

A Guide to Using SHELXTL

Gregory S. Girolami, Julia L. Brumaghim, James G. Priepot, and Jon P. Goveia
24 January 2000
SHELXTL NT Version

1.	SHELXTL programs and files	2
2.	General Procedures	4
3.	Preparing for X-ray work-up.	4
4.	Setting up for the refinement – XPREP	5
5.	Solving by direct methods – XS.	8
6.	Assessing the solution – XP	10
7.	Solving by Patterson methods – XS	13
8.	What to do when you can't find a solution	15
9.	Editing the .ins file generated by XP.	15
10.	Least squares refinement – XL	17
11.	Checking for missed crystallographic symmetry – PLATON	21
12.	Turning atoms anisotropic and adding hydrogen atoms	22
13.	Correcting for absorption – XPREP.	24
14.	Extinction, weighting schemes, and absolute configuration.	28
15.	Final details	29
16.	Generating graphical displays – XP	32
17.	Printing diagrams generated by XP	33
18.	Miscellaneous XP commands	34
	Appendix 1. Sample structure report	35

Copyright © 2000 by G. S. Girolami.

Reproduction or translation of any part of this work beyond that permitted by Section 107 or 108 of the 1976 United States Copyright Act without the permission of the copyright owner is not permitted.

1. SHELXTL programs and files

1.1) SHELXTL programs

SHELXTL is a software package that is useful for solving and refining single-crystal X-ray diffraction data sets. SHELXTL consists of four major programs, which are:

- XPREP – Space group determination, absorption corrections, unit cell transformations, and reciprocal space plots. XPREP reads the raw data file `filename.raw` and the parameter file `filename.p4p` written by the diffractometer control program, and writes the instruction file `filename.ins` and reflection data file `filename.hkl` for use by the programs XS and XL.
- XS – Structure solution by direct methods or Patterson methods. XS reads `filename.hkl` and `filename.ins` (generated by XPREP) and writes the best solution (a list of atomic coordinates plus other information) to the results file `filename.res`. A detailed listing of the program's activities are written to the file `filename.lst`.
- XL – Least-squares structure refinement. XL reads `filename.ins` (obtained by renaming the `filename.res` file generated by XS or by a previous XL refinement cycle) and writes a new results file `filename.res` and listing file `filename.lst`.
- XP – Interactive molecular graphics and publication-quality diagrams. XP reads the `filename.res` from XS or XL and (if so desired) can (re)write the `filename.ins` file for the next series of refinements.

1.2) SHELXTL files

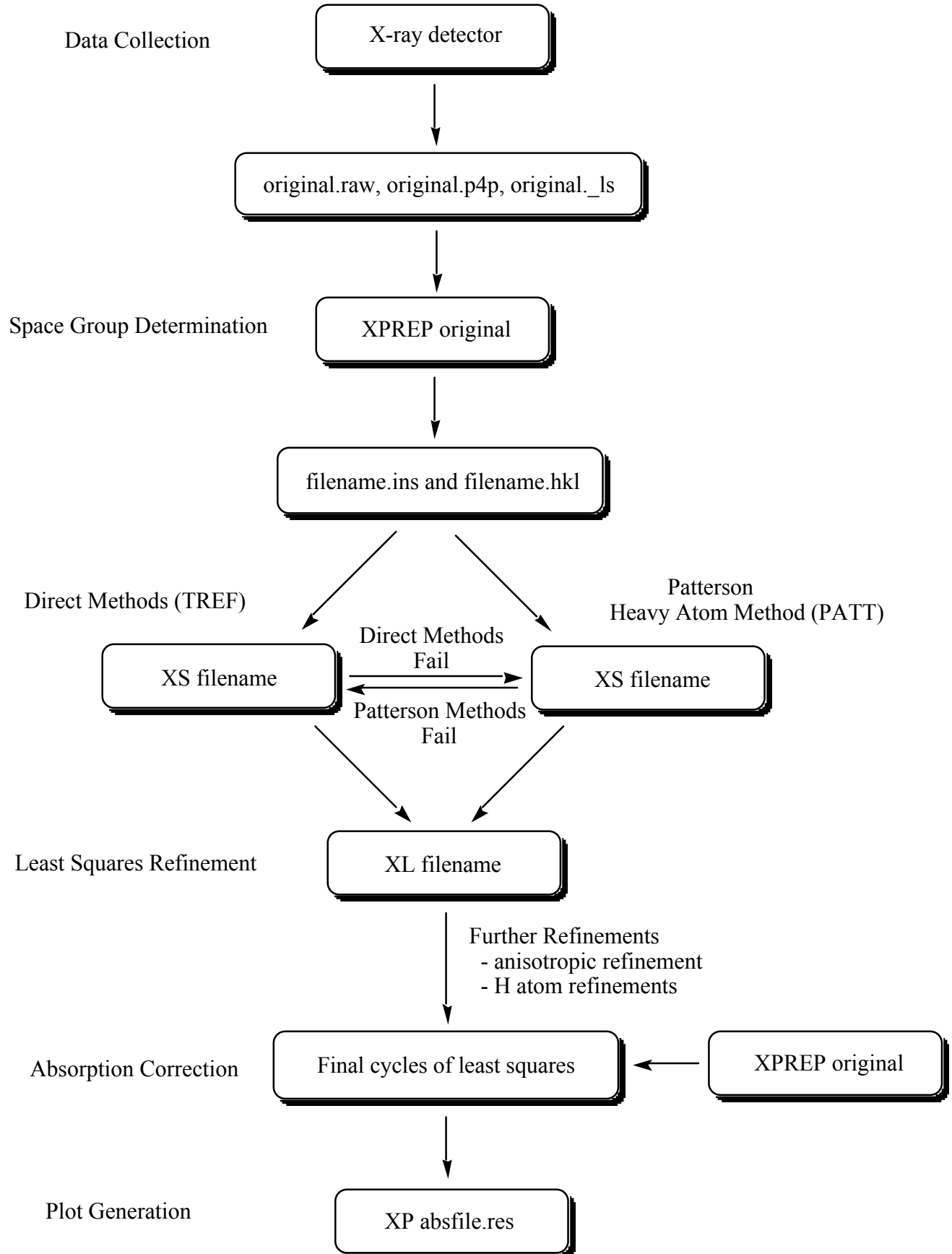
SHELXTL uses files of the type '`filename.ext`', where *filename* is a name of up to eight alphanumeric characters, and *ext* is a three character extension. Unless you change them, files for a given problem will always have the same *filename* but different extensions. The extensions define the function of each file and are generated or recognized automatically by the SHELXTL programs. The file extensions used by SHELXTL are:

- `.raw` – raw reflection data; used as input file for XPREP
- `.ls` – statistical analysis of raw reflection data
- `.prp` – a listing of what was done while running XPREP
- `.p4p` – data collection parameter file; used as input file for XPREP
- `.hkl` – reflection intensity table created by XPREP.
- `.ins` – instruction file containing cell parameters, etc. created by XPREP; input file for XS, XL, and XP
- `.res` – results file created by running XS or XL; updated form of `.ins` file
- `.lst` – a listing of what was done in the last computer run.

1.3) Solving a crystal structure

The flowchart on the following page briefly outlines the steps necessary to solve an X-ray crystal structure by the methods discussed in this handout. It includes everything from X-ray data collection to printing out a plot of the newly characterized molecule.

This handout does not, however, discuss how to handle disorder, non-integral site occupancy factors, or restraints. A separate handout is available that deals with these issues. Consult an expert for help if you encounter these problems.



2. General Procedures

- 2.1) Almost all commands that you type in must be followed by hitting the <enter> key. In this handout, this is implicit; you will not be reminded that hitting the <enter> key is necessary.
- 2.2) **It is useful to make regular backup copies of key files (especially .ins files).** To do this, go to the Edit pull-down menu, and open the file of interest. Then go to the File pull down menu, click on Save As, and type in a new name for the backed-up file (e.g., bkup1.ins). Note: the filename must be eight characters or shorter and must consist only of letters and numbers.

NOTE: The SHELXTL software is unforgiving of mistakes in file handling – it is very easy to corrupt the .ins file and undo all the work you have done. If you make backups regularly, then you will be able to recover easily from data handling mishaps: simply open the most recent backed-up file, use the Save As command to rename it filename.ins, and continue with your refinement. If you do not make regular backups, then you risk having to redo steps you have already done.

- 2.3) Some editing of text files will be necessary. The editor used in this laboratory is Wordpad; below is a summary of some of the keyboard-activated commands used in this text editor:

↑, ↓, ←, and → keys	Moves cursor around the document
<delete> and <backspace>	Deletes characters
<alt> f	Opens File pull-down menu; commands listed below: <ul style="list-style-type: none"> o Open file s Save file x Exit file p Print file
<alt> e	Opens Edit pull-down menu; commands listed below: <ul style="list-style-type: none"> u Undo last action a Select all text t Cut selected text c Copy selected text p Paste selected text l Delete selected text

To highlight a certain portion of text to edit, hold down the <shift> key and use the ↑, ↓, ←, and → keys to move the cursor. Alternatively, you may use a mouse.

3. Preparing for X-ray work-up

- 3.1) The diffraction data you will solve has already been transferred to your computer. All files are stored in the Data folder on the desktop. To view a list of files in your folder, double click on the Data icon.
- 3.2) Look in your Data folder and make sure that the original.raw and original.p4p files are present, where *original* is a name that has been assigned to the files by the crystallographer who collected the data (e.g., w25r). The .raw file is a list of the reflection intensities sorted by Miller index. The .p4p file (and ._ls file, if present) contain summaries of the data collection parameters and include the unit cell constants.
- 3.3) Next, you must tell the software which file contains your data. You do this by defining a

“project”. First, open the SHELXTL window by double clicking the SHELXTL icon.

- 3.4) Start a new project by clicking Project, dragging to New, and then releasing. Navigate through the folders (if necessary) until you see the original.raw file displayed in the window. Type in a name for the project and write it down - we recommend using the same name as the name of the .raw file. Then highlight the original.raw file with your mouse, and click Ok.
- 3.5) You should now see a window that has your project name and path. To view this window at any other time, click Project in the SHELXTL window, drag to Open, and release.

4. Setting up for the refinement – XPREP

- 4.1) Now you are ready to run XPREP, which is a program that will help you decide on a space group and set up the files necessary to solve your structure. With the SHELXTL window open, click on XPREP, drag to XPREP, and release.
- 4.2) XPREP will read the cell parameters (a b c α β γ) from the .p4p file. If no .p4p file is present, XPREP will ask you to enter these parameters manually. To do this, simply enter the appropriate values for a b c α β γ with spaces in between.
- 4.3) XPREP will next read and analyze the diffraction data in the .raw file. The first analysis is a table of lattice conditions. The table is headed P (primitive), A, B, C (end-centered), I (body-centered), F (face-centered), Obv, and Rev. Look at the last number in each column, which gives the average value of the intensities for that lattice condition divided by the error (I/σ). If, for any column, the value of I/σ is 3 or less (or significantly smaller than the values in all the other columns), the condition is a likely one for your crystal (the smaller the number, the better).

Usually, XPREP recommends which lattice condition is best; this recommendation is enclosed in brackets, e.g. [P]. If this is the case, hit <enter> to accept XPREP’s choice and go to step 4.4, unless you have reasons to do otherwise.

