Identification of Novel Androgen Receptor Antagonists Using Structure- and Ligand-based Methods

Huifang Li, Xin Ren, Eric Leblanc, Kate Frewin, Paul S. Rennie and Artem Cherkasov

Content:

Supplementary Figure S1. Structures of steroids and current non-steroidal anti-androgens. Red color marks the common substructure of the compounds that is responsible for the AR HBS target binding.

Supplementary Figure S2. Structures of 37 selected compounds from the structure- and ligand-based virtual screening for biological testing

Supplementary Table S1. *ROC AUC of different methods based on different number of descriptors*

Supplementary Table S2. *Descriptors and their corresponding meanings*

Figure S1

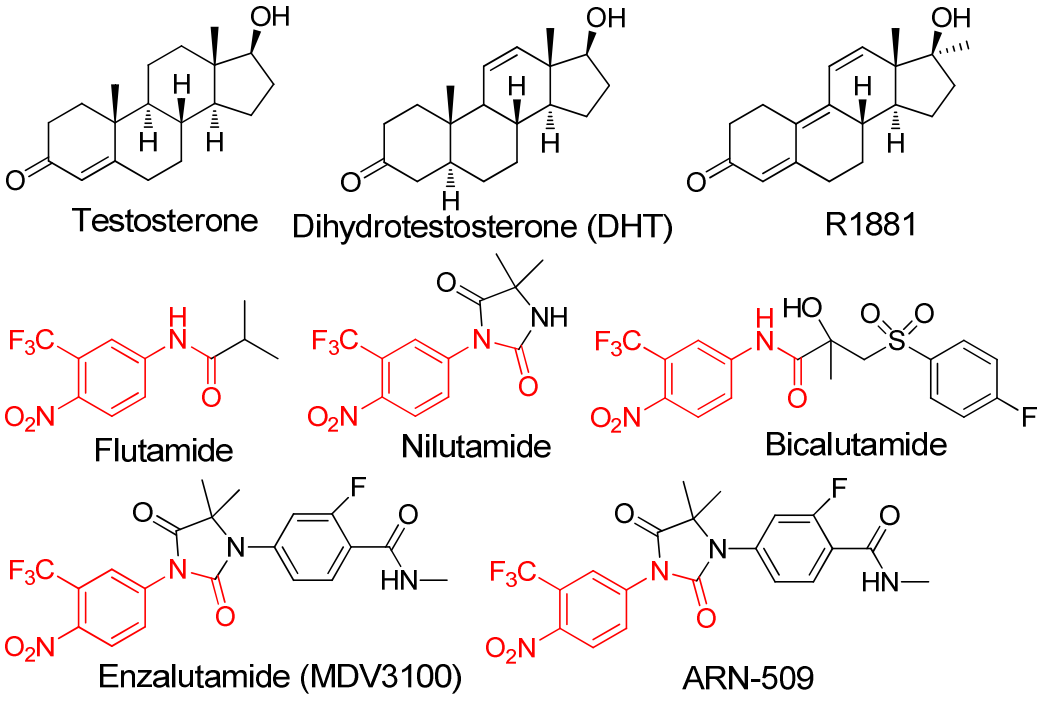

Figure S2

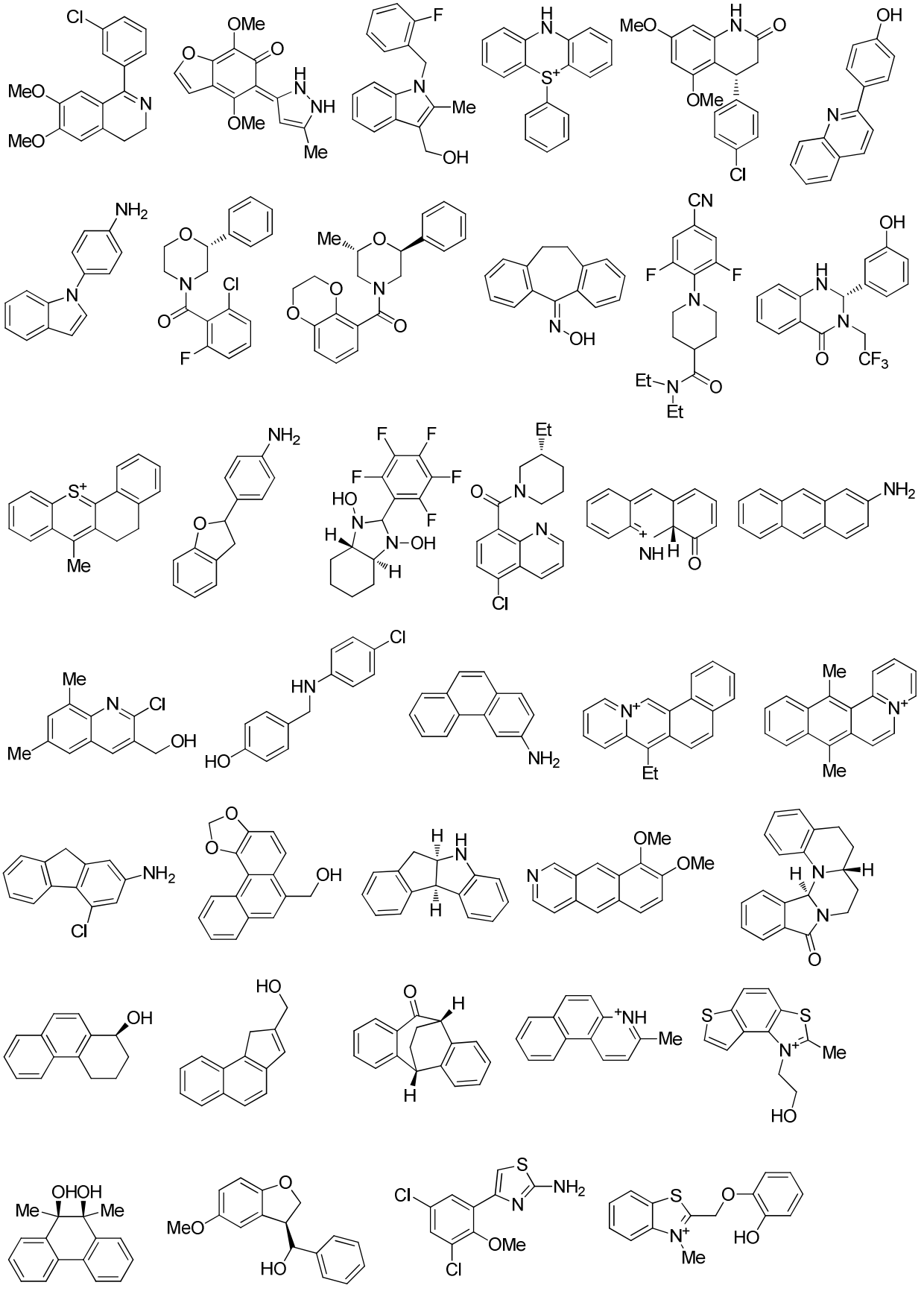

Table S1. ROC AUC of different methods based on different number of descriptors

Desc. No.	ADTree	ANN	Bagging	Decorate	kNN	K-star	Lazy-IB1	Logitboost	RandomForest
15	0.705	0.731	0.787	0.785	0.788	0.783	0.711	0.777	0.784
16	0.702	0.712	0.781	0.787	0.787	0.779	0.718	0.782	0.799
17	0.710	0.73	0.797	0.789	0.793	0.785	0.717	0.778	0.802
18	0.695	0.745	0.800	0.792	0.8	0.794	0.722	0.790	0.810
19	0.699	0.746	0.804	0.799	0.798	0.790	0.713	0.776	0.809
20	0.703	0.746	0.808	0.777	0.801	0.793	0.716	0.781	0.806
21	0.700	0.767	0.806	0.776	0.804	0.796	0.715	0.787	0.802
22	0.705	0.759	0.807	0.784	0.806	0.798	0.72	0.794	0.804
23	0.716	0.747	0.814	0.802	0.802	0.805	0.717	0.791	0.814
24	0.722	0.769	0.813	0.812	0.806	0.807	0.716	0.804	0.812
25	0.715	0.737	0.811	0.808	0.804	0.809	0.716	0.807	0.822
26	0.717	0.736	0.807	0.813	0.812	0.807	0.712	0.795	0.816
27	0.710	0.754	0.823	0.793	0.805	0.807	0.72	0.807	0.824
28	0.712	0.767	0.822	0.796	0.801	0.808	0.711	0.795	0.817
29	0.721	0.75	0.818	0.800	0.803	0.809	0.713	0.796	0.813
30	0.724	0.752	0.818	0.805	0.803	0.809	0.715	0.802	0.815
