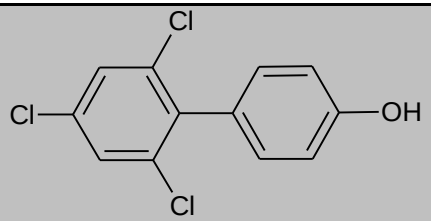
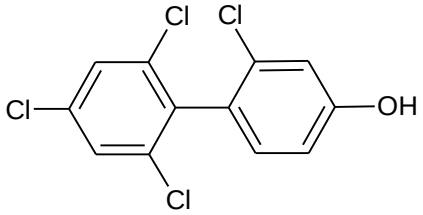
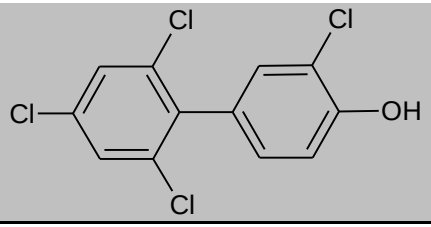
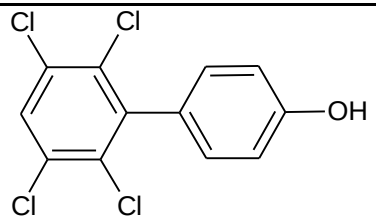
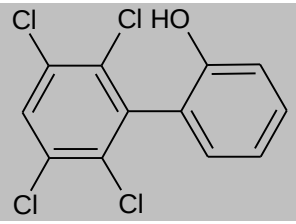
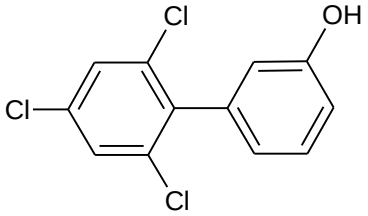
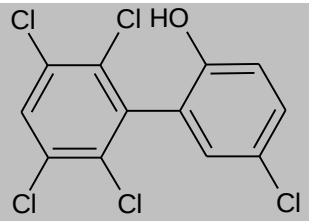
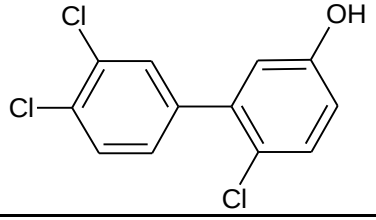
Supplementary Information

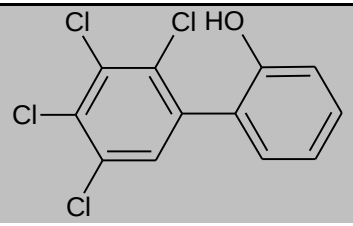
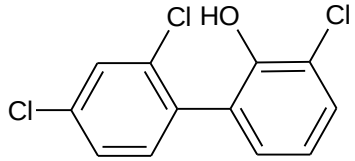
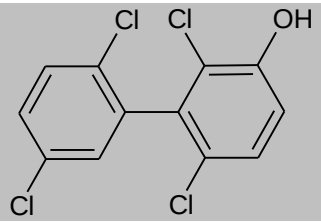
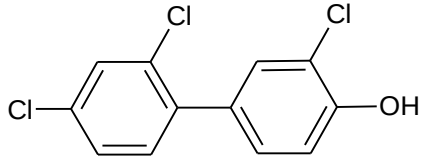
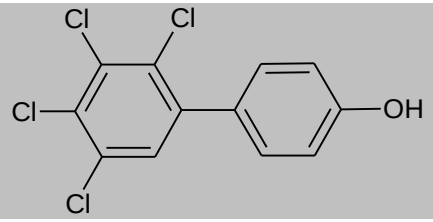
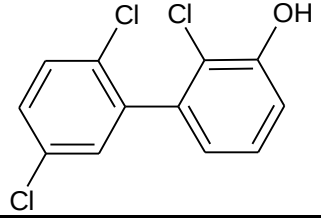
Quantitative structure-activity relationship modeling of hydroxylated polychlorinated biphenyls as constitutive androstane receptor agonists

