

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

## 1. 結晶構造データベース

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

(2013 年 1 月 24 日受理)

現在、物質科学の分野において物質設計・探査・開発の速度は著しく急速と成って来ている。同分野において結晶学（結晶構造）に関する情報は、物性の起源を理解するのにも欠かせない手法である。ここに各種結晶構造データベースに関して、それらの概観とアクセス法に関して紹介を行う。

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

### 1. Crystal structure database

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: January 24, 2013)

Recently, speeds of material design, search, and development, are drastically increasing in material science field. The crystal structure database is one of the most important tools in the field, because the crystallographic (including crystal structure) information is always base knowledge, to understand origins of various properties. Brief information of the databases is given here.

#### 1. 序文

結晶学は現代、物質科学における基礎学問であり、古くは有史以前から発達した学問の 1 つと言えます。古代の人は水晶を始めとした天然鉱物の美しい外形が何故、形成しているかを色々な推論の元、考えを巡らせていたと思われます。その後、近代科学・技術の発展と共に単なる肉眼を用いた結晶外形の観察

から、当時の物理学や化学的な知見のみならず、幾何・有理数学などの数学的概念の導入を経て結晶が呈している謎を解明して来ました。その結晶学に大きな転機を示したのが、1912 年のドイツ・Max Theodor Felix von Laue 教授によって発見された X 線回折現象です。[1] von Laue 教授は 1895 年に Wilhelm Conrad Röntgen 教授によって発見された謎の放射線

(X線)[2] を硫酸銅の単結晶に照射する事により、X線の正体が電磁波である事と同時に結晶が空間格子構造(結晶構造)を持っている事を明らかにしました。1913年にはWilliam Henry Bragg教授とWilliam Lawrence Bragg(当時、学生)のBragg父子によって、結晶内の原子間隔とX線との回折の関係を定式化した“Braggの式( $2d \sin\theta = n\lambda$ ,  $d$ : 面間隔,  $\theta$ : 回折角,  $\lambda$ : 波長)”が報告されました。[3, 4] この“Braggの式”により、X線回折による結晶構造解析に理論的裏付けが与えられました。この式はX線のみに関わらず、中性子線・電子線と言った他の線源を用いた回折でも適用出来ます。この様に、X線を用いた結晶の観察手法で初めて、“実は結晶は原子が基本的に規則配列する事によって構成されている”事が判明しました。これら成果が現代結晶学の祖とも言える発見・研究で、これらの成果によりRöntgen教授は1901年に第1回ノーベル物理学賞を受賞し、von Laue教授は1914年に、Bragg父子は1915年に共にノーベル物理学賞を受賞しています。[5] その後、分析手法としてのX線回折法は時代と共に革新的な技術的進歩を経て、現在でも結晶構造を決定するのに最も身近な方法として化学・物理・鉱物学・薬学・生物学などのあらゆる科学分野において数々の成果の発見に貢献しており、毎年、年間数万を超える数の構造解析データが公表されて来ています。

この様にX線(もしくは中性子線, 電子線)を用いて毎年膨大な数の結晶構造データが蓄積されて来ています。それら結晶構造データは単にその物質の結晶構造を明らかにするだけでなく、その物質が有する物性・特性の起源を演繹・帰納するのに欠かせない情報で有ります。又、より高い物性・特性値を示す物質の開発において、類縁もしくは他の物質の情報を知る事も非常に重要な知見でも有ります。その為、既知の結晶構造データの共有化も早くから図られても来ています。1990年代後半までは回折・構造解析技術が現在と比べて未熟だったが故に公表される構造解析データ数も少なく、各年で公表された構造データを纏めて書籍化していました。[6] しかし、目覚ましいコンピューターの進化と共に、効率的かつ革新的な回折・構造解析技術が発達し、前述の様に膨大な数の物質の結晶構造の解析が毎年なされる様になると、構造データを纏めて数年後に書籍化すると言った悠長な事は言うておられず、コンピューターを利用したデータの共有化が図られる事になって来ました。しかしながら、データベースが作られ始めた黎明期はデータの共通フォーマットが

無かった為に、色々なデータベースを利用するユーザーは各フォーマットの詳細を知る必要が有り、それらデータを利用して結晶構造描画、電子状態などの計算を行うプログラム作者は自身のプログラムを各データ・フォーマットに対応させる必要が有る言う煩雑さが生じました。この煩雑さを解消させる為に国際結晶学連盟(International Union of Crystallography: IUCr)が主導と成って開発されたのが結晶学共通データ・フォーマット CIF(Crystallographic Information File)です。[7] このデータ・フォーマットの標準化を推し進める事により、現在、主要なデータベースではCIFをデータ入出力ファイルの1つとして採用しております。結晶構造描画プログラムを始め数多くの科学計算プログラムでもデータ入出力フォーマットの1つとしてCIFを採用しており、入力時の煩雑さやミスを低減させています。加えて、アメリカ化学会系の各学術誌を始めとして、Nature, Science, Elsevier, Wiley, Springerなど主たる科学論文出版社も論文投稿時に、結晶構造データは今回紹介する各データベースに登録し、論文中にその登録番号を付記するか、投稿原稿と共にCIFとして提出する事を義務付けています。

