NAOMI ChemBio Suite **25**

an academic software collection

Prof. Dr. Matthias Rarey

University of Hamburg

ZBH – Center for Bioinformatics

Research Group for Computational Molecular Design

09 – 2025

OVERVIEW: CHEMINFORMATICS

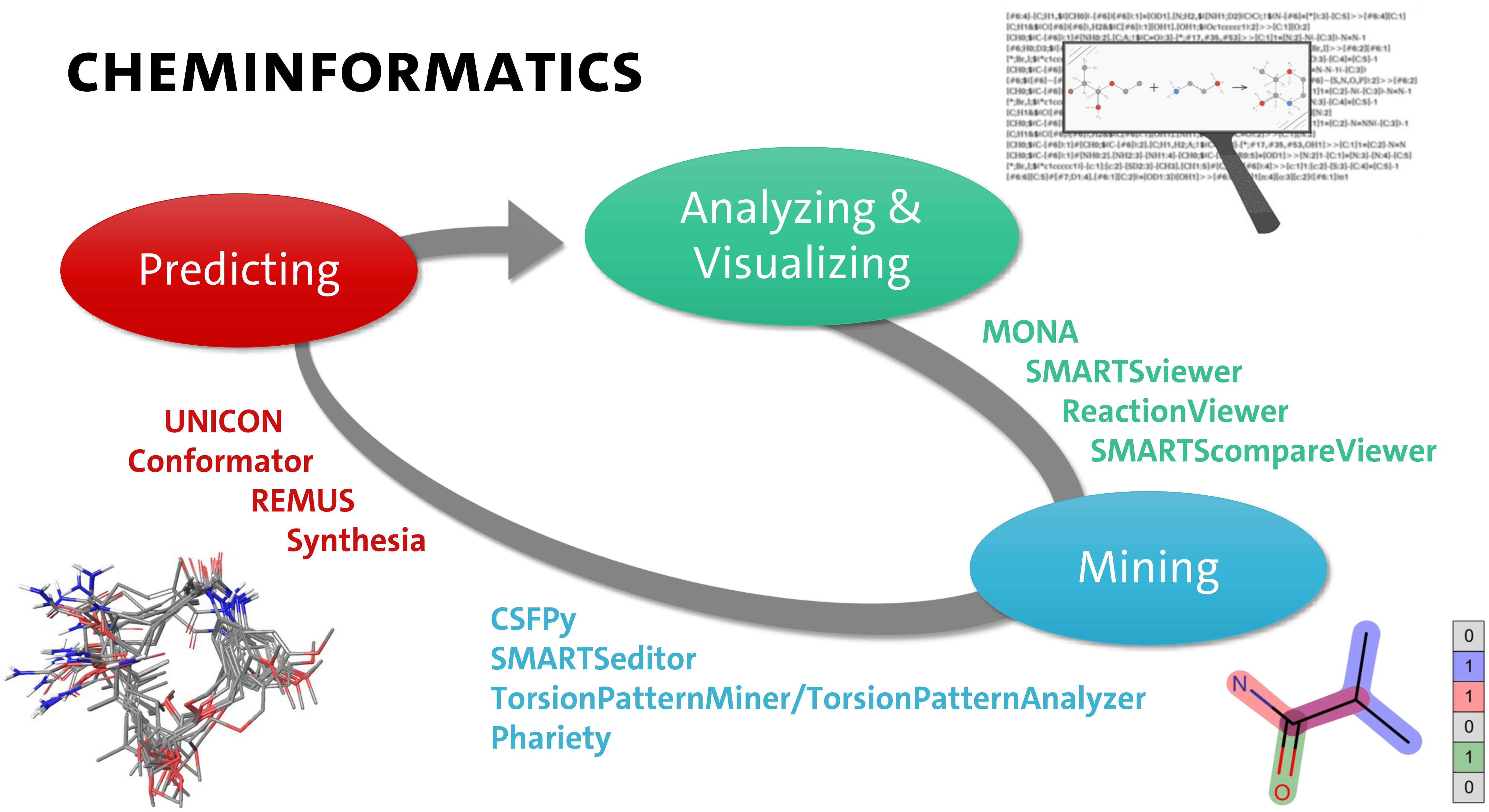
OVERVIEW: CHEMINFORMATICS

Deprecated Tools

OVERVIEW: CHEMICAL SPACES

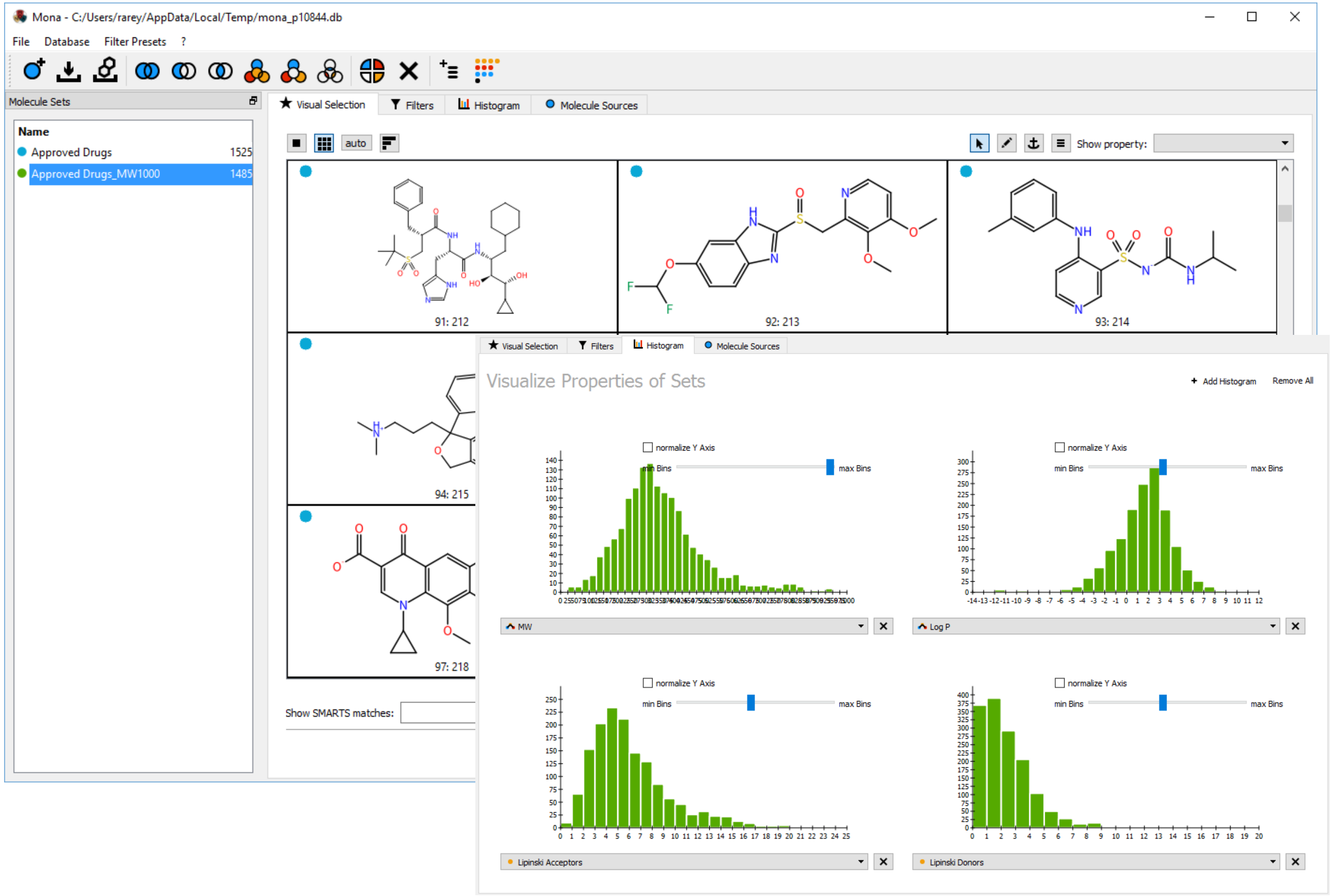
OVERVIEW: STRUCTURE-BASED DESIGN

OVERVIEW: STRUCTURE-BASED DESIGN


CHEMINFORMATICS
Predicting
Analyzing &
Visualizing
Mining
UNICON
Conformator
REMUS
Synthesia
MONA
SMARTSviewer
ReactionViewer
SMARTScompareViewer
CSFPy
SMARTSeditor
TorsionPatternMiner/TorsionPatternAnalyzer
Phariety
N
O
0
1
1
0
1
0

MONA: MANAGING COMPOUND COLLECTIONS

Cheminformatics in an Interactive Fashion. Visualize, analyze, filter, and cluster compound collections; Easily spot the differences between molecule sets; Create your own sub-collections by property filters of SMARTS patterns, by inclusion and exclusion rules, or just manually; No database installation, no scripting, no pipelining - just an easy-to-use graphical interface

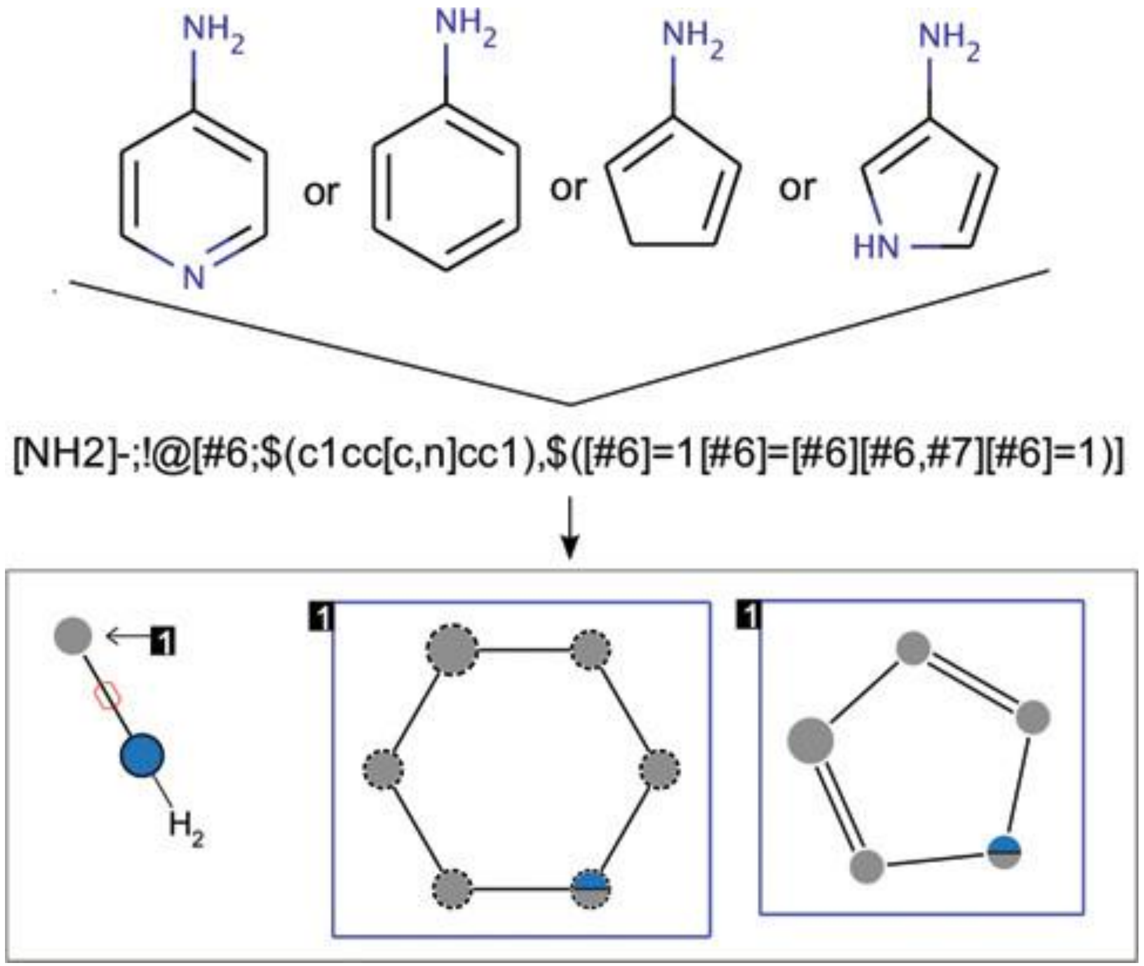


Main developer: *Matthias Hilbig*

Hilbig et al., J Chem Inf Model (2015), 55(10):2071-8 | Hilbig et al., J Cheminf (2013), 5(1):38

SMARTSviewer: INTUITIVELY COMMUNICATING CHEMICAL PATTERNS

Visualizing Complex SMARTS Patterns. Visualize your individual SMARTS strings; Translate a comprehensive but cryptic linear representation into a human-readable, explainable, and visually appealing 2D diagram; The depiction is based on structure diagrams as known from standard molecule visualization; Complex SMARTS patterns are split into simpler parts by extracting the recursive parts in separate structure diagrams; A legend explains the graphic symbols as well as the meaning of the SMARTS in words; This technology is now implemented in the even more versatile ReactionViewer package

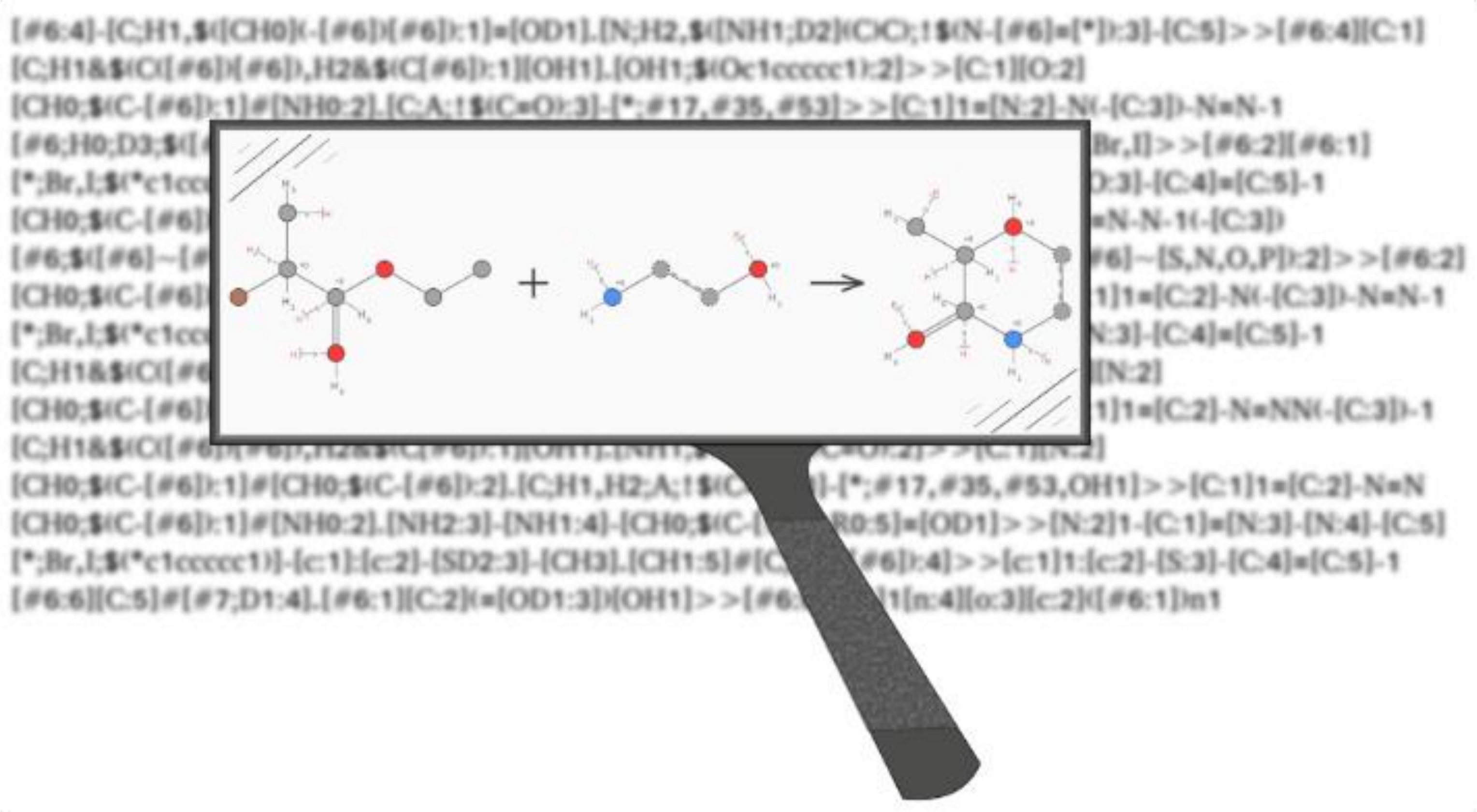


Main developer: *Karen Schomburg*

Schomburg et al., J Chem Inf Model (2010), 50(9):1529-35

ReactionViewer: VISUALIZING CHEMICAL REACTIONS WITH HIGH PRECISION

Visualizing Generic Reaction Patterns. Create easy-to-understand images for generic reaction patterns using the SMARTSviewer technology; Supports Reaction SMILES, Reaction SMARTS, and SMIRKS; Inspired by the IUPAC's Compendium of Chemical Terminology for reaction equations; Features direct image output in PNG, SVG, or PDF format; Enables single command-based conversion of large reaction pattern collections to single PDF documents

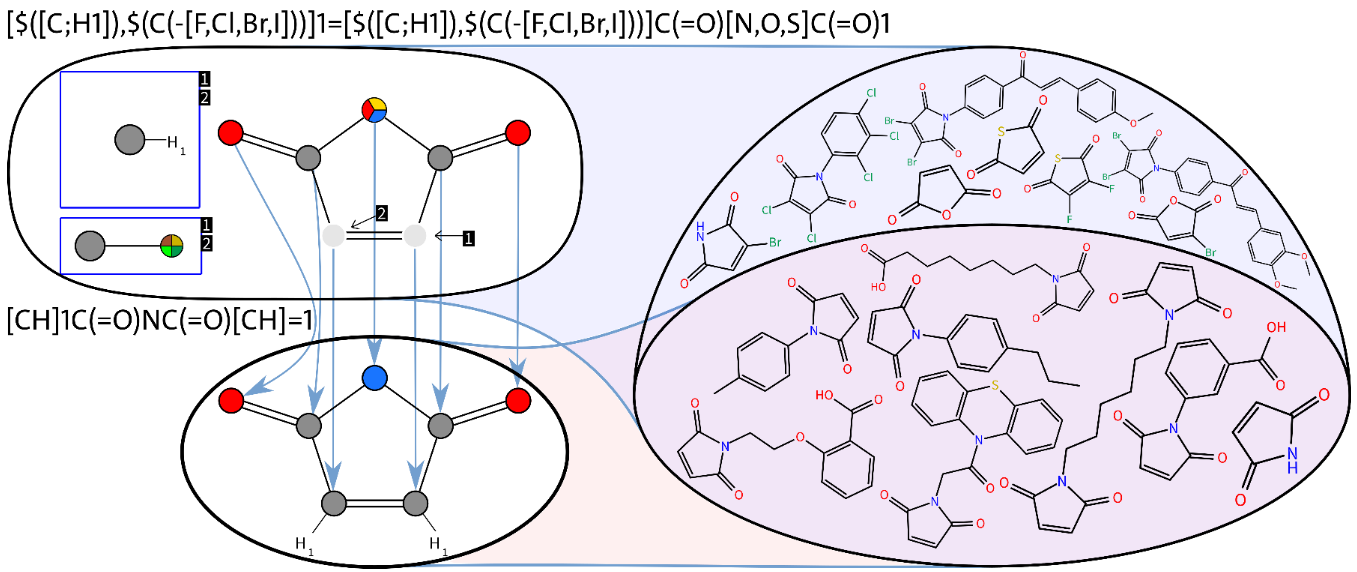


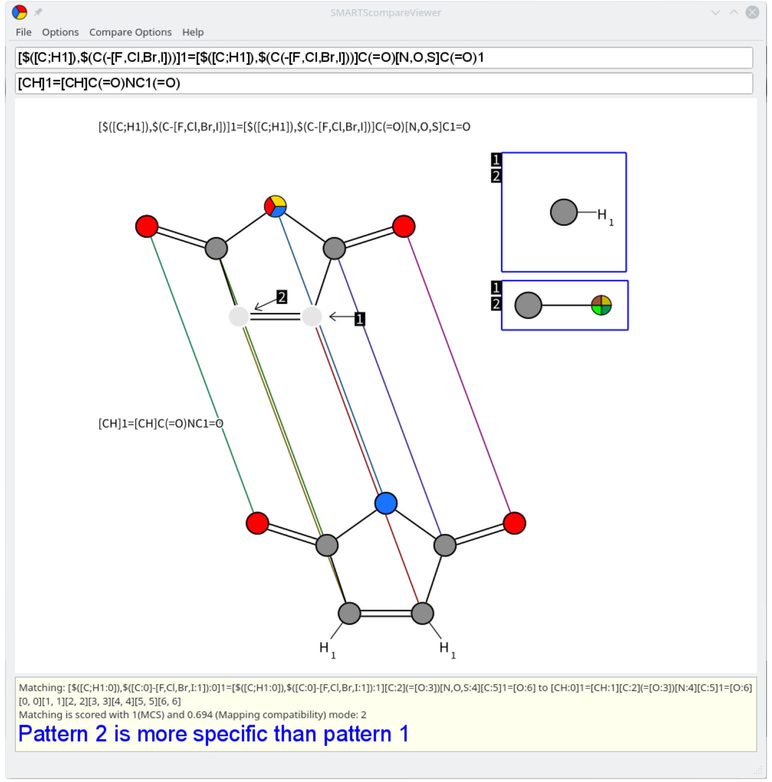
Main developers: *Uschi Dolfus, Robert Schmidt*

