Supporting Information

Machine Learning Model Analysis and Data Visualization with Small Molecules Tested in a Mouse Model of *Mycobacterium tuberculosis* Infection (2014-2015)

Sean Ekins,[1,2*#] Alexander L. Perryman,[3#] Alex M. Clark,[4] Robert C. Reynolds [5] and Joel S. Freundlich [3,6]

[1] Collaborative Drug Discovery, 1633 Bayshore Highway, Suite 342, Burlingame, CA 94010, USA.

[2] Collaborations in Chemistry, 5616 Hilltop Needmore Road, Fuquay-Varina, NC 27526, USA.

[3] Department of Pharmacology, Physiology and Neuroscience, Rutgers University-New Jersey Medical School, Newark, New Jersey 07103, United States

[4] Molecular Materials Informatics, Inc., 1900 St. Jacques #302, Montreal H3J 2S1, Quebec, Canada

[5] University of Alabama at Birmingham, College of Arts and Sciences, Department of Chemistry, 1530 3rd Avenue South, Birmingham, AL 35294-1240, USA.

[6] Division of Infectious Diseases, Department of Medicine, and the Ruy V. Lourenço Center for the Study of Emerging and Re-emerging Pathogens, Rutgers University–New Jersey Medical School, Newark, New Jersey 07103, United States

*To whom correspondence should be addressed. E-mail: ekinssean@yahoo.com

Phone 215-687-1320

[#] Contributed equally.

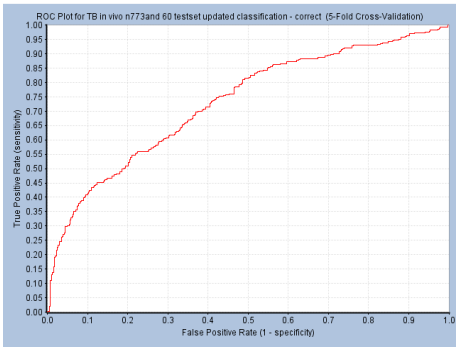
Supplemental data 1. 60 molecules curated in this study. Molecule identifier, closest distance to the N784 model and SMILES – This file is a separate xl file

Supplemental data 2. Original N = 773 Bayesian model used to predict 60 new molecules

ROC score is 0.767 (leave-one-out).
Best cutoff for this model is -1.159.
Test set validation: ROC score = 0.553915275994865
Model Rating: Quality 0.554: Fail
Confusion Matrix: True Positives = 29, False Negatives = 12, False Positives = 12, True Negatives = 7

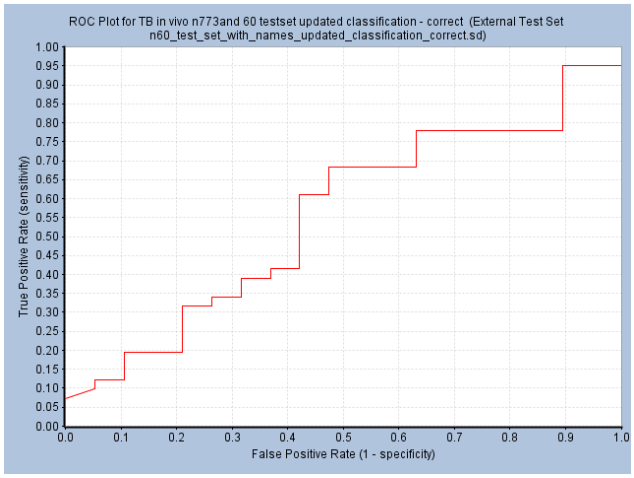
Training set

5-Fold Cross-Validation Result									
Model Name	**ROC Score**	**ROC Rating**	**True Positive**	**False Negative**	**False Positive**	**True Negative**	**Sensitivity**	**Specificity**	**Concordance**
TB in vivo n773and 60 testset updated classification - correct	0.731	Fair	247	115	41	370	0.682	0.900	0.798



Test set

Validation Result Using External Test Set n60_test_set_with_names_updated_classification_correct.sd									
Model Name	**ROC Score**	**ROC Rating**	**True Positive**	**False Negative**	**False Positive**	**True Negative**	**Sensitivity**	**Specificity**	**Concordance**
TB in vivo n773and 60 testset updated classification - correct	0.554	Fail	29	12	12	7	0.707	0.368	0.600

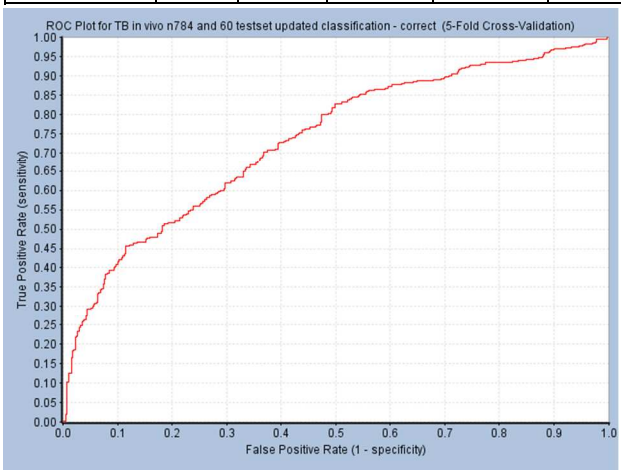

ROC Plot for TB in vivo n773and 60 testset updated classification - correct (External Test Set
n60_test_set_with_names_updated_classification_correct.sd)
True Positive Rate (sensitivity)
False Positive Rate (1 - specificity)
1.00
0.95
0.90
0.85
0.80
0.75
0.70
0.65
0.60
0.55
0.50
0.45
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0

Supplemental data 3. N=784 Bayesian model used to predict 60 new molecules

ROC score is 0.772 (leave-one-out).
Best cutoff for this model is -2.361.
Test set validation: ROC score = 0.557766367137356
Model Rating: Quality 0.558: Fail
Confusion Matrix: True Positives = 32, False Negatives = 9, False Positives = 17, True Negatives = 2

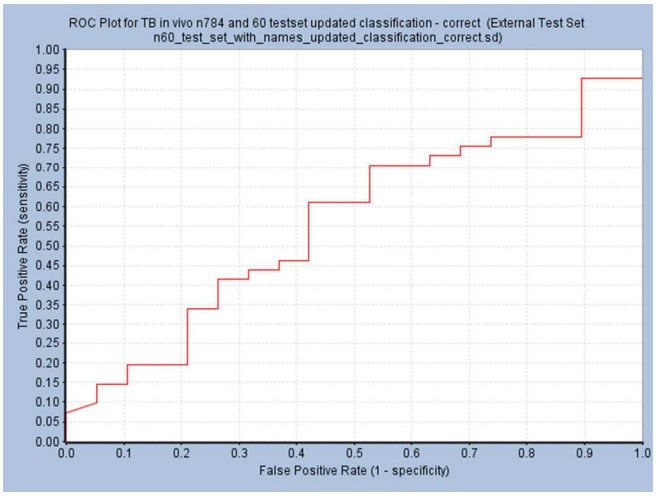
Training set

5-Fold Cross-Validation Result									
Model Name	**ROC Score**	**ROC Rating**	**True Positive**	**False Negative**	**False Positive**	**True Negative**	**Sensitivity**	**Specificity**	**Concordance**
TB in vivo n784 and 60 testset updated classification - correct	0.733	Fair	275	98	51	360	0.737	0.876	0.810



Test set

Validation Result Using External Test Set n60_test_set_with_names_updated_classification_correct.sd									
Model Name	**ROC Score**	**ROC Rating**	**True Positive**	**False Negative**	**False Positive**	**True Negative**	**Sensitivity**	**Specificity**	**Concordance**
TB in vivo n784 and 60 testset updated classification - correct	0.558	Fail	32	9	17	2	0.780	0.105	0.567


ROC Plot for TB in vivo n784 and 60 testset updated classification - correct (External Test Set
n60_test_set_with_names_updated_classification_correct.sd)
True Positive Rate (sensitivity)
False Positive Rate (1 - specificity)
1.00
0.95
0.90
0.85
0.80
0.75
0.70
0.65
0.60
0.55
0.50
0.45
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0

Supplemental data 4 – N=784 SVM model information output from R.

