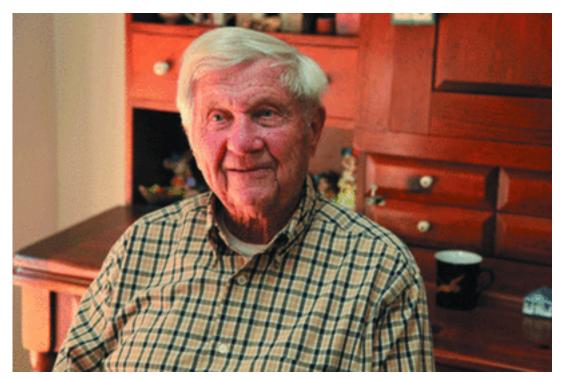
Fourty Years of Marshing: Is the Missed Symmetry Problem now Solved?



Richard E. Marsh (1922-2017) Ton Spek, Utrecht University, NL. 22-7-2018, ACA-Toronto

My Connection with Dick

- My early interest in the 'Missed Symmetry Problem' was aroused by Dick's 1979 paper: Richard E. Marsh & Verner Schomaker, Inorg. Chem. (1979) 18, 2331-2336, Some Incorrect Space groups in Inorg. Chem, Volume 16 (1977)
- That paper inspired me, as a service crystallographer, and others to work on software to be routinely aware of this space group problem.
- On the basis of that software I had a nice one-time joint 'internet based' collaborative paper in Acta Cryst. with corrections to the literature.

The Missed Symmetry Problem

- Correct space group assignment to a crystal structure is not always trivial (in particular concerning the centro/non-centro choice case).
- As a result, many (and still uncorrected) errors can be found in the literature and databases.
- Some errors are harmful in having their impact on the presented chemistry, others can be of a more esthetic nature.
- Pseudo-symmetry, disorder, poor data and twinning may confuse the issue.

Space Group Corrections

Space group corrections include:

- Change of crystal system
- Change of Laue group
- Non-Centro to centrosymmetric
- Missed translation symmetry

The last two cases are the more problematic ones since they may lead to structure artefacts in addition to less optimal refinement results.

The Problem

- The problem is that the correct space group assignment is often only possible after the structure is preliminary 'solved' in a lower symmetry group, e,g. P1.
- Authors do often not know how to manage a transformation to higher symmetry space groups. E.g. the needed origin shift (P-1) or origin choice (C2/c) problems. Education & Experience needed!
- In addition, for proper evaluation of a proposed space group assignment, we will need the primary reflection data on which the study is based.
- Those data are not always archived for published papers, making 'correction reports' based purely on coordinates tricky.

More Correction Papers since 1979

- More erroneous structure reports, due to inexperience, started to appear in the literature, in particular in non-crystallographic journals (Inorg Chem, JACS). Often no crystalographic referees involved.
- The 1979 'missed symmetry' paper was the first in a series of similar papers by Dick, often in collaboration with other 'senior' scientists. Other authors soon picked up the same subject as well.
- The procedure that was used by Dick was to look by hand for special relations in the coordinates of similar 'independent' atoms. Published ORTEPS are helpfull to spot problems.

How Large is the Problem ?

- An under-bound can be gleaned from the searchable information available in the CSD
- Searching the CSD for the tekst 'REINT_OF', that stands for 'reinterpretation', reports 1800 entries (89 between 2013-2017) (2018 version)
- 1061 of those indicated as re-interpretations are (co)authored by Dick Marsh
- However, there are many more not marked yet (and corrected) in the CSD (e.g. those reported in Dick's last Acta Cryst. (2014). C70, 834 paper).

The 156 updated structures, including the CSD refcode, the reported space group, and the revised space group.

