enCIFer User Guide & Tutorials

EnCIFer User Guide & Tutorials

2022.3 CSD Release

Registered Charity No 800579

Conditions of Use

Cambridge Crystallographic Data Centre

12 Union Road

Cambridge CB2 1EZ

United Kingdom

Web: www.ccdc.cam.ac.uk

Telephone: +44-1223-336408

Email: admin@ccdc.cam.ac.uk

Contacting User Support

If you have any technical or scientific queries concerning this CCDC product, please contact User Support who will try to help.

Email: support@ccdc.cam.ac.uk

Website: www.ccdc.cam.ac.uk/support

Telephone: +44-1223-336022

A list of frequently asked questions (FAQs) is available at the website address given above. This resource is continually being updated with answers to common questions. Please scan the archive for the relevant product before making use of our email and telephone support service.

If you need to contact User Support, please try to provide the following information:

- The name and version number of the product with which you are having problems.
- The make, model and operating system of the workstation you are using.
- A clear description of the problem and the circumstances under which it occurred.

Also be prepared to email error messages, input files and other output. This information is always useful when trying to determine the cause of a problem.

We try to deal with User Support queries within one working day but sometimes problems can take longer to solve. When this happens, we will keep you informed of our progress and try to provide you with an answer as quickly as possible.

If you are a licensed user of the Cambridge Structural Database (CSD) we will try to deal with your User Support query on the working day that it arrives. However, problems can sometimes take longer to solve. When this happens, we will keep you informed of our progress and try to provide you with an answer as quickly as possible.

If you downloaded EnCIFer but are not a user of the CSD itself, we will still try to help but will have to give priority to CSD users.

Activating CSD-Community Access for EnCIFer

Starting the CCDC Software Activation Tool Graphical User Interface

You will need to activate EnCIFer in order to access features available in CSD-Community, but you can do so for free. To activate CSD-Community in a copy of encipher:

- Use the CCDC Activation tool to select the **CSD-Community** tab and then click **Activate**.

The CCDC Activation tool can be opened in the following ways on each supported platform.

Windows

The CCDC Software Activation tools will be located in the `<INSTALLDIR>\CSD_2022\Software Activation` folder of a CSD Release installation, or `<INSTALLDIR>\Software Activation` in other CCDC product installers.

To start the graphical user interface, double-click on **ccdc_activator_gui.exe**

Linux

The CCDC Software Activation tools will be located in the `<INSTALLDIR>\CSD_2022\bin` folder of a CSD Release installation, or `<INSTALLDIR>\Software_Activation`

in other CCDC product installers. To start the graphical user interface, run **ccdc_activator_gui**

MacOS

The CCDC Software Activation tools will be located in the `<INSTALLDIR>\CSD_2022` folder of a CSD Release installation, or `<INSTALLDIR>\Software_Activation` in other CCDC product installers. To start the graphical user interface, double-click the **CCDC-Software-Activation** application.

CIF - The Crystallographic Information File

Introduction to the CIF

The small-molecule Crystallographic Information File (CIF: Hall, Allen & Brown, Acta Crystallographica, **A47**, 655-685, 1991) is a universal format for the electronic storage and exchange of crystallographic information. It has been adopted as the international standard for this purpose by the International Union of Crystallography (IUCr), and is used in:

- Assembling laboratory archives.
- Transferring crystallographic information between laboratories.
- Depositing crystallographic information with most major journals.

- Depositing data with the crystallographic databases.
- As a database output format, e.g. for entries retrieved from the Cambridge Structural Database.

A CIF is an electronic ASCII file, which is intended to be human readable and editable. It consists of a set of data items which may appear individually or in looped lists. The lines in a CIF must not exceed 2048 characters in length, although a soft limit of around 80 characters per line is often used for easy readability and to ensure facile transmission via e-mail.

Full details of the small-molecule CIF, including leading references, can be found on the IUCr website at www.iucr.org, and only a brief general introduction is provided here.

Data Names, Data Values and Data Blocks in a CIF

Every data item is represented by a unique data name followed by its associated data value.

Data items:

- Are described in a dictionary which defines meaning and usage.
- Must start with an underscore (underline) character.
- Can be any type of string (text, numeric or mixed), ranging from a single character to many lines of text.

Data values are delimited by:

- Spaces.
- Double or single quotes, to delimit a data value that contains spaces but does not span lines.
- Pairs of lines beginning with a semicolon, which delimit multi-line data items.

Data values may be set to:

- `?` (unknown).
- `.` (inapplicable).

Related data items, e.g. those which relate to an individual crystal structure are grouped together in a data block:

- The beginning of a block is designated by the string `data_` prefixing the name of the block.
- The end of a block is recognized by another `data_` record, or by an end-of-file mark.

These principles are illustrated in the following examples:

```
data_structure_1
_cell_length_a              5.959(1)
_chemical_formula_moiety    'C23 H36 O7'
_publ_contact_author
;
Dr S. Motherwell
CCDC
12 Union Road
Cambridge CB2 1EZ
England
;
```

Looped Lists of Data Items in a CIF

If a data name is preceded by `loop_` then a series of data values may be associated with it. The following example shows four data values, individual symmetry operators, associated with a single data name:

```
loop_
_symmetry_equiv_pos_as_xyz
'x, y, z'
'-x, y+1/2, -z+1/2'
'-x, -y, -z'
'x, -y-1/2, z-1/2'
```

It is also possible to group a series of data values together under a series of different data names, i.e. to express a table of data values within the CIF, by use of the `loop_` construction. This is illustrated below for the atom labels and x,y,z-coordinates for four atoms, where the data names can be regarded as the column headings for a conventional printed table:

```
loop_
_atom_site_label
_atom_site_fract_x
_atom_site_fract_y
_atom_site_fract_z
I1 0.26639(7) 0.61557(3) 0.94292(3)
I2 0.64548(7) 0.36488(3) 0.56299(3)
P3 0.0438(2) 0.27607(11) 0.74432(13)
O1 -0.0989(7) 0.2488(3) 0.6619(4)
```

The end of a loop is recognized by:

- Any new data name, beginning with an underscore.
- Another `loop_` statement.
- A new block starting with a `data_` statement.
- An end-of-file mark.

The CIF Dictionary

Valid CIF data names and the permitted data value type(s) for each name are expressed in computer-readable dictionaries, where the dictionary syntax is defined in a separate Dictionary Definition Language (DDL). Thus, the dictionary entry for `_cell_length_a` specifies that the data value will be in Angstroms and that a standard uncertainty (ESD) is allowed, placed in parentheses immediately following the value.

EnCIFer is able to load dictionaries which conform to the DLL1.4.1 format (referred to as DDL1) including the small molecule core dictionary and the powder diffraction dictionary. The current DDL1 dictionaries available from the IUCr are included in the EnCIFer distribution with the permission of the IUCr, who hold the

copyright. EnCIFer supports neither DDL2 nor the proposed DDLm, CIF files or dictionaries, e.g. the macromolecular CIF dictionary mmCIF.

The latest small-molecule core and other DDL1 dictionaries and details of the DDL can be found on the IUCr website: www.iucr.org. This site also describes the operation of COMCIFS, the IUCr Committee that oversees CIF development and the approval of additional CIF data names.

Generating and Editing a CIF

The crystallographic part of a CIF is normally generated automatically by the software package used for structure refinement. These CIF generators have been tried and tested over the past decade, and it is seldom necessary to amend any of these data items. Obviously, the software package can only output data items that it knows about, i.e. those items that are input to, or are generated by, the refinement process, or which can reliably be deduced from these data.

For most purposes, it is necessary to edit into the CIF those additional data items that are needed for archival purposes, or for transmission of the CIF to a colleague, journal or database. These data items might include authors' names and addresses, chemical or physical data, such as chemical compound names, the melting point, etc.

Some journals permit the text of the manuscript to be incorporated into the CIF using specific data names. The CIF standard includes support for simple typesetting features such as subscripts, superscripts and consequences of multiple ASCII characters (digraphs or trigraphs) to express non-Roman characters, mathematical symbols etc.

If a CIF is amended or enhanced by manual editing, it is quite easy to destroy the integrity of the CIF format, so that it cannot be parsed and interpreted by the many programs (structure validation software, structure visualizers, etc.) that are now designed to read CIF data. It is because of these pitfalls in manual editing that the EnCIFer program has been written.

Depositing CIFs with Journals and the CCDC

Nearly all journals now expect or require that crystal structure data associated with published papers shall be deposited in CIF format. Authors should consult the Notes for Authors for their chosen journal to see their most recent arrangements for depositing CIF data.

Many major journals have arrangements with the Cambridge Crystallographic Data Centre, whereby the CIF is deposited with the CCDC before the paper is submitted for publication. The CCDC supplies a CCDC Deposition Number and authors include this number in their submitted manuscript. This number is printed in the published paper and is stored in the Cambridge Structural Database, where it affords a valuable link to electronic versions of the journal and allows scientists to request individual datasets free of charge from the CCDC.

The CCDC also accepts CSD Communications (previously called Private Communications to the CSD), i.e. crystal structures for which formal journal publication is not envisaged. The data will be checked and evaluated before inclusion in the CSD, and any problems clarified with the depositor. The author(s) name(s) and contact address will be included in the CSD.

Full details about CIF deposition to the CCDC, for both journal publications and for CSD Communications, can be found on the CCDC website at www.ccdc.cam.ac.uk.

General Features of EnCIFer

Acknowledgements

The CCDC gratefully acknowledges the following copyright works, which are used under licence in EnCIFer:

- Qt, the platform-independent GUI application framework: www.qt.io
- Mesa C 1999-2001 Brian Paul: www.mesa3d.org
- PCRE library package, by Philip Hazel, and copyright by the University of Cambridge, England: www.pcre.org

Scope of the Program

The EnCIFer program operates on single or multi-block CIFs to permit:

- Location and reporting of syntax/format violations using the currently enabled dictionary.
- Correction of these syntax/format violations.
- Editing of existing individual data items or looped data items.
- Addition of new individual data items or looped data items.
- Addition of certain standard additional information via two data entry wizards:
- Publication wizard - basic bibliographic information required by most journals and databases that accept CIF deposition documents.
- Data wizard - chemical and physical property information that enhances a CIF for journal or database deposition.
- Visualization of structure(s) in the CIF.
- Consistency checks on a small number of data fields.

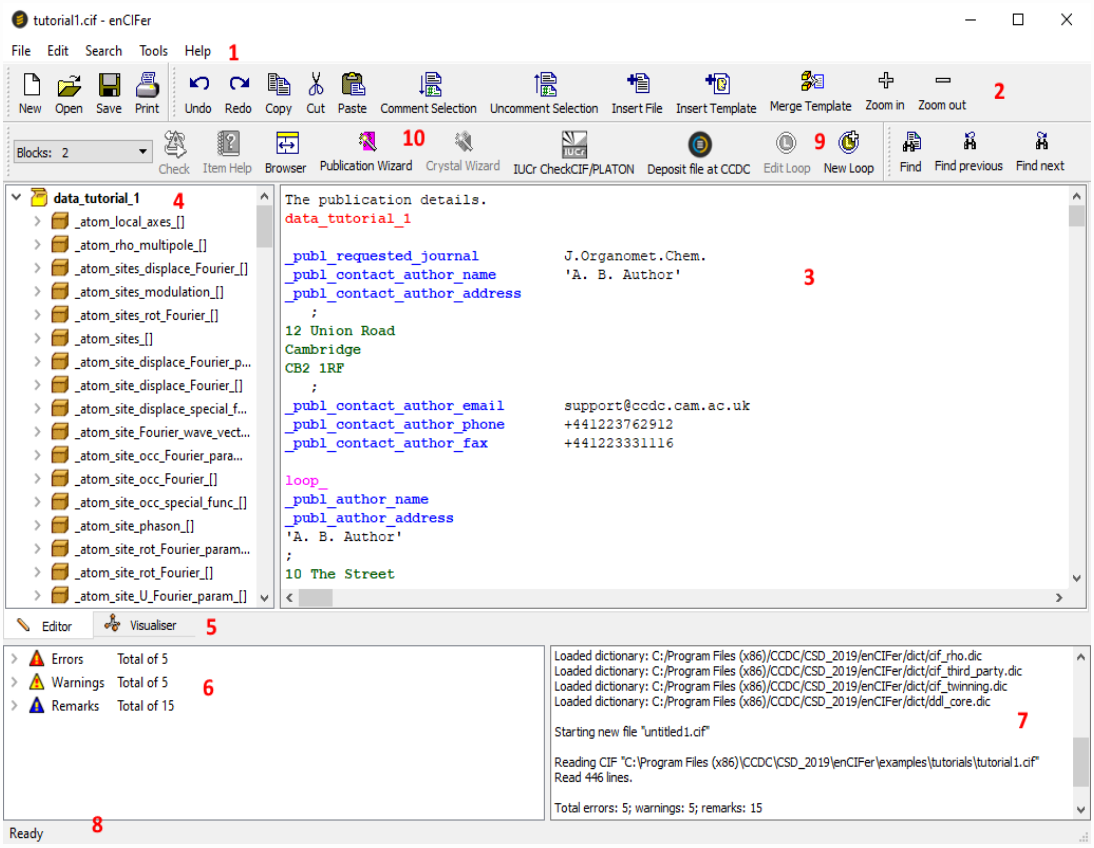
- Additional crystallographic and publication checks via a link to the checkCIF/PLATON service provided by the IUCr.

In all cases where data is edited or added, EnCIFer will check the format integrity of the amended file.

Thus, EnCIFer is designed to ensure that CIFs are syntactically correct before they are archived, transmitted or used as input to other software, e.g. data validation programs or structure visualizers. The visualization and other checking features also aid users in validating the content and semantics of the CIFs.

Overview of the Interface

The EnCIFer interface consists of the following:

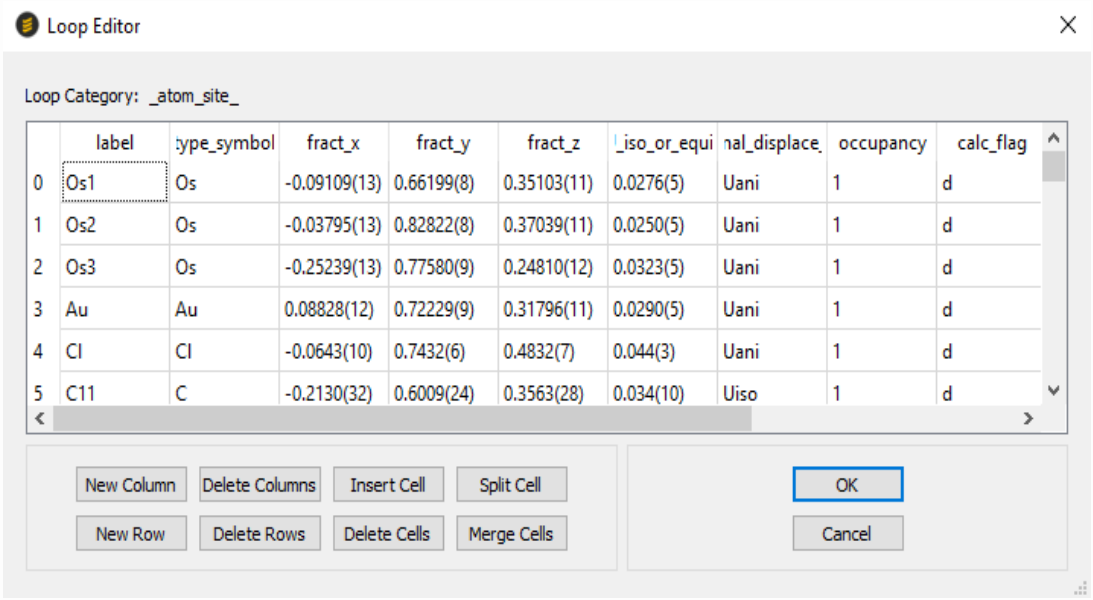


1. Top-level menu - comprising **File**, **Edit**, **Search**, **Tools** and **Help** menus.

2. Tool bar - containing many common program options.

3. Text editor - for CIFs with syntax highlighting.

4. Browser - allowing CIF dictionary navigation.
5. Structure Visualiser - for displaying the crystal structure(s) in CIFs.
6. Error list view - for displaying and navigating error, warning and remark messages.
7. Scrolling log - message window.
8. Status bar - displaying help messages and line/column numbers.
9. Loop editor - spreadsheet view of loop constructs
10. Wizards - for entering crystal, chemical and publication data:



Command Line Options

EnCIFer recognizes the following command line options:

```
encifer -style=<style>

encifer -directrendering

encifer -nodirectrendering

encifer <filename>
```

- `-style=<windows/motif/cde/plastique/cleanlooks>`, which changes the user interface style.
- `-directrendering` and `-nodirectrendering`: which enables or disables hardware acceleration. The default is `-directrendering` on all platforms. If problems are encountered with EnCIFer's 3D display, try both values.
- `<filename>`, which opens the named file. Multiple files may be specified by separating the names with spaces.

- `-font <fontname>` — LINUX only — defines the application font.

Note: If EnCIFer is run for a second or subsequent time whilst it is already running, then new files will be added to the existing instance, rather than opening a new instance. In this case, any non-filename command line arguments will be ignored. Also, if EnCIFer is set to reuse a single window then only the first file specified will be opened.

File Operations

Opening a CIF

When EnCIFer is launched, by default the **Editor** pane shows an empty CIF, named `untitled1.cif`.

An existing CIF can be opened in a number of ways:

- Via the **Open** dialog box which is obtained either by selecting **File** > **Open**, by clicking the **Open** icon on the toolbar, or by using the **Ctrl+O** keyboard shortcut. Choose the file type (CIF `*.cif` or all files) you want and then choose the file you wish to open.
- When typing the `encifer` command (typically on UNIX), you can supply a file name on the command line, and EnCIFer will open the file when it starts.
- If you are using a computer that supports Drag and Drop, you can drag a file icon and drop it onto the EnCIFer program icon (or a shortcut to it). EnCIFer will launch and open the dropped file. Alternatively, you can drop a file icon onto the EnCIFer editor window, and the dropped file will be opened.
- A list of the most recently opened files is maintained by EnCIFer. These may be reopened by clicking **File** in the top-level menu, selecting **Recent Files** and then choosing the filename from the resulting pull-down list. A warning will result if the file is no longer found on the system.

- On Windows, the installer can associate files ending with `.cif` with EnCIFer, so that opening such a file will automatically load the file into EnCIFer.
- On macOS, EnCIFer will associate itself with `.cif` files, so that such files can be automatically opened in EnCIFer, or dragged and dropped onto an EnCIFer icon in the tray.

The CIF is read as a plain text file using the character set of the native system (e.g. ISO-8859-1). The end-of-line sequences for Windows (<CR><LF>), macOS (<CR>) and UNIX/Linux (<LF>) are all interpreted as indicating a new line irrespective of the native system. As a result, all versions of EnCIFer can read plain text files generated on Windows, macOS and UNIX/Linux systems.

Single Window and Multiple Window Operation

By default, each CIF is opened in a new EnCIFer window. Alternatively, EnCIFer may be configured to reuse the current window when opening files. To configure this behaviour:

1. Hit **Edit** on the top-level menu and select **Preferences**. In the **Preferences** dialog box, click the **General** tab.
2. Click the radio buttons **Use multiple windows** or **Reuse single window** to set the desired behaviour.
3. Click the **OK** button to apply changes, **Cancel** to discard changes or **Restore Defaults** to restore the default setting.

Creating a New CIF

A new empty CIF `untitledn.cif` may be opened by:

- Clicking **File** on the top-level menu and selecting **New**.
- Clicking on the **New** icon on the toolbar.
- Using the **Ctrl+N** keyboard shortcut.

