Structure Deposition Workshop (DEP-001)

Mar 2023

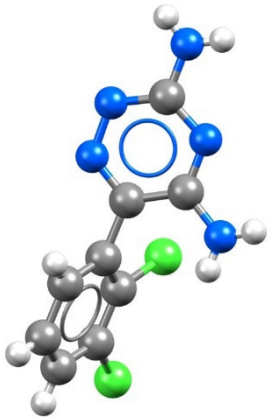


Table of Contents

Introduction

Structure deposition to the CCDC can be conducted through our CIF deposition and validation service. The web services enable you to submit CIF and associated structure factor files to the joint CCDC and FIZ Karlsruhe deposition service. During the deposition process, you can correct syntax errors, check the integrity and novelty of your data, and add metadata. You can also generate checkCIF reports and 2D chemical diagrams that can be downloaded at the end of the deposition process.

Learning Outcomes

This tutorial will guide you through the use of the CIF deposition and validation process. At the end of this tutorial, you will be able to:

- Deposit CIFs.
- Generate validation and checkCIF reports.
- Enhance your crystallographic data.

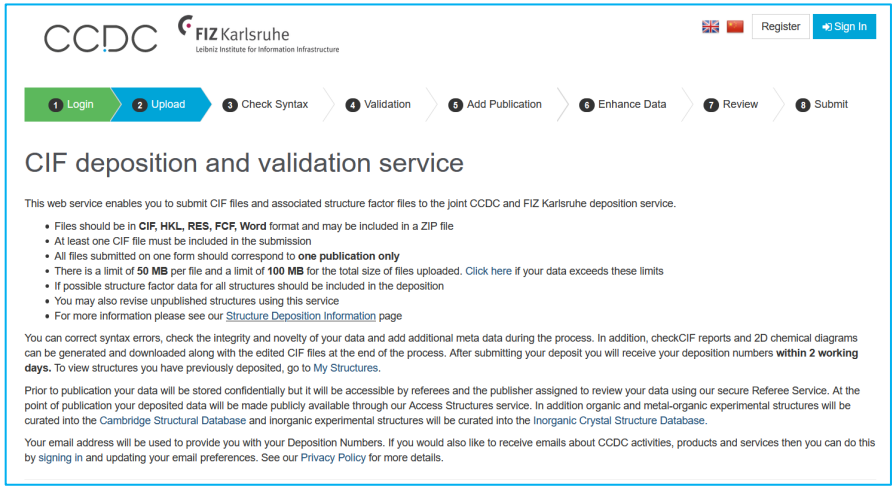
It will take approximately **25 minutes** to complete deposition of the first example file. If you decide to complete the further exercises, the total workshop time will be approximately **40 minutes**. The words in *Blue Italic* in the text are reported in the Glossary at the end of this handout.

Pre-required Skills

There are no pre-required skills for this workshop.

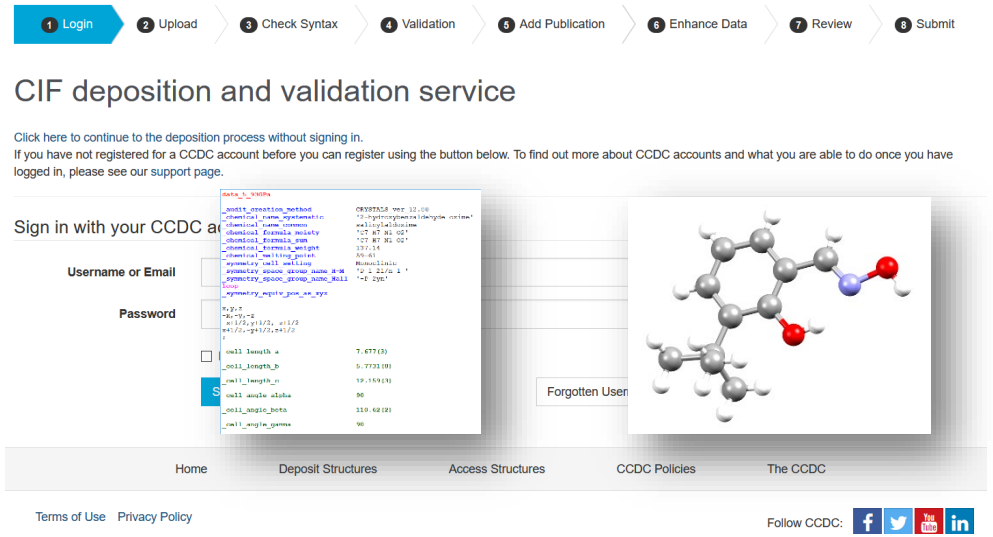
Materials

In this workshop, you will use two example CIF files; one with correct syntax, and one where there are some syntax errors which need to be corrected during the deposition process. These files can be obtained from our website. You will need a CCDC account to download the files. The instructions to create an account and to download the materials are part of the workshop.



Depositing CIF files

In the following example, you will learn how to register for a free account for the CCDC website. This will allow you to deposit data more easily and to access and share your data using the MyStructures service. You will then look at depositing crystal data with the CCDC in the CIF file format. To do this, you will use the joint CCDC and FIZ Karlsruhe web deposition service.



