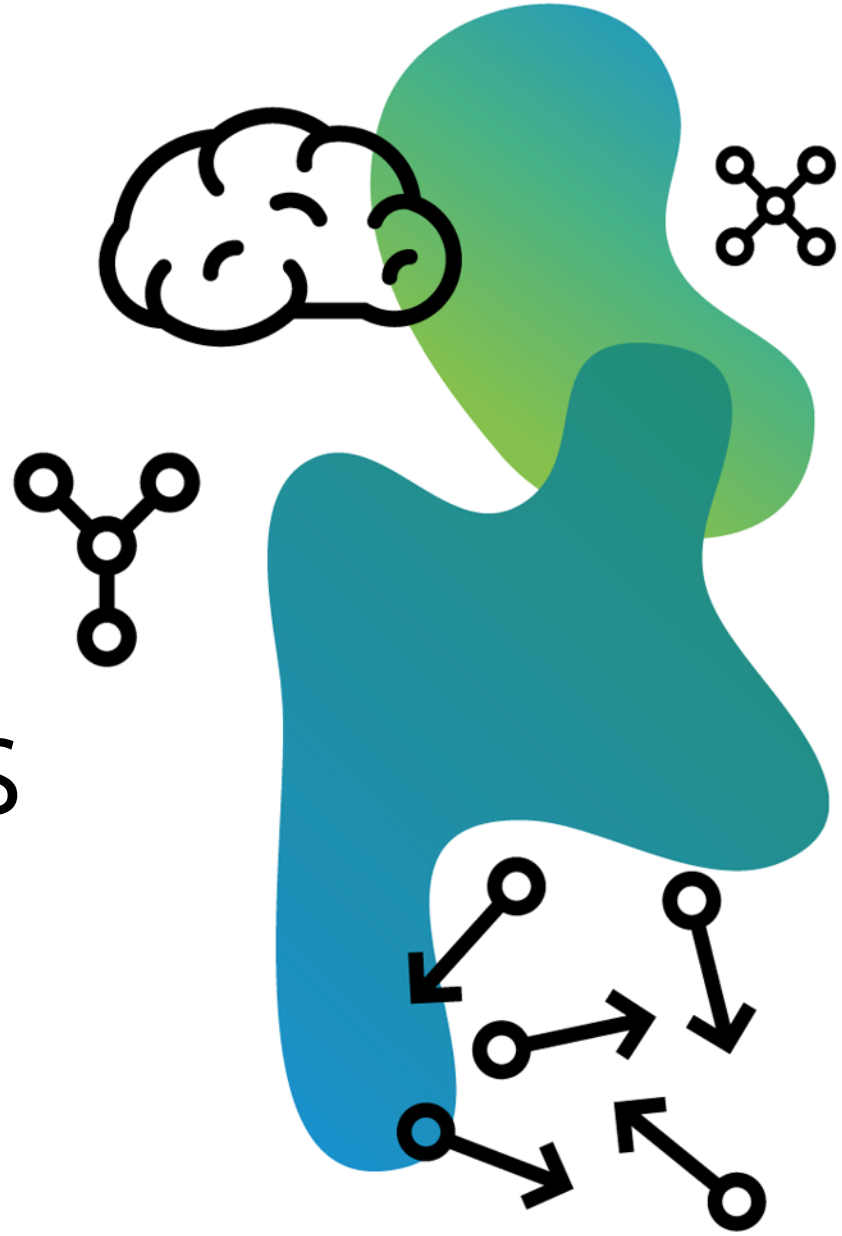
idorsia

Beyond software: CADD behind the scenes



SCS Spring School on Digital Chemistry –
Applied to Drug & Crop Protection Discovery

Les Diablerets

16.-20.04.2023

Outline

- Part 1: *(2:30 h)*
  - General introduction:
    - Most common Protein Ligand interactions
    - Small molecule conformational preferences
    - Electron densities
  - Introduction to the target Sars-Cov2 main protease MPro
  - Introduction to PyMol and coot
  - Exercise 1.1-5 and Discussions
- Part 2: *(1:30h)*
  - General introduction: Halogen bonds
  - Exercise 2.1
  - Halogen bond example from Idorsia, other interactions
  - Exercise 2.2

idorsia

Ligand-Protein Interactions

- General Concepts of Binding
- Intermolecular Ligand-Protein Interactions
  - Hydrophobic interaction
  - Hydrogen bonding
  - Salt Bridge interactions
  - Cation-Aryl interactions
  - Aryl-Aryl interactions
  - X-Bonds

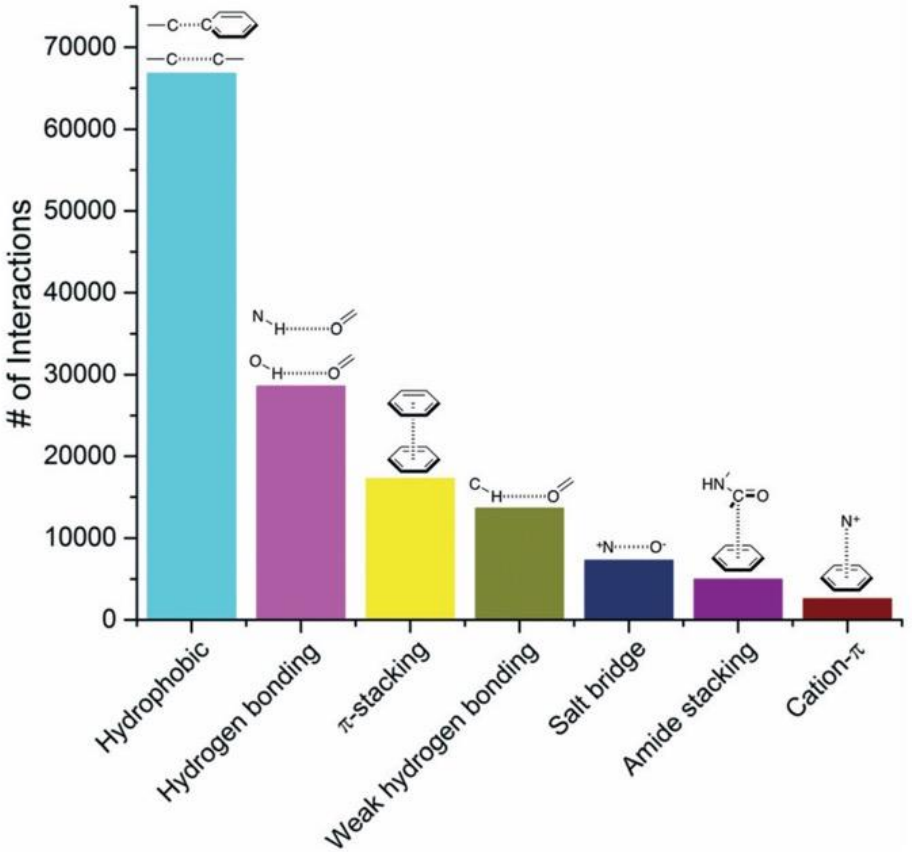
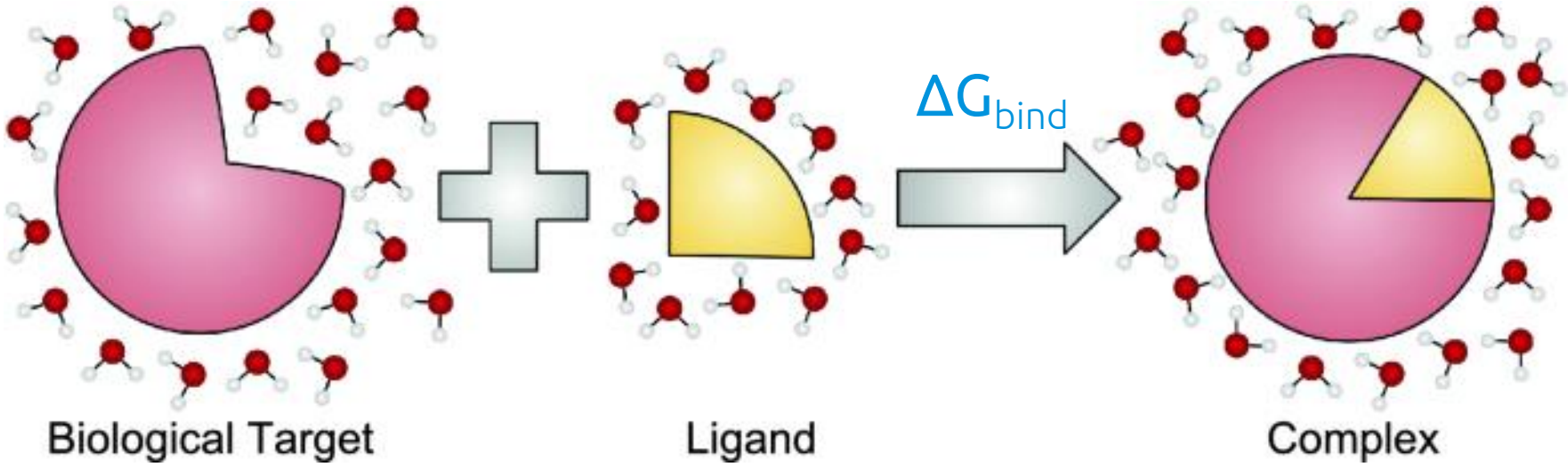


Fig. 1 Frequency distribution of the most common non-covalent interactions observed in protein–ligands extracted from the PDB.

Med. Chem. Commun., **2017**, 8, 1970

Gibbs Free Energy of binding: ΔG_{bind}

$$\Delta G_{bind} = \Delta H_{bind} - T\Delta S_{bind}$$

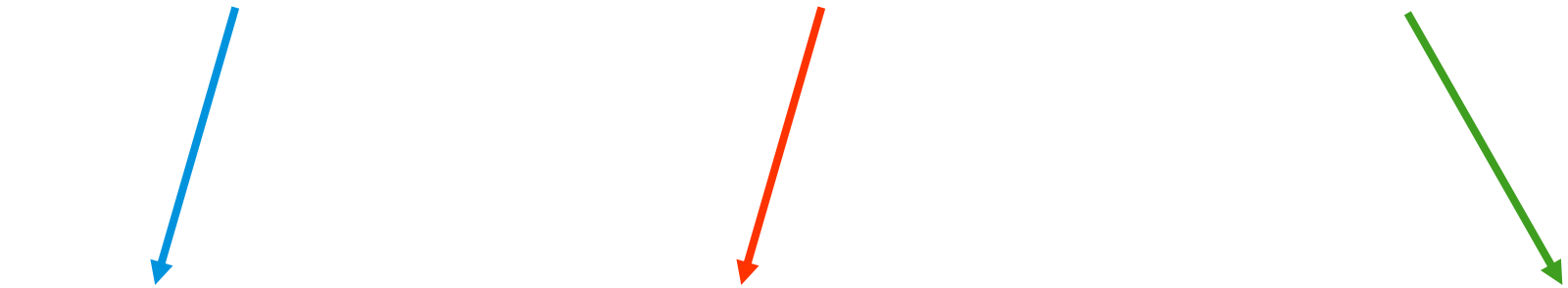


Med. Chem. Commun., **2017**, 8, 1970

idorsia

Gibbs Free Energy of binding: ΔG_{bind}

$$\Delta G_{bind} = \Delta H_{bind} - T\Delta S_{bind}$$



Affinity

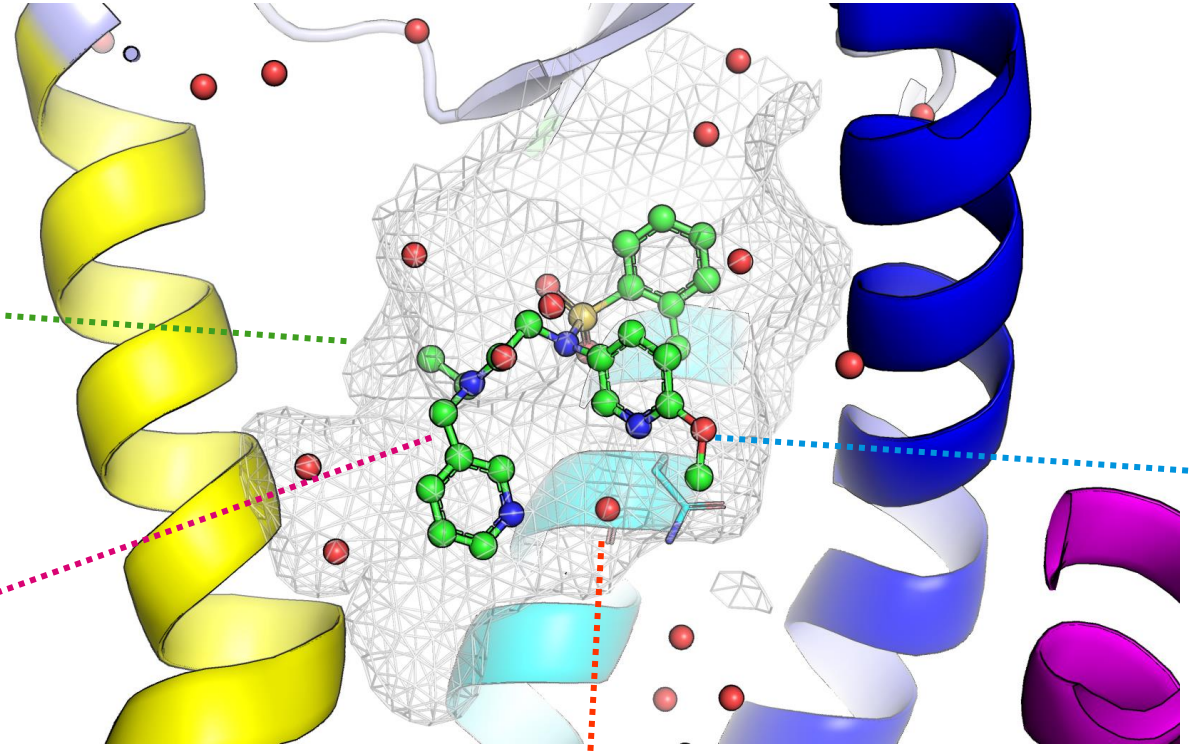
Enthalpy:
Intermolecular interactions
Desolvation

Entropy:
Desolvation
Hydrophobic effect
Conformational Changes

idorsia

ΔG_{bind} & visual inspection?

… one should simply ask ourselves: at what cost?



