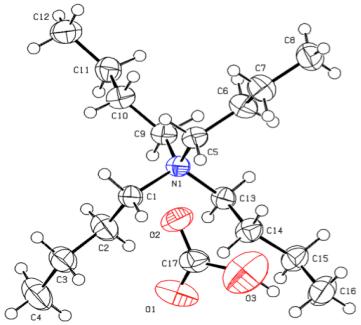
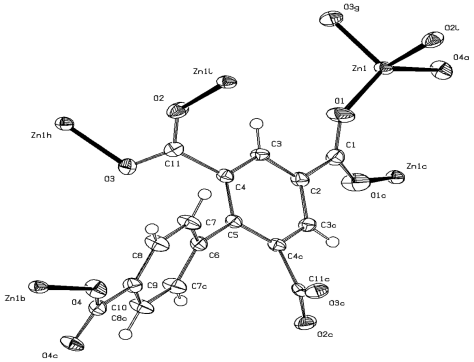
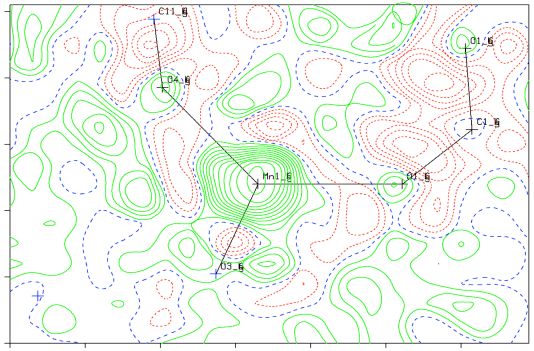






On the Proper Reporting and Archival of Crystal Structure Data

Ton Spek

Utrecht University, NL

(ACA2015-Philadelphia)

Summary Message

- Crystal structures should be published with their experimental (reflection) data archived

- Two supporting examples follow

Boring Structure ?

- Let me start with a story about a BORING looking published structure.
- The authors think that it is hot.
- Published in the high-impact journal

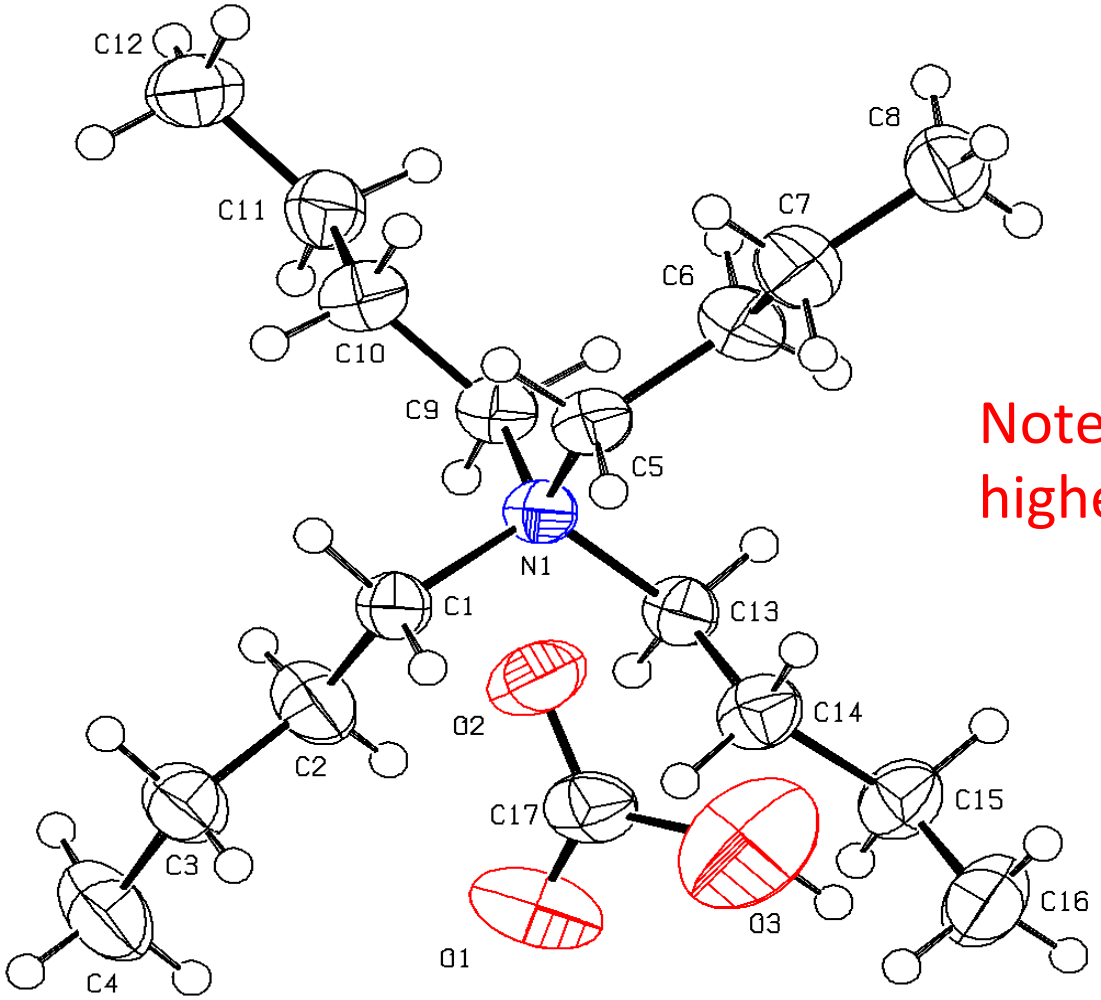
  Angew. Chem. Int. Ed.

  2015, 54, 164-168

Why an Interesting Structure ?

- Title: “Crystallographic Snapshot of an Arrested Intermediate in the Biomimetic Activation of CO_2”
- Product of the bubbling of CO_2 through a solution of n-*tetrabutylammonium hydroxide*
- $[N(C_4H_9)_4]^+[OH]^- + CO_2$ ➔ $[N(C_4H_9)_4]^+[HCO3]^-$
- Study motivated by CO_2 fixation/conversion research.

