Improved Reporting of Crystal Structures: the Impact of Publishing Policy on Data Quality

Brian McMahon[1], Peter R. Strickland[1] and John R. Helliwell[2]

[1]International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, UK; [2]School of Chemistry, University of Manchester, Oxford Road, ManchesterM13 9PL, UK and CCLRC Daresbury Laboratory, Warrington WA4 4AD, UK

bm@iucr.org

Crystallography Journals Online

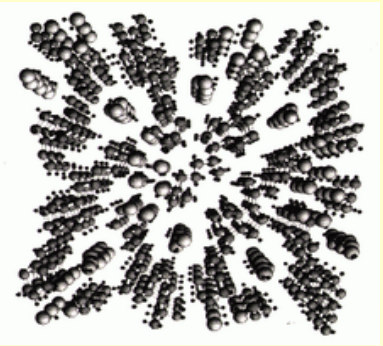
Structure of presentation

- Publication of crystal structure reports
- Data exchange/archive standards
- Publication workflow for small-unit-cell structures
- Community consensus for biological macromolecules
- Data publication at source

Crystallography Journals Online

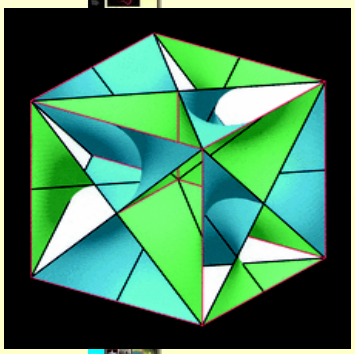
Publication of crystal structure reports

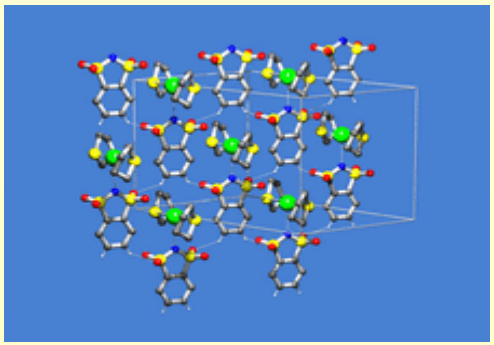
Crystallography Journals Online

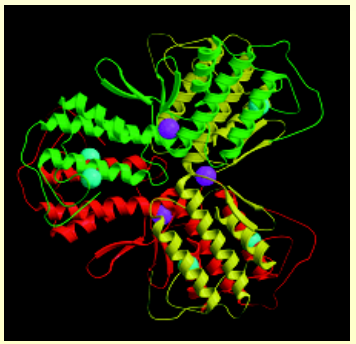


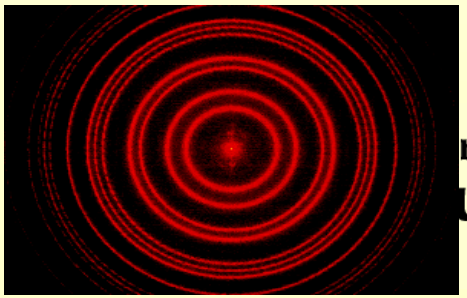
Crystallography

- The branch of science devoted to the study of molecular and crystalline structure
- Far-reaching applications in chemistry, physics, mathematics, biology and materials science







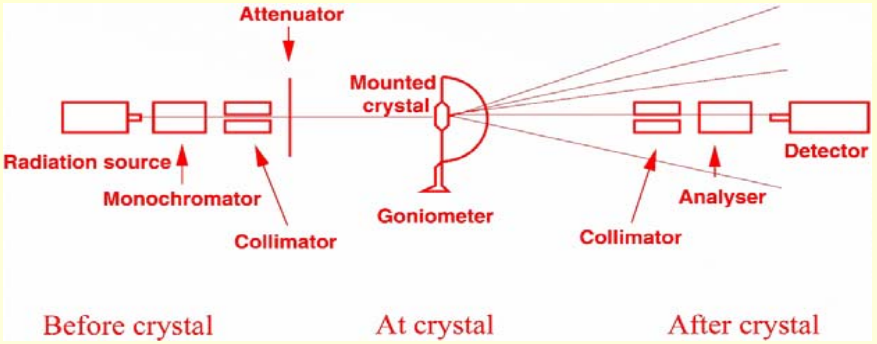


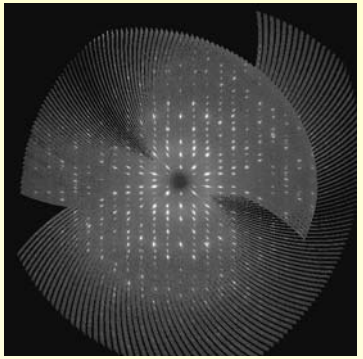
Crystal structures ‘published’

- Curated databases
  - Cambridge Structural Database
    - Small organic/metal-organic: 335,280 : 29,000/yr
  - Protein Data Bank
    - Biological macromolecules: 34,506 : 5,500/yr
  - Inorganic Crystal Structure Database (82,676), CrystMet (99,893), Powder Diffraction File (240,050)
- IUCr journals
  - *Acta Crystallographica Sections C, E*
    - Small-molecule, inorganic: 2357 articles/year
  - *Acta Crystallographica Sections D, F*
    - Biological macromolecules: ~ 120+ structural articles/year

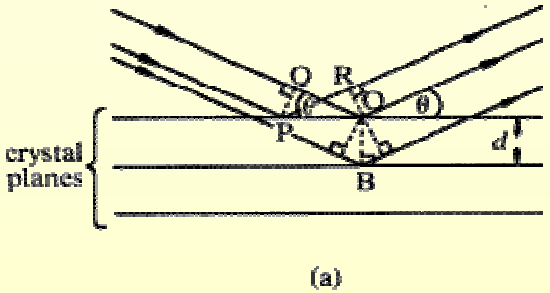
Crystallography Journals Online

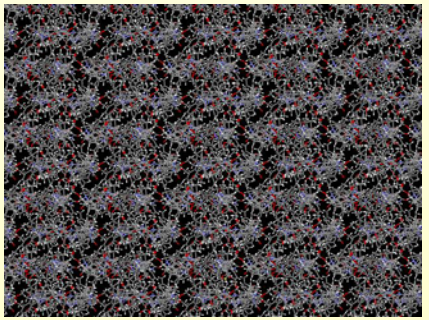
The crystallographic experiment





- Bench diffractometer, synchrotron, area detector, photographic film, space shuttle…





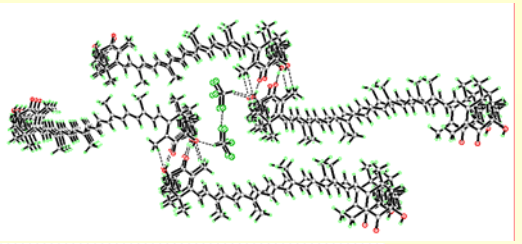
- Bragg's law: $n\lambda = 2d \sin\theta$

