

**CHEMISTRY QUALITY AND THE MEDICINAL CHEMISTRY - BIOLOGY INTERFACE.** Lipinski, C. A.\* Pfizer Global Research and Development, Groton New London Labs (retired), Eastern Point Road, mail stop 8200-36, Groton, CT 06340, Lipinskica@groton.pfizer.com.

High throughput screening (HTS) works because it is a process and not simply a screen. The “apparent” quality of an individual HTS screen superficially may appear inferior to that of a traditional biology experiment. But in fact the inherent biology quality in an HTS screen is almost always greater than in a traditional biology experiment. The HTS “apparent quality” is the product of three terms: 1) the inherent biology quality which is always high for an HTS screen with an acceptable z factor; 2) the stochastic false positive rate which simply results from the statistics of screening a lot of compounds and 3) a term which relates to errors in the infrastructure of compound distribution and to false positives that are inherent to the chemical structure being screened.

The third term is the most troublesome and one where the medicinal chemistry biology interface is particularly important. False positive errors related to biology are the easiest to detect because the false positives do not reproduce. False positive errors related to chemistry do reproduce because the apparent positive signal is related to the chemical structure per se. A successful HTS process uses a staged triage of experiments and computation to whittle the initial positives down to set of apparently active hits that are numerically small enough to be examined by an experienced medicinal chemist. But why is the medicinal chemist so good at picking out the garbage structures from the potential winners? Is it intuition or something else? And how does the medicinal chemist explain this process to the biologist in such a manner as to retain a productive cooperative team environment?

Virtually without exception HTS screens are run industry wide without knowing compound concentration. This can result in scrambled IC-50's and messed up selectivity conclusions. A big component in this is compound precipitation from DMSO stock solutions. The chemist has an important role here. Chemists know how to keep reagents dry. By contrast biologists run everything in water. DMSO used in the distribution of compound stocks should be bone dry. Even a little water in the DMSO used in compound stocks is really really bad from a precipitation viewpoint. This happens because DMSO and water exhibit very non-ideal behavior. A little bit of water in DMSO does not reduce compound solubility by a little bit. The little bit of water in DMSO reduces solubility by a whole lot. So DMSO used in compound distribution should be kept extremely dry and treated as if it were butyl lithium in hexane.

# CHEMISTRY QUALITY AND THE MEDICINAL CHEMISTRY - BIOLOGY INTERFACE.

Christopher A. Lipinski

Adjunct Senior Research Fellow

Pfizer Global R&D (retired)

Lipinskica@groton.pfizer.com

# HTS “apparent” quality

- Apparent HTS quality is the product of 3 factors
- 1) Inherent biology quality
  - very good in an HTS
- 2) Stochastic error
  - primarily in biology
  - statistical effect
- 3) Compound distribution infrastructure
  - Chemistry related false positives

# Inherent HTS biology quality

- HTS does not mean sloppy biology
- Traditional biology  $p = 0.05$ 
  - 50,000 false positives in a 1 M HTS
- HTS 0.5% hit rate
  - 5,000 false positives in a 1 M HTS
- Good HTS with z factor  $> 0.5$ 
  - good understanding of sd of negative controls
  - good understanding of sd of positive controls

# Biology and Chemistry False Positives

- Biology False Positive
- Assay related error
- Compound independent
- Stochastic. Random vs. order of testing
- Non reproducible
- Hits will fail in retest
- Chemistry False Positive
- Compound related error
- Compound dependent
- Non random vs. order of testing
- Reproducible
- Hits will confirm

# Biology False Positives (Stochastic)

- Liquid handling errors
- Reader errors
- Reagent errors
- Temperature errors
- Time errors
- Plate based errors

# Chemistry False Positives

- Promiscuous aggregators
- Covalent modifiers
- Detergents
- Optical interferers
- Reducing agents
- Cytotoxic agents
- Chelators

# False Positive Biology Error

- 1,000,000 compounds tested
- 0.1% false positives
- 1000 apparent “hits”
- Retest
- 0.1% false positives
- 1 apparent “hit”
- Reconfirm rate is  $1/1000 = 0.1\%$
- Sounds bad
- But very easy to weed out false positives

# False Positive Chemistry Error

- 1,000,000 compounds tested
- 0.1% false positives
- 1000 apparent “hits”
- Retest
- 1000 apparent “hits”
- Reconfirm rate is  $1000/1000 = 100\%$
- Sounds wonderful
- But very difficult to weed out false positives

# Trends in screening library quality

- Less than 20% of commercially available compounds are drug-like
- 3.5% of virtual combichem compounds are drug-like
- Wyeth                      66% of screening library junked
- AstraZeneca            50% of screening library junked
- Pfizer more than 600,000 compounds junked

# Library Cleanup Rationale

- Remove chemistry false positives
- HTS is a process NOT just a screen
- Primary HTS
- Retest
- Computational triage
- Medicinal chemistry pattern match triage
- Secondary assay triage

# Consequences of a dirty library

- For a very difficult screen
- No true positives
- Dirty screening library of chemistry false positives
- Primary HTS gives “hits”
- Retest HTS gives “hits”
- 100% retest rate
- Naive screener is happy
- Chemist is unhappy

# Consequences of library cleanup

- For a very difficult screen
- No true positives
- Clean library with no chemistry false positives
- Primary HTS gives biology false positive “hits”
- Retest HTS gives almost no “hits”
- Close to 0% retest rate
- Naive screener is unhappy
- Chemist is happy

# Information screeners exist

- Academics
- Goal is to renew research funding
- Work product is a publication
- Reproducible false positives can be published in biology journal
  - e.g JBC, BBRC