If XPREP has trouble recommending a lattice condition, it will show a question mark enclosed in brackets [?]. If this happens, look to see which lattice condition has the smallest number in the last line (smallest value for I/σ). Type in the letter corresponding to your choice (P, A, B, C, etc.), unless you know from an earlier attempt to solve the structure that this choice may not be correct. Then go to step 4.4

- 4.4) You should now be back in the XPREP main menu, and the suggested option should be to look for higher symmetry [H]. Hit <enter> to accept this option.

XPREP will first print out some information that you can ignore, and then it will display one or more options for the crystal system; each choice is printed in between -----dashed lines. (The seven possible crystal systems are triclinic, monoclinic, orthorhombic, tetragonal, trigonal, hexagonal, and cubic). Record the possibilities with low FOM values (FOM = figure of merit; note: triclinic *always* has FOM = 0). Ignore the comment below the table about the original cell.

- 4.5) If XPREP makes a recommendation for a crystal system, the recommendation will be enclosed in brackets, e.g. [A], which stands for choice **A**. Hit <enter> to accept its choice unless you know it to be wrong, and then go to step 4.6.

If XPREP has difficulty choosing a crystal system for the unit cell (it will show [?])

instead of a letter), it is best to pick the highest symmetry crystal system listed among the possible choices (for example, pick orthorhombic over monoclinic, and monoclinic over triclinic), unless you know that the higher symmetry choice does not lead to a solution of the data. Type in the appropriate letter of the best option, hit <return>, and go to step 4.6.

- 4.6) After the search for higher symmetry is done, you will be back in the main menu and the suggested option will be [S] to enter the space group determination routine. Hit <enter>.
- 4.7) The next suggested option will also be [S] to determine the space group. Hit <enter>. XPREP will then suggest a crystal system (triclinic, etc.). Choose the same crystal system as you did in step 4.4.
- 4.8) XPREP will then go through the lattice condition routine and make a suggestion (primitive, etc.). Choose the same option as you did in step 4.3.
- 4.9) Next, XPREP will give a table of possible space groups. If at least one space group appears in the table, go to step 4.10.

Occasionally, XPREP will be unable to find any space group that agrees with the systematic absences it finds. In such cases, go back to the XPREP main menu by hitting <enter> and change the tolerances (choice **T**). Make the minimum I/σ gap (choice **G**) half of its current size and the maximum mean I/σ (choice **A**) twice its current size. Choose the **E** option to exit to the XPREP main menu and then continue with step 4.4.

In some cases, XPREP will refuse to identify a space group even after adjusting the tolerances, and you will have to decide yourself what the systematic absences are. Look at the last two rows of the systematic absence exceptions table. A systematic absence is likely if $\langle I \rangle$ or $\langle I/\sigma \rangle$ is significantly less than "normal". From these decisions, you should be able to guess a space group. Consult an expert if in doubt. Enter the space group manually by typing its symbol (use a minus signs and parentheses to indicate rotoinversion and screw axes, respectively: thus P-1 or P2(1)/c). Then go to step 4.10.

- 4.10) In the table of possible space groups, the last number in each row is the CFOM number (which stands for combined figure of merit) for each space group. Usually, the smaller the CFOM, the better. If CFOM is 2 or less, the space group choice is a good one (although not necessarily the right one!). If it is 12 or greater, you will probably have problems solving the data set in that space group.

Write down the three or four space groups that have the lowest values of CFOM (if there are that many). Record the CFOM values for each.

Look at the space group with the lowest value of CFOM. If it is a centrosymmetric space group, go to step 4.11.

If the space group with the lowest value of CFOM is chiral or non-centrosymmetric, then answer the following two questions:

- a) Is there a centrosymmetric space group whose CFOM is within 3 units of the CFOM of the "best" space group?
- b) Does your crystal contain heavy scatterers (for Mo radiation, atoms with Z greater than 18)?

If the answer to either question is no, go to step 4.11.

If the answer to both questions is yes, it is likely that the actual space group is the centrosymmetric space group with the lowest CFOM. Skip step 4.11 and override XPREP's choice in step 4.12. If it proves impossible to find a solution in the centrosymmetric space group, you can always re-run XPREP to choose a different one.

- 4.11) Above the table of systematic absence exceptions is a number, labeled Mean $|E^*E-1|$, which has been calculated from the data in the .raw file. Compare the value with those expected for centrosymmetric and non-centrosymmetric space groups. If the Mean $|E^*E-1|$ value is closer (say) to that expected for a centrosymmetric space group, then you should check to make sure that the space group chosen by XPREP (or by you) is also centrosymmetric. If it isn't, it may be the wrong space group.

NOTE: Heavy scatterers (atoms with $Z > 18$) can distort the Mean $|E^*E-1|$ value and make it an unreliable indicator of whether the cell is centrosymmetric or not.

- 4.12) XPREP will give a recommended space group in brackets: e.g. [F] for option **F**. If you like XPREP's recommended choice, accept it by hitting <enter> and go to step 4.13.

If you wish to choose a different space group than the one recommended by XPREP (or if XPREP can't decide and shows [?]) as its choice), type the letter corresponding to the option you think is best (see steps 4.10 and 4.11), then hit <enter>, and go to step 4.13.

- 4.13) After you have chosen a space group, a screen will come up with a table of information about the current data set. Write down the space group, and from the line that reads "current cell", write down the values of a, b, c, α , β , γ , and the volume of the unit cell V, which appear in that order.
- 4.14) Hit <enter>. XPREP will then display a chemical formula for the compound or ask you to provide one. If the formula is not present, enter it manually (the program understands abbreviations such as Me, Et, Ph, etc.). If the formula is incorrect, type **F** and enter the correct formula.
- 4.15) Check the value that XPREP has calculated for Z (the number of formula units per unit cell). For organic and many inorganic molecules, a rough calculation for Z can be done manually:

$$Z \approx (\text{volume of unit cell in } \text{\AA}^3) / (10 \text{ times the number of atoms in molecule})$$

More sophisticated methods for estimating Z have also been discussed (see Girolami, G. S. *J. Chem. Educ.* **1994**, *71*, 962-964). Z is always an integer and very often is a power of 2 (1, 2, 4, 8, or 16), and more rarely other numbers (e.g. a multiple of 3 for trigonal and hexagonal space groups).

XPREP sometimes miscalculates this number and gives an unreasonable value such as 5; if this is the case, override it by typing **Z** and then enter in a new value. When you are happy with the value of Z, hit <enter> to return to the XPREP main menu.

- 4.16) **If this is your first time through XPREP, continue with step 4.17 below. If you are doing an absorption correction, stop here and return to section 13, step 13.4.**
- 4.17) Hit <enter> until XPREP asks for an output filename. The filename you enter should be *different* from that of the original input file; you can pick any name you like up to eight alphanumeric characters. Type the name (without an extension) and hit <enter>.
- 4.18) **Answer Y (for yes) when it asks whether to (over)write the intensity data file**

(default is no). This causes XPREP to create a file called filename.hkl.

NOTE: XPREP also creates a file called original.prp, where *original* is the original name of the .raw and .p4p files you started with. This .prp file contains a summary of all the information calculated and all the decisions made while running XPREP.

4.19) Quit the XPREP program (option **Q**).

4.20) Last, change the Project. Click on Project, drag to New, and then release. Navigate through the folders (if necessary) until you see the filename.hkl file displayed in the window. (If this file is not there, re-run XPREP beginning with step 4.1 and make sure to follow the instructions in steps 4.17 and 4.18.) Type in a name for the project and write it down - we recommend using the same name as the name of the .hkl file. Then highlight the filename.hkl file with your mouse, and click Ok. Go to Section 5.

5. Solving by direct methods – XS

The program XS helps you find a partial molecular structure that is consistent with the reflection intensities. The input files for this program are the two files created by XPREP: filename.ins (which contains instructions telling XS what to do) and filename.hkl (which contains the reflection data). NOTE: the .hkl file differs from the .raw file in several ways, but you do not need to worry about the differences. Both the .ins and .hkl files are required to run XS.

5.1) First, inspect the .ins file. To do this, click Edit, drag to Edit .ins, and click once more.

5.2) Make sure the **TREF** card (instruction) is present in the .ins file. Its presence tells the program XS to use direct methods to try to solve the structure. Direct methods should be used in the first attempt to solve a structure because, if successful, most of the atoms in the structure will be found all at once (vs. Patterson methods, which find only a few atoms).

5.3) Close the .ins file by clicking the box at the upper right corner of the window.

5.4) Run XS by clicking XS at the top of the SHELXTL window.

NOTE: If an error message appears on the screen when you try to run the program XS, open the .lst file by clicking Edit, dragging to Edit .lst, and clicking once more. Scroll down to the end of the file and try to determine what the program was doing when it aborted. This usually will give you an indication of what mistake you have made in the .ins file. To fix the problem, you will have to make a suitable change in the .ins file. If in doubt, consult an expert. Go to step 5.1 to begin this process.

5.5) As the program XS runs, it will print out some of its findings. Record the following values (if these numbers disappear from the screen before you can write them down, you can find them in the .lst file):

R(int): Measures how well the diffraction pattern matches the Laue group. R(int) should be less than 0.2 (less than 0.1 is better).

R(sigma): The average error in the intensities of the reflections. R(sigma) should be less than 0.2 (less than 0.1 is better).

If either of the above values is greater than 0.2, there are two possibilities: (a)

there are lots of weak data, or (b) the selected space group may not be correct. For case (a) you may be able to use the program SAINT to remove weak high-angle data from the .hkl file, while for case (b) you should consider running XPREP again to select another space group. Consult an expert if you encounter these problems.

Number of Systematic Absence Violations: Reflections which should have zero intensities for the selected space group but did not. This number should be less than 10. If it is greater than ten, check the rejected reflection table in the .lst file after XS finishes running (see steps 5.6 and 5.7).

Number of Bad (Inconsistent) Equivalents: Pairs of observed reflections which should have the same intensity but do not. This number should be less than 100 for most crystal structures. If it is greater than 100, you do not need to do anything if your crystal contains heavy scatterers (for Mo radiation, atoms with atomic numbers greater than 18), or if the crystal was a needle or a very thin plate. If the number of bad equivalents is greater than 100 and neither of these situations applies, then (a) you have done something wrong in XPREP and should re-run XPREP more carefully, or (b) the diffraction pattern is weak, and you should ask an expert about whether to use the SAINT program to delete high angle data.

CFOM (for the best solution): The lower the CFOM (combined figure of merit), the better. If CFOM is less than 0.1, then there is a good chance that a correct solution has been found, but some incorrect solutions can give surprisingly low values for CFOM. If CFOM is greater than 0.3, then it is unlikely (but not impossible) that a correct solution has been found.

- 5.6) When XS has finished, open the .lst file by clicking Edit, dragging to Edit .lst, and clicking once more. The .lst file contains a description of information that was created by running XS. Here is a brief summary of the file contents in order of appearance:

A copy of the .ins file
 Rejected reflection table (contains information about systematic absences)
 Inconsistent equivalents table (contains information about bad equivalents)
 Results of the direct methods (or Patterson) calculations
 Locations of found atoms
 Two-dimensional electron density map
 Distance and angle table for found atoms

Use the cursor arrows or <page up>/<page down> keys to scroll through the file.

- 5.7) If the number of absence violations was less than 10, skip to step 5.8. If the number was more than 10, scroll down until you find a table with the headings:

h k l F*F Sigma Why rejected

This is the rejected reflections table (although unfortunately it is not titled as such). Check that all of the values for **F*F** (the intensities of the “absent” reflections) satisfy one of two criteria: (a) F*F is less than 10, or (b) F*F is less than 5 times the **sigma** value. If some reflections satisfy neither criterion, then these reflections violate the systematic absences for your space group, and the space group may be wrong. You should consider running XPREP again to select another. If all reflections satisfy one of the two criteria, continue with the next step.

