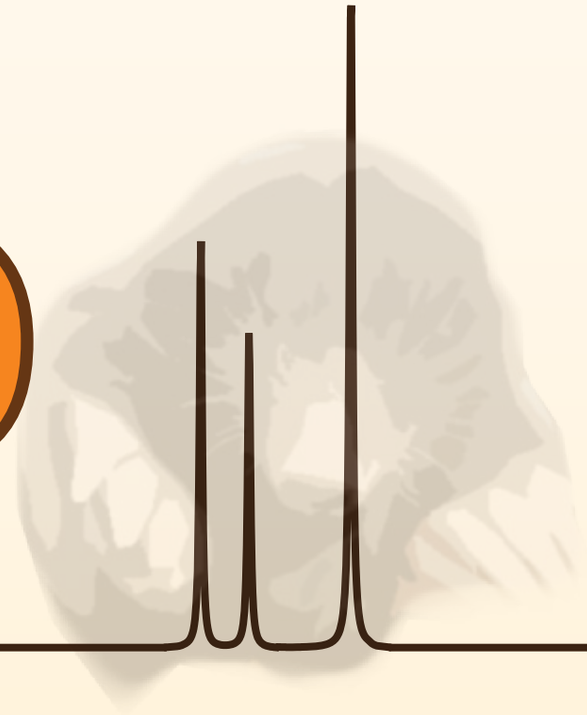
TALP, a multisolution direct-space strategy for solving molecular crystals from powder diffraction data based on restrained LS

PowderXRD 2013

TALP



Oriol Vallcorba
Jordi Rius
Carlos Frontera
Carles Miravitlles

ovallcorba@icmab.com

Part I

Introduction & ***Practical Aspects***

CSIC
Consejo Superior de Investigaciones Científicas

ICMAB
Institut de Ciència de Materials de Barcelona
www.icmab.es
Campus de la UAB
08193-Bellaterra (Spain)

ALBA
CELLS - ALBA Synchrotron Light Facility

Crystal Structure Solution

X-ray Diffraction

Single Crystal (100-15μm)

Powder [1] (<100nm)

Direct Methods [2]

- Reciprocal space
- High resolution data

XLENS

Direct-Space Methods [3]

- Molecular model
- Agreement obs/calc

Patterson-search

TALP

Global optimization

- FOX (Favre-Nicolin & Cerny, 2002)
- DASH (David et al., 2006)
- ESPOIR (Le Bail, 2001)
- PSSP (Pagola, 2010)
- TOPAS (Coelho, 2000)
- PowderSolve (Engel, 1999)
- EXPO (Altomare et al., 2004)
- …

MonteCarlo (Sim. Annealing)

[1] David W.I.F., Shankland, K. (**2008**). Acta Cryst. A64, 52-64.

[2] Rius, J. (**2011**). Acta Cryst. A67, 63-67.

[3] Černý, R., Favre-Nicolin, V. (**2007**). Z. Kristallogr. 222, 105-113.

TALP approach

Pattern Matching
(Le Bail method)

AJUST
(Rius *et al.*, 1996)

RIBOLS
(Rius, 1989)

Restrained
Rietveld
refinement

Search Algorithm

Direct space
(G. Opt & Patt. Search)

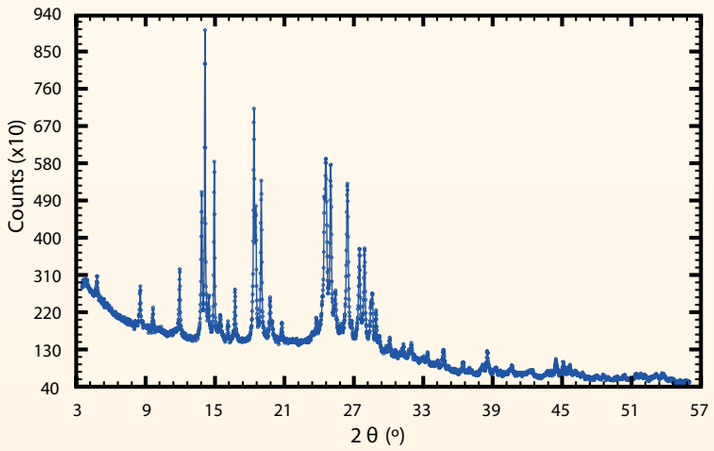
Rot function

TALP

- Fast Least-Squares (**FLS**) minimizations of atomic coordinates
- Restrained-based
- User-friendly

From data to crystal structure

Powder diffraction data



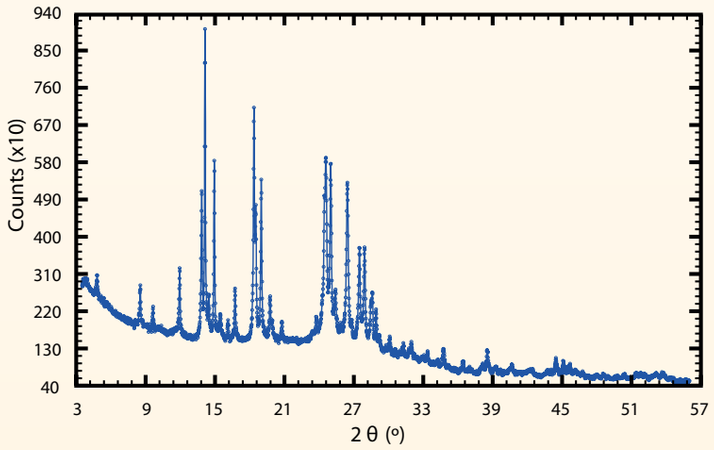
Indexing

TALP

Crystal Structure

From data to crystal structure

Powder diffraction data



Indexing

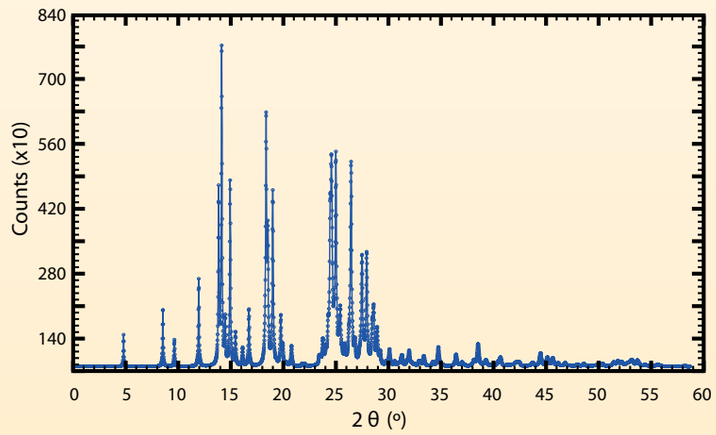
DAJUST software

Space group and refined cell parameters

Intensity and profile information (QCK file)

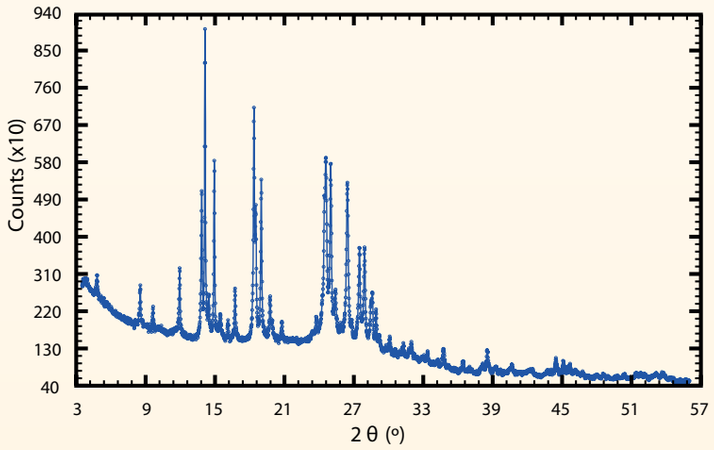
TALP

Crystal Structure



From data to crystal structure

Powder diffraction data



Indexing

DAJUST software

Space group and refined cell parameters

Intensity and profile information (QCK file)

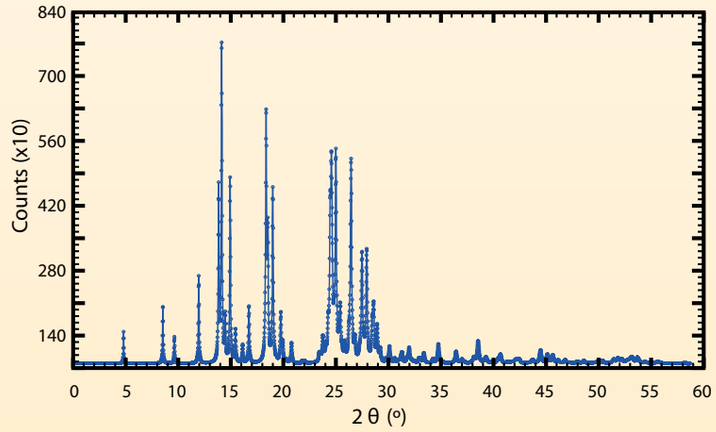
Molecular Model and Restraints

- Completeness
- Bonds, angles, FRB
- Atomic coordinates

Known structures or calculations

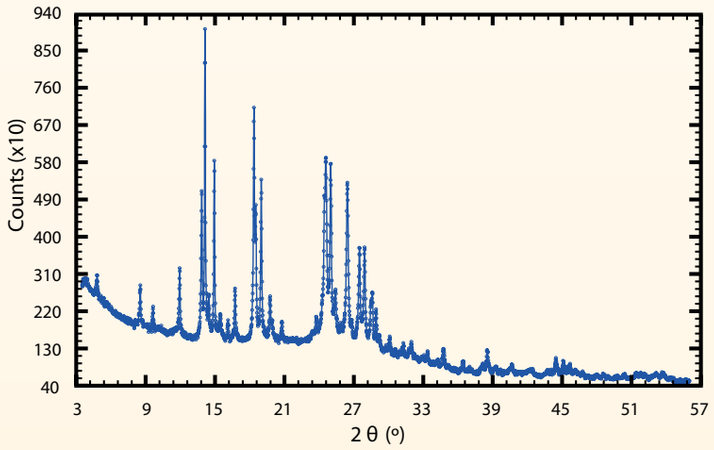
TALP

Crystal Structure



From data to crystal structure

Powder diffraction data



Indexing

DAJUST software

Space group and refined cell parameters

Intensity and profile information (QCK file)

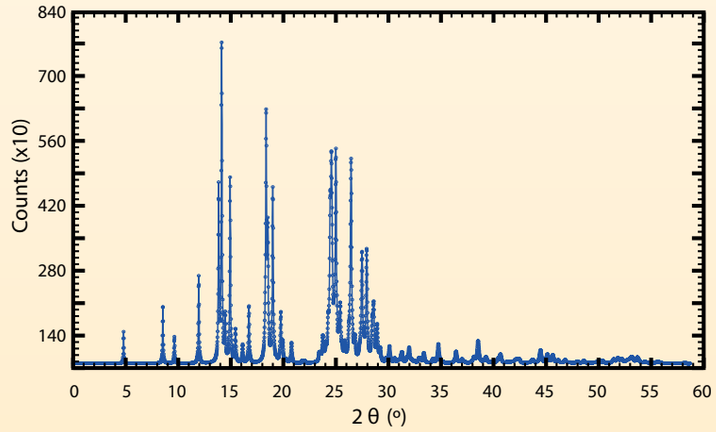
Molecular Model and Restraints

- Completeness
- Bonds, angles, FRB
- Atomic coordinates

Known structures or calculations

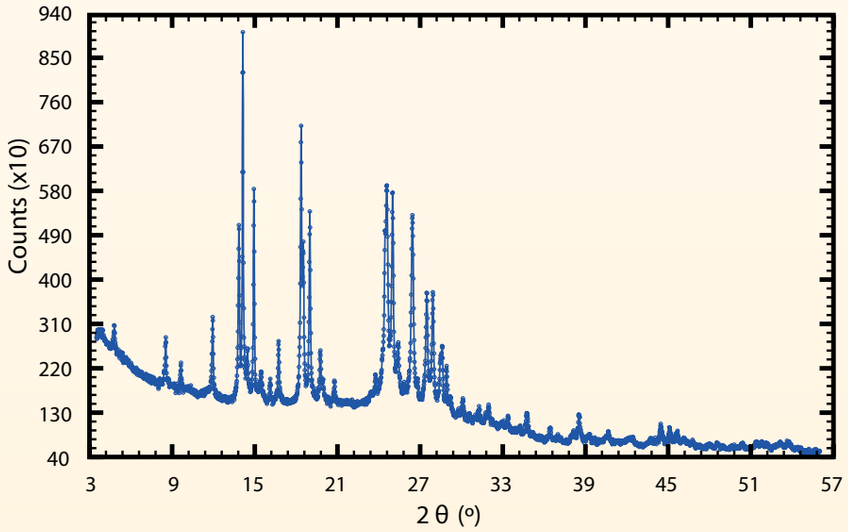
TALP

Crystal Structure



- **Data quality** (overlap, resolution, preferred orientation, instrumental aspects,...)
- **Structural complexity** (torsion angles, Z', heavy atoms, solvents,...)
- Construction and description of the **molecular model**

Intensity extraction: DAJUST software



DAJUST software

Prepare PDD for structure solution

```
h k l Yobs ... 1
h k l Yobs ... 1
h k l Yobs ... 1
h k l Yobs ... 2
h k l Yobs ... 2
h k l Yobs ... 1
```

- Easy/fast intensity extraction
- 1 phase only
- Output files prepared for XLENS (.POW) & TALP (.QCK)

Intensity extraction: DAJUST software

$y_{obs,\mathbf{H}}$

&

Profile information

3-step extraction procedure:

1. Resolved reflections ***h***
   $(y_{obs,h}; j_h; \Omega_h)$
2. Substraction
3. Non-resolved reflection ***k***
   $(y_{obs,k}; j_k; \Omega_k)$
   $(j_l; \Omega_{l,k})$

For each reflection:

n_h H K L $2\theta_h$ NREST y_{obs} D_h J_h Ω_h w_h η_h m $(y_{obs}-y_{bkg})$ y_{bkg} LP AC

(if non-resolved) → $n_{h(k)}$ n_l $2\theta_l$ $\Omega_{l,h(k)}$
…
…

NREST lines

Vallcorba, O., Rius, J., Frontera, C., Peral, I., Miravitlles, C. (**2012**). J. Appl. Cryst. 45, 844–848.

Molecular model

Generated by calculations (i.e. MM) or from similar structures

- Completeness
- Atomic coordinates
- Bonds, angles, FRB

`code.TXT` *Contains:* Atomic coordinates of the molecular model

Use: To generate the TALP input file (starting model and restraints)

Line 1:	**Title**
Line 2 (cell param.):	**`a b c alpha beta gamma`**
Line 3..n (at. coord.):	**`Label AtomType x y z`**

Molecular model

Generated by calculations (i.e. MM) or from similar structures

- Completeness
- Atomic coordinates
- Bonds, angles, FRB

`code.TXT` *Contains:* Atomic coordinates of the molecular model

Use: To generate the TALP input file (starting model and restraints)

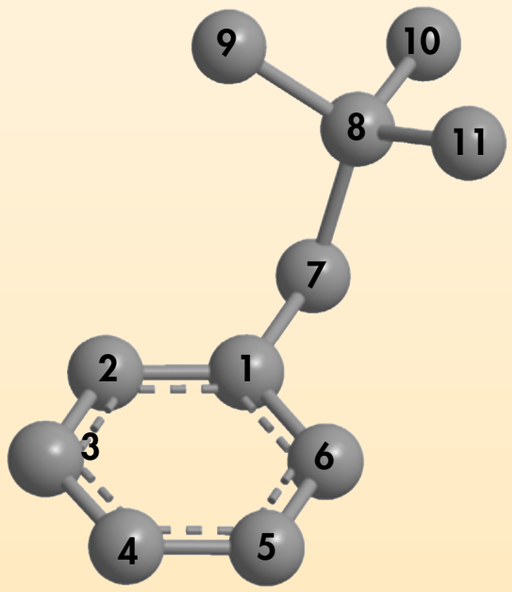
Line 1: **Title**

Line 2 (cell param.): **`a b c alpha beta gamma`**

Line 3..n (at. coord.): **`Label AtomType x y z`**

Restraints: Bond distance, Bond angle, Atomic planes

Free Rotation Bonds:



Molecular model

Generated by calculations (i.e. MM) or from similar structures

- Completeness
- Atomic coordinates
- Bonds, angles, FRB

`code.TXT` *Contains:* Atomic coordinates of the molecular model

Use: To generate the TALP input file (starting model and restraints)

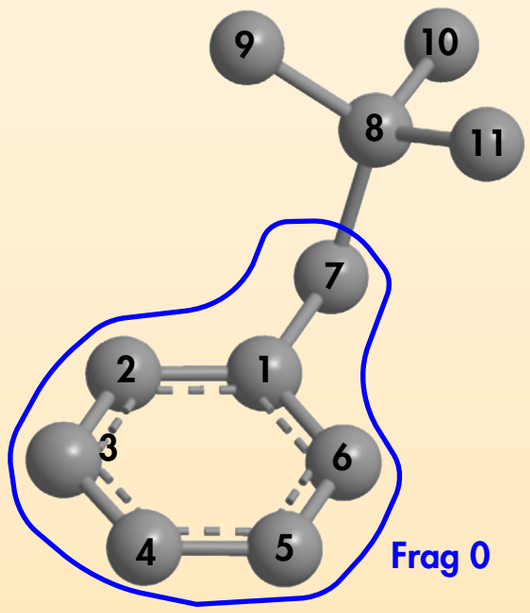
Line 1: **Title**

Line 2 (cell param.): **`a b c alpha beta gamma`**

Line 3..n (at. coord.): **`Label AtomType x y z`**

Restraints: Bond distance, Bond angle, Atomic planes

Free Rotation Bonds:



Frag 0 = Base fragment

Molecular model

Generated by calculations (i.e. MM) or from similar structures

- Completeness
- Atomic coordinates
- Bonds, angles, FRB

`code.TXT` *Contains:* Atomic coordinates of the molecular model

Use: To generate the TALP input file (starting model and restraints)

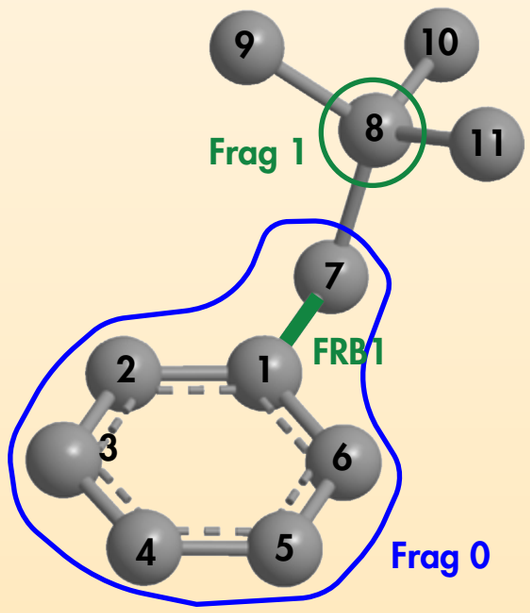
Line 1: **Title**

Line 2 (cell param.): **`a b c alpha beta gamma`**

Line 3..n (at. coord.): **`Label AtomType x y z`**

Restraints: Bond distance, Bond angle, Atomic planes

Free Rotation Bonds:



Frag 0 = Base fragment

FRB1 (C1–C7) = Free rotation bond that moves Frag 1. It uses Frag 0 as pivot.

Molecular model

Generated by calculations (i.e. MM) or from similar structures

- Completeness
- Atomic coordinates
- Bonds, angles, FRB

`code.TXT` *Contains:* Atomic coordinates of the molecular model

Use: To generate the TALP input file (starting model and restraints)

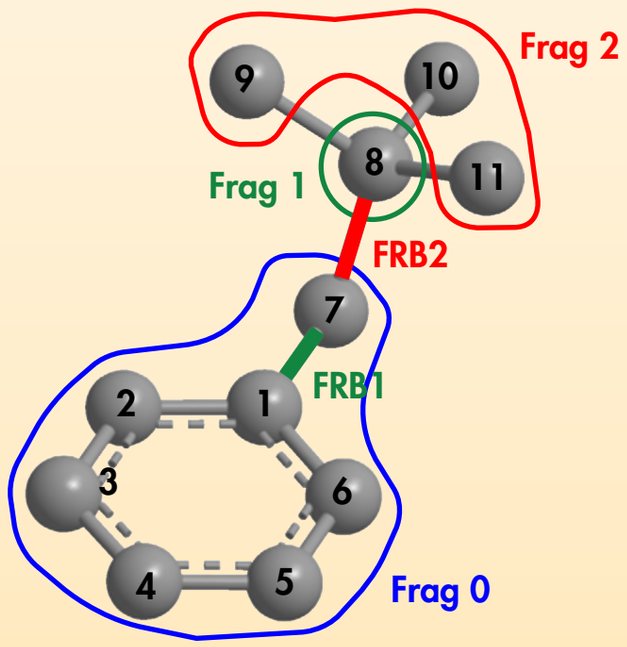
Line 1: **Title**

Line 2 (cell param.): **`a b c alpha beta gamma`**

Line 3..n (at. coord.): **`Label AtomType x y z`**

Restraints: Bond distance, Bond angle, Atomic planes

Free Rotation Bonds:



Frag 0 = Base fragment

FRB1 (C1–C7) = Free rotation bond that moves Frag 1. It uses Frag 0 as pivot.

FRB2 (C7–C8) = Free rotation bond that moves Frag 2. It uses Frag 1 as pivot.