Dolfus et al., J Chem Inf Model (2022), 62(19):4680-9

SMARTScompareViewer: ANALYZING AND COMPARING MOLECULAR PATTERNS

Manage and Analyze Large Filter Pattern Collections. Test whether a pattern A is included in another pattern B, i.e., A matches whenever B matches; Construct and verify a pattern hierarchy; Calculate the similarity between patterns; Search for similar, more specific, or more generic patterns in pattern collections like Pfizer LINT, SMARTCyp, or PAINS; Test for incompatible atom or bond type combinations; Visualize node mappings between similar SMARTS expressions; Supports logic and recursive SMARTS expressions; No stereochemistry support so far



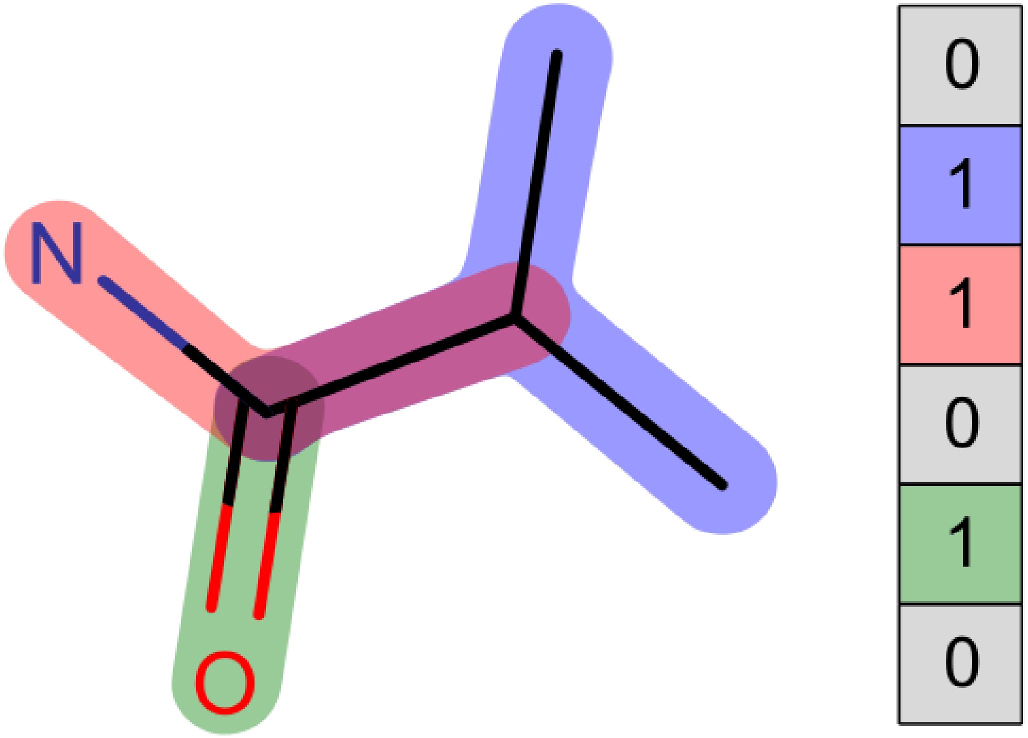


Main developers: *Robert Schmidt, Emanuel Ehmki; Farina Ohm, Andriy Mashychev, Hans-Christian Ehrlich*

Schmidt et al., J Chem Inf Model (2019), 59(6):2560-71 | Ehmki et al., J Chem Inf Model (2019), 59(6):2572-86

CSFPy: THE ALL-CONNECTED-SUBGRAPH TOPOLOGICAL FINGERPRINT

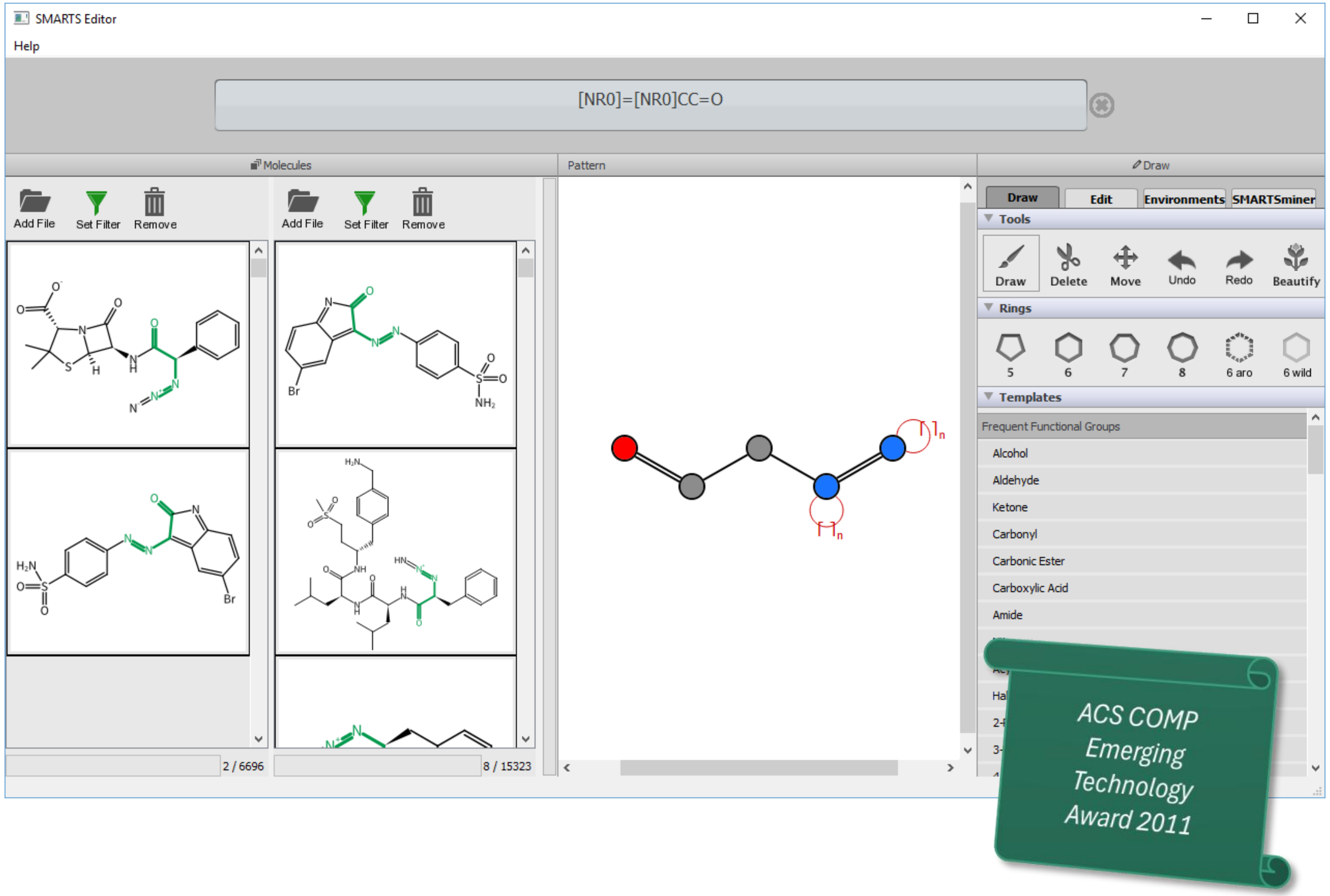
Improved Similarity Searching and Machine Learning. Create a topological fingerprint representing all connected subgraphs within specified size ranges (ECFPs consist only of circular substructures); Covering the complete feature space; Precisely control the level of encoded chemical detail; Convert large collections of molecules and calculate their similarity; Easy-to-integrate thanks to an easy-to-use Python interface



Main developer: *Louis Bellmann*

SMARTSeditor: INTERACTIVE AND SEMI-AUTOMATED GRAPHICAL CREATION OF SMARTS EXPRESSIONS

Simply Draw Your Molecular Pattern. Visualize and understand SMARTS expressions; Draw scaffolds and interactively modify atomic properties; Navigate and specify atomic environments graphically; Match and validate SMARTS expressions on the fly; Calculate and select discriminative patterns separating two molecule sets fully automatically

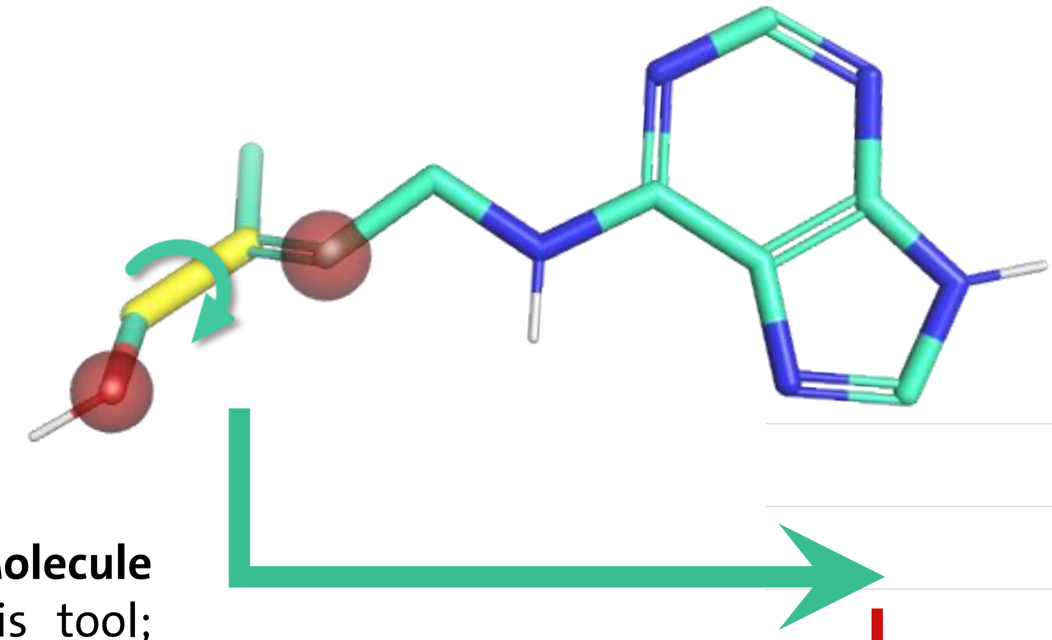


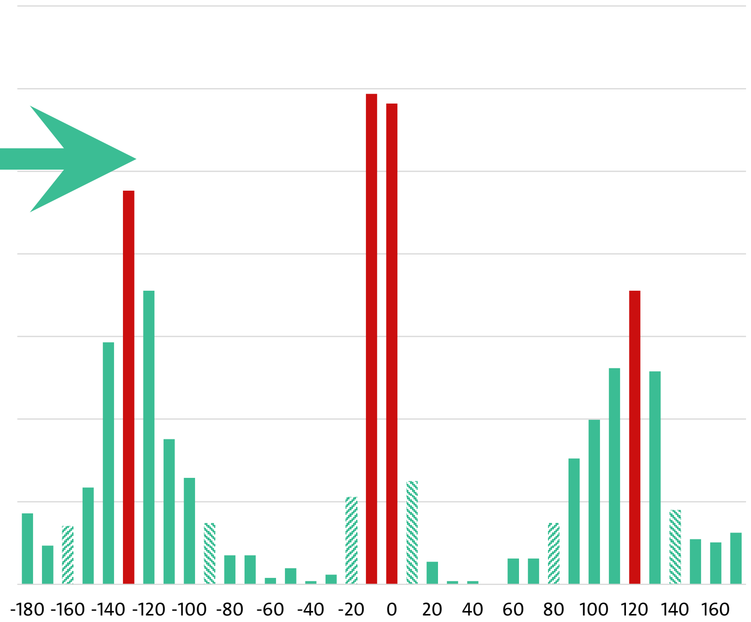
Main developers: *Karen Schomburg, Stefan Bietz, Lars Wetzer*

Bietz et al., J Chem Inf Model (2015), 55(8):1535-48
Schomburg et al., Drug Discov Today (2013), 18(13-14):651-8 | Schomburg et al., J Chem Inf Model (2010), 50(9):1529-35

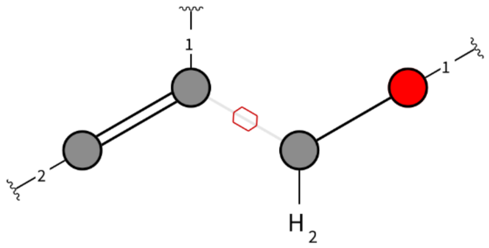
TorsionPatternMiner/TorsionAnalyzer: CREATING KNOWLEDGE-BASED TORSION PROFILES & ANALYZING CONFORMATIONS

TorsionPatternMiner - Creating Torsion Libraries. All-in-one tool to populate torsion angle statistics for SMARTS torsion patterns; Cutting-edge SMARTS pattern analysis with SMARTScompare; Created torsion libraries can be used by many conformer generators like Conformator, RDKit, or OMEGA





TorsionAnalyzer - Analyzing Torsions in Molecule Conformations. Easy-to-use torsion analysis tool; Enables high-throughput torsion analysis via the command line; GUI mode to visualize torsions, torsion patterns, and their statistics



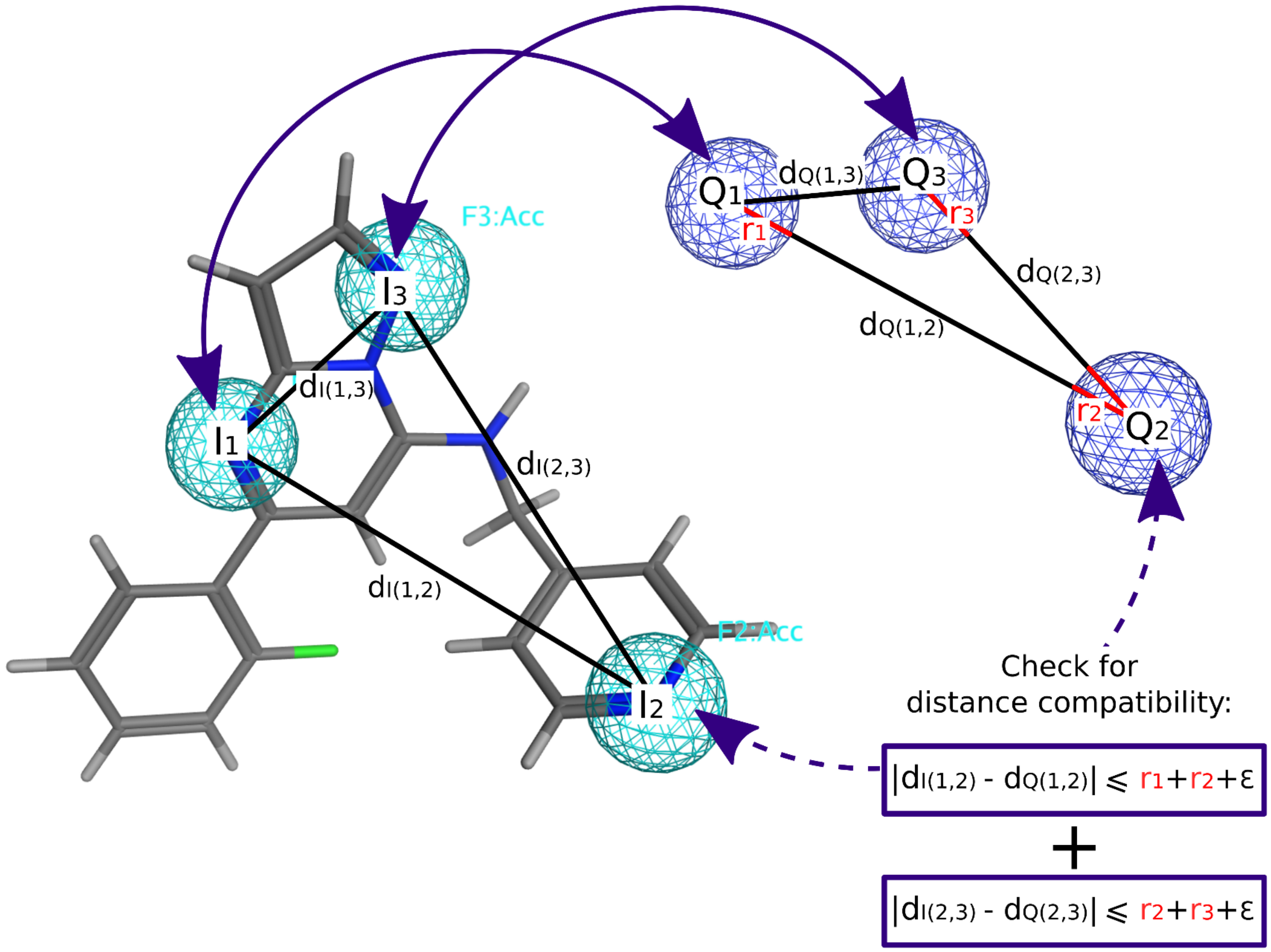
SMARTS:
[CX3:1]=[CX3:2]!@[CH2:3][OX2:4]

Main developers: *Patrick Penner, Robert Schmidt, Agnes Meyder*

Penner et al., J Chem Inf Model (2022), 62(7):1644-53 | Guba et al., J Chem Inf Model (2016), 56(1):1-5

Phariety: PHARMACOPHORE MATCHING MADE EASY

Pharmacophore Matching. Implements a well-established, precise and efficient backtracking algorithm; Enables fast and reliable pharmacophore screening with a command-line tool; Handles all standard pharmacophore features including interaction directions and exclusion volumes; Parses pharmacophore models in PH4 (MOE) or JSON (Pharmer) format and molecules in SMILES or SDF format; A fully integrated structure and conformer generator enables on-the-fly searching

