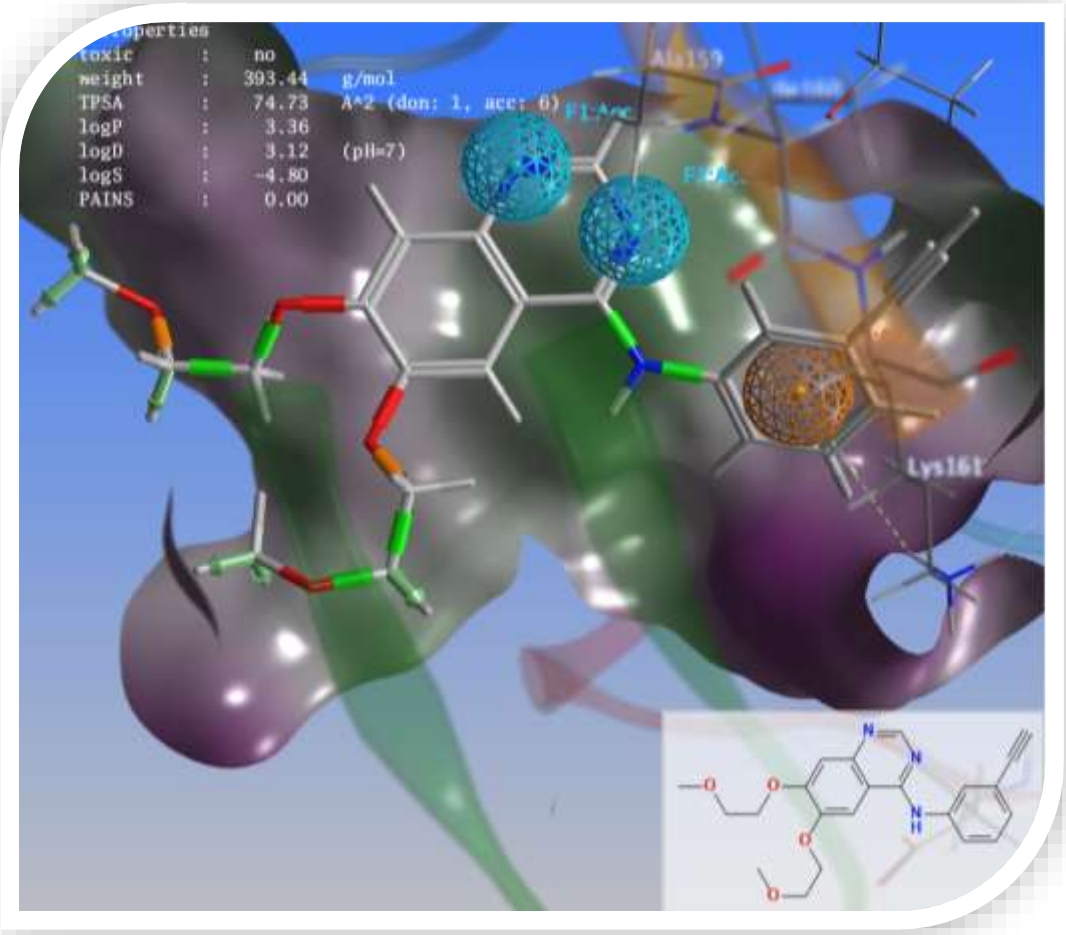
MOE

Molecular Operating Environment



- Molecular Modeling and Simulations
- Protein Modeling and Bioinformatics
- Structure/Fragment-Based Drug Design
- Pharmacophore Modeling
- Cheminformatics
- Development Environment

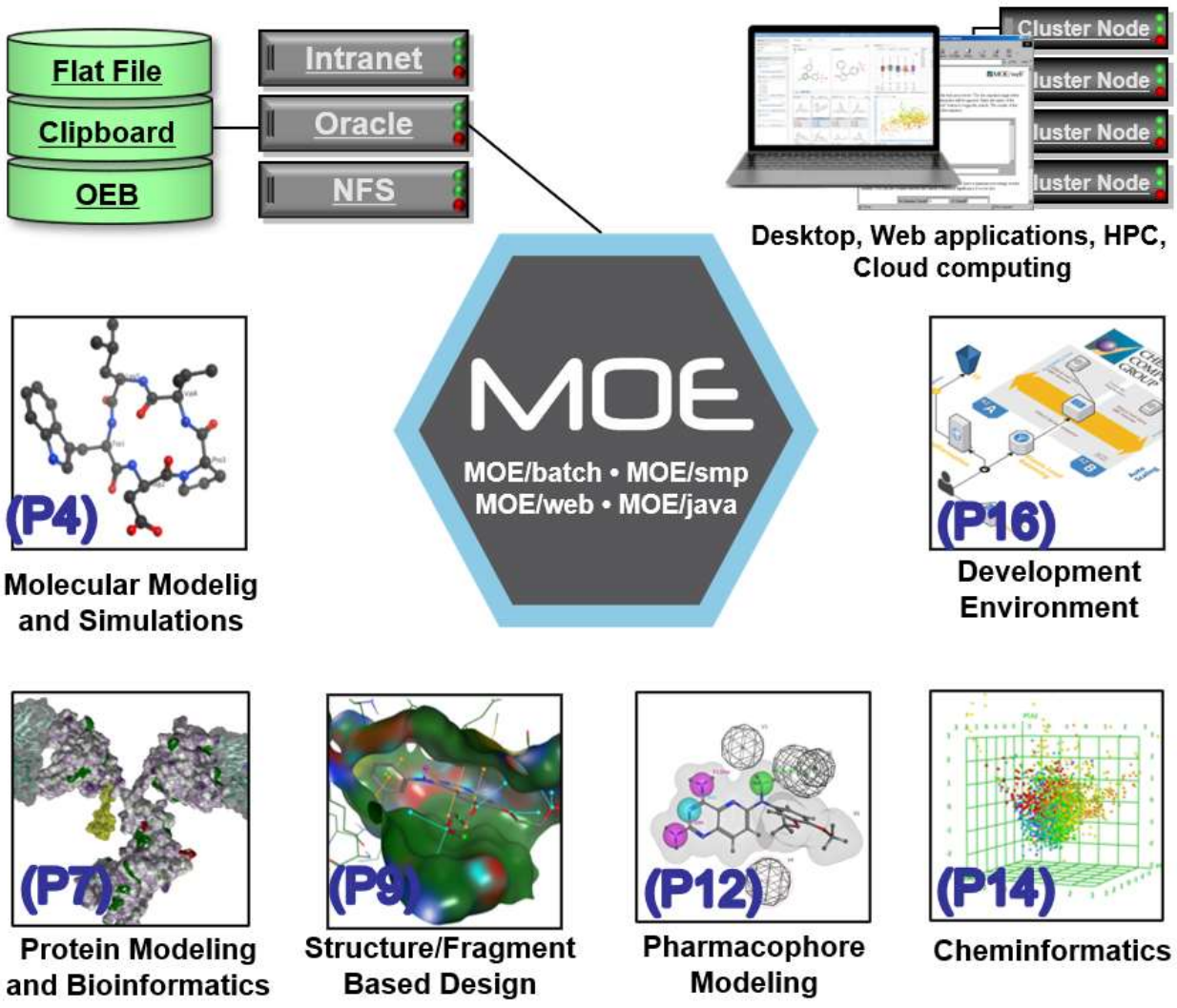
2025/04

Comprehensive Software System for Life Science

Molecular Operating Environment (MOE)

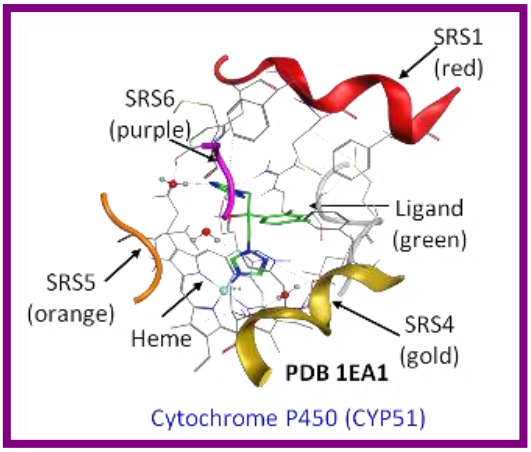
MOE, Molecular Operating Environment, developed by Chemical Computing Group ULC. (CCG), is a computational chemistry and molecular modelling tool that provides a comprehensive and interactive visualisation interface that allows multiple views of a molecular system and has many scientific applications.

