

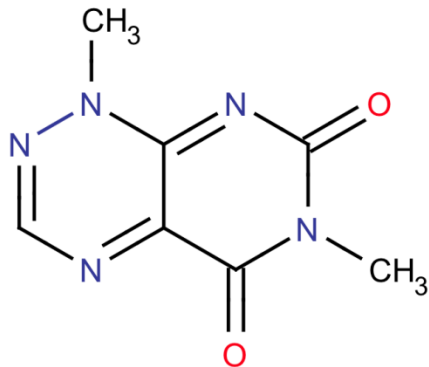
Mining promiscuous substructures in PubChem using molecular matching pairs

Oleg Ursu

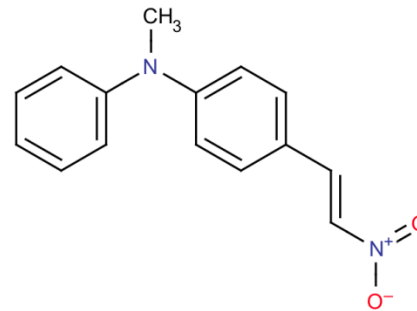
University of New Mexico

Frequent hitters

- Hits in many unrelated HTS assays
 - Promiscuous compounds (aggregation, reactive, etc.)
 - Assay interference (fluorescence, quenchers, etc.)
- Experts knowledge/SMARTS rules
- Not available in public domain



