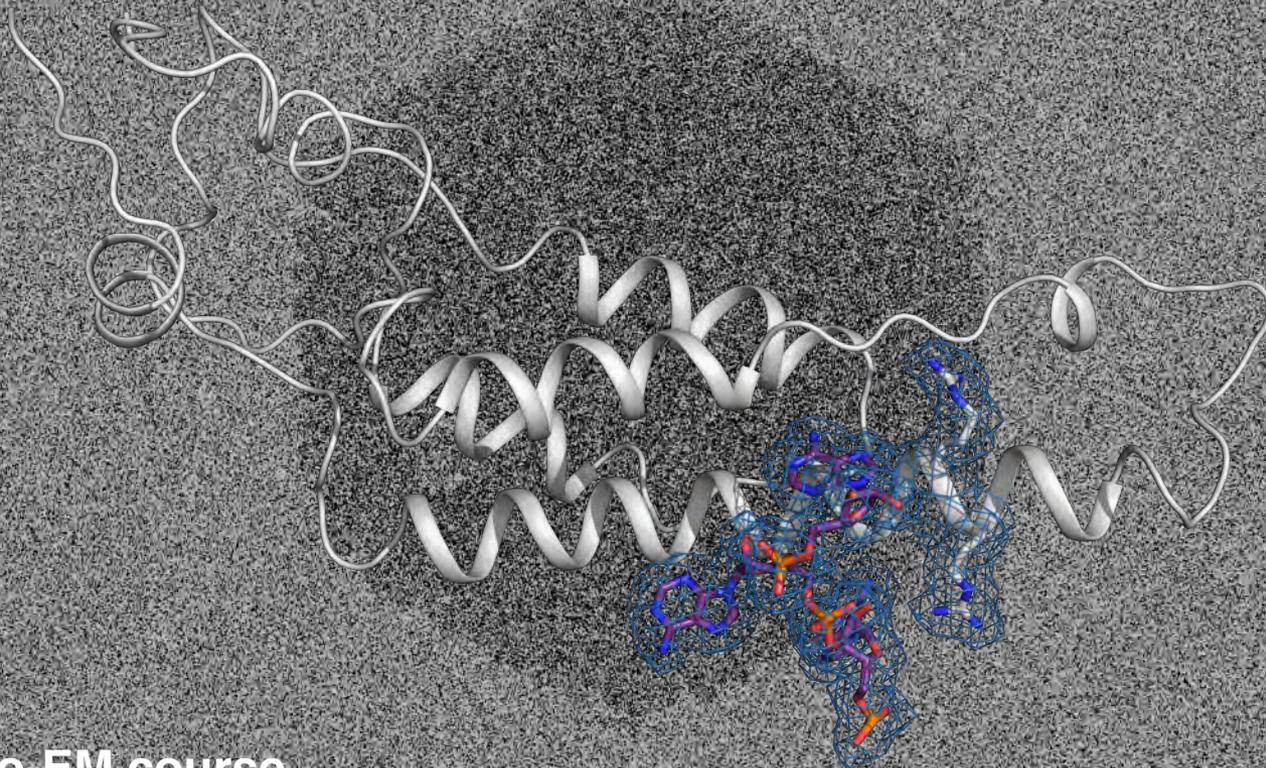
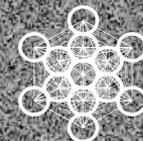


# Model building and refinement with high-resolution cryo-EM data



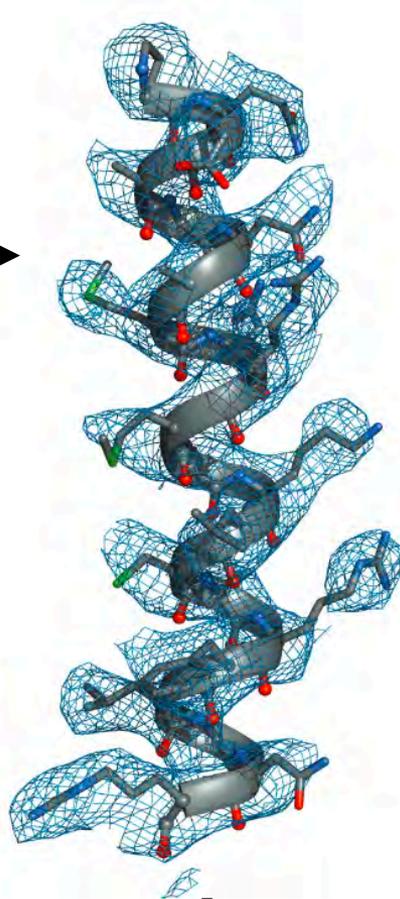
**EMBO cryo-EM course**

Birkbeck College, London, UK  
11 September 2019



# Why do we need model refinement?

**Map**



**Model building**



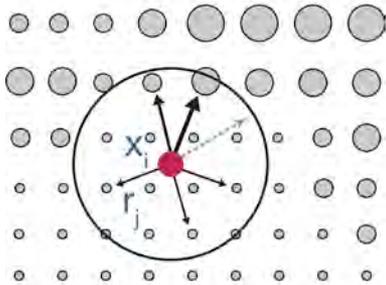
Atomic model guides our mechanistic interpretation of cryo-EM structures

# General aspects of fitting strategies

- **Conformational search** ( $x, y, z, \psi, \theta, \rho$ )
  - orient components in EM density (rigid)
  - sample and score conformations different from template structure (flexible)

- **Fitting potentials**

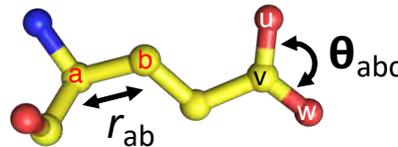
## Density gradient



$$\frac{\delta}{dr} \rho(r)$$

## Energy functional (stereochemistry)

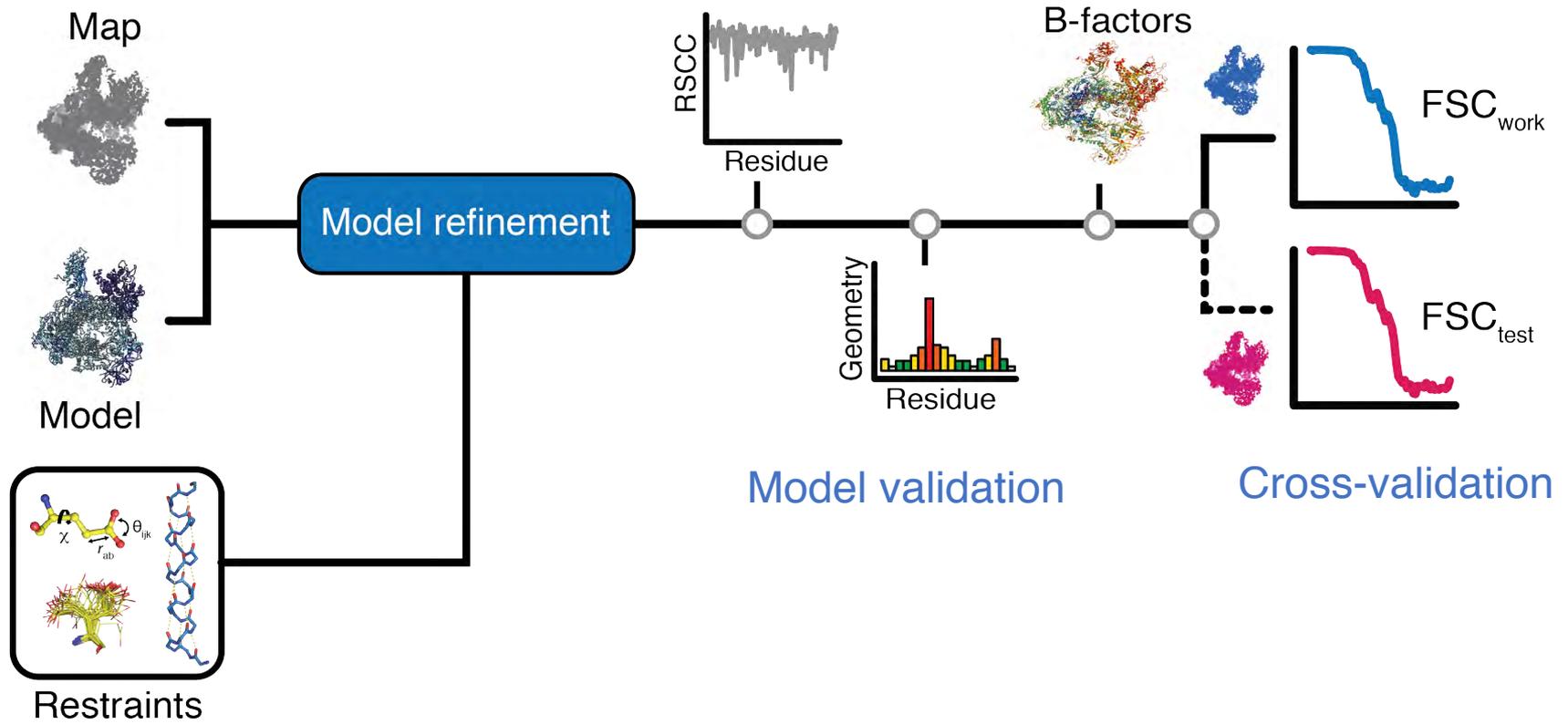
$$E_{tot} = E_{bond} + E_{angle} + E_{dihedral} + E_{vdW} + E_{coulomb} + \dots$$



- **External restraints**

- constraints from alternative methods (SAXS, FRET, mass spectrometry)
- subunit composition and stoichiometry
- interaction restraints
- structural quality parameters

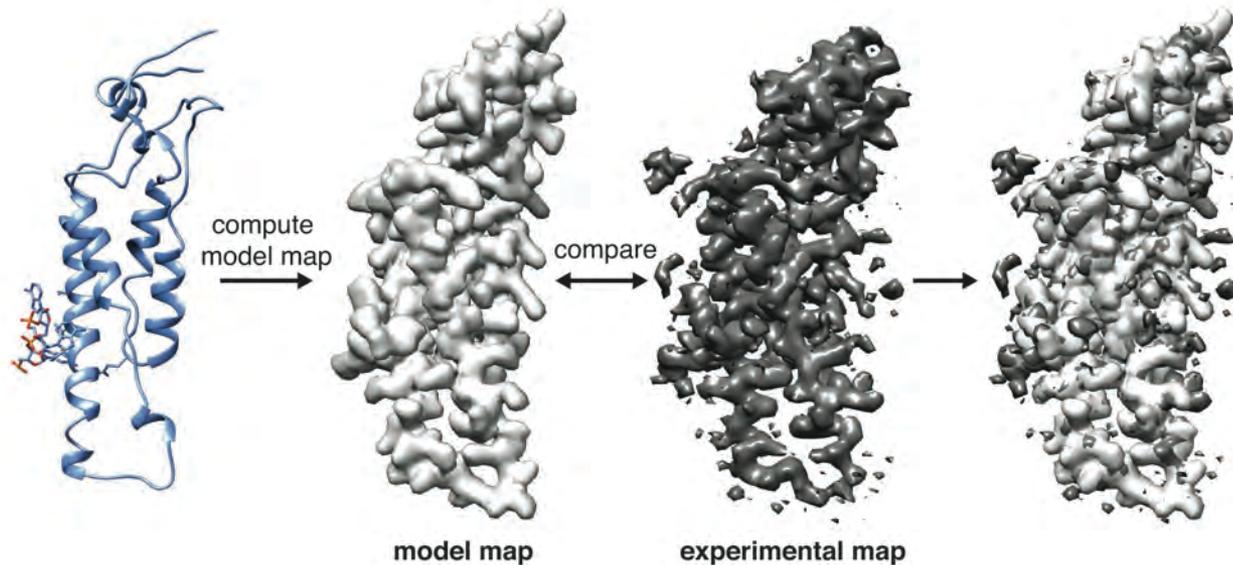
# Atomic model refinement in a nutshell



**Goal:** Fit chemically and structurally consistent models into Coulomb potential maps

# Scoring of fitted models

Correlation of simulated and experimental density



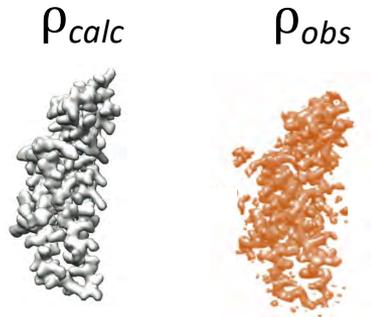
**Quality-of-fit measure:**

$$CC(T) = \frac{\sum_{ijk} \tilde{\rho}_{\text{model}}(\vec{a}_{ijk}) \tilde{\rho}_{\text{target}}(\vec{a}_{ijk})}{\sqrt{\sum_{ijk} (\tilde{\rho}_{\text{model}}(\vec{a}_{ijk}))^2 \sum_{ijk} (\tilde{\rho}_{\text{target}}(\vec{a}_{ijk}))^2}}$$

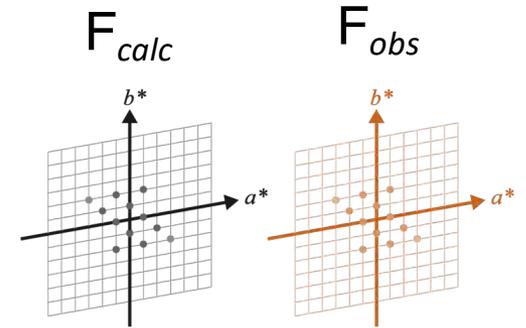
**Refinement target:**

$$E_{\rho} = 1 - CC = \min \int (\rho_o - \rho_c)^2 dv$$

# Real space or Fourier space?



$$E_{exp}(\rho) = \int_V [|\rho_{calc}(\mathbf{x})| - |\rho_{obs}(\mathbf{x})|]^2 d\mathbf{x}$$



$$E_{exp}(F) = \sum_{\mathbf{h}} [ |F_{calc}(\mathbf{h})| - |F_{obs}(\mathbf{h})| ]^2$$

**Real (direct) space**

**Fourier (reciprocal) space**

$$\rho(x, y) = \frac{1}{A} \sum_{h=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \mathbf{F}_{hk} \exp[-2\pi i(hx + ky)]$$

Refinement in real-space and Fourier (reciprocal) space are equivalent

# Restraints – what do we know about macromolecules\*

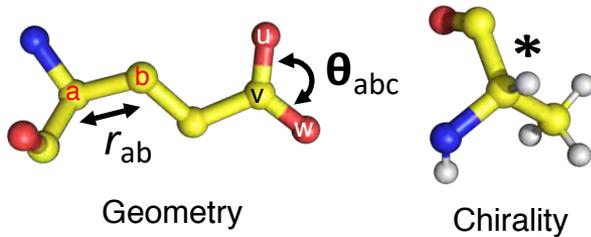
- Proteins consist of atoms that are bonded to each other in a specific way
- Proteins tend to form secondary structures
- Two (or more) molecules with sufficiently high sequence identity are likely to be similar to each other in 3D
- It is likely that if there are two copies of the same molecule they will be similar to each other (at least locally)
- Oscillation of atoms close to each other in 3D cannot be dramatically different
- DNA/RNA tend to form base-pairs and stacked bases tend to be parallel

**All this information can be translated into restraints to restrain refinement**

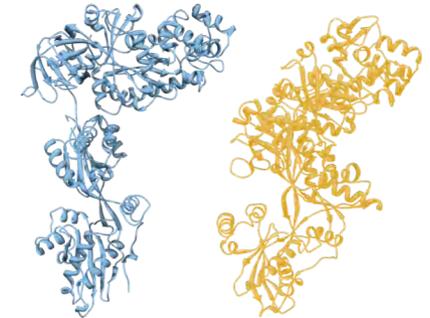
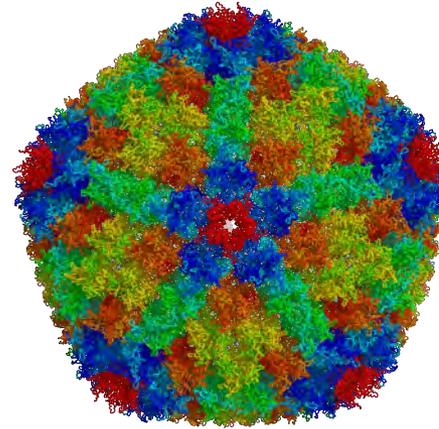
Restraints stabilize refinement, ensure that the final model is consistent with prior knowledge and reduce the chance of overfitting.

# Summary: sources of prior information (restraints)

## Stereochemistry

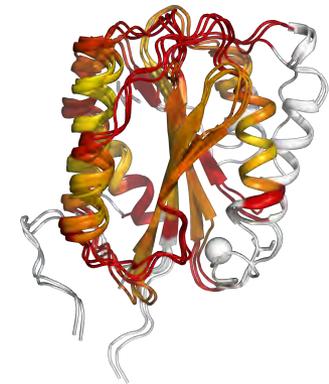
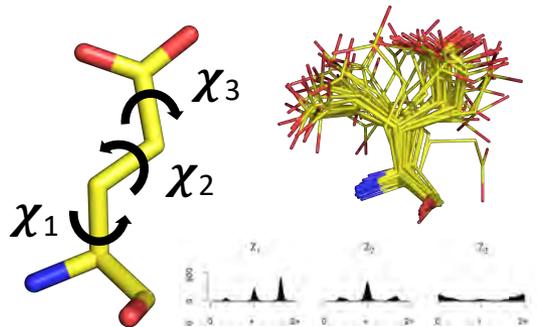


## (NCS) symmetry

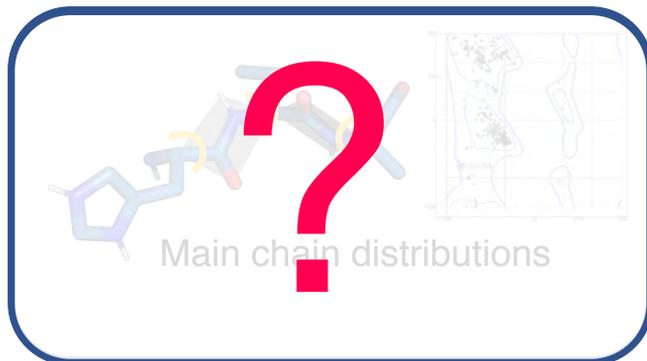


Reference restraints  
(Jelly Body, DEN, ...)

## Rotamer distributions

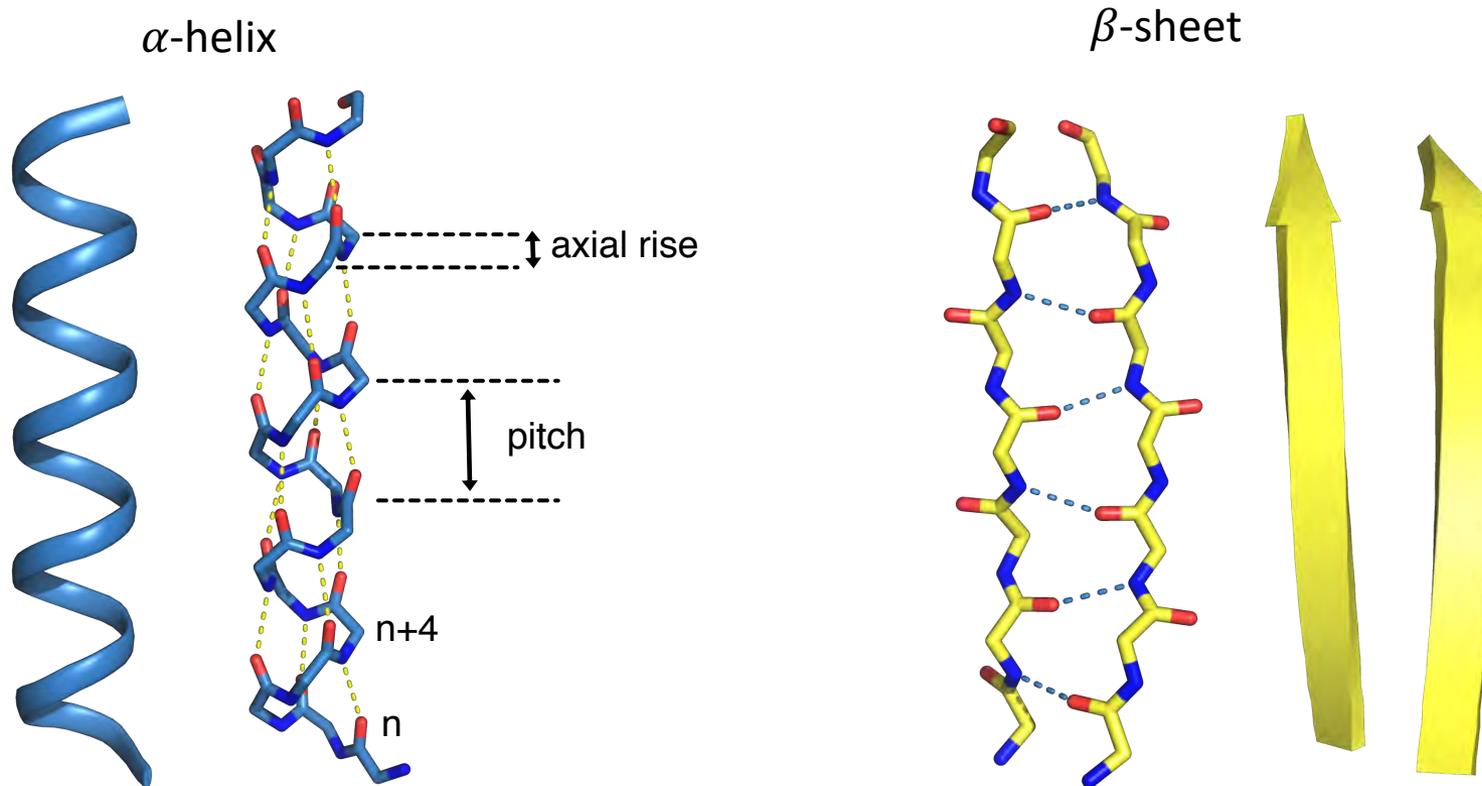


Similarity restraints  
(ProSMART, ...)



# Secondary structure restraints

Distances between hydrogen-bonding atoms in protein helices and sheets or nucleic acid base pairs can be restrained.



