SUPPLEMENTARY MATERIALS

Identification of Small Molecule Frequent Hitters of GST–Glutathione interaction

Jara K. Brenke[1,#], Elena S. Salmina[2,#,‡], Larissa Ringelstetter[1], Scarlett Dornauer[1], Maria Kuzikov[3], Ina Rothenaigner[1], Kenji Schorpp[1], Fabian Giehler[4,5], Jay Gopalakrishnan[6], Arnd Kieser[4,5], Sheraz Gul[3], Igor V. Tetko[7,8,‡] and Kamyar Hadian[1,§]

[1]Helmholtz Zentrum München für Gesundheit und Umwelt (HMGU), Assay Development and Screening Platform, Institute of Molecular Toxicology and Pharmacology, Ingolstädter Landstrasse 1, 85764 Neuherberg, Germany
[2]Institute for Organic Chemistry, Technical University Bergakademie Freiberg, Leipziger Str. 29, Freiberg, D-09596, Germany
[3]Fraunhofer Institute for Molecular Biology and Applied Ecology, ScreeningPort (Fraunhofer-IME SP), Schnackenburgallee 114, D-22525 Hamburg, Germany
[4]Helmholtz Zentrum München für Gesundheit und Umwelt, Research Unit Gene Vectors, Marchioninistr. 25, 81377 Munich, Germany
[5]German Center for Infection Research (DZIF), Partner Site Munich, Munich, Germany
[6]Laboratory for Centrosome and Cytoskeleton Biology, CMMC, Robert-Koch-Str. 21, 50931 Cologne, Germany
[7]Helmholtz Zentrum München für Gesundheit und Umwelt (HMGU), Institute of Structural Biology, Ingolstädter Landstrasse 1, G. 60w, 85764 Neuherberg, Germany
[8]BigChem GmbH, Ingolstädter Landstrasse 1, G. 60w, 85764 Neuherberg, Germany

[#] Equal contribution

[‡] Chemoinformatic analysis

[§] Corresponding author

Address correspondence to:
Kamyar Hadian, Ph. D.
Helmholtz Zentrum München GmbH,
Assay Development and Screening Platform,
Institute of Molecular Toxicology and Pharmacology,
Ingolstädter Landstr. 1,
85764 Neuherberg, Germany
Phone: +49 89 3187 2664
E-mail: kamyar.hadian@helmholtz-muenchen.de

CONTENTS

Table S1 PubChem primary screen assay p. 3
Table S2 PubChem confirmatory screen assays p. 4
Table S3 GST-FHs recognized by PAINS filters p. 6
Table S4 GST-FHs recognized by Promiscuity filters p. 8
Table S5 Substructural filters developed for identification of AlphaScreen^{TM}- GST-FHs p. 10
Table S6 Aromatic sulfonylamides in "PC-Artifacts" dataset
Table S7 GST-FHs recognized by Potential Electrophilic Agents (PEA) filters
Table S8 Assay counts for PAINS-, Promiscuous- and PEA-filters recognized the studied GST-FHs
Table S9 Assay counts for GST-FH-filter "Cyanothioazinones" and its oxo-analogue
Table S10 Examples of compounds recognized by GST-FH-filter "Cyanothioazinones" and their "clean" oxo-analogues found in HMGU collection
Table S11 Example of compounds for which substituents significantly modulate sensitivity to GST/GSH-system (based on HMGU collection screening results)

Fig. S1 GST-FHs containing azafluorenone moieties: a) recognized by PAINS-filter "Keto_phenone_A" ; b) and c) do not recognized by PAINS-filters p. 9

Explanation of the developed filters

GST-FH containing quinone moiety **p. 27**
Fig. S2 Examples of studied compounds containing a quinone moiety: a) GST-FH; b) and c) "clean" compounds that are structurally similar to the GST-FH. p. 27
Table S12 Assay counts for filters identifying a dialkoxyethane fragment p. 27

GST-FHs containing catechol moiety **p. 29**
Fig. S3 Compounds containing benzophenanthrene-like molecular scaffold a) GST-FH; b) benzophenanthrene-like molecular scaffold.
Fig. S4 Examples of "clean" compounds containing the benzophenanthrene-like molecular scaffold.

GST-FHs containing sulfonylamide moiety
Fig. S5 GST-FHs containing aromatic sulfonylamide moiety: a) secondary heterylsulfonylamide; b) tertiary sulfonylpiperazine. Dashed lines show applied structural fragmentation.
Table S13 Assay counts for halogen substituted aromatic sulfonylamides

GST-FHs containing azaspironone moiety
Fig. S6 Compounds containing azaspironone-like molecular scaffold: a) GST-FH; b) azaspironone-like molecular scaffold.
Fig. S7 "Clean" compounds from HMGU-collection that contains azaspironone-like molecular scaffold.
Fig. S8 Examples of "clean" compounds from "PC-Primary screen" collection that contain azaspironone-like molecular scaffold.

Table S1

PUBCHEM PRIMARY SCREEN ASSAY

(Query line: <alphascreen gst AND (pcassay_protein_target[filt]) AND (Screening[filt])AND 1000:10000000[Total Sid Count]>)

AID	Donor beads	Acceptor beads	Tested compounds	Tested substances	Comment
623870	Glutathione Donor Beads/GST-TACC3 E629A dimer	Ni^{2+} acceptor beads/His-ARNT PAS-B	All (391,165) Active (2,589) Inactive (388,573) Inconclusive (9)*	All (392,905) Active(2,595) Inactive(390,301) Inconclusive(9)*	**A dataset name:** "PC-Primary screen" **Individual tested compounds:** Active: 2,580 Inactive: 388,520 **Description:** AlphaScreen™ primary screen assay **Interpretation:** Active compounds can be either true or false positive

* inconclusive compounds were skipped

Table S2

PUBCHEM CONFIRMATORY SCREEN ASSAYS

(Query: <alphascreen gst AND (pcassay_protein_target[filt]) AND (Confirmatory[filt])>)

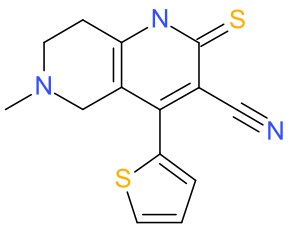
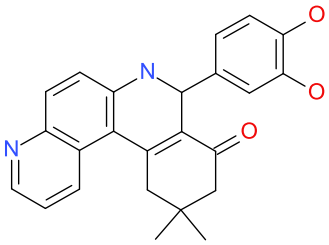
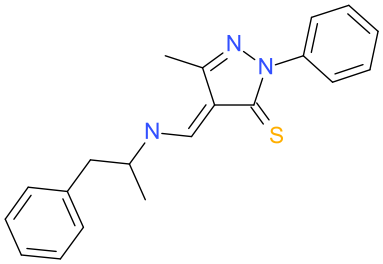
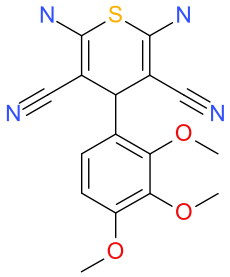
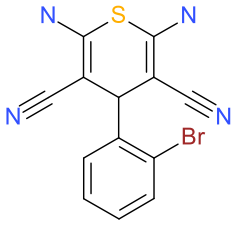
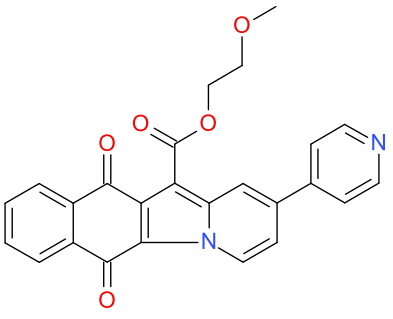
AID	Donor beads	Acceptor beads	Tested compounds	Tested substances	Comment
720564	Glutathione Donor Beads/ dual tagged His-ARNT PASB-GST	Ni^{2+} acceptor beads/ dual tagged His-ARNT PASB-GST	All (42) Active (28) Inactive (14)	All (43) Active (29) Inactive (14)	**A dataset name:** "PC-Artifacts" **Individual tested compounds:** Active: 629 Inactive: 637 **Description:** Assays to detect artifacts of AlphaScreen™ technology **Interpretation:** Active compounds are false positive
651705	Glutathione Donor Beads/dual tagged His-ARNT PASB-GST	Ni^{2+} acceptor beads/ dual tagged His-ARNT PASB-GST	All (1,242) Active (617) Inactive (625)	All (1,246) Active (621) Inactive (625)	
720563	Glutathione Donor Beads/ GST-TACC3 E629A dimer	Ni^{2+} acceptor beads/ His-ARNT PAS-B	All (42) Active (37) Inactive (4) Inconclusive (1)*	All (43) Active (38) Inactive (4) Inconclusive (1)*	**A dataset name:** "PC-Confirmatory" **Individual tested compounds:** Active: 995 Inactive: 288 **Description:** AlphaScreen™ confirmatory screen assay **Interpretation:** Inactive compounds are false positive
651703	Glutathione Donor Beads / GST-tagged fragment of the TACC3 coactivator	Ni^{2+} acceptor beads/ His-ARNT PAS-B	All (1,242) Active (967) Inactive (275)	All (1,246) Active (971) Inactive (275)	
651557	core106 protein tagged with Glutathione-S-transferase	core106 protein tagged with Flag peptide	All (14) Probe (1)* Active (8)	All (14) Probe (1)* Active (8)	

