```
\structno
\noindent{\nineit Crystal data}\nobreak\par
\vskip2pt\begindoublecolumns\twocoltrue\defaultfont
\raggedright
\everypar={\global\parindent=0pt\hangindent=1em
          \hangafter=1 }\noindent

\chemformiupac{C$_{10}$H$_{8}$BrNS}
\chemform{C$_{10}$H$_{8}$BrNS}
\chemformsum{C$_{10}$H$_{8}$BrNS}
\molwt{254.14}
\system{Orthorhombic}\def\sgsetno{2}
\defaultfont
\sgHM{$P2{}_{1}2{}_{1}2{}_{1}$}
\cella{5.7339 (7)}
\cellb{14.8229 (15)}
\cellc{23.469 (2)}
\cellalpha{90.00}
\cellbeta{90.00}
\cellgamma{90.00}
\cellvol{1994.7 (4)}
\cellz{8}
\dx{1.693}
\dm{missing}
\densmetha{not measured}
\radiationtype{Mo {\it K}$\alpha$}
\wavelength{0.71073}
\cellrefl{58}
\cellthetamin{4.8}
\cellthetamax{11.6}
\absorpmu{4.28}
\celltemp{296 (2)}
\shape{Plate}
\colour{Colourless}
\sizemax{0.40}
\sizemid{0.32}
\sizemin{0.04}
\sizerad{missing}
\origin{see text}
```

Fig. 5.7.3.2. Part of a TEX file used to print the article shown in Fig. 5.7.3.1(*b*).

```
<sec id="sec2.1">
 <sec id="sec2.1.1">
  <st><?print tpct=0pt>Crystal data</st>
  <p>
   <l id="l1" type="unord">
    <li><p>C<inf arrange="stagger">10</inf>H<inf
        arrange="stagger">8</inf>BrNS</p></li>
    <li><p><it>M</it><inf
        arrange="stagger"><it>r</it></inf> =
        254.14</p></li>
    <li><p>Orthorhombic, 
     <fi type="tex" print-info="tth" img.type="tth"
        img.data="teximages/ga1014fi3.tth">
        P2_{1}2_{1}2_{1}</fi></p></li>
    <li><p><it>a</it> =
        5.7339 (7) &Aring;</p></li>
    <li><p><it>b</it> =
        14.8229 (15) &Aring;</p></li>
    <li><p><it>c</it> =
        23.469 (2) &Aring;</p></li>
    <li><p><it>V</it> =
        1994.7 (4) &Aring;<sup
        arrange="stagger">3</sup></p></li>
    <li><p><it>Z</it> =
        8</p></li>
    <li><p><it>D</it><inf
        arrange="stagger"><it>x</it></inf> =
        1.693 Mg m<sup
        arrange="stagger">&minus;3</sup></p></li>
    <li><p>Mo <it>K</it>&alpha; radiation</p></li>
    <li><p>Cell parameters from 58 <?print show
        &softreturn;>reflections</p></li>
    <li><p>&thetas; = 4.8&ndash;11.6&deg;</p></li>
    <li><p>&mu; = 4.28 mm<sup
        arrange="stagger">&minus;1</sup></p></li>
    <li><p><it>T</it> =
         296 (2) K</p></li>
    <li><p>Plate, colourless</p></li>
    <li><p>0.40 &times; 0.32 &times;
         0.04 mm</p></li>
   </l>
  </p>
</sec>
```

Fig. 5.7.3.3. Part of the SGML file used to print the article shown in Fig. 5.7.3.1(*b*).

Section 5.3.8.2.1 describes one transformation to XML in the biological structures field, designed primarily for database interchange rather than publication. This transformation preserves the underlying data model of an mmCIF very closely, and one might anticipate similar XML transformations for small-molecule CIF applications and for publications. It is even possible that the XML transformations referred to in Chapter 5.3 could be used for publishing articles if suitable style transformations are developed, but this has not been tested yet.