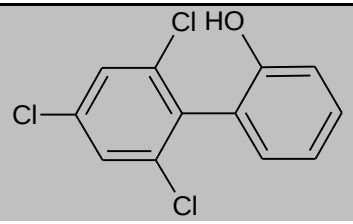
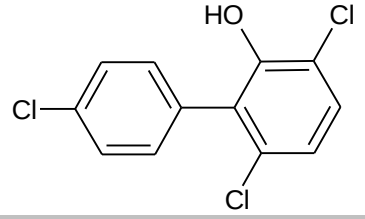
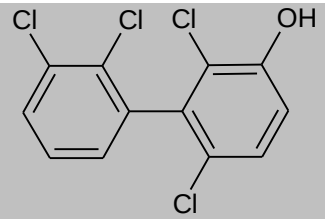
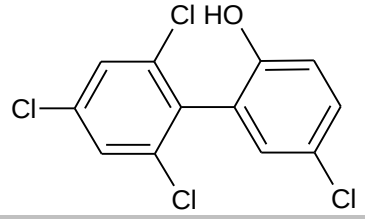
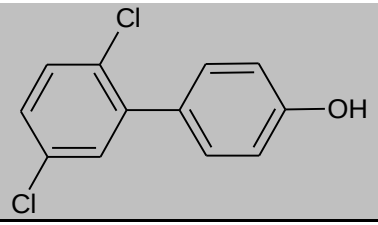
Lukman K. Akinola, Adamu Uzairu, Gideon A. Shallangwa, Stephen E. Abechi

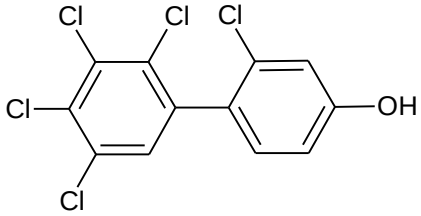
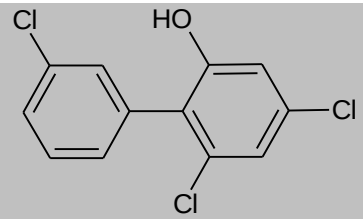
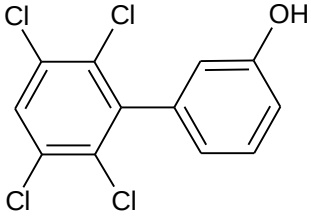
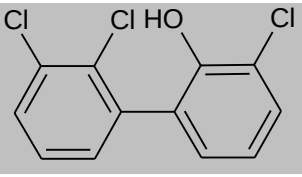
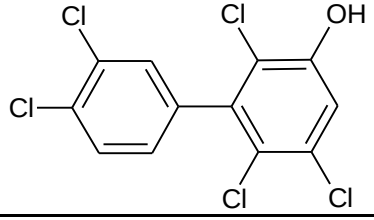
Table S1: Structural formulae of hydroxylated polychlorinated biphenyls and their agonistic activities to constitutive androstane receptor*

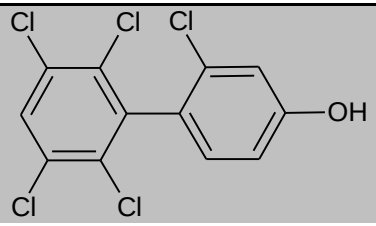
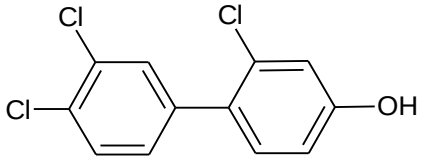
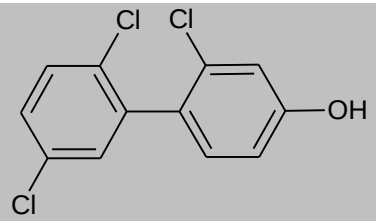
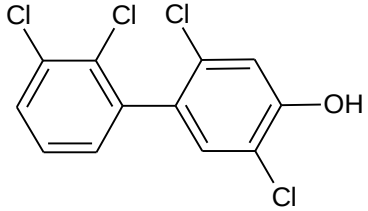
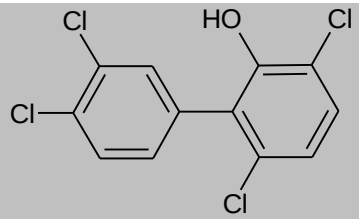
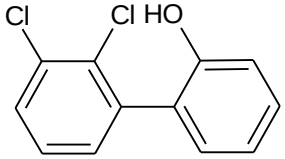
Compound ID	Compound name	Structural formula	EC x 10 (nM)
C1*	2',4',6'-trichlorobiphenyl-4-ol		0.59 ± 0.046
C2	2,2',4',6'-tetrachlorobiphenyl-4-ol		3.3 ± 0.17
C3	2',3,4',6'-tetrachlorobiphenyl-4-ol		26 ± 1.6

