TLS REFINEMENT WITH REFMAC5

Martyn Winn
CCP4, Daresbury Laboratory, U.K.

Florence, August 23rd 2005

REFMAC5

*REFMAC*5 is a program for the refinement of macromolecular structures, written by Garib Murshudov et al., York, UK.
Distributed as part of the CCP4 suite:

http://www.ccp4.ac.uk/download.php
http://www.ysbl.york.ac.uk/~garib/refmac/latest_refmac.html

Some points about the program:

It is strongly based on ML and Bayesian statistics
It is very easy to use (CCP4*i*)
It has an extensive built-in dictionary
It allows various tasks (model idealisation, rigid-body refinement, phased and non-phased restrained and unrestrained refinement)
It allows a flexible model parameterization (iso-,aniso-, mixed-ADPs, TLS, bulk solvent)
It exploits a good minimization algorithm

What's new in Refmac?

Current version is 5.2

Recent major changes are in the dictionary and ligand handling.

Refmac publications:

- Murshudov, G.N. & al. (1997), Refinement of macromolecular structures by the maximum-likelihood method, Acta Cryst. D**53**, 240-255
- Murshudov, G.N. & al. (1999), Efficient anisotropic refinement of macromolecular strcutures using FFT, Acta Cryst. D**55**, 247-255
- Winn, M.D. & al. (2001), Use of TLS parameters to model anisotropic displacement parameters, Acta Cryst. D**57**, 122-133
- Steiner, R.A. & al. (2003), Fisher's information in maximum-likelihood macromolecular crystallographic refinement, Acta Cryst. D**59**, 2114-2124
- Vagin, A.A. & al. (2004), REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use, Acta Cryst. D**60**, 2184-2195

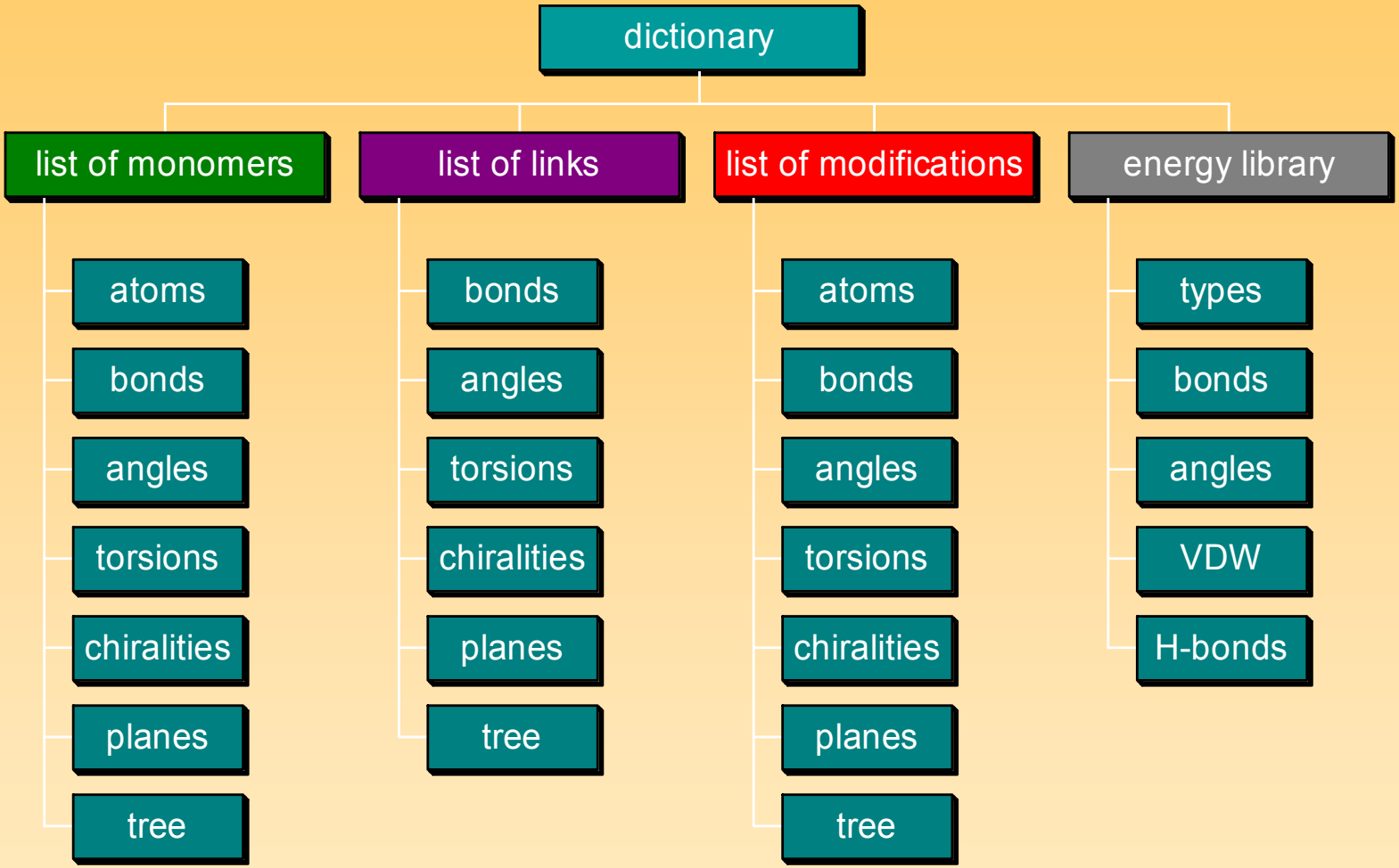
See Roberto Steiner's tutorial from ACA05:

http://www.ccp4.ac.uk/courses/ACA2005/PPT/ACA05_Refmac5.ppt

Covers:

- Maximum likelihood refinement
- Dictionary
- TLS refinement

Organization of dictionary



Links e.g. trans or cis peptide bond, disulphide bond, sugar links, etc
Modifications e.g. NH_3 terminus, phosphorylation, renaming, etc.

Monomer library

$CCP4/lib/data/monomers/

ener_lib.cif	Definition of atom types
mon_lib_list.html	Lists monomers, links, modifications with hyperlinks to detailed descriptions
0/,1/,...a/,b/,...	Definition of various monomers

~1000 monomers have complete description

Others have “minimal” description

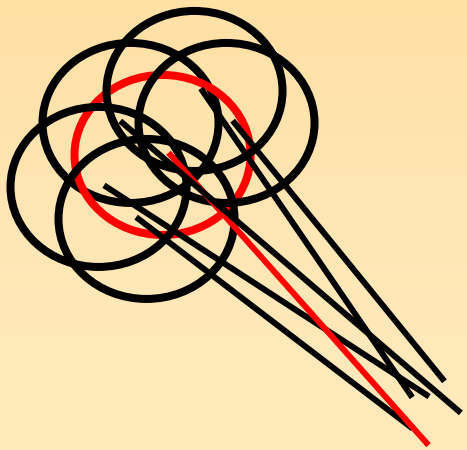
If you have monomer(s) in your coordinate file for which there is no/minimal description, REFMAC5 generates a complete library description (monomer.cif) and then it stops so you can check the result.

Also: Sketcher in ccp4i, PRODRG2 server

TLS refinement: Aims

- Experiment measures time- and space-averaged structure

- In addition to mean **atomic positions**, mean square **atomic displacements** from mean position (static and dynamic) are an important part of the model of a protein.



TLS refinement: Aims

- Atomic displacements are likely anisotropic, but rarely have luxury of refining individual anisotropic Us. Instead isotropic Bs.

- TLS parameterisation allows an intermediate description

T = translation
L = libration
S = screw-motion

Contributions to atomic U

$$U = U_{crystal} + U_{TLS} + U_{internal} + U_{atom}$$

$U_{crystal}$: overall anisotropic scale factor w.r.t. crystal axes.

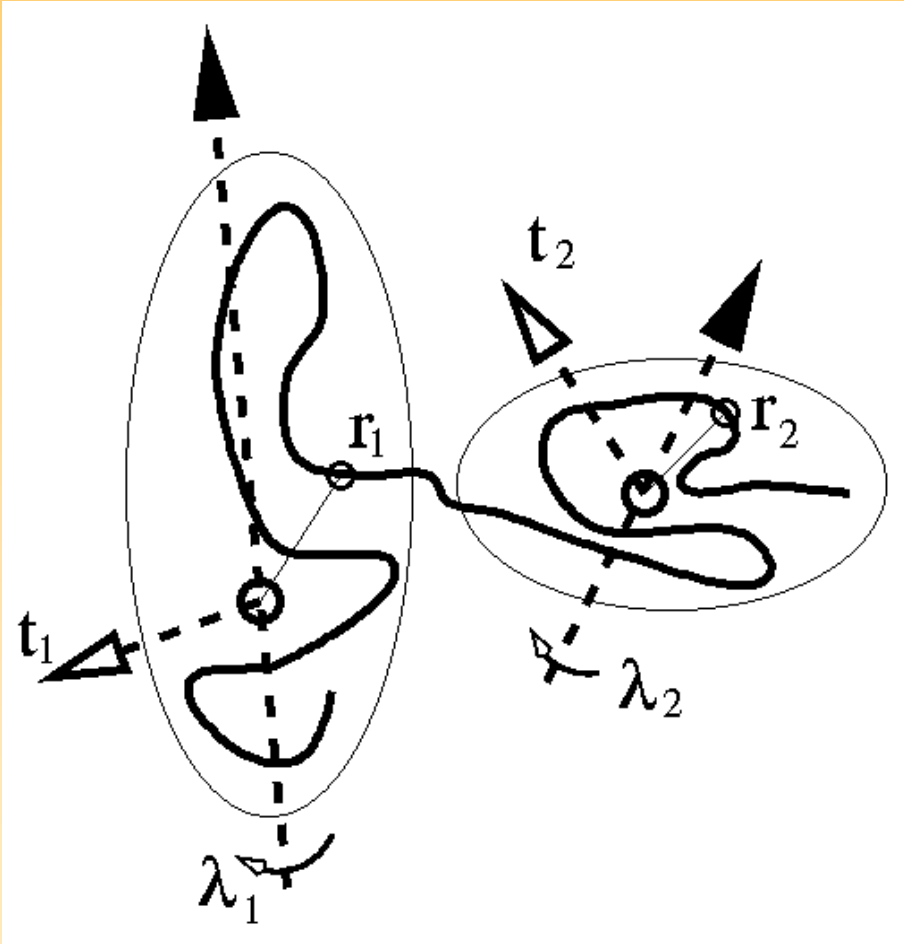
U_{TLS} : pseudo rigid body displacements e.g. of molecules, domains, secondary structure elements, side groups, etc.

