Location of this talk:
www.ysbl.york.ac.uk/refmac/Presentations/refmac_Osaka.ppt

Refinement of Macromolecular structures using REFMAC5

Garib N Murshudov
York Structural Laboratory
Chemistry Department
University of York

Contents

1) Introduction
2) Considerations for refinement
3) Overall parameters
4) TLS
5) Dictionary and alternative conformations
6) Twin
7) Some new features in REFMAC
8) Conclusions

Available refinement programs

- SHELXL
- CNS
- REFMAC5
- TNT
- BUSTER/TNT
- Phenix.refine
- RESTRAINT
- MOPRO

What can REFMAC do?

- Simple maximum likelihood restrained refinement
- Twin refinement
- Phased refinement (with Hendrickson-Lattmann coefficients)
- SAD/SIRAS refinement
- Structure idealisation
- Library for more than 8000 ligands (from the next version)
- Covalent links between ligands and ligand-protein
- Rigid body refinement
- NCS local, restraints to external structures, jelly body
- TLS refinement
- Map sharpening
- etc

Considerations in refinement

- Function to optimise (link between data and model)
  - Should use experimental data
  - Should be able to handle chemical (e.g bonds) and other (e.g. NCS, structural) information
- Parameters
  - Depends on the stage of analysis
  - Depends on amount and quality of the experimental data
- Methods to optimise
  - Depends on stage of analysis: simulated annealing, conjugate gradient, second order (normal matrix, information matrix, second derivatives)
  - Some methods can give error estimate as a by-product. E.g second order.

Two components of target function

Crystallographic target functions have two components: one of them describes the fit of the model parameters into the experimental data and the second describes chemical integrity (restraints).

Currently used restraints are: bond lengths, angles, chirals, planes, ncs if available, some torsion angles, reference structures

Crystallographic refinement

The function in crystallographic refinement has a form:

$$L(p)=wL_X(p)+L_G(p)$$

Where $L_X(p)$ is -loglikelihood and $L_G(p)$ is -log of prior probability distribution - restraints.

It is one of many possible formulations. It uses Bayesian statistics. Other formulation is also possible. For example: stat physical energy, constrained optimisaton, inverse problem etc.

Various forms of X-ray component

- SAD functions uses observed F+ and F- directly without any preprocessing by a phasing program (It is not available in the current version but will be available soon)
- SIRAS – uses native and derivative anomalous data directly
- MLHL - explicit use of phases with Hendrickson Lattman coefficients
- Rice - Maximum likelihood refinement without phase information

-loglikelihood

-loglikelihood depends on assumptions about the experimental data, crystal contents and parameters. For example with assumptions that all observations are independent (e.g. no twinning), there is no anomolous scatterers and no phase information available, for acentric reflections it becomes:

$$L_X(p) = \sum \frac{|F_o|^2 + |F_c|^2}{\Sigma} - \log(I_0(2|F_o||F_c|/\Sigma)) + \log(\Sigma) + const$$

And for centric reflections:

$$L_X(p) = \sum \frac{|F_o|^2 + |F_c|^2}{2\Sigma} - \log(\cosh(|F_o||F_c|/\Sigma)) + 0.5\log(\Sigma) + const$$

All parameters (scale, other overall and atomic) are inside $|F_c|$ and Σ

Note that these are loglikelihood of multiples of chi-squared distribution with degree of freedom 2 and 1

Twin refinement in REFMAC

Twin refinement in refmac (5.5 or later) is automatic.

- Identify “twin” operators
- Calculate “Rmerge” ($\Sigma|I_h$-$<I>_{twin}|$ /ΣI_h) for each operator. If Rmerge>0.50 keep it: Twin plus crystal symmetry operators should form a group
- Refine twin fractions. Keep only “significant” domains (default threshold is 5%): Twin plus symmetry operators should form a group

Intensities can be used

If phases are available they can be used

Maximum likelihood refinement is used

Likelihood

$$
\begin{aligned}
P(I_o;F) &= \int_F P(I_o,F;F_c)dF &&= \int_F P(I_o;F)P(F;F_c)dF \\
P(I_o;F) &= N_o e^{-\frac{\sum(I_{oj}-\sum\alpha_{ij}|I_{ij}|)^2}{(2\sigma_{oj}^2}} \\
P(F;F_c) &= N_c \prod e^{-\frac{|F-DF_c|^2}{\Sigma}}
\end{aligned}
$$

The dimension of integration is in general twice the number of twin related domains. Since the phases do not contribute to the first part of the integrant the second part becomes Rice distribution.

The integration is carried out using Laplace approximation.

These equations are general enough to account for: non-merohedral twinning (including allawtwin), unmerged data. A little bit modification should allow handling of simultaneous twin and SAD/ MAD phasing, radiation damage

Twin: Few warnings about R factors

For acentric case only:

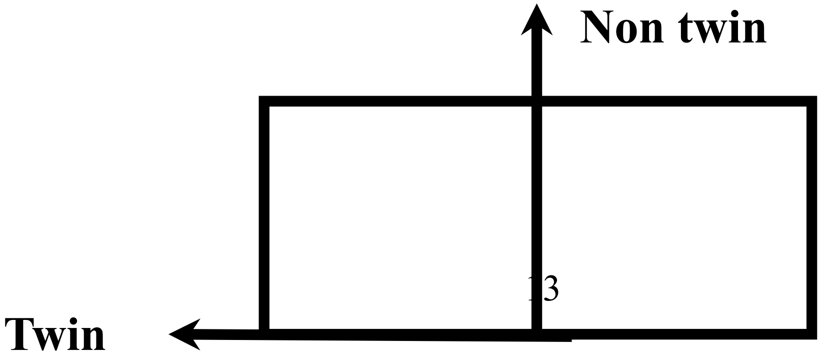
For random structure

Crystallographic R factors

No twinning	58%
For perfect twinning: twin modelled	40%
For perfect twinning without twin modelled	50%

R merges without experimental error

No twinning	50%
Along non twinned axes with another axis than twin	37.5%



Map calculation

- After refinement programs usually give coefficients for two type of maps: 1) 2Fo-Fc type maps. They try to represent the content of the crystal. 2) Fo-Fc type of maps. They try to represent difference between contents of the crystal and current atomic model. Both these maps should be inspected and model should be corrected if necessary.
- Refmac gives coefficients:

  $2\ m\ F_o\ -\ D\ F_c$ – to represent contents of the crystal

  $m\ F_o\ –D\ F_c$ - to represent differences

m is the figure of merit (reliability) of the phase of the current reflection and D is related tomodel error. m depends on each reflection and D depends on resolution

If phase information is available then map coefficients correspond to the combined phases.

Parameters

Usual parameters (if programs allow it)

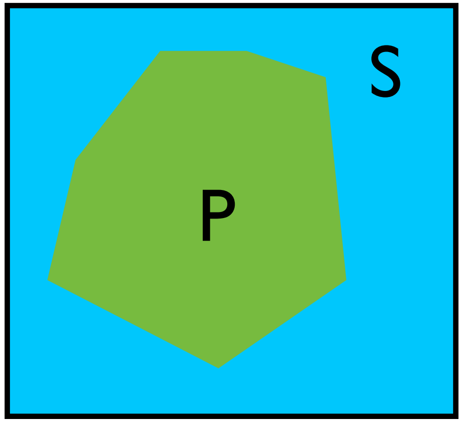
1) Positions x,y,z
2) B values – isotropic or anisotropic
3) Occupancy

Derived parameters

4) Rigid body positional
   - After molecular replacement
   - Isomorphous crystal (liganded, unliganded, different data)
5) Rigid body of B values – TLS
   - Useful at the medium and final stages
   - At low resolution when full anisotropy is impossible
6) Torsion angles

Bulk solvent
Method 1: Babinet's bulk solvent correction

At low resolution electron density is flat. Only difference between solvent and protein regions is that solvent has lower density than protein. If we would increase solvent just enough to make its density equal to that of protein then we would have flat density (constant). Fourier transformation of constant is zero (apart from F000). So contribution from solvent can be calculated using that of protein. And it means that total structure factor can calculated using contribution from protein only

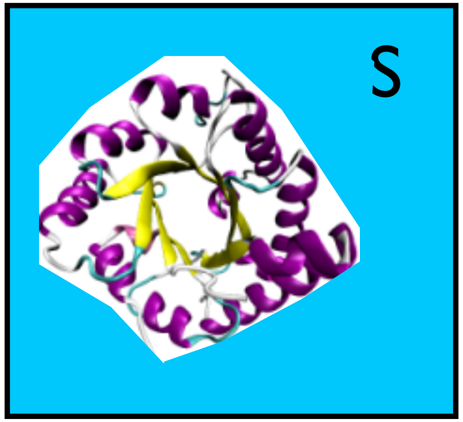


$$\rho_s+\rho_p=\rho_T \quad <==> \quad F_s+F_p=F_T$$
$$\rho_s+k\rho_p=c \quad <==> \quad F_s+kF_p=0$$
$$F_s=-kF_p \quad ==> \quad F_T=F_p-kF_p=(1-k)F_p$$

k is usually taken as $k_b \exp(-B_b s^2)$. k_b must be less than 1. k_b and B_b are adjustable parameters

Bulk solvent
Method 2: Mask based bulk solvent correction

Total structure factor is the sum of protein contribution and solvent contribution. Solvent region is flat. Protein contribution is calculated as usual. The region occupied by protein atoms is masked out. The remaining part of the cell is filled with constant values and corresponding structure factors are calculated. Finally total structure factor is calculated using



$$\boxed{F_T=F_p+k_sF_s}$$

k_s is adjustable parameter.

