

解説

# 結晶構造データベースと結晶学共通データ・フォーマット CIF について

## 2. 結晶学情報共通データ・フォーマット CIF

松下能孝\*

(独) 物質・材料研究機構 中核機能部門 材料分析ステーション  
〒305-0047 茨城県つくば市千現 1-2-1

\*Matsushita.Yoshitaka@nims.go.jp

(2014年10月1日受理: 2014年10月16日掲載決定)

結晶学(結晶構造)に関する情報は、各物性の起源を理解するのにも欠かせない手法である。ここに国際結晶学連盟が推奨している結晶学情報共通データ・フォーマット(CIF)に関して紹介を行う。このCIFの普及に伴って各結晶構造データベース中のデータや各種計算プログラムにおける入力パラメーターの共有化が可能と成っている。加えて、CIFを用いる事によって論文中の誤植を減らす事にも役立っている。

## Brief introduction about crystal-structure database and crystallographic data format, CIF

### 2. Crystallographic Information File (CIF)

Yoshitaka Matsushita\*

Research Network and Facility Services Division,  
National Institute for Materials Science (NIMS)  
1-2-1 Sengen, Tsukuba, Ibaraki 305-0047, Japan

\*Matsushita.Yoshitaka@nims.go.jp

(Received: October 1, 2014; Accepted: October 16, 2014)

Crystallographic information including crystal structure is one of the most important knowledge to understand physical and chemical properties of the compounds. Brief information of the "Crystallographic Information File" (CIF) format developed by committee of International Union of Crystallography, is given here. Using this CIF formatted data, we can access more easily to various kinds of databases and also to many programs to calculate properties of the compounds. Moreover, CIF is very effective tool to reduce misprinting on the scientific papers and reports.

#### 1. 序文

前拙文で示しました様に、現在では年間に数万もの結晶構造が解析され、公表されて来ています。[1] これら膨大な数の結晶構造データはその物質の結晶化学的知見を与えるだけではなく、その物質が有す

る物性・特性の起源を解明する為や、より高い物性・特性値を示す物質の開発において、欠かせない情報で有ります。故に、解析された結晶構造データは順次、所定の結晶構造データベースに登録されて行く訳です。結晶構造データのデータベースへの登録は、

自己申告，出版社から所定のデータベースへのデータ転送，またはデータベースを統括している事務局から直接著者への登録依頼などの方法が採られています。この際に主として用いられているデータ形式が本拙文で紹介する結晶学情報共通データ・フォーマット (Crystallographic Information File: CIF)です。[2, 3] 本データ・フォーマットは国際結晶学連盟 (International Union of Crystallography: IUCr) の分科会である Working Party on Crystallographic Information (WPCI) が主導と成って開発されました。本 CIF は単に結晶構造データのみを記録したデータベース登録用のデータ・フォーマットではなく，結晶構造データを得る為に行われた実験・計算条件なども内包出来る上に，今後，新規な記載事項が生じた際も簡単に追加出来る非常に柔軟性・拡張性が高い物と成っています。加えて，文法も非常に簡単なので最低限の結晶学の知識が有れば容易に中身を理解する事が出来るデータ・フォーマットと成っています。以下にその概観を紹介して行きます。

## 2. 開発経緯および変遷

どの分野でも同じだと思いますが，データベース構築創成期において各データベース間で共通のフォーマットを持ちません。勿論，結晶学においても同様でした。この事は各データベースに登録されたデータの共有を困難にするばかりか，その登録されたデータを利用した各プログラム(結晶構造描画，バンド計算，各種物性値計算など)の入出力時に多大なる煩雑さを生じておりました。又，1990年代半ばまでは各データも紙媒体での投稿・出版の後にデータベースに登録されておりましたので，この過程で生じる誤植に依るデータのミスも数多く有りました。

そこで IUCr では，上記の問題を極力防ぐ為に作業部会 WPCI を立ち上げ，結晶学情報の共通フォーマットの作成に掛かりました。そして，1988年にウィーンで開催された第11回欧州結晶学会議において第1回 WPCI 会議が開かれ，大まかな方針が定められました。この際の大きな成果が，結晶学情報における共通データ・フォーマットとして，STAR (Self-defining Text Archive and Retrieval) を採用する事を決めた事でした。[4, 5] この STAR フォーマットの基礎構造はシンプルなテキスト形式であり，定義を予め決め，それを周知しておけば如何様にも拡張性が高くする事が可能で，汎用性も非常に高いものです。加えて，クロス・プラットフォームをも意

識されたフォーマットですので，前述の様に基本，中身はテキスト・フォーマットです。(寧ろ開発当初はクロス・プラットフォームと言う概念では無く，UNIX や大型計算機ベースだったので，バイナリー形式よりもテキスト・ベースの方がメモリー容量は取りましますし，コンピューターの認識速度は落ちますが，人間として解りやすい，扱いやすいフォーマットにしたと言うのが正しいかも知れません。) 又，化学式などで使われる“下付き文字”，“上付き文字”，“ギリシャ文字”などの特殊文字形式の扱いは TeX と似た方式を採っており，O<sub>2</sub> などの下付き文字は“O<sub>2</sub>”，O<sup>2</sup> などの上付き文字は“O<sup>2+</sup>”，α などのギリシャ文字は“ $\alpha$ ”と言った具合です。一方，予め指定された文字列(データ測定方法など)以外の表記は，“”もしくは‘ ’で括る事で文字列として認識されます。

