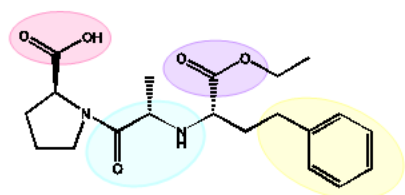
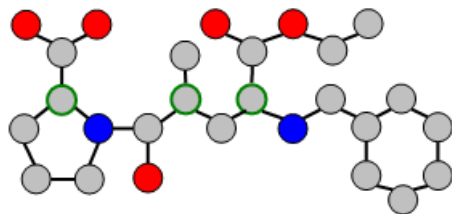


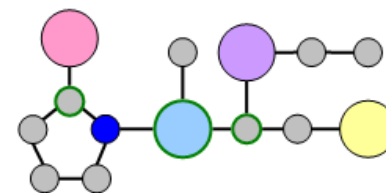
Topological Summaries: Using Graphs for Chemical Searching and Mining



Kekule structure

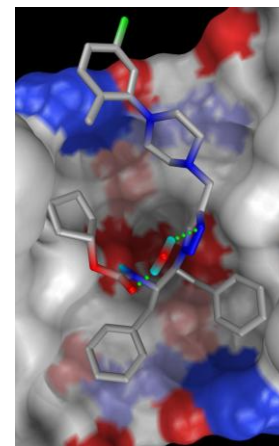


Graph



Graph with abstract nodes

- Graphs are a flexible & unifying model
- Scalable similarity searches through novel index structure
- Mining of significant fragments in collections
- Classification of compounds based on significant fragments
- Development of SAR models through geometric mining