本稿では、先ずは、現在、利用されている各種結晶構造データベースに関して紹介し、次号で結晶学共通データ・フォーマット CIFに関する紹介をさせて頂きます。

## 2. 各種結晶構造データベース

結晶構造に関するデータベースは大別して以下の3つが存在しています。

1. 結晶構造自体
2. 粉末回折パターン
3. 各種物性値と共に併載

以下に代表的なデータベースに関してその概略を順番に紹介して行きます。

### 2-1. 結晶構造自体

本データベースに収録されたデータは、単結晶試料もしくは粉末(多結晶)試料と(X線, 中性子線もしくは電子線)回折法を用いて得られたデータから解明された結晶構造の詳細(格子定数, 空間群, 各原子座標および温度因子, 各元素の席占有率, 結晶構造から演繹された化学構造式など)を基本としています。これらの結晶学データが有れば、単なる結晶構造の描画のみならず、対象物質の(単結晶お

よび粉末) 回折パターン・シミュレーション, 電子構造の計算, 各種物性値などの計算が可能と成ります。

本データベース群は, 大まかに収録されている物質群によって大別されています。

1. 無機物質 (含む合金・金属間化合物) 系
2. 有機物質 (含む有機金属錯体) 系
3. タンパク質系
4. 複合系

各データベースの概観は以下の通りです。

### 2-1-1. 無機物質 (含む合金・金属間化合物) 系

#### A. Inorganic Crystal Structure Database (ICSD)

項目	内容
制作	FIZ Karlsruhe (独), NIST (米)
内容	無機化合物の名称, 分子式, 三次元原子座標値, 結晶学データ, 書籍事項
件数	15 万件以上 (2012 年秋版)
更新	年 2 回, 年約 6,000 件追加
利用形態	DVD 版 (ローカル・インストール) Windows XP/Vista/7 Intranet 版 Linux Web 版 Windows, Mac OS, Linux (Internet Explorer 7/8, Firefox 3.0 以降) Java Runtime Environment (JRE) 1.6 のインストールが必要
契約形態	年間定額 (大学割引があります)
対応するオンライン・ファイル	ICSD (STN)

ICSD は元々 1978 年にドイツ・FIZ Karlsruhe - Leibniz Institute for Information Infrastructure で開発が開始されたデータベースを基にしており, 後年アメリカ・National Institute of Standards and Technology (NIST) が参画し, NIST が独自に収集していたデータ (NIST Crystal Data: NCD) を ICSD に加えると共に, 現在は両研究所が共同で開発している無機結晶構造データベースです。[8] 収録物質の範囲は, 鉱物・セラミックス・合金・金属間化合物などです。近年は, 実験的に得られた結晶構造のみならず, 第一原理計算などで予測された結晶構造のデータも追加収録しています。

市販の ICSD には, 簡単な結晶構造表示や粉末パターンのシミュレーション等のプログラムも含んでおり, 検索キーは含有元素, 組成式, 結晶学データなどと成っています。

ICSD への自身のデータの登録や最新のデータの

入手依頼は直接 ICSD オフィスへのメール (CrysDATA@fiz-karlsruhe.de) で行います。この際, データの受け渡しは CIF にて行われます。

データ更新: 年 2 回

件数 (2012 年 10 月現在): 約 16 万件

開発元:

<http://icsd.fiz-karlsruhe.de/icsd/>

日本での販売元 (化学情報協会):

[http://www.jaici.or.jp/wcas/wcas\\_icsd.htm](http://www.jaici.or.jp/wcas/wcas_icsd.htm)

#### B. CRYSTMET

項目	内容
制作	Toth Information Systems, Inc. (カナダ)
内容	金属, 合金の名称, 分子式, 三次元原子座標, 結晶学データ, 書籍事項
件数	15 万件以上 (2012 年秋版)
更新	年 2 回, 年約 6,000 件追加
頒布形態	CD (検索・表示ソフト付)
使用可能機種	Windows XP/Vista/7
契約形態	年間定額 (大学割引があります)