この CIF の基本構文文法の詳細は International Tables for Crystallography Vol. G [6] に纏められています。又，CIF は前述の様に非常に拡張性に飛んだ柔軟なフォーマットを採用したので，現在でも第一原理や ab-initio 計算などの手法を用いて理論予測された結晶構造や，3次元空間(構造)では表記出来ない結晶(変調構造・準結晶など)にも対応すべく絶えず進化を続けています。以下にこの CIF の概観，特に結晶構造を理解する上や，計算科学などに必須なパラメーターを中心に紹介します。

## 3. CIF (Crystallographic Information File)

各々の化合物データは“data\_”から始まり，基本，次の“data\_”が出て来た所までが1つのデータとして認識されます。故に，基本，CIF は1つの化合物(結晶構造)データごとに作成される物ですが，この“data\_”宣言文を挿入する事によって，幾つかの化合物のデータを連結して1つのファイルとする事も可能です。

唯，最近では，CIF 中に後述する様に結晶構造そのもののデータのみならず，実験・解析結果や生データそのものなど色々な情報を全て内包していますので，1つの化合物で数千行のデータと成ります。とすると，人間から見ると1つ1つの化合物データに分けておいた方が解り易く，読み飛ばしなどの間違いが無くて済みます。勿論，コンピューターで自動認識させる場合はその限りでは有りません。

各々の化合物データにおける CIF は基本，以下の構造体(ブロック)から成っています。

- A. データ公表者および出版情報
- B. 基礎結晶学的情報
- C. 実験・解析方法の詳細
- D. 得られた結晶構造データおよび結晶化学的情報
- E. 解析時に使用したパラメーター・ファイル
- F. 解析に用いられたデータ本体

各論に入る前に、CIF 文法上および結晶学上の一般的な約束事を示しておきます。各々の CIF 構文の記述は常に “\_” から始まる形式になっています。以下は格子定数を例に示しています。

```
_cell_length_a          6.3244(7)
```

ここで結晶学における“長さの単位”に関して、記述しておきます。結晶学では、格子定数などの単位として、現在においても非 SI 単位である “Å” を伝統的に使用しています。故に前述の格子定数のみならず、使用 X 線 (含む放射光)・中性子線・電子線の波長、各結合長などは全て Å 単位と成っています。

加えて、各値の標準偏差値の取り扱い法 (表記法) が分野で若干異なるので、ここで結晶学における表記法についても述べておきます。上記の格子定数  $a$  の例で見られる様に結晶学では伝統的に標準偏差値は標準偏差が乗る値の後に ( ) を付けて値を示します。前述の格子定数を例にとると、6.3244(7) は 6.3244±0.0007 という事に成ります。標準偏差値が “±0.00011” の場合は、例えば、略して 6.3244(1) とは書かずに、値の桁数を 1 桁下げて、6.32441(11) 等と標準偏差値自身が内包する偏差値を明確化します。この様に CIF では結晶学のルールに準拠しています。

次に CIF 中における繰り返しデータ列の取扱法を示します。対称操作、原子座標、温度因子など繰り返し同じ構文データ (パラメーター) が続く場合は、最初に “loop\_” で宣言し、その後に繰り返すデータの構文が続きます。各行中の各パラメーターは 1 文字分以上の “空スペース” を入れる事によって区別されます。

```
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_occupancy
```

```
C1 C 0.4312(5) 0.1751(3) 0.3397(2) 0.0190(8) 1.0
C2 C 0.4096(5) 0.0862(3) 0.4065(3) 0.0213(9) 1.0
C3 C 0.4624(5) 0.1328(3) 0.4976(3) 0.0198(8) 1.0
```

上記では、結晶構造解析の結果、C1 サイトは炭素原子が結晶学的に独立なサイト (原子座標)  $x = 0.4312(5)$ ,  $y = 0.1751(3)$ ,  $z = 0.3397(2)$  に位置に存在し、この原子は等方性温度因子 (剛体モデル) 0.0190(8) Å<sup>2</sup> で熱振動している事を示しています。最後の 1.0 はこの位置に 100% の存在確率でこの原子が存在している事を示しています。C2 サイト以降もこの配列に準拠しています。

又、1 行当たりの最大文字数は、初期は Fortran に準拠して 1 バイト・コードで 80 文字とされていましたが、現在は 2048 文字まで拡張されています。[6] 複数行に渡る情報を示す場合、“;” で括って上げます。

```
_publ_section_abstract
;
```

加えて、CIF は前述の構造体の中、もしくは異なる構造体中でも宣言する文法を間違えていなければ、各データが前後しようとも問題は有りません。又、“#” で宣言された行は、コンピューターは読み飛ばす事に成っています。

では、以下に実際のデータを示しながら、CIF の概観を示します。

### 3.1. データ公表者や出版情報

これは言うまでも無く、データを公表した本人の名前、所属先、公表先などの出版情報を記載した物で、最初に data\_global と宣言した後、情報を示します。ここでは大きく以下の 4 つの副構造体で構成されています。

1. Submission details
2. Processing summary (IUCr office use only)
3. Title and author list
4. Text

最初の投稿・出版情報は “\_publ\_” で宣言された文法で各情報を記述します。2 番目は “\_journal\_” で宣言され、国際結晶学連盟 (IUCr) が発刊している学術誌 Acta Crystallography (A, B, C, D, E など) に投稿した際にのみ活用される物で、編集者・出版社側が記述する部分です。通常、我々投稿者は記述す