Mask based bulk solvent is a standard in all refinement programs. In refmac it is default.

Overall parameters: Scaling

There are several options for scaling:

1) Babinet’s bulk solvent assumes that at low resolution solvent and protein contributors are very similar and only difference is overall density and B value. It has the form: $k_b = 1 - k_b\ e(-B_b\ s^2/4)$

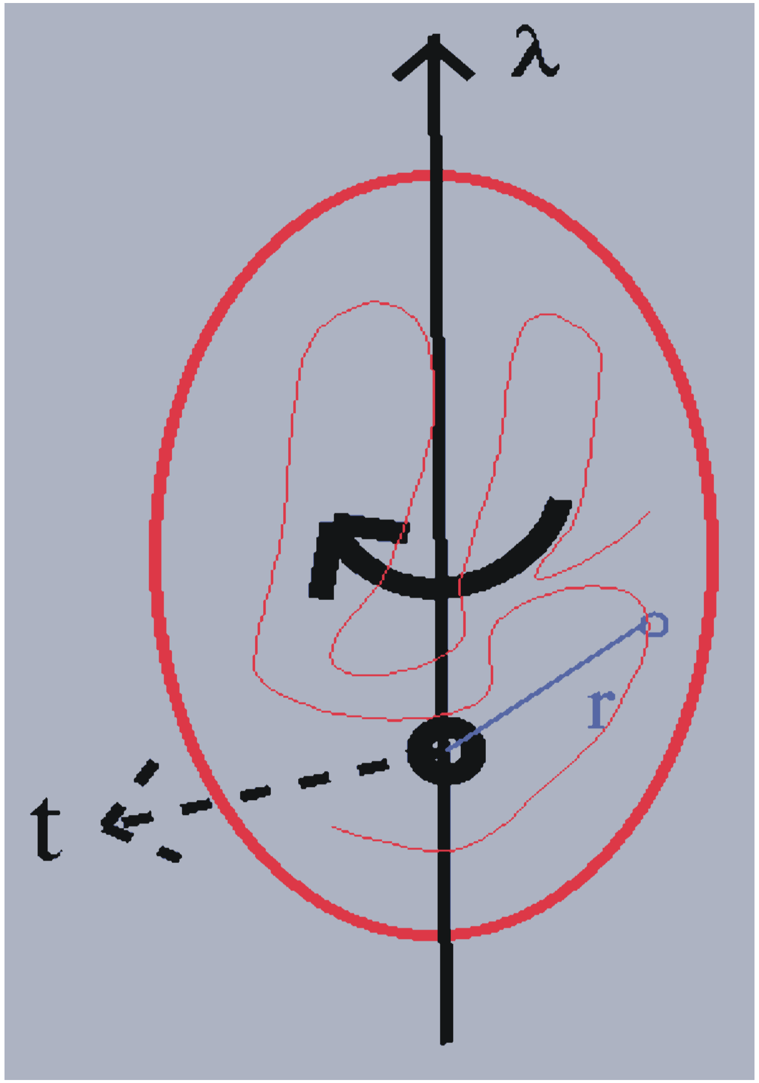
2) Mask bulk solvent: Part of the asymmetric unit not occupied by atoms are asigned constant value and Fourier transformation from this part is calculated. Then this contribution is added with scale value to “protein” structure factors. Total structure factor has a form: $F_{tot} = F_p + s_s \exp(-B_s\ s^2/4) F_s$.

3) The final total structure factor that is scaled has a form:

$$s_{aniso} s_{protein}\ k_b F_{tot}$$

TLS

Rigid-body motion



General displacement of a rigid-body point can be described as a rotation along an axis passing through a fixed point together with a translation of that fixed point.

$$\underline{\mathbf{u}} = \underline{\mathbf{t}} + \mathbf{D}\underline{\mathbf{r}}$$

for small librations

$$\underline{\mathbf{u}} \approx \underline{\mathbf{t}} + \underline{\lambda}\times\underline{\mathbf{r}}$$

D = rotation matrix

λ = vector along the rotation axis of magnitude equal to the angle of rotation

TLS parameters

Dyad product:

$$\underline{uu}^{T} = \underline{tt}^{T} + \underline{t\lambda}^{T}\times\underline{r}^{T} - \underline{r}\times\underline{\lambda t}^{T} - \underline{r}\times\underline{\lambda\lambda}^{T}\times\underline{r}^{T}$$

ADPs are the time and space average

$$U_{TLS} = \langle\underline{uu}^{T}\rangle = T + S^{T}\times\underline{r}^{T} - \underline{r}\times S - \underline{r}\times L\times\underline{r}^{T}$$

$T = \langle\underline{tt}^{T}\rangle$ **6 parameters, TRANSLATION**

$L = \langle\underline{\lambda\lambda}^{T}\rangle$ **6 parameters, LIBRATION**

$S = \langle\underline{\lambda t}^{T}\rangle$ **8 parameters, SCREW-ROTATION**

TLS groups

Rigid groups should be defined as TLS groups. As starting point they could be: subunits or domains.

If you use script then default rigid groups are subunits or segments if defined.

In ccp4i you should define rigid groups (in the next version default will be subunits).

Rigid group could be defined using TLSMD webserver:

http://skuld.bmsc.washington.edu/~tlsmd/

Use of available knowledge

1) NCS local
2) Restraints to known structure(s)
3) Restraints to current inter-atomic distances (implicit normal modes or “jelly” body)
4) Better restraints on B values

These are available from the version 5.6

Note
Buster/TNT has local NCS and restraints to known structures
CNS has restraints to known structures (they call it deformable elastic network)
Phenix has B-value restraints on non-bonded atom pairs and automatic global NCS
Local NCS (only for torsion angle related atom pairs) was available in SHELXL since the beginning of time

NCS: Three approaches

NCS may improve refinement and produce “better” model. There are three different approaches in refinement:

1) Strict NCS

   NCS related molecules are exact copy of each other. You need to give only one independent copy. You need also to give NCS operators.

2) NCS global restraints

   Entire molecules are similar to each other

3) Local NCS

   Molecules are only locally similar

Auto NCS: local and global

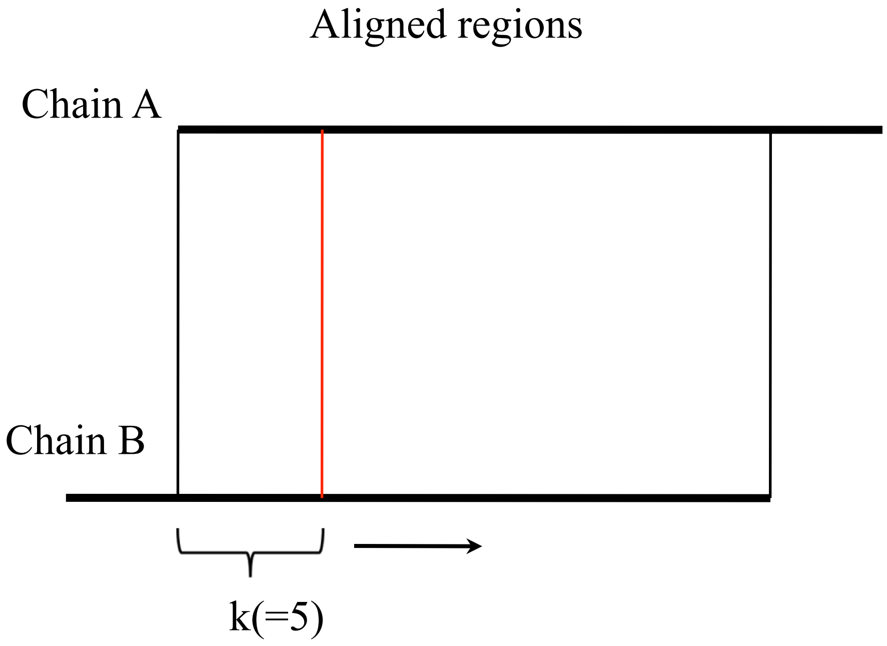
1. Align all chains with all chains using Needleman-Wunsh method
2. If alignment score is higher than predefined (e.g.80%) value then consider them as similar
3.Find local RMS and if average local RMS is less than predefined value then consider them aligned
4. Find correspondence between atoms
5. If global restraints (i.e. restraints based on RMS between atoms of aligned chains) then identify domains
6.For local NCS make the list of corresponding interatomic distances (remove bond and angle related atom pairs)
7.Design weights

The list of interatomic distance pairs is calculated at every cycle

Auto NCS

Global RMS is calculated using all aligned atoms.