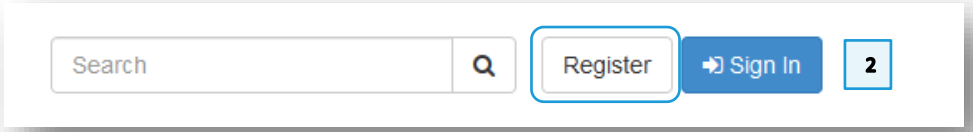
Registering for a CCDC account

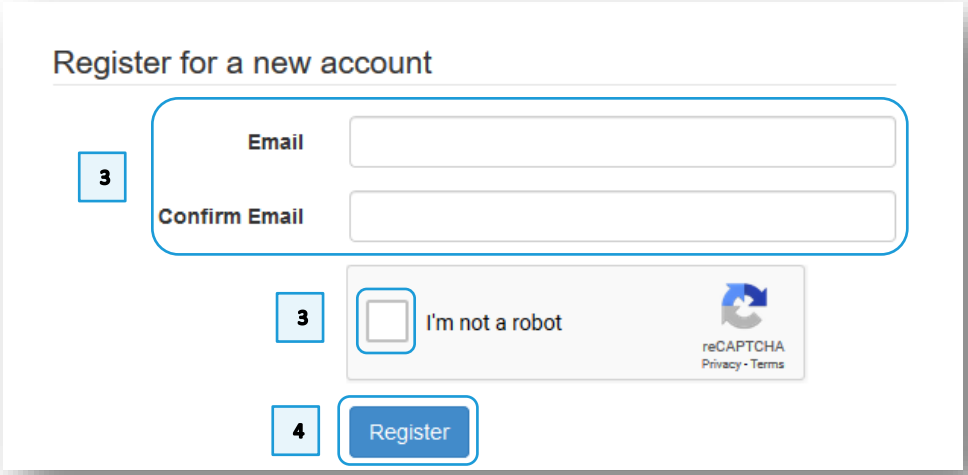
1. Open a web browser and navigate to https://www.ccdc.cam.ac.uk/deposit.

2. We recommend creating a CCDC account, which makes depositing structures easier and allows you to access and share your data through our MyStructures service. If you need to create an account, click on **Register** at the top right of the page. Alternatively, if you already have an account, you can click the link to continue the deposition process and proceed to the section of the workshop 'Obtaining example CIF files'.

3. Enter your email address in the boxes and click the reCAPTCHA to indicate you are not a robot. On this page you may also view our privacy policy before setting up an account.

4. Click **Register** in the blue box at the bottom of the page.





Redeeming an invitation

5. Check your inbox for an email containing your invitation.

6. Click on the **Redeem invitation** tab and enter the invitation code you received via email.

7. Tick the reCAPTCHA box to indicate you are still not a robot.

8. Click **Register**.

Obtaining example CIF files

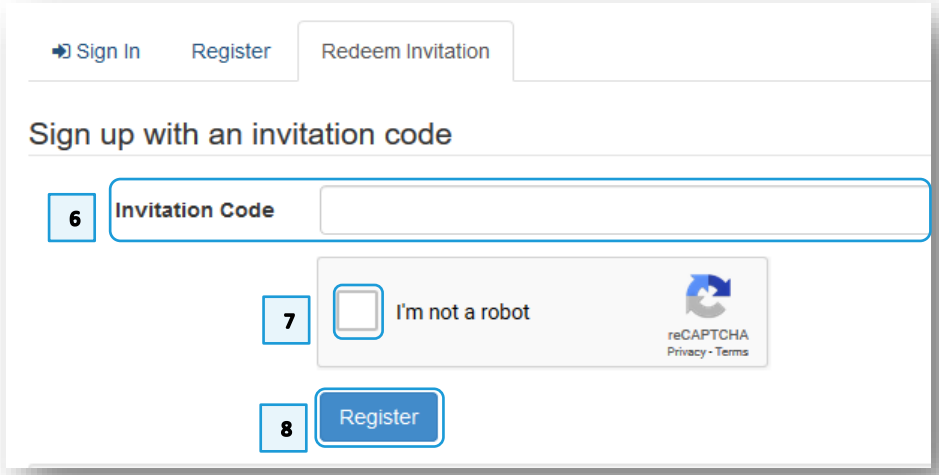
9. Open a web browser and navigate to https://www.ccdc.cam.ac.uk/support-and-resources/downloads/

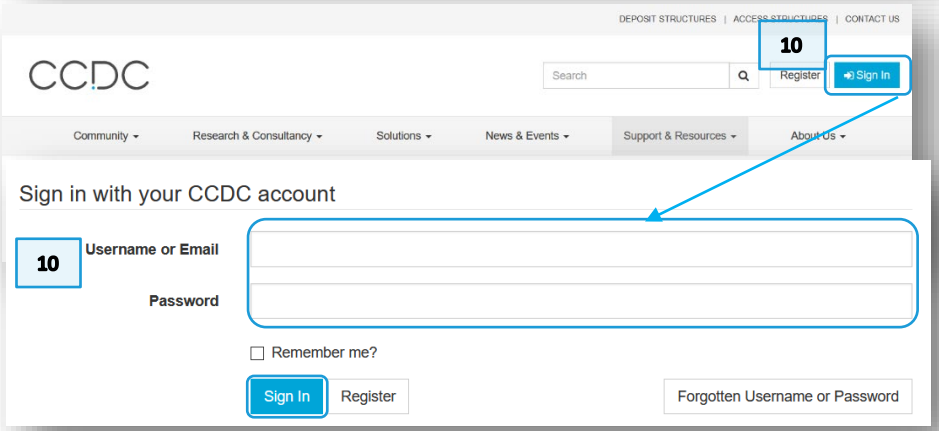
10. A CCDC account is required for downloading the example files. On the downloads page, click on **Sign In**. Enter your login details and click Sign In to return to the Downloads page.