best	0.724	0.769	0.823	0.813	0.812	0.809	0.722	0.807	0.824

Table S2. Descriptors and their corresponding meanings

Symbol	Definition	Class
a_nN	Number of nitrogen atoms	2D
ARR	aromatic ratio	constitutional descriptors
Average_Softness	Arithmetic mean of softnesses of all atoms of a molecule	IND
b_double	Number of double bonds. Aromatic bonds are not considered to be double bonds.	2D
B02[N-O]	presence/absence of N - O at topological distance 02	2D binary fingerprints

B02[O-O]	presence/absence of O - O at topological distance 02	2D binary fingerprints
B09[C-O]	presence/absence of C - O at topological distance 09	2D binary fingerprints
BCUT_SLOGP_0	LogP BCUT (0/3)	2D
BCUT_SMR_0	Molar Refractivity BCUT (0/3)	2D
EEig09r	Eigenvalue 09 from edge adj. matrix weighted by resonance integrals	edge adjacency indices
ESpm08d	Spectral moment 08 from edge adj. matrix weighted by dipole moments	edge adjacency indices
F04[C-C]	frequency of C - C at topological distance 04	2D binary fingerprints
F05[O-F]	frequency of O - F at topological distance 05	2D binary fingerprints
GCUT_SLOGP_3	logP GCUT (3/3)	2D
Kier2	Second kappa shape index	2D
MSD	mean square distance index (Balaban)	topological descriptors
nCb-	number of substituted benzene C(sp2)	functional group counts
nCconj	number of non-aromatic conjugated C(sp2)	functional group counts
nDB	number of double bonds	constitutional descriptors
PCR	ratio of multiple path count over path count	walk and path counts
PEOE_VSA_PNEG	Total polar negative VDW surface area	2D
R7u	R autocorrelation of lag 7 / unweighted	GETAWAY descriptors
SlogP	Log Octanol/Water Partition Coefficient	2D
SMR	Molar Refractivity	2D
SRW05	self-returning walk count of order 05	walk and path counts
std_dim1	Standard dimension 1	i3D
Sum_Hardness	Sum of hardnesses of atoms of a molecule	IND
vsurf_A	Ampiphilic moment	i3D
vsurf_CW1	Capacity factor at -0.2	i3D
vsurf_IW6	Hydrophilic integy moment at -4.0	i3D