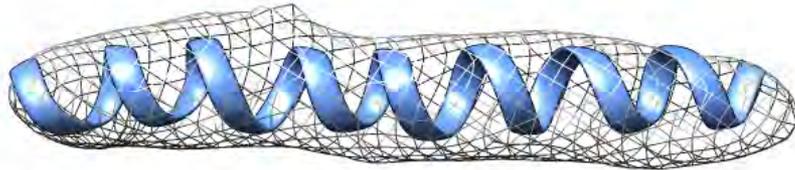
## Hydrogen bonding pattern

$\alpha$ -helix	n+4
$3_{10}$ -helix	n+3
pi-helix	n+5

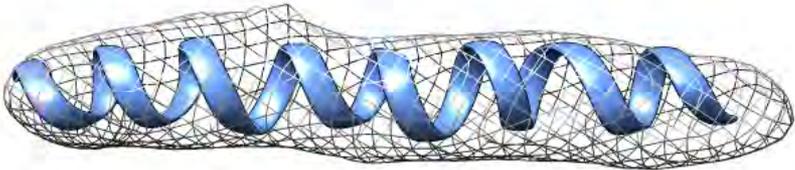
Helps keep regular structure from unravelling during refinement

# Secondary structure restraints

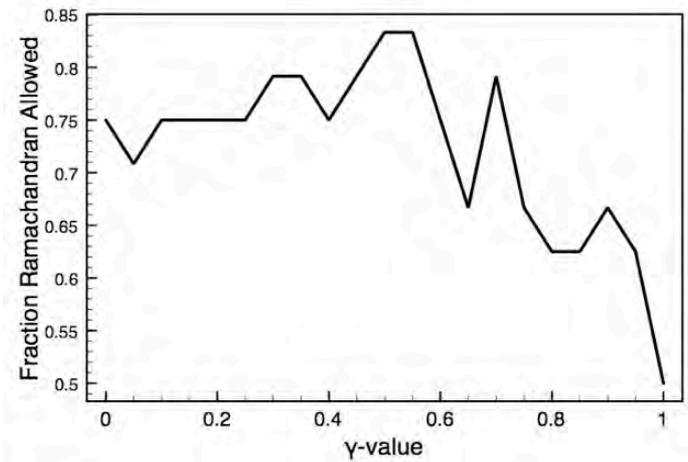
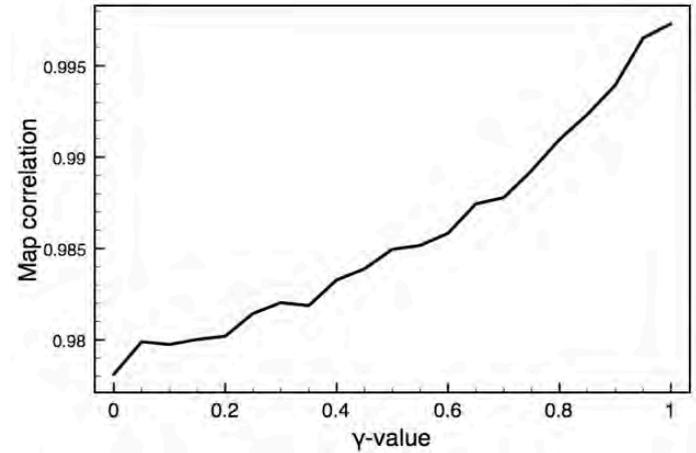
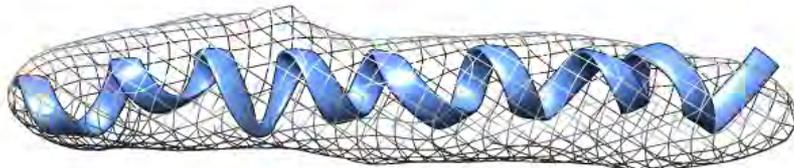
$k = 0.0$



$k = 0.5$



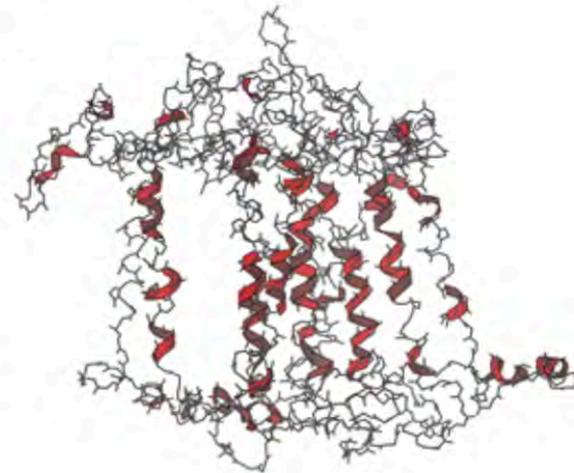
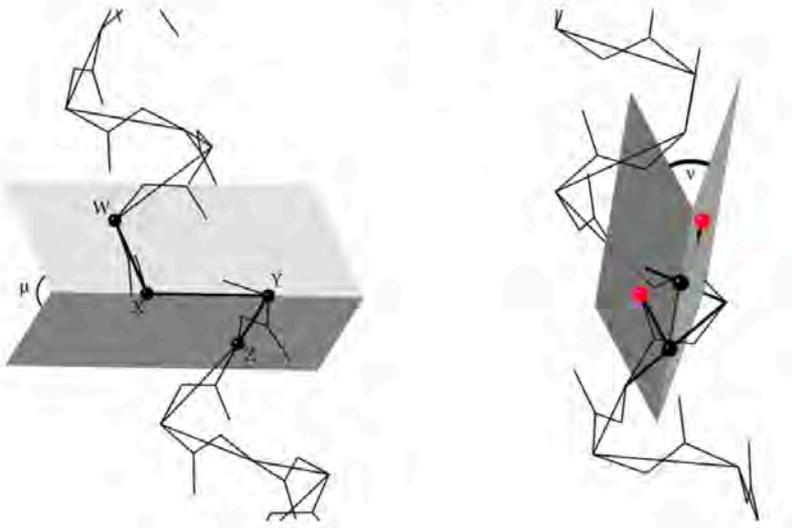
$k = 1.0$



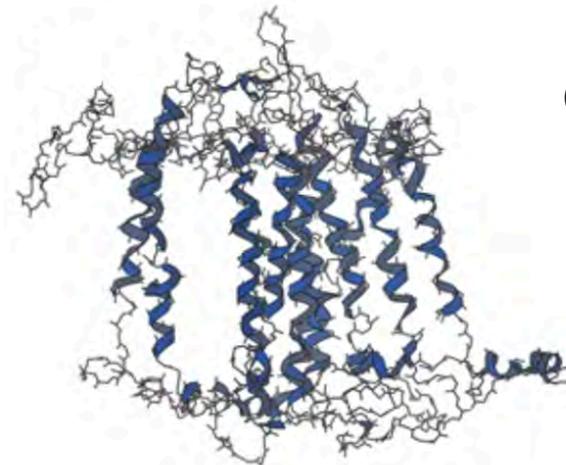
# CaBLAM

Low-resolution structures are vulnerable to errors in peptide plane orientation.

SSE detection with pseudodihedrals

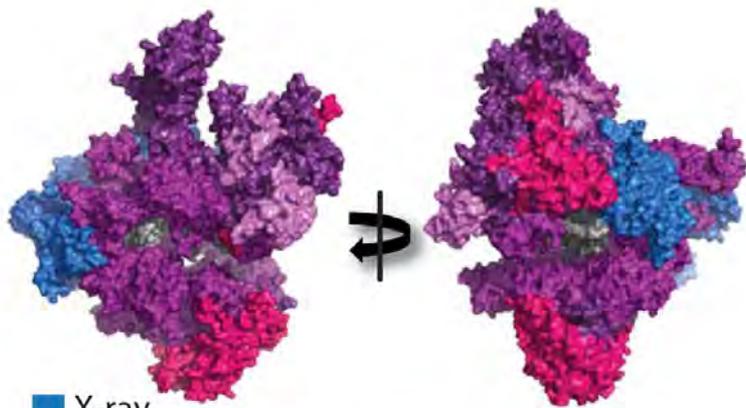


kdssp



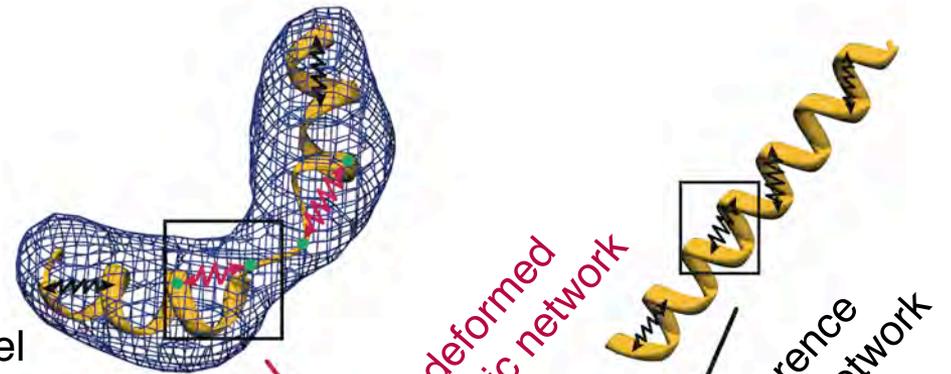
CaBLAM

# Fuzzy constraints



- X-ray
- Homology (high)
- Homology (medium)
- Homology (low)
- De novo

$$E_{\text{DEN}}(n) = k \sum_{\text{pairs } i,j} \left( d_{ij}(n) - d_{ij}^0(n) \right)^2$$



reference model

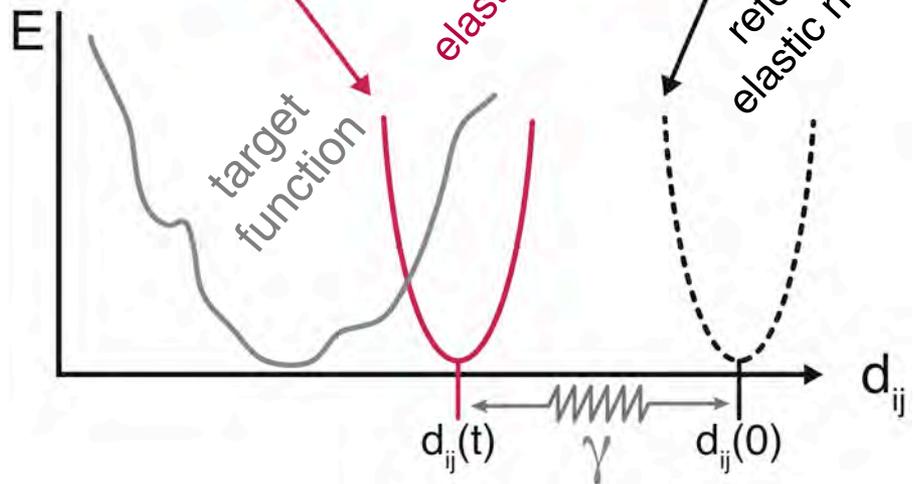
$\gamma = 0$

deformability

$\gamma = 1$

experiment

Similar to “Jelly Body”  
refinement in Refmac



# Tools

## Manual model building

Coot

iSOLDE (interactive molecular dynamics)

Paul Emsley

Tristan Troll

## Automatic model building

Buccaneer

ArpWarp

Phenix.map\_to\_model

Rosetta

Kevin Cowtan

Victor Lamzin

Tom Terwilliger

Frank di Maio

## Model refinement

Refmac

Phenix.real\_space\_refine

Rosetta

Garib Murshudov

Pavel Afonine

Frank di Maio/David Baker

## Other Tools

ProSMART

LibG

ACEDRG

Molprobity

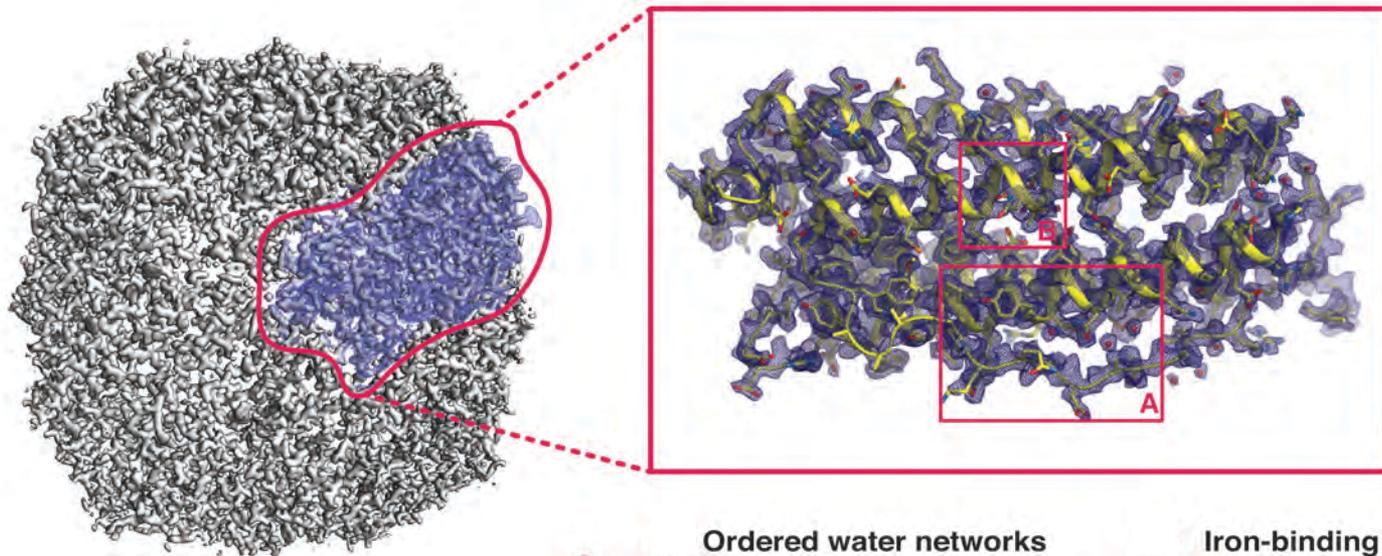
Rob Nicholls

Fei Long

Fei Long

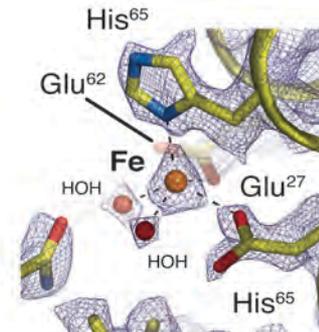
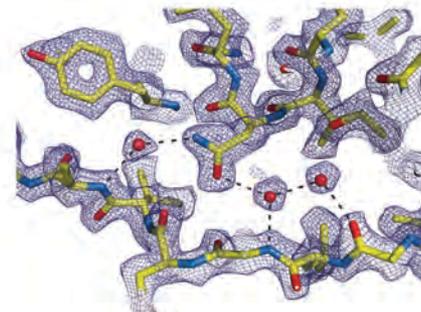
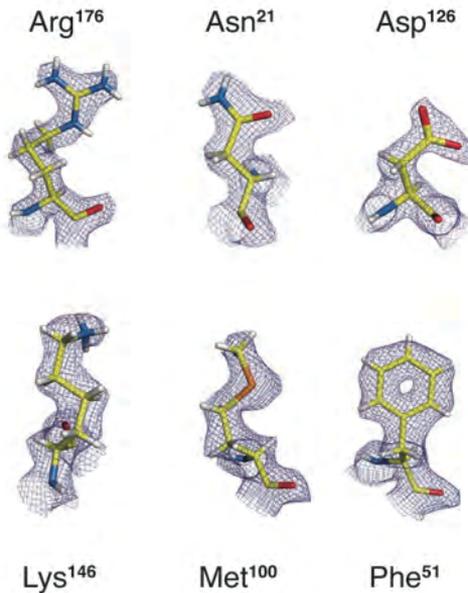
Jane Richardson

When likely everything will just work...



Ordered water networks

Iron-binding site

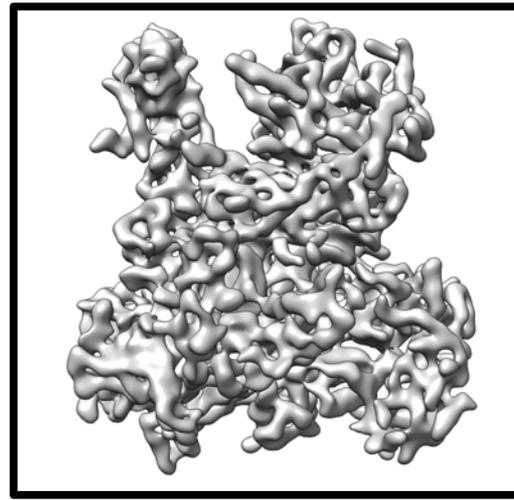
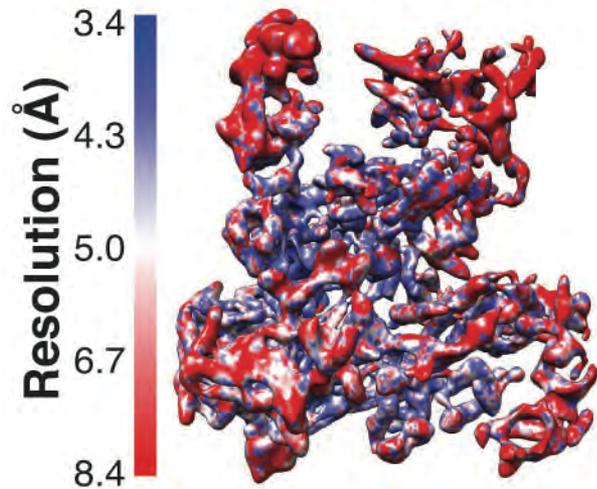


EMD-0144 | PDB 5fja | EMPIAR-10200

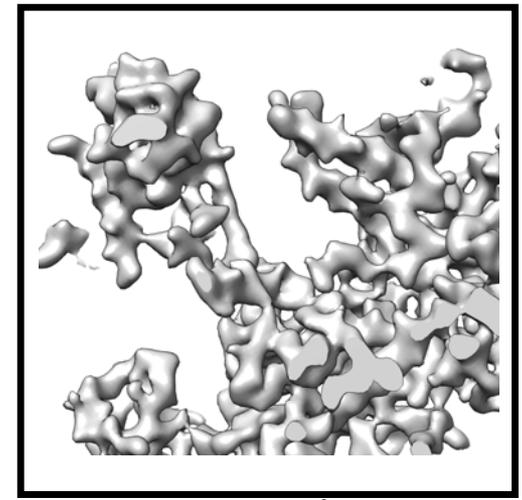


Wim Hagen

...and when it may be more challenging

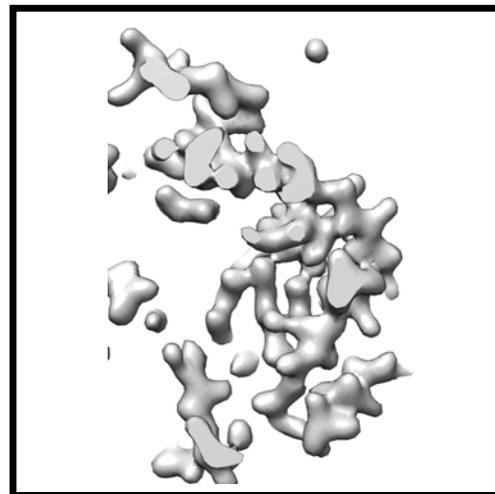


9 - 7 Å

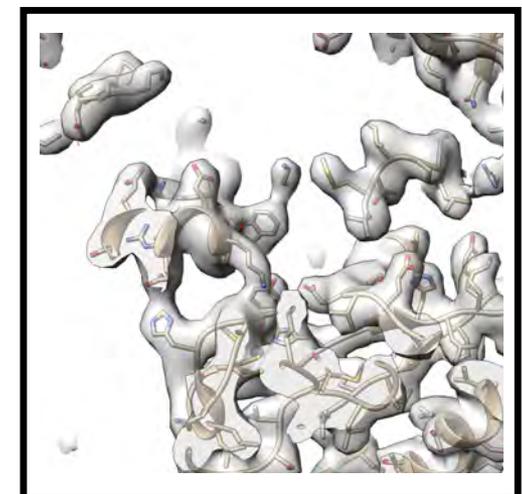


5 - 6 Å

RNA Pol III (open)  
EMD-3180 | PDB ID 5fja



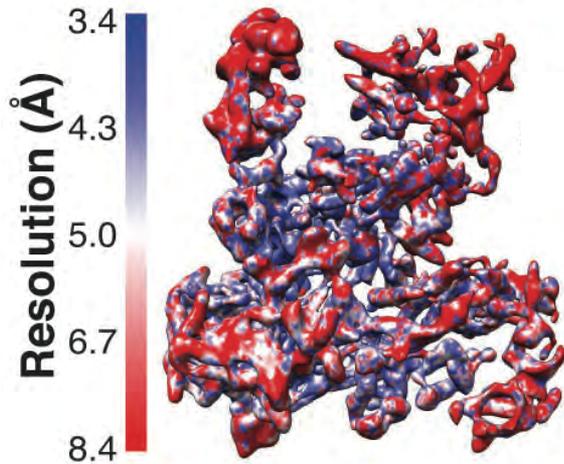
< 4.8 Å



< 4.0 Å

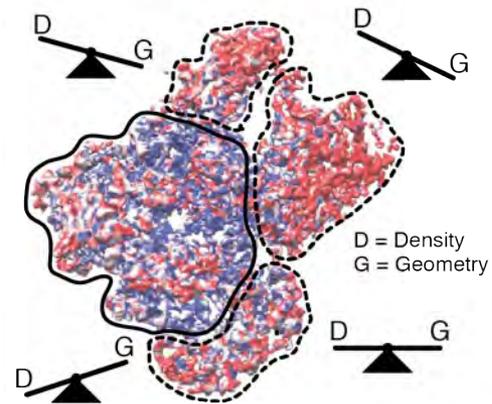
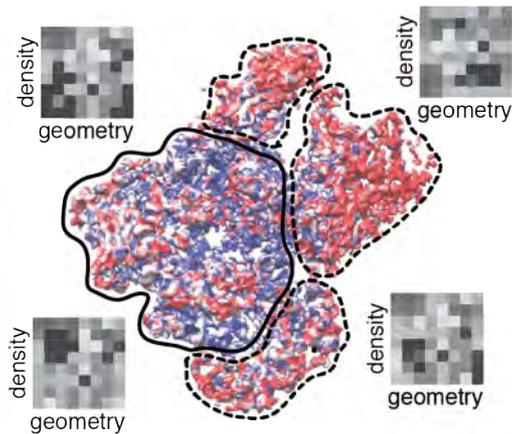
# Common challenges for cryo-EM structures

