Closing the loop between synthesis and design: Balancing optimisation of potency with selectivity

Peter Hunt, Tamsin Mansley, Edmund Champness, Nicholas Foster & Matthew Segall
Optibrium Ltd., Cambridge, UK., peter.hunt@optibrium.com

Introduction

Drug discovery is a multi-parameter optimisation (MPO) process, in which the goal is to simultaneously optimise target potency, selectivity and a broad range of Absorption, Distribution, Metabolism, Excretion and Toxicity (ADMET) properties.

We present a truly MPO approach to *de-novo* design, using Probabilistic Scoring [1] and quantitative structure-activity relationship (QSAR) models to generate high quality compound ideas. This is exemplified with optimisation of selective dipeptidyl peptidase (DPP) inhibitors.

Anagliptin: A Dipeptidyl Peptidase-4 Inhibitor

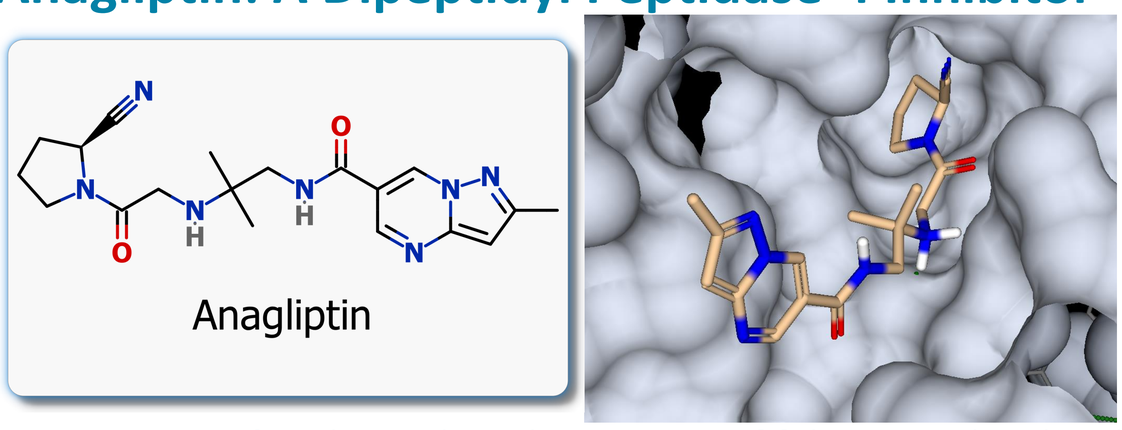

Figure 1. Structure of Anagliptin and crystal structure 3WQH [2] showing Anagliptin bound to human Dipeptidyl Peptidase-4 (DPP-4).

Anagliptin is a drug for the treatment of type 2 diabetes mellitus, belonging to the "gliptin" class of DPP-4 inhibitors. It is approved for use in Japan but has been withdrawn elsewhere due to animal toxicity [3] raising concerns about selectivity over DPP-2, DPP-8 and DPP-9, which have since been disproved.

We chose the non-selective Anagliptin analogue, CHEMBL1929395, which has measured IC_{50} data at three of the receptors of interest, as the starting template for our *de-novo* design strategy.

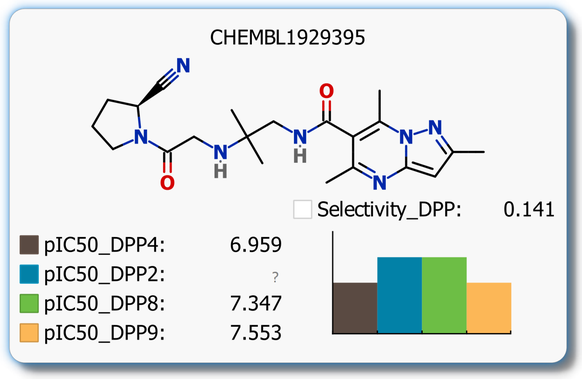

Figure 2. Structure of CHEMBL1929395 with measured pIC50 at DPP-4, DPP-8 and DPP-9 [4]. The low selectivity score (0.141 on a scale of 0 to 1) indicates low selectivity for DPP-4.

Potency prediction using QSAR models

To prioritise *de-novo* design ideas it is necessary to predict potencies of the virtual compounds. Models were built to predict pIC_{50} at DPP-2, DPP-4, DPP-8 and DPP-9 using training, test and validation sets with measured pIC_{50}s from ChEMBL [4]. QSAR model building techniques available in StarDrop's Auto-Modeller™ module [5] were employed.

