



COMPUTATIONAL CRYSTALLOGRAPHY INITIATIVE

Refinement and *phenix.refine*: 34 random questions and answers

Pavel Afonine

Computational Crystallography Initiative
Physical Biosciences Division
Lawrence Berkeley National Laboratory, Berkeley CA, USA

What is the best citation for phenix.refine?

- Latest paper:

Towards automated crystallographic structure refinement with phenix.refine

P. V. Afonine, R. W. Grosse-Kunstleve, N. Echols, J. J. Headd, N. W. Moriarty, M. Mustyakimov, T. C. Terwilliger, A. Urzhumtsev, P. H. Zwart and P. D. Adams

Acta Cryst. (2012). D68, 352-367

I can't find bulk-solvent k_{sol} and B_{sol} , as well as overall anisotropic matrix B_{cart} in phenix.refine output...

- Since April 2012 phenix.refine and other tools use a new better bulk-solvent model and overall anisotropic scaling.
- It is faster and almost always produces lower R-factors compared to previous model.
- More details:

Bulk-solvent and overall scaling revisited: faster calculations, improved results

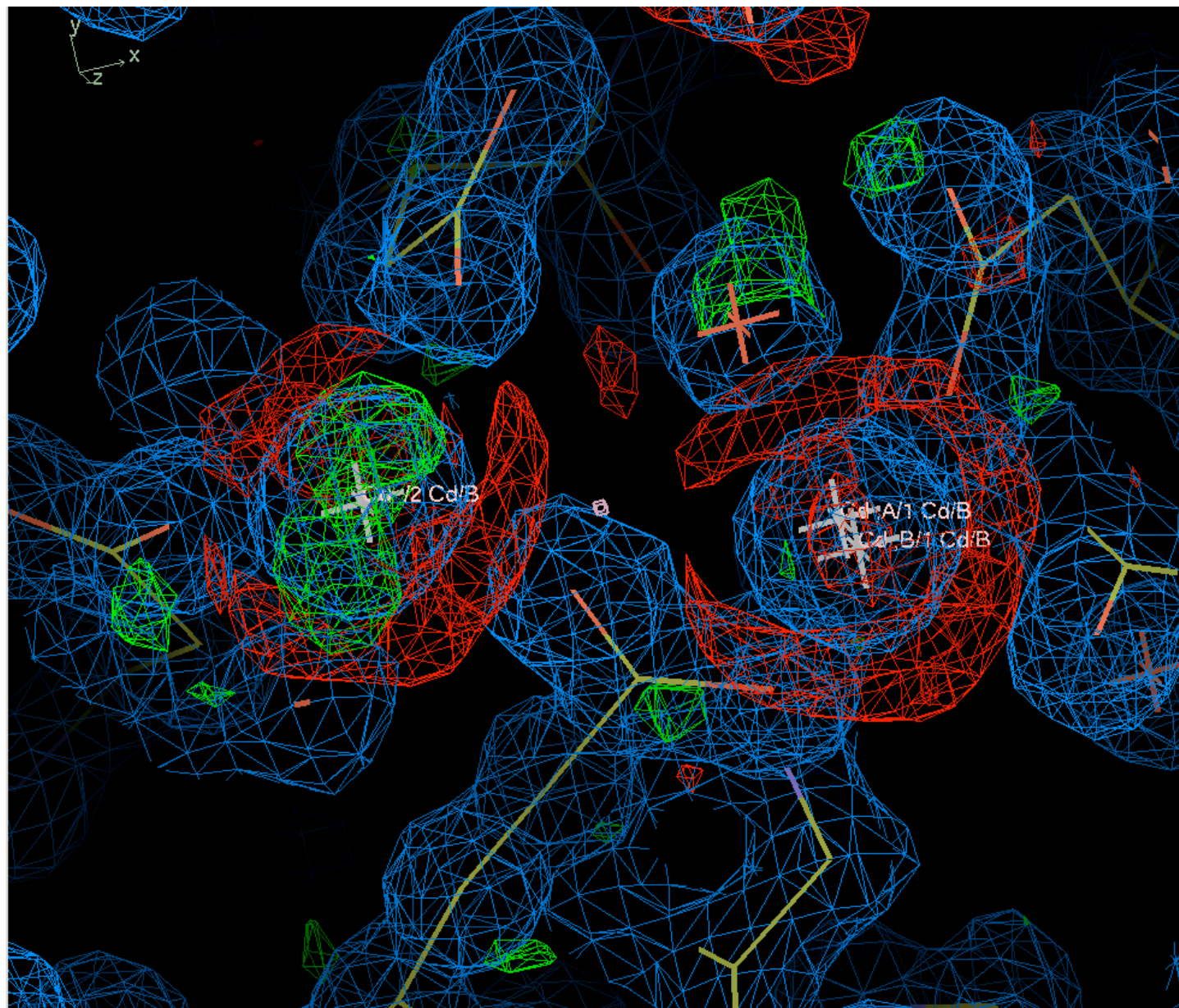
P. V. Afonine, R. W. Grosse-Kunstleve, Adams & A. Urzhumtsev
Acta Cryst. (2013). D69

**BUSTER-TNT produced R-factors that are different from
phenix.refine... Why?**

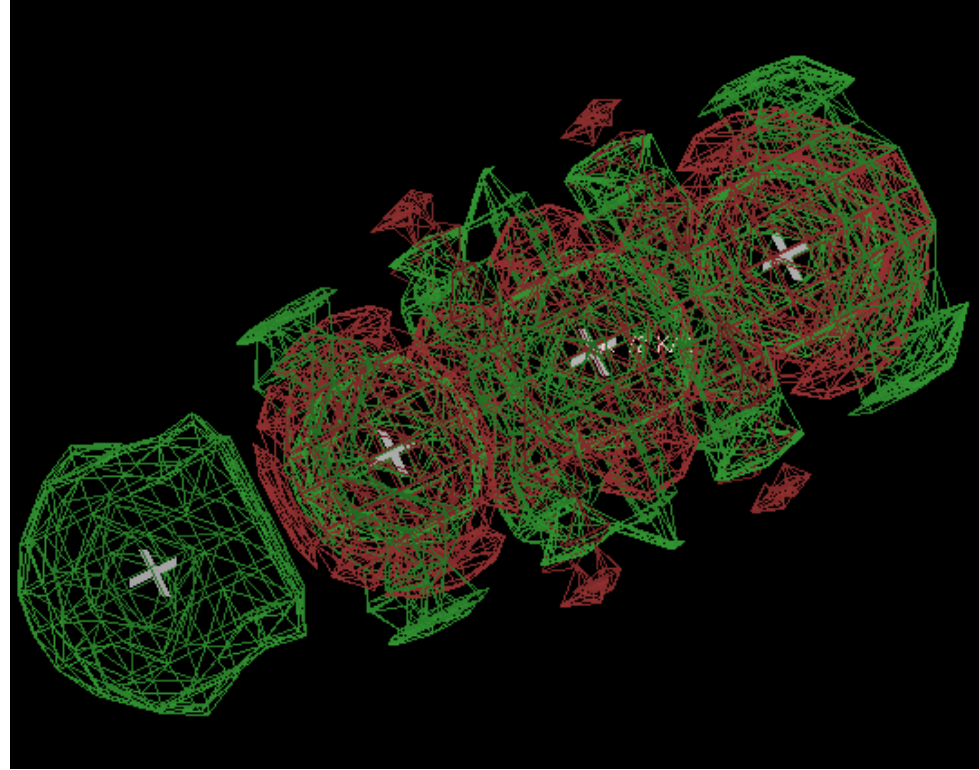
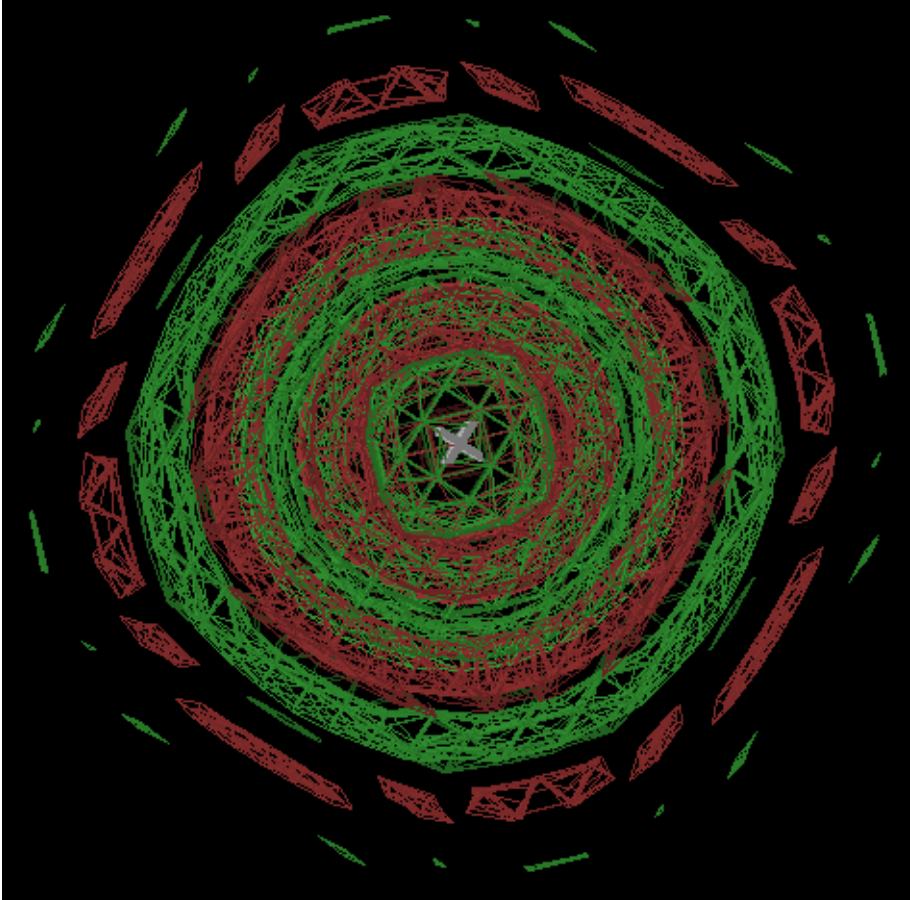
- In part, this is because BUSTER-TNT uses a different formula to compute R-factor, which makes comparison of R-factors between two programs nonsensical.

