Fragment-based Lead Discovery Conference 2010
Philadelphia, 10th - 13th October 2010

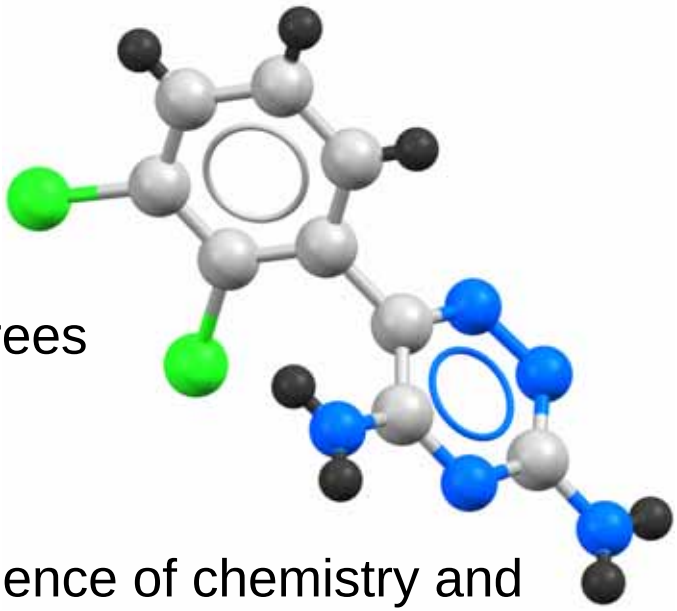
From Druggability, to Ligand Efficiency, to the Universe of Heterocyclic Rings

Colin R Groom

Cambridge Crystallographic Data Centre

Monday 11th October
Fragment Libraries and Chemical Space

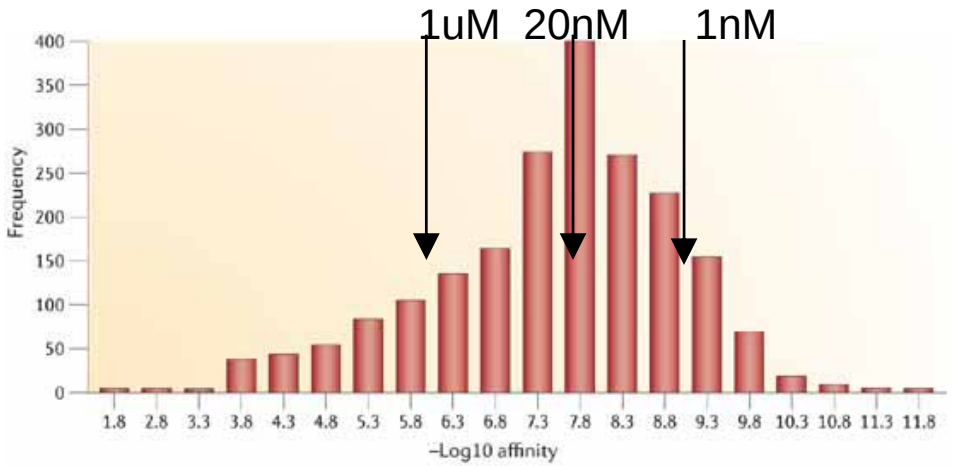
The Cambridge Crystallographic Data Centre



- A non-profit, charitable institution
- Self financing and self administering
- 51 employees
- Recognised institute for postgraduate degrees of the University of Cambridge
- Objectives
  - “advancement and promotion of the science of chemistry and crystallography for the public benefit”
- Provides the Cambridge Structural Database System
- Associated software
  - GOLD, Relibase

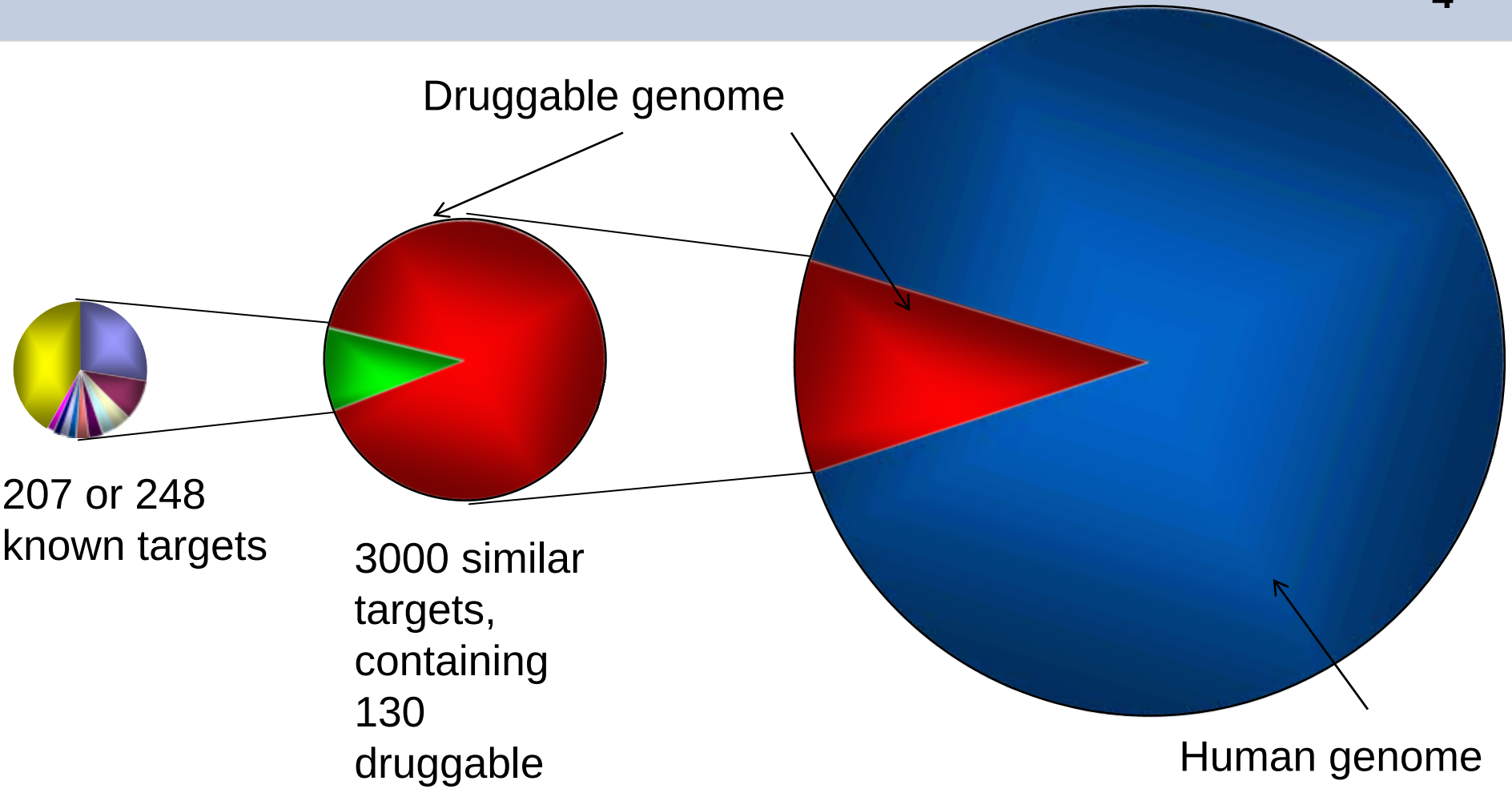
Drug candidate attrition – the primary motivation

Percentage values	Oral	Non oral	**Total**
Ro5 pass	52	22	74
Ro5 fail	13	13	26
Total	65	35	**(1194)**



Hopkins, A.L., and Groom, C.R. The Druggable Genome. Nature Reviews Drug Discovery (2002), 1,727-730
Overington, J.L., Al-Lazakani, B. and Hopkins, A.L. How many drug targets are there? Nature Reviews Drug Discovery (2006), 5, 993-996

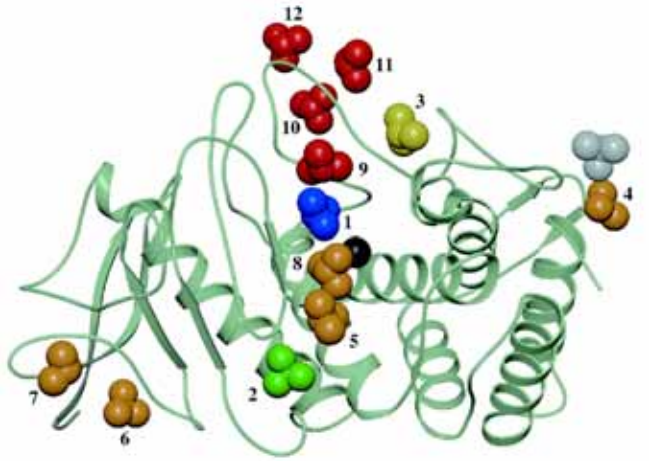
Target attrition – the second motivation



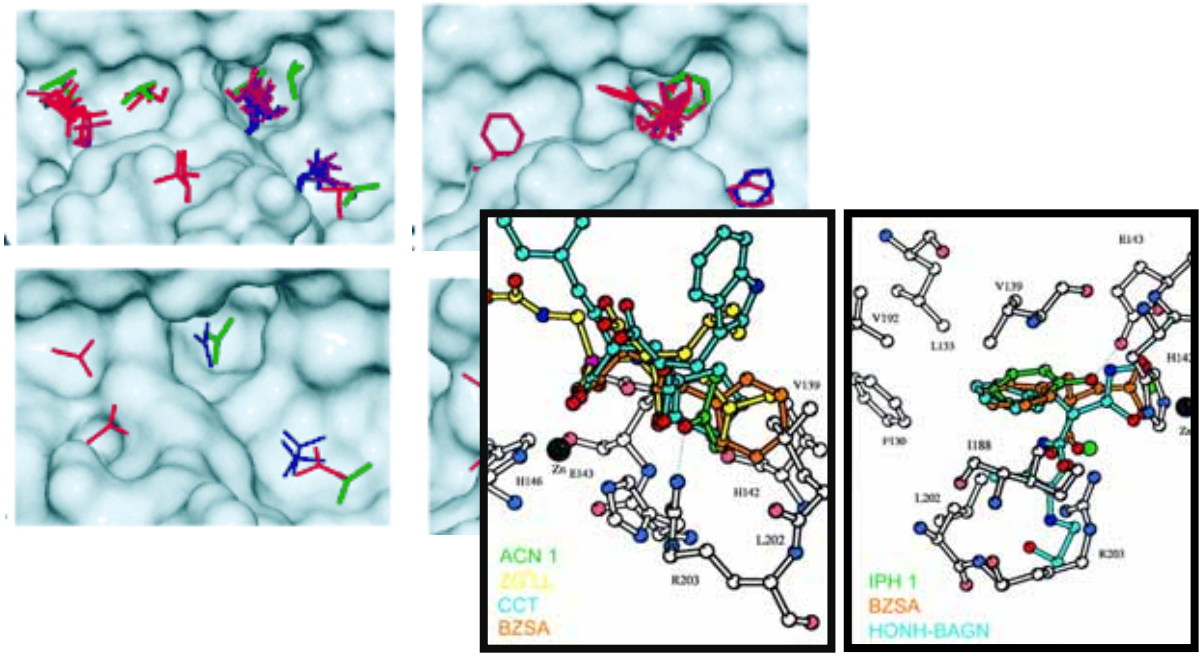
Hopkins, A.L., and Groom, C.R. The Druggable Genome. Nature Reviews Drug Discovery (2002), 1,727-730

Can we start from a better place?

Can we identify small ligands?



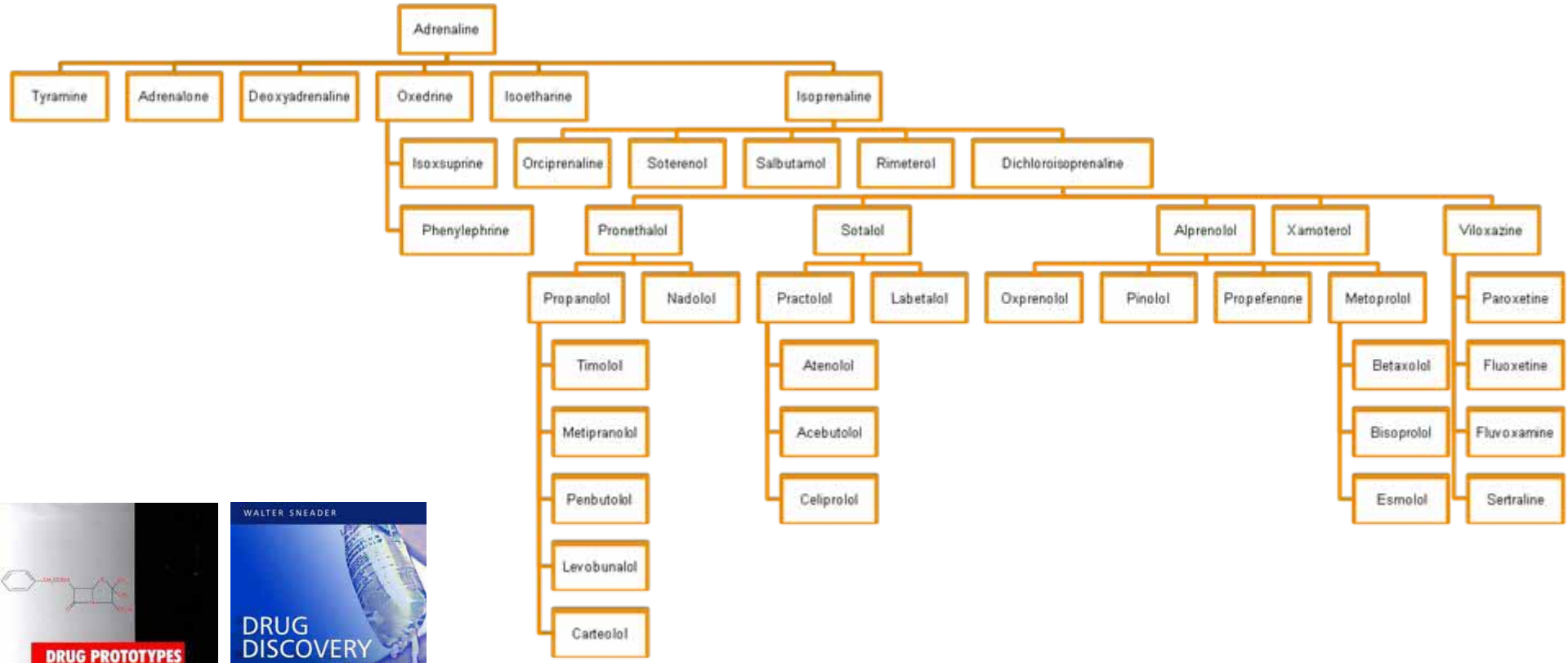
Can we do crystallographic fragment screening