# Work product, compound or information

- Industry
  - useful compound is the work product
  - tolerance for 35-50% failure in an HTS
  - pressure for library cleanup
- Academic
  - publication is the work product
  - no tolerance for failure, “publish or perish”
  - no pressure for library cleanup
- Compound work product imposes discipline
- HTS works in industry because it is a process

# Retest rate paradox

- As screening libraries are cleaned up HTS reconfirm rates will fall
- This effect will be most pronounced for the most difficult targets as the true positive hit rate approaches zero

# False HTS positives and chemistry quality

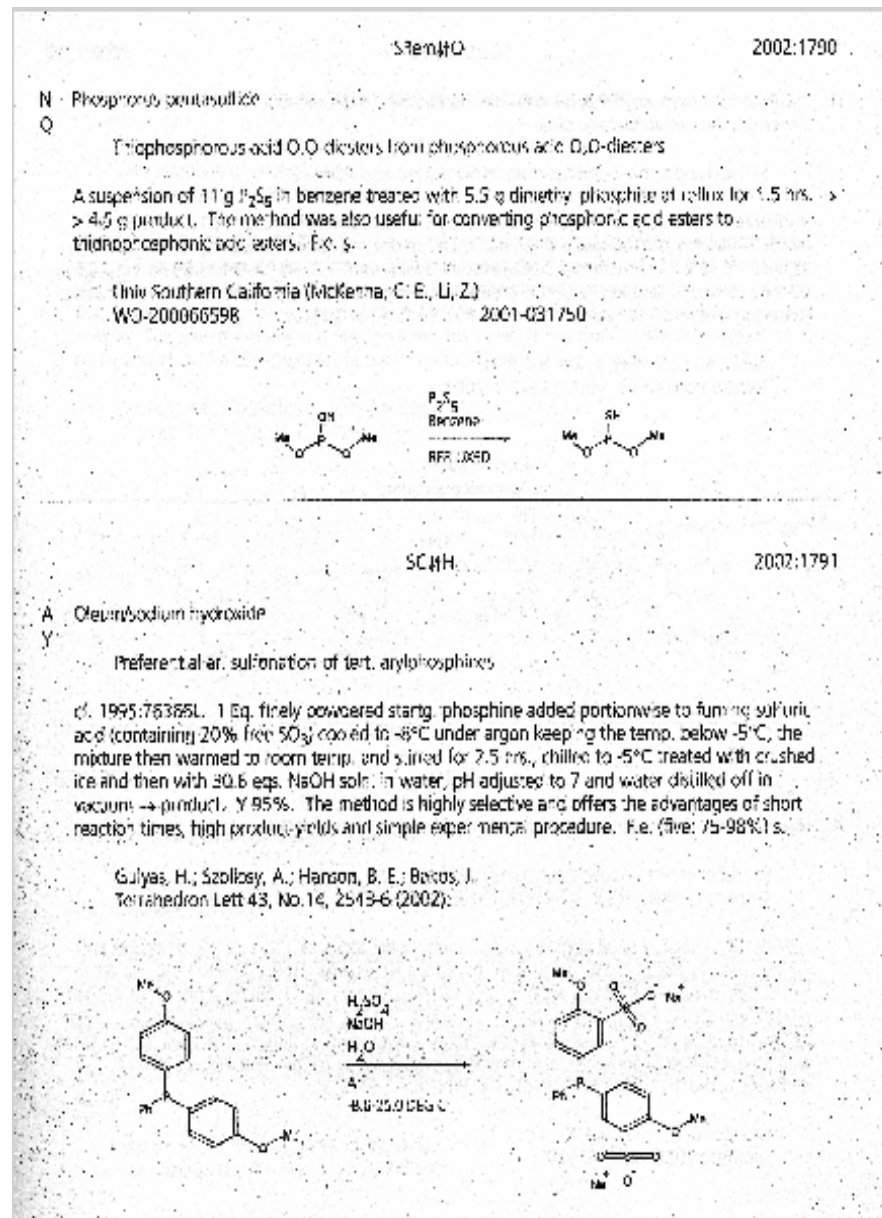
- Find and talk to an experienced medicinal chemist
- But what happens if this communication is approached naively?

# Incomprehensible

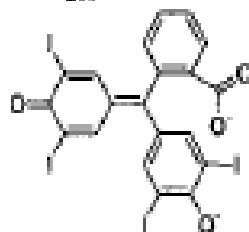
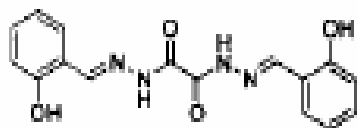
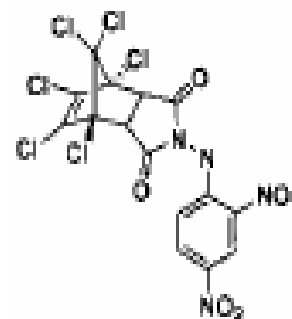
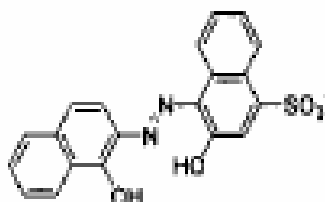
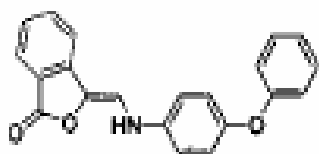
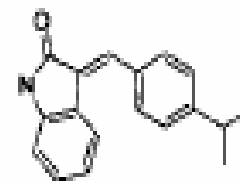
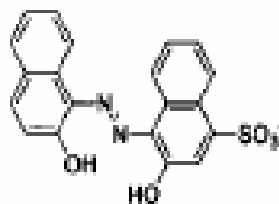
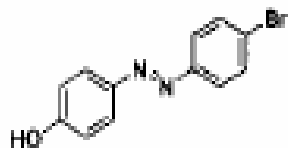
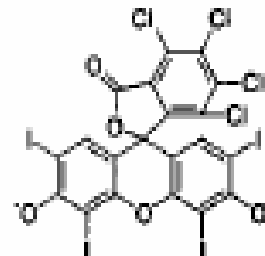
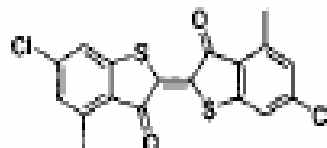
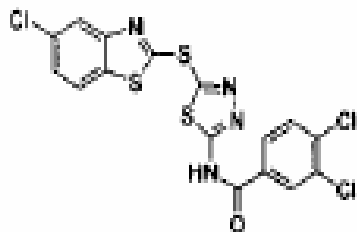
“Reality” depiction in chemistry is incomprehensible to biology / genomics and vice versa.

