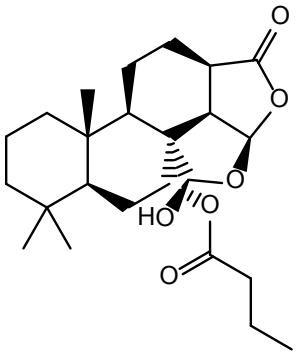
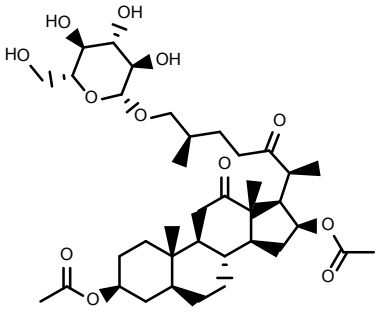
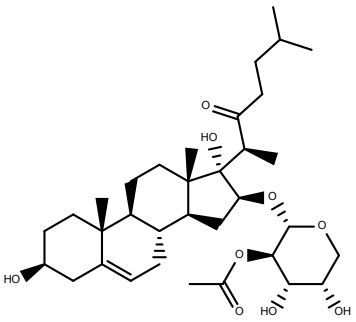
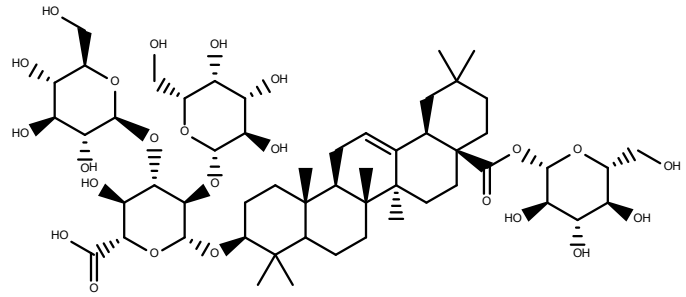
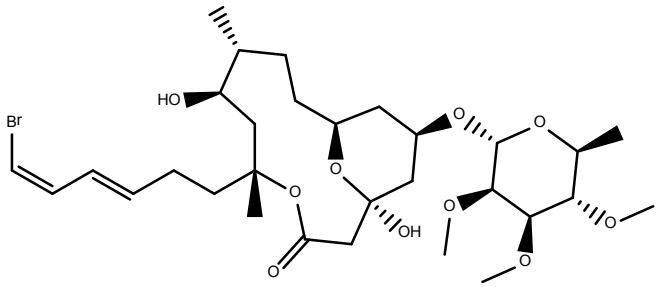
Supporting Information. Electronic Supplementary Information (ESI) available: Those additional tables and figures indicated within the text are supported.

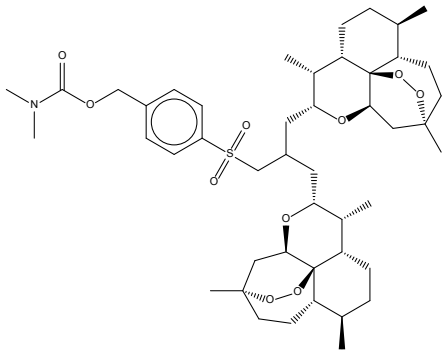
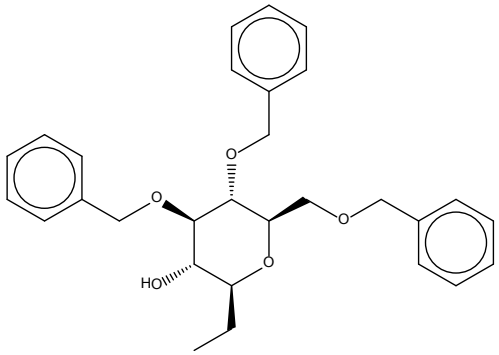
Table S1. 18 Most represented targets (2.01 % of the training dataset) within the output of the Naïve Bayesian Classifier run over the dataset composed of 1066 cytotoxic compounds to HeLa cells. The percentage of compounds for which these targets appeared among the top fifteen ranked predicted targets is depicted in the third column.

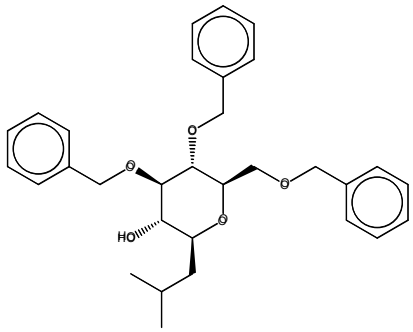
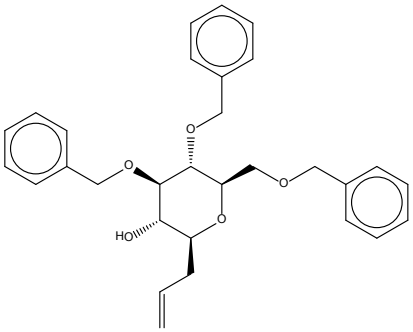
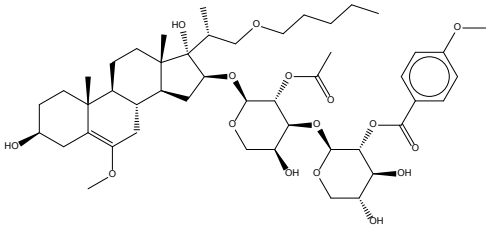
CHEMBL ID	Target	Target frequency
CHEMBL299	Protein kinase C alpha	17,45
CHEMBL4302	P-glycoprotein 1	13,41
CHEMBL3775	Dual specificity phosphatase Cdc25	13,23
CHEMBL5399	Pituitary adenylate cyclase-activating polypeptide type I receptor	11,73
CHEMBL5378	P-selectin	11,54
CHEMBL4894	Galanin receptor	11,26
CHEMBL4069	Corticotropin releasing factor receptor	11,26
CHEMBL3522	Cytochrome P450 17A1	10,98
CHEMBL1856	Steroid 5-alpha-reductase 2	9,94
CHEMBL1978	Cytochrome P450 19A1	9,66
CHEMBL4625	Apoptosis regulator Bcl-X	9,38
CHEMBL5409	G-protein coupled bile acid receptor 1	8,44
CHEMBL253	Cannabinoid CB2 receptor	7,97
CHEMBL2899	Brain adenylate cyclase 1	7,60
CHEMBL3746	11-beta-hydroxysteroid dehydrogenase 2	7,60
CHEMBL3142	DNA-dependent protein kinase	6,75
CHEMBL1781	DNA topoisomerase I	6,66
CHEMBL1804	Somatostatine Receptor 2	6,47

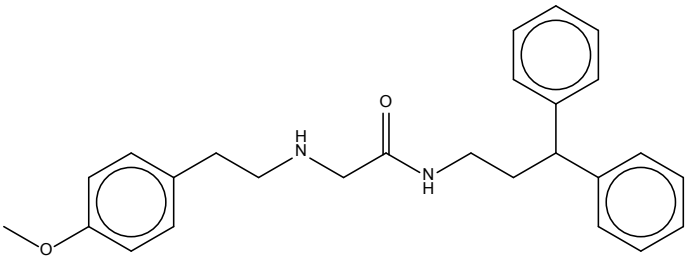
Table S2. 5 nearest neighbors from the cytotoxicity dataset for compounds selected for their putative inhibitory activity against Topoisomerase and P-gp in terms of the Tanimoto score (Tc. Openbabel fingerprint FP2). On this basis, an overall dissimilarity between the cytotoxicity dataset and the selected compounds is demonstrated, assuring the novelty of the chemical scaffolds herein tested.

