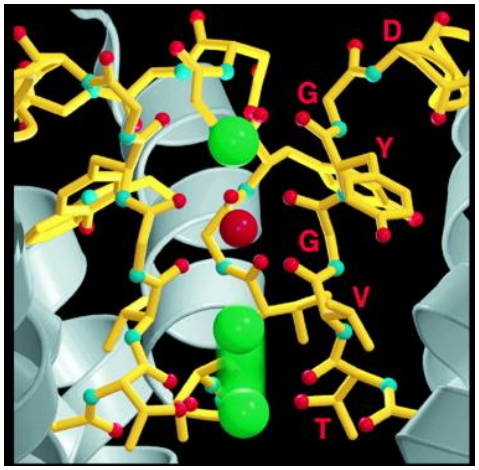
Large scale analysis of electronic effects in protein structure

CSGF Program Review

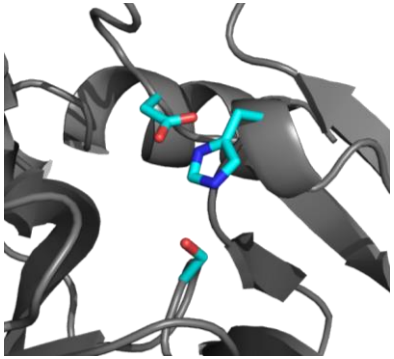
July 15, 2019

Helena Qi

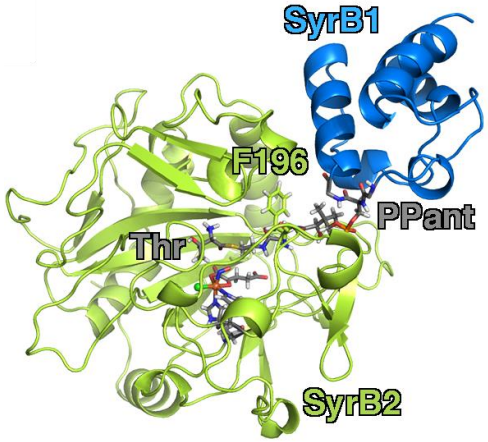
Why do we care about proteins?



Potassium channel[1]



Chymotrypsin[2]

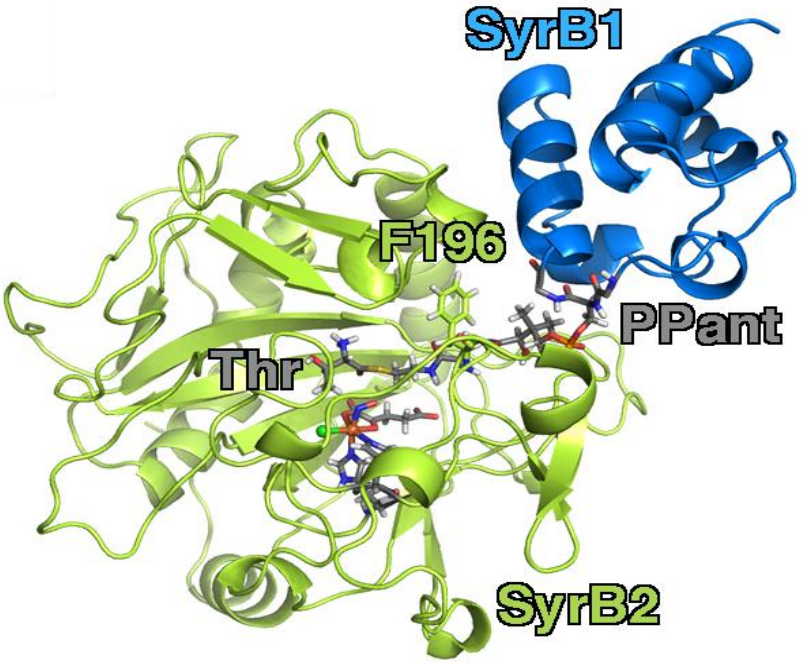


Aliphatic Halogenase SyrB2[3]

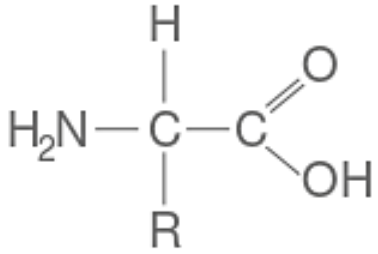
- Structure guides function, but is not the whole story

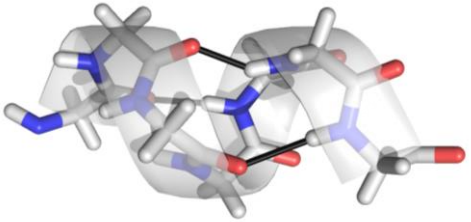
[1]Doyle, D.A.; *et al.* *Science* **1998**, *280(5360)*, 69-77
[2]1AB9
[3]Mehmood, R.; *et al. ACS Catal.*, **2019**, 9, 4930–4943

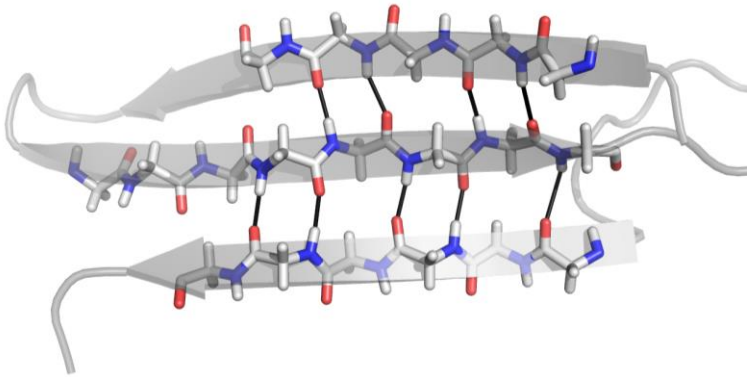
Protein structure



Aliphatic Halogenase SyrB2[3]



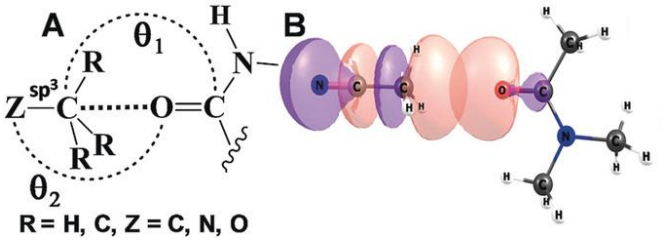




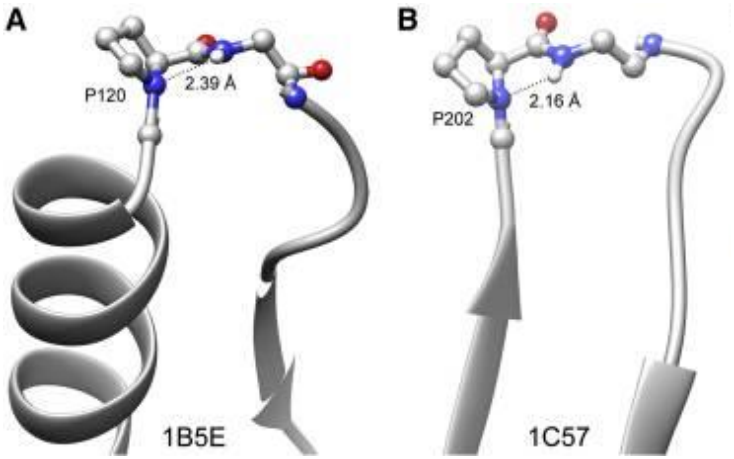
- X-ray crystallography allows us to examine protein structures

[3]Mehmood, R.; *et al. ACS Catal.*, **2019**, 9, 4930–4943

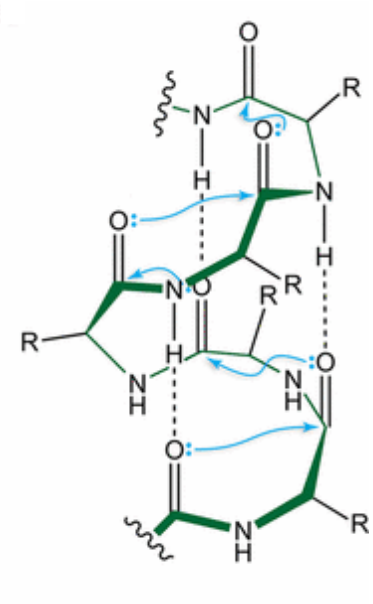
Mining the Protein DataBank for Unusual Chemistry