SVM model

```
FitSummary


Call:
svm(formula = form, data = xy, type = type, kernel = tolower("Radial"),
    gamma = gamma, cost = cost, probability = prob, fitted = TRUE,
    epsilon = epsilon, nu = nu, coef0 = coef0, degree = degree, scale = TRUE)



Parameters:
   SVM-Type:  C-classification
 SVM-Kernel:  radial
       cost:  2
      gamma:  0.007352941

Number of Support Vectors:  647

 ( 328 319 )



Number of Classes:  2

Levels:
 0 1
```

Cross-validation results (5-fold):

	Gamma	Cost	ROC Score	Best
1	0.007353	1	0.7688	
2	0.007353	2	**0.7745**	***

Contingency Table (best CV model):

	Predicted	
Actual	0	1
0	308	103
1	126	247

All-data model results (non-cross-validated):

Settings used:

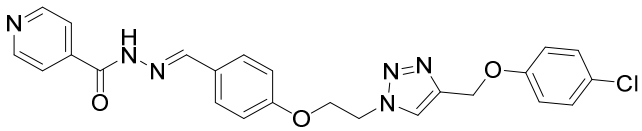
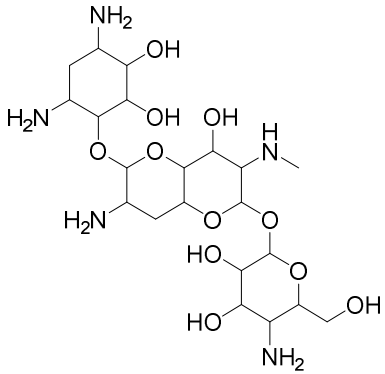
Gamma	Cost
0.007352941	2

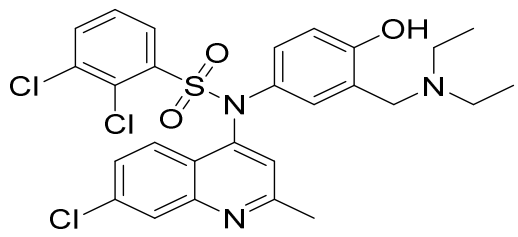
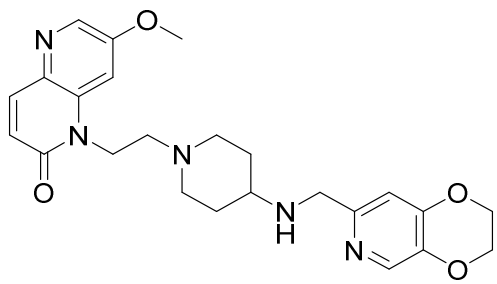
ROC Score**: 0.9766**

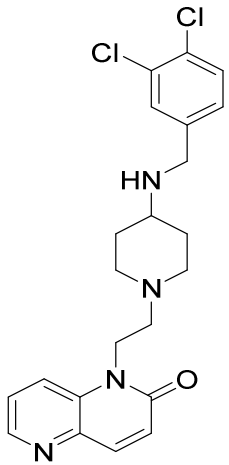
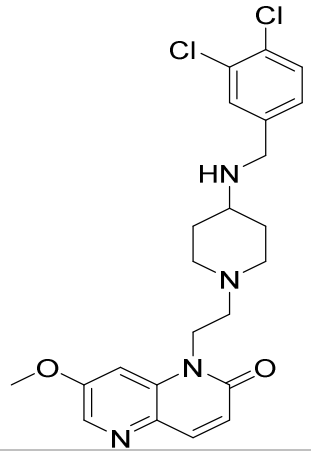
Contingency Table (all-data model):

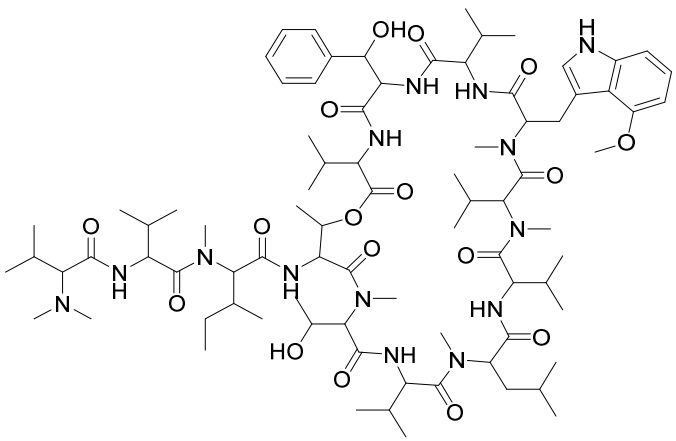
	Predicted	
Actual	0	1
0	389	22
1	44	329

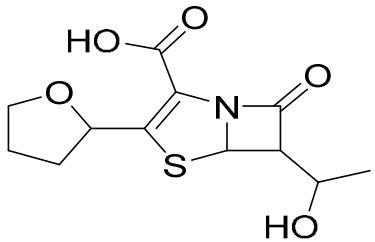
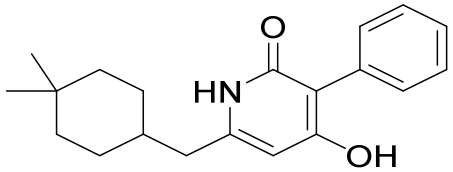
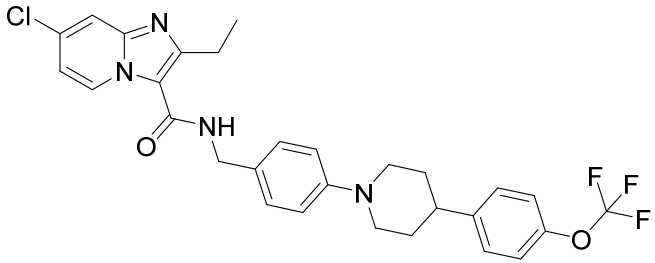
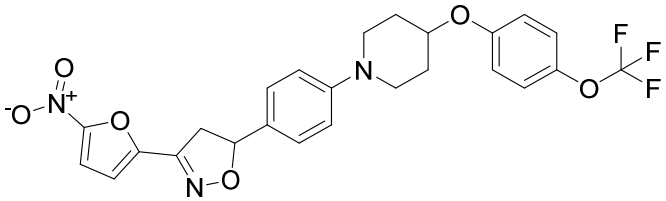
Table S1. 60 Molecules tested in the *in vivo* mouse infection model reported between 2014-2015

Molecule	Identifier and IUPAC Name	*In vitro* activity	*In vivo* Activity ~1 log CFU	Possible Target	References	Lung Log CFU reduction
	Compound 47 N'-[(1E)-(4-{2-[4-(4-chlorophenoxymethyl)-1H-1,2,3-triazol-1-yl]ethoxy}phenyl)methylidene]pyridine-4-carbohydrazide	MIC 0.195-0.39 μM	0	InhA	[59]	No effect on lung of BALB/c, dosed orally 25, 50 and 100 mg/kg once daily 5 days /week – 3, 6, 10 week treatment
	Apramycin 5-amino-2-({7-amino-6-[(4,6-diamino-2,3-dihydroxycyclohexyl)oxy]-4-hydroxy-3-(methylamino)-octahydropyrano[3,2-b]pyran-2-yl}oxy)-6-(hydroxymethyl)oxane-3,4-diol		1		[60]	2.4 log reduction in CFU in Female gamma interferon knockout mice, dosed 200 mg/kg subcutaneously for 9 days

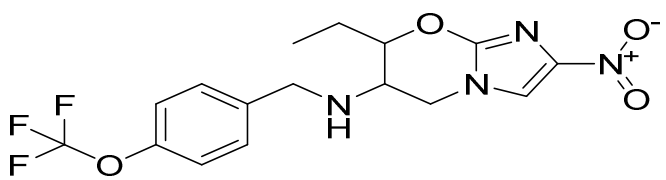
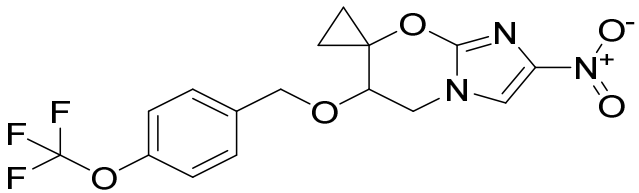
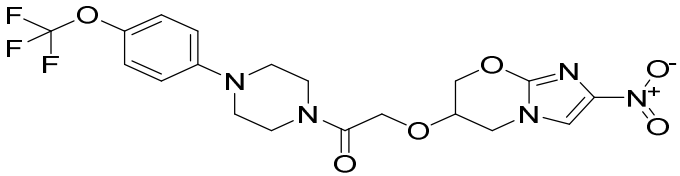
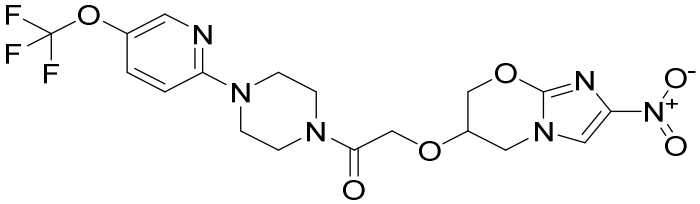
	Compound 9d 2,3-dichloro-N-(7-chloro-2-methylquinolin-4-yl)-N-{3-[(diethylamino) methyl]-4-hydroxyphenyl}benzene-1-sulfonamide	MIC 3.12 µg/mL	1	ATP synthase	[61]	2.12 log CFU reduction in BALB/c mic treated with 100 mg/kg orally, 6 times a week for 4 weeks
	Compound 1 1-{2-[4-({2H,3H-[1,4]dioxino[2,3-c]pyridin-7-ylmethyl}amino)piperidin-1-yl]ethyl}-7-methoxy-1,2-dihydro-1,5-naphthyridin-2-one	MIC 0.29 µM	0	Gyrase	[62]	No change in C57BL/6 mice dosed for 8 days at 100 mg/kg