Crystallography Journals Online

Consistent data pipeline

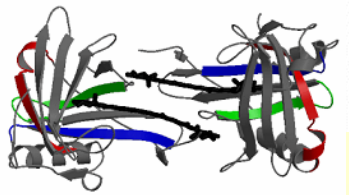
Characteristics of apparatus

Characteristics of sample and specimen



Data reduction techniques

Solution and refinement strategies

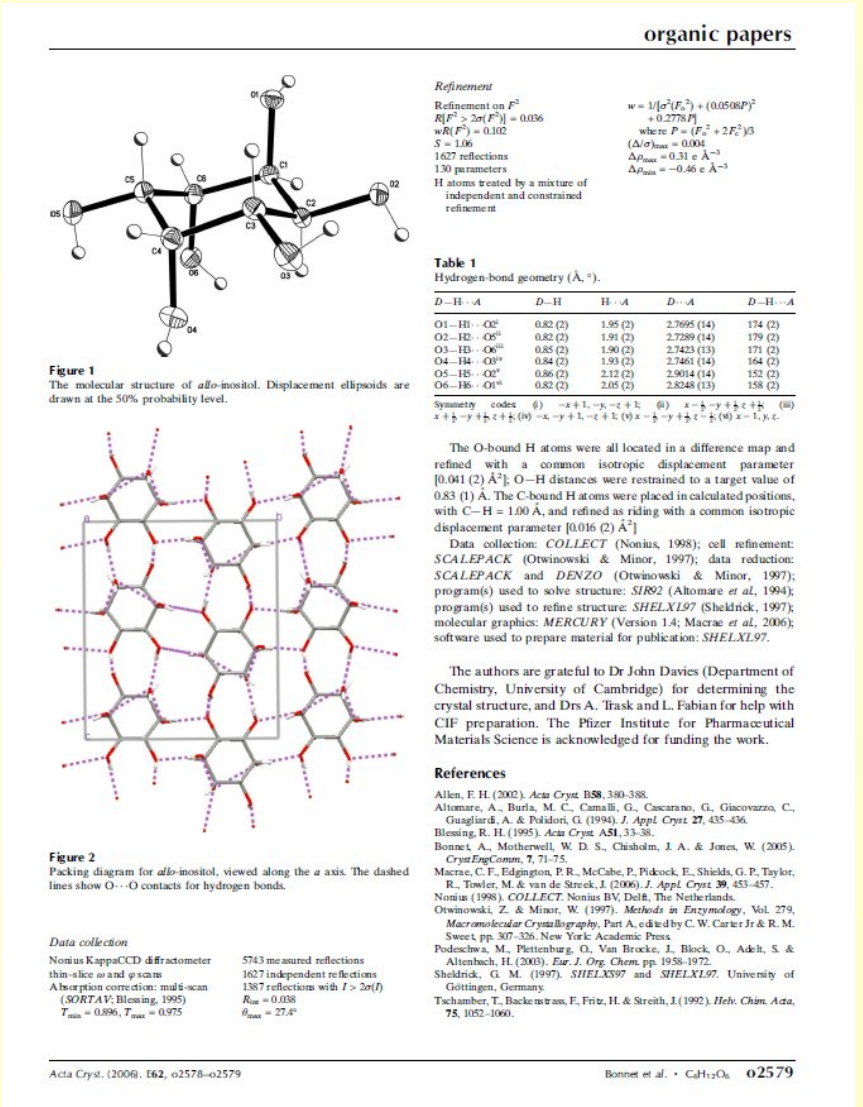


Crystallography Journals Online

Crystal Structure reports - data-rich scientific articles

- 3-d positional coordinates
- Atomic motions
- Molecular geometry
- Chemical bonding
- Crystal packing
- Chemical behaviour arising from structure

- Two dedicated IUCr journals: *Acta Cryst. C, E*
- Important part of scientific discussion in many other titles: *Acta Cryst. B, D, F*



organic papers

Figure 1
The molecular structure of *allo*-inositol. Displacement ellipsoids are drawn at the 50% probability level.

Figure 2
Packing diagram for *allo*-inositol, viewed along the *a* axis. The dashed lines show O···O contacts for hydrogen bonds.

Data collection

Nonius KappaCCD diffractometer
thin-slice ω and φ scans
Absorption correction: multi-scan (*SORTAV*; Blessing, 1995)
$T_{min} = 0.896$, $T_{max} = 0.975$
5743 measured reflections
1627 independent reflections
1387 reflections with $I > 2\sigma(I)$
$R_{int} = 0.038$
$\theta_{max} = 27.4°$

Refinement

Refinement on F^2
$R[F^2 > 2\sigma(F^2)] = 0.036$
$wR(F^2) = 0.102$
$S = 1.06$
1627 reflections
130 parameters
H atoms treated by a mixture of independent and constrained refinement
$w = 1/[\sigma^2(F_o^2) + (0.0508P)^2 + 0.2778P]$ where $P = (F_o^2 + 2F_c^2)/3$
$(\Delta/\sigma)_{max} = 0.004$
$\Delta\rho_{max} = 0.31$ e Å^{-3}
$\Delta\rho_{min} = -0.46$ e Å^{-3}

Table 1
Hydrogen-bond geometry (Å, °).

D—H···A	D—H	H···A	D···A	D—H···A
O1—H1···O2^{i}	0.82 (2)	1.95 (2)	2.7695 (14)	174 (2)
O2—H2···O6^{ii}	0.82 (2)	1.91 (2)	2.7289 (14)	179 (2)
O3—H3···O6^{iii}	0.85 (2)	1.90 (2)	2.7423 (13)	171 (2)
O4—H4···O3^{iv}	0.84 (2)	1.93 (2)	2.7461 (14)	164 (2)
O5—H5···O2^{v}	0.86 (2)	2.12 (2)	2.9014 (14)	152 (2)
O6—H6···O1^{vi}	0.82 (2)	2.05 (2)	2.8248 (13)	158 (2)

Symmetry codes: (i) $-x+1, -y, -z+1$; (ii) $x-\frac{1}{2}, -y+\frac{1}{2}, z+\frac{1}{2}$; (iii) $x+\frac{1}{2}, -y+\frac{1}{2}, z+\frac{1}{2}$; (iv) $-x, -y+1, -z+1$; (v) $x-\frac{1}{2}, -y+\frac{1}{2}, z-\frac{1}{2}$; (vi) $x-1, y, z$.