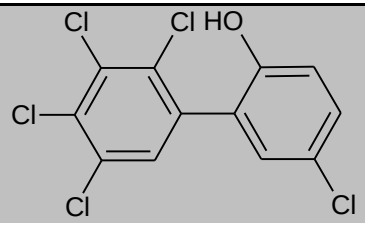
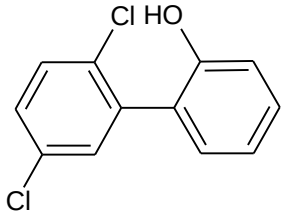
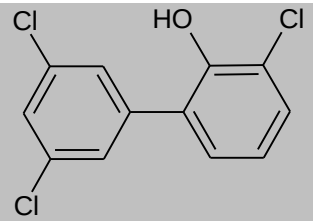
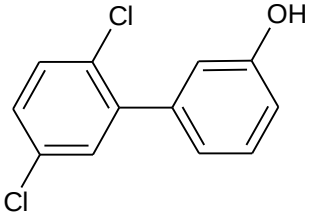
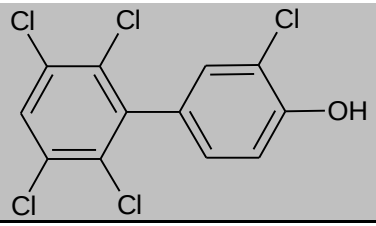
C4	2',3',5',6'-tetrachlorobiphenyl-4-ol		34 ± 1.2
C5	2',3',5',6'-tetrachlorobiphenyl-2-ol		41 ± 5.6
C6	2',4',6'-trichlorobiphenyl-3-ol		76 ± 5.0
C7	2',3',5,5',6'-pentachlorobiphenyl-2-ol		120 ± 0.0
C8	3',4',6-trichlorobiphenyl-3-ol		130 ± 12

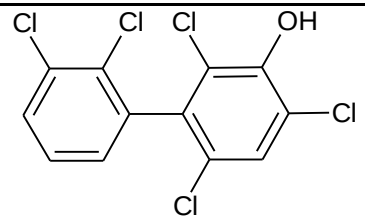
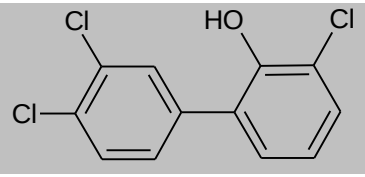
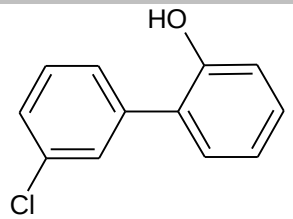
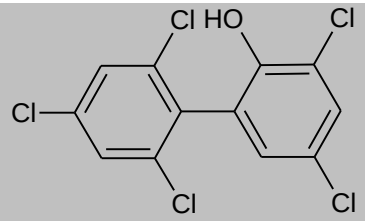
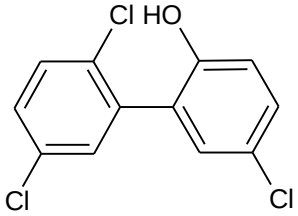
C9	2',3',4',5'-tetrachlorobiphenyl-2-ol		160 ± 21
C10	2',3,4'-trichlorobiphenyl-2-ol		180 ± 12
C11*	2,2',5',6-tetrachlorobiphenyl-3-ol		230 ± 16
C12	2',3,4'-trichlorobiphenyl-4-ol		230 ± 17
C13*	2',3',4',5'-tetrachlorobiphenyl-4-ol		250 ± 12
C14	2,2',5'-trichlorobiphenyl-3-ol		260 ± 26

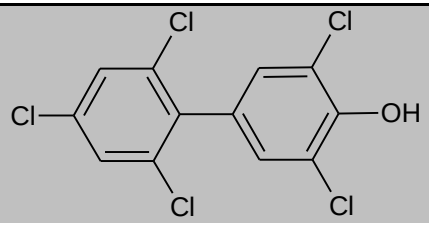
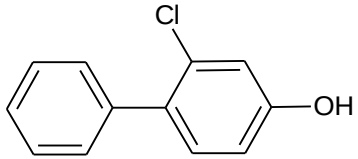
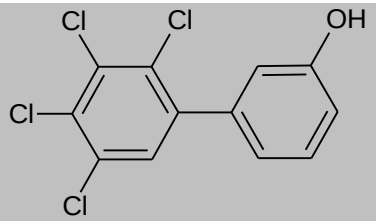
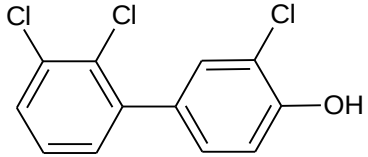
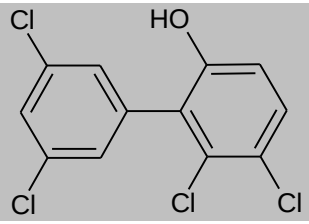
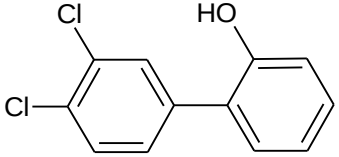
C15	2',4',6'-trichlorobiphenyl-2-ol		270 ± 26
C16*	3,4',6-trichlorobiphenyl-2-ol		270 ± 25
C17	2,2',3',6-tetrachlorobiphenyl-3-ol		280 ± 21
C18	2',4',5,6'-tetrachlorobiphenyl-2-ol		310 ± 33
C19	2',5'-dichlorobiphenyl-4-ol		390 ± 22

