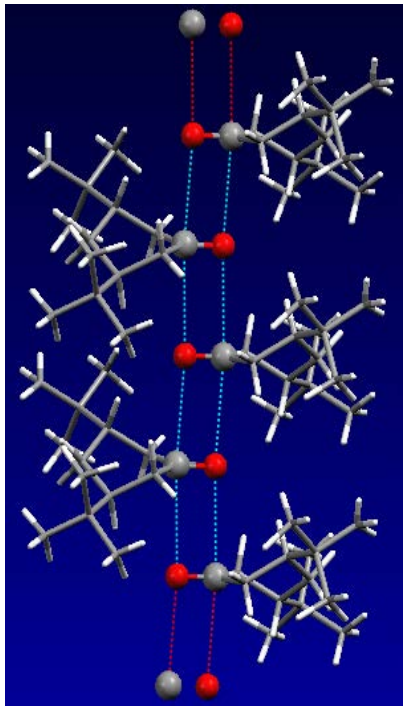
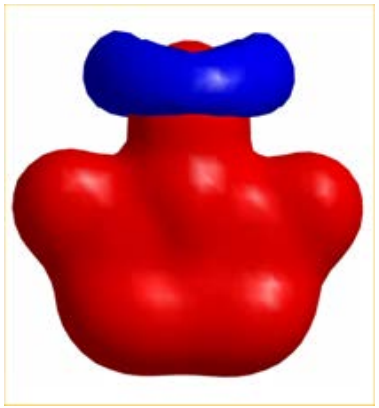




Experiment and Theory in the study of Intermolecular Interactions

Frank H Allen

Cambridge Crystallographic Data Centre (CCDC)

12 Union Road, Cambridge

[allen@ccdc.cam.ac.uk]

SCI, Cambridge, 21 March 2013

Standing on the shoulders …. etc.

"There are Agents in Nature able to make the Particles of Bodies stick together by very strong Attractions, and it is the Business of experimental Philosophy to find them out"

Isaac Newton, Optiks (1718)

*[cited by H.A.Bent, Chem. Rev., **16**, 588-648, 1968]*

"Those things whose textures fall so aptly contrary to one another that hollows fit solids, each in the one and the other, make the best joining."

Lucretius (99BC – 55BC) De Rerum Natura

*[cited by J.D.Dunitz & A.Gavezzotti, Chem.Soc.Rev., **8**, 2622-2633, 2009]*

Some ‘Agents’ stick better than others – when, where, how and why?

Experimental observations – the when and where

- Information principally from crystallography
- Large databases available: CSD ~ 650,000 structures, PDB ~ 90,000 structures. Both are used here

Computational modelling – the how and why

- Wide variety of methodologies available
- IMPT (Univ. Cambridge) is frequently used here

“Materials and Methods”

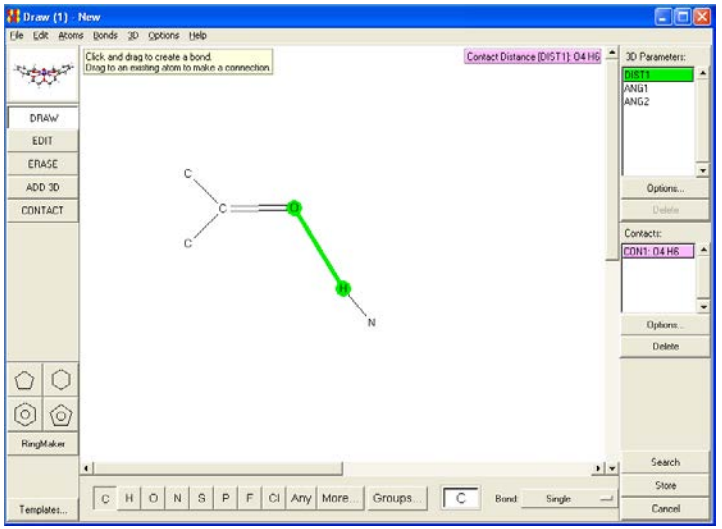
Experimental geometry – CSD Searches

[Bruno et al., Acta Cryst., B58, 389-397, 2002]

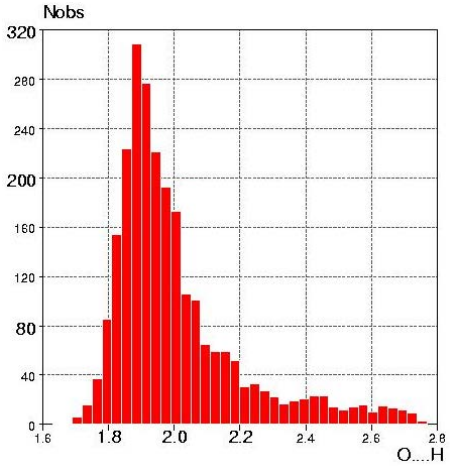
❖ Intermolecular Search:

Draw fragments

Define non-bonded contact and d-limits



❖ Data Analysis



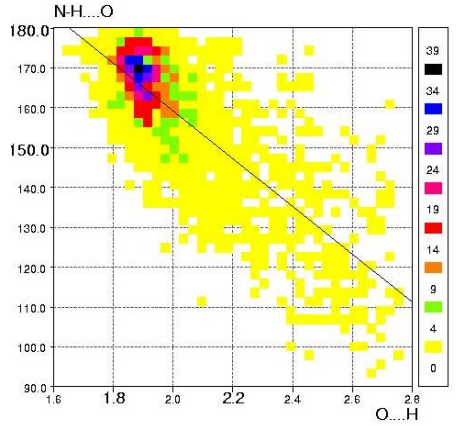
❖ Plots

d(H•••O)

d(H•••O)

vs.

θ(N–H•••O)



Intermolecular PerturbationTheory

[Hayes & Stone, Mol. Phys., 53, 84-98, 1984]

- Calculates interaction energy for fixed mutual orientations of (small) model molecules
- Use CSD to indicate preferred mutual orientations for exploration of the energy hypersurface
- Basis sets: 6-31-G, G* and G**, results free of basis set superposition errors
- Total energy E(t) as sum of individual components:

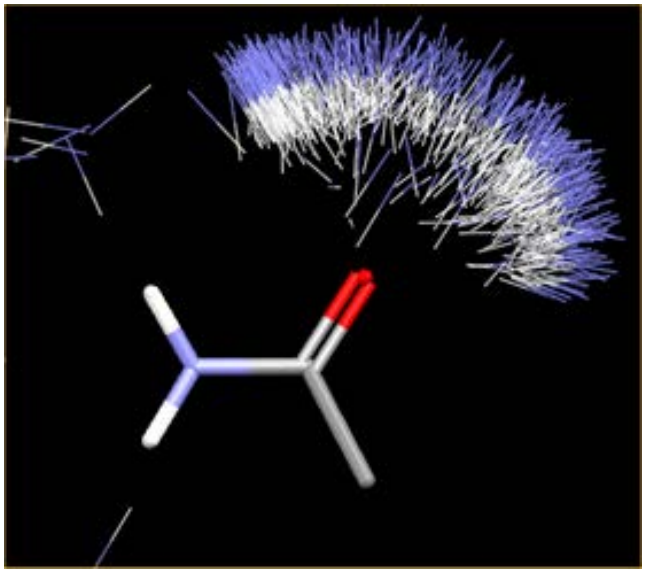
electrostatics **exchange repulsion**

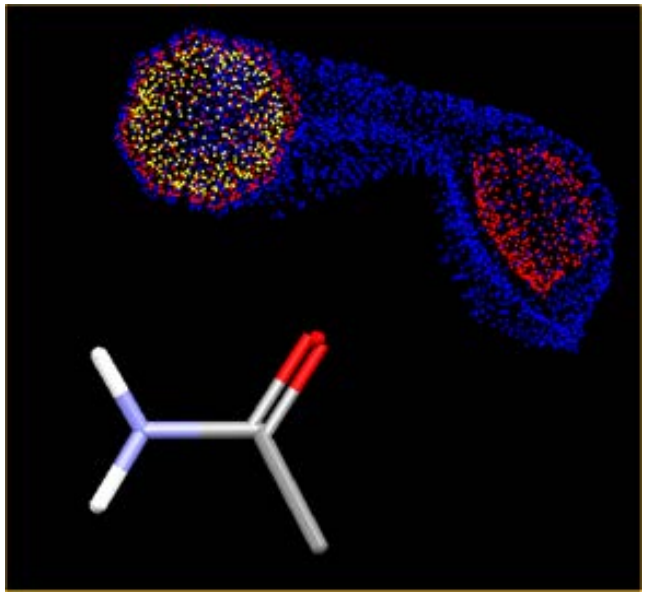
polarisation **charge transfer** **dispersion**

IsoStar Knowledge Base

[Bruno et al., JCAMD, 11, 525-537, 1997]

