CIF を出発点とする第一原理計算支援用ユーティリティー

A Utility to Assist First-Principles Calculations Starting from CIFs

Key-words：CIF, File converter, Shell script, First-principles calculations

泉　富士夫・宮崎　晃平

Fujio IZUMI[*1,*2] and Kohei MIYAZAKI[*2]

([*1]Japan Fine Ceramics Center, [*2]Kyoto University)

1. CIF とは何か

Crystallographic Information File（CIF）は X 線・中性子回折データの解析で得た結晶構造データを一定のフォーマットに従いテキストファイルとして保存するために，国際結晶学連合（International Union of Crystallography: IUCr）の主導により制定された．CIF の主要定義を編纂した Core dictionary[1]（coreCIF）はネットからも入手できる．CIF は結晶データの共通フォーマットとして確固たる地位を占めており，結晶構造解析プログラムはたいてい CIF を保存する機能をもつ．構造解析結果を含む論文を投稿する際，CIF の提出を義務づけている学術誌も多い．

各種データベースで検索した化合物データの内，結晶構造データを CIF として保存すれば，種々の科学技術計算プログラムで直接入力し，処理・利用できる．また CIF を他フォーマットのファイルに変換するコンバーターも多い．Open Babel[2] はとりわけ有名である．

2. CIF を巡る混乱と無秩序

プログラマーとしての観点からは，厳密な仕様に従って記述されたテキストファイルである CIF は互換性が申し分なく，入出力ファイルとして扱いやすいように見える．ところが実際に手に入る CIF は不完全なものが多い．単純なミスに留まらないところがなんとも不条理である．

これまで幾度となく遭遇し，憤慨したのは，結晶構造を収録したデータベースや CIF をデータ源とする科学技術計算ソフトが CIF の仕様を遵守していないことだった．たとえば“_symmetry_…”という廃絶された定義に執着し続けているデータベース（ICSD, PDF-4+，COD，Materials Project，MatNavi 等）が多数派を占めている．遺憾ながら，材料設計支援統合システムを謳う某商用ソフトや DFT コードの入力ファイル作成用補助ツールは三次元可視化プログラム VESTA[3] が出力した CIF をことごとく門前払いする．なんと由緒正しい定義“_space_group_…”，たとえば“_space_group_name_H-M_alt”を拒絶するのである．Open Babel は排除こそしないものの，“_symmetry_equiv_pos_as_xyz”を出力する．そこで，CIF 中の“_space_group_…”を相当する“_symmetry_…”に置換するマクロを作ったこともあった．また PDF-4+ と Materials Project から得た CIF では，Hermann-Mauguin の空間群記号にスペースが入っていない．Bond_Str[4] は二重映進面 e を含む空間群（直方晶系）の正式な Hermann-Mauguin 名を誤記と見なす．

こういう無秩序をもたらした個々の原因は知る由もないが，データベースの構築組織やアプリケーションの製作者が CoreCIF に明記されているルールを理解し遵守すべきなのは言うまでもない．cifconv はこのような嘆かわしい現状を十分把握し，十分な対策を講じた上で開発した．

3. 自作ソフトにおける CIF 活用の歴史

筆者（泉）はこれまで CIF を活用するユーティリティーを次々に開発してきた．RIETAN-FP[5] の入力ファイル中に CIF の結晶データを導入する cif2ins，リートベルト法で求めた結晶データから CIF を作成する lst2cif が代表例である．VESTA[3] が CIF を入出力できるのは言うまでもない．さらに(a)CIF のコンテンツ，(b)結晶構造図，(c)Dysnomia[6] により最大エントロピー法で決定した電子密度の等値曲面イメージ，(d)gnuplot でプロットしたグラフを合併して LaTeX 文書化するユーティリティー cif2pdf（英文）と E2J（和文）も作成した．(b)と(c)はもちろん VESTA で視覚化する．VESTA と Dysnomia を中核とする VENUS システムは結晶構造解析と電子状態計算とをつなぐ架け橋として貢献することを願って開発した．これらの自作・外部プログラムを秀丸エディタ（Windows）または Jedit Ω（macOS）から起動し，入出力ファイルを編集・閲覧できるようにしたのが RIETAN-FP・VENUS 統合支援環境[7)~9)] である．