元々は 1960 年にアメリカ・Los Alamos National Laboratory で合金・金属間化合物・鉱物を対象に結晶構造データの収集が開始され, その後, カナダ・National Research Council of Canada に引き継がれたデータベースで, 1996 年にカナダ・Toth Information Systems 社にその開発・販売が移管された物です。[9] 収録対象物質は, 合金・金属間化合物・鉱物を含む金属系ですが, 一部, 酸化物・ハロゲン化物・硫化物などの無機化合物も含んでいます。CRYSTMET の検索キーは含有元素, 組成式, 粉末パターンなどが有ります。

データ更新: 年 2 回

件数 (Ver. 4.9.0, 2012 年 11 月): 15 万件以上

開発元:

<http://www.tothcanada.com/databases.htm>

日本での販売元 (化学情報協会):

[http://www.jaici.or.jp/wcas/wcas\\_crystmet.htm](http://www.jaici.or.jp/wcas/wcas_crystmet.htm)

## C. Pearson's Crystal Data (PCD)

PCD は序文で示した書籍版結晶学データベース (*Strukturbericht, Structure Reports* [6]) の編纂にも参画され、早くから結晶構造データベースの構築を行い、合金・金属間化合物系の結晶構造に関して分類を行って来た William Burton Pearson 教授に敬意を払い命名されたデータベースで有り、アメリカ・ASM International 社が出版した "Pearson's Handbook: Crystallographic Data for Intermetallic Phases" [10] や後述する Pauling File がベースと成っています。

現在は ASM International 社とスイス・Materials Phases Data System 社が共同開発しており、金属間化合物のみならず無機化合物全般に渡って収録しています。

データ更新：年1回

件数 (2013 年版)：23 万件以上

開発元 (online)：

<http://www.asminternational.org/portal/site/www/info/online-databases/pearsons/>

開発元 (CD-ROM)：

<http://www.crystalimpact.com/pcd/Default.htm>

日本での販売元 (ライトストーン社)：

<http://www.lightstone.co.jp/pearsonscd/index.html>

## 2-1-2. 有機物質 (含む有機金属錯体) 系

## A. Cambridge Structural Database (CSD)

CSD の開発元はイギリス・Cambridge Crystallographic Data Centre (CCDC) で、X線・中性子線回折法などの手法を用いて得られたデータを解析して得られた有機化合物・有機金属化合物の結晶構造データベースです。1965年から開発が開始され、世界で最も権威ある結晶構造データベースとされています。[11]

市販データベースには書誌事項、分子構造、結晶構造などを検索するプログラムのほか、結晶構造表示、データ解析などのプログラムをも含んでいます。

又、CSDへのデータの登録や最新のデータの依頼は下記のCSDのホームページから行います。この際、データの受け渡しはCIFです。

データ更新：年1回+随時更新

件数 (2013年1月現在)：57万件以上

開発元：

<http://www.ccdc.cam.ac.uk/products/csd/>

日本での販売元

(会社など営利団体向け：化学情報協会)：

[http://www.jaici.or.jp/wcas/wcas\\_ccdc.htm](http://www.jaici.or.jp/wcas/wcas_ccdc.htm)

日本での販売元

(大学など教育機関向け：大阪大学・蛋白質研究所)：

<http://www.protein.osaka-u.ac.jp/csd/csd.html>

## 2-1-3. タンパク質系

## A. Protein Data Bank (PDB)

1971年にアメリカ・Brookhaven National Laboratoryにて開発されたデータベースで有り、1998年にアメリカ・Research Collaboratory for Structural Bioinformatics (RCSB)に移管された後、2003年、RCSBを含むヨーロッパ (Protein Data Bank in Europe: PDBe) および日本 (Protein Data Bank in Japan: PDBj) の3つの研究組織により、Worldwide Protein Data Bank (wwPDB) が結成され、PDBのデータの登録、処理、配布が行われています。

唯、物質が巨大分子であると言う特殊性からその出力ファイルのフォーマットは後述の通常のCIFでは無く、PDB Format, PDB Exchange Format (mmCIF), PDBML Format (xml) などと成っています。しかし、2013年中には共通データ・フォーマットが確立される予定に成っています。[12]

又、PDBへのデータの登録や最新のデータの検索・入手は下記のwwPDBやPDBjなどの各機関のホームページから行います。

件数 (2013年1月現在) : 8万件以上  
開発元 (wwPDB) :

<http://www.wwpdb.org/index.html>

日本での配布元 (PDBj) :

[http://www.pdbj.org/index\\_j.html](http://www.pdbj.org/index_j.html)