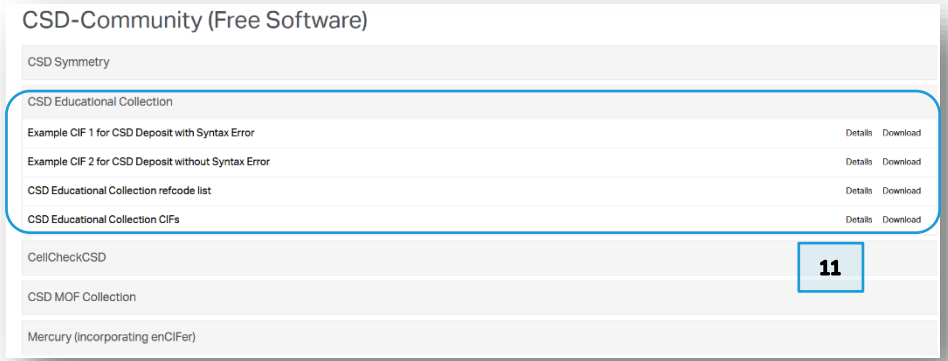
11. Under *CSD-Community (Free Software)*, click on **CSD Educational Collection** and click **Download** beside *Example CIF 1 for CSD Deposit with Syntax Error*.

12. Click **Agree & Download** to accept the agreement displayed and save the "example_1.cif" file.

13. Repeat **Steps 11** and **12** for *Example CIF 2 for CSD Deposit without Syntax Error* to download the "example_2.cif" file. Remember the location where you save these files.







Personal details and CIF upload

1. Go to the web deposition service at https://www.ccdc.cam.ac.uk/deposit

2. You should still be logged in but if you are not, click Sign in, fill in your Username or Email and Password and click on **Sign In**.

3. The details you entered when setting up a CCDC account should appear automatically in the appropriate fields. The fields marked with '*' are mandatory. You can also link your *ORCID iD* here if you have one. Hover your mouse on the "?" icon for more information on each field.

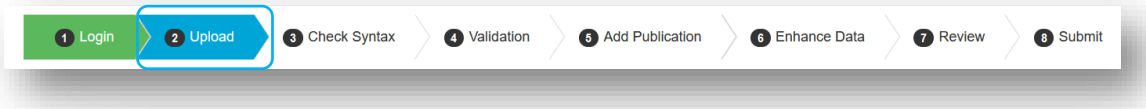
4. In the *More Information* section, you can enter any information here but for this workshop, type "This is a test".

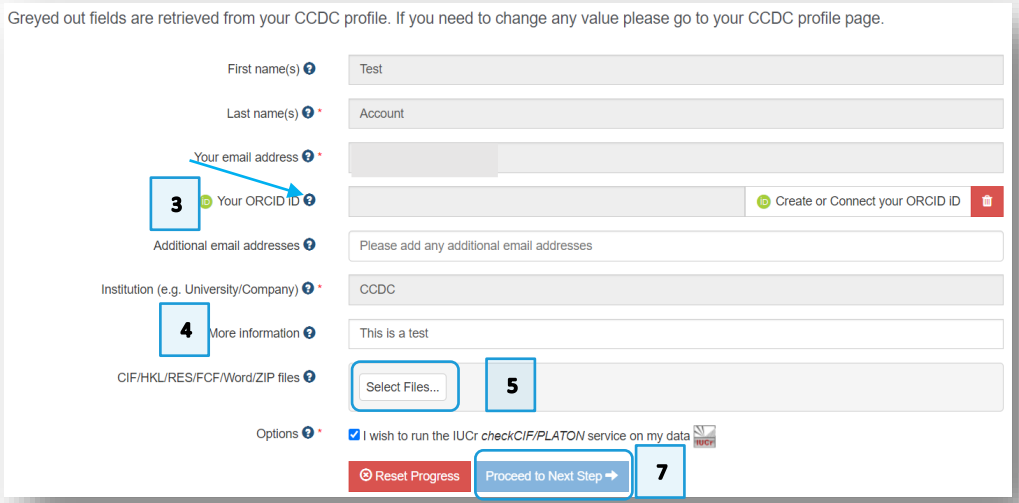
5. To add a *CIF file* to be uploaded, click on **Select Files…** and select "example_1.cif" and click **Open**. If the file is not in this folder, navigate to the appropriate folder.

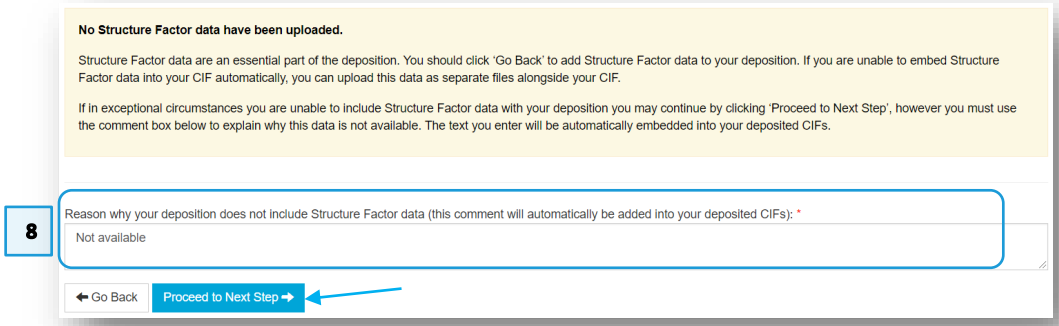
6. In order to run the IUCr *checkCIF* program on your CIF file to check for syntax and other errors, ensure the box is ticked for the option "I wish to run the IUCr checkCIF/PLATON service on my data".

