

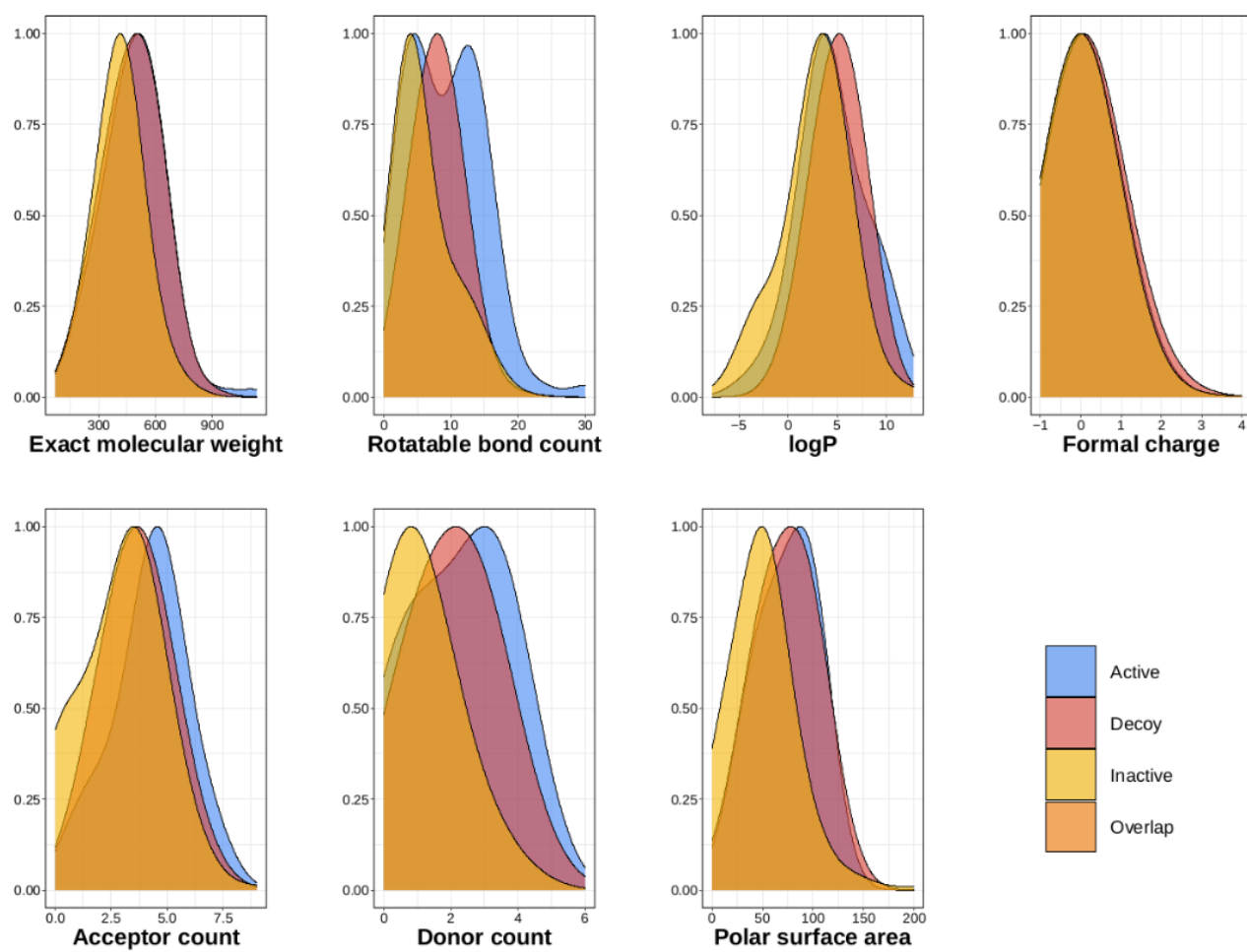
A practical guide to machine-learning scoring for structure-based virtual screening

In the format provided by the
authors and unedited

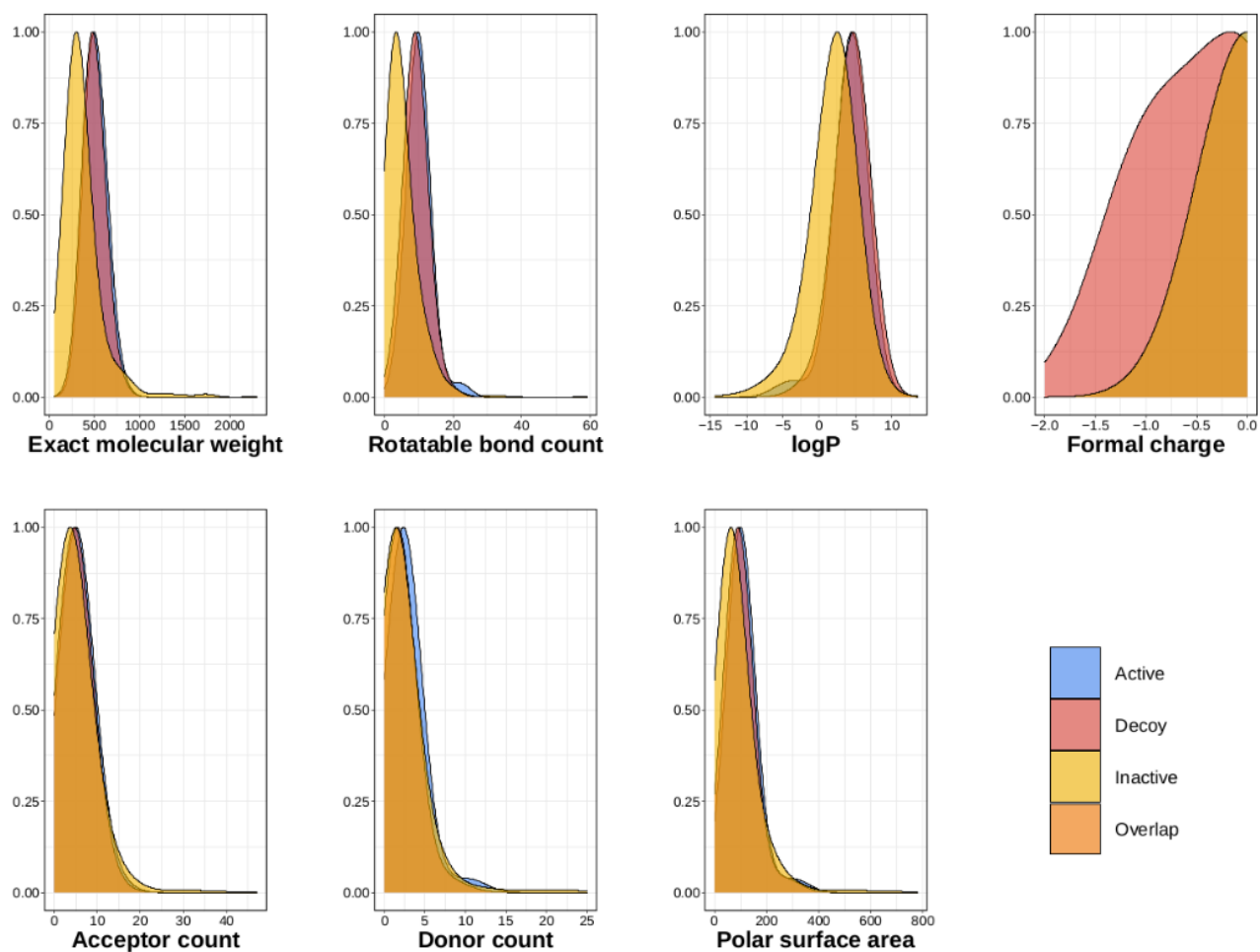
Supplementary Table 1. SBVS performance of four generic SFs (Smina, IFP, CNN-Score, RF-Score-VS) on three DEKOIS 2.0 benchmarks (ACHE, HMGR, PPARA).