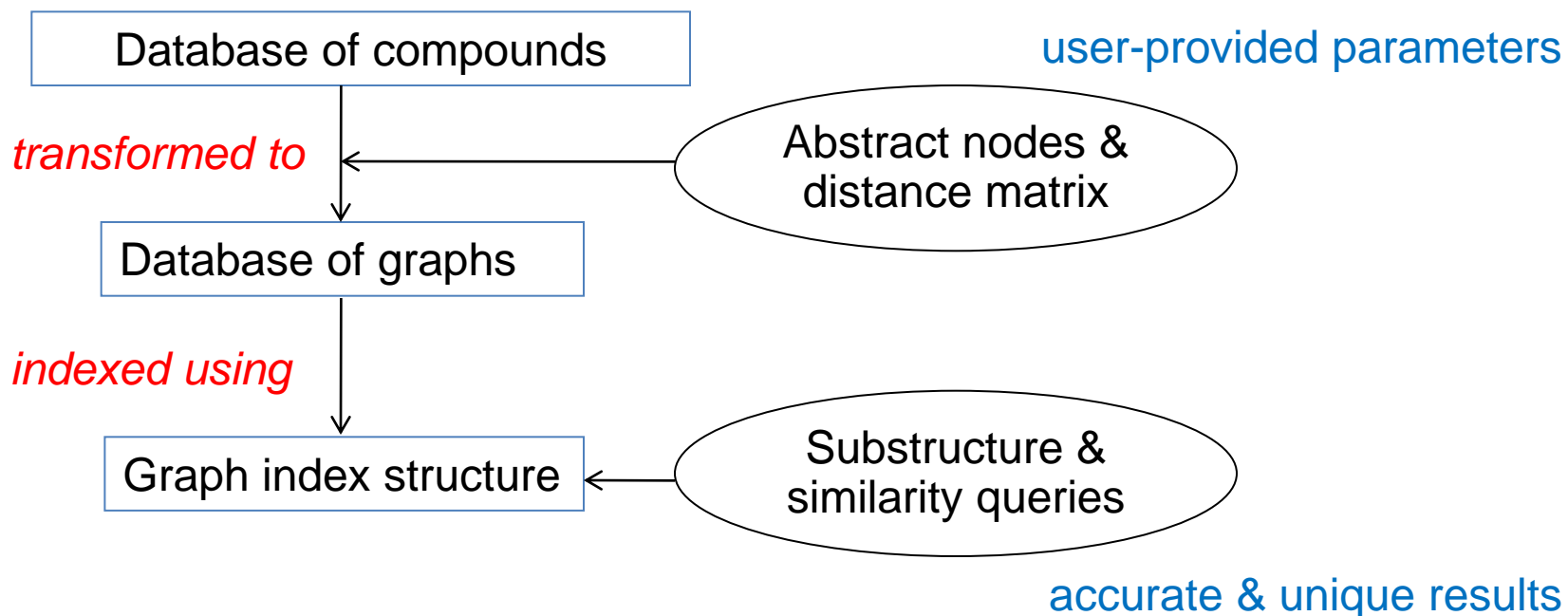


Protein-ligand interaction

Feature/Benefit Summary

| Feature | Benefits | Novel technology |
|--|--|--|
| Similarity Searching (SimFinder tool) | <ul style="list-style-type: none">• Improved hit identification at lower cost• Better comparison to known drugs | <ul style="list-style-type: none">• Topology-based searching• Scalable index structure |
| Active substructure identification (SigFinder tool) | <ul style="list-style-type: none">• Identification of active substructures for lead identification and optimization• Mining of protein-ligand complexes | <ul style="list-style-type: none">• Significance model based on graph analysis |
| Pharmacophore analysis & development (PharmaMiner tool) | <ul style="list-style-type: none">• Better pharmacophore analysis & models• Fragment based lead discovery | <ul style="list-style-type: none">• Selective queries on multiple targets• Joint 3D pharmacophore space |

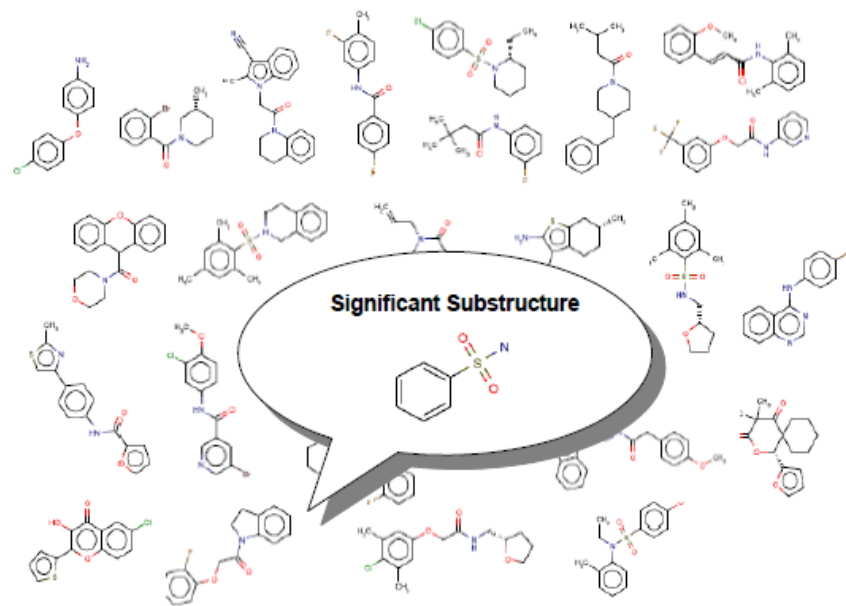
SimFinder (Similarity Searching)