$U_{internal}$: internal displacements of molecules, e.g. normal modes of vibration, torsions, etc.

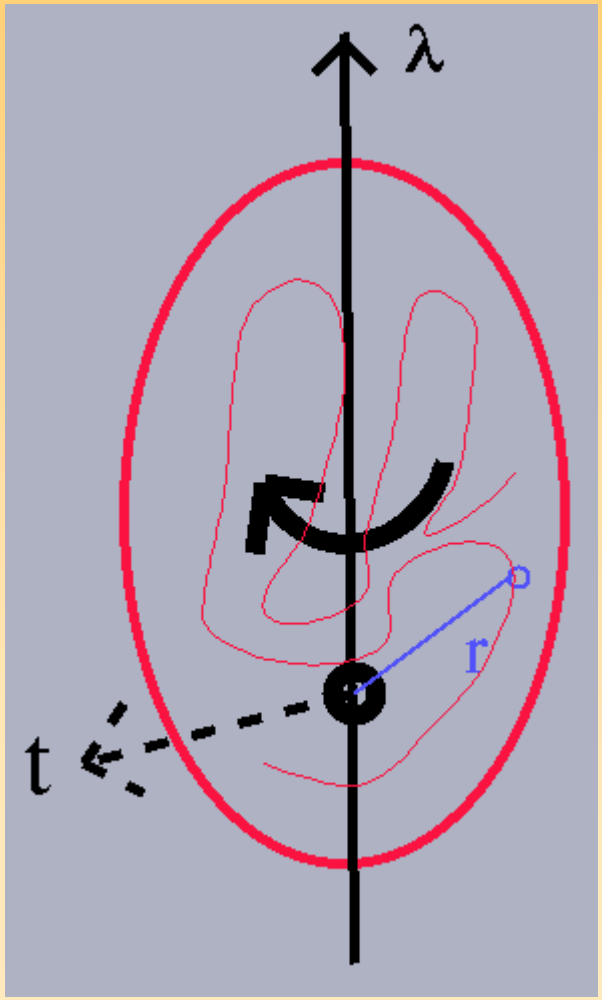
U_{atom} : anisotropy of individual atoms

U are 3x3 tensors. $B = (8\pi^2/3)\,\mathrm{Tr}(U)$

Rigid body model



Rigid body motion



General displacement of atom (position **r** w.r.t. origin O) in rigid body**:**

$$\mathbf{u} = \mathbf{t} + \mathsf{D}.\mathbf{r}$$

For small libration **λ:**

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

TLS parameters

Experiment yields **mean square** displacements

- Corresponding dyad:

$$\mathbf{uu} = \mathbf{tt} + \mathbf{t}\lambda \times \mathbf{r} - \mathbf{r} \times \lambda\mathbf{t} - \mathbf{r} \times \lambda\lambda \times \mathbf{r}$$

- Average over dynamic motion and static disorder gives atomic anisotropic displacement parameter (ADP):

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

- T, L and S describe mean square translation and libration of rigid body and their correlation.
- T $\Rightarrow$ 6 parameters, L $\Rightarrow$ 6 parameters, S $\Rightarrow$ 8 parameters (trace of S is undetermined)

Use of TLS

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

- Given refined atomic U's, fit TLS parameters

  - analysis

  - Harata, K. & Kanai, R., (2002) Crystallographic dissection of the thermal motion of protein-sugar complex, Proteins, **48**, 53-62
  - Wilson, M.A. & Brunger, A.T.., (2000) The 1.0 Å crystal structure of Ca(2+)-bound calmodulin: an analysis of disorder and implications for functionally relevant plasticity, J. Mol. Biol. **301**, 1237-1256

- Use TLS as refinement parameters

  TLS $\Rightarrow$ U's $\Rightarrow$ structure factor

  - refinement

  – Winn et al., (2003) Macromolecular TLS refinement in REFMAC at moderate resolutions, *Methods Enzymol*., **374**, 300-321

TLS in refinement

- TLS parameters are contribution to displacement parameters of model
- Can specify 1 or more TLS groups to describe contents of asymmetric unit (or part thereof)
- 6 + 6 + 8 = 20 parameters per group (irrespective of number of atoms in the group)

- **Number of extra refinement parameters depends on how many groups used!**

At what resolution can I use TLS?

Any! Resolution only affects level of detail:

- Resolution < **1.2 Å** - full anisotropic refinement
- Resolution ~ **1.5 Å** - marginal for full anisotropic refinement. But can do detailed TLS, e.g. Howlin *et al*, Ribonuclease A, **1.45 Å**, 45 side chain groups; Harris *et al*, papain, **1.6Å**, 69 side chain groups.
- Resolution **1.5 Å** - **2.5 Å** $\Rightarrow$ model molecules/domains rather than side chains.
- T.Sandalova *et al (*2001) *PNAS*, ***98***, 9533-9538 - thioredoxin reductase at **3.0 Å** - TLS group for each of 6 monomers in asu

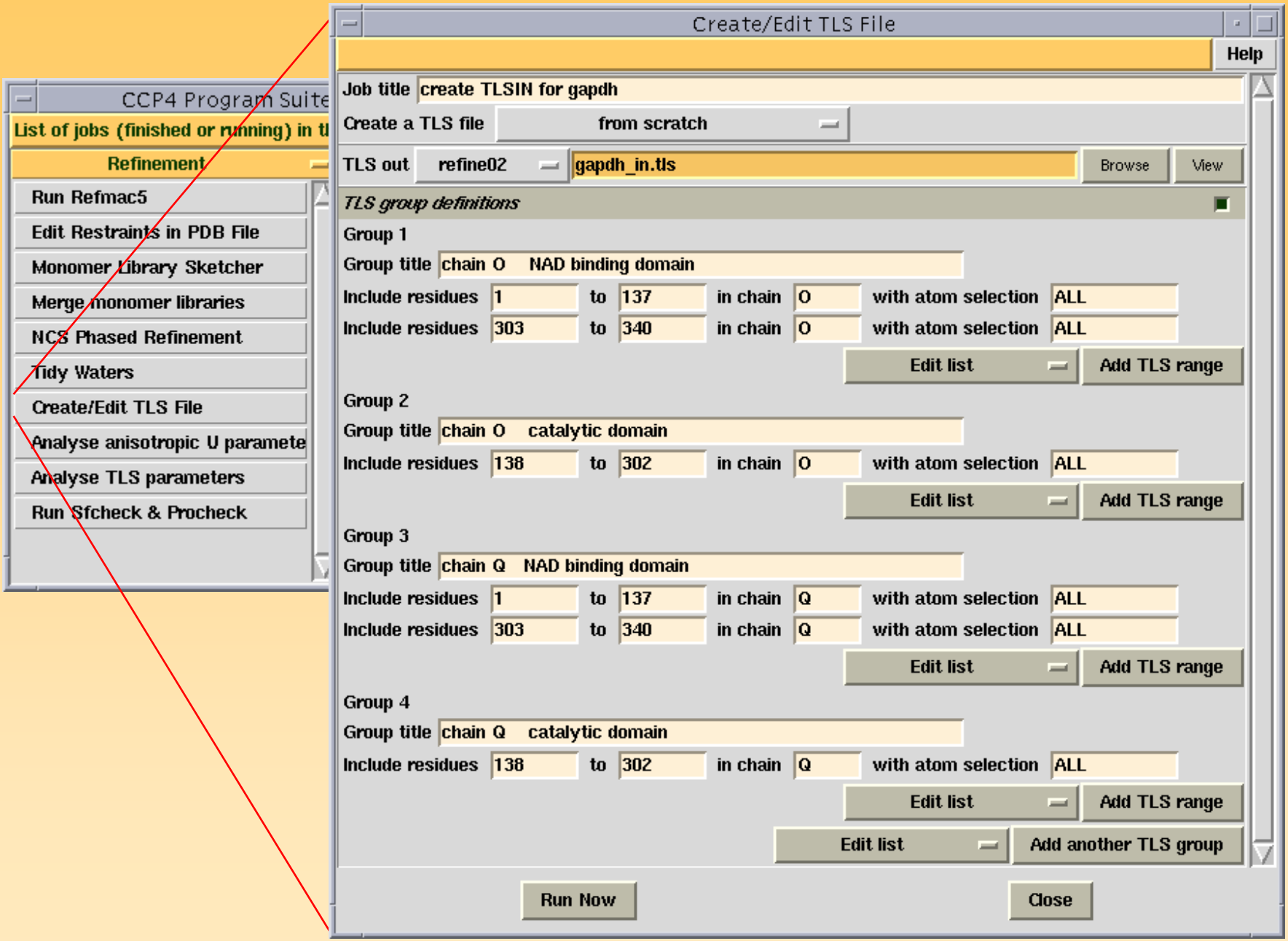
In fact, rigid-body assumption works better at low resolution

Implementation in REFMAC

Refinement parameters: scaling parameters, TLS parameters and residual B factors.

Typical procedure:

- Specify TLS groups (currently via TLSIN file).
- Use anisotropic scaling.
- Set B values to constant value
- Refine TLS parameters (and scaling parameters) against ML residual.
- Refine coordinates and residual B factors.


CCP4 Program Suite
List of jobs (finished or running) in t
Refinement
Run Refmac5
Edit Restraints in PDB File
Monomer Library Sketcher
Merge monomer libraries
NCS Phased Refinement
Tidy Waters
Create/Edit TLS File
Analyse anisotropic U paramete
Analyse TLS parameters
Run Sfcheck & Procheck
Create/Edit TLS File
Help
Job title create TLSIN for gapdh
Create a TLS file from scratch
TLS out refine02 gapdh_in.tls Browse View
TLS group definitions
Group 1
Group title chain O NAD binding domain
Include residues 1 to 137 in chain O with atom selection ALL
Include residues 303 to 340 in chain O with atom selection ALL
Edit list Add TLS range
Group 2
Group title chain O catalytic domain
Include residues 138 to 302 in chain O with atom selection ALL
Edit list Add TLS range
Group 3
Group title chain Q NAD binding domain
Include residues 1 to 137 in chain Q with atom selection ALL
Include residues 303 to 340 in chain Q with atom selection ALL
Edit list Add TLS range
Group 4
Group title chain Q catalytic domain
Include residues 138 to 302 in chain Q with atom selection ALL
Edit list Add TLS range
Edit list Add another TLS group
Run Now
Close

Initial TLSIN file

```
TLS  Chain O NAD binding
RANGE  'O   1.' 'O 137.' ALL
RANGE  'O 303.' 'O 340.' ALL

TLS    Chain O Catalytic
RANGE  'O 138.' 'O 302.' ALL

TLS    Chain Q NAD binding
RANGE  'Q   1.' 'Q 137.' ALL
RANGE  'Q 303.' 'Q 340.' ALL

TLS    Chain Q Catalytic
RANGE  'Q 138.' 'Q 302.' ALL
```

Choice of TLS groups

- Chemical knowledge, e.g. aromatic side groups of amino acids, secondary structure elements, domains, molecules

- Best fit of TLS to ADPs of test structure, or rigid-body criterion applied to ADPs. Both implemented in CCP4 program ANISOANL.