る必要ありません。3番目は具体的な投稿・出版情報で、これも前述2番目の情報と同様に“\_journal\_”で宣言され、論文タイトル、雑誌名、出版情報(巻号、ページ、出版年など)、各著者の名前、所属先、住所などの情報が続きます。但し、この副構造体はIUCrが発刊する学術雑誌以外の全雑誌に使用されます。最後の副構造体“Text”は、“\_publ\_section\_”で宣言した後、論文要旨のみならず、図・表を除いた論文本体をCIF中に内包させる事も可能です。前述の様にCIFは単なるテキスト・ファイルですので、図・表はJPEGなどの一般的なファイル形式で別途添付する事を義務付けられており、CIF中では各図・表のcaptionのみを記します。CIF本体と図・表との関連付けは、CIF内にファイル名を記載する事によってなされ、CIFを受理した機関(データベース作成機関、出版社など)が責任を持ってCIF本体と図・表を保管する事になっています。唯、この煩雑性は近い将来、何らかの形で解消される筈です。

以下に例として、“1. Submission details”における著者(corresponding author)の部分のみを示します。

```
_publ_contact_author_name
'Yoshitaka Matsushita'
_publ_contact_author_address
;
Research Network and Facility Services Division
National Institute for Materials Science (NIMS)
1-2-1 Sengen, Tsukuba,
Ibaraki 305-0047
Japan
;
_publ_contact_author_phone '81-29-851-3354'
_publ_contact_author_fax ' '
_publ_contact_author_email
'Matsushita.Yoshitaka@nims.go.jp'
_publ_requested_journal ?
_publ_requested_coeditor_name ?
```

現在、Elsevier や Willey などから発刊されている学術誌では、投稿中の論文においても指定された各結晶構造データベース [1] に、CIF形式で結晶構造を登録し、登録番号を得、原稿中にその番号を記す事が義務付けられている雑誌も多くあります。加えて、データのpriorityだけを取る為に、投稿する前に結晶構造を各データベースに登録し、論文公表後にデータベース中の当該CIFに出版情報を追加するという方法を取っている研究者も数多くいます。

### 3.2. 基礎結晶学的情報

ここから具体的な結晶構造情報で、化合物ごとに“\_data\_”で始まるデータが示されます。通常、先頭行にCIFを作成した際に用いたプログラム名を記します。以下の例では、IUCrがほぼ単結晶構造解析用標準プログラムとしているSHELXLの2014/6 release版 [7]を使用してCIF作成をした事を示しています。

```
_audit_creation_method SHELXL-2014/6
```

次に“\_chemical\_”で始まる構文で、国際純正・応用化学連合(IUPAC, International Union of Pure and Applied Chemistry)が規定した命名法による化合物名、化合物の一般的な名称、その物質の融点、分子式、各元素ごとの組成式、分子量と続きます。

```
_chemical_name_systematic
;
?
;
_chemical_name_common ?
_chemical_melting_point ?
_chemical_formula_moiety ?
_chemical_formula_sum
'C16 H12 O4 S10'
_chemical_formula_weight 588.86
```

この構造体中においては、上記の例で示した様に著者が正確な情報を有しない場合もしくは正確な命名が困難な場合、“?”として記述しない事も可能です。

この部分の後には、“\_atom\_type\_”で宣言される各元素の異常分散項(X線および放射光のみ)、“\_space\_group\_”で示される空間群・対称操作情報、“\_cell\_”から始まる格子定数などの基礎結晶学的データが続きます。以下に“\_space\_group\_”および“\_cell\_”の構造体を示します。

```
_space_group_crystal_system monoclinic
_space_group_IT_number 14
_space_group_name_H-M_alt 'P 21/n'
_space_group_name_Hall '-P 2yn'
loop_
_space_group_symop_operation_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 15.311(2)
_cell_length_b 9.1532(14)
_cell_length_c 17.005(2)
```

```
_cell_angle_alpha          90
_cell_angle_beta           113.733(6)
_cell_angle_gamma          90
_cell_volume                2181.6(6)
_cell_formula_units_Z      4
```

この空間群・対称操作情報、格子定数情報と後述の各原子座標の情報があれば、結晶化学的情報（各原子間距離、角度など）の導出、結晶構造の描画、回折パターンのシミュレーションが可能となり、電子密度のバンド計算などの計算科学にも必須のパラメーター群と成ります。

### 3.3. 実験・解析方法の詳細

ここからが実際にデータを収集した際の実験内容およびデータ解析の詳細を示すブロックと成ります。実験的に結晶構造を得るにはX線（含む放射光）、中性子線、電子線などの量子ビームを利用した回折実験が主たる実験手法と成りますので、各々実験条件文が若干異なります。ここでは実験室系単結晶X線回折法を用いた例を示します。

まずは、“\_exptl\_”で始まる構文で、試料の形態、色、大きさ（単位：mm）などの実験に使用した試料自体の情報を示されています。試料の形態や色は文字列ですので、' 'で括ります。

```
_exptl_crystal_description  'Plate'
_exptl_crystal_colour       'Red'
_exptl_crystal_size_max     0.2200
_exptl_crystal_size_mid     0.0700
_exptl_crystal_size_min     0.0200
```

続いては、“\_diffrn\_”から始まるセクションです。このセクションでは、回折データを収集した際の実験装置系の情報、データ解析時のパラメーター・最終的に構造を決定した際のR値などの情報が続きます。以下では、このセクションから一部を抜粋して示します。最初は、データ測定時の温度（単位：K）、使用したX線波長、X線ターゲットの種類、測定に用いた装置を示しています。次は、総観測回折点数、観測した等価回折点のピーク強度の平均R値、観測した回折点hklの最大・最小値（範囲）です。

```
_diffrn_ambient_temperature  110(2)
_diffrn_radiation_wavelength  0.710747
_diffrn_radiation_type       MoKαa
_diffrn_measurement_device
;
Rigaku AFC11 Saturn724+ (2x2 bin mode)
;
```

```
_diffrn_reflns_number        21278
_diffrn_reflns_av_R_equivalents  0.0791
_diffrn_reflns_limit_h_min    -11
_diffrn_reflns_limit_h_max    12
_diffrn_reflns_limit_k_min    -13
_diffrn_reflns_limit_k_max    12
_diffrn_reflns_limit_l_min    -36
_diffrn_reflns_limit_l_max    36
```