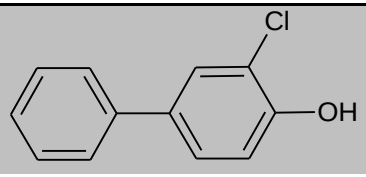
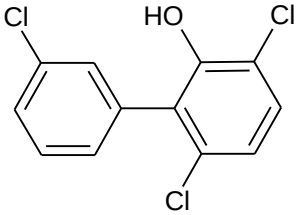
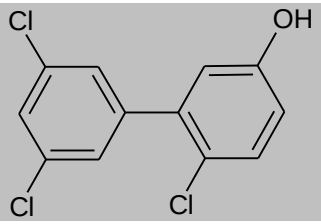
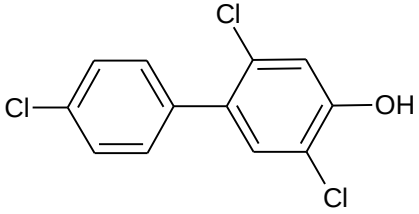
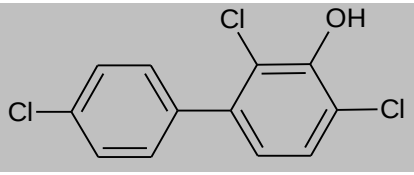
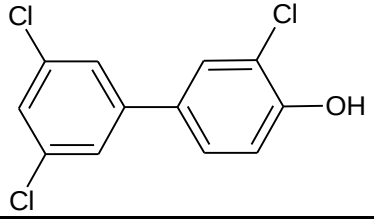
C20	2,2',3',4',5'-pentachlorobiphenyl-4-ol		410 ± 82
C21*	3',4,6-trichlorobiphenyl-2-ol		420 ± 34
C22*	2',3',5',6'-tetrachlorobiphenyl-3-ol		500 ± 17
C23	2',3,3'-trichlorobiphenyl-2-ol		540 ± 43
C24*	2,3',4',5,6-pentachlorobiphenyl-3-ol		540 ± 70

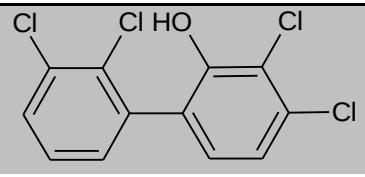
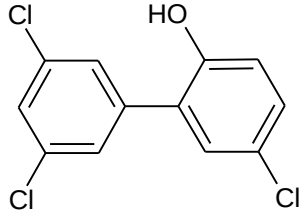
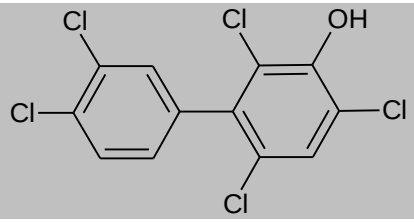
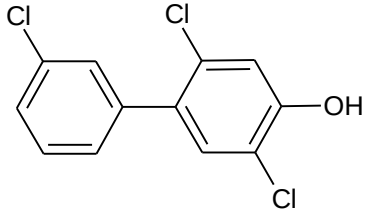
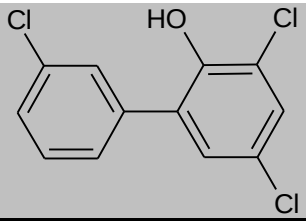
C25	2,2',3',5',6'-pentachlorobiphenyl-4-ol		620 ± 69
C26	2,3',4'-trichlorobiphenyl-4-ol		640 ± 61
C27	2,2',5'-trichlorobiphenyl-4-ol		660 ± 56
C28*	2,2',3',5-tetrachlorobiphenyl-4-ol		930 ± 62
C29	3,3',4',6-tetrachlorobiphenyl-2-ol		960 ± 130
C30	2',3'-dichlorobiphenyl-2-ol		1000 ± 120

