Crystal Structure Validation with PLATON

Ton Spek

National Single Crystal Service Facility,

Utrecht University,The Netherlands

ACA2007, July 21, 2007

Overview of the Talk

- Some Examples of Recently Published Structures with a Problem that Apparently Escaped the Attention of the Referees.
- Promote CheckCif as the Current (Partial) IUCr Solution to this problem.
- Details of what PLATON can do in this Context.
- Discussion of some more Examples of Avoidable Errors.
- Some Concluding Remarks

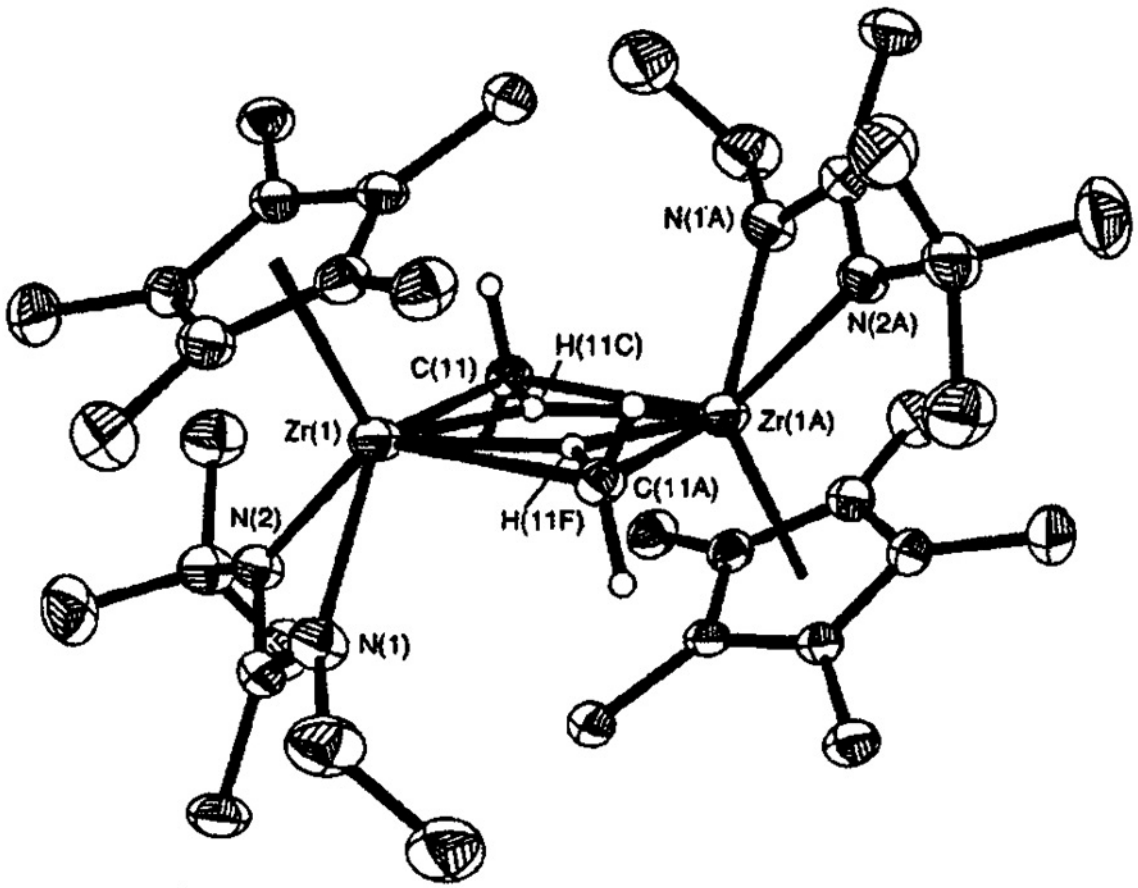


Figure 2. Molecular structure of **2** (30% thermal ellipsoids). The two $[B(C_6F_5)_4]$ anions and all but the bridging methyl hydrogen atoms, represented by spheres of arbitrary size, have been omitted for the sake of clarity. For each bridging methyl group, only the labeled hydrogen atom has been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

Structure of an Interesting 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. !

Comment

- The methyl hydrogen atoms are expected outside the Zr_2C_2 ring (and indeed have been found in similar structures)
- Referees likely had no access to (or did not access) the primary data other than the ORTEP illustration in the paper.
- General problem: A limited number of experts is available to referee too many structural papers that offer only limited primary (deposited) data.

Dalton Trans. (2001), 729-735

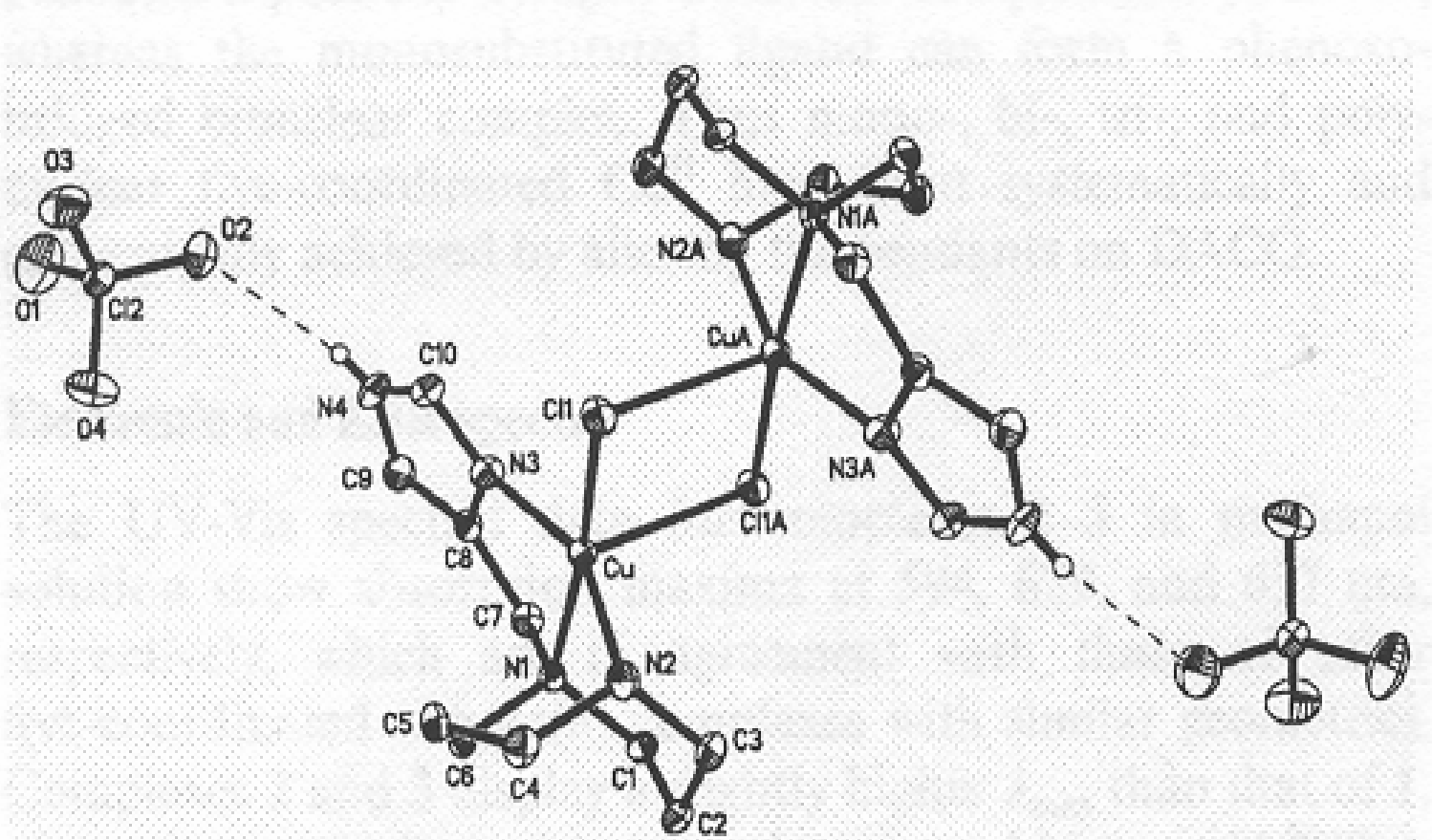
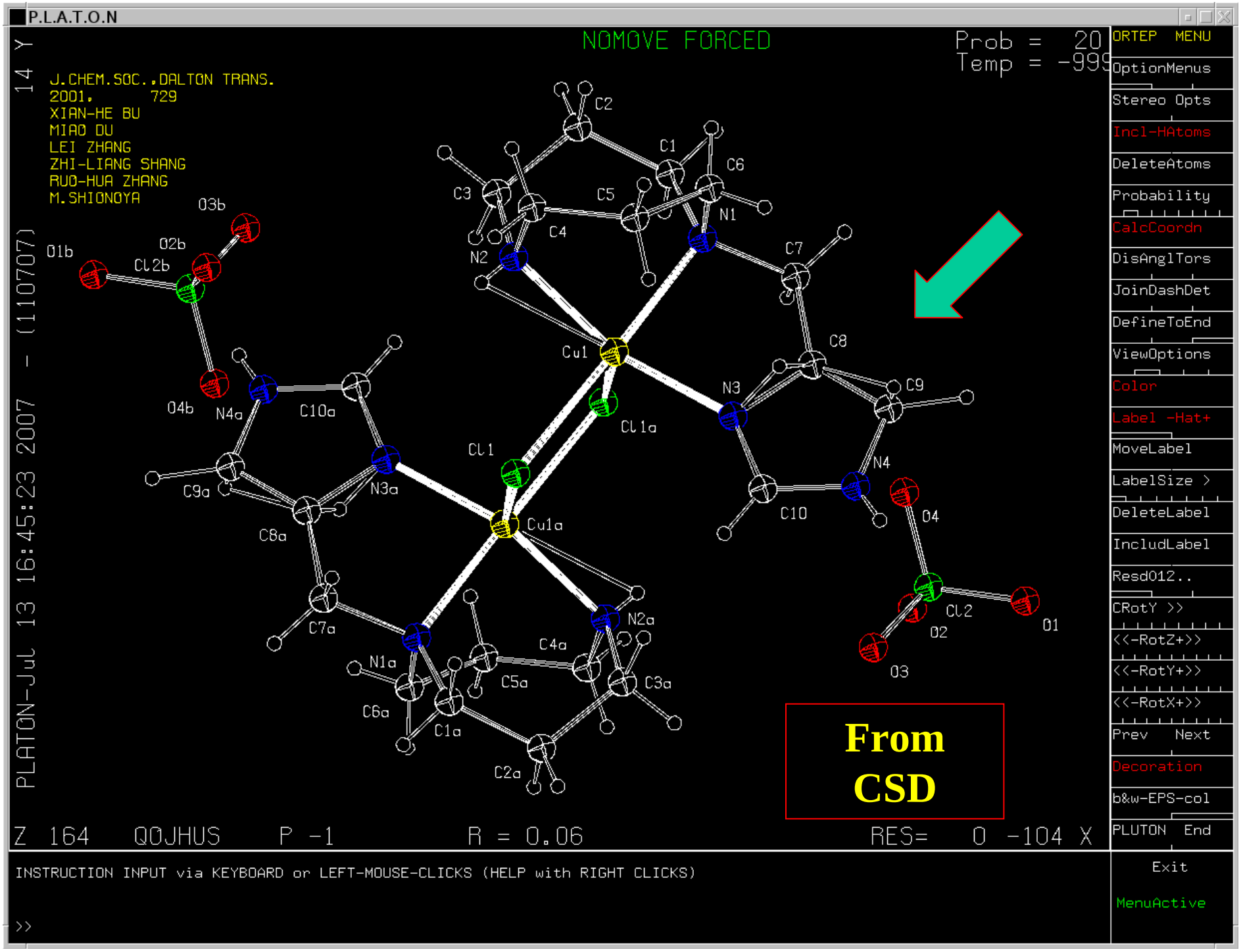