次の“\_computing\_”セクションでは、データ測定および構造解析に用いたプログラムを示します。

```
_computing_data_collection    'CrystalClear
(Rigaku Inc., 2008)'
_computing_cell_refinement     'CrystalClear
(Rigaku Inc., 2008)'
_computing_data_reduction      'CrystalClear
(Rigaku Inc., 2008)'
_computing_structure_solution  'SHELXS-2014/6
(Sheldrick, 2014)'
_computing_structure_refinement 'SHELXL-2014/6
(Sheldrick, 2014)'
```

続く、“\_atom\_sites\_solution\_”では、結晶構造モデルを求めた方法を示しています。

```
_atom_sites_solution_primary   direct
_atom_sites_solution_secondary difmap
```

ここでは、直接法を用いて初期結晶構造モデルを求め、続いて差フーリエ法にて直接法では求められなかった残りの原子位置を見出し、最終結晶構造モデルを決めて行った事を示しています。

最後の“\_refine\_ls\_”では、最小二乗法を用いた結晶構造の精密化を行った際の過程・結果を示します。以下では、結晶構造の精密化では、観測構造因子の2乗の $F^2$ を用い、完全行列法で構造の精密化を行い、その際の観測データと計算結果を合わせる際の重み付け（Scale）値を計算で求めた事を示しています。その後は、最終的にどの程度まで精密化が収束したかを示す結果群です。この例では、R値が5.5%まで収束し、その際のGOF(=S)値（期待値：1.00）は0.844であった事を示しています。加えて、最終的な最小二乗法における各パラメーターの最大・平均変動値をも示し、この解析結果では最終的に完全に各パラメーターが収束しきっており、変動をしなかった事を示しています。

```
_refine_ls_structure_factor_coef  Fsqd
_refine_ls_matrix_type             full
_refine_ls_weighting_scheme        calc
```

```
_refine_ls_R_factor_gt      0.0550
_refine_ls_goodness_of_fit_ref  0.844
_refine_ls_shift/su_max      0.000
_refine_ls_shift/su_mean     0.000
```

### 3.4. 得られた結晶構造データおよび結晶化学的情報

ここから具体的な結晶構造情報と成ります。このパラメーターと前述の空間群および格子定数の情報が有れば、結晶構造描画、バンド計算、各種物性値計算などが可能と成る物で、取りあえず結晶構造を知りたいと言う多くのユーザーにとって必須な情報です。本章冒頭で示した物は実験もしくは計算を用いて得られた結果の必須事項のみを示した簡略化した物ですが、以下では現段階での実験結果を反映した完全版を示します。各行の内、解析時に構造モデル（原子座標、温度因子など）に束縛を掛けたパラメーター群にはその旨を記す必要が有るのですが、以下の例では特段、束縛を掛けておりませんので、この様な場合は、全て“.”で処理します。これは完全版ですので、色々とパラメーターが存在しますが、繰り返しに成りますが、結晶構造描画や各計算には前述の簡略化した物だけで十分です。

```
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
  _atom_site_occupancy
  _atom_site_site_symmetry_order
  _atom_site_calc_flag
  _atom_site_refinement_flags_posn
  _atom_site_refinement_flags_adp
  _atom_site_refinement_flags_occupancy
  _atom_site_disorder_assembly
  _atom_site_disorder_group
S1 S 0.55716(10) 0.47925(9) 0.30362(3) 0.0175(2)
Uani 1 1 d . . . . .
S2 S 1.11857(10) 0.58352(9) 0.30094(4) 0.0183(2)
Uani 1 1 d D . . . . .
S3 S 1.13938(10) 0.86633(9) 0.36904(3) 0.0180(2)
Uani 1 1 d . . . . .
```

上記では各原子の熱振動モード（U）は等方性（剛体モデル）でしたが、以下では三行三列の2次の非等方性温度因子（回転楕円体モデル）まで展開させて解析した結果を示しています。各“U\_”後の数字

は三行三列行列中の各独立成分を示しており、単位は等方性同様に Å<sup>2</sup>です。現在の結晶構造解析においては特殊な事情が無い限り、2次の非等方性温度因子までの解析が求められます。

```
loop_
  _atom_site_aniso_label
  _atom_site_aniso_U_11
  _atom_site_aniso_U_22
  _atom_site_aniso_U_33
  _atom_site_aniso_U_23
  _atom_site_aniso_U_13
  _atom_site_aniso_U_12
S1 0.0217(5) 0.0134(4) 0.0170(5) -0.0013(4)
0.0024(4) -0.0010(4)
S2 0.0204(5) 0.0133(4) 0.0205(5) 0.0007(4)
-0.0031(4) 0.0025(4)
S3 0.0189(5) 0.0155(5) 0.0185(5) -0.0020(4)
-0.0021(4) 0.0033(4)
```