	ACHE	HMGR	PPARA
- Number of actives	40	40	40
- Number of decoys	1200	1200	1200
- Total number of molecules	1240	1240	1240
- Maximal EF1%	30.00	30.00	30.00
Smina:			
- EF1%	15.00	2.50	0.00
- NEF1%	0.500	0.083	0.000
IFP:			
- EF1%	15.00	10.00	0.00
- NEF1%	0.500	0.333	0.000
CNN-Score:			
- EF1%	0.00	7.50	10.00
- NEF1%	0.000	0.250	0.333
RF-Score-VS:			
- EF1%	2.50	30.00	12.50
- NEF1%	0.083	1.000	0.417

Supplementary Figure 1. Distribution plots of seven physico-chemical properties calculated from all molecules of the ACHE ligand set.



Supplementary Figure 2. Distribution plots of seven physico-chemical properties calculated from all molecules of the HMGR ligand set.



Supplementary Table 2. List of all active molecules of the users' own PPARA benchmark ($n = 1115$) and their partition into the test set and the training set according to their clusters (OTS option).

Cluster group	ID	Data set
118	CHEMBL260627	Test
125	CHEMBL400625	Test
127	CHEMBL2237300	Test
128	CHEMBL4083589	Test
132	CHEMBL1934468	Test
144	CHEMBL280897	Test
145	CHEMBL252995	Test
150	CHEMBL198579	Test
155	CHEMBL205777	Test
156	CHEMBL472451	Test
157	CHEMBL192500	Test
158	CHEMBL278994	Test
159	CHEMBL410478	Test
160	CHEMBL3780740	Test
161	CHEMBL2037182	Test
162	CHEMBL2037008	Test
163	CHEMBL4571975	Test
164	CHEMBL4104138	Test
165	CHEMBL4101577	Test
166	CHEMBL4100619	Test
167	CHEMBL3982782	Test
168	CHEMBL3735011	Test
169	CHEMBL3601096	Test
170	CHEMBL3397563	Test
171	CHEMBL3354988	Test
172	CHEMBL3354987	Test
173	CHEMBL2069935	Test
174	CHEMBL1934308	Test
175	CHEMBL1927476	Test
176	CHEMBL1760555	Test
177	CHEMBL1720494	Test
178	CHEMBL1494365	Test
179	CHEMBL1089633	Test
180	CHEMBL1080361	Test
181	CHEMBL1079489	Test
182	CHEMBL1075683	Test
183	CHEMBL584704	Test
184	CHEMBL572795	Test
185	CHEMBL526740	Test
186	CHEMBL438346	Test
187	CHEMBL426883	Test
188	CHEMBL401811	Test
189	CHEMBL401562	Test
190	CHEMBL374948	Test
191	CHEMBL318696	Test
192	CHEMBL306229	Test
193	CHEMBL272017	Test
194	CHEMBL250673	Test
195	CHEMBL249865	Test

196	CHEMBL219156	Test
197	CHEMBL213355	Test
198	CHEMBL207569	Test
199	CHEMBL202631	Test
200	CHEMBL198929	Test
201	CHEMBL198786	Test
202	CHEMBL194698	Test
203	CHEMBL190739	Test
204	CHEMBL190645	Test
205	CHEMBL39687	Test
206	CHEMBL496744	Test
207	CHEMBL491039	Test
208	CHEMBL391085	Test
209	CHEMBL391084	Test
210	CHEMBL593087	Test
211	CHEMBL257345	Test
212	CHEMBL221529	Test
213	CHEMBL148797	Test
214	CHEMBL2113219	Test
215	CHEMBL383011	Test
216	CHEMBL255872	Test
217	CHEMBL3707395	Test
218	CHEMBL592054	Test
219	CHEMBL426878	Test
220	CHEMBL242007	Test
221	CHEMBL148319	Test
222	CHEMBL214921	Test
223	CHEMBL221044	Test
224	CHEMBL424133	Test
225	CHEMBL280065	Test
226	CHEMBL24458	Test
0	CHEMBL107789	Training
0	CHEMBL118939	Training
0	CHEMBL120009	Training
0	CHEMBL202118	Training
0	CHEMBL375471	Training
0	CHEMBL120246	Training
0	CHEMBL120456	Training
0	CHEMBL255392	Training
0	CHEMBL270587	Training
0	CHEMBL331206	Training
0	CHEMBL332465	Training
0	CHEMBL382655	Training
0	CHEMBL387914	Training
0	CHEMBL402161	Training
0	CHEMBL427071	Training
0	CHEMBL116540	Training
0	CHEMBL120611	Training
0	CHEMBL201803	Training
0	CHEMBL201832	Training
0	CHEMBL202609	Training
0	CHEMBL229525	Training
0	CHEMBL229573	Training
0	CHEMBL229778	Training

0	CHEMBL229785	Training
0	CHEMBL229786	Training
0	CHEMBL229836	Training
0	CHEMBL229837	Training
0	CHEMBL229888	Training
0	CHEMBL262558	Training
0	CHEMBL270801	Training
0	CHEMBL271877	Training
0	CHEMBL376165	Training
0	CHEMBL389903	Training
0	CHEMBL390116	Training
0	CHEMBL425666	Training
0	CHEMBL426483	Training
0	CHEMBL455957	Training
0	CHEMBL462982	Training
0	CHEMBL1222408	Training
0	CHEMBL1779473	Training
0	CHEMBL1779474	Training
0	CHEMBL1779476	Training
0	CHEMBL1779478	Training
0	CHEMBL1779479	Training
0	CHEMBL1779480	Training
0	CHEMBL1779481	Training
0	CHEMBL1779482	Training
0	CHEMBL1956149	Training
1	CHEMBL1089501	Training
1	CHEMBL1089210	Training
1	CHEMBL3828718	Training
1	CHEMBL1088910	Training
1	CHEMBL1089164	Training
1	CHEMBL1089500	Training
1	CHEMBL1090189	Training
1	CHEMBL1090211	Training
1	CHEMBL1090878	Training
1	CHEMBL1091196	Training
1	CHEMBL1091218	Training
1	CHEMBL1091980	Training
1	CHEMBL1092235	Training
1	CHEMBL1092236	Training
1	CHEMBL1092237	Training
1	CHEMBL1092931	Training
1	CHEMBL3827820	Training
1	CHEMBL1088911	Training
1	CHEMBL1090877	Training
1	CHEMBL1087466	Training
1	CHEMBL1087467	Training
1	CHEMBL1087580	Training
1	CHEMBL1087582	Training
1	CHEMBL1089969	Training
1	CHEMBL1090187	Training
1	CHEMBL1092932	Training
1	CHEMBL1097193	Training
1	CHEMBL1097194	Training
1	CHEMBL1097195	Training

