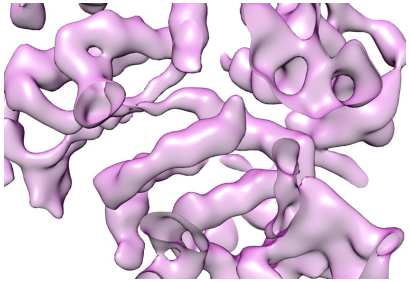
Phenix Map Tools

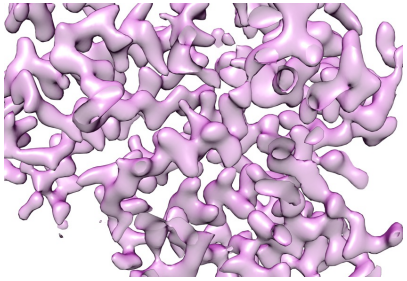
Pavel Afonine

October 28th, 2025
UTMB, Galveston, TX

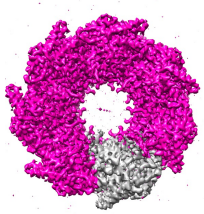
Phenix tools for cryo-EM



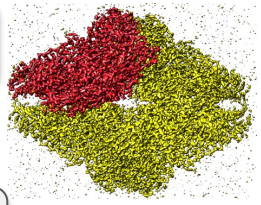
Starting map



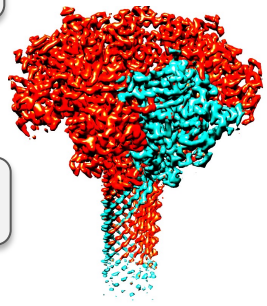
Map improvement



Map symmetry

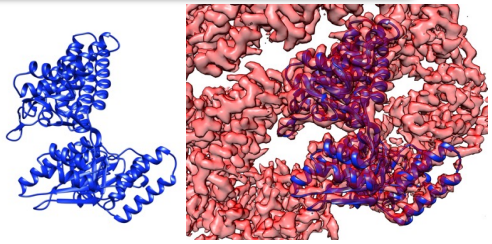


Map manipulation

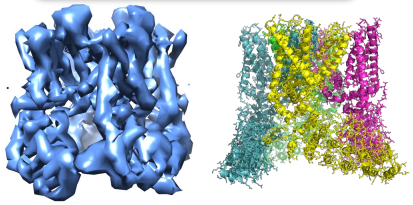


Extract unique part

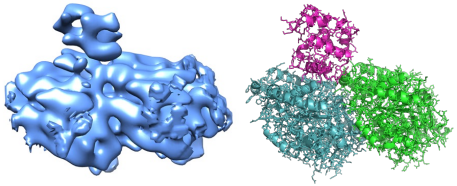
Docking, model building



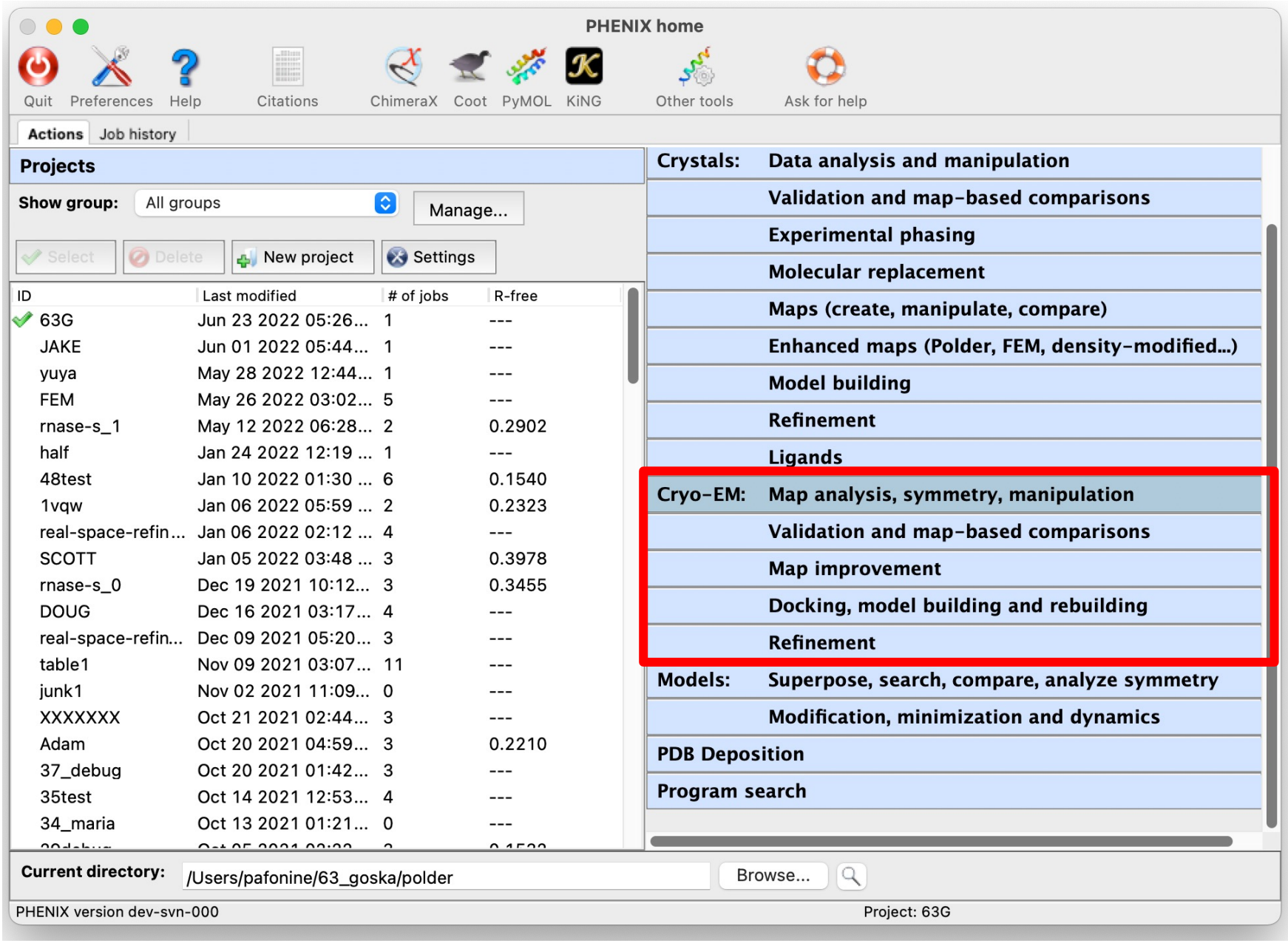
Refinement



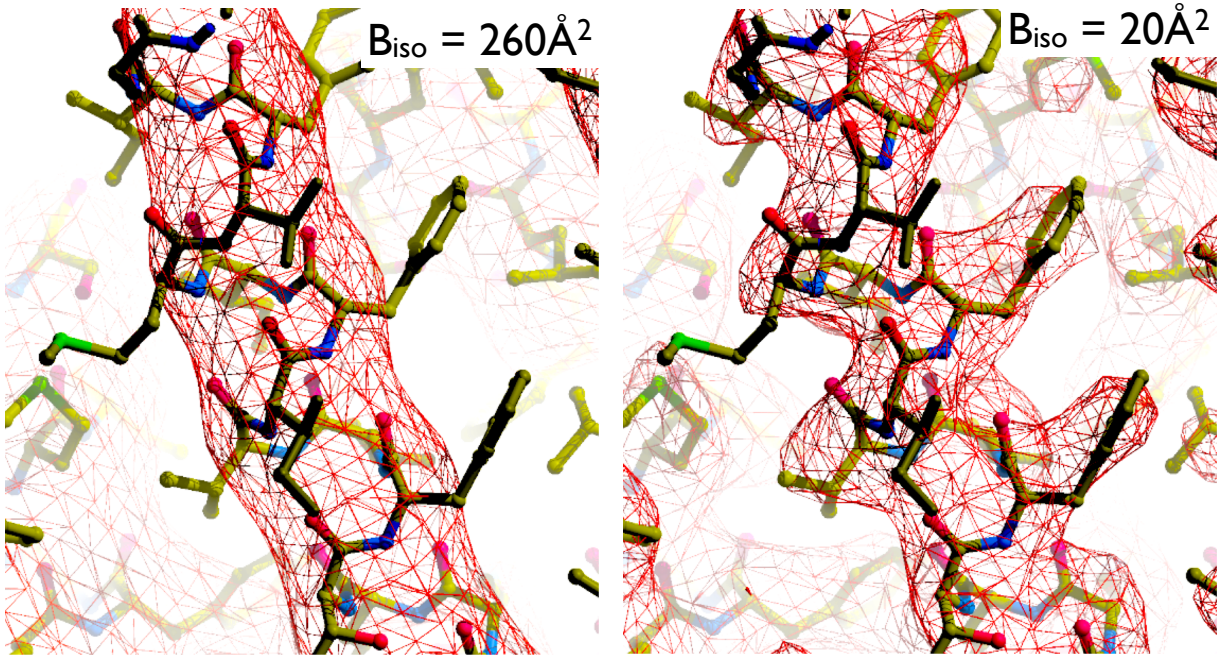
Validation



Phenix tools for cryo-EM: *GUI and command line*



Automated map sharpening: *phenix.auto_sharpen*



Deposited Map *Autosharpened Map*

EMDB: 8414, PDB: 5tji

Fully automatic:

No manual trial-and-error | No parameters to adjust | Only inputs: map and resolution

Acta Cryst D STRUCTURAL BIOLOGY
ISSN 2059-7983

Automated map sharpening by maximization of detail and connectivity

Thomas C. Terwilliger,[a,b]* Oleg V. Sobolev,[c] Pavel V. Afonine[c,d] and Paul D. Adams[d,e]

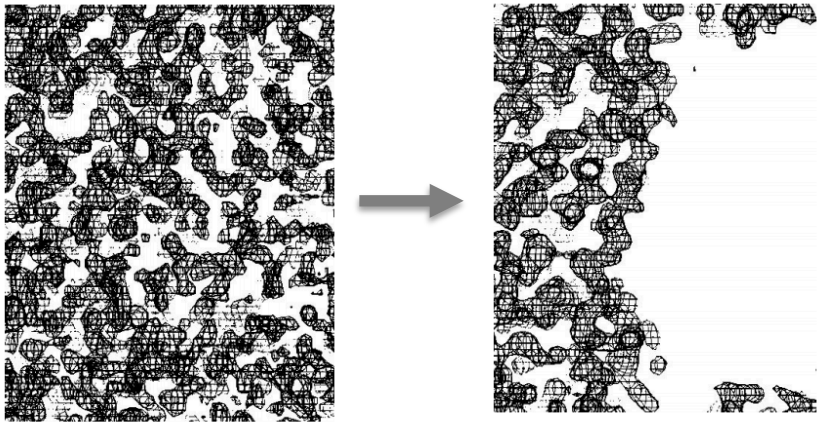
Density modification: *phenix.density_modify_cryo_em*

Similar principals for crystallography and cryo-EM:

change the map so that it is most consistent with what we know about macromolecules