Local RMS is calculated using k (default is 5) residue sliding windows and then averaging of the results



$$Ave(RmsLoc)_k = \frac{1}{N-k+1}\sum_{i=1}^{N-k+1} RmsLoc_i$$

$$RMS = Ave(RmsLoc)_N$$

Auto NCS: Iterative alignment

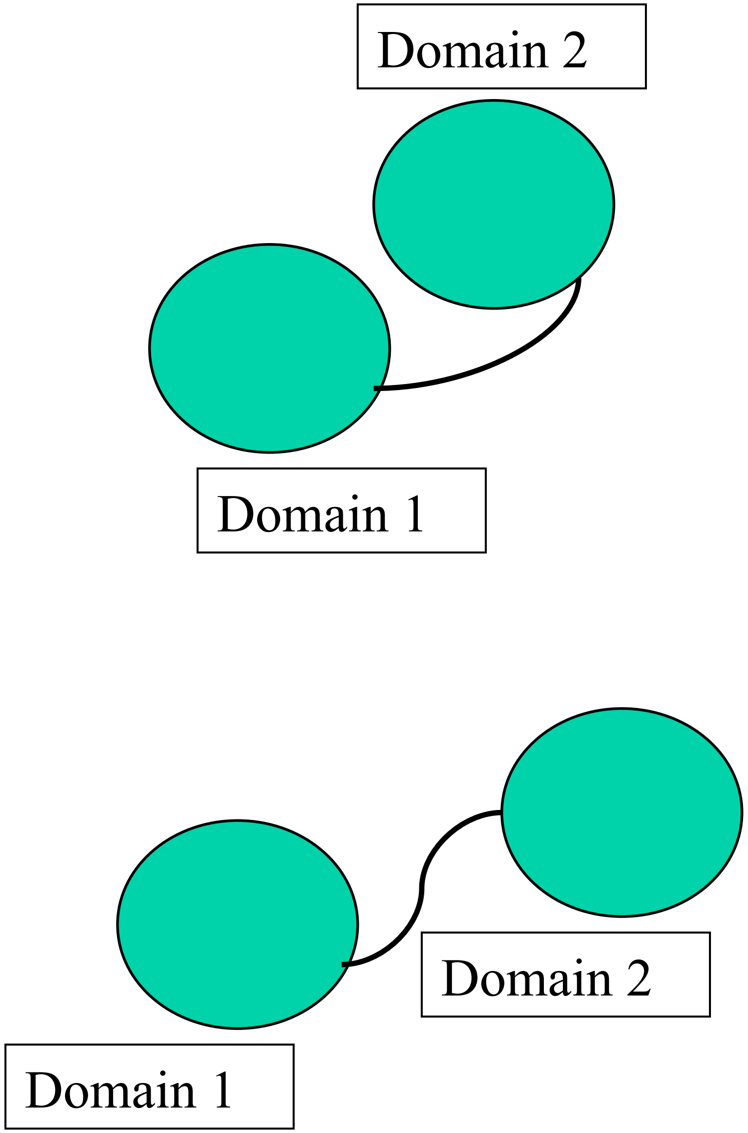
Example of alignment: 2vtu.
There are two chains similar to each other. There appears to be gene duplication

RMS – all aligned atoms
Ave(RmsLoc) – local RMS

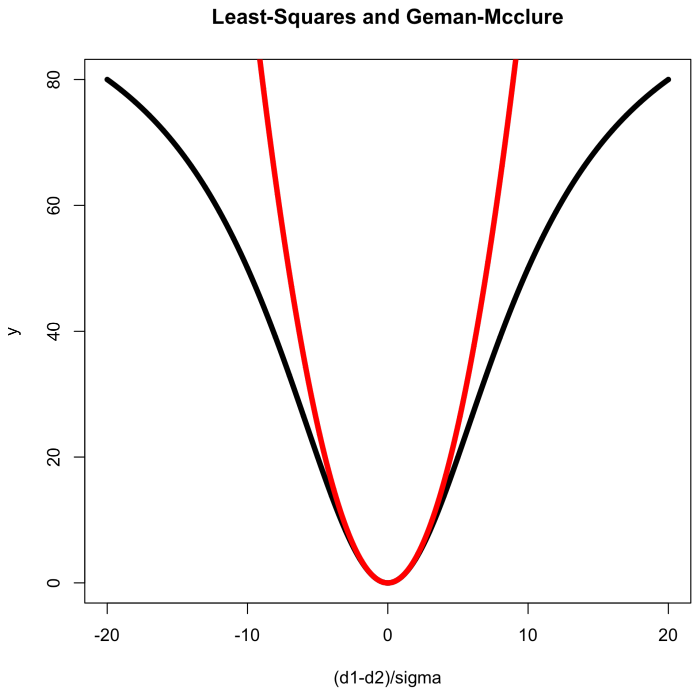
```
                     ********* Alignment results *********
 -------------------------------------------------------------------------------
: N:     Chain 1 :       Chain 2 :  No of aligned :Score :    RMS   :Ave(RmsLoc):
 -------------------------------------------------------------------------------
:  1 : J( 131 - 256 ) : J(   3 - 128 ) :   126 :  1.0000 :   5.2409 :   1.6608 :
:  2 : J(   1 - 257 ) : L(   1 - 257 ) :   257 :  1.0000 :   4.8200 :   1.6694 :
:  3 : J( 131 - 256 ) : L(   3 - 128 ) :   126 :  1.0000 :   5.2092 :   1.6820 :
:  4 : J(   3 - 128 ) : L( 131 - 256 ) :   126 :  1.0000 :   3.0316 :   1.5414 :
:  5 : L( 131 - 256 ) : L(   3 - 128 ) :   126 :  1.0000 :   0.4515 :   0.0464 :
-------------------------------------------------------------------------------
```

Auto NCS: Conformational changes

In many cases it could be expected that two or more copies of the same molecule will have (slightly) different conformation. For example if there is a domain movement then internal structures of domains will be same but between domains distances will be different in two copies of a molecule



Robust estimators



One class of robust (to outliers) estimators are called M-estimators: maximum-likelihood like estimators. One of the popular functions is Geman-Mcclure.

Essentially when distances are similar then they should be kept similar and when they are too different they should be allowed to be different.

This function is used for NCS local restraints as well as for restraints to external structures

Red line: x^2
Black line: x^2/(1+w x^2)

where $x=(d1-d2)/\sigma$, $w=0.1$

Restraints to external structures
It is done by Rob Nicholls

ProSmart

Compares Two Protein Chains

- Conformation-invariant structural comparison
- Residue-residue alignment
- Superimposition
- Residue-based and global similarity scores

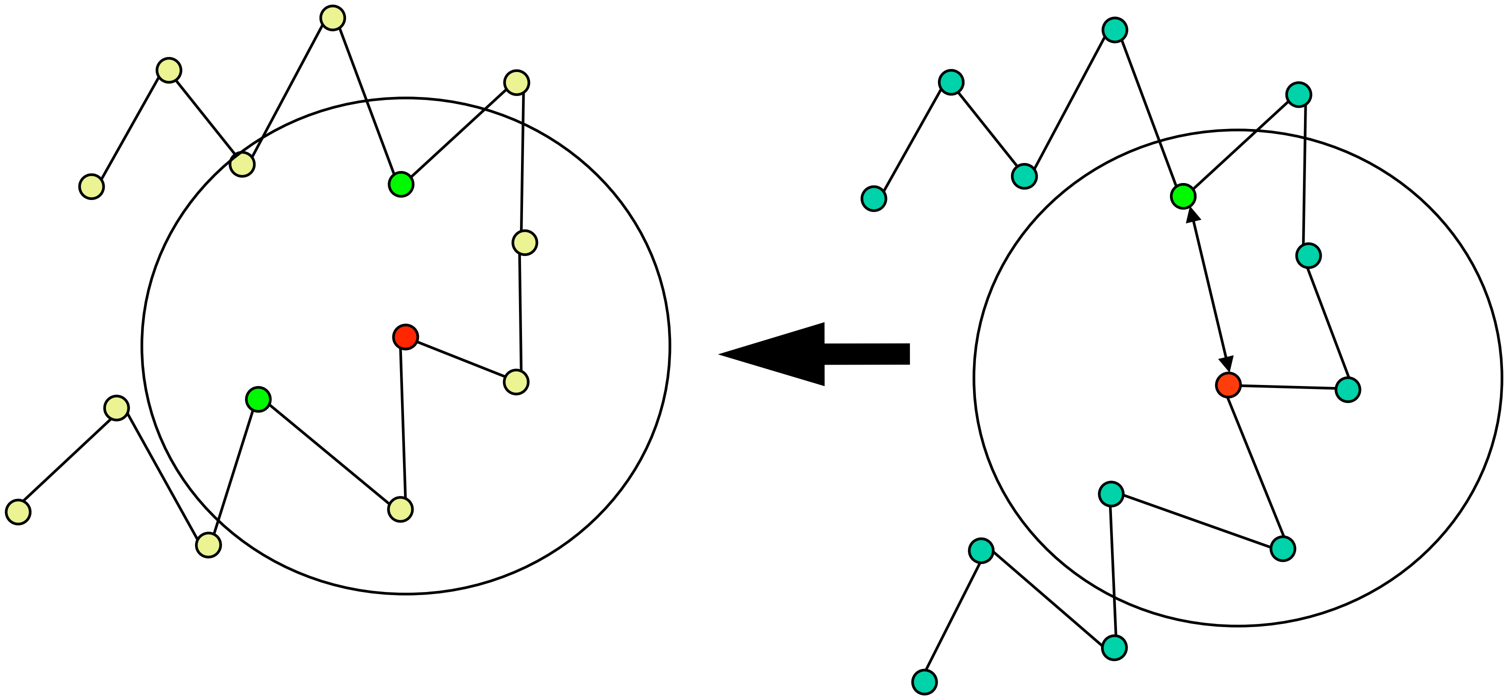
Produces local atomic distance restraints

- Based on one or more aligned chains
- Possibility of multi-crystal refinement

ProSmart Restrain

structure to be refined

known similar structure (prior)



Remove bond and angle related pairs

Restraints to current distances

The term is added to the target function:

$$\sum_{pairs} w(|d| - |d_{current}|)^2$$

Summation is over all pairs in the same chain and within given distance (default 4.2A). $d_{current}$ is recalculated at every cycle. This function does not contribute to gradients. It only contributes to the second derivative matrix.

It is equivalent to adding springs between atom pairs. During refinement inter-atomic distances are not changed very much. If all pairs would be used and weights would be very large then it would be equivalent to rigid body refinement.

It could be called “implicit normal modes”, “soft” body or “jelly” body refinement.

