User manual to program DISORDER

Author: Alexander Borovinskiy
E-mail: borovin@cmpharm.ucsf.edu
Version of user manual: 1.2
Date: 12/2/2006

Describes program release: 4.0.0
Release date: 12/2/06

1 Introduction

Program DISORDER provides tools for simulation of the diffraction patterns from fibrillar assemblies of biomolecules and optimization of the fibrillar models with respect to experimental diffraction data. A key feature of the program is the ability to simulate diffraction patterns from fibrillar assemblies with orientation disorder.

2 Theory

DISORDER implements the following methods to simulate fiber diffraction pattern from a given model.

2.1 Simulation of fiber diffraction patterns

For the asymmetric unit of atoms repeated on a helix given by their coordinates (r_j,φ_j,z_j) the cylindrically averaged intensity distribution along layer line l is calculated according to Franklin and Klug[1] as

$$I(R,Z=\frac{l}{c}) = \sum_{i,j}\sum_{n} f_i f_j J_n(2\pi R r_i) J_n(2\pi R r_j)\cos[n(\varphi_i-\varphi_j)-\frac{2\pi l}{c}(z_i-z_j)] \quad (1),$$

where f_i are atomic scattering factors, (R, Z) are the cylindrical coordinates in reciprocal space, and c is the repeat distance of the fiber in the Z direction.

The order n of the Bessel functions contributing to the layer line l is subject to the selection rule

$$\frac{n}{PN}+\frac{m}{p}=\frac{l}{c} \quad (2),$$

where P is the pitch of the helix, p is the axial translation per asymmetric unit, N is the order of rotational symmetry of the fiber, and m is an integer. The spacing between layer lines in reciprocal space is equal to $1/c$. The summation of the Bessel functions in equation (1) is performed using the method of Klug *et al.*[2] Solvent-corrected atomic scattering factors f_i were used to calculate the diffraction intensities (1)[3]

$$f_i^{(solv)} = \left(1-k_{solv}\exp\left(-B_{solv}\frac{D^2}{4}\right)\right) f_i^{(vac)} \quad (3),$$

where k_{solv} is a scale factor used to adjust average solvent scattering intensity, and B_{solv} is a large artificial temperature factor applied to account for scattering from the disordered solvent.

For a coherently diffracting crystallite of length L, the distribution of intensity across the layer line is well approximated by the Gaussian form

$$f(\Delta Z) = \exp[-L^2\pi(\Delta Z)^2] \quad (4),$$

where ΔZ is the distance from the center of the layer line. The intensity distribution in the diffraction pattern of a single fiber oriented along the z-axis is then given as

$$I(R,Z) = I\left(R, \frac{l}{c} \pm \Delta Z\right) = I\left(R, \frac{l}{c}\right) f(\Delta Z) \quad (5).$$

The diffracting sample is considered as an assembly of fibrous particles, which are randomly disoriented with respect to the *z*-axis according to a Gaussian distribution. The probability of finding particles in an element of solid angle $d\Omega$ at an angle α to the *z* axis is $N(\alpha)d\Omega/4\pi$, where

$$N(\alpha) = \frac{2}{\alpha_0^2} \exp\left[\frac{-\alpha^2}{2\alpha_0^2}\right] \quad (6)$$

and α_0 is the disorientation parameter. The intensity distribution in the diffraction pattern generated by the assembly of the disoriented fibrous particles is then given by the integral[4]

$$I_{dis}(R,Z) = \int_{\Omega} I(R,Z) N(\alpha) \frac{d\Omega}{4\pi} \quad (7)$$

Following Holmes and Barrington Leigh[4] this integral is calculated in DISORDER as

$$I_{dis}(R,Z) = \frac{I(R,Z)}{4\pi\sqrt{R^2+Z^2}} \int_0^{2\pi} d\varphi \int_0^{\pi} d\gamma \sin\gamma \exp\left[-\frac{1}{2}\left(\frac{\arccos((Z\cos\gamma + R\sin\gamma)/\sqrt{R^2+Z^2})}{\alpha_0}\right)^2\right] \quad (8),$$

where φ,γ are the angles describing the orientation of the single fibrous particle in the sample.

For the comparison with experimental fiber diffraction data intensities (8) are multiplied by isotropic temperature factor $\exp[-B(R^2+Z^2)/2]$, or anisotropic factor $\exp[-(B_R R^2 + B_Z Z^2)/2]$.

2.2 Comparison of simulated and observed diffraction patterns

The fiber diffraction residual was calculated for every model as a measure of the similarity of the simulated pattern to the observed in the form

$$R = \frac{\sum_{R,Z} \left| f\left(\left| kI_{dis}(R,Z) - I_{obs}(R,Z) \right| \right) \right|^2}{\sum_{R,Z} I_{obs}(R,Z)^2} \quad (9),$$

where

$$f(x) = \begin{cases} x, \text{if } x < x_0 \\ x_0, \text{if } x \geq x_0 \end{cases} \quad (10).$$

Here $I_{obs}(x,y)$ is the observed diffraction intensity at the point of reciprocal space with coordinates (R,Z). The factor

$$k = \frac{\sum_{R,Z} I_{dis}(R,Z) I_{obs}(R,Z)}{\sum_{R,Z} I_{dis}(R,Z)^2} \quad (11)$$

was applied to bring the calculated diffraction intensities to the same scale as the observed. It should be noted that the residual (9) is not equivalent to the "traditional" fiber diffraction R-factor[5], since it is calculated by summation across the whole diffraction pattern and is not limited to the layer lines.

2.3 Optimization of the model parameters by simulated annealing

To obtain a better fit of the simulated diffraction patterns to the experimental data the bulk solvent parameters k_{solv} and B_{solv} and isotropic model B-factors were optimized for every model by simulated annealing minimization of the whole pattern residuals. The limits of the parameters variation during the minimization are shown in Table 1.

2.4 Model positioning

The fiber axis was aligned with the direction of the z-axis of the model coordinate system. The models of the asymmetric unit were initially placed into the coordinate system so that their centers of masses were located at the origin and the average direction of the H-bonding coincided with the z-axis. The models were positioned by two rotation transformations followed by translation along the x-axis:

$$\mathbf{r}_i' = \mathbf{d} + \mathbf{T}_\psi \mathbf{T}_\varphi \mathbf{r}_i \qquad (12),$$

where $\mathbf{r}_i$, $\mathbf{r}_i'$ are the coordinates of the model before and after transformation, $\mathbf{T}_\varphi$, $\mathbf{T}_\psi$ are the matrices of rotation about the z- and x-axis, and $\mathbf{d}=(d,0,0)$ is a translation vector.

3 Input/Output files

This section discusses required and optional files used or generated by DISORDER, the file formats understood by the program, tips for preparation of the input files.

3.1 Input files

1. Model coordinates (required)
   User must provide the coordinates of a model asymmetric unit in PDB format. Tip: the model can be prepared using programs O, Insight2, MODELLER

2. Processed experimental image (optional)
3. Experimental image mask (optional)
4. Pixel weights files for scaling of the simulated image and for residual calculations (optional)
5. Data files for calculations of disoriented fiber diffraction patterns (optional)
6. Commands script file (required)

3.2 Output files

1. Simulated diffraction pattern in quadrant or full view
2. Compared simulated and experimental diffraction patterns
3. Experimental diffraction pattern in quadrant or full view
4. Compared simulated and experimental equatorial intensity profiles
5. Compared simulated and experimental meridional radial scans
6. Compared simulated and experimental layer-lines intensities
7. Reports in user-defined format
8. Coordinates of the fibrillar assembly model

4 Other programs used by DISORDER

4.1 Fit2D

4.2 MDL Chime plug-in to Internet Explorer

5 Data flow

This section discusses use of DISORDER in combination with other programs in a context of modeling against fiber diffraction data.

