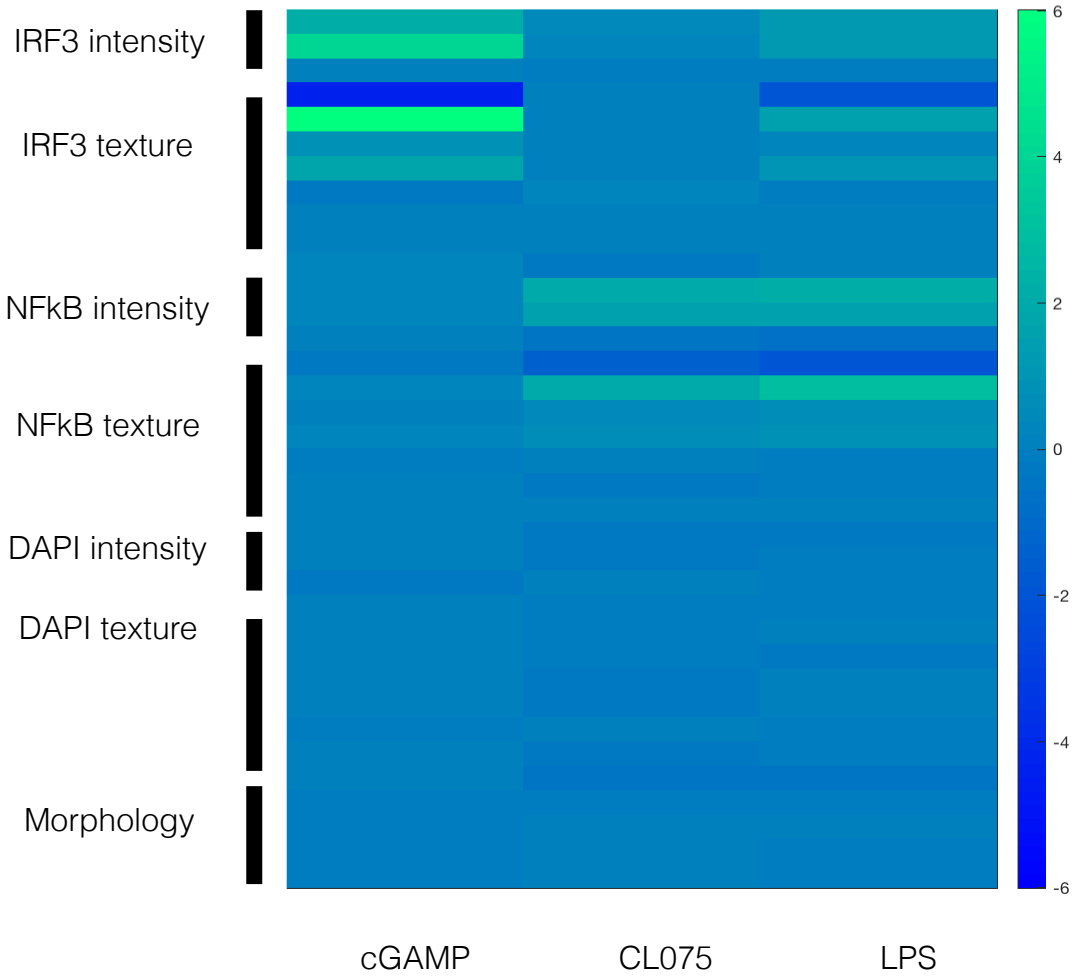
Supporting Information for:

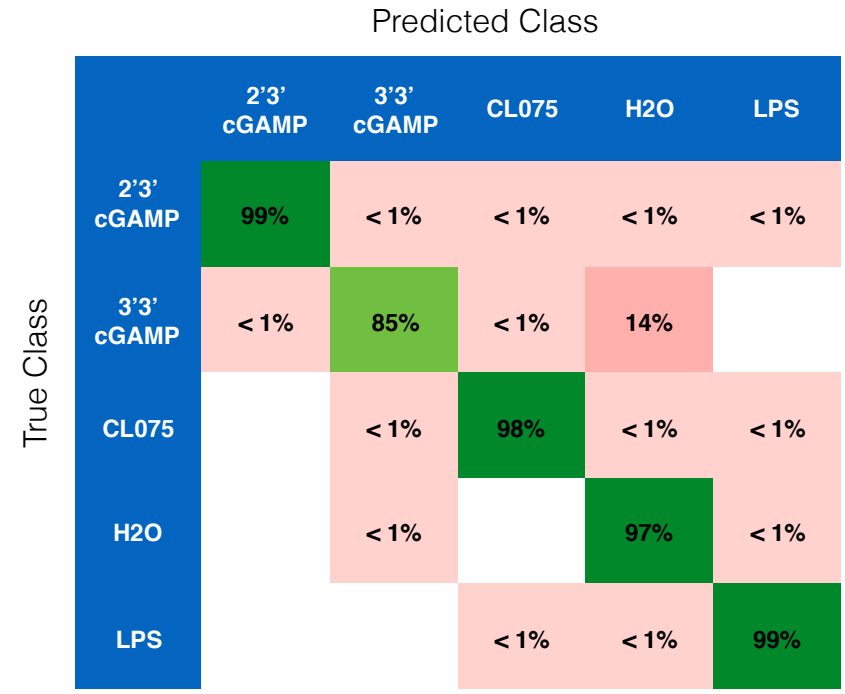
"A high content screen in macrophages identifies small molecule modulators of STING-IRF3 and NFkB signaling"

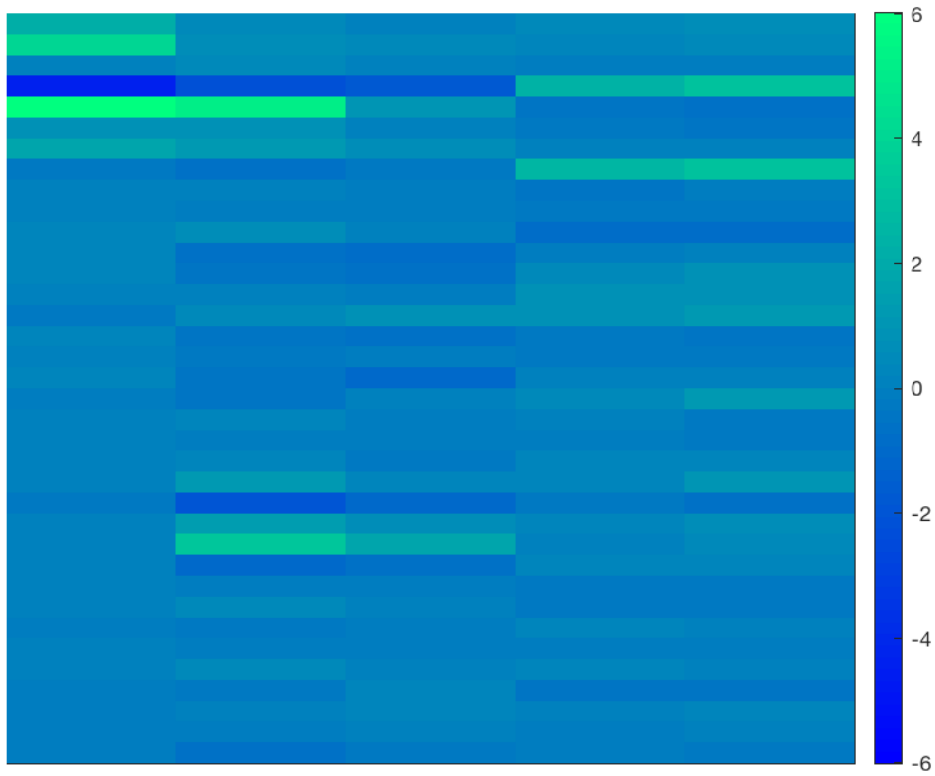
Peter David Koch[1,2], Howard Miller[3], Gary Yu[3], John Tallarico[3], Peter K. Sorger[1,2], Yuan Wang[3], Yan Feng[3], Jason Ray Thomas[3], Nathan Thomas Ross[3], Timothy Mitchison[1,2]

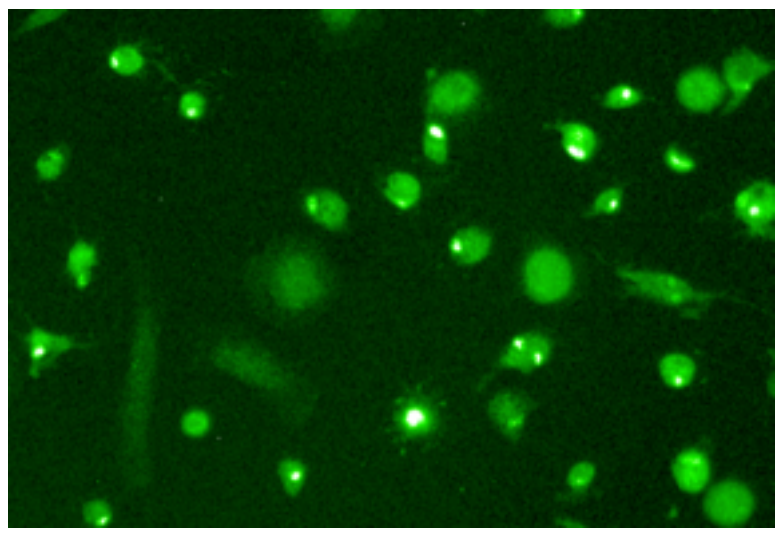
1. Department of Systems Biology, Harvard Medical School, 200 Longwood Ave., Boston, MA 02115
2. Laboratory of Systems Pharmacology, Harvard Medical School, 200 Longwood Ave., Boston, MA 02115
3. Chemical Biology and Therapeutics, Novartis Institutes for Biomedical Research, 181 Massachusetts Ave., Cambridge, MA 02139

This file contains supplemental figures 1-11, as well as structures of small molecules used in secondary screening (tables 1 and 2).









TPEN

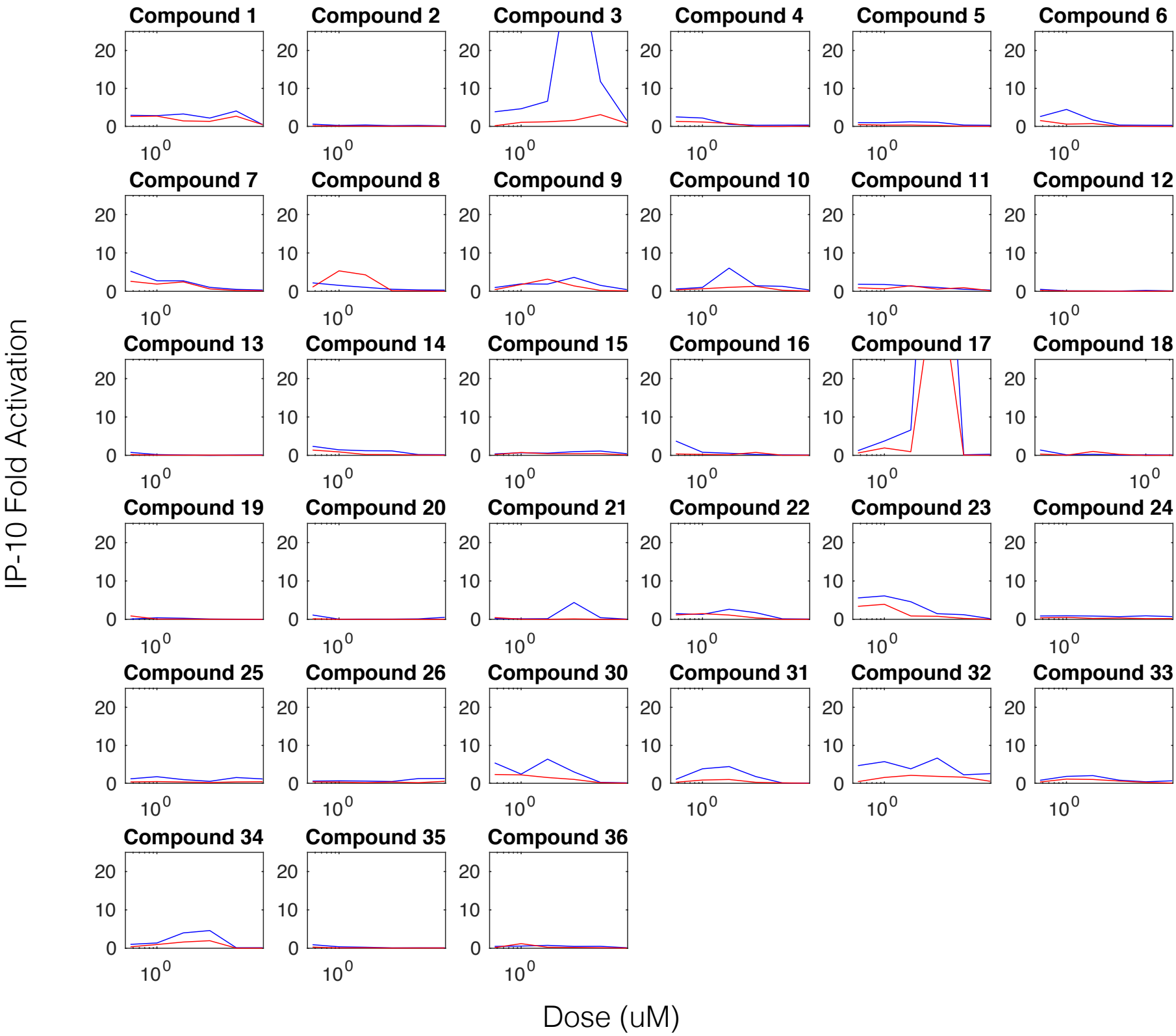
Supplement figure 1.

