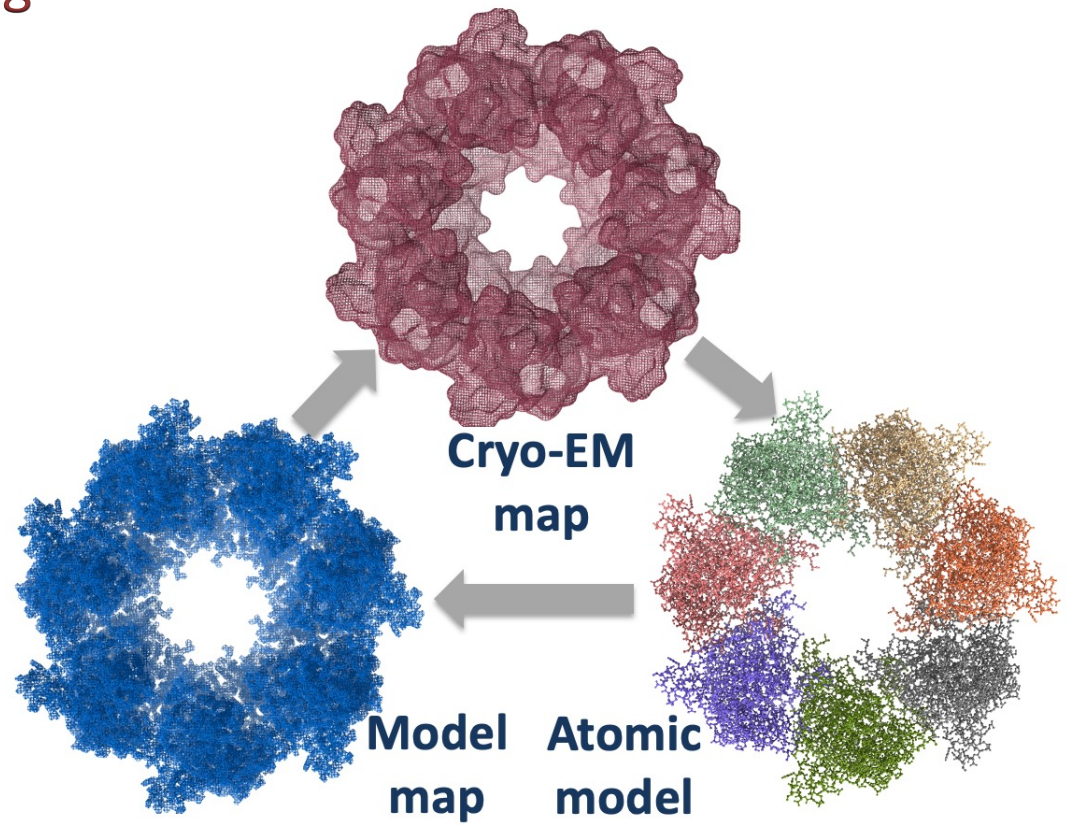
VRM: variable resolution maps

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

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

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

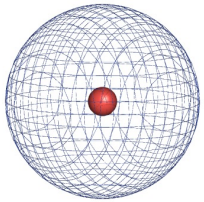
$$\rho(\mathbf{x}) = \sum_{i}^{M} \theta Z_i e^{-k|\mathbf{x}-\mathbf{r}_i|^2}$$

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

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

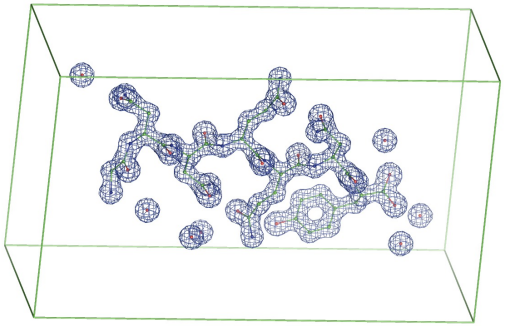
- Resolution is not accounted for

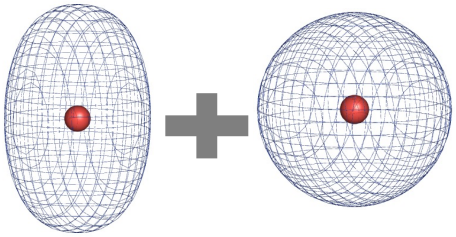
Model map: Independent Atom Model (IAM)



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

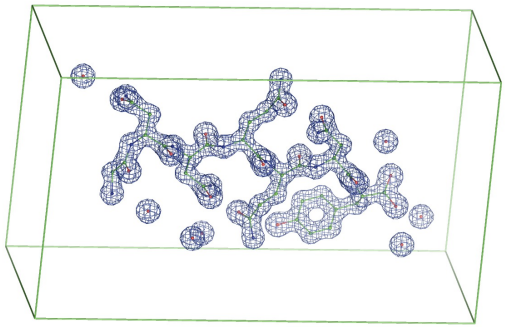
```
ATOM     25  CA  PRO A   4      31.309  29.489  26.044  1.00 57.79           C
```

