

# Structure Validation: Automation, Vigilance, New Tools

Anthony Linden

Institute of Organic Chemistry  
University of Zürich

Sandy Blake

School of Chemistry  
University of Nottingham UK



University of Zurich

Journals Commission Meeting, Madrid August 2011

## What is validation?

---

Comparison against normally expected values or conditions

- Are all the usual information and data present?
- Do related or derived parameters match?
- Do bonded atoms have compatible  $U^j$  values?
- Has the refinement converged?
- Is the space group correct?
- Are the assigned atom types correct?
- etc, etc, etc...

# Valid-ation

---

Correct

Appropriate

Defensible

## Why do we need *checkCIF*?

---

*checkCIF* introduced by the IUCr in 1997

Ongoing development by Ton Spek in *PLATON*

- Throughput of labs exploded in the CCD era
- Nice GUIs, but people often no longer look at output/log files
- More non-experts determining structures
- Help people avoid simple errors and oversights
- Encourage maintenance of quality standards (best practice)
- Increase publication success rate for authors (less revisions)
- Decrease publication times for journals

## Are validation and vigilance still needed?

---

- Many avoidable mistakes still appear in submitted or published papers
  - Inexperience
  - Complacency
  - Ignoring (lesser) validation Alerts
  - Do not understand Alerts
  - Blind reliance on *checkCIF* – if no Alert, then it must be OK
  - Conversely, blind reliance by reviewers – if there is an Alert, there must be a problem!

## *checkCIF* is...

---

- **A tool to help YOU...**
  - efficiently check your work
  - avoid blunders
  - follow best practice ideals
  - achieve the best result possible
- Not intended as a hurdle to make life tough
- Not intended to hinder publication of correct results
- Not intended to make you write long explanations for everything
  - scientists always document (non-routine) experimental procedures, don't they...?
- **Also a useful tool for (knowledgeable) reviewers**

## Current *checkCIF* and *PLATON* tests

---

- CIF syntax, missing information, data consistency and quality
- Unit cell & space-group symmetry
- (An)isotropic displacement parameters
- Intramolecular & intermolecular contacts
- Coordination-related issues
- Solvent-accessible voids
- Consistency of geometric parameters & s.u.'s
- Reflection data consistency, completeness, twinning
- and much more...

## Sources of outlier parameters

---

- Incorrect structure (*e.g.*, wrong space group or atom)
- Unresolved feature (*e.g.*, untreated disorder)
- Non-optimal procedures (*e.g.*, poor disorder modelling)
- Artefact resulting from limited data quality
- Special experimental conditions (document them)
- A genuinely unusual observation – worthy of discussion!



## Too many tests and Alerts?

---

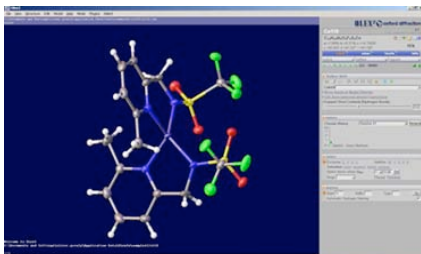
- When is an outlier important, when is it not?
- *E.g.*, use of *SQUEEZE*
  - Should formula should include estimate of the omitted solvent?
  - Alert A about voids, formula/model mismatch, molecular weight, F(000), density and absorption co-efficient.
  - OK if proper details in the CIF and/or experimental section.
  - In other cases, a formula/model mismatch might truly indicate forgotten atoms or a mistyped formula – important things.
- One Alert C might be insignificant. Several related C Alerts might indicate problems. Should all these be set to Alert A to gain attention?
- More CIF definitions needed for special cases – twinning, *SQUEEZE*.

## Automation

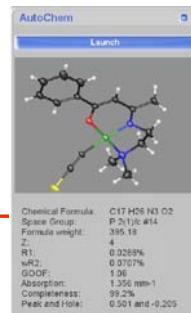
---

- New generation of fully-automated diffractometers
- Progress in automatic structure solution & refinement
- Manufacturers promise:
  - “No or little crystallographic knowledge required”
  - “Routine small molecule structure determination is accessible to students and scientists of other disciplines”

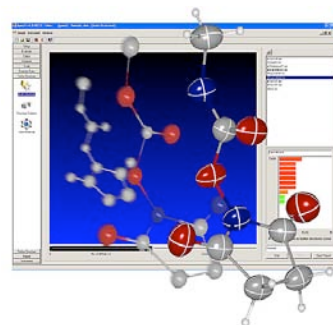




## Automation



- Drop in a crystal, push a button, sit back, and ...
- Pretty picture without further ado – if there are no Alerts, it must be OK ... right?
- Can a person with “no crystallographic knowledge” rely on that (yet)?
- Further checking of results seems essential (*e.g.*, element assignments)
- If the result is not the expected molecule, what happens then?



