ZINC Is … Not Commercial

John Irwin
University of California San Francisco
#chemdbopenchem 2011-08-26

ZINC is … A Team

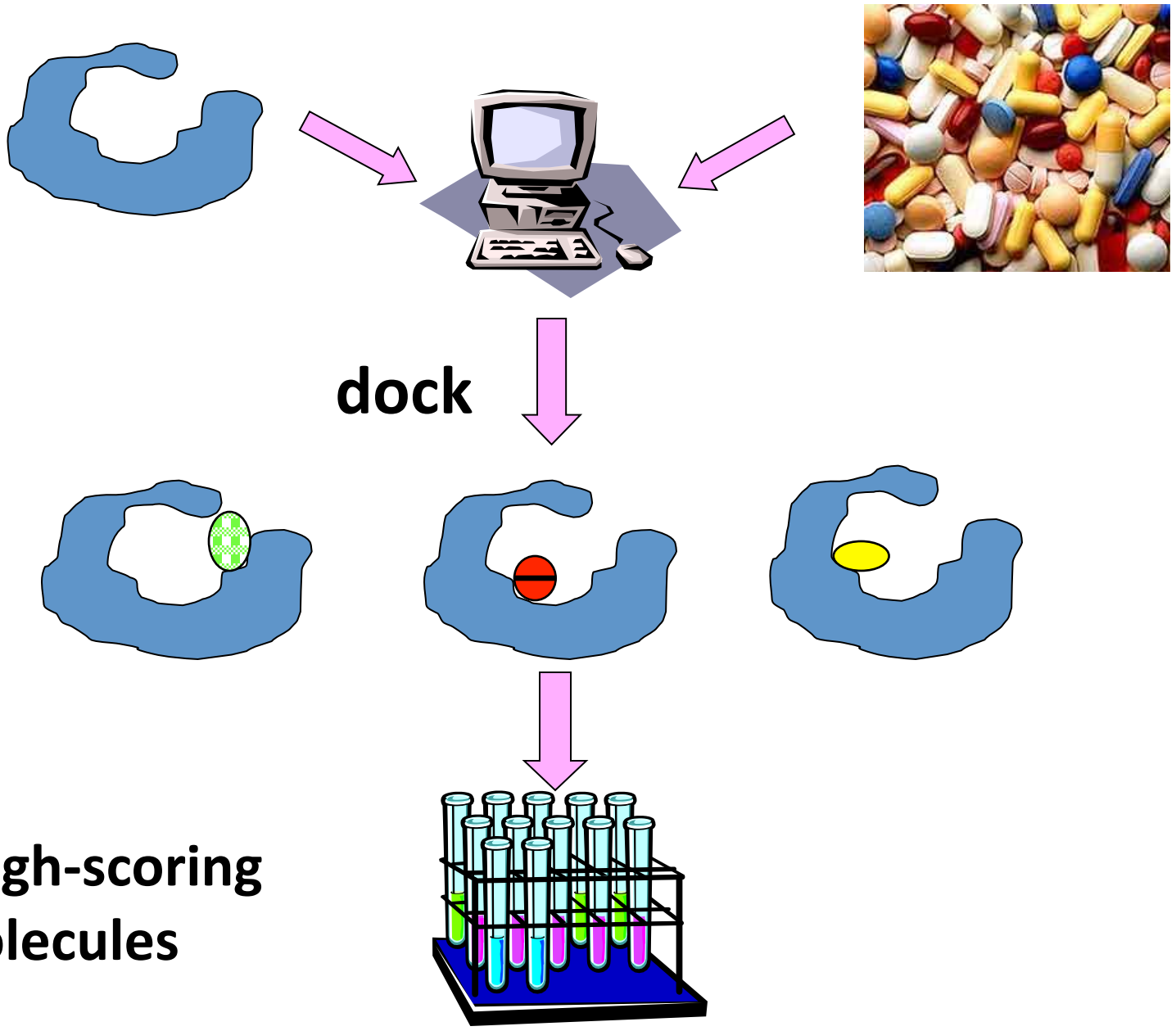
Teague Sterling

Erin Bolstad

Shoichet Lab for support, NIH for funds

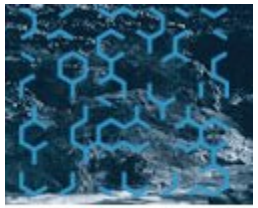
ZINC is … Chemistry for Biology

ZINC is … Focused on Docking



ZINC is … Purchasable

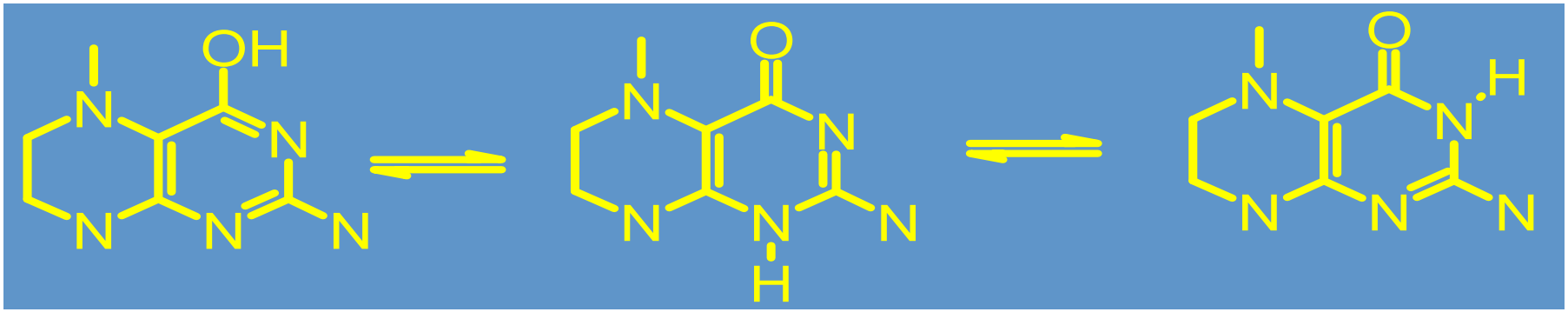
Key Organics



Oakwood

ZINC is … Biologically Relevant

- 3D
  - Enumerate up to 4 implied stereoisomers
- Protonation/charge
  - Histidine is 2 molecules: charge=0 and charge=1
  - pH-dependent subsets available
- Tautomers



Relevant tautomers enumerated

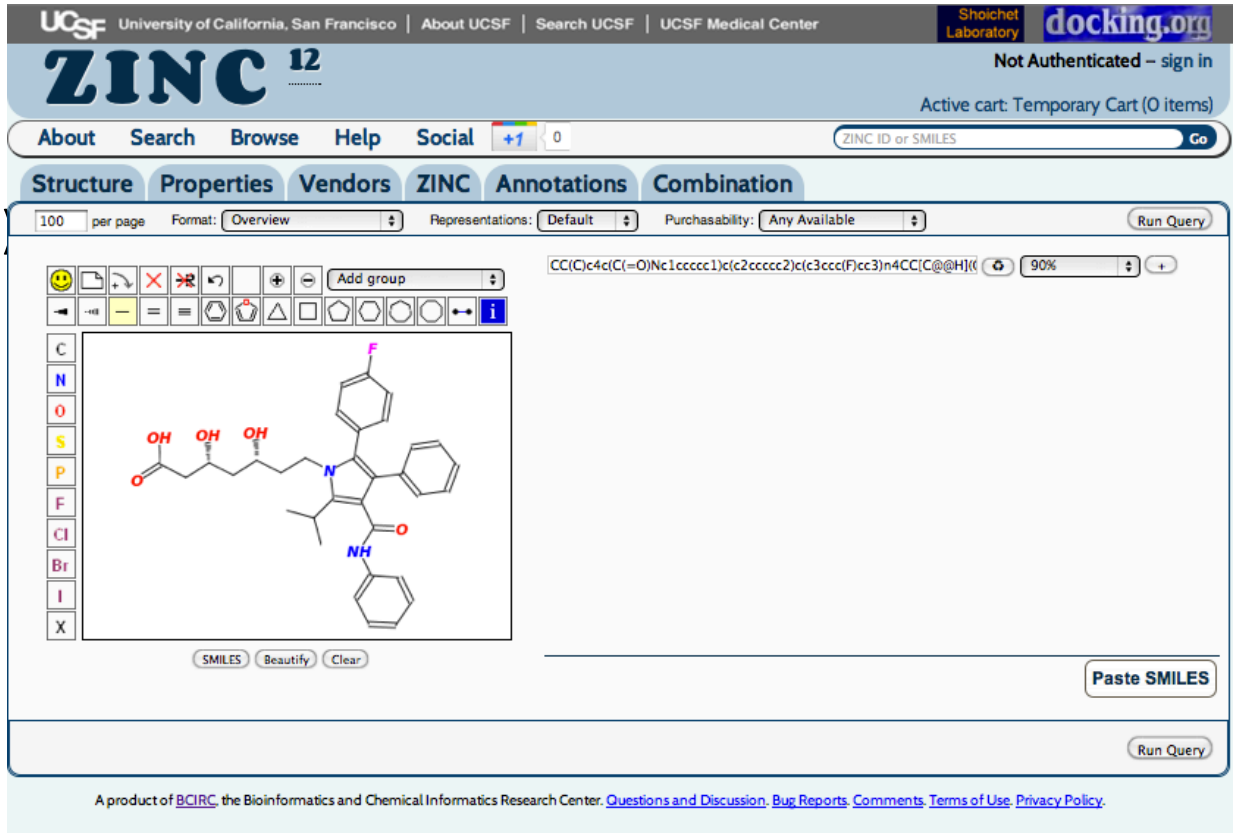
ZINC is … Pragmatic

20+ million compounds
- commercially available
- structures calculated
- multiple conformations
- properties (charge, solv, etc…)
- links to suppliers