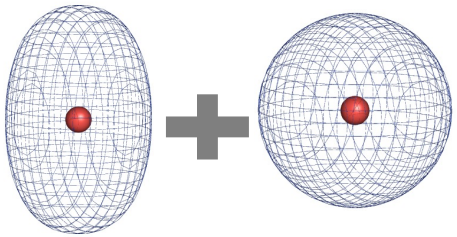




$$\rho_{MODEL}(\mathbf{r}) = \sum_{i=1}^{Natoms} \rho_{atoms}(\mathbf{r})$$

Model map





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

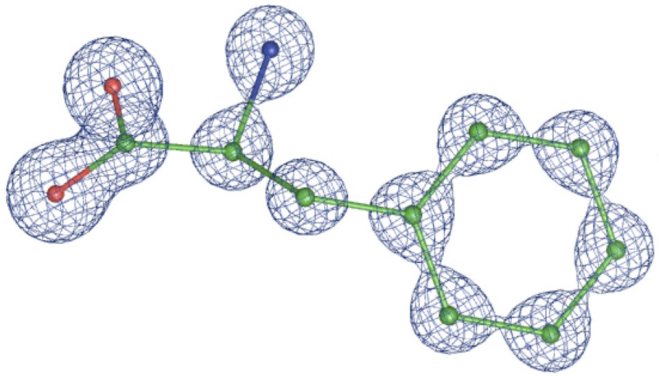
This is not the map we need for refinement, validation, model building..

In practice, we rarely need this map at all!

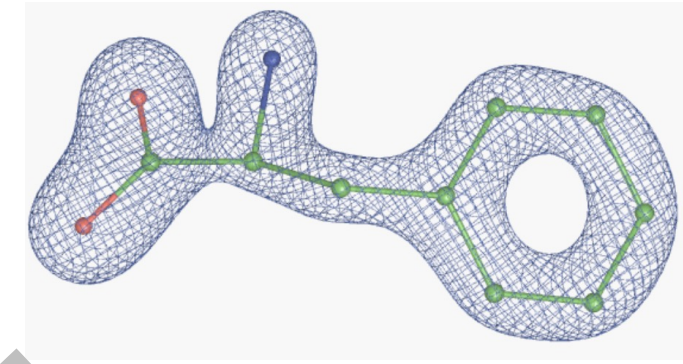
This is because it is an “infinite resolution” (exact) map, which cannot be meaningfully compared with the experimental map.

Model map suitable for calculations

Exact model map ρ_{calc}

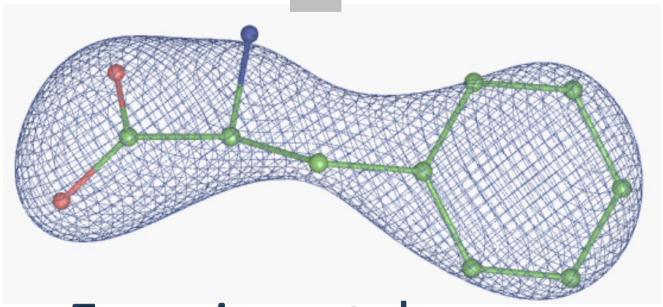


Finite resolution model map $\widetilde{\rho_{calc}}$



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

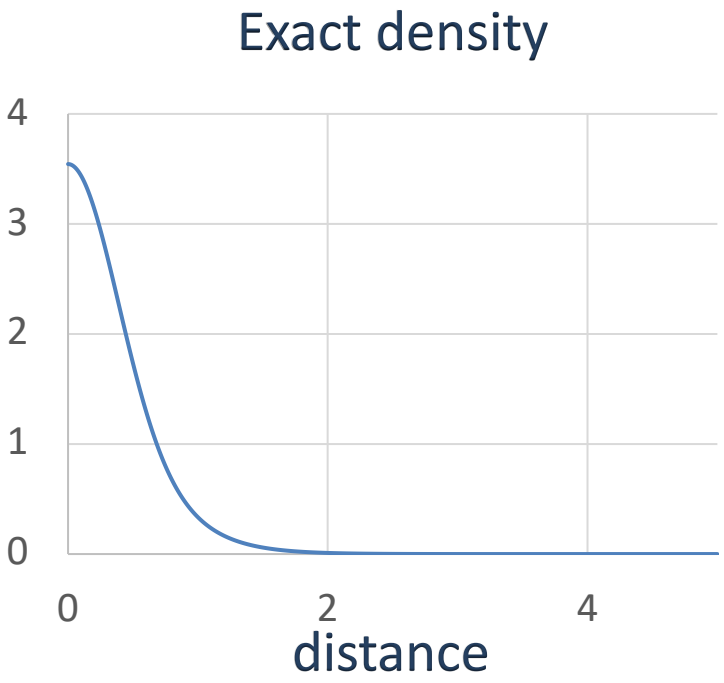
$$\mathrm{CC} = \frac{\sum \rho_{obs} \widetilde{\rho_{calc}}}{(\sum \rho_{obs}^2 \sum \widetilde{\rho_{calc}}^2)^{1/2}}$$