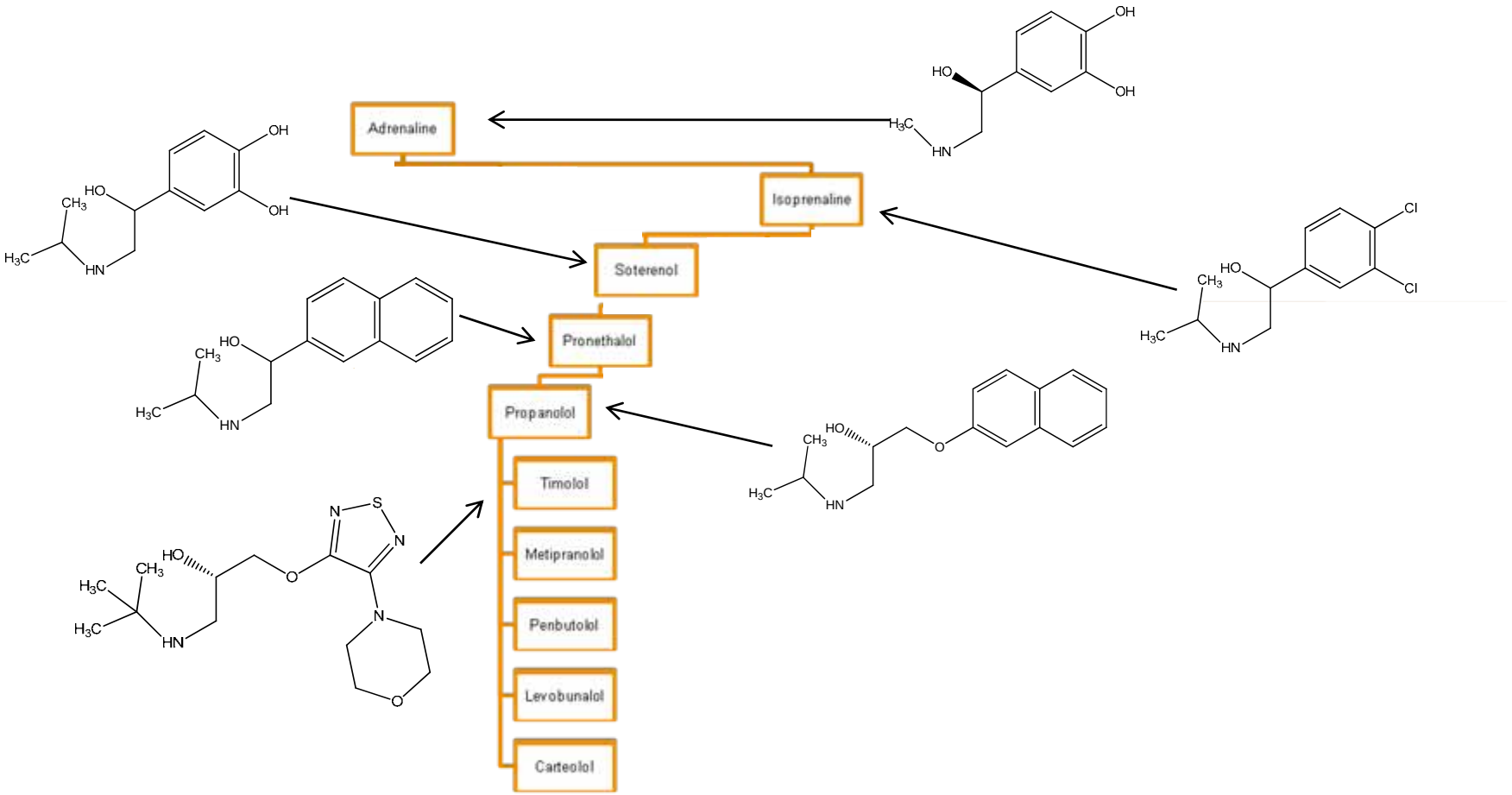
Can we do fragment-based SAR analysis?

Locating interaction sites on proteins: The crystal structure of thermolysin soaked in 2% to 100% isopropanol
Proteins: Structure, Function and Genetics, 1999, 37,4, 628-640 A. C. English, S.H. Done, L.S.D. Caves, C. R. Groom, R.E. Hubbard
Experimental and computational mapping of the binding surface of a crystalline protein. Protein Engineering, 2001 14, 47-59 Andrew C. English, Colin R. Groom, and Roderick E. Hubbard

Early fragment optimisation



Early fragment optimisation



Assessing our starting point: Ligand efficiency

- The binding energy of a compound as a function of any of its properties
- Originally expressed as function of number of heavy atoms
- Can be surface area, lipophilicity etc
- What efficiency do drugs bind with?
- How efficient is our hit / lead?
- What ligand efficiencies have been observed against target X

Ligand binding efficiency

- We can produce figures for the binding efficiency for any ligand
  - As

    $\Delta G = -RT \ln (K_i)$

  - Then free energy per atom

    ligand efficiency = ΔG / N

Where N = number of non-hydrogen atoms

N: A surrogate for logP, logD, volume, metabolic liability, etc

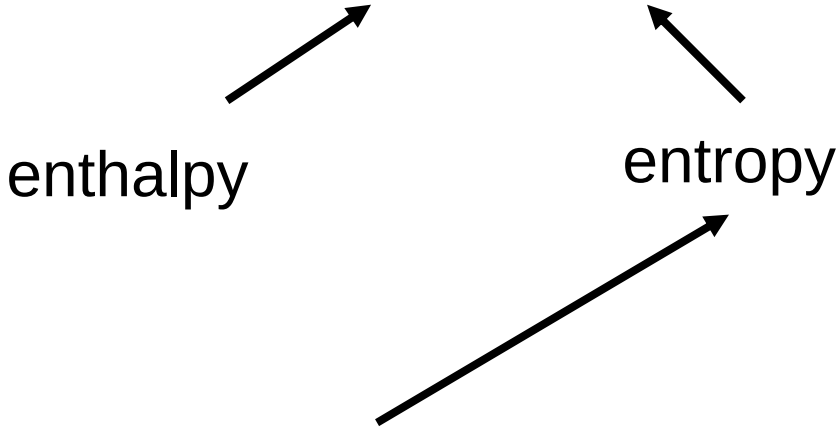
Ligand efficiency: a useful metric for lead selection. Hopkins, A.L., Groom, C.R. and Alex, A. DDT, (2004) 9, 430-431

Comparing ligand efficiencies

- For ligands of the same size
  - i.e. alternative leads
    - Very illuminating
- For ligands of different sizes
  - Requires a little more care…

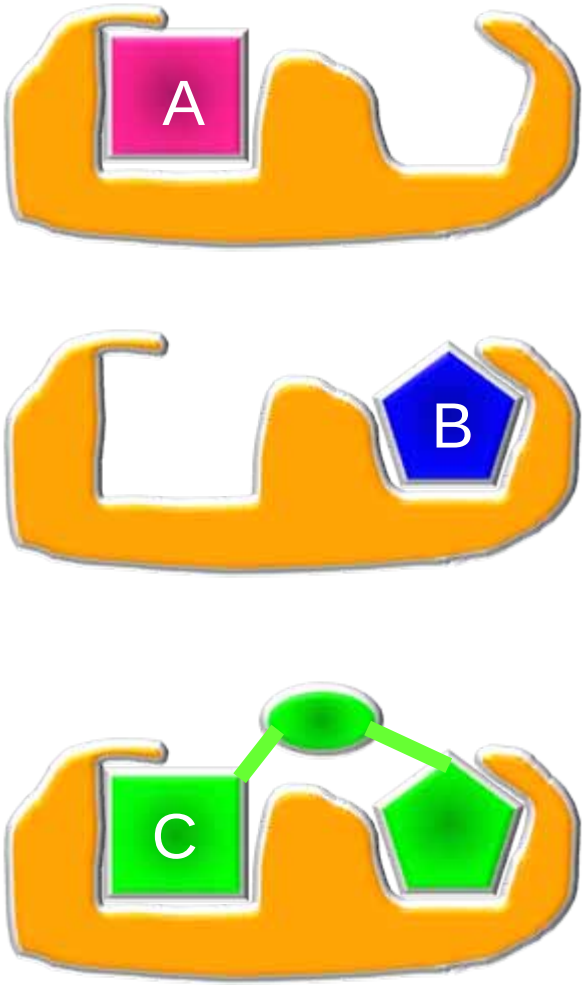
Entropy and enthalpy

$$\Delta G = -RT \ln (K_i) = \Delta H - T\Delta S$$



- Solvent displacement from ligand and protein
- Internal conformational entropy of ligand and protein
- Combining two independently tumbling molecules into one

Additivity of fragments



$\Delta G^A = \Delta H^A - T\Delta S^A$

$\Delta G^B = \Delta H^B - T\Delta S^B$

$\Delta G^C = \Delta H^C - T\Delta S^C$

$\Delta H^C \cong \Delta H^A + \Delta H^B$

$T\Delta S^C \cong T\Delta S^A \cong T\Delta S^B$

$\Delta G^C = \Delta H^A + \Delta H^B - T\Delta S^C$

$\Delta G^C >> \Delta G^A + \Delta G^B$

Intrinsic entropy loss

- The entropic penalty due to stopping a molecule tumble is proportional to log molecular weight
  - A 500 mw ligand, binding to a 30,000 mw protein
    - ligand looses log (500) = 2.700 units
    - The protein target looses log (30,000) – log (30,500) = 0.007
      - Total entropic penalty (tumbling only) = 2.707

- When you sort out all the constants this is about +4 kcal/mol

The consequences of translational and rotational entropy lost by small molecules on binding to proteins.
Murray, C.W and Verdonk, M.L. J. Comp-Aid Mol. Des. (2002). 16, 741-753

Here's why its tricky

- A is a 300 mw ligand binding to a 30,000 mw protein
  - Entropic penalty (tumbling only) about +4 kcal/mol

- C is a 500 mw ligand, binding to a 30,000 mw protein
  - Entropic penalty (tumbling only) about +4

- A and C pay roughly the same entropic penalty

Comparing ligand binding efficiencies

- A = 300 mw, 10 μM ΔG^A = -6 kcal/mol
  - Therefore efficiency = 0.30 kcal/mol/atom

- C = 500 mw, 10 nM ΔG^C = -10
  - Therefore efficiency = 0.30

- Conclusion
  - A and C have the same ligand efficiency?

Lets go back to efficiencies

- A = 300 mw, 10 μM ΔG^A = -6 kcal LE= 0.30
- C = 500 mw, 10 nM ΔG^C = -10 LE= 0.30

- Efficiencies considering the constant entropic penalty

ΔG^A = -6 kcal = -10 + 4

  - Enthalpy= -10 kcal/mol
  - Enthalpic efficiency = -10 / 300 = 0.49

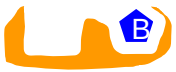
ΔG^C = -10 kcal = -14 + 4

  - Enthalpy= -14
  - Enthalpic efficiency = -14 / 500 = 0.37

- Revised conclusion = A is actually more efficient than C in *enthalpic terms*

Comparing Fragment A with lead C

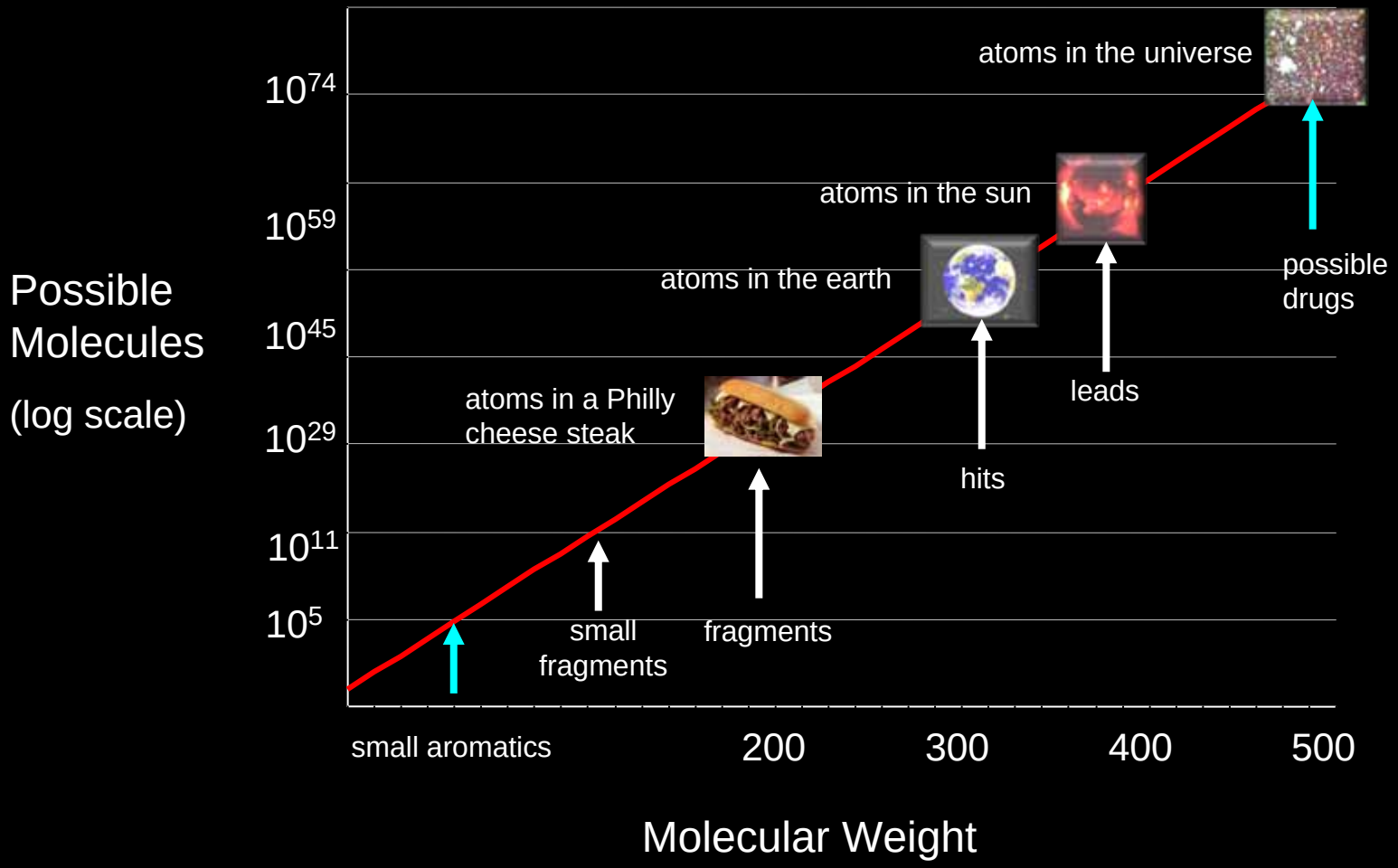
- Say A was our fragment
  - A is 300 mw, 10 μM ΔG = -6 kcal (-10 + 4)
  - Intrinsic efficiency = 0.49 (was 0.3)
- Say C is our lead
  - C is 500 mw, 10 nM ΔG = -10 kcal (-14 + 4)
  - Intrinsic efficiency = 0.37 (was 0.3)
- To get from A to C we added B
  - B is 200 mw, it contributes –4 kcal to ΔG
    - there is no change to rigid body entropy
  - Its intrinsic efficiency is only 0.27 (was 0.3)
- The 200 mw piece we have added (B) is much less efficient than the original 300mw piece (A)