6 Examples of common use

This section describes a number of examples of application of DISORDER to the usual tasks in fiber diffraction simulations. Tobacco Mosaic Virus models and fiber diffraction data[6] were chosen to illustrate the program capabilities, since these constitute a most well known example of structure determination by fiber diffraction. The experimental diffraction data were kindly provided by Dr. Gerald Stubbs. The models of TMV were downloaded from Protein Data Bank (PDB Ids: `1vtm and 2tmv)`.

The model of Sup35 heptamer peptide[7] GNNQQNY (PDB ID: 1yjp) was used in Example 13 to illustrate the representation of the fibrillar model as a small crystallite.

The commands that are important part for a particular example are highlighted in bold font.

6.1 Calculate layer lines for a given model

The layer lines are calculated using atomic scattering factors in vacuum and stored in a data file.

```
REGIME DISPLAY
RSIZE 250
ZSIZE 250
PIXEL_RESOLUTION 0.001

ROTATIONAL_SYMMETRY 1
UNITS 49
TURNS_X_PROTOFILAMENTS 3
AXIAL_TRANSLATION 1.40816
MAX_NUM_LINES 24
MAX_BESSEL_ORDER 100
READ_MODEL ./PDB/1vtm.pdb
DMAX 250
REPORT_FILE ./Examples/Output_reports/tmv-u2_lines.log
WRITE_LAYER_LINES ./Examples/Output_reports/tmv-u2_lines.dat
```

6.2 Simulate oriented diffraction pattern

The layer-lines are calculated, stored as an image in SMV format and displayed in a program FIT2D.

```
REGIME DISPLAY
RSIZE 250
ZSIZE 250
PIXEL_RESOLUTION 0.001
ROTATIONAL_SYMMETRY 1
UNITS 49
TURNS_X_PROTOFILAMENTS 3
AXIAL_TRANSLATION 1.40816
MAX_NUM_LINES 24
MAX_BESSEL_ORDER 100
READ_MODEL ./PDB/1vtm.pdb
REPORT_FILE ./Examples/Output_reports/tmv-u2_lines.log
```

```
BACKSTOP 10
DMAX 250
WRITE_SIMULATED_PATTERN ./Examples/Output_images/tmv-u2_sim.smv
SHOW_COMPARISON ON
```

6.3 Calculate layer-lines for a given model and compare them with experimental layer-lines intensities

Experimental layer-lines are read from a data file, the compared simulated and experimental lines are stored in a data file.

```
REGIME DISPLAY
RSIZE 250
ZSIZE 250
PIXEL_RESOLUTION 0.001

ROTATIONAL_SYMMETRY 1
UNITS 49
TURNS_X_PROTOFILAMENTS 3
AXIAL_TRANSLATION 1.40816
MAX_NUM_LINES 24
MAX_BESSEL_ORDER 100

READ_MODEL ./PDB/2tmv.pdb
READ_LAYER_LINES .\IMAGES\TMV_lines_obs.dat
AUTO_MASK ON

REPORT_FILE ./Examples/Output_reports/tmv.log

BACKSTOP 5
DMAX 250
WRITE_COMPARED_SLICES 69.0 ./Examples/Output_reports/tmv-cmprd_lines.dat
```

6.4 Simulate disoriented diffraction pattern

Orientation disorder applied in calculation of the diffraction pattern. The parameter `ALPHA0` specifies the degree of the disorientation in the model.

```
REGIME DISPLAY
RSIZE 250
ZSIZE 250
PIXEL_RESOLUTION 0.001
ROTATIONAL_SYMMETRY 1
UNITS 49
TURNS_X_PROTOFILAMENTS 3
AXIAL_TRANSLATION 1.40816
MAX_NUM_LINES 24
MAX_BESSEL_ORDER 100

READ_MODEL ./PDB/1vtm.pdb
REPORT_FILE ./Examples/Output_reports/tmv.log
```

```
DISORDER ON
ALPHA0 2.6
COHERENCE_LENGTH 250.

BACKSTOP 30
DMAX 245
WRITE_SIMULATED_PATTERN ./Examples/Output_images/tmv-u2_sim.smv
SHOW_COMPARISON ON
```

6.5 Simulate disoriented diffraction pattern and compare it with experimental image

```
REGIME DISPLAY
RSIZE 250
ZSIZE 250
PIXEL_RESOLUTION 0.002
ROTATIONAL_SYMMETRY 1
UNITS 49
TURNS_X_PROTOFILAMENTS 3
AXIAL_TRANSLATION 1.40816
MAX_NUM_LINES 24
MAX_BESSEL_ORDER 100

READ_MODEL ./PDB/1vtm.pdb
REPORT_FILE ./Examples/Output_reports/tmv.log
READ_IMAGE .\IMAGES\F03000.206
BSL_FRAME 0
AUTO_MASK ON

DISORDER ON
ALPHA0 2.6
COHERENCE_LENGTH 350.
TEMPERATURE_FACTOR ISOTROPIC 2.0 #B_iso
SOLVENT_CONTRAST 0.99 200.0 #ksol bsol

BACKSTOP 20
DMAX 185
WRITE_COMPARED_PATTERNS ./Examples/Output_images/tmv_u2_cmprd.smv
SHOW_COMPARISON ON
```

6.6 Position model and simulate diffraction pattern

The next example requires some introductory notes. The TMV(U2) model stored in the file 1vtm.pdb has a center of mass located at (55.221; 13.932; 34.345). To illustrate application of the commands for model positioning a different model file was prepared (1vtm_cm.pdb). That model was obtained from 1vtm.pdb by rotating it −14° around Z-axis (so the center of mass (CM) moves to the XZ plane) and translating it, so the CM positioned at the origin of the coordinate system.

The command script listed below reads model 1vtm_cm.pdb and translates it 56.9513 Å along X-axis. Then the diffraction pattern is calculated for the positioned model and compared to the experimental image. Please consult "Model positioning" in "Theory" section to make sure that such transformed 1vtm_cm.pdb model is equivalent to 1vtm.pdb

```
REGIME DISPLAY
RSIZE 250
ZSIZE 250
PIXEL_RESOLUTION 0.002

ROTATIONAL_SYMMETRY 1
UNITS 49
TURNS_X_PROTOFILAMENTS 3
AXIAL_TRANSLATION 1.40816
MAX_NUM_LINES 24
MAX_BESSEL_ORDER 50

READ_MODEL ./PDB/1vtm_cm.pdb
MOVE_MODEL 56.9513 0.0 0.0 0.0

READ_IMAGE .\IMAGES\F03000.206
BSL_FRAME 0
AUTO_MASK ON

REPORT_FILE ./Examples/Output_reports/tmv.log
REPORT_COLUMNS ALPHA0 K_SOL B_SOL B_ISO

DISORDER ON
ALPHA0 2.6

COHERENCE_LENGTH 350.
TEMPERATURE_FACTOR ISOTROPIC 2.0
SOLVENT_CONTRAST 0.99 200.0

BACKSTOP 20
DMAX 185
WRITE_COMPARED_PATTERNS ./Examples/Output_images/tmv_u2_cmprd.smv
SHOW_COMPARISON ON
```