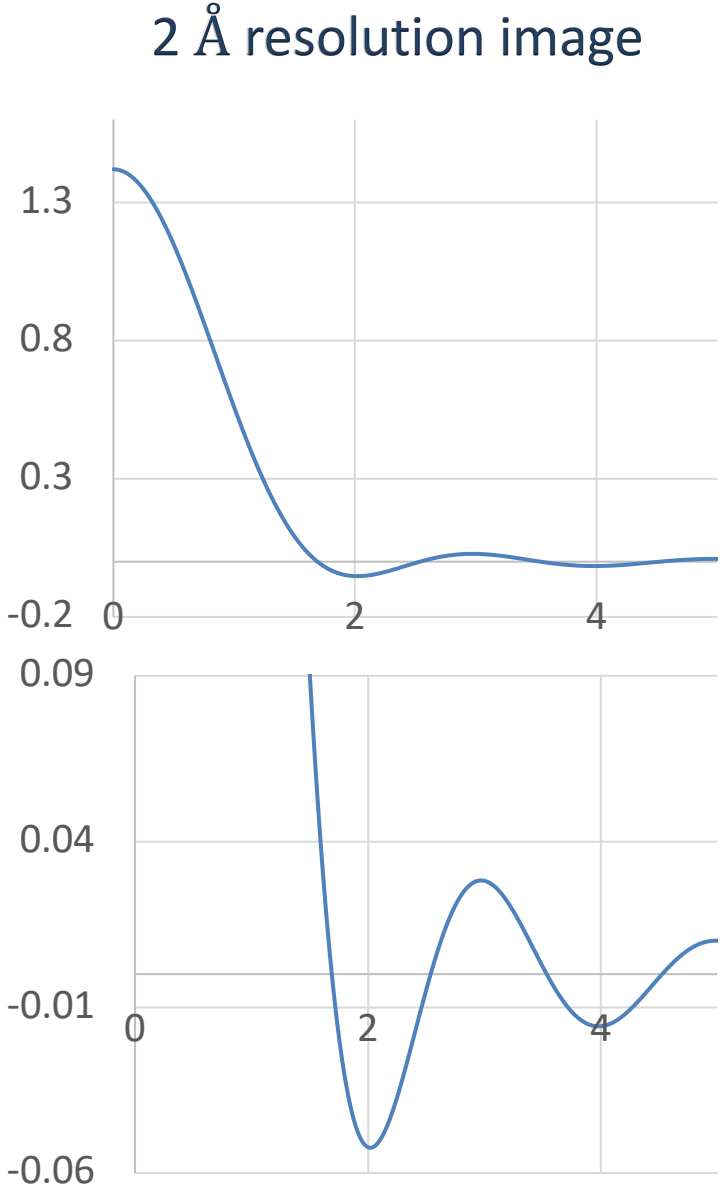


Experimental map

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

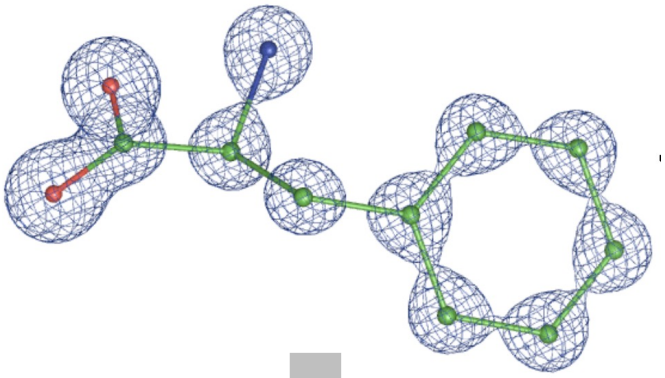
Model map illustrations for Carbon atom



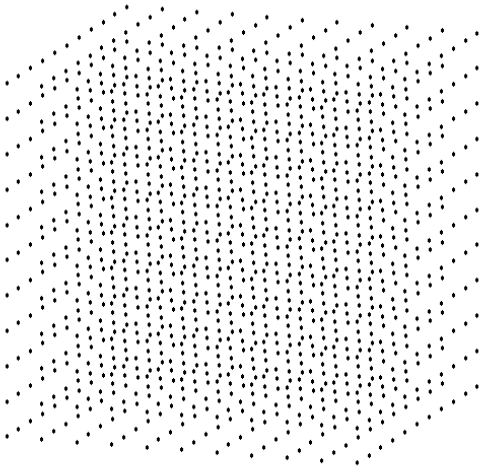


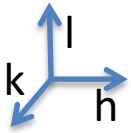
Finite resolution model map calculation (Fourier map)