Shape complementarity

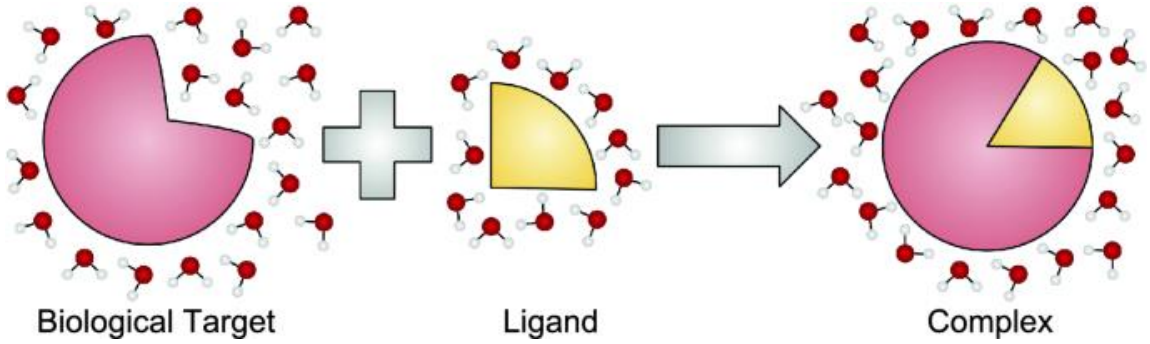
(un)satisfied polar atoms

Ligand conformation (strain)

Water displacement

Desolvation & Hydrophobic effect

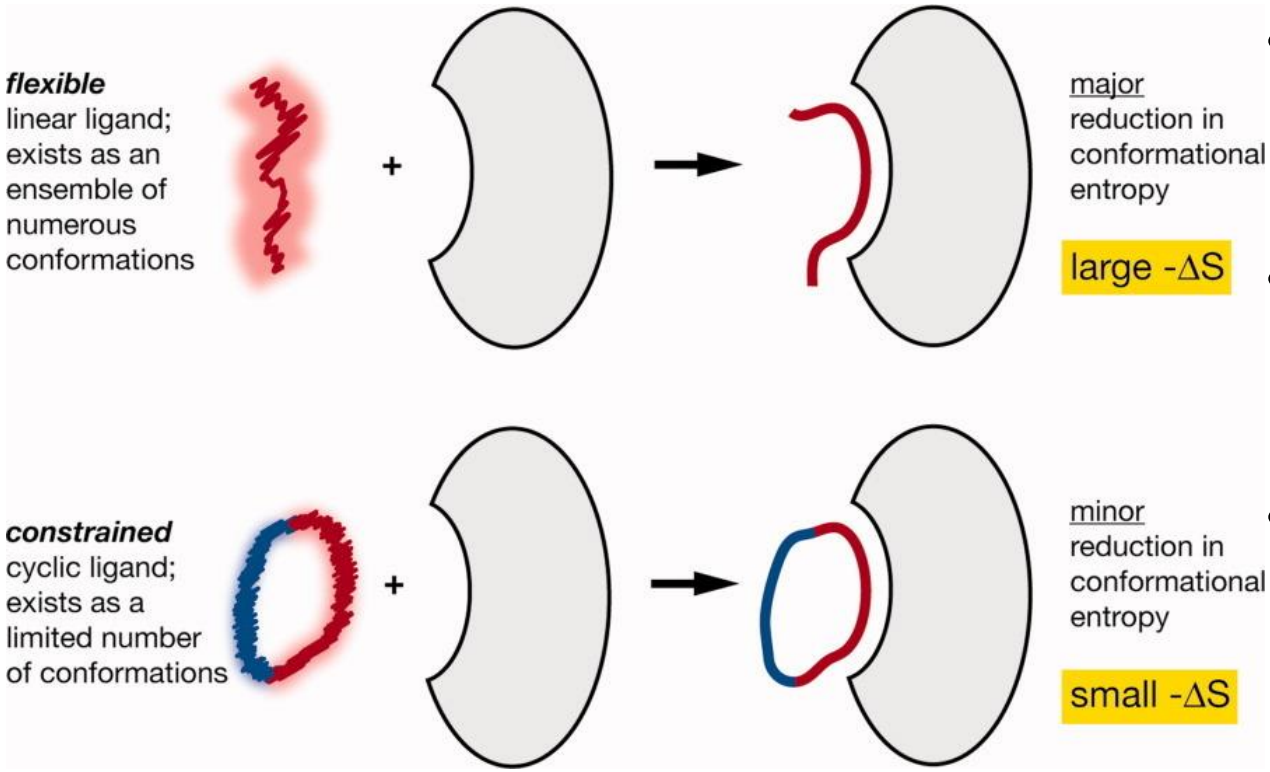
Ligand binding displaces waters around the ligand (water shell) and in the binding site.



- Desolvation of polar compounds is less favorable due to breaking of strong electrostatic interactions with the water shell (enthalpy driven). To compensate this the ligand should make strong interactions with the receptor.

- Desolvation of lipophilic compounds is more favorable due to less electrostatic interactions and the hydrophobic effect: compounds will aggregate to minimize contact with water (entropy driven)

Conformational Entropy

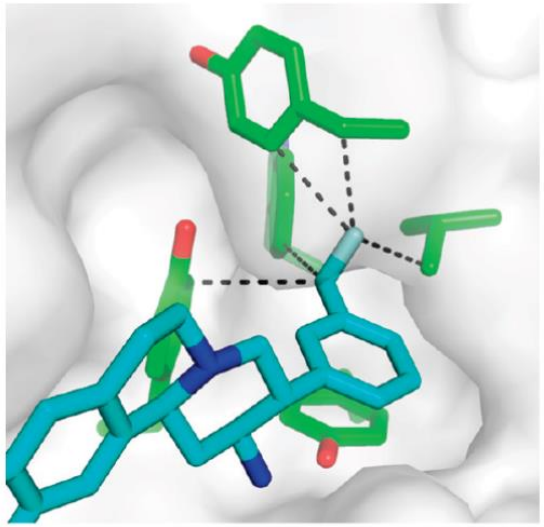


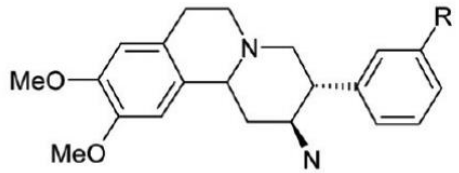
- Refers to the change in the flexibility or degree of freedom of the protein and/or ligand upon binding
- The more conformational states are possible (i.e. flexible molecules), the higher the entropic penalty
- Reducing the degree of freedom of one molecule can help to reduce the affinity (e.g. cyclization, internal strain…)

Biopolymers, **2008**, 8, 653-667

idorsia

Hydrophobic interactions





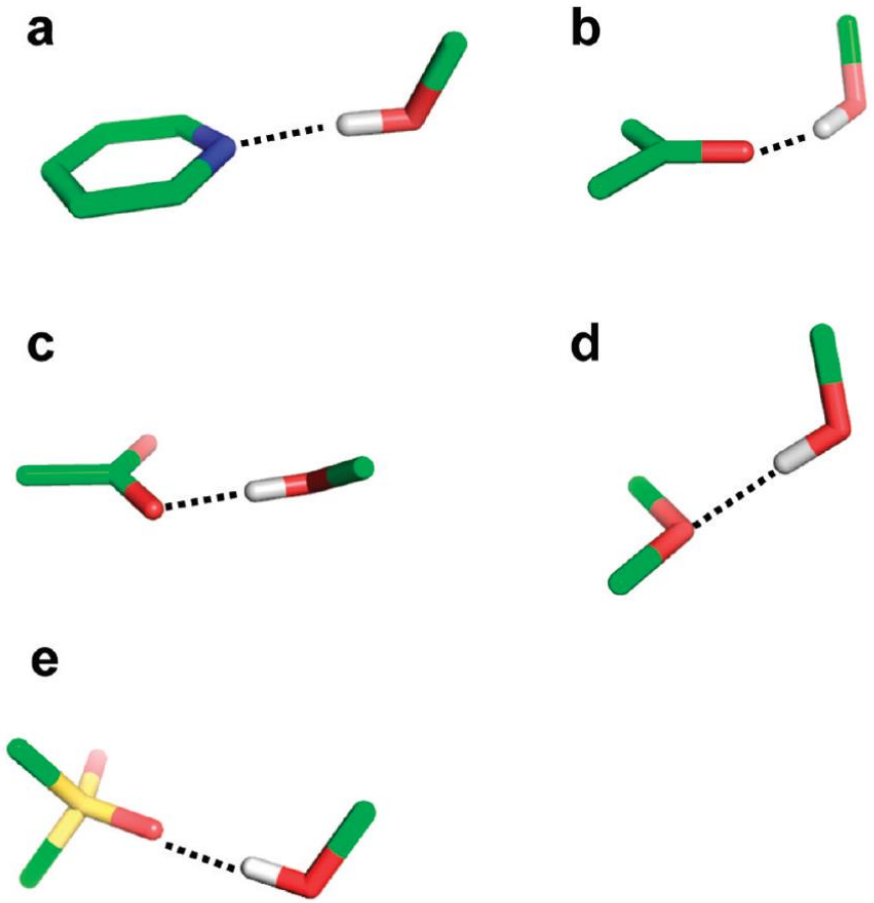
R	K_d (nM)
H	200
CH_3	4.6
CH_2F	0.5

- By far, most common interactions in protein–ligand complexes
- Aliphatic-aliphatic or aliphatic-aryl contacts
- Lower desolvation cost due to hydrophobic effect
- Unfavorable water (*unhappy*) nearby lipophilic sites are easily replaced
- *Magic methyl* is a famous illustration of optimized hydrophobic interaction

J. Med. Chem., **2010**, 53, 5061-5084

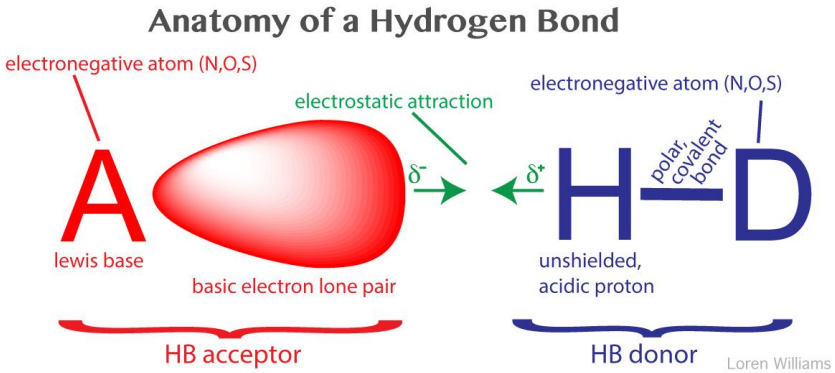
idorsia

Hydrogen bonding

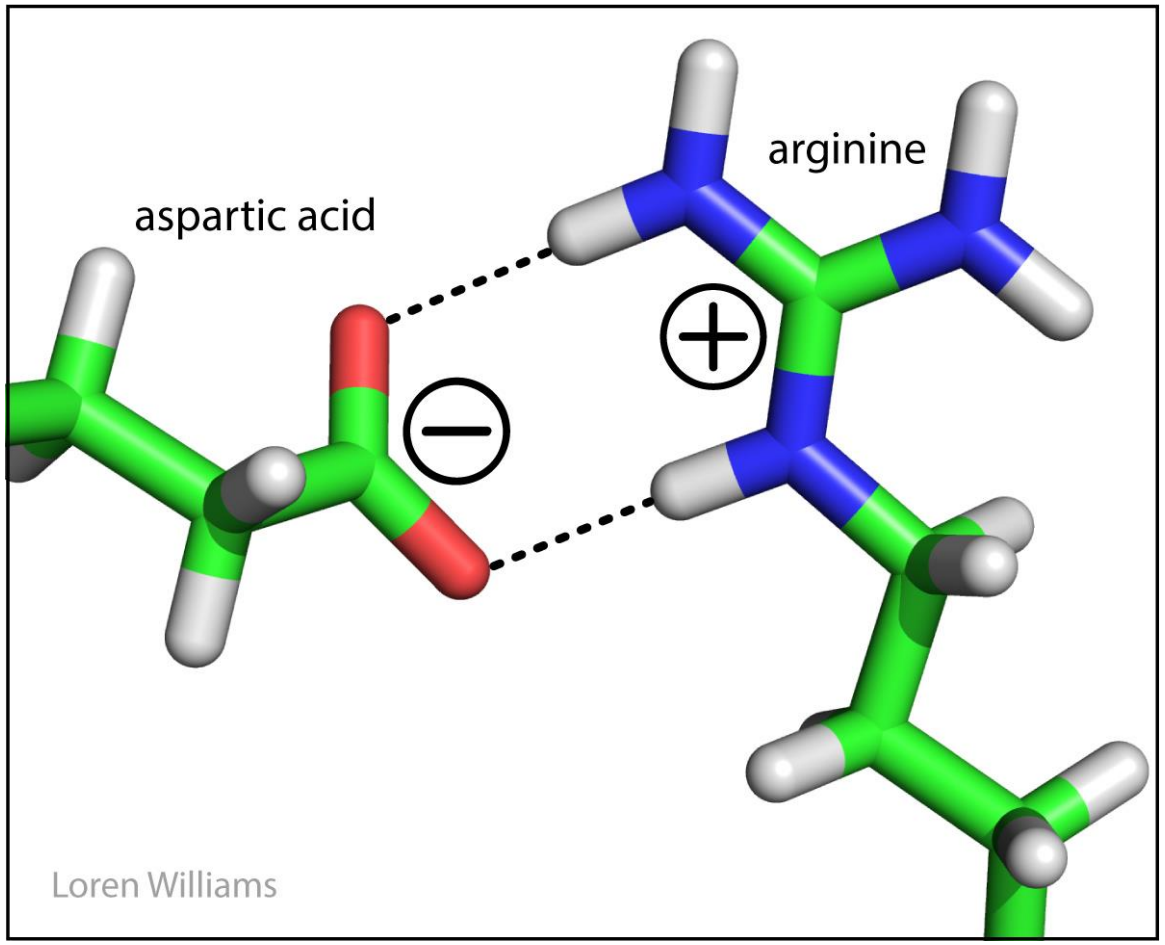


J. Med. Chem., **2010**, 53, 5061-5084

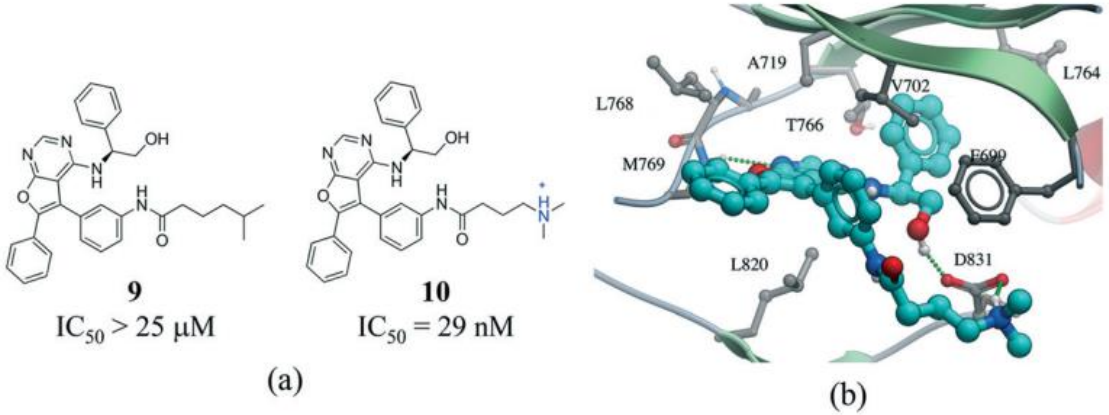
- Stabilizing interaction between a Lewis base (atom with a lone pair) and an electron depleted hydrogen atom
- Highly directional constraints
- Length ranges from 2.6 to 3.1 Angstroms
- Strength of a H-bond correlates with the donor acidity and the acceptor basicity



Salt bridges



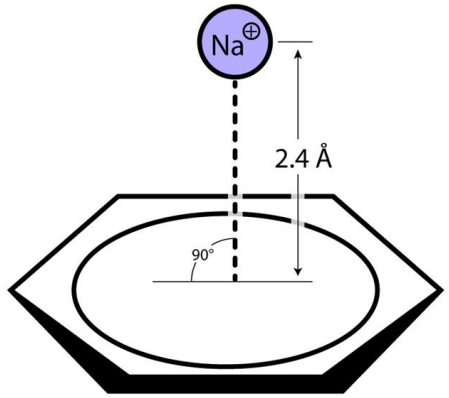
- Combination of H bonding and ionic interaction: ‘charge assisted H bond’
- **High** desolvation cost
- Buried salt bridge tends to be stronger than solvent exposed ones (*i.e.* water competition).