4. CIF コンバーター cifconv

最近は，CIF を種々のテキストファイル，すなわち

1. サイト・ポテンシャルとマーデルング・エネルギー（MADEL）

© 日本セラミックス協会

2. 原子間距離，結合角，bond valence sum（Bond_Str[4]）
3. VASP[10] や Quantum ESPRESSO 等の第一原理計算ソフト用の入力データ（cif2cell[11]，C-Tools[12]）
4. 固溶・部分占有サイトを含む不定比化合物をモデル化した誘導構造の CIF（supercell[13]）
5. RIETAN-FP[5] による多相試料のリートベルト解析用の入力データ

を記録したファイルに変換するためのシェルスクリプト cifconv.command（以下，cifconv と略す）を鋭意開発してきた．ただし（　）内はファイル出力に使う外部アプリケーションを示す．上記のような多目的変換ツールである cifconv は，多角的な視点から結晶構造を理解・評価するのに有効な上，構造解析と電子状態計算にすこぶる役立つ．cifconv を同梱した RIETAN-FP・VENUS システムは同システム配布ファイルの Web ページ[14] で無償公開している．

cifconv は BusyBox に包含されている ash（Windows）または bash（macOS）で書いた．macOS 版はわずかな変更で Linux へ移植できる．cifconv は千数百行に達しているが，十分な数の注釈を散りばめている上，20 を超える関数からなるモジュール構造をもつため，処理手続きの見通しが良く，アルゴリズムを理解しやすい．UNIX コマンドによる古典的テキスト処理を採用した理由は，さほどメモリーを消費しない上，パイプを頻用した並列処理と慣れ親しんだストリームエディター sed の徹底利用が図れるためである．

cifconv では，同一フォルダーに cifconv と複数の CIF を置いて実行するバッチ処理，支援環境上で一つの CIF に対し cifconv マクロを実行する単一処理のどちらかを選択できる．cifconv は変則的な Hermann-Mauguin 記号を修正するだけでなく，Hall 記号の行を挿入することにより CIF を確実に読み込むよう努めている．

以下，cifconv の第一原理計算関連機能に焦点を絞って解説するが，その全貌については英文マニュアルをお読みいただきたい．

5. VASP 用入力ファイルの作成

POSCAR と KPOINTS の作成は cif2cell[11] が受け持つ．デフォールトのオプションは“--setup-all --vasp-format=5 --vasp-cartesian-lattice-vectors”だが，CIF の先頭に注釈行として“#VASP --...”という形の注釈行を置けば cif2cell 用オプションを追加できる．POTCAR は基本的には cifconv 中のコードによって出力される．すなわち，550 eV のカットオフ・エネルギーに対応する擬ポテンシャル・ファイルを記録した PAW_potentials.txt から POTCAR のフォルダー名を入力した後，POTCAR に連結出力する．擬ポテンシャルデータを収めたファイルを直接指定するには，CIF の冒頭に“#VASP pbe Ba_sv O Ti_sv”というように各元素の POTCAR が入っているフォルダーの名前を入力すればよい．いずれにせよ，上記の機能は VASP の正規ユーザーだけが享受できる．

cif2cell が CIF から作成する INCAR はごく短く，実用からは程遠い．そこで独自の文法に則って雛形ファイル INCAR.ins を記述するという仕組みを導入した．INCAR.ins を CIF と同じ階層に置き，cifconv をバッチモードまたは単一処理モードで実行すると，INCAR が生成する．

INCAR.ins → INCAR 変換は RIETAN-FP という伝家の宝刀を抜くことにより，あっさり実現した．RIETAN-FP が Tink と呼ばれる入力ファイル *.ins のプリプロセッサーを含んでいるためである．具体的には，If・Select 構文の入力制御用変数の入力，If 構文，Select 構文を INCAR.ins 中に記述することにより VASP 用データの入力あるいはスキップを制御する（**図 1**）．