Refcode	Reported space group	Revised space group	Refcode	Reported space group	Revised space group	Refcode	Reported space group	Revised space grou	ıp
APONOJ	<i>P</i> 1	P1 P1 P1 P1 P1	IGEKIP	P1 P1	<i>P</i> 1	PUBTEM	Pī	C2/c	_
APOPAX	P1	P1	IMOMUT	P1	$P2_1/n$	PUBTOW	P21	$\frac{P2_1}{P1}$	
APOPOL	P1	P <u>1</u>	IQATEA	P1	$\frac{P2_1}{P1}$	PUBTUC	P1 P1	P1	Last 2014
APORIH	P1		ISECOZ	P1		QAMLEX		$P2_1/n$	Last 2014
AQINIY	$P\overline{1}$	C2/c	ITEXUB01	Cc	C2/c	QANCEP	P1	$P2_1/n$ $P\overline{1}$ $P\overline{1}$	
ATANEP	P21	P212121	ITIREJ	P1 P1	P21212	QAPWUB	P1	P1	Acta Cryst.
ATIWIK	Cc	$P2_1/c$	ITUMIU	P1	$P2_1/n$	QASLUT	P43	P4322	
AVATIB	Cc P1	C2/c	IVEMUS	P1 P1	$P2_{2}/n$	QUDNAF	Cc	C2/c P1 P1	C70, 834-836
AWEMEV		C2/c	IXUGUE	P1	C2/c	QUHNUD	P1	P1	$C/0, 034^{-}030$
AYEJUK	Cc	C2/c	IZIGII	$P2_12_12_1$	Pbca	QUJHOT	P1	P1	and a second state
AZEKEW	$P2_1/c$	Pecn	KAZRIO	P1	$P\overline{1}$	QUMKEP	Cc	C2/c	paper with
BAHHUP	P1	$P\overline{1}$	KEFNIU	P21	C2221	RUGDON	$P\overline{1}$	$P\overline{1}$	• •
BAZYOS	Pn P1	$P2_1/n$	KETDEU	Pc P1	$P2_1/c$	RUHGAD	Pn	$P2_1/n$	156 new
BOXNOS		C2/c	KUBCOA		Cmca	RUHKOV	$P2_1$	Pbca	13011010
BOYFIF	P1	$P\overline{1}$	KUFLED	P1	$P\overline{1}$	SAQHOJ	P1	$P\overline{1}$	compostions.
BOYZIZ	Cc P1	C2/c	KUGPOS	P2/c	Pbcm	SAWKUY	Cc P1	C2/c	corrections
BULCIV	$P\overline{1}$	P2/c	KUJLUX	14	14/m	SEFQEB	$P\overline{1}$	$P2_1/c$	
CACMEA	P1	$P\overline{1}$	LAYNOQ	$P2_1$	P212121	SELGIB	$P\overline{1}$	$P2_1/c$	In recent
CAFWIR	$P\overline{1}$	C2/c	LEPMEZ03	$\frac{P2_1}{P1}$	$P2_1/n$	SUJWIE	$P\overline{1}$	$P2_1/n$	millecent
CAJCAT	P43	P4322 P1	LUKDEB	$P\overline{1}$	$P2_1/n$	TASTEO	P31	P3121	Development
CASXEB	P1		LULSUH	P21 P1	P212121	TAWYOH	$P2_1/c$	Pbca	Papers.
CEBYIT	Cc	C2/c P1	MAMXEF	$P\overline{1}$	P2./c	TECLEU	$P\overline{1}$	$P2_1/c$	•
CEBZAM	P1 P1	$P\overline{1}$	MAMXOP	P1	$\frac{P2_{1}/c}{P1}$	TEJLUR	$P\overline{1}$	C2m P1	
CUGKAR	$P\overline{1}$	C2/c	MAMXUV	Cc	C2/c	TUFTUK	P1	$P\overline{1}$	
CUMSEJ	P1	$P\overline{1}$	MAMYAC	P21	P2,/c	TUHBEE	P21	$P2_1/c$	Nowaraing
CUNBET	P1	P21 P1	MAVDAQ	P1	$\frac{P2_{1}/c}{P1}$	UCABOR	P1	$P2_1/c$ $P\overline{1}$ $R\overline{3}c$	No warning
DABVIN	$P\overline{1}$	$P\overline{1}$	MAVZUG	P1	P21/c	UHANUN	P1 P1	R3c	
DAQWOJ	Cc	C2/c	MEBGEH	$P2_1/c$	Pbcn	UQUDUG	$P\overline{1}$	$P2_1/n$	Info added
DEDVEO	Pa	Pca21	MEDRIY	Cc	C2/c	UROPIB	$P2_1/n$	Pnnm	into added
DELFIL	P1	$P\overline{1}$	MEDXAW	C2	Aba2	UROPUN	Pc	$P2_1/c$	in the 2018
DUKYIS	Cc	C2/c	MOVWIE	C2 P1	$P\overline{1}$	UWAVOE	$P2_1/n$	Pnma	111 the 2010
EMOCEP	$P2_1ab$	Pbca	MOXDUZ	PI	C2/c	UWIKUH	C2	Fdd2	
ENEKUE	PI	$P2_1/n$	MUKGEF	P1 P1	C2/c	UWUFEY	$P\overline{1}$	C2/c	CSD
EQEDAG	P1	PI	NARMAW	$P\overline{1}$	P2/c	VASBUO	$P\overline{1}$	$P2_1/n$	•••
ERAZAZ	$P2_1/m$	Pnma	NUCSUA	P1	P2/c P1	VATHIJ	Pn	P2/c	
EZIQEK	P43	P43212	NUCTAH	P1	$P\overline{1}$	VAVHEH	P1	Pna21	
EZOZEZ	Pna21	Pccn	OHAPOD	P1	Pī	VUHBEG	Cc	C2/c	A 4 4 1 1 1 1 1 1
FAMYUP	PI	$P\overline{1}$	OMOCAV	P1	P21212	WEHPUW	P1	C2/c P1	Will be handled
FAZGIY	P3	P321	OMOCUP	$P\overline{1}$	P2./n	WUGKAL	P1	C2	
FEPGOY	P1	Pī	ONILIH	P1	P1 P1	WULCEM	Pa		In the 2019
FOXTIW	P21	C2221	OQEVOW	P1	PI	WUNLUN	P1	$\frac{P2_1}{P1}$	III the 2019
FUDMIB	P21	P21212	OTAPOP	PI	P2,/c	WUQQEF	P21	P212121	Delesse
FUFCEP	P21 P1	$P2_1/c$	OWEDAW	PI	P2:/a	WUSVUC	141	14,22	Release.
FUNHIG	R3	R32	OXEJIL	P1 P1 P1 P1 P1	P2./n	WUSWAJ	14	14,22	
GEGVAR	R3 P1	$P2_1/n$	OXERAL	P21	Pmn21	XEHYUG	PI	P21/c	
GENHUE	PT	$P2_1/c$	PAKMUL	Cc	C2/c	XESPOC	P21	$P2_1/n$	
GUHTAF	P21	P2.	PAZOXN12		P2:/n	XOTYOV	Pc	P21/c	
HATCAI	P1	P21 P1	PECPIY	Pn P1	P2./c	XOTYUB	P21	P21/n	
HOGDUD	Cc	C2/c	POPPAM	Ia	12/a	YOSFES	141	14.22	
HORTEO01	P1	PI	POSWOK	P1	PI	YUHCAG	P1	14122 P1	
HOXLEM	P1		POTBAC	P1	P2:	YUNWIO	Cc	C2/c	
IBIXEY	P1	12 P1	POYXIL	PI	P2./c	ZAQKAF	Cc	C2/c	