"Carbon-bonds"
n→σ* electron delocalization[1]



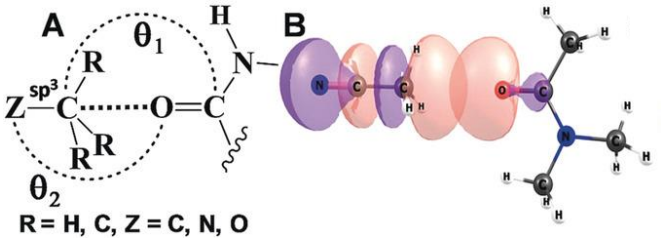
Proline N-H···N hydrogen bonds[3]



Backbone carbonyl
n→π* interactions[2]

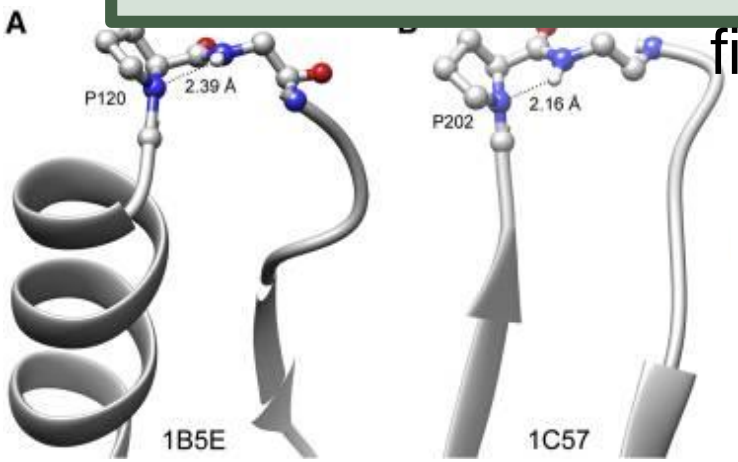
[1]Mundlapati, V. R. *et al.*, *Angew. Chem. Int. Ed.* **2018**, *57*, 16496-16500
[2]Newberry, R. W.; Raines, R. T., *Accounts of Chemical Research* **2017**, *50*, 1838-1846
[3]Deepak, R.N.V.K.; Sankararamakrishnan, R., *Biophys J.* **2016,** *110(9)*, 1967-1979

Mining the Protein DataBank for Unusual Chemistry

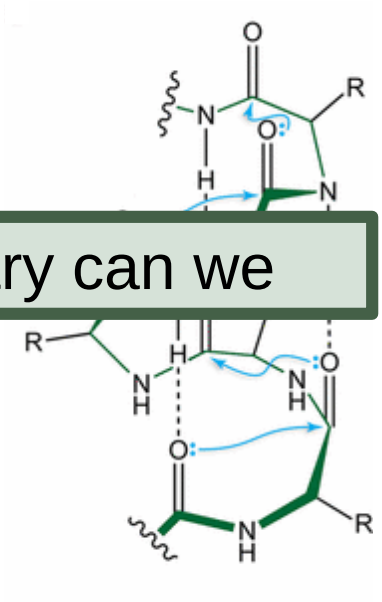


"Carbon-bonds"
n→σ*

What kinds of unusual chemistry can we find?



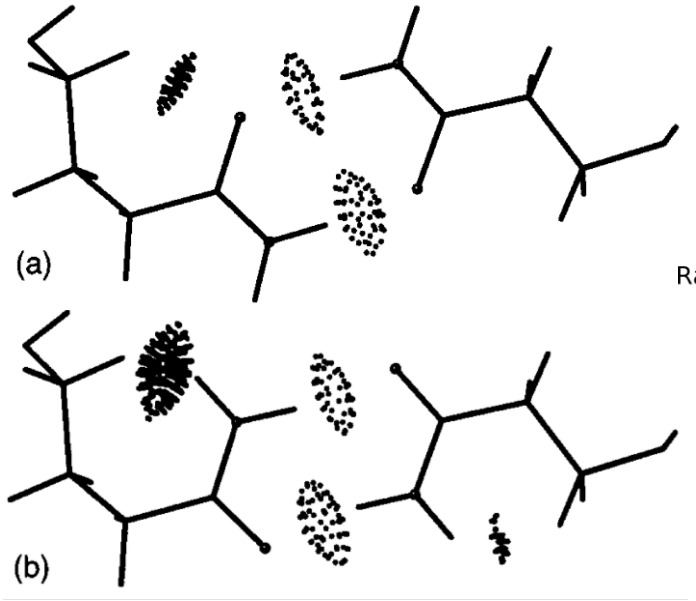
Proline N-H···N hydrogen bonds[3]

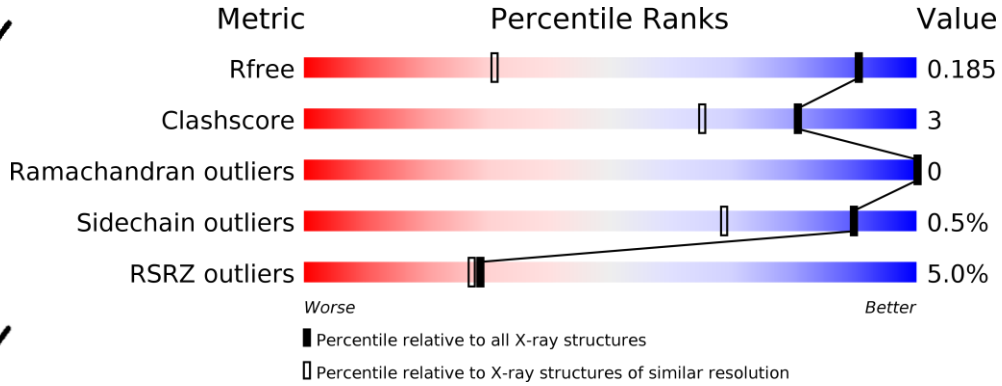


Backbone carbonyl n→π* interactions[2]

[1]Mundlapati, V. R. *et al.*, *Angew. Chem. Int. Ed.* **2018**, *57*, 16496-16500
[2]Newberry, R. W.; Raines, R. T., *Accounts of Chemical Research* **2017**, *50*, 1838-1846
[3]Deepak, R.N.V.K.; Sankararamakrishnan, R., *Biophys J.* **2016,** *110(9)*, 1967-1979

“Clashes” as a source of unexpected chemistry?

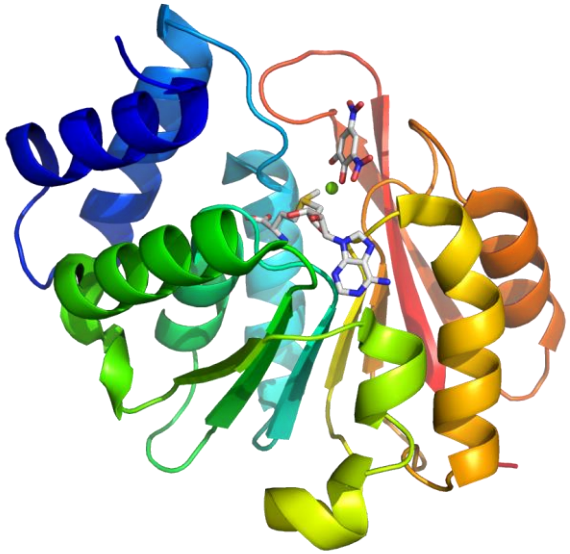




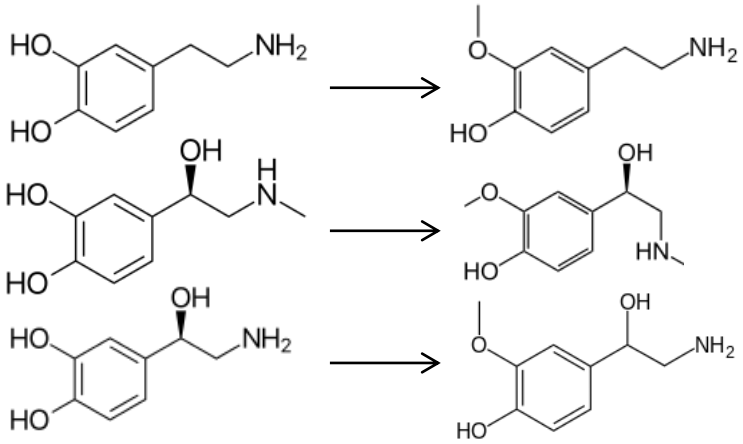
- Clashes are a powerful tool to diagnose local fitting problems
- Clashscore used to judge quality of structure