Tetrabutylammonium Hydrogencarbonate



Note the slightly higher Uij for O3

Author's interpretation of the experimental data

Selected Validation Issues

```
375_ALERT_2_A Strange C-O-H Geometry (C-O  >   1.45 Ang) .....         O3 Check
#=============================================================================
420_ALERT_2_B D-H Without Acceptor        O3     -   H3     ..     Please Check
#=============================================================================
222_ALERT_3_C Large Non-Solvent  H     Uiso(max)/Uiso(min) ...         6.1 Ratio
244_ALERT_4_C Low   'Solvent' Ueq as Compared to Neighbors of         C17 Check
340_ALERT_3_C Low Bond Precision on  C-C Bonds ...............     0.0056 Ang.
351_ALERT_3_C Long    C-H (X0.96,N1.08A)  C4     -   H4B    ..       1.13 Ang.
#=============================================================================
005_ALERT_5_G No _iucr_refine_instructions_details  in the CIF     Please Do !
008_ALERT_5_G No _iucr_refine_reflections_details   in the CIF     Please Do !
860_ALERT_3_G Number of Least-Squares Restraints .............          1 Note
#=============================================================================
```

U_{iso} (H3) = 0.9 ANG**2 !!

Interpretation Issues

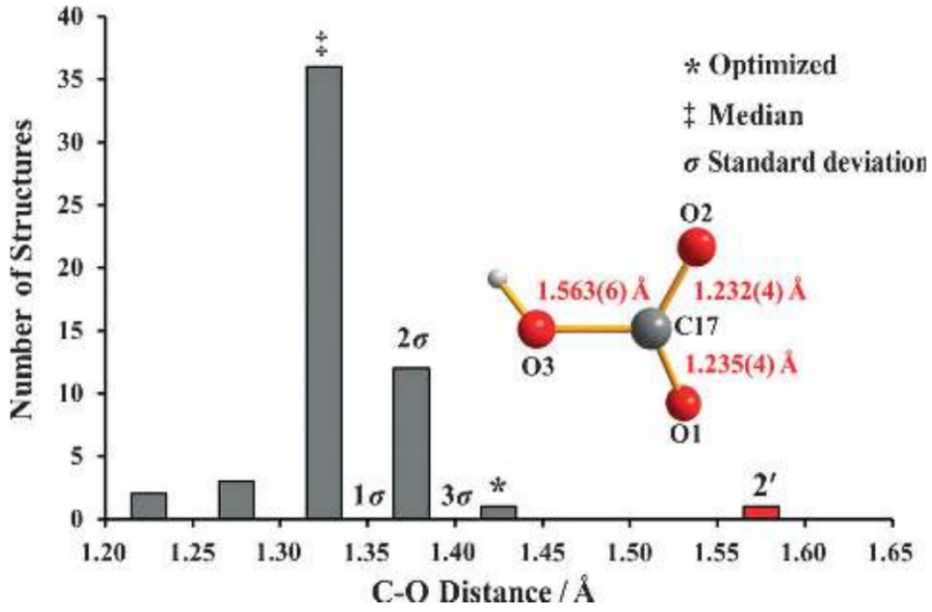


Figure 2. Distribution of C–O(H) bond lengths in all structures containing non-ligating HCO_3^- ions (72 hits from 55 structures) obtained from a survey of the CSD, along with the CCSD/6-311G(d,p) optimized HCO_3^- ion.

The authors noticed that:

- O3-H without acceptor
- The unusually long C17-O3 bond of 1.563(6) Ang.

But interpreted that as:

"H-O…CO_2 , stabilized by interactions with the cation scaffold"

Referees accepted that strange argument

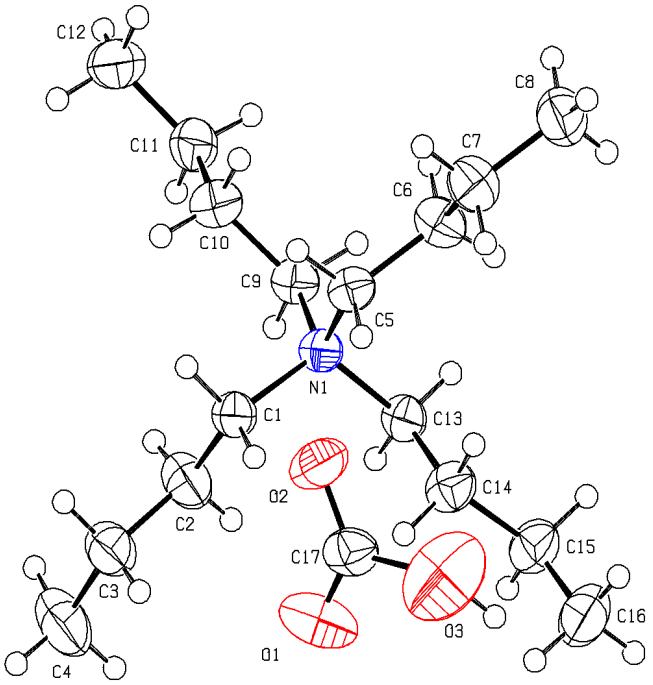
No reflection data were deposited for independent evaluation

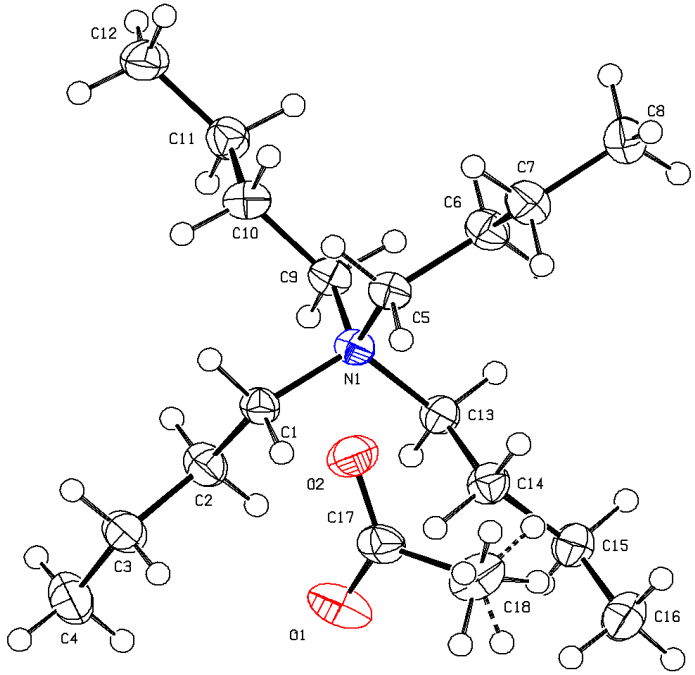
Acetate anion might fit as well

A few month's later …

- A paper appeared entitled: 'Comment on "Crystallographic Snapshot of an Arrested Intermediate in the Biomimetic Activation of CO_2" by J.A.C. Clyburn et al. (2015). Angew. Chem., 54, 7484-7487.
- New data collection with acetate as anion => The previously described mystery species is likely a common acetate ion (O-H ➔ CH_3)
- Now supported by deposited reflection data in SHELXL2014 format