S.R.Holbrook & S.H.Kim, *J.Mol.Biol.*, **173**, 361 (1984)

T.R.Schneider, in *Proc. CCP4 Study Weekend*, 133 (1996)

M.J.Bernett *et al*, *Proteins Struct. Funct. Genet.*, **57**, 626 (2004)

Choice of TLS groups (cont.)

- Fit TLS groups to refined ADPs or isotropic B factors. TLSMD server finds best single group, then best split into 2 groups, etc. See later slide.

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

- Dynamic domains identified from multiple configurations, e.g. more than one crystal form (DYNDOM), difference distance matrices (ESCET), MD simulations.

S.Hayward and H.J.C.Berendsen, *Proteins Struct. Funct. Genet*., **30**, 144, (1998)

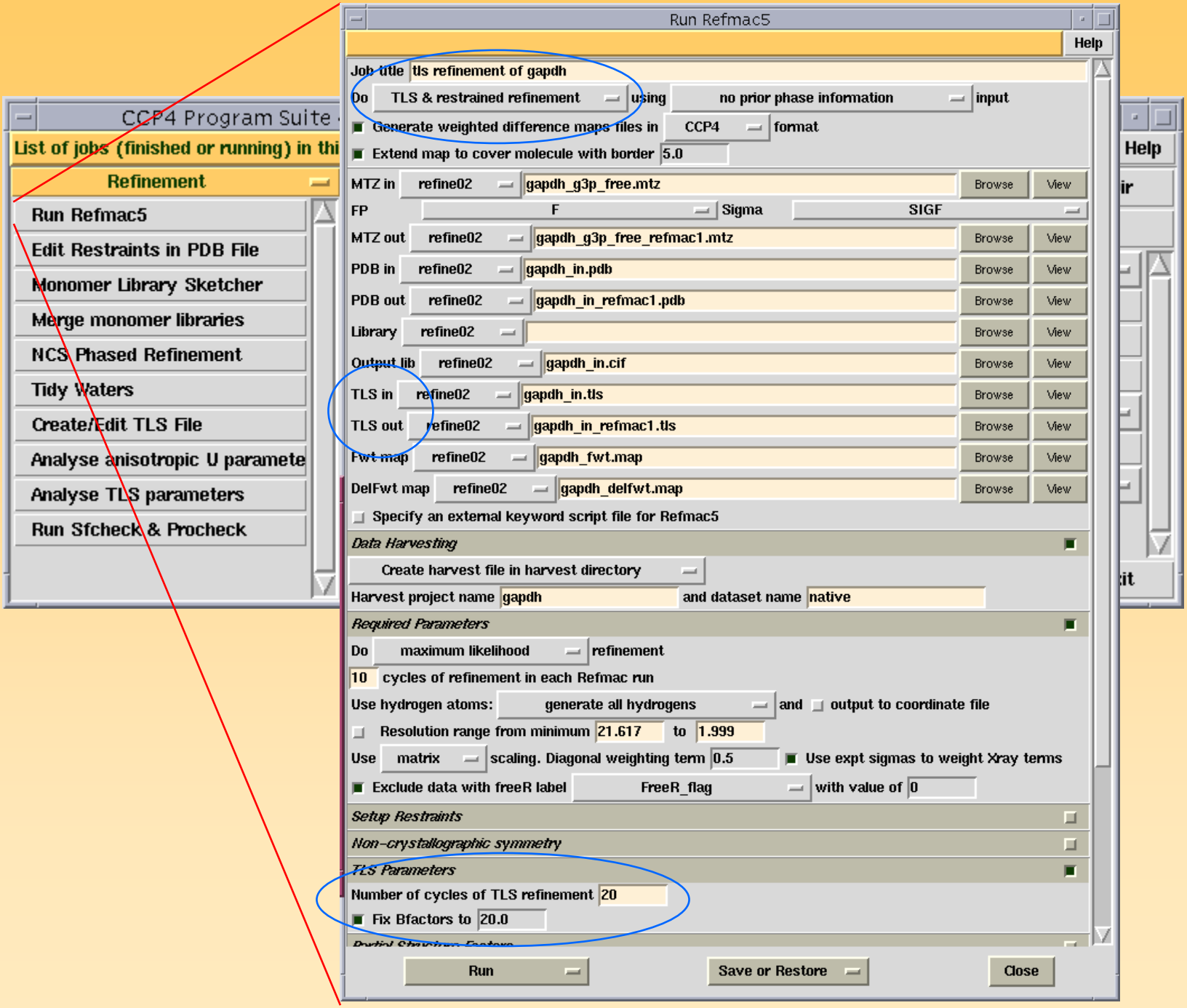
T.R.Schneider, *Acta Cryst*. **D60**, 2269 (2004)

Initialisation of B factors

1. First TLS refinement.
Set B values to constant value (precise value irrelevant) - allows TLS parameters to describe coarse-grained features

2. Subsequent TLS refinement.
TLSIN contains TLS parameters from earlier cycle. Can then keep residual Bs from earlier cycle.

3. Problematic TLS refinements (many `MAKE_U_POSITIVE` warnings).
Re-refine TLS parameters from initial zero values, but recycling earlier residual Bs. Aids stability in some cases.

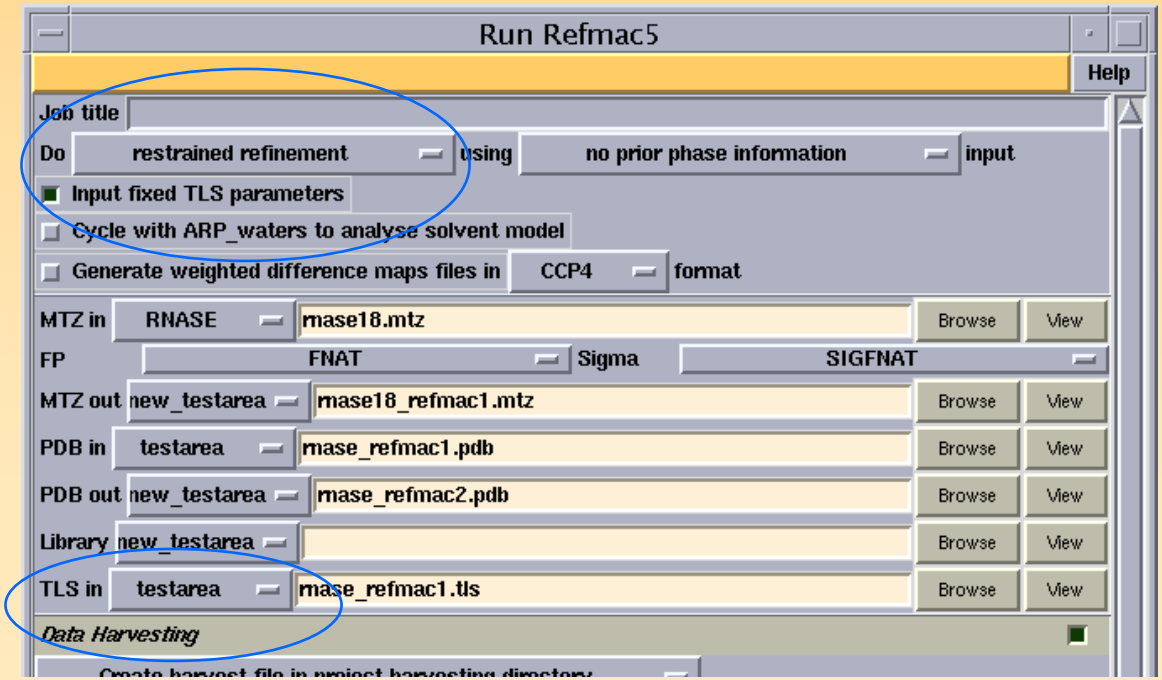

CCP4 Program Suite
List of jobs (finished or running) in thi
Refinement
Run Refmac5
Edit Restraints in PDB File
Monomer Library Sketcher
Merge monomer libraries
NCS Phased Refinement
Tidy Waters
Create/Edit TLS File
Analyse anisotropic U paramete
Analyse TLS parameters
Run Sfcheck & Procheck
Run Refmac5
Help
Job title tls refinement of gapdh
Do TLS & restrained refinement using no prior phase information input
Generate weighted difference maps files in CCP4 format
Extend map to cover molecule with border 5.0
MTZ in refine02 gapdh_g3p_free.mtz Browse View
FP F Sigma SIGF
MTZ out refine02 gapdh_g3p_free_refmac1.mtz Browse View
PDB in refine02 gapdh_in.pdb Browse View
PDB out refine02 gapdh_in_refmac1.pdb Browse View
Library refine02 Browse View
Output lib refine02 gapdh_in.cif Browse View
TLS in refine02 gapdh_in.tls Browse View
TLS out refine02 gapdh_in_refmac1.tls Browse View
Fwt map refine02 gapdh_fwt.map Browse View
DelFwt map refine02 gapdh_delfwt.map Browse View
Specify an external keyword script file for Refmac5
Data Harvesting
Create harvest file in harvest directory
Harvest project name gapdh and dataset name native
Required Parameters
Do maximum likelihood refinement
10 cycles of refinement in each Refmac run
Use hydrogen atoms: generate all hydrogens and output to coordinate file
Resolution range from minimum 21.617 to 1.999
Use matrix scaling. Diagonal weighting term 0.5 Use expt sigmas to weight Xray terms
Exclude data with freeR label FreeR_flag with value of 0
Setup Restraints
Non-crystallographic symmetry
TLS Parameters
Number of cycles of TLS refinement 20
Fix Bfactors to 20.0
Run
Save or Restore
Close

Re-running Refmac

1. If major re-building or changes in model - start again from zero TLS parameters. Ensures realistic set of TLS parameters.