- 5.8) Continue scrolling down (a long way) until you come to a table with the following heading:

Try Ralpa Nqual Sigma-1 M(abs) CFOM Seminvariants

This table lists the various direct methods attempts. An asterisk will appear next to the CFOM value of the solution with the lowest CFOM. This is the attempt that XS will look

at more closely, because this attempt is most likely to give the correct solution to your structure. If it turns out NOT to be the correct solution, however, you may have to come back to this table and look at it more closely. For now, however, continue with the next step.

- 5.9) The rest of the .lst file contains a two-dimensional projection of the electron density map, and several tables of bond distances and angles between the peaks in this map. Although it is possible to inspect this information to determine whether a correct solution has been found, it is usually easier to look at the results in three dimensions using the program XP.
- 5.10) Close the .lst file by clicking the box at the upper right corner of the window.
- 5.11) **Make back-up copies of the .ins and .res files** by following the instructions in step 2.2. Go on to the next section.

6. Assessing the solution – XP

The program XP is a powerful graphics package that greatly assists in deciding whether a correct solution has been found by XS (and in identifying Q peaks found by XL).

- 6.1) In the SHELXTL window, click “XP” to run the XP program. A prompt should appear (usually after some messages which you can ignore).
- 6.2) Type “reap filename.res” where *filename* is the name of the .res file, and then type “fmol”. An atomic connectivity list should appear. The first atom on each line is connected to the atoms that follow the dash.

If no atomic connectivity list appears after typing “fmol”, your .ins and .res files are corrupted. Open your most recent uncorrupted backup copy of an .ins file, and rename it filename.ins, where *filename* is the name you have been using. Then go to step 10.1.

- 6.3) If this is your first time through XP, go to step 6.4.

If instead you have already been through XP once (or more times) before, and already know how to use the proj, diag, and kill commands, go ahead and use these commands to kill unwanted peaks. If some Q peaks need to be given labels, go to step 6.12. If all the non-hydrogen atoms have been given labels in previous cycles, go to step 6.16.

- 6.4) Type “proj”. An image of the structure should come up. This structure shows the locations of the major peaks in the calculated electron density map. You can rotate the structure by using the cursors to highlight the options at the right of the screen and hitting <enter>. Hitting <enter> again stops the rotation. Hitting <spacebar> will rotate the structure faster; hitting <spacebar> again will make it rotate slower.

NOTE: if you solved the structure by Patterson methods, only a few atoms will appear. If some of these atoms seem to form a reasonable local structure, then continue with step 6.10. If they do not form a reasonable structure, then you could try a cycle of least-squares refinement (exit XP, copy the filename.res file to filename.ins, and go to section 10), or you could try some of the “desperate measures” in section 8.

- 6.5) You can look at the labels for the peaks in the displayed structure by choosing the All Labels command in the menu. In the early stages of solving a crystal structure, most of the peaks are labeled with the letter Q and a number; this means that the chemical identity of the peak (i.e., what element it is) has not yet been assigned. (NOTE: In a

Patterson solution, there are no Q atoms.)

Realize that some of the unassigned peaks will be noise even if the solution found by XS is the right one. Q peaks with large numbers are more likely to be noise, but sometimes even a few of the Q peaks with small numbers (including Q1) can be spurious.

- 6.6) Look at the displayed structure to see if you can find chemically sensible groups of peaks (phenyl groups, a central atom with ligands around it, etc., if these are expected). *Even if you can identify only one or two atoms in the molecule at this point, this is often enough to proceed with the solution, especially if these are “heavy” atoms.*

Your goal is to try to pick out the real peaks from the noise peaks if possible. If you think you see how to do this (i.e., some of the displayed peaks make chemical sense), skip to step 6.10.

If the displayed structure does not make immediate sense, it is possible that your molecule sits on a symmetry element and only half (or less) of your molecule is shown. To test this possibility, exit the proj routine by hitting <esc>, and type “grow”. Then type “proj” again to see if the displayed structure is more recognizable. If it is, go to step 6.10; if not, go to the next step. If using the “grow” command made the display *extremely* complicated, go back to step 6.2 and start again, this time without using the grow command.

- 6.7) If the displayed structure still doesn't make sense, try to identify peaks that are clearly NOT part of the structure. Eliminating these peaks often makes “seeing” the structure easier. Use the cursor buttons to highlight the All Labels option. Generally, the Q peaks worth deleting fall into one of two categories:
- a) Q peaks that form chemically non-sensible distances or angles to nearby “real” atoms. This includes peaks that are “off by themselves”. To find the distances and angles around a particular peak, type “envi atomname”, where *atomname* is the label of the peak you are interested in (for example, “envi Q1”).
 - b) Q peaks with large numbers; these are more likely to be “noise” in the electron density map.

If possible, write down the Q numbers of the peaks that clearly fall into one of these two categories. If in doubt, do not write its number down.

- 6.8) Exit the proj routine, and then type “diag”. This brings up an image of the structure. Hit any key; the structure will contract into a smaller window, and the XP command prompt will reappear.

Use the kill command to delete the peaks you have selected in step 6.6. For example, if you wish to delete peaks Q2, Q14, and Q17, type “kill Q2 Q14 Q17” with spaces between each label. If you wish to delete peaks Q30 through Q39, you can use the command “kill Q30 to Q39” or “kill q3?”, where the question mark serves as a wild card.

If you accidentally delete peaks that are part of your structure, it may be wise to go back to step 6.2 and begin again from there. If the accidentally-deleted peak is a “light” atom, however, this is not necessary; the peak will re-appear later when you refine the structure in section 10.

- 6.9) Type “proj” and look at the structure again. You can repeat steps 6.7 and 6.8 if desired. If the structure makes sense now, go to step 6.10.

If the structure still does not make sense, exit `proj` and try killing all but the top 10 or 20 Q atoms in the list. Then use “`proj`” to see what is left. If some of the remaining atoms form a possible core structure, proceed to step 6.10.

If no part of the structure makes sense and if heavy atoms (larger than silicon) are present, go to section 7 and try the Patterson heavy atom method to solve the structure. If the structure doesn't make sense and heavy atoms are not present (or if the Patterson solution doesn't make sense either), go to section 8.

- 6.10) When you have made some sense of the displayed structure, the next thing you must do is to devise a systematic and rational atom labeling scheme for *all* atoms which you expect to find in the structure (*except* for symmetry-related atoms), even if they have not yet been located. You are allowed to change the labels of the atoms assigned by XS if you wish.

All atom labels should be the chemical symbol for its element followed by a number (or numbers plus letters). For example, N3 or OS4A are fine, but CP1 is not (because Cp is not the symbol for any chemical element).

Physically adjacent (i.e., mutually bonded) atoms should have numerically adjacent labels, if at all possible. For example, a five-membered ring should have its carbon atoms labeled in order from (for example) C1 to C5, and a second five-membered ring (for example) from C11 to C15. Make sure that no two atoms have the same label.

When you have finished devising your labeling scheme, rotate the molecule in `proj` until you get a convenient view of the molecule in which none of the atoms overlap (i.e., all atoms are visible). Then exit the `proj` routine by hitting `<esc>` and go to the next step.

- 6.11) Your next job is to assign labels to the real atoms in the structure, and to delete any remaining noise peaks. Before doing this, however, please read the following information.

IMPORTANT: Do *not* assign labels to symmetry-related atoms that have been generated by the `grow` command (these are usually indicated by a trailing letter; for example P1A or Q13C). Instead, when these atoms are highlighted, delete them by typing `<enter>`.

It will be helpful to know what “chemically sensible” bond distances are. For carbon, nitrogen, and oxygen atoms, at this stage in solving a structure, typical bond lengths are:

bond order	typical length
single	1.3 to 1.7 Å
double	1.2 to 1.5 Å
triple	1.1 to 1.3 Å

Other typical bond distances can be found in Pauling's *The Nature of the Chemical Bond*, or many chemistry textbooks. If any two peaks are closer than 1.0 Å to each other, one of the two peaks is not real (or there is disorder). Now go to step 6.12.

- 6.12) Type “`pick`”. (If you have already gone through one or more cycles of the least-squares program XL and are looking for atoms not found in previous cycles, you may wish to use the command “`pick $q`”, which will highlight only the unassigned Q peaks.)

XP will display a line diagram showing all peaks. Your job is either to give each peak a name from the labeling scheme you devised in step 6.10, or to delete the peak. Each peak

will be highlighted in turn (it will blink). The current name of the peak and the distances to nearby peaks will be shown near the bottom of the screen.

- To rename the peak, type the new name followed by <enter>.
- To delete the peak from the refinement, type <enter>.
- To leave the peak unchanged, type <spacebar>.
- To go back and correct an earlier mistake, type <backspace>.
- To zoom in on a region, type <tab>; type <shift><tab> to zoom out.
- To exit the routine without making any changes, type <esc>.
- To exit the routine but apply all changes made, type / (the slash symbol).

Note: It is often useful to write down all peak assignments made.

After you have typed in labels for the desired peaks, hit <spacebar> until you leave the pick routine (do not use <backslash>!).

- 6.13) If you wish to change the label of an atom after exiting the pick routine, there are several ways you can do this. Suppose that atom C10 was mistakenly given the label C11, and Q peak Q2 should be given the label C33. You can type “pick C11 Q2”; the pick routine will then only ask you for new labels for atoms C11 and Q2. Alternatively, you can use the **name** command: e.g., type “name C11 C10” and “name Q2 C33”.
- 6.14) Type “**kill \$q**” to delete any Q peaks that you might have left in accidentally. **Then type “fuse”** to remove symmetry-related atoms that you might accidentally have kept.
- 6.15) Next, **reorder the atom list** so that the atoms are ordered by element type and in increasing numerical order. Suppose you have ruthenium, phosphorus, and carbon atoms in your structure. Type “sort all” and then “sort \$ru \$p \$c” to order the list. It is not necessary to include \$h for the hydrogen atoms; they will automatically appear in the right place (after the atoms they are attached to). See also step 18.2 for more information about the sort command.
- 6.16) Type “fmol”. Look at the list of connections to see if they make sense.
- If any Q peaks are still listed, go back to step 6.14. If the atoms are not listed in some rational order, go back to step 6.15. Otherwise, go on to the next step.
- 6.17) Using the same *filename* as for the .hkl file, write an updated .ins file (with the new atoms you have assigned). To do this, type “file filename.ins”.
- The program will ask for the name of the .res file so that it can use the information to write the updated .ins file; change the default name only if necessary. Ignore the message about **MOLE** instructions, if such a message appears.
- 6.18) Exit XP by typing “exit”, and go to section 9.

7. Solving by Patterson methods – XP

If direct methods fail to give a correct solution and if heavy atoms (larger than silicon) are present, one can try Patterson methods to solve the structure.

- 7.1) First, recall backed-up version of the original .ins file generated by XS (the one that contains the **TREF** card); you made this backup in step 5.11. To open this file, click Edit, drag to Edit, release, and type in the name of the backed-up .ins file.