Fig. 2 The ORTEP view of complex **2** with 30% probability thermal ellipsoids.

Next Slide: ORTEP with downloaded CIF data ➔


P.L.A.T.O.N
NOMOVE FORCED
Prob = 20
Temp = -999
J.CHEM.SOC.,DALTON TRANS.
2001, 729
XIAN-HE BU
MIAO DU
LEI ZHANG
ZHI-LIANG SHANG
RUO-HUA ZHANG
M.SHIONOYA
PLATON-Jul 13 16:45:23 2007 - (110707)
14 Y
Z 164 QOJHUS P -1 R = 0.06 RES= 0 -104 X
From CSD
INSTRUCTION INPUT via KEYBOARD or LEFT-MOUSE-CLICKS (HELP with RIGHT CLICKS)
>>
ORTEP MENU
OptionMenus
Stereo Opts
Incl-HAtoms
DeleteAtoms
Probability
CalcCoordn
DisAnglTors
JoinDashDet
DefineToEnd
ViewOptions
Color
Label -Hat+
MoveLabel
LabelSize >
DeleteLabel
IncludLabel
Resd012..
CRotY >>
<<-RotZ+>>
<<-RotY+>>
<<-RotX+>>
Prev Next
Decoration
b&w-EPS-col
PLUTON End
Exit
MenuActive

Organometallics (2006) 25, 1511-1516

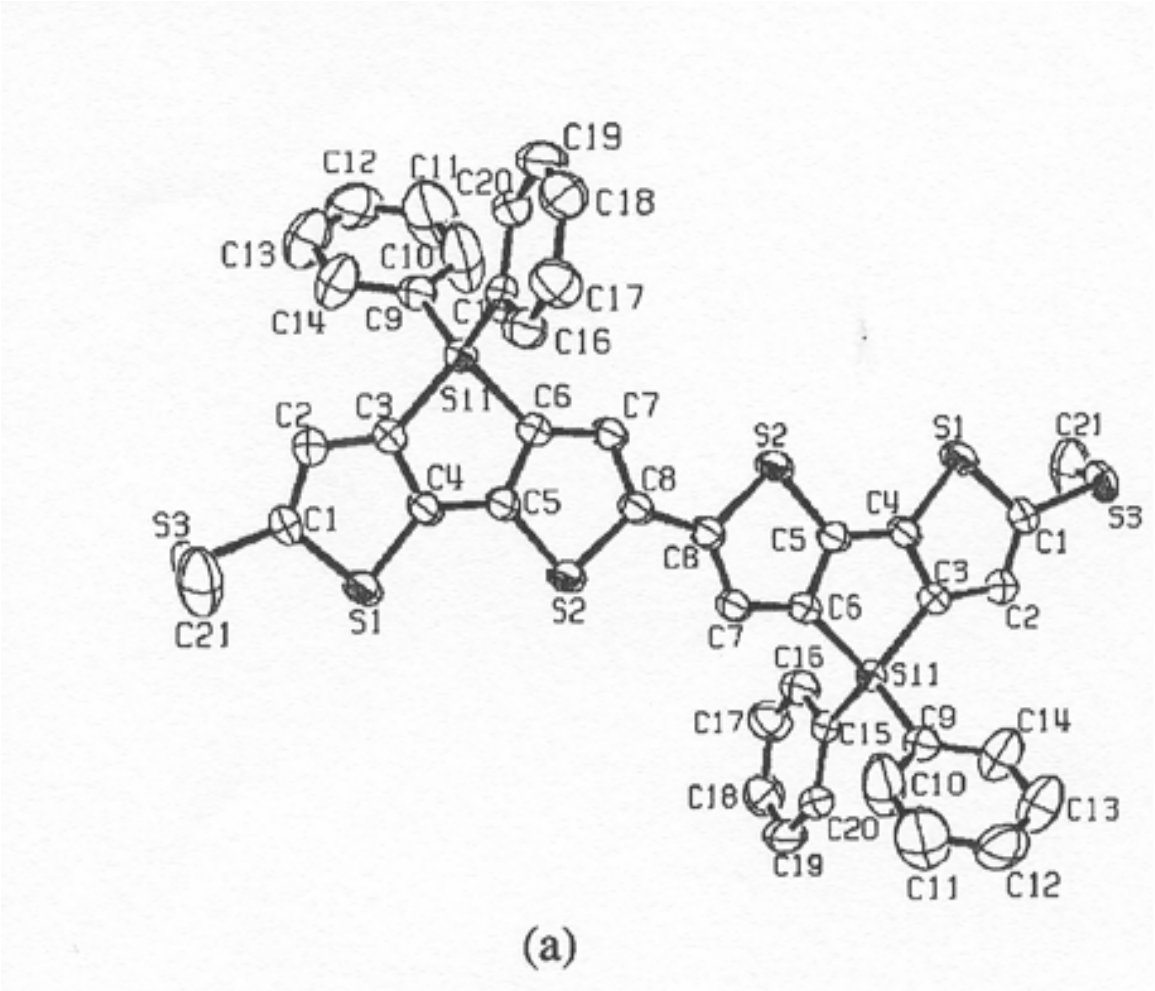
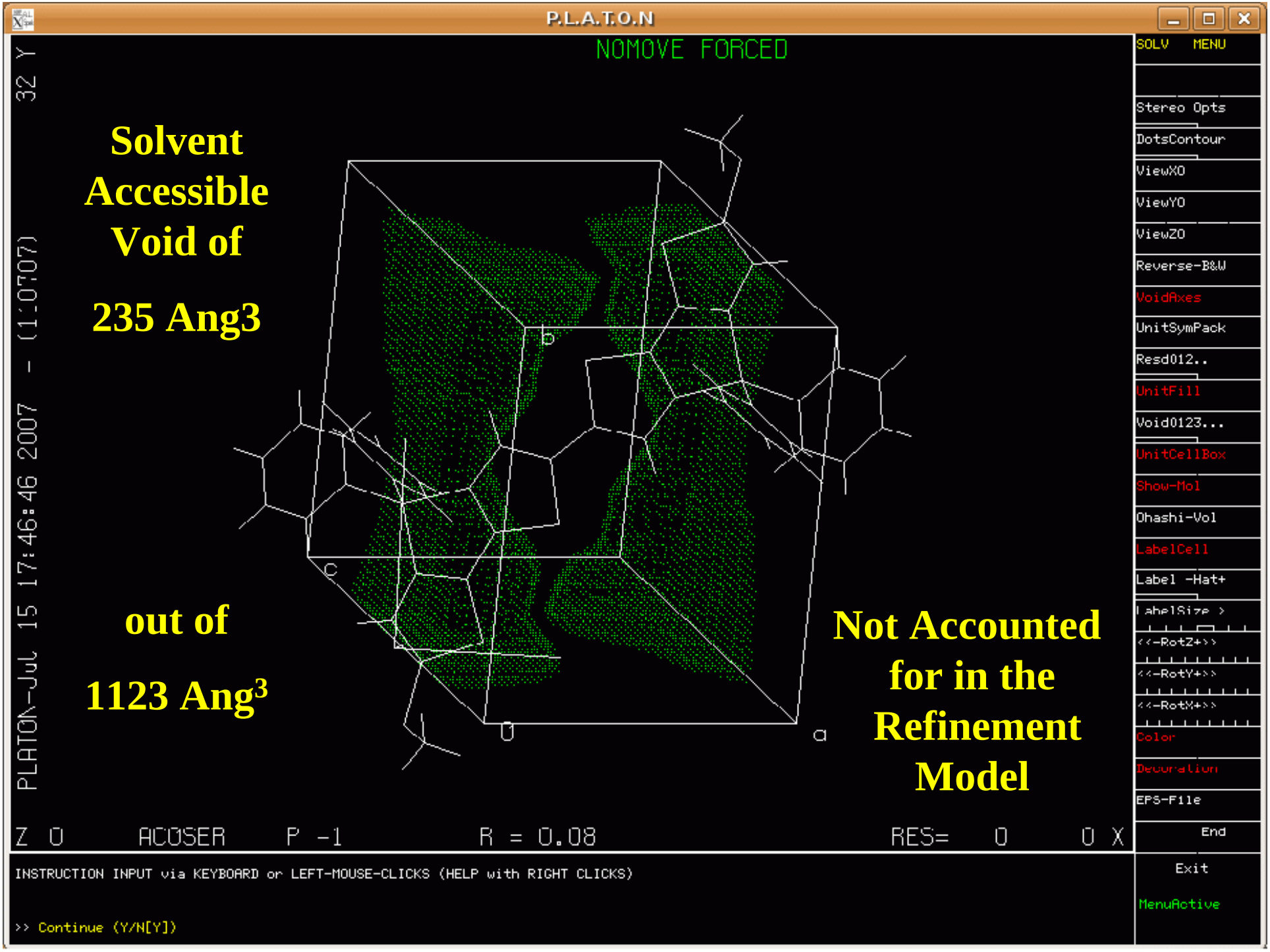