## Alert indicators


```
380 ALERT 4 C Likely Unrefined X(sp2)-Methyl Moiety ..... C18
412 ALERT 2 C Short Intra XH3 .. XHn : H19B .. H30A = 1.81 Ang.
720 ALERT 4 C Number of Unusual/Non-Standard Label(s) .... 1
```

Alert numbers 1-5 indicate the type of issue.

Alert levels A, B, C indicate the severity of the issue.

G is a general issue to check, not necessarily an error.

## Alert types



---

```
380 ALERT 4 C Likely Unrefined X(sp2)-Methyl Moiety ..... C18
412 ALERT 2 C Short Intra XH3 .. XHn : H19B .. H30A = 1.81 Ang.
720 ALERT 4 C Number of Unusual/Non-Standard Label(s) .... 1
```

ALERT Type 1 = CIF construction/syntax error, inconsistent or missing data

ALERT Type 2 = Indicator that the structure model may be wrong or deficient

ALERT Type 3 = Indicator that the structure quality may be low

ALERT Type 4 = Improvement, methodology, query or suggestion

ALERT Type 5 = Informative message, check

## Vigilance – additional to validation

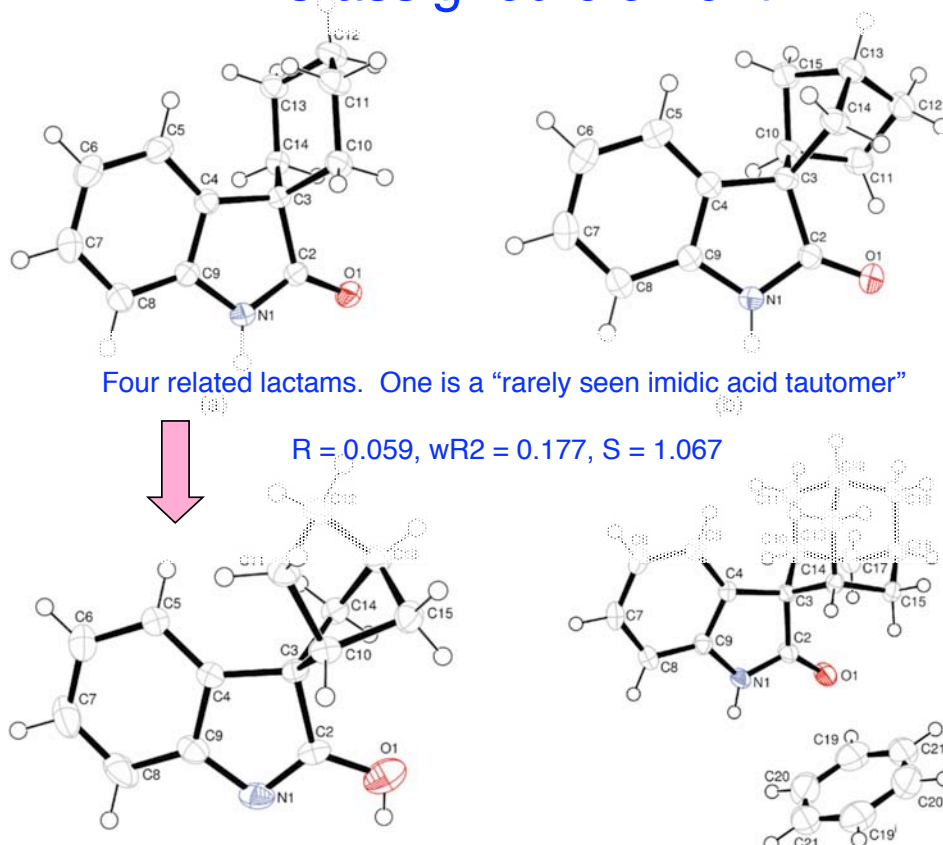
---

- Does the structure make sense to you?
- Does the structure look right and is it geometrically logical?
- Must be able to rationalise structure with the expected or plausible chemistry, etc.
- Don't force (restrain) a structure to be that which it is not.
- Does the geometry agree with similar structures in databases?
- Unusual geometry or other features are rarely a new property – more likely to be the effect of an inadequacy of the model
- Look critically at the output files (*e.g.*, .lst file)

## Possible limits to validation

- Test not (yet) implemented: high ADPs on isolated atom
- Test not practical: C–C range is 1.49 – 1.60 Å
- Error not a validation issue: “needle, 0.28 x 0.24 x 0.03 mm”
- Mistake cannot be detected from CIF data: wrong elements
- Nonsense entries in the CIF: see *Acta Cryst.* 2003, E59, e2