Crystallography

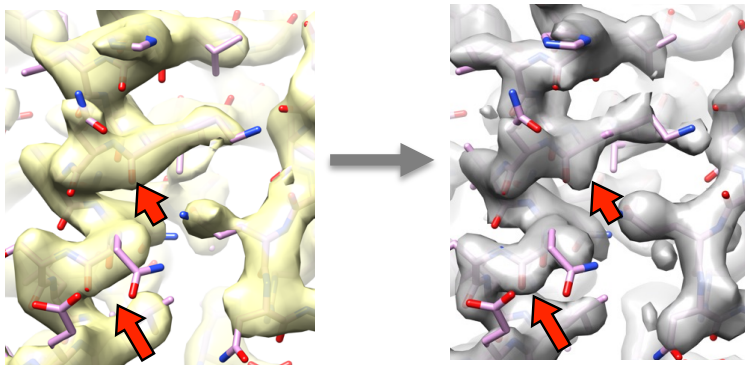
From uninterpretable to interpretable map



Cryo-EM

Effect is less dramatic as in crystallography

- Can increase map resolution (0.05-0.3 Å)
- Can improve map clarity for interpretation



nature methods

ARTICLES

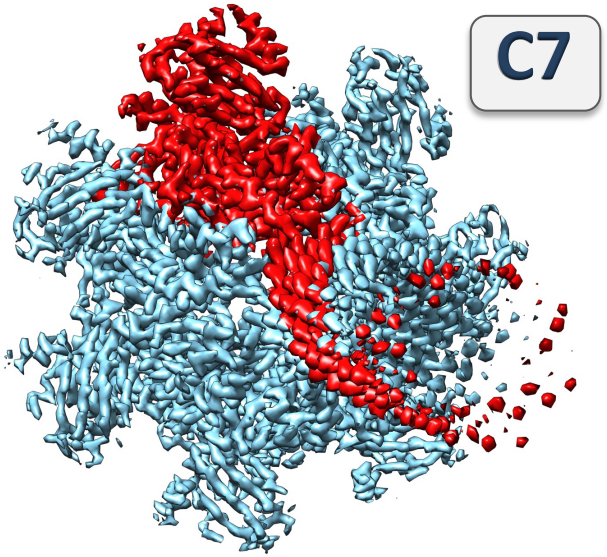
https://doi.org/10.1038/s41592-020-0914-9

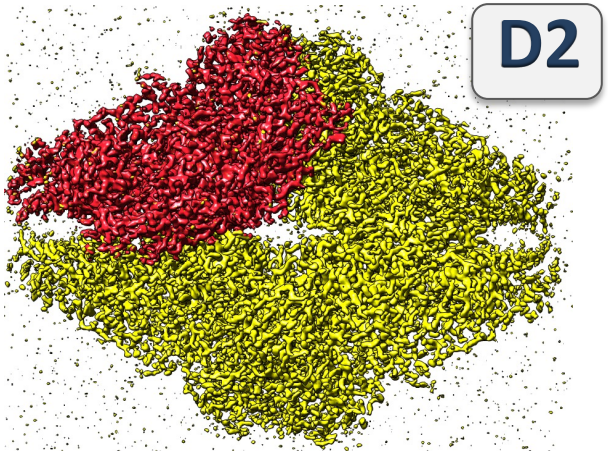
Check for updates

Improvement of cryo-EM maps by density modification

Thomas C. Terwilliger[1,2]✉, Steven J. Ludtke[3], Randy J. Read[4], Paul D. Adams[5,6] and Pavel V. Afonine[5]

Finding map symmetry: *phenix.symmetry_from_map*

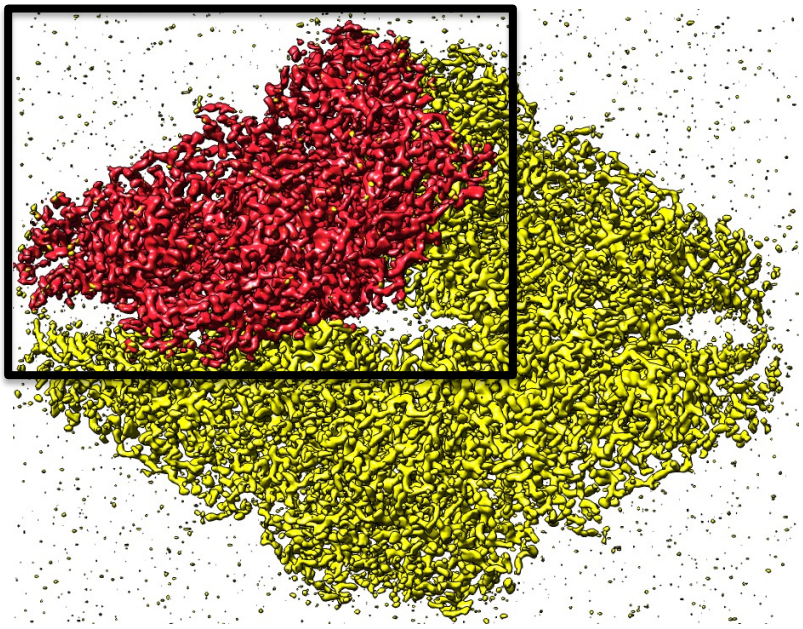




http://phenix-online.org/newsletter/

Tools for interpreting cryo-EM maps using models from the PDB

Extracting unique part of map using *phenix.map_box*

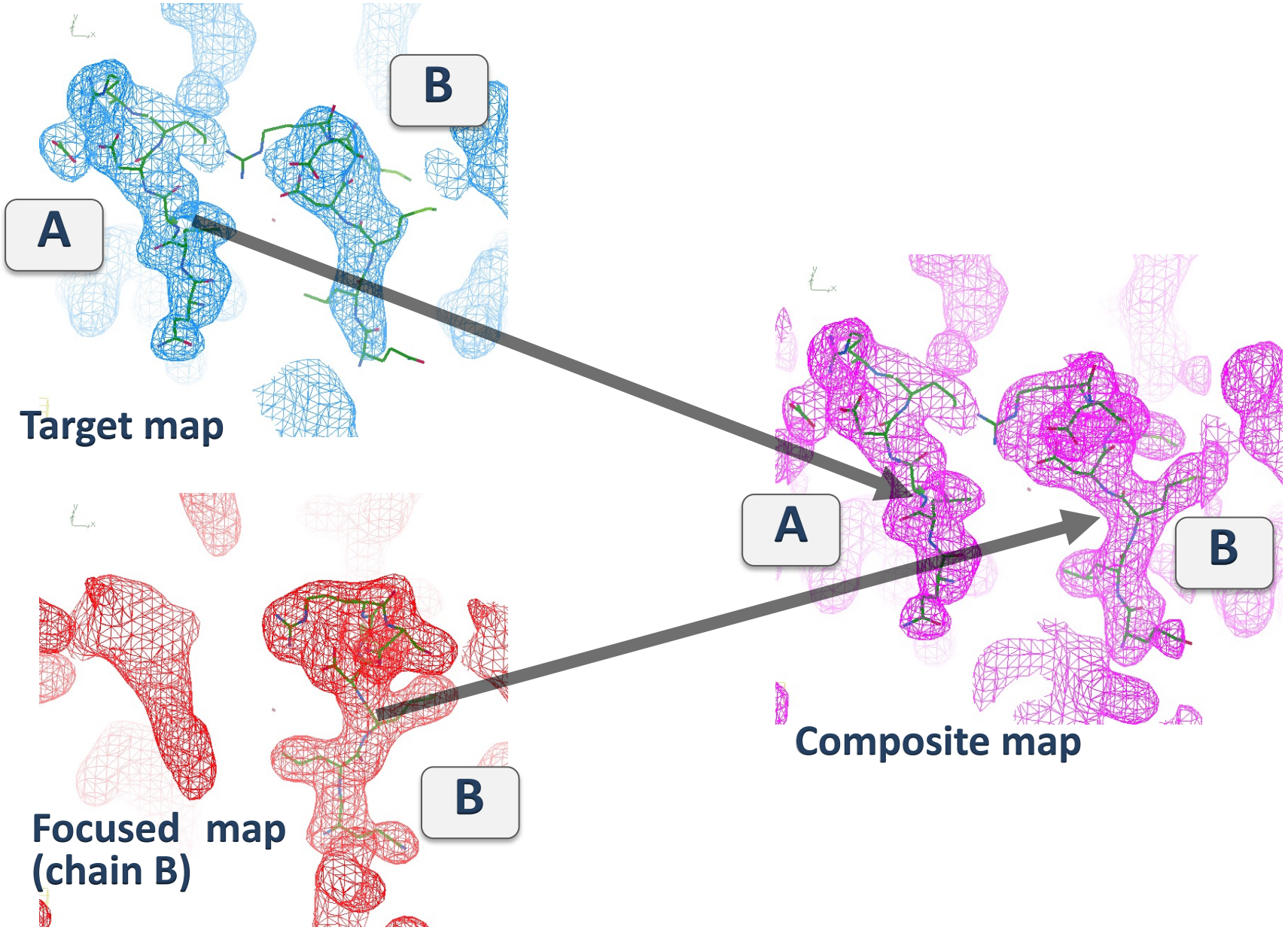


Lots of options: use map only, use model, use symmetry, mask boxed map, and many more!

http://phenix-online.org/newsletter/

Tools for interpreting cryo-EM maps using models from the PDB

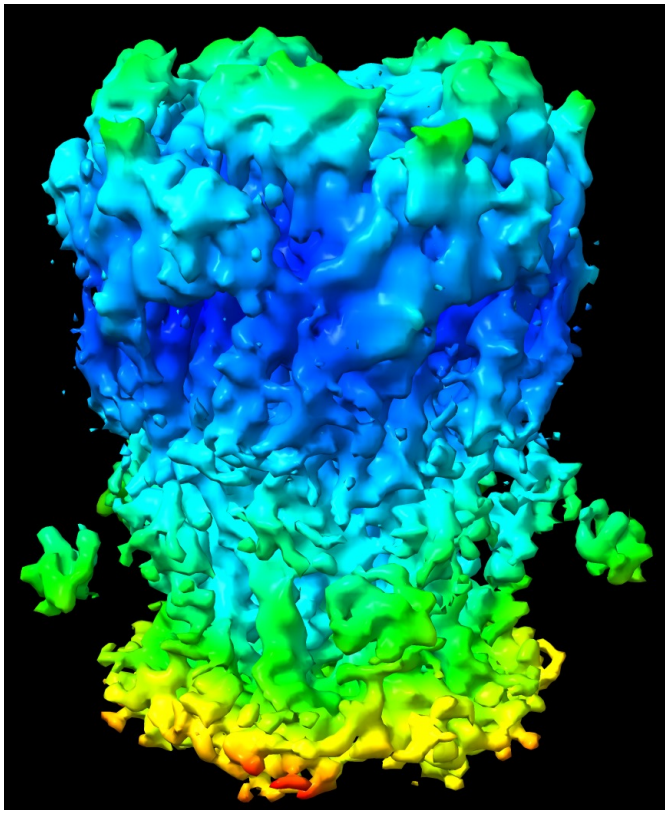
Combining maps with *phenix.combine_focused_maps*

