

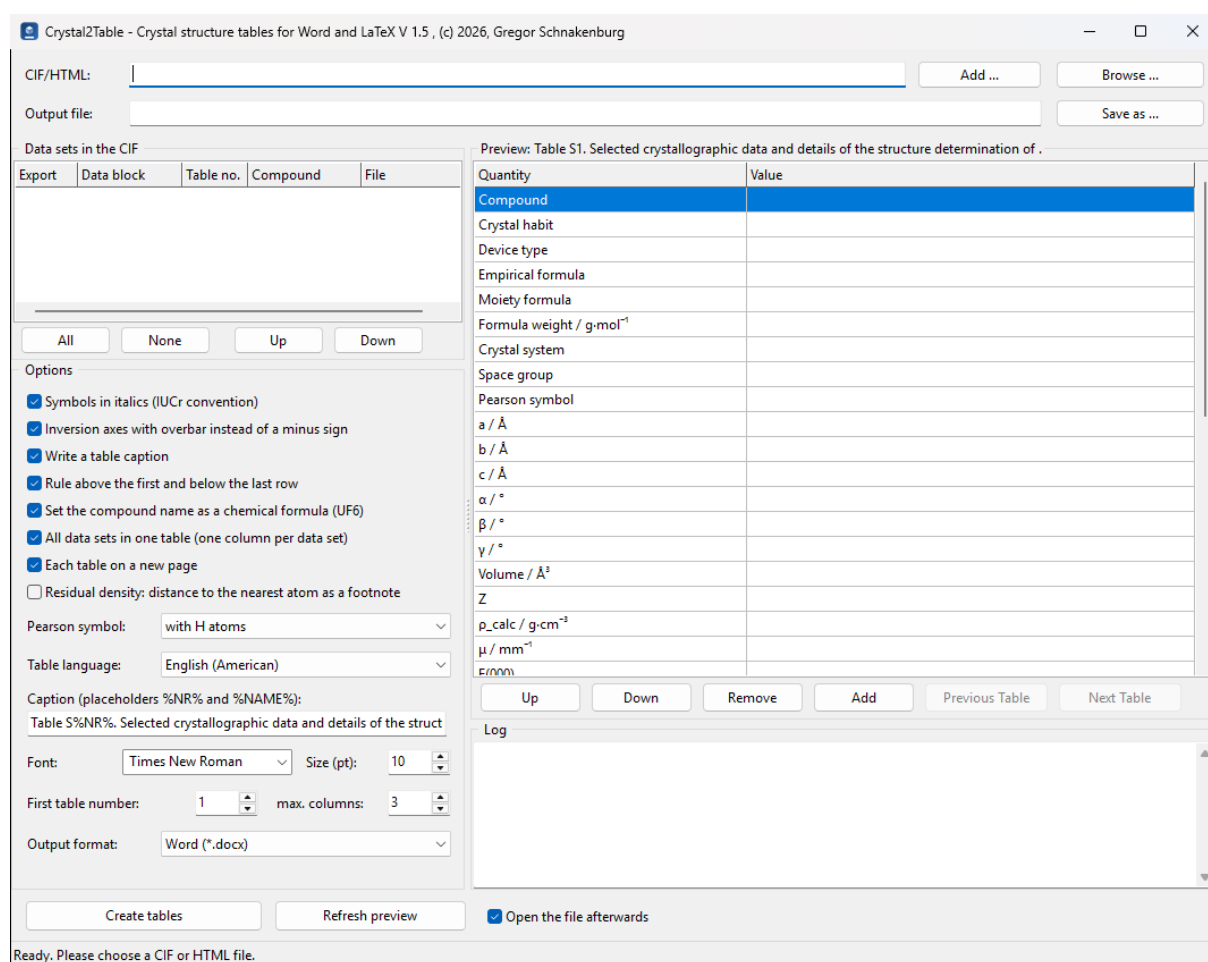
# Crystal2Table – Short Manual

For program version V 1.5

Crystal2Table turns crystallographic data into the table “Selected crystallographic data and details of the structure determination” – as a Word file (.docx) or as LaTeX source (.tex.txt). It reads CIF files and structure reports in HTML format (see below). If a file holds several data\_ blocks, or if several files are loaded, you get either one table per data set or a single table with one column per data set.