The O-bound H atoms were all located in a difference map and refined with a common isotropic displacement parameter [0.041 (2) Å^2]; O—H distances were restrained to a target value of 0.83 (1) Å. The C-bound H atoms were placed in calculated positions, with C—H = 1.00 Å, and refined as riding with a common isotropic displacement parameter [0.016 (2) Å^2].

Data collection: *COLLECT* (Nonius, 1998); cell refinement: *SCALEPACK* (Otwinowski & Minor, 1997); data reduction: *SCALEPACK* and *DENZO* (Otwinowski & Minor, 1997); program(s) used to solve structure: *SIR92* (Altomare *et al.*, 1994); program(s) used to refine structure: *SHELXL97* (Sheldrick, 1997); molecular graphics: *MERCURY* (Version 1.4; Macrae *et al.*, 2006); software used to prepare material for publication: *SHELXL97*.

The authors are grateful to Dr John Davies (Department of Chemistry, University of Cambridge) for determining the crystal structure, and Drs A. Trask and L. Fabian for help with CIF preparation. The Pfizer Institute for Pharmaceutical Materials Science is acknowledged for funding the work.

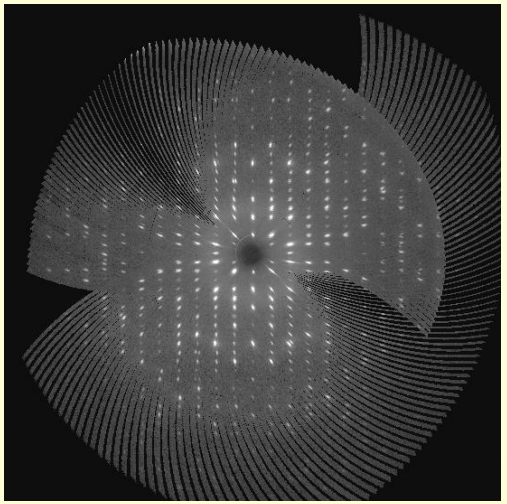
References

Allen, F. H. (2002). *Acta Cryst.* B**58**, 380–388.
Altomare, A., Burla, M. C., Camalli, G., Cascarano, G., Giacovazzo, C., Guagliardi, A. & Polidori, G. (1994). *J. Appl. Cryst.* **27**, 435–436.
Blessing, R. H. (1995). *Acta Cryst.* A**51**, 33–38.
Bonnet, A., Motherwell, W. D. S., Chisholm, J. A. & Jones, W. (2005). *CrystEngComm*, **7**, 71–75.
Macrae, C. F., Edgington, P. R., McCabe, P., Pidcock, E., Shields, G. P., Taylor, R., Towler, M. & van de Streek, J. (2006). *J. Appl. Cryst.* **39**, 453–457.
Nonius (1998). *COLLECT*. Nonius BV, Delft, The Netherlands.
Otwinowski, Z. & Minor, W. (1997). *Methods in Enzymology*, Vol. 279, *Macromolecular Crystallography*, Part A, edited by C. W. Carter Jr & R. M. Sweet, pp. 307–326. New York: Academic Press.
Podeschwa, M., Plettenburg, O., Van Brocke, J., Block, O., Adelt, S. & Altenbach, H. (2003). *Eur. J. Org. Chem.* pp. 1958–1972.
Sheldrick, G. M. (1997). *SHELXS97* and *SHELXL97*. University of Göttingen, Germany.
Tschamber, T., Backenstrass, F., Fritz, H. & Streith, J. (1992). *Helv. Chim. Acta*, **75**, 1052–1060.

Data that inform the discussion

Raw data

(image plate, diffractometer, film)



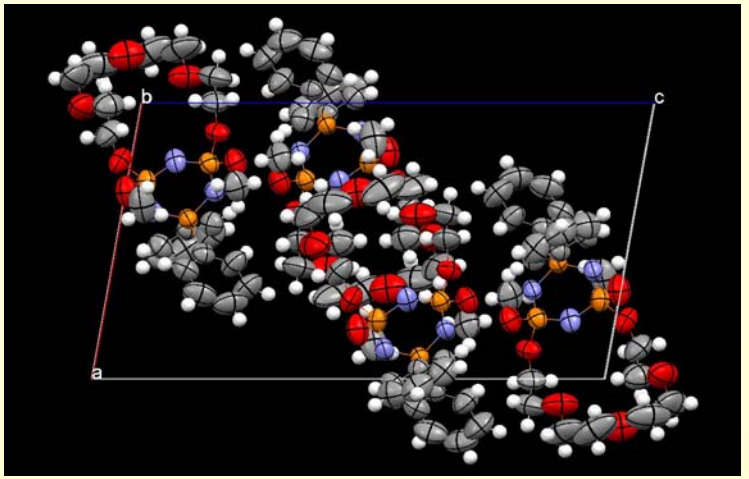
Primary data

(structure factors)

```
# h,k,l, Fc-squared, Fo-squared, sigma(Fo-squared) and status flag
#
data_6
_shelx_title ' 01SRC413 in P2(1)/n'
_shelx_refln_list_code          4
_shelx_F_calc_maximum      183.83
_exptl_crystal_F_000      1144.00
_reflns_d_resolution_high  0.7705

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

_cell_length_a    11.8293
_cell_length_b    10.3312
_cell_length_c    21.6318
_cell_angle_alpha  90.000
_cell_angle_beta  100.203
_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
   2   0   0       772.37      856.47      28.20 o
   4   0   0      1445.15     1446.80      39.55 o
   6   0   0      1130.79     1097.08      30.62 o
   8   0   0      1347.13     1490.27      55.41 o
  10   0   0      3273.01     3545.64     154.91 o
  12   0   0        48.20       40.50       4.56 o
  14   0   0        79.87       63.02       7.91 o
   2   1   0      2093.70     1975.83      47.36 o
   3   1   0     33795.10    34884.29    1287.71 o
   4   1   0      2298.16     2035.72      38.24 o
   5   1   0         9.73       36.06       5.59 o
   6   1   0       449.80      506.89      11.92 o
   7   1   0         1.81        7.91       5.59 o
   8   1   0        43.36       28.81       6.79 o
   9   1   0        64.18       48.51       6.02 o
  10   1   0      1412.22     1628.54      45.96 o
  11   1   0       242.68      279.96       9.70 o
  12   1   0        14.96       10.52       3.84 o
  13   1   0        16.87       15.76       4.56 o
  14   1   0        16.46        7.91       7.91 o
  15   1   0         0.00        3.95       5.59 o
   0   2   0      2443.71     2679.14      61.27 o
   1   2   0     23397.80    23770.90     546.30 o
   2   2   0     20572.37    19502.51     520.01 o
   3   2   0      8854.88     8282.53     169.57 o
```

Derived data

(six-dimensional structural model)