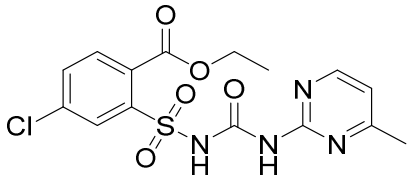
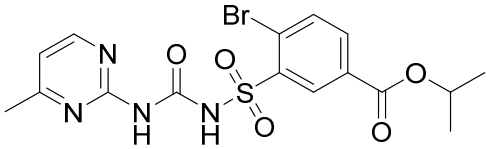
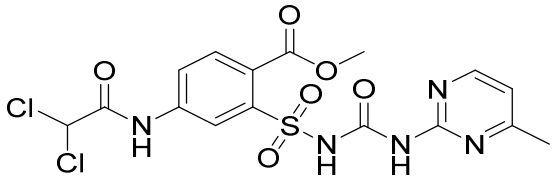
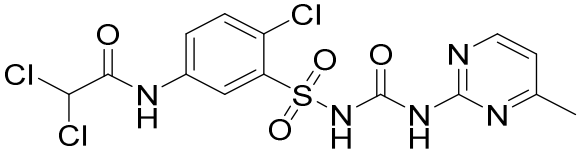
	Compound 2 1-[2-(4-{[(3,4-dichlorophenyl)methyl]amino}piperidin-1-yl)ethyl]-1,2-dihydro-1,5-naphthyridin-2-one	MIC 0.04 μM	1	Gyrase	62	>2 log CFU change in C57BL/6 mice dosed for 8 days at 50 mg/kg
	Compound 3 1-[2-(4-{[(3,4-dichlorophenyl)methyl]amino}piperidin-1-yl)ethyl]-7-methoxy-1,2-dihydro-1,5-naphthyridin-2-one	MIC <0.01 μM	1	Gyrase	62	>3 log CFU change in C57BL/6 mice dosed for 8 days at 50 mg/kg

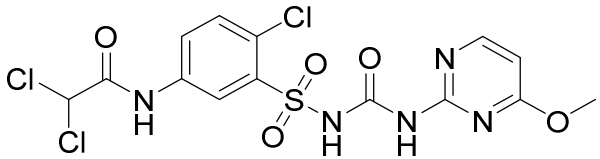
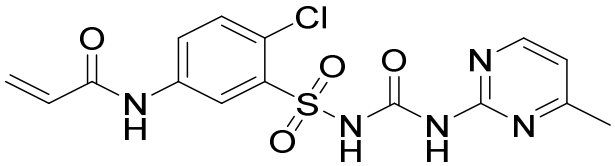
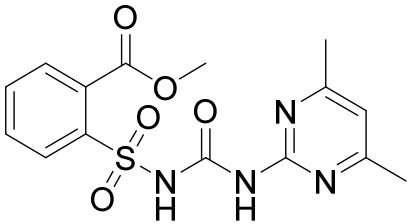
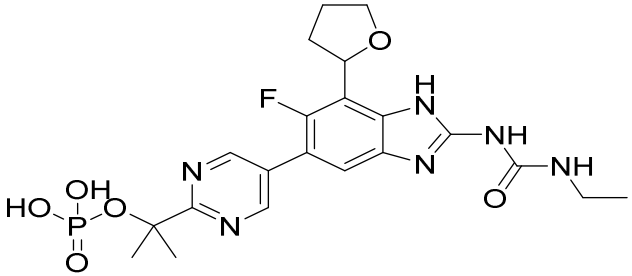
	Ecumicin 2-{2-[2-(dimethylamino)-3-methylbutanamido]-N,3-dimethylbutanamido}-N-{6-[hydroxy(phenyl)methyl]-27-(1-hydroxyethyl)-12-[(4-methoxy-1H-indol-3-yl)methyl]-13,16,22,28,31-pentamethyl-21-(2-methylpropyl)-2,5,8,11,14,17,20,23,26,29-decaoxo-3,9,15,18,24-pentakis(propan-2-yl)-1-oxa-4,7,10,13,16,19,22,25,28-nonaazacyclohentriacontan-30-yl}-3-methylpentanamide	MIC 0.16 μM	1	ClpC1 ATPase complex	63	1.3 log CFU reduction at 20 mg/kg dose subcutaneously in female BALB/c mice treated at intervals for 25 days.

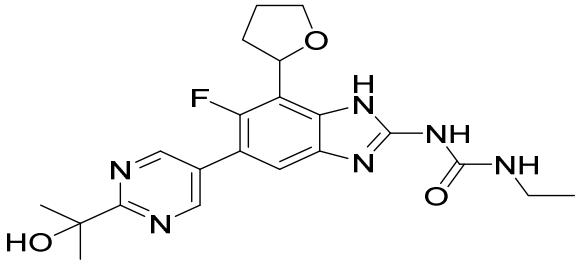
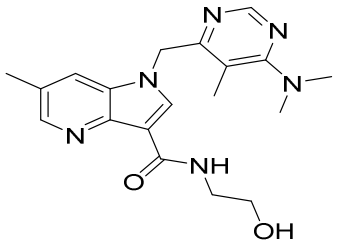
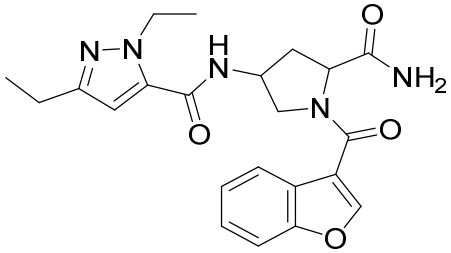
	Faropenem 6-(1-hydroxyethyl)-7-oxo-3-(oxolan-2-yl)-4-thia-1-azabicyclo[3.2.0] hept-2-ene-2-carboxylic acid	MIC 1.3 µg/mL	0	L,D-transpeptidases	64	No change with 500 mg/kg dosed intratracheal in C57BL/6J mice four times a day for 7 days
	NITD-916 6-[(4,4-dimethylcyclohexyl)methyl]-4-hydroxy-3-phenyl-1,2-dihydropyridin-2-one	MIC 0.05 µM, InhA IC_{50} 0.57 µM	1	InhA	65	1 log at 30 mg/kg, BALB/c mice, 4 week treatment.
	Compound 49 7-chloro-2-ethyl-N-[(4-{4-[4-(trifluoromethoxy)phenyl]piperidin-1-yl}phenyl)methyl]imidazo[1,2-a]pyridine-3-carboxamide	MIC_{80} 4 nM	1	QcrB	66	1.52 log CFU reduction in BALB/c mice dosed orally for 28 days
	Compound 9a 1-{4-[3-(5-nitrofuran-2-yl)-4,5-dihydro-1,2-oxazol-5-yl]phenyl}-4-[4-(trifluoromethoxy)phenoxy]piperidine	MIC 0.024 µg/mL	1		67	>1 log CFU reduction in GKO mouse 300 mg/kg dosed orally for 9 days

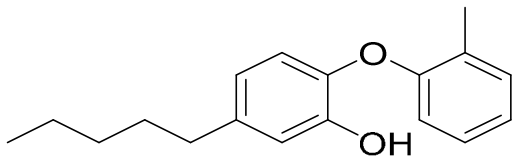
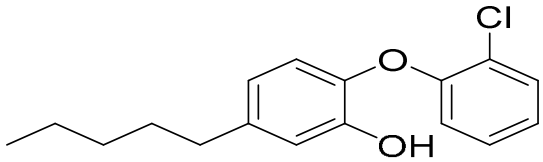
	Amino-824 2-nitro-N-{[4-(trifluoromethoxy)phenyl]methyl}-5H,6H,7H-imidazo[2,1-b][1,3]oxazin-6-amine	MIC 0.14 mg/L	1		68	1.08 log CFU reduction at 100 mg/kg in BALB/c mice dosed orally for 4 weeks daily
	AminoEthyl-824 2-nitro-N-{2-[4-(trifluoromethoxy)phenyl]ethyl}-5H,6H,7H-imidazo[2, 1-b][1,3]oxazin-6-amine	MIC 0.06 mg/L	1		68	1.69 log CFU reduction at 100 mg/kg in BALB/c mice dosed orally for 4 weeks daily
	NI-135 N-{[3-chloro-4-(trifluoromethoxy)phenyl]methyl}-2-nitro-5H,6H,7H-imidazo [2,1-b][1,3]oxazin-6-amine	MIC 0.03 mg/L	1		68	1.88 log CFU reduction at 100 mg/kg in BALB/c mice dosed orally for 4 weeks daily
	NI-136 N-{[3-fluoro-4-(trifluoromethoxy)phenyl]methyl}-2-nitro-5H,6H,7H-imidazo[2,1-b][1,3]oxazin-6-amine	MIC 0.03 mg/L	1		68	2.06 log CFU reduction at 100 mg/kg in BALB/c mice dosed orally for 4 weeks daily

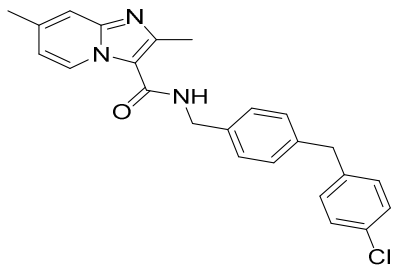
	NI-182 7-ethyl-2-nitro-N-{[4-(trifluoromethoxy)phenyl]methyl}-5H,6H,7H-imidazo [2,1-b][1,3]oxazin-6-amine	MIC 0.02 mg/L	1		68	1.26 log CFU reduction at 100 mg/kg in BALB/c mice dosed orally for 4 weeks daily
	NI-302 2'-nitro-6'-{[4-(trifluoromethoxy)phenyl]methoxy}-5',6'-dihydrospiro[cyclopropane-1,7'-imidazo[2,1-b][1,3]oxazine]	MIC 0.03 mg/L	1		68	2.17 log CFU reduction at 100 mg/kg in BALB/c mice dosed orally for 4 weeks daily
	NI-622 2-({2-nitro-5H,6H,7H-imidazo[2,1-b][1,3]oxazin-6-yl}oxy)-1-{4-[4-(tri fluoromethoxy)phenyl]piperazin-1-yl}ethan-1-one	MIC 0.18 mg/L	0		68	0.81 log CFU reduction at 100 mg/kg in BALB/c mice dosed orally for 4 weeks daily
	NI-644 2-({2-nitro-5H,6H,7H-imidazo[2,1-b][1,3]oxazin-6-yl}oxy)-1-{4-[5-(trifluoromethoxy)pyridin-2-yl]piperazin-1-yl}ethan-1-one	MIC 0.09 mg/L	0		68	0.57 log CFU reduction at 100 mg/kg in BALB/c mice dosed orally for 4 weeks daily