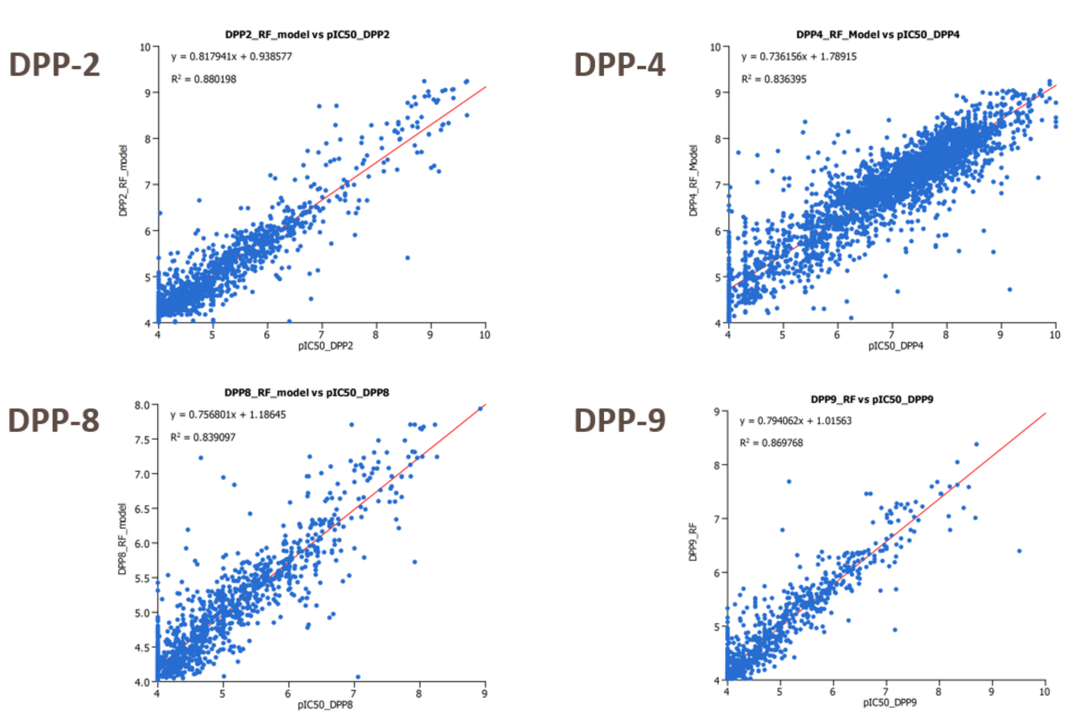

Figure 3. Validation of models to predict potency of compounds for DPP-2, DPP-4, DPP-8 and DPP-9.

Multi-Parameter Optimisation

To design potential DPP-4 inhibitors, selective over DPP-2, DPP-8 and DPP-9 and with a balanced ADMET profile we took an MPO approach; optimising for high potency at DPP-4 and low potency at DPP-2, DPP-8 and DPP-9 whilst simultaneously optimising key ADMET properties for oral bioavailability. Compounds are scored using the Probabilistic Scoring [1] approach in StarDrop™ [6], which uniquely accounts for the uncertainties in the data.

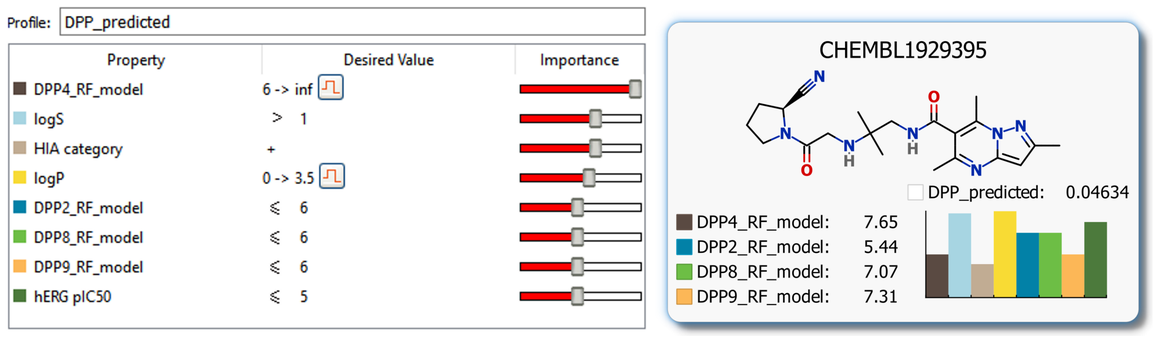

Figure 4. MPO scoring profile for scoring DPP-4 compounds. Scores range between 0-1 (low-high) with CHEMBL1929395 having a low score = 0.0463. The coloured histogram indicates the impact of each property on the score.

De-Novo Design: Med. Chem. Idea Generation

By combining *in-silico* design with predictive models and MPO many virtual compounds can be evaluated and prioritised against project objectives. We applied a 'transformation rules' approach, using the Nova™ module of StarDrop [6]. A library of >200 functional group and framework transformations, derived from medicinal chemistry experience [7], were applied to the seed compound, CHEMBL1929395, to generate synthetically accessible virtual compounds.