## Resolution variation

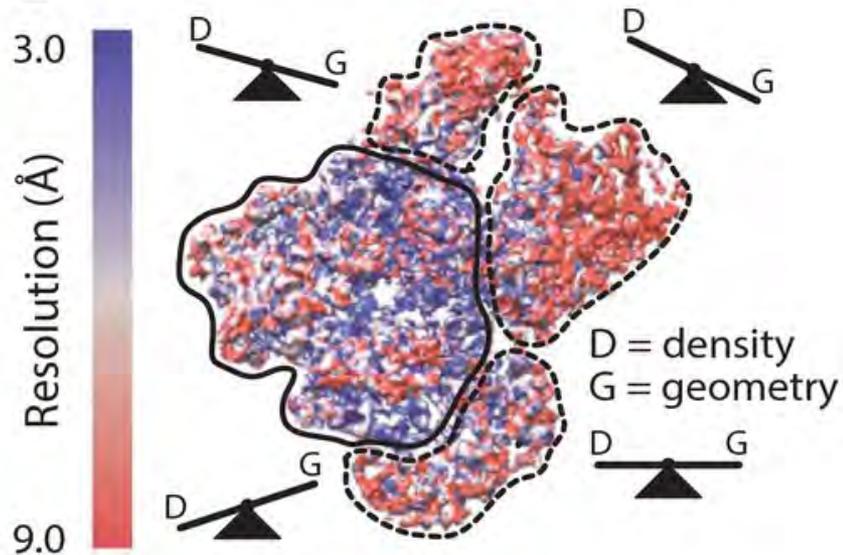


Single refinement strategy may not be appropriate

- Segmented refinement
- Local weight refinement



# Coupling restraint weighting to local resolution



## Refinement statistics

	global	local
Ramachandran (%)		
favoured	78.32	85.52
allowed	13.36	13.25
disallowed	8.23	1.23
Rotamer outliers (%)	17.4	2.1
C-beta deviations	68	0
Clash score*	28.6	13.4
MolProbity score	3.24	2.38

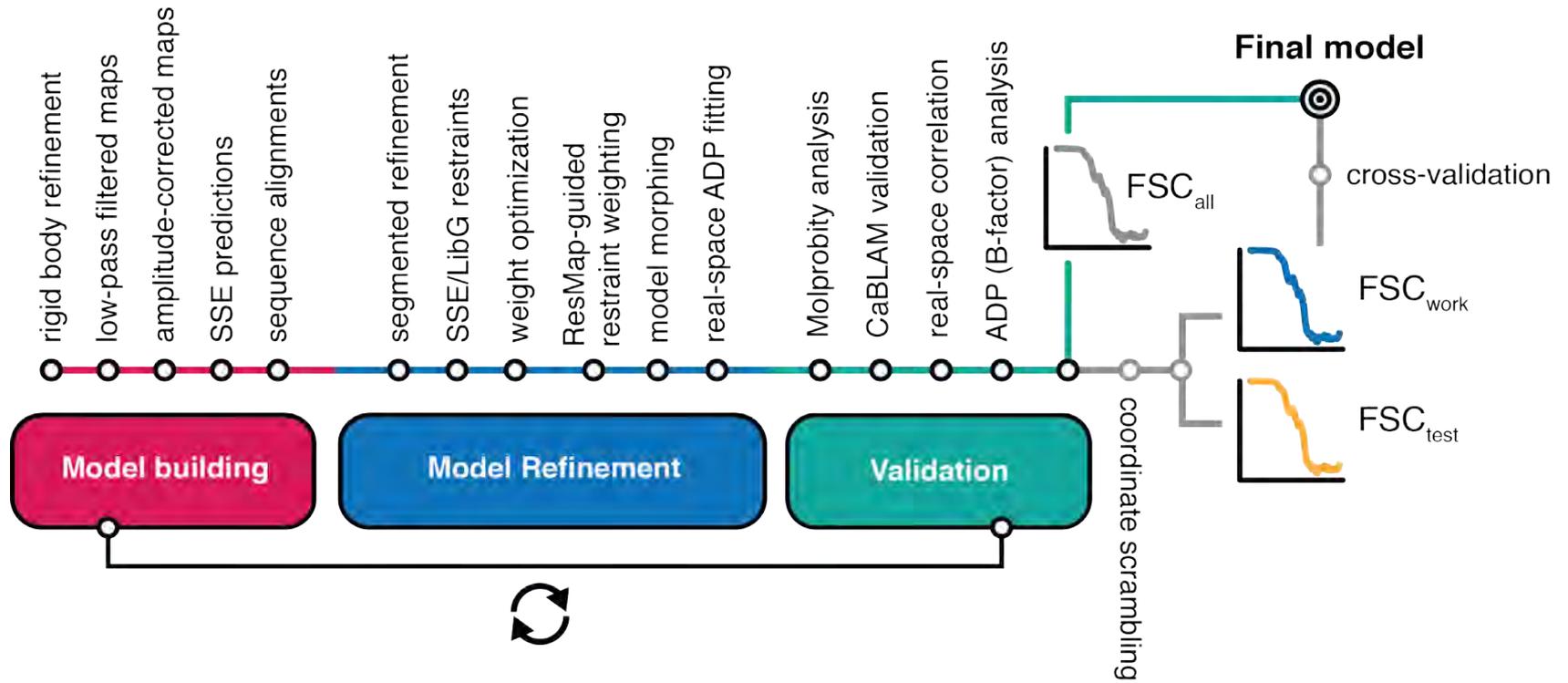
\*( $\Sigma$ vdw overlaps)/1000 atoms

- Local determination of refinement target weights improves model geometry
- Currently done in “area-mode”, but could be done on per-residue basis

*Nature* **528**: 231-239 (2015)

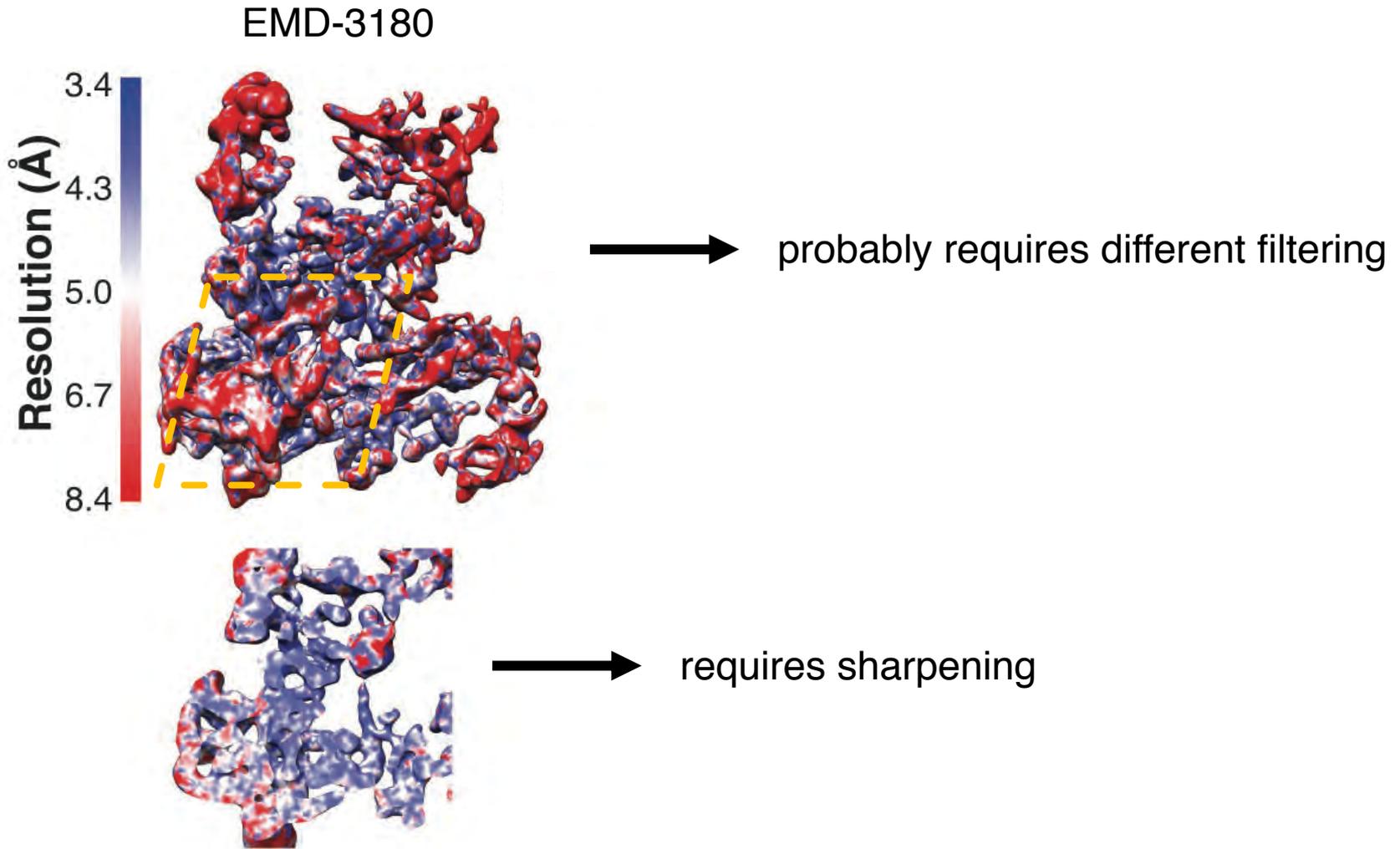
*FEBS* **283**: 2811-2819 (2016)

# Defining a strategy for model refinement



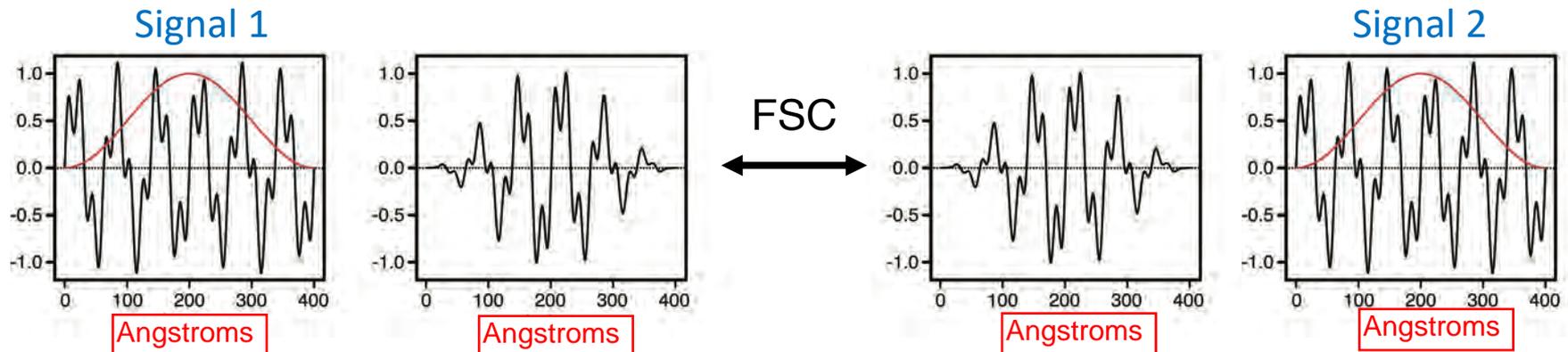
Optimal strategy may (and probably will) differ in each case!

# Resolution variation in cryo-EM maps

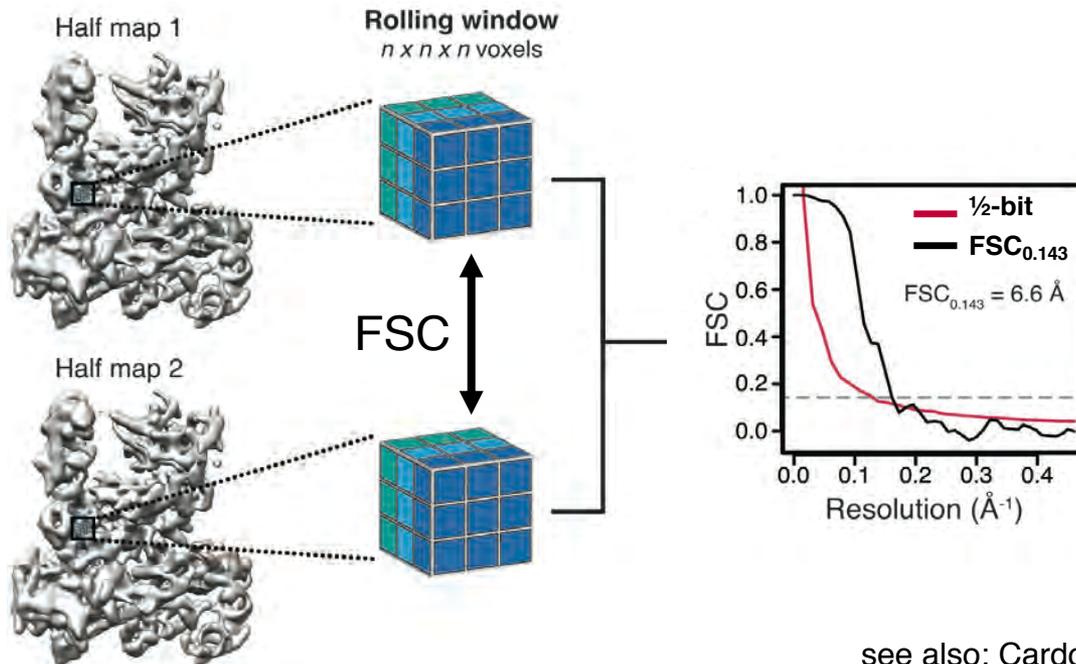


# Local filtering

## Localized Fourier correlations



## 3D case

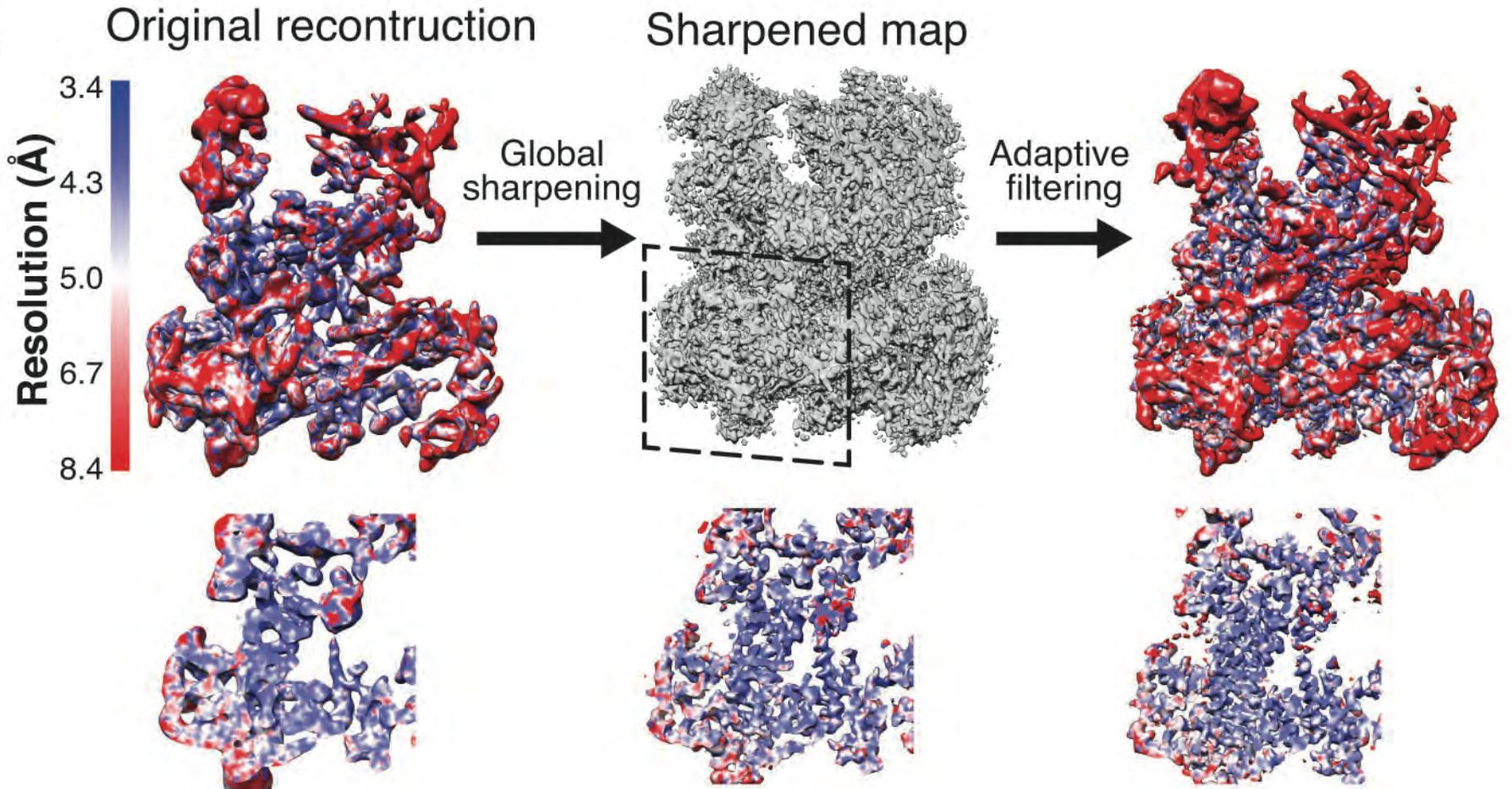


Use voxel-wise resolution estimate to low-pass filter map at local resolution

<https://gitlab.tudelft.nl/jakobi/locres>

see also: Cardone et al., *J Struct Biol* **184**, 226–236 (2013)

# Generation of adaptively filtered maps



Not taking care of varying contrast loss

# Amplitudes and image contrast

Image

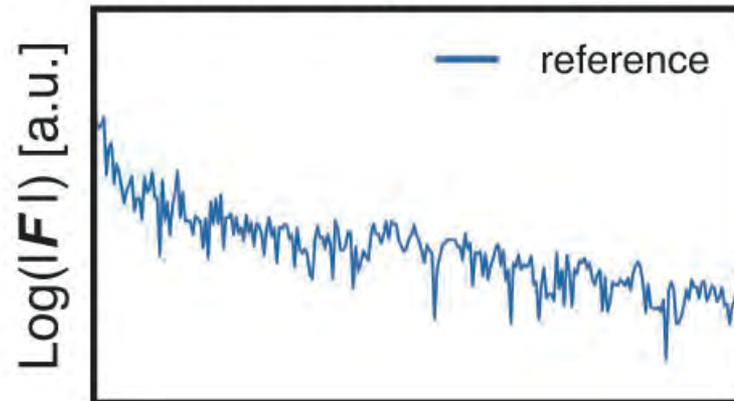


Relative scaling of low vs. high frequency amplitudes determines image contrast

Amplitudes

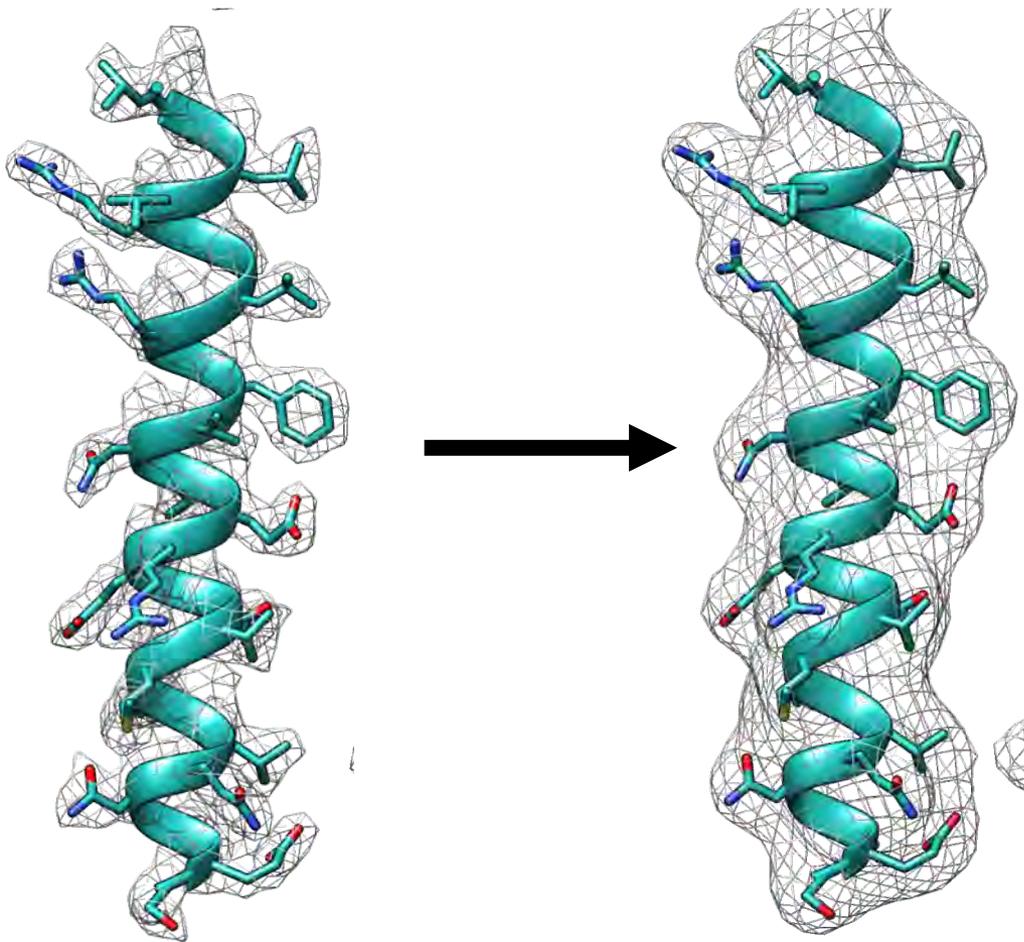


Radially averaged amplitudes

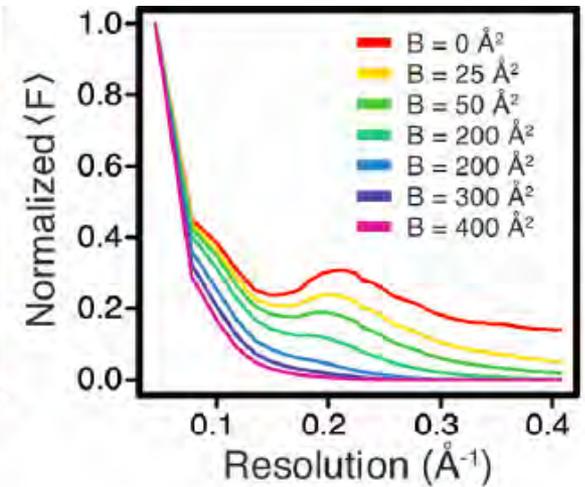


# Contrast loss

Variability owing to heterogeneity and computational inaccuracies during reconstruction cause blurring of the signal in the map → **contrast loss**