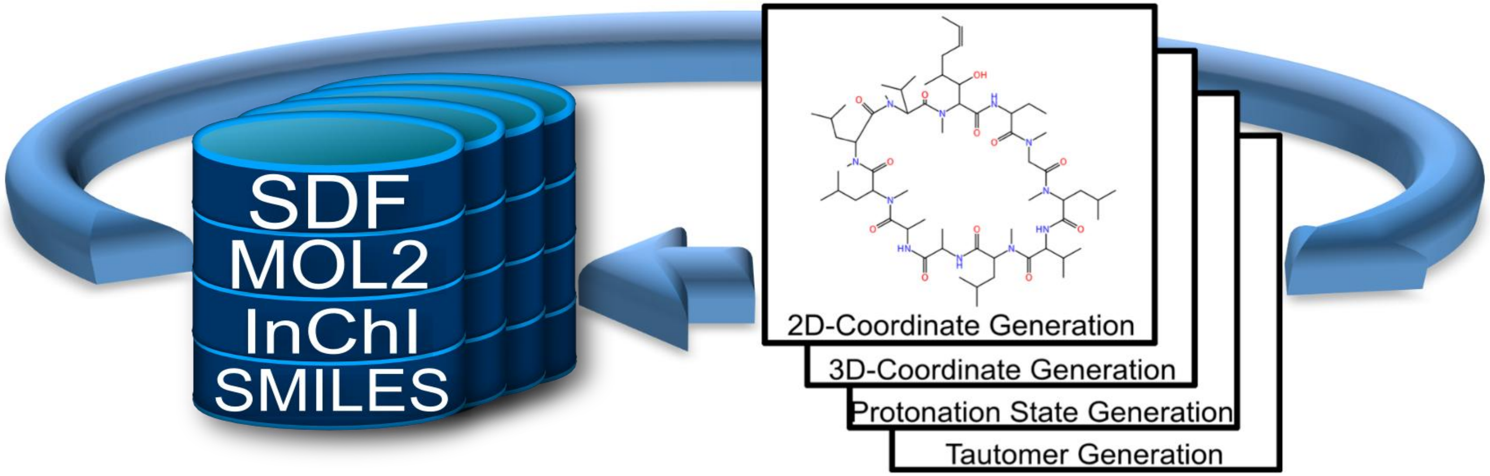


Main developer: *Uschi Dolfus*

Meyenburg et al., J Comput Aided Mol Des (2023), 37(1):1-16

UNICON: THE UNIVERSAL FILE FORMAT CONVERTER

Convert Between Various File Formats Including Coordinate Generation. Convert compound files between SDF, MOL2, InChI, and SMILES; Import from PDB and mmCIF; Export to InChI Key; Avoid conversion errors with a consistent chemical model; Generate 2D coordinates with high layout quality; Generate low-energy 3D conformations with Conformator using the initial conformation as starting point; Enumerate tautomeric and protomeric forms; Generate low-energy conformer ensembles; Do it all with high speed in a single step with an easy-to-use command-line tool

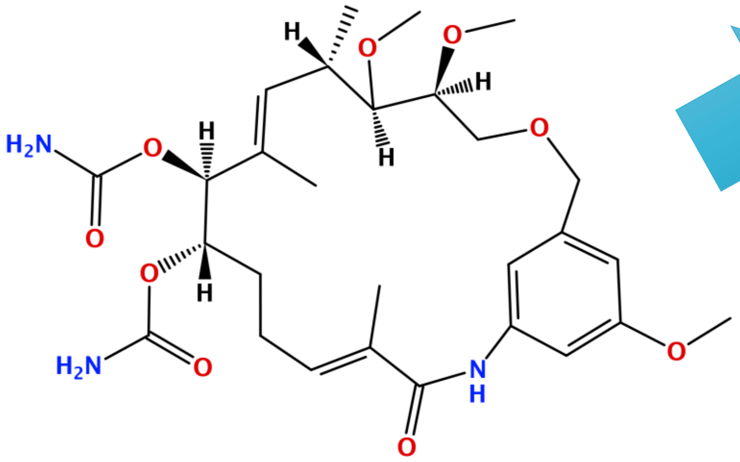


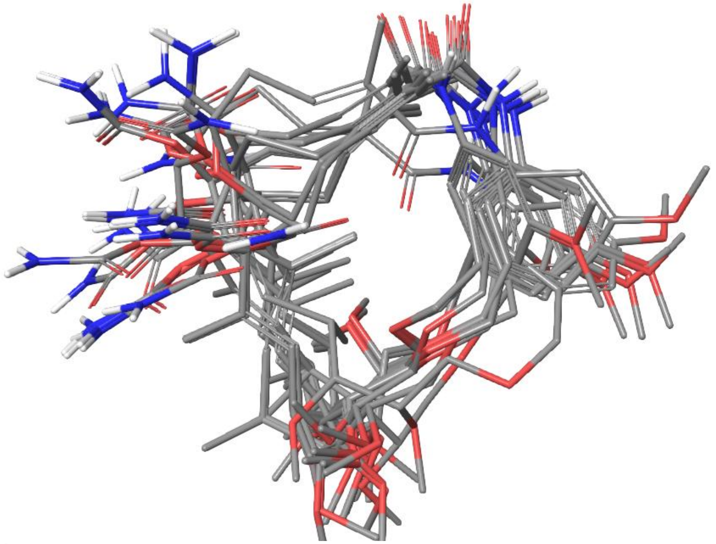
Main developers: *Kai Sommer, Sascha Urbaczek, Adrian Kolodzik*

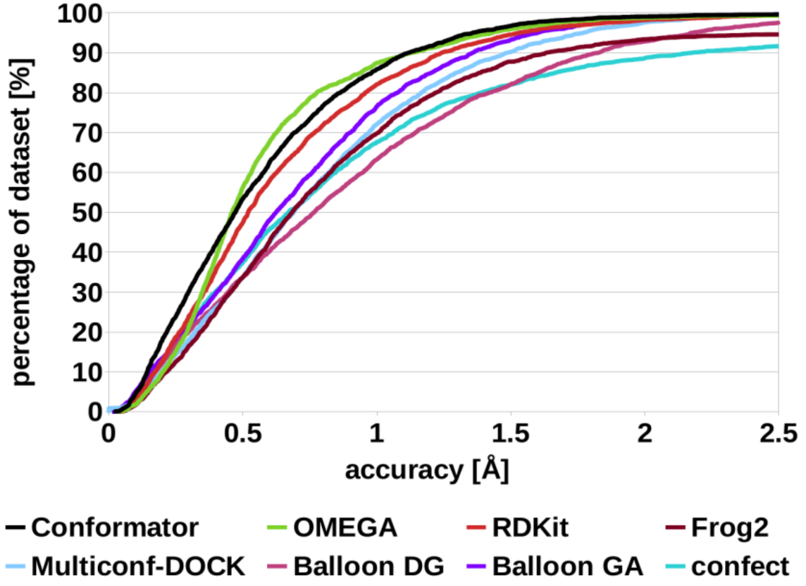
Sommer et al., J Chem Inf Model (2016), 56(6):1105-11
Urbaczek et al., J Chem Inf Model (2014), 54(3):756-66 | Urbaczek et al., J Chem Inf Model (2011), 51(12):3199-207

Conformator: GENERATION OF CONFORMER ENSEMBLES

Generate High-Quality Conformer Ensembles for Small Molecules. Read molecular structures from SDF and MOL2 files as well as from SMILES and InChI notations; Sample the conformational space of small molecules (including macrocycles) with a fast, accurate, and effective knowledge-based algorithm without PDB bias; Create diverse conformer ensembles for any given ensemble size by automated RMSD thresholding







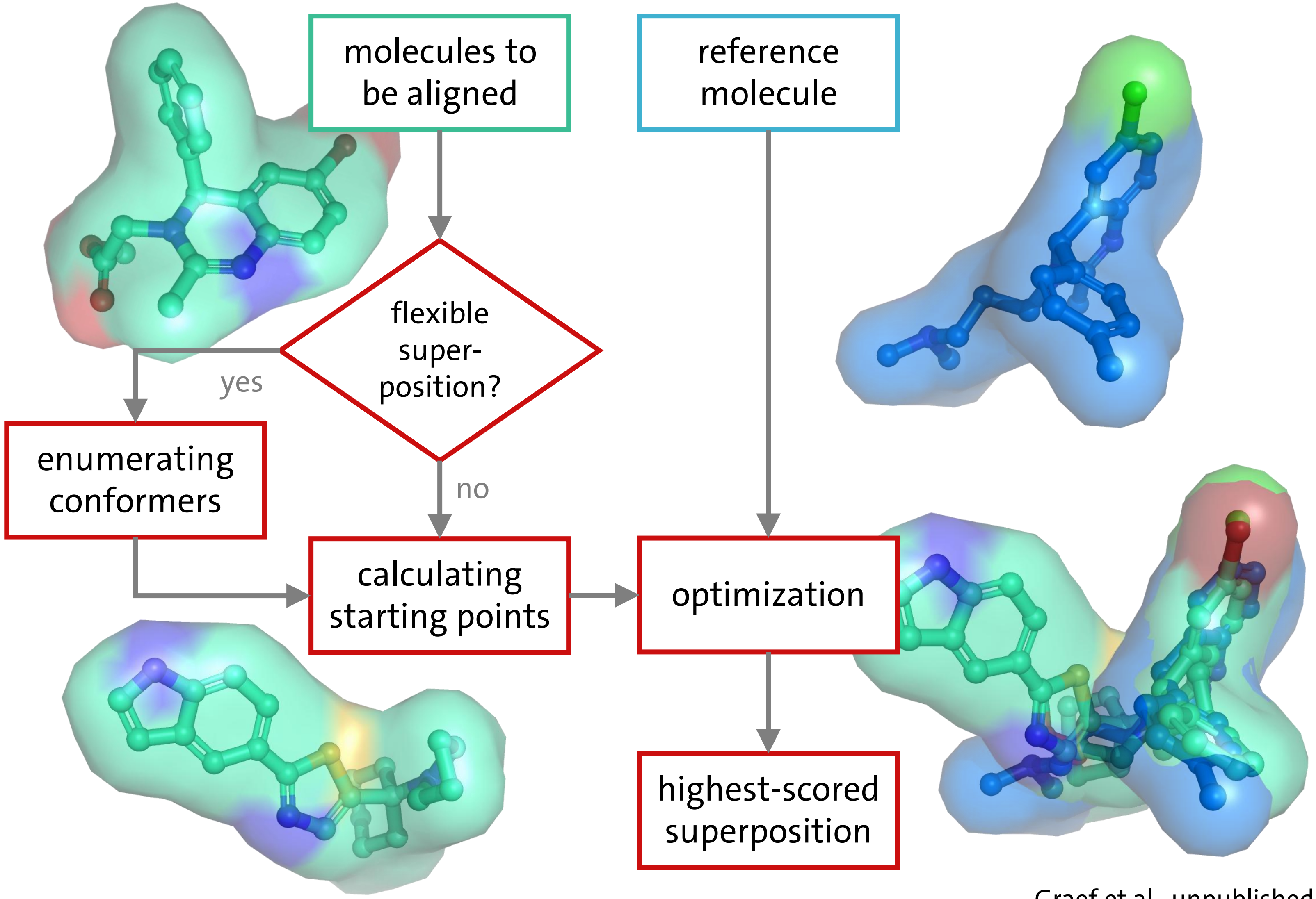
Main developers: *Nils-Ole Friedrich, Florian Flachsenberg, Kai Sommer*

Friedrich et al., J Chem Inf Model (2017), 57(3):529-39 | Friedrich et al., J Chem Inf Model (2019), 59(2):731-42

REMUS: SHAPE-BASED ALIGNMENTS - INTERACTIVE EXPLORATION AND SCREENING

Search for 3D Similarity Within Small and Medium-Sized Compound Collections. Calculate molecular similarity based on classical Gaussian (colored) shape descriptions; Handle molecular flexibility by a conformer generator (Conformator) and conformer fine tuning during alignment; Employ a brand-new step-limited Broyden-Fletcher-Goldfarb-Shanno (BFGS) numerical optimizer; Visually explore molecular alignments and perform them semi-automated by RMSD fitting for user-specified atom pairs

Main developer: *Joel Graef*



Graef et al., unpublished

Synthesia: FULL MODIFICATION CONTROL OVER RETROSYNTHETIC ROUTES FOR OPTIMIZING LEAD STRUCTURES

Generating Structural Analogues in a Synthesis-Aware Process. Integrates the preservation of the synthesizability of the target compounds into the lead structure modification process; Creates structural diversity for a lead structure that matches user-defined molecular properties without losing the applicability of a particular synthetic pathway; Maximizes the synthetic efficiency and provides an initial estimation of the effort of synthesizing the entire compound series; Requires a retrosynthetic route of the lead structure in JSON format (chemical nodes need to be provided in SMILES Format and reaction nodes in Reaction SMARTS format, and a children object if it is not a leaf) and a list with possible substitute candidates

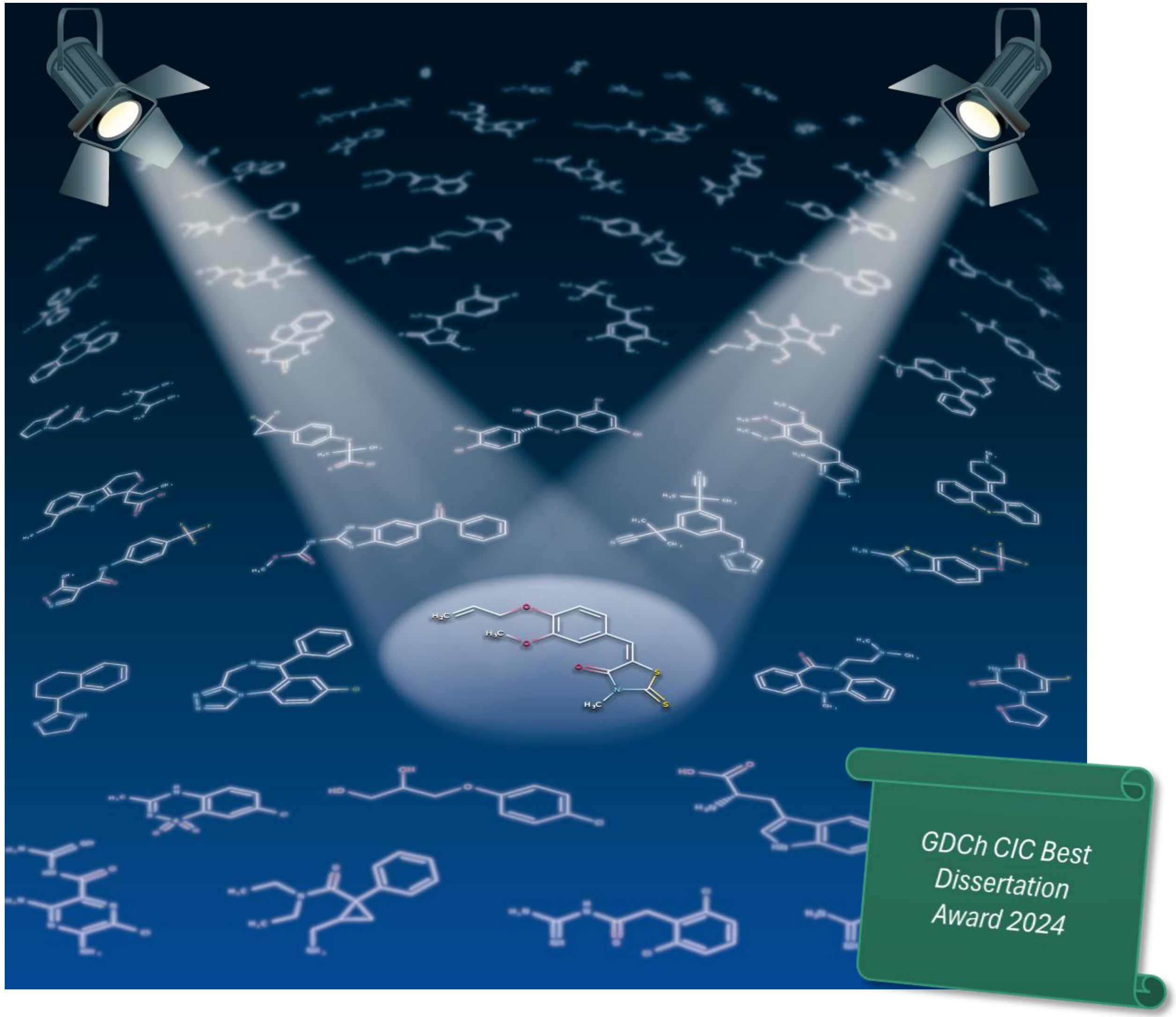
Main developer: *Uschi Dolfus, Robert Schmidt*

Dolfus et al., J Chem Inf Model (2022), 62(15):3565-76

CHEMICAL SPACES

SpaceLight: TOPOLOGICAL SIMILARITY SEARCHES IN LARGE CHEMICAL SPACES

Explore Chemical Spaces Containing 10^{15++} Compounds in Seconds. Search for structurally similar compounds in chemical spaces far too large for state-of-the-art substructure-driven search methods; Retrieve results similar to a classical fingerprint (e.g., using the ECFP) search; Detect scaffolds from large building block collections; Create your own large reaction-driven chemical space; Use your own laptop and finish search runs within seconds



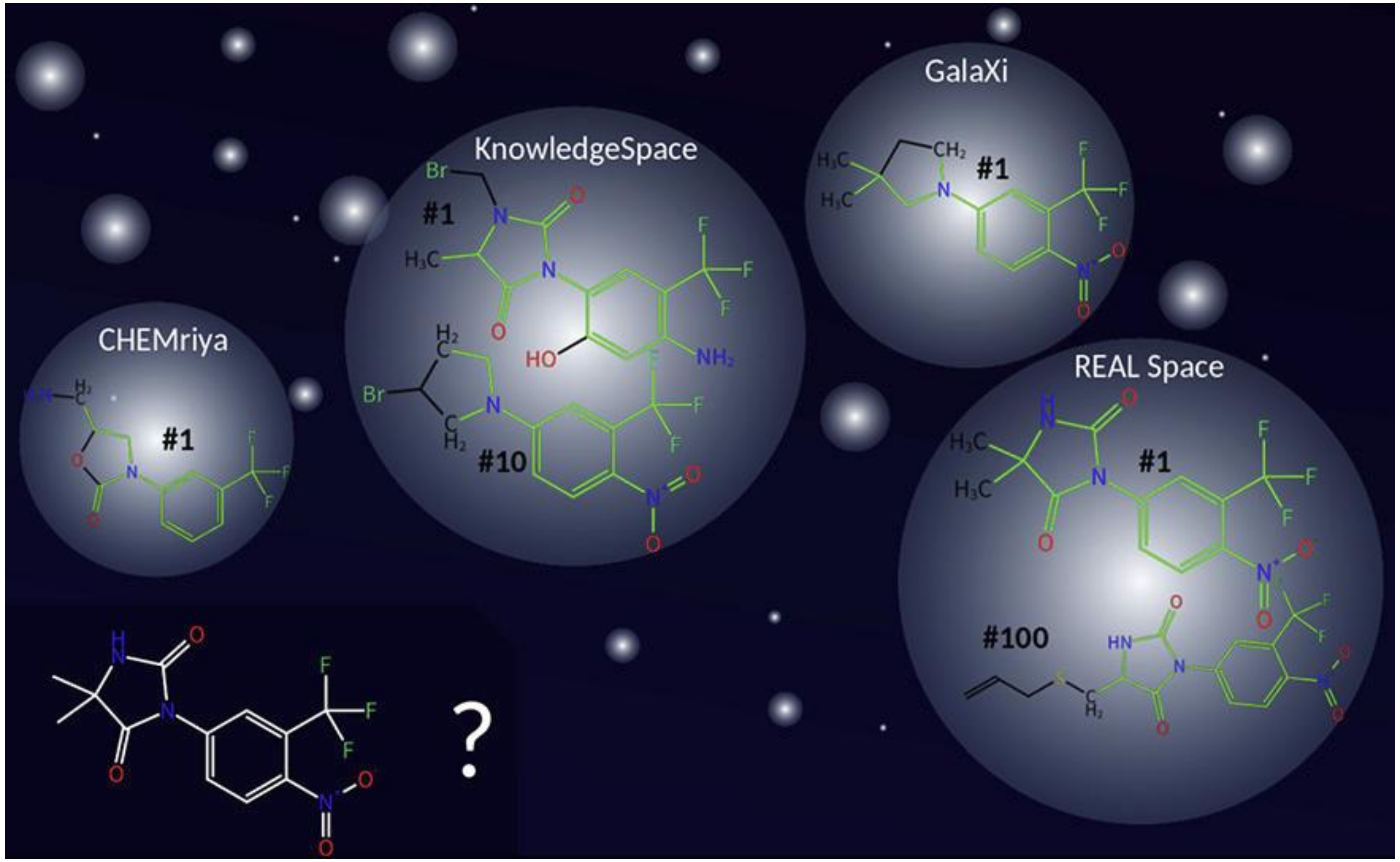
Note: The ChemBio Suite SpaceLight version works with public spaces only. SpaceLight for commercial spaces can be licensed exclusively from BioSolveIT GmbH.

