

# Pathfinder-driven chemical space exploration and multiparameter optimization in tandem with Glide/IFD and QSAR-based active learning approach to prioritize design ideas for FEP+ calculations of SARS-COV-2 PL<sup>pro</sup> inhibitors

N.J. Gumede\*

Department of Chemistry, Mangosuthu University of Technology, PO Box 12363, Jacobs, 4026, South Africa.

## Corresponding Author

\*E-mail: ngumede@mut.ac.za

## Supplementary Materials

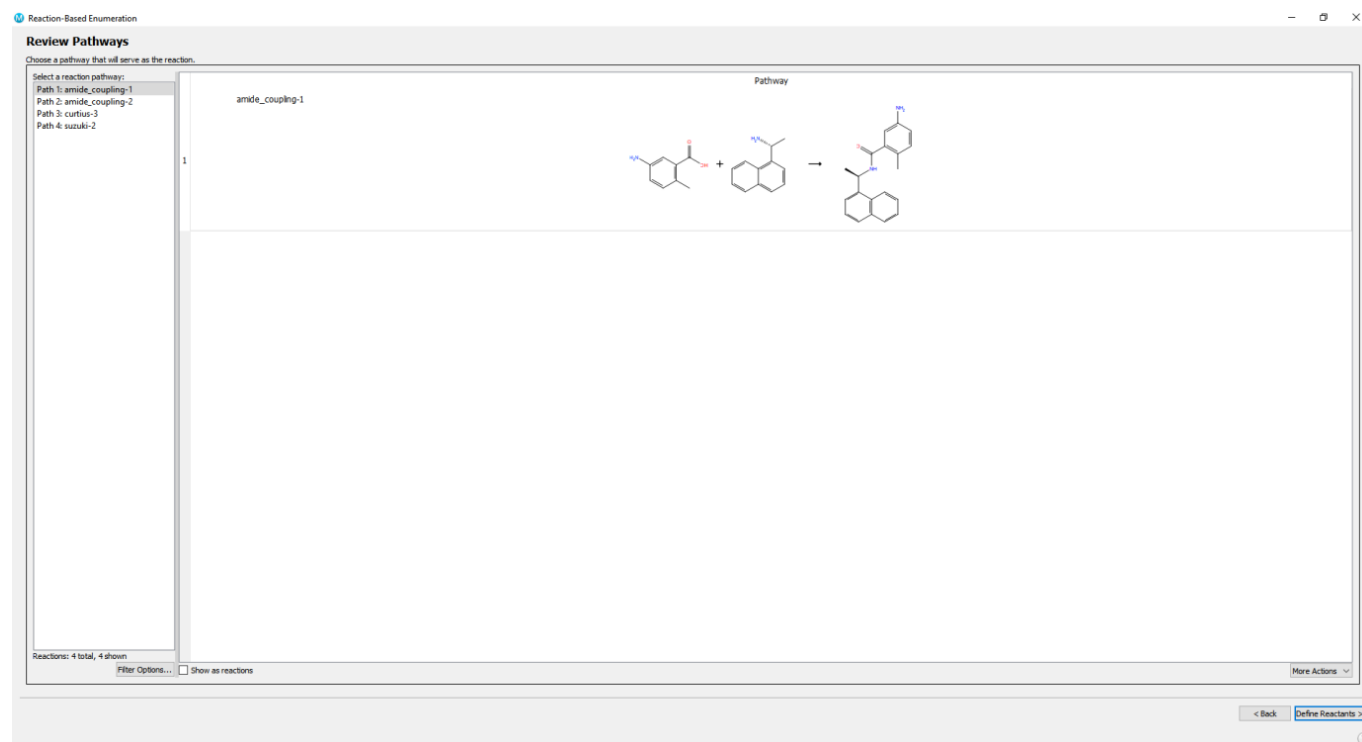
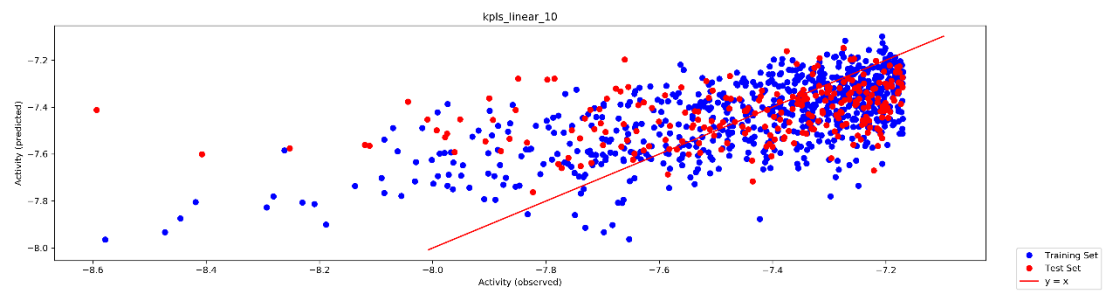
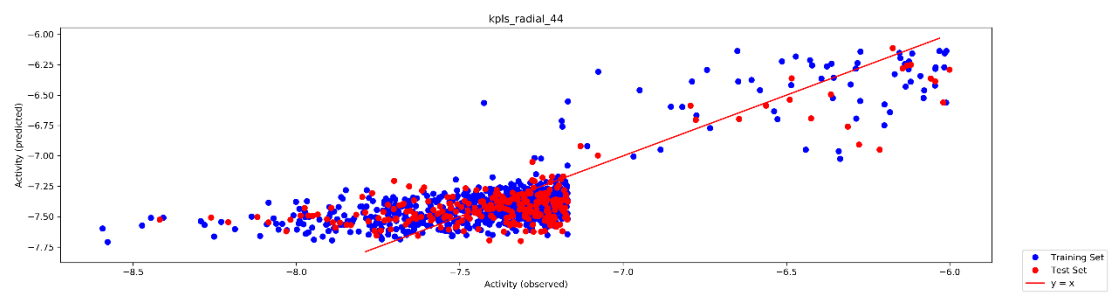