Free to the community

Multiple subsets
- 8.8 M drug-like (Lipinski)
- 3.4 M lead-like (Oprea…)
- 450 K fragment-like (Astex, …)

Availlable in popular formats
- SMILES, SDF, mol2, flexibase



Irwin & Shoichet JCIM 2005

Updated continuously (10,000 new today)
Over 2 million new compounds per year
Over 1 million depletions per year

ZINC is NOT … a procurement agent

ZINC is NOT … encyclopedic

But these are all in ZINC

Oxymoronecules

A free database for virtual screening

ZINC is not commercial

Home | Subsets | HELP! | Mailing Lists

Oxymoronecules are purchasable compounds from normally catalogs that are themselves non-purchasable, such as FDA, ChEMBL, etc. These databases have been provided to us under collaborative agreements with each respective vendor. Molecules in the original catalogs may not appear in ZINC because they do not pass our filters, or they have not been processed yet. ZINC may be used free of charge for research by individuals and institutions. More about ZINC subsets is here. **Whereas you are free to share the results of a ZINC search or a screen of molecules from ZINC, you may not redistribute major portions of ZINC without the express written permission of John Irwin.** Additional usage notes may be found below the table.

Logo (click to browse ZINC)	Vendor Information Name Email Phone Fax	Catalog Information Source entries Version/Date	ZINC Information Loaded Filtered Sole source Depleted	Diversity Information Cluster representatives at a given Tanimoto cutoff
BindingDB	BindingDB.org none p.not for sale f.not for sale	287799 2011-05-04	49719 5066 30652 NYA	90%:22421 80%:13853 70%:8673 60%:4642
ChEBI	ChEBI no email p.no phone f.no fax	19640 2011-06-27	5302 2835 369 NYA	90%:N/A 80%:N/A 70%:N/A 60%:N/A
ChEMBL	ChEMBL10 no email p.no phone f.no fax	999881 2011-07-14	452655 63394 NYA	90%:N/A 80%:81530 70%:31290 60%:10127
DrugBank Experimental	DrugBank-experimental not for sale p.phone f.not for sale	4854 2011-02-15	3438 897 122 NYA	90%:1685 80%:1334 70%:1022 60%:717

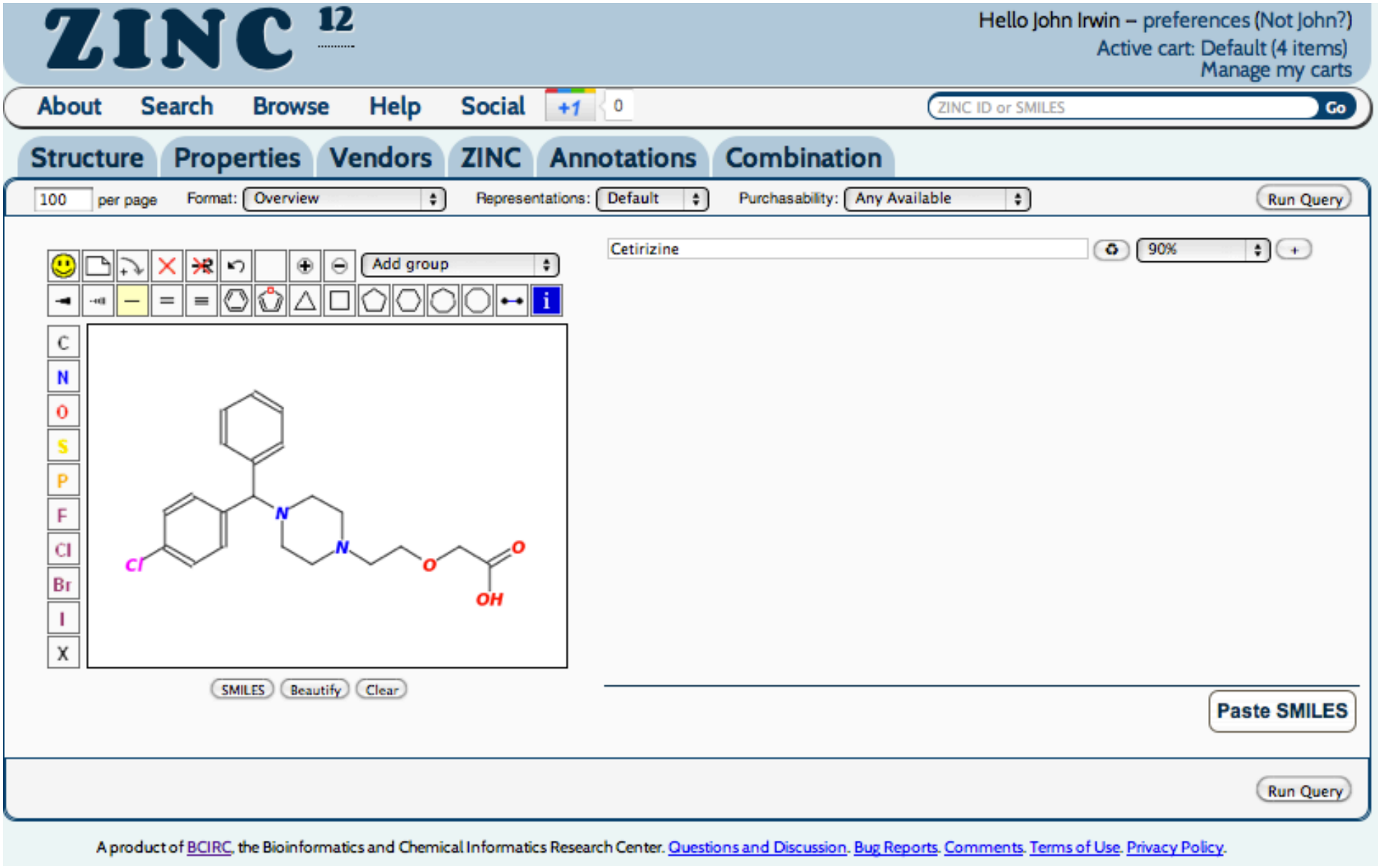
ZINC is … Maintained

- Continuously updated
- Typically
  - 85% purchase success for “stock”
  - 65% purchase success for “make on demand”
- In the last 6 months:
  - 50% of vendors fully updated
  - 80% of subsets refreshed
- 8 years old and going strong (NIH GM071896)