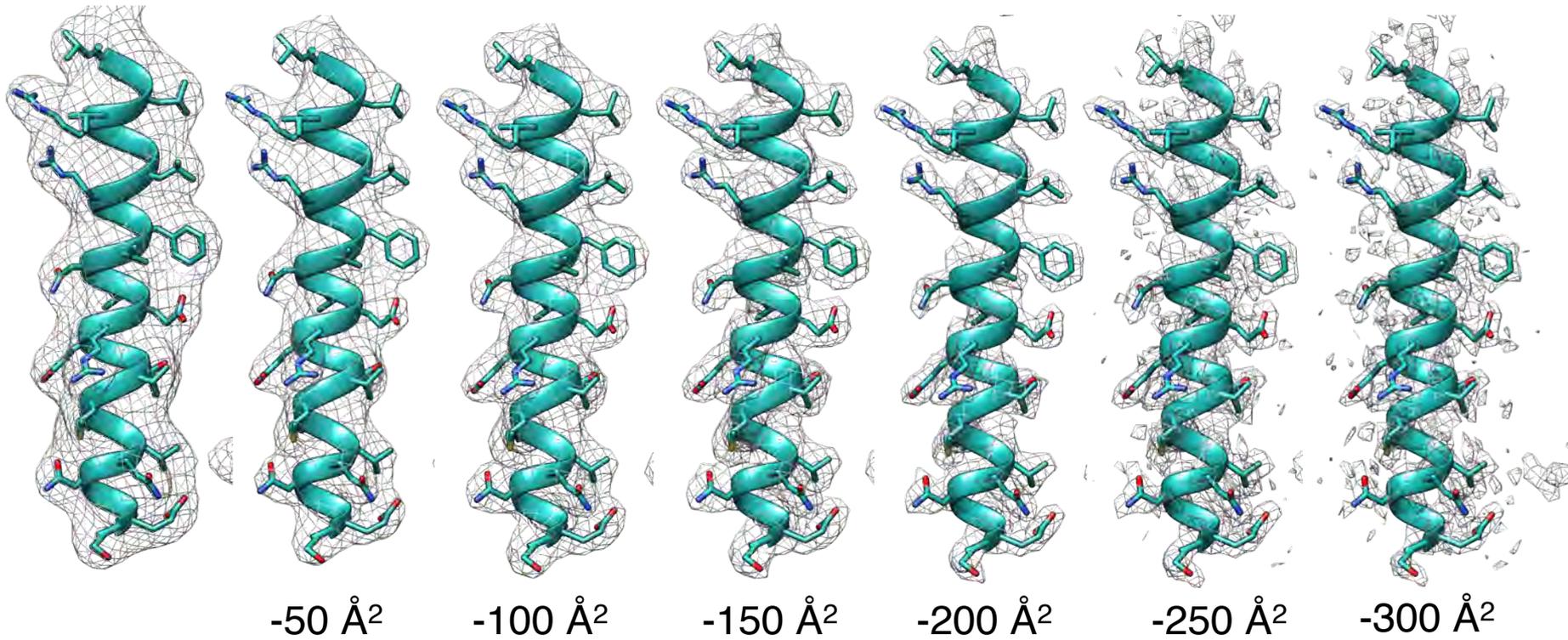


Contrast loss



# Map sharpening

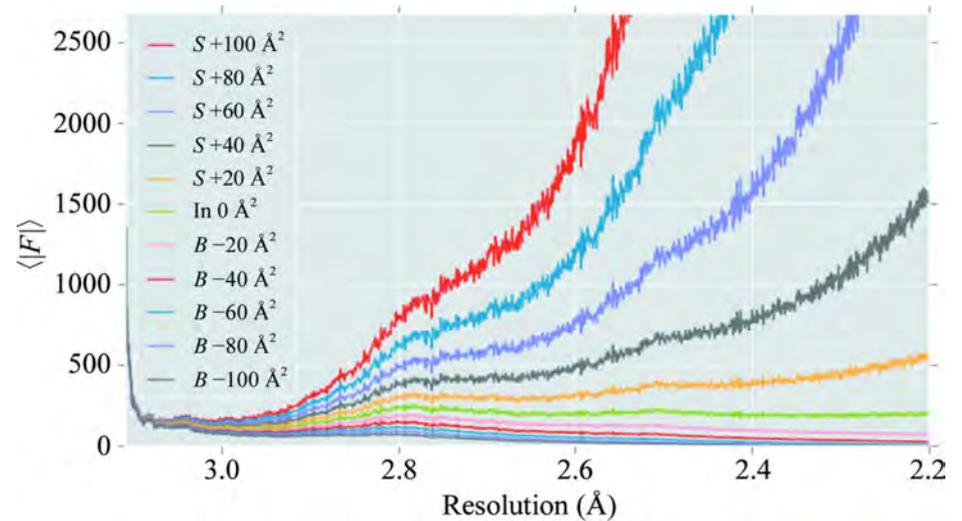
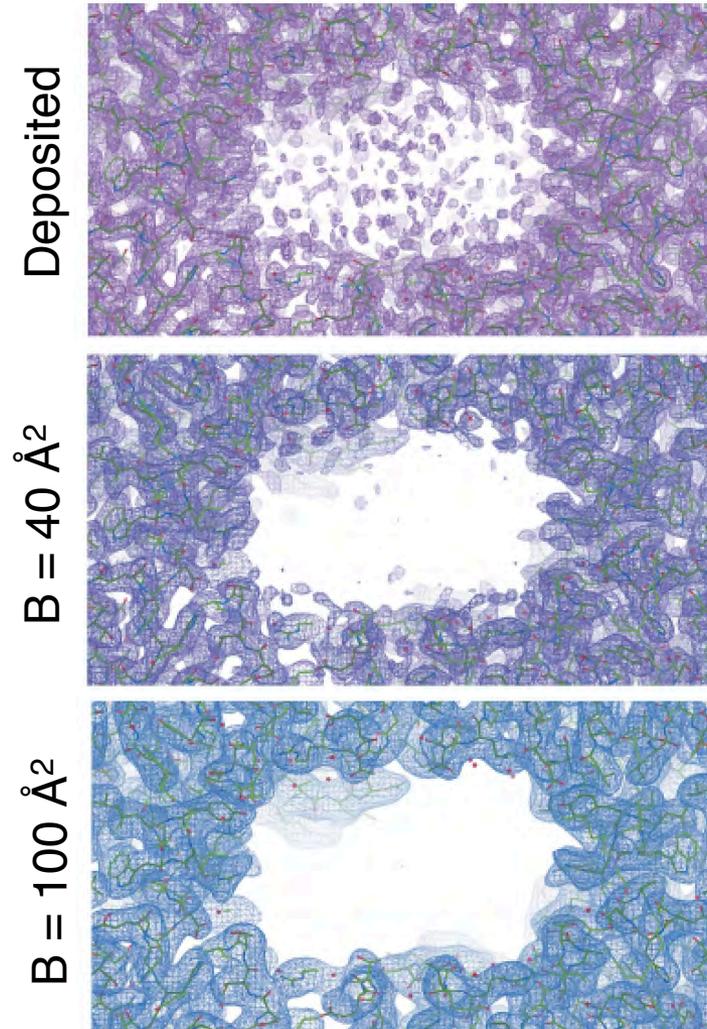
Restore amplitude contrast by sharpening with:  $F_{sharp} = F_{obs} \cdot e^{-B(1/d \min)^2}$



# Map blurring

Maps can be over-sharpened.

Blurring can be used to improve over-sharpened maps.

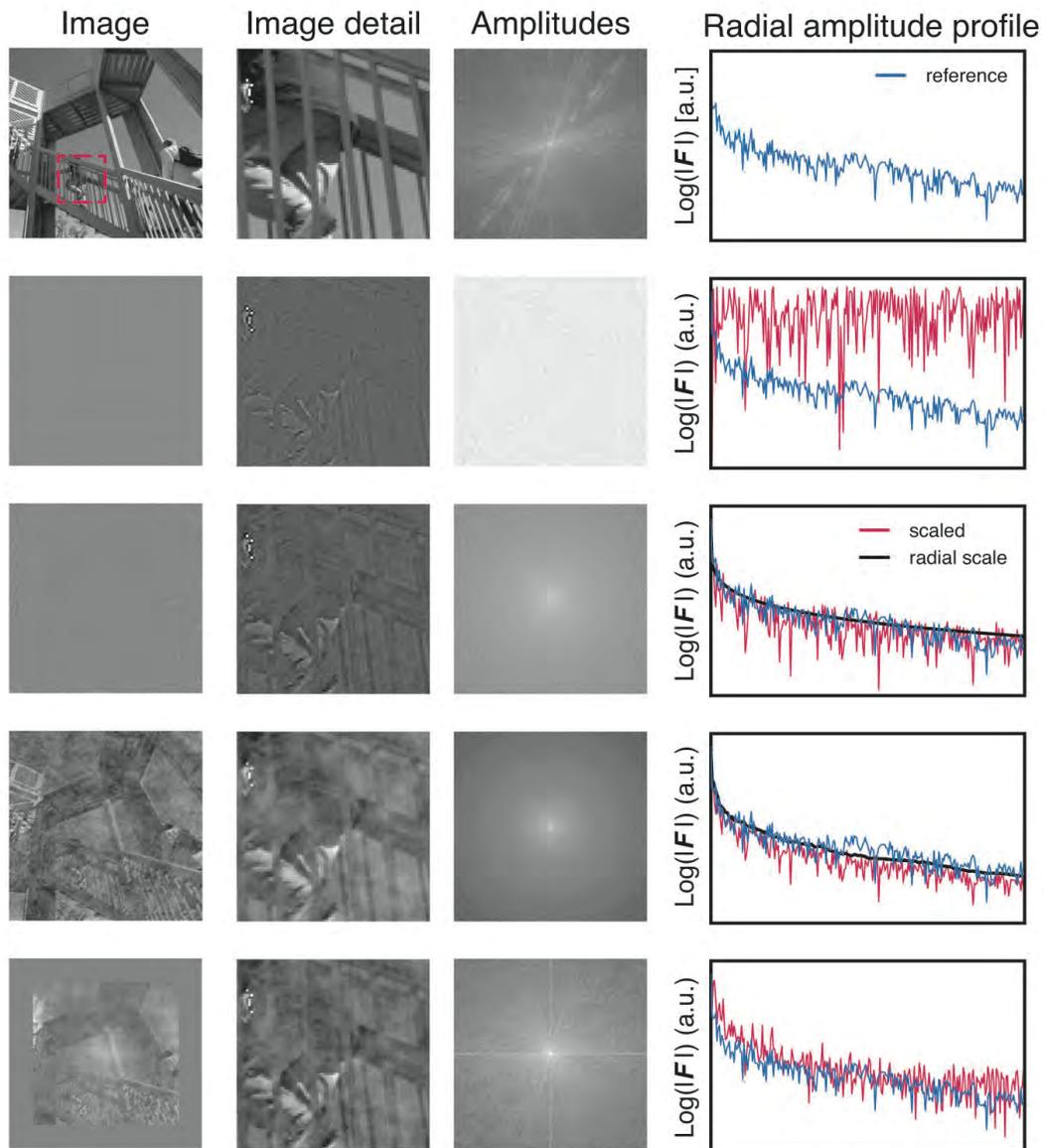


## Useful tools:

MRCtoMTZ in CCP-EM

Interactive sharpening in Coot

# Effects of amplitude scaling on image contrast



**No fall-off correction**

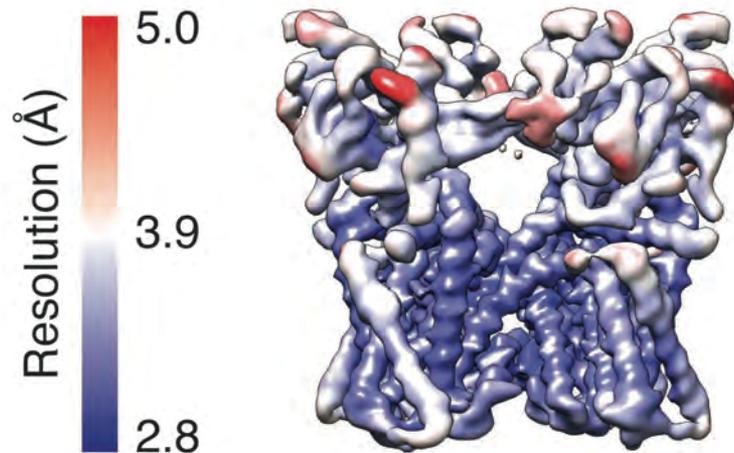
**Exponential fall-off correction**

**Reference fall-off correction**

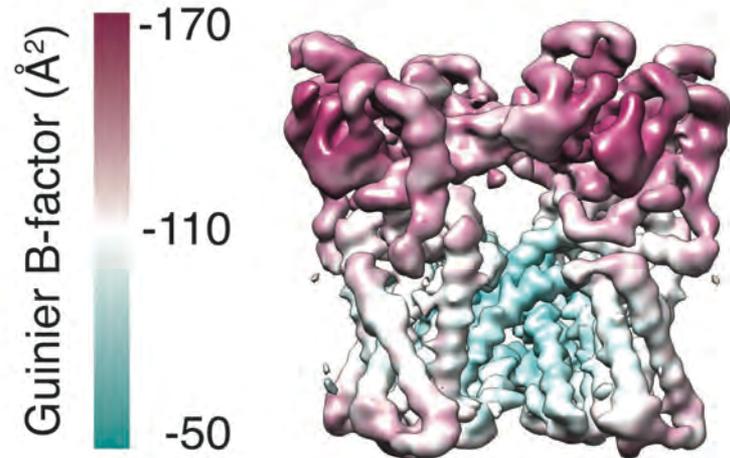
**Local reference fall-off correction**

# Local variation of map B-factors

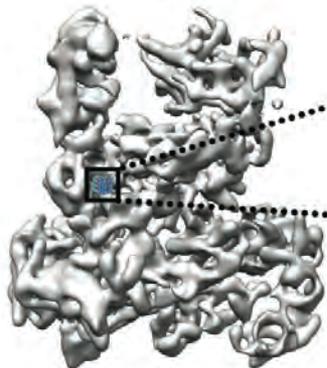
## Local resolution



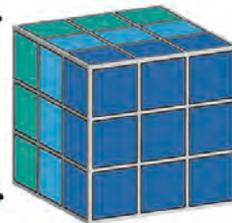
## B-factor



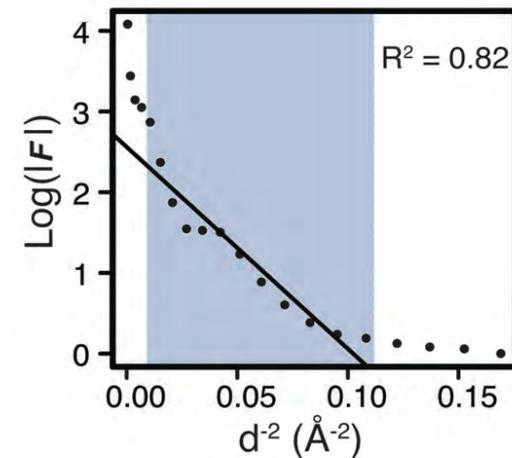
## original reconstruction



## Rolling window $n \times n \times n$ voxels

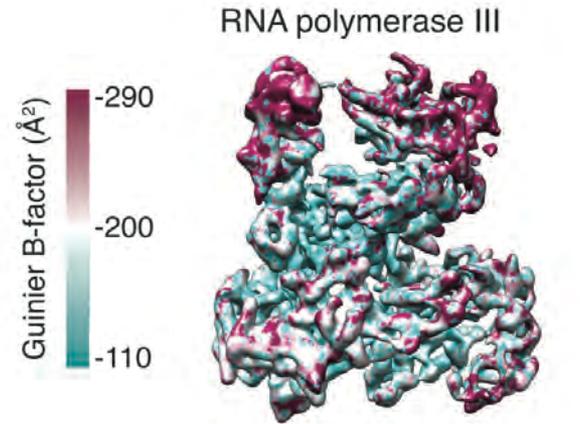
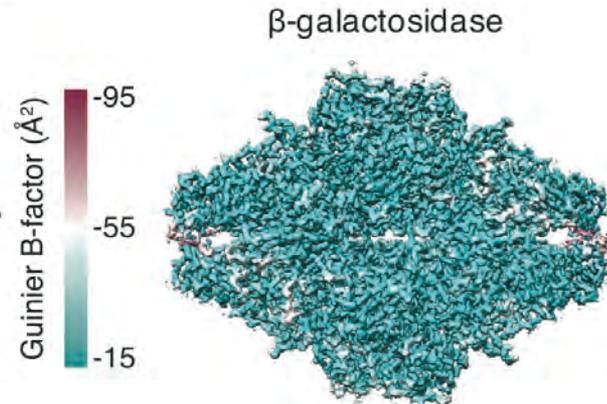
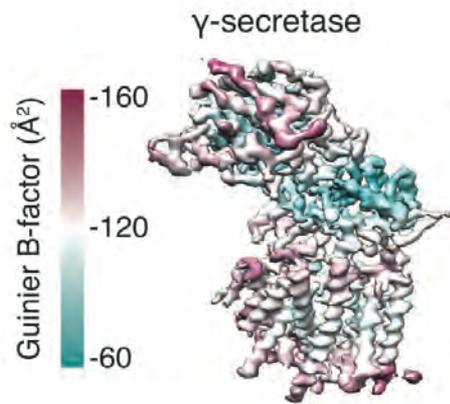


## Guinier plot

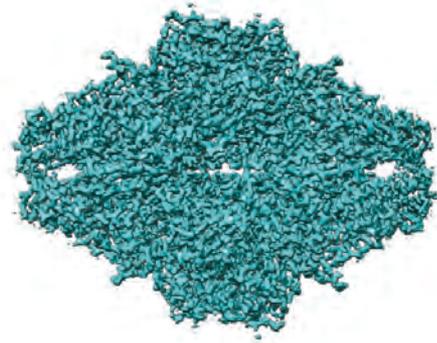
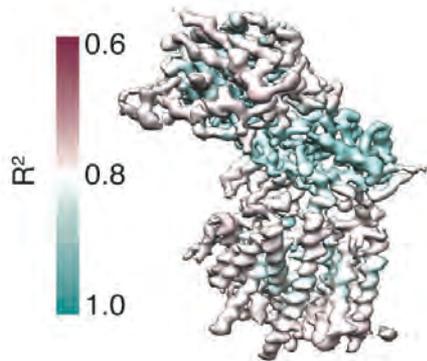