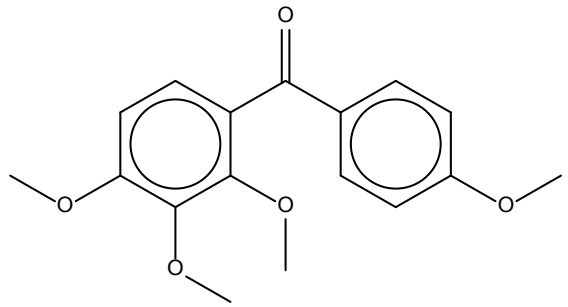
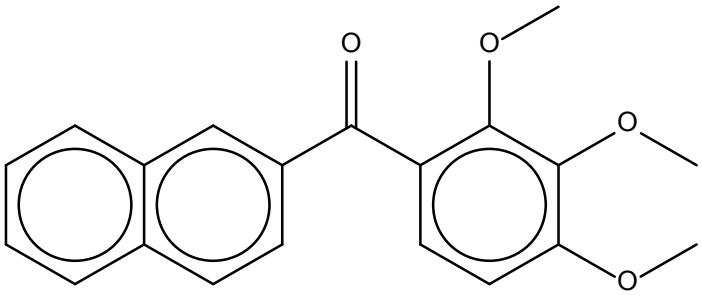
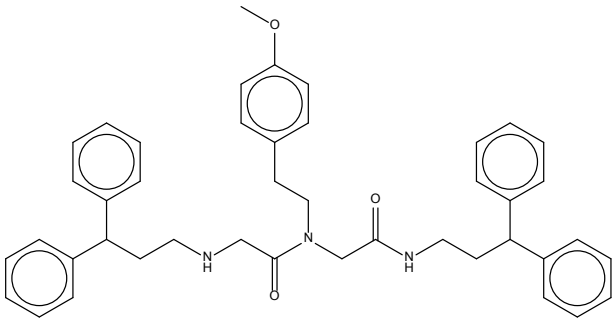
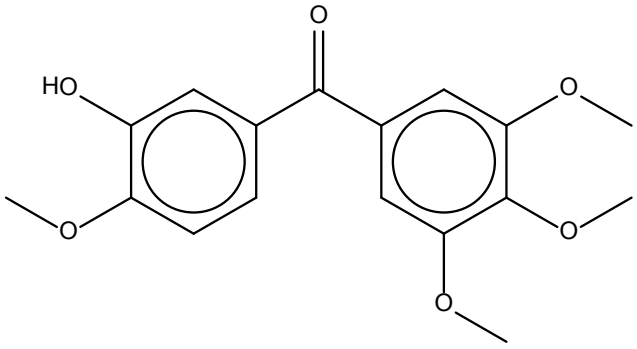
5 nearest neighbours for compound P-gp 1	Tanimoto coefficient (Tc)
	0.61
	0.57
	0.51

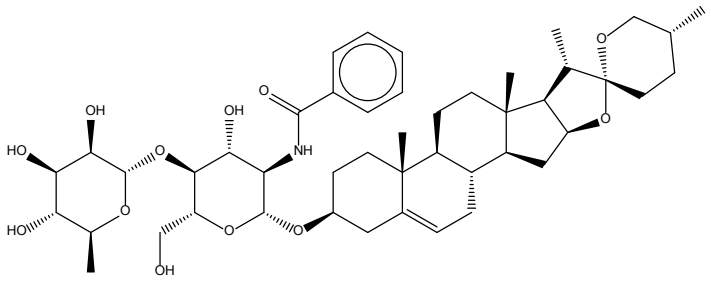
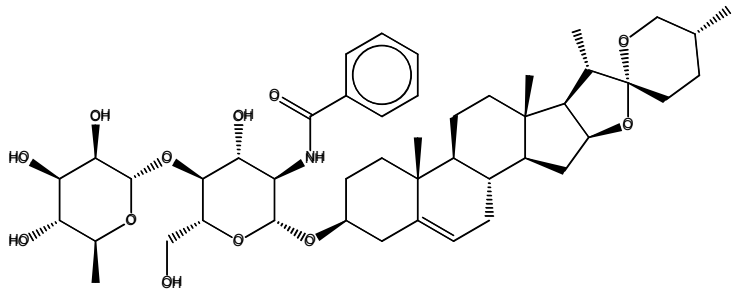
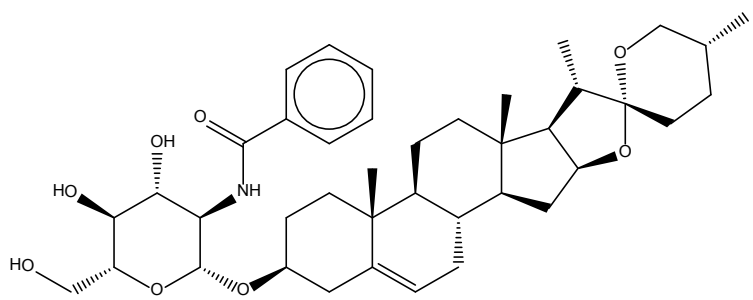
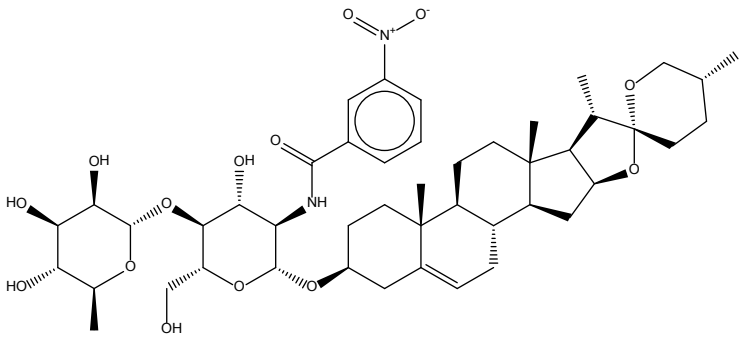
	0.50
	0.48

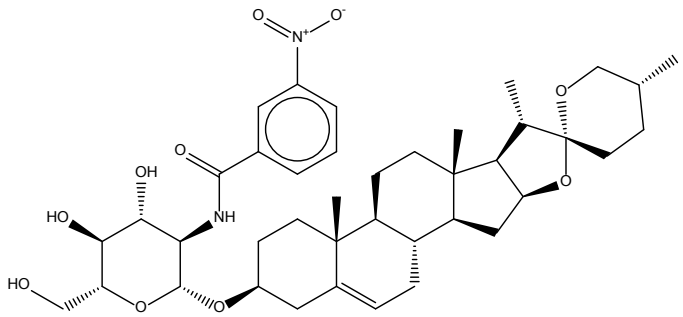
5 nearest neighbours for compound P-gp 2	Tanimoto coefficient (Tc)
	0.41
	0.32

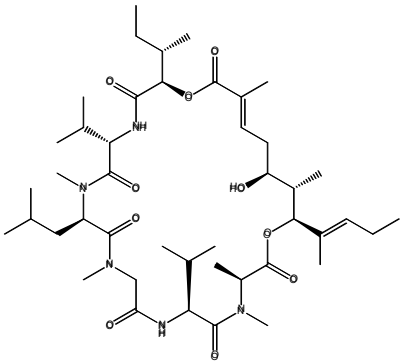
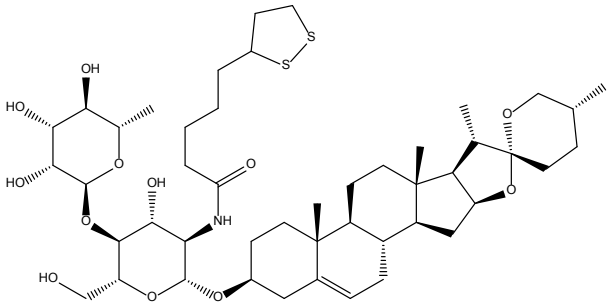
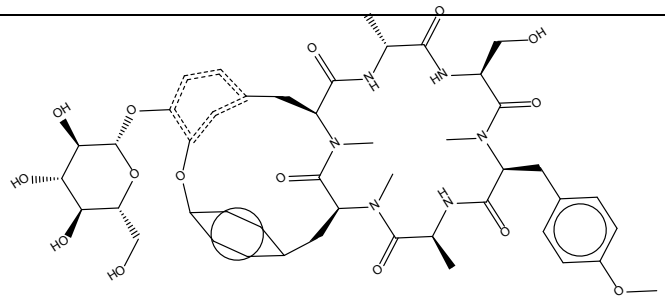
	0.32
	0.30
	0.29