1. Start with a compound

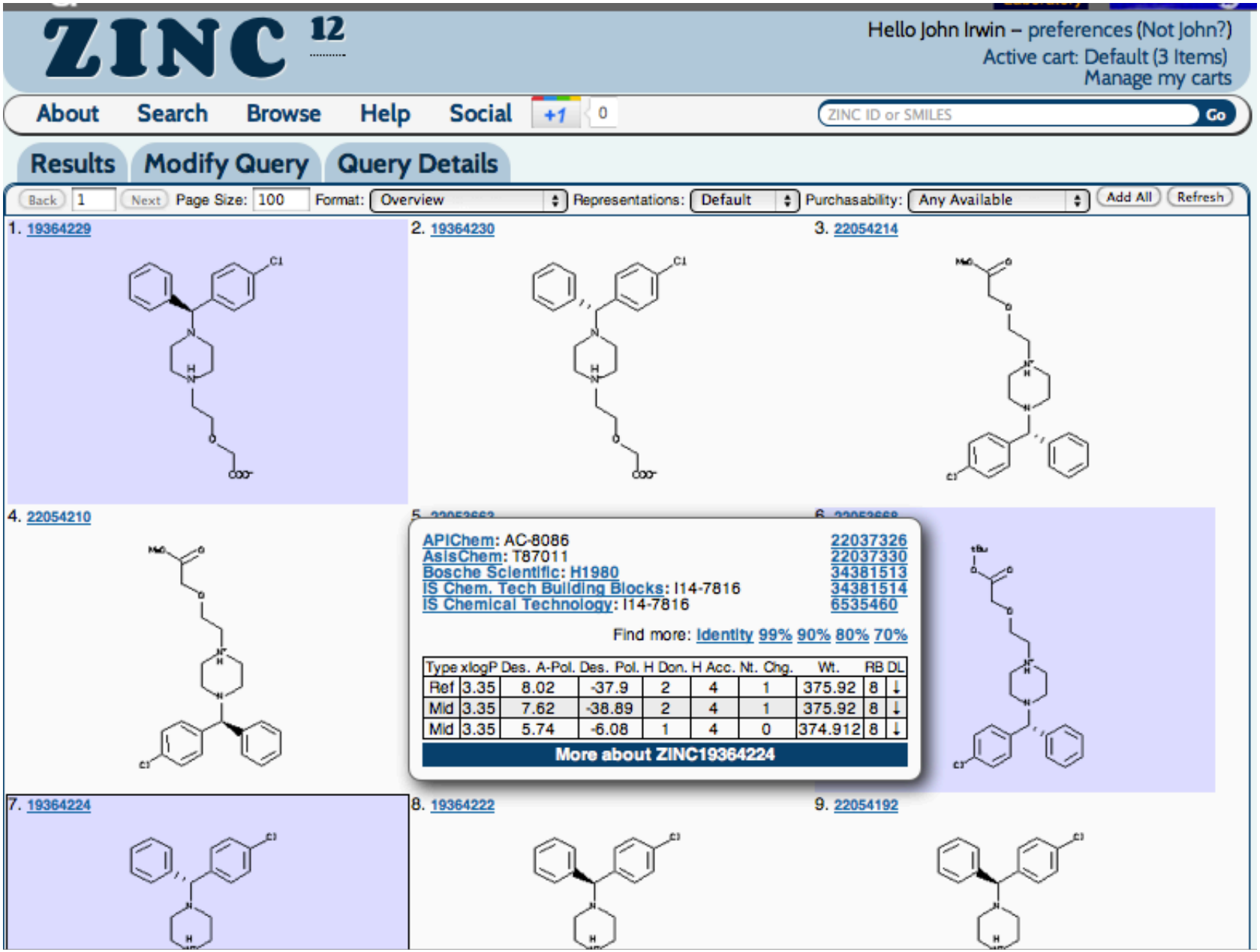
What analogs can I buy?