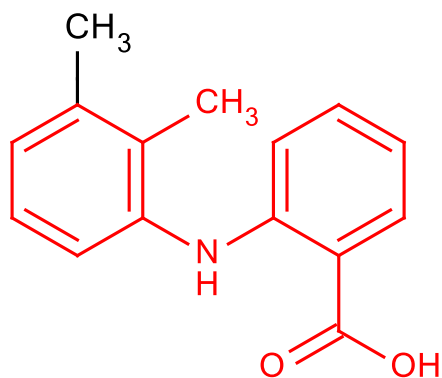
558/194



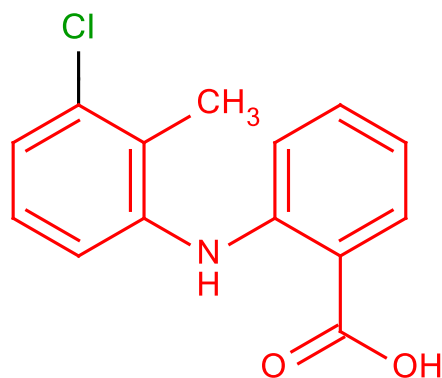
452/62

Molecular matching pairs

Molecular pair

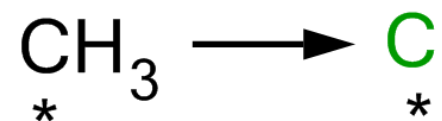


Mefenamic acid

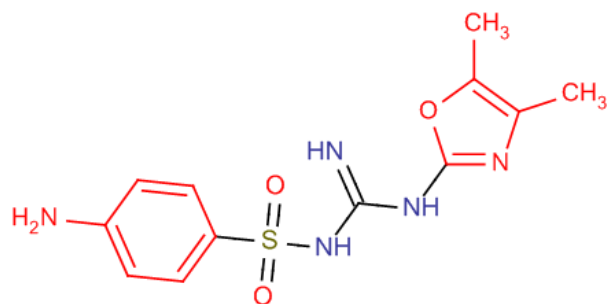


Tolfenamic acid

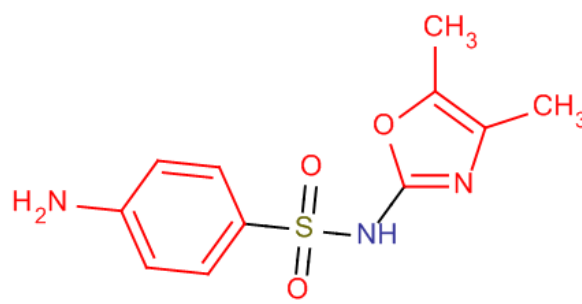
Molecular transformation



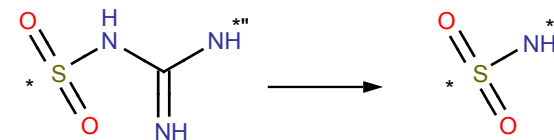
Single cut



Sulfaguanole



Sulfamoxole



Double cut

MMPs applications and implementations

- Identify molecular transformations in a series of chemical structures
- Map molecular transformations to changes in biological/chemical/physical properties
- Study bioisosterism, activity cliffs, SAR
- Guide optimization of: aqueous solubility, plasma protein and hERG binding, P450 metabolism
- Leatherface¹, WizePairZ², Hussain³

1. Cheminformatics in Drug Discovery, 2004, 271-285

2. J. Chem. Inf. Model. 2010, 50, 1350-1357

3. J. Chem. Inf. Model. 2010, 50, 339-348

PubChem database

- Chemical and biological activity database, NIH funded
- Data deposited by 9 Molecular Libraries centers
- Additional data from NCI, ChEMBL, Pharma, etc.

	Total	Molecular Libraries
Substances	~ 85 mil	~ 400k
Assays	504,228	3,860 (0.77%)
Screening data points	~142 mil	~ 134 mil (94%)