TALP Fast Least Squares (FLS) minimizations

Intensities → **TALP** ← Molecular model

$$M = S_Y + kS_R$$

$$S_Y = \sum_{\mathbf{H}} w_{\mathbf{H}} (y_{obs,\mathbf{H}} - y_{\mathbf{H}})^2 \qquad S_R = \sum_j \sigma_j^{-2} (d_{obs,j} - d_j)^2$$

DAJUST software

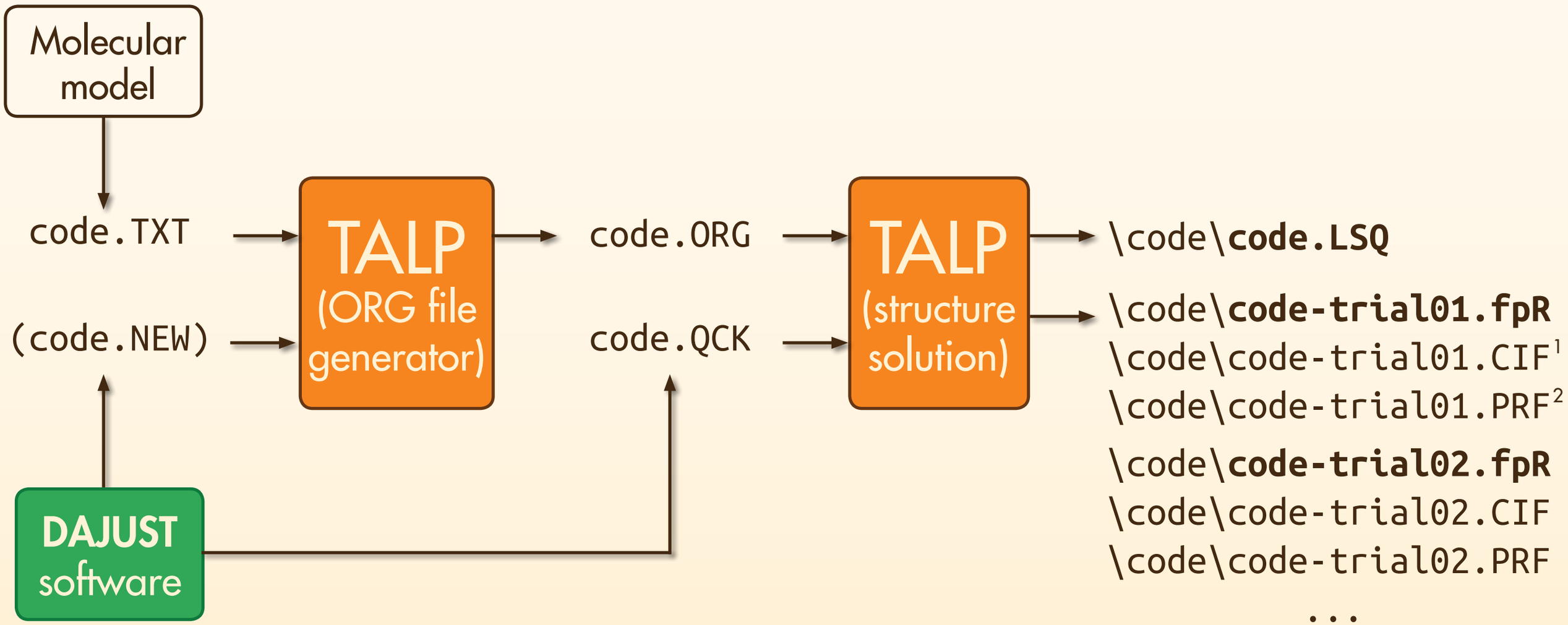
h ***k***

Resolved reflections ***h***: $y_{\boldsymbol{h}} = c j_{\boldsymbol{h}} \Omega_{\boldsymbol{h}} |F_{\boldsymbol{h}}|^2$

non-resolved reflections ***k***: $y_{\boldsymbol{k}} = c j_{\boldsymbol{k}} \Omega_{\boldsymbol{k}} |F_{\boldsymbol{k}}|^2 + c \sum_{\boldsymbol{l}(\boldsymbol{k})} j_{\boldsymbol{l}} \Omega_{\boldsymbol{l},\boldsymbol{k}} |F_{\boldsymbol{l}}|^2$

non-resolved only!

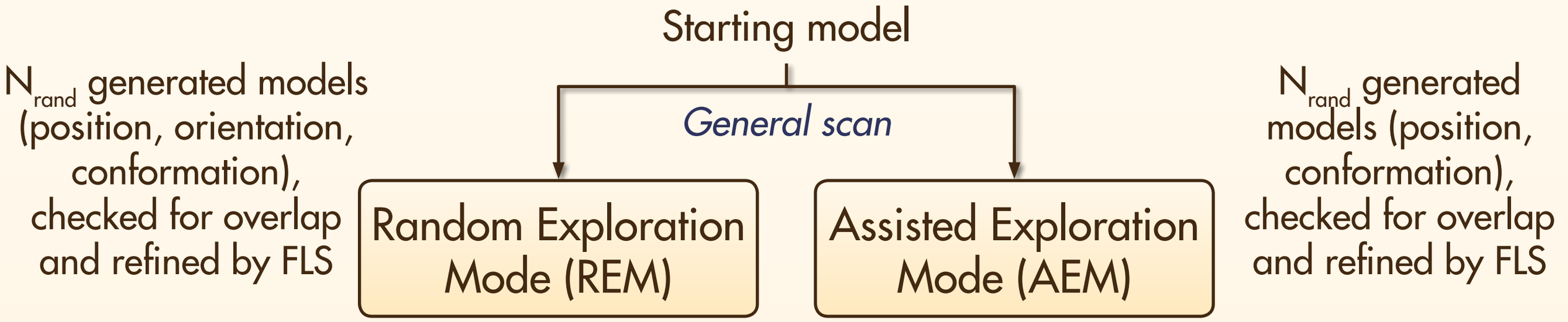
Input and output files in TALP



[1] Crystallographic Information File. Hall, S.R., Allen, F.H., Brown, I.D. (**1991**). *Acta Cryst. A*, *47*, 655.

[2] FULLPROF observed and calculated profile file. Rodriguez-Carvajal J. (**1993**). *Physica B*. 192, *55*

TALP search strategy (trial flow-sheet)



$$\chi_{TALP} = \left(\frac{S_Y}{N_H + N_{REST} - P} \right)^{1/2}$$

Quality of intermediate models after FLS

TALP search strategy (trial flow-sheet)

Starting model

N_{rand} generated models (position, orientation, conformation), checked for overlap and refined by FLS

General scan

Random Exploration Mode (REM)

Assisted Exploration Mode (AEM)

N_{rand} generated models (position, conformation), checked for overlap and refined by FLS

$$\chi_{TALP} = \left(\frac{S_Y}{N_H + N_{REST} - P} \right)^{1/2}$$

Quality of intermediate models after FLS

Local Incremental Scan (LIS)

N_{CI} increments of limited size to the best model (position, orientation, conformation) followed by FLS refinement

TALP search strategy (trial flow-sheet)

Starting model

N_{rand} generated models (position, orientation, conformation), checked for overlap and refined by FLS

General scan

Random Exploration Mode (REM)

Assisted Exploration Mode (AEM)

N_{rand} generated models (position, conformation), checked for overlap and refined by FLS

$$\chi_{TALP} = \left(\frac{S_Y}{N_H + N_{REST} - P} \right)^{1/2}$$

Quality of intermediate models after FLS

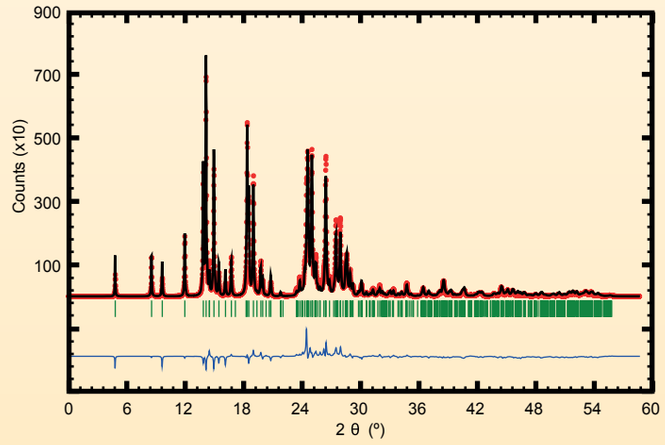
Local Incremental Scan (LIS)

N_{CI} increments of limited size to the best model (position, orientation, conformation) followed by FLS refinement

$$Q_{trial} = \frac{R_{wp,FPR}}{R_{wp,model-free}}$$

Indicator for candidate solutions (usually below 1.5 for correct solutions)

Fixed-Profile Rietveld refinement (FPR)

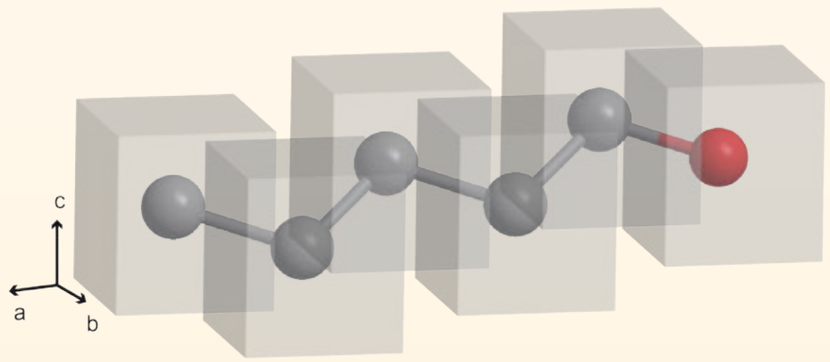


Final model (Q_{trial})

Global Overlap Criterion (GOC)

Controls the amount of molecular overlap

1. Initial molecular model



Minimum internal overlap

Cubes centered to atomic positions
(density function: inside=1, outside=0)

Sum of all density values (for 1 molecule):

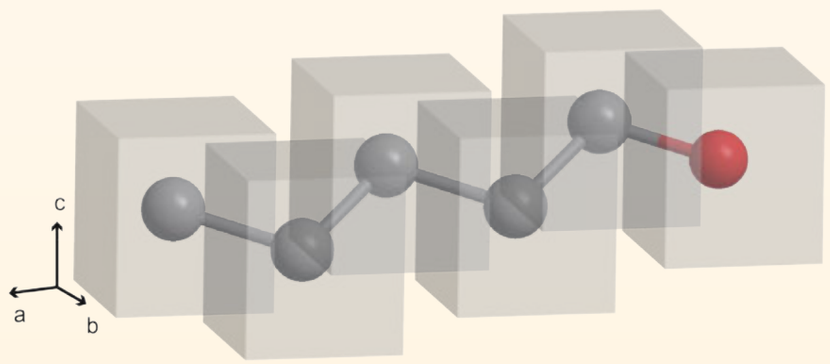
$$O_{imol} = \int_V g_{imol}^2 dV$$

reference value

Global Overlap Criterion (GOC)

Controls the amount of molecular overlap

1. Initial molecular model



Minimum internal overlap

Cubes centered to atomic positions
(density function: inside=1, outside=0)

Sum of all density values (for 1 molecule):

$$O_{imol} = \int_V g_{imol}^2 dV$$

reference value

2. For each randomly generated model

$$q_{intra} = \frac{O_{mol}}{O_{imol}}$$

intramolecular overlap

$$O_{all} = \int_V g_{all}^2 dV \longrightarrow q_{inter} = \frac{O_{all}}{ZO_{mol}}$$

(for all symmetry-related molecules)

intermolecular overlap

IF

q_{intra}>~1.1 OR q_{inter}>1.05

Generation of a new model
(limit 100)

Very fast and non-blocking (no distance calculations)

Assisted Exploration Mode (AEM): Rotation Function

How it works:

1. Identify the search model: fragment with known geometry
2. Calculation of Rot(Ω) for all possible orientations of the fragment

$$Rot(\mathbf{\Omega}) = \int_V P'_{obs} \cdot P_{mod}(\mathbf{\Omega}) \cdot dV$$

   High Rot values are probable orientations of the fragment

3. Most probable orientations are used as initial models in TALP.

   The orientation of the molecule is not a “free” (completely random) parameter in the model generation.

Assisted Exploration Mode (AEM): Rotation Function

How it works:

1. Identify the search model: fragment with known geometry
2. Calculation of Rot(Ω) for all possible orientations of the fragment

$$Rot(\mathbf{\Omega}) = \int_V P'_{obs} \cdot P_{mod}(\mathbf{\Omega}) \cdot dV$$

High Rot values are probable orientations of the fragment

3. Most probable orientations are used as initial models in TALP.

   The orientation of the molecule is not a "free" (completely random) parameter in the model generation.

Possible applications:

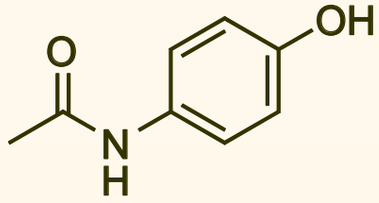
1. Molecules with a rigid part and a flexible chain (with several FRBs).
2. To reduce the problem complexity when Z'=2, by "fixing" the orientation of one of the molecules

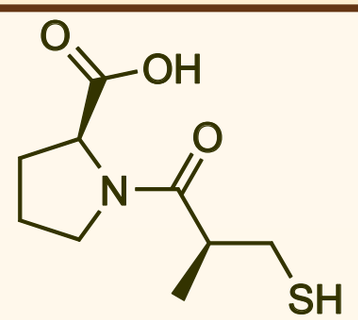
Performance (REM)

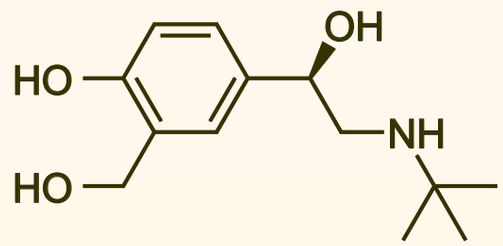
- Results of **100 trials**
- **Calculated models** (MM)
- N_{sol} = No. of solutions
- t_{trial} = Average time per trial
- t_{sol} = Estimated time for obtaining a solution:

$$t_{sol} = \frac{100 t_{trial}}{N_{sol}}$$

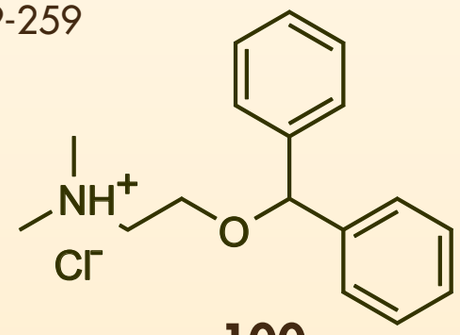
- *FRB* = Free rotation bonds
- Conditions:
  N_{rand}= 1500(20 FLS), REM
  N_{Cl}= 3000(15 FLS)
- Data Florence *et al.* 2005
  J. Appl. Cryst. **2005**, 38, 249-259

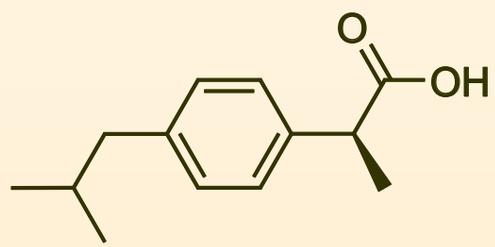


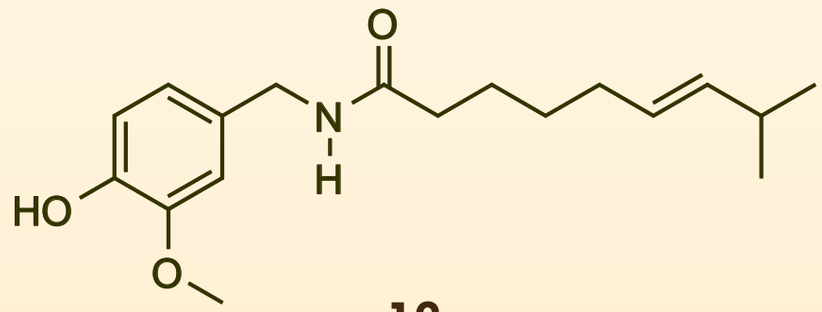




***N*$_{sol}$**	**100**	**50**	**74**
t_{trial} (min)	0.5	1.6	1.5
***t*$_{sol}$** (min)	**0.5**	**3.2**	**2.0**
***Q*$_{trial}$**	1.200	0.950	0.950
r.m.s.d. (Å)	0.057	0.091	0.151
FRB, N_{rest}, N_{par}	**2**, 37, 35	**4**, 42, 44	**5**, 50, 53
s.g., *V* (Å^3)	$P2_1/n$, 772	$P2_12_12_1$, 1077	*Pbca*, 2823







***N*$_{sol}$**	**100**	**20**	**13**
t_{trial} (min)	1.4	3.6	2.6
***t*$_{sol}$** (min)	**1.4**	**18.0**	**20.0**
***Q*$_{trial}$**	1.008	1.094	0.953
r.m.s.d. (Å)	0.281	0.103	0.323
FRB, N_{rest}, N_{par}	**6**, 59, 44	**8**, 86, 91	**11**, 66, 68
s.g., *V* (Å^3)	$Pna2_1$, 1640	$P2_1$, 1245	$P2_1/c$, 1422

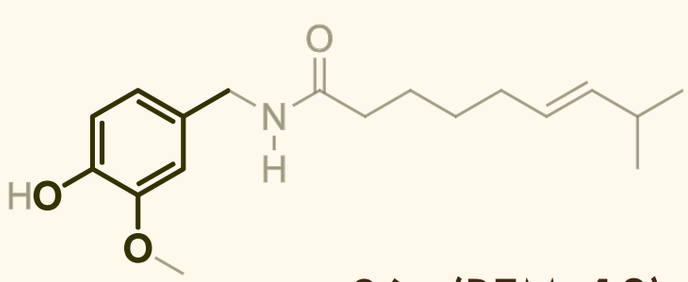
Performance (AEM)

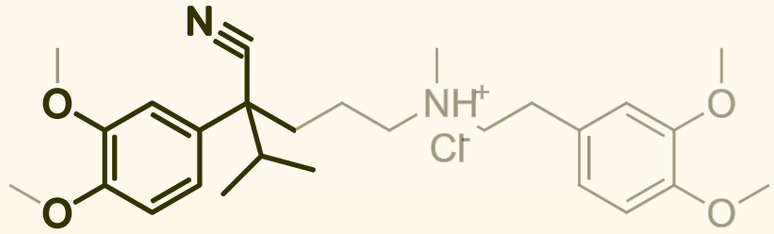
- Results of **100 trials**
- N_{sol} = No. of solutions
- t_{trial} = Average time per trial
- t_{sol} = Estimated time for obtaining a solution:

$$t_{sol} = \frac{100 t_{trial}}{N_{sol}}$$

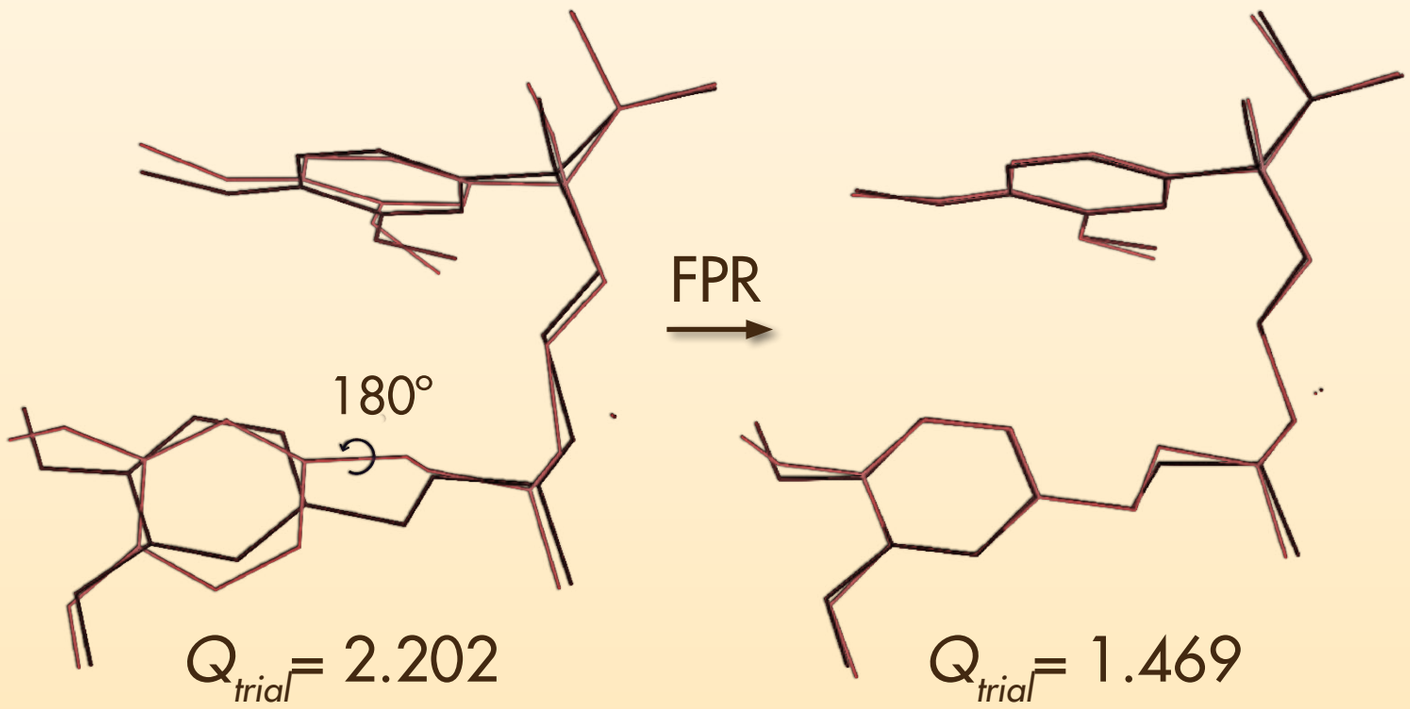
- *FRB* = Free rotation bonds
- Data Florence *et al.* 2005