1	CHEMBL3826990	Training
1	CHEMBL3827050	Training
1	CHEMBL3827241	Training
1	CHEMBL3827335	Training
1	CHEMBL3827378	Training
1	CHEMBL3827510	Training
1	CHEMBL3827584	Training
1	CHEMBL3827617	Training
1	CHEMBL3827740	Training
1	CHEMBL3827940	Training
1	CHEMBL3828133	Training
1	CHEMBL3828201	Training
1	CHEMBL3828232	Training
1	CHEMBL3828330	Training
1	CHEMBL3828487	Training
2	CHEMBL112725	Training
2	CHEMBL361424	Training
2	CHEMBL435343	Training
2	CHEMBL367891	Training
2	CHEMBL360246	Training
2	CHEMBL367549	Training
2	CHEMBL190603	Training
2	CHEMBL192582	Training
2	CHEMBL192638	Training
2	CHEMBL366145	Training
2	CHEMBL111480	Training
2	CHEMBL112973	Training
2	CHEMBL176222	Training
2	CHEMBL178513	Training
2	CHEMBL178630	Training
2	CHEMBL179180	Training
2	CHEMBL179452	Training
2	CHEMBL179657	Training
2	CHEMBL180912	Training
2	CHEMBL360231	Training
2	CHEMBL360240	Training
2	CHEMBL360992	Training
2	CHEMBL361166	Training
2	CHEMBL426592	Training
2	CHEMBL447190	Training
2	CHEMBL109638	Training
2	CHEMBL176075	Training
2	CHEMBL178309	Training
2	CHEMBL179622	Training
2	CHEMBL326873	Training
2	CHEMBL2037079	Training
2	CHEMBL111161	Training
2	CHEMBL178659	Training
2	CHEMBL369289	Training
2	CHEMBL370024	Training
3	CHEMBL191275	Training
3	CHEMBL489074	Training
3	CHEMBL192518	Training
3	CHEMBL489075	Training

3	CHEMBL191060	Training
3	CHEMBL189116	Training
3	CHEMBL470145	Training
3	CHEMBL472244	Training
3	CHEMBL513244	Training
3	CHEMBL572565	Training
3	CHEMBL572566	Training
3	CHEMBL572794	Training
3	CHEMBL573015	Training
3	CHEMBL573252	Training
3	CHEMBL574386	Training
3	CHEMBL574607	Training
3	CHEMBL578860	Training
3	CHEMBL2312700	Training
3	CHEMBL2312701	Training
3	CHEMBL2312702	Training
3	CHEMBL2312703	Training
3	CHEMBL2312704	Training
3	CHEMBL2312705	Training
3	CHEMBL2312706	Training
3	CHEMBL2312707	Training
3	CHEMBL2312708	Training
3	CHEMBL2312711	Training
3	CHEMBL2312712	Training
3	CHEMBL2312713	Training
3	CHEMBL2312714	Training
3	CHEMBL2312715	Training
3	CHEMBL2312716	Training
3	CHEMBL2312717	Training
3	CHEMBL2312718	Training
3	CHEMBL2375951	Training
4	CHEMBL247951	Training
4	CHEMBL245464	Training
4	CHEMBL245465	Training
4	CHEMBL245664	Training
4	CHEMBL245865	Training
4	CHEMBL245866	Training
4	CHEMBL246069	Training
4	CHEMBL246070	Training
4	CHEMBL246281	Training
4	CHEMBL246282	Training
4	CHEMBL246510	Training
4	CHEMBL247748	Training
4	CHEMBL247749	Training
4	CHEMBL247952	Training
4	CHEMBL248117	Training
4	CHEMBL248324	Training
4	CHEMBL248325	Training
4	CHEMBL248326	Training
4	CHEMBL248526	Training
4	CHEMBL248527	Training
4	CHEMBL391620	Training
4	CHEMBL391622	Training
4	CHEMBL392810	Training

4	CHEMBL392811	Training
4	CHEMBL394296	Training
4	CHEMBL395924	Training
4	CHEMBL398117	Training
4	CHEMBL398118	Training
4	CHEMBL427701	Training
4	CHEMBL429824	Training
4	CHEMBL477119	Training
4	CHEMBL477312	Training
4	CHEMBL479207	Training
5	CHEMBL247919	Training
5	CHEMBL248118	Training
6	CHEMBL318351	Training
6	CHEMBL116947	Training
6	CHEMBL119791	Training
6	CHEMBL122070	Training
6	CHEMBL3764076	Training
7	CHEMBL482740	Training
7	CHEMBL483341	Training
7	CHEMBL484330	Training
7	CHEMBL484409	Training
7	CHEMBL484606	Training
7	CHEMBL484607	Training
7	CHEMBL484725	Training
7	CHEMBL485187	Training
7	CHEMBL485188	Training
7	CHEMBL485537	Training
7	CHEMBL519032	Training
7	CHEMBL519351	Training
7	CHEMBL519802	Training
7	CHEMBL521170	Training
7	CHEMBL521462	Training
7	CHEMBL484000	Training
7	CHEMBL484001	Training
7	CHEMBL484137	Training
7	CHEMBL484962	Training
7	CHEMBL519504	Training
7	CHEMBL519974	Training
7	CHEMBL521007	Training
7	CHEMBL482993	Training
7	CHEMBL484138	Training
7	CHEMBL1813008	Training
7	CHEMBL1813009	Training
7	CHEMBL1813012	Training
8	CHEMBL428574	Training
9	CHEMBL209546	Training
9	CHEMBL209749	Training
9	CHEMBL210017	Training
9	CHEMBL210163	Training
9	CHEMBL211371	Training
9	CHEMBL213487	Training
9	CHEMBL213602	Training
9	CHEMBL213675	Training
9	CHEMBL213678	Training