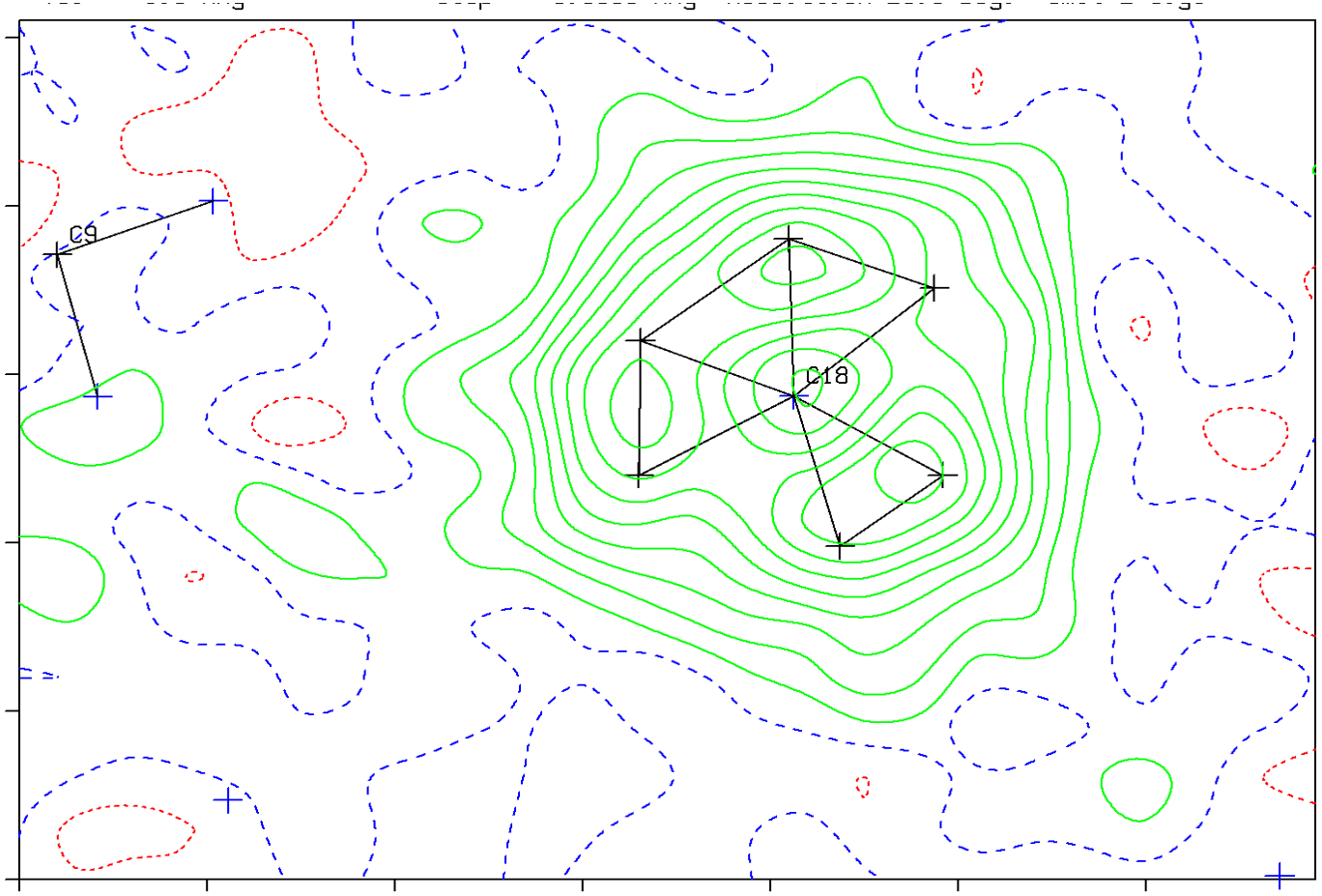
Comparison of the ‘hydrogen carbonate’ and acetate structures





Observe the close resemblance of the displacement ellipsoids

Contoured CH_3 density section in the acetate structure



We would expect to see a similar section prepared with the "HCO_3" data

Last Month 'Update'

Angewandte
Corrigendum

Crystallographic Snapshot of an Arrested Intermediate in the Biomimetic Activation of CO_2

S. L. Ackermann, D. J. Wolstenholme,*
C. Frazee, G. Deslongchamps,
S. H. M. Riley, A. Decken,
G. S. McGrady* _____ **164–168**

Angew. Chem. Int. Ed. **2015**, *54*

DOI: 10.1002/anie.201407165

The authors of this Communication now believe that a single crystal of a $[(n\text{-}C_4H_9)_4N]^+[CH_3CO_2]^-$ impurity in their bulk sample was misidentified as $[(n\text{-}C_4H_9)_4N]^+[O_2C\cdots OH]^-$ in their original report. This was interpreted as an arrested intermediate in the base-mediated activation of CO_2, indicating an egregiously long C—O(H) bond that is in fact the C—C bond of the $[CH_3CO_2]^-$ salt. However, the biomimetic nature of the hydrophobic pockets in $[(n\text{-}C_4H_9)_4N]^+[CH_3CO_2]^-$ still offers valuable insights into the substrate binding sites in human carbonic anhydrase II, since the geometry of the weak hydrogen bonding is identical to the interactions that tether CO_2 to the pockets of this metalloenzyme. Moreover, the computational component of the original study revealed a three-stage process for the base-mediated activation of CO_2, which is independent of the experimental structural study.

Afterall a dull structure …..
but perfectly illustrating my message
By-the-way, where did the acetate anion come from?

