Supplemental Material

Physicochemical properties and structural parameters contributing to the antibacterial activity and efflux susceptibility of small molecule inhibitors of *Escherichia coli*

Sara S. El Zahed, Shawn French, Maya A. Farha, Garima Kumar, Eric D. Brown

Table of Contents

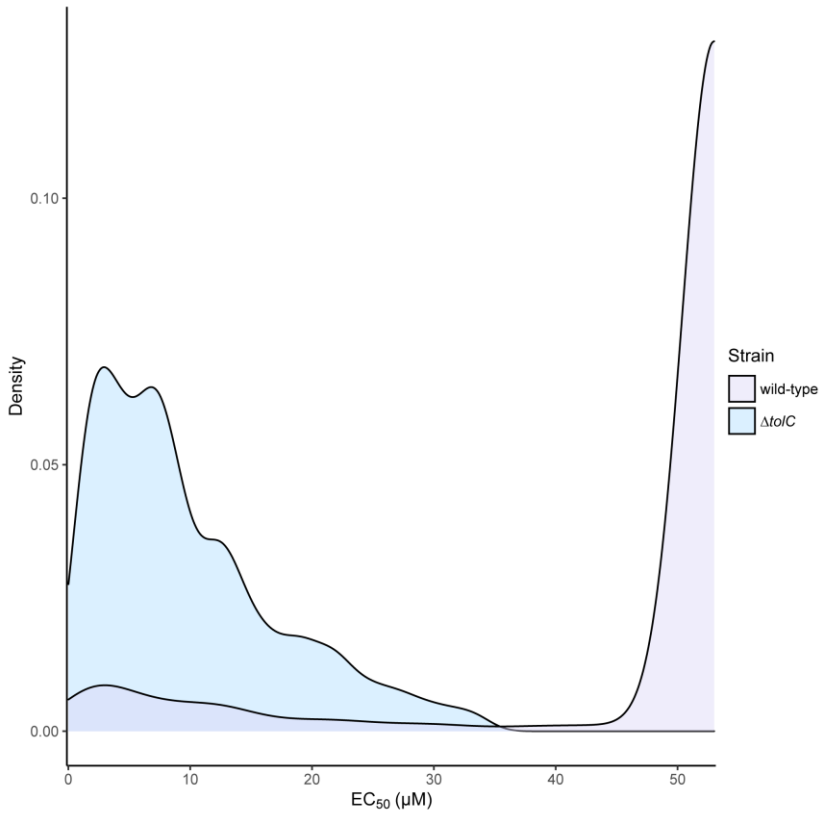


Fig. S1. The distribution of the measured potency values of the 4,500 actives in wild-type *E. coli* BW25113 and the mutant strain *E. coli* BW25113 Δ*tolC*. Density plot comparing the distribution of the measured potency values of the 4,500 actives from the primary screen in the efflux-proficient strain wild-type *E. coli* BW25113 (purple) and the efflux-compromised strain *E. coli* BW25113 Δ*tolC* (blue).

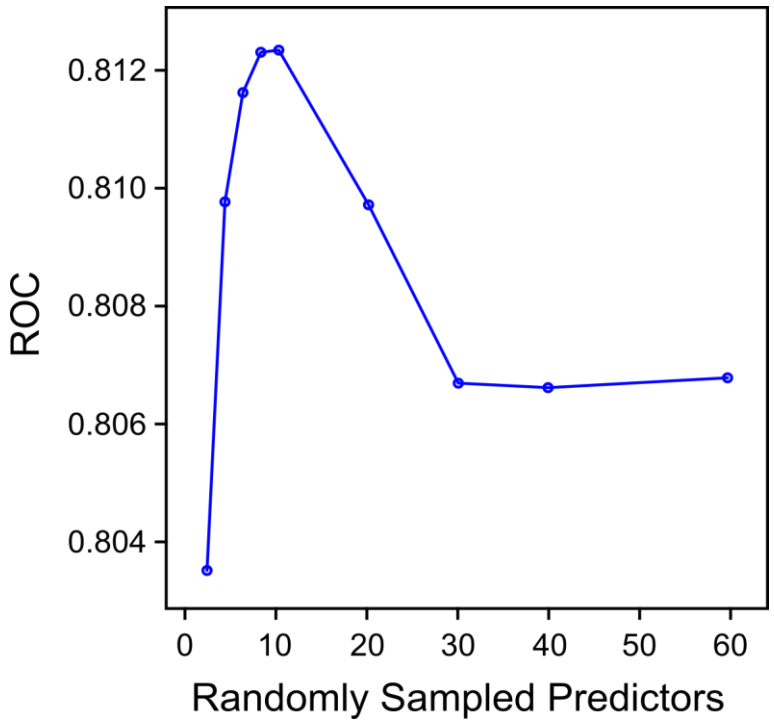


Fig. S2. The receiver operating characteristic curve for the random forest model classifying efflux-dependent actives and inactive molecules from the primary screen. The training set of molecules used to build this model was composed of 3,780 efflux-dependent actives and a random set of 3,780 inactive molecules from the primary screen. Tuning the number of molecular descriptors possible at each node in the decision trees, the number of randomly selected descriptors (features) were compared to a receiver operating characteristic curve (ROC). Here, the number of molecular descriptors at each tree node that resulted in the most accurate model was observed to be 10.