9	CHEMBL213902	Training
9	CHEMBL214272	Training
9	CHEMBL265364	Training
9	CHEMBL378705	Training
9	CHEMBL380070	Training
9	CHEMBL384203	Training
9	CHEMBL387023	Training
9	CHEMBL211513	Training
9	CHEMBL211839	Training
9	CHEMBL377898	Training
9	CHEMBL380404	Training
9	CHEMBL2037078	Training
9	CHEMBL2037080	Training
9	CHEMBL2037081	Training
9	CHEMBL2037083	Training
10	CHEMBL81248	Training
10	CHEMBL82034	Training
10	CHEMBL81592	Training
10	CHEMBL412371	Training
10	CHEMBL246272	Training
10	CHEMBL246473	Training
10	CHEMBL247925	Training
10	CHEMBL247926	Training
10	CHEMBL394967	Training
10	CHEMBL395191	Training
10	CHEMBL396694	Training
10	CHEMBL428748	Training
10	CHEMBL429823	Training
10	CHEMBL245656	Training
10	CHEMBL245674	Training
10	CHEMBL246270	Training
10	CHEMBL394947	Training
10	CHEMBL396696	Training
10	CHEMBL557580	Training
10	CHEMBL394948	Training
10	CHEMBL394968	Training
10	CHEMBL397120	Training
11	CHEMBL180019	Training
11	CHEMBL369725	Training
11	CHEMBL190968	Training
11	CHEMBL178909	Training
11	CHEMBL182200	Training
12	CHEMBL76676	Training
12	CHEMBL78442	Training
12	CHEMBL78476	Training
12	CHEMBL78523	Training
12	CHEMBL80301	Training
12	CHEMBL80364	Training
12	CHEMBL80671	Training
12	CHEMBL80772	Training
12	CHEMBL80791	Training
12	CHEMBL80905	Training
12	CHEMBL80979	Training
12	CHEMBL305723	Training

12	CHEMBL305854	Training
12	CHEMBL309778	Training
12	CHEMBL310426	Training
12	CHEMBL310962	Training
12	CHEMBL311429	Training
12	CHEMBL312791	Training
12	CHEMBL433306	Training
12	CHEMBL433316	Training
12	CHEMBL189146	Training
12	CHEMBL2237303	Training
13	CHEMBL402840	Training
13	CHEMBL272137	Training
13	CHEMBL270451	Training
13	CHEMBL270664	Training
13	CHEMBL270665	Training
13	CHEMBL402841	Training
13	CHEMBL456355	Training
13	CHEMBL456752	Training
13	CHEMBL456753	Training
13	CHEMBL456973	Training
13	CHEMBL456974	Training
13	CHEMBL457405	Training
13	CHEMBL457623	Training
13	CHEMBL514256	Training
13	CHEMBL514264	Training
13	CHEMBL2158959	Training
14	CHEMBL173159	Training
14	CHEMBL435523	Training
14	CHEMBL172492	Training
14	CHEMBL172931	Training
14	CHEMBL173236	Training
14	CHEMBL173285	Training
14	CHEMBL173386	Training
14	CHEMBL173409	Training
14	CHEMBL173494	Training
14	CHEMBL174764	Training
14	CHEMBL367019	Training
14	CHEMBL369053	Training
14	CHEMBL369607	Training
14	CHEMBL172490	Training
14	CHEMBL175332	Training
14	CHEMBL263029	Training
14	CHEMBL367311	Training
14	CHEMBL368371	Training
14	CHEMBL435715	Training
14	CHEMBL197496	Training
14	CHEMBL198529	Training
15	CHEMBL77502	Training
15	CHEMBL25259	Training
15	CHEMBL312717	Training
15	CHEMBL82602	Training
15	CHEMBL81676	Training
15	CHEMBL321634	Training
15	CHEMBL79426	Training

15	CHEMBL264385	Training
15	CHEMBL66206	Training
15	CHEMBL78529	Training
15	CHEMBL189712	Training
15	CHEMBL309882	Training
15	CHEMBL222121	Training
15	CHEMBL363018	Training
15	CHEMBL2204610	Training
15	CHEMBL2237301	Training
15	CHEMBL2037007	Training
16	CHEMBL230304	Training
16	CHEMBL230305	Training
16	CHEMBL230409	Training
16	CHEMBL230410	Training
16	CHEMBL230411	Training
16	CHEMBL230517	Training
16	CHEMBL230518	Training
16	CHEMBL230622	Training
16	CHEMBL230623	Training
16	CHEMBL230705	Training
16	CHEMBL230706	Training
16	CHEMBL230707	Training
16	CHEMBL230814	Training
16	CHEMBL231425	Training
16	CHEMBL231523	Training
16	CHEMBL388008	Training
16	CHEMBL388422	Training
16	CHEMBL388645	Training
16	CHEMBL389297	Training
16	CHEMBL389298	Training
17	CHEMBL4463341	Training
17	CHEMBL4245047	Training
17	CHEMBL4437601	Training
17	CHEMBL4443452	Training
17	CHEMBL4448458	Training
17	CHEMBL4448658	Training
17	CHEMBL4454550	Training
17	CHEMBL4455529	Training
17	CHEMBL4455560	Training
17	CHEMBL4465958	Training
17	CHEMBL4466987	Training
17	CHEMBL4468237	Training
17	CHEMBL4468359	Training
17	CHEMBL4540905	Training
17	CHEMBL4550448	Training
17	CHEMBL4561177	Training
17	CHEMBL4564731	Training
17	CHEMBL4591150	Training
17	CHEMBL4592059	Training
18	CHEMBL1089176	Training
18	CHEMBL253952	Training
19	CHEMBL3398436	Training
19	CHEMBL3398444	Training
19	CHEMBL3398438	Training

19	CHEMBL3398443	Training
19	CHEMBL3398439	Training
19	CHEMBL3398445	Training
19	CHEMBL3398459	Training
19	CHEMBL257164	Training
19	CHEMBL3398440	Training
19	CHEMBL3398442	Training
19	CHEMBL3398455	Training
19	CHEMBL3398456	Training
19	CHEMBL3398458	Training
19	CHEMBL3398460	Training
19	CHEMBL3398461	Training
19	CHEMBL3398462	Training
19	CHEMBL3398463	Training
20	CHEMBL454678	Training
20	CHEMBL455157	Training
20	CHEMBL457062	Training
20	CHEMBL460278	Training
20	CHEMBL462554	Training
20	CHEMBL462580	Training
20	CHEMBL470322	Training
20	CHEMBL471533	Training
20	CHEMBL472399	Training
20	CHEMBL472571	Training
20	CHEMBL473300	Training
20	CHEMBL513035	Training
20	CHEMBL513594	Training
20	CHEMBL515667	Training
20	CHEMBL516824	Training
20	CHEMBL516964	Training
20	CHEMBL518348	Training
21	CHEMBL87307	Training
21	CHEMBL87380	Training
21	CHEMBL421729	Training
21	CHEMBL420973	Training
21	CHEMBL144489	Training
21	CHEMBL245454	Training
21	CHEMBL246279	Training
21	CHEMBL246485	Training
21	CHEMBL247110	Training
21	CHEMBL247111	Training
21	CHEMBL247314	Training
21	CHEMBL247937	Training
21	CHEMBL248112	Training
21	CHEMBL248113	Training
21	CHEMBL248318	Training
21	CHEMBL396154	Training
21	CHEMBL397331	Training
22	CHEMBL373874	Training
22	CHEMBL483351	Training
22	CHEMBL508299	Training
22	CHEMBL484329	Training
22	CHEMBL521131	Training
23	CHEMBL1090188	Training