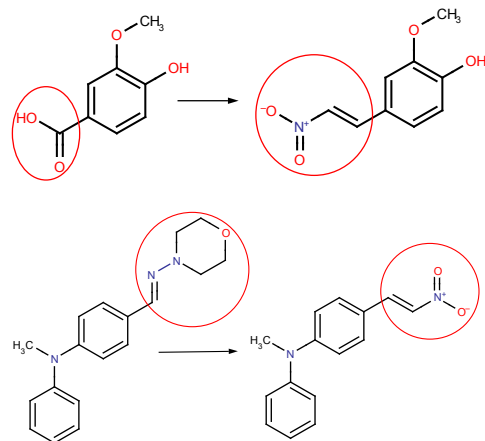
August 2011

Frequent hitters to promiscuous substructures

Download and clean MLSMR library from PubChem

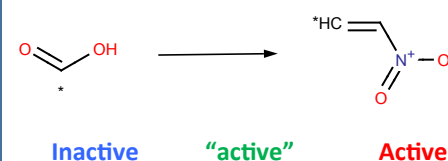
- Salt and solvent removal
- Tautomers

Identify all non-redundant single, double, and triple cut MMPs

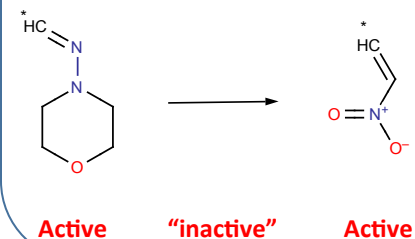


Map “active” MMPs to HTS assays

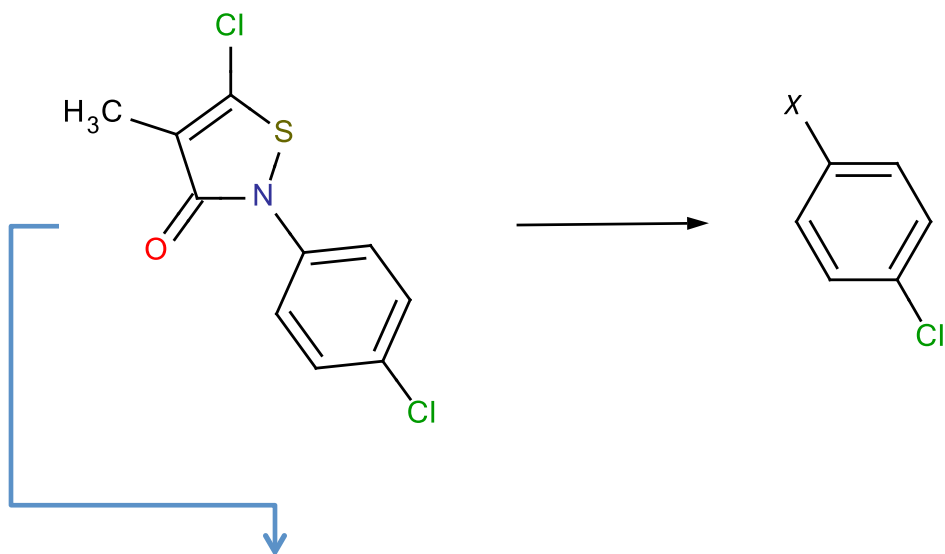
AID 493098



AID 504408



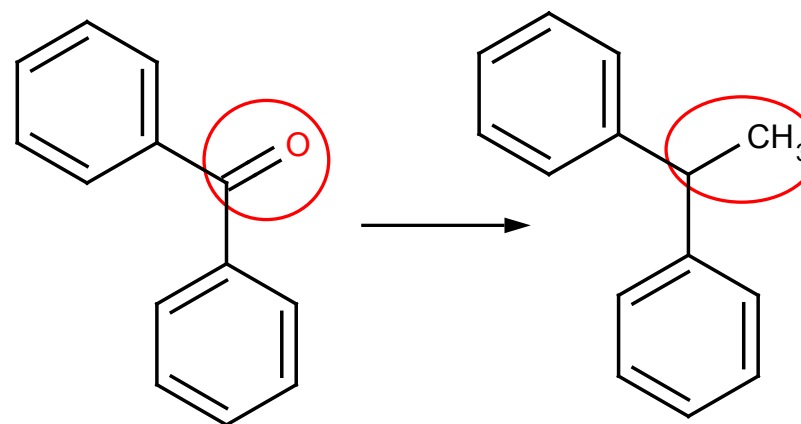
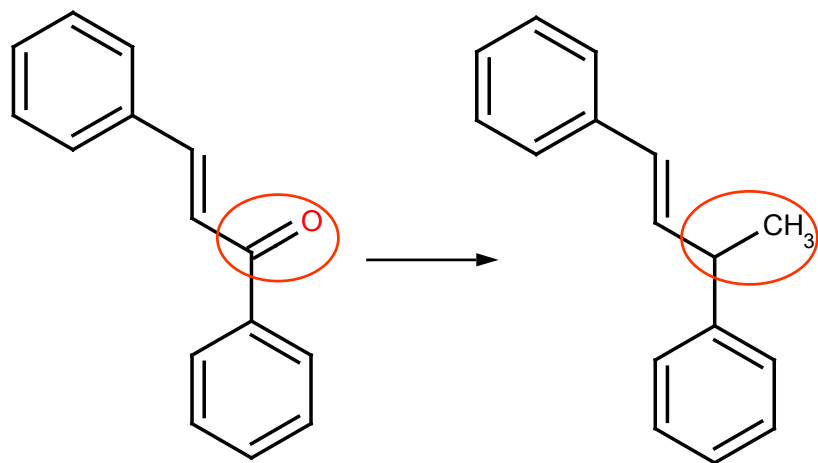
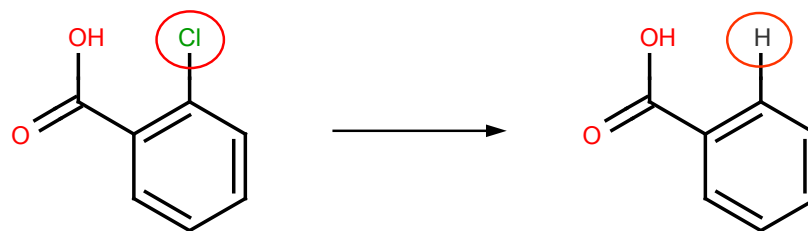
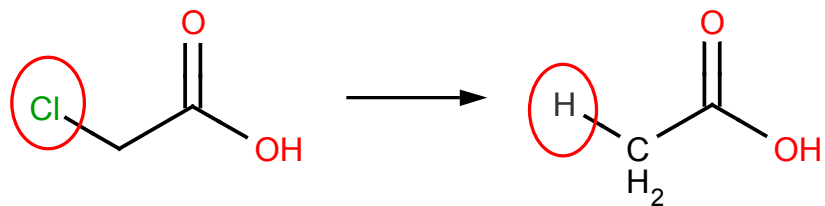
“Promiscuous” fragments