これらに続くのが、上記の各原子座標および前述の対称操作、格子定数から計算された各原子間距離（`_geom_bond_`、単位：Å）と角度（`_geom_angle_`、単位：°）のテーブルと成ります。今回の両テーブルでは、基準原子位置（距離では左の原子、角度では真ん中の原子）以外の原子位置の対称操作の情報（`_site_symmetry_2` など）や、第一近接原子間距離など特に重要な距離を有する場合のフラッグ（`_publ_flag_`）はおのおの“.”と“?”として CIF 作成プログラムは出力しています。これは、今回の例が有機物で、結晶の対称性が低く、格子が大きいので、結晶学的に独立な原子座標が非常に多く、対称操作の情報やフラッグを1つ1つ示すと何が重要に成るか煩雑化する為の処理です。

```
loop_
  _geom_bond_atom_site_label_1
  _geom_bond_atom_site_label_2
  _geom_bond_distance
  _geom_bond_site_symmetry_2
  _geom_bond_publ_flag
S1 C11 1.750(3) . ?
S1 C1 1.768(3) . ?

loop_
  _geom_angle_atom_site_label_1
  _geom_angle_atom_site_label_2
  _geom_angle_atom_site_label_3
  _geom_angle
  _geom_angle_site_symmetry_1
  _geom_angle_site_symmetry_3
  _geom_angle_publ_flag
C11 S1 C1 94.53(16) . . ?
```

```
C7 S2 C6 92.77(16) . . ?
```

### 3.5. 解析時に使用したパラメーター・ファイル

単結晶構造解析プログラム・SHELXL のバージョン 20xx 以降などでは、解析時のパラメーター・ファイルや解析に用いた実験データが CIF 中に標準出力されております。一般に実験データは数百～数万回折点と言う非常に膨大な量の回折点データから構成されていますので、CIF 自体のファイル容量は非常に肥大しますが、これらのデータが有れば、誰でもこの構造解析の内容を追試（再解析）することが可能と成ります。

CIF 中に出力されているパラメーター・ファイルは一般的には SHELXL 用データ形式ですが、先の“\_computing\_”セクションで示した他の解析用プログラムでも構いません。今回は例として、SHELXL フォーマットを上げます。最初に“\_shelxl\_version\_number\_”として、もう1度、使用したプログラムのバージョン情報を宣言しておく必要があります。その後、SHELXL の解析結果ファイルである res ファイルが添付されます。この SHELXL の res ファイルには残渣電子密度ピークの位置・高さの情報も入っているので、慣れてくれば、これらの情報だけでも使用された試料の質や解析具合が或程度、判断出来る場合があります。

```
_shelxl_version_number 2014/6
_shelx_res_file
TITL r51305Ba P-1(2) 2013-3-22
CELL 0.71073 8.7604 10.0725 25.8815 96.591
92.220 90.392
```

### 3.6. 解析に用いられたデータ本体

次は解析に用いられた観測回折データと成ります。以下の例では、SHELXL 用にフォーマット変換された回折データを示し、回折指数 *hkl*、観測構造因子 ( $F_o$ )、その誤差 ( $\sigma(F_o)$ )、Scale 値、三行三列の変換マトリックスの順番と成っています。唯、この変換マトリックス項は、生データの結晶軸の変換を行った為に生じた物で、通常、この項は必要ありません。又、ここにおける各観測構造因子は、生データに Lorentz・偏光および経験的吸収補正を行った後の値と成っています。

```
_shelx_hkl_file
;
0 0 -1 0.52 1.11 1
0.96714-0.96825-0.25312 0.24994-0.02362-0.00388
```

```
0 0 1 -0.00 1.29 1
0.97320-0.97210-0.22991 0.23309 0.00137 0.02614
0 0 2 -8.12 14.45 1
0.97580-0.97359-0.21820 0.22455 0.01388 0.04115
0 0 -2 -4.64 6.76 1
0.96367-0.96589-0.26461 0.25825-0.03612-0.01890
```

## 4. CIF 作成, チェック・プログラム

前章までにおいて駆け足ではありましたが、CIF の概観を紹介致しました。唯、結晶学を生業にしている方で有っても膨大なパラメーター・ファイルである CIF の内容を作成・精査する事は非常に困難です。故に CIF の作成、内容のチェックをプログラムに判断させる方が賢明です。CIF の作成は SHELXL のみならず、市販・無料問わず現在、使用されている数多くの結晶構造解析プログラム本体および構造解析パッケージ・プログラムにおいて、結果の出力時に同時に（もしくは付属する CIF 作成専用プログラムを利用して）CIF を出力することが可能です。加えて、結晶構造描画プログラムでもファイル・フォーマットを変換し、CIF として出力することが出来ます。勿論、これら結晶構造描画プログラムから出力された CIF には実験・解析上のデータ（測定・解析条件）は含んでおりません。

以下に CIF の出力が可能な代表的なプログラム名とそれらの HP 情報を順不同で列記しておきます。

(構造解析プログラム)

#### A. SHELXL

<http://shelx.uni-ac.gwdg.de/SHELXL/>

#### B. JANA

<http://jana.fzu.cz/>

#### C. RIETAN

<http://fujioizumi.verse.jp/download/download.html>

#### D. FullProf

<http://www-llb.cea.fr/fullweb/powder.htm>

#### E. GSAS

<http://www.ncnr.nist.gov/programs/crystallography/software/gsas.html>

(結晶構造描画プログラム)

#### F. VESTA

<http://jp-minerals.org/vesta/jp/>

#### G. CrystalMaker

<http://www.crystalmaker.com/>

## H. Diamond

<http://www.crystalimpact.com/diamond/>

これらのプログラムで出力された CIF は CIF チェック専用の HP もしくはプログラムを利用して CIF 構文の誤りのみならず、実験上ならびに解析上、誤りや不十分がないかをも同時にチェックします。各チェック内容の規定は IUCr で厳格に規定されています。我々ユーザーは、これらで CIF の内容をチェックした後、そのチェック結果ファイルを添付した上で、多くのデータベースへの登録時、出版社への投稿することが義務付けられています。以下に代表的なプログラムを紹介致します。