Main developer: *Louis Bellmann*

Bellmann et al., J Chem Inf Model (2021), 61(1):238-51

SpaceMACS: MAXIMUM COMMON SUBSTRUCTURE SEARCHES IN LARGE CHEMICAL SPACES

Search Chemical Spaces Containing 10^{15++} Compounds by Substructures in Seconds.
Search for structurally similar compounds in chemical spaces by maximum common substructure similarity; Retrieve hundred thousands of compounds by maximum common substructures within tenths of seconds on a single, standard server; Apply a chemically motivated similarity measure finding the closest analog available in chemical spaces



Note: The ChemBio Suite SpaceMACS version works with public spaces only. SpaceMACS for commercial spaces can be licensed exclusively from BioSolveIT GmbH.

Main developer: *Robert Schmidt*

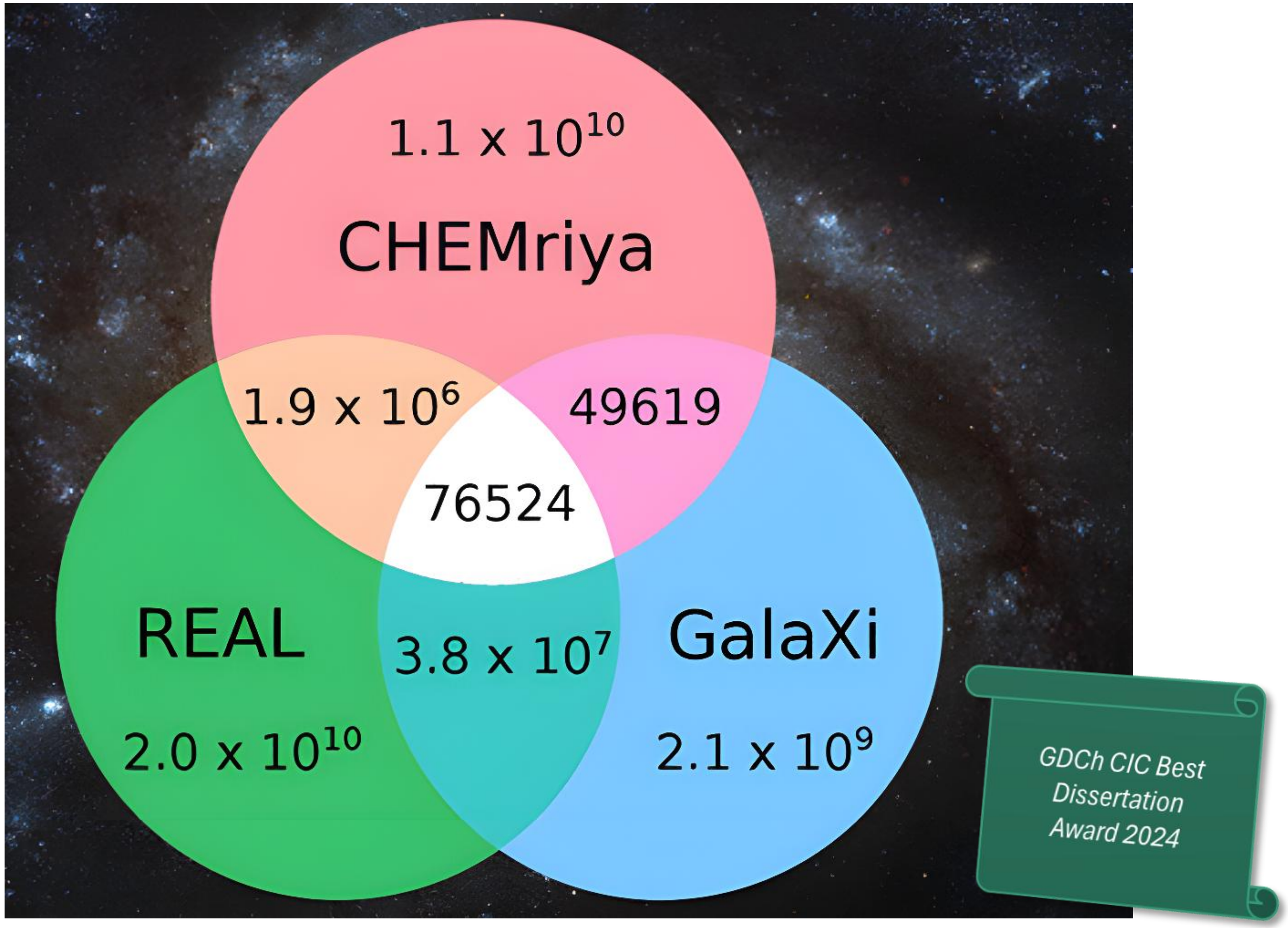
Schmidt et al., J Chem Inf Model (2021), 62(9): 2133-50

SpaceCompare: ANALYZING, COMPARING, AND OPTIMIZING ULTRA-LARGE CHEMICAL SPACES

Analyze Chemical Spaces. Calculate the exact overlap of products shared by very large chemical fragment spaces (provided that the overlap is enumerable); Calculate exact physicochemical property distributions for ultra-large chemical spaces (heavy atom count, molecular weight, clogP, number of hydrogen bond donors and acceptors); Generate chemical subspaces with optimized product properties

Note: SpaceCompare works with public, self-created, and commercial spaces available from BioSolveIT GmbH.

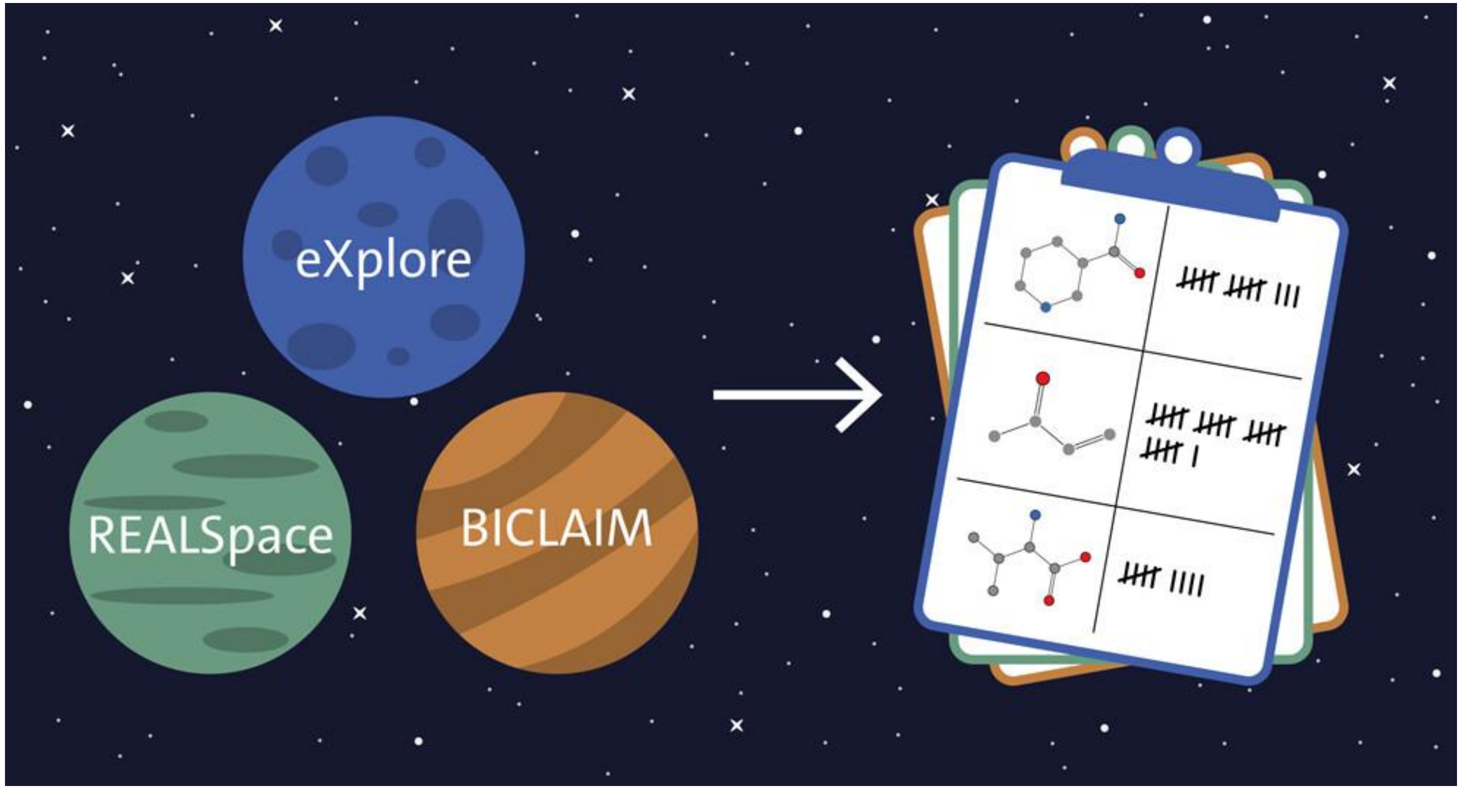
Main developers: *Louis Bellmann, Patrick Penner*



Bellmann et al., J Chem Inf Model (2022), 62(3):553-66 | Bellmann et al., J Chem Inf Model (2022), 62(11):2800-10

SpaceProp: COMPUTING MOLECULAR PROPERTY DISTRIBUTIONS IN LARGE CHEMICAL FRAGMENT SPACES

A Telescope for Chemical Spaces. Calculate exact chemical property distributions for ultra-large chemical spaces (heavy atom count, molecular weight, clogP, number of hydrogen bond donors and acceptors, TPSA, number of rotatable bonds, number of molecules containing user-defined molecular patterns as defined by SMARTS patterns) and get example molecules for every value or pattern



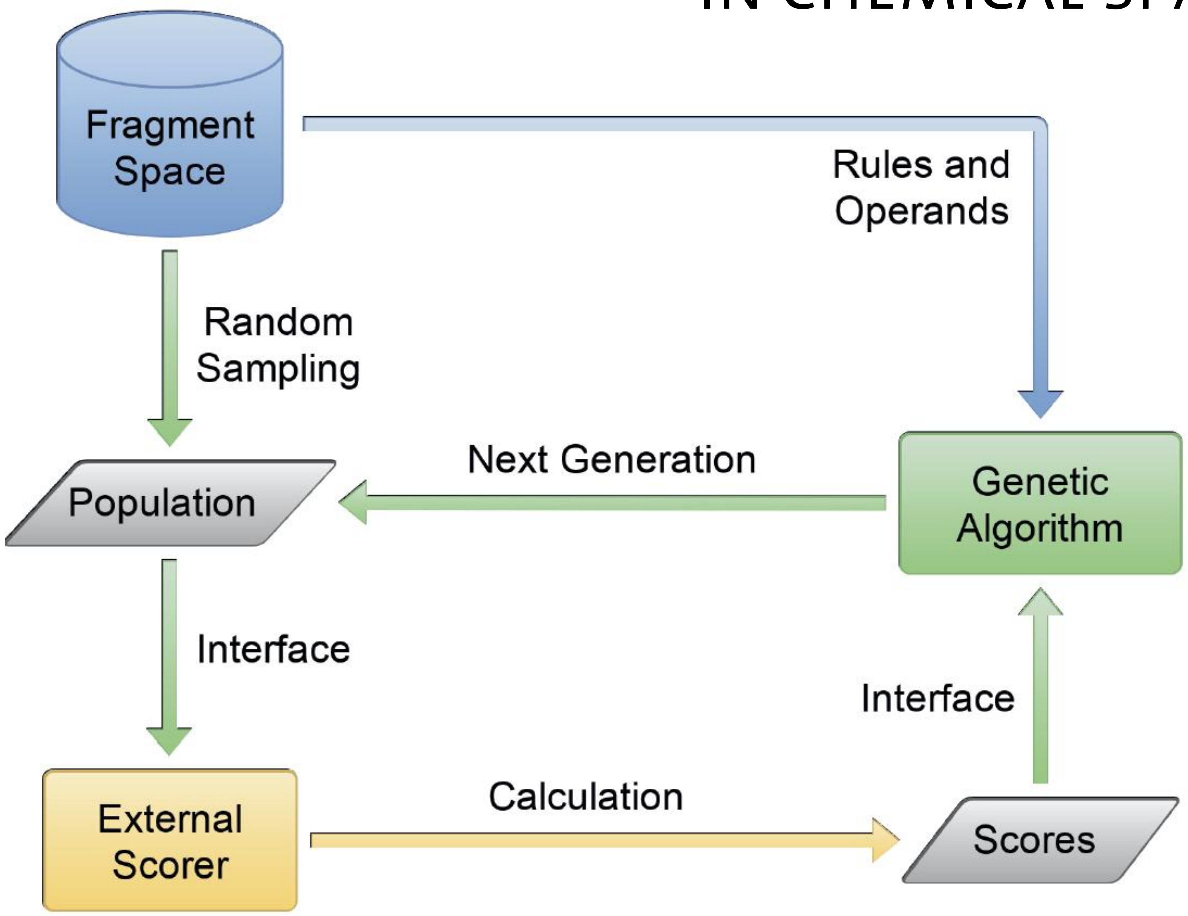
Note: SpaceProp works with public, self-created, and commercial spaces available from BioSolveIT GmbH.

Main developers: *Justin Lübbers, Louis Bellmann*

Bellmann et al., J Chem Inf Model (2022), 62(11):2800-10 | Lübbers et al., J Chem Inf Model (2024), 64(6):2008-20

Galileo: GENETIC ALGORITHM FOR GENERAL PURPOSE SEARCHING IN CHEMICAL SPACE

Chemical Space Search. Searching for molecules in chemical fragment spaces by user-defined scoring functions; Genetic algorithm directly operating in fragment spaces (so each created molecule is from the space); Combinable with external scoring functions via system call; Pharmacophore search based on Phariety integrated; First 3D search engine for fragment spaces, however, quite resource-demanding

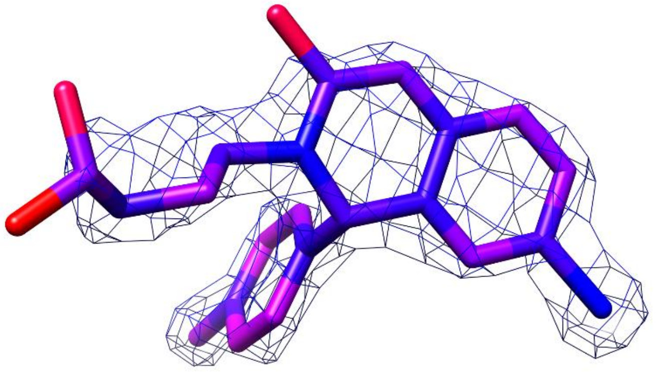


Note: Galileo works with public, self-created, and commercial spaces available from BioSolveIT GmbH.

Main developers: *Christian Meyenburg, Uschi Dolfus*

Meyenburg et al., J Comput Aided Mol Des (2023), 37(1):1-16

STRUCTURE-BASED DESIGN



Predicting

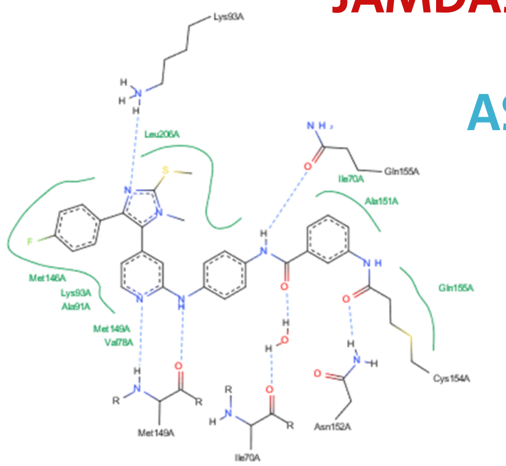
Analyzing & Visualizing

Mining

DoGSite3
WarPP
JAMDA
JAMDAscorer

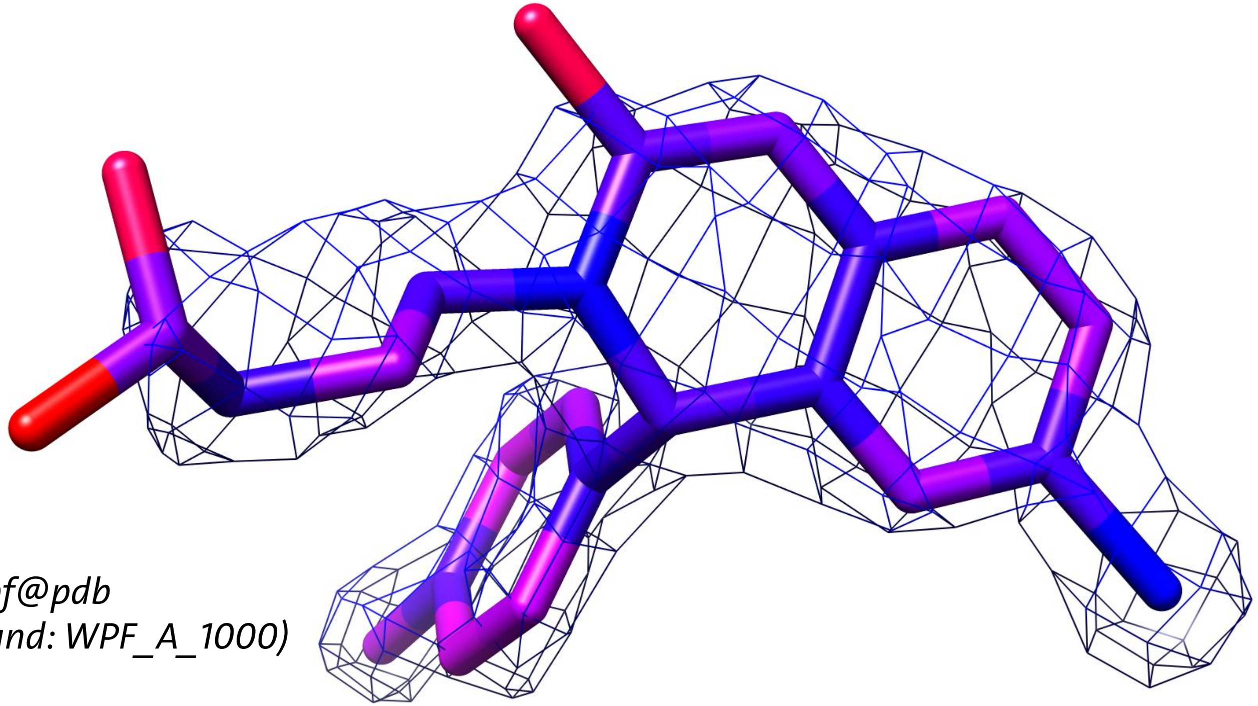
EDIAscorer
StructureProfiler
AltLocEnumerator
LiifeSoaks

