Deposit your crystallographic data in the CSD: how, when, why, and what!

CCDC Virtual Workshop Autumn 2021 – Session 1

Suzanna Ward, Vij Menon, Natalie Johnson, Clare Tovee, Ilaria Gimondi

2nd November 2021

CCDC
advancing structural science

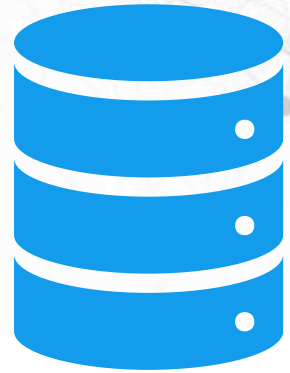
Learning outcomes for the session

- Become familiar with the CSD.
- Know which data and which data files you can submit to the CSD.
- Learn how to deposit data to the CSD.
- Know how to check the novelty and integrity of your data.
- Learn how you can enhance your deposited data.
- Discover what happens to your data after deposition and how you can manage them.
- What the benefits of sharing your structural data with the community are.

CCDC

What is a crystallographic database?

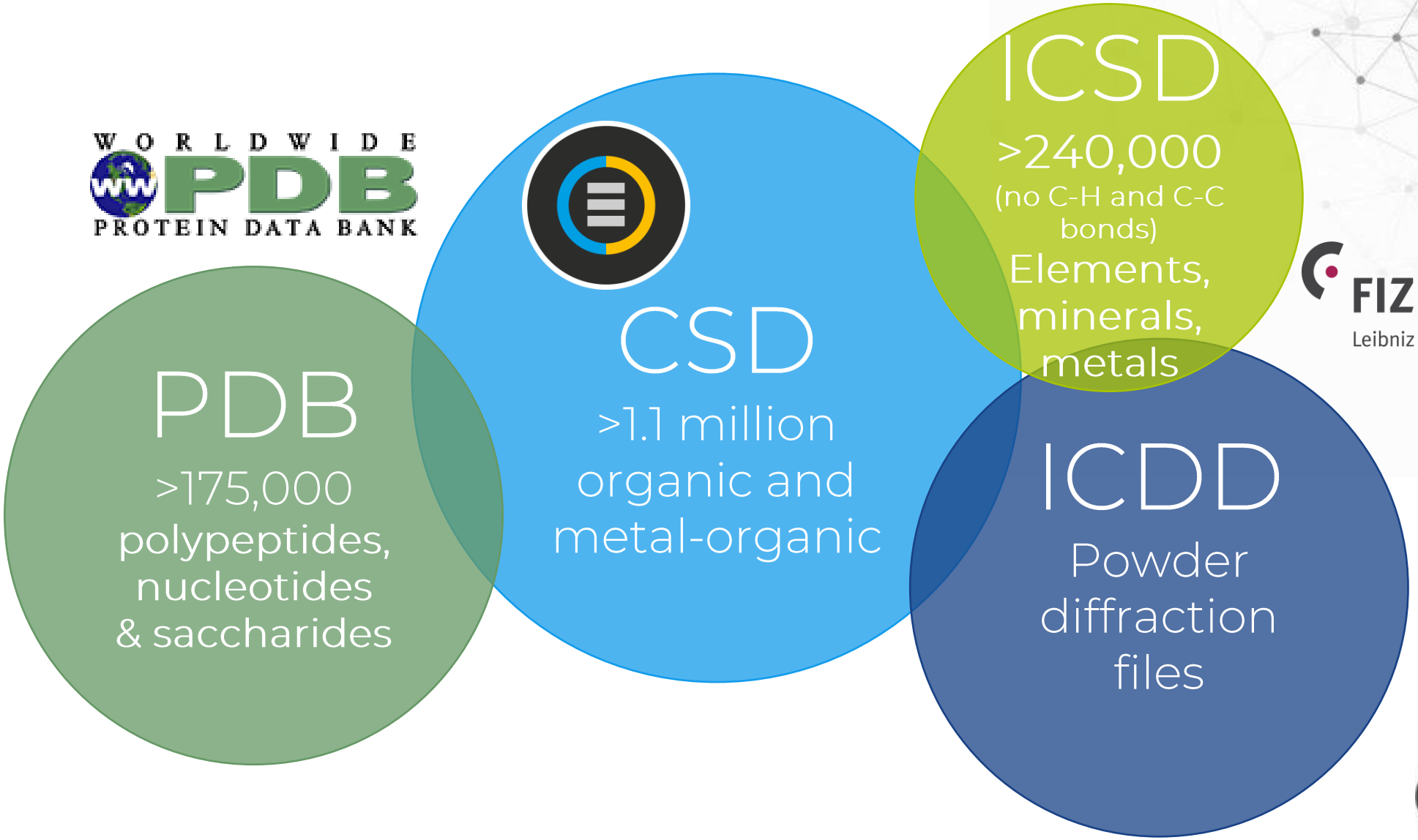
- They usually include:
  - Bibliographic data
  - Crystallographic, chemical and experimental information
  - Atomic coordinates



- What makes a crystallographic data special?
  - Standard agreed file format
  - Every published structure is added to the appropriate database
  - There are established curated databases rather than just collections of data with some level of quality control through processing & validation
  - They enable you to learn from the wealth of data they contain

CCDC

Structural databases

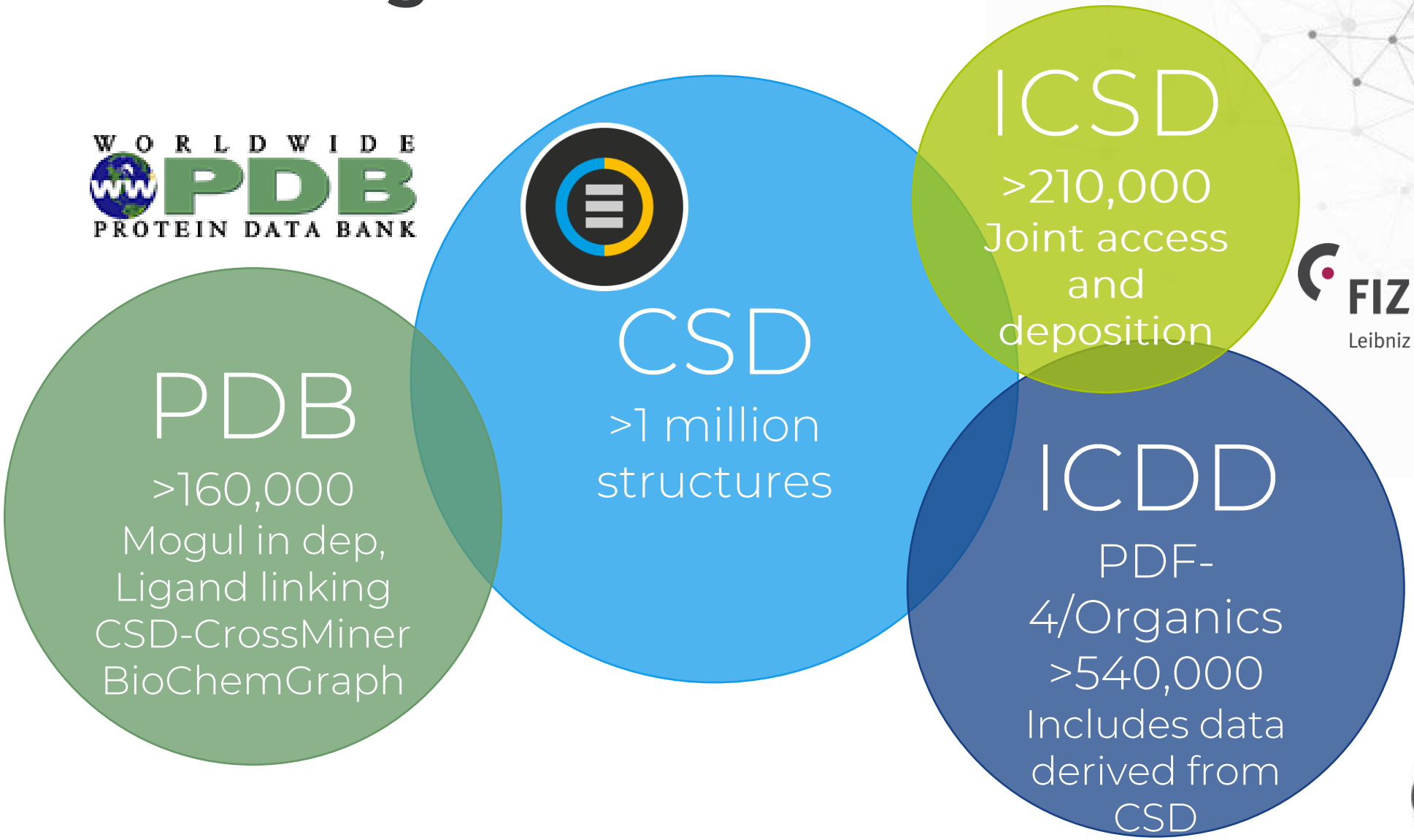


FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

ICDD

CCDC

More integrated structural databases



FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

ICDD

CCDC

The vision

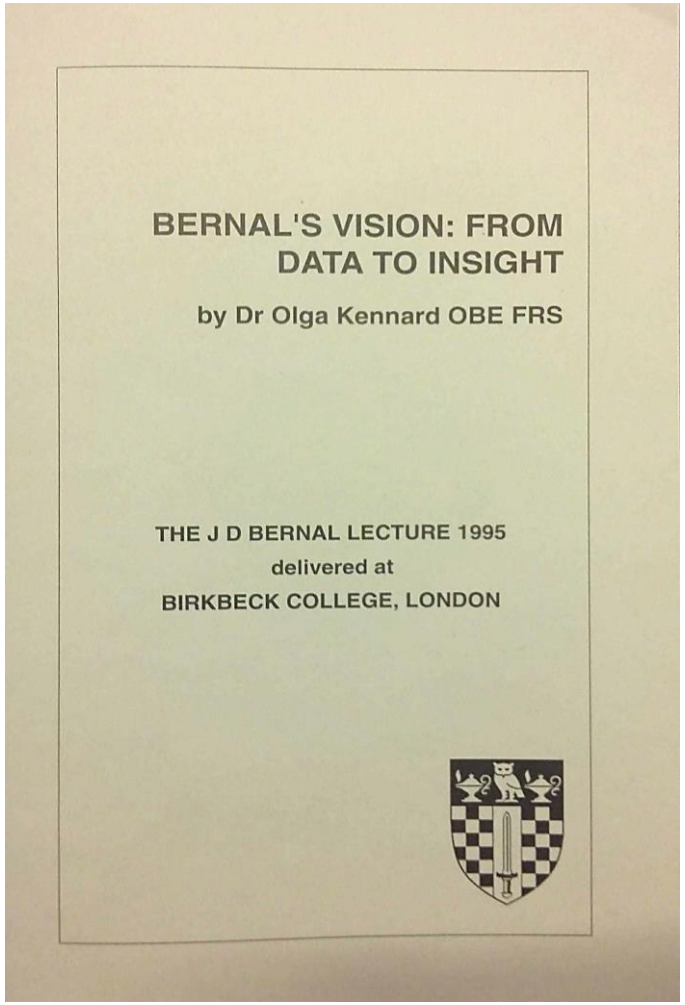
J.D. Bernal and research group including Olga Kennard at Stonehenge in 1948

- Established in 1965 by Olga Kennard
- Olga and J.D. Bernal had a vision that a collective use of data would lead to new knowledge and generate insights

Kennard, O. “From Private Data to Public Knowledge.” The Impact of Electronic Publishing on the Academic Community. Ed. I Butterworth. Portland Press Ltd, 1997. 159–166.

CCDC

The vision



BERNAL'S VISION: FROM DATA TO INSIGHT

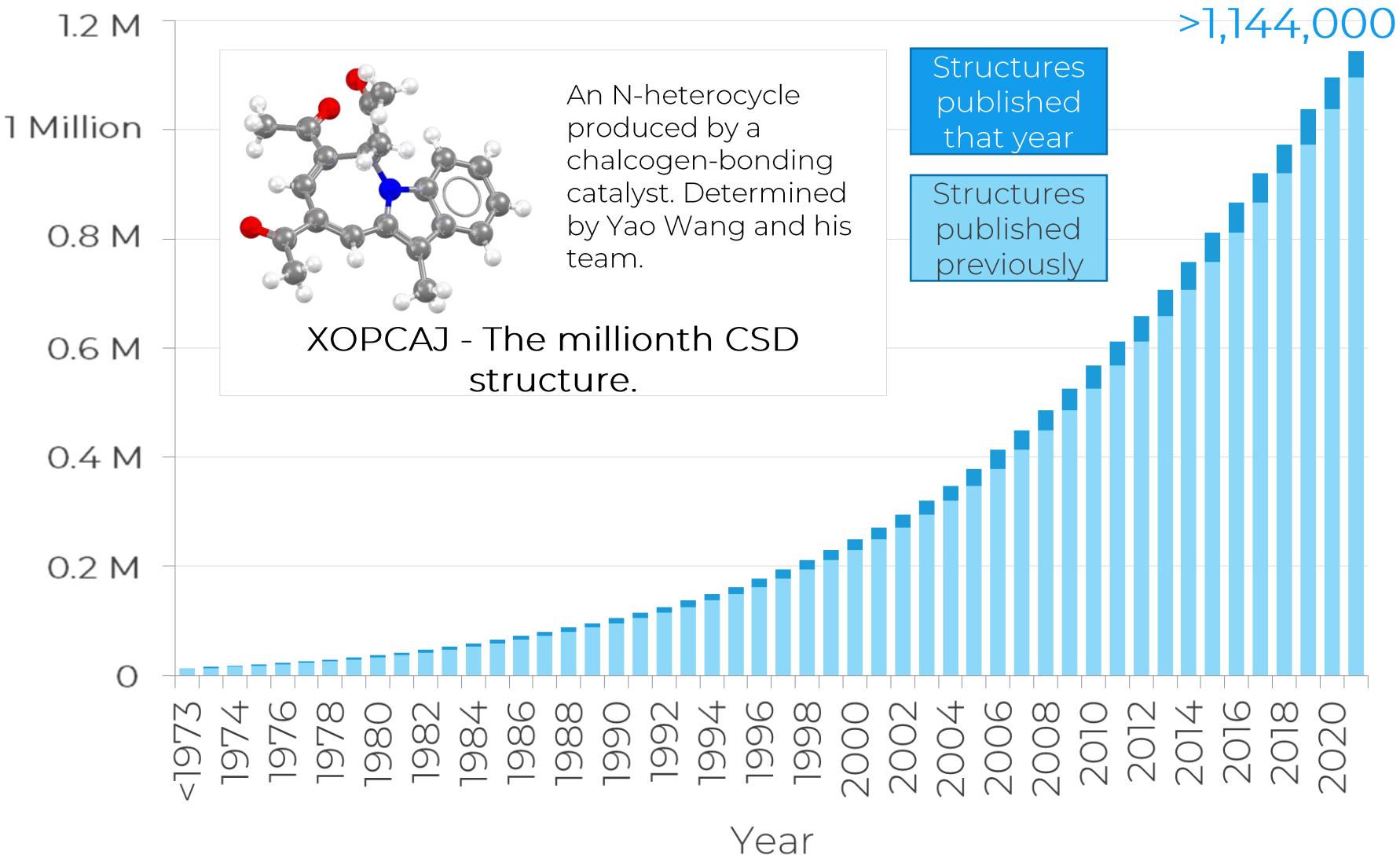
by Dr Olga Kennard OBE FRS

THE J D BERNAL LECTURE 1995
delivered at
BIRKBECK COLLEGE, LONDON

We clearly recognised even in those early days, that data banks have three principal functions. Firstly they must gather together existing knowledge and make it readily available to the scientific community. Secondly they can be used to reduce a large number of observations to a small set of constants and rules, and in this way transform a data base to a knowledge base. Such a knowledge base may obviate the need for further individual experiments in specific areas. Thirdly, they facilitate the comparison and collective analysis of individual results to gain insight into new or as yet unexplained phenomena. These ideas have been at the heart of the work of the Cambridge Crystallographic Data Centre and the driving force for improving methods of data collection, storage and dissemination. Most importantly they influenced development of computer programs and methodologies which are needed for the analysis and transformation of the accumulated information. (5)

CCDC

The Cambridge Structural Database (CSD)

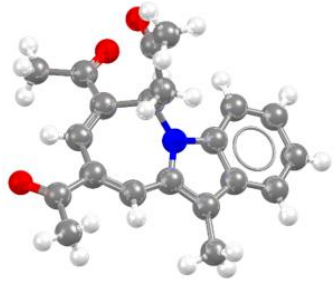


- Every published structure
  - Inc. ASAP & early view
  - CSD Communications
  - Patents
  - University repositories
  - Thesis
- Every entry enriched and annotated by experts
- Discoverability of data and knowledge
- Sustainable for over 55 years
- A trusted CoreTrustSeal repository

CORE TRUST SEAL

CCDC

CSD Refcodes



CSD Refcode - XOPCAJ

What is XOPCAJ?

