Why
Crystal Structure Validation ?

Ton Spek,

National Single Crystal Facility,
Utrecht University,
Utrecht, The Netherlands

Slovenia, 17-june-2010.

Overview of this Lecture

- Why Crystal Structure Validation ?
- What are the Validation Questions ?
- How is Validation Implemented ?
- What key problems are addressed ?
- Some Examples of Detected Problems.
- Evaluation and Performance.
- Summary.

Why Crystal Structure Validation ?

- The explosion of Reported Structure Determinations every year.
- Many analyses are done nowadays Black-Box style by non-specialists.
- There is a limited number of experts/referees trained and available to detect common pitfalls in publications.
- Validation offers a list of ALERTed (i.e. unusual) issues that require special attention of the analyst, the specialist and the referee.
- Validation tries to be helpful and sets quality standards.
- New and sadly: Detection of clear fraud and fraudulent practices.
- ➔ Recent Literature Example of Poor Refereeing

Has any Referee Looked at this Recent Paper ?

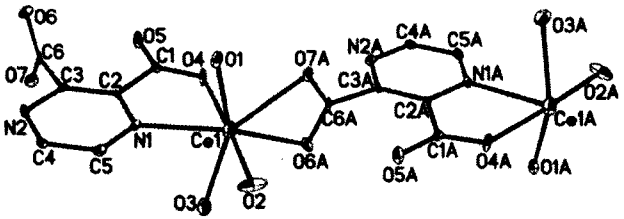
Z. Kristallogr. NCS **225** (2010) 71-73 / **DOI** 10.1524/ncrs.2010.0029

Crystal structure of triaqua-(pyrazine-2,3-dicarboxylato)cerium(III), $Ce(H_2O)_3(C_6H_2N_2O_4)$

Xun Feng*,I, Jian-She ZhaoII, Xin-Ge ShiI, Fang RuanI and Li-Ya WangI

I Luoyang Normal University, College of Chemistry and Chemical Engineering, Luoyang 471022, Henan, P. R. China
II Northwest University, Key Laboratory of Synthetic and Natural Functional Molecule Chemistry, Ministry of Education, College of Chemistry and Materials Science, Xi'an 710069, P. R. China

Received May 12, 2009, in revised form June 17, 2009, accepted and available on-line November 17, 2009; CCDC no. 1267/2654



Abstract

$C_6H_8CeN_2O_7$, orthorhombic, $P2_12_12_1$ (no. 19), $a = 5.7479(8)$ Å, $b = 10.909(2)$ Å, $c = 15.370(2)$ Å, $V = 963.7$ Å^3, $Z = 4$, $R_{gt(F)} = 0.045$, $wR_{ref(F^2)} = 0.121$, $T = 296$ K.

lanthanide complexes based on 2,3-pyrazonedicarboxylic acid, just several examples were previously reported [9]. Because the lanthanide ions generally have higher coordination numbers than the *d*-block metal ions, it is more likely to incorporate ancillary ligands into lanthanide coordination networks. Such ancillary ligands may be removed keeping the lanthanide-organic framework intact, thereby generating porous solids with coordinatively unsaturated lanthanide ions that may have potential catalytic activity.

The asymmetric unit of the title complex consists of one Ce(III) cation, one 2,3-pyrazinedicarboxylate ligand, and three coordinated water molecules. 4,4′-Bipy is not present in the final product. Ce(III) is a hepta-coordinated [CeN_1O_6], and exhibits a slightly distorted pentagonal bipyramidal coordination sphere. Two oxygen atoms of chelated deprotonated carboxylate group

Ce(III), Ce(II) or Something else ?

The CSD Reports an Isomorphous Cd(II) Complex
Ma et al. (2006). Acta Cryst. E62, m32528-m2529

$[Ce(C_6H_2N_2O_4)(H_2O)_3]n$

$P2_12_12_1$

a = 5.7479(8)

b = 10.909(2)

c = 15.370(2)

0.37 x 0.25 x 0.17 mm

R1 = 0.045, wR2 = 0.121

Ce1-O7a = 2.399(6)

Ce1-O6a = 2.406(6)

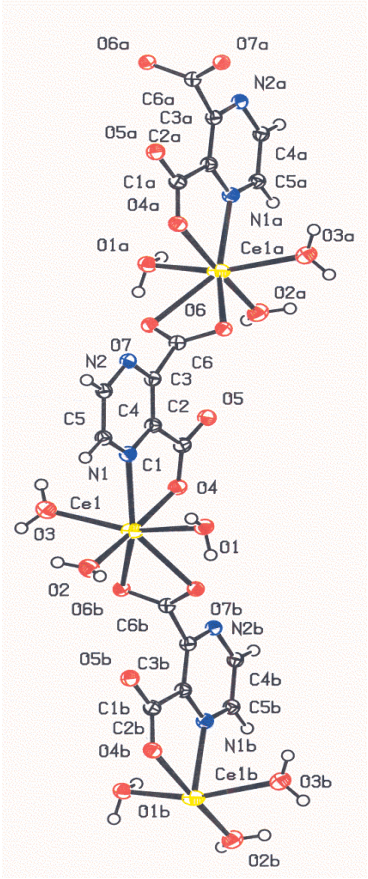
Ce1-O4 = 2.295(7)

Ce1-N1 = 2.359(8)

Ce1-O1 = 2.520(6)

Ce1-O2 = 2.212(8)

Ce1-O3 = 2.397(8)



$[Cd(C_6H_2N_2O_4)(H_2O)_3]n$

$P2_12_12_1$

a = 5.7365(12)

b = 10.903(3)

c = 15.362(3)

0.37 x 0.35 x 0.27 mm

R1 = 0.034, wR2 = 0.090

Cd1-O1a = 2.398(5)

Cd1-O2a = 2.398(5)

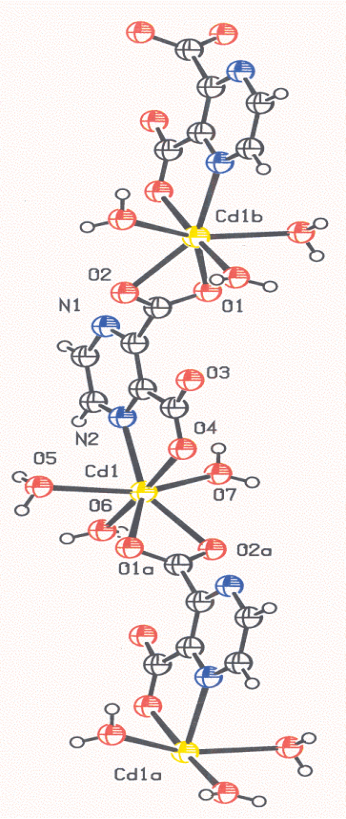
Cd1-O4 = 2.287(5

Cd1- N2 = 2.346(5)

Cd1-O7 = 2.530(5)

Cd1-O6 = 2.210(5)

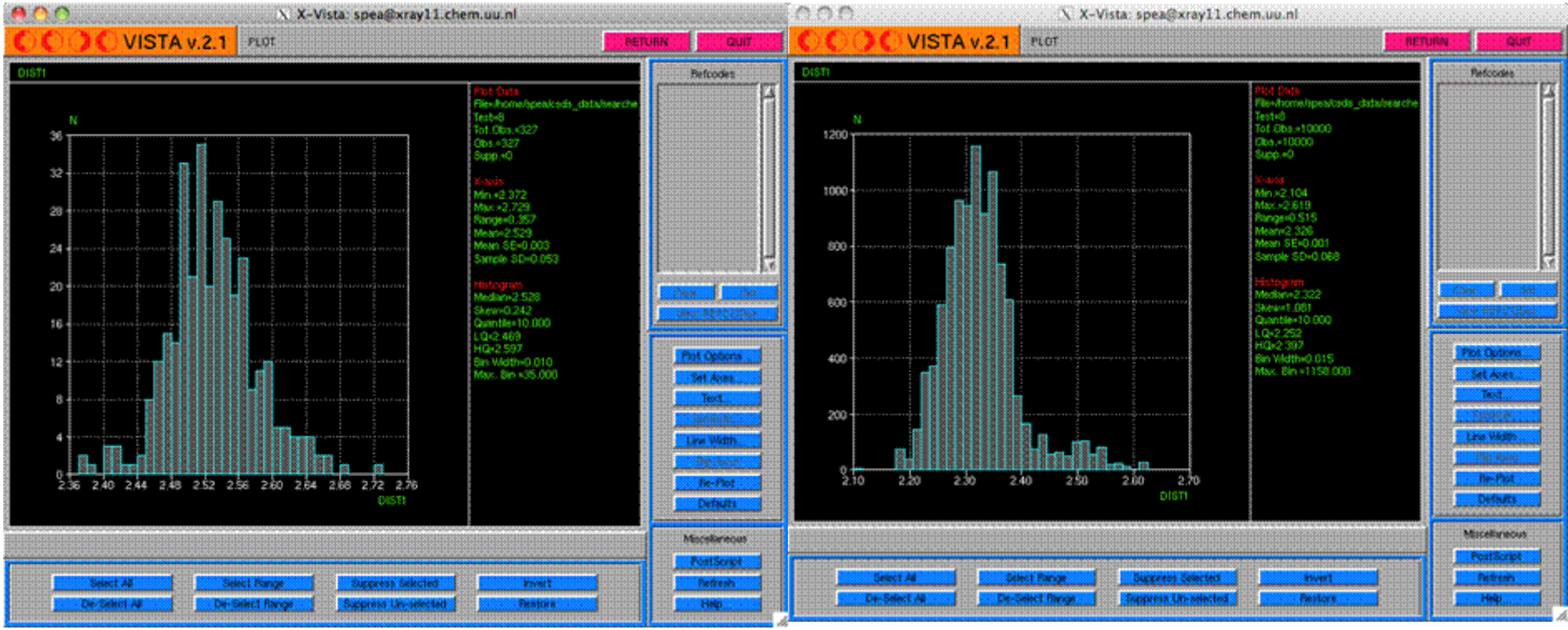
Cd1-O5 = 2.387(5)



Consult CSD/VISTA for Expected X-Ow Distances

Ce-Ow Distances

Cd-Ow Distances



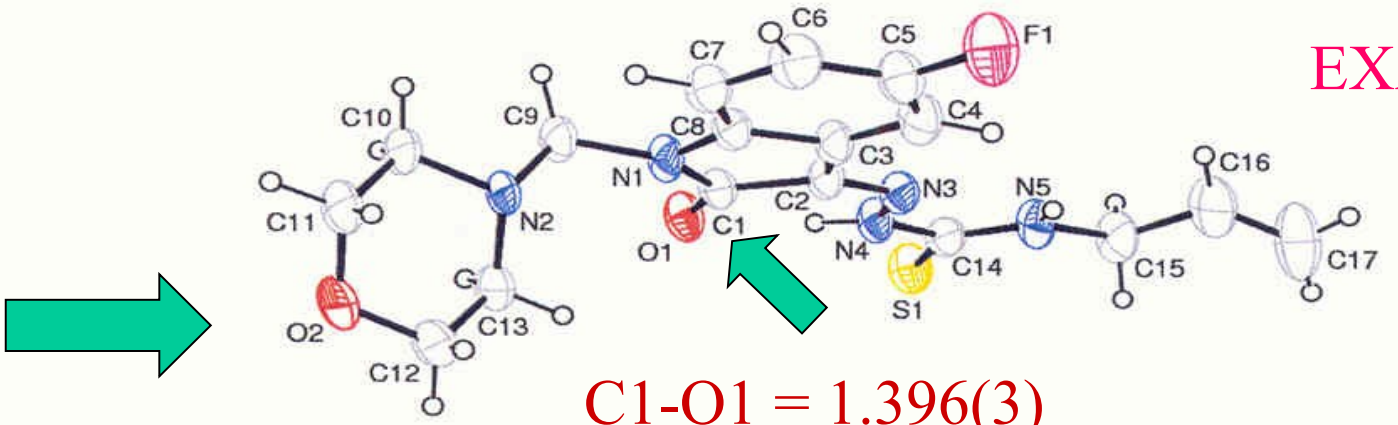
Conclusion: This is a Cd complex and not a Ce complex ?
The authors claim the use of Ce(NO3) in the synthesis !?
We need the (absent) reflection data for an absolute proof !