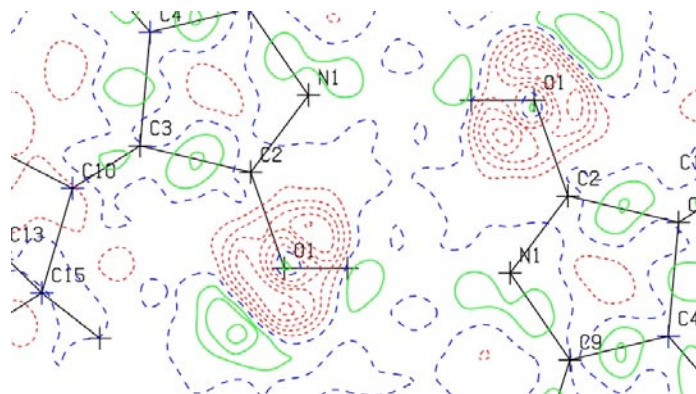
### Mis-assigned element



## 230\_ALERT\_2\_B Hirshfeld Test Diff for O1 -- C2 .. 11.83 su

### Peaks list

Q1	0.54	1.07	O1
Q2	0.28	0.77	C3
Q3	0.26	0.73	C3
Q4	0.25	0.76	C10



Contoured difference maps are very useful – easy in *PLATON*

## Refine as an amine

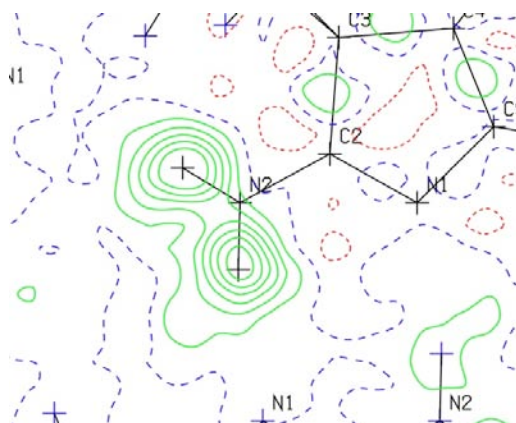
R = 0.046, wR2 = 0.117  
(formerly 0.059, 0.177)

No relevant Alerts

Q1 0.22      0.77 C3

**N2 is pyramidal.**

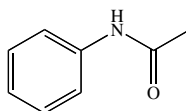
Do not fall into the trap of thinking it is planar (imine) and use AFIX 93!



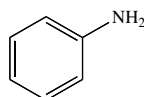
Now the chemist has work to do!

## Geometry of –NH and –NH<sub>2</sub> groups

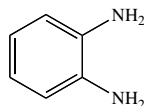
Amides: planar



Phenylamines: usually planar



Phenyldiamines: one of the amine groups may be pyramidal



Validation is not usually revealing.

Be careful about auto-calculation of H with amines and hydroxy groups.

Test the H-atom positions: refine the H-atoms, or refine their Uiso values.

Look at contoured difference maps.

## Missing H atom

The issue raises only an Alert G

343\_ALERT\_2\_G Check sp? Angle Range  
in Main Residue for .. C18

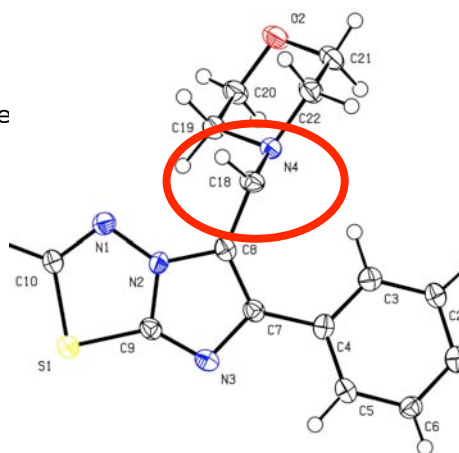
Largest peak: 0.84 e/Å<sup>3</sup>

H-atoms from diff. map and refined.

So one H was missed, but...

No mismatched formula!

Author claims that structure is fine because  
there is no serious *checkCIF* Alert



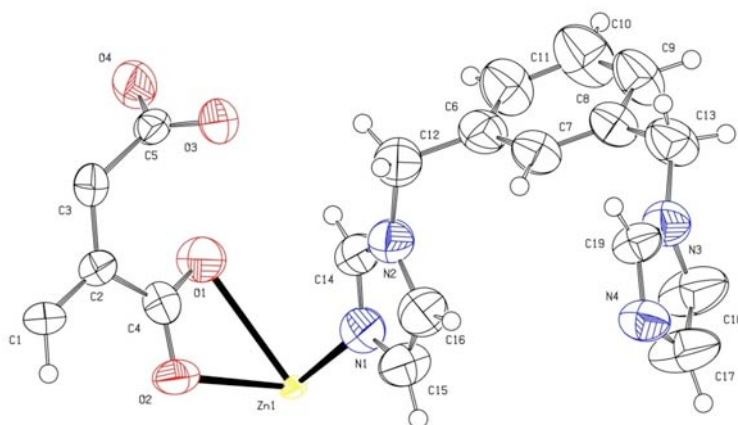
LOOK at and understand the structure AND the chemistry