- A CSD Refcode
- A database reference code
- Containing 6-8 characters
- Used to identify entries in the CSD

Refcode families

- The same substances are assigned the *same* 6 letter code plus an additional 2 numbers
  - Polymorphs
  - New determinations or re-refinements of the same substance
  - Determinations at different temperatures/pressures
- Stereoisomers or different solvates, co-crystals, etc are assigned *different* refcode families

Some of my favourite refcodes are: KITTEN, BATMAN, DISNEY, GAUTAM, GLYCIN

CCDC

Inside the Cambridge Structural Database

The CSD is a database of all the published organic and metal-organic experimental crystal structures

Organic
43%

Metal-Organic
57%

At least one transition metal, lanthanide, actinide or any of Al, Ga, In, Tl, Ge, Sn, Pb, Sb, Bi, Po

Not Polymeric
89%

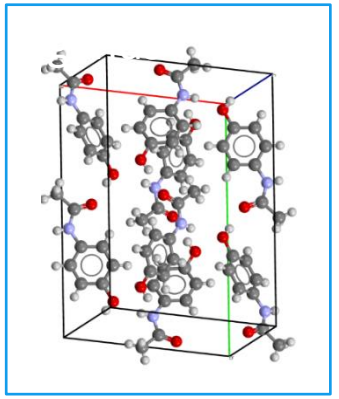
Polymeric: 11%

Single Component
56%

Multi Component
44%

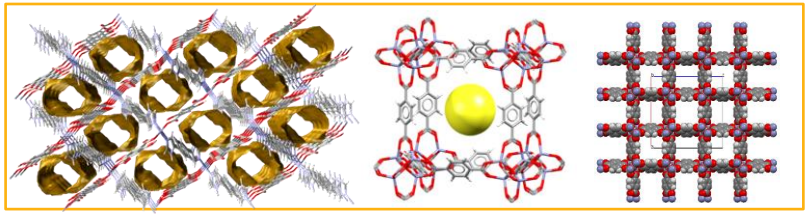
Organic

- Drugs
- Agrochemicals
- Pigments
- Explosives
- Protein ligands



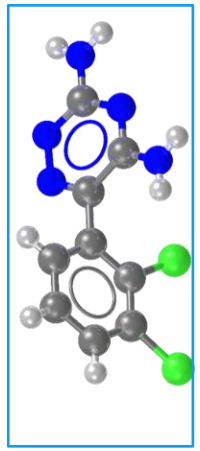
Metal-Organic

- Metal Organic Frameworks
- Models for new catalysts
- Porous frameworks for gas storage
- Fundamental chemical bonding



Additional data

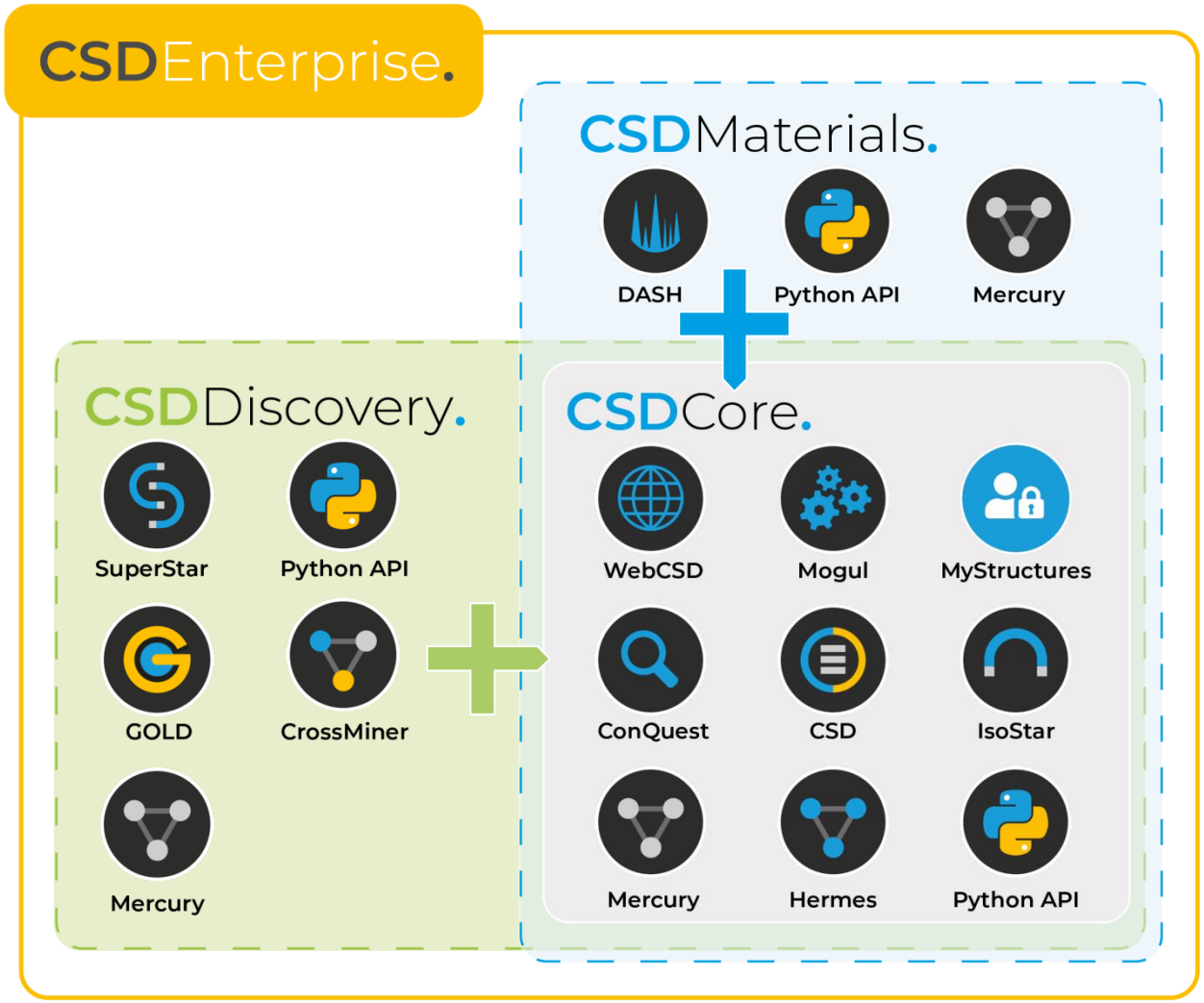
- 11,525 polymorph families
- 171,683 melting points
- 909,992 crystal colours
- 778,663 crystal shapes
- 24,916 bioactivity details
- 11,387 natural source data
- > 250,000 oxidation states



Links and subsets

- DrugBank
- Druglike
- MOFs
- PDB ligands
- PubChem
- ChemSpider
- Pesticide PDB

The CSD software

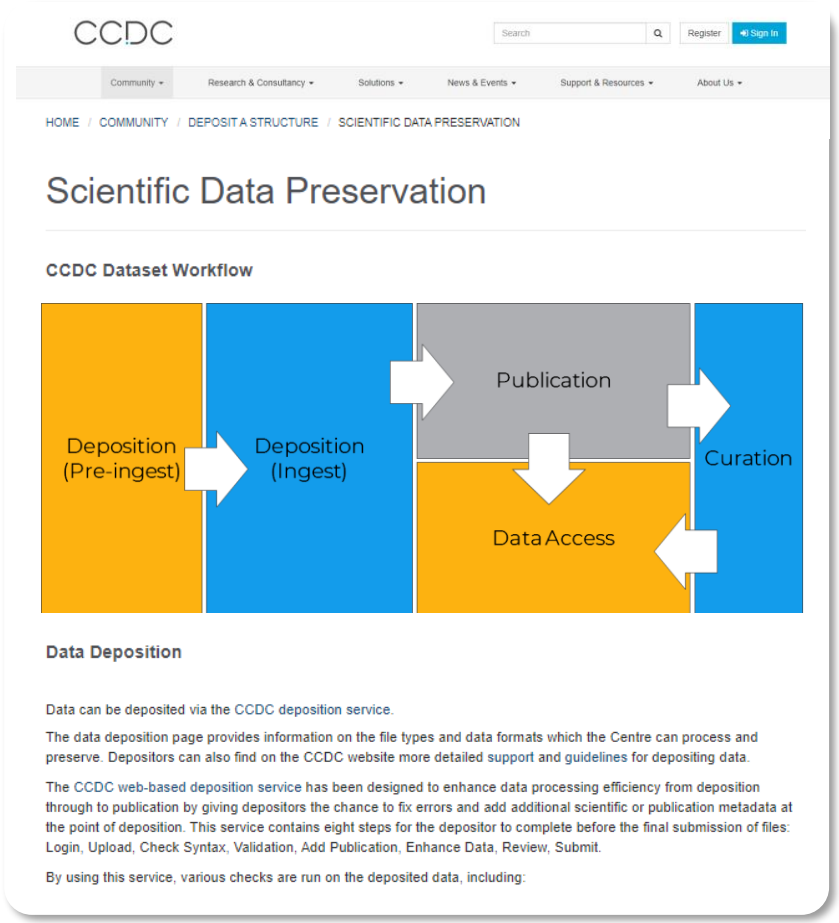


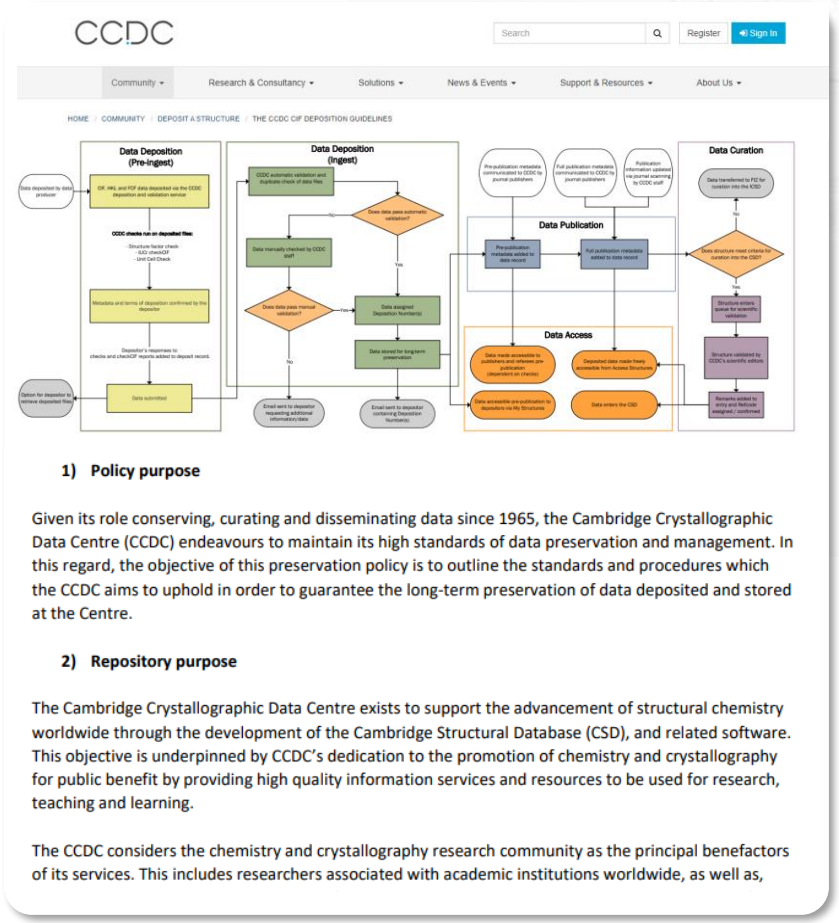
Professional Services

Research & Knowledge partnerships

CCDC

A data preservation policy





CORE TRUST SEAL

CCDC

https://www.ccdc.cam.ac.uk/Community/depositastructure/scientific-data-preservation/

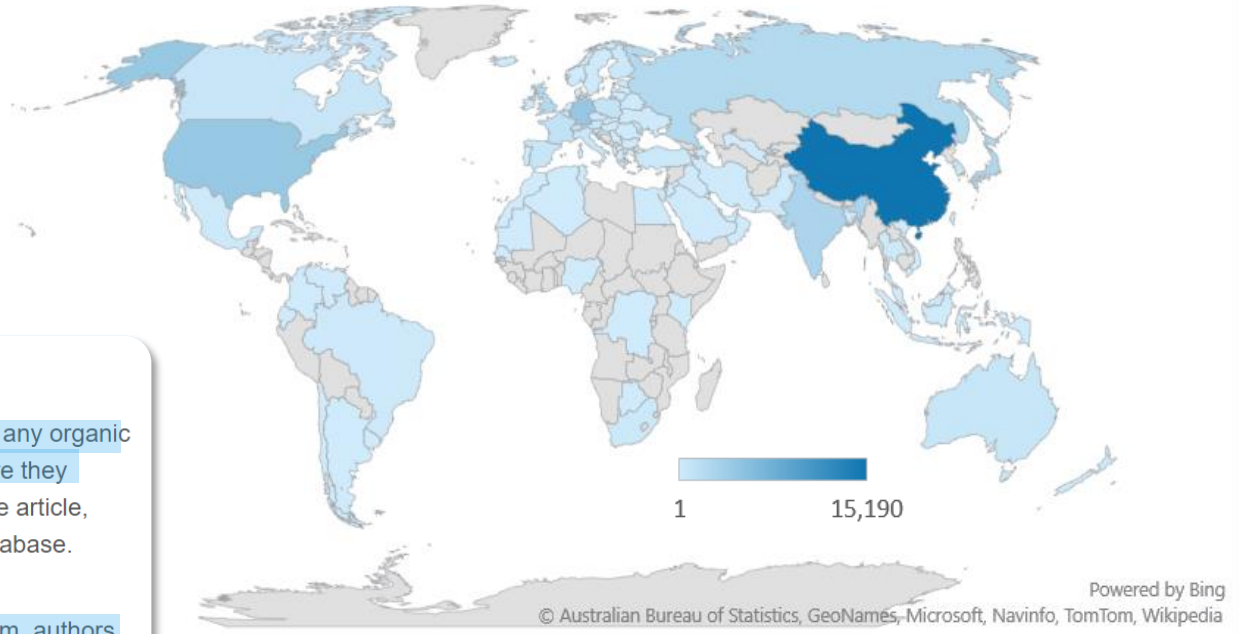
Deposition

- Data deposited pre-publication
  - Enables links to data sets at the point of publication
  - Tailored deposition service

Small molecule single crystal data

Authors should present their crystal data in a CIF (Crystallographic Information File) format and deposit any organic or organometallic structural information with the Cambridge Crystallographic Data Centre (CCDC) before they submit their manuscript to us. Data will be held in the CCDC's confidential archive until publication of the article, when data for organic and organometallic compounds will be entered into the Cambridge Structural Database. Authors are encouraged to deposit inorganic crystal structures with the ICSD, hosted by FIZ Karlsruhe.

During submission of a manuscript to the Royal Society of Chemistry using our online submission system, authors will be asked to provide CCDC reference numbers; CIFs should not be submitted with the manuscript (these should have already been deposited with the CCDC/ICSD, see above). Any revised CIFs obtained subsequently should be deposited directly with the CCDC before the revised manuscript is submitted to us. CCDC or ICSD numbers should be included in the manuscript prior to submission.



2021 deposited structures by crystallographer country

>30% of depositions from China

www.ccdc.cam.ac.uk/deposit

When should you deposit data?

- Pre-publication – when you have finished a structure is when you know most about the dataset
- Enables you to check and correct data prior to inclusion in a manuscript
- Deposited data is stored privately at the CCDC prior to publication
- Most publishers require data to be deposited before manuscript submission
  - Enables publishers and referees secure access to the data during the peer review process
  - Enables links to the data to be added at the point of publication

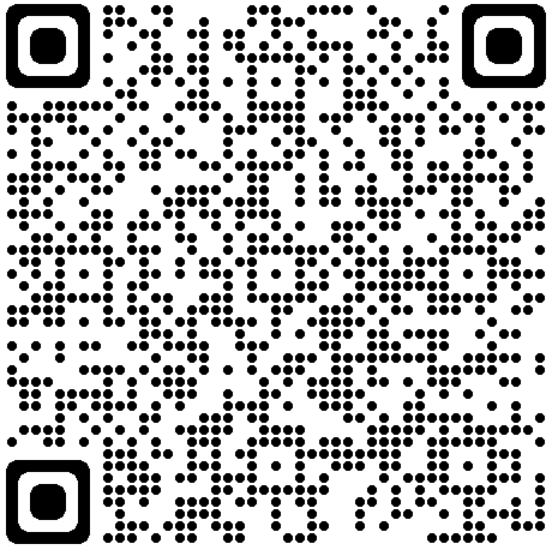
CCDC

CSD criteria

Organic and metal-organic experimental crystal structures in CIF format

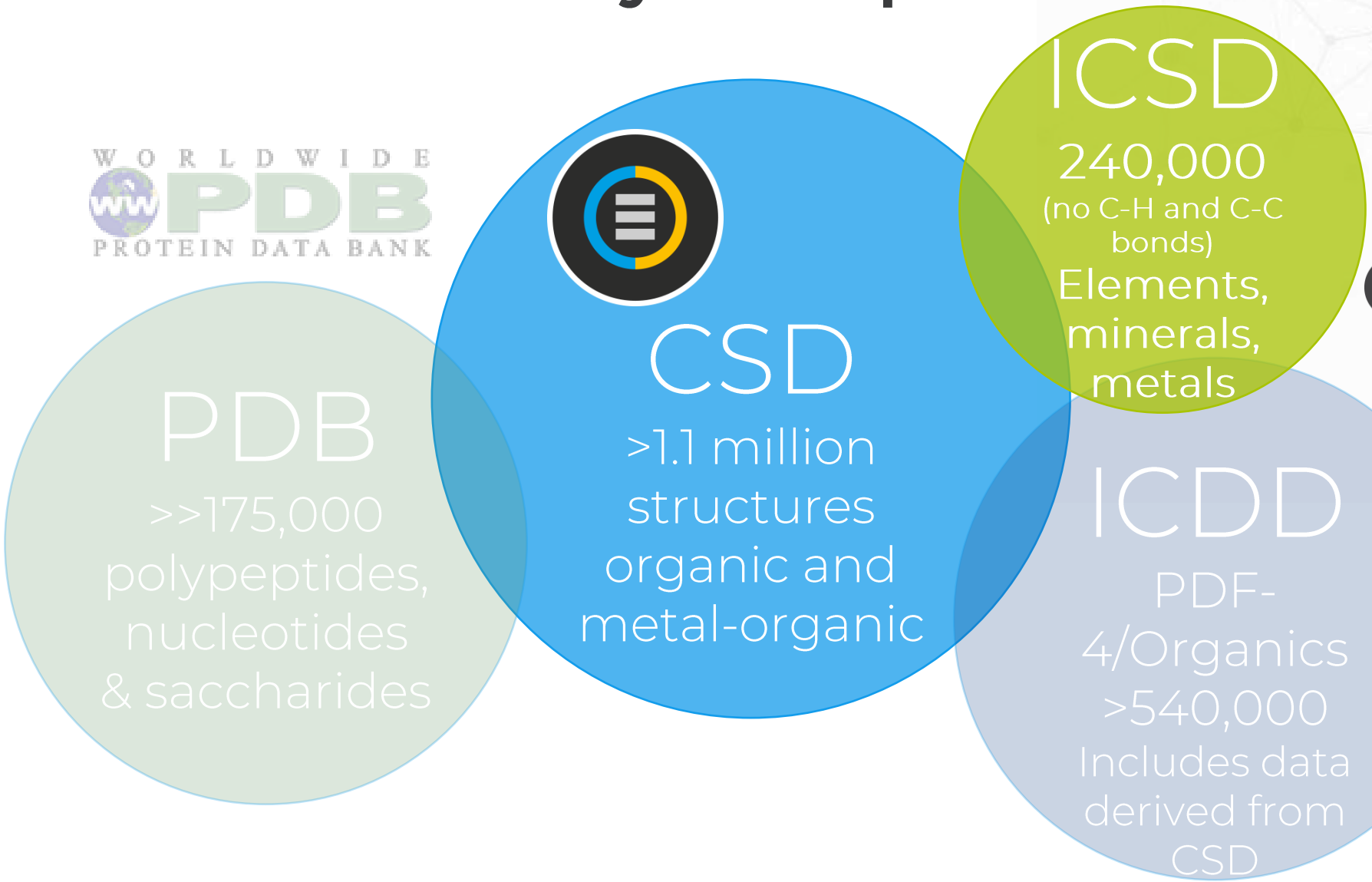
- Types of experimental determination include:
  - Single crystal studies
    - where cell parameters are reported
  - Powder studies
    - where cell parameters, atomic coordinates and constrained refinement
  - from x-ray, neutron or electron diffraction

- Organic and metal-organic chemistry includes:
  - peptides and saccharides of up to 24 residues
  - mono-, di- and tri-nucleotides
  - metal carbonyls
  - boron compounds containing one or more B-H or B-OH bond and borazines
  - ring compounds containing any two of: N, P, S, Se and Te



CCDC

What data can you deposit with us?