Maps interactions between contact groups and central groups





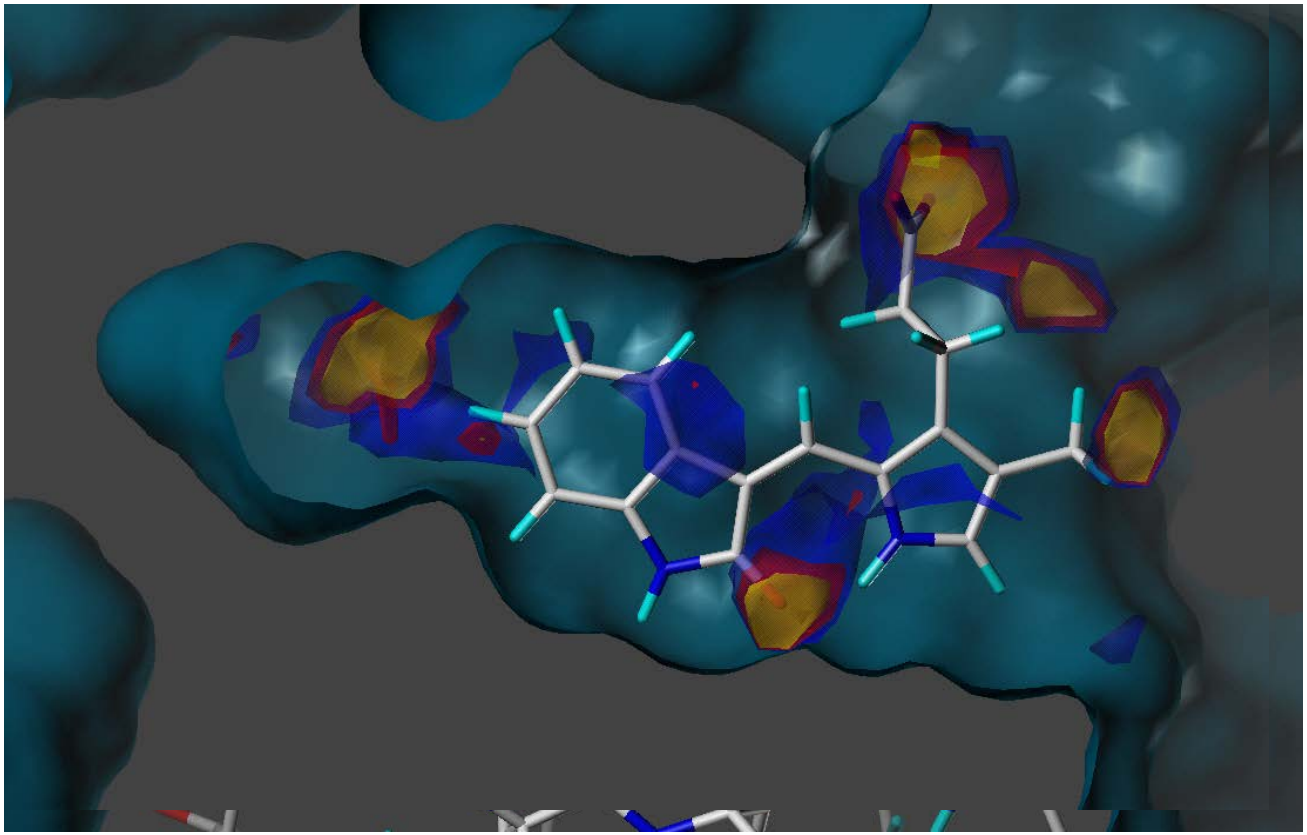
Experimental data

- From CSD and PDB protein–ligand complexes
- >300 central groups,
- >50 contact groups
- >22,000 CSD scatterplots
- >7,400 PDB scatterplots

Theoretical data (DMA/IMPT)

- >1,500 energy minima

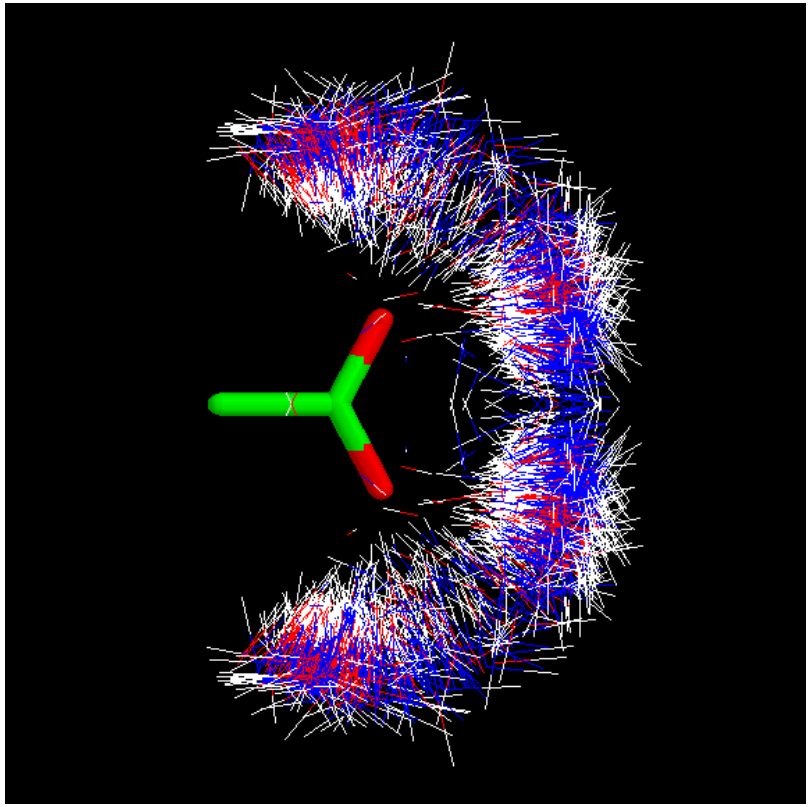
Superstar program uses IsoStar knowledge to predict functional group binding e.g. -C=O binding to Tyrosine kinase (1fgi)



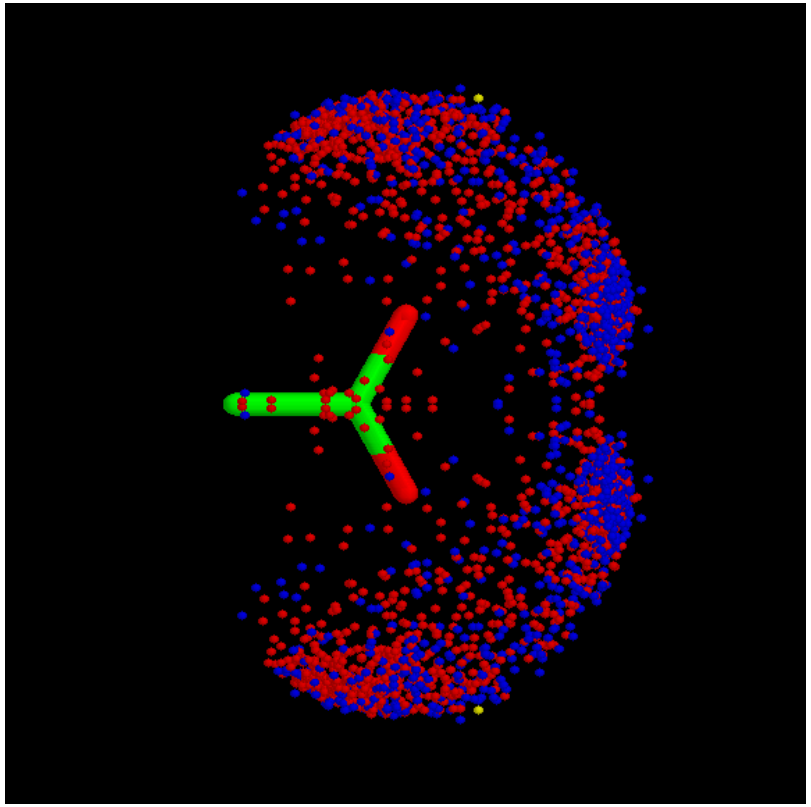
What can we learn from IsoStar

IsoStar KB: CSD *vs.* PDB data
N-H and O-H interactions with carboxylate

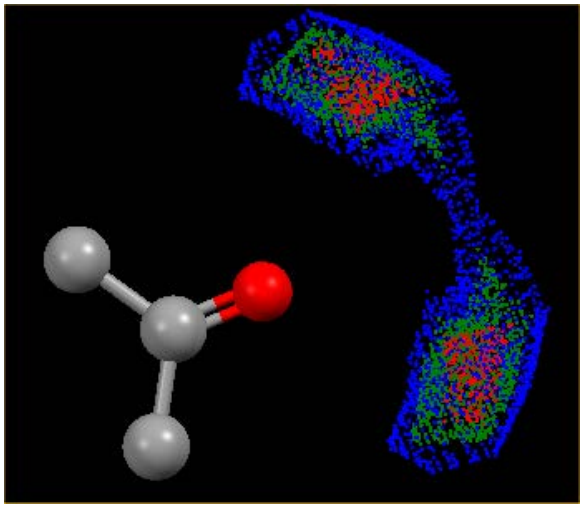
CSD Data



Protein-Ligand Complexes

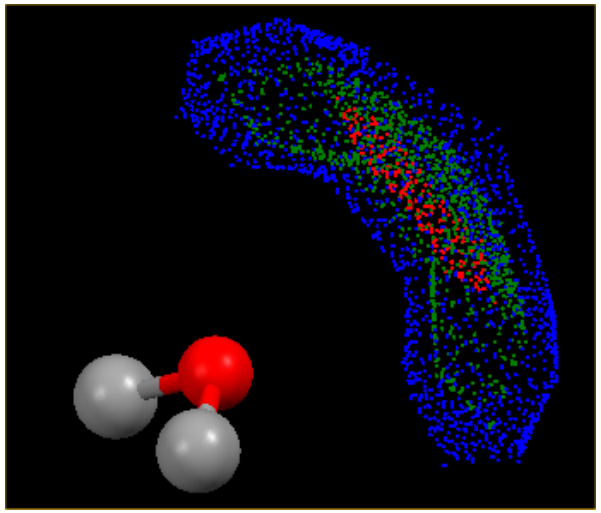


Not all O-acceptors are the same:
O-H•••O bonding to ketones, ethers and esters



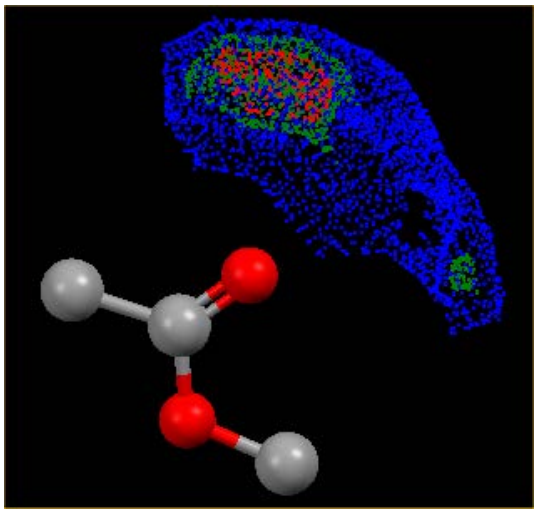
d(C=O•••H) 1.91(1)

E(C=O•••H) -24.7



d(C–O•••H) 1.94(2)

E(C–O•••H) -21.1



d(C=O•••H)1.92(2)

d(C-O•••H) 2.19(2)