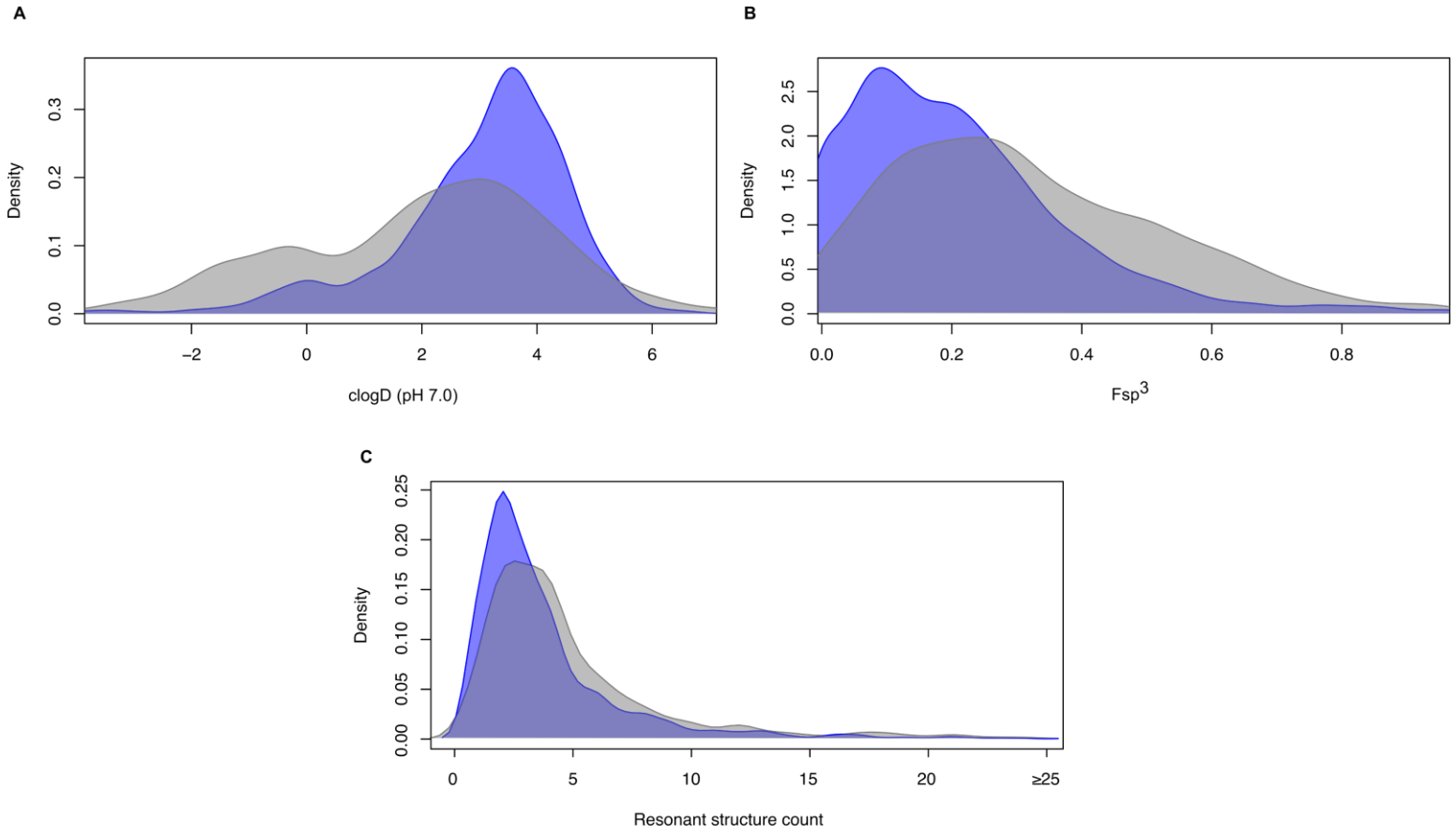


Fig. S3. Physicochemical and molecular space occupied by efflux-dependent actives and inactive molecules from the primary screen. Density plots comparing some molecular descriptors of the 3,780 efflux-dependent actives (blue) and the random set of 3,780 inactive molecules (grey) from the primary screen. This set of molecules and their associated descriptors were used to build the random forest model in Fig. 3. Shown are the top three descriptors contributing to this model's accuracy in classification: **(A)** clogD (pH 7.0), **(B)** Fsp^3, and **(C)** resonant structure count.