This will open a new window if the **Use multiple windows** preference is set. If the **Reuse single window** preference is set, then any of the above actions will replace the current window contents with an empty file.

CIF Template File

EnCIFer may be configured to open a template CIF rather than showing an empty file when starting up and when **New** or **New Window** is selected. This may be useful when publishing CIFs in order to avoid editing the same values into multiple files.

This template file may also be inserted into the editor at the current cursor position either by:

- Hitting **Edit** in the top-level menu and selecting **Insert Template**.
- Clicking the **Insert Template** icon on the toolbar.

To set the template file:

- Hit **Edit** on the top-level menu and select **Preferences**.
- In the **Preferences** dialog box, select the **Editor** tab, click the **Browse** button and select the template file in the **Open** dialog box.
- Check the **Always start with template** box for the template to be automatically used whenever a new file is created, instead of an empty file.

Merge CIF Template

Data items in the CIF template may be merged into the current CIF block, with the option to preserve or overwrite existing data values. This may be useful in order to set standard data item values in an automatically generated `.cif` e.g. publication details or diffractometer parameters.

The template file may be merged into the block at the current cursor position either by:

- Hitting **Edit** in the top-level menu and selecting **Merge Template**.
- Clicking the **Merge Template** icon on the toolbar:

In the resulting dialog, choose the block(s) in the template CIF which you wish to merge into the current block.

If the **Overwrite data items in current block** checkbox is unchecked, data items will only be copied from the template file if the data items are not present in the current block. Looped data items will only be copied if none of the data items in the loop are set in the current block.

If the **Overwrite data items in current block** checkbox is checked, data items from the template file will replace any data items, including looped data items, already present in the current block.

Insert File

Text may be inserted into the current **Editor** pane from a CIF or plain text file using the **Insert File** option in the top-level **Edit** menu or by clicking the **Insert File** icon on the toolbar.

In the resulting **File** dialog box, choose the file type (CIF `*.cif` or all files) you want and then select the file you wish to open. The contents of the file are inserted into the editor at the current cursor position.

If the text contains characters which are not part of the CIF character set, they will be converted to CIF digraph or trigraph sequences where possible. For example, accented characters in the ISO-8859-1 (Latin 1) character set are converted to the corresponding CIF character sequences.

Close File

CIFs may be closed either by:

- Hitting **File** on the top-level menu and selecting **Close File**.
- Using the **Ctrl+W** keyboard shortcut.

If there are other EnCIFer windows open from the same session, the current window is closed. Otherwise, if there is only one EnCIFer window, a new `untitledn.cif` file is opened in this window.

If the current file has been modified, a dialog box will prompt whether to save changes to the current file. Click **Yes** to save changes, **No** to discard changes or **Cancel** to abort the operation.

Saving Files

To save the current CIF:

- Hit **File** in the top-level menu and **Save** in the next menu.
- Click the **Save** icon on the toolbar.
- Use the **Ctrl+S** keyboard shortcut.
- If the current editor contains a new CIF, the **Save As** dialog box opens which allows you to specify the name and location of the file to be created.

To save an automatic backup copy of the current CIF each time the CIF is saved, select the **Save backup CIFs** option in the **Preferences** dialog box:

1. Hit **Edit** in the top-level menu and select **Preferences**.
2. In the **Preferences** dialog box, click the **General** tab.
3. Check the **Save backup CIFs** box to enable this feature.

The backup of CIF `filename.ext` will be saved as `filename_backup.ext`.

To save the current CIF with another filename:

1. Hit **File** in the top-level menu, **Save As** in the next menu.
2. This opens the **Save As** dialog box which allows you to specify the name and location of the CIF to be created or replaced.
3. CIFs are saved as plain text with the end-of-line termination characters appropriate to the native system, i.e. <CR><LF> for Windows, <CR> for macOS and <LF> for UNIX/Linux, irrespective of the end-of-line sequence(s) present in the file when it was opened in EnCIFer.

Printing Files

The CIF in the current editor may be printed by:

- Hitting **File** in the top-level menu and selecting **Print**.
- Clicking the **Print** icon on the toolbar.
- By using the **Ctrl+P** keyboard shortcut.

In each case, a **Print** dialog box appears which enables the printer to be set up. The options available in this dialog box depend on the operating system. If you have a colour printer, you can print the CIF in colour to show the syntax highlighting as displayed in the **Editor** pane.

Text Editor

CIFs are edited in a plain-text **Editor** window. If the **Visualiser** is displayed, click on the **Editor** tab to return to the **Text Editor**.

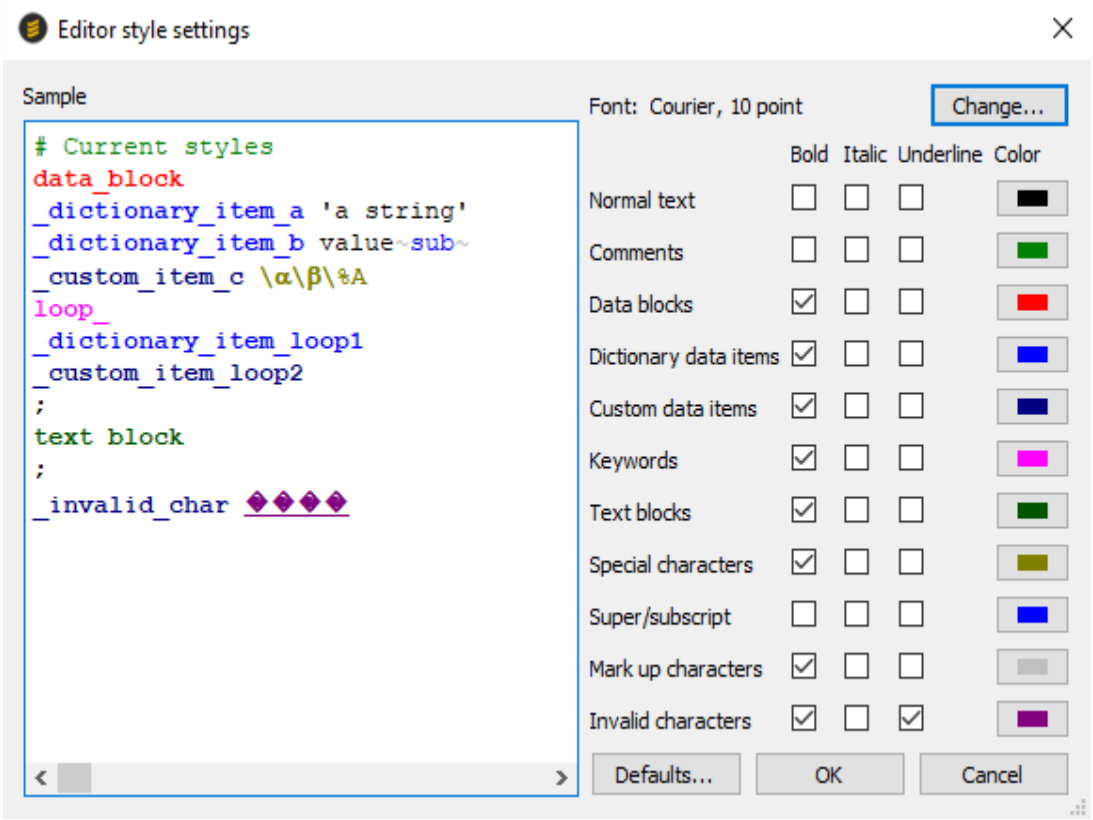
Syntax Highlighting

Text is highlighted according to the CIF syntax. This helps to give a visual indication of potential errors. The default scheme is as follows:

- Data block header: bold red text.
- Loop keyword: bold magenta text.
- Data names in any of the currently enabled dictionaries: bold blue text.
- Data names not in any of the currently enabled dictionaries: bold dark blue text.
- Comments: italic green text.
- Multi-line text blocks (semicolon delimited): bold dark green text.
- Special characters and mark-up sequences: bold dark yellow text.
- Superscripts/subscripts: blue text.
- Superscript/subscript markers: bold grey text.
- Other (Normal) text: black.
- Invalid CIF characters (e.g. non-breaking space characters): bold underlined purple text.

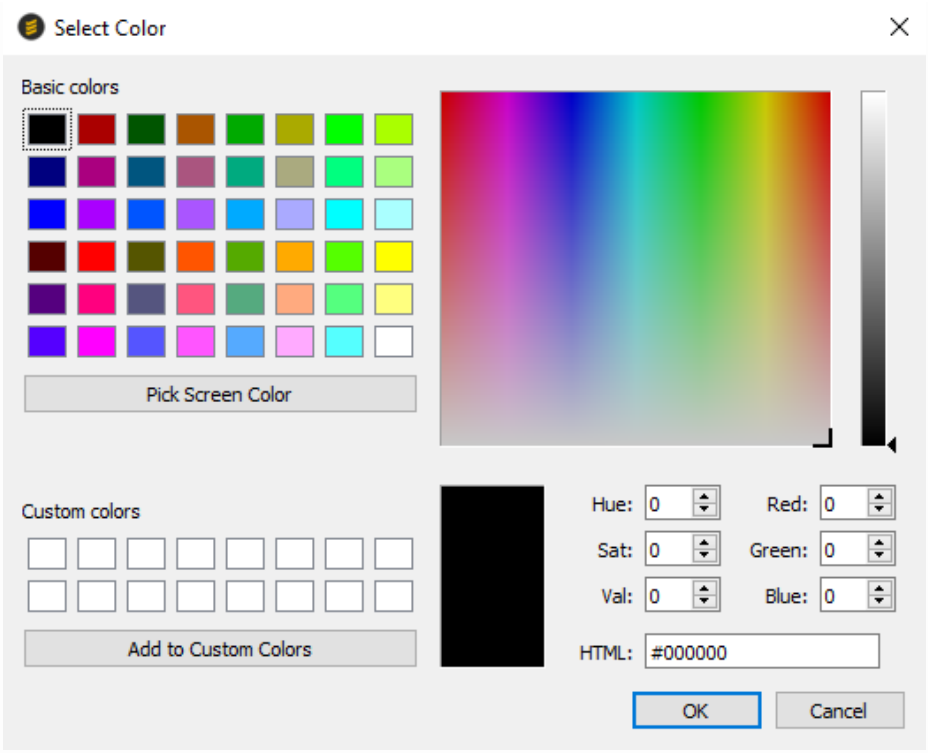
To change the style settings:

1. Hit **Edit** on the top-level menu and select **Preferences**.
2. In the resulting dialog box, click the **Editor** tab and then select the button alongside click **to change style settings**.
3. The **Editor style settings** dialog box should now appear. This shows the current style and color settings along with a sample of each type of text highlighting in a pane on the left.



At the top, the font name and size used for all styles (except special characters) is displayed, special characters will use a 'Symbol' font of the same size. The **Change** button to the right can be used to set this font. When adjusting only the size and name will be used, any bold, italic or other style settings will be ignored.

Next to each style are checkboxes to set **bold**, italic and underline styles. There is also a button displaying the currently used color, pressing this brings up a **Color selection** dialog:



To restore the default settings:

- Click on the **Defaults** button. You will be asked to choose between a **Variable Width** or **Fixed Width** font. Variable Width will use the default application font for your system; Fixed Width will use a 'Courier' font of the same size. It may be more convenient to use a fixed-width font to maintain column alignment, especially in large loops.

As settings are adjusted, the text in the **Sample** pane is immediately updated to reflect the style changes. Text in the editor windows will only be updated when the preferences dialogs are closed by selecting **OK** in both the **Editor style settings** and **Preferences** dialog boxes. Selecting **Cancel** will discard any changes.

Scaling the Editor Text

To enlarge the text in the editor:

- Click **the Zoom-in (+)** icon on the toolbar.

To reduce the size of the text in the editor, click the **Zoom-out (-)** icon on the toolbar.

These actions are a shorthand way of opening the **Preferences** dialog and adjusting the font size.

Data Item Dictionary Help

To show the CIF dictionary information for a data item, place the cursor on the data name in the **Text Editor**, and either:

- Right-click on the data name and select **Item help** from the resulting menu.
- Use the **Ctrl+I** keyboard shortcut.
- Click **Item Help** on the toolbar:

The data item name and the dictionary in which the data name is defined are shown in the title box, and the scrolling window provides the definition, recommended values (enumeration list) and examples taken from the dictionary.

Editing Text

Text may be typed into the **Editor** pane in the same way as for a regular plain text editor. Due to the arbitrary width nature of the CIF format, there is no line wrapping and hard line breaks must be entered as appropriate. For an explanation of the different line limits (see Soft Line-Length Limit).

Double-clicking in the text selects the word (text delimited by white space) at the current cursor position. To select all text in the editor, hit **Edit** on the top-level menu and click **Select All**, or use the **Ctrl+A** keyboard shortcut.

The status bar at the bottom of the EnCIFer window reports the line and column number at the current cursor position, and whether the file has been modified since it was last saved.

CIFs must be composed in a limited character set. Other special characters (e.g. Greek letters and mathematical symbols) must be input according to the CIF markup conventions. The special character markup sequences may be shown by hitting **Help** on the top-level menu and selecting **Special characters**. Click the **Close** button to dismiss the **Special characters** dialog box.

Copy, Cut and Paste

To copy selected text to the clipboard, either:

- Hit **Edit** on the top-level menu and select **Copy**.
- Click the **Copy** icon on the toolbar.
- Use the **Ctrl+C** keyboard shortcut.

To cut selected text to the clipboard, either:

- Hit **Edit** on the top-level menu and select **Cut**.
- Click the **Cut** icon on the toolbar.
- Use the **Ctrl+X** keyboard shortcut.

To paste text from the clipboard, either:

- Hit **Edit** on the top-level menu and select **Paste**.
- Click the **Paste** icon on the toolbar.
- Use the **Ctrl+V** keyboard shortcut.

EnCIFer uses the system clipboard for copy, cut and paste operations, so text may be cut and pasted to and from other applications, e.g. word processors or text editors. Plain text clipboard content will be used. It may be necessary to insert hard line breaks in the pasted text in EnCIFer.

If the text contains characters which are not part of the CIF character set, they will be converted to CIF digraph or trigraph sequences where possible. For example, accented characters in the ISO-8859-1 (Latin 1) character set are converted to the corresponding CIF character sequences.

Cutting and pasting text in EnCIFer containing such characters provides an automatic way of converting the special characters to CIF characters sequences where possible. The sequence **Select all** > **Copy** > **Paste** can be used to force all special characters in a CIF to be adjusted.

Undo and Redo

The last editor operation may be reversed by either:

- Hitting **Edit** on the top-level menu and then selecting **Undo**.
- Clicking the **Undo** icon on the toolbar.
- Clicking the **Undo** icon on the toolbar.
- Using the **Ctrl+Z** keyboard shortcut.

The last editor operation to be undone may be redone by either:

- Hitting **Edit** on the top-level menu and selecting **Redo**.
- Clicking the **Redo** icon on the toolbar.
- Using the **Ctrl+Y** keyboard shortcut.

Certain operations may require two undo or two redo steps.

Find and Replace

Find

To find text in the editor, either:

- Hit **Search** on the top-level menu and select **Find**.
- Click the **Find** icon on the toolbar.
- Use the **Ctrl+F** keyboard shortcut.

A **Find Text** dialog box should appear:

1. Type the required text into the **Find** edit box. Tick any of the options **Whole words only**, **Case sensitive**, **Start at Beginning**, and select the search direction (**Forward** or **Backward**) by clicking the appropriate radio button. Then click the **Find** button to find the text.
2. If the text is found, the cursor will move to the text which will be highlighted.
3. To repeat the find operation, click the **Find** button again whilst the dialog box is still displayed.
4. To dismiss the **Find Text** dialog box, click the **Close** button.

If the **Find** dialog is reopened the previous search text will be displayed. Other previously searched for items are available via a drop-down list.

Find Next and Find Previous

To find the next match either:

- Hit **Search** on the top-level menu and select **Find next**.
- Click the **Find next** icon on the toolbar.
- Use the **F3** keyboard shortcut.

To find the previous match either:

- Hit **Search** on the top-level menu and select **Find previous**.
- Click the **Find previous** icon on the toolbar.
- Use the **Ctrl+F3** keyboard shortcut.

Replace

To replace text in the editor either:

- Hit **Search** on the top-level menu and select **Replace**.
- Use the **Ctrl+R** keyboard shortcut.

A **Replace Text** dialog box should appear:

1. Type the text to be replaced in the **Find** edit box and the replacement text into the **Replace** edit box. Tick any of the options **Whole words only**, **Case sensitive**, **Start at Beginning**, and select the search direction (**Forward** or **Backward**) by clicking the appropriate radio button.
2. Click **Find Next** to find the next match, **Replace** to replace the text for the currently highlighted match, or **Replace All** to replace all occurrences of the text.
3. Click the **Close** button to dismiss the dialog box.

As with the **Find** dialog, if the **Replace** dialog is reopened the previous search and replace texts will be displayed. Other previously searched for and replacement texts are available via a drop-down list.

Line and Block Navigation

To move to a particular line either:

- Hit **Search** from the top-level menu and select **Go to line**.
- Use the **Ctrl+G** keyboard shortcut.
- In the resulting dialog box, type the desired line number and click on the **Go** button.

To move to the start of a particular block, select the required block from the **Blocks** pull-down menu on the toolbar.

To show line numbers alongside the **Editor**, select the **Show line numbers** option in the **Preferences** dialog box:

1. Hit **Edit** on the top-level menu and select **Preferences**.
2. In the **Preferences** dialog box, click the **Editor** tab.
3. Check the **Show line numbers** box to enable this feature.

Comment and Uncomment

In CIF, comments are begun by a # character and continue until the end of the line.

Lines can be quickly commented (a # prepended) out by selecting **Comment Selection** from the **Edit** menu or by using the **Toolbar** button.

Lines can be uncommented (a single # removed) by selecting **Uncomment Selection** from the **Edit** menu or by using the **Toolbar** button.

When no text is selected these commands will affect only the current line. If a block of text is selected, then all selected lines will be adjusted. Only a single # will be prepended or removed for each use of these commands, so if a selection contains a mixture of commented and uncommented lines the differences will be preserved.

Browser

The browser provides a hierarchical view of the CIF in terms of the data blocks and their data items, the hierarchy being defined by the arrangement of data categories, sub-categories and data items in the currently enabled dictionaries. It provides a means of navigating the blocks and data items in the CIF and allows additional data names to be inserted into the editor text.

Enabling the Browser

To toggle the display of the browser:

- Hit **Tools** in the top-level menu and select **Browser**.
- Click the **Browser** icon on the toolbar:

- Use the **Ctrl+B** keyboard shortcut.

To adjust the width of the browser relative to the editor, drag the vertical splitter bar (the <-||-> cursor should appear).

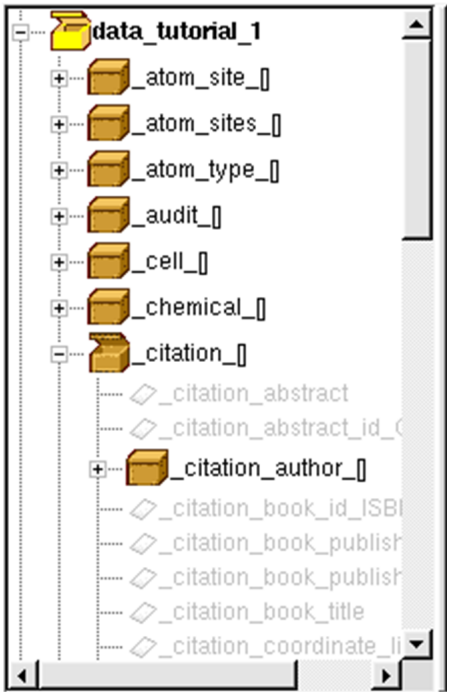
Expanding and Contracting the Browser

When the **Browser** is first displayed, it shows the CIF data blocks as **Closed yellow box** icons. Double-clicking on the **Box** icon, or clicking the **[+]** or **[>]** icon, expands the view to show the data item categories in black text with closed orange boxes, the data block now shows an **Open yellow box** icon.