J. Appl. Cryst. **2005**, 38, 249-259





N_{sol}	**36** (REM=13)	**13** **(+7)**
t_{trial} (min)	2.6	37.2
t_{sol} (min)	**7.2**	**286.2** **416.5**
Q_{trial}	0.953	1.469 2.202
r.m.s.d. (Å)	0.323	0.141
FRB, N_{rest}, N_{par}	**11**, 66, 68	**13**, 99, 104
s.g., *V* (Å^3)	$P2_1/c$, 1422	P-1, 1382
Conditions	N_{rand}= 1500(20 FLS), AEM N_{CI}= 3000(15 FLS)	N_{rand}= 10000(30 FLS), AEM N_{CI}= 15000(15 FLS)



- Significant improvement
- Long atomic chains anchored on terminal aryl
- Model (fragment) accuracy

TALP, a multisolution direct-space strategy for solving molecular crystals from powder diffraction data based on restrained LS

TALP

Oriol Vallcorba
Jordi Rius
Carlos Frontera
Carles Miravitlles

ovallcorba@icmab.com

Part 2

Software demonstration

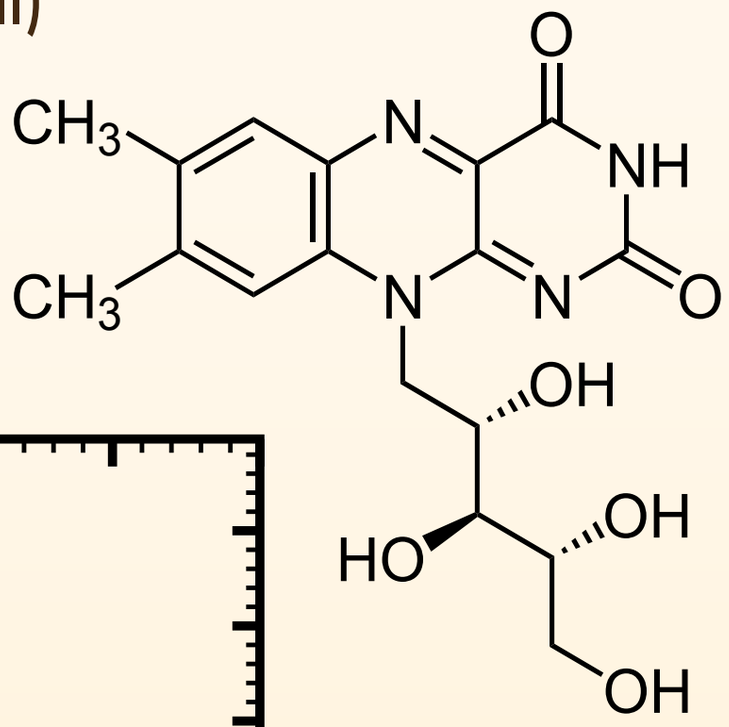
CSIC
Consejo Superior de Investigaciones Científicas

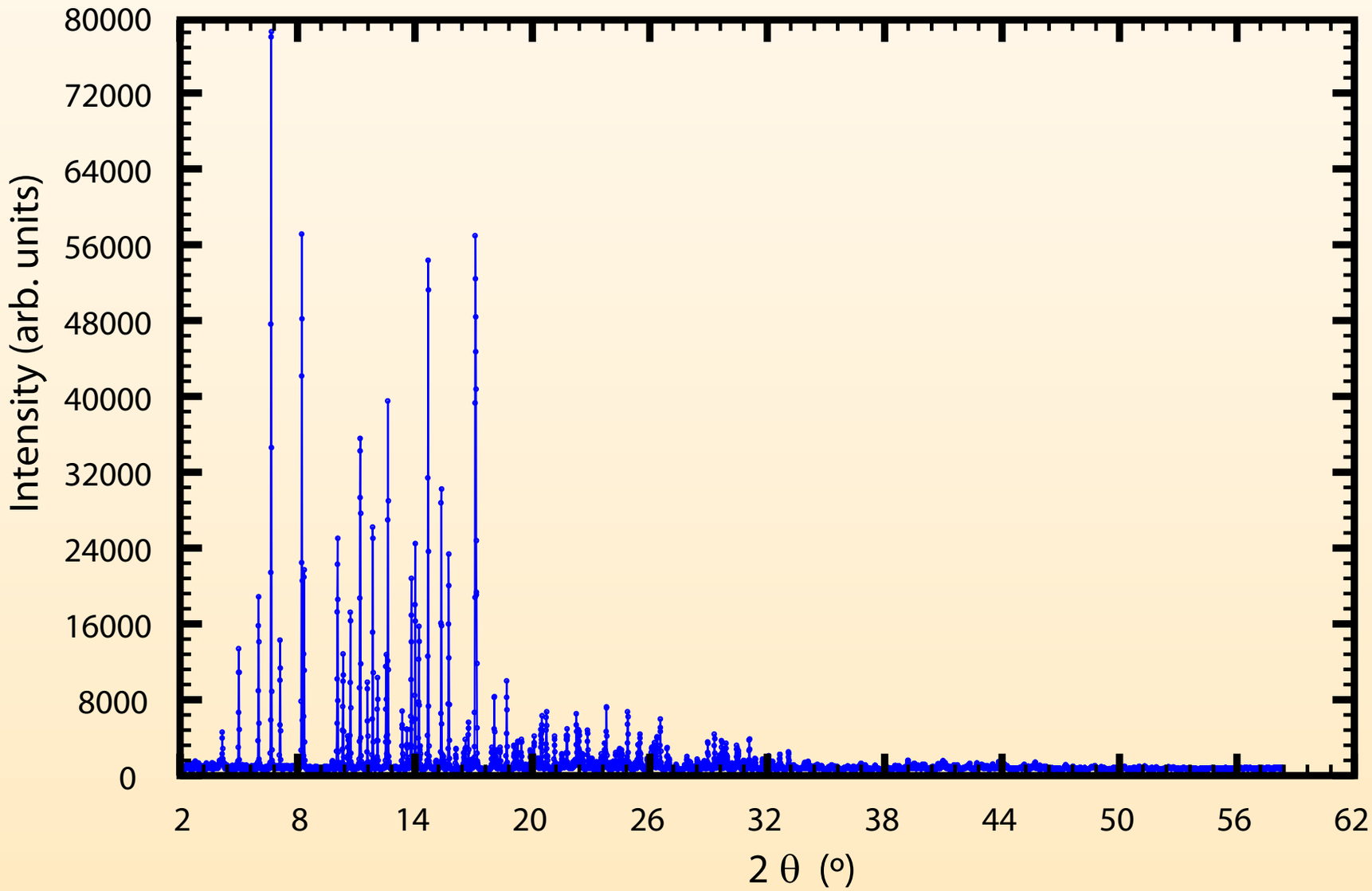
ICMAB
Institut de Ciència de Materials de Barcelona
www.icmab.es
Campus de la UAB
08193-Bellaterra (Spain)

ALBA
CELLS - ALBA
Synchrotron Light Facility

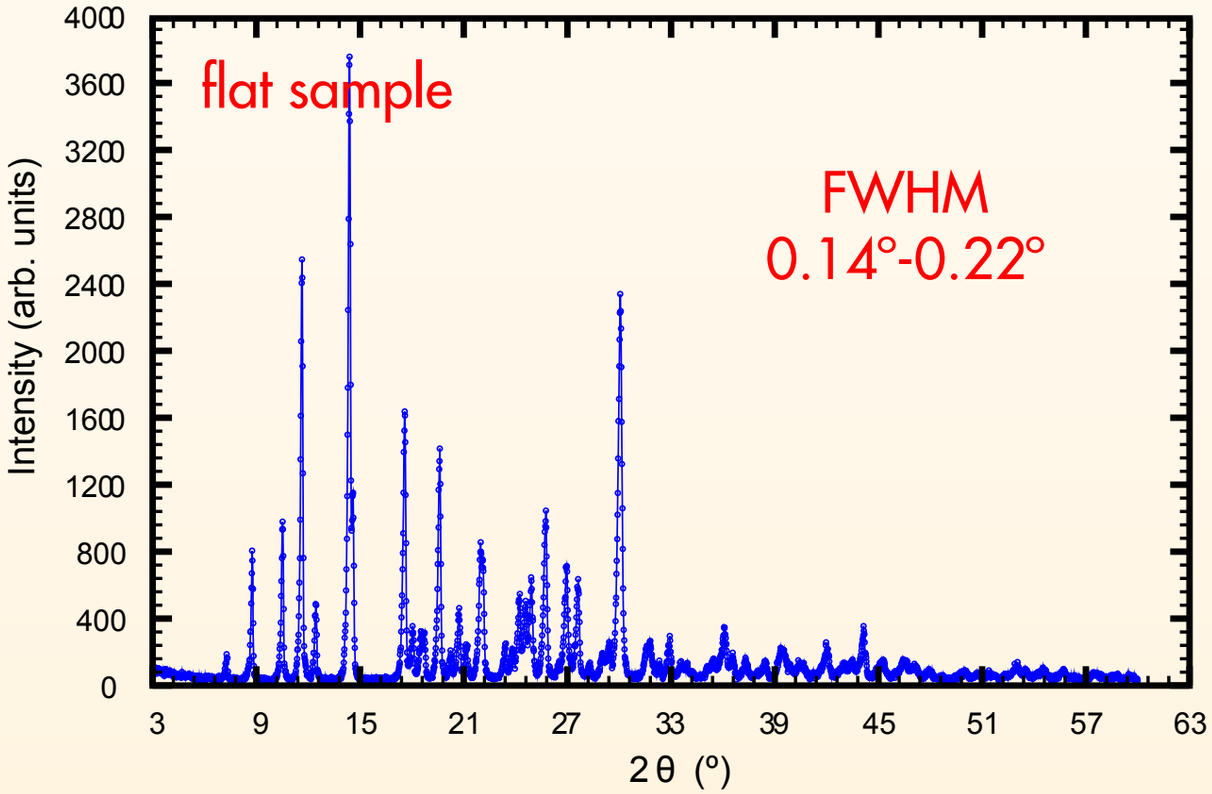
(-)-Riboflavin

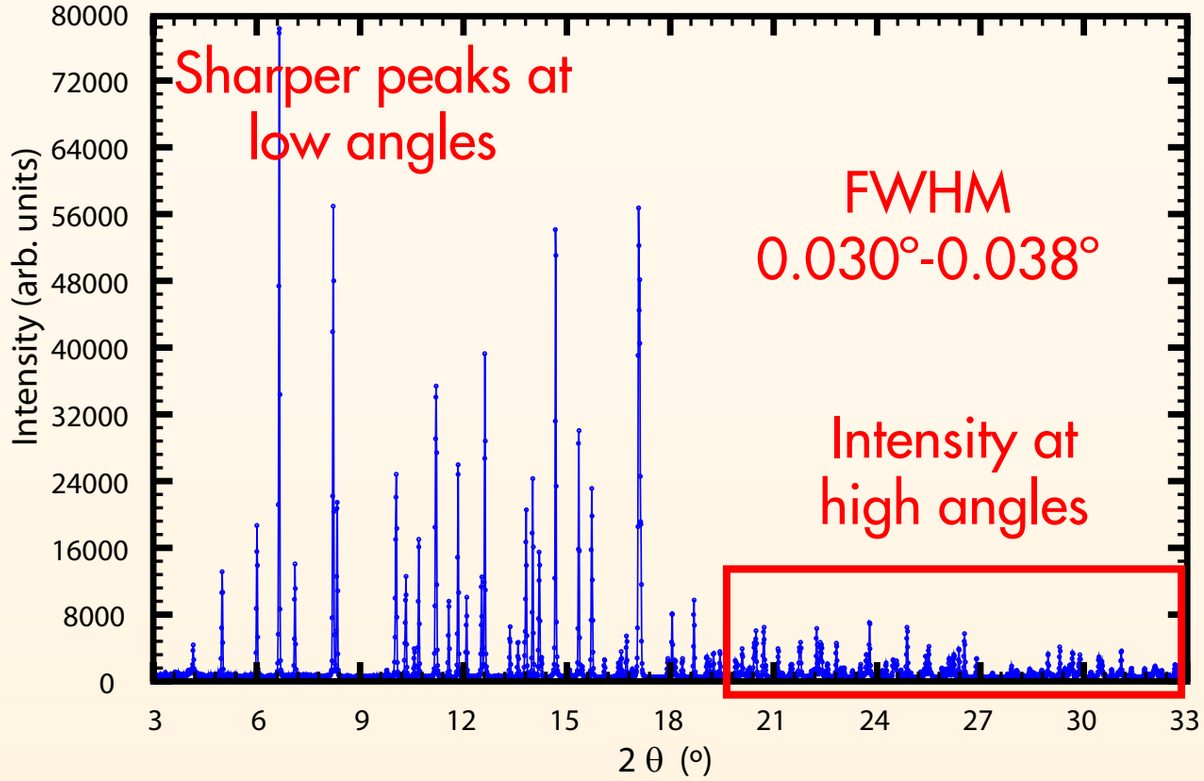
- Capillary (0.5mm) synchrotron data (ALBA-MYTHEN II)
- 8 patterns collected (180s/pattern)
- λ = 0.88320 Å
- $C_{17}H_{20}N_4O_6$



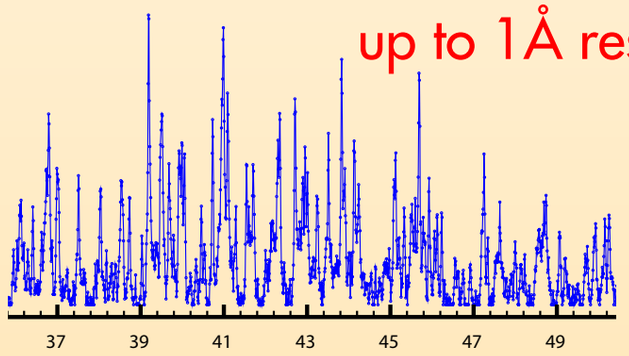


Lab vs Synchrotron PDD



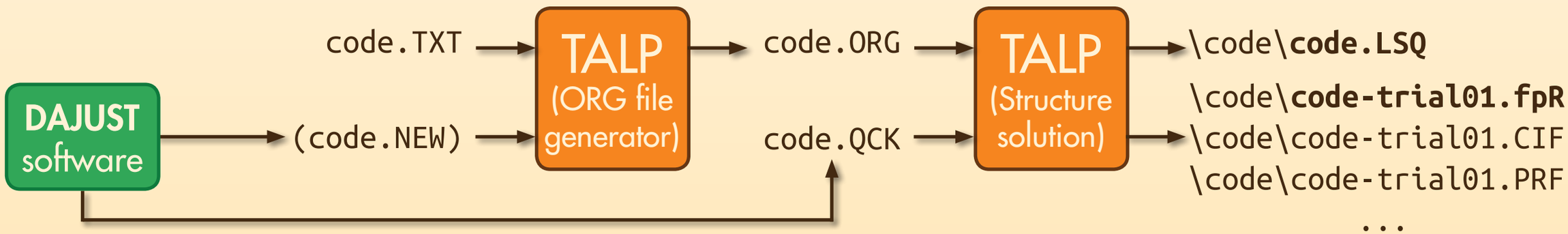


~	Indexing	✓
✓	**TALP structure solution**	✓ ✓
~	Structure Refinement	✓



Steps

1. Indexing
2. Pattern matching
3. Molecular model
4. Restraints, FRB
5. TALP application
6. Check results
7. Rietveld Refinement



Indexing

- e.g. DICVOL

```
(...)

   THE SOLUTION IS NOW USED TO TRY TO INDEX ALL INPUT  56 LINES :
   ============================================================
  DIRECT PARAMETERS :    A= 20.17036  B= 15.17811  C=  5.35121   VOLUME= 1638.26
  STANDARD DEVIATIONS :       .00034    .00029    .00010
  REFINED ZERO-POINT SHIFT : 0.0181 deg. 2-theta

    H   K   L    DOBS     DCAL    DOBS-DCAL 2TH.OBS  2TH.CAL  DIF.2TH.

    1   1   0  12.17640 12.18077  -0.00438    4.157    4.155   0.001
    2   0   0  10.12641 10.12167   0.00473    4.999    5.001  -0.002
    2   1   0   8.42660  8.42523   0.00137    6.008    6.009  -0.001
    0   2   0   7.61019  7.60968   0.00051    6.653    6.654   0.000

(...)

          * NUMBER OF LINES
               .- LINES INPUT      =   56
               .- LINES INDEXED    =   56
               .- LINES CALCULATED =  224
          * AVERAGE 2-Theta DIFFERENCE FOR INDEXED LINES =-0.0007
          * MEAN ABSOLUTE DISCREPANCIES
                                          <Q> =0.2522E-04
                              <DELTA(2-THETA)> =0.1880E-02
              MAX. ERROR ACCEPTED (DEG. 2-THETA) =0.4500E-01
          * FIGURES OF MERIT
                1.- M( 20) =   82.2
                2.- F( 20) =  345.4(0.0016,   36)
                3.- F( 56) =  133.0(0.0019,  224)
```

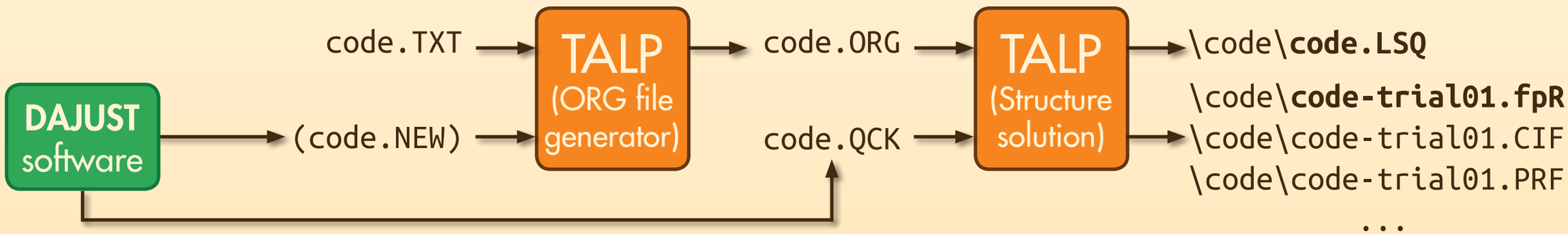
DAjust requires the approximate cell parameters:

a	20.17 Å	α	90°
b	15.18 Å	β	90°
c	5.35 Å	γ	90°

V 1638 Å^3

Orthorhombic

$Z=4$ $Z'=1$



DAJUST pattern matching

a. Initial cell parameters:
   20.17Å 15.18Å 5.35Å 90.0° 90.0° 90.0°

b. PXRD data format: Pattern=1 (NYLINE=10)
   $2\theta_{sup}$=55°, Synchrotron (λ = 0.88320 Å)

c. Pattern matching parameters (pV, zero, L G coef.)

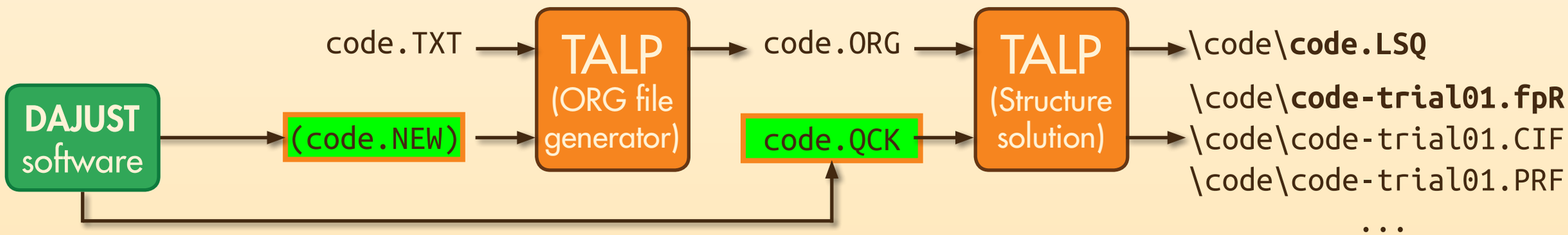
d. Run DAJUST (highest symmetry)

e. Run SGAID (sg. $P2_12_12_1$)

f. Run DAJUST

g. Copy QCK file to TALP folder

Seen yesterday in Data Processing talk (DAjust practical examples)



DAJUST pattern matching

ICMAB CSIC
Consejo Superior de Investigaciones Científicas

DAjust_UI_v1209

File Options Help

General

Name (-)-Riboflavin
Formula C17 H20 N4 O6
Z 4
Cycles 20

Cell parameters

a 20.1714 α 90.000
b 15.1786 β 90.0
c 5.351 γ 90.000
Metric orthorombic
Fix cell param.
Use existing LAT file

Symmetry

SG.Num:Symbol 19 : P 21 21 21 set Lattice P
Centrosimetric
No centrosimetric
Highest symmetry
Symmetry matrices unlock

```
X+1/2,-Y+1/2,-Z
-X,Y+1/2,-Z+1/2
-X+1/2,-Y,Z+1/2
```

Instrumental & Data parameters

Pattern type 1 (Yobs) nYline 10 2θsup 55.0000
Rad. source Synchrotron λ1 0.88320 λ2 0.00000
Intensity Extraction NFWHM 20 tolerance 0.5000
Sample Flat Capillary
Primary monochromator
Yes No cos 2θ 1.00000
Divergence correction
Yes No Fixed slit Variable slit
slit-w (º) 0.0000 sample size 0.0000 armlenght 0.0000
Absorption correction
Yes No
R*mu 0.0000 2θRefInt 0.0000 Transmi. 0.0000

Profile Function

Type Pseudo-Voigt T... zero -0.018162 exp 1.0000
Lorentzian coef. X 0.265E-01 Y 0.247E-02
Gaussian coef. W 0.793E-03 V 0.000E+00 U 0.000E+00
Pref. Orientation: March-coef 1.000 direction (h k l) 0 0 1

Excluded zones

2θinf Add
2θsup Del.