## Another misassigned element?

---

Calculated  $\rho(\min) = -0.50$ ,  $\rho(\max) = 1.35$  e/Å<sup>3</sup>  
R = 0.0466, wR<sup>2</sup> = 0.1318, S = 1.042, Npar = 263, Flack = 0.20(3)

```
232_ALERT_2_B Hirshfeld Test Diff (M-X) Zn1 -- O2 .. 11.83 su
232_ALERT_2_B Hirshfeld Test Diff (M-X) Zn1 -- N1 .. 10.32 su
232_ALERT_2_C Hirshfeld Test Diff (M-X) Zn1 -- N4_b .. 8.65 su
094_ALERT_2_C Ratio of Maximum / Minimum Residual Density .. 2.47
```



## Another misassigned element?

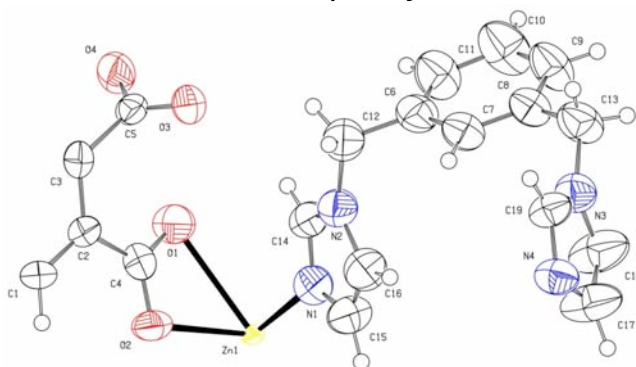
---

Zn complexes known to generate Hirshfeld Alerts  
– Lutz, M. & Spek, A. L. (2009). *Acta Cryst.* **C65**, m69

Replace Zn with Cd \_\_ R = 0.033, Hirshfeld Alerts now C level

Refine occupancy for Cd \_\_ 0.89. A lighter element?

With Rh, R = 0.030, no Alerts, occupancy 0.97



## Zn, Cd, Rh or Ru?

---

Chemically, a mix-up of Cd with Zn seems more likely than Rh or Ru

Chemist swears that it is a Zn complex! So should we believe it?

What else can be checked?

M–N and M–O bond lengths – compare with related structures in CSD

In structure: M–N = 2.25, M–O = 2.25–2.44 Å

In CSD: Zn–N = 2.0, Zn–O = 2.1–2.3 Å

Cd–N = 2.3, Cd–O = 2.3 Å

## What's wrong here?

---

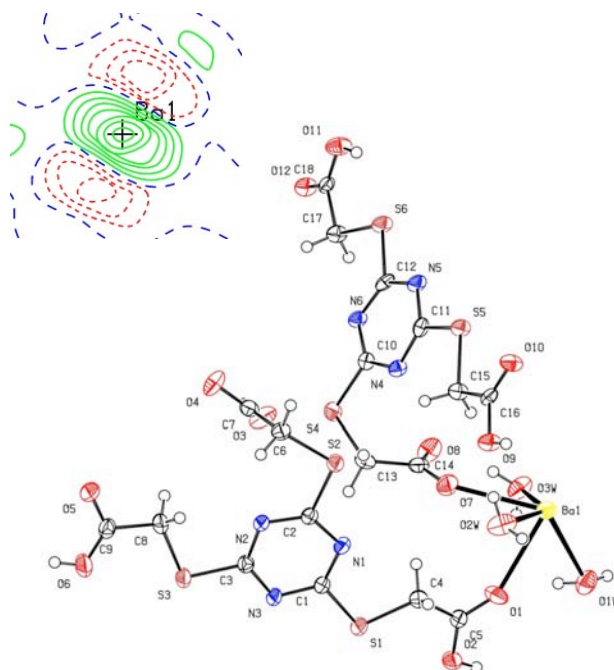
R = 0.047, wR = 0.088,  
shift/error 0.000

Looks reasonable visually, but...

Large peak near Ba: 3.4 e/Å<sup>3</sup>  
**Alert A**

Hard to be sure about hydroxy &  
water H-atoms – diff. maps quite  
noisy

Alert C about poor Ba–O–H angles  
for one water



## Refine again

---

R = 0.035, down from 0.047 !!