## 2-1-4. 複合系

## A. Crystallography Open Database (COD)

前述までのデータベースはその殆どが市販されています。唯、ICSD, CRYSTMET, PCDなどのデータベースは初期投資および年間維持費がかかる為、異分野の方や学生を含め初心者の方や財政的に厳しい発展途上国の方にはそれらデータベースの購入・利用が非常に困難です。そこでフランス・Université du MaineのArmelle Le Bail教授が中心となり、“**無料で、開かれた結晶構造データベースの開発**”というコンセプトの下で開発されているのが本データベースです。[13, 14] 本データベースは有志によって運営されており、筆者もAdvisory Boardとして登録されています。対象化合物は無機(含む合金系)・有機(含む有機金属錯体)を問わず収録されており、全てweb上でデータの検索、データの登録が可能と成っています。

本データベースは無料で公開されている為、Rigaku社やPANalyticals社と言った主要なX線回折計メーカーの粉末回折装置付属のデータベースの1つとして活用されていると共に、市販の粉末回折データ・マッチング・プログラム (Match! 2) などの標準データベースとしても使用されています。

件数 (2013年1月現在) : 21万件以上  
開発元 :

<http://www.crystallography.net/>

## 2-2. 粉末回折パターン

このデータベースは物質開発を行っている方々に最も身近な物でしょう。前述の“結晶構造のみ”のデータベースとは異なり、基本的に実験で得られた生データを収録した形に成っています。このデータベースの趣旨は、物質の粉末回折パターンは物質が有する結晶構造固有の物（指紋データ）ですので、既知のパターンを収集し、新規物質のパターンと照合する事により、ザックリとその物質の結晶構造を推測しようと言う物です。基本的にパターン・マッチングですので、データベースに収録されている物質のデータは各原子座標までの結晶構造が必ずしも明らかに成っている必要は無く、結晶構造が分かっていない物は、未知結晶構造物質のデータとして収録されます。故に物質を合成した後に得られた物質を先ず分析する手法が粉末回折法で、本データベースは基本、その結果の解釈・解析に用いられます。

### A. Powder Diffraction File (PDF)

このデータベースは、現在、国際回折データ・センター (International Centre for Diffraction Data: ICDD) が中心と成ってデータを収集・公表している物で、別名 ICDD データとも呼ばれます。読者の皆さんも一度は御聞きに成られたり、使用された事が有られる方も多いのではないのでしょうか？このデータベースは収集・公表機関の変遷が有り、昔は各収録データがカード形式でしたので、古くは ASTM (American Society for Testing and Materials) カードもしくは JCPDS (Joint Committee for Powder Diffraction Standards) カードと呼ばれており、図書館などでカード・ファイルや書籍として保存されていました。[15]

本データベースに収録されるデータはその内容により、レベル1 (PDF-1) からレベル4 (PDF-4) と分類されます。具体的なデータの収録内容は、各面間隔： $d$  値、最強回折ピーク強度との相対ピーク観測強度： $I/I_{max}$ 、物質名・鉱物名・化学式や測定条件や分かれば格子定数、空間群、各回折ピークの指数  $hkl$  などの結晶学的情報を含んだ ID 情報がその基本データで、レベル1 (PDF-1) と呼ばれるデータセットと成ります。レベル2 (PDF-2) は、レベル1の内容に加えて、前述の NCD および ICSD データベースに収録されている結晶構造データから逆算された粉末回折パターン・データを含みます。その為、収録データ数が膨大と成り、CD-ROM (もしくは DVD-ROM) などの記録媒体にデータベースを収録し、コンピューター検索を容易にした物と成っています。そのレベル2データに生の回折パターンをも収録した物がレベル3 (PDF-3) ですが、これは現在 (2013年) に到っても公表されていません。その代わりに公表されたのが、レベル4 (PDF-4) です。このレベルのデータセットは基本的にはレベル2データのアップデート版であるのですが、仕様がちょっと複雑で、以下の3つの派生型が公表・販売されています。

1. PDF-4+ : 無機物系のみ
2. PDF-4 Organics : 有機物のみ
3. PDF-4 Minerals : 鉱物系のみ

PDF-4+は、無機化合物のみ (含む合金系) のデータを収録すると共に、後述する Pauling File に収録されている結晶構造データから計算されたデータ・回折パターンが追加収録されています。[15, 16]

PDF-4 Organics は、有機物のみを対象とし、前述の CCDC に収録されている結晶構造データから計算されたデータ・回折パターンが追加収録されています。

PDF-4 Minerals は、PDF-4+から鉱物系のデータのみを抜粋した物と成っています。

新規データを論文などで公表した場合、そのデータの PDF への登録は開発元 ICDD へ自己申告するか、ICDD から登録依頼メールが登録用紙と共に送られて来るので、その用紙に登録必要事項を記入し、生データと共にメールで返信する事によりなされます。