**I see positive and negative peaks around heavy atoms?
What's wrong?**

Spurious peaks around heavy atoms

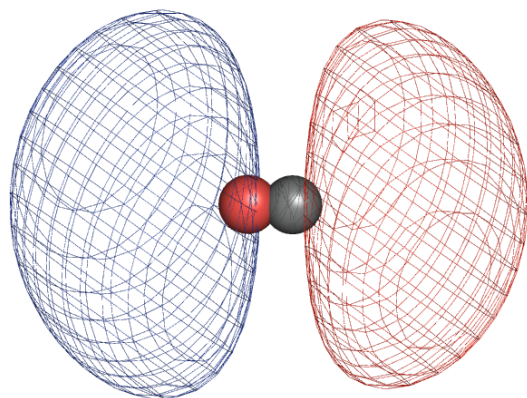


Fourier truncation ripples

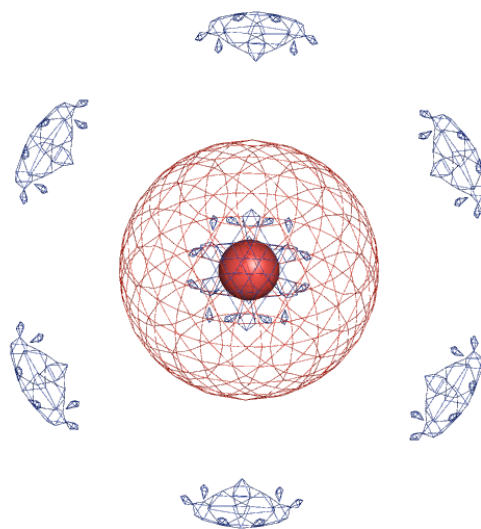


Errors in position or/and in occupancy or/and in B-factor

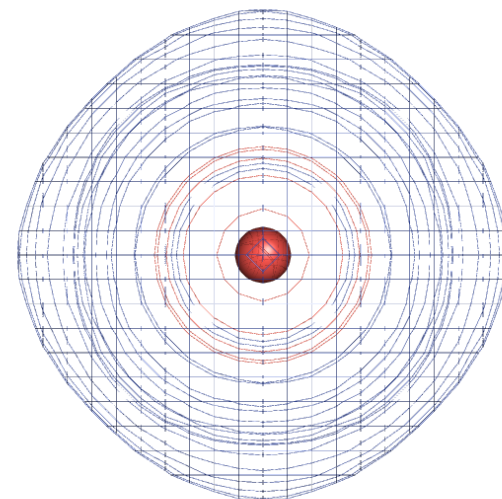
Error in position



Error in occupancy

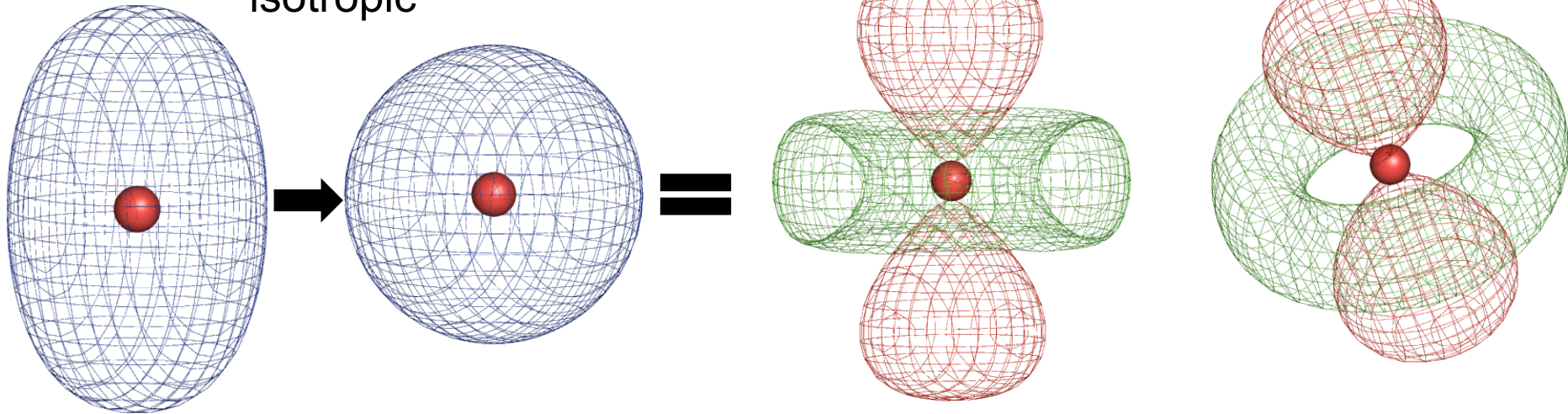


Error in B-factor



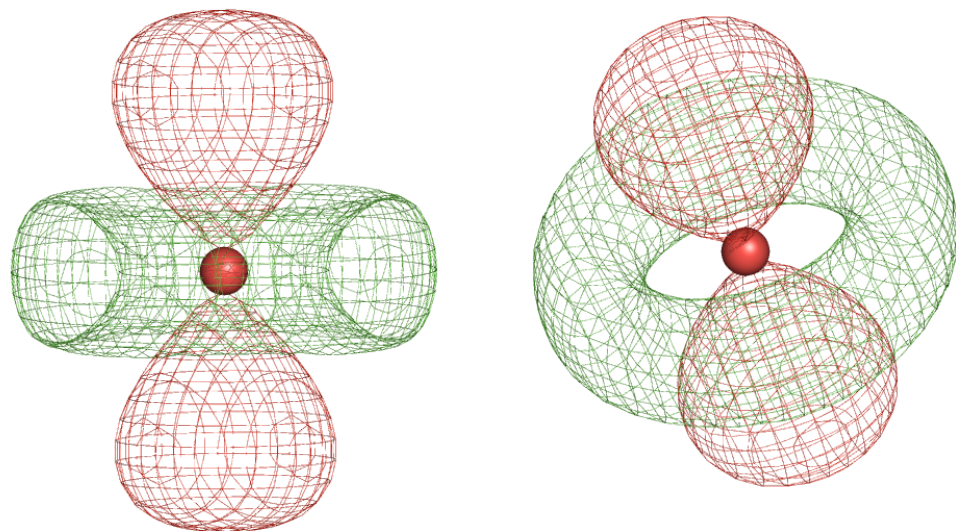
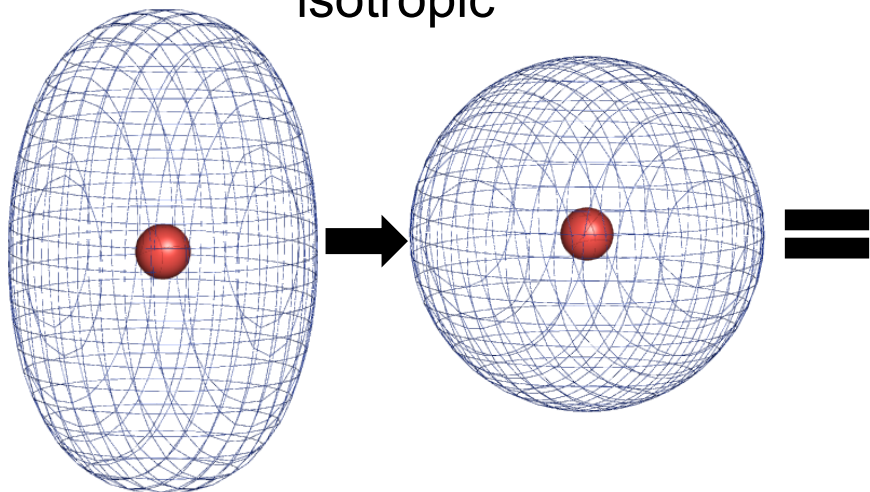
Errors in position or/and in occupancy or/and in B-factor

Model anisotropic atom with isotropic

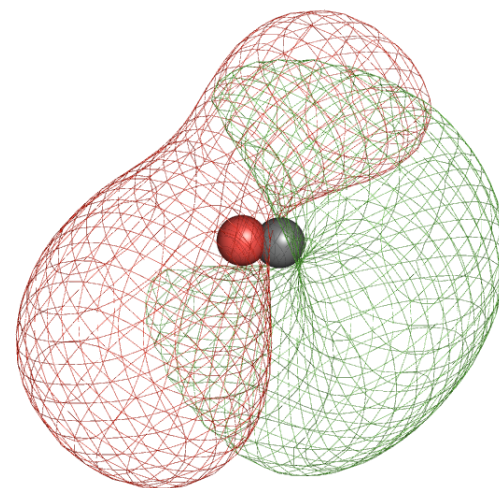


Errors in position or/and in occupancy or/and in B-factor

Model anisotropic atom with isotropic



Add positional error

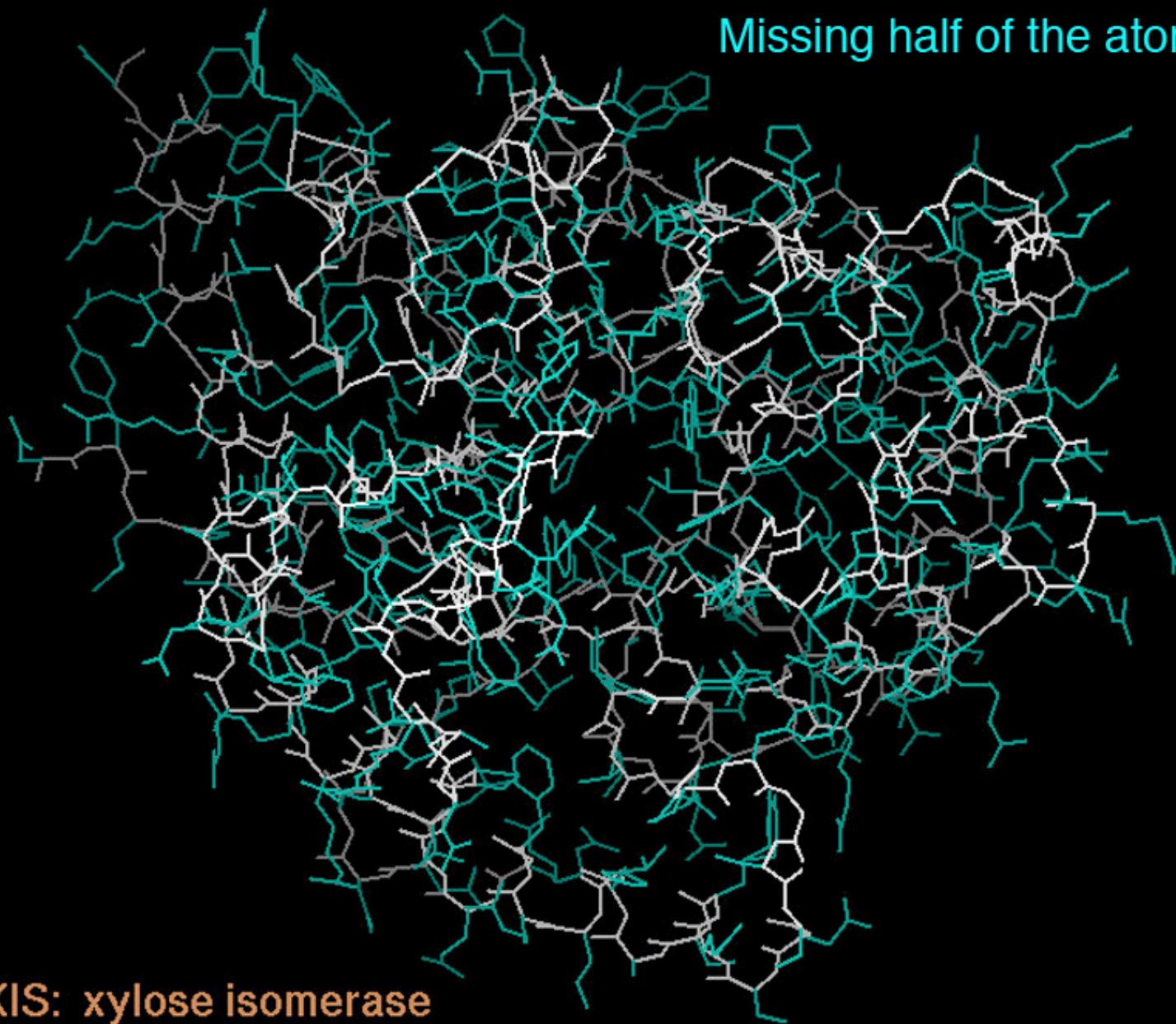


Summary

- If this is Fourier truncation effect – there is nothing one can do.
- If these are errors in atomic parameters:
 - Do more refinement macro-cycles
 - Refine occupancy of the metal
 - Refine anisotropic ADP of metal only
 - If it is anomalous scatterer: refine f' and f''

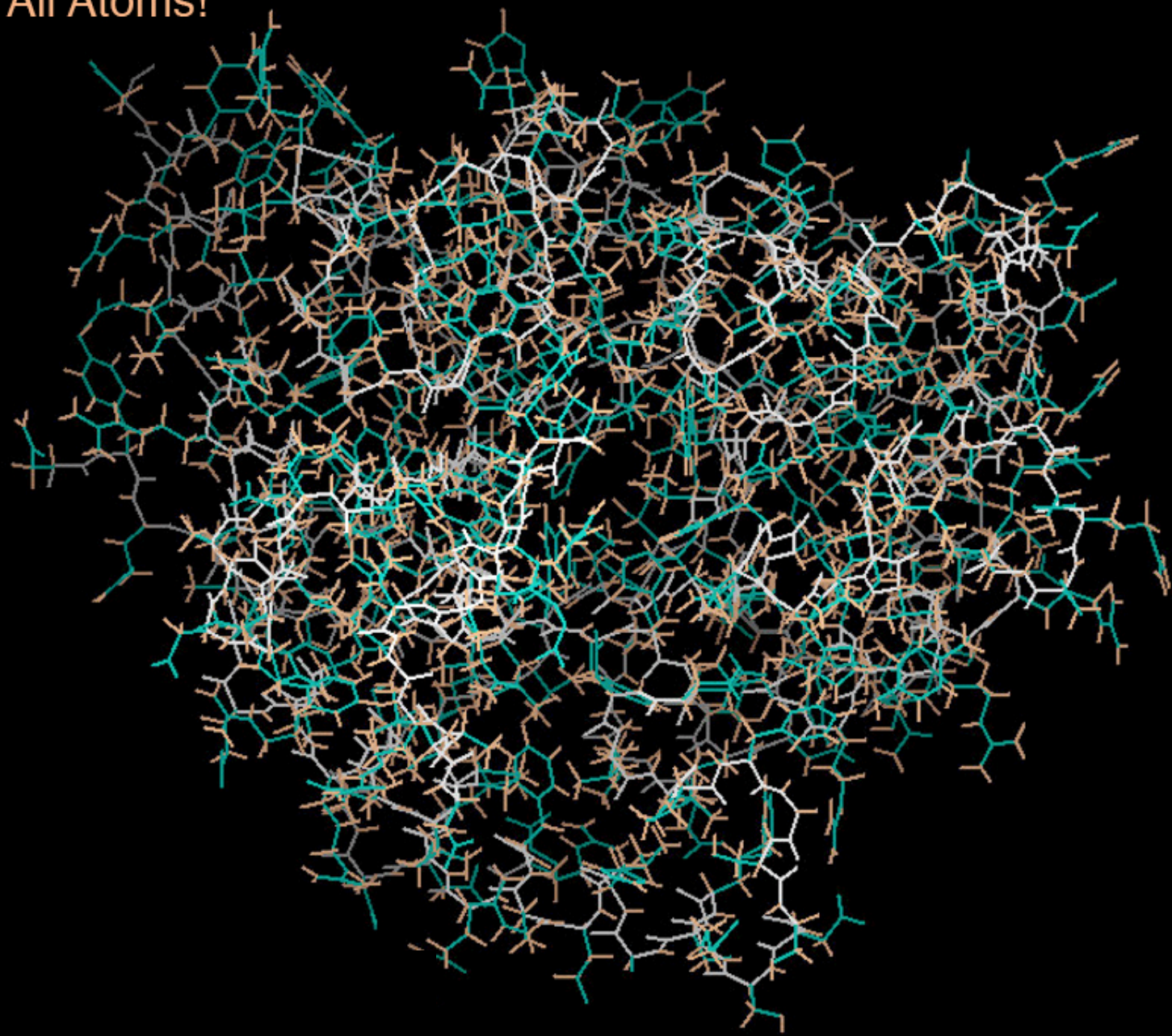
We can't see hydrogen atoms in X-ray map at typical resolutions..
Why use them in refinement and keep in output PDB file?