If・Select 構文中のブロックを入力・スキップするか否かを制御するための変数は I，J，K，L，M，N で始まり，大文字，数字，‘@’からなる名前をもち，

変数名＝整数：注釈

変数名＝整数！注釈

という形式で入力する（図 1a）．前者では変数名と値を前処理に使うのに対し，後者の場合，行全体が注釈と見なされる．図 1a の場合，計算モードは最適化と

a. 構文制御用変数の入力

```
MODE = 1! Intrinsic and total energies
MODE = 2! Band diagram
MODE = 3! Density of states
MODE = 4: Optimization
```

b. If 構文

```
If MODE = 1 then
  ISMEAR = -5
else if MODE = 2 then
  ICHARG = 11
  ISMEAR = 0
else if MODE = 3 then
  ICHARG = 11
  ISMEAR = -5
else if MODE = 4 then
  ISMEAR = 0
  NSW = 100
  IBRION = 2
end if
```

c. Select 構文

```
Select case MODE
case 1
  ISMEAR = -5
case 2
  ICHARG = 11
  ISMEAR = 0
case 3
  ICHARG = 11
  ISMEAR = -5
case 4
  ISMEAR = 0
  NSW = 100
  IBRION = 2
end select
```

図 1　INCAR.ins を INCAR に変換するのに使う計算モード関連部分の例．

なる．

MODEの値に応じて場合分けするIf構文を図1bに例示した．関係式には＝，＞，＞＝，＜，＜＝，＜＞（≠）を含めることができる．二つの関係式は“If MODE ＞ 0 and MODE ＜ 4 then”というように論理演算子andまたはorで並べられる．字下げ行の先頭にVASPのタグ[15]が置かれている．上記のIf構文に相当するSelect構文が図1cである．

Ifブロックでelse，Select構文でcase defaultが使えるのは言うまでもない．If・Select構文は一段だけ入れ子にできる．ただし，内側の構文は一つ以上字下げしなければならない．

Tinkは複雑な手続きによりテキストを処理するが，長年にわたり使い込んできた結果，枯れ切っている．わずかな変更により独創的なcharacter user interfaceを再利用できたのは大きな喜びだった．構造解析プログラムをINCARの作成に使うのは奇手といって過言でないが，簡便性と使い勝手の良さを兼ね備えた機能を実現できたことを誇りに思う．

6．他の第一原理計算プログラムの実行

VASPを購入する資力を持ち合わせていない人達のために，無料ソフトであるQuantum ESPRESSO，OpenMX，xTAPP，RSDFTの入力ファイルも作成できるようにした．具体的にはC-Tools[12]のGUIを経由せずに，VASP用入力ファイルINCAR等をc-toolsコマンドで直接変換する．C-ToolsはCIFの入力が苦手な反面，安定かつ高速なファイル変換ツールとして利用価値が高い．上記4つのDFT計算プログラム用の入力ファイルはかなり多くのデータを含む上，C-ToolsのUDF setupダイアログで手軽に変更できるので，第一原理計算の出発点として十分使える．xTAPPの場合は，入力支援・可視化ツールTAPIOCAの併用も推奨したい．

7．Supercellの活用

VASPは基本的に占有率が1のサイトだけからなる化合物を扱う．固溶体や一部のサイトが欠損した不定比化合物の場合，$\boldsymbol{a}$，$\boldsymbol{b}$，$\boldsymbol{c}$方向に整数倍の周期をもち，全サイトの占有率が1であり，なおかつ電気的中性が保たれている誘導構造（derivative structure）を対象とする近似計算が有効である．最新の誘導構造導出プログラムsupercell[13]は入出力ファイルをCIFに統一しているため，cifconvとの相性が良い．CIFの入出力はOpen Babel[2]に任せており，安定している．

cifconvはhoge.cifという名前のCIFからhoge_scというフォルダー中にsupercellによる計算用の同名ファイルhoge.cifを生成する．その先頭付近にたとえば“#supercell 1x1x2”という$\boldsymbol{a}$，$\boldsymbol{b}$，$\boldsymbol{c}$方向に沿った単位胞の長さを指定する注釈行を挿入した後，支援環境からsupercellマクロを実行すると，サブフォルダー1x1x2中に誘導構造のCIFがsupercellコマンドの引数−n lNで指定した数Nだけ出力される．さらにN個の構造のマーデルング・エネルギーを昇順に記録したファイルも生成するが，これは構造モデル候補のスクリーニングに使える．比較的安定な誘導構造のCIFをcifconvで処理すれば，VASPを始めとするDFT計算プログラム用の入力ファイルが得られる．