ASCONA/SIENA
MicroMiner
SiteMine
PiMine

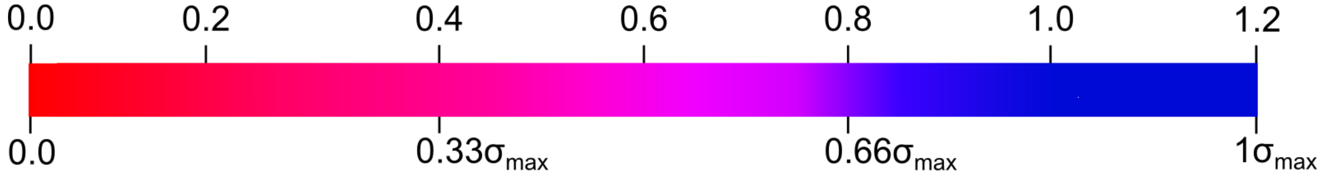


EDIAscorer: ELECTRON DENSITY SCORE FOR INDIVIDUAL ATOMS

Estimate the Experimental Uncertainty of Individual Atoms in X-Ray Structures. Create a map showing the electron density support of each individual atom in a crystallographically solved structure; Automatically compare a structure model to its electron density in CCP4 format; Calculate an $EDIA_m$ value for each molecule emphasizing individual atoms with low EDIA values to quantify structural inconsistencies; Get error analysis information for atoms with low electron density support values; Fully-automated reliable prefiltering of large structure collections



2wpf@pdb
(ligand: WPF_A_1000)

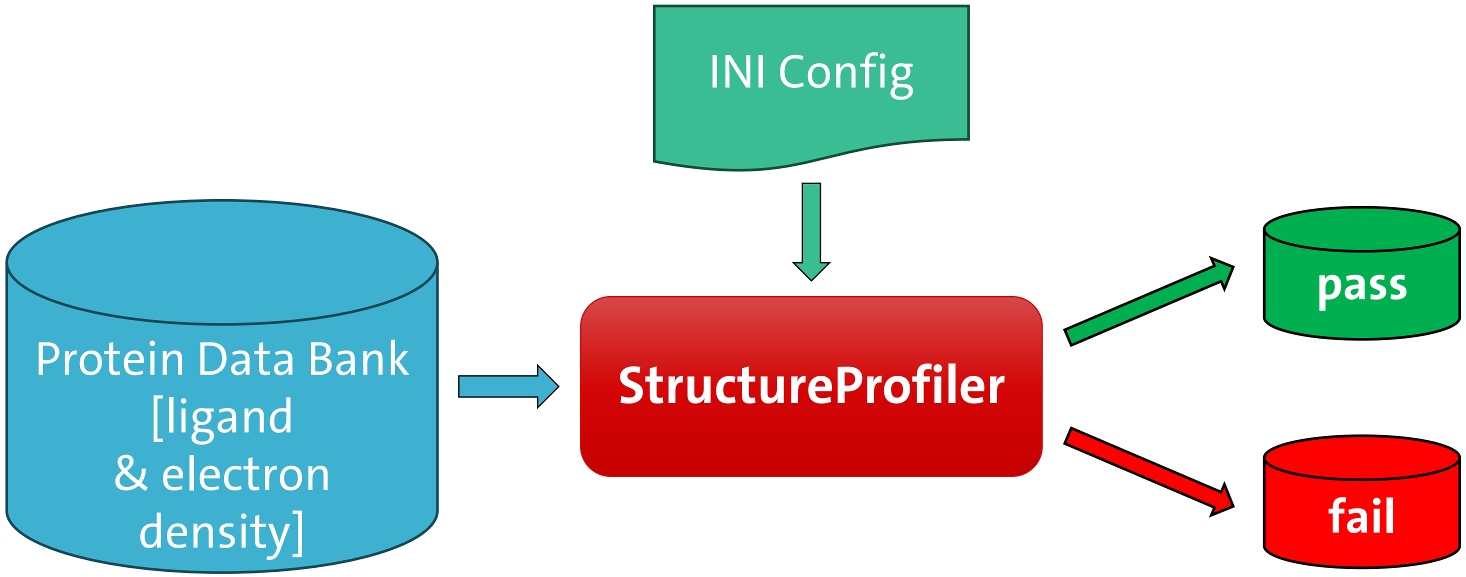


Main developers: *Agnes Meyder, Eva Nittinger, Florian Flachsenberg*

Meyder et al., J Chem Inf Model (2017), 57(10):2437-47

StructureProfiler: FILTERING PDB STRUCTURES BY PROPERTIES AND EXPERIMENTAL CRITERIA

Automatically Profile X-Ray Structures with the Most Frequently Applied Quality Criteria. All-in-one tool for PDB entry selection; Fully configurable via INI parameter files; Preconfigured filters highly similar to Astex, Iridium, Platinum, and a combination thereof; Supports PDB and mmCIF format with optional electron density files; Includes EDIAscorer results and further crystallographic quality criteria; Includes torsion angle and clash checks; Includes ligand descriptor filters, e.g., Lipinski's Rule of Five; Enables fast and efficient updating of benchmark datasets with new PDB files

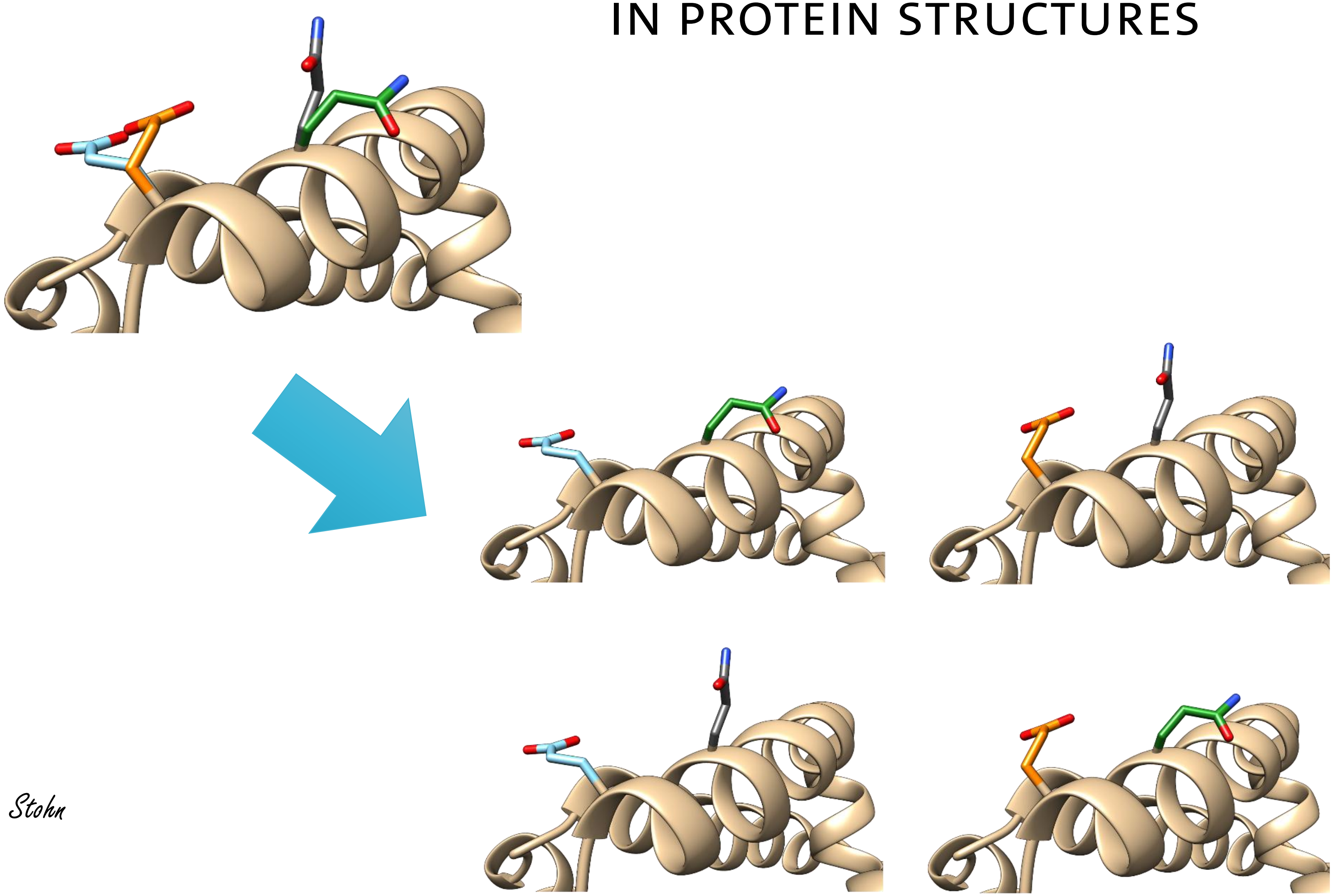


Main developers: *Agnes Meyder, Stefanie Kampen*

Meyder et al., Bioinformatics (2019), 35(5):874-6

AltLocEnumerator: DEALING WITH EXPERIMENTAL ALTERNATE LOCATIONS IN PROTEIN STRUCTURES

Modeling Structural Ambiguities. Automatic handling of alternate locations (AltLocs) in experimentally solved protein structures in PDB or mmCIF format; Exploring and evaluating protein conformational space through experimentally validated AltLocs; Using a graph-based algorithm to enumerate all valid structure conformations that do not lead to clashes or introduce chain breaks; Generating all AltLoc-induced structural variants and writing them to separate PDB files

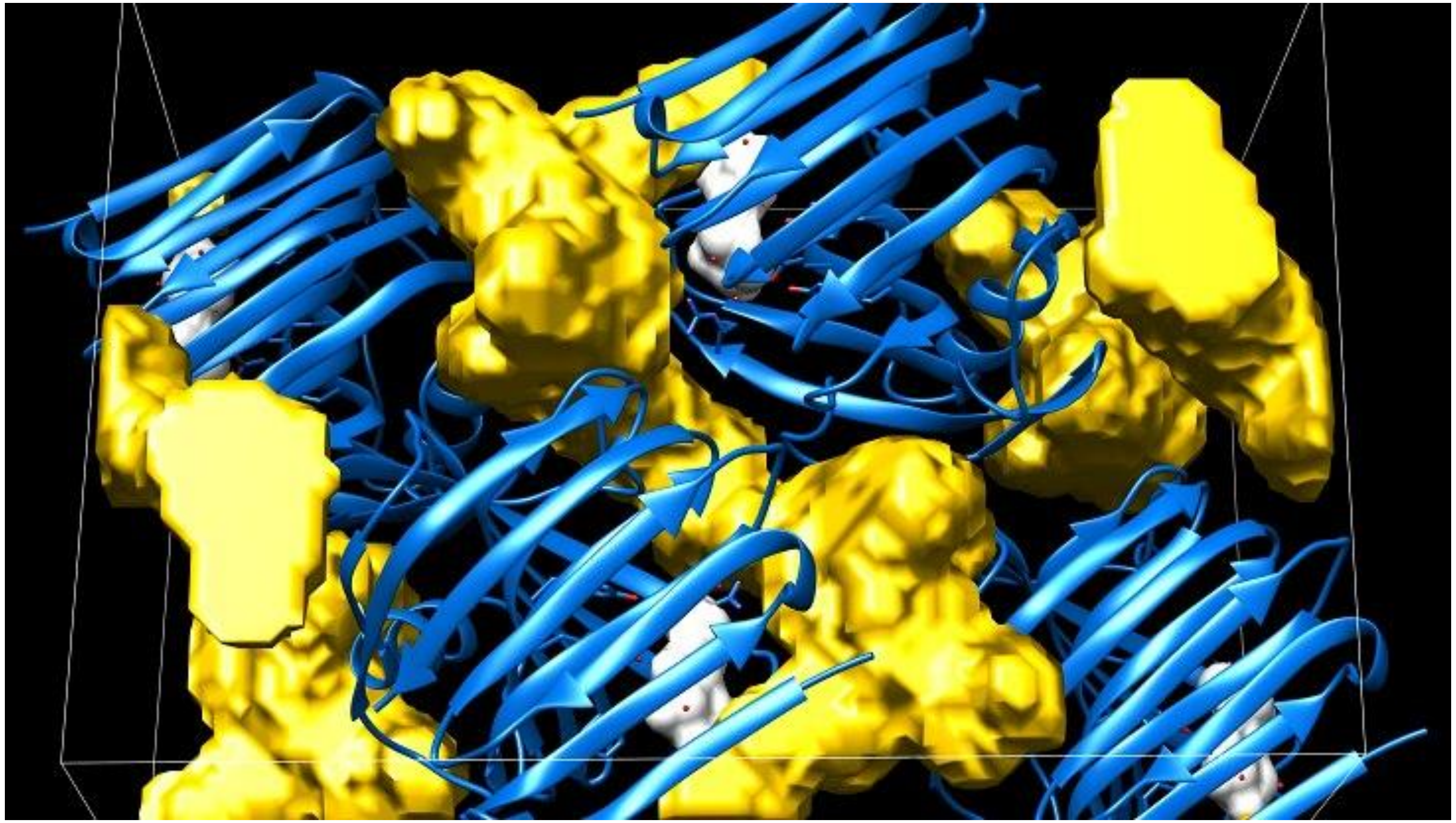


Main developers: *Jochen Sieg, Torben Gutermuth, Tim Stohn*

Gutermuth et al., J Chem Inf Model (2023), 63(8):2573-85

LifeSoaks: CALCULATING SOLVENT CHANNELS IN CRYSTALS AND THEIR CORRESPONDING BOTTLENECK RADII

Find Solvent Channels in Crystallographic Structures and Predict their Accessibility through an Automated Annotation of Bottleneck Radii. Simplifies the process of manually checking a crystal structure for solvent channels; Assigns bottleneck radii for solvent channels and small molecule binding sites and generates solvent maps in CCP4 format for visualizations; Ideally suited as analysis tool prior to soaking experiments to select the most suitable experimental conditions and crystal forms; Runs in an automated fashion and finishes within seconds to minutes for moderately sized unit cells



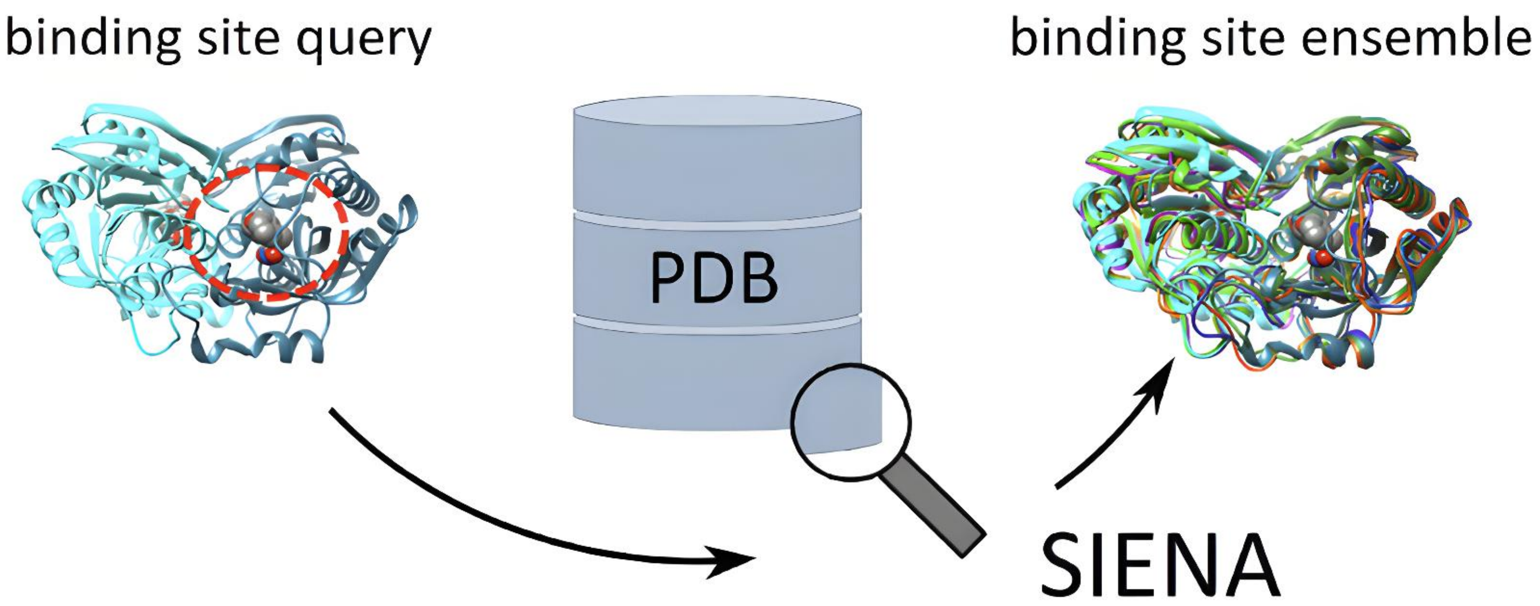
Main developer: *Jonathan Pletzer-Zelgert*

Pletzer-Zelgert et al., Acta Crystallogr D Struct Biol (2023), 79(Pt 9):837-56

ASCONA/SIENA: BINDING SITE SEARCH AND ALIGNMENT

Create Aligned Binding Site Ensembles. Search for sequence-similar binding sites in the whole PDB; Control sequence identity and coverage of the detected similar sites; Find related binding sites in dimeric and multimeric structures consisting of identical and different chains; Consider structural variation during the search and alignment; Create a multiple structure alignment of binding sites; Reduce binding site collections to small ensembles with high structural variance

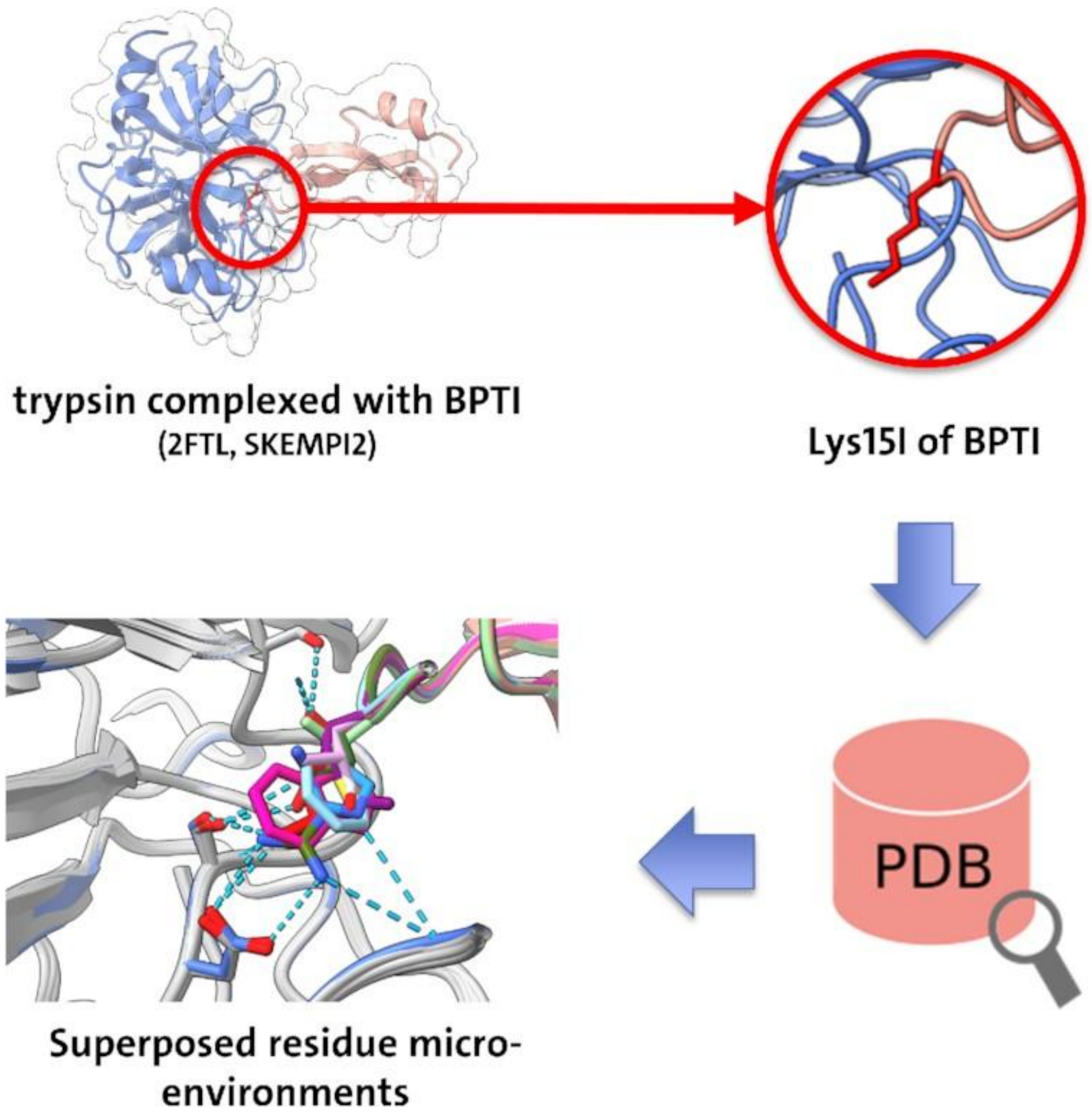
Main developer: *Stefan Bietz*



MicroMiner: SEARCHING SIMILAR 3D RESIDUE MICROENVIRONMENTS

Mining for Single-Residue Substitutions in Protein Structure Databases. Searches protein residue environments with local sequence and structure similarity; Structural mutation search in the entire PDB, your inhouse structure collection, or (subsets of) the AlphaFold Database; Explore the mutation landscape of proteins with experimental or predicted structures; Applicable in single domains or even on protein-protein or protein-ligand interfaces; Several filter options to simplify downstream analysis