Just two Examples of problems with entries archived in the CSD

- The CSD is a rich source of chemical information.
- However: An analysis of the > 500000 structures in the CSD learns that a not insignificant number of the entries has undetected serious errors.
- Nearly all searches in the CSD for statistical info show (e.g. with VISTA) outliers that, when inspected closely, can be shown to be erroneous.
- The following two problem cases were detected as part of a search for short S…S contacts.

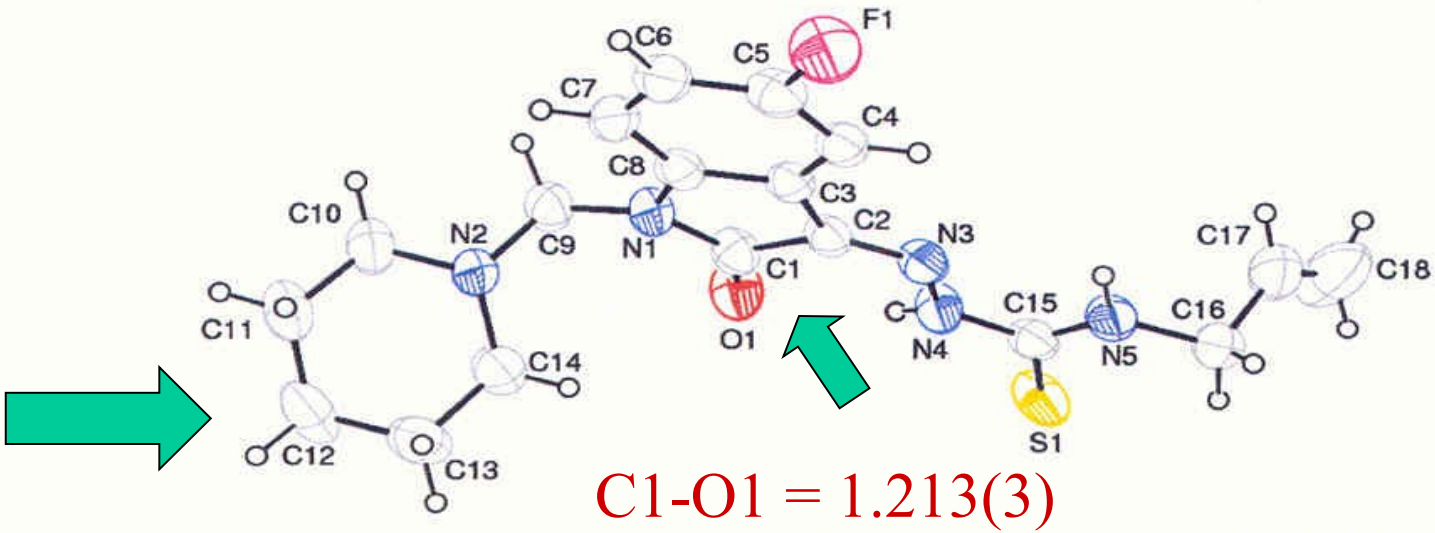
Two Related Structures – Strange Metrical Differences

EXAMPLE 1



C1-O1 = 1.396(3)

Figure 2. Molecular structure of **3e** showing the atom labeling scheme. Displacement ellipsoids are drawn as 30% probability level.



C1-O1 = 1.213(3)

Figure 3. Molecular structure of **3f** showing the atom labeling scheme. Displacement ellipsoids are drawn as 30% probability level.

Huge Geometry Differences !?

Table 2. Selected bond lengths and angles (°, Å) for **3e** and **3f**

3e		3f	
C1–O1	1.396(3)	C1–O1	1.213(3)
C1–N1	1.313(3)	C1–N1	1.367(3)
C1–C2	1.612(4)	C1–C2	1.509(4)
C2–N3	1.163(3)	C2–N3	1.291(3)
C2–C3	1.575(4)	C2–C3	1.448(3)
C5–F1	1.567(4)	C5–F1	1.355(3)
C8–N1	1.630(4)	C8–N1	1.409(3)
C9–N2	1.253(3)	C9–N2	1.436(3)
C9–N1	1.478(3)	C9–N1	1.459(3)
C10–N2	1.423(3)	C10–N2	1.463(3)
C10–C11	1.302(4)	C10–C11	1.486(4)
C11–O2	1.481(4)	C11–C12	1.503(5)
C12–O2	1.357(4)	C12–C13	1.516(5)
C12–C13	1.311(4)	C13–C14	1.500(4)
C13–N2	1.532(4)	C14–N2	1.471(3)
C14–N4	1.245(3)	C15–N5	1.323(3)
C14–N5	1.451(4)	C15–N4	1.369(3)
C14–S1	1.864(3)	C15–S1	1.669(3)
C15–C16	1.645(5)	C16–C17	1.440(4)
C15–N5	1.327(3)	C16–N5	1.456(3)
C16–C17	1.238(5)	C17–C18	1.201(5)
N3–N4	1.492(3)	N3–N4	1.341(3)

EXAMPLE 1

There is obviously a problem with 3e:
Where were the referees of this paper ?

TIKRUB

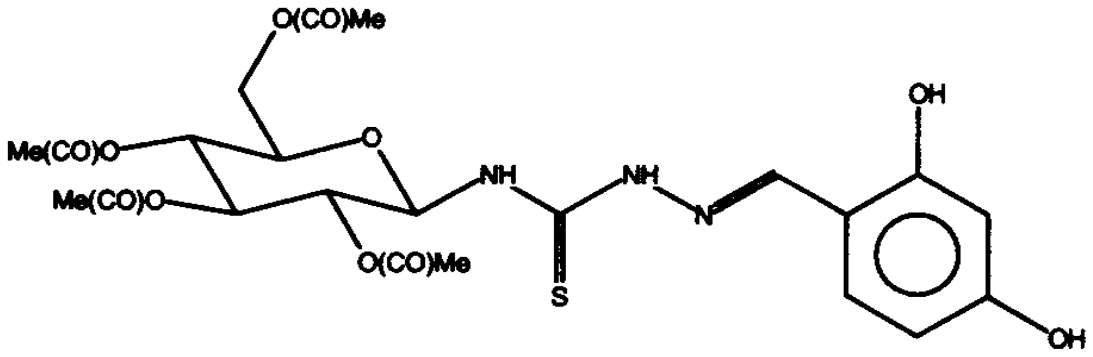
Reference:	Xue-Mei Li, Su-Juan Ye, Xu-Li Tang, Cai-Feng Ding, Shu-Sheng Zhang (2006) *Asian J.Chem.* ,18,491
Formula:	$C_{22}\ H_{27}\ N_3\ O_{11}\ S_1, H_2\ O_1$
Compound Name:	N-(2,3,4,6-Tetra-O-acetyl-β-D-glucopyranosyl) thiocarbamic 2,4-dihydroxybenzoyl hydrazine monohydrate

Space Group:	P43212	***Cell:***	***a***	12.020(0)	***b***	12.020(0)	***c*** 38.877(3)
Space Group No.:	96	***(Å,°)***	α	90.00	β	90.00	γ 90.00
R-Factor (%):	6.47	***Temperature(K):***	293		***Density(g/cm³):***		1.323

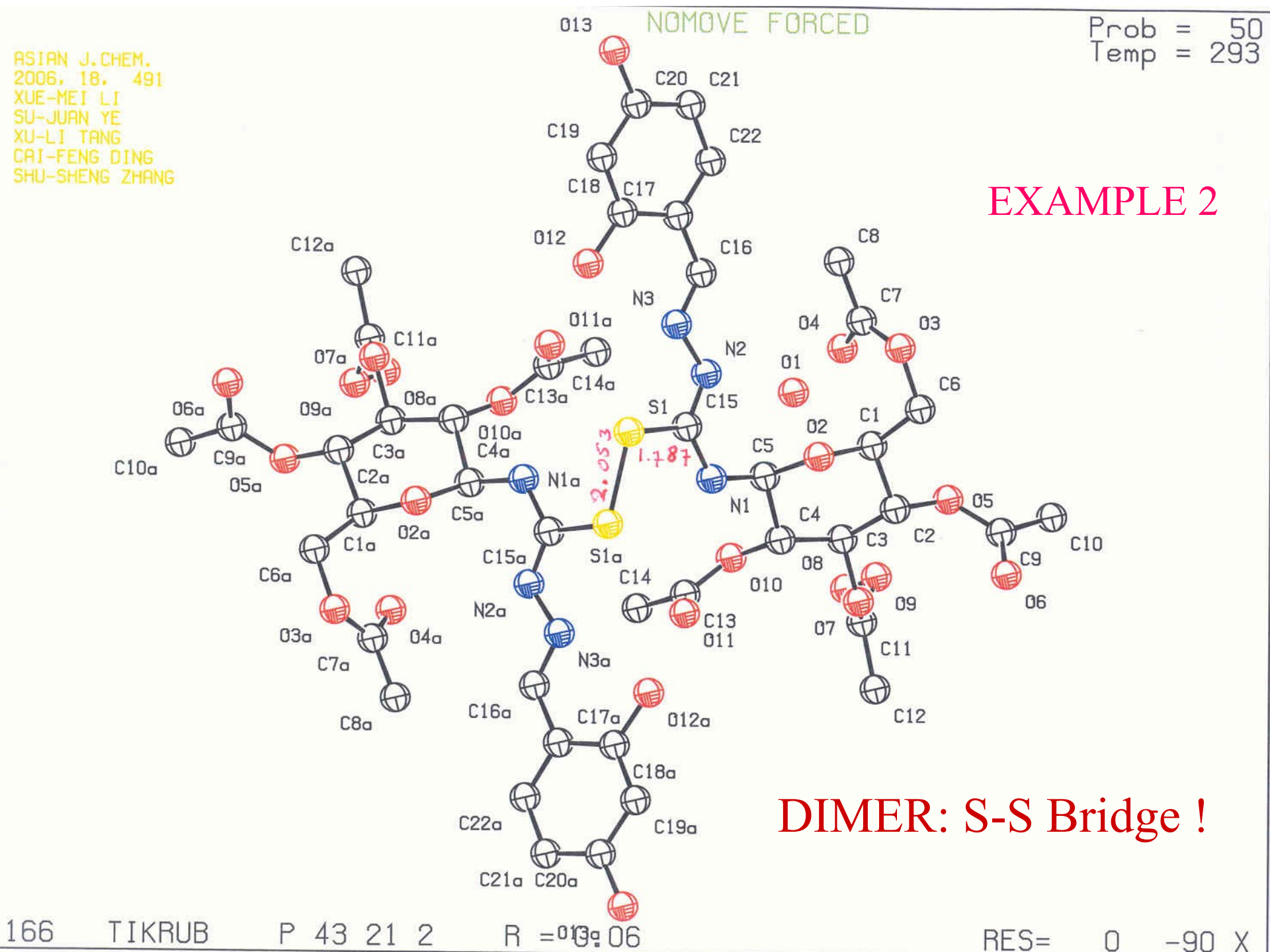
EXAMPLE 2

Reported as Monomer

BUT ➔



H_2O


ASIAN J.CHEM.
2006, 18, 491
XUE-MEI LI
SU-JUAN YE
XU-LI TANG
CAI-FENG DING
SHU-SHENG ZHANG
NOMOVE FORCED
Prob = 50
Temp = 293
EXAMPLE 2
DIMER: S-S Bridge !
2.053
1.787
166 TIKRUB P 43 21 2 R = 0.06
RES= 0 -90 X

WHAT ARE THE VALIDATION QUESTIONS ?

