HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

Interpretation of structural models

CRC/RTG Spring School – X-Ray Crystallography

Alexander Minges

Institute of Biochemical Plant Physiology

From data to a structural model

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

Where to get the slides

`http://www.biochemplant.hhu.de`

Will be uploaded today evening!

From data to a structural model

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

Software used

PyMOL `https://pymol.org`
`https://pymolwiki.org/index.php/Windows_Install`

Coot `https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot`

From data to a structural model

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

Where we left off before lunch

- Diffraction experiments yield data in form of spot intensities

From data to a structural model

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

Where we left off before lunch

- Diffraction experiments yield data in form of spot intensities
- Intensity of each reflection is proportional to the squared amplitude of a corresponding structure factor

From data to a structural model

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

Where we left off before lunch

- Diffraction experiments yield data in form of spot intensities
- Intensity of each reflection is proportional to the squared amplitude of a corresponding structure factor
- Structure factors contain information of both the phase and amplitude of a wave

From data to a structural model

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

Where we left off before lunch

- Diffraction experiments yield data in form of spot intensities
- Intensity of each reflection is proportional to the squared amplitude of a corresponding structure factor
- Structure factors contain information of both the phase and amplitude of a wave
- Phase problem may be solved using various techniques (e.g. molecular replacement)

From data to a structural model

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

Where we left off before lunch

- Diffraction experiments yield data in form of spot intensities
- Intensity of each reflection is proportional to the squared amplitude of a corresponding structure factor
- Structure factors contain information of both the phase and amplitude of a wave
- Phase problem may be solved using various techniques (e.g. molecular replacement)

Structure factor amplitudes and phase information can be combined to reconstruct the electron density.

From data to a structural model

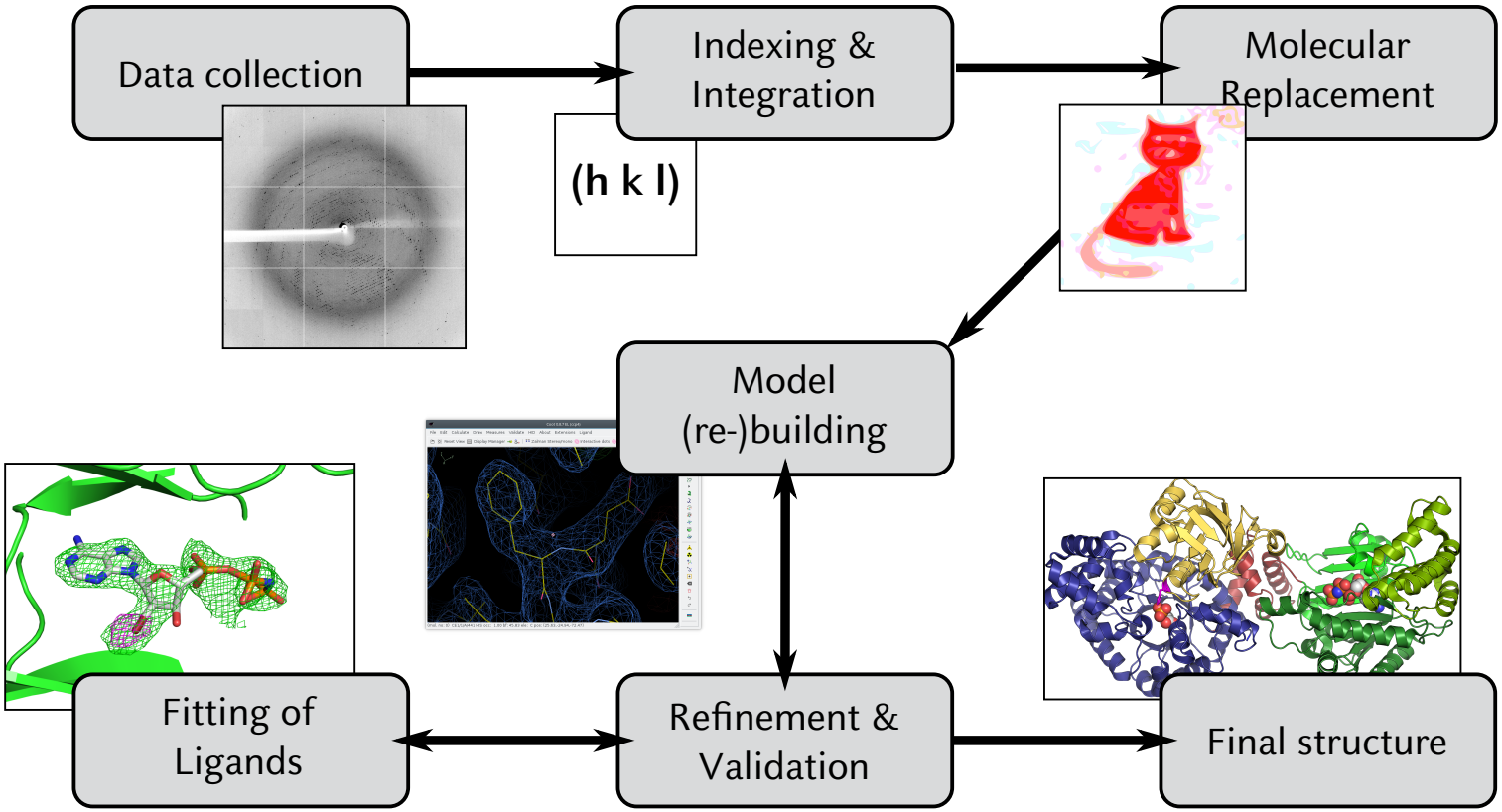


Figure: Scheme of X-Ray structure determination

Where to find structural data?

Where to find structural data?

RCSB PDB
PROTEIN DATA BANK

https://www.rcsb.org

PDBe
Protein Data Bank in Europe

https://www.pdbe.org

The Protein Data Bank (PDB)

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

What is the PDB?

- A database/collection of structural models…
- …from various sources (e.g. X-Ray, NMR, Cryo-EM)…
- …and associated experimental data.

RCSB PDB
PROTEIN DATA BANK

https://www.rcsb.org

PDBe
Protein Data Bank in Europe

https://www.pdbe.org

The Protein Data Bank (PDB)

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

What is the PDB?

- A database/collection of structural models...
- ...from various sources (e.g. X-Ray, NMR, Cryo-EM)...
- ...and associated experimental data.

RCSB PDB
PROTEIN DATA BANK

https://www.rcsb.org

PDBe
Protein Data Bank in Europe

https://www.pdbe.org

The Protein Data Bank (PDB)

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

What is the PDB?

- A database/collection of structural models...
- ...from various sources (e.g. X-Ray, NMR, Cryo-EM)...
- ...and associated experimental data.

RCSB PDB
PROTEIN DATA BANK

https://www.rcsb.org

PDBe
Protein Data Bank in Europe

https://www.pdbe.org

The Protein Data Bank (PDB)

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

What is the PDB?

- A database/collection of structural models...
- ...from various sources (e.g. X-Ray, NMR, Cryo-EM)...
- ...and associated experimental data.

RCSB PDB
PROTEIN DATA BANK

https://www.rcsb.org

PDBe
Protein Data Bank in Europe

https://www.pdbe.org

The Protein Data Bank (PDB)

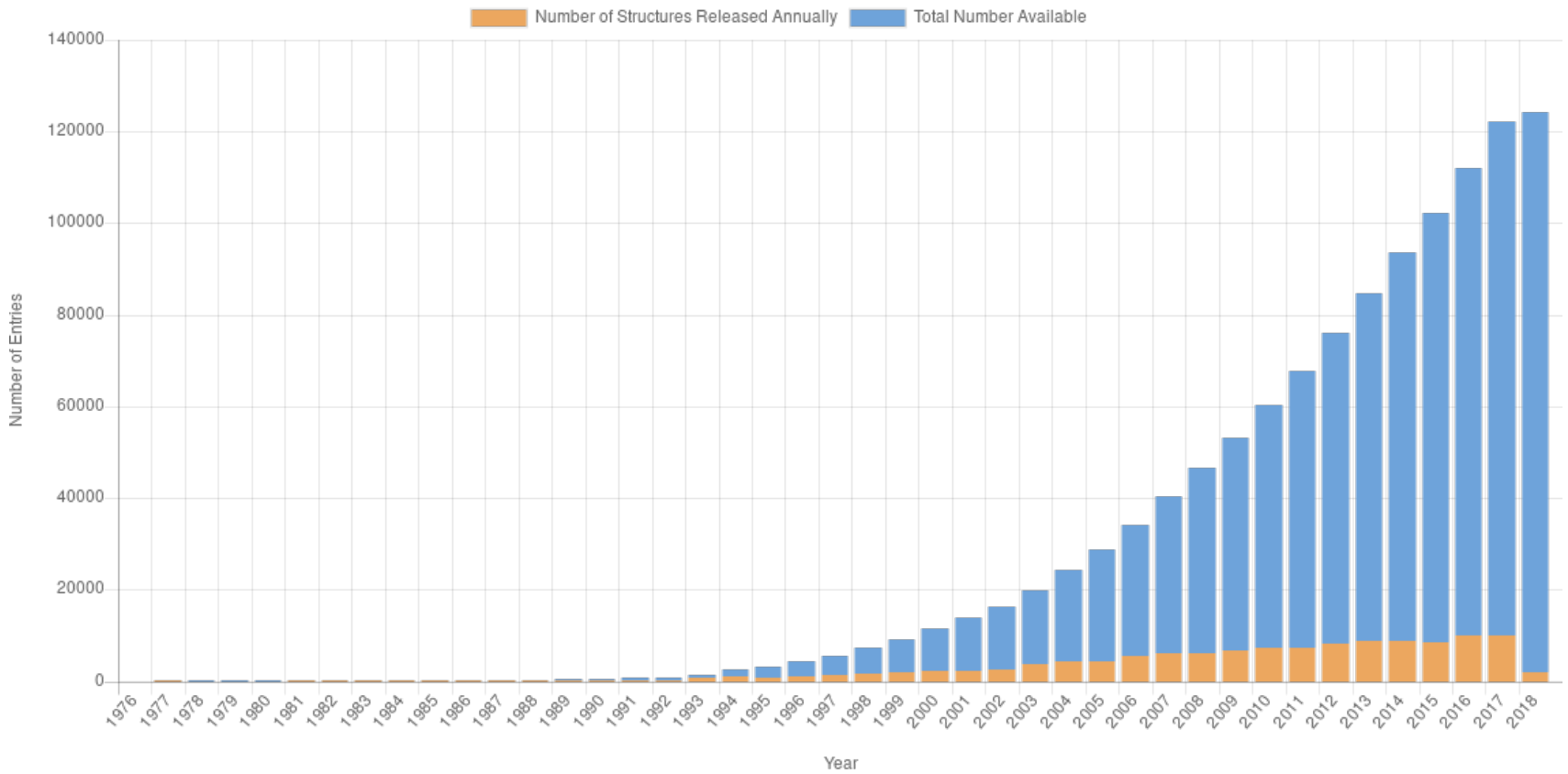


Figure: Growth of PDB (only X-Ray)

https://www.rcsb.org/stats

Structure of PDB files

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

Classic PDB files are plain text files which are human-readable!

Header

Contains information about:

- authors/publication
- experimental procedure
- data collection
- refinement and model statistics
- biological assembly
- primary/secondary structure
- missing residues/atoms
- ...

Structure of PDB files

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

Header

```
HEADER    HYDROLASE                                17-JAN-16   5HMV
TITLE     RE REFINEMENT OF 4MWK.
COMPND    MOL_ID: 1;
COMPND   2 MOLECULE: LYSOZYME C;
COMPND   3 CHAIN: A;
COMPND   4 SYNONYM: 1,4-BETA-N-ACETYLMURAMIDASE C,ALLERGEN GAL D IV;
COMPND   5 EC: 3.2.1.17
SOURCE    MOL_ID: 1;
SOURCE   2 ORGANISM_SCIENTIFIC: GALLUS GALLUS;
SOURCE   3 ORGANISM_COMMON: CHICKEN;
SOURCE   4 ORGANISM_TAXID: 9031;
SOURCE   5 TISSUE: EGG WHITE
KEYWDS    STRUCTURAL DYNAMICS CISPLATIN HISTIDINE PROTEIN, HYDROLASE
EXPDTA    X-RAY DIFFRACTION
AUTHOR    J.R.HELLIWELL
REVDAT   7   13-SEP-17 5HMV    1       REMARK
REVDAT   6   17-AUG-16 5HMV    1
REVDAT   5   10-AUG-16 5HMV    1       JRNL
```

Structure of PDB files

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

Body

`ATOM`	protein
`HETATM`	small compounds, ions, solvent, ...
`CONNECT`	optional; bonds between atoms
`ANISOU`	optional; anisotropic B-factors

Structure of PDB files

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

Body

```
ATOM      1  N   LYS A   1      -3.371 -10.156  -8.720  1.00 15.86           N
ATOM      2  CA  LYS A   1      -2.483 -10.521  -9.827  1.00 16.67           C
ATOM      3  C   LYS A   1      -2.518 -12.032 -10.037  1.00 19.82           C
ATOM      4  O   LYS A   1      -2.506 -12.813  -9.073  1.00 15.15           O
ATOM      5  CB  LYS A   1      -1.053 -10.037  -9.545  1.00 17.67           C
ATOM      6  CG  LYS A   1       0.017 -10.569 -10.498  1.00 18.43           C
ATOM      7  CD  LYS A   1       1.392 -10.047 -10.097  1.00 22.48           C
ATOM      8  CE  LYS A   1       2.517 -10.653 -10.939  1.00 22.32           C
ATOM      9  NZ  LYS A   1       2.338 -10.354 -12.381  1.00 30.88           N
ATOM     10  N   VAL A   2      -2.600 -12.449 -11.295  1.00 19.96           N
ATOM     11  CA  VAL A   2      -2.494 -13.860 -11.611  1.00 17.16           C
ATOM     12  C   VAL A   2      -1.067 -14.107 -12.085  1.00 21.51           C
ATOM     13  O   VAL A   2      -0.624 -13.516 -13.071  1.00 24.51           O
ATOM     14  CB  VAL A   2      -3.514 -14.293 -12.665  1.00 16.56           C
ATOM     15  CG1 VAL A   2      -3.329 -15.760 -13.007  1.00 18.51           C
ATOM     16  CG2 VAL A   2      -4.938 -14.050 -12.163  1.00 17.76           C
```

Excercise – Spot the difference

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

Finding the right structural model

Compare the given models of chicken egg lysozyme. Imagine you want to predict ligand interactions and need precise information on e.g. side chain conformations. Which model suits your needs best?

Excercise – Spot the difference

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

Finding the right structural model

Compare the given models of chicken egg lysozyme. Imagine you want to predict ligand interactions and need precise information on e.g. side chain conformations. Which model suits your needs best?

1. **Launch *PyMOL***

Excercise – Spot the difference

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

Finding the right structural model

Compare the given models of chicken egg lysozyme. Imagine you want to predict ligand interactions and need precise information on e.g. side chain conformations. Which model suits your needs best?

1. **Launch *PyMOL***
2. **Load both models:**

```
fetch 5hmv
fetch 4gcf
```

Excercise – Spot the difference

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

Finding the right structural model

Compare the given models of chicken egg lysozyme. Imagine you want to predict ligand interactions and need precise information on e.g. side chain conformations. Which model suits your needs best?

