3.5. Classification and use of electron density data

By P. R. Mallinson and I. D. Brown

3.5.1. Introduction

Modern high-resolution X-ray diffraction methods enable the electron density distributions in many crystals to be accurately determined, allowing one to explore the interactions that hold crystals and molecules together (Coppen, 1997; Koritsanszky & Coppen, 2001). For example, Bader’s (1990) theory of atoms in molecules gives a description of chemical bonding in terms of the topology of the electron density, but it requires an accurate knowledge of the electron density. [Note that the terms ‘electron density’ (or distribution) and ‘charge density’ tend to be used interchangeably in the literature. The term ‘electron density’ is used here since ‘charge density’ strictly includes both positive (nuclear) and negative (electronic) charge. Properties such as electrostatic potential, electric fields and their gradients, and molecular electric moments are derived from the charge density distribution, which includes the contribution of the nuclear charge (Spackman & Brown, 1994).] The advent of area detectors, more powerful laboratory computing facilities and developments in software have combined to produce a rapid increase in the number of accurate electron density determinations now available. While such studies are not yet routine, they are sufficiently common to justify defining a structure for a secure and systematic archive. This in turn requires a method of representing the electron density acceptable to the crystallographic community.

The CIF standard which has been adopted by the IUCr and is in common use in macromolecular and small-molecule crystallography is the obvious choice for such a file structure. At the Sagamore XIII conference held at Stare Jablonki, Poland, in 2000, the IUCr Commission on Charge, Spin and Momentum Densities decided to prepare a CIF dictionary for reporting electron densities in crystals. The result is the rhoCIF dictionary described in this chapter. It was prepared in conjunction with members of the Commission and other members of the community. [The name rhoCIF for this dictionary was chosen over the more obvious choice of edCIF to avoid any possible confusion with electron diffraction. The distribution of electron density in a molecule or crystal can also be calculated theoretically, and the results are usually compared with experiment. The Hohenberg–Kohn theorem (Hohenberg & Kohn, 1964) shows that the ground-state energy is a unique functional of the electron density, thus providing a theoretical justification for relating chemical and physical properties directly to the observable electron density. Density functional theory has been applied to chemical problems for some time (see, for example, Labanowski & Andzelm, 1991; Ziegler, 1991; Kohn et al., 1996; Nagy, 1998). Although the current version of the rhoCIF dictionary is intended to record the results of experimental measurements, it can also be used to represent a theoretical electron density that has been fitted with an appropriate model.]

3.5.2. Dictionary design considerations

The rhoCIF dictionary contains the definitions of data items used to report accurate electron densities in crystals. The structure of these definitions conforms to the CIF rules and conventions, and selected data items reflect those currently needed in the field. In any field of science, the nature of the data used to describe the experiment and the results changes with time, and accurate electron density studies are no exception. The items defined in this dictionary describe the density representation approach used most often today – the atom-centred multipole expansion model introduced by Stewart (1973) and refined by Hansen & Coppen (1978). In this model, the electron density is expressed as the sum of three components. Two are the partial contributions of the core and valence shell of the spherical free atom, usually taken from the paper by Clementi & Roetti (1974). The third component is an aspherical multipole expansion of the remaining electron density. This is an atom-centred function, flexible enough to describe most electron density distributions likely to be encountered in a crystal, and with components that have the same symmetries as the atom-centred wave functions (orbitals) frequently used by chemists when describing chemical bonding.

3.5.3. Classification of data definitions

The rhoCIF dictionary is intended to be used in combination with the core CIF dictionary, or, after adaptation to DDL2, the mmCIF dictionary. The two dictionaries can be combined to form a composite dictionary by concatenation. Alternatively, the merging procedures described in Section 3.1.9 can be used to create a virtual dictionary at run time. Because the dictionary is not self-contained, its structure can be kept very simple, comprising only two categories: one to define a local axis system around each atom and one to contain the various parameters of the Hansen & Coppen (1978) multipole function.