Why Bother ?

- Space group re-interpretation papers have always been controversial (and sometimes considered pedantic ..) Dick's early rebuttal on this is already given in the Acta Cryst. (1979) B35, 1094-1099 paper:
- "Some will argue that the mishap was harmless; afterall, the reported structure is essentially correct, as far as we know, except for the symmetry and the divergent lengths and angles. We can understand, but hardly agree."

New Computer Based Algorithms

- 1982: CREDUC, Y Le Page J.Appl.Cryst.,**15**,255.
- Determination of the Metrical Symmetry of the lattice from the number of compatible two-fold axes.
 E.g. 1 → monoclinic, 5 → tetragonal, 9 → cubic
- Two-axis = coinciding low-index (0,1,2) direct and reciprocal lattice vectors within a tolerance.
- Actual symmetry can be lower, not higher than the metrical symmetry.
- The PLATON/LEPAGE tool is Modelled on CREDUC
- Does the cell content support the lattice symmetry ?

The MISSYM Algorithm

- In 1987, Y. Le Page published an extention to CREDUC called MISSYM (J.Appl.Cryst.,**20**,264).
- MISSYM Looks for support in the actual structure (I.e. coordinate set) for the proposed metrical symmetry elements of the lattice found by the CREDUC algorithm.
- Liberal tolerances are used in general to catch outliers due to refinement divergence and pseudo-symmetry cases.