MOE is equipped with a wide range of applications and rich data content to support the research activities of not only computational chemists, but also synthetic chemists, biologists and X-ray crystallographers. MOE is available in various usage modes such as graphical, command line, web application and workflow, and can be flexibly customised to create the most suitable molecular modelling environment for each user and research purpose.



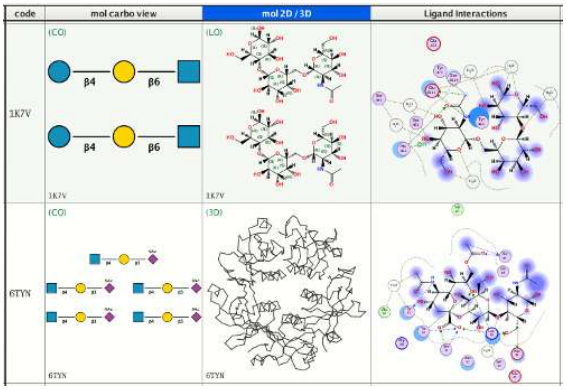
❑ Graphical User Interface (GUI)

The GUI of MOE provides the common working environment on various platforms (Microsoft Windows, macOS and Linux) from PCs to workstations.
Various analysis functions such as molecular simulation, chemoinformatics, protein modeling, receptor structure-based molecular design, and pharmacophore analysis are provided. Furthermore, more advanced molecular design is possible by linking these functions.



❑ Data Contents

Molecular structure databases organized by purpose are provided, including lead-like compounds, molecular fragments, and MOE Project Databases (family protein databases for antibodies, kinases, GPCRs, etc.). By utilizing these databases, users do not need to spend time collecting and organizing the databases themselves, and can use them for virtual screening and other applications immediately after installation.

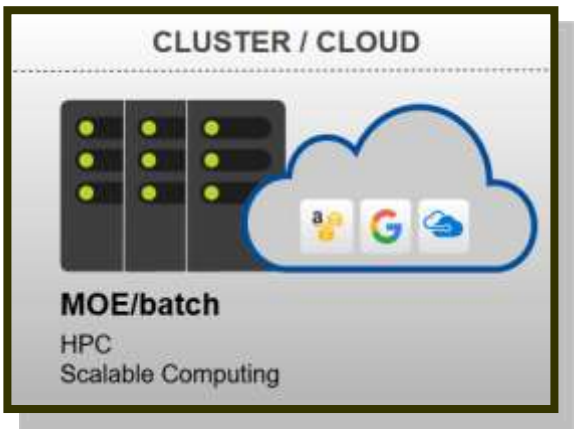


❑ Development Environment

MOE applications written in SVL are basically open-source. With the use of the development tools, *e.g.*, SVL Text Editor, users can easily refer to the contents of programs and modify the algorithms or parameters.
The programs and scripts written in SVL run on various hardware; Applications developed on Windows PC's can be run on Linux PC's, workstations or Mac's without any changes in codes, and *vice versa*. SVL codes are automatically loaded and compiled at the startup of MOE in a short time.

❑ High Performance Computing Framework

Large, parallelizable calculations can be submitted to external queueing systems directly from MOE through MOE's high performance computing (HPC) framework. Using this framework, MOE applications are able to provide a simple interface for submitting batch jobs to a job management system, and for monitoring their status. At the same time, the ability to run locally as a batch process or to save a job directory for calculation at a later time are also easily accessed.



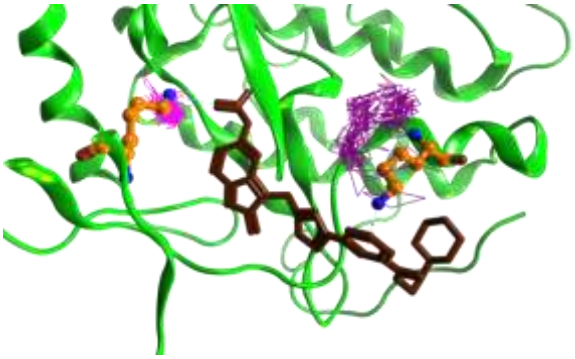
❑ MOE GPU-accelerated applications

The SVL GPU API provides a unified interface to integrate CUDA® and OpenCL™ kernel source code into SVL programs. It allows to send SVL vector data to the GPU and use it for the computation for some applications.

MOE's representation of organic chemical structures and flexible architecture provide a solid foundation for molecular modeling and computational chemistry. User-friendly molecule editors, wide variety of available file formats, choice of validated forcefields, powerful modeling applications, and customizability all make MOE the most flexible molecular modeling environment.

❑ Molecular Builders and Data Import/Export

Create or edit small molecules, peptides, proteins, carbohydrates, DNA/RNA and crystal structures using a variety of molecular builders. Import and export molecular structure information, and other data, in many standard file formats. Use SD pipeline command-line tools for compound filtering, descriptor calculation and structure depiction directly on SD files.



❑ Molecular Mechanics and Dynamics

Molecular dynamics simulations numerically solve the classical equations of motion for molecules and sample thermodynamic ensembles under specified thermodynamic conditions. Three large-scale energy minimization algorithms are available, and distance, angle and torsion angle constraints can be imposed. Atomistic NVE, NVT, NPT and NPH ensemble molecular dynamics simulations can be conducted, optionally including atomic wall restraints and tethers. To utilize HPC, NAMD and AMBER interface is also available. Free Energy can be calculated with Thermodynamic Integration(AMBER TI) using AMBER from the interface. AMBER TI interface generates the submission scripts to run AMBER TI simulations on a cluster or in the cloud directly or through a job management system. And support for Hydrogen Mass Repartitioning (HMR), which redistributes mass from heavy atoms to light atoms to slow down motion without changing the overall mass of the system, has been added to AMBER TI, Dynamics and Low Mode MD.

❑ An all-atom forcefield combining EHT and AMBER force field

An all-atom forcefield combining EHT and Amber10/19. Parameterized for proteins, nucleic acids and lipids using AMBER force field, and parameterized for small molecules using 2D Extended Hückel Theory (EHT). A bond angle and torsion parameterization methodology based on percentage s hybridization on directional bonds is used for all forcefield EHT calculations in MOE. This parameters have been found to lead to good agreement with QM geometry predictions.
The AMBER CMAP is saved to PRMTOP files, to correct backbone torsion psi-phi angles in proteins during Amber MD simulations. MMFF94 force field can be selected as a force field for small molecules.

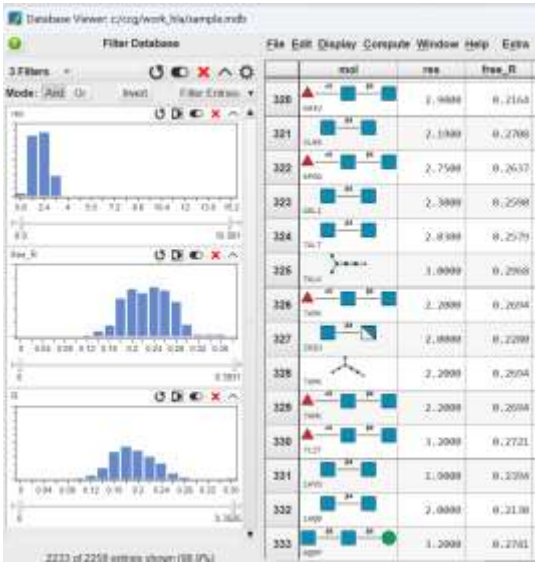
❑ Preparing 3D macromolecular structures

This application for preparing macromolecular structures for the subsequent simulation works is common structural problems in X-ray crystal structure data (lack of terminal caps, missing atoms, missing loops, *etc.*) are automatically diagnosed and presented in a list. Each diagnosed problem can be inspected and fixed from the graphical user interface.