Single Crystal Structure Validation addresses three simple but important questions:

1 – Is the reported information complete?

2 – What is the quality of the analysis?

3 – Is the Structure Correct?

Implementation Problems of Structure Validation Before 1990

- Multiple Data Storage Types (often listing files).
- No Standard Computer Readable Format for data exchange.
- Data entry for publication via retyping in the manuscript.
- Thus: multiple typo's in Published Data.
- CSD Database Archival by Retyping from the published paper.
- Published data often incomplete.
- No easy numerical checking by referees etc.

The CIF Standard Solution

- CIF-Standard Proposal for Data Archival and Exchange:

  S.R. Hall, F.H. Allen, I.D. Brown (1991). Acta Cryst. A47, 655-685.
- Pioneered and Adopted by the International Union for Crystallography and Syd Hall (XTAL-System)
- Early adoption by the author of the now most used software package SHELXL97 (G.M.Sheldrick)
- Most current software now reads & writes CIF

CIF File Structure

- Both Computer and Human Readable Ascii encoded file
- Free Format
- Mostly 80 columns wide (maximum 2048)
- Parsable in units (Data names and Values)
- Data Order Flexible
- Dataname and Value associations
- loops

Constructs

- **data_*name***
  where *name* the choosen identifier of the data
- Data associations e.g.
  _ cell_length_a 16.6392(2)
- Repetition (loop)
  loop_
  __symmetry_equiv_pos_as_xyz
  'x, y, z'
  '-x, y+1/2, -z'

Construct for Text

- Text can be included between semi-columns
- Used for Acta Cryst. Section C & E submissions
- Example

_publ_section_comment

;

This paper presents to the best of our knowledge the first example of a very important MOF contruct.

;

CIF Example File

xterm

```
data_100K
_audit_creation_method            SHELXL-97
_chemical_name_systematic
;
Butane-1,4-diaminium sulfate
;
_chemical_name_common             'putrescinium sulfate'
_chemical_melting_point           ?
_chemical_formula_moiety          '2(C2 H7 N), O4 S'
_chemical_formula_sum              'C4 H14 N2 O4 S'
_chemical_formula_weight          186.23
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'
 'S'  'S'   0.1246   0.1234
 'International Tables Vol C Tables 4.2.6.8 and 6.1.1.4'
_symmetry_cell_setting            orthorhombic
_symmetry_space_group_name_H-M    'P n m a'
_symmetry_space_group_name_Hall   '-P 2ac 2n'
loop_
 _symmetry_equiv_pos_as_xyz
 'x, y, z'
 '-x+1/2, -y, z+1/2'
 'x+1/2, -y+1/2, -z+1/2'
 '-x, y+1/2, -z'
 '-x, -y, -z'
 'x-1/2, y, -z-1/2'
 '-x-1/2, y-1/2, z-1/2'
 'x, -y-1/2, z'
                                   1,1           To
```

xterm

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

_cell_length_a                    9.9722(4)
_cell_length_b                    9.4675(4)
_cell_length_c                    8.6532(4)
_cell_angle_alpha                 90.00
_cell_angle_beta                  90.00
_cell_angle_gamma                 90.00
_cell_volume                      816.96(6)
_cell_formula_units_Z             4
_cell_measurement_temperature     100(1)
_cell_measurement_reflns_used     2655
_cell_measurement_theta_min       2.0340
_cell_measurement_theta_max       27.6992
_exptl_crystal_description        'prismatic'
_exptl_crystal_colour             'colourless'
_exptl_crystal_size_max           0.6
_exptl_crystal_size_mid           0.4
_exptl_crystal_size_min           0.4
_exptl_crystal_density_meas       ?
_exptl_crystal_density_diffrn     1.514
_exptl_crystal_density_method     'not measured'
_exptl_crystal_F_000              400
_exptl_absorpt_coefficient_mu     0.370
_exptl_absorpt_correction_type    none
_exptl_absorpt_correction_T_min   ?
_exptl_absorpt_correction_T_max   ?
_exptl_absorpt_process_details    ?

_exptl_special_details
;

;
-- INSERT --                      53,42         5
```

CIF Completion

- CIF files are mostly created by the refinement program (e.g. SHELXL).
- Missing data can be added with a Text Editor, The Program **enCIFer** (from the CCDC) or **publCIF** (From the IUCr).
- The syntax can be checked with a locally installed version of the program **enCIFer**

  (Freely Available: www.ccdc.cam.ac.uk).

sk3290.cif – enCIFer

File Edit Search Tools Help

Blocks: 3

```
_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'
'S'  'S'   0.1246   0.1234
'International Tables Vol C Tables 4.2.6.8 and 6.1.1.4'

_symmetry_cell_setting          orthorhombic
_symmetry_space_group_name_H-M   'P n m a'
_symmetry_space_group_name_Hall  '-P 2ac 2n'

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

_cell_length_a                  9.9722(4)
_cell_length_b                  9.4675(4)
_cell_length_c                  8.6532(4)
_cell_angle_alpha
_cell_angle_beta                90.00
_cell_angle_gamma               90.00
_cell_volume                    816.96(6)
_cell_formula_units_Z           4
_cell_measurement_temperatur    100(1)
_cell_measurement_reflns_used   2655
_cell_measurement_theta_min     2.0340
```

EDITOR

Error detected with PROGRAM enCIFer

Missing Data

Editor | Visualiser

- Errors — Total of 1
  - Line 439 Data name _cell_angle_alpha immediately followed by another data name
- Warnings — Total of 7
- Remarks — - none

Total errors: 0; warnings: 9; remarks: 0

Checking CIF ...

Total errors: 1; warnings: 10; remarks: 0

How is Validation Currently Implemented ?

- The results of a structure analysis are now required to be available in the computer readable CIF format.
- Validation checks can be executed at any time both in-house or through the WEB-based IUCr CHECKCIF server.
- A file (Check.def) defines the issues that are tested with levels of severity and associated explanation and advise. (www.cryst.chem.uu.nl/platon/CIF-VALIDATION.pdf)
- Most non-trivial tests are executed with routines in the program PLATON

VALIDATION ALERT LEVELS

CheckCIF/PLATON creates a report in the form of a list of ALERTS with the following ALERT levels:

- ALERT A – Serious Problem
- ALERT B – Potentially Serious Problem
- ALERT C – Check & Explain
- ALERT G – Verify or Take Notice

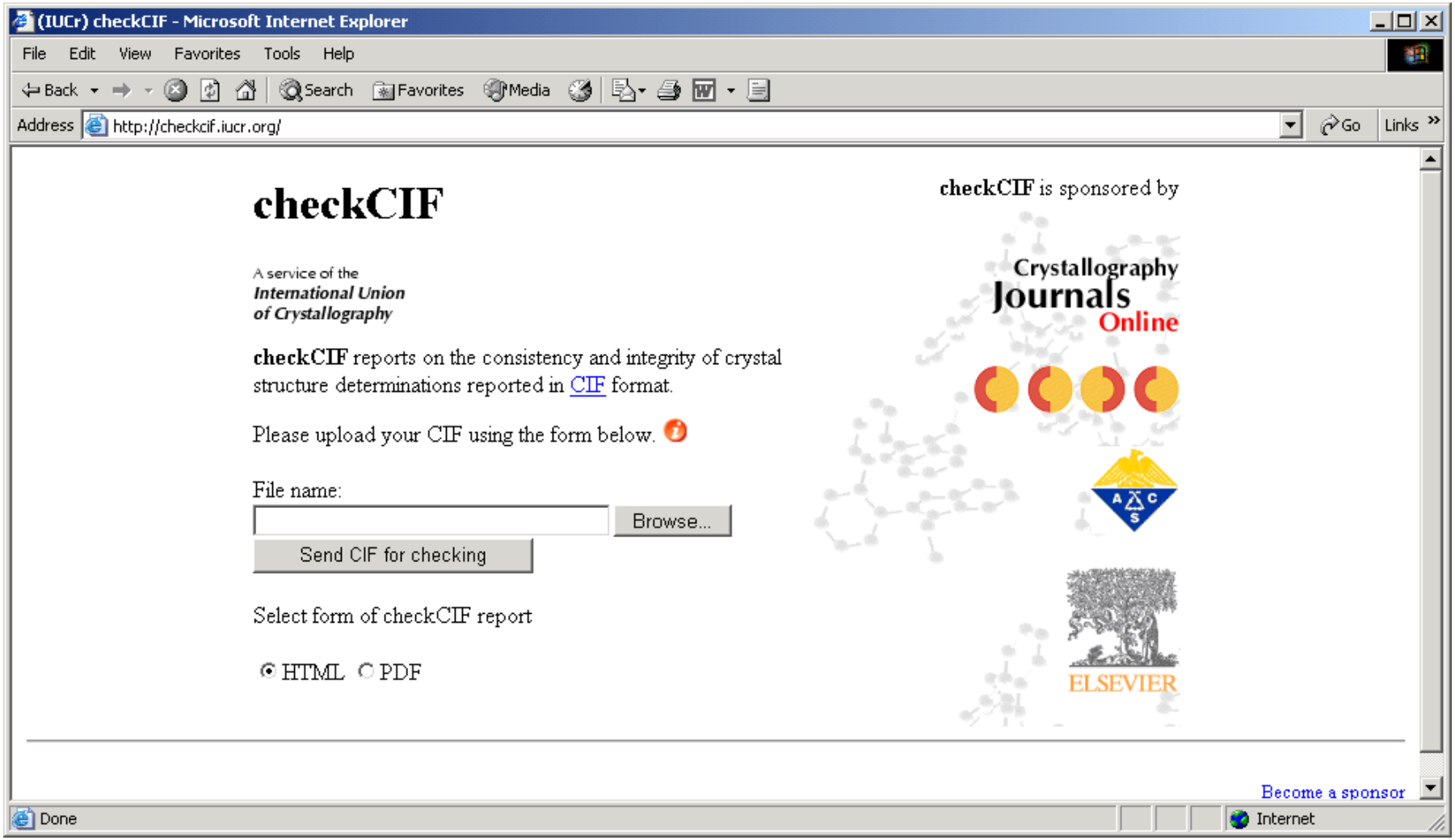
VALIDATION ALERT TYPES

1 - CIF Construction/Syntax errors,
Missing or Inconsistent Data.

2 - Indicators that the Structure Model
may be Wrong or Deficient.

3 - Indicators that the quality of the results
may be low.

4 – Info, Cosmetic Improvements, Queries and
Suggestions.



The Chester CHECKCIF Server http://checkcif.iucr.org

checkCIF/PLATON report (publication check)

No syntax errors found.
Please wait while processing ....

CIF dictionary
Interpreting this report

Datablock: I

Bond precision: C-C = 0.0157 A Wavelength=0.71073
Cell: a=7.6336(15) b=27.725(6) c=12.051(2)
alpha=90 beta=98.80(3) gamma=90
Temperature: 153 K

	Calculated	Reported
Volume	2520.5(9)	2520.5(9)
Space group	P 21/n	P 21/n
Hall group	-P 2yn	-P 2yn
Moiety formula	C28 H32 Br0.97 Cl1.03 N2 Ru	C28 H32 Br0.97 Cl1.03 N2 Ru
Sum formula	C28 H32 Br0.97 Cl1.03 N2 Ru	C28 H32 Br0.97 Cl1.03 N2 Ru
Mr	611.69	611.69
Dx,g cm-3	1.612	1.612
Z	4	4
Mu (mm-1)	2.290	2.290
F000	1237.9	1237.9
F000'	1232.67	
h,k,lmax	9,33,14	9,33,14
Nref	4452	4449
Tmin,Tmax	0.726,0.955	0.581,0.955
Tmin'	0.546	

Correction method= MULTI-SCAN
Data completeness= 0.999 Theta(max)= 25.030
R(reflections)= 0.0891(3288) wR2(reflections)= 0.2547(4449)
S = 1.037 Npar= 306

The following ALERTS were generated. Each ALERT has the format
test-name_ALERT_alert-type_alert-level.
Click on the hyperlinks for more details of the test.

Alert level B
PLAT083_ALERT_2_B SHELXL Second Parameter in WGHT unusually Large. 37.00

Alert level C
RFACR01_ALERT_3_C The value of the weighted R factor is > 0.25
Weighted R factor given 0.255
PLAT301_ALERT_3_C Main Residue Disorder ..... 6.00 Perc.

PLATON/CHECK CIF + FCF Results