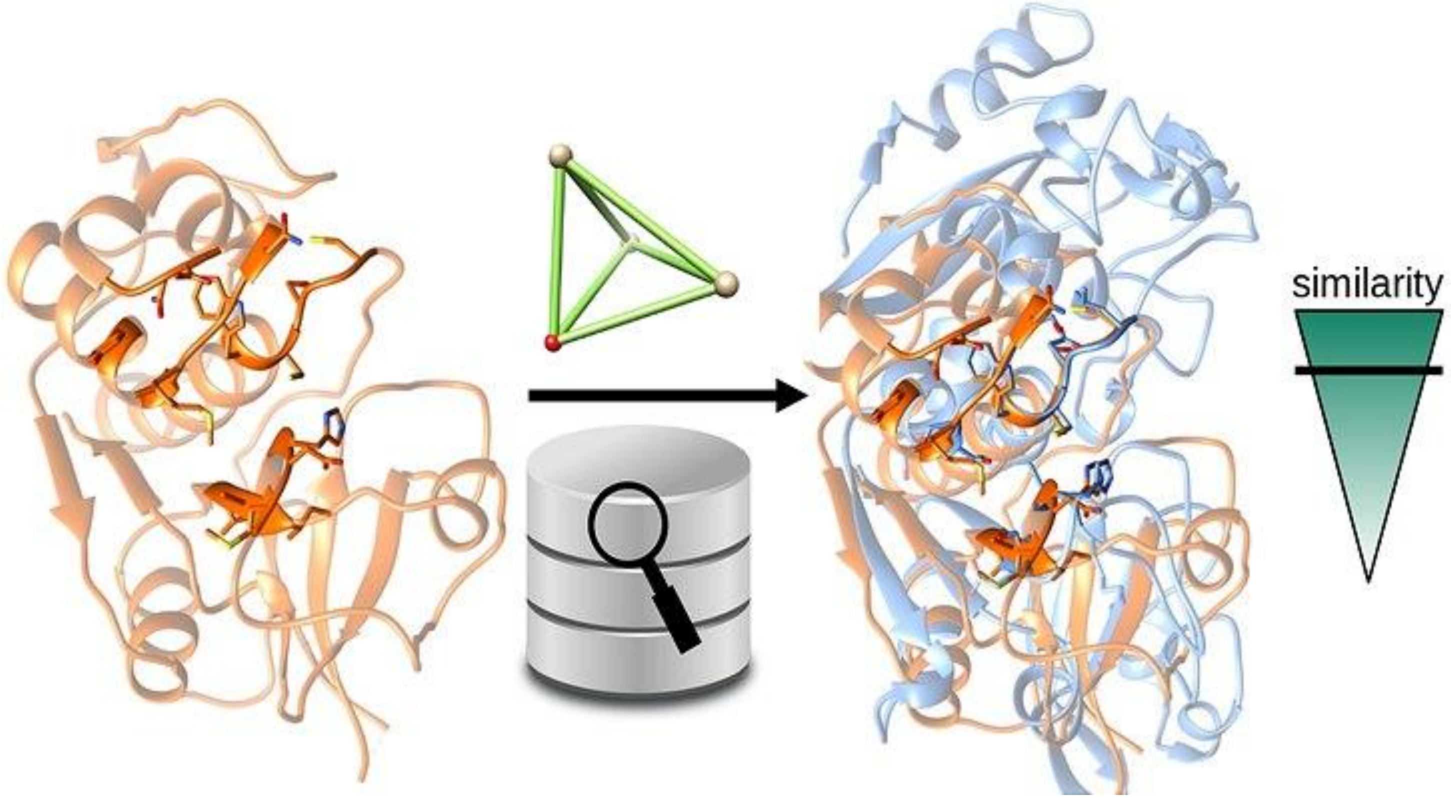
Main developers: *Jochen Sieg, Stefan Bietz*



Sieg et al., Brief Bioinform (2023), 24(6):bbad357

SiteMine: BINDING SITE SIMILARITY SEARCHES IN LARGE COLLECTIONS OF PROTEIN POCKETS

Compare Small Molecule Binding Sites in Large Databases and Calculate their Physico-chemical and Shape Similarity. Screen large GeoMine databases of ligand-based and/or predicted binding sites for similar sites to a query site; Databases of individual experimental or predicted structures can be created automatically; Prefiltering cascade based on several pocket properties and subsets of structure identifiers prior to screening; Includes two scoring functions accounting for the physicochemical and shape similarity of the respective binding sites; Runs approx. 3 minutes on a database of 10,000 pockets

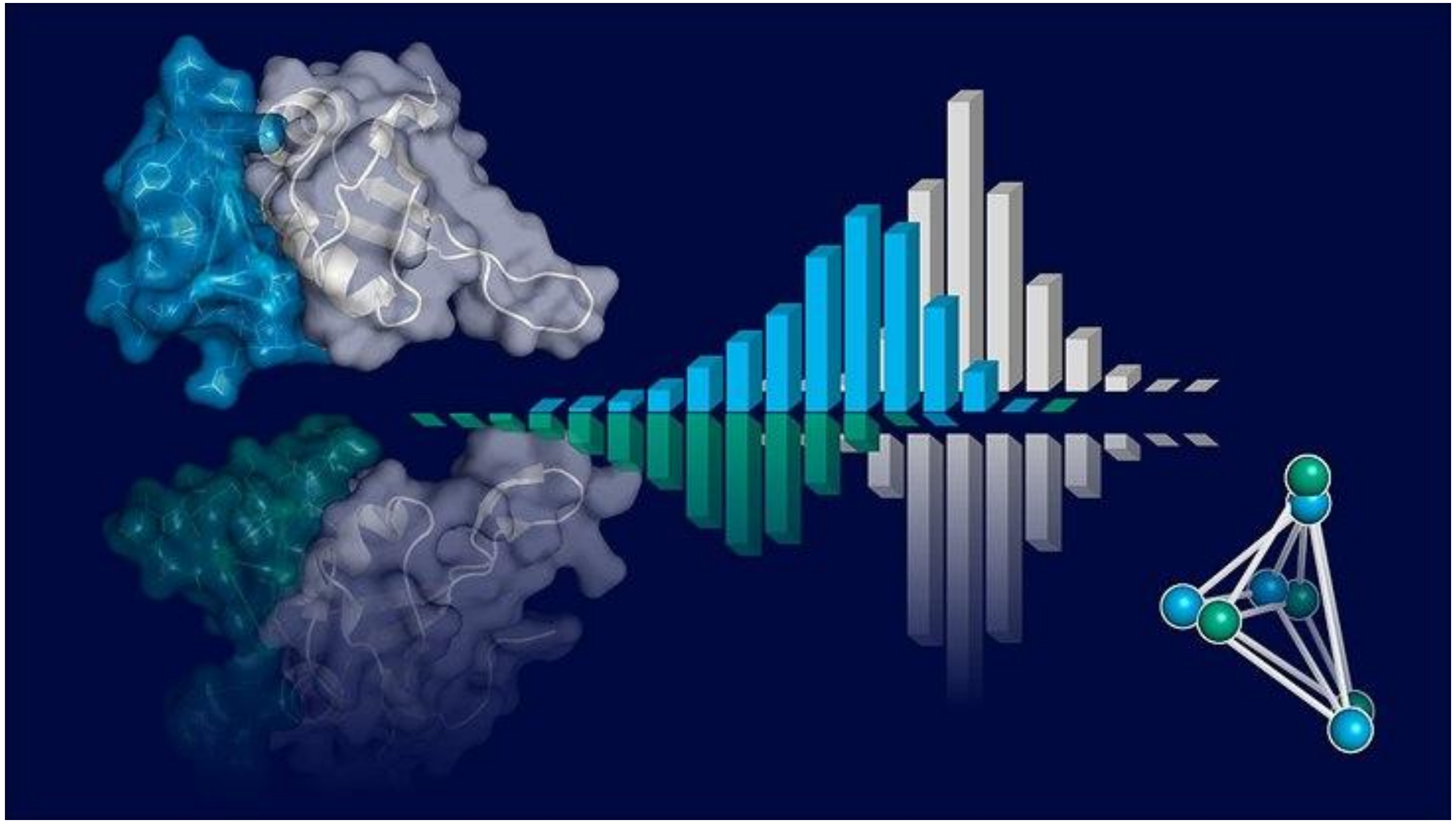


Main developers: *Thorben Reim, Joel Graef*

Reim et al., Arch Pharm (2024), 357(5):2300661

PiMine: PROTEIN INTERFACE COMPARISON IN COLLECTIONS OF PROTEIN-PROTEIN INTERFACES

Mine for Similarities of a Predicted or Known Protein-Protein Interface in Relevant Protein-Protein Complexes. Predicted single-chain interfaces or known protein-protein interfaces can be used in this mining approach based on GeoMine databases; Databases of individual experimental or predicted structures can be created automatically; Reliably calculate the shape and physicochemical similarities to known interfaces with a completely sequence-independent approach; A comparison to approx. 170k biologically relevant interfaces takes about one day

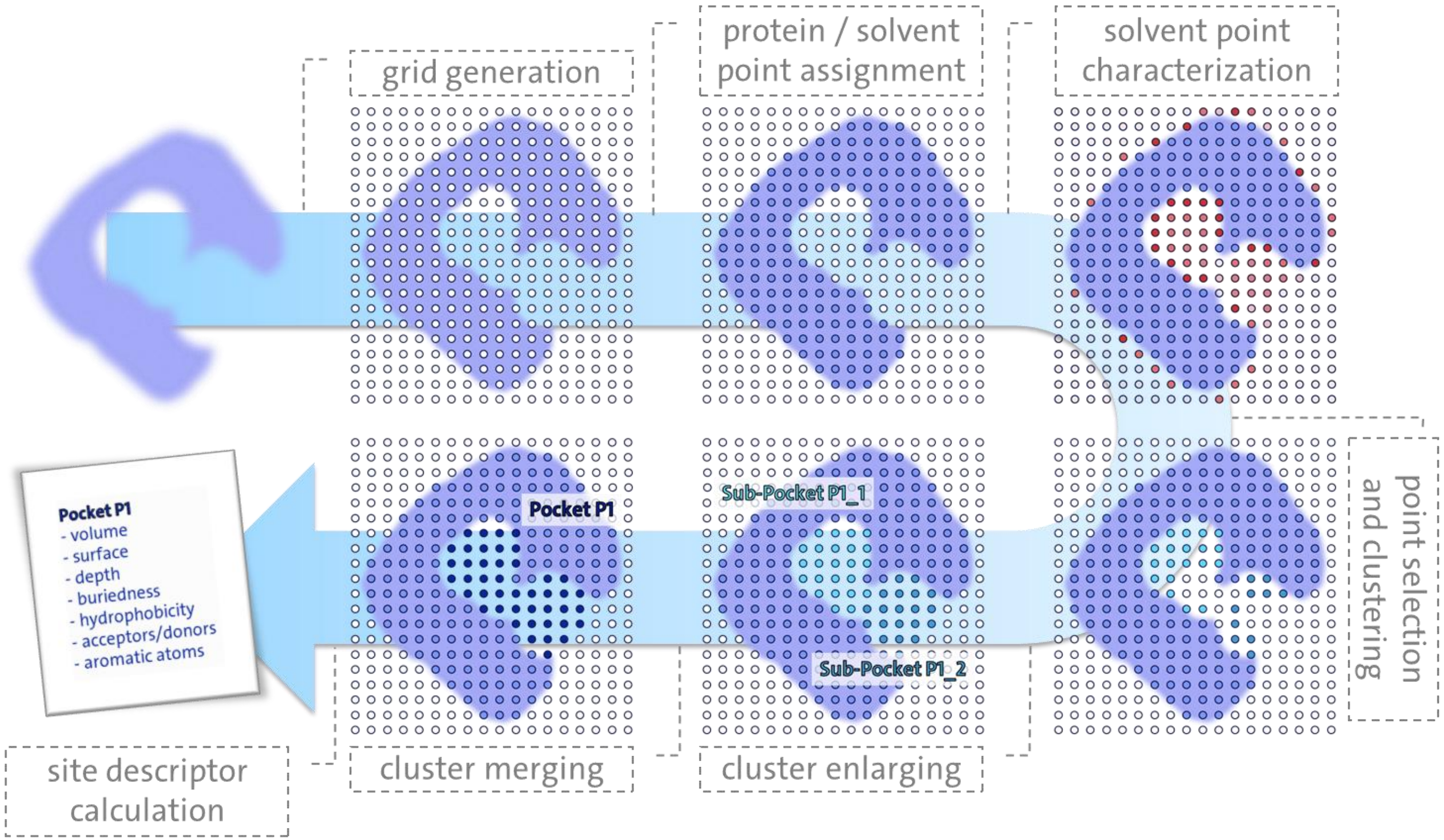


Main developers: *Joel Graef, Thorben Reim*

Graef et al., J Chem Inf Model (2024), 64(8):3332-49

DoGSite3: BINDING POCKET PREDICTION AND SITE DESCRIPTOR CALCULATION

Predict Robust Reliable Small Molecule Binding Sites and Their Geometrical and Chemical Descriptors. Based on the grid-based DoGSite algorithm for the prediction of pockets and their sub-pockets; Largely rotation- and translation-invariant due to a normalization procedure before binding site prediction; Known ligands in the structure can be used to bias the grid by sufficiently buried fragments; Novel chemical binding site descriptors considering solvent-accessibility; Increased robustness through comprehensive parameter optimization; Runs finish within seconds



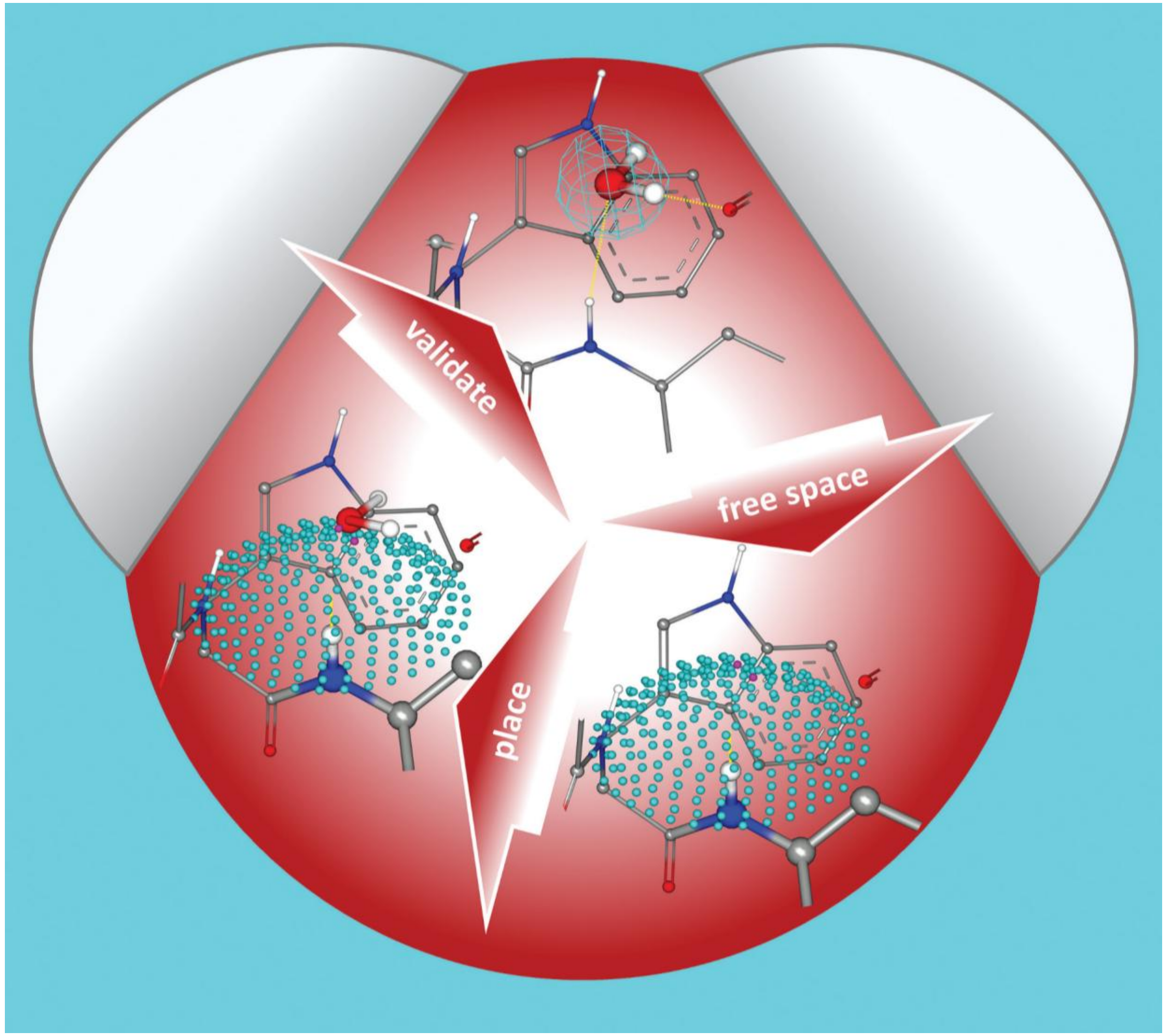
Main developers: *Joel Graef, Andrea Volkamer*

Graef et al., J Chem Inf Model (2023), 63(10):3128-37

WarPP: PLACING WATER MOLECULES IN THE BINDING SITES OF PDB FILES

Predict the Position and Orientation of Water Molecules in Small Molecule Binding Sites. WarPP places and scores water molecules in binding sites of crystallographic structures based on EDIAscorer results and interaction geometries as known from experimentally solved protein structures; Validated on a high-quality set of 1,500 protein-ligand complexes with 20,000 crystallographically observed water molecules; Sufficiently fast for high-throughput analyses; Correct placement in approx. 80% of the cases; Predictions can be exported as PDB files for, e.g., molecular docking with JAMDA