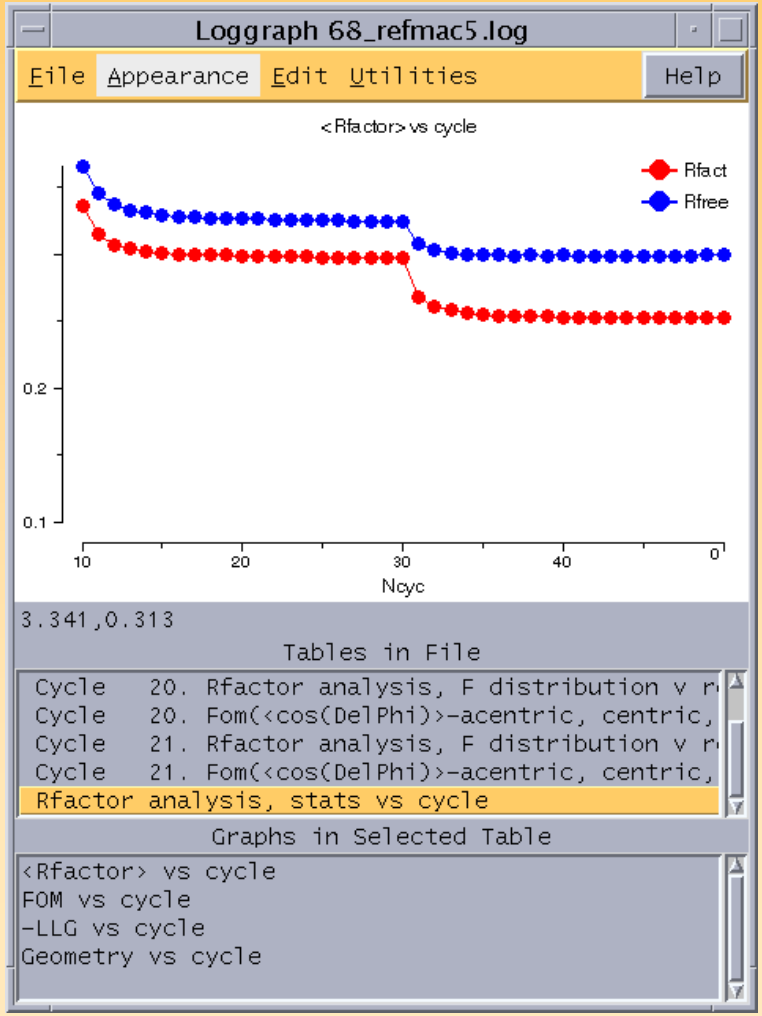
2. If minor re-building - TLSIN is TLSOUT from previous cycle

3. Can input fixed TLS parameters, and do restrained refinement only.



What to look for in output

- Usual refinement statistics.
- Check R_free and TLS parameters in log file for convergence.
- Check TLS parameters to see if any dominant displacements.
- Pass `XYZOUT` and `TLSOUT` through `TLSANL` for analysis
- Consider alternative choices of TLS groups



Affect on electron density

- In general, produces “cleaner” electron density.
- Often allows additional minor model building.
- Occasionally allows major model building.

TLSOUT from Refmac

```
TLS  Chain O NAD binding
RANGE  'O   1.' 'O 137.' ALL
RANGE  'O 303.' 'O 340.' ALL
ORIGIN 94.142 6.943 75.932
T  0.5200 0.5278 0.5031 -0.0049 -0.0040 0.0081
L  0.11   0.08   0.04   -0.03    0.02   0.06
S -0.014  0.020  0.002   0.022   0.000 -0.016  -0.005 0.022
```

.

.

.

Running TLSANL

`XYZIN:` output coordinates from refmac with residual B factors (BRESID keyword)

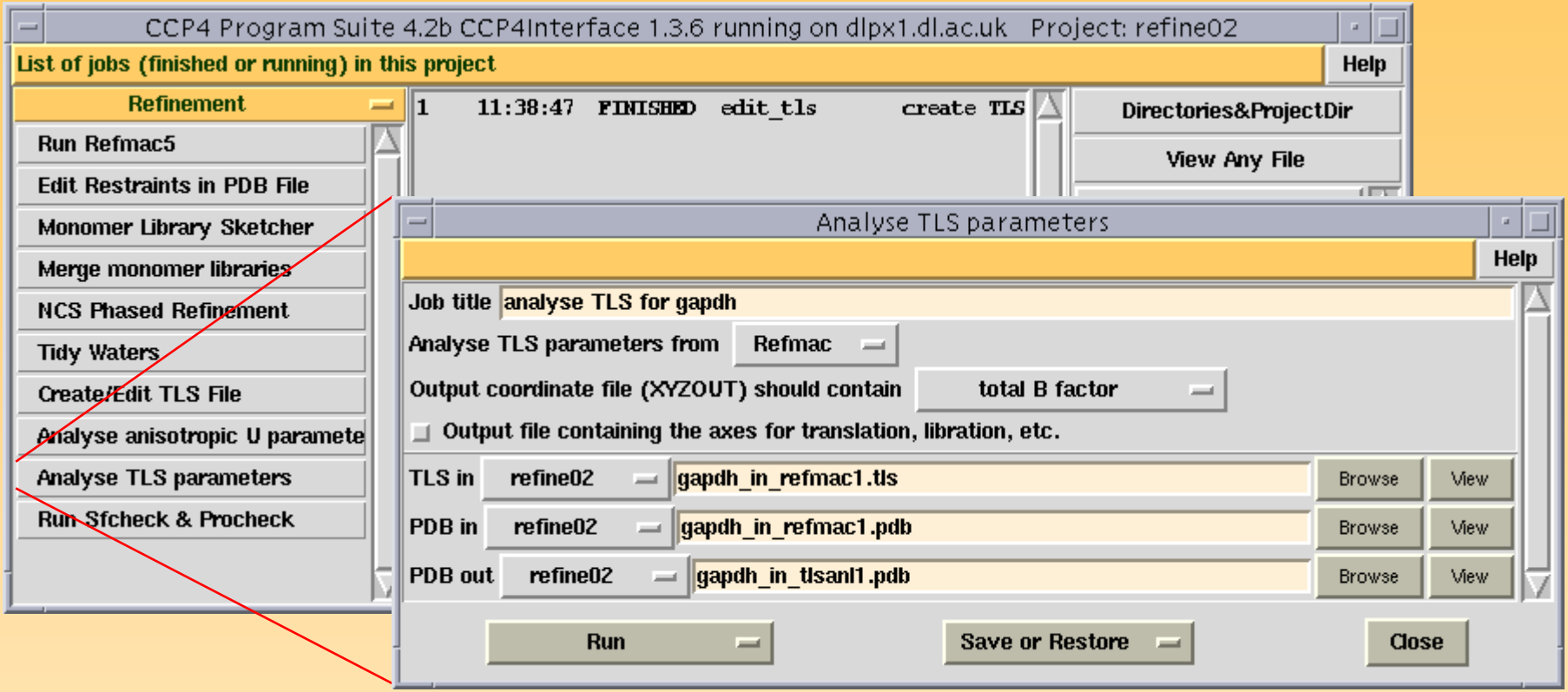
`TLSIN:` output TLS parameters from refmac

`XYZOUT:` ANISOU records including TLS and residual B contributions

ATOM records containing choice of B (ISOOUT keyword)

`AXES:` If AXES keyword set, file of principal axes in `mmCIF` format (for ccp4mg) or in `molscript` format

Howlin, B. et al. (1993) TLSANL: TLS parameter-analysis program for segmented anisotropic refinement of macromolecular structures, *J. Appl. Cryst.* **26**, 622-624


CCP4 Program Suite 4.2b CCP4Interface 1.3.6 running on dlpx1.dl.ac.uk Project: refine02
List of jobs (finished or running) in this project
Help
Refinement
Run Refmac5
Edit Restraints in PDB File
Monomer Library Sketcher
Merge monomer libraries
NCS Phased Refinement
Tidy Waters
Create/Edit TLS File
Analyse anisotropic U paramete
Analyse TLS parameters
Run Sfcheck & Procheck
1 11:38:47 FINISHED edit_tls create TLS
Directories&ProjectDir
View Any File
Analyse TLS parameters
Help
Job title analyse TLS for gapdh
Analyse TLS parameters from Refmac
Output coordinate file (XYZOUT) should contain total B factor
Output file containing the axes for translation, libration, etc.
TLS in refine02 gapdh_in_refmac1.tls Browse View
PDB in refine02 gapdh_in_refmac1.pdb Browse View
PDB out refine02 gapdh_in_tlsanl1.pdb Browse View
Run
Save or Restore
Close

Log output of TLSANL

Origins (T and S, but not L, origin-dependent):

- Origin of calculation
- Centre of Reaction

Axial systems for each tensor:

- orthogonal
- librational

"Simplest" description:

3 non-intersecting screw axes + 3 reduced translations

GAPDH: dimer as 1 TLS group

Eigenvalues of reduced translation tensor:

0.267 Å^2

0.112 Å^2

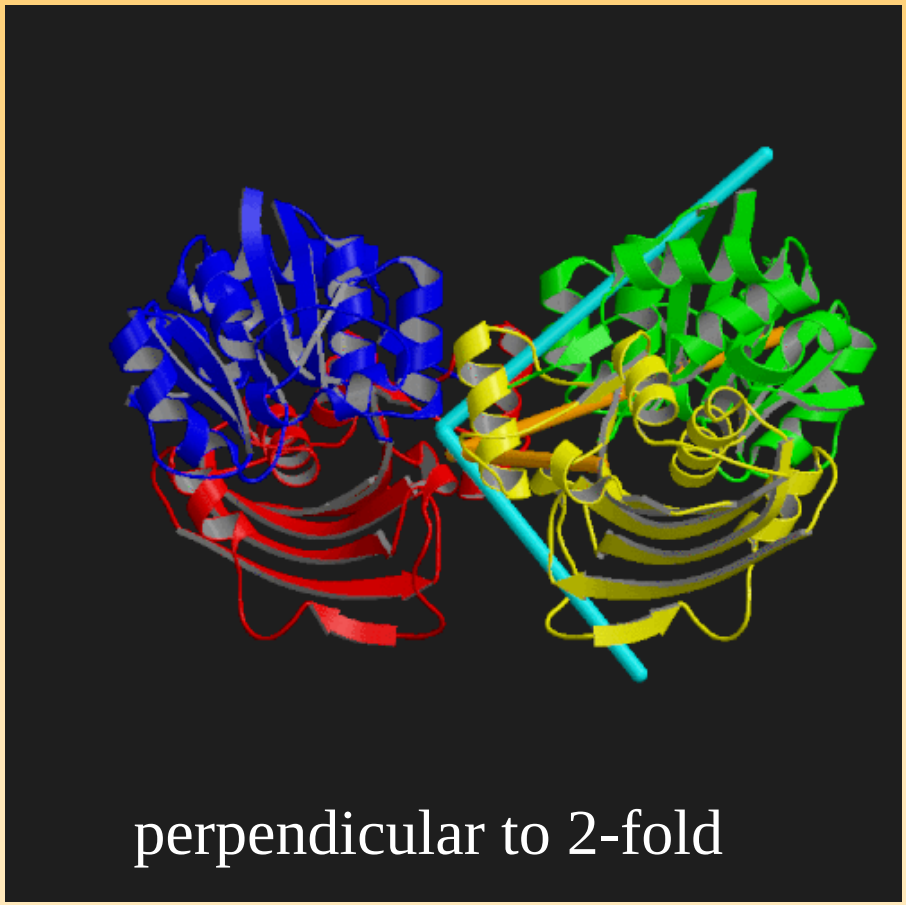
$^-$0.005 Å^2

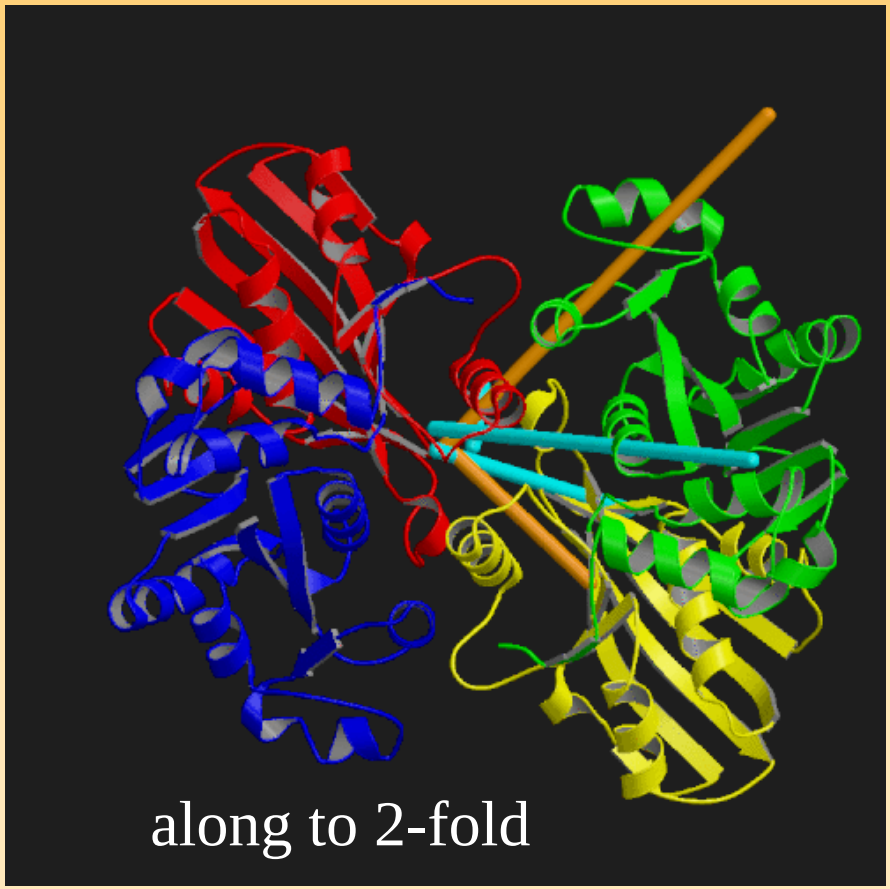
Eigenvalues and pitches of screw axes:

1.403 (°)^2	0.888Å
1.014 (°)^2	1.472Å
0.204 (°)^2	1.214Å

Isupov, M. et al. (1999), *J. Mol. Biol.*, **291**, 651-660

GAPDH: dimer as 1 TLS group





orange = reduced translation, cyan = non-intersecting screw

Atomic parameters from TLSANL

- TLS tensors $\rightarrow U_{TLS}$ for atoms in group (XYZOUT)
- $U_{TLS} \rightarrow B_{TLS}$ and anisotropy A
- Also have individually refined B_{res} from Refmac
- $B_{TOT} = B_{TLS} + B_{res}$ (XYZOUT, keyword ISOOUT)

Displaying derived ADPs

ccp4mg: http://www.ysbl.york.ac.uk/~ccp4mg

Coot: http://www.ysbl.york.ac.uk/~emsley/coot

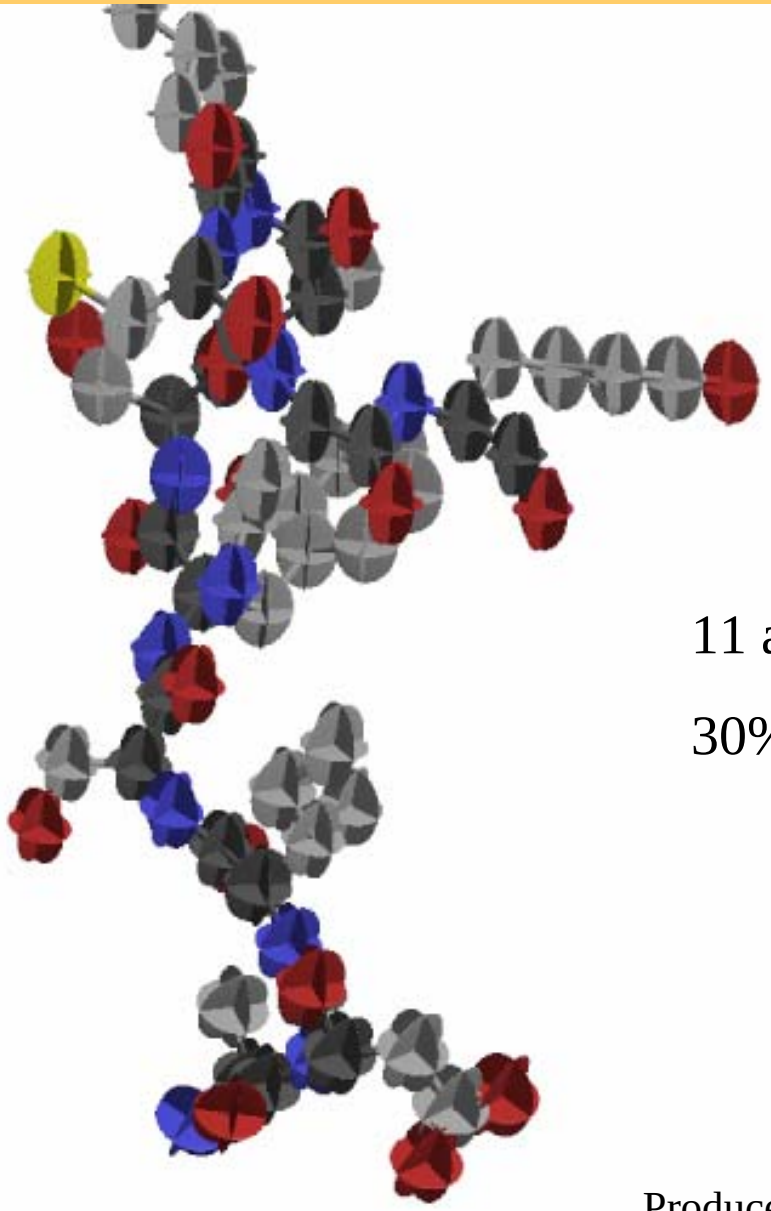
rastep, render from RASTER3D package:
http://www.bmsc.washington.edu/raster3d/raster3d.html

ORTEP: http://www.ornl.gov/ortep/ortep.html

xtalview: http://www.scripps.edu/pub/dem-web/toc.html

Xfit: http://www.ansto.gov.au/natfac/asrp7_xfit.html

povscript: http://people.brandeis.edu/~fenn/povscript



11 a.a. (8% of TLS group)