One difficulty with a simple CIF-to-XML transformation is that it could be easily adapted to the publication of structure reports in dedicated journals, but would not necessarily be compatible with other XML implementations developed by an unspecialized publishing house. This could be avoided by the registration of an XML name space covering transformed CIF data and the production of portable stylesheet transformations that could be adopted and modified to meet the requirements of different publishing houses. As yet, we know of no initiatives in this direction.

XML name spaces have been registered to safeguard the development of subject-specific methods of representation as part of a project by the International Union of Pure and Applied Chemistry (Becker, 2001). One markup language that falls within the scope of this project is Chemical Markup Language (CML) (Murray-Rust & Rzepa, 1999, 2001).

Further discussions of the relationship between CIF and XML representations and a proposal for extensions to certain CIF data values to accommodate the wider range of data structures permitted in XML are given by Bernstein (2000).

We acknowledge the guidance, enthusiasm and dedication of past and present members of the editorial boards of *Acta Crystallographica Sections C* and *E* in developing the journals along the path described in this chapter. Particular tribute must be paid to Syd Hall, George Ferguson, Bill Clegg, David Watson and Tony Linden. We are very grateful to Ton Spek for his close involvement with the development of checking software, and also wish to acknowledge George Sheldrick, Mario Nardelli, Eric Gabe, Peter White, Yvon Le Page, Alan Mighell, Vicky Karen, Doug du Boulay, Mike Dacombe and Charlie Bugg for their help in the early days of automated structure checking. We wish also to pay tribute to the dedication and effort of our colleagues in the IUCr editorial office: Gillian Holmes, Sean Conway, Amanda Berry, Sarah Froggatt and Lisa Stephenson; and we thank the many authors who have been willing to test new approaches through the years.

Appendix 5.7.1
Request list for *Acta Crystallographica Section C*

Table A5.7.1.1 contains the request list for *Acta Crystallographica Section C* as given in the 2005 *Notes for Authors*. This list is appropriate for a single-crystal X-ray diffraction study and gives all the data items that are displayed in an article *if they are present in the CIF*. In principle, a smaller set of mandatory data items could be supplied as a separate request list. However, certain items may be

considered mandatory or not depending on the nature of the study and on the presence of other data items in the CIF, so checking for mandatory items is performed through higher-level algorithmic checks during the pre-submission validation stage.