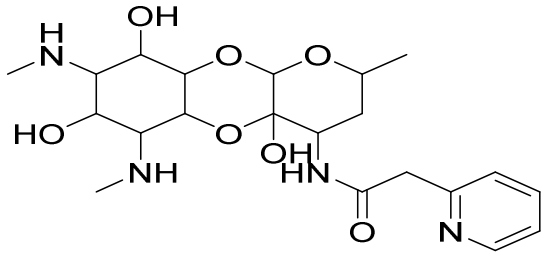
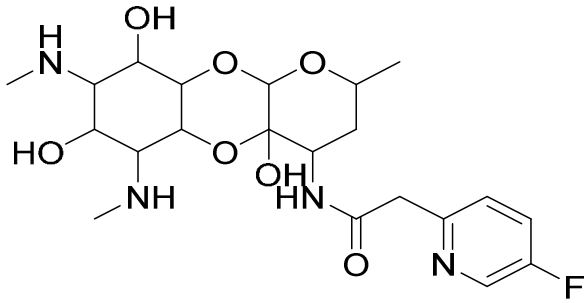
	Compound 3 ethyl 4-chloro-2-({[(4-methylpyrimidin-2-yl)carbamoyl]amino}sulfonyl)benzoate	MIC 10 mg/L	1	acetohydroxyacid synthase (AHAS)	[69]	1 log CFU reduction in Female BALB/c mice orally Dosed at 500 mg/kg
	Compound 15 propan-2-yl 4-bromo-3-({[(4-methylpyrimidin-2-yl)carbamoyl]amino}sulfonyl)benzoate	MIC 10 mg/L	1	acetohydroxyacid synthase (AHAS)	[69]	>1 log CFU reduction in Female BALB/c Dosed at 500 mg/kg
	Compound 20 methyl 4-(2,2-dichloroacetamido)-2-({[(4-methylpyrimidin-2-yl)carbamoyl]amino}sulfonyl)benzoate	MIC 20 mg/L	0	acetohydroxyacid synthase (AHAS)	[69]	<1 log CFU reduction in Female BALB/c Dosed at 500 mg/kg
	Compound 30 2,2-dichloro-N-[4-chloro-3-({[(4-methylpyrimidin-2-yl)carbamoyl]amino }sulfonyl)phenyl]acetamide	MIC 10 mg/L	1	acetohydroxyacid synthase (AHAS)	[69]	>2 log CFU reduction in Female BALB/c Dosed at 500 mg/kg

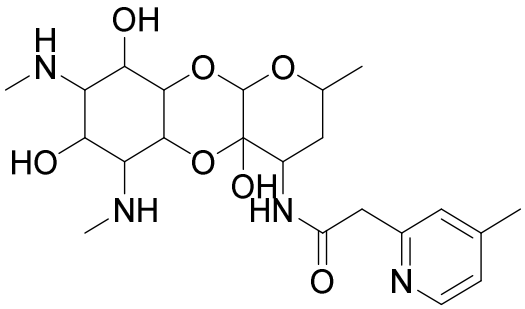
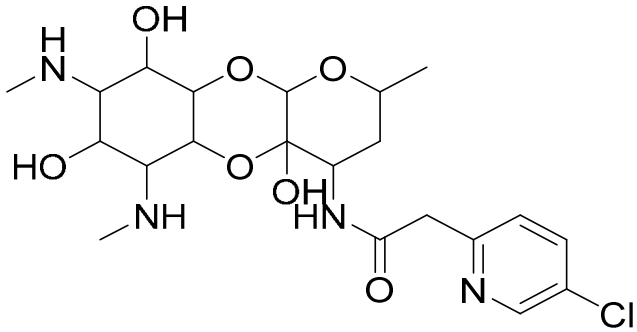
	Compound 31 2,2-dichloro-N-[4-chloro-3-({[(4-methoxypyrimidin-2-yl)carbamoyl]amino}sulfonyl)phenyl]acetamide	MIC 10 mg/L	1	acetohydroxyacid synthase (AHAS)	69	>2 log CFU reduction in Female BALB/c Dosed at 500 mg/kg
	Compound 32 N-[4-chloro-3-({[(4-methylpyrimidin-2-yl)carbamoyl]amino}sulfonyl)phenyl]prop-2-enamide	MIC 20 mg/L	0	acetohydroxyacid synthase (AHAS)	69	<1 log CFU reduction in Female BALB/c Dosed at 500 mg/kg
	Sulfometuron Methyl methyl 2-({[(4,6-dimethylpyrimidin-2-yl)carbamoyl]amino}sulfonyl)benzoate	MIC 10 mg/L	1	acetohydroxyacid synthase (AHAS)	69	>2 log CFU reduction in Female BALB/c Dosed at 500 mg/kg
	pVXc-486 {[2-(5-{2-[(ethylcarbamoyl)amino]-6-fluoro-7-(oxolan-2-yl)-1H-1,3-benzodiazol-5-yl}pyrimidin-2-yl)propan-2-yl]oxy}phosphonic acid		1	Gyrase B	70	1 log CFU reduction at 30 and 100 mg/kg in female BALB/c mice dosed orally 5 times a week for 4 weeks

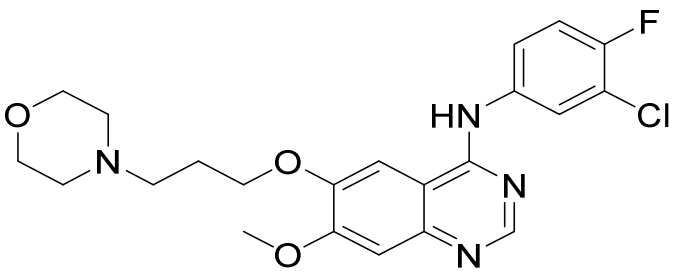
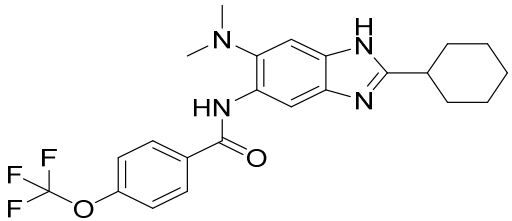
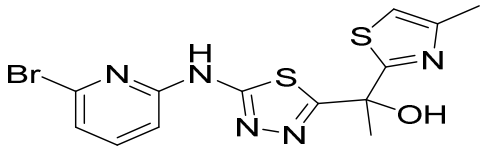
	VXc-486 3-ethyl-1-{6-fluoro-5-[2-(2-hydroxypropan-2-yl)pyrimidin-5-yl]-7-(oxolan-2-yl)-1H-1,3-benzodiazol-2-yl}urea	MIC 0.12-0.25 μg/mL	1	Gyrase B	[70]	1 log CFU reduction at 30 and 100 mg/kg in female BALB/c mice dosed orally 5 times a week for 4 weeks
	Compound 2 1-{[6-(dimethylamino)-5-methylpyrimidin-4-yl]methyl}-N-(2-hydroxyethyl)-6-methyl-1H-pyrrolo[3,2-b]pyridine-3-carboxamide	IC_{50} 0.032 μM MIC 0.5-1.56 μM	1	DprE1	[71]	~ 1 log CFU reduction at 300 mg/kg in BALB/c mice, treated 6 days a week for 4 weeks.
	Compound 65 N-[1-(1-benzofuran-3-carbonyl)-5-carbamoylpyrrolidin-3-yl]-1,3-diethyl-1H-pyrazole-5-carboxamide	IC_{50} 4 nM, MIC 0.5 μM	0	InhA	[72]	No significant reduction in log CFU Female C57BL/6, 8 day dosing

	PT070 2-(2-methylphenoxy)-5-pentylphenol	IC_{50} 50.7 nM, MIC 3.125 µg/mL	0	InhA	[73]	No significant reduction in log CFU GKO mouse dosed at 50 mg/kg twice a day, 10 day treatment
	PT091 2-(2-chlorophenoxy)-5-pentylphenol	IC_{50} 45 nM, MIC 1.56 µg/mL	0	InhA	[73]	No significant reduction in log CFU GKO mouse dosed at 25 mg/kg twice a day, 10 day treatment