7. Click **Proceed to Next Step**.

8. A warning message should appear to advise you that *No Structure Factor data have been uploaded*. It is important to include *structure factor* data in the deposition if available (read here why). The CCDC accepts structure factor data in *.fcf* or *.hkl* format. In this example, however, there are no structure factors included. In the text box enter "Not available – This is a test" and click **Proceed to Next Step**. The reason provided for the absence of structure factors will be included in the deposited CIF.





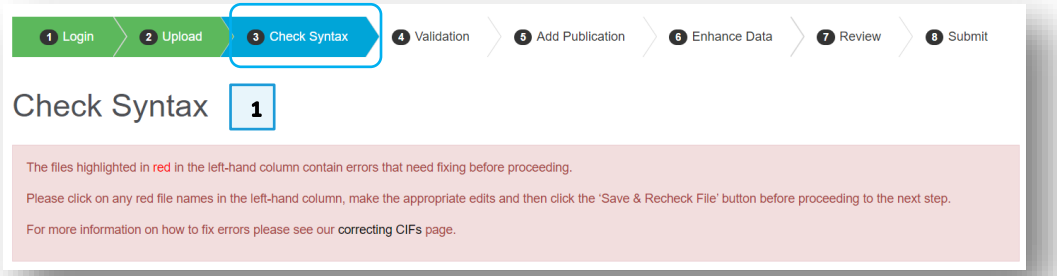


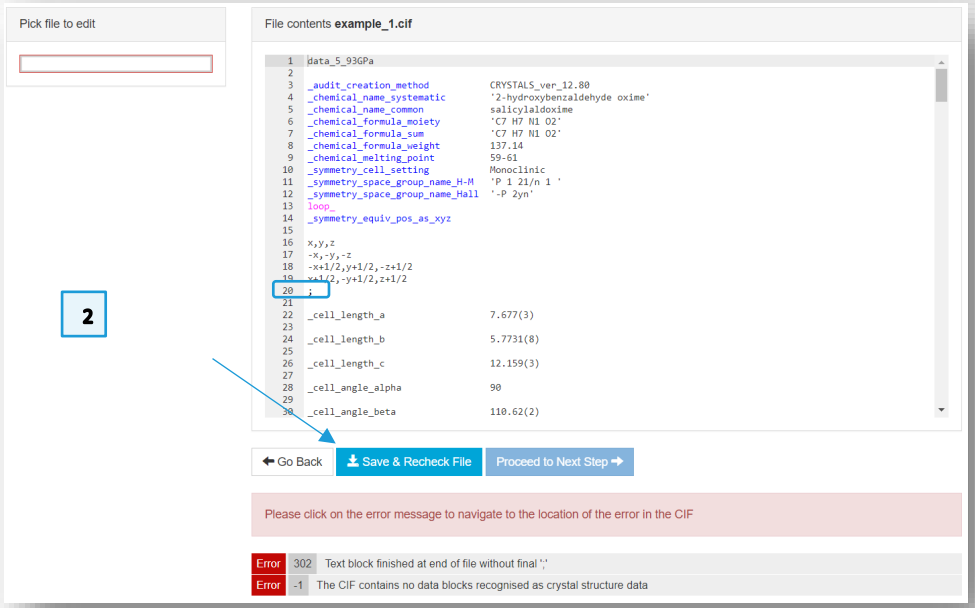
Syntax checking

The CIF format has a specific structure, or syntax, that needs to be followed to allow the file to be read by crystallographic programs. This stage in the deposition process checks the CIFs for any potential errors.

1. You will see a warning in red at the top of the page indicating there are syntax errors with your CIF file which need to be corrected before CheckCIF can be run. Descriptions of the errors are shown at the bottom of the page. In this case there is an extra semi-colon in the file. This is indicated by the error **'Text block finished at end of file without final ';''**. Semi colons are used in CIFs to enclose longer data responses that span multiple lines in the CIF. The CIF file display is colour-coded, which can help you find the location of the error.

2. Near the top of the CIF you will notice that the colour coding changes for the unit cell values. Directly above this section, on line 20, there is a semi-colon that is out of place. Delete the semicolon on line 20 and click **Save & Recheck File**. Once you delete the semi-colon, you see the colour-coding re-appears (blue). This should result in no syntax errors in the file and you can click **Proceed to Next Step**.