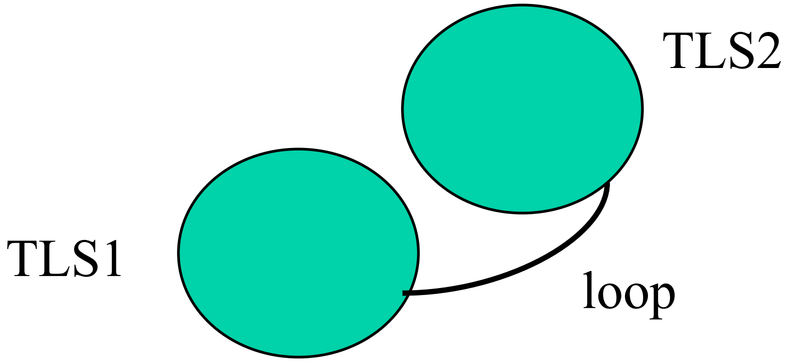
B value restraints and TLS

Designing restraints on B values is much more difficult.
Current available options to deal with B values at low resolutions

1) Group B as implemented in CNS
2) TLS group refinement as implemented in refmac and phenix.refine

Both of them have some applications. TLS seems to work for wide range of cases but unfortunately it is very often misused. One of the problems is discontinuity of B values. Neighbouring atoms may end up having wildly different B values

In ideal world anisotropic U with good restraints should be used. But this world is far far away yet. Only in some cases full aniso refinement at 3Å gives better R/Rfree than TLS refinement. These cases are with extreme ansiotropic data.



Kullback-Leibler divergence

If there are two densities of distributions – *p(x) and* q(x) then symmetrised Kullback-Leibler divergence between them is defined (it is distance between distributions)

$$\frac{1}{2}(\int_{-\infty}^{\infty} p(x)\log(\frac{p(x)}{q(x)})dx + \int_{-\infty}^{\infty} p(x)\log(\frac{q(x)}{p(x)})dx)$$

If both distributions are Gaussian with the same mean values and U_1 and U_2 variances then this distance becomes:

$$tr(U_1 U_2^{-1} + U_2 U_1^{-1} - 2I)$$

And for isotropic case it becomes

$$3(\frac{B_1}{B_2} + \frac{B_2}{B_1} - 2) = 3\frac{(B_1 - B_2)^2}{B_1 B_2}$$

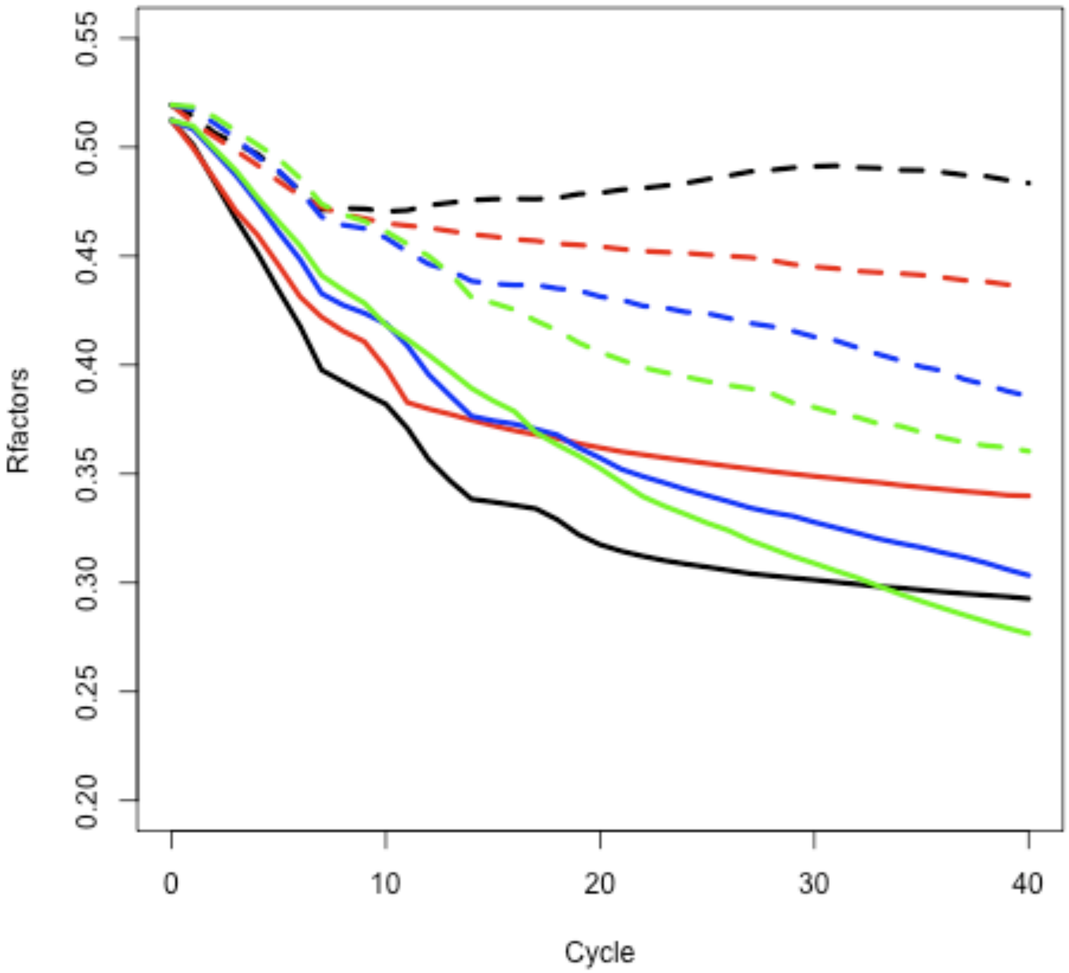
Restraints for bonded pairs have more weights more than for non-bonded pairs. For nonbonded atoms weights depend on the distance between atoms.

This type of restraint is also applied for rigid bond restraints in anisotropic refinement

Example, after molecular replacement
3A resolution, data completeness 71%

Rfactors vs cycle
Black – simple refinement
Red – Global NCS
Blue – Local NCS
Green – “Jelly” body

Solid lines – Rfactor
Dashed lines - Rfree



Example: 4A resolution, data from pdb 2r6c

Starting R/Rfree = 36.0/35.6

R/Rfree after 40 cycles of refinement

	None	Ncs local	Jelly body	External structure
R	20.80	21.44	23.72	23.38
Rfree	32.69	31.79	29.82	28.69

MAP SHARPENING: INVERSE PROBLEM

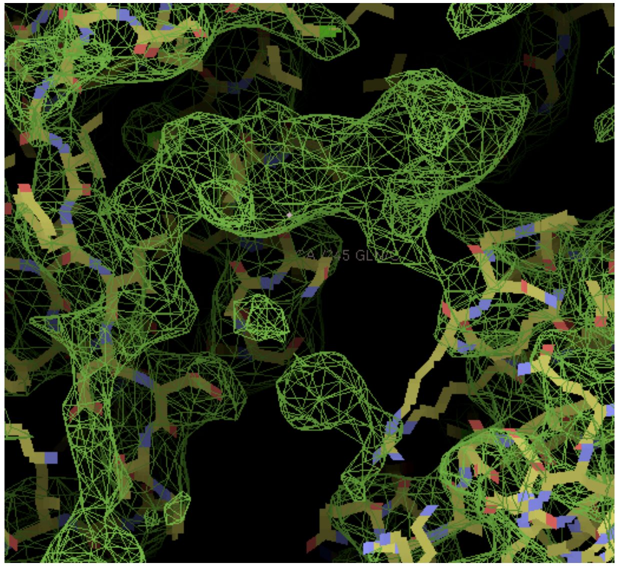
Simple map sharpening may increase signal as well as noise. Sometimes noise amplification may be very large and mask out the signal completely.
Noises in electron density: series termination, errors in phases and noises in experimental data.

Regularised map sharpening tries to reduce noise amplification while increasing signal

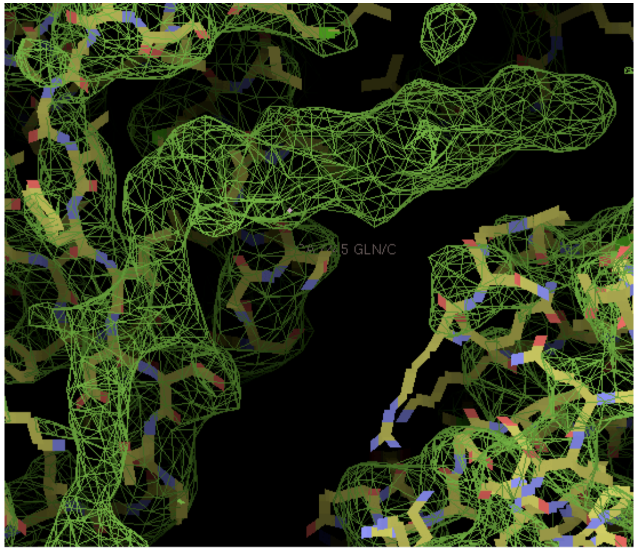
$$F_{deblurred} = \frac{e^{-B|s|^2/4}}{e^{-2B|s|^2/4} + \alpha\,|\,s\,|^2} F = K_\alpha(s,B)F$$