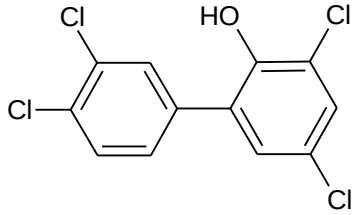
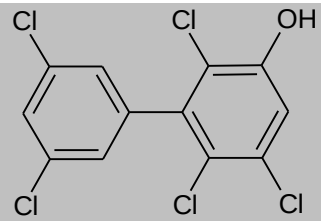
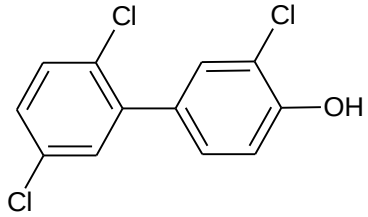
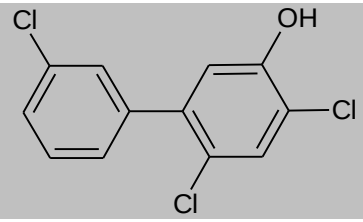
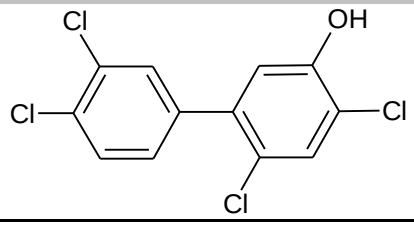
C31	2',3',4',5,5'-pentachlorobiphenyl-2-ol		1200 ± 210
C32	2',5'-dichlorobiphenyl-2-ol		1300 ± 170
C33*	3,3',5'-trichlorobiphenyl-2-ol		1300 ± 140
C34*	2',5'-dichlorobiphenyl-3-ol		1400 ± 120
C35	2',3,3',5',6'-pentachlorobiphenyl-4-ol		1400 ± 340

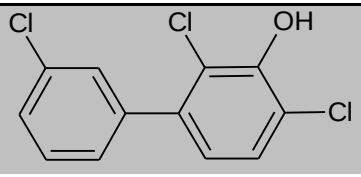
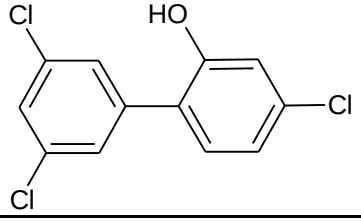
C36*	2,2',3',4,6-pentachlorobiphenyl-3-ol		1500 ± 120
C37*	3,3',4'-trichlorobiphenyl-2-ol		1600 ± 120
C38	5-chlorobiphenyl-2-ol		1700 ± 160
C39	2',3,4',5,6'-pentachlorobiphenyl-2-ol		1900 ± 210
C40	2',5,5'-trichlorobiphenyl-2-ol		2100 ± 370

C41	2',3,4',5,6'-pentachlorobiphenyl-4-ol		2100 ± 210
C42	2-chlorobiphenyl-4-ol		2200 ± 120
C43*	2',3',4',5'-tetrachlorobiphenyl-3-ol		2200 ± 240
C44	2',3,3'-trichlorobiphenyl-4-ol		2400 ± 120
C45	3',5,5',6-tetrachlorobiphenyl-2-ol		2700 ± 94
C46	3',4'-dichlorobiphenyl-2-ol		2900 ± 210

C47	3-chlorobiphenyl-4-ol		2900 ± 82
C48*	3,3',6-trichlorobiphenyl-2-ol		3000 ± 120
C49	3',5',6-trichlorobiphenyl-3-ol		3200 ± 330
C50	2,4',5-trichlorobiphenyl-4-ol		3700 ± 94
C51*	2,4,4'-trichlorobiphenyl-3-ol		4100 ± 190
C52	3,3',5'-trichlorobiphenyl-4-ol		5100 ± 220

C53	2',3,3',4-tetrachlorobiphenyl-2-ol		5300 ± 610
C54*	3',5,5'-trichlorobiphenyl-2-ol		5900 ± 460
C55	2,3',4,4',6-pentachlorobiphenyl-3-ol		6200 ± 340
C56	2,3',5-trichlorobiphenyl-4-ol		6400 ± 650
C57*	3,3',5-trichlorobiphenyl-2-ol		6600 ± 740

C58	3,3',4',5-tetrachlorobiphenyl-2-ol		6900 ± 660
C59	2,3',5,5',6-pentachlorobiphenyl-3-ol		7100 ± 490
C60*	2',3,5'-trichlorobiphenyl-4-ol		7800 ± 860
C61	3',4,6-trichlorobiphenyl-3-ol		7900 ± 400
C62*	3',4,4',6-tetrachlorobiphenyl-3-ol		8000 ± 750