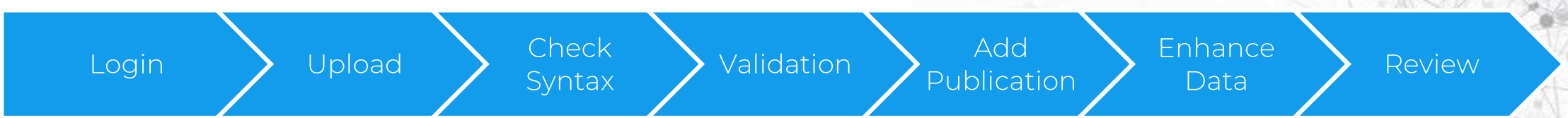
FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

ICDD

CCDC

Joint ICSD/CSD deposition service

Deposit your crystal structures ready for publication using our joint deposition service – we'll handle entry into the CSD or ICSD



Our deposition service aims to make it easy for you to:

- Follow community recommendations
- Provide reliable data and metadata
- Deposit all organic, inorganic and metal-organic data
- Publish directly through a database

Components include:

- Identification of contributors
- Use of standard formats & syntax checking
- Generation of validation report
- Capture of publication, experimental and chemical metadata
- Additional enrichment of data by CCDC/FIZ Karlsruhe

www.ccdc.cam.ac.uk/deposit

CCDC

Deposition

CIF deposition and validation service

Click here to continue to the deposition process without signing in.

If you have not registered for a CCDC account before you can register using the button below. To find out more about CCDC accounts and what you are able to do once you have logged in, please see our support page.

Sign in with your CCDC account

Username or Email

sword

Remember me?

Sign In | Register | Forgotten Username or Password

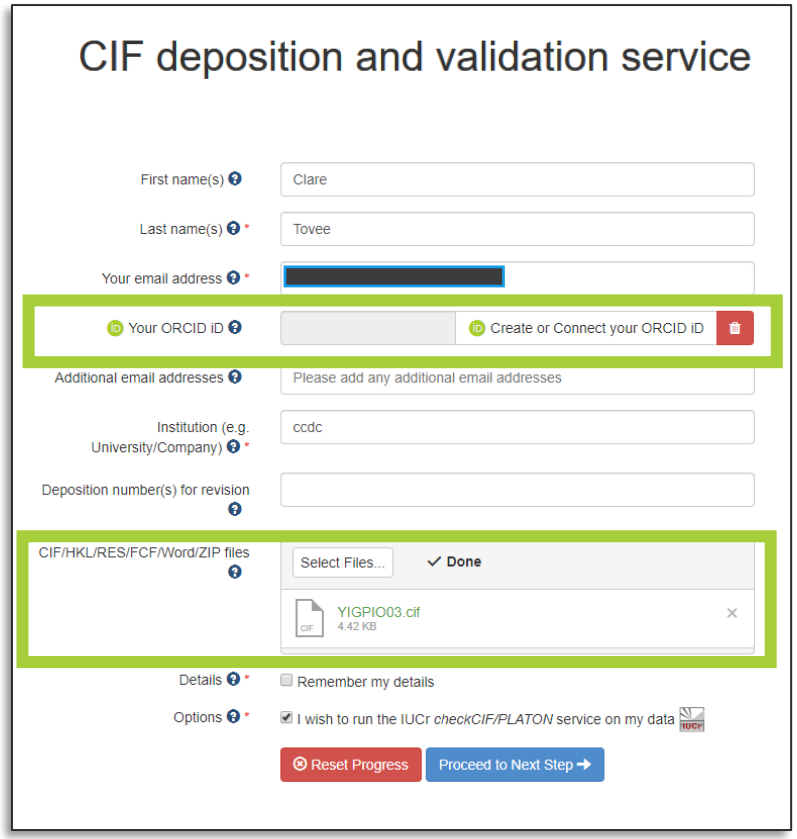
If you don't want an account on our website you can click here

Why register?

- Simplifies deposition
- Gives you access to see, edit and share your data
- Enables you to link your CSD licence to use WebCSD

Deposition

Login → Upload → Check Syntax → Validation → Add Publication → Enhance Data → Review



Along with your personal details you are asked to add:

- Your ORCID ID
  - A persistent digital identifier that distinguishes you from every other researcher
  - Enable you to add your data into your ORCID ID
  - Displayed on our Access Structures service alongside your data
- Your files
  - In CIF format – the standard file format for crystallographic data
- Two options
  - Remember your details
  - Run checkCIF – a service to check the integrity of your data

Deposition

Login → Upload → Check Syntax → Validation → Add Publication → Enhance Data → Review

CIF deposition and validation service

No Structure Factor data have been uploaded.

Structure Factor data are an essential part of the deposition. You should click 'Go Back' to add Structure Factor data to your deposition. If you are unable to embed Structure Factor data into your CIF automatically, you can upload this data as separate files alongside your CIF.

If in exceptional circumstances you are unable to include Structure Factor data with your deposition you may continue by clicking 'Proceed to Next Step', however you must use the comment box below to explain why this data is not available. The text you enter will be automatically embedded into your deposited CIFs.

Reason why your deposition does not include Structure Factor data (this comment will automatically be added into your deposited CIFs): *

← Go Back | Proceed to Next Step →

Structure Factor data check

- The experimental data which were used to determine the 3D structure
- Structure factor data enables
  - Rigorous validation of the structure-determination results
  - Their preservation and continued accessibility
- Data should either be appended to the CIF or in .FCF or .HKL format
- If data is missing you need to add a comment explaining why

https://www.ccdc.cam.ac.uk/Community/depositastructure/cif-deposition-guidelines/structure-factor/

CCDC

Structure factors

Structure factors: what

The CCDC and FIZ Karlsruhe strongly encourage the incl
Structure Database (ICSD), in line with recommendations

What are structure factors?

Structure factors are created from experimental crystallog
in the structure. For detailed information, see the IUCr's c

Currently, there are two types of information that CCDC a
intensities (.hkl). They can also be appended to the CIF fi

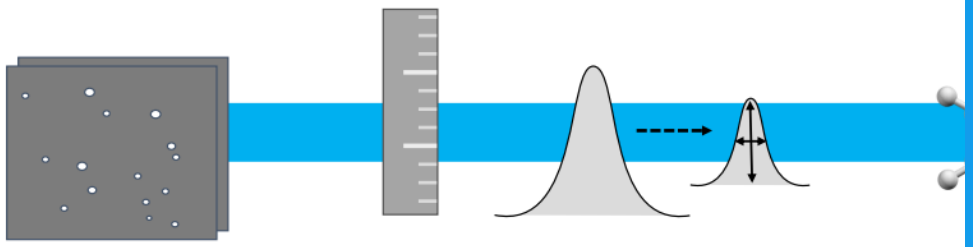
- **The .hkl file** contains the intensity and standard ur
values: h, k and l - to identify the reflection. This inf
- **The .fcf file** contains the structure factors, which a
be recreated using the .hkl and .res file. This last o
constraints that have been applied). The .fcf file is

These files provide more information about the crystal str
crystallographic data, as part of the publications standard

Why is it important to share struct

- It's good data practice - by sharing the data used t
later, then it is already stored, especially if you mov
- CheckCIF (the IUCr's CIF checking service) can d
.res in the CIF allows such checks to be performed
- Several journals require the deposition of structure

A number of different files are created at each stage
crystallographic experiment – here we have listed so
the main ones. Files may have different names/some s
may be combined in different crystallographic softw



Data Collection

Calculating reflection intensities and applying corrections (Integration & Scaling)

Struct
mo

- Diffraction frames (.sfrm, .cbf, .img, etc.)

- .hkl

If the diffraction frames have been deposited in another repository, the DOI can be associated with your dataset during deposition under "Raw Data DOI"

Files for CCDC deposition:

- Mandatory
- Recommended

If no structure factor data has been uploaded, you should:

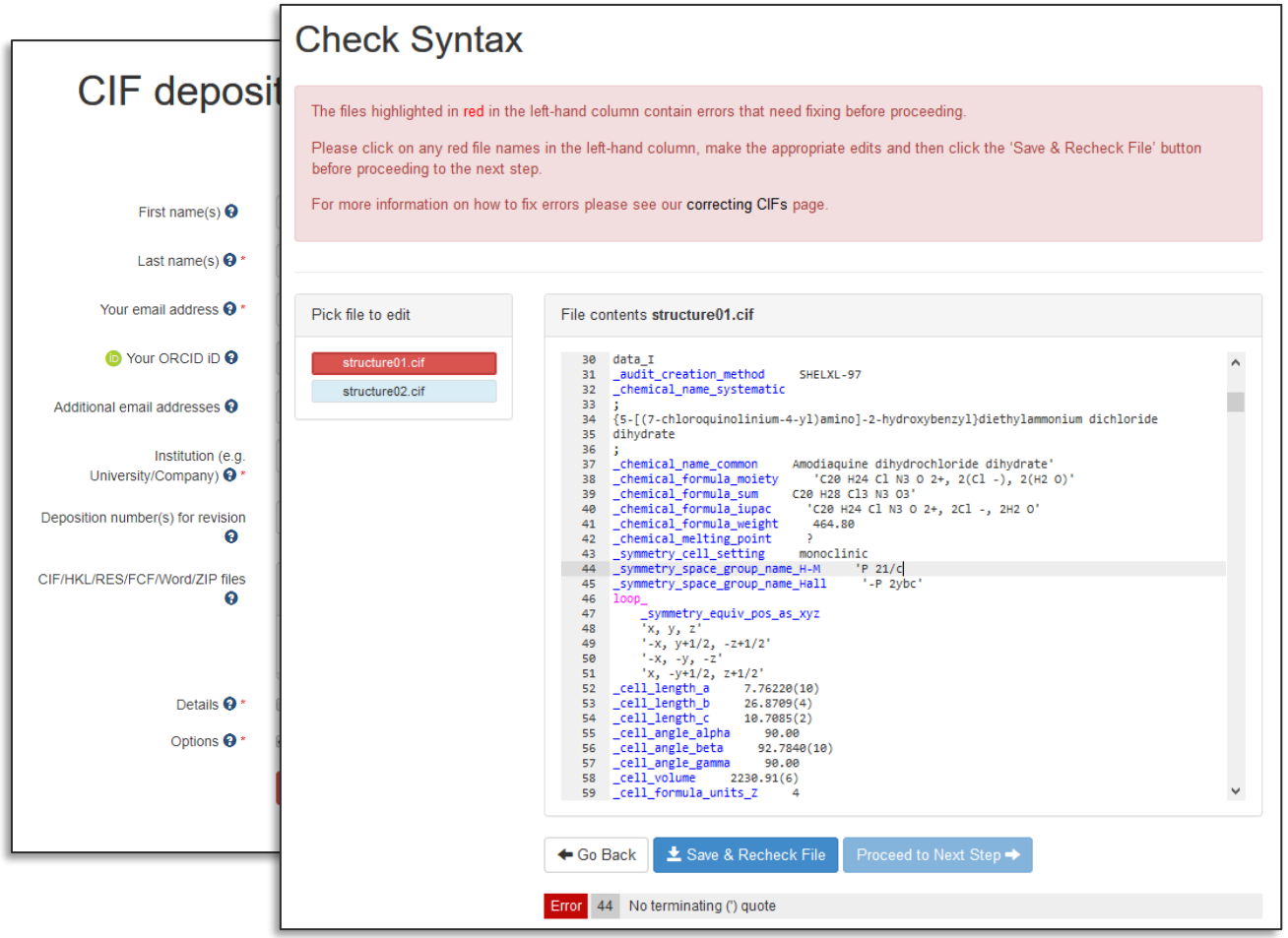
- Go back and upload a new dataset which includes structure factors
- Most refinement software now embeds this into a CIF
- If it is historic data and no structure factor data is available, you need to add a comment before proceeding.

www.ccdc.cam.ac.uk/Community/depositastructure/cif-deposition-guidelines/structure-factor/

CCDC

Deposition

Login → Upload → Check Syntax → Validation → Add Publication → Enhance Data → Review



Checking the syntax

- Uses enCIFer syntax checking
- Syntax colour coded
  - Field names (blue)
  - Loops (pink)
  - Data (black)
- Errors highlighted
- Syntax errors need fixing before proceeding
- Once fixed save and recheck errors

CCDC

```
1  data_5_93GPa
2
3  _audit_creation_method           CRYSTALS_ver_12.80
4  _chemical_name_systematic        '2-hydroxybenzaldehyde oxime'
5  _chemical_name_common            salicylaldoxime
6  _chemical_formula_moiety         'C7 H7 N1 O2'
7  _chemical_formula_sum            'C7 H7 N1 O2'
8  _chemical_formula_weight         137.14
9  _chemical_melting_point          59-61
10 _symmetry_cell_setting           Monoclinic
11 _symmetry_space_group_name_H-M   'P 1 21/n 1 '
12 _symmetry_space_group_name_Hall  '-P 2yn'
13 loop_
14 _symmetry_equiv_pos_as_xyz
15
16 x,y,z
17 -x,-y,-z
18 -x+1/2,y+1/2,-z+1/2
19 x+1/2,-y+1/2,z+1/2
20 ;
21
22 _cell_length_a                  7.677(3)
23
24 _cell_length_b                  5.7731(8)
25
26 _cell_length_c                  12.159(3)
27
28 _cell_angle_alpha               90
29
30 _cell_angle_beta                110.62(2)
```

You can easily see if the CIF format is broken by looking at the colour coding

```
loop_
_symmetry_equiv_pos_as_xyz

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

_cell_length_a                  7.677(3)
_cell_length_b                  5.7731(8)
_cell_length_c                  12.159(3)
_cell_angle_alpha               90
_cell_angle_beta                110.62(2)
```

← Go Back | Save & Recheck File | Proceed to Next Step →

Please click on the error message to navigate to the location of the error in the CIF

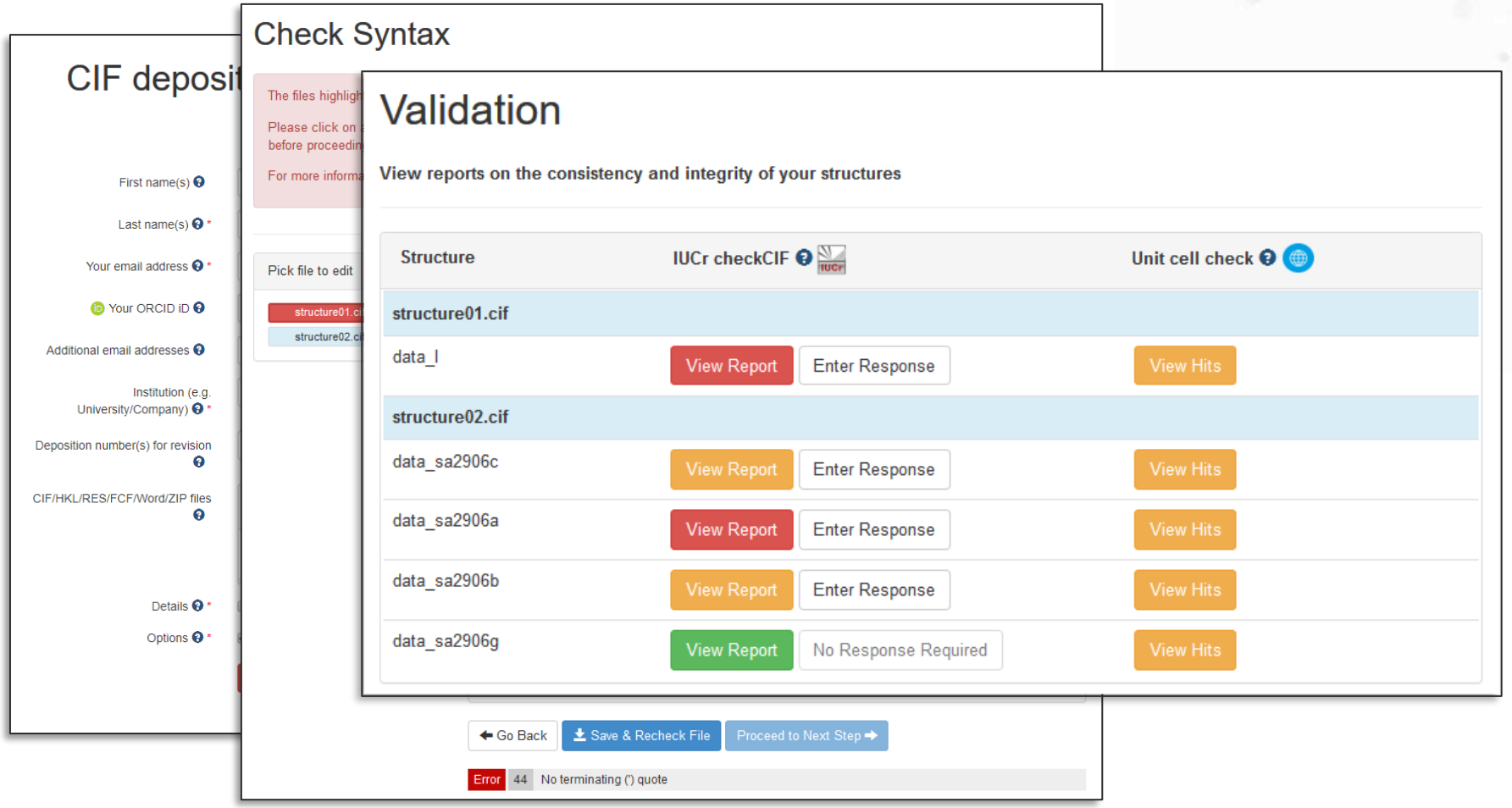
Error	302	Text block finished at end of file without final ';'
Error	-1	The CIF contains no data blocks recognised as crystal structure data

You can easily navigate to errors

CCDC

Deposition

Login → Upload → Check Syntax → Validation → Add Publication → Enhance Data → Review