24792062
Tested assays 359
Active assays 75

X	Tested assays	Active assays
	354	1
26731534		
	354	1
26728462		
	348	1
26728988		

MMP environments



Promiscuity score

$$\text{score} = \frac{\text{active subst.}}{\text{testedsubst.} + \text{median}(\text{testedsubst.})} \times \frac{\text{active mmeps}}{\text{tested mmeps} + \text{median}(\text{tested mmeps})} \times \frac{\text{active assays}}{\text{testedassays} + \text{median}(\text{testedassays})} \times \frac{\text{active samples}}{\text{testedsamples} + \text{median}(\text{testedsamples})} \times 1600$$

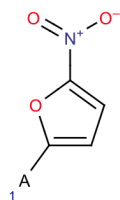
of substances with 1 or more "active" MMPs

of "active" MMPs for each component

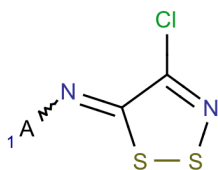
distinct assays with 1 or more "active" MMPs

of samples mapped to each "active" MMP

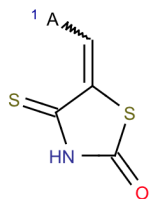
Top promiscuous fragments



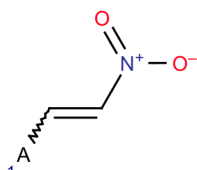
25/24
61/59
421/97
7436/395
7.16



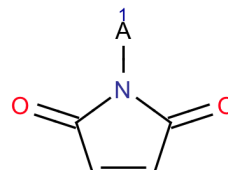
23/23
1113/1101
576/79
7736/410
5.83



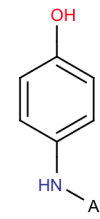
6/6
268/264
416/76
2195/140
5.53



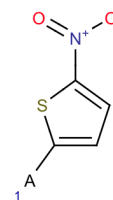
32/24
208/198
453/110
10729/460
5.48



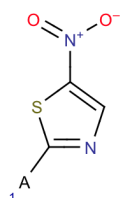
29/24
378/336
470/79
6275/340
4.92