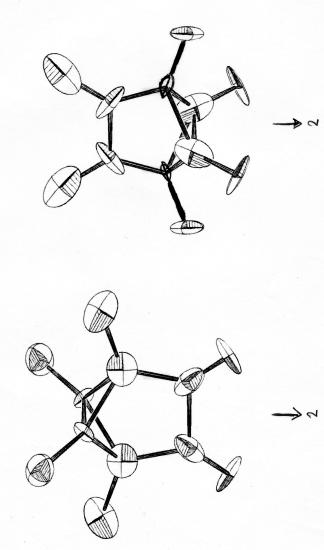
MISSYM Implementation

- Originally as part of the NRCVAX package.
- That version reports additional symmetry elements but leaves their evaluation and implementation into the proper space group to the educated researcher.
- In PLATON as the ADDSYM routine along with a work-up to the suggested new space group.
- ADDSYM is also part of the IUCr/checkCIF tool

PLATON/ADDSYM Extensions

- Suggests automatically a revised space group, unit cell and origin transformation.
- May Generate a new '.res' file for subsequent SHELXL refinement.
- Allows for a percentage of non-fitting atoms in order to catch interesting pseudo-symmetry cases.
- Handles sub-cells (originating from structure prediction)
- Accepts input data in '.res' and '.cif' format.
- Runs on files with multiple entries (CSD/CIF).
- Display of the revised (averaged) structure.

Retro Check of my own Early Publications with the ADDSYM Tool



One Hit ! in papers published in 1977, prior to the Marsh & Schomaker (1979) paper.

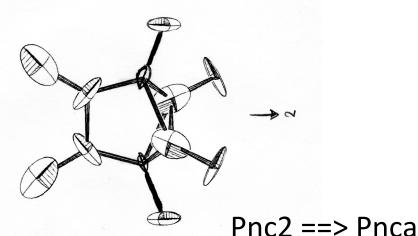
Cryst. Struct. Commun. (1977) 6, 259-262 (Journal was the forerunner of Acta Cryst. C)

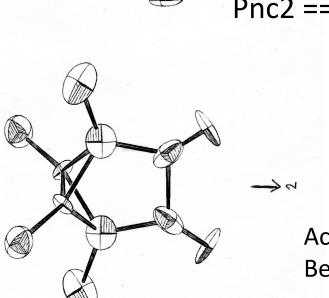
Refined with XRAY76 in the Non-Centro Space Group Pnc2. Poor structure: $R_F = 0.08$

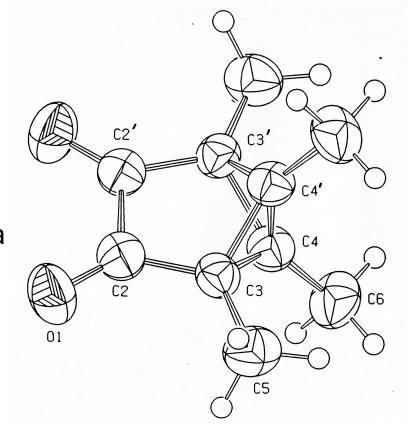
No inversion implemented in view of hk0, h=2n+1 outliers up to 30 sigma(I)

Re-Refined with SHELX76 in Space Group Pnca. $R_F = 0.049$, $wR_F = 0.077$

Refinement Result with the Proper Space Group Assignment

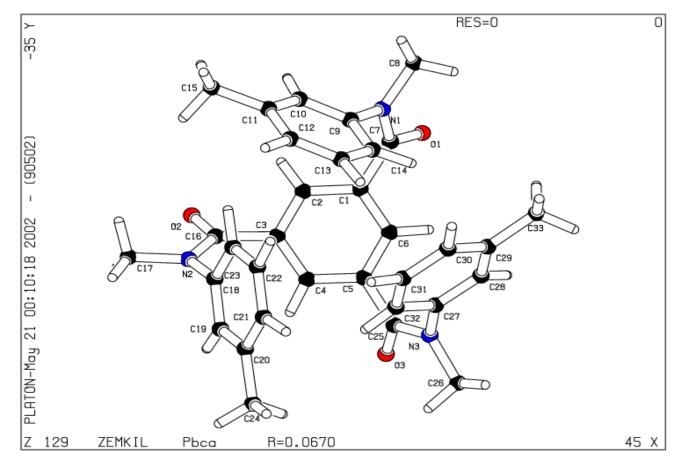






Acta Cryst. (1990). C46, 1357-1358 Better geometry and Displacement Parameters

Change of Crystal System Example proposed by ADDSYM



Asymmetric Unit in Pbca



PLATON/ADDSYM for ZEMKIL Pbca

ADDSYM Search on ALL NON-H Chemical Types [Max NonFit 20 Perc] The Structure Implies the Following Symmetry Elements Subject to the Criteria: 1.00 Deg (Metric), 0.25 Ang (Rotations), 0.45 Ang (Inv), 0.45 Ang (Transl)