Figure S1. Reaction based enumeration generated pathway for GRL-0617.

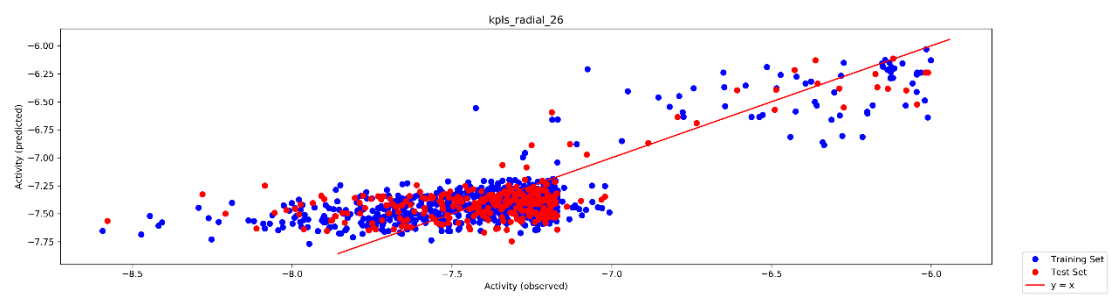
(a)



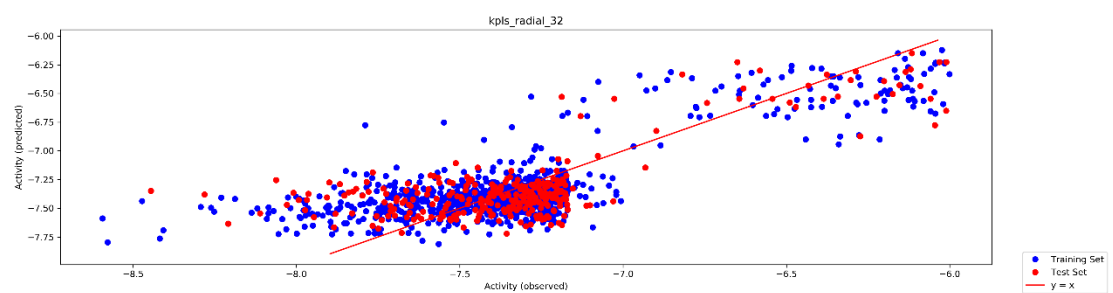
(b)