Word, J.M. *et al. J. Mol Bio.* **1999.** *285(4)*, 1735-1747

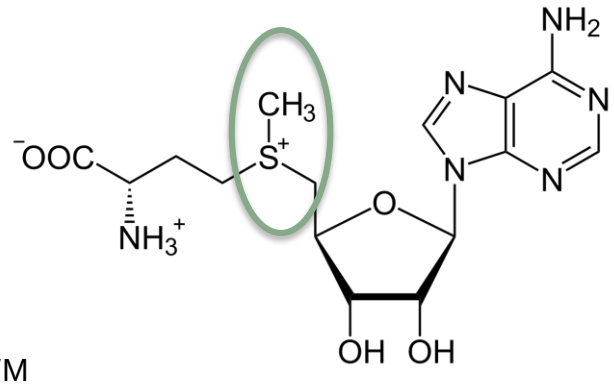
The case of catechol-*O*-methyltransferase

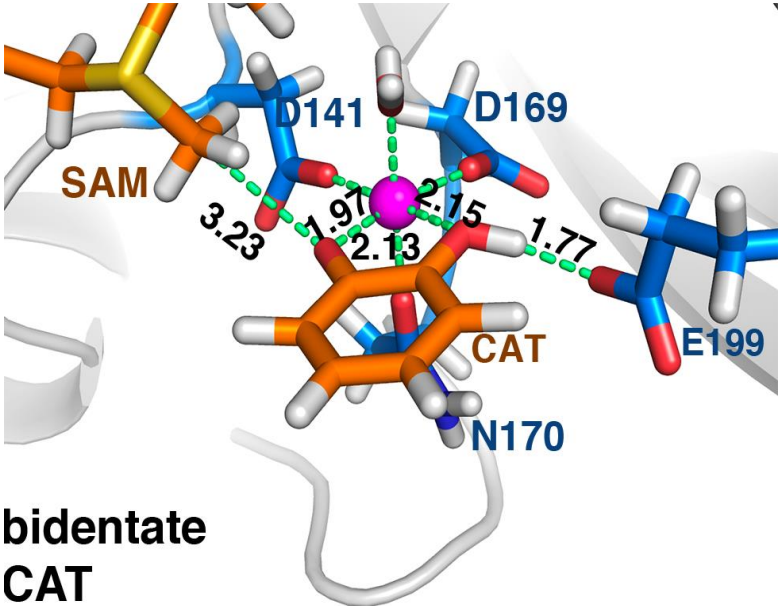


Inactivates catecholamine neurotransmitters



In the active site, Mg^{2+} positions catecholate so it can react with SAM

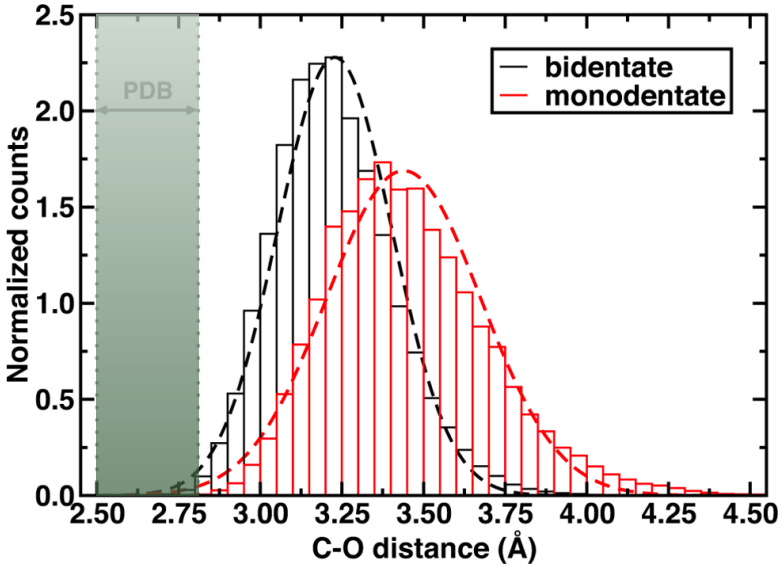


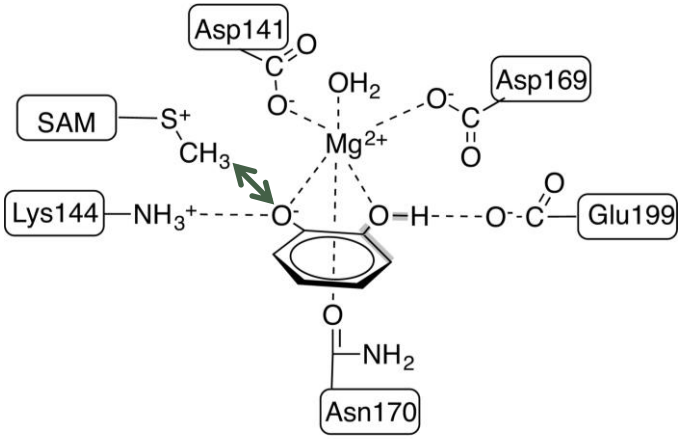


PDBID 3BWM

Patra, N.; Ioannidis, I.I.; Kulik, H.J., **2016**. PLOS ONE 11(8): e0161868

The case of catechol-*O*-methyltransferase



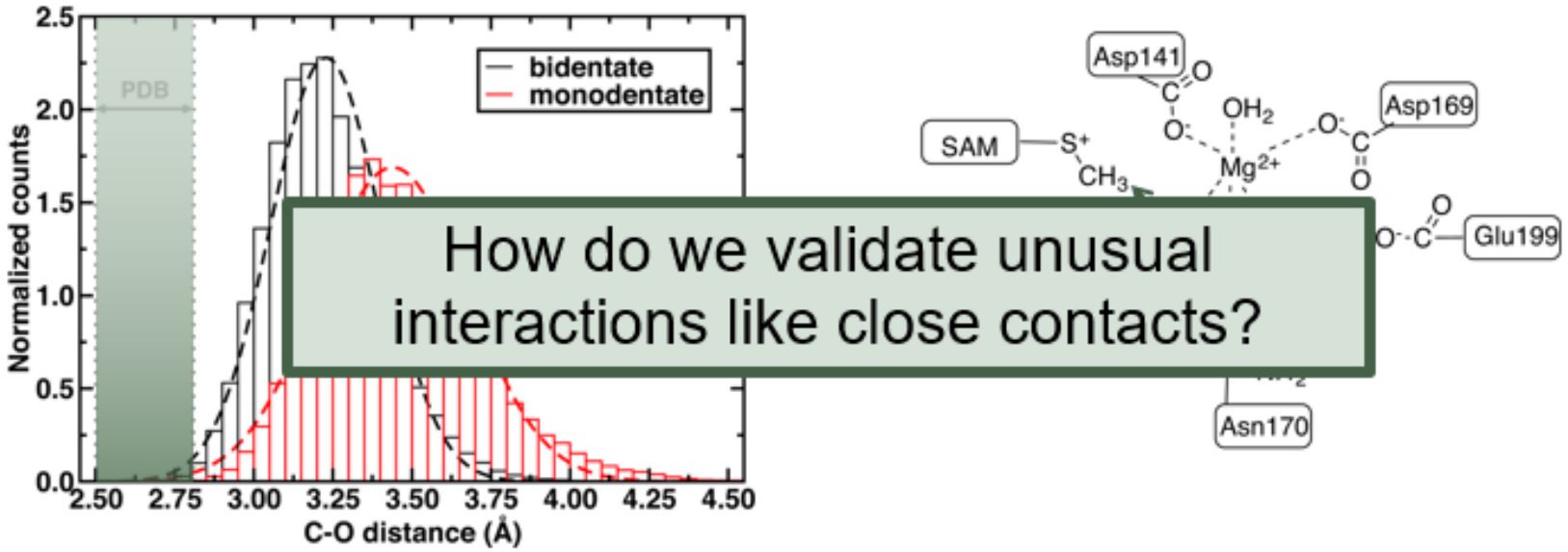


- Experimental crystal structure shows unusually short SAM-catechol distance

- Unable to be replicated with standard methods used to study proteins

Patra, N.; Ioannidis, I.I.; Kulik, H.J., **2016**. PLOS ONE 11(8): e0161868

The case of catechol-*O*-methyltransferase

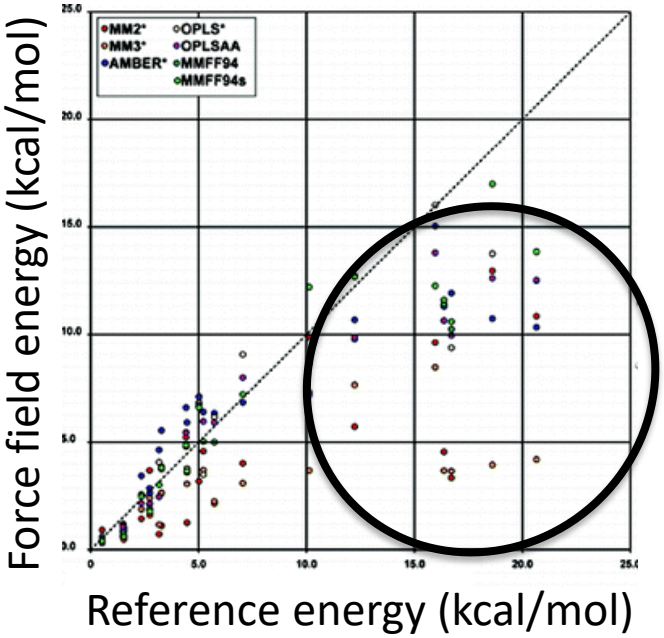


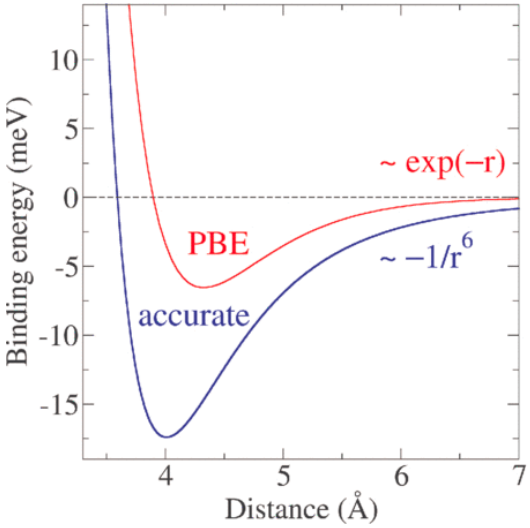
- Experimental crystal structure shows unusually short SAM-catechol distance