1. **Launch *PyMOL***
2. **Load both models:**
   ```
   fetch 5hmv
   fetch 4gcf
   ```
3. **Align both models:**
   ```
   cealign 5hmv, 4gcf
   ```

Excercise – Spot the difference

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

Finding the right structural model

Compare the given models of chicken egg lysozyme. Imagine you want to predict ligand interactions and need precise information on e.g. side chain conformations. Which model suits your needs best?

1. **Launch *PyMOL***
2. **Load both models:**
   ```
   fetch 5hmv
   fetch 4gcf
   ```
3. **Align both models:**
   ```
   cealign 5hmv, 4gcf
   ```
4. **Visualize as cartoon:**
   ```
   hide everything
   show cartoon
   ```

Excercise – Spot the difference

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

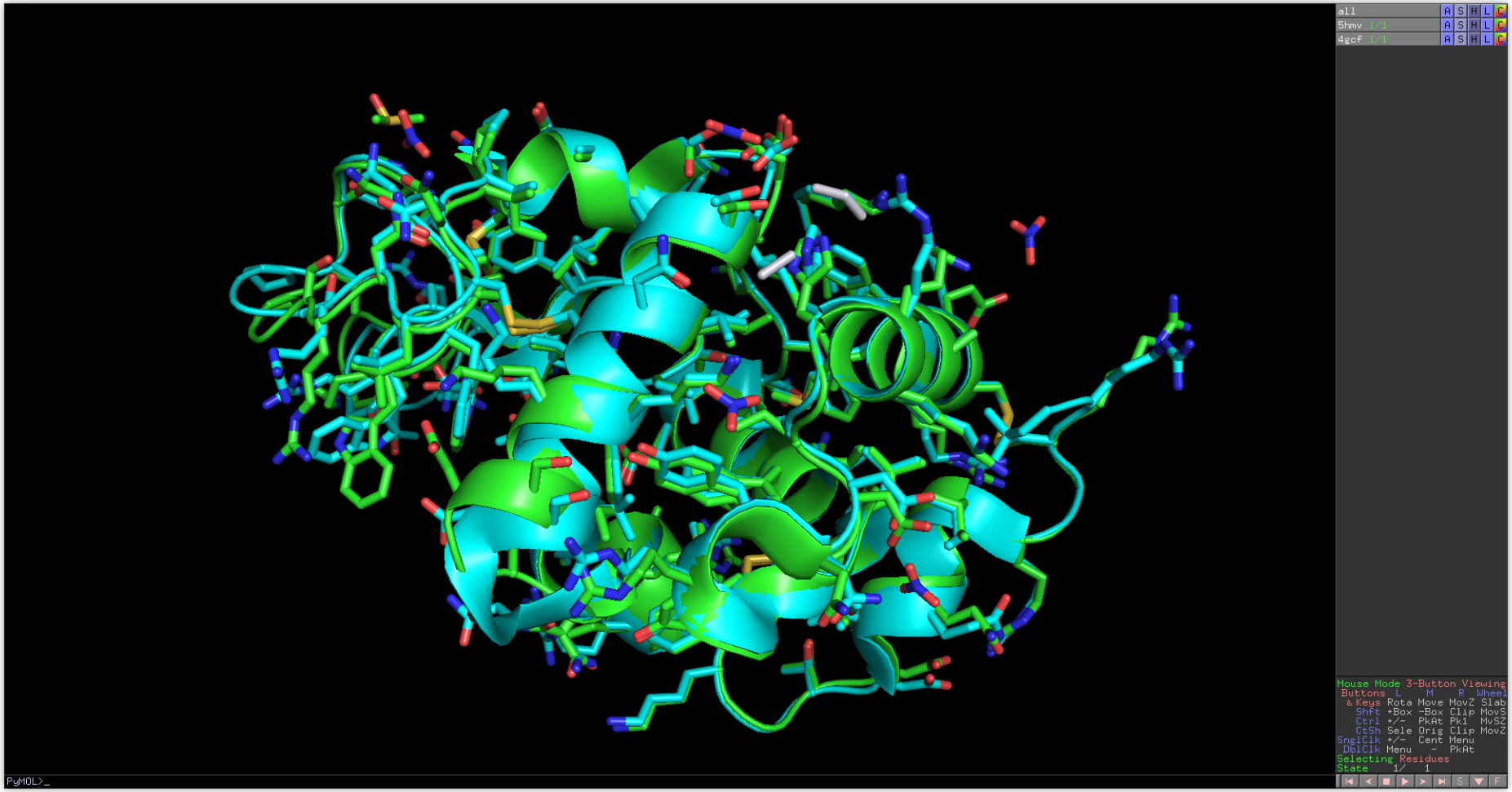
Finding the right structural model

Compare the given models of chicken egg lysozyme. Imagine you want to predict ligand interactions and need precise information on e.g. side chain conformations. Which model suits your needs best?

1. **Launch *PyMOL***
2. **Load both models:**
   ```
   fetch 5hmv
   fetch 4gcf
   ```
3. **Align both models:**
   ```
   cealign 5hmv, 4gcf
   ```
4. **Visualize as cartoon:**
   ```
   hide everything
   show cartoon
   ```
5. **Show side chains:**
   ```
   set cartoon_side_chain_helper, 1
   show sticks
   ```

Excercise – Spot the difference

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF



Excercise – Spot the difference

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

Finding the right structural model

Compare the given models of chicken egg lysozyme. Imagine you want to predict ligand interactions and need precise information on e.g. side chain conformations. Which model suits your needs best?

- Only subtile differences in side chain conformations
- Overall fold virtually identical

Excercise – Spot the difference

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

Finding the right structural model

Compare the given models of chicken egg lysozyme. Imagine you want to predict ligand interactions and need precise information on e.g. side chain conformations. Which model suits your needs best?

- Only subtile differences in side chain conformations
- Overall fold virtually identical

Both models are equally good!

Excercise – Spot the difference

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

Finding the right structural model

Compare the given models of chicken egg lysozyme. Imagine you want to predict ligand interactions and need precise information on e.g. side chain conformations. Which model suits your needs best?

- Only subtile differences in side chain conformations
- Overall fold virtually identical

Really?

Excercise – Spot the difference

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

Finding the right structural model

Compare the given models of chicken egg lysozyme. Imagine you want to predict ligand interactions and need precise information on e.g. side chain conformations. Which model suits your needs best?

- Only subtile differences in side chain conformations
- Overall fold virtually identical

Resolution is the key!

Resolution of data

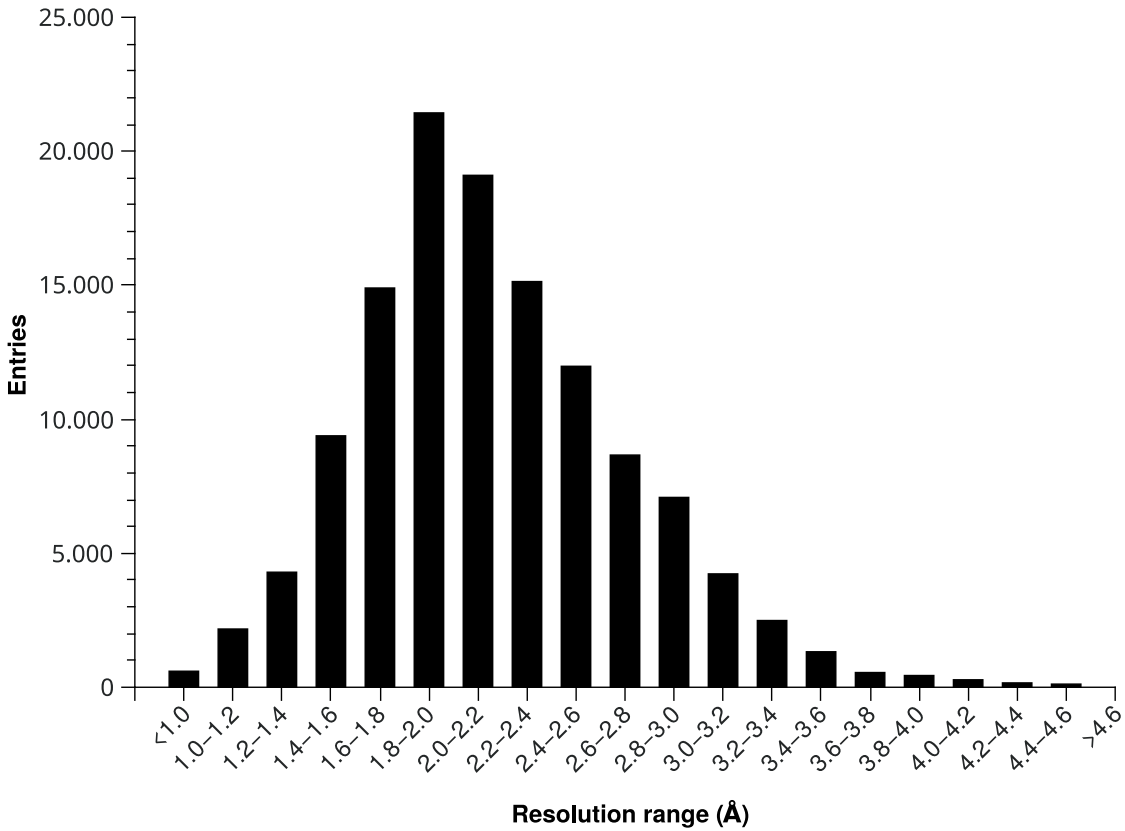


Figure: Resolution of PDB entries (March 2018)

https://www.rcsb.org/stats/distribution_resolution

Resolution of data

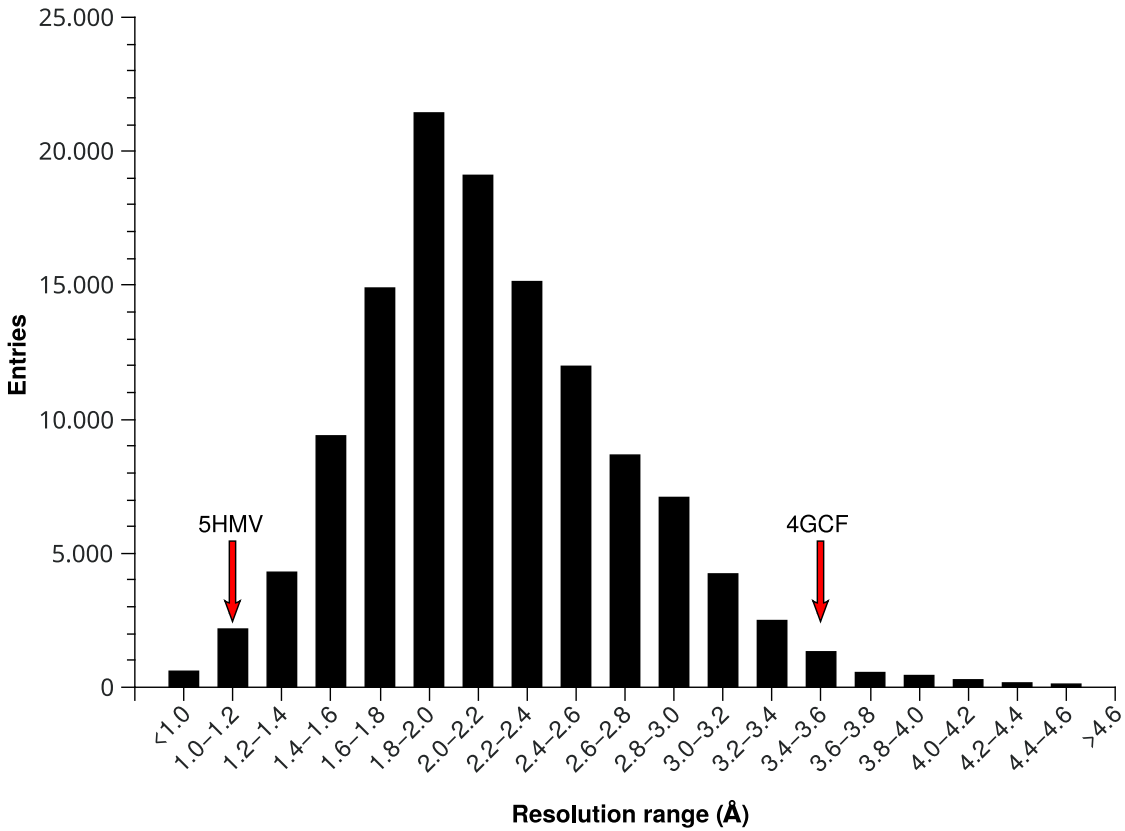


Figure: Resolution of PDB entries (March 2018)

https://www.rcsb.org/stats/distribution_resolution

Resolution of data

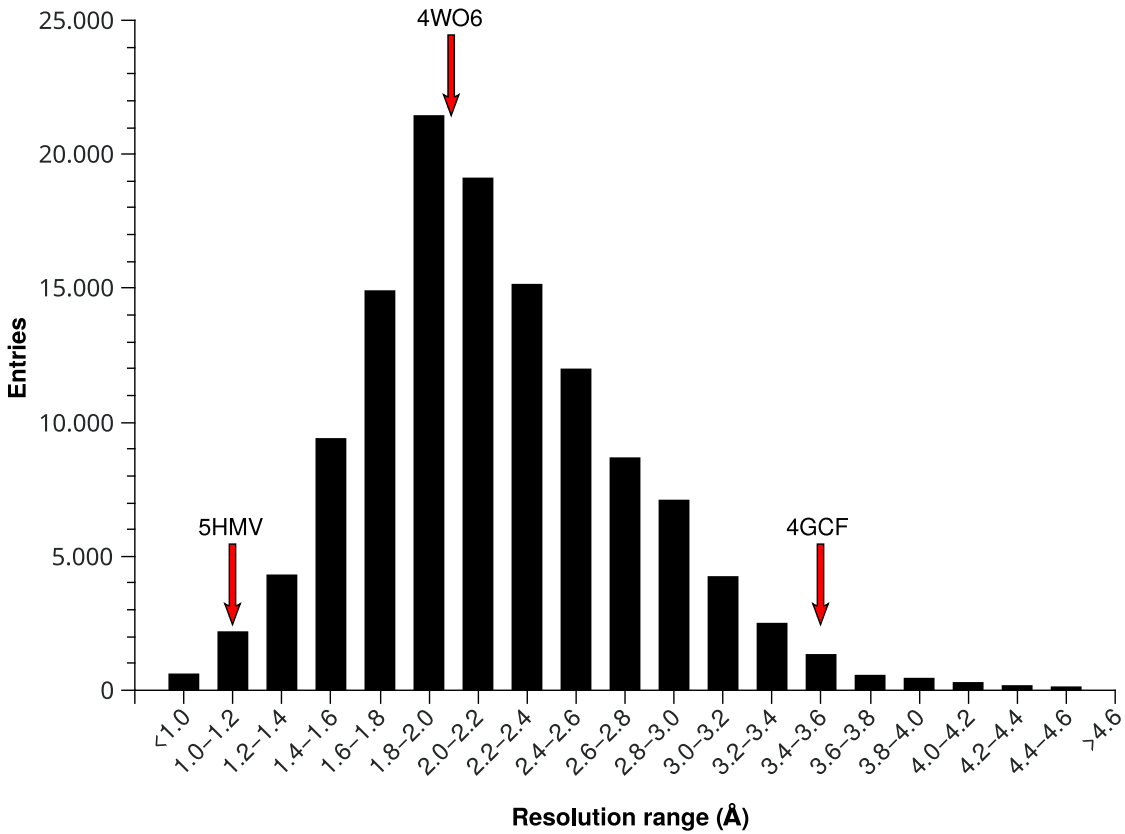


Figure: Resolution of PDB entries (March 2018)

https://www.rcsb.org/stats/distribution_resolution

Resolution of data

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

- Higher resolution: More observations (data-to-parameter ratio)
- More observations: More accurate atomic positions
- Sharper and cleaner ED maps
- Easier interpretation and more reliable model building