# Local B-factor variation

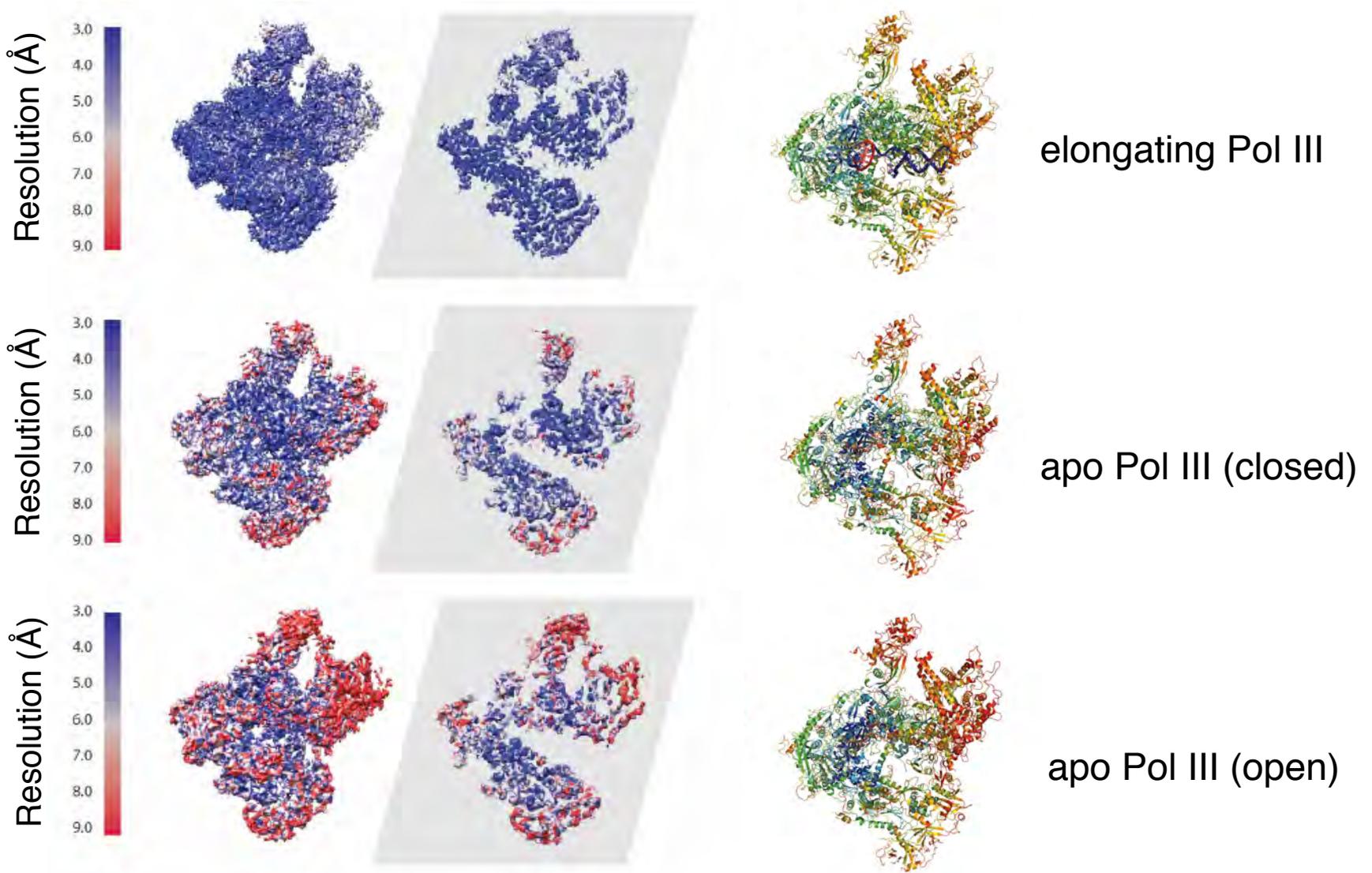
B-factor



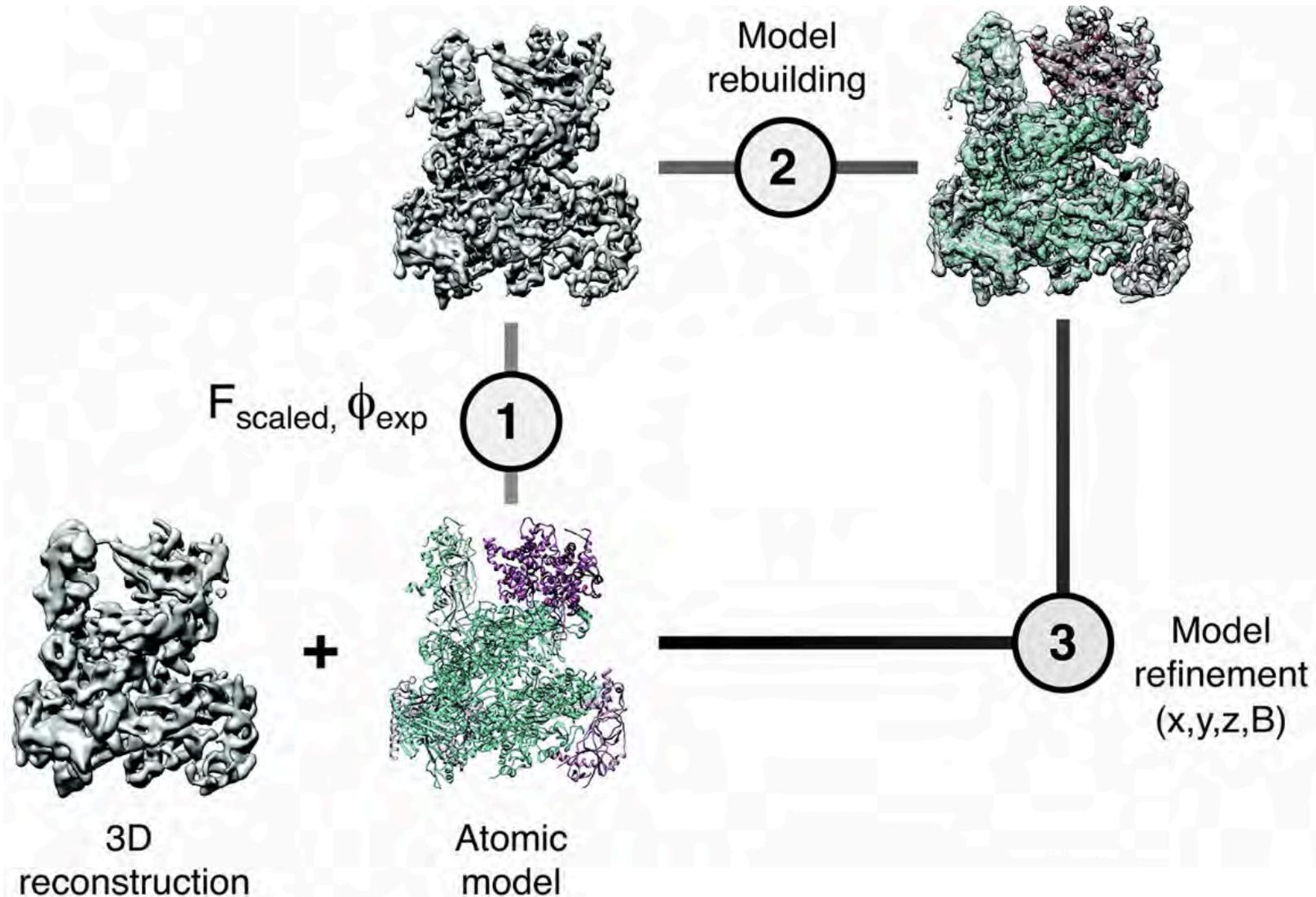
Goodness-of-fit



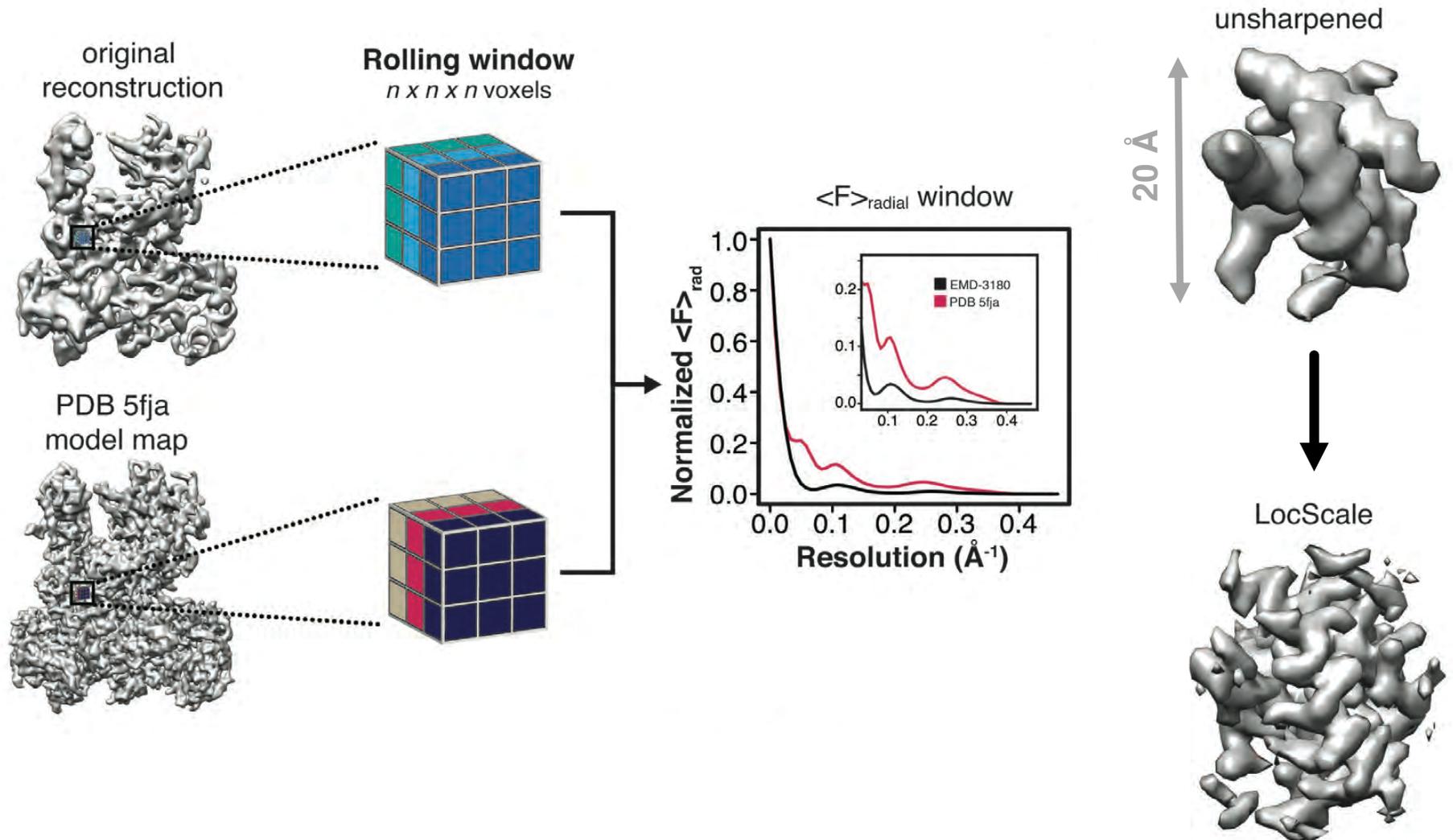
# Local resolution correlates with ADPs



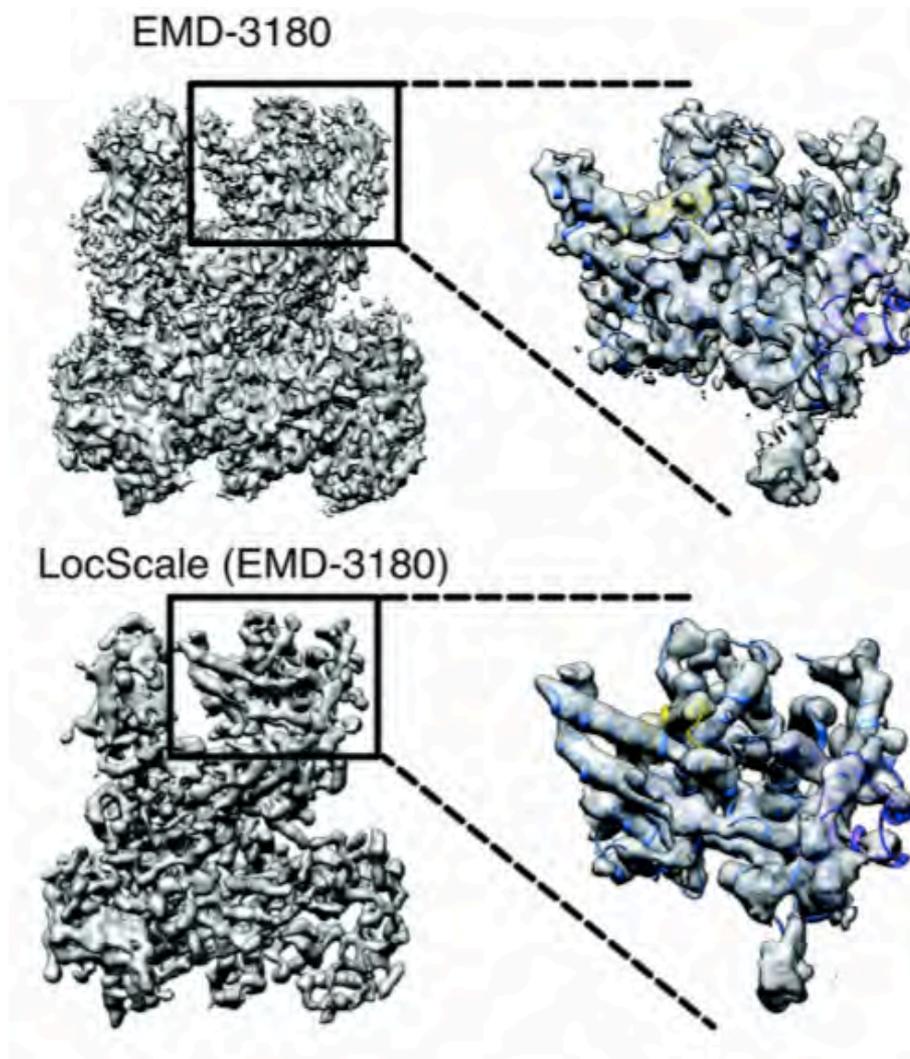
# Map sharpening by reference-based amplitude scaling



# Local sharpening by reference-based amplitude scaling



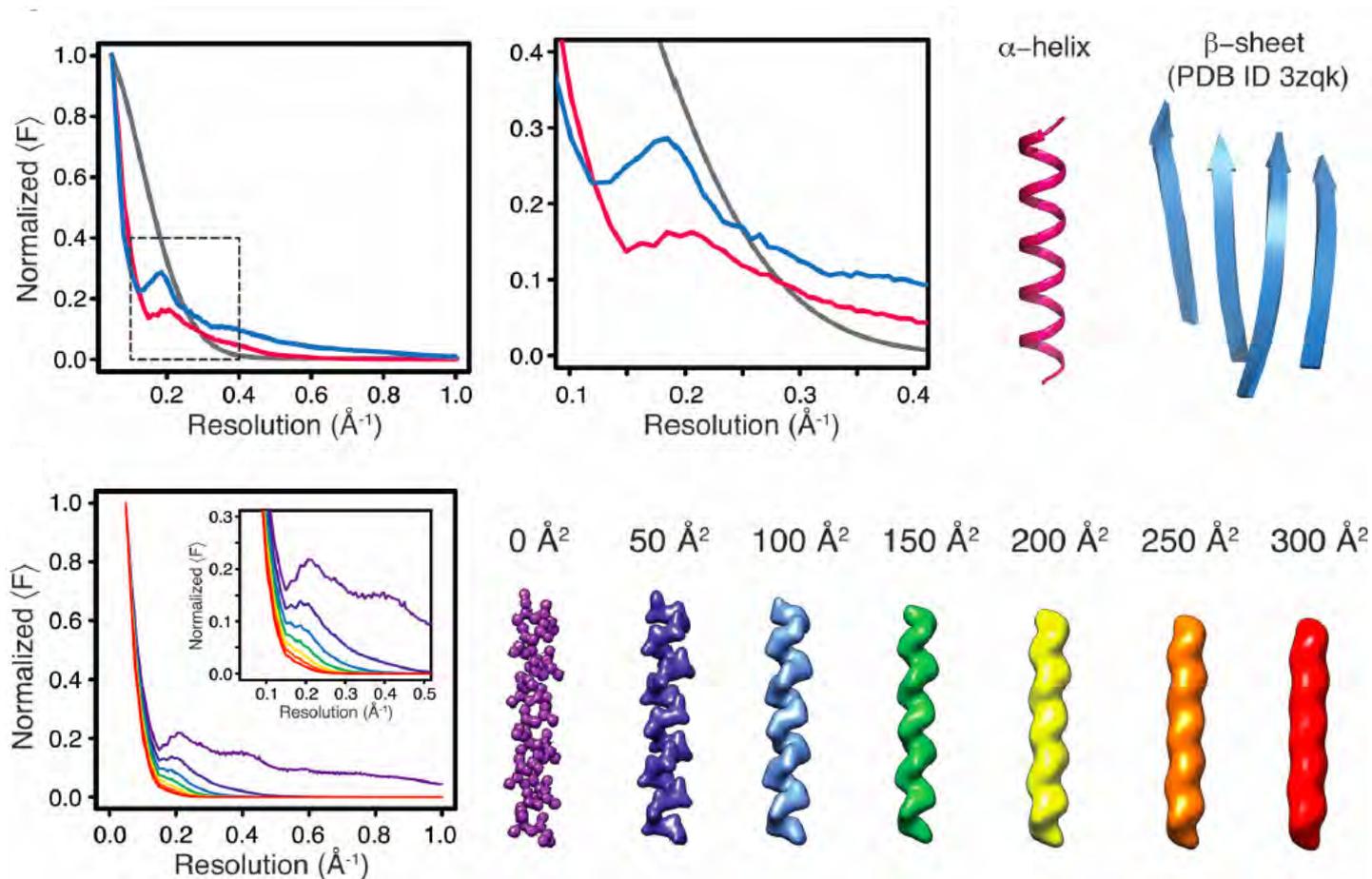
# Local sharpening by reference-based amplitude scaling



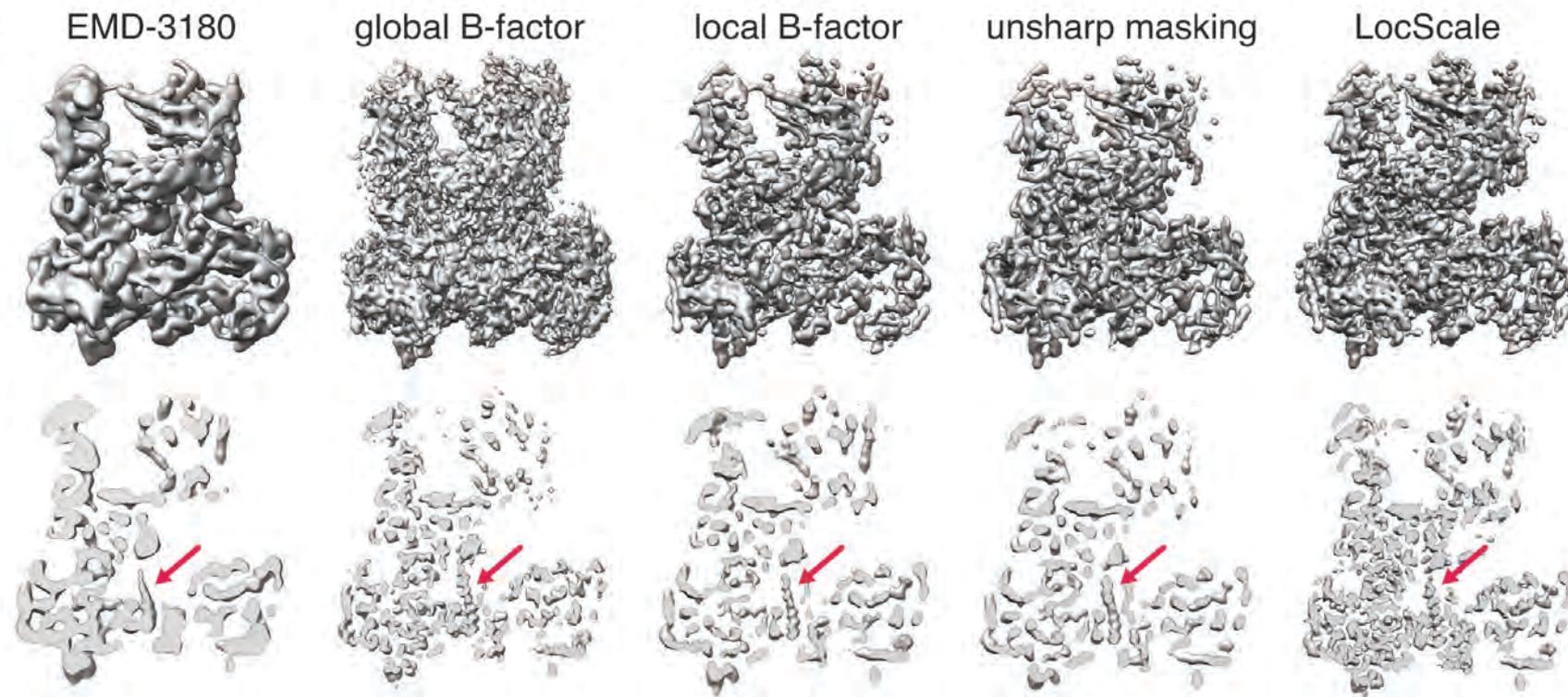
# How do we evaluate optimal sharpening?

- The map should provide maximum level of detail
- The map should show expected features of macromolecular structures

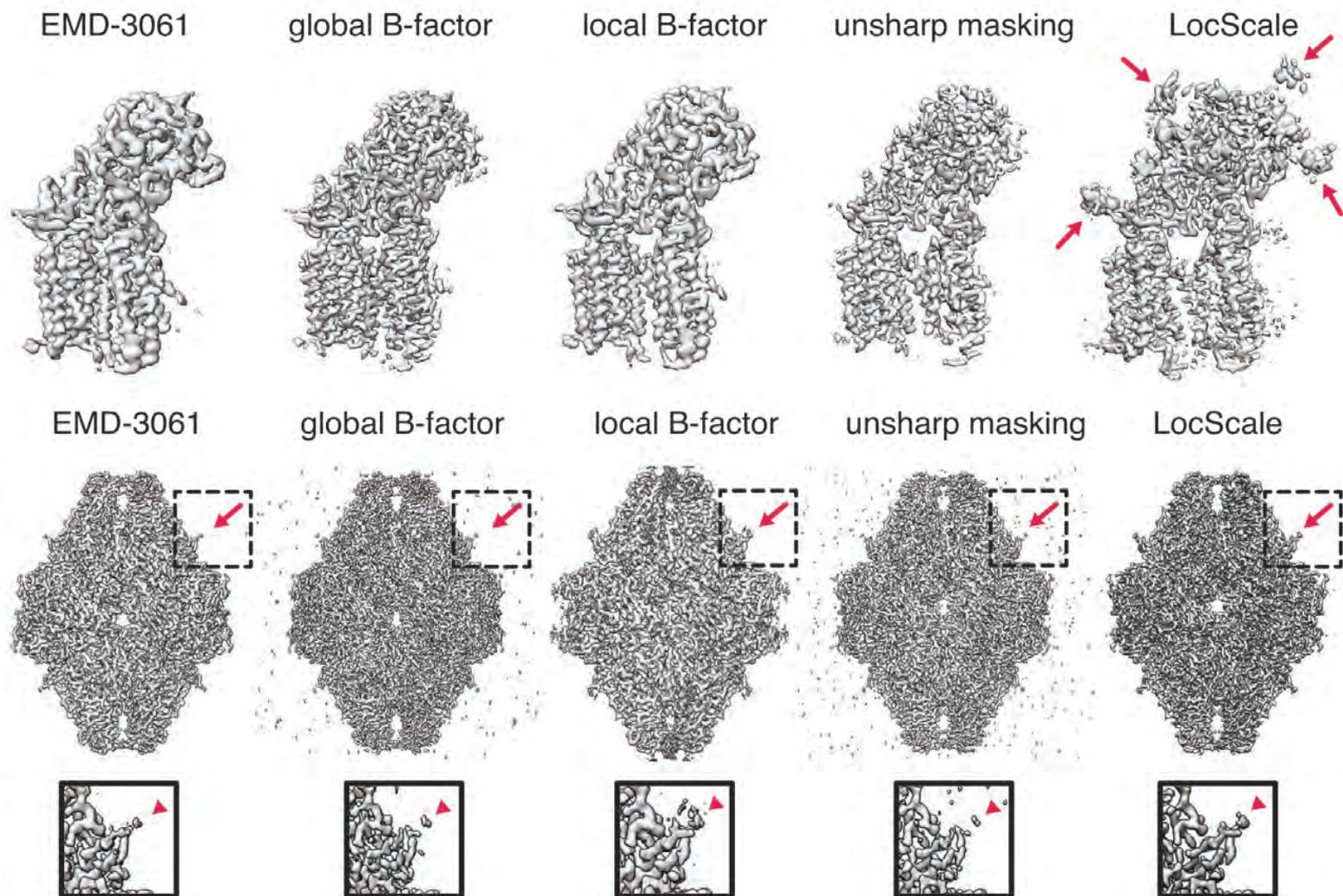
Secondary structure imposes characteristic deviations on amplitude profiles