- Unable to be replicated with standard methods used to study proteins

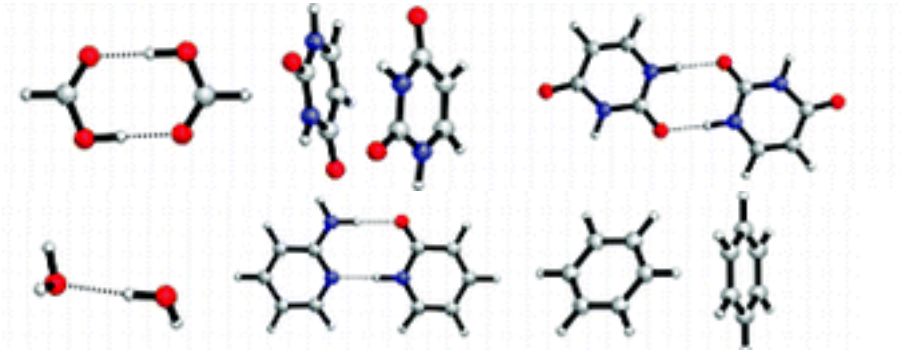
Patra, N.; Ioannidis, I.I.; Kulik, H.J., **2016**. PLOS ONE 11(8): e0161868

Non-covalent interactions are difficult to model



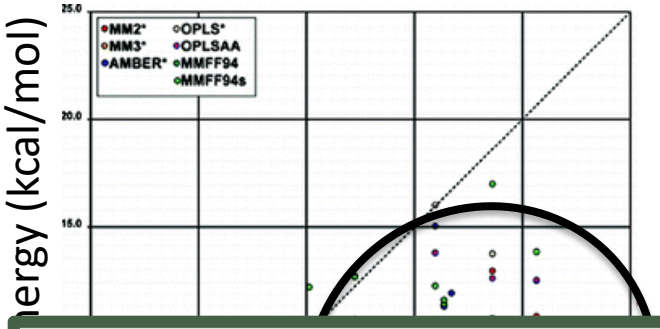


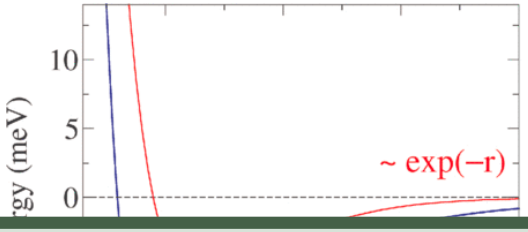
Kr dimer binding curves[2]



[1]Paton, R.S; Goodman, J.M., *J. Chem. Inf. Model.* **2009**, *49(4)*, 944-955
[2]Klimes, J.; Michaelides, A., *J. Chem. Phys.* **2012,** *137,* 120901

Non-covalent interactions are difficult to model

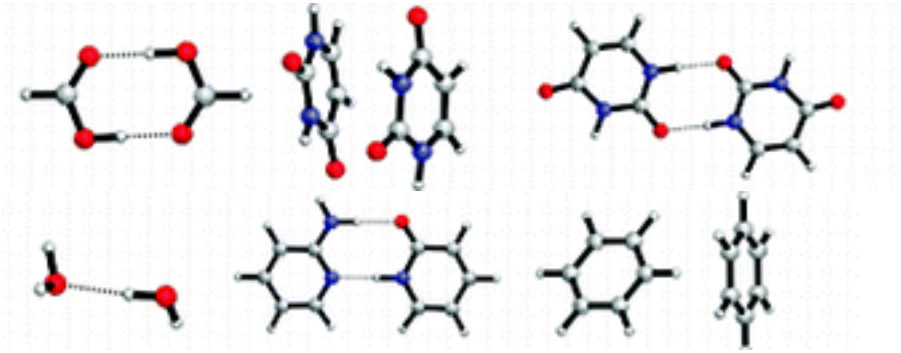




How well do standard methods capture unusual interactions?
When does it get it "right"? And when does it fail?

Reference energy (kcal/mol)

Distance (Å)



Kr dimer binding curves[2]

[1]Paton, R.S; Goodman, J.M., *J. Chem. Inf. Model.* **2009**, *49(4)*, 944-955
[2]Klimes, J.; Michaelides, A., *J. Chem. Phys.* **2012,** *137,* 120901