Electron density

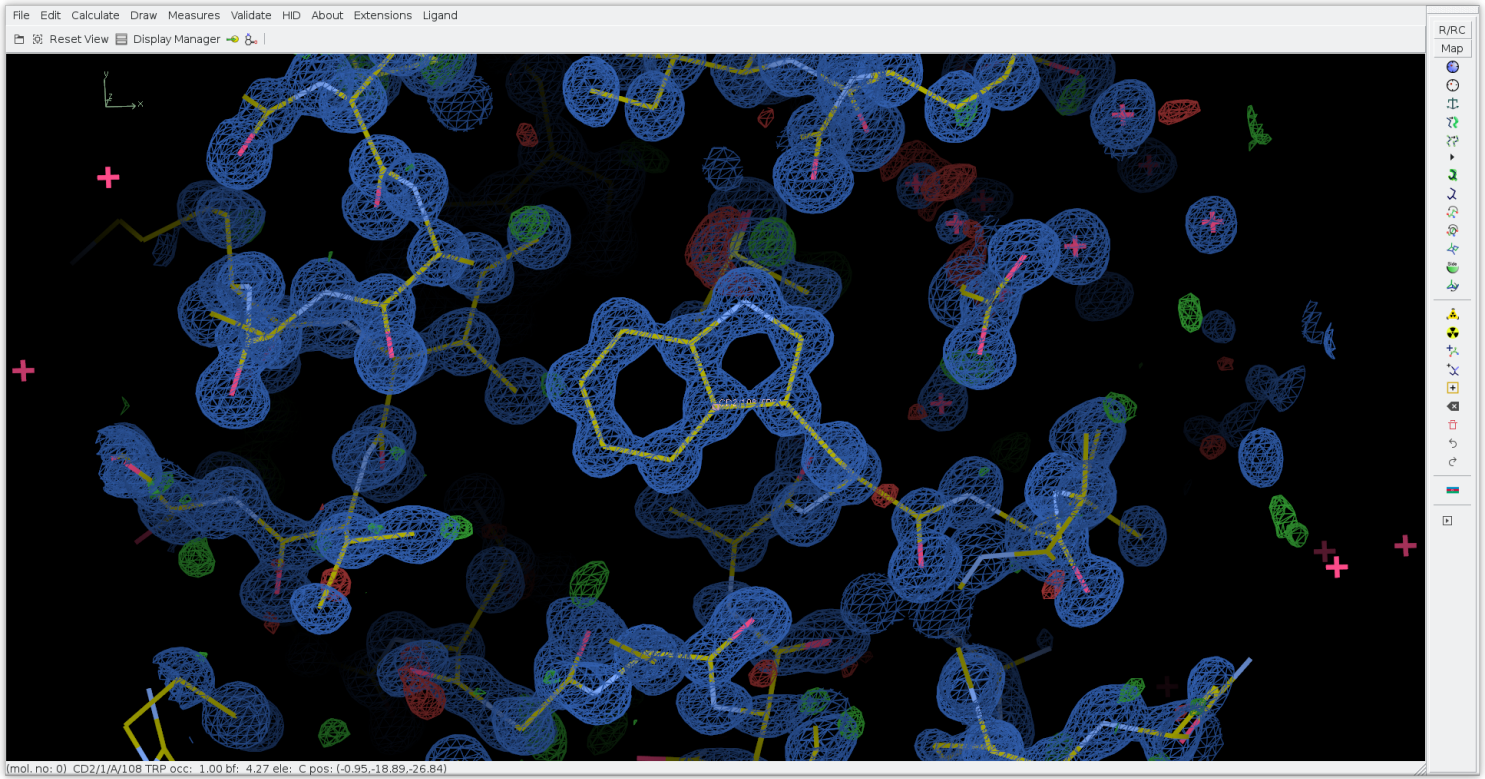


Figure: Electron density map of PDB 5HMV (0.98 Å, 1.5 σ)

Electron density

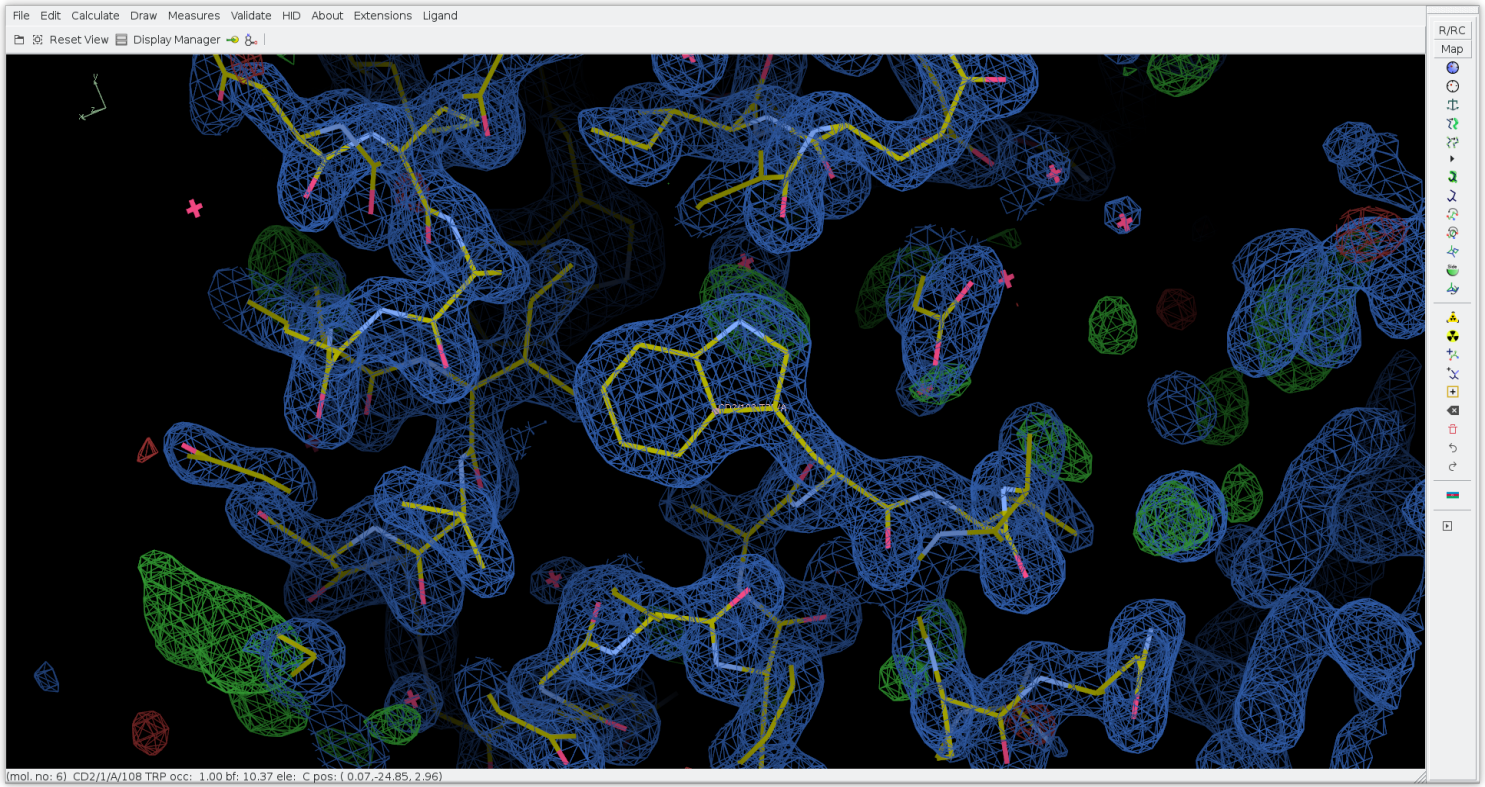


Figure: Electron density map of PDB 4WO6 (2.0 Å, 1.5 σ)

Electron density

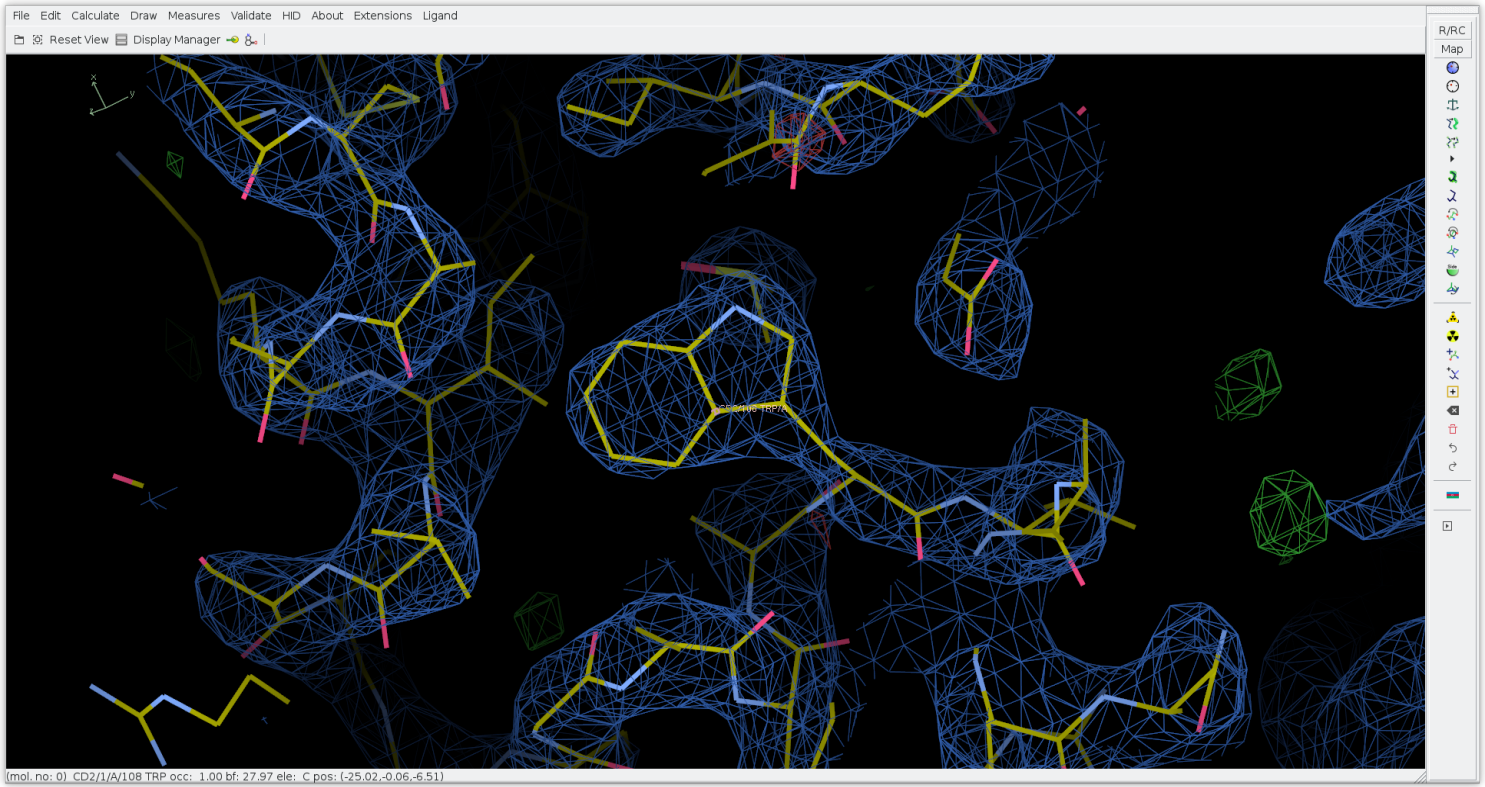
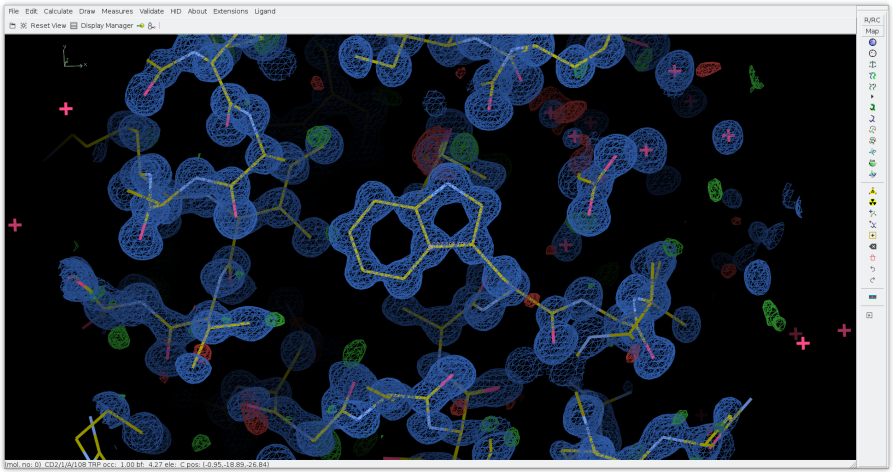


Figure: Electron density map of PDB 4GFC (3.5 Å, 1.5 σ)

Electron density

- Density around model
- Missing/excess model



Excercise

Model building

We will complete a lysozyme structure that is missing a patch of 14 aa. Try to complete the structure. Do you recognize the secondary structure element these residues form?

Excercise

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

Model building

We will complete a lysozyme structure that is missing a patch of 14 aa. Try to complete the structure. Do you recognize the secondary structure element these residues form?

1. **Copy files to desktop:**
   `map.ccp4` and `seed.pdb`

Excercise

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

Model building

We will complete a lysozyme structure that is missing a patch of 14 aa. Try to complete the structure. Do you recognize the secondary structure element these residues form?

1. **Copy files to desktop:**
   `map.ccp4` and `seed.pdb`
2. **Launch *Coot***

Excercise

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

Model building

We will complete a lysozyme structure that is missing a patch of 14 aa. Try to complete the structure. Do you recognize the secondary structure element these residues form?

1. **Copy files to desktop:**
   `map.ccp4` and `seed.pdb`
2. **Launch *Coot***
3. **Load model:**
   `File: Open Coordinates (seed.pdb)`

Excercise

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

Model building

We will complete a lysozyme structure that is missing a patch of 14 aa. Try to complete the structure. Do you recognize the secondary structure element these residues form?

1. **Copy files to desktop:**
   `map.ccp4` and `seed.pdb`
2. **Launch *Coot***
3. **Load model:**
   `File: Open Coordinates (seed.pdb)`
4. **Load map:**
   `File: Open Map (map.ccp4)`

Quality of X-Ray data

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

Quality indicators of structural models

- Resolution
- Completeness/Redundancy
- Data merging statistics ($R_{\text{merge/meas/pim}}$)
- Fit of model and data (R/R_{free}, $RSCC$)
- B-factors (atomic displacement parameters)
- Correct chemistry/geometry
  - Bond/torsion angles (e.g. Ramachandran plot)
  - Bond lengths
  - (Binding) geometry of ligands
  - Clashes

Quality of X-Ray data

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

Quality indicators of structural models

- Resolution
- Completeness/Redundancy
- Data merging statistics ($R_{merge/meas/pim}$)
- Fit of model and data (R/R_{free}, $RSCC$)
- B-factors (atomic displacement parameters)
- Correct chemistry/geometry
  - Bond/torsion angles (e.g. Ramachandran plot)
  - Bond lengths
  - (Binding) geometry of ligands
  - Clashes

Most of these can be found in *Table 1*!