Important Message

- Unusual structural features are most likely wrong
- A reported crystal structure (i.e. coordinates and atom type assignment) is just the authors interpretation of the data and open for debate
- Deposition of the data is needed to prove/disprove an authors interpretation (by referees and readers)
- Not all chemical journals enforce the good scientific practice of the deposition of the reflection data yet
- Programs like SHELXL2014 now automatically include the reflection data in the deposited CIF. Those data should not be deleted for whatever reason

Deposition/Archival

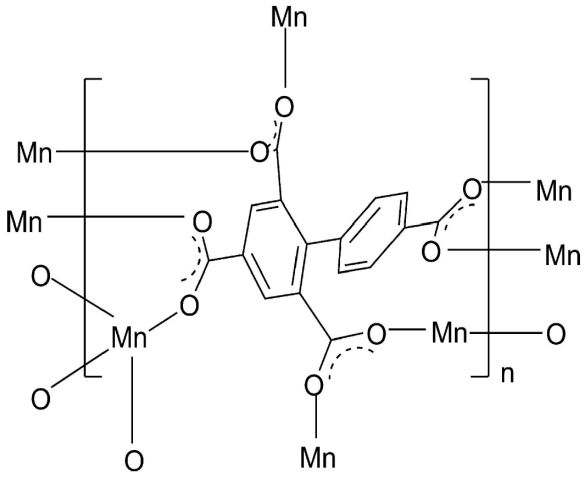
- The IUCr journals require as a minimum the deposition of Fo/Fc data. Better now: include the unmerged reflection data into the CIF.
- Some programs do that now automatically, including the refinement instructions.
- The CCDC is now fully equipped and happy to accept reflection data for archival.
- Private communications to the CCDC should include reflection data to make more sense.

Structure Validation

- A crystal structure needs to be fully validated before publication on the basis of both the refinement results and experimental (reflection) data on which they are based.
- Tools available to authors, referees and readers should include checkCIF, the Cambridge Crystallographic database, a re-refinement program and very important: chemical knowledge
- All non-standard procedures used (e.g. constraints/restraints, disorder, SQUEEZE) should be sufficiently detailed in a structure report
- All unusual structural features should be investigated and discussed in sufficient detail.
- A refereeing example of an actual structure report follows.

A Structure Report submitted to Acta Cryst. C (2015)

- Title: ”The synthesis and DFT calculation of a four-coordinated manganese compound Mn_2(BPTC)”.



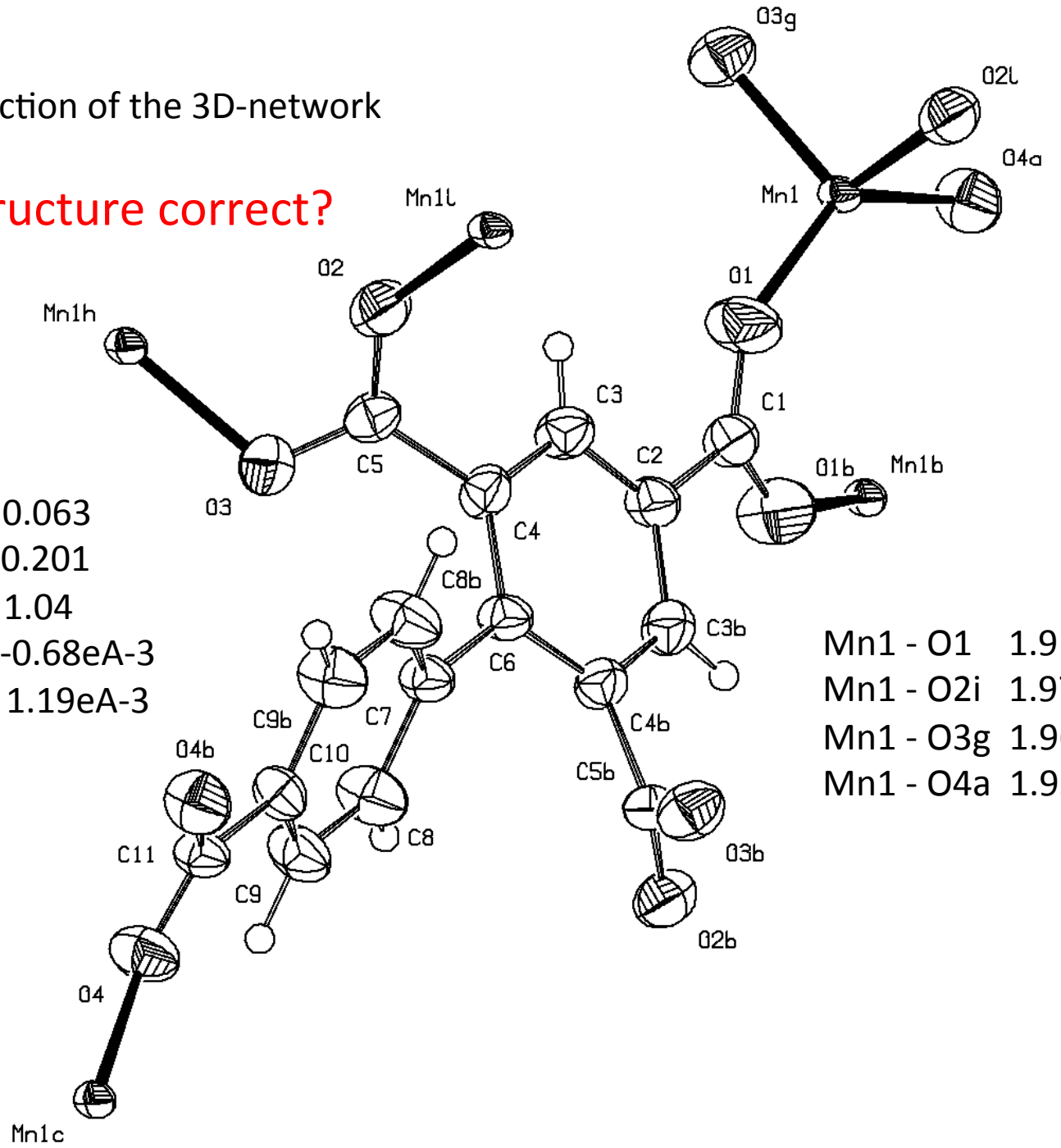
- 3D-Network

- Prepared in a stainless steel reactor at 443K from $Mn(ClO_4)_2.6H_2O$ + H_4BPTC ligand.