The following script produces the same result:

```
REGIME BROWSE_MODEL_SPACE
RSIZE 250
ZSIZE 250
PIXEL_RESOLUTION 0.002
```

```
ROTATIONAL_SYMMETRY 1
UNITS 49
TURNS_X_PROTOFILAMENTS 3
AXIAL_TRANSLATION 1.40816
MAX_NUM_LINES 24
MAX_BESSEL_ORDER 50

READ_MODEL ./PDB/1vtm_cm.pdb
MOVE_MODEL RADIUS 56.9513 56.9513 0.1

READ_IMAGE .\IMAGES\F03000.206
BSL_FRAME 0
AUTO_MASK ON

REPORT_FILE ./Examples/Output_reports/tmv.log
REPORT_COLUMNS ALPHA0 K_SOL B_SOL B_ISO

DISORDER ON
ALPHA0 2.6

COHERENCE_LENGTH 350.
TEMPERATURE_FACTOR ISOTROPIC 2.0
SOLVENT_CONTRAST 0.99 200.0

BACKSTOP 20
DMAX 185
WRITE_COMPARED_PATTERNS ./Examples/Output_images/tmv_u2_cmprd.smv
SHOW_COMPARISON ON
```

6.7 Scale the simulated image intensity and compare it with experimental image

By default, all the pixels of the diffraction pattern where mask=1 are used to scale the intensities of the simulated diffraction pattern (see equation (11) in "Theory"). It is possible to scale the intensities of the simulated image the level of the experimental one using only a part of the diffraction pattern, that presents the most interest to the user. In this example the simulated image is scaled by a sector of the diffraction pattern that includes pixels, which have resolution in the range between 0.1 $Å^{-1}$ and 0.25 $Å^{-1}$. To do that the scaling weights outside this region are set to 0.

```
REGIME DISPLAY
RSIZE 250
ZSIZE 250
PIXEL_RESOLUTION 0.002
ROTATIONAL_SYMMETRY 1
UNITS 49
TURNS_X_PROTOFILAMENTS 3
AXIAL_TRANSLATION 1.40816
MAX_NUM_LINES 24
MAX_BESSEL_ORDER 50

READ_MODEL ./PDB/1vtm.pdb
```

```
READ_IMAGE .\IMAGES\F03000.206
BSL_FRAME 0
AUTO_MASK ON
SET_SCALE_WEIGHTS SECTOR 5 0 0 50 0.0
SET_SCALE_WEIGHTS SECTOR 125 0 0 185 0.0
REPORT_FILE ./Examples/Output_reports/tmv.log

DISORDER ON
ALPHA0 2.6
COHERENCE_LENGTH 350.
TEMPERATURE_FACTOR ISOTROPIC 2.0
SOLVENT_CONTRAST 0.99 200.0
BACKSTOP 20
DMAX 185
WRITE_COMPARED_PATTERNS ./Examples/Output_images/tmv_u2_cmprd.smv
SHOW_COMPARISON ON
```

6.8 Subtract circularly symmetric or flat background from experimental image

Example will be described in future versions. Please, refer to CIRC_BACKGROUND command description.

6.9 Optimize solvent and model parameters

As a general rule, simulation of a diffraction pattern with vacuum atomic scattering factors produces poor fit to the experimental data, since the intensities in the center of the simulated image overestimated. Usually, atomic scattering factors are corrected to account for the diffraction of bulk solvent. Several methods implemented in DISORDER to improve the fit of the simulated diffraction patterns by optimization of the solvent contrast parameters and uniform model B-factors.

It was demonstrated in examples 5-7 how to set up model isotropic B-factor and solvent parameters manually.

The solvent parameters and the uniform model isotropic B-factor can be optimized using simulated annealing method:

```
REGIME DISPLAY
RSIZE 250
ZSIZE 250
PIXEL_RESOLUTION 0.002
ROTATIONAL_SYMMETRY 1
UNITS 49
TURNS_X_PROTOFILAMENTS 3
```

```
AXIAL_TRANSLATION 1.40816
MAX_NUM_LINES 24
MAX_BESSEL_ORDER 50

READ_MODEL ./PDB/1vtm.pdb
READ_IMAGE .\IMAGES\R03000.208
BSL_FRAME 0
AUTO_MASK ON

SET_SCALE_WEIGHTS SECTOR 5 0 0 50 0.0
SET_SCALE_WEIGHTS SECTOR 125 0 0 185 0.0

REPORT_FILE ./Examples/Output_reports/tmv.log
REPORT_COLUMNS ALPHA0 K_SOL B_SOL B_ISO

DISORDER ON
ALPHA0 2.6
COHERENCE_LENGTH 350.

NUM_ITERATIONS 3000
ANNEALING_TEMPERATURE 0.5
RANDOM_START ON
SEARCH_B_FACTORS_ISO 0.8 0.999 0.01 400 1500 50 2 100 2

BACKSTOP 50
DMAX 172
WRITE_COMPARED_PATTERNS ./Examples/Output_images/tmv_u2_cmprd.smv
SHOW_COMPARISON ON
```

Alternatively, these parameters can be optimized using a grid-search procedure:

```
REGIME DISPLAY
RSIZE 250
ZSIZE 250
PIXEL_RESOLUTION 0.002
ROTATIONAL_SYMMETRY 1
UNITS 49
TURNS_X_PROTOFILAMENTS 3
AXIAL_TRANSLATION 1.40816
MAX_NUM_LINES 24
MAX_BESSEL_ORDER 50

READ_MODEL ./PDB/1vtm.pdb
READ_IMAGE .\IMAGES\R03000.208
BSL_FRAME 0
AUTO_MASK ON
SET_SCALE_WEIGHTS SECTOR 5 0  0 50 0.0
SET_SCALE_WEIGHTS SECTOR 125 0  0 185 0.0

REPORT_FILE ./Examples/Output_reports/tmv.log
REPORT_COLUMNS ALPHA0 K_SOL B_SOL B_ISO
```

```
DISORDER ON
ALPHA0 2.6
COHERENCE_LENGTH 350.
```

BROWSE_B_FACTORS_ISO 0.96 0.99 0.01 200 500 50 2 20 1

```
BACKSTOP 50
DMAX 172
WRITE_COMPARED_PATTERNS ./Examples/Output_images/tmv_u2_cmprd.smv
SHOW_COMPARISON ON
```

Sometimes, a better fit of the simulated diffraction pattern to the experimental data can be obtained, if anisotropic uniform model B-factors are used (in R and Z directions in reciprocal space). However, the execution of this command is time consuming, so try to avoid large searches. Here is an example of a grid-search optimization with anisotropic model B-factors:

```
REGIME DISPLAY
RSIZE 250
ZSIZE 250
PIXEL_RESOLUTION 0.002
ROTATIONAL_SYMMETRY 1
UNITS 49
TURNS_X_PROTOFILAMENTS 3
AXIAL_TRANSLATION 1.40816
MAX_NUM_LINES 24
MAX_BESSEL_ORDER 50

READ_MODEL ./PDB/1vtm.pdb
READ_IMAGE .\IMAGES\R03000.208
BSL_FRAME 0
AUTO_MASK ON

SET_SCALE_WEIGHTS SECTOR 5 0  0 50 0.0
SET_SCALE_WEIGHTS SECTOR 125 0  0 185 0.0

REPORT_FILE ./Examples/Output_reports/tmv.log
REPORT_COLUMNS ALPHA0 K_SOL B_SOL B_ISO RESIDUAL

DISORDER ON
ALPHA0 2.6
COHERENCE_LENGTH 350.
```