Word does not have to be installed: the program writes the .docx file itself. Started without parameters it opens its window; with --cif=... it runs in batch mode (see the last section).

## The window at a glance



- **Top:** the two paths – CIF file and Word file – with the **Browse ...** and **Save as ...** buttons
- **Upper left:** the list of data sets in the CIF
- **Lower left:** the options
- **Upper right:** the preview of the table
- **Lower right:** the log
- **Bottom:** **Create Word tables**, **Refresh preview** and **Open the file afterwards**; the status line is at the very bottom

The divider between the left and right half can be dragged; window size and position are remembered until the next start. If the window gets too small for all the options, the options box grows a scroll bar, so every option stays reachable.

## Step by step

### 1. Choose the file

**Browse ...** opens the file dialog, where several files can be picked at once – CIF files (`.cif`), HTML reports (`.html`) or a mixture of both. Their data blocks end up in one list and the “File” column shows where each came from. Paths can also be typed into the field by hand, separated by semicolons and confirmed with the Tab key. Files still encoded in Windows-1252 (ANSI) are converted to UTF-8 on reading.

**Add ...** loads further files **in addition**: the rows already set up keep their order, ticks and edited names, and the new data sets are appended to the list. **Browse ...** starts over instead.

Files can also be dragged into the window with the mouse – one at a time or several at once, anywhere on the window. If nothing is loaded yet they are opened straight away; otherwise the program asks whether to **add** them or to let them **replace** what is loaded. Folders are skipped and noted in the log; the files enter the list in the order of their names.

Blocks without structure data (e.g. `data_global`) are recognised and not preselected; the log says so.

### 2. Set up the data sets

The list has five columns:

- **Export** – ticked: this data set goes into the Word file
- **Data block** – its name in the CIF (read-only)
- **Table no.** – table number, numbered automatically
- **Compound** – the name used in the caption; the data block name by default. Click the cell and type to change it
- **File** – which CIF file the block came from

Below the list, **All** and **None** select every or no data set. **Up** and **Down** move the selected data set and thereby fix the order of the tables – or, in the combined table, the order of the value columns. Table numbers are reassigned as you go.

### Equal cell parameters

All six cell parameters appear in the table. What symmetry makes equal is not spelled out twice, though: it refers to the first parameter of its group. Monoclinic cells give  $\alpha = 90^\circ$  and  $\gamma = \alpha$ ; orthorhombic  $\beta = \alpha$  and  $\gamma = \alpha$ ; tetragonal in addition  $b = a$ ; cubic also  $c = a$ . Trigonal and hexagonal cells give  $b = a$  and  $\beta = \alpha$ , while  $\gamma = 120^\circ$  is written out; in a rhombohedral setting all three edges and all three angles are equal.

The reference is only made when the CIF really gives the same value. If a value differs despite the crystal system stated, it is written out – the crystal system is then more likely wrong than the measured value.

### 3. Read the preview

On the right the table appears as it will be written. Superscripts and subscripts are hinted at with Unicode characters ( $\text{g}\cdot\text{mol}^{-1}$ ,  $\rho_{\text{calc}}$ ). The preview shows **all** rows currently set up – including those this data set has no value for. Empty rows are dropped from the Word file. The exception is the Flack parameter: in non-centrosymmetric space groups its cell stays in the table even when empty, so that a missing value catches the eye during proofreading.

With “All data sets in one table” switched on, the preview shows the combined table with one column per data set; otherwise the data set selected in the list.