							3					nobisorder
Symm. Elem	. Input Cell Row	Reduced Cell Row	(Ang) d Typ	Dat	(Deg) Angle	() Ftt 1	(Ang) MaxDev.		I ×	nput C .y	ell z	Organic Round
Ь	[1 0 0]	[1 0 0]			0.00	100	0	through Glide =	1/4 0	0 1/2	0	Parentheses
С	[0 1 0]	[0-1 0]	17.86 4	1	0.00	100	۵	through	ů o	1/4	Ō	Label-Alias
a	[0 0 1]	[001]	17.86 4	1	0.00	100	0	Gllde = through	Ο	0 0	1/2 1/4	R/S-Determ
3 *	Г <u>1_1_</u> 17	[-1-1-1]	30,93 3	3	0.00	100	0,009	Gllde =	1/2 1/2		0 1/2	NoSubCell
					U .UU	C15	-033	through	172	U	172	Norm-H-bond
3 *	[1 1 1]	[1-1-1]	30.93 3	3	0.00	100	0.009	through	2/3	1/6	1/6	Join-Expand
3 *	[1-1 1]	[1 1-1]	30,93 3	3	0.00	C33 100	-C15 0,009	through	1/3	2/3	1/3	LstARU RCel
0	F 4 4 47				0.00	C15	-033					LstCellSymm
J *		[-1 1-1]	30.93 3	3	0.00	100 C33	0.009 -C15	through	1/6	1/6	1/3	ListAtoms
-1	=======	=========	========	====	=====	100	0	at	0	0	0	ListBonds
		^					т т		(-		Ŧ	LstFlagRadi
Fie (duced-to-		Inpu	it-to	-Reduc			put-to-Co			• T a	X-LineWidth
((1 O 0 1		(1 X (C		0 -1	() () () () () () () () () () () () () ((()	1 O O -1	() ()	Ue	et(T) =	Reverse-B&W
ì	õ õ	1)	î		Ō.	-1)	Ì	ō ō	-1 j	1.	000	Q-Peak-Incl
Cell	Lattice	a	b c	;	Alpha	Beta	Gamma	Volume C	rystalS	iystem	Laue	EPS HGL TEK
Input					90.00	90.00	90.00		rthorho	mblc	mmm	NoSymm
Reduc	ced P 17 ent cP 17				90.00 90.00	90.00 90.00			ſ	ublc	m-3	Prev Next
	ent cr 17 In shifte						ansform				m-J	SAVE-InstrS
		a ca: u.u. onal Symm						atton (Na 20!	51			ENTRY-LIST
10860		onde ognini		3300								Reset End

INSTRUCTION INPUT via KEYBOARD or LEFT-MOUSE-CLICKS (HELP with RIGHT CLICKS)

Exit

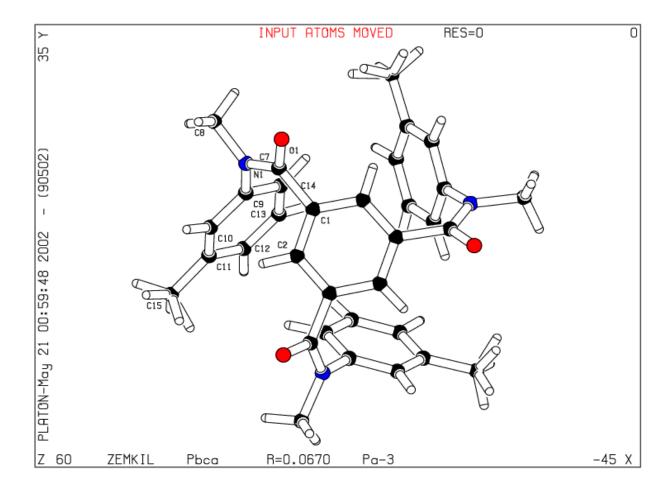


OptionMenus

NoMove

tra

Orthorhombic Pbca 🗲 Cubic Pa-3



Molecule on three fold axis

Early Survey's of the Missed Symmetry Issue

- A survey by Dick Marsh in 1997 of all structures published in space group Cc showed that about 10% of the assignments were wrong.
- A new survey by Dick Marsh in 2004 surprisingly showed that this percentage was still around 10%.
- Note: none of the additional 164 corrected assignments were from Acta Cryst. journals.
- Common revisions were to C2/c, Fdd2, R-3c etc.

Centro or Non-centrosymmetric ?