Main developers: *Eva Nittinger, Florian Flachsenberg*

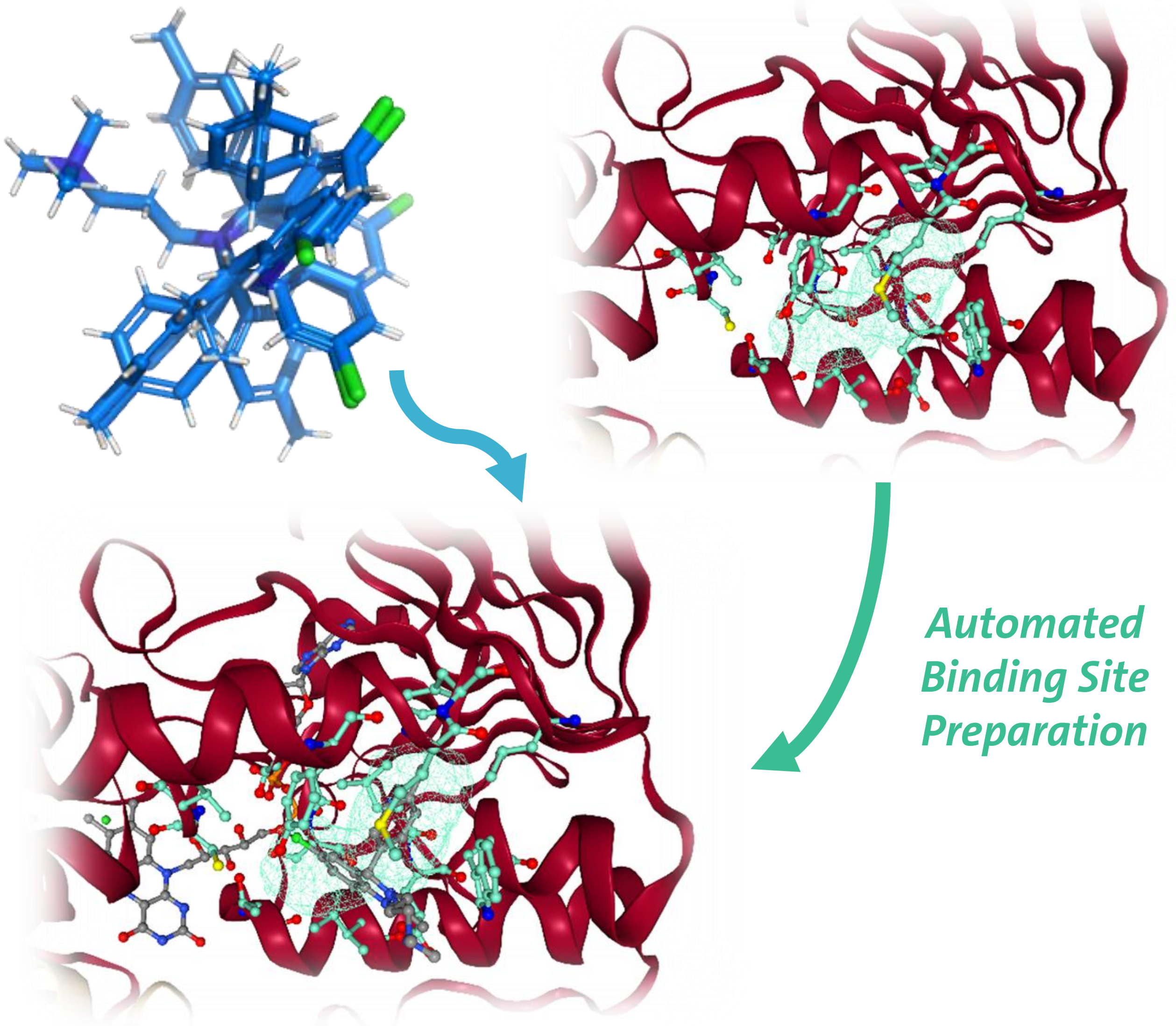


Nittinger et al., J Chem Inf Model (2018), 58(8):1625-37

JAMDA: TARGET PREPARATION AND PROTEIN-LIGAND DOCKING

Prepare Individual Protein Structures and Dock Small Molecules in a Pre-processed Binding Site of Choice. JAMDA simplifies the process of protein-ligand docking by automatic pre-processing protocols for the protein and binding sites of interest; The JAMDA scoring function enables the retrieval of 75% of the native poses in the three highest ranked solutions for high-quality complexes; Individual configurations for protein preparation are available, e.g., considering protein ensembles, binding site water molecules, or cofactors; An individual number of input conformations for the ligands of interest can be generated fully automated using the Conformator

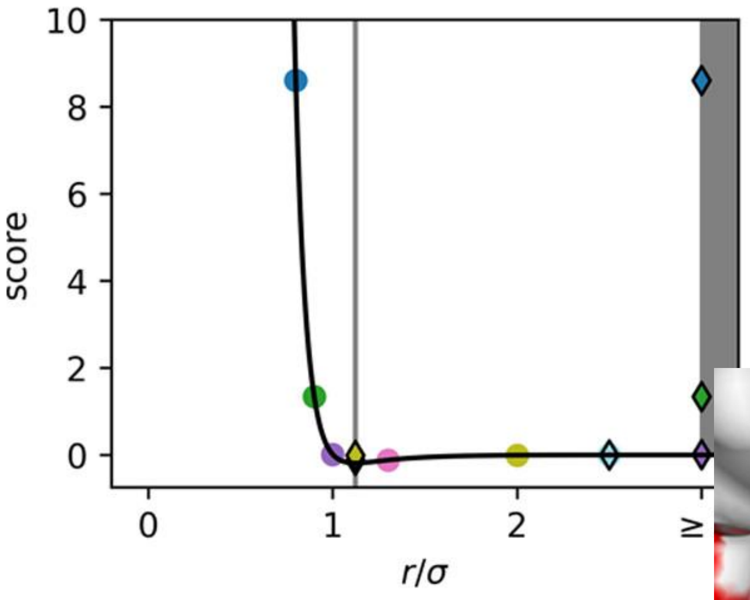
Main developer: *Florian Flachsenberg*

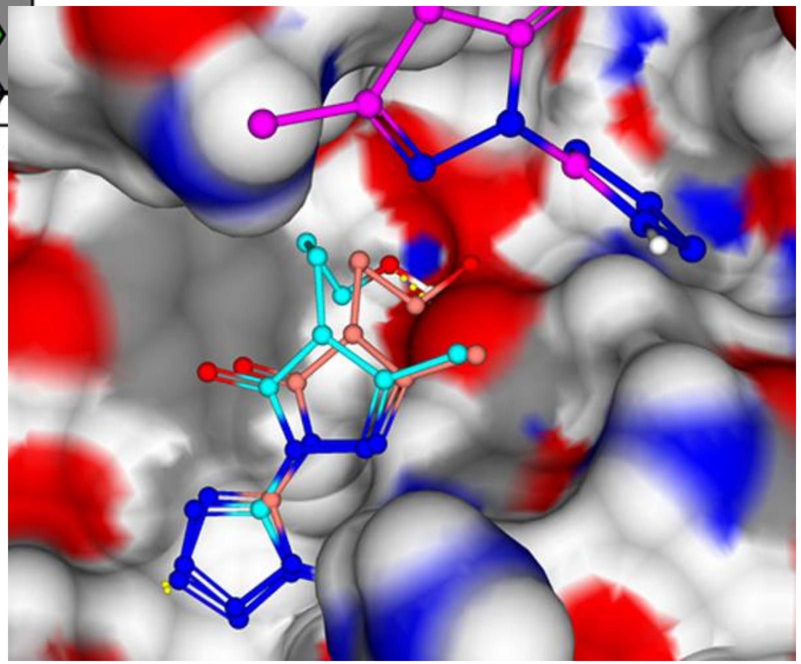


Flachsenberg et al., J Chem Inf Model (2023), 64(1):219-37

JamdaScorer: GRADIENT-BASED OPTIMIZATION OF PROTEIN-LIGAND POSES

Post-Optimize Protein-Ligand Complexes into Precise Local Energy Minima with the JAMDA Scoring Function. JamdaScorer is a new empirical scoring function for use with gradient-based optimizers. Combined with a newly developed optimizer, LSL-BFGS, fast convergence, locality and precise detection of local minima can be guaranteed; JamdaScorer considers hydrogen bond geometries, hydrophobic contacts, clashes, and torsion angles therefore balancing the most important terms in protein-ligand scoring functions



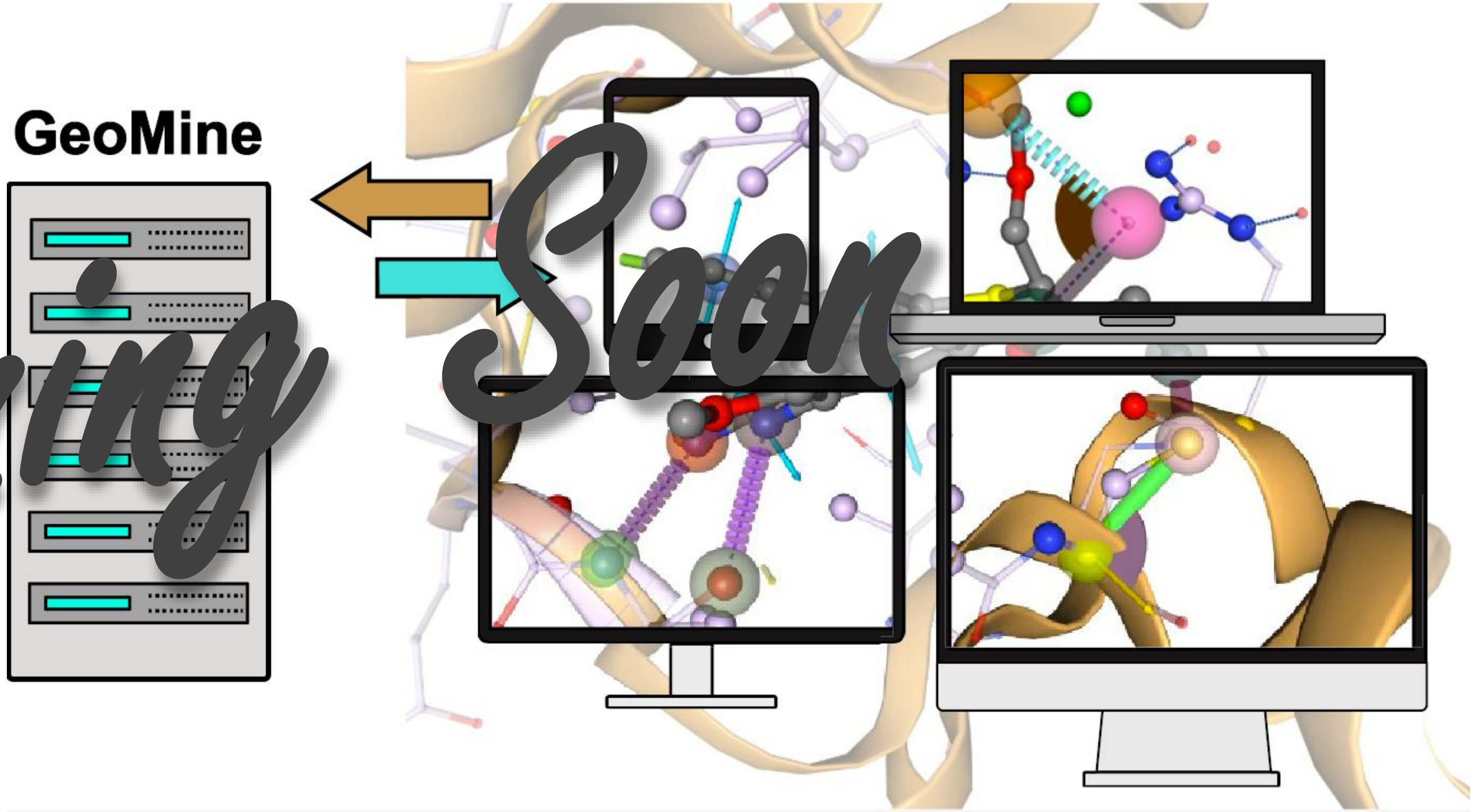


Main developers: *Florian Flachsenberg, Agnes Meyder, Kai Sommer, Patrick Penner*

Flachsenberg et al., J Chem Inf Model (2020), 60(12):6502-22 | Flachsenberg et al., J Comput Chem (2021), 42(15): 1095-100

GeoMine-C: SEARCHING GEOMETRIC PATTERNS IN PROTEIN-LIGAND INTERFACES

Mine Protein-Ligand Binding Sites. Rapid searching for spatial interaction patterns in large collections of protein-ligand complexes and binding pockets; Fully-automated database generation from PDB and mmCIF; Built on the free database systems SQLite and PostgreSQL; supports radius-based pockets (based on ligands and predicted pockets (based on DoGSite3); Flexible query management based on XML (for the REST service) or JSON in the GUI mode; Outputs statistics about matching and hit structures superimposed on the query in PDB format



Main developers: *Joel Graef, Konrad Diedrich, Martin Poppinga*

Diedrich et al., J Comput Aided Mol Des (2024), 38(1):23 | Graef et al., J Med Chem (2022), 65(2):1384-95
Diedrich et al., Bioinformatics (2021), 37(3):424-5

PoseEdit-C: GENERATING AND EDITING PROTEIN-LIGAND INTERACTION DIAGRAMS IN 2D

Generate and Modify 2D Diagrams of Protein-Ligand Complexes. Automated PoseView-based protocol for depicting protein-ligand interactions (hydrophobic contacts, polar contacts, aromatic interactions, metal coordination) by high-quality 2D images within seconds; Easy-to-use browser interface; Also supports nucleic acid-ligand complexes; Editing options include translation, rotation, the addition of residues and interactions, the rotation along bonds and user-specified axes, changing and moving labels, adding water-mediated contacts; Diagrams can be exported in SVG format

Main developers: *Konrad Diedrich, Katrin Schöning-Stierand*

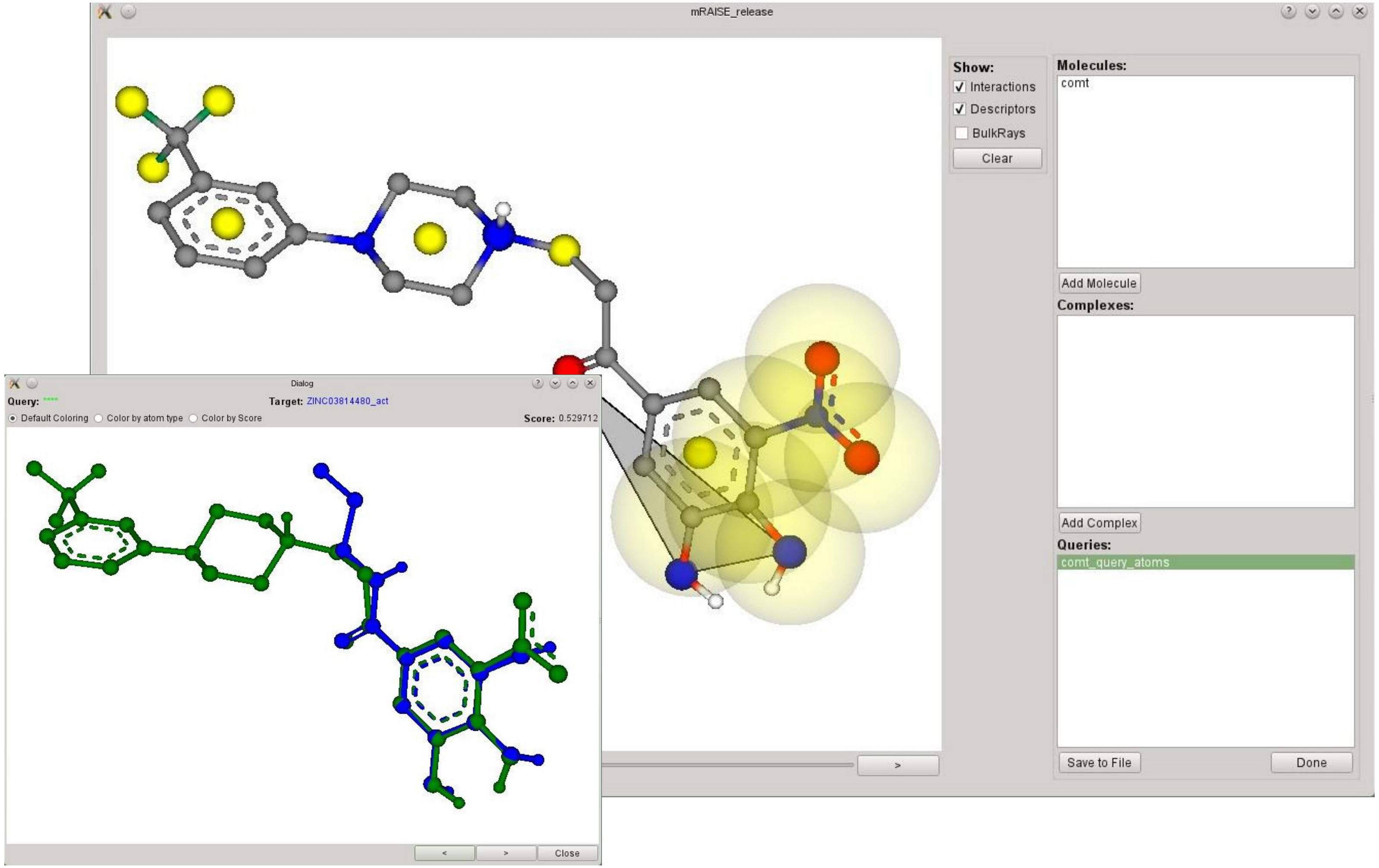
Diedrich et al., J Comput Aided Mol Des (2023), 37(10):491-503

DEPRECATED TOOLS

These software tools remain available for download. However, they will not be updated anymore. The reasons could be manifold, in most cases: the dependence on deprecated external libraries, the replacement by newer software tools already available or on their way, low usage, extremely high maintenance efforts.

mRAISE: LIGAND-BASED VIRTUAL SCREENING WITH USER-DEFINED PARTIAL SHAPE CONSTRAINTS

3D Similarity Searches with Individual Constraints. Fast ligand-based virtual screening of large compound collections using the RAISE index technology; Screen compound libraries for molecules, e.g., with a certain percentage of shape similarity, fulfilling complex derived constraints, or matching your own manual selection of partial shape constraints; Find similar structures and generate accurate alignments