Expanding these categories shows the data names within that category, the category boxes change to show open box icons. The data names may not be expanded further and are displayed differently depending on their status:

- Data names which are present in the current block are shown in black text and have item icons.
- Data names which are not set are shown in grey and have grey item icons.
- Items which are looped appear in black text with a double item icon.
- Items which have errors or warnings related to their value appear in black with a warning icon.

To contract the display for a category or data block, double-click on the **Open box** icon or click the **[-]** icon alongside:



Note: If the browser becomes difficult to navigate because of too many dictionary items, it is possible to disable unused dictionaries (see CIF Dictionary Options).

Browser Navigation

Block Navigation

To move to a particular **Block** in the editor, either:

- Right-click on the corresponding data block name and select **Go to block** from the resulting menu.
- Select the data block name from the pull-down **Blocks** list on the toolbar.

Both the **Browser** and **Text Editor** will scroll to ensure that the data block names are displayed in the viewing areas and are highlighted.

Item Navigation

Double-clicking on a data name which has been set (shown in black), or right-clicking and selecting **Go to item** will move the cursor to the data name in the editor and highlight the matching text. This is not possible for data items which have not been set in the current block.

Inserting and Copying Data Names

Right-clicking on a data name which has not been set (shown in grey) and selecting **Paste into editor** inserts the data name into the editor at the current cursor position. This is not possible for data items which are already set in the current block.

A data name may be copied to the clipboard by right-clicking on the data name and selecting **Copy item name** from the resulting menu.

Data Item Dictionary Help

CIF dictionary help is available both for data names and data categories by right-clicking on the text in the browser and selecting **Item help** from the resulting menu.

Setting Data Items

Data items may be set or reset by:

1. Right-clicking on the data name text in the browser and selecting **Set data item value** from the resulting menu. The resulting dialog box shows the dictionary data item help and allows the data value to be input.

2. If the data item has an enumeration list in the dictionary, the enumerated values are shown in a pull-down list, in addition to `?` (unknown) and `.` (not applicable). Select the appropriate data value from the list.

3. Otherwise, an **Text** edit box is shown. Type the data value in the box. No quotes are required, these are added automatically by EnCIFer.

4. Click the **OK** button to apply the changes. The CIF will be re-parsed automatically and the new data highlighted. Else, click the **Cancel** button to abort the change.

Syntax and Dictionary Checking

Error, Warning and Remark Display Windows

When a CIF is opened in EnCIFer, it is parsed to check for CIF syntax and CIF dictionary compliance. Optionally, it may be checked for the presence of mandatory data items and data consistency.

Error, warning, and remark messages are displayed in an expanding list view on the lower left, and a summary is written to a scrolling log at the lower right of the window.

The relative widths of the two windows may be altered by clicking and dragging the vertical splitter between the windows (the <-||-> cursor should appear).

The scrolling log may be cleared by right-clicking in the **Log** window and selecting **Clear** from the resulting pop-up menu.

Updating Error, Warning and Remark Messages

After the CIF has been edited, the file may be re-parsed to re-check for errors and warnings by:

- Hitting **Tools** on the top-level menu and selecting **Check**.
- Clicking the **Re-check** icon in the toolbar.

- Using the **Ctrl+K** keyboard shortcut.

Any errors or warnings are added to the scrolling log. The expanding list view is updated to show the current error and warning messages.

Navigating Error, Warning and Remark Messages

To expand the list view:

- Click the **Expand** icon alongside the red errors, yellow warnings, or blue remarks triangles.

Double-clicking on an error, warning or remark message in the list view moves the cursor to the corresponding line in the CIF, provided that the CIF has not been edited so as to change the line numbering since the CIF was last parsed. The corresponding line is then highlighted in yellow. For a few errors, such as those for duplicate data item names, two lines will be highlighted, one in yellow and the other in green.

Right-clicking on the error, warning or remark message shows the documentation help for the message. This includes advice on common causes of the error, warning or remark and how to correct them.

Configuring the Number of Error and Warning Messages

The number of error and warning messages displayed can be configured in the **Preferences** dialog box:

1. Hit **Edit** in the top-level menu, select **Preferences**, and click the **CIF checking** tab.
2. The **Maximum consecutive error lines** setting controls how many consecutive error lines may be encountered before error checking is abandoned. It also controls the number of

consecutive warning lines which are shown, further warnings being omitted until after a line without warnings is encountered.

3. To change the setting, type in the **Text** edit box or use the arrows to increment or decrement the number.

4. Click the **OK** button to apply a change, **Defaults** to return to the default setting or **Cancel** to abandon changes.

CIF Dictionary Options

EnCIFer loads DDL1.4.1 dictionary files with the extension .dic in the dict directory of the distribution on startup (see The CIF Dictionary). The **Output** pane reports which dictionaries have been loaded and any errors in reading dictionaries.

Additional DDL1.4.1 dictionaries with the extension .dic may be added to this directory if you have write permission, or the dictionaries can be updated with more recent versions.

By default, all dictionaries which are loaded successfully are enabled and used to check the CIF data items and values. Individual dictionaries may be disabled in the **Preferences** dialog box:

1. Hit **Edit** in the top-level menu, select **Preferences** and click the **Dictionaries** tab.

2. The available dictionaries are listed in a spreadsheet, with associated version numbers and dates.

3. Click the appropriate checkboxes in the spreadsheet to enable or disable particular dictionaries.

Soft Line-Length Limit

The current CIF 1.1 specification permits lines up to 2048 characters in length, whereas CIF 1.0 only permitted lines 80 characters or less. EnCIFer will give an error if a line exceeds a hard limit of 2048

characters or a warning if a line exceeds a soft limit which can be set in the range 72-2048 characters (default 80) in the **Preferences** dialog box:

1. Hit **Edit** in the top-level menu, select **Preferences** and click the **CIF checking** tab.
2. To set the desired limit, type in the **Text** edit box or use the arrows to increment or decrement the number.

Mandatory Data Items

Overview of Mandatory Data Items

EnCIFer can check for the presence of data items, listed in configurable files, required by journal publishers or crystallographic databases (see Format of the Mandatory Data Items File and Setting the Mandatory Data Items File).

A distinction is made between blocks comprising publication information, (`_journal`, `_publ`) chemical and crystallographic data, (`_atom_site`, `_atom_sites`, `_cell`, `_chemical`, `_diffrn`, `_exptl`, `_geom`, `_refine`, `_symmetry`) or both. A remark message is given if a data name is not present in an appropriate data block or if the data value is `?` (unknown). No remark is given if the data value is `.` (inapplicable). If there is no block in the CIF of a particular type, remarks are given for the first block.

If a data item has an alternate or replace relation to another data item in the CIF dictionary, the related data item is also checked. If a related item is present in the data block with a value other than `?` (unknown) no remark is generated, e.g. no remark will be given for: `_refine_ls_R_factor_gt` if the alternate term `_refine_ls_R_factor_obs` is set.

Format of the Mandatory Data Items File

This file is in plain text CIF format, with up to three blocks:

- `data_publication` - items required in publication data blocks.

- `data_crystal` - items required in crystal data blocks.
- `data_all` - items required in all blocks.

The required data items are listed in each text block (the data item value is ignored), e.g.

```
# Data required in publication information blocks
data_publication
_publ_contact_author_name ?
_publ_contact_author_address ?

# Data required in crystal data blocks
data_crystal
_cell_length_a
...
_symmetry_equiv_pos_as_xyz
```

Setting the Mandatory Data Items File

In order to set the mandatory data items in the file:

1. Hit **Edit** on the top-level menu and select **Preferences**.
2. In the **Preferences** dialog box, click the **CIF checking** tab, hit the **Browse** button and select the template file in the **Open** dialog box.
3. Check the **Check for Mandatory Data Items** box to enable checks using the contents of this file.
4. A file, `recommended.cif`, is provided with the EnCIFer distribution. This contains a list of items that are commonly required by chemistry journals which publish the results of crystal structure determinations, and for deposition at the CCDC. This file may be customized to suit the requirements of a particular journal.
5. Some additional mandatory data items files are provided which are already tailored to the requirements of particular journal publishers e.g., `mandatory_rsc.cif`. Please refer to the comments in these files, detailing the conditions of use, and to the notes for authors for the particular journal.

6. As these files are themselves in CIF format, they can also be opened and edited using EnCIFer.

Data Item Consistency Checks

EnCIFer performs some consistency checks between related data item values. Failures of these checks appear as remarks or warnings, as appropriate. By default, the checks are enabled, to adjust this setting:

1. Hit **Edit** on the top-level menu and select **Preferences**.
2. In the **Preferences** dialog box, click the **CIF checking** tab, check or uncheck the **Perform data consistency checks** box.

Details follow on each of the checks made.

If any of the data items mentioned in the checks are not present or have `?` or `.` values, the check is not performed, and no messages generated.

Symmetry Operator-Space group Name Consistency

This checks that if a space group name is provided in a `_symmetry_space_group_name_H-M` item it is consistent with symmetry operators provided in `_symmetry_equiv_pos_as_xyz`. Related data item names will be used as appropriate.

Atom Label Uniqueness

Looped `_atom_site_label` values are checked to ensure they are unique.

Formula Weight Consistency

A `_chemical_formula_weight` value is checked to be within 1.1 Daltons of the weight calculated from the value of `_chemical_formula_sum`.

Summary of Error Messages

These usually indicate serious errors in the CIF syntax. The error messages include:

- Data item <data name> has already been set in this data_ block.
- No data items in loop_.
- Too many or too few data values in the loop.
- Data name <data name> immediately followed by another data name.
- Data name <data name> more than 75 characters long.
- Data name <data name> not followed by data value.
- CIF contains more than one data block called <data block>.
- CIF contains no data_ blocks.
- No terminating (‘) quote.
- No terminating (“) quote.
- Text block finished at end of file without final ‘;’ .
- Text block finished at end of file without final ‘;’ .
- (Line) More than 2048 characters long.
- Invalid non-printable character <octal value>.
- Invalid printable character <octal value>.
- Closing (;) semicolon not followed by white space.

Data item <data name> has already been set in this data_ block

This error occurs if the same data name appears more than once in the same data block. This can happen if a data block header has been omitted or incorrectly formatted, such that two blocks are interpreted as one.

No data items in loop_

This error occurs if the `loop_` keyword is present, but it is not followed by any data item names. This can happen if the `loop_` keyword is followed immediately by data values or another CIF keyword.

Too many or too few data values in the loop

This error indicates that the number of data values in the loop is not an exact multiple of the number of data item names given in the loop. Typical causes of these problems include:

- No data values in the loop, just data names. Every data name must have at least one data value in the loop.
- Spacing is omitted between values in the loop, e.g. `. . .` should be `. . .`
- Quotes are omitted around data values such that they are interpreted as multiple values rather than a single value, e.g. `J. A. Smith` should be `'J. A. Smith'`.
- Text after the loop is interpreted as data values, e.g. because a data name has been omitted, a comment (#) character omitted or a keyword has been misspelt.

Data name <data name> immediately followed by another data name

This error occurs when there are two consecutive data names. The first data name should be followed by a data value.

Data name <data name> more than 75 characters long

Data names may be no more than 75 characters in length.

Data name <data name> not followed by data value

This error occurs when a data name is followed by another keyword (e.g. `loop_` or `data_`). The data name should be followed by a data value before the keyword.

CIF contains more than one data block called <data block>

This error occurs when more than one data blocks have the same name in the data block header. All data blocks must be given a unique name.

CIF contains no data_ blocks

This error occurs if there are no data block headers (`data_<name>`).

No terminating (“) quote

These errors occur if a data value has an opening but not a closing single or double quote on the same line, or if there is a spurious opening quote (e.g. `‘psi-scan` should be `psi-scan`, `‘psi-scan’` or `“psi-scan”`). This can occur if a data value delimited by single or double quotes has been split over two lines. This can sometimes happen if email or other software wraps text lines at less than 80 characters.

Either the text should be moved to the same line, or be treated as a multiple line value and delimited with semicolons.

Opening (;) semicolon not first character on line & Terminating (;) semicolon not first character on line

These errors occur if semicolons delimiting a text string are the first non-space characters on a line but are not the first character as required. Any spaces before the semicolons should be deleted.

Text block finished at end of file without final ';'

This error indicates that a closing semicolon has been omitted, causing the remainder of the CIF to be interpreted as a runaway multiple line data value, or there is an additional spurious semicolon in the CIF which has the same effect. It is important to ensure that all semicolons occur in pairs around data values. The syntax highlighting in the editor should help to show the cause of this problem.

(Line) More than 2048 characters long

The CIF 1.1 specification requires lines to be no more than 2048 characters in length. Line breaks should be inserted to ensure all lines are 2048 characters or less in length. Where necessary, single or double-quoted strings should be converted to semicolon-delimited multiple-line strings to conform to the line length limit.

Invalid non-printable character <octal value> & Invalid printable character <octal value>

These messages indicate that the CIF contains characters which are not included in the allowed character set. The characters have their own syntax highlighting style so that they can be easily seen on the lines containing these errors.

Non-printable characters are usually control characters, whereas printable characters may be accented characters or special symbols. Printable characters should be converted to the CIF digraph or trigraph sequences.

Select the **Special characters** dialog box from the top-level **Help** menu. If the printable character displays correctly in the **Editor**, it is usually possible to do the conversion automatically by cutting and pasting the text containing the invalid characters. Non-printable characters should be converted to spaces or be deleted altogether.

Closing (;) semicolon not followed by white space

When a semicolon delimited text block is closed, it must be followed by at least one white space character. This white space character will often be a new line, but spaces are sometimes used within loops.

This error is most often found within a loop where all values are quoted, for example author names and addresses. It can also be found where values are immediately followed by comments. For example, the following will generate this error as there is no space between `;` and `#`.

```
_publ_contact_author_address
;
12 Union Road
Cambridge
CB2 1RF
;# comment
```

Summary of Warning Messages

These include warnings:

- For less serious CIF syntax errors.
- Where data names are not found in the currently enabled CIF dictionaries.
- Where data values do not conform to the values permitted for the corresponding data names in the CIF dictionaries.

CIF Syntax Warning Messages

This section contains a list of all commonly occurring CIF syntax warning messages:

- Missing identifier for data_ block.
- Data item found before the first data_ block.
- Ignored string <string>.

- Ignored string (possible semi-colon mismatch).
- Ignoring uncommented string(s) before first data block.
- (Line) More than <soft line length limit> characters long.
- Repeated keyword <keyword>.
- Data value contains mismatched subscript (or superscript) markup.
- End of file marker <marker> is non-standard.

Missing identifier for data_ block

This warning indicates that the block name has been omitted from the data block header. Each data block should have a unique identifier appended to the `data_` keyword, e.g. `data_compound_1`.

Data item found before the first data_ block

This warning occurs if data names are encountered before the first `data_` keyword. A data block, consisting of the keyword `data_` plus some unique identifier e.g. `data_compound_1` should appear before any data items.

Ignored string <string>

This warning occurs if text is encountered which is not interpreted as a data value, comment or keyword. This may be due to a number of common problems:

- A data value containing spaces is not quoted correctly. All single-line data values containing spaces must be delimited with single or double quotes, otherwise text after the first space is not interpreted as part of the data value, e.g. `rotating anode generator` should be `'rotating anode generator'`.
- A data value delimited by single or double quotes has been split over two lines (this can sometimes happen if email or other software wraps text at less than 80 characters). Either the text should be moved to the same line, or treated as a multiple line value and delimited with semicolons, e.g.

```
'a long
string'
```

should be:

```
'a long string'
```

or:

```
;
a long
string
;
```

- The opening semicolon of a multiple-line data value has been omitted, such that text after the closing semicolon is interpreted as another data value, e.g.

```
chemical_name_systematic
A compound name
;
```

should be:

```
chemical_name_systematic
;
A compound name
;
```

The syntax highlighting in the editor should help to show this.

- Multiple data values are associated with a data item name, but the `loop_` keyword has been omitted. Without the `loop_` keyword, a data name may only be associated with one data value, e.g.

```
_publ_author_name 'J. Soap' 'J. Bloggs'
```

should be:

```
loop_

 _publ_author_name 'J. Soap' 'J. Bloggs'
```

- A comment has not been preceded by a hash (#) character, e.g.

```
_publ_contact_author_name ‘J. Soap’ Name of contact author
```

should be:

```
_publ_contact_author_name ‘J. Soap’ # Name of contact author
```

Ignored string (possible semi-colon mismatch)

This warning usually occurs if a semicolon closing a multiple-line data value has been omitted, or if an extra semicolon is present at the start of a line somewhere in the CIF. It is important to ensure that all semicolons occur in pairs around data values as the first character on a line. The syntax highlighting in the editor should help to show the cause of these problems.

Ignoring uncommented string(s) before first data block

This warning indicates that there is text before the first data block header (`data_` keyword) in the CIF which is not formatted correctly as CIF. This may comprise email headers, textual comments or similar. All such text should be removed from the CIF or be commented out by placing hash (#) characters at the start of each line.

(Line) More than <soft line length limit> characters long

The line is longer than the current soft line length limit (default 80) (see Soft Line-Length Limit). Line breaks should be inserted to ensure all lines are shorter than the current soft line length limit. Where necessary, single or double-quoted strings should be converted to semicolon-delimited multiple-line strings to conform to the line length limit.

Repeated keyword <keyword>

This warning most commonly occurs when the `loop_` keyword occurs on consecutive lines. The duplicate `loop_` keyword should be deleted.

Data value contains mismatched subscript (or superscript) markup

Superscript text is indicated by a pair of ^ markers, for example x^2 is written x^2^ and subscript text is indicated by a pair of ~ markers, for example H_2O is written H~2~O.

Superscript and subscript markers must be correctly paired and arranged to return to normal text by the end of each data value, although they may span multiple lines within a single semicolon delimited text string. Superscripted and subscripted text, and the markers, are colored differently to normal text; this may help in seeing where the problem occurs.

~ and ^ may also appear when defining accented characters, these uses are ignored for this check.

End of file marker <marker> is non-standard

Some early CIF files were terminated with an uncommented sequence beginning `_eof`. This sequence is not valid according to the version 1.1 specification. These sequences should be commented or deleted.

CIF Dictionary Compliance Warning Messages

This section contains a list of all commonly occurring CIF dictionary compliance warning messages:

- Created non-standard data item <data name>.
- Data value is not in dictionary enumeration list.
- Data value should be a number.
- Data value is not a correctly formatted number.
- Data value should be greater than a defined limit.
- Data value should be smaller than a defined limit.
- Data value should not have a standard uncertainty (ESD).

Created non-standard data item <data name>

This warning occurs if the data name cannot be found in a currently enabled dictionary. This may indicate a number of possible problems:

- A data name has been spelt incorrectly.
- A custom data item has been created.
- A data name from a dictionary which is not currently enabled has been used (see CIF Dictionary Options).
- The CIF dictionary has been updated since this version of EnCIFer was released. The IUCr website can be checked for new versions to download (see CIF Dictionary Options).

Data value is not in dictionary enumeration list

Some data items have an enumeration list of recommended values in the dictionary which are included in the data item dictionary help (see Data Item Dictionary Help).

This warning indicates that the data value given does not correspond to any of the recommended values.

Data value should be a number

Some data items should have a numeric value (integral or floating-point without units, but with a standard uncertainty in some cases).