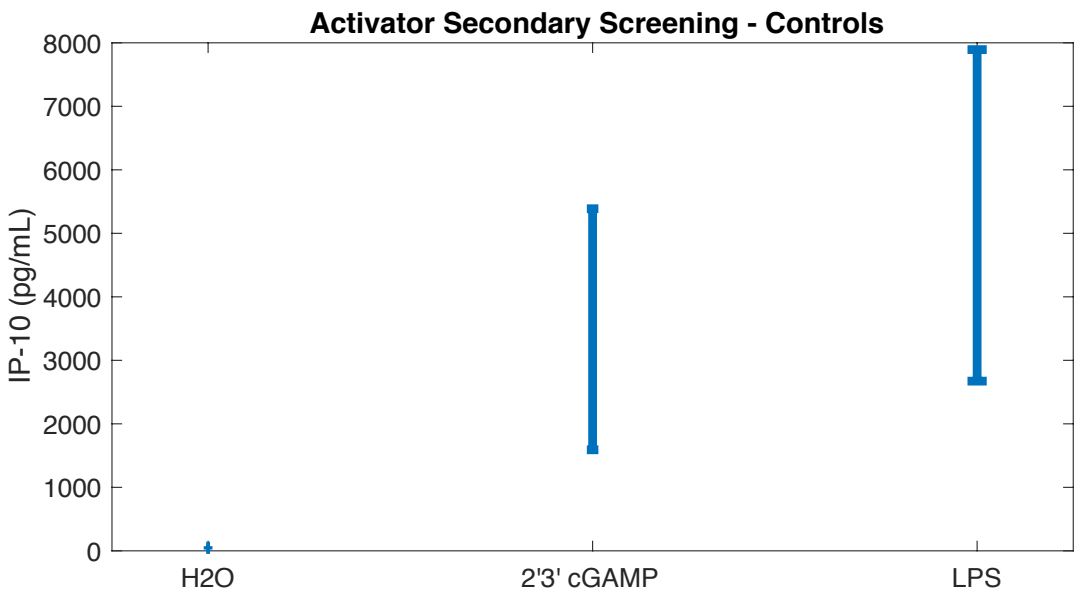
a) HeatMap of features affected by immune stimuli. In the activator screen, a total of 36 features were computed from the high-content analysis (see methods). The immune stimuli controls mainly affected nuclear intensities and two texture parameters. These features indicate nuclear translocation. Note that 2'3' cGAMP affects IRF3 features, CL075 affects NFkB features, and LPS affects both.

b) A quadratic kernel support vector machine was trained using data from our control wells (H2O, 2'3' cGAMP, 3'3' cGAMP, CL075, and LPS), with 10-fold cross validation, using the Machine Learning Toolbox in MATLAB. Confusion matrix for SVM shows that most groups are accurately classified with the exception of 3'3' cGAMP which had a weak effect in our assay. We note though that in developing this classifier, we found that the accuracy of 3'3' cGAMP classification was greatly improved by inclusion of more features (data not shown). This SVM was subsequently used to classify every well in the screen. For every compound, 12 wells were classified (6 doses ranging from 10 uM to 31.6 nM, in duplicate). Of the 12 wells, any compound that registered 2'3' cGAMP at least twice was picked for follow up.

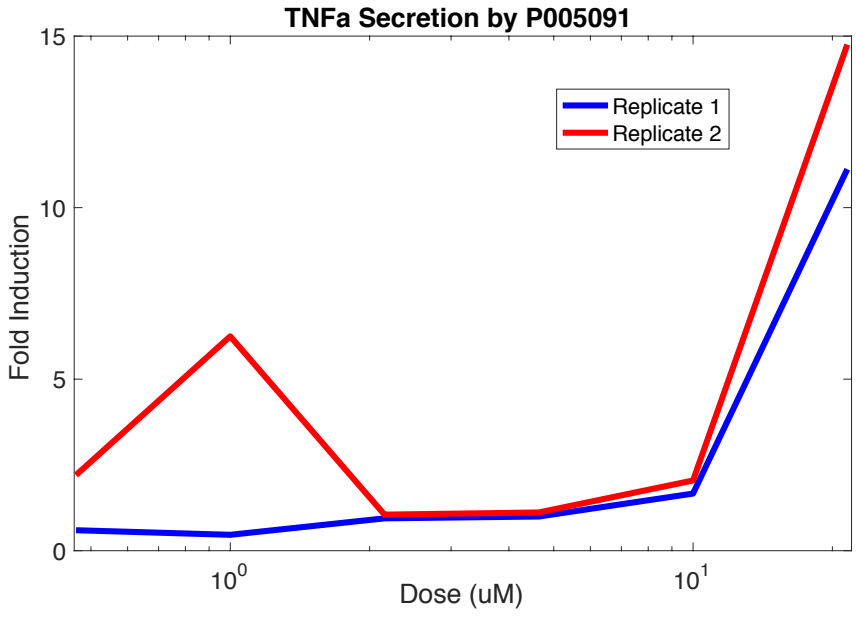
c) TPEN and colchicine both affected IRF3 nuclear intensity and localization to similar levels. However, TPEN more closely mimicked the effect on texture, and colchicine affected other parameters that cGAMP did not. This resulted in TPEN being closer to cGAMP in the multidimensional space, and being selected for follow up. Thus, inclusion of additional features did offer some benefit to our analysis.

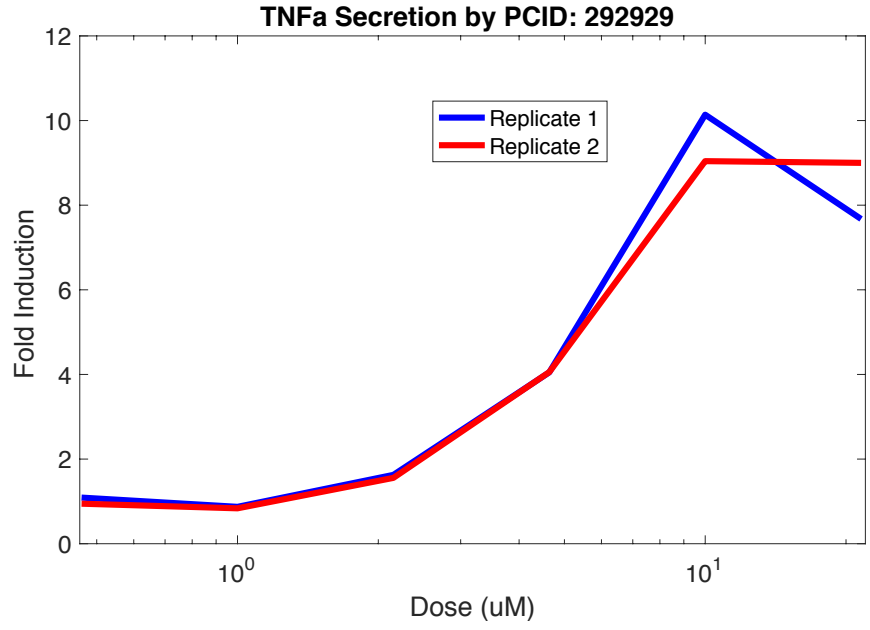
d) Representative image of TPEN treated cells shows nuclear puncta in the IRF3 channel.

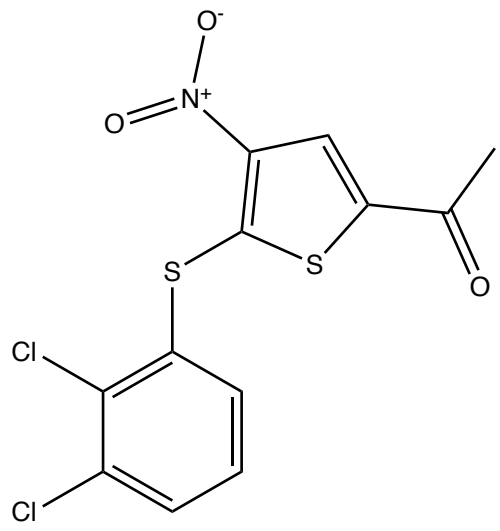




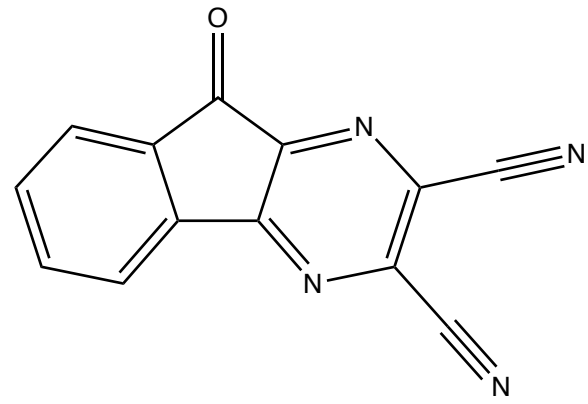
Supplement figure 2. Re-test of selected agonists in IP-10 ELISA assay. a) Most compounds do not reliably induce IP-10 secretion. Only compound 17 (TPEN) scores significantly in duplicate Compound numbering correspond to table 1. b) Controls for activator secondary screening. Both 2'3' cGAMP (34 uM) and LPS (2.5 ug/mL) induce IP-10 secretion. Error bars reflect median absolute deviation. Error bars reflect technical replicates, N = 16 for H2O, and N=12 for 2'3' cGAMP and LPS.