Validation and CheckCIF reports

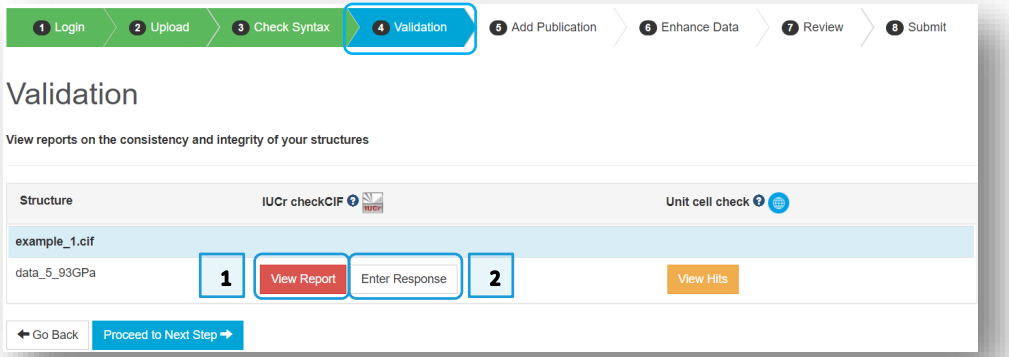
IUCr checkCIF/PLATON is a service which allows CIF files to be checked for problems. The responses are reported with a Level A, B, C or G (from most serious to least) and a type depending on what kind of alert it is. You can find more information on correcting CIFs here and other resources linked from the page.

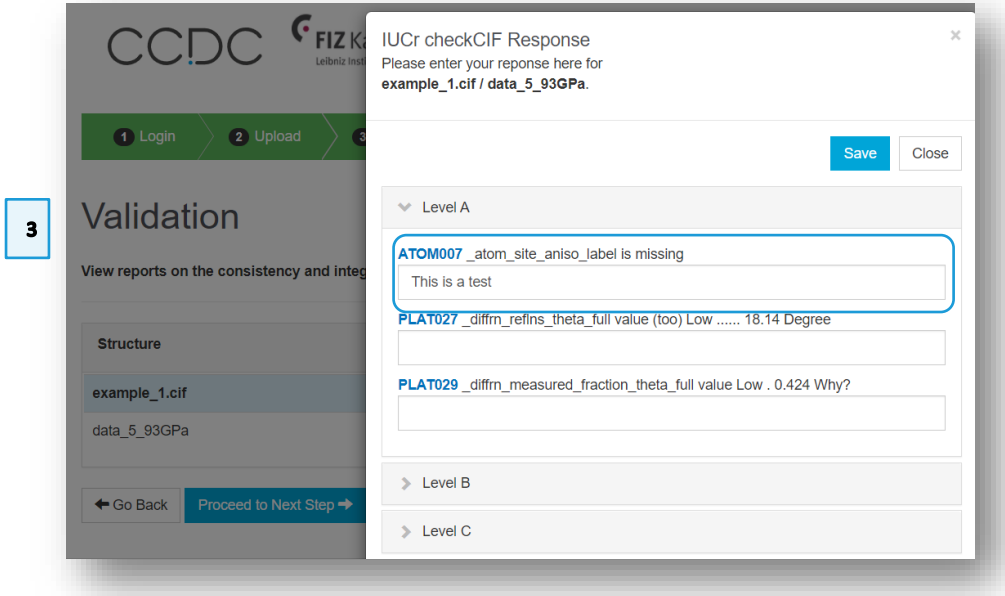
1. The IUCr checkCIF/PLATON service should run, and you should be able to click **View Report** to see a report on the consistency and integrity of the structure.

2. Any responses you may wish to provide to the Level A, B or C alerts found by the checkCIF report can be provided by clicking **Enter Response** on the main page. These responses will be added to the deposited CIF and may also be downloaded along with the checkCIF report at the end of the deposition process. The alert titles are hyperlinked to the IUCr website and will give explanations about the nature of the error and possible solutions.

3. For this example, we will address the first of the A-level alerts. In practice you should first try to resolve the alerts by fixing the underlying problems with the structure refinement and upload a revised CIF file. If it is not possible to fix the issues, then enter a response to be included in the deposited CIF. For the purposes of this workshop, for the level A alert "ATOM007 _atom_site_aniso_label is missing", enter the response "This is a test" and click **Save**.