Chemistry much closer to pathology than to biology / genomics.

Graphical pattern recognition is the forte of chemists



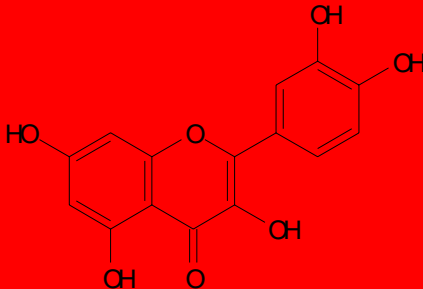
# Aggregation false positives in HTS assays



**These look good to biologists**

**Remove these types of compounds from any assays**

# Chemistry “quality” and biology effort

CHEMISTRY	CAS	NAME	Promiscuous	Refs
				
117-39-5	117-39-5	Quercetin	1	<b>9644</b>
482-89-3	482-89-3	Indigo	1	<b>1170</b>
479-41-4	479-41-4	Indirubin	1	<b>263</b>
133052-90-1	133052-90-1	Bisindolylmaleimide I	1	<b>155</b>
109511-58-2	109511-58-2	U0126	1	<b>127</b>
125314-64-9	125314-64-9	Bisindolylmaleimide IX	1	<b>99</b>
82-08-6	82-08-6	Rottlerin	1	<b>95</b>
85753-43-1	85753-43-1	K-252c	1	<b>74</b>
145-63-1	145-63-1	Suramin	0	<b>1349</b>
19545-26-7	19545-26-7	Wortmannin	0	<b>510</b>
103745-39-7	103745-39-7	HA-1077	0	<b>140</b>
146986-50-7	146986-50-7	Y-27632	0	<b>113</b>
152121-47-6	152121-47-6	SB203580	0	<b>262</b>
167869-21-8	167869-21-8	PD 98059	0	<b>310</b>
318480-82-9	318480-82-9	SC68376	0	<b>8</b>

> 80%

flawed  
compounds?

# The amygdala and emotional memory

- Tale of the tiger
- Chemists are particularly social
- Chemists are superb at pattern recognition
- Structural context of “negative” compounds is locked into the amygdala
- Rapid effortless chemistry context memory retrieval

# Biology, chemistry and structures

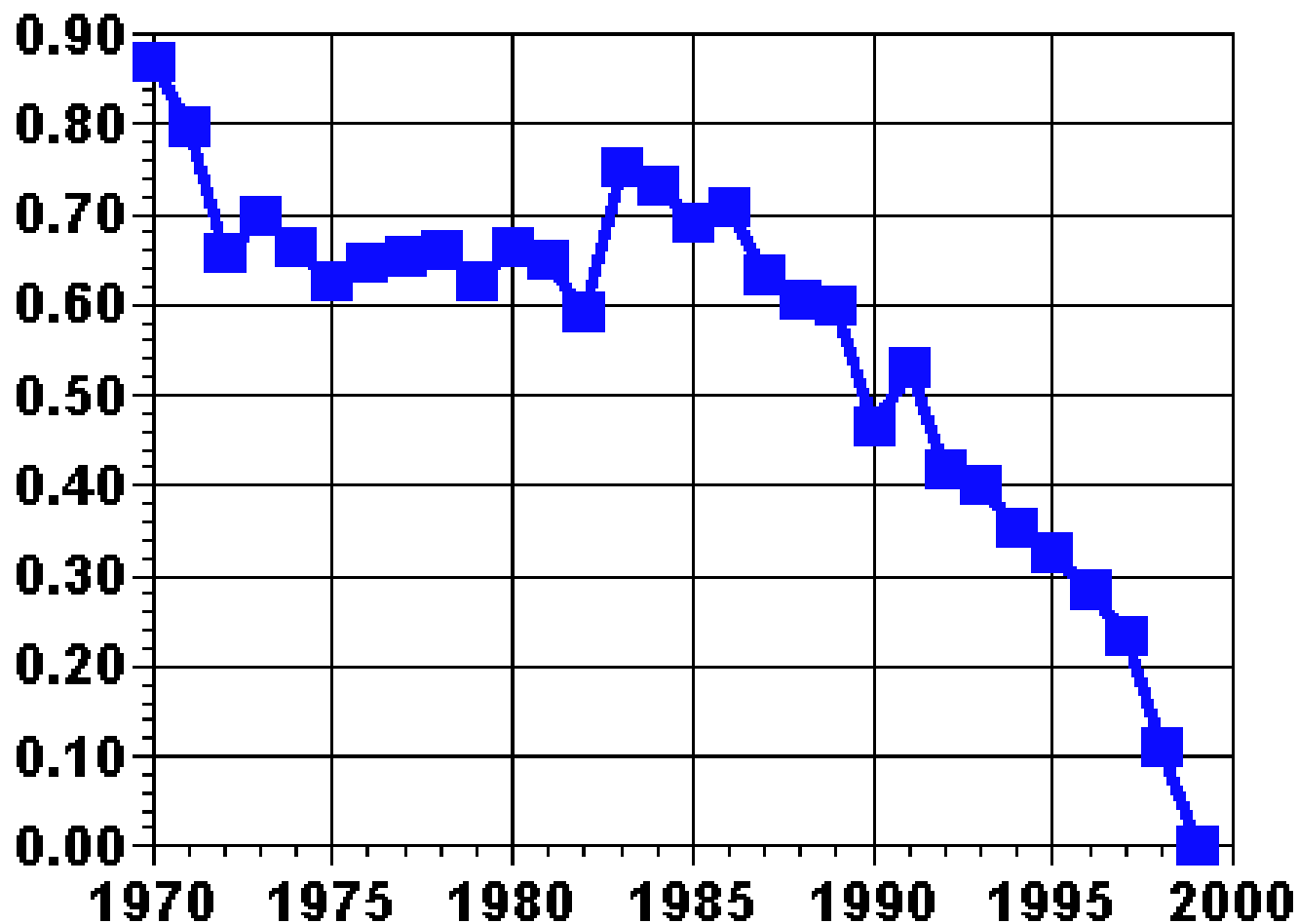
- Exercise tact!