Table 2. Crystal Data, Experimental Conditions, and Summary of Structural Refinement for 5

molecular formula	$C_{42}H_{30}S_6Si_2$
molecular weight	783.23
cryst style	triclinic
space group	$P\bar{1}$ (#2)
cell dimens	
a, Å	9.3886(5)
b, Å	10.5077(1)
c, Å	12.4065(6)
α, deg	68.413(2)
β, deg	81.359(5)
γ, deg	83.288(3)
V, Å^3	1122.64(8)
Z	1
D_{calcd}, g/cm^3	1.158
F_{000}	406.00
cryst size, mm^3	0.35 × 0.35 × 0.15
cryst color	red
μ, cm^{-1}	3.84
diffractometer	Rigaku RAXIS-RAPID
temp, °C	−50
wavelength, Å	0.71069 (Mo Kα)
monochromator	graphite
no. of obsd rflns ($I > 3\sigma(I)$)	8452
R	0.079
R_w[a]	0.157

[a] Weighting scheme is $[\sum w(|F_o| - |F_c|)^2/\sum wF_o^2]^{1/2}$.

Next Slide: This is why the reported density is low and the R and Rw high ➔


P.L.A.T.O.N
NOMOVE FORCED
Solvent Accessible Void of
235 Ang3
out of
1123 Ang3
Not Accounted for in the Refinement Model
PLATON-Jul 15 17:46:46 2007 - (1.0707)
32 Y
Z 0 ACOSER P -1 R = 0.08 RES= 0 0 X
INSTRUCTION INPUT via KEYBOARD or LEFT-MOUSE-CLICKS (HELP with RIGHT CLICKS)
>> Continue (Y/N[Y])
SOLV MENU
Stereo Opts
DotsContour
ViewX0
ViewY0
ViewZ0
Reverse-B&W
VoidAxes
UnitSymPack
Resd012..
UnitFill
Void0123...
UnitCellBox
Show-Mol
Ohashi-Vol
LabelCell
Label -Hat+
LabelSize >
<<-RotZ+>>
<<-RotY+>>
<<-RotX+>>
Color
Decoration
EPS-File
End
Exit
MenuActive

SOLUTION

A solution for the structure validation problem was pioneered by the *International Union of Crystallography*

- Provide and archive crystallographic data in the computer readable CIF standard format.
- Offer Automated validation, with a computer generated report for authors and referees.
- Have journals enforce a structure validation protocol.
- The IUCr journals and most major journals now indeed implement some form of validation procedure.

THE CIF DATA STANDARD

- Driving Force: Syd Hall (IUCr/ Acta Cryst C)
- Early Adopted by XTAL & SHELX(T)L.
- Currently: WinGX,Crystals,Texsan, Maxus etc.
- Acta Cryst. C/E – Electronic Submission
- Acta Cryst.:Automatic Validation at the Gate
- CIF data available for referees for detailed inspection (and optional calculations).
- Data retrieval from the WEB for published papers
- CCDC – Deposition in CIF-FORMAT.

VALIDATION QUESTIONS

Single crystal 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?

IUCr CHECKCIF WEB-Service

http://checkcif.iucr.org reports the outcome of:

- IUCr standard tests

  Consistency, Missing Data, Proper Procedure, Quality etc.

- + Additional PLATON based tests

  Missed Symmetry, Twinning, Voids, Geometry, Displacement Parameters, Absolute Structure etc.

ALERT LEVELS

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

ALERT TYPES

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

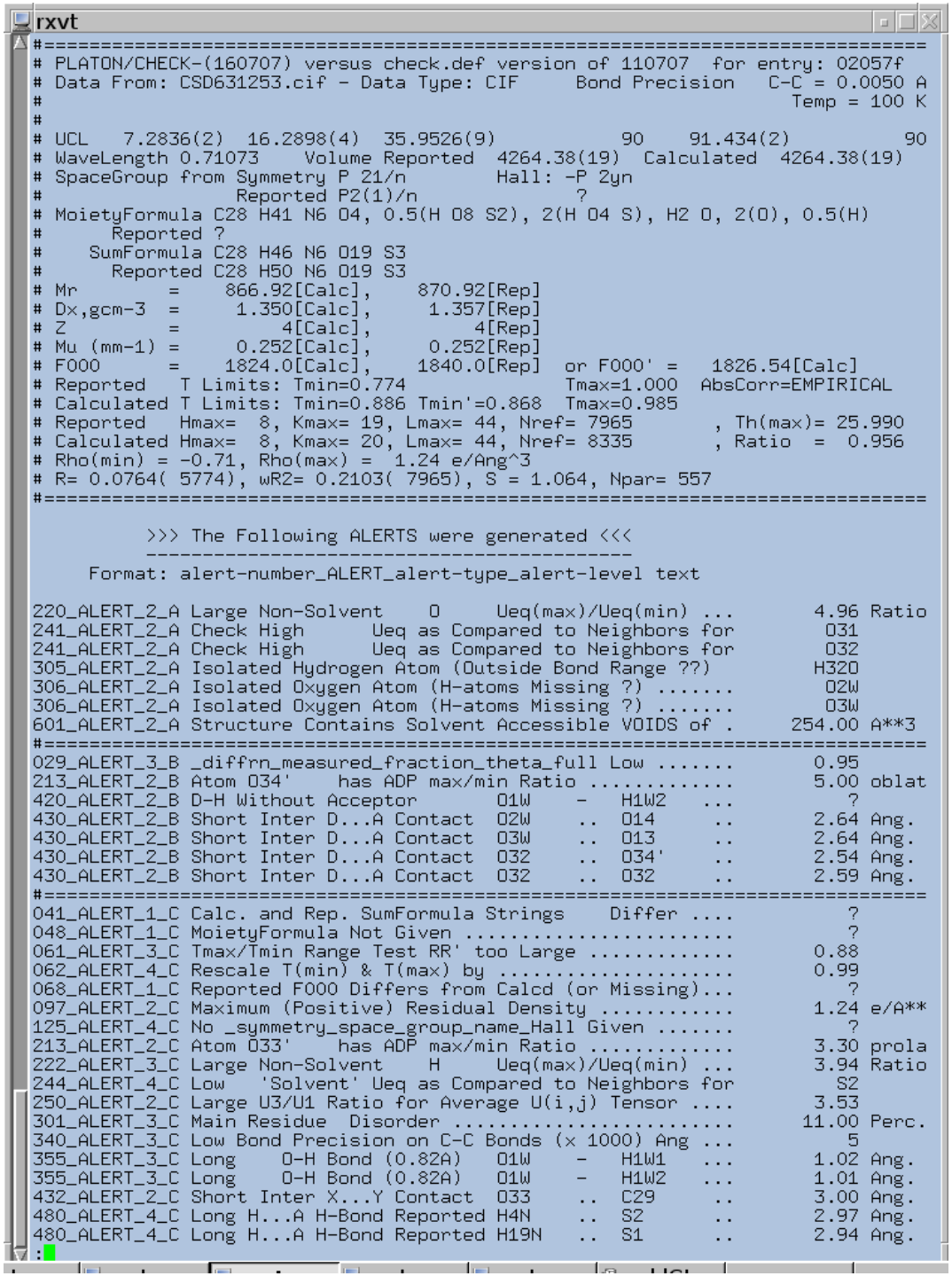
2 - Indicators that the Structure Model
may be Wrong or Deficient.

