Journal of Cheminformatics

POSTER PRESENTATION Open Access

Kernel learning for ligand-based virtual screening: discovery of a new PPARγ agonist

Matthias Rupp[1*], T Schroeter[2], R Steri[1], Ewgenij Proschak[1], K Hansen[2], H Zettl[1], O Rau[1], M Schubert-Zsilavecz[1], K-R Müller[2], Gisbert Schneider[1]

From 5th German Conference on Cheminformatics: 23. CIC-Workshop
Goslar, Germany. 8-10 November 2009

We demonstrate the theoretical and practical application of modern kernel-based machine learning methods to ligand-based virtual screening by successful prospective screening for novel agonists of the peroxisome proliferator-activated receptor γ (PPARγ) [1]. PPARγ is a nuclear receptor involved in lipid and glucose metabolism, and related to type-2 diabetes and dyslipidemia. Applied methods included a graph kernel designed for molecular similarity analysis [2], kernel principle component analysis [3], multiple kernel learning [4], and, Gaussian process regression [5].

In the machine learning approach to ligand-based virtual screening, one uses the similarity principle [6] to identify potentially active compounds based on their similarity to known reference ligands. Kernel-based machine learning [7] uses the "kernel trick", a systematic approach to the derivation of non-linear versions of linear algorithms like separating hyperplanes and regression. Prerequisites for kernel learning are similarity measures with the mathematical property of positive semidefiniteness (kernels).

The iterative similarity optimal assignment graph kernel (ISOAK) [2] is defined directly on the annotated structure graph, and was designed specifically for the comparison of small molecules. In our virtual screening study, its use improved results, e.g., in principle component analysis-based visualization and Gaussian process regression. Following a thorough retrospective validation using a data set of 176 published PPARγ agonists [8], we screened a vendor library for novel agonists. Subsequent testing of 15 compounds in a cell-based transactivation assay [9] yielded four active compounds.

The most interesting hit, a natural product derivative with cyclobutane scaffold, is a full selective PPARγ agonist (EC50 = 10 ± 0.2 μM, inactive on PPARα and PPARβ/δ at 10 μM). We demonstrate how the interplay of several modern kernel-based machine learning approaches can successfully improve ligand-based virtual screening results.

Author details
[1]University of Frankfurt, Siesmayerstr. 70, 60323 Frankfurt am Main, Germany. [2]Berlin, Germany.

Published: 4 May 2010

References

1. Henke B: *Progress in Medicinal Chemistry* 2004, **42**:1-53.
2. Rupp M, Proschak E, Schneider G: *J Chem Inf Model* 2007, **47(6)**:2280-2286.
3. Schölkopf B, Smola A, Müller K-R: *Neural Comput* 1998, **10(5)**:1299-1319.
4. Sonnenburg S, Rätsch G, Schäfer C, Schölkopf B: *J Mach Learn Res* 2006, **7(7)**:1531-1565.
5. Rasmussen C, Williams C: MIT Press, Cambridge 2006.
6. Johnson M, Maggiora M: Wiley, New York 1990.
7. Schölkopf B, Smola A: MIT Press, Cambridge 2002.
8. Rücker C, Scarsi M, Meringer M: *Bioorg Med Chem* 2006, **14(15)**:5178-5195.
9. Rau O, Wurglics M, Paulke A, Zitzkowski J, Meindl N, Bock A, Dingermann T, Abdel-Tawab M, Schubert-Zsilavecz M: *Planta Med* 2006, **72(10)**:881-887.

doi:10.1186/1758-2946-2-S1-P27
Cite this article as: Rupp *et al.*: Kernel learning for ligand-based virtual screening: discovery of a new PPARγ agonist. *Journal of Cheminformatics* 2010 **2**(Suppl 1):P27.

[1]University of Frankfurt, Siesmayerstr. 70, 60323 Frankfurt am Main, Germany