

GALAHAD™

RAPID, HIGH QUALITY PHARMACOPHORIC PERCEPTION AND MOLECULAR ALIGNMENTS



GALAHAD aligns a set of molecules that share a common mode of biological activity, and develops a corresponding pharmacophore hypothesis. Using a sophisticated new genetic algorithm and a multi-objective scoring function, GALAHAD takes into account energetics, steric similarity, and pharmacophoric overlap, while accommodating conformational flexibility, ambiguous stereochemistry, alternative ring configurations, multiple partial match constraints, and alternative feature mappings among molecules. Pharmacophore models are returned as hypermolecules, which contain information from every molecule in the training set, as well as a 3D search query that can be used to probe databases for new structures that match the model. New target molecules that were not included in the training set can be fit to the model, yielding scores for energy as well as steric and pharmacophoric similarity that relate directly to ligand affinity.

Advantages

- Pareto multi-objective optimization is used to simultaneously balance steric, pharmacophoric, and energy information to build the most valuable hypermolecule models required, so models are unbiased
- Run time scales linearly with the number of ligands, unlike other methods
- Partial match and partial coverage models can be created in a timely manner

GALAHAD allows researchers to automatically develop pharmacophore hypotheses and structural alignments from a set of molecules that bind at a common site. No prior knowledge of pharmacophore elements, constraints, or molecular alignment is required, making it ideal for exploring new targets and new modes of action.

GALAHAD uses a sophisticated new genetic algorithm (GA) that defines each molecule as a core structure plus a set of torsions. To overcome limitations in existing pharmacophore tools, GALAHAD's genetic algorithm was developed on real-world data sets. Pre-processing (or pre-generated conformers) provide torsional biases that speed up calculations, while yielding less strained ligand conformations in output pharmacophore models.

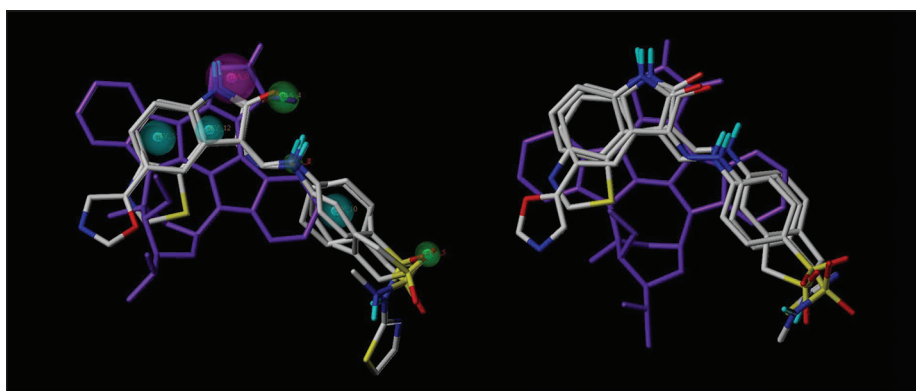
Torsions are then applied to the core structure, and 3D similarity among the ligands is rapidly measured using pharmacophoric and steric multiplets. Since each molecule is simultaneously compared to every other molecule, specification of a template molecule is not required and the results are not biased towards any structure. Using pharmacophoric and steric similarity instead of one-to-one feature mapping allows the calculations to be carried out in internal coordinate space, which produces significantly faster convergence of the GA.

Each set of conformers generated is then evaluated for corresponding features, and correspondences are filtered based on geometric consistency for subsequent alignment in Cartesian space. These steps of pharmacophore detection and conformation selection, followed by alignment and generation of a 3D structure

query, are carried out using an extension of the LAMDA algorithm².

One of GALAHAD's advantages is its ability to generate queries containing partial match constraints. Instead of rigid requirements that must all be met simultaneously, GALAHAD's hypothesis can include constraints where most of the features are required, or where only some are required for the molecule to be considered a hit. This flexibility in pharmacophore description more accurately models the real world, and allows significantly more flexibility in retrieving new potentially active compounds from structural libraries.

Pareto scoring within the GA is used to balance the tradeoff among the conflicting demands of maximizing pharmacophore consensus, maximizing steric consensus, and minimizing energy. In general, this



GALAHAD model derived from four cyclin-dependent kinase (CDK2) inhibitors (left) vs. the overlay based on the corresponding X-ray crystal structures (right).