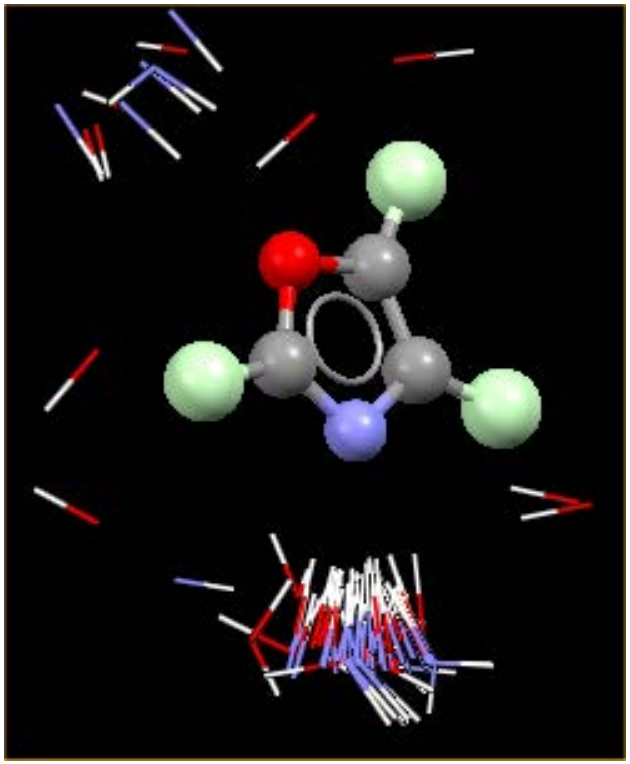
E(C=O•••H) -25.4

E(C–O•••H) -15.0

(E-values in kJ.mol^{-1})

Not all O-acceptors are the same:
H-bonding to oxazole (l) and isoxazole (r)

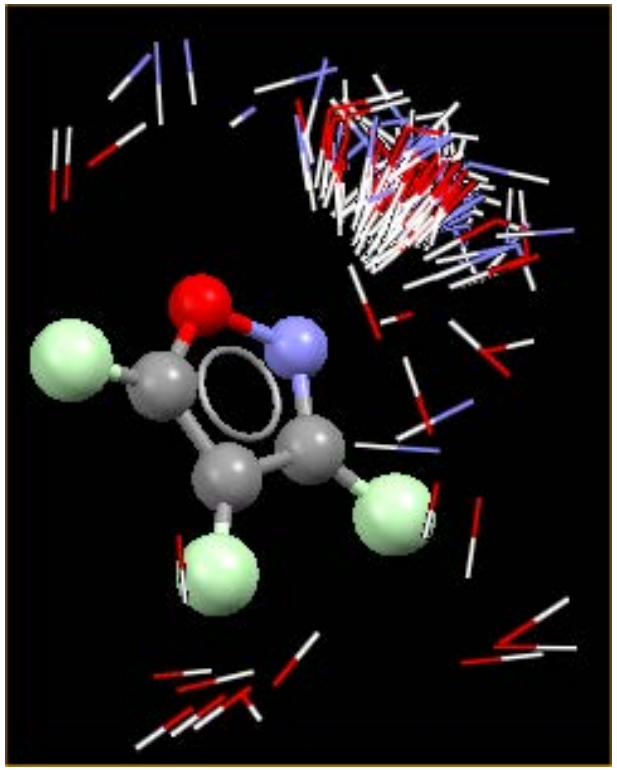
[Nobeli et al., J.Comp.Chem., 18, 2060-2074, 1997]



Oxazole

E(OH…N) -24.5

E(OH…O) -12.5



Isoxazole

E(OH…N) -21.9

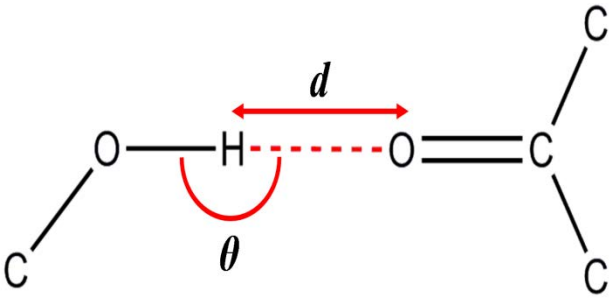
E(OH…O) -17.5

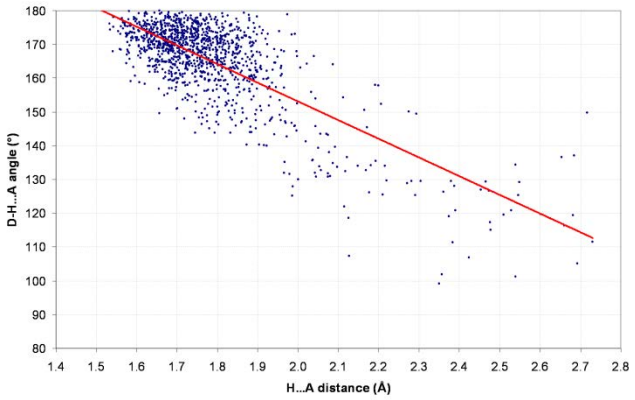
More detailed studies

Directionality of H-bonds at donor-H
and
Directionality of interactions at C-Halogen

H-bond directionality at donor-H

[Wood et al., CrystEngComm, 11, 1563-1571, 2009]





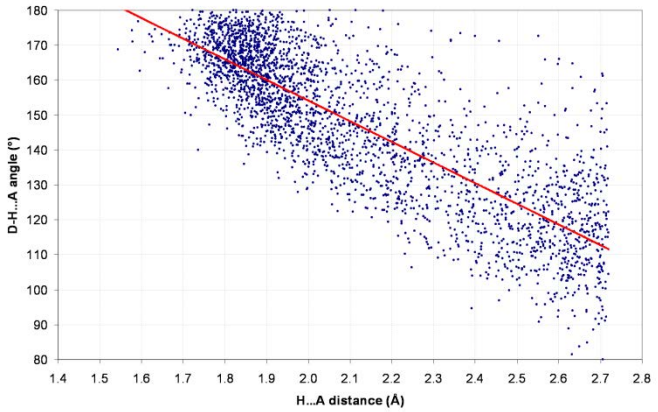
Plots of d vs θ for

O-H•••N(pyridine) [top] and

O-H•••O(ether) [bottom]

Neutron-normalised H-atoms

θ cut-off at 90°



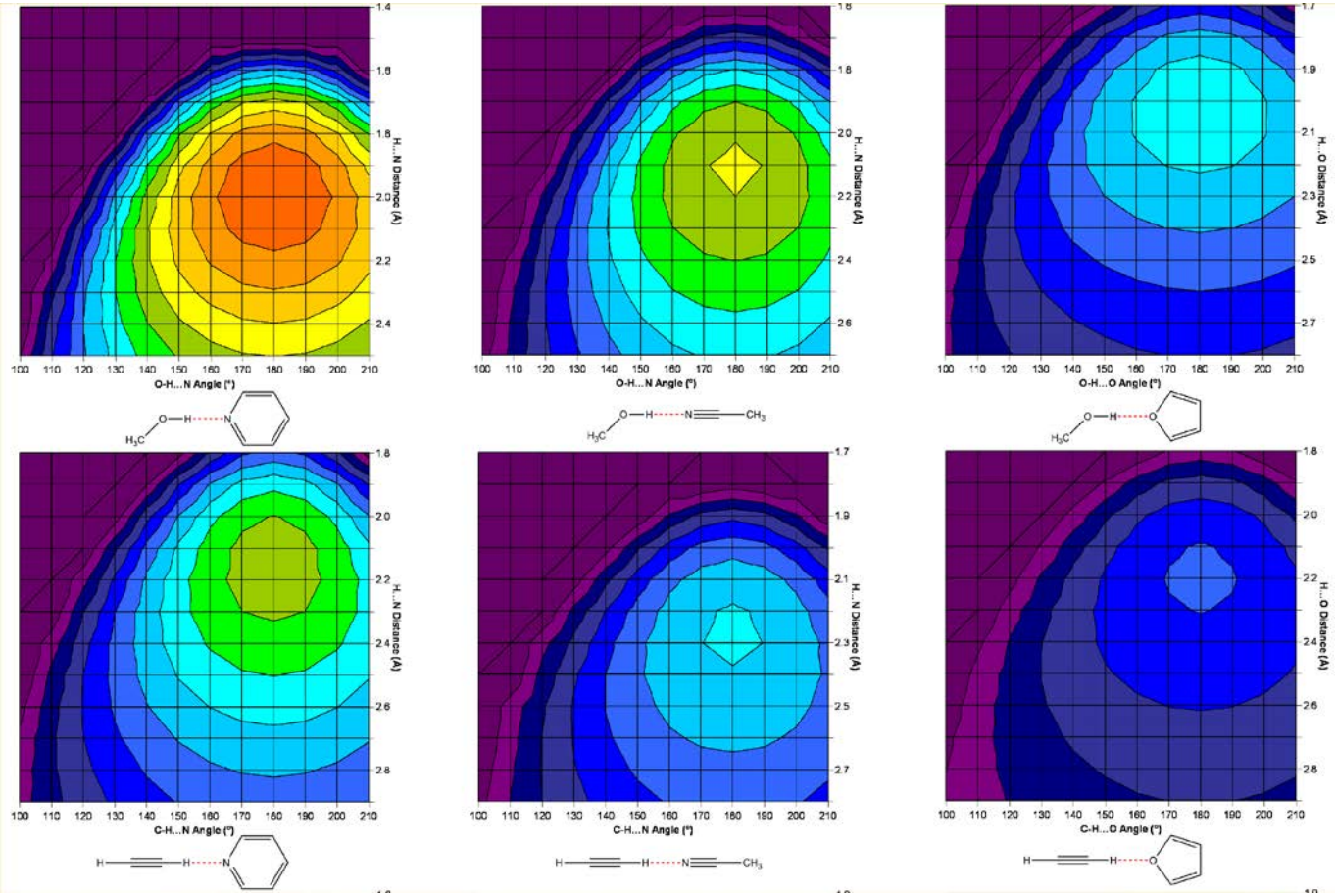
H-bond d vs. θ plots, coloured by IMPT interaction energy for six donor-acceptor pairs

Acceptor: pyridine-N CH_3–CN furan-O

Donor:

CH_3–OH

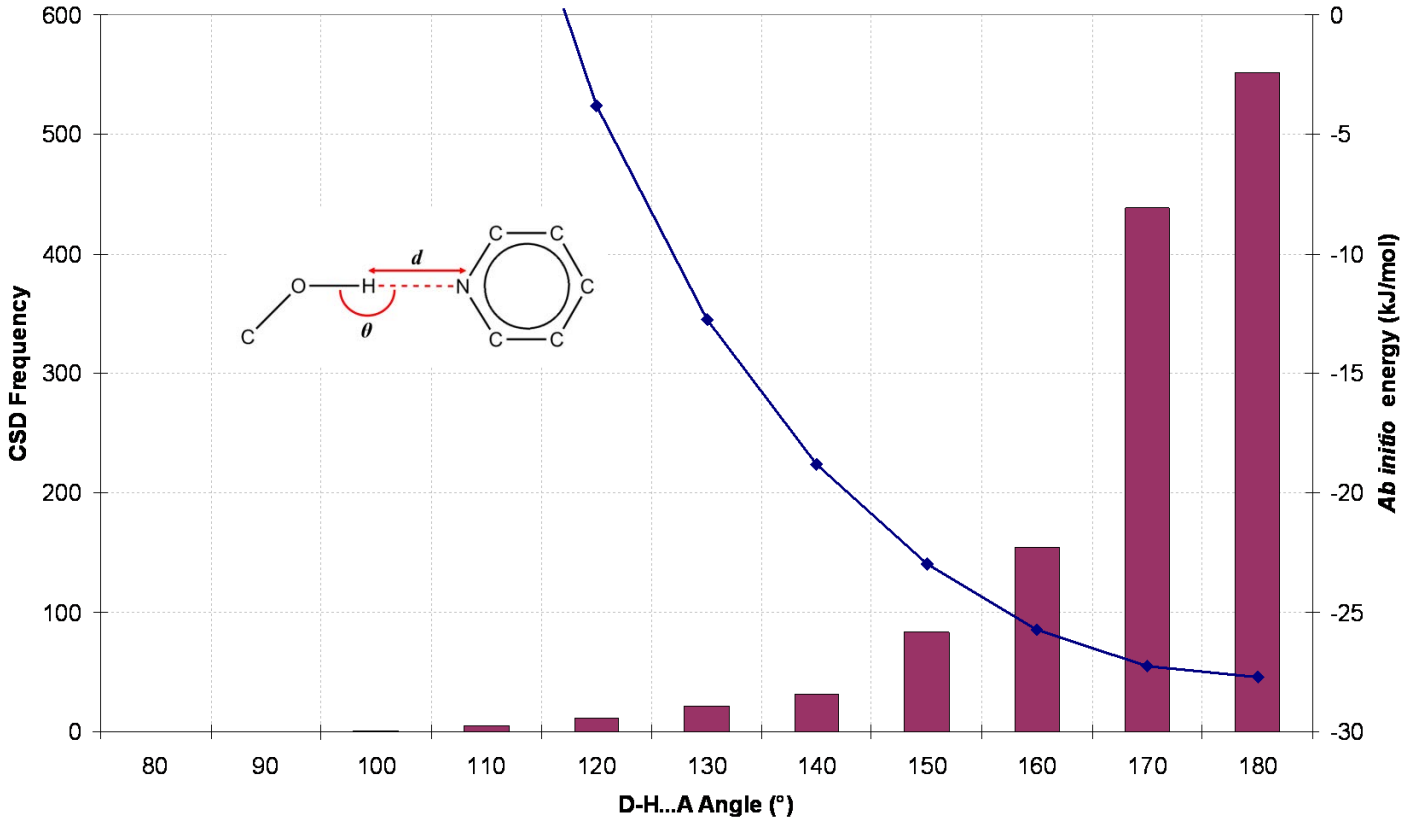
C≡C–H



Histogram of θ for O–H•••N(py) H-bonds with interaction energy curve superimposed

E_t-values
(kJ.mol^{-1})

θ	E_t
180°	-28
150°	-23
140°	-19
130°	-13
120°	-4



H-bond directionality at donor-H: summary

[Wood et al., CrystEngComm, 11, 1563-1571, 2009]

Nine H-bonds of varying strengths studied

- At θ = 140° E_t is ~50% of E_t (linear)
- At θ = 120° E_t is ~15% of E_t(linear)

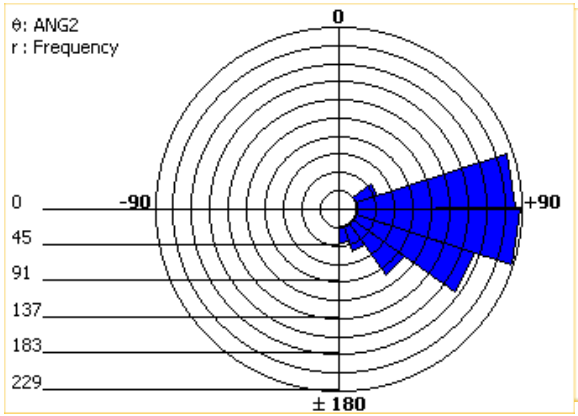
However

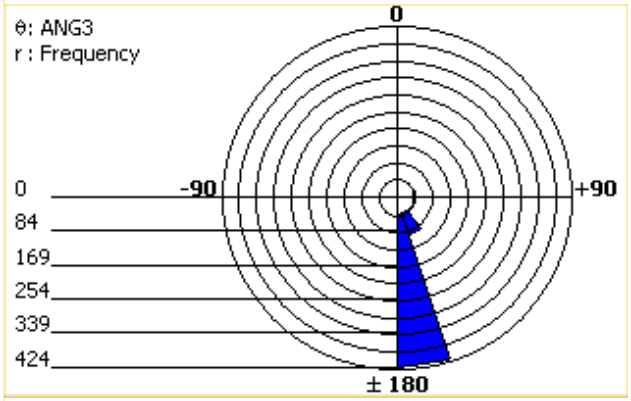
- For linear bonds, decrease in E_t with d(HB) is slower
- E_t at vdW limit can still be ~50% of E_t(linear)

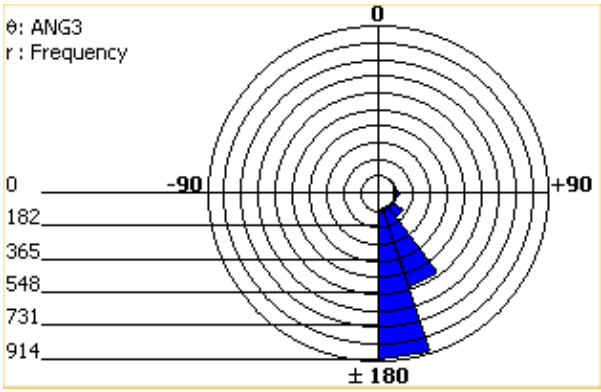
Recommend

- θ-limit of 120° for automated H-bond recognition, but
- Be aware of intramolecular H-bonds and H-bifurcation
- Look for near-linear H-bonds just above the vdW limit

Directionality of interactions at carbon-bound halogens (Hal = Cl, Br, I)





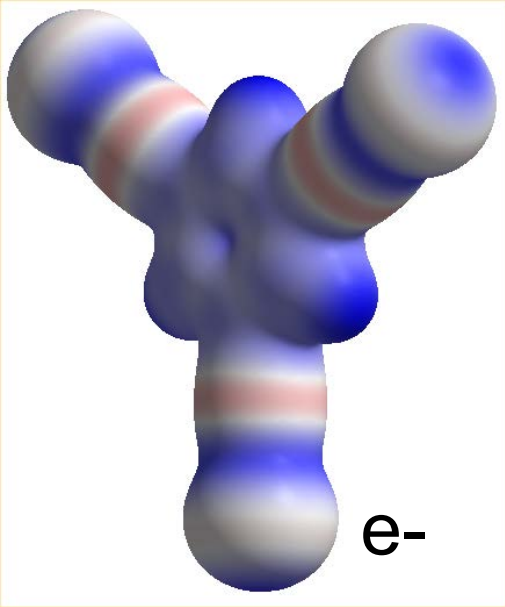


C
|
Hal•••H-N,O

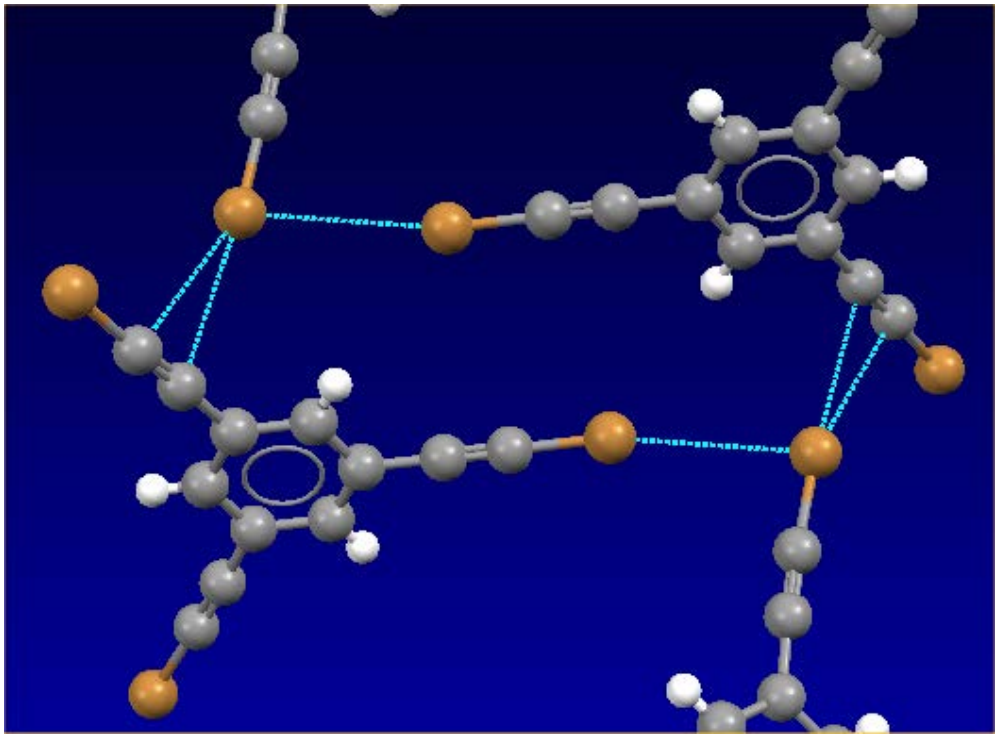
C
|
Hal
•
•
Hal^-

C
|
Hal
•
•
O(=C)

Directionality of C-Br interactions in tri(bromoethynyl)-benzene



EP Map (MOPAC)

