

Supporting Information

Using Molecular Initiating Events to Generate 2D Structure Activity Relationships for Toxicity Screening

Timothy E. H. Allen, Sonia Liggi, Jonathan M. Goodman, Steve Gutsell, and Paul J. Russell

CONTENTS

SMARTS/SMILES OF STRUCTURAL ALERTS S2

SMARTS/SMILES OF STRUCTURAL ALERTS

Receptor	MIE	SMARTS/SMILES
G-Protein Coupled Receptors		
Adenosine A2a Receptor		
	Alert AA2aR 1 - 2-Pyrimidine	<chem>c1cnc(nc1)N</chem>
	Alert AA2aR 2 - 4-Pyrimidine	<chem>c1cncnc1N</chem>
	Alert AA2aR 3 - Adenine	<chem>n1cnc2c1ncnc2N</chem>
	Alert AA2aR 4 - AA2aR Frag 1	<chem>c1cnc(N)n2c1nc(c3ccco3)n2</chem>
Adrenergic Receptors		
Alpha	Alert A-ADR 1 - Phenylpiperazine-like	<chem>a1aaaaa1N2CCNCC2</chem>
	Alert A-ADR 2 - Tolazoline-like	<chem>a1aaaaa1AC2=NCCN2</chem>
Alpha-1a	Alert A-1aADR 1 - Phenylethanolamin-like	<chem>a1aaaaa1C(O)CN</chem>
	Alert A-1aADR 2 - 1-Ethyl-4-phenylpiperazine-like	<chem>a1aaaaa1N2CCN(CC)CC2</chem>
Alpha-2a	Alert A-2aADR 1 - CID 5145436	<chem>OCC1=NCCN1</chem>
	Alert A-2aADR 2 - CID 13001	<chem>CCC1=NCCN1</chem>
	Alert A-2aADR 3 - A2aADR Frag 2	<chem>CCCNc1nc(N)ccn1</chem>
	Alert A-2aADR 4 - 4-Benzyl-1H-imidazole-like	<chem>a1aaaaa1Ac2cncn2</chem>
	Alert A-2aADR 5- Guanidine	<chem>N=C(N)N</chem>
	Alert A-2aADR 6 - 2-(1-Hydroxyl ethyl)-2-imimidazole	<chem>CC(O)C1=NCCN1</chem>
Beta	Alert B-ADR 1 - 1-(Phenethylamino)propan-2-ol-like	<chem>a1aaaaa1CCNCC(O)*</chem>
	Alert B-ADR 2 - 2-(Ethylamino)-1-phenylethanol-like	<chem>a1aaaaa1C(O)CNCC</chem>
	Alert B-ADR 3 - 1-(Ethylamino)-3-phenoxy-2-propanol-like	<chem>a1aaaaa1OCC(O)CNCC</chem>

Receptor	MIE	SMARTS/SMILES
Cannabinoid Receptor		
CB1	Alert CCB1R 1 - 2-Phenylpyrazole-like	a1aaaaa1n2nccc2
	Alert CCB1R 2 - 5-Phenyl-1H-pyrazole-like	a1aaaaa1c2ccnn2
	Alert CCB1R 3 - Pyrazole-3-carboxamide	NC(=O)c1ccnn1
CB2	Alert CCB2R 1 - CCB2R Frag 1	CCCCNC(=O)C(=C)C=O
	Alert CCB2R 2 - 1-(1-Ethyl-1H-pyrrol-3-yl)ethanone	Cn1ccc(c1)C(=O)C
CB1 & CB2	Alert CR 1 - Indole	c1cccc2c1ccn2
Cholecystokinin A Receptor		
	Alert CCKAR 1 - 2-Acetamido-N-methylacetamide	CNC(=O)CNC(=O)C
	Alert CCKAR 2 - N-[2-(methylamino)ethyl]acetamide	CNCCNC(=O)C
	Alert CCKAR 3 - CID 9957635	CN1CCN(CC1)C(=O)c2nc(c3ccccc3)n(c4ccccc4)c2
Dopamine Receptors		
D1	Alert DD1R 1 - Dihydroxine-like	a1aaaa1*c2cc(O)c(O)cc2
		OR a1aaaaa1*c2cc(O)c(O)cc2
	Alert DD1R 2 - Benzazepine-like	c1cccc2c1CCNCC2
	Alert DD1R 3 - Benzazepine-like with aromatic	c1cccc2c1C(a3aaaaa3)CNCC2
	Alert DD1R 4 - CID 15288	c1cccc2c1nc3ccccc3nc2N4CCN(C)CC4
D2	Alert DD2R 1 - Piperazine	N1CCNCC1
	Alert DD2R 2 - 1,4 Dimethylpiperazine	CN1CCN(C)CC1
	Alert DD2R 3 - Phenylpiperazine	c1cccc1N2CCNCC2