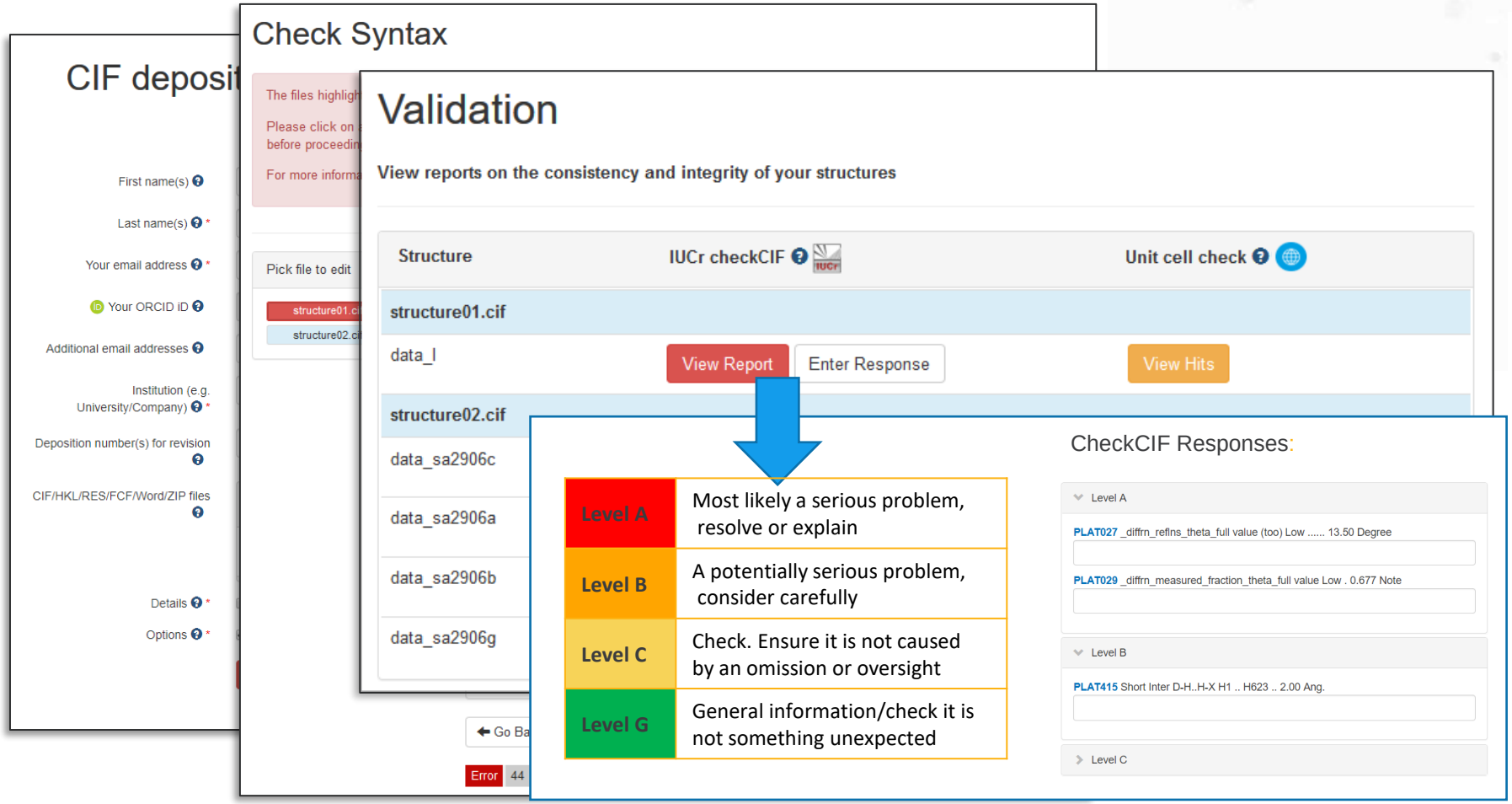


CheckCIF

- Checks consistency and integrity of the data
- Generates alerts that should either be corrected or explained
- Report and explanations can be embedded in a CIF

Deposition

Login → Upload → Check Syntax → Validation → Add Publication → Enhance Data → Review



CheckCIF

- Checks consistency and integrity of the data
- Generates alerts that should either be corrected or explained
- Report and explanations can be embedded in a CIF

What to do if you have a checkCIF alert?

Alerts provide potential errors, unusual findings and suggestions for improvement

- All alerts should be checked
- Then decide if you need to improve your model/dataset or add a comment providing more information about the reason for the alert

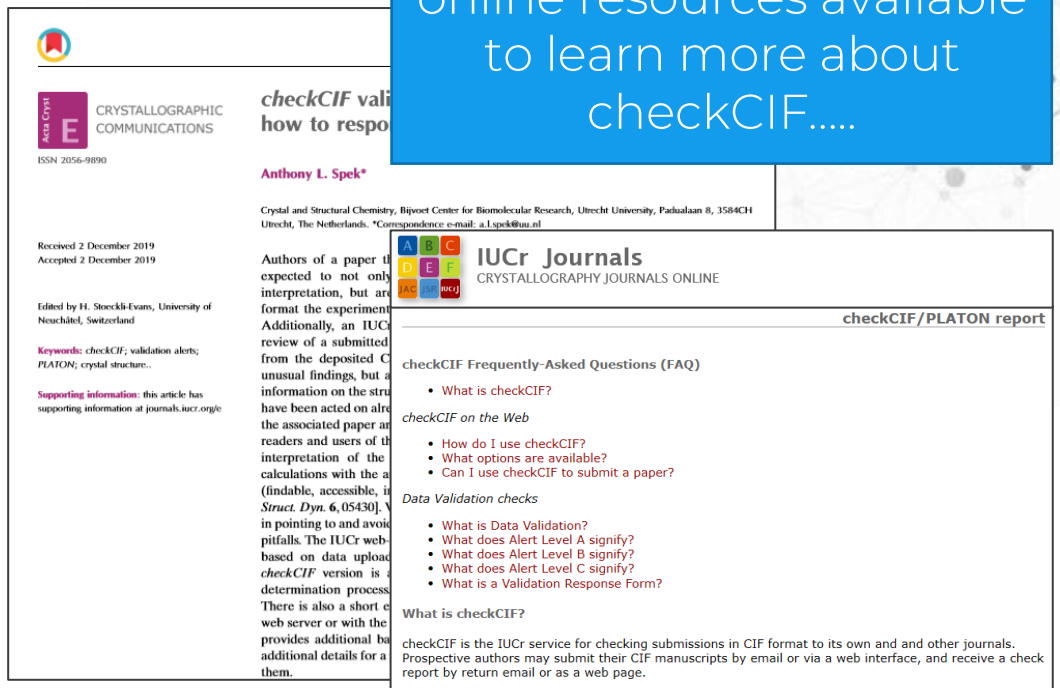
Alert level A
PLAT183_ALERT_1_A Missing _cell_measur
PLAT184_ALERT_1_A Missing _cell_measur
PLAT185_ALERT_1_A Missing _cell_measur

Alert level B
PLAT029_ALERT_3_B _diff

Alert level C
ABSMU01_ALERT_1_C The

Alerts are hyperlinked to more information about the alert and what action should be taken

There are a number of online resources available to learn more about checkCIF.....


CRYSTALLOGRAPHIC COMMUNICATIONS
ISSN 2056-9890

checkCIF vali
how to respo

Anthony L. Spek*

Crystal and Structural Chemistry, Bijvoet Center for Biomolecular Research, Utrecht University, Padualaan 8, 3584CH Utrecht, The Netherlands. *Correspondence e-mail: a.l.spek@uu.nl

Received 2 December 2019
Accepted 2 December 2019

Edited by H. Stoeckli-Evans, University of Neuchâtel, Switzerland

Keywords: *checkCIF*; validation alerts; *PLATON*; crystal structure..

Supporting information: this article has supporting information at journals.iucr.org/e

Authors of a paper t
expected to not only
interpretation, but ar
format the experiment
Additionally, an IUCr
review of a submitted
from the deposited C
unusual findings, but a
information on the stru
have been acted on alre
the associated paper ar
readers and users of th
interpretation of the
calculations with the a
(findable, accessible, i
Struct. Dyn. **6**, 05430].
in pointing to and avoi
pitfalls. The IUCr web-
based on data upload
checkCIF version is
determination process
There is also a short e
web server or with the
provides additional ba
additional details for a
them.

IUCr Journals
CRYSTALLOGRAPHY JOURNALS ONLINE

checkCIF/PLATON report

checkCIF Frequently-Asked Questions (FAQ)

- What is checkCIF?

checkCIF on the Web

- How do I use checkCIF?
- What options are available?
- Can I use checkCIF to submit a paper?

Data Validation checks

- What is Data Validation?
- What does Alert Level A signify?
- What does Alert Level B signify?
- What does Alert Level C signify?
- What is a Validation Response Form?

What is checkCIF?

checkCIF is the IUCr service for checking submissions in CIF format to its own and and other journals. Prospective authors may submit their CIF manuscripts by email or via a web interface, and receive a check report by return email or as a web page.

checkCIF on the web

How do I use checkCIF?

Web interface
checkCIF is at

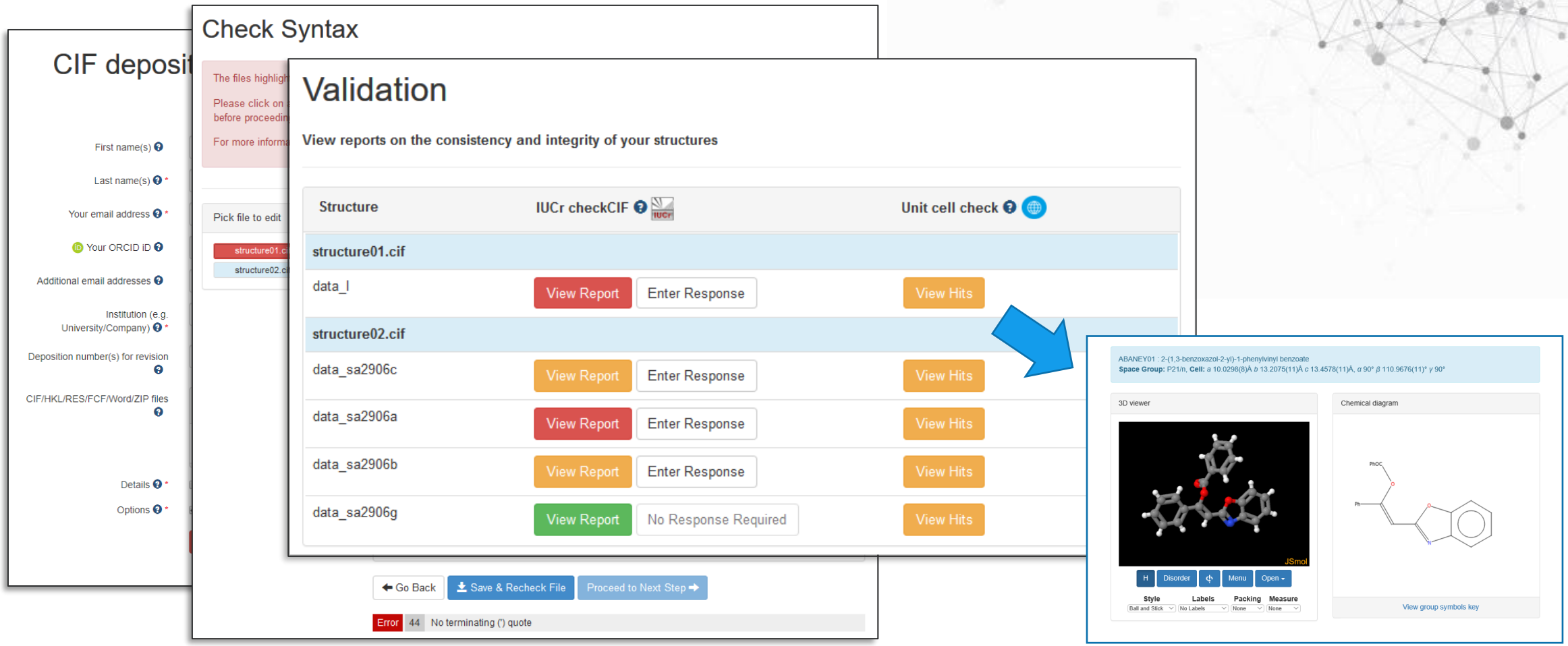
- checkCIF for full publication checks
- checkCIF for basic structural checks.

Enter the full path name of a CIF on your local computer system in the text entry box, or use the **Browse...** button to locate the desired file. Send the file for processing by clicking on the **Send CIF for checking** button. Processing and return of the results may take several minutes, depending on file sizes, available network bandwidth and load on the server computer.

- Acta E, Volume 76| Part 1| January 2020| Pages 1-11 https://doi.org/10.1107/S2056989019016244
- https://journals.iucr.org/services/cif/checking/checkfaq.html

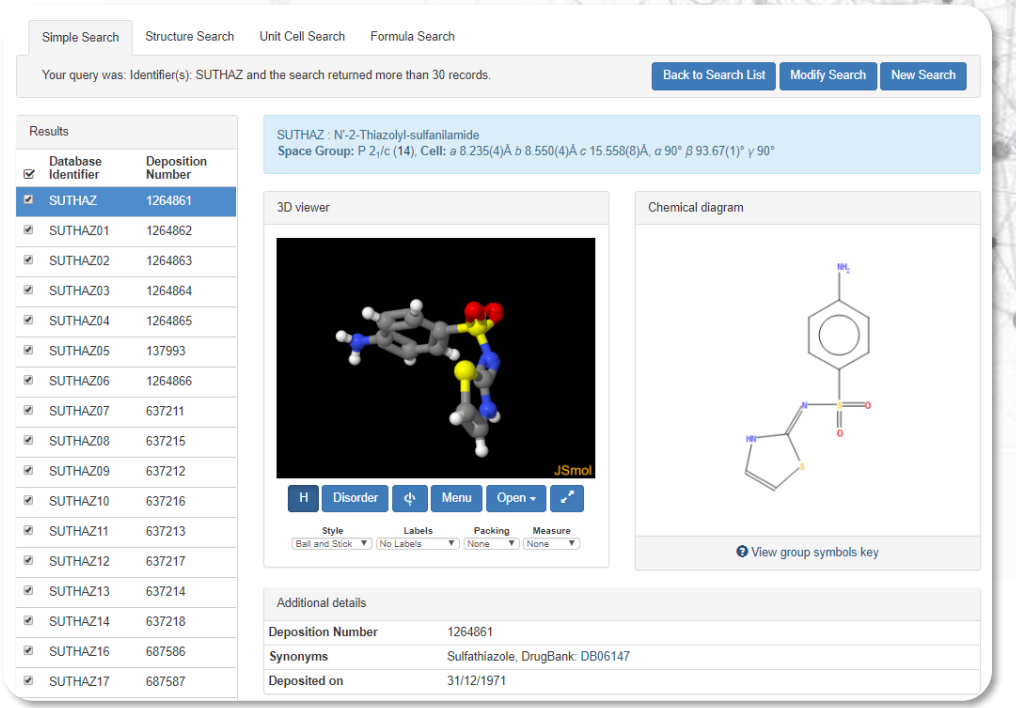
Deposition

Login → Upload → Check Syntax → Validation → Add Publication → Enhance Data → Review



What if my compound is already in a database?

- You can still deposit your data – it's still useful for research!
- Once deposited you can either:
  - Publish your new structure in an associated scientific article
  - Publish directly through the CSD without an associated scientific article, known as a ***CSD Communication***



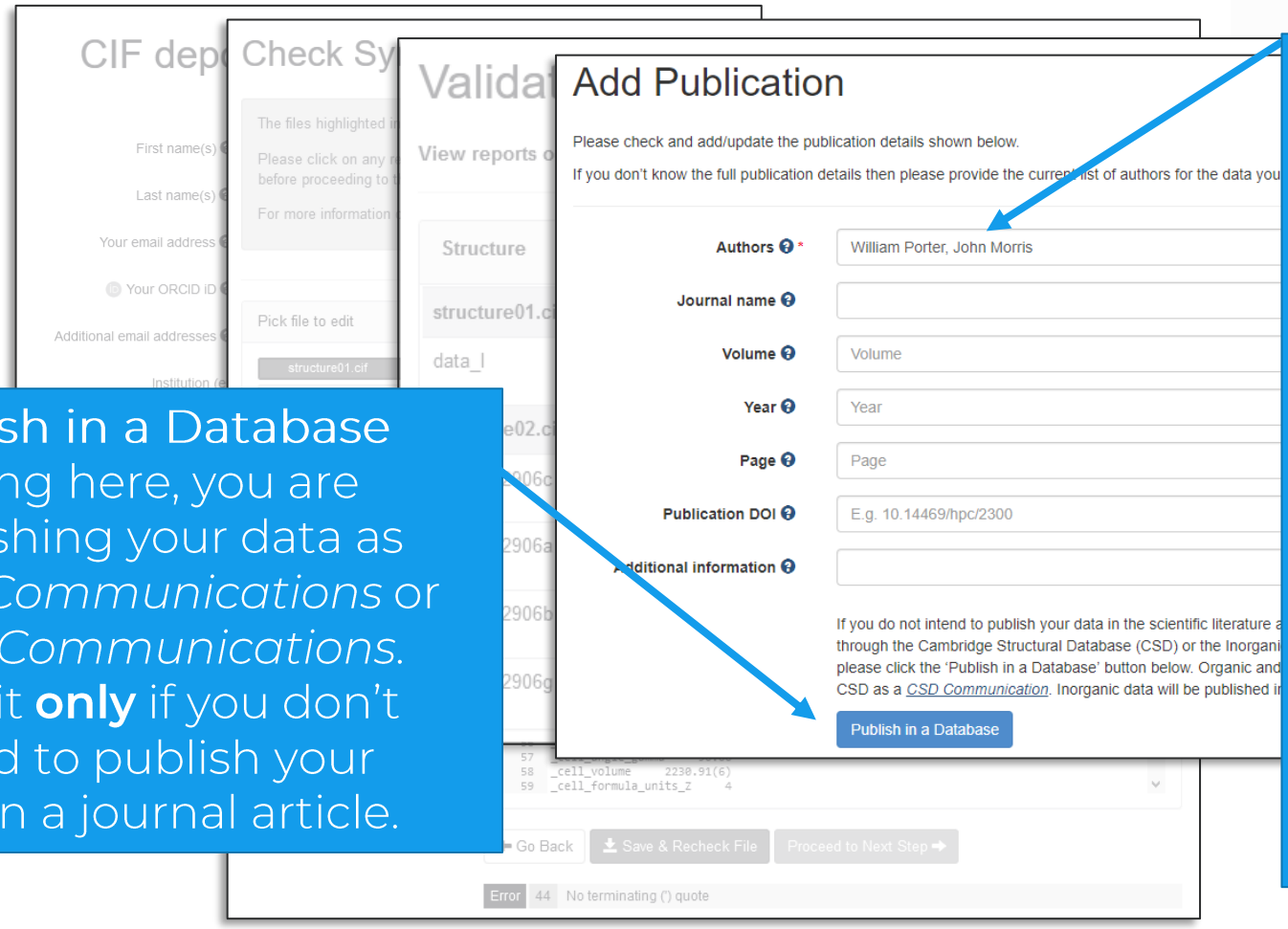
Did you know glycine is one of the most independently determined substance in the CSD? The structure of glycine has been determined so many times there are two different refcode families!