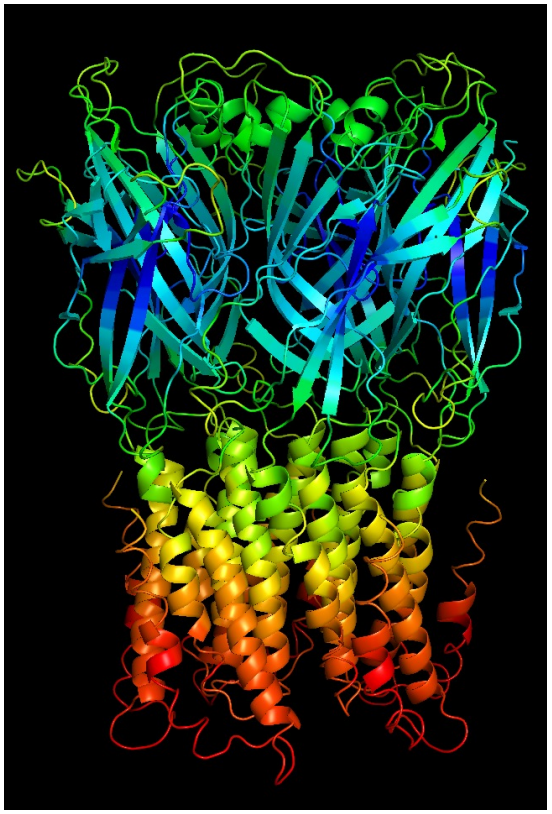


Local resolution: *phenix.local_resolution*

EMDB: 20986

PDB: 6v0b





```
phenix.local_resolution half_map_1.mrc half_map_2.mrc
```

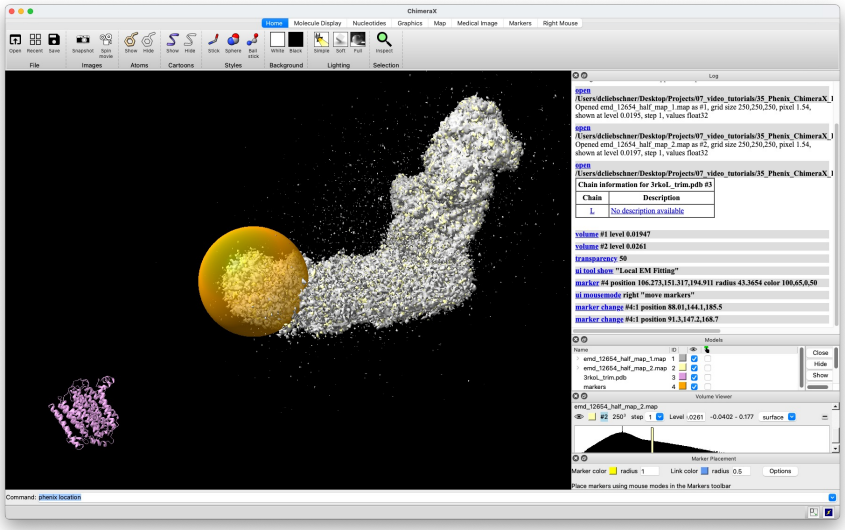
In ChimeraX load both maps and use

```
color sample #1 map #2 palette rainbow
```

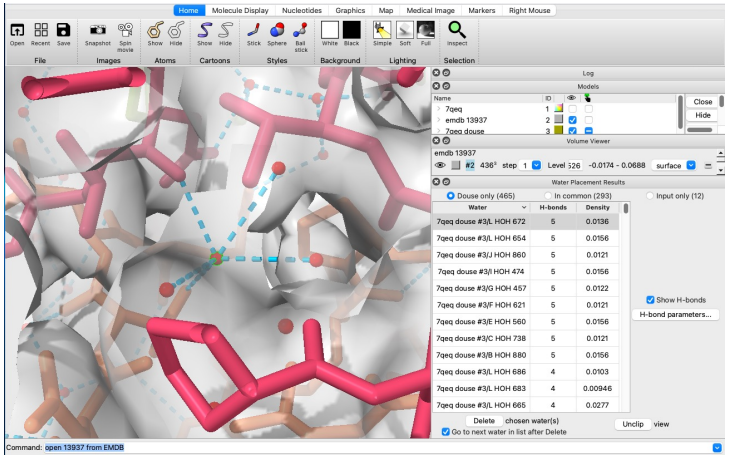
to color map #1 based on the resolution values in map #2

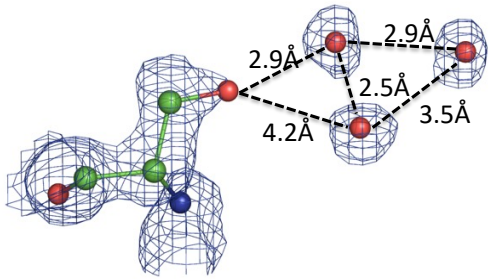
Phenix + ChimeraX

Local EM map fitting

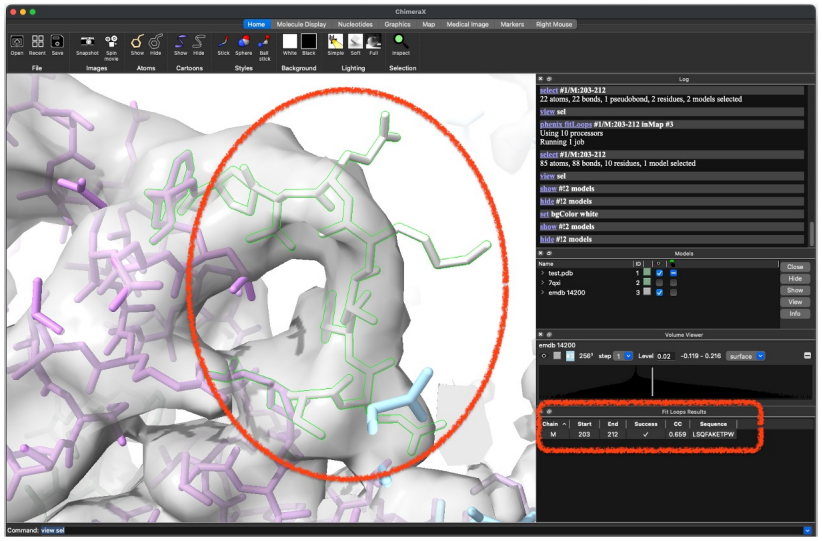


Water building into EM maps





Loop Fitting

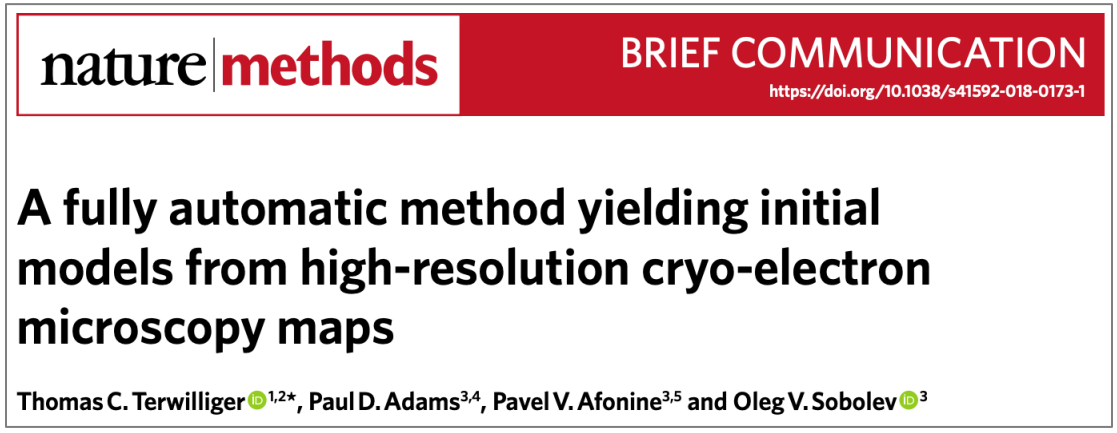


More is coming !

Automated model building: *phenix.map_to_model*

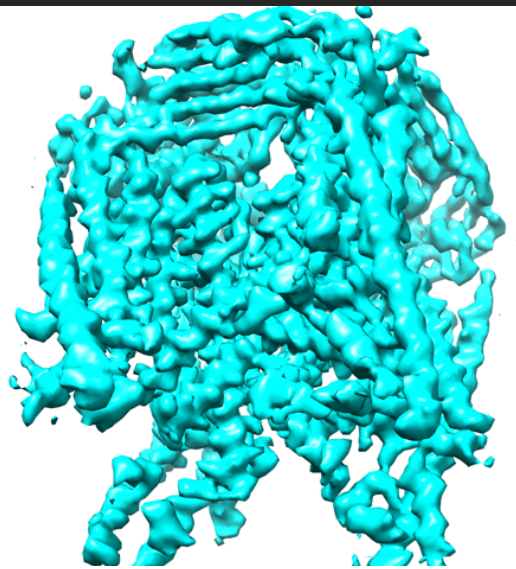
Automated model building, facts:

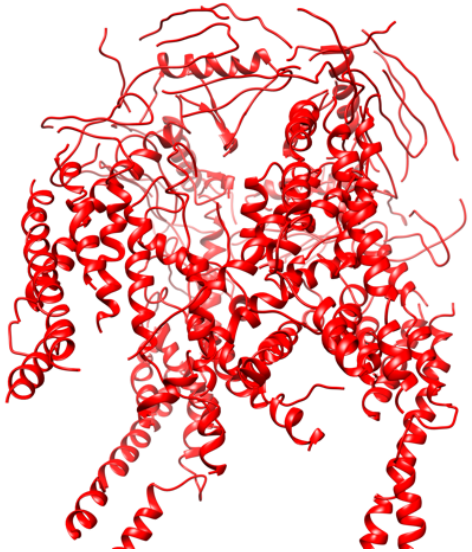
- No automated model building produces 100% complete and accurate model
- Produces initial model for further manual building
- The lower the resolution, the less complete and accurate the auto-built model