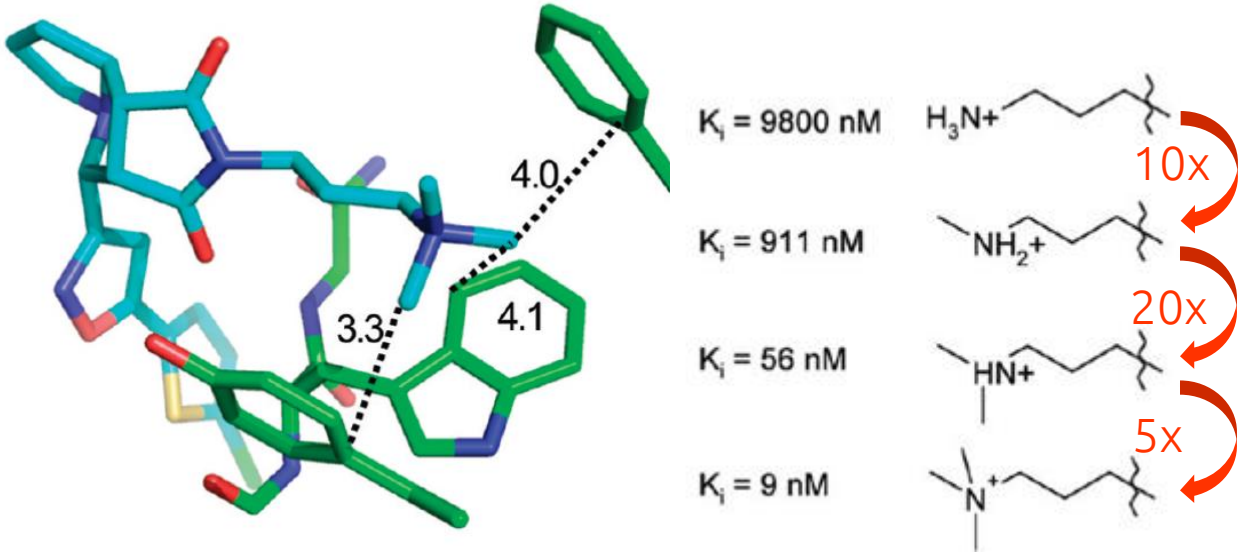


Med. Chem. Commun., **2017**, 8, 1970

Molecular Interactions (gatech.edu)

Cation-Aryl Interaction



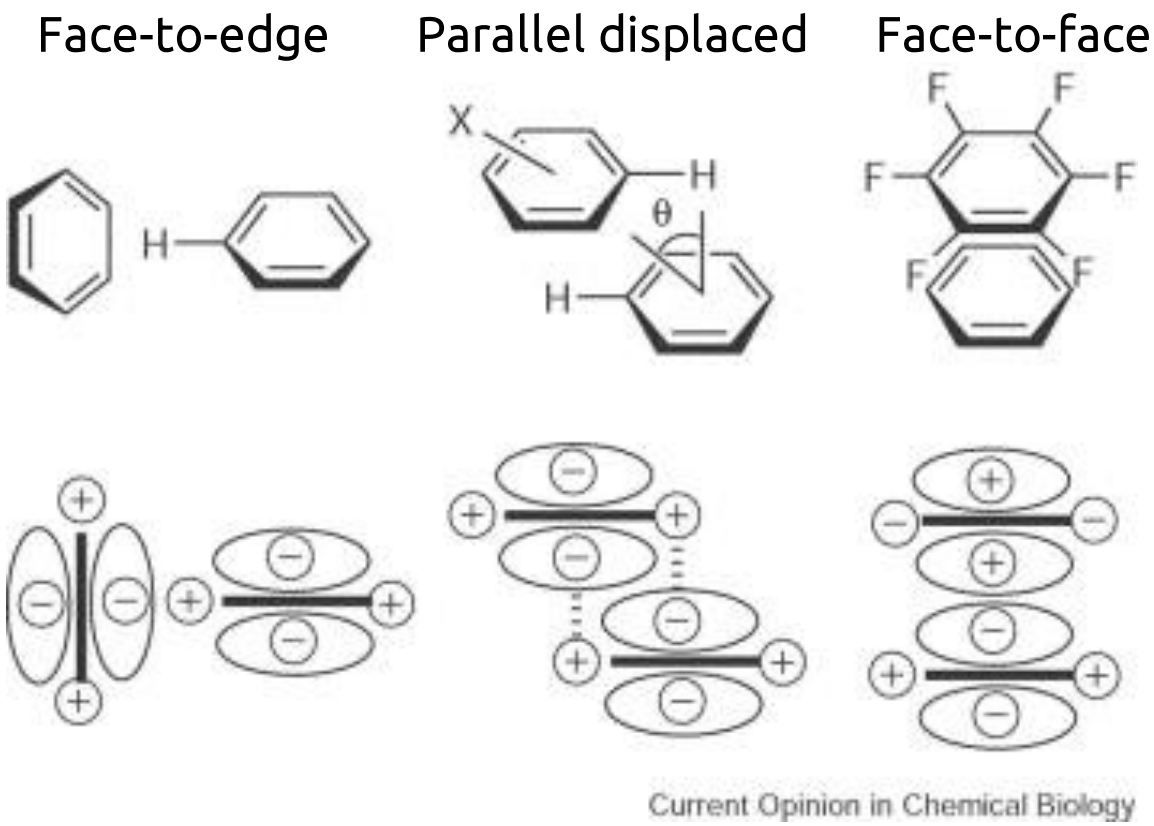


J. Med. Chem., **2010**, 53, 5061-5084

- Ideally, the cation is centered directly over the pi-system
- Majority of cation-aryl interactions occur at the protein surface
  - Arg > Lys
  - Trp > Phe > Tyr
- Lower (but still high) desolvation cost than salt bridge
- Methylation of the basic amine can drastically lower the desolvation cost, thus improving affinity

idorsia

Aryl-Aryl Interaction

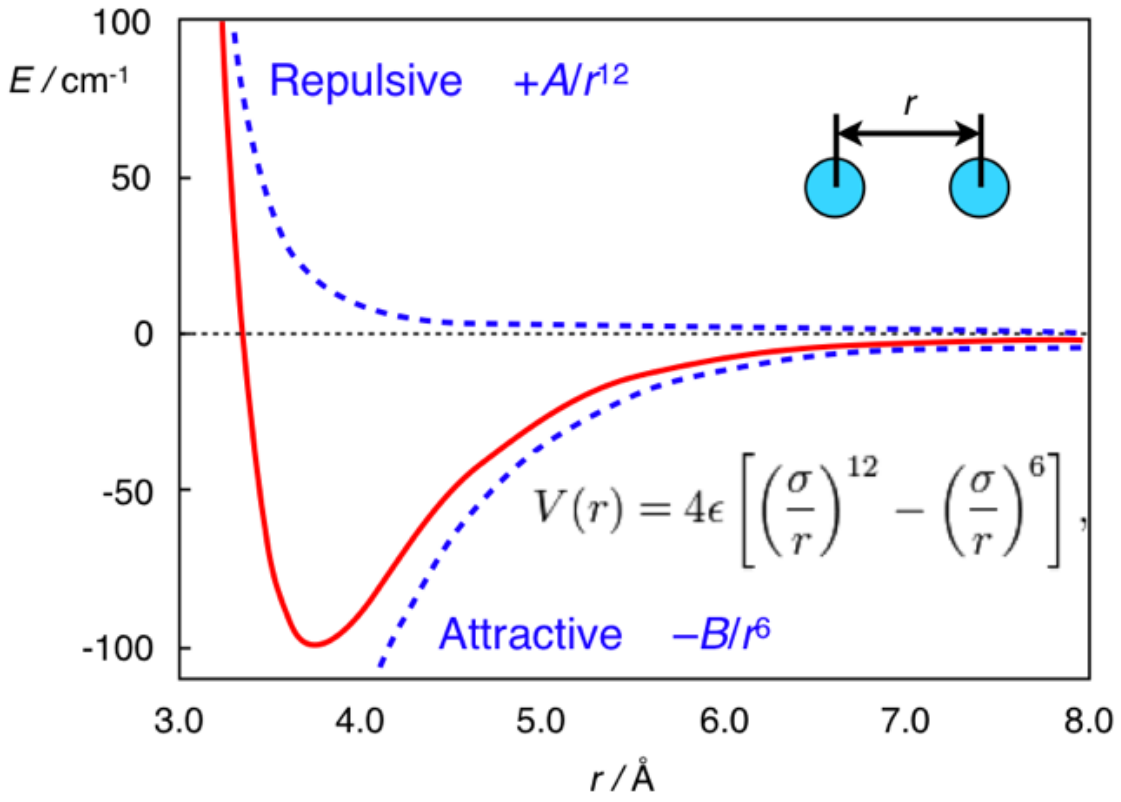


- Interaction between two aromatic rings:
  - Face-to-edge
  - Parallel-displaced
  - Face-to-face: between electron deficient/rich rings, benefits from charge transfer.
- Lower desolvation cost

Curr. Op. Chem. Biol., **2002**, 6, 736-741

Repulsive Interactions

The Lennard-Jones potential describes the interaction between two atoms



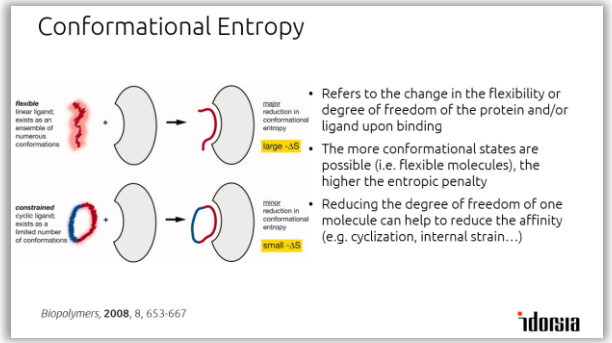
- Repulsion occurs when atoms come too close to each other, and their electron clouds start to overlap
- Small steric clash can kill the affinity
- Surface depiction can be used to quickly assess clashes.
- Distances below the shortest experimentally observed ones are to be avoided.

idorsia

Small molecule conformational preferences

Why look closer?

- Small molecules can adopt various different conformation in solution and when bound to a protein target (=bioactive conformation).
- Energetically, there are preferred conformational states
- **Pre-organization** of small molecules in solution can help reducing the entropic cost upon binding to the protein target



- For rather rigid small molecules, there is less entropic cost when binding to a protein target, but also less flexibility to perfectly interact with the protein within the given geometry
- Conformational preferences can be modulated during the design process by changing an atom, introducing or removing a substituent
  → a molecule can be optimized not only by rigidification but also by adapting its conformational preference according to the bioactive conformation

Literature:

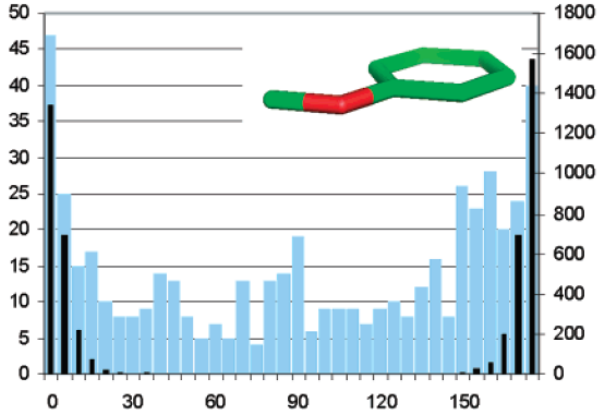
- J. Chem. Inf. Model. 2008, 48, 1-24, K. Brameld et al.:
  "Small Molecule Conformational Preferences Derived from Crystal Structure Data. A Medicinal Chemistry Focused Analysis"

idorsia

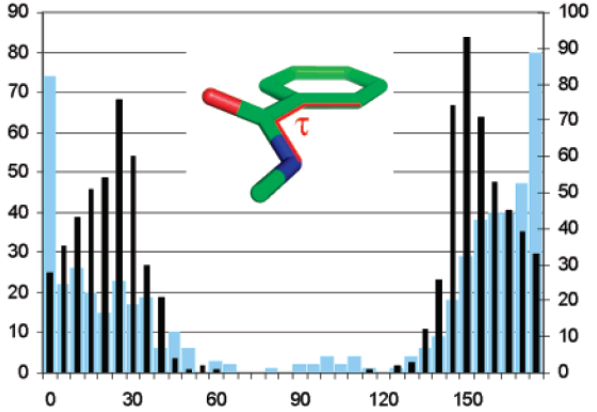
Small molecule conformational preferences

Aryl substituents (planar, 30°, and perpendicular preferences)

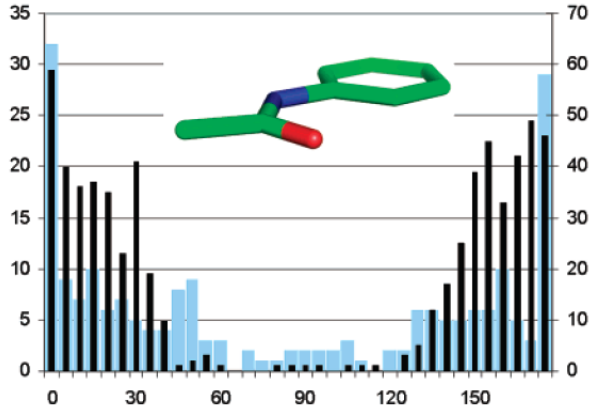
Primary alkoxy substituents



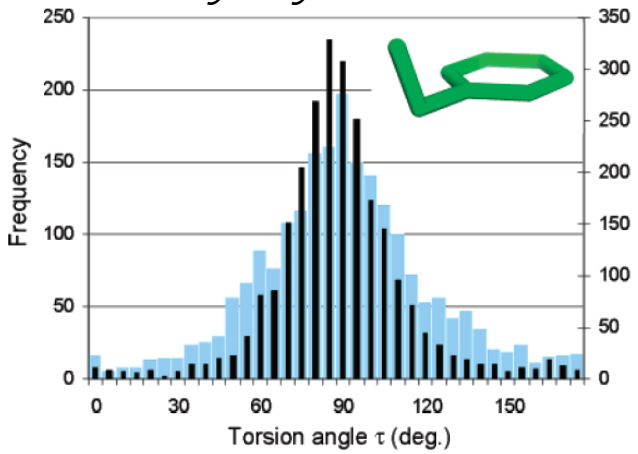
Primary benzamides



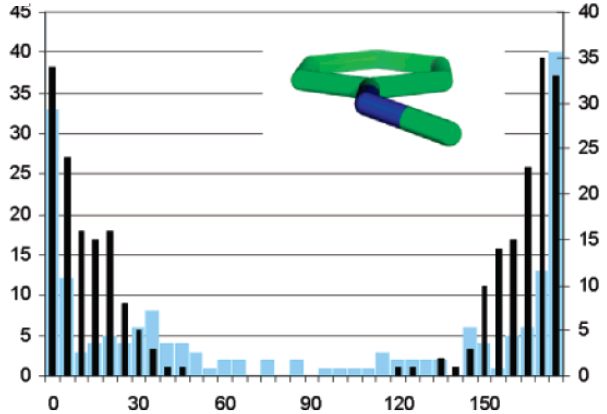
Acylated anilines



Primary alkyl substituents



Substituted anilines



CSD data
PDB data

Figure 12. Torsion histograms (occurrence vs torsion angle bins) for the most frequent aryl substituents in the absence of ortho substituents. In each case, the PDB histogram (light blue, *y*-axis to the left) is overlaid with the corresponding histogram derived from CSD ligands (black, *y*-axis to the right): (a) primary benzamides, (b) acylated anilines, (c) aryl sylfonyl fragments (sulfonamides or sulfones), (d) a subset of (c) which are either methyl sulfones or sulfonamides carrying no N substituents, (e) primary alkoxy substituents, (f) a subset of (e) where methoxy substituents are omitted, (g) primary alkyl substituents, and (h) N,N-disubstituted anilines.

