Contents

[Non Crystallographic Symmetry (NCS) refinement 1](#_Toc384739315)

[Search terms for implementation in Phenix: 2](#_Toc384739316)

[Read papers: 2](#_Toc384739317)

[Tests 2](#_Toc384739318)

[Work notes 2](#_Toc384739319)

[Rotation convention 3](#_Toc384739320)

[Cost function gradient 3](#_Toc384739321)

[Choosing origin 6](#_Toc384739322)

[Choosing coordinate system 6](#_Toc384739323)

[Parameters 6](#_Toc384739324)

[References 8](#_Toc384739325)

# Non Crystallographic Symmetry (NCS) refinement

Refinement of the rotation and translation operation used to produce the complete Asymmetric Unit (ASU) from a single NCS copy.

* Consider using least-squares, maximum likelihood and log likelihood target functions
* Check the linear optimization methods

Need to get a cost function and gradient (in relation to the refined parameters.

* Hypothesis:
* Parameters: Rotation and translation matrices.   
  (Need to maintain rotation characteristics during refinement and when shaking parameters for testing)
* Cost Function:
  + For least-squares (LS) use (Lunin & Skovoroda 1995):



Where



* For Maximum Likelihood estimation I need to work out the likelihood function, the probability of the transformation, given the coordinates. Get the probability from summing over all observation… need to create energy function.   
  [Refinement of X-ray Crystal Structures](file:///C:\Users\Youval\Downloads\brunger_adams_2012.pdf)  
  (Pannu & Read 1996)



* Consider refinement strategy if possible within the PHENIX machinery
* For high resolution data, consider ways to musk the part of the NCS copy that are not the same.
* Test refinement on PDB files that include all NCS copies, not only the 150 that do not.
* Check if there is some symmetry that specify what the transformation should be. For example, viruses biomolecule many times have a define symmetry, does this symmetry translate to analytical transformations for the ASU?

## Search terms for implementation in Phenix:

Target auto switch resolution

# Read papers:

(Afonine et al 2009), (Afonine et al 2005), (Lunin & Skovoroda 1995)

(Afonine & Urzhumtsev 2004)

[Note1](rigid_body_Euler.doc), [Note2](file:///C:\Phenix\Dev\WORK\WORK\NCS\REFINE_NCS_OPERATORS\rigid_body_Euler2.doc), [Refinement of X-ray Crystal Structures](../../../../../../Users/Youval/Downloads/brunger_adams_2012.pdf).

# Tests

* Model shape, number of NCS copies
* Shaking of R and T

# Work notes

Our “origin”, the center of rotation is not necessarily the center of mass or center of coordinates of the given NCS, perhaps it is so for the complete ASU. **It is important to know what is the reference coordinate system, since different rotation and translation will results from a different choice of system.**

I assume that the single NCS copy coordinates are the coordinates for which the Rotation and Translation operation are given to.

Let the coordinates of a single atom , in the initial single NCS copy, in relation to the origin be:   
And for the NCS copy in a way that Where

And are the rotation matrix and translation vector.



## Rotation convention



Try to avoid the angle, since this cause the rotation matrix to be a function of and we will not be able to tell differences between changes in and.

To avoid those issue, consider changing the coordinate system for the refinement and then convert to the standard expected in the PDB. Consider what would an optimal representation for the PDB.

## Cost function gradient

Let be the cost function

To minimize we need it’s derivatives in respect to the parameters we want to refine, and.

We do the derivation in relation to the angle and not to the component of to reduce the number if independent parameters and to keep the rotation matrix properties.

The gradient of in relation of the transformation, for each atom, is:



For the rotation part:

Let 



Base on equation is:







If we consider that the cost function is a function of then we can calculate





Now rewrite equation in a matrix notation



Let’s rewrite equation:



Using the following notation:

Let be a matrix where all components are 1 and with the dimensions, where is the number of atoms.

Let be a matrix with the dimensions, where is the number of atoms, of the form:



The matrix is a matrix of the form:



Let be



Where  and is the transformation

So now we can write the gradient of the cost function in relation to the angles and the translation as:



**OF COURSE do not use matrix multiplication when doing calculation in the program. This is extremely not efficient. Just directly calculate only the diagonal, the Trace.**

## Choosing origin

The natural and the default choice can be the origin of the coordinate of the single NCS copy.

To allow origin change, write a function that convert the translation vectors according to the rotation matrices, when moving the origin, so that we will have 







## Choosing coordinate system

To avoid the issue caused in the case of , when the rotation derivative becomes a function of and we can’t tell them apart, consider rotating the coordinate system around the axis, so and

## Parameters

Our data from the PDB file isnot

So to get from we can do the following:

Calculate the possible angles:



Choose and calculate

If the signs of both the original and the calculated using the new angles are correct then either or are the correct pair

If not then either or are the correct pair

Now choose and calculate

If the signs of both original and calculated are correct then either or are the correct pair

If not then either or are the correct pair.

Now you can just choose angles that work

This will give as two possible solutions: or , which mean that the rotation matrix is not unique in respect to the rotation angles. It does not matter to our purpose, as long as the angles produce the desire rotation matrix.

There is a unique cases of, in those cases there is not ambiguity in,, and. In those cases the rotation matrix will become:





So there is not unique solution and we can just choose and that obey equation.

A simple choice can be



**Maybe** we should consider a rotation of the coordinate system as described in choosing coordinate system

# Target and Gradient for transformation refinement

Input data (CIF file) and model (PDB)

Update model for Bulk-solvent and overall scaling

Get weights

Get Miller arrays

Average Bijvoet mates

Add function that give rotation angles from rotation matrices

Select strategy  
FFT or Direct  
ml or ls\_wunit\_k1   
(Max Likelihood or Least Squares)  
Use hydrogen or not

Get F-obs and F-free

Get restraints manager

Get initial F-model

# References

Afonine PV, Grosse-Kunstleve RW, Adams PD. 2005. A robust bulk-solvent correction and anisotropic scaling procedure. *Acta crystallographica. Section D, Biological crystallography* 61: 850-5

Afonine PV, Grosse-Kunstleve RW, Urzhumtsev A, Adams PD. 2009. Automatic multiple-zone rigid-body refinement with a large convergence radius. *J Appl Crystallogr* 42: 607-15

Afonine PV, Urzhumtsev A. 2004. On a fast calculation of structure factors at a subatomic resolution. *Acta Crystallogr A* 60: 19-32

Lunin VY, Skovoroda TP. 1995. R-Free Likelihood-Based Estimates of Errors for Phases Calculated from Atomic Models. *Acta Crystallogr A* 51: 880-87

Pannu NS, Read RJ. 1996. Improved structure refinement through maximum likelihood. *Acta Crystallogr A* 52: 659-68