For example, there are over 46 sulfathiazole structures in the CSD Refcode family SUTHAZ

CCDC

Deposition

Login → Upload → Check Syntax → Validation → Publication → Enhance Data → Review



Publish in a Database

Clicking here, you are publishing your data as *CSD Communications* or *ICSD Communications*. Click it **only** if you don't intend to publish your data in a journal article.

Authors' names format

- Same order as on publication.
- Each name should be separated by a comma and space.
- No any additional characters, titles, or salutation.
- You can use initials or expand the first name; if you are using initials, then you do not need spaces between them.

Examples of correct authors' names in the CSD e.g.: *C.A.Tovee, C.A.Kilner, J.A.Thomas, M.A.Halcrow*
or *Clare A. Tovee, Colin A. Kilner, Jim A. Thomas, Malcolm A. Halcrow.*

Deposition

www.ccdc.cam.ac.uk/community/csd-communications/

Login → Upload → Check Syntax → Validation → Add Publication → Enhance Data → Review

Add Publication

Please check and add/update the publication details shown below.

If you don't know the full publication details then please provide the current list of authors for the data you are deposi

Authors * William Porter, John Morris

Journal name

Volume Volume

Year Year

Page Page

Publication DOI E.g. 10.14469/hpc/2300

Additional information

If you do not intend to publish your data in the scientific literature and would li
through the Cambridge Structural Database (CSD) or the Inorganic Crystal St
please click the 'Publish in a Database' button below. Organic and metal-orga
CSD as a *CSD Communication*. Inorganic data will be published in the ICSD

Publish in a Database

CSD Communications

ISSN 2631-9888

***CSD Communications* is a collection of small molecule crystallographic data which has been shared by depositors through the Cambridge Structural Database (CSD) without an associated scientific article.**

This may include new data collections of novel structures or new determinations or re-refinements of previously published compounds. *CSD Communications* have an ISSN number and each entry contains the author list and crystallographer details as supplied by the depositor. This allows authors and data producers to get credit for their data, as their structure receives a full CCDC citation including DOI and publication year. The objective of *CSD Communications* is to help scientists contribute to the quantity of crystallographic data available to the public. So far *CSD Communications* have provided the scientific community with access to more than 28,000 crystallographic structures which may never have otherwise been made public.

More information about the benefits of *CSD Communications* is available here and we also have a series of blog posts about *CSD Communications* and data sharing in general if you want to learn more about this initiative.

Archive

The *CSD Communications* archive contains all of the data published directly through the Cambridge Structural Database (CSD) by the Cambridge Crystallographic Data Centre (CCDC) as *CSD Communications* (previously known as *Private Communications*). Select a year to view all *CSD Communications* published in that year.

- 2021
- 2020
- 2019
- 2018
- 2017
- 2016
- 2015

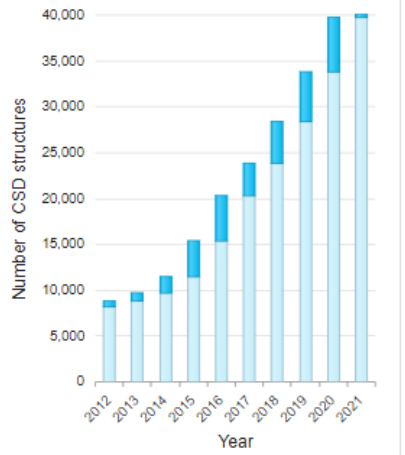
In This Section

CSD Communications Information

CSD Communications

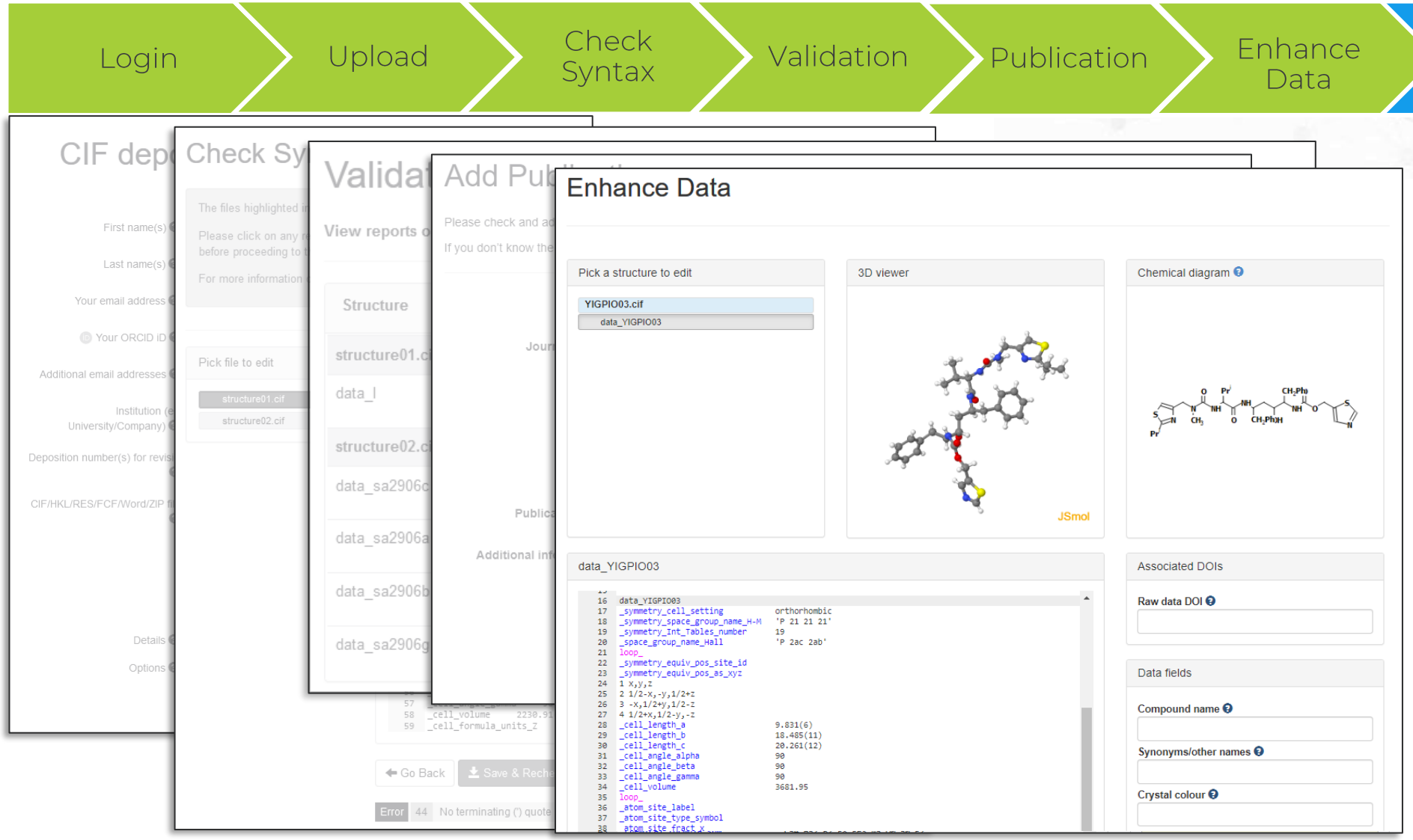
This Year: 290 **To Date:** 40,077

CSD Communications growth



Deposition

Login → Upload → Check Syntax → Validation → Publication → Enhance Data → Review



During this stage:

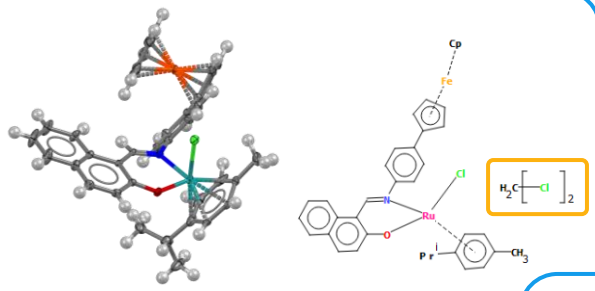
- Check
- Enhance

Each data set by updating info in the data fields

CCDC

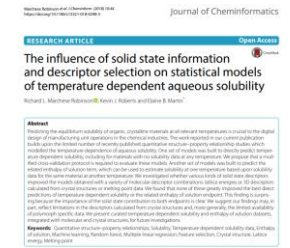
Enhancing key metadata

- Depositors are asked to check and enhance key metadata
- Helps to improve the quality of key information
- Helps to ensure data is correctly curated and represented in the CSD
- Helps to improve the availability of data
- Ultimately improves the discoverability and re-use of datasets and enables scientists worldwide to learn more from the data

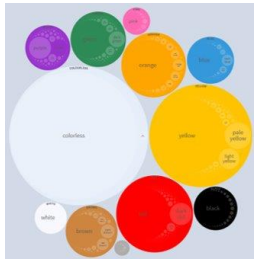


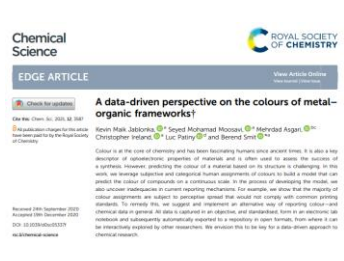
CSD BAGNEG Example of SQUEEZED data being represented correctly

>170K melting points in the CSD



Additional metadata being used to help predict solubilities



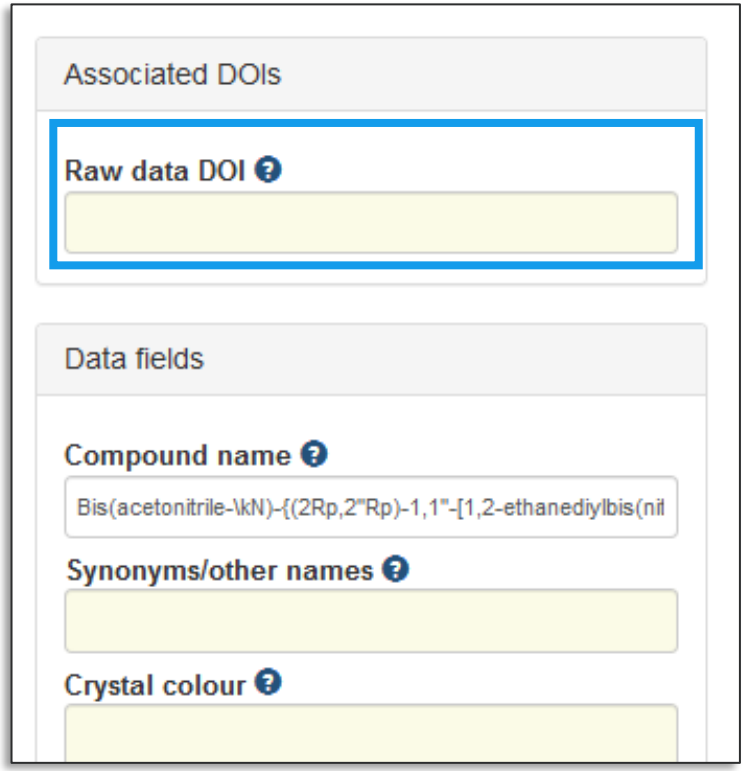


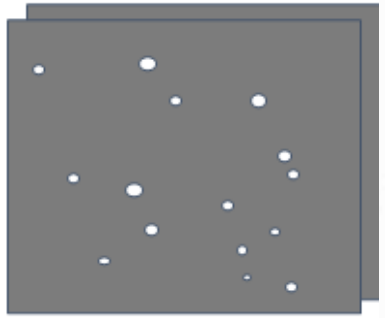
Example of using CSD data to assess the colours of Metal-Organic Frameworks

Chem. Sci., 2021,12, 3587-3598, https://doi.org/10.1039/D0SC05337F,
J Cheminform 10, 44 (2018). https://doi.org/10.1186/s13321-018-0298-3

CCDC

Raw data DOIs





Raw data DOIs are for data created during data collection

- Diffraction frames e.g. .sfrm, .cbf, .img, etc

- We do not accept diffraction frame data during deposition
- But you can choose to upload your diffraction frames in another repository or institution repository
  - If you do publish your diffraction frames you should add the DOI to “Raw data DOI”
- Raw data DOI should not include the DOI of associated articles

CCDC

Review / Submit

Login → Upload → Check Syntax → Validation → Publication → Enhance Data → Review → **Submit**

Review

You are about to submit 3 structures.
Press 'Submit' to finalise the deposition.

First name(s)	Albert
Last name(s)	Einstein
Your email address	[redacted]
Your ORCID iD	
Additional email addresses	
Institution (e.g. University/Company)	CCDC
Authors	Albert Einstein
Files uploaded	Total of 1 file • ExampleCIF.cif

← Go Back | Submit →

Submit

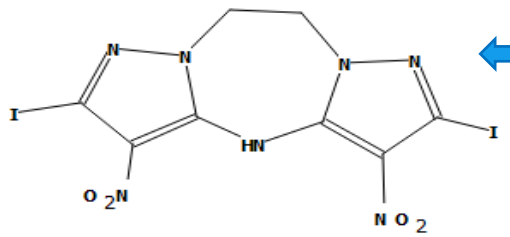
Thank you for your deposition.

Your deposition number(s) will soon (usually within 2 working days) be sent to [redacted]

First name(s)	Albert
Last name(s)	Einstein
Your email address	[redacted]
Your ORCID iD	
Additional email addresses	
Institution (e.g. University/Company)	CCDC
Authors	Albert Einstein
Files uploaded	Total of 1 file • ExampleCIF.cif

If any of the data here are not correct please contact our web-deposition team via email at deposit@ccdc.cam.ac.uk
You can now download all deposited files, generated checkCIF reports and generated chemical diagrams by clicking the Retrieve Deposited Files button below.

View My Structures | Retrieve Deposited Files | Deposit More Structures | Return to CCDC Homepage



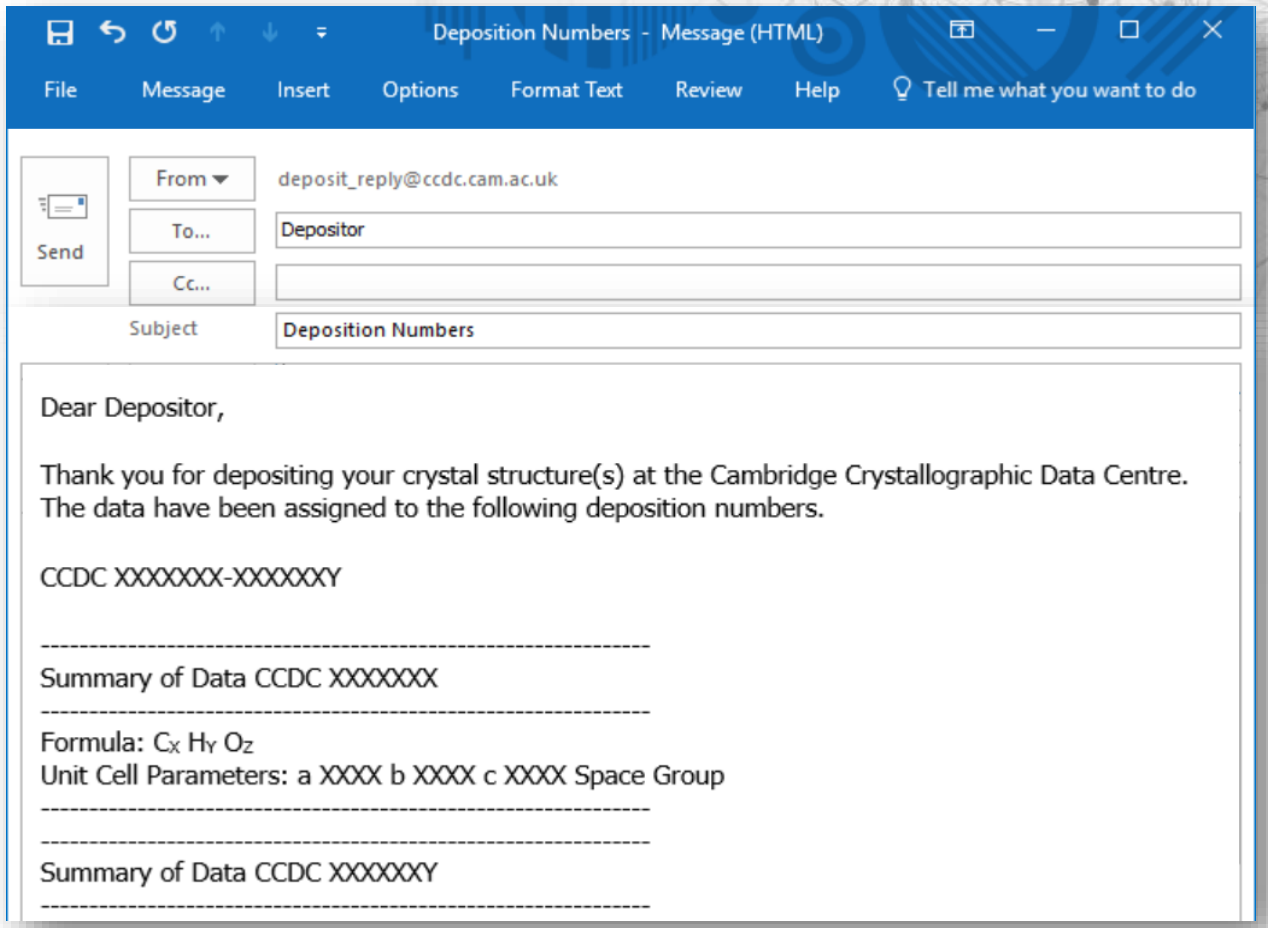
download (3) — Search download (3)

Name	Type	Compressed size	Password ...	Size
ExampleCIF_file001	CIF File	12 KB	No	51 KB
ExampleCIF-data_shrv307_file002	PNG File	7 KB	No	8 KB
ExampleCIF-data_shrv307_file003.mol	MOL File	1 KB	No	3 KB
ExampleCIF-data_shrv328_file006	PNG File	5 KB	No	6 KB
ExampleCIF-data_shrv328_file007.mol	MOL File	1 KB	No	4 KB
ExampleCIF-data_shrv334_file004	PNG File	7 KB	No	7 KB
ExampleCIF-data_shrv334_file005.mol	MOL File	1 KB	No	3 KB

CCDC

Receiving your Deposition Numbers

- Upon submission your data is processed at the CCDC
- Deposition Numbers are assigned to each dataset
- You will receive email confirmation of your Deposition Numbers
  - This could take a few minutes up to a few days for us to process



Dataset identifiers

Deposition Numbers should be included in manuscripts

Deposition Numbers - Message (HTML)

File | Message | Insert | Options | Format Text | Review | Help | Tell me what you want to do

Send

From ▾ deposit_reply@ccdc.cam.ac.uk

To... Depositor

Cc...