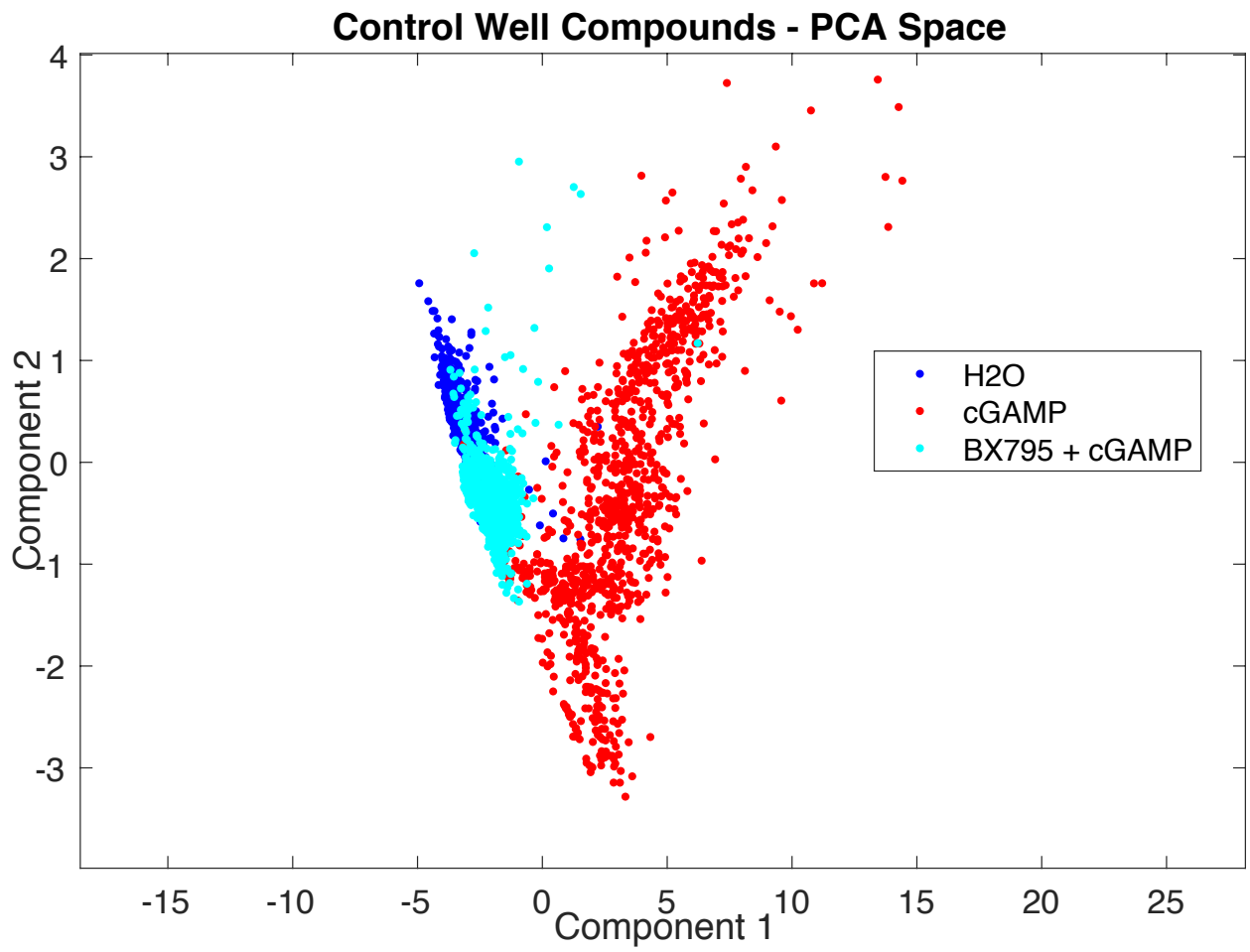


P005091
PCID: 2819993



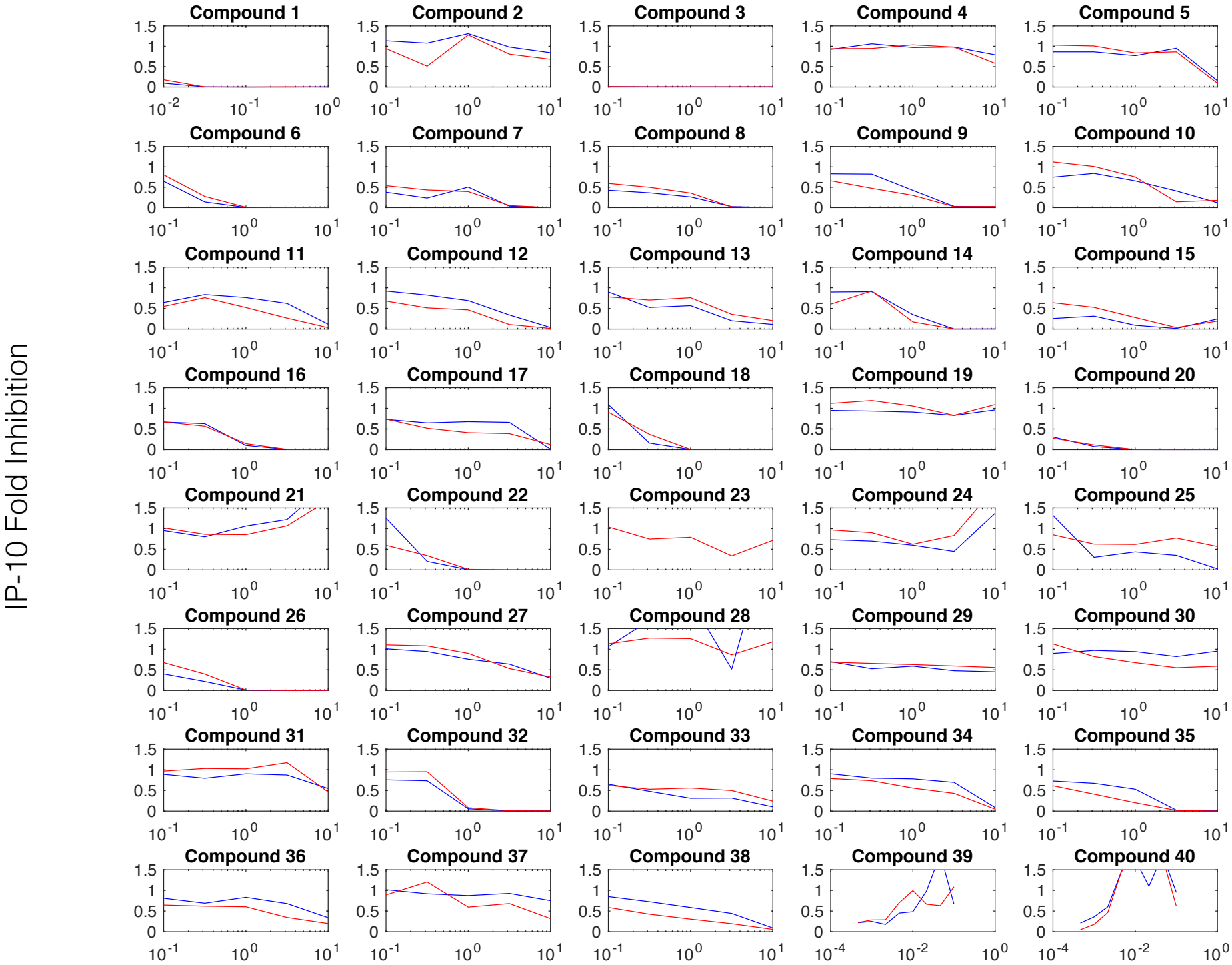
PCID: 292929

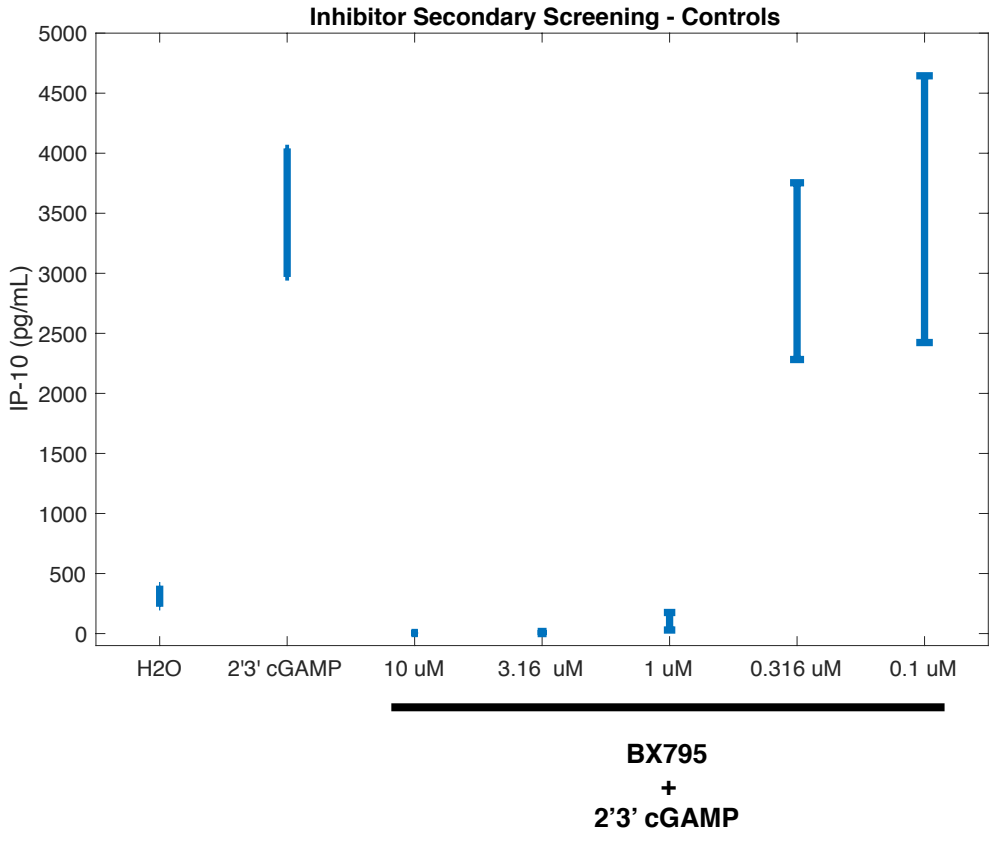
Supplement figure 3: Two molecules that have inhibitory activity against ubiquitin hydrolyses induced TNFa secretion significantly, in duplicate at high doses.



Supplement figure 4. We also used high content data in the inhibitor screen analysis. We examined data from negative control (H2O) and cGAMP treated wells, looking at the same 36 computed features, and performed principal component analysis (PCA) to reduce the dimensionality to 6 components. Above is a plot of the H2O wells and cGAMP wells, and also BX795 + cGAMP wells. As expected, BX795 wells overlap with the H2O wells.

The separation between H2O and cGAMP wells may not seem strong, but note that this plot only shows 2 components. Additionally, in selecting hits, we computed the Mahalanobis distance between each well and the H2O control. This distance incorporates the standard deviation of each component. The H2O wells are tightly constrained in a band, so even if a well is slightly displaced, it is penalized depending on the direction.





Supplement figure 5. Re-test of selected antagonists in IP-10 ELISA assay. a) Most compounds block IP-10 secretion in secondary assays Compounds 1-38 correspond to table 2. Compound 39 and 40 are bafilomycin A1 and concanomycin A. b) Controls for secondary assay. Cells were treated with 30 uM of 2'3' cGAMP and/or BX795 (concentrations indicated on plot), a TBK1 inhibitor. Error bars correspond to technical replicates, N = 6 for all TBK1 inhibitor doses, N = 18 for 2'3' cGAMP, and N = 24 for H2O

Enriched Targets (binary)

	Gene name (All	p-value
1	JAK3	1.13E-41
2	JAK2	2.47E-34
3	AURKB	1.25E-28
4	AURKA	3.29E-27
5	TYK2	4.04E-27
6	PKN1	7.02E-26
7	**ALK**	1.21E-25
8	**TBK1**	2.13E-25
9	JAK1	1.03E-24
10	PHKG1	5.46E-23
11	MAP3K19	1.26E-21
12	AAK1	1.26E-21
13	SRPK1	1.65E-21
14	FLT3	5.94E-21
15	GRK1	1.21E-20
16	STK4	1.27E-20
17	**PRKAA1**	1.87E-20
18	LRRK2	2.54E-20
19	SRPK3	3.33E-20
20	MYLK	5.90E-20

	Gene name (Remove Tbk1 inhibitors)	p-value	Number of inhibitors
1	JAK3	1.88E-19	22
2	JAK1	1.55E-13	15
3	JAK2	1.36E-11	22
4	AURKA	3.53E-11	23
5	TYK2	1.18E-10	11
6	WEE1	3.70E-09	6
7	ZAP70	4.23E-09	10
8	MYLK	4.41E-09	11
9	AXL	7.57E-09	15
10	BMPR1B	7.65E-09	6
11	**ALK**	2.32E-08	18
12	PHKG1	9.31E-08	6
13	PRKCH	1.19E-07	26
14	LYN	1.23E-07	22
15	PRKCD	1.66E-07	26
16	PTK2	2.07E-07	12
17	ACVR1	3.39E-07	9
18	**MAPKAPK5**	4.70E-07	5
19	AURKB	5.06E-07	25
20	PRKCE	8.08E-07	27

Supplement figure 6.

a) Top 20 genes enriched using hypergeometric enrichment as described in fig. 7. Shown are genes enriched with all compounds included (left) and genes enriched with Tbk1-active inhibitors removed (right). The number of inhibitors column corresponds to the number of small molecules in the screen that biochemical activity against the target, with IC50 < 1 uM

b) Partial correlation scores. MAPKAPK5 and ALK show strong correlations

Correlated Targets (continuous)

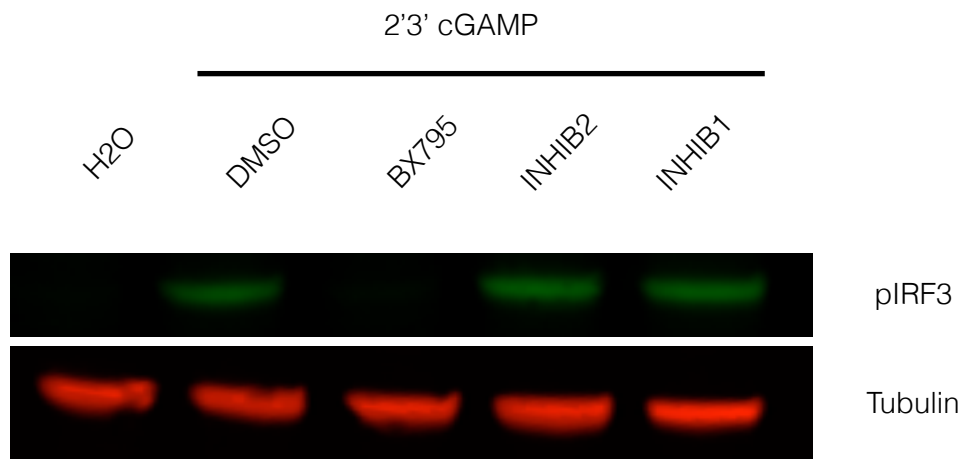
	Gene name	p-value
1	**ALK**	7.13E-22
2	BTK	2.02E-12
3	CHUK	3.25E-11
4	PRKCH	3.55E-10
5	PRKCA	4.97E-10
6	TGFBR1	1.42E-08
7	ITK	6.05E-07
8	AURKB	2.63E-06
9	PRKCB	3.00E-06
10	**MAPKAPK5**	7.30E-06
11	EIF2AK2	1.59E-05
12	PKN1	3.50E-05
13	SYK	8.83E-05
14	FGR	1.80E-04
15	MKNK1	2.38E-04
16	FLT3	3.49E-04
17	PRKD3	4.94E-04
18	DDR1	5.81E-04
19	FGFR3	6.02E-04
20	MAP2K6	8.79E-04

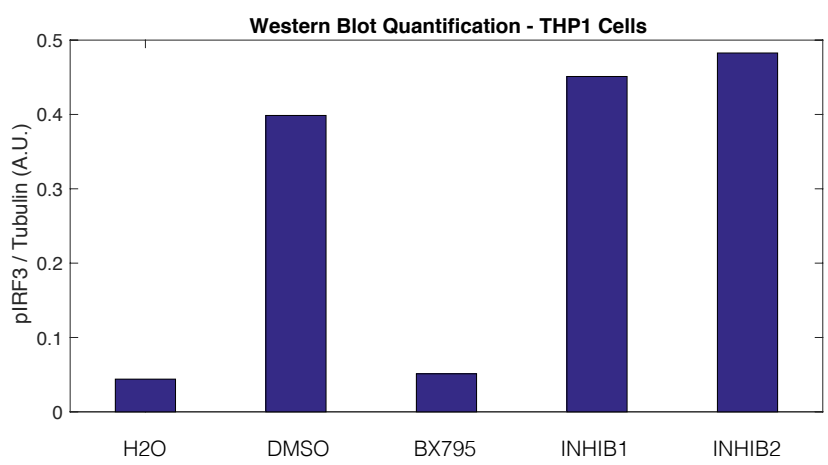
Description of method

We also computed linear correlations between IRF3 nuclear score and biochemical inhibition profile, for every kinase gene.