Table A5.7.1.1. *Request list for Acta Crystallographica Section C*

(*a*) Data names relating to the text of an article	
`_publ_contact_author_name`	Contact author's name
`_publ_contact_author_address`	Contact author's address
`_publ_contact_author_email`	E-mail address to be published
`_publ_contact_author_fax`	For editorial communications
`_publ_contact_author_phone`	For editorial communications
`_publ_contact_letter`	Letter of submission, with date
`_publ_requested_journal`	'Acta Crystallographica Section C'
`_publ_requested_category`	Publication choice (FI, FM, FO, AD)
`_publ_section_title`	Title of paper
`_publ_section_title_footnote`	Footnote to title of paper
`_publ_author_name`	List of author(s) name(s)
`_publ_author_footnote`	Footnote(s) to author(s) name(s)
`_publ_author_address`	Author(s) address(es)
`_publ_section_synopsis`	Synopsis for compounds that cannot be shown as a chemical diagram
`_publ_section_abstract`	Abstract of paper in English
`_publ_section_comment`	Discussion of study
`_publ_section_acknowledgements`	Acknowledgements
`_publ_section_references`	References
`_publ_section_figure_captions`	Legends to figures
(*b*) Data names relating to the experimental data	
`_publ_section_exptl_prep`	Compound preparation details
`_chemical_formula_sum`	Chemical formula as sum of elements
`_chemical_formula_moiety`	Chemical formula in moieties
`_chemical_formula_weight`	Chemical formula mass (Da)
`_chemical_melting_point`	Melting point (K)
`_symmetry_cell_setting`	Code for cell setting
`_symmetry_space_group_name_H-M`	Space-group symbol, including unique axis
`_symmetry_equiv_pos_as_xyz`	Equivalent positions in order used by `_geom_`
`_cell_length_a`	Unit-cell lengths (Å)
`_cell_length_b`	
`_cell_length_c`	
`_cell_angle_alpha`	Unit-cell angles (°)
`_cell_angle_beta`	
`_cell_angle_gamma`	
`_cell_volume`	Unit-cell volume (Å^3)
`_cell_formula_units_Z`	Number of formulae per unit cell
`_exptl_crystal_density_diffrn`	Density calculated from unit cell and contents (Mg m^{-3})
`_exptl_crystal_density_meas`	Density measured experimentally (Mg m^{-3})
`_exptl_crystal_density_method`	Method used to measure density experimentally
`_diffrn_radiation_type`	Radiation type (*e.g.* neutron or Mo $K\alpha$)
`_diffrn_radiation_wavelength`	Radiation wavelength (Å)
`_cell_measurement_reflns_used`	Number of reflections used to measure unit cell
`_cell_measurement_theta_min`	Minimum θ of reflections used to measure unit cell (°)
`_cell_measurement_theta_max`	Maximum θ of reflections used to measure unit cell (°)
`_cell_measurement_temperature`	Measurement temperature (K)
`_exptl_absorpt_coefficient_mu`	Linear absorption coefficient (mm^{-1})
`_exptl_crystal_description`	Crystal habit description
`_exptl_crystal_size_max`	Maximum dimension of crystal (mm)
`_exptl_crystal_size_mid`	Medial dimension of crystal (mm)
`_exptl_crystal_size_min`	Minimum dimension of crystal (mm)
`_exptl_crystal_size_rad`	Radius of spherical or cylindrical crystal (mm)
`_exptl_crystal_colour`	Crystal colour
`_diffrn_measurement_device_type`	Diffractometer make and type
`_diffrn_measurement_method`	Mode of intensity measurement and scan
`_diffrn_detector_area_resol_mean`	Resolution of area detector (pixels mm^{-1})
`_exptl_absorpt_correction_type`	Code for absorption correction
`_exptl_absorpt_process_details`	Literature reference for absorption correction [*e.g.* '(North et al., 1968)']
`_exptl_absorpt_correction_T_min`	Minimum transmission factor from corrections
`_exptl_absorpt_correction_T_max`	Maximum transmission factor from corrections
`_diffrn_reflns_number`	Total number of reflections measured
`_reflns_number_total`	Number of symmetry-independent reflections
`_reflns_number_gt`	Number of reflections > σ threshold
`_reflns_threshold_expression`	σ expression for F, F^2 or I threshold
`_diffrn_reflns_theta_max`	Maximum θ of measured reflections (°)
`_diffrn_reflns_theta_full`	θ to which available reflections are 'complete' (°)
`_diffrn_measured_fraction_theta_max`	Fraction of unique reflections measured to θ_{max}