# HTS and compound concentration

- Nobody knows compound concentration
- IC-50 orders gets scrambled
- Selectivity panels get messed up
- Compounds precipitate from DMSO
- Compounds precipitate from buffer

**Streamlined System for Purifying and Quantifying a Diverse Library of Compounds and the Effect of Compound Concentration Measurements on the Accurate Interpretation of Biological Assay Results.** *Anal. Chem.* (2004), 76(24), 7278-7287.

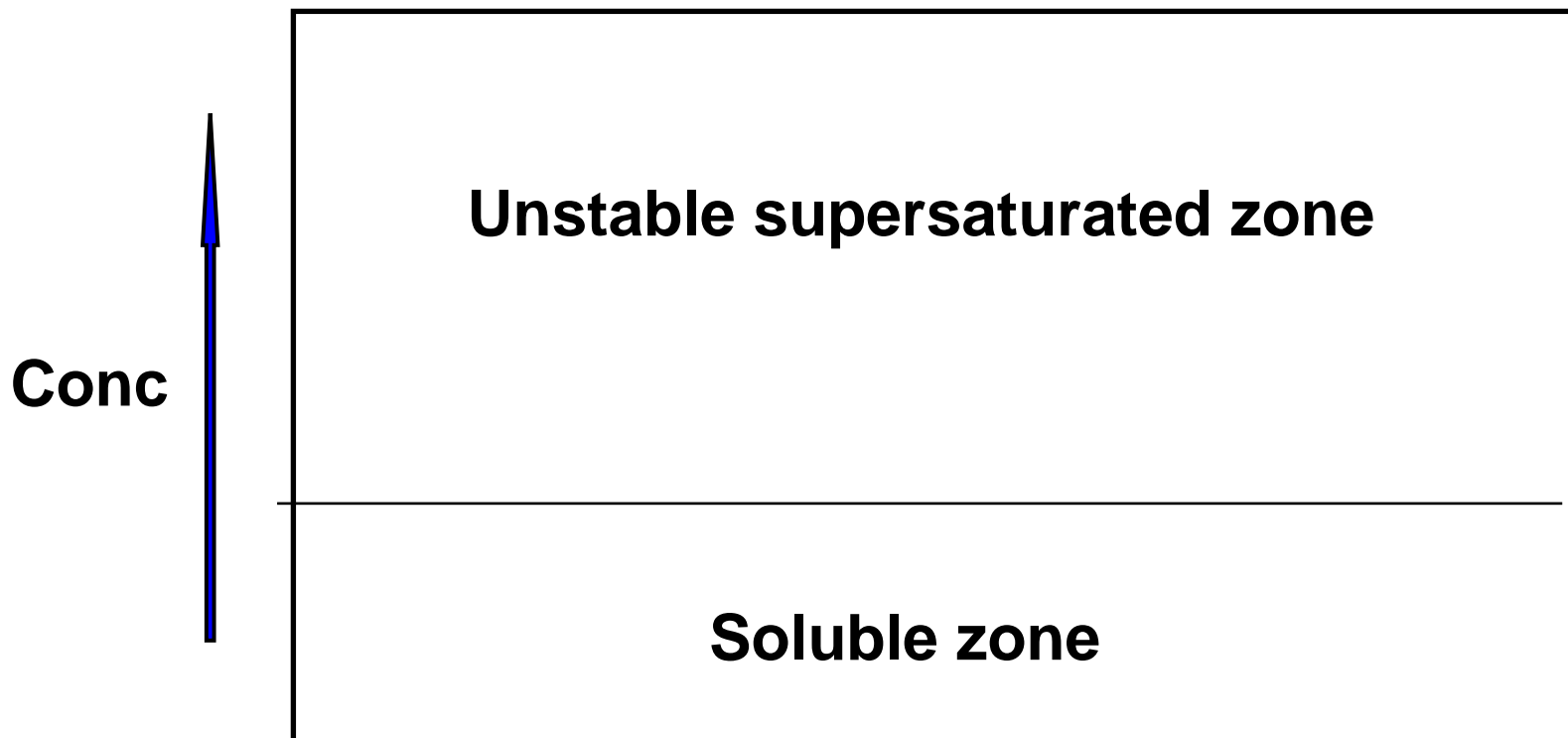
# Fraction of Pfizer Groton compounds having melting point field information