	(GST)/ Glutathione Donor Beads	tag/Anti-Flag acceptor beads	Inactive (6)	Inactive (6)	
463085	core106 protein tagged with Glutathione-S-transferase (GST)/ Glutathione Donor Beads	core106 protein tagged with Flag peptide tag/Anti-Flag acceptor beads	All (3) Active (1) Inactive (2)	All (3) Active (1) Inactive (2)	
1054073 592913 755356 657291 1077871 1077870 684154 684153 665206 707987 425673 755336 1075143					These assays were skipped since description of the protocol was not sufficient on PubChem

* inconclusive compounds and probe were skipped

Table S3

GST-FHs recognized by PAINS filters[1]

OCHEM ID	Structure	Filter name
M416273		Cyano_pyridone_A
M2679239		Catechol_A
M1040098		Ene_five_het_H
M1108836		Dhp_bis_amino_CN
M1108832		Dhp_bis_amino_CN
M1109524		Quinone_A

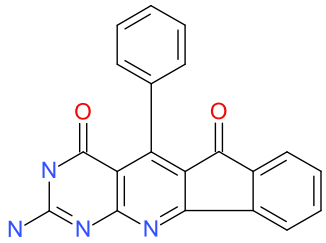
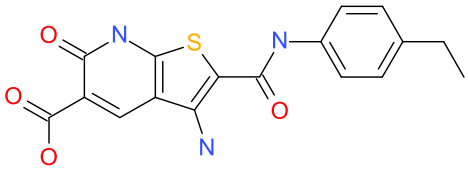
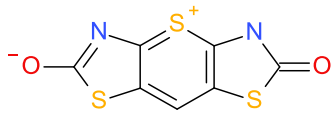
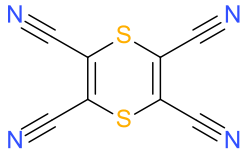
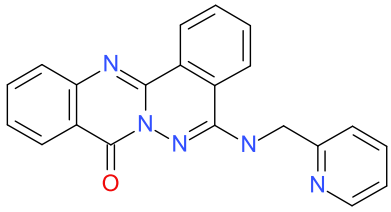
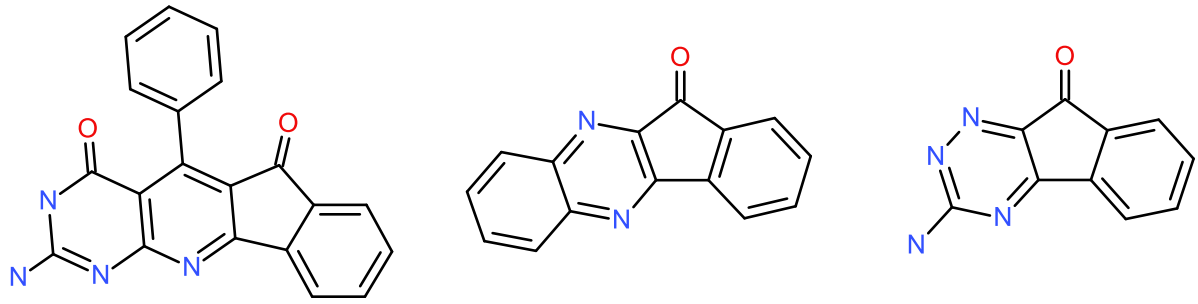
M1027192		Keto_phenone_A
M1110421		Het_65_E

Table S4

GST-FHs recognized by Promiscuity filters[2]

OCHEM ID	Structure	Filter name
M1024320		thiopyrylium
M1024891		four_nitriles
M771355		linear_polycyclic_aromatic_I

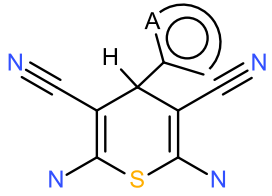
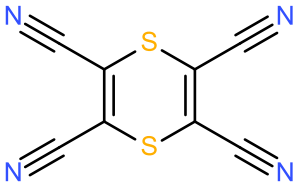
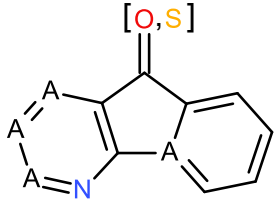


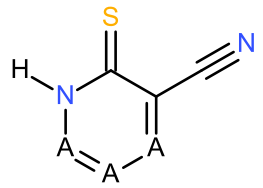
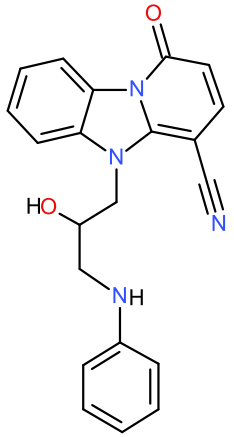
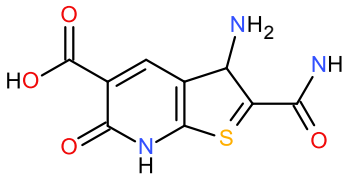
a) b) c)

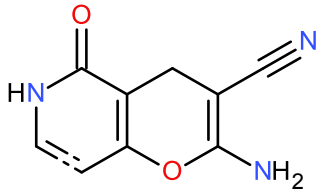
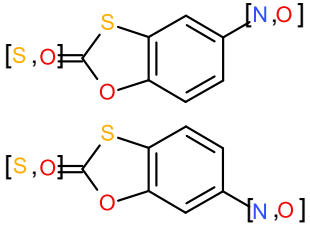
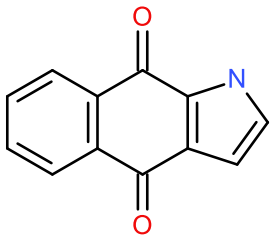
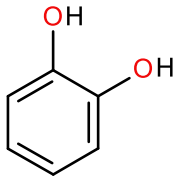
Fig. S1. GST-FHs containing azafluorenone moieties: a) recognized by PAINS-filter "Keto_phenone_A";[1] b) and c) not recognized by PAINS-filters.[1]

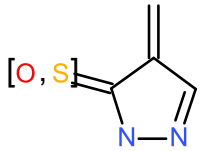
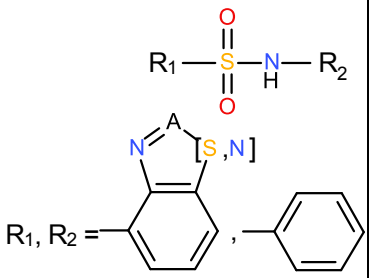
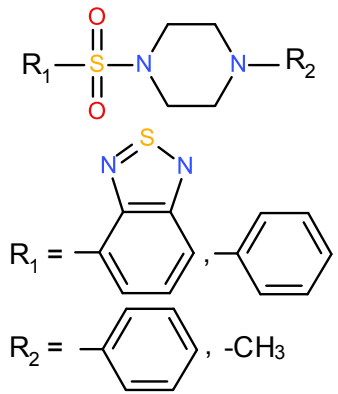
Table S5