Exact model map ρ_{calc}



FT $^{-1}$

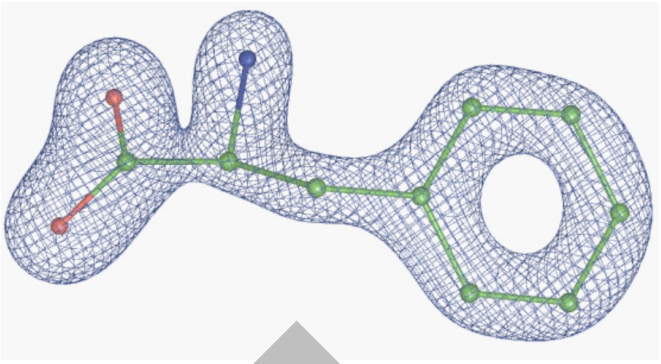




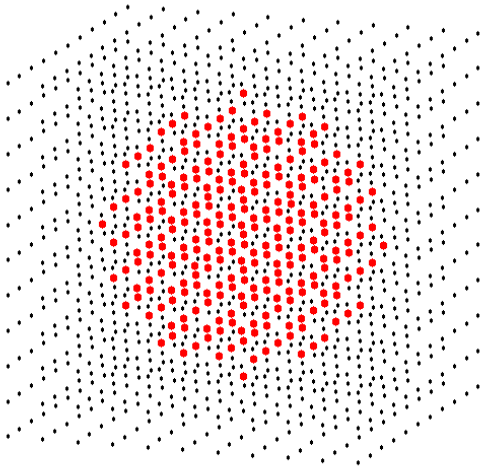
Fourier map coefficients

Truncate resolution

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

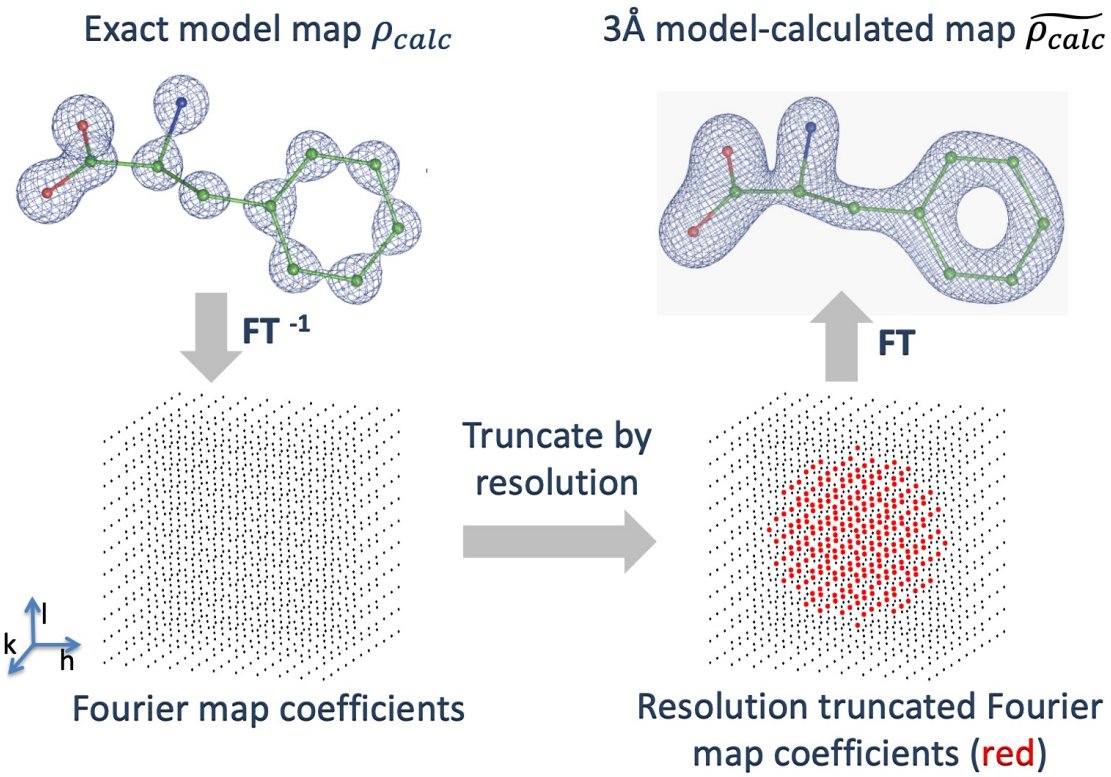


FT



Resolution truncated Fourier map coefficients (red)