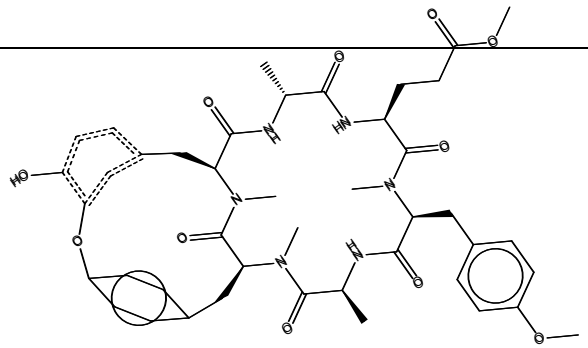
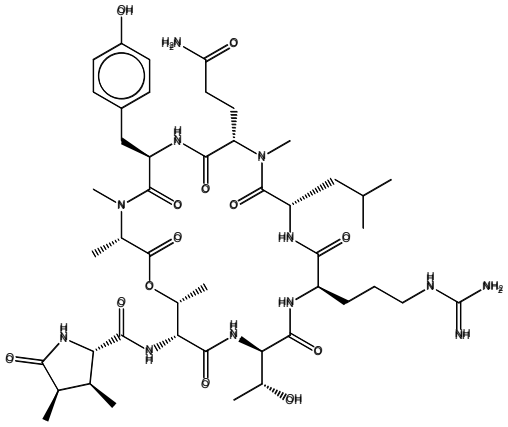
5 nearest neighbours for compound P-gp 3	Tanimoto coefficient (Tc)
	0.46

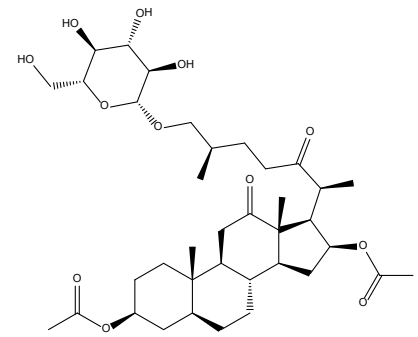
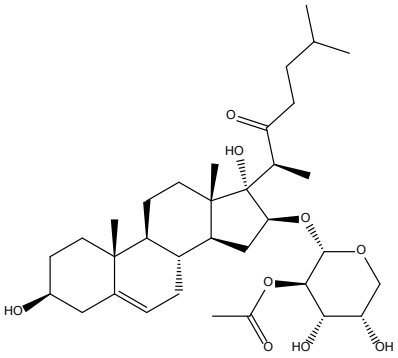
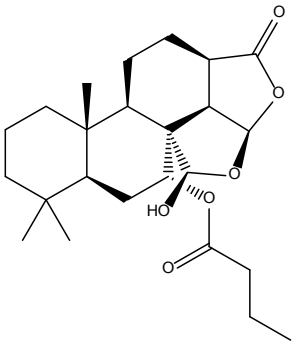
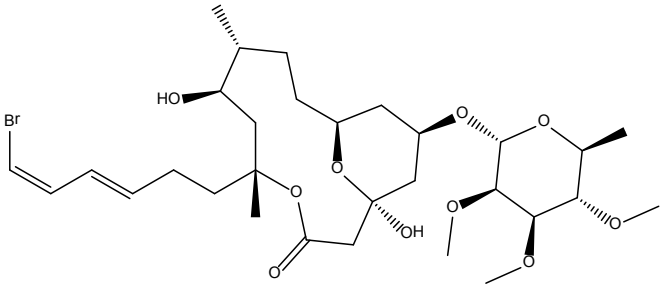
	0.46
	0.45
	0.43
	0.43

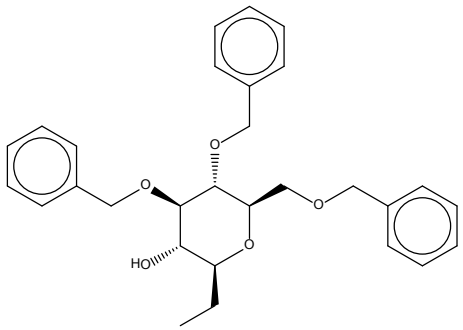
5 nearest neighbours for compound P-gp 4	Tanimoto coefficient (Tc)
	0.58
	0.58
	0.57
	0.50

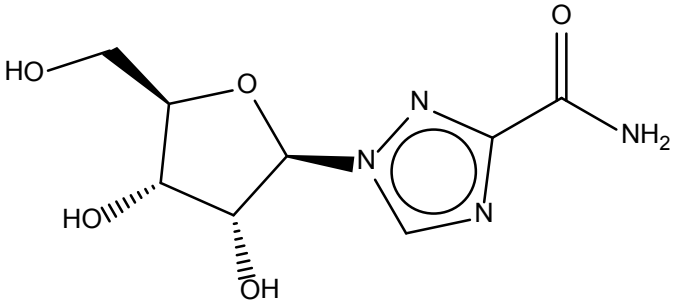
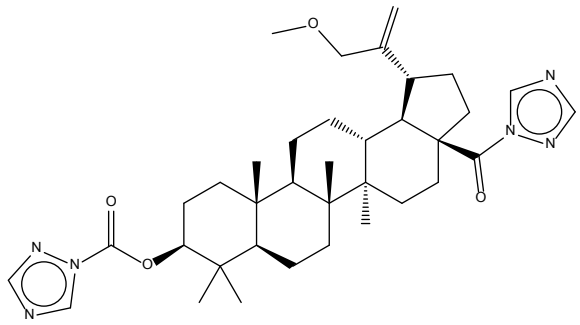
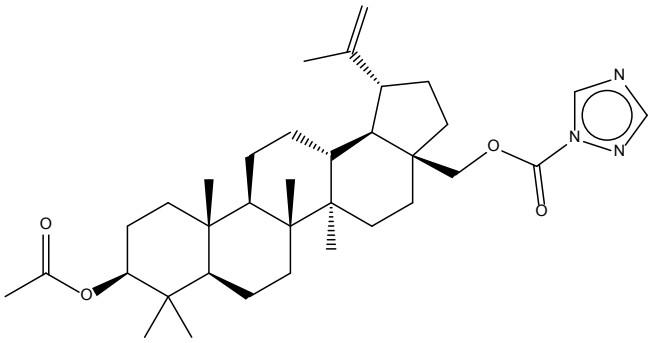
	0.48