Missing half of the atoms!



4XIS: xylose isomerase

All Atoms!



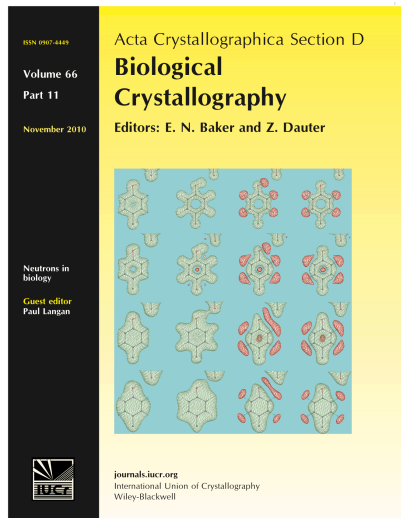
Hydrogen atoms in refinement

■ phenix.refine: options for handling H atoms at **any** resolution:

- Riding model (low-high resolution)
- Individual atoms (ultrahigh resolution or neutron data)
- Account for scattering contribution or just use to improve the geometry

■ Using H atoms in refinement:

- Improve R-factors
- Improve model geometry (remove bad clashes)
- Model residual density at high resolution or in neutron maps



www.phenix-online.org

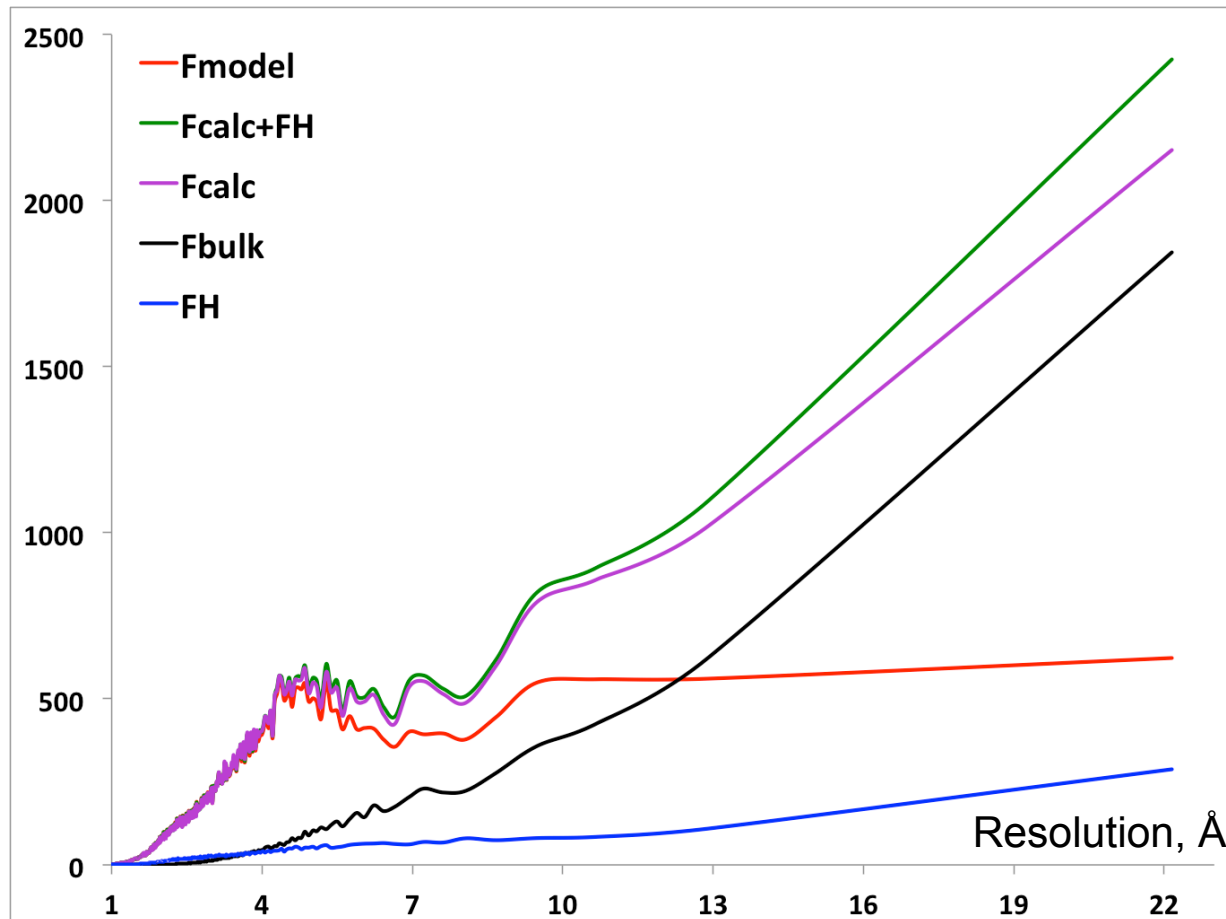
Review and developments:
Afonine, et al. (2010). Acta Cryst. D66, 1153-1163.

Afonine P.V. & Adams P.D. (2012). On the contribution of hydrogen atoms to X-ray scattering

Contribution of hydrogen atoms to Fcalc

- Total model structure factor:

$$F_{\text{model}} = \underbrace{F_{\text{calc}}}_{\text{Non-H atoms}} + \underbrace{F_{\text{H}}}_{\text{H atoms}} + \underbrace{F_{\text{bulk}}}_{\text{Bulk-solvent}}$$



I now know using H atoms in refinement clearly is a good idea.
When should I add them to my model?

- There is no definitive answer.
- Using hydrogen atoms towards the end may be better:
 - Faster refinements
 - Easier model building in Coot
 - Less chances model traps in a local minimum

I get $R_{\text{WORK}}=18\%$ and $R_{\text{FREE}}=23\%$: are they good? Am I ready to PDB deposit the structure?

How to tell if R-factor is *good*

- **Question:** “ $R_{\text{WORK}}=18\%$ and $R_{\text{FREE}}=23\%$: are they good?”
 - Question does not make sense unless data resolution is specified
- **Answer:**
 - Yes, it’s likely a good result if the data resolution is around 2.5 Å.
 - No, it is very bad result, if the data resolution is 1.0 Å or higher.
- One can ask similar questions about other parameters, such as bond/angles RMSDs, average B-factors, etc...

Rwork and Rfree: typical values depend on resolution

- Say you are refining a structure at 1.0 Å resolution and the R-factors are:
 $R_{\text{WORK}} = 18\%$ and R_{FREE} is 23%.

– Are these values good? Am I done with refinement?

- PDB statistics: histograms for R_{WORK} , R_{FREE} , $R_{\text{FREE}} - R_{\text{WORK}}$ for all similar structures:

R_{WORK} at 0.9-1.1Å

| | |
|---------------------|------------------------|
| 0.10 - 0.12: | 68 |
| 0.12 - 0.14: | 94 |
| 0.14 - 0.16: | 73 |
| 0.16 - 0.18: | 17 <<< |
| 0.18 - 0.20: | 12 |
| 0.20 - 0.21: | 3 |
| 0.21 - 0.23: | 5 |
| 0.23 - 0.25: | 0 |
| 0.25 - 0.27: | 0 |
| 0.27 - 0.29: | 2 |

R_{FREE} at 0.9-1.1Å

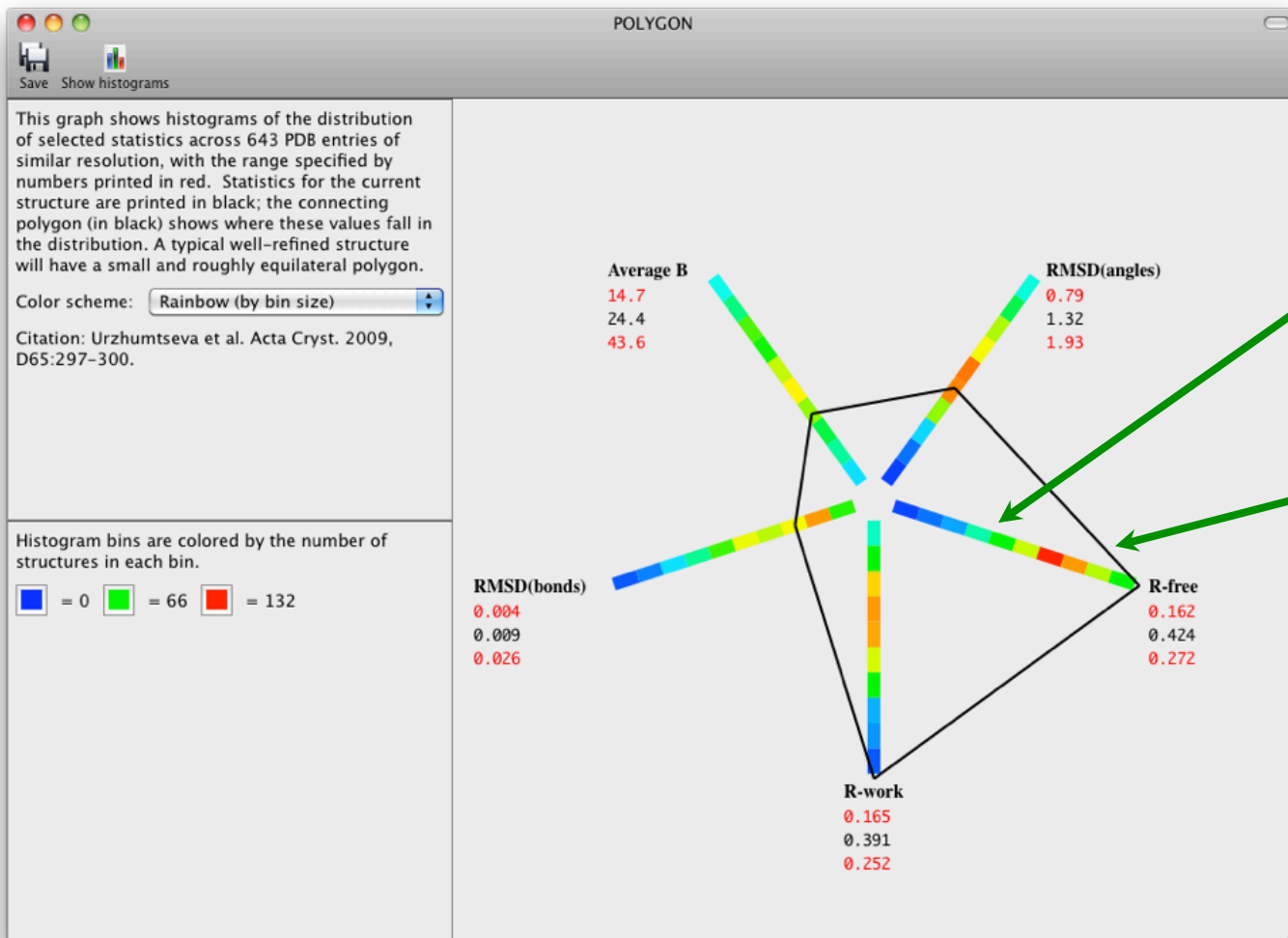
| | |
|---------------------|-----------------------|
| 0.11 - 0.13: | 16 |
| 0.13 - 0.15: | 56 |
| 0.15 - 0.17: | 97 |
| 0.17 - 0.18: | 69 |
| 0.18 - 0.20: | 14 |
| 0.20 - 0.22: | 12 |
| 0.22 - 0.24: | 3 <<< |
| 0.24 - 0.26: | 4 |
| 0.26 - 0.28: | 1 |
| 0.28 - 0.30: | 2 |

$R_{\text{FREE}} - R_{\text{WORK}}$ at 0.9-1.1Å

| | |
|---------------------|------------------------|
| 0.00 - 0.01: | 8 |
| 0.01 - 0.01: | 22 |
| 0.01 - 0.02: | 56 |
| 0.02 - 0.03: | 62 |
| 0.03 - 0.03: | 58 |
| 0.03 - 0.04: | 29 |
| 0.04 - 0.04: | 14 |
| 0.04 - 0.05: | 10 <<< |
| 0.05 - 0.06: | 6 |
| 0.06 - 0.06: | 9 |

- **Answer:** the R-factors are not good, the structure needs some more work.