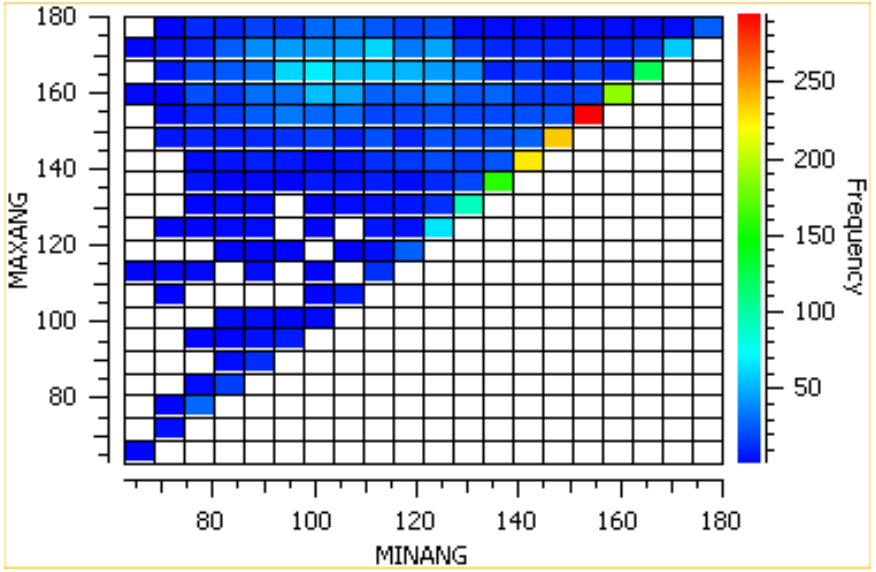


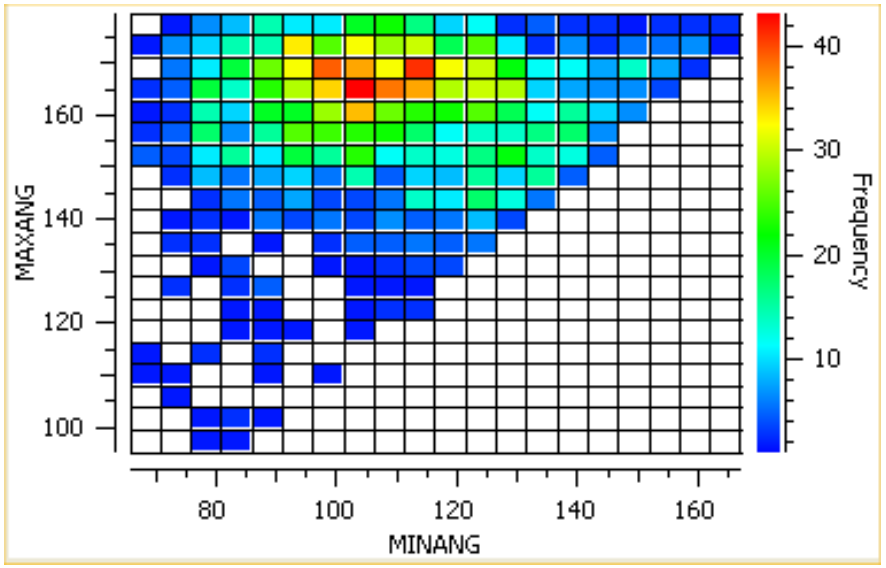
Leading references:

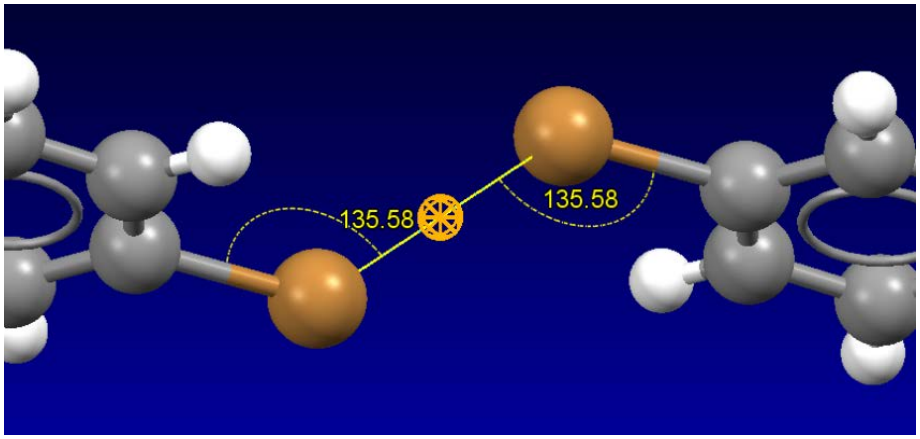
JACS, 118, 3108-3116, 1996; *Crystal Growth Des. 1, 277-290, 2001;*

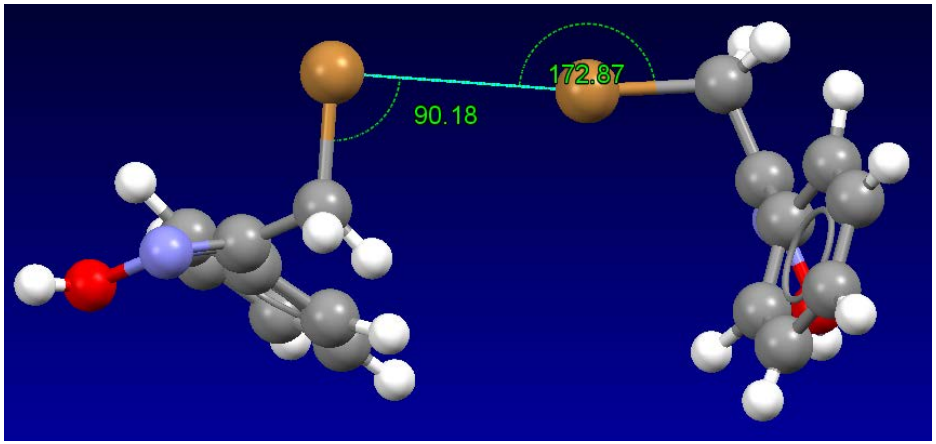
Int.J.Quantum Chem., 107, 3046-3052, 2007

Directionality of Hal•••Hal interactions between carbon-bound halogens (Hal = Cl, Br, I)

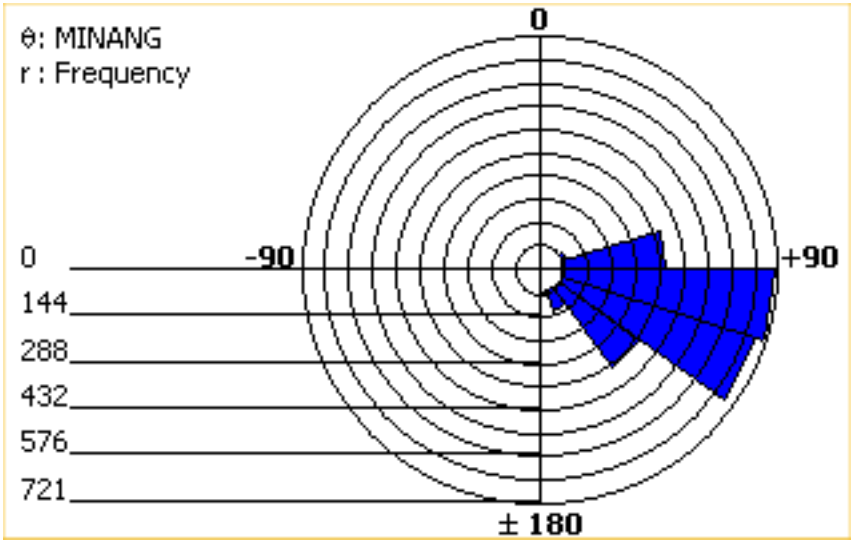


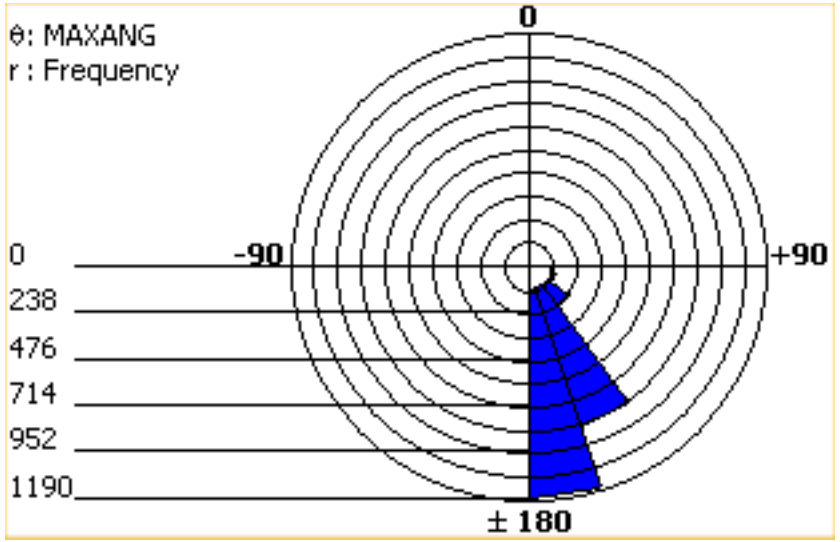






Directionality of Hal••Hal interactions between carbon-bound halogens (Hal = Cl, Br, I) [Inversion-related examples removed]