Literature:

- J. Chem. Inf. Model. 2008, 48, 1-24, K. Brameld et al.:
  "Small Molecule Conformational Preferences Derived from Crystal Structure Data. A Medicinal Chemistry Focused Analysis"

Small molecule conformational preferences

Aryl substituents (lone pair repulsion or intramolecular H-bonds)

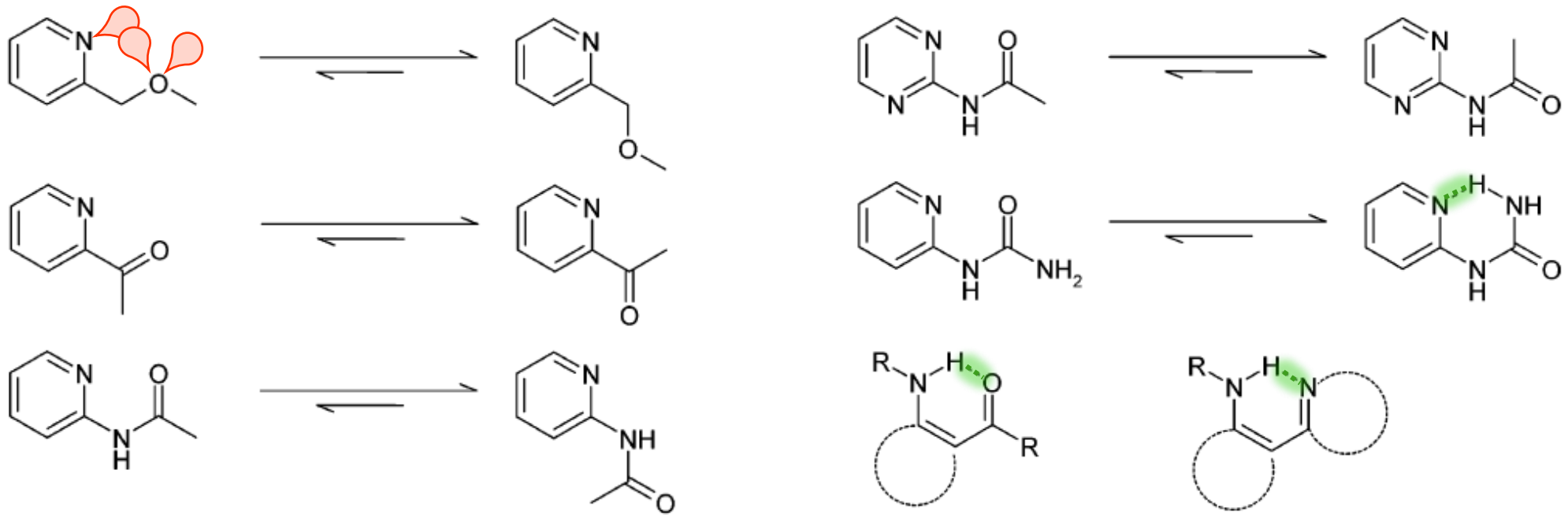


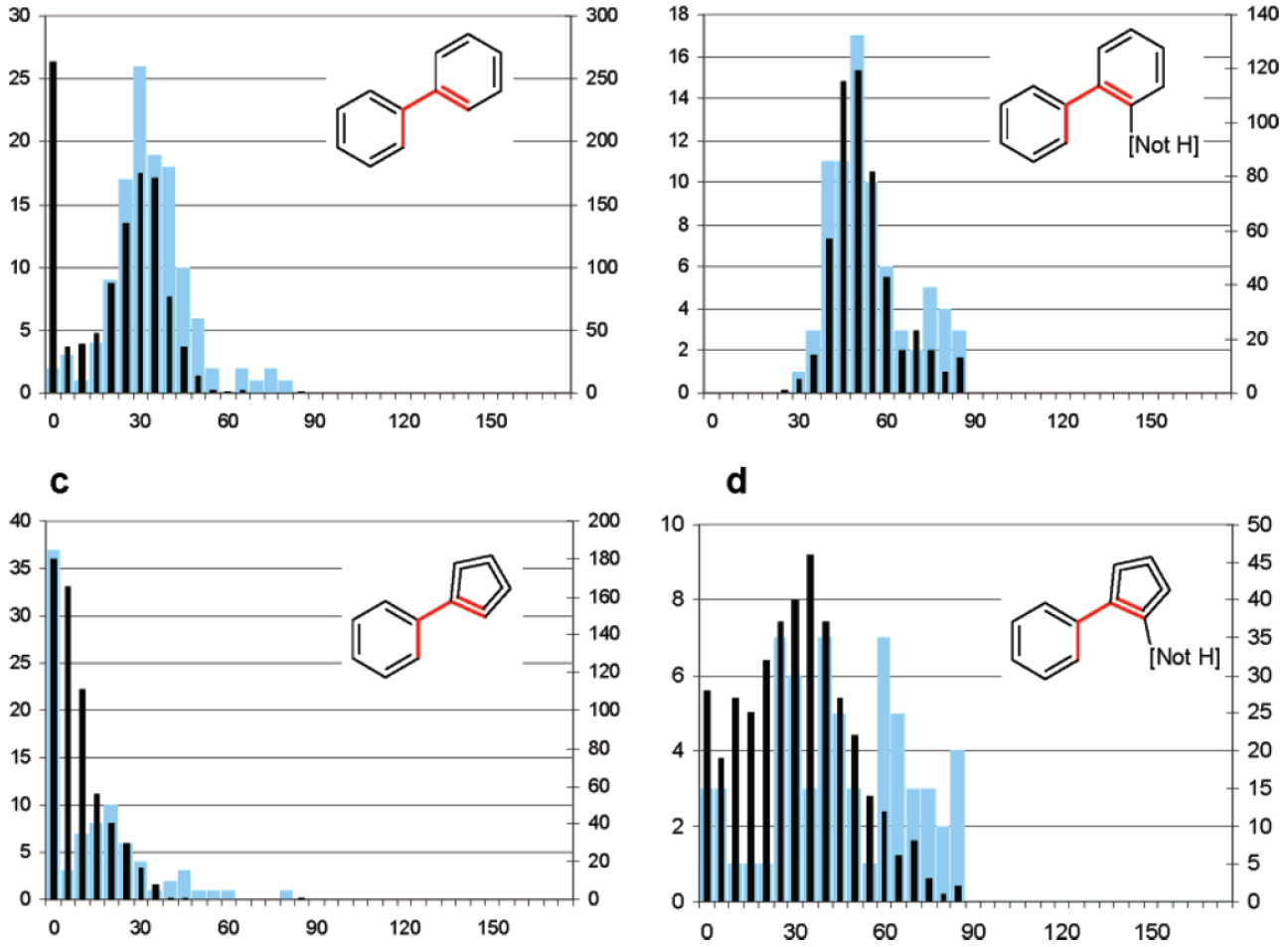
Figure 15. Structural elements adopting preferred conformations due to lone pair repulsion (a–d), intermolecular hydrogen bonds (f, g), or a combination thereof (e).

Literature:

- J. Chem. Inf. Model. 2008, 48, 1-24, K. Brameld et al.:
  "Small Molecule Conformational Preferences Derived from Crystal Structure Data. A Medicinal Chemistry Focused Analysis"

Small molecule conformational preferences

Biaryl systems (effect of ortho substituents)



CSD data
PDB data

Figure 17. Torsion histograms for common biaryl ring systems. The histograms reflect the angle between the ring planes. In each case, the PDB histogram (light blue – *y*-axis to the left) is overlaid with the corresponding histogram derived from CSD ligands (black – *y*-axis to the right): (a) biphenyl systems lacking *ortho* substituents, (b) biphenyl derivatives with one *ortho*-substituent, (c) results of a composite query for various phenyl-substituted planar five-membered heterocycles lacking substituents *ortho* to the biaryl axis, and (d) query analogous to (c) with one *ortho*-substituent on the five-membered ring.

Literature:

- J. Chem. Inf. Model. 2008, 48, 1-24, K. Brameld et al.:
  "Small Molecule Conformational Preferences Derived from Crystal Structure Data. A Medicinal Chemistry Focused Analysis"

Small molecule conformational preferences

Sulfur oxygen contacts

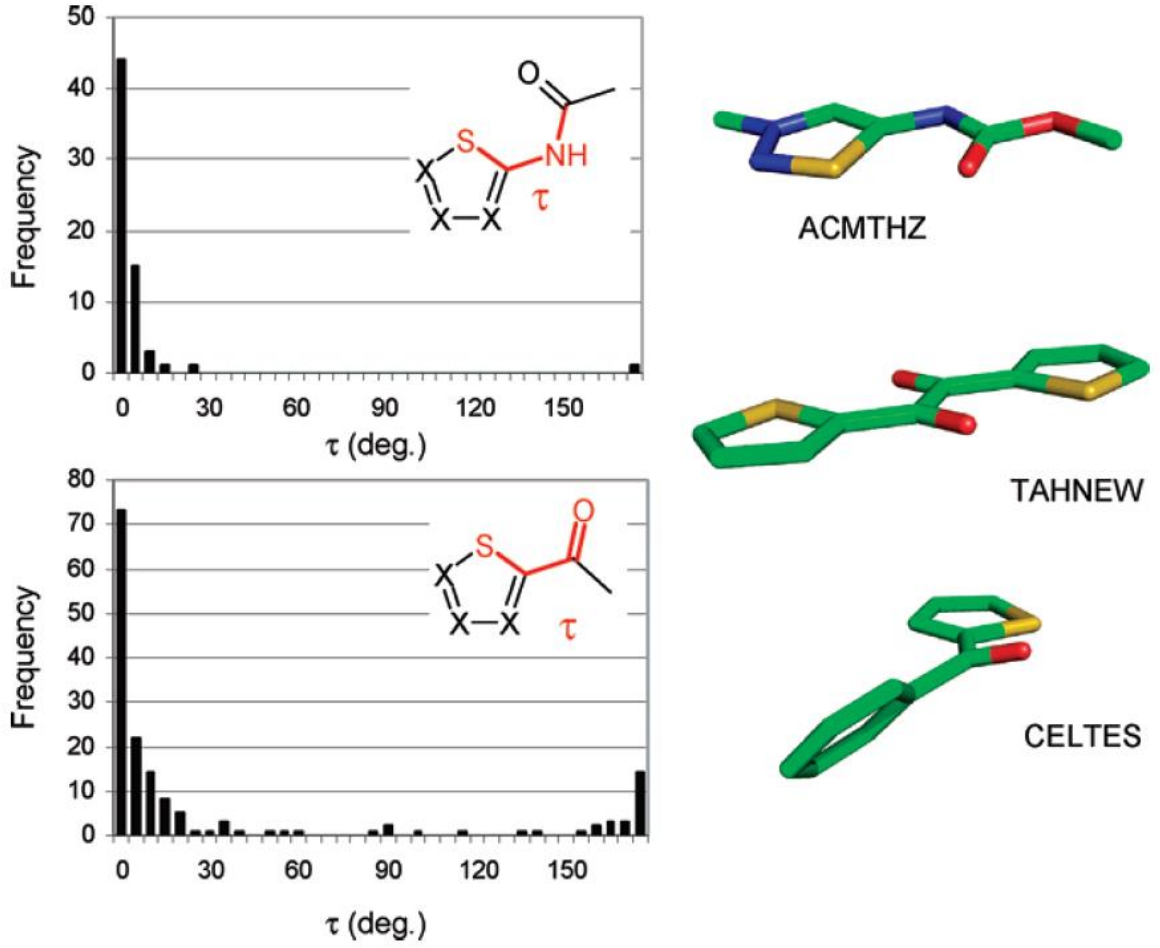


Figure 23. Torsion angle histograms and CSD examples illustrating the attractive interaction between carbonyl oxygen and sulfur in thiophene and related heterocycles.

Literature:

- J. Chem. Inf. Model. 2008, 48, 1-24, K. Brameld et al.:
  "Small Molecule Conformational Preferences Derived from Crystal Structure Data. A Medicinal Chemistry Focused Analysis"

Small molecule conformational preferences

Sulfonamides (strong preference for conformation I)

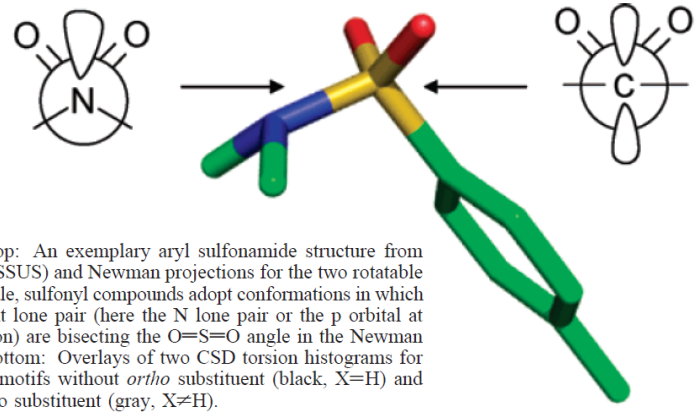
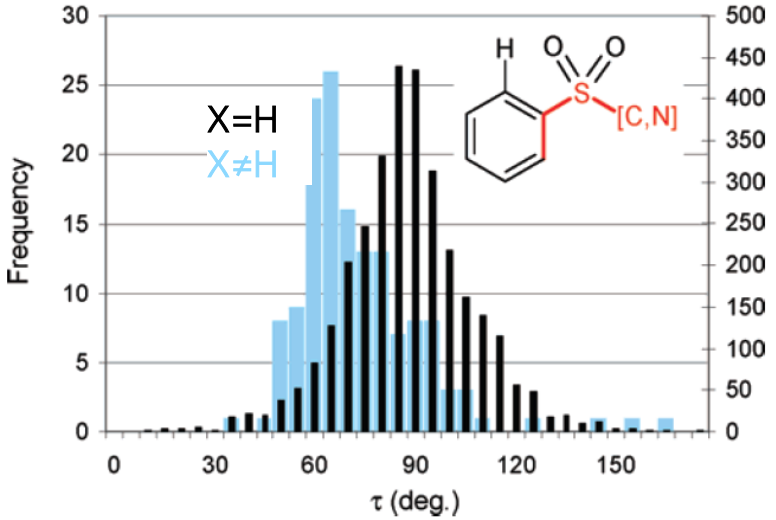
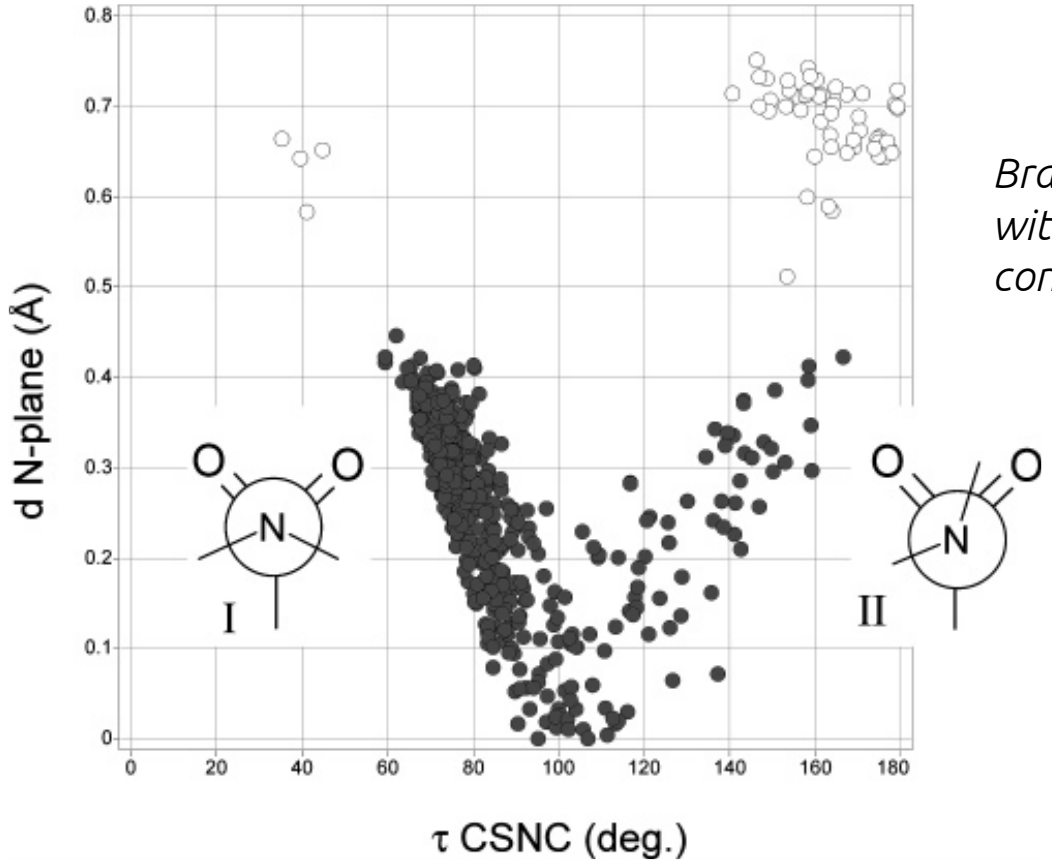
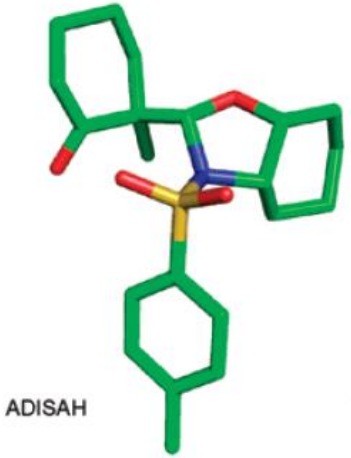


Figure 9. Top: An exemplary aryl sulfonamide structure from the CSD (GESSUS) and Newman projections for the two rotatable bonds. As a rule, sulfonyl compounds adopt conformations in which the substituent lone pair (here the N lone pair or the p orbital at the *ipso* carbon) are bisecting the O=S=O angle in the Newman projection. Bottom: Overlays of two CSD torsion histograms for aryl sulfonyl motifs without *ortho* substituent (black, X=H) and with one ortho substituent (gray, X≠H).