## A. CheckCIF

<http://checkcif.iucr.org/>

<http://journals.iucr.org/services/cif/checking/checkfull.html>

本プログラムは IUCr が作成・公開している CIF チェック・プログラムで、Web を介してターゲット CIF をアップロードし、内容をチェックさせます。CIF 構文・規定を決定している機関である IUCr が作成・公開しているが故に常に最新の CIF 構文および規定に準拠しています。加えて、本プログラムは IUCr が出版している各種学術誌 (Acta Crystallogr. シリーズ, J. Appl. Crystal., J. Synchrotron Rad. など) のみならず、米国、英国、日本と言った主要化学会や Elsevier や Wiley と言った主要学術出版社からもサポートされている点が特徴です。

## checkCIF

A service of the  
International Union of Crystallography

checkCIF reports on the consistency and integrity of crystal structure determinations reported in CIF format.

Please upload your CIF using the form below.

File name:

選択されていません

Select form of checkCIF report

HTML  PDF

Select validation type

Full validation of CIF and structure factors  
 Validation of CIF only (no structure factors)

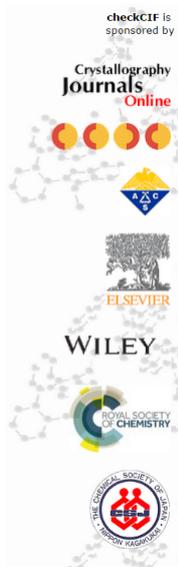
Output Validation Response Form

Level A alerts only  
 Level A and B alerts  
 Level A, B and C alerts  
 None

Information about this version of checkCIF ...

## Useful links

Prepublication check for submissions to IUCr Journals  
Details of checkCIF/PLATON tests  
CIF dictionary  
Download CIF editor (pubCIF) from the IUCr  
Download CIF editor (enCIFer) from the CCDC



上記にサイトからチェックを行いたい CIF を送付すると自動的に次ページの様なチェック結果が非常に簡単な結晶 (分子) 構造図をも添付した形で帰って参ります。この出力形式は上記の HP で HTML もしくは PDF を選択する形に成ります。このファイルを前述の如く、論文投稿時に添付します。

ここでそのチェック結果の重要な点を幾つか紹介します。CIF 内の各パラメーター値を検討し、実験上もしくは解析上、明確な問題がある場合、その重大さから Alert level A から C の3段階、およびタイプ・ミスなどの一般的な問題 Alert level G の計4つの Alert level が存在します。Alert level A が提示された場合、その実験もしくは解析上重大なミスが有りますので、再実験もしくは再解析が必要と成り、一般にはその論文の投稿 (結晶構造解析関係の部分) も認められません。Alert level B の場合は Alert level A よりも事の重大さは無いが、何らかの問題の可能性が有るので、実験・解析共に再検討をする必要が有ります。この Alert level の場合、この CIF チェック時に指摘された項目に関して明確な回答がなされれば、その論文の投稿は認められます。次のレベルである Alert level C は軽微なミスなどの“可能性”があるレベルで、単に注意を喚起しているだけです。次ページに示す例で提示されているアラートの内容は2つです。最初の物は、測定したデータの  $\theta$  レンジが 97.5%とちょっと少ないですよと言う喚起で、もう1つは C-C 結合長の精度がちょっと悪いかも知れませんよと言う内容を示しています。この例の場合、試料の結晶性が単結晶回折実験用には完全にベストな物で無く、高角のデータが充分に測定出来なかったと言う点と、各原子の熱振動を抑える (熱雑音を減らす) ために回折データを測定した温度が 110 K と室温では無かったので、室温データを基にしている C-C 平均結合長よりもわずかに短い距離を報告しているので注意を喚起されています。唯、この様にこれらアラートに明確に原因を回答出来るので、仮にアラートが出ても何ら問題無くデータ・論文の公表が出来ます。

## checkCIF/PLATON report

You have not supplied any structure factors. As a result the full set of tests cannot be run.

THIS REPORT IS FOR GUIDANCE ONLY. IF USED AS PART OF A REVIEW PROCEDURE FOR PUBLICATION, IT SHOULD NOT REPLACE THE EXPERTISE OF AN EXPERIENCED CRYSTALLOGRAPHIC REFEREE.

No syntax errors found CIF dictionary Interpreting this report

## Datablock: shelx

Bond precision:	C-C = 0.0063 Å	Wavelength=1.54178
Cell:	a=8.814 (5)    b=6.263 (4)    c=26.385 (11)	alpha=90    beta=95.862 (11)    gamma=90
Temperature:	163 K	
Volume	Calculated 1448.9 (14)	Reported 1449.0 (14)
Space group	P 21/n	P 21/n
Hall group	-P 2yn	-P 2yn
Moiety formula	C7 H3 O2 S4, C6 H8 N	C7 H3 O2 S4, C6 H8 N
Sum formula	C13 H11 N O2 S4	C13 H11 N O2 S4
Mr	341.51	341.47
Dx, g cm <sup>-3</sup>	1.566	1.565
Z	4	4
Mu (mm <sup>-1</sup> )	6.029	6.028
F000	704.0	704.0
F000'	710.78	
h,k,lmax	10, 7, 31	10, 7, 31
Nref	2649	2579
Tmin, Tmax	0.731, 0.835	0.716, 1.000
Tmin'	0.588	
Correction method=	MULTI-SCAN	
Data completeness=	0.974	Theta(max)= 68.192
R(reflections)=	0.0550 ( 1860)	wR2(reflections)= 0.1347 ( 2579)
S =	0.997	Npar= 182
The following ALERTS were generated. Each ALERT has the format test-name ALERT alert-type alert-level. Click on the hyperlinks for more details of the test.		