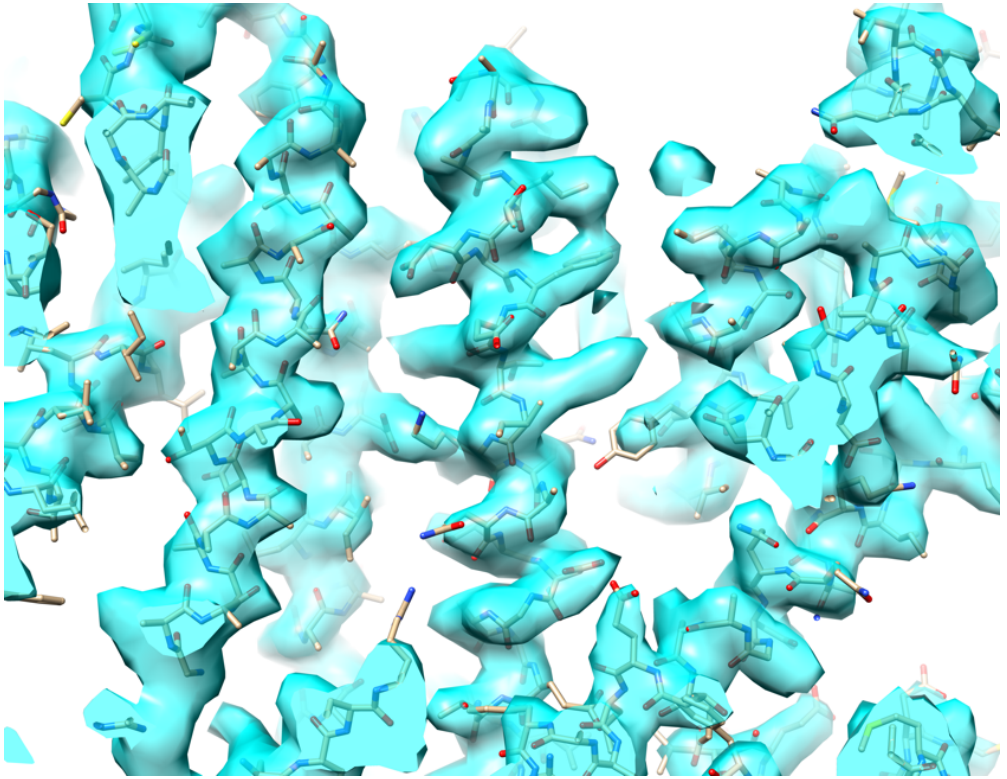


Automated model building: *phenix.map_to_model*

TRPML3 channel (4.1 Å, 78% built, 1.3 Å rmsd)



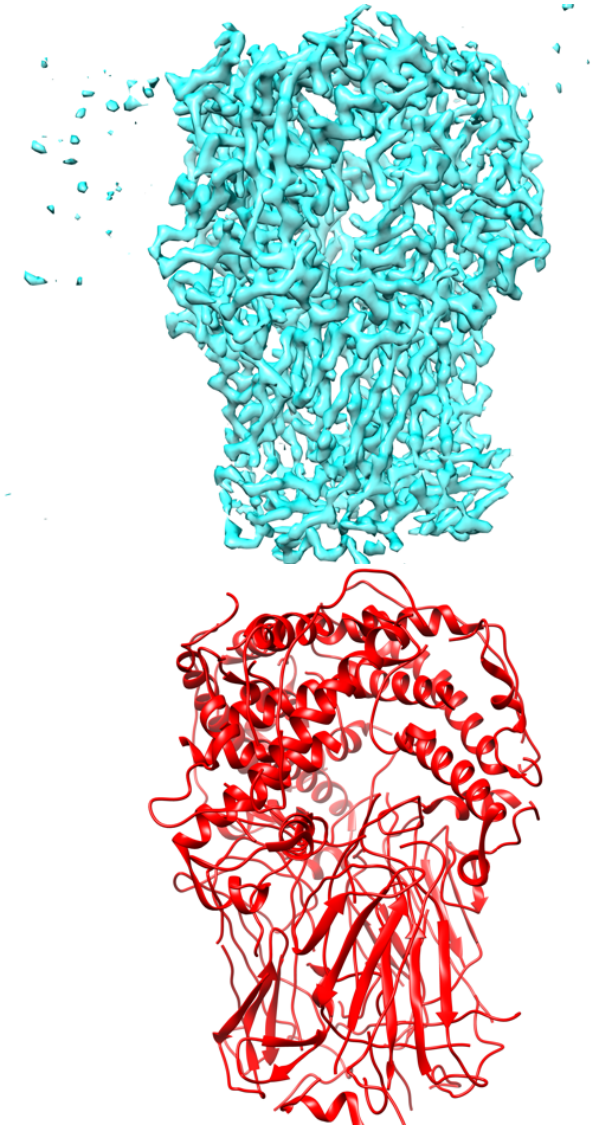


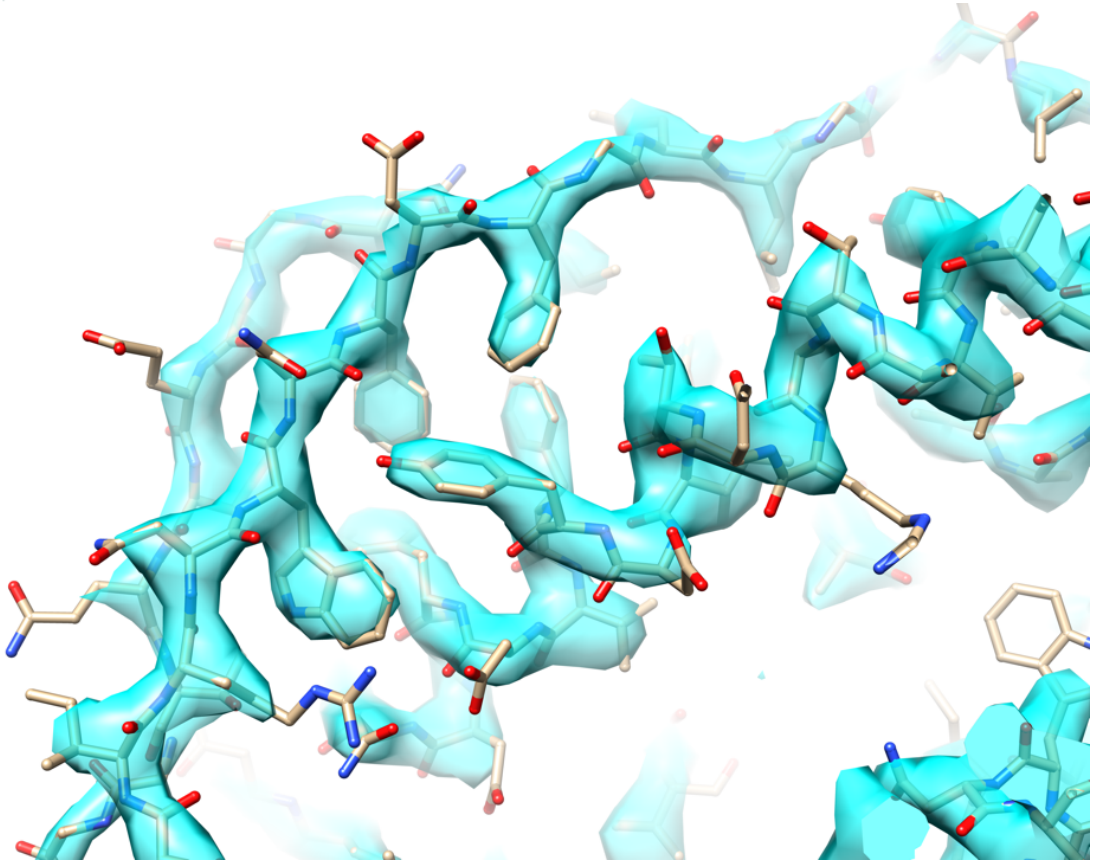


Data from Zhou, X. et al. (2017) Nat. Struct. Mol. Biol. 24: 1146

Automated model building: *phenix.map_to_model*

Rotavirus VP6 (2.6 Å, 100% built, 0.9 Å rmsd)





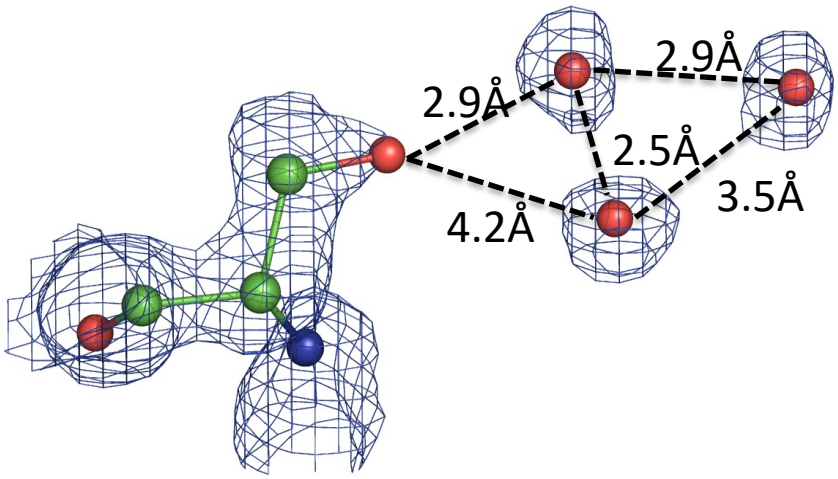
Data from Grant and Grigorieff, eLife 2015;4:e06980

Manual model building steps

If *phenix.map_to_model* fails or model is too big or else:

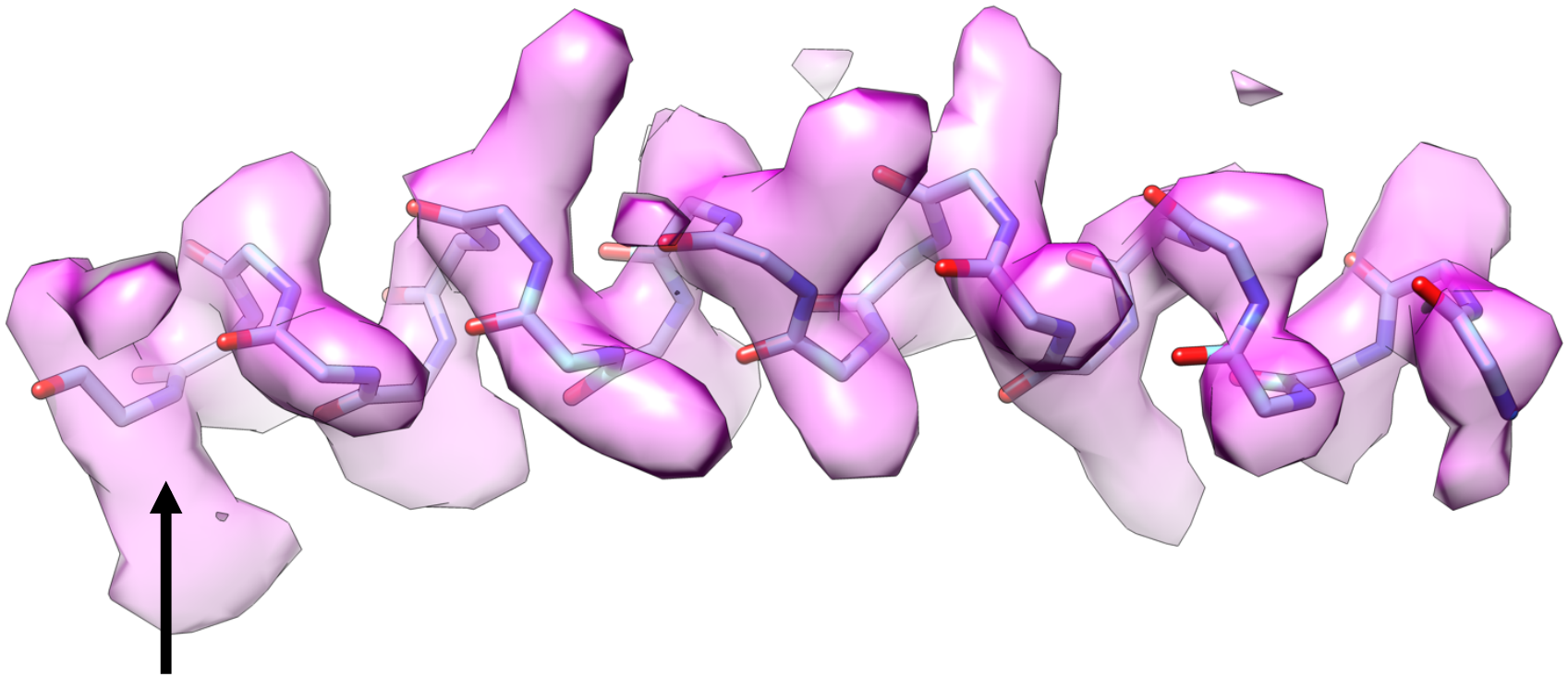
- Auto-sharpen the map
- Run Map Symmetry to obtain symmetry
- Run Map Box to obtain asymmetric unit (using symmetry)
- Run Map to Model on asymmetric unit
- Run Apply NCS Operators on model, with the trim overlap option (supplying the full map)

Automated water building: *phenix.douse*



Available in ChimeraX!