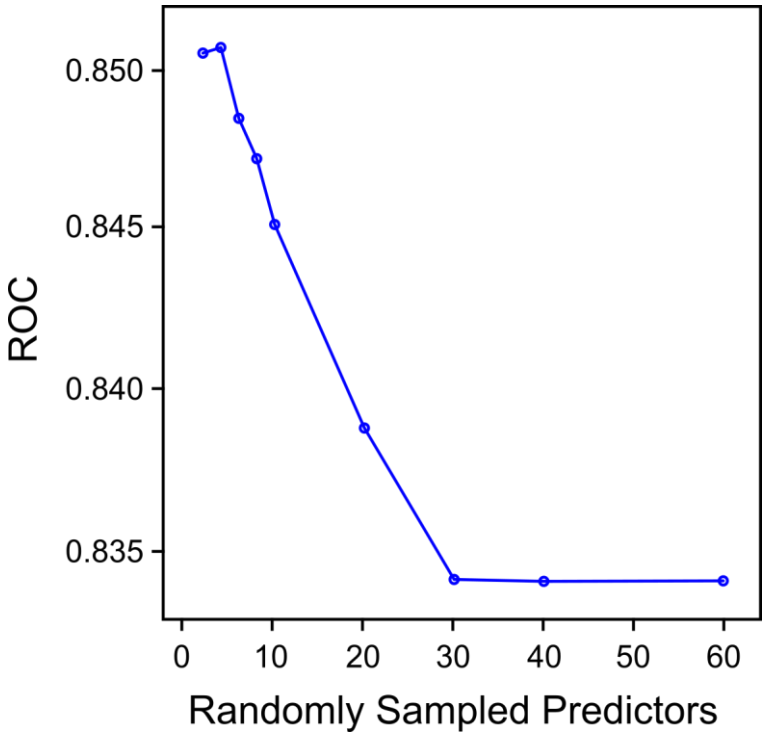


Fig. S4. The receiver operating characteristic curve for SERF. The training set of molecules used to build this model was composed of a random set of ~290 pumped molecules and ~290 non-pumped molecules. Tuning the number of molecular descriptors possible at each node in the decision trees, the number of randomly selected descriptors (features) were compared to a receiver operating characteristic curve (ROC). Here, the number of molecular descriptors at each tree node that resulted in the most accurate model was observed to be 6.