Data value is not a correctly formatted number

This warning indicates that the data value does not conform to the CIF definition of a numeric data value, i.e. integral or floating-point number with a standard uncertainty in some cases (see Summary of Warning Messages). This may be explained by one of the following problems:

- There should be no spaces between the value and the standard uncertainty.

- The standard uncertainty should be surrounded by opening and closing parentheses, e.g. 12.934(3).
- A range of values, e.g. 212–215, is not permitted.
- Units or percentage signs are not permitted in the dictionary.

Data value should be greater than a defined limit

Some data items which should have a numeric value have a recommended numeric range. If the data item is allowed to have a standard uncertainty (ESD) the range is extended by +/–3u where u is the standard uncertainty given in the data value.

These warnings indicate that the data value does not fall within the recommended range, which are included in the data item dictionary help (see Data Item Dictionary Help).

Data value should not have a standard uncertainty (ESD)

Some data items should not have a standard uncertainty associated with them. This warning indicates that a standard uncertainty (ESD) has been given in a numeric data value for such an item, i.e. a number in parentheses after the numeric value.

CCDC Data Consistency Check Warning Messages

These messages relate to CCDC Checks (see Setting the Mandatory Data Items File).

If the relevant data items are not present or have values ? or . the related check will not be performed. If you wish this fact to be reported, use the mandatory data items feature (see Overview of Mandatory Data Items).

This section contains the following messages:

- Error parsing_chemical_formula_sum ‘value’.
- Ignoring invalid symmetry operator text <text>.
- Duplicate atom label ‘<label>’.

Error parsing_chemical_formula_sum 'value'

This indicates the check of formula weight failed as the sum formula value could not be parsed.

Ignoring invalid symmetry operator text <text>

This indicates that whilst attempting to perform space group consistency checks an unreadable symmetry operator was encountered. This can occur if comments in a symmetry operator loop are erroneously wrapped when sending a CIF via email.

Check that each value of data items `_space_group_symop_operation_xyz` or `_symmetry_equiv_pos_as_xyz` contains x, y and z values, separated by commas (,).

Duplicate atom label <label>

This is a failure of the duplicate atom label check as atom labels should be unique. It means that there are duplicate values for the `_atom_site_label` data item, which will usually appear in a loop.

Summary of Remarks Messages

Mandatory Data Item Remarks Messages

These messages relate to the checking for mandatory data items (see Mandatory Data Items):

- Data item <data item> not found in block <block>.
- Data item <data item> not set in block <block>.
- Related item <related item> not set in block <block>.

Data item <data item> not found in block <block>

This indicates that the data item, and any related data item, is not found in the block.

Data item <data item> not set in block <block>

This indicates that the data item is found but has an unknown value in the block, and no related item is found with a value other than unknown, e.g. mandatory data items file contains: `_refine_ls_R_factor_gt` and CIF contains either: `_refine_ls_R_factor_gt ?`, `_refine_ls_R_factor_gt ?` or `_refine_ls_R_factor_obs ?`

Related item <related item> not set in block <block>

This indicates that the data item is not found in the data block and the related item has an unknown value, e.g. mandatory data items file contains: `_refine_ls_R_factor_gt` and CIF contains: `_refine_ls_R_factor_obs ?`

CCDC Data Consistency Check Remarks Messages

These relate to CCDC Checks (see Setting the Mandatory Data Items File).

If the relevant data items are not present or have values `?` or `.` the related check will not be performed. If you wish this fact to be reported, use the mandatory data items feature (see Overview of Mandatory Data Items).

Messages are given as remarks when items individually fulfill the dictionary requirements but are inconsistent with respect to each other:

- Ignoring supplied space group <name>; using space group generated from symmetry operators <name>.
- Ignoring supplied international tables number <number>; using space group generated from symmetry operators <name>.
- Formula weight <weight> different to weight calculated from chemical formula <weight>.

Ignoring supplied space group <name>; using space group generated from symmetry operators <name>

This is a failure of the space group consistency check, which means that the text in `_symmetry_space_group_name_H-M` does not match the name of the space group generated from the operators in `_space_group_symop_operation_xyz` or `_symmetry_equiv_pos_as_xyz`.

When loading the structure in the **Visualiser**, the space group specified by the operators will always be preferred.

Ignoring supplied international tables number <number>; using space group generated from symmetry operators <name>

This is a failure of the space group consistency check, which means that the value in `_space_group_IT_number` or `_symmetry_int_tables_number` does not match the number of the space group generated from the operators in `_space_group_symop_operation_xyz` or `_symmetry_equiv_pos_as_xyz`.

When loading the structure in the **Visualiser**, the space group specified by the operators will always be preferred.

Formula weight <weight> different to weight calculated from chemical formula <weight>

This is a failure of the formula weight consistency check, which means that the value of the `_chemical_formula_weight` data item does not match a weight calculated from the value of the `_chemical_formula_sum` data item.

The values may differ by up to 1.1 Daltons without this message being generated.

Loop Editor

Overview of Loops

Loops may be constructed or modified in a loop editor spreadsheet as an alternative to using the **Text Editor** pane. The **Text Editor** is disabled whilst the **Loop Editor** is in use.

It is strongly recommended that syntax errors are corrected before using the **Loop Editor**. Otherwise, if there are CIF syntax errors which affect the loop, some data may be lost if edits made in the **Loop Editor** are applied to the CIF.

The **Loop Editor** preserves comments associated with the data names. Any comments interspersed between the data values in the loop will be moved to between the last data name and the first data value.

Creating a New Loop

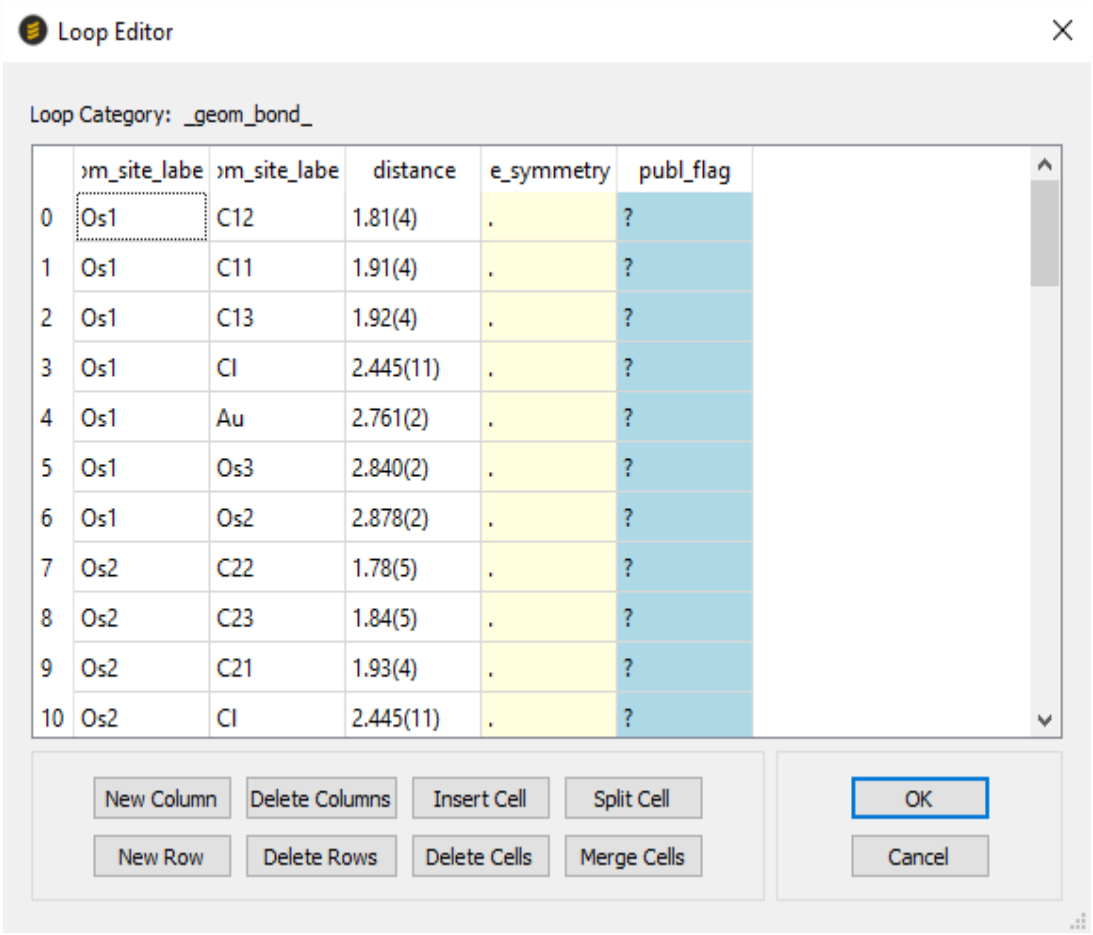
A new loop may be inserted at the current cursor position in the **Editor** pane by hitting **Tools** on the top-level menu and selecting **New Loop** or by selecting the **(L)+** icon from the toolbar. The new loop initially has no rows or columns.

A new loop may not be inserted on lines comprising a semicolon-delimited multi-line string, including those lines with the opening and closing semicolons, and the icon is disabled in this case.

Editing an Existing loop

An existing loop may be opened in the **Loop Editor** by placing the cursor on a line containing the `loop_` keyword in the **Editor** pane, which enables the **(L)** toolbar icon. Selecting this icon, or hitting **Tools** on the top-level menu and selecting **Edit Loop**, starts the **Loop Editor**, which provides a spreadsheet view of the loop data.

The CIF will be automatically re-parsed if necessary before the spreadsheet view is shown:



Loop Display

The common prefix of all the data names in the loop is displayed next to the **Loop Category** label, the remaining text of each data name are shown as column headings in the spreadsheet. Rows are numbered sequentially. Spreadsheet cells are colour-coded according to data content: grey for empty cells, yellow for cells containing . and blue for cells containing ?.

Cells that give CIF dictionary errors (i.e. where the data value is incompatible with the CIF dictionary definition of the data item) show a yellow warning triangle.

Data item help may be obtained by right-clicking on a spreadsheet cell. This shows a dialog box containing the dictionary definition for the data item in the column, provided that the data name is present in a currently enabled dictionary.

Editing Cells

Cells may be edited by first double-clicking in the spreadsheet cell and then editing the contents, or by highlighting a cell and typing. Pressing **Return** or **Enter** applies changes to the spreadsheet cell and revalidates the data value.

Applying and Discarding Changes

Changes made in the **Loop Editor** are applied to the CIF in the **Text Editor** by clicking on the **OK** button. This button is deactivated whilst a cell is being edited.

Changes made in the **Loop Editor** may be discarded by clicking the **Cancel** button.

Changes applied to the **Text Editor** may be undone by:

- Hitting **Edit** in the top-level menu and selecting **Undo**.
- Using the **Ctrl+Z** keyboard shortcut.

This is done in one step for a new loop or two steps for an existing loop (the first undo step removes the modified loop text, and the second undo restores the original loop text).

Resizing Columns and Rows

Columns may be resized by left-clicking and dragging the right-hand boundary of the column header (a <-||-> cursor should appear) to change the column width. Alternatively, double-clicking whilst the <-||-> cursor is shown resizes the column to accommodate the longest text in the column, including the data item names in the column header, when these are not highlighted in bold.

Rows may be resized by left-clicking and dragging the lower boundary of the row header (a <-||-> cursor should appear) to change the row height. Alternatively, double-clicking whilst the <-||-> cursor is shown adjusts the row height to accommodate the current text in that row.

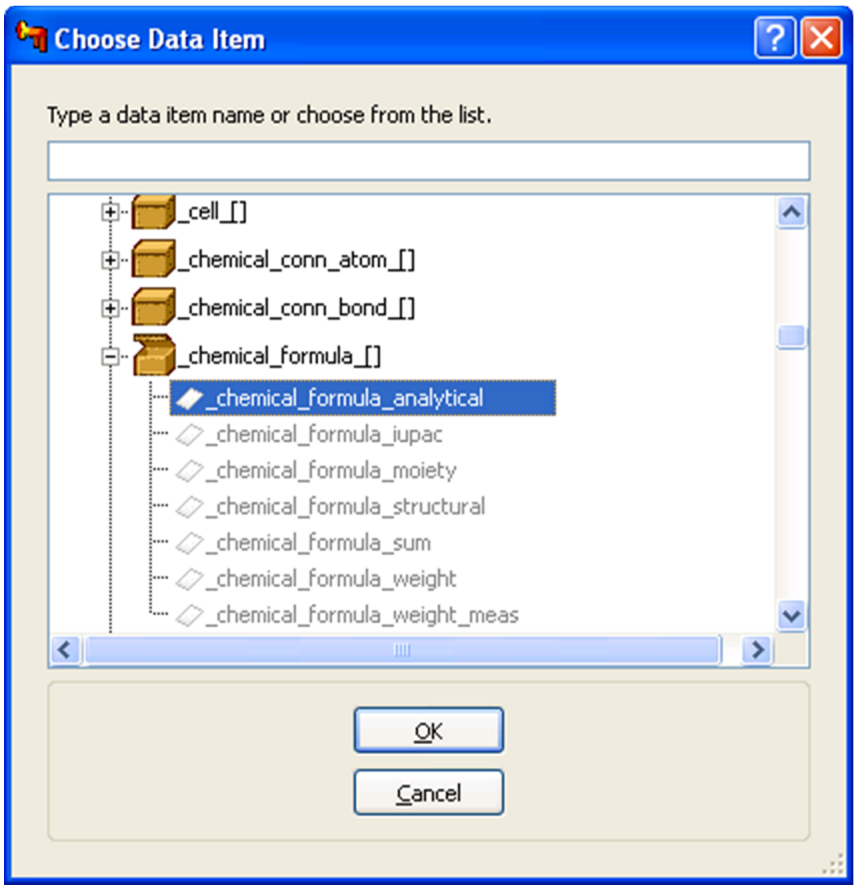
Moving Columns and Rows

Individual columns may be moved by **Ctrl**+**Left-Clicking** a column header (a <--> cursor should appear) and dragging the cursor to a column boundary. The column will be moved to a position to the left of the column boundary.

Individual rows may be moved by **Ctrl**+**Left-Clicking** a numbered row header (a <--> cursor should appear) and dragging the cursor to a row boundary. The row will be moved to a position below the row boundary, and the rows will be renumbered to preserve the sequential numbering, i.e. rows do not retain their original numbering when moved.

Adding New Columns, Rows and Cells

Adding Columns



A column may be added by clicking on the **New Column** button to show the **Choose Data Item** dialog box. A data name must be provided in order to create a new column for that data item. The data name may be:

- Typed into the **Text** edit box.
- Entered into the **Text** edit box by double-clicking a data item in the currently enabled dictionary from the `data_dictionary` expandable list view.
- Entered into the **Text** edit box by right-clicking on the data item in the `data_dictionary` list view and selecting **Paste into editor** from the resulting menu.

Initially, the list view shows all data item categories. The list view may be expanded by clicking the **[+]** boxes to show data item names. The browser expansion is retained between successive new column operations, to facilitate construction of loops with related data items.

Data item dictionary help is available for data categories and data items in the browser by right-clicking and selecting **Item help** from the resulting menu:

- Clicking the **OK** button adds a column to the loop spreadsheet at the right of the spreadsheet, provided that a data item name has been entered in the **Text** edit box.
- Clicking the **Cancel** button aborts the process of adding a new column.

All cells in the new column are given initial values of . so that the existing cell alignment is unchanged.

Adding Rows

A new row may be added to the spreadsheet, after the last row, by clicking the **New Row** button. Rows cannot be added to a spreadsheet which has no columns (e.g. when creating a new loop).

Cells in the new row are not given values and unless given values will not be written out when the **Loop Editor** exits.

Adding Cells

An individual cell may be added before the current cell with the **Insert Cell** button. This provides a means of fixing loop alignment where spacing has been omitted between data values in the loop.

Selecting Columns, Rows and Cells

Columns are selected by left-clicking column headers displaying the data item names. Multiple contiguous column selections are made by **Shift+Left-Clicking** a range of column headers.

Rows or groups of rows are selected by left-clicking or **Shift+Left-Clicking** on the numbered row header sections.

Cells are selected by left-clicking. Contiguous blocks of cells may be selected by **Shift+Left-Clicking** and individual cells added to the selection by **Ctrl+Right-Clicking**.

Blocks of cells can also be selected by holding the left mouse button and dragging.

Deleting Columns, Rows and Cells

Selected columns may be deleted by clicking the **Delete Columns** button. If all columns are deleted, all the rows are also deleted as it is not possible to have a table with columns but no rows in the **Loop Editor**.

Selected rows may be deleted by clicking the **Delete Rows** button. The remaining rows are renumbered to preserve sequential numbering. If all rows are removed, the column headers are retained.

A selected range of cells may be deleted with the **Delete Cells** button (the cells need not be contiguous). The remaining cells are moved to the left, and up to the end of the row above where necessary, to occupy spaces vacated by the empty cells. Incomplete rows remain incomplete, and the deleted cells will not be written out when the **Loop Editor** exits. The number of values in the loop will need to be corrected or syntax errors will result. Any empty rows which result when cells are removed are also deleted.

Merging Cells

Adjacent cells spanning any number of rows may be merged into a single cell. This can be useful when correcting loops where spaces have been mistakenly added, or a value containing a space has not been quoted. To merge cells:

- Select the cells you wish to merge (see Selecting Columns, Rows and Cells) and click the **Merge Cells** button. All but one of the

selected cells will be deleted, the remaining cell will contain the concatenated text of all the merged cells. The remaining cells will move as described above in **Deleting Columns**.

Splitting Cells

The text in a single cell may be divided between the existing cell and a new cell. This can be useful when correcting loops where spaces have been mistakenly missed out between values.

To split a cell:

1. Double-click in the cell to obtain the editing cursor, as in **Editing Cells** above.
2. Use the cursor keys or click again to place the cursor at the position in the text you would like the cell to be split and click **Split Cell.**
3. A new cell will be inserted adjacent to the cell being edited with other cells being moved as in **Inserting Cells** above. Text to the left of the cursor will be placed in the left cell, text to the right of the cursor will be placed in the right cell.

Data Wizards

Overview of Data Wizards

Two wizards are provided in EnCIFer to assist with entering both publication details and additional chemical and crystallographic data into a CIF data block, with syntactically correct formatting. These wizards prompt for data values which are frequently required for publication, but which are not always generated automatically by crystal structure solution software.

The publication data wizard and crystal data wizard both operate on the current CIF block, as determined by the current editor cursor position. They are activated by either:

- Clicking the **Publication Wizard** or **Crystal Wizard** icons respectively on the toolbar.
- Selecting **Tools** on the top-level menu and clicking **Publication Wizard** or **Crystal Wizard**.

The **Text Editor** may not be used whilst a wizard is in operation.

The publication data wizard may be used to create a new data block by placing the cursor before the first block, or be applied to any existing block.

The crystal data wizard may only be used on a data block which already contains crystal data (`_chemical`, `_cell` or `_atom_site` data items).

Structure of the Wizards

Each wizard comprises a number of pages:

- To move to the next wizard page, click the **Next** button.
- To return to the previous page in the wizard, click the **Back** button.
- To abort the wizard from any page, discarding any changes already input, click the **Cancel** button.

The first page provides introductory information. The second page shows any CIF error and warning messages and the other pages of each wizard allow input of a number of CIF data items:

- Where the data item is already set in the CIF block, the data value is shown in the wizard, even where it is set to `.` or `?`.
- If data items are left unset (empty) in the wizard, the corresponding data name will be omitted from the data block if the changes made in the wizard are applied to the editor.