Messages

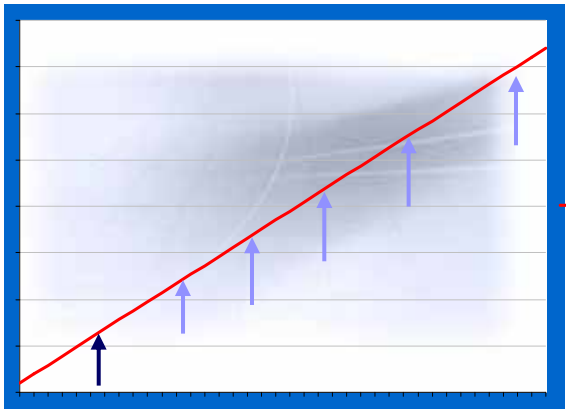
- Compare ligand efficiencies of different size ligands with caution
- If one ‘maintains ligand efficiency’ during fragment to hit process
  - The atoms added are less efficient in terms of enthalpy than the original fragment
- ITC data can contribute a tremendous amount

The chemical universe



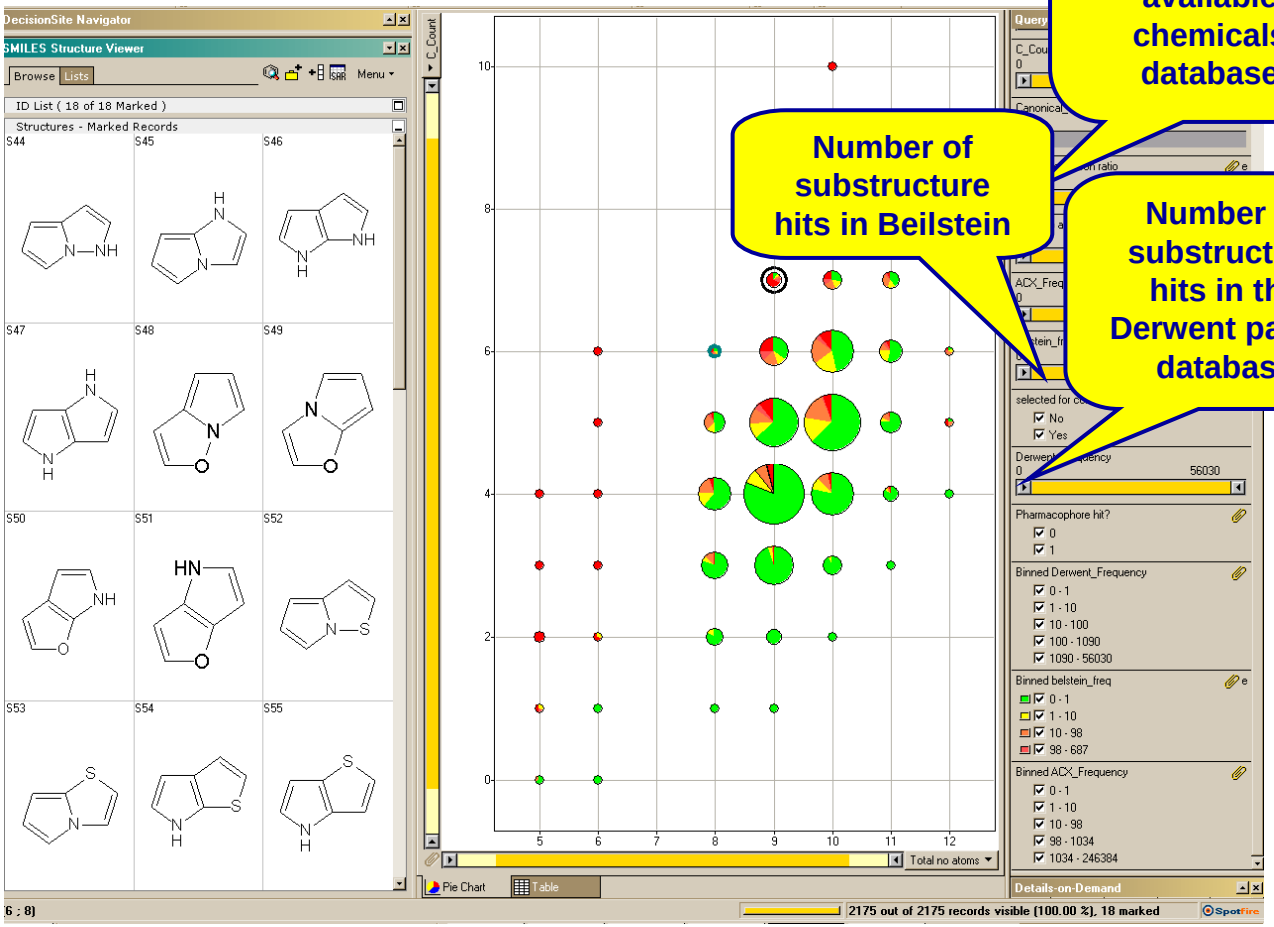
Why Planar Heteroaromatics ?

- Tractable number
  - Allows one to see the full picture
- Pharmaceutical interest
  - Compounds get larger during ‘optimisation’
  - Smaller compounds more likely to show activity
  - Focus on ligand efficiency
  - May be more ‘novelisable’
  - Scaffold of molecule important for IP
  - Source of novelty of chemical series
  - Aryl bond formation allows modifications
- Planar heteroaromatics are key to a medicinal chemists thinking
  - Complete enumeration of all aromatic monocycles and bicycles
  - 5 and 6 membered rings, C,N,S,O, Neutral, Obey Hückel’s rule, Only exocyclic carbonyls



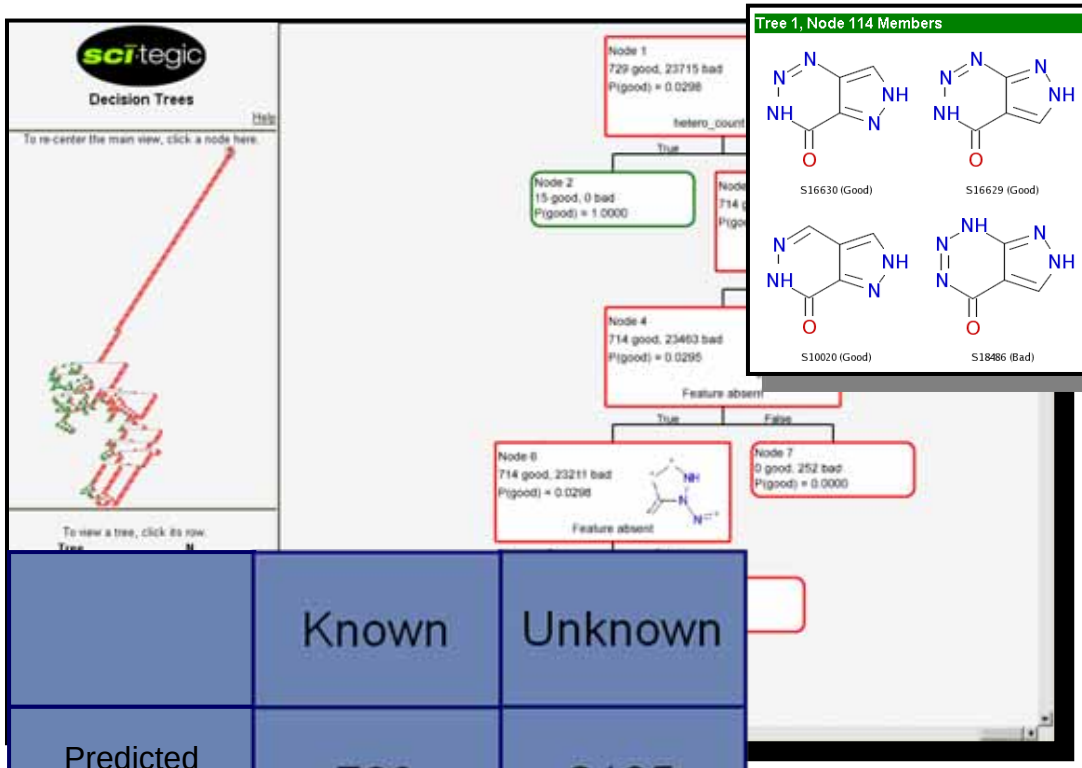
VEHICLe summary

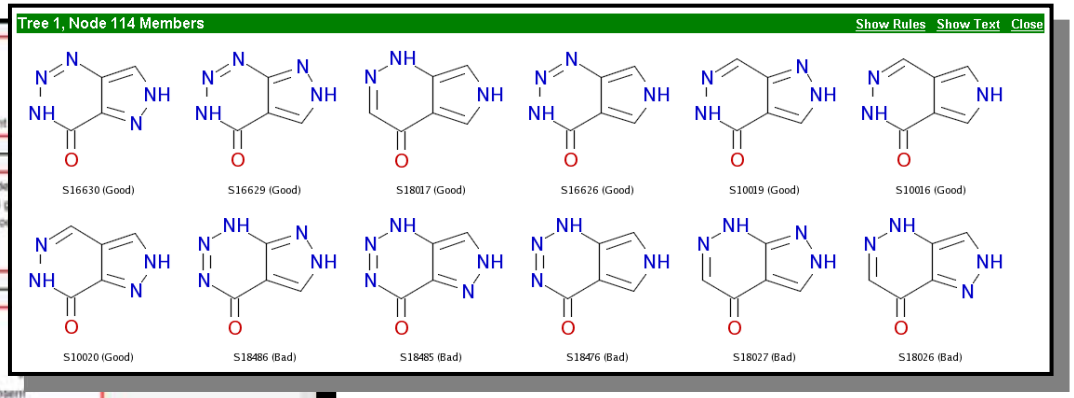
- 24,867 rings
- About 500 found in drugs



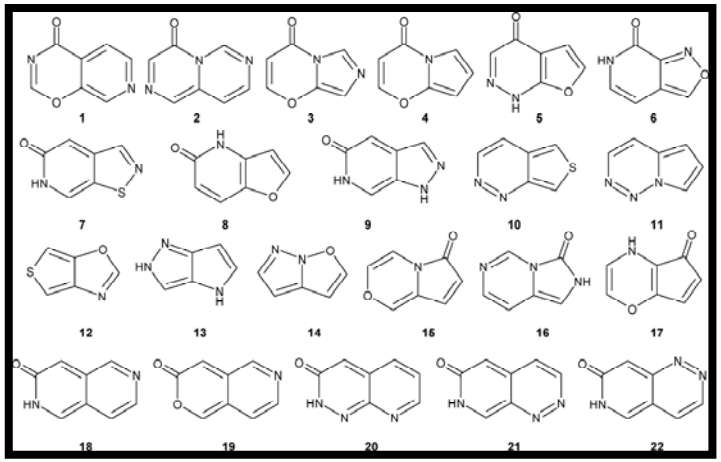
Heteroaromatic Rings of the Future. William R. Pitt, David M. Parry, Benjamin G. Perry and Colin R. Groom
J. Med. Chem., 2009, 52 (9), pp 2952–2963 DOI: 10.1021/jm801513z

Synthetic Accessibility



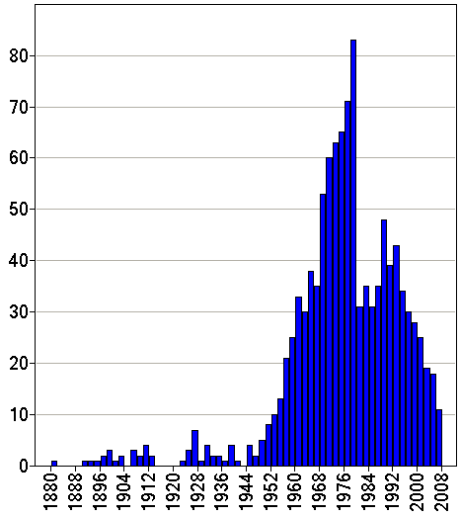


	Known	Unknown
Predicted synthesisable	769	2185
Predicted difficult	0	21913



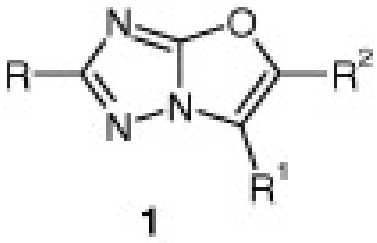
Rings of the future

- Peak of heterocyclic chemistry in 1971
  - Now 1701 examples
- Number of new heterocycles published declining
- The remaining ones are very precious

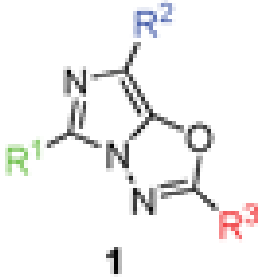



Tetrahedron Letters

[1,3]Oxazolo[3,2-b][1,2,4]triazoles: a versatile synthesis of a novel heterocycle



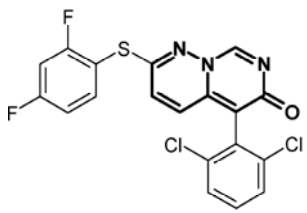

Robust preparation of novel imidazo[5,1-*b*][1,3,4]oxadiazoles†

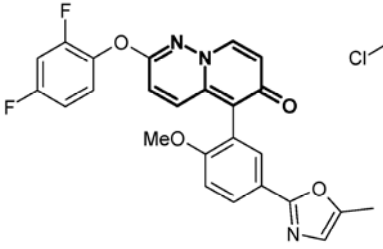


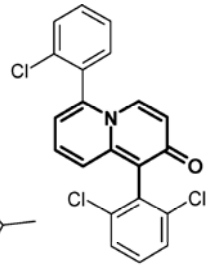
WO 9827098 A1 19980625
Vertex (VX-745)

US 2007129372
Merck & Co. Inc.

