2023年度 PX-UG中級者講習会
＠東京理科大学葛飾キャンパス

DIALSの紹介と演習

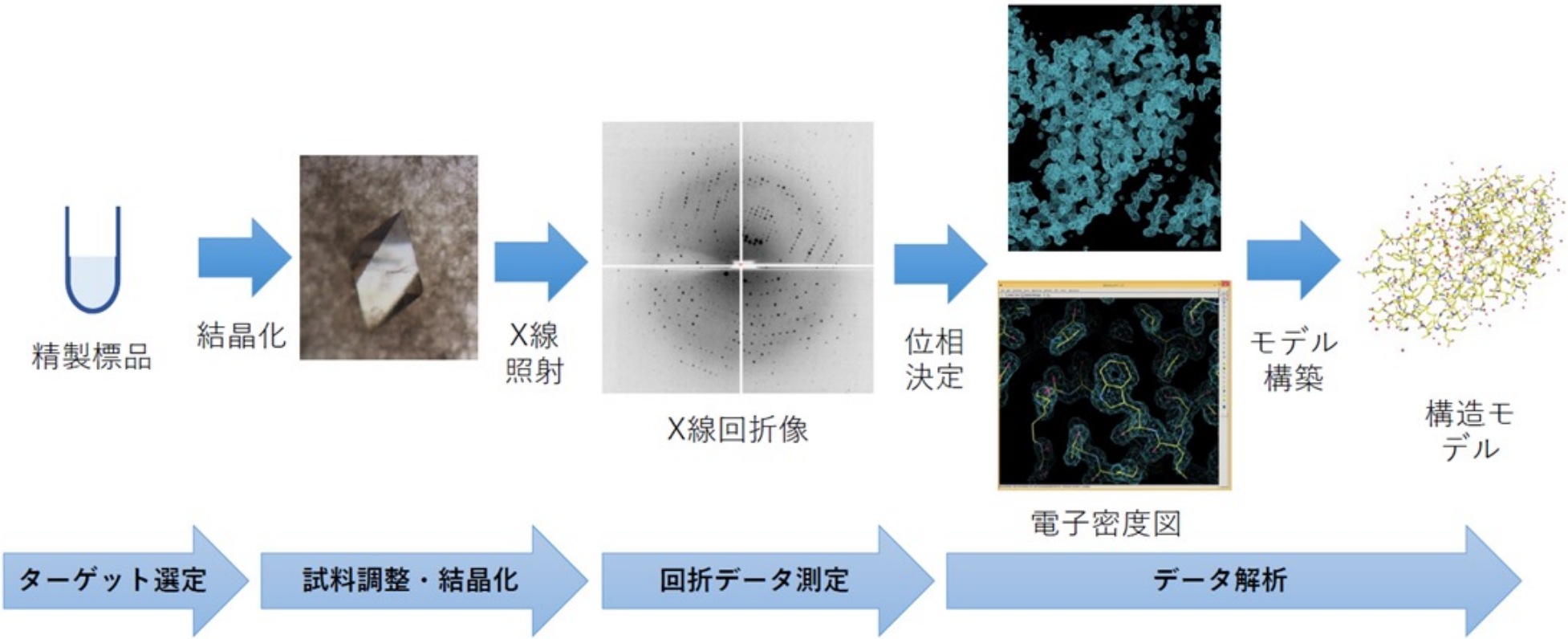
東京理科大学　生命システム工学科
白石　充典
伊藤　翔

本日の内容

1，DIALSの紹介　（白石）

2，演習　（伊藤）

タンパク質X線結晶構造解析の流れ



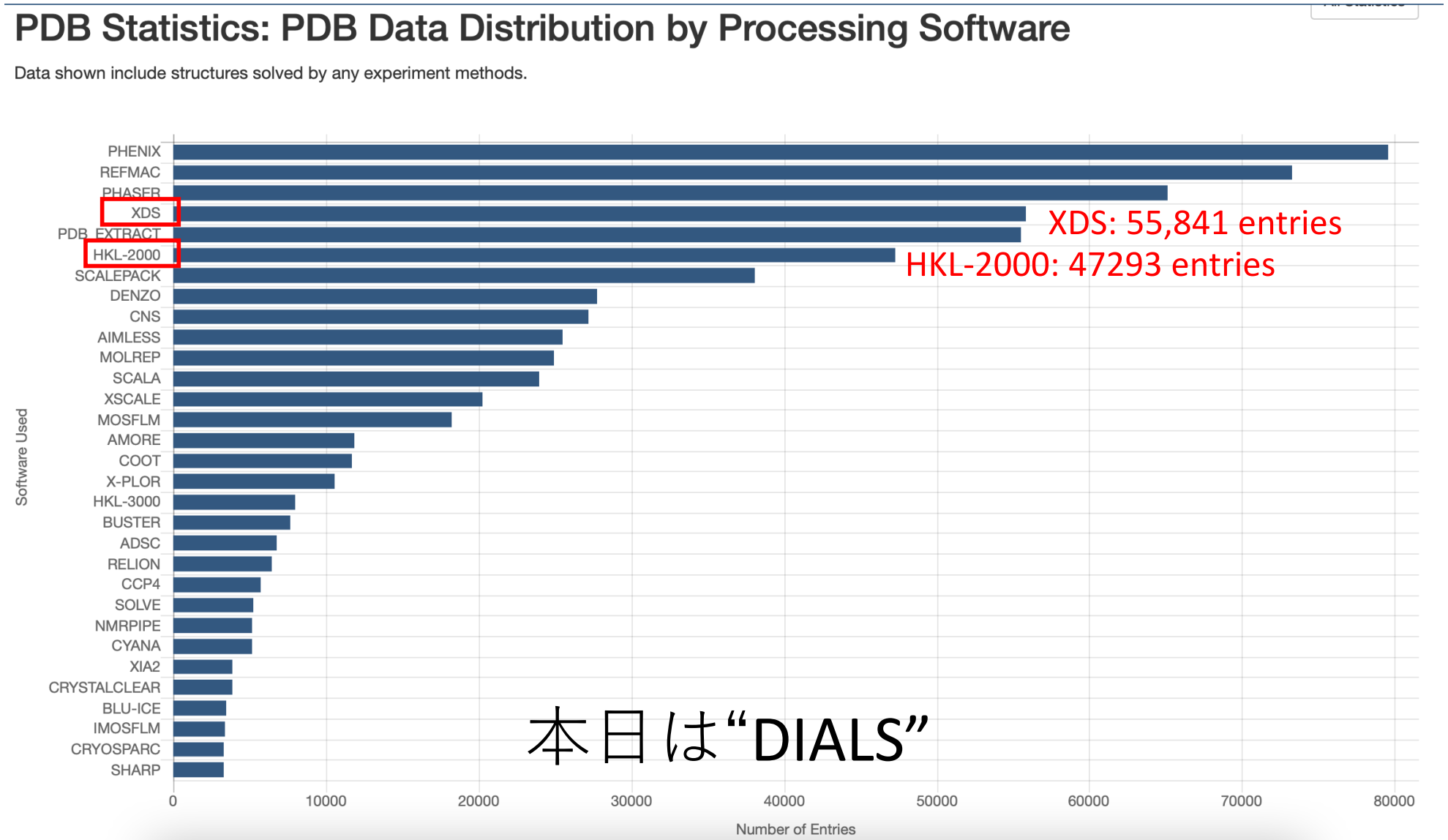
回折データの処理

- 格子定数，空間群の決定
- 回折強度の積分
- 強度データの統計処理

より良い処理が，より良い構造解析につながる。

「KEK構造生物学研究センターにおけるクラウドの試み」山田悠介氏のスライドより
https://www.hpcwire.jp/archives/47366

回折データ処理にどのソフトを使ってきましたか？



DIALS: Diffraction Integration for Advanced Light Sources

DIALSのトップページ　https://dials.github.io/index.html#

DIALS

Diffraction Integration for Advanced Light Sources

- Home
- About
- Installation
- Documentation
- Tutorials
- How-to
- Knowledge Base
- Workshops
- Publications
- Projects
- License

DIALS: Diffraction Integration for Advanced Light Sources

X-ray crystallography for structural biology has benefited greatly from a number of advances in recent years including high performance pixel array detectors, new beamlines capable of delivering micron and sub-micron focus and new light sources such as XFELs. The DIALS project is a collaborative endeavour to develop new diffraction integration software to meet the data analysis requirements presented by these recent advances. There are three end goals: to develop an extensible framework for the development of algorithms to analyse X-ray diffraction data; the implementation of algorithms within this framework and finally a set of user facing tools using these algorithms to allow integration of data from diffraction experiments on synchrotron and free electron sources.