cifconvはMADEL用入力ファイルhoge.pmeを作成できるようにsupercellの出力した酸化数に関する部分を整形するとともに，RIETAN-FPやVESTAで入力できるようサイト名を変更する機能も備えている．たとえば，supercellの出力したCIFからRIETAN-FPで粉末回折パターンをグラフ化して実測パターンと比較し，VESTAで結晶構造を表示することができる．

8．正極材料への応用

最後に，cifconvを活用してVASP用入力ファイルを作成する手続きを略述する．まず，全サイトの占有率が1である$LiCoO_2$を取り上げる．ICSDのデータベースから出力したCIF[16]（LiCoO2.cif）の中身を**表1**に抜粋した．三方晶系に属するが，六方格子の軸設定を採用している．雛形ファイルINCAR.insを開き，VASPの計算に必要なカットオフ・エネルギー（ENCUT），対称性利用の有無（ISYM），スピン分極の有無（ISPIN）等のパラメーターを入力する．不要なタグは#を付けて注釈にする．上述のように，計算モード（変数MODE）として①エネルギー一点計算，②バンド計算，③DOS計算，④構造最適化を数字で指定する．もち

表1　結晶構造に関する情報

	$LiCoO_2$	$LiNi_{1/3}Mn_{1/3}Co_{1/3}O_2$
格子定数 a/Å	2.8161(5)	2.860(2)
格子定数 c/Å	14.0536(5)	14.227(8)
空間群	$R\bar{3}m$ (No. 166)	$R\bar{3}m$ (No. 166)
サイト名 (酸化数) 分率座標 x，y，z 占有率	Li1（+1): 0 0 0.5 1.0 Co1（+3): 0 0 0 1.0 O1（−2): 0 0 0.2592 1.0	Li1（+1): 0 0 0 0.975 Co1（+3): 0 0 0.5 0.33333 Ni1（+2): 0 0 0.5 0.309 Mn1（+4): 0 0 0.5 0.33333 O1（−2): 0 0 0.2411 1.0 Li2（+1): 0 0 0.5 0.025 Ni2（+2): 0 0 0 0.025

ろんMODEはINCAR.ins内でのみ有効で，INCARには残らない．

変数ILDAUを1に設定すると，自己相互作用の補正法DFT+Uが使える．磁気モーメントの初期値（MAGMOM），+U補正を加える電子軌道の方位量子数（LDAUL），*U*パラメーター（LDAUU），*J*パラメーター（LDAUJ）を各元素に対して与える．その際には，格子定数と原子位置の初期値を与えるPOSCARと擬ポテンシャル・データを記録したPOTCARの場合と同じ元素順で入力しなければならない．しかし，cif2cellがどういう元素順で両ファイルを作成するのかを事前に知るのは無理である．そこで，試しにcifconvを実行してコンソールに出力された元素名の順番を参照し，INCAR.insに反映させるとよい．

LiCoO2.cifとINCAR.insを同じフォルダーに置いてcifconvを実行すると，コンソール出力から元素の順番がCo，O，Liであることがわかる．INCAR.ins中のMAGMOM等ではその順にデータを入力する．引き続きcifconvを再実行すると，LiCoO2_VASPフォルダー中のVASP用入力ファイルINCAR，POSCAR，POTCAR，KPOINTSが更新される．