23	CHEMBL1089970	Training
24	CHEMBL401616	Training
24	CHEMBL270835	Training
24	CHEMBL270961	Training
24	CHEMBL271394	Training
24	CHEMBL271775	Training
24	CHEMBL271776	Training
24	CHEMBL271873	Training
24	CHEMBL272175	Training
24	CHEMBL272466	Training
24	CHEMBL273035	Training
24	CHEMBL273036	Training
24	CHEMBL407950	Training
24	CHEMBL410704	Training
24	CHEMBL410705	Training
24	CHEMBL447533	Training
24	CHEMBL513145	Training
25	CHEMBL217907	Training
25	CHEMBL220533	Training
25	CHEMBL221123	Training
25	CHEMBL221184	Training
25	CHEMBL221474	Training
25	CHEMBL221525	Training
25	CHEMBL221641	Training
25	CHEMBL221649	Training
25	CHEMBL374448	Training
25	CHEMBL386957	Training
25	CHEMBL425639	Training
25	CHEMBL426550	Training
25	CHEMBL221388	Training
25	CHEMBL218487	Training
25	CHEMBL220655	Training
25	CHEMBL221489	Training
26	CHEMBL232597	Training
26	CHEMBL244573	Training
26	CHEMBL244574	Training
26	CHEMBL244575	Training
26	CHEMBL395560	Training
26	CHEMBL396496	Training
26	CHEMBL428037	Training
26	CHEMBL244786	Training
26	CHEMBL258832	Training
26	CHEMBL259055	Training
26	CHEMBL259856	Training
26	CHEMBL261555	Training
26	CHEMBL408941	Training
26	CHEMBL409355	Training
26	CHEMBL409513	Training
26	CHEMBL411045	Training
27	CHEMBL219586	Training
27	CHEMBL221074	Training
27	CHEMBL218094	Training
27	CHEMBL218637	Training
27	CHEMBL218687	Training

27	CHEMBL218691	Training
27	CHEMBL218784	Training
27	CHEMBL219587	Training
27	CHEMBL219626	Training
27	CHEMBL220281	Training
27	CHEMBL220926	Training
27	CHEMBL221201	Training
27	CHEMBL375760	Training
27	CHEMBL375761	Training
27	CHEMBL376641	Training
27	CHEMBL3354989	Training
28	CHEMBL549723	Training
28	CHEMBL550332	Training
28	CHEMBL551420	Training
28	CHEMBL551479	Training
28	CHEMBL551884	Training
28	CHEMBL552347	Training
28	CHEMBL557342	Training
28	CHEMBL558072	Training
28	CHEMBL558141	Training
28	CHEMBL560205	Training
28	CHEMBL560484	Training
28	CHEMBL561544	Training
28	CHEMBL563351	Training
28	CHEMBL563546	Training
28	CHEMBL564017	Training
29	CHEMBL201880	Training
29	CHEMBL201036	Training
29	CHEMBL437479	Training
29	CHEMBL216395	Training
29	CHEMBL384746	Training
29	CHEMBL414868	Training
29	CHEMBL446489	Training
29	CHEMBL451216	Training
29	CHEMBL454508	Training
29	CHEMBL492324	Training
29	CHEMBL492534	Training
29	CHEMBL492535	Training
29	CHEMBL494142	Training
29	CHEMBL506730	Training
29	CHEMBL522366	Training
30	CHEMBL189110	Training
30	CHEMBL190334	Training
30	CHEMBL192609	Training
30	CHEMBL363780	Training
30	CHEMBL425725	Training
30	CHEMBL426269	Training
30	CHEMBL190976	Training
31	CHEMBL94306	Training
31	CHEMBL96143	Training
31	CHEMBL96451	Training
31	CHEMBL318684	Training
31	CHEMBL328615	Training
31	CHEMBL420441	Training

31	CHEMBL117738	Training
31	CHEMBL119496	Training
31	CHEMBL122418	Training
31	CHEMBL2037009	Training
31	CHEMBL2037010	Training
31	CHEMBL1783034	Training
31	CHEMBL2037181	Training
31	CHEMBL2037183	Training
32	CHEMBL2204593	Training
32	CHEMBL2204594	Training
32	CHEMBL2204595	Training
32	CHEMBL2204596	Training
32	CHEMBL2204597	Training
32	CHEMBL2204598	Training
32	CHEMBL2204599	Training
32	CHEMBL2204600	Training
32	CHEMBL2204602	Training
32	CHEMBL2204603	Training
32	CHEMBL2204604	Training
32	CHEMBL2204605	Training
33	CHEMBL1934476	Training
33	CHEMBL1934484	Training
33	CHEMBL1934310	Training
33	CHEMBL1934312	Training
33	CHEMBL1934313	Training
33	CHEMBL1934317	Training
33	CHEMBL1934472	Training
33	CHEMBL1934480	Training
33	CHEMBL1934481	Training
33	CHEMBL1934482	Training
33	CHEMBL1934483	Training
33	CHEMBL1934485	Training
33	CHEMBL1934486	Training
34	CHEMBL1087587	Training
34	CHEMBL1088098	Training
34	CHEMBL1088099	Training
34	CHEMBL1088114	Training
34	CHEMBL1088246	Training
34	CHEMBL1088257	Training
34	CHEMBL1088515	Training
35	CHEMBL205665	Training
35	CHEMBL203409	Training
35	CHEMBL424674	Training
35	CHEMBL424685	Training
35	CHEMBL202739	Training
35	CHEMBL202973	Training
35	CHEMBL203338	Training
35	CHEMBL203408	Training
35	CHEMBL203621	Training
35	CHEMBL205803	Training
35	CHEMBL206202	Training
35	CHEMBL206759	Training
35	CHEMBL382178	Training
36	CHEMBL255388	Training