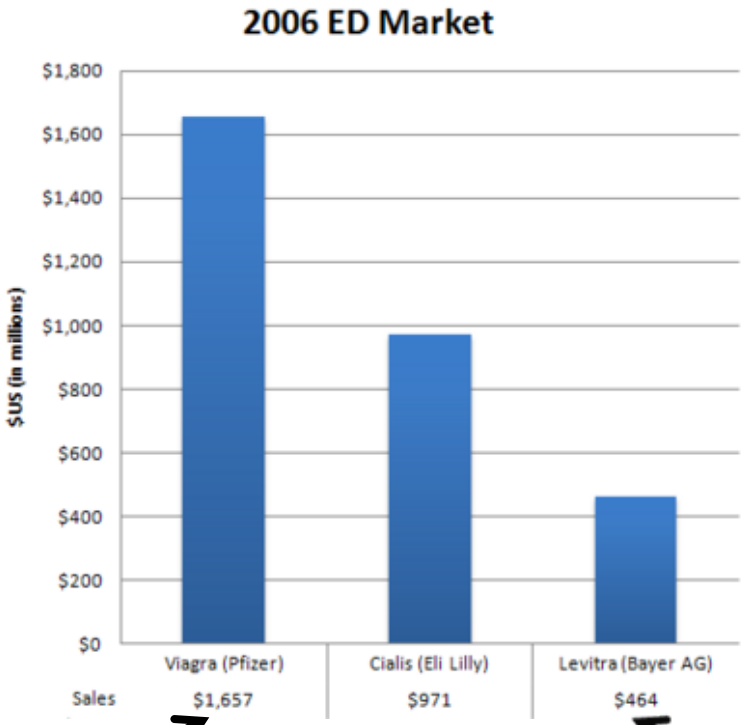
WO 2007021710

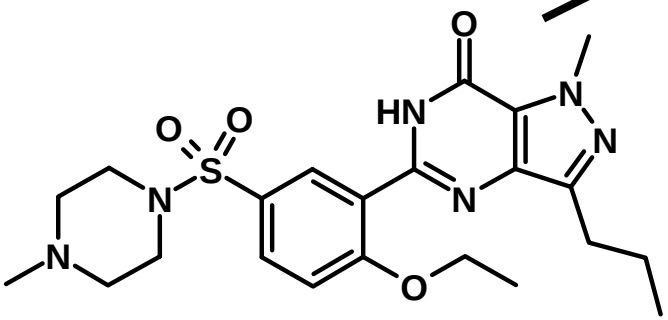


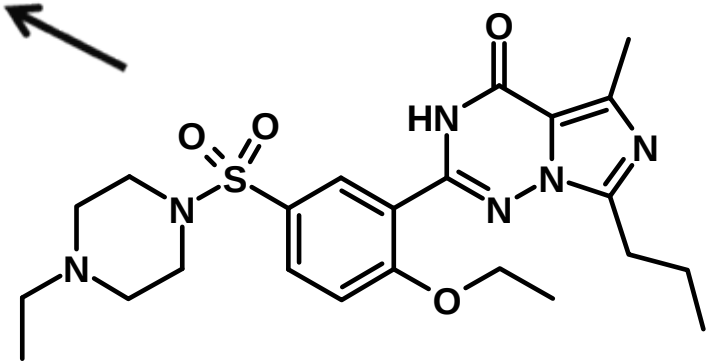




Impact of analogues

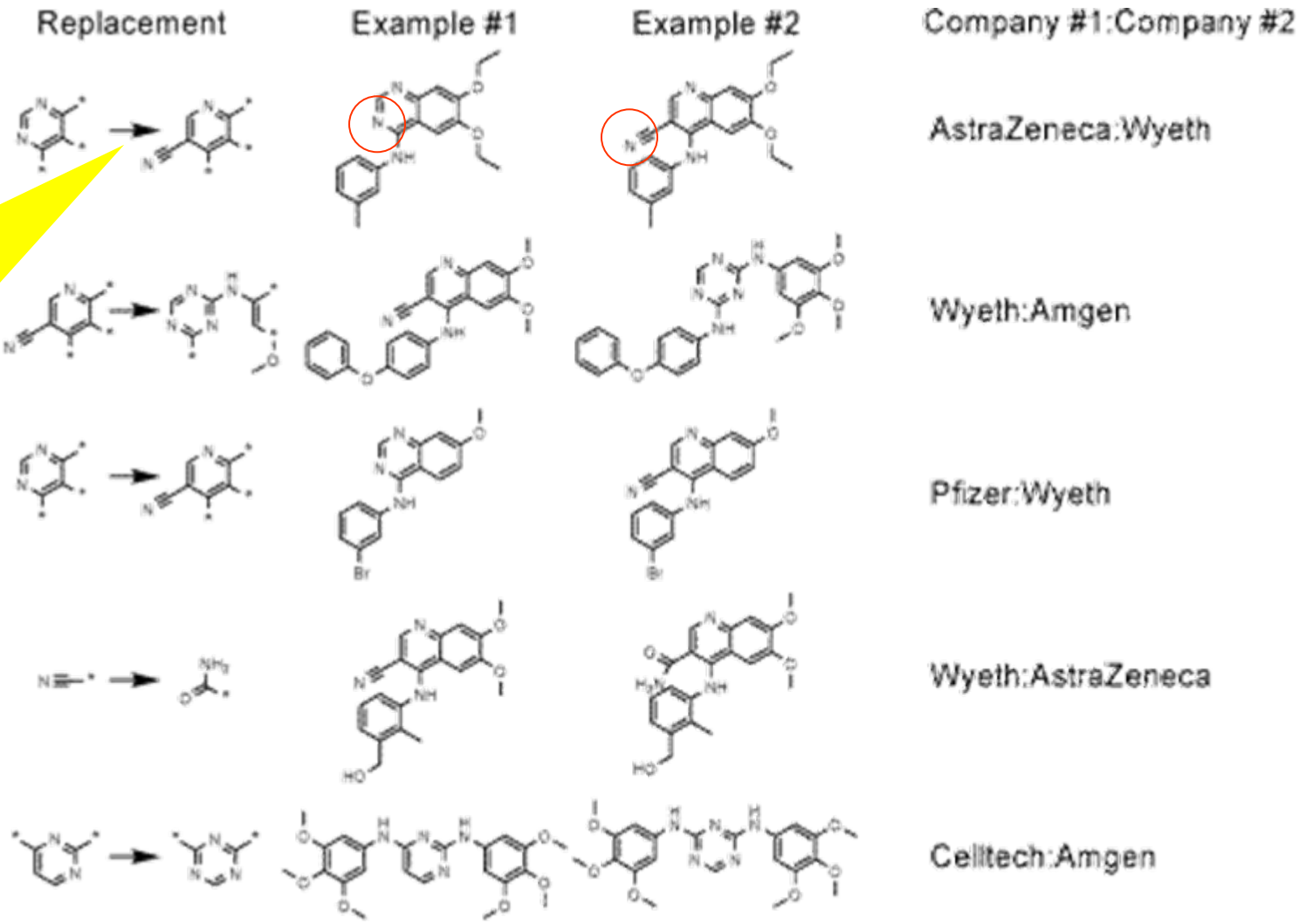






Frolicking in Patent Space - 1

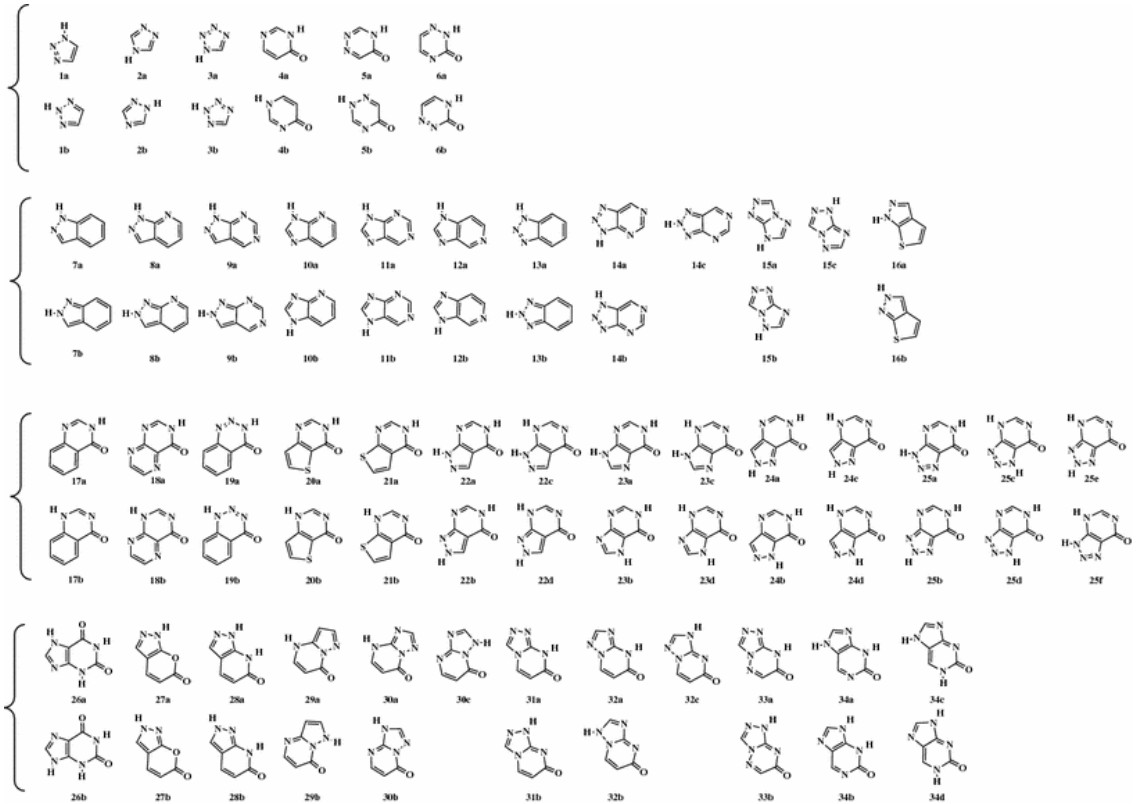
70 examples, across 4 patents, of Wyeth doing this mod to an AZ quinazoline



Southall and Ajay J. Med. Chem. 2006, 49, 6, 2103-2109

Tautomeric Heterocyclic ‘Space’

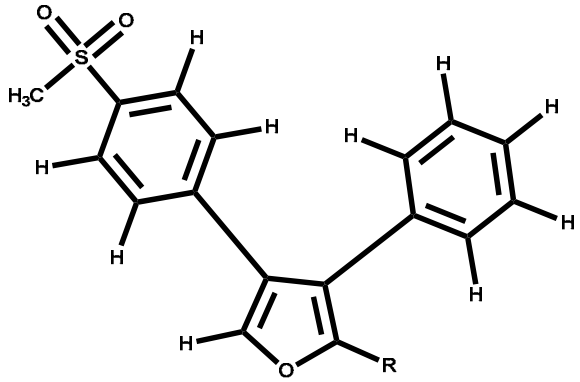
- Space is larger when we think of tautomers



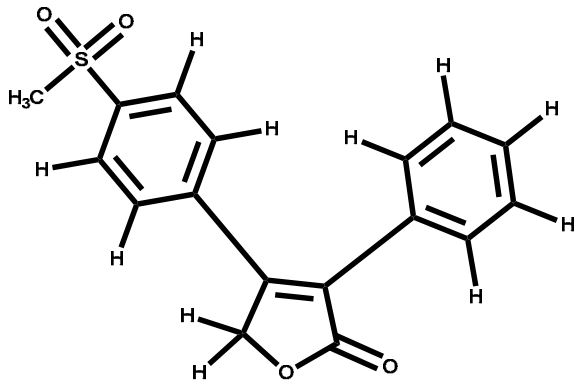
A.J. Cruz-Cabeza, A. Schreyer and W.R. Pitt Annular tautomerism: experimental observations and quantum mechanics calculations
J. CAMD 24, 6-7, 575-586, DOI: 10.1007/s10822-010-9345-5

