Supplementary Material

Active learning FEP using 3D-QSAR for prioritizing bioisosteres in medicinal chemistry

Venkata K. Ramaswamy, Matthew Habgood, Mark D. Mackey*

Cresset, New Cambridge House, Bassingbourn Road, Litlington SG8 0O5, Cambridgeshire, United Kingdom

Figure SM 1. *2D representation of molecules with reported activity data[1,2] (pIC50 spanning 5.21 to 9) within the pool of 500 Spark results.*

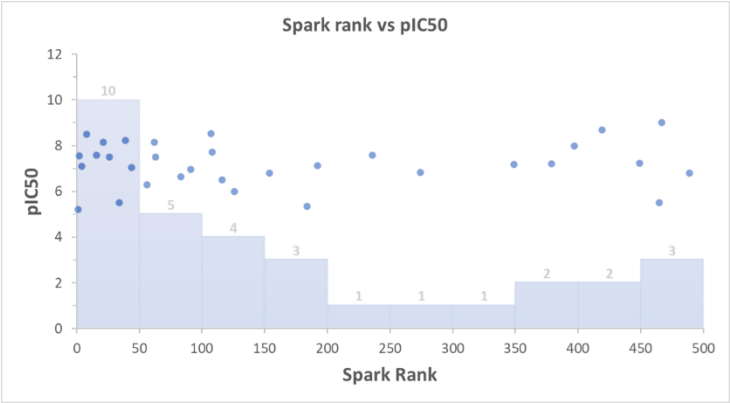

Figure SM 2. *Scatter plot of Spark rank vs. activity (pIC50) of known hALR2 inhibitors among top 500 Spark suggestions. The overlaid frequency histogram represents the binned Spark rank (bin size 50) with the count of known molecules in each bin annotated at the top of each column.*

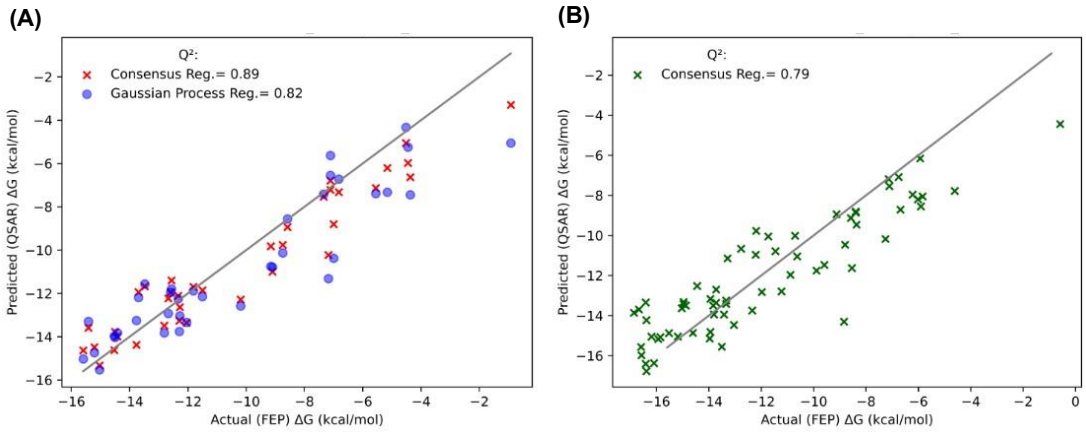

Figure SM 3. *Plot of binding affinities (ΔG in kcal/mol) predicted by* ***(A)*** *Consensus and Gaussian process Regression models of QSAR round-1 (explore mode) and* ***(B)*** *Consensus model of QSAR round-2 (exploit mode) against FEP activities for cross-validation of the training set. The diagonal line is an ideal fit with zero intercept and unit slope.*

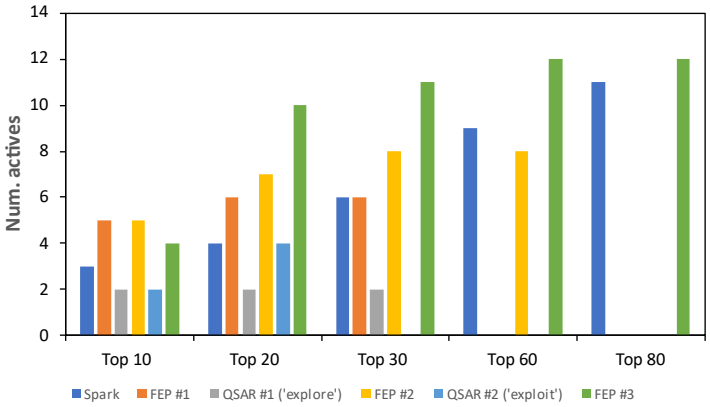

Figure SM 4. *Histogram plot of the number of known actives among the top-ranked molecules retrieved from Spark and different stages of the active learning workflow.*