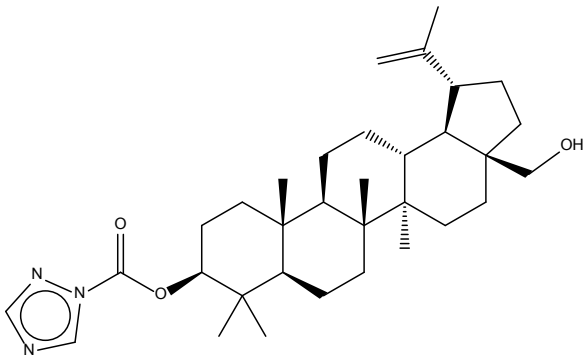
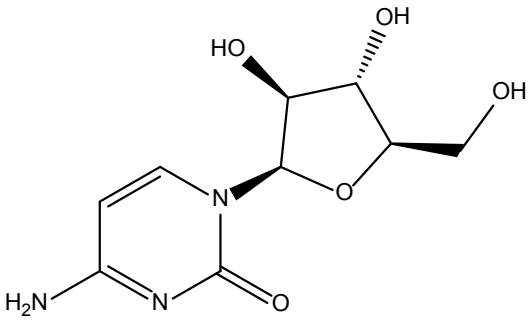
5 nearest neighbours for compound P-gp 5	Tanimoto coefficient (Tc)
	0.34
	0.34
	0.34

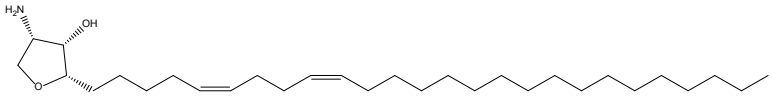
	0.33
	0.33

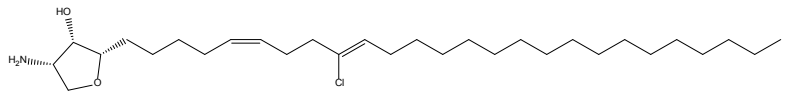
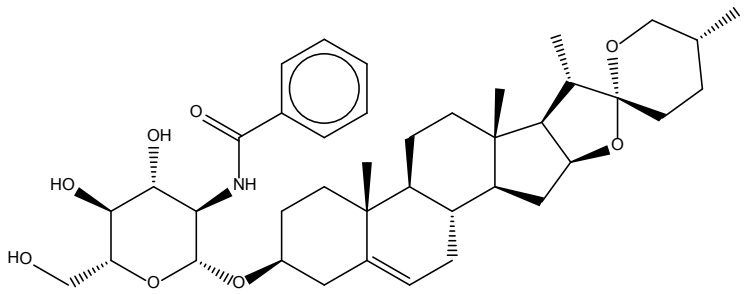
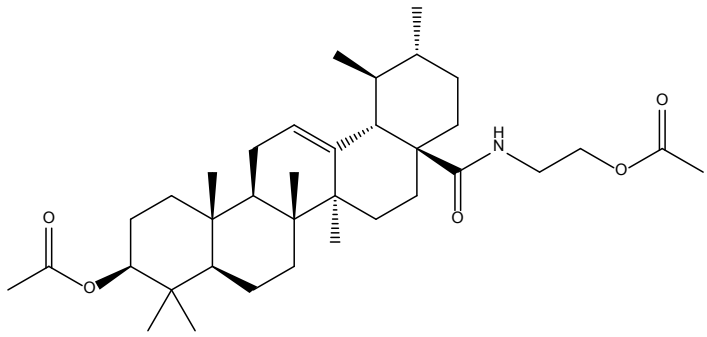
5 nearest neighbours for compound Topo 1	Tanimoto coefficient (Tc)
	0.68
	0.49
	0.49
	0.48

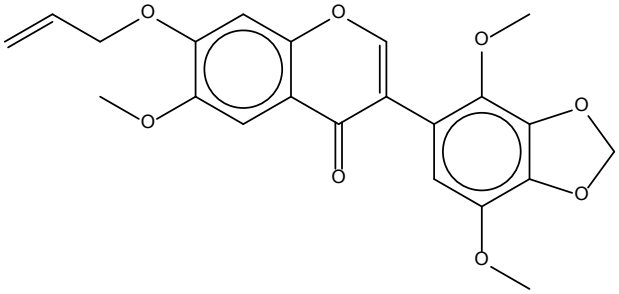
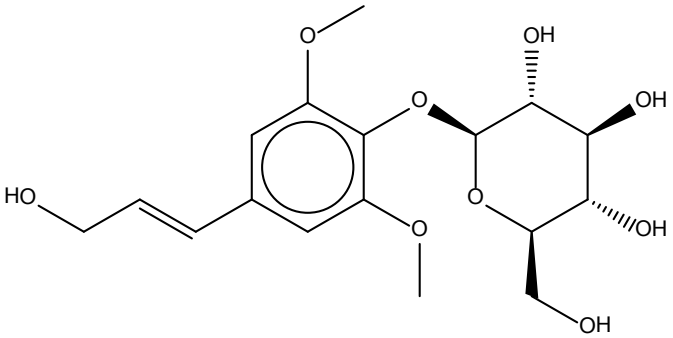
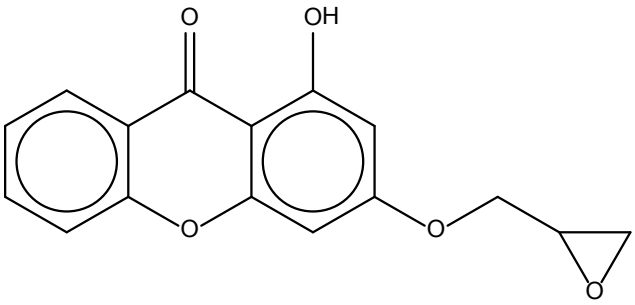
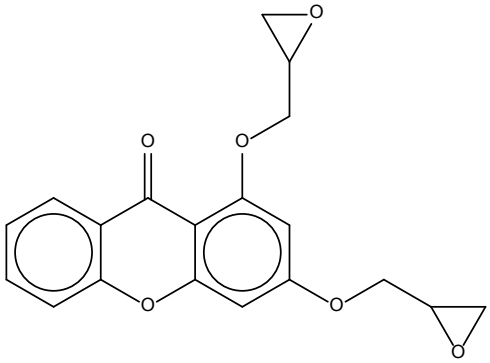
	0.46

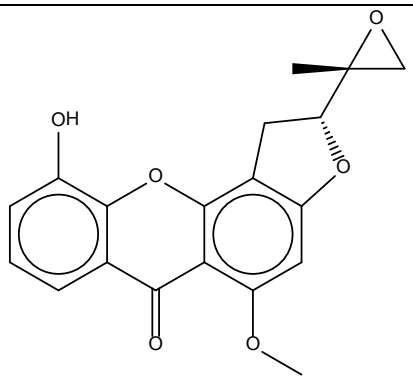
5 nearest neighbours for compound Topo 2	Tanimoto coefficient (Tc)
	1
	0.44
	0.43

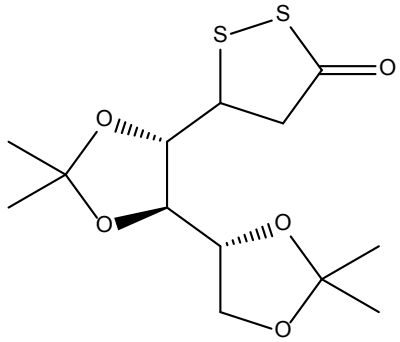
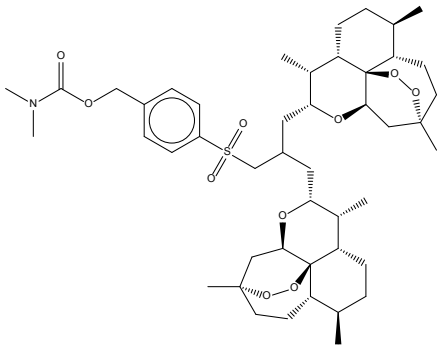
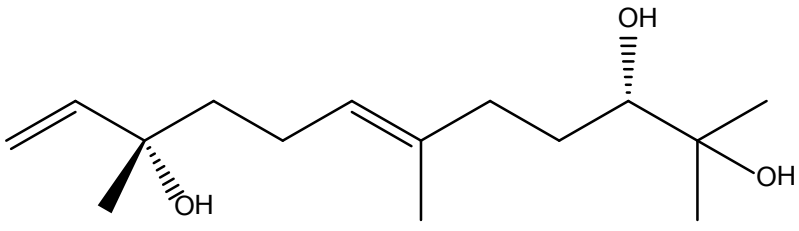
	0.41
	0.40