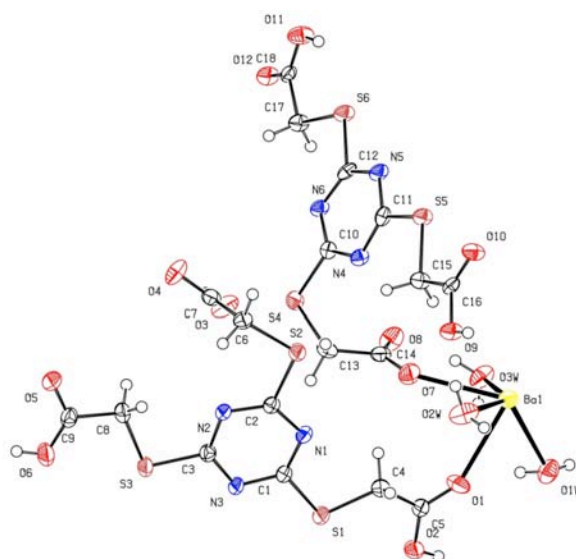
Ba ellipsoid shrinks (large shift/error)

Residual peaks gone

Diff. maps clean and H atoms clear

Using element other than Ba did not reproduce author's "converged" result

Why???



## Improper use of DAMP 0 0

---

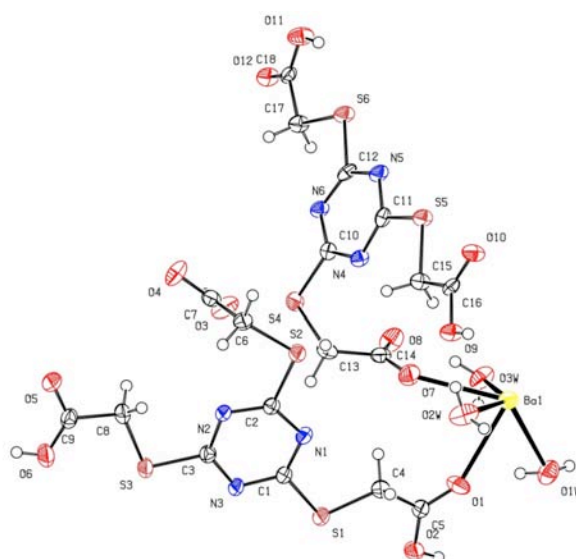
Structure had not yet converged

DAMP 0 0 sets shifts to zero!

Use ONLY for GCLS refinement  
AFTER full convergence  
(to generate s.u.s)

NEVER for L.S. refinement

GCLS should not usually be  
needed for final refinement of  
small-molecule structures



# Structure Factor Validation

(fcf validation)

## What can fcf validation detect?

---

- Mismatch between the data block names in the CIF and .fcf file
- Mismatch between cell parameters in the CIF and .fcf file
- The .fcf file is not from the refinement that produced the CIF
- Incomplete updating of a CIF (*e.g.*, weighting scheme)
- Overlooked twinning
- Atomic coordinates transformed, but not the  $U^j$
- Incorrect element assignment (supplements other tests)
- Element reassignment without re-refining
- Modifying atomic and displacement parameters in the CIF (cheating!)

## When might fcf validation not work?

- Weighting scheme cannot be interpreted by PLATON
  - e.g., Refining with JANA (sometimes), CRYSTALS, RAEELS
- Non-merohedral twins
  - SHELXL HKLF5 type input cannot be reconstructed from the .fcf file
- .fcf file is in a format not understood by PLATON

The screenshot shows the PLATON software interface. At the top, the title 'PLATON' is displayed in large, colorful letters, followed by the subtitle 'A Multipurpose Crystallographic Tool' and the copyright information '(C) 1980-2010 A.L. Spek - 40M-Version: 30310'. Below this is a menu with several categories: GRAPHICS, GEOM-CALC, VOIDS FLIP, SYMMETRY, ABSORPTION, REPORT, and MISC-TOOLS. Each category lists various sub-options. On the right side, there is a 'PLATON MENU' with a list of options including OptionMenu, NoMove, Join-Expand, Organic, Round, Parentheses, Label-Alias, R/S-Determ, Norm-H-bond, NoSymm, NoDisorder, ListARU RCell, ListCellSymm, ListAtoms, ListBonds, ListFlagRadi, Exclude H, MinQPeakHgt, MinQPeakDis, Q-Peak-Incl, KeyInstruct, Prev Next, SAVE-InstrS, ENTRY-LIST, and Reset End. At the bottom, there is a status bar with the text 'WORKING'.