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

4 - Cosmetic Improvements, Queries and
Suggestions.

In-House Validation with PLATON

- Details: www.cryst.chem.uu.nl/platon
- Available for UNIX/LINUX, Windows, Mac-OSX
- Driven by the file **CHECK.DEF** with criteria, ALERT messages and advice.
- Unix: **platon –u structure.cif**
- Result on file: **structure.chk**
- Applicable on CIF's and CCDC-FDAT



EXAMPLE OF PLATON GENERATED ALERTS FOR A RECENT PAPER PUBLISHED IN J.Amer.Chem.Soc. (2007)

Attracted special attention in Chemical and Engineering News

Problems Addressed by PLATON/CIF-CHECK

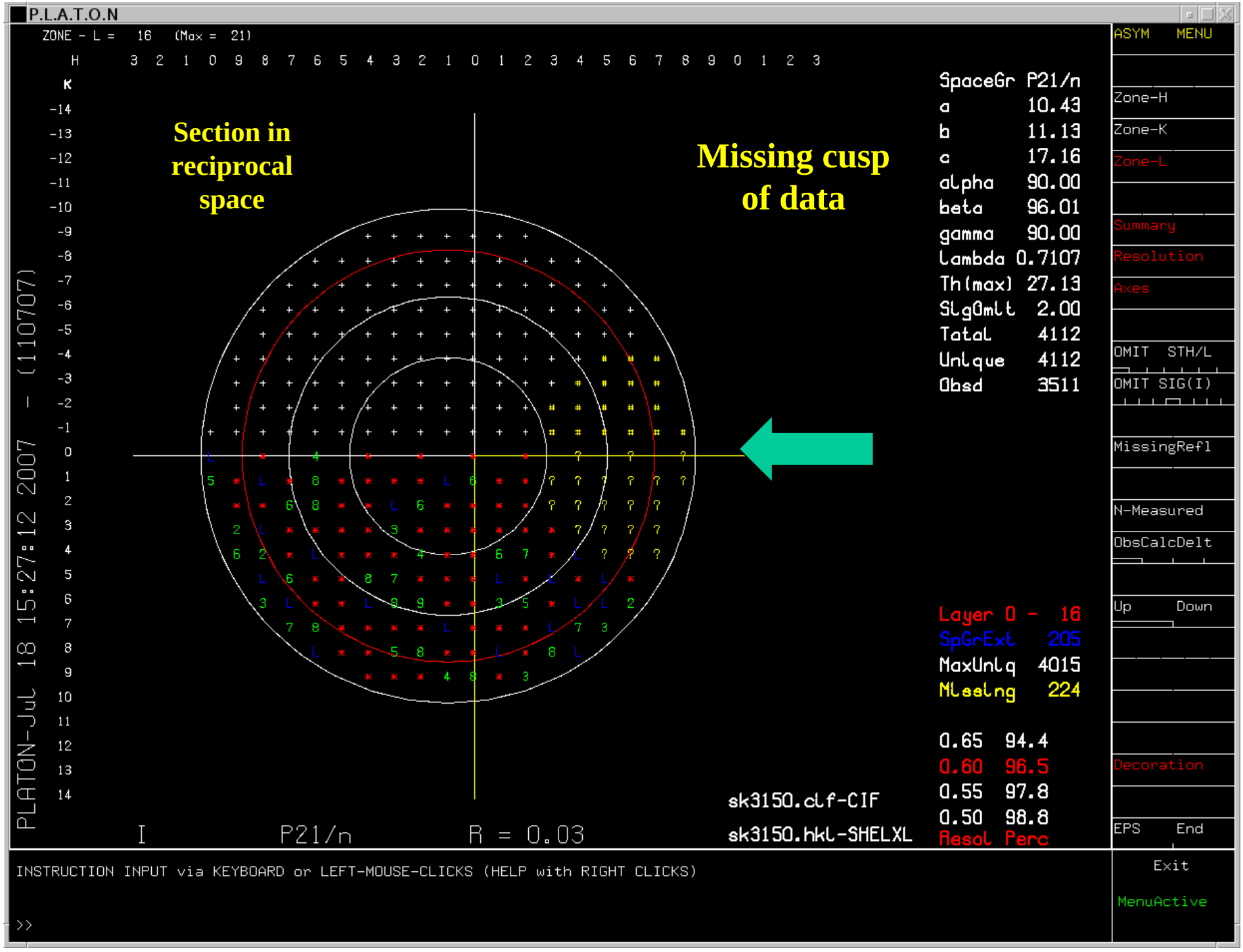
- Missed Higher Space Group Symmetry
- Solvent Accessible Voids in the Structure
- Unusual Displacement Parameters
- Hirshfeld Rigid Bond test
- Misassigned Atom Type
- Population/Occupancy Parameters
- Mono Coordinated/Bonded Metals
- Isolated Atoms (e.g. O, H, Transition Metals)

More Problems Addressed by PLATON

- Too Many Hydrogen Atoms on an Atom
- Missing Hydrogen Atoms
- Valence & Hybridization
- Short Intra/Inter-Molecular Contacts
- O-H without Acceptor
- Unusual Bond Length/Angle
- CH3 Moiety Geometry
- To be extended with tests for new problems ‘invented’ by authors.

Problems Addressed by PLATON/FCF-CHECK

- Information from .cif and .fcf files
- Report on the resolution of the data
- Report about randomly missing data
- Check the completeness of the data (e.g. for missing cusps of data
- Report on Missed (Pseudo) Merohedral Twinning
- Report on Friedel Pairs and Absolute Structure
- Next Slide: ASYM VIEW Display ➔


P.L.A.T.O.N
ZONE - L = 16 (Max = 21)
H 3 2 1 0 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 0 1 2 3
K
-14 -13 -12 -11 -10 -9 -8 -7 -6 -5 -4 -3 -2 -1 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14
PLATON-Jul 18 15:27:12 2007 - (110707)
Section in reciprocal space
Missing cusp of data
SpaceGr P21/n
a 10.43
b 11.13
c 17.16
alpha 90.00
beta 96.01
gamma 90.00
Lambda 0.7107
Th(max) 27.13
SigOmit 2.00
Total 4112
Unique 4112
Obsd 3511
Layer 0 - 16
SpGrExt 205
MaxUniq 4015
Missing 224
0.65 94.4
0.60 96.5
0.55 97.8
0.50 98.8
Resol Perc
sk3150.clf-CIF
sk3150.hkl-SHELXL
I P21/n R = 0.03
ASYM MENU
Zone-H
Zone-K
Zone-L
Summary
Resolution
Axes
OMIT STH/L
OMIT SIG(I)
MissingRefl
N-Measured
ObsCalcDelt
Up Down
Decoration
EPS End
Exit
MenuActive
INSTRUCTION INPUT via KEYBOARD or LEFT-MOUSE-CLICKS (HELP with RIGHT CLICKS)
>>