### 4. Set up the rows of the table

- **Up / Down** move the selected row
- **Remove** takes every selected row out – select several with the shift key held down or by dragging the mouse
- **Add** opens a list of every further quantity the program can pull out of the CIF: colour, crystal morphology,  $R_{\text{int}}/R_{\sigma}$ , number of parameters/restraints/constraints, the individual Miller ranges, wavelength, the programs used, the CCDC number and more. The rows you tick are inserted after the currently selected row. **Select all** and **Deselect all** tick or untick the whole list
- **Previous Table / Next Table** step through the tables when the data make more than one: with the data sets combined, through the blocks of at most “max. columns” data sets, otherwise through the data sets themselves

### 5. Change the labels

The label of a row can be overwritten directly in the “Quantity” column: click the cell and type. CIF notation applies:

- $\wedge 3 \wedge$  superscript, e.g.  $\text{\AA}^3$
- $\sim \text{calc} \sim$  subscript, e.g.  $\rho_{\text{calc}}$
- $\backslash a \backslash b \backslash g \dots$  for Greek letters ( $\alpha$ ,  $\beta$ ,  $\gamma$ )
- $\$V\$$  italic – between two section signs
- $\$Z\$$  bold – between two dollar signs

These can be combined with each other and with super- and subscript:  $\$p\$ \sim \text{calc} \sim$  gives an italic  $\rho$  with a subscript calc. In values taken from the CIF,  $\$$  and  $\$$  stay ordinary text – only labels you type yourself know the markup. With the option “Symbols in italics” switched off, your own italics go as well, along with all the others.

In the preview the label appears as it will be set ( $V / \text{\AA}^3$ ). As soon as the cell is opened for editing – double click or F2 – the notation with its control characters comes back ( $\$V\$ / \text{\AA}^3$ ), even for rows never named by hand. If editing is cancelled or nothing is changed, the label stays as it was.

An empty entry restores the original label.

### 6. Write the file

**Create tables** writes the file to the path in the “Output file” field. If that field is empty, the name of the first input file with the matching extension is used. If the file already exists you are asked first. With **Open the file afterwards** ticked, Word starts as soon as the file is written – or the text editor for LaTeX.

The **output format** decides what is written: a Word file (.docx) or LaTeX source (.tex.txt). With LaTeX the surrounding document sets the type, so font and size are switched off. Only the table is written – table with tabular and the caption in \caption, so LaTeX numbers the tables itself; the number from the template is left out. For “S1” numbering put `\renewcommand{\thetable}{S\arabic{table}}` in the preamble (the file says so in a comment). Without “Write a table caption” only the bare tabular is left. Special characters become LaTeX commands; the file needs no extra package.

When an option is changed that alters the preview – overbar, caption and its template, compound name as a formula, the footnote on the residual density, Pearson symbol, language, combined table, number of columns, first table number – the preview is rebuilt at once. Rules, page break, font and font size only take effect in the file that is written and leave the preview untouched.

**Refresh preview** rebuilds the preview without writing anything – needed only when the input file has changed on disk.

## HTML reports instead of a CIF

From a structure report in HTML format the first table is read – the one with the crystallographic data. Everything after that works as it does for a CIF: the same rows, the same languages, the same formatting.

The **compound name** is taken from “Identification code”, before the //. Where two names separated by a comma are given, the one that does **not** start with the first four characters of the file name is used: in GCLU677.html, “GCLU677, APA-030 // GXraymo\_9769f” gives APA-030. It can be overwritten in the list as usual.

If GCLU677.res sits next to GCLU677.html, the values the report does not carry are taken from it: residual electron density, the number of observed reflections, the weighting scheme, extinction and twin fraction. Whatever the report states is left untouched – the .res file only fills gaps.