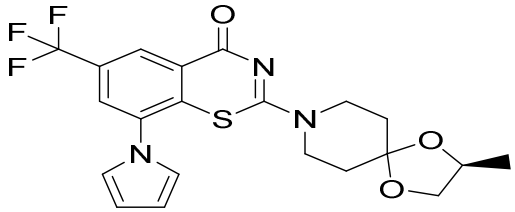
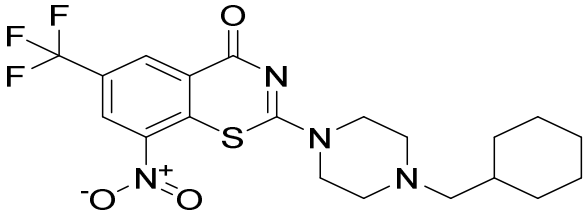
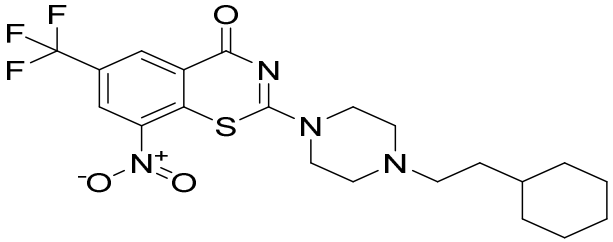
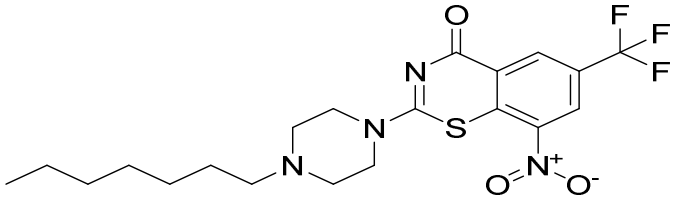
	ND 09759 N-({4-[(4-chlorophenyl)methyl]phenyl}methyl)-2,7-dimethylimidazo[1,2-a]pyridine-3-carboxamide	MIC ≤ 0.006 μM	0	cytochrome C reductase (qcrB)	[74]	0.5 log decrease in CFU in female BALB/c mice dosed orally 30 mg/kg

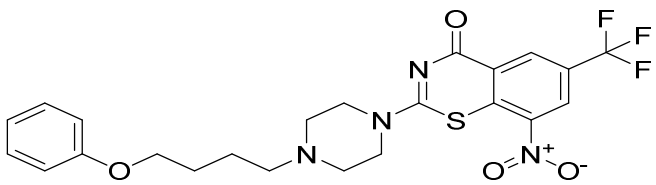
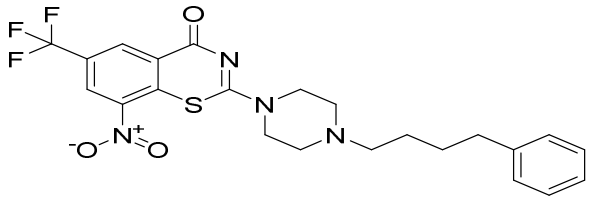
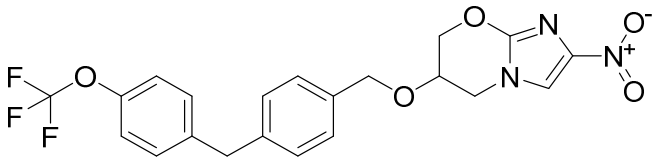
	spectinamide 1329 2-(pyridin-2-yl)-N-(4a,7,9-trihydroxy-2-methyl-6,8-bis(methylamino)decahydro-2H-benzo[b]pyrano[2,3-e][1,4]dioxin-4-yl)acetamide	MIC 1.6 µg/mL, mycobacterial ribosomes IC_{50} 1.16 µg/mL	1	RpsE	[75]	>1 log CFU , GKO mice 200 mg/kg twice daily, 22 day treatment
	spectinamide 1445 2-(5-fluoropyridin-2-yl)-N-(4a,7,9-trihydroxy-2-methyl-6,8-bis(methylamino)decahydro-2H-benzo[b]pyrano[2,3-e][1,4]dioxin-4-yl)acetamide	MIC 0.8 µg/mL, mycobacterial ribosomes IC_{50} 0.55 µg/mL	1	RpsE	[75]	>1 log CFU , GKO mice 200 mg/kg twice daily, 22 day treatment

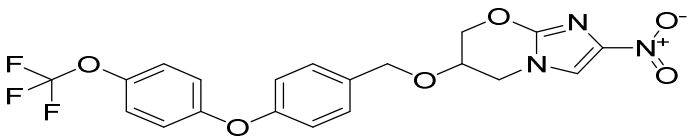
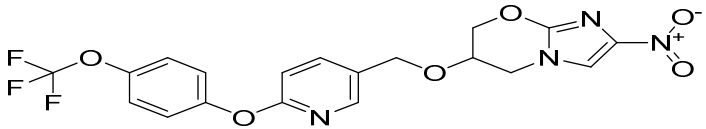
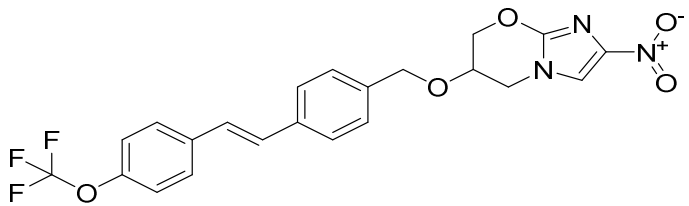
	spectinamide 1544 2-(4-methylpyridin-2-yl)-N-(4a,7,9-trihydroxy-2-methyl-6,8-bis(methylamino)decahydro-2H-benzo[b]pyrano[2,3-e][1,4]dioxin-4-yl)acetamide	MIC 0.8 µg/mL, mycobacterial ribosomes IC_{50} 0.54 µg/mL	1	RpsE	[75]	>1 log CFU , GKO mice 200 mg/kg twice daily, 22 day treatment
	spectinamide 1599 2-(5-chloropyridin-2-yl)-N-(4a,7,9-trihydroxy-2-methyl-6,8-bis(methylamino)decahydro-2H-benzo[b]pyrano[2,3-e][1,4]dioxin-4-yl)acetamide	MIC 1.6 µg/mL, mycobacterial ribosomes IC_{50} 0.53 µg/mL	1	RpsE	[75]	>1 log CFU , GKO mice 200 mg/kg twice daily, 22 day treatment

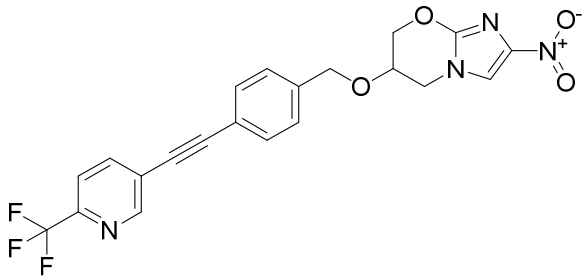
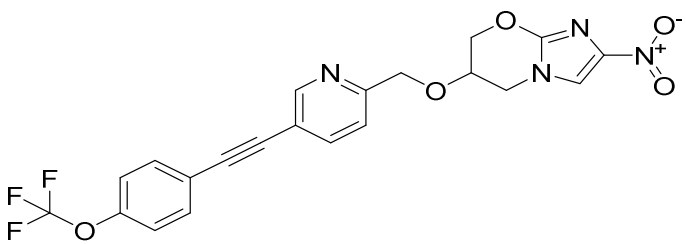
	gefitinib N-(3-chloro-4-fluorophenyl)-7-methoxy-6-(3-morpholinopropoxy)quinazolin-4-amine		0	human kinases	76	Statistically significant (but < 1 log CFU) in BALB/c mice IP injection with 100 mg/kg for 6 days
	SB-PG17G-A20 N-[2-cyclohexyl-6-(dimethylamino)-1H-1,3-benzodiazol-5-yl]-4-(trifluoromethoxy)benzamide	MIC 0.16 µg/mL	1	FtsZ	77	1.73 log CFU reductionin C57BL/6 mouse dosed 50 mg/kg IP twice a day for 15 days
	Compound 8d 1-{5-[(6-bromopyridin-2-yl)amino]-1,3,4-thiadiazol-2-yl}-1-(4-methyl-1,3-thiazol-2-yl)ethan-1-ol		0	InhA	78	0.7 log reduction in Female C57BL/6, 4 day dosing