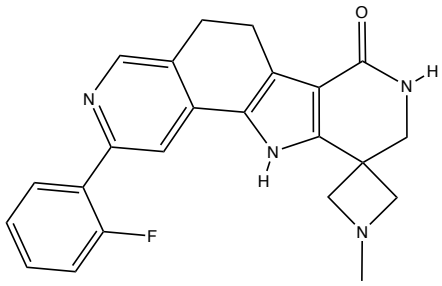
The first challenge is to compute a biochemical inhibition profile. We use biological data at each dose point, but only have a single IC50 for biochemical information. To compute a biochemical inhibition profile at every dose point, we assume that the percent activity remaining of a kinase, in the presence of an inhibitor, f = 1/(1+(alpha*dose/IC50)), where alpha is free parameter, accounting for the fact that kinase inhibitors need to be used at doses greater than their biochemical IC50 (we set alpha = 0.3).

We next compute the correlation between IRF3 score and biochemical inhibition profile, using ALL of our kinase inhibitors, for every kinase gene. However, to minimize transitive effects, we compute the partial correlation. To illustrate this effect, suppose Kinase A controls IRF3 activity. We should then ideally see a correlation between IRF3 score and the biochemical inhibition profile for Kinase A. Next, consider Kinase B, which is not in the pathway. However, suppose Kinase B, is commonly inhibited by the same molecules that inhibit Kinase A. Then it is possible to see a correlation between IRF3 score and the biochemical inhibition score for Kinase B. Thus, Kinase B may show some correlation just through a transitive effect. To minimize transitive effects, we compute a correlation, in which we the calculate the correlation between IRF3 and Kinase A over all inhibitors, controlling for all other inhibition profiles of all other kinases.

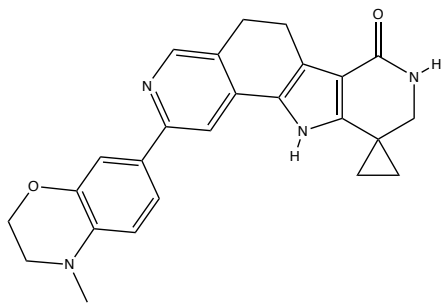




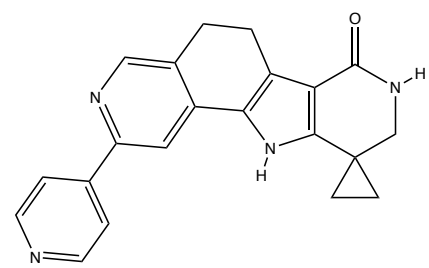
Supplement figure 7. Western blot and quantification for pIRF3 in THP1- cells, as in fig. 6d. Inhibitor doses are 10 uM and 2’3’ cGAMP dose is 62.5 uM. Cells were pre-treated with inhibitor for 1 hour and with cGAMP for 4 hours.



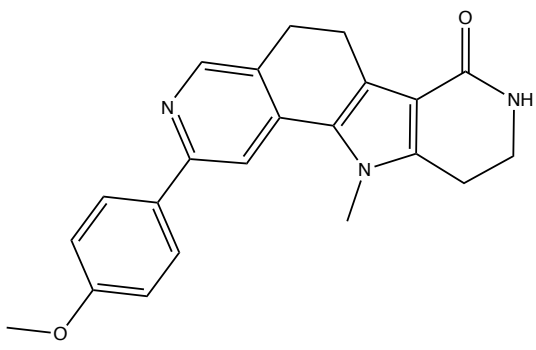
INHIB1X



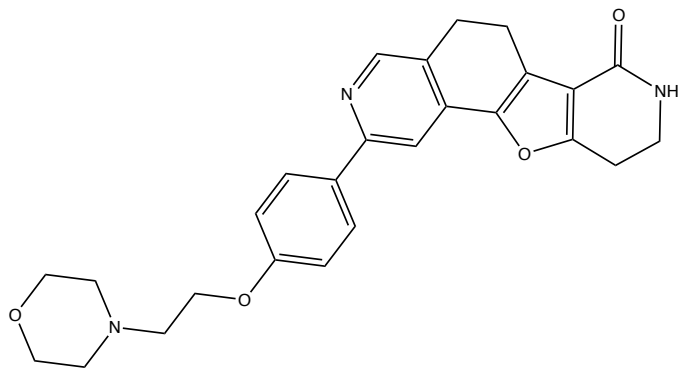
INHIB2



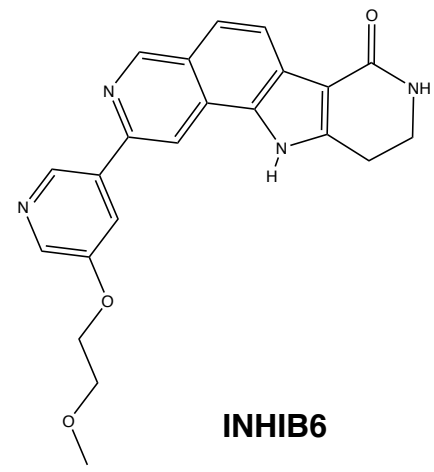
INHIB3



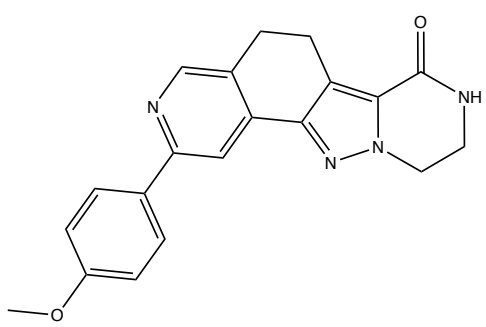
INHIB4



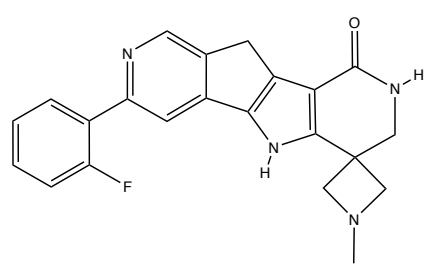
INHIB5



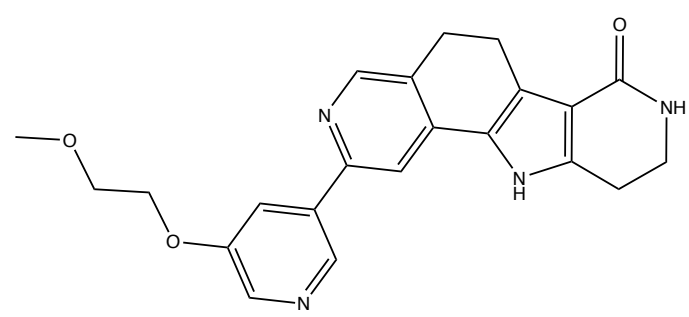
INHIB6



INHIB7



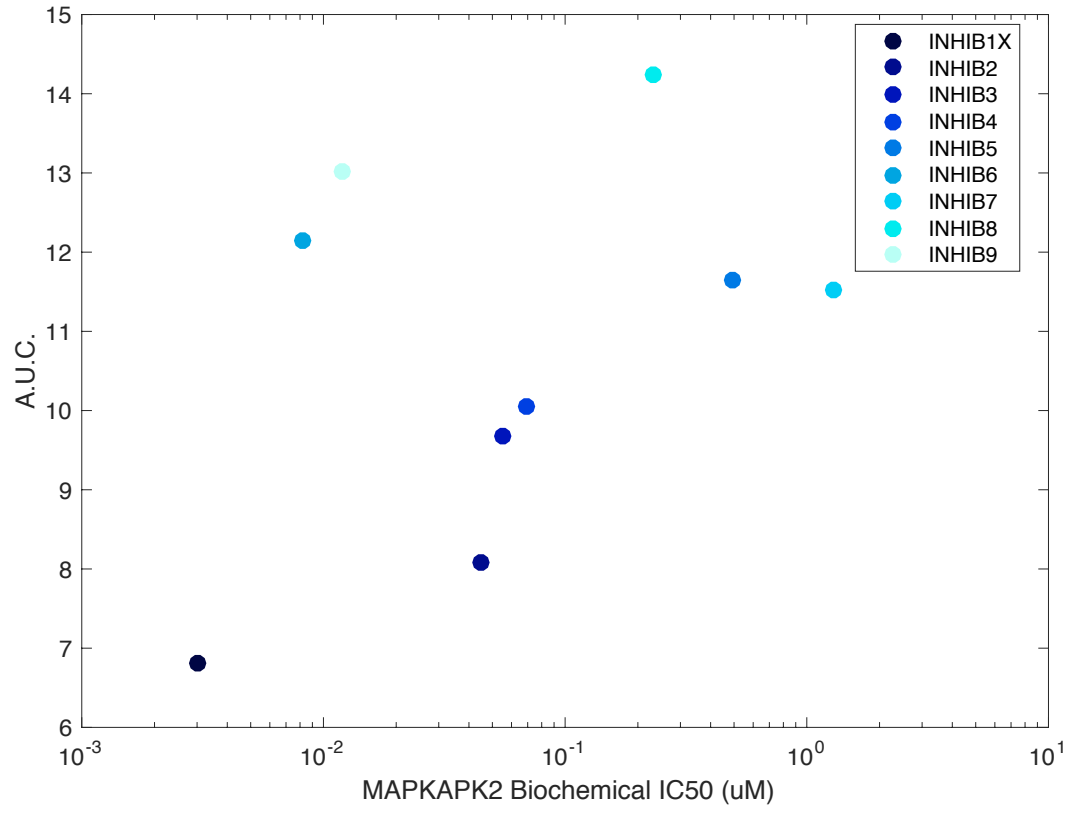
INHIB8



INHIB9

Compound	MAPKAPK5 IC50 (nM)	MAPKAPK2 IC50 (nM)
INHIB1X	< 3	< 3
INHIB2	3.8	45
INHIB3	62	55
INHIB4	5.6	69
INHIB5	30	490
INHIB6	5.3	8.2
INHIB7	33	1300
INHIB8	820	230
INHIB9	4.2	12