GRAPHICS	GEOM-CALC	VOIDS FLIP	SYMMETRY	ABSORPTION	REPORT	MISC-TOOLS
PLUTONauto	Calc ALL	Calc Solv	ADDSYM	MULscanABS	ValidatLon	SYSTEM-S
GATEP/ADP	Calc Intra	Calc K.P. I	ADDSYM-EQL	ABSPslScan	ASYM-VIEW	FCF2HKL
NewmanPlat	Calc Inter	SQUEEZE	ADDSYM-EXT	ABSTempa	FCF-Valld	Expand2P1
ALng-Plots	Calc Coord	CALC-FCF	ADDSYM-PLT	ABSGauss	Dl fFaurLer	FCF-Gener
PLane-Plat	Calc Metal	Contour-SQ	ADDSYM-SHX	ABSXtal	ANALofVAR	HKL-Gener
Pol yhedra	Calc Geom	SOLV F3D	NEWSYM	ABSSphere	ByvoetPatr	HKL-Transf
ContourDlf	Calc Hbond	SOLV PLOT	NONSYM	SHXABS	ASYM-EXPCT	EXOR-RES
Contour-Fa	Calc TMA	CavLtyPlat	LePage		ASYM-Valld	ANIS-RES
AutoMolFlt	L.S.-PLANE		DelRed		SupplMater	Rename-RES
HKL2Powder	DlhedAngle		MOLSYM		EXPECT-HKL	Auto-Renum
SLmPowderP	AngleLLnes	FLIP MENU	SPGRframEX		CSD-CELL	SPF -el d
RadDlstFun	AngLsplLLn	FLIP SHOW	ASYM		CSD-QUEST	SHELXL-res
Patterson	CremerPopl	FLIP PATT	ASYMaverFR		StructTldy	CIF -acc
PLUTONatLv	BondValenc	FLIPPER 25	LePageTwn		StralnAnal	PDB -pdb
	HFIX - RES	STRUCTURE?	TwnRotMat	Xtal Hablt	LocCIF-acc	CIF2SHELXL
Xtal Data (CIF ) gg3227.cif- Set 1( 1): I						
RefL Data (SHELXL ) gg3227.hkl [ NO-DIAC] ( 1): I						


# checkcif.iucr.org

---

**checkCIF**

A service of the  
International Union of Crystallography

**checkCIF reports** on the consistency and integrity of crystal structure determinations reported in CIF format.

Please upload your CIF using the form below. 

File name:

Select form of checkCIF report:

HTML  PDF

Select validation type:

Full validation of CIF and structure factors  
 Validation of CIF only (no structure factors)

---


Information about this version of checkCIF ...

**Useful links**

- Details of checkCIF/PLATON tests
- CIF dictionary
- Full publication check
- Download CIF editor (pubCIF) from the IUCr
- Download CIF editor (enCIFer) from the CCDC

checkCIF is sponsored by


**Crystallography Journals Online**




**ELSEVIER**

**WILEY-BLACKWELL**


**RSC** Advancing the Chemical Sciences








# journals.iucr.org/services/cif/checking/checkfull.html

 **checkCIF/PLATON (standard) - Mozilla Firefox**

File Edit View History Bookmarks Tools Help

 checkCIF/PLATON (standard)

   [http://vm02.iucr.org/cgi-bin/checkcif\\_hkl.pl](http://vm02.iucr.org/cgi-bin/checkcif_hkl.pl)

 UoN Chemistry News  RSS Feed for <http://for...>

---

**checkCIF/PLATON (standard)**

---

File name of structure factor file (in CIF format) on your local filesystem:

---

Download CIF editor (pubCIF) from the IUCr  
Download CIF editor (enCIFer) from the CCDC  
Test a new CIF entry

[journals.iucr.org/services/cif/checking/checkcifhkl.html](http://journals.iucr.org/services/cif/checking/checkcifhkl.html)

checkCIF/PLATON (standard) - Mozilla Firefox

checkCIF/PLATON (standard)

Structure factors have been supplied for datablock(s) I

checkCIF

A service of the  
International Union of Crystallography

No syntax errors have been found. A PDF version of the  
checkCIF report may be downloaded from the link that will  
appear below shortly.

Download checkCIF report

Structure factor report

Useful links  
Details of checkCIF/PLATON tests  
CIF dictionary  
Full publication check

checkCIF is sponsored by  
Crystallography  
Journals  
Online  
ELSEVIER  
WILEY-BLACKWELL  
RSC  
Advancing the  
Chemical Sciences

Download CIF editor (pubCIF) from the IUCr  
Download CIF editor (enCIFer) from the CCDC  
Test a new CIF entry