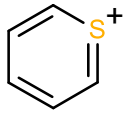
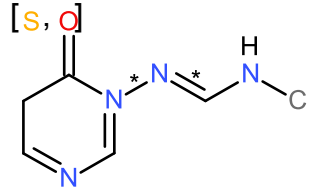
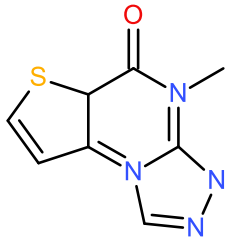
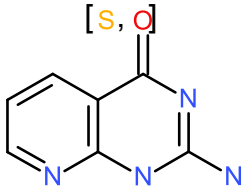
Substructural filters developed for identification of AlphaScreen™-GST-FHs

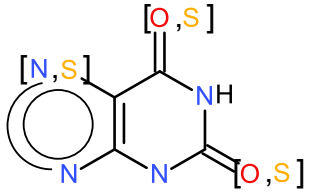
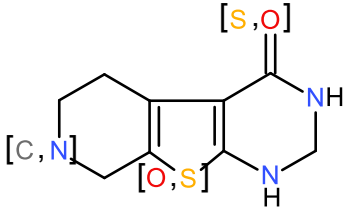
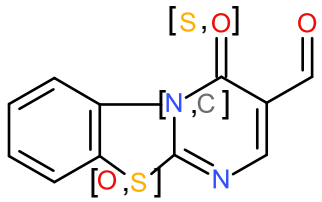
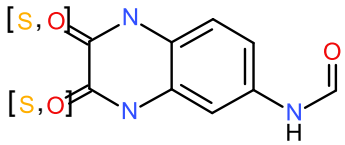
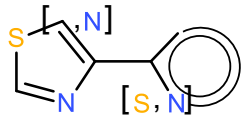
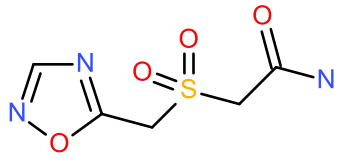
N	Name	Depiction/*Comment*	EV (HMGU)	PC-Artifacts		PC-Confirmatory		PC-Primary screen		SMARTS
				active	inactive	active*	inactive	active	inactive	
1	Cyanothiopyrans	 A = any aromatic atom/ *A copy of PAINS-filter "Dhp_bis_amino_CN"*[1]	233	-	-	-	-	-	15	[NH2]C1=C(C#[NX1])[C!H0](-cc)C(C#[NX1])=C([NH2])S1
2	Cyanodithiine	 A singleton	NA	-	-	-	-	-	-	[#16X2]1-,:[#6X3]([CX2]#[NX1])=,:[#6X3]([CX2]#[NX1])-,:[#16X2]-,:[#6X3]([CX2]#[NX1])=,:[#6X3]1([CX2]#[NX1])
3	Azafluorenones	 A = aromatic C or N/ *Expanded PAINS-filter "Keto_phenone_A"*[1]	57	-	-	-	-	-	23	[OX1,SX1]=[#6X3]1c2c([nX2][cX3,nX2,nX3][cX3,nX2,nX3][cX3,nX2,nX3]2)[c,n]2c1[cX3!$([#6X3]=[A])][cX3!$([#6X3]=[A])][cX3!$([#6X3]=[A])][cX3!$([#6X3]=[A])]2

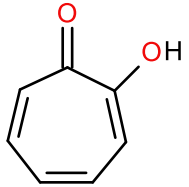
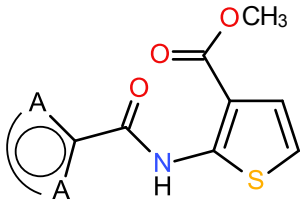
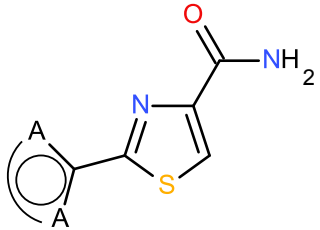
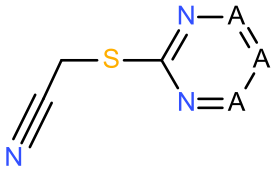
4	Cyanothioazinones	 A = aromatic C or N/ *A copy of PAINS-filter "Cyano_pyridone_A"*[1]	50	-	-	-	-	-	44	[SX1]=[#6X3]1-,:[#7X3H1]-,:[#7X2,#6X3!$([#6X3]=[A])]=,:[#7X2,#6X3!$([#6X3]=[A])]-,:[#7X2,#6X3!$([#6X3]=[A])]=,:[#6X3]1([CX2]#[NX1])
5	Cyanopyridinones		20	-	-	-	-	-	-	[OX1]=[#6X3]1-,:[#6X3]=,:[#6X3]-,:[#6X3]([CX2]#[NX1])=,:[#6X3]2-,:[#7X3]([CX4][CX4]([OX2H1])[CX4][NX3H1]c5ccccc5)-,:[#6]3ccccc3-,:[#7X3]12
6	Thienopyridinones	 *A copy of PAINS-filter "Het_65_E"*[1]	NA	1	-	-	-	-	-	[SX2]1C(=C(C4=C1[NX3!H0]C(C(=[C!H0]4)C(=[OX1])[OX2H1])=[OX1])[NX3H2])C(=[OX1])[NX3!H0]

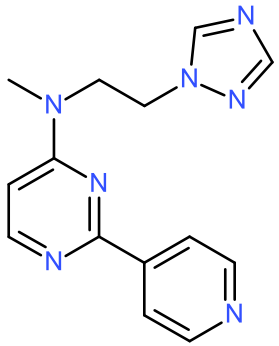
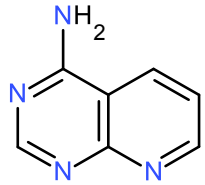
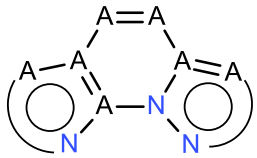
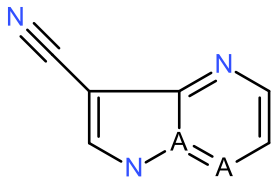
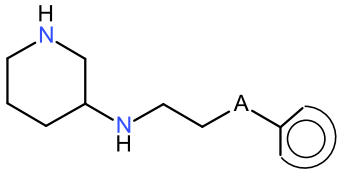
7	Pyranopyridinones	 A dashed line indicates double or aromatic bond	50	-	-	-	-	-	10	[OX2]1[CX3]([NX3H2])=[CX3]([CX2]#[NX1])[CX4][#6X3]2[#6X3](=[OX1])-,:[#7X3H1]-,:[#6X3]=,:[#6X3]-,:[#6X3]12
8	Benzoxathiolones	 *An expanded PAINS-filter "Thio_carbonate_A"*[1]	33	1	-	-	-	1	30	[$([#8]1-,:[#6](-,:[#16]c4cc(ccc14)[#7,#8])=[OX1,SX1]),$([#8]1-,:[#6](-,:[#16]c4ccc(cc14)[#7,#8])=[OX1,SX1])]
9	Indoloquinones		50	-	-	-	-	-	1	[#6X3]1-,:[#6X3]=,:[#6X3]-,:[#6X3]=,:[#6X3]2-,:[#6X3]1-,:[#6X3](=[OX1])-,:[#6X3]3-,:[nX3]4-,:[#6X3](-,:[#6X3]=,:[#6X3]-,:[#6X3]=,:[#6X3]4)=,:[#6X3]-,:[#6X3]3-,:[#6X3]2=[OX1]
10	Catechols	 *A copy of PAINS-filter "Catechol_A"*[1]	17	1	3	2	1	8	281	[c!$([#6]=[A])]1[c!$([#6]=[A])][c!$([#6]=[A])][c!$([#6]=[A])]([c!$([#6]=[A])]([c!$([#6]=[A])]1)[OH1])[OH1]