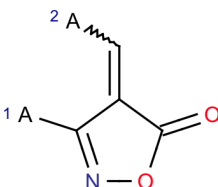
12/12
96/96
564/72
4242/221
4.40



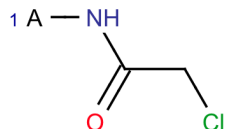
21/19
50/46
397/70
5649/268
4.01



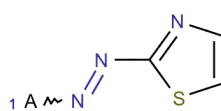
42/35
256/185
470/100
11104/394
3.49



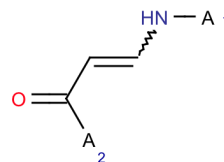
6/6
520/498
465/60
2070/109
3.32



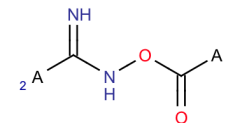
70/56
344/323
472/92
24694/587
2.81



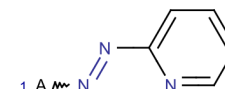
5/4
124/68
444/79
1232/109
2.75



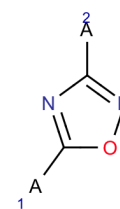
40/35
890/769
568/80
13417/388
2.64



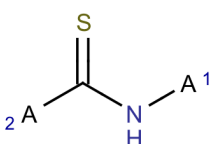
53/47
568/537
474/65
17155/475
2.56



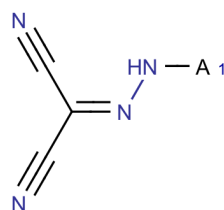
4/3
189/114
446/69
976/100
2.53



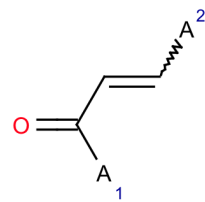
59/46
499/456
578/101
22394/509
2.50



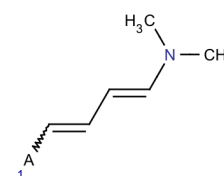
10/10
329/296
552/65
3625/128
2.48



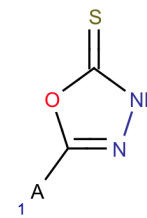
9/9
254/249
452/52
3283/116
2.32



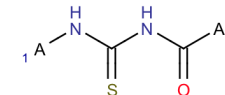
39/35
204/154
454/80
11226/250
1.99



2/2
16/15
396/59
692/63
1.93



36/35
1008/843
575/71
13681/224
1.41

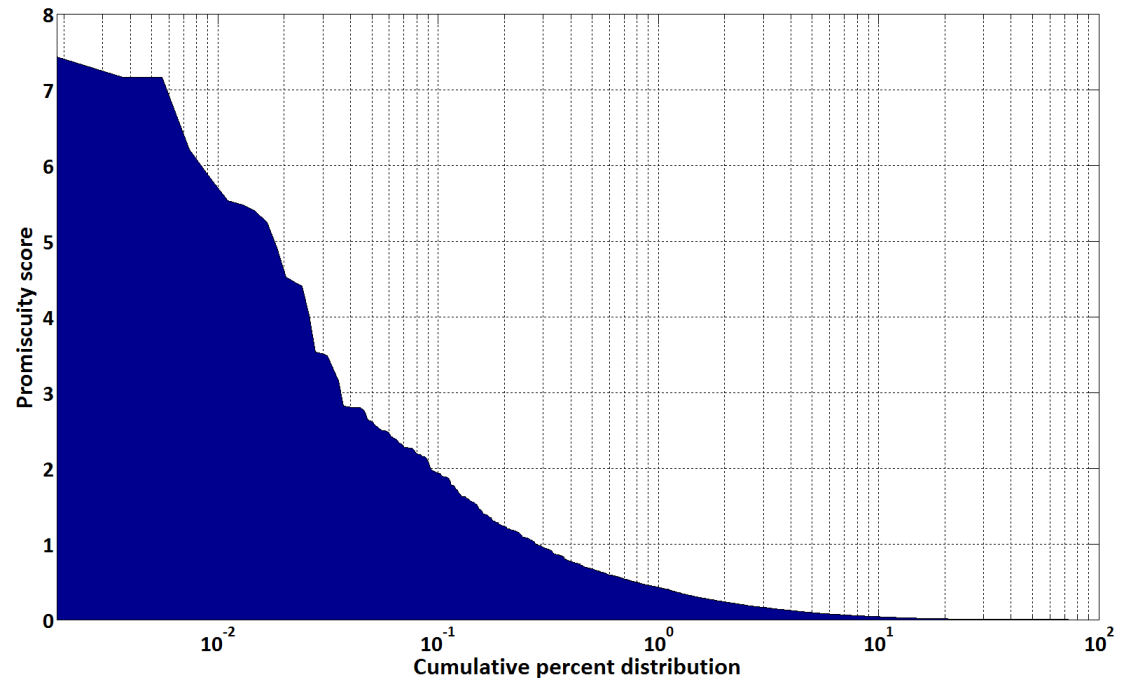


91/51
424/347
454/70
28318/454
0.91

tested substances/active substances
tested MMPS/active MMPS
tested assays/active assays
tested samples/active samples
promiscuity score