`_diffrn_measured_fraction_theta_full`	Fraction of unique reflections measured to θ_{full}
`_diffrn_reflns_av_R_equivalents`	R factor for symmetry-equivalent intensities
`_diffrn_reflns_limit_h_min`	Minimum/maximum h index of measured data
`_diffrn_reflns_limit_h_max`	
`_diffrn_reflns_limit_k_min`	Minimum/maximum k index of measured data
`_diffrn_reflns_limit_k_max`	
`_diffrn_reflns_limit_l_min`	Minimum/maximum l index of measured data
`_diffrn_reflns_limit_l_max`	
`_diffrn_standards_number`	Number of standards used in measurement
`_diffrn_standards_interval_count`	Number of measurements between standards
`_diffrn_standards_interval_time` *	Time (min) between standards
`_diffrn_standards_decay_%`	Percentage decrease in standards intensity
`_refine_ls_structure_factor_coef`	Code for F, F^2 or I used in least-squares refinement
`_refine_ls_R_factor_gt`	R factor of F for reflections > threshold
`_refine_ls_wR_factor_ref`	R factor of coefficient for refinement reflections
`_refine_ls_goodness_of_fit_ref`	Goodness of fit S for refinement reflections
`_refine_ls_number_reflns`	Number of reflections used in refinement
`_refine_ls_number_parameters`	Number of parameters refined
`_refine_ls_weighting_scheme`	Code for weight type
`_refine_ls_weighting_details`	Weighting expression
`_refine_ls_hydrogen_treatment`	Code for H-atom treatment
`_refine_ls_shift/su_max`	Maximum shift/s.u. ratio after final refinement cycle
`_refine_diff_density_max`	Maximum/minimum values of final difference map (e Å^{-3})
`_refine_diff_density_min`	
`_refine_ls_extinction_method`	Description of extinction methods applied
`_refine_ls_extinction_coef`	Extinction coefficient applied in corrections
`_refine_ls_abs_structure_details`	Absolute structure method and Friedel-pair number
`_refine_ls_abs_structure_Flack`	Measure of absolute structure
`_refine_ls_abs_structure_Rogers`†	Measure of absolute structure
`_publ_section_exptl_refinement`	Special details of the refinement
`_computing_data_collection`	Reference to data-collection software
`_computing_cell_refinement`	Reference to cell-refinement software
`_computing_data_reduction`	Reference to data-reduction software
`_computing_structure_solution`	Reference to structure-solution software
`_computing_structure_refinement`	Reference to structure-refinement software
`_computing_molecular_graphics`	Reference to visualization software
`_computing_publication_material`	Reference to publication preparation software
`loop_`	
`_atom_type_symbol`	Atom type symbol (usually element symbol)
`_atom_type_description`	Description of atom type
`_atom_type_scat_source`	Reference to scattering factors applied
`_atom_type_scat_dispersion_real`	Real anomalous-dispersion value applied
`_atom_type_scat_dispersion_imag`	Imaginary anomalous-dispersion value applied
`loop_`	
`_atom_site_label`	Unique label identifying the atom site
`_atom_site_fract_x`	Fractional coordinates of atom site
`_atom_site_fract_y`	
`_atom_site_fract_z`	
`_atom_site_U_iso_or_equiv`	Isotropic atomic displacement parameter, or equivalent from anisotropic atomic displacement parameters
`_atom_site_occupancy`	Occupancy fraction for site (default is 1.0)
`_atom_site_disorder_assembly`	Code that identifies functional group suffering disorder
`_atom_site_disorder_group`	Code that identifies disorder group
`_atom_site_adp_type`	Atomic displacement parameter type
`loop_`	
`_atom_site_aniso_label`	Unique label identifying the atom site
`_atom_site_aniso_U_11`	Elements of anisotropic atomic displacement parameter tensor
`_atom_site_aniso_U_22`	
`_atom_site_aniso_U_33`	
`_atom_site_aniso_U_12`	
`_atom_site_aniso_U_13`	
`_atom_site_aniso_U_23`	
`loop_`	
`_geom_bond_atom_site_label_1`	Labels identifying the atom sites 1 and 2
`_geom_bond_atom_site_label_2`	
`_geom_bond_site_symmetry_1`	Symmetry codes (*e.g.* 2_554) for atom sites 1 and 2
`_geom_bond_site_symmetry_2`	
`_geom_bond_distance`	Distance between atom sites 1 and 2 (Å)
`_geom_bond_publ_flag`	Flag for print request (yes or no)