Relevant section of the 3D-network

293 K

Is this structure correct?



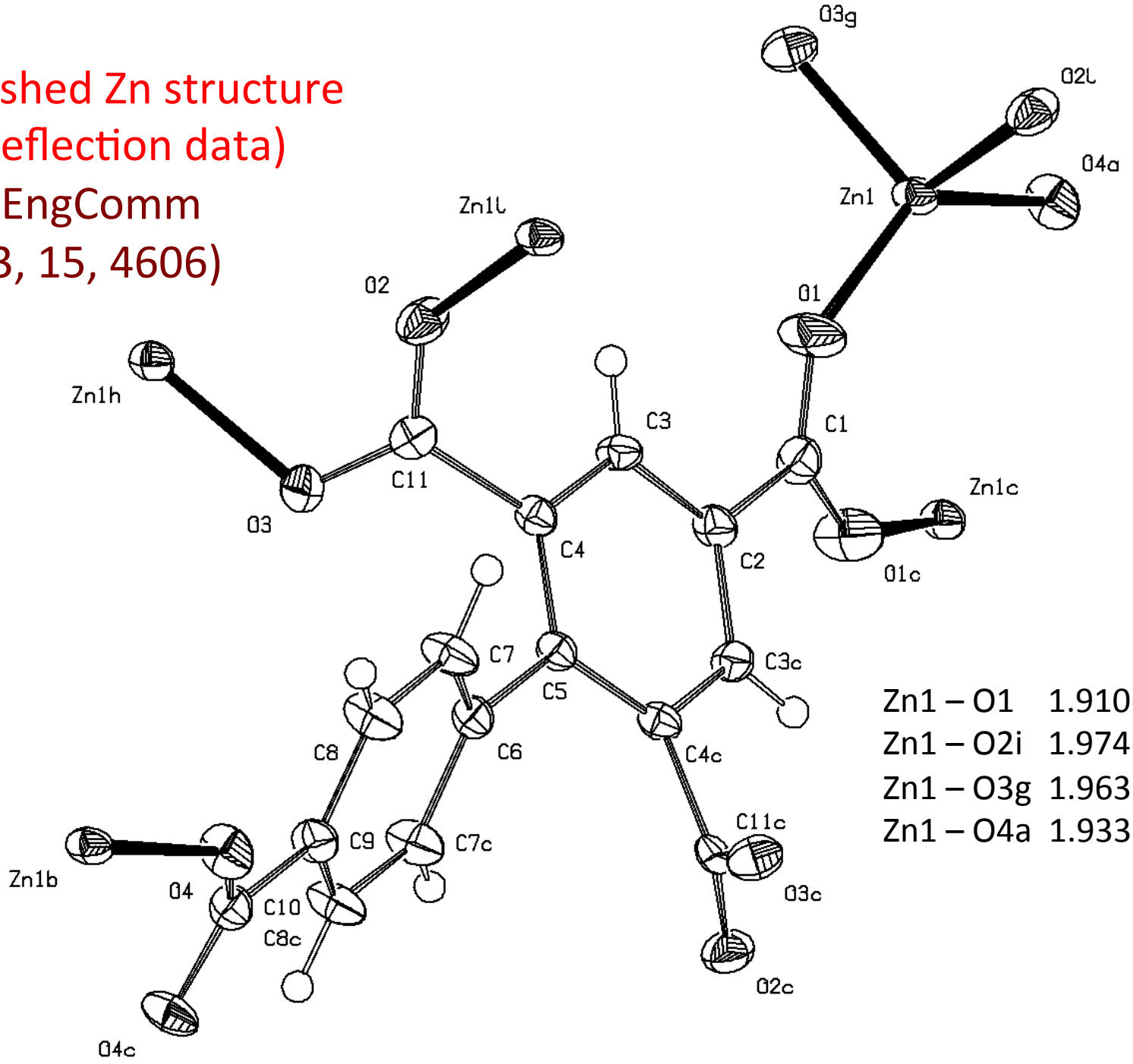
R_1 = 0.063
wR_2 = 0.201
S = 1.04
Rho(min) = -0.68eA-3
Rho(max) = 1.19eA-3

Mn1 - O1 1.914(4) Ang
Mn1 - O2i 1.974(5)
Mn1 - O3g 1.960(4)
Mn1 - O4a 1.935(4)

CSD-Search

- The Mn 4-coordination is somewhat unusual:
- Coordination generally square planar rather than tetrahedral.
- An isomorphous Zn complex was reported previously (no deposited reflection data, thus no comparison of reflection data possible)
- The coordination around Zn closely fits the one of the Mn complex

Published Zn structure
(No reflection data)
CrystEngComm
(2013, 15, 4606)