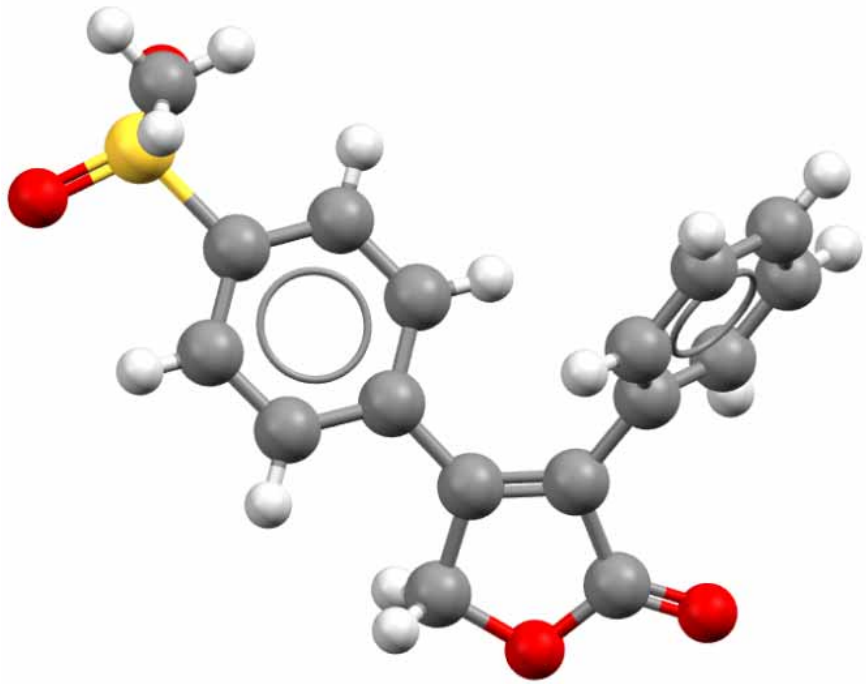
Impact of tautomers

EPO 679-157 (Searle, Pfizer, Celebrex, $2b in yr 1)



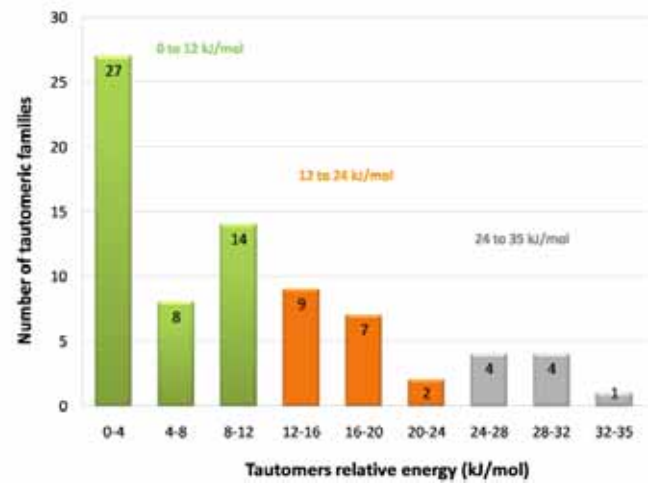
EPO 705-254 (Merck, Vioxx, $1.5b in yr 1)

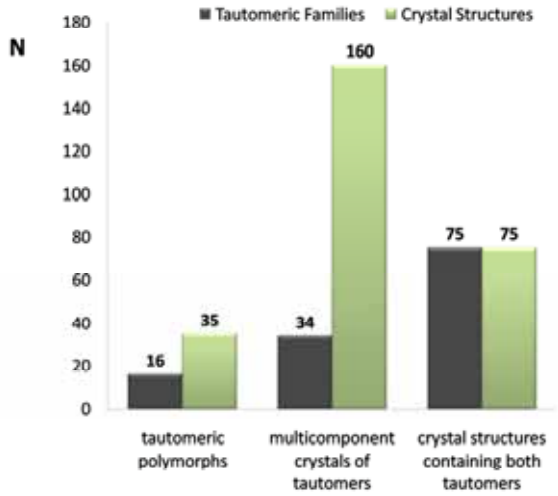


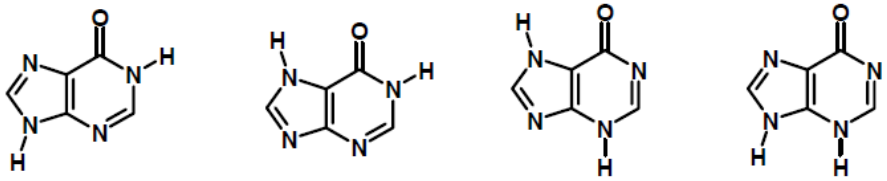


CAXMUJ

The CSD and tautomers



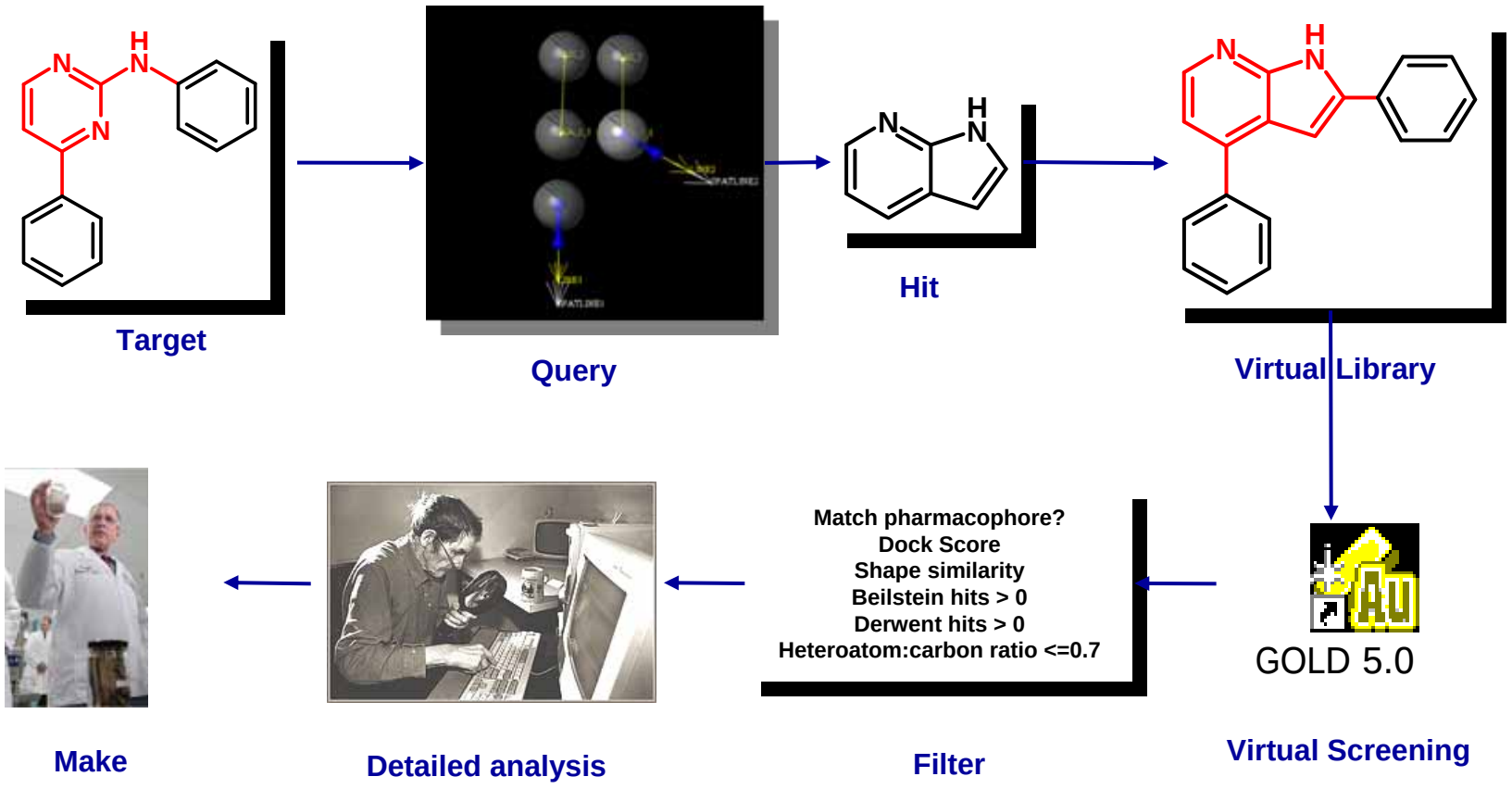




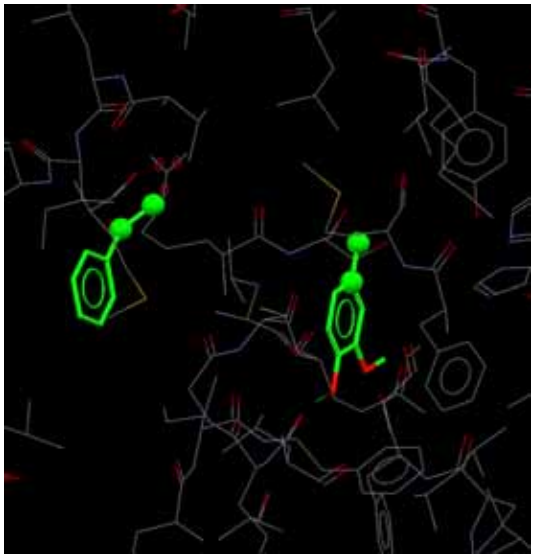
CSD occurrences	178	11	4	0
ΔE (kJ/mol) MP2/6-311++G** +PCM	0	2.14	20.85	34.79

A. J. Cruz-Cabeza and C. R. Groom Identification, Classification and Relative Stability of Tautomers in the Cambridge Structural Database. Cryst. Eng. Comm. 2010, DOI: 10.1039/C0CE00123F

Using VEHICLe fragments to Scaffold Hop

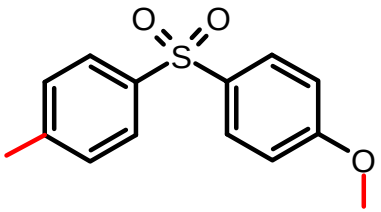


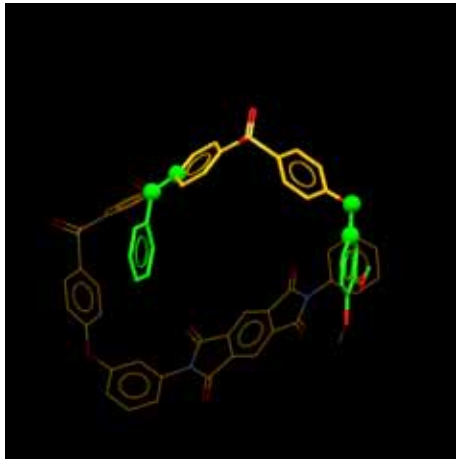
Fragment linking using the CSD

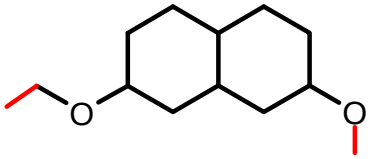


Two fragments bound to sub-pockets of PPARG

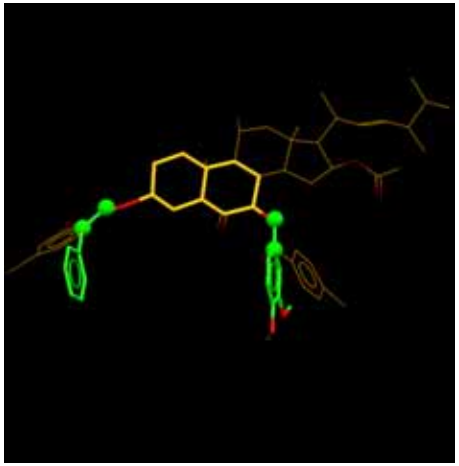
BALVEQ





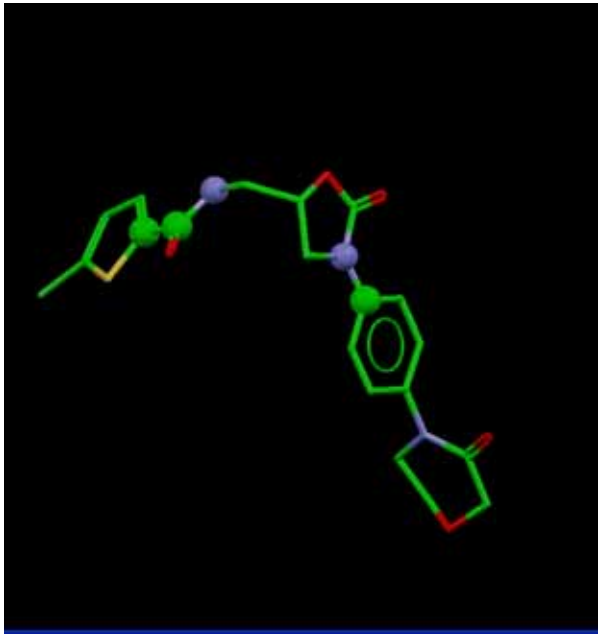


VAHKEV

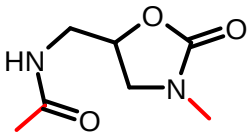


Fragment replacement using the CSD

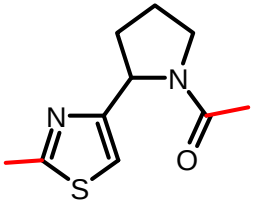
- Select atoms and bonds between which Bioisostere is sought
- Select a level of alignment quality
- Search

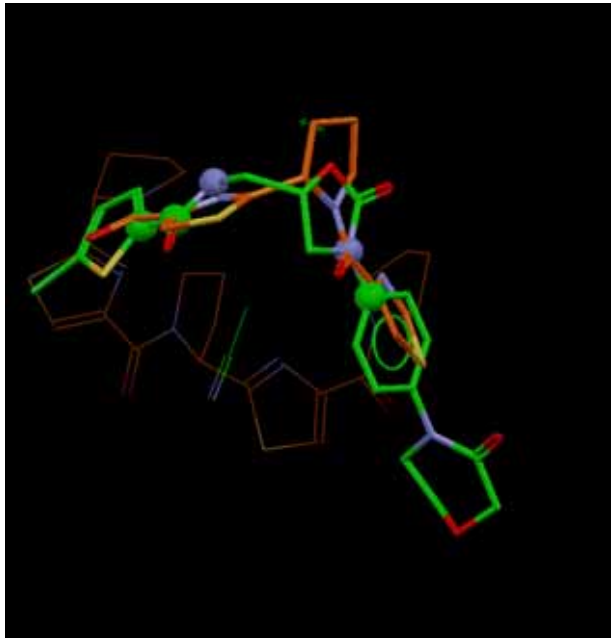


Potent Factor Xa Ligand (from 1w26)