- Correct assignment is not always trivial.
- Refinement of (close to) centrosymmetric structures in P1 is mathematically unstable.
- Mis-assignment -> geometry artifacts.
- **Dick Marsh**: Borderline cases are best treated as centro-symmetric (static/dynamic disorder)
- Inspect Fo/Fc of weak reflection data.
- In general: chiral molecules-> non-centro SpGr.

Interesting Story

- ADDSYM can automatically analyse a CIF file containing multiple entries exported from the CSD.
- In 1999, in response to the paper "P1 or P-1 or something else" I send Dick my larger automatically created "P1" list. He immediately started to work on that list and indeed found some more missed symmetry cases.
- Interestingly, one of the entries (NAXQAM) that he had previously updated to Cc was 'Marshed' again to Fdd2. I did not realize that when I supplied my list.
- I was told that he enjoyed being Marshed.

Joint Paper

- As a follow-up, Dick asked me to send him my automatically generated list for space group corrections to structures published in C2.
- This resulted in a one time long distance internet based collaboration and the paper:
- "Use of software to seach for higher symmetry: space group C2". Acta Cryst. (2001). B57, 800-805.
- It was an honor and an enjoyable experience

ADDSYM runs for all 2013-2018 CSD Entries: Is the problem solved ?

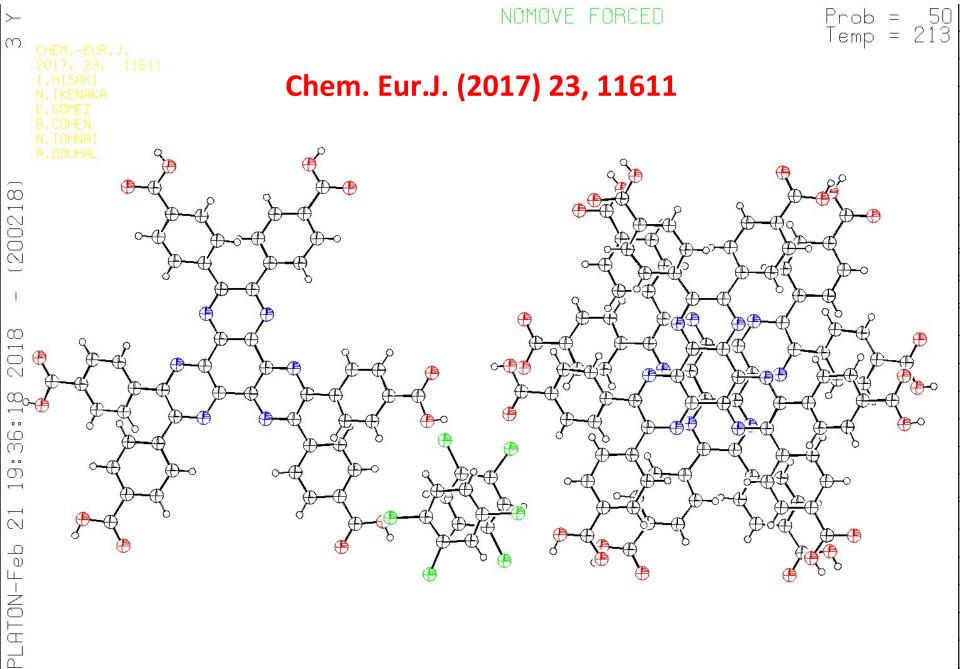
Year #New ENTRIES ADDSYM HITS 100% FIT 2013 47233 797 (1.7%) 364 (0.77%) 890 (1.8) 416 (0.84) 2014 49546 2015 390 (0.74) 52652 860 (1.6) 2016 343 (0.63) 832 (1.5) 54333 253 (0.62) 618 (1.5) 2017 40496 93 (0.56) 258 (1.6) 2018 16546

'HITS' = Number of entries satisfying the default ADDSYM criteria for closer inspection (Not necessarily an Error)

ADDSYM Criteria

- Default criteria are set to values to also catch large outliers (in cell dimensions or near singular refinement parameters), pseudosymmetry, incorrectly assigned atom types etc.
- All reported cases of additional symmetry should be investigated, acted upon, reported and discussed in a structure report.
- Next: subset of hits for 2017 as illustration