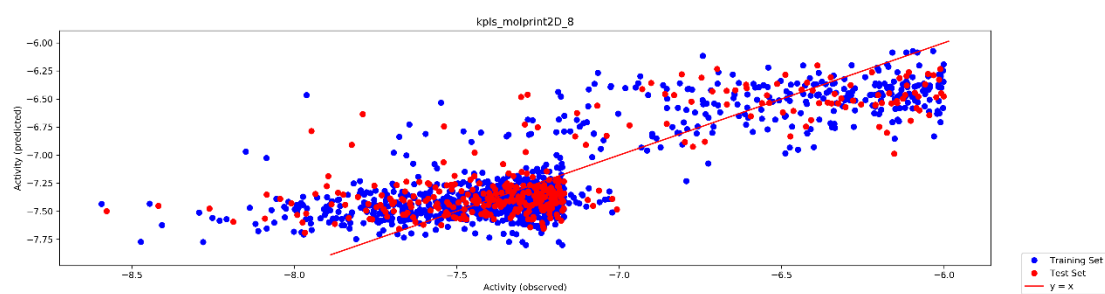
(c)



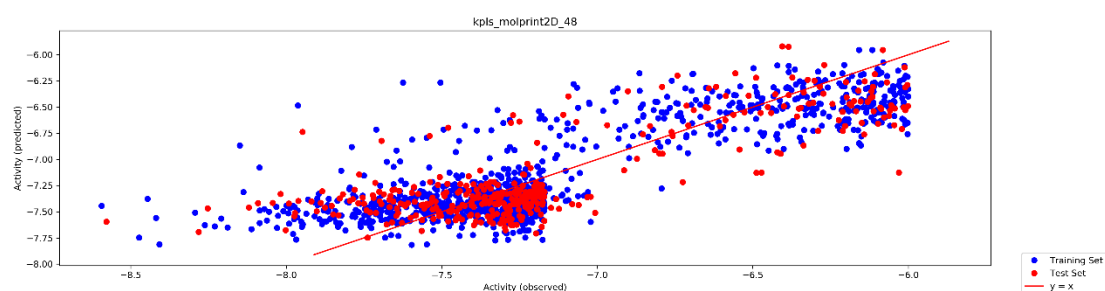
(d)



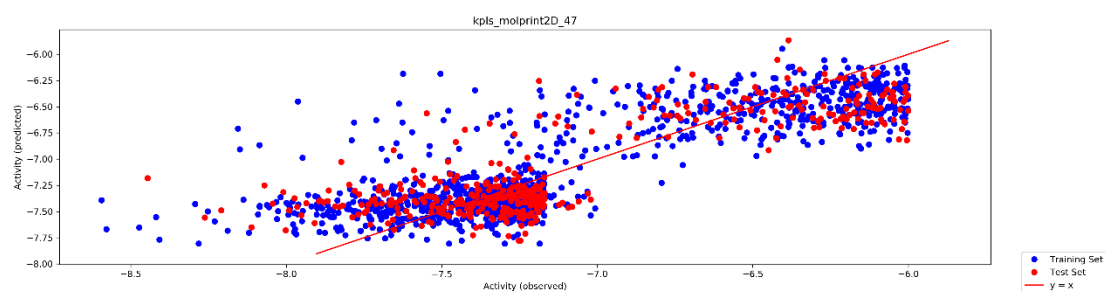
(e)



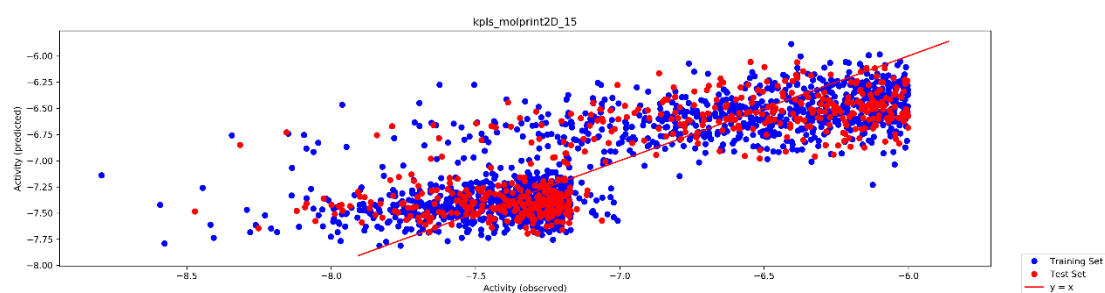
(f)