Branched N substituents with steric hindrance in conformation I



Literature:

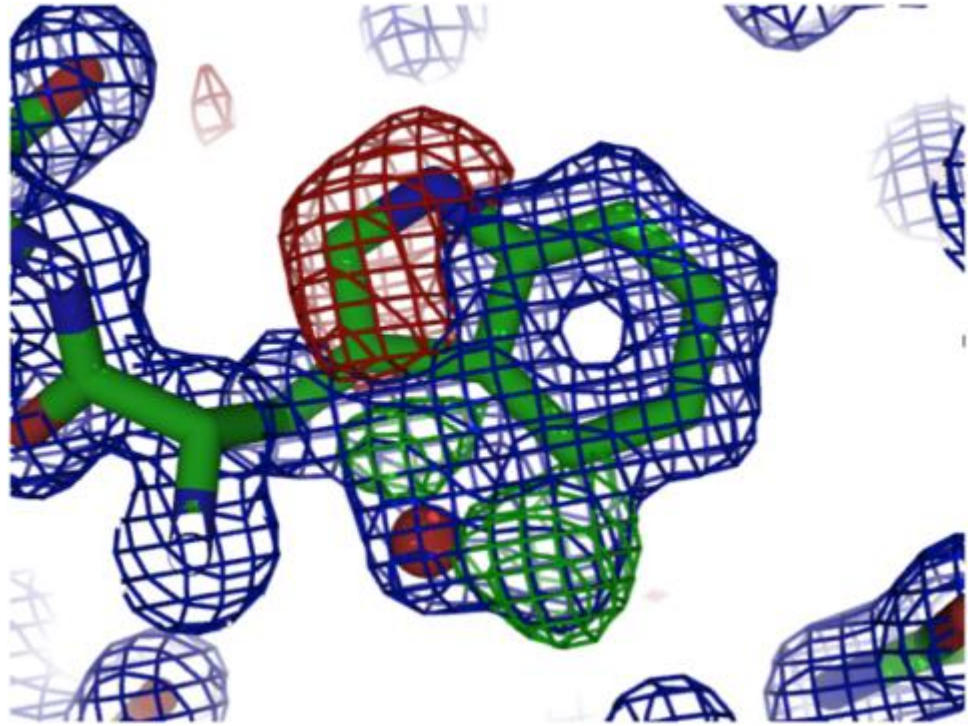
- J. Chem. Inf. Model. 2008, 48, 1-24, K. Brameld et al.:
  "Small Molecule Conformational Preferences Derived from Crystal Structure Data. A Medicinal Chemistry Focused Analysis"

When working with xray structures…

The structure model is an interpretation of the electron density maps

- PDB structure is a model: experiment is observing electron density within a protein crystal (averaging disordered segments)
- Protein amino acid sequence is known and can be fitted into the observed electron density – remaining electron density belongs to ligands, waters, salts, alternative conformations of amino acid side chains or entire loops, buffer molecules, etc.
- The structural model is iteratively refined by adding more and more atoms into the density
- Experimental electron density map "$F_{observed}$" can't be looked at directly
- 2Fo-Fc map: combines observed diffraction data (Fo) with more weight and diffraction data calculated from the atomic model (Fc). The atomic model should explain as much observed density as possible.
- Fo-Fc map – difference electron density map: shows where the experimentally observed density and the built atomic model disagree:
  - Positive (green) Fo-Fc density: experimentally observed density, that is not present in the atomic model (atoms are missing in the model) → further refine model by adding atoms into this density
  - Negative (red) Fo-Fc density: less or no observed density at the position of an atom in the model (atom position or occupancy in the model is wrong)
    → refine model by either removing atoms in this position or by lowering the atom's occupancy
    → for electron-rich elements such as sulfur: check for radiation damage

– Examining the correspondence between the electron density map and the published molecular model reveals regions of uncertainty in the model that are relevant to the modeler working with the structure



Literature:

- Electron density maps - Proteopedia, life in 3D
- PDBe brings electron-density viewing to the masses | Protein Data Bank in Europe (ebi.ac.uk)

Electron density maps

When working with xray structures – have a look!

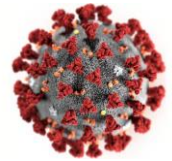
- Some residues might need to be flipped (e.g. His, Asn, Gln):
  → check whether orientation and possible protonation state makes sense by looking at the density and the surrounding amino acids
- The ligand binding mode or a part of a ligand within the binding pocket can be shaky, alternative binding modes might not have been modeled
- Water molecules might be missing in the model or could also be less occupied than indicated in the model
- Take these findings into account when starting to derive key interactions from your structure before you use these for your prospective modeling

Outline

- Part 1: *(2:30 h)*
  - General introduction:
    Most common Protein Ligand interactions
    small molecule conformational preferences
    electron densities
  - Introduction to the target Sars-Cov2 main protease MPro
  - Introduction to PyMol and coot
  - Exercise 1.1-5 and Discussions
- Part 2: *(1:30h)*
  - General introduction: Halogen bonds
  - Exercise 2.1
  - Halogen bond example from Idorsia, other interactions
  - Exercise 2.2

idorsia

SARS-CoV-2 pandemic



First scientific publication in February 2020

The NEW ENGLAND JOURNAL of MEDICINE

BRIEF REPORT

A Novel Coronavirus from Patients with Pneumonia in China, 2019

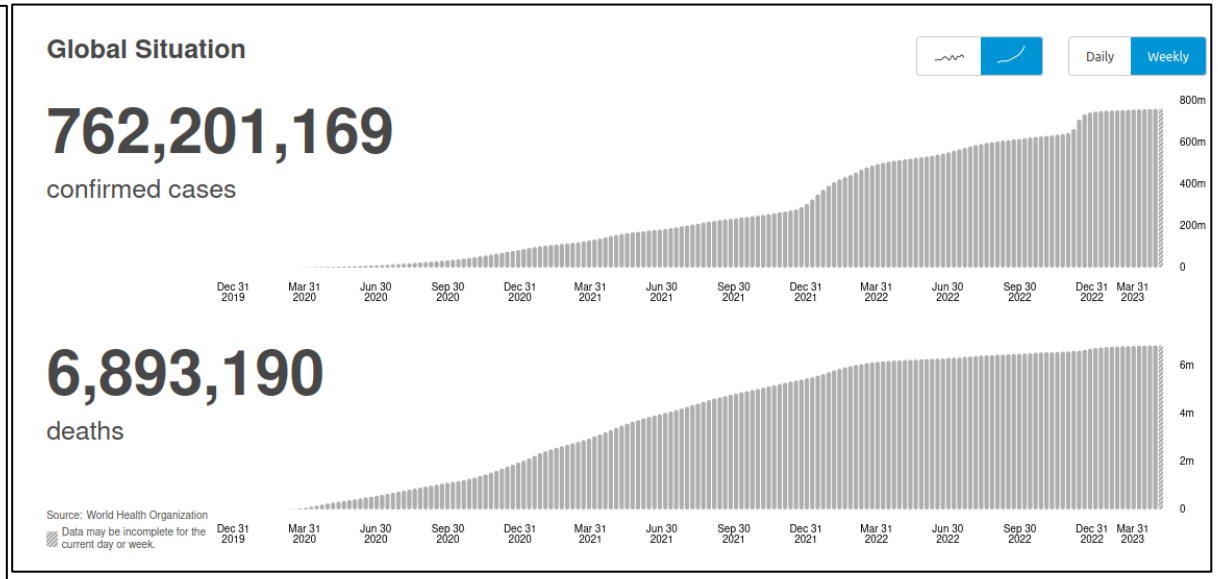
Na Zhu, Ph.D., Dingyu Zhang, M.D., Wenling Wang, Ph.D., Xingwang Li, M.D., Bo Yang, M.S., Jingdong Song, Ph.D., Xiang Zhao, Ph.D., Baoying Huang, Ph.D., Weifeng Shi, Ph.D., Roujian Lu, M.D., Peihua Niu, Ph.D., Faxian Zhan, Ph.D., Xuejun Ma, Ph.D., Dayan Wang, Ph.D., Wenbo Xu, M.D., Guizhen Wu, M.D., George F. Gao, D.Phil., and Wenjie Tan, M.D., Ph.D., for the China Novel Coronavirus Investigating and Research Team

SUMMARY

In December 2019, a cluster of patients with pneumonia of unknown cause was linked to a seafood wholesale market in Wuhan, China. A previously unknown betacoronavirus was discovered through the use of unbiased sequencing in samples from patients with pneumonia. Human airway epithelial cells were used to isolate a novel coronavirus, named 2019-nCoV, which formed a clade within the subgenus sarbecovirus, Orthocoronavirinae subfamily. Different from both MERS-CoV and SARS-CoV, 2019-nCoV is the seventh member of the family of coronaviruses that infect humans. Enhanced surveillance and further investigation are ongoing. (Funded by the National Key Research and Development Program of China and the National Major Project for Control and Prevention of Infectious Disease in China.)

From the NHC Key Laboratory of Biosafety, National Institute for Viral Disease Control and Prevention, Chinese Center for Disease Control and Prevention (N.Z., W.W., J.S., X.Z., B.H., R.L., P.N., X.M., D.W., W.X., G.W., G.F.G., W.T.), and the Department of Infectious Diseases, Beijing Ditan Hospital, Capital Medical University (X.L.) — both in Beijing; Wuhan Jinyintan Hospital (D.Z.), the Division for Viral Disease Detection, Hubei Provincial Center for Disease Control and Prevention (B.Y., F.Z.), and the Center for Biosafety Mega-Science, Chinese Acade-

Situation as of today



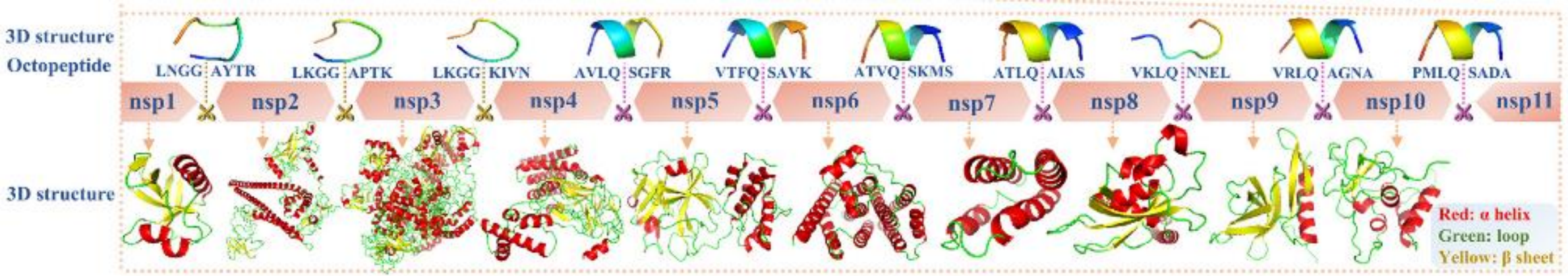
Source: WHO

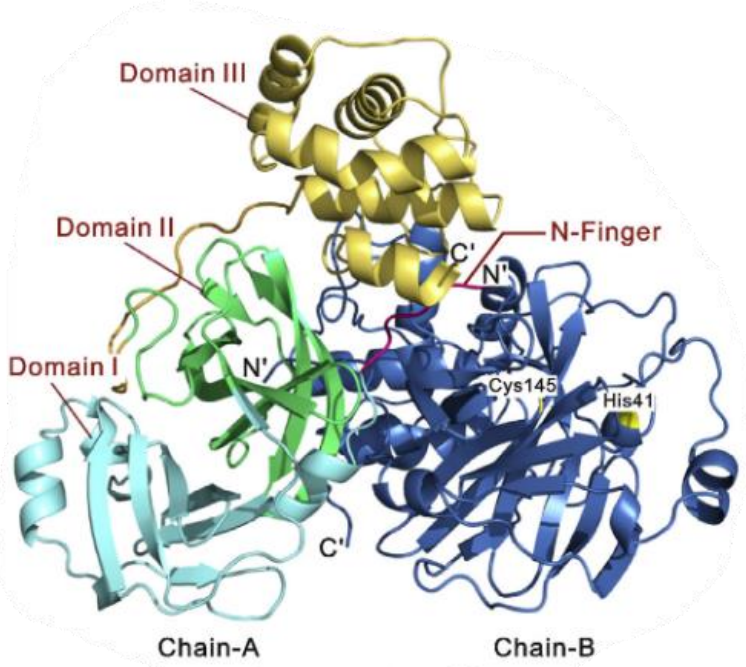
Over the last 20 years, the world has experienced 3 coronaviruses outbreaks:

- 2003: SARS
- 2012: MERS
- 2019: SARS-CoV-2

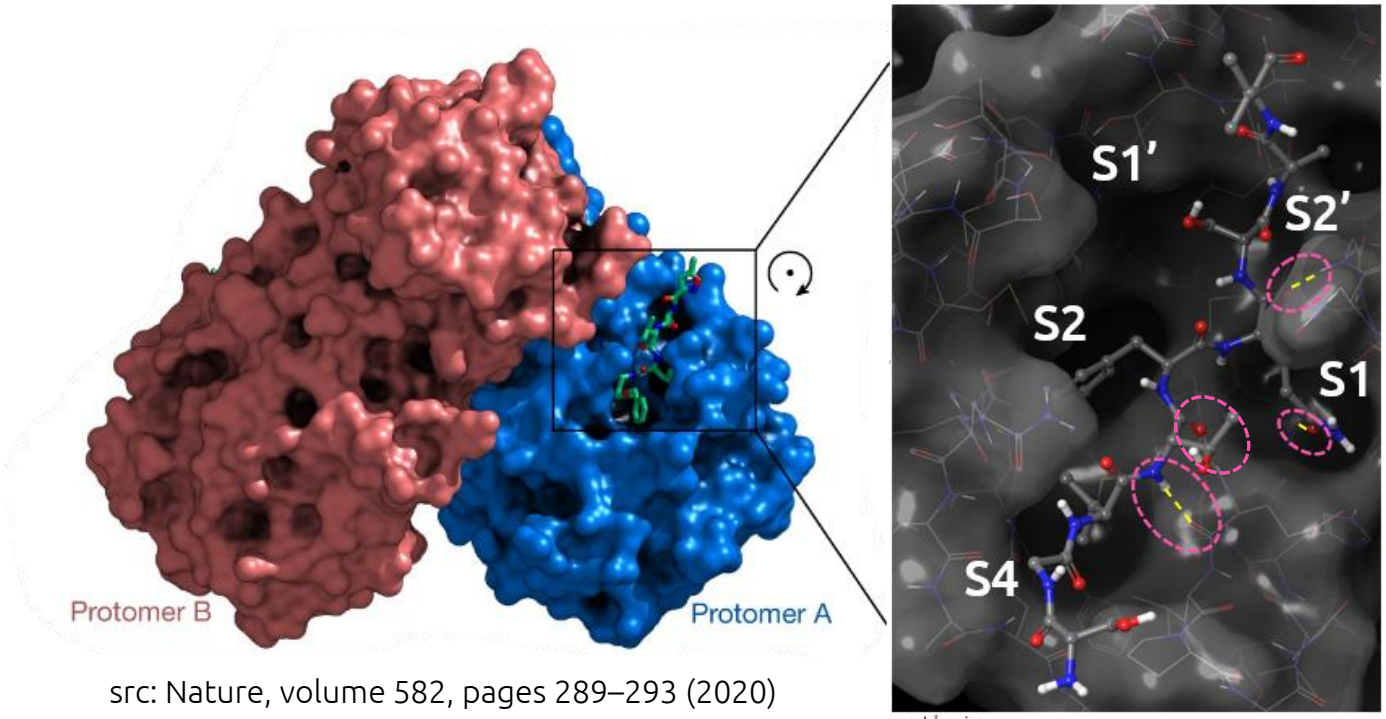
idorsia

Mpro structure in complex with the substrate





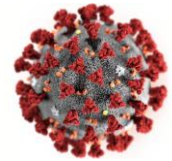
Journal of Pharmaceutical Analysis
Volume 10, Issue 4, August 2020, Pages 313-319