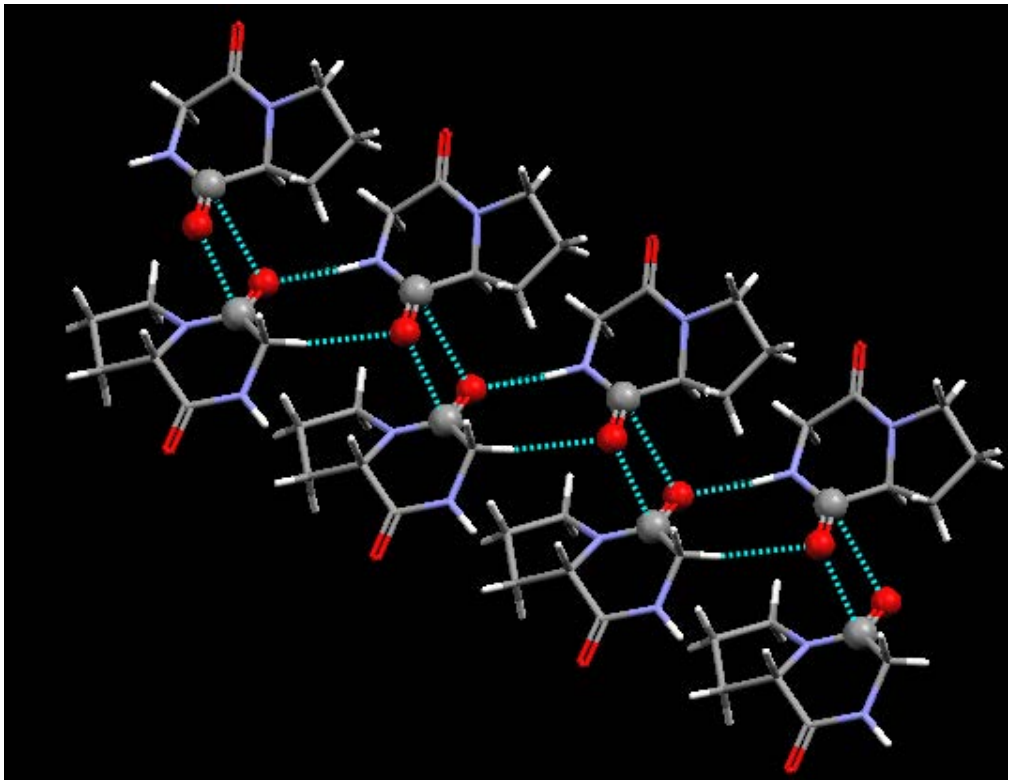


Wilcken et al., J.Med.Chem. 2013 (ASAP) [Focus on drug design]

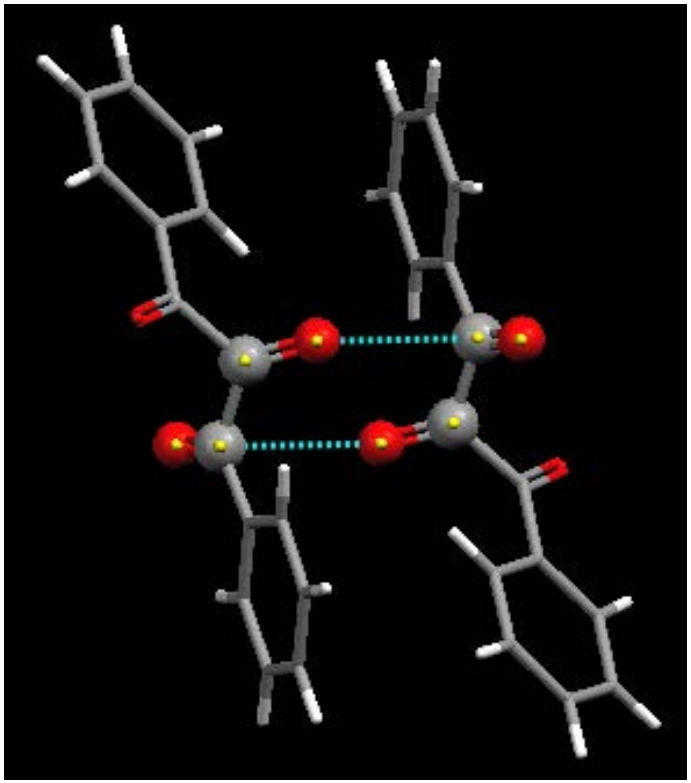
Other interactions not mediated by hydrogen

Interactions not mediated by hydrogen
C=O•••C=O Interactions in small molecules

LPROGL:
Cyclo(L-prolyl-L-glycyl)

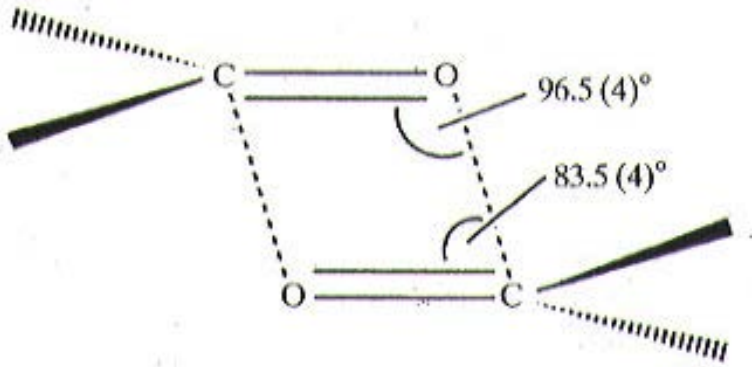


BIGXAG: 1,3-diphenyl-propane-1,2,3-trione

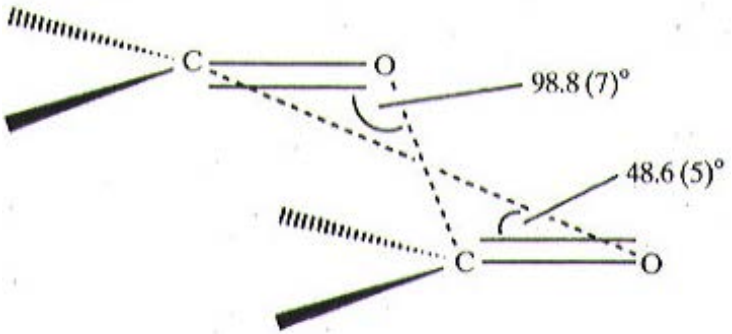


Interactions not mediated by hydrogen
Carbonyl•••Carbonyl interactions and energies

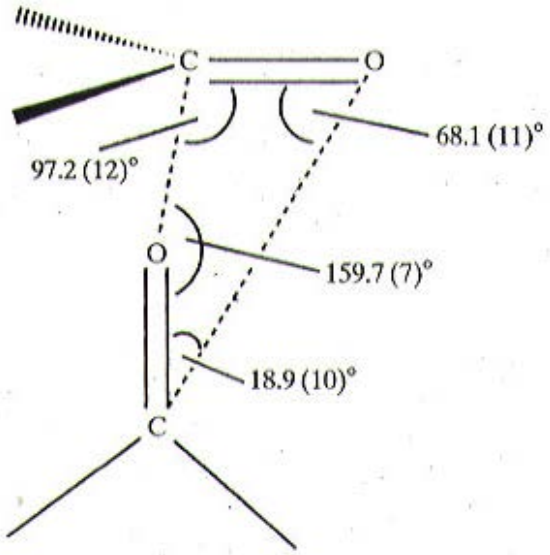
Acta Cryst., B54, 320, 1998



Antiparallel (-22.4 kJ mol^{-1})



Sheared parallel



Perpendicular and variants

(-7.6 kJ.mol^{-1})

C=O•••C=O Interactions in protein-ligand complexes

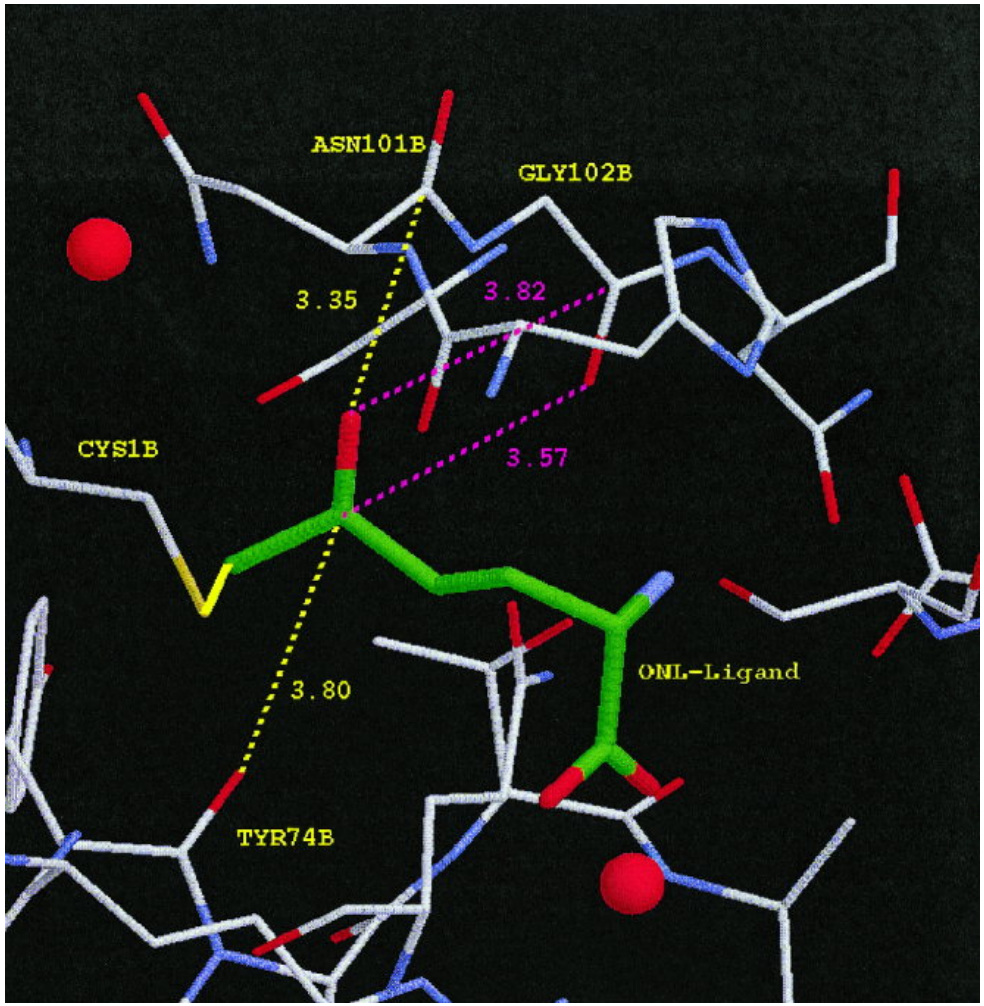
[Bergner et al., Biopolymers, 61, 99-110,2002]

Relibase Search

42 hits with O…C <3.8Å
13 ↑↓ , 29 ┤ & ‖

Backbone and Asn side-chain carbonyls often involved

1ecc: E-coli amido transferase complex of 5-oxo-norleucine



C=O•••C=O Interactions in proteins

- *Maccallum et al., J.Mol. Biol., 248, 374 & 361, 1995*

  CO•••CO attractions have a substantial influence on β-strand, α-helix & β-sheet conformations