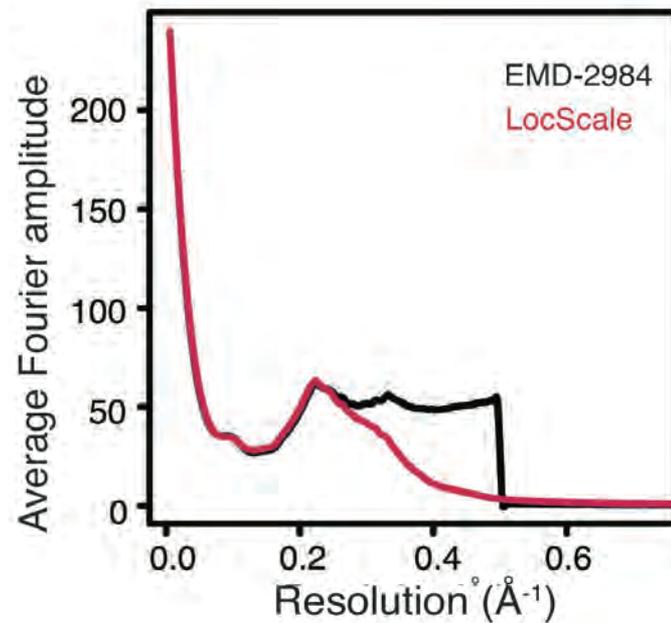
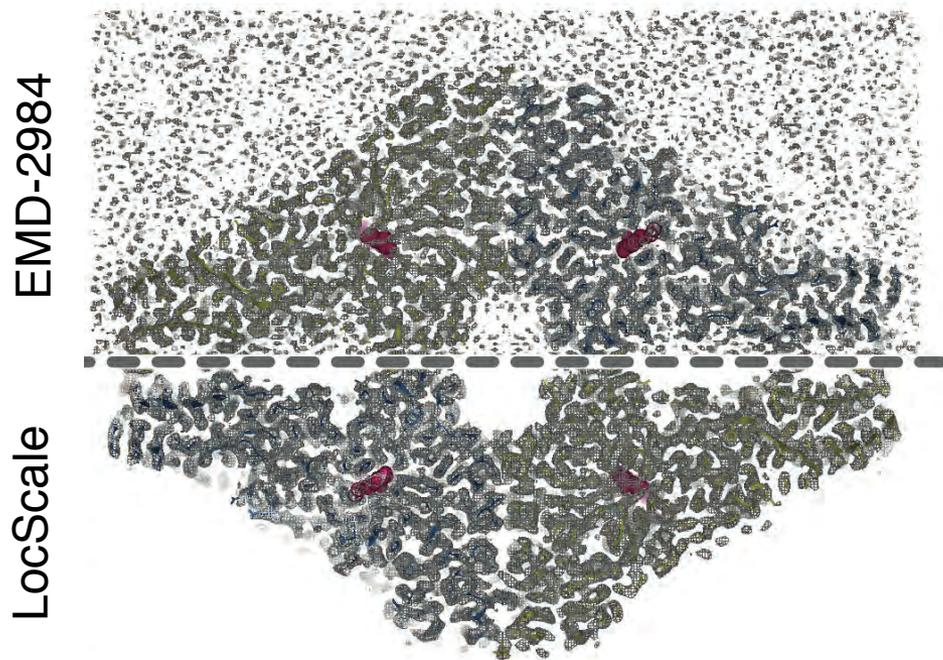
# Comparing different sharpening methods



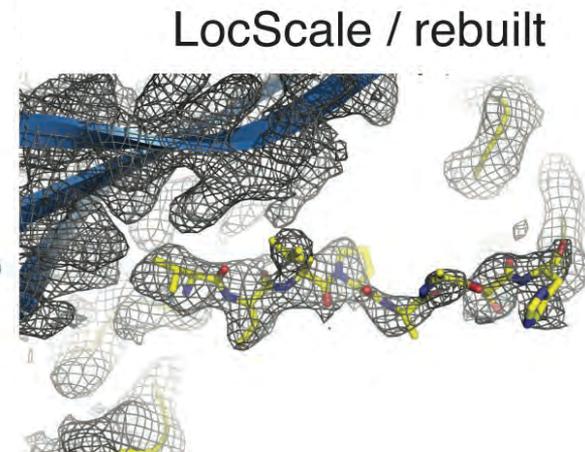
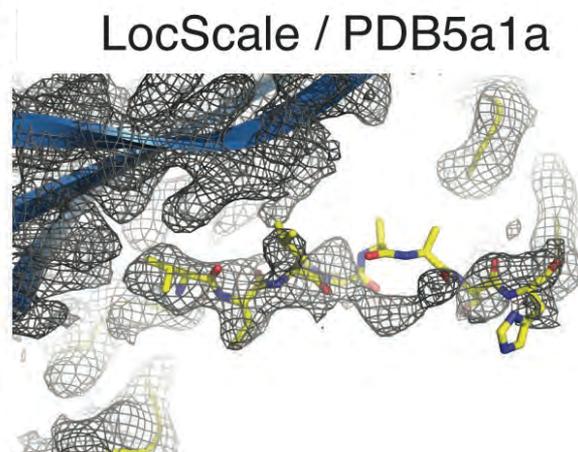
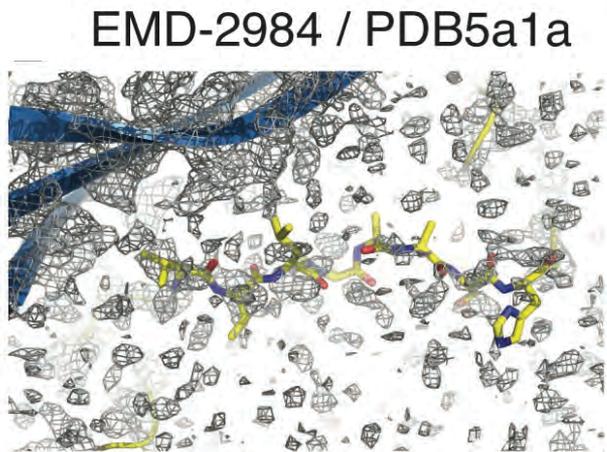
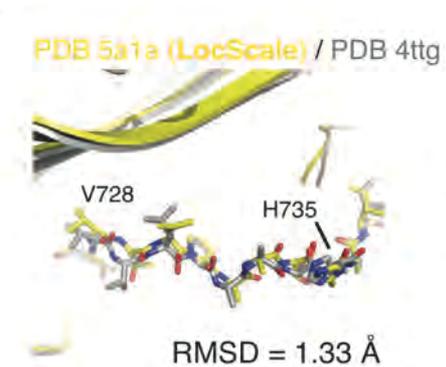
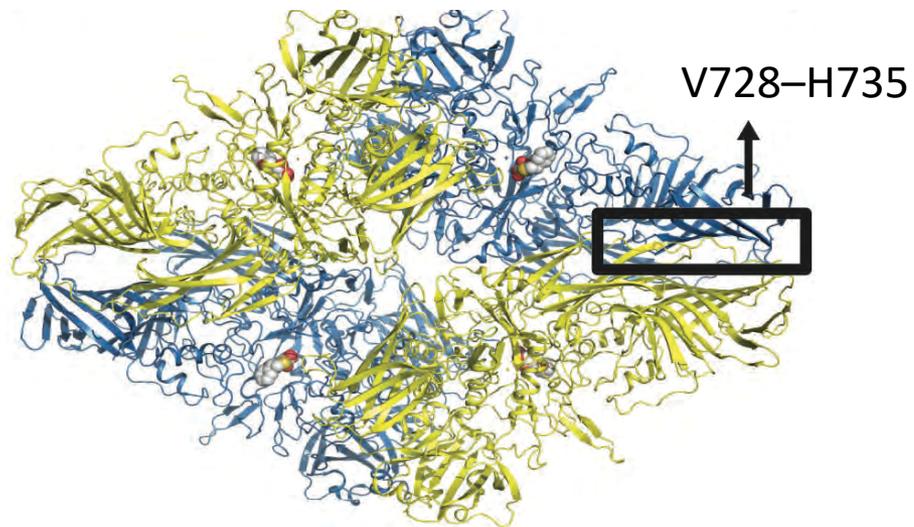
# Comparing different sharpening methods



# Example 1 – $\beta$ -galactosidase at 2.2 Å

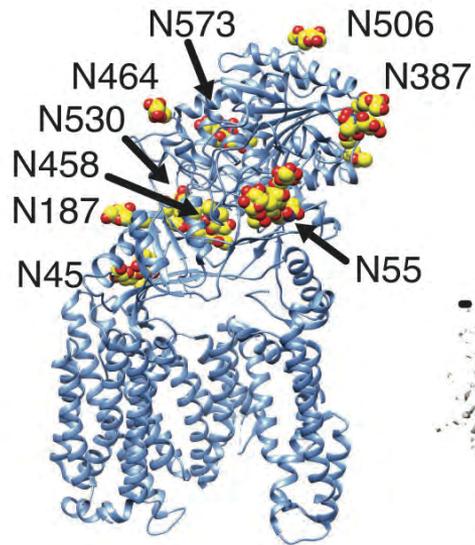


# Example 1 – $\beta$ -galactosidase at 2.2 Å

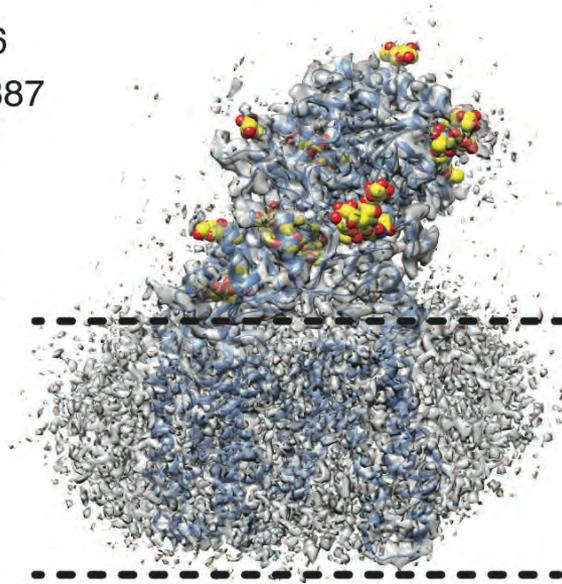


## Example 2 – $\gamma$ -secretase at 3.4 Å

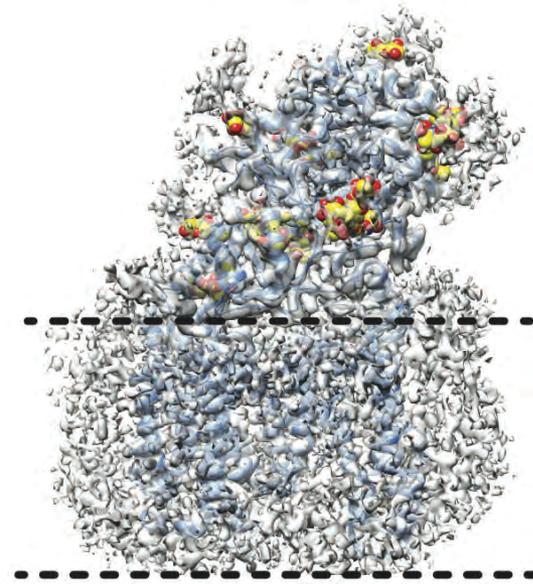
PDB 5a63  
( $\gamma$ -secretase)



EMD-3061  
( $\gamma$ -secretase)

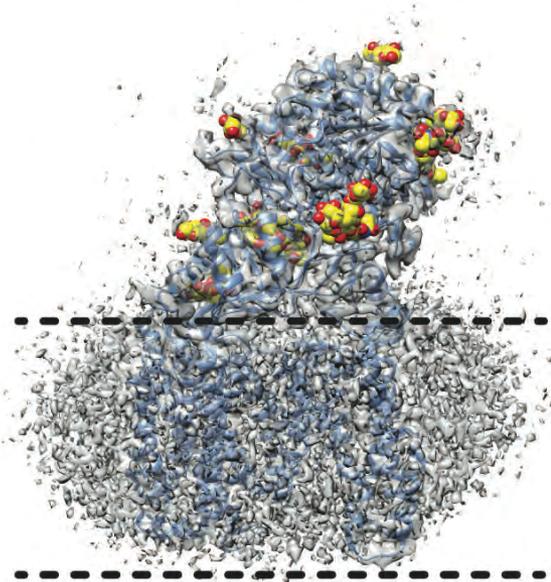


LocScale  
( $\gamma$ -secretase)

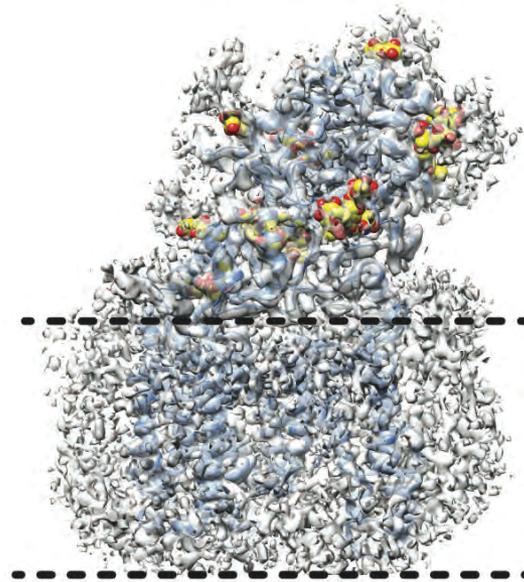


## Example 2 – $\gamma$ -secretase at 3.4 Å

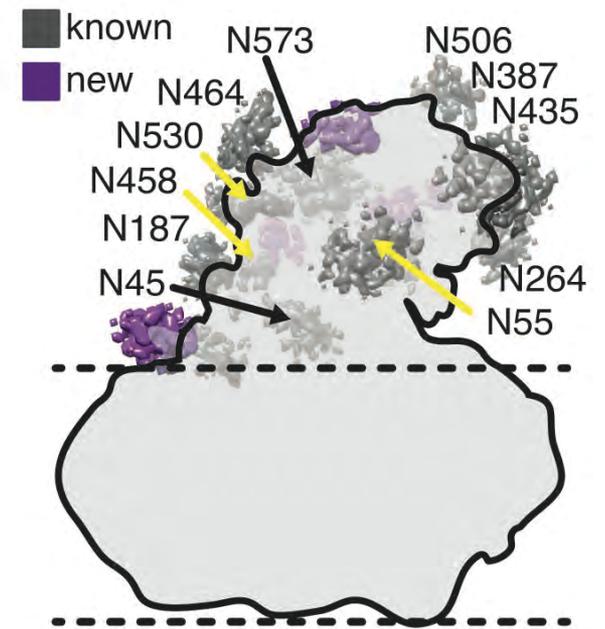
EMD-3061  
( $\gamma$ -secretase)



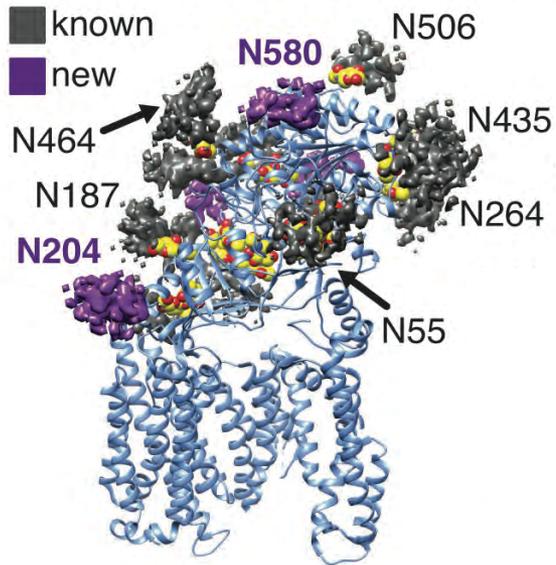
LocScale  
( $\gamma$ -secretase)



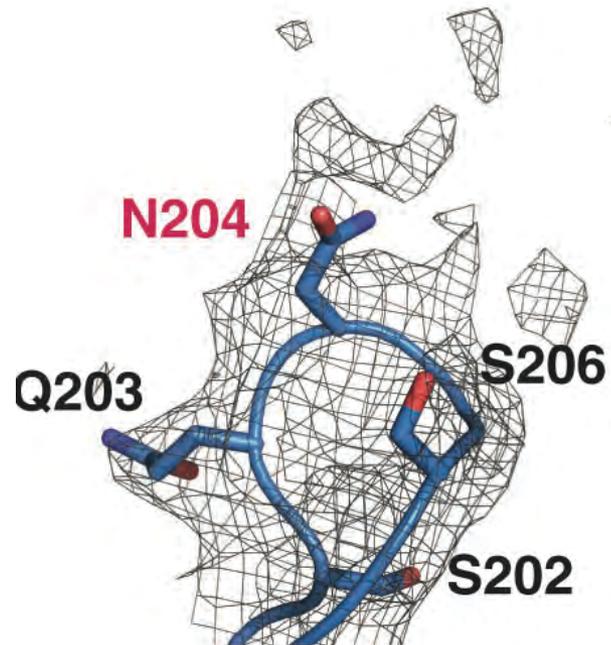
Difference density



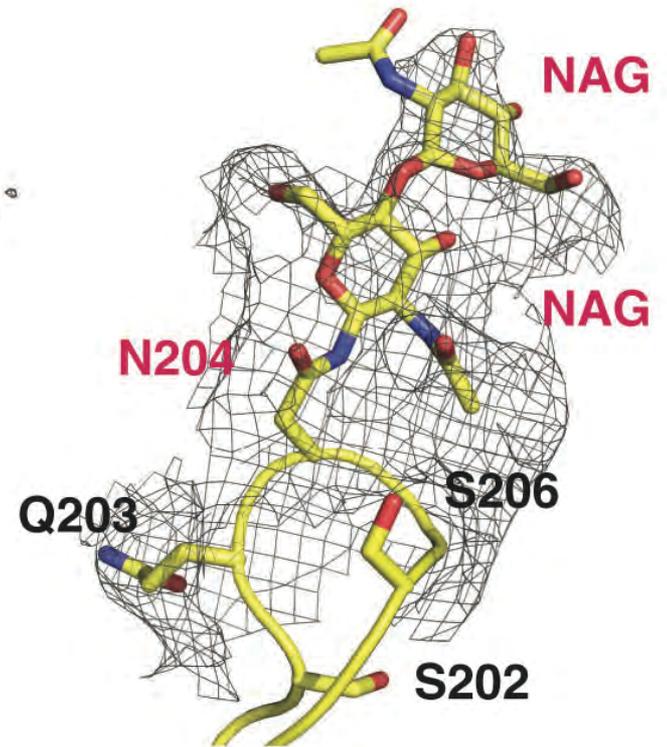
## Example 2 – $\gamma$ -secretase at 3.4 Å



EMD-3061



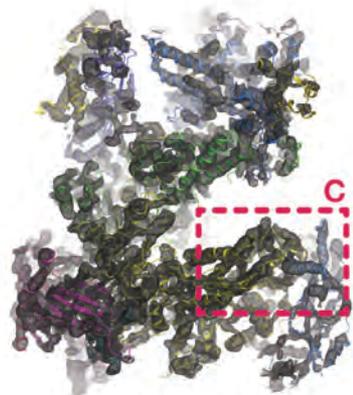
LocScale



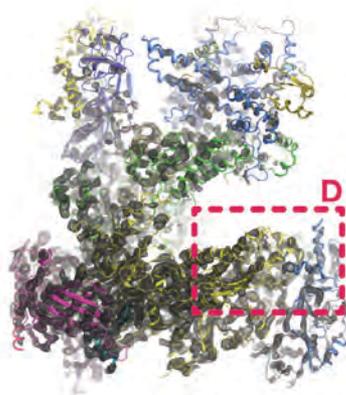
# Visualization & automatic model building

EMD-3180  
(LocScale)

EMD-3180



3.5 $\sigma$



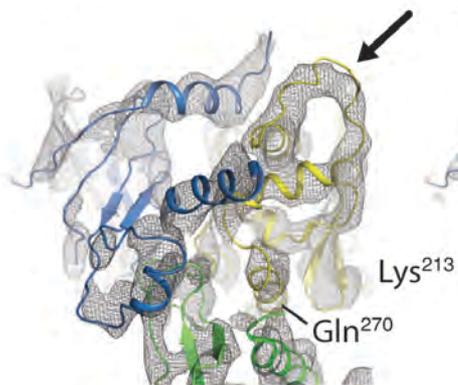
4.0 $\sigma$



3.0 $\sigma$



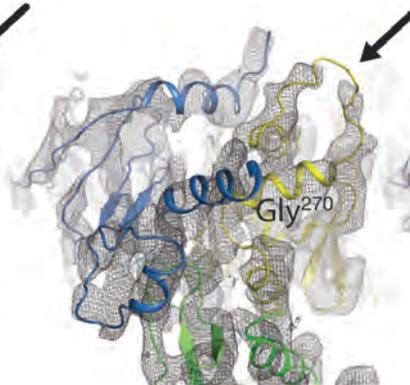
2.0 $\sigma$



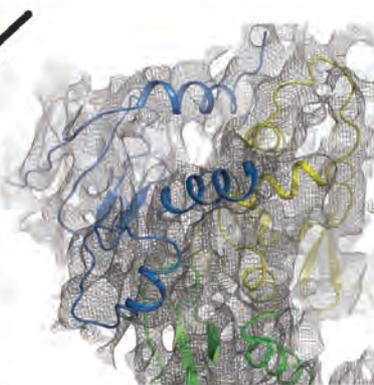
3.5 $\sigma$



4.0 $\sigma$



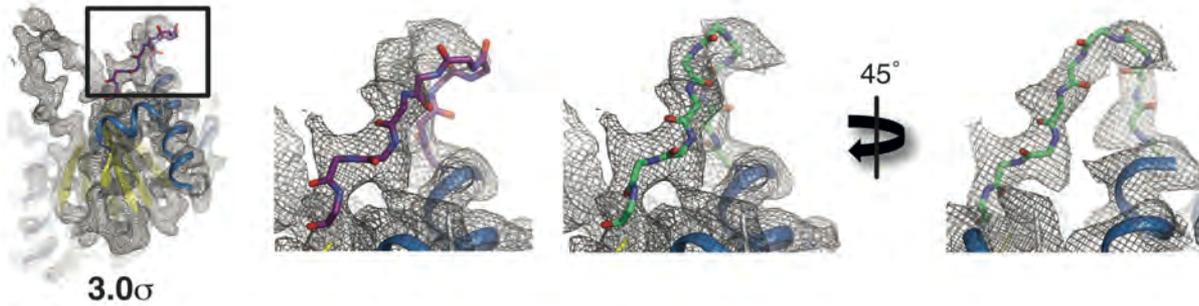
3.0 $\sigma$



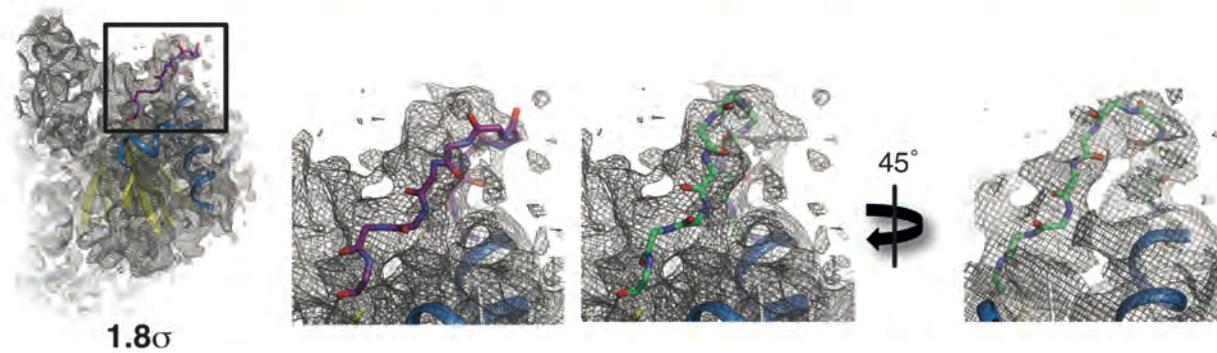
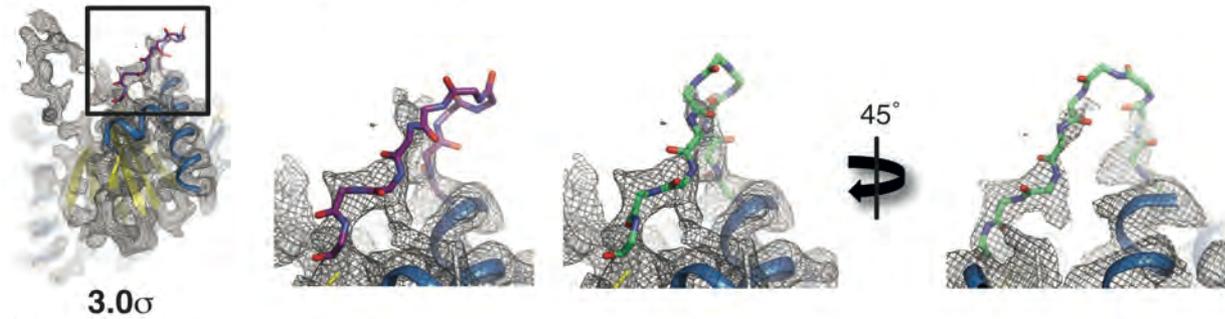
2.0 $\sigma$