MAP SHARPENING: 2R6C, 4Å RESOLUTION

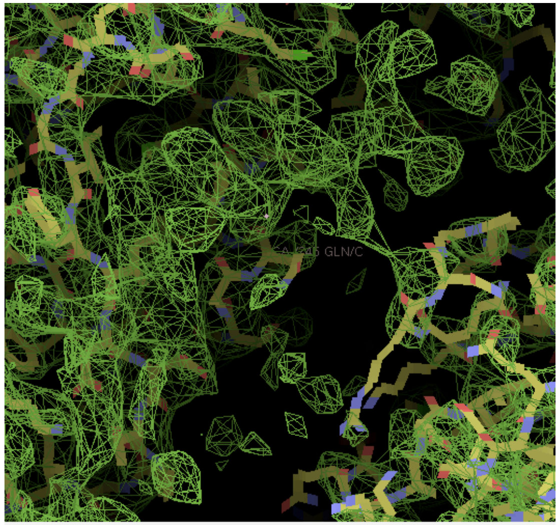
Original



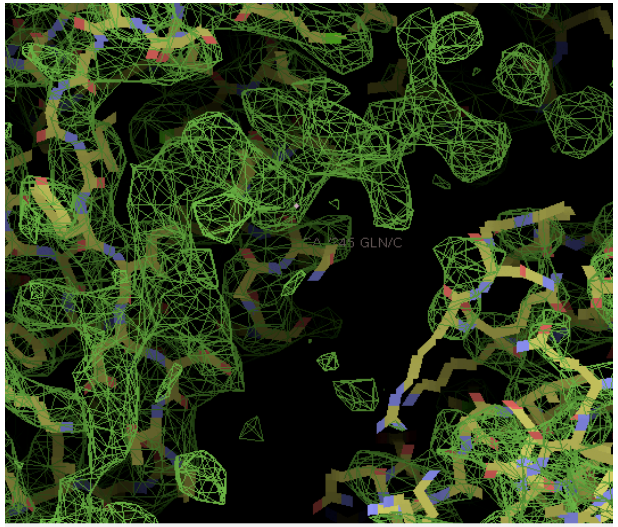
No sharpening



Sharpening, median B
α 0



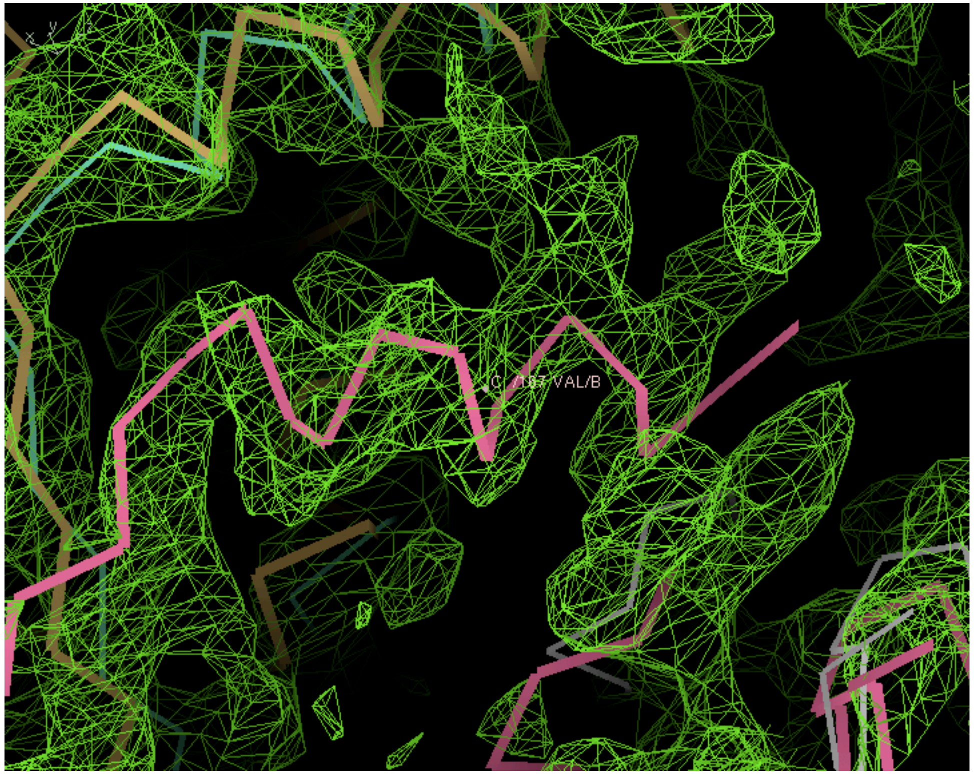
Sharpening, median B
α optimised



Top left and bottom: After local NCS refinement

Example, 2r6c, Electron density

Known structure (2r6a) superimposed to 2r6c structure. There is a helix. Side chains are visible to some degree



Alternative conformation of links: how to handle

Description

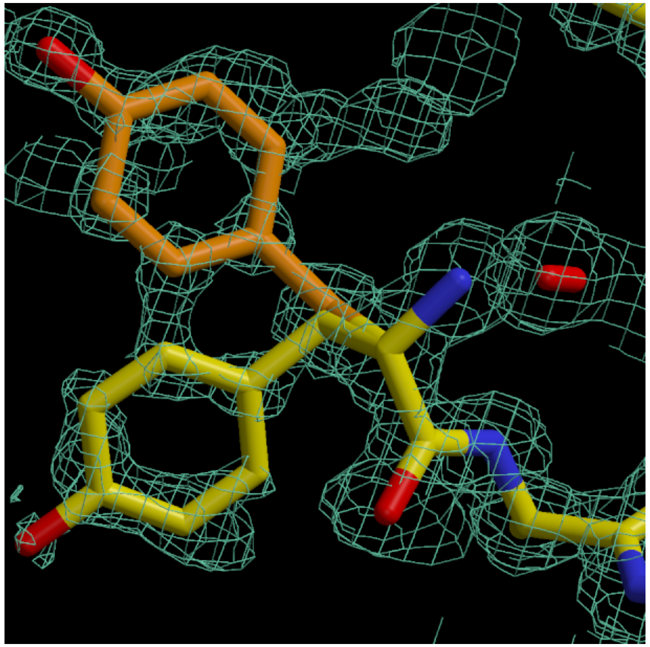
Description of link(s) should be added to the library. When residues make link then each component is usually modified. Description of Link should contain it also

PDB

```
LINK     C6  BBEN B   1                O1  BMAF S   2   BEN-MAF
LINK     OE2 AGLU A 320                C1  AMAF S   2   GLU-MAF
```

Alternative conformations

Example from 0.88Å catalase structure:Two conformations of Tyrosine. Ring is clearly in two conformation. To refine it properly CB also needs to be split. It helps adding hydrogen atom on CB and improves restraints in anisotropic U values



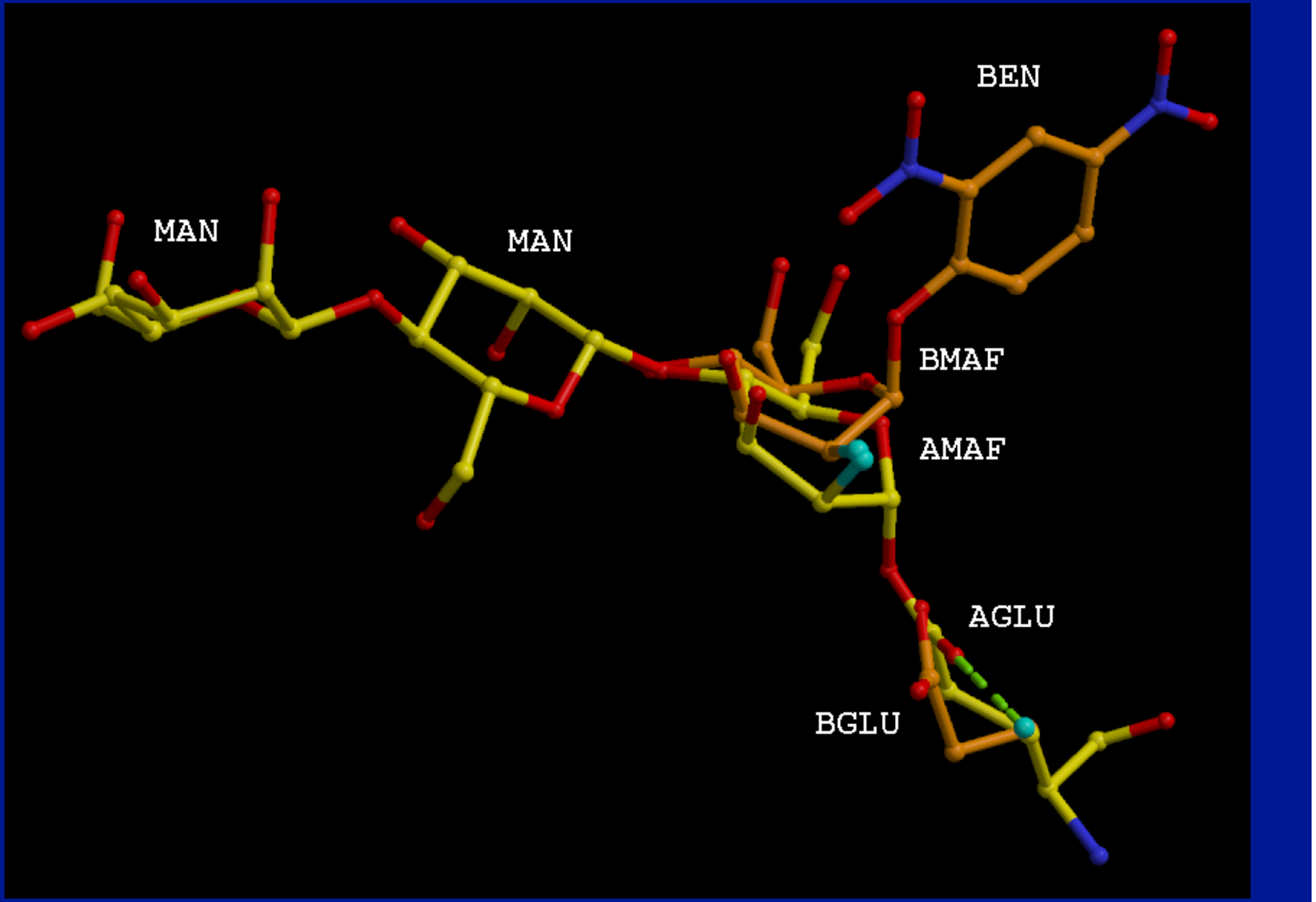
Alternative conformation: Example in pdb file

```
ATOM    977  N   GLU A  67     -11.870   9.060   4.949  1.00 12.89           N
ATOM    978  CA  GLU A  67     -12.166  10.353   4.354  1.00 14.00           C
ATOM    980  CB AGLU A  67     -13.562  10.341   3.738  0.50 14.81           C
ATOM    981  CB BGLU A  67     -13.526  10.285   3.654  0.50 14.35           C
ATOM    986  CG AGLU A  67     -13.701   9.400   2.573  0.50 16.32           C
ATOM    987  CG BGLU A  67     -13.876  11.476   2.777  0.50 14.00           C
ATOM    992  CD AGLU A  67     -15.128   9.179   2.134  0.50 17.17           C
ATOM    993  CD BGLU A  67     -15.237  11.332   2.110  0.50 15.68           C
ATOM    994  OE1AGLU A  67     -15.742  10.153   1.644  0.50 20.31           O
ATOM    995  OE1BGLU A  67     -15.598  12.213   1.307  0.50 16.68           O
ATOM    996  OE2BGLU A  67     -15.944  10.342   2.389  0.50 18.94           O
ATOM    997  OE2AGLU A  67     -15.610   8.027   2.235  0.50 21.30           O
ATOM    998  C   GLU A  67     -12.110  11.473   5.386  1.00 13.40           C
ATOM    999  O   GLU A  67     -11.543  12.528   5.110  1.00 12.98           O
```