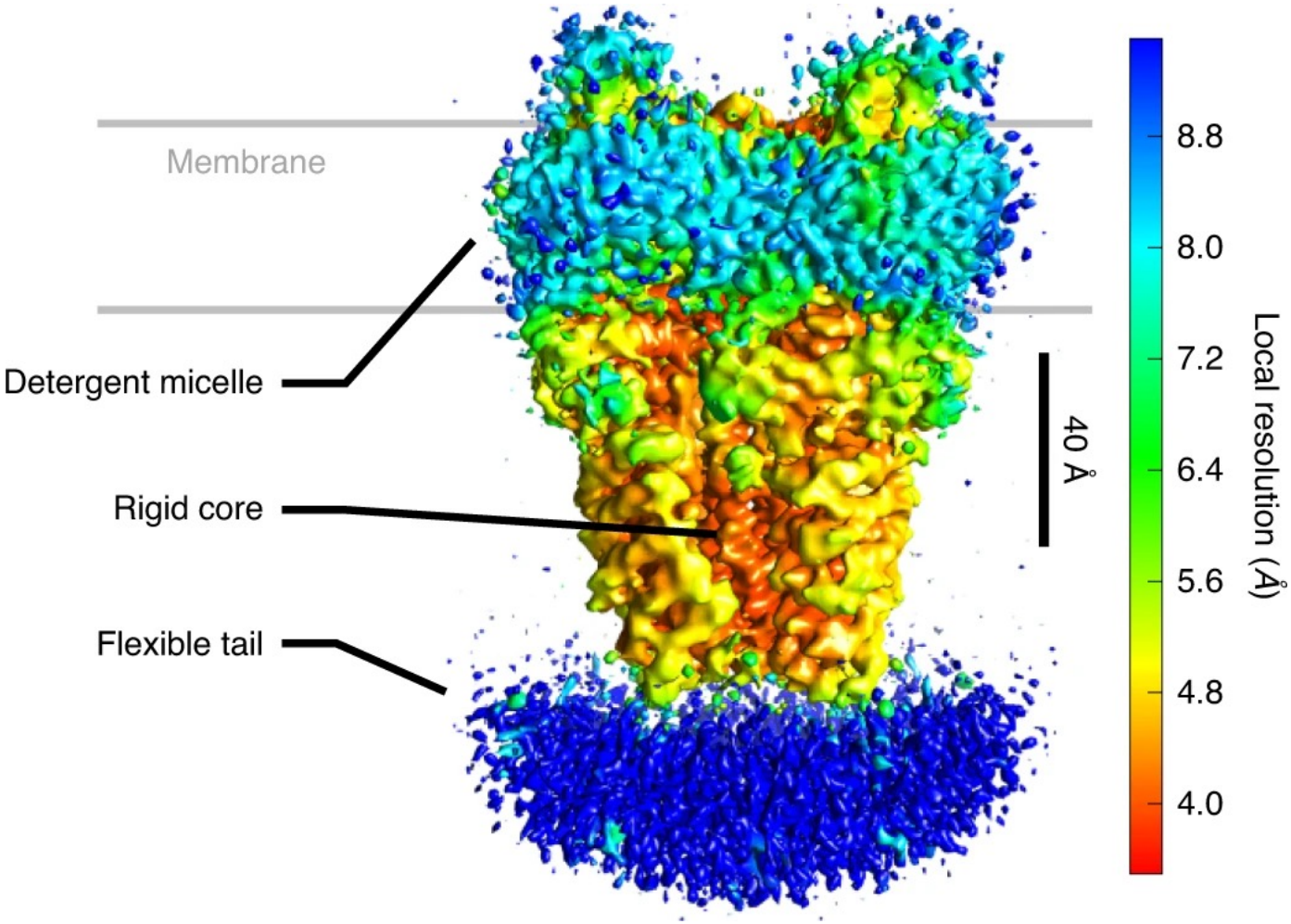
Finite resolution model map calculation (Fourier map)



Two problems with this approach:

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

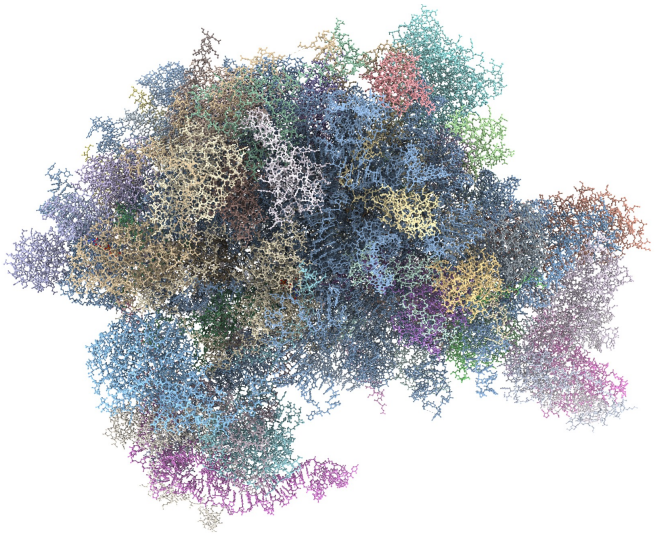
CryoEM: local resolution can vary a lot



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

Time: single function calculation

150k atoms



11 sec

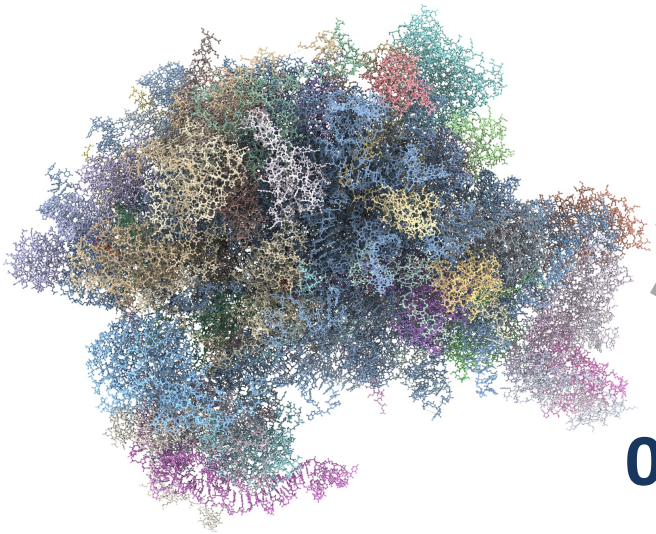
$$LS = \sum_{map} (\rho_{obs} - \widetilde{\rho_{calc}})^2$$

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

Typical refinement: ~1000 evaluations

Time: single function calculation

150k atoms



11 sec

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

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

0.02 sec

$$T = -\sum_{atoms} \rho_{obs}(x_{atom}, y_{atom}, z_{atom})$$

Used in **`phenix.real_space_refine`**

Typical refinement: ~1000 evaluations

Refinement target

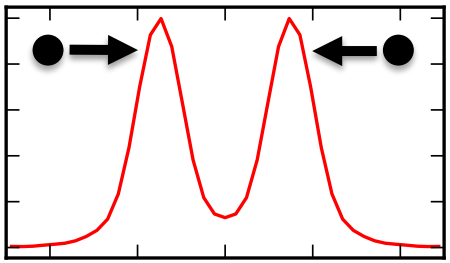
$$T = -\sum_{atoms} \rho_{obs}(x_{atom}, y_{atom}, z_{atom})$$

is much less accurate than $$LS = \sum_{map} (\rho_{obs} - \widetilde{\rho_{calc}})^2$$

because it aims at moving atoms towards nearest map peaks without assuming shape of the map:

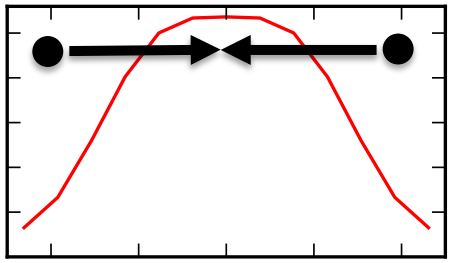
High resolution

Atom 1



Atom 2

Low resolution



- Moving atoms to nearest peaks ≠ making correct model
- Can't refine B-factors and occupancies
- Lower resolution = less accurate
- Need to use a lot of geometric restraints in refinement

Solution comes from this series of publications

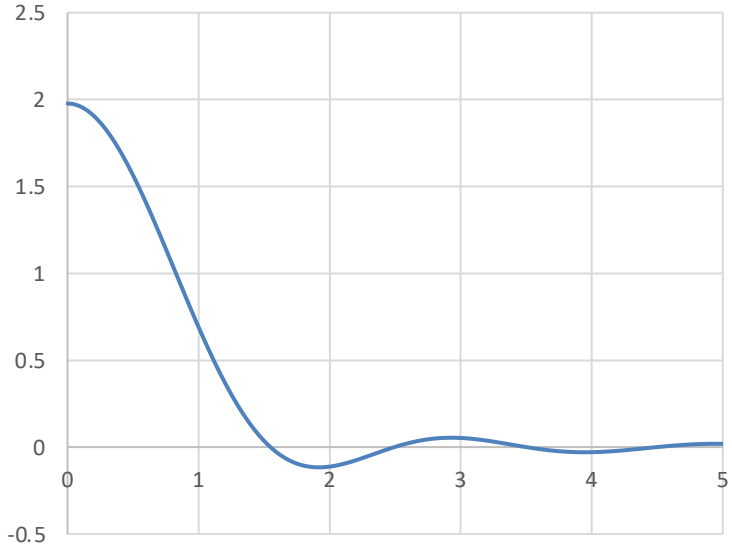
Urzhumtsev, A., Lunin, V.Y. (2022) "**Analytic modeling of inhomogeneous-resolution maps in cryo-electron microscopy and crystallography**". *IUCr Journal*, **9**, 728-734.

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.