# Consequences of amorphous compounds

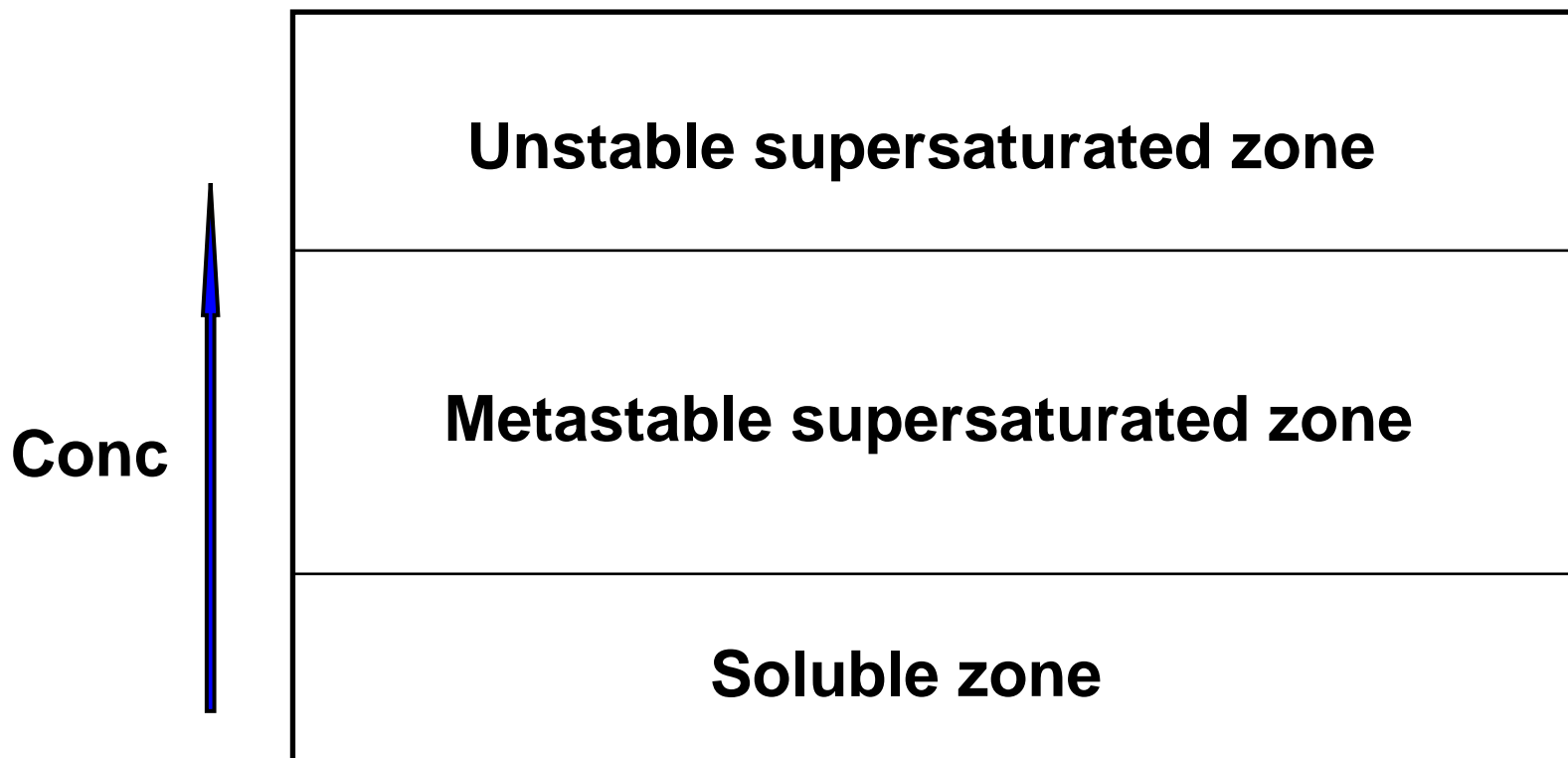
- Amorphous DMSO solubility is always higher than when compound is crystalline
- Amorphous compounds from combichem or medchem initially easily dissolve in DMSO
  - allows preparation of DMSO stocks
- Sets stage for later precipitation problems