## Options

| Option                      | Effect                                                                                                                                                                                                                                                                                                                                 |
|-----------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Symbols in italics          | $\alpha$ , $b$ , $c$ , $\alpha$ , $\beta$ , $\gamma$ , $Z$ , $\rho$ , $\mu$ , $F$ , $T$ , $\Theta$ , $\lambda$ , $h$ $k$ $l$ , $R_1$ , $wR_2$ , $l$ , $\sigma$ , the series letter of the radiation (Mo-K $\alpha$ ) and the letters of the space group and Pearson symbols are set in italics; digits stay upright (IUCr convention). |
| Inversion axes with overbar | P -1 is set as $P\overline{1}$ – the overbar is a Word equation object, just as if typed by hand. Switched off, $P-1$ appears instead.                                                                                                                                                                                                 |
| Write a table caption       | A line above the table, following the template in the “Caption” field.                                                                                                                                                                                                                                                                 |
| Rule above/below            | A horizontal rule above the first and below the last row of the table.                                                                                                                                                                                                                                                                 |
| Alignment                   | The labels are set flush left, the values from the second column on are centred – in Word as well as in LaTeX.                                                                                                                                                                                                                         |
| Compound name as a formula  | Sets the name in the caption as a chemical formula: UF6 becomes UF <sub>6</sub> . Numbers that do not follow an element symbol are left alone – so names such as 10 or 1a are too.                                                                                                                                                     |
| All data sets in one table  | Instead of one table per data set, a single table: column 1 holds the quantities, then one column per data set with the compound name as its header.                                                                                                                                                                                   |

| Option                                                       | Effect                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|--------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                                              | Quantities only some data sets have (the Flack parameter) stay empty in the other columns.                                                                                                                                                                                                                                                                                                                                                |
| Each table on a new page                                     | A page break before every further table.                                                                                                                                                                                                                                                                                                                                                                                                  |
| Residual density: distance to the nearest atom as a footnote | Adds superscript letters to the values in the row “Largest diff. peak/hole” and writes the distances to the nearest atom under the table (see below).                                                                                                                                                                                                                                                                                     |
| Pearson symbol                                               | Do not print, with H atoms, without H atoms, or both variants. With “both” they share one row, separated by “,” ( <i>mP222</i> , <i>mP82</i> ). The row is dropped if the formula, Z or the crystal system is missing.                                                                                                                                                                                                                    |
| Table language                                               | German, American or British English. This switches the labels in the first column, the caption template, the crystal system (monoklin/monoclinic), the type of absorption correction (numerisch/numerical), the crystal shape and colour (schwarzer Block/black block), “Nr.”/“No.”, <i>R</i> indices/indexes, colour/color, the language tag of the Word file (spell checking) – and the user interface itself. Numbers are not touched. |
| Caption                                                      | Template with the placeholders %NR% (table number) and %NAME% (compound). In the combined table the names are joined the way the table language does it: “1, 2, and 3” in English, „1, 2 und 3“ in German. Everything up to the first full stop or colon, plus %NAME%, is set in bold – with several data sets only the names, not the separators between them.                                                                           |
| Output format                                                | Word (.docx) or LaTeX (.tex.txt). Switching it carries the extension of the output file along.                                                                                                                                                                                                                                                                                                                                            |
| Font / size                                                  | Times New Roman, 10 pt by default. The size is given in whole points. The width of the first column follows the longest label and the font size; with a smaller font it gets narrower and the space freed goes to the value columns.                                                                                                                                                                                                      |
| First table number                                           | Number of the first table; the rest follow consecutively.                                                                                                                                                                                                                                                                                                                                                                                 |
| max. columns                                                 | Largest number of value columns per table (3 by default). If more data sets are selected, a further table follows with its own number and caption. All partial tables get the same column widths so that they line up underneath each other.                                                                                                                                                                                              |

A minus sign in front of a number is always a true minus sign (U+2212) – in exponents ( $\text{g}\cdot\text{mol}^{-1}$ ,  $\text{g}\cdot\text{cm}^{-3}$ ,  $\text{mm}^{-1}$ ,  $\text{e}\cdot\text{\AA}^{-3}$ ), in index ranges, in residual electron densities and in the Flack parameter. Hyphens inside words (multi-scan, goodness-of-fit) are left alone.