Promiscuity score distribution

- 651,299 unique single cut MMPs
- 777,931 unique double cut MMPs
- 73,191 unique triple cut MMPs
- 53,993 number of unique fragments



Tyrosine Phosphatase Inhibitor-3 Sensitizes Melanoma and Colon Cancer to Biotherapeutics and Chemotherapeutics

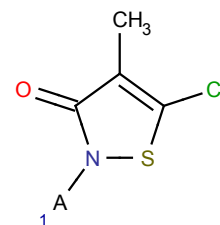
S:
S:

ELSEVIER

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry

journal homepage: www.elsevier.com/locate/bmc

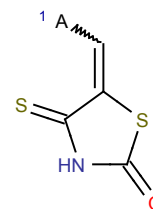


Synthesis of isothiazol-3-one derivatives as inhibitors of histone acetyltransferases (HATs)

Stephen G 6716 *J. Med. Chem.* 2009, 52, 6716–6723
 Michael Ja DOI: 10.1021/jm901016k

Cancer Research |

Journal of
**Medicinal
 Chemistry**
 Article



Multidentate Small-Molecule Inhibitors of *Vaccinia* H1-Related (VHR) Phosphatase Decrease Proliferation of Cervix Cancer Cells[†]

Shuangding Wu,^{†,§} Sofie Vossius,^{1,†} Soudad Rahmouni,^{†,‡} Ana V. Miletic,[§] Torkel Vang,[§] Jesus Vazquez-Rodriguez,[§]
 Fabio Cerignoli,[§] Yu
 Tomas Mustelin,[§] and

OPEN ACCESS Freely available online

PLoS one

The ERAD Inhibitor Eeyarestatin I Is a Bifunctional Compound with a Membrane-Binding Domain and a p97/VCP Inhibitory Group

Qiuyan Wang¹,
 Wiestner², Willi

¹Laboratory of Molecular
 America, ²Hematology &
 Biochemistry and Biology,
⁴The Genomics Core Lab
 America

1692 *J. Med. Chem.* 1996, 39, 1692–1703

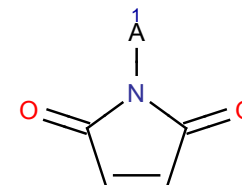
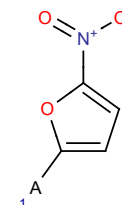
Design, Synthesis, and Biochemical Evaluation of *N*-Substituted Maleimides as Inhibitors of Prostaglandin Endoperoxide Synthases[†]

Amit S. Kalgutkar, Brenda C. Crews, and Lawrence J. Marnett*

A. B. Hancock, Jr., Memorial Laboratory for Cancer Research, Department of Biochemistry, Center in Molecular Toxicology and Vanderbilt Cancer Center, Vanderbilt University School of Medicine, Nashville, Tennessee 37232

Received November 29, 1995[§]

N-(Carboxyalkyl)maleimides are rapid as well as time-dependent inhibitors of prostaglandin endoperoxide synthase (PGHS). The corresponding *N*-alkylmaleimides were only time-dependent inactivators of PGHS, suggesting that the carboxylate is critical for rapid inhibition. Several *N*-substituted maleimide analogs containing structural features similar to those of the nonsteroidal anti-inflammatory drug aspirin were synthesized and evaluated as inhibitors of PGHS. Most of the aspirin-like maleimides inactivated the cyclooxygenase activity of purified ovine PGHS-1 in a time- and concentration-dependent manner similar to that of aspirin. The peroxidase activity of PGHS was also inactivated by the maleimide analogs. The cyclooxygenase



Conclusions

- Automated tool to identify promiscuous patterns in a database of biological activities
- Uses ChemAxon JChem/Marvin API for:
 - I/O; fragmentation; SMARTS match/generation; SMIRKS generation; chemical structure/patterns depiction
- Filter unwanted HTS hits
- Questions?