Query Guide

Organic chemistry grew rapidly during the 20th century, as evidenced by the increase in the number of known organic compounds from approximately 12,000 in 1880 to one million in 1960. Today, this number has grown over tens of millions. The big bang of information technology starting at the end of the last century brought about revolutionary changes in chemical data handling methods. The enormous rise in the number of known chemical structures has been multiplied by the new virtual chemistry methods which erased barriers limiting chemists' imaginations. Applying fast and intelligent structure searching software programs has become vital for the competing pharmaceutical and biotech industries.

The molecule structures found in the examples of this document are drawn using MarvinSketch, the integrated structure drawing tool of JChem.

- Full structure search, substructure search
- Other search types
- Similarity search
- Searching in the database

- Tautomer search
- Vague bond search levels
- Checking sp-hybridization state
- Search Options Guide

- Aromatization
- Standardization in the database

- Hit display-coloring

- Query atoms
  - Atom lists, not lists
  - Generic query atoms
  - Atom properties
  - Isotopes, charges, radicals
  - Link nodes (link atoms)
  - Repeating units
  - Group matching of undefined R atoms
  - SMARTS atoms
  - Pseudo atoms
  - Homology groups
  - Lone pairs
- Querying bonds
  - Generic bonds
  - Note on aromatic bonds
  - Chain/ring bond attributes
  - SMARTS bonds
  - Coordinate bonds
  - Position variation bonds
- Components
  - Smarts component level grouping
  - Component, Mixture and Formulation brackets
  - Other component features
- Explicit hydrogens
- Searching with empty structures
- Filtering expressions
- Attached data

- Differences in matching Daylight and MDL formats
  - ANY and not list atoms
  - H query property
  - Double bond stereo matching mode
- SMARTS feature matrix
  - Supported SMARTS features
  - NOT supported SMARTS features
- Molfile(ISIS) query feature matrix
  - Supported mol file (ISIS) query features
  - NOT supported mol file(ISIS) features

- Tetrahedral stereochemistry
- Relative configuration of tetrahedral stereo centers
  - MDL Enhanced Stereo Representation
  - MDL old stereo representation (chiral flag)
- E/Z stereochemistry of double bonds
- Special types of E/Z stereoisomerism
  - Double bond stereo matching near aromatic rings
  - Cumulated system with odd numbers of double bonds
- Axial isomerism
- Syn-anti, endo-exo isomerism of bicyclic molecules
- Stereo models

- R-group query structures
  - Root structure
  - R-groups
  - Conditions
  - Remarks
- R-group decomposition
- Searching in Markush targets tables
  - Generic Markush features in Markush targets
  - Querying Markush targets
  - Markush structure reduction to the hit
  - Markush aromatization
  - Creation of R-group definitions from a structure file
- Reaction search
  - Component types
  - Atom maps
  - Reacting center bond query feature
  - Reacting center stereo query feature
  - Component identification
- Polymer search
  - Polymer representation
  - Cyclization and phase-shifting
  - End group matching
  - Further polymer types
  - Attached data
  - Polymer mixtures
  - Relation to MDL polymer search types
- Sophisticated formula search