The two categories in the electron density CIF dictionary lie within the single ATOM category group and relate to the structural model. They are as follows:

- Specification of local axes at each atom (§3.5.3.1)
  - ATOM_LOCAL_AXES
- Atom-centred multipole expansion functions (§3.5.3.2)
  - ATOM_RHO_MULTIPOLE

3.5.3.1. Specification of local axes at each atom

Data items in this category are as follows:

- ATOM_LOCAL_AXES
- _atom_local_axes_atom label
- _atom_site_label
- _atom_local_axes_atom0
- _atom_local_axes_atom1
- _atom_local_axes_atom2
- _atom_local_axes_ax1
- _atom_local_axes_ax2

The bullet (•) indicates a category key. The arrow (→) is a reference to a parent data item.
3. CIF DATA DEFINITION AND CLASSIFICATION

Example 3.5.3.1. Definition of the local Cartesian axis system for the atoms in a proton sponge complex.

```
loop
  _atom_local_axes_atom_label
  _atom_local_axes_atom0
  _atom_local_axes_atom1
  _atom_local_axes_atom2
  _atom_local_axes_atom_label
N(1)  C(1)  X  N(1)  N(2)  Y
C(2)  C(1)  X  C(1)  C(2)  Y
C(3)  C(1)  X  C(1)  C(2)  Y
C(4)  C(10) X  C(4)  C(3)  Y
C(5)  C(10) X  C(5)  C(6)  Y
C(6)  C(10) X  C(6)  C(7)  Y
C(7)  C(8) X  C(7)  C(6)  Y
C(8)  C(9) X  C(8)  N(2)  Y
C(9)  C(10) X  C(9)  C(1)  Y
C(10) C(9) X  C(10) C(5)  Y
C(11) N(1) X  C(11) H(111) Y
C(12) N(2) X  C(12) H(141) Y
CL(1A) C(1A) X  CL(1A) CL(2A) Y
CL(2A) C(1A) X  CL(2A) CL(1A) Y
O(1A) C(1A) X  O(1A) H(100) Y
O(2A) C(1A) X  O(2A) CL(1A) Y
O(3A) C(1A) X  O(3A) CL(2A) Y
O(4A) C(1A) X  O(4A) H(100) Y
C(1A) O(1A) X  C(1A) O(2A) Y
C(2A) C(1A) X  C(2A) CL(1A) Y
C(3A) C(1A) X  C(3A) CL(2A) Y
O(4A) O(1A) X  O(4A) H(100) Y
C(1A) C(8) X  C(1A) N(2) Y
C(2) C(1) X  C(2) C(3) Y
C(1) N(1) X  C(1) C(2) Y
N(2) C(8) X  N(2) N(1) Y
N(3) C(3) X  N(3) H(3) Y
N(4) N(1) X  N(4) H(141) Y
N(5) C(5) X  N(5) H(4) Y
H(6) C(6) X  H(6) H(7) Y
H(7) C(7) X  H(7) H(6) Y
```

The bullet (•) indicates a category key. The arrow (→) is a reference to a parent data item.

Data items in this category are also looped. The category key (_atom_rho_multipole_atom_label) is also a child of _atom_site_label in the ATOM_SITE category, thereby linking the electron density multipole coefficients to a specific atom site in the list of atom coordinates. This category contains all the parameters of the Hansen & Coppens (1978) function

\[
\rho(r) = \rho_{\text{core}}(r) + \rho_{\text{valence}}(r) \\
+ \sum_{l \leq l_{\text{max}}} \sum_{m \leq l} \{P(l,m)d(l,m,\theta,\varphi)\},
\]

where \(r\) is a position vector with respect to the atomic nucleus, and \(\rho_{\text{core}}(r)\) and \(\rho_{\text{valence}}(r)\) are the spherical core and valence electron densities, respectively. They are obtained from atomic orbital analytic wavefunctions such as those tabulated by Clementi & Roetti (1974). They are also the Fourier transforms of the X-ray scattering factors given in _atom_rho_multipole_scat_core and _atom_rho_multipole_scat_valence.