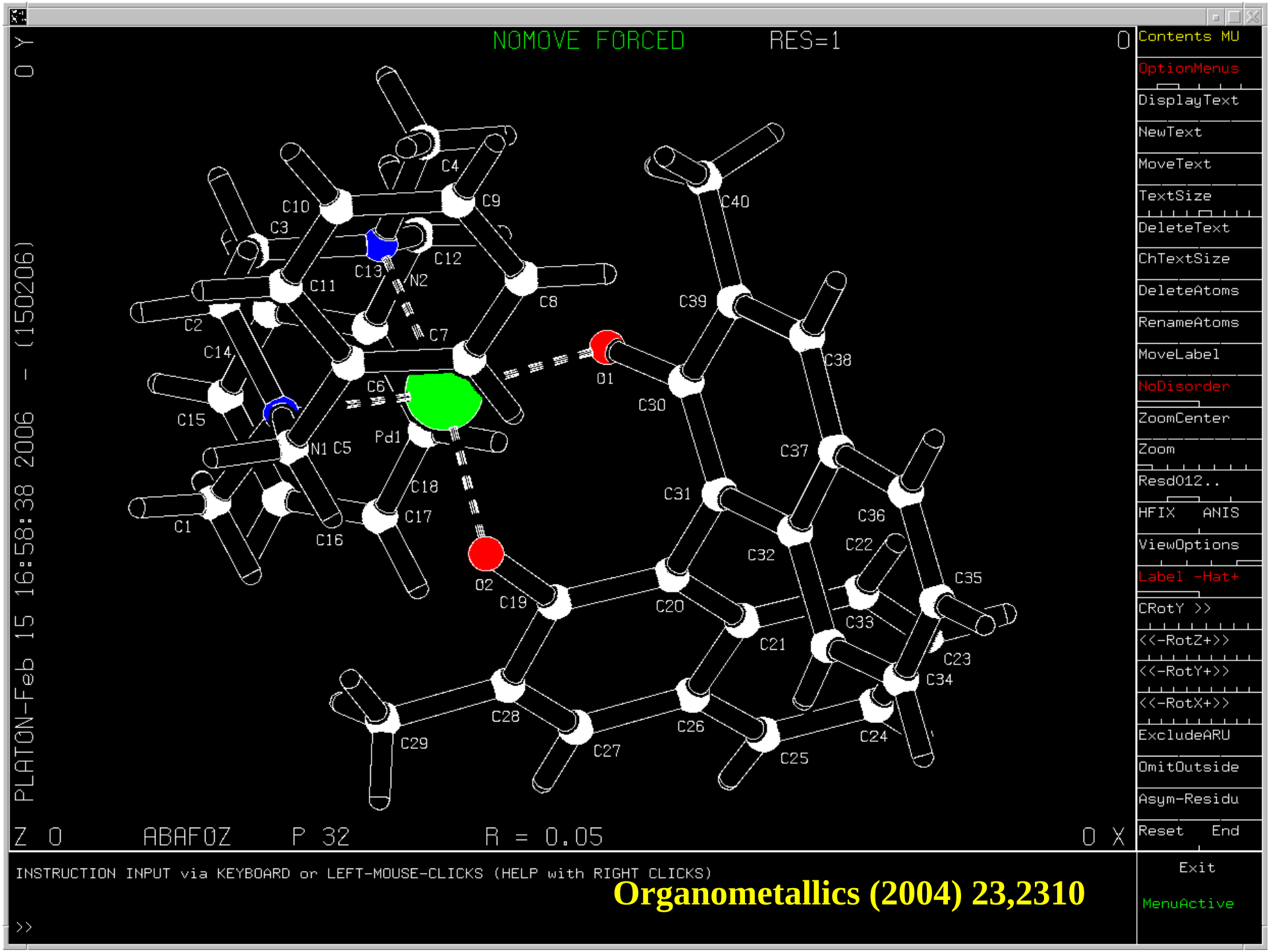
The Missed Symmetry Problem

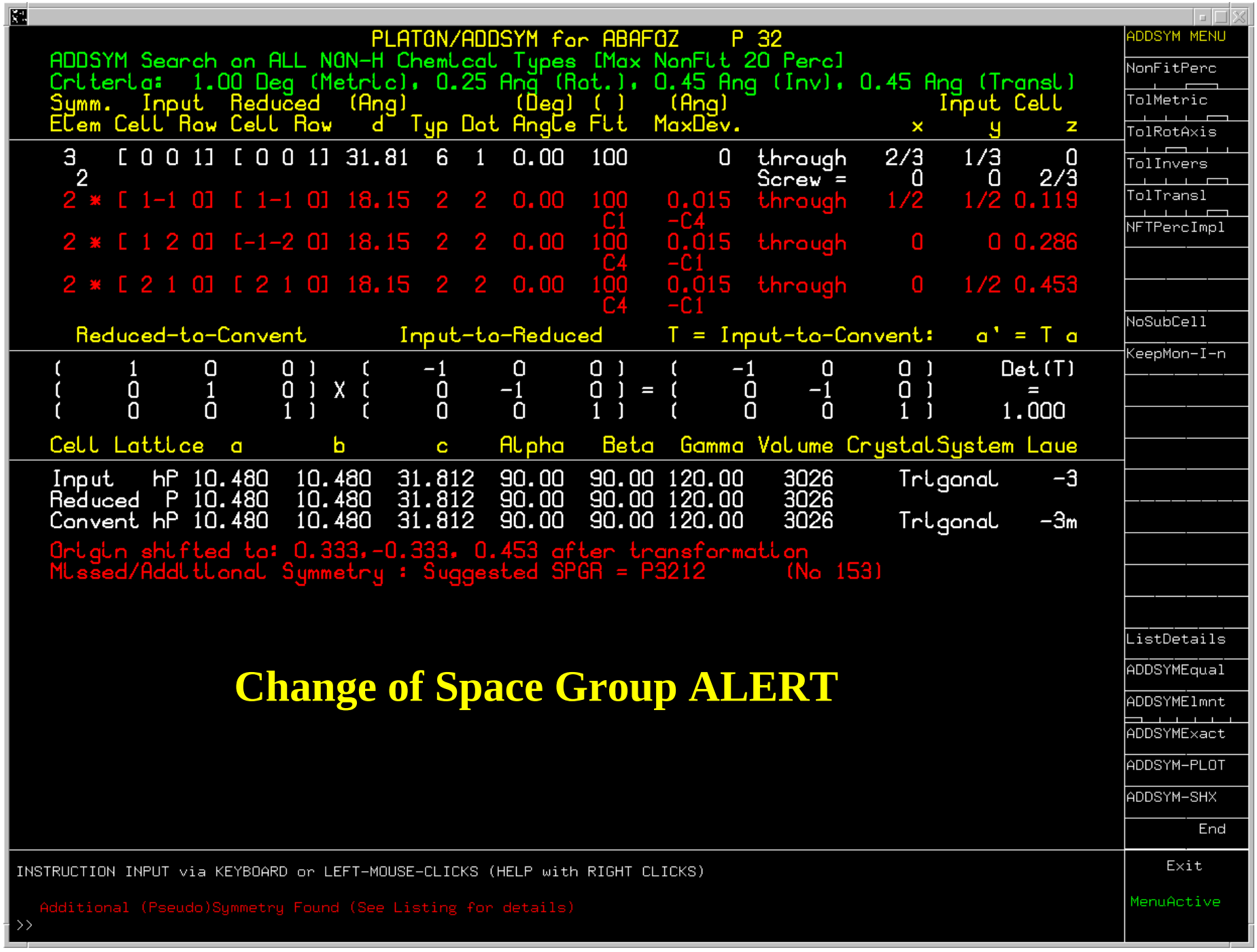
- Up to 10 % of the structures in space groups such as Cc have higher symmetry (e.g. C2/c, R-3c, Fdd2 etc.) than was originally reported.

  (To be Marshed is not good for your scientific reputation as a crystallographer).

- MISSYM (Y. LePage) & PLATON/ADDSYM algorithm addresses the problem.

