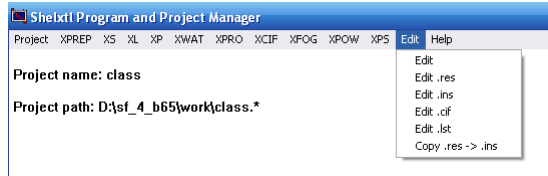


## Structure Determination using Program Shelxtl

Use routine **XPREP** in program Saint, you can prepare the input files, \*.ins and \*.HKL for structure determination

You can view or edit these files in Shelxtl by clicking menu **Edit**



**Step 1.** Get initial structural either use Patterson or Direct Method, by entering Patt or Tref in input file. Click menu XS to run structure determination. The output files from the structure determination is \*.Lst and \*.res files. \*.res file can be used as input file for structure refinement and electron density calculation \

- **Example input file for structure solution**

```
TITL sf_4_b6m in C2/c           name of your crystal
CELL 0.71073 12.6651 19.3272 11.1318 90.000 112.922 90.000  wavelength and cell
ZERR 4.00 0.0013 0.0019 0.0011 0.000 0.003 0.000  # of molecule in unit cell and cell parameters
LATT 7                               code of Lattice (negative means non centro-space group)
SYMM -X, Y, 0.5-Z                   symmetry operations
SFACC H O F U                       elements in the crystal
UNIT 80 80 24 40 4                  # of elements in an unit cell
Patt (or Tref)                       run Patterson or Direct Method
HKLF 4                               format of data
END
```

- **Result of Patterson Vectors**

•	X	Y	Z	Weight	Peak	Sigma	Length	
•	1	0.0000	0.0000	0.0000	4.	999.	110.19	0.00
•	2	0.0000	0.0000	0.5000	4.	499.	54.99	5.57
•	3	0.0099	0.0000	0.9041	2.	59.	6.49	1.12
•	4	0.0915	0.0000	0.6753	2.	59.	6.46	4.20
•	5	0.0952	0.0000	0.1776	2.	52.	5.72	1.87
•	6	0.0041	0.0489	0.0539	1.	46.	5.09	1.11
•	7	0.0478	0.0155	0.9416	1.	43.	4.70	1.09
•	8	0.1511	0.0345	0.4244	1.	42.	4.58	4.40
•	9	0.0000	0.0593	0.0000	2.	40.	4.36	1.15
•	10	0.0696	0.0424	0.0446	1.	39.	4.35	1.16
•	11	0.0588	0.0151	0.0981	1.	38.	4.19	1.10
•	12	0.0438	0.1011	0.0260	1.	35.	3.87	2.02
•	13	0.5000	0.9776	0.5000	2.	34.	3.70	6.62
•	14	0.1525	0.1224	0.3448	1.	33.	3.62	4.28
•								

- **Atomic Positions from Patterson Search**

- TITL sf\_4\_b6m in C2/c
- CELL 0.71073 12.6651 19.3272 11.1318 90.000 112.922 90.000
- ZERR 4.00 0.0013 0.0019 0.0011 0.000 0.003 0.000
- LATT 7
- SYMM -X, Y, 0.5-Z
- SFAC C H O F U
- UNIT 80 80 24 40 4
- L.S. 4
- BOND
- FMAP 2
- PLAN 20
- U1 5 0.50000 0.50000 0.50000 10.50000 0.04
- U2 5 0.59470 0.50000 0.67677 11.00000 0.04
- F3 4 0.43907 0.60134 0.46791 11.00000 0.04
- F4 4 0.56100 0.60134 0.53209 11.00000 0.04
- F5 4 0.64392 0.62373 0.33405 11.00000 0.04
- F6 4 0.35615 0.62373 0.66595 11.00000 0.04
- F7 4 0.12131 0.63799 0.34994 11.00000 0.04
- F8 4 0.87876 0.63799 0.65006 11.00000 0.04
- HKLF 4

- **Example input file for Direct Method**

```
TITL sf_4_b6m in C2/c
CELL 0.71073 12.6651 19.3272 11.1318 90.000 112.922 90.000
ZERR 4.00 0.0013 0.0019 0.0011 0.000 0.003
LATT 7 *code of lattice type. A negative value means non-centrosymmetric
SYMM -X, Y, 0.5-Z
SFACC H O F U
UNIT 80 80 24 40 4
TREF np[100] nE[#] kapschal[#] ntan[#] wn[#] *input parameters for direct method
HKL 4
END
```

**np**: the number of direct methods attempts

**nE**: reflections are employed in the full tangent formula phase refinement

**Kapschal**: multiplies the products of the three E-values used in triplet phase relations;

**ntan**: the number of cycles of full tangent formula refinement, which follows the phase annealing stage and involves all nE reflections;