- 7.2) When the backed-up .ins file opens, delete the **TREF** card and in its place type **PATT** (which stands for Patterson).
- 7.3) Click File, drag to Save As, and release. Rename the file filename.ins, where *filename* is the name of the .hkl file. Close the window by clicking the box at the upper right corner.
- 7.4) Run XS again by clicking “XS” at the top of the SHELXTL window.
- 7.5) When XS stops running, open the .lst file by clicking Edit, dragging to Edit .lst, and clicking again. Scroll down toward the end of the file until you reach a table of what the program believes are the heavy atoms in the molecule. (To scroll down in Edit, use the Page Down or down cursor keys, or use the mouse to drag the button in the bar on the right of the window.) The table you are looking for looks like this:

Name	At.No.	x	y	z	s.o.f.	Min. dist. / PATSMF	
MO1	44.9	0.5000	0.5611	0.2500	0.5000	9.42	391.5
CU2	23.4	0.6525	0.5619	0.3391	1.000	5.09	2.55
						120.3	299.1

Most of this table is obvious, except perhaps for the pairs of numbers in the last several columns under the Min. dist. heading. The *upper* number of each pair is an interatomic distance. The upper numbers in the *first* column give the minimum distance between that atom and its nearest symmetry-related partner. Thus, for MO1, the nearest symmetry-related MO1 atom (which may be in the next unit cell or a symmetry-related partner in the same unit cell) is 9.42 Å away. Similarly, the shortest CU2•••CU2 distance is 5.09 Å.

The upper numbers in the second, third, etc., columns also give interatomic distances. Thus, the shortest MO1•••CU2 distance is 2.55 Å.

HELPFUL HINT: an easy way to read the distances. Consider the 2.55 Å distance in the table above. Start at the 2.55 Å distance, which is on the CU2 line, scan *up* the column until you reach a blank space, then look left. You see MO1. Therefore, the distance given is the CU2•••MO1 distance.

The *lower* number in each pair is a vector amplitude. The number should become bigger as the two atoms involved become better scatterers. Thus, MO1-MO1 (391.5) > MO1-CU2 (299.1) > CU2-CU2 (120.3).

- 7.6) Check the estimated atomic numbers for the heavy atoms found. If the program makes an atom assignment that is incompatible with the atomic number (for example it labels the atom OS2 but the peak actually has an atomic number of 14), make a note of this. If the program has misidentified atoms with similar atomic numbers (for example carbon and oxygen), this will not be a problem in the early stages of the refinement and you can sort this out later if the true identities are not obvious now. If the program has come up with atoms that you suspect are not real (because their distances and angles look wrong), make a note of those atoms too.

Check the interatomic distances. They sometimes can help you decide whether the Patterson solution is a reasonable one or not. See section 6, step 6.11 for some “chemically sensible” bond distances.

- 7.7) Sometimes, the .lst file will show more than one Patterson solution. If this is the case, it is worth inspecting the other solutions to see if they make better sense than the first one. If one of the latter solutions does make better sense, you will have to modify the .res file by hand so that the correct atoms and coordinates are included. See an expert if you are not sure how to do this.
- 7.8) Close the .lst file by clicking the box at upper right corner of the window. If the first Patterson solution seems like it is worth a closer look (or if you manually inserted another solution into the .res file), then go to section 6 and follow the instructions to inspect the structure.

If all the Patterson solution(s) are clearly bogus, then go to the next section.

8. What to do when you can't find a solution.

What do you do if neither the direct methods nor Patterson methods worked? Here are some things to try:

- 8.1) Open the backed-up version of the original .ins file from step 5.11 by clicking Edit, dragging to Edit, and typing in the name of the backed-up file. Replace the **TREF** (or **PATT**) card with a **TREF 1000** card, and then run XS by clicking the "XS" button at the top of the SHELXTL window. This will make XS try 1000 different solutions, rather than only 64. It may find a better solution than it did the first time. If the **TREF 1000** card doesn't work, try **TREF 2000**.
- 8.2) If the method in step 8.1 doesn't work, investigate other direct methods solutions that had different phase assignments than the "best" solution. In the .lst file from a direct methods run of XS, these alternative possible solutions will have slightly higher values of CFOM its seminvariant signs will be *different* from those of the "best" solution. Write down the six- or seven-number "Try" code for several alternative solutions, and modify the .ins file by changing the **TREF** card to **TREF -nnnnnn** (be sure to include the minus sign), where nnnnnn is the code number of one of the alternative solutions. Then run XS again. Investigate the solution to see if it makes sense; if it doesn't, try **TREF** with the code number of the third-best solution with seminvariant signs that differ from the first two.
- 8.3) If the methods in step 8.1 and 8.2 don't work, use XPREP to set up .ins and .hkl files in a lower-symmetry space group or lower-symmetry crystal system. You may have to manually override the XPREP recommendations for crystal system and space group to do this. It is sometimes effective to try solving the structures in the triclinic space groups P1 or P-1. If a solution is found by direct or Patterson methods, the atomic positions will have to be transformed to the correct space group. Consult an expert for this.
- 8.4) If all else fails, consult an expert.

9. Editing the .ins file generated by XP

The next step is to make sure that the filename.ins file is suited for use by the least-squares refinement program XL.

- 9.1) If this is the FIRST time you have come to this section, go to step 9.2. If this is the SECOND or subsequent time you have come to this section, skip to step 9.9.
- 9.2) Open the .ins file by clicking Edit, dragging to Edit .ins, and clicking again. Change the

PLAN 20 card to **PLAN -20** (note the minus sign); the minus sign changes the format of one of the tables you will inspect later. The absolute value of the number on this card indicates how many Q peaks will be included in the next .res file generated by XL. If more than 20 atoms are still missing in your structure, it is probably a good idea to increase the number on the **PLAN** card accordingly.

- 9.3) The sixth line from the top of the file is a line containing the **SFAC** card. This line should be recorded. For example:

SFAC C H O P O S

The order in which the elements appear determines some very important numbers called the **atom code numbers**, which are used by the least-squares refinement program XL. From the example above, carbon would be assigned the code number one, hydrogen would be assigned the code number two, etc.

- 9.4) The card following the **SFAC** card is the **UNIT** card, which gives the number of each type of atom within the unit cell (this can be calculated by multiplying the formula by Z). If the **UNIT** card looks like this:

UNIT 92 192 4 8 4

then there are 92 carbon, 192 hydrogen, 4 oxygen, 8 phosphorus, and 4 osmium atoms per unit cell (the atom counts are listed in the same order as on the **SFAC** card).

- 9.5) Farther down the file is a list of atoms. A typical atom line looks like this:

Os1 5 0.72716 0.95872 0.82487 11.0000 0.04

Key: **atom label** **atom code** **x** **y** **z** **site** **displacement occupancy factor** **parameter**

Check to make sure that the names of the atoms agree with your devised labeling scheme. For example, consider the following chart:

OS1	5	(other peak data)
P1	4	
P2	4	
C1	1	

If it is known that P2 is actually an oxygen atom, you should rename P2 to (say) O1 by moving the cursor down to the label you want to change, deleting the old name, and typing in the new name. **It is also crucial to make sure that the atom code number for that atom is correct.** For the present example, this means that for the new atom O1, the number in the second column of that line must be changed from 4 to 3;

O1	3	(other peak data)
-----------	----------	--------------------------

IMPORTANT: the atom code number in the second column tells SHELX what kind of atom it is, not the atom label in the first column.

- 9.6) All assigned atoms should appear *before* the **HKLF 4** and **END** cards, so move them if necessary using the cut and paste commands. (Anything in the file after the **END** card is ignored by all the SHELXTL programs; you do not need to delete these lines.)

- 9.7) It is very helpful (though not strictly necessary) to list the atoms in some rational order, so that physically adjacent atoms are near one another in the list. If the atoms are out of order, either reorder them manually using the cut and paste commands, or reorder them using the “sort” command in XP (see step 6.15).
- 9.8) Save the modified .ins file (click File, drag to Save, and click. Click on Save File) and then close the window by clicking the box at the upper right corner of the window.
- 9.9) **Make a backup of the .ins file** using the instructions in step 2.2. Then go to the next section.

10. Least squares refinement – XL

After a solution has been found by XS, the program XL can be used to find missing atoms and to refine the structure further. Several XL cycles are usually needed.

- 10.1) Run XL by clicking “XL” at the top of the SHELXTL window. This will cause the program to execute four cycles of least-squares refinement.

NOTE: If an error message appears on the screen when you try to run the program XL, open the .lst file by clicking Edit, dragging to Edit .lst, and clicking once more. Scroll down to the end of the file and try to determine what the program was doing when it aborted. This usually will give you an indication of what mistake you have made in the .ins file. To fix the problem, you will have to make a suitable change in the .ins file. After you have made your change and saved the modified .ins file, begin again with step 10.1.

- 10.2) As XL runs, a brief description of the results are printed out on the screen. The **wR2** value, which is the least-squares residual, should become smaller for each successive cycle. The magnitudes of **Mean shift/esd** and the **maximum** (shift/esd) should also become smaller each cycle; these values tell you how much the various parameters (atomic coordinates and displacement parameters) are changing from least-squares cycle to least-squares cycle.

If **wR2** or the **shift/esd** values increase significantly from one cycle to the next, the least-squares refinement has probably diverged. If this happens, check the **U11** values for the identified atoms to see whether any of them have become unrealistically large (see steps 10.5 and 10.6). If this is not the problem, try another method to refine your model, modify the .ins file accordingly, and rerun XL.

Sometimes you will get a message warning you that an EXTI or SWAT correction may be needed; ignore this message for now.

Record the values for **wR2** before cycle 5, number of **data**, and number of **parameters**.

- 10.3) Inspect the values of **maximum** (shift/esd) before cycle 5. If this value is between +15 and -15, go to step 10.4. If this value is *exactly* +15 or *exactly* -15 and **wR2** is still dropping rapidly, then you should carry out a few more cycles of least-squares refinement. Do this by **copying the .res file to the .ins file** (Click Edit, drag to Copy .res-.ins, and click again) and then returning to step 10.1.
- 10.4) **If you have accomplished the following three things:** (a) you have found all the non-hydrogen atoms in your structure, (b) you have typed in labels (according to your labeling

scheme) for all them, and (c) the **wR2** factor is essentially constant, then **go to Section 11**. If not, continue with step 10.5.

10.5) Open the .lst file by clicking Edit, dragging to Edit .lst, and clicking again. The .lst file contains the following information in order:

The contents of the .ins file that was used for that set of refinements
 A table of atomic radii and connectivities
 Tables of rejected reflections and bad equivalents.
 A successive report of results for the last cycles of least squares refinements.
 A list of atoms with their newly-refined parameters
 A table of disagreeable reflections
 A table of bond distances and angles for “found” atoms only
 An electron density map
 A table of bond distances and angles for both “found” and Q atoms

10.6) The first thing in the .lst file is a list of the contents of the .ins file. Scroll down past this list and past several following tables until you come to the list of atoms with their newly-refined parameters. Each line in the list will look like this:

OS1	5	0.78725	0.95964	0.82595	1.0000	0.01482
		0.00003	0.00005	0.00003	0.0000	0.00037

Key: **Label Code** **x** **y** **z** **s.o.f.** **U11**

The second line (if present) gives the errors in the parameters listed in the first line.

Inspect the displacement parameter **U11** for each atom. The **U11** values for all the non-hydrogen atoms should be equal within a factor of 2 or 3, and most atoms should have a **U11** that is less than 0.10. If **U11** for any atom is greater than 0.10, there could be large thermal motion or disorder. Such effects are usually seen for atoms or groups near the periphery of the structure with rotational or conformational degrees of freedom. For example, cyclopentadienyl rings, CF₃ groups, long alkyl chains, and solvent molecules often show large displacement parameters.

If **U11** for any atom is more than 0.30, there are several possibilities that you should consider.