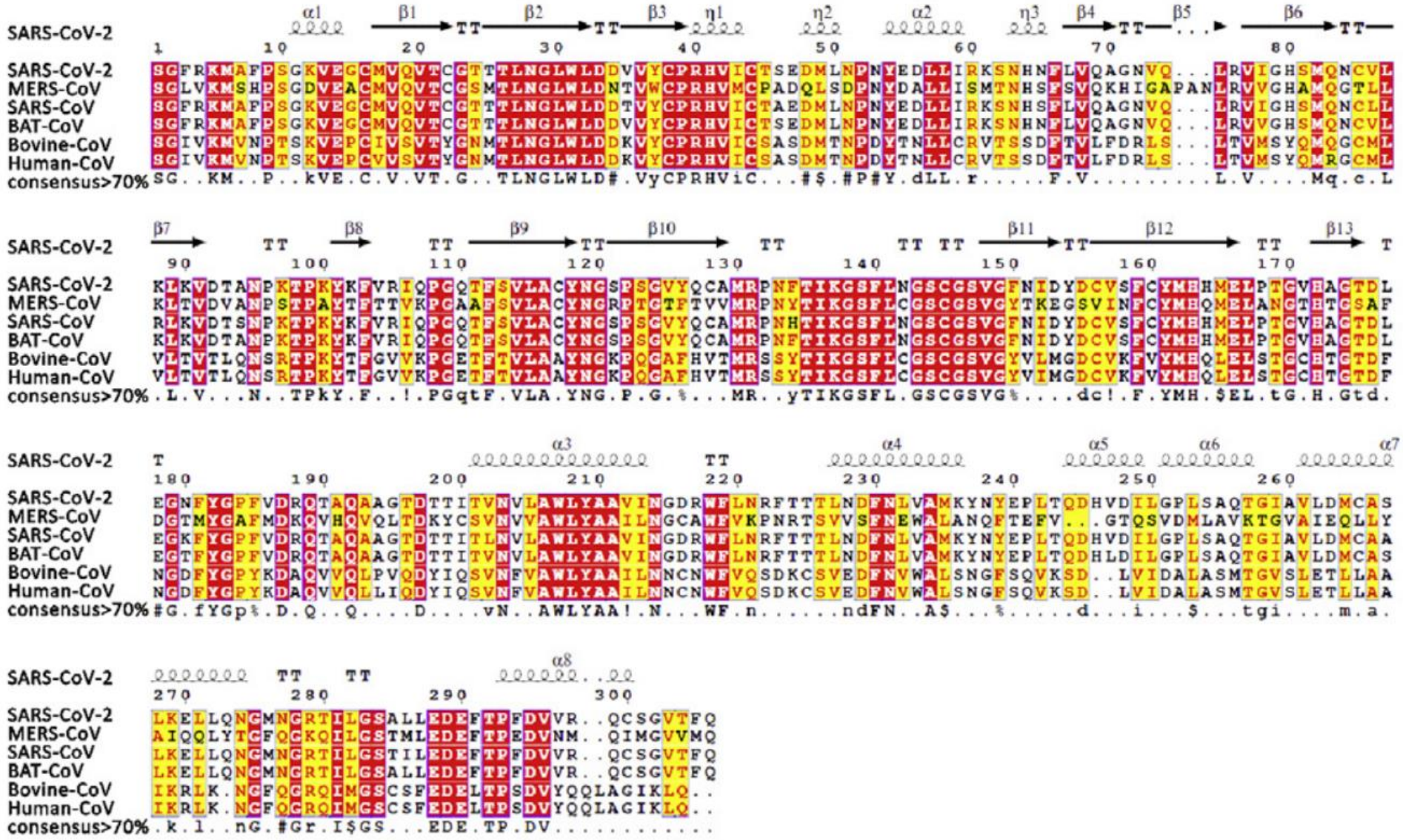


src: Nature, volume 582, pages 289–293 (2020)

src: Idorsia

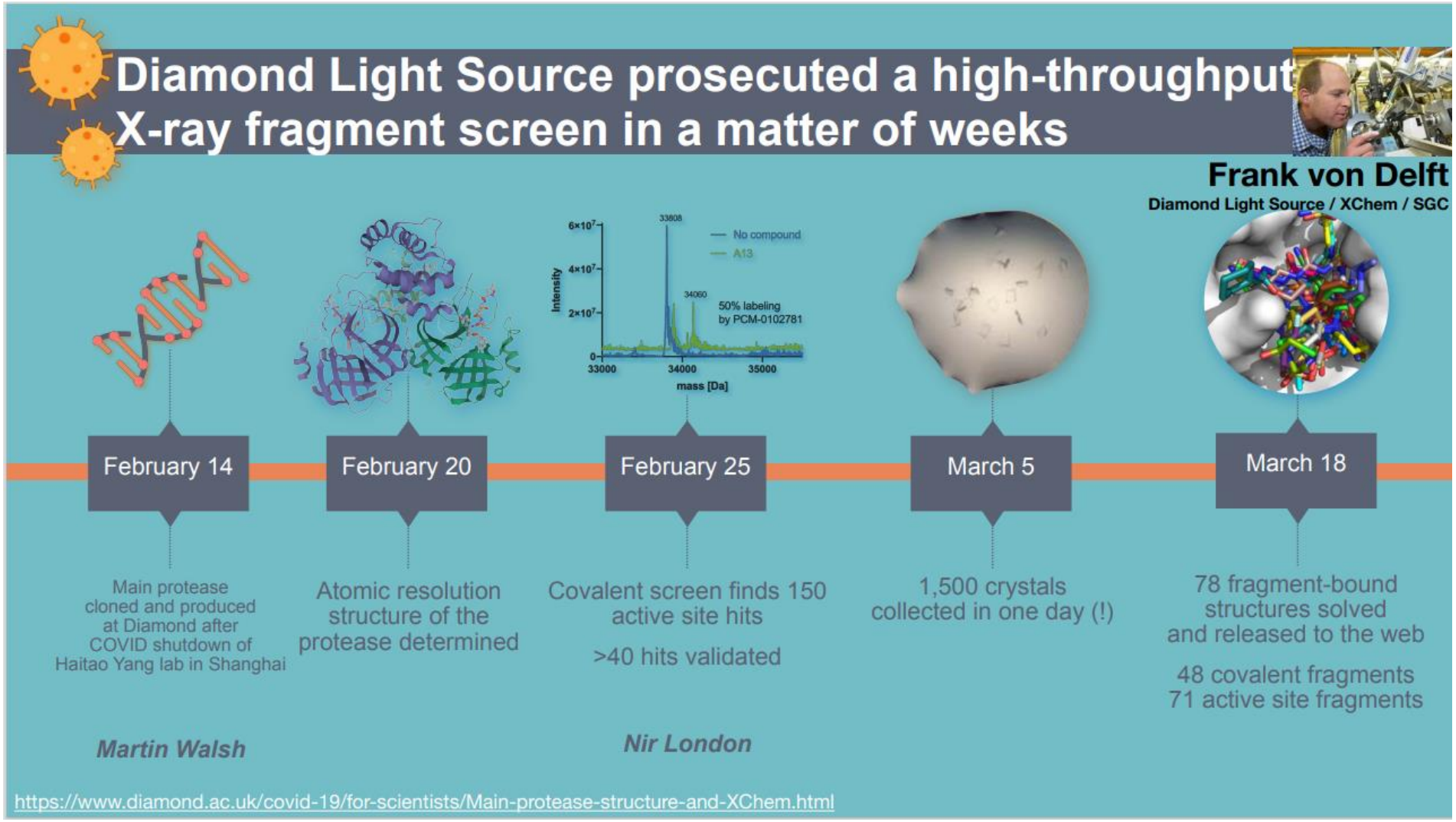
Mpro sequence analyses across various coronaviruses





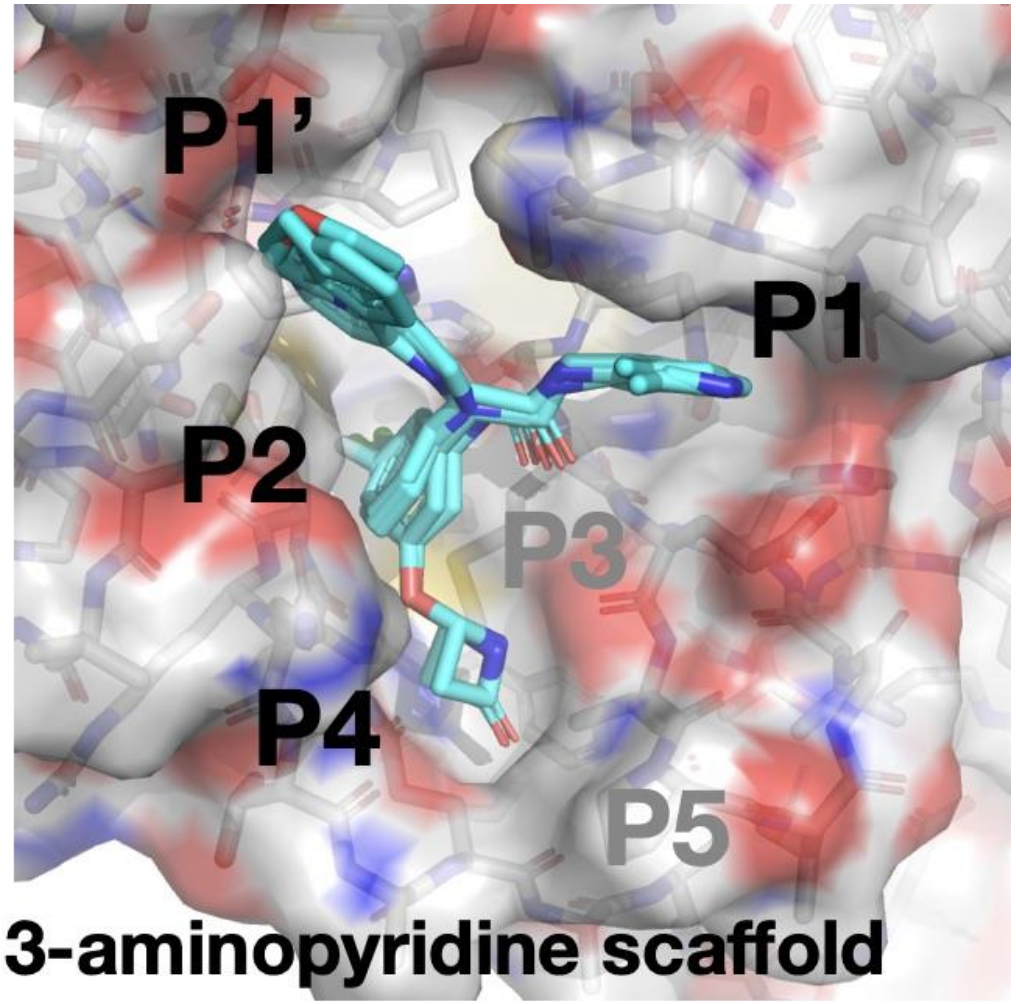
The COVID Moonshot

Crystallographic screen in 2020



The COVID Moonshot

Drug Discovery



COVID Moonshot
516 Tweets

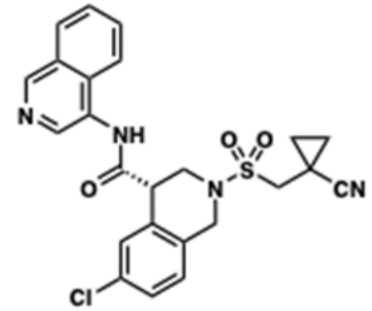
Follow

Tweets | Replies | Media | Likes

Pinned Tweet

COVID Moonshot @covid_moonshot · Feb 1, 2022

Today we're delighted to share a significant milestone in Moonshot's journey towards an open-science COVID antiviral: We have identified compounds meeting our criteria for progression to preclinical evaluation!



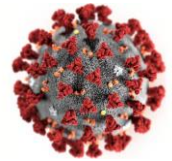
MAT-POS-e194df51-1

ALT

37 nM SARS-CoV-2 Mpro IC_{50} (enzymatic)
64 nM SARS-CoV-2 antiviral EC_{50} (A549 cells)

6 157 429

SARS-CoV-2 Mpro inhibitor Drug Discovery



Outlook

- Mpro represent an ideal target for the following reasons:
  - Absence in human biology**
  - Highly conserved three-dimensional structure among various coronaviruses***
  - Essential role in viral replication****
- Risk of decrease in vaccine protection overtime due to virus mutations*
- Possible outbreaks of new coronaviruses in the coming years
- An oral inhibitor could provide prophylaxis following exposure or treat illness at onset of symptoms

*Madhi et al. (2021). Efficacy of the ChAdOx1 nCoV-19 Covid-19 Vaccine against the B.1.351 Variant. *New England Journal of Medicine*, 1–14.
**Goyal, B et al. (2020). Targeting the Dimerization of the Main Protease of Coronaviruses: A Potential Broad-Spectrum Therapeutic Strategy. In *ACS Combinatorial Science* (Vol. 22, Issue 6, pp. 297–305).
***Tahir ul Qamar et al. (2020). Structural basis of SARS-CoV-2 3CLpro and anti-COVID-19 drug discovery from medicinal plants. *Journal of Pharmaceutical Analysis*, *10*(4), 313–319.
****De Wit et al. (2016). SARS and MERS: Recent insights into emerging coronaviruses. *Nature Reviews Microbiology*, *14*(8), 523–534.

idorsia

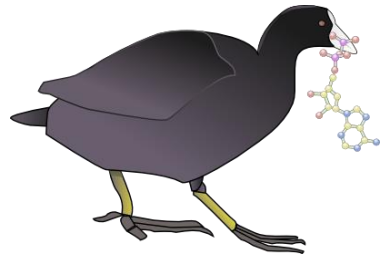
Outline

- Part 1: *(2:30 h)*
  - General introduction:
    Most common Protein Ligand interactions
    small molecule conformational preferences
    electron densities
  - Introduction to the target Sars-Cov2 main protease MPro
  - Introduction to PyMol and coot
  - Exercise 1.1-5 and Discussions
- Part 2: *(1:30h)*
  - General introduction: Halogen bonds
  - Exercise 2.1
  - Halogen bond example from Idorsia, other interactions
  - Exercise 2.2

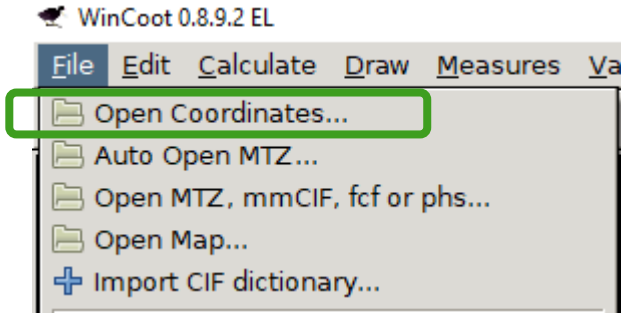
idorsia

Mini intro Coot 1

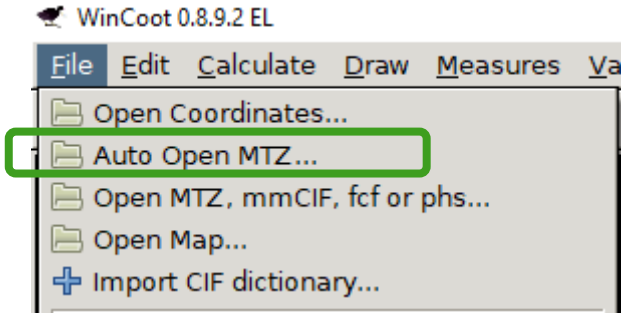
Coot displays maps and models



- Open your .pdb file (coordinates):
  File → Open Coordinates…



- Open your .mtz file (reflection data from the xray experiment):
  File → Auto Open MTZ…



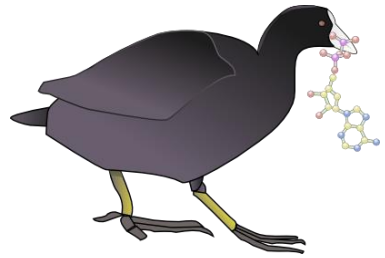
This will automatically create the 2FoFc and the FoFc maps. By default the contour level for sigma is 1.5 for 2FoFc and ±3.0 for FoFc maps.