```
=================================================================================
>>> The Following Model and Quality ALERTS were generated     - (Acta-Mode) <<<
=================================================================================
     Format: alert-number_ALERT_alert-type_alert-level text

148_ALERT_3_B su on the      a  - Axis is Too Large (x 1000) .         10 Ang.
148_ALERT_3_B su on the      b  - Axis is Too Large (x 1000) .          8 Ang.
148_ALERT_3_B su on the      c  - Axis is Too Large (x 1000) .         30 Ang.
230_ALERT_2_B Hirshfeld Test Diff for    O1     --  C1       ..      26.27 su
230_ALERT_2_B Hirshfeld Test Diff for    O2     --  C1       ..      11.10 su
242_ALERT_2_B Check Low       Ueq as Compared to Neighbors for         C1
420_ALERT_2_B D-H Without Acceptor       O1     -   H1      ...        ?
#================================================================================
230_ALERT_2_C Hirshfeld Test Diff for    C1     --  C2       ..       6.87 su
242_ALERT_2_C Check Low       Ueq as Compared to Neighbors for        C11
911_ALERT_3_C Missing # FCF Refl Between THmin & STh/L=  0.595         11
913_ALERT_3_C Missing # of Very Strong Reflections in FCF ....         1
#================================================================================
860_ALERT_3_G Note: Number of Least-Squares Restraints .......         2
#================================================================================
#================================================================================