Sequence from map: *phenix.sequence_from_map*

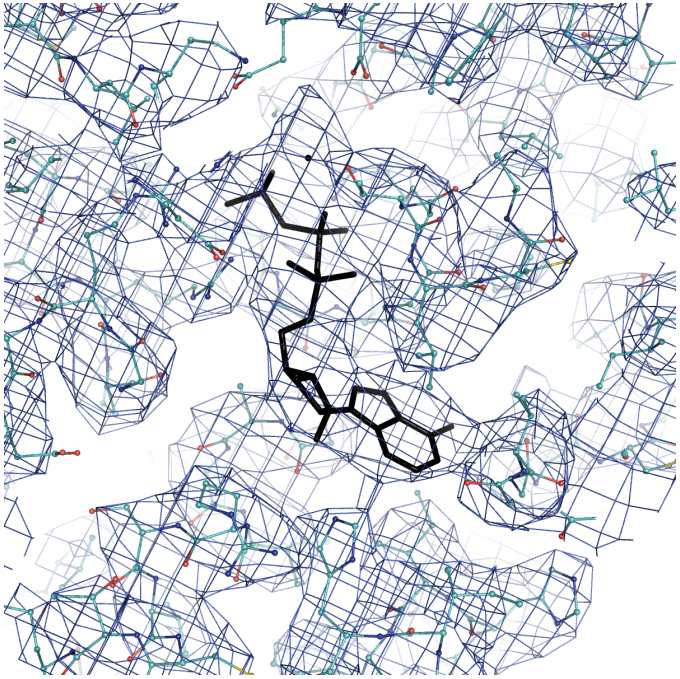


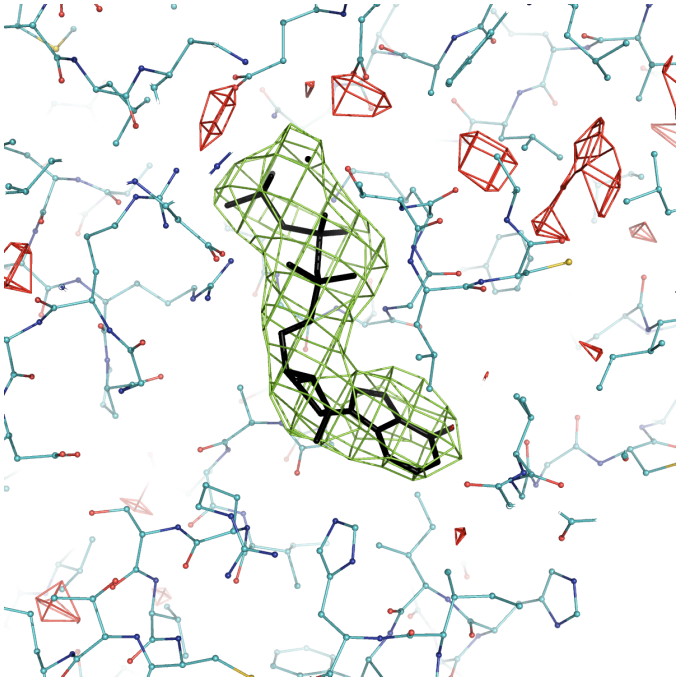
Residue	G	A	S	V	I	L	M	C	F	Y	K	R	W	H	E	D	Q	N	P	T
CC	0.30	0.50	0.53	0.47	0.58	0.62	0.68	0.59	0.83	0.77	0.71	0.69	0.70	0.82	0.65	0.64	0.60	0.60	0.35	0.47
Prob	3	0	0	0	0	0	1	0	40	23	5	5	4	9	2	2	1	0	2	0

- Determine probability of side chain at each C**α**
- Align sequence to maximize total probability for the chain

Difference maps: *phenix.real_space_diff_map*

5L4g, EMDB 4002

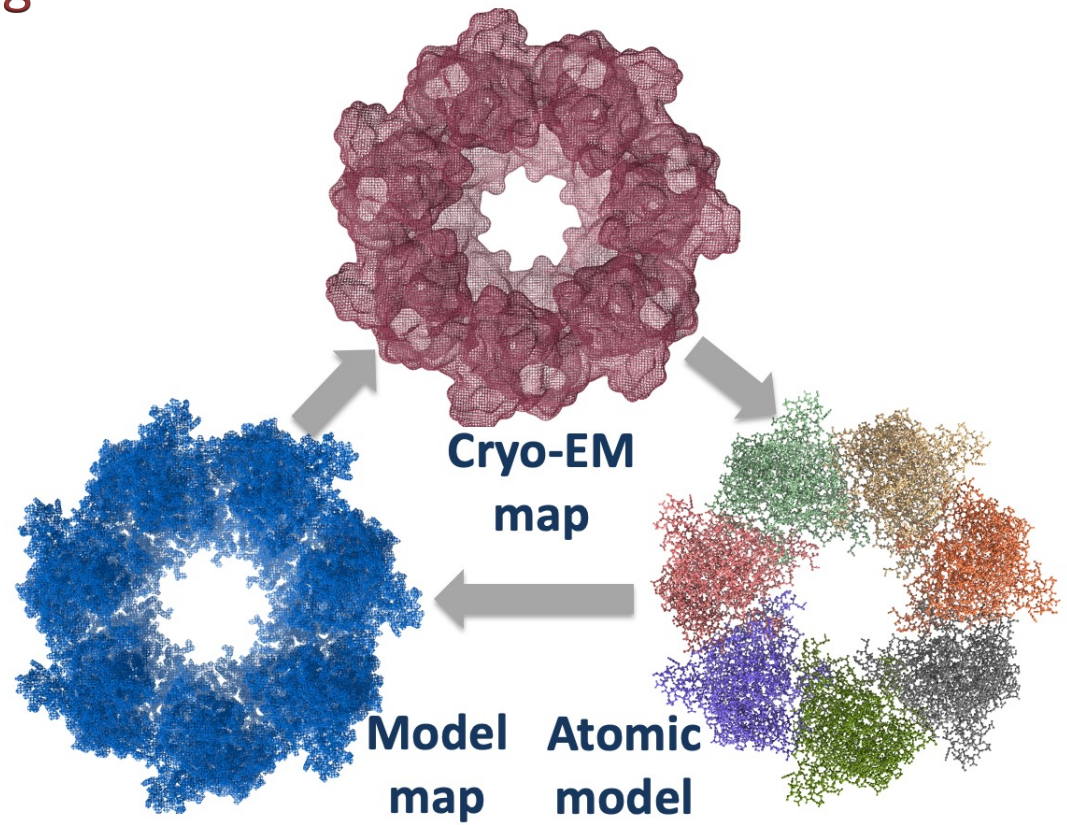




- Analogue of crystallographic Fo-Fc map
- Requires well-refined model (including B factors)

Model map

- Computing an adequate map from an atomic model is the key for:
  - Refinement
  - Model building
  - Validation
  - Ligands



Model map

- ChimeraX: (https://www.rbvi.ucsf.edu/chimerax/docs/user/commands/molmap.html)
  - Each atom is a 3D Gaussian distribution of width proportional to the resolution and amplitude proportional to the atomic number.
  - Tom Goddard (August 2025): *"I don't have a paper reference, I believe I copied what EMAN did 20 years ago"*

- CryoFit:

$$\rho\left(k, \mathbf{r}_j\right) = \exp\left[-\frac{1}{2}\left(\frac{\mathbf{r}_k - \mathbf{r}_j}{\sigma}\right)^2\right]$$

2σ = resolution

- DEMO-EMfit:

$$\rho_M(\nu_i) = \sum_{j=1}^{L} m \sqrt{\left(\frac{\pi}{(2.4 + 0.8R)^2}\right)^2} \exp\left(-\left(\frac{\pi}{2.4 + 0.8R}\right)^2 |\nu_i - x_j|^2\right)$$

m = atomic mass
R = resolution

- Struc2mapGAN, EMReady:

$$\rho(\mathbf{x}) = \sum_i^M \theta Z_i e^{-k|\mathbf{x}-\mathbf{r}_i|^2}$$

$k = (\pi / (1.2 + 0.6R))^2$
$\Theta = (k / \pi)^{1.5}$
R = resolution

- Q-score

$$y = Ae^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} + B$$

σ = 0.6
A,B ~ map mean and s.d.

Model map: two fundamental limitations

$$\rho\left(k, \mathbf{r}_j\right) = \exp\left[-\frac{1}{2}\left(\frac{\mathbf{r}_k - \mathbf{r}_j}{\sigma}\right)^2\right]$$

$$\rho_M(\nu_i) = \sum_{j=1}^{L} m_j \sqrt{\left(\frac{\pi}{(2.4 + 0.8R)^2}\right)^2} \exp\left(-\left(\frac{\pi}{2.4 + 0.8R}\right)^2 |\nu_i - x_j|^2\right)$$

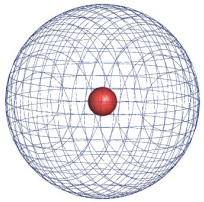
$$\rho(\mathbf{x}) = \sum_{i}^{M} \theta Z_i e^{-k|\mathbf{x}-\mathbf{r}_i|^2}$$

$$y = Ae^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} + B$$

- Ad hoc, intuitive formulations to mimic experimental density
  - No chemical element type
  - No charge
  - No occupancy
  - No B factors (isotropic, anisotropic)

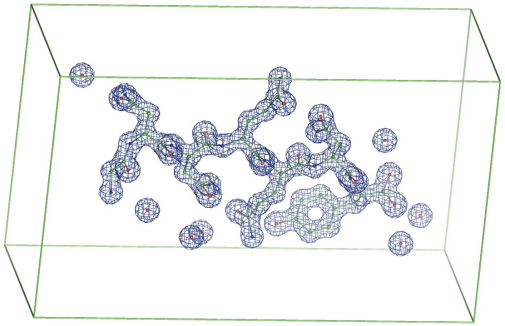
- Resolution is not accounted for

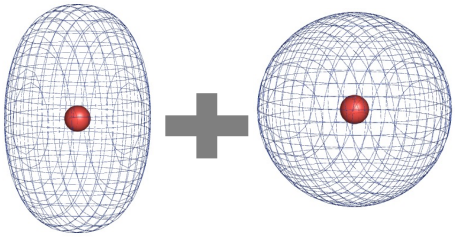
Model map: Independent Atom Model (IAM)



$$\rho_{atom}(\mathbf{r},\mathbf{r}_0,B,q) = q\sum_{k=1}^{5} a_k \left(\frac{4\pi}{b_k+B}\right)^{3/2} \exp\left(-\frac{4\pi^2|\mathbf{r}-\mathbf{r}_0|^2}{b_k+B}\right)$$