36	CHEMBL256468	Training
36	CHEMBL257517	Training
36	CHEMBL271240	Training
36	CHEMBL271739	Training
36	CHEMBL272336	Training
36	CHEMBL272962	Training
36	CHEMBL403847	Training
36	CHEMBL404451	Training
36	CHEMBL404646	Training
36	CHEMBL411118	Training
36	CHEMBL429734	Training
37	CHEMBL253416	Training
37	CHEMBL253417	Training
37	CHEMBL253851	Training
37	CHEMBL254054	Training
37	CHEMBL398677	Training
37	CHEMBL399629	Training
37	CHEMBL252996	Training
37	CHEMBL537921	Training
37	CHEMBL554562	Training
37	CHEMBL556643	Training
37	CHEMBL562659	Training
37	CHEMBL565198	Training
38	CHEMBL3901479	Training
38	CHEMBL3905544	Training
38	CHEMBL3920527	Training
38	CHEMBL3957862	Training
38	CHEMBL3981588	Training
38	CHEMBL4093975	Training
38	CHEMBL4444610	Training
38	CHEMBL4454087	Training
38	CHEMBL4475061	Training
38	CHEMBL4528723	Training
38	CHEMBL4558447	Training
39	CHEMBL589256	Training
39	CHEMBL591055	Training
39	CHEMBL591347	Training
39	CHEMBL591348	Training
39	CHEMBL591533	Training
39	CHEMBL591534	Training
39	CHEMBL592234	Training
39	CHEMBL592235	Training
39	CHEMBL592236	Training
39	CHEMBL603612	Training
39	CHEMBL606462	Training
40	CHEMBL258766	Training
40	CHEMBL259973	Training
40	CHEMBL260839	Training
40	CHEMBL260840	Training
40	CHEMBL409215	Training
41	CHEMBL1813003	Training
41	CHEMBL1813004	Training
41	CHEMBL1813005	Training
41	CHEMBL1813006	Training

41	CHEMBL1813007	Training
41	CHEMBL1813010	Training
41	CHEMBL1813011	Training
42	CHEMBL469937	Training
42	CHEMBL470144	Training
42	CHEMBL1956364	Training
42	CHEMBL2375947	Training
42	CHEMBL2375948	Training
43	CHEMBL200840	Training
43	CHEMBL383986	Training
43	CHEMBL434945	Training
43	CHEMBL246203	Training
43	CHEMBL246204	Training
43	CHEMBL246205	Training
43	CHEMBL246206	Training
43	CHEMBL246208	Training
43	CHEMBL248261	Training
43	CHEMBL400309	Training
44	CHEMBL187557	Training
44	CHEMBL189066	Training
44	CHEMBL190938	Training
44	CHEMBL188751	Training
44	CHEMBL191235	Training
44	CHEMBL191276	Training
44	CHEMBL364467	Training
44	CHEMBL370021	Training
44	CHEMBL371195	Training
44	CHEMBL426451	Training
45	CHEMBL253109	Training
45	CHEMBL253310	Training
45	CHEMBL253520	Training
45	CHEMBL254744	Training
45	CHEMBL254946	Training
45	CHEMBL402100	Training
45	CHEMBL403637	Training
45	CHEMBL404467	Training
45	CHEMBL427890	Training
46	CHEMBL21241	Training
46	CHEMBL20392	Training
46	CHEMBL20280	Training
46	CHEMBL20495	Training
46	CHEMBL21105	Training
46	CHEMBL21706	Training
46	CHEMBL281791	Training
46	CHEMBL2088421	Training
46	CHEMBL2088422	Training
47	CHEMBL3103200	Training
47	CHEMBL3103193	Training
47	CHEMBL3103197	Training
47	CHEMBL3103198	Training
47	CHEMBL3103194	Training
47	CHEMBL3103195	Training
47	CHEMBL3103196	Training
47	CHEMBL3103199	Training

48	CHEMBL454651	Training
48	CHEMBL455158	Training
48	CHEMBL461145	Training
48	CHEMBL461533	Training
48	CHEMBL509680	Training
48	CHEMBL516510	Training
49	CHEMBL363006	Training
50	CHEMBL201581	Training
50	CHEMBL201653	Training
50	CHEMBL425766	Training
50	CHEMBL202617	Training
51	CHEMBL1080470	Training
51	CHEMBL1081354	Training
51	CHEMBL1081355	Training
51	CHEMBL1086861	Training
52	CHEMBL4448919	Training
53	CHEMBL4062472	Training
53	CHEMBL4072395	Training
53	CHEMBL4074391	Training
53	CHEMBL4093146	Training
53	CHEMBL4094877	Training
53	CHEMBL4101596	Training
53	CHEMBL4103575	Training
54	CHEMBL3318038	Training
54	CHEMBL3318039	Training
54	CHEMBL3318040	Training
54	CHEMBL3318041	Training
54	CHEMBL518038	Training
54	CHEMBL3318037	Training
54	CHEMBL456792	Training
55	CHEMBL1630380	Training
55	CHEMBL1630381	Training
55	CHEMBL1630382	Training
55	CHEMBL1630383	Training
55	CHEMBL1630384	Training
55	CHEMBL1630386	Training
56	CHEMBL1083110	Training
56	CHEMBL1084320	Training
56	CHEMBL1084321	Training
56	CHEMBL1085062	Training
56	CHEMBL1085795	Training
56	CHEMBL1085796	Training
56	CHEMBL1086515	Training
57	CHEMBL116671	Training
57	CHEMBL443910	Training
57	CHEMBL455853	Training
57	CHEMBL476197	Training
57	CHEMBL477644	Training
57	CHEMBL514568	Training
57	CHEMBL515666	Training
58	CHEMBL473952	Training
58	CHEMBL474149	Training
58	CHEMBL474157	Training
58	CHEMBL491142	Training

58	CHEMBL514328	Training
58	CHEMBL474148	Training
58	CHEMBL514007	Training
59	CHEMBL481551	Training
59	CHEMBL467349	Training
59	CHEMBL468160	Training
59	CHEMBL512084	Training
59	CHEMBL512789	Training
60	CHEMBL461571	Training
60	CHEMBL549746	Training
60	CHEMBL552240	Training
60	CHEMBL560830	Training
60	CHEMBL562864	Training
60	CHEMBL564036	Training
60	CHEMBL564902	Training
61	CHEMBL239352	Training
61	CHEMBL239353	Training
61	CHEMBL239561	Training
61	CHEMBL240437	Training
61	CHEMBL240459	Training
61	CHEMBL392474	Training
61	CHEMBL430182	Training
62	CHEMBL375370	Training
63	CHEMBL119911	Training
64	CHEMBL4445084	Training
64	CHEMBL4466709	Training
64	CHEMBL4471421	Training
64	CHEMBL4474407	Training
64	CHEMBL4571498	Training
64	CHEMBL4586557	Training
65	CHEMBL3923046	Training
65	CHEMBL3945367	Training
65	CHEMBL3948805	Training
65	CHEMBL3957467	Training
65	CHEMBL3962832	Training
65	CHEMBL4110590	Training
66	CHEMBL3398446	Training
66	CHEMBL3398449	Training
66	CHEMBL3398451	Training
67	CHEMBL2206302	Training
67	CHEMBL2206303	Training
67	CHEMBL2206304	Training
68	CHEMBL1927473	Training
68	CHEMBL2035057	Training
68	CHEMBL2035068	Training
68	CHEMBL2035069	Training
68	CHEMBL2035070	Training
68	CHEMBL2035071	Training
69	CHEMBL1087471	Training
69	CHEMBL1087590	Training
69	CHEMBL1087593	Training
69	CHEMBL1087594	Training
70	CHEMBL20125	Training
70	CHEMBL279211	Training