次の試験には，固溶体の例として三元系正極材料$LiNi_{1/3}Mn_{1/3}Co_{1/3}O_2$（NMC）を選んだ．$LiCoO_2$と同じ空間群に属するが，3*b*サイトをNi，Mn，Coが1/3ずつ占める．また，Li^+とNi^{2+}のイオン半径が互いに近いため，LiとNiが少し混ざり合った構造を持つ．ICSDデーターベースから出力したCIF[17]（NMC.cif）における主要データは表1の通りである．NMC.cifの冒頭に“#supercell 2x2x1”と追記してからcifconvで処理すると，NMC_scフォルダー中に同一名のNMC.cifが生成する．それを秀丸エディタかJedit Ωで開き，supercellマクロを実行すると，2x2x1フォルダーが生成し，誘導構造のCIFファイルが20個出力される．そのフォルダーにINCAR.insをコピーし，supercellが出力したCIFファイルの一つを支援環境で表示させてからcifconvマクロを実行し，上述のように元素の順序を調べる．NMCの場合，Ni，Mn，Co，O，Liの順となる．INCAR.ins中の元素データをその順に変更してからcifconvをバッチモードで実行すると，20個のCIFファイルごとにVASP用入力ファイル4つがサブフォルダーに生成する．POSCARに記録されたNi，Mn，Co原子の座標からVESTAで描画した二つのイメージを**図2**に示す．

引き続き，マーデルング・エネルギーを判定基準とするスクリーニングにより有力候補を選び，VASPで計算する．今回は最も安定な構造で収束を確認した．

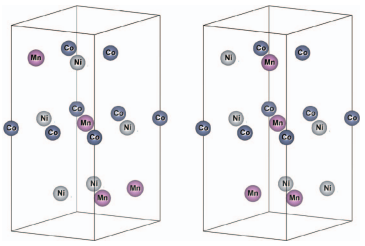

図2　POSCARに出力された誘導構造のイメージ（LiとOは省略）．左が静電エネルギー的に最も安定な構造．

RIETAN-FP・VENUSシステムに新メンバーcifconvが加わったのは望外の喜びだった．今後も拙作ソフトの改良と無料ハンズオンの開催を続行し，研究環境の格差を少しでも是正するよう微力を尽くしていきたい．

謝　辞　第一原理計算関連機能をcifconvに追加するためのモチベーションと貴重な助言を与えていただいた森分博紀，田口綾子（JFCC）の両氏に感謝する．

文　献

1）S. R. Hall, F. H. Allen and I. D. Brown, “International Tables for Crystallography”, Vol. G（2006）4.1.
2）http://openbabel.org/wiki
3）K. Momma and F. Izumi, *J. Appl. Crystallogr.*, **44**, 1272-1276（2011）.
4）https://bit.ly/2VJzL5P
5）F. Izumi and K. Momma, *Solid State Phenom.*, **130**, 15-20（2007）.
6）K. Momma, T. Ikeda, A. A. Belik and F. Izumi, *Powder Diffr.*, **28**, 184-193（2013）.
7）泉　富士夫，まてりあ，**56**, 393-396（2017）.
8）泉　富士夫，まてりあ，**56**, 453-457（2017）.
9）泉　富士夫，まてりあ，**56**, 503-507（2017）.
10）https://www.vasp.at/
11）T. Björkman, *Comput. Phys. Commun.*, **182**, 1183-1186（2011）.
12）https://sourceforge.net/projects/c-tools/
13）K. Okhotnikov, T. Charpentier and S. Cadars, *J. Cheminform.*, **8**, 17（2016）.
14）http://fujioizumi.verse.jp/download/download.html
15）https://bit.ly/2DIijnW
16）J. Akimoto, Y. Gotoh and Y. Oosawa, *J. Solid State Chem.*, **141**, 298-302（1998）.
17）S.-C. Yin, Y.-H. Rho, I. Swainson and L. F. Nazar, *Chem. Mater.*, **18**, 1901-1910（2006）.

筆 者 紹 介

泉　富士夫（いずみ　ふじお）
ファインセラミックスセンター客員研究員，京都大学大学院工学研究科研究員．
［連絡先］ファインセラミックスセンター，京都大学大学院工学研究科
E-mail：fizumi3776@gmail.com

宮崎　晃平（みやざき　こうへい）
京都大学大学院地球環境学堂准教授．
［連絡先］〒615-8510　京都府京都市西京区京都大学桂 Aクラスター京都大学大学院工学研究科
E-mail：myzkohei@elech.kuic.kyoto-u.ac.jp