❑ Spectral Analysis

The Spectral Analysis application supports computational studies of NMR and VCD spectra for small molecules and peptides. In addition, chemical shift, coupling constant and/or pairwise distance data can be used to determine the relative populations within a set of conformers, yielding the in-solution conformer distribution. This information can be used for compound analysis as well as for understanding the energetics of binding.

❑ Database Viewer



The Database Viewer has molecule display features and also supports the display of graphic objects. Carbohydrates are automatically detected and displayed using their SNFG symbols. Atom stereochemistry can be displayed with bond wedges and chirality labels.

Filter Database provides a quick interactive way to filter database entries by property. The distributions of database field values are displayed as histograms or lists of categorical values, in which a range of values can be specified. The data display is updated dynamically when any changes are made to the database contents.

❑ Conformational Analysis

The Conformational Search application contains three methods for generating conformations, the output of each of which is subjected to energy minimization:

Systematic Search: Systematic generation of rotatable bond dihedral angle combinations.

Stochastic Search: Random assignment of dihedral angle combinations, even for rings, and stereochemistry inversions, where permitted.

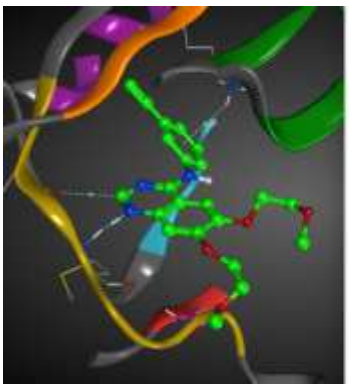
LowModeMD Search: A short molecular dynamics simulation using velocities with little kinetic energy on the high-frequency vibrational modes.

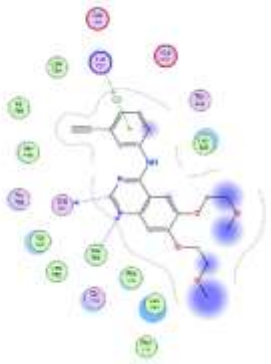
The LowModeMD Search method generates conformations using a short ~1 ps run of Molecular Dynamics (MD) at constant temperature followed by an all-atom energy minimization. The resulting conformation is saved to the output database provided it meets the energetic and geometric criteria. Note! LowModeMD Search is intended for large, possibly disconnected, and complex structures like macrocycles and protein loops but it is also very efficient for detailed small molecule analysis. However, chiral inversions as well as amide and double bond rotations will very rarely result.

❑ Quantum Chemical Calculations

MOE provides interfaces to popular *ab initio*, semi-empirical and density functional codes such as GAMESS, Gaussian, ADF, MOPAC v22.1.1.. Uses a graphical interface to configure a calculation and view results such as electron density contours, molecular orbital plots, energy level diagrams and other properties. Use MOE/smp to conduct the calculations on collections of compounds in parallel.

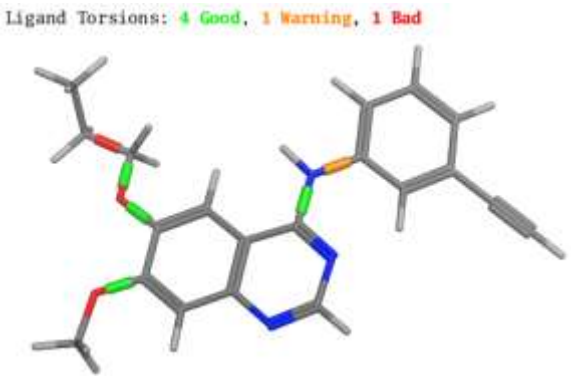
❑ Nonbonded Interaction Visualization





The molecular interactions are calculated based on the Extended Hückel Theory (EHT). Hydrogen bond contacts and van der Waals contacts (clashes) can be rendered in the MOE Window in a variety of styles. Separate controls are provided for hydrogen contacts and clashes.The Ligand Interactions application provides a means to visualize an active site of a complex in diagrammatic form. The content is presented in a concise and stylized manner, and summarizes a large amount of spatial data making use of mnemonic glyphs where appropriate. The panel can be used as a complement to 3D visualization, as the 2D and 3D representations can be displayed alongside one another. The content can also be exported in graphical or tabular form, as well as printed. Use the protein-ligand interaction 2D diagrams to easily identify

polar, hydrophobic, acidic and basic residues. Visualize solvent-exposed ligand atoms and residues in close contact with ligand atoms as well as sidechain and backbone acceptor and donor interactions.

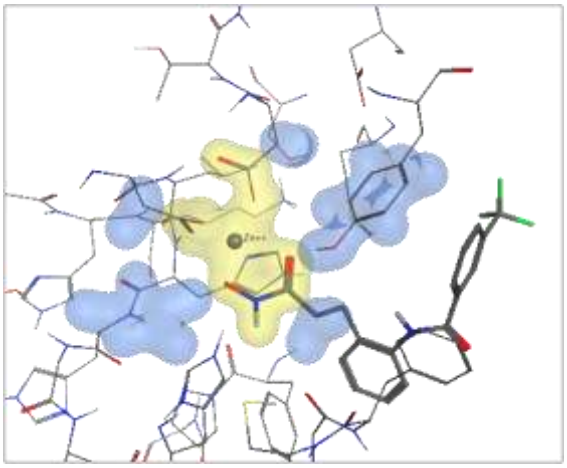


❑ Ligand Torsion Analyzer

This is an application for dynamically assessing the quality of the dihedral angles in the current system based on their consistency with CSD statistics.Bond sleeves with traffic light coloring are displayed in the MOE 3D window as quality indicators, along with a text summary. Additionally, a Ligand Torsion Analyzer color bar can be displayed in the MOE footer Dihedral Plot for the currently selected torsion.

❑ ONIOM (QM/MM)

Large systems, such as ligand/pocket complexes, can be treated by QM/MM methods. In this, a (small) region of interest (the inner layer) – for example, the locus of charge transfer or bond making/breaking, i.e. "where chemistry occurs" – is treated quantum-chemically, while the rest of the system (the outer layer) is modeled by a forcefield. In complex systems, a middle layer might be included that is treated quantum-chemically using a lower level of theory. Determining the complex energy or minimized geometry using this heterogeneous approach can gain a significant reduction in computation time over treating all atoms with QM.



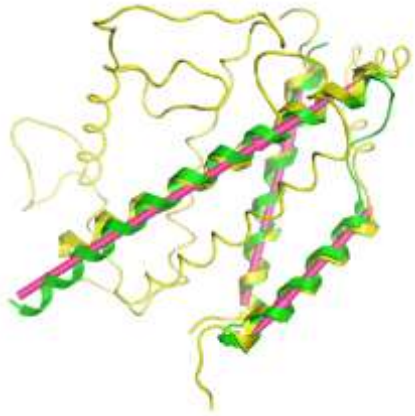
Applications for protein structure prediction are powerful and easy to use, both for experts and occasional users. Powerful homolog search, alignment algorithm and refinement methodology make high-quality sequence-to-structure predictions possible.

❑ Protein Structure and Family Databases

Protein Database: MOE provides a database of non-redundant protein domains from the Protein Data Bank. Entries consist of single protein domains. Nearby solvent, ions and ligands are also included. In MOE 2024, the database contains over 50,000 non-redundant domains obtained from the over 222,000 PDB entries containing protein molecule models in July 2024.

MOE Project Databases: A number of protein family databases — including antibodies, kinases and GPCRs — are provided. Each is superposed into a common frame of reference based on the core region of the reference files. The databases can all be used with the Project Search panel as the repository for all relevant information used in this application.

❑ Remote Homology and Fold Identification

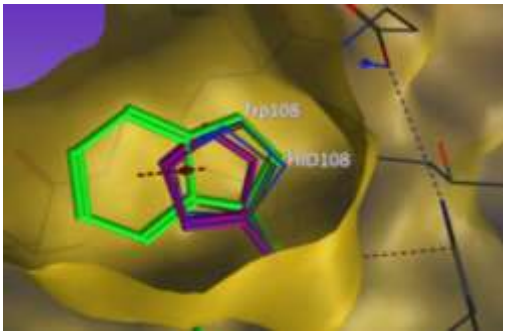


Search the Structural Family Database to identify protein families relevant to structure prediction. The search uses a FASTA-type local alignment followed by a family membership test based upon full multiple alignments and Z-score significance testing. Folds of even distantly related homologs can be reliably identified with few false positives (unlike pairwise searches). Use the Domain Motif Search to identify structural homologs with low sequence identity.