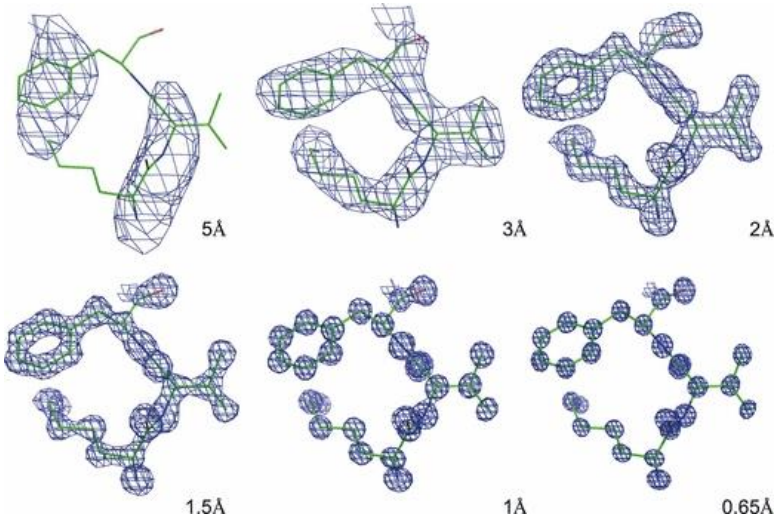
Overview

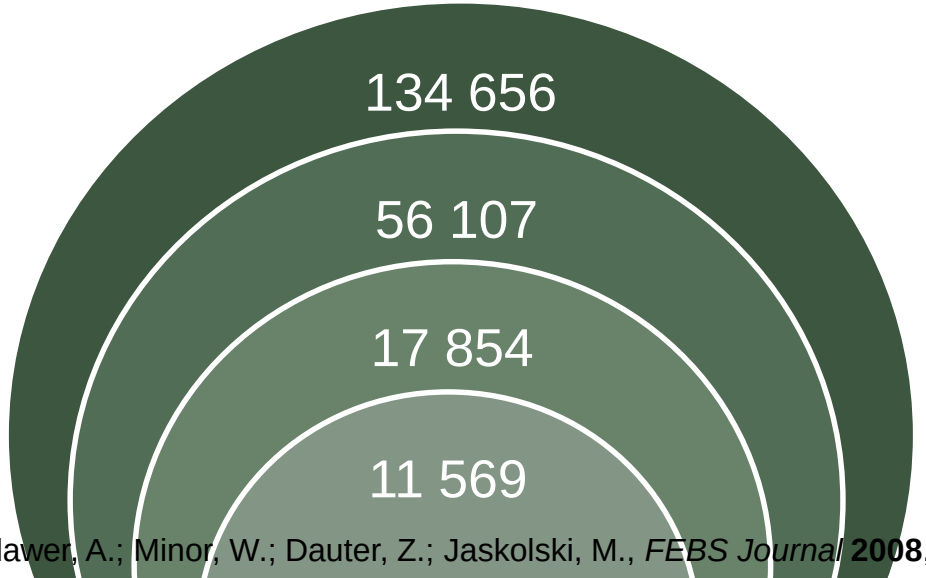
What are the statistics of close contacts in protein crystal structures?

What do minimal models say about the stability of the close contacts?

What is the role of the protein environment?

Curating the protein dataset

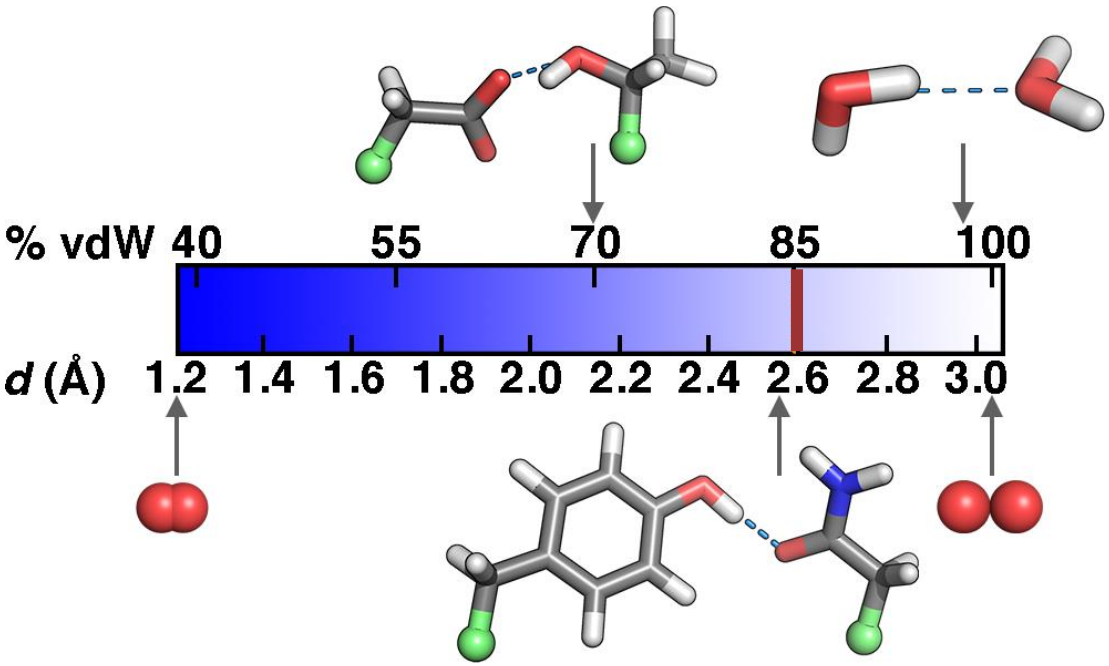




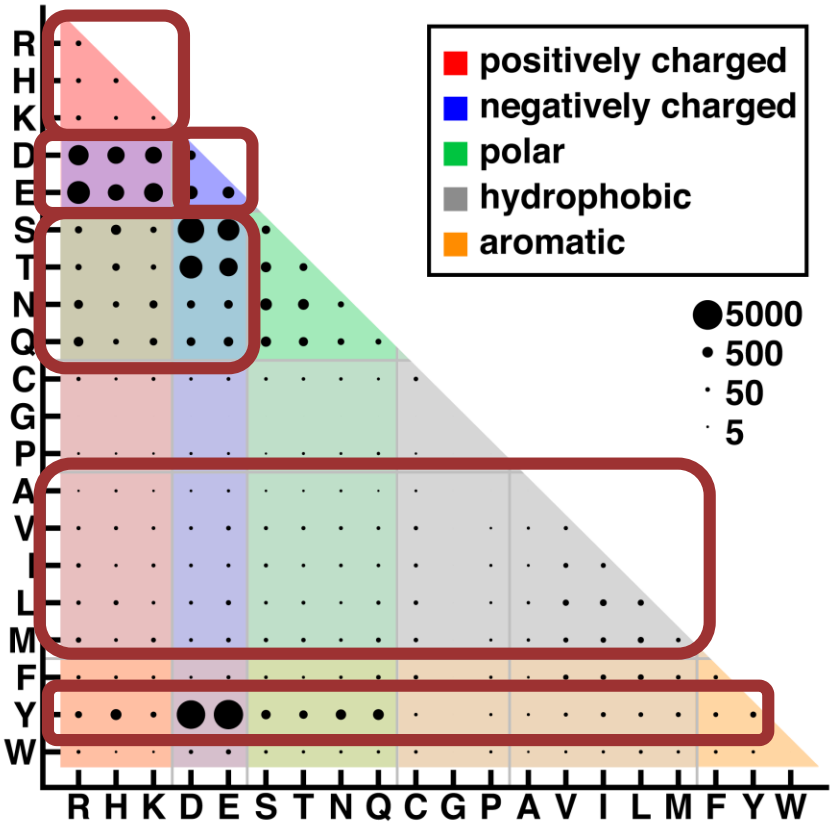
- The PDB is a repository of all solved protein structures
- X-ray crystal structures at high resolution
- Only non-redundant proteins
- Check for quality of structure
  - Data completeness
  - Assess quality of fit between experiment and model
  - Do not use metrics that may exclude unusual features

Wlodawer, A.; Minor, W.; Dauter, Z.; Jaskolski, M., *FEBS Journal* **2008**, *275*, 1-21.

Close contact curation



Which residues interact with each other?



Overview

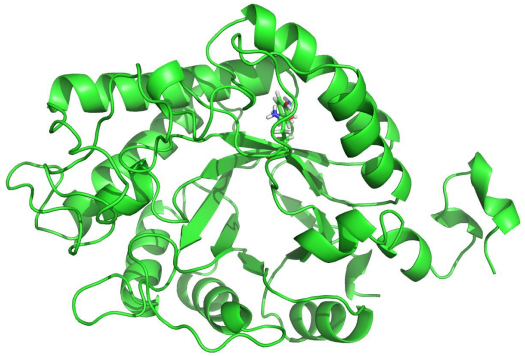
What are the statistics of close contacts in protein crystal structures?

What do minimal models say about the stability of the close contacts?

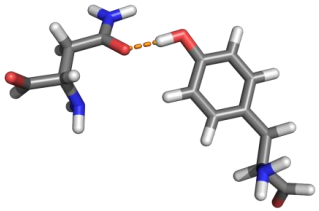
What is the role of the protein environment?

Minimal models

- Instead of modeling the whole protein...