* Alternative to `_diffrn_standards_interval_count`. † Alternative to `_refine_ls_abs_structure_Flack`.

Table A5.7.1.1 (*cont.*).

`loop_`	
`_geom_angle_atom_site_label_1`	Labels identifying the atom sites 1, 2 and 3
`_geom_angle_atom_site_label_2`	
`_geom_angle_atom_site_label_3`	
`_geom_angle_site_symmetry_1`	Symmetry codes for atom sites 1, 2 and 3
`_geom_angle_site_symmetry_2`	
`_geom_angle_site_symmetry_3`	
`_geom_angle`	Angle between atom sites 1, 2 and 3 (°)
`_geom_angle_publ_flag`	Flag for print request (yes or no)
`loop_`	
`_geom_torsion_atom_site_label_1`	Labels identifying the atom sites 1, 2, 3 and 4
`_geom_torsion_atom_site_label_2`	
`_geom_torsion_atom_site_label_3`	
`_geom_torsion_atom_site_label_4`	
`_geom_torsion_site_symmetry_1`	Symmetry codes for atom sites 1, 2, 3 and 4
`_geom_torsion_site_symmetry_2`	
`_geom_torsion_site_symmetry_3`	
`_geom_torsion_site_symmetry_4`	
`_geom_torsion`	Torsion angle between atom sites 1, 2, 3 and 4 (°)
`_geom_torsion_publ_flag`	Flag for print request (yes or no)
`loop_`	
`_geom_hbond_atom_site_label_D`	Donor-atom label in hydrogen bond
`_geom_hbond_atom_site_label_H`	H-atom label in hydrogen bond
`_geom_hbond_atom_site_label_A`	Acceptor-atom label in hydrogen bond
`_geom_hbond_site_symmetry_D`	Symmetry code for donor site
`_geom_hbond_site_symmetry_H`	Symmetry code for H-atom site
`_geom_hbond_site_symmetry_A`	Symmetry code for acceptor site
`_geom_hbond_distance_DH`	Donor atom-to-H-atom distance (Å)
`_geom_hbond_distance_HA`	H-atom-to-acceptor atom distance (Å)
`_geom_hbond_distance_DA`	Donor atom-to-acceptor atom distance (Å)
`_geom_hbond_angle_DHA`	Donor to H to acceptor angle (°)
`_geom_hbond_publ_flag`	Flag for print request (yes or no)
(*c*) Data names for adding items to the standard request list	
`loop_`	
`_publ_manuscript_incl_extra_item`	Additional CIF item submitted for publication
`_publ_manuscript_incl_extra_defn`	Is item defined in core dictionary? (yes or no)
(*d*) Data names for structure-factor lists	
`loop_`	
`_refln_index_h`	Miller indices *h*, *k* and *l*
`_refln_index_k`	
`_refln_index_l`	
`_refln_F_meas`	Measured F
`_refln_F_squared_meas`‡	Measured F^2
`_refln_F_sigma`	Standard uncertainty of F
`_refln_F_squared_sigma`‡	Standard uncertainty of F^2
`_refln_F_calc`	Calculated F
`_refln_F_squared_calc`‡	Calculated F^2

‡ Alternative to the corresponding data name without 'squared'.

Appendix 5.7.2
Data validation using *checkcif*

Table A5.7.2.1 lists the *checkcif* tests concerned primarily with the completeness and self-consistency of individual or closely related data items. These tests were developed from the routines of *PREPUB* (du Boulay & Hall, 1996) and in the IUCr Editorial Office. Table A5.7.2.2 lists the tests applied specifically by the program *PLATON* (Spek, 2003), which performs a more detailed crystallographic analysis of the structure itself.

Each entry in each table has an identifying code and a numeric type. The type is used to categorize the alert messages generated when the tested values deviate from assigned norms. Type 1 refers to syntactic or other errors of construction in the CIF, or to inconsistent or missing data. Type 2 alerts indicate that the structure model may be wrong or deficient. Type 3 alerts indicate that the quality of the structure may be low, owing to limited or incomplete data coverage. Alerts of type 4 are indicative of deviations from style or suggested good practice, or may offer suggestions for improvement in presentation. The alerts within each category may be of varying levels of severity.

Full details of the tests and algorithms applied for the *checkcif* tests may be found at http://journals.iucr.org/services/cif/datavalidation.html or on the CD-ROM accompanying this volume. These include comments which provide help in interpreting the results of the tests and suggest ways in which the author can improve the data. The comments were provided by A. Linden and other members of the IUCr journal editorial boards.

The tests listed in Tables A5.7.2.1 and A5.7.2.2 are appropriate for small-unit-cell single-crystal structure determinations. More discriminating tests are being introduced for powder diffraction studies and for modulated structures.