Start with a drug: Cetirizine



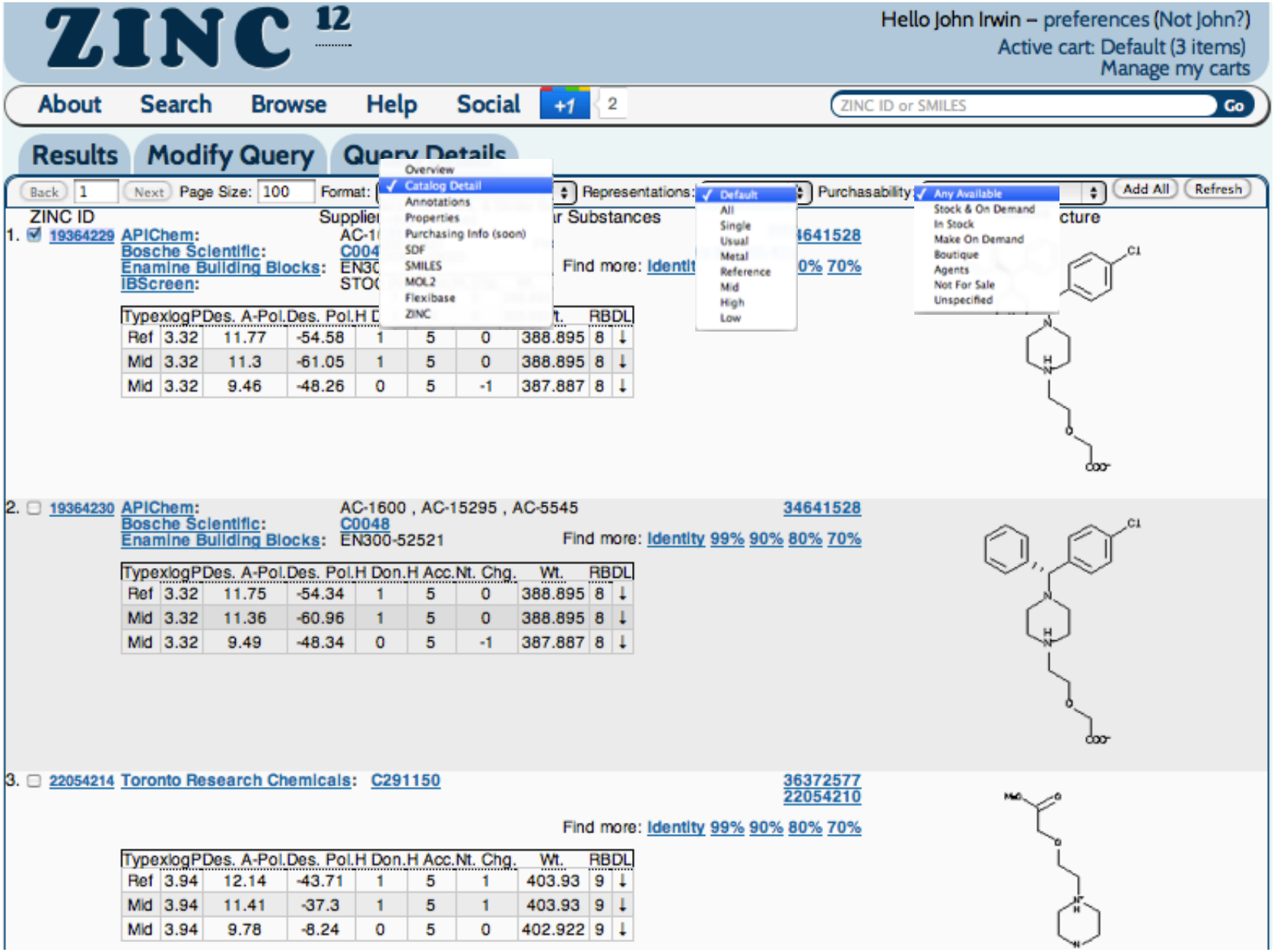
Case 1: Start with a drug: Cetirizine

Save molecules to shopping cart



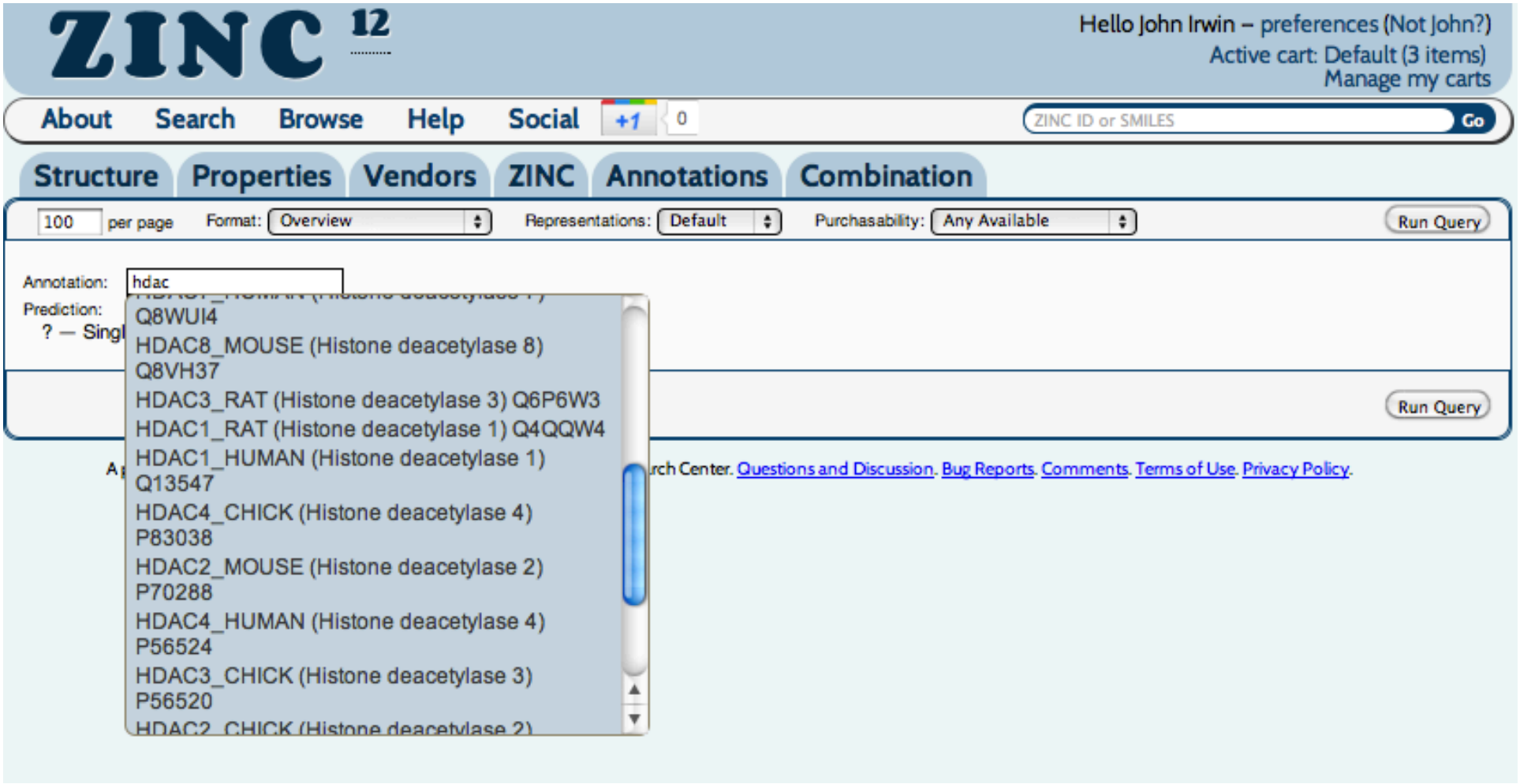
Case 1: Start with a drug: Cetirizine

Select report format options



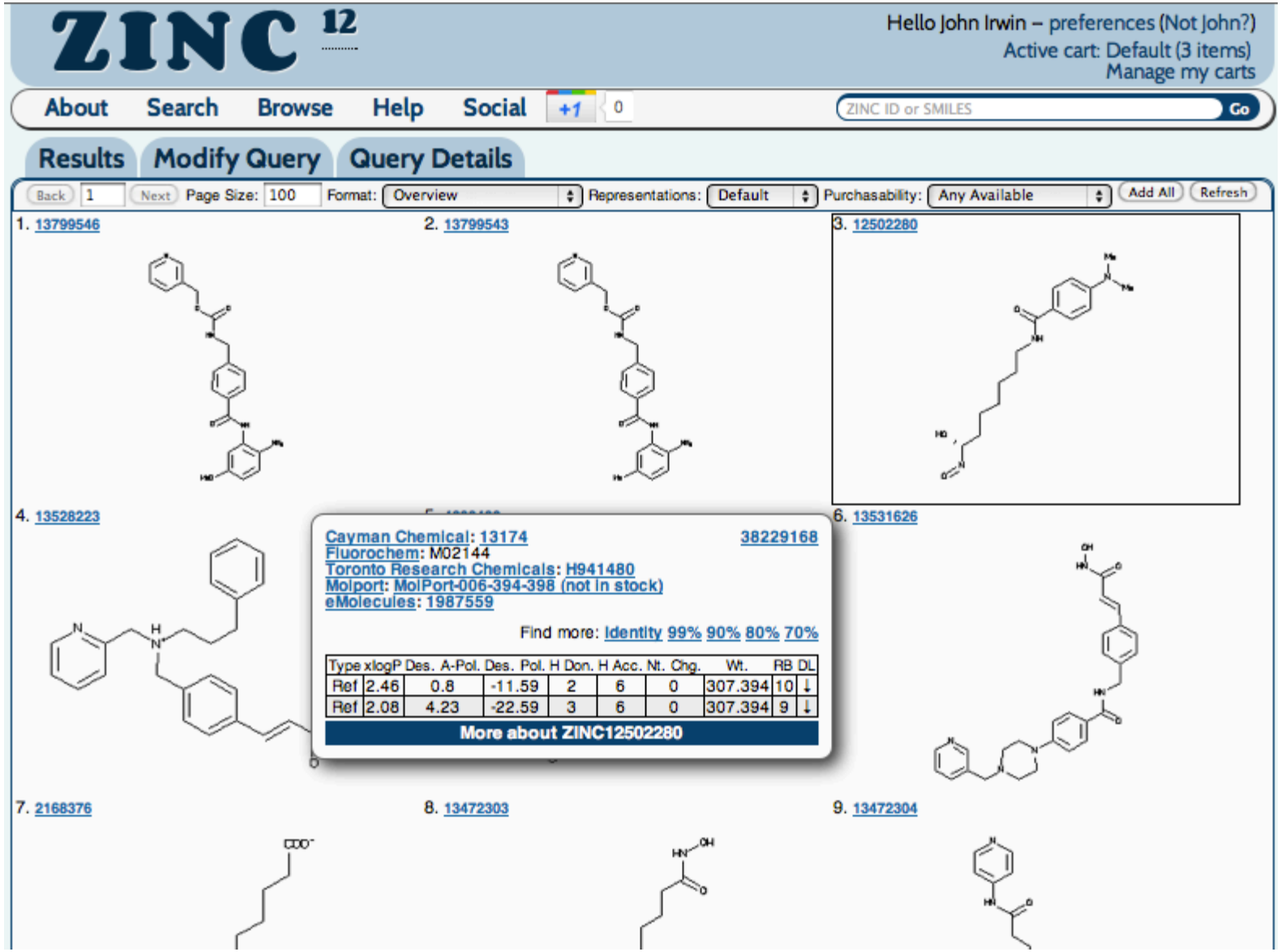
2. Start with a target