- Data item dictionary help is available for some data items by clicking on the appropriate icon on the wizard pages.

The final page of each wizard allows data entered or modified in the wizard to be applied to the CIF block in the **Text Editor** by clicking **Finish**. Previous pages may be reviewed by clicking **Back**, and the edits may be discarded by clicking **Cancel**.

Changes applied to the **Text Editor** may be undone, by hitting **Edit** in the top-level menu and selecting **Undo,** or by using the **Ctrl+Z** shortcut. This will be one step for a new publication block or two steps for an existing block (the first undo step removes the modified data block text, and the second undo restores the original data block text).

Using the Wizard when there are CIF Errors or Warnings

Before a wizard can be used to enter data the CIF must be free from errors, to prevent serious data loss from occurring if changes made in the wizard are applied to the editor.

When starting a wizard, the CIF will be re-parsed if necessary, and any errors or warnings are shown on the second page of the wizard. If there are fatal errors, wizard operation cannot continue, and the errors must be corrected first. If there are warnings, the wizard can be used but data loss may occur for some warnings, in particular Ignored string warnings.

Publication Wizard Outline

Page 1. Introduction

Introduction to the wizard.

Page 2. Errors and Warnings

Summary of current error, warning and remarks messages.

Page 3. Contact Information

This page allows entry of contact author details. If `_publ_contact_author` is present in the CIF data block, the wizard will interpret the first line as the contact name and the remainder as the contact address, unless the `_publ_contact_author_name` is set in which case it is all interpreted as the contact address.

It is also possible to enter these details in the preferences, such that they are automatically entered into the wizard if the CIF block has no contact author data items set.

To set the contact author preference:

1. Hit **Edit** on the main menu > **Preferences** > **Wizard**.
2. The contact author details may then be entered in the edit boxes.
3. Click the **OK** button to apply the changes, or **Cancel** to discard changes.

Page 4. Publication Information

This page asks whether the structure is being submitted to a journal for publication, is being submitted to a database as a CSD Communication, or has already been published (the year, page and/or volume are known). Clicking on the appropriate radio button determines which pages are subsequently displayed.

Page 5. Requested Journal

A list of possible abbreviated journal titles is given in a scrolling list view. These are journals from which crystal structure data has previously been abstracted by the CCDC.

Double-clicking on a journal title enters it into the **Text** edit box. Alternatively, a journal title may be typed into the **Text** edit box. The scrolling list is updated to show only those journals which match what has been typed, ignoring case and punctuation. The journal manuscript code as assigned by the journal publisher may be entered in the **Text** edit box if this is already known.

Page 6. Journal Details

A scrolling list of journals is given as for the Requested journal name in page 5 of the wizard. Additionally, page numbers, volume and year may be entered.

Page 7. Author Details

Only one author in the CIF loop is displayed at any time. Initially, the first author in the loop is displayed. Family and first author name and the address corresponding to the currently displayed author are shown in separate edit windows. The values may be edited by typing in the edit windows.

If there are no author details in the CIF, the contact author set on Page 3 of the wizard is inserted automatically as the first author.

Clicking the **Previous** and **Next** buttons allows the previous and next author in the loop to be displayed in turn.

To add authors:

1. Click the **Add** button.
2. Type in the author's details in the **Edit** window.
3. The new author is added to the end of the loop. If the previous author in the list has an address set, this address is copied for the new author.

The current author may be deleted by clicking the **Delete** button.

Page 8. Finish

Click the **Finish** button to apply the changes to the CIF in the editor, or **Back** to review changes.

Crystal Data Wizard Outline

Page 1. Introduction

Introduction to the wizard.

Page 2. Errors and Warnings

Summary of current error, warning and remark messages.

Page 3. Physical and Chemical Information

Contains:

- Systematic name.
- Common name.
- Moiety formula.
- Sum formula.
- Compound source.
- Physical Properties. A list of common properties is provided in a drop-down list. These can be appended to the text by selecting an item then pressing the **Add** button. Free text can also be entered in the box to the right.

Page 4. Physical and Chemical Information

Contains:

- Crystallization solvent.
- Melting point.

- Crystal habit.
- Crystal colour.
- Diffraction temperature.
- Diffraction pressure.

Page 5. Symmetry Information

The crystal system may be set from a pull-down list containing the allowed CIF values. The wizard tries to ensure consistency between the space group number in International Tables for Crystallography (IUCr/Kluwer, 1993) and the Hermann-Mauguin (H-M) symbol:

- The wizard shows an editable pull-down list of common space group settings corresponding to the International Tables number.
- Other H-M symbols may be entered by clicking and typing in the **H-M symbol** edit box, pressing **Return** to complete the edit.
- The International Tables space group number will be updated if the H-M symbol is recognized.
- Setting the International Tables number updates the list of possible space groups.

EnCIFer tries to ensure consistency between the Hermann-Mauguin symbols and the symmetry equivalent positions:

- A warning is given if they do not agree, suggesting an H-M symbol derived from the equivalent positions where possible.
- If the H-M symbol is not given in the CIF, EnCIFer tries to derive it from the symmetry equivalent positions.
- Absolute Configuration. Allowed values can be selected from a drop-down list. Press the **Item help** icon for a full description of the values.

Page 6. Diffraction Information

Details of:

- Radiation probe. Selection of the probe is made from the drop-down list.
- Radiation type.
- Radiation wavelength.
- Radiation source class.
- Radiation source type.

Page 7. Finish

Click the **Finish** button to apply the changes to the CIF in the editor, or **Back** to review changes.

IUCr CheckCIF/PLATON

There is a link between EnCIFer and the CheckCIF/PLATON service provided by the IUCr: www.iucr.org.

The CheckCIF/PLATON service provides an additional level of checks of CIF data to those present in EnCIFer. Performing these checks is recommended before submission of a CIF to the CCDC. The checks work by uploading your CIF to an IUCr web page which then displays the results. The details of these checks are constantly being updated so please refer to the relevant web pages for details.

To use the link, either select the CheckCIF/PLATON item on the Tools menu, or click the appropriate toolbar icon.

- It is recommended to use EnCIFer to correct all error, warning and remarks messages before performing these additional checks.

- If the current CIF has been edited, then you will be first asked if you wish to save your edits. This is done because only the saved CIF data is sent for checking, not the current state of the **Text Editor**.
- You will then receive a prompt that the current CIF file name has been placed on the clipboard, click **OK**, or type **Esc** or **Return** to clear this prompt.

The appropriate IUCr web page should then be displayed in your default web browser. Click on the **File name** box and then either select **Paste** from the **Edit** menu, or use the **Ctrl+V** key combination (**Command key+V** on macOS). Then press the **Send CIF for checking** button. The report should now be displayed.

For web security reasons it is not possible for EnCIFer to automatically insert in the correct file name, it has to be pasted in by the user.

Visualizer

Visualizer Overview

Crystal structure visualizer windows may be displayed by clicking on the **Visualiser** tab under the **Text Editor** window. These are regenerated if the CIF has been re-checked since the visualizer windows were last shown, to take into account any changes to the crystal data blocks.

By default, a 2x3 grid of visualizer windows is shown, with one window for each data block comprising crystal structure data. The caption under each visualizer window gives the corresponding data block name.

If there are more than six crystal data blocks, you may page through the blocks by clicking on the **Back <-** and **Forward ->** icons. If there is only one crystal structure block, a single visualizer window is shown. And if there are no crystal structure blocks, an empty grid of visualizer windows is shown.

To show a single visualizer window for one block:

1. Click the **Zoom-in (+) magnifier** icon, or select the desired data block name from the **Blocks** pull-down toolbar list.
2. Page through the crystal data blocks by clicking on the **Back <-** and **Forward ->** icons.
3. The block displayed is changed by selecting the desired data block name from the **Blocks** pull-down toolbar list.
4. To return to the 2x3 grid of visualizer windows, click the **Zoom-out (-)** magnifier icon.

Visualizer Warnings

As structures are loaded into the **Visualiser** additional messages will be displayed in the Scrolling Log. Normally this will consist of the text “Loading visualiser for block: <data block name>” for each block loaded. It is possible for some extra warnings to be displayed here relating to unknown element types and various cross-checks of atomic displacement parameters.

Right-Clicking in the Display Area

Right-clicking in the **Display** area generates menus which provide access to many display options. The options available will depend on whether you right-click on:

- A blank area in the **Display** window, away from objects such as atoms, bonds, centroids, planes or nonbonded contacts.
- Specific objects in the **Display** window, i.e. an atom, bond, centroid, plane or nonbonded contact.

Menu items will be greyed out if they are inapplicable given the current state of the display.

Picking Modes

The **Picking Mode** controls what happens when you left-click on items in the **Display** area:

- Use of the default **Select** mode permits the selection of atoms by clicking on them in the **Display** area. Once selected, a set of atoms can be subjected to an operation, e.g. colouring, or used for a calculation, e.g. of a centroid.
- Picking atoms in **Label** mode will display their atom labels - pick again to switch off the label.

Use of **Measure Distance**, **Measure Angle** or **Measure Torsion** modes permits the measurement of geometrical parameters by picking two, three or four atoms, respectively.

To set the picking mode, right-click and select the **Picking Mode** menu option, a sub-menu will be shown containing the various modes. The working of these modes is explained below.

Undoing Commands and Resetting the Display

Undoing Commands

Undo and redo are available when a single structure is being displayed; these actions are disabled when the 2x3 grid of structures is shown:

- Undo and redo are available via the **Edit** menu, toolbar buttons and **Ctrl-Z** and **Ctrl-Y** keyboard shortcuts.
- Undo and redo only affect changes made to the **Visualizer** display; they are entirely separate to the undo and redo options in the **Text Editor**.
- When the CIF content is changed and the visualizer updated, any previous undo and redo history is lost.

- If the view is switched from a single structure to the 2x3 display and back again, the undo and redo history is maintained.
- Each structure has its own separate undo and redo history.
- Actions performed whilst viewing the 2x3 display can be undone and redone if the view is switched to the relevant single structure.

Resetting the Display to the Default View Direction, Scale and Perspective

The display area can be returned to the default view direction (viz. looking down the crystallographic b axis) and default scale and perspective:

- Right-click in the **Display** background, select **View** from the resulting menu and then hit **Reset View** in the next menu.

Any crystal packing or intermolecular contact information will be retained in the resulting default view.

Selecting and Deselecting Atoms and Molecules

The Purpose of Selecting Atoms and Molecules

Selection of atoms or molecules is useful for changing properties such as display style, calculating objects such as mean planes, etc. The basic idea is that once a set of atoms has been selected, then one or more operations can be performed on those atoms.

Ways of Selecting Atoms and Molecules

Atoms may be selected or deselected in several ways:

- Right-click anywhere in the **Display** area (atom, bond or background) and choose **Selection** from the resulting menu.

- Set the **Picking Mode** to **Select** and click on individual atoms with the left mouse button to select them. Once selected, an atom can be deselected by clicking on it again.
- When the **Picking Mode** is set to **Select**, all atoms become deselected if you left-click anywhere in the display-area background.
- Press the **Shift** key and then click on any atom to select or deselect the entire molecule containing that atom.

Identifying Selected Atoms

Atoms which are selected are indicated by small yellow wireframe spheres.

Moving the Display and Setting the View Direction

Translating the Display

The contents of the **Display** area can be translated in several ways:

- By holding the middle mouse button down while moving the cursor in the display area (three-button mouse only).
- By moving the cursor in the display area while keeping both the left-hand mouse button and the **Ctrl** key pressed down.

Rotating the Display

The contents of the display area can be rotated by:

- Moving the cursor in the **Display** area while keeping the left-hand mouse button pressed down (x and y rotation).
- Moving the cursor in the **Display** area while keeping both the left-hand mouse button and the **Shift** key pressed down (z rotation).

Scaling the Display

The contents of the **Display** area can be scaled (i.e. zoomed in or out) by:

- Moving the cursor up and down in the **Display** area while keeping the right-hand mouse button pressed down.

Viewing Along Crystallographic Axes

The contents of the **Display** area can be viewed along real cell axes or reciprocal cell axes by:

- Right-clicking in the background, selecting **View** from the resulting menu, and then choosing the view direction from the resulting menu.

Changing the Display Perspective

The display perspective can be changed by:

- Moving the cursor from side to side in the **Display** area while keeping the right-hand mouse button and the **Shift** key pressed down.

Atom, Molecule and Background Display

Types of Atom and Bond Display Properties

Atoms and bonds have several display properties that you can control:

- Their display style: wireframe, stick, ball-and-stick, spacefill or ellipsoid (displays the ADPs).
- Whether atoms are labelled or not. You have no control over what the label is, you can only control whether it is shown or not.

- Their colour: atoms can be coloured according to their element type (see Standard Element Colours) or their symmetry equivalence (see Symmetry-Equivalence Colouring), or can be assigned a user-specified colour.
- Whether atoms and bonds are displayed or hidden.

Display properties can be set:

- Globally, i.e. for all atoms in the display.
- For any specified set of atoms (see Setting Display Properties for Particular Atom(s) or Bond(s)).

Setting a Global Display Style

To set all atoms to a new display style (wireframe, stick, ball and stick, spacefill or ellipsoid), you must first ensure that no atoms are selected (see Identifying Selected Atoms). You can then:

- Right-click in the **Display** area background, click **Styles** from the pull-down menu, and select the required style (**Wireframe, Capped sticks, Ball and stick, Spacefill** or **Ellipsoid**).
- Alter the ellipsoid settings (see Setting Ellipsoid Display Options).
- Switch the display of bond-types on or off by right-clicking in the **Display** area background, picking **Styles** from the pull-down menu, and selecting **Display Bond Types** from the resulting pull-down menu.
- Identify aromatic rings by displaying a circle within each ring by right-clicking anywhere in the **Display** area background, picking **Styles** from the pull-down menu, and selecting **Display Aromatic Rings** from the resulting menu.

Display styles can also be set for individual atoms, bonds and molecules (see Setting Display Properties for Particular Atom(s) or Bond(s)).

Turning All Atom Labels On or Off

To switch all atom labels on or off, you must first ensure that no atoms are selected (see Identifying Selected Atoms). You can then:

- Right-click in the **Display** area background, pick **Labels** from the pull-down menu, and select the required option (**Show labels, Hide labels**).

Atom labels can also be turned on and off individually (see Setting Display Properties for Particular Atom(s) or Bond(s)).

Setting a Global Colouring Scheme

To colour all atoms by element type (see Standard Element Colours) or symmetry equivalence (see Symmetry-Equivalence Colouring), you must first ensure that no atoms are selected (see Identifying Selected Atoms). You can then right-click in the **Display** area background, pick **Colours** from the pull-down menu, and select the required option (**Colour By Element, Colour by Symmetry Equivalence**).

Colours can also be set for individual atoms and molecules (see Setting Display Properties for Particular Atom(s) or Bond(s)).

Setting the Background Colour

You can switch between the default black background and an alternative colour by right-clicking in the **Background** area and hitting **Draw Backdrop**. The alternative colour will be a blue gradient.

Setting Display Properties for Particular Atom(s) or Bond(s)

There are two methods for setting the display properties of a particular atom, bond, molecule, or set of atoms:

- Select the atoms whose display properties you wish to change (see Ways of Selecting Atoms and Molecules). right-click in the **Display** area background, select the appropriate option from

the pull-down menu (**Styles, Colours, Labels, Show/Hide**) and then choose the desired display-property setting from the next menu.

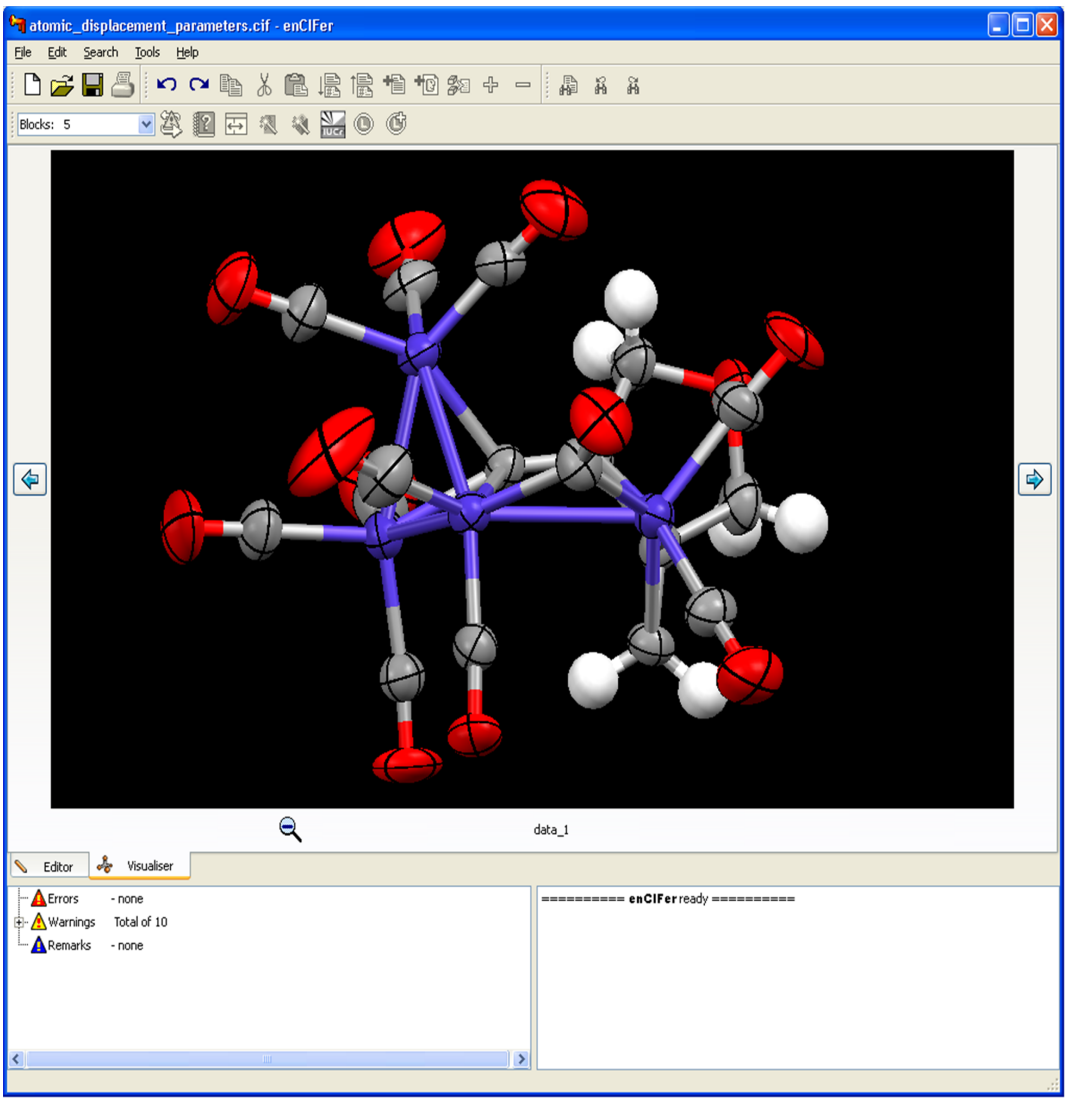
- Right-click on an individual atom or bond, select the appropriate option from the pull-down menu (**Styles, Colours, Labels, Show/Hide**) and then select the desired display-property setting from the next menu. The chosen setting will be applied to the atom or bond on which you clicked.

In addition:

- You can switch individual atom labels on and off by setting the **Picking Mode** to **Label** in the menu which results when the **Picking Mode** option is selected from the right-click background menu. left-clicking on an atom will then toggle its label on and off.