30% probability level

Produced with Raster3D

GAPDH - TLS-derived aniso-U ellipsoids

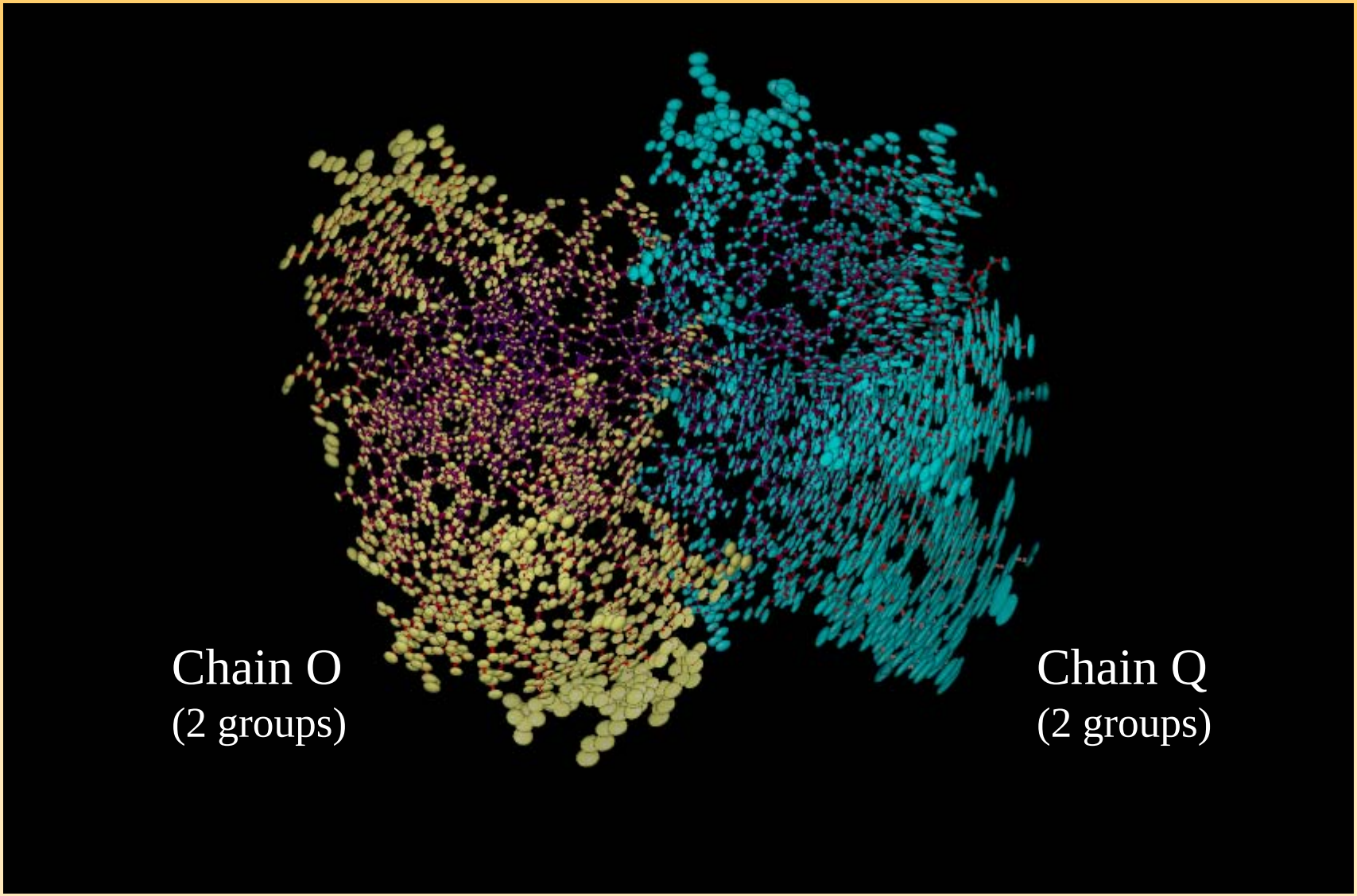


Figure produced by ccp4mg

Contributions to total B factor

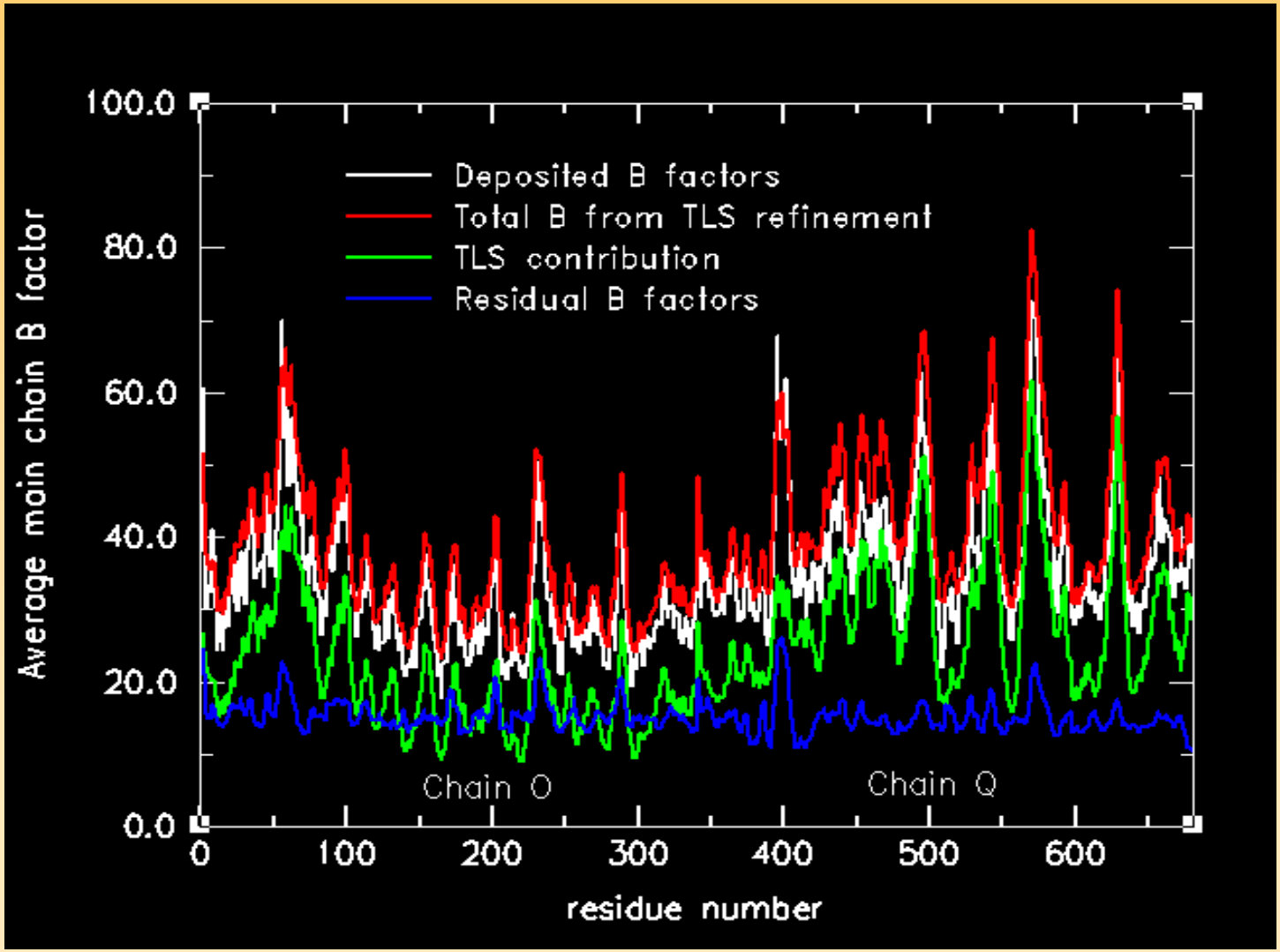
$B_{TOT} = B_{TLS} + B_{res}$

B_{TLS} describes overall displacements of molecules or domains

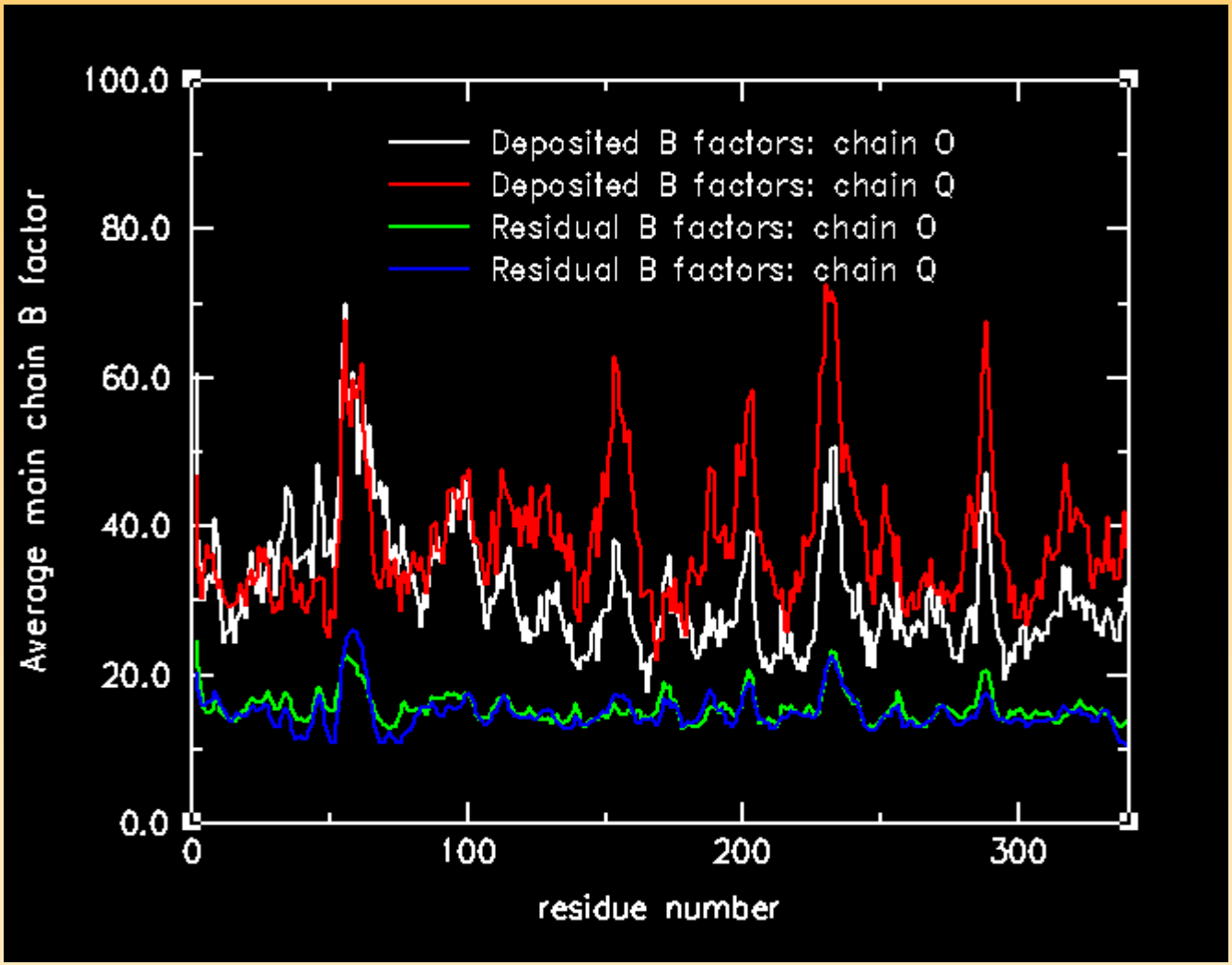
B_{res} describes local displacements, and expected to be similar between molecules

Some ambiguity in the division between BTLS and Bres. It is total that is meaningful. Individual contributions can even be negative!

Example: GAPDH



Example: GAPDH



Example - mannitol dehydrogenase

Hörer *et al.*, *J.Biol.Chem.* **276**, 27555 (2001)

1.5 Å data

3 tetramers in a.s.u.

TLS refinement with 1 group per monomer

Free-R 23.6% $\rightarrow$ 20.9%

Tetramer	***B's before TLS***	***B's after TLS***	***crystal contacts***
ABCD	27.1	13.3	38
EFGH	18.0	13.3	50
IJKL	18.6	13.3	49

Example of problematic refinement

Round 1:
Initialise TLS to zero, and Bs to constant
Later rounds:
Re-refine TLS from zero, but recycle earlier residual Bs

Round	Monomer	av(B_total)	av(B_TLS)	av(B_res)	non-pos
1	B	0.421	-8.479	8.900	6187
	E	8.994	-0.106	9.100	
2	B	12.621	-6.306	18.927	2
	E	21.555	2.198	19.357	
3	B	13.468	-5.320	18.787	0
	E	22.438	3.144	19.294	

Example: octamer in a.s.u., data to 1.8Å, refines to R/Rfree = 0.153/0.199 (round 1), 0.162/0.197 (rounds 2 & 3)

See also Jay Painter's TLS damp patch

Depositing TLS parameters

• TLS parameters are included in the header of the PDB file, and also in the mmCIF-format data harvesting file. Either or both can be used for deposition.
• The residual B factors are the parameters of your model, and should be deposited rather than any derived quantity. The PDB header should include:

```
REMARK   3    ATOM RECORD CONTAINS RESIDUAL B FACTORS ONLY
```

to confirm this.