POLYGON: Graphical comparison of statistics versus the PDB



Colored bars are one-dimensional histograms showing distribution of values for structures at similar resolution

The black polygon shows where the statistics for the user's structure fall in each histogram

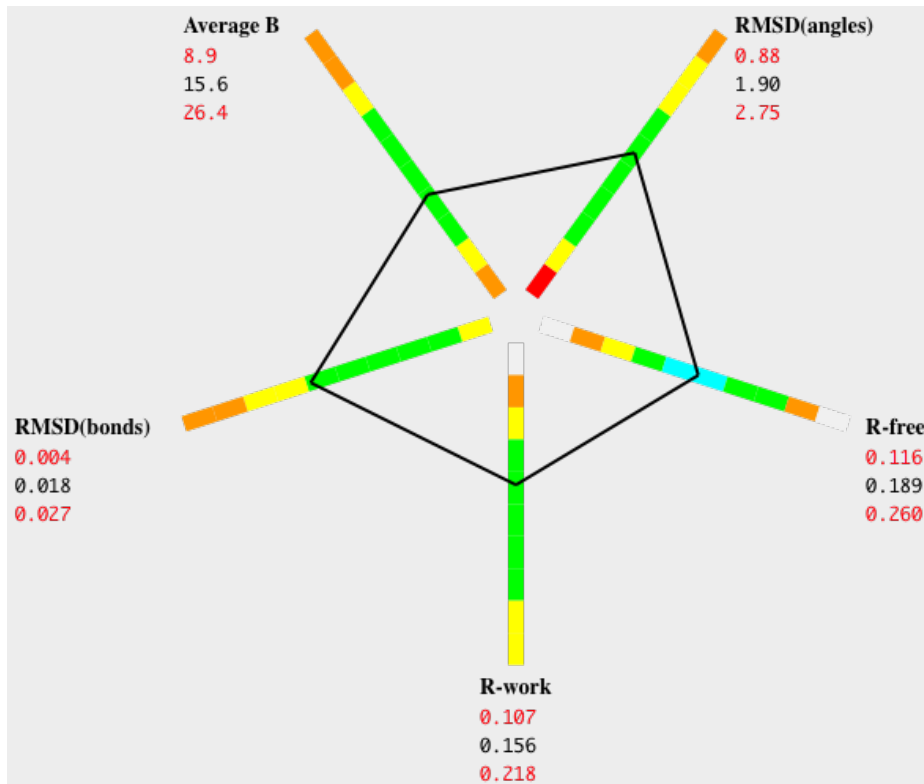
The structure used to generate this figure has good geometry relative to the PDB, but very poor R-factors.

Crystallographic model quality at a glance.

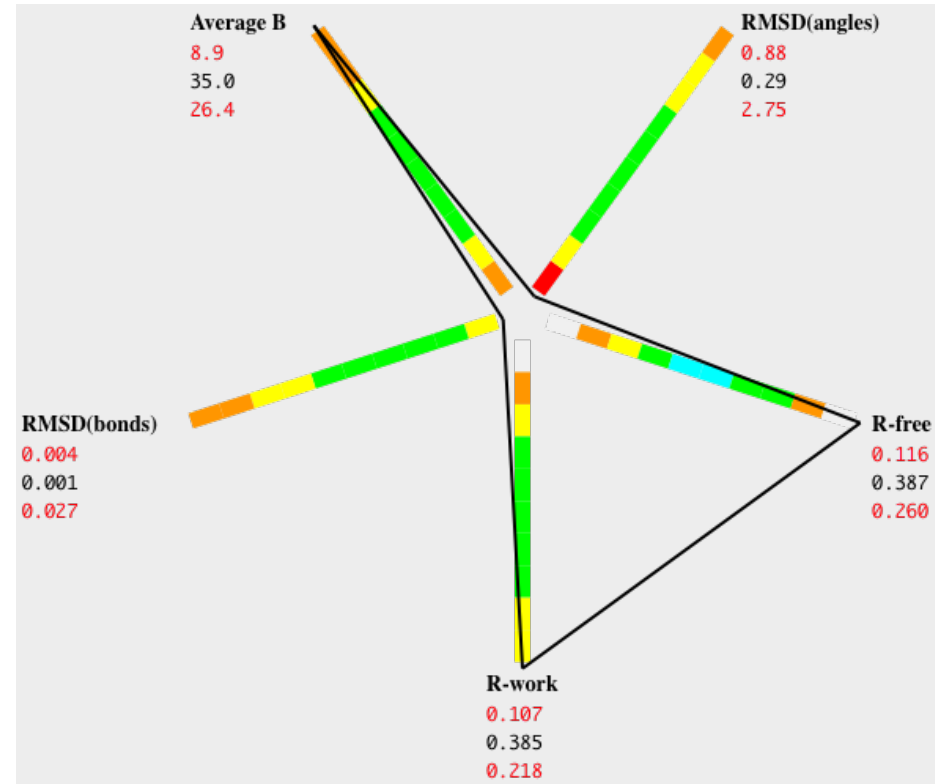
L.Urzhumtseva, P.V.Afonine, P.D.Adams & A.Urzhumtsev. *Acta Cryst.* D65, 297-300 (2009)

POLYGON

Likely good model



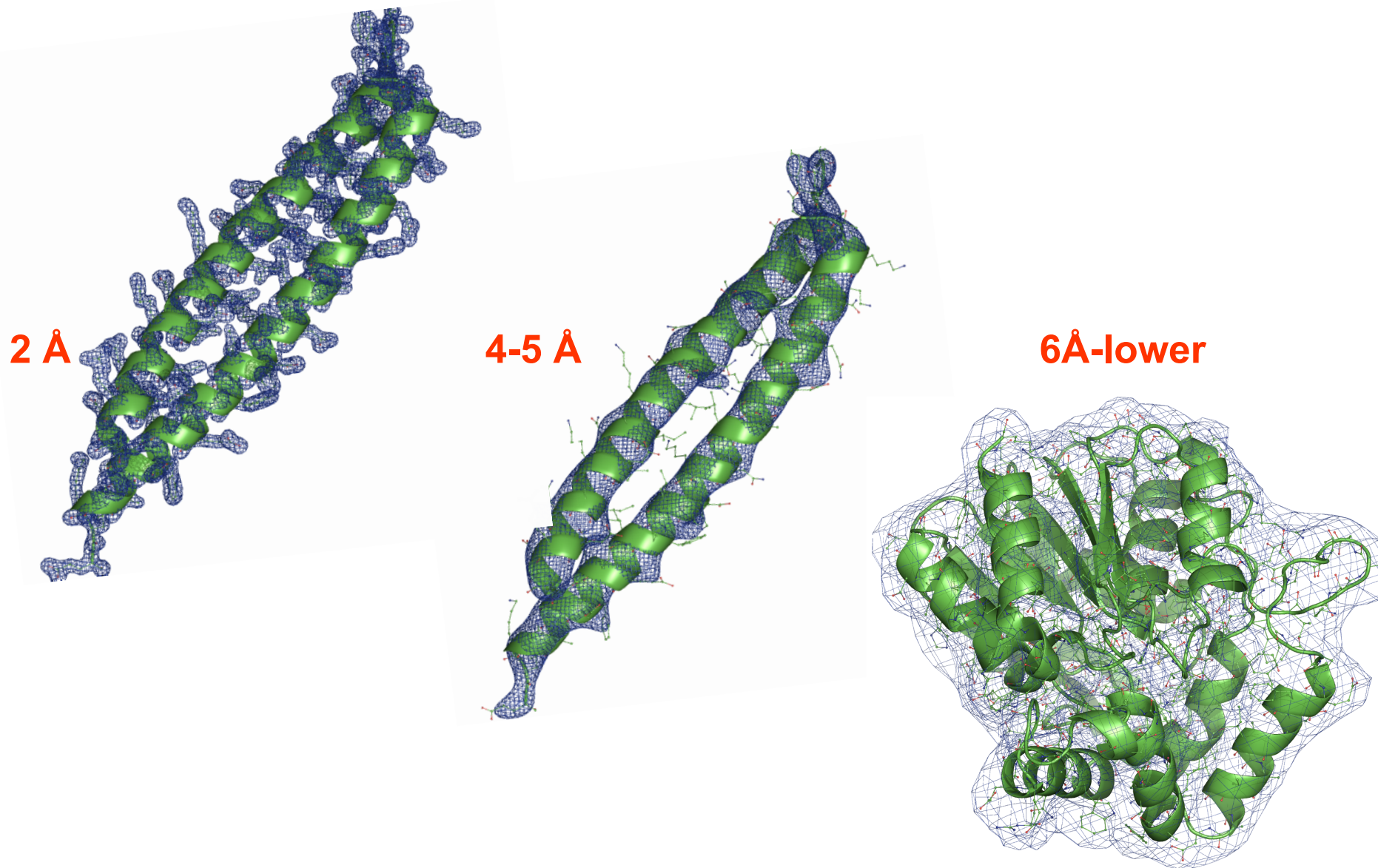
This model needs some attention



Ramachandran plot is like Rfree for geometry validation, so why should I ever use it as restraints?

Specific restraints for refinement at low and very low resolution

- At low resolution electron density map is not informative enough to maintain secondary and higher level structural organization...

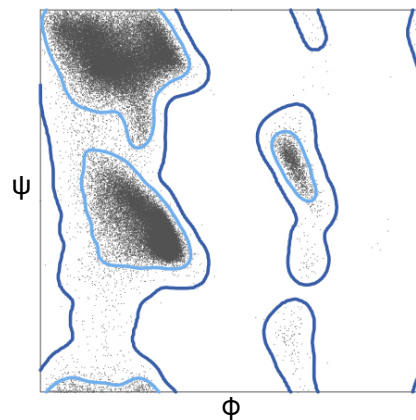


Specific restraints for refinement at low and very low resolution

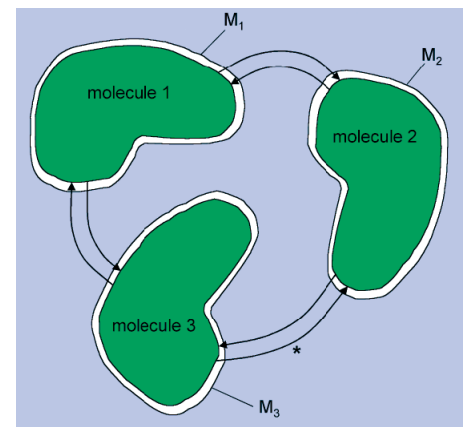
... need to use even more information:



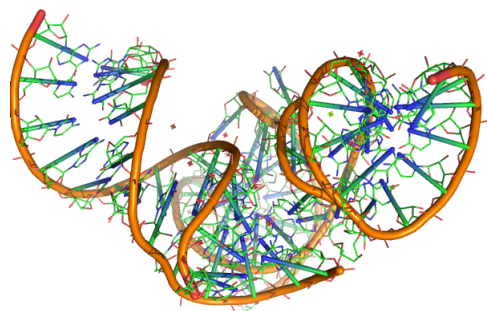
– Reference model



– Ramachandran plot restraints



– NCS restraints or constraints



– Secondary-structure restraints

$$T_{\text{RESTRAINTS}} = T_{\text{BOND}} + T_{\text{ANGLE}} + \dots + T_{\text{NCS}} + T_{\text{RAMACHANDRAN}} + T_{\text{REFERENCE}} + \dots$$

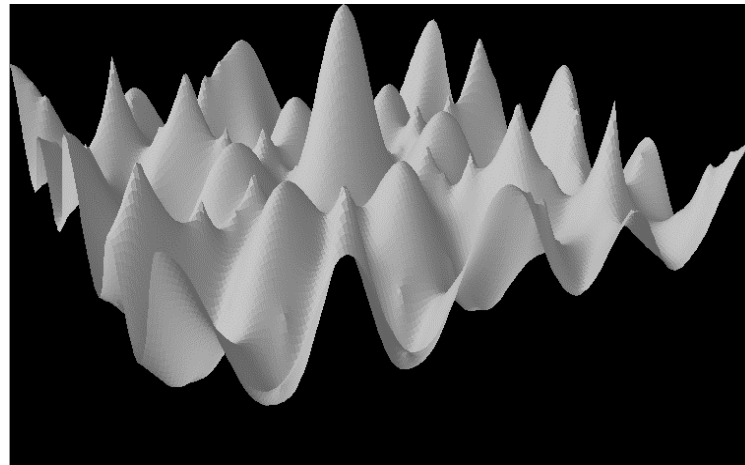
- Normally one should not use Ramachandran restraints
- In case of low resolution some residues may notoriously become Ramachandran plot outliers after refinement. In this case:
 - Fix those residue manually first, then
 - Enable Ramachandran plot restraints which will keep them from becoming outliers

phenix.refine produced a model with Rfree=24.3, then I tried program X and it gave me much better result: Rfree=23.9%.