Quality of X-Ray data

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

Read the paper!

Important aspects of the structure in question may be explained there...

Understanding *Table 1*

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

TABLE I. X-ray crystallographic data and protein model refinement statistics.

	Cisplatin	Cisplatin	Cisplatin	Carboplatin
Unit cell parameters (Å)/(deg)	a = 26.99	a = 26.77	A = 27.34	a = 26.96
	b = 31.81	b = 31.38	B = 32.13	b = 31.79
	c = 34.07	c = 33.86	C = 34.29	c = 34.05
	$\alpha = 89.08$	$\alpha = 88.90$	A = 88.04	$\alpha = 88.76$
	$\beta = 72.00$	$\beta = 72.31$	B = 71.17	$\beta = 71.99$
	$\gamma = 67.81$	$\gamma = 68.46$	$\gamma = 68.35$	$\gamma = 68.33$
PDB id's	4mwk[a]	4mwm[a]	4mwn[a]	4oxe[a]
Data collection temperature (K)	150	200	295	200
Crystal size (mm)	0.6	0.3	0.25	0.2
Total absorbed X-ray dose (MGy)[b]	0.31	0.37	0.48	0.31
Crystal growth time	5 weeks	8 days	6 weeks	11 days
Observed reflections	202732	118846	112029	156662
Unique reflections	51605	37288	19160	35817
Resolution (Å)	29.28–0.98 (1.02–0.98)	32.09–1.12 (1.15–1.12)	32.31–1.42 (1.51–1.42)	32.21–1.13 (1.16–1.13)
Completeness (%)	90.6 (51.7)	95.1 (84.6)	99.2 (98.4)	94.1 (77.0)
Rmerge (%)	0.045 (0.209)	0.087 (0.179)	0.145 (0.554)	0.086 (0.310)

Understanding *Table 1*

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

Mean I/sig(I)	15.6 (3.2)[c]	5.9 (2.1)	6.8 (1.1)[d]	7.2 (2.0)
Redundancy	3.5 (0.6)	2.2 (0.9)	5.7 (2.0)	2.9 (1.1)
Cruickshank diffraction precision index (Å) for coordinate error	0.022	0.037	0.084	0.046
Number of protein atoms	1007	998	992	998
Average B factor (Å^2) for protein atoms	7.2	11.4	14.7	9.4
Number of water molecules	142	94	41	98
Average B factor (Å^2) for water molecules	16.4	19.7	22.6	30.2
Number of cisplatin/carboplatin atoms	15	11	4	13
Average B factor (Å^2) for cisplatin and carboplatin atoms	33.5	38.8	27.1	36.3
Number of other bound atoms	85	72	28	54
Average B factor (Å^2) for other bound atoms	18.9	35.4	29.3	32.0
R factor/R free (%)	11.7/14.5	14.7/18.7	20.8/23.5	17.7/22.1
R factor all	11.9	14.9	21.0	17.9
root mean square deviation bond lengths (Å)/ Angles (deg)	0.039/2.907	0.023/2.323	0.021/2.247	0.031/2.592
Ramachandran favoured	97.5	97.5	96.0	95.8
Ramachandran allowed	2.5	2.5	3.2	4.2
Ramachandran disallowed	0	0	0.8[e]	0

Protein backbone angles

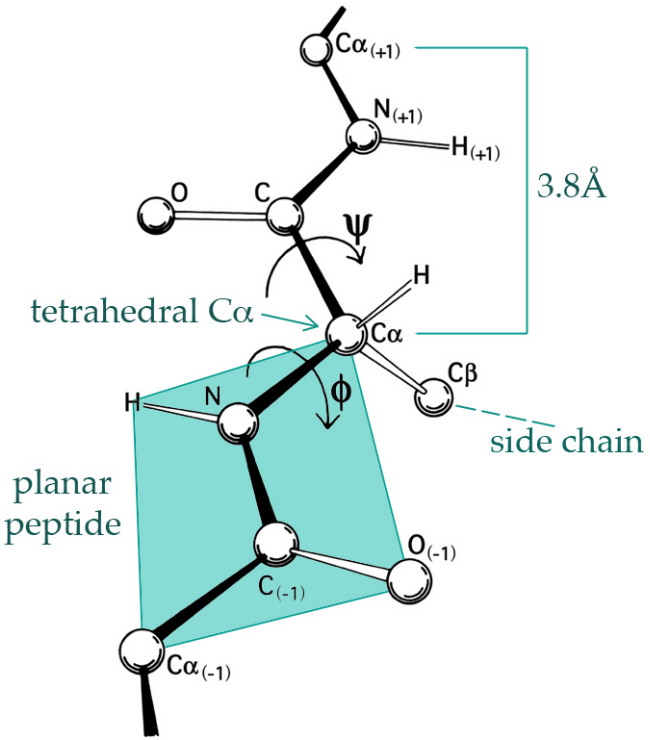


Figure: Protein backbone angles[1]

[1] Jane S. Richardson, adapted from *The Anatomy and Taxonomy of Protein Structure*

Protein backbone angles

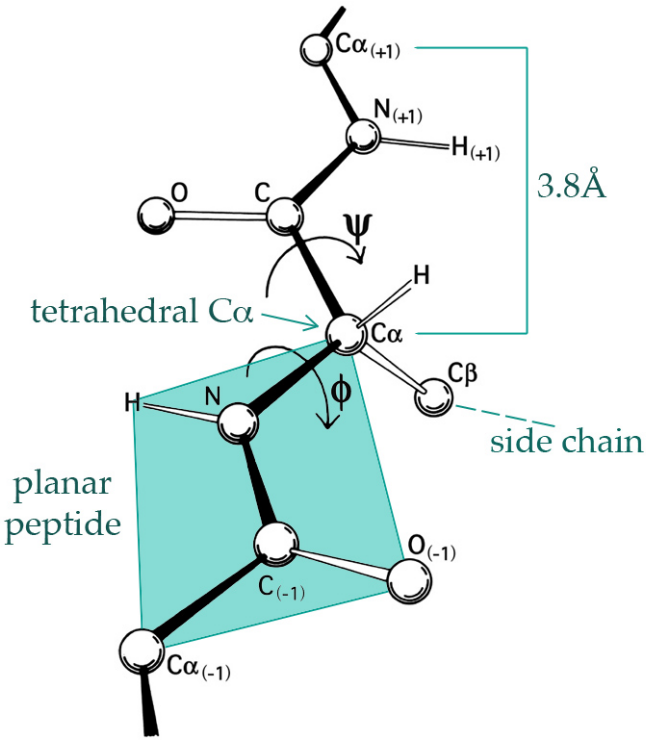


Figure: Protein backbone angles[1]

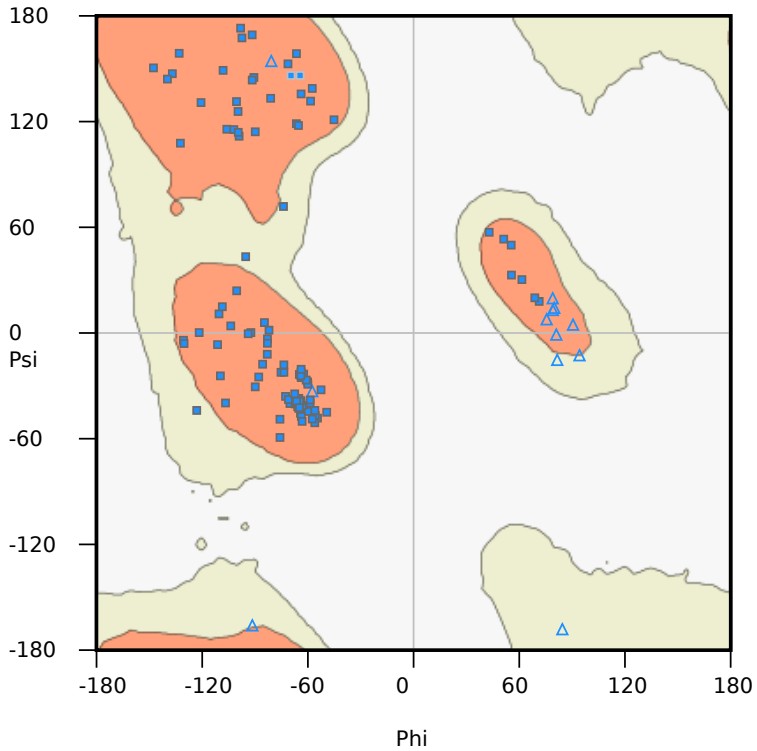


Figure: Ramachandran plot of PDB 5HMV

[1] CC BY Jane S. Richardson, adapted from *The Anatomy and Taxonomy of Protein Structure*

Bond lengths

Table: Common bond lengths found in protein structures

Peptide bond	Bond length (Å)	Single bond	Bond length (Å)
C_α – C	1.525 ± 0.026	C – C	1.540 ± 0.027
C – N	1.336 ± 0.023	C – N	1.489 ± 0.030
N – C_α	1.459 ± 0.020	C – O	1.420 ± 0.020
C = O	1.229 ± 0.019	C – S	1.807 ± 0.026
C_α – C_β	1.530 ± 0.020	S – S	2.033 ± 0.016

Bernhard Rupp (2010) *Biomolecular Crystallography* p. 629

Secondary structure elements

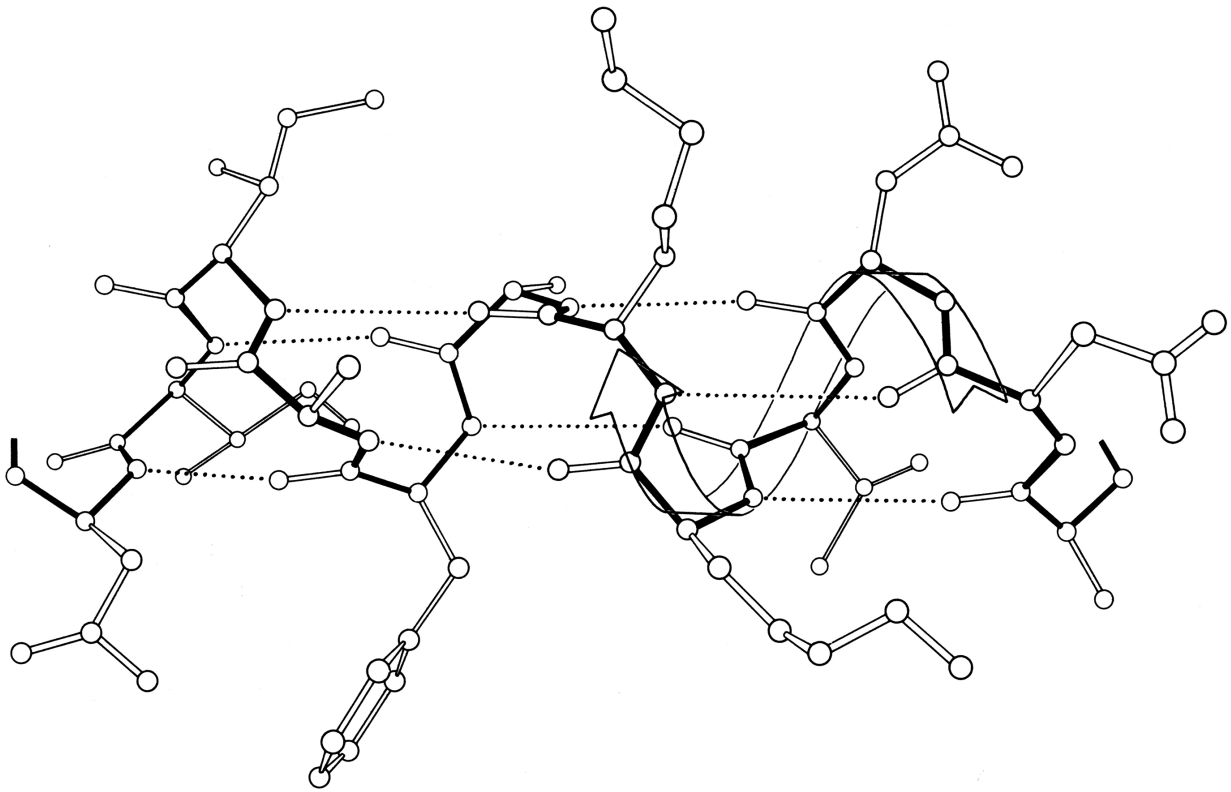


Figure: α-helix with backbone (black) and sidechains (white).

Jane S. Richardson (2007) *The Anatomy and Taxonomy of Protein Structure*

Secondary structure elements

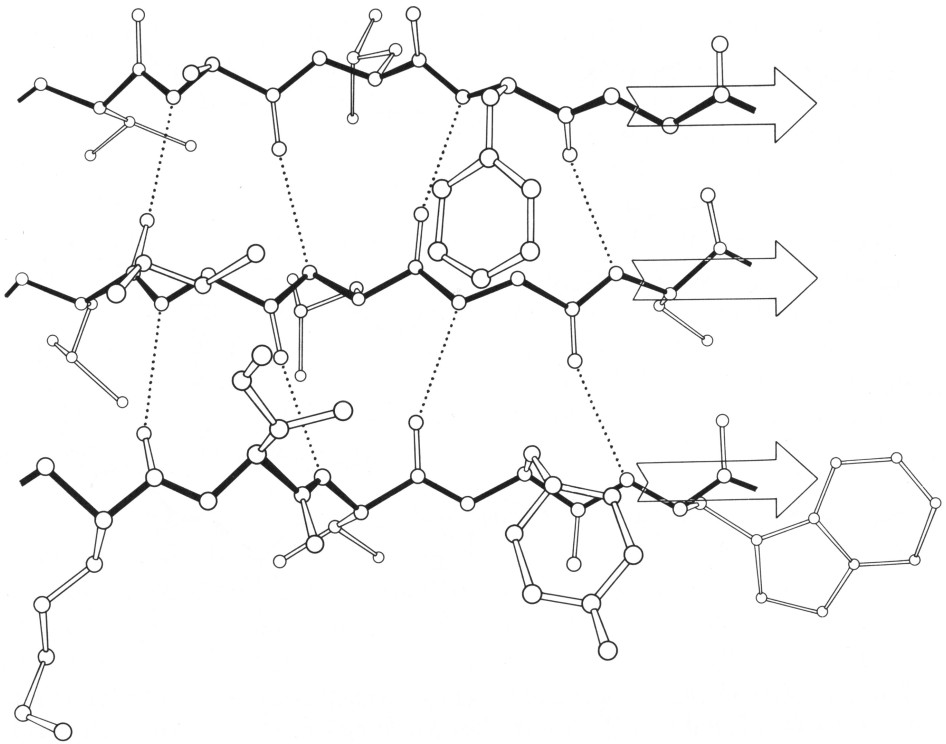


Figure: Parallel β-sheet with backbone (black) and sidechains (white).