# More examples

EMD-6287 (LocScale)



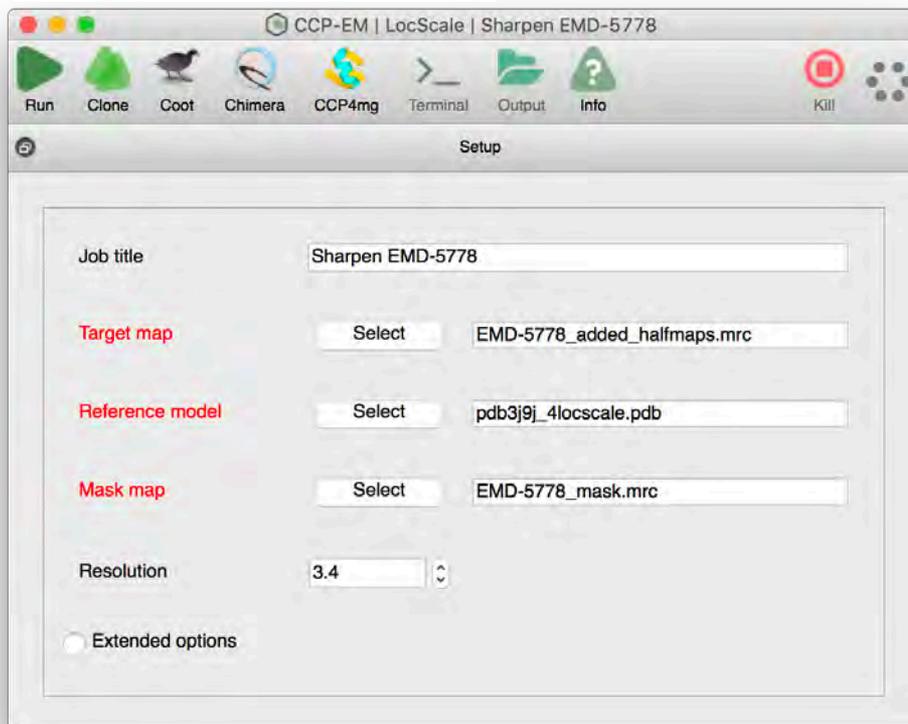
EMD-6287 (global sharpening)



# Availability and CCP-EM

Source and command line tool: <http://git.tudelft.nl/jakobi/locscale>

Much nicer (with GUI): <http://www.ccpem.ac.uk/download.php>



Thanks to:



**Tom Burnley**

GUI & pipeline



**Colin Palmer**

OpenMPI support  
mrcfile.py

## Other map sharpening tools

**Global sharpening:** Any 3D refinement program  
(Relion, cryoSPARC, EMAN2, SPHIRE, XMIPP, ...)

**Local sharpening:** `phenix.auto_sharpen` (part of PHENIX)  
Terwilliger et al., *Acta Cryst D Struct Biol* **74** (2018)

`LocalDeblur` (part of Scipion)  
Ramirez-Aportela et al., *BioRxiv* (2018)

**Interactive sharpening:** `MRC2MTZ` (part of CCP-EM)  
Burnley et al. *Acta Cryst D Struct Biol* **74** (2018)

`Coot`  
Emsley et al. *Acta Cryst D Struct Biol* **66** (2010)

# Validation of map features

- We want to determine which density features result from true signal and not from amplified noise

When is signal really signal?

**Hypothesis testing**

This can be problematic if you have to do many tests...

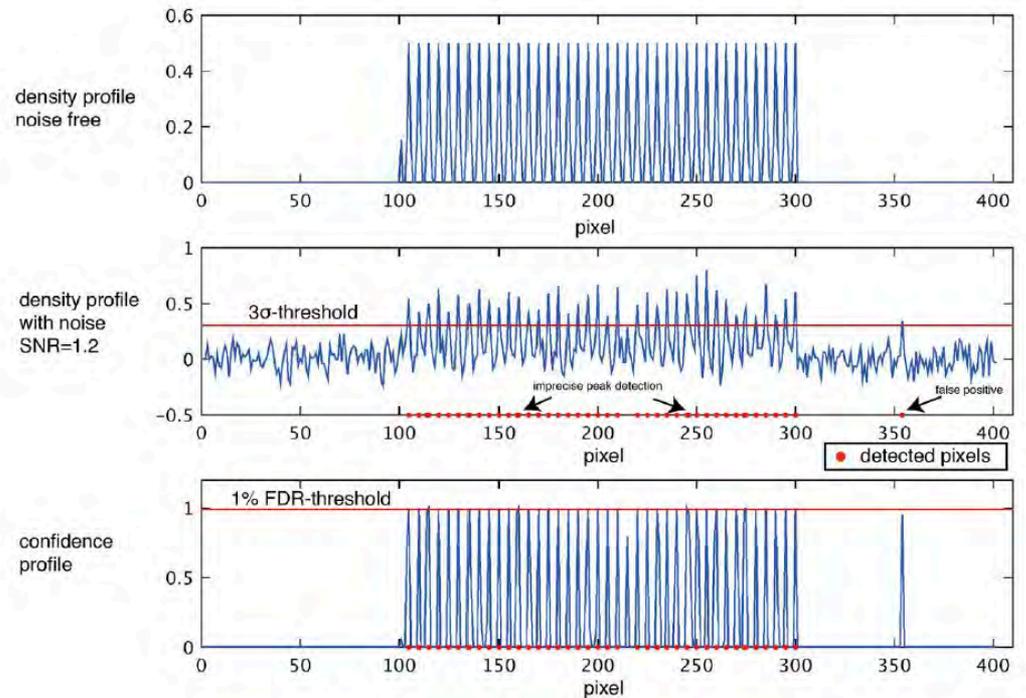
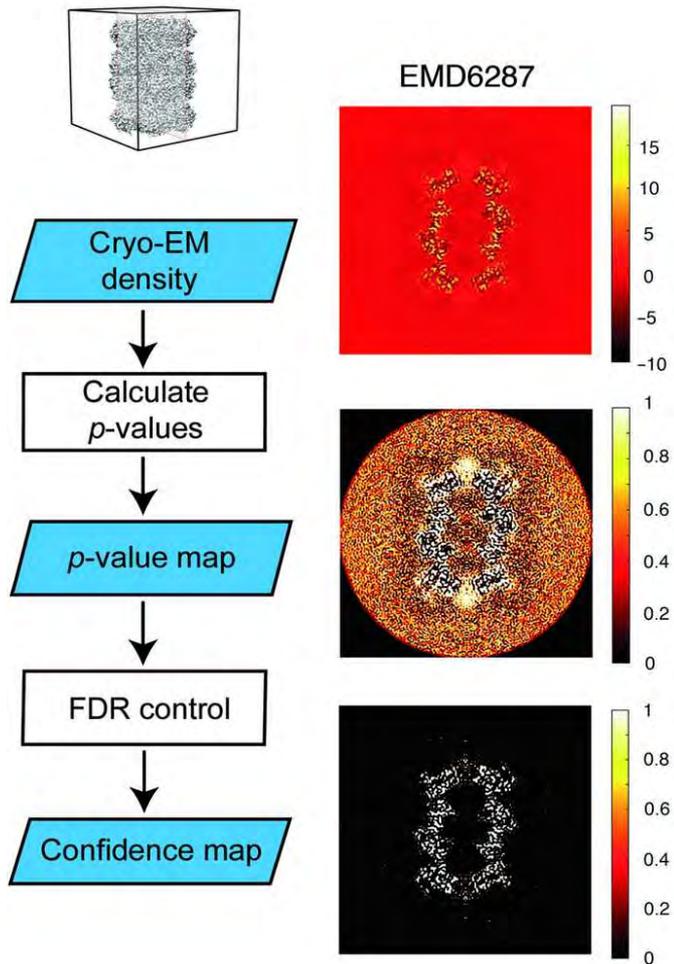
**Example:** 20 tests,  $p = 0.05$

What is the probability of observing at least one significant event just due to chance?

$$\begin{aligned} P(\text{at least one significant event}) &= 1 - P(\text{no significant event}) \\ &= 1 - (1-0.05)^{20} \\ &\approx \mathbf{0.64} \end{aligned}$$

→ There is a 64% chance to obtain one significant result just due to chance

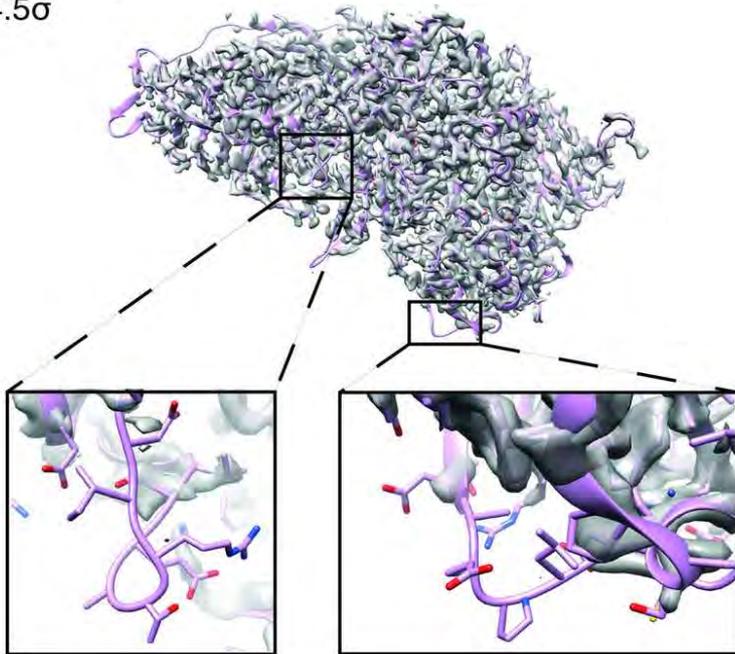
# Confidence maps by False Discovery Rate control



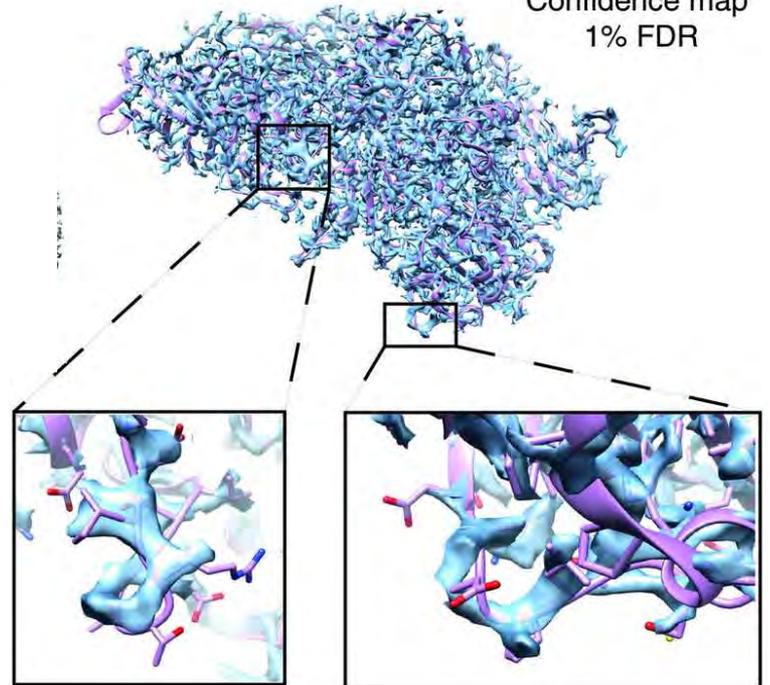
Maximilian Beckers

# Confidence maps allow detecting weak signal

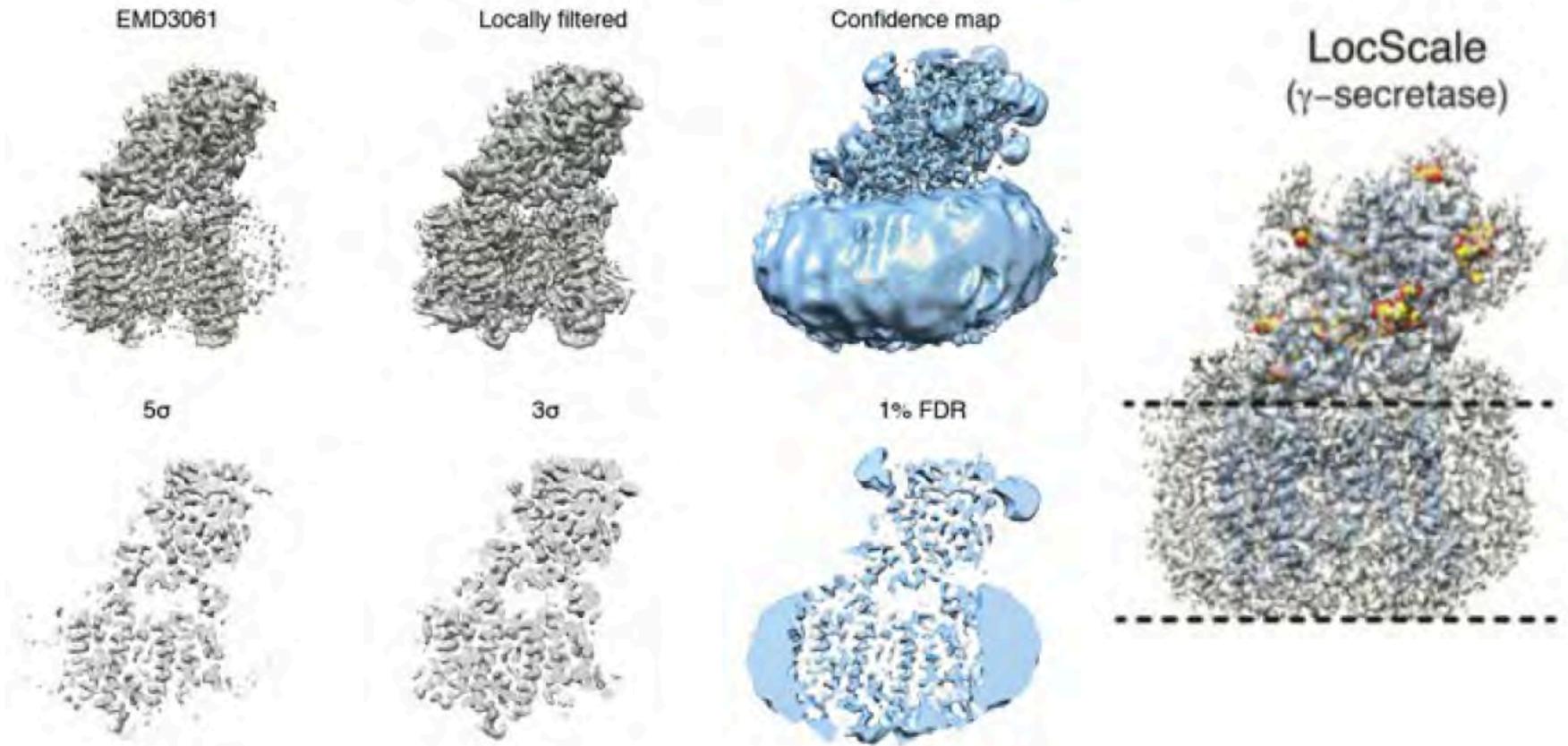
EM map  
 $4.5\sigma$



Confidence map  
1% FDR



# Confidence maps – $\gamma$ -secretase



# LAFTER maps: map denoising

- **Local Agreement Filter for Transmission EM Reconstructions**
- Compares half maps to identify shared features
- Preserves shared signal, suppresses noise
- Two-pass real space filter
- For map visualisation & model building – not refinement

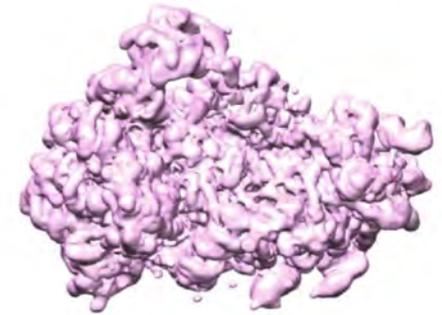
**Christopher Aylett**

EMD-3048

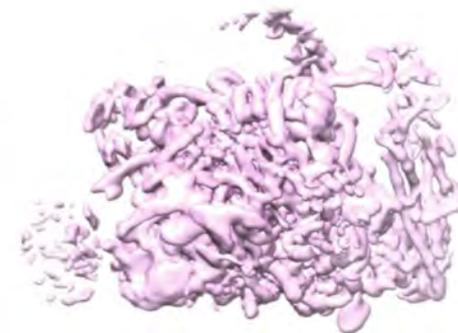
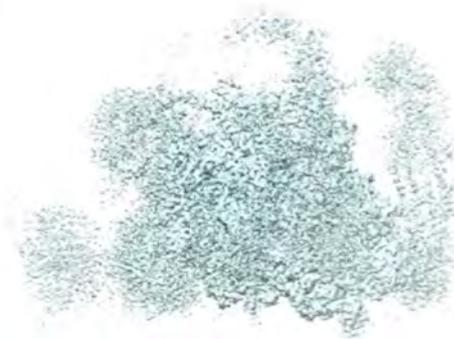
Original map



LAFTER map

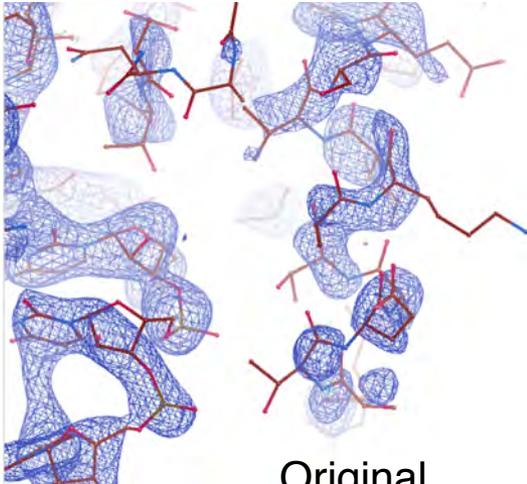


EMD-6721

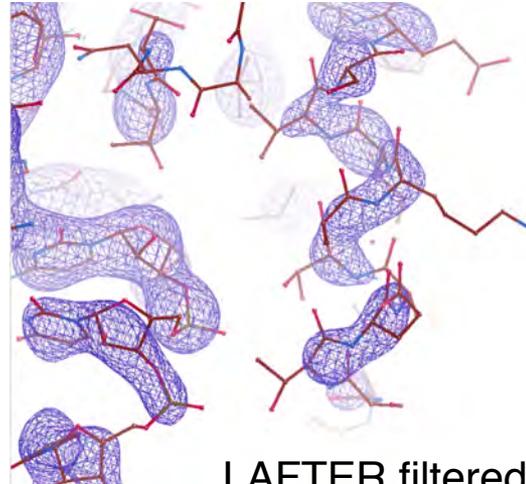


# Effect of LAFTER

**High contour:** strong features remain similar

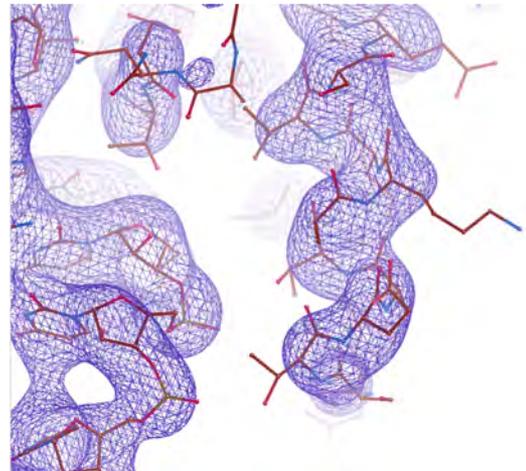
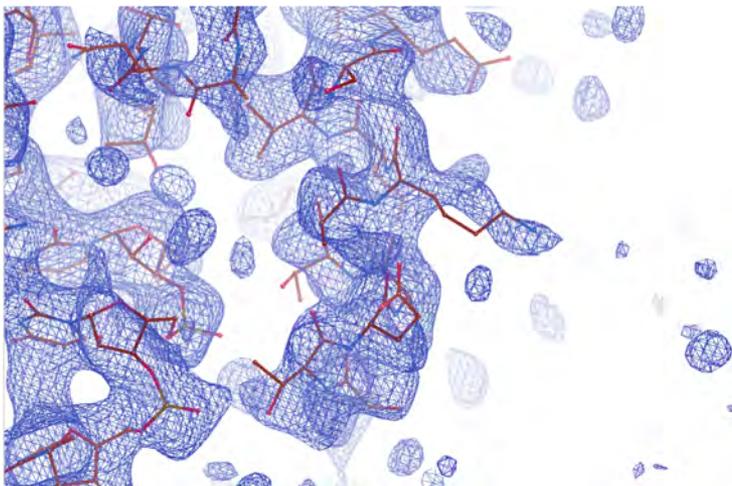


Original



LAFTER filtered

**Low contour:** weak features are very different. LAFTER removes noise



EMD-2847

Slide courtesy Colin Palmer

# Validation

With great insight comes great responsibility...