C63	2,3',4-trichlorobiphenyl-3-ol		8700 ± 660
C64*	3',4,5'-trichlorobiphenyl-2-ol		9500 ± 450

*Compounds assigned to test set are marked with single asterisks

Table S2: Computed molecular descriptors for OH-PCBs in the training set

Compound ID	MATS1s	VE3_DzZ	VE1_Dzp	SpMin8_Bhv	SpMax5_Bhi	topoRadius	RDF95u	RDF45m
C2	-0.29653	0.42349	1.10631	-0.09553	-0.50333	2.06534	-0.64477	1.38925
C3	-0.29653	0.30846	0.92603	-0.25216	-0.22549	2.06534	-0.45197	1.14070
C4	-0.18560	1.18157	2.48959	1.40477	-1.45427	-0.47662	3.14915	1.13218
C5	-0.07468	0.94021	0.90486	-0.67153	0.95216	-0.47662	-0.60273	1.07687
C6	-0.33947	0.82649	1.26820	0.41729	1.12190	-0.47662	-0.64477	0.69972
C7	-0.13787	0.54335	0.40020	-2.34800	1.93657	-0.47662	-0.58037	1.62830
C8	-0.33947	-1.06419	-0.20564	0.23737	-0.40326	-0.47662	-0.64477	-0.64683
C9	-0.07468	0.91707	0.71475	-0.68427	1.04373	-0.47662	-0.64477	-0.08710
C10	3.82327	-0.66709	-0.90549	0.01110	-0.78379	-0.47662	-0.25233	-0.59118
C12	-0.33947	-1.02702	0.28546	0.01109	-0.78892	2.06534	-0.39630	-0.59795
C14	-0.23075	-1.19056	0.23078	-0.21431	0.38989	-0.47662	-0.60350	-0.62731
C15	-0.23075	0.63458	0.24941	-0.69732	1.01069	-0.47662	-0.64477	0.56617
C17	-0.07468	-0.31062	-0.35314	-0.21687	0.48532	-0.47662	-0.60151	1.45843
C18	-0.29653	-0.05066	-0.27400	-2.17774	1.12190	-0.47662	-0.38430	1.12443
C19	-0.38006	0.75495	1.49559	1.41527	-1.48278	-0.47662	1.90530	-1.12015
C20	-0.13787	0.74369	1.45229	-0.06004	-0.50251	2.06534	0.50743	0.28991
C23	-0.12202	-2.48398	-0.61333	-0.21270	0.48529	-0.47662	-0.59690	-0.58067
C25	-0.13787	0.80094	1.77183	-0.03172	-0.50461	-0.47662	3.39030	1.99394
C26	-0.33947	-0.31613	0.29569	0.24442	-0.51701	2.06534	-0.45923	-0.59770
C27	-0.33947	-0.09256	0.83690	0.22388	-0.51010	-0.47662	1.51709	-0.64606
C29	-0.18560	0.14602	-1.09541	0.22780	-0.36522	-0.47662	-0.21808	-0.10426
C30	-0.16684	0.33238	-0.04438	-0.19862	0.48421	-0.47662	-0.61899	-0.92798
C31	-0.13787	0.49244	0.20402	-2.65979	1.98891	-0.47662	-0.44109	0.26504
C32	-0.27345	0.25147	-0.08714	0.09025	0.15397	-0.47662	-0.61420	-0.94533
C35	-0.13787	0.73869	1.60960	-0.25036	-0.24623	-0.47662	3.61533	1.68789
C38	-0.31385	-1.40079	-0.67527	0.68544	-0.48286	-0.47662	-0.62589	-1.42701
C39	-0.25109	-1.92430	-0.83300	0.23165	1.12190	-0.47662	-0.22848	1.68920
C40	-0.33947	-3.30464	-0.64593	0.08420	0.25276	-0.47662	-0.59756	-0.49140
C41	-0.25109	-1.48216	0.15839	0.23165	1.12190	2.06534	-0.30440	1.66772
C42	-0.31385	0.77940	-0.11277	1.54469	-1.88230	-0.47662	0.86949	-1.50904
C44	-0.23075	-0.35106	0.60322	-0.21035	-0.25863	-0.47662	1.05597	-0.60212
C45	-0.29653	0.17296	-1.11480	-2.27063	1.12190	-0.47662	-0.59191	0.65163
C46	-0.27345	0.27872	-0.35501	0.24152	-0.43313	-0.47662	-0.64477	-1.27575
C47	-0.31385	0.93609	-0.51377	1.54469	-1.88230	-0.47662	0.85420	-1.63664
C49	-0.44820	-0.79985	0.00428	-2.44136	1.12190	-0.47662	-0.19087	-0.65145
C50	-0.33947	0.66756	-0.99237	1.44244	-1.72560	2.06534	-0.54848	-0.59846
C52	-0.44820	-0.58007	0.31652	-0.24763	-0.27126	-0.47662	1.50721	-0.79098
C53	-0.07468	-0.06923	-1.08822	-0.21282	0.48544	-0.47662	-0.43817	-0.15702
C55	-0.13787	0.41577	-1.21718	0.23013	-0.36559	2.06534	-0.47213	1.54540
C56	-0.33947	0.55046	-0.65299	0.61631	-0.43408	-0.47662	1.28735	-0.60429
C58	-0.29653	0.15414	-1.08302	0.22855	-0.37715	-0.47662	-0.00405	-0.36057
C59	-0.25109	0.43095	-1.29427	0.23165	1.12190	-0.47662	-0.58594	1.76518
C61	-0.33947	0.67837	-1.12519	0.64906	-0.46063	-0.47662	-0.12523	-0.64330
C63	-0.23075	0.69216	-1.09741	0.63445	-0.44871	-0.47662	0.13427	-0.67115