2Tinf	2Tsup

Background

type Removed 2θ
<back> 60.00000 bg-value
FWHM 0.15000
Iterations 20

2T

Asymmetry correction

coef. Add
2θsup Del.

coef	2Tsup

Ghost peaks

value 2θ
η asym

value	2T	width

Run AJUST | Stop run | open PRF | Output window
Run SGAid | Reload JST | SGAid window

Work file: O:\ovallcorba\riboflavin\Riboflavin.jst
Data file: O:\ovallcorba\riboflavin\Riboflavin.net

SGAid Output

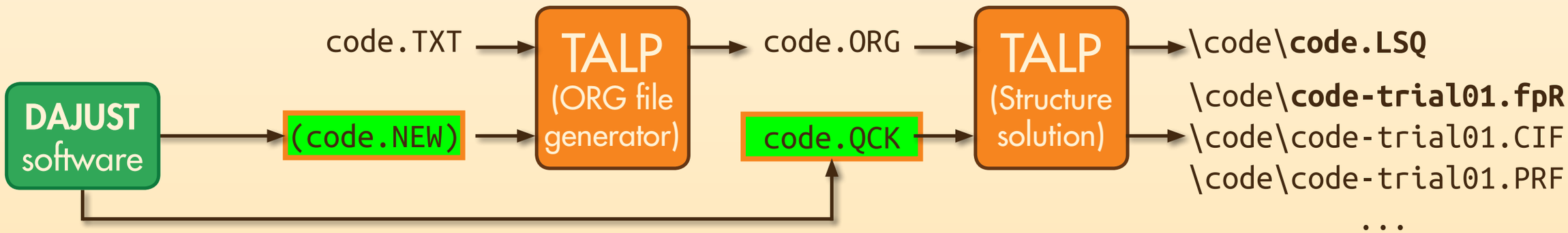
Click on a row to load symmetry matrices for the S.G into main window
on top

SGNum	H-M symbol	RW (%)	NREF	Setti...
25	Pmm2	**18,96**	1205	abc
47	Pmmm	**18,96**	1205	abc
16	P222	**18,96**	1205	abc
17	P222(1)	**18,97**	1197	acb
18	P2(1)2(1)2	**18,99**	1194	bca
19	P2(1)2(1)2(1)	**19,06**	1183	abc
51	Pmma	**19,38**	1165	bca
26	Pmc2(1)	**19,38**	1165	cab
28	Pma2	**19,38**	1165	bac
31	Pmn2(1)	21,99	1165	bac
59	Pmmn	21,99	1165	bca

DAjust_UI Output

Font Size: + -
on top

```
[12:13]   17  0.1903E+00  0.1017E+00  0.5543E-01  0.5404E+01
[12:13]   18  0.1903E+00  0.1016E+00  0.5543E-01  0.5404E+01
[12:13]   19  0.1903E+00  0.1016E+00  0.5543E-01  0.5404E+01
[12:13]   20  0.1903E+00  0.1016E+00  0.5543E-01  0.5404E+01
[12:13]
[12:13]   ZERO-SHIFT:
[12:13]   -0.01843   0.00011
[12:13]   FINAL LATTICE PARAMETERS:
[12:13]   20.17060  15.17828   5.35091   90.0000   90.0000   90.0000
[12:13]
[12:13]      Volume=     1638.209
[12:13]
[12:13]   ST. DEVS. OF LATTICE PARAMETERS:
[12:13]     1     0.00021
[12:13]     2     0.00016
[12:13]     3     0.00005
[12:13]   PEAK WIDTH PARAMETERS:
[12:13]     0.270E-01   0.138E-02  -0.206E-05
[12:13]     0.242E-02   0.179E-03   0.362E-06
[12:13]     0.792E-03   0.812E-05  -0.137E-07
[12:13]   SCALA FACTOR:
[12:13]    0.9998E+00  0.2864E-02 -0.2065E-03
[12:13]  ------------------------------------------------------------------
[12:13]  Cycle   Rwp  Rp   RB   Chi(planes) Chi(dist)   Chi(I)
[12:13]   20  0.1903E+00  0.1016E+00  0.5543E-01  0.5404E+01

[12:13] New χ = 5.404  (previous χ = 5.381)
[12:13] Worse χ, fields not updated. To load new values click on load NEW.
[12:13] dAjust finished
```



Molecular Model

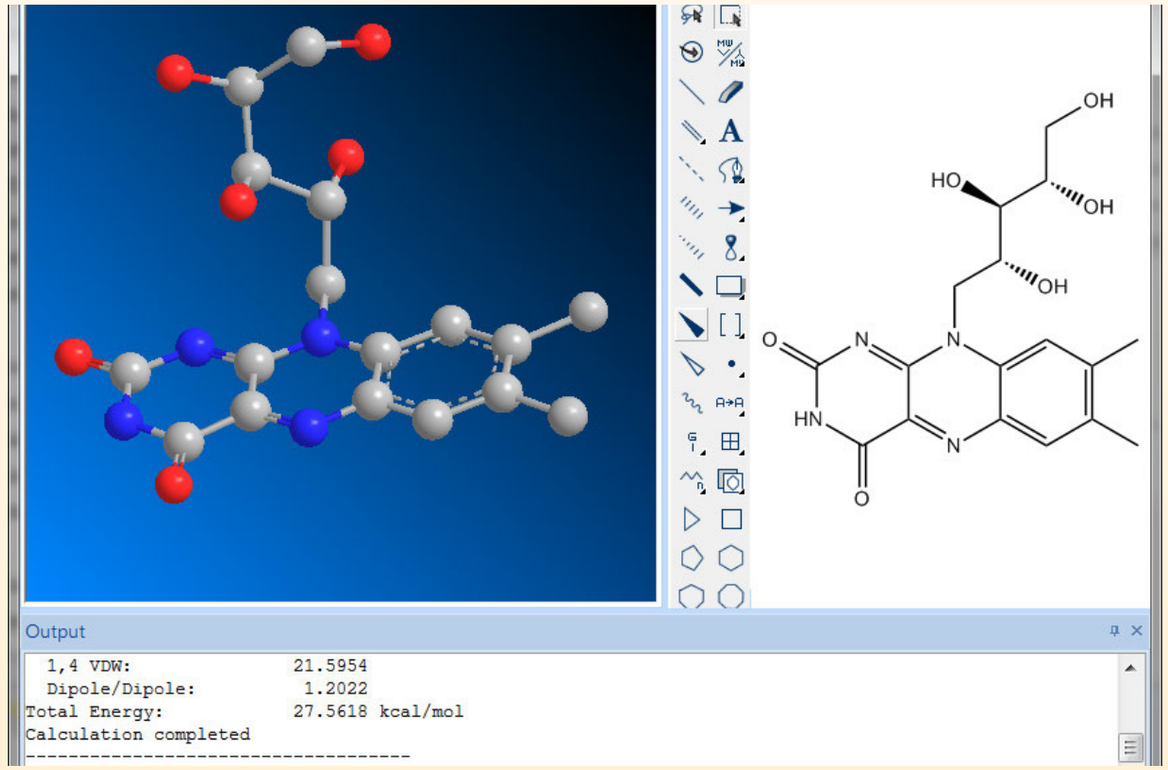
Molecular Model from calculations (MM)

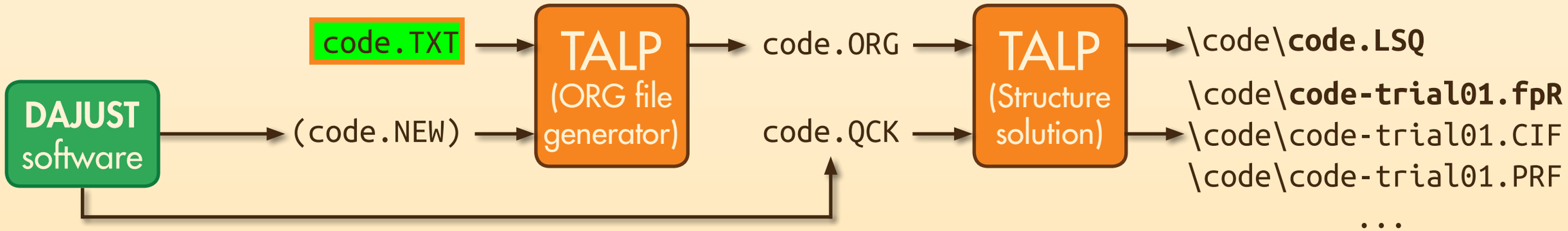
i.e. Avogadro (MMFF94, UFF), **Chem3D (MM2)**,...

code.TXT

```
!(-)riboflavin
1 1 1 90 90 90
C1    C    -0.00858    1.30554    0.51270
C2    C     0.56154   -2.39206   -0.06907
C3    C     2.71448    1.18584   -2.69982
C4    C    -1.38964    1.62446   -0.05307
C5    C    -1.98857    2.84068    0.59021
...
```

Recomendation: first atom near mass center

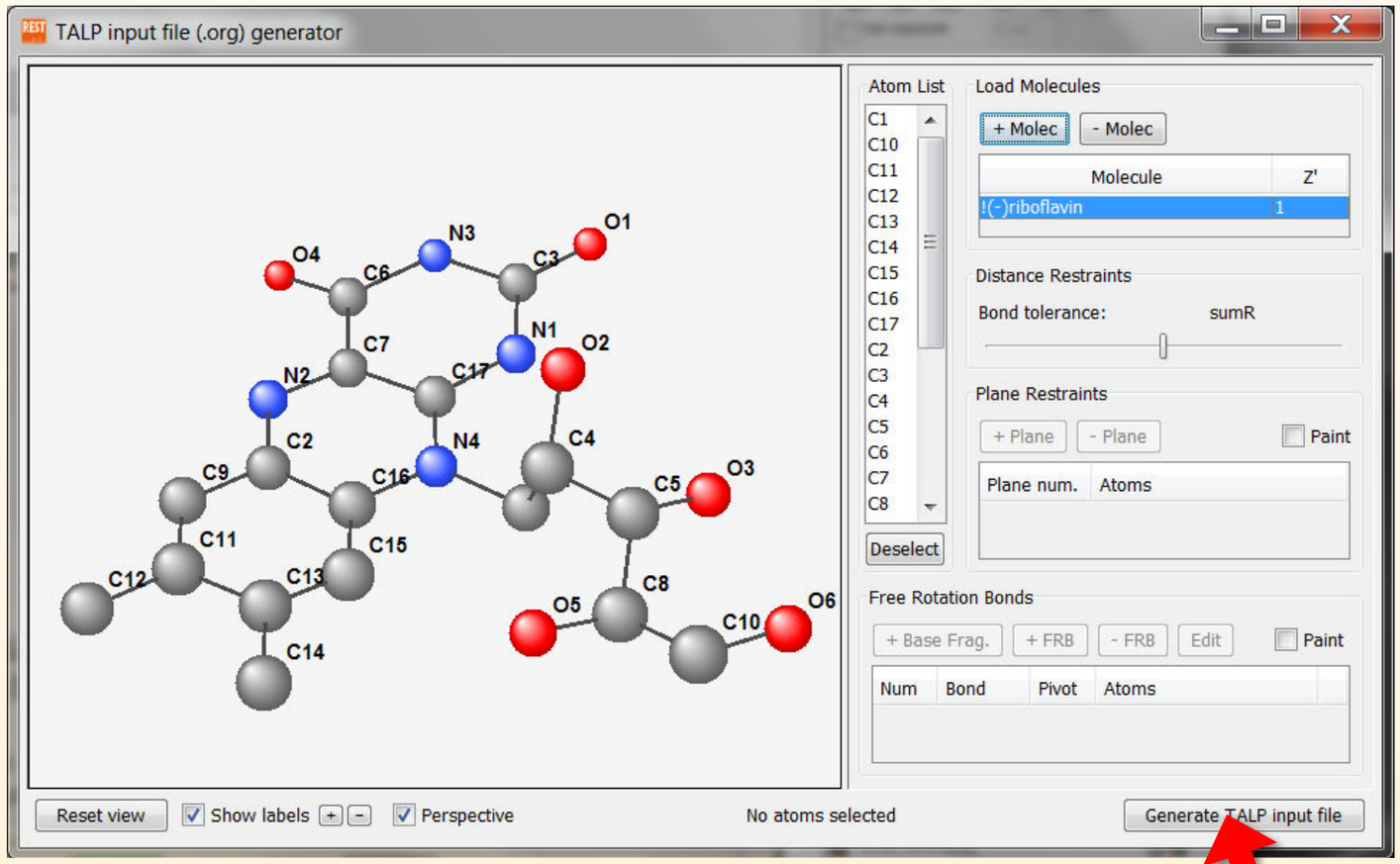




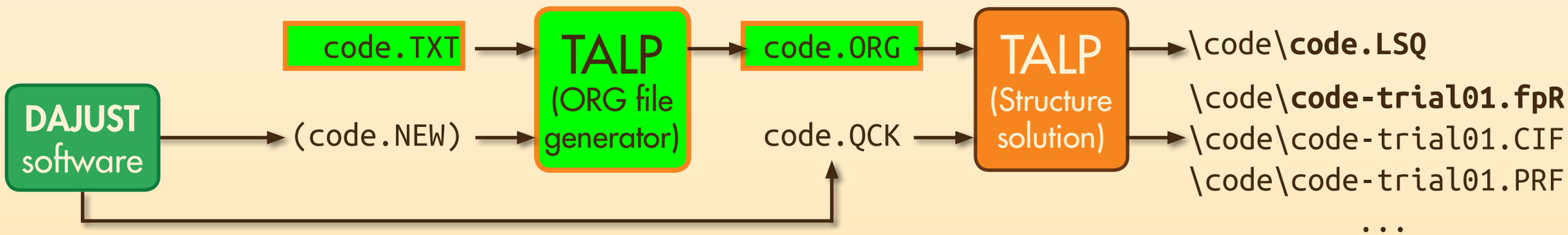
TALP input file (.ORG) generator

ICMAB CSIC
Consejo Superior de Investigaciones Científicas

Indexing - Pattern matching - Molecular model - Restraints, FRB - TALP application - Check results - Rietveld Refinement

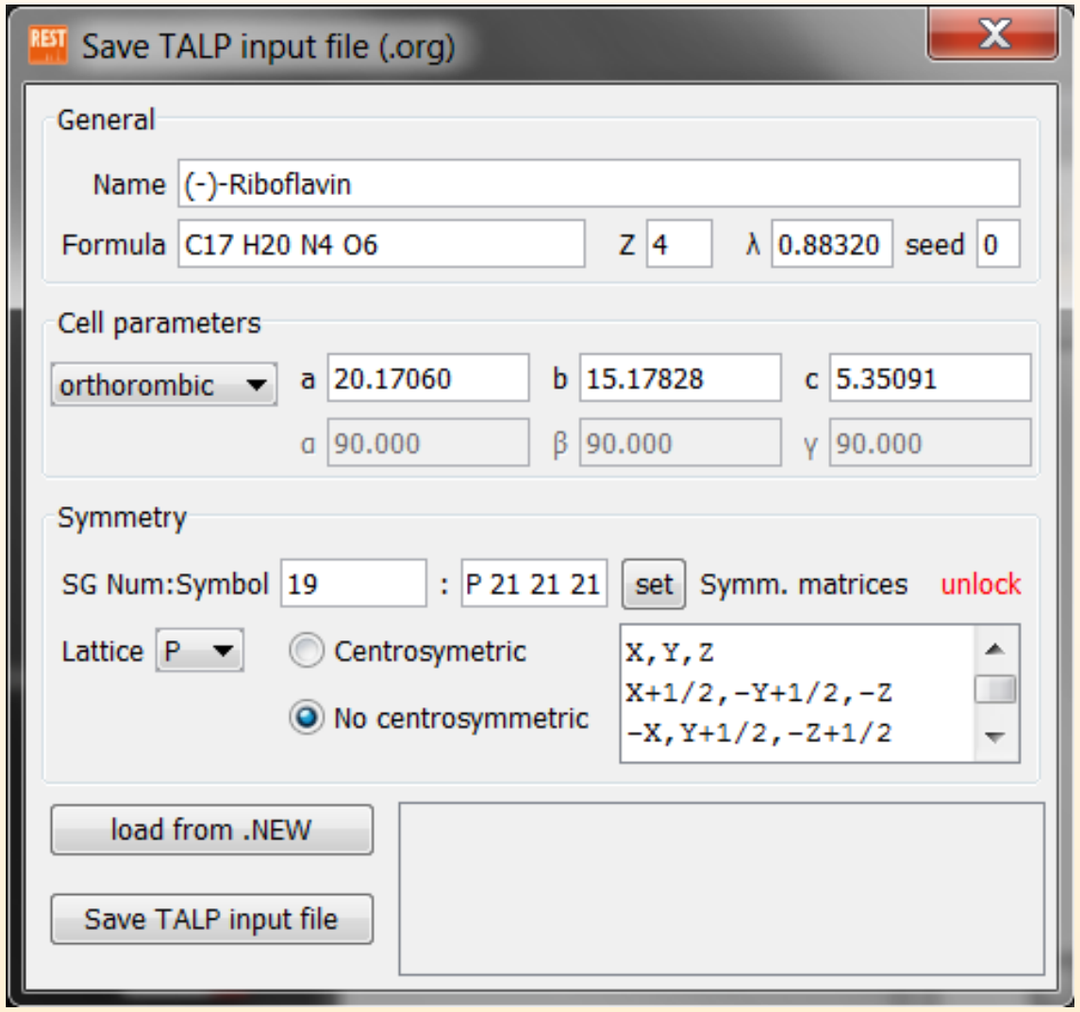


a. Import model coordinates (Z'=1)
b. Distance Restraints (bond tolerance)
c. Plane restraints
d. Free Rotation Bonds
e. Generate TALP input file

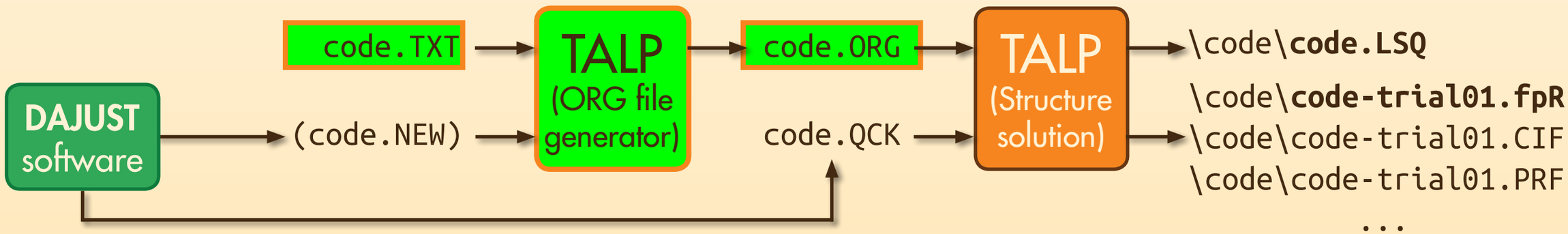


TALP input file (.ORG) generator

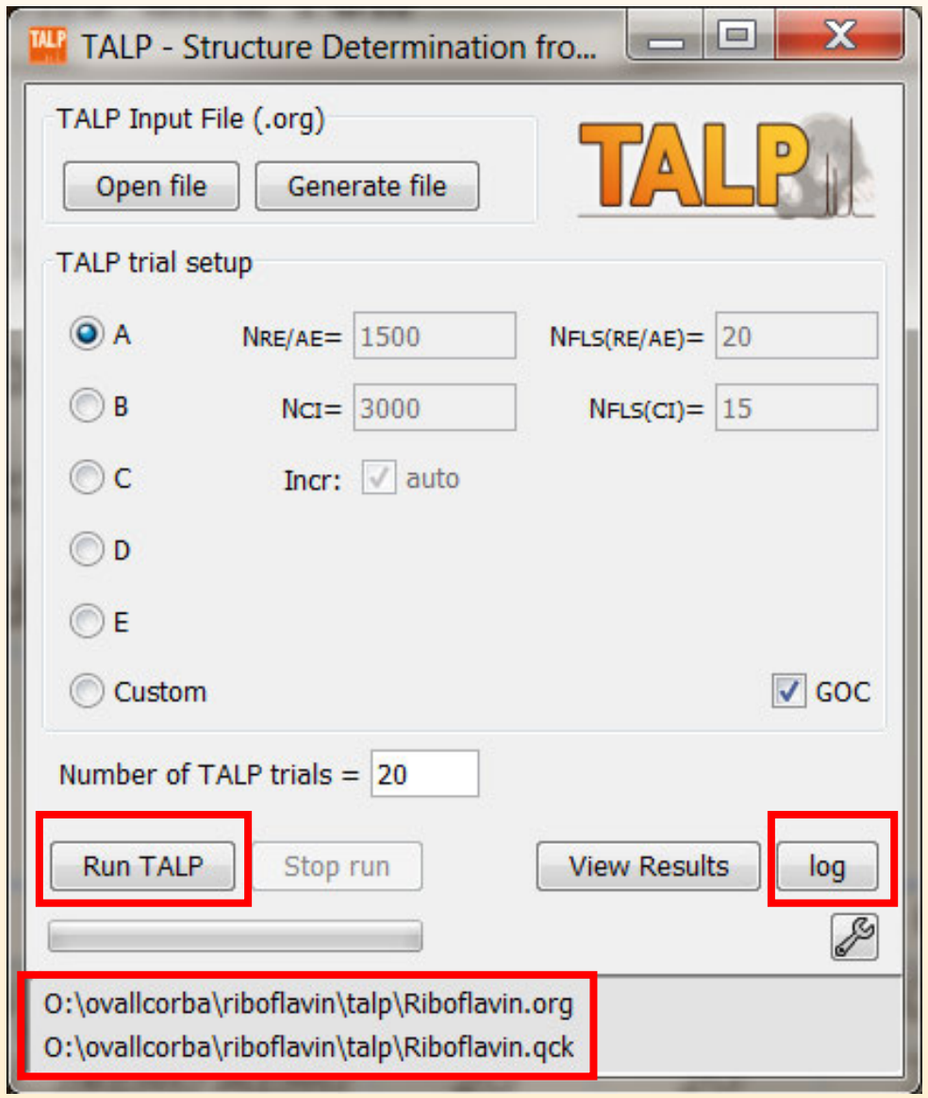
Indexing - Pattern matching - Molecular model - Restraints, FRB - TALP application - Check results - Rietveld Refinement