Note that pdb is strictly formatted. Every element has its position

Link between residues in double conformation

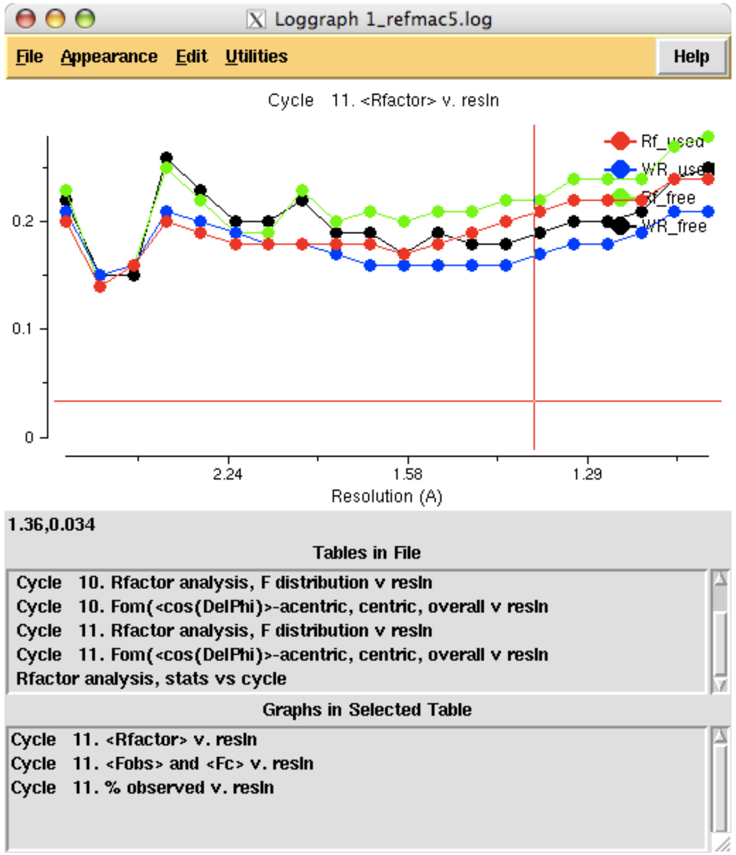
Fluro-modified sugar MAF is in two conformation. One of them is bound to GLU and another one is bound to ligand BEN



Things to look at

- R factor/Rfree: They should go down during refinement
- Geometric parameters: rms bond and other. They should be reasonable. For example rms bond should be around 0.02
- Map and coordinates using coot
- Logggraph outputs. That is available on the cpp4i interface
- Twin operators and twin fractions

Behaviour of R/Rfree, average Fobs vs resolution should be reasonable. If there is a bump or it has an irregular behaviour then something is wrong with your data or refinement.



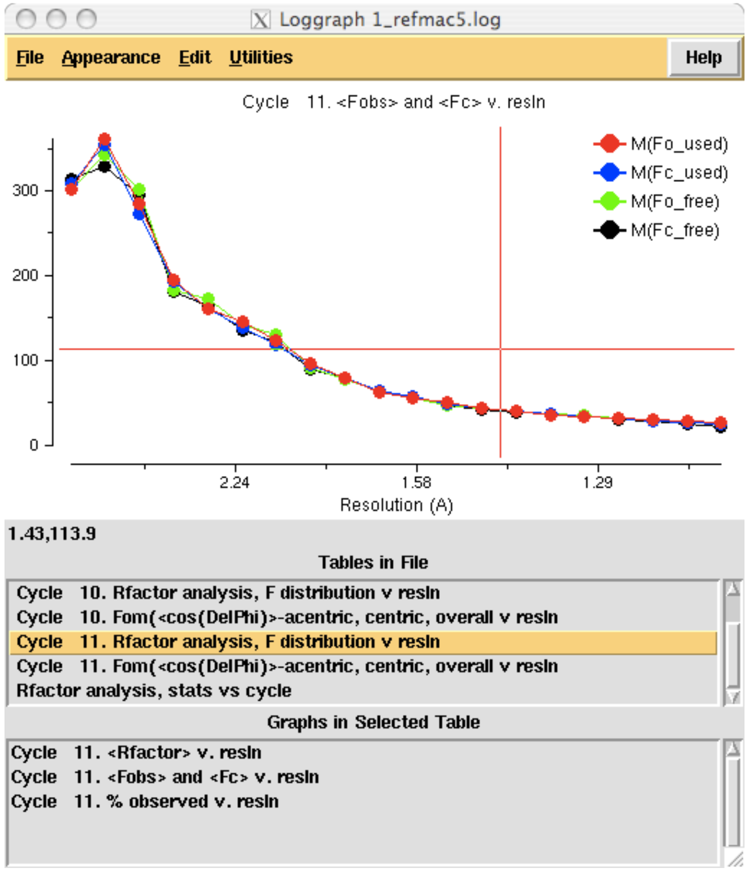
1.36,0.034

Tables in File

Cycle 10. Rfactor analysis, F distribution v resln
Cycle 10. Fom(<cos(DelPhi)>-acentric, centric, overall v resln
Cycle 11. Rfactor analysis, F distribution v resln
Cycle 11. Fom(<cos(DelPhi)>-acentric, centric, overall v resln
Rfactor analysis, stats vs cycle

Graphs in Selected Table

Cycle 11. <Rfactor> v. resln
Cycle 11. <Fobs> and <Fc> v. resln
Cycle 11. % observed v. resln



1.43,113.9

Tables in File

Cycle 10. Rfactor analysis, F distribution v resln
Cycle 10. Fom(<cos(DelPhi)>-acentric, centric, overall v resln
Cycle 11. Rfactor analysis, F distribution v resln
Cycle 11. Fom(<cos(DelPhi)>-acentric, centric, overall v resln
Rfactor analysis, stats vs cycle

Graphs in Selected Table

Cycle 11. <Rfactor> v. resln
Cycle 11. <Fobs> and <Fc> v. resln
Cycle 11. % observed v. resln

What and when

- Rigid body: At early stages - after molecular replacement or when refining against data from isomorphous crystals
- “Jelly” body – At early stages and may be at low resolution
- TLS - at medium and end stages of refinement at resolutions up to 1.7-1.6A (roughly)
- Anisotropic - At higher resolution towards the end of refinement
- Adding hydrogens - Higher than 2A but they could be added always
- Phased refinement - at early and medium stages of refinement
- SAD – at the early srages
- Twin - always (?). Be careful at early stages
- NCS local – always?
- Ligands - as soon as you see them
- What else?