Now I'm switching to program X, but I would like to know why phenix.refine produced a worse result?

Refinement convergence

- Profile of a refinement function is very complex

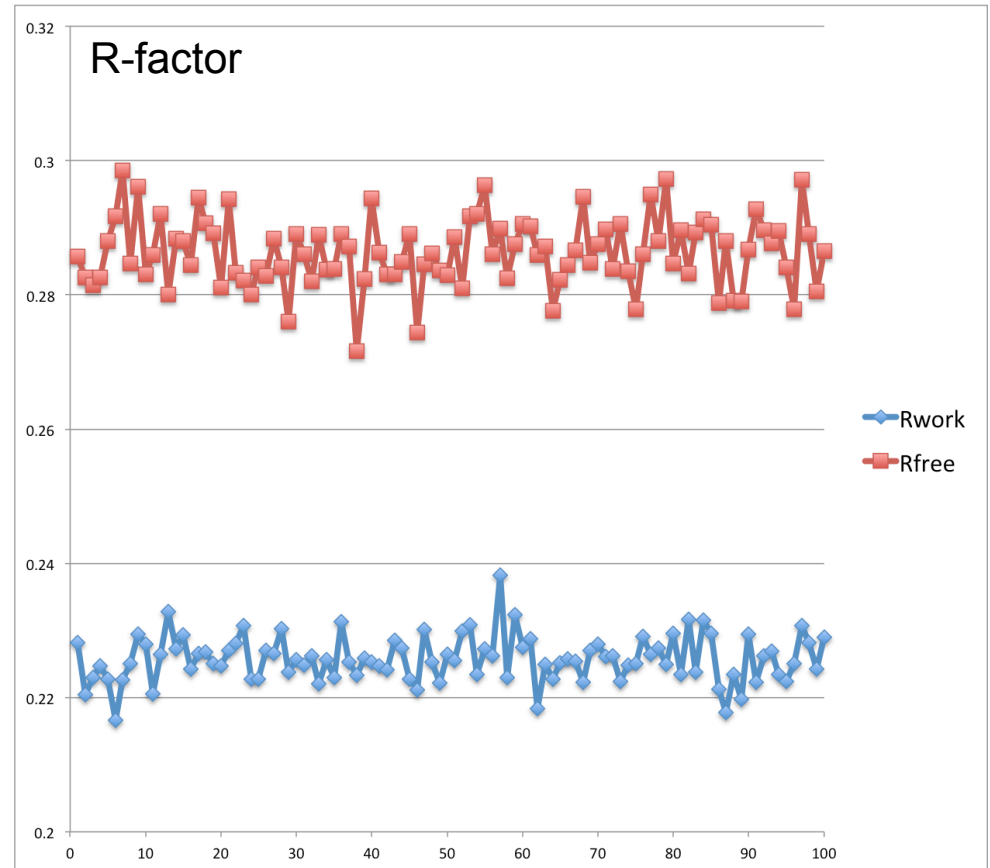
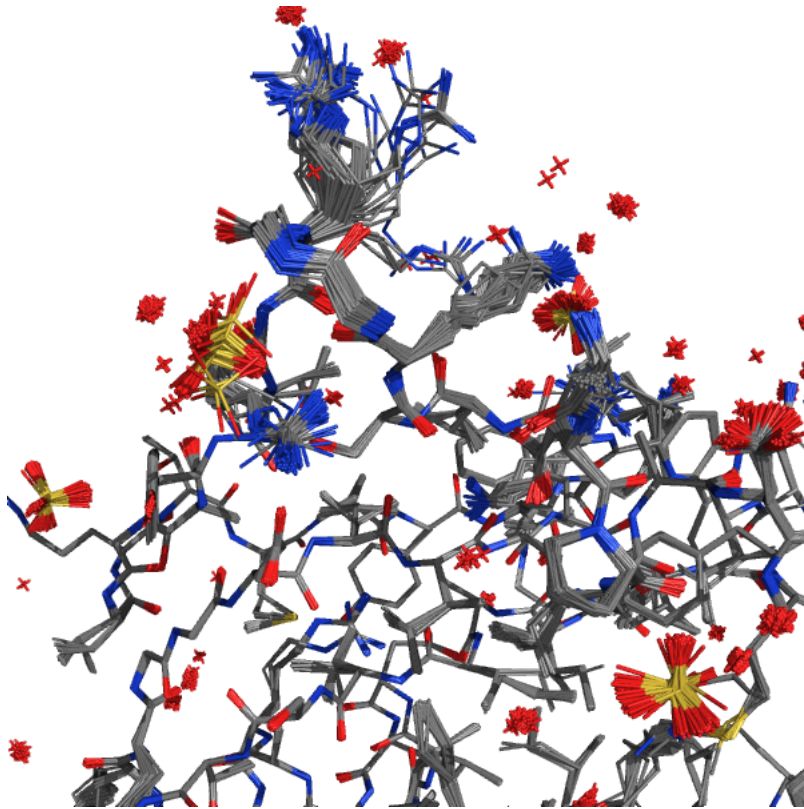


Picture: Dale Tronrud

- Refinement programs have very small convergence radii compared to the size of the function profile
 - Refinement result highly depends on starting point

Result of many refinements with slightly different starting conditions

- Ensemble of slightly different structures having small deviations in atomic positions, B-factors, etc... R-factors deviate too.



Refinement convergence

- **Interpretation of the ensemble:**
 - The variation of the structures in the ensemble reflects:
 - Refinement artifacts (limited convergence radius and speed)
 - Some structural variations
 - Spread between the refined structures is the function of resolution (lower the resolution – higher the spread), and the differences between initial structures
 - Obtaining such ensemble is very useful in order to assess the degree of uncertainty that comes from refinement alone

Data resolution is 3.5Å. How do I enable group B-factor refinement (1 or 2 isotropic B per residue)?

- *phenix.refine* uses a better type of B-factor restraints for individual isotropic B-factor refinement.
 - This allows to refine isotropic individual B-factors at low and very low resolutions
- If there is a reason to suspect there is a problem due to refining individual B-factors at low resolution:
 - Report a problem,
 - Try group B-factor refinement while waiting for a response.

Which restraints *phenix.refine* uses for isotropic B-factor refinement and why they can be used even at very low resolution?

- Isotropic ADP restraints in *phenix.refine* (Afonine *et al*, 2005):
 - Covalent bond is rigid: ADPs of bonded atoms are similar (Hirshfeld, 1976);
 - ADPs of spatially close (non-bonded) atoms are similar (Schneider, 1996);
 - Variation of ADPs of bonded atoms is related to the absolute values of ADPs. Atoms with higher ADPs can have larger differences (Ian Tickle, CCP4 BB, March 14, 2003).

$$T_{\text{adp}} = \sum_{i=1}^{N_{\text{atoms}}} \left[\sum_{j=1}^{M_{\text{atoms}}} \frac{1}{r_{ij}^p} \frac{\left(U_{\text{local},i} - U_{\text{local},j} \right)^2}{\left(U_{\text{local},i} + U_{\text{local},j} \right)^q} \right]$$

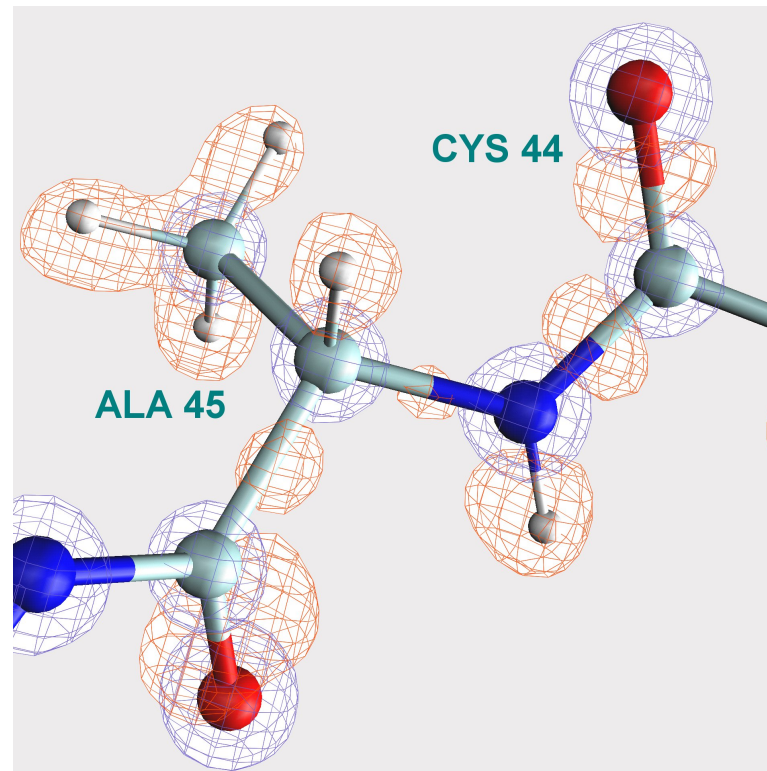
- Since these restraints better model local B-factor variation (compared to traditional restraints) they can be tightened more, which allows using them at lower resolutions

Is SHELX the only program good for high-resolution refinement?

- SHELX is a good program indeed...
 - ...but it is not the only good one for high resolution refinement

Refinement at subatomic resolution

Aldose Reductase (0.66 Å resolution)



Fo-Fc (orange)

2Fo-Fc (blue)

✓ *phenix.refine* has unique set of tools to correctly refine such structures

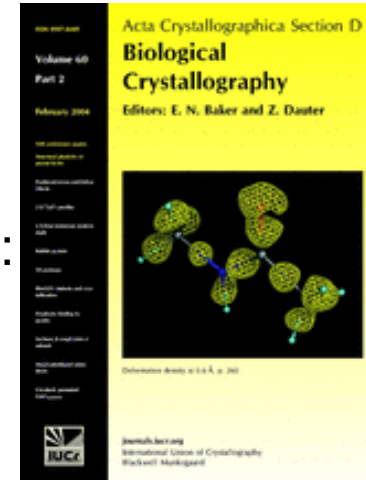
Modeling at subatomic resolution: IAS model

- Basics of IAS model:

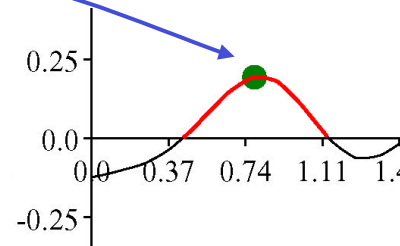
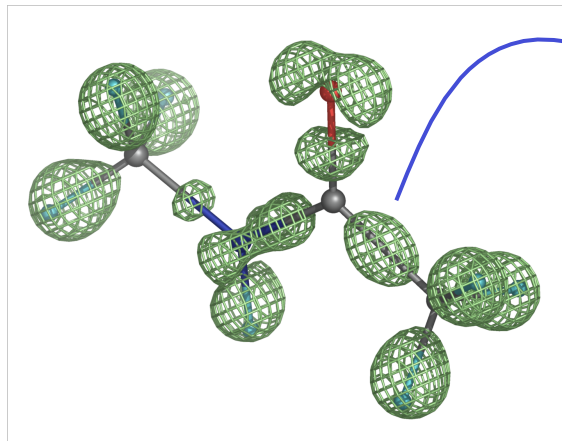
Afonine et al, Acta Cryst. D60 (2004)