- Option load from NEW file (DAjust)
- Same code/location QCK
  (if not the program asks for and copy it to the save location)

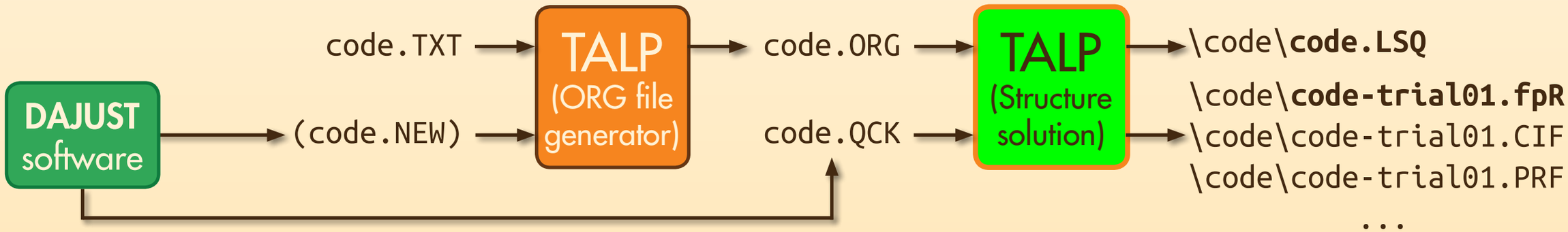


TALP experiment conditions and run



a. Preset/Custom conditions
b. N trials
c. GOC
d. Log window

	A	B	C	D	E
N_{rnd}	1500	3000	6000	10000	20000
$N_{FLS}(REM/AEM)$	20	20	30	30	30
N_{CI}	3000	6000	8000	15000	30000
$N_{FLS}(CI)$	15	15	15	15	15



TALP search strategy (trial flow-sheet)

Starting model

N_{rand} generated models (position, orientation, conformation), checked for overlap and refined by FLS

General scan

Random Exploration Mode (REM)

Assisted Exploration Mode (AEM)

N_{rand} generated models (position, conformation), checked for overlap and refined by FLS

$$\chi_{TALP} = \left(\frac{S_Y}{N_H + N_{REST} - P} \right)^{1/2}$$

Quality of intermediate models after FLS

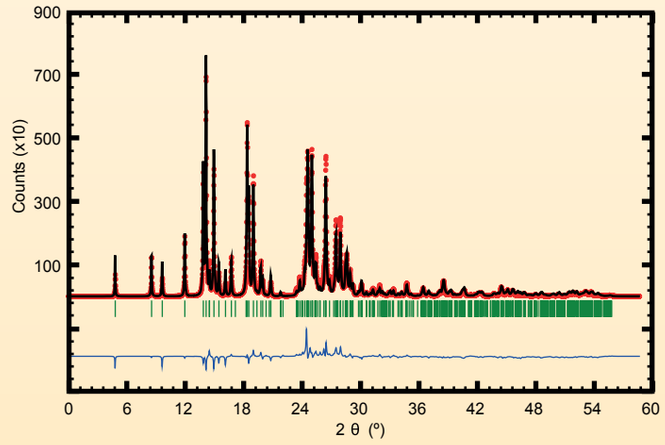
Local Incremental Scan (LIS)

N_{CI} increments of limited size to the best model (position, orientation, conformation) followed by FLS refinement

$$Q_{trial} = \frac{R_{wp,FPR}}{R_{wp,model-free}}$$

Indicator for candidate solutions (usually below 1.5 for correct solutions)

Fixed-Profile Rietveld refinement (FPR)



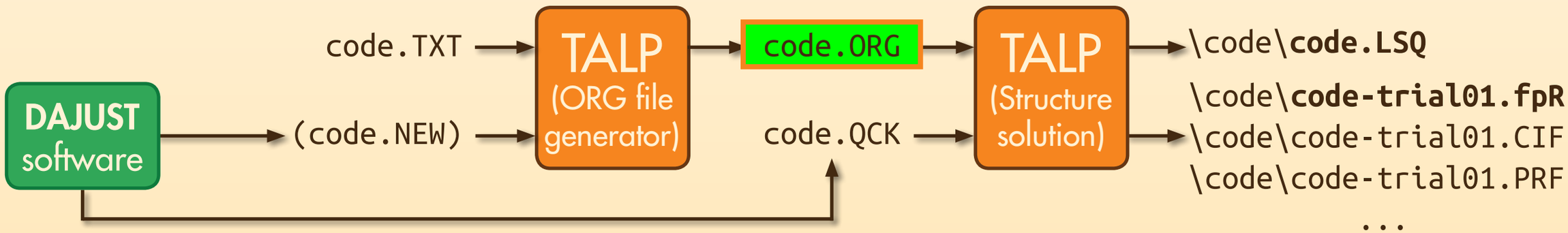
Final model ($\boldsymbol{Q_{trial}}$)

ORG file details

Additional control/options

- Cell, symmetry, content,...
- Scale, Bover
- Atoms/Fragments (parameters, link, rigid bodies, FR,...)
- Molecular groups
- FRB: Free Rotation Bonds (limits REM, initial value,link,...)
- REST: Restraints (sdev, symmetry related,...)
- DIST: Distance calculation
- Options: RadType, HKL, FR, Seed
- 2TEX: Excluded zones

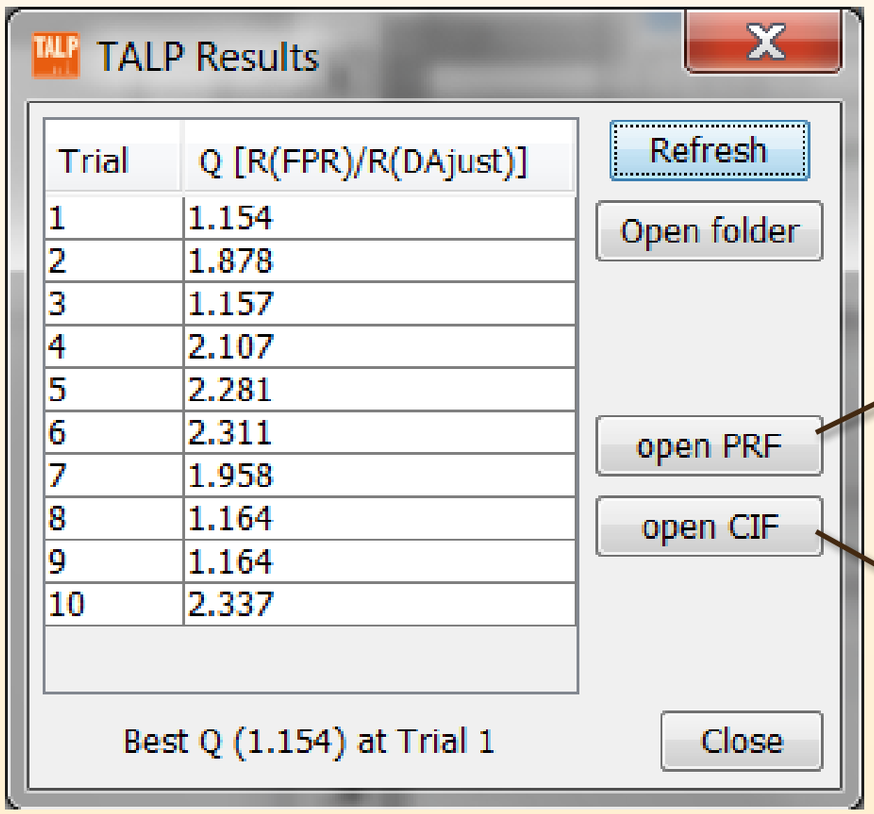
Automatically generated by TALP (standard)

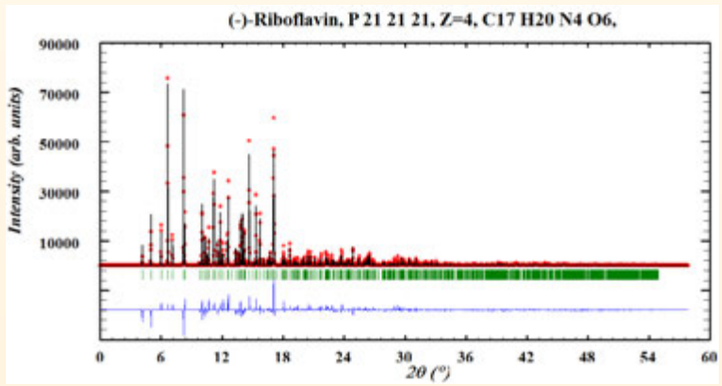


TALP output

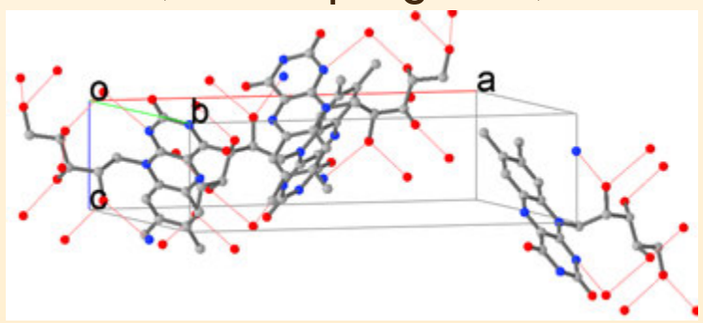
Results window and output files

- LSQ file: general
- fpR files: trial
- CIF/PRF visual inspection





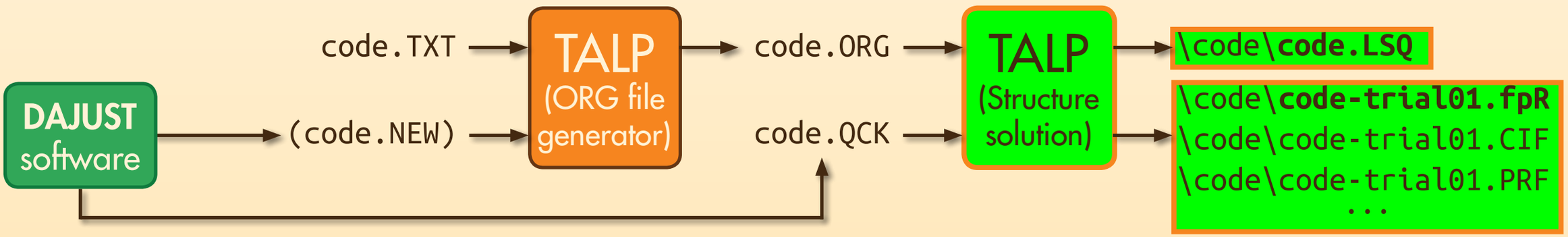
(default programs)



Remember!

$$Q_{trial} = \frac{R_{wp,FPR}}{R_{wp,model-free}}$$

Indicator for candidate solutions (usually near 1.0 for correct solutions)



Check results

SGAid Output

Click on a row to load symmetry matrices for the S.G into main window

on top

SGNum	H-M symbol	RW (%)	NREF	Setti...
25	Pmm2	**18,96**	1205	abc
47	Pmmm	**18,96**	1205	abc
16	P222	**18,96**	1205	abc
17	P222(1)	**18,97**	1197	acb
18	P2(1)2(1)2	**18,99**	1194	bca
19	P2(1)2(1)2(1)	**19,06**	1183	abc
51	Pmma	**19,38**	1165	bca
26	Pmc2(1)	**19,38**	1165	cab
28	Pma2	**19,38**	1165	bac
31	Pmn2(1)	21,99	1165	bac
59	Pmmn	21,99	1165	bca
32	Pba2	22,51	1113	abc
55	Pbam	22,51	1113	abc
62	Pnma	23,65	1112	bca
33	Pna2(1)	23,65	1112	bac
57	Pbcm	25,10	1115	bac
29	Pca2(1)	25,10	1115	abc
58	Pnnm	25,53	1115	abc
34	Pnn2	25,53	1115	abc
53	Pmna	26,08	1117	cba
30	Pnc2	26,08	1117	bac
27	Pcc2	27,55	1121	abc
49	Pccm	27,55	1121	abc
50	Pban	34,21	991	abc

a=20.1714 b=15.1786 c=5.351 α=90.0 β=90.0 γ=90.0

P mc2(1)

TALP Results

Trial	Q [R(FPR)/R(DAjust)]
1	4.501
2	4.075
3	4.285
4	4.129
5	4.513
6	4.152
7	4.152
8	4.158
9	4.524
10	4.084

Refresh | Open folder | open PRF | open CIF

Best Q (4.075) at Trial 2 | Close

P ma2

TALP Results

Trial	Q [R(FPR)/R(DAjust)]
1	3.342
2	3.008
3	3.173
4	3.256
5	2.958
6	3.044
7	3.378
8	3.098
9	3.139
10	3.324

Refresh | Open folder | open PRF | open CIF

Best Q (2.958) at Trial 5 | Close

P 2(1)2(1)2(1)

TALP Results

Trial	Q [R(FPR)/R(DAjust)]
1	1.154
2	1.878
3	1.157
4	2.107
5	2.281
6	2.311
7	1.958
8	1.164
9	1.164
10	2.337

Refresh | Open folder | open PRF | open CIF

Best Q (1.154) at Trial 1 | Close

Sometimes you have to go back...

Numeric result files (LSQ and fpR)

.LSQ

General results file (text):

- Copy of input file
- Data read by TALP
- TALP experiment info
- Trial Results (Qtrial)

```
...
TALP EXPERIMENT INFO:
 Program version= TALP1209
 Filename= Riboflavin
 Experiment start date:time= [130326:1258]
 nRE= 1500 (nFLS= 20)
 nCI= 3000 (nFLS= 15)
 Automatic Increments
-------------------------------------------------
[130326:1258] TRIAL: 1 of 20
 Chi(Trial) =  1.132 (Experiment best=  1.132)
 Q(Trial)   =  1.975 (Experiment best=  1.975)
 time(trial)=    3.9 min
-------------------------------------------------
[130326:1302] TRIAL: 2 of 20
 Chi(Trial) =  0.904 (Experiment best=  0.904)
 Q(Trial)   =  1.275 (Experiment best=  1.275)
 time(trial)=    3.9 min
-------------------------------------------------
...
```

.fpR

Trial result files (text):

- Figures of Merit
- Refined parameters (coordinates,...)
- Dist/Angle Calculations

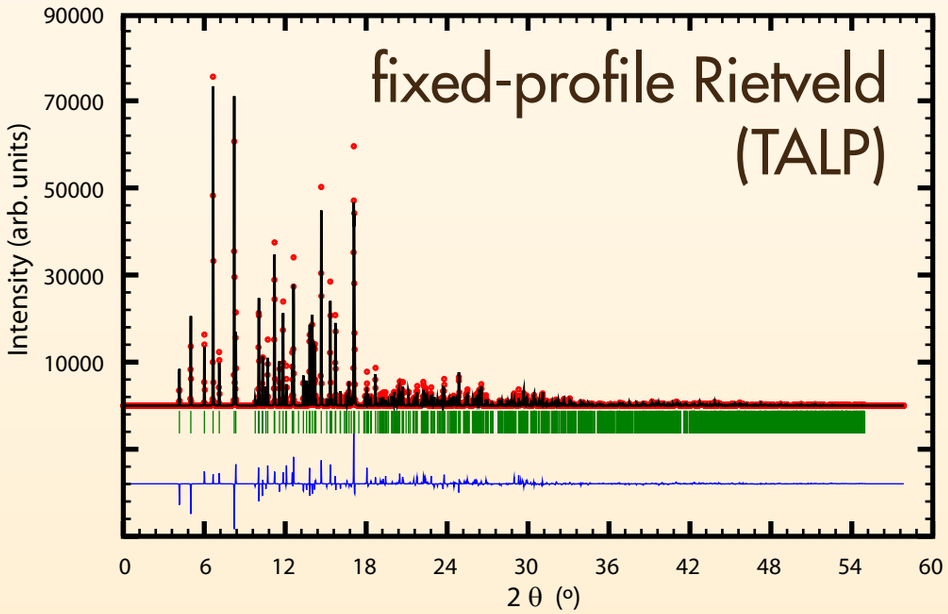
```
-------------------------------------------------------------
FIXED-PROFILE RIETVELD REFINEMENT OUTPUT (trial solution):
-------------------------------------------------------------
(-)-Riboflavin, P 21 21 21, Z=4, C17 H20 N4 O6,
-------------------------------------------------------------
  Rwp         Q(Trial)
  0.2426E+00  0.1275E+01
  Rexp        Rwm(planes) Rwd(dist)
  0.3742E-01  0.3712E+01  0.8833E+00
  Scale       Bover
  0.2993E+00  0.0000E+00
-------------------------------------------------------------
    1   1.1  1.1  1.1         0    0    0         0    0
C1a      1   0.96710   0.60087   1.56000   1.00000   3.295
    1   1.1  1.1  1.1         0    0    0         0    0
C2a      1   1.13120   0.54129   1.22363   1.00000   3.295
...
-------------------------------------------------------------
 DISTANCES AND ANGLES:
 (no calc options found)
-------------------------------------------------------------
```

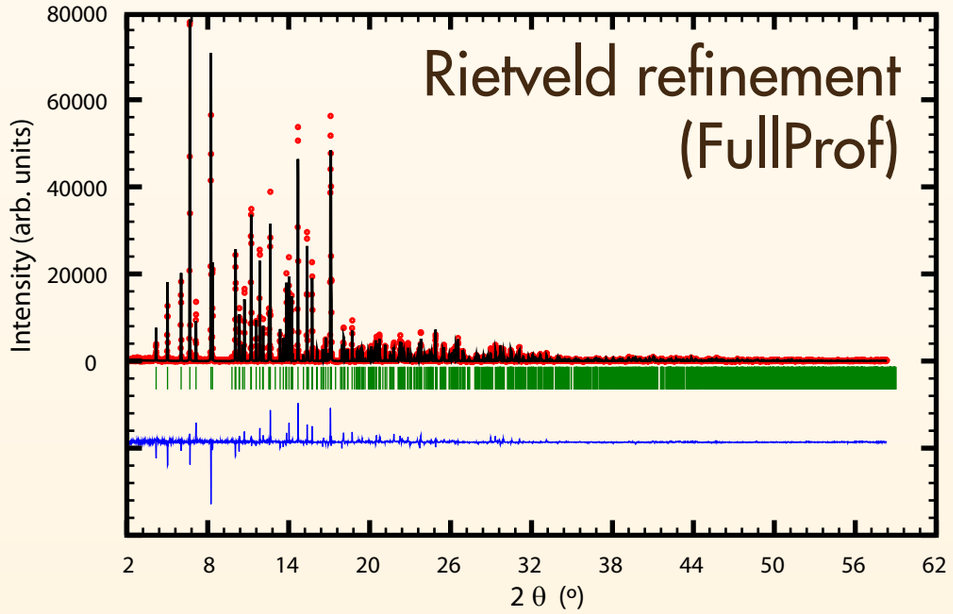
Structure refinement

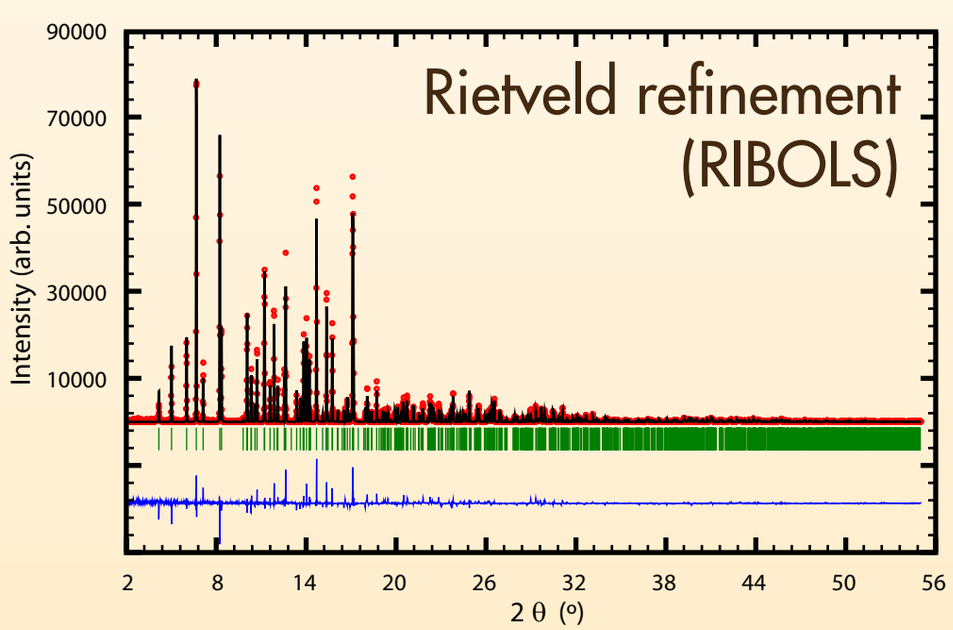
TALP = Structure solution

↓

A final Rietveld refinement is needed





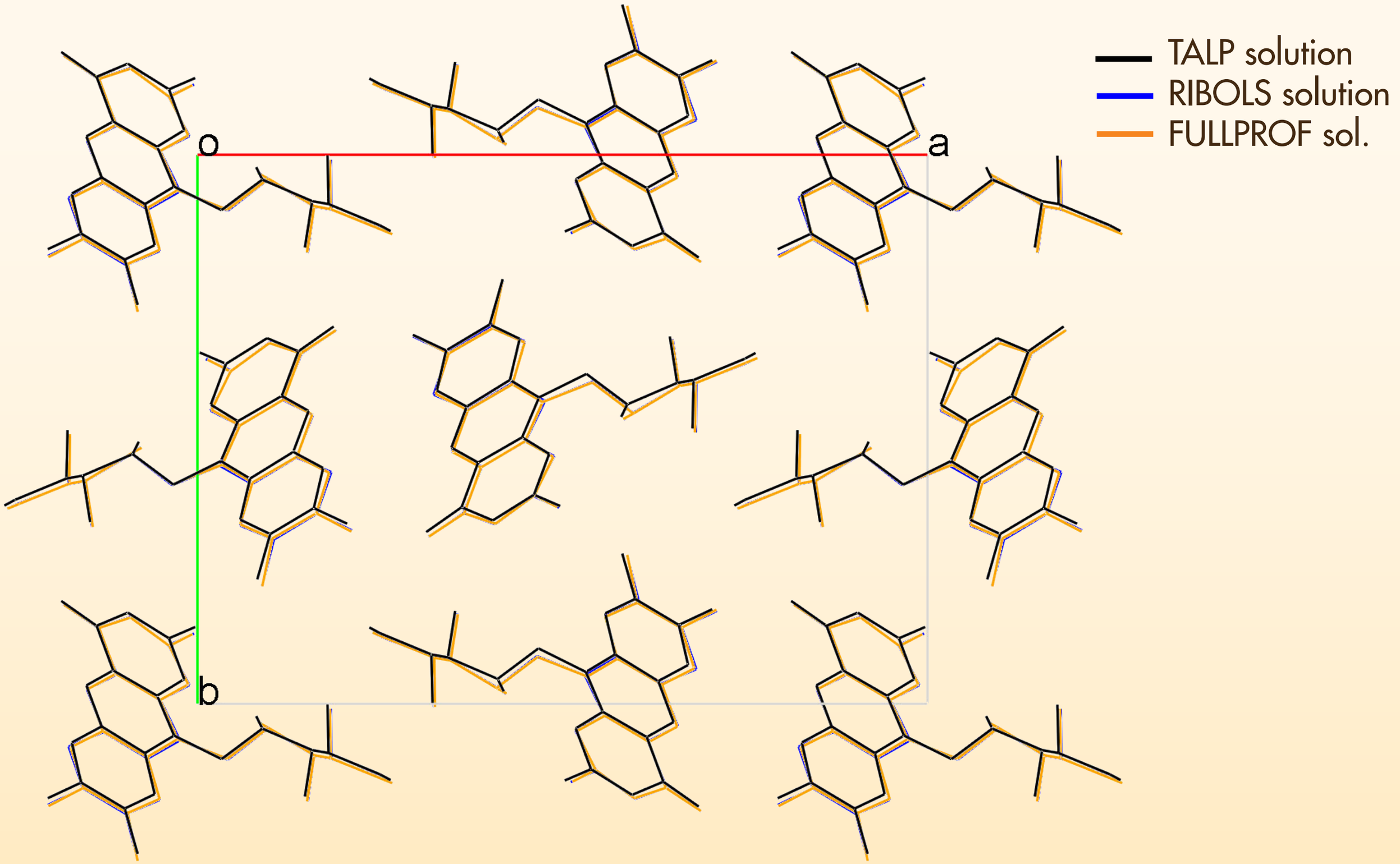


Cell parameters:

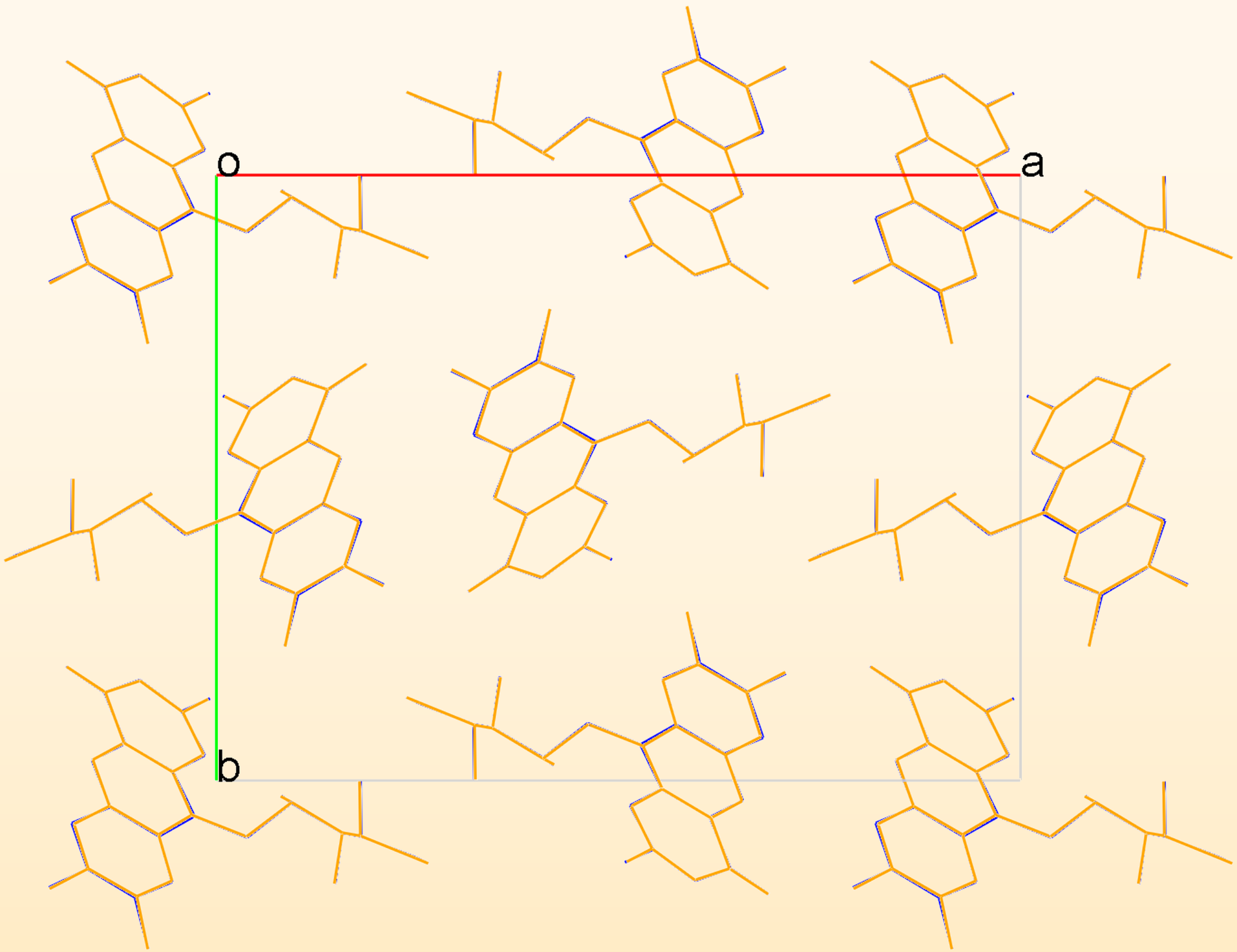
DAjust:	20.17060(21)	15.17828(16)	5.35091(5)	90.0	90.0	90.0
RIBOLS:	20.17194(28)	15.17888(22)	5.35120(7)	90.0	90.0	90.0
fullProf:	20.17326(22)	15.17961(17)	5.35155(6)	90.0	90.0	90.0

(fast refinements, could be improved...)

(-)-Riboflavin

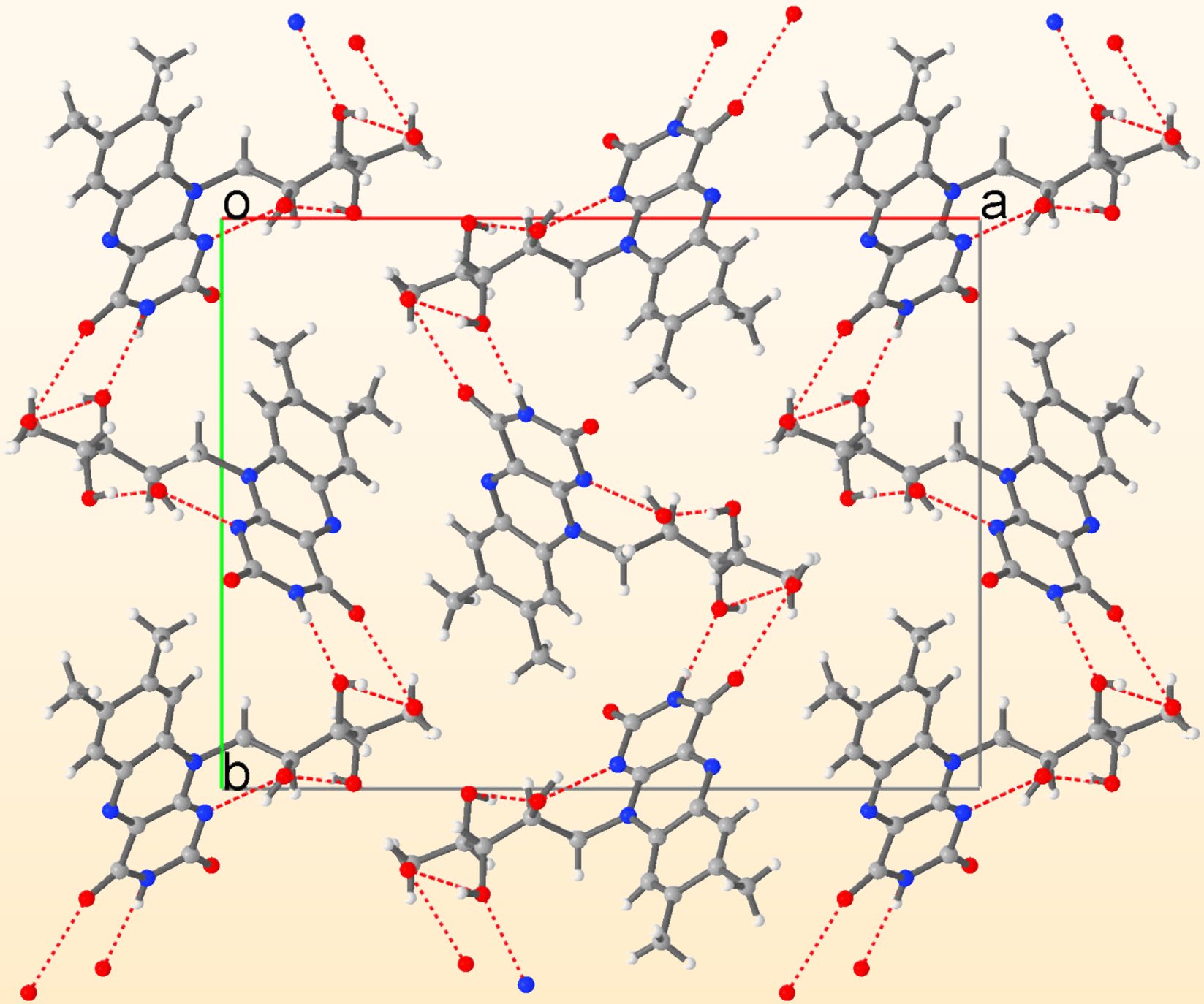


(-)-Riboflavin



RIBOLS solution
FULLPROF sol.

(-)-Riboflavin



- CCDC 906129

TALP

Oriol Vallcorba
Jordi Rius
Carlos Frontera
Carles Miravitlles

J. Appl. Cryst. 2012, 48. 1270-1277

ovallcorba@icmab.com

www.icmab.es/crystallography/software

THANK YOU FOR YOUR ATTENTION!!

Thanks to the ICMAB, the Crystallography Dpt. and specially to Prof. Jordi Rius, Prof. Carles Miravitlles, Dr. Carles Frontera & Dr. Inma Peral for their help and support.

PowderXRD 2013

CSIC
Consejo Superior de Investigaciones Científicas

ICMAB
Institut de Ciència de Materials de Barcelona
www.icmab.es
Campus de la UAB
08193-Bellaterra (Spain)

CELLS - ALBA
Synchrotron Light Facility
ALBA

TALP, a multisolution direct-space strategy for solving molecular crystals from powder diffraction data based on restrained LS

TALP

Oriol Vallcorba
Jordi Rius
Carlos Frontera
Carles Miravitlles

ovallcorba@icmab.com

Part 2

Software demonstration

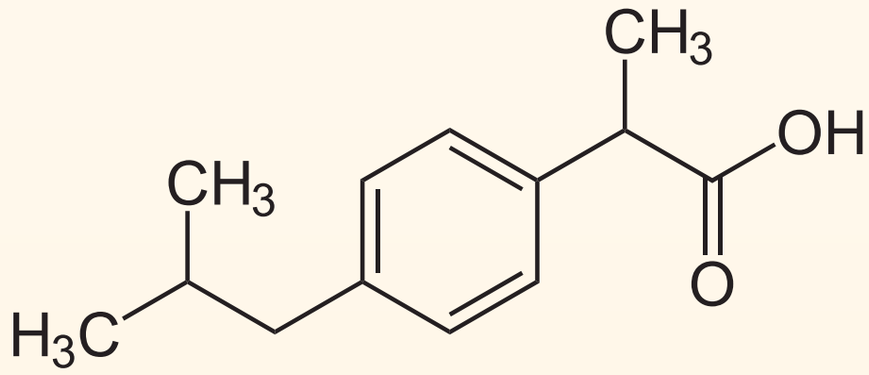
CSIC
Consejo Superior de Investigaciones Científicas

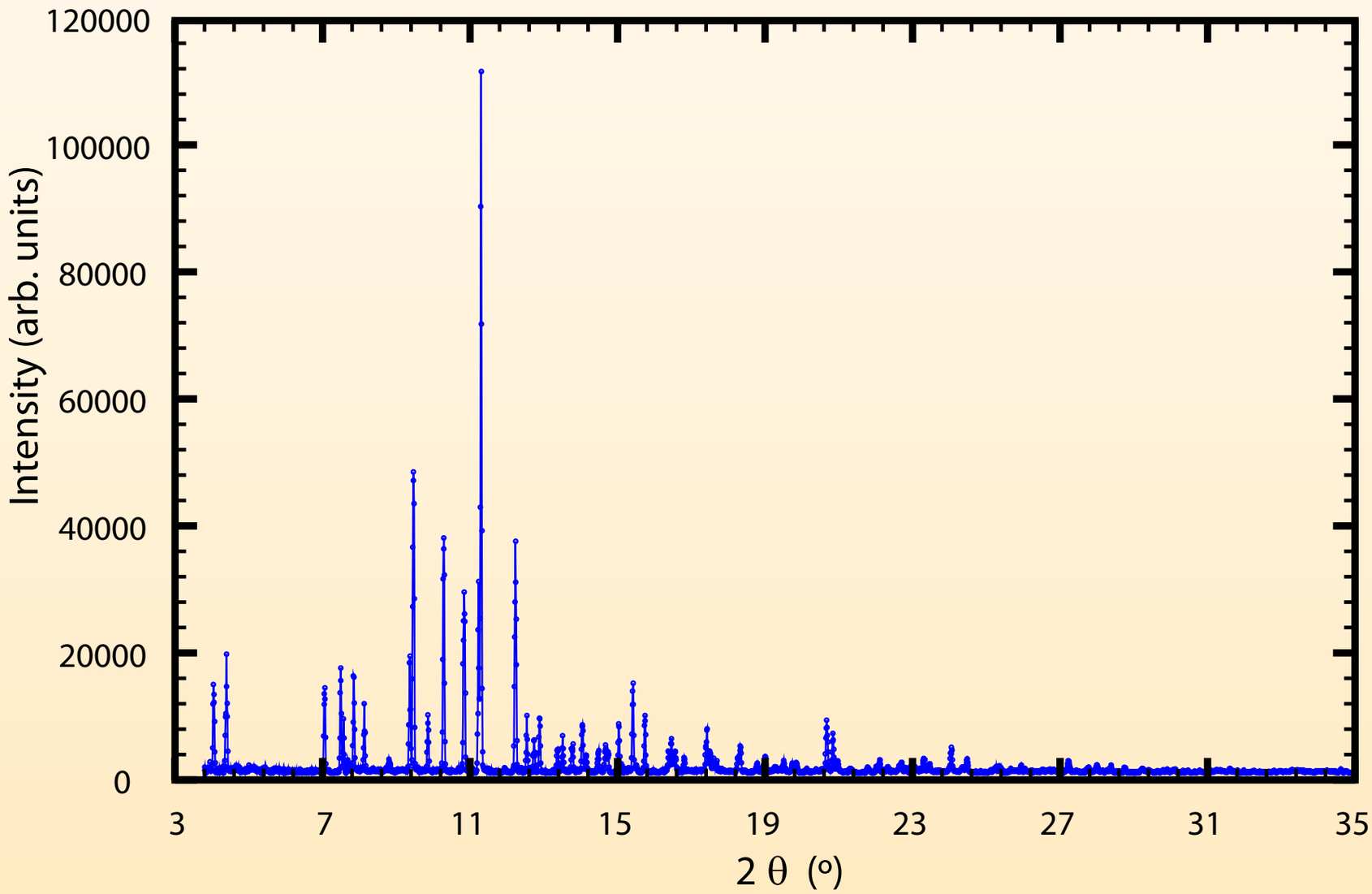
ICMAB
Institut de Ciència de Materials de Barcelona
www.icmab.es
Campus de la UAB
08193-Bellaterra (Spain)

ALBA
CELLS - ALBA
Synchrotron Light Facility

(*S*)-(–)-Ibuprofen

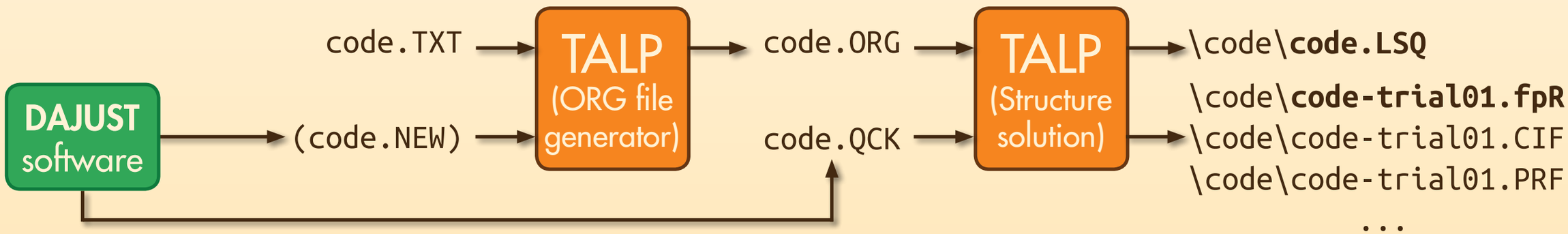
- Capillary Synchrotron data (MYTHEN II)
- $\lambda = 0.88368$ Å
- $C_{13}H_{18}O_2$





Steps

1. Indexing
2. Pattern matching
3. Molecular model
4. Restraints, FRB
5. TALP application
6. Check results
7. Rietveld Refinement



Indexing

- e.g. DICVOL

```
(...)

  THE SOLUTION IS NOW USED TO TRY TO INDEX ALL INPUT  38 LINES :
  ============================================================
DIRECT PARAMETERS : A=13.5406 B= 8.0320 C=12.4797 BETA=112.971 VOLUME= 1249.64
STANDARD DEVIATIONS:    .0011      .0009      .0016         .009
REFINED ZERO-POINT SHIFT : 0.0124 deg. 2-theta

REDUCED CELL : A=12.4797 B= 8.0320 C=13.5406 BETA=112.971 VOLUME=1249.64

   H   K   L    DOBS     DCAL    DOBS-DCAL 2TH.OBS  2TH.CAL  DIF.2TH.

   1   0   0  12.51406 12.50487   0.00919    4.047    4.050  -0.003
   0   0   1  11.52393 11.52238   0.00156    4.395    4.395  -0.001
   1   0   1   7.18061  7.18152  -0.00091    7.056    7.055   0.001
   1   1   0   6.76248  6.76311  -0.00062    7.492    7.492   0.001
(...)

        * NUMBER OF LINES
            .- LINES INPUT      =   38
            .- LINES INDEXED    =   38
            .- LINES CALCULATED =  167
        * AVERAGE 2-Theta DIFFERENCE FOR INDEXED LINES = 0.0000
        * MEAN ABSOLUTE DISCREPANCIES
                                        <Q> =0.2396E-04
                              <DELTA(2-THETA)> =0.1915E-02
            MAX. ERROR ACCEPTED (DEG. 2-THETA) =0.4500E-01
        * FIGURES OF MERIT
              1.- M( 20) =  167.4
              2.- F( 20) =  661.0(0.0008,   38)
```

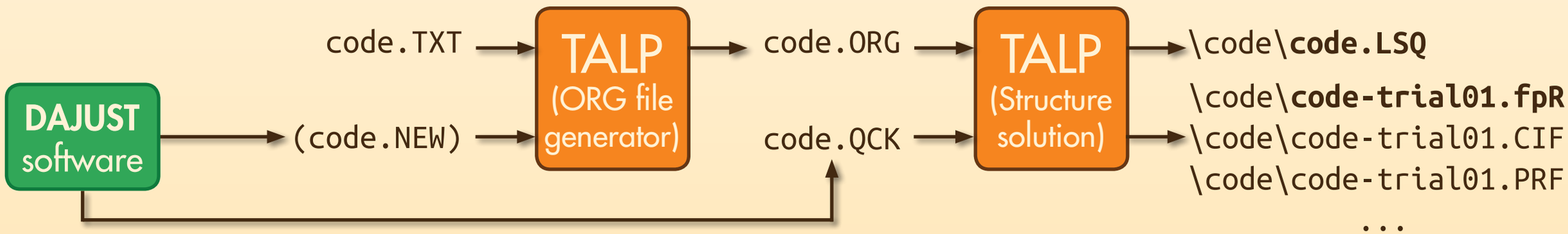
DAjust requires the approximate cell parameters:

a	12.48 Å	α	90°
b	8.03 Å	β	112.97°
c	13.54 Å	γ	90°

V 1250 Å^3

Monoclinic

Z=4 *Z′*=2



DAJUST pattern matching

a. PXRD data format: Pattern=1 (NYLINE=10)
   $2\theta_{sup}$=35°, Synchrotron (λ = 0.88368 Å)

b. Initial cell parameters:
   12.48Å 8.03Å 13.54Å 90.0° 112.97° 90.0°

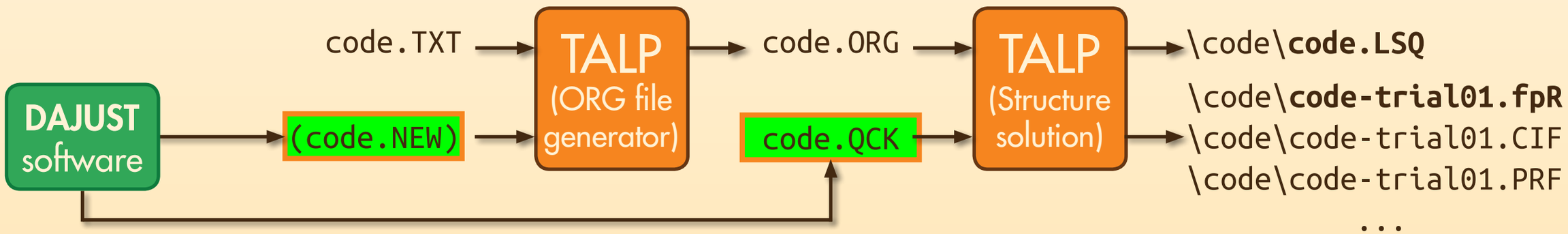
c. Pattern matching parameters (pV, zero, Lorentzian)

d. Run DAJUST (highest symmetry)

e. Run DAJUST & Run SGAID (sg. $P2_1$)

f. Run DAJUST

g. Copy QCK file to TALP folder



DAJUST pattern matching

ICMAB CSIC
Consejo Superior de Investigaciones Científicas

DAjust_UI_v1209

File Options Help

General

Name: S-Ibuprofen
Formula: C13H18O2
Z: 4
Cycles: 40

Cell parameters

a: 12.48264 α: 90.000
b: 8.03489 β: 112.9720
c: 13.54240 γ: 90.000
Metric: monoclinic
Fix cell param.
Use existing LAT file

Symmetry

SG.Num:Symbol: :P2(1) set Lattice: P
Centrosimetric
No centrosimetric
Highest symmetry
Symmetry matrices unlock
+X,+Y,+Z
-X,1/2+Y,-Z

Instrumental & Data parameters

Pattern type: 1 (Yobs) nYline: 10 2θsup: 35.0000
Rad. source: Synchrotron λ1: 0.88367 λ2: 0.00000
Intensity Extraction NFWHM: 25 tolerance: 0.5000
Sample: Flat Capillary
Primary monochromator: Yes No cos 2θ: 1.00000
Divergence correction: Yes No Fixed slit Variable slit
slit-w (°): 0.0000 sample size: 0.0000 armlenght: 0.0000
Absorption correction: Yes No
R*mu: 0.0000 2θRefInt: 0.0000 Transmi.: 0.0000

Profile Function

Type: Pseudo-Voigt T... zero: -0.009073 exp: 1.0000
Lorentzian coef. X: 0.368E-01 Y: 0.000E+00
Gaussian coef. W: 0.115E-02 V: 0.668E-02 U: 0.617E-02
Pref. Orientation: March-coef: 1.000 direction (h k l): 0 0 0

Excluded zones

2θinf: Add
2θsup: Del.