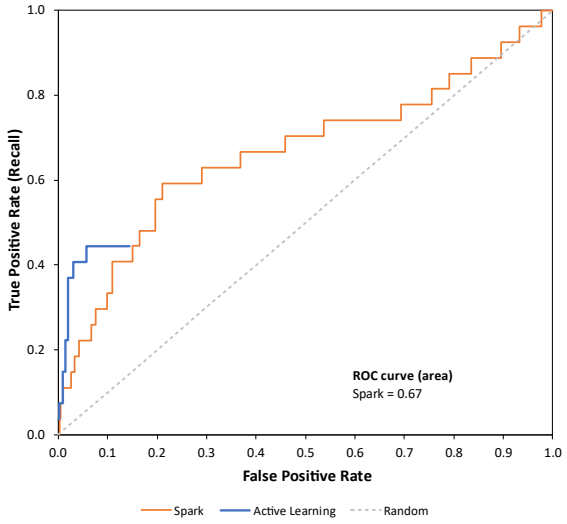

Figure SM 5. *ROC AUC curves showing an early enrichment of actives in our active learning workflow compared to Spark. Molecules with reported pIC50 > 6 are considered actives.*

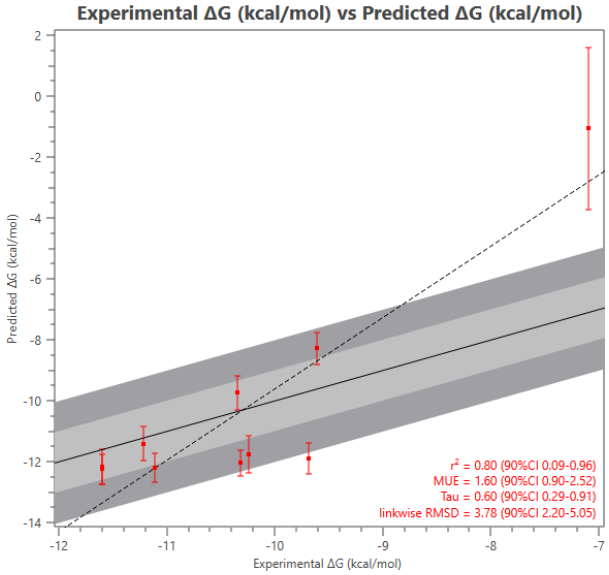

Figure SM 6. *Comparison of experimental and predicted ΔG (kcal/mol) on a subset of 10 ALR2 inhibitors with known activities from Mylari et al.[1,2] demonstrating the suitability and reliability of our Flare FEP protocol for this case study. The error values are bootstrapped estimates including the MBAR error, hysteresis and cycle closure. The corresponding coefficient of determination (R^2), Mean Unsigned Error (MUE), Kendall's tau and linkwise RMSD (RMSD errors on the links) statistics are shown in the bottom-right corner of the Activity Plot.*

References

(1) Mylari, B. L.; Larson, E. R.; Beyer, T. A.; Zembrowski, W. J.; Aldinger, C. E.; Dee, M. F.; Siegel, T. W.; Singleton, D. H. Novel, Potent Aldose Reductase Inhibitors: 3,4-Dihydro-4-Oxo-3-[[5-(Trifluoromethyl)-2-Benzothiazolyl]Methyl]-1-Phthalazineacetic Acid (Zopolrestat) and Congeners. *J. Med. Chem.* **1991**, *34* (1), 108–122. https://doi.org/10.1021/jm00105a018.

(2) Mylari, B. L.; Beyer, T. A.; Scott, P. J.; Aldinger, C. E.; Dee, M. F.; Siegel, T. W.; Zembrowski, W. J. Potent, Orally Active Aldose Reductase Inhibitors Related to Zopolrestat: Surrogates for Benzothiazole Side Chain. *J. Med. Chem.* **1992**, *35* (3), 457–465. https://doi.org/10.1021/jm00081a006.