## “Normal” feedback in the absence of a .fcf file (or if .fcf file not recognised)

```
PLATON/CHECK-( 30310) versus check.def version of 260210 for entry: I
Data From: x.cif - Data Type: CIF          Bond Precision   C-C = 0.0040 A
                                           Temp = 93 K
UCL  3.7797(17)   10.362(3)   11.297(3)   62.92(2)   82.30(3)   87.60(4)
WaveLength 0.71075   Volume Reported   390.3(2)   Calculated   390.3(2)
SpaceGroup from Symmetry P -1           Hall: -P 1
                Reported P -1           -P 1
MoietyFormula C12 H12 N2, 2(F3 H2 O2 V)
                Reported C12 H12 N2 2+ , 2H2 F3 O2 V -
SumFormula C12 H16 F6 N2 O4 V2
                Reported C12 H16 F6 N2 O4 V2
Mr           =   468.15[Calc],   468.15[Rep]
Dx,gcm-3    =   1.992[Calc],   1.992[Rep]
Z           =   1[Calc],       1[Rep]
Mu (mm-1)  =   1.288[Calc],   1.290[Rep]
F000       =   234.0[Calc],   234.0[Rep]   or F000' =   234.76[Calc]
Reported T Limits: Tmin=0.864           Tmax=1.000   AbsCorr=MULTI-SCAN
Calculated T Limits: Tmin=0.857 Tmin'=0.857 Tmax=0.975
Reported Hmax= 4, Kmax= 12, Lmax= 13, Nref= 1385 , Th(max)= 25.340
Calculated Hmax= 4, Kmax= 12, Lmax= 13, Nref= 1435 , Ratio = 0.965
Reported Rho(min) = -0.34, Rho(max) = 0.36 e/Ang**3 (From CIF)
w=1/[sigma**2(Fo**2)+(0.0393P)**2+ 0.0941P], P=(Fo**2+2*Fc**2)/3
R= 0.0329( 1215), wR2= 0.0800( 1385), S = 1.081, Npar= 126
```

## Enhanced feedback in the presence of a .fcf file

---

```
***
# Reported      Hmax=  4, Kmax= 12, Lmax= 13, Nref= 1385      , Th(max)= 25.340
# Obs in FCF   Hmax=  4, Kmax= 12, Lmax= 13, Nref= 1385      , Th(max)= 25.341
# Calculated   Hmax=  4, Kmax= 12, Lmax= 13, Nref= 1435      , Ratio  = 0.965
# Reported     Rho(min) = -0.34, Rho(max) =  0.36 e/Ang**3 (From CIF)
# Calculated   Rho(min) = -0.35, Rho(max) =  0.30 e/Ang**3 (From CIF+FCF data)
# w=1/[sigma**2(Fo**2)+(0.0393P)**2+ 0.0941P], P=(Fo**2+2*Fc**2)/3
# R= 0.0329( 1215), wR2= 0.0800( 1385), S = 1.081          (From CIF+FCF data)
# R= 0.0329( 1215), wR2= 0.0800( 1385), S = 1.081          (From FCF data only)
# R= 0.0329( 1215), wR2= 0.0800( 1385), S = 1.081, Npar= 126
```

## Structure Factor Validation Output

---

```
Reported      Hmax=  4, Kmax= 12, Lmax= 13, Nref= 1385, Th(max)= 25.340
Obs in FCF    Hmax=  4, Kmax= 12, Lmax= 13, Nref= 1385, Th(max)= 25.341
Calculated    Hmax=  4, Kmax= 12, Lmax= 13, Nref= 1435, Ratio  = 0.965

Reported     Rho(min) = -0.34, Rho(max) =  0.36 e/Ang**3 (From CIF)
Calculated   Rho(min) = -0.35, Rho(max) =  0.30 e/Ang**3 (From CIF+FCF data)

w=1/[sigma**2(Fo**2)+(0.0393P)**2+ 0.0941P], P=(Fo**2+2*Fc**2)/3

R= 0.0329( 1215), wR2= 0.0800( 1385), S = 1.081          (From CIF+FCF data)
R= 0.0329( 1215), wR2= 0.0800( 1385), S = 1.081          (From FCF data only)
R= 0.0329( 1215), wR2= 0.0800( 1385), S = 1.081, Npar= 126

From CIF + FCF data = recalculated from Fobs and atomic coordinates + U's
in CIF

From FCF data only = recalculated solely from Fobs, Fcalc and weights

From CIF = reported values in the CIF
```

## Structure Factor Validation Output

---

### Mismatch between the data block names in the CIF and .fcf file

```
900_ALERT_1_A No Matching Reflection File Found ..... !
902_ALERT_1_A No (Interpretable) Reflections found in FCF .... !
```

### Mismatch between cell parameters in the CIF and .fcf file

```
901_ALERT_1_A Cell Parameters in CIF and FCF do not Match .... !
902_ALERT_1_A No (Interpretable) Reflections found in FCF .... !
```

In both cases, no extended summary of R-factors, delta-Rho, etc.

## Weighting scheme in CIF not understood

---

```
808_ALERT_5_G No Parsable SHELXL Style Weighting Scheme Found !
929_ALERT_5_G No Weight Pars,Obs and Calc R1,wR2,S not checked !
```

May cause mismatches in the extended summary of R-factors, residual electron density peaks, etc., but no other Alerts in this regard.

Little to be done if not using SHELXL.

New CIF datanames to robustly document weights are needed;  
the current dataname is a free text item.