2Tinf	2Tsup

Background

type: Calcula... 2θ: + -
<back>: 60.00000 bg-value: File
FWHM: 0.15000
Iterations: 20

2T	bg-value

Asymmetry correction

coef.: Add
2θsup: Del.

coef	2Tsup
0.600	5.000

Ghost peaks

value: 2θ:
η: asym:

value	2T	width

Run AJUST | Stop run | open PRF | Output window
Run SGAid | Reload JST | SGAid window

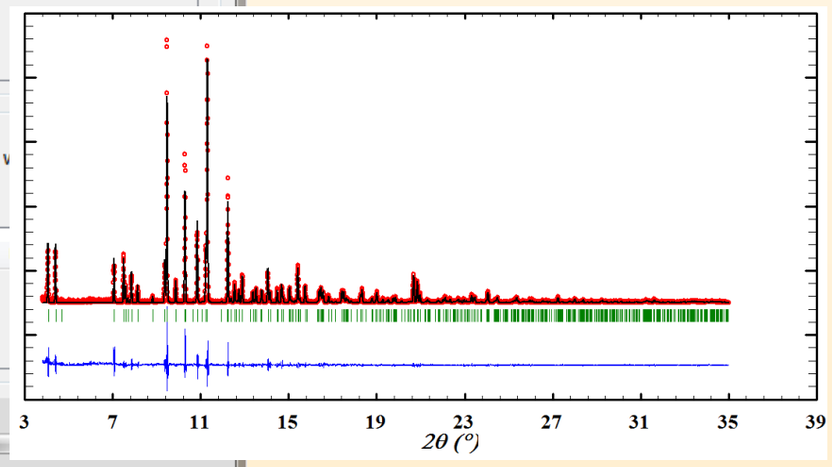
Work file: O:\ovallcorba\ibuprofen\dajust\ibuprofen@0.88368.jst
Data file: O:\ovallcorba\ibuprofen\dajust\ibuprofen@0.88368.dat

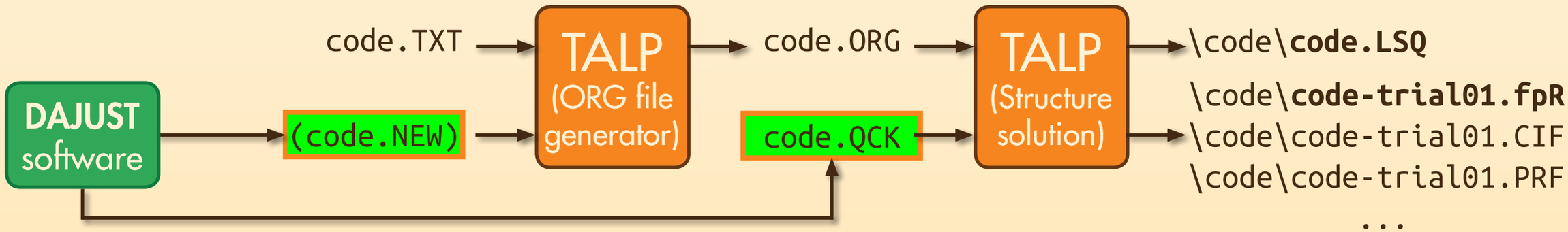
SGAid Output

Click on a row to load symmetry matrices for the S.G into main window
on top

SGNum	H-M symbol	RW (%)	NREF	Setting
6	Pm	**4,85**	409	abc
3	P2	**4,85**	409	abc
10	P2/m	**4,85**	409	abc
11	P2(1)/m	**4,85**	406	abc
4	P2(1)	**4,85**	406	abc
233	P2/n	8,04	358	abc
236	Pn	8,04	358	abc
234	P2(1)/n	8,04	355	abc
232	P2(1)/a	12,03	353	abc
231	P2/a	12,03	356	abc
235	Pa	12,03	356	abc
14	P2(1)/c	12,11	356	abc
13	P2/c	12,11	359	abc
7	Pc	12,11	359	abc

a=12.4657 b=8.0298 c=13.5385 α=90.0 β=112.928 γ=90.0





Molecular Model

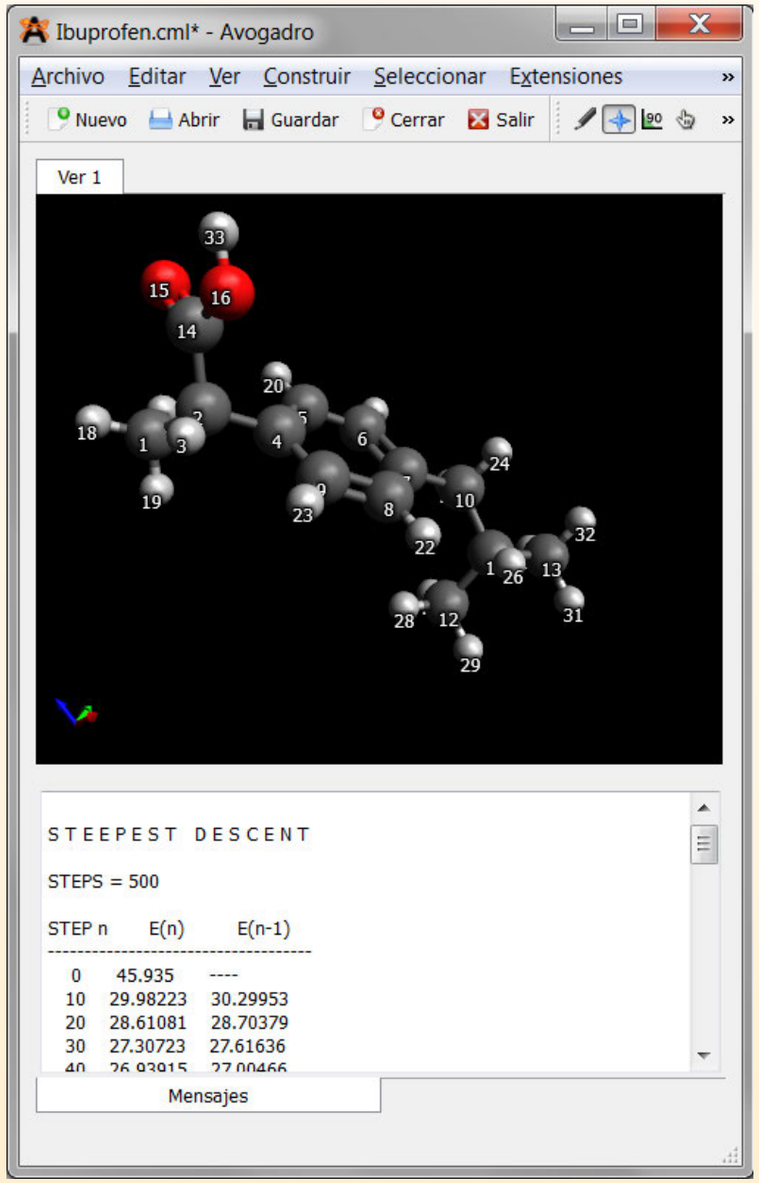
Molecular Model from calculations (MM)

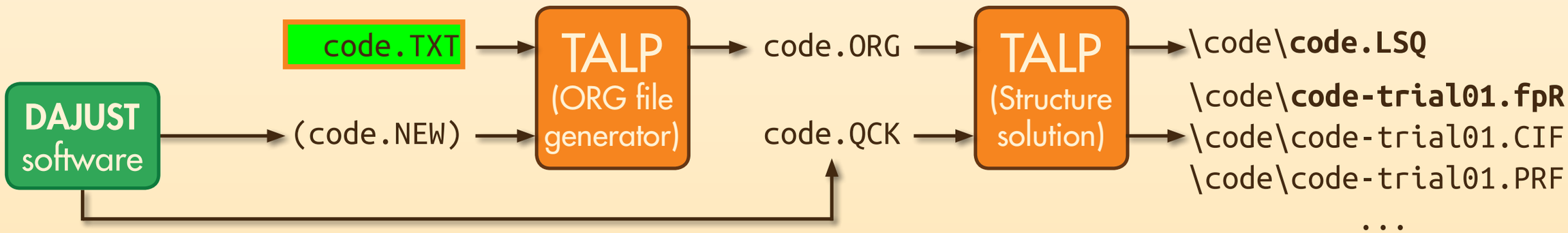
i.e. **Avogadro** (**MMFF94**, UFF), Chem3D (MM2),...

code.TXT

```
(S)-Ibuprofen
1 1 1 90 90 90
C6      C        3.4590    2.3496   -3.4027
C1      C        1.1118   -0.2872    0.3516
C2      C        2.6065   -0.0446    0.1267
C3      C        2.8853    0.7971   -1.1050
C4      C        2.3018    2.0589   -1.2814
...
```

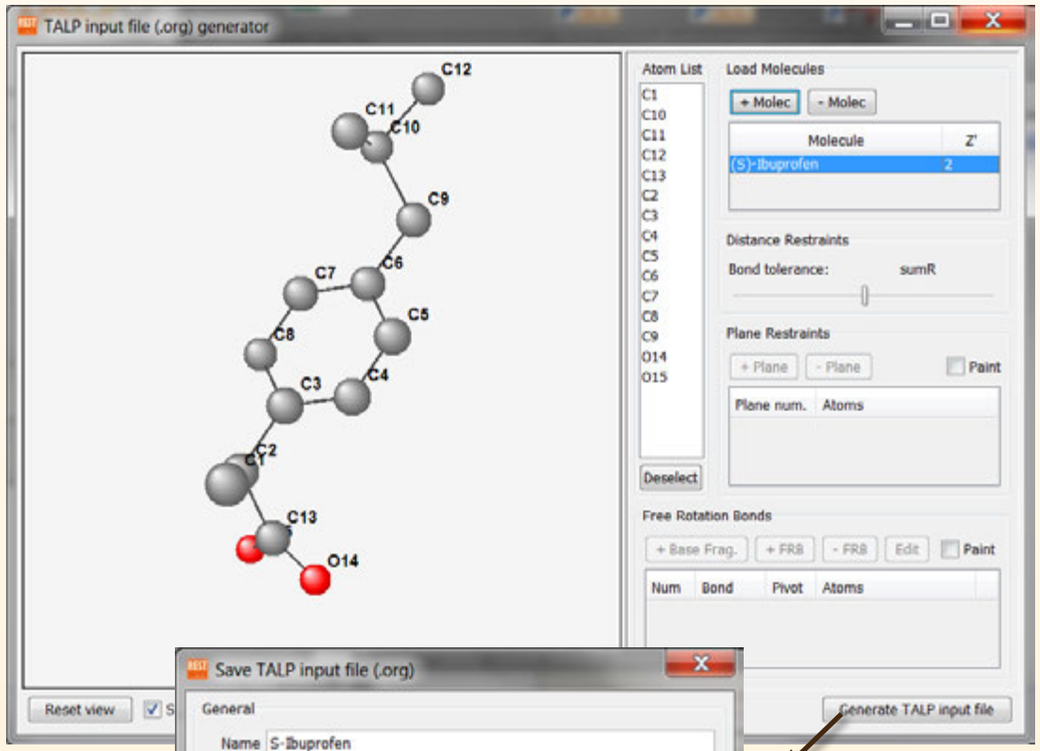
Recomendation: first atom near mass center

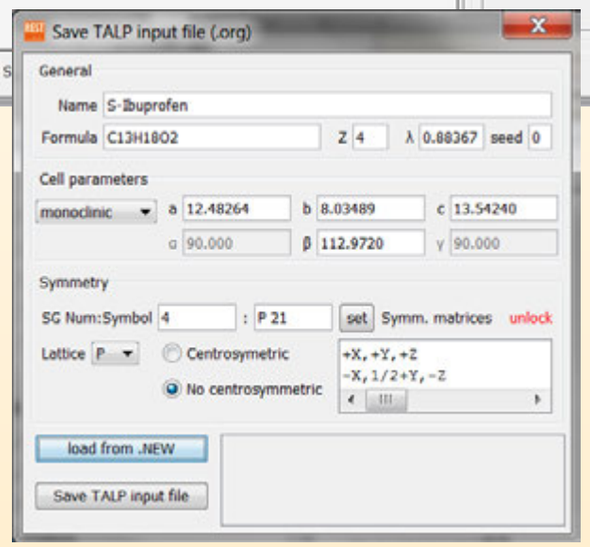




TALP input file (.ORG) generator

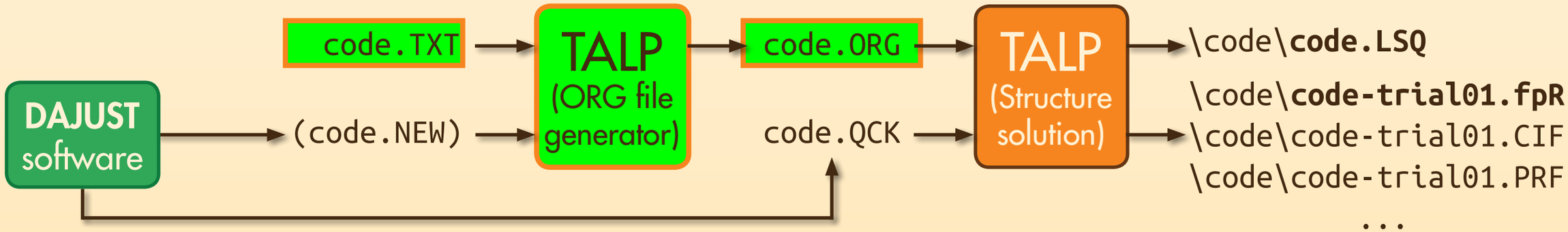
Indexing - Pattern matching - Molecular model - Restraints, FRB - TALP application - Check results - Rietveld Refinement



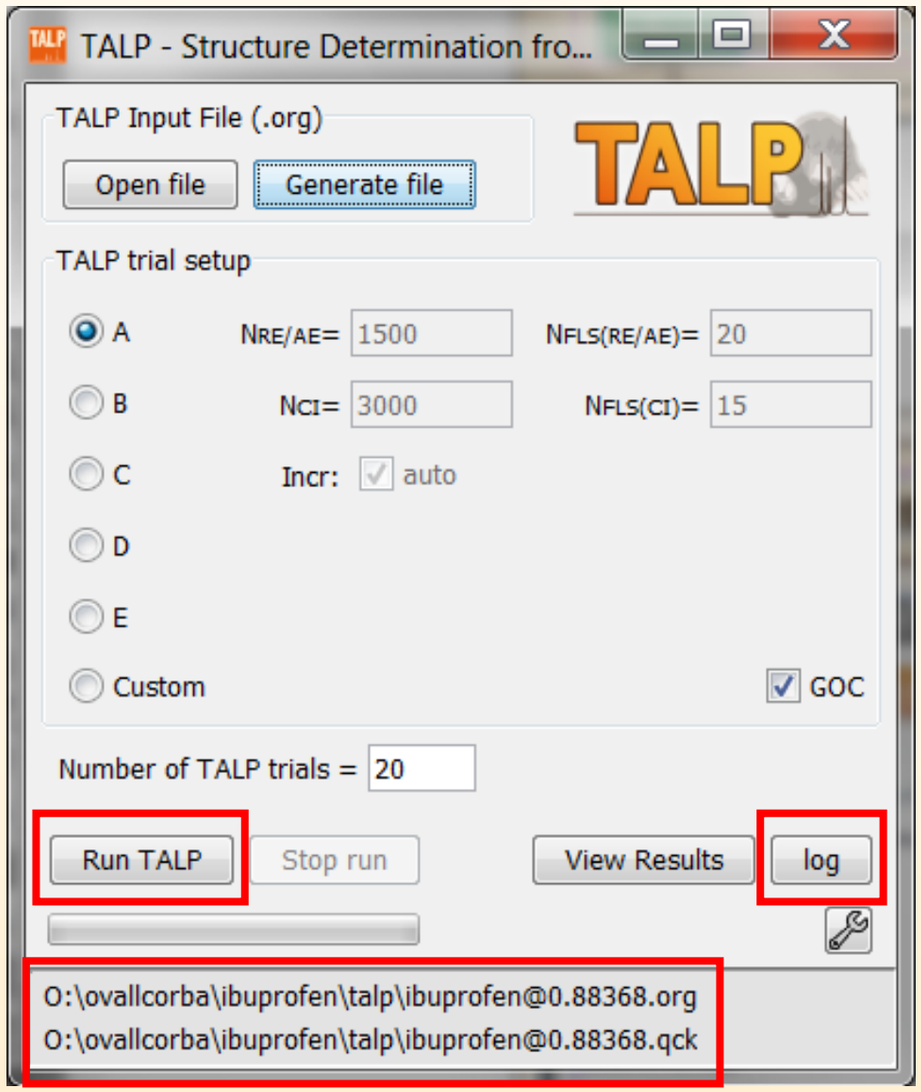


a. Import model coordinates (Z'=2)
b. Distance Restraints (bond tolerance)
c. Plane restraints
d. Free Rotation Bonds
e. Generate TALP input file

- Option NEW file
- Same code/location QCK (if not the program asks for and copy it to the save location)

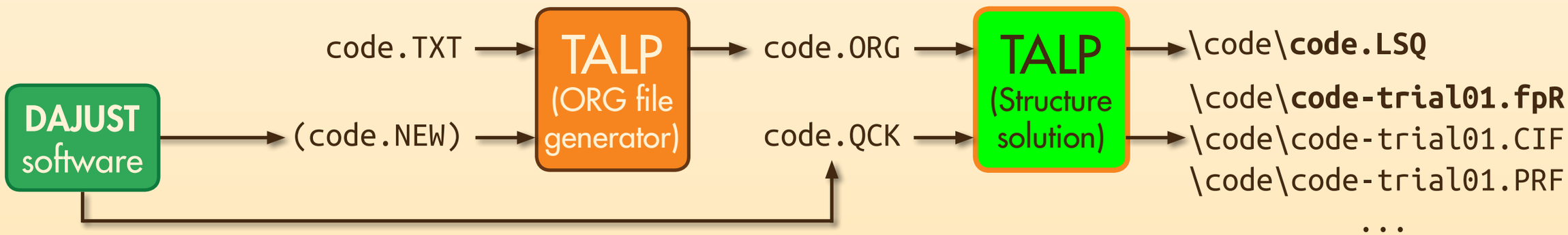


TALP experiment conditions and run



a. Preset/Custom conditions
b. N trials
c. GOC
d. Log window

	A	B	C	D	E
N_{rnd}	1500	3000	6000	10000	20000
$N_{FLS}(REM/AEM)$	20	20	30	30	30
N_{CI}	3000	6000	8000	15000	30000
$N_{FLS}(CI)$	15	15	15	15	15



TALP search strategy (trial flow-sheet)

Starting model

N_{rand} generated models (position, orientation, conformation), checked for overlap and refined by FLS

General scan

Random Exploration Mode (REM)

Assisted Exploration Mode (AEM)