4. Click **Proceed to Next Step**.





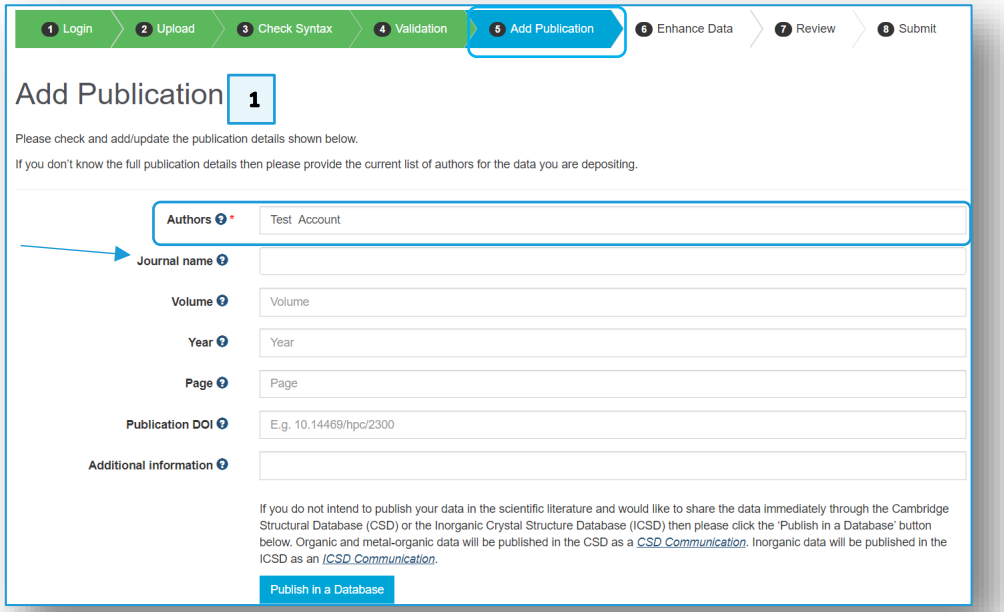
Adding publication details

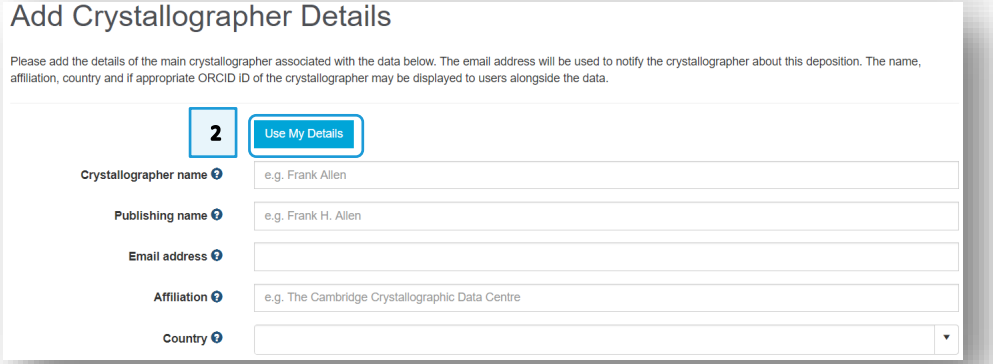
1. Your name should be automatically entered in the *Authors* field. Further publication details can also be entered on this page; namely *Journal name*, *volume*, *year*, *page* and *DOI* if applicable, as well as any *additional information* you may wish to include.
   The author's names should be added in the same order as on the publication. Each name should be separated by a comma and space, without any additional characters or salutation. You can use initials (no spacing between them) or expand the first name; if you are using initials, then you do not need spaces between them.
   For this example, change the *Authors* field to "P.A.Wood, R.S.Forgan, D.Henderson, S.Parsons, E.Pidcock, P.A.Tasker, J.E.Warren" and enter "Acta Crystallographica, Section B: Structural Science [1983-2012]" in the *Journal name* field.

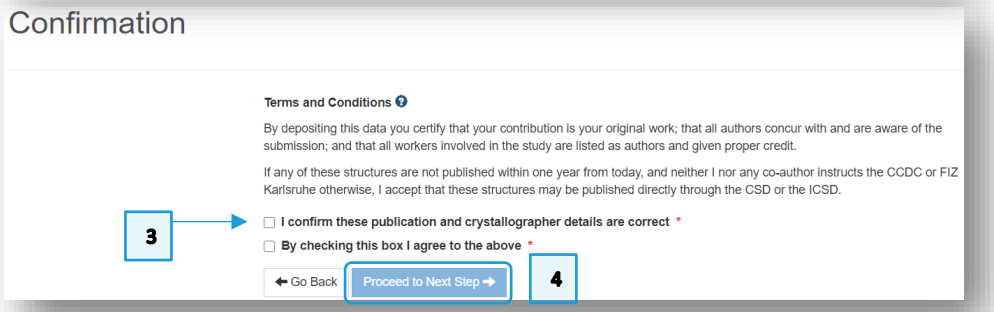
2. The crystallographer's details are required and should be entered in the *Add Crystallographer Details* section. If the crystallographer is not included in the list of authors for a publication, their details can be included here as a way for their contribution to be recognised when the data are available in the database. If you are the crystallographer, click the **Use My Details** button to automatically fill in the details from your profile. For this example, click the Use My Details button. If nothing is filled in, click Use Different Details and enter your information.

3. Tick the two boxes in the *Confirmation* section to confirm the details are correct and that you accept the Terms and Conditions.