## Language of the user interface

The “Table language” setting also switches the user interface: window title, labels, button captions, drop-down lists, column headers, the file dialogs, the “Add” window, the status line and the messages in the log. Choosing “German” takes everything back to German. American and British English differ only in the table, not in the interface. Log lines already written stay in the language they were written in.

## Settings

All options, the row order, changed labels and the size and position of the window are saved when the program is closed and loaded again at the next start. They live in `crystal2table.ini` next to the program; if that folder is read-only, in the user profile under `AppData\Local\crystal2table`.

The order of the data sets is not saved – it belongs to the file currently loaded.

## Absorption correction

In a German table the entries from `_exptl_absorpt_correction_type` are translated where German uses a different word: numerical becomes numerisch, integration Integration, analytical analytisch, empirical empirisch, semi-empirical semi-empirisch, none keine, sphere Kugel and cylinder Zylinder. Upper and lower case make no difference.

multi-scan, psi-scan and gaussian stay as they are, as does any entry that does not match one of the translated words exactly – a whole phrase such as Semi-empirical from equivalents, for instance. In an English table the value always stays as the CIF has it.

## Crystal habit: shape and colour

In a German table the shape and the colour of the crystal are translated – in the rows “Crystal morphology” and “Colour” and together in “Crystal habit”. The shapes: block becomes Block, plate Platte, needle Nadel, plank Planke, prism Prisma and rod Stäbchen.

The colour comes first and takes the ending the shape asks for: black block becomes “schwarzer Block”, black plate “schwarze Platte”, black prism “schwarzes Prisma”. With no shape behind it – in the row “Colour”, or when the CIF gives a colour only – it stays uninflected.

Translated are colourless farblos, white weiß, black schwarz, grey grau, brown braun, red rot, pink rosa, orange orange, yellow gelb, green grün, blue blau, violet violett, metallic metallisch, lustrous glänzend and clear klar. The qualifiers light, dark, pale, deep, dull and the forms in -ish merge with the colour into one word: dark red block gives “dunkelroter Block”, pale yellow needle “blassgelbe Nadel”.

clear describes the crystal itself and therefore takes an ending of its own: clear yellow block gives “klarer gelber Block”, clear yellow needle “klare gelbe Nadel”. metallic and lustrous qualify the colour instead and stay uninflected: “metallisch schwarzer Block”.

The translation only happens when every part is known. If the colour contains a word that is not in the list (mauve), or the shape is none of the six (shard, the plural blocks, thin needle), the whole entry stays as the CIF has it – better English than half translated.

## Footnote on the residual electron density

The option **Residual density: distance to the nearest atom as a footnote** adds superscript letters to the values in the row “Largest diff. peak/hole” and writes under the table how far the peak lies from the nearest atom, for example “a) 0.64 Å from N3; b) 0.88 Å from Ga”. The letters run through the whole table – in a combined table that means column by column.

If a **SHELXL listing file** (.lst) sits next to the input file, peak and hole are taken from there; SHELXL has already worked out the distance. Otherwise the highest Q peak of the res file is used and its

distance to the nearest atom is computed here – with every symmetry operation, the centring and the lattice translations. The deepest hole is not in the res file, so it gets no footnote.

Nearest atom means the nearest one of all, hydrogen included – just as in the SHELXL listing. Without a listing or a res file the row stays as it would be without the option.

## Extinction and twin fraction from the listing file

When a structure is loaded from an HTML report and its res file, the extinction coefficient (EXTI) and the twin fraction (BASF) appear there as bare numbers without a standard uncertainty. If the **SHELXL listing file** (.lst) sits next to the input, the uncertainty is taken from there: after every refinement cycle SHELXL lists the parameters that belong to no atom, each with its uncertainty. 0.06678 and 0.00127 become 0.0668(13) – the same notation SHELXL writes into its own CIF.