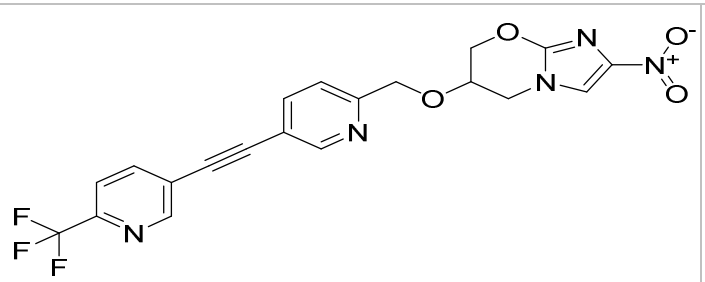
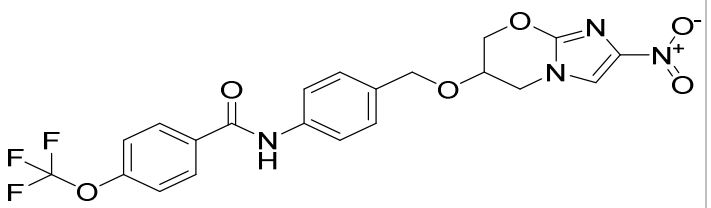
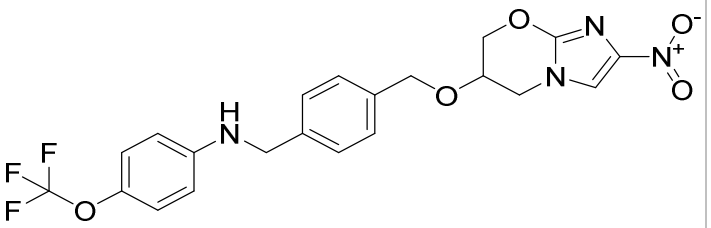
	Cyclogriselimycin (4R)-N,4-dimethyl-1-[(2S)-3-methyl-2-(N-methylacetamido)butanoyl]-N-[(3R,6S,9R,11R,15S,18S,19R,25R,28S,30S)-6,15,25-tris(butan-2-yl)-30-cyclohexyl-4,11,19,26-tetramethyl-2,5,8,14,17,21,24,27-octaoxo-3-(propan-2-yl)-20-oxa-1,4,7,13,16,23,26-heptaazatricyclo[26.3.0.0□,[13]]hentriacontan-18-yl]pyrrolidine-2-carboxamide	MIC 0.06 µg/mL	1	DnaN	[79]	~2log CFU decrease in mice dosed 100 mg/kg, 4 week treatment
	lansoprazole 2-(((3-methyl-4-(2,2,2-trifluoroethoxy)pyridin-2-yl)methyl)sulfinyl)-1H-benzo[d]imidazole	MIC 1.13 uM	1	QcrB cytochrome bc ATP synthase	[80]	(not 1 log CFU) Statistically significant CFU reduction. Female BALB/c mice dosed 300 mg kg twice a day for 9 days

	PyrBTZ01 2-[(2S)-2-methyl-1,4-dioxa-8-azaspiro[4.5]decan-8-yl]-8-(1H-pyrrol-1-yl)-6-(trifluoromethyl)-4H-1,3-benzothiazin-4-one	MIC 0.35 µM	0	DprE1	[81]	Not active in female BALBc treated 50 mg/kg, 5 days a week for 4 weeks
	PBTZ169 2-[4-(cyclohexylmethyl)piperazin-1-yl]-8-nitro-6-(trifluoromethyl)-4H-1,3-benzothiazin-4-one	MIC ≤ 0.19 µM	1	DprE1	[82]	Approx. 1 log reduction Female BALBc treated 50mg/kg, 5 days a week for 4 weeks
	PBTZ129 2-[4-(2-cyclohexylethyl)piperazin-1-yl]-8-nitro-6-(trifluoromethyl)-4H-1,3-benzothiazin-4-one	MIC 0.19 µM	1	DprE1	[82]	Approx. 1 log reduction Female BALBc treated 50mg/kg, 5 days a week for 4 weeks
	PBTZ134 2-(4-heptylpiperazin-1-yl)-8-nitro-6-(trifluoromethyl)-4H-1,3-benzothiazin-4-one	MIC 0.19 µM	1	DprE1	[82]	Approx. 1 log reduction Female BALBc treated 50 mg/kg, 5 days a week for 4 weeks

	PBTZ137 8-nitro-2-[4-(4-phenoxybutyl)piperazin-1-yl]-6-(trifluoromethyl)-4H-1, 3-benzothiazin-4-one	MIC 1.5 μM	1	DprE1	[82]	Approx. 1 log reduction Female BALBc treated 50 mg/kg, 5 days a week for 4 weeks
	PBTZ139 8-nitro-2-[4-(4-phenylbutyl)piperazin-1-yl]-6-(trifluoromethyl)-4H-1,3-benzothiazin-4-one	MIC 0.37 μM	1	DprE1	[82]	Approx. 1 log reduction Female BALBc treated 50 mg/kg, 5 days a week for 4 weeks
	Palmer 2015 11 2-nitro-6-[(4-{[4-(trifluoromethoxy)phenyl]methyl}phenyl)methoxy]-5H,6H,7H-imidazo[2,1-b][1,3]oxazine	MABA MIC 0.02 μM	1		[83]	Ratio relative to compound 1 = 1.8 Acute in vitro efficacy in BALB/c mice dosed at 100 mg/kg 5 days a week for 3 weeks

	Palmer 2015 15 2-nitro-6-({4-[4-(trifluoromethoxy)phenoxy]phenyl}methoxy)-5H,6H,7H-imidazo[2,1-b][1,3]oxazine	MABA MIC 0.03 μM	1		83	Ratio relative to compound 1 = 2.1 Acute in vitro efficacy in BALB/c mice dosed at 100 mg/kg 5 days a week for 3 weeks
	Palmer 2015 24 5-[({2-nitro-5H,6H,7H-imidazo[2,1-b][1,3]oxazin-6-yl}oxy)methyl]-2-[4-(trifluoromethoxy)phenoxy]pyridine	MABA MIC 0.06 μM	1		83	Ratio relative to compound 1 = 0.97 Acute in vitro efficacy in BALB/c mice dosed at 100 mg/kg 5 days a week for 3 weeks
	Palmer 2015 28 2-nitro-6-({4-[(E)-2-[4-(trifluoromethoxy)phenyl]ethenyl]phenyl}methoxy)-5H,6H,7H-imidazo[2,1-b][1,3]oxazine	MABA MIC 0.02 μM	1		83	Ratio relative to compound 1 = 5.8 Acute in vitro efficacy in BALB/c mice dosed at 100 mg/kg 5 days a week for 3 weeks

	Palmer 2015 35 5-(2-{4-[({2-nitro-5H,6H,7H-imidazo[2,1-b][1,3]oxazin-6-yl}oxy)meth yl]phenyl}ethynyl)-2-(trifluoromethyl)pyridine	MABA MIC 0.13 μM	1		[83]	Ratio relative to compound 1 = 5.4 Acute in vitro efficacy in BALB/c mice dosed at 100 mg/kg 5 days a week for 3 weeks
	Palmer 2015 42 2-[({2-nitro-5H,6H,7H-imidazo[2,1-b][1,3]oxazin-6-yl}oxy)methyl]-5-{2-[4-(trifluoromethoxy)phenyl]ethy nyl}pyridine	MABA MIC 0.02 μM	1		[83]	Ratio relative to compound 1 = >933 Acute in vitro efficacy in BALB/c mice dosed at 100 mg/kg 5 days a week for 3 weeks

	Palmer 2015 45 5-(2-{6-[({2-nitro-5H,6H,7H-imidazo[2,1-b][1,3]oxazin-6-]pyridin-3-yl}ethynyl)-2-(trifluoromethyl)pyridine yl}oxy)methyl	MABA MIC 0.035 μM	0		83	Ratio relative to compound 1 = 0.23 Acute in vitro efficacy in BALB/c mice dosed at 100 mg/kg 5 days a week for 3 weeks
	Palmer 2015 65 N-{4-[({2-nitro-5H,6H,7H-imidazo[2,1-b][1,3]oxazin-6-yl}oxy)methyl] phenyl}-4-(trifluoromethoxy)benzamide	MABA MIC 0.045 μM	0		83	Ratio relative to compound 1 = <0.01 Acute in vitro efficacy in BALB/c mice dosed at 100 mg/kg 5 days a week for 3 weeks
	Palmer 2015 66 N-({4-[({2-nitro-5H,6H,7H-imidazo[2,1-b][1,3]oxazin-6-yl}oxy)methyl] phenyl}methyl)-4-(trifluoromethoxy)aniline	MABA MIC 0.06 μM	0		83	Ratio relative to compound 1 = 0.02 Acute in vitro efficacy in BALB/c mice dosed at 100 mg/kg 5 days a week for 3 weeks

	Palmer 2015 76 1-{4-[({2-nitro-5H,6H,7H-imidazo[2,1-b][1,3]oxazin-6-yl}oxy)methyl]phenyl}-4-[4-(trifluoromethoxy)phenyl]piperazine	MABA MIC 0.02 μM	1		83	Ratio relative to compound 1 = 2.8 Acute in vitro efficacy in BALB/c mice dosed at 100 mg/kg 5 days a week for 3 weeks
	Palmer 2015 77 1-{4-[({2-nitro-5H,6H,7H-imidazo[2,1-b][1,3]oxazin-6-yl}oxy)methyl]phenyl}-4-[5-(trifluoromethyl)pyridin-2-yl]piperazine	MABA MIC 0.02 μM	0		83	Ratio relative to compound 1 = 0.24 Acute in vitro efficacy in BALB/c mice dosed at 100 mg/kg 5 days a week for 3 weeks
	Palmer 2015 84 2-nitro-6-((4-(1-(4-(trifluoromethoxy)phenyl)-1l3-morpholino)benzyl) oxy)-6,7-dihydro-5H-imidazo[2,1-b][1,3]oxazine	MABA MIC 0.02 μM	0		83	Ratio relative to compound 1 = 0.4 Acute in vitro efficacy in BALB/c mice dosed at 100 mg/kg 5 days a week for 3 weeks

Table S2. Leave out 50% x 100 cross-validation for the Bayesian N = 784 model (mean and standard deviation).