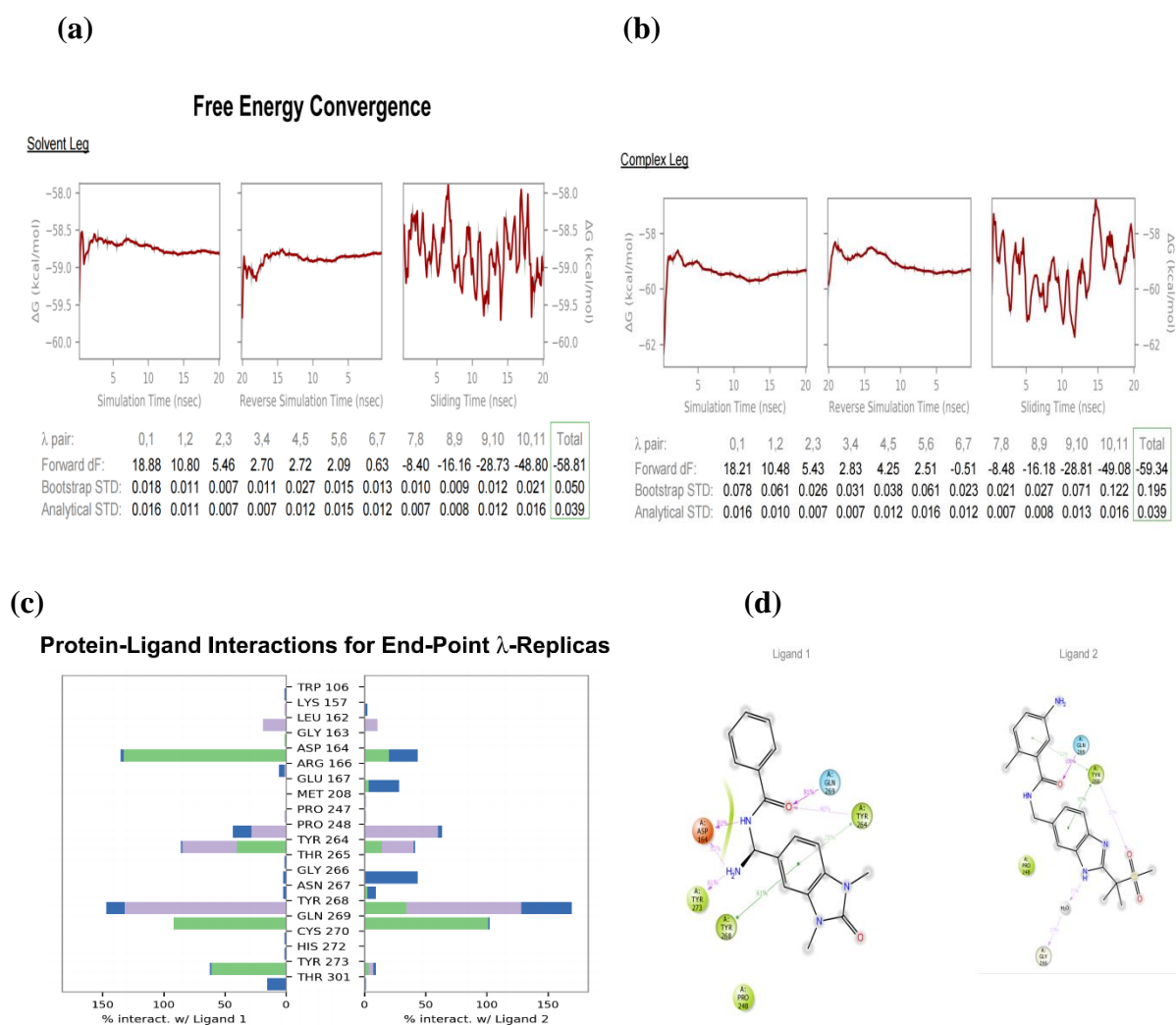
(g)