=================================================================================
>>> The Following Improvement and Query ALERTS were generated - (Acta-Mode) <<<
=================================================================================
926_ALERT_1_B Reported and Calculated   R1 * 100.0 Differ by .      -0.81
927_ALERT_1_B Reported and Calculated  wR2 * 100.0 Differ by .      -2.26
#================================================================================
042_ALERT_1_C Calc. and Reported MoietyFormula Strings  Differ          ?
790_ALERT_4_C Centre of Gravity not Within Unit Cell: Resd.  #          2
              C8 H6 O4
928_ALERT_1_C Reported and Calculated    S value   Differ by .      -0.13
#================================================================================
128_ALERT_4_G Non-standard setting of Space-group P21/c   ....      P21/n
#================================================================================
```

Which Key Validation Issues are Addressed

- Missed Space Group symmetry (“being Marshed”)
- Wrong chemistry (Mis-assigned atom types).
- Too many, too few or misplaced H-atoms.
- Unusual displacement parameters.
- Hirshfeld Rigid Bond test violations.
- Missed solvent accessible voids in the structure.
- Missed Twinning.
- Absolute structure
- Data quality and completenes.

Simple Validation Issues

- Many data sets are apparently collected at either 293(2) or 273 K
- Improper program defaults or values from previous papers are retained.
- Data collected with a CCD system and corrected for absorption with Psi-scans ! ?

Examples of Correctable Issues

- Following are some examples of the type of problems addressed.

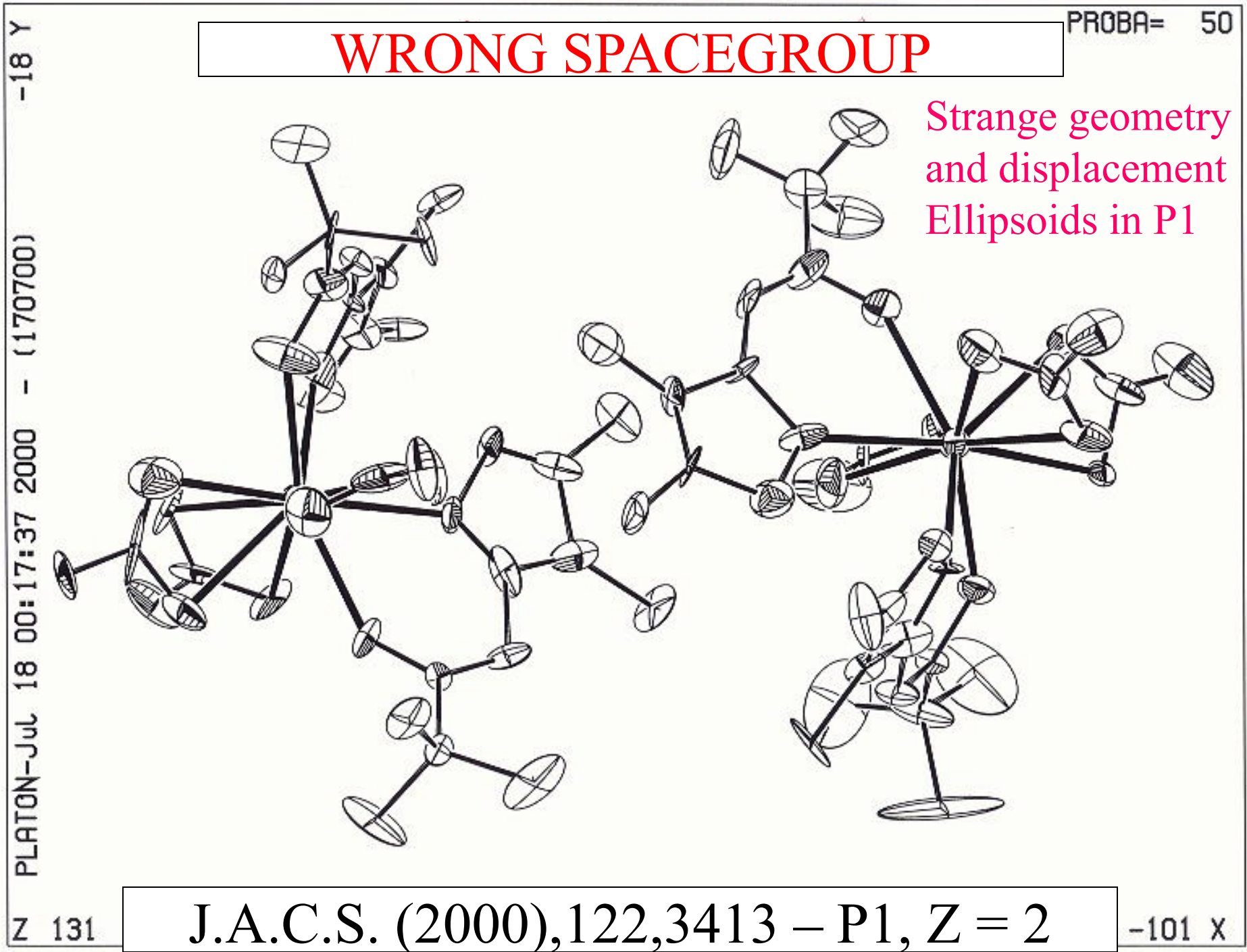
1 – Refinement in the Wrong Space group.

2 – Wrong Atom Type Assignment.

3 – Misplaced H-Atoms.

4 – Missing H-Atoms.

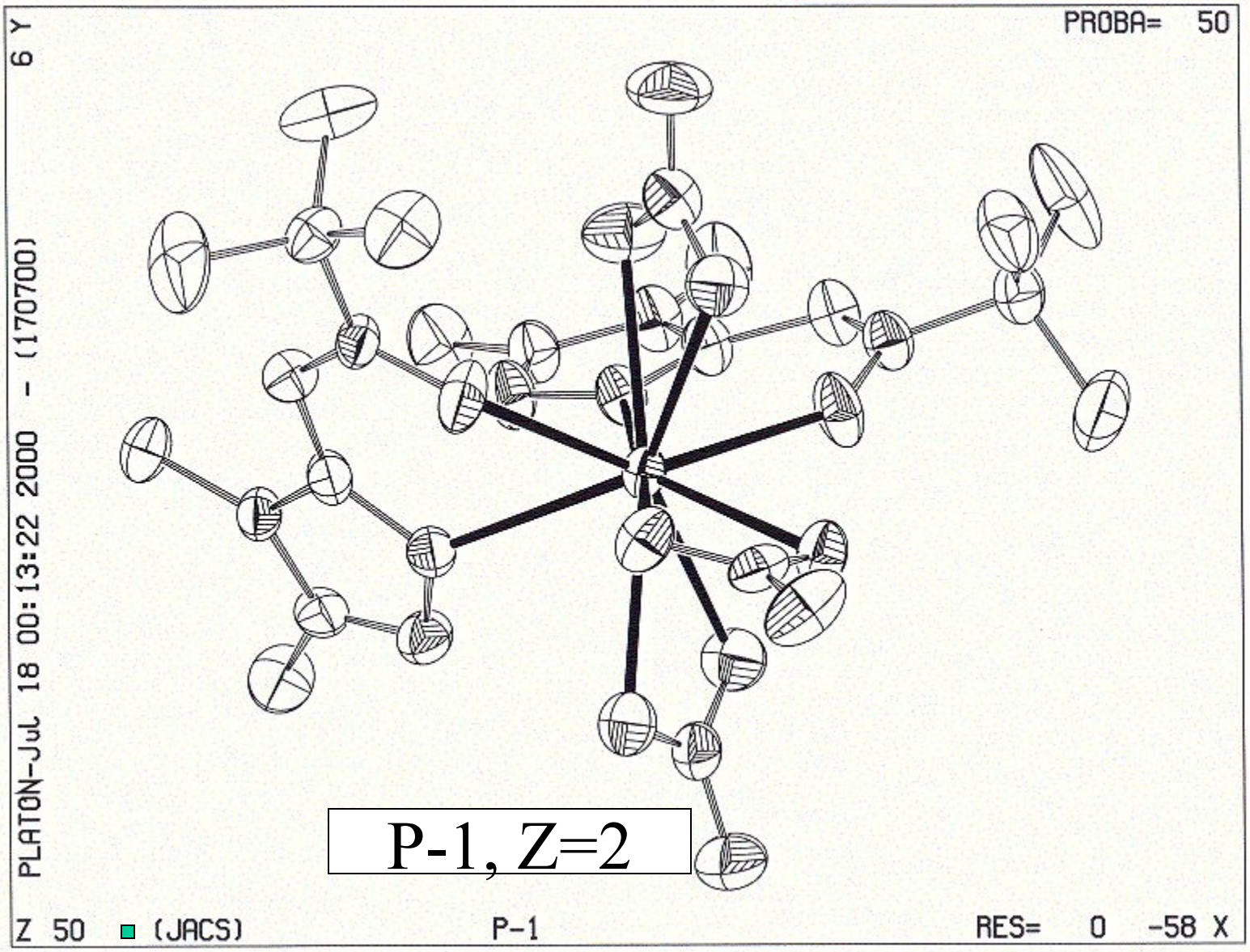
WRONG SPACEGROUP



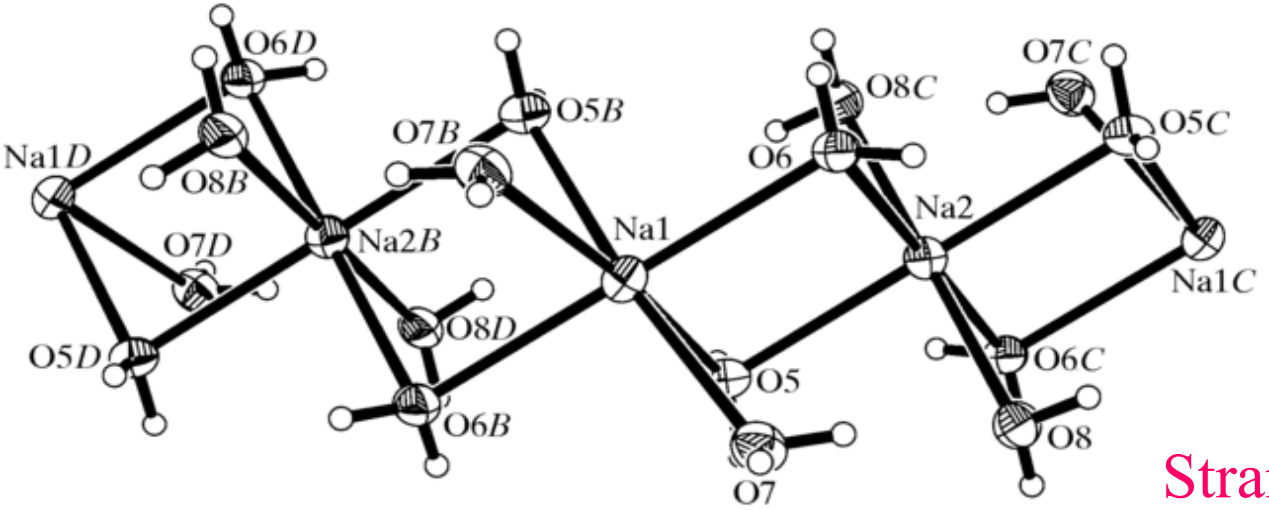
Strange geometry and displacement Ellipsoids in P1

J.A.C.S. (2000),122,3413 – P1, Z = 2

CORRECTLY REFINED STRUCTURE

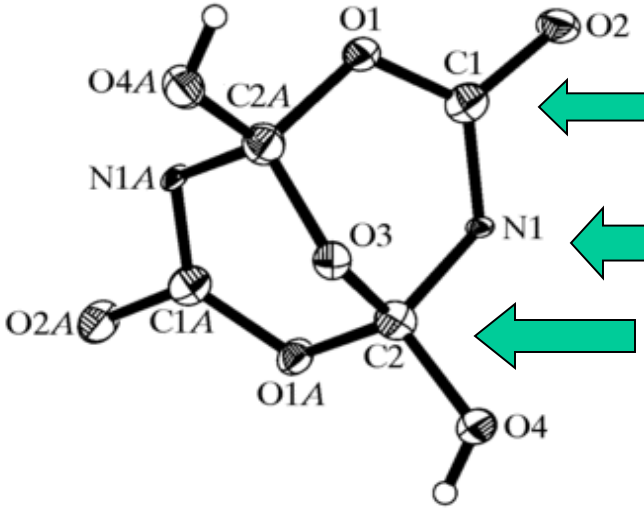


Published with Wrong Composition



Unexpected Result !

Strange Ellipsoids



C → B

N → O

C → B

Corrected Structure BORAX !

=> Retracted

Searching for structures with a Methyl Moiety bridging two metals …

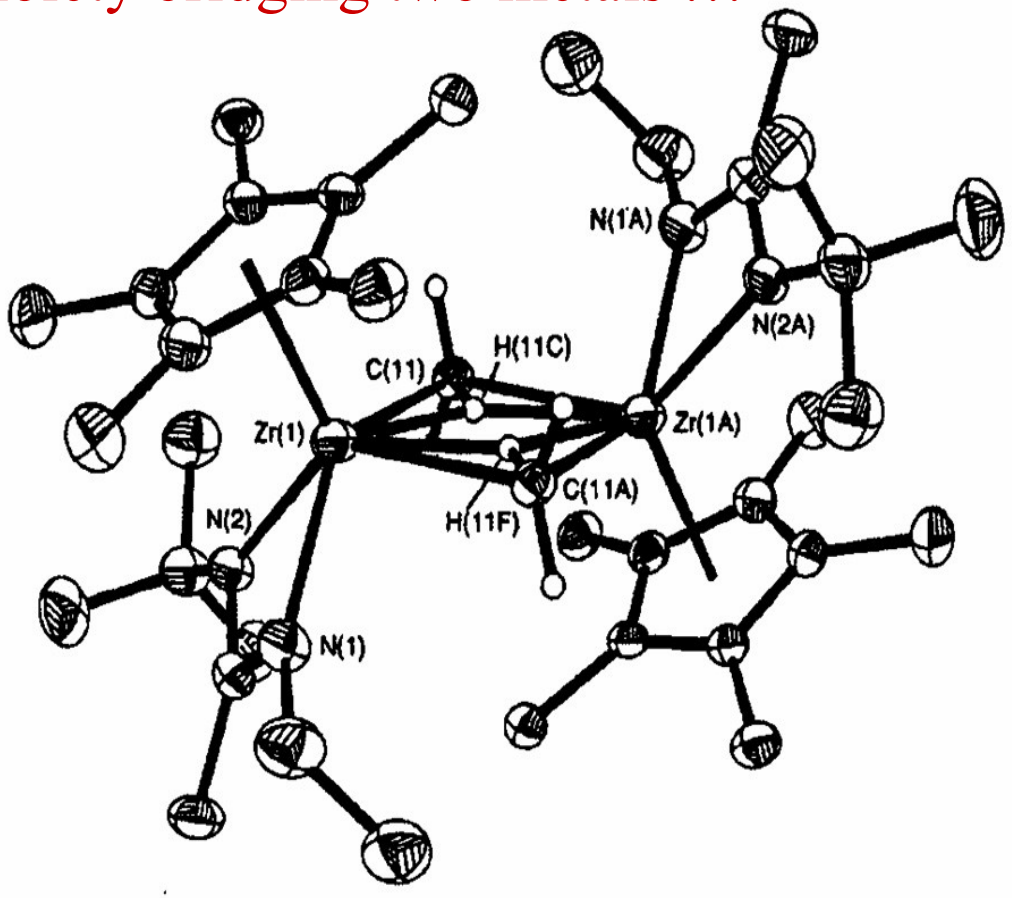


Figure 2. Molecular structure of **2** (30% thermal ellipsoids). The two $[B(C_6F_5)_4]$ anions and all but the bridging methyl hydrogen atoms, represented by spheres of arbitrary size, have been omitted for the sake of clarity. For each bridging methyl group, only the labeled hydrogen atom has been crystallographically located with the other two being placed in logical positions.

Structure of a strange CH_3 Bridged Zr Dimer

Paper has been cited

47 times !

So can we believe this structure?

The Referees did …!

But …

H .. H = 1.32 Ang. !

HOT STRUCTURE – FAST LANE PUBLICATION

THE NEWSMAGAZINE OF THE CHEMICAL WORLD ONLINE | HOW TO LOG IN | HOW TO REACH US | SITE MAP

CHEMICAL & Engineering News

Search Go!

HOME | THIS WEEK'S CONTENTS | C&EN CLASSIFIEDS | JOIN ACS

- Table of Contents
- C&EN Classifieds
- Today's Headlines
- Cover Story
- Editor's Page
- Business
- Government & Policy
- Science & Technology
- ACS News
- Calendars
- Books
- Career & Employment
- Special Reports
- Nanotechnology
- What's That Stuff?

Back Issues
2002 Go!

Safety Letters
Chemcyclopedia

ACS Members can sign up to receive C&EN e-mail newsletter.
Join

Join ACS

SCIENCE & TECHNOLOGY

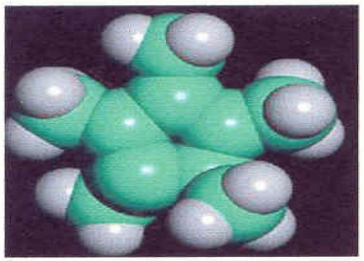
April 29, 2002
Volume 80, Number 17