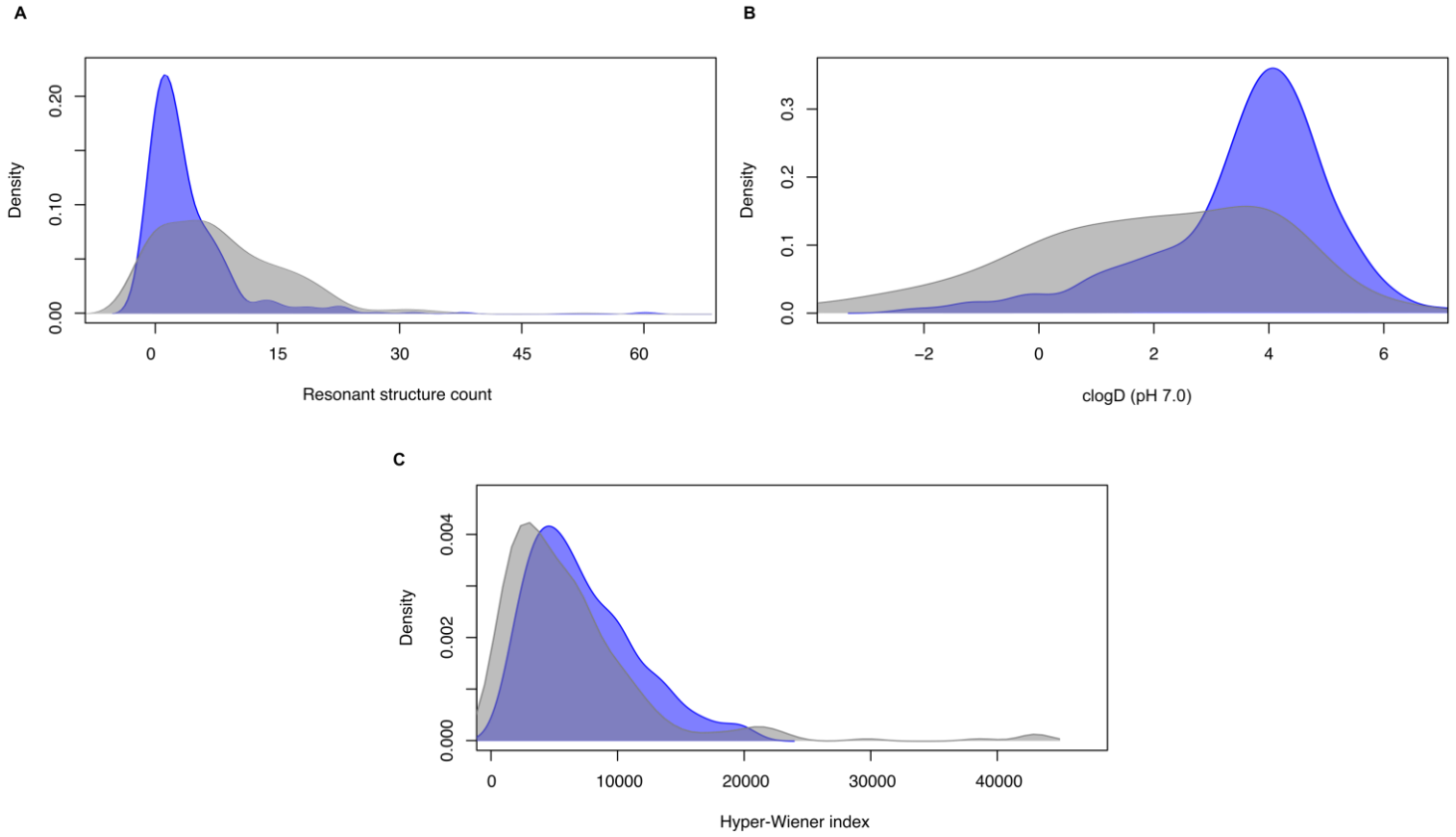
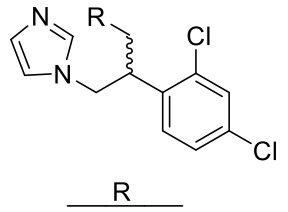
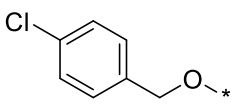


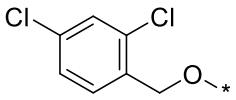
Fig. S5. Physicochemical and molecular space occupied by pumped molecules and non-pumped molecules. Density plots comparing some molecular descriptors of the random set of ~290 pumped molecules (blue) and ~290 non-pumped molecules (grey). This set of molecules and their associated descriptors were used to build SERF. Shown are the top three descriptors contributing to SERF's accuracy in classification: **(A)** resonant structure count, **(B)** clogD (pH 7.0), and **(C)** hyper-Wiener index.

Table S5. Structure, activity, and molecular descriptors of an antifungal compound series

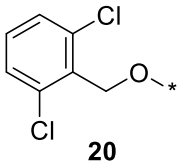




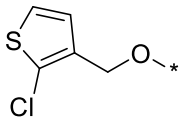
18



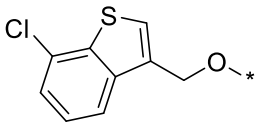
19



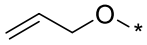
20



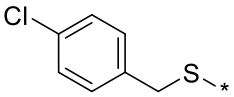
21



22



23



24

Cmpd[a]	Cmpd Name	W.T.[b] EC_{50} (μM)	Δ*tolC* EC_{50} (μM)	F.C.[c]	M.W. (g mol^{-1})	clogP	PSA[d] ($Å^2$)	Fsp^3 [e]	RSC [f]	HWI [g]
18	Econazole	>50	4.34	>23.1	381.68	4.24	27.1	0.17	2	5,290
19	Miconazole	>50	4.80	>20.8	416.12	4.85	27.1	0.17	2	5,856
20	Isoconazole	>50	6.44	>15.5	416.12	4.85	27.1	0.17	2	5,561
21	Tioconazole	>50	6.50	>15.4	387.7	4.23	55.3	0.19	2	4,297
22	Sertaconazole	>50	7.45	>13.4	437.76	5.08	55.3	0.15	2	7,432
23	Enilconazole	>50	24.6	>4.07	297.18	2.90	27.1	0.21	2	2,273
24	Sulconazole	11.98	5.30	2.3	397.74	5.52	43.1	0.17	2	5,290

[a]Cmpd, compound

[b]W.T., wild-type *E. coli*

[c]F.C., fold-change

[d]PSA, polar surface area

[e]Fsp^3, ratio of sp^3 hybridized carbon atoms/total carbon atoms

[f]RSC, resonant structure count

[g]HWI, hyper-Wiener index