Start from a target: HDACs

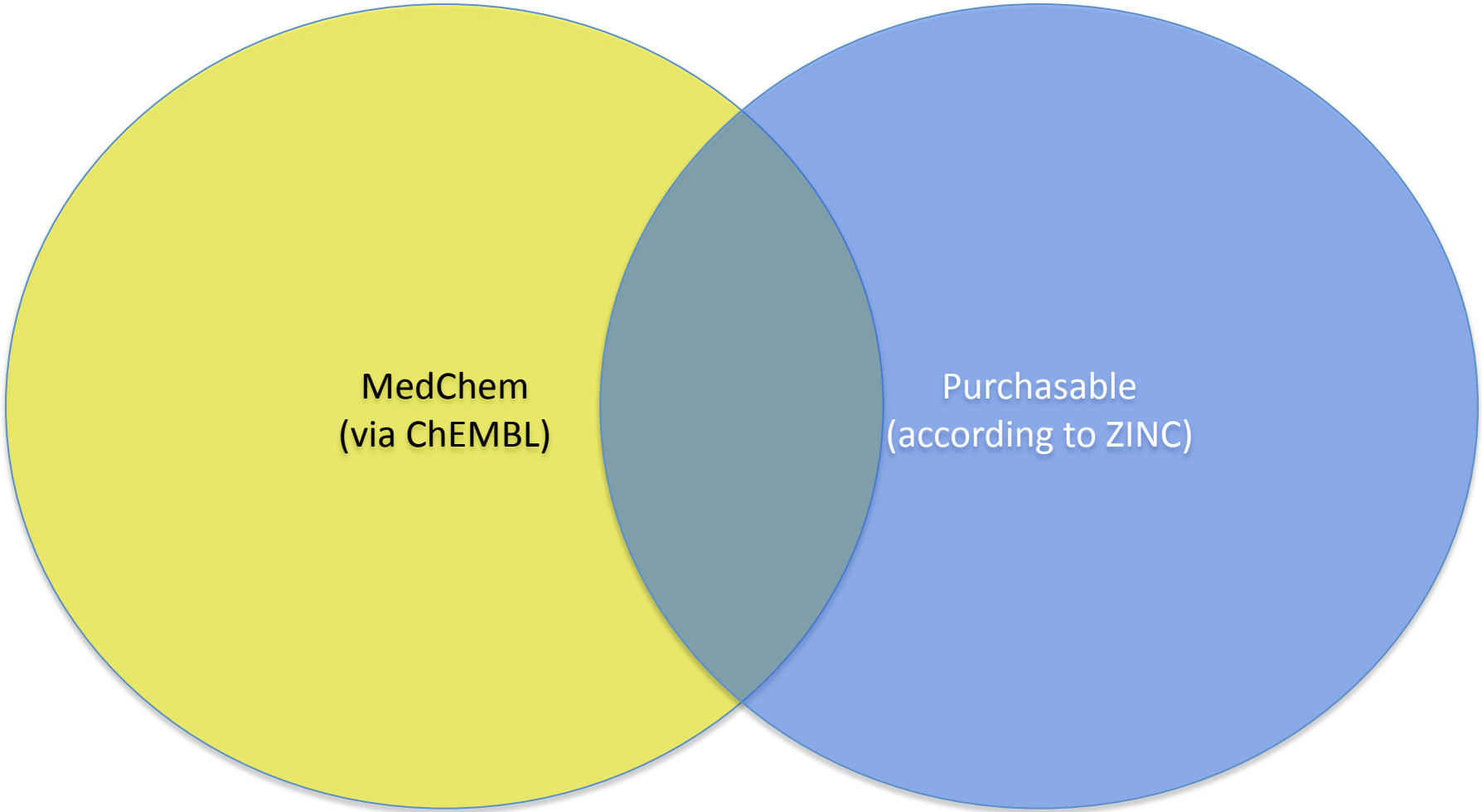


Case 2: Start from a target: HDACs

HDAC4 compounds you can buy



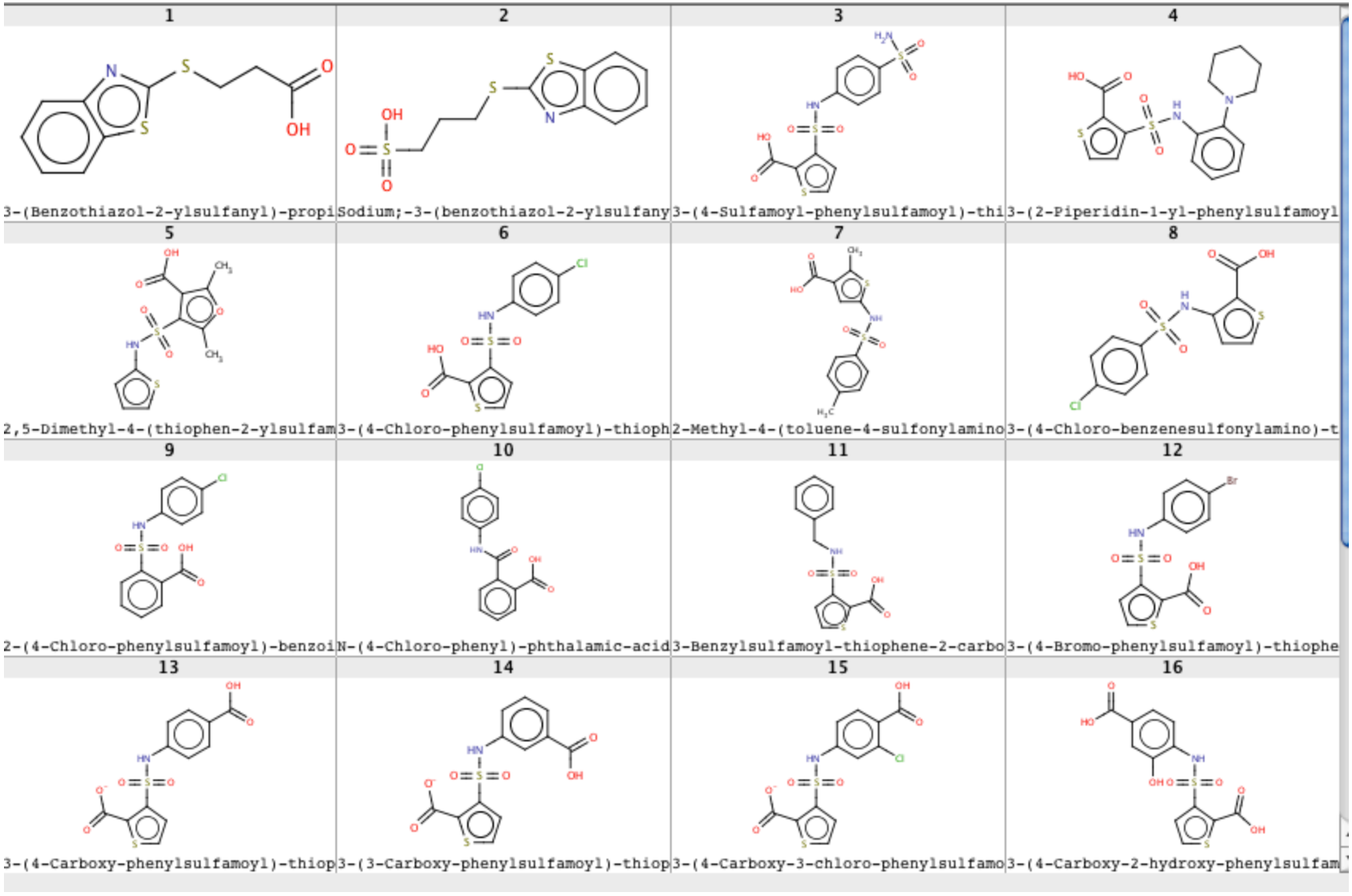
Purchasable Med Chem Space



3. Start with a set of actives

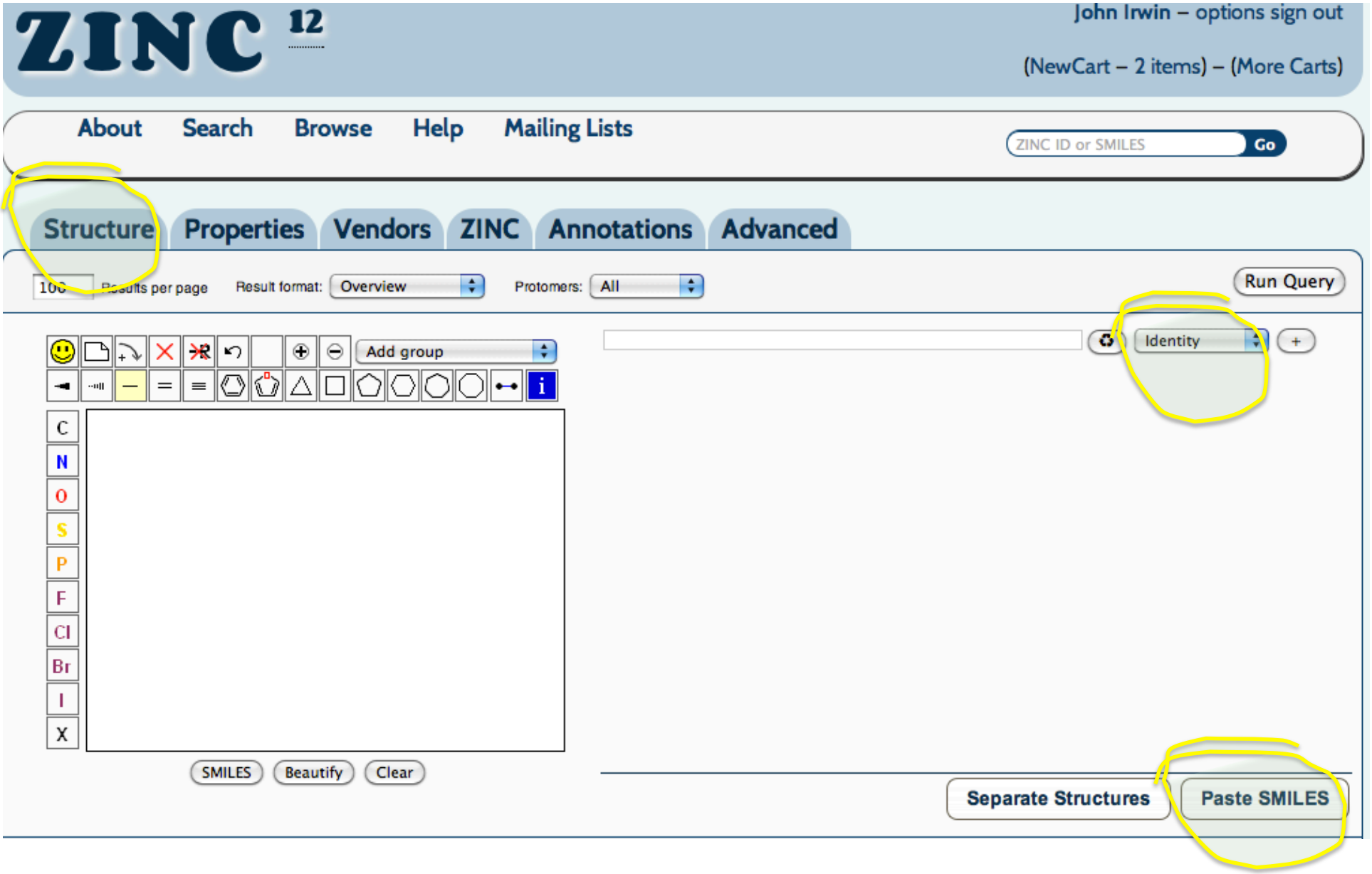
Case 3: Start from a set of actives: AmpC ligands