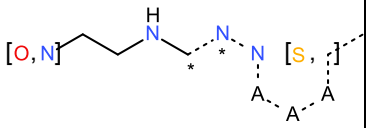
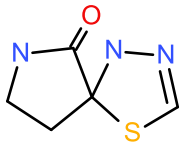
11	Pyrazolones	 *An expanded PAINS-filter "Ene_five_het_H"*[1]	300	4	-	2	2	22	180	[#7X2]1=,:[#6X3]-,:[#6X3](-,:[#6X3](-,:[#7X3]1)=[SX1,OX1])=,:[#6X3H1](-,:[#7X3H1])
12	Sulfonylamide_A	 A = any atom/ *An expanded PAINS-filter "Diazox_sulfon_A"*[1]	100	-	1	-	1	1	74	[$([cX3!$([#6X3]=[A])]3[cX3!$([#6X3]=[A])][cX3!$([#6X3]=[A])][cX3!$([#6X3]=[A])][cX3!$([#6X3]=[A])][cX3]3[SX4](=[OX1])(=[OX1])[NX3H1][cX3]1[cX3!$([#6X3]=[A])][cX3!$([#6X3]=[A])][cX3!$([#6X3]=[A])]2s[cX3!$([#6X3]=[A])]n[cX3!$([#6X3]=[A])]2[cX3!$([#6X3]=[A])]1),$([NX3]([cX3]1[cX3!$([#6X3]=[A])][cX3!$([#6X3]=[A])][cX3!$([#6X3]=[A])][cX3!$([#6X3]=[A])][cX3!$([#6X3]=[A])]1)[SX4](=[OX1])(=[OX1])[cX3]2c3n[o,s]n[cX3!$([#6X3]=[A])]3[cX3!$([#6X3]=[A])][cX3!$([#6X3]=[A])][cX3!$([#6X3]=[A])]2)]
13	Sulfonylamide_B	 *An expanded PAINS-filter "Diazox_sulfon_B"*[1]	17	1	-	-	1	2	544	[NX3]1([CH3,$(c5[cX3!$([#6X3]=[A])][cX3!$([#6X3]=[A])][cX3!$([#6X3]=[A])][cX3!$([#6X3]=[A])]c5)])[CX4][CX4][NX3]([SX4](=[OX1])(=[OX1])[$(c2[cX3R1;!$([#6X3]=[OX1])][cX3R1;!$([#6X3]=[OX1])][cX3R1;!$([#6X3]=[OX1])][cX3R1;!$([#6X3]=[OX1])][cX3R1;!$([#6X3]=[OX1])]2),$(c7[cX3!$([#6X3]=[A])][cX3!$([#6X3]=[A])][cX3!$([#6X3]=[A])]c8nsnc78)])[CX4][CX4]1

14	Thiopyrylium	 *Copy of Promiscuous-filter "Thiopyrylium"*[2]	NA	-	-	-	-	-	-	c1[#16+]cccc1
15	Pyrimidinones_A	 * indicates that a bond also can be aromatic	71	-	-	-	-	2	21	[OX1,SX1]=[#6X3]1[#7X3](-,:[#7X2]=,:[#6X3][NX3H1][#6])-,:[cX3!$([#6X3]=[A])]=,:[#7X2]-,:[cX3!$([#6X3]=[A])]=,:[cX3!$([#6X3]=[A])]1
16	Pyrimidinones_B		25	-	-	-	-	-	18	[OX1]=[#6X3]1-,:[#6X3]2scc-,:[#6X3]2-,:[#7X3]3-,:[#6X3](=,:[#7X2]-,:[#7X2]=,:[#6X3]3[CX4])-,:[#7X3]1[CX4H3]
17	Pyrimidones_C		17	-	-	-	-	-	-	[OX1,SX1]=[#6X3]1[#7X3]-,:[#6X3!$([#6X3]=[A])](-[NX3])=,:[#7X2]-,:[#6X3]2([nX2][#6X3!$([#6X3]=[A])][#6X3!$([#6X3]=[A])][#6X3!$([#6X3]=[A])]c12)

18	Pyrimidinodiones		133	-	-	-	-	2	16	[#6X3]1(=[OX1,SX1])-,:[#7X3H1]-,:[#6X3](=[OX1,SX1])-,:[#7X3,#7X2]~[#6X3]([#7X2,#7X3;R1])-,:[#6X3]1[sX2r5R1,nX2r5R1,$([nx2r6R1]([cX3])[cX3])]
19	Thienopyrimidinone_A		38	1	-	-	-	1	24	[OX1,SX1]=[#6X3]1-,:[#7X3H1]-,:[#6X4]-,:[#7X3H1]-,:[#6]2-,:[#16X2,#8X2]-,:[#6]3[CX4][NX3,CX4][CX4][CX4]-,:[#6]3-,:[#6]12
20	Thienopyrimidinone_B		100	-	-	-	-	-	-	[OX1,SX1]=[#6X3]1-,:[#6X2]([#6X3]=[OX1])=,:[#6X3]-,:[#7X2]=,:[#6X3]2-,:[#16X2,#8X2]-,:[#6]3[cX3!$([#6X3]=[A])][cX3!$([#6X3]=[A])][cX3!$([#6X3]=[A])][cX3!$([#6X3]=[A])][cX3!$([#6X3]=[A])]3-,:[#7X3,#6]12
21	Piperazinediones		50	-	-	-	-	-	-	[#7X3]1-,:[#6X3](=[OX1,SX1])-,:[#6X3](=[OX1,SX1])-,:[NX3]-,:[#6X3]2[cX3!$([#6X3]=[A])][cX3](=[NX3H1][CX3]=[OX1])[cX3!$([#6X3]=[A])][cX3!$([#6X3]=[A])][cX3]12
22	4-Heteryl-thiazoles		87	43	2	2	-	61	287	s1[c,nX2][c,nX2](-c(c)[nX2,nX3,sX2])[nX2]c1
23	Oxadiazoles		50	-	-	-	-	-	-	o1nc([CX4])nc1[CX4H2][SX4](=[OX1])(=[OX1])[CX4H2][CX3](=[OX1])[NX3!H0]

24	Cycloheptenones		100	-	2	2	-	4	12	[#6]1-,:[#6]=,:[#6]-,:[#6X3](=[OX1])-,:[#6]([OX2H1])=,:[#6]-,:[#6]=,:1
25	Thiophenes	 A = any aromatic atom	33	-	-	-	-	-	75	[sX2]1[cX3]([NX3H1][CX3](=[OX1])-c)[cX3]([CX3](=[OX1])[OX2][CX4H3])[cX3][cX3]1
26	Thiazoles		100	-	-	-	-	-	1	[nX2]1[cX3](-[cX3])[sX2][cX3][cX3]1[CX3](=[OX1])[NX3H2]
27	Cyanothioethers	 A = aromatic C or N	33	-	-	-	-	1	49	[NX1]#[CX2][CX4H2][SX2]c1n[c,n][c,n][c,n]n1

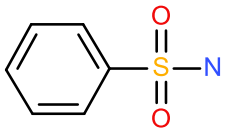
28	Aminopyrimidinones_A		NA	-	-	-	-	-	-	c1[nX2]c[nX2][nX3]1[CX4][CX4][NX3]([CX4])-c2cc[nX2]c(-c3cc[nX2]cc3)[nX2]2
29	Aminopyrimidinones_B		NA	-	-	-	-	-	16	[NX3H2]c1ncnc2ncccc21
30	Fused azacycles_A	 A = aromatic C or N	50	-	-	-	-	-	4	[nX2R1][c,n;R2]1[c,n;R2]([c,n;R1])[c,n][c,n][c,n;R2]([c,n;R1])[nR2]1[nX2R1]
31	Fused azacycles_B	 A = aromatic C or N	29	-	-	-	-	-	83	[NX1]#[CX2]c1[c!$([#6]=[OX1])]n[n,c!$([#6]=[OX1])]2[c,n][c!$([#6]=[OX1])][c!$([#6]=[OX1])]nc12
32	Ethylenediamines		50	-	-	-	-	-		[NX3H1]1[CX4][CX4][CX4][CX4]([NX3H1][CX4][CX4]Ac)[CX4]1

		A = any aliphatic atom								
33	Arylethanolamines	 A dashed line indicates aromatic bonds; * indicates atoms that must belong to 6-membered ring	29	-	-	-	-	1	12	[$([NX3]([CX4,H])([CX4,H])[CX4]),$([OX2]([H,CX4])[CX4])][CX4!H0][CX4H2][NX3H1]c[nX2r6R1][nX2,nX3][nX2,cX3;R1][cX3,nX2][cX3,nX2][sX2,cX3][cX3]
34	Azaspirononones		33	1	-	-	-	2	11	[X4;R2]12(-,:;@[#6X3]-,:;@[#6X3]-,:;@[#7X3]-,:;@[#6X3]-,:;@1)-,:;@[#16X2]-,:;@[#6X3]=,:;@[#7X2]-,:;@[#7X3]-,:;@2

* The number of active “PC-Confirmatory”-compounds was reduced by the number of identical compounds found among “Artifacts”-compounds.