❑ Loop/Linker Modeler and Browser

This application in MOE is used to find and select loop conformations. Interactive searching and browsing of potential loop candidates make it possible to bring expert knowledge of a system to bear when selecting loop candidates for further optimization. Ready access to useful tools such as affinity and electron density scoring make it possible to use additional criteria when selecting a loop conformation.

❑ Build Rotamer Library



In MOE, the Rotamer Library is used to define the allowed set of amino acids for multiple applications, such as Homology Modeler, Protein Design, Rotamer Explorer, and Protein Builder. As well as defining the set of nucleic acids for the DNA/RNA Builder. The Build Rotamer Library application aids in the creation of new non-natural amino/nucleic acids for use in these applications.

❑ Mutation and Rotamer Exploration

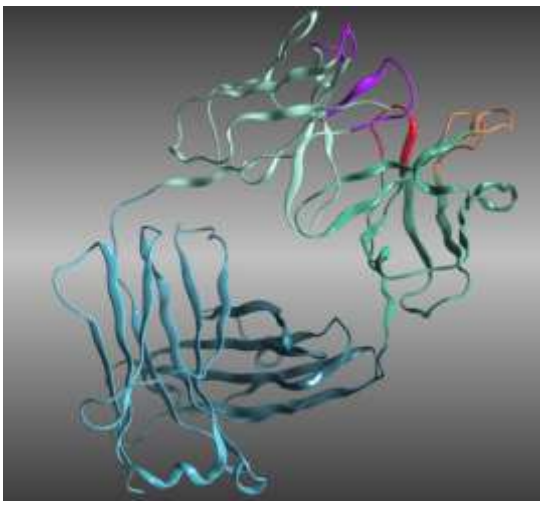
Discover accessible natural/non-natural amino acid side chain conformations with MOE's Rotamer Explorer. This predicts the structure of amino acid mutations in a 3D protein structure. Rank candidate rotamers using an energy-based scoring function and visually analyze them using MOE's graphical interface.

❑ Protein Geometry

The Protein Geometry application provides a variety of stereochemical measurements for inspection of the structural quality in a given protein:

1. Backbone bond lengths, angles and dihedrals.
2. Ramachandran $\phi-\psi$ dihedral plots.
3. Sidechain rotamer quality.
4. Nonbonded contact quality.

❑ Homology Modeling and Antibody Modeling



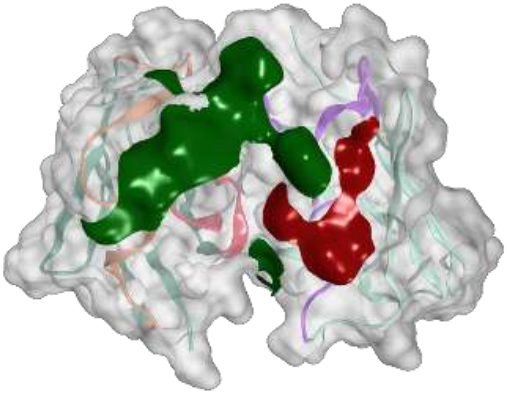
The Homology Model application is to build and refine all-atom 3D models for protein sequences, based on homologous template protein structures. This application supports MOE/cluster (MOE/smp) cluster computing. If run under a MOE/smp session, Homology Model will run in parallel.

The MOE Antibody Homology Modeling application offers a collection of procedures, as well as an interface, to meet the demands of effective antibody homology modeling. A knowledge-based approach is applied with an underlying database of antibody structures currently in the Protein Data Bank (PDB), clustered by class, species, subclass and framework sequence identity. This database may be enriched with in-house antibody structures and can be continually updated and re-clustered. MOE implements a variety of numbering schemes: Kabat, Chothia, IMGT, and CCG's original assign residue UIDs and define CDR boundaries.

❑ Protein Engineering/Protein Design

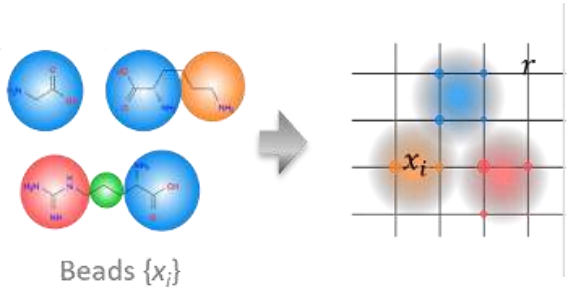
The Protein Design application permits unlimited simultaneous residue mutations of target structures, and then calculates the properties of the resultant mutant, with the purpose of examining how mutations of a protein modulate its protein properties. A streamlined workflow for protein design is added and enhances MOE's offering for computer aided biologics design (Alanine scanning, Cysteine scanning, Resistance scanning etc.). An ensemble of conformations can be automatically generated and scored for thermostability or binding affinity. The results of the calculations are written to a database and can be browsed. The Database Browser is Protein Design aware.

❑ Protein Properties



The Ensemble Protein Properties application in MOE provides a common interface for all calculations of properties and descriptors specialized for proteins with ensemble sampling and pH-dependence built-in as a central feature of the calculation. a comprehensive set of sequence- and structure-based physical properties such as pI, zeta potential, mobility, dipole moment, *etc.* for building QSPR model or helping rationalize protein solubility, solution viscosity, stability, and aggregation. Use the predicted properties in conjunction with preliminary experimental data to rationalize stability and aggregation at a given pH. Calculate properties for an ensemble of mutants to identify and predict physical property trends on a relative scale.

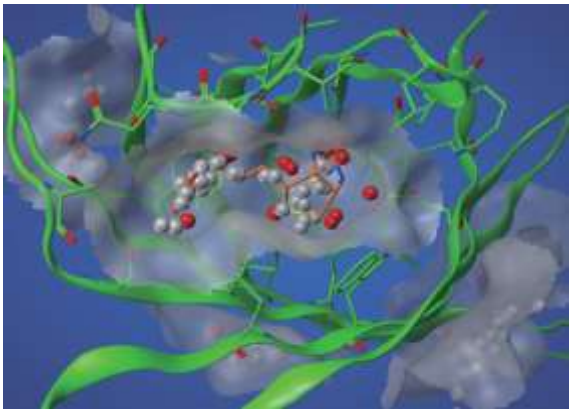
❑ Protein-Protein Docking



In Protein-Protein Docking, a coarse-grained bead model is used to reduce the computational search space. In the coarse-grained minimization step, the calculation of pairwise bead interaction energies, which represents a significant amount of the processing time, has been reimplemented for GPU. This improvement have resulted in a 6–10x speedup in cases dominated by bead refinement, and a 3x speedup in structures that are faster to minimize.

Macromolecular crystallographic data, when available, can be a valuable source of information for active ligand discovery. MOE provides several applications for visualising and understanding details of receptor active sites and receptor-ligand interactions. These applications can be used to suggest improvements to known ligands or to screen ligand databases for ligand candidates.

❑ Active Site Detection



Detect candidate protein-ligand and protein-protein binding sites using a fast geometric algorithm based on Edelsbrunner's Alpha Shapes. Each site on a macromolecular structure is ranked according to its PLB index*. Visualize individual sites or populate them with "dummy atoms" for docking calculations or as starting points for *de novo* ligand design efforts.

*PLB index: Soga, S., Shirai, H., Kobori, M., Hirayama, N.; Use of Amino Acid Composition to Predict Ligand-Binding Sites; *J. Chem. Inf. Model.* **47** (2007) 400-406.