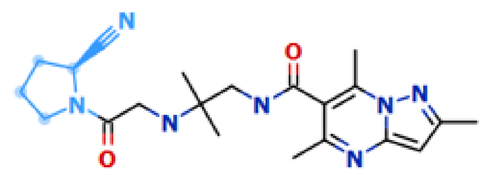

Figure 5. The cyanopyrrollidine, highlighted in blue for CHEMBL192935, binds deep in the binding pocket (Figure 1) and was conserved during the *de-novo* design process. Transformations were applied over 4 generations, with the 15 compounds best satisfying an MPO profile (Figure 4) taken forward to seed the next generation.

Results

The compounds were scored using Probabilistic scoring, to prioritise those with the highest chance of success, prior to further evaluation, e.g. by docking, or synthesis and testing. The designed compound library includes high-scoring members similar to known DPP-4 ligands and also adds new structural diversity for exploration (Figure 6).

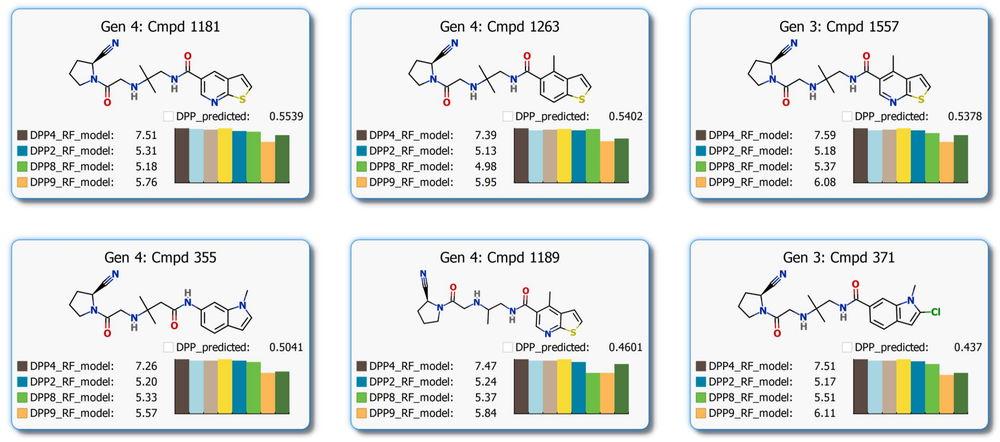

Figure 6. High-scoring *de-novo* compound ideas exhibiting 2 log unit predicted selectivity for DPP4 over DPP2, DPP8 and DPP9, with an overall balance of ADMET properties. The highest scoring idea, 'Gen 4: Cmpd 1181', scores 0.55 *vs* 0.05 for the seed, CHEMBL1929395.

Using an MPO approach, we have shown that it is possible to optimise both predicted potency and selectivity over multiple receptors in a single step, whilst simultaneously optimising a balanced ADMET profile.

References

[1] M.D. Segall (2012) *Curr. Pharm. Des.* 18(9) pp. 1292-1310

[2] PDB ID: 3WQH, DOI 10.2210/pdb3wqh/pdb; Y.S. Watanabe, Y. Yasuda, S. Okada, T. Motoyama and M. Oka (2015) *J. Enzyme Inhib. Med. Chem.* pp. 1-8

[3] R.N. Kushwaha, W. Haq, and S.B. Katti (2014) Curr. Med. Chem. 21(35) pp. 4013-4045

[4] ChEMBL ID: CHEMBL1929395; ChEMBL Database: A.P. Bento, A. Gaulton, A. Hersey, L.J. Bellis, J. Chambers, M. Davies, F.A. Krüger, Y. Light, L. Mak, S. McGlinchey, M. Nowotka, G. Papadatos, R. Santos and J.P. Overington (2014) *Nucleic Acids Res*., 42 pp. 1083-1090.

[5] O. Obrezanova, J.M.R. Gola, E.J. Champness and M.D. Segall (2008) *J. Comput. Aided Mol. Design* 22(6-7) pp. 431-440

[6] StarDrop v.6.4, Optibrium Ltd; http://www.optibrium.com/stardrop

[7] K.D. Stewart, M. Shiroda and C.A. James (2006) *Bioorg. Med. Chem*. 14(20) pp. 7011-7022; M. Segall, E. Champness, C. Leeding, R. Lilien, R. Mettu and B. Stevens *J. Chem. Inf. Model*. (2011) 51 pp. 2967-2976

© Optibrium Ltd. 2017. Optibrium™, StarDrop™, Auto-Modeller™ and Nova™ are trademarks of Optibrium Ltd.