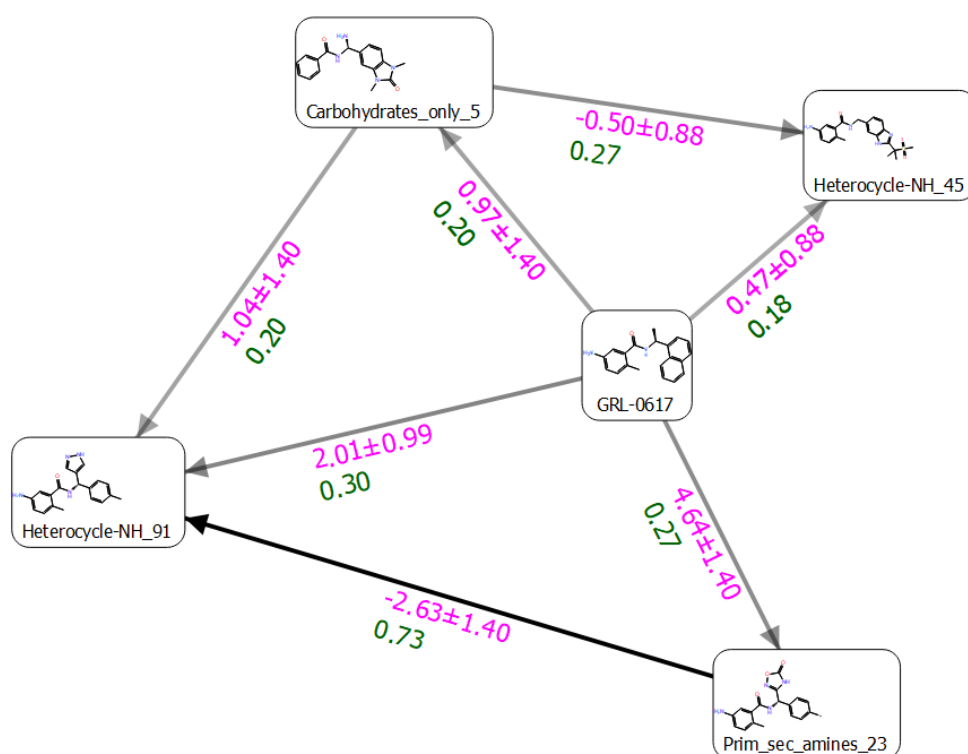
(h)



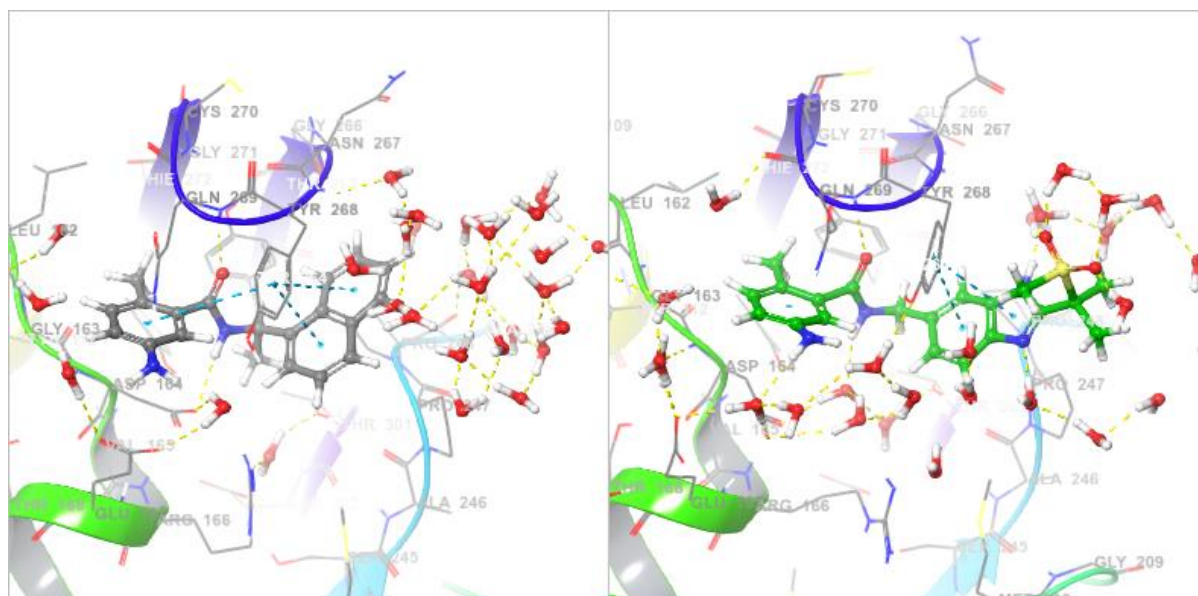
**Figure S2.** Plot of Active learning Glide SP and Auto QSAR models to predict the binding affinities of enumerated compounds **(a)** Auto QSAR\_model\_1 with an  $R^2$  of 0.4310 and a  $Q^2$  of 0.1571 **(b)** Auto QSAR\_model\_2 with an  $R^2$  of 0.6262 and a  $Q^2$  of 0.5929 **(c)** Auto QSAR\_model\_3 with an  $R^2$  of 0.6090 and a  $Q^2$  of 0.5630 **(d)** Auto QSAR\_model\_4 with an  $R^2$  of 0.6227 and a  $Q^2$  of 0.5793 **(e)** Auto QSAR\_model\_5 with an  $R^2$  of 0.6980 and a  $Q^2$  of 0.6450 **(f)** Auto QSAR\_model\_6 with an  $R^2$  of 0.7084 and a  $Q^2$  of 0.6676 **(g)** Auto QSAR\_model\_7 with an  $R^2$  of 0.7086 and a  $Q^2$  of 0.6910 **(h)** Auto QSAR\_model\_9 with an  $R^2$  of 0.6718 and a  $Q^2$  of 0.6476.