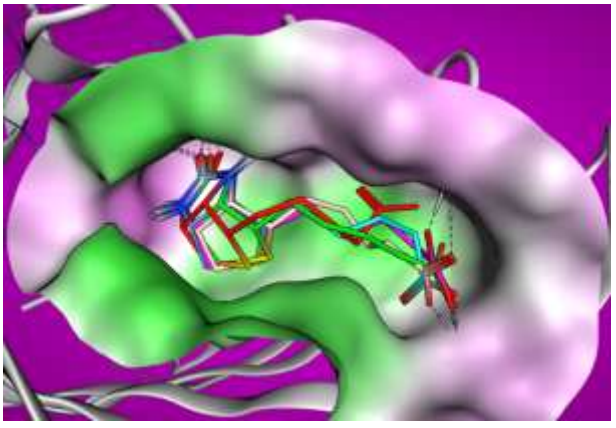
❑ Electrostatic Maps

Determine favorable locations of neutral, positive and negative features in an active site. The electrostatic maps are calculated by solving the non-linear Poisson-Boltzmann equation for receptor atoms and pseudo-ionic species. The isocontour levels of the generated predicted map are expressed in kcal/mol. The advantage of the present method is that fully screened electrostatic potentials are used and the domination of the potential by ionic groups is avoided.

❑ Probabilistic Contact Potentials

Visualize and understand directional details of hydrophobic or hydrophilic contact preferences of a receptor or ligand using probabilistic contact maps. CCG has developed a suite of analytical probability distributions that correspond very well to inter-atomic distance, angle and out-of-plane angle histograms derived from a large collection of crystallographic structures. These distributions are then used to form a composite "preference map" for a given macromolecular structure.

❑ Docking simulation



MOE's Dock application searches for favorable binding modes between small- to medium-sized ligands and a not-too-flexible macromolecular target, which is usually a protein. For each ligand, a number of placements called poses are generated and scored. If desired, the poses can be covalently bound to a receptor side-chain atom. The score can be calculated as either free energy of binding including, among other contributions, solvation, entropy, and enthalpy terms based on polar interaction energies (including metal ligation) or qualitative shaped-based numerical value. Poses can be optionally constrained to fit a pharmacophore query, and final filtering by pharmacophore query is also possible. The atoms of the active site can be permitted to move: sidechains can be tethered, but backbone atoms are always fixed. The final highest-scoring poses, along with their scores and conformation energies, are written to a database where they are ready for further analysis.

❑ GBVI/WSA dG scoring function

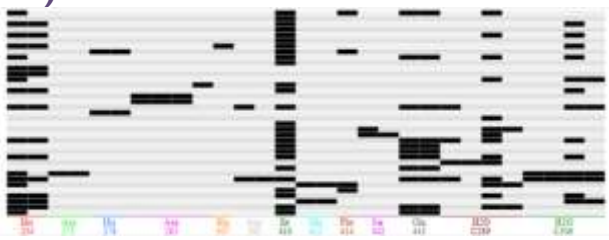
The GBVI/WSA dG scoring function is based on the forcefield, GB/VI implicit solvent model and a weighted surface area term. There are only three fit parameters in the scoring function each of which is related to

Linear Interaction Energy principles and were fit to experimental binding free energies.

❑ Multi-Fragment Search

Multi-Fragment Search is an ensemble-based methodology for mapping the preferred locations of specific chemical groups in a receptor structure. An active site of a macromolecular structure is populated with a large number of chemical fragments, which are subjected to an energy minimization protocol. The resulting group locations are clustered, scored (including solvation effects) and output to a database for subsequent visualization and analysis.

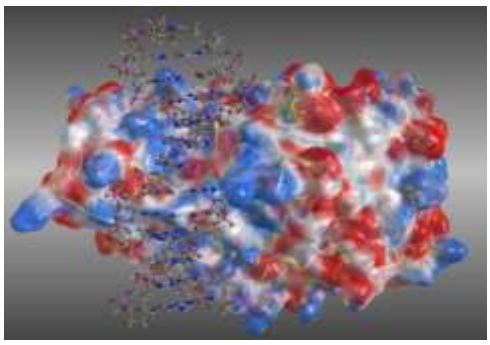
❑ Protein-Ligand Interaction Fingerprints (PLIF)



This tool is a method for summarizing the interactions between ligands and proteins using a fingerprint scheme. There are two categories of interaction in which a residue may participate: potential (energy-based) contacts and surface (patch) contacts. For potential contacts, the value is that of the strongest interaction between any pair of atoms in the residue and ligand, whereas for surface contacts the value is the total contact area of each type between a residue and the other molecule.

❑ QuickPrep

QuickPrep is a streamlined interface for interactive 3D ligand optimization in the active site. A dedicated button bar provides commands for structure preparation, active site analysis, molecular property/binding affinity calculations and ligand modification and optimization in the active site. QuickPrep can be used by both computational and medicinal chemists.



❑ Molecular Surface and Maps

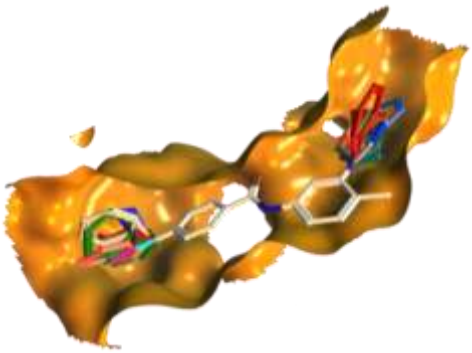
Molecular Surface and Maps is an integrated application for active site analysis. Create molecular surfaces, predict contact preferences and calculate electrostatic maps. Color molecular surfaces by choosing from a variety of schemes such as temperature factor, pocket, lipophilicity, and electrostatic potential.

❑ Scaffold Replacement

Link annotations denote substitution points on a candidate scaffold molecule and the locations of potential R-group substituents. Search standard 3D conformational databases or special scaffold and linker databases to find novel chemical scaffolds that preserve substituent's geometry. Add additional pharmacophore features to preserve known scaffold interactions or volume constraints to satisfy shape requirements.

❑ Combinatorial Builder

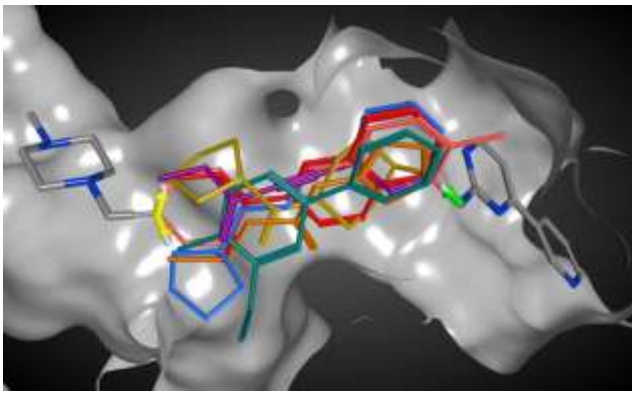
Multiple attachment points and 3D R-group libraries can be systematically enumerated around a scaffold, refined and filtered in a receptor pocket.



❑ BREED: Ligand Hybridization

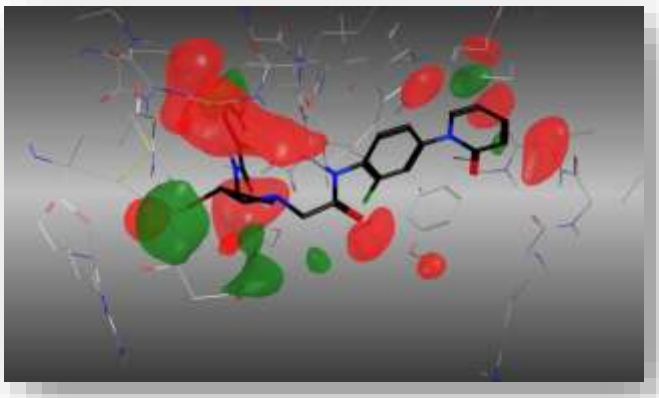
BREED is a method for the generation of novel inhibitors from structures of known ligands bound to a common target. New ligands are generated by combining parts of existing ligands. The user can specify QuaSAR descriptor, Model file and/or Pharmacophore query filters to limit the results. If a receptor is present, structures can be refined in the binding cleft as with docking calculations. Thus, BREED can be used as part of a ligand-based or structure-based discovery methodology.