# Thermodynamic aqueous solubility

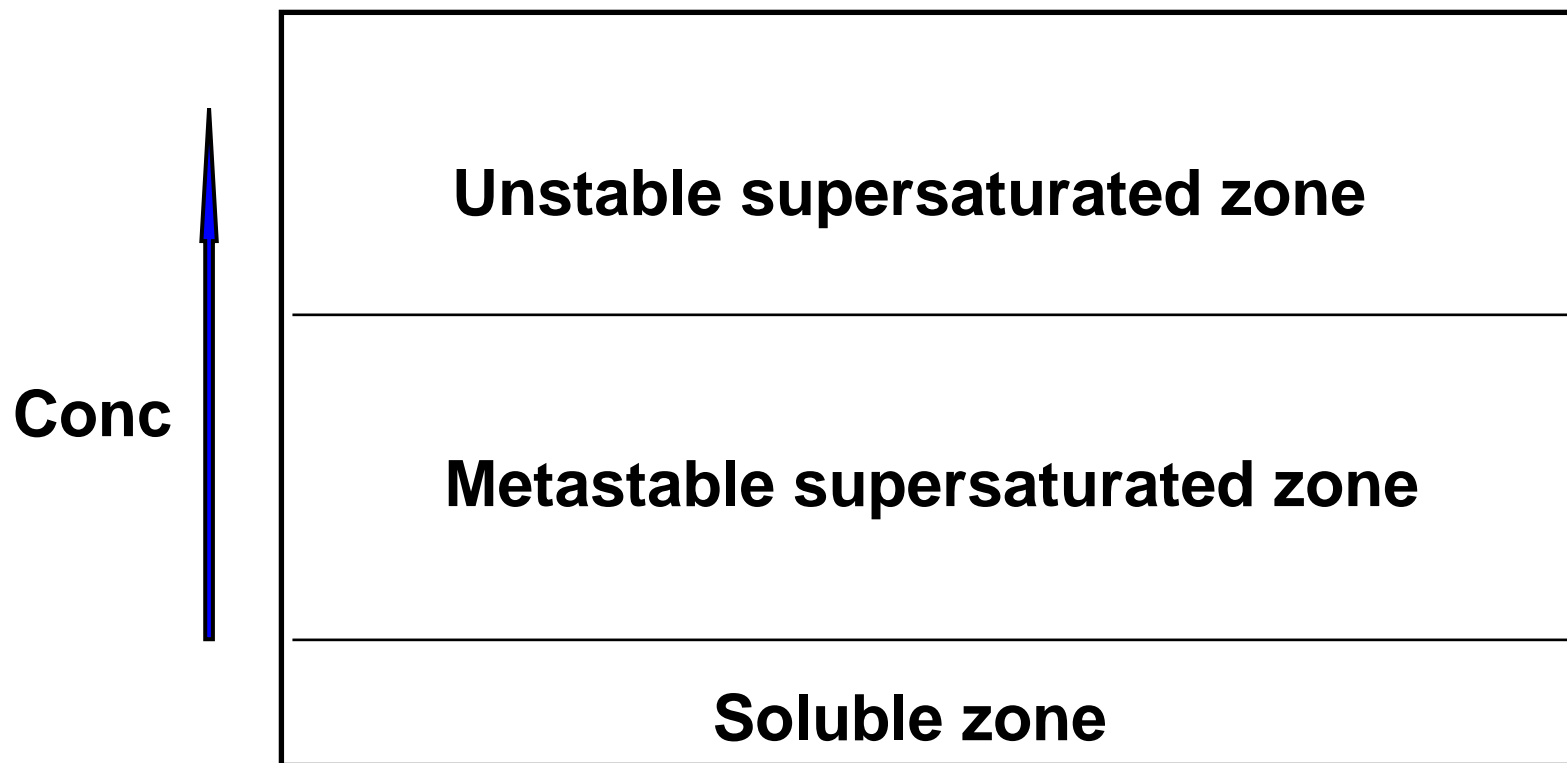


**If solid is present, crystallization (precipitation) will occur if concentration rises into the supersaturated zone**

# Metastable supersaturated zone



**If solid is not present, precipitation cannot occur from the metastable supersaturated zone even though the solution is thermodynamically supersaturated**



**If solid is not present, precipitation can occur from the unstable supersaturated zone. One way to enter the unstable zone is through freeze thaw cycles in wet DMSO**

# Water and DMSO solubility

- Large, lipophilic compound
  - aqueous insoluble
  - DMSO greatly helps aqueous solubility
- Very crystalline compound
  - may show no computational problem
  - no “rule of 5” violation
  - aqueous insoluble
  - high melting point
  - strong intermolecular crystal lattice
  - DMSO does not help aqueous solubility

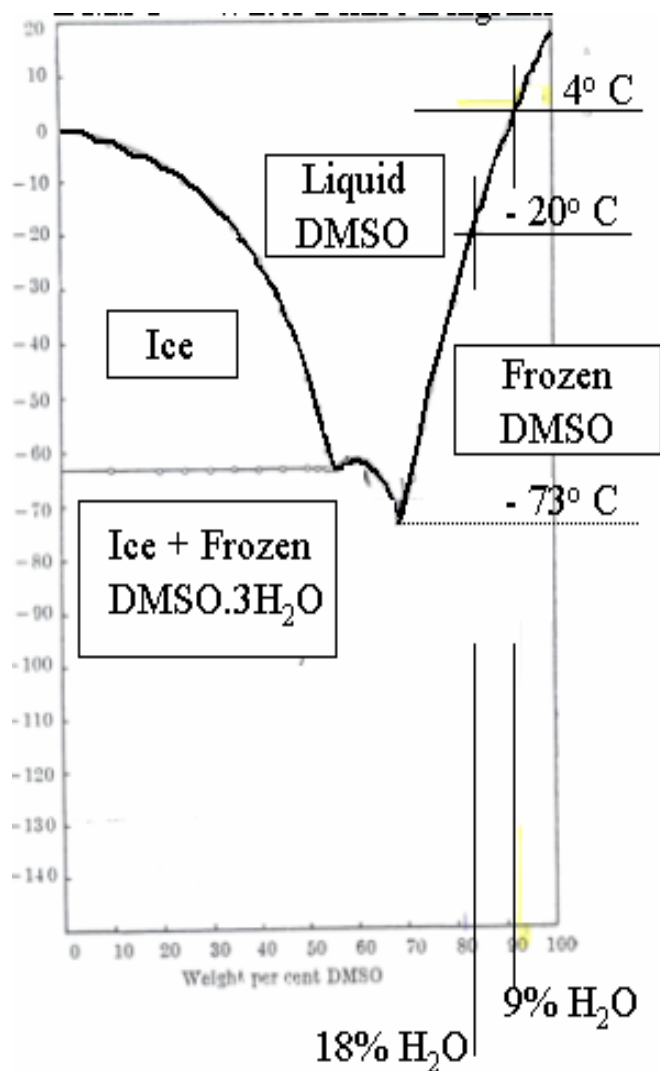
# DMSO helps aqueous solvation when the problem is size / lipophilicity