BROWSE_B_FACTORS_ANISO 0.96 0.99 0.01 200 500 50 2 20 1 2 20 1

```
BACKSTOP 50
```

```
DMAX 172
WRITE_COMPARED_PATTERNS ./Examples/Output_images/tmv_u2_cmprd.smv
SHOW_COMPARISON ON
```

6.10 Quantitative comparison of the simulated and experimental images

The least-squares residuals for the whole diffraction pattern and user-defined regions can be calculated and stored into the report file. In this example the command IO_THRESHOLD is used to exclude the pixels of low intensities from the calculation of the residuals. That helps to improve the sensitivity of the residual as a measure of the simulated patterns fit. The whole pattern residual and region residual for a rectangular area corresponding to layer-line 6 are calculated in this example:

```
REGIME DISPLAY
RSIZE 250
ZSIZE 250
PIXEL_RESOLUTION 0.002
ROTATIONAL_SYMMETRY 1
UNITS 49
TURNS_X_PROTOFILAMENTS 3
AXIAL_TRANSLATION 1.40816
MAX_NUM_LINES 24
MAX_BESSEL_ORDER 50

READ_MODEL ./PDB/1vtm.pdb
READ_IMAGE .\IMAGES\R03000.208
BSL_FRAME 0
AUTO_MASK ON

SET_SCALE_WEIGHTS SECTOR 5 0  0 50 0.0
SET_SCALE_WEIGHTS SECTOR 125 0  0 185 0.0

REPORT_FILE ./Examples/Output_reports/tmv.log
REPORT_COLUMNS ALPHA0 K_SOL B_SOL R_REGIONS RESIDUAL

DISORDER ON
ALPHA0 2.6
COHERENCE_LENGTH 350.
TEMPERATURE_FACTOR ISOTROPIC 2.0
SOLVENT_CONTRAST 0.97 200.0
BACKSTOP 50
DMAX 172

CALCULATE_R WHOLE_PATTERN
CALCULATE_R REGION RECTANGLE 22 41 170 49
IO_THRESHOLD 50
```

```
WRITE_COMPARED_PATTERNS ./Examples/Output_images/tmv_u2_cmprd.smv
SHOW_COMPARISON ON
```

6.11 Optimize model orientation by grid-search

REGIME BROWSE_MODEL_SPACE
RSIZE 250
ZSIZE 250
PIXEL_RESOLUTION 0.002
ROTATIONAL_SYMMETRY 1
UNITS 49
TURNS_X_PROTOFILAMENTS 3
AXIAL_TRANSLATION 1.40816
MAX_NUM_LINES 24
MAX_BESSEL_ORDER 100

READ_MODEL ./PDB/1vtm_cm.pdb

MOVE_MODEL RADIUS 50.0 60. 0.5
MOVE_MODEL ALPHA 0 355 5
CHECK_MODEL ON
CLASH_DISTANCE 3.5# dist. of close contact between backbone atoms
CONTACT_CHAINS –17 -16 -15 -1 1 15 16 17# for CHECK_MODEL

READ_IMAGE .\IMAGES\R03000.208
BSL_FRAME 0
AUTO_MASK ON

SET_SCALE_WEIGHTS SECTOR 5 0 0 50 0.0
SET_SCALE_WEIGHTS SECTOR 125 0 0 185 0.0

REPORT_FILE ./Examples/Output_reports/tmv_orientation.log
REPORT_COLUMNS ALPHA0 K_SOL B_SOL R_REGIONS RESIDUAL

DISORDER ON
ALPHA0 2.6
COHERENCE_LENGTH 350.
TEMPERATURE_FACTOR ISOTROPIC 2.0
SOLVENT_CONTRAST 0.97 200.0
BACKSTOP 50
DMAX 172

CALCULATE_R WHOLE_PATTERN
CALCULATE_R REGION RECTANGLE 22 41 170 49
IO_THRESHOLD 50

6.12 Optimize model by simulated annealing

This example will be included in a future version.

6.13 Fiber diffraction calculations for a small crystallite model

```
#GNNQQNY
REGIME DISPLAY
RSIZE 250
ZSIZE 250
PIXEL_RESOLUTION 0.001  # Angstrom^-1

#parameters of the model used in selection rule:
#l=k*(c*N/P)+m*(c/p) <=> l=k*TURNS_X_PROTOFILAMENTS + m*UNITS
#Bessel functions of orders J_(k*N) are used in calculation of
#intensities

ROTATIONAL_SYMMETRY 1 #number of protofilaments N
UNITS 168 #u=c/p, units in period
TURNS_X_PROTOFILAMENTS 167 #t=c*N/P, P is the pitch of the helix
AXIAL_TRANSLATION 4.8 #p
MAX_NUM_LINES 250
MAX_BESSEL_ORDER 50

#parameters for reading input PDB file
NO_H_NO_HOH ON
READ_MODEL ./PDB/GNNQQNY_ab.pdb # 7mer chains A&B

#description of the lattice in terms of Eisenberg's model unit cell
BUILD_CRYSTALITE 22. 23.5  4.87 90 90 72.92 4 3 1 #21.94 23.48 # P111

REPORT_FILE ./Examples/Output_reports/7mer.log # log file
REPORT_COLUMNS ALPHA0 K_SOL B_SOL B_ISO# R_REGIONS RESIDUAL #

#parameters for simulation of fiber diffraction pattern
DISORDER ON #OFF #
ALPHA0 8.5 #degree of disorientation, deg


#parameters for diffraction calculation:
COHERENCE_LENGTH 500. # Angstroms

#parameters used in solvent contrast method
TEMPERATURE_FACTOR ISOTROPIC 2.0 #B_iso
SOLVENT_CONTRAST 0.97 200.0 #ksol bsol
BACKSTOP 10
DMAX 250
WRITE_SIMULATED_PATTERN ./Examples/Output_images/7mer_sim.smv
SHOW_COMPARISON ON
```

7 List of commands used in DISORDER scripting language

Symbol '#' is used to comment part of the line that follows it in a script file.

7.1 Fiber diffraction calculation parameters

REGIME <regime keyword> selects the regime for the current program run. Must be the first command in the script file.

Several keywords are accepted:

DISPLAY allows user to simulate the diffraction pattern for a single orientation of a model provided by user and display the pattern with a program 'fit2d'

BROWSE_MODEL_SPACE allows user to simulate diffraction patterns for a range orientations of a model defined by a command MOVE_MODEL and compare them to the experimental diffraction pattern by means of calculation of a least squares residuals;

SEARCH_MODEL_SPACE allows user to search for optimal values of helical symmetry, orientation of a model by simulated annealing algorithm. (*The code that implements this regime is under construction.*);

BROWSE_UNIT_CELL in this regime an asymmetric unit of a model is constructed as a small crystallite. This regime allows user to calculate diffraction patterns for a range of unit cell parameters of the crystallite and compare them to the experimental diffraction pattern by means of calculation of a least squares residuals. See commands CELL_* for more details;

The following commands define the helical symmetry of the model:

ROTATIONAL_SYMMETRY <int> defines rotational symmetry of the model fibril.