Crystallography Journals Online

Data exchange/archive standards

Crystallography Journals Online

Examples of CIF data

Formulae, coordinates

```
data_99107abs
_chemical_name_systematic
; 3-Benzo[b]thien-2-yl-5,6-dihydro-1,4,2-
     oxathiazine 4-oxide
;
_chemical_name_common                 ?
_chemical_formula_iupac              'C11 H9 N O2 S2'
_chemical_formula_moiety             'C11 H9 N O2 S2'
_chemical_formula_sum                'C11 H9 N O2 S2'
_chemical_formula_weight              251.31
loop_
 _atom_site_label
 _atom_site_type_symbol
 _atom_site_fract_x
 _atom_site_fract_y
 _atom_site_fract_z
 _atom_site_U_iso_or_equiv
 _atom_site_adp_type
S4   S   0.32163(7)    0.45232(6) 0.52011(3)
     0.04532(13) Uani
S11  S   0.39642(7)    0.67998(6) 0.29598(2)
     0.04215(12) Uani
O1   O   -0.00302(17) 0.67538(16) 0.47124(8)
     0.0470(3) Uani
O4   O   0.2601(2)     0.28588(16) 0.50279(10)
     0.0700(5) Uani
H5A  H   0.1284        0.4834 0.6221 0.060 Uiso
H5B  H   0.1861        0.6537 0.5908 0.060 Uiso
```

Raw (image) data

```
data_CXVT_0132
 loop_
 _array_data.array_id
 _array_data.binary_id
 _array_data.data
 image_1 1
;
--CIF-BINARY-FORMAT-SECTION--
Content-Type: application/octet-stream; conversions="x-
     CBF_PACKED"
Content-Transfer-Encoding: BASE64
X-Binary-Size: 3745758
X-Binary-ID: 1
X-Binary-Element-Type: "signed 32-bit integer"
Content-MD5: 1zsJjWPfol2GYl2V+QSXrw==

ELhQAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAADHcRzHcRxGQQwC
     ZsGuAKUFAIhS93U8
/91rMvpiEXw1pwoceMIBYHj78x7u9nszkeh7qm3XK6jk/Aa4x3Ecx3Ec
     x3Ecx3EcBzEEgApW
/y8xGar1BaqZXkcCow74Aw77fp8W5Sf2vP6O6A/SD8ZnixLf4/WMOzCg
     EAhqVnnv3wsk8oO9
EFa5G/3Gfq94GwLjHNE+gd8ndgf1foI+GN2LQIAneVRf9rXyCkwIyc/y
     /ILuHsdxHMdxHMdx
HMdxGAURAGraeHhJh6ccYbjVMKH5Xxjq3wx6VNQ4sPPCAxNNMrb/BEsA
     DjdtYzBLtS+VJMtJ
```

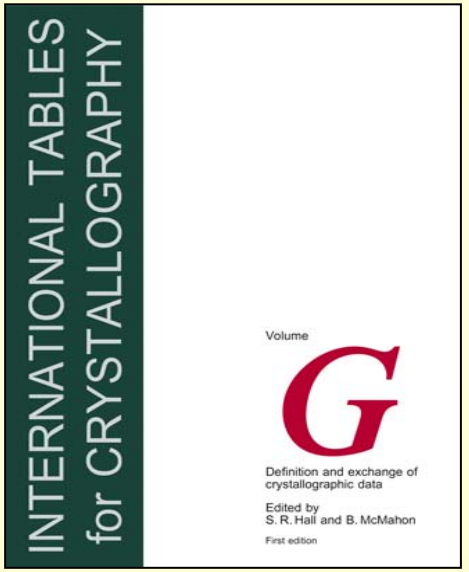
Data dictionary definition

```
data_chemical_formula_weight
     _name          '_chemical_formula_weight'
     _category              chemical_formula
     _type                  numb
     _enumeration_range     1.0:
     _units                 Da
     _units_detail        'daltons'
     _definition
;  Formula mass in daltons. This mass
   should correspond to the formulae given
   under _chemical_formula_structural,
   *_iupac, *_moiety or *_sum and, together
   with the Z value and cell parameters,
   should yield the density given as
   _exptl_crystal_density_diffrn.
;
```

Crystallography Journals Online

Standard description of data

- Crystallographic Information Framework
  - *International Tables for Crystallography* (2005). Vol. G, *Definition and exchange of crystallographic data*, edited by S. R. Hall & B. McMahon, 1st ed. Berlin: Springer.
- CIF file structure
  - Hall, S. R., Allen, F. H. & Brown, I. D. (1991). *The Crystallographic Information File (CIF): a new standard archive file for crystallography. Acta Cryst.* A**47**, 655-685
- Dictionary definition language
  - Hall, S. R. & Cook, A. P. F. (1995). *STAR dictionary definition language: initial specification. J. Chem. Inf. Comput. Sci.* **35**, 819-825.
- Data dictionaries



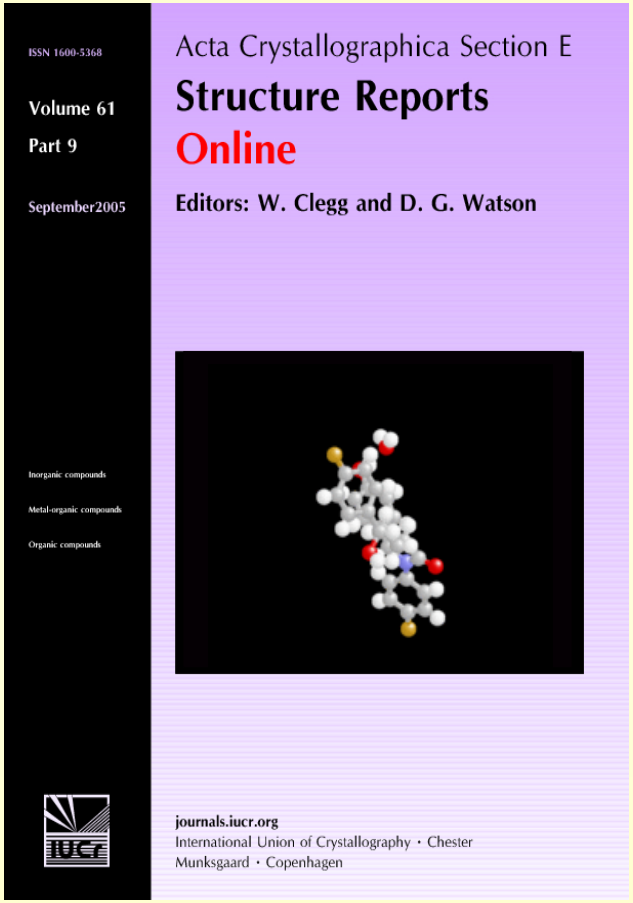
Crystallography Journals Online

Publication workflow for small-unit-cell structures

Crystallography Journals Online

Peer-reviewed structure-reports journals

- Data submitted as CIF
- Automated checking on submission
- Reviewer reports
- Automated page composition
- Key indicators
- Supplementary data sets