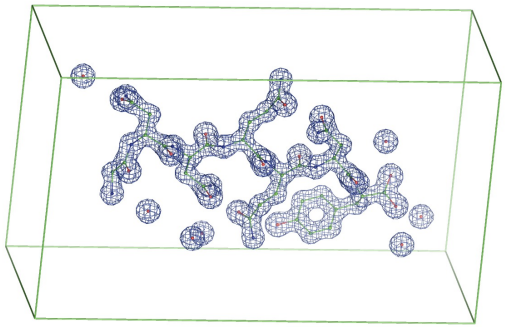
```
ATOM     25  CA  PRO A   4      31.309  29.489  26.044  1.00 57.79           C
```

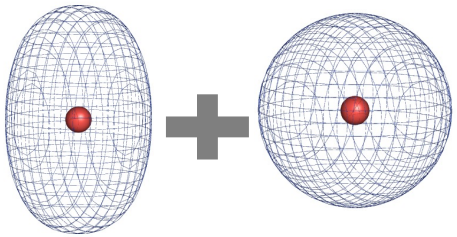




$$\rho_{MODEL}(\mathbf{r}) = \sum_{i=1}^{Natoms} \rho_{atoms}(\mathbf{r})$$

Model map





$$\rho_{MODEL}(\mathbf{r}) = \sum_{i=1}^{Natoms} \rho_{atoms}(\mathbf{r})$$

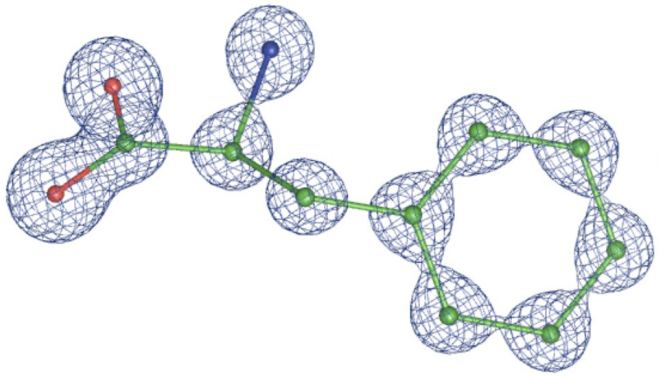
This is not the map we need for refinement, validation, model building..

In practice, we rarely need this map at all!

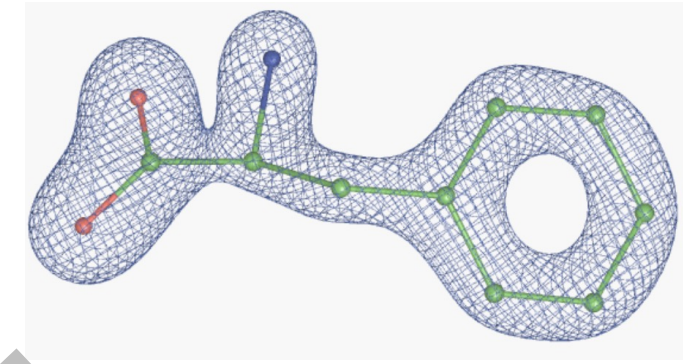
This is because it is an “infinite resolution” (exact) map, which cannot be meaningfully compared with the experimental map.

Model map suitable for calculations

Exact model map ρ_{calc}

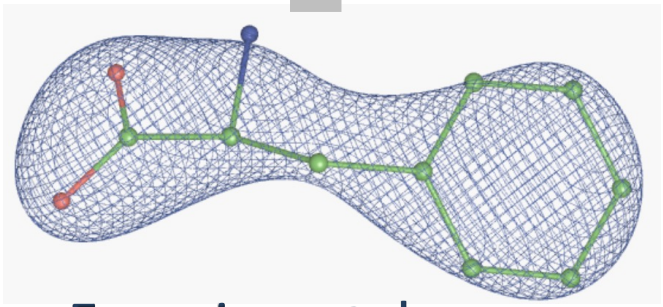


Finite resolution model map $\widetilde{\rho_{calc}}$



$$LS = \sum_{map} (\rho_{obs} - \widetilde{\rho_{calc}})^2$$

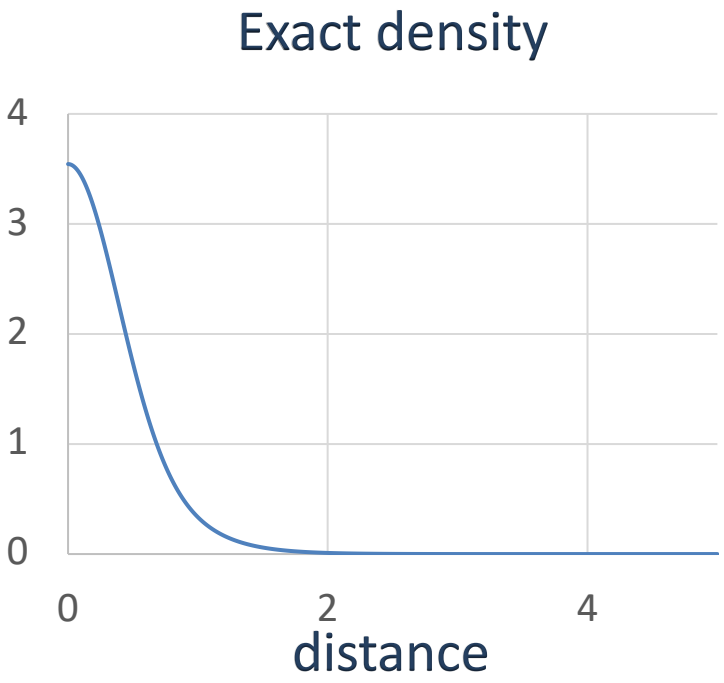
$$\mathrm{CC} = \frac{\sum \rho_{obs} \widetilde{\rho_{calc}}}{(\sum \rho_{obs}^2 \sum \widetilde{\rho_{calc}}^2)^{1/2}}$$

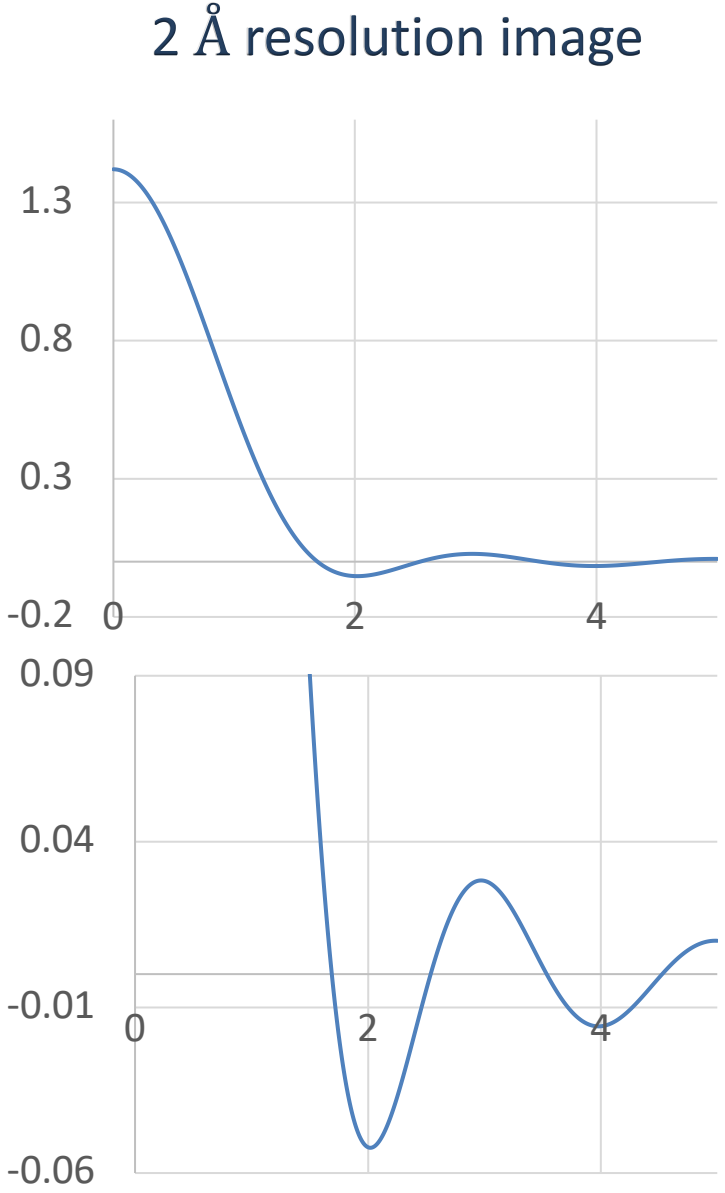


Experimental map

ρ_{obs} and $\widetilde{\rho_{calc}}$ must be the same resolution

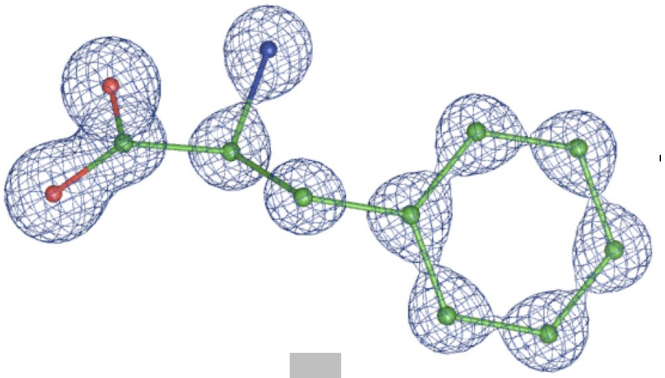
Model map illustrations for Carbon atom



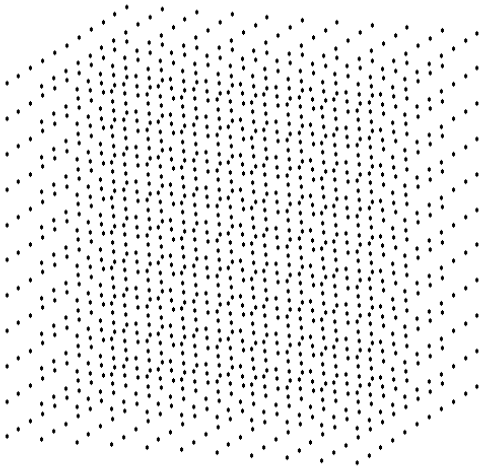


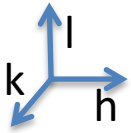
Finite resolution model map calculation (Fourier map)