External ROC	Internal ROC	Concordance (%)	Specificity (%)	Sensitivity (%)
0.73 ± 0.02	0.75 ± 0.02	67.02 ± 2.23	74.15 ± 9.11	59.05 ± 9.22

Table S3 PA-824 analogs in test set predicted with Bayesian N = 784 model (mean closest distance = 0.26 with a standard deviation of 0.07

Molecule name and IUPAC ID	Closest distance to N784 model	In vivo classification	predicted
Amino-824 2-nitro-N-{[4-(trifluoromethoxy)phenyl]methyl}-5H,6H,7H-imidazo[2,1-b][1,3]oxazin-6-amine	0.32	1	1
AminoEthyl-824 2-nitro-N-{2-[4-(trifluoromethoxy)phenyl]ethyl}-5H,6H,7H-imidazo[2, 1-b][1,3]oxazin-6-amine	0.34	1	1
NI-135 N-{[3-chloro-4-(trifluoromethoxy)phenyl]methyl}-2-nitro-5H,6H,7H-imidazo [2,1-b][1,3]oxazin-6-amine	0.37	1	1
NI-136 N-{[3-fluoro-4-(trifluoromethoxy)phenyl]methyl}-2-nitro-5H,6H,7H-im idazo[2,1-b][1,3]oxazin-6-amine	0.37	1	1
NI-182 7-ethyl-2-nitro-N-{[4-(trifluoromethoxy)phenyl]methyl}-5H,6H,7H-imidazo [2,1-b][1,3]oxazin-6-amine	0.36	1	1
NI-302 2'-nitro-6'-{[4-(trifluoromethoxy)phenyl]methoxy}-5',6'-dihydrospiro[cyclopropane-1,7'-imidazo[2,1-b][1,3]oxazine]	0.24	1	1

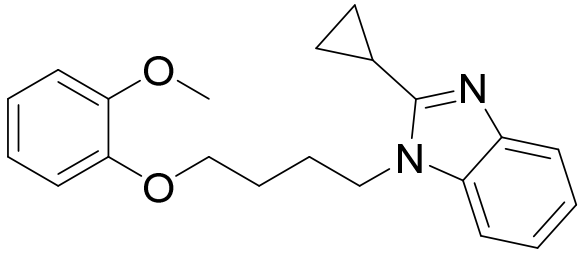
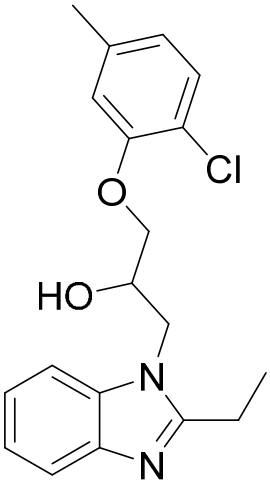
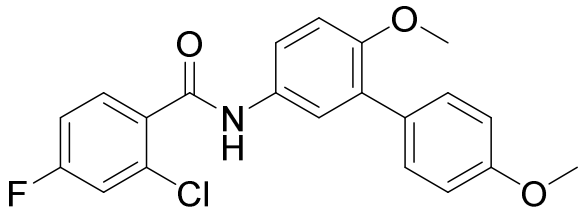
NI-622 2-({2-nitro-5H,6H,7H-imidazo[2,1-b][1,3]oxazin-6-yl}oxy)-1-{4-[4-(tri fluoromethoxy)phenyl]piperazin-1-yl}ethan-1-one	0.18	0	1
NI-644 2-({2-nitro-5H,6H,7H-imidazo[2,1-b][1,3]oxazin-6-yl}oxy)-1-{4-[5-(trifluoromethoxy)pyridin-2-yl]piperazin-1-yl}ethan-1-one	0.30	0	1
Palmer 2015 11 2-nitro-6-[(4-{[4-(trifluoromethoxy)phenyl]methyl}phenyl)methoxy]-5 H,6H,7H-imidazo[2,1-b][1,3]oxazine	0.18	1	1
Palmer 2015 15 2-nitro-6-({4-[4-(trifluoromethoxy)phenoxy]phenyl}methoxy)-5H,6H,7H-imidazo[2,1-b][1,3]oxazine	0.18	1	1
Palmer 2015 24 5-[({2-nitro-5H,6H,7H-imidazo[2,1-b][1,3]oxazin-6-yl}oxy)methyl]-2-[4-(trifluoromethoxy)phenoxy]pyridine	0.18	1	1
Palmer 2015 28 2-nitro-6-({4-[(E)-2-[4-(trifluoromethoxy)phenyl]ethenyl]phenyl}methoxy)-5H,6H,7H-imidazo[2,1-b][1,3]oxazine	0.18	1	1
Palmer 2015 35 5-(2-{4-[({2-nitro-5H,6H,7H-imidazo[2,1-b][1,3]oxazin-6-yl}oxy)meth yl]phenyl}ethynyl)-2-(trifluoromethyl)pyridine	0.20	1	1
Palmer 2015 42 2-[({2-nitro-5H,6H,7H-imidazo[2,1-b][1,3]oxazin-6-yl}oxy)methyl]-5-{2-[4-(trifluoromethoxy)phenyl]ethynyl}pyridine	0.21	1	1
Palmer 2015 45 5-(2-{6-[({2-nitro-5H,6H,7H-	0.21	0	1

imidazo[2,1-b][1,3]oxazin-6-]pyridin-3-yl}ethynyl)-2-(trifluoromethyl)pyridine yl}oxy)methyl			
Palmer 2015 65 N-{4-[({2-nitro-5H,6H,7H-imidazo[2,1-b][1,3]oxazin-6-yl}oxy)methyl] phenyl}-4-(trifluoromethoxy)benzamide	0.26	0	1
Palmer 2015 66 N-({4-[({2-nitro-5H,6H,7H-imidazo[2,1-b][1,3]oxazin-6-yl}oxy)methyl] phenyl}methyl)-4-(trifluoromethoxy)aniline	0.27	0	1
Palmer 2015 76 1-{4-[({2-nitro-5H,6H,7H-imidazo[2,1-b][1,3]oxazin-6-yl}oxy)methyl]phenyl}-4-[4-(trifluoromethoxy)phenyl]piperazine	0.26	1	1
Palmer 2015 77 1-{4-[({2-nitro-5H,6H,7H-imidazo[2,1-b][1,3]oxazin-6-yl}oxy)methyl]phenyl}-4-[5-(trifluoromethyl)pyridin-2-yl]piperazine	0.30	0	1
Palmer 2015 84 2-nitro-6-((4-(1-(4-(trifluoromethoxy)phenyl)-1l3-morpholino)benzyl) oxy)-6,7-dihydro-5H-imidazo[2,1-b][1,3]oxazine	0.27	0	1

Table S4. External ROC values for predictions on the new compounds (N = 60) generated with the N = 784 *in vivo* models

Bayesian	SVM	Single Tree	Forest
0.558	0.549	0.526	0.533

Table S5. Intersection of top consensus predictions on the set of 177 *Mtb* leads from GSK, using the Combined TB dual event Bayesian + Full $t_{1/2}$ MLM Bayesian and the TAACF-CB2 dual event Bayesian + Full $t_{1/2}$ MLM Bayesian (*i.e.*, a consensus of both consensus approaches).

Compound Name	2D Structure
GSK994258A	
GSK889423A	
GSK1733953A	

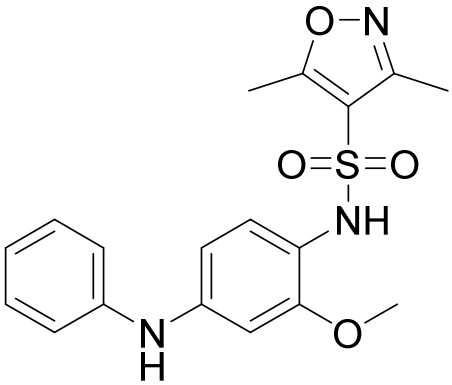
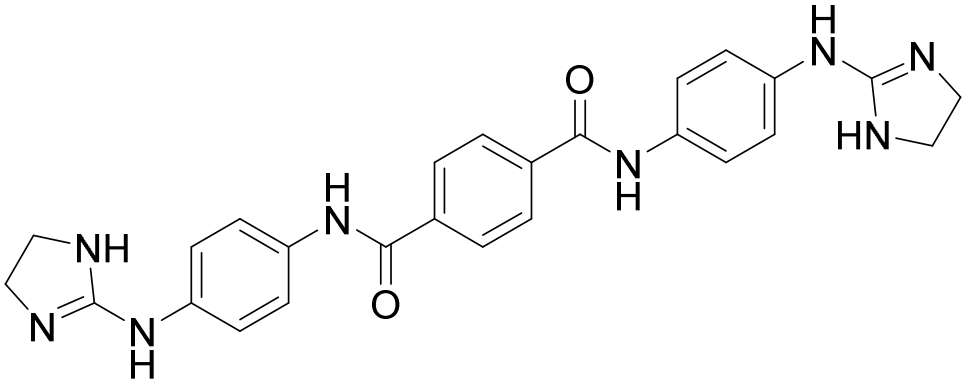
GSK270670A	
GW356807A	

Table S6. Consensus *in vivo* active predictions on the set of 177 *Mtb* leads from GSK, using the Combined TB dual event Bayesian + Full $t_{1/2}$ MLM Bayesian

Compound Name [a]	Full $t_{1/2}$ MLM Bayesian (just FCFP_6) Score [b]	Combined TB dual-event Bayesian Score
GSK994258A	11.48	3.75
GSK889423A	4.31	9.68
BRL-51091AM	2.96	7.20
GSK1733953A	2.29	3.25
GV187303X	2.27	3.84
GSK270670A	1.82	4.16
GW356807A	1.77	2.94
BRL-10988SA	0.30	5.49
BRL-51093AM	-0.55	3.50
GSK1180781A	-0.58	13.92
SB-435634	-0.66	1.47
GSK1925843A	-0.91	3.56
GSK353071A	-0.91	2.20
GSK731389A	-1.12	6.26
GSK1742694A	-1.39	2.80
GR223839X	-2.19	4.36
GR153167X	-2.19	2.89
GSK1385423A	-2.78	1.86
GSK1310678A	-2.97	1.35
GSK749336A	-3.08	1.83
GW859039X	-3.21	18.92

GW857165X	-3.32	4.79
GSK426032A	-3.43	4.23
GSK358607A	-3.63	3.78
GSK1829728A	-3.65	1.14
GSK1589671A	-3.94	20.34
BRL-8903SA	-4.20	3.64

[a] Compound names in bold-faced type correspond to molecules that were also present in the consensus approach that utilized the TAACF-CB2 dual event Bayesian + the Full $t_{1/2}$ MLM Bayesian (*i.e.*, they are part of the consensus of both consensus approaches; see **Table S3**).