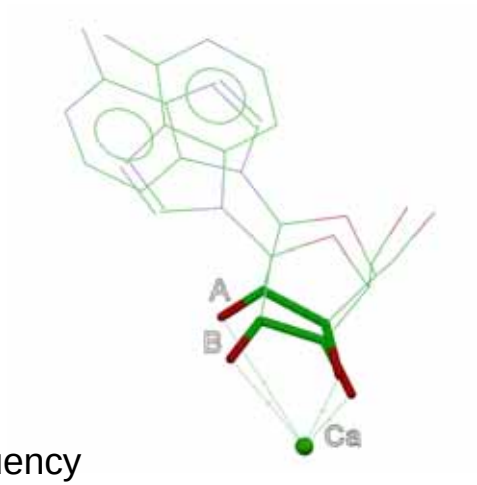


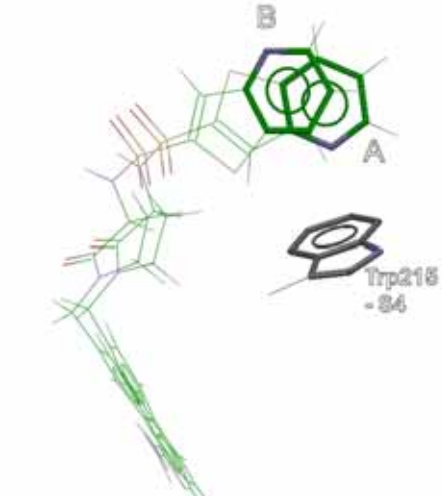
EKOLUL

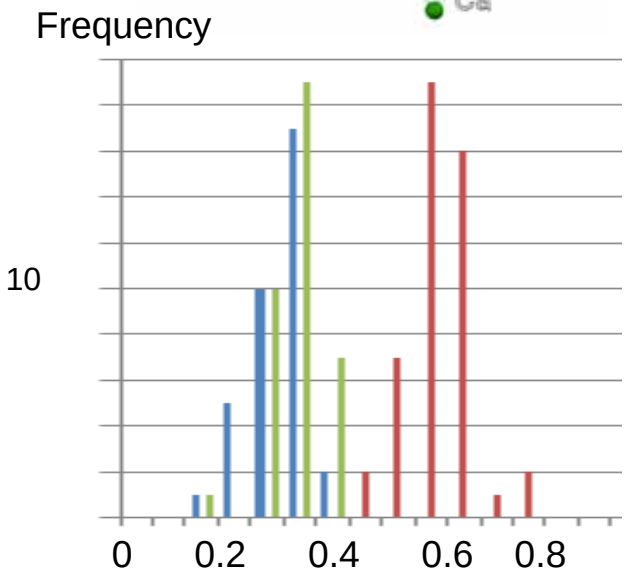




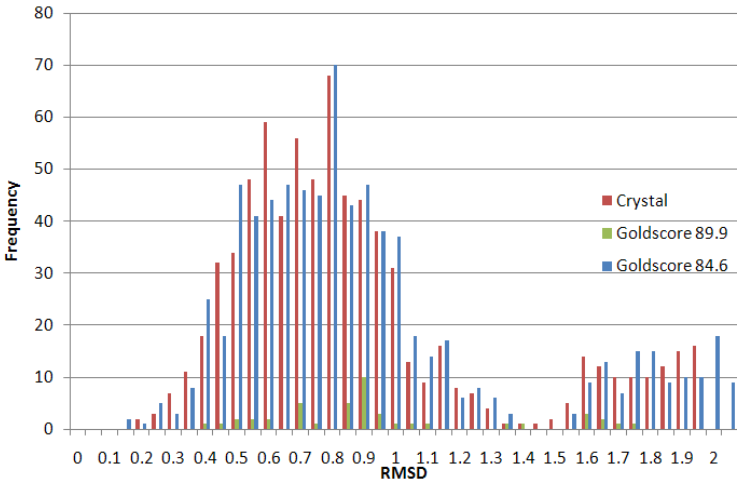
Validation of fitted fragments using the CSD



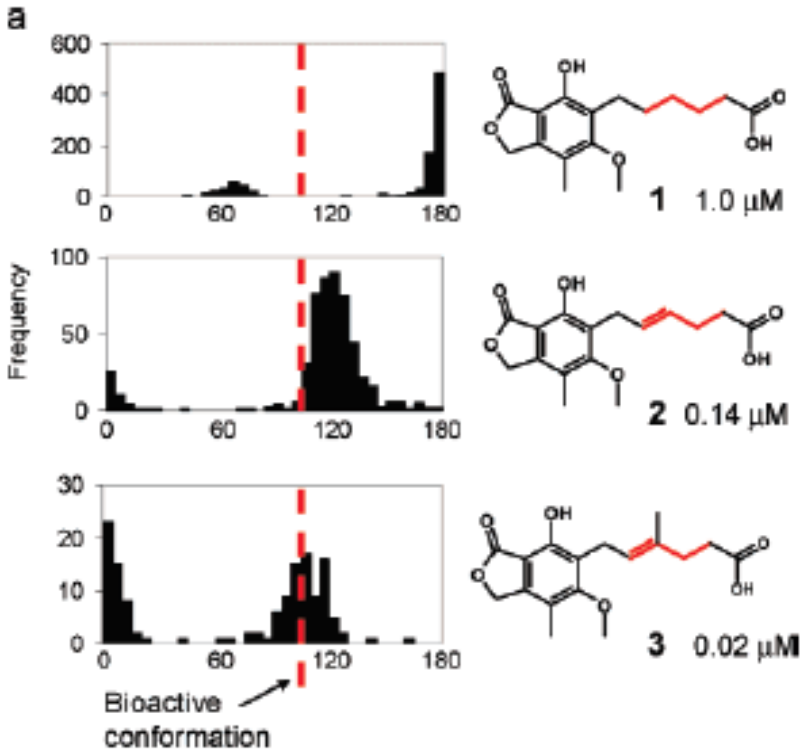


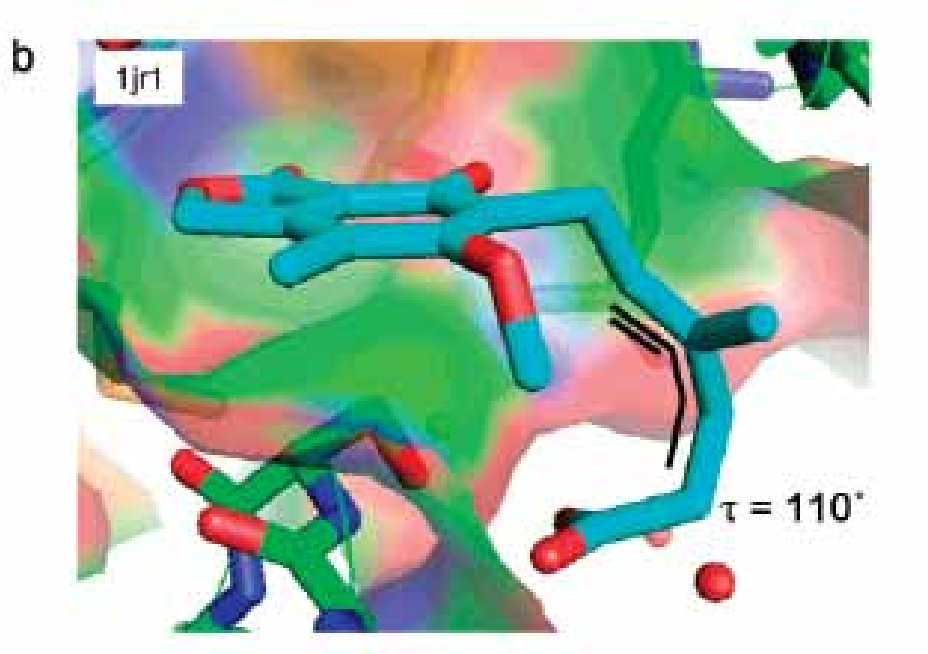


Histogram of RMSD's for CSD instances of this coordination geometry



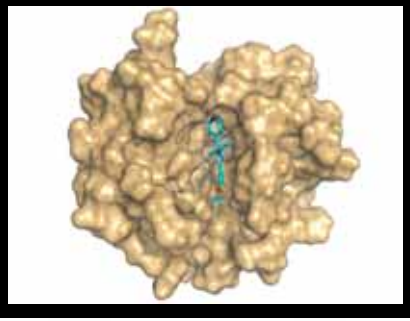
Designing in the right linker conformation





Brameld. K.A., Kuhn, B., Reuter, D.C. and Stahl, M. *J. Chem. Inf. Mod*, 48(1), 1-24 (2008)