5 nearest neighbours for compound Topo 3	Tanimoto coefficient (Tc)
	0.26
	0.25

	0.24
	0.24
	0.23

5 nearest neighbours for compound Topo 4	Tanimoto coefficient (Tc)
	0.57
	0.49
	0.45
	0.45

	0.45

5 nearest neighbours for compound Topo 5	Tanimoto coefficient (Tc)
	0.21
	0.21
	0.19

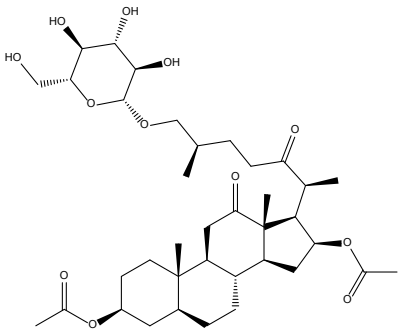
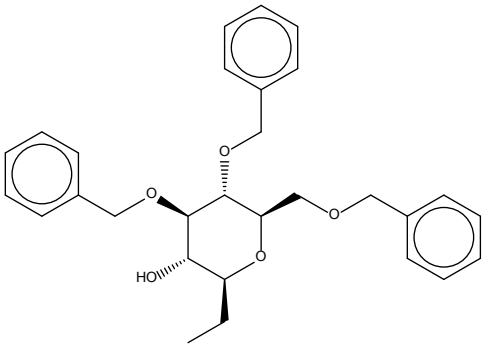
	0.19
	0.19

Table S3. Statistics analysis of compounds´ Topo 1-5 and P-gp 2, 4 synergism in HeLa cells. Normality (depicted as sample A / sample B in the table) and homoscedasticity of the normalized absorbance values of the MTT (cytotoxicity) assay for i) compounds Topo 1-5 at 120 μM and ii) compounds Topo 1-5 at 120 μM in the presence of either P-gp 2 (first row) or P-gp 4 (second row) at 50 μM was evaluated. A two-tailed t-test of independent samples was applied (α=0.05).

	Topo 1			
	Normality	Sig. (Levene)	t	P-value
P-gp 2	0,293 / 0.709	0.063	0.632	**0.56155**
P-gp 4	0.624 / 0.293	0.047	0.980	**0.38262**

	Topo 2			
	Normality	Sig. (Levene)	t	P-value
P-gp 2	0.211 / 0.376	0.459	9.312	**0.00074**
P-gp 4	0.396 / 0.211	0.914	10.817	**0.00041**

	Topo 3			
	Normality	Sig. (Levene)	t	P-value
P-gp 2	0.173 / 0.176	0.964	23.057	**0.00002**
P-gp 4	0.186 / 0.176	0.079	11.371	**0.00034**

	Topo 4			
	Normality	Sig. (Levene)	t	P-value
P-gp 2	0.359 / 0.279	0.083	-1.238	**0.28343**
P-gp 4	0.279 / 0.556	0.370	-4.574	**0.01023**

	Topo 5			
	Normality	Sig. (Levene)	t	P-value
P-gp 2	0.490 / 0.169	0.390	10.160	**0.00053**
P-gp 4	0.279 / 0.166	0.864	0.082	**0.00079**

Table S4. PubChem BioAssay IDs of the 186 assays from which the cytotoxicity dataset (1066 compounds) was derived. Only those assays describing compounds which had proved cytotoxic to HeLa cells in cell-based assays were selected.

294847	381933	398539	464559	479358	88185
294834	88679	377626	450675	471495	88184
294831	536120	361454	449889	434865	88181
270163	346969	355718	430990	353125	88073
81036	346800	344236	426954	87614	88072
81035	320480	335543	426417	87608	88070
81031	365322	329726	404114	407951	88067
81029	402790	327127	401978	380586	86512
303724	402787	276856	387757	379896	86509
303703	356397	248913	381783	379439	88666
274814	392419	88689	381781	376360	538234
274810	353190	88677	381780	343443	528360
569084	311630	536124	381766	333399	528359
569083	310228	349477	355864	329152	528358
564804	307643	320481	338960	329150	528357
564575	322407	87636	284030	329148	81461
427932	88535	569123	264148	329146	467906
488585	568658	540172	260418	375386	467902
474627	549982	539013	260417	297473	447524
464541	538235	426719	247881	291746	437448
458901	422519	423636	244765	259369	437447
449887	346248	370457	87597	259358	437433
399054	88537	504287	88526	248527	436447
387749	502767	502353	365321	247998	378183
385102	486643	492147	502427	247772	344241
385097	471521	488530	88526	209179	332255
381790	455467	484874	383191	105086	332254
381775	447340	482938	383189	95233	332252
381770	437362	474646	307644	88188	329730
357113	430358	474643	249265	88187	329727
294250	410367	469272	87598	88186	305460

Figure S1. Heatmap representation of the similarity analysis (Tanimoto coefficient, OpenBabel FP2) against known Topoisomerase inhibitors from CHEMBL version 10 (448 compounds). Each row of the figure corresponds to one of the 448 known inhibitors. The majority of compounds bear Tanimoto coefficients lower than 0.4, guaranteeing therefore the novelty of the Topoisomerase I inhibitors tested in this work.