データ更新：年1回

件数 (2013年1月現在) :

32 万件以上 (PDF-4+),

47 万件以上 (PDF-4 Organics),

約 4 万件 (PDF-4 Minerals)

開発元 :

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

日本での販売元 (ライトストーン社) :

<http://www.lightstone.co.jp/icdd/index.html>

### 2-3. 各種物性値と共に併載

ここからのデータベース群は結晶構造データのみを含んだ物では無く、状態図・電気特性・磁性・光学特性など各種物性値と共に収録されており、総合物性データベースとも言えます。

### A. Pauling File



このデータベースは化学・結晶学に多大なる影響を与えた Linus Carl Pauling 教授に肖って命名されたデータベースで、1995年に日本・科学技術振興機構 (JST)、スイス・Material Phases Data System 社 (MPDS) および日本・東京大学が中心になってスタートした”the Pauling File project”が元と成っています。2002年からは同プロジェクト終了と共に主導権は MPDS に委譲され、二元系化合物のデータに限ったデータベースを公表しました。[17] 本 Pauling File に収録されているデータは無機化合物で、それらの状態図、結晶構造、回折パターン、物性データなどを収録しています。次から紹介するデータベース (SpringerMaterials 中の Inorganic Solid Phases および AtomWork) における結晶構造に関するデータは本データベースがその源と成っています。

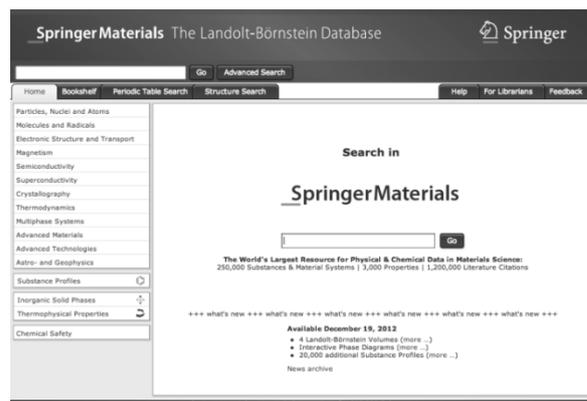
データ更新 : 定期更新無し

件数 (2008年時点) : 約 25 万件

開発元 :

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

### B. SpringerMaterials: Landolt-Börnstein Database



出版社 Springer 社が物理学・化学・工学分野において体系的かつ包括的に厳選した物性情報を収録した世界最大のデータベースです。元は同社出版の膨大な物性情報を纏めた Landolt-Börnstein (通称: LB) 全巻 [18] であり、本データベースは以下の4つから構成されています。

1. Landolt-Börnstein:  
物理科学・化学・工学のデータ集
2. Inorganic Solid Phases:  
固体無機化合物に関するデータベース
3. Thermophysical Properties:  
熱物性に関するデータベース
4. Chemical Safety:  
化学物質安全性データシート

唯、残念ながら、現状では結晶構造基本データ・フォーマットである CIF として本データベースから結晶構造データを出力させる事は出来ません。

データ更新 : 随時

開発元 :

<http://www.springermaterials.com/navigation/>

### C. AtomWork



本データベースは 2010 年に公開されたデータベースで無機材料を対象としており、収録データは日本・物質・材料研究機構 (NIMS), スイス・Material Phases Data System 社 (MPDS), 日本・科学技術振興機構 (JST) が共同で製作した物です。[19]

収録内容は、科学技術文献から抽出した無機材料の結晶構造, X 線回折, 特性, 状態図に関するデータで、以下の 3 つ方法で検索が出来ます。

#### 1. Search Material

化学成分系・化学式・物質名・プロトタイプ・ピアソン記号・空間群番号などを指定して物質を検索

#### 2. Search Materials having specified property

材料特性を指定して物質を検索

#### 3. Search phase diagrams

成分系を指定して状態図を検索

このデータベースにおける結晶構造の部分は前述の Pauling File がその元データと成っています。2010 年における登録データ数は以下の通りです。

データ更新：随時

登録数 (2010 年時点)：

結晶構造：8 万件以上,

特性：5 万件以上,

状態図：1 万以上

開発元：

<http://crystdb.nims.go.jp/>

### 3. おわりに

現代における物質探索・開発スピードはこれまでに無く早く成っています。その中において各種データベースの重要性は増すばかりです。我々が普段、実験結果を議論する際に使用しているイオン半径・金属半径・Van der Waals 半径や Bond valence sum (結合距離からターゲット・イオンの価数を見積もる方法) も元々は 1938 年に刊行された Linus Carl Pauling 教授の名著 (*The Nature of the Chemical Bond*, 化学結合論) [20] で述べられているコンセプトをベースにしており、後年、数多くの物質の結晶構造が解明された事によって、それら結晶構造から得られた各結合距離や結合角度を基に前述のイオン半径 [21] や Bond valence sum 法の初期値 ( $R_0$ ) [22, 23] などの各値が精密化されて来ました。この際に結晶構造データベースが活用された事は言うまでも有りません。