Receptor	MIE	SMARTS/SMILES
Endothelin Receptor A		
	Alert ER ET-A 1 - 5-Ethyl-1,3-benzodioxole	<chem>CCc1ccc2c(c1)OCO2</chem>
	Alert ER ET-A 2 - N-Methylbenzenesulfonamide-like	<chem>a1aaaaa1S(=O)(=O)N*</chem>
	Alert ER ET-A 3 - Tryptamine	<chem>NCCc1cnc2c1cccc2</chem>
	Alert ER ET-A 4 - ER ET-A Frag 1-like	<chem>a1aaaaa1S(=O)(=O)Nc2ccno2</chem>
	Alert ER ET-A 5 - N-Methyl-2-biphenylsulfonamide-like	<chem>a1aaaaa1a2aaaaa2S(=O)(=O)N*</chem>
Histamine Receptors		
H1	Alert HH1R 1 - Doxepine-like	<chem>a1aaaaa1*(a2aaaaa2)***NC</chem>
		OR <chem>a1aaaaa1*(a2aaaaa2)****NC</chem>
		OR <chem>a1aaaaa1*(a2aaaaa2)*****NC</chem>
	Alert HH1R 2 - 4-Phenoxypiperidine	<chem>c1ccccc1OC2CCNCC2</chem>
H2	Alert HH2R 1 - Imidazole	<chem>n1ccnc1</chem>
	Alert HH2R 2 - Guanidine	<chem>N=C(N)N</chem>
	Alert HH2R 4 - Indole	<chem>c1cccc2c1ccn2</chem>
Muscarinic Acetylcholine Receptors		
M1 and M2	Alert MAR 1 - Formanilide	<chem>c1ccccc1NC=O</chem>
M2 and M3	Alert MAR 2 - N-Ethyl-N,N-dimethylpropanaminum	<chem>CCCN(C)(C)CC</chem>
M3	Alert MAR 3 - Tetramethylamonium	<chem>CN(C)(C)C</chem>
All	Alert MAR 4 - MAR Pharmacophore	<chem>[N+]</chem>
		OR <chem>c1ccccc2Cc2ccccc2</chem>
		OR <chem>C(=O)CCO</chem>