**wn**: a parameter used in calculating the combined figure of merit CFOM:

CFOM = Ra (NQUAL < wn) or

CFOM = Ra + (wn - NQUAL)<sup>2</sup> (NQUAL >= wn);

**If you don't give these values, the program will set default values. When you fail to get the correct solution, you should consider increasing np, nE and ntan. For a difficult structure, try following command: TREF 1000 200 0.85 300 -0.25**

- **Selected output of direct methods**

```
.
.
INCONSISTENT EQUIVALENTS
.
h k l F*F Sigma(F*F) Esd of mean(F*F)
.
2 0 0 1376.33 8.82 309.02
.
6 0 0 442.58 6.49 43.86
.
-5 1 0 426.80 6.09 38.98
.
-3 1 0 1009.79 9.22 207.91
.
-1 1 0 2133.92 9.38 244.06
.
1 1 0 2042.89 11.88 217.46
.
3 1 0 926.55 7.97 127.81
.
5 1 0 359.27 6.72 39.87
.
-6 2 0 461.33 7.07 38.32
.
-4 2 0 732.81 7.22 66.32
.
-2 2 0 435.40 4.96 37.07
.
.....
** Etc. **
.
4910 Unique reflections, of which 2837 observed
.
R(int) = 0.0916 R(sigma) = 0.0991 Friedel opposites merged
.
R(int) = Σ | Fo2 - Fo2(mean) | / Σ [ Fo2 ]
.
R(sigma) = Σ [ σ(Fo2) ] / Σ [ Fo2 ]
```

- If you cannot solve the structure, you need to check these inconsistent diffractions. If necessary, you can delete them from HKL list (.HKL file). Programs, like **SORTAV** implemented in **WinGX**, are available to remove outliers (<http://www.chem.gla.ac.uk/~louis/software/wingx>)

```

.....
Observed E .GT. 1.200 1.300 1.400 1.500 1.600 1.700 1.800 1.900 2.000 2.100
Number      853 564 323 158 74 35 22 12 12 7
              Centric Acentric HKL Okl h0l hk0 Rest
Mean Abs(E*E-1) 0.968 0.736 0.995 0.646 0.393 0.278 0.607
.....

```

ONE-PHASE SEMINVARIANTS

```

h k l E P+ Phi
2 0 2 1.651 0.14
-4 0 2 1.492 0.56
-4 8 2 1.738 0.44
6 2 2 1.861 1.00
.....

```

Expected value of Sigma-1 = 0.815

Following phases held constant with unit weights for the initial 4 weighted tangent cycles (before phase annealing):

```

h k l E Phase/Comment
3 1 1 4.635 random phase
2 0 2 1.651 180 sigma-1 = 0.144
6 2 2 1.861 0 sigma-1 = 1.000
1 9 0 1.703 random phase
0 8 0 1.460 0 sigma-1 = 1.000
-5 1 5 1.805 random phase
1 3 1 1.775 random phase
.....

```

“Although the original versions of *MULTAN* assigned up to four fixed phases to define the origin and enantiomorph, and then generated starting sets by permuting numerical values for several further phases, it appears that assigning random starting values to all phases is at least as effective.”

**“Phase Annealing in *SHELX-90*: Direct Methods for Larger Structures”**

*George Sheldrick Acta Cryst. (1990). A46, 467-473*



## Finally

- Phase set with the smallest CFOM is used to calculate an E-Fourier synthesis and E map
- Search the E map to find potential atoms.
- Use an iterative procedure (Sheldrick, 1982), in which potential atoms are eliminated to reduce an R index (calculated for E values assuming point atoms)
- Use remaining atoms to phase the next E map, obtain a further figure of merit (RE), and facilitate a solution that has a good chemical interpretation

```
• Solution from Direct Method
• .....
• E-Fourier for sf_4_b6m in C2/c
• Maximum = 891.51, minimum = -155.07   highest memory used = 8783 / 14135   0.0 seconds elapsed time
• Heavy-atom assignments:
•   x   y   z   s.o.f. Height
• U1  0.0000 0.5000 0.0000 0.5000 891.5
• Peak list optimization
• RE = 0.351 for 18 surviving atoms and 853 E-values   Highest memory used = 1598 / 7677
• .....
• Atom Peak  x   y   z   SOF Height Distances and Angles
• .....
• U1  0.  0.0000 0.5000 0.0000 0.500 3.14 0 1  2.389
•           03  3.876 125.9
•           06  2.114 88.9 43.7
•           08  3.408 92.0 37.4 8.5
•           13  2.430 147.8 23.4 59.8 55.9
•           14  3.614 15.2 114.8 84.2 85.3 138.1
•           16  2.045 62.3 63.6 36.1 34.2 86.1 52.5
•           18  3.965 24.7 105.4 64.9 69.0 124.6 27.5 44.7
•           20  3.678 40.3 94.7 76.9 74.6 117.6 25.1 40.8 43.9
•           21  2.705 67.4 71.4 27.9 35.5 84.5 68.2 39.8 42.8 72.5
•           23  2.934 82.3 100.5 72.4 80.7 93.9 94.6 93.4 70.6 114.4 54.0
•           24  3.513 139.0 27.7 50.6 49.0
• .....
• .....
```