現代における具体的なデータベースの活用法とし

ては、物質を合成する前の物質設計段階において関連物質の検索、合成後はその合成物の評価において用いられている事でしょう。それ以外にも大学などの教育現場においても具体的に結晶構造を提示して講義される前のデータベースを用いた結晶構造の検索がなされている事でしょう。具体的なデータベースの活用法は文献 2 4, 2 5 を御参考下さい。

これらデータベースの重要性・汎用性が増したのは、P C やインターネットの発達も有りますが、国際結晶学連盟 (IUCr) が積極的に推し進めた結晶学共通データ・フォーマット CIF の普及が大きな要因の 1 つです。[7] この CIF に関しては次の機会に紹介させていただきます。

### 付記

本稿では、前述の様に結晶構造に関する主要なデータベースを概略と共に紹介して参りましたが、これらのデータベース以外にも結晶学、結晶構造および結晶成長に関するデータベースは数多く存在します。加えて、結晶構造データベース用に発刊されている専門誌も有ります。それらの内、代表的なものに関し、以下に web address のみを列記しますので、御参考下さい。

#### (データベース)

結晶学総合データベース：

<http://www.cryst.ehu.es/>

ゼオライト結晶構造データベース：

<http://www.iza-structure.org/databases/search/ncd/>

鉱物結晶構造データベース：

1. <http://ruff.geo.arizona.edu/AMS/amcsd.php>

2. <http://database.iem.ac.ru/mincryst/index.php>

3. <http://ruff.info/>

カルコゲナイド化合物結晶構造データ集：

<http://crystalmaker.com/library/chalcogenides.html>

非周期結晶構造データベース：

<http://158.227.21.14/incstrdb/index.php>

ナノ結晶データベース：

<http://nanocrystallography.research.pdx.edu/>

結晶外形データベース：

<http://nanocrystallography.research.pdx.edu/search/cmd/>

仮想 MOF (Metal-Organic Frameworks) データベース：

<http://hmofs.northwestern.edu/hc/crystals.php>

MOF (Metal-Organic Frameworks) 物質データベース：

<http://helios.princeton.edu/mofomics/>

核酸データベース

<http://ndbserver.rutgers.edu/>

生化学系巨大分子結晶化データベース :

<http://xpdb.nist.gov:8060/BMCD4/>

#### (雑誌)

IUCr journals (*Acta Crystallographica* A~F, 特にCとE) :

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

*Zeitschrift für Kristallographie - New Crystal Structures* :

<http://www.oldenbourg-verlag.de/wissenschaftsverlag/zeitschrift-kristallographie-new-crystal-structures/14337266>

X-ray Structure Analysis Online :

<http://www.jsac.or.jp/cgi-bin/xraystruct/toc/>

#### 4. 謝辞

本稿を執筆するに当たって貴重な機会を与えて下さった(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・材料研究機構の徐一斌博士にはAtomWorkの詳細に関しまして、教授頂きました。心より感謝致します。加えて、化学情報協会の桜井尋海博士には同協会から発売されている各種データベースの情報並びにフォーマット類などに関する情報および同協会のホームページの転載許可を頂きました。心より御礼申し上げます。又、ライトストーン社には同社のホームページの転載を許可して頂きました事、御礼申し上げます。