Subject **Deposition Numbers**

Dear Depositor,

Thank you for depositing your crystal structure(s) at the Cambridge Crystallographic Data Centre. The data have been assigned to the following deposition numbers.

CCDC XXXXXXX-XXXXXXY

Summary of Data CCDC XXXXXXX

Formula: $C_X H_Y O_Z$

Unit Cell Parameters: a XXXX b XXXX c XXXX Space Group

Summary of Data CCDC XXXXXXY

---

16 M. A. Meador, H. Hart, *J. Org. Chem.* 1989, **54**, 2336–2341.

CrossRef | CAS | Web of Science® Times Cited: 16 | 'ejournals@cambridge - find full text'

17 CCDC 1543805 (**16**) contains the supplementary crystallographic data for this paper. These data can be obtained free of charge from The Cambridge Crystallographic Data Centre.

McKervey, A. R. Maguire, S. M. Tuladhar and M. Fiona Twohig, *J. Chem. Sc*
1047–1054 **DOI**: 10.1039/P19900001047; (*b*) H. Duddeck, *J. Chem. Soc.*, month.
1055–1063 **DOI**: 10.1039/P19900001055; (*c*) P. Panne and J. M. Fox, *J. Ar*
External Links.

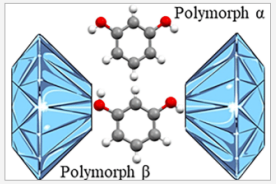
Footnote

† Electronic supplementary information (ESI) available: Experimental procedures and spectros
615536. For ESI and crystallographic data in CIF or other electronic format see DOI: 10.1039/b82

This journal is © The Royal Society of Chemistry 2009

Abstract

Polymorph α of resorcinol, at ambient pressure stable to 365 K when it transforms to polymorph β, is exceptionally resistant to high pressure. The crystals of polymorph α can be compressed to over 4 GPa without transforming to the β phase. We have performed high-pressure recrystallization of resorcinol aqueous and methanol solutions, and they yielded polymorph α below 0.5 GPa and polymorph β above this pressure. Our single-crystal X-ray diffraction studies on resorcinol polymorphs in a diamond-anvil cell reveal the structural origins of the phase transition. The high pressure changes the angular dimensions of bistable hydrogen bonds OH···O, which destabilizes the H-atoms and the structure of polymorph α above 0.5 GPa, consistent with the calorimetric and NMR results. The high-temperature, high-pressure polymorph β achieves the more dense packing through the changed conformation of one of the hydroxyl groups and the considerable twisting of the hydrogen bonds necessary for the formation of additional C–H···π bonds. The large temperature and pressure hysteresis of the polymorphs α and β are connected with the different topologies of their O–H···O networks.



Synopsis

The pressure- and temperature-dependent transition between resorcinol polymorphs α and β has been rationalized in terms of transforming H-bonds and their networks.

Introduction

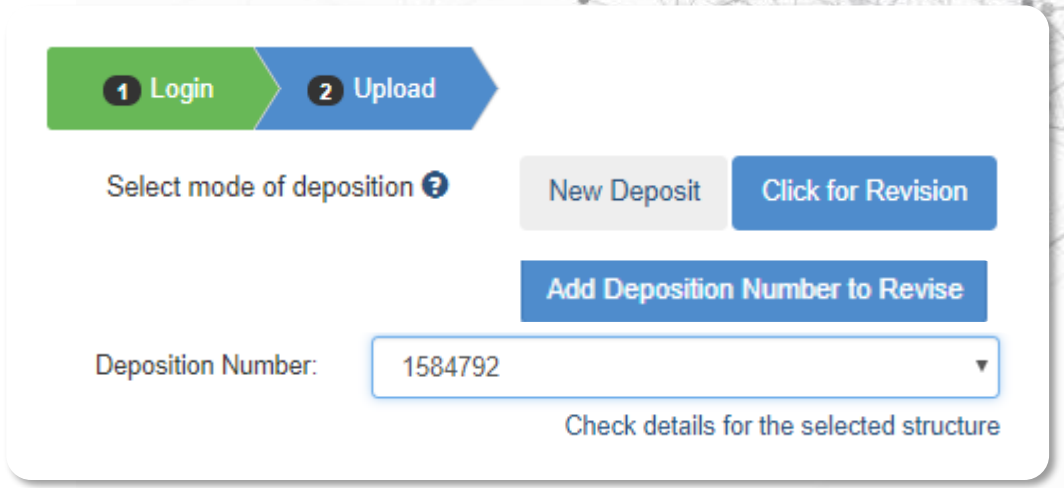
ARTICLE SECTIONS Jump To

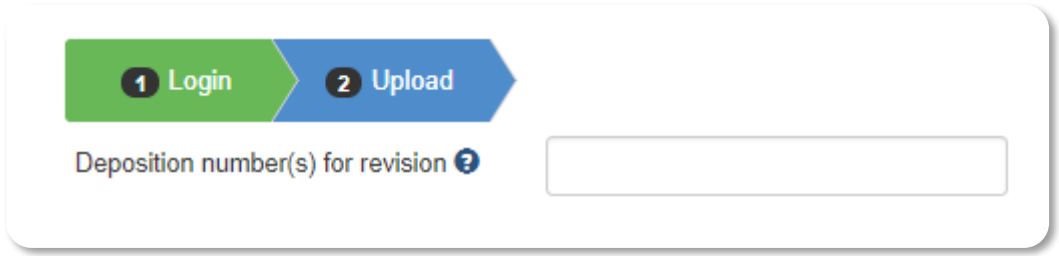
Resorcinol, an intermediate often used in chemical practice and a pharmaceutical agent, was one of the first organic compounds for which the phenomenon of polymorphism was described and the first organic compound for which the structures of both polymorphs were determined in 1938 by Robertson and Ubbelohde.(1,2) Until today, the resorcinol crystals belong to the best known examples of polymorphs.(3)

CCDC: 1913232
CCDC: 1913233
CCDC: 1913234
CCDC: 1913235
CCDC: 1913236
CCDC: 1913237
CCDC: 1913238
CCDC: 1913239
CCDC: 1913240
CCDC: 1913241
CCDC: 1913242

Revising data

- You can revise your data prior to publication
- During deposition:
  - If logged in:
    1. Click for revision
    2. Add deposition number to revise / select deposition number you wish to revise
  - If not logged in:
    1. Add deposition number to Deposition Number(s) for revision box
- Revised dataset will be assigned the same deposition number
- Old revisions will no longer be visible

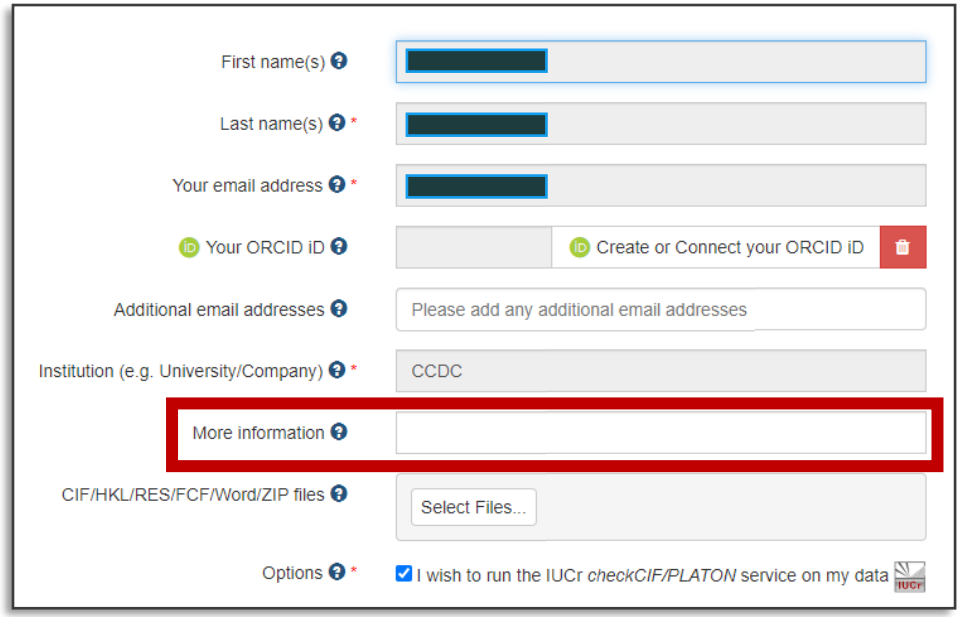




CCDC

An extra tip for logged-in users

Login → Upload



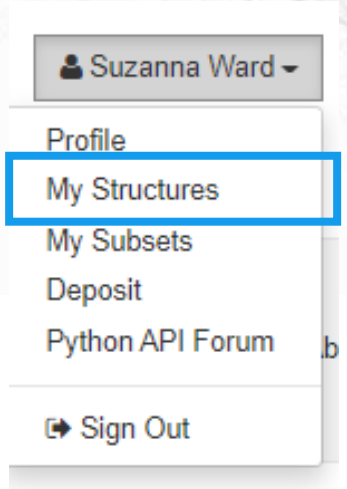
Text should only be added in the “More Information” box if you are resubmitting data or there are any reasons to consider before assigning you a deposition number.

Adding other information in this field will result in delays in processing your data.

CCDC

What happens to your data after deposition?

- It is stored privately until published in an associated scientific article or you decide to publish directly through the CSD or ICSD
- If you have registered on our website you can access your data at any time using our My Structures service
  - Enables you to view and retrieve datasets
  - Enables you to edit and update datasets
  - Enables you to share your datasets with co-workers
  - Enables you to create your own institutional database
  - Enables you to publish your data directly through the CSD/ICSD



- If you have not published your data after one year, we will email you to ask if you would like to publish it

CCDC

My Structures – view, edit, retrieve, share

My Structures

Search by Deposition Number | Clear Filters | Save Settings | Reset Settings

Suzanna Ward

- Profile
- My Structures
- My Subsets
- Deposit
- Python API Forum
- Sign Out

depositing structures they may take a few minutes to appear in the table. If you are still unable to see your structures after you have received your Deposition Numbers please act deposit@ccdc.cam.ac.uk

displayed columns or filter results you should click on the down arrow of the relevant column heading.
results by a particular column click on the column heading you wish to order your results by.
can select which structures are displayed by using the dropdown box below. Different subsets can be created using the 'Add to Subset' button and you can view or export these sets.

structures | Default Data View | My Downloads | Associate More Structures | Add to Subset | Share Structures

		Deposition …	Datablock	Deposited On	Deposited By	Refcode	Formula	Embargoed…	Status
☐	Details		data_tBu10kbar	28/04/2021			C11 H15 N1 O2	28/02/2022	Unpublished
☐	Details		data_5_93GPa	28/04/2021			C7 H7 N1 O2	28/02/2022	Unpublished
☐	Details	1584792	data_tBu10kbar	09/11/2017			C11 H15 N1 O2	29/11/2019	Unpublished
☐	Details	825801	data_2007may0…	06/06/2011	ward@ccdc.cam…	YIFTUD01	C9 H10 O3		Published in the CSD
☐	Details	825800	data_2007may0…	06/06/2011	ward@ccdc.cam…	EPAHOT	C8 H5 N1 O3		Published in the CSD

My Structure Details

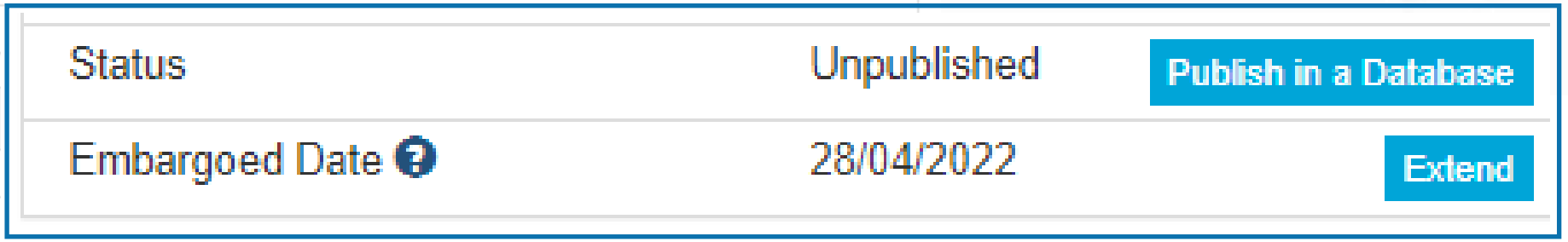
Datablock: tBu10kbar
Space Group: I2/a, **Cell:** *a* 14.811Å *b* 6.456Å *c* 19.759Å, *α* 90° *β* 94.060(8)° *γ* 90°
Formula: C11 H15 N1 O2, Temperature: 298 K

User Details

Reliability Score	3
2D Diagram	True
2D/3D Match	Full
2D/3D Last Edited	
User Compound Name	2-t-butyl-6-[(hydroxyimino)methyl]phenol
User Identifier	

Additional Details

Deposition Number		
Refcode		
Compound Name	2-t-butyl-6-[(hydroxyi ethyl]phenol	
Deposited On	28/04/2021	
Additional Depositors		
Status	Unpublished	Publish in a Database
Embargoed Date	28/04/2022	Extend



CCDC

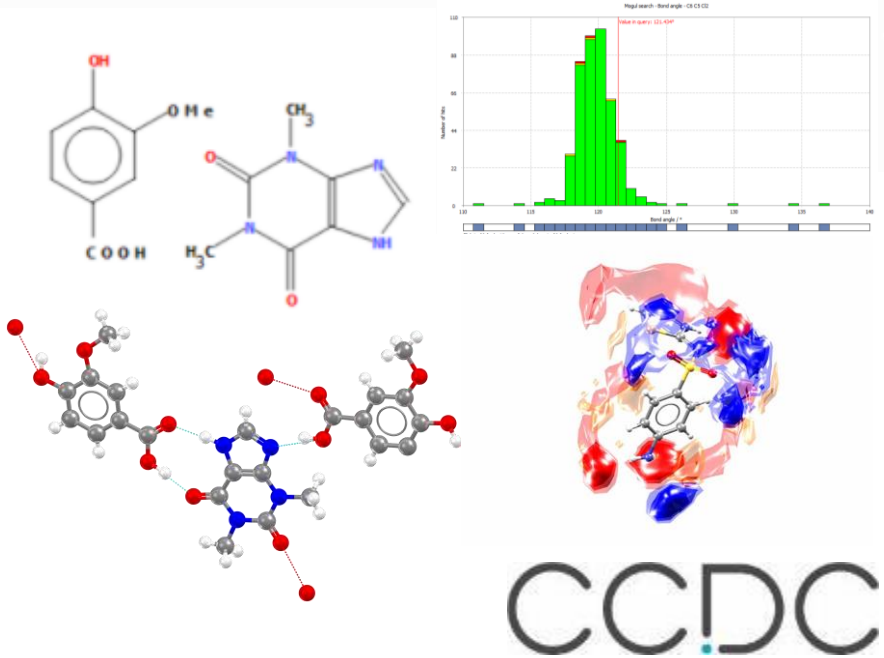
Curating the data into the CSD

After publication your data is enhanced and curated

- Every entry undergoes both automated and manual processes
- Enhanced and curated by experts at CCDC
- Ensuring accessibility, discoverability and re-usability of the data

```
data_b:

_audit_creation_method            SHELXL-97
_chemical_name_systematic         ?
_chemical_name_common             ?
_chemical_melting_point           ?
_chemical_formula_moiety          ?
_chemical_formula_sum             'C23 H23 N4 O10'
_chemical_formula_weight          515.45

loop_
_atom_type_symbol
_atom_type_description
_atom_type_scat_dispersion_real
_atom_type_scat_dispersion_imag
_atom_type_scat_source
C C 0.0033 0.0016 'International Tables Vol C '
H H 0.0000 0.0000 'International Tables Vol C '
N N 0.0061 0.0033 'International Tables Vol C '
O O 0.0106 0.0060 'International Tables Vol C '

_symmetry_cell_setting            Monoclinic
_symmetry_space_group_name_H-M    P21/c

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'

_cell_length_a                    11.2741(6)
_cell_length_b                    15.7712(10)
_cell_length_c                    13.3746(8)
_cell_angle_alpha                 90.00
_cell_angle_beta                  90.572(2)
_cell_angle_gamma                 90.00
_cell_volume                      2378.0(2)
_cell_formula_units_Z             4
_cell_measurement_temperature     173(2)
_cell_measurement_reflns_used     29773
_cell_measurement_theta_min       1.81
_cell_measurement_theta_max       28.33

loop_
_atom_site_label
_atom_site_type_symbol
_atom_site_fract_x
_atom_site_fract_y
_atom_site_fract_z
_atom_site_U_iso_or_equiv
_atom_site_adp_type
_atom_site_occupancy
_atom_site_symmetry_multiplicity
_atom_site_calc_flag
_atom_site_refinement_flags
_atom_site_disorder_assembly
_atom_site_disorder_group
O5 O 0.89248(12) 0.23961(7) 0.36204(8) 0.0332(3) Uani 1 1 d . . .
C17 C 0.80189(14) 0.28620(9) 0.21135(11) 0.0254(3) Uani 1 1 d . . .
O3A O 0.41130(12) 0.68432(8) -0.38659(9) 0.0385(3) Uani 1 1 d . . .
N14 N 0.71955(13) 0.32448(9) 0.06384(10) 0.0312(3) Uani 1 1 d . . .
O6 O 0.86638(12) 0.52322(7) 0.29698(9) 0.0373(3) Uani 1 1 d . . .
N12 N 0.79157(12) 0.43468(8) 0.17806(10) 0.0273(3) Uani 1 1 d . . .
N10 N 0.87464(12) 0.38185(8) 0.32964(10) 0.0270(3) Uani 1 1 d . . .
N16 N 0.77053(13) 0.21410(8) 0.15937(10) 0.0292(3) Uani 1 1 d . . .
O2A O 0.54917(16) 0.33725(9) -0.18177(12) 0.0606(5) Uani 1 1 d . . .
O2A O 0.62375(15) 0.42935(9) -0.07228(11) 0.0530(4) Uani 1 1 d . . .
C13 C 0.76968(14) 0.35252(9) 0.15071(11) 0.0248(3) Uani 1 1 d . . .
C5A C 0.50784(15) 0.63428(11) -0.23618(12) 0.0315(4) Uani 1 1 d . . .
C9 C 0.85852(14) 0.29649(9) 0.30524(11) 0.0254(3) Uani 1 1 d . . .
C11 C 0.84449(15) 0.45099(9) 0.26871(12) 0.0274(3) Uani 1 1 d . . .
C15 C 0.72239(16) 0.24010(11) 0.07303(12) 0.0329(4) Uani 1 1 d . . .
H15 H 0.6931 0.2024 0.0232 0.039 Uiso 1 1 calc R . .
C22 C 0.93191(17) 0.40270(10) 0.42587(13) 0.0351(4) Uani 1 1 d . . .
H59 H 0.9575 0.3503 0.4591 0.053 Uiso 1 1 calc R . .
H61 H 0.8754 0.4327 0.4684 0.053 Uiso 1 1 calc R . .
H60 H 1.0010 0.4390 0.4142 0.053 Uiso 1 1 calc R . .
O4A O 0.36611(14) 0.53273(9) -0.44926(10) 0.0331(4) Uani 0.812(3) 1 d P A 1
C3A C 0.42736(15) 0.53758(10) -0.35989(12) 0.0290(3) Uani 1 1 d . . .
C2A C 0.46563(15) 0.46975(10) -0.30318(12) 0.0307(4) Uani 1 1 d . A .
H2A H 0.4509 0.4135 -0.3255 0.037 Uiso 1 1 calc R . .
C6A C 0.54612(15) 0.56586(11) -0.17955(12) 0.0313(4) Uani 1 1 d . A .
H6A H 0.5563 0.5752 -0.1177 0.038 Uiso 1 1 calc R . .
C4A C 0.44943(15) 0.62071(10) -0.32675(12) 0.0281(3) Uani 1 1 d . A .
C1A C 0.52603(15) 0.48357(11) -0.21283(12) 0.0308(4) Uani 1 1 d . . .
C34 C 0.76722(17) 0.50614(10) 0.11077(14) 0.0374(4) Uani 1 1 d . . .
H50 H 0.7951 0.4924 0.0435 0.056 Uiso 1 1 calc R . .
H52 H 0.8086 0.5567 0.1355 0.056 Uiso 1 1 calc R . .
```