N_{rand} generated models (position, conformation), checked for overlap and refined by FLS

$$\chi_{TALP} = \left(\frac{S_Y}{N_H + N_{REST} - P} \right)^{1/2}$$

Quality of intermediate models after FLS

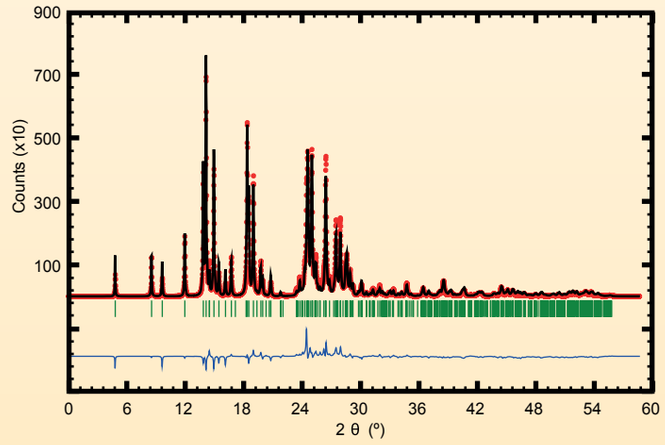
Local Incremental Scan (LIS)

N_{CI} increments of limited size to the best model (position, orientation, conformation) followed by FLS refinement

$$Q_{trial} = \frac{R_{wp,FPR}}{R_{wp,model-free}}$$

Indicator for candidate solutions (usually below 1.5 for correct solutions)

Fixed-Profile Rietveld refinement (FPR)



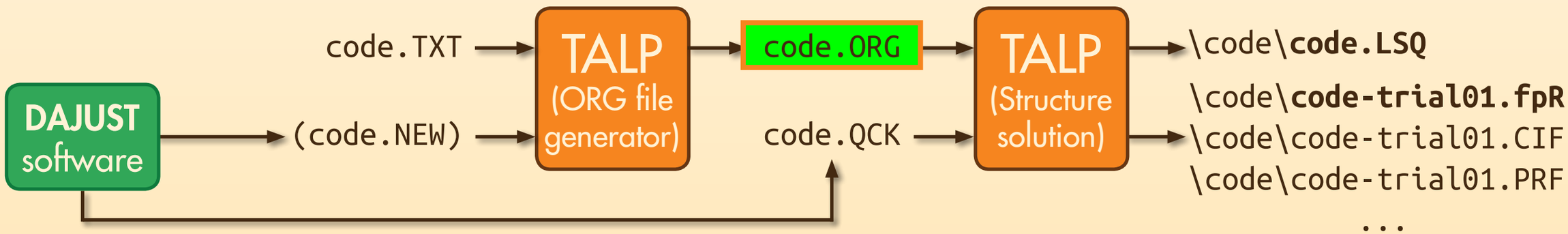
Final model (Q_{trial})

ORG file details

Additional control/options

Automatically generated by TALP

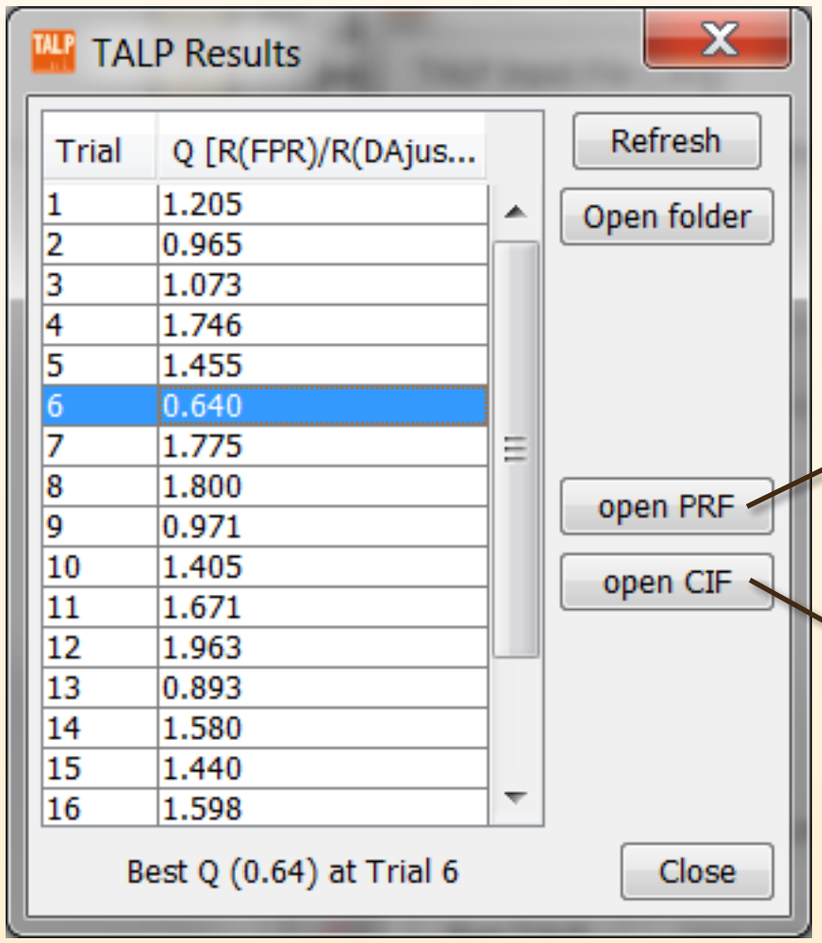
- Cell, symmetry, content,...
- Scale, Bover
- Atoms/Fragments (parameters, link, rigid bodies, FR,...)
- Molecular groups
- FRB: Free Rotation Bonds (limits REM, initial value,link,...)
- REST: Restraints (sdev, symmetry related,...)
- DIST: Distance calculation
- Options: RadType, HKL, FR, Seed
- 2TEX: Excluded zones

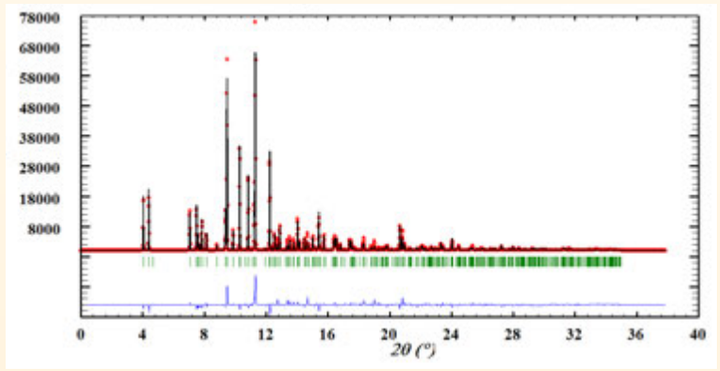


TALP output

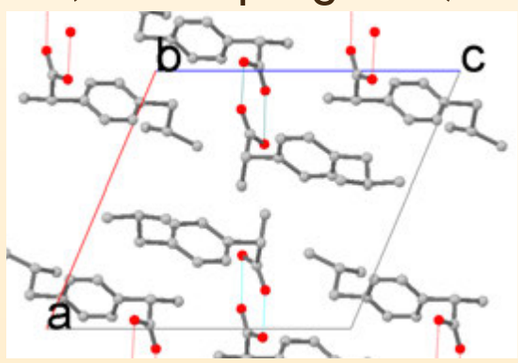
Results window and output files

- LSQ file: general
- fpR files: trial
- CIF/PRF visual inspection





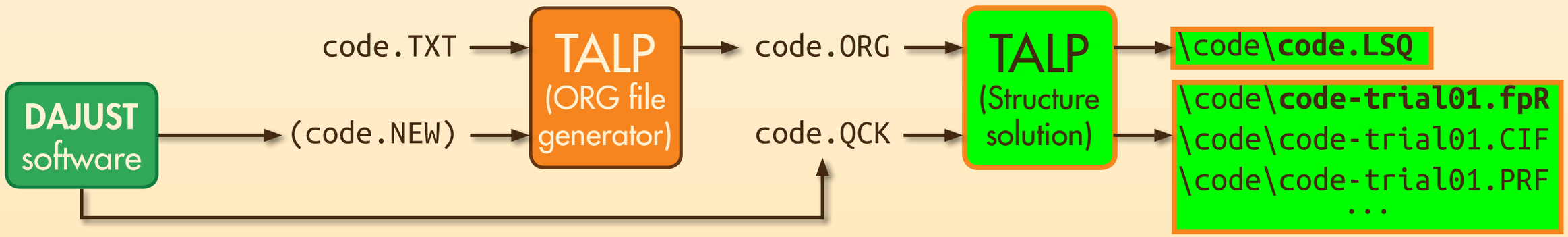
(default programs)



Remember!

$$Q_{trial} = \frac{R_{wp,FPR}}{R_{wp,model-free}}$$

Indicator for candidate solutions (usually near 1.0 for correct solutions)



Numeric result files (LSQ and fpR)

.LSQ

General results file (text):

- Copy of input file
- Data read by TALP
- TALP experiment info
- Trial Results (Qtrial)

```
(...)
TALP EXPERIMENT INFO:
 Program version= TALP1209
 Filename= ibuprofen@0.88368
 Experiment start date:time= [130327:1403]
 nRE= 1500 (nFLS= 20)
 nCI= 3000 (nFLS= 15)
 Automatic Increments
-------------------------------------------------
[130327:1403] TRIAL: 1 of 20
 Chi(Trial) =  0.664 (Experiment best=  0.664)
 Q(Trial)   =  1.205 (Experiment best=  1.205)
 time(trial)=    3.1 min
-------------------------------------------------
[130327:1406] TRIAL: 2 of 20
 Chi(Trial) =  0.630 (Experiment best=  0.630)
 Q(Trial)   =  0.965 (Experiment best=  0.965)
 time(trial)=    3.1 min
-------------------------------------------------
(...)
```

.fpR

Trial result files (text):

- Figures of Merit
- Refined parameters (coordinates,...)
- Dist/Angle Calculations

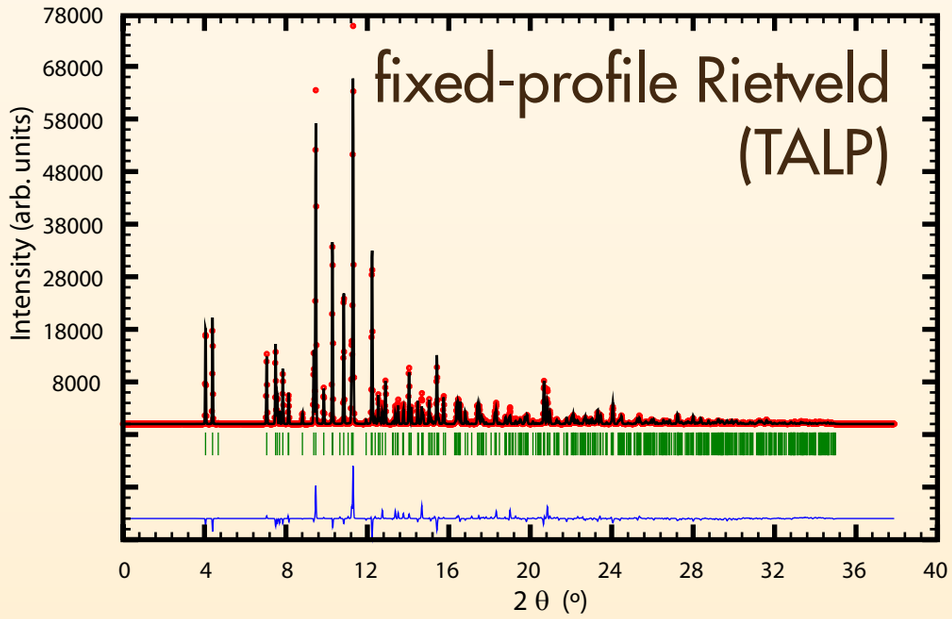
```
---------------------------------------------------------
FIXED-PROFILE RIETVELD REFINEMENT OUTPUT (trial solution):
---------------------------------------------------------
S-Ibuprofen, P 21, Z=4, C13 H18 O2,
---------------------------------------------------------
  Rwp         Q(Trial)
  0.1218E+00  0.6395E+00
  Rexp        Rwm(planes) Rwd(dist)
  0.2670E-01  0.1802E+00  0.6834E+00
  Scale       Bover
  0.6144E+00  0.0000E+00
---------------------------------------------------------
    1   1.1  0.1  1.1         0    0    0         0    0
C4a      1   0.11641  -0.08688   0.80549   1.00000   3.422
    1   1.1  1.1  1.1         0    0    0         0    0
C5a      1   0.18985  -0.14843   0.89969   1.00000   3.422
...
---------------------------------------------------------
 DISTANCES AND ANGLES:

    1  C4a      0.11641  -0.08688   0.80549
C5a     1.343  0.001   0.18985  -0.14843   0.89969  1  0.000  0.000  0.000
C6a     2.336  0.001   0.19110  -0.05562   0.99148  1  0.000  0.000  0.000
(...)
```

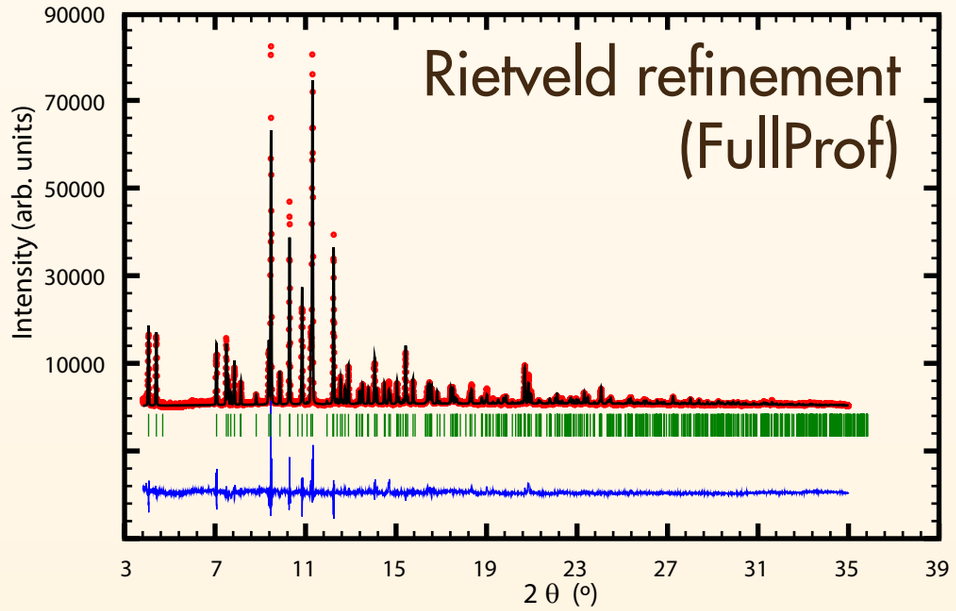
Structure refinement

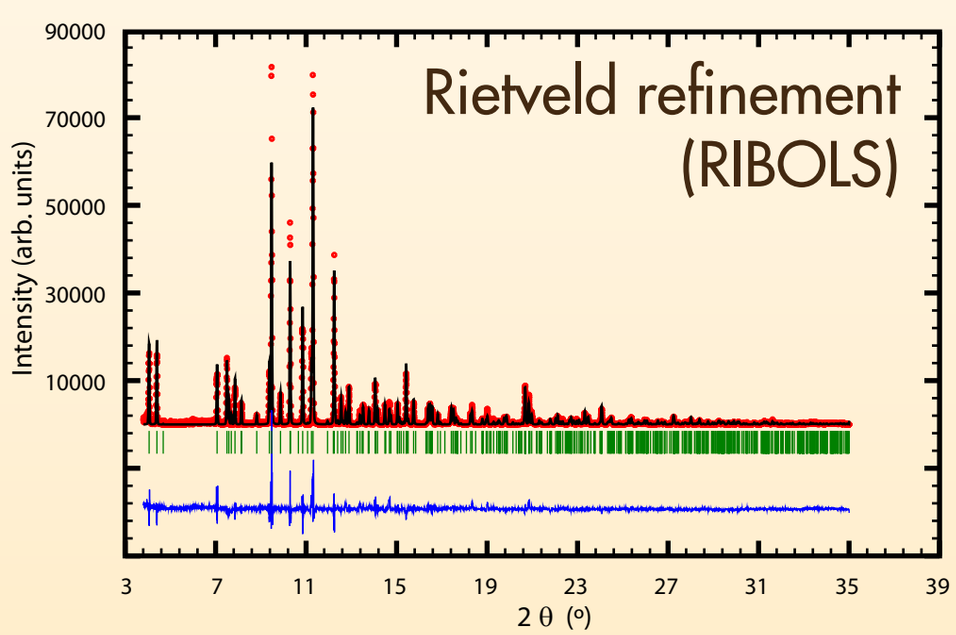
TALP = Structure solution

↓

A final Rietveld refinement is needed







Cell parameters:

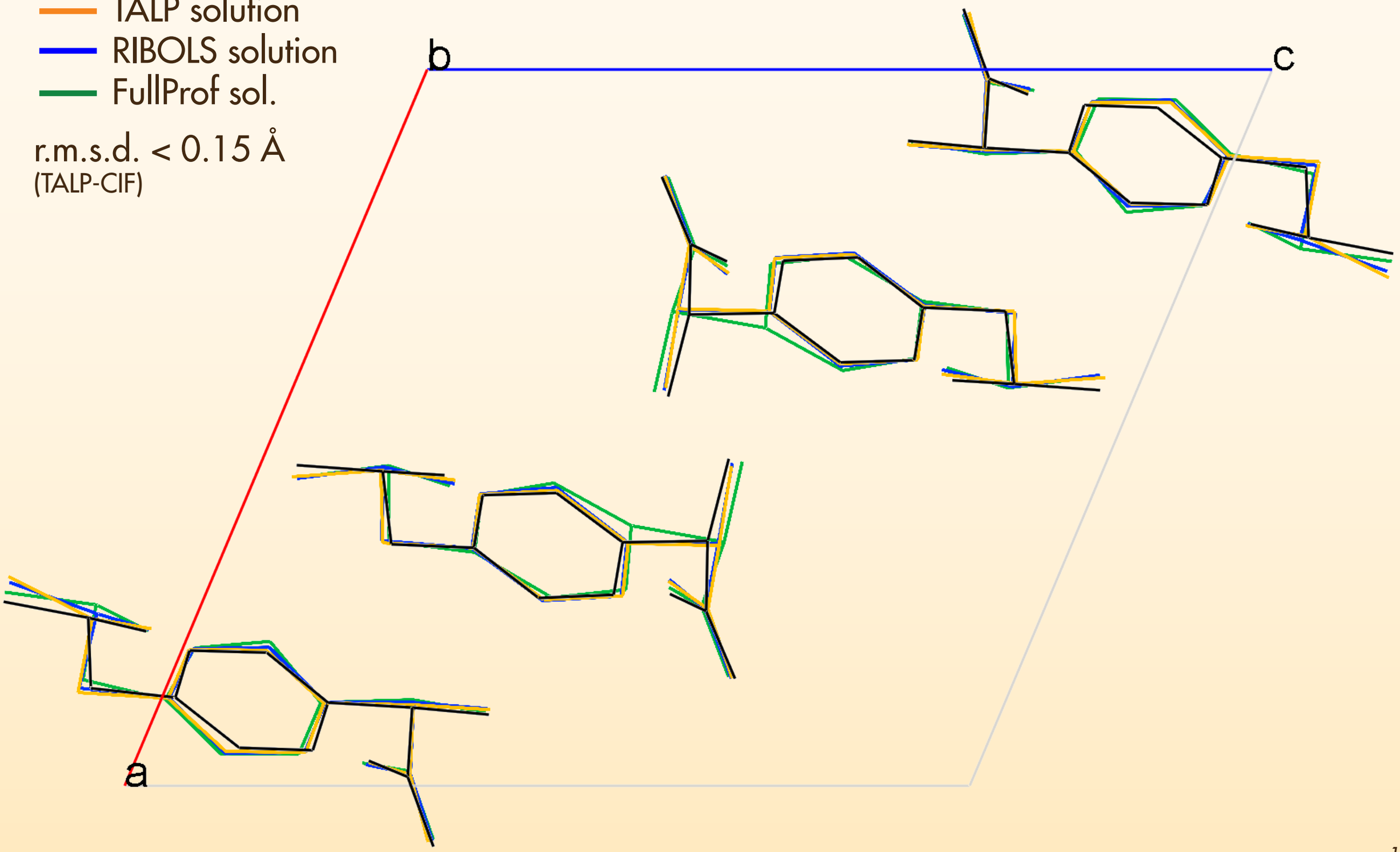
Published:	12.462(3)	8.035(3)	13.539(4)	90.0	112.89(3)	90.0
DAjust:	12.4826(3)	8.0349(2)	13.5424(4)	90.0	112.972(2)	90.0
Ribols:	12.4829(3)	8.0350(2)	13.5425(4)	90.0	112.971(2)	90.0
fullProf:	12.4827(4)	8.0345(2)	13.5428(4)	90.0	112.971(1)	90.0

(*S*)-(–)-Ibuprofen

ICMAB CSIC
Consejo Superior de Investigaciones Científicas

published CIF
TALP solution
RIBOLS solution
FullProf sol.

r.m.s.d. < 0.15 Å
(TALP-CIF)



TALP

Oriol Vallcorba
Jordi Rius
Carlos Frontera
Carles Miravitlles

J. Appl. Cryst. 2012, 48. 1270-1277

ovallcorba@icmab.com

www.icmab.es/crystallography/software

THANK YOU FOR YOUR ATTENTION!!

Thanks to the ICMAB, the Crystallography Dpt. and specially to Prof. Jordi Rius, Prof. Carles Miravitlles, Dr. Carles Frontera & Dr. Inma Peral for their help and support.

PowderXRD 2013

CSIC
Consejo Superior de Investigaciones Científicas

ICMAB
Institut de Ciència de Materials de Barcelona
www.icmab.es
Campus de la UAB
08193-Bellaterra (Spain)

CELLS - ALBA
Synchrotron Light Facility
ALBA