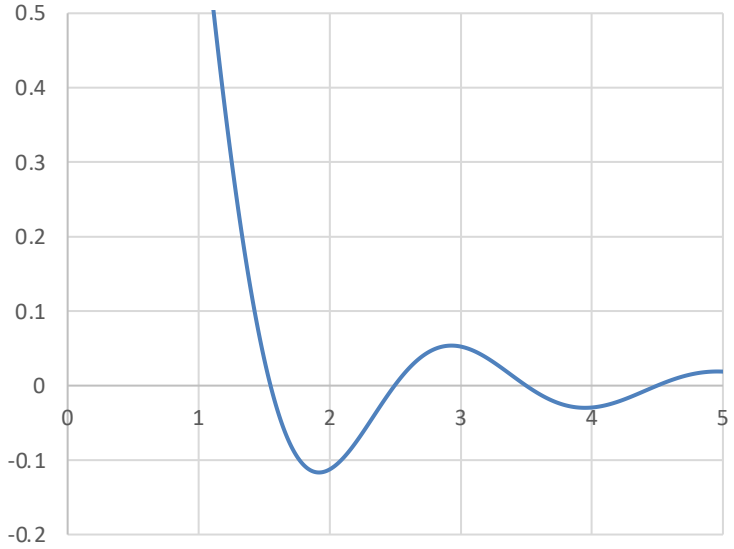
Urzhumtseva, L., Lunin, V.Y., Urzhumtsev, A. (2023) "**Algorithms and programs for the shell decomposition of oscillating functions in space**". *J. Appl.Cryst.,* **56**, 302-311.

Idea at a glance

3 Å resolution image of C atom (2D)



Zooming on ripples



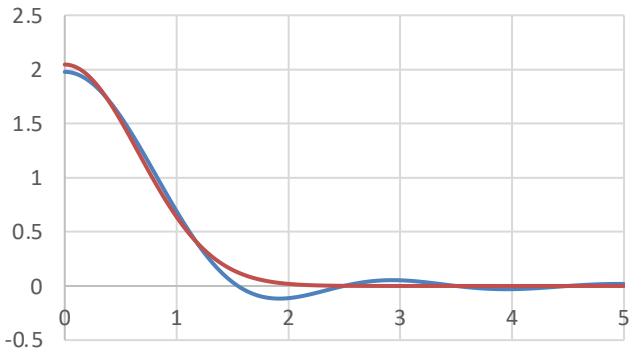
This image can be approximated with desired accuracy by this function

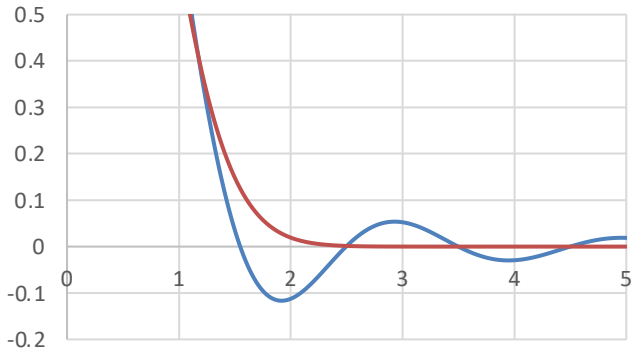
$$\sum_{m=1}^{M} C_m \Omega(\mathbf{r};\, R_m, B_m)$$

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

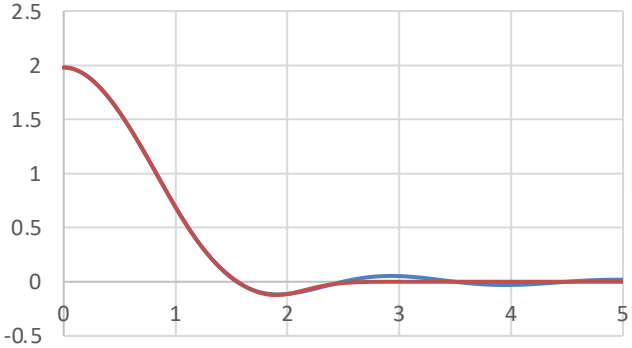
More terms M = more accurate approximation

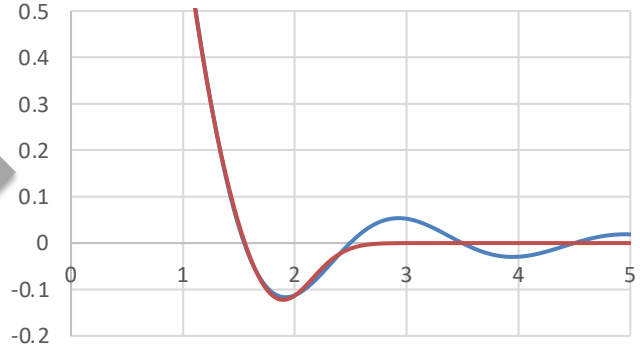
$$\sum_{m=1}^{M} C_m \Omega(\mathbf{r}; R_m, B_m)$$