LEJCUB	Ρ	n	mP = >mP	0.0	0	0	0.000	0.00	0.204		100%	P21/n
LEJDAI	Ρ	n	mP = >mP	0.0	0	0	0.000	0.00	0.236		100%	P21/n
MANNUN	Ρ	21	mP = >mP	0.0	0	0	0.000	0.00	0.000		82%	P21/c
MANQEA	Ρ	21/n	mP=>mP	0.0	0	0	0.000	0.00	0.239	S	100%	P21/m
MAPFER	Ι	4	tI=>tI	0.0	0	0	0.000	0.00	0.138		100%	I422
MAPXEJ	Ρ	ba2	oP=>oP	0.0	0	0	0.000	0.00	0.000		84%	Pcca
MAQHEU \star	Ρ	-1	aP=>aP	0.0	0	0	0.000	0.00	0.000	S	100%	P-1
MASWIP	Ρ	21	mP = >mP	0.0	0	0	0.000	0.00	0.248		87%	P21/c
MAVYAM	Ρ	21	mP = >mP	0.0	0	0	0.000	0.00	0.000		87%	P21/c
MAVZER	Ρ	65	hP=>hP	0.0	0	0	0.000	0.00	0.174		100%	P6522
MAWLII	Ρ	21	mP=>oP	0.0	0	0	0.000	0.15	0.032		100%	Pmc21
MAXVEP	Ρ	-1	aP=>mP	0.0	0	0	0.000	0.11	0.007		100%	P21/c
MAXWUG	Ρ	1	aP=>aP	0.0	0	0	0.000	0.00	0.000		89%	P-1
MAXXER	Ρ	21/c	mP=>oP	0.0	0	0	0.000	0.09	0.012		100%	Pbca
MAZDAV	Ρ	21/n	mP=>oP	0.0	0	0	0.000	0.36	0.241		95%	Pnma
MAZSOY	Ι	2	mI = mI	0.0	0	0	0.000	0.00	0.000		93%	I2/c
MAZTAL	Ρ	1	aP=>aP	0.0	0	0	0.000	0.00	0.000		97%	P-1
MAZTOZ	Ρ	1	aP=>aP	0.0	0	0	0.000	0.00	0.000		92%	P-1
MEBBIH	Ρ	21	mP=>mP	0.0	0	0	0.000	0.00	0.000		96%	P21/n
мевкем *	Ρ	-3	hP=>hP	0.0	0	0	0.000	0.00	0.006		100%	P-3c1
MEBKIQ	R	-3	hR=>hR	0.0	0	0	0.000	0.00	0.010		100%	R-3c
MEBRET	Ρ	21/c	mP=>oP	0.0	0	0	0.000	0.86	0.248		95%	Pnma
MEDKIS	Ρ	с	mP=>mP	0.0	0	0	0.000	0.00	0.249		99%	P21/c
MEDNER	Ρ	с	mP=>mP	0.0	0	0	0.000	0.00	0.249		99%	P21/c
MEDXAX	Ρ	-3	hP=>hP	0.0	0	0	0.000	0.00	0.217		100%	P-3m1
MEGMAP	Ρ	-4	tP=>tP	0.0	0	0	0.000	0.00	0.231	į.	98%	P-42m
MEHQAU	С	2	mC=>mC	0.0	0	0	0.000	0.00	0.042		100%	C2/c
MEHQEY01	С	2	mC=>mC	0.0	0	0	0.000	0.00	0.051		100%	C2/c
MEJZIL03	Ρ	-1	aP=>mP	0.0	0	0	0.000	0.15	0.240	S	100%	P21/c
MEJZIL04	Ρ	-1	aP=>mP	0.0	0	0	0.000	0.05	0.244	S	100%	P21/c
MEJZIL05	Ρ	-1	aP=>mP	0.0	0	0	0.000	0.15	0.247	S	100%	P21/c