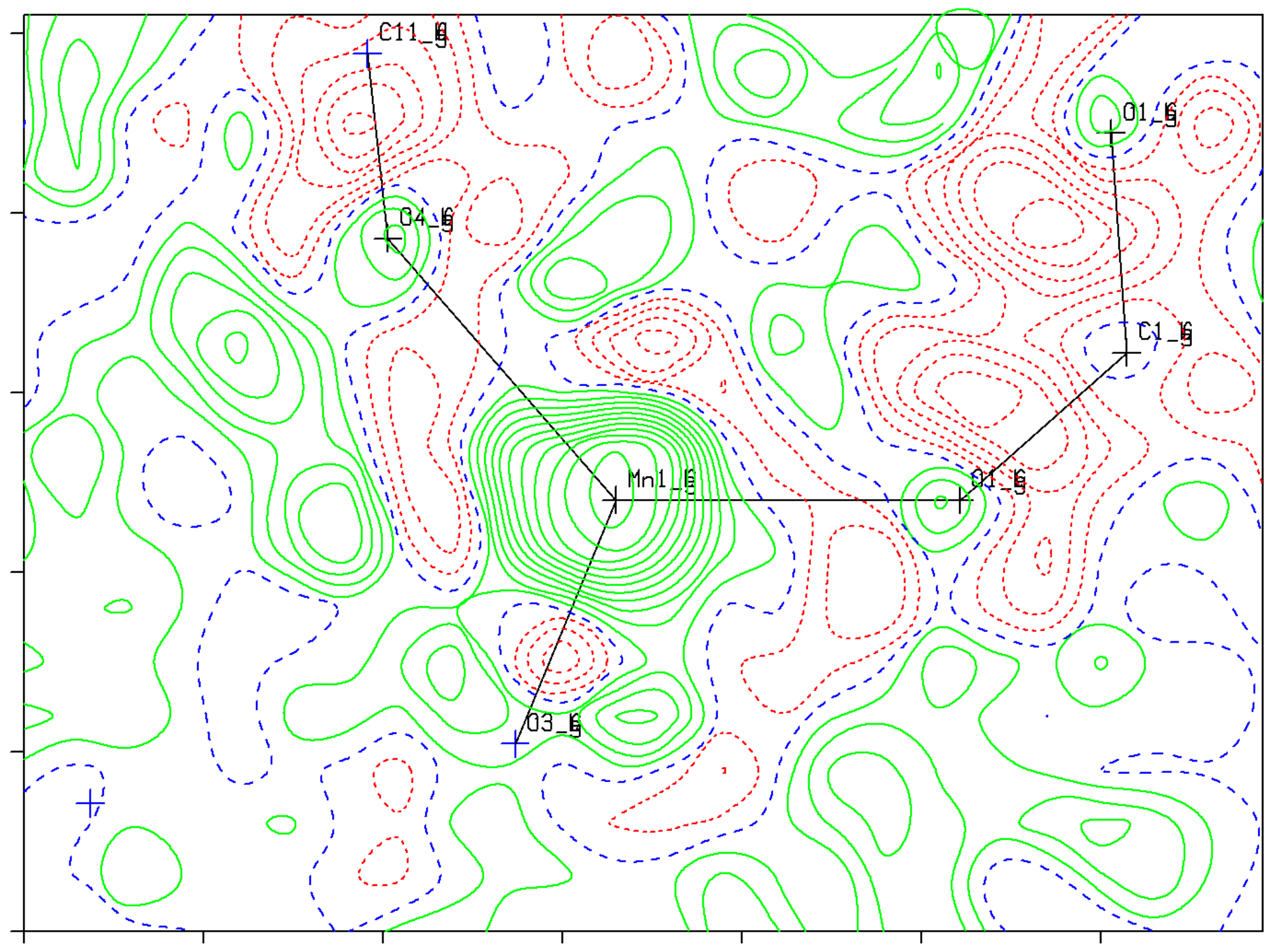


Zn1 – O1 1.910(2)
Zn1 – O2i 1.974(2)
Zn1 – O3g 1.963(2)
Zn1 – O4a 1.933(2)