Table S6

Aromatic sulfonylamides identified in "PC-Artifacts" dataset

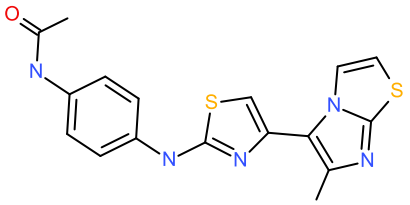
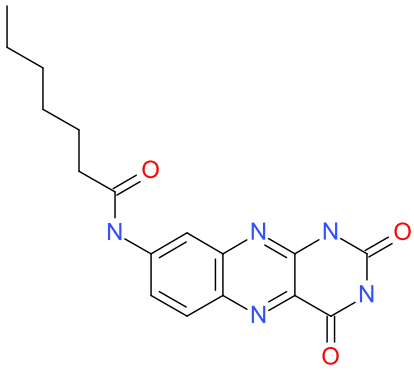
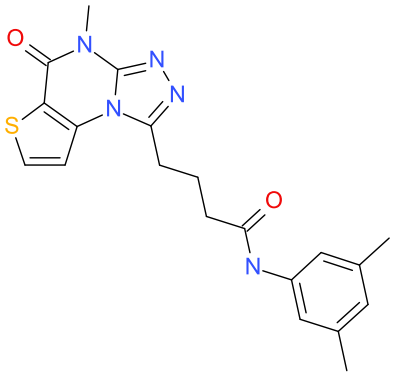
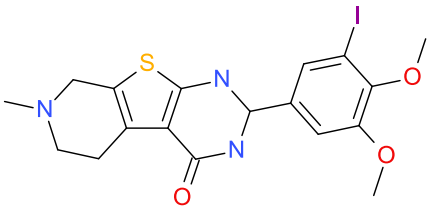
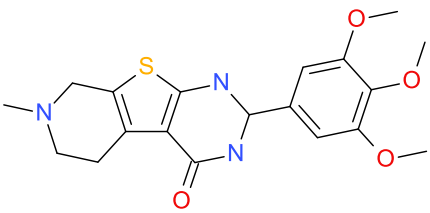
Searched fragment	Active compounds	Inactive compounds
	70 (30)[1]	58 (4)[1]

[1]The number of total and (in parenthesis) compounds identified by AlphaScreen-HIS-FHs[3] filters are shown.

The PC-Artifacts dataset included compounds tested for their ability to be false positive. The screening assays incorporated both GST-tagged and His-tagged proteins. Thus their promiscuity could be due to their interference with His/Ni^{2+}- or GST/GSH-systems. The AlphaScreen-HIS-FHs[3] filters identified 30 (43%) of compounds responsible for the first mechanism while promiscuity of the remaining 40 (57%) of compounds can be assumed to be GST specific.

Table S7

GST-FHs recognized by Potential Electrophilic Agents (PEA) filters[4]

OCHEM ID	Structure	Filter name
M3383345		Acetates
M1111066		Acetates
M1215709		Acetates
M1108012		Hydroquinones
M2675581		Hydroquinones

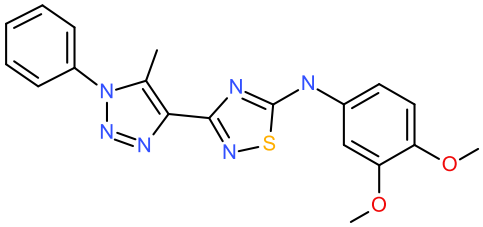
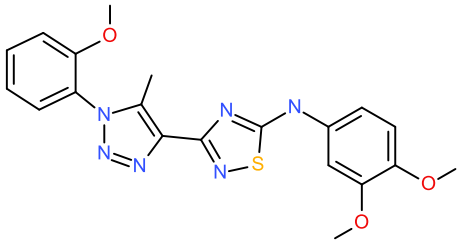
M1217802		Hydroquinones
M3200831		Hydroquinones

Table S8

Assay counts for the PAINS-, Promiscuous- and PEA-filters recognized the studied GST-FHs[1, 2, 4]

Filter name	HMGU*		"PC-Artifacts"		"PC-Confirmatory"		"PC-Primary screen"	
	active	inactive	active	inactive	active	inactive	active	inactive
PAINS-filters[1]								
Cyano_pyridone_A	2	4	-	-	-	-	-	44
Catechol_A	3	12	1	3	3	1	8	281
Ene_five_het_H	2	-	-	-	-	-	-	-
Dhp_bis_amino_CN	14	3	-	-	-	-	-	15
Quinone_A	31	27	5	1	6	-	63	449
Keto_phenone_A	1	2	-	-	-	-	-	21
Het_65_E	1	-	-	-	-	-	-	-
Thio_carbonate_A	-	8	-	-	-	-	-	27
Promiscuous-filters[2]								
Thiopyrylium	1	-	-	-	-	-	-	-
Four_nitriles	2	1	-	-	-	-	-	11
Linear_polycyclic_aromatic_I	4	5	-	-	-	-	-	23
Potential electrophilic agents (PEA) - filters[4]								
Acetates	62	2,025	90	79	136	35	277	56,311
Hydroquinones	44	510	14	31	31	14	96	13,129
Pyranones	15	513	2	9	6	5	40	8,724
Poly aromatic hydrocarbons	35	216	15	-	22	8	61	7,251
Azlactones	8	80	-	-	-	-	-	672

* Compounds from HMGU-collection that are active at least in one GST/GSH-comprising assay were considered as "active" (in total 775 compounds) and compounds that are inactive in both GST/GSH-comprising assays were considered as inactive (in total 24,222).

Acceptance/rejection of a filter was based on both, the ratio of active and active compounds for a given dataset and mechanistical background for the filter. According to the received data PAINS- and promiscuous filters are more selective than PEA-filters. In average for the formers inactive/active ratio is ≤ 4 while for PEA-filters is ≥ 20. The activity of the compounds in PubChem dataset can result from both, a true protein-protein disruption effect as well as an interference with AlphaScreen™ technology. Therefore, we paid more attention to assay counts for HMGU-collection and if inactive/active ratio for a filter was > 4 we skipped the filter from the further consideration.