<b>Alert level C</b>		
PLAT029_ALERT_3_C	diff measured fraction theta full Low	0.975
PLAT340_ALERT_3_C	Low Bond Precision on C-C Bonds	0.0063 Ang.
<b>Alert level G</b>		
PLAT005_ALERT_5_G	No _iucr_refine_instructions_details in the CIF	Please Do!
PLAT007_ALERT_5_G	Number of Unrefined Donor-H Atoms	3 Why?
PLAT720_ALERT_4_G	Number of Unusual/Non-Standard Labels	3
<b>Alert level A</b>	Most likely a serious problem - resolve or explain	
<b>Alert level B</b>	A potentially serious problem, consider carefully	
<b>Alert level C</b>	Check. Ensure it is not caused by an omission or oversight	
<b>Alert level G</b>	General information/check it is not something unexpected	
<b>Alert type 1</b>	CIF construction/syntax error, inconsistent or missing data	
<b>Alert type 2</b>	Indicator that the structure model may be wrong or deficient	
<b>Alert type 3</b>	Indicator that the structure quality may be low	
<b>Alert type 4</b>	Improvement, methodology, query or suggestion	
<b>Alert type 5</b>	Informative message, check	

It is advisable to attempt to resolve as many as possible of the alerts in all categories. Often the minor alerts point to easily fixed oversights, errors and omissions in your CIF or refinement strategy, so attention to these fine details can be worthwhile. In order to resolve some of the more serious problems it may be necessary to carry out additional measurements or structure refinements. However, the purpose of your study may justify the reported deviations and the more serious of these should normally be commented upon in the discussion or experimental section of a paper or in the "special details" fields of the CIF. checkCIF was carefully designed to identify outliers and unusual parameters, but every test has its limitations and alerts that are not important in a particular case may appear. Conversely, the absence of alerts does not guarantee there are no aspects of the results needing attention. It is up to the individual to critically assess their own results and, if necessary, seek expert advice.

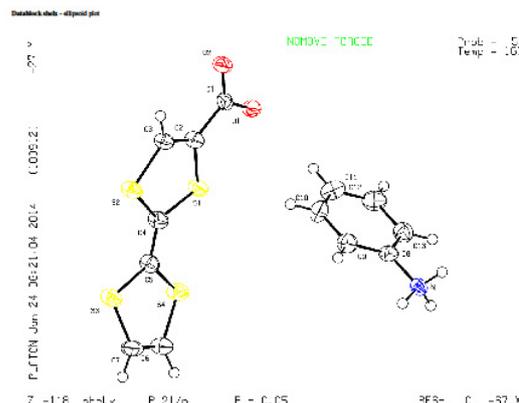
## Publication of your CIF in IUCr journals

A basic structural check has been run on your CIF. These basic checks will be run on all CIFs submitted for publication in IUCr journals (*Acta Crystallographica*, *Journal of Applied Crystallography*, *Journal of Synchrotron Radiation*); however, if you intend to submit to *Acta Crystallographica Section C* or *E*, you should make sure that full publication checks are run on the final version of your CIF prior to submission.

## Publication of your CIF in other journals

Please refer to the *Notes for Authors* of the relevant journal for any special instructions relating to CIF submission.

PLATON version of 18/09/2013; check.def file version of 12/09/2013



## B. EnCIFer

<http://www.ccdc.cam.ac.uk/Solutions/FreeSoftware/Pages/EnCIFer.aspx>

このプログラムは Cambridge Crystallographic Data Centre (CCDC)が作成し、無料で配布しているプログラムで、CIF の作成、内容のチェックも可能です。又、MacOSX, Windows, Linux と行った主要な OS 上で動作させる事が出来ます。

## 5. おわりに

本稿では非常に駆け足で CIF の概観ならびに設計思想を示して来ました。序文でも示しましたが、今後も CIF は更なる進化を遂げ、その対象を単なる 3 次元結晶から準結晶などの高次の対称性を有する結晶にも拡張するでしょうし、さらには非晶質物質も対象とすると思われます。唯、重要なのは基本的な

CIF の設計思想を変えない事です。この設計思想が前もって確立していれば、仮に対象物質の幅を拡げたとしても基本設計部からの枝葉を拡張するだけで対応が可能と成って来ます。貴表面分析研究会でも電子分光やイオン分光・データに関する基本データ・フォーマットの確立やデータベース構築に関心を持っておられる方が多くいらっしゃると思います。特にフォーマットを確立される場合、他の異分野の方の意見・動向を見極め、独りよがりやその場限りに成る様な物では無く、世界中の方々が幅広く使用される物を作成して下さる事を切に願います。

## 6. 謝辞

今年 (2014) は前拙文 [1] でも示しました様に、100 年前の 1914 年に von Laue 教授が X 線回折現象発見の功績によってノーベル物理学賞を受賞され、翌年の 1915 年には Bragg 父子が “Bragg の式” の導出と X 線による結晶構造解析に関する研究によって同じくノーベル物理学賞を受賞されており、これらの功績があった上で現代の回折結晶学が成り立っております。そこで本年 (2014 年) は、国際結晶学連合 (IUCr)、ユネスコ (UNESCO) と国際科学会議 (ICSU) の支援および国際連合の承認のもと、Laue 教授および Bragg 父子の業績の 100 周年を記念するため、2014 年を世界結晶年 (IYCr2014) として制定されており、結晶学に関係する各国際・国内学会において数多くの記念行事がなされ、世界各国において記念切手の発行もなされております。この様な結晶学を記念する貴重な年に本拙文を発行して下さる様に取り計らって下さった (独) 物質・材料研究機構の吉川英樹博士に心より感謝致します。(独) 物質・材料研究機構の道上勇一博士には非周期系物質に関する CIF の詳細に関しまして、御教授頂きました。心より感謝致します。