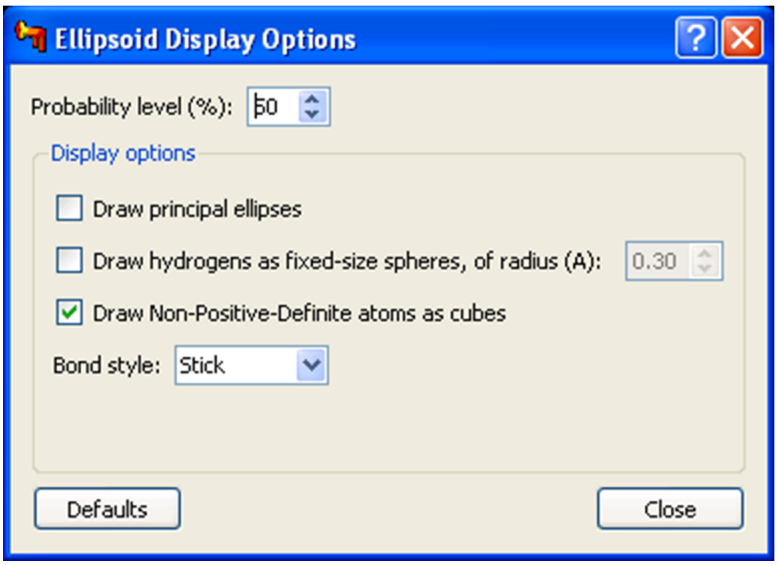
Setting Ellipsoid Display Options

Displacement ellipsoids can be displayed for structures where the CIF file contains Atomic Displacement Parameter values (Uequiv and Uij values) (see Atomic Displacement Parameters (ADPs)):



To display the Ellipsoid settings:

- Right-click in the **Display** area background, then select **Styles** > **Ellipsoid Settings**. This will open the **Ellipsoid Display Options** window which can be used to customize various settings.



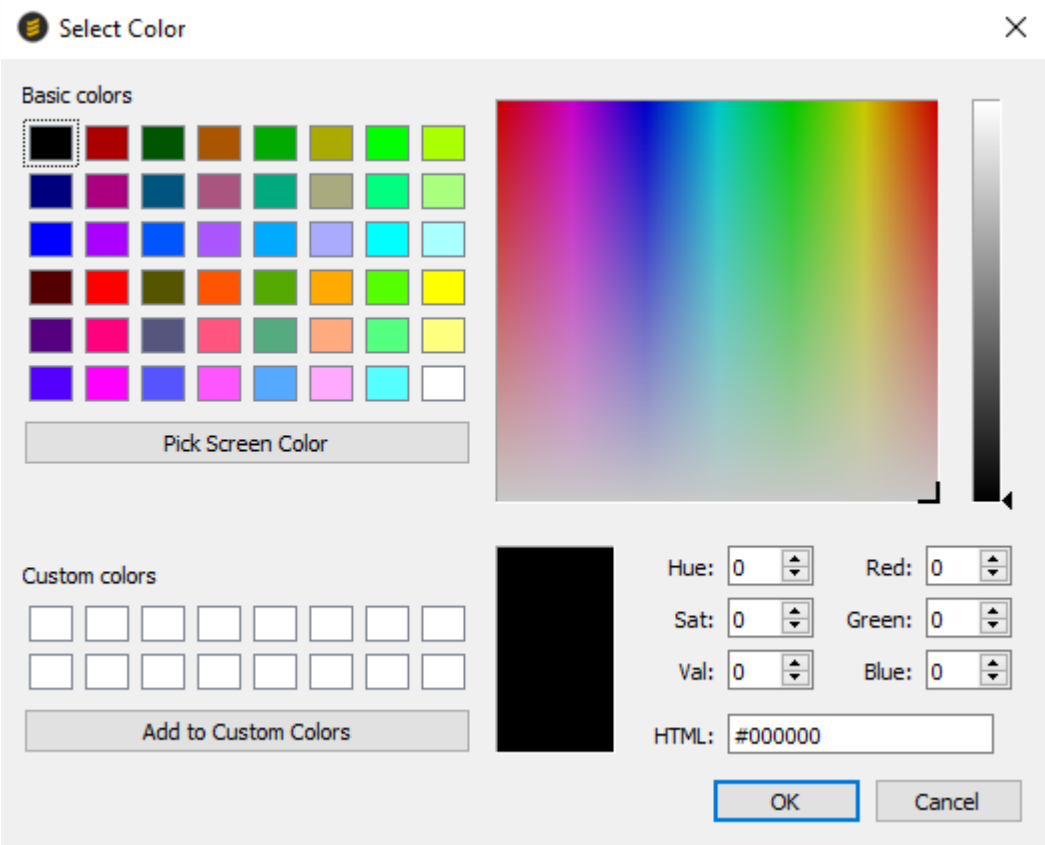
The display options are as follows:

- **Probability level (%)**: use this to alter the probability level (thus size) the ellipsoids are displayed at. For crystal structures determined from neutron diffraction data or at low temperature, you may want to increase the probability level; for structures determined at high temperature, you may want to reduce the probability levels.
- **Draw hydrogens as fixed-size spheres**: use this tick box to control the display of H atoms. The radius used for the spheres is customizable.
- The **Draw principal ellipses** and **Draw Non-Positive-Definite atoms as cubes** tick boxes can be checked or unchecked to enable or disable the specified functionality. A non-positive-definite atom is termed as such when the corresponding ellipsoid cannot be represented in real space, e.g. the ellipsoid may have a negative volume. Non-positive-definite atoms therefore highlight a problem with the data, e.g. a particular element may have been determined as a carbon atom but it is in fact a nitrogen atom.
- Bonds can either be represented as **Wireframe** or **Stick**.
- You can use the **Defaults** button to return to the standard settings.

Further information about ADPs is provided (see Atomic Displacement Parameters (ADPs)).

Using the Select Color Dialog Box

You may need to use the **Select Color** dialog box when choosing a colour for an atom, molecule or other object:



There are several ways to select a colour:

- Click on one of the “Basic colors”.
- Click on the **Colour** palette.
- Click on one of the “Custom colors” (only possible if you have defined some earlier in the EnCIFer session).
- Click in the vertical **Lightness** slider.
- Type your required Red, Green and Blue values.

- Type your required Hue, Saturation and Lightness (i.e. **Val**.) values:
    - The **Hue** defines the colour, e.g. 0 is red, 120 is green, and 240 is blue.
    - The **Saturation** defines the amount of colour; 255 is fully saturated, whereas a colour with a saturation of zero is always white.
    - The **Lightness** can also be controlled; a Lightness of zero always gives black.

Once you have chosen the required colour, hit **OK** to continue.

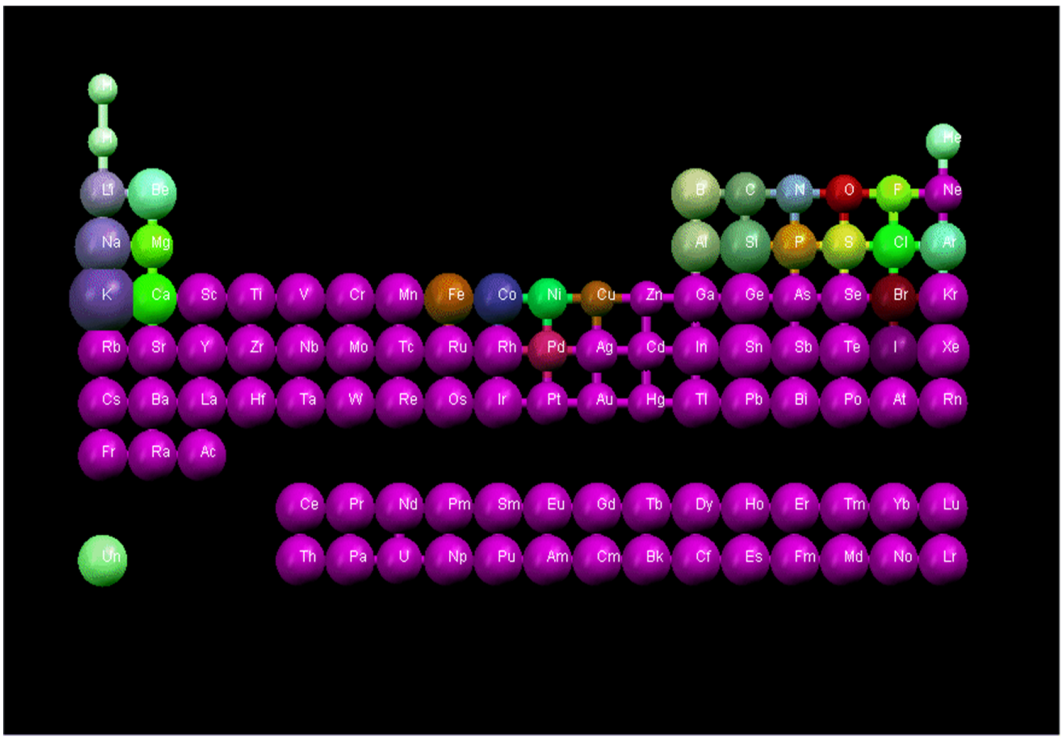
To create a customized palette, you can choose a colour and then hit **Add To Custom Colors**. This adds the current colour to the "Custom colors" area for use later in the same session.

Atom and Molecule Colouring Conventions

Standard Element Colours

When atoms are coloured by element (see Setting a Global Colouring Scheme), each atom is assigned a colour depending on its element type.

Colours used for the common elements are:



- All other elements are coloured purple.

Element colours can be changed by right clicking in the visualizer window and selecting the option **Colours -> Element Colours**. This will display an **Element Colour Option**s dialog in which colours can be customized.

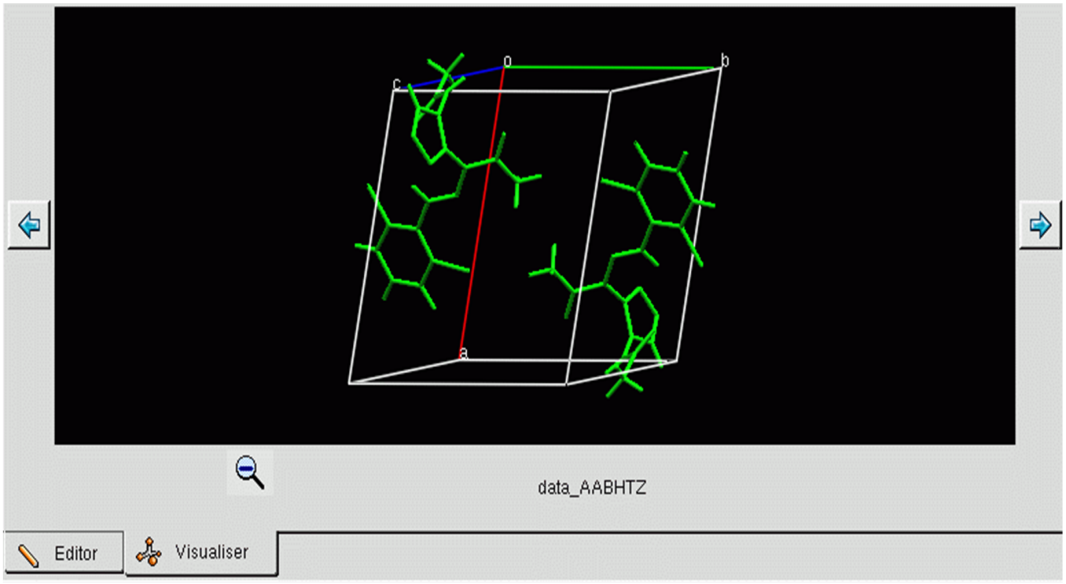
Symmetry-Equivalence Colouring

If all atoms are coloured **by Symmetry equivalence** (see Setting a Global Colouring Scheme), each different molecule (or ion) in the crystal chemical unit is assigned a different colour (the crystal chemical unit is the same as the asymmetric unit in the majority of structures).

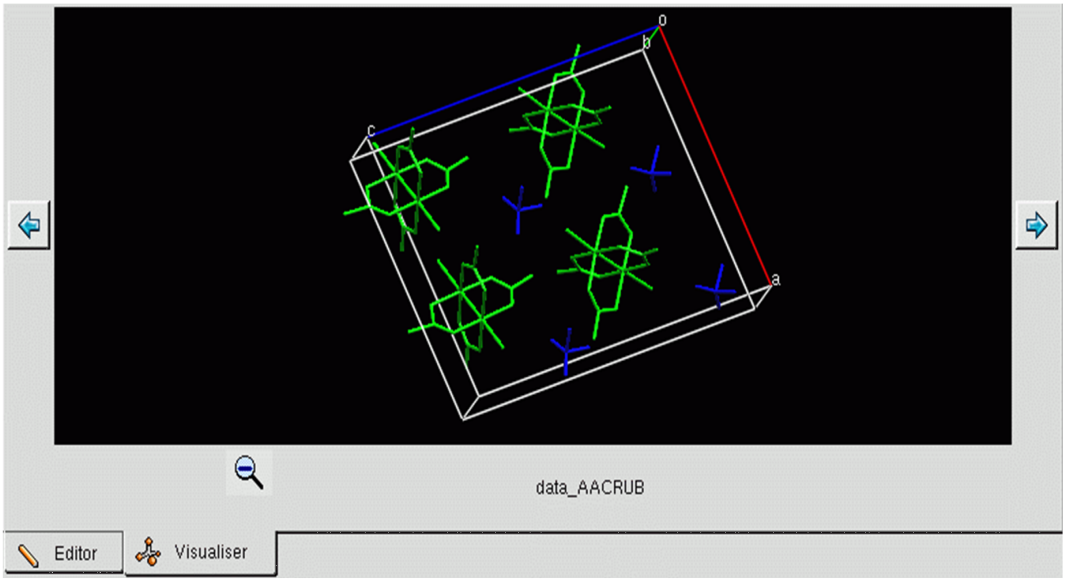
When a packing diagram is then constructed, any given molecule is assigned the same colour as that of the molecule in the crystal chemical unit to which it is related by crystallographic symmetry. This means that molecules of the same colour are crystallographically (and therefore chemically and geometrically) identical to each other.

For example:

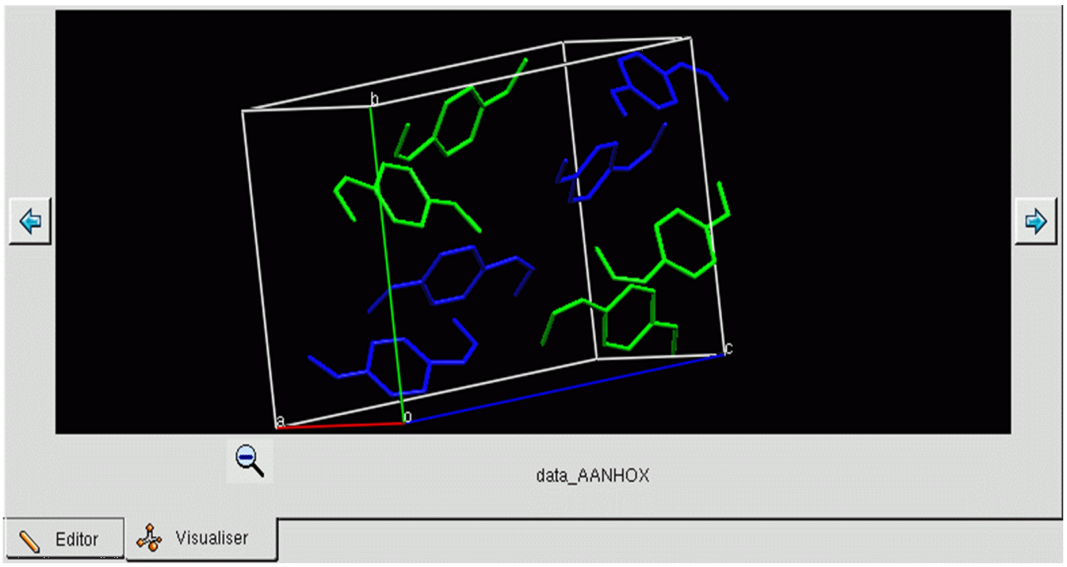
- If the crystal chemical unit of a crystal contains only one molecule, and a packing diagram of several unit cells is constructed, all molecules will have the same colour, e.g.



- If the crystal chemical unit contains two chemically different molecules or ions, each will be assigned a different colour and the molecules in the packing diagram will have two colours, e.g.



- If the crystal chemical unit contains two molecules which are chemically identical but not related by crystallographic symmetry, each will be coloured differently and, again, the packing diagram will contain molecules of two colours, e.g.



Measurement

Measuring Distances, Angles and Torsions

Measurement mode can be set in two ways:

- By right-clicking in the **Display** area background, selecting **Picking Mode** and setting the **Picking Mode** in the resulting menu, viz. **Measure Distance, Measure Angle** or **Measure Torsion**.
- By right-clicking in the **Display** area (but not on an atom or bond), selecting **Measure from the pull-down menu**, and then choosing **Measure Distances, Measure Angles** or **Measure Torsions**.

Geometrical measurements (intramolecular or intermolecular) can now be displayed by clicking on two atoms for a distance, three atoms for an angle or four atoms for a torsion angle.

Clearing Measurements

All geometrical measurements can be removed from the **Display** by:

- Clicking **Clear Measurements** in the tool bar near the top of the main window.

- Right-clicking in the **Display** area background and then hitting **Clear Measurements** in the resulting pull-down menu.
- Right-clicking in the **Display** area background, selecting **Measure** in the resulting pull-down menu, and then hitting **Clear Measurements** in the next menu.
- An individual measurement can be removed by right-clicking on any part of the green dotted line that does not lie on a chemical bond and selecting **Delete Measurement** from the resulting pull-down menu.

Displaying Packing Diagrams (Unit Cells)

Colouring Conventions for Unit Cell Axes

The crystallographic a, b and c axes are coloured red, green and blue, respectively.

Switching Packing On and Off

You can switch between displaying a packing diagram (the contents of a single unit cell) or a single molecule - or, more precisely, the molecule(s) in the crystal chemical unit by:

- Right-clicking in the **Display** area background, selecting **Packing** from the pull-down menu, and then using the **Molecule** or **Packing** options to switch between representations.

Appendices

Keyboard Shortcuts

Action	PC & UNIX	macOS
Select All Text	**Ctrl+A**	**Command key+A**
Show/Hide CIF Browser	**Ctrl+B**	**Command key+B**

Action	PC & UNIX	macOS
Copy Text	**Ctrl+C**	**Command key+C**
Delete next character	**Ctrl+D**	**Command key+D**
Go to end of line	**Ctrl+E**	**Command key+E**
Find Text	**Ctrl+F**	**Command key+F**
Find Next	**F3**	**F3**
Find Previous	**Ctrl+F3**	**Command key+F3**
Go to Line	**Ctrl+G**	**Command key+G**
Delete previous character	**Ctrl+H**	**Command key+H**
Data Item Dictionary Help	**Ctrl+I**	**Command key+I**
Re-Check CIF	**Ctrl+K**	**Command key+K**
Edit Loop	**Ctrl+L**	**Command key+L**
New File	**Ctrl+N**	**Command key+N**
Open File	**Ctrl+O**	**Command key+O**
Print	**Ctrl+P**	**Command key+P**
Quit/Exit	**Ctrl+Q**	**Command key+Q**
Replace Text	**Ctrl+R**	**Command key+R**
Save	**Ctrl+S**	**Command key+S**
Paste Text	**Ctrl+V**	**Command key+V**
Close File	**Ctrl+W**	**Command key+W**
Cut Text	**Ctrl+X**	**Command key+X**
Redo	**Ctrl+Y**	**Command key+Y**
Undo	**Ctrl+Z**	**Command key+Z**

Atomic Displacement Parameters (ADPs)

Atomic Displacement Parameters (also known as atomic vibration parameters and thermal parameters) are used to generate displacement ellipsoids. Displacement ellipsoids represent atomic motion and can be either isotropic or anisotropic. The shape and size of ADPs can be used to highlight potential errors with the data; the smaller and more spherical the ellipses for anisotropic atoms,

the better the data (see Troublesome Crystal Structures: Prevention, Detection and Resolution, R. L. Harlow, J. Res. Natl. Inst. Stand. Technol., **101(3)**, 327, 1996).