Designing to the right target



cross-docking

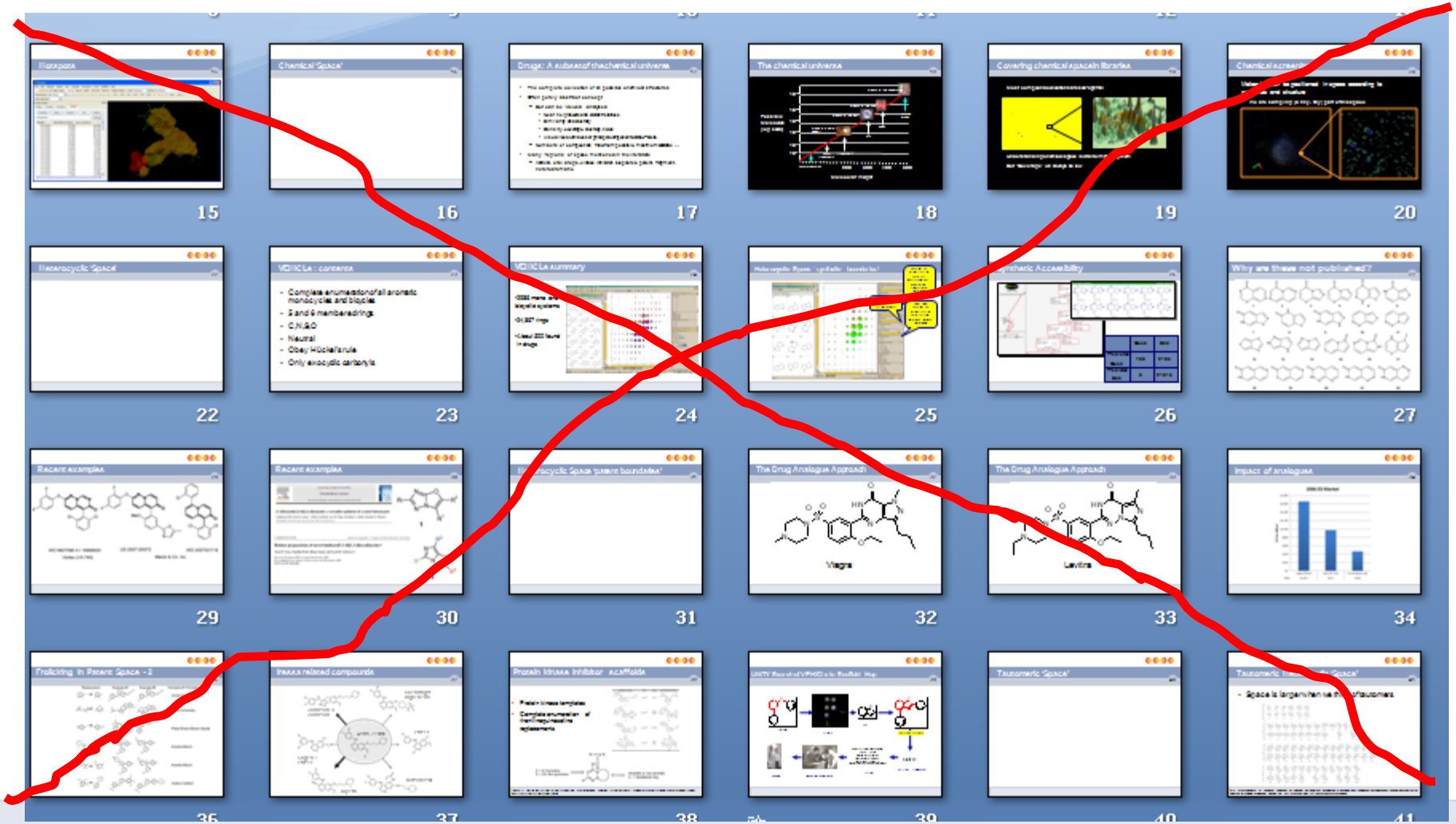
correct | wrong

ensemble docking

correct | wrong

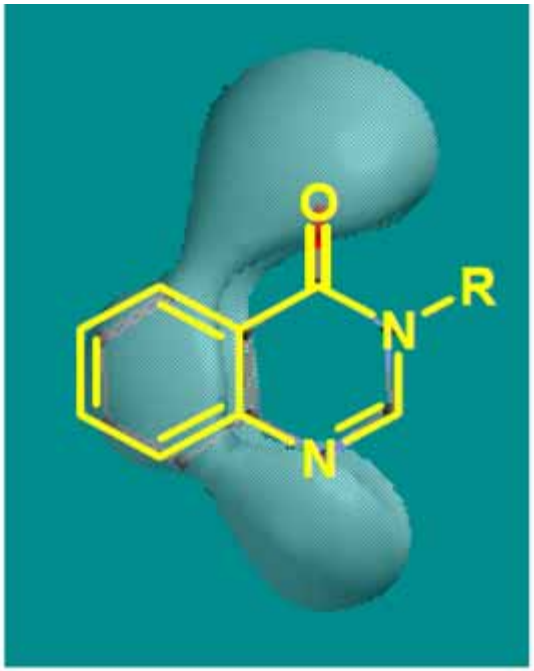
ligand	protein														
	1a4q	1b9t	1b9v	1f8b	1f8c	1f8d	1f8e	1inf	1inv	1inw	1mwe	1nnc	2qwi	2qwj	2qwk
1a4q	**0.30 (106.92)**	2.20 (66.42)	1.11 (69.69)	11.82 (83.27)	1.07 (91.99)	1.28 (83.86)	1.05 (88.94)	1.88 (65.40)	1.47 (74.38)	1.37 (83.16)	1.08 (88.88)	0.98 (90.77)	0.52 (83.51)	0.41 (86.58)	0.47 (96.74)
1b9t	0.91 (73.98)	**0.81 (69.57)**	0.93 (67.12)	1.32 (74.46)	1.43 (70.09)	1.34 (71.34)	1.20 (68.76)	1.19 (60.51)	1.50 (67.40)	1.01 (67.04)	1.18 (73.36)	1.13 (69.36)	3.95 (61.63)	1.10 (59.16)	1.11 (63.16)
1b9v	1.24 (81.18)	0.84 (71.87)	**0.93 (74.31)**	1.37 (74.19)	1.28 (73.23)	1.33 (72.40)	1.06 (68.63)	1.32 (70.99)	1.18 (68.45)	0.99 (69.86)	0.99 (76.20)	1.36 (74.23)	1.18 (66.22)	1.18 (67.91)	1.09 (74.24)
1f8b	0.55 (75.13)	0.57 (57.41)	0.71 (58.67)	**0.42 (69.12)**	0.47 (71.61)	0.52 (66.35)	0.50 (68.51)	4.02 (52.51)	0.77 (57.95)	0.90 (60.01)	0.37 (71.63)	0.40 (69.21)	2.13 (61.23)	0.46 (63.81)	1.09 (66.46)
1f8c	0.59 (84.36)	1.14 (62.93)	0.53 (63.59)	0.34 (71.37)	**0.42 (80.32)**	0.45 (69.30)	0.44 (76.92)	1.21 (62.56)	1.10 (63.67)	0.87 (67.94)	0.38 (74.62)	0.37 (76.55)	0.55 (64.82)	0.54 (73.86)	1.08 (74.95)
1f8d	1.20 (75.46)	4.13 (63.04)	0.65 (63.86)	1.66 (72.40)	0.42 (74.17)	**0.33 (70.56)**	0.51 (71.49)	4.18 (61.69)	0.78 (61.65)	0.82 (62.65)	0.47 (75.00)	0.54 (72.63)	1.15 (63.17)	0.54 (67.77)	1.18 (69.78)
1f8e	1.17 (84.95)	1.40 (66.56)	0.59 (66.19)	0.51 (74.83)	0.40 (82.94)	0.31 (72.76)	**0.48 (80.40)**	1.53 (62.74)	0.85 (67.27)	0.99 (70.55)	0.50 (77.60)	0.45 (79.54)	1.45 (66.15)	0.56 (75.92)	1.14 (78.96)
1inf	0.67 (57.45)	3.59 (51.69)	6.03 (49.64)	0.54 (56.39)	0.59 (55.04)	0.54 (55.53)	0.57 (53.26)	**3.65 (49.75)**	0.74 (51.12)	6.26 (47.32)	0.58 (58.88)	3.22 (47.43)	0.98 (45.81)	4.17 (48.31)	0.60 (51.80)
1inv	0.59 (77.25)	0.67 (59.65)	0.69 (62.75)	0.73 (74.62)	0.67 (76.32)	0.69 (73.76)	0.73 (75.95)	3.88 (63.81)	**1.16 (66.43)**	0.65 (66.14)	0.69 (76.65)	0.82 (74.12)	0.92 (70.39)	0.80 (70.98)	0.81 (72.36)
1inw	2.19 (82.59)	2.16 (62.15)	0.77 (64.72)	0.86 (73.99)	0.83 (78.13)	0.93 (74.62)	0.95 (79.08)	3.97 (66.97)	0.92 (66.61)	**0.77 (70.66)**	0.89 (75.76)	1.10 (76.26)	1.18 (68.61)	0.99 (72.91)	1.13 (72.91)
1mwe	0.62 (82.67)	0.69 (64.17)	0.72 (69.40)	0.41 (75.17)	0.42 (76.35)	0.54 (76.13)	0.51 (75.71)	1.13 (58.37)	0.81 (62.86)	0.69 (64.70)	**0.41 (79.73)**	0.50 (75.33)	0.63 (66.59)	0.51 (70.47)	0.92 (69.27)
1nnc	0.47 (88.18)	0.56 (64.32)	0.45 (68.42)	0.31 (75.66)	0.32 (83.90)	0.34 (77.83)	0.35 (82.55)	2.70 (66.71)	0.46 (64.15)	0.51 (64.03)	0.26 (78.66)	**0.77 (82.31)**	0.42 (78.11)	0.41 (80.05)	1.02 (81.03)
2qwi	1.17 (83.54)	1.66 (58.56)	1.29 (59.95)	0.56 (72.97)	1.47 (81.51)	0.66 (73.32)	0.47 (80.58)	2.02 (65.07)	1.57 (62.39)	8.92 (65.07)	0.65 (76.43)	0.65 (80.85)	**1.05 (78.49)**	1.03 (78.50)	1.04 (82.78)
2qwj	0.40 (86.52)	0.83 (62.10)	0.86 (60.88)	0.67 (73.22)	0.68 (81.05)	0.64 (70.85)	0.78 (77.46)	1.17 (64.23)	0.96 (65.54)	1.21 (65.59)	0.75 (75.16)	0.69 (74.06)	0.96 (67.21)	**0.31 (74.91)**	0.46 (81.61)
2qwk	0.37 (86.54)	0.90 (63.74)	0.76 (62.88)	1.03 (72.94)	1.04 (80.18)	1.03 (70.65)	1.05 (77.23)	1.06 (66.25)	1.18 (66.44)	1.18 (67.39)	0.91 (74.18)	1.06 (74.23)	0.78 (70.11)	0.39 (74.18)	**0.50 (82.17)**

Forget all this rubbish



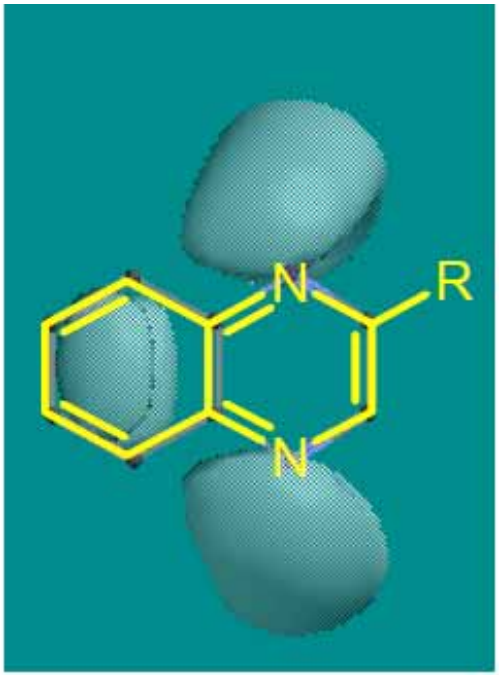
Interaction Space

- It’s what’s on the outside that matters



Active

Inactive



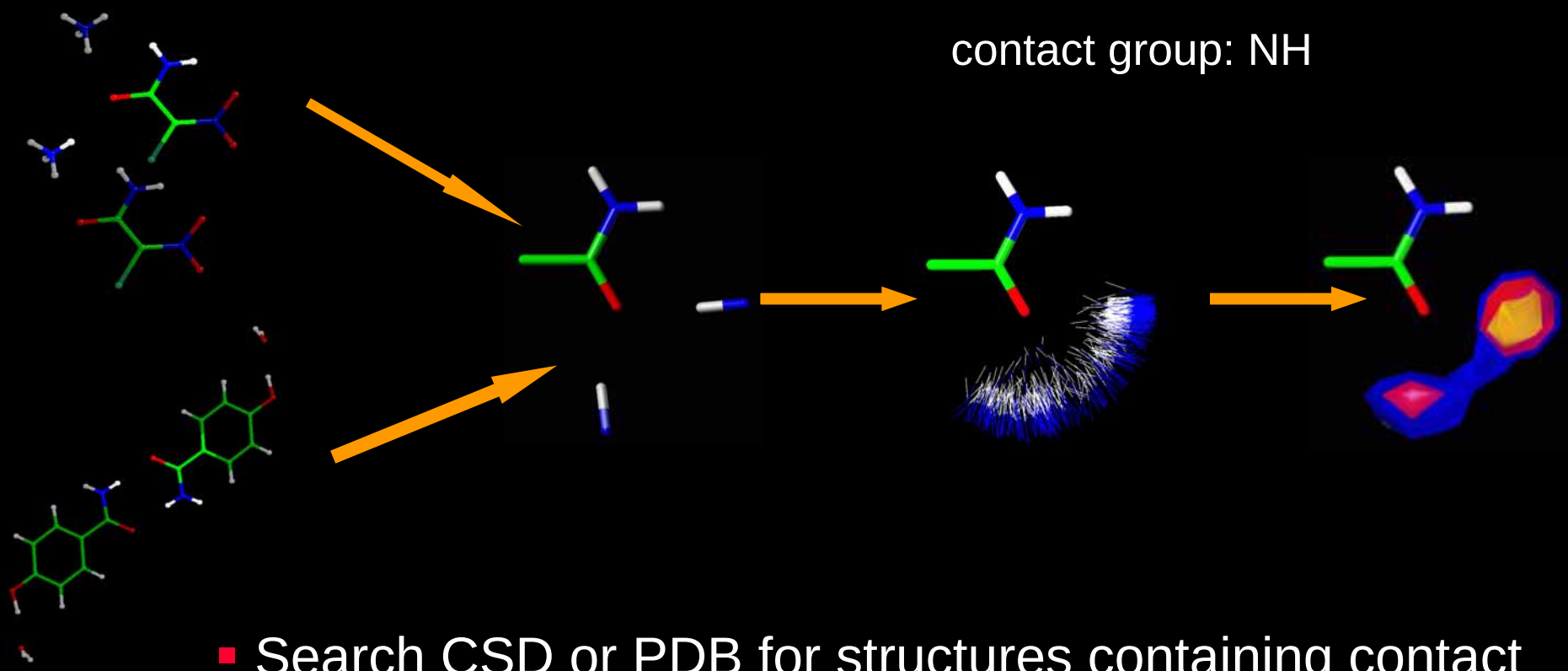
Active

IRAK-4 inhibitors. Part II: A structure-based assessment of imidazo[1,2-a]pyridine binding
George M. Buckley, Thomas A. Ceska, Joanne L. Fraser, Lewis Gowers, Colin R. Groom, Alicia Perez Higueruelloa, Kerry Jenkinsa, Stephen R. Mack, Trevor Morgan, David M. Parry, William R. Pitt, Oliver Rausch, Marianna D. Richard and Verity Sabin Organic & Medicinal Chemistry Letters (2008), 18, 11, 1, 3291-3295