How to use new features

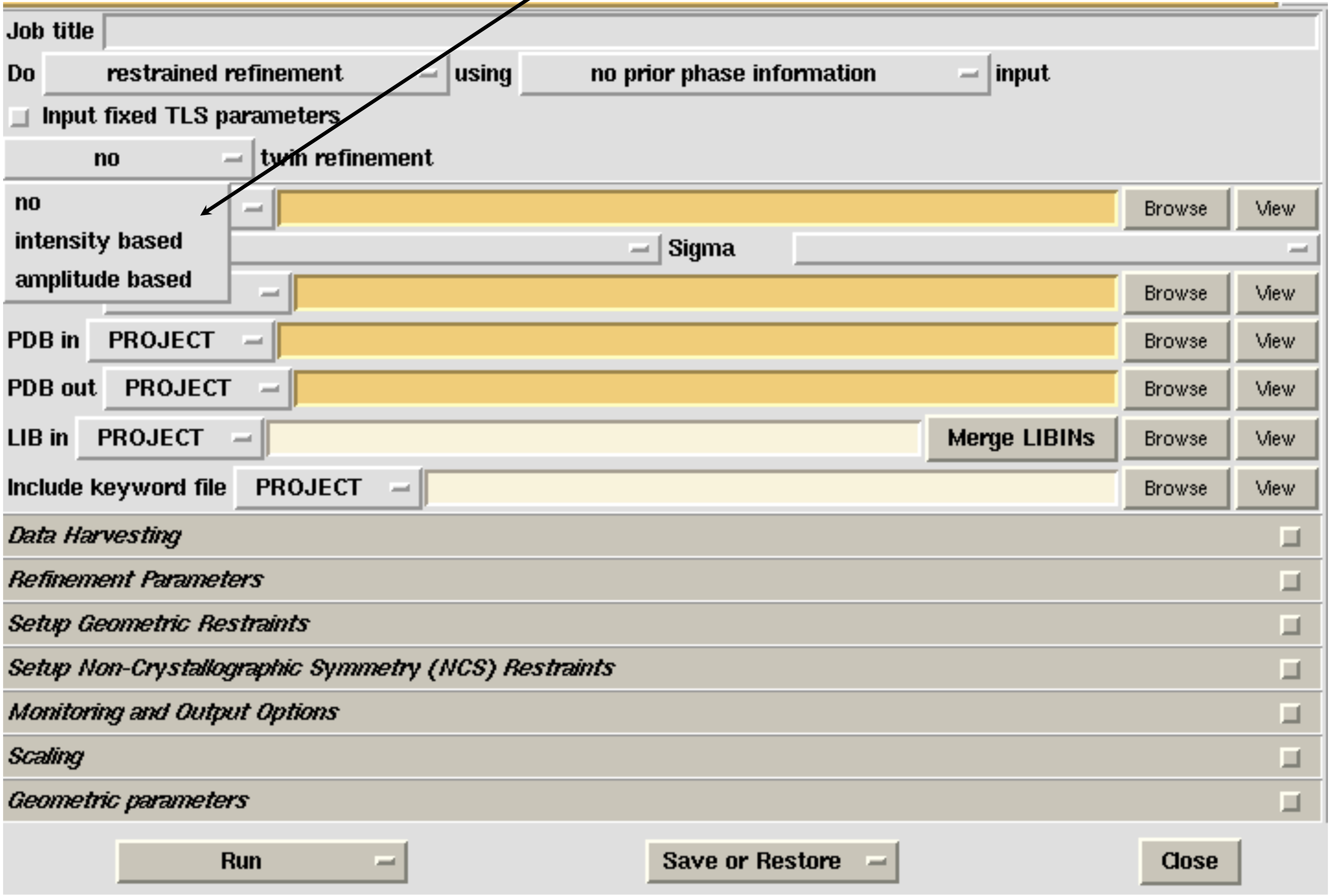
Download refmac from the website
www.ysbl.york.ac.uk/refmac/data/refmac_experimental/refmac5.6_linux.tar.gz
www.ysbl.york.ac.uk/refmac/data/refmac_experimental/refmac5.6_macintel.tar.gz

Download the dictionary:
www.ysbl.york.ac.uk/refmac/data/refmac_experimental/refmac5.6_dictionary_v5.18.gz

Change atom names using molprobity (optional: important if you have dna/rna)
http://molprobity.biochem.duke.edu/

Refmac refmac5 with the new one and you are ready for the new version.

Twin refinement (it works with older version also



Job title

Do restrained refinement using no prior phase information input

□ Input fixed TLS parameters

no twin refinement

Sigma

PDB in PROJECT

PDB out PROJECT

LIB in PROJECT

Include keyword file PROJECT

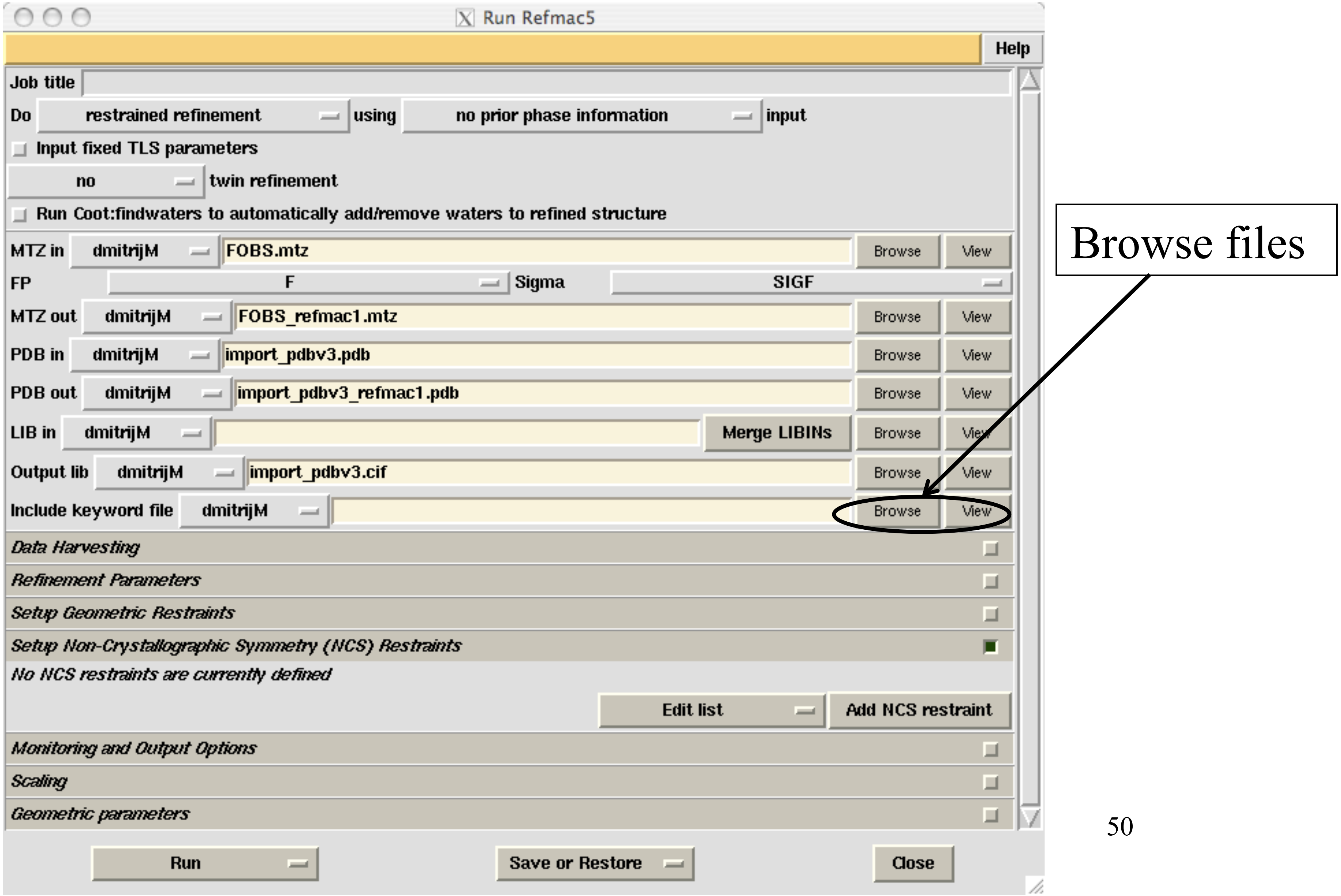
Adding external keywords

- Add the following command to a file:

```
ncsr local          # automatic and local ncs
ridg dist sigm 0.05  # jelly body restraints
mapcalculate shar    # regularised map sharpening
```

Save in a file (say keyw.dat)