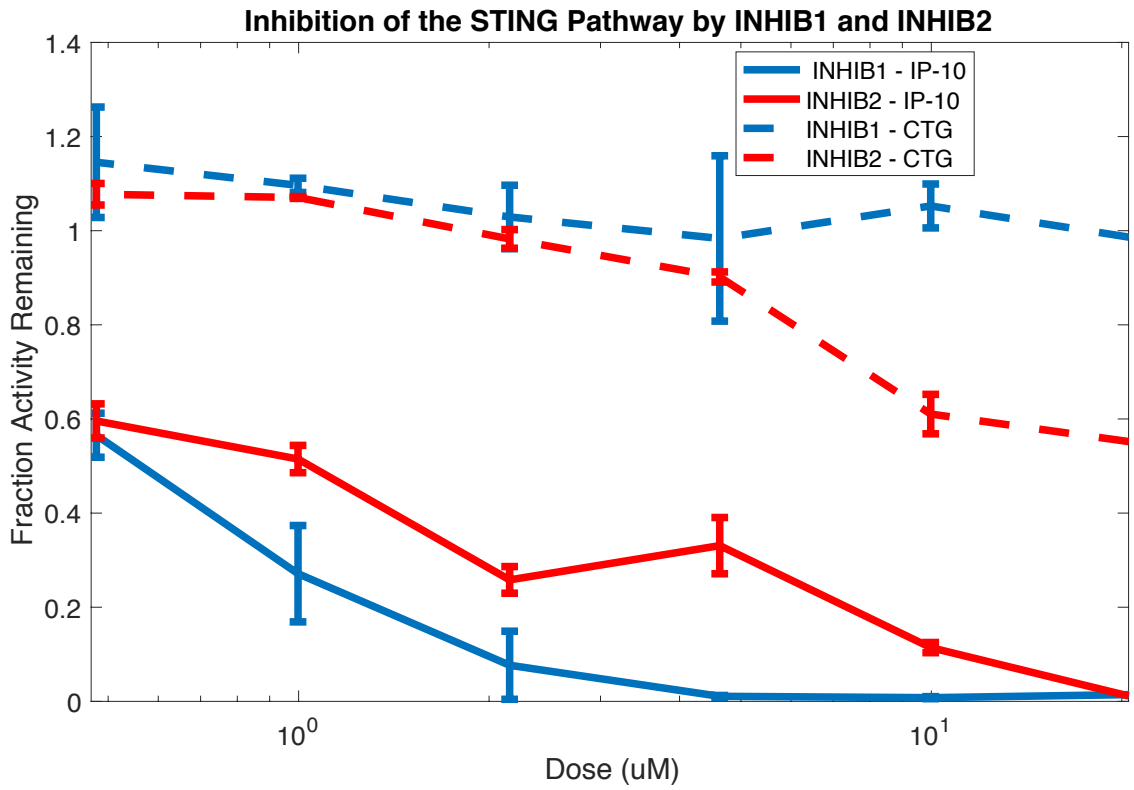
Supplement figure 8. Structure of 9 tetracycle compounds with IC50s for MAPKAPK2 and MAPKAPK5. The first tetracycline is called INHIB1X, so as not to confuse with the non-teracycle MAPKAPK2/5 inhibitor, INHIB1



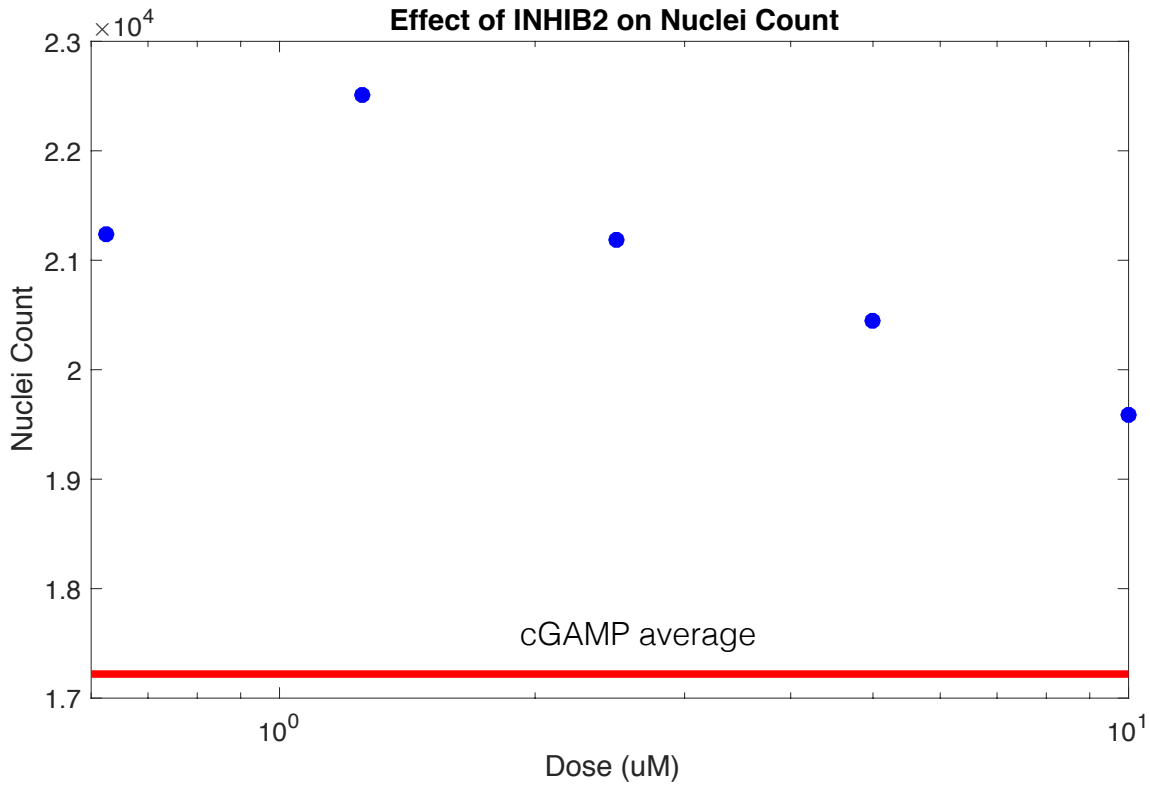
Spearman correlation = 0.28

Supplement figure 9. Activity of 9 tetracycle inhibitors plotted against MAPKAPK2 IC50s. There also is a weak correlation. The correlation is stronger for MAPKAPK5, but it is difficult to discriminate between the two genes with this data.

a

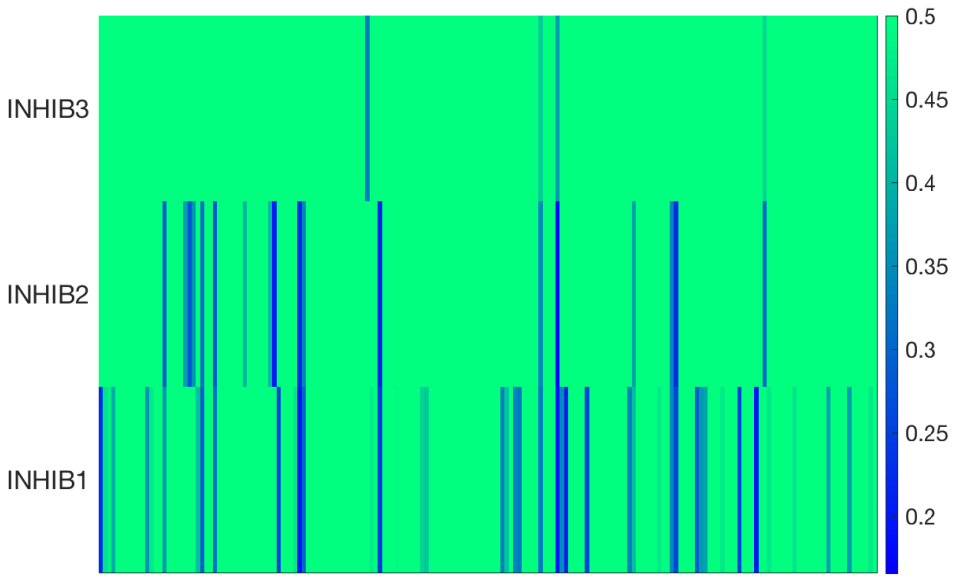


b

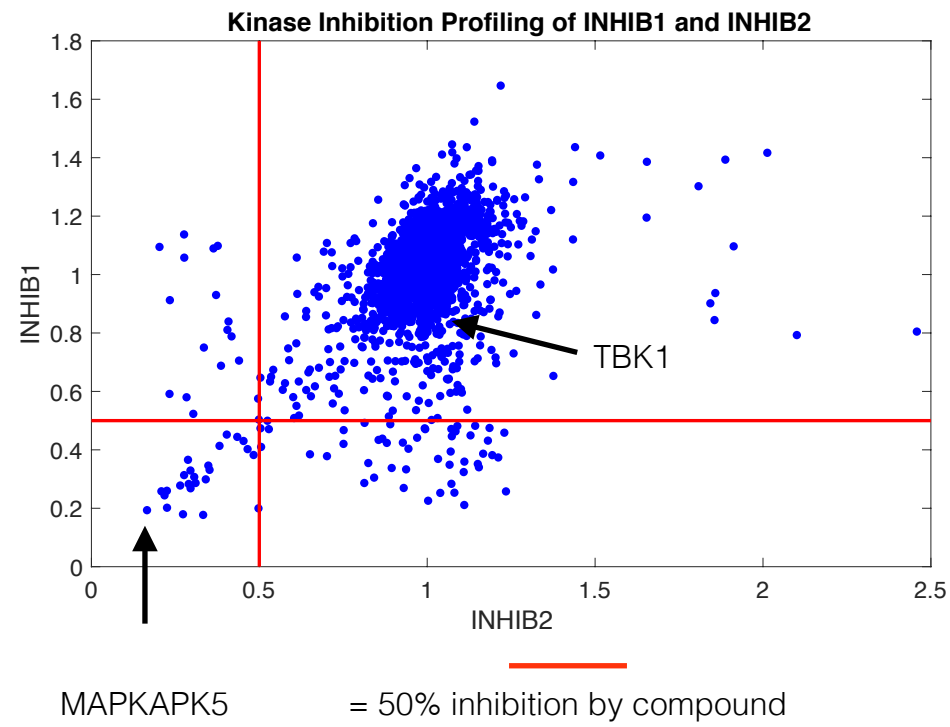


Supplement figure 10. Toxicity data for INHIB1 and INHIB2. a) IP-10 levels of primary macrophages treated with 2'3' cGAMP (20 uM) in combination with either INHIB1 and INHI2 were measured. Plotted is the fraction of IP-10 levels compared to 2'3' cGAMP-only controls (solid lines). Also plotted is the Cell-Titer-Glo (CTG) count to measure cellular viability. CTG counts were also normalized to 2'3' cGAMP-only controls. INHIB1 shows no toxicity as measured by CTG. INHIB2 shows some toxicity, but not enough to account for inhibition of the STING pathway. b) Nuclei count of primary macrophages treated with INHIB2 in combination with 2'3' cGAMP (experiment reported in fig. 6e -mini-SAR of 9 tetracycles). Nuclear count of INHIB2 treated wells is above cGAMP average. There is a reduction in nuclei count as dose of INHIB2 is increased, but by < 10%, not enough to account for pathway inhibition.

a



b



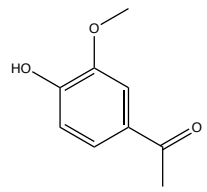
c

KINASE GENE	INHIB1 SCORE	INHIB2 SCORE	INHIB3 SCORE
'BRD2'	0.3659	0.2876	0.7066
'CAMKK2'	0.2866	0.3103	0.8989
'CDK11B'	0.3135	0.2760	0.9779
'CSNK2A1'	0.2019	0.2252	0.7190
'CSNK2A2'	0.2990	0.3405	0.7837
'GAK'	0.2442	0.2177	0.6907
'MAPK6'	0.3315	0.3518	0.4353
'MAPKAPK5'	0.1934	0.1656	0.3648
'PHKG2'	0.4136	0.3815	0.6578
'PRKD2'	0.3461	0.3482	0.7857
'PRKD3'	0.2599	0.2248	0.8458
'STK10'	0.1998	0.4975	0.9009

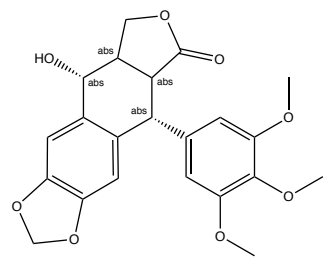
Supplement figure 11. KinoMatrix experiment for INHIB1, 2, and 3. a) The mass spectrometry identified 235 kinases in the pulldown, but only 184 had scores for all 3 inhibitors. Any kinase with a score > 0.5 (50% inhibition with 25 uM of compound) should be inactive in cell-based assays. Any kinase with a score < 0.5 is *potentially* active in cell-based assays. b) Plot of kinase inhibition scores for INHIB1 and INHIB2. 24 kinases are inhibited by both INHIB1 and INHIB2, shown in the left lower quadrant. MAPKAPK5 has the highest inhibition score for both INHIB1 and INHIB2 in the quadrant. c) List of inhibition scores for 12 kinases (it is debatable whether BRD2 has kinase activity), actively competed by INHIB1 and INHIB2, and also with scores for INHIB3. INHIB3, the weak analog, inhibits some kinases, including MAPKAPK5, but to much a weaker degree than INHIB1 and INHIB2. In this particular experiment, a kinase cannot be ruled out as a target if a drug inhibits it biochemically, but is weak in cell-based assays, as a number of reasons, such as cell-permeability, dosing, etc., can explain the effect. Thus, this data cannot rule out MAPKAPK5. Importantly, TBK1 is not inhibited by any of the small molecules.