CENEAR 80 17 p. 30
ISSN 0009-2347

ELUSIVE CARBOCATION ISOLATED AS A SOLID

Pentamethylcyclopentadienyl cation is found to be a stable singlet with a distorted structure

RON DAGANI

Chemists like to study molecules they can store in a bottle on the shelf. If they can't put the molecules in a bottle, they will, of course, study them any way they can--in solution, in the gas phase, in a frozen matrix. But there's a particular satisfaction in bottling a molecule that no one thought could be bottled.



MISSHAPEN A space-filling model of the pentamethylcyclopentadienyl cation shows its distorted, nonplanar ring.

Chemistry professor Joseph B. Lambert of Northwestern University knows that satisfaction. Earlier this month, he and graduate student Lijun Lin reported the first isolation and X-ray structural characterization of a cyclopentadienyl cation--specifically, the pentamethyl-substituted $C_5Me_5^+$ cation [*Angew. Chem. Int. Ed.*, **41**, 1429 (2002)].

Related People
Joseph B. Lambert
Vitaly Rassolov
Ronald Breslow

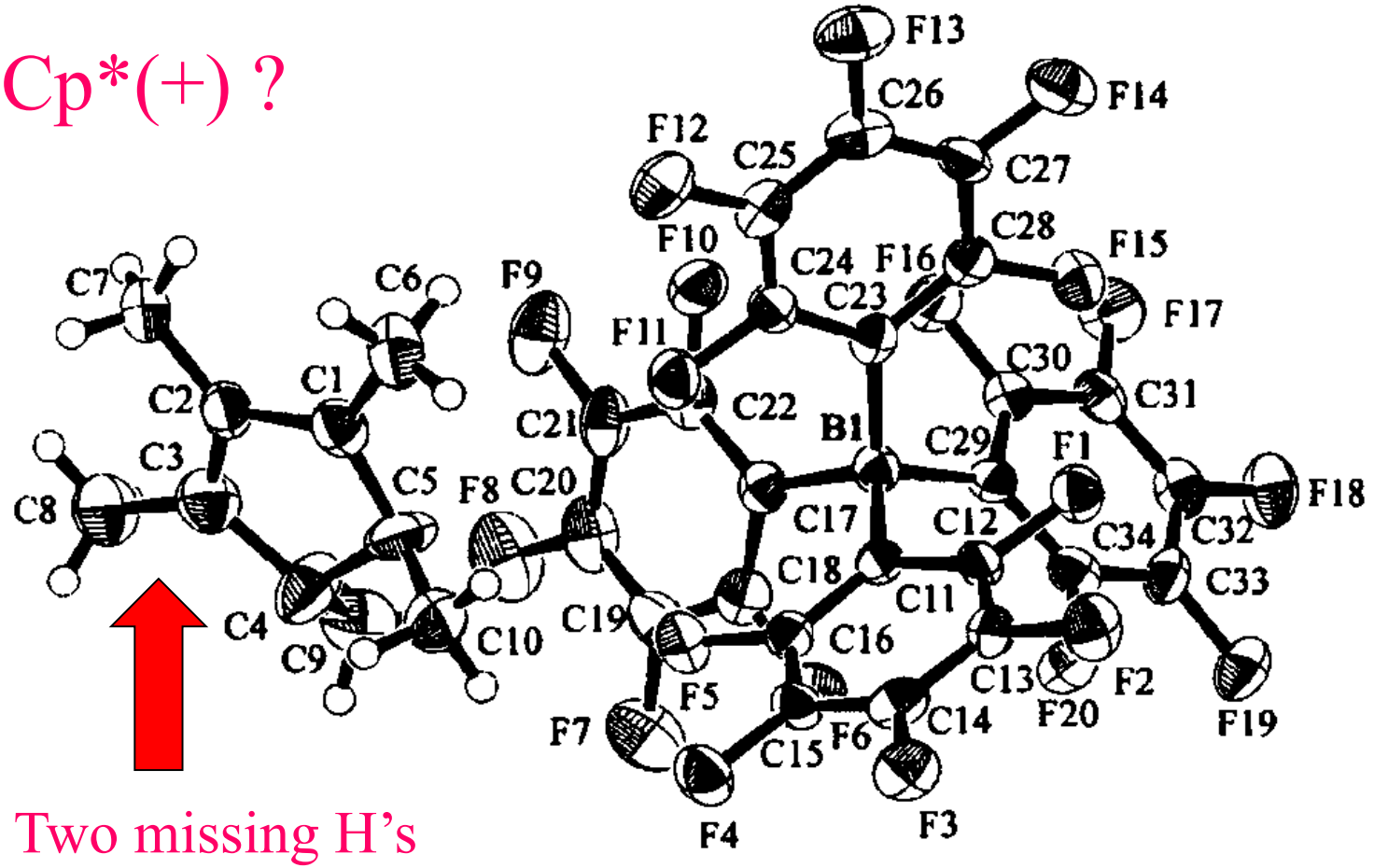
E-mail this article to a friend
Print this article
E-mail the editor

Cp*(+) !! . ?

THE STABLE PENTAMETHYLCYCLOPENTADIENYL CATION

J.B.Lambert et al. Angew. Chem. Int. Ed. 2002, 41, 1429-1431

Cp*(+) ?



No ! Two missing H's

Figure 1. The crystal structure of pentamethylcyclopentadienyl tetrakis-(pentafluorophenyl)borate. There is no covalent bonding between the cation on the left and the anion on the right.

NOT SO HOT AFTER ALL !!

Editors Note in the next issue of Angewandte Chemie

CORRIGENDUM

Note from the Editors: unfortunately the results reported in the communication "The Stable Pentamethylcyclopentadienyl Cation" by Joseph B. Lambert et al. in issue 8/2002 (pp. 1429–1431) must be corrected. Guy Bertrand et al. quickly discovered that not the pentamethylcyclopentadienyl cation but the pentamethylcyclopentenyl cation was prepared and characterized (the corresponding communication will be published in issue 13, and will appear earlier on the *Angewandte Chemie* homepage).

Evaluation and Performance

- The validation scheme has been very successful for Acta Cryst. C & E in setting standards for quality and reliability.
- The missed symmetry problem has been solved for the IUCr journals (unfortunately not generally yet: There are still numerous 'Marshable' structures).
- Most major chemical journals currently have now some form of a validation scheme implemented.
- But, has it solved all problems … ?

Problems to be Addressed

- Synthetic Chemist View: 'Addressing Crystallographic Details holds up the Publication of Important Chemistry' (but see previous example in Angew. Chemie !)
- Interesting Author Question in response to referee issue:

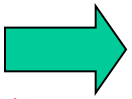
  What does it mean "Space group Incorrect"
- Crystallographic Education (beyond Pushbutton training and Black Box operation) is getting scarce nowadays.
- Sadly: Referees who do not understand or do not know how to respond adequately to ALERTS
- Recently: The need to Detect Fraud and Fraudulous manipulation ….

Note on Editing the CIF

- The Idea of editing the CIF is to add missing (experimental) information to the CIF.
- However: Some authors have now been found to polish away less nice numerical values.
- This leaves traces and is generally detected sooner or later by the validation software and is not good for the scientific career of the culprit…
- The recently implemented FCF-Checking now addresses this issue in even more detail.

Reflection CIF (FCF)

Cell Data Should correspond with CIF data



Terminal — ssh — 73×43