UNITS <int> the command defines number of asymmetric units in one period of the model fibril.

TURNS_X_PROTOFILAMENTS <int> if **ROTATIONAL_SYMMETRY**=1, the parameter defines the number of helix turns in one period of the fiber. If **ROTATIONAL_SYMMETRY**>1, the parameter defines a number of helix turns in the segment of the fibril that has length (c*N), where c is a period of the fibril and N is the order of its rotational symmetry. In other words, the selection rule: $l=k*(c*N/P)+m*(c/p)$ is equivalent to $l=k$***TURNS_X_PROTOFILAMENTS** + m***UNITS.** Bessel functions of orders J_{k*N} are used in calculation of intensities.

AXIAL_TRANSLATION <float> defines the axial translation of symmetrically related units on the helix in (Å).

MAX_BESSEL_ORDER <int> defines maximal order of Bessel functions used in calculation of the Furier transforms.

MAX_NUM_LINES <int> allows user to limit the number of calculated layer lines. If this parameter is set to high values, only the lines visible at the current values of **DMAX** and **PIXEL_RESOLUTION** are calculated.

RSIZE <101/250>
ZSIZE <101/250> These parameters define the size of the calculated diffraction patterns in pixels. Currently the disordered diffraction patterns can be calculated only for two sizes of the image either 101x101, or 250x250. Use smaller image size for time consuming calculations, and larger image size, if the resolution of fine details of the diffraction pattern is desired. It is possible to use different image sizes, if necessary, without modification of the program code. However, the disorientation integrals must be recalculated for the new image size and this task may take several days.
The strange image size 101x101 is inherited from the program *Cerius*^{2}. The output from *Cerius*^{2} was originally used to test the developed code.

PIXEL_RESOLUTION <float> pixel resolution, $Å^{-1}$

BACKSTOP <int> defines the radius of the beamstop in pixels.

DMAX <int> defines the maximum resolution of the image in pixels.

DISORDER <ON/OFF> The command defines whether to calculate oriented (OFF), or disoriented (ON) diffraction pattern. If set ON, define the disorientation parameter **ALPHA0**.

ALPHA0 The command defines disorientation parameter. Several variants of syntax are accepted:

ALPHA0 <*alpha0*, (deg), float> The disorientation parameter is fixed to *alpha0* degrees. This syntax is accepted in the regime **DISPLAY**.

ALPHA0 <*alpha0_min*, (deg), float> <*alpha0_max*, (deg), float> <*step*, (deg), float>
The disorientation parameter is varied in the range between *alpha0_min* and *alpha0_max* degrees with increment *step*. If the user sets *alpha0* to the values for which the disorientation integrals were not precalculated, the program displays warning message. This syntax is accepted in the regime **BROWSE_MODEL_SPACE**.

COHERENCE_LENGTH The command defines the coherence length of a model. The layer lines are assumed to have Gaussian cross-shape, and the parameter *length* defines the width of the corresponding Gaussian distribution.
COHERENCE_LENGTH <*length*, (Å), float>

The coherence length is fixed to *length* Å. This syntax is accepted in the regime **DISPLAY**.

COHERENCE_LENGTH <*length_min*, (Å), float> <*length_max*, (Å), float> <*step*, (Å), float>
This syntax is accepted in the regime **BROWSE_MODEL_SPACE**.

FIBERS_PACKING <*number of fibers*, int> <*period*, (Å), float>
This command applies the correction to the equatorial layer line for the hexagonal packing of the fibrils. The first parameter defines the number of the fibrils in a bundle and the second is the period of the hexagonal lattice.

TEMPERATURE_FACTOR The command defines the uniform B-factors applied to all the atoms of the model. There are two variants of the command syntax accepted:

TEMPERATURE_FACTOR ISOTROPIC <B_{iso}, (Å^2), float>
Uniform isotropic B-factor B_{iso} is applied in diffraction calculations.
TEMPERATURE_FACTOR ANISOTROPIC <B_R, (Å^2), float> <B_Z, (Å^2), float>
Anisotropic B-factors B_R, B_Z are applied in radial and meridional directions in diffraction calculations.

The following commands define the regime of calculation of atomic scattering factors:

VACUUM The atomic scattering factors in vacuum are applied in calculation of the diffraction patterns.

SOLVENT_CONTRAST <k_{sol}, float> <B_{sol}, float>
Solvent-corrected atomic scattering factors are used to calculate the diffraction intensities:

$$f_i^{(solv)}(D) = \left(1 - k_{solv} \exp\left(-B_{solv}\frac{D^2}{4}\right)\right) f_i^{(vac)}(D)$$

BROWSE_B_FACTORS_ISO <k_{sol}_min, float> <k_{sol}_max, float> <k_{sol}_step, float> <B_{sol}_min, float> <B_{sol}_max, float> <B_{sol}_step, float> <B_{iso}_min, float> <B_{iso}_max, float> <B_{iso}_step, float>
The parameters k_{sol}, B_{sol} and B_{iso} are optimized to obtain lower whole pattern least squares residuals for every model considered. This command currently works in the regimes DISPLAY and BROWSE_MODEL_SPACE.

BROWSE_B_FACTORS_ANISO <k_{sol}_min, float> <k_{sol}_max, float> <k_{sol}_step, float> <B_{sol}_min, float> <B_{sol}_max, float> <B_{sol}_step, float> <B_R_min, float> <B_R_max, float> <B_R_step, float> <B_Z_min, float> <B_Z_max, float> <B_Z_step, float>
The parameters k_{sol}, B_{sol} B_R and B_Z are optimized to obtain lower whole pattern least squares residuals for every model considered.

This command currently works in the regimes DISPLAY and BROWSE_MODEL_SPACE.

SEARCH_B_FACTORS_ISO The parameters k_{sol}, B_{sol} and B_{iso} are optimized by simulated annealing (SA) to obtain lower whole pattern least squares residuals for every model considered. This command currently works in the regimes DISPLAY and BROWSE_MODEL_SPACE. There are two variants of the command syntax, which depend on whether the (SA) optimization is started from a random values of the parameters k_{sol}, B_{sol} and B_{iso}, or from the fixed initial set of values:

SEARCH_B_FACTORS_ISO <k_{sol}_min, float> <k_{sol}_max, float> <k_{sol}_step, float> <B_{sol}_min, float> <B_{sol}_max, float> <B_{sol}_step, float> <B_{iso}_min, float> <B_{iso}_max, float> <B_{iso}_step, float>

This syntax is applied when the command RANDOM_START is ON. For each parameter, the initial parameter value is chosen randomly between *_min* and *_max* values. The initial step of optimization is set to *_step* value. During the optimization, the parameters are allowed to vary between the defined *_min* and *_max* values.

SEARCH_B_FACTORS_ISO <k_{sol}_min, float> <k_{sol}_max, float> <k_{sol}_step, float> <k_{sol}_init, float> <B_{sol}_min, float> <B_{sol}_max, float> <B_{sol}_step, float> <B_{sol}_init, float> <B_{iso}_min, float> <B_{iso}_max, float> <B_{iso}_step, float> <B_{iso}_init, float>

This syntax is applied when the command RANDOM_START is OFF. For each parameter, the initial parameter value is set to *_init* value, which must be chosen between *_min* and *_max* values. The initial step of optimization is set to *_step* value. During the optimization, the parameters are allowed to vary between the defined *_min* and *_max* values.