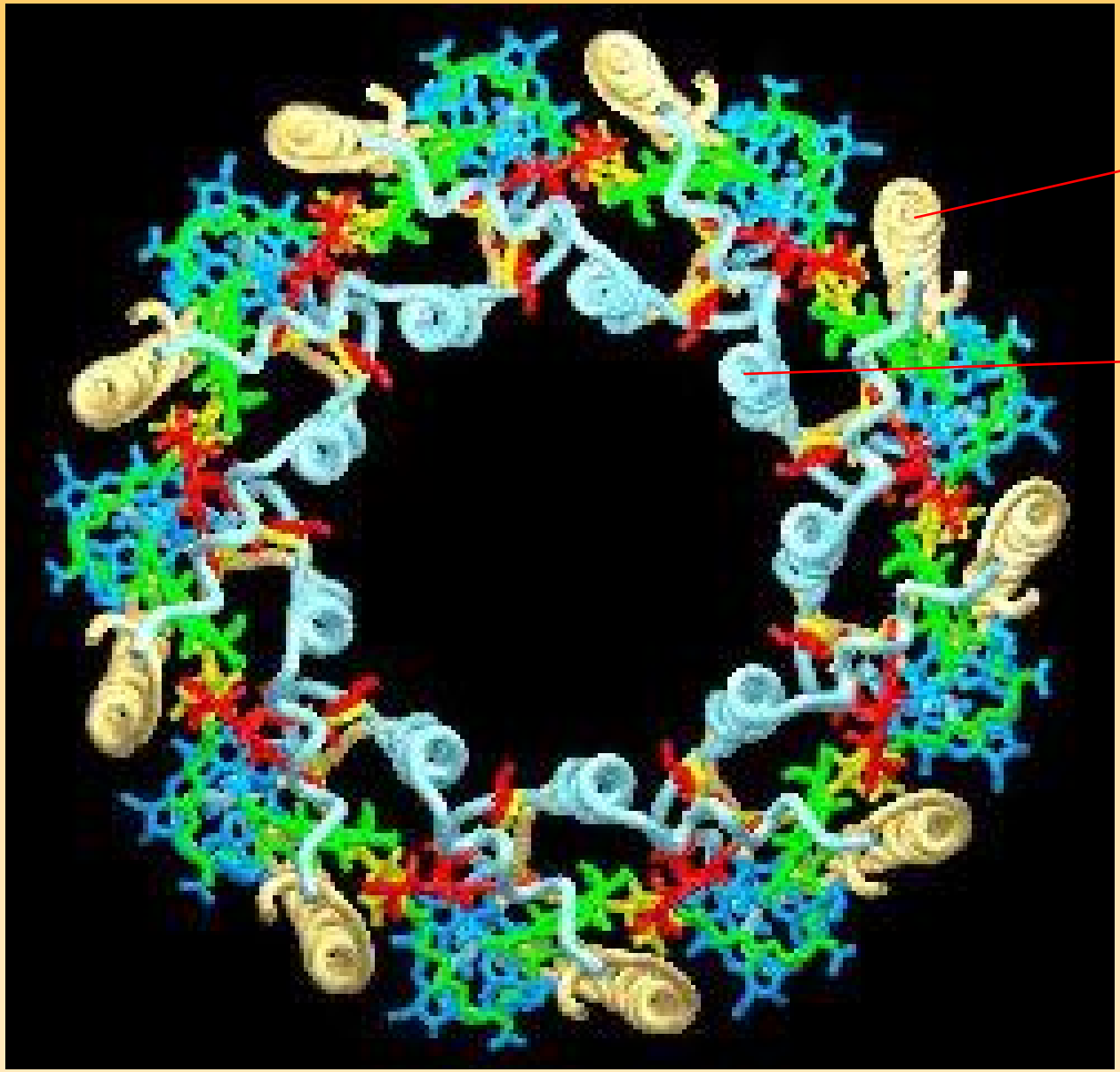
N.B. TLSEXTRACT (new in CCP4 5.0) can extract TLS parameters from the head of a PDB file, and write to TLSOUT.

Large example - light harvesting complex

Complex is nonamer. Each monomer contains:

- α peptide
- β peptide
- 2 x B850 bacteriochlorophyll
- 1 x B800 bacteriochlorophyll
- 2 x carotenoids

Crystallographic asu = 3 monomers


β
α

TLS models

a) 1 group for a.s.u. (20 parameters)

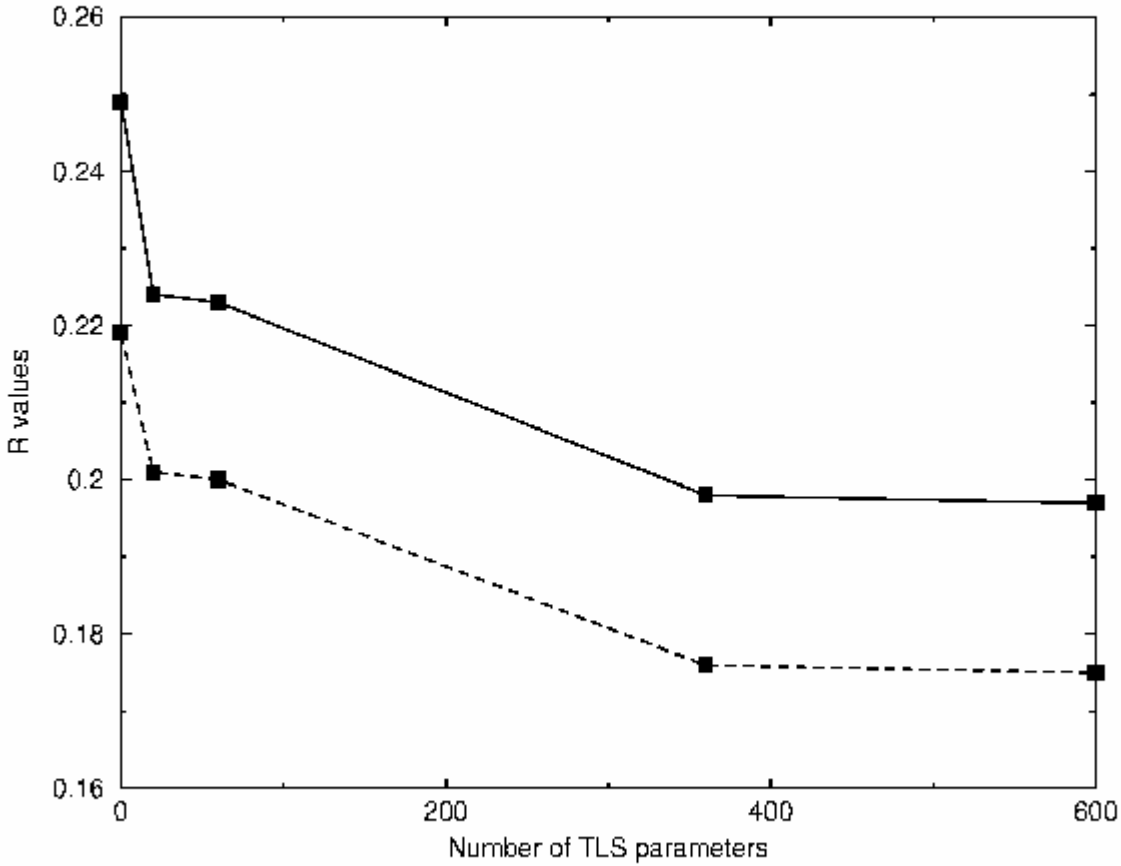
b) 1 group per NCS unit (3 x 20 pars)

c) 1 group per molecule (18 x 20 pars)

d) 3 groups per peptide

+ 1 group per pigment

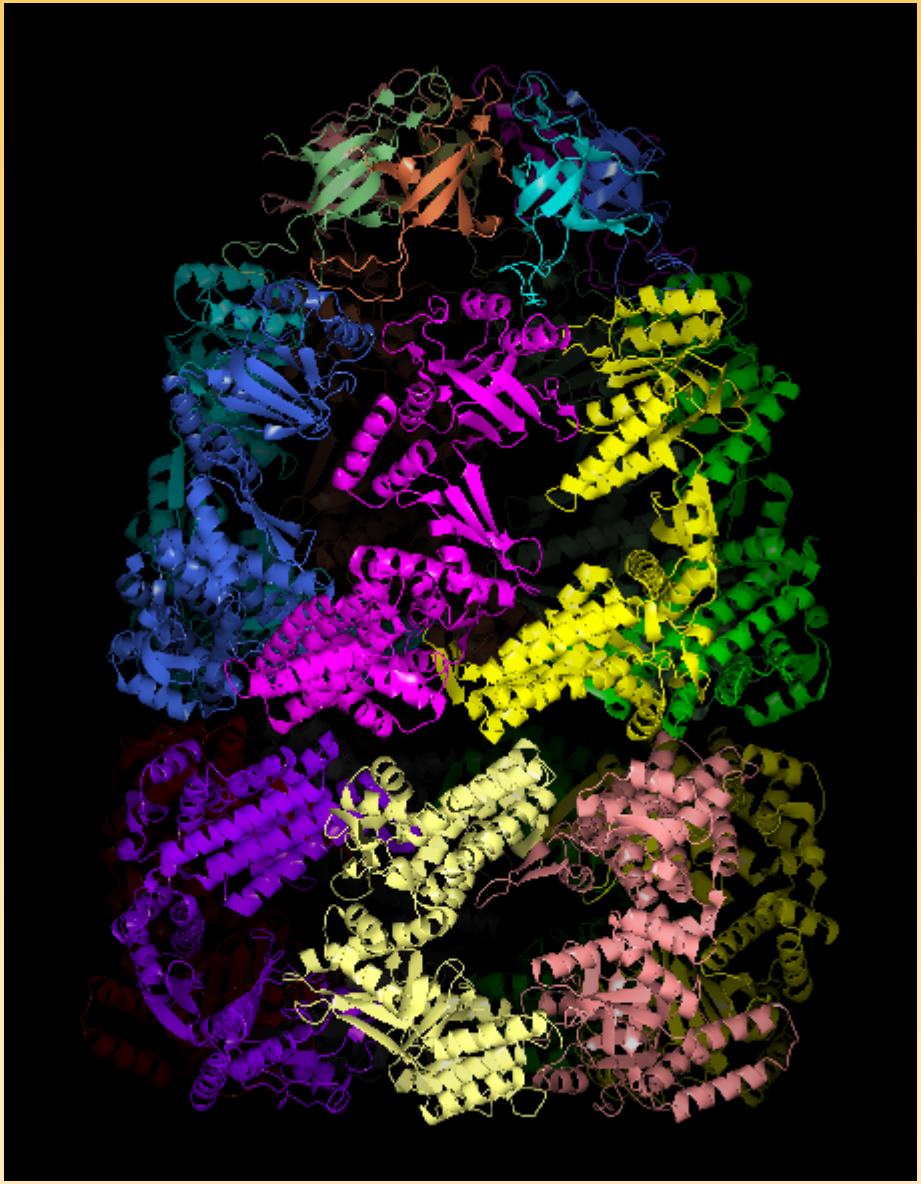
(total 30 x 20 parameters)


0.26
0.24
0.22
0.2
0.18
0.16
R values
0
200
400
600
Number of TLS parameters

Another large example - GroEL

C.Chaudhry et al., JMB, 342, 229-245 (2004)

- TLS refinement of unliganded GroEL, GroEL-ATPγS complex, GroEL-GroES-ADP.AlF_x complex, and GroEL-GroES-ADP complex.
- Best results with one TLS group per GroEL domain “indicating that the inclusion of relative domain displacements significantly improves the quality of the model”
- Show that binding of ligands and GroES causes large changes in dynamic properties
- Can correlate TLS results with changes between states of the machine.