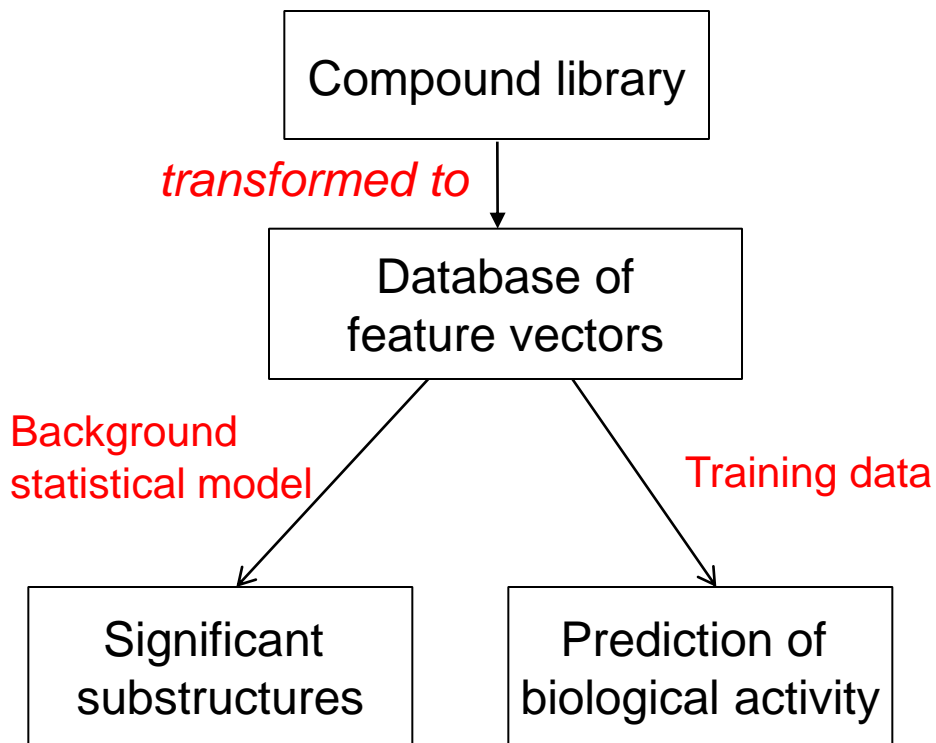
- Accuracy and scalability validated through retrospective studies
- Ongoing assays on beta-secretase (Alzheimer's)
- Comparison of hits with alternative methods such as fingerprints

SigFinder (Mining Significant Substructures)

- Statistics-based modeling of substructure significance
- Validation by identification of substructures specific to biological activity



Pictorial view of substructure mining



PharmaMiner (Joint 3D Analysis of Pharmacophore Space)

- We define a Joint Pharmacophore Space by extracting 3D descriptors from diverse libraries/targets
- Analysis of pharmacophores in the Joint Space
 - Classification and clustering for understanding biological activity
 - Screening for compounds based on presence and/or absence of activity
- Fragment-based discovery and scaffold hopping
- Flexible definition of pharmacophoric features (donor, acceptor, hydrophobic core, etc.)

References

- Huahai He; Ambuj K. Singh; Closure-tree: An Index Structure for Graph Queries. Proceedings of the 22nd International Conference on Data Engineering (ICDE), April, 2006, pp 38 - 50 [DOI :10.1109/ICDE.2006.37](https://doi.org/10.1109/ICDE.2006.37).
- Huahai He; Ambuj K. Singh; *GraphRank: Statistical Modeling and Mining of Significant Subgraphs in the Feature Space*. Proceedings of the 6th IEEE International Conference on Data Mining (ICDM), December, 2006, [doi:10.1109/ICDM.2006.79](https://doi.org/10.1109/ICDM.2006.79)
- Sayan Ranu and Ambuj Singh, [GraphSig: A Scalable Approach to Mining Significant Subgraphs in Large Graph Databases](#) in *25th International Conference on Data Engineering (ICDE)*, 2009.
- Sayan Ranu; Ambuj Singh; *Mining Statistically Significant Molecular Substructures for Efficient Molecular Classification.*, J. Chem. Inf. Model., 2009, 49(11), pp 2537–2550 [DOI : 10.1021/ci900035z](https://doi.org/10.1021/ci900035z)
- Huahai He; Ambuj K. Singh; [Graphs-at-a-time: Query Language and Access Methods for Graph Databases](#), SIGMOD, June, 2008, pp 405-418. ISBN: 978-1-60558-102-6