- First practical examples of implementation and use in PHENIX:

Afonine et al, Acta Cryst. D63, 1194-1197 (2007)



IAS modeling in PHENIX



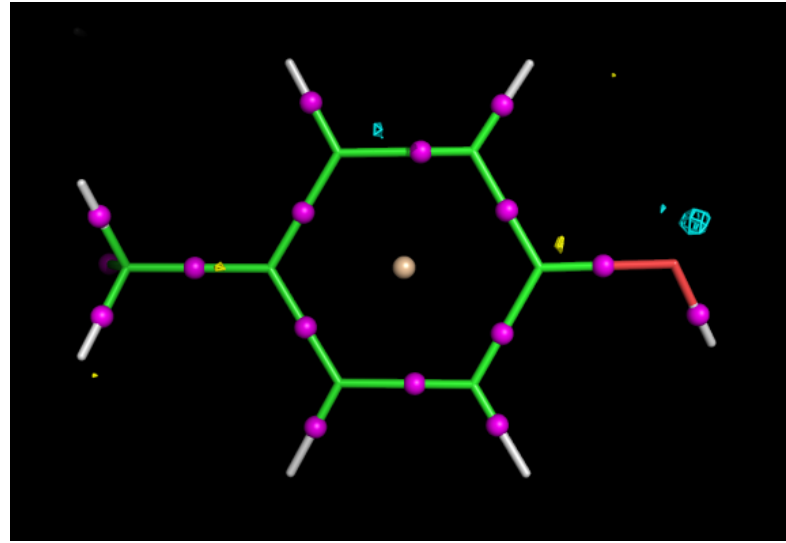
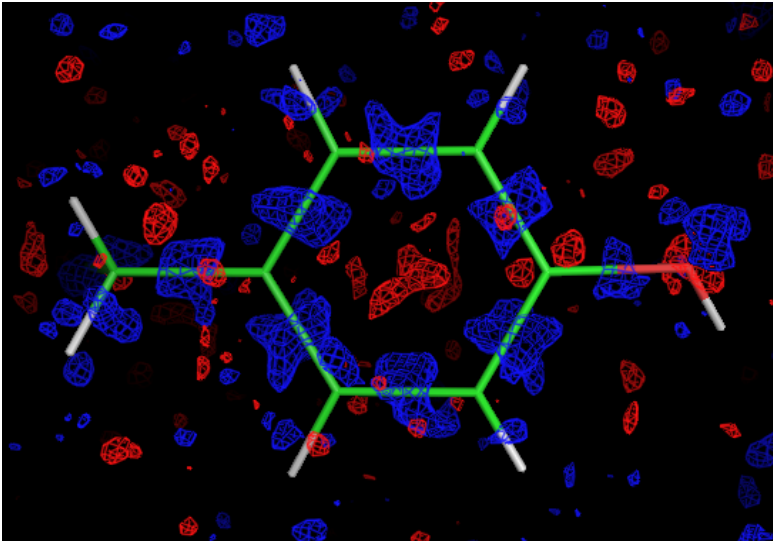
Simple Gaussian is good enough:

$$f_{bond_scatterer}(s) = a \exp(-b s^2)$$

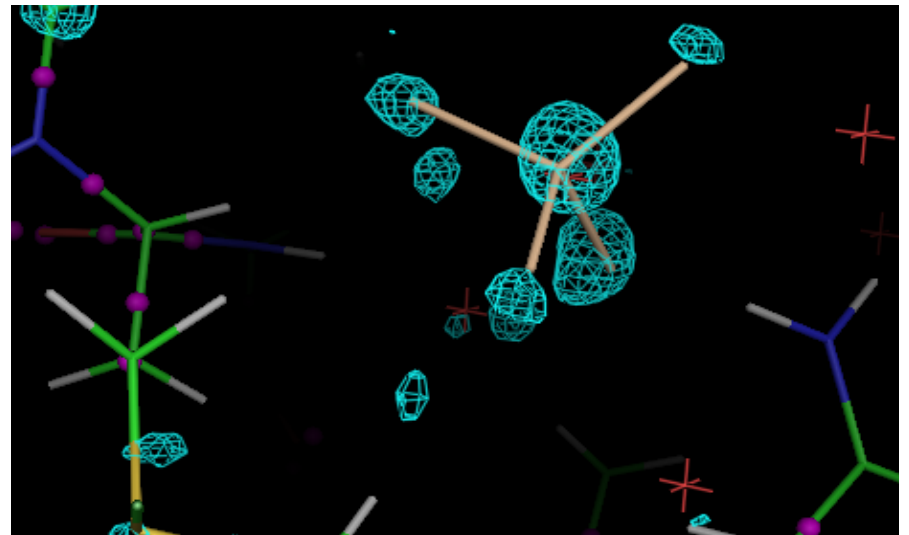
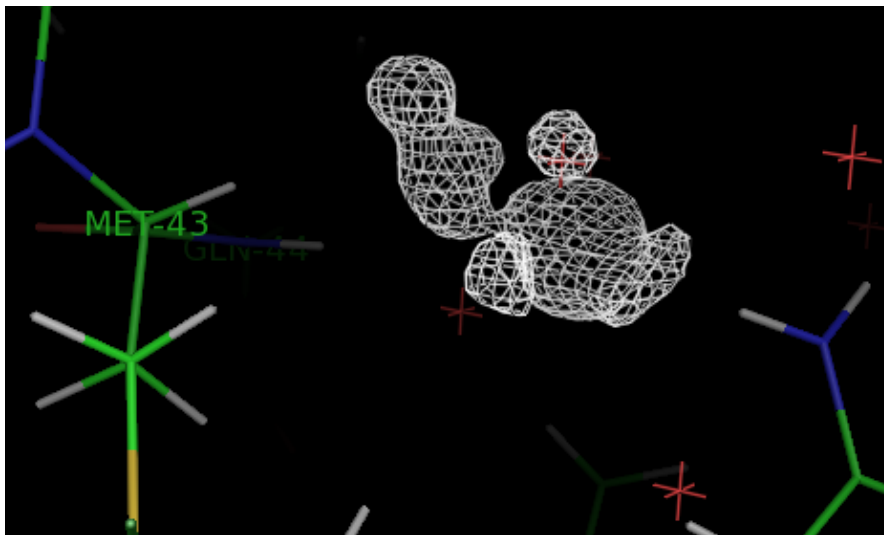
a and **b** are pre-computed library for most bond types

IAS modeling: benefits

- Improve maps: reduce noise. Before (left) and after (right) adding of IAS.



- Find new features: originally wrong water (left) replaced with SO₄ ion (right) clearly suggested by improved map after adding IAS



Is SHELX the only program good for refinement of alternative conformations?

- SHELX is good indeed, but one can do most of occupancy refinements in phenix.refine too

Occupancy refinement

- Automatic constraints for occupancies of atoms in alternate locations
- Any user defined selections for individual and/or group occupancy refinement can be added on top of the automatic selection.

| | | | | | | | | | | | |
|------|---|----|------|---|-----|---------|--------|--------|------|-------|---|
| ATOM | 1 | N | AARG | A | 192 | -5.782 | 17.932 | 11.414 | 0.72 | 8.38 | N |
| ATOM | 2 | CA | AARG | A | 192 | -6.979 | 17.425 | 10.929 | 0.72 | 10.12 | C |
| ATOM | 3 | C | AARG | A | 192 | -6.762 | 16.088 | 10.271 | 0.72 | 7.90 | C |
| ATOM | 7 | N | BARG | A | 192 | -11.719 | 17.007 | 9.061 | 0.28 | 9.89 | N |
| ATOM | 8 | CA | BARG | A | 192 | -10.495 | 17.679 | 9.569 | 0.28 | 11.66 | C |
| ATOM | 9 | C | BARG | A | 192 | -9.259 | 17.590 | 8.718 | 0.28 | 12.76 | C |

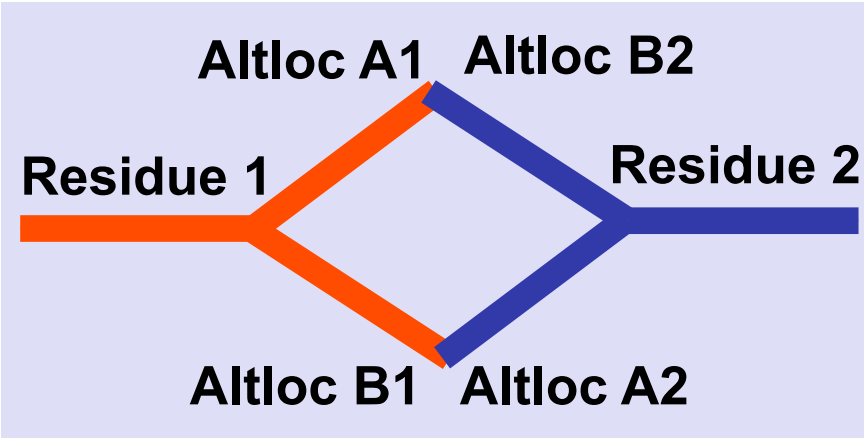
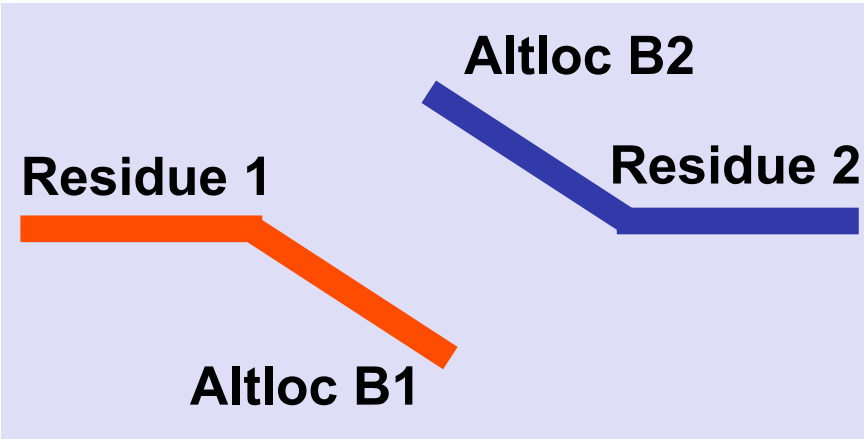
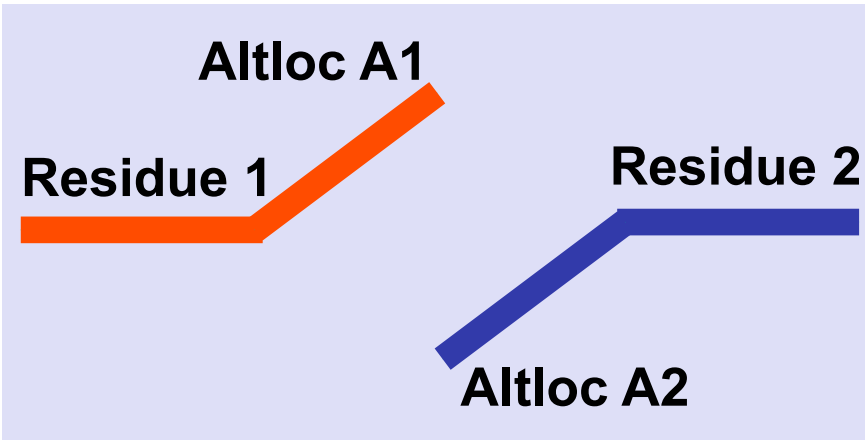
| | | | | | | | | | | | |
|------|-----|----|--|---|----|---------|-------|---------|------|-------|----|
| ATOM | 549 | AU | | A | 34 | -23.064 | 7.146 | -23.942 | 0.78 | 15.44 | Au |
|------|-----|----|--|---|----|---------|-------|---------|------|-------|----|

| | | | | | | | | | | | |
|------|-----|-----|------|---|----|---------|-------|---------|------|-------|---|
| ATOM | 549 | HA3 | ARG | A | 34 | -23.064 | 7.146 | -23.942 | 1.00 | 15.44 | H |
| ATOM | 550 | H | AARG | A | 34 | -24.447 | 7.644 | -21.715 | 0.15 | 8.34 | H |
| ATOM | 551 | D | BARG | A | 34 | -24.447 | 7.644 | -21.715 | 0.85 | 7.65 | D |
| ATOM | 552 | N | ARG | A | 35 | -22.459 | 9.801 | -22.791 | 1.00 | 8.54 | N |

| | | | | | | | | | | | |
|------|----|----|-----|--|---|-------|-------|-------|------|-------|--|
| ATOM | 6 | S | SO4 | | 1 | 1.302 | 1.419 | 1.560 | 0.70 | 13.00 | |
| ATOM | 7 | O1 | SO4 | | 1 | 1.497 | 1.295 | 0.118 | 0.70 | 11.00 | |
| ATOM | 8 | O2 | SO4 | | 1 | 1.098 | 0.095 | 2.140 | 0.70 | 10.00 | |
| ATOM | 9 | O3 | SO4 | | 1 | 2.481 | 2.037 | 2.159 | 0.70 | 14.00 | |
| ATOM | 10 | O4 | SO4 | | 1 | 0.131 | 2.251 | 1.823 | 0.70 | 12.00 | |

| | | | | | | | | | | | |
|------|------|----|------|---|---|--------|--------|--------|------|-------|---|
| ATOM | 3690 | O2 | AEDO | C | 1 | 23.106 | -3.999 | -8.239 | 0.58 | 15.69 | O |
| ATOM | 3691 | C2 | AEDO | C | 1 | 21.710 | -4.102 | -8.630 | 0.58 | 15.43 | C |
| ATOM | 3692 | C1 | AEDO | C | 1 | 20.965 | -2.841 | -8.282 | 0.58 | 16.78 | C |
| ATOM | 3693 | O1 | AEDO | C | 1 | 21.111 | -2.587 | -6.901 | 0.58 | 19.33 | O |
| ATOM | 3687 | I | BIOD | C | 1 | 21.798 | -3.596 | -7.915 | 0.42 | 34.88 | I |