7.2 Model manipulation

READ_MODEL <filename, string> read model in PDB format.

READ_ADDITIONAL_MODEL <filename, string >

read additional model from PDB file and place to the end of *initial model*. The command used for building random prion fiber model.

BUILD_RANDOM_FIBER The command used for building random prion fiber model. Internal use only.

READ_CHAIN_ID <ON/OFF> this command defines whether to read chain ID, when the program reads in a PDB file with a model. *The command is deprecated. Reading of chain ID from pdb file is performed automatically.*

WRITE_CHAIN_ID <ON/OFF> this command defines whether to write chain ID, when the program writes a PDB file with a model.

The following commands determine which atoms should be read in from a PDB file:

IGNORE_HYDROGENS ON hydrogen atoms are not read in.

NO_H_NO_HOH ON — hydrogen atoms and water are not read in.

READ_BACKBONE_ONLY ON — only backbone atoms, C_β and atoms of proline sidechains are read.

READ_CA_ONLY ON — only C_α atoms are read in.

READ_HETATM ON — read in heteroatoms.

READ_ALL ON — read all atoms. By default, if no directives are given in the script file, all atoms are read in from PDB file.

USE_B_FACTORS <ON/OFF> — the command defines whether the individual B-factors should be read in from PDB file and used in calculations. The use of this command discontinued.

CHECK_MODEL <ON/OFF> — check close contacts between backbone atoms of the adjacent asymmetric units

CLASH_DISTANCE <float> — The command defines the distance of close contact, (Å).

CONTACT_CHAINS <index1> [<index2>] [<index3>]…
The command defines which symmetrically related units should be checked for close contacts with the original unit under current helical parameters. The command parameters include the indexes of the particular symmetrically related units on the helix, which should be checked.
Example: in case of TMV, the neighboring units that might have close contacts with the original unit are –17, -16, -15, -1, 1, 15, 16, 17. The command syntax in this case is
CONTACT_CHAINS –17 -16 -15 -1 1 15 16 17

MAX_NUM_CLASHES <int> — Maximal number of close contacts allowed.

PRINT_CLASHES <ON/OFF> — The command prints on screen atoms of the adjacent units that are too close.

CHECK_SEGMENTS — currently not used.

There are to structures that are used in DISORDER to store the model. They are referenced as *initial model* (**_INIT_MODEL** in the command names) and *current model* (**_CUR_MODEL**). When the program reads in a model from a PDB file, it is stored in the *initial model*. It is recommended to design user scripts in such a way that after coordinates transformations the model is placed into *current model,* since the *current model* is used in fiber diffraction calculations. Many commands are design to behave in such a manner (**MOVE_MODEL**, for example). It is possible, however, to do coordinates transformations on *initial/current* model and store result

back into it, as well as to copy protein structures between these to models to provide user some flexibility. The following commands describe the coordinates transformations applied to the *initial* and *current models.*

ORIENT_INIT_MODEL
ORIENT_CUR_MODEL
These commands orient the chosen model by positioning 3 atoms. The first atom placed into the origin of the coordinate system, the second atom is placed onto x axis, and the third atom is placed into the xz plane. The command syntax has two variants. If the parameter READ_CHAIN_ID is ON, then the syntax is

ORIENT_###_MODEL <*chain ID1*, char> <*residue number1*, int> <*atom name1*, string> <*chain ID2*, char> <*residue number2*, int> <*atom name2*, string> <*chain ID3*, char> <*residue number3*, int> <*atom name3*, string>

If the parameter READ_CHAIN_ID is OFF, then the syntax is

ORIENT_###_MODEL <*residue number1*, int> <*atom name1*, string> <*residue number2*, int> <*atom name2*, string> <*residue number3*, int> <*atom name3*, string>

CM_INIT_TO_ORIGIN
CM_CUR_TO_ORIGIN
The commands translate the center of mass of the chosen model into the origin of the coordinate system.

COPY_INIT_TO_CUR
Copy *initial model* to *current model*.
COPY_CUR_TO_INIT
Copy *current model* to *initial model*. The original model read from the PDB file is deleted in this operation.

ADD_TRANSLATED <*to_x*, (Å), float> <*to_y*, (Å), float> <*to_z*, (Å), float>
The command adds *initial model* translated by vector {*to_x, to_y, to_z*} to the end of the *current model*.

ROTATE_X_INIT <*angle*, (deg), float>
ROTATE_X_CUR <*angle*, (deg), float>
ROTATE_Y_INIT <*angle*, (deg), float>
ROTATE_Y_CUR <*angle*, (deg), float>
ROTATE_Z_INIT <*angle*, (deg), float>
ROTATE_Z_CUR <*angle*, (deg), float>
These commands rotate the chosen model around the chosen axis by an *angle*.

ADD_ROT_X180_TRANSLATED <*to_x*, (Å), float> <*to_y*, (Å), float> <*to_z*, (Å), float>
ADD_ROT_Y180_TRANSLATED <*to_x*, (Å), float> <*to_y*, (Å), float> <*to_z*, (Å), float>
ADD_ROT_Z180_TRANSLATED <*to_x*, (Å), float> <*to_y*, (Å), float> <*to_z*, (Å), float>
These commands add *initial model* rotated by 180˚ around the chosen axis and translated by a vector {*to_x, to_y, to_z*} to the end of the *current model*.

TRANSLATE_INIT_MODEL <*to_x*, (Å), float> <*to_y*, (Å), float> <*to_z*, (Å), float>
TRANSLATE_CUR_MODEL <*to_x*, (Å), float> <*to_y*, (Å), float> <*to_z*, (Å), float>
These commands translate the chosen model by a vector {*to_x, to_y, to_z*}. The result is stored in the same model.

RANDOM_MODEL <ON/OFF> if ON, the atoms of a model provided by user are randomly distributed in the cylindrical sector, corresponding to a volume occupied by an asymmetric unit under the current helical parameters. The command is used to test the sensitivity of the least squares residuals.

RANDOMIZE_MODEL <x_0, float> <y_0, float> <z_0, float>
the atomic coordinates of the model are disturbed by adding Gaussian distributed random numbers with σ_x, σ_y, σ_z determined by x_0, y_0, z_0. The command is used to test the sensitivity of the least squares residuals.

MAX_FIBER_RADIUS <float> maximal radius, (Å), of the model fibril allowed, when the model orientation parameters are varied.

MOVE_MODEL this command is used for the positioning of the asymmetric unit relative to the fibrillar axis. Recommended use: first, the center of mass of the model is placed into the origin of the coordinate system; then MOVE_MODEL used to orient the model using rotation by Eiler angles and to displace it in radial direction. Several variants of the command syntax are acceptable:

MOVE_MODEL OFF the coordinates of the asymmetric unit are not modified.

MOVE_MODEL <*radius*, (Å), float> <*alpha*, (deg), float> <*beta*, (deg), float> <*gamma*, (deg), float>
This syntax is accepted in the regime **DISPLAY.** The model is rotated by angle *alpha* around *z* axis, then by angle *beta* around *x* axis, then by angle *gamma* around *z* axis again (Eiler rotation angles). Then model is translated along *x* axis, in positive direction, by *radius*.

MOVE_MODEL <keyword> <*value_min*, float> <*value_max*, float> <*increment*, float>
This syntax is accepted in the regimes **BROWSE_MODEL_SPACE** and **BROWSE_UNIT_CELL.** The keywords accepted are **RADIUS, ALPHA, BETA** and **GAMMA**. The parameter of the model orientation is varied in the range between *value_min* and *value_max* with a step *increment*. The diffraction patterns for these orientations will be calculated.