❑ MedChem Transformations



MedChem Transformations is an approach used to discover novel chemical structures by applying a set of transformation rules to existing ligands. Typical transformations might exchange functional groups, or alter all or part of individual rings while preserving the rest of ligand. These can be applied iteratively, resulting in cumulative changes to the starting structures. If a receptor is available, Forcefield Refinement can be used to generate bound poses along with various docking scores.

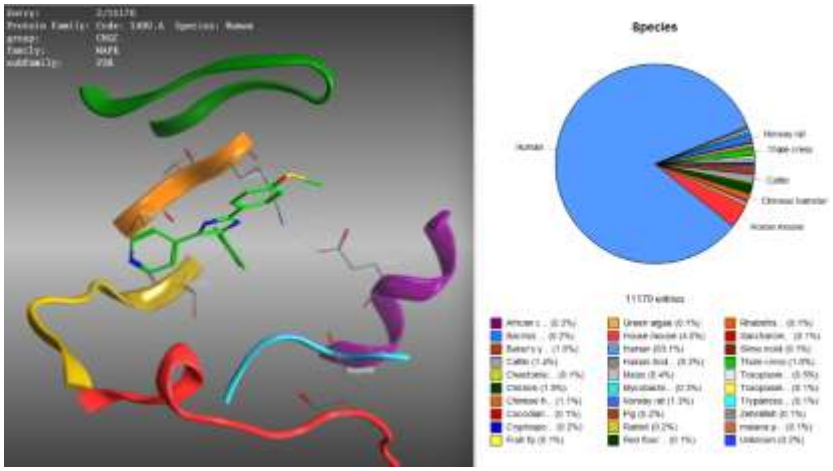
❑ Solvent Analysis with 3D-RISM



This application uses the 3D-RISM (three-dimensional reference interaction site model) method to analyze the role of solvent in macromolecular systems. The application computes a time-averaged distribution of water H and O densities, along with free-energy maps for analyzing solvent stability and solvation contributions to binding free-energy. A unique feature of 3D-RISM is its ability to include various concentrations of salt or hydrophobic molecules as part of the solvent, thereby producing density maps for ions and hydrophobic groups as well as for water. The Solvent Analysis tool can be applied to various water-related tasks such as water placement in low- resolution X-ray protein structures, estimating solvent effects in protein-ligand binding, and designing ligands to exploit solvation effects such as water displacement.

❑ MOE project applications

The MOE Project is a highly configurable meta-protocol for data organization. It takes a set of standard format (PDB, mmCIF, SDF, MTZ, etc.) data files as input, processes this data, and outputs a single .mdb file. Structures can be aligned and superposed, optionally with electron density around the active site and/or ligand topology correction. Various forms of Structure Preparation can also be carried out, and associated data assembled or calculated, possibly using SOAP, HTTP or a number of other common access methods.
The MOE Project application is used to manage this information and keep it up-to-date by assembling and organizing the data into a single MOE database (MDB) file. The MOE Project framework supports aligning and superposing all structures into a common reference frame, performing structure preparation work, and more. By centralizing data into a single repository, MOE Project helps to ensure data integrity and consistency. The use of a single integrated structure also facilitates data dissemination and maintenance..

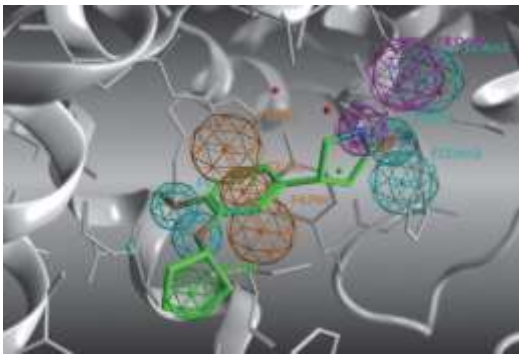


Pharmacophore Modeling

MOE's pharmacophore modeling is a powerful means to generate and use 3D geometric information to search for novel active compounds, particularly when no receptor geometry is available. Pharmacophore methods use a generalized ligand representation and geometric constraints to bypass the structural or chemical class bias of 2D methods.

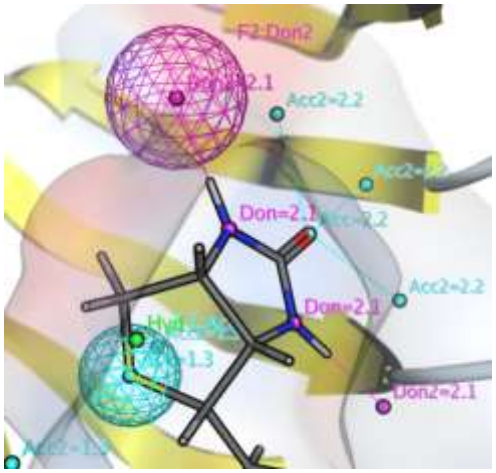
❑ Pharmacophore Annotation Scheme

A pharmacophore annotation scheme, or more simply a scheme, is an automatic procedure to assign pharmacophore annotation points (such as H-bond donor, H-bond acceptor, hydrophobe, etc.) to a 3D conformation of a molecule. The pharmacophore annotation points are not part of the molecule proper, even though some of the annotation points may be located directly on top of a particular atom of the molecule. MOE has a variety of built-in schemes which provide different types of annotations and use different policies to locate the annotations about a particular conformation.



❑ Extended Hückel Theory for Pharmacophore Discovery (EHT scheme)

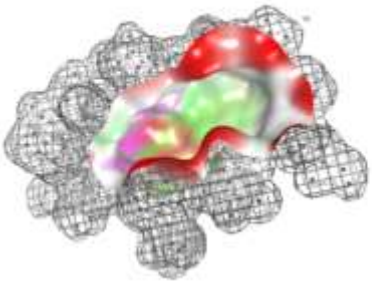
The EHT scheme provides a more nuanced means of investigating non-bonded interactions than the Unified or other traditional alternatives. These older methods use a SMARTS-pattern based atom/environment mapping approach similar to that used by classical force fields. The EHT scheme replaces this with annotation points derived from 2D EHT calculation which provides both the location and interaction type along with a direct measure of the H-bond interaction energies (strengths) for each interacting atom pair. This approach provides a better representation of the specific chemical environment of each ligand and can represent non-standard interactions not amenable to rule-based methods.



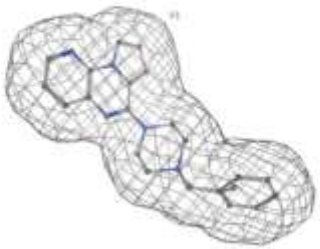
❑ Volume and Shape Filters

Volume constraints are spatial constraints imposed on particular atoms. These constraints serve both to confine hits to spatial regions in which steric clashes with the receptor are avoided as well as to ensure that favorable contact regions will be populated with appropriate atoms.

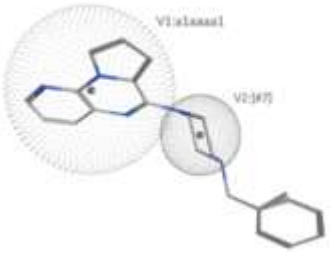
Excluded Volume:



Ligand Shape:



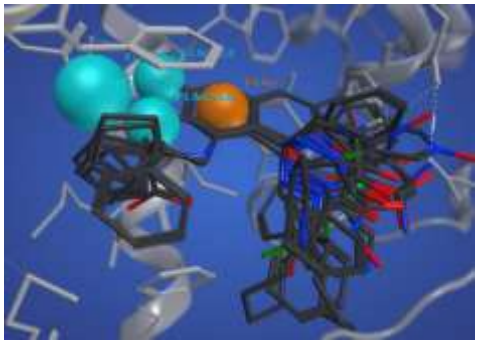
Occupied Volume:



Pharmacophore Query Editor

Use an interactive editor to construct a 3D query from a molecular alignment or macromolecular structure. Use the query to filter a conformational database in an effort to determine candidate active compounds that satisfy the pharmacophore model. Refine the query to contain locations of Boolean expressions of pharmacophore features as well as restrictions on shape by using union-of-spheres for included, excluded and exterior volumes each containing an optional SMARTS chemical pattern.