AmpC ligands: pictures



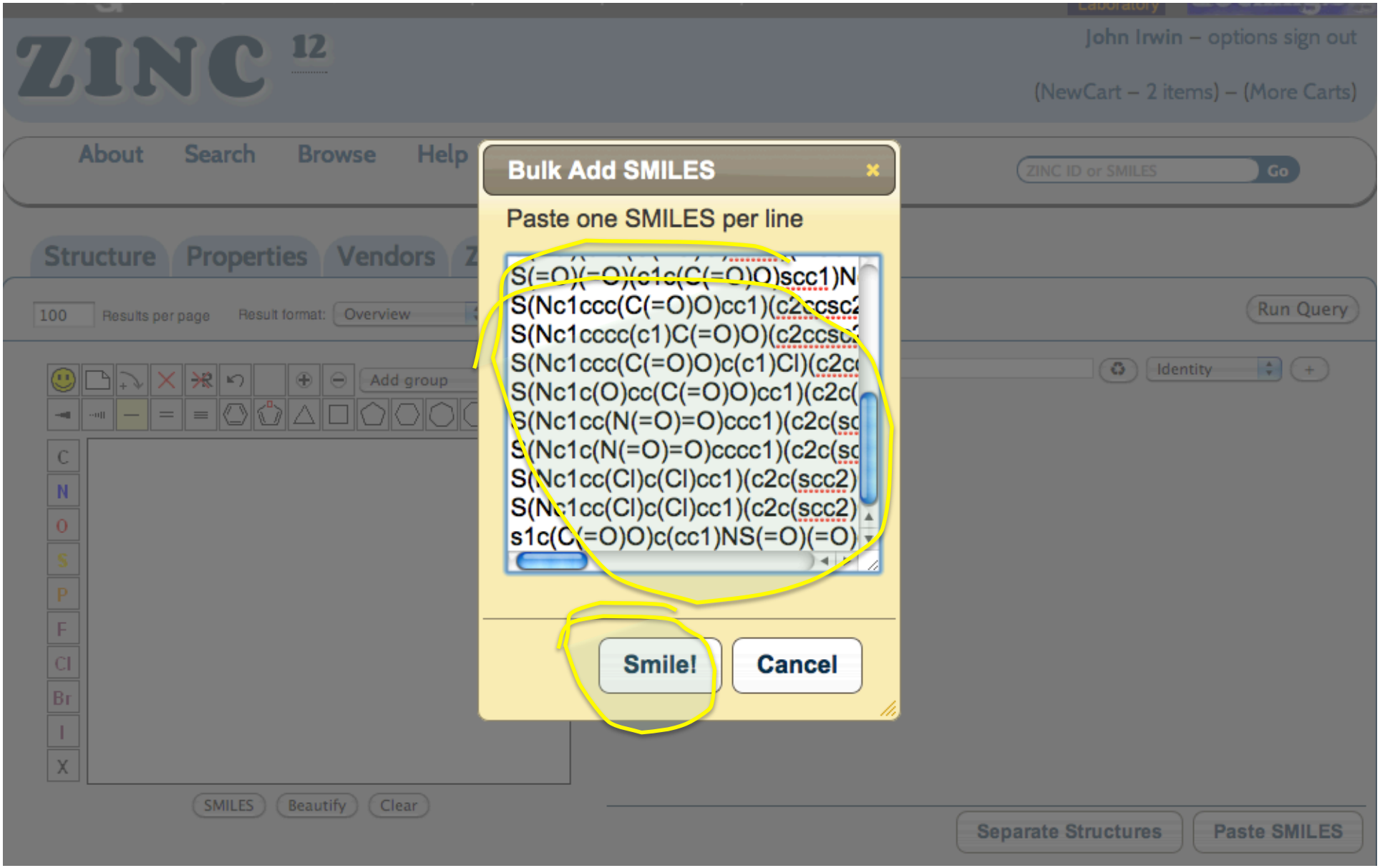
Case 3: Start from a set of actives: AmpC ligands

Look for a perfect match



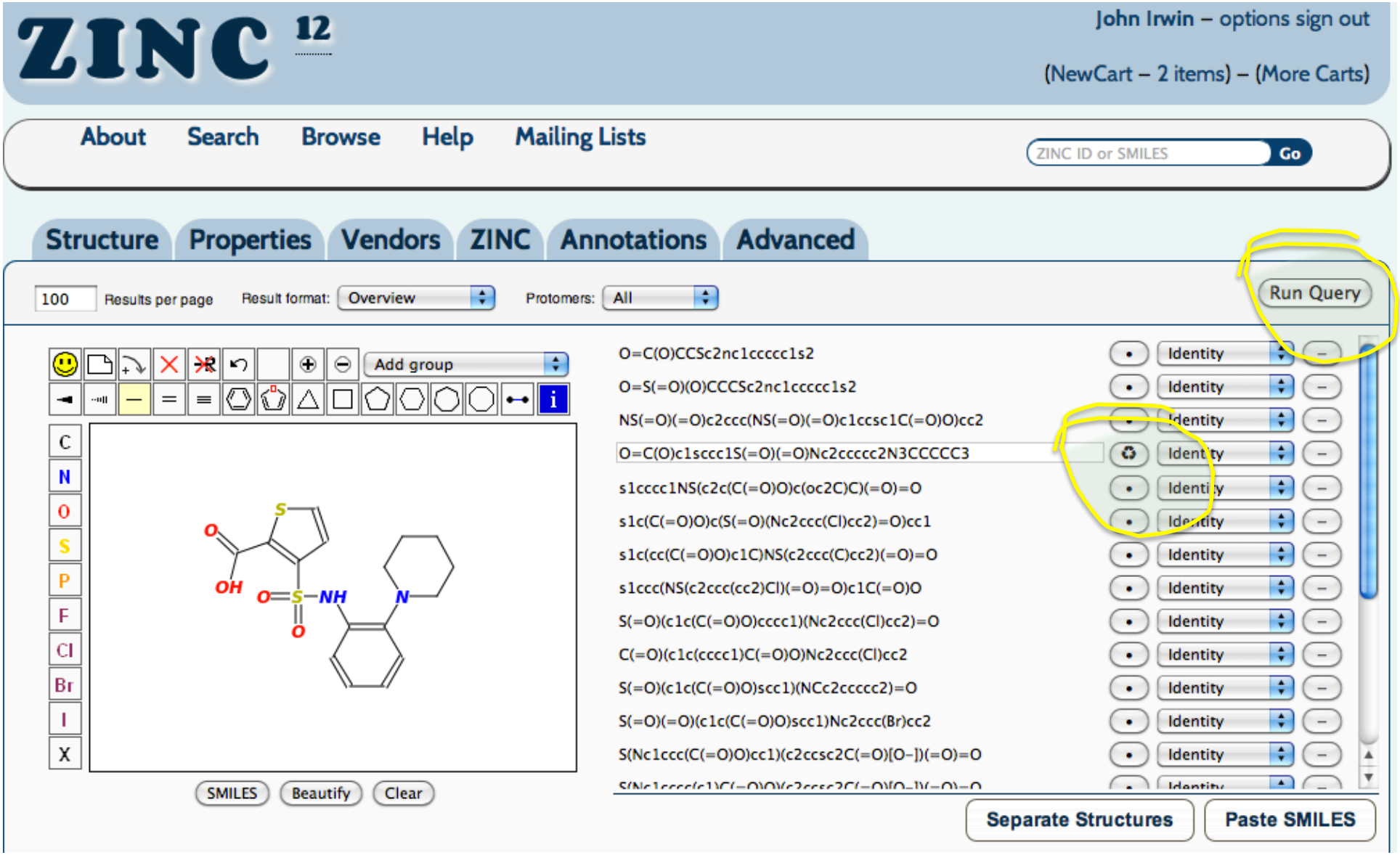
Case 3: Start from a set of actives: AmpC ligands