- The atom is disordered over two or more locations. Consult an expert.
- There is actually no atom at that location. To check this, delete the atom from the .res file, **copy the .res file to a .ins file**, and run XL again to see if no large difference peak appears at the coordinates of the removed atom.
- There is an atom at that location but it has been misidentified and actually is an atom with a smaller atomic number. Check to see that the atom code for the atom is correct (e.g. 1 for a carbon atom, etc.).

If **U11** for any atom is less than 0.01, there are several possibilities.

- There is no atom at that location. Often, “ghost peaks” appear at special positions in the difference map. Remove the atom, **copy the .res file to a .ins file**, and rerun XL. If **wR2** remains the same or decreases, then it is likely that the atom was not real.

b) There is an atom at that location but it has been misidentified and actually is an atom with a larger atomic number. Check to see that the atom code is correct.

10.7) Continue scrolling down the .lst file until you come to a map of labels and numbers. (Note: if this map is missing, you forgot to add the minus sign to the PLAN card; see step 9.2.) This map is a two-dimensional projection of the contents of the asymmetric unit; it is an electron density map. Each label and number stands for a peak in the electron density map.

NOTE: an asterisk (*) means that two peaks are projected on top of each other and their labels overlap.

The atoms whose locations have already been found will be assigned the labels you gave them (e.g. OS1 for osmium atom 1). The rest of the peaks in the electron density map are not currently assigned a label and are represented by a number (e.g. 2 for the second largest unassigned peak in the electron density map).

10.8) Scroll down some more until you come to a table headed:

Atom Peak x y z Sof Height Distances and Angles:

This table is very useful, but it takes some practice to read it properly.

Atom: This column gives the atom labels. The first entries in the table are atoms that have already been given labels by you. These atoms are followed by *unassigned* peaks in the electron density difference map. (The PLAN card dictates how many of these unassigned peaks will be listed.) The program gives these unassigned peaks a temporary label, which is just a number: 1, 2, 3, etc. These unassigned peaks (which are the Q peaks that appear in XP) are listed in order of *decreasing* electron density.

Peak: for unassigned atoms (i.e., peaks labeled 1, 2, 3, etc.), this column gives the electron density of the peak in electrons per cubic Ångstrom.

x, y, z: atomic coordinates of the peak.

Sof: site occupancy factor (normally = 1, except for atoms on special positions or disordered atoms).

Height: in the 2-D map, the elevation of the atom above or below the plane of the paper.

Distances and Angles: best explained by an example. Suppose we have:

Atom	Distances and Angles				
OS1	0	P2	2.281		
	0	P3	2.291	92.5	
	0	2	2.162	90.4	87.0

This table is interpreted as follows:

The OS1-P2 distance is 2.281 Å.

The OS1-P3 distance is 2.291 Å.

The OS1-2 distance is 2.162 Å (the identity of peak 2 is presently unassigned)

The P2-OS1-P3 angle is 92.5°.
 The P2-OS1-2 angle is 90.4°.
 The P3-OS1-2 angle is 87.0°.

You should look at these to make sure you understand how the table is written (especially how the angles are listed).

HELPFUL HINT: an easy way to read the angles. Consider the 90.4° angle in the table above. The three atoms involved are: atom 2 (the line the 90.4° angle is on), OS1 (the atom at the extreme upper left of the table), and P2 (start at the 90.4° angle, scan up the column until you reach a blank space, then look left). Therefore, the 90.4° angle is the 2-OS1-P2 angle.

If a number other than 0 had appeared in front of (say) P2, it would mean that the distances and angles listed are for a symmetry-related partner of P2.

- 10.9) NOTE: This table only contains one entry for each set of symmetry-related atoms in the unit cell. This means, for example, that if your unit cell contains 4 molecules (i.e., $Z = 4$) but all the molecules are related by some symmetry element in the space group chosen, then only one molecule will be displayed. If your molecule sits ON a symmetry element, then only half (or less!) of the molecule will be shown.

The table is sometimes divided up into several separate sections; each section gives the distances and angles for one molecule in the asymmetric unit. The program defines a molecule as a collection of atoms separated in space from other collections of atoms. If a peak you are looking for is not in the first section (which is for Molecule 1), search down in the .lst file until you come to another section for Molecule 2 (or 3 etc.). You may be able to find your atom there.

- 10.10) The latter sections of the bond distance and angle table give information about the largest unassigned peaks in the electron density map (peaks 1, 2, 3, etc.). **These are the Q peaks that appear in the XP graphics routine, but the letter Q is omitted in the .lst file.** There are several possible assignments of these peaks.

- a) The peak is an atom that you should incorporate into your structure. The electron density in electrons per cubic Angstrom (given in the second column, labeled 'Peak') gives a clue to the atomic number of the atom. Typical electron densities are (don't take these too seriously): hydrogen atom (0.3 to 0.8 e/Å³); carbon, nitrogen, or oxygen atom (2 to 12 e/Å³). A more general rule: if the electron density (in e/Å³) is x, then the atomic number is usually between x and 3x.
- b) The peak is an artifact. Two types of artifact peaks are most common: those that lie near a heavy scatterer, and those that lie on a symmetry element. Peaks located within 1 Å of a heavy scatterer are commonly seen and are caused principally by errors in the absorption correction.
- c) The peak is noise. Noise peaks generally have electron densities of less than 0.5 e/Å³.

If peaks with significant electron density are not near the other molecules in your unit cell, they may be due to solvent molecules.

If any two non-hydrogen atoms are closer than 1.0 Å to each other, then there is disorder. Consult an expert.

10.11) Go to step 6.1 to inspect the structure by using the graphics package XP.

11. Checking for missed crystallographic symmetry – PLATON

Sometimes you can find a solution to the X-ray data set in the wrong space group. Almost always, the true space group has more symmetry elements than the one you are in. If you have solved the structure in a wrong (i.e., lower symmetry) space group, several warning signs may appear. First, variables that should be independent turn out to be highly correlated (this is shown by the presence of numerous correlation coefficients of over 0.8 in the list of strong correlations in the .lst file). Second, chemically equivalent bond distances may vary over an unusually wide range (for example, the C-C bonds in a phenyl ring are not equal to 1.40 ± 0.03 Å but vary from 1.2 to 1.6 Å). Third, chemically similar groups may not behave similarly in the refinement (i.e., it is easy to find all the atoms in one phenyl group but difficult to find them in another phenyl group).

To check whether you have misassigned the space group, use the program PLATON (if you do not have a copy, download it from the following website: www.cryst.chem.uu.nl/platon). PLATON looks at the coordinates of the atoms found in the structure to see whether any symmetry relationships between the atoms have been missed. PLATON can find symmetry elements that relate different molecules, and it can find symmetry elements that relate different parts of the same molecule, if these symmetry elements are actually present.

To analyze your structure with PLATON, it is helpful if you have located and assigned the *same number and same kinds* of potentially symmetry-equivalent atoms in each independent molecule (or in each half of the same molecule). In other words, do not run PLATON if you have found many more atoms in one molecule than in another potentially symmetry-equivalent molecule.

- 11.1) Click on the PLATON icon. From the Files pull-down menu, click on Select Data File, and navigate through the folders to open the file filename.res.
- 11.2) From the Platon menu, select Command Line. At the prompt, type “calc missym”.
- 11.3) Ignore the caveat and the message about standard settings. If no “missed” symmetry is found, PLATON will tell you this. In this case, type “exit” to exit PLATON, click on the box at the upper right to close the window, and continue with section 12.

If PLATON *does* find missed symmetry, it will give you a message to that effect. The missed symmetry elements will also be labeled with an asterisk (*). Go to step 11.4.

- 11.4) If PLATON finds missed symmetry, it will also tell you what Laue class of space groups should be investigated. Write this information down. PLATON will also tell you the locations of missed symmetry elements; write this information down also. Then type “exit” to close PLATON.
- 11.5) Save a backup copy of your old .res file (see step 2.2) just in case the following step does not work.
- 11.6) You may be able to transform your old solution into the correct one by using the “push” command in XP to move the origin, writing a new .ins file, closing XP, adding the appropriate **SYMM** cards to the .ins file, and then using the “fuse” command in XP to eliminate duplicate atoms. Ask an expert for help if you want to do this.

If this procedure is successful, go to step 10.1. If not, go to step 11.7.

- 11.7) If it proves difficult or impossible to modify the old solution, you can try re-solving the data set from scratch in the correct space group. To do this, click on Project, drag to Open, and then release. Highlight the project corresponding to the original.raw file (see step 3.2), and then click OK.
- 11.8) Use XPREP to select the appropriate higher-symmetry space group, and finally use XS to try to find a solution in the new space group. Begin this procedure by going to step 4.1.

12. Turning atoms anisotropic and adding hydrogen atoms

Until now, the identified atoms in the crystal structure have been refined isotropically (i.e., as spheres). Owing to thermal motion, however, the electron density distribution of an atom usually has an ellipsoidal shape. Letting the atoms refine as ellipsoids is called anisotropic refinement.

Also at this point, it is often worthwhile to begin adding hydrogen atoms to your model. If you wish to locate hydrogen atoms and refine their positions, you can do this just as if the hydrogen atom were like any other atom: find it in the difference map, rename the Q peak, and add the peak to the list of atoms above the **HKLF 4** card. It is usually easier to locate and refine hydrogen atom positions after you have made the non-hydrogen atoms anisotropic and after you have carried out an absorption correction (see section 13).

If there are hydrogen atoms whose positions are hard to find or are relatively unimportant (such as hydrogen atoms bound to carbon atoms in organic groups), it is preferable to place them in idealized positions calculated from the locations of nearby atoms. This approach is often advantageous because it keeps the number of parameters to a minimum (which is a good thing): putting hydrogen atoms in idealized positions requires 0 or 1 parameter, whereas refining their locations requires 3 or 4.

- 12.1) Open the .res file by clicking Edit, dragging to Edit .res, and clicking.
- 12.2) If you would like to turn all the non-hydrogen atoms anisotropic, insert an **ANIS** instruction on a new line between the UNIT and FVAR cards. Often, however, it is risky to turn all the non-hydrogen atoms anisotropic at once. You can turn selected atoms anisotropic by using commands like **ANIS \$OS \$P** (which would turn all the osmium and phosphorus atoms anisotropic), or **ANIS C1** (which would turn only atom C1 anisotropic). Hydrogen atoms are never refined anisotropically.

Turning an atom anisotropic adds five parameters to the least-squares refinement. If the number of parameters exceeds about one tenth of the total number of reflections (or about one fifth of the number of observed reflections with $F_o > 4 F_c$), then the least squares refinement tends to become unstable and divergent. For this reason, one should not indiscriminately turn atoms anisotropic. If forced to limit the number of anisotropic atoms, leave some or all of the lighter atoms isotropic.

Once an atom is turned anisotropic, it stays anisotropic for all subsequent least-squares cycles. To turn an anisotropic atom isotropic, delete the last five displacement parameters (**U22** etc.) manually from the .res file or use the “isop” command in XP.

- 12.3) Save the .res file (click File, drag to Save, release, then click Save File), and then close the file by clicking the box at the upper right corner of window. **Copy the .res file to .ins file** (click Edit, drag to Copy .res-.ins, and release). Make a backup copy of the .ins file

by following the instructions in step 2.2.

- 12.4) Run XL by clicking “XL” at the top of the SHELXTL window.