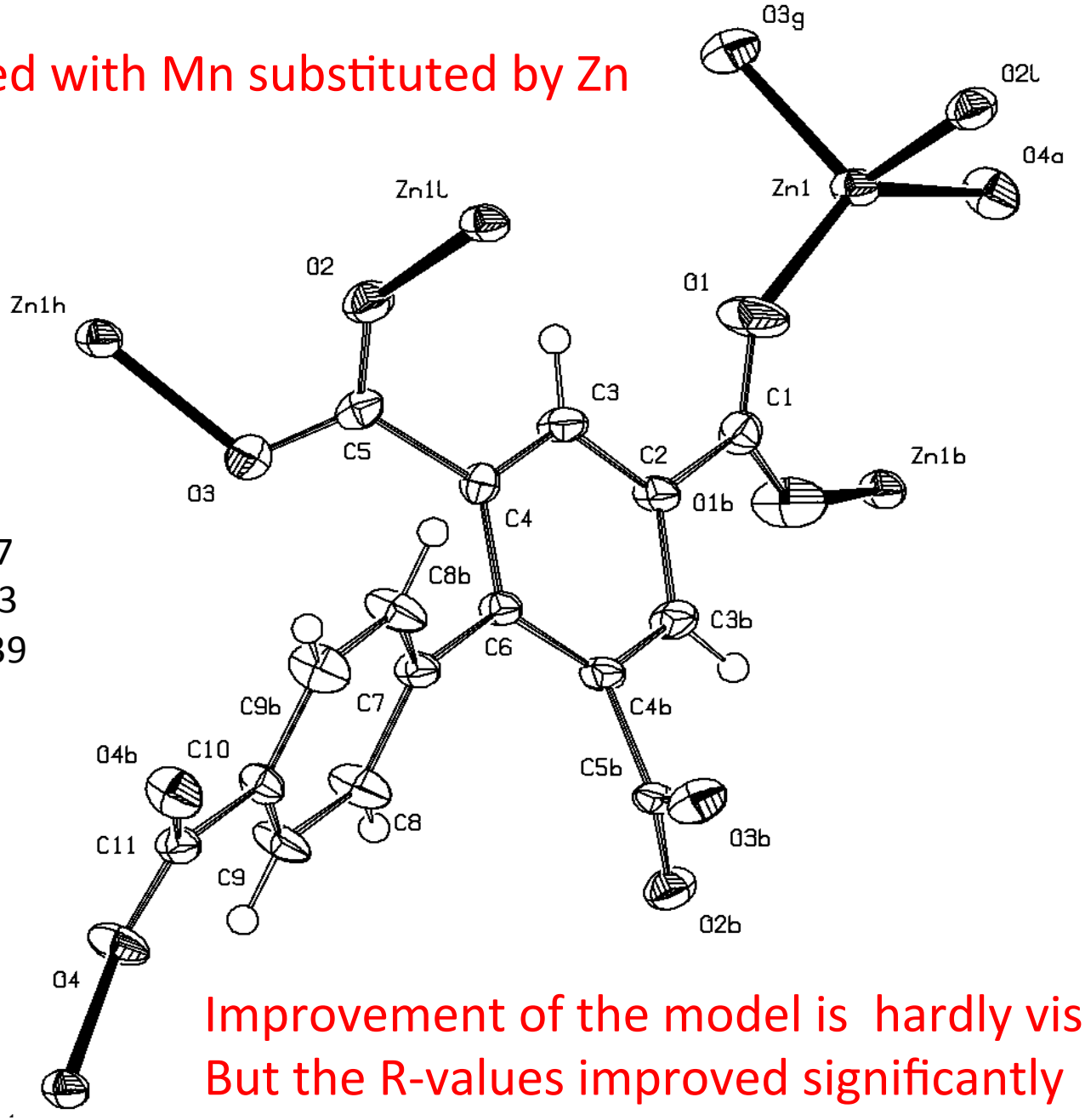
checkCIF

- What can checkCIF tell us about the Mn complex?
- Also, always inspect the difference density map for unusual features

```
973_ALERT_2_B Check Calcd Positive Residual Density on     Mn1        1.55 eA-3
#=============================================================================
241_ALERT_2_C High      Ueq as Compared to Neighbors for .....          01 Check
341_ALERT_3_C Low Bond Precision on  C-C Bonds ...............      0.0080 Ang.
430_ALERT_2_C Short Inter D...A Contact  O2      ..  O4       ..       2.88 Ang.
790_ALERT_4_C Centre of Gravity not Within Unit Cell: Resd.  #           1 Note
              C8 H3 Mn O4
976_ALERT_2_C Check Calcd Residual Density  0.63A From      O2       -0.56 eA-3
976_ALERT_2_C Check Calcd Residual Density  0.77A From      O3       -0.54 eA-3
#=============================================================================
004_ALERT_5_G Polymeric Structure Found with Maximum Dimension           3 Info
008_ALERT_5_G No _iucr_refine_reflections_details   in the CIF     Please Do !
042_ALERT_1_G Calc. and Reported MoietyFormula Strings  Differ     Please Check
045_ALERT_1_G Calculated and Reported Z Differ by .............       2.00 Ratio
072_ALERT_2_G SHELXL First  Parameter in WGHT Unusually Large.        0.11 Report
164_ALERT_4_G Nr. of Refined C-H H-Atoms in Heavy-Atom Struct.           1 Note
199_ALERT_1_G Reported _cell_measurement_temperature ..... (K)         293 Check
200_ALERT_1_G Reported   _diffrn_ambient_temperature ..... (K)         293 Check
232_ALERT_2_G Hirshfeld Test Diff (M-X)  Mn1    --  O1       ..        5.3 su
232_ALERT_2_G Hirshfeld Test Diff (M-X)  Mn1    --  O4_a     ..        8.5 su
232_ALERT_2_G Hirshfeld Test Diff (M-X)  Mn1    --  O3_g     ..        8.5 su
232_ALERT_2_G Hirshfeld Test Diff (M-X)  Mn1    --  O2_i     ..        6.0 su
764_ALERT_4_G Overcomplete CIF Bond List Detected (Rep/Expd) .        1.35 Ratio
910_ALERT_3_G Missing # of FCF Reflection(s) Below Th(Min) ...           3 Report
912_ALERT_4_G Missing # of FCF Reflections Above STh/L=  0.600         253 Note
#-----------------------------------------------------------------------------
```