[b] Full $t_{1/2}$ MLM Bayesian scores highlighted in green correspond to compounds that had a Bayesian score >0 for both this Bayesian and also the Pruned $t_{1/2}$ MLM Bayesian (which used all 9 descriptors). Thus, according to the published histogram-based strategy, these compounds have a higher probability of being stable than the ones listed below them.

Table S7. Consensus *in vivo* active predictions on the set of 177 *Mtb* leads from GSK, using the TAACF-CB2 dual event Bayesian + Full $t_{1/2}$ MLM Bayesian

Compound Name [a]	Full $t_{1/2}$ MLM Bayesian (just FCFP_6) Score [b]	TAACF-CB2 dual-event Bayesian Score
GSK994258A	11.48	3.48
GSK829969A	6.64	3.81
GSK316438A	4.98	4.60
GSK889423A	4.31	3.47
GSK3011724A	3.37	0.74
GSK1650514A	3.12	6.18
GSK468214A	2.69	3.71
GSK1733953A	2.29	0.81
GSK353496A	2.18	0.88
GSK270670A	1.82	3.16
GW356807A	1.77	0.53
GSK1955236A	0.93	0.68
GSK437009A	0.49	1.30
GSK353069A	-0.25	2.47
GSK991960A	-0.32	4.48
GSK130506A	-0.46	3.51
GSK1180781A	-0.58	0.92
SB-435634	-0.66	1.64
GSK1925843A	-0.91	2.88
GSK353071A	-0.91	2.91
GSK1051703A	-0.91	2.07

GSK731389A	-1.12	0.86
GSK735816A	-1.28	1.09
GSK1742694A	-1.39	2.81
GSK636544A	-2.28	1.40
GSK547487A	-2.71	1.59
GSK1385423A	-2.78	1.10
GSK735826A	-2.97	2.16
GSK749336A	-3.08	0.91
GW859039X	-3.21	3.11
GW857165X	-3.32	2.39
GSK426032A	-3.43	1.90
GSK358607A	-3.63	3.76
GSK1829728A	-3.65	1.00
GSK920703A	-3.85	2.21
GSK888636A	-3.87	1.12
GSK1859936A	-3.92	1.14
GSK1589671A	-3.94	6.03
GSK690382A	-4.20	2.82
GSK1570606A	-4.37	0.84

[a] Compound names in bold-faced type correspond to molecules that were also present in the consensus approach that utilized the Combined TB dual event Bayesian + the Full $t_{1/2}$ MLM Bayesian (*i.e.*, they are part of the consensus of both consensus approaches; see **Table S3**).

[b] Full $t_{1/2}$ MLM Bayesian scores highlighted in green correspond to compounds that had a Bayesian score >0 for both this Bayesian and also the Pruned $t_{1/2}$ MLM Bayesian

(which used all 9 descriptors). Thus, these compounds have a higher probability of being stable than the ones listed below them.

Figure S1. PDF file of Honeycomb clustering – separate file

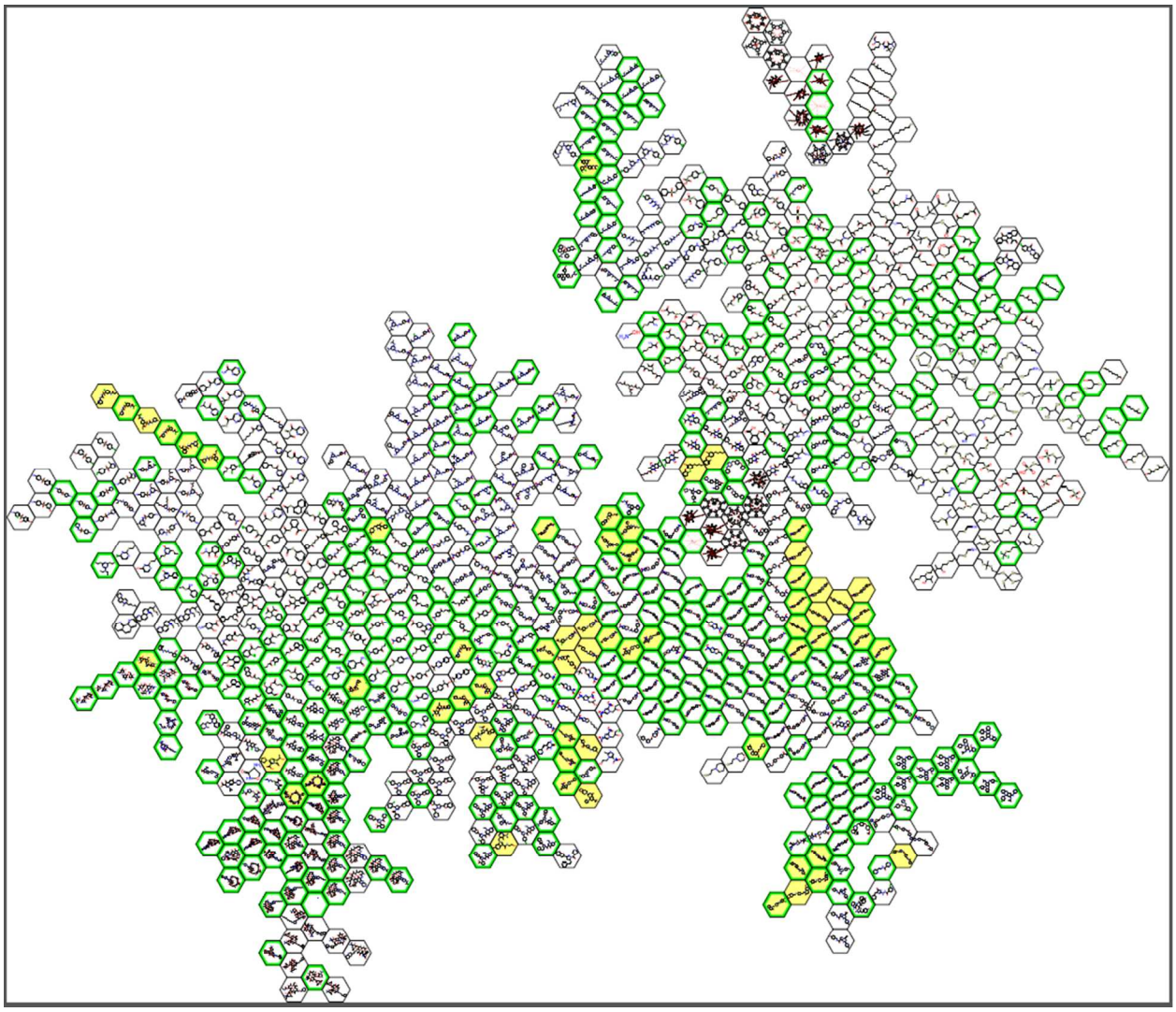
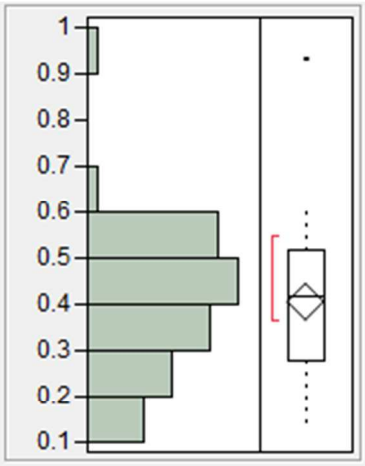


Figure S2. Distribution of Closest Distance Scores and statistics calculated in JMP version 8.0.1 (SAS Institute, Cary, NC).



Quantiles

100.0%	maximum	0.92945
99.5%		0.92945
97.5%		0.75889
90.0%		0.57274
75.0%	quartile	0.51522
50.0%	median	0.41859
25.0%	quartile	0.27724
10.0%		0.18637
2.5%		0.15251
0.5%		0.12325
0.0%	minimum	0.12325

Moments

Mean	0.4040845
Std Dev	0.1479119
Std Err Mean	0.0190953
Upper 95% Mean	0.4422942
Lower 95% Mean	0.3658747
N	60