Table S3: Computed molecular descriptors for OH-PCBs in the test set

Compound ID	MATS1s	VE3_DzZ	VE1_Dzp	SpMin8_Bhv	SpMax5_Bhi	topoRadius	RDF95u	RDF45m
C1	3.82327	0.97358	1.83587	1.39963	-1.44283	2.06534	-0.64477	0.58102
C11	3.82327	-0.20342	-0.38381	-0.53711	0.39347	-0.47662	-0.59575	1.42727
C13	-0.18560	1.15443	2.21068	1.40230	-1.43588	2.06534	0.75914	-0.19070
C16	3.82327	0.87324	-0.47018	1.43955	-1.72986	-0.47662	-0.25416	-0.55977
C21	-0.33947	0.71835	-0.85352	0.59444	-0.44510	-0.47662	-0.44817	-0.08025
C22	-0.18560	1.07490	1.93434	0.43865	1.95748	-0.47662	-0.08048	1.24748
C24	-0.13787	0.49011	-1.06616	0.23148	-0.36915	-0.47662	-0.49262	0.38570
C28	-0.18560	-1.00074	0.12683	-0.22017	0.48557	-0.47662	0.84073	-0.14505
C33	-0.33947	-1.56137	-0.80762	-1.44013	1.12190	-0.47662	-0.58270	-0.82401
C34	-0.38006	0.54443	0.94825	0.53476	0.29753	-0.47662	-0.24463	-1.12067
C36	-0.02465	0.23600	-0.86247	-0.21901	0.48540	-0.47662	-0.39013	1.96239
C37	-0.23075	-1.18646	-0.95924	0.23327	-0.37788	-0.47662	0.02094	-0.81118
C43	-0.18560	1.05042	1.67761	0.42856	1.98644	-0.47662	-0.64277	-0.18019
C48	-0.23075	0.79370	-0.73589	0.60575	-0.44819	-0.47662	-0.59722	-0.55323
C51	-0.23075	0.78487	-1.22661	1.44613	-1.73017	2.06534	-0.64477	-0.65710
C54	-0.44820	-2.68413	-0.79332	-1.60996	1.12190	-0.47662	-0.57431	-0.84538
C57	-0.33947	0.80529	-0.69035	0.60929	-0.45398	-0.47662	-0.58549	-0.80757
C60	-0.33947	-0.49111	0.57437	-0.21347	-0.35137	-0.47662	1.44744	-0.66181
C62	-0.29653	-0.13094	-0.82454	0.23420	-0.39019	2.06534	-0.51056	-0.18239
C64	-0.44820	-0.91858	-0.57844	-2.08405	1.11243	-0.47662	-0.16962	-0.86489

Table S4: Observed activities, predicted activities, residuals, standardized residuals and leverages of compounds in the training set