70	CHEMBL281608	Training
71	CHEMBL230519	Training
72	CHEMBL278590	Training
72	CHEMBL248906	Training
72	CHEMBL124842	Training
72	CHEMBL125230	Training
72	CHEMBL340011	Training
72	CHEMBL435278	Training
73	CHEMBL87928	Training
73	CHEMBL89996	Training
74	CHEMBL110743	Training
74	CHEMBL111271	Training
75	CHEMBL220597	Training
75	CHEMBL221850	Training
75	CHEMBL222117	Training
75	CHEMBL374296	Training
76	CHEMBL220708	Training
77	CHEMBL237519	Training
77	CHEMBL237307	Training
77	CHEMBL236876	Training
77	CHEMBL237085	Training
77	CHEMBL237518	Training
77	CHEMBL392717	Training
78	CHEMBL206794	Training
78	CHEMBL203940	Training
78	CHEMBL205725	Training
78	CHEMBL377508	Training
78	CHEMBL380837	Training
79	CHEMBL107367	Training
79	CHEMBL147384	Training
79	CHEMBL356382	Training
79	CHEMBL2282523	Training
79	CHEMBL474295	Training
79	CHEMBL147095	Training
80	CHEMBL589747	Training
81	CHEMBL481683	Training
81	CHEMBL482334	Training
82	CHEMBL331469	Training
83	CHEMBL200495	Training
83	CHEMBL118887	Training
83	CHEMBL121206	Training
84	CHEMBL381477	Training
85	CHEMBL477095	Training
85	CHEMBL478725	Training
85	CHEMBL593796	Training
85	CHEMBL593801	Training
85	CHEMBL594034	Training
86	CHEMBL118756	Training
86	CHEMBL118251	Training
86	CHEMBL122417	Training
86	CHEMBL2037184	Training
87	CHEMBL4449760	Training
87	CHEMBL4453610	Training
87	CHEMBL4461949	Training

87	CHEMBL4471853	Training
88	CHEMBL4073499	Training
88	CHEMBL4092600	Training
88	CHEMBL4064900	Training
88	CHEMBL4100336	Training
89	CHEMBL3906547	Training
89	CHEMBL3916538	Training
89	CHEMBL3919950	Training
89	CHEMBL3982893	Training
90	CHEMBL1934471	Training
90	CHEMBL1934473	Training
91	CHEMBL1934314	Training
92	CHEMBL472071	Training
92	CHEMBL485155	Training
92	CHEMBL485179	Training
92	CHEMBL520508	Training
93	CHEMBL468159	Training
94	CHEMBL233003	Training
94	CHEMBL234819	Training
94	CHEMBL393381	Training
94	CHEMBL394031	Training
95	CHEMBL446893	Training
95	CHEMBL147721	Training
95	CHEMBL345009	Training
96	CHEMBL86658	Training
96	CHEMBL355695	Training
96	CHEMBL166637	Training
96	CHEMBL166628	Training
97	CHEMBL117796	Training
97	CHEMBL118577	Training
97	CHEMBL118816	Training
97	CHEMBL324027	Training
98	CHEMBL491750	Training
98	CHEMBL491751	Training
99	CHEMBL25710	Training
99	CHEMBL23874	Training
99	CHEMBL1783033	Training
100	CHEMBL4519542	Training
100	CHEMBL4535730	Training
100	CHEMBL4593987	Training
101	CHEMBL4464538	Training
101	CHEMBL4519376	Training
102	CHEMBL4436335	Training
102	CHEMBL4470753	Training
102	CHEMBL4516239	Training
103	CHEMBL4454999	Training
103	CHEMBL4457812	Training
103	CHEMBL4543257	Training
104	CHEMBL4160945	Training
104	CHEMBL4167789	Training
104	CHEMBL4174996	Training
105	CHEMBL4060007	Training
106	CHEMBL1828826	Training
106	CHEMBL3752643	Training

106	CHEMBL3753965	Training
107	CHEMBL3601087	Training
107	CHEMBL3601089	Training
107	CHEMBL3601093	Training
108	CHEMBL2022403	Training
108	CHEMBL2022404	Training
108	CHEMBL2022405	Training
109	CHEMBL1935493	Training
109	CHEMBL1935495	Training
109	CHEMBL1935608	Training
110	CHEMBL1927474	Training
110	CHEMBL1927475	Training
111	CHEMBL1075895	Training
111	CHEMBL1080365	Training
111	CHEMBL1080699	Training
112	CHEMBL562786	Training
113	CHEMBL489178	Training
113	CHEMBL514544	Training
113	CHEMBL522754	Training
114	CHEMBL481549	Training
114	CHEMBL481550	Training
114	CHEMBL520272	Training
115	CHEMBL101123	Training
115	CHEMBL322690	Training
115	CHEMBL443218	Training
116	CHEMBL197684	Training
116	CHEMBL198163	Training
116	CHEMBL383607	Training
117	CHEMBL117402	Training
117	CHEMBL325844	Training
117	CHEMBL331388	Training
119	CHEMBL252675	Training
119	CHEMBL253486	Training
119	CHEMBL253923	Training
120	CHEMBL118329	Training
120	CHEMBL120243	Training
121	CHEMBL1090212	Training
121	CHEMBL1091664	Training
122	CHEMBL401534	Training
122	CHEMBL256142	Training
122	CHEMBL256358	Training
123	CHEMBL364378	Training
123	CHEMBL192521	Training
123	CHEMBL193082	Training
124	CHEMBL427530	Training
124	CHEMBL191149	Training
124	CHEMBL191556	Training
126	CHEMBL232395	Training
126	CHEMBL391479	Training
126	CHEMBL232396	Training
129	CHEMBL3899254	Training
129	CHEMBL3890301	Training
130	CHEMBL3601098	Training
130	CHEMBL3601099	Training

131	CHEMBL3335660	Training
131	CHEMBL3335661	Training
133	CHEMBL1778839	Training
133	CHEMBL1778840	Training
134	CHEMBL1711502	Training
134	CHEMBL1732818	Training
135	CHEMBL1644305	Training
135	CHEMBL1644310	Training
136	CHEMBL1644299	Training
136	CHEMBL1644301	Training
137	CHEMBL1079961	Training
137	CHEMBL1079962	Training
138	CHEMBL1075723	Training
138	CHEMBL1075732	Training
139	CHEMBL456912	Training
139	CHEMBL514713	Training
140	CHEMBL489768	Training
140	CHEMBL489769	Training
141	CHEMBL448996	Training
141	CHEMBL450050	Training
142	CHEMBL401560	Training
142	CHEMBL403909	Training
143	CHEMBL102004	Training
143	CHEMBL317424	Training
146	CHEMBL246209	Training
146	CHEMBL248411	Training
147	CHEMBL240868	Training
147	CHEMBL241269	Training
148	CHEMBL2113263	Training
148	CHEMBL220223	Training
149	CHEMBL214289	Training
149	CHEMBL214295	Training
151	CHEMBL495904	Training
151	CHEMBL498171	Training
152	CHEMBL234659	Training
152	CHEMBL391478	Training
153	CHEMBL234588	Training
153	CHEMBL234589	Training
154	CHEMBL255309	Training
154	CHEMBL403432	Training