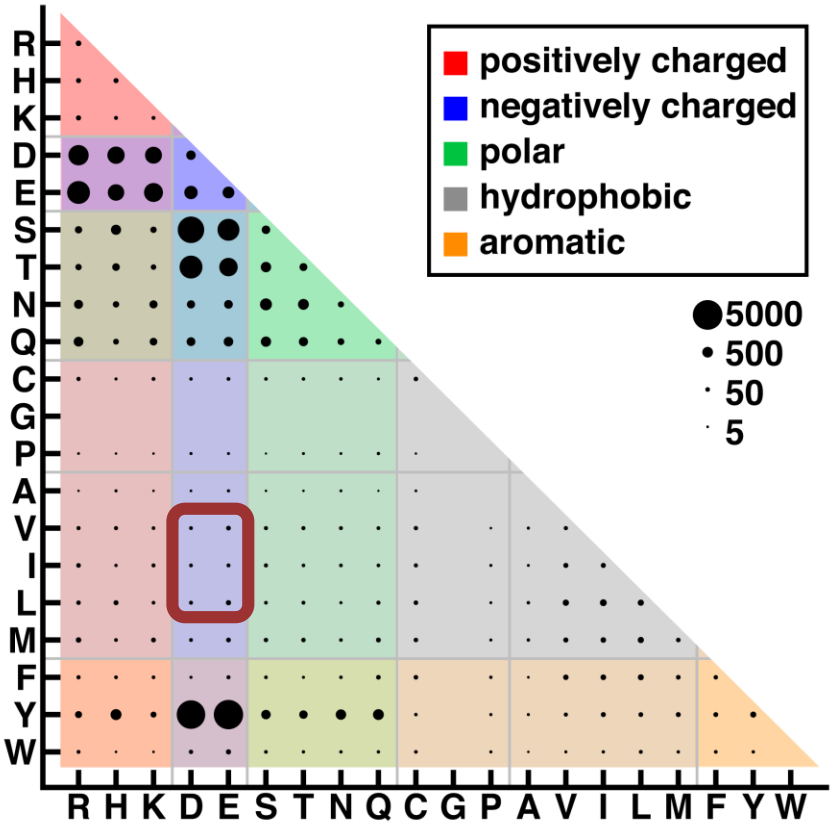


- First screen 5,289 close contacts with a minimal model

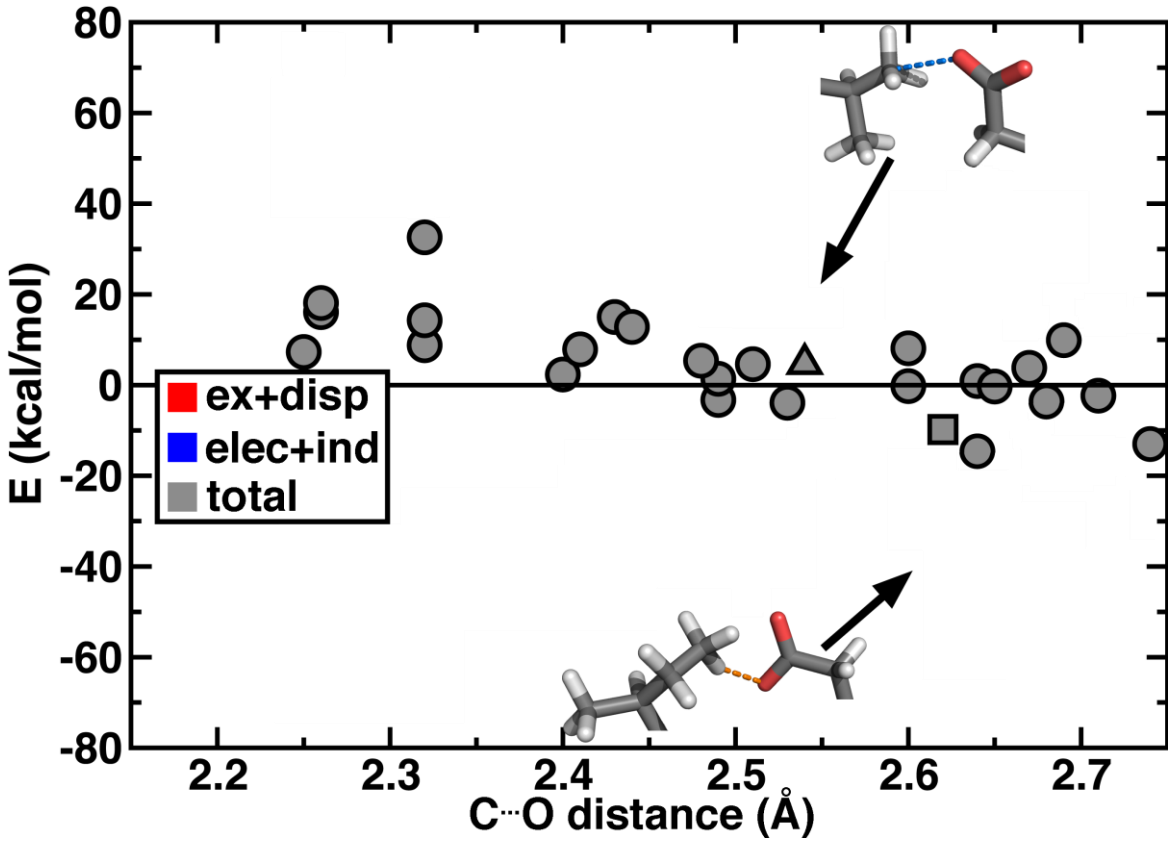


- Considerations: must add hydrogen atoms back in
- Compare results from classical and quantum mechanical methods

Which residues interact with each other?

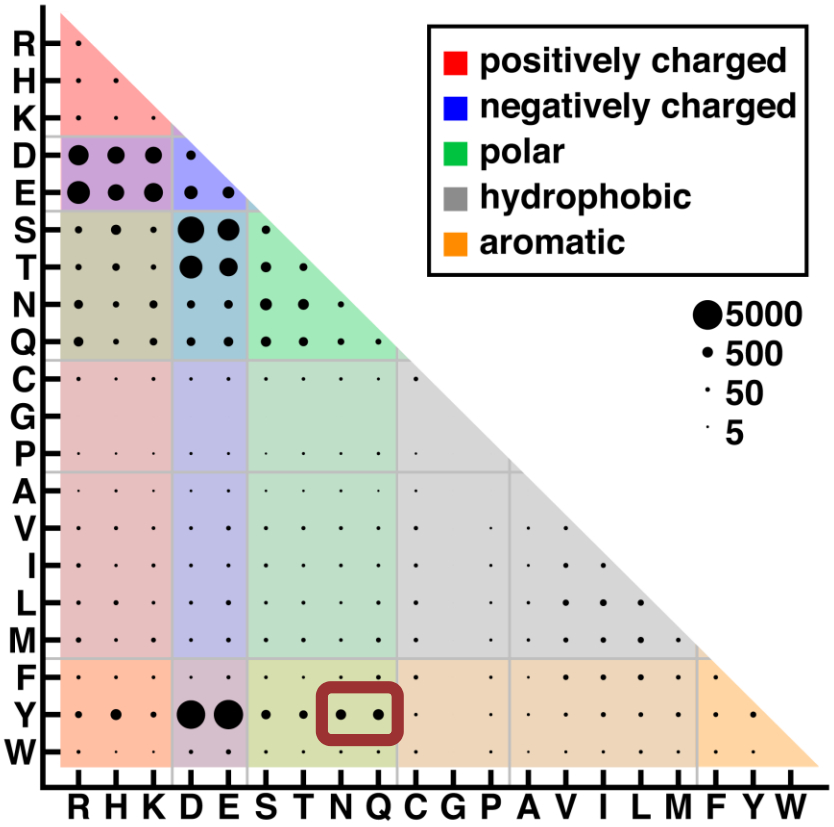


Hydrophobic (Val, Ile, Leu) – negative charge (Asp, Glu) interactions

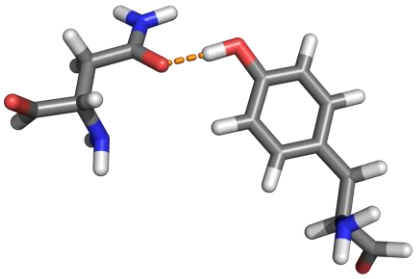


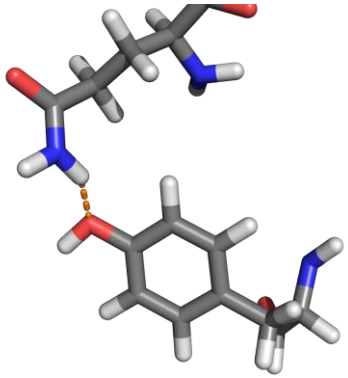
- 29 of 100 cases modeled
- C-H···O interactions are observed in the favorable cases, with angles >130°
- Unfavorable interactions are either:
  - Small angles
  - Small distance

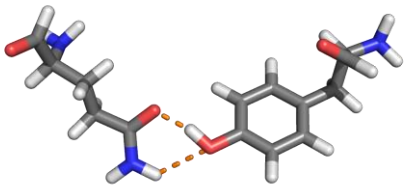
Which residues interact with each other?