Table A5.7.2.1. *List of data-validation tests applied by checkcif*

Test name	Type	Purpose
ABSMU01	1	Check that μ is consistent with the cell contents
ABSTM01	1	Check that $T_{\min}$ is less than $T_{\max}$
ABSTM02	3	Check that $T_{\min}$ and $T_{\max}$ are appropriate to the crystal size and μ
ABSTY01	1	Check that `_exptl_absorpt_correction_type` is a recognized keyword
ABSTY02	1	Check that `_exptl_absorpt_correction_type` contains some reference text
CELLK01	1	Check that temperature is in Kelvin
CELLT01	1	Check that $\theta_{\min}$ is less than $\theta_{\max}$
CELLV01	1	Check that the `_cell_volume` matches `_cell_length_` and `_cell_angle_` values
CELLV02	1	Check that the `_cell_volume` s.u. matches `_cell_length_` and `_cell_angle_` s.u. values
CELLZ01	1	Check consistency between formula, *Z*, atom list and symmetry
CHEMS01	1	Check that the `_chemical_formula_sum` is properly constructed
CHEMS02	1	Check that the stated category is consistent with the formula of the compound
CHEMW01	1	Check consistency between `_chemical_formula_weight` and `_chemical_formula_sum`
CHEMW03	2	Check consistency between weight, *Z*, symmetry and atom list
CRYSC01	1	Check that colour of crystal is consistent with expected colour code combinations
CRYSR01	1	Check that the radius of the crystal is given for a spherical or cylindrical crystal
CRYSS01	1	Check consistency of crystal dimensions
CRYSS02	3	Check that the values of `_exptl_crystal_size_`* are not larger than expected
DENSD01	1	Check consistency of density, cell volume and weight
DENSX01	1	Check that `_exptl_crystal_density_meas` matches `_exptl_crystal_density_diffrn`
DIFMN01	1	Check that `_refine_diff_density_min` is less than `_refine_diff_density_max`
DIFMN02	2	Check that `_refine_diff_density_min` is within expected limits
DIFMN03	1	Check for adjacent site information if `_refine_diff_density_min` is outside expected limits
DIFMX01	2	Check that `_refine_diff_density_max` is within expected limits
DIFMX02	1	Check for adjacent site information if `_refine_diff_density_max` is outside expected limits
FCOEF01	1	Check that the value of `_refine_ls_structure_factor_coef` is recognized
FORMU01	1/2	Check consistency between formulae and atom site data
GOODF01	1/2	Check that `_refine_ls_goodness_of_fit_ref` is within expected limits
HYDTR01	1	Check that the value of `_refine_ls_hydrogen_treatment` is recognized
RADNT01	1	Check that the radiation type is recognized
RADNW01	1	Check that `_diffrn_radiation_wavelength` matches `_diffrn_radiation_type`
REFLE01	3	Check that `_reflns_threshold_expression` contains a multiplier which is below the limit
REFLG01	1	Check that `_reflns_number_gt` is less than or equal to `_diffrn_reflns_number`
REFLL01	1	Check that `_diffrn_reflns_limit_` values are in the correct order
REFLT01	1	Check that `_reflns_number_total` is less than or equal to `_diffrn_reflns_number`
REFLT02	1	Check that `_reflns_number_total` is greater than or equal to `_reflns_number_gt`
REFLT03	1/3/4	Check consistency of `_reflns_number_total` with cell volume, symmetry and $\theta_{\max}$
REFNR01	3	Check the ratio of `_refine_ls_number_reflns` and `_refine_ls_number_parameters`
RFACG01	3	Check that `_refine_ls_R_factor_gt` is within expected limits
RFACR01	3	Check that `_refine_ls_wR_factor_ref` is within expected limits
RINTA01	3	Check that `_diffrn_reflns_av_R_equivalents` is within expected limits
SHFSU01	2	Check that `_refine_ls_shift/su_max` is within expected limits
STRDE01	1	Check that `_refine_ls_abs_structure_details` is present if necessary
STRVA01	2/4	Check that `_refine_ls_abs_structure_Flack` is within expected limits
STRVA02	2/3/4	Check that `_refine_ls_abs_structure_Rogers` is within expected limits
SYMMG01	1	Check that the `_symmetry_space_group_name_H-M` value is recognized