Paste SMILES



Case 3: Start from a set of actives: AmpC ligands

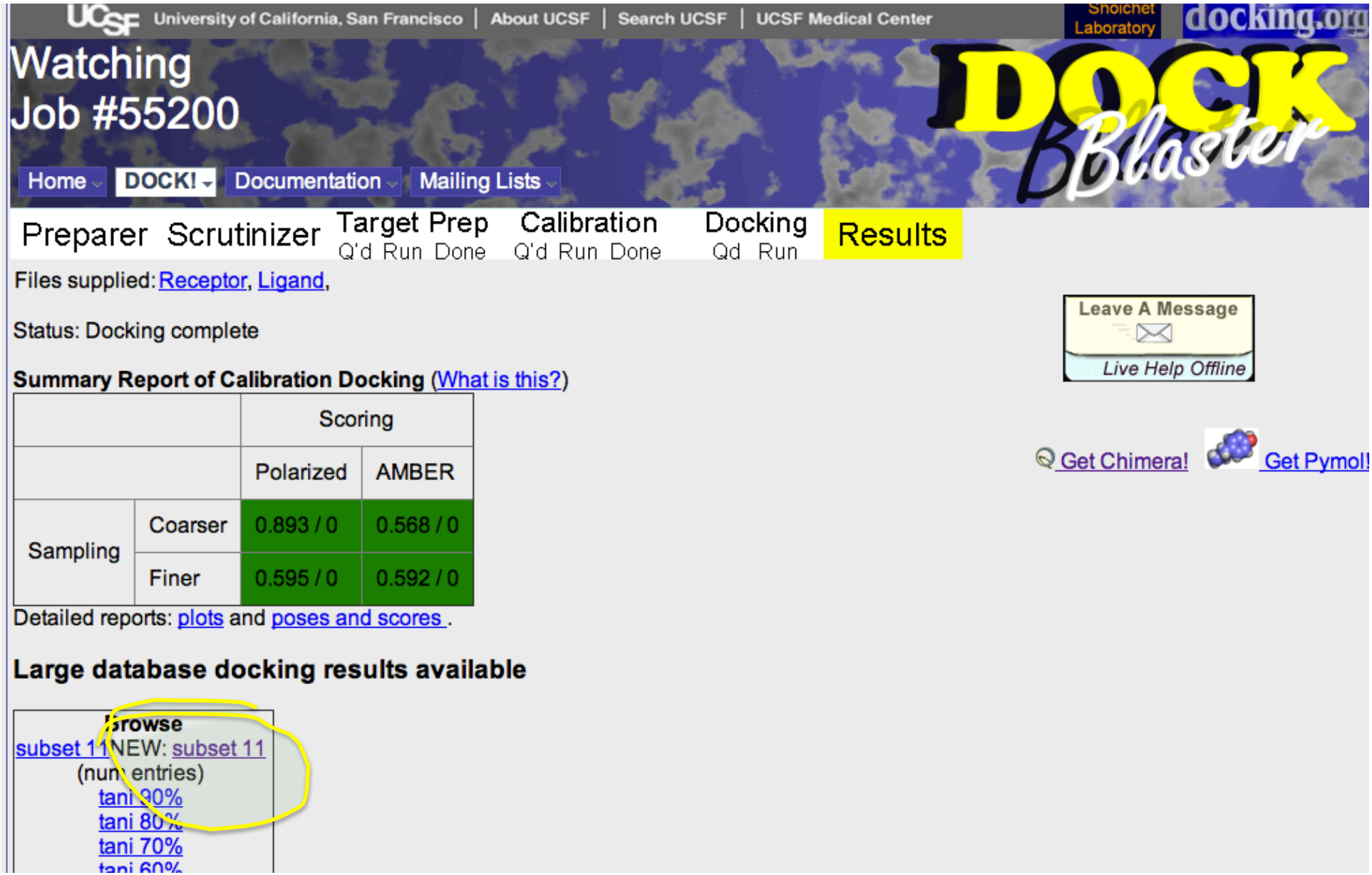
Review uploaded SMILES



4. Start with docking

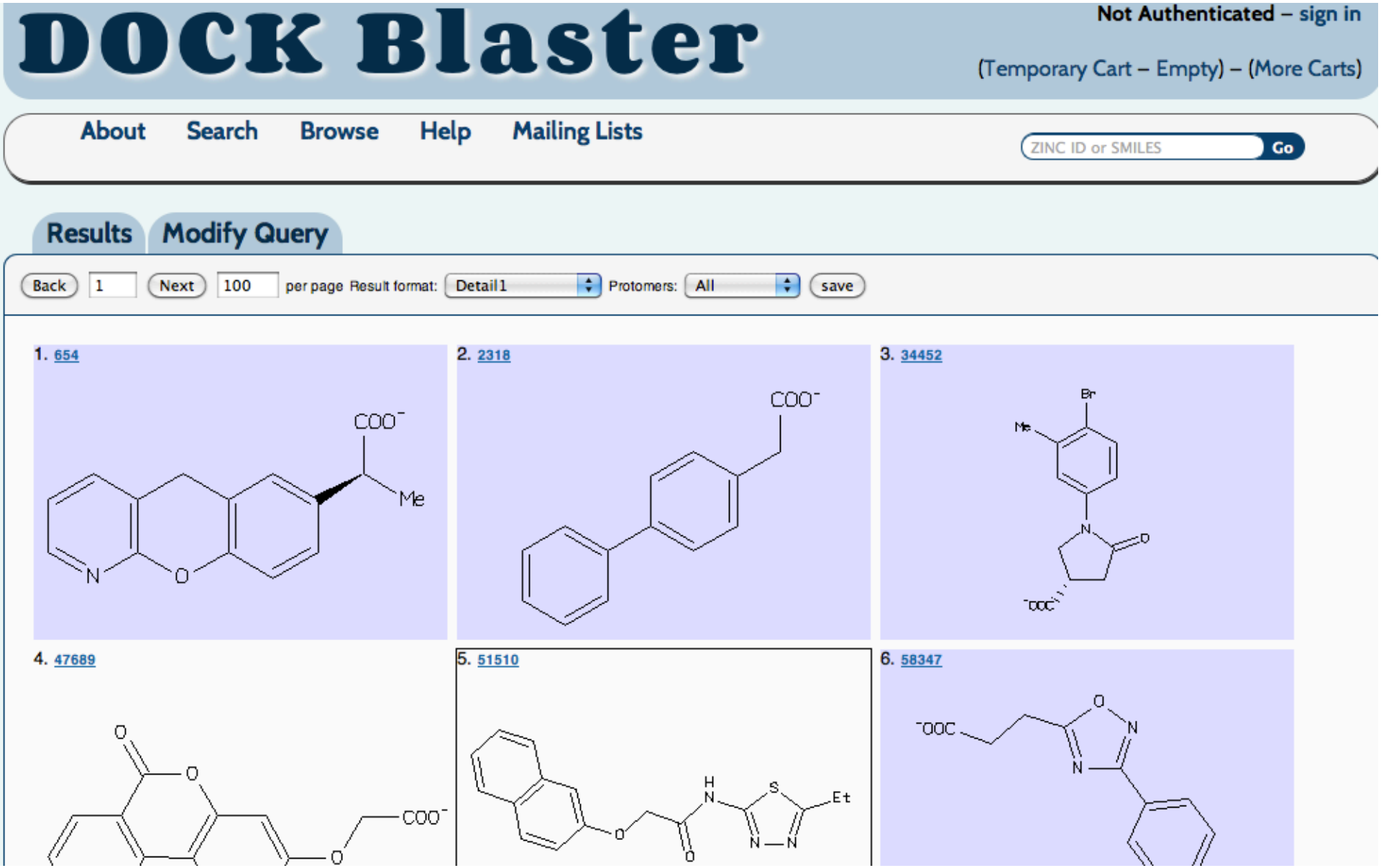
Case 4: Start with docking

Docking results

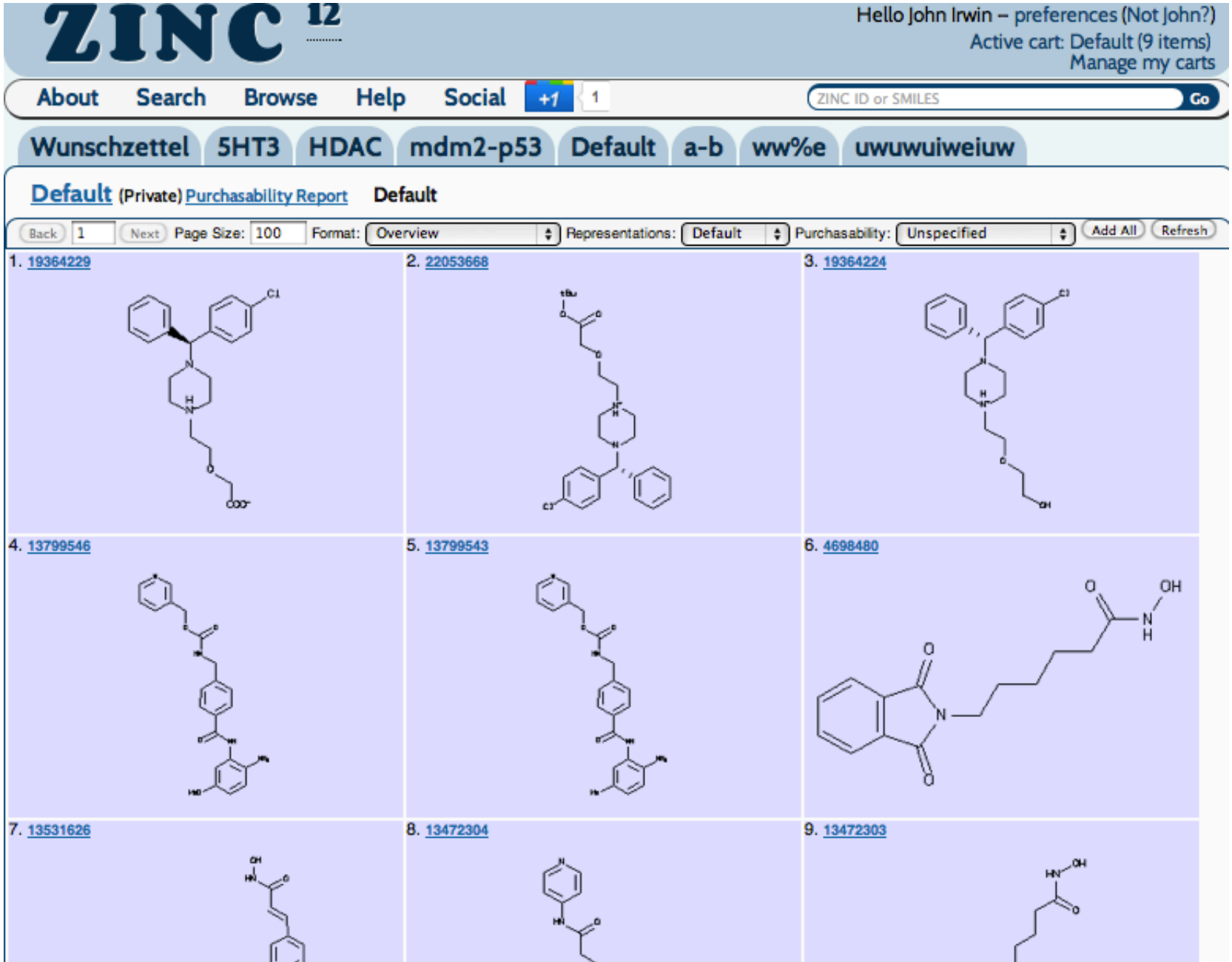


Case 4: Start with docking

Docking results



Shopping cart

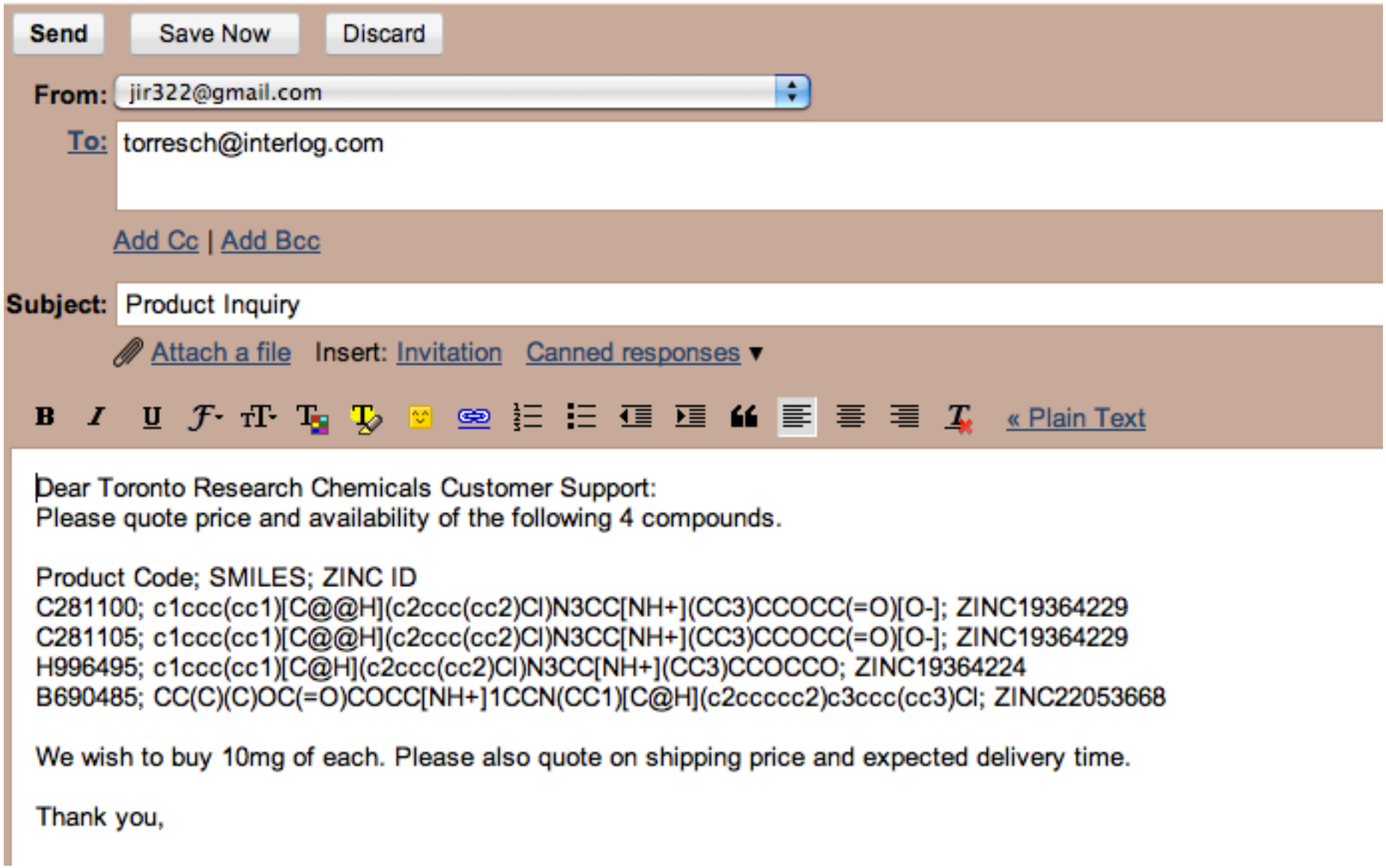


Purchasability Report

Download

Catalog ZINC ID	Available Supplier Code(s)	Email SMILES
Toronto Research Chemicals	3	Email Inquiry torresch@interlog.com
19364229	C281100;C281105	c1ccc(cc1)[C@@H](c2ccc(cc2)Cl)N3CC[NH+](CC3)CCOCC(=O)[O-]
19364224	H996495	c1ccc(cc1)[C@H](c2ccc(cc2)Cl)N3CC[NH+](CC3)CCOCCO
22053668	B690485	CC(C)(C)OC(=O)COCC[NH+]1CCN(CC1)[C@H](c2ccccc2)c3ccc(cc3)Cl
eMolecules	3	Email Inquiry info@emolecules.com
19364224	27521121;539401	c1ccc(cc1)[C@H](c2ccc(cc2)Cl)N3CC[NH+](CC3)CCOCCO
19364229	25035671;11021211;532399	c1ccc(cc1)[C@@H](c2ccc(cc2)Cl)N3CC[NH+](CC3)CCOCC(=O)[O-]
22053668	25035441	CC(C)(C)OC(=O)COCC[NH+]1CCN(CC1)[C@H](c2ccccc2)c3ccc(cc3)Cl
APAC Sourcing Solutions	2	Email Inquiry anusuabag@apacss.com
19364224	812046	c1ccc(cc1)[C@H](c2ccc(cc2)Cl)N3CC[NH+](CC3)CCOCCO
19364229	550560	c1ccc(cc1)[C@@H](c2ccc(cc2)Cl)N3CC[NH+](CC3)CCOCC(=O)[O-]
APIChem	2	Email Inquiry sales@apichemistry.com
19364229	AC-5545;AC-1600	c1ccc(cc1)[C@@H](c2ccc(cc2)Cl)N3CC[NH+](CC3)CCOCC(=O)[O-]
19364224	AC-8086	c1ccc(cc1)[C@H](c2ccc(cc2)Cl)N3CC[NH+](CC3)CCOCCO
Ambinter	2	Email Inquiry contact@ambinter.com
19364224	SBB012597	c1ccc(cc1)[C@H](c2ccc(cc2)Cl)N3CC[NH+](CC3)CCOCCO
19364229	PHAR056425;BB_SC-2161;ST077146	c1ccc(cc1)[C@@H](c2ccc(cc2)Cl)N3CC[NH+](CC3)CCOCC(=O)[O-]