C11_6
O1_6
O4_6
C1_6
Mn1_6
C1_6
O3_6

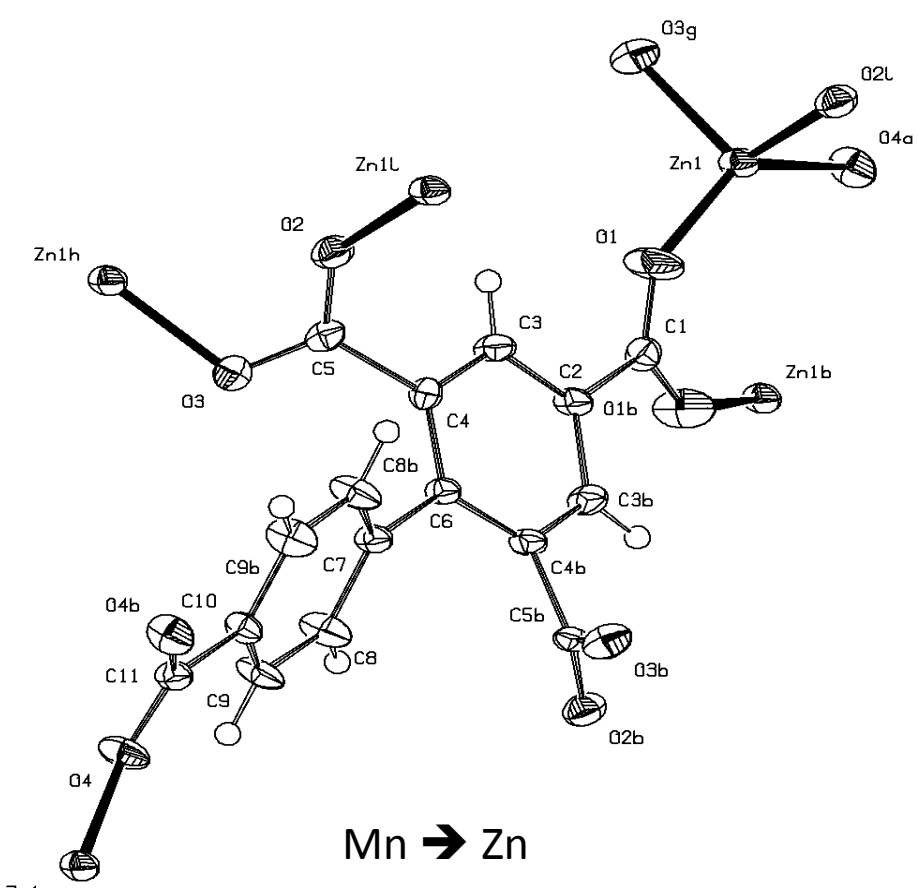
Refined with Mn substituted by Zn



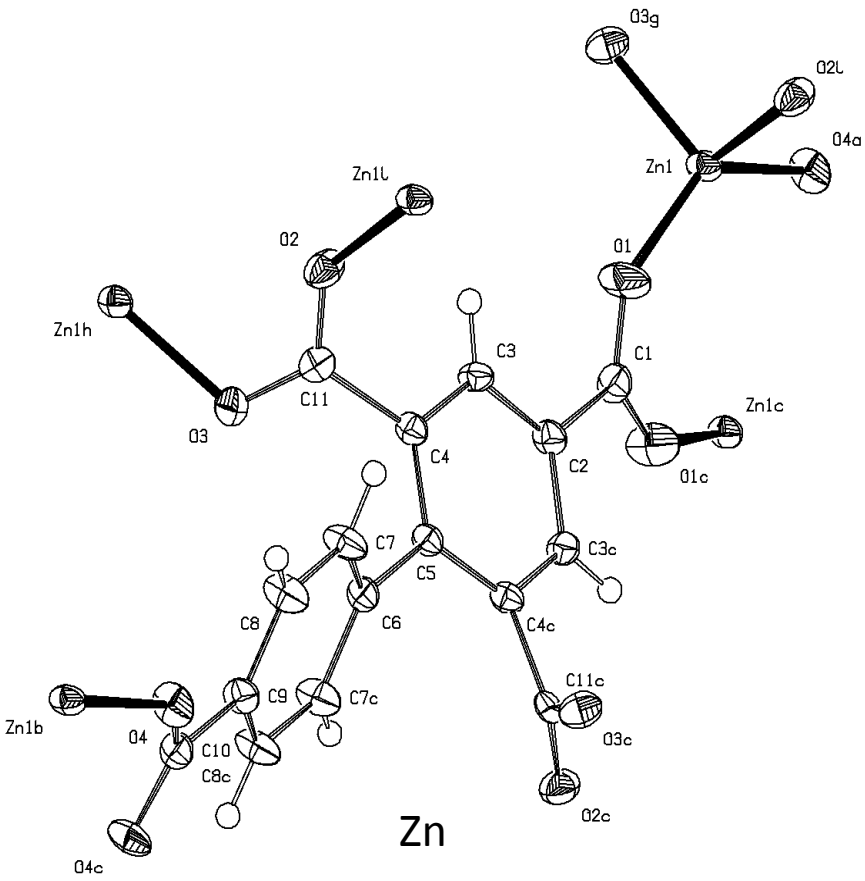
R_1 = 0.047
wR_2 = 0.113
S = 1.039

Improvement of the model is hardly visible,
But the R-values improved significantly

Structure Comparison



Mn ➔ Zn



Zn

The displacement ellipsoids are surprisingly similar (both RT)
Unfortunately, the two reflection sets can not be compared
Paper rejected !

Common checkCIF Issues

- ADDSYM is designed to warn for possible higher symmetry between the supplied coordinate set with a suggestion for a revised space group. The tolerances used are relatively high in order to catch most known mis-assignments. An ALERT not necessarily means an error, just something to investigated and reported on.
- With multiple components in the unit cell or with polymers it is not always clear what the best choise is for the formula unit. CheckCIF might make a different choise than an author on the basis of chemical insight. The main issue is that the reported Z and related quantities are consistent with the reported formula unit.
- CheckCIF uses its own f' and f'' anomalous dispersion parameter values for a synchrotron based wavelength as a check against those supplied. The actual values can be listed with the ANOM <wavelength> instruction in PLATON.
- Details on missing reflection and outlier ALERTS can be found in the FCF validation report.

Missing Reflections (Asym. Refl. Unit) below sin(th)/lambda = 0.5

Nr	H	K	L	sin(th)/lambda	Theta	I(calc)	I(calc)/I(max)
1	0	2	0	0.051	2.10*	23.09	0.00029
2	1	3	0	0.097	3.95	168.56	0.00209
3	-1	5	1	0.148	6.02	307.55	0.00381
4	6	10	1	0.443	18.37	2760.21	0.03422
5	-1	1	2	0.113	4.59	146.91	0.00182
6	0	1	2	0.103	4.19	1183.75	0.01468
7	-1	3	2	0.134	5.47	1582.16	0.01962
8	-4	14	3	0.446	18.49	30.64	0.00038
9	-5	14	4	0.492	20.47	2085.30	0.02585
10	-3	1	7	0.373	15.35	894.30	0.01109
11	-2	3	7	0.363	14.94	93.44	0.00116

** Note: I(max) is the maximum I(obs) encountered in the fcf-file **

Starred Reflections have a Theta below Theta(Min) = 2.28

From CIF: Theta(Min) = 2.28

Section 4:

Resolution & Completeness Statistics (Cumulative and Friedel Pairs Averaged)

Theta	sin(th)/Lambda	Complete	Expected	Measured	Missing	
20.82	0.500	0.994	1748	1737	11	
23.01	0.550	0.994	2340	2327	13	
25.24	0.600	0.994	3035	3016	19	
						ACTA Min. Res. ---
26.01	0.617	0.994	3300	3280	20	

```
PLATON-(Version  270714)-Mode=2 FCF-File Validation for:I
================================================================================
For Documentation: http://www.platonsoft.nl/FCF-VALIDATION.pdf

Section 1
================================================================================
General Data
================================================================================
Crystal Data From: zl2039.cif
Fo/Fc   Data From: zl2039.fcf FCF-TYPE=LIST4
Space Group      : P21/n
Wavelength (Ang) :   0.71073
Unit Cell (CIF)  :    8.5814  19.4260  10.1180   90.000   96.620   90.000
SHELX WGHT Pars. :    0.0336   0.0000

Section 2
================================================================================
Reflections with abs((I(obs) - I(calc)) / SigW(I) .GT.  3.0
================================================================================
   Nr    H    K    L  Theta     I(obs)    I(calc) Sigma(I)    Ratio  SigW(I) RatioW
--------------------------------------------------------------------------------
    1    8    6    0  20.56      89.20     155.04    20.77    -3.17    21.25  -3.10
    2    4    7    0  12.14     102.88      64.92    11.25     3.37    11.55   3.29
    3   -1    0    3   6.28     782.69     674.72    26.62     4.06    35.76   3.02
    4    0    4    3   7.40     485.61     590.60    24.33    -4.32    30.67  -3.42
    5    8    6    4  23.12     -11.16      29.85    13.62    -3.01    13.64  -3.01
    6    4   13    5  20.48     528.19     400.54    39.70     3.22    42.40   3.01
    7    1   15    7  22.03      94.71     177.93    26.92    -3.09    27.39  -3.04
    8    6    2    9  25.37     723.80    1021.52    91.48    -3.25    96.59  -3.08
                                                             --------         --------
                                                  Average =    -0.77            -0.79
```

Concluding Remarks

- CheckCIF is intended to be helpful and to protect against oversights rather than being annoying
- Many ALERTS have their origin in previous ‘incidents’
- Some ALERTS can point to interesting features in a structure and relevant to be discussed
- Some ALERTS may be less important in a given context
- Sometimes, a combination of minor ALERTS can point to a problem that needs to be resolved
- The set of ALERTS is still fine tuned and extended on the bases of reported new ‘incidents’.
- Please send suggestions and examples of annoying issues to a.l.spek@uu.nl

Thanks !

a.l.spek@uu.nl
www.platonsoft.nl