For ellipsoids to be displayed in Encifer, it is necessary for the relevant U_{equiv} and U_{ij} values to be present in the CIF file. U_{equiv} and U_{ij} values are not stored in the CSD, thus it is not possible to view ellipsoids for entries output from the CSD. CIFs obtained through the CCDC's online request form (www.ccdc.cam.ac.uk/products/csd/request) may contain these data where provided by the publishing author.

Ellipsoids can be displayed in Encifer by clicking on the **Visualiser** background, selecting the **Styles** submenu and then selecting **Ellipsoid**. Settings for the Ellipsoids display style may be controlled using the **Ellipsoid setting** window, available from the same menu.

Relevant Program Options:

- Displaying ellipsoids (see Setting a Global Display Style).
- Setting ellipsoid display options (see Setting Ellipsoid Display Options).

Tutorials

Tutorial 1: Viewing a CIF and Correcting Errors, Warnings and Remarks

Objectives

To identify errors in an input CIF.

To correct the errors and save the resulting file.

Steps Required

1. Open the CIF.

2. Fix all the syntax errors in the CIF.
3. Fix warnings.
4. Investigate remarks and fix where necessary.
5. Save the resulting error-free CIF.

The Example

For the purpose of this tutorial errors have deliberately been introduced into the example CIF, taken from the following journal reference: C.M.Hay, N.E.Leadbeater, J.Lewis, P.R.Raithby, K.Burgess, New.J.Chem., **22**, 787, 1998 (CSD refcode SUDQUD).

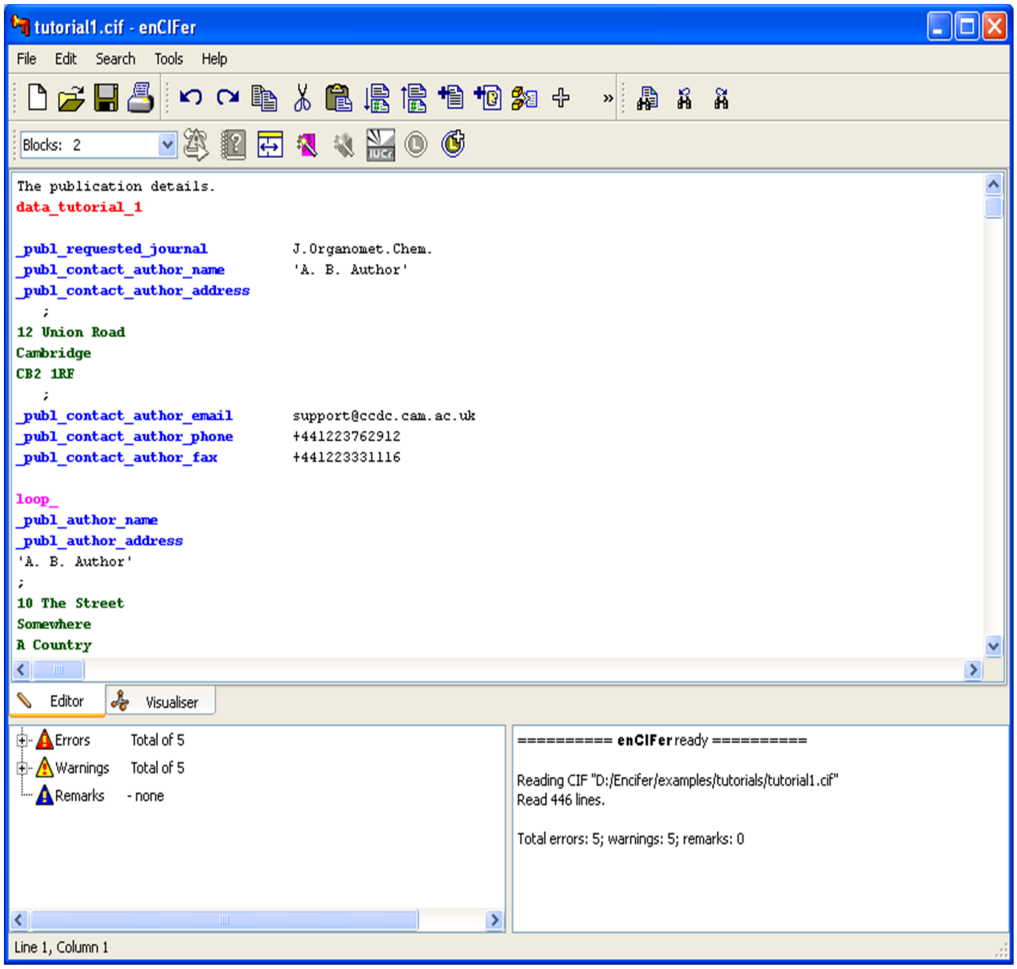
Menu Commands Required

Enable checking of mandatory data items

1. Select **Edit** on the top-level menu and select **Preferences**.
2. In the **Preferences** dialog box, click the **CIF** checking tab, then click the **Browse** button and select the template file `<InstallationDirectory>/Examples/Mandatory/recommended.cif` in the **Open** dialog box, and click the **Open** button.
3. Check the **Check for Mandatory Data Items box** to enable checking for mandatory data items.
4. Click **OK** to close the **Preferences** dialog box.

Open the CIF and save a copy

1. Click **File** on the top-level menu, and select **Open**. In the resulting **Open** dialog box, select the file: `<Installation Directory>/Examples/Tutorials/tutorial1.cif` and click the **Open** button.
2. The CIF should now be loaded in the editor with syntax highlighting. The error list view in the lower left of the window should report five errors, five warnings and fifteen remarks.



Fixing errors

1. Click the expand tab alongside the “Errors” red triangle icon, or double-click on the “Errors” list item. The list should expand to show five line numbers and associated error messages.

2. Double-click on the first error message, “Opening (;) semicolon not first character on line”. The cursor should move to the corresponding line in the editor, now highlighted in blue. To fix the error, delete the spaces before the semicolon so that it is the first character on the line (Do not re-check the CIF at this point).

3. Double-click on the second error message, “Closing (;) semicolon not the first character on line”, and edit the line to remove the leading spaces before the semicolon. The multi-line `_publ_contact_author_address` string should now be formatted correctly. Re-check the CIF now that you have corrected both the opening and terminating semicolons.

4. Double-click on the third error message, “CIF contains more than one data block called data_tutorial_1”. The second occurrence of `data_tutorial_1` in the editor should now be highlighted. Since data block names should be unique, change this data block name to `data_compound_1`.

5. Double-click on the next error message, “No terminating (‘) quote”. The problem is that a data value with spaces has an opening but not a closing quote. Edit the `_chemical_formula_moiety` to add the missing quote, i.e. `‘C28 H15 Au Cl O10 Os3 P’`.

6. Double-click on the error message, “Data name _chemical_compound_source not followed by data value”. Since each data name must be followed by a data value, and the source of the chemical is unknown, the data value should be edited to `?` (unknown).

7. To check that the errors are now fixed, click the check **CIF** icon on the toolbar or use the **Ctrl+K** keyboard shortcut (see Keyboard Shortcuts). In order to expand the text on the toolbar select **Edit** from the top-level menu and then **Preferences**. From the resulting **Preferences** dialog box, click the **General** tab, and check the **Show text on toolbar buttons** checkbox, then hit the **OK** button to close; the icons should now be clearly labelled on the toolbar. The error list view should be updated to show no errors, five warnings and fifteen remarks.

Fixing warnings

1. Expand the list view to show all the warning messages, all of these should be listed with their corresponding line numbers. Double-click on the warning message “Ignoring uncommented string(s) before first data block”.

2. The first line of the CIF should be highlighted in the editor. The problem is that no hash character precedes the comment (this is necessary even for comments before the first data block). Select the **Comment Selection** command from either the **Edit** menu or toolbar. The line should then read: `# The publication details.` and be shown in green italic style (if the default text styles are being used).

3. Double-click on the warning “Data value is not in the dictionary enumeration list”. To determine the enumeration list of values for this data item, right-click on the text `_exptl_absorpt_correction_type`, select Item help from the resulting menu and a data item **Dictionary help** dialog box should open. Note that the recommended values include `psi-scan`, not `psiscan`. Click the **Close** button to dismiss the data item help dialog box. Edit the data value `psiscan` to `psi-scan`.

4. Double-click on the warning message “Ignored string R3mV”. The problem with this line is that the data value `Siemens R3mV` is not single or double-quoted, so the data value `Siemens` is associated with the data name `_diffrn_measurement_device` and the other text `R3mV` has been ignored. To fix the problem, add single quotes around the data value. The line should now read: `_diffrn_measurement_device ‘Siemens R3mV’`

5. Double-click on the warning message “Data value is not a correctly formatted number”. The line with the data item name `_diffrn_standards_decay_%` should be highlighted. The warning indicates that a numeric data value is expected for the data item according to the CIF dictionary but the value given does not conform to the CIF definition of a numeric value. In this case, the problem is that a percent sign has been included in the value where this is not necessary, as the data name indicates that the value is to be interpreted as a percentage. To fix the warning, delete the percent sign from the value.

6. Double-click on the last warning message, “More than 80 characters long” (this assumes the default soft line length limit is being used). To fix the warning, insert a line break between the data name `_refine_ls_weighting_details` and the quoted data value, such that the two lines are both within the 80 character soft line length limit.

7. To check that the warnings have now been fixed, re-check the CIF by clicking the **Check CIF** icon on the toolbar or by using the **Ctrl+K** keyboard shortcut. The error list view should now report no errors, no warnings and fifteen remarks.

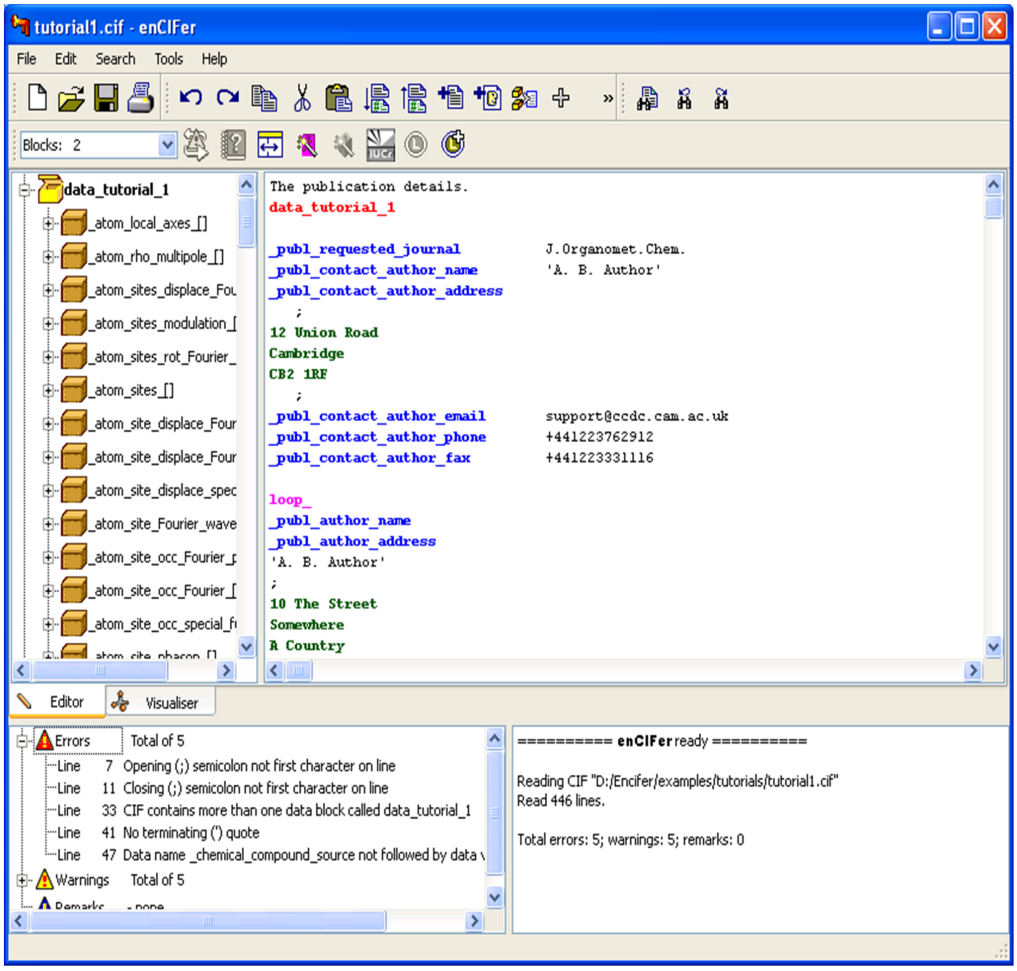
Investigating remarks

1. Expand the list view to show the fifteen Remarks messages.

2. Double-click on the following message “Data item _diffrn_radiation_probe not found in block data_compound_1”. The data block name should be highlighted.

3. Click on the **Browser** icon on the toolbar, or use the **Ctrl+B** keyboard shortcut. A browser showing the names of the blocks `data_tutorial_1` and `data_compound_1` should appear to the left of the editor pane.

4. To adjust the relative widths of the browser and editor, drag the vertical splitter which divides them.



5. Expand the browser for “data_compound_1” by clicking the **Expand** icon or by double-clicking on the **Box** icon. The currently enabled dictionary categories should now be shown.

6. Expand the `_diffrn_radiation[]` category. Data items which are set in the block data_compound_1 are shown in black, data items which are not set are shown in grey.

7. Double-click on `_diffrn_radiation_monochromator`. The data item should now be highlighted in the editor. Insert a blank line under this data item and place the cursor at the beginning of this line.

8. In the browser, right-click the data name `_diffrn_radiation_probe` and select **Paste into editor**. The text should be pasted into the editor at the current cursor position, and the data item should now be shown in blue in the browser.

9. To see the allowed values for the data item, right-click on the data item in the browser and select **Item Help** from the resulting menu (alternatively, right-click on the data item in the editor). The appropriate value in this case is `x-ray`, so dismiss the data item help dialog box by clicking on the **Close** button, and type `x-ray` after the data name in the **Text Editor**. The line should then read `_diffrn_radiation_probe x-ray`

10. Alternately you can copy text from the **Item Help** dialog. In the above case this would mean clicking and dragging to highlight the text 'x-ray' in the dialog, then right-clicking and selecting **Copy**. Then after dismissing the dialog position the cursor next to the new data item and select **Paste**.

11. Re-check the CIF by clicking the check CIF icon on the toolbar or by using the **Ctrl+K** keyboard shortcut.

12. Double-click on the message "Data item _refine_ls_hydrogen_treatment not found in block data_compound_1". The data block name should be highlighted.

13. Expand the `_refine_[]` category. Once again, data items which are set in the block "data_compound_1" are shown in black, data items which are not set (including `_refine_ls_hydrogen_treatment`) are shown in grey.

14. Double-click on the data item `_refine_ls_number_restraints`. This data item should now be highlighted in the editor.

15. In the browser, right-click the data name `_refine_ls_hydrogen_treatment` and select **Set data item value**.

16. The allowed values for the data item are shown in the **Data value edit** dialog box. The appropriate value in this case is `constr`, so select the data value `constr` from the pull-down enumeration list and click **OK**. The line should then read `_refine_ls_hydrogen_treatment constr`

17. Double-click on the remark: “Data item _symmetry_Int_tables_number not found in block data_compound_1”. The data block name should be highlighted.

18. Expand the `_symmetry_[]` category.

19. In the **Text Editor**, insert a blank line under the data item: `_symmetry_space_group_name_H-M` and place the cursor at the beginning of this line.

20. In the browser, right-click the data name `_symmetry_Int_tables_number` and select **Set data item value**.

21. The appropriate value in this case is `14`, so type this in the lower window in the edit data dialog. Click the **OK** button. The line should then read `_symmetry_Int_tables_number 14`

22. Double-click on the remark: “Data item _chemical_name_common not set in block data_compound_1” As the trivial name for this compound is unknown, the data value should be left as `?` (unknown).

23. Double-click on the remark “Data item _chemical_melting_point not set in block data_compound_1”. The melting point is also unknown, so the data value should be left as `?` (unknown).

24. Double-click on the remark “Data item _exptl_crystal_colour not set in block data_compound_1”. In this case, the crystal is known to be red in colour, so the data value `?` should be changed to `red`.

25. Re-check the CIF. The error list view should show no errors, no warnings and eleven remarks.

26. You may review the remaining remarks as a further exercise.

Save corrected CIF

1. Clear the scrolling list of messages in the lower right pane, by right-clicking on the pane and selecting **Clear** from the resulting pull-down menu.
2. Save the file by either hitting **File** on the top-level menu and selecting **Save**, by clicking the **Save** icon on the toolbar or by using the **Ctrl+S** keyboard shortcut.
3. Exit EnCIFer by hitting **File** on the top-level menu and selecting **Exit**, or by using the **Ctrl+Q** keyboard shortcut.

This ends the tutorial.

Tutorial 2: Preparing a CIF for Deposition using the Data Entry Wizards

Objectives

To add additional chemical, crystallographic and experimental data to a CIF.

To add author and journal information to the CIF.

Steps Required

1. Open the example CIF and save a copy.
2. Add chemical, crystallographic and experimental data to the CIF using the Crystal **Data Wizard**.
3. Set your contact author details in the preferences.
4. Add author and journal data to a CIF using the **Publication Data Wizard**.
5. Save the augmented CIF.

The Example

For the purpose of this tutorial, errors have been deliberately introduced into the example CIF, taken from the following journal reference: J.Lewis, Chi-Keung Li, M.R.A.Al-Mandhary, P.R.Raithby, J.Chem.Soc., Dalton Trans., 1915, 1993 (CSD refcode YARDAW).

Menu Commands Required

Open the required CIF and save a copy

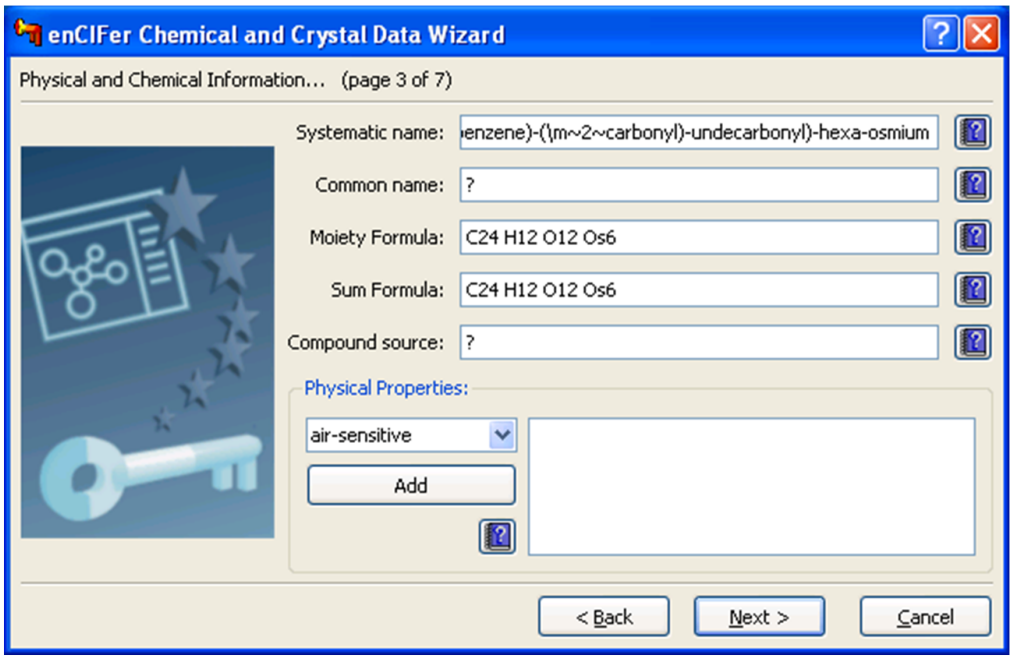
1. From the top-level file menu, select **Open** and browse to locate the CIF, which is located in: `<Installation Directory>/examples/tutorials/tutorial2.cif`.
2. Save a personal copy of this file. From the top-level file menu, select **Save As**, choose a file name in the resulting **Save As** dialog box and click the **Save** button to save the copy.
3. The CIF containing the data block `data_tutorial_2` should be loaded.