Sometimes, problems are encountered with anisotropic refinements. Occasionally the value for **mean shift/esd** and **maximum** shift/esd will increase (the least-squares refinement is diverging). If this is the case, check the .res file to make sure that all the identified atoms have the correct code number (see section 9, step 9.3). If this is not the problem, try refining only some of the non-hydrogen atoms isotropically. To do this, open the .ins file by clicking Edit, dragging to Edit .ins and clicking. Modify the **ANIS** instruction appropriately.

NOTE: The .res file generated by XL will no longer show the **ANIS** card, but the atoms that have been turned anisotropic will have six displacement parameters instead of one.

- 12.5) Open the .lst file by clicking Edit, dragging to Edit .lst and clicking. Scroll down until you come to the table headed Principal Mean Square Atomic Displacements.

If the message “non-positive definite” appears after one or more of the atoms, then one of the axes defining the thermal ellipsoid for that atom has a negative (or zero) value. This is a bad sign; perhaps you are in the wrong space group, or have misidentified an atom. Consult an expert.

If the message “may be split into ...” appears after one or more of the atoms, then the atoms has refined to a very thin sausage or tube shape. This is a sign of possible disorder. Consult an expert for what to do next.

If neither message appears, go on to the next step.

- 12.6) If you wish to turn additional atoms anisotropic, return to step 12.1. If not, go to the next step.
- 12.7) In the last stages of solving the structure, you can go “hydrogen hunting”. **Before you do this, however, make sure that the atoms have been sorted properly (see step 6.15).**

If you have a very good data set, it may be possible to spot the hydrogen atoms in the difference map and assign the Q atoms just as you would for any other atom.

Usually, however, it is not worth adding lots of parameters to your solution to define the positions of the hydrogen atoms. Instead, it is often possible to add hydrogen atoms in “idealized” positions calculated based on simple ideas about bonding and minimizing steric repulsions.

There are two ways to add hydrogen atoms in idealized positions. One way is to use the “hadd” command in the graphics package XP. This method is quick and easy, but doesn’t always work. The second way is to insert the hydrogens by modifying the .res file by hand. This is also pretty easy.

- 12.8) Run XP by clicking “XP” at the top of the SHELXTL window. Type “fmol less \$q”, followed by “hadd”. Use the “proj” command to inspect the structure, making sure that every methyl group has three hydrogen atoms, etc. The “hadd” routine uses bond distance to decide how many hydrogen atoms to attach. Therefore, if the bond distances are wrong, then the program will attach the wrong number of hydrogens.
- 12.9) If the program has not added the hydrogen atoms exactly correctly, then go to step 12.10

and insert the hydrogen atoms by hand.

If the program has added hydrogen atoms in all the right places (check closely!), type “file filename.ins” to write a new .ins file with the hydrogen atoms added, and then exit XP. Use the Edit .ins command to open the .ins file, and change all the AFIX 33 commands (if any) to AFIX 137. This change tells the program to let methyl groups rotate to find the best rotameric conformation. Then save the .ins file and go to step 12.12.

- 12.10) Type “exit” to exit XP, and open the .res file (click Edit, drag to Edit .res and click). Add the following command(s) anywhere between the UNIT and WGHT cards:

HFIX n atoms

where n = idealized position code for H atoms (for a full list, see the Siemens manual), and atoms = atom name(s), in sequence, separated by a space. For example:

HFIX 137 C6 C7 C8 C9

This card puts H atoms on atoms C6-C9, making them all idealized methyl groups, and allowing them to refine as a rotating group. The thermal parameters for the H atoms are defined to be a certain number (1.0 - 1.5) times the thermal parameter of the carbon the H's are attached to (the exact number is determined by the value of n). For example, the thermal parameter **U** for the H atoms in a methyl group is set equal to **1.5 times U** of the carbon on which the hydrogen atoms reside. If more than one type of hydrogen atom is present, use multiple **HFIX** lines.

Commonly used **HFIX** values (for n):

137 = idealized rotating CH ₃ group	43 = idealized aromatic CH
33 = idealized non-rotating CH ₃ group	13 = idealized tertiary CH
(to be used if 137 doesn't work)	23 = idealized secondary CH ₂

NOTE: Upon running XL, the **HFIX** commands will be translated into **AFIX** commands; the latter will be automatically placed in the atom list by SHELXTL, and idealized H atoms will be added to the atom listing.

- 12.11) **Copy the .res file to the .ins file** (click Edit, drag to copy .res-.ins, and click)
- 12.12) Run XL by clicking “XL” at the top of the SHELXTL window.
- 12.13) Check the **maximum** shift/esd for the last cycle. If this number is less than 1, then go to section 13.

If the maximum shift/esd is greater than 1, then copy the .res file to the .ins file (click Edit, drag to copy .res-.ins, and click), and run XL again. If the maximum shift/esd is still greater than 1, then note which parameter(s) are not converging by looking at the summaries of the least squares cycles in the .lst file. If these parameters are the locations of hydrogen atoms in a methyl group, then change the **AFIX 137** card(s) for the appropriate methyl group(s) to **AFIX 33**.

- 12.14) Go to step 12.12.

13. Correcting for absorption – XPREP

Because most crystals are not shaped like spheres, some crystal dimensions will be larger than others. X-rays that travel longer distances through the crystal will be absorbed and scattered more by the crystal than those traveling shorter distances. This effect, if not corrected for, will skew the reflection intensities. The absorption correction corrects the intensity of each reflection for the distance the X-rays must travel through the crystal for that reflection.

Absorption corrections can only be done when the complete formula of the unit cell is known, including solvent molecules (if these are present). Generally, this information is known only after all of the non-hydrogen atoms in the unit cell have been located.

The steps in this section will guide you through the procedure, but a few comments about what you will be doing are appropriate. The program XPREP contains the software needed to correct the observed reflection intensities for absorption. (The *first* time you went through the program XPREP, you did not do an absorption correction, and the .hkl file that was written contained intensities that were *not* corrected for absorption.) You will rerun XPREP on the *original* raw data (original.raw; the file from which filename.hkl was created). After correcting the measured intensities for absorption, XPREP will write a new .hkl file (with a new name – different from the first one) that contains the absorption-corrected reflection intensities. You will then have to rename your old .ins file (which contains the atomic coordinates of the atoms in your structure) with this new name, change the Project, and run XL on the new .hkl file. The absorption-corrected data should give you fewer systematic absence violations, fewer bad equivalents, and a lower wR2 factor (all good things!).

- 13.1) Open the original .raw file by clicking Project, dragging to Open, and choosing the original.raw dataset, where *original* is the name of the file you opened in step 3.2.
- 13.2) Run XPREP by going to the SHELX window, clicking XPREP, dragging down to XPREP, and clicking again.
- 13.3) **Now go to step 4.2**, and continue in that section up to step 4.15, making sure to select the same space group in which the crystal structure was solved.
- 13.4) You should have just completed step 4.15; if not, go back to step 13.3 and begin again.
- 13.5) In the lower right corner of the table on the XPREP main menu page is listed the value of μ , the absorption coefficient, in mm^{-1} . Write this number down. (Note: if $\mu = 0.000$, you have done something wrong, probably not doing steps 4.14 and 4.15 properly. In this case, quit XPREP and start again at step 13.1.)
- 13.6) Select option **A** to enter the absorption correction routine.

The preferred method of correcting for absorption is the face-indexed method, and normally you should go to step 13.7 next. Only if a face-indexed correction is not possible for some reason should you skip to the psi-scan correction in step 13.16.

- 13.7) Face-indexed absorption corrections require that crystal faces have been assigned Miller indices and the crystal dimensions have been measured. Choose **F** from the menu to start the face-indexed correction.
- 13.8) XPREP will ask if the data are indexed on the original cell; hit <enter>
- 13.9) XPREP will next read the .p4p file and display the h k l values for each crystal face, and the distance d between each crystal face and the center of the crystal. If these values are not in the .p4p file, you will have to enter them by hand.

Record the h k l and d values for the *shortest* distance (this information may be useful later), and then hit <enter> to continue.

- 13.10) The X-ray wavelength will be displayed; hit <enter> unless you have reason to do otherwise.

- 13.11) The value of μ will be displayed (it should be the same as before). Hit <enter>.
- 13.12) Accept the defaults for the next three questions (the last one asking about an input file).
- 13.13) XPREP will then begin the correction process, which will take a few minutes to complete. When it is finished, record the values for **min** and **max** transmission.
- 13.14) Hit <enter> to return to the absorption correction menu, and then **W** to write a new .hkl file containing the absorption-corrected data. Give the .hkl output file a new name (e.g. face) that is different from the one you have been using.
- 13.15) To perform a psi-scan correction, go to step 13.16. Otherwise, skip to step 13.36.
- 13.16) Unlike face-indexed methods, psi-scan absorption corrections require only a general knowledge of the size of the crystal; The method is hard to defend theoretically, and tends to “correct” for much more than absorption. The psi-scan method should be avoided whenever possible, and a face-indexed correction applied instead. Proceed with the following at your own risk.

In the absorption correction menu, choose option **C** (change current dataset). Enter 1 for the current dataset index. The current dataset arrow near the top of the page should now be pointing to the original.raw file. Then choose **P** to start the psi-scan absorption correction.

- 13.17) XPREP will ask if the data are indexed on the original cell; hit <enter>.
- 13.18) The X-ray wavelength will be displayed; hit <enter> unless you have reason to do otherwise.
- 13.19) **When prompted for the filename for psi-scan data, enter original.raw.**
- 13.20) If your crystal is chiral, XPREP will ask if you wish to treat Friedel opposites as equivalent; answer yes unless you have reason to do otherwise.
- 13.21) XPREP will then ask for the tolerance for the direction cosines, the I/σ threshold, and the resolution threshold; again accept the defaults by hitting <enter>.
- 13.22) After a short delay, you will be asked to choose either **E** for an Ellipsoidal or **L** for a Laminar correction. Laminar corrections are better for platy (thin) crystals, whereas ellipsoidal corrections are better for 3D-like crystals. If you are in doubt about which type of correction is best, run both types to see which one gives the better **R(int)** value.
- Choose **E** to run the ellipsoidal correction first. (If you are doing a laminar correction instead, type **L** and skip forward to step 13.31.)
- 13.23) Calculate value for $\mu*r$ (μ times r), where μ is the linear absorption coefficient in mm^{-1} and r is the radius of the shortest dimension of the crystal in mm (these values should have been recorded in step 13.5 and 13.9). Typical values for r are between 0.01 and 0.15. Round off the $\mu*r$ product to one significant figure and enter this number.
- 13.24) XPREP will then carry out the absorption calculation, which will take a few seconds. Below is an example of the absorption correction results:

Merging R(int) = 0.0980 before correction
Merging R(int) = 0.0289 after correction

Maximum Transmission = 0.967. minimum transmission = 0.437

Record the value of **R(int) after correction** and **maximum transmission**.

After the correction is done, hit <return> so that the intensities are not printed.

- 13.25) Select **R** (repeat) from the menu and repeat steps 13.22 through 13.24, entering two times μ^*r and 0.5 times μ^*r (rounded to one significant figure) when asked for this number.

In some cases, a message appears saying that a laminar correction might work better. In this case, write down the indices (h k l values) the program suggests. Even when this message appears, the ellipsoidal correction might give better corrected data.

- 13.26) For the three values of μ^*r that you entered, see which one gives the lowest value for **R(int)** but which has a **maximum transmission** value that is less than 1.000. **Repeat the ellipsoidal correction one more time** with this value of μ^*r , because the absorption correction is always performed with the last value entered into the computer. When the correction is finished, process the data (select **P**), accepting the default input data filename [original.raw].