- *Deane et al., Protein Eng., 12, 1025-1028, 1999*

  CO•••CO interactions stabilise partially allowed Ramachandran conformations of Asp and Asn

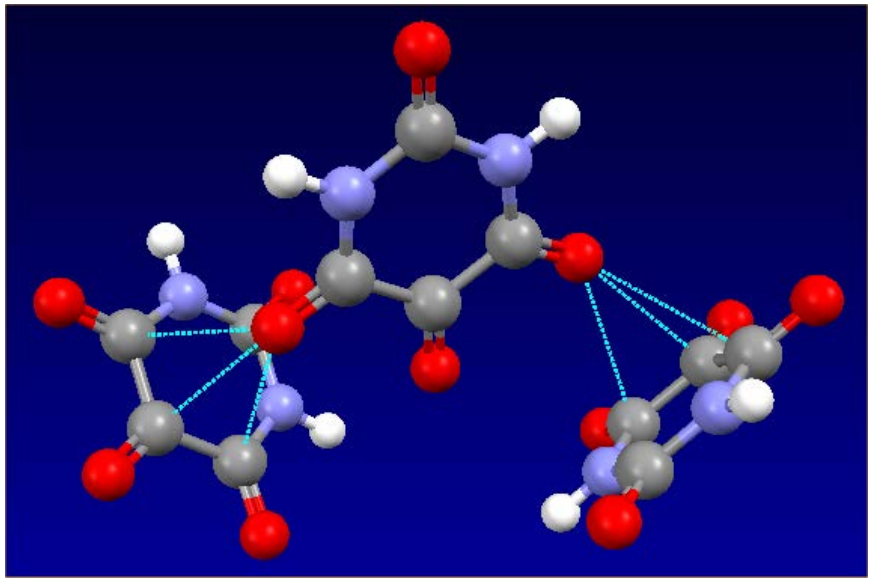
- *Hinderaker & Raines, Protein Science 12, 1118, 2002*

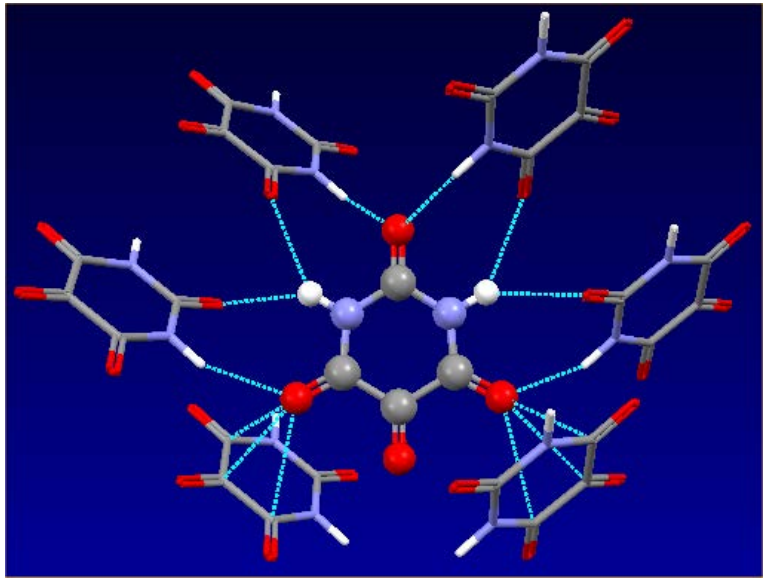
  n-π* Interactions between carbonyl groups (DFT)

[In small molecules, CΞN••• CΞN have similar motifs and energies *Acta Cryst., B**64**, 393-396, 2008]*

Alloxan has C=O•••C=O, but does it have N-H•••O?

[Dunitz & Schweizer, CrystEngComm., 9, 266-269, 2007]





C=O•••C=O 2.79, 2.96, 3.01Å

N•••O 3.21, 3.28Å

N•••H 2.36, 2.44Å

But, attractive energy of N-H •••O ≈ energy of C=O•••C=O!

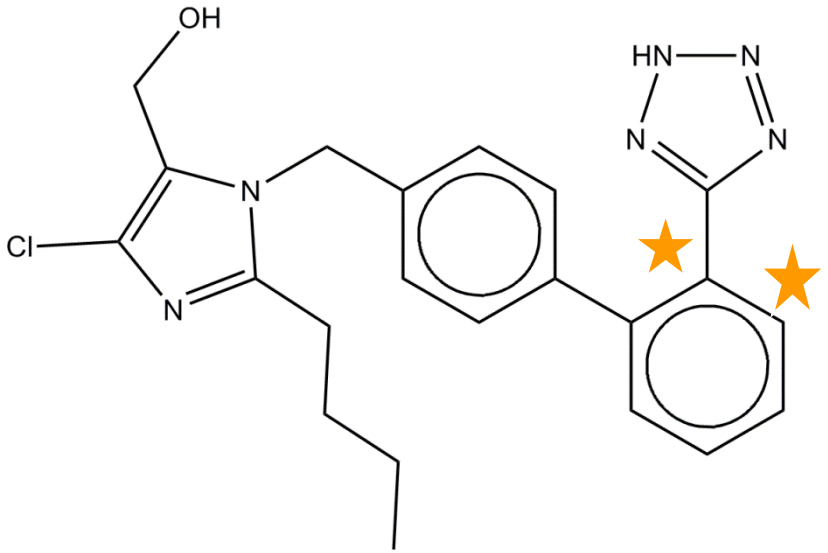
Using crystallography and theory to understand isosterism

Carboxylic acid – tetrazole isosterism

[Allen et al., JCIM, 52, 857-866, 2012]

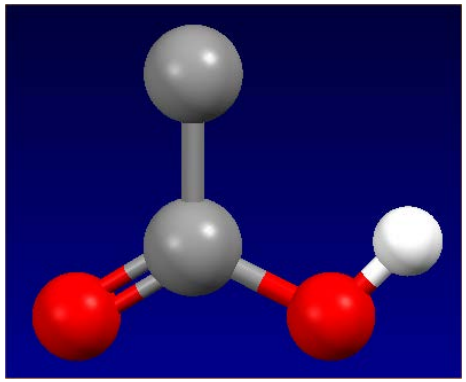
Losartan

- Imidazole-based antihypertensive
- Originally designed with –COOH at C_2 or C_3 (★)
- Effective by injection, but has limited oral bioavailability
- C_2-COOH replaced by 1*H*-tetrazole
- 10-fold increase in oral bioavailibility
- Other similar examples



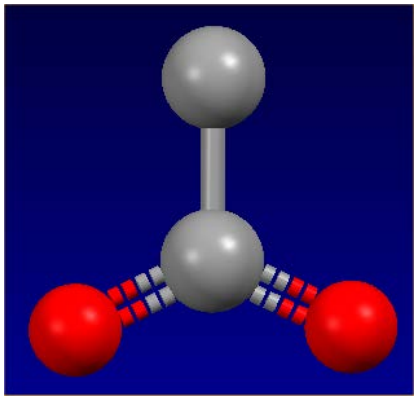
Hydrogen bond propensities (%,CSD) and energies (kJ.mol^{-1}, IMPT)

70%
(-24)



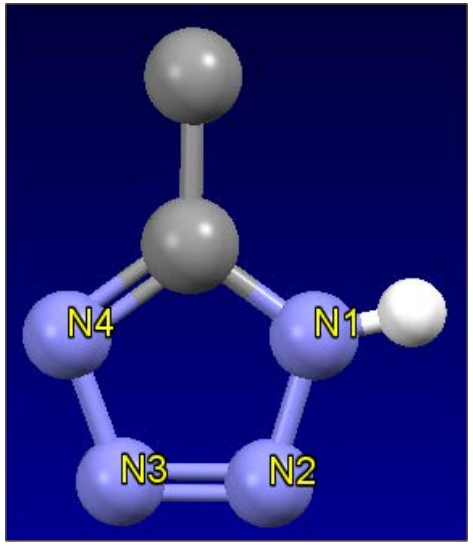
91%
(-34)

93%
(-78)



63%
(-27)

41%
(-19)

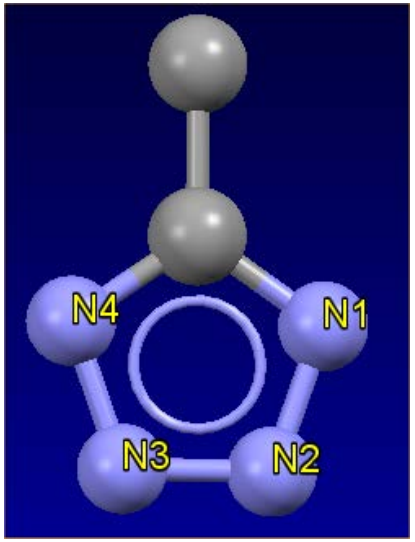


100%
(-44)