Crystallography Journals Online

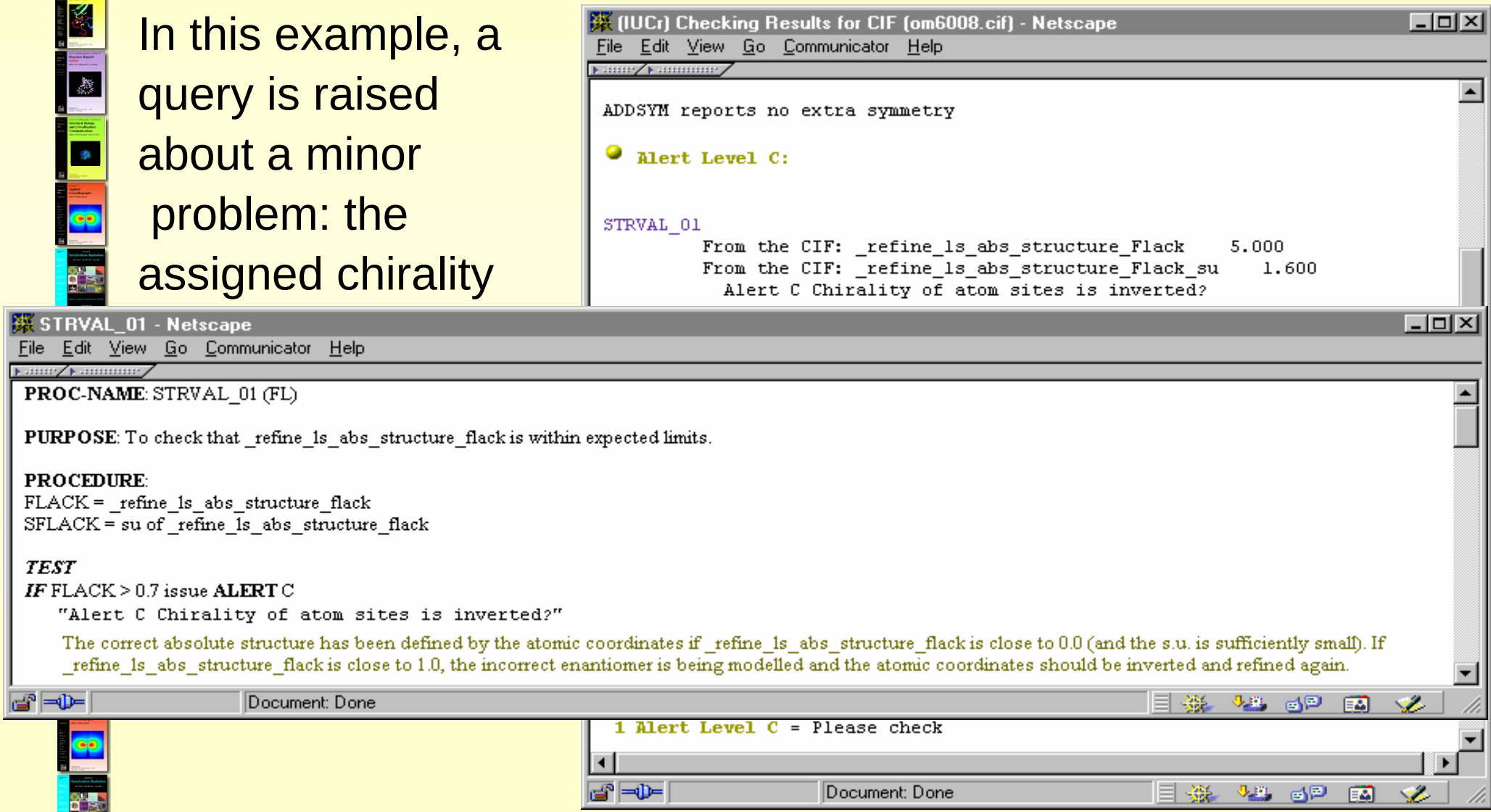
Technical aspects of peer review

- Check internal consistency of data dependencies (CIF dictionary)
- Check scientific reasonableness of model
- Check completeness of experimental metadata
- Check quality of derived structural model
- Consistency checks between raw, primary and derived data

Crystallography Journals Online

Feedback to submitting author (1)

In this example, a query is raised about a minor problem: the assigned chirality



Feedback to submitting author (2)

In this example, some mandatory information is missing: the author must explain or supply

(IUCr) Checking Results for CIF (a.cif) - Netscape

File Edit View Go Communicator Help

```
 2 Alert Level A = Potentially serious problem
 0 Alert Level B = Potential problem
 9 Alert Level C = Please check
```

VALIDATION ISSUES

The validation checking software has detected some potential problems with your CIF.

If you intend to submit this CIF for publication in an IUCr journal (*Acta Crystallographica*, *Journal of Applied Crystallography* or *Journal of Synchrotron Radiation)*, you should attempt to resolve the more serious problems (level A or B) before submission. This may involve additional measurements or structure refinements. However, the nature of your study may justify the reported deviations from the submission requirements of the journal. If this is the case, you can insert an explanation in your CIF using the Validation Reply Form (VRF) below. Your explanation will be assessed as part of the review process.

If you wish to submit your CIF for publication in *Acta Crystallographica Section C*, you should send your CIF to cifpub@iucr.org; submissions to *Acta Crystallographica Section E* should be made via the web (Submission Form). If your CIF is to form part of a submission to *Acta Crystallographica Section B*, the Co-editor handling your submission will ask you to send your CIF to Chester during the review of your paper.

```
_vrf_SYMM_004_bkl380
;
PROBLEM: Alert A _symmetry_equiv_pos_as_xyz loop is missing.
RESPONSE: ...
;
```

Document: Done

Example review report (1)

```
Bond precision:    C-C = 0.0036 A   Wavelength=0.71073
Cell: a=18.120(4)  b=11.317(2)  c=19.777(4)
      alpha=90  beta=90  gamma=90

                          Calculated                Reported
Volume                    4055.6(14)                4055.6(14)
Space group               P b c a                   P b c a
Hall group                -P 2ac 2ab                -P 2ac 2ab
Moiety formula            C22 H27 Cu N3 O2          C22 H27 Cu N3 O2
Sum formula               C22 H27 Cu N3 O2 C22      H27 Cu N3 O2
Mr                        429.02                    429.01
Dx,g cm-3                 1.405                     1.405
Z 8 8 Mu (mm-1)           1.099                     1.099
F000                      1800.0                    1800.0
F000'                     1803.09
h,k,lmax                  24,15,27                  24,15,27
Nref                      5559                      5497
Tmin,Tmax                 0.768,0.874               0.824,0.903
Tmin'                     0.644
Correction method= 'INTEGRATION'
Data completeness=  Ratio = 0.99
Theta(max)= 29.33
R(reflections)= 0.0465( 4147) wR2(reflections)= 0.0973( 5497)
S = 1.091 Npar= 255
```