Interaction Space

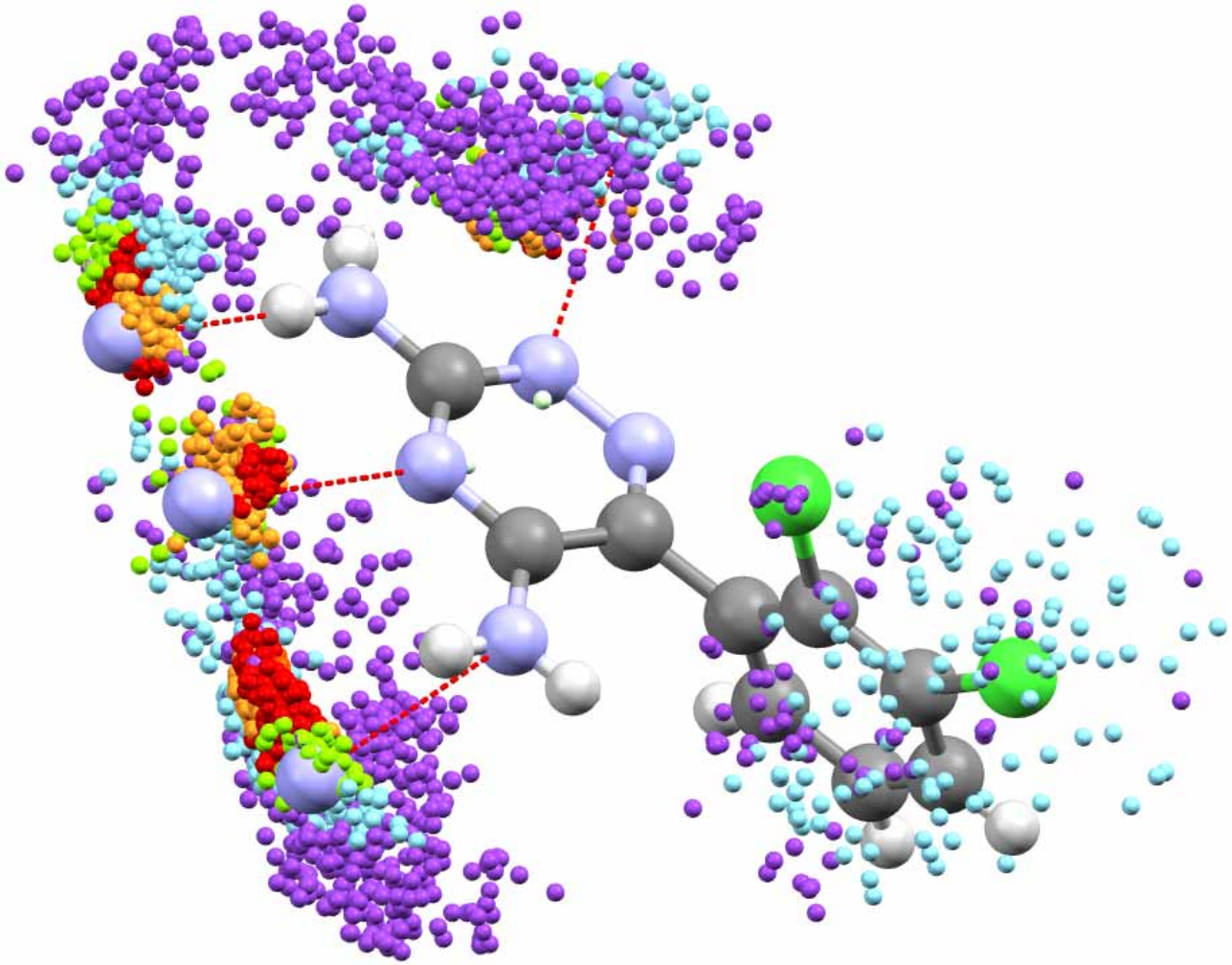
central group: -$CONH_2$

contact group: NH

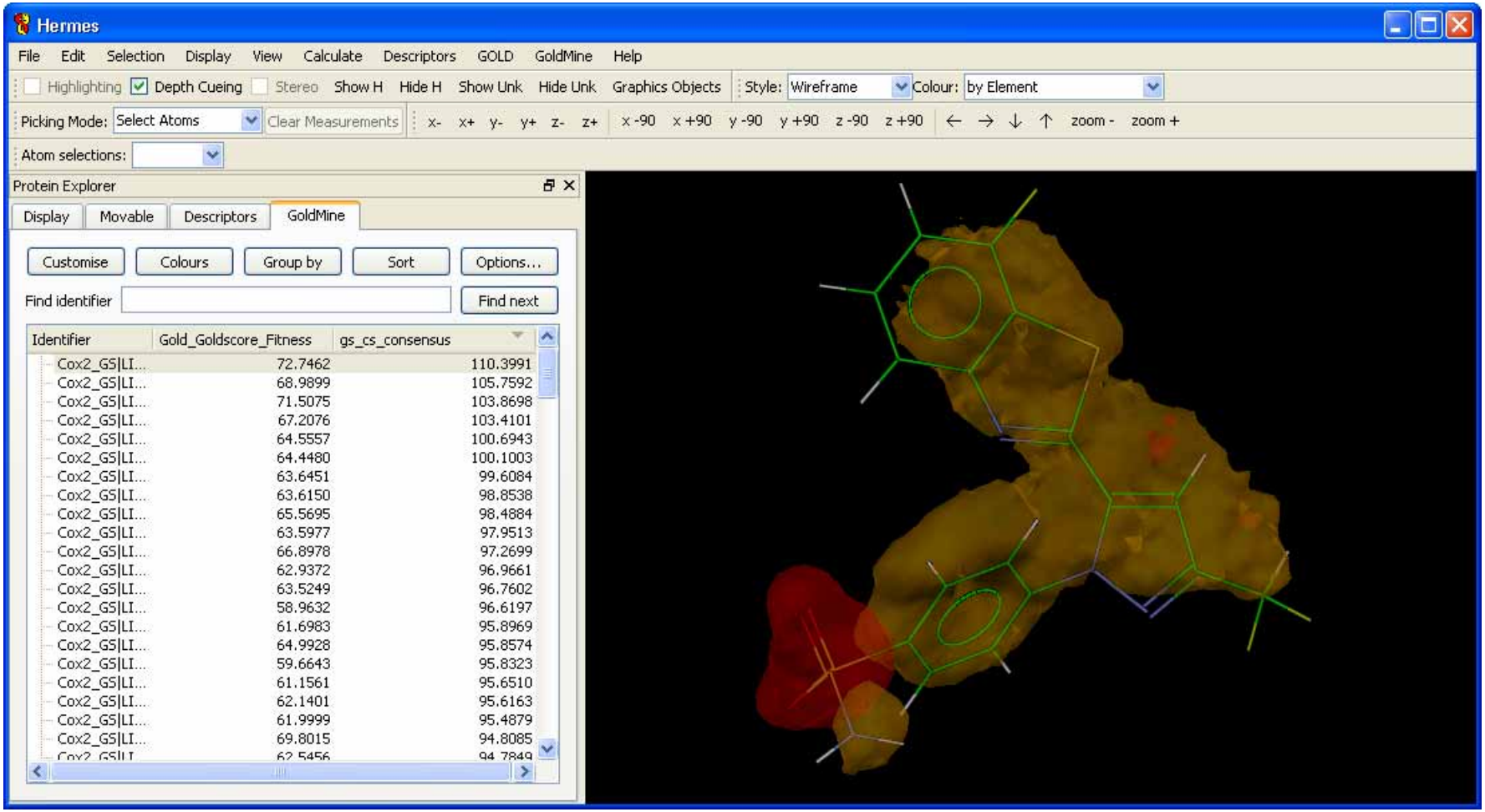


- Search CSD or PDB for structures containing contact
- Superimpose hits and display distribution

Interactions propensities can be normalised



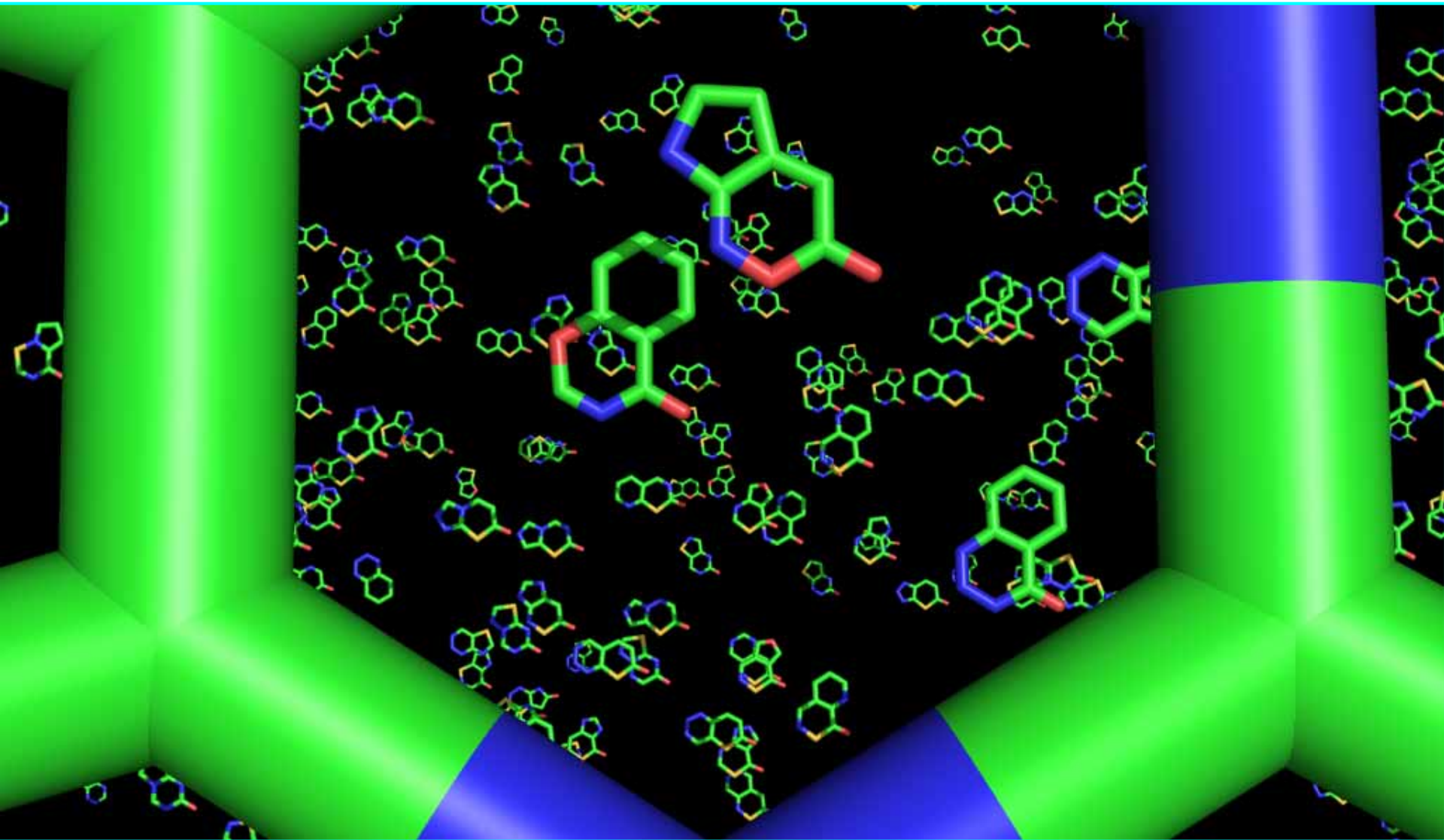
Hotspots


Hermes
File Edit Selection Display View Calculate Descriptors GOLD GoldMine Help
Highlighting Depth Cueing Stereo Show H Hide H Show Unk Hide Unk Graphics Objects Style: Wireframe Colour: by Element
Picking Mode: Select Atoms Clear Measurements x- x+ y- y+ z- z+ x -90 x +90 y -90 y +90 z -90 z +90 zoom - zoom +
Atom selections:
Protein Explorer
Display Movable Descriptors GoldMine
Customise Colours Group by Sort Options...
Find identifier Find next
Identifier Gold_Goldscore_Fitness gs_cs_consensus
Cox2_GS|LI... 72.7462 110.3991
Cox2_GS|LI... 68.9899 105.7592
Cox2_GS|LI... 71.5075 103.8698
Cox2_GS|LI... 67.2076 103.4101
Cox2_GS|LI... 64.5557 100.6943
Cox2_GS|LI... 64.4480 100.1003
Cox2_GS|LI... 63.6451 99.6084
Cox2_GS|LI... 63.6150 98.8538
Cox2_GS|LI... 65.5695 98.4884
Cox2_GS|LI... 63.5977 97.9513
Cox2_GS|LI... 66.8978 97.2699
Cox2_GS|LI... 62.9372 96.9661
Cox2_GS|LI... 63.5249 96.7602
Cox2_GS|LI... 58.9632 96.6197
Cox2_GS|LI... 61.6983 95.8969
Cox2_GS|LI... 64.9928 95.8574
Cox2_GS|LI... 59.6643 95.8323
Cox2_GS|LI... 61.1561 95.6510
Cox2_GS|LI... 62.1401 95.6163
Cox2_GS|LI... 61.9999 95.4879
Cox2_GS|LI... 69.8015 94.8085

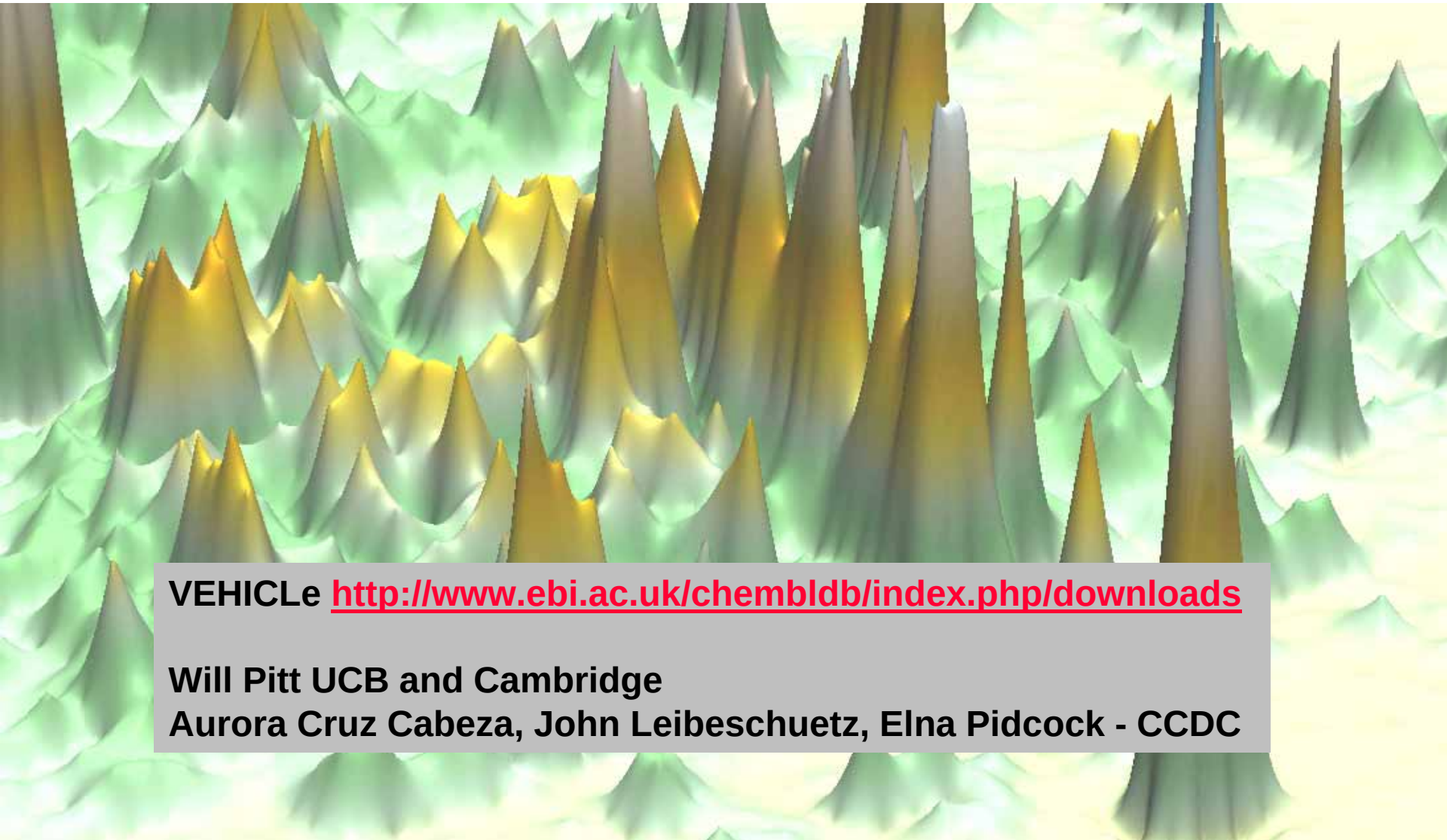
Final thoughts

- Be careful comparing ligand efficiency values
- But the most interesting regions of fragment space are computationally accessible
- Use available data
- The problems and opportunities aren’t just technical and scientific
  - Significant IP implications
- Interactions matter

The End



Acknowledgements



VEHICLe http://www.ebi.ac.uk/chembldb/index.php/downloads

Will Pitt UCB and Cambridge
Aurora Cruz Cabeza, John Leibeschuetz, Elna Pidcock - CCDC