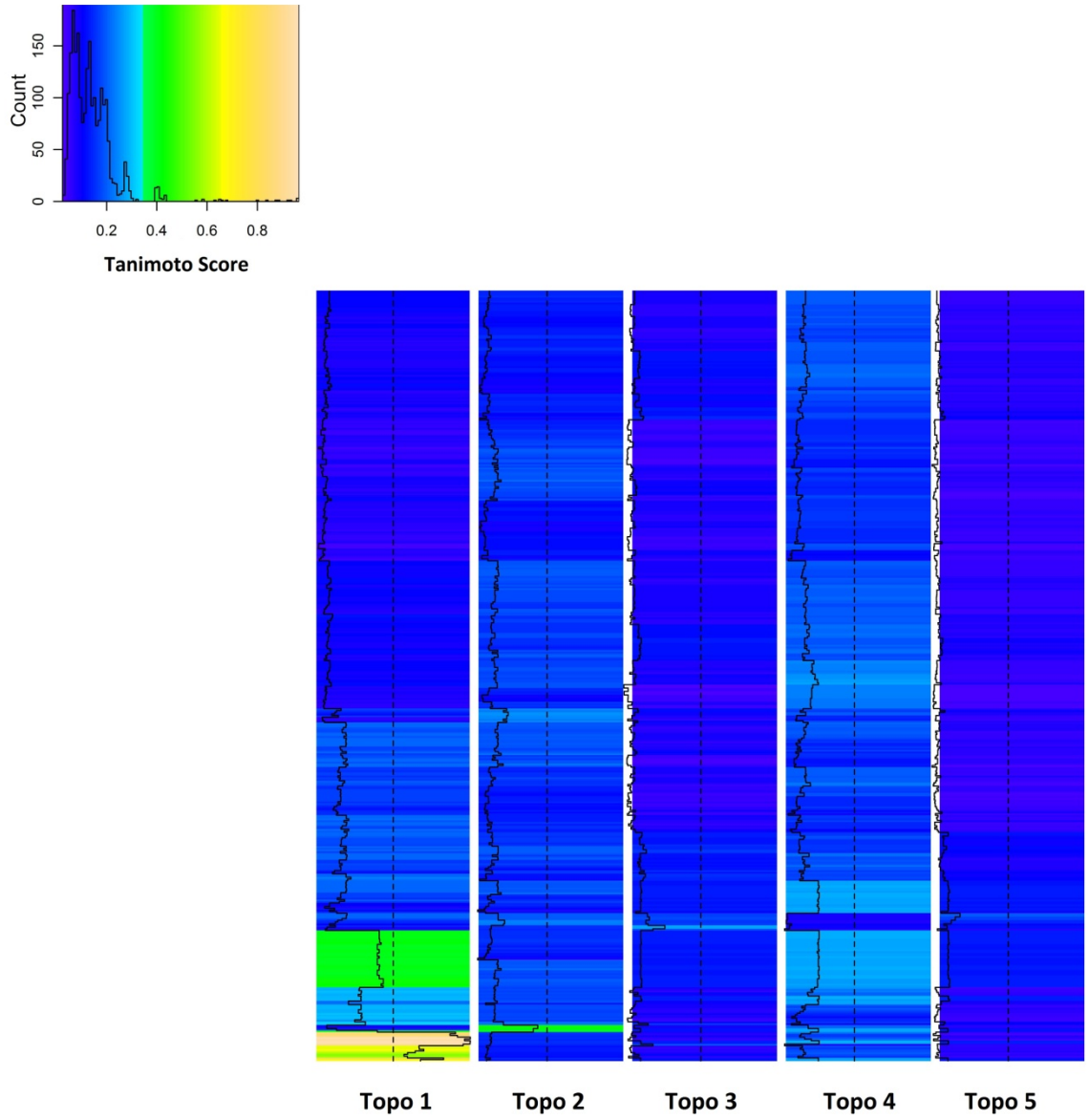


Figure S2. Heatmap representation of the similarity analysis (Tanimoto coefficient, OpenBabel FP2) against known P-gp inhibitors described by Chen et al. (797 compounds). Each row of the figure corresponds to one of the 797 P-gp inhibitors. The five compounds selected for P-glycoprotein 1 bear Tanimoto coefficients lower than 0.5, except for compound P-gp 3 where Tanimoto coefficients between 0.5 and 0.7 can be found. This evidence guarantees the novelty of the P-glycoprotein 1 inhibitors presented in this work.

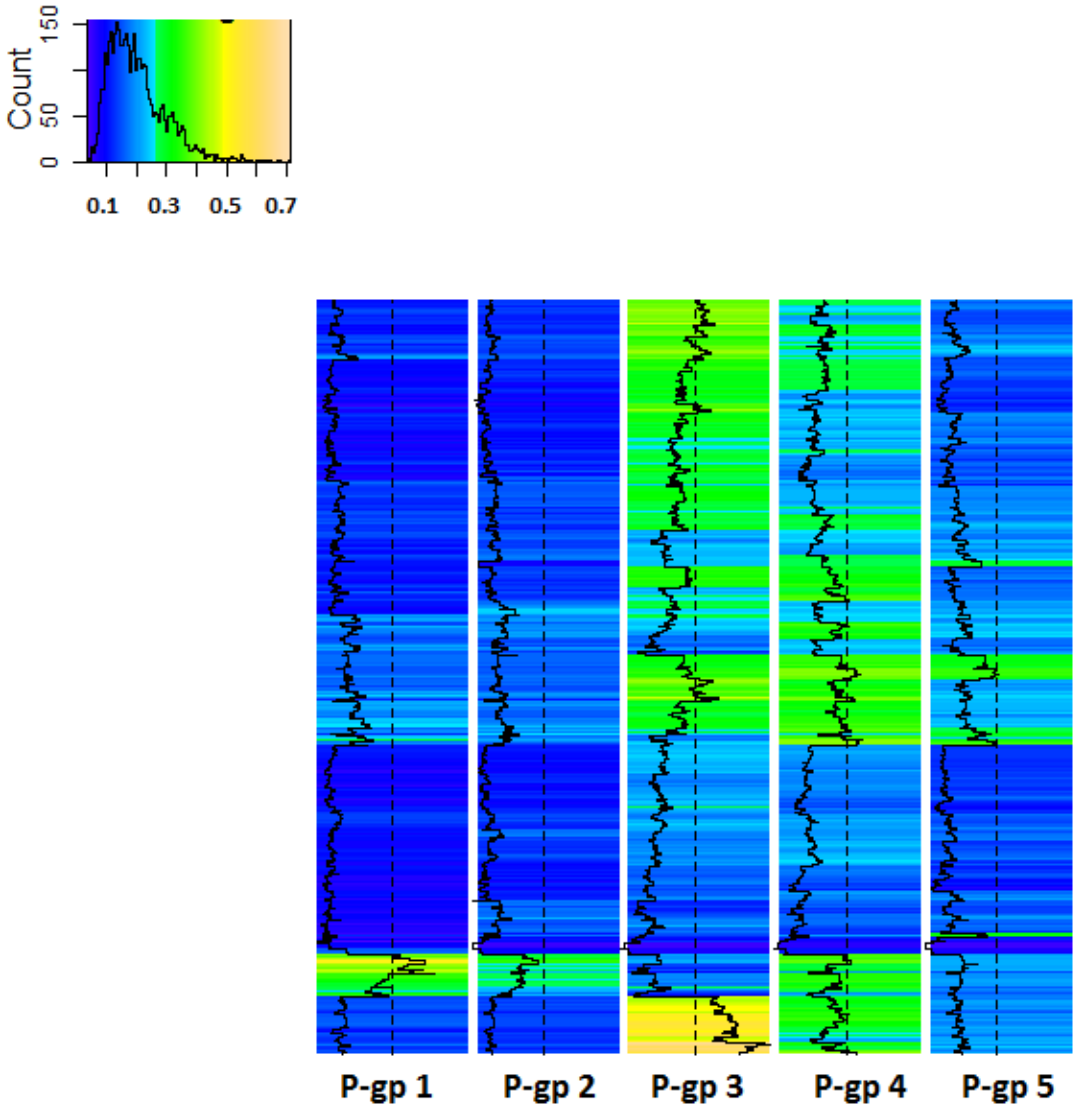


Figure S3. Correlation matrix (Tanimoto coefficient, OpenBabel FP2) for compounds P-gp 1-5 and Topo 1-5**.** The absence of Tanimoto coefficients higher than ~0.6 points to the chemical diversity of the ten compounds subjected to experimental assessment.

