesizing its StereoPlex "parent." Synthesizing the stereoisomer is thus easier than starting from scratch.

Hardware and Software Requirements

StereoPlex requires a separate license, and runs on workstations operating under IRIX® (SGI®) or Linux® (x86).

Complementary Software

Integration of StereoPlex in the SYBYL® expert molecular modeling environment allows users to access SYBYL's powerful molecular design and analysis tools.

- **ClogP/CMR** for including molar refractivity and logP in QSAR and ADME models.
- **Concord** for generating accurate 3D coordinates.
- **Confort™** for generating sets of diverse, low energy conformers.
- **DiverseSolutions®** for designing, comparing, or selecting compound libraries.
- **Legion™/CombiLibMaker™** for building virtual combinatorial libraries in cSLN format.
- **Molconn-Z™** for computing a wide range of topological indices based on molecular structure.
- **OptDesign®** for designing and editing combinatorial libraries.
- **ProtoPlex™** for multiplexing and normalizing protonation states and tautomeric forms of structures.
- **Selector™** for characterizing and sampling compound libraries.



WWW.TRIPOS.COM				CONTACT_US@TRIPOS.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 0065 Printed in USA All trademarks are the property of their respective owners.						