Receptor	MIE	SMARTS/SMILES
Opioid Receptors		
	Alert OR 1 - Morphine-like (4 or more)	<chem>c1c2c3c(cc1)~*~*(~N(A)A)~*~*3~*(~[O,N]2)~*~[OH,NH]</chem> OR <chem>[OH,NH]~*~*~[O,N]~*(~*~[OH,NH])~*~*~*~N(A)A</chem> OR <chem>[OH,NH]c1c2c3c(cc1)~*~*~*~*3~*(~[O,N]2)~*~[OH,NH]</chem> OR <chem>[OH,NH]c1ccc(cc1)~*~*(~N(A)A)~*~*~*~*~[OH,NH]</chem> OR <chem>[OH,NH]c1c2c3c(cc1)~*~*(~N(A)A)~*~*3~*~[O,N]2</chem>
	Alert OR 1 - Morphine-like (3 or more)	<chem>[OH,NH]~*~*~[O,N]~*~*~[OH,NH]</chem> OR <chem>[OH,NH]~*~*~[O,N]~*~*~*~*~N(A)A</chem> OR <chem>[OH,NH]~*~*~*~*(~*~*~[OH,NH])~*~*~*~N(A)A</chem> OR <chem>[O,N]~*(~*~[OH,NH])~*~*~*~N(A)A</chem> OR <chem>c1c2c3c(cc1)~*~*~*~*3~*(~*~[OH,NH])~[O,N]2</chem> OR <chem>c1c2c3c(cc1)~*~*(~N(A)A)~*~*3~*~[O,N]2</chem> OR <chem>c1cc3c(cc1)~*~*(~N(A)A)~*~*3~*~*~[OH,NH]</chem> OR <chem>[OH,NH]c1c2c3c(cc1)~*~*~*~*3~*~[O,N]2</chem> OR <chem>[OH,NH]c1cc3c(cc1)~*~*~*~*3~*~*~[OH,NH]</chem> OR <chem>[OH,NH]c1cc3c(cc1)~*~*(~N(A)A)~*~*3</chem>
	Alert OR 2 - 1-Methyl-4-phenylpiperidine	<chem>c1ccccc1C2CCN(C)CC2</chem>
Serotonin Receptors		
All	Alert SR 1 - 3-Ethyl Indole	<chem>c1ccccc2c1c(CC)cn2</chem>
1A, 1B, 2A	Alert SR 1a1b2a 1 - 1-Methyl-4-Phenylpiperazine-like	<chem>a1aaaaa1N2CCN(C)CC2</chem>
1A	Alert SR 1a 1 - SR 1a Frag 2	<chem>NC(=O)c1cccc2c1CC(CO2)N(C)CCCC</chem>
1B	Alert SR 1b 1 - 2-Methyl-1,2-dihydroquinoline	<chem>CC1Nc2ccccc2C=C1</chem>
2B	Alert SR 2b 1 - CID 15206310 Alert SR 2b 2 - SR 2b Frag 2	<chem>c1ccccc2c1c(=O)N3C2CNCCC3</chem> <chem>c1ccccc2C3CC(C)=NN3C4(Oc12)CCN(C)CC4</chem>

Receptor	MIE	SMARTS/SMILES
Vasopressin V1a		
	Alert VV1a 1 - VV1a Frag 1	<chem>CC(=O)N(CC)CCC</chem>
	Alert VV1a 2 - n-benzyl-n-ethylmethylamine	<chem>c1ccccc1CN(CC)C</chem>
	Alert VV1a 3 - VV1a Frag 3	<chem>NC(=O)C(CC)NCC</chem>
	Alert VV1a 4 - VV1a Frag 4	<chem>CCNC(=O)C(CC)NC(=O)CN(C)C=O</chem>
Enzymes		
Acetylcholinesterase		
	Alert AC 1a - Tacrine	<chem>c1ccccc2c1c(N)c3c(n2)CCCC3</chem>
	Alert AC 1b - 4-Quinolinamine	<chem>c1ccccc2c1c(N)ccn2</chem>
	Alert AC 1c - AC Frag 1	<chem>c1cc(N)c2c(n1)CCCC2</chem>
	Alert AC 1d - Dimethyl benzylamine	<chem>c1ccccc1CN</chem>
	Alert AC 1e - Dimethyl benzylamine-like SMARTS	<chem>a1aaaaa1CN</chem>
Cyclooxygenases		
	Alert COX 1 - Sulfonated aromatic rings	<chem>a1aaaaa1S(=O)(=O)</chem>
	Alert COX 2 - Cinnamaldehyde-like	<chem>a1aaaaa1C=CC=O</chem>
	Alert COX 3 - 5-Phenyl-1H-pyrazole-like	<chem>a1aaaaa1c2ccnn2</chem>
Dihydrofolate Reductase		
	Alert DHFR 1 (SMILES) - Diaminopyrimidine	<chem>Nc1ccnc(N)n1</chem>
	Alert DHFR 1 (SMARTS) - Diaminopyrimidine-like	<chem>a1aaana1[NH,OH]</chem>
GAR Transformylase		
	Alert GART 1 - N-Acetyl-DL-glutamic acid	<chem>CC(=O)NC(C(=O)O)CCC(=O)O</chem>