## Twining detected

---

```
931_ALERT_5_G Check Twin Law ( 1 0 0)[      ] Estimated BASF 0.18
931_ALERT_5_G Check Twin Law (      ) [ 3 0 1] Estimated BASF 0.17
```

Alert may be generated even if twinning has been handled.

Twining may also cause mismatches in the extended summary of R-factors, e.g. where non-merohedral twinning is treated with the HKLF5 method.

CIF datanames to enable proper reporting and validation of twins urgently needed – currently under development.

## The .fcf file is not from the same refinement as the CIF

---

A water molecule was omitted from the refinement used to generate the .fcf file, but the finished model is in the CIF

```
Reported Rho(min) = -0.34, Rho(max) = 0.36 e/Ang**3 (From CIF)
Calculated Rho(min) = -1.18, Rho(max) = 10.08 e/Ang**3 (From CIF+FCF data)
w=1/[sigma**2(Fo**2)+(0.0393P)**2+ 0.0941P], P=(Fo**2+2*Fc**2)/3

R= 0.1442( 1215), wR2= 0.2787( 1385), S = 4.255 (From CIF+FCF data)
R= 0.2189( 1215), wR2= 0.5046( 1385), S = 7.612 (From FCF data only)
R= 0.0329( 1215), wR2= 0.0800( 1385), S = 1.081, Npar= 126

973_ALERT_2_A Large Calcd. Positive Residual Density on V1 10.08 eA-3
971_ALERT_2_B Large Calcd. Non-Metal Positive Residual Density 3.14 eA-3
921_ALERT_1_A R1 * 100.0 in the CIF and FCF Differ by ..... -18.60
922_ALERT_1_A wR2 * 100.0 in the CIF and FCF Differ by ..... -42.46
923_ALERT_1_A S values in the CIF and FCF Differ by ..... -6.53
925_ALERT_1_A The Reported and Calculated Rho(max) Differ by . 9.72 eA-3
926_ALERT_1_A Reported and Calculated R1 * 100.0 Differ by . -11.13
927_ALERT_1_A Reported and Calculated wR2 * 100.0 Differ by . -19.87
928_ALERT_1_A Reported and Calculated S value Differ by . -3.17
```

## Weights not updated in CIF after a new refinement

### A common fault!!

---

```
R= 0.0329( 1215), wR2= 0.0640( 1385), S = 1.849 (From CIF+FCF data)
R= 0.0329( 1215), wR2= 0.0640( 1385), S = 1.848 (From FCF data only)
R= 0.0329( 1215), wR2= 0.0800( 1385), S = 1.081, Npar= 126
```

```
923_ALERT_1_A S values in the CIF and FCF Differ by ..... -0.77
922_ALERT_1_B wR2 * 100.0 in the CIF and FCF Differ by ..... 1.60
927_ALERT_1_B Reported and Calculated wR2 * 100.0 Differ by . 1.60
928_ALERT_1_B Reported and Calculated S value Differ by . -0.77
```

## Weights not updated in CIF after a new refinement

### A common fault!!

---

```
_refine_ls_structure_factor_coef Fsqd
_refine_ls_matrix_type full
_refine_ls_weighting_scheme calc
_refine_ls_weighting_details
'w = 1/[\s^2^(Fo^2^)+(0.0393P)^2^+0.0941P] where P=(Fo^2^+2Fc^2^)/3'
_atom_sites_solution_primary direct
_atom_sites_solution_secondary difmap
_atom_sites_solution_hydrogens geom
_refine_ls_hydrogen_treatment mixed
_refine_ls_extinction_method none
_refine_ls_extinction_coef ?
_refine_ls_number_reflns 1385
_refine_ls_number_parameters 126
_refine_ls_number_restraints 2
_refine_ls_R_factor_all 0.0395
_refine_ls_R_factor_gt 0.0329
_refine_ls_wR_factor_ref 0.0800
_refine_ls_wR_factor_gt 0.0765
_refine_ls_goodness_of_fit_ref 1.081
```

## Improper editing of a CIF

---

- Atomic coordinates transformed through a symmetry operation other than inversion, but not the  $U^j$ 
  - always re-refine and generate a new CIF, avoid piecemeal cut/paste or hand-editing of the CIF itself.
- Element reassignment without re-refining
- Modifying atomic and displacement parameters in the CIF (to hide things)

Such manipulations lead to mismatches of R-factors, goodness-of-fit and residual electron density.

## Summary

---

- *checkCIF* is a tool for YOU as author/referee/co-editor
- Be vigilant – do not rely solely on *checkCIF*
- Structure factor validation is also very important
- Watch out for errors that validation may not detect

For proper review, referees need the fcf files!

How many non-IUCr journals require their submission?  
(this used to be a rhetorical question)

How many wrong structures are missed because a journal does not require structure factor submission?

*checkCIF* development: Ton Spek (Utrecht) & Mike Hoyland (IUCr Chester office)



**Portland Building, University of Nottingham**