Literature:

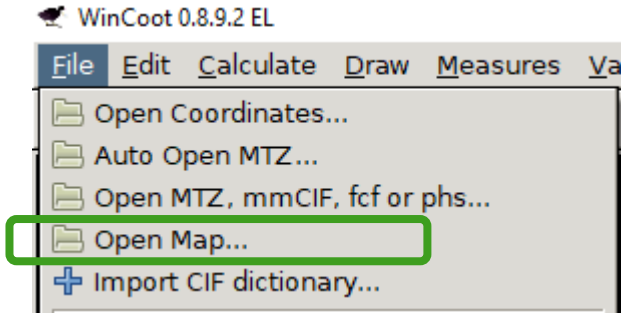
- Coot website
- Acta Crystallographica Section D66, 2010, 486-501, P. Emsley et al.: “Features and development of Coot”

Mini intro Coot 2

In case you don’t have the .mtz

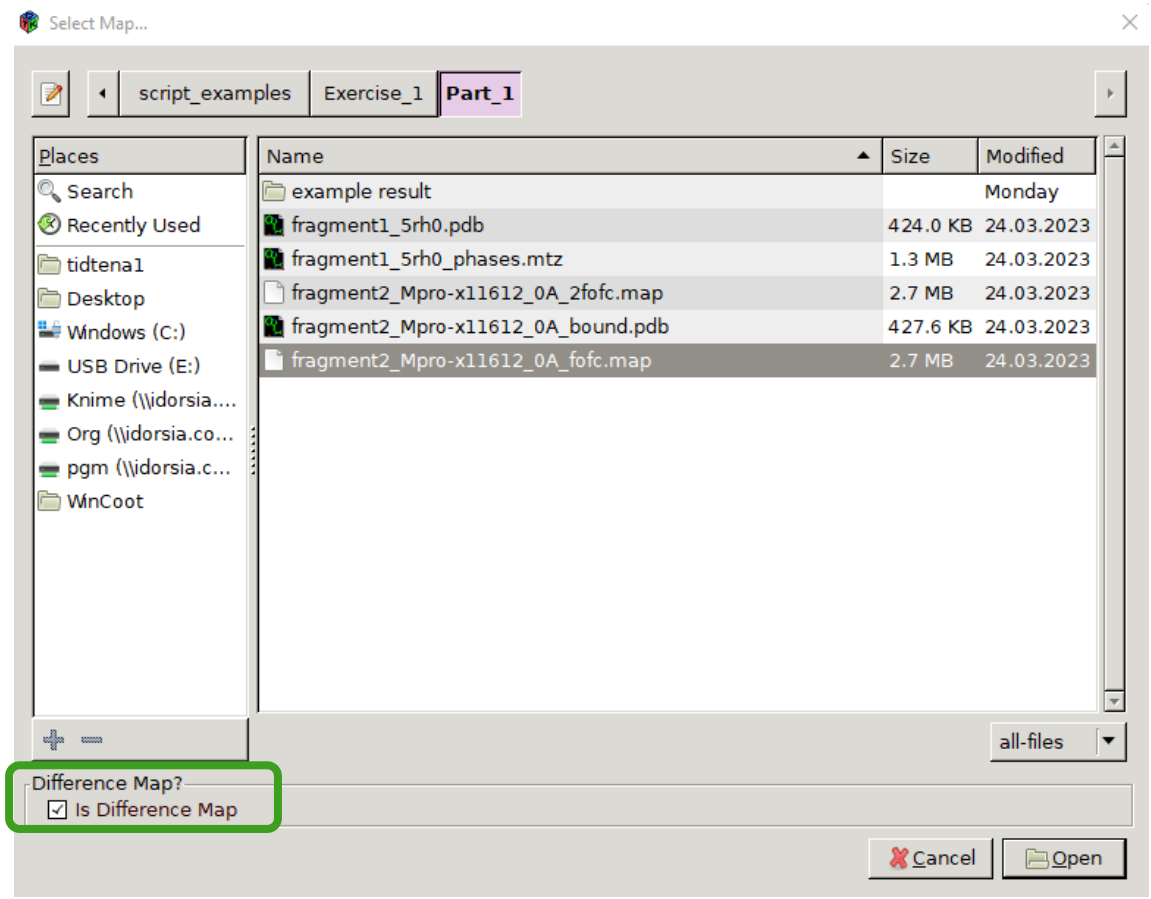


- Open your .map files (2FoFc and FoFc maps):
  File → Open Map…



Make sure to tick the box for the FoFc map in the next dialog.

By default the contour level for sigma is 1.5 for 2FoFc and ±3.0 for FoFc maps.

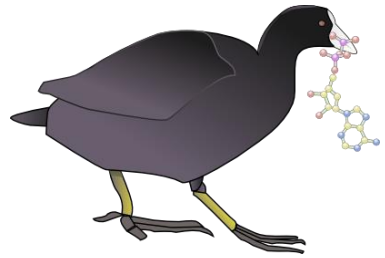


Literature:

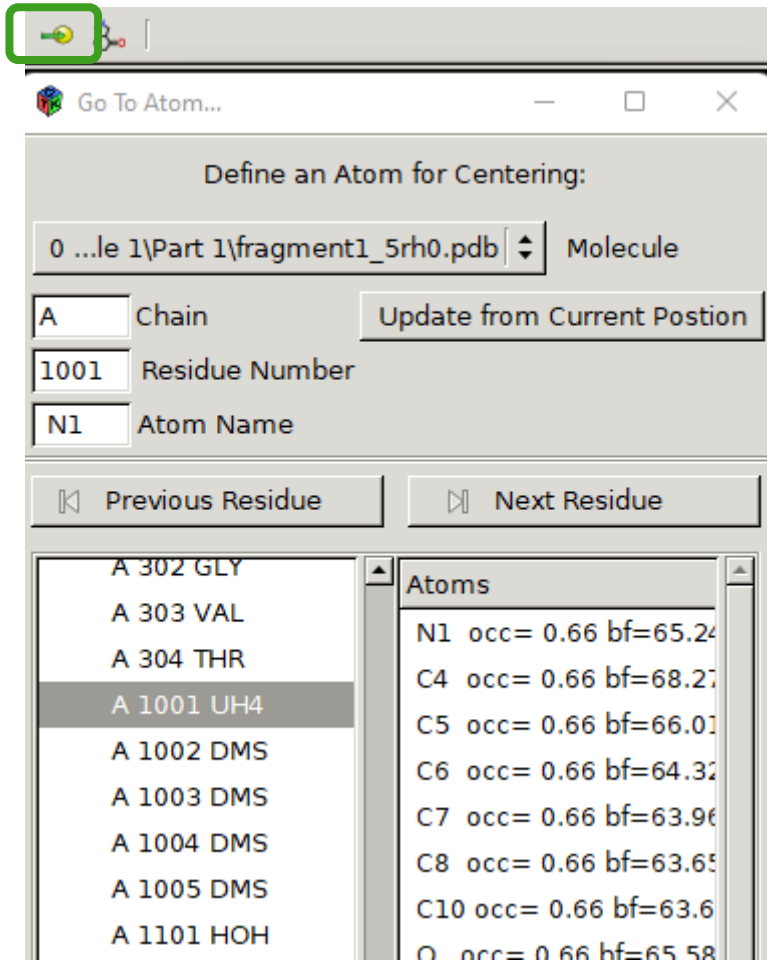
- Coot website
- Acta Crystallographica Section D66, 2010, 486-501, P. Emsley et al.: “Features and development of Coot”

Mini intro Coot 3

How to navigate



- Jump to your ligand in the viewer:
  - ➢ click the little symbol for “Go To Atom”  or press F6
  - ➢ in the new window “open” the chains by clicking on the triangle and scroll down through the residues. Click on the residue that is the ligand (hint: usually the ligand comes after all amino acids and before solvents/ions)
  - ➢ in the viewer window the center will jump to the selected molecule
- Rotate: Hold the left mouse button and move
- Zoom: Hold the right mouse button and move
- Jump: Middle-button-click on an atom and the viewer center will jump
- Drag: Hold the middle mouse button and move

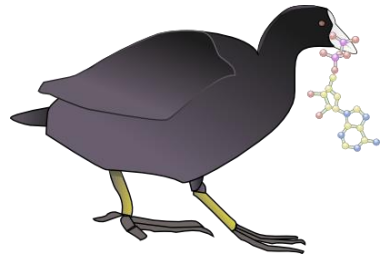


Literature:

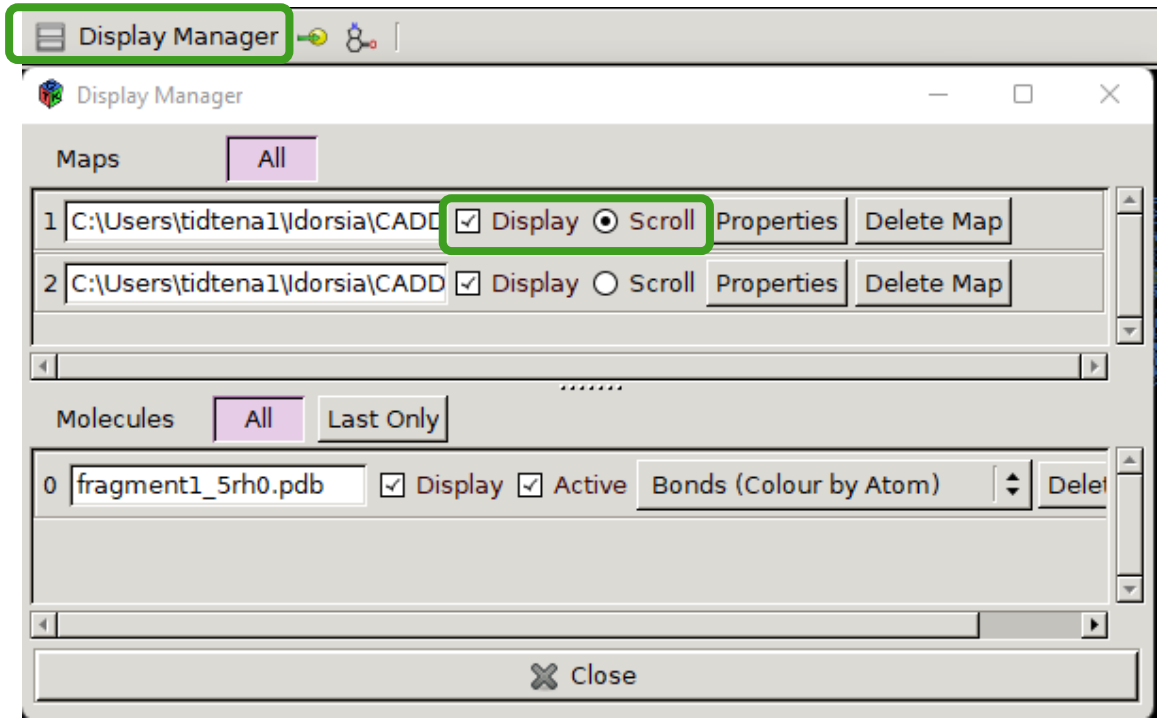
- Coot website
- Acta Crystallographica Section D66, 2010, 486-501, P. Emsley et al.: “Features and development of Coot”

Mini intro Coot 4

How switch objects on and off, scroll through sigma contour levels



- Open the Display Manager (shortcut F7)
- All opened maps and coordinate files are listed and can be switched on and off (tick or untick the “Display” box)
- Only one map at a time can be selected for scrolling through sigma contour levels
- Scroll: use the mouse wheel
  (the sigma level will be displayed on the top of the viewer window)



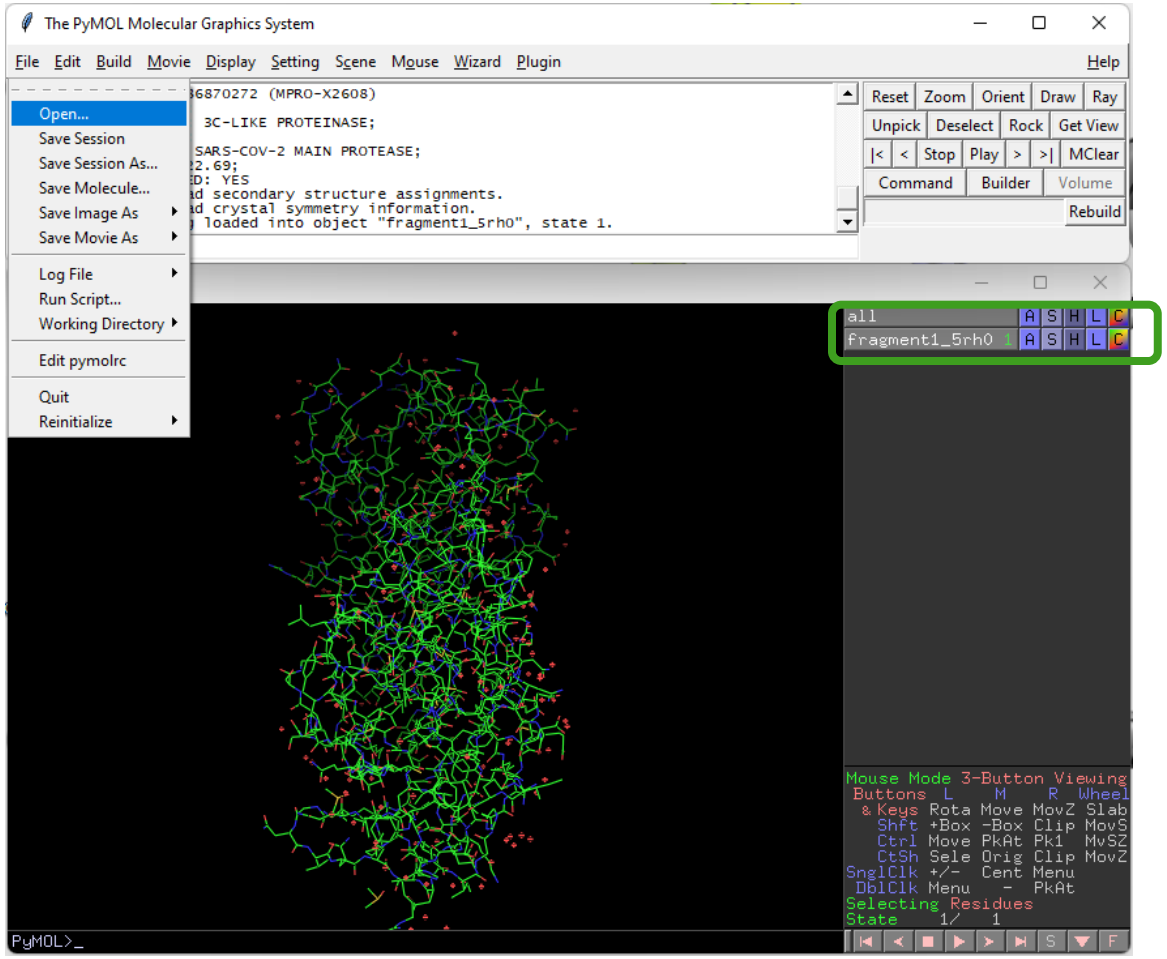
Literature:

- Coot website
- Acta Crystallographica Section D66, 2010, 486-501, P. Emsley et al.: “Features and development of Coot”

Mini intro PyMol 1

PyMOL is a user-sponsored molecular visualization system on an open-source foundation, maintained and distributed by Schrödinger.