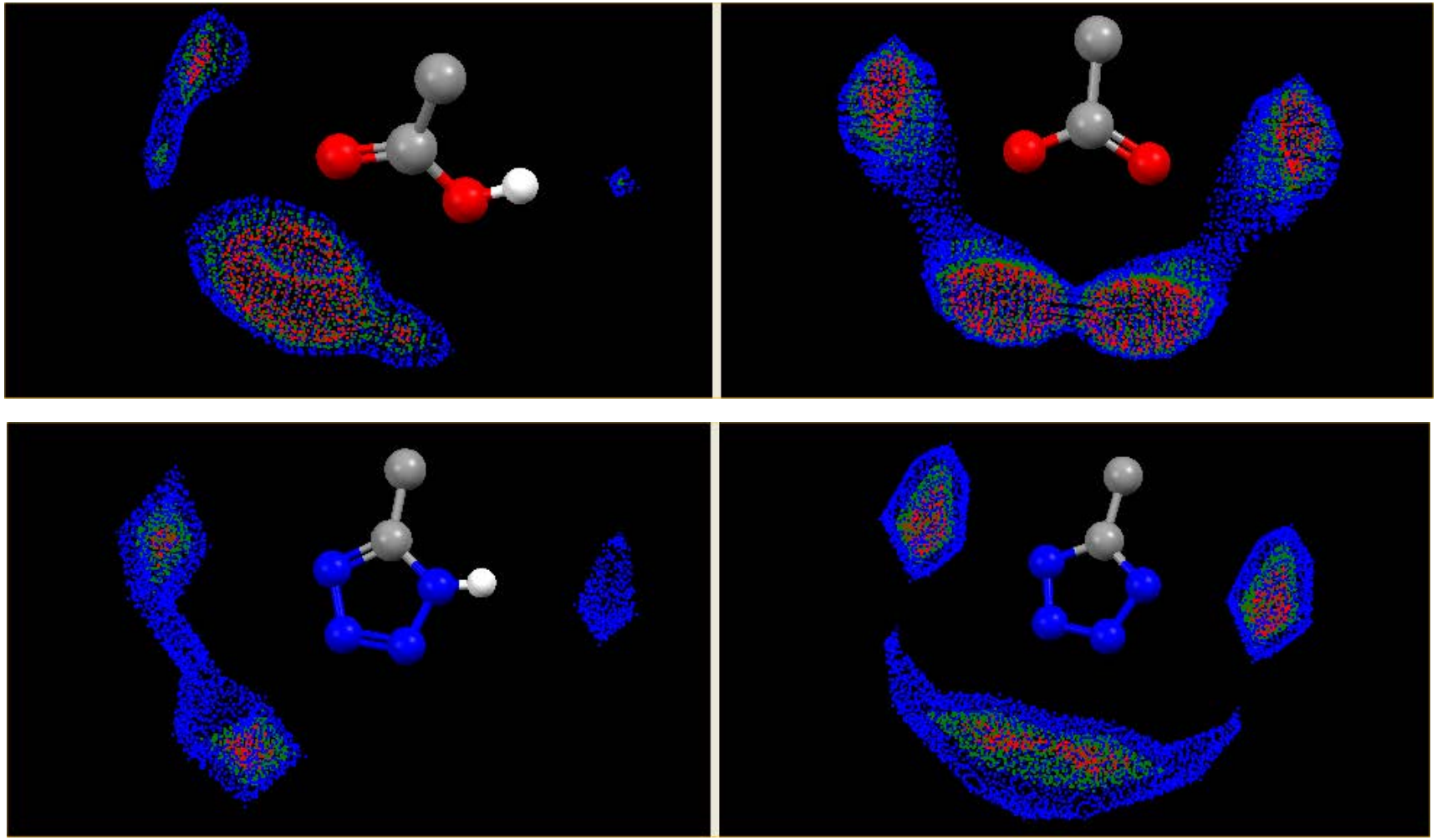
15%
(-17)

98%
(-62)

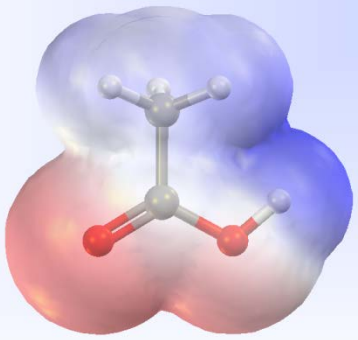
94%
(-59)



IsoStar plots for N,O-H donors for carboxyl, carboxylate, tetrazole, tetrazolate

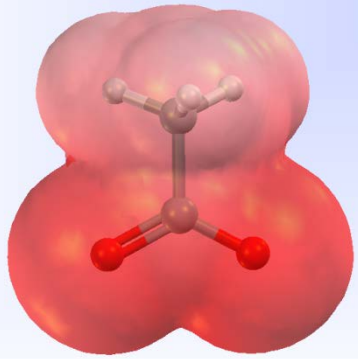


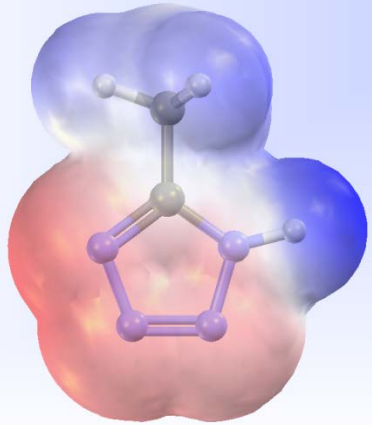
Electrostatic potential surfaces (MOPAC/AM1)



carboxyl

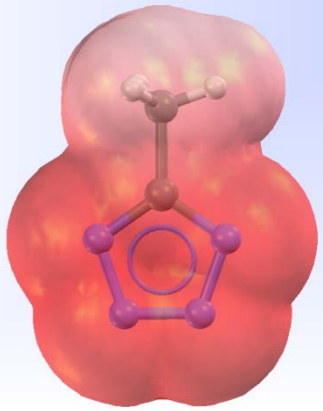
carboxylate





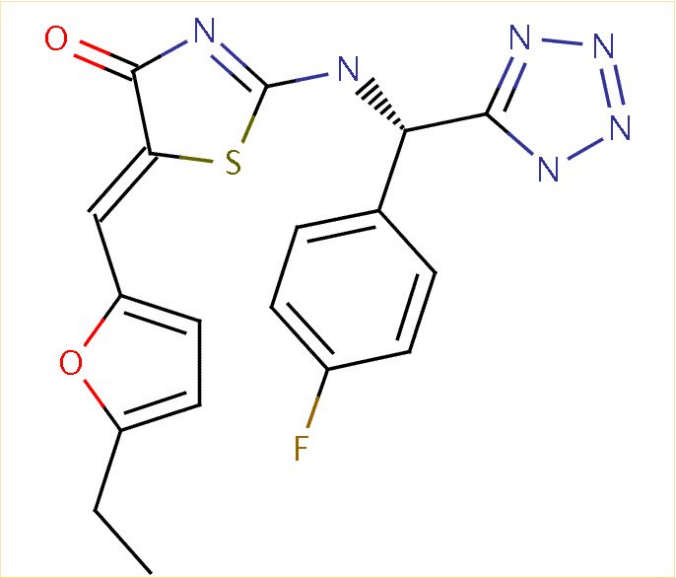
1H-tetrazole

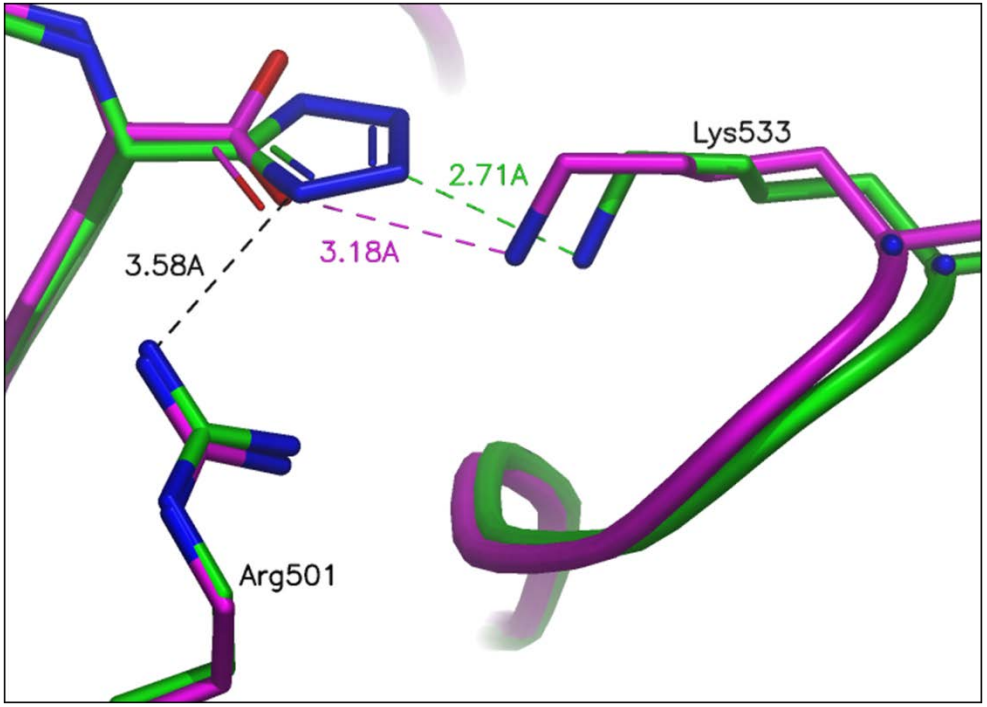
tetrazolate



Comparative binding of tetrazolate and carboxylate ligands to same protein

- Three -COO/tetrazolate ligand pairs each bound to same protein in PDB. Ligands VRX (-COOH) and **VXR** (tetrazole, below) both bind to Hepatitis C NS5B polymerase.
- Protein flexes to accommodate larger tetrazole ring in **VXR**
- **VXR** has stronger binding





Experiment and Theory in the study of Intermolecular Interactions

Conclusions

- Crystal structures tell us when and where molecules ‘stick’ together
- Computational chemistry can tell us how and why these ‘adhesions’ occur
- It is good for crystallographers to understand the origins of their observations
- It is good for modellers to have experimental confirmation of their computational explanations
- Everyone benefits, including drug designers and the developers of solid-form deliverables

Acknowledgements

All CCDC Staff, past and present, particularly:

Colin Groom, Elna Pidcock, Pete Wood, Peter Galek,

John Liebeschuetz, Tjelvar Olsson , Aurora Cruz-Cabeza

Robin Taylor, Sam Motherwell

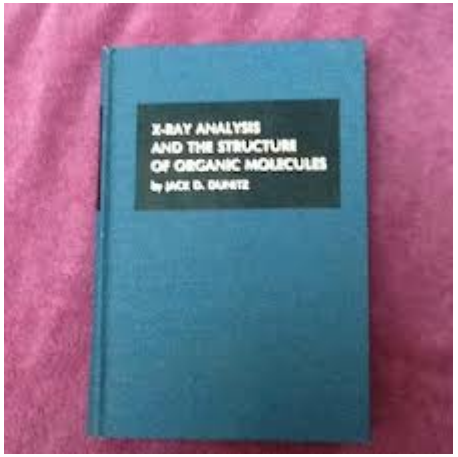
Help and advice with IMPT:

Anthony Stone (Cambridge), Sally Price (UCL)

The Cambridge Structural Database:

All those who determine and publish crystal structures and the CCDC Database Team who maintain the CSD

A Special Acknowledgement



Professor Jack D Dunitz

[Born 29 March 1923, Glasgow]

Lab. für Organische Chemie

ETH Zurich HCI H333

Switzerland

CCDC Governor & Trustee, 1986-1999

Annual Visitor 1999 – present (next visit in May 2013)

You have 8 days to send a 90th birthday card!