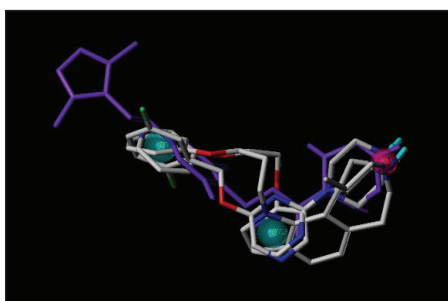
Applications

- Generate pharmacophore hypotheses from a set of ligands
- Determine functional group correspondences for molecules bound to the same receptor
- Align groups of molecules for 3D-QSAR studies
- Develop database queries to find new molecules that fit the pharmacophore model
- Test the fit of ligands to a pharmacophore model
- Align molecules to other single molecules, hypermolecules, UNITY queries, or combinations

involves exploring alternative feature mappings and conformations. Pareto scoring decouples competing measures of model quality, making exploration of solution space much more efficient. A single Pareto run produces better, more diverse models than an entire series of GA runs using a range of fitness term weights.³

The output of a GALAHAD run is a set of hypermolecules and pharmacophore models, each of which contains data from all molecules in the input set. This data is implicit in the ligand alignment and overlapping pharmacophoric features. The models returned are ranked by their Pareto scores, and can then be examined and refined further. The calculations can be resumed if preliminary runs show promising results.

Each hypermolecule includes a ligand alignment and common pharmacophoric features in the form of a UNITY[®] query that can be used to search 3D databases for new molecules that fit the model. The new ligands identified can then be fit to the model, or be used to further refine it.



GALAHAD alignment for four 5-HT2C antagonists.

Features

- Intuitive control and straightforward use
- Extensive parameter control for more sophisticated users
- Flexible and partial constraints included in generated models
- Alternative ring configurations and stereochemistry are supported
- Extensive control of genetic algorithm and alignment parameters
- Multi-objective Pareto scoring function
- Goodness-of-fit descriptors including energy, steric and pharmacophoric similarity, and conformance to the model query are produced for each ligand

Hardware and Software Requirements

GALAHAD is accessible through the SYBYL[®] expert molecular modeling environment and requires GALAHAD and SYBYL licenses. SYBYL and GALAHAD run on workstations operating under IRIX[®] (SGI[®]) or Linux[®] (x86).

Acknowledgements

Components of GALAHAD were developed in Tripos collaborations with the University of Sheffield, Novo Nordisk A/S, and Biovitrum AB.

Complementary Software

- **UNITY** for rapid, flexible 3D searching of databases to identify lead compounds based on a pharmacophore hypothesis.
- **QSAR with CoMFA[®]** for constructing predictive structure-activity models from sets of aligned molecules.
- **Selector[™]** for evaluating the diversity of compound databases and selecting diverse subsets.

- **Tuplets[™]** for pharmacophore-based virtual screening without a 3D model.
- **Confort[™]** for generating sets of molecular conformers.
- **ProtoPlex[™]** for rapid generation of protomers and tautomers.
- **StereoPlex[®]** for expanding the stereochemical diversity of a chemical database.
- **Legion[™]/CombiLibMaker[™]** for building virtual combinatorial libraries in cSLN format.
- **VolSurf[™]** for predicting ADME properties of compounds based on pre-calculated models.
- **OptDesign[®]** for designing and editing combinatorial libraries.
- **Concord[®]** for rapidly converting 2D chemical structures into high quality 3D structures.

References

1. Clark, R.D.; Abrahamian, E.; Abrams, C.; Brandt, P.; Gustavsson, A.-L.; Homan, E.; Strizhev, A.; Wirstam, M.; Wolohan, P. "Genetic Algorithm with Linear Assignment for Hypermolecular Alignment of Datasets." manuscript in preparation.
2. Richmond, N.J.; Willett, P.; Clark, R.D. "Alignment of three-dimensional molecules using an image recognition algorithm." *J. Mol. Graph. Model.* **2004**, *23*, 199-209.
3. Cottrell, S.J.; Gillet, V.J.; Taylor, R.; Wilton, D.J. "Generation of multiple pharmacophore hypotheses using multiobjective optimisation techniques." *J. Comput.-Aided Drug Design* **2004**, *18*, 665-682.

WWW.TRIPOS.COM

CONTACT_US@TRIPOS.COM



AUSTRALIA
+61 (7) 5439 9775

CANADA
+1 450 4334500

FRANCE
+33 1 69 59 29 49

GERMANY
+49 89 45 10 300

JAPAN
+81 3 5166 1721

UNITED KINGDOM
+44 1 908 650000

UNITED STATES
800 323 2960
+1 314 647 1099