Jane S. Richardson (2007) *The Anatomy and Taxonomy of Protein Structure*

Secondary structure elements

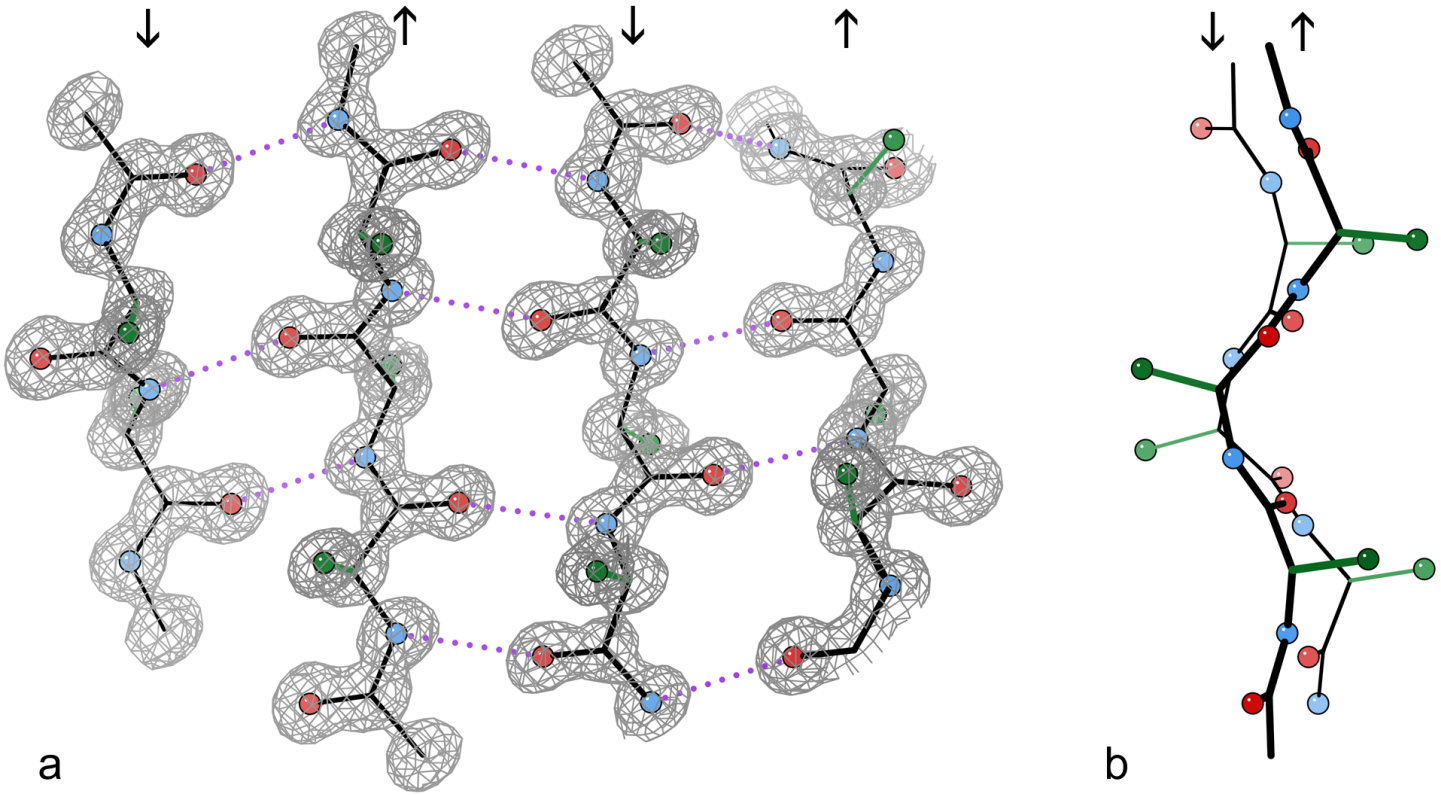


Figure: Antiparallel β-sheet in top (a) and side view (b).

B-factors

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

Definition

B-factors describe the uncertainty in atomic positions.

- For rigid parts usually below 20 Å²
- Dependent on the resolution (data-to-parameter ratio) different models are used:
  - isotropic model (low resolution)
  - isotropic + TLS (low to medium resolution)
  - anisotropic (high resolution)
- Effect in practice hard to distinguish from occupancy

B-factors

- Termini and loops naturally tend to have higher B-factors
- May hint to functional protein movements
- Atoms not supported by enough density will have extremely high B-factors

A word of warning

Don’t take the exact position of a certain feature (e.g. side chain) for granted!
Always check the corresponding B-factors!

B-factors

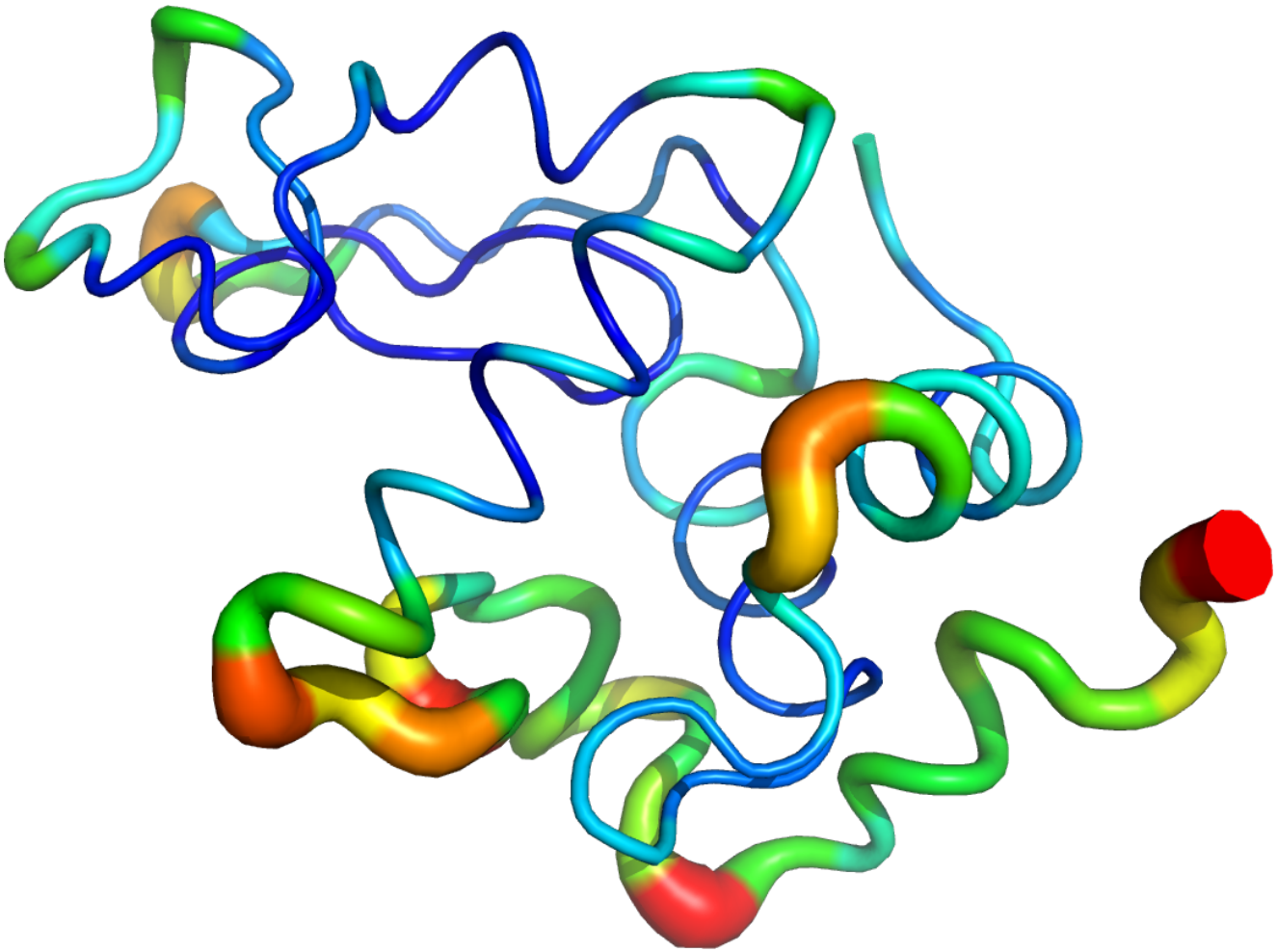


Figure: PDB 5HMV with B-factors represented by color and thickness.

B-factors

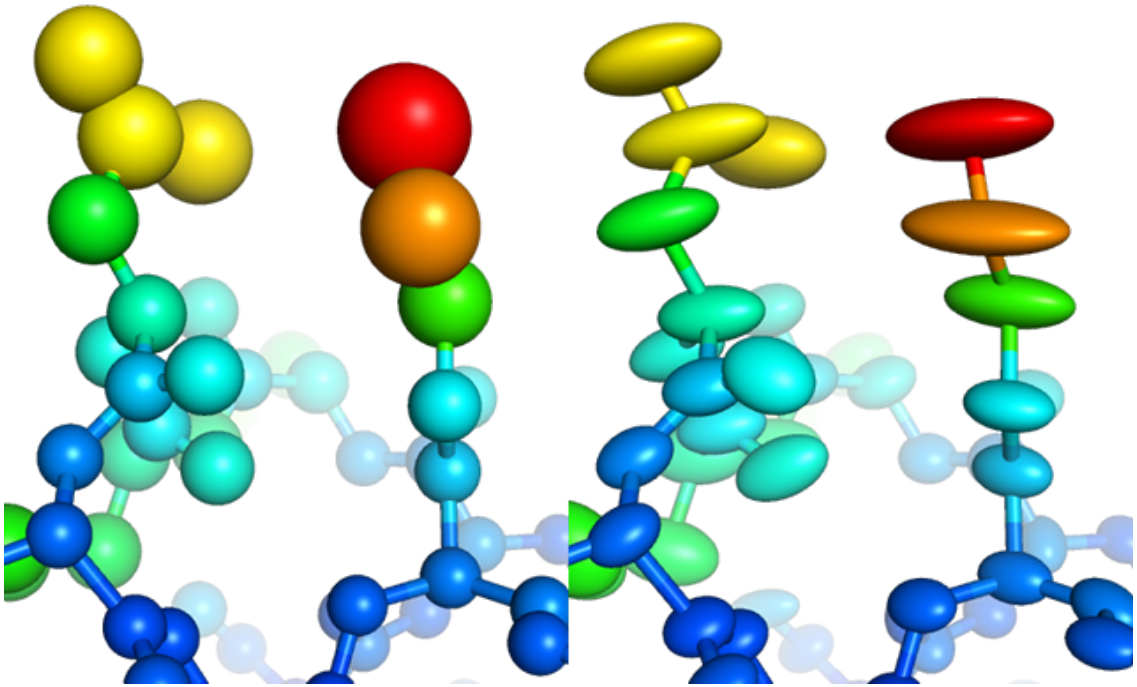


Figure: Example of isotropic (left) and anisotropic (right) B-factors.

`https://www.phenix-online.org/documentation/dictionary.html`

B-factors

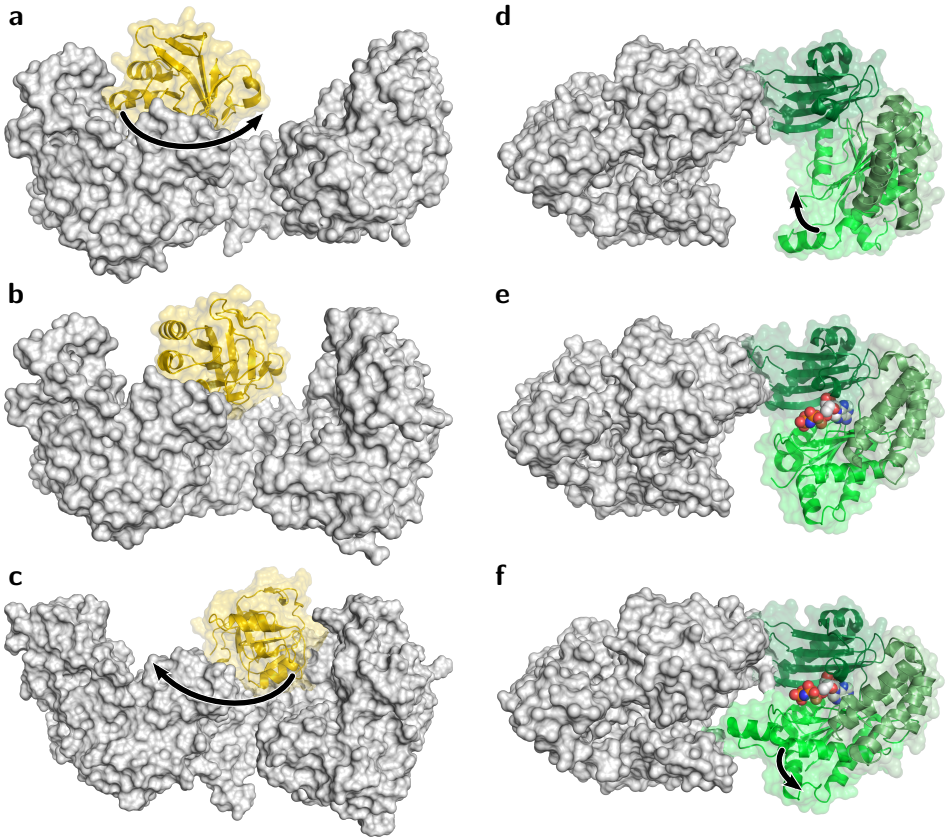


Figure: Dynamics of pyruvate phosphate dikinase (PPDK).

Minges et al. (2017) doi:10.1038/srep45389

B-factors

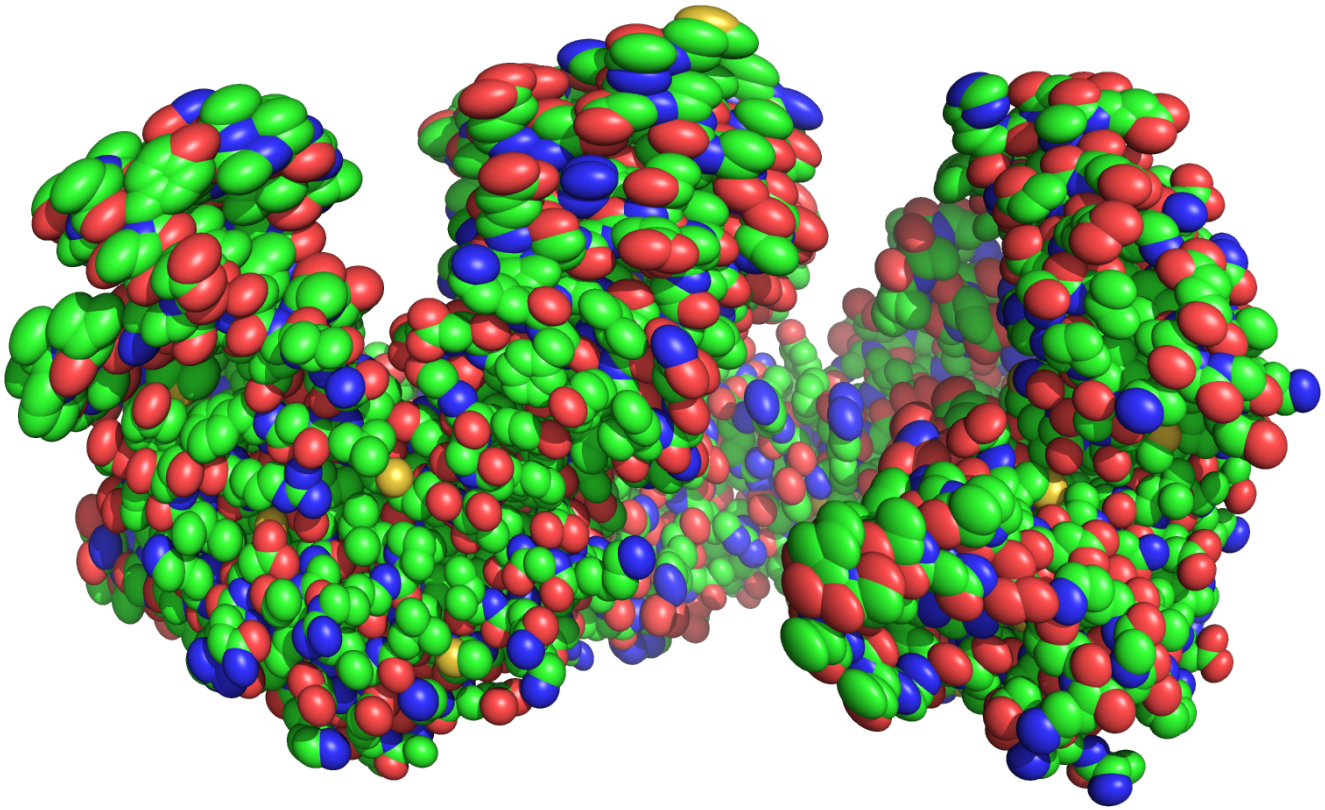


Figure: B-factors are related to protein dynamics (PDB 5JVN).