- The only one scenario that is possible to do in SHELX and not yet in phenix.refine:



- Constrained system:
 - $A1+B1=1$
 - $A2+B2=1$
 - $B1+A2=1$

Why don't phenix.refine uses 10 (or 20) resolution bins to show statistics such as R-factors, completeness etc.?

| Bin no. | 3gk8 | | | | | |
|---------|------------|-------|------|------------|-------|------|
| | d^{-3} | | | $\ln(d)$ | | |
| 1 | 22.18–5.00 | 0.906 | 1938 | 22.18–8.16 | 0.610 | 300 |
| 2 | 5.00–3.98 | 0.994 | 2052 | 8.15–7.00 | 0.993 | 300 |
| 3 | 3.98–3.48 | 0.997 | 2060 | 7.00–6.01 | 0.996 | 452 |
| 4 | 3.48–3.16 | 0.995 | 2051 | 6.01–5.16 | 0.994 | 700 |
| 5 | 3.16–2.93 | 0.976 | 1988 | 5.16–4.43 | 0.993 | 1087 |
| 6 | 2.93–2.76 | 0.968 | 1973 | 4.43–3.81 | 0.996 | 1735 |
| 7 | 2.76–2.62 | 0.958 | 1902 | 3.81–3.27 | 0.996 | 2716 |
| 8 | 2.62–2.51 | 0.952 | 1961 | 3.27–2.81 | 0.979 | 4149 |
| 9 | 2.51–2.41 | 0.954 | 1941 | 2.81–2.41 | 0.955 | 6410 |
| 10 | 2.41–2.33 | 0.941 | 1876 | 2.41–2.07 | 0.931 | 9748 |
| 11 | 2.33–2.26 | 0.933 | 1897 | 2.07–1.85 | 0.827 | 9681 |
| 12 | 2.26–2.19 | 0.940 | 1881 | | | |
| 13 | 2.19–2.13 | 0.931 | 1876 | | | |
| 14 | 2.13–2.08 | 0.914 | 1838 | | | |
| 15 | 2.08–2.03 | 0.897 | 1834 | | | |
| 16 | 2.03–1.99 | 0.891 | 1766 | | | |
| 17 | 1.99–1.95 | 0.865 | 1765 | | | |
| 18 | 1.95–1.92 | 0.825 | 1645 | | | |
| 19 | 1.91–1.88 | 0.767 | 1537 | | | |
| 20 | 1.88–1.85 | 0.732 | 1497 | | | |

Bulk-solvent and overall scaling revisited: faster calculations, improved results

P. V. Afonine, R. W. Grosse-Kunstleve, Adams & A. Urzhumtsev; Acta Cryst. (2013). D69

How do I know if refinement converged?

- Check PDB file header or log file or R vs macrocycle plot in the GUI to see how R-factor changes between cycles:
 - If R does not drop anymore then refinement converged.
 - Sometimes it takes several cycles before R starts dropping
 - If R keeps dropping, do more refinement

**Data resolution is 1.57Å: should I refine individual B-factors
as isotropic or anisotropic?**

- At such a “corner resolution” it isn’t known a priori which parameterization would work best.
- A general suggestion: If in doubt – try! This is the most robust way of finding the solution.
- Try both, isotropic and anisotropic, and see what gives better R-factors.
- “corner resolution” : approximately 1.5 ... 1.7Å

A metal ion keeps drifting out of density peak (or/and its coordination geometry gets distorted after refinement)... Is there a way to keep the ion in place?

- Use coordination restrains that can be created by command
phenix.metal_coordination
Or an equivalent in the GUI
- In the future this will be done automatically

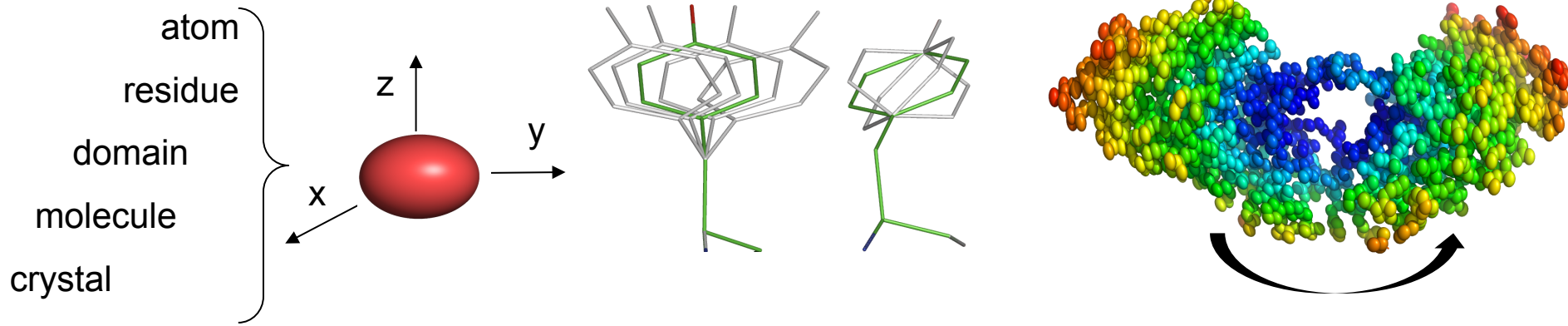
I suspect twinning... Is R-factor drop after introducing twin information into refinement enough of a reason for switching to twin-refinement?

- No. R-factors are not comparable.
 - Garib Murshudov, Applied Computational Mathematics, Vol10, No2, 2011
- Check reflection statistics (use Xtrriage in Phenix GUI for this)
- Use Xtrriage to proper diagnose twinning

I used TLS in refinement. Now I have ANISOU records in PDB file. Why? Did I accidentally do individual anisotropic B-factor refinement?

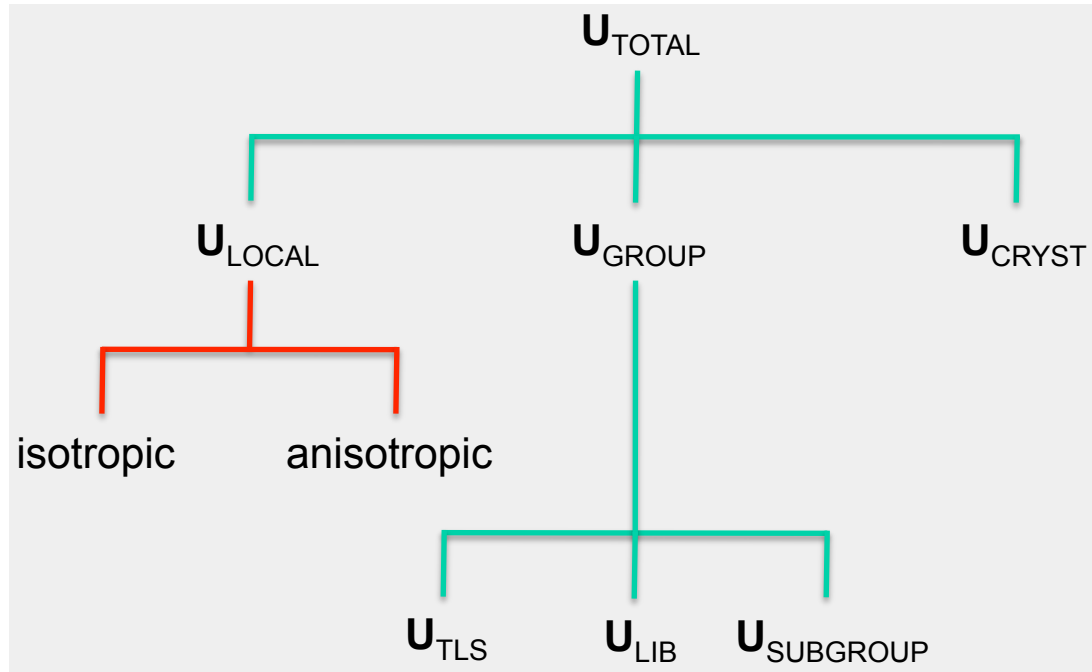
Atomic Displacement Parameters (B-factors)

- Hierarchy and anisotropy of atomic displacements



Total ADP:

$$\mathbf{U}_{\text{TOTAL}} = \mathbf{U}_{\text{CRYST}} + \mathbf{U}_{\text{GROUP}} + \mathbf{U}_{\text{LOCAL}}$$



ADP refinement: what goes into PDB

phenix.refine outputs TOTAL B-factor (iso- and anisotropic):

$$U_{\text{TOTAL}} = U_{\text{ATOM}} + U_{\text{TLS}} + \dots$$

Isotropic equivalent

| | | | | | | | | | | | |
|--------|---|----|-----|---|--------|--------|--------|------|-------|------|---|
| ATOM | 1 | CA | ALA | 1 | 37.211 | 30.126 | 28.127 | 1.00 | 26.82 | C | |
| ANISOU | 1 | CA | ALA | 1 | 3397 | 3397 | 3397 | 2634 | 2634 | 2634 | C |

$$U_{\text{TOTAL}} = U_{\text{ATOM}} + U_{\text{TLS}} + \dots$$

Are there any data resolution restrictions for using TLS in refinement?

- Conceptually – no:
 - Atomic motions in the crystal are not aware of diffraction experiment specifics or limitations
- Practically, yes: there is a (*technical*) high resolution limit in phenix.refine:
 - If data resolution is high enough to use individual anisotropic B-factors, then TLS cannot be used.
 - Individual anisotropic B-factor refinement cannot be combined with TLS

Is Ethan Merritt's TLSMD server the only way to define TLS groups?

TLS groups for refinement automatically

The screenshot shows the 'Configure' dialog box in phenix.refine. The 'Input data' tab is active, showing 'Refinement settings' and 'Output' sub-tabs. Under 'Input files', the 'File path' is set to '/Users/afonine/PHENIX-dev-625/phenix_001.pdb', the 'Format' is 'PDB', and the 'Data type' is 'Input model'. The 'Space group' is set to 'P 62 2 2'. Below this, there are sections for 'X-ray data and experimental phases' and 'Neutron data', each with 'Data labels' and 'High resolution' fields.

The 'TLS group selections' dialog box is open, showing a list of 'TLS refinement groups'. The text explains that phenix.find_tls_groups can be used to identify suitable TLS groups. The list contains three entries, with the first one selected:

- chain 'A' and (resseq 295:394)
- chain 'A' and (resseq 395:434)
- chain 'A' and (resseq 435:569)

Buttons for '+ Add', '- Delete', 'Update item', and 'View/pick' are visible at the bottom. The 'Edit selected' field shows 'chain 'A' and (resseq 295:394)'. A URL is provided: (* http://skuld.bmsc.washington.edu/~tlsmd). The 'Number of processors for phenix.find_tls_groups' is set to 2.

The main window of phenix.refine displays a 3D molecular model of a protein structure. The model is rendered in a stick representation with different colors for different parts of the structure. The interface includes a toolbar with icons for 'Close window', 'Models', 'Maps', 'Settings', 'Help', 'Select atoms', 'Zoom selection', 'Clear selection', and 'Revert'. At the bottom, the 'Object' is 'phenix_001.pdb' and the 'Selection' is 'chain 'A' and (resseq 295:394)'. The 'Mouse' action is 'Rotate view'. A status bar at the bottom right indicates '827 atoms selected'.