```
#
# h,k,l, Fc-squared, Fo-squared, sigma(Fo-squared) and status flag
#
data_s4223a
_shelx_title ' s4223a - SHELXL'
_shelx_refln_list_code          4
_shelx_F_calc_maximum      687.61
_exptl_crystal_F_000      3008.00
_refIns_d_resolution_high  0.7696

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

_cell_length_a    28.5187
_cell_length_b    11.2036
_cell_length_c    19.4833
_cell_angle_alpha  90.000
_cell_angle_beta  101.351
_cell_angle_gamma  90.000

_shelx_F_squared_multiplier     1.000

loop_
 _refln_index_h
 _refln_index_k
 _refln_index_l
 _refln_F_squared_calc
 _refln_F_squared_meas
 _refln_F_squared_sigma
 _refln_observed_status
   3   1 -25     3134.30     3933.22    696.84 o
   5   1 -25     9060.39     9137.18    788.13 o
   7   1 -25      308.45      307.35    552.52 o
   9   1 -25     1727.00     2193.83    543.48 o
  11   1 -25     1591.10     1053.00    587.08 o
   6   2 -25    11770.63    12530.49    809.23 o
   8   2 -25     4369.66     3470.53    667.51 o
  -4   0 -24    19927.68    20382.10    939.86 o
"s4223a.fcf" 13988L, 698435C                          1,1           Top
```

FCF-VALIDATION

- Check of CIF & FCF data Consistency
- Check of completeness of the reflection data set.
- Automatic Detection of ignored twinning
- Detection of Applied Twinning Correction without having been Reported in the paper.
- Validity check of the reported Flack parameter value against the Hooft parameter value.
- Analysis of the details of the Difference Density Fourier Map for unreported features.

Sloppy, Novice or Fraudulent ?

- Errors are easily made and unfortunately not always discernable from fraud.
- Wrong element type assignments can be caused as part of an incorrect analysis of an unintended reaction product.
- Alternative element types can be (and have been) substituted deliberately to create a ‘new publishable’ structure.

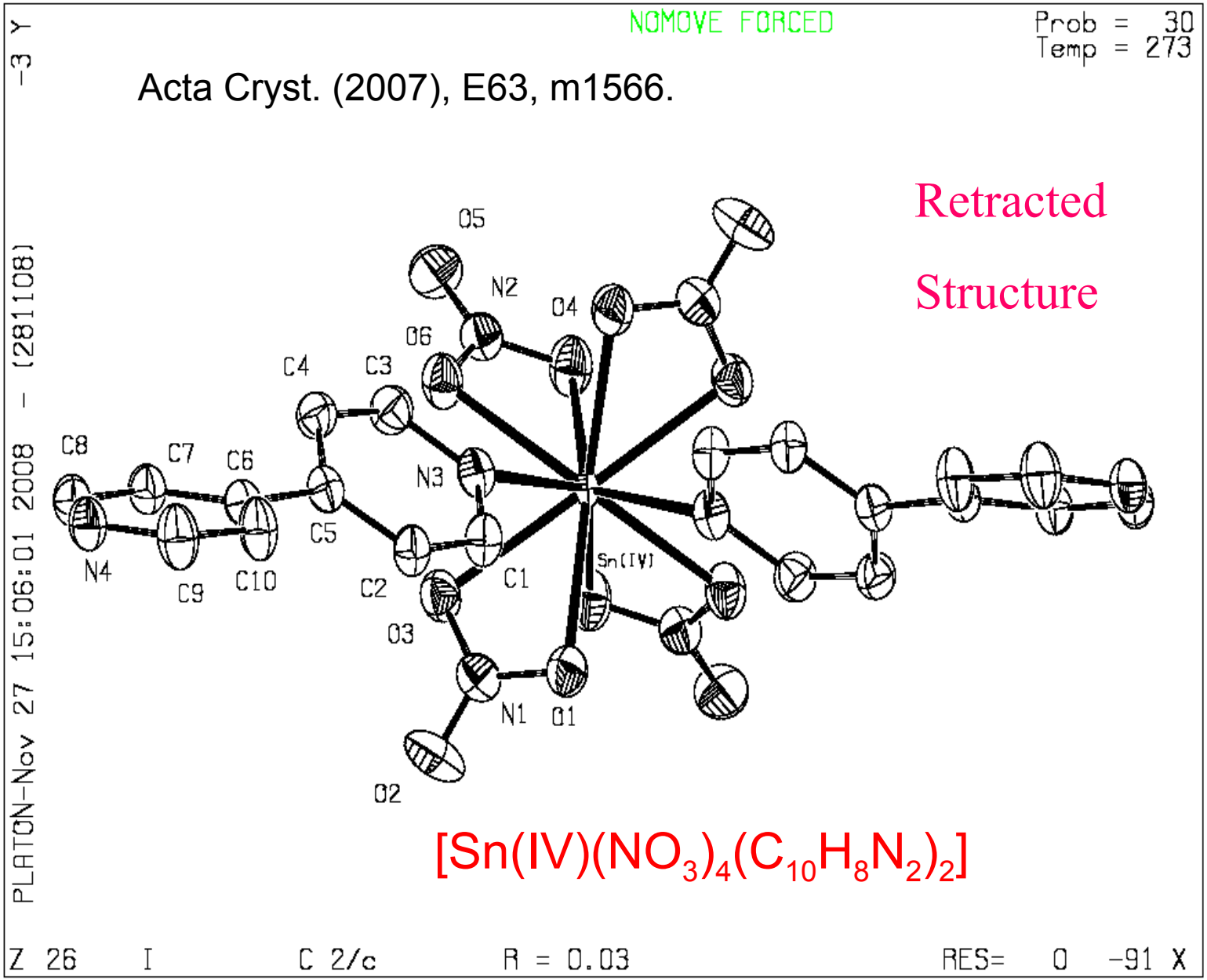
The need of serious validation by knowledgeable Referees

- The validation issues and tools are probably best illustrated with an analysis of a few fraudulous papers that reached the recent literature and (unfortunately) the CSD.
- Early warning signs are generally: troublesome displacement parameters and unusual short inter-molecular contacts.

Some Relevant ALERTS

Wrong atom type assignments generally cause:

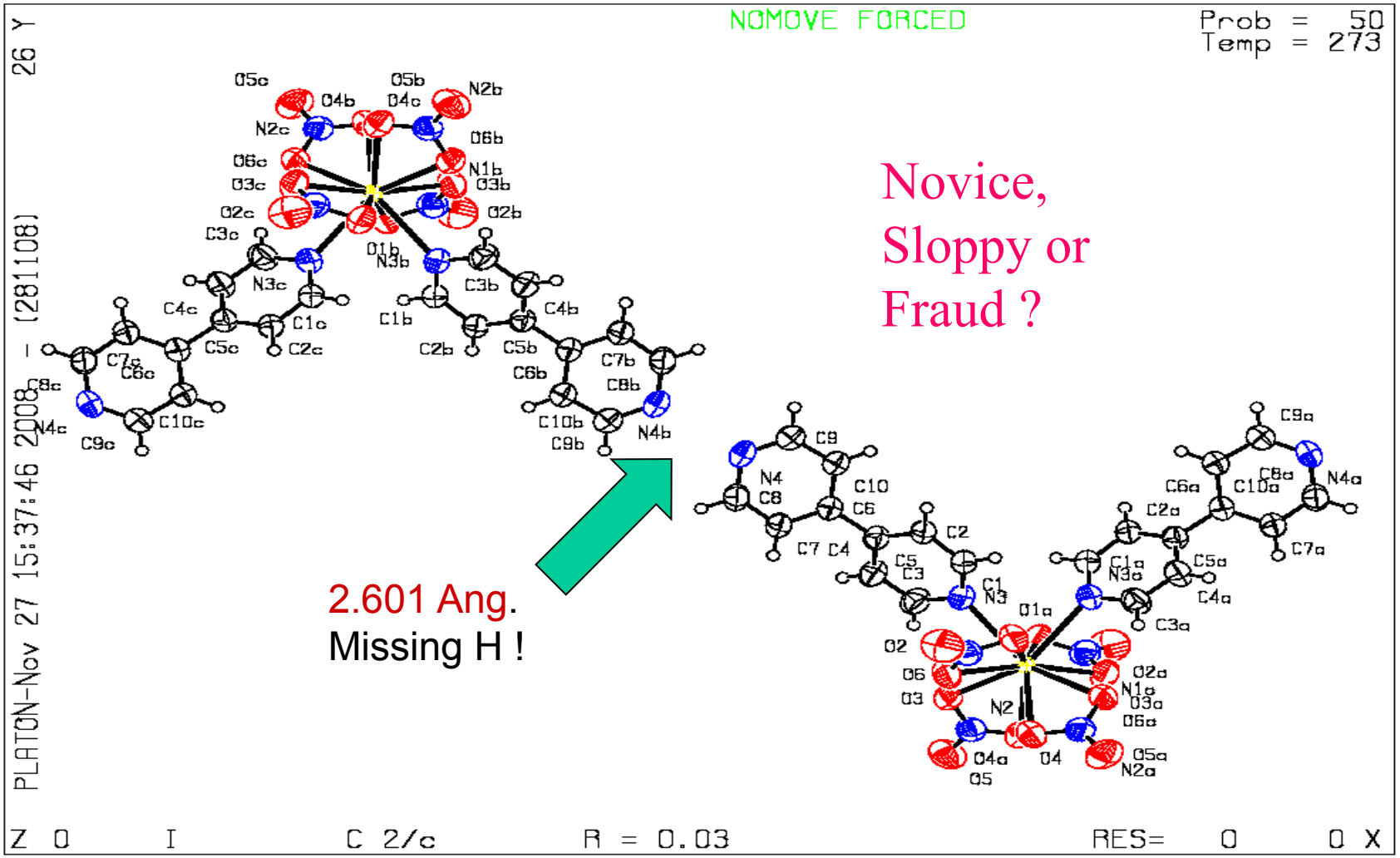
- Serious Hirshfeld Rigid Bond Violation ALERTS
- Larger than expected difference map minima and maxima.
- wR2 >> 2 * R1
- High values for the SHELXL refined weight parameter



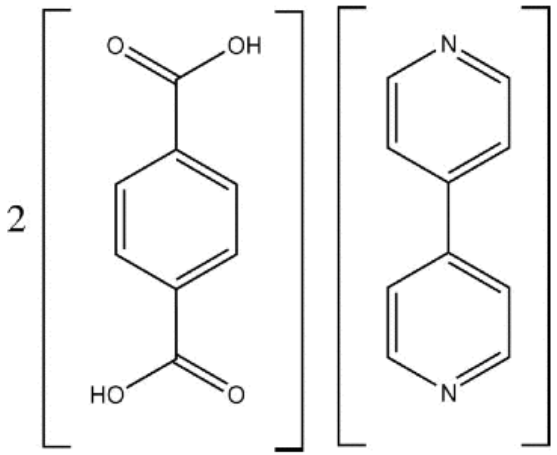
Acta Cryst. (2007), E63, m1566.

Retracted Structure

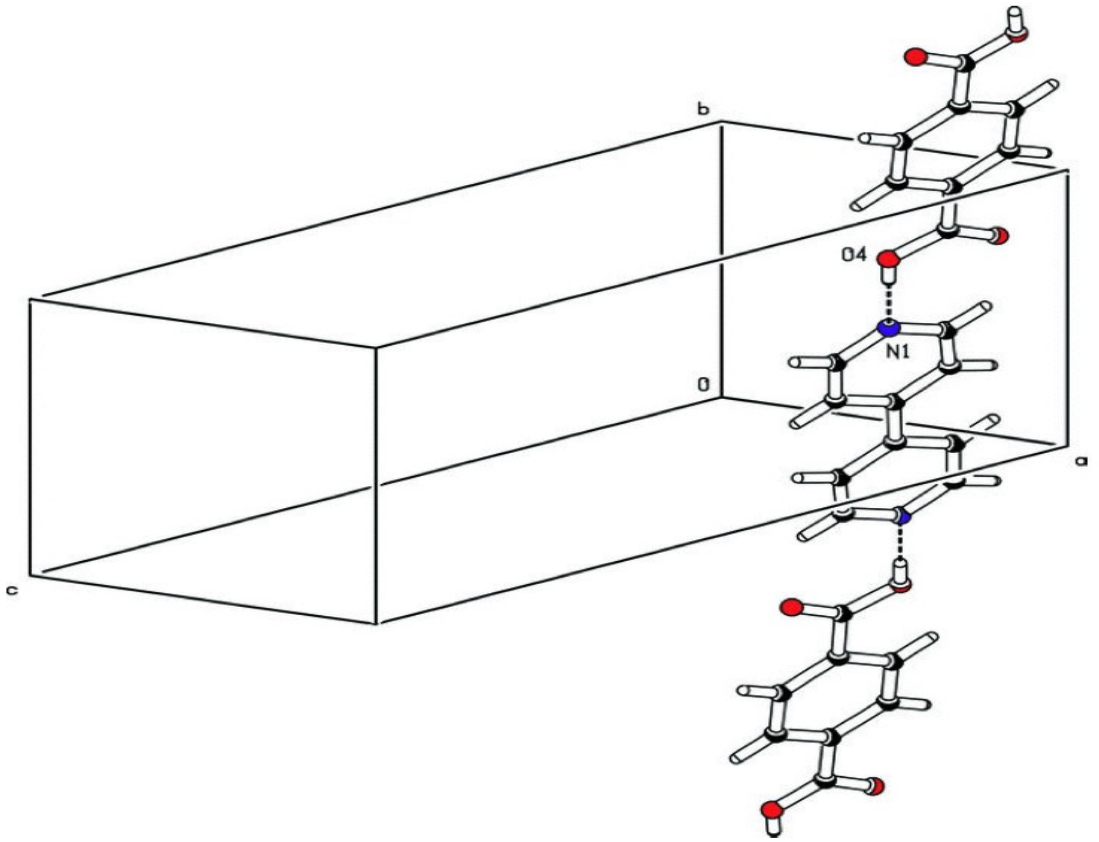
$[Sn(IV)(NO_3)_4(C_{10}H_8N_2)_2]$



Missing H in bridge & Sn(IV) => Lanthanide(III)

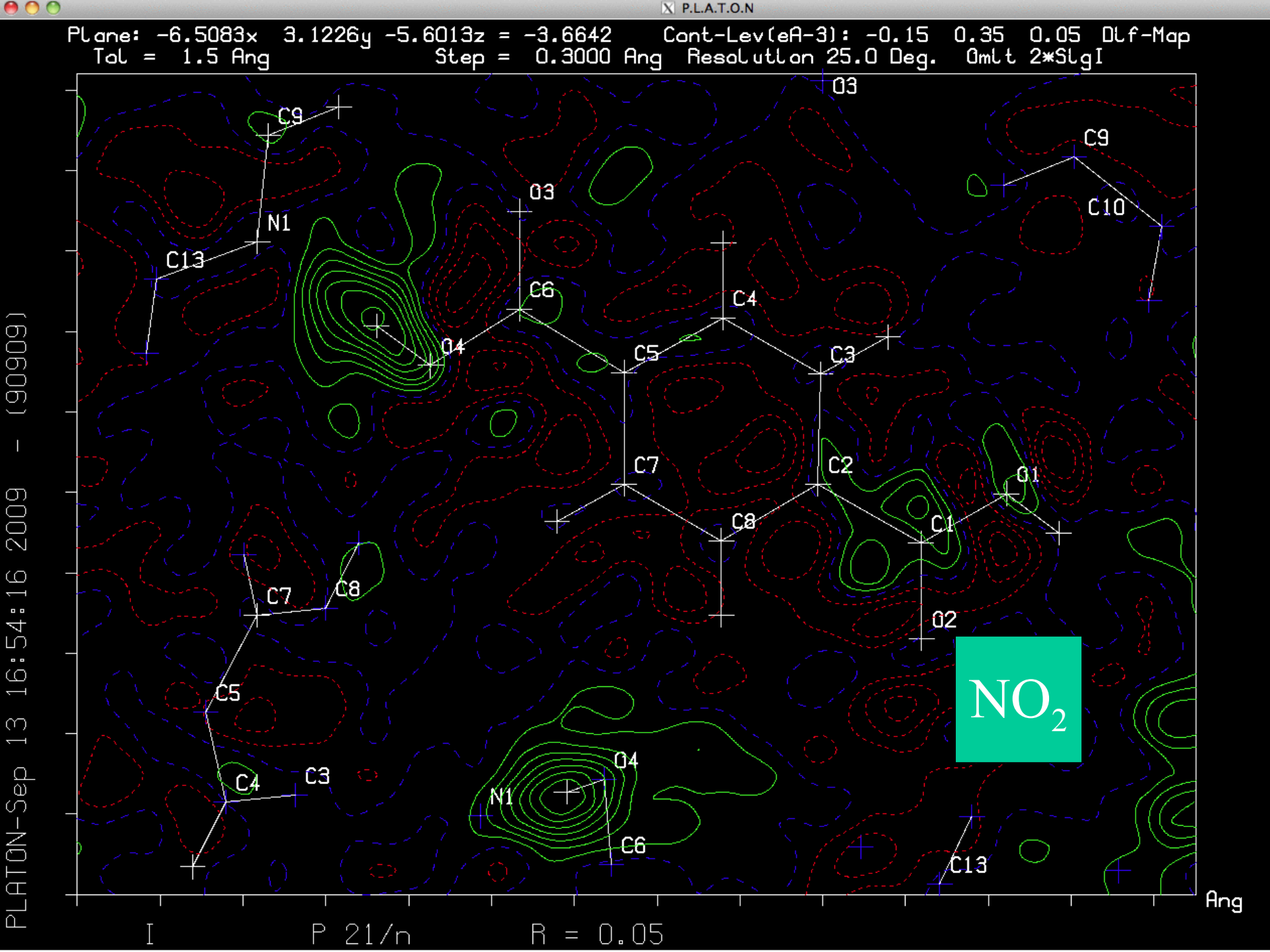


Published structure is claimed to form an infinite hydrogen bonded chain



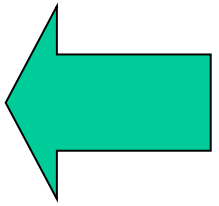
However: This structure does not include a dicarboxylic acid but the previously published para-nitrobenzoic acid.

PROOF: Difference map calculated without the 2 carboxylic H-atoms


P.L.A.T.O.N
Plane: -6.5083x 3.1226y -5.6013z = -3.6642 Cont-Lev(eA-3): -0.15 0.35 0.05 Dlf-Map
Tol = 1.5 Ang Step = 0.3000 Ang Resolution 25.0 Deg. Omit 2*SigI
PLATON-Sep 13 16:54:16 2009 - (90909)
O3
C9
C10
N1
C13
C6
C4
O4
C5
C3
C7
C2
O1
C8
C1
O2
NO_2
Ang
I P 21/n R = 0.05

There are clear ALERTS ! But apparently ignored

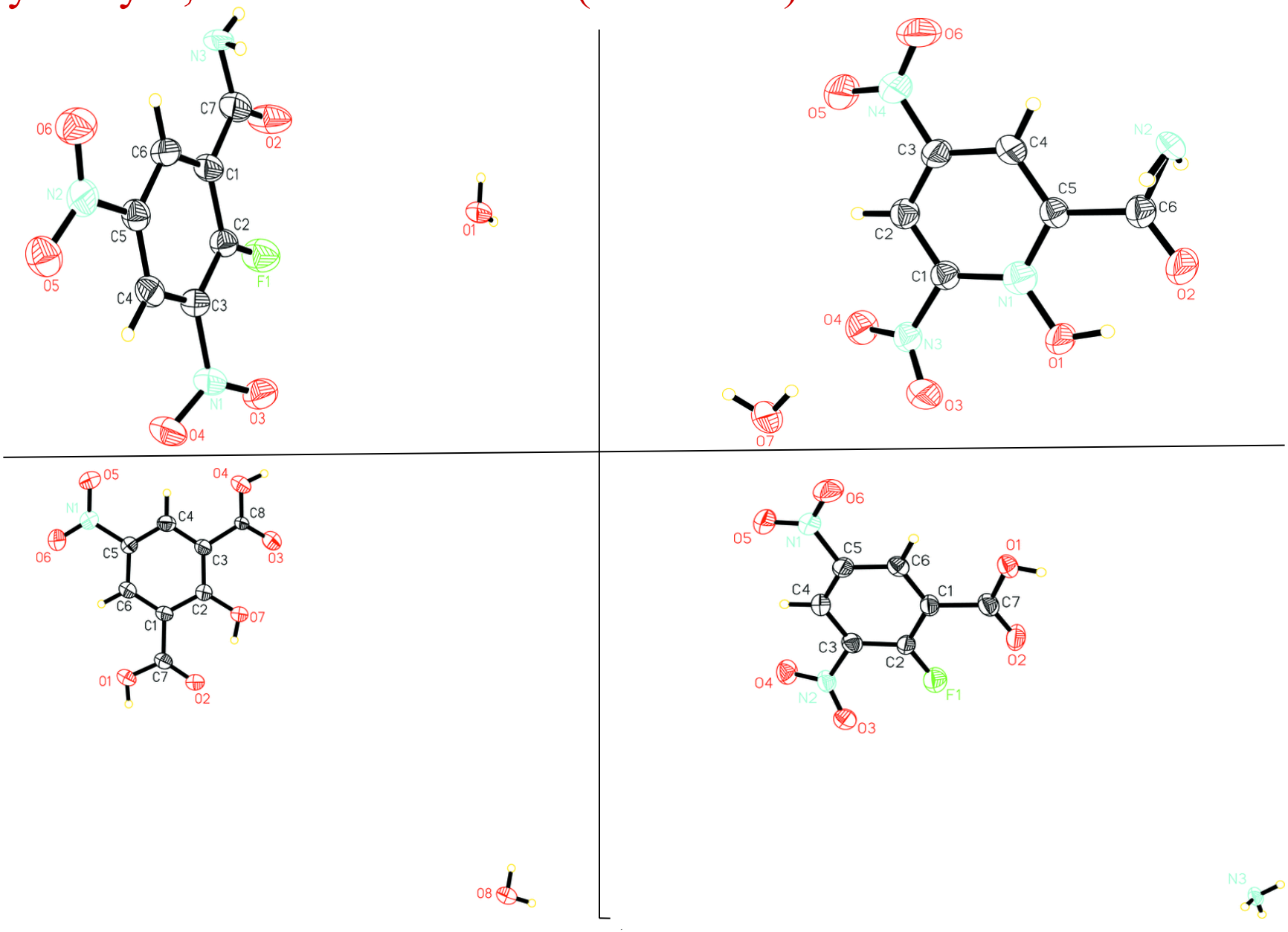
```
=================================================================================
>>> The Following Model and Quality ALERTS were generated     - (Acta-Mode) <<<
=================================================================================
     Format: alert-number_ALERT_alert-type_alert-level text