- 13.27) Write a new .hkl file for the ellipsoidally-corrected data by choosing **W** from the absorption correction menu. Enter a new output file name (e.g. psiell).

- 13.28) If XPREP never recommended doing a laminar absorption correction, go to step 13.36. If at any time a message appeared recommending a laminar correction, go to step 13.29.

- 13.29) In the absorption correction menu, choose **C** to change the current dataset index (enter 1, not the filename, when prompted). Check to make sure the original.raw dataset is selected by the "current dataset" arrow.

- 13.30) Select **P** to do a laminar psi-scan correction and follow the instructions in steps 13.17-13.22, selecting **L** when asked which type of correction to apply.

- 13.31) Enter the h k l values (with spaces in between) for the shortest crystal dimension (recorded in step 13.4) or for the indices recommended by XPREP when running the ellipsoidal corrections. (If these are different, carry out both laminar corrections separately and see which gives the lowest value for **R(int)** after correction).

- 13.32) When prompted for the minimum glancing angle, enter 3. This number tells XPREP to de-emphasize reflections for which the incident or diffracted beams lie within 3 degrees of the laminar plane.

- 13.33) Record the number of data suppressed. If the number of data suppressed is large compared to the total number of reflections in the data set (more than 15%), the laminar absorption correction should not be used.

- 13.34) XPREP will then carry out the laminar absorption correction, which will take a few minutes. When the correction is finished, record the **min** and **max** transmissions (don't print the intensities).

- 13.35) If R(int) is larger than that for the best ellipsoidal correction with **max** less than 1.000, skip to step 13.36.

Otherwise, process the data (select option **P**), accepting the default filename. Hit <enter> to return to the absorption correction menu and choose **W** to write a new .hkl file for the

laminar psi-corrected data. Enter a new output filename (e.g. psilam) when asked for it.

- 13.36) Exit XPREP. There now should be an .hkl file containing absorption-corrected data for each of the methods you used (face.hkl, psiell.hkl, or psilam.hkl).
- 13.37) To use XL to refine against the absorption-corrected .hkl files you have just generated, you must create .ins files with the same name: i.e., you need a face.ins file to go along with the face.hkl file. Since you have already solved the structure, it is not necessary to solve the structure again; instead, the best thing to do is to start with the .res file that you were using before carrying out the absorption correction.

Go to the Edit pull-down menu, and click Edit .res, to open the filename.res file. Then go to the File pull-down menu, click Save As, and type in the new name “absfile.ins”, where *absfile* is one of the names of the absorption-corrected data (e.g., face.ins).

Repeat this procedure (with different names for absfile) for any other absorption-corrected .hkl files you generated above. This step creates .ins files containing the proper information for continuing the least-squares refinement on the absorption-corrected data.

- 13.38) Start a new project by clicking Project, dragging to New, and then releasing. Navigate through the folders (if necessary) until you see one of the absfile.ins files displayed in the window. Type in a name for the project and write it down - we recommend using the same name as the name of the absfile.ins file. Then highlight the absfile.ins file with your mouse, and click Ok.
- 13.39) Run XL again by clicking “XL” at the top of the SHELXTL window. Normally, the number of bad equivalents should be lower than the number seen before the absorption correction. If it is not, consult an expert.
- 13.40) Continue with the next section.

14. Extinction, weighting schemes, and absolute configuration

This section describes three aspects of solving a structure that now need attention. 1) You must test for the presence of extinction, a phenomenon in which some of the most intense reflections have reduced intensities owing to the “mosaicity” of the crystal. 2) You must choose a weighting scheme, which helps the refinement by putting the most weight on the data that have the smallest errors. 3) If your space group is chiral or non-centrosymmetric, you must check to see if the correct absolute configuration has been chosen.

- 14.1) In some cases, a warning that an EXTI (i.e., extinction) correction is recommended may appear on the screen at the end of an XL run. It is useful, however, to check for extinction even if such a message does not appear. To add an extinction parameter, insert the card **EXTI .001** after the **PLAN** card in the .res file. This parameter will refine automatically with subsequent XL runs.
- 14.2) Toward the end of the refinement one can change the weighting scheme used in the refinement to the new value suggested by the XL program. The new value for the **WGHT** card is listed in the .res file just before the Q atoms. Copy the WGHT card at the bottom of the file (highlight the line with mouse, click Edit, drag to Copy, and click) and paste over the existing WGHT card near the top of the file (highlight top WGHT card with mouse, click Edit, drag to Paste, and click). Note: copy the entire WGHT card (both numbers!).

- 14.3) **Copy the .res file to the .ins file** (click Edit, drag to Copy .res-.ins, and click), then run XL by clicking “XL” at the top of the SHELXTL window.
- 14.4) Repeat steps 14.2 and 14.3 until the EXTI and recommended WGHT cards do not change and the value of **wR2** no longer decreases.

The extinction parameter will refine automatically to a non-zero number if it is needed. If the **EXTI** number refines to 0, remove the **EXTI** card from the .ins file.

The first number on the WGHT card should be less than 0.1, and the second number on the WGHT card should be less than 5. If either of the numbers is larger than these values, consult an expert. If the second number is large, then it is possible that your structure is disordered or that you have missed assigning some electron density (especially solvent molecules). For a good discussion of weights, see Dunitz “X-ray Analysis and the Structures of Organic Molecules.”

- 14.5) If your space group is centrosymmetric, skip to step 14.7.
- 14.6) If your space group is chiral or non-centrosymmetric, XL will calculate a parameter known as the Flack parameter and will write its value in the .lst file just after the analysis of the least-squares cycles. If the Flack parameter is close to zero, then the proper absolute configuration has been chosen, and you can skip to step 14.7.

If the Flack parameter is close to 1.0, then the incorrect absolute configuration has been chosen. Invert the structure by adding the card **MOVE 1 1 1 -1** to the .ins file between the **UNIT** and **WGHT** cards. Then re-run XL and re-check the value of the Flack parameter. If it is now close to zero, go to step 14.7. If not, consult an expert.

(NOTE: inverting one of the 11 pairs of enantiomorphous space groups or the space groups *Fdd2*, *I4₁*, *I4₁22*, *I4₁md*, *I4₁cd*, *I-42d*, and *F4₁32* is carried out differently: see Section 8.10 of the Siemens manual for details.)

If the Flack parameter is between 0.3 and 0.7, it is possible that your crystal was twinned. Consult an expert.

- 14.7) If you have completed sections 3 to 14, your crystal structure is now refined (with any luck) as much as it can be. Continue on to the next section.

15. Final details

- 15.1) Near the beginning of the .lst file, there may be a table that compares two things: (a) the number of atoms of each element that are expected in the unit cell, as calculated from the molecular formula and Z, and (b) the number of atoms of each element that are actually present in the unit cell, according to the list of atoms you have found. If this table is absent, skip to the next step. If this table is present, check to see whether there is a difference in the calculated and found molecular formulas. If there is a difference, you should have a reason for omitting some of the atoms from your model (e.g., hydrogen atoms whose locations cannot be determined or calculated).

If you find at this point that you have made a mistake in the formula of your unit cell, you must go back and redo the absorption correction using the correct formula.

- 15.2) In the .lst file just after the summary of the least squares cycles, there will be a table called “Largest correlation matrix elements.” This table shows correlation coefficients

between pairs of parameters. If any correlation coefficient is above 0.8, you should consult an expert. If there are lots of correlation coefficients above 0.9, you have probably chosen a space group that is too low in symmetry (such as refining a centrosymmetric crystal in a non-centrosymmetric space group). Use PLATON (section 11) to look for missed symmetry.

- 15.3) Double check that no “non-positive definite” atoms are present (see step 12.5).
- 15.4) Check to see whether any reflections are statistical outliers. To do this, inspect the $\Delta F^2/\text{esd}$ values for the peaks at the top of the Most Disagreeable Reflections table in the .lst file. If, for one or more reflections at the top of the list, $\Delta F^2/\text{esd}$ is two or more times greater than the other values near the top of the list, the reflections are statistical outliers. Make a note of the Miller indexes hkl of the statistical outliers.

A common cause of statistical outliers is that the reflection was obscured by the beam stop. If this is the reason, the resolution of the reflection will be the largest of all the reflections in the Most Disagreeable Reflections table, and F_o^2 will be much smaller than F_c^2 .

If you have statistical outliers, ask an expert to help you delete these reflections from the .hkl file. Then re-run XL.

- 15.5) In the .lst file, look at the value of the electron density for the largest unassigned (Q) peak in the table of bond distances and angles at the end of the file (refer back to step 10.8; the electron density is the number in the ‘Peak’ column). For an excellent structure, the largest unassigned peak will have an electron density of under $0.5 \text{ e}/\text{\AA}^3$. If any peak is more than $1.5 \text{ e}/\text{\AA}^3$, you should look to see where this peak is. If it is within 1 \AA of a heavy scatterer (e.g. Z larger than 30), the peak is due to inaccuracies in the absorption correction and can be ignored. If it is not close to a heavy scatterer, it may indicate that the structure is disordered. If the largest peak is larger than $4 \text{ e}/\text{\AA}^3$, you have a severe problem (perhaps you missed an atom!). Consult an expert.
- 15.6) A word about the least-squares residual **wR2**. If the final value of **wR2** (for all data) is less than 0.1, that is a low value and means that the observed intensities agree reasonably well with the intensities calculated from your structural model. A low value of **wR2** is no guarantee, however, that you have done everything correctly. There are many examples in which **wR2** is less than 0.1 but the structure has been solved in the wrong space group!

If the final value of **wR2** exceeds 0.2, however, you should consult an expert.

- 15.7) Finally, check the value of the **maximum** shift/esd for the last cycle of least-squares refinement (this number is printed as XL runs and can also be found in the .lst file near where the results of the least-squares cycles are summarized). Your refinement has converged when the **maximum** shift/esd is between 0.01 and -0.01. If convergence is slow, change the **L.S.** card in the .ins file from **L.S. 4** to **L.S. 10**.
- 15.8) If these final details check out, then you will want to prepare tables of the final results. To do this, insert an **ACTA** card into the .ins file anywhere between the **UNIT** and **WGHT** cards. **ACTA** tells XL to prepare several output tables, such as a list of atomic coordinates, a list of displacement parameters, a list of bond distances and angles, and a list of observed and calculated structure factors.
- 15.9) OPTIONAL: You may also add other cards if you wish XL to calculate additional structural parameters (see the Siemens manual for more information about these options):

BOND \$H (Optional). Tells XL to add bond distances and angles involving the hydrogen atoms to the distance and angle tables.

CONF (Optional). Tells XL to generate a table of torsional angles.

MPLA n atoms (Optional). Tells XL to calculate a best least-squares plane through the first n atoms of the list of atoms following the number n. **MPLA** also calculates the distances of each of the named atoms from the least-squares plane. **MPLA** can be used more than once to calculate several different planes, and it generates angles between the previous plane, if more than one plane is defined.

RTAB bond/ang/tors atoms (Optional). Tells XL to calculate non-bonding distances, bond angles, or torsional angles between the atoms listed in the card.