Receptor	MIE	SMARTS/SMILES
Histone Deacetylase 1		
	Alert HDAC 1 1 - Hydroxamic acids	<chem>CC(=O)NO</chem>
	Alert HDAC 1 2a - Benzamide	<chem>c1ccccc1C(=O)N</chem>
	Alert HDAC 1 2b - Benzamide-like SMARTS	<chem>a1aaaaa1C(=O)N</chem>
	Alert HDAC 1 2c - Acetanilide	<chem>c1ccccc1NC(=O)C</chem>
	Alert HDAC 1 2d - Acetanilide-like SMARTS	<chem>a1aaaaa1NC(=O)C</chem>
	Alert HDAC 1 2e - Benanilide	<chem>c1ccccc1NC(=O)c2ccccc2</chem>
	Alert HDAC 1 2f - Benanilide-like SMARTS	<chem>a1aaaaa1NC(=O)a2aaaaa2</chem>
Monoamine Oxidase		
	Alert MAO A 1 - 3-Phenyl-2-oxazolidone	<chem>c1ccccc1N2C(=O)OCC2</chem>
	Alert MAO A 2a - 4-Methylphenol	<chem>Oc1ccc(cc1)C</chem>
	Alert MAO A 2b - 3-Methylphenol	<chem>Oc1cccc(c1)C</chem>
	Alert MAO A 2c - Methyl-3,4-diphenol	<chem>Cc1cc(O)c(O)cc1</chem>
	Alert MAO A 3a - Aminomethaziale	<chem>Cc1csc(N)n1</chem>
	Alert MAO A 3b - 4-Phenyl-1,3-thiazole-2-amine	<chem>c1ccccc1c2csc(N)n2</chem>
	Alert MAO A 3c - CID 7958070	<chem>c1ccccc1c2csc(NN=CC)n2</chem>
Phosphodiesterases		
3A	Alert PDE 3A 1 - PDE 3A Frag 2	<chem>COc1c(OC)cc2c(c1)ncnc2N3CCC(C3)OC</chem>
	Alert PDE 3A 2a - PDE Frag 3	<chem>COc1ccc(c2n1nc(C)c2)C3=NNC(=O)CC3C</chem>
	Alert PDE 3A 2b - PDE Frag 3-like SMARTS	<chem>COa1aaa(a2a1aa(C)a2)C3=NNC(=O)CC3C</chem>
	Alert PDE 3A 3 - PDE 3A Frag 4	<chem>CC=NNC(=O)CC</chem>
	Alert PDE 3A 4 - Veratrol-like	<chem>COa1aaaaa1OC</chem>
4D	Alert PDE 4D 1 - Homoveratrol-like	<chem>COa1aa(C)aaa1OC</chem>
	Alert PDE 4D 2 - 8-Isopropyl-1H-purine Frag	<chem>CC(C)c1nc2cncnc2n1</chem>
	Alert PDE 4D 3 - PDE 4D Frag 3	<chem>nC1=CCNC=N1</chem>

Receptor	MIE	SMARTS/SMILES
Thymidylate Synthase		
	Alert TS 1 - N-Acetyl-DL-glutamic acid	<chem>CC(=O)NC(C(=O)O)CCC(=O)O</chem>
	Alert TS 2 - Pyrimidine	<chem>c1ncccn1</chem>
	Alert TS 3 - Uracil	<chem>N1C=CC(=O)NC(=O)1</chem>
Tyrosine Protein Kinase		
	Alert LCK 1 - Pyrimidine-like	<chem>*c1ccncn1</chem>
Ion Channels		
Glutamate (NMDA) Receptor		
	Alerts GRSZ1 1 - GRSZ1 Frag 1	<chem>CS(=O)(=O)N(CC)c1cccc2c1NC(=O)C(=O)N2</chem>
Potassium Voltage Gated Channel Subunit H Member 2		
	Alert HERG 1a - Flex Aromatic Amine 1-5	<chem>a1aaaaa1~*~N(C)C</chem>
		OR <chem>a1aaaaa1~*~*~N(C)C</chem>
		OR <chem>a1aaaaa1~*~*~*~N(C)C</chem>
		OR <chem>a1aaaaa1~*~*~*~*~N(C)C</chem>
		OR <chem>a1aaaaa1~*~*~*~*~*~N(C)C</chem>
	Alert HERG 1b - Flex Aromatic Amine 2-5	<chem>a1aaaaa1~*~*~N(C)C</chem>
		OR <chem>a1aaaaa1~*~*~*~N(C)C</chem>
		OR <chem>a1aaaaa1~*~*~*~*~N(C)C</chem>
		OR <chem>a1aaaaa1~*~*~*~*~*~N(C)C</chem>
Serotonin 3A Receptor		
	Alert S3AR 1 - S3AR Frag 1	<chem>CC(=O)NC1CC2CCCC(C1)N2C</chem>
	Alert S3AR 2 - 1-[2-(Phenylsulfanyl)phenyl]piperazine	<chem>c1ccc(cc1)Sc2ccccc2N3CCNCC3</chem>