4. Click **Proceed to Next Step**.







Enhancing your data

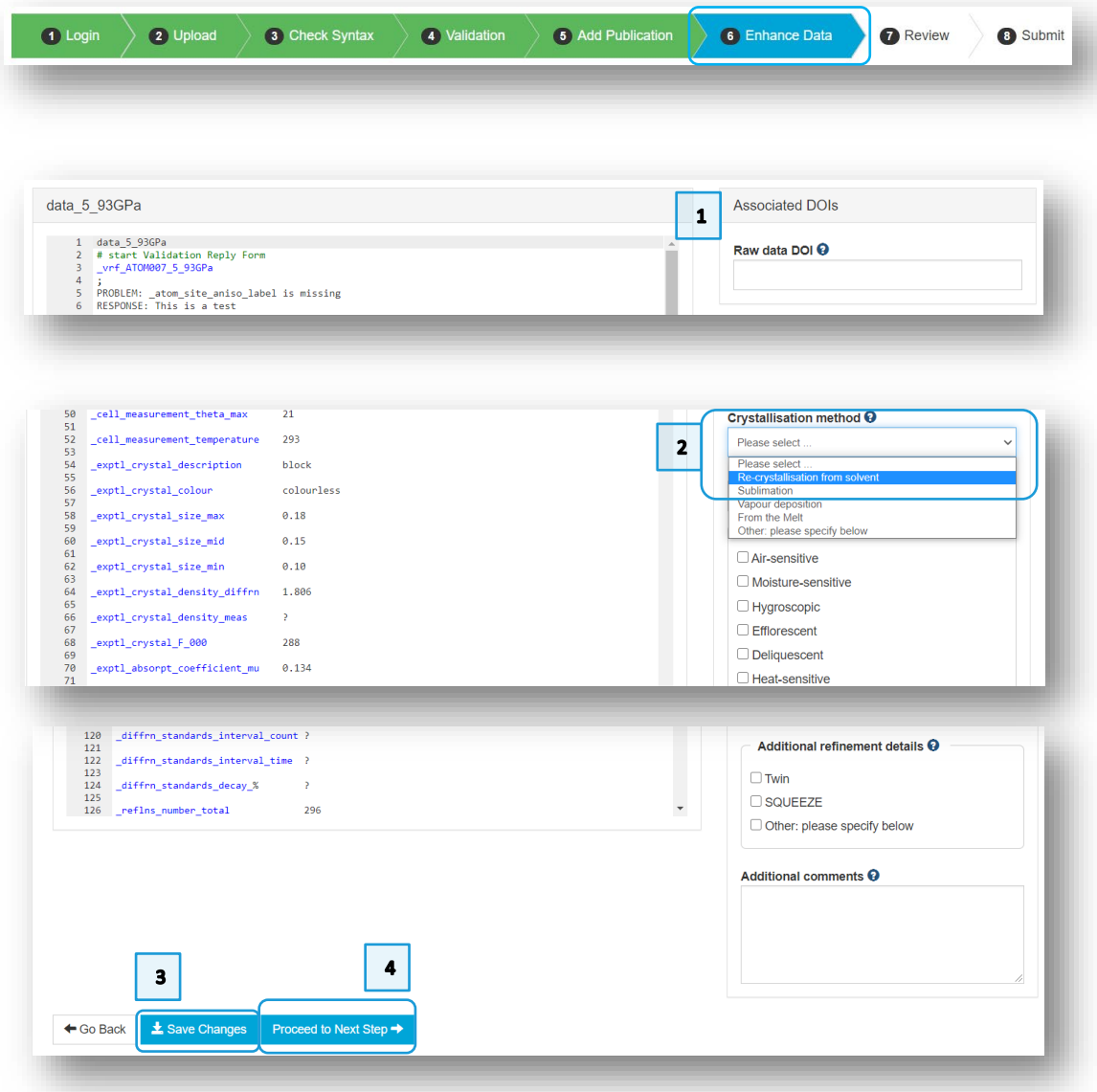
The Enhance Data page allows you to preview how the structure(s) to be deposited will look as a 2D chemical diagram and as viewed in 3D, as well as add extra information to the CIF file prior to deposition. The data fields on the right of the page can be edited and enhanced, the changes saved, and the modifications to the CIF can be seen in the pane on the left.

1. Other information can be added to the structure, such as a DOI link to deposited raw diffraction data files. These are the files measured during experimental data collection. The CCDC does not currently store *raw data* files. However, if you have deposited your raw data files with another repository and have a DOI link for this then you can associate this DOI link with the relevant files you have deposited with CCDC. You can do this during the deposition process by adding the DOI at the 'Enhance Data' stage of the process, found under 'Associated DOIs'. We will leave this blank for this workshop.

2. From the *Crystallisation method* drop-down menu, select "Re-crystallisation from solvent".

3. Click **Save Changes**.

4. Click **Proceed to Next Step**.



Reviewing and submitting your data

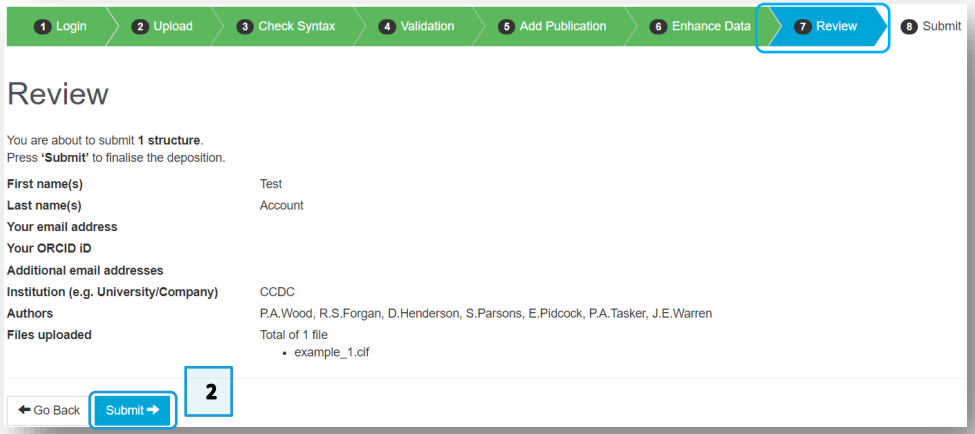
A final check on the data before deposition

1. Check the details on the Review page, where you can click **Go Back** if you want to make any changes.

2. Click **Submit**.

3. A message should appear to confirm your file has been deposited and you will receive deposition numbers by email. The buttons at the bottom of the page then allow you to either start the process again to deposit more CIF files, view your structures in the *My Structures* section of your profile, or download a copy of the data as deposited.

4. Click **Retrieve Deposited Files**; this will allow you to download a zip file containing the deposited CIF, a html-format checkCIF report generated during the deposition process, and the automatically generated 2D diagram as a .png and .mol file.



How to cite your own data in a publication

To reference your data in the corresponding publication, you should do so by reporting the *CCDC number* (the 6 or 7-digit number that would be assigned to the structure by CCDC after deposition).

Conclusion

In conclusion, you have learned how to deposit a CIF using our online deposition service, correct CIF syntax, generate a checkCIF report and enhance the data in the deposited CIF file in the process.