Author(s)	A.Jacobs, F.M.A.Noa
Reference	CrystEngComm (2015), 17, 98
Publication DOI	10.1039/C4CE01795A
Deposition	CCDC 1022107
Formula	$2(C_8H_8O_4)$, $C_7H_8N_4O_2$
Compound	bis(4-Hydroxy-3-methoxybenzoic acid) 1,3-dimethyl-3,7-dihydro-1H-purine-2,6-dione
Synonym	bis(Vanilic acid) theophylline
Spacegroup	**Name:** P21/c **Number:** 14
Cell	**a:** 11.274(<1) **b:** 15.771(1) **c:** 13.375(<1) **alpha:** 90.00 **beta:** 90.57(<1) **gamma:** 90.00 **Volume:** 2377.967
Reduced Cell	**a:** 11.274 **b:** 13.375 **c:** 15.771 **alpha:** 90.00 **beta:** 90.00 **gamma:** 90.57 **Volume:** 2377.967
Molecular Volume	594.492
Chemical Units	2
Z, Z'	**Z:** 4.0 **Z':** 1.0
R-Factor (%)	4.7
Disorder	C7A,O4A and C9A,O5A disordered over two sites with occupancies 0.812:0.188.
Temperature (K)	173
Density	**CCDC:** 1.443 **Author:** 1.44
Intensity Meas	diffractometer
Average Sigma (C-C)	0.001-0.005Å
Colour	colorless
Habit	rectangle
Recryst. Solvent	water



FAIR data and interoperability

InChI for Chemical Structures

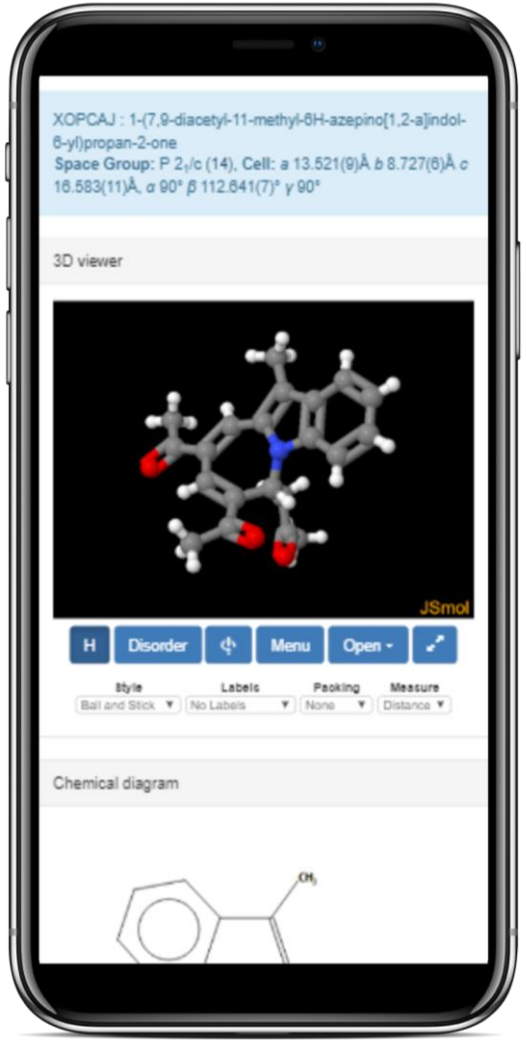
DOIs for Digital Objects

ORCID iDs for Researchers

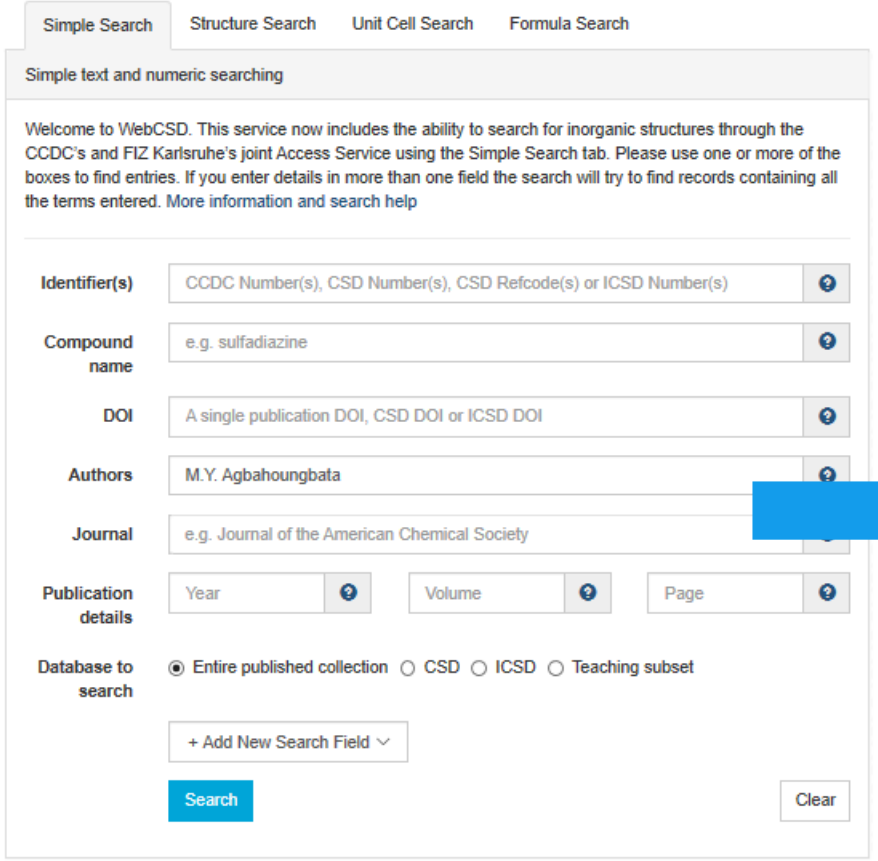
Wilkinson, M. D. et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci. Data 3:160018 DOI: 10.1038/sdata.2016.18 (2016)

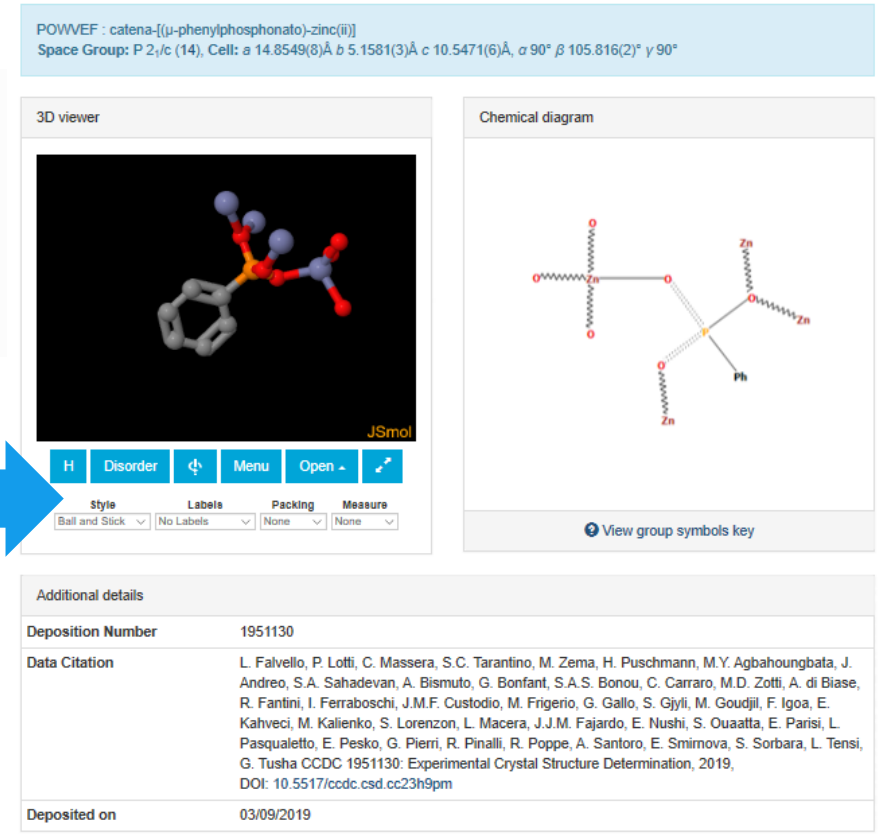
Access Structures

www.ccdc.cam.ac.uk/structures



- Access Structures allows you to search CSD and ICSD
- Every individual dataset is free to view and download





CCDC FIZ Karlsruhe Leibniz Institute for Information Infrastructure

CSD Entry: ZOYBEW

Sign In

Licensed to: CCDC Main Site

Simple Search | Structure Search | Unit Cell Search | Formula Search

Your query was: DOI: 10.1039/C4CE01795A and the search returned 7 records.

Modify Search | New Search

Results

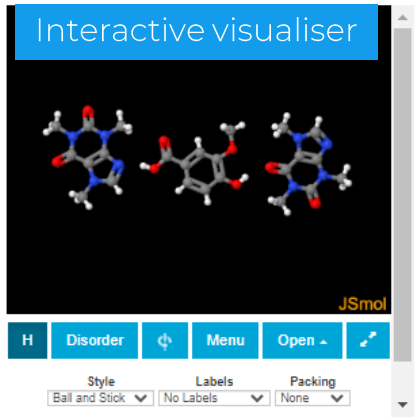
Database Identifier	Deposition Number
ZOYBEW	1022106
ZOYBIA	1022107
ZOYBOG	1022108
ZOYCAT	1022109
ZOYCEX	1022110
ZOYCIB	1022111
ZOYCOH	1022112

Download

ZOYBEW : 4-hydroxy-3-methoxybenzoic acid bis(1,3,7-trimethyl-3,7-dihydro-1H-purine-2,6-dione)
Space Group: P $\bar{1}$ (2), **Cell:** *a* 10.110(2)Å *b* 10.525(2)Å *c* 12.221(2)Å, *α* 77.69(3)° *β* 81.26(3)° *γ* 82.69(3)°

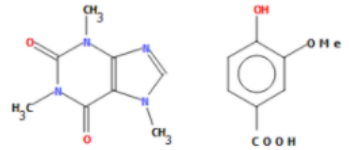
3D viewer

Interactive visualiser



Chemical diagram

Curated diagram



View group symbols key

Additional details

Deposition Number	1022106
Data Citation	Ayesha Jacobs, Francoise M. Amombo Noa CCDC 1022106: Experimental Crystal Structure Determination, 2014, DOI: 10.5517/cc139l55
Synonyms	vanillic acid bis(caffeine)
Deposited on	30/08/2014

Associated publications

Ayesha Jacobs, Francoise M. Amombo Noa, *CrystEngComm*, 2015, 17, 98, DOI: 10.1039/C4CE01795A

Identifier and key details. Banner colour coded depending on database

Additional chemical, crystal and experimental details

Chemical details

Formula	$2(C_8 H_{10} N_4 O_2),C_8 H_8 O_4$

Crystal details

Space group	P $\bar{1}$ (2)
Unit cell	*a* 10.110(2)Å *b* 10.525(2)Å *c* 12.221(2)Å *α* 77.69(3)° *β* 81.26(3)° *γ* 82.69(3)°
Cell volume	1249.69
Reduced cell	*a* 10.110Å *b* 10.525Å *c* 12.221Å *α* 77.690° *β* 81.260° *γ* 82.690°
Z, Z'	2, 1
Habit	rectangle
Recrystallisation solvent	Re-crystallisation from solvent: ethyl methylketone
Colour	colorless