Crystallography Journals Online

Example review report (2)

- **Alert level A**

```
PLAT725_ALERT_1_A D-H     Calc     0.91000, Rep 1.01000 Dev... 0.10 Ang.
                  N3   -H3       1.555   1.555
PLAT725_ALERT_1_A D-H     Calc     0.97000, Rep 1.09000 Dev... 0.12 Ang.
                  C19  -H19B     1.555   1.555
PLAT725_ALERT_1_A D-H     Calc     0.97000, Rep 1.09000 Dev... 0.12 Ang.
                  C29  -H29B     1.555   1.555
PLAT726_ALERT_1_A H...A   Calc     2.25000, Rep 2.16000 Dev... 0.09 Ang.
                  H3   -O11      1.555   5.665
```

- **Alert level C**

```
PLAT199_ALERT_1_C Check the Reported _cell_measurement_temperature 293 K
PLAT200_ALERT_1_C Check the Reported _diffrn_ambient_temperature . 293 K
PLAT728_ALERT_1_C D-H..A  Calc      118.00, Rep 116.00 Dev... 2.00 Deg.
                  C19 -H19B -O21      1.555   1.555   1.555
```

Crystallography Journals Online

Reader assessment

Mozilla Firefox

File Edit View Go Bookmarks Tools Help

Back Forward Reload Stop Home http://jou Go

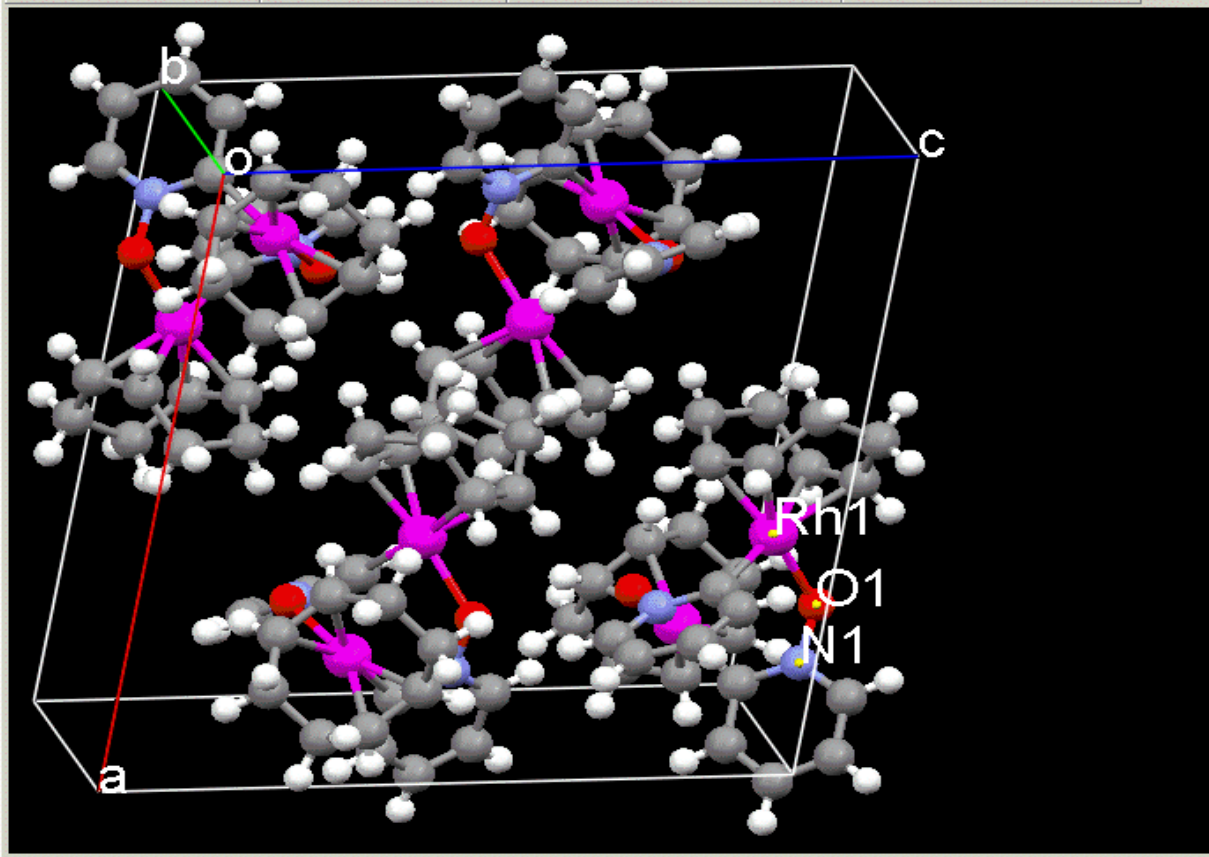
```
#
# h,k,l, Fc-squared, Fo-squared, sigma(Fo-squared) and status
```

I (P21/c)

File Edit Selection Display View Calculate Modify Databases Help

Style: Ball and Stick Colour: by Element Picking Mode: Select Atoms Clear Measurements

a b c a* b* c* x- x+ y- y+ z- z+ x -90 x +90 y -90 y +90 z -90 z +90 zoom - zoom +



Display

Packing

Short Contact < (sum of vdw radii)

H-Bond Default definition

Contacts...

More Info

Powder...

Reset

Options

Show hydrogen

Show cell axes

Label atoms

Press the left mouse button and move the mouse to rotate the structure

(IUCr) Structure Reports Online Contents - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

Back Forward Reload Stop Home http://journals.iucr

Structure Reports Online

Journals Onlin search help subscribe

html pdf abstract ci 3d view structure factors checkCIF buy

Acta Cryst. (2006). E**62**, m264-m266 [doi:10.1107/S16

Bis(μ-pyridinyl-1-oxide)-1κ*O*^{1}:2κ*C*^{2};1κ*C*^{2}:2κ*O*

R. Fandos and M. D. Walter

Online 11 January 2006

html pdf abstract cif 3d view structure factors checkCIF buy

Acta Cryst. (2006). E**62**, m267-m268 [doi:10.1107/S16

***catena*-Poly[[[diaquacalcium(II)]-di-μ-3-carb**

Community consensus for biological macromolecules

Crystallography Journals Online

Extending the approach

- Consensus in small-molecule crystallographic community
- Emerging standards in macromolecular crystallography
- *actabiostandards*