How does Tyr interact with Asn/Gln?

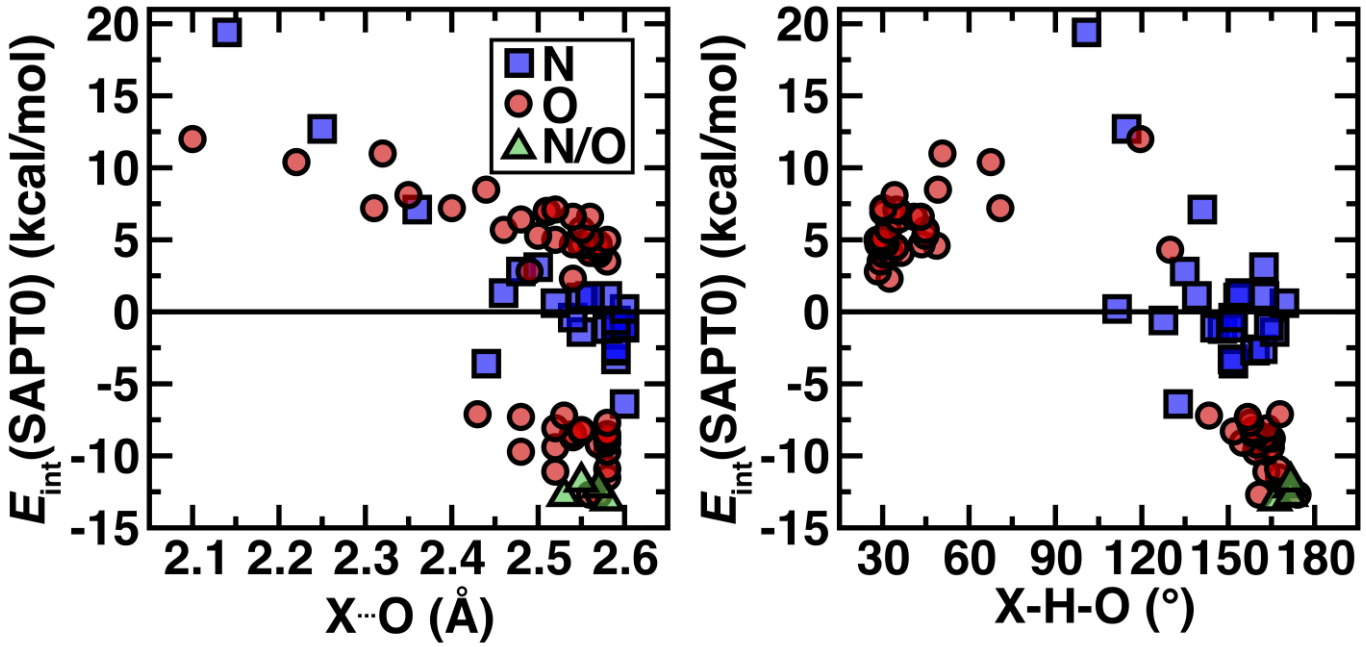






- 452 of 619
- 167 of 619

Modeled 73 interactions



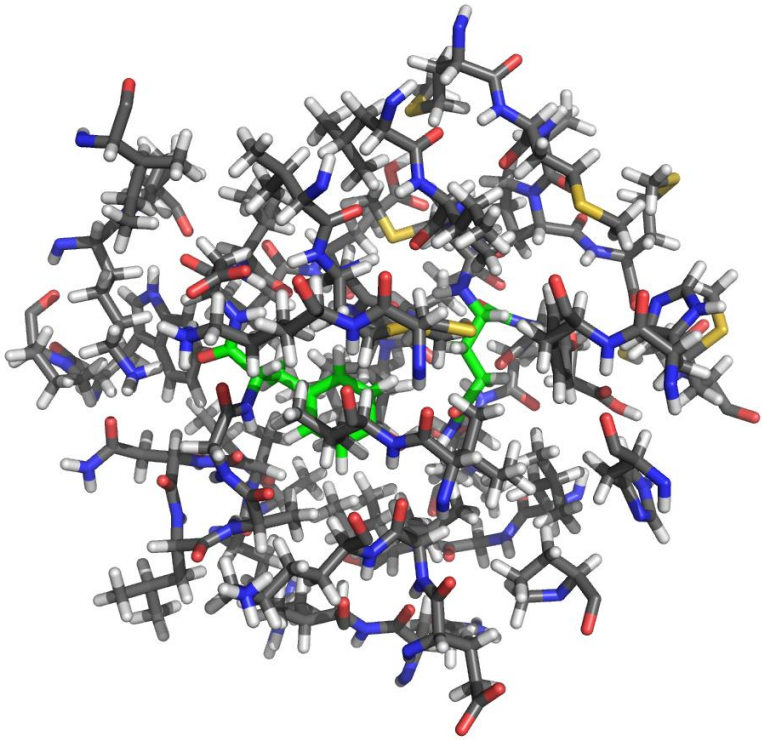
Overview

What are the statistics of close contacts in protein crystal structures?

What do minimal models say about the stability of the close contacts?

What is the role of the protein environment?

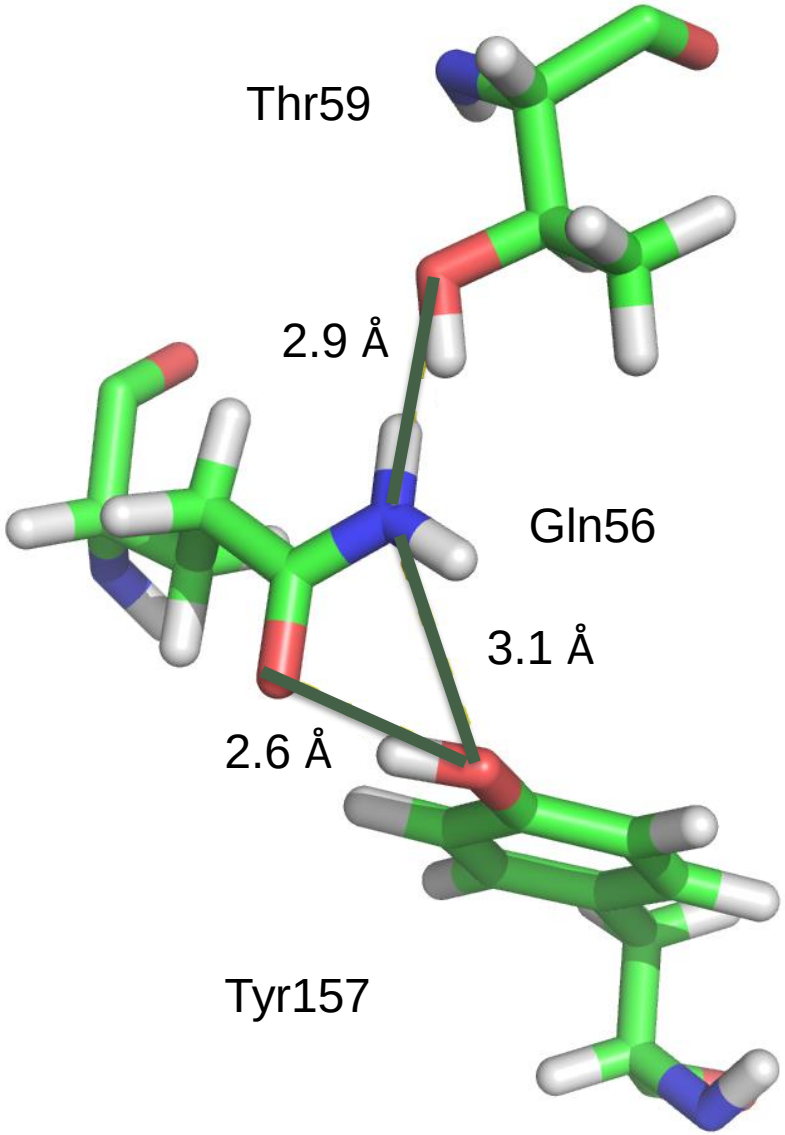
Cluster preparation



- Close contact
- Select residues within 7 Å of close contact
  - 1000s of atoms! Only possible with GPU acceleration
- Freeze all heavy atoms except close contact