#### 5. 参考文献

- [1] M. von Laue, *Physikalische Zeitschrift*, **14**, 1075 (1913).
- [2] W. C. Röntgen, *Sitzungsberichte der Physikalisch-medicinischen Gesellschaft zu Würzburg*, 132 (1895).
- [3] W. L. Bragg, *Proc. Cambridge Phil. Soc.*, **17**, 43 (1913).
- [4] W. H. Bragg and W. L. Bragg, *Proc. Royal Soc. London.*, **88**, 428 (1913).
- [5] [http://www.nobelprize.org/nobel\\_prizes/physics/](http://www.nobelprize.org/nobel_prizes/physics/)
- [6] “*Strukturbericht*, Vols. 1~7.” Akademische Verlagsgesellschaft, Germany (1931~1943), and “*Structure Reports*, Vols. 8~58”, Kluwer Academic Publishers, Germany (1956~1990).
- [7] “*International Tables for Crystallography, Vol. G: Definition and Exchange of Crystallographic Data*”, ed. by S. Hall and B. McMahon, Springer, Dordrecht, the Netherlands (2005).
- [8] G. Bergerhoff and I. D. Brown, in “*Crystallographic Databases*”, ed. by F. H. Allen, G. Bergerhoff, and R. Sievers, Chapter 2.2, pp. 77-95, International Union of Crystallography, Chester, England (1987).
- [9] P. S. White, J. R. Rodgers, and Y. Le Page, *Acta Crystallogr.* **B58**, 343 (2002).
- [10] P. Villars L. D. Calvert, “*Pearson's handbook of crystallographic data for intermetallic phases*, Vols. 1~3”, American Society for Metals, Metals Park, Ohio, U. S. A. (1986).
- [11] F. H. Allen, *Acta Crystallogr.*, **B58**, 380 (2002).
- [12] 日本蛋白質構造データバンク *Newsletter*, **14**, 1 (2012).
- [13] S. Gražulis, D. Chateigner, R. T. Downs, A. F. T. Yokochi, M. Quirós, L. Lutterotti, E. Manakova, J. Butkus, P. Moeck, and A. Le Bail, *J. Appl. Crystallogr.*, **42**, 726 (2009).
- [14] S. Gražulis, A. Daškevič, A. Merkys, D. Chateigner, L. Lutterotti, M. Quirós, N. R. Serebryanaya, P. Moeck, R. T. Downs, and A. Le Bail, *Nucleic Acids Research*, **40**, D420 (2012).
- [15] J. Faber and T. Fawcett, *Acta Crystallogr.*, **B58**, 325 (2002).
- [16] S. N. Kabekkodu, J. Faber, and T. Fawcett, *Acta Crystallogr.*, **B58**, 333 (2002).
- [17] P. Villars, M. Berndt, K. Brandenburg, K. Cenzual, J. Daams, F. Hulliger, T. Massalski, H. Okamoto, K. Osaki, A. Prince, H. Putz, and S. Iwata S. *J. Alloys Compd.*, **367**, 293 (2004).
- [18] “*Landolt-Börnstein series*”, Springer Publishers, Heidelberg, Germany (1883~2012).
- [19] Y. Xu, M. Yamazaki, and P. Villars, *Jpn. J. Appl. Phys.*, **50**, 11RH02 (2011).
- [20] L. Pauling, “*The nature of the chemical bond and the structure of molecules and crystals: an introduction to modern structural chemistry*”, Cornell University Press, Ithaca, U. S. A. (1939: 1<sup>st</sup> edition, 1959: 3<sup>rd</sup> edition)”, and “*化学結合論 (改訂版)*”, 小泉正夫 (訳) 共立出版, 東京 (1962).
- [21] R. D. Shannon, *Acta Crystallogr.*, **A32**, 751 (1976).
- [22] I. D. Brown, “*The chemical bond in inorganic chemistry*” IUCr Monographs in Crystallography 12, Oxford Science Publications, Oxford, U. K.

(2002).

[23] <http://www.iucr.org/resources/data/data-sets/bond-valence-parameters>

[24] 熊田伸弘, セラミックス **35**, 881, (2000).

[25] 桜井尋海, 化学経済 (3) 54 (2012).

### (後述)

本稿に関する元々の依頼内容は, “結晶構造データベースの紹介と作成時における経験談” でした. 故に最後に経験談を述べさせていただきます.

私は学部4年(1992年)の卒論研究として, 結晶中の結合様式として共有結合とイオン結合のみならず, 孤立電子対由来の立体活性状態を有すると考えられている非常に複雑な結晶構造を有する硫化物系鉱物の結晶構造解析を与えられました. 当時は PC も各個人が有する物とは言い難く, PC 自体の能力も現在の物とは比較に成らないくらい低い物でした.

(私は PC が嫌いなので, 個人的に PC を購入したのは博士論文を書かなければ成らないと言う理由から博士課程3年の時でした!) 又, 当時は PC 上で検索出来る結晶構造データベースも無く, 結晶構造データを基にその結晶構造を描画させる様なプログラムもまともな物が無い時代でも有りました. 唯, 私の研究を遂行する上において, 既知の関連物質の結晶構造の知見は必須でしたので, 実験の合間を縫って足しげく図書館に通い, 関連物質の結晶構造が掲載されている雑誌を片っ端から検索, 孫引き&コピーして読み漁っていました. しかも, 雑誌に掲載されている結晶構造データは全てが数値データ(各原子が結晶中に存在する位置の三次元座標)です. それを研究室に有る共用の PC にデータ入力し, その結晶構造図を描画させて初めてその構造の全体像が解ると言う状態でした.(当時はスキャナー自体無かったので, 当然の事ながら全て手入力でした.) この結晶構造図を描画させる過程で, 各原子(イオン)間の結合を描画させる訳ですが, 幾つものデータで印刷ミスを発見し, それらの印刷ミスを自分で各原子間距離・角度から逆算して修正し, 正解を得る作業も行っていました.