Add external keywords file in refmac interface



Run Refmac5

Job title

Do restrained refinement using no prior phase information input

□ Input fixed TLS parameters

no twin refinement

□ Run Coot:findwaters to automatically add/remove waters to refined structure

MTZ in dmitrijM FOBS.mtz

FP F Sigma SIGF

MTZ out dmitrijM FOBS_refmac1.mtz

PDB in dmitrijM import_pdbv3.pdb

PDB out dmitrijM import_pdbv3_refmac1.pdb

LIB in dmitrijM

Output lib dmitrijM import_pdbv3.cif

Include keyword file dmitrijM

Data Harvesting □

Refinement Parameters □

Setup Geometric Restraints □

Setup Non-Crystallographic Symmetry (NCS) Restraints ■

No NCS restraints are currently defined

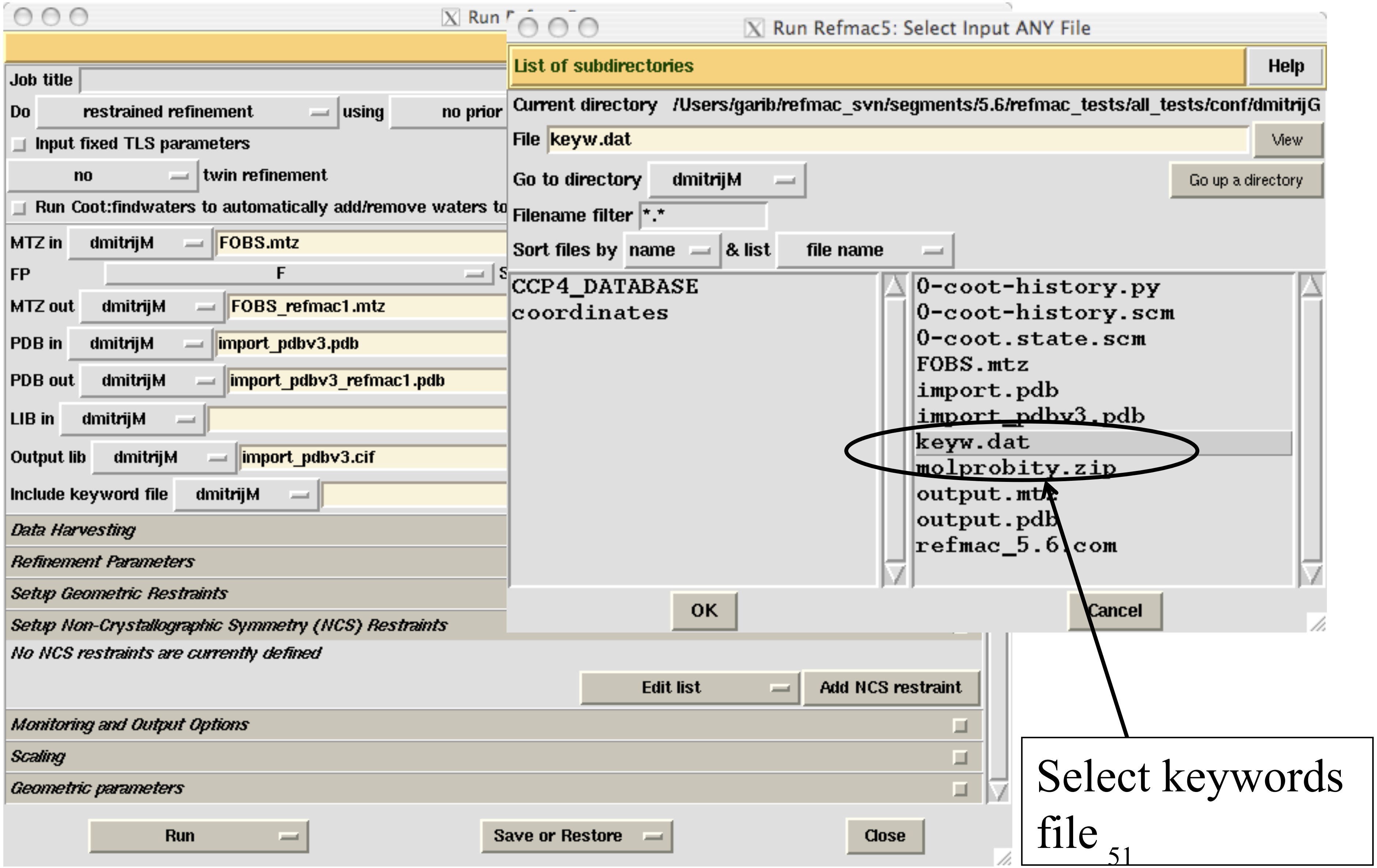
Monitoring and Output Options □

Scaling □

Geometric parameters □

Browse files

Add external keywords file in refmac interface



Run Refmac5: Select Input ANY File

Current directory /Users/garib/refmac_svn/segments/5.6/refmac_tests/all_tests/conf/dmitrijG

File keyw.dat

Go to directory dmitrijM

Filename filter *.*

Sort files by name & list file name

```
CCP4_DATABASE            0-coot-history.py
coordinates              0-coot-history.scm
                         0-coot.state.scm
                         FOBS.mtz
                         import.pdb
                         import_pdbv3.pdb
                         keyw.dat
                         molprobity.zip
                         output.mtz
                         output.pdb
                         refmac_5.6.com
```

Job title

Do restrained refinement using no prior

□ Input fixed TLS parameters

no twin refinement

□ Run Coot:findwaters to automatically add/remove waters to

MTZ in dmitrijM FOBS.mtz

FP F

MTZ out dmitrijM FOBS_refmac1.mtz

PDB in dmitrijM import_pdbv3.pdb

PDB out dmitrijM import_pdbv3_refmac1.pdb

LIB in dmitrijM

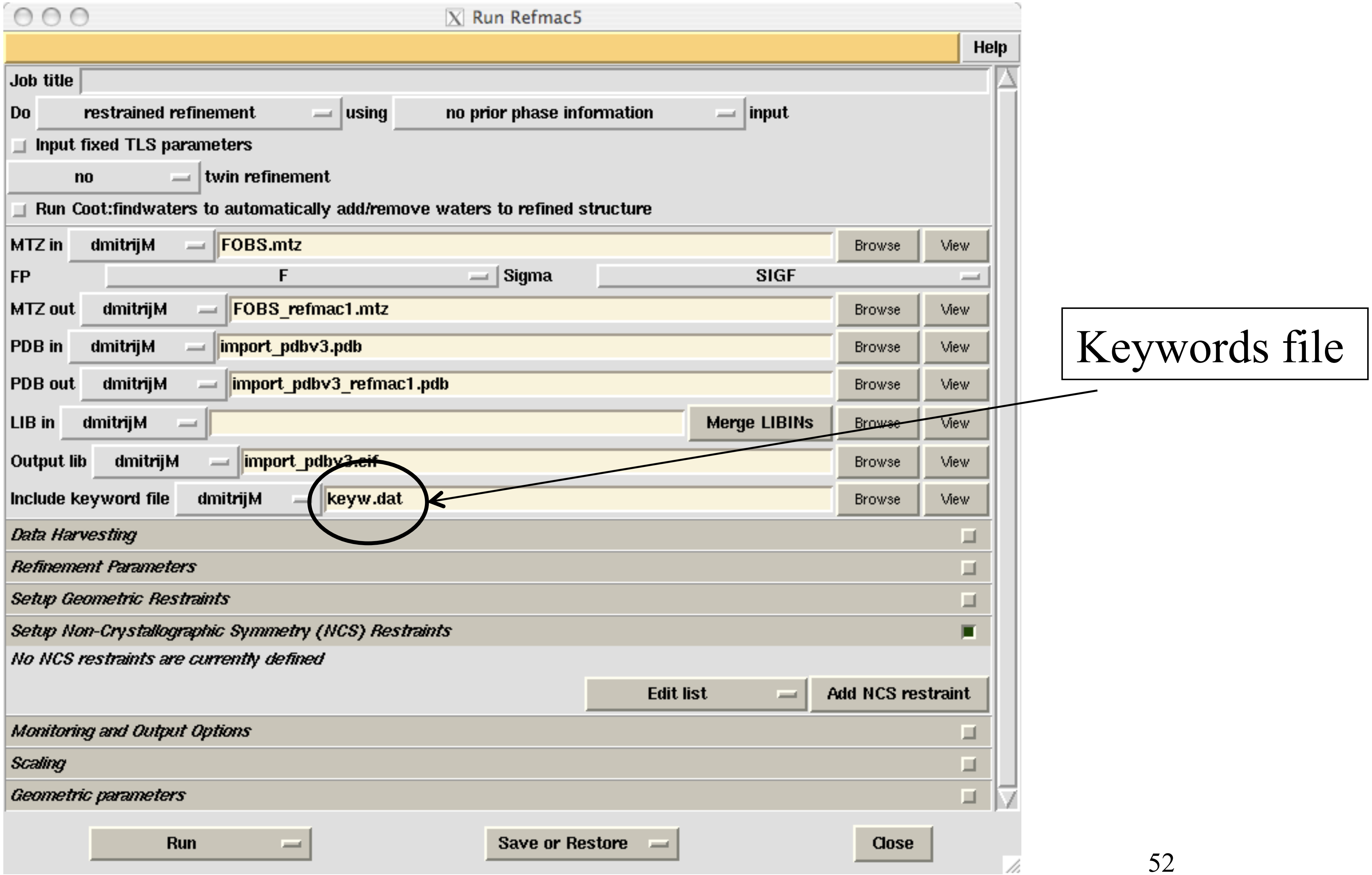
Output lib dmitrijM import_pdbv3.cif

Include keyword file dmitrijM

No NCS restraints are currently defined

Select keywords file

Add external keywords file in refmac interface



Run Refmac5

Job title

Do: restrained refinement using no prior phase information input

□ Input fixed TLS parameters

no twin refinement

□ Run Coot:findwaters to automatically add/remove waters to refined structure

MTZ in: dmitrijM FOBS.mtz

FP: F Sigma: SIGF

MTZ out: dmitrijM FOBS_refmac1.mtz

PDB in: dmitrijM import_pdbv3.pdb

PDB out: dmitrijM import_pdbv3_refmac1.pdb

LIB in: dmitrijM

Output lib: dmitrijM import_pdbv3.cif

Include keyword file: dmitrijM keyw.dat

Data Harvesting □

Refinement Parameters □

Setup Geometric Restraints □

Setup Non-Crystallographic Symmetry (NCS) Restraints ■

No NCS restraints are currently defined

Monitoring and Output Options □

Scaling □

Geometric parameters □

Keywords file

Conclusion

- SAD/SIRAS can improve behaviour of model building
- Twin refinement improves statistics and occasionally electron density
- Use of similar structures should improve reliability of the derived model: Especially at low resolution
- NCS restraints must be done automatically: but conformational flexibility must be accounted for
- “Jelly” body works better than I thought it should
- Regularised map sharpening looks promising. More work should be done on series termination and general sharpening operators

Acknowledgment

York

Alexei Vagin

Andrey Lebedev

Rob Nocholls

Fei Long

Leiden

Pavol Skubak

Raj Pannu

CCP4, YSBL people

REFMAC is available from CCP4 or from York's ftp site:

www.ysbl.york.ac.uk/refmac/latest_refmac.html

This and other presentations can be found on:

www.ysbl.york.ac.uk/refmac/Presentations/