Exercises

- Click on **Deposit More Structures**. Repeat the exercise using example_1.cif up until you reach the *Validation* page. Click **View Hits** to examine the results of the Unit cell check. The results should include the example structure itself, SALOXM09, as well as a structure from the ICSD.

- Repeat the worked example, but instead deposit example_2.cif.

Summary

This workshop introduced the process of depositing your structure to the CSD. You should now be familiar with:

- Depositing CIF files.
- Generating validation and checkCIF reports.
- Enhancing your crystallographic data.

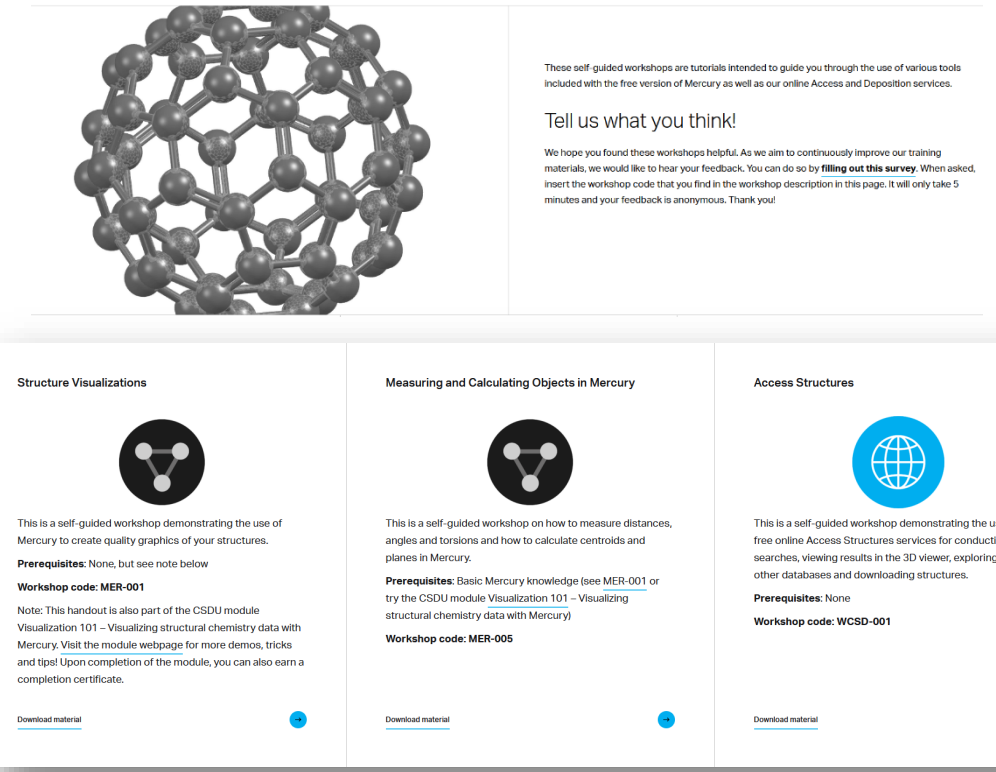
Next Steps

Other CSD-Community self-guided workshops can be found here (https://www.ccdc.cam.ac.uk/community/training-and-learning/workshop-materials/csd-community-workshops/), including an introduction to *CSD Communications* and how to publish them (available in English and Chinese).

For more information, visit our website: https://www.ccdc.cam.ac.uk/community/access-deposit-structures/deposit-a-structure/.

Feedback

We hope this workshop improved your understanding of the Deposition process and you found it useful for your work. As we aim to continuously improve our training materials, we would love to get your feedback. Click on this link to a survey, it will take less than **5** minutes to complete. The feedback is anonymous. You will be asked to insert the workshop code, which for this self-guided workshop is DEP-001. Thank you!



Glossary

CCDC number
A 6-7-digit deposition number provided by CCDC upon structure deposition. A CCDC number can then been included in associated manuscripts and is often used to link articles to associated datasets.

checkCIF
A service run by the International Union of Crystallography (IUCr) that checks the consistency and integrity of CIFs. It is available to run during deposition and from the IUCr website. A list of the tests carried out and further information about what they mean can be found here.

CIF
A Crystallographic Information File. This is the standard file format for crystallographic data. A list of CIF dictionary provided by the IUCr can be found on the IUCr website.

DOI
The digital object identifier (DOI) is a unique string of numbers, letters, or symbols used to identify objects online. The CCDC uses DOIs to provide links to the data, e.g. DOI: 10.5517/ccspp8d or the associated publication DOI: 10.1107/S0021889809008450. More information on DOIs can be found here.

ORCID iD
An Open Research and Contributor ID is a unique, persistent identifier for individuals to use as they engage in research, scholarship, and innovation activities. More information can be found at https://orcid.org/.

Raw data
Raw data refers to the data collected during a crystallographic experiment by measuring the diffraction pattern of a crystal using a diffractometer. They may also be called diffraction images.

Structure factors
Structure factors are created from experimental crystallographic data during the structural solution process. The information can be used to describe the distribution of electron density in the structure. Currently there are two types of information that CCDC accepts as 'structure factor' information, these are the structure factors themselves (often found in a .fcf file) and the reflection intensities (.hkl). They can also be appended to the CIF file, as is the default in many crystallographic programs. See more on our website at https://www.ccdc.cam.ac.uk/Community/depositastructure/cif-deposition-guidelines/structure-factor/.