Add chemical, crystallographic and experimental data

1. Click on the data block name `data_tutorial_2` in the editor and click the Crystal Data Wizard icon. The introductory wizard page should appear.
2. Click on the **Next** button. The **CIF syntax checking** page should be shown, showing no Warnings or fatal errors. It may show several remarks if checking of mandatory data items has been enabled as detailed in tutorial 1. Click on the **Next** button again.
3. **Chemical Information** Enter the name `bis(eta6-benzene)-(mu2-carbonyl)-undecacarbonyl-hexa-osmium` into the **Systematic name** box. To conform with CIF conventions, this should be entered as:

   ```
   bis(\h^6^-benzene)-(\m\~2\~carbonyl)-undecarbonyl-hexa-osmium
   ```

4. There is not a known trivial name, so the **Common Name** box can be left as `?` (unknown).



5. Click on the help icon to the right of the **Compound source** box. This should display the data item dictionary help dialog box for `_chemical_compound_source`. Click the **Close** button to dismiss the dialog box. As the source is unknown, the existing value of `?` can be left unchanged.

6. The **Physical Properties** pane may be used to enter any relevant information about the crystal. To aid this, a list of common terms is provided in the drop-down list. These common terms can be appended to the text by first selecting a term and then pressing the **Add** button.

7. Click the **Next** button to show **Physical and Chemical Information** Enter `dark red` in the **Crystal Colour** box. Quotes should not be entered as the wizard will add these automatically.

8. Similarly, enter `parallelepiped` in the **Crystal Habit** box.

9. The **Recrystallisation Method** item should be used to list the solvents used in the crystal preparation. See the **Data item help** for full details of all the text which may be entered here.

10. Click the **Next** button to show the **Symmetry Information** page. The **Crystal system** pull-down box should show `orthorhombic`, the **International Tables space group number** `61` and **Hermann-Mauguin space group symbol** `P b c a`. Although

these data are not set in the CIF, they are derived automatically by the wizard from the symmetry operators `_symmetry_equivalent_positions_as_xyz` in this case.

11. Click on the arrow on the **Hermann-Mauguin space group symbol** pull-down. This should display three possible settings for space group number 61; ? (unknown), P b c a and P c a b. If you select P c a b, the wizard should report: “Warning: Space group symbol: P c a b does not agree with the given symmetry operators: P b c a.*” *To correct this error, change the symbol back to P b c a.

12. Click on the **Next** button to show the **Diffraction information.** The experiment was performed with an X-ray source, so click on the **Radiation Probe** pull-down list and select **X-ray**.

13. The radiation type and wavelength have been set correctly already so can be left unchanged.

14. If the **Absolute Configuration** of the structure is known, it may be selected here: see the **Data item help** for definitions of available terms.

15. Click **Next** to show the **Apply changes to CIF** page. Click **Finish** to apply the changes to the CIF in the editor.

Adding contact details to the preferences

1. As it is useful to include the same author contact details in any CIF which you may publish, these may be entered once in the preferences and saved for future use.

2. Hit **Edit** on the top-level menu and select **Preferences**. The **Preferences** dialog box should appear; click the **Wizard** tab to show the **Contact author details** pane.

3. Enter your name, email, address, telephone and fax numbers in the appropriate boxes. There is no need to enter any quotes as the Wizard will add these automatically.

4. Click the **OK** button to apply the settings.

Adding publication data

1. To add a publication data block above the data block `data_tutorial_2`, insert a few blank lines above the `data_tutorial_2` line in the editor.

2. Place the cursor at the beginning of the first blank line and hit the **Publication Wizard** icon on the toolbar (if there is more one or more crystal structure blocks in the CIF, it is usually more convenient to include publication data which is common to all structures in a single block above the crystal structure blocks).

3. The **Publication Data Wizard** should start. Click **Next** to show the errors and warnings page (there should be none). Again, there may be some remarks depending on preference settings.

4. Click the **Next** button again to show the **Contact Information** page. As there were no contact details in the CIF, the wizard fills in the boxes from the information set in the Wizard preferences automatically. The fields may be edited further if desired.

5. Click the **Next** button to show the **Publication Information** page. Assuming the structure is not yet published, click the button **is being submitted to a journal for publication**, and click the **Next** button again.

6. Start to type in the (abbreviated) requested journal name, `Dalton Trans.` into the **Text** edit box. As you type, the scrolling list box is updated to show matching journal titles. Double-clicking on the matching title completes the text in the **Text** edit box.

7. Click the **Next** button to show the author information page. As there are no authors given in the CIF, the contact details are automatically shown in the name and address boxes.

8. Click on the data item dictionary help icon for **Author name** to find out the correct format for entering author names, then click the **Close** button to dismiss the dialog box.

9. In the **Author address** box the address you entered into the preferences dialog should be shown, although you can edit the text if you wish.

10. Try adding another author name by clicking on the **Add** button. Enter the name in the **Family name and First name(s) or initial(s)** boxes (no quotes are required). The address of the first author is copied for the second author. Amend the address so that it is correct for the second author (this should span several lines). If you now click the **Previous** button, the boxes are updated to show the name and address of the first author.

11. To add further authors, click the **Add** button again.

12. Click the **Next** button to show the final page, and **Finish** to apply the changes to the CIF.

Save the augmented CIF

1. Save the file by hitting **File** on the top-level menu and selecting **Save**, by clicking the **Save** icon on the toolbar or by using the **Ctrl+S** keyboard shortcut.

2. Exit EnCIFer by hitting **File** on the top-level menu and selecting **Exit**, or by using the **Ctrl+Q** keyboard shortcut.

This ends the tutorial.

Tutorial 3: Using the Loop Editor and Visualizer

Objectives

To identify and correct loop errors in an input CIF.

Add a loop containing the symmetry operators.

To view the corrected structure and save the resulting file.

Steps Required

1. Open the CIF and save a personal copy.

2. Fix loop errors in the spreadsheet **Loop Editor**.

3. Add a loop using the spreadsheet **Loop Editor**.
4. Check the structure in the **Visualizer**.
5. Save the CIF.

The Example

For the purpose of this tutorial errors have deliberately been introduced into the example CIF, taken from the following journal reference: A.J.Amoroso, L.P.Clarke, J.E.Davies, J.Lewis, H.R.Powell, P.R.Raithby, G.P.Shields, J.Organomet.Chem., **635**, 119, 2001 (CSD refcode CAFMOM).

Menu Commands Required

Open the CIF and save a copy

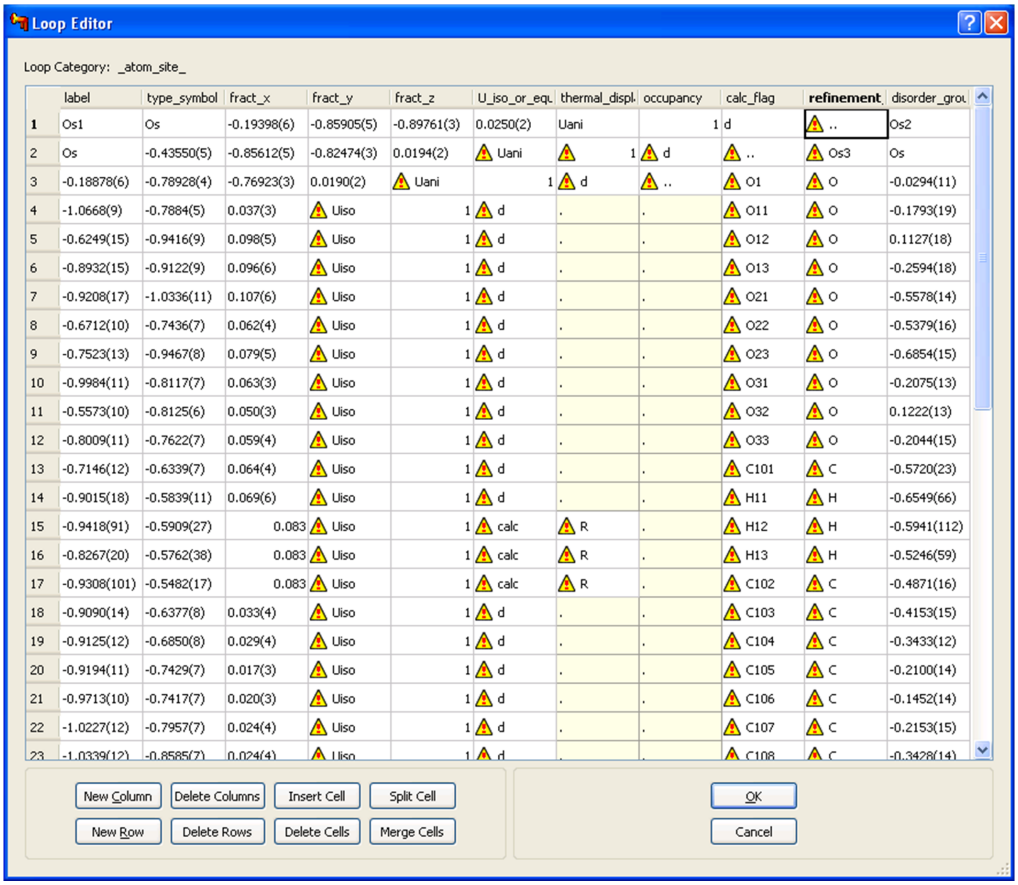
1. Hit **File** on the top-level menu, and select **Open**. In the resulting **Open** dialog box, select the file: `<Installation Directory>/Examples/Tutorials/Tutorial3.cif`.
2. Save a personal copy of this file. From the top-level file menu, select **Save As**, choose a file name in the resulting **Save As** dialog box and click the **Save** button to save the copy.
3. The CIF should now be loaded in the editor with syntax highlighting. The error list view in the lower left of the window should two errors, sixteen warnings and twenty-one remarks. You may a different number of warnings if you have adjusted the **Maximum Consecutive Error Lines** preference setting. The presence of remarks assumes you have enable checking of mandatory data items as detailed in tutorial one.

Fixing loop errors

1. Click to expand the “Errors” in the error list view in the lower left corner of the EnCIFer window. Double-click on the first error message: “Too many or too few data values in the loop”. The `loop_` line above the looped data should be highlighted and the

Edit Loop icon in the toolbar should become active. Click the **Loop Editor** button in the toolbar to show the **Loop Editor** spreadsheet.

2. You may wish to adjust the horizontal size of the spreadsheet to view all columns; this can be achieved by clicking and dragging the bottom right corner of the dialog. Alternatively, the view can be scrolled as required.



3. The spreadsheet shows that some of the data values are misaligned. Cells with unknown values (?) are coloured blue, inapplicable values yellow (.), empty cells grey and other cells white. Yellow warning triangles indicate that the data value does not conform to the recommended values for this data item (these also give rise to warnings in the **Editor**).

4. Scroll the spreadsheet view to examine the first row of the loop. The penultimate column contains the data value `..` and the last column an atom label. The error in this case is that both the loop columns `_atom_site_refinement_flags` and `_atom_site_disorder_group` should both contain the data value `.`

but the space has been omitted, such that the first atom label in the next row is interpreted as the data value for `_atom_site_disorder_group`.

5. To fix this problem, double click in the penultimate cell on the first row containing the value `..` move the cursor between the two `.` characters (either by clicking again or using the cursor keys) and click the **Split Cell** button. The cell at this position will now contain a single `.` character and a new cell inserted also containing a single `.` character. Subsequent values will have moved forward by one position in the table.

6. Repeat these steps for the second and third rows of the table. Examine the spreadsheet to check that the data values are now aligned correctly in columns. The cells in both columns should now be coloured yellow.

7. The loop is now free from warning triangles and empty cells, click the **OK** button to apply the changes to the CIF displayed in the **Editor**.

8. The CIF is checked automatically upon exiting the **Loop Editor**. The error list view should now report one error, two warnings and twenty-one remarks.

9. Expand the error list and double-click on the remaining error message: *“*Too many or too few data values in the loop”. Click on the **Edit Loop** button in the toolbar to open this loop in the **Spreadsheet editor**.

10. Examine the spreadsheet view. The fourth row should contain four missing cells and three cells with data values. The second, third and fourth rows contain spurious inapplicable `.` data values which disrupt the loop alignment.

11. Click on first of the `.` data values, and then click on the **Delete Cells** button to remove the cell from the table. Repeat for the remaining 2 cells. You can also click in the first of these cells and then use **Ctrl+Left-Click** in the cells to select the three cells. The positions of the other cells should be adjusted to correct the column alignment, and the fourth row is deleted (as it would now contain only empty cells).

12. Click on the **OK** button to apply the changes to the CIF in the **Text Editor**. The error list view should now report no errors, no warnings and twenty-one remarks.

Adding a loop for the symmetry operators

1. Click on the **Visualiser** tab below the **Editor** pane. Right-click in the visualizer background and select **Packing** from the resulting menu. In the sub-menu, select **Packing**. This will produce a 'Packing view' pop-up box stating "Packing display will not correctly display structures without a valid cell and space group". The problem is that no symmetry equivalent positions are given in the CIF nor any space group symbol. Return to the editor by clicking the **Editor** tab.

2. To add the symmetry equivalent positions, place the cursor in the editor on a blank line below the data item `_symmetry_cell_setting`, and click the **New Loop** icon on the toolbar. An empty **Spreadsheet editor** window should appear.

3. Click on the **New Column** button to show the **Choose Data Item** dialog box. This shows a CIF browser of data items in the currently enabled dictionary, expanded to show categories. Scroll down to show `_symmetry_equiv_[]` and expand the category to show data items (double-click on `_symmetry_equiv_[]` or click the expand box alongside).

4. Double-click on the data name `_symmetry_equiv_pos_site_id` to add it to the **Text** edit box, and click the **OK** button (alternatively, the data item name may be typed in the **Text** edit box). A column heading with this data item name should be added to the spreadsheet.

5. Click on the **New Column** button again, the **Choose Data Item** dialog box should still be expanded to show the `_symmetry_equiv_[]` data items. Double-click on the data name `_symmetry_equiv_pos_as_xyz` and click the **OK** button. The spreadsheet should now have two columns headed `site_id` and `as_xyz`, the common initial part of the data names (`_symmetry_equiv_pos_`) having moved to the **Loop Category** indicator.

6. As the space group is P212121 with four symmetry equivalent positions, click the **New Row** button four times to add four rows. An empty 2 x 4 table should now be shown in the spreadsheet.

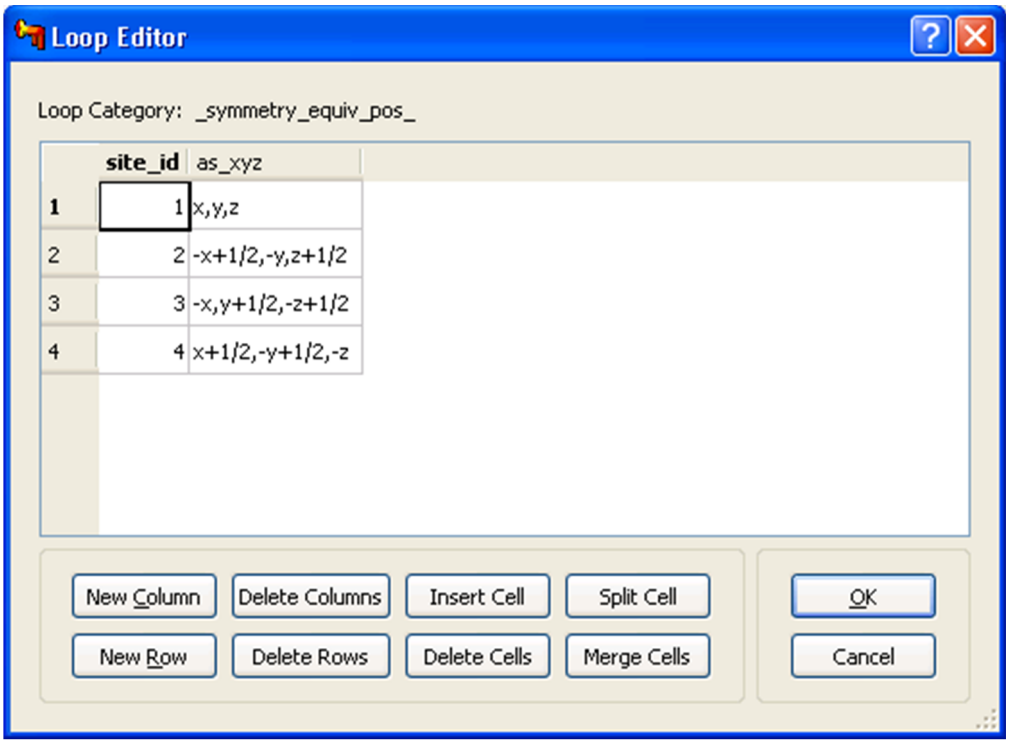
7. Enter the data for the `site_id` and `as_xyz` columns as follows into the spreadsheet, by clicking on each cell, typing in the value, and pressing **Enter**. For convenience the `_symmetry_equiv_pos_as_xyz` values follow:

```
x,y,z

-x+1/2,-y,z+1/2

-x,y+1/2,-z+1/2

x+1/2,-y+1/2,-z
```



8. Once all the data have been entered, click the **OK** button to add the loop to the CIF in the **Text Editor**.

9. Set the data item `_symmetry_space_group_name_H-M` to `P 21 21 21`. To find the data name, hit **Search** on the top-level menu and select **Find**, or use the **Ctrl+F** keyboard shortcut. In the resulting **Find Text** dialog box, type in the data item `_symmetry_space_group_name_H-M` (`_symmetry_sp` is adequate) and

click the **Find** button. The corresponding text should be highlighted in blue in the editor. Click the **Close** button to dismiss the **Find** dialog. Change the data value from `?` to `P 21 21 21`, remembering to add single or double quotes around the value.

10. Re-check the CIF (**Ctrl+K**). The error list view should now report no errors, no warnings and nineteen remarks.

11. Investigate the remaining remarks yourself and fix them where appropriate. To help do this it may be useful to refer back to tutorial one.

Viewing the crystal structure

Click on the **Visualiser** tab. As before in step 3, attempt to show the contents of one unit cell by right-clicking in the visualizer background and selecting **Packing** from the resulting menu and then clicking **Packing** in the sub-menu. Now the symmetry operators have been correctly assigned, the visualizer will show one unit cell with the cell axes colour-coded as a (red), b (green) and c (blue).

Saving the file

Save the file either by hitting **File** on the top-level menu and selecting **Save**, by clicking the **Save** icon on the toolbar or by using the **Ctrl+S** keyboard shortcut.

This ends the tutorial.