148_ALERT_3_B su on the      a  - Axis is Too Large (x 1000) .          10 Ang.
148_ALERT_3_B su on the      b  - Axis is Too Large (x 1000) .           8 Ang.
148_ALERT_3_B su on the      c  - Axis is Too Large (x 1000) .          30 Ang.
230_ALERT_2_B Hirshfeld Test Diff for    O1     --  C1       ..       26.27 su
230_ALERT_2_B Hirshfeld Test Diff for    O2     --  C1       ..       11.10 su
242_ALERT_2_B Check Low       Ueq as Compared to Neighbors for          C1
420_ALERT_2_B D-H Without Acceptor       O1     -   H1      ...         ?
#=================================================================================
230_ALERT_2_C Hirshfeld Test Diff for    C1     --  C2       ..        6.87 su
242_ALERT_2_C Check Low       Ueq as Compared to Neighbors for         C11
911_ALERT_3_C Missing # FCF Refl Between THmin & STh/L=  0.595          11
913_ALERT_3_C Missing # of Very Strong Reflections in FCF ....          1
#=================================================================================
860_ALERT_3_G Note: Number of Least-Squares Restraints .......         2
#=================================================================================
#=================================================================================

=================================================================================
>>> The Following Improvement and Query ALERTS were generated - (Acta-Mode) <<<
=================================================================================
926_ALERT_1_B Reported and Calculated   R1 * 100.0 Differ by .       -0.81
927_ALERT_1_B Reported and Calculated  wR2 * 100.0 Differ by .       -2.26
#=================================================================================
042_ALERT_1_C Calc. and Reported MoietyFormula Strings  Differ          ?
790_ALERT_4_C Centre of Gravity not Within Unit Cell: Resd.  #          2
              C8 H6 O4
928_ALERT_1_C Reported and Calculated    S value   Differ by .       -0.13
#=================================================================================
128_ALERT_4_G Non-standard setting of Space-group P21/c   ....      P21/n
#=================================================================================
```



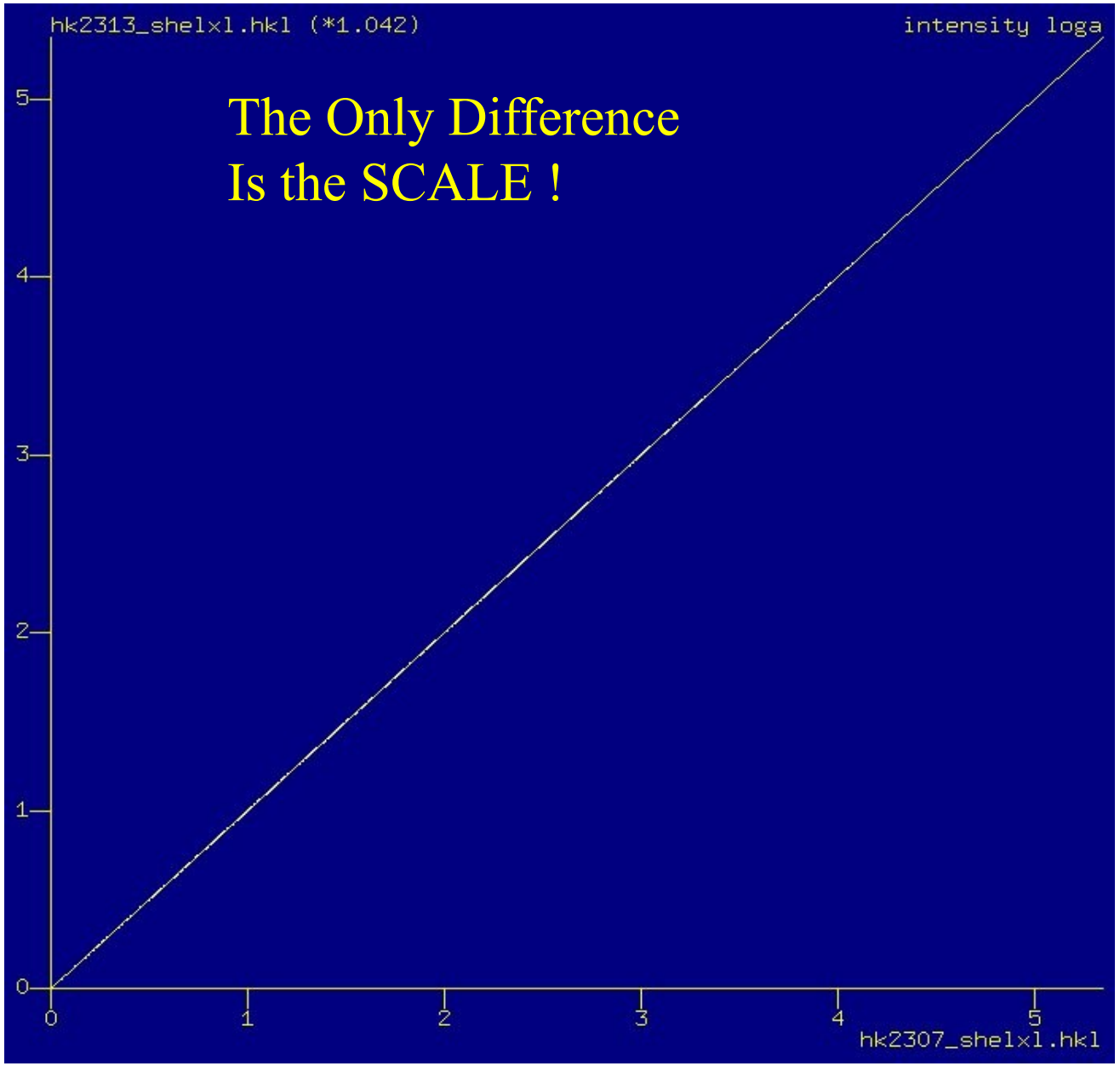
The Ultimate Shame

- Recently a whole series of ‘isomorphous’ substitions was detected for an already published structure.
- Similar series have now been detected for coordination complexes (Transition metals and lanthanides)
- How could referees let those pass ?
- Over 100 structures now retracted
- Fraud detected by looking at all papers of the same authors of a ‘strange’ structure (and their institutions)

BogusVariations (with Hirshfeld ALERTS) on the Published Structure 2-hydroxy-3,5-nitrobenzoic acid (ZAJGUM)



Comparison of the Observed data for two ‘isomorphous’ compounds.



Same
Data !

SLOPPY

Or

FRAUD ?

Summary & Conclusions

Validation Procedures:

- May save a lot of Time in Checking, both by the Investigators and by the Journals (referees).
- Often surface problems that only an experienced crystallographer might be able to detect/address.
- May point at Interesting Structural Features (Pseudo-Symmetry, short Interactions etc.) to be investigated/discussed.
- Set Quality Standards (Not just on R-Value).
- May provide Proof of a GOOD or Fraud structure.
- Deposition of structure factors is strongly advised

Thanks !

For your attention

www.cryst.chem.uu.nl/platon
www.cryst.chem.uu.nl/xraysoft
www.cryst.chem.uu.nl/platon/CIF-VALIDATION.pdf
www.cryst.chem.uu.nl/platon/FCF-VALIDATION.pdf

- Papers on structure validation:
- A.L.Spek (2003). J. Appl. Cryst. 36, 7-13.
- A.L.Spek (2009). Acta Cryst. D65, 148-155.