Bosche Scientific	2	Email Inquiry sales@boschesci.com
19364224	H1980	c1ccc(cc1)[C@H](c2ccc(cc2)Cl)N3CC[NH+](CC3)CCOCCO
19364229	C0048	c1ccc(cc1)[C@@H](c2ccc(cc2)Cl)N3CC[NH+](CC3)CCOCC(=O)[O-]
IS Chemical Technology	2	Email Inquiry sales@ispharm.com
19364229	I06-0389;I14-0479	c1ccc(cc1)[C@@H](c2ccc(cc2)Cl)N3CC[NH+](CC3)CCOCC(=O)[O-]
19364224	I14-7816	c1ccc(cc1)[C@H](c2ccc(cc2)Cl)N3CC[NH+](CC3)CCOCCO
Prestwick Chemical	2	Email Inquiry
19364229	Prestw-503	c1ccc(cc1)[C@@H](c2ccc(cc2)Cl)N3CC[NH+](CC3)CCOCC(=O)[O-]
19364224	Prestw-133	c1ccc(cc1)[C@H](c2ccc(cc2)Cl)N3CC[NH+](CC3)CCOCCO
Princeton BioMolecular Research	2	Email Inquiry info@princetonbio.com

Email vendor



Send | Save Now | Discard

From: jir322@gmail.com

To: torresch@interlog.com

Add Cc | Add Bcc

Subject: Product Inquiry

Attach a file Insert: Invitation Canned responses ▾

« Plain Text

Dear Toronto Research Chemicals Customer Support:
Please quote price and availability of the following 4 compounds.

Product Code; SMILES; ZINC ID
C281100; c1ccc(cc1)[C@@H](c2ccc(cc2)Cl)N3CC[NH+](CC3)CCOCC(=O)[O-]; ZINC19364229
C281105; c1ccc(cc1)[C@@H](c2ccc(cc2)Cl)N3CC[NH+](CC3)CCOCC(=O)[O-]; ZINC19364229
H996495; c1ccc(cc1)[C@H](c2ccc(cc2)Cl)N3CC[NH+](CC3)CCOCCO; ZINC19364224
B690485; CC(C)(C)OC(=O)COCC[NH+]1CCN(CC1)[C@H](c2ccccc2)c3ccc(cc3)Cl; ZINC22053668

We wish to buy 10mg of each. Please also quote on shipping price and expected delivery time.

Thank you,

ZINC Facilitates Ligand Discovery

- Find compounds to buy starting from
  - A compound
  - A target (med chem literature via ChEMBL)
  - A set of hits
  - Docking

- Features
  - Shopping Carts (Create, maintain, share)
  - Prioritize and purchase docking hits
  - One click™* ordering