BUILD_CRYSTALITE <*A*, (Å), float> <*B*, (Å), float> <*C*, (Å), float> <*alpha*, (deg), float> <*beta*, (deg), float> <*gamma*, (deg), float> <*number of cells A*, int> <*number of cells B*, int> <*number of cells C*, int>
This command builds a small crystallite from a unit cell model stored in the initial model. The command parameters include dimensions and angles of the crystal lattice and the numbers of unit cells that should be built in three lattice directions. The center mass of the crystallite model built is placed into the origin of the coordinates system.

7.3 Output options

PRINT_SELECTION_RULE<ON/OFF> the command prints to the screen the results of the application of the selection rule, i.e. which Bessel orders appear on which layer lines.

PRINT_INPUT_MODEL <ON/OFF> the command prints to the screen the coordinates of a model read from PDB file. It is used to control the correctness of the program input.

WRITE_RADIAL_SCAN <*R1*, (pix), int> <*Z1*, (pix), int> <*R2*, (pix), int> <*Z2*, (pix), int> <filename, string>
The use of this command is discontinued. Use WRITE_COMPARED_SCANS instead.

WRITE_COMPARED_SCANS <*R1*, (pix), int> <*Z1*, (pix), int> <*R2*, (pix), int> <*Z2*, (pix), int> <filename, string>
The command writes into a file the compared intensity profiles of the simulated and observed diffraction patterns averaged across the sector defined by its diagonal corners {*R1, Z1*} and {*R2, Z2*}.

WRITE_ASSEMBLY_MODEL <filename, string> <*number of units*, int>
Write the fibril model that corresponds to the chosen helical symmetry and orientation parameters into a PDB file. The segment of the fibril that contains *number of units* asymmetric units in it is written.

WRITE_CURRENT_MODEL <filename, string>
Writes the current model into a PDB file. The command is commonly used to control the results of the coordinates transformations applied to the model.

WRITE_CRYSTAL_MODEL <filename, string>
Writes the model of the small crystallite into a PDB file. The command is commonly used to control the results of crystallite building.

WRITE_SEQUENCE <filename, string> Writes sequence of the model into a filein a one-letter code format.

SHOW_MODEL <full path to the model> Displays the current model in the Internet Explorer window. The command requires installation of Chime plug-in for viewing molecular structures.

SHOW_ASSEMBLY_MODEL <full path to the model>
Displays the model fibril in the Internet Explorer window. The command requires installation of Chime plug-in for viewing molecular structures.

REPORT_FILE <filename, string> the command writes the log of the program run into a text file.

WRITE_SIMULATED_PATTERN <filename, string>

This command writes the simulated diffraction pattern in 4 quadrants view into a file in .spr format

WRITE_EXP_DATA <filename, string> This command writes the observed diffraction pattern in 4 quadrants view into a file in .spr format.

WRITE_COMPARED_PATTERNS <filename, string>

This command writes the compared observed and simulated diffraction patterns into a file in .spr format. The observed pattern appears in upper right and lower left quadrants of the image, the simulated pattern appears in the upper left and lower right quadrants.

WRITE_COMPARED_ICALC_IDIS <filename, string>

This command writes the compared observed, oriented and disoriented diffraction patterns into a file in .spr format. The observed pattern appears in upper right and lower left quadrants of the image, the simulated patterns appear in the upper left and lower right quadrants.

WRITE_I_OBSERVED <filename, string>

This command writes the observed diffraction pattern into a file in .spr format.

WRITE_I_CALCULATED <filename, string>

This command writes the oriented diffraction pattern into a file in .spr format.

WRITE_I_DISORDERED <filename, string>

This command writes the disoriented diffraction pattern into a file in .spr format.

WRITE_COMPARED_EQUATORS <filename, string>

This command writes the equatorial intensity profiles of the observed and simulated diffraction patterns into a file in ASCII format.

WRITE_COMPARED_STRIPES <*R1*, (pix), int> <*Z1*, (pix), int> <*R2*, (pix), int> <*Z2*, (pix), int> <filename, string>

This command writes into a file the compared simulated and observed intensity profiles averaged across the stripe, defined by its diagonal corners

WRITE_COMPARED_SLICES <*period*, (Å), float> <filename, string>

This command writes into a file compared intensity profiles of the slices made across the observed and simulated diffraction patterns at *Z=l/period*, where *l=0,1,2…*

SHOW_COMPARISON <ON/OFF> show the simulated and observed diffraction patterns in fit2d.

7.4 Image manipulation

READ_IMAGE <filename, string> The command reads the experimental image from file. Currently, only the images in .spr format generated by program fit2d are accepted. The size of the image in pixels must by **RSIZE** by **ZSIZE**.

The mask defines which pixels to ignore and which to include in calculation of the diffraction pattern. If the pixel has mask value 1, it will be included in calculations, if 0, it will be ignored and the intensity of this pixel in the calculated diffraction pattern will be set to 0. It will be also ignored when scaling the simulated pattern and calculating least squares residuals. Typically, the regions of the diffraction pattern which are masked out include beamstop, shadowed region near the meridian and the pixels that have D>=RSIZE, where $D^2=R^2+Z^2$. The user may also decide to mask out other pixels. The following commands are used to determine the mask:

READ_MASK <filename, string> The command reads the mask from file. Currently, only the masks in .spr format generated by program fit2d are accepted. The size of the mask in pixels must be **RSIZE** by **ZSIZE**.

DEFAULT_MASK ON The command defines the mask that has pixels on the beamstop of radius 5 pix and pixels with R=0 masked out.

AUTO_MASK ON The command defines the mask from a user provided experimental image file. If the intensity of a pixel is less, or equal to 0 in the experimental pattern, the pixel is masked out. The use of this command is recommended, if user cannot provide the mask.

READ_SCALE_WEIGHTS <filename, string>
The command reads the weights for scaling the simulated diffraction pattern from file. Currently, only the files in .spr format generated by program fit2d are accepted. The size of the weights matrix must be **RSIZE** by **ZSIZE**.

READ_R_WEIGHTS <filename, string>
The command reads the weights for the calculation of the whole pattern least squares residuals from file. Currently, only the files in .spr format generated by program fit2d are accepted. The size of the weights matrix must be **RSIZE** by **ZSIZE**.

FLAT_BACKGROUND <*intensity*, float> subtract the *intensity* of the flat background from the experimental image.

CIRC_BACKGROUND <*radius1*, int> <*intensity1*, float> <*radius2*, int> <*intensity2*, float> [<*radius3*, int> <*intensity3*, float>] …
The command defines the intensity profile of the circularly symmetric background subtracted from the experimental image. The profile is defined by values of *intensities* at certain *radii* (pix)

PEAK_PENALTY_R <ON/OFF>

PEAK_PENALTY_S <ON/OFF>
INTENSITY_WEIGHTS <ON/OFF> The command defines weights for the calculation of the least squares residuals from the experimental image intensity statistics. The pixels that have high intensity are assigned lower weights.