Compound ID	$log\ 1/(EC\ x\ 10)$ (Observed)	$log\ 1/(EC\ x\ 10)$ (Predicted)	Residual	Standardized residual	Leverage
C2	8.48149	7.966203	0.515287	2.145471	0.284691
C3	7.58503	7.347388	0.237642	0.920130	0.172844
C4	7.46852	7.365296	0.103224	0.440642	0.319494
C5	7.38722	7.411666	-0.024446	-0.094058	0.162347
C6	7.11919	7.201262	-0.082072	-0.350685	0.320802
C7	6.92082	6.777131	0.143689	0.557318	0.175709
C8	6.88606	6.663666	0.222394	0.849106	0.149323
C9	6.79588	6.659911	0.135969	0.522666	0.160785
C10	6.74473	6.734922	0.009808	0.262365	0.982671
C12	6.63827	6.427436	0.210834	0.826975	0.193997
C14	6.58503	6.542039	0.042991	0.159579	0.099994
C15	6.56864	6.553141	0.015499	0.057087	0.085946
C17	6.55284	6.804269	-0.251429	-0.973778	0.173292
C18	6.50864	6.407851	0.100789	0.396573	0.199023
C19	6.40894	6.419116	-0.010176	-0.039398	0.172771
C20	6.38722	6.844968	-0.457748	-1.797527	0.195838
C23	6.26761	5.836159	0.431451	1.673544	0.175806
C25	6.20761	6.487348	-0.279738	-1.167294	0.287830
C26	6.19382	6.139271	0.054549	0.211415	0.174458
C27	6.18046	5.899550	0.280910	1.037183	0.090369
C29	6.01773	5.596041	0.421689	1.556196	0.089465
C30	6.00000	5.924227	0.075773	0.282502	0.107865
C31	5.92082	5.883764	0.037056	0.150663	0.249848
C32	5.88606	6.065797	-0.179737	-0.660134	0.080713
C35	5.85387	5.857308	-0.003438	-0.014556	0.308382
C38	5.76955	5.848196	-0.078646	-0.295874	0.123855
C39	5.72125	5.679553	0.041697	0.173037	0.279921
C40	5.67778	6.019430	-0.341650	-1.453936	0.315280

C41	5.67778	5.739072	-0.061292	-0.285617	0.428942
C42	5.65758	5.853015	-0.195435	-0.748867	0.155429
C44	5.61979	6.008286	-0.388496	-1.420896	0.072981
C45	5.56864	5.557225	0.011415	0.045857	0.231683
C46	5.53760	6.124280	-0.586680	-2.167012	0.091088
C47	5.53760	5.409837	0.127763	0.493424	0.168597
C49	5.49485	5.810539	-0.315689	-1.279852	0.245532
C50	5.43180	5.598578	-0.166778	-0.674190	0.241153
C52	5.29243	5.292523	-0.000093	-0.000352	0.139053
C53	5.27572	5.205187	0.070533	0.258627	0.077676
C55	5.20761	5.540103	-0.332493	-1.369421	0.268976
C56	5.19382	4.473613	0.720207	2.805894	0.183015
C58	5.16115	5.307314	-0.146164	-0.536516	0.079643
C59	5.14874	5.314924	-0.166184	-0.681589	0.262817
C61	5.10237	5.145157	-0.042787	-0.159423	0.106759
C63	5.06048	4.960480	0.100000	0.373975	0.113335

Table S5: Observed activities, predicted activities, residuals, standardized residuals and leverages of compounds in the test set

Compound ID	$log\ 1/(EC\ x\ 10)$ (Observed)	$log\ 1/(EC\ x\ 10)$ (Predicted)	Residual	Standardized residual	Leverage
C1	9.22915	9.354426	-0.125276	0.220046	0.691101
C11	6.63827	7.553221	-0.914951	-1.271137	0.076381
C13	6.60206	7.518333	-0.916273	-1.273634	0.073563
C16	6.56864	7.448096	-0.879456	-1.204110	0.071103
C21	6.37675	5.917256	0.459494	1.324297	0.059603
C22	6.30103	6.947388	-0.646358	-0.763939	0.056295
C24	6.26761	6.044336	0.223274	0.878231	0.055057
C28	6.03152	5.415000	0.616520	1.620817	0.050168
C33	5.88606	5.151675	0.734385	1.843388	0.050524
C34	5.85387	6.403547	-0.549677	-0.581372	0.050950
C36	5.82391	6.334598	-0.510688	-0.507747	0.051459
C37	5.79588	5.349913	0.445967	1.298753	0.052034
C43	5.65758	6.479917	-0.822337	-1.096249	0.056268
C48	5.52288	5.941157	-0.418277	-0.333243	0.062624
C51	5.38722	5.435909	-0.048689	0.364669	0.071250
C54	5.22915	5.287427	-0.058277	0.346564	0.084120
C57	5.18046	5.841892	-0.661432	-0.792404	0.088695
C60	5.10791	5.710937	-0.603027	-0.682115	0.096046
C62	5.09691	5.197164	-0.100254	0.267296	0.097216
C64	5.02228	5.083032	-0.060752	0.341890	0.105544