Exact model map ρ_{calc}



FT $^{-1}$

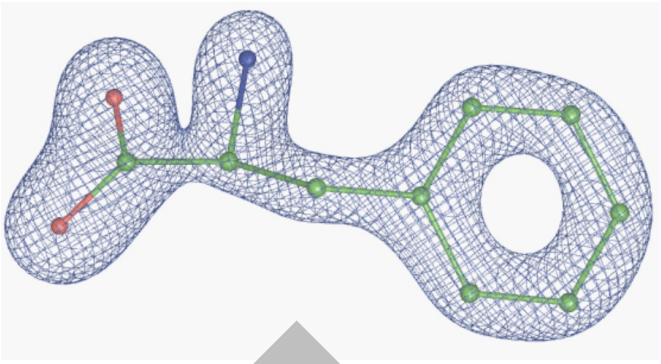




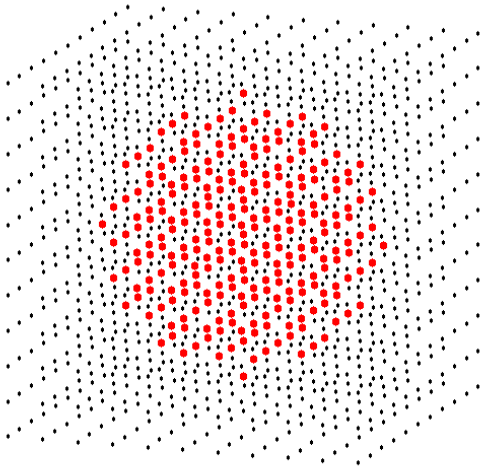
Fourier map coefficients

Truncate resolution

Model-calculated Fourier map $\widetilde{\rho_{calc}}$

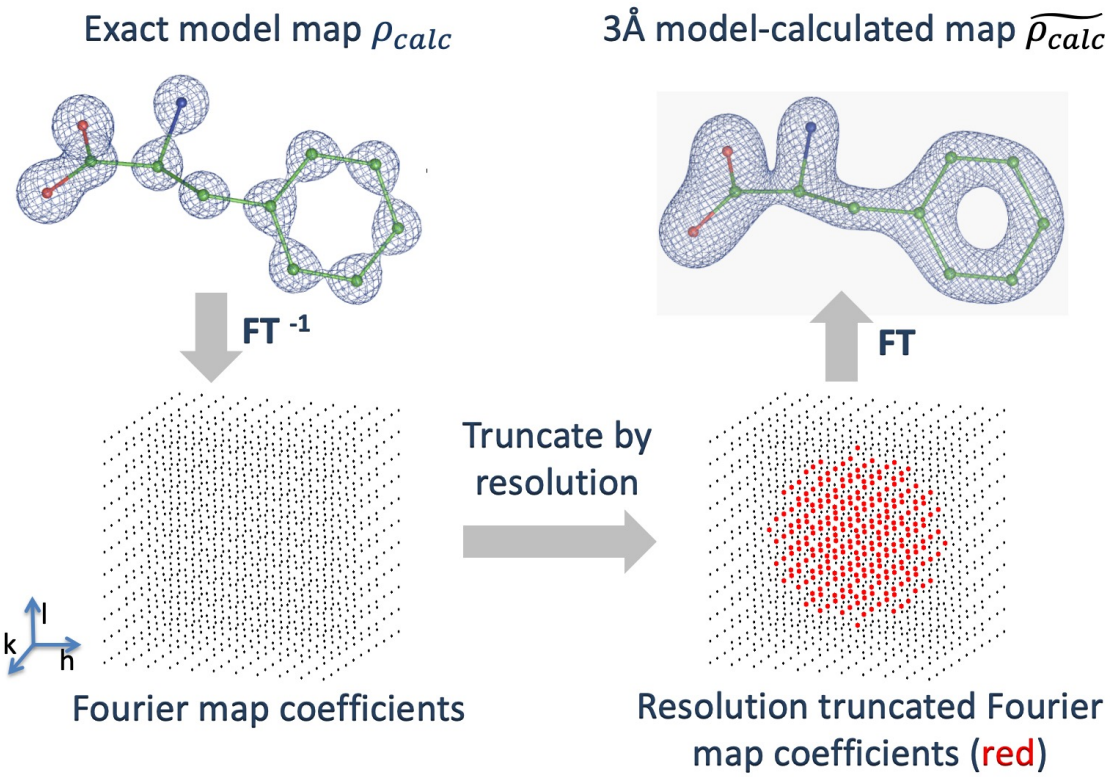


FT



Resolution truncated Fourier map coefficients (red)

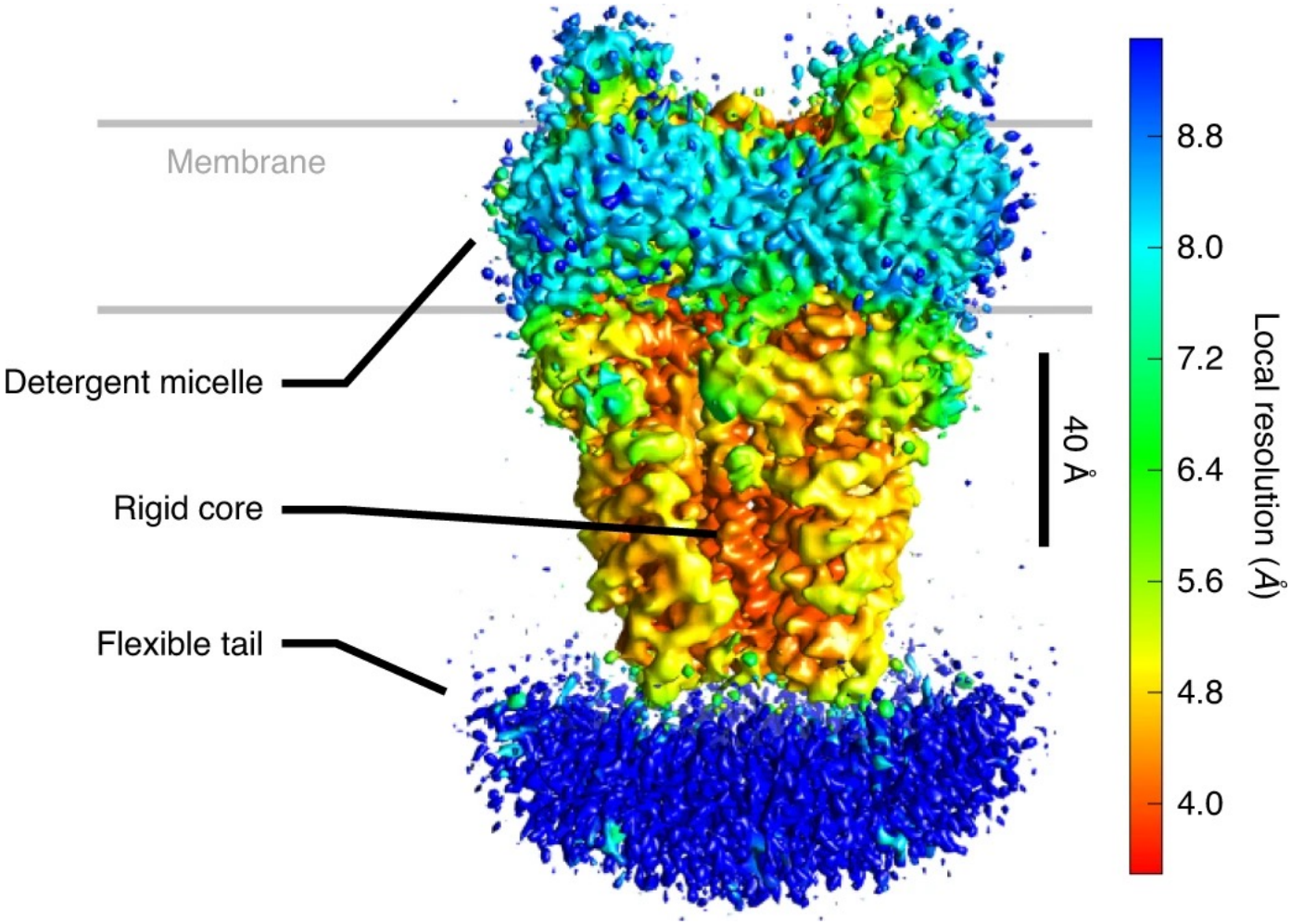
Finite resolution model map calculation (Fourier map)



Two problems with this approach:

1. This is a very indirect and heavy calculation
2. Resolution for all atoms is exactly the same

CryoEM: local resolution can vary a lot



Adopted from Zhang et al. *Nat Methods* **17**, 1214–1221 (2020).

Solution comes from this series of publications

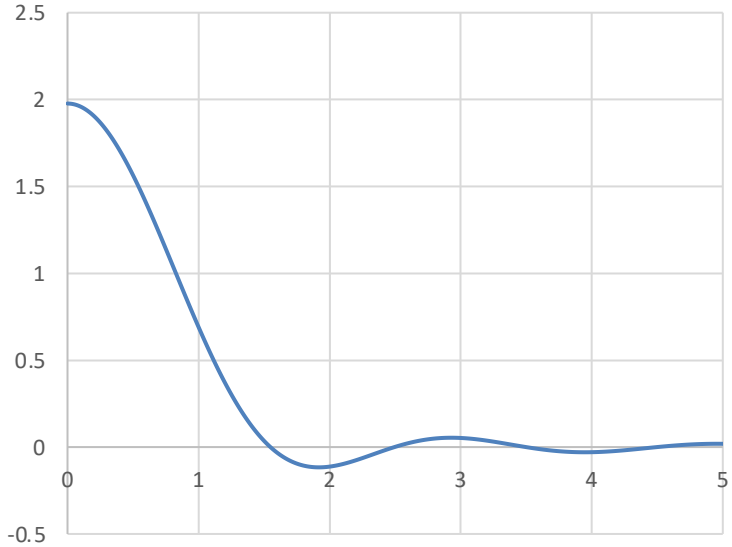
Urzhumtsev, A., Lunin, V.Y. (2022) "**Analytic modeling of inhomogeneous-resolution maps in cryo-electron microscopy and crystallography**". *IUCr Journal*, **9**, 728-734.

Urzhumtsev, A., Urzhumtseva, L.M., Lunin, V.Y. (2022) "**Direct calculation of cryo-EM and crystallographic model maps for real-space refinement**". *Acta Cryst.,* D**78**,1451-1468.

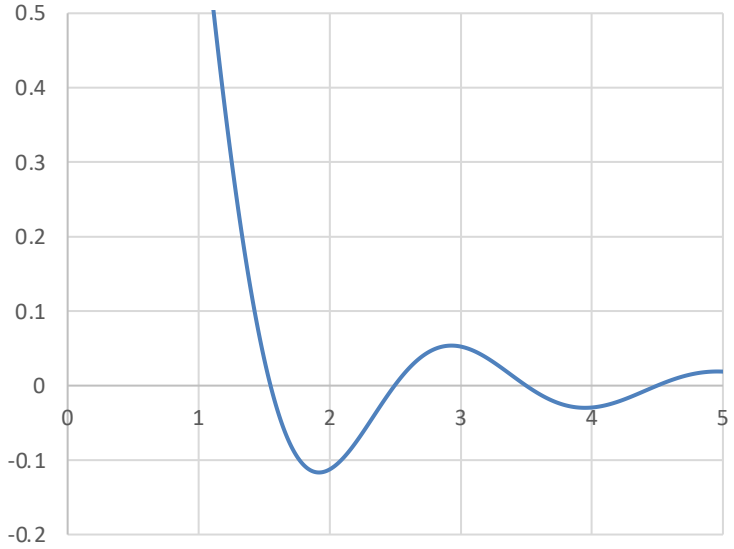
Urzhumtseva, L., Lunin, V.Y., Urzhumtsev, A. (2023) "**Algorithms and programs for the shell decomposition of oscillating functions in space**". *J. Appl.Cryst.,* **56**, 302-311.