Minges et al. (2017) doi:10.1038/srep45389

Occupancy

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

Definition

Fraction of unit cells in which a specific atom occurs at a given position.

- Usually 1.0 (100 %) for protein backbone and most side chains (see alternate conformations)
- Lower values quite common for ligands
- In practice its effects are hard to tell apart from B-factor influence

Alternate conformations

- Usually 1.0 (100 %) for protein backbone and most side chains (see alternate conformations)
- Lower values quite common for ligands

Ligand geometry

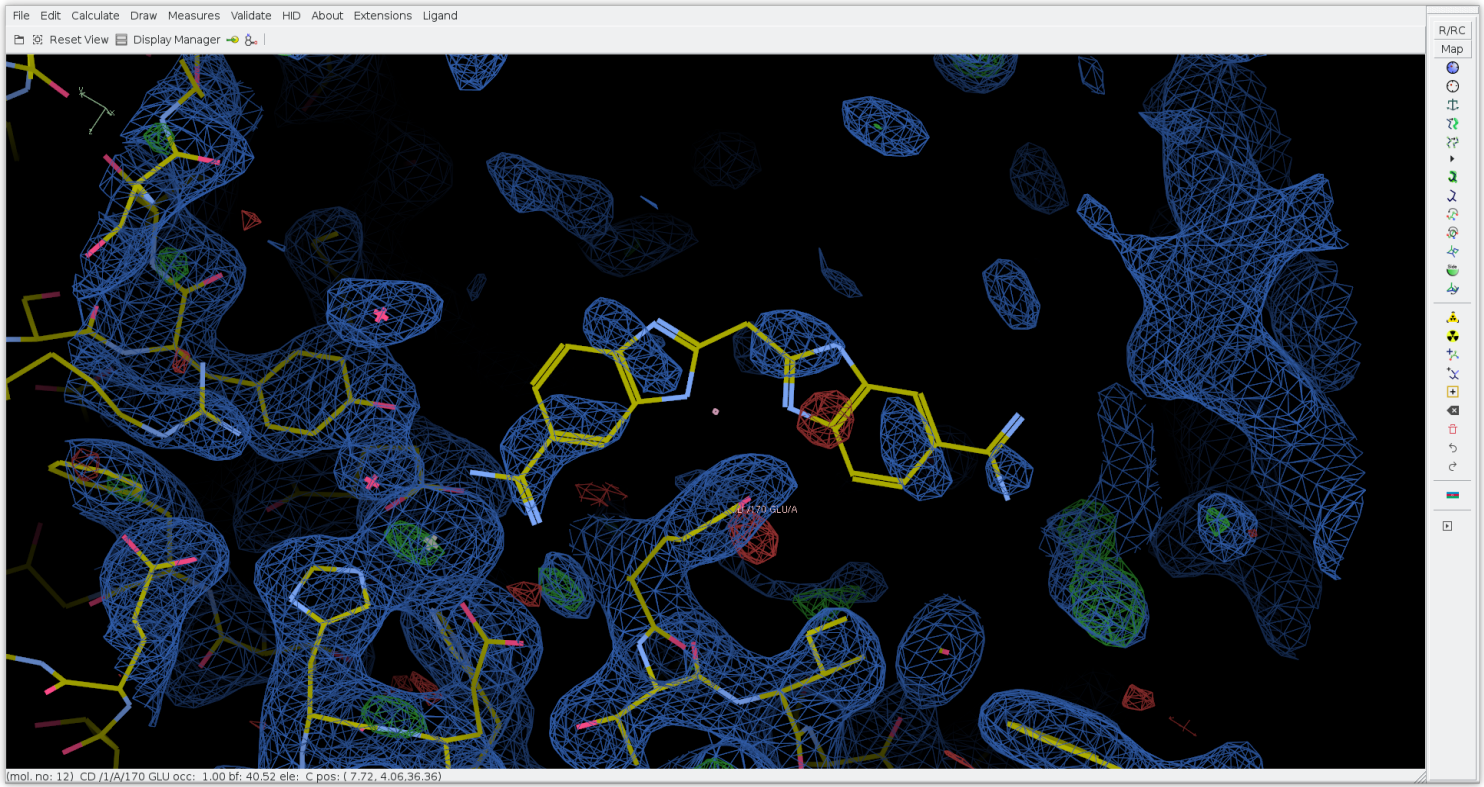


Figure: Ligand bound to PDB 1FQH (retracted, contoured at 0.8 σ).

Ligand geometry

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

- Ligand geometry often neglected
- Unusual or chemically unlikely conformations
- Especially for “uncommon” ligands
- Use ligand validation tools in Coot!

Ligand geometry

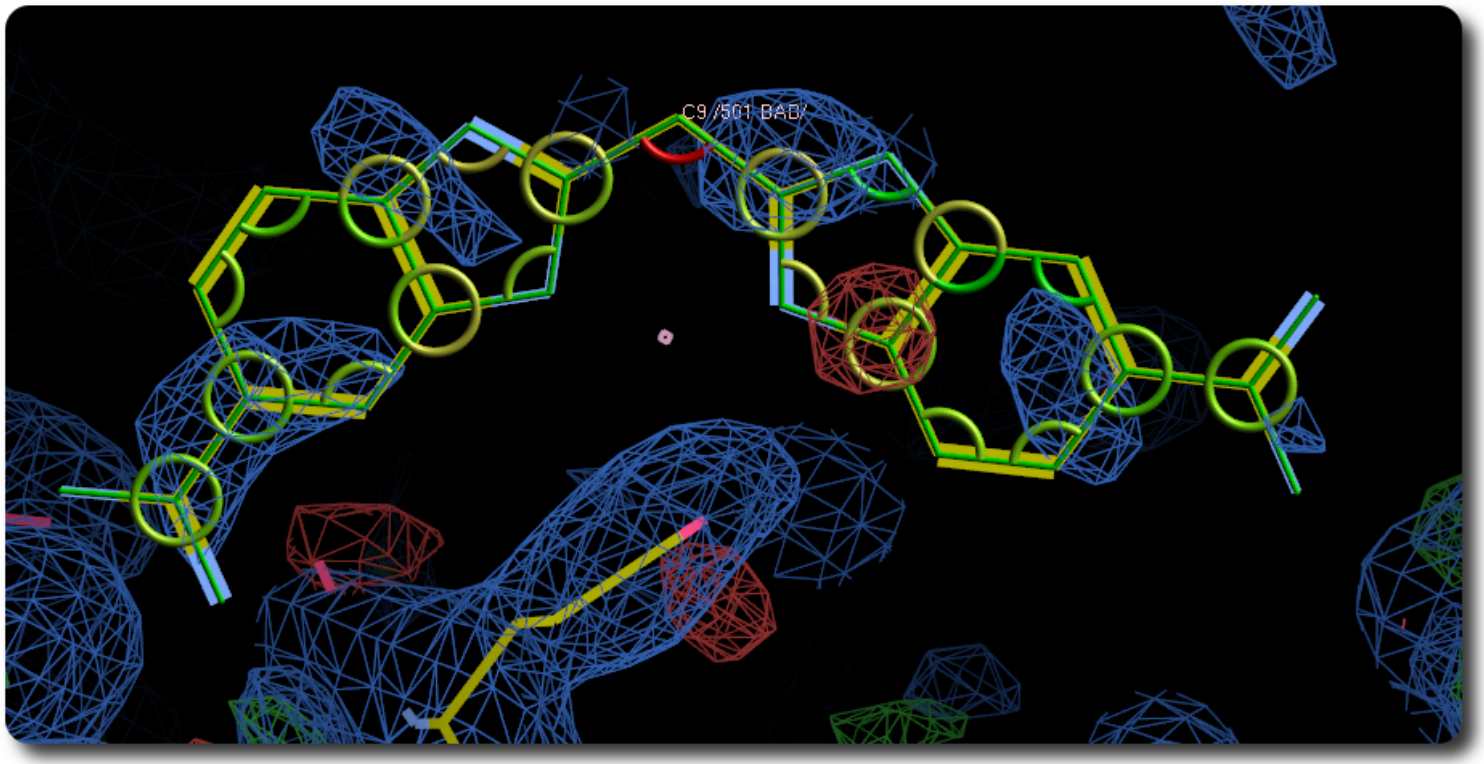


Figure: Ligand bound to PDB 1FQH (retracted, contoured at 0.8 σ).

Contacts & Clashes

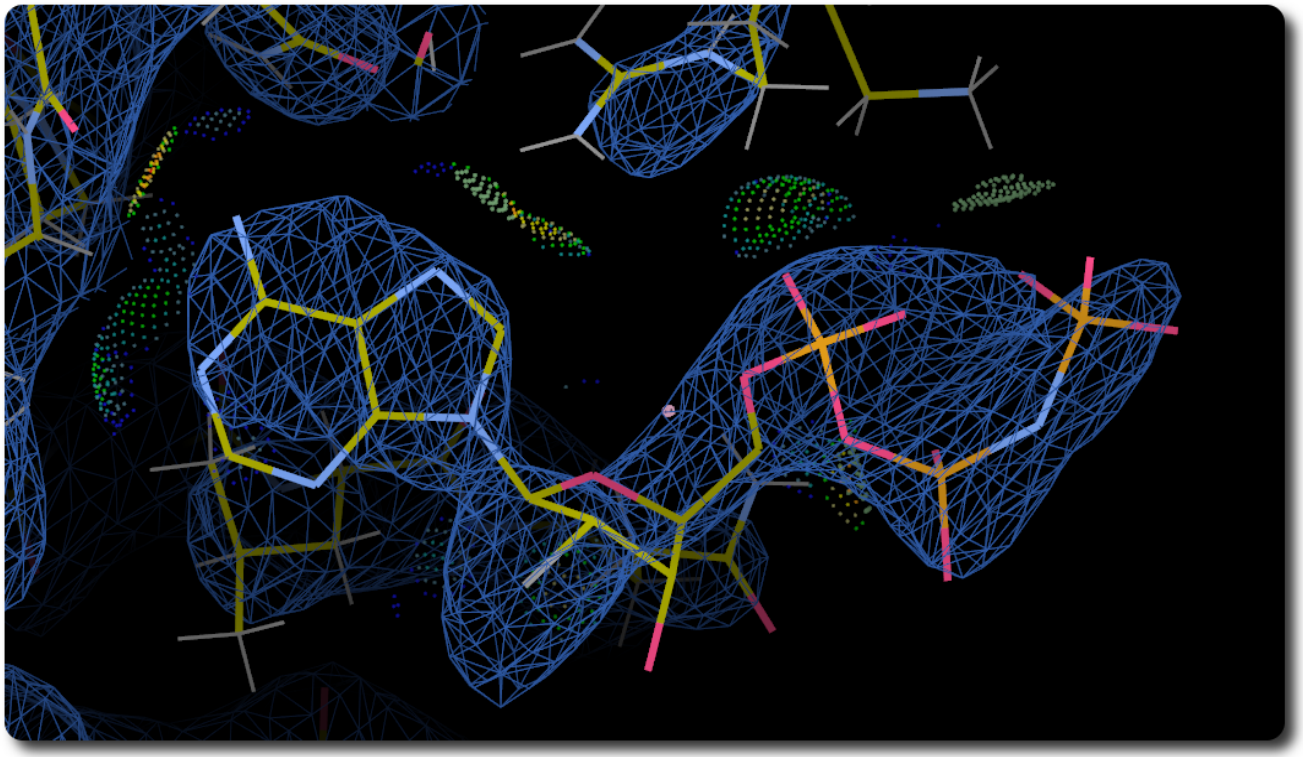


Figure: Contact visualization in Coot (PDB 5JVN).