Activator Compounds for Follow Up

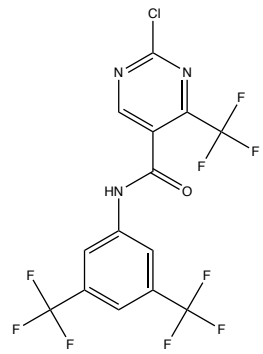
1



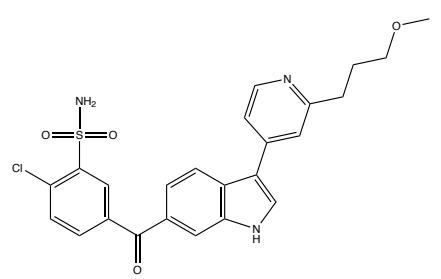
2



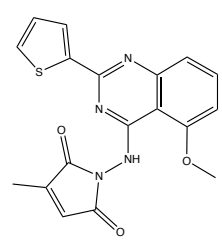
3



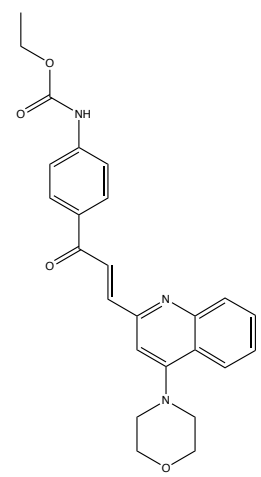
4



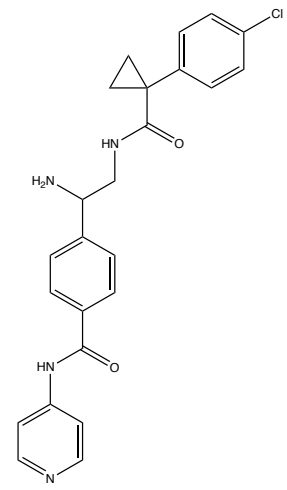
5



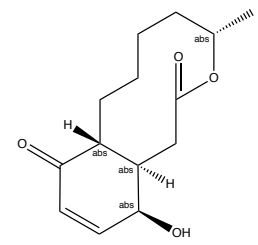
6



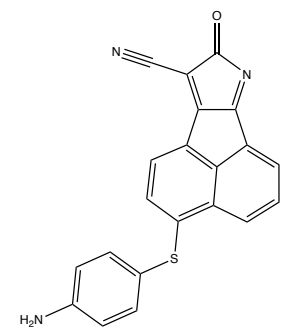
7



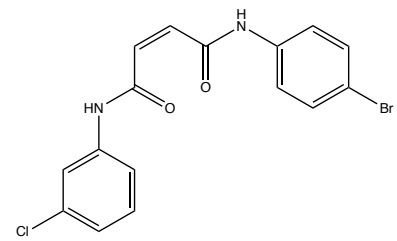
8



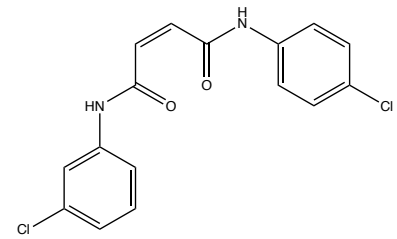
9



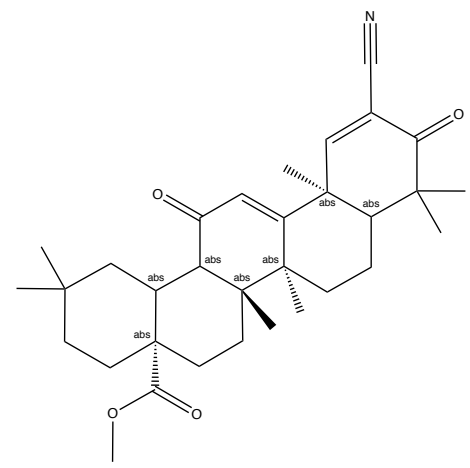
10



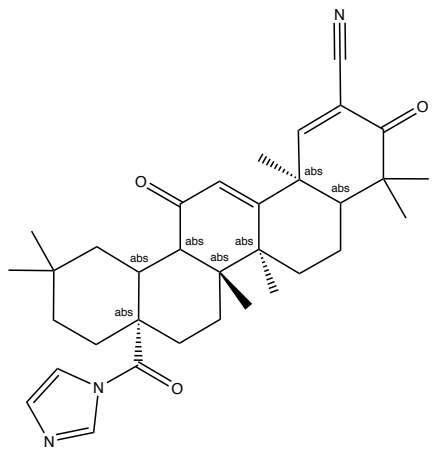
11



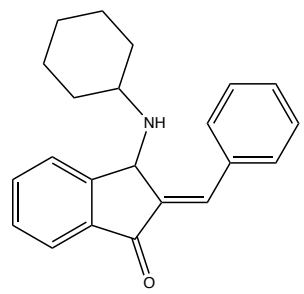
12



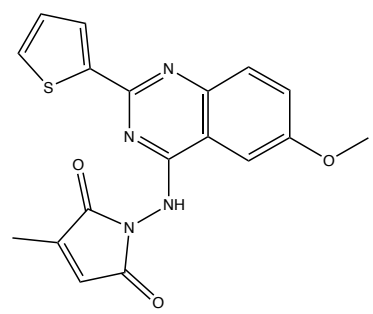
13



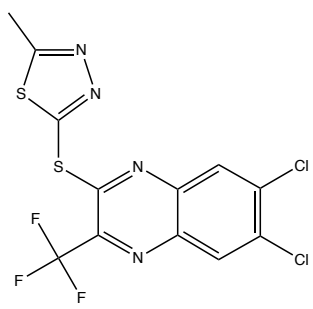
14



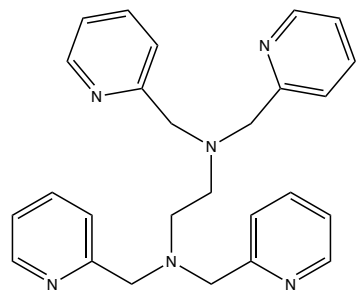
15



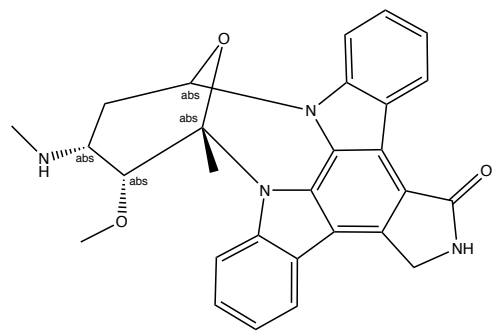
16



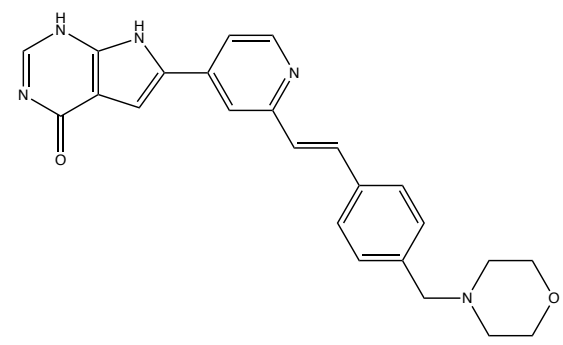
17



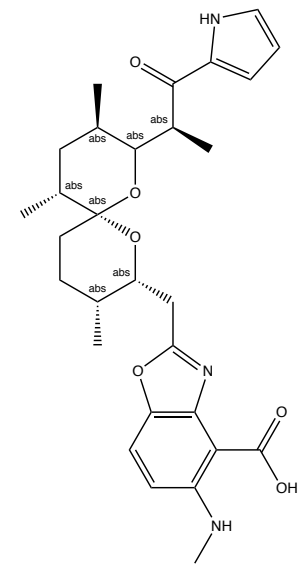
18



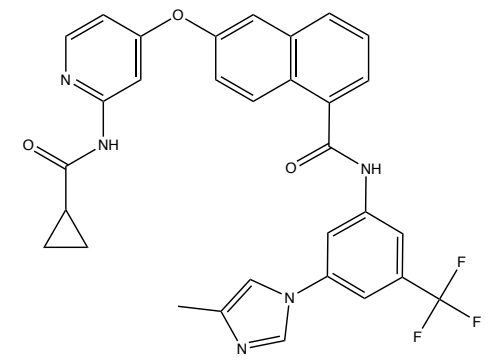
19



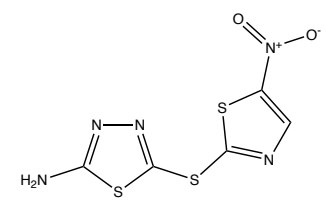
20



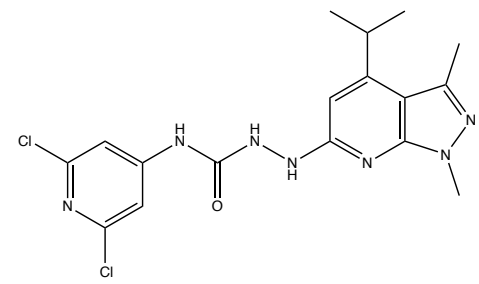
21



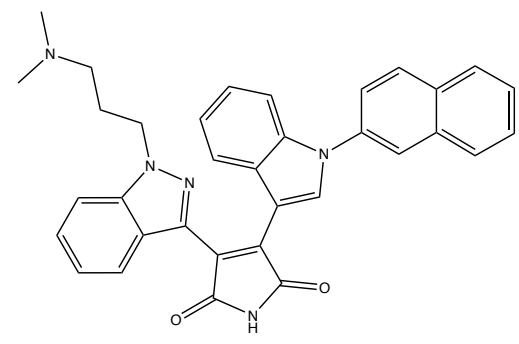
22

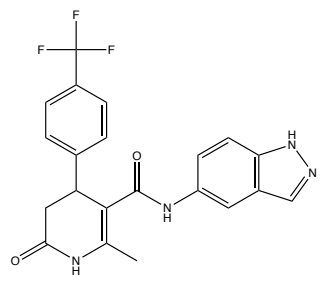


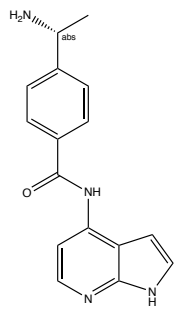
23

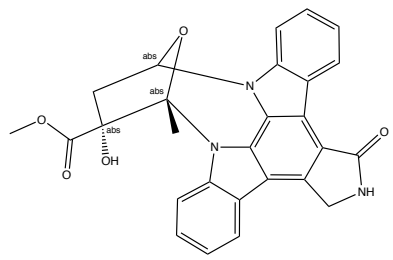


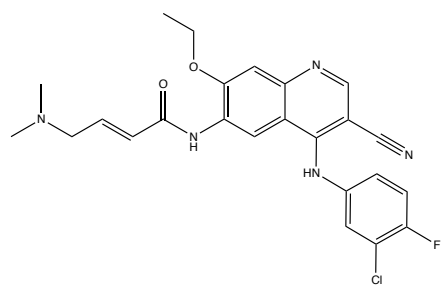