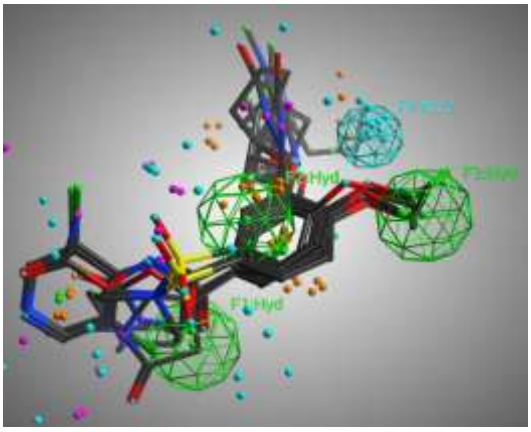
❑ Pharmacophore Search



Rapidly search a conformational database for compound conformations that satisfy a pharmacophore query. Search multiple databases, a sub-range of molecules or a database of docked compounds. Output data consists of molecules that satisfy the 3D pharmacophore query (either all conformations or just the conformations that satisfy the query). Partial matches, the output of all symmetric matches and the specification of essential features are supported. Selection expressions can be used to limit the pharmacophore search to a subset of atoms of the system. For example, it is possible to quickly find protein-protein docking poses with specific interactions.

❑ High Throughput Conformational Analysis

Use MOE's High Throughput Conformational Search methodology to construct conformation databases for virtual screening. Conformational databases are constructed using a parallelized fragment-based approach. Molecules are subdivided into overlapping fragments each of which is subjected to a rigorous stochastic search. The fragment conformations are rapidly assembled by superposing the overlapping atoms. A database of fragments is maintained (and augmented as the search proceeds) making conformation generation of combinatorial libraries very fast.

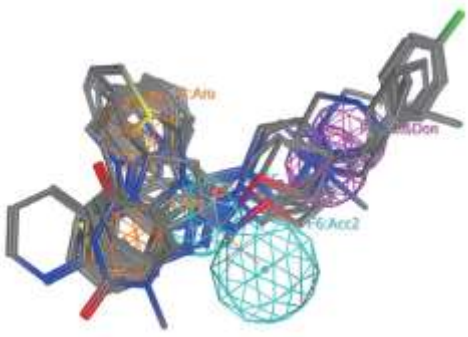


❑ Pharmacophore Elucidation

Generate pharmacophore queries from a collection of input compounds (possibly with activity data) by considering all possible discrete geometries and all possible combinations of feature query expressions. A build-up strategy is used to avoid the combinatorial explosion. Enforce limits on feature counts and add custom query expressions. Each pharmacophore query is scored based on known active compound coverage, statistical activity enrichment and atomic overlap of matching conformations. The resulting scores and induced molecular alignments of high-scoring pharmacophore queries are written to a MOE molecular database for further analysis.

❑ Pharmacophore Consensus

This application suggests possible pharmacophore queries based on a set of aligned active compounds. A consensus calculation requires a set of aligned input molecules, a tolerance radius, the consensus score threshold, and the consensus score mode. Pharmacophore consensus is particularly useful when starting from a few highly active compounds.



MOE provides a suite of applications for manipulating and analyzing large collections of compounds, building property models, consensus models, and SD pipeline command-line tools.

❑ MOEsaic

MOEsaic is a web-based application for analyzing SAR data, visualizing trends, exploring new virtual leads, and documenting results. Its streamlined interface features interactive MMP analysis and R-group profiling for fast assessment of property cliffs and SAR transferability of fragments, an integrated sketcher for designing novel structures, and models for property prediction.

Matched Molecular Pair Search

A Matched Molecular Pair (MMP) search algorithm finds chemical structures that differ in some substructures. MMP Search is a special kind of substructure search that defines the query by the portion of the molecule that is changing, as opposed to the constant portion query of the traditional substructure search. MMPs are important in SAR/SPR analysis because they can detect activity cliffs: small changes in structure that lead to large changes in activity or selectivity. Alternatively, changes in structure that preserve activity or selectivity are potential bioisosteres.

Properties Calculation

Perform automatic MMP fragmentation and calculation of descriptors HBA, HBD, tPSA, sLogP, MW. Real-time filtering/formatting and plotting of simple rule-of-thumb properties (Oral drugs / Sweet Spot / Golden Triangle) with conditional formatting. MOEsaic is a web-based application for analyzing SAR data, visualizing trends, exploring new virtual leads, and documenting results. Its streamlined interface features interactive MMP analysis and R-group profiling for fast assessment of property cliffs and SAR transferability of fragments, an integrated sketcher for designing novel structures, and models for property prediction.

❑ Molecular Descriptors

Calculate over 400 molecular descriptors including topological indices, structural keys, E-state indices, physical properties (such as *logP*, molecular weight, and molar refractivity), topological polar surface area and CCG's VSA descriptors with wide applicability to both biological activity and ADME property prediction. Use descriptors for classification, clustering, filtering, and predictive model construction. Add custom descriptors using MOE's built-in Scientific Vector Language.

❑ HTS-Binary QSAR

MOE's patented Binary QSAR methodology is ideal for building pass/fail models from high error content data and standard molecular descriptors. Use the resulting probabilistic models (based on Bayesian statistical inference) as a biasing agent in the design of focused combinatorial libraries.

❑ The Structure-Activity Report (SAReport)

SAReport contains sophisticated analysis methods to help scientists identify important groups and make more effective choices for synthesis. As experimental data builds up on discovery projects, it is often difficult for scientists and managers to keep track of the information. SAReport puts the information into a consistent context for better decision making.

❑ QSAR/QSPR Predictive Modeling

Build QSAR/QSPR models using linear, probabilistic and decision-tree methodologies. CCG's unique Binary QSAR methodology is ideal for building pass/fail models from high error content data. Linear models include PCR and PLS methodologies and can support biological activity or ADME assessments.

❑ Similarity, Diversity, and Fingerprints

Perform similarity searching and diverse subset selection using Descriptor, Conformation, and Molecular Fingerprint methodologies. Choose between a number of fingerprint systems including 2-, 3-, and 4-point pharmacophore fingerprints in 2D or 3D and MACCS key fingerprints.

❑ Combinatorial Library Design

Enumerate both reaction-based or R-group-based virtual libraries either in 2D or 3D. Filter the virtual compounds by properties or pharmacophores. Build combinatorial libraries by combining scaffold and R-group databases or with MOE's reaction-based library generation method. Reactions can be chosen from a list or sketched in standard sketchers and applied to enumerate libraries of compounds. The default reagent database contains over 3,000 reagents curated from commercial vendors, however, custom reagent databases can be specified.

❑ RECAP Analysis and Synthesis

Analyze large collections of compounds to produce fragments resulting from retrosynthetic rules. Use the resulting fragments in a *de novo* synthesis methodology to produce novel chemical structures that have an increased likelihood of synthetic accessibility. Specify heavy atom mean and variance to control the size distribution on the randomly generated structures. Apply leadlike/druglike, QSAR/QSPR predictive models or 3D pharmacophore filters for *de novo* virtual screening applications.

❑ rsynth : Synthetic Score

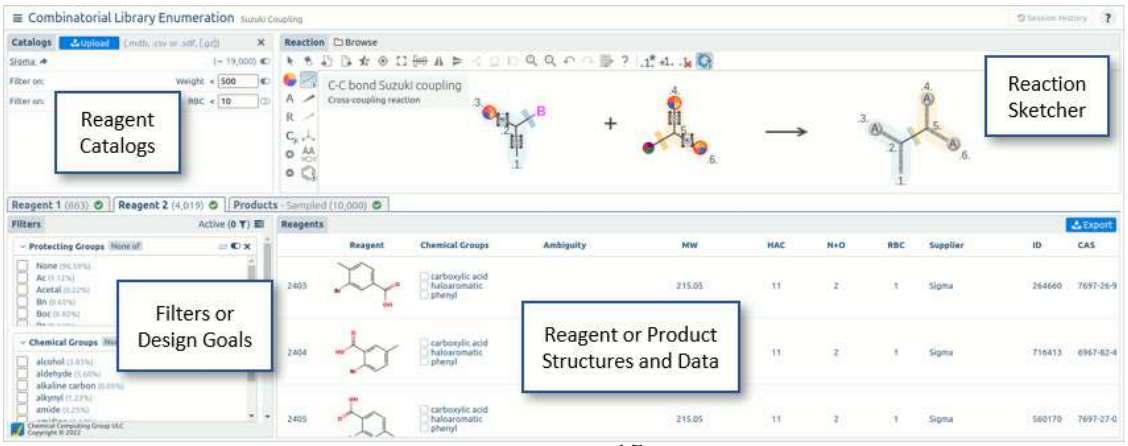
The synthetic feasibility score, which is the fraction of the atoms of each new structure that ultimately appear in a retrosynthetic fragment found in the starting materials database. A value of one indicates that the molecule is very likely synthesizable.