Crystallography Journals Online

Setting the standards

(IUCr) Data for structural communications - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

Back Forward Reload Stop Home http://journals.working.iucr.org/f/services/structuralcommunications/mmcifreqditems

1. Sample information

Description	mmCIF items
1.1. Macromolecule and source information	
Example 1: complex of *E. coli* glutamate decarboxylase α with glutarate Example 2: a zinc-induced heterodimer of two isoforms of phospholipase A_2 Example 3: mutation Example 4: mutation and modification	
Structure name	_struct.title
Component molecules	_entity.pdbx_description
Biological functional unit (BFU) or macromolecular assembly, numbers and types of chains	_struct_biol.details
Mass of BFU (Da)	_struct_biol.pdbx_formula_weight _struct_biol.pdbx_formula_weight_method
Macromolecule sequence and chemical configuration	
Sequence database reference code	_struct_ref.db_name _struct_ref.db_code
Polymers (one-letter code sequence)	_entity_poly.pdbx_seq_one_letter_code_can _entity_poly.pdbx_seq_one_letter_code
or Polymer sequence as list of residues	_entity_poly_seq.num _entity_poly_seq.mon_id

Validation of macromolecule structures

validation report

Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091
Editors: H. M. Einspahr and J. M. Guss

Crystal Structure of SodA-1 (BA4499) fror
Resolution

I.W. Boucher, V.M. Levdikov, E.V. Blagova, M.J. Wilkinson and K.S. Wilson

CONFIDENTIAL – NOT TO BE REPRODUCED, QUOTED NOR S

This document contains a summary of the results of the validation checks c
submitted for deposition.

It includes a brief summary of the compound name, sequence, cell para
PDB checking on distances and angles, close contacts, torsion angles, c
and a number of graphical indicators of geometric characteristics.

References

PROCHECK: Laskowski *et al.* (1993). *J. Appl. Cryst.* **26**, 283–291.

NUCHECK: Feng *et al.* (1998). *NUCHECK*. Report NDB-407, Rutgers University, New

SFCHECK: Vaguine *et al.* (1999). *Acta Cryst.* D**55**, 191–205

Co-editor:
Professor F. Frolow
Department of Molecular Microbiology and Biotechnology, Tel-Av
Telephone: 972(3)6408723
Fax: 972(3)6409407
Email: mbfrolow@post.tau.ac.il

validation report Tue Apr 26 11:43:36 2005

The following checks were made on :

DISTANCES AND ANGLES

We have checked your intra and in
procedures currently in place at

==> The following solvent molecul
macromolecule atoms which are
asymmetric unit.

HETATM 3440 O HOH 243
HETATM 3586 O HOH 389
HETATM 3612 O HOH 415
HETATM 3634 O HOH 437

The coordinates for water mol
the asymmetric unit are liste
it is the policy of the PDB t
entry with the ones below:

none

==> Close contacts in same asymme
Angstroms are considered as c

none

==> Close contacts based on cryst
Angstroms are considered as c

none

==> Bond and angle checks are per
error for all bonds and angle
units [L. Clowney et al., Geo
Bases, J.Am.Chem.Soc. 1996, 1
Parameters in Nucleic Acids:

1996, 118, 519-529] and amino
Bond and Angle Parameters for
Crystallogr. 1991, A47, 392-4
dictionary values by more tha
identified as an outlier.

*** Covalent Bond Lengths:

The RMS deviation for covalent bo
dictionary is 0.017 Angstroms

All covalent bonds lie within a 6
standard dictionary values.

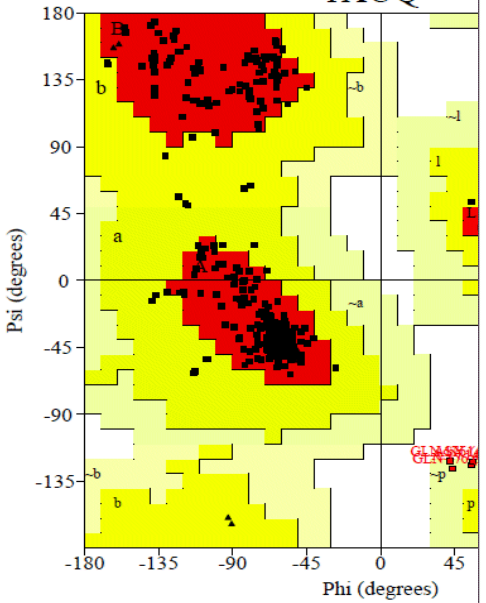
*** Covalent Angle Values:

The RMS deviation for covalent an
dictionary is 1.4 degrees.

All covalent bond angles lie with
standard dictionary values.

PROCHECK

Ramachandran
1XUQ



Plot statistics

Residues in most favoured regions [A,B,L]
Residues in additional allowed regions [a,b,l,p]
Residues in generously allowed regions [~a,~b,~l,~p]
Residues in disallowed regions

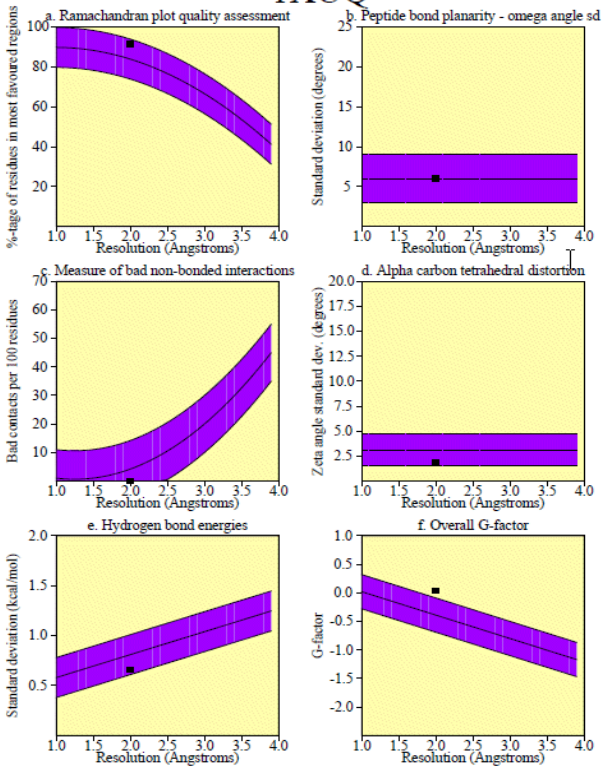
Number of non-glycine and non-proline residues
Number of end-residues (excl. Gly and Pro)
Number of glycine residues (shown as triangles)
Number of proline residues