Idea at a glance

3 Å resolution image of C atom (2D)



Zooming on ripples



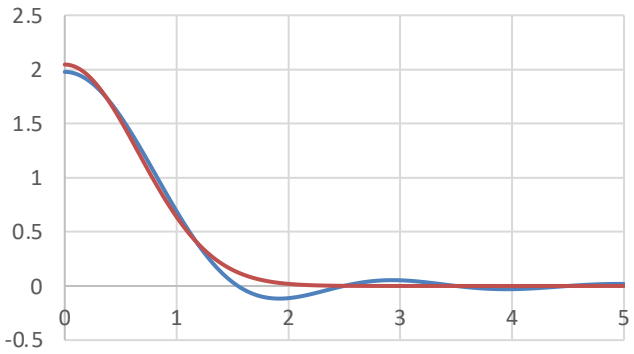
This image can be approximated with desired accuracy by this function

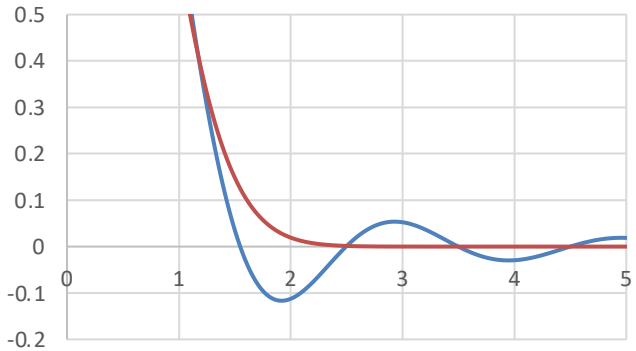
$$\sum_{m=1}^{M} C_m \Omega(\mathbf{r};\, R_m, B_m)$$

$$\Omega(\mathbf{r};\, R, B) = \frac{1}{|\mathbf{r}|R}\left(\frac{1}{4\pi B}\right)^{1/2}\left[\exp\left(-\frac{4\pi^2(|\mathbf{r}|-R)^2}{B}\right) - \exp\left(-\frac{4\pi^2(|\mathbf{r}|+R)^2}{B}\right)\right]$$

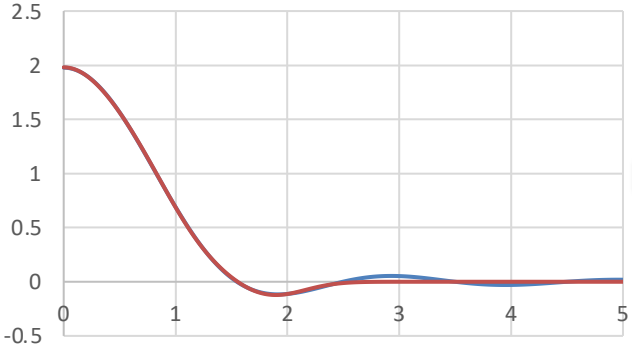
More terms M = more accurate approximation

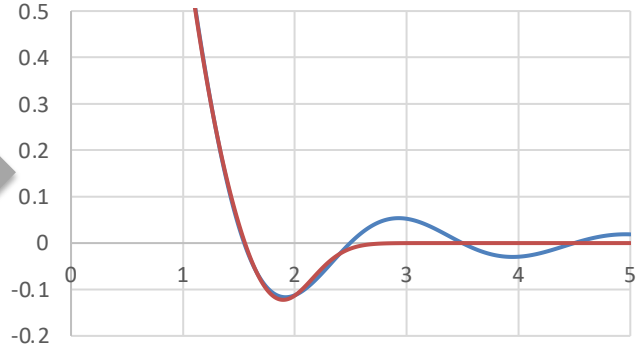
$$\sum_{m=1}^{M} C_m \Omega(\mathbf{r}; R_m, B_m)$$



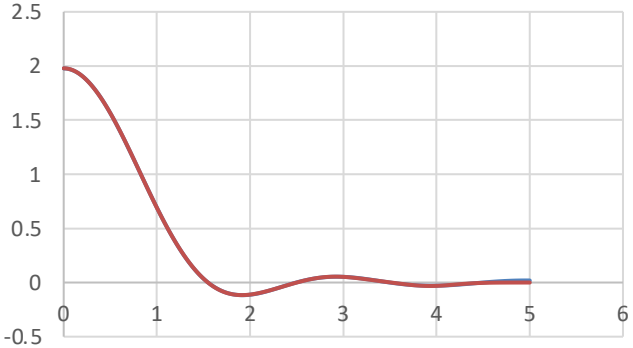


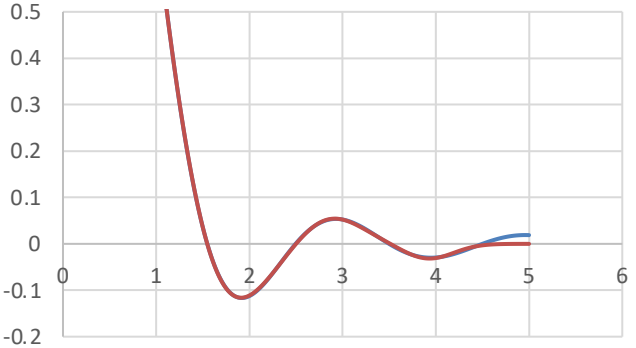
1 term





2 terms





4 terms

Practicalities

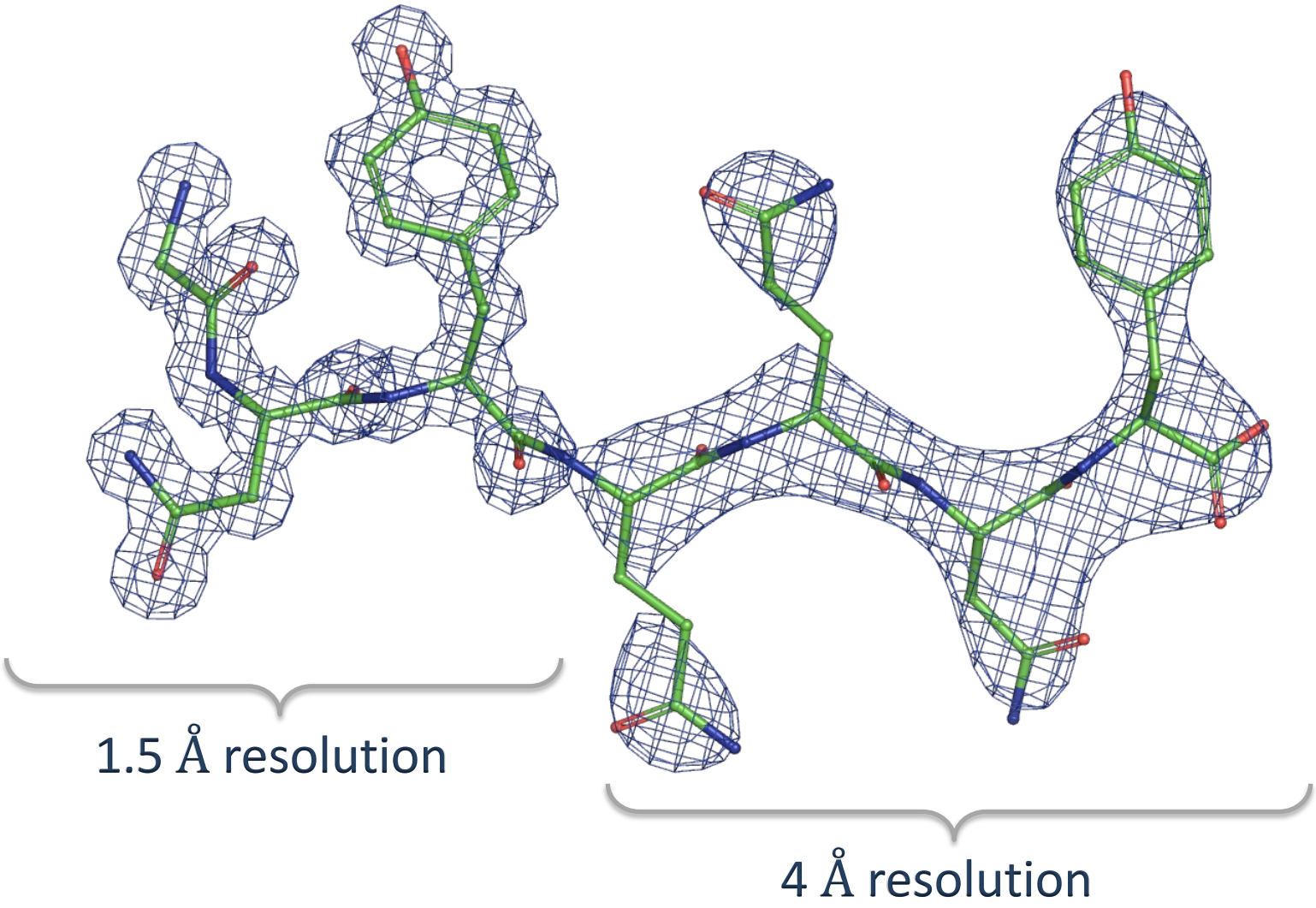
$$\rho_{calc} = \sum_{m=1}^{M} C_m \Omega(\mathbf{r}; R_m, B_m)$$

$$\Omega(\mathbf{r}; R, B) = \frac{1}{|\mathbf{r}|R}\left(\frac{1}{4\pi B}\right)^{1/2}\left[\exp\left(-\frac{4\pi^2(|\mathbf{r}|-R)^2}{B}\right) - \exp\left(-\frac{4\pi^2(|\mathbf{r}|+R)^2}{B}\right)\right]$$

- Precompute and tabulate B, C and R constants for all atoms from Periodic Table and all resolutions
  - This is done by fitting to actual Fourier maps
- To calculate the map, assign resolution to each atom in your model, then look up its parameters in precomputed tables, and compute the map

What can we do with this?

Fourier maps can be computed analytically for a model, allowing different atoms or regions (chains, domains, loops, etc.) to have distinct resolutions



What can we do with this?

- Difference maps (aka Fo-Fc in Xtal) can be calculated more accurately
- Refinement:
  - Analytic derivatives w.r.t. coordinates, occupancies, B factors
  - More accurate refinement target function
  - Local resolution aware
- Validation: CC_{MASK} aware of local resolution
- AI/ML based model building, ligand identification:
  - Fast and accurate generation of realistic density maps of given resolution for training purposes
- Accurate map-to-model fit evaluation (rotamer fitting)