- PyMol has 2 windows (viewer & commands menu)
- File → Open… and select your .pdb file
- A new object appears on the object panel
- Switch the object on and off in the viewer by clicking on it in the object panel
- Click on the “S” (show) next to the object to change visualization styles (lines, sticks, cartoon, surface…)



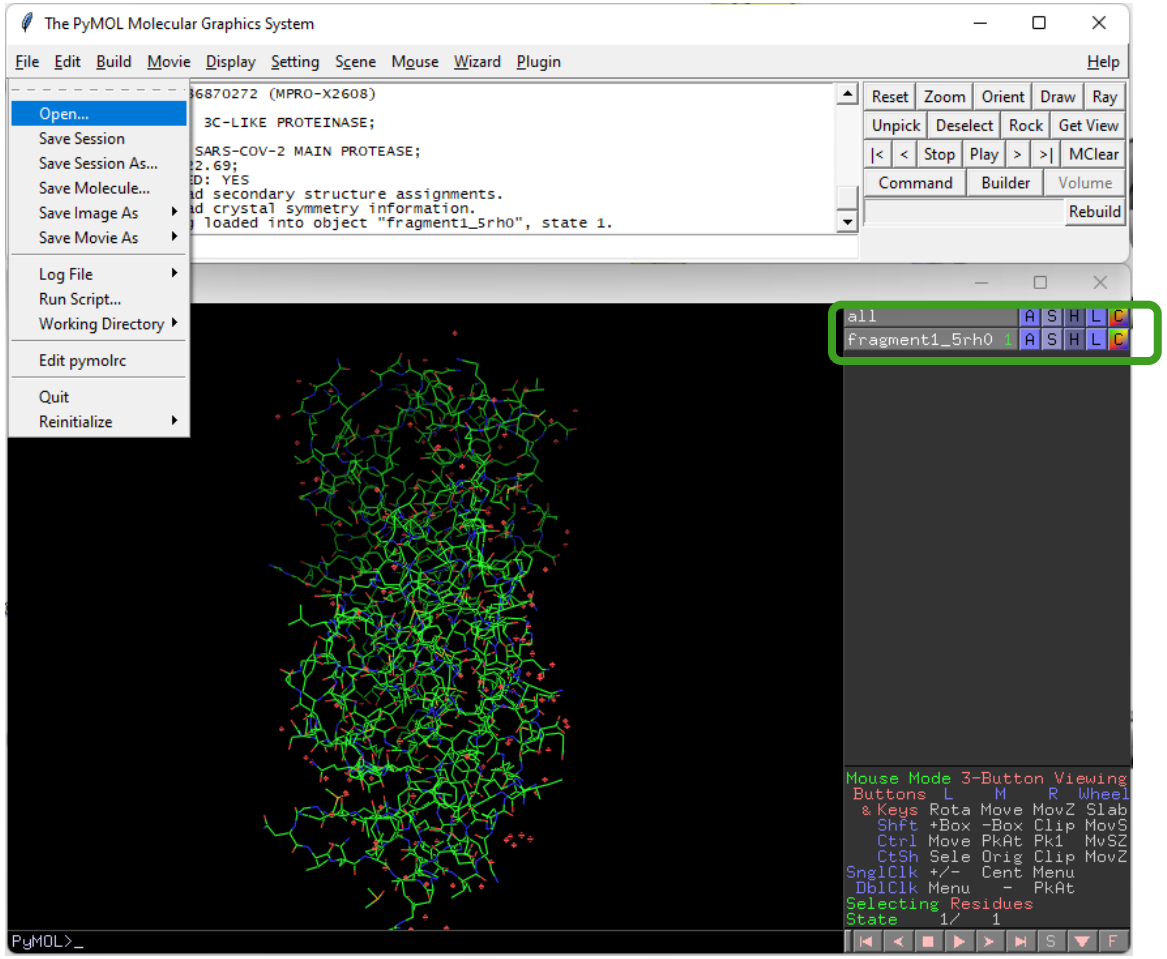
Literature:

- PyMol website

Mini intro PyMol 1

PyMOL is a user-sponsored molecular visualization system on an open-source foundation, maintained and distributed by Schrödinger.

- PyMol has 2 windows (viewer & commands menu)
- File → Open… and select your .pdb file
- A new object appears on the object panel
- Switch the object on and off in the viewer by clicking on it in the object panel
- Click on the “S” (show) next to the object to change visualization styles (lines, sticks, cartoon, surface…)
- Click on the “H” (hide) to remove visualization styles
- Click on the “C” (color) to change coloring. If you select “by element” only carbons will be colored differently.



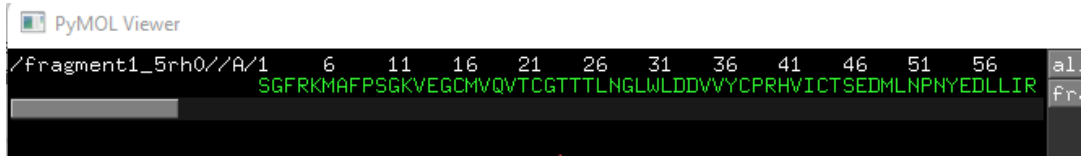
Literature:

- PyMol website

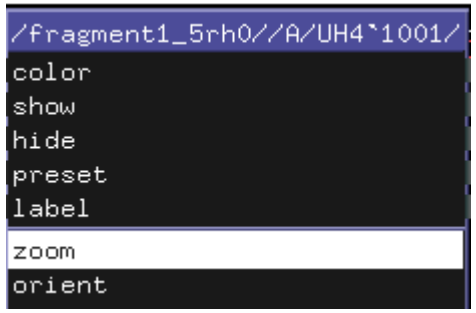
Mini intro PyMol 2

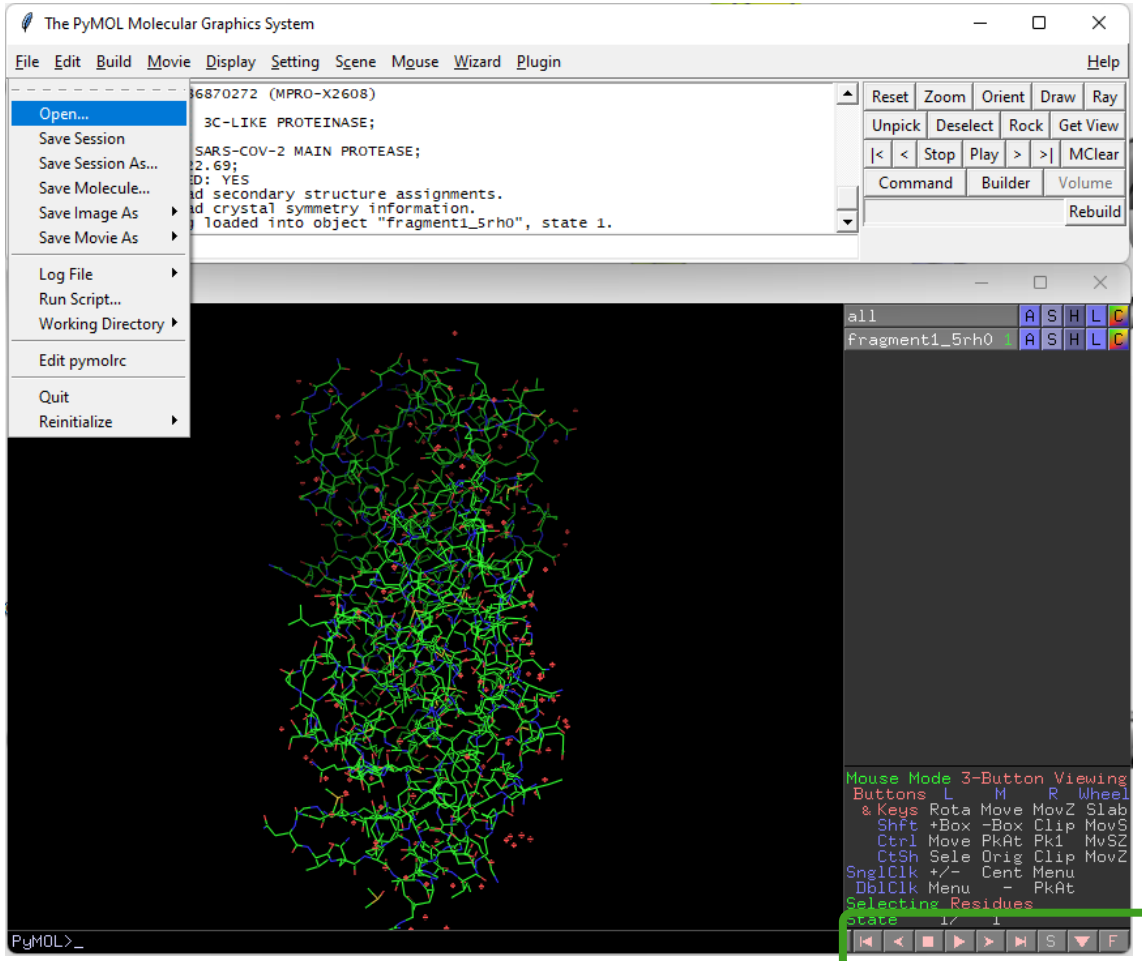
PyMOL is a user-sponsored molecular visualization system on an open-source foundation, maintained and distributed by Schrödinger.

- Zoom: Hold the right mouse button and move
- Mouse wheel: change the depths of your view
- Rotate: Hold the left mouse button and move
- Jump: middle-button-click on atom
- “S” on the far bottom right in the viewer window displays the sequence of your visible objects. It will be shown in the top of the viewer window:



scroll it to the right, right-click on the ligand, “zoom” to the ligand





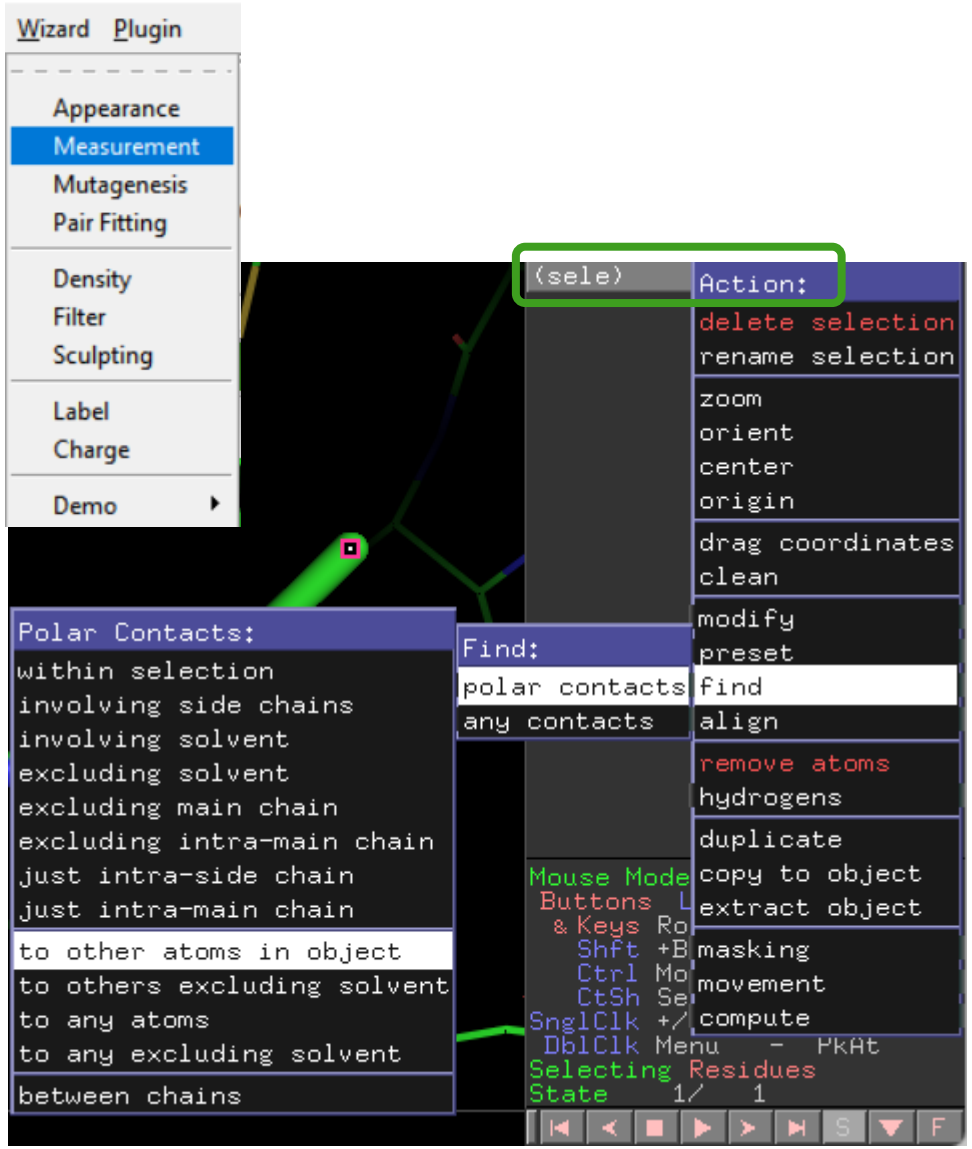
Literature:

- PyMol website

Mini intro PyMol 3

PyMOL is a user-sponsored molecular visualization system on an open-source foundation, maintained and distributed by Schrödinger.

- Measure distances between 2 atoms:
  - Wizard → Measurement
  - Left-click on 2 atoms
  - A dashed line object labeled with the distance is created
- Find polar contacts:
  - left-click on ligand (it will be selected: "(sele)"-object appears on the right)
  - Click on "A" next to the sele-object → find → polar contacts → to other atoms in object: a new object with all dotted lines will be created. **Caveat**: it just considers distances not angles – not all of these dotted lines will be good H-bonds!



Literature:

- PyMol website

Mini intro PyMol 4

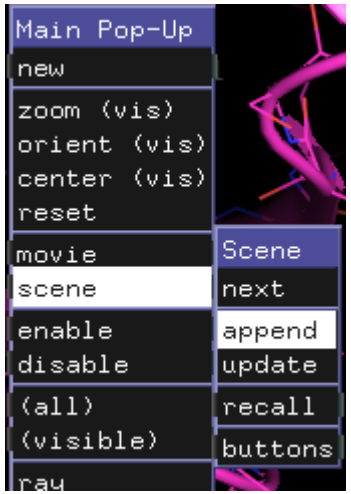
PyMOL is a user-sponsored molecular visualization system on an open-source foundation, maintained and distributed by Schrödinger.

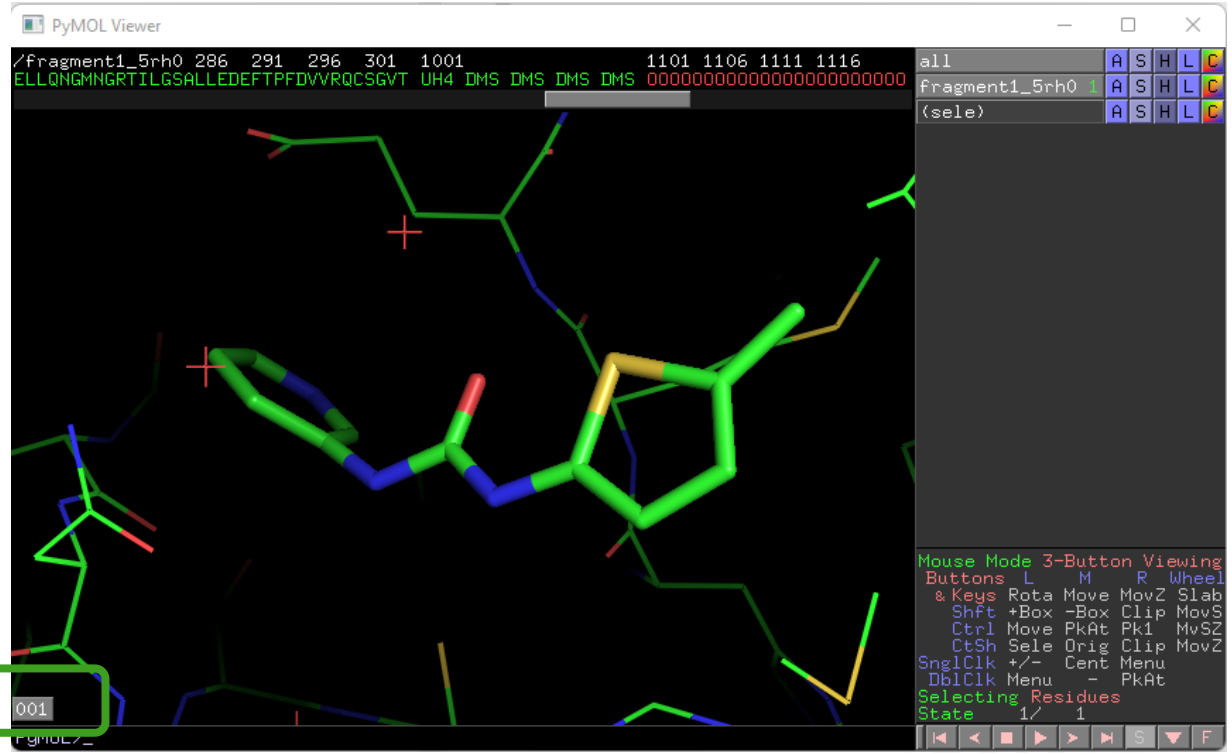
- Align two protein structures on their Cα:
  - Type “align” into the bottom line of the window:

```
PyMOL>align fragment2_Mpro-x0107, fragment1_5rh0_
```

```
align object-name, reference-object-name ↵
```

- Scenes: Save a particular view as a scene
  - Right click into the background → scene → append
  - A scene “001” will be saved in the bottom left of the viewer
  - Right click on it and rename or update it (in case of changes)
  - Left click on it to see your pre-defined view





Literature:

- PyMol website

Outline

- Part 1: *(2:30 h)*
  - General introduction:
    Most common Protein Ligand interactions
    small molecule conformational preferences
    electron densities
  - Introduction to the target Sars-Cov2 main protease MPro
  - Introduction to PyMol and coot
  - Exercise 1.1-5 and Discussions
- Part 2: *(1:30h)*
  - General introduction: Halogen bonds
  - Exercise 2.1
  - Halogen bond example from Idorsia, other interactions
  - Exercise 2.2

idorsia