❑ SD Pipeline Command Line Tools

Operate directly on SD files for structure depiction, acid/base protonation state, database filtering, sorting and descriptor calculations. Remove records that do not satisfy a series of filters (*e.g.* lead-like, reactive groups, drug-like, *etc.*), sort records and remove duplicate entries from SD files. Calculate descriptors and write the output to SD or ASCII formats.

❑ Combinatorial Library Enumeration.

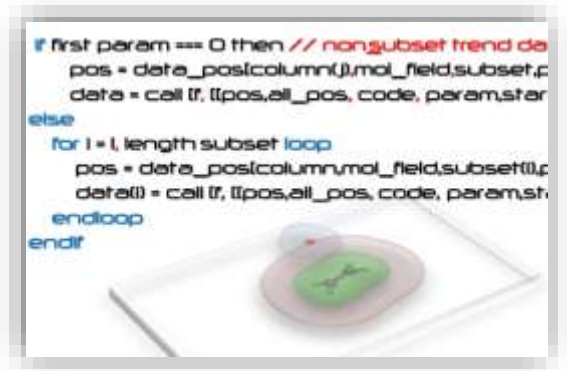
The Combinatorial Library Enumeration (CLE) web application is a reaction-based library enumerator. It uses a novel probabilities-based virtual compound enumeration algorithm that allows for browsing and tuning the fully-enumerated library interactively without the library actually having to be fully-enumerated.



In addition to the suite of graphical applications, MOE contains a tool for making changes to the existing applications or creating new applications. With MOE, expert modelers, application developers, and occasional users can benefit from sharing the same software system. Methodology written by application developers can be validated by expert modelers and then deployed to occasional users using either the MOE graphical interface or a Web interface.

❑ Scientific Vector Language (SVL)

SVL is the built-in command language, scripting language and application development language of MOE. SVL is a "chemistry aware" vectorized computer programming language with over 1,000 specific functions for analyzing and manipulating chemical structures and related molecular objects. SVL is a concise, high-level language and SVL programs are typically 10 times smaller than equivalent programs written in C or FORTRAN. SVL source code is compiled to a "byte code" representation, which is then interpreted by the base run-time environment making SVL programs inherently portable across different computer hardware and operating systems.



❑ Background Computing

MOE/batch is an adaptation of the MOE run-time environment intended for batch or background calculations that do not require a graphical interface. All non-graphical MOE functionality is accessible. MOE/web technology distributed with MOE is used to create Web Browser interfaces to MOE applications. SVL scripts and programs can be used to automate workflow, and are convenient for performing repetitive tasks or for modifying options.

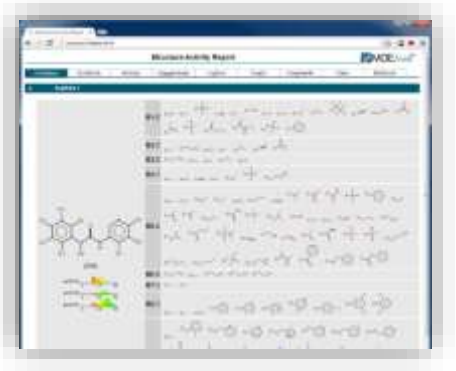
❑ Cloud Computing

Cloud service providers such as Amazon Web Services (AWS) provide rapidly scalable on-demand compute and storage capacity allowing large scale, parallelizable computations to be carried out quickly and in real-time. In addition to AWS, there are numerous alternative cloud computing providers. You can select a provider which meets the needs of your organization.

❑ HPC Job Management Framework

Large, parallelizable calculations can be submitted to external queueing systems directly from MOE through MOE's high performance computing (HPC) framework. Using this framework, MOE applications are able to provide a simple interface for submitting batch jobs to a job management system, and for monitoring their status. At the same time, the ability to run locally as a batch process or to save a job directory for calculation at a later time are also easily accessed. Furthermore, calculation using GPU is also possible. The new gpu resource specification keyword is used to specify GPU usage on jobs submitted to job management systems via MOE's HPC framework.

❑ Cluster Computing



A standard part of MOE is the MOE/smp distributed computing technology. With MOE/smp, multiple cooperating computers can be used to perform large-scale calculations. A heterogeneous collection of computers including laptops, workstations and multi-processor clusters, all running different operating systems, can be easily harnessed together in a single MOE session. The MOE/smp programming model in SVL makes it easy to parallelize SVL applications.

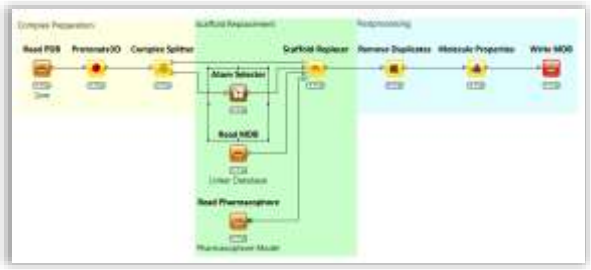
❑ MOE/web: Web Browser and SOAP Application Framework

MOE/web is an application environment that allows for large scale deployment of MOE applications or custom SVL programs to occasional users such as medicinal chemists via a simple web interface and/or to SOAP applications. MOE/web is distributed with MOE and customized applications can be incorporated.

❑ URL (HTTP/FTP) & TCP-IP Sockets

Access data from the internet via FTP, download files from web servers via HTTP and directly access third party database servers such as Oracle and DB2 via JDBC. MOE contains built-in applications to download new entries from the RCSB Protein Data Bank and a relational database browser. It is also possible to write SVL programs that interface with existing applications and libraries.

❑ MOE Extensions for KNIME



KNIME is a free data analytics and reporting platform. Through the MOE Extensions for KNIME, you can extend state-of-the-art Life Sciences Modeling into the KNIME platform. The largest and most complete plugin containing more than 180 components is based on MOE.

License Style

MOE is offered under a floating license. It is possible to be used freely on any coexisting computers and workstations under the network environment with which the license of MOE is common to all the hardware supported by MOE.
The license of MOE is managed with network floating license with the token as a unit, and can be used from any machine within the limitation of the license. The number of tokens MOE uses at a time is different depending on the mode:

Program (Mode of usage)	Number of necessary tokens
MOE (Graphics mode)	3 tokens
MOE/batch (Non-graphics mode)	1 token
MOE/web (Webservice mode)	1 token and more
MOE/smp (Parallel computation mode)	1 token x number of processes (When a non-graphics mode is used)

MOE has no limit of the number of install machines.

Support Platform

Architecture	OS	Graphics
Intel	Windows 10/11 64bit	OpenGL*
	Linux 64bit (glibc version 2.7 or later)	OpenGL* / X11
	macOS 10.12 and newer	OpenGL* / X11
Apple Silicon (M1 and higher)	macOS 11 and newer, FlexNet v11.18 required	OpenGL * / X11

As for the supported platform, please inquire of MOLSIS Inc.

*To use full graphics effects and anaglyph stereo, OpenGL 2.1 or later with GLSL shader is needed.

MOLSIS Inc.

3-19-9, Hatchobori, Chuo-ku, Tokyo
104-0032, Japan

URL: https://www.molsis.co.jp/en/

support@molsis.co.jp

Chemical Computing Group

1010 Sherbrooke St. W, Suite 910 Montreal,
Quebec, Canada H3A 2R7

http://www.chemcomp.com/