SET_SCALE_WEIGHTS <RECTANGLE/SECTOR> <*R1*, (pix), int> <*Z1*, (pix), int> <*R2*, (pix), int> <*Z2*, (pix), int> <*weight*, float>

This command allows user to modify the pixels *weights* used for scaling of the calculated diffraction pattern in a certain region of the diffraction pattern. The region may have a shape of RECTANGLE or SECTOR and is defined by the pixel coordinates of its diagonal corners {*R1, Z1*} and {*R2, Z2*}. By default, all the *weights* are equal to 1.0. If the weights are set to zero, the region is ignored during scaling.

SET_R_WEIGHTS <RECTANGLE/SECTOR> <*R1*, (pix), int> <*Z1*, (pix), int> <*R2*, (pix), int> <*Z2*, (pix), int> <*weight*, float>

This command allows user to modify the pixels *weights* used for the calculation of the whole pattern least squares residual in a certain region of the diffraction pattern. The region may have a shape of RECTANGLE or SECTOR and is defined by the pixel coordinates of its diagonal corners {*R1, Z1*} and {*R2, Z2*}. By default, all the *weights* are equal to 1.0. If the weights are set to zero, the region is ignored during the residual calculation.

CALCULATE_R The command allows the user to define the parts of the diffraction pattern for which the least squares residuals should be calculated. The lower values of the residuals indicate the better fit of the simulated diffraction pattern. Several variants of the command syntax are accepted:

CALCULATE_R WHOLE_PATTERN all the pixels of the simulated diffraction pattern, except those masked out are used in calculation of the residuals. The residual is printed into the REPORT_FILE.

CALCULATE_R BANDS A set of the residuals calculated for the different resolution bands. By default, the widths of all the resolution bands are set to ten pixels. The residuals are printed into the REPORT_FILE.

CALCULATE_R RESOLUTION The residuals are calculated up to various resolution limits. By default, the step between the resolution limits is set to ten pixels. The residuals are printed into the REPORT_FILE.

CALCULATE_R REGION <RECTANGLE/SECTOR> <*R1*, (pix), int> <*Z1*, (pix), int> <*R2*, (pix), int> <*Z2*, (pix), int>

The residual is calculated for the user-defined region that might have a shape of a rectangle or sector. The shape is defined by the coordinates of two diagonal corners. Several regions may be defined by a user. The residuals are printed into the REPORT_FILE.

IO_THRESHOLD <*intensity*, float>

This command defines the threshold *intensity* in the experimental diffraction pattern, such that pixels with low intensity are excluded from the calculation of the least squares residuals. The command is used to exclude the regions of the diffraction pattern, which contain only the background.

CUTOFF_HIGH <*intensity*, float>

This command allows to suppress the influence of highly over- or underestimated calculated intensities on the values of the least squares residuals. If the command is applied, then the least squares residuals are calculated as

$$R = \frac{\sum_{R,Z} f(|I_{calc}(R,Z) - I_{obs}(R,Z)|)^2}{\sum_{R,Z} I_{obs}(R,Z)^2},$$

where $f(x)=x$, if $x<intensity$, and $f(x)=intensity$, if $x>intensity$.

IC_IO_RATIO <RECTANGLE/SECTOR> <*R1*, (pix), int> <*Z1*, (pix), int> <*R2*, (pix), int> <*Z2*, (pix), int>

This command writes into the REPORT_FILE the information about the ratio of the total simulated to the total observed intensities in a user-defined region of the diffraction pattern. This region may have a shape of a RECTANGLE or SECTOR. It is defined by the coordinates of its diagonal corners. More than one region may be defined by a user.

7.5 Commands for SEARCH_MODEL_SPACE regime

To be described in future versions.

7.6 Commands for BROWSE_UNIT_CELL regime

CRYSTAL_SYMMETRY <symmetry group, string>

Defines symmetry group of the crystallite model.

NUM_CELLS_A <int>
NUM_CELLS_B <int>
NUM_CELLS_C <int>

These commands define how many unit cells must be built in each lattice direction, when the small crystallite model is constructed.

CELL_A FREEZE <*A*, (Å), float>
CELL_B FREEZE <*B*, (Å), float>
CELL_C FREEZE <*C*, (Å), float>

These command define the unit cell dimensions. The keyword FREEZE means that the corresponding parameter has the same value for all the models for which the diffraction patterns are calculated.

CELL_A BROWSE <*A_min*, (Å), float> <*A_max*, (Å), float> <*A_step*, (Å), float>
CELL_B BROWSE <*B_min*, (Å), float> <*B_max*, (Å), float> <*B_step*, (Å), float>

CELL_C BROWSE <*C_min*, (Å), float> <*C_max*, (Å), float> <*C_step*, (Å), float>

These commands define the limits of variation of the unit cell dimensions in the small crystallite models. The keyword BROWSE means that the corresponding parameter will be varied between the defined minimal and maximal values with an increment *step*.

CELL_ALPHA FREEZE <*alpha*, (deg), float>
CELL_BETA FREEZE <*beta*, (deg), float>
CELL_GAMMA FREEZE <*gamma*, (deg), float>

These command define the unit cell angles. The keyword FREEZE means that the corresponding parameter has the same value for all the models for which the diffraction patterns are calculated.

CELL_ALPHA BROWSE <*alpha_min*, (deg), float> <*alpha_max*, (deg), float> <*alpha_step*, (deg), float>
CELL_BETA BROWSE <*beta_min*, (deg), float> <*beta_max*, (deg), float> <*beta_step*, (deg), float>
CELL_GAMMA BROWSE <*gamma_min*, (deg), float> <*gamma_max*, (deg), float> <*gamma_step*, (deg), float>

These commands define the limits of variation of the unit cell angles in the small crystallite models. The keyword BROWSE means that the corresponding parameter will be varied between the defined minimal and maximal values with an increment *step*.

RANDOM_STACK The use of this command discontinued.

7.7 Commands used for simulated annealing optimization

NUM_ITERATIONS <int>

Command defines the number of steps of the simulated annealing optimization.

ANNEALING TEMPERATURE <double**>**

Command defines the temperature for the simulated annealing optimization. Recommended range of temperature values is between 0.05 and 0.5.

RANDOM_START <ON/OFF> defines whether the simulated annealing optimization should start from randomly chosen, or from a user-defined set of parameters.

8 References

1. Franklin, R. E. & Klug, A. (1955). The splitting of layer lines in X-ray fibre diagrams of helical structures: application to tobacco mosaic virus. *Acta Crystallogr.* 8, 777-780.
2. Klug, A., Crick, F. H. C. & Wyckoff, H. W. (1958). Diffraction By Helical Structures. *Acta Crystallogr.* 11, 199-213.
3. Badger, J. (1997). Modeling and Refinement of Water Molecules and Disordered Solvent. *Methods in Enzymology* 277, 344-353.
4. Holmes, K. C. & Barrington Leigh, J. (1974). The Effect Of Disorientation On The Intensity Distribution Of Non-Crystalline Fibres. I. Theory. *Acta Crystallogr.* A30, 635-645.
5. Stubbs, G. (1989). The probability distributions of X-ray intensities in fiber diffraction: largest likely values for fiber diffraction R factors. *Acta Crystallogr.* A45, 254-258.
6. Wang, H. & Stubbs, G. (1993). Molecular Dynamics in Refinement against Fiber Diffraction Data. *Acta Cryst.* A49, 504-513.
7. Nelson, R., Sawaya, M. R., Balbirnie, M., Madsen, A. Ø., Riekel, C., Grothe, R. & Eisenberg, D. (2005). Structure of the cross-b spine of amyloid-like fibrils. *Nature* 435, 773-778.

9 Index