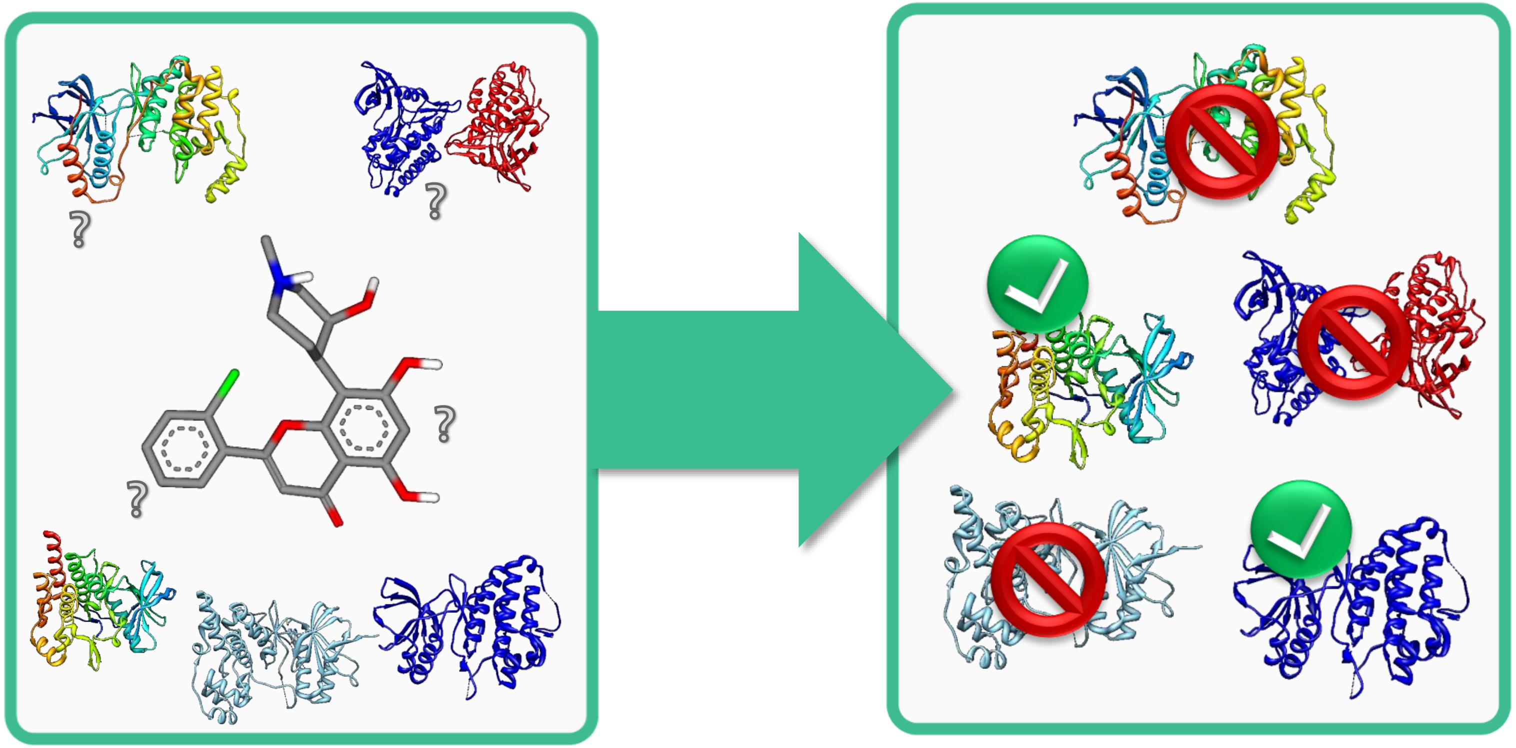


Main developer: *Matthias von Behren*

von Behren et al., J Comput Aided Mol Des (2016) 30(8):583-94

iRAISE: INVERSE VIRTUAL SCREENING AND STRUCTURE-BASED TARGET PREDICTION

Search through Thousands of Protein Structures for Potential Binding Sites of Small Molecules. Find potential binding sites for a small molecule to explain off-target effects or eestablish polypharmacology; Prepare and use a relational protein structure database for fast access to structures; Prepare proteins in a fully-automated fashion; Score potential small molecule binding sites with an innovative multi-step pose scoring procedure



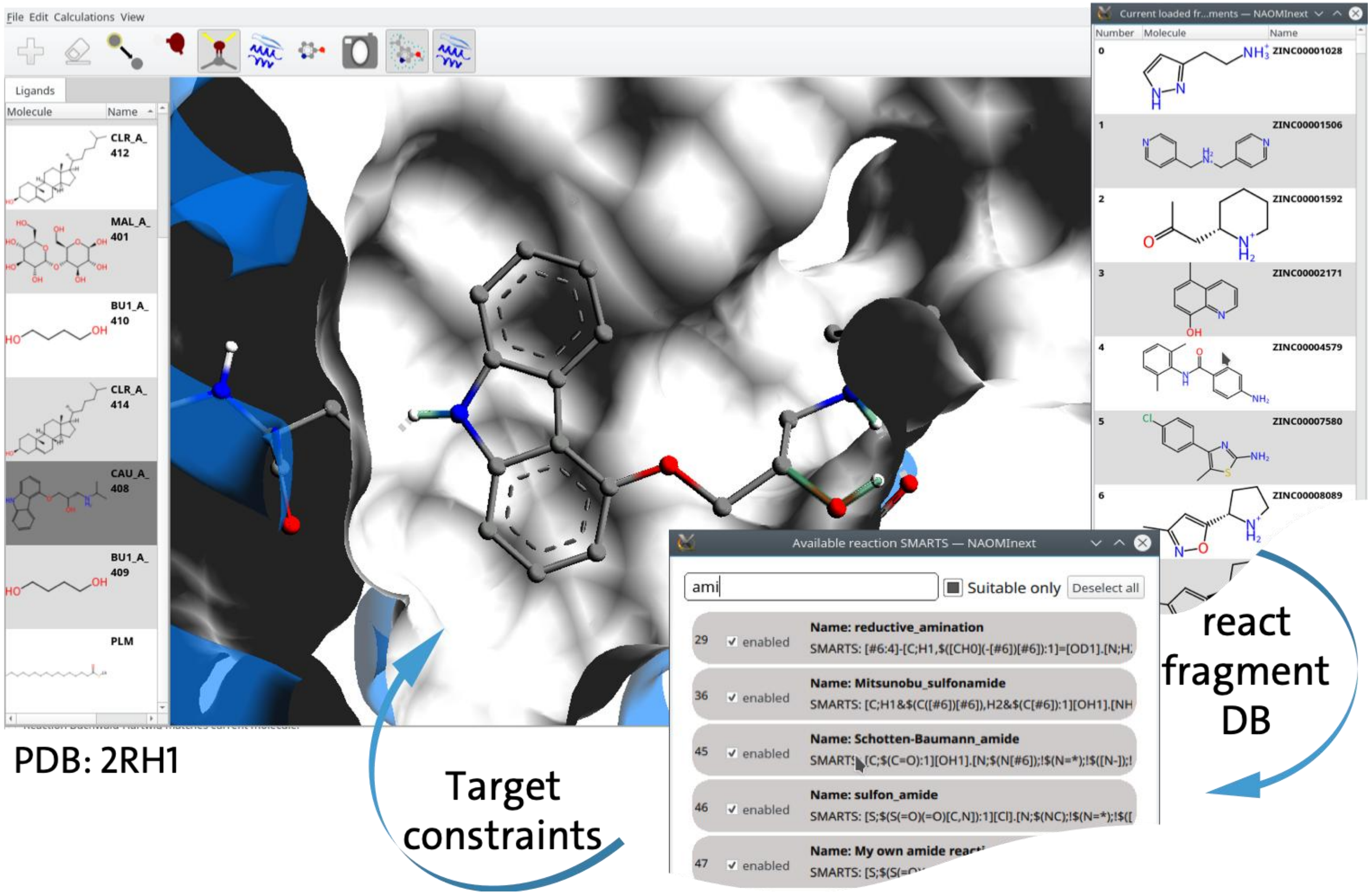
Main developer: *Karen Schomburg*

Schomburg et al., J Chem Inf Model (2014), 54(6):1676-86 | Schomburg et al., J Chem Inf Model (2014), 54(8):2261-74

NAOMInext: SYNTHETICALLY FEASIBLE FRAGMENT GROWING

Forward Synthesis Planning and High-Throughput Growing. Grow ligands into binding sites based on crystallized or docked fragments; Create a target-focused library of your in-house or vendor catalogue of fragments; Use 58 published and incorporated robust reaction rules in SMIRKS format; Provide your own reactions which are checked for validity and consistency prior to usage; Visually inspect reaction vectors; Add residue constraints to guide the growing process into a specific sub-pocket

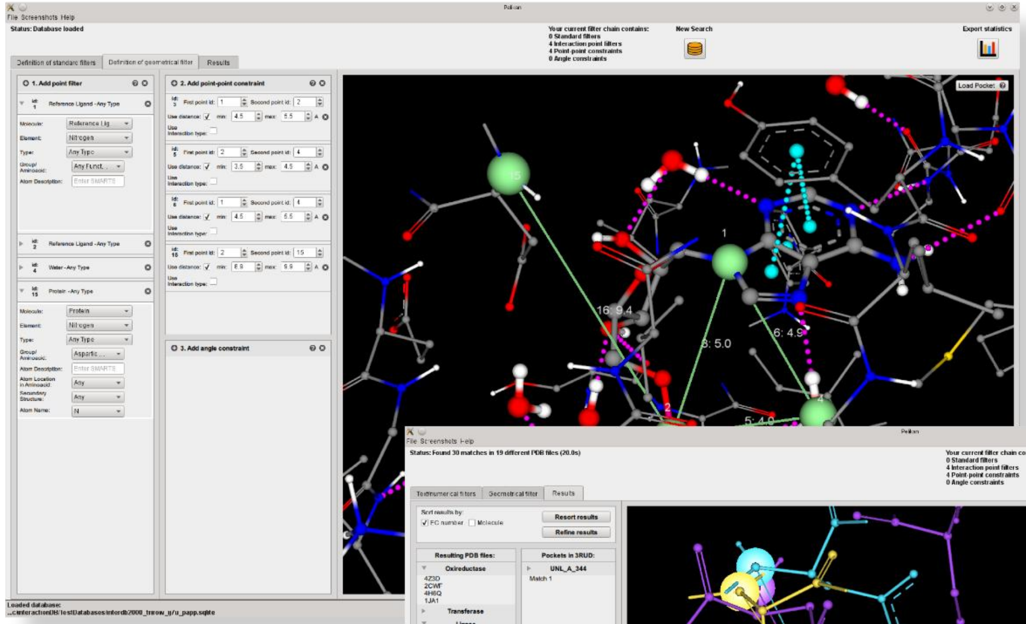
Main developer: *Kai Sommer*

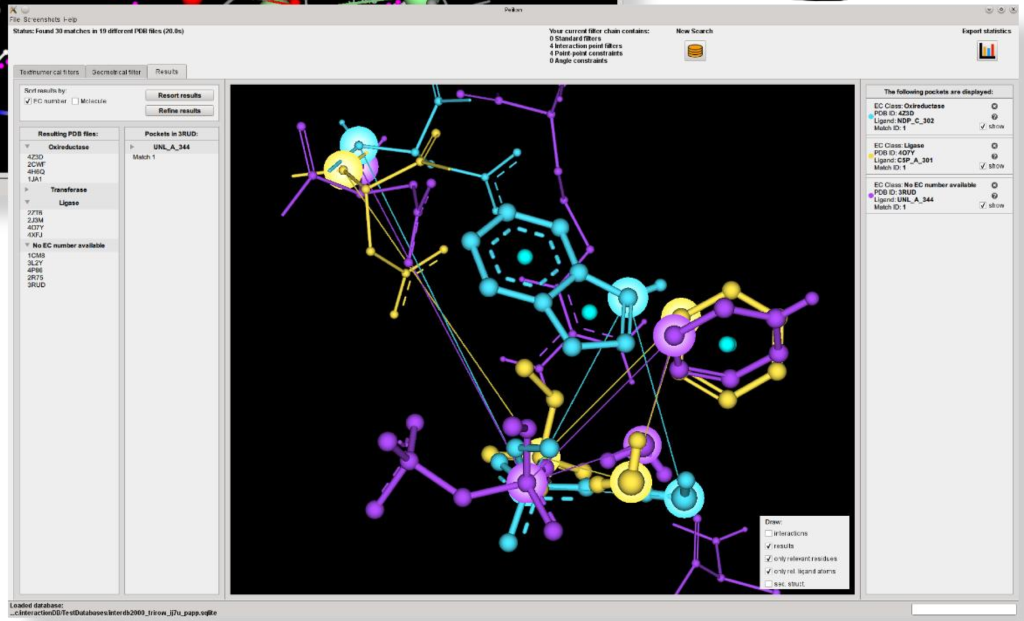


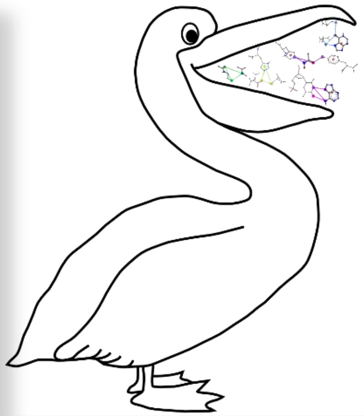
Sommer et al., Eur J Med Chem (2019), 163, 747-62

PELIKAN: SEARCHING INTERACTION PATTERNS

Find and Analyze All Occurrences of Individual Interaction Patterns. Create a database of your collection of protein-ligand complexes or search in the whole PDB; Define your spatial constraints in 3D starting from a binding site of interest or from scratch; Combine 3D searches with textual and numerical filters for various ligand, protein pocket, and complex properties; Rapidly find all occurrences of interaction patterns; browse and refine your results in a 3D viewer; Precompiled PELIKAN SQLite databases available for PDB, sc-PDB, and several target classes





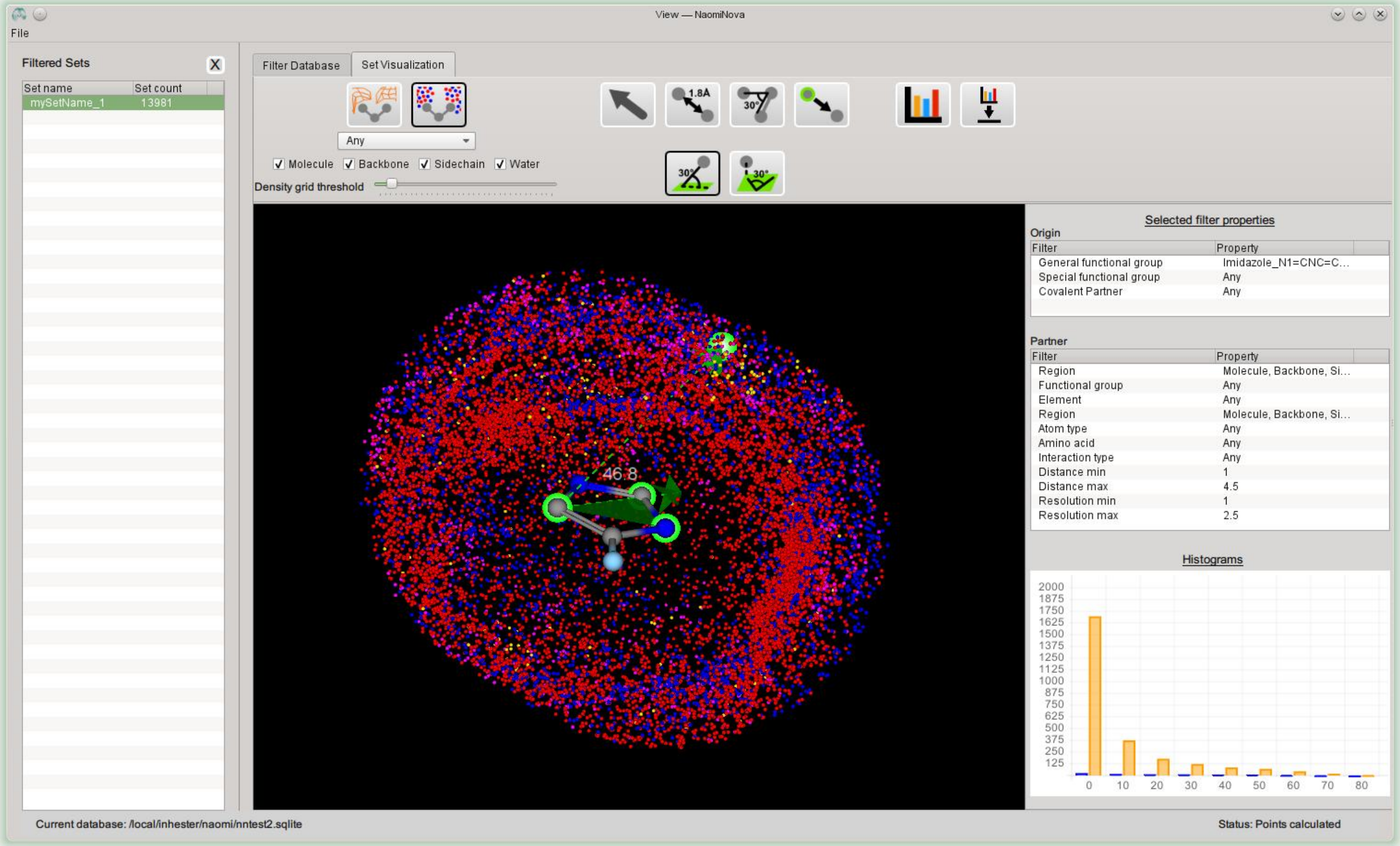


Main developer: *Therese Inhester*

Inhester et al., J Chem Inf Model (2017), 57(2):148-58 | Inhester et al., Wiley Interdiscip Rev: Comput Mol Sci (2014), 4:562-75

NAOMInova: INTUITIVE ANALYSIS OF INTERACTION GEOMETRIES

Determine Optimal Geometries for Non-Covalent Interactions. Create a database of your collection of protein-ligand complexes; Query the database for atoms found in the vicinity of functional groups; Visualize all detected atoms or density clouds; Filter and inspect all detected atoms in the proximity of the functional group (incl. electron density support values from EDIAscorer); Click on any data point to inspect the original protein-ligand complex and measure distances and angle distributions; Add your individual functional group descriptions



Main developers: *Therese Inhester, Eva Nittinger*

Inhester et al., J Chem Inf Model (2017), 57(9):2132-42 | Nittinger et al., J Med Chem (2017), 60(10):4245-57

ABOUT

NAOMI ChemBio Suite – FAQ

Will older software become part of the collection? We are integrating successful, older software tools into our collection as we already did for TorsionAnalyzer and DoGSite, developed with Merck and BioSolveIT. PoseEdit and GeoMine (available at https://proteins.plus) will also become standalone tools. Tools based on the FlexX and FTrees software cannot be integrated into the suite for legal reasons.

Will you charge a substantially higher fee once the collection is extended? We do not intend to increase the license fees. We do this as an academic group and calculate prices to build and maintain a fund that assists in keeping the software alive. The licensing of our suite is not a for-profit business.

Will software tools disappear from the collection? We might remove tools if we cannot keep them running. Reasons might be that 3rd-party libraries are no longer available, substantial problems occur with OS updates, or we find a capital scientific flaw we cannot fix (which has never happened so far).

What level of support can you offer? We will help wherever we can. However, we are not a software company. Therefore, we cannot guarantee short response times or promise to fix problems and bugs that might occur in the future. We offer the software on an as-is level. Therefore, you can test everything before you decide.

Will the collection grow over time? Yes, we will continuously extend the tool portfolio. We want to add all software resulting from our research without exclusivity constraints (which sometimes happens in the case of 3rd-party funding). We have a lot of cool stuff in the pipeline in all three research areas, i.e., cheminformatics, structure-based molecular design, and visual analytics in chemistry.

Why is our software not available for free?

We are strong believers in the policy that academic software should be free for academic research and evaluation purposes but should not be free for commercial applications. You might disagree, but here are our arguments:

- PhD students are not paid to create software. They are not even paid for research most frequently. Salaries, even in rich countries like Germany, are much below that of a software developer. License fees for our software are exclusively used to support academic research in the AMD group, i.e.:
  - To improve the salary of PhD students
  - To give travel grants to PhD students for participating in international conferences
- Creating and licensing software is a valid business. Thousands of computational chemists and computer scientists worldwide work for enterprises offering software and related services. Free software destroys, or at least harms, this business.
- Academic software should be designed and developed sustainably for several reasons. First and foremost, scientific results should be reproducible. Second, the software production should be economic. Third and most important, software development during a PhD project is an important training on the job. Producing sustainable software is, however, not for free. Substantial efforts are necessary to establish the required infrastructure for reviewing, testing, cross-compiling, and achieving the necessary code quality and documentation. License fees help to support this endeavor.

License Model

- Tools are distributed as installable software packages for various Linux platforms, Windows and Mac OS
- Single registration, access to all tools via a unified web service
- Continuously extended by new software tools
- Regularly updated
- All tools available for free for academic use and evaluation purposes

All incoming fees are exclusively used to support PhD students and Postdocs participating in the development of the NAOMI ChemBio Suite.

License model for non-academic use:

- Single licenses for all tools
- Unlimited site license (unlimited number of users and number of CPUs)
- Annual license fee depending on company size:

License Type	# of employees	Fee (in k€)
very small	1-9	2
small	10-99	4
medium	100-999	8
large (single site)	$\geq$1000	16
large (all sites)	$\geq$1000	25

- Discounts for project partners

Contact

Prof. Dr. Matthias Rarey [matthias.rarey@uni-hamburg.de]

University of Hamburg
ZBH - Center for Bioinformatics [https://uhh.de/zbh]
Research Group for Computational Molecular Design [https://uhh.de/amd]

Software Server: https://uhh.de/naomi
Web Server: https://proteins.plus
https://smarts.plus
Source Code Libraries: https://github.com/rareylab