Only the parenthesis comes from the listing: the value itself stays the one from the report or the res file, and a row that would not be there without the listing never appears. An uncertainty already present in the table file or in the CIF is left alone. Nothing changes either when the input file holds several data blocks, when the listing does not match the state of the res file, or when SHELXL could not determine one of the uncertainties – after a CGLS refinement, for instance, the listing holds no such table at all.

## Batch mode (command line)

If --cif= is given, no window appears: the program writes the table file – .docx or .tex.txt, whichever is set – and exits. Every data block that holds structure data is written.

```
crystal2table.exe "--cif=One.cif;Two.cif" --out=Tables.docx --start=2 --onetable
```

In PowerShell and Bash the semicolon separates commands, so the argument with several files must be quoted (not needed in cmd.exe).

The starting point is the settings saved from the user interface, row order and custom labels included; command line switches take precedence. So --lang=us changes only the language and leaves everything else as set.

| Switch                             | Meaning                                                                                              |
|------------------------------------|------------------------------------------------------------------------------------------------------|
| --cif=File.cif                     | input file (.cif or .html); several separated by semicolons (required, and what triggers batch mode) |
| --out=File.docx                    | output file; without it, the name of the first input file with the extension .docx or .tex.txt       |
| --onetable                         | all data sets in one combined table                                                                  |
| --maxcols=3                        | largest number of value columns per table (1 to 99)                                                  |
| --start=1                          | number of the first table (1 to 9999)                                                                |
| --pearson=withh withouth both none | Pearson symbol; --no-pearson is the same as none                                                     |
| --lang=de us gb                    | table language                                                                                       |
| --latex, --word                    | output format; without it the setting from the user interface applies                                |
| --font=Arial                       | font                                                                                                 |
| --size=10                          | font size in whole points (4 to 72)                                                                  |

| Switch           | Meaning                                                                                                      |
|------------------|--------------------------------------------------------------------------------------------------------------|
| --pagebreak      | page break before every further table                                                                        |
| --peaknotes      | footnote with the distance of the largest residual density to the nearest atom; --no-peaknotes leaves it out |
| --no-italic      | no italics for symbols                                                                                       |
| --no-overbar     | minus sign instead of an overbar                                                                             |
| --no-caption     | no table caption                                                                                             |
| --no-nameformula | do not set the name as a chemical formula                                                                    |
| --no-ini         | ignore the saved settings and use the built-in defaults only                                                 |
| --help           | prints an overview of all switches and exits; also --?, -h and /?                                            |

A folder named by --out is created if it does not exist; if --out ends with a path separator it is taken as a folder and the file name comes from the input. The log takes the name of the output file with the extension .log (Tables.docx becomes Tables.log); its first line says which settings file was used. Return codes: 0 fine, 1 no --cif, 2 input file not readable, 3 no usable data block, 4 output file not writable, 5 unexpected error. If even the log cannot be written, the message goes to the error stream (2> errors.txt). A value outside the permitted range, an unknown switch and a typing error are noted in the log; the default is used instead.

## Licence

Crystal2Table is published under the BSD 3-Clause Licence. Copyright (c) 2026, Gregor Schnakenburg. The full licence text comes with the program as the file LICENSE; it allows redistribution and modification, but requires the copyright notice, the licence text and the disclaimer to be kept.

The program contains statically linked parts of the Lazarus LCL and of the Free Pascal run-time library. Both are under the LGPL with a linking exception that expressly allows the finished program to be distributed under terms of your choice. The file THIRD-PARTY-NOTICES.md supplied with the program gives the details.

The program comes with no warranty of any kind. Check the tables it produces against the original data before publishing them.