Table S9

Assays counts for "Cyanothioazinones" GST-FH filter and its oxo-analogue

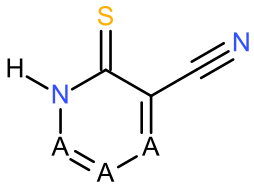
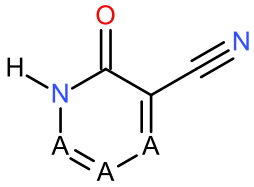
Filter	GST-FHs	HMGU*		"PC-Artifacts"		"PC-Confirmatory"		"PC-Primary screen"	
		active	inactive	active	inactive	active	inactive	active	inactive
	1	2	4	-	-	-	-	-	44
	-	-	18	-	-	-	-	1	

Table S10

GST-FH recognized by "Cyanothioazinones" filter and its "clean" oxo-analogues found in HMGU-collection

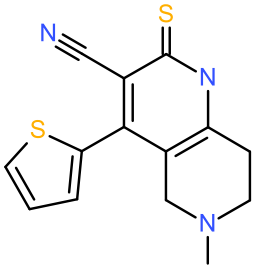
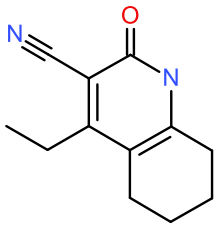
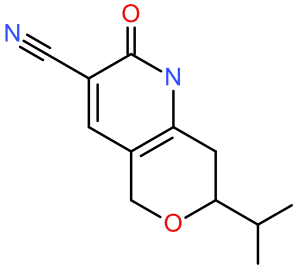
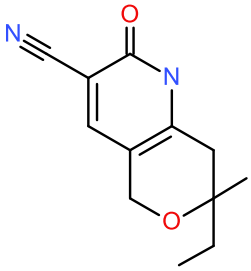
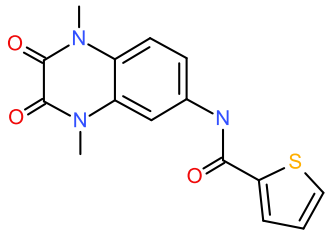
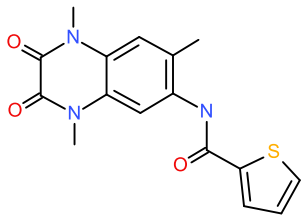
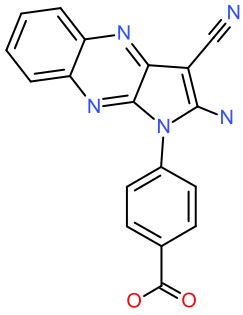
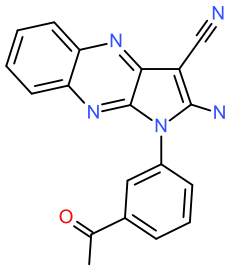
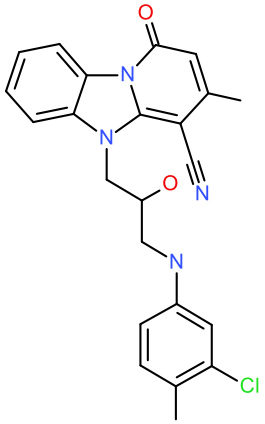
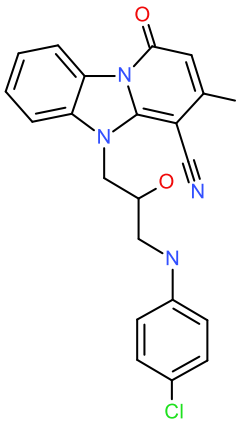
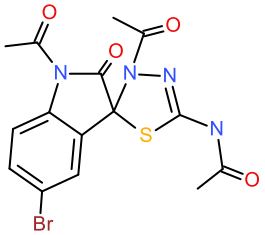
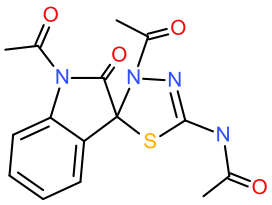
GST-FH	"Clean" oxo-analogues		
			
M416273	M1025582	M1027169	M1027170

Table S11

Example of compounds for which substituents greatly modulate sensitivity to GST/GSH-system (based on the HMGU-collection screening results)

GST-FH	"Clean" compound
	
M1111171 (28\ 5\ 103\ 95\ 104) *	M1111172 (86\ 84\ 88\ 101\ 104)
	
M1109005 (50\ 65\ 124\ 109\ 113)	M1951244 (105\ 89\ 110\ 104\ 98)
	
M1113194 (47\ 51\ 96\ nd\ 94)	M1113190 (95\ 94\ 101\ 90\ 102)
	
M1040939 (59\ 12\ 104\ 97\ 102)	M1108669 (103\ 106\ 107\ 104\ 103)

***** Each molecule is characterized as follow:

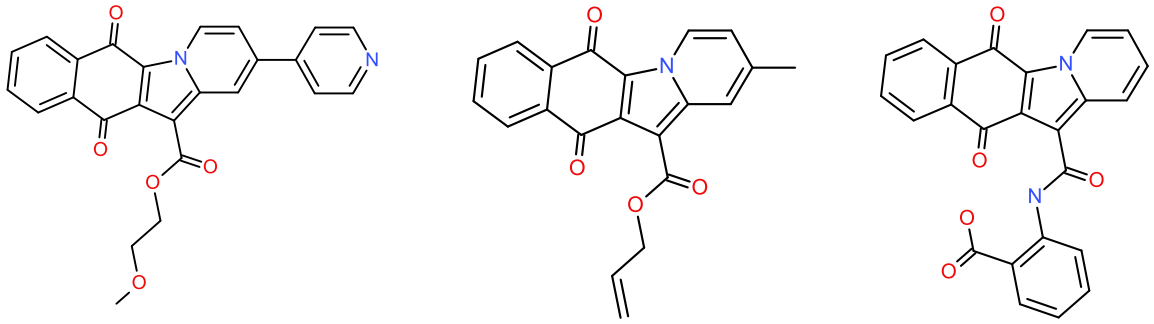
<*OCHEMID* (*E* (assay 1) / *E* (assay 2) / *E* (assay 3) / *E* (assay 4) / *E* (assay 5))> where *OCHEMID* is the ID-number of a compound in OCHEM database;[5]

E (assay *i*) is the measured light emission in the studied i^{th} AlphaScreen™ assay, % (*i* = 1…5). For detailed information on assays please refer to "Materials and Methods" section in the corresponding publication.

Explanation of the developed filters

GST-FH CONTAINING QUINONE MOIETY

The only GSH-FH containing quinone moiety is shown on Fig. S2a. In this compound a quinone structure is a part of the anthraquinone aromatic system and so does not possess double bonds activated for electrophilic attacks. We think that the quinone moiety alone does not provoke promiscuity of this class of compounds. Additionally, a structural examination showed that highly structurally similar quinones appear among "clean" substances (Fig. S2b,c).



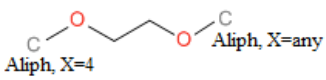
a) b) c)

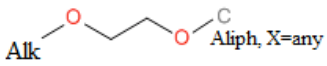
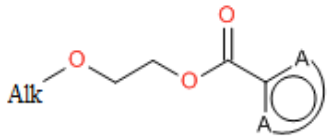
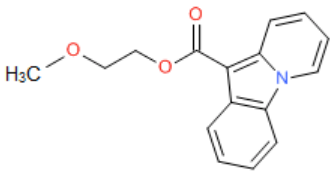
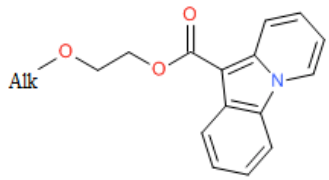
Fig. S2. Examples of studied compounds containing a quinone moiety: a) GST-FH; b) and c) "clean" compounds that are structurally similar to the GST-FH.

Therefore we assumed that promiscuity of this compound in GST/GSH assays is due to presence of a dimethoxymethane-like "tail". Overall statistics for this fragment is presented in Table S12. It is clear that the dialkoxyethane-tail alone also does not allow distinguishing between GST-FHs and "clean" compounds. Its combination with an aromatic scaffold is required for the promiscuity and this effect is also structure specific. Due to these reasons we created a singleton-like filter. It can be extended once new data will be collected or there will a better mechanistic understanding of interference of compounds with GST/GSH assays.