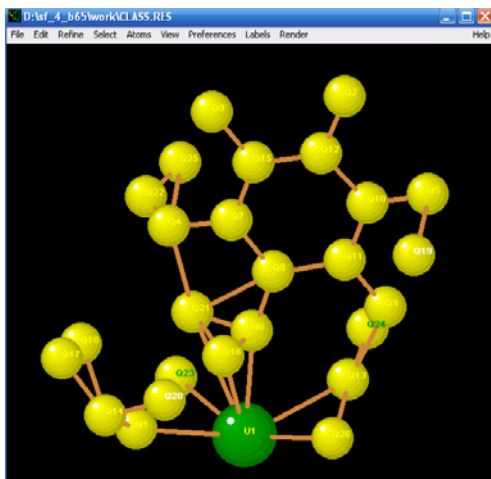
- **\*.res file from Direct method**
- TITL sf\_4\_b6m in C2/c
- CELL 0.71073 12.6651 19.3272 11.1318 90.000 112.922 90.000
- ZERR 4.00 0.0013 0.0019 0.0011 0.000 0.003 0.000
- LATT 7
- SYMM -X, Y, 0.5-Z
- SFAC C H O F U
- UNIT 80 80 24 40 4
- L.S. 4
- BOND
- FMAP 2
- PLAN 20
- MOLE 1
- U1 5 0.0000 0.5000 0.0000 10.500000 0.05
- Q1 1 -0.1465 0.4670 0.0736 11.000000 0.05 128.80
- Q2 1 0.1282 0.1384 -0.1388 11.000000 0.05 116.45
- Q3 1 0.1505 0.3789 -0.1616 11.000000 0.05 116.18
- Q4 1 0.0203 0.2769 0.1367 11.000000 0.05 115.50
- Q5 1 0.1777 0.2522 -0.2434 11.000000 0.05 105.44
- Q6 1 0.0576 0.3963 0.0238 11.000000 0.05 104.38
- Q7 1 0.0534 0.2718 0.0399 11.000000 0.05 99.06
- Q8 1 0.0787 0.3322 -0.0107 11.000000 0.05 97.39
- Q9 1 0.0505 0.1503 0.0579 11.000000 0.05 94.93

**Step 2. Use Graphic program Xshell, you can display the structure from Patterson or Direct method solution**

In this example, both Patterson and Direct method suggest U is located at special position (0,0,0), and Direct method gives most atomic positions of the molecule.

Frequently, the program cannot automatically assign element to a found position (labeled as Q) or assigns them incorrectly, so you need apply your synthesis knowledge and the sense of chemical bonding to assign the elements correctly. The partial structure from Patterson or Direct method can be very “noisy”, you need remove these peaks, typically having low height or very close to the heavy atom.

## $C_{20}H_{20}F_{10}O_6U$ structure from Direct Method.



Because U atom is at inversion center, only half of the molecule is independent and to be listed.

### Step 3 Structure Refinement and Fourier Electron Density Map

Once the initial structure is derived from Patterson or Direct method, the complete structure can be obtained by repeating the process of **refinement + difference Fourier calculation**.

By default, the program runs Refinement-DF calculation (based on the optimized Structure factors) automatically. But you can change the input file to fit your needs.

```
TITL sf_4_b6m in C2/c
CELL 0.71073 12.6651 19.3272 11.1318 90.000 112.922 90.000
ZERR 4.00 0.0013 0.0019 0.0011 0.000 0.003 0.000
LATT 7
SYMM -X, Y, 0.5-Z
SFAC C H O F U
UNIT 80 80 24 40 4
L.S. 4          # of cycle of refinement. If # is 0, only do the Fc calculation and Fourier Calculation
BOND
ACTA          generate CIF file
FMAP 2        Define the Type of Fourier map. 2: Fo-Fc; 3: Fo
PLAN 20      # of peaks from Fourier map to be listed
MOLE 1
U1 5 0.0000 0.5000 0.0000 10.500000 0.05      atom position(x,y,z), occupancy, temperature *
MOLE 2
HKLF 4
END
```

\* Because U is at inversion center, its occupancy is 0.5, 10.5 means the occupancy is fixed during the refinement