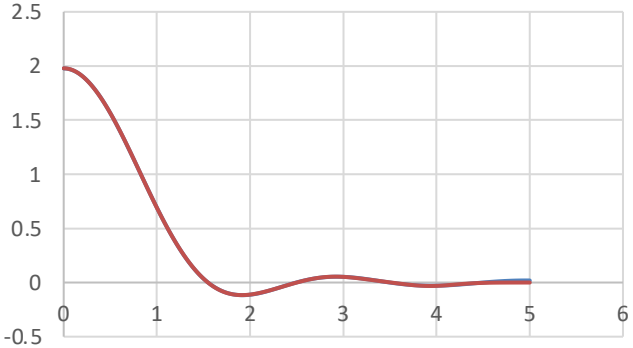


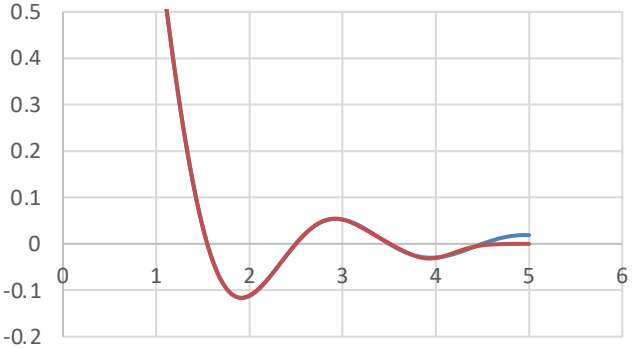
1 term





2 terms





4 terms

Practicalities

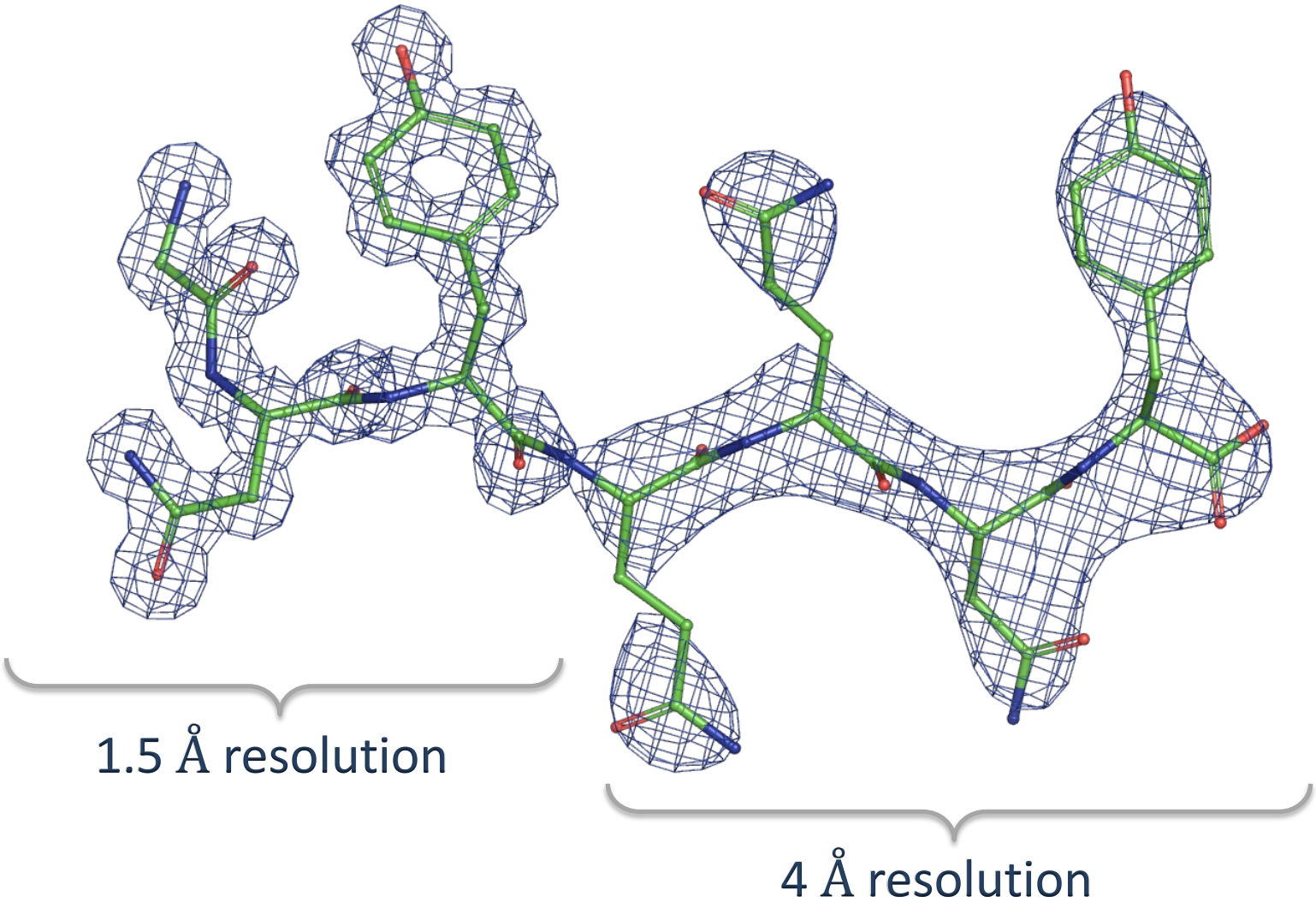
$$\rho_{calc} = \sum_{m=1}^{M} C_m \Omega(\mathbf{r}; R_m, B_m)$$

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

- 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)

Implementation

- CCTBX

  `cctbx_project/cctbx/maptbx/bcr`