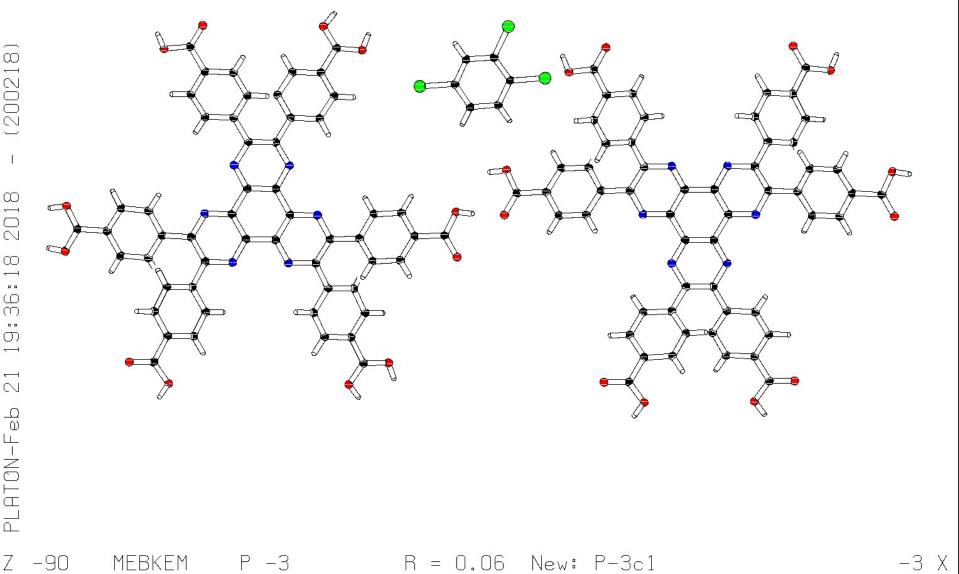


Z - 89 MEBKEM P - 3 R = 0.06

RES= 0 10 X

INPUT ATOMS MOVED RES=0 0

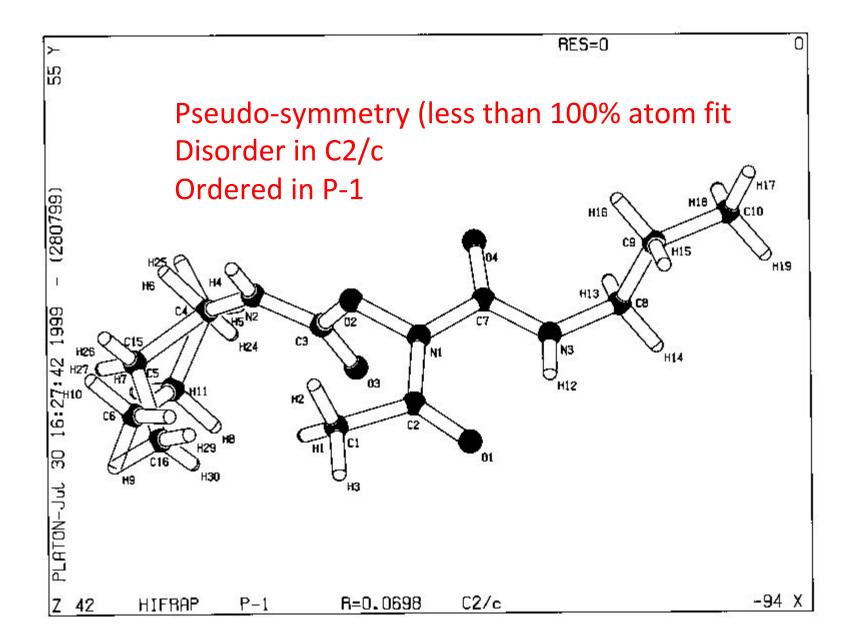
Change of LAUE class: Low to high trigonal $-3 \rightarrow -3m1$, P-3 \rightarrow P-3c1



 \succ

The Problem of too High Symmetry

- In particular with point detector data, it was easy to miss weak diffraction spots, resulting in artificial disorder (i.e. averaged structures)
- Pseudo-symmetry may lead to structure solutions in too high symmetry groups and artificial disorder.
- Twinning may lead to similar issues
- A pseudo-symmetry case next:



Comments

- Hydrogen atoms and disordered atoms are not included in the ADDSYM analysis.
- Pseudo-symmetry can be both problematic and/or interesting (Inorganic Structures)
- Missed symmetry claims should always be supported with the primary reflection data.
- In the past, those data were deposited (and sometimes retyped in support of Dick's early symmetry corrections).
- Currently embedded unmerged reflection data in SHELXL cif's should not be deleted

40 Years of Marshing

• Obvious Question:

Is the Missed Symmetry Issue now Solved after 40 years ?

- Answer:
 - -Yes: For Publications in the IUCr Journals.
 - No: For Publications in Chemical Journals.

It is amazing that, given the availability of tools like IUCr/checkCIF, the issue is still often ignored by authors, editors and referees.

Thanks Dick for Keeping our Science to Standards and You for Listening !

- http://www.cryst.chem.uu.nl/spek
- <u>http://www.platonsoft.nl/spek</u>
- Email: a.l.spek@uu.nl