- Next Slide: Example ➔


NOMOVE FORCED
RES=1
0
0 Y
PLATON-Feb 15 16:58:38 2006 - (150206)
Z 0
ABAFOZ
P 32
R = 0.05
0 X
C4
C9
C10
C3
C12
C13
N2
C11
C8
C2
C7
C14
C6
C15
Pd1
N1
C5
C18
C17
C1
C16
O2
C19
C20
C21
C28
C29
C27
C26
C25
C24
C23
C34
C35
C33
C22
C36
C32
C31
C37
C30
O1
C39
C38
C40
Contents MU
OptionMenus
DisplayText
NewText
MoveText
TextSize
DeleteText
ChTextSize
DeleteAtoms
RenameAtoms
MoveLabel
NoDisorder
ZoomCenter
Zoom
Resd012..
HFIX ANIS
ViewOptions
Label -Hat+
CRotY >>
<<-RotZ+>>
<<-RotY+>>
<<-RotX+>>
ExcludeARU
OmitOutside
Asym-Residu
Reset End
Exit
MenuActive
INSTRUCTION INPUT via KEYBOARD or LEFT-MOUSE-CLICKS (HELP with RIGHT CLICKS)
>>
Organometallics (2004) 23,2310


PLATON/ADDSYM for ABAFOZ P 32
ADDSYM Search on ALL NON-H Chemical Types [Max NonFit 20 Perc]
Criteria: 1.00 Deg (Metric), 0.25 Ang (Rot.), 0.45 Ang (Inv), 0.45 Ang (Transl)
Symm. Input Reduced (Ang) (Deg) ( ) (Ang) Input Cell
Elem Cell Row Cell Row d Typ Dot Angle Flt MaxDev. x y z
3 [0 0 1] [0 0 1] 31.81 6 1 0.00 100 0 through 2/3 1/3 0
2 Screw = 0 0 2/3
2 * [1-1 0] [1-1 0] 18.15 2 2 0.00 100 0.015 through 1/2 1/2 0.119
C1 -C4
2 * [1 2 0] [-1-2 0] 18.15 2 2 0.00 100 0.015 through 0 0 0.286
C4 -C1
2 * [2 1 0] [2 1 0] 18.15 2 2 0.00 100 0.015 through 0 1/2 0.453
C4 -C1
Reduced-to-Convent Input-to-Reduced T = Input-to-Convent: a' = T a
(1 0 0) (-1 0 0) (-1 0 0) Det(T)
(0 1 0) X (0 -1 0) = (0 -1 0) =
(0 0 1) (0 0 1) (0 0 1) 1.000
Cell Lattice a b c Alpha Beta Gamma Volume CrystalSystem Laue
Input hP 10.480 10.480 31.812 90.00 90.00 120.00 3026 Trigonal -3
Reduced P 10.480 10.480 31.812 90.00 90.00 120.00 3026
Convent hP 10.480 10.480 31.812 90.00 90.00 120.00 3026 Trigonal -3m
Origin shifted to: 0.333,-0.333, 0.453 after transformation
Missed/Additional Symmetry : Suggested SPGR = P3212 (No 153)
Change of Space Group ALERT
ADDSYM MENU
NonFitPerc
TolMetric
TolRotAxis
TolInvers
TolTransl
NFTPercImpl
NoSubCell
KeepMon-I-n
ListDetails
ADDSYMEqual
ADDSYMElmnt
ADDSYMExact
ADDSYM-PLOT
ADDSYM-SHX
End
Exit
MenuActive
INSTRUCTION INPUT via KEYBOARD or LEFT-MOUSE-CLICKS (HELP with RIGHT CLICKS)
Additional (Pseudo)Symmetry Found (See Listing for details)
>>

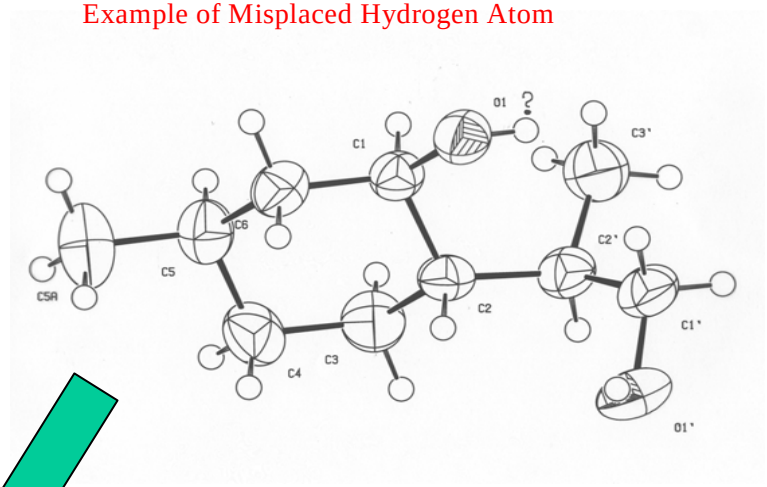
Misoriented O-H

- The O-H moiety is generally, with very few exceptions, part of a D-H..A system.
- An investigation of structures in the CSD brings up many ‘exceptions’.
- Closer analysis shows that misplacement of the O-H hydrogen atom is in general the cause.
- Molecules have an environment in the crystal !
- Example ➔

Example of a PLATON/Check Validation Report: Two ALERTS related to the misplaced Hydrogen Atom

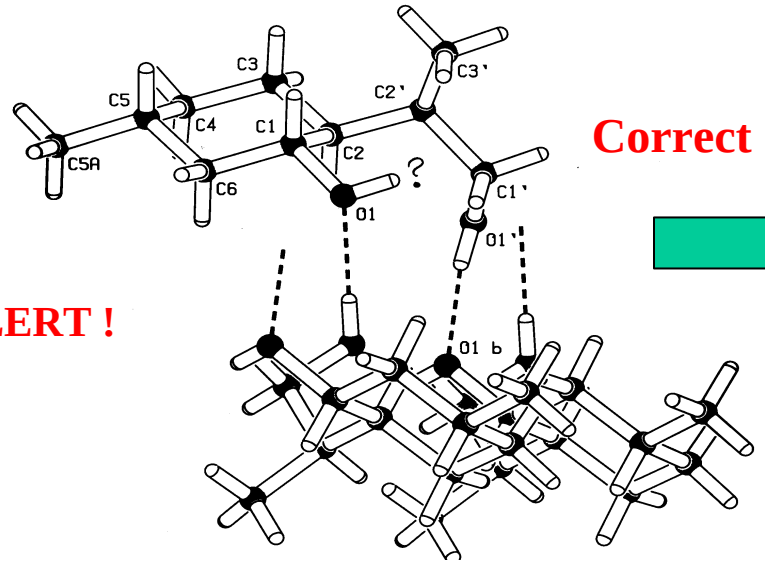
```
#===============================================================================
# PLATON/CHECK run versus check.def version of 21/06/00      for entry:I
# Data From: publ.cif - Data Type: CIF

# CELL  0.71069   8.571   6.466   9.850   90.00  106.78   90.00    522.69
# SpaceGroup P21                  Rep: P 21
# MoietyFormula C10 H20 O2
#      Reported ?
# SumFormula C10 H20 O2   Rep: C10 H20 O2
# Mr         =    172.26[Calc],    172.26[Rep]
# Dx,gcm-3   =     1.094[Calc],     1.094[Rep]
# Z          =         2[Calc],         2[Rep]
# Mu (mm-1)  =     0.074[Calc],     0.074[Rep]
# Calculated T limits: Tmin=0.993 Tmin'=0.978  Tmax=0.996
# Reported   Hmax= 10, Kmax=  7, Lmax= 11, Nref= 1935        , Th(max)= 25.76
# Calculated Hmax= 10, Kmax=  7, Lmax= 12, Nref= 1090( 1991), Ratio= 1.78( 0.97)
# R= 0.0500( 1299), wR2= 0.1870( 1935), S = 0.721, Npar= 112
#===============================================================================

>>> The Following ALERTS were generated <<<

420_ALERT B D-H without acceptor       O(1)   -   H(1)              ?         <<
#===============================================================================
048_ALERT C MoietyFormula not Given .....                        ?
084_ALERT C High R2 value ..............................      0.19
086_ALERT C Unsatisfactory S value (Too Low or not given)      0.72
142_ALERT C su on b - axis small or missing (x 100000) ...       30 Ang.
145_ALERT C su on beta     small or missing  (x 10000) ...       30 Deg.
414_ALERT C Short intra D-H..H-X       H(1)    ..  H(12') =    1.93 Ang.    <<
708_ALERT C D-H..A  Calc    170.5(5), Rep      170.00, Dev.    1.05 Sigma
              O1'  -H1'  -O1      1.555   1.555   2.646

  1 ALERT Level B = Potentially Serious Problem
  7 ALERT Level C = Check & Explain
#===============================================================================
```

Example of Misplaced Hydrogen Atom



Validation Looks at inter-molecular contacts

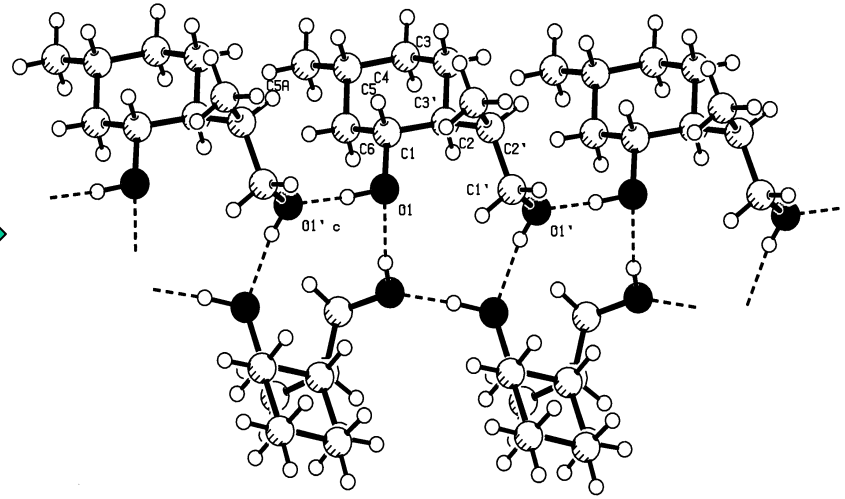
Unsatisfactory Hydrogen Bond Network



ALERT !