こうした作業を行っている際に強く感じたのが, “正しい情報の入ったデータベースの重要性” と “日本発のデータベースの作成” でした. 前者は言うまでも無いのですが, 後者の “日本発のデータベース” は本稿でも紹介しましたが, 現在においても完全に日本で作成された公のデータベースは有りません.

そこで私は, “みんなが作らないなら自分で作る! しかも, 公表された基礎データは万人が共有すべき物であるので, 無料で配布する! 1日に5つの構造を入力して行ったら, 個人でも年間に2000化合物近いデータが入られる!” と PC でゲームをしている友人達を横目にスタートさせたのが, 付記で示しました “カルコゲナイド化合物結晶構造データ集” です. 都合, 現在は公式的にはアップデートをしておりませんが, 15,000種以上の化合物のデータが登録されており, 個人ベースのデータ集としては現在でも世界最大級であると自負しております.

このデータ集作成時には色々障害が有りました. 最大の障害は自身で殆どプログラムが書けないと言う自己の能力の無さが有りました. それ以外にも周りから, “**何でそんな無駄な事をしているの?**” とか “**そんな事, 外国に任せておけば良いじゃない?**” と言った意見が, 教官の方々からも有りました. (尤も, 教官からの言葉の真意は, “**もっと実験して成果を出せ!**” と言う意味だったのでしょうが, , , ,) しかし, これらの意見は私には前時代的な物と感じられましたので, 時間を見てはデータの整合性のチェック&データの入力を行っていました. その後, 形になるに従って, 完成品としてフロッピー・ディスクにコピーして, 外国の同様の物質系を研究している研究者を中心に無料で配布し始めました. 国内はともかく, 諸外国からは概ね好評で, “書籍中に使いたい!”, “論文中で使いたい!”, “授業に使いたい” と言う私のデータ集の許可を得る為の手紙・メールが舞い込み始めました.

この時に感じたのが,

1. **例えプログラム書きが不得手で, 経験値の少ない4年生で有っても, 目的意識とその分野におけるしっかりとした基礎知識が有れば, データ集の作成は可能である.**
2. **土壌として, 欧米の方々の方が過去に発表されたデータに関して真摯に向き合っている.**

と言う事でした. 特に2に関してですが, 本稿で紹介させて頂いた様に各データベースは古くから欧米(特にドイツは構造解析法が確立され, その成果が公表され始めると直ぐに集積開始!)で構築された物です. この様に欧米では, 結晶構造などの様な基礎データは自分達の研究のみならず, “**各分野へ貢献しえる物**” であると言う認識と, 例え, 特定の研究室から発表されたデータで有っても, そのデータが公表された以上は “**万人が共有すべきデータ**” であると言う強い認識が有るものだと思います.

唯、私のデータ集には欠点があります。それは私が本データ集を作成するに当って結晶学共通データ・フォーマット CIF では無く、特定の市販結晶構造描画プログラム用のバイナリー・データ・フォーマット（最初は ATOMS, 現在は CrystalMaker）を採用してしまった点です。私がデータ集を構築し始めた当時（1992 年頃），“CIF は未だフォーマットが完全に確立しておらず、普及もしていない。例え私のデータ集を使用される方がこれらのプログラムの正規版を持っていないともデモ版が手に入るので良いであろう”と私は考えていました。後年、Crystallography Open Database (COD)の提唱者である Le Bail 教授から同データベースへの参画を打診された際に私のデータ集のデータの全部を CIF に変換して COD に収録してくれと要望されました。これには非常に手間取り、私には詳細が未公開の ATOMS および CrystalMaker バイナリー・データ・フォーマットが解析出来ず、断念しました。

最後に、表面分析研究会では電子分光やイオン分光・データに関する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ので、僭越では有りますが、先駆的な結晶構造データベースに何らかの関与した身からの以下の提言をさせていただきます。

1. 開発前に最終的なデータの利用のされ方を想像する。
2. 国際標準的なデータのフォーマットを確立する。
3. 国際的なデータ集積所を作る。
4. 確立した標準フォーマットで測定・解析データが出力出来る様に装置メーカー・プログラム開発者に周知徹底する。
5. 各関連学会や出版社でも、同フォーマットでのデータの取り扱いを可能にする様に協力を御願ひする。

と言った事が肝要かと思います。これらを念頭にデータベースを構築すれば、分野が異なる方も含めた万人が利用しやすいデータベースが構築出来るのではないのでしょうか？