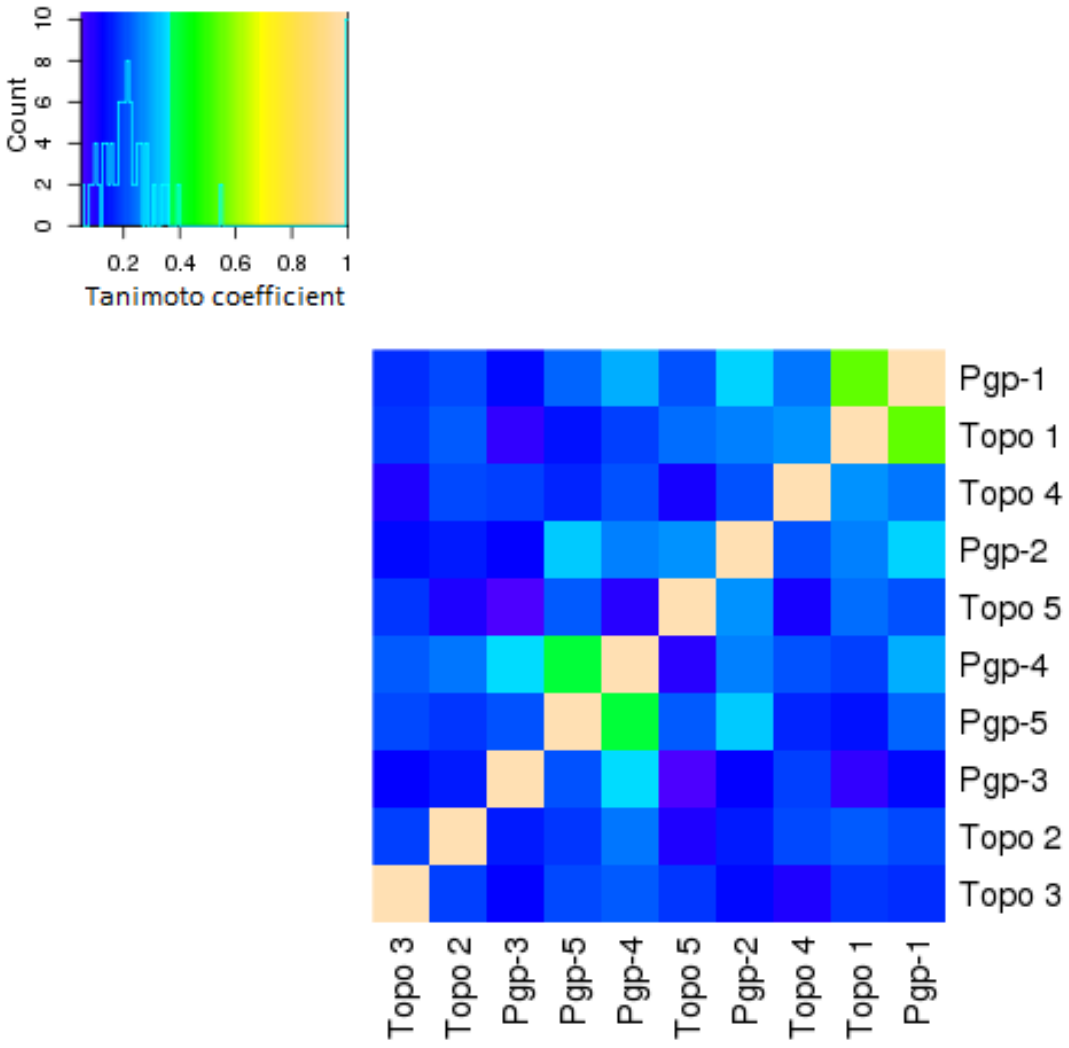


Figure S4. Maximum Common Subgraph (MCS) for compounds selected for a) Topoisomerase I and b) P-glycoprotein 1. Except for the methyl hexopyranoside (2-(hydroxymethyl)-6-methoxyoxane-3,4,5-triol) present in compounds Topo 1,2,4, compounds within each of the two considered activity classes are considerably dissimilar (see Figure S4).

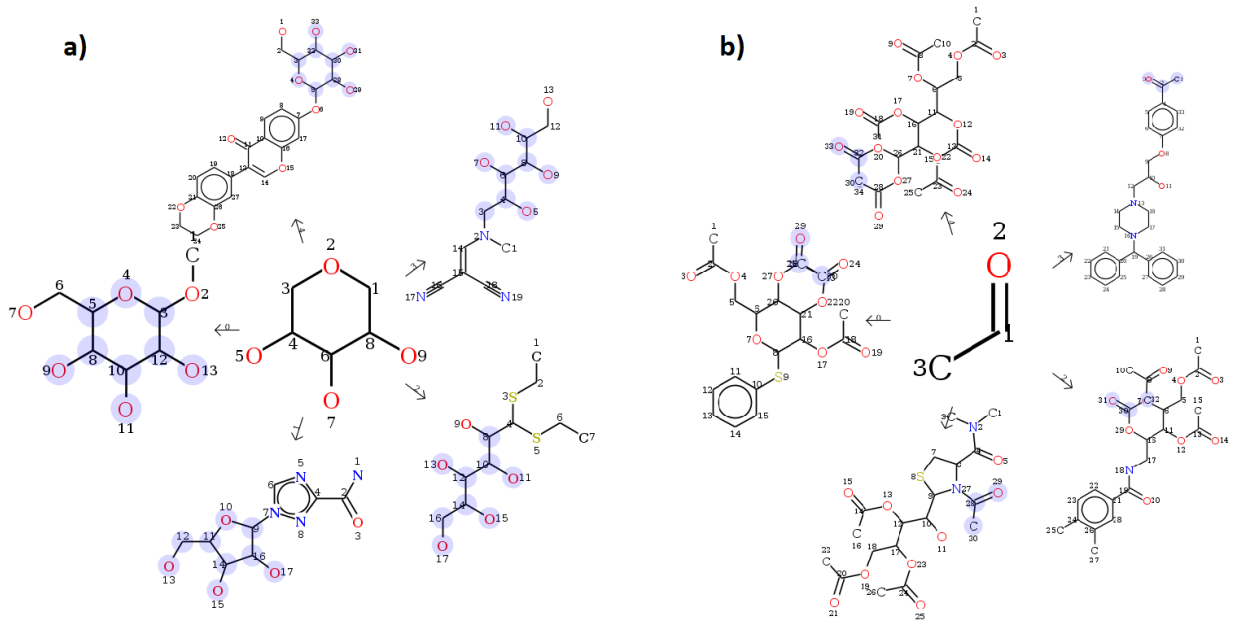
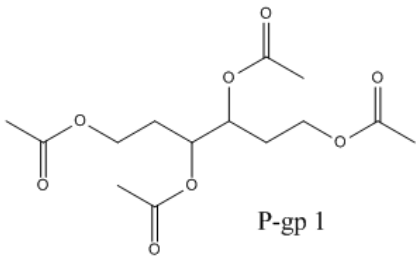
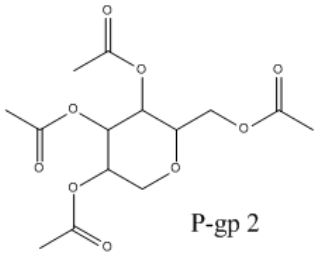


Figure S5. Maximum Common Subgraph (MCS) compounds selected for P-glycoprotein 1 and Topoisomerase I and their respective inhibitors databases described in the text. In the case of compounds Topo 3 and Topo 5, it was not possible to determine a MCS between them and known inhibitors from CHEMBL, concurring with the similarity analysis depicted in Figure S1.

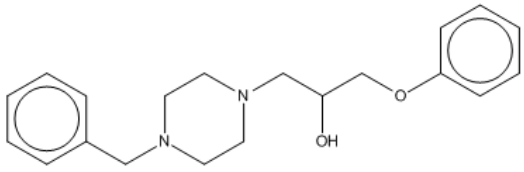


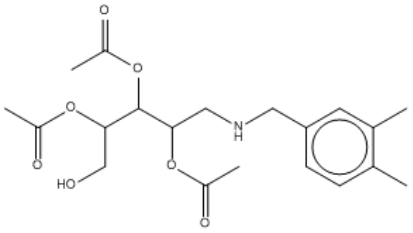
P-gp 1



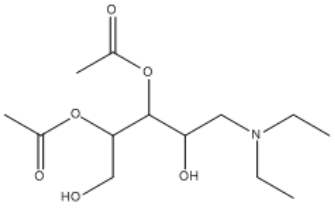
P-gp 2

P-gp 3

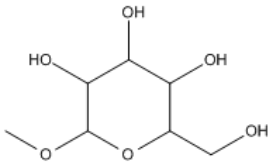




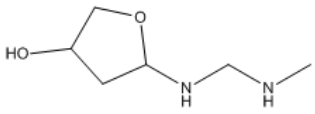
P-gp 4



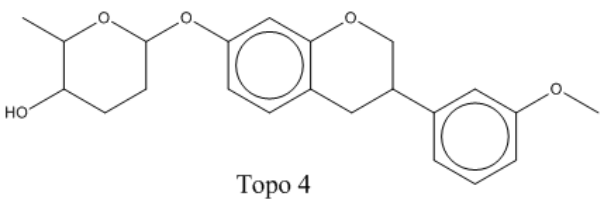
P-gp 5



Topo 1



Topo 2



Topo 4

Figure S6. P-gp biochemical assay performed in the presence of the five compounds predicted to inhibit P-gp, Verapamil and sodium orthovanadate. The five compounds inhibit the activity of the P-gp, albeit compounds P-gp 2 and P-gp 4 produce an effect comparable to that of Verapamil (t-test of independent samples, $p < 0.05$).