## 7. 参考文献

- [1] 松下能孝, *J. Surface Analysis*, **19**, 177 (2013).
- [2] S. R. Hall, F. H. Allen, and I. D. Brown, *Acta Crystallogr.*, **A47**, 655 (1991).
- [3] I. D. Brown and B. McMahon, *Acta Crystallogr.*, **B58**, 317 (2002).
- [4] S. R. Hall, *J. Chem. Inf. Comput. Sci.*, **31**, 326 (1991).
- [5] S. R. Hall and N. Spadaccini, *J. Chem. Inf. Comput. Sci.*, **34**, 505 (1994).

[6] “*International Tables for Crystallography, Vol. G: Definition and Exchange of Crystallographic Data*” ed. by S. Hall and B. McMahon, Springer, Dordrecht, the Netherland (2005).

[7] G. M. Sheldrick, *Acta Crystallogr.*, **A64**, 112 (2008).

## 査読コメント

査読者 1. 吉原一紘 (オミクロンナノテクノロジー ジャパン)

結晶学の分野で使用されているデータ・フォーマットに関して、必要十分な内容が丁寧に解説されており、このままで掲載することを薦めます。但し、解説の内容には直接は関係しないのですが、同様のシステムを表面分析に取り入れようとするときの参考のために、可能ならば、以下のような観点からのご意見を記載いただけたら、幸いです。

## [査読者 1-1]

結晶構造データを論文に発表する場合には、雑誌によっては必ず CIF で記述したデータを付けなくてはならないとされていますが、投稿者は CIF を理解していることが前提でしょうか。通常の解析ソフトには CIF への変換ソフトは組み込まれているようですが、中にはそれらを使わない研究者もおられるのではないのでしょうか。

## [著者]

現在、基本的に主要なるデータ測定および解析プログラムにおいて半自動的に CIF を出力します。ここで“半自動的に”と表現するのは、プログラムによってはユーザーが CIF 出力命令をプログラムに送る必要があるからです。

この様に CIF を自動的に生成させる事は可能な訳ですが、その一方、CIF のフォーマットや中身を理解出来ない方は、その出力ファイルを一方的に信じる、もしくは詳しい方の助力を得て、改訂を行う必要があります。特に Alert 対策には詳しい方の助力が絶対に必要と成ってきます。もし、周りにそのような方が居らっしゃらない場合、最悪、そのデータの公表は CIF の添付を要しない雑誌に投稿される事と成ると思われれます。

唯、データベース収録において、単に基本的な結晶学的なパラメーターだけ少なくとも登録するという面も有りますので、CIF の添付を要しない雑誌に投稿された結果はデータベース作成者があらゆる雑誌を見て、データの収録を行っています。

次に、CIF の自力作成の件ですが、詳細に熟知していられる方の場合、自力で CIF を完成させる事も可能ですが、実験～解析結果まで不足無く CIF を作るのは非常に面倒かつ大変な作業ですので、或る程度以上に CIF を熟知している我々も前述の如く、データ測定および解析プログラムにおいて半自動的に CIF を出力させ、それを雛形にして、改訂すべき点は自力で行っています。

[査読者 1-2]

結晶構造の解析では、Alert Level によりデータの信頼性が検証できるようですが、結晶構造データベースにとっては、結晶構造の情報だけではなく、試料情報も重要だと思います。しかし、CIF では不明な場合には「?」と記述することが許されていますので、試料情報の信頼性に問題があるものが含まれる可能性があります。データベースへの登録の可否の判定はどのようになされるのでしょうか。

[著者]

データベースへの登録の可否は、最終 CIF チェックにおいて、Alert A が存在している場合、基本、登録&原稿の投稿は認められません。B 以下の Alert を出しているデータに関しては、投稿者がその Alert を消せない理由に対する明確な回答を示し、且つ査読者が納得した場合、そのデータは受理されますことと成っています。

又、試料そのものの情報（色、形態など）は付帯情報ですので、今の所、余り重視されておりません。故に、現状では“?”でも構いません。唯、実は色や形態はその試料の本質（物性）と密接な関係を有している場合が多いので、今後、これらの情報も重視されるかも知れません。

（例）色=>バンド構造、形態=>結晶系・構造

査読者 2. 渡部大介（アルバック・ファイ）

本投稿原稿は結晶学においてデータベース間の共通データフォーマットの概観・設計思想が分かりやすく紹介されており、CIF を利用する者にとってはもちろんのこと、電子分光におけるデータベースや共通フォーマットを検討する場合においても実用的で価値が高いと考えます。

[査読者 2-1]

第 3 章 3-A.項で“CIF 中では各図・表の caption のみを記します”とありますが、別途添付されている

ファイルと CIF 中の図・表との関連づけはどのように行っているのでしょうか？

この点をご説明いただけると、より CIF に関して理解が進むと考えます。

[著者]

御示唆の程、ありがとうございます。確かに御指摘の通りですので、本件に関しまして、4 ページ目に加筆致しました。

[査読者 2-2]

第 3 章 3-E.項で SHELXL の res ファイルの情報を記載する例が記述されていますが、TITLE r51305Ba ... の行の前行に";"を付けて複数行であることを示す必要はないでしょうか？

[著者]

本行が 1 バイト・コードで 2048 文字を超える場合、";"を利用して複数行である事を示します。一方、本例の場合、同文字数を超えませんが、敢えて、";"を入れる必要は有りません。