Are Torsion angle NCS restraints are always better than Cartesian NCS?

- Not always. Cartesian NCS may be better:
 - In case of many NCS copies
 - Very low resolution
 - Refinement in lower than actual symmetry

I see some density that looks very much like a molecule but I cannot identify it.. Should I use dummy atoms UNK, UNL, UNX etc to model it?

- No. Records in PDB files like these are useless:

| | | | | | | | | | | |
|------|----|-----|-----|---|-------|---------|--------|------|------|---|
| ATOM | 10 | O | UNK | 2 | 6.348 | -11.323 | 10.667 | 1.00 | 8.06 | X |
| ATOM | 11 | O | UNK | 2 | 6.994 | -12.600 | 10.740 | 1.00 | 7.16 | X |
| ATOM | 12 | O | UNK | 2 | 6.028 | -13.737 | 10.607 | 1.00 | 6.58 | X |
| ATOM | 13 | DUM | UNK | 2 | 6.796 | -15.043 | 10.583 | 1.00 | 8.28 | |
| ATOM | 14 | DUM | UNK | 2 | 5.099 | -13.727 | 11.792 | 1.00 | 7.15 | |

When should I add water?

- It's not about when, it's about how.
 - Automatically: you are at the mercy of the program
 - Manually: you are on your own (may be tedious – inefficient)
- Facts:
 - Add water improves overall density – more model can be built
 - The whole idea of ARP/wARP is adding dummy atoms into density peaks with following refinement
 - Interpretation of ligand density with water is not a big problem as long as ligand building tools
 - Interpret water as just density peaks in $2mF_o-DF_c$ map
 - Use water-omit mF_o-DF_c maps
 - Adding wrong water into noise peaks may introduce bias
 - Not adding water till last moment also introduces bias since refined parameters of existing atoms tend to compensate for missing water:
 - Lunin, V.Y., Afonine P.V. & Urzhumtsev, A.G. (2002). "Likelihood-based refinement. I. Irremovable model errors". *Acta Cryst.* A58, 270–282.

What should I deposit into PDB?

- PDB file from your last phenix.refine run
- MTZ file from your last phenix.refine run

- Phenix.refine has an option to output CIF files (both model and data)

I cut data by resolution (sigma) and got better R. This is fantastic! Should I always do it? Should I cut some more data to get even better R?

- Obviously not!
 - Fitting the same amount of parameters against less data is easier than against more data
 - Better fit (lower R) without model improvement
 - Comparing R-factors computed using different sets (amount) of reflections does not make sense

I have anomalous data $F_{obs}(+)$ and $F_{obs}(-)$, but also have F_{mean} (or corresponding I_{mean}). What to use in refinement?

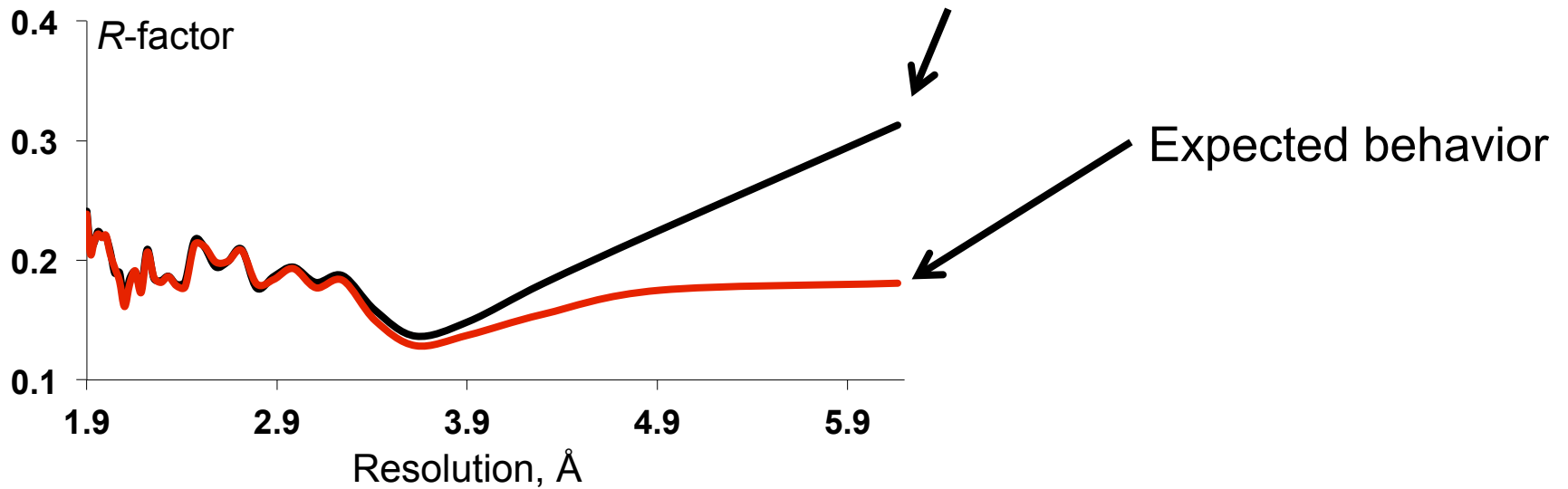
- Refinement against Fobs(+) and Fobs(-) is refinement against less manipulated data compared to $F_{\text{mean}} = (F_{\text{obs}(+)} + F_{\text{obs}(-)})/2$
- Refinement against Fobs(+) and Fobs(-) may be slower (since there are almost twice more data)
- If refine against Fobs(+) and Fobs(-) phenix.refine will create anomalous difference map by default
- If refine against Fobs(+) and Fobs(-) then deposit into PDB Fobs(+) and Fobs(-), and not Fmean!

I see negative density blobs in hydrophobic cores..

- This is a footprint of bulk-solvent mask being set in hydrophobic areas where there is no solvent at all
- This problem is addressed in recent version of Phenix

I see sharp increase of R-factor in lowest resolution bins

Bulk solvent isn't flat or something is wrong with the low resolution data



When should I use Simulated Annealing refinement?

- When model is expected to have gross errors, for example, right after molecular replacement or extensive crude manual rebuilding

When should I use rigid body refinement?

- When model or its pieces are expected to undergo concerted moves in order to fix the data
- For good near to final models SA refinement may do more harm than good

I encountered a problem using phenix.refine: is switching to program X is the only solution?

- **Something didn't work as expected?... program crashed?... missing feature?...**

phenixbb@phenix-online.org

bugs@phenix-online.org

help@phenix-online.org

- **Reporting a problem / bug:**

Send at least:

- 1) PHENIX version;
- 2) Command and parameters I used;
- 3) Input and output files (at least logs).

Best:

Send all input files and command that resulted in problem

Subscribe to PHENIX bulletin board: www.phenix-online.org

How I should **NOT** report a problem?

Real life example

From **CCP4** mailing list:

On 3/5/13 10:11 PM, XXX wrote:

Hi, Everybody,

For refinement of ion in my structure, I used phenix.metal_coordination to produce a geometry restraints file elbow.edits. But after including the file in the phenix refinement, the refined pdb has clash between metal and one of the coordinated atoms (...). Is this a bug in phenix or I did something wrong?

Thanks!

Real life example

From **CCP4** mailing list:

On 3/5/13 10:11 PM, XXX wrote:

Hi, Everybody,

For refinement of ion in my structure, I used phenix.metal_coordination to produce a geometry restraints file elbow.edits. But after including the file in the phenix refinement, the refined pdb has clash between metal and one of the coordinated atoms (...). Is this a bug in phenix or I did something wrong?

Thanks!

On 3/5/13 11:21 PM, Pavel Afonine wrote:

Hi XXX,

if you send me the inputs (data, model and any parameter files) I will tell you what's wrong. If you choose to send the files, please do so to my email address (not the whole mailing list).

FYI: there is Phenix mailing list for questions like this.

Pavel

Real life example

From **CCP4** mailing list:

On 3/5/13 10:11 PM, XXX wrote:

Hi, Everybody,

For refinement of ion in my structure, I used phenix.metal_coordination to produce a geometry restraints file elbow.edits. But after including the file in the phenix refinement, the refined pdb has clash between metal and one of the coordinated atoms (...). Is this a bug in phenix or I did something wrong?

Thanks!

On 3/5/13 11:21 PM, Pavel Afonine wrote:

Hi XXX,

if you send me the inputs (data, model and any parameter files) I will tell you what's wrong. If you choose to send the files, please do so to my email address (not the whole mailing list).

FYI: there is Phenix mailing list for questions like this.

Pavel

On 3/5/13 11:23 PM, XXX wrote:

Can you just tell me how to solve the problem?

Thanks

Real life example

On 3/5/13 11:36 PM, Pavel Afonine wrote:

No, sorry. I cannot suggest a fix since I do not know what the problem is.

Debugging involves 1) reproducing the problem, 2) finding what causes it and 3) fixing it or suggesting the user a work-around. Step 1 requires to have the data, model and other parameters – that's why I asked to send me the files.

Pavel

Real life example

On 3/5/13 11:36 PM, Pavel Afonine wrote:

No, sorry. How can I suggest a fix if I do not know what the problem is?

Debugging involves 1) reproducing the problem, 2) finding what causes it and 3) fixing it or suggesting the user a work-around. Step 1 requires to have the data, model and other parameters – that's why I asked to send me the files.

Pavel

On 3/6/13 12:07 AM, XXX wrote:

its a bug of phenix! and i will fix the problem myself. Thanks

Real life example

On 3/5/13 11:36 PM, Pavel Afonine wrote:

No, sorry. How can I suggest a fix if I do not know what the problem is?

Debugging involves 1) reproducing the problem, 2) finding what causes it and 3) fixing it or suggesting the user a work-around. Step 1 requires to have the data, model and other parameters – that's why I asked to send me the files.

Pavel

On 3/6/13 12:07 AM, XXX wrote:

its a bug of phenix! and i will fix the problem myself. Thanks

On 3/6/13 6:11 AM, Pavel Afonine wrote:

Hi XXX,

Great! I'm glad you can do it, and good luck!

If it is Phenix bug though, then it would be helpful if we fix it on our end so no one else runs into the same problem again.

Pavel

Where to find more information?

Email me your questions: **PAfonine@lbl.gov**

Or send it to Phenix mailing list:

phenixbb@phenix-online.org

Phenix
<http://www.phenix-online.org/> Google

Apple Yahoo! Google Maps YouTube Wikipedia News (1731) Popular

Phenix NEW [Development release of PHENIX version 1.4 now available](#)
Python-based Hierarchical ENvironment for Integrated Xtallography

PHENIX is a new software suite for the automated determination of macromolecular structures using X-ray crystallography and other methods.

Citing PHENIX:
 PHENIX: building new software for automated crystallographic structure determination P.D. Adams, R.W. Grosse-Kunstleve, L.-W. Hung, T.R. Ioerger, A.J. McCoy, N.W. Moriarty, R.J. Read, J.C. Sacchettini, N.K. Sauter and T.C. Terwilliger. *Acta Cryst.* D58, 1948-1954 (2002)

Download the latest development release (1.4-3) [First request download password]

Help: [FAQ](#) [Mailing List Subscription](#) [List Archives](#) [Report a Bug](#) [Email for Help](#)

Using PHENIX (release 1.4-3): [Full Documentation](#) [PDF](#)

- Assessing data quality with [phenix.xtriage](#)
- Automated structure solution with [AutoSol](#)
- Automated molecular replacement with [AutoMR](#)
- Automated model building and rebuilding with [AutoBuild](#)
- Automated ligand fitting with [LigandFit](#)
- Structure refinement with [phenix.refine](#)
- Generation of ligand coordinates and restraints with [elbow](#)
- The [PHENIX Graphical User Interface](#)

[Documentation for 1.3-final](#)

The PHENIX system also includes SOLVE/RESOLVE, Phaser, Textal, the CCI Applications (phenix.xtriage, phenix.refine, elbow and many more), components from Molprobit, and the Computational Crystallography Toolbox in a Python framework.




Funding for PHENIX: [Protein Structure Initiative \(NIH General Medical Sciences\)](#)

The PHENIX Industrial Consortium [Information](#)

For-profit groups can obtain access to PHENIX through a Consortium agreement. This provides a license to use PHENIX and research funds to develop new features in PHENIX tailored to the needs of commercial users. [Members](#)
[Download](#)
[Contact Us](#)

Groups developing PHENIX:

[Paul Adams](#) [Randy Read](#) [Jane & Dave Richardson](#) [Tom Terwilliger](#) [Tom Ioerger & Jim Sacchettini](#)

[Privacy and Security Notice](#) [About this website](#)

This presentation (PDF file) and much more