Table S12

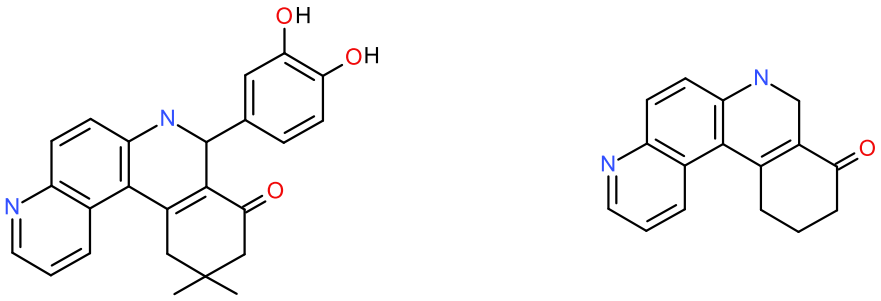
Assay counts for filters identifying a dialkoxyethane fragment

Filter	HMGU*		"PC-Artifacts"		"PC-Confirmatory"		"PC-Primary screen"	
	active	inactive	active	inactive	active	inactive	active	inactive
"Dialkoxyethane tail_1"  Aliph = aliphatic atom; X = the number of total connection of the carbon SMARTS:<[CX4][OX2][CX4H2][CX4H2][OX2]C>	**2**	**49**	**1**	**7**	**2**	**6**	**15**	**721**
"Dialkoxyethane tail_2"	**2**	**33**	**1**	**1**	**-**	**2**	**3**	**291**

 Alk = alkyl substituent; Aliph = aliphatic atom; X = the number of total connection of the carbon SMARTS:<[$([CH3]),$([CH2][CH3]),$([CH2]([CH3])[CH3]),$([CH2][CH2][CH3]),$([CH2][CH2][CH2][CH3])][OX2][CX4H2][CX4H2][OX2]C>								
“Dialkoxyethane tail_3”  Alk = alkyl substituent; A = any aromatic atom SMARTS:<[CX4H3][OX2][CX4H2][CX4H2][OX2][CX3](=[OX1])c>	**1**	**1**	**1**	**1**	**-**	**2**	**2**	**70**
“Dialkoxyethane tail_4”  SMARTS:<[CX4][OX2][CX4H2][CX4H2][OX2][CX3](=[OX1])c1c[nX3]c2[c,n][c,n]3[c,n][c,n][c,n][c,n][c,n]3[c,n]c12>	**1**	**-**	**1**	**1**	**-**	**2**	**2**	**4**
“Dialkoxyethane tail_5”  SMARTS:<[$([CH3]),$([CH2][CH3]),$([CH2]([CH3])[CH3]),$([CH2][CH2][CH3]),$([CH2][CH2][CH2][CH3])][OX2][CX4H2][CX4H2][OX2][CX3](=[OX1])c1c[nX3]c2[c,n][c,n]3[c,n][c,n][c,n][c,n][c,n]3[c,n]c12>	**1**	**-**	**1**	**1**	**-**	**2**	**2**	**4**

GST-FH CONTAINING CATECHOL MOIETY

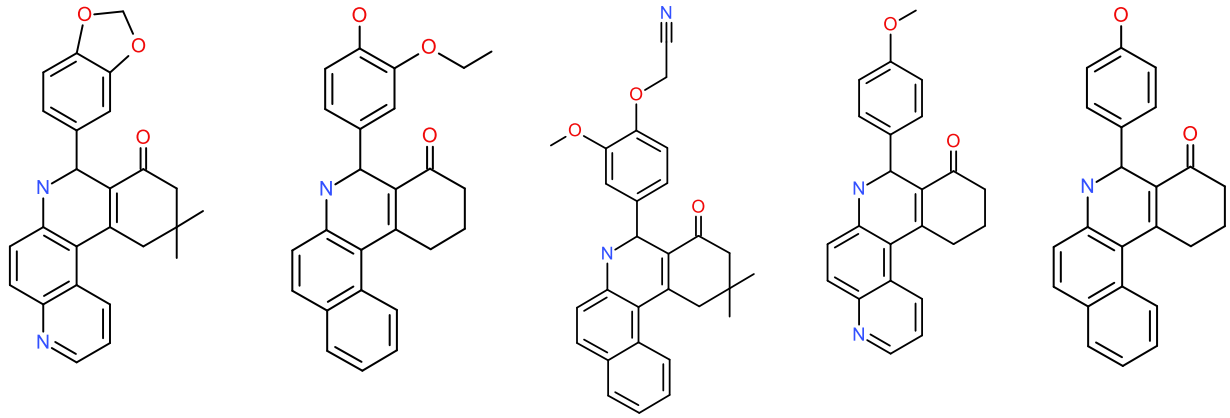
The only GSH-FH containing a catechol moiety is shown on Fig. S3a:



a) b)

Fig. S3. Compounds containing benzophenanthrene-like molecular scaffold a) a GST-FH identified in HMGU collection; b) benzophenanthrene-like molecular scaffold.

A benzophenanthrene-like molecular scaffold (Fig. S3b) to which the catechol moiety is attached can be easily found among "clean" substances (Fig. S4). Thus this scaffold should not be considered alone as a reason for the compound promiscuity in GST/GSH-assays.



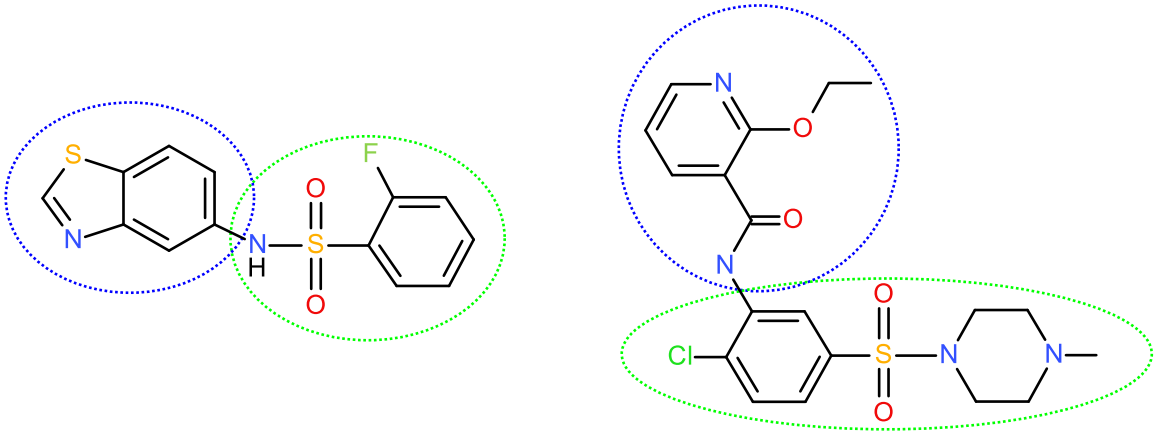
a) b) c) d) e)

Fig. S4. Examples of "clean" compounds containing the benzophenanthrene-like molecular scaffold.

A catechol fragment is incorporated in 12 clean substances of very high diversity and met among 281 inactive versus 8 active compounds in the primary screen. Despite of its known problematic nature it is difficult to assume that this fragment is the reason for GST-promiscuity. The latter is more likely caused by simultaneous contributions of the scaffold, catechol moiety, overall molecular shape and hydrogen bond donating/accepting properties. So we restricted "Catechol_A" filter by inclusion of the aromatic molecular scaffold (Fig. S3b). The filter might be refined once more structural information appear.

GST-FH CONTAINING SULFONYLAMIDE MOIETY

Two sulfonylamide derivatives were presented among the studied GST-FHs: secondary heterylsulfonylamide (Fig. S5a) and tertiary sulfonylpiperazine (Fig. S5b).



a) b)

Fig. S5. GST-FHs containing aromatic sulfonylamide moiety: a) secondary heterylsulfonylamide; b) tertiary sulfonylpiperazine. Dashed lines show applied structural fragmentation.

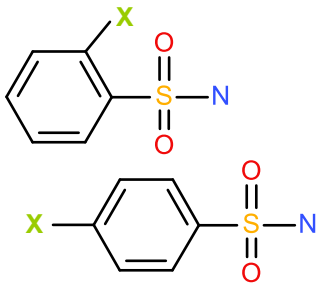
Sulfonylpiperazine chemical moiety is a relatively unreactive scaffold used as a basic of several medicines. Structural analysis showed that more than 100 “clean” substances comprise this moiety what allowed assuming that this fragment is not responsible for GST/GSH-interfering potential of the substances. Hence, it can be assumed that promiscuity of the sulfonylpiperazine is caused by fragment contoured by a blue dashed line on Fig. S5b. However, this fragment was present in 20 “clean” substances. So either both these fragments define the compound promiscuity synergistically or there is another fragment that causes interference with GST/GSH-system.