Minges et al. (2017) doi:10.1038/srep45389

Contacts & Clashes

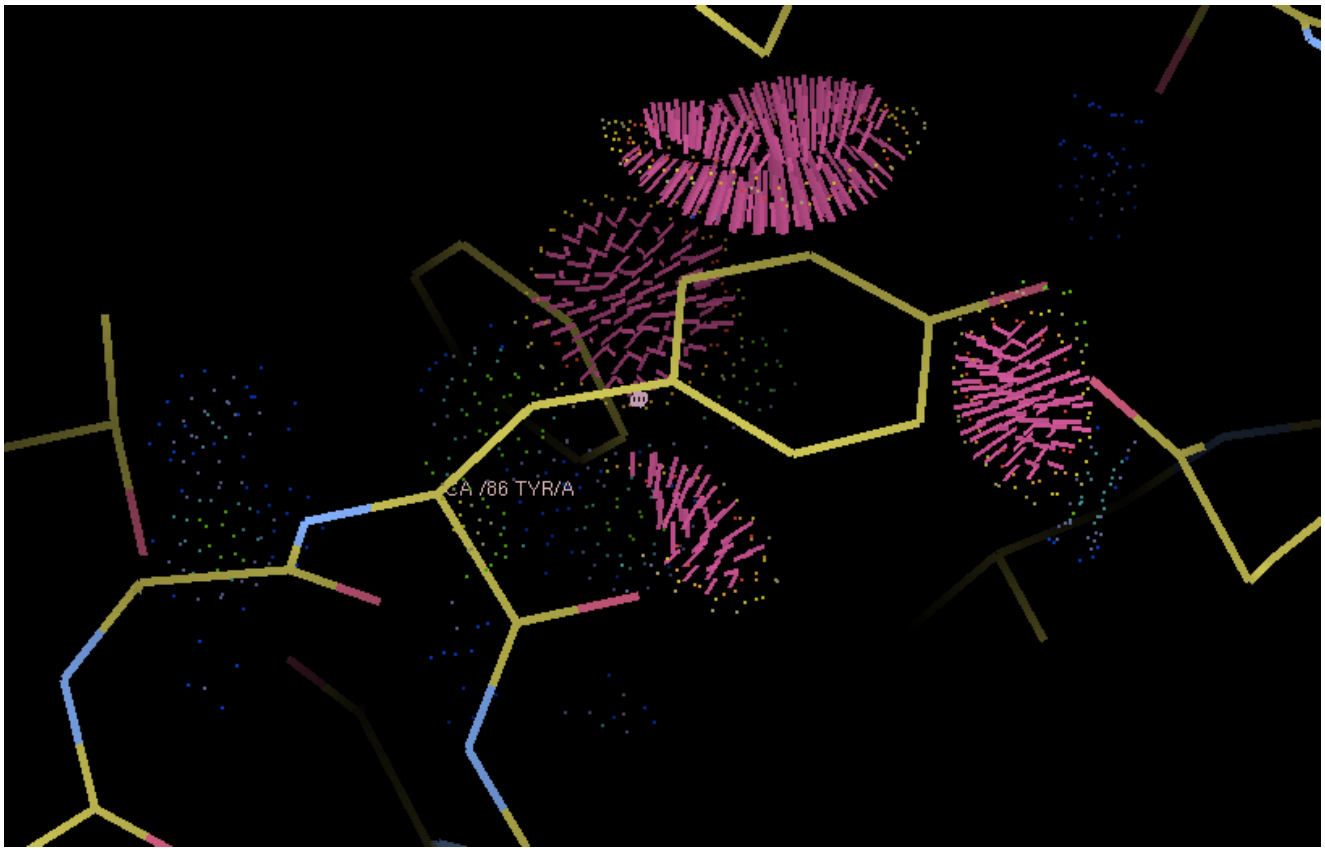


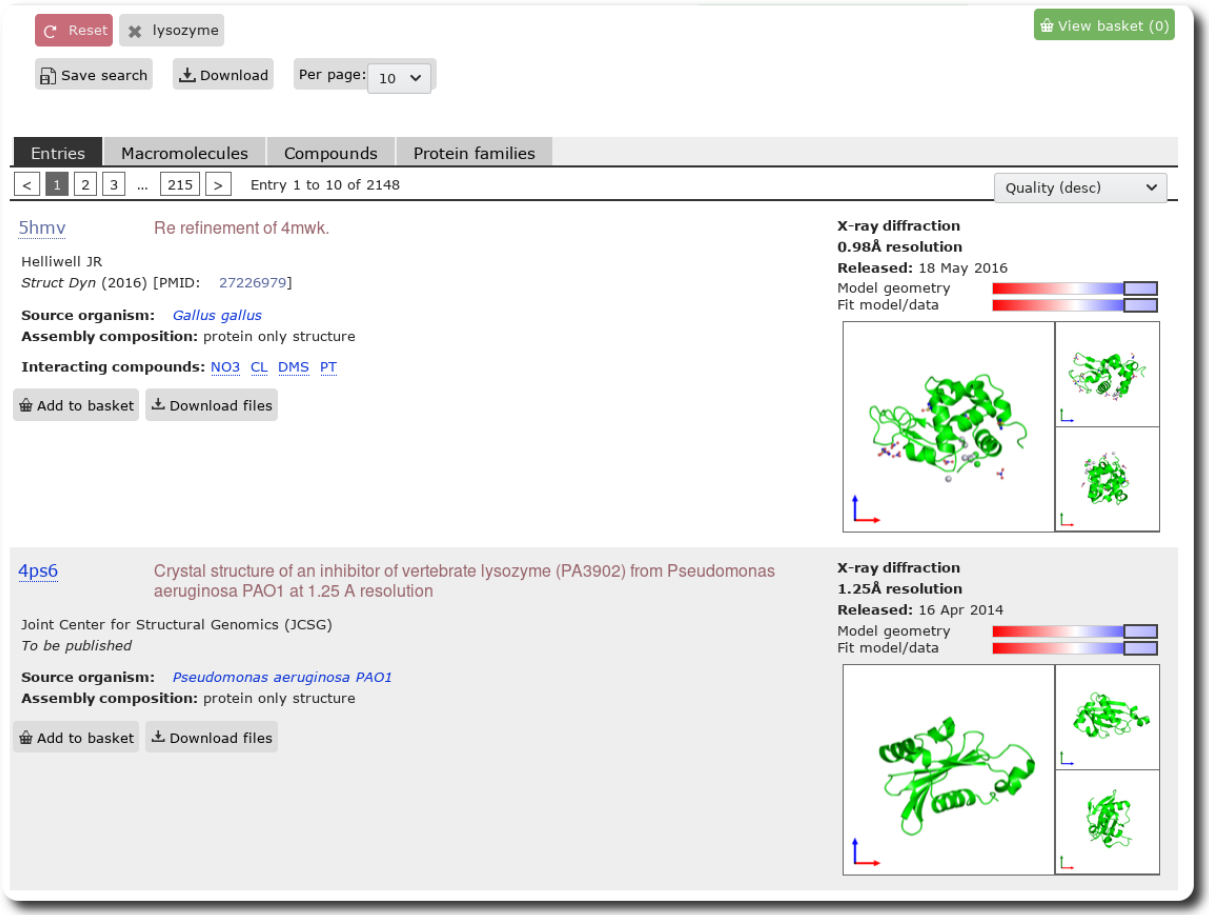
Figure: Visualization of severe clashes in Coot.

https://www.phenix-online.org/documentation/dictionary.html

Validation provided by the PDB

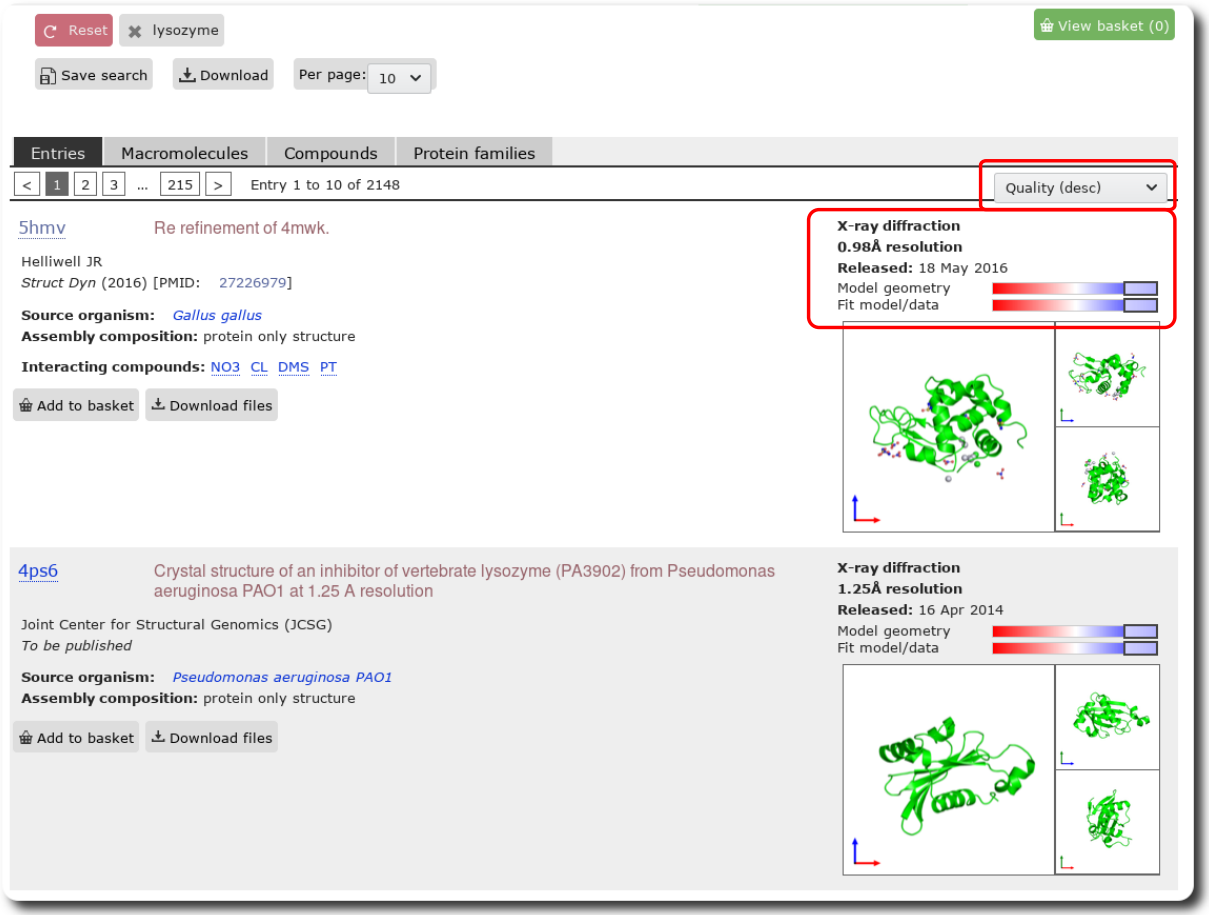
- Validation reports for each structure
- Annual update of reports
- Comprehensive summary of quality indicators