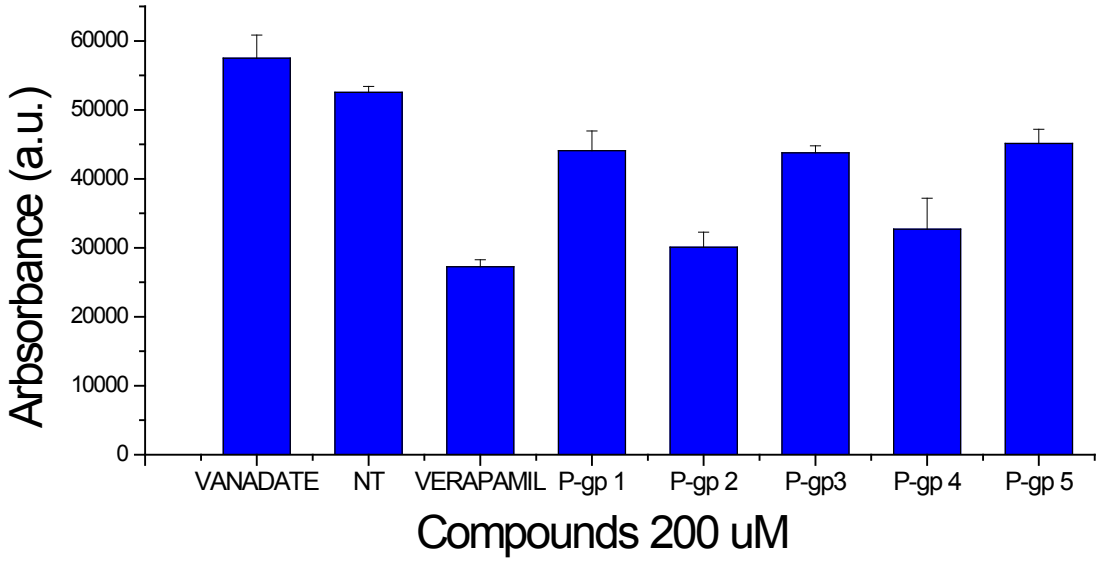


Figure S7. Dose-dependent curve for the inhibition of P-gp in the presence of compounds a) P-gp 2 and b) P-gp 4. Luminescence values are proportional to the concentration of ATP, which is in turn inversely proportional to the activity of P-gp. IC50 values were determined, being 37 ± 5 μM and 28 ± 2 μM for compounds P-gp 2 and P-gp-4 respectively.

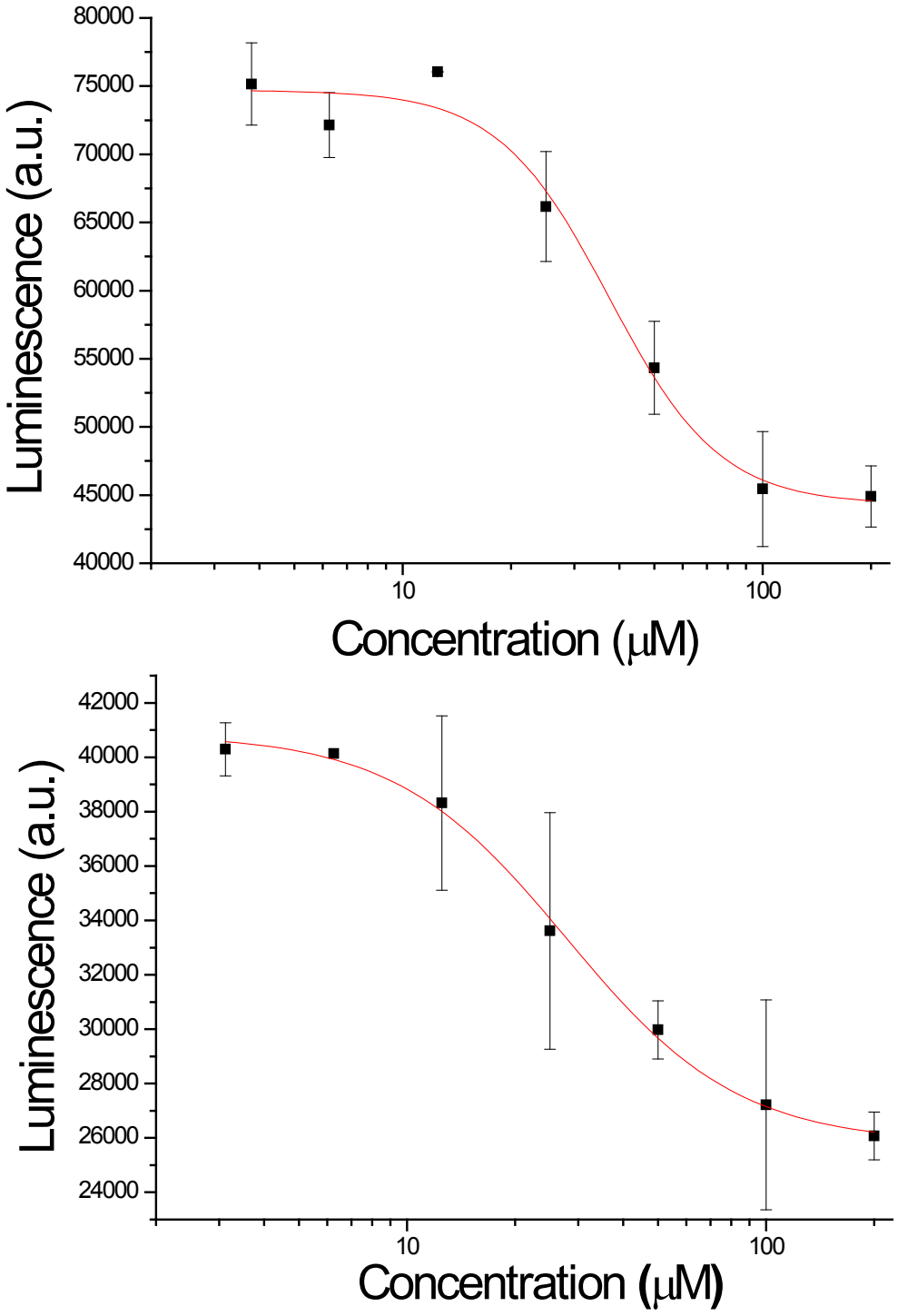


Figure S8. Morphology of HeLa cell cultures in the presence of compounds a) P-gp 2 and b) P-gp 4, after a period of exposure of 48 hours. P-gp 2 and P-gp 4 do not induce apoptosis, as inferred from the morphology of the cells, fact in accordance with the targets predicted for these two compounds (membrane channels). Apoptosis would suggest off-target interactions of these compounds with the DNA or apoptosis-related proteins.