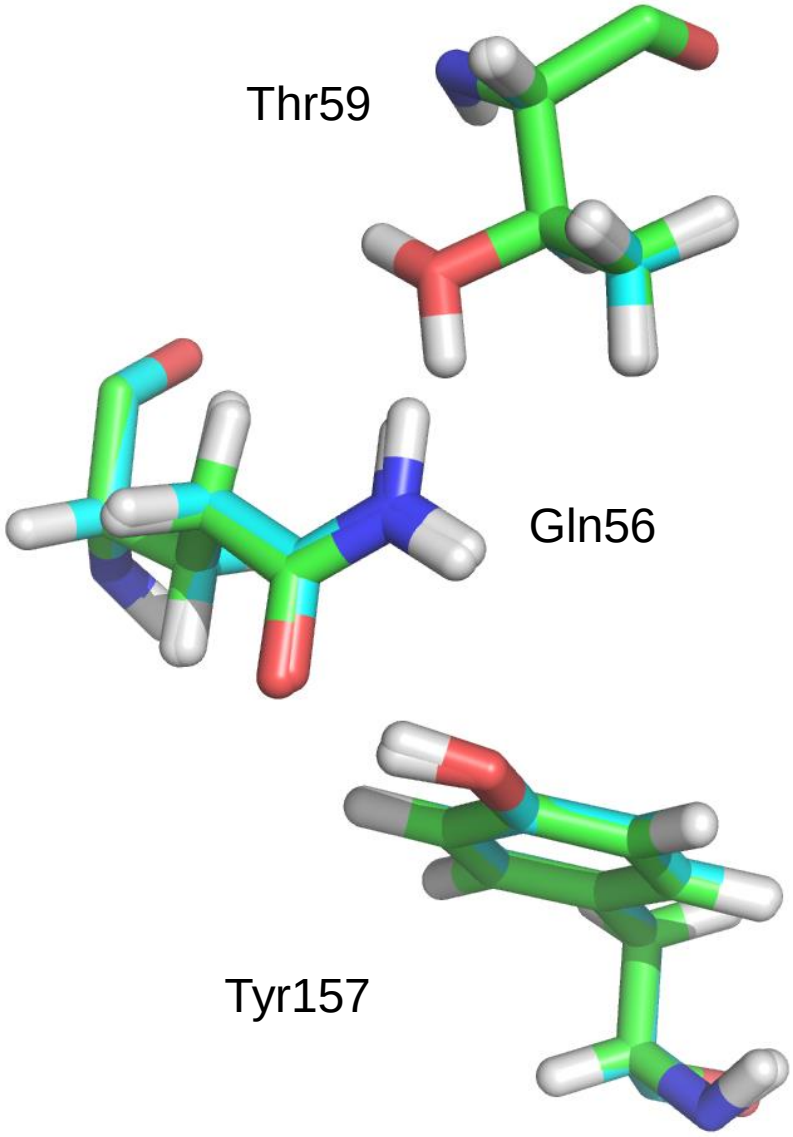
Optimization results

- Initial structure



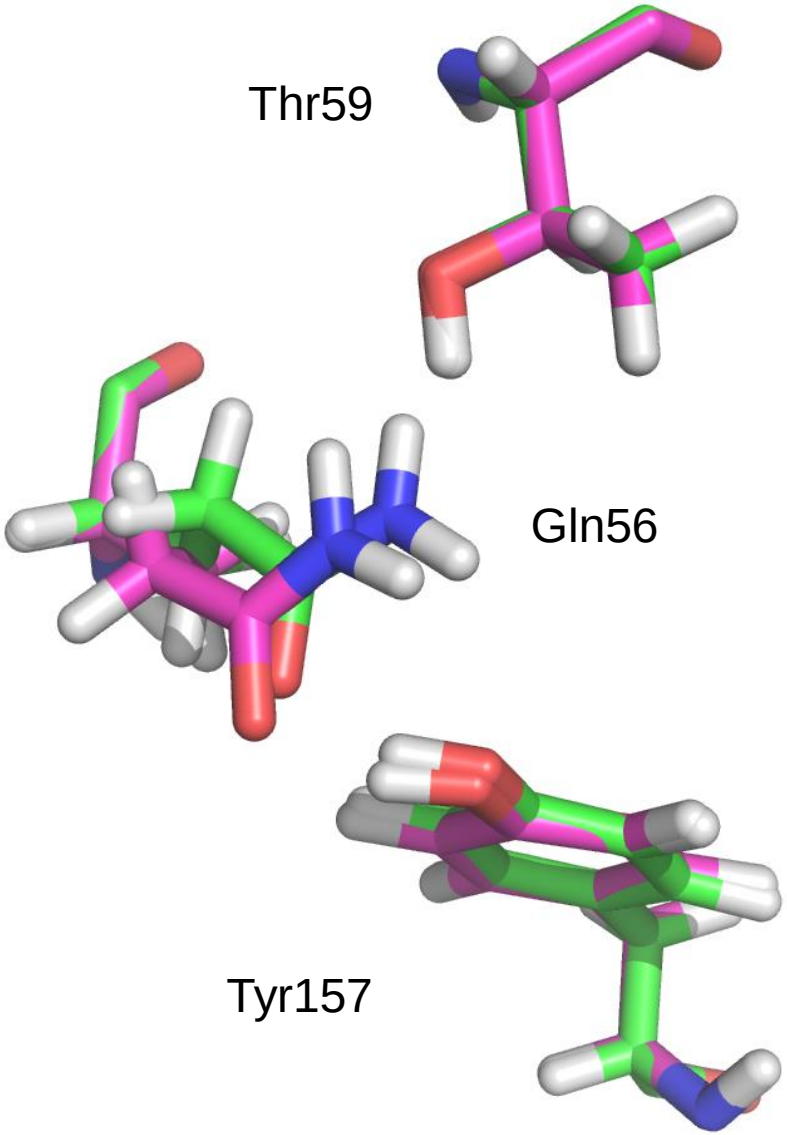
Optimization results

- Initial structure
- QM optimization

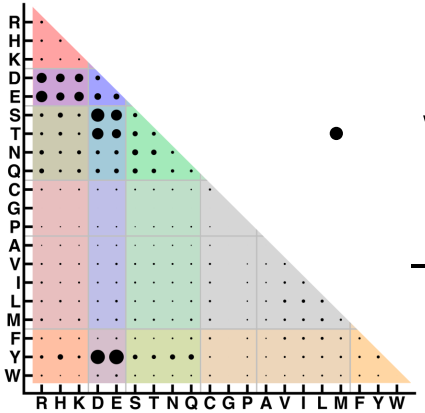


Optimization results

- Initial structure
- QM optimization
- MM optimization

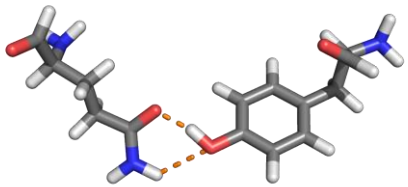


Conclusions



- We searched for short distances that could be considered mistakes
  - Close contacts are ubiquitous in the PDB, even in well resolved structures

- Some overrepresented residues (e.g., Tyr) form interesting interactions that have not been previously observed
  - Future work: understanding the double hydrogen bond motif



- Large-scale electronic structure validates close contacts not favored with classical simulation

Acknowledgements

DOE CSGF
BURROUGHS WELLCOME FUND
XSEDE
Extreme Science and Engineering Discovery Environment
NIH
MIT CEHS