Validation provided by the PDB



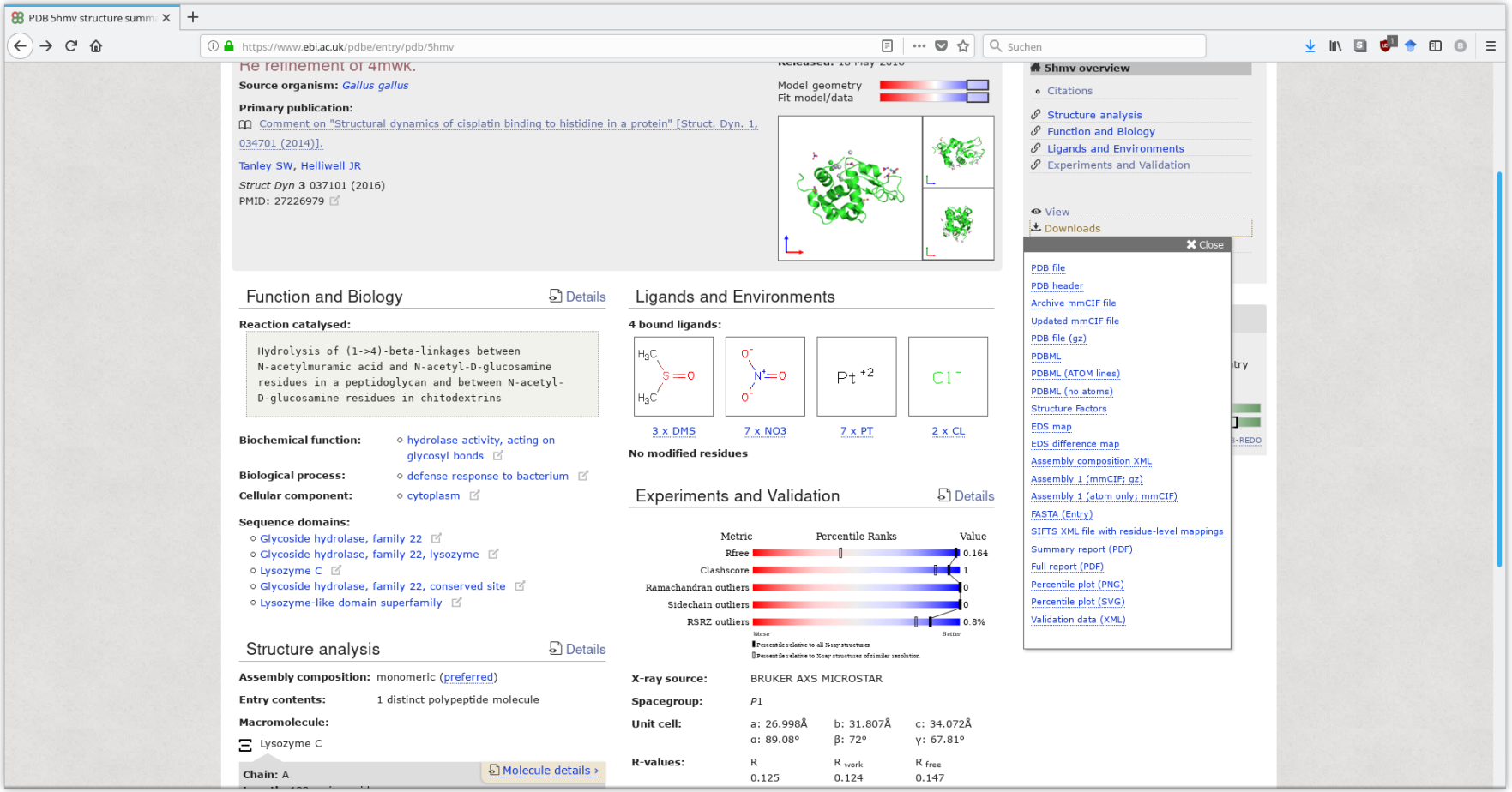
https://www.pdbe.org

Validation provided by the PDB



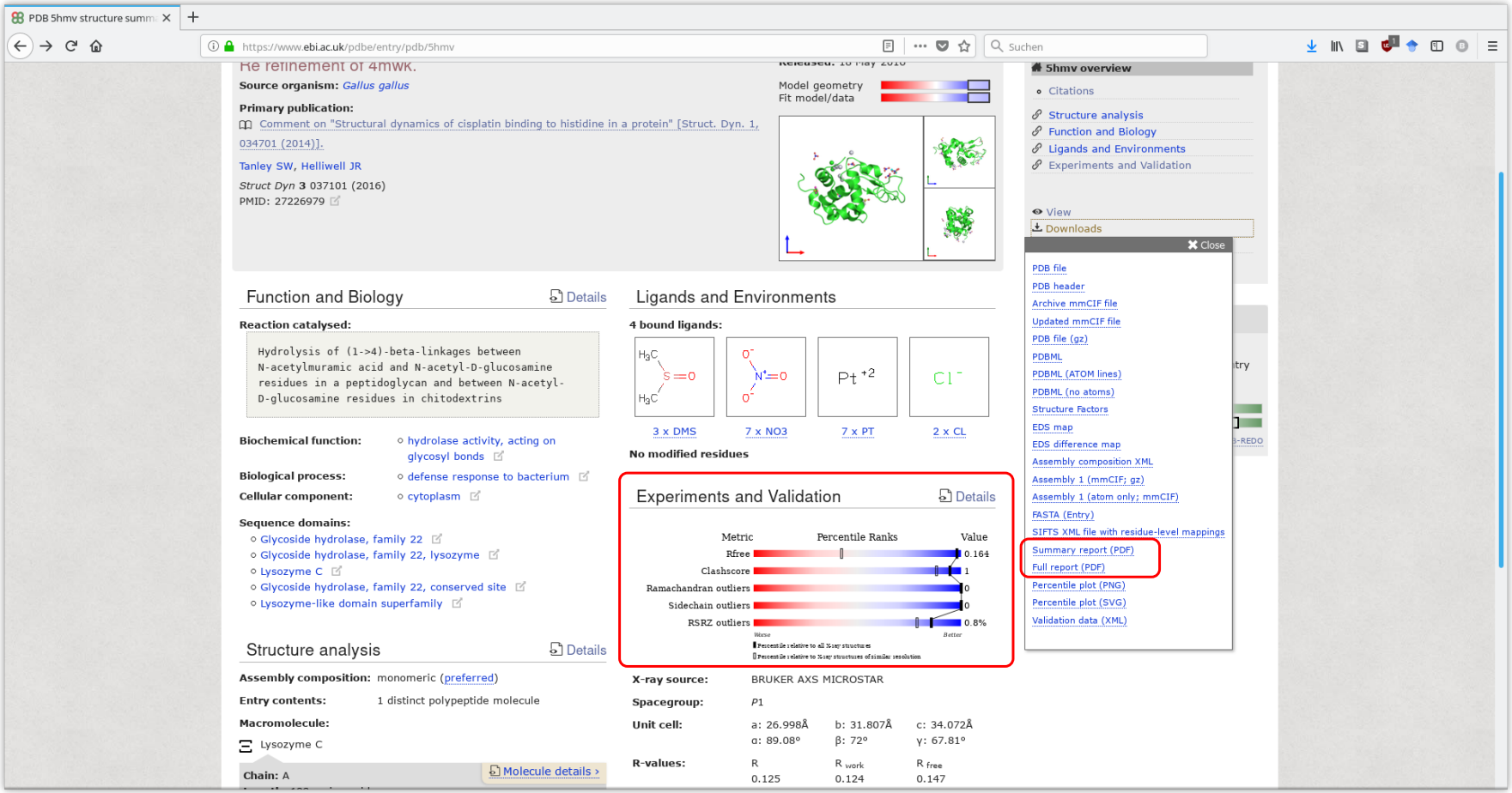
https://www.pdbe.org

Validation provided by the PDB



https://www.pdbe.org

Validation provided by the PDB



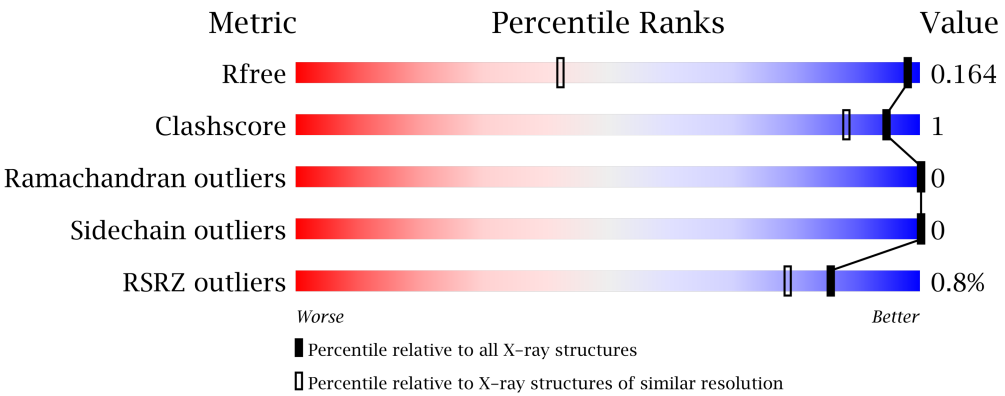
https://www.pdbe.org

1 Overall quality at a glance

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

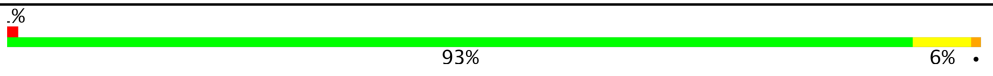
The reported resolution of this entry is 0.98 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



PDB validation report

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	128	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

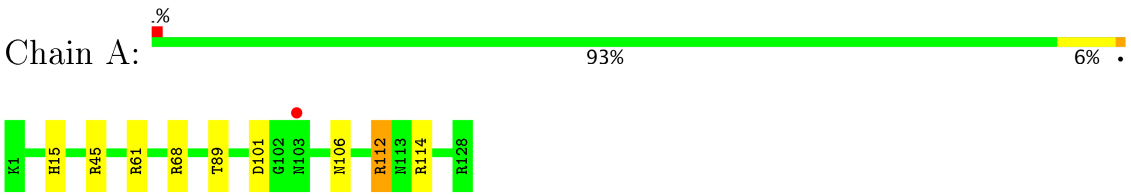
Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	DMS	A	201	-	-	-	X
3	NO3	A	204	-	-	-	X

Continued on next page...

3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

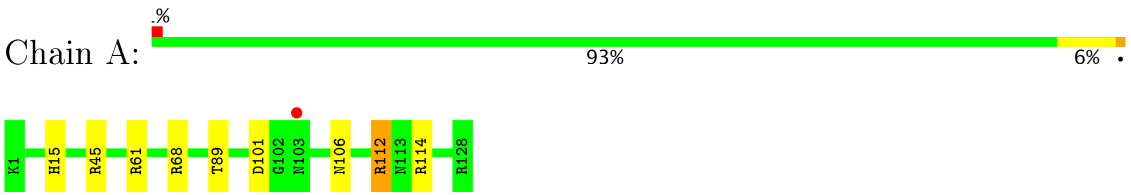
- Molecule 1: Lysozyme C



3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: Lysozyme C



PDB-REDO

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

- Provides automated re-refinement pipeline of existing PDB entries

PDB-REDO

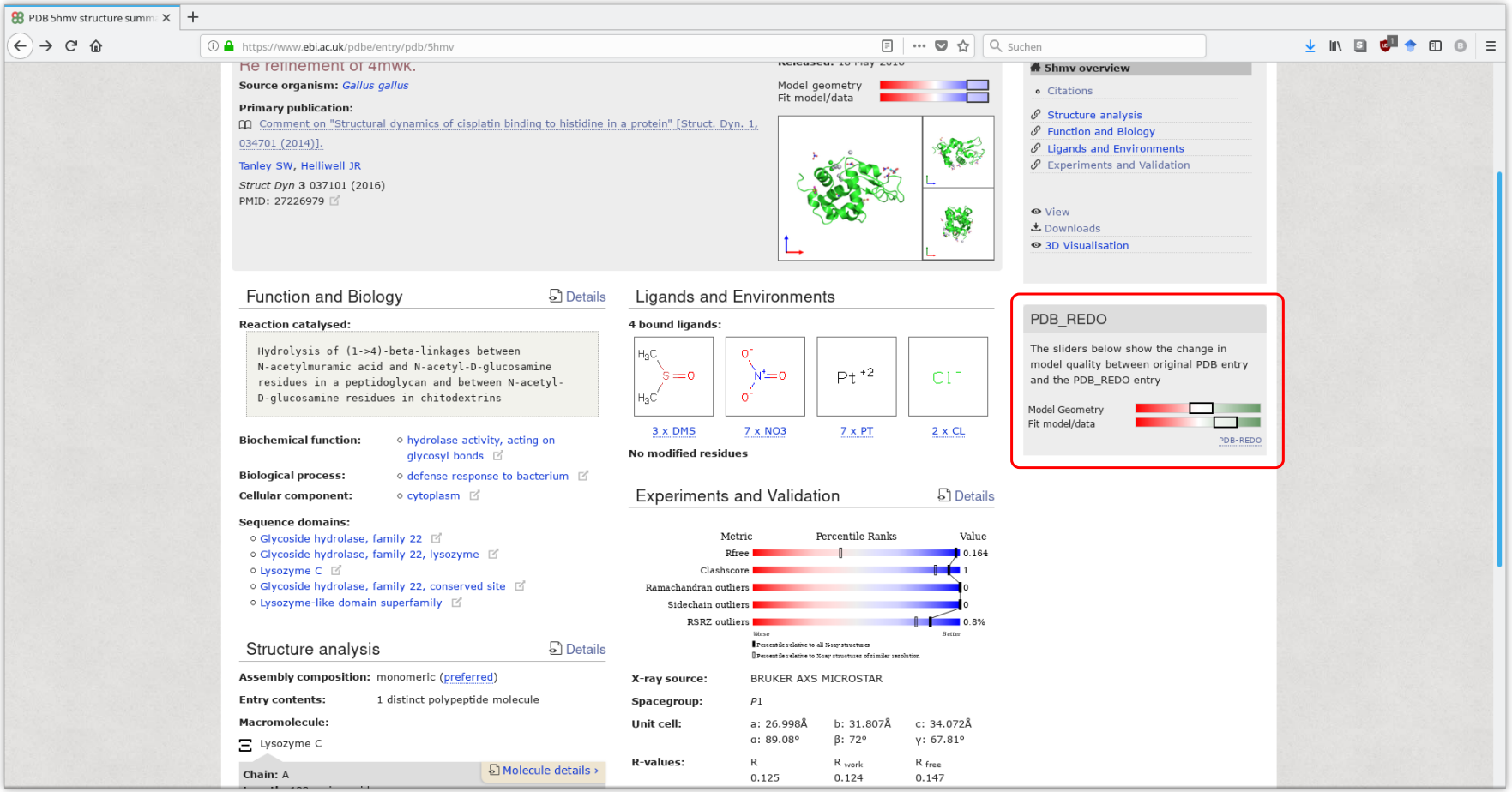
HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

- Provides automated re-refinement pipeline of existing PDB entries
- Older entries may take advantage of new software developments

PDB-REDO

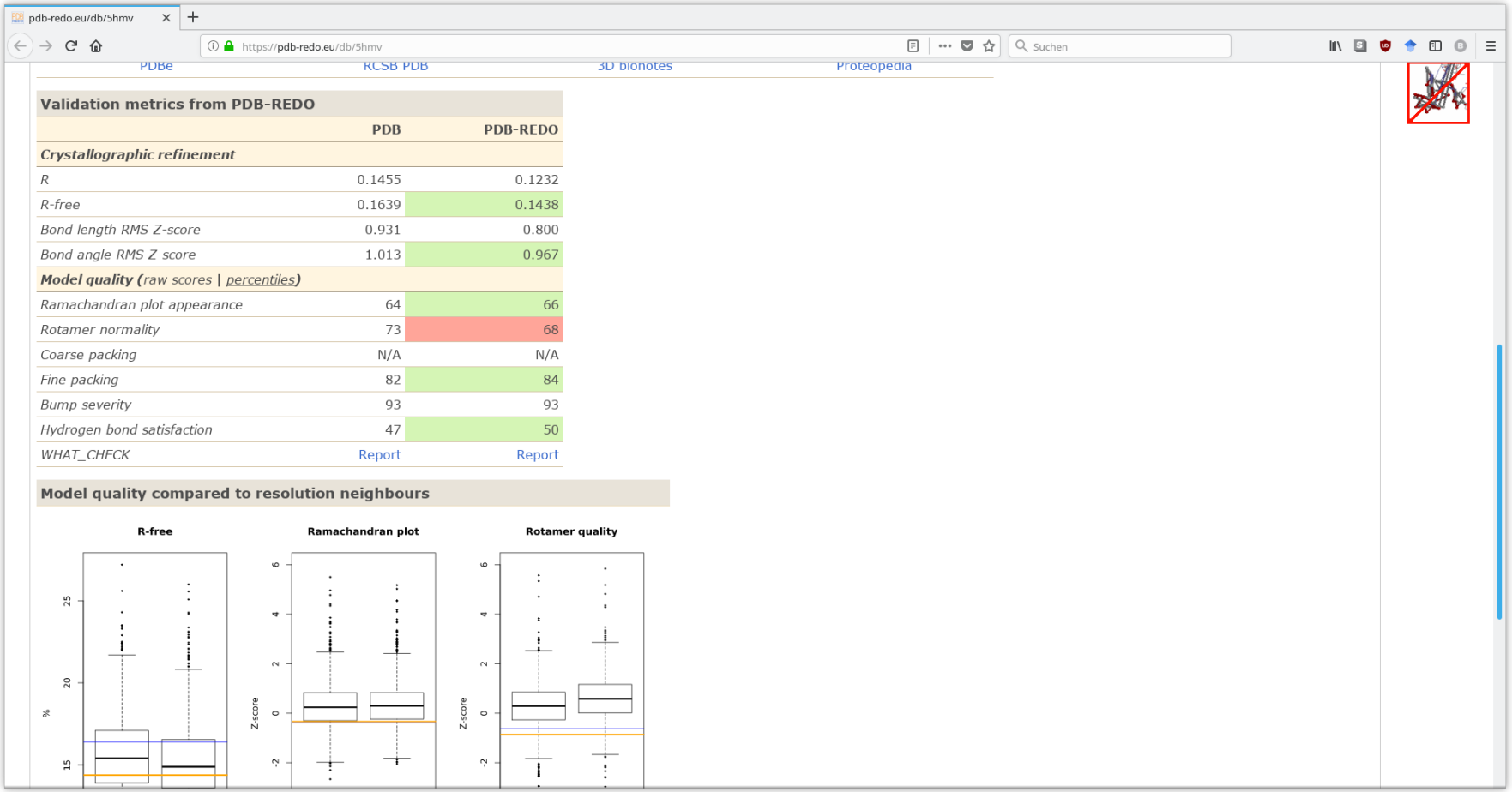
- Provides automated re-refinement pipeline of existing PDB entries
- Older entries may take advantage of new software developments
- Often leads to significant improvement of model quality and/or completeness

PDB-REDO



https://www.pdbe.org

PDB-REDO



https://pdb-redo.eu/db/5hmv

Excercise

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

Assessment of model quality

Choose a model from the PDB (you may choose freely or use one of the examples below). Use the tools discussed in this course to explore the model and evaluate its quality and possible limitations.

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

Assessment of model quality

Choose a model from the PDB (you may choose freely or use one of the examples below). Use the tools discussed in this course to explore the model and evaluate its quality and possible limitations.

- **Example IDs:**
  `4BHX, 5D1F, 1KBL, 5JVN, 2XKG`

Excercise

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

Assessment of model quality

Choose a model from the PDB (you may choose freely or use one of the examples below). Use the tools discussed in this course to explore the model and evaluate its quality and possible limitations.

- **Example IDs:**
  `4BHX`, `5D1F`, `1KBL`, `5JVN`, `2XKG`
- Ligands bound?
- Fit to electron density
- Geometry
- Flexibility

Concluding remarks

- Use your common sense when working with structural models
- Keep in mind that all structures in the PDB are merely **models**

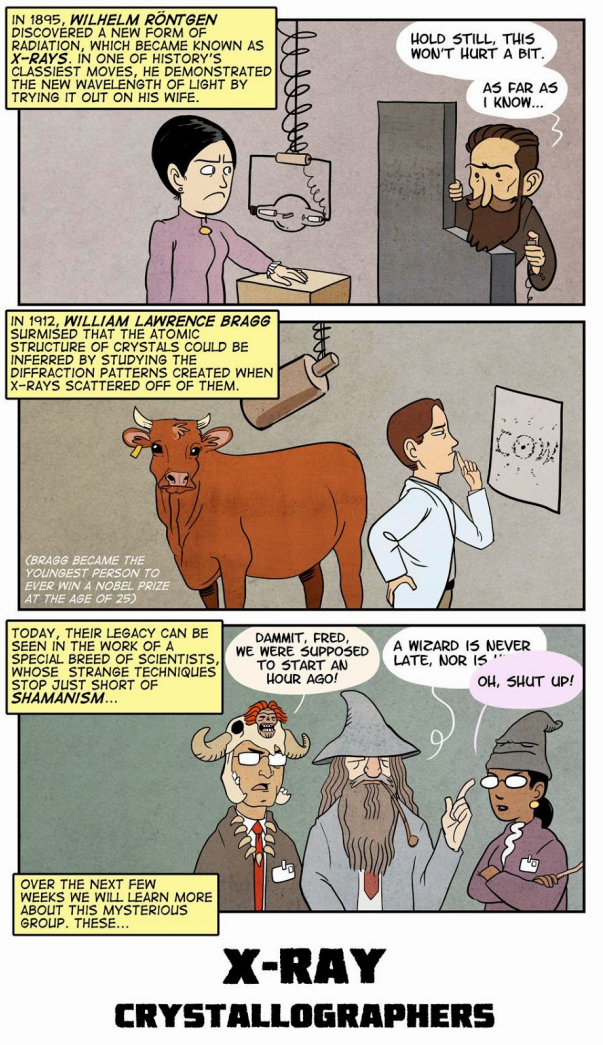
Concluding remarks

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

- Use your common sense when working with structural models
- Keep in mind that all structures in the PDB are merely **models**

Thank you for your attention!

Concluding remarks



Maki Noro: http://www.popsci.com/blognetwork/tags/crystallography