**Figure S3.** FEP+ simulation results for relative binding affinity prediction for the perturbation cycle of compound 5 (ligand 1) and compound 45 (ligand 2). **(a)** Free energy convergence of the perturbation cycle in solvent leg. **(b)** Free energy convergence of the perturbation cycle in complex leg. **(c)** Histogram depicting protein-ligand interactions for endpoint  $\lambda$ -replicas. **(d)** Ligand interaction diagram detailing the type of interactions observed between the two ligands and the receptor.

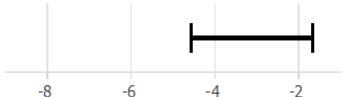
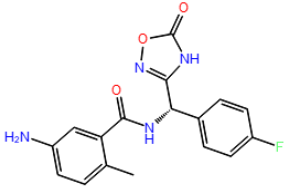
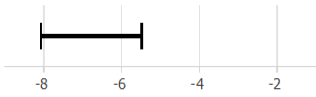
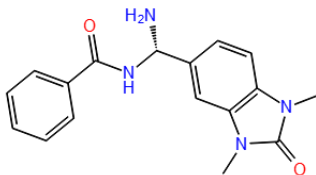
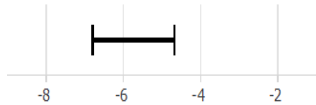
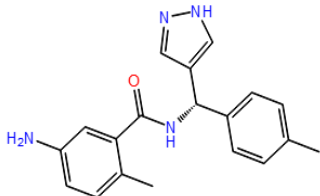


**Figure S4.** A circle closure perturbation map of design ideas and GRL-0617 detailing the difference in binding free energy,  $\Delta\Delta G$  (pink) and its associated error estimate between two ligands in an edge/node, as well as their associated ligand similarities scores (green).



**Figure S5.** FEP+ node representation for GRL-0617 (grey carbons) on the left and compound 45 (green carbons) on the right frame in the active site cavity of PL<sup>pro</sup> with favourable and unfavourable waters shown.

**Table S1.** Relative binding affinity prediction and its associated predicted error in kcal/mol of the design ideas and GRL-0617 as a reference compound against SARS-COV-2 PLpro using FEP+.

No.	Ligand	Estimated $\Delta G$	Pred. $\Delta G$ kcal/mol	Pred. Error kcal/mol	Exp. $\Delta G$ kcal/mol	Exp. Error	Structure
1	Primary & secondary amines_23		-3,11	1,45	N/A	N/A	
2	Carboxylates_only_5		-6,78	1,30	N/A	N/A	
3	N-Heterocycles_91		-5,74	1,06	N/A	N/A	

4

GRL-0617

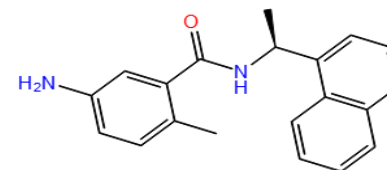


-7,75

0,40

-7,75

0



5

N-Heterocycles\_45



-7,28

0,96

N/A

N/A