24

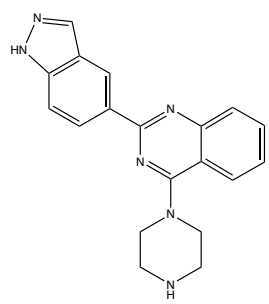


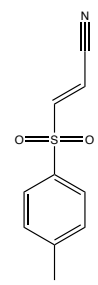
25 

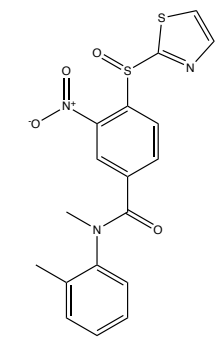
26 

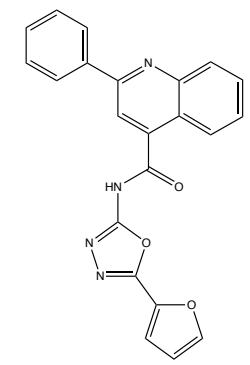
27 

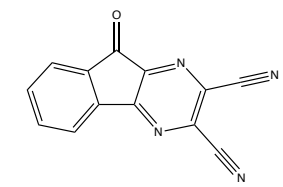
28 

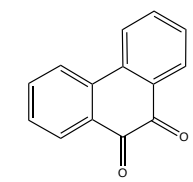
29 

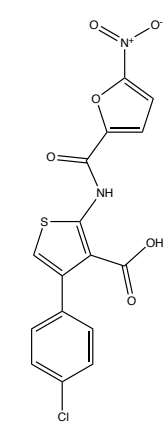
30 

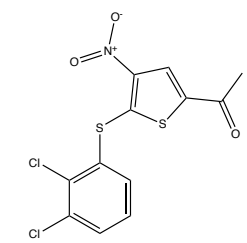
31 

32 

33 

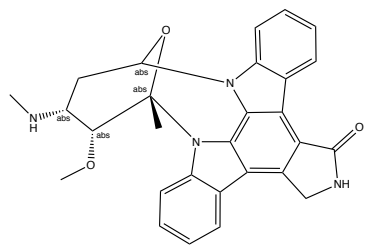
34 

35 

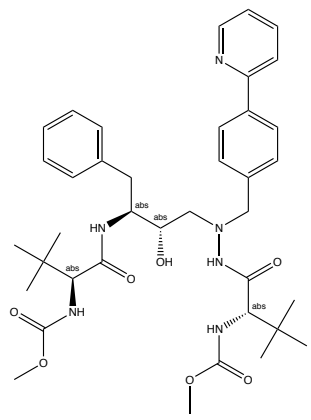
36 

Inhibitor Compounds for Follow Up

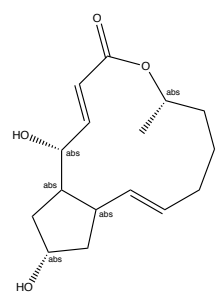
1



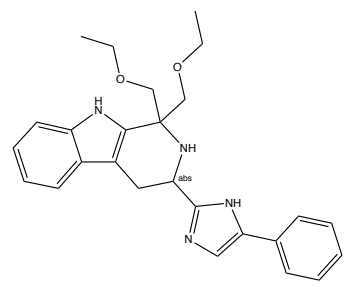
2



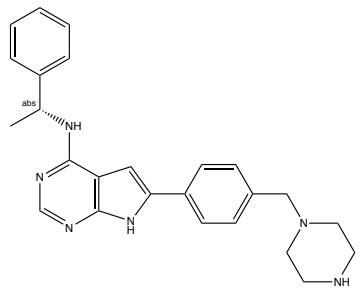
3



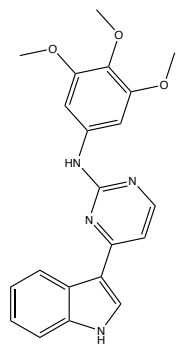
4



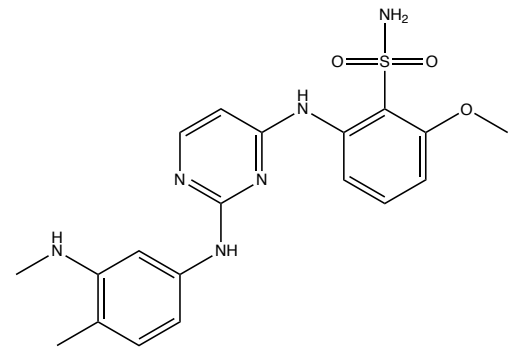
5



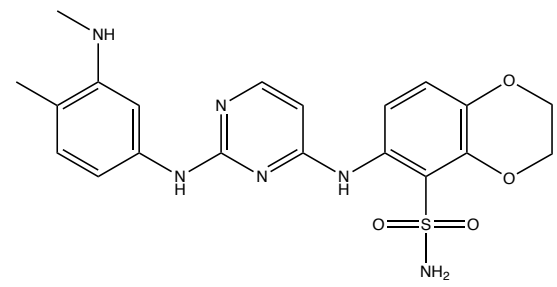
6



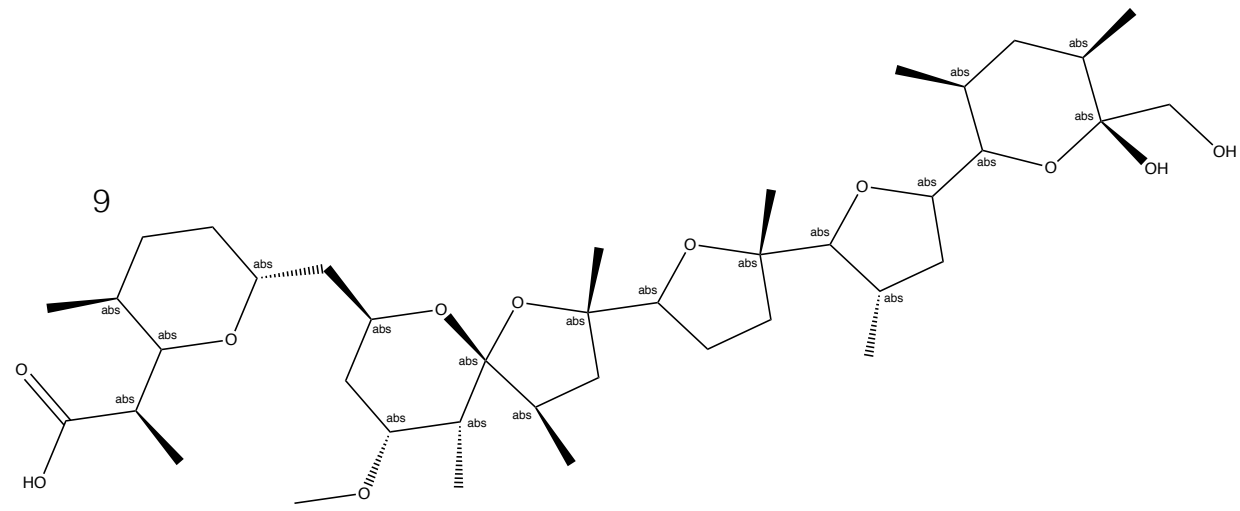
7



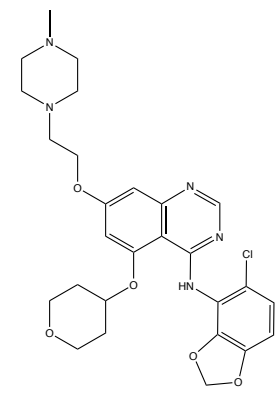
8



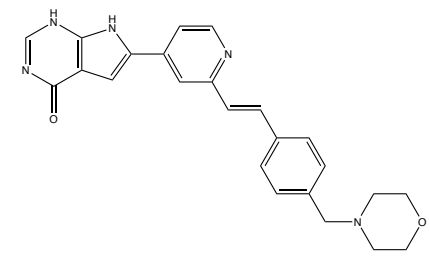
9



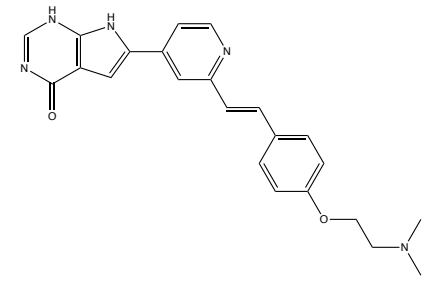
10



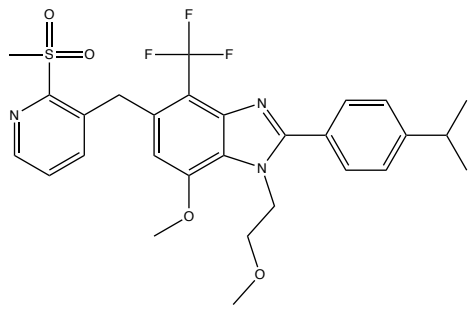
11



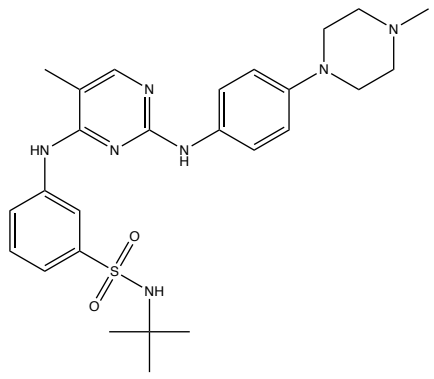
12



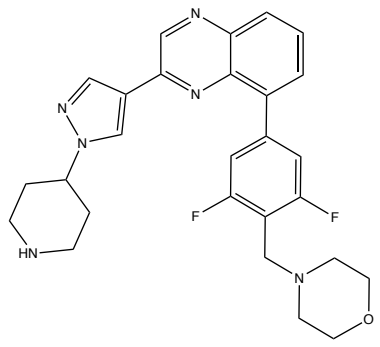
13



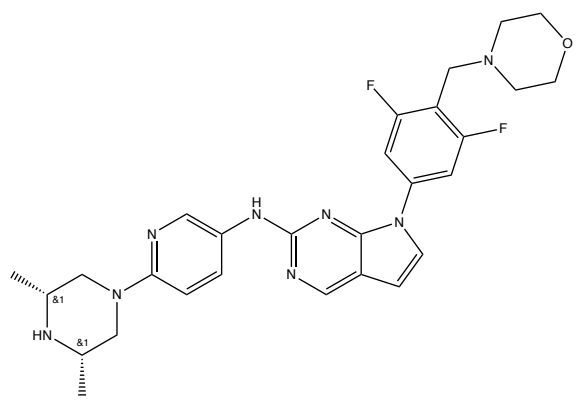
14



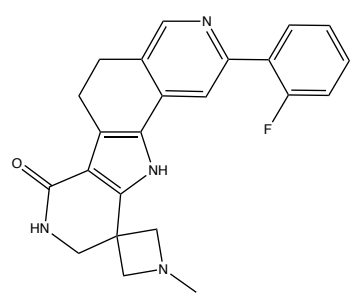
15