Correct !

Satisfactory Hydrogen Bond Network with new H-position



Wrong Structures

- Sometimes (semi) automatic structure determination procedures can come up with ‘reasonably looking’ but wrong structures.
- Structure validation software should send out proper ALERTS to the investigators

  (e.g. IUCr Checkcif)

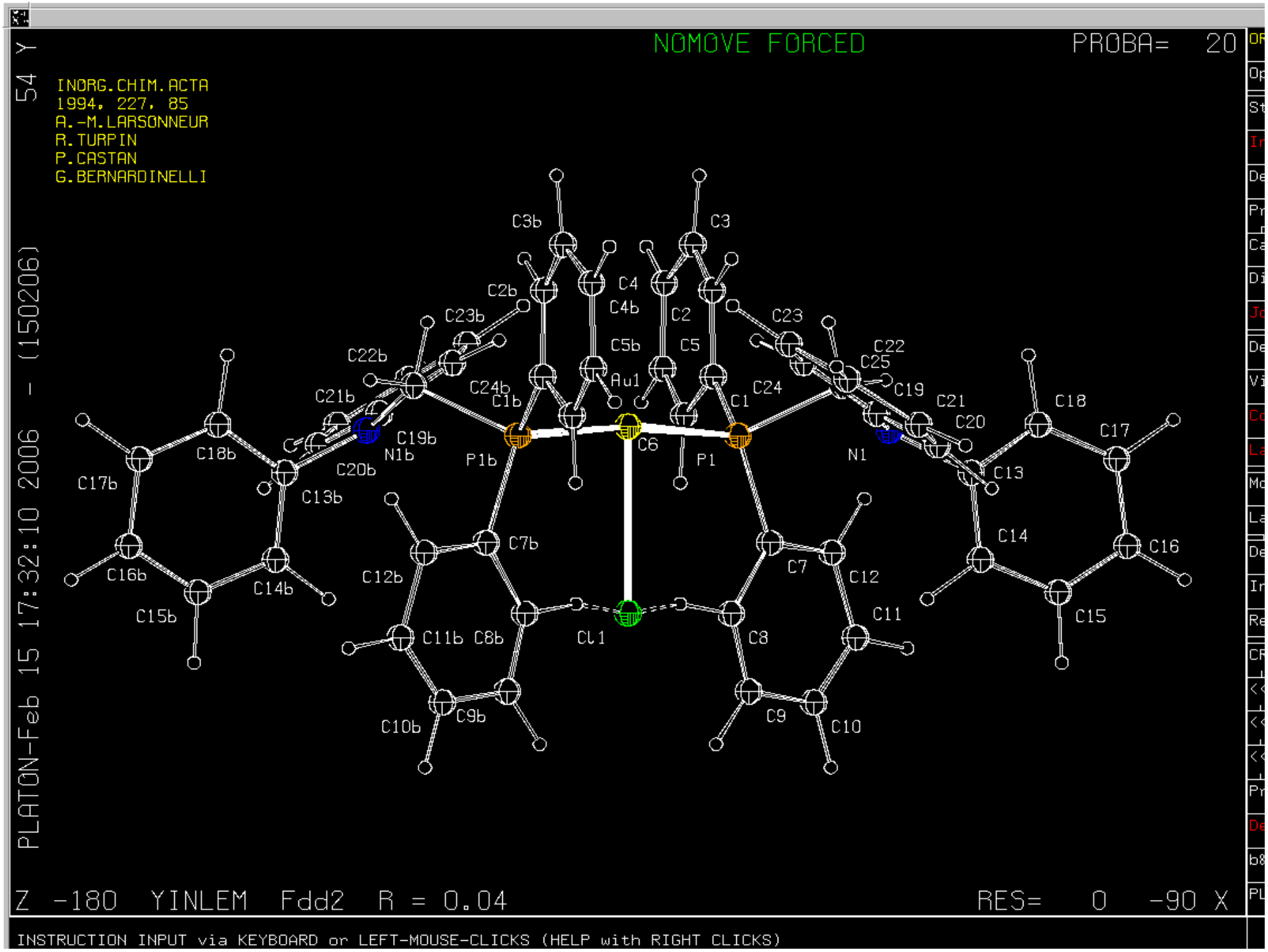
Structure Determination Artifacts

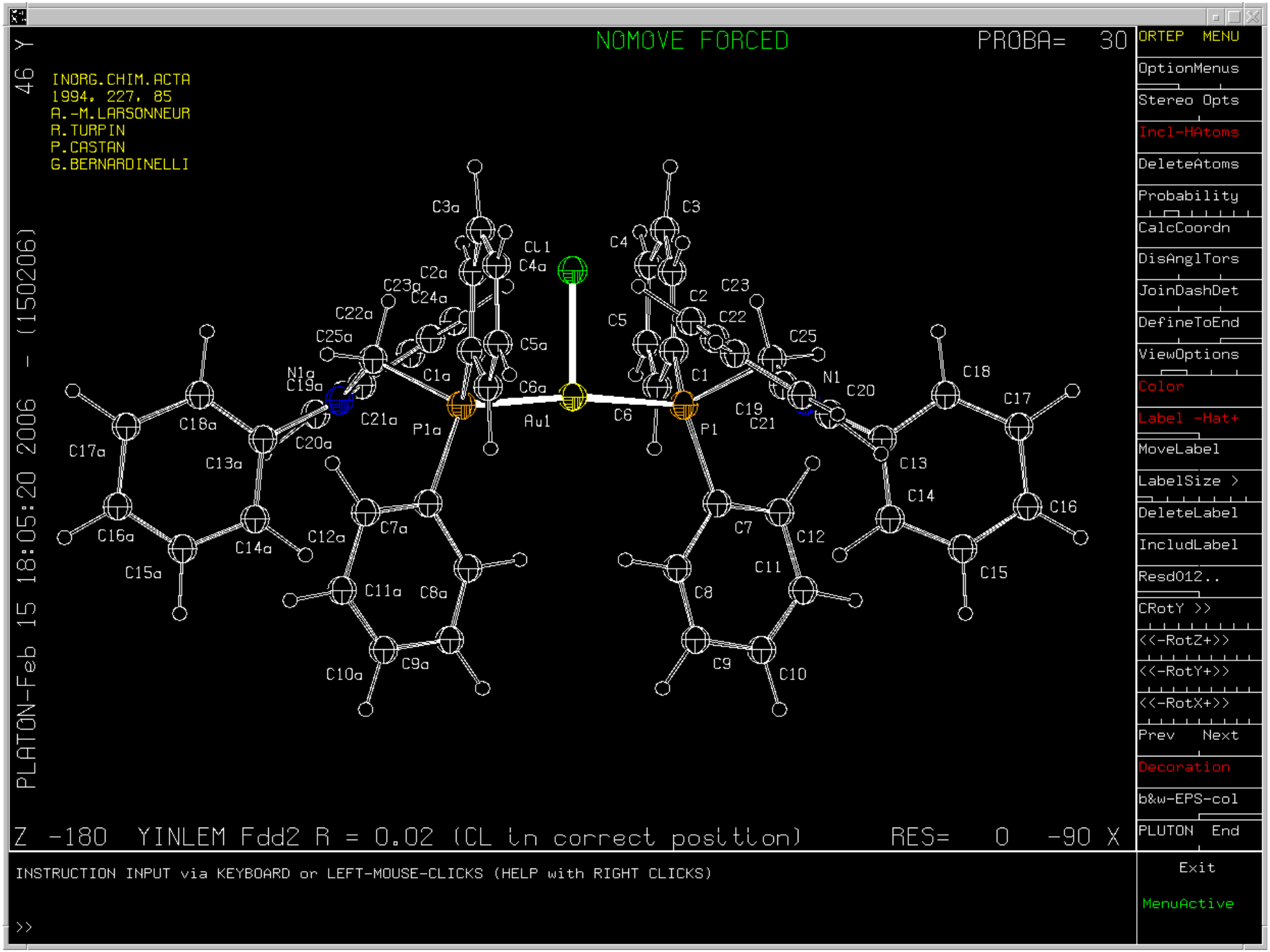
- Pseudo-symmetry easily results in false structures (often requiring a disorder model).
- Example: An Organometallic-AuCl compound from the CSD with the Cl in the

  wrong position ➔ Very Short C-H..Cl ?!

  ALERTED by validation (C..Cl = 2.19 Ang)
- Moving Cl to the correct position drops

  R from 4 to 2 % (see next slides).