- Compound has to make a “hole” in DMSO to dissolve
- easier to do this in DMSO than in water
  - no H-bond donor / acceptor networks to disrupt
- DMSO has a high dielectric constant
- solvates compound dipoles
  - almost all drugs have dipole

# Sample in DMSO lifetime

- Compound disappears from DMSO solution
- What is the explanation?
- Chemical integrity in DMSO
  - keep cold and frozen
  - avoid oxygen
  - keep dry
- Compound solubility in DMSO
  - cold and / or frozen is the worst choice possible
  - avoid freeze thaw cycles
  - Avoid water uptake into DMSO

# DMSO – water phase diagram



**DMSO containing 9% water is unfrozen in the typical lab refrigerator**

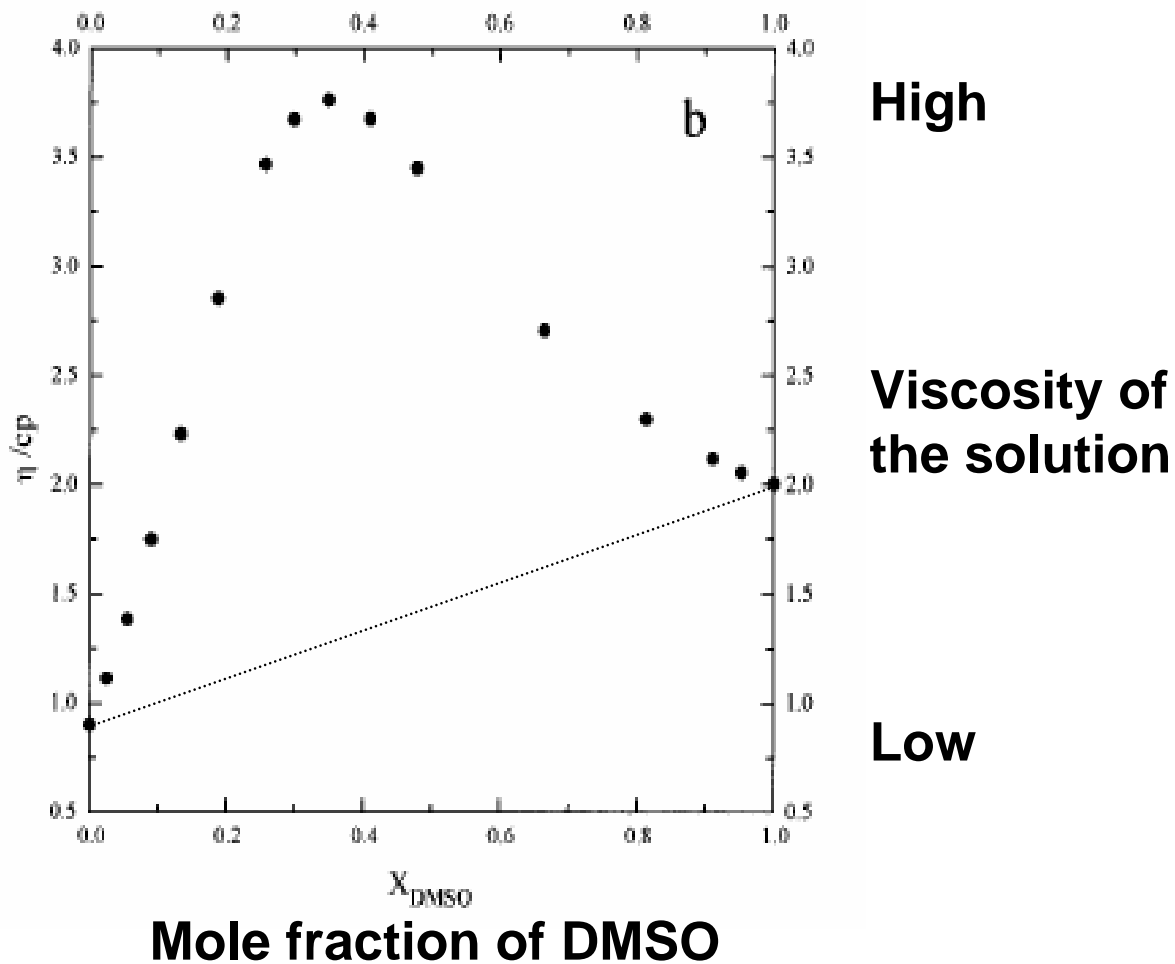
**Very bad practice. Typical biology procedure of storing samples in DMSO in the non-freezer part of a lab refrigerator**

Nature (1969), 220, 1315-1317

# DMSO and water, non-ideal behavior

- DMSO plus water exhibits very non ideal behavior
  - change maximum at 33% by weight water
  - corresponds to molar ratio 1 DMSO : 2 H<sub>2</sub>O
  - liquid phase is very structured
  - ice-1 like lattice structure
  - more H-bonding, more order than in pure water
- Melting point declines from 18 deg C to -73 deg C
- Viscosity increases as water is added to DMSO
- Solution gets colder as water is added to DMSO
  - (J Org Chem 2001, 66, 5846-5852)