3.5.3.2. Atom-centred multipole expansion functions

Data items in this category are as follows:

- **ATOM_RHO_MULTIPOLe**
  - _atom_rho_multipole_atom_label
  - _atom_site_label
  - _atom_rho_multipole_coeff_P0
  - _atom_rho_multipole_coeff_Pv
  - _atom_rho_multipole_coeff_P00

The ATOM_LOCAL_AXES category is used to define a local Cartesian axis system for each atom. Data items in these categories are looped, with the category key (_atom_local_axes_atom_label) being a pointer to the _atom_site_label defined in the ATOM_SITE category of the core CIF dictionary. The orientation of the Cartesian basis is defined in terms of vectors between atoms, and any arbitrary orientation can be described by including dummy atoms if necessary. The spherical coordinate system used in the description of the multipoles is defined in terms of the local Cartesian basis: \(z\) is the normal axis of the spherical coordinate system, \(\varphi\) is measured from the \(x\) axis in the \(xy\) plane, and \(\theta\) is measured from the \(z\) axis having a value of \(\varphi = +90^\circ\).

Example 3.5.3.1 shows how the category is used in part of an electron density CIF relating to a charge density study of dichloromaleic acid (a ‘proton sponge’) (Mallinson et al., 1997).

Data items in this category are also looped. The category key (_atom_rho_multipole_atom_label) is also a child of _atom_site_label in the ATOM_SITE category, thereby linking the electron density multipole coefficients to a specific atom site in the list of atom coordinates. This category contains all the parameters of the Hansen & Coppens (1978) function

\[
\rho(r) = \rho_{\text{core}}(r) + \rho_{\text{valence}}(r) \\
+ \sum_{l \leq l_{\text{max}}} \sum_{m \leq l} \{P(l,m)d(l,m,\theta,\varphi)\},
\]

where \(r\) is a position vector with respect to the atomic nucleus, and \(\rho_{\text{core}}(r)\) and \(\rho_{\text{valence}}(r)\) are the spherical core and valence electron densities, respectively. They are obtained from atomic orbital analytic wavefunctions such as those tabulated by Clementi & Roetti (1974). They are also the Fourier transforms of the X-ray scattering factors given in _atom_rho_multipole_scat_core and _atom_rho_multipole_scat_valence.

**P\text{r}** is the weight applied to the free-atom core (_atom_rho_multipole_coeff_Pc), **P\text{v}** is the weight applied to the free-atom valence shell (_atom_rho_multipole_coeff_Pv), \(P(0,0)\) is the number of remaining electrons (_atom_rho_multipole_coeff_P00) and \(P + P + P(0,0) = Z\) (the atomic number) for
a neutral atom. \( \kappa \) is the valence electron expansion factor (\( \text{atom}_\rho \text{mult} \text{ipole}_\text{atom}_\text{label} \)), \( R^{(l',m',l',r)} \) is the radial function (Slater or equivalent) (\( \text{atom}_\rho \text{mult} \text{ipole}_\text{radial}_* \)), \( l' \) is the multipole function expansion factor (\( \text{atom}_\rho \text{mult} \text{ipole}_\text{kappa}_\text{prime}_1 \)), \( P_l^m \) are the spherical harmonic coefficients (\( \text{atom}_\rho \text{mult} \text{ipole}_\text{coeff}_P^m[l,m] \) and \( d(l,m,\theta,\varphi) \) is the spherical harmonic of order \( l \), \( m \) at the angle \((\theta,\varphi)\). The summations are performed over the index ranges \( 0 \leq l \leq l_{\text{max}}, -l \leq m \leq l \), where \( l_{\text{max}} \) is the highest order of multipole applied.

Example 3.5.3.2 demonstrates how the category is used in the proton sponge complex of Example 3.5.3.1. Only the first atom is shown in the example.

3.5.4. Development of the dictionary and supporting software

The first implementation of the dictionary appeared in a version of the XD program package (Koritsanszky et al., 2003). Thus XD can read and write CIFs which include data items in the two categories described above (as well as, of course, items in the core CIF dictionary). It is envisaged that future developments to the rhoCIF dictionary will add features that are relevant to other widely used program packages. In the XD implementation, use was made of the CIFbx library of Fortran functions for programming CIF applications (Hall & Bernstein, 1996).

We thank the users of the program package XD (Koritsanszky et al., 2003) who provided feedback during the development of the rhoCIF dictionary and Brian McMahon for helpful comments on its presentation.

References