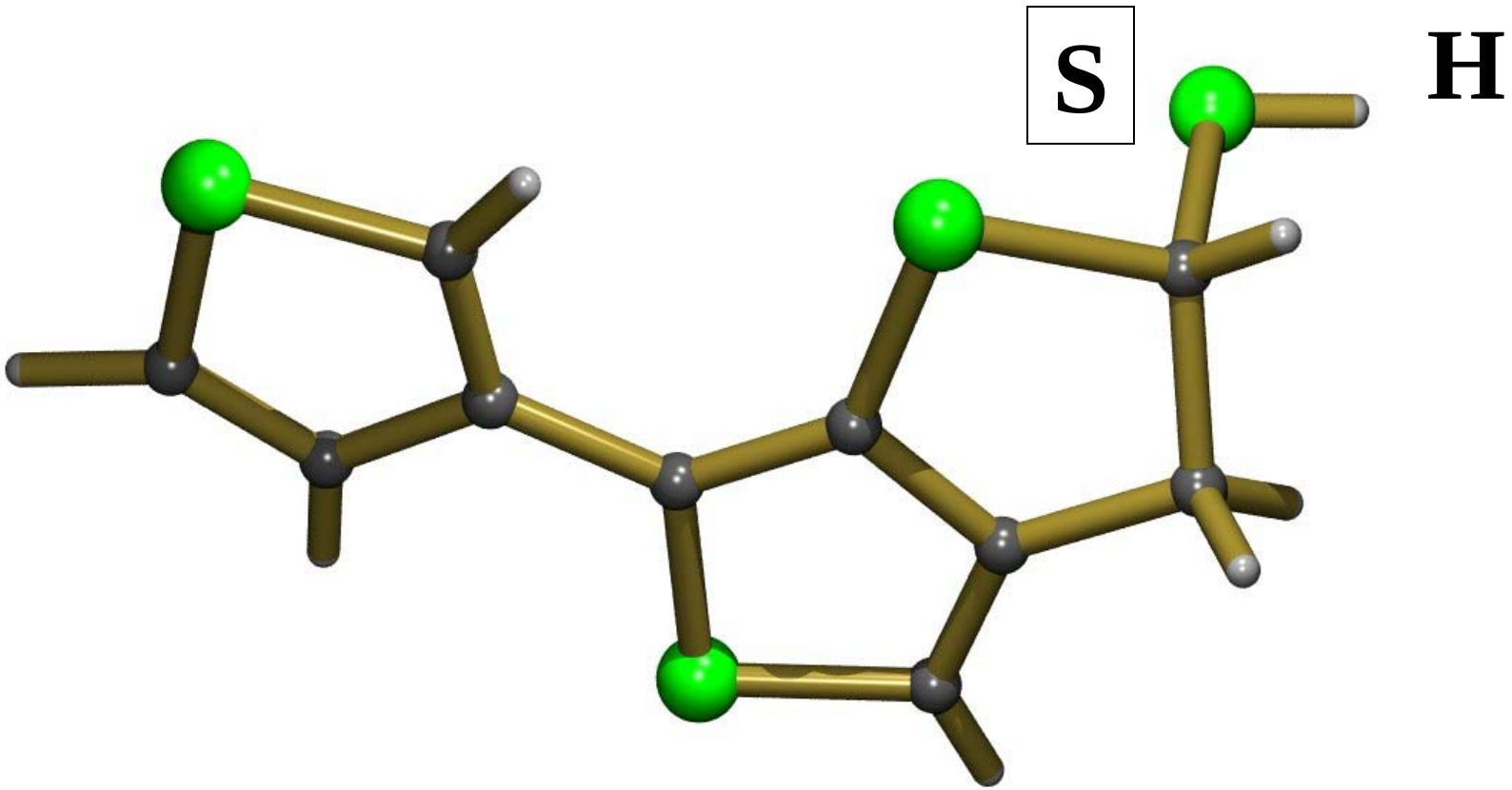

NOMOVE FORCED
PROBA= 20
INORG.CHIM.ACTA
1994, 227, 85
A.-M.LARSONNEUR
R.TURPIN
P.CASTAN
G.BERNARDINELLI
54 Y
PLATON-Feb 15 17:32:10 2006 - (150206)
C3b
C3
C4
C4b
C2b
C2
C23b
C23
C22b
C5b
C5
C22
C25
Au1
C24b
C1b
C1
C24
C19
C21
C21b
C19b
C18
C18b
N1b
C20b
P1b
C6
P1
N1
C20
C17
C17b
C13b
C13
C7b
C14
C16
C16b
C14b
C12b
C7
C12
C15b
C11
C15
C11b
C8b
Cl1
C8
C9b
C9
C10b
C10
Z -180 YINLEM Fdd2 R = 0.04
RES= 0 -90 X
INSTRUCTION INPUT via KEYBOARD or LEFT-MOUSE-CLICKS (HELP with RIGHT CLICKS)


NOMOVE FORCED
PROBA= 30
INORG.CHIM.ACTA
1994, 227, 85
A.-M.LARSONNEUR
R.TURPIN
P.CASTAN
G.BERNARDINELLI
46 Y
PLATON-Feb 15 18:05:20 2006 - (150206)
Z -180 YINLEM Fdd2 R = 0.02 (CL in correct position)
RES= 0 -90 X
INSTRUCTION INPUT via KEYBOARD or LEFT-MOUSE-CLICKS (HELP with RIGHT CLICKS)
>>
ORTEP MENU
OptionMenus
Stereo Opts
Incl-HAtoms
DeleteAtoms
Probability
CalcCoordn
DisAnglTors
JoinDashDet
DefineToEnd
ViewOptions
Color
Label -Hat+
MoveLabel
LabelSize >
DeleteLabel
IncludLabel
Resd012..
CRotY >>
<<-RotZ+>>
<<-RotY+>>
<<-RotX+>>
Prev Next
Decoration
b&w-EPS-col
PLUTON End
Exit
MenuActive

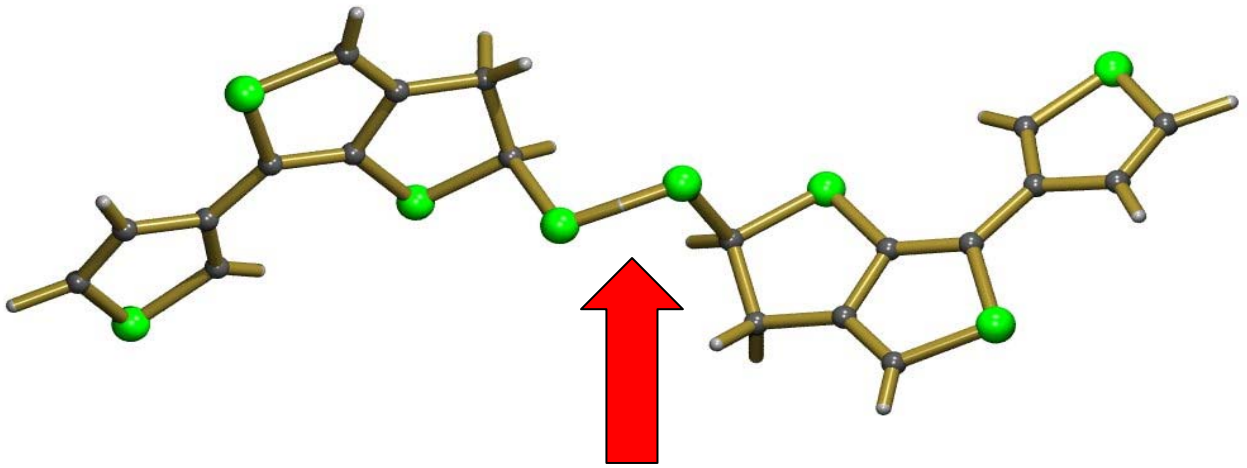
Consult the CSD

- It is a good idea to always consult the CSD for previous reports on structures related to the one at hand (in particular when the result looks unique).
- The statistics provided by VISTA (CCDC) can be very helpful for this.
- However, be aware, such an analysis often shows outliers. Many of those appear to be errors.
- Example: A search for short S..S contacts gave:

Entry from the CSD



But with Space Group Symmetry



=> Different structure with S-S Bond !

Some Statistics

- Validation CSD Entries 2006 + 2007
- Number of entries: 35760
- # of likely Space Group Changes: 384
- # of structures with voids: 3354
- Numerous problems with H, O, OH, H_2O etc.

THE MESSAGE

- Validation should not be postponed to the publication phase. All validation issues should be taken care of during the analysis.
- Everything unusual in a structure is suspect,

  mostly incorrect (artifact) and should be investigated and discussed in great detail and supported by additional independent evidence.
- The CSD can be very helpful when looking for possible precedents (but be careful)

CONCLUSION

Validation Procedures are excellent Tools to:

- Set Quality Standards (Not just on R-Value)
- Save a lot of Time in Checking, both by the Investigators and the Journals (referees)
- Point at Interesting Features (Pseudo-Symmetry, short Interactions etc.) to be discussed.
- Surface a problem that only an experienced Crystallographer might be able to Address
- Proof of a GOOD structure.

Additional Info

http://www.cryst.chem.uu.nl

(including a copy of this powerpoint presentation)

Thanks

for your attention !!