The heterylsulfonylnamide comprises a benzothiazole moiety (countered by a blue dashed line on Fig. S5a). This fragment does not possess any reactive properties too. The secondary sulfonylamide moiety (countered by a green dashed line on Fig. S5a) is incorporated by more than 500 “clean” compounds.

One of distinguishing structural features of these two GST-FHs is that both of them contain halogen substituted benzene ring with electron withdrawing sulphonylamido group in *ortho*- or *para*- positions. Structural analysis of HMGU and PubChem datasets revealed next assay counts (Table S13):

Table S13

Assay counts for halogen substituted aromatic sulfonylamides

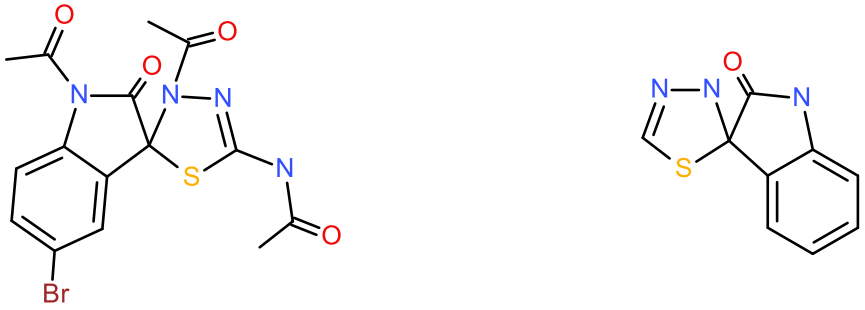
Filter	HMGU*		"PC-Artifacts"		"PC-Confirmatory"		"PC-Primary screen"	
	active	inactive	active	inactive	active	inactive	active	inactive
 X = F, Cl, Br, I	5	108	8	7	2	5	8	7,836

According to the results this filter did not distinguish between "clean" compounds and GST-FHs too. Therefore, PAINS-filters "Diazox_sulfon_A" and "Diazox_sulfon_B" that recognize structurally highly similar sulfonylamides were slightly modified to include the identified GST-FHs and added to GST-FH-S as filters with high selectivity.

GST-FH CONTAINING AZASPIRONONE MOIETY

For all studied azaspironones chirality of spirocarbon was not provided for the analyzed datasets. Azaspironones belong to a compound class where the nature of substituents greatly modifies their GST/GSH-promiscuity.

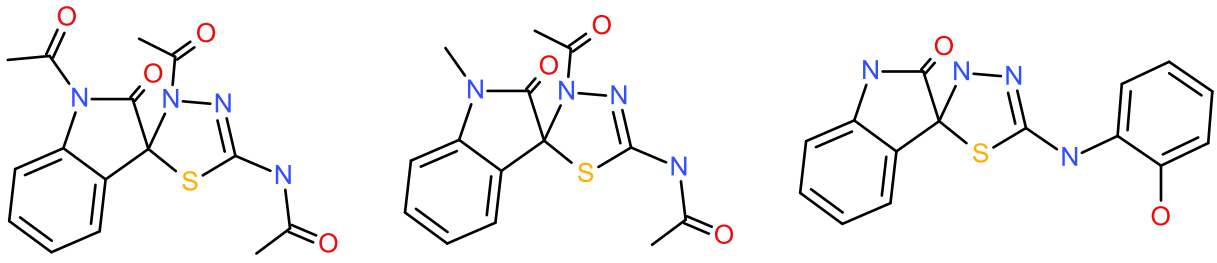
Among GST-FHs only one spironone was presented (Fig. S6a). The structural analysis showed that in PubChem and HMGU collections there were other spirocompounds containing a common scaffold depicted on Fig. S6b.



a) b)

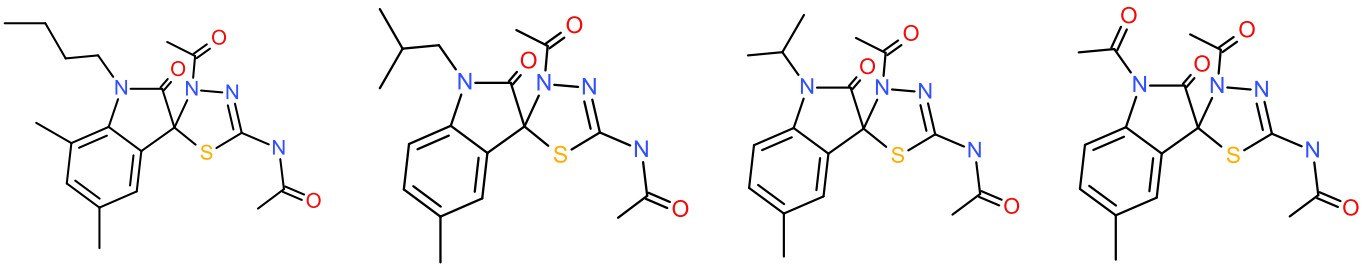
Fig. S6. Compounds containing azaspironone-like molecular scaffold: a) GST-FH; b) azaspironone-like molecular scaffold.

HMGU collection comprised three (Fig. S7) and "PC-Primary screen" collection had 11 (Fig. S8) "clean" compounds containing azaspironone-like molecular scaffold. The only difference between GST-FH and "clean" compounds (cmp. Fig. S7a and Fig. S8a) is a bromine substituent in benzene ring. However, we were not able to check the influence of halogen-substituents onto GST-promiscuity of compounds due to the limited structural data. Meanwhile, analysis of PubChem data showed that all inactive compounds contain alkyl-group at the amide nitrogen while in active substances (shared data from HMGU and PubChem collections) this position is occupied either by hydrogen or acetyl-group (Fig. S8b-d). However, the developed filter does not include the restriction for the amide nitrogen substituent because more structural information is necessary to confirm it.



a) b) c)

Fig. S7. "Clean" compounds from HMGU-collection that contain azaspironone-like molecular scaffold.



a) b) c) d)

Fig. S8. Examples of “clean” compounds from “PC-Primary screen” collection that contain azaspironone-like molecular scaffold.

Supplementary references

1. Baell, J.B.; Holloway, G.A. New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. *J. Med. Chem.* **2010**, 53, 2719-40.
2. Pearce, B.C.; Sofia, M.J.; Good, A.C.; Drexler, D.M.; Stock, D.A. An empirical process for the design of high-throughput screening deck filters. *J. Chem. Inf. Model.* **2006**, 46, 1060-8.
3. Schorpp, K.; Rothenaigner, I.; Salmina, E.; Reinshagen, J.; Low, T.; Brenke, J.K.; Gopalakrishnan, J.; Tetko, I.V.; Gul, S.; Hadian, K. Identification of Small-Molecule Frequent Hitters from AlphaScreen High-Throughput Screens. *J. Biomol. Screen.* **2014**, 19, 715-726.
4. Enoch, S.J.; Madden, J.C.; Cronin, M.T. Identification of mechanisms of toxic action for skin sensitisation using a SMARTS pattern based approach. *SAR QSAR Environ. Res.* **2008**, 19, 555-78.
5. Sushko, I.; Novotarskyi, S.; Korner, R.; Pandey, A.K.; Rupp, M.; Teetz, W.; Brandmaier, S.; Abdelaziz, A.; Prokopenko, V.V.; Tanchuk, V.Y.; Todeschini, R.; Varnek, A.; Marcou, G.; Ertl, P.; Potemkin, V.; Grishina, M.; Gasteiger, J.; Schwab, C.; Baskin, I.I.; Palyulin, V.A.; Radchenko, E.V.; Welsh, W.J.; Kholodovych, V.; Chekmarev, D.; Cherkasov, A.; Aires-de-Sousa, J.; Zhang, Q.Y.; Bender, A.; Nigsch, F.; Patiny, L.; Williams, A.; Tkachenko, V.; Tetko, I.V. Online chemical modeling environment (OCHEM): web platform for data storage, model development and publishing of chemical information. *J. Comput.-Aided Mol. Des.* **2011**, 25, 533-54.