Structure  
Resource

## High-Resolution Cryo-EM Maps and Models: A Crystallographer's Perspective

Alexander Wlodawer,<sup>1,4,\*</sup> Mi Li,<sup>1,2</sup> and Zbigniew Dauter<sup>3</sup>

<sup>1</sup>Protein Structure Section, Macromolecular Crystallography Laboratory, National Cancer Institute, Frederick, MD 21702, USA

<sup>2</sup>Basic Science Program, Leidos Biomedical Research, Frederick National Laboratory for Cancer Research, Frederick, MD 21702, USA

<sup>3</sup>Synchrotron Radiation Research Section, Macromolecular Crystallography Laboratory, NCI, Argonne National Laboratory, Argonne, IL 60439, USA

<sup>4</sup>Lead Contact

\*Correspondence: [wlodawer@nih.gov](mailto:wlodawer@nih.gov)

<http://dx.doi.org/10.1016/j.str.2017.07.012>

The screenshot shows the EMDDataResource website interface. At the top left is the EMDDataResource logo with the tagline "Unified Data Resource for 3DEM". To the right of the logo is the text "EM Validation Challenges". Further right are "Login" and social media icons for Facebook and RSS. Below this is a dark blue navigation bar with the following menu items: "ABOUT", "MODEL METRICS CHALLENGE", "PREVIOUS CHALLENGES", and "NEWS". The main content area has a "Home" link. The primary heading is "2019 Model Metrics Challenge", with a sub-heading "last update: July 17, 2019". Below this is a "Quick Links" section divided into two columns: "Info" and "Results".

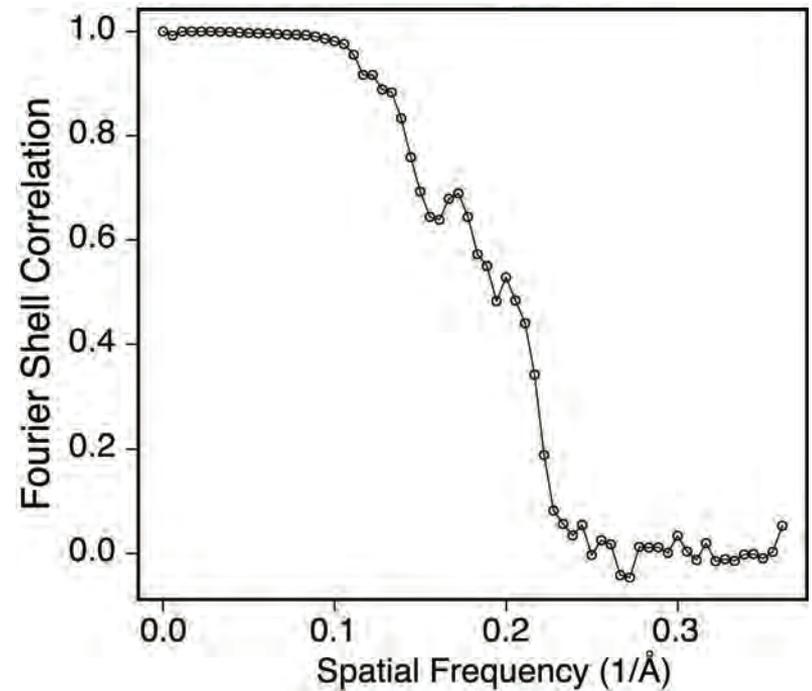
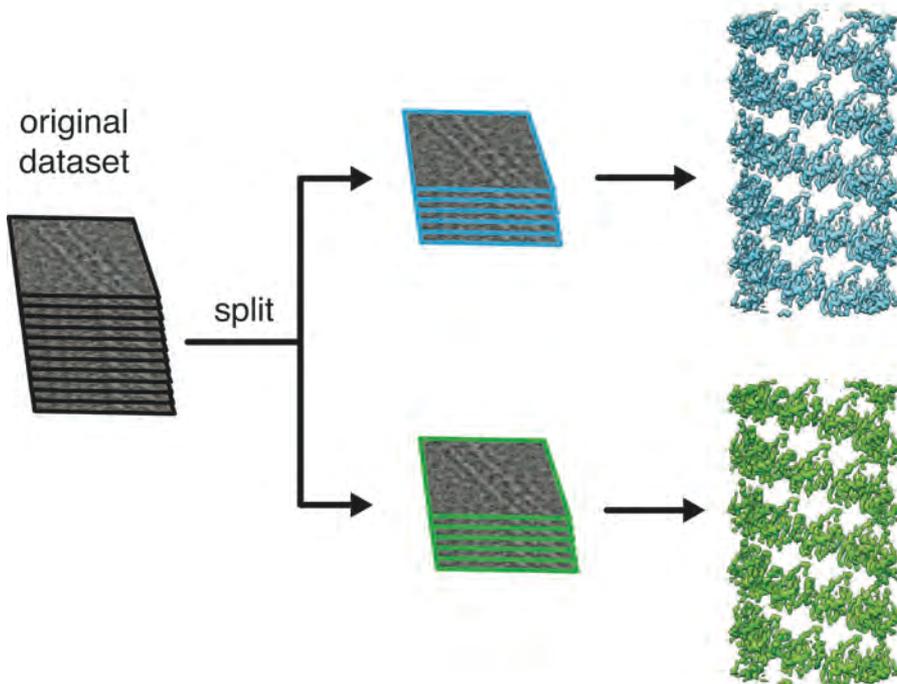
Info	Results
<ul style="list-style-type: none"><li>Announcement</li><li>Timeline</li><li>FAQ</li><li>Deposition form (closed)</li><li>Modeller Teams with codes</li><li>June 13-15 Face-to-Face meeting</li></ul>	<ul style="list-style-type: none"><li>Submitted Models Access via: <a href="#">FTP</a> /// <a href="#">Dropbox</a></li><li>Summary statistics</li><li>Model Compare Pipeline Analysis</li><li>Model Compare Data Repositorium (files used in pipeline)</li><li>Submitted models full metadata and team codes (xlsx)</li></ul>

# Model vs. map cross-validation

Random free R set approximation (MX) is not valid for EM

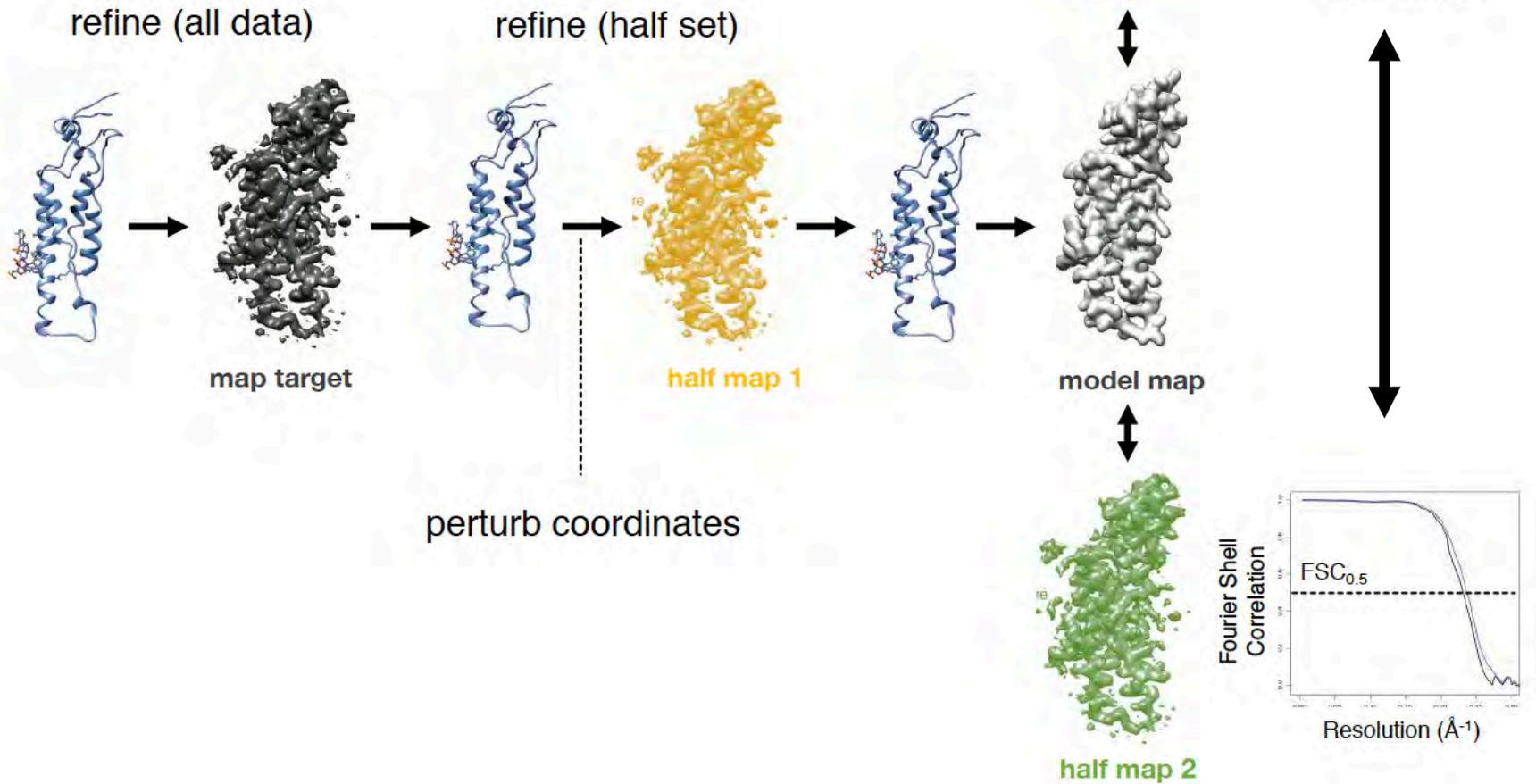
→ select “free band” not used in EM model refinement

Validate model against  
“free” band



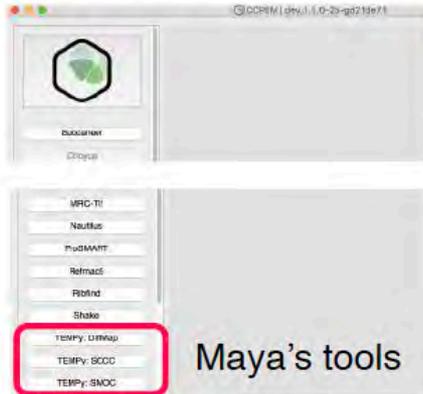
# Model vs. map cross-validation

FSC validation against independent half maps



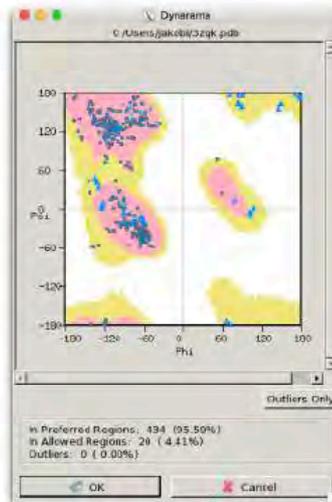
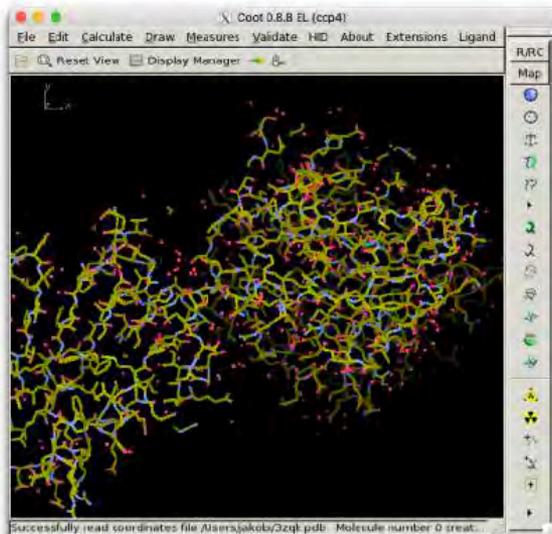
# Useful tools for model validation

## TemPy/CCP-EM ([www.ccpem.ac.uk](http://www.ccpem.ac.uk))



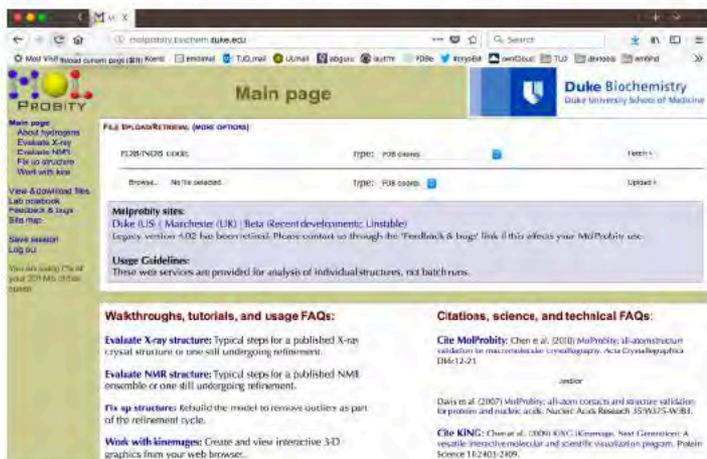
- Many useful validation tools
- Difference maps

## Coot (<https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/cool/>)



# Useful tools for model validation

Molprobrity (<http://molprobrity.biochem.duke.edu/>)



## Comprehensive stereochemistry validation

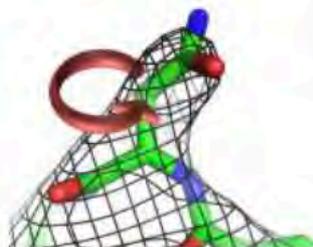
Helps you to resolve:

- Steric clashes
- Geometry outliers
- NQ flips

EMRinger (<http://emringer.com/>)



Assesses quality of model-to-map fit of models from cryo-EM using map value interpolation of side chain rotameric positions.



# Reporting validation statistics

“Table 1”

Extended Data Table 1 | Refinement statistics

	Elongating Pol III	Apo Pol III (closed clamp)	Apo Pol III (open clamp)
<b>Model composition</b>			
No. of chains	17+3	17	17
Non-hydrogen atoms	38276	38677	39427
Protein residues	4839	4882	4845
Nucleic acid	47	—	—
Ligand (Zn <sup>2+</sup> )	6	6	8
<b>Refinement</b>			
PDB ID	5j8	5j9	5ja
Resolution (Å)	290.2-3.8	290.2-4.0	290.2-4.7
Map sharpening B-factor (Å <sup>2</sup> )	-100	-136	-140
Average B-factor (Å <sup>2</sup> )			
<i>Protein</i>	64.9	161.8	182.0
<i>Nucleic acid</i>	70.9	—	—
<i>Ligand (Zn<sup>2+</sup>)</i>	58.6	62.4	107.9
Molprobability score	2.6	2.58	2.48
Clashscore (all atoms)	14.5	13.41	13.84
Rotamer outliers (%)	2.00	2.08	1.6
<b>Ramachandran statistics</b>			
Favored (%)	82.46	82.52	82.58
Disallowed (%)	1.15	1.22	1.03
RMS (bonds, Å)	0.0032	0.0032	0.0030
RMS (angles, °)	1.03	1.01	0.89
<b>Nucleic Acid</b>			
Connect sugar pucker (%)	81.5	—	—
Good backbone conform. (%)	88.7	—	—

Be specific if reality is complex

Extended Data Table 2 | Model statistics for elongating Pol III

Subunit	Protein	Chain ID	Mw (kDa)	No. of residues	Residues built	Chain breaks	All-atom clashscore	Molprobability score	Average B-factor (Å <sup>2</sup> )	Local resolution (Å)
<b>Core</b>										
<i>RPC1</i>	C100	A	182.9	1480	1422 (97.4%)	3	10.42	2.24	50.35	3.8
<i>RPC2</i>	C126	B	129.5	1149	1115 (97.0%)	—	12.61	2.40	60.10	3.8
<i>RPC4d</i>	A140	C	37.7	335	335 (100%)	—	10.60	2.40	60.42	3.8
<i>BBP5</i>	ABC27	E	25.1	215	215 (100%)	—	16.07	2.81	67.37	4.1
<i>BBP6</i>	ABC29	F	17.9	155	83 (53.5%)	—	3.87	2.34	57.87	3.6
<i>BBP8</i>	ABC14.5	H	16.5	145	140 (96.5%)	1	3.95	2.15	62.34	4.2
<i>RPC11</i>	C11	I	12.5	110	43 (38.1%)	—	16.00	2.73	73.40	4.3
<i>BBP10</i>	ABC10B	J	8.3	70	63 (97%)	—	5.40	2.59	56.79	3.6
<i>RPC19</i>	A1019	K	16.1	142	101 (71.1%)	—	8.85	2.41	59.88	3.7
<i>RPC10</i>	ABC10a	L	7.2	70	40 (57.1%)	—	12.13	2.65	67.80	4.2
<b>Stalk</b>										
<i>RPC9</i>	C17	D	18.6	161	119 (73.9%)	2	15.82	2.51	73.39	4.4
<i>RPC8</i>	C25	G	34.3	212	191 (90.1%)	2	16.88	2.87	70.33	4.3
<b>Heterodimer</b>										
<i>RPC3</i>	C82	O	74.0	654	539 (82.4%)	2	15.99	2.66	73.38	4.5
<i>RPC6</i>	C34	P	36.1	317	80 (26.1%)	2	26.41	3.54	75.34	5.3
<i>RPC7</i>	C31	Q	27.7	251	63 (25.1%)	2	10.13	2.13	76.31	4.6
<b>Heterodimer</b>										
<i>RPC5</i>	C37	M	82.1	280	164 (58.2%)	1	11.72	2.59	74.18	4.3
<i>RPC4</i>	C53	N	46.7	422	110 (26.1%)	1	12.10	2.68	71.51	4.2
<b>RNA</b>	RNA	R		18	9 (50.0%)	—			75.30	3.9
<b>DNA</b>	Non-Templatic	S		38	13 (38.5%)	—			75.31	3.9
<b>DNA</b>	Templatic	T		38	23 (60.5%)	—			78.43	3.9

## Help your reviewers...



**wwPDB Validation Service**

[FAQ](#)

**Existing validator:**

Username:

Password:

[Log In](#)

[Forgot Password](#)

[Deposition server](#)

Deposit your data to PDB, BMRB and EMDB at [deposit.wwpdb.org](#)

**wwPDB news and announcements**

**Compliance with GDPR legislation**

wwPDB has revised its privacy policy in line with the requirements of the EU's GDPR legislation.

[Start a new validator](#)

Welcome to the wwPDB validation system.

This server runs the perform the same validation as you would observe during the deposition process. This service is designed to help you check your model and experimental files prior to their deposition.

To continue with an existing validator, please login to the site.

To start a new validator, please complete the form below. Upon completion, you will be emailed the information specific to your new validator.

Your email address:

Password (optional):  (it will provide one) if it is blank:  (to be emailed to you)

Country:

Experimental method:

- X-Ray Diffraction
- Electron Microscopy
- Cryo-EM
- Neutron Diffraction
- Cryo-electron tomography
- Microbeam XRD

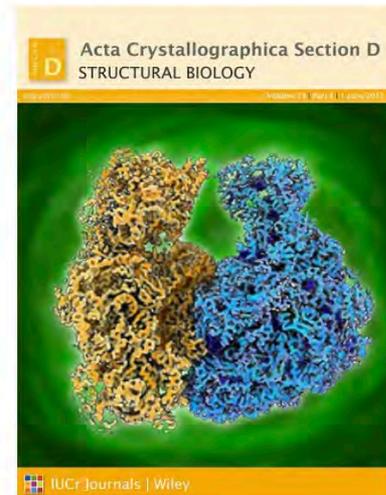
# Useful resources

<http://www.ccpem.ac.uk/courses/>

The screenshot shows the CCP-EM website for the Icknield Workshop. The header includes the CCP-EM logo, the text "Collaborative Computational Project for Electron cryo-Microscopy", and the MRC logo. A navigation bar contains links for Home, About CCP-EM, CCP-EM Projects, Downloads, Resources / Documentation, Workshops / Courses, and Symposium. The main heading is "Icknield Workshop: Model Building and Refinement using cryo-EM Maps". Below this, it states "Course kindly hosted at RAL by DLS | April 2019". A "Schedule" section lists the following:

- Schedule**
- Lectures**
- CCP-EM: Tom Burnley
- Introduction
- LoScale: Arjen Jakobi
- Map sharpening tools for cryo-EM density maps
- Buccaneer: Kevin Cowtan
- Model building in Buccaneer and Nautilus
- MeIProbit: Jane Richardson
- Validating CryoEM models from diagnosis to healing
- ARP/wARP: Grzegorz Chojnowski
- Model building into high resolution cryo-EM maps
- Coot: Paul Emsley

## CCP-EM symposium proceedings *Acta D*



<http://cryoem.tudelft.nl/software>

<http://gitlab.tudelft.nl/aj-lab>

The screenshot shows the rsref software documentation page. It includes the following text:

**rsref** is a scripted modular workflow for the refinement of atomic models against high-resolution cryo-EM density maps. It is not a standalone refinement software, but rather provides a set of tools for model and/or map manipulation, refinement protocols, the analysis of the refinement cycles and validation of the resulting coordinate models.

**rsref** makes use of the libraries from the **octbx** project.

**Availability:**

**rsref** is an ongoing development and not officially released. If you are fine with limited support, please see the **Wiki** pages for download and usage instructions and tutorial.

If **rsref** is useful for your work please acknowledge the **octbx** project (doi).

The diagram illustrates the rsref workflow. It starts with "Map" and "Model building" (labeled "Model building"). The "Map" path leads to "FSC" and "Residual". The "Model building" path leads to "Residual". Both paths converge at "Refinement", which leads to "B factors" and "Cross-validation". "Refinement" also leads to "Geometry" and "R-factor". "Cross-validation" leads to "FSC<sub>cut</sub>" and "FSC<sub>max</sub>". "B factors" leads to "FSC<sub>cut</sub>". "Geometry" leads to "R-factor". "R-factor" leads to "Validation". "Validation" leads to "FSC<sub>max</sub>". The entire process is labeled "rsref" at the bottom.

# Summary

- Atomic model building and refinement are now an important part of the cryo-EM structure determination process
- Resolution variation in cryo-EM maps still poses many challenges for model building (and refinement)
- Optimal map sharpening and/or filtering is not trivial
- Local filtering or sharpening (and LAFTER denoising) can serve to visually improve poorly resolved map regions for model building (and refinement)
- FDR thresholding may provide more objective way for map thresholding
- Many challenges remain, but tools are becoming better
- Validation is still an open field, but important initiatives have started

## Availability of tools shown:

<https://git.tudelft.nl/jakobi/>

<http://cryoem.tudelft.nl/software>



# Acknowledgements

## Müller Gang @ EMBL



Carsten  
Sachse



Maximilian  
Beckers



## MRC-LMB

Garib Murshudov



## CCP-EM

Tom Burnley  
Colin Palmer  
Agnel Joseph



Participants of the CCP-EM  
Icknield & Madrid workshops

<https://pypi.python.org/pypi/mrcfile>

<https://pypi.python.org/pypi/clipper-python/>

<http://cctbx.sourceforge.net/>



EMBL