Supplementary Table 3. SBVS performance of target-specific ML SFs trained on the users' own data sets ACHE and HMGR and tested on the corresponding DEKOIS 2.0 benchmarks (PETS option, see Table S1 for the performance of four generic SFs). The values indicated here are median values obtained after 10 training-test runs of each ML algorithm on each training-test partition. Detailed results for each run can be downloaded from <https://github.com/vktrannguyen/MLSF-protocol> (Supporting Information folder).

	ACHE	HMGR
RF:		
- EF1%	30.00	30.00
- NEF1%	1.000	1.000
XGB:		
- EF1%	10.00	30.00
- NEF1%	0.333	1.000
SVM:		
- EF1%	30.00	30.00
- NEF1%	1.000	1.000
ANN:		
- EF1%	5.00	30.00
- NEF1%	0.167	1.000
DNN:		
- EF1%	7.50	30.00
- NEF1%	0.250	1.000

Supplementary Table 4. SBVS performance of target-specific ML SFs trained on the users' own training sets ACHE, HMGR, PPARA and tested on the corresponding users' own test sets (OTS option). The values related to target-specific ML SFs are the medians obtained after 10 training-test runs of each ML algorithm on each training-test partition. Detailed results for each run can be downloaded from <https://github.com/vktrannguyen/MLSF-protocol> (Supporting Information folder). The performance of four generic SFs (Smina, IFP, CNN-Score, RF-Score-VS) is also reported for comparison (only one run is performed on each test set for each of these SFs).

	ACHE	HMGR	PPARA
- Number of actives	37	31	80
- Number of inactives/decoys	2170	2077	2399
- Total number of molecules	2207	2108	2479
- Maximal EF1%	59.46	67.74	30.00
RF:			
- EF1%	43.24	54.84	0.00
- NEF1%	0.727	0.810	0.000
XGB:			
- EF1%	45.95	51.61	1.25
- NEF1%	0.773	0.762	0.042
SVM:			
- EF1%	40.54	54.84	1.88
- NEF1%	0.682	0.810	0.063
ANN:			
- EF1%	37.84	58.06	2.50
- NEF1%	0.636	0.857	0.083
DNN:			
- EF1%	35.14	54.84	30.00
- NEF1%	0.591	0.810	1.000
Smina:			
- EF1%	8.11	3.23	1.25
- NEF1%	0.136	0.048	0.042
IFP:			
- EF1%	0.00	16.13	0.00
- NEF1%	0.000	0.238	0.000
CNN-Score:			
- EF1%	8.11	0.00	0.00
- NEF1%	0.136	0.000	0.000
RF-Score-VS:			
- EF1%	5.41	9.68	1.25
- NEF1%	0.091	0.143	0.042

Supplementary Table 5. SBVS performance of target-specific ML SFs issued from the five “Dissimilar” (D) training-test partitions. A “Dissimilar” partition includes a full training set for a target, and a part of its corresponding test set composed only of molecules whose similarity (Morgan fingerprints, 2048 bits, radius 2) to any training molecule is lower than 0.70. The values related to target-specific ML SFs are the medians obtained after 10 training-test runs of each ML algorithm on each training-test partition. The performance of four generic SFs (Smina, IFP, CNN-Score, RF-Score-VS) is also reported for comparison (only one run is performed on each test set for each of these SFs).

	ACHE-PETS-D	ACHE-OTS-D	HMGR-PETS-D	HMGR-OTS-D	PPARA-OTS-D
- Number of actives	39	12	8	7	65
- Number of inactives/decoys	1200	2136	1199	2026	1698
- Total number of molecules	1239	2148	1207	2033	1763
- Maximal EF1%	30.77	100.00	100.00	100.00	26.15
RF:					
- EF1%	30.77	41.67	87.50	71.43	1.54
- NEF1%	1.000	0.417	0.875	0.714	0.059
XGB:					
- EF1%	7.69	41.67	62.50	71.43	0.00
- NEF1%	0.250	0.417	0.625	0.714	0.000
SVM:					
- EF1%	30.77	41.67	87.50	71.43	1.54
- NEF1%	1.000	0.417	0.875	0.714	0.059
ANN:					
- EF1%	5.13	33.33	100.00	71.43	3.08
- NEF1%	0.167	0.333	1.000	0.714	0.118
DNN:					
- EF1%	5.13	33.33	68.75	71.43	26.15
- NEF1%	0.167	0.333	0.688	0.714	1.000
Smina:					
- EF1%	15.38	16.67	0.00	0.00	1.54
- NEF1%	0.500	0.167	0.000	0.000	0.059
IFP:					
- EF1%	15.38	0.00	37.50	14.29	1.54
- NEF1%	0.500	0.000	0.375	0.143	0.059
CNN-Score:					
- EF1%	0.00	16.67	25.00	0.00	0.00
- NEF1%	0.000	0.167	0.250	0.000	0.000
RF-Score-VS:					
- EF1%	2.56	0.00	62.50	28.57	0.00
- NEF1%	0.083	0.000	0.625	0.286	0.000

Supplementary Discussion. Comparing the ligand binding sites of the ACHE human-extracted and *Tetronarce californica*-extracted variants.

The UniProt ID of the ACHE human-extracted variant is P22303 (representative PDB ID: 4M0E).

The UniProt ID of the ACHE *Tetronarce californica*-extracted variant is P04058 (representative PDB ID: 1E66).

The binding site of each PDB structure is composed of all amino acid residues that contain at least one heavy atom found within 5 Å from any heavy atom of the co-crystallized ligand.

The “Align” tool of UniProt is used to align the two amino acid sequences that correspond to the two binding sites of P22303 and P04058. The aligned sequences are as follows (the different amino acids are marked in bold red):

P22303 (4M0E)	P04058 (1E66)
D74	D72
G82	G80
T83	S81
W86	W84
G120	G117
G121	G118
G122	G119
Y124	Y121
S125	S122
Y133	Y130
E202	E199
S203	S200
F297	F290
Y337	F330
F338	F331
Y341	Y334
W439	W432
M443	M436
P446	I439
H447	H440
G448	G441
Y449	Y442

There are 19 (out of 22) similar positions after sequence alignment. The percentage sequence identity for the binding sites is therefore 86.36%.