GroES - 7 groups

GroEL - 42 groups

Figure produced by ccp4mg

Problems

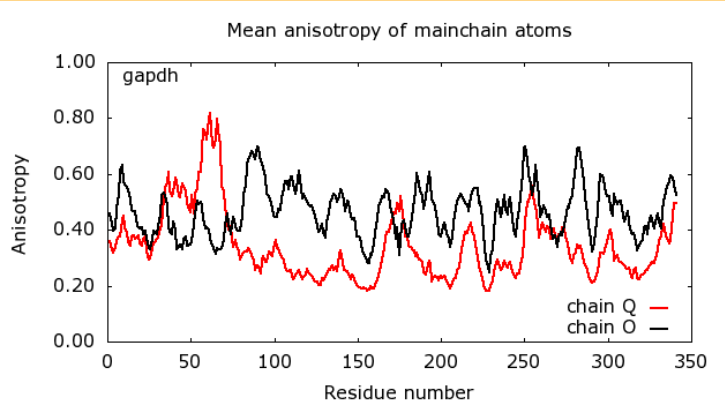
See http://www.ccp4.ac.uk/martyn/tls_research.html

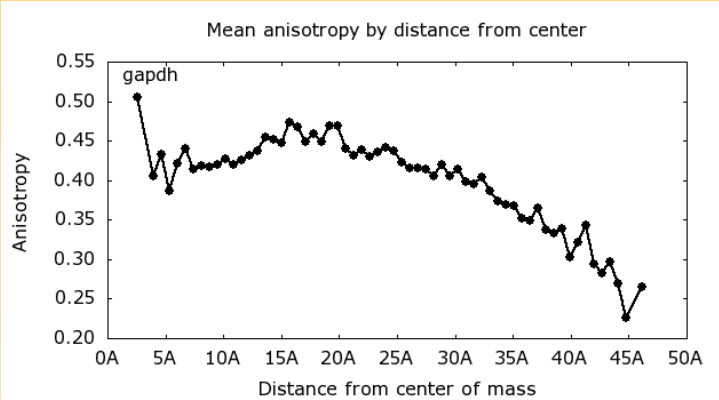
- TLS refinement can be unstable when the model is incomplete.
- Look for atoms or groups of atoms that don't fit quasi-rigid-body assumption
- Try different scaling models

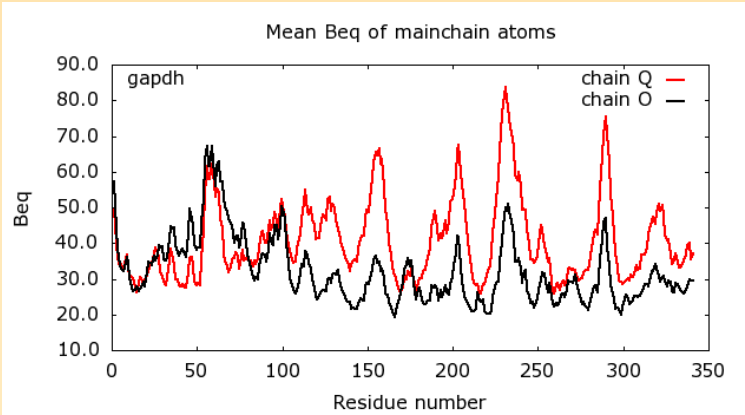
PARVATI server

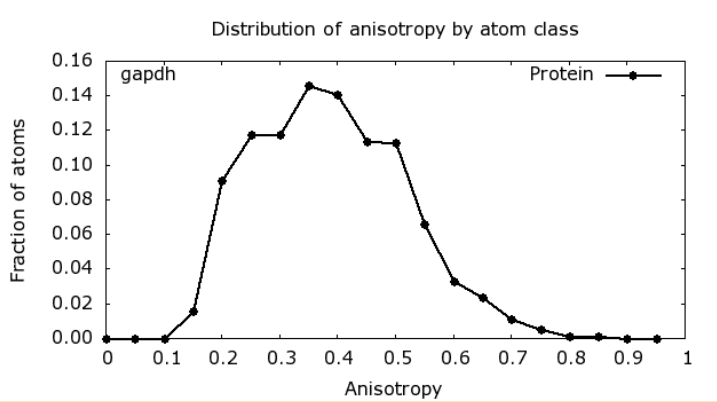
http://www.bmsc.washington.edu/parvati/parvati.html

Protein Anisotropic Refinement Validation and Analysis Tool





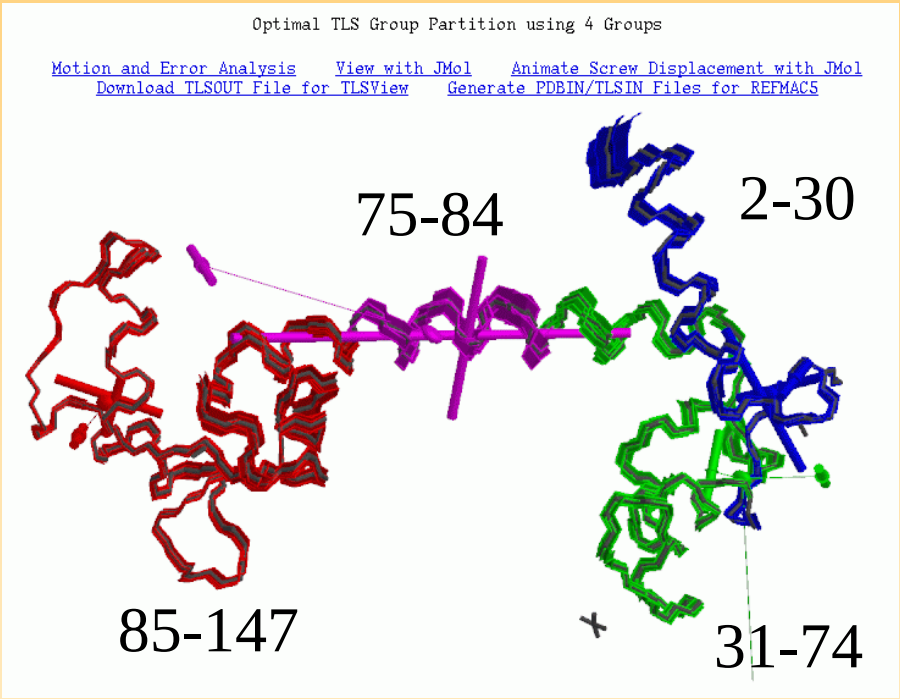




TLS Motion Determination (TLSMD)

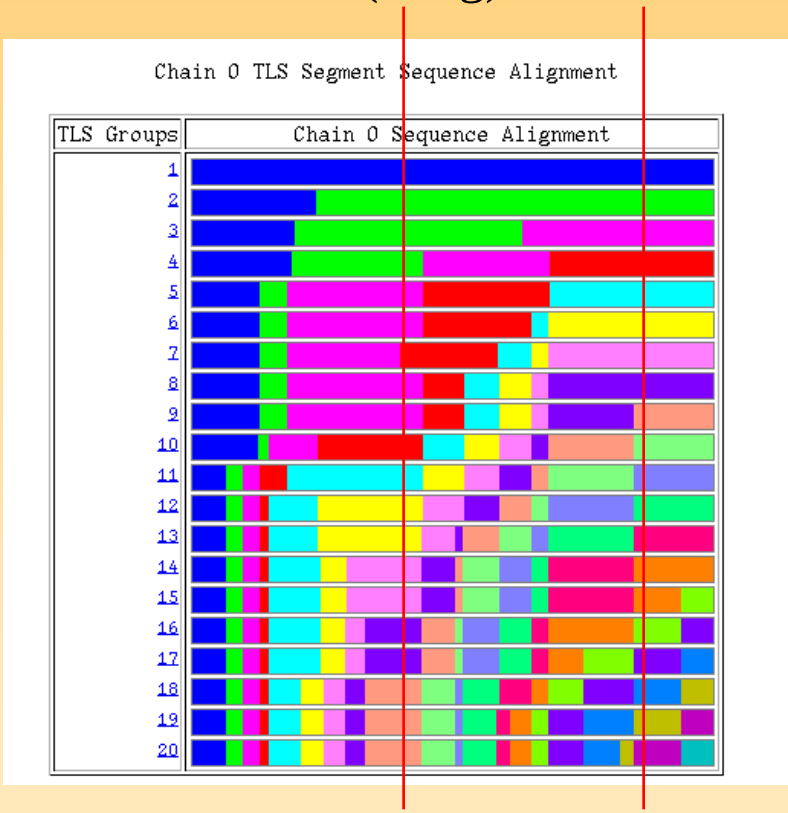
http://skuld.bmsc.washington.edu/~tlsmd
Uses TLSView: J.Painter and E.A.Merritt, *Acta Cryst* **D61**, 465 (2005)

calmodulin (1exr) 1.0Å



N.B. 3 groups combines red and pink into one
cf. M.A.Wilson & A.T.Brunger, *JMB*, **301**, 1237 (2000)
TLS groups 2-70, 71-90, 91-147

GAPDH (1b7g) 2.05Å



NAD-binding domain 1-137 & 303-340
catalytic domain 138-302

Summary TLS

- **TLS parameterization allows to partly take into account anisotropic motions at modest resolution (> 3.5 Å)**

- **TLS refinement might improve refinement statistics of several percent**

- **TLS refinement in *REFMAC*5 is fast and therefore can be used routinely**

- **TLS parameters can be analyzed to extract physical significance**

Acknowledgements

- BBSRC (CCP4 grant)

- Garib Murshudov, York
- Miroslav Papiz, Daresbury
- All users for feedback