Total number of residues

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected

PROCHECK

Main-chain parameters
1XUQ



Plot statistics

Stereochemical parameter	No. of data pts	Parameter value	Comparison values Typical value	Band width	No. of band widths from mean	
a. %-tage residues in A, B, L	344	91.6	83.8	10.0	0.8	Inside
b. Omega angle st dev	396	6.0	6.0	3.0	0.0	Inside
c. Bad contacts / 100 residues	0	0.0	4.2	10.0	-0.4	Inside
d. Zeta angle st dev	368	1.8	3.1	1.6	-0.8	Inside
e. H-bond energy st dev	264	0.7	0.8	0.2	-0.8	Inside
f. Overall G-factor	400	0.0	-0.4	0.3	1.4	BETTER

Data publication at source

Crystallography Journals Online

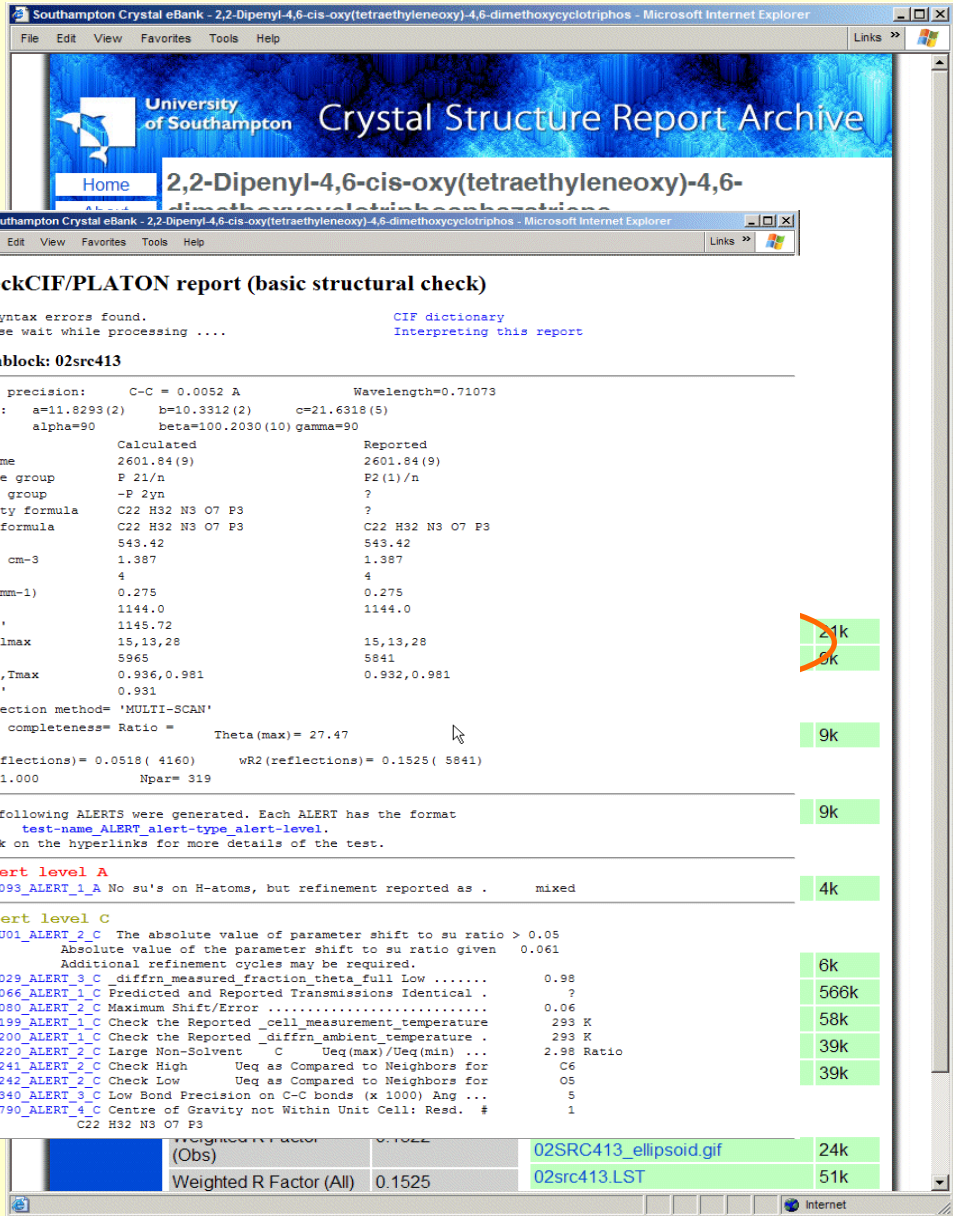
Making public the data

- Small-molecule crystallography ‘routine’
- Burden of writing full report articles in the literature
- Crystal structures by-products of chemistry research
- Valuable results never enter public domain
- Rise of laboratory ‘repositories’

Crystallography Journals Online

Extending the scholarly publication paradigm

- ePrints repository
- OAI-PMH
- Standard metadata
- All data
- Links to publication
- Rights
- Quality



ALPSP Award 2006

ALPSP Award for Publishing Innovation

This year, the panel reviewed 12 applications from which they selected a shortlist of three. The judges considered the originality and innovative qualities of the projects submitted, together with their utility and long term development prospects.

This year's award was made to the **International Union of Crystallography (IUCr)** for their **Data Exchange, Quality Assurance and Integrated Data Publication (CIF and *check*CIF)**.

The judges were impressed with the way in which CIF and *check*CIF are easily accessible and have served to make critical crystallographical data more consistently reliable and accessible at all stages of the information chain, from authors, reviewers and editors through to readers and researchers. In doing so, the system takes away the donkeywork from ensuring that the results of scientific research are trustworthy without detracting from the value of human judgement in the research and publication process.

The development and maintenance of CIF and *check*CIF is sponsored by several publishers, but it is freely accessible to all. IUCr already works closely with other related structural science communities and is looking to extend this cooperation. The judges felt that in developing CIF and *check*CIF, the IUCr has established an important example of data quality assurance with potential applications in other scientific, medical, and indeed social sciences publishing.

"The IUCr is honoured by the 2006 ALPSP Award for Publishing Innovation, which recognises the hard work and dedication of our publishing staff and academic collaborators, and the role that learned societies can play in introducing novel and valuable contributions to scientific information exchange. The Crystallographic Information Framework owes much to the special nature of crystallography and its relatively compact community of practitioners; but we hope that this award will encourage other scientific disciplines to follow similar approaches to integrating research data and literature, and to extending the tradition of peer review more deeply into the supporting data."

Peter Strickland, Managing Editor, IUCr Publications

Crystallography Journals Online

Summary

- Standard data format
- Automated checking/quality assessment
- Objective publication standards
- Adoption of standards in wider community
- Improvement in quality
- Potential to extend consistency checking even further

Crystallography Journals Online