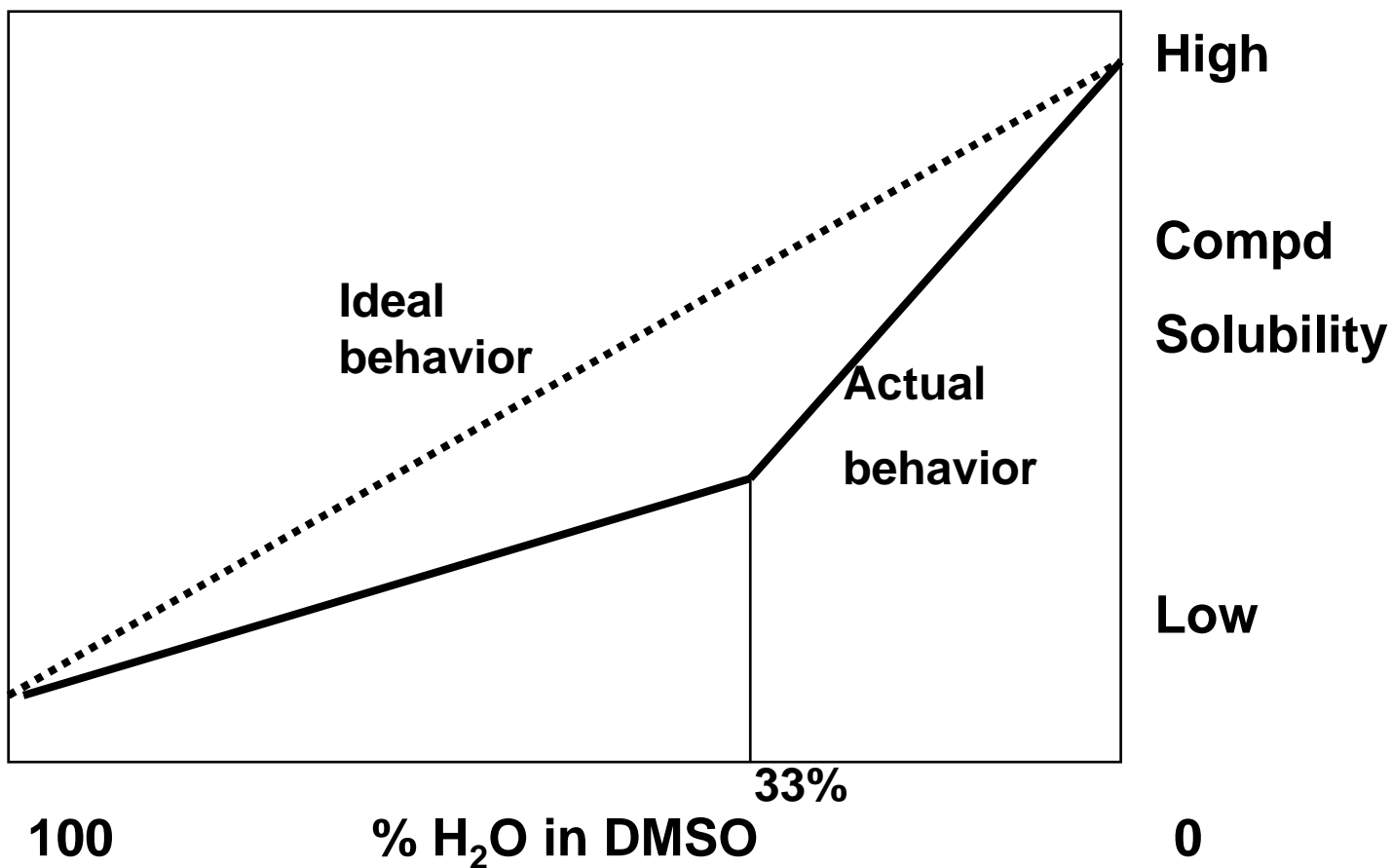
# Viscosity and water addition to DMSO



# Compound solubility in DMSO and H<sub>2</sub>O

- As H<sub>2</sub>O is added solvent is more structured
- More difficult to form a cavity
- More difficult to dissolve a large lipophilic compound
- Solubility non ideality is expected to be worst at 33% by weight H<sub>2</sub>O

# DMSO solubility dependence on H<sub>2</sub>O



# Timing factor in DMSO solubility

- Once a compound crystallizes from DMSO it will not easily re-dissolve
  - crystallized compound is in a lower energy, higher melting point, less DMSO soluble form
- Narrow working window (time window) for keeping most compounds dissolved in DMSO
  - 1 to 2 days at room temperature
  - explains why compounds are active when freshly made but not when stored
- Freeze thaw cycles in synergy with water uptake increases the probability of crystallization

# Summary

- Crystalline state is important to aqueous solubility
- Crystalline state is important to DMSO solubility
  - be alert for compound precipitation from DMSO
  - do not store liquid DMSO stocks in the refrigerator
  - minimize time once DMSO stocks are diluted
  - expect erratic HTS screening results dependent on minor compound handling differences
  - minimize freeze thaw cycles
- Keep the DMSO bone dry
- Talk to biologists about keeping DMSO dry

# Acknowledgements

The generous support of Pfizer Global R&D, Groton New London Labs in my post retirement activities is gratefully acknowledged