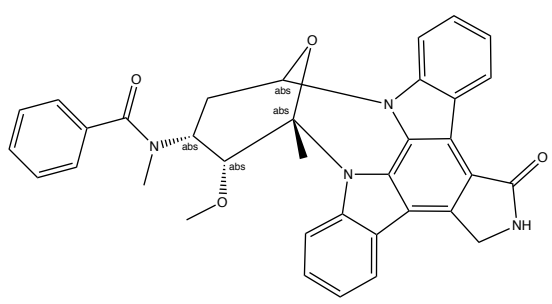
16



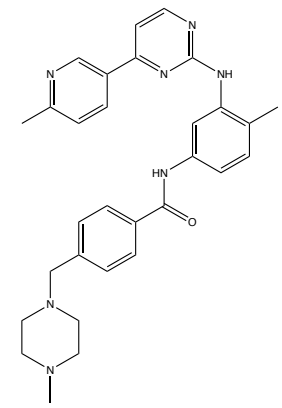
17



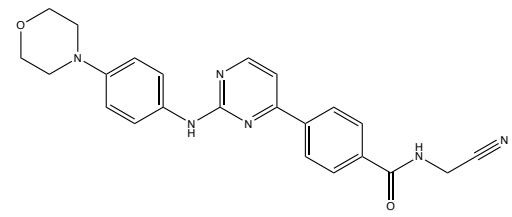
18



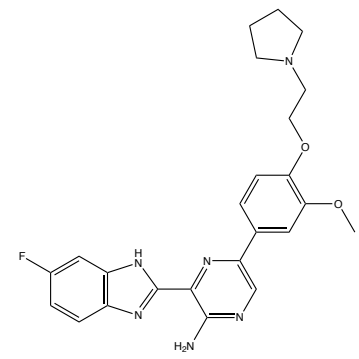
19



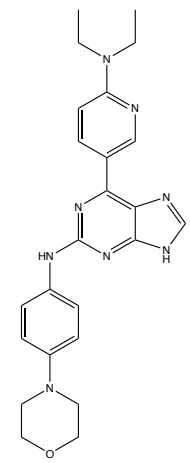
20



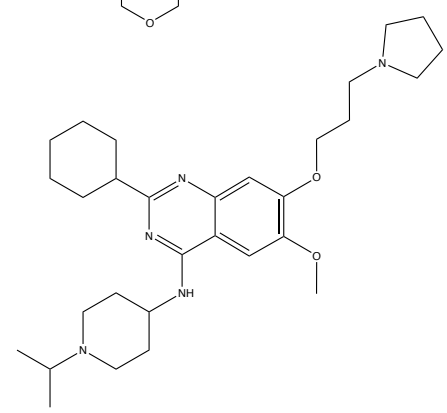
21



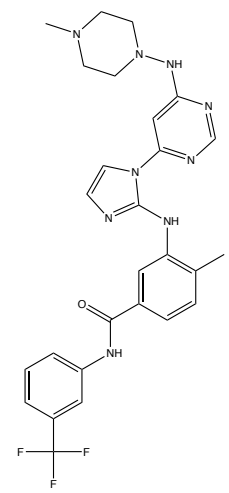
22



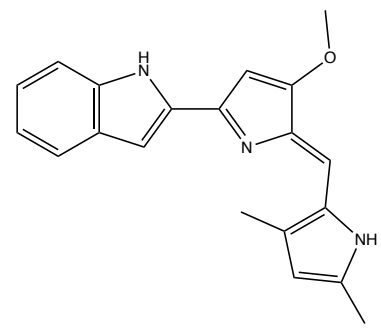
23



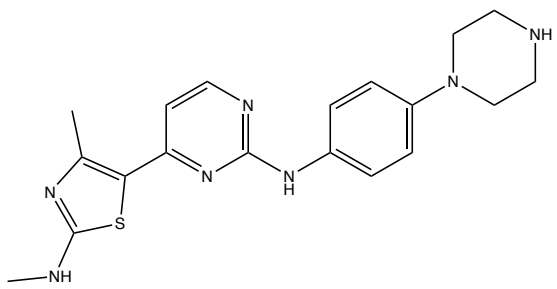
24



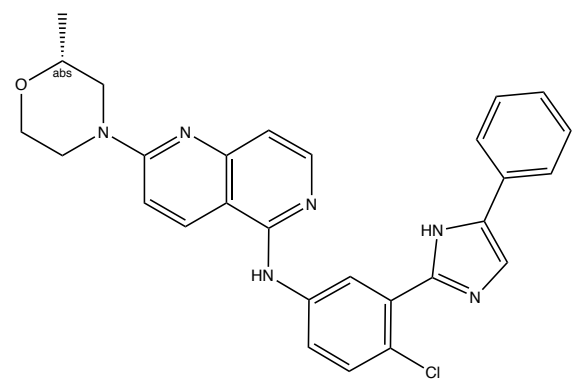
25



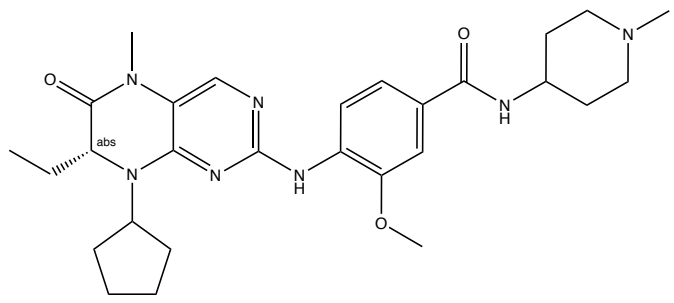
26



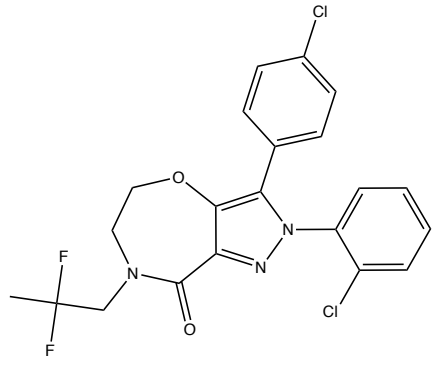
27



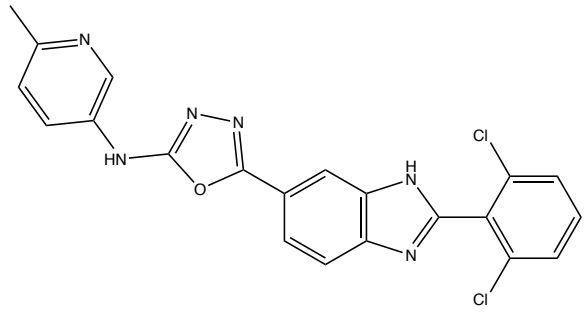
28



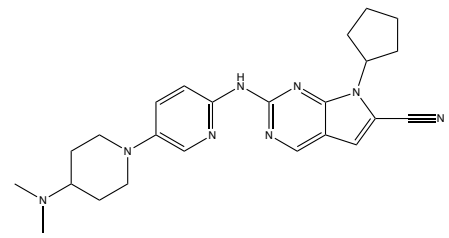
29



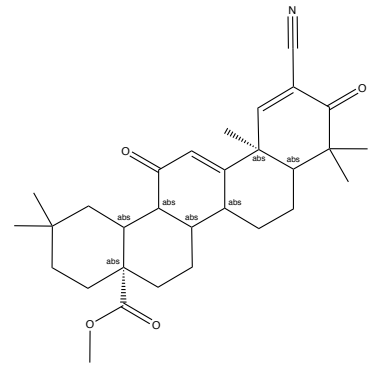
30



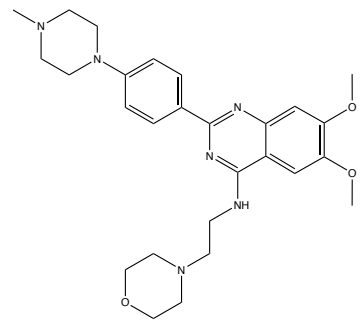
31



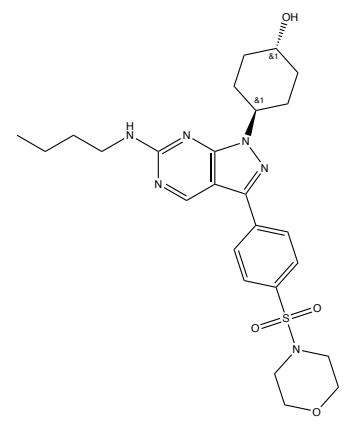
32



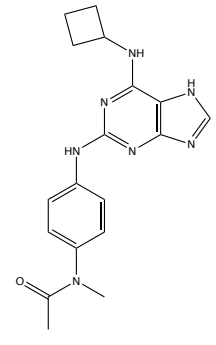
33



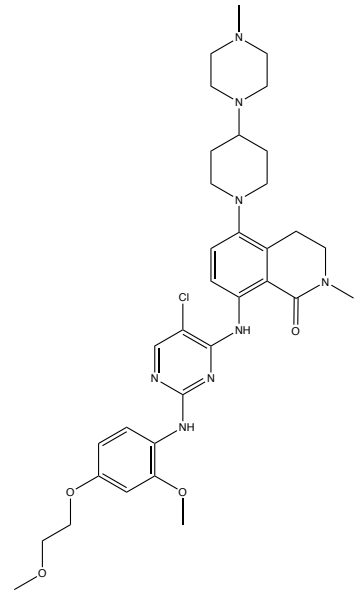
34



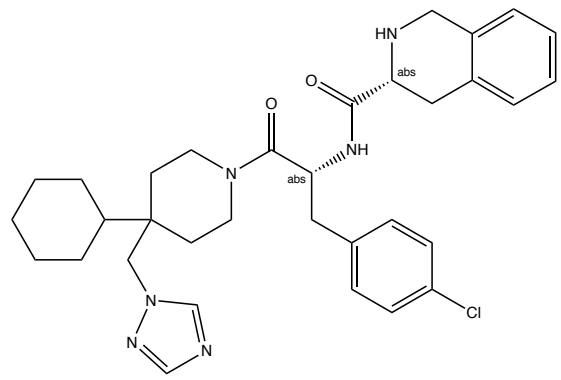
35



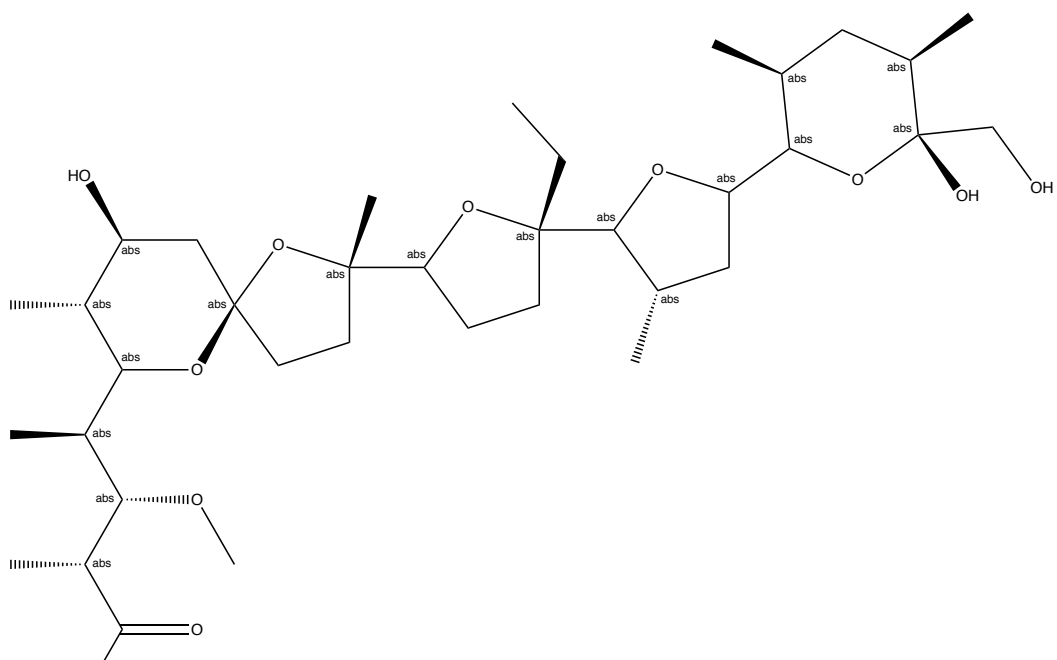
36



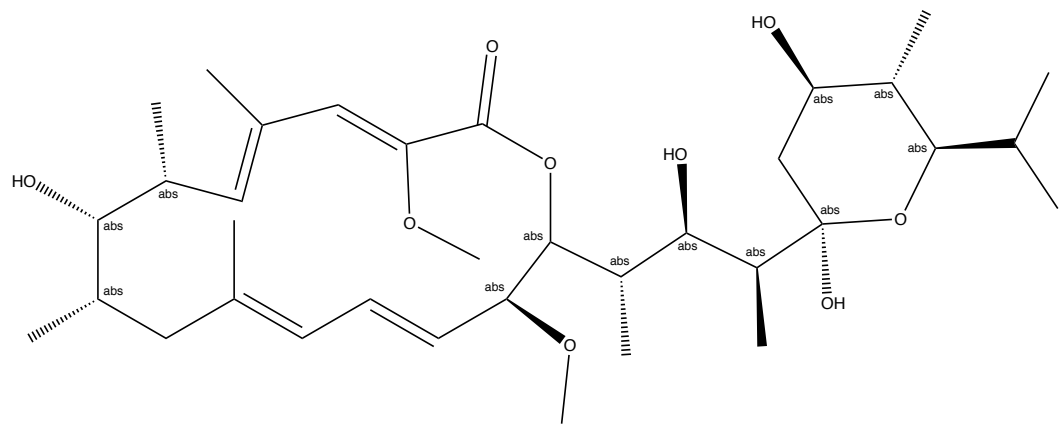
37


HN
abs
O
O
NH
N
abs
N
N
N
Cl

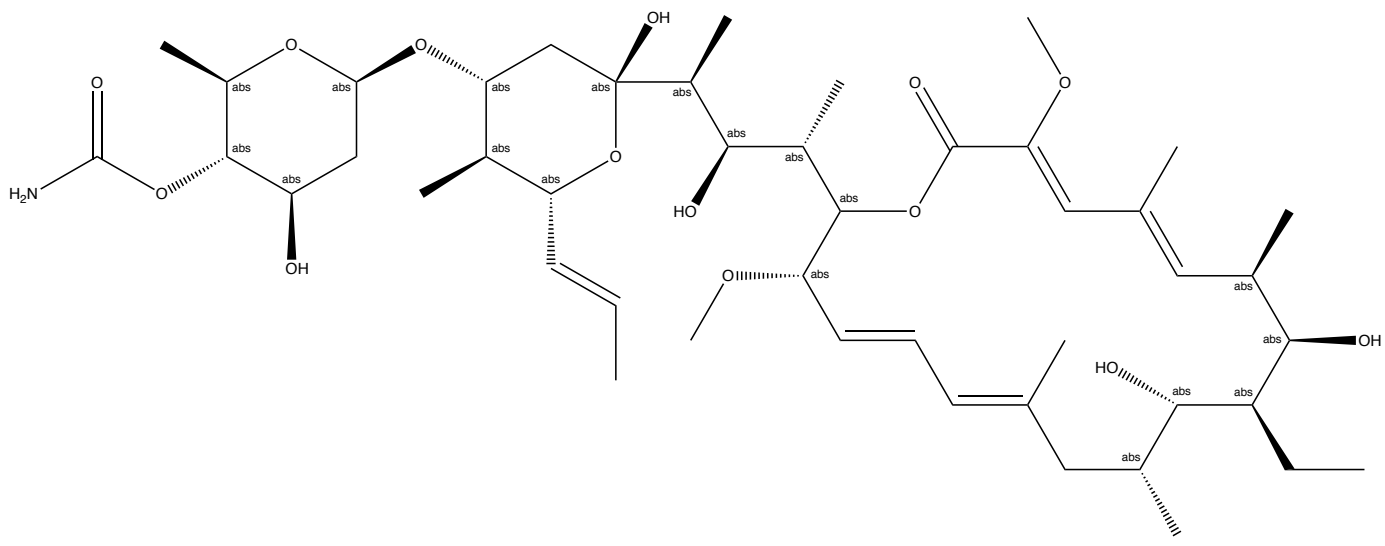
38


HO
abs
O
O
abs
O
abs
O
abs
O
abs
O
abs
OH
OH
HO

39


HO
O
O
HO
abs
O
abs
OH
O

40


O
O
O
abs
OH
HO
O
O
O
OH
HO
OH