Receptor	MIE	SMARTS/SMILES
Sodium Channel V Subunit alpha		
	Alert SCV-A 1a - Pyrimidine-like aromatic	a1naaan1
	Alert SCV-A 1b - Pyrimidine-like wildcard	*1n***n1
	Alert SCV-A 2 - 4-(4-Piperidineoxy)pyrimidine	C1CNCCC2Oc2ccncn2
	Alert SCV-A 3 - SCV Frag 1	CCN1CCC(CC1)Oc2ncnc3c2ccn3Cc4cccc4
	Alert SCV-A 4 - N-(2-Phenylethyl)acetamide	CC(=O)NCCc1cccc1
Voltage Gated K Channel Sub Unit Kv7.1		
	Alert VGKC Kv7.1 1 - VGKV Kv7.1 Frag 1	c1cccc1c2csc(n2)NC(=O)C(NC)C
Nuclear Receptors		
Androgen Receptor		
	Alert AR 1a - Benzonitrile	c1cccc1C#N
	Alert AR 1b - Nitrobenzene	c1ccc(cc1)[N+](=O)[O-]
	Alert AR 1c - Quinolone	c1cccc2c1NC(=O)C=C2
Glucocorticoid Receptor		
	Alert GR 1a - GR Frag 1	C12(C)CCCC1CCC2C
	Alert GR 1b - tert-Butylcyclohexane	C1CCCCC1C(C)(C)C
	Alert GR 1c - GR Frag 4	CC=CCCCC1CCC(C)(O)C1(C)C
	Alert GR 1d - GR Frag 5	CC=CCCC1CCCC(C)(C)C1CC
Transporters		
Dopamine Transporter		
	Alert DT 1 - 3-Phenyl-8-azabicyclo[3.2.1]octane	c1cccc1C2CC(N3)CCC3C2
	Alert DT 2 - Diphenylmethane-like	c1cccc1*c2cccc2

Receptor	MIE	SMARTS/SMILES
Norepinephrine Transporter		
	Alert NT 1a - Cocaine-like 1	<chem>c1cccc1C2C(C(=O)OC)C(N3)CCC3C2</chem>
	Alert NT 1b - Cocaine-like 2	<chem>C1C(C(=O)OC)C(N2)CCC2C1</chem>
	Alert NT 1c - Cocaine-like 3	<chem>c1cccc1C2CC(N3)CCC3C2</chem>
	Alert NT 1d - Cocaine-like 4	<chem>C1CC(N2)CCC2C1</chem>
	Alert NT 2 - Amphetamines-like	<chem>c1cccc1**N</chem>
	OR	<chem>c1cccc1***N</chem>
Serotonin Transporter		
	Alert ST 1 - 3-Ethyl-indole	<chem>CCc1cnc2c1cccc2</chem>
	Alert ST 2 - DMEA	<chem>CN(CC)C</chem>
	Alert ST 3 - Benzyloxybenzene	<chem>c1cccc1COc2ccccc2</chem>
	Alert ST 4 - Diphenylmethane	<chem>c1cccc1Cc2ccccc2</chem>

Table S1. SMILES/SMARTS strings for each structural alert developed in this work.