- 15.10) After the ACTA (and any optional) cards have been added to the .ins file, run XL through a final series of least-squares cycles. The tables will be written to the files absfile.cif and absfile.fcf.
- 15.11) The program XCIF is used to transform the data output files (created by inserting the ACTA card into the .ins file) into tables suitable for printing by word processing programs. To run XCIF, click "XCIF" at the top of the SHELXTL window.
- 15.12) Choose **T** from the menu to type the crystal/atom tables from the .cif file. Accept the default name for the .cif file, and type **Y** to select data_absfile. When asked for the new filename, type "absfile.tex", where *absfile* is the name of your absorption-corrected data.
- 15.13) When the program asks for the file to resolve unresolved information, override the [ang] default setting by typing "def". If you get an error message here, and the program asks you for another file name, type "c:\saxi\sxtl\xcif.def".
- 15.14) Answer <return> to the following five or six questions, but answer **N** to the full torsion angles and **N** for the hydrogen bond table.
- 15.15) From the list of options, choose **F** to create a structure factor table from the .fcf file. Accept the default name for the .fcf file. When asked for a filename for the tables, type "absfile.sft", where *absfile* is the name of your absorption-corrected data. Accept the default paper size and number of lines. When you return to the list of options, type **Q** to quit XCIF.
- XCIF writes the tables of crystallographic parameters and structure details, atomic coordinates, anisotropic displacement parameters, and bond lengths and angles to the file absfile.tex. It also writes the structure factor table to the file absfile.sft.
- 15.16) Open the absfile.tex file in Wordpad. Add any missing information such as the lattice type (triclinic, etc.), space group, and dimensions of the crystal (some of this information can be found in the .lst file). The dimensions of the crystal (in mm) can be found in the original.p4p file on the line beginning "CSIZE" (use Wordpad to view this file).
- 15.17) If you used the **MPLA** or **RTAB** card(s), the tables of information generated from these cards should be cut from the .lst file and pasted into a new table called Geometric Details in the .tex file. If you used a **CONF** card, the generated table of torsional angles must be cut from the .lst file and pasted into a new table called Torsional Angles in the .tex file.
- 15.18) Change the font of the absfile.tex file to **Courier 10** and set the top and bottom margins to **0.75 inch**. When you have the tables in their final form, print the file (click File, drag

to Print, and click. In the print window, click Print).

- 15.19) Then use Wordpad to open the absfile.sft file. Delete the line of nonsense characters at the top of the file. Change the font to **Courier 7** and set the left and right margins to **0.5 inch** and the top and bottom margins to **0.75 inch**. Print this file also.
- 15.20) You should write a few paragraphs or so of details about how the diffraction data were collected, solved, and refined. A template document that you can modify is included in Appendix 1.
- 15.21) Once you have finished printing out the crystal data tables, you will want to print out anisotropic displacement (“thermal ellipsoid”) diagrams of the newly-characterized molecule using XP (see sections 16 -18).

16. Generating graphical displays – XP

- 16.1) First, open Paintbrush (or other graphics package, if available), and then minimize its window. In the SHELXTL window, click “XP” to run the XP program.
- 16.2) At the XP command prompt, type “reap filename.res” where *filename* is the name of the most recent .res file.
- 16.3) Type “fmol”, and then “kill \$q”.
- 16.4) If your molecule sits on a symmetry element, use the command “grow” to add the symmetry-related portion of the molecule.
- 16.5) Type “proj”. An image of the structure should come up. You can rotate the structure by highlighting the options at the right of the screen and hitting <enter>. Hitting <enter> again stops the rotation. Hitting <spacebar> will rotate the structure faster; hitting <spacebar> again will make it rotate slower. Rotate the molecule until you get the view you want.
- 16.6) Type “telp 0 -35” to get a thermal ellipsoid view of the molecule. (The 0 instructs XP not to display a stereo view, and the -35 tells XP to draw the a 35% probability density surfaces; the minus sign is necessary to display the thermal ellipsoids.)
- 16.7) Continue to hit <enter> until XP asks for “Plotfile:”, then type in a name (any name is fine), the program then shows a display of your molecule on the screen.
- 16.8) A small box appears on screen near one of the atoms; the box shows where the label for that atom will appear. **The top-left corner of the box is directly on the atom to be labeled.** This can save confusion in a cluttered area of atoms.

It is usually necessary to adjust the position of the box slightly to prevent the label from being printed over an existing atom. Adjust the box with the mouse or with the arrow keys. Hit <space bar> to “apply” a label. Hit <enter> to skip to the next atom without applying a label. If you make a mistake, hit the <backspace> key, and the labels will be erased one at a time. When all labels have been applied, stop with the molecule still displayed on the screen.

If you go too far and accidentally exit telp, then type “view plotfile.plt”, where *plotfile* is the name you entered in step 16.7.

- 16.9) To print a view of your molecule, go to section 17.
- 16.10) Type "e" to exit the telp routine.
- 16.11) To create a view of the unit cell contents, type "matr x" where x is 1, 2, or 3 depending on which axis you want to look down. (1 = a-axis, 2 = b-axis, 3 = c-axis).
- 16.12) Type "pbox x y" where x and y are numbers that describe the width/height and depth of the packing box in Ångstroms, respectively. Normally, these numbers are chosen to fill the unit cell (i.e., to display Z molecules in or just outside the box). Good starting numbers are pbox 15 5.
- 16.13) Type "pack" to pack the box. If the view is satisfactory, go to step 16.14. If the view is unsatisfactory, exit pack, go back to step 16.11, and use different parameters in the pbox command.
- 16.14) Hit SGEN/fmol to exit pack.
- 16.15) Type "telp cell"
- 16.16) Type in a name for the plot file (any name will do)
- 16.17) Hit the enter key four times to label the origin and the three cell axes.
- 16.18) To print a view of the unit cell, go to section 17.
- 16.19) Hit the <escape> key to exit pack.
- 16.20) Type "read filename.res", where filename is the name of the latest .res file.
- 16.21) Type "fmol" followed by "kill \$q".
- 16.22) Repeat steps 16.11 to 16.21 for the other axes.
- 16.23) You are done!

17. Printing diagrams generated by XP

Note: the following section assumes that Paintbrush will be used as your graphics package. This procedure, however, does not give publication-quality images. If you have the capability, use the **draw** command in XP either to print directly to a printer, or to generate an HPGL file that can be read by Coreldraw or other graphics software.

- 17.1) When you are in XP and have a view that you wish to print, move the cursor (if still present) to one of the corners of the screen. Then hold down the <alt> key and press the <print screen> key.
- 17.2) Hold down the <control> key and press <escape>. This will bring up a menu. Use the mouse to open the minimized Paintbrush window.
- 17.3) After Paintbrush opens, hold down the <control> key, and press the "v" key. The picture of your molecule should appear after a few seconds.
- 17.4) Click the cursor once outside of the black region.

- 17.5) Click on Image, then Invert colors (this is important to get a good printout). The image should now be blue on white instead of yellow on black.
- 17.6) Click on Image, then Attributes. In the colors box, click "Black and White"
- 17.7) Click File, drag to Page setup, and click. Click Landscape and then click ok.
- 17.8) Click File, drag to Print Preview, and click. Make sure the structure only covers one page.
- 17.9) Click File, drag to Print, and click. Select Print 1 page and click print.
- 17.10) Minimize the Paintbrush window, and return to step 16.9 or step 16.18.

18. Miscellaneous XP commands

- 18.1) To generate pictures without hydrogen atoms, after the reap command add "fmol less \$H". This will give a new view with no hydrogen atoms present.
- 18.2) It is possible to sort selected atoms in the atom list. Suppose that atom C10 is in the correct place in the list, but atoms C11 and C12, which should follow C10, are instead located elsewhere in the atom list. Type "sort C10 C11 C12" will move atoms C11 and C12 (and their appended hydrogen atoms, if any) to follow atom C10.
- 18.3) To eliminate the bonds between two atoms, use the "undo" command. For example "undo OS1 C12 OS2 C13" eliminates the bonds pairwise so OS1 and C12 will no longer have a bond and OS2 and C13 will not either.
- 18.4) To join to atoms with a bond, use the "join" command. For example "join C1 C2" will join those two atoms with a bond.
- 18.5) To change the labels of all atoms including hydrogen atoms, type "pick/h".
- 18.6) To find out what atoms are attached to a given atom, type "bang atomname", where *atomname* is the atom of interest (e.g., bang P1). This command will give a list of the bond distances and angles of all the atoms bonded to the atom entered. If you want to see relationships between atoms which are not bonded, type "envi atomname" (e.g., envi C1).
- 18.7) To create a "dummy" centroid for a cyclopentadienyl ring (or other group), type "cent/x atomnames", where *atomnames* is a list in cyclic order of the atoms in the ring (e.g. cent/x C1 C2 C3 C4 C5). Draw a new "bond" between the metal and the centroid (which is given the label X1 or X1A) by using the "join" command. The drawing of the molecule will look nicer if you use the "undo" command to erase the bonds between the metal and the atoms that were used to form the centroid.
- 18.8) To change the size of the labels in the graphics, type "labl 0 s" where *s* specifies the size of the label. When *s* = 400, the characters are printed in 8 font.
- 18.9) To view previous .plt files, type "view name.plt", where *name* is what you typed in XP when asked for a plotfile name.
- 18.10) To display a space-filling diagram of the molecule, type "sfil". Label the atoms as usual.

18.11) To get a description of how to use any command, type “help command” where *command* is the name of the command you want explained (e.g. help join).

The Siemens manual can be consulted for other XP commands.

Appendix 1. Sample structure report.

Below is an example of a final report for a crystal structure. If you did anything unusual to solve your crystal structure, include it in your summary even if it is not mentioned below. The references to Table 1 in the summary refer to the first table in the .tex file. For this reason, the following report should be accompany the .tex file tables; in published journal articles, the information in this structure report usually appears in the Experimental Section.

The **shadowed type** indicates sections that you should modify by inserting your results.

Crystallographic Summary for My Molecule

General crystallographic data are listed in Table 1. Systematic absences for ***h0l* ($l \neq 2n$)** were consistent with the space group(s) **Pc** and **P2/c**. **The centrosymmetric space group P2/c was chosen based on the value of $|E^2-1|$.** A **face-indexed (or psi scan)** absorption correction was applied (absorption coefficient $\mu = 1.32 \text{ mm}^{-1}$), and the maximum and minimum transmission factors were **0.9253** and **0.2480**. After the correction, there were 11 systematic absence violations and **14** bad equivalents. Systematically absent reflections were deleted and symmetry equivalent reflections were averaged to yield the set of unique data. **The reflection 0 -2 5 was found to be a statistical outlier ($\Delta F^2/\text{esd} = 14.0$) and was deleted from the data set.** The resulting 4218 data were used in the least squares refinement.

The structure was solved using **direct methods** by using the SHELXTL software package. The correct positions for the **mercury and molybdenum** atoms were deduced from an **E-map (or a sharpened Patterson map)**. Subsequent least-squares refinement and difference Fourier calculations revealed the positions of the remaining non-hydrogen atoms. At this point, a calculation by PLATON showed that there was no missed crystallographic symmetry. Non-hydrogen atoms were refined with independent anisotropic displacement parameters. Hydrogen atoms were **placed in "idealized" positions (or located in the difference Fourier map)** and their displacement parameters were fixed to be **20-50 % larger than those of the attached non-hydrogen atom (or refined independently)**. An isotropic extinction parameter (see the SHELXTL manual for the definition of the EXTI command) was refined to a **final value of 0.0013 (or was not needed)**. The weighting parameters (see the SHELXTL manual for the definition of the WGHT command) were **0.0211** and **3.0421**. Successful convergence was indicated by the maximum shift/error of **0.001** for the last cycle of least squares refinement. The largest peak in the final Fourier difference map (**$0.62 \text{ e}\text{\AA}^{-3}$**) was located **0.71 Å** from the mercury atom.