Experimental details

R-factor (%)	5.42
Temperature (K)	173

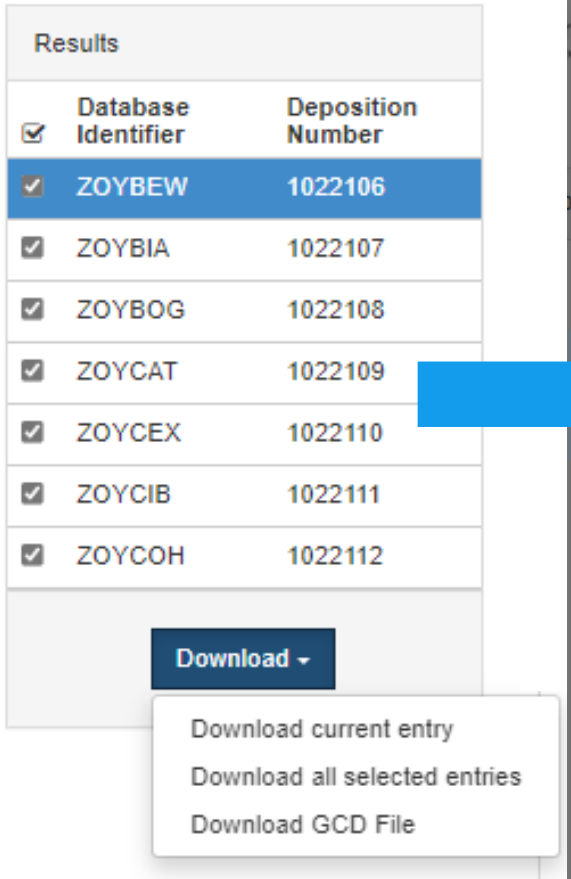
Click through identifiers

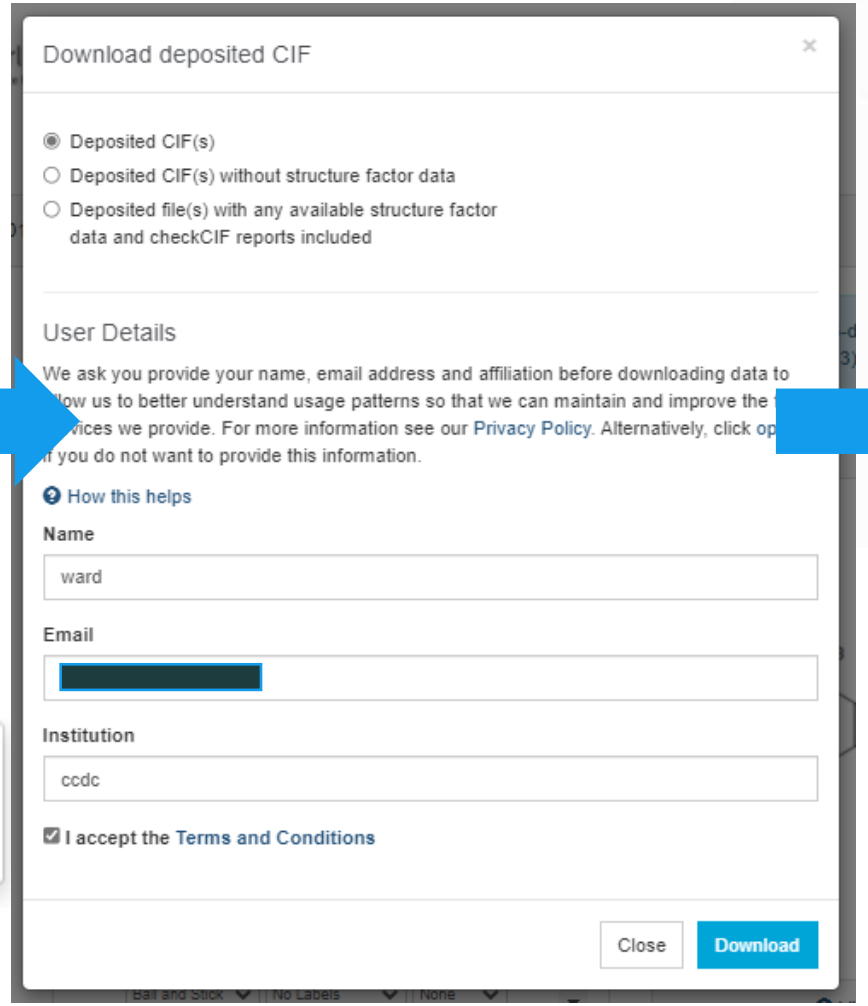
Data citation and DOI

Query highlighted and DOI links to publication

CCDC

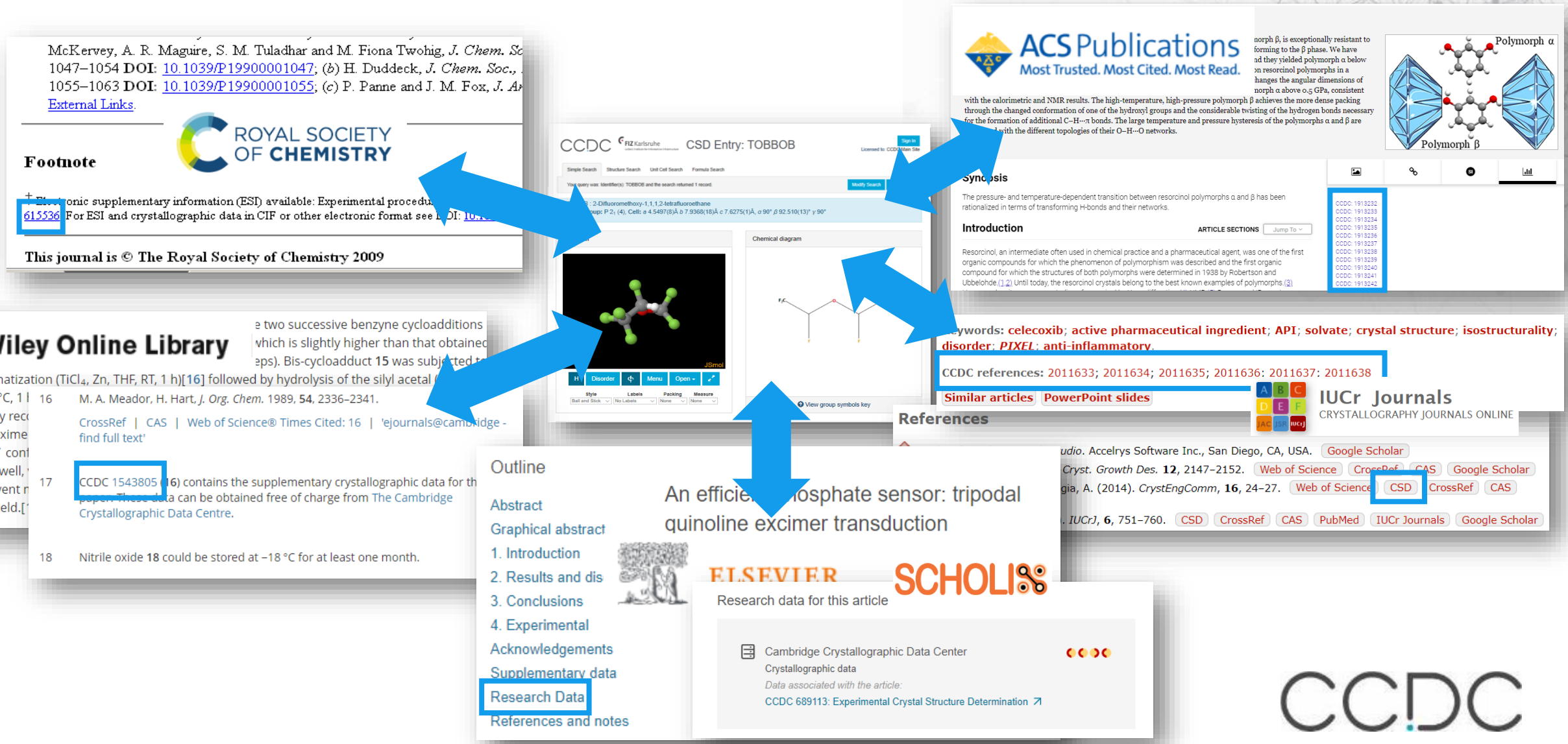
Retrieving and downloading files



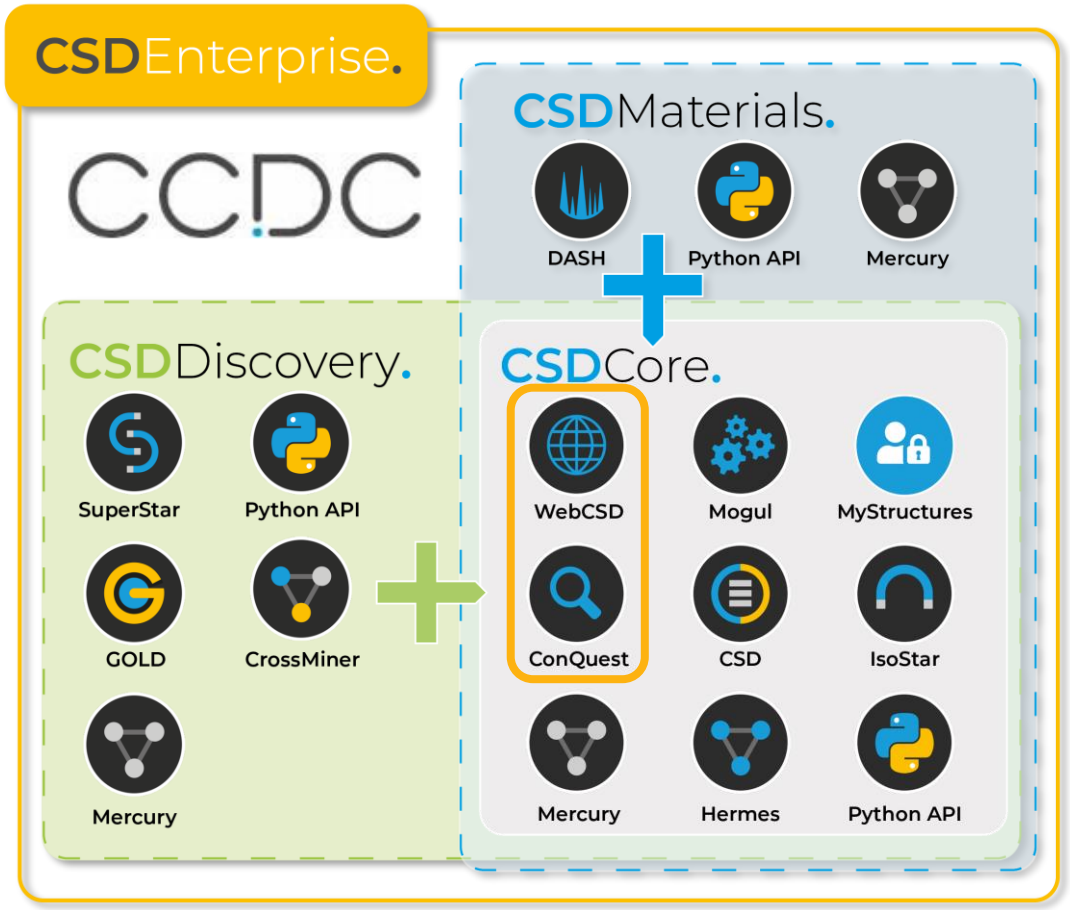


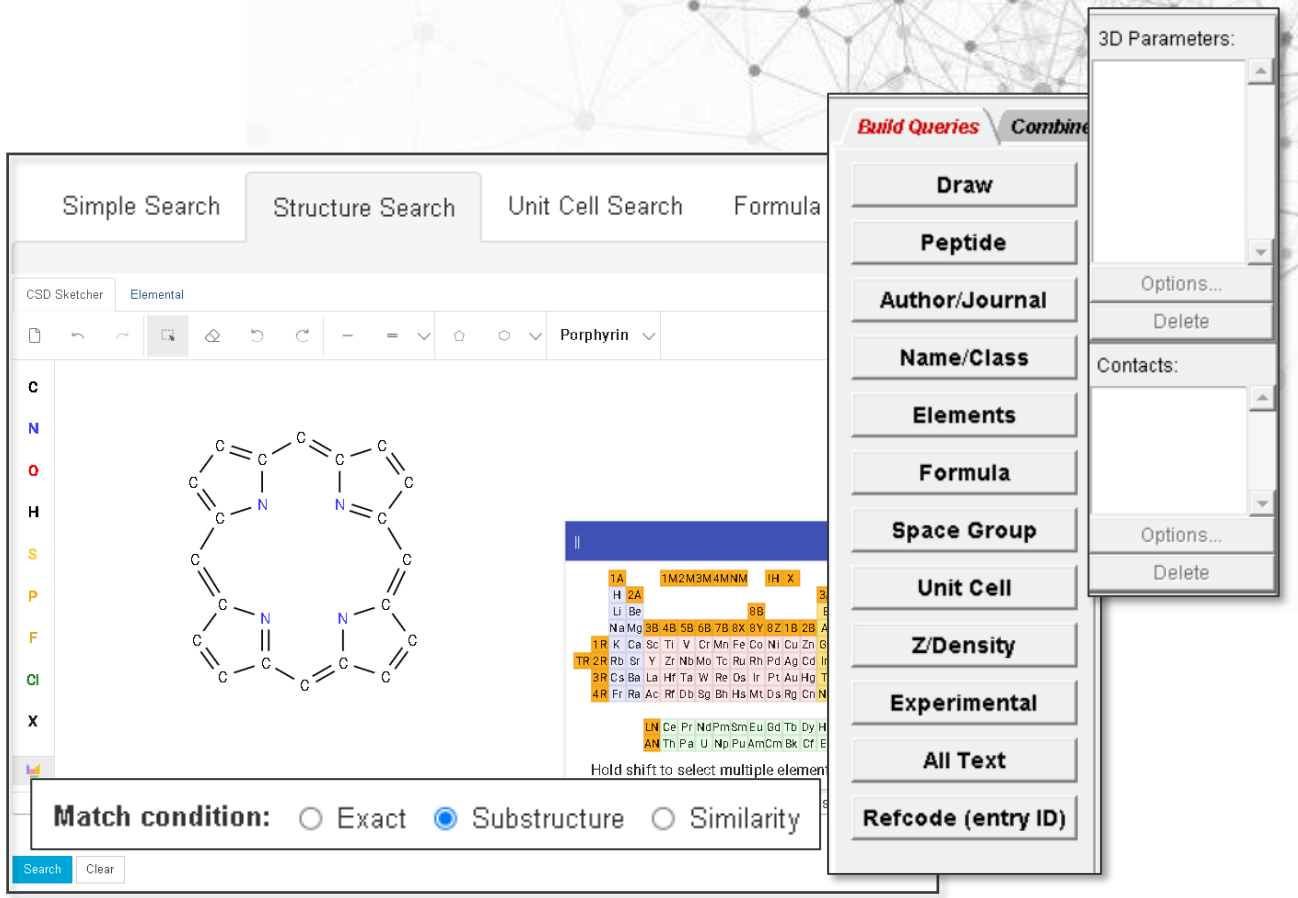
```
##############################################################################
#
# This file contains crystal structure data downloaded from the
# Cambridge Structural Database (CSD) hosted by the Cambridge
# Crystallographic Data Centre (CCDC).
#
# Full information about CCDC data access policies and citation
# guidelines are available at http://www.ccdc.cam.ac.uk/access/V1
#
# Audit and citation data items may have been added by the CCDC.
# Please retain this information to preserve the provenance of
# this file and to allow appropriate attribution of the data.
#
##############################################################################

data_b:
_audit_block_doi                 10.5517/cc139166
_database_code_depnum_ccdc_archive 'CCDC 1022107'
loop_
itation_id
ation_doi
ation_year
 10.1039/C4CE01795A 2015
_audit_update_record
;
2014-08-30 deposited with the CCDC.     2021-01-22 downloaded from the CCDC.
;

_audit_creation_method
_chemical_name_systematic
;
 (bis) vanillic acid theop
;
_chemical_name_common
_chemical_melting_point
_chemical_formula_moiety
_chemical_formula_sum
_chemical_formula_weight

loop_
_atom_type_symbol
_atom_type_description
_atom_type_scat_dispersion_real
_atom_type_scat_dispersion_imag
_atom_type_scat_source
C C 0.0033 0.0016 'International Tables Vol C Tables 4.2.6.8 and 6.1.1.4'
H H 0.0000 0.0000 'International Tables Vol C Tables 4.2.6.8 and 6.1.1.4'
N N 0.0061 0.0033 'International Tables Vol C Tables 4.2.6.8 and 6.1.1.4'
O O 0.0106 0.0060 'International Tables Vol C Tables 4.2.6.8 and 6.1.1.4'

_symmetry_cell_setting           Monoclinic
_symmetry_space_group_name_H-M   P21/c

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'

_cell_length_a                   11.2741(6)
_cell_length_b                   15.7712(10)
_cell_length_c                   13.3746(8)
_cell_angle_alpha                90.00
_cell_angle_beta                 90.572(2)
_cell_angle_gamma                90.00
_cell_volume                     2378.0(2)
_cell_formula_units_Z            4
_cell_measurement_temperature    173(2)
_cell_measurement_reflns_used    29773
_cell_measurement_theta_min      1.81
_cell_measurement_theta_max      28.33
```

From data to publication



Advanced search and analysis





CCDC

CSD Communications

Small molecule crystal structure data published just through the Cambridge Structural Database without a journal article

If you do not intend to publish your data in the scie
through the Cambridge Structural Database (CSD)
please click the 'Publish in a Database' button belc
CSD as a *CSD Communication*. Inorganic data wil

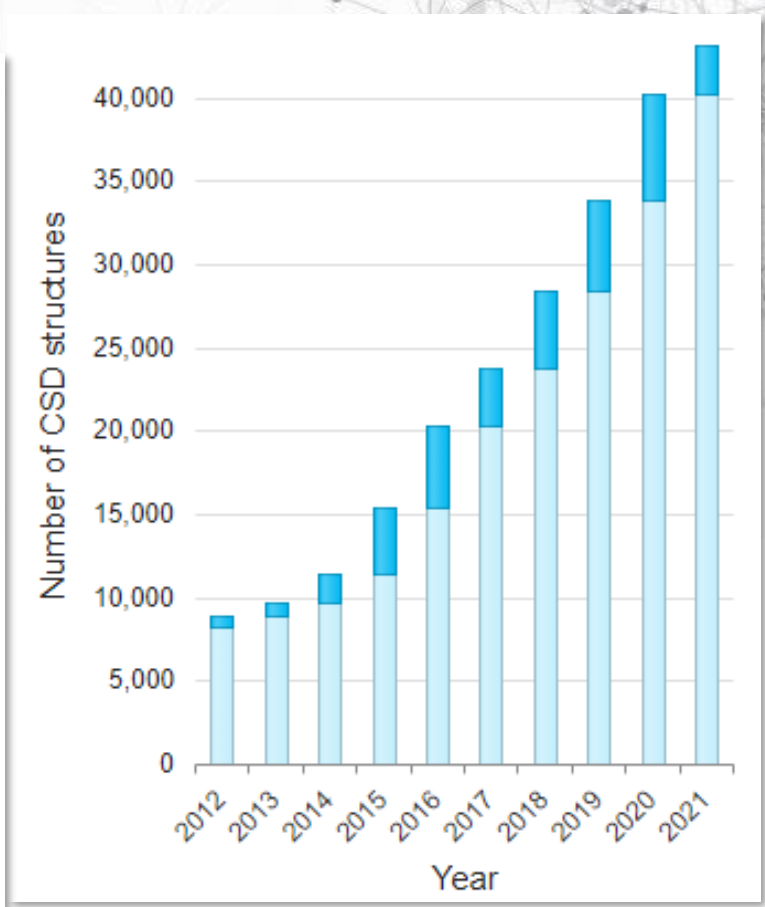
Publish in a Database

ISSN 2631-9888

Archive

The *CSD Communications* archive contains all of the data publish
(CSD) by the Cambridge Crystallographic Data Centre (CCDC) as
Communications). Select a year to view all *CSD Communications*

- 2021
- 2020
- 2019
- 2018



https://www.ccdc.cam.ac.uk/Community/csd-communications/

CCDC

Where to find more guidance

The CCDC CIF Deposition Guidelines

This page is designed to help you deposit your data, check your CIFs and provide you more information as you go through our deposition process. Click on the boxes below to navigate to the section of interest.

Basic guidelines	Required CIF fields	Authorship guidelines	Checking CIFs
Correcting CIFs	Useful resources	Structure Factors	Translated guidelines

Basic Guidelines

When preparing your CIF (What is a CIF?) for deposition please include as much information as possible and check it carefully. This is especially true for *CSD Communications* where there is no paper to describe the chemistry and experimental details leading to your structure.

If we are unable to validate your structure from the information you have provided we may contact you. If we cannot resolve the issue, unfortunately, we may not be able to add your structure to the CSD.

In this section

- The CCDC CIF Deposition Guidelines.pdf
- Structure factors: what are they and why share them?
- The Benefits of Data Sharing

Related Topics

- Deposit a Structure
- Structure Deposition Information
- Correcting CIFs
- CSD Communications

https://www.ccdc.cam.ac.uk/Community/depositastructure/cif-deposition-guidelines/

CCDC

Want to explore more?

Training and Educational Resources

The wealth of information contained within the Cambridge Structural Database (CSD) extends far beyond a collection of crystal structures. Knowledge derived from these materials informs much of chemistry, biochemistry, and biology. Chemical and structural concepts are often difficult to grasp without real world, interactive examples for students to explore.

The CCDC and our colleagues continually produce educational materials for use in classroom and computer lab settings, or as independent study modules. Many of these materials make use of the Teaching Subset - a freely available set of over 750 structures that can be investigated with the free version of our Mercury visualisation and analysis program. Of course, our database of over one million entries are available for free through our Access Structures portal.

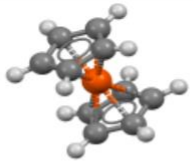
If you are an educator looking for supplementary teaching materials, find out more about the Teaching Database **here**. If you have developed your own modules using the CSD and would like to share them with the broader community, please contact us at education@ccdc.cam.ac.uk.

To keep up to date with the latest news from education and outreach at the CCDC, sign up for the Education and Outreach Newsletter **here**.

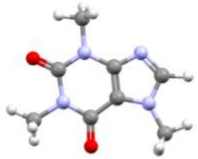
Register for E&O newsletter

CSD-Community

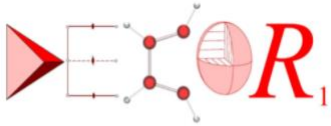
Freely accessible tools from the CCDC.



Information on the Teaching Subset



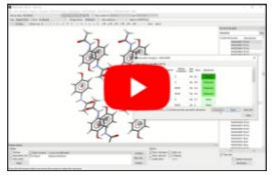
Access a series of teaching modules for use in the classroom



DECOR: Educational Resources for Teaching Crystallography

Self-guided workshops

Download a series of self-guided workshop materials for CCDC tools and features

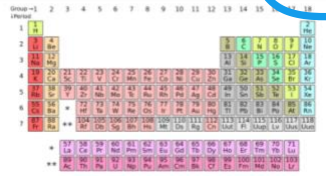


Watch software training and support videos

CSD University modules

On-demand modules with completion certificate

Access fun science activities for kids through the CCDC Home learning page



Explore the Periodic Table through Crystal Structures

YouTube and LabTube channels

CCDC