Contact

For feature requests, bug reports or any other information, please contact the DIALS developers at dials-support@lists.sourceforge.net, or post an issue.

Mirror sites: DIALS-East, DIALS-West, dials.github.io.

Funding

DIALS
Diffraction Integration for Advanced Light Sources

“共同開発による”，<u>最新技術に対応した</u>回折データ処理・解析のための統合ソフトウェア

開発における最終目標

- ✓日々の進歩に対応できる，拡張可能なフレームワークの開発
- ✓フレームワークへのアルゴリズムの実装
- ✓シンクロトロンや自由電子源でのデータ処理が可能な一連のユーザー向けツールの開発

開発チーム

DIALS East　diamond　CCP4　UKRI Science and Technology Facilities Council

DIALS West　BERKELEY LAB　SSRL

https://dials.github.io/より

DIALS; About

完全なオープンソース。

（開発チーム以外の貢献もウエルカム）

the DIALS project builds on knowledge accumulated over many decades in the field of crystallographic data processing. We benefit greatly from the altruism of experts who contribute their ideas and advice, either directly or via their detailed publications on existing algorithms and packages such as XDS [2] and MOSFLM [3].

無駄に手間をかけないように，既存のソフトウエア（XDSやMOSFLM）を大いに参考にしている。

Continuing in the spirit of these workshops we held our own series of meetings, with talks from invited speakers, and code camps in which specific problems are addressed by intensive effort across the collaboration. Summaries of these meetings and copies of slides given as presentations are available here.

ミーティング，招待講演，コードキャンプを重ねて開発していった。

https://dials.github.io/about.htmlより

Workshops

- DIALS 1: 12th-13th June 2012 (Diamond Light Source, UK)
  - Organising committee
  - Program and presentations
- DIALS West-1: July 2012 (Boston, MA)
- DIALS-2: February 2013 (Berkeley, CA)
- DIALS-3: 22nd May 2013 (Cambridge, UK)
  - Attendees
  - Program and presentations
- DIALS-4A: September 2013 (Diamond Light Source, UK)
- DIALS-4B: October 2013 (Berkeley, CA)
- DIALS-5: November 2014 (Diamond Light Source, UK)
- DIALS-6: May 2015 (Berkeley, CA)

Program and presentations

- Historical Perspectives: “what did you do at the time and how would you do things differently today and why?” 4 x 30 minute presentations and discussion:
    - Andrew Leslie: 30 years of Mosflm development, what have we learnt?
    - Jim Pflugrath: d*TREK from 1995 to 2012: An Adventure in Object Oriented Diffraction Image Processing
    - Wolfgang Kabsch: XDS 1986 - ???? : retrospective
    - Kay Diederichs: Quality Control in Data Processing
- Challenging cases: “what are the challenges and why do you think they can be beaten?” 4 x 30 miniute presentations and discussion:
    - Tom White: CrystFEL: data processing for X-ray free electron lasers
    - Anastassis (Tassos) Perrakis: TBC
    - Frank Vondelft: TBC
    - Dave Stuart: TBC
- Ongoing developments: “how will these help to address the challenges?” 5 x 30 minute presentations followed by critical discussion:
    - Nick Sauter: Collaborative Software for Data Reduction
    - Graeme Winter: TBC
    - Gerard Bricogne: Meeting challenges through collaborative (vs. individualistic) software development: silver bullet, or waste of time?
    - Wolfgang Kabsch: XDS 1976 - ???? : future work
    - Loes Kroon-Batenburg: A different look at diffraction data with EVAL: towards understanding the complete diffraction pattern.
- Discussion and wrap-up

DIALS References

Background modelling of diffraction data in the presence of ice rings
Parkhurst JM, Thorn A, Vollmar M, Winter G, Waterman DG, Fuentes-Montero L, Gildea RJ, Murshudov GN, Evans G.
IUCrJ 4, 626-638 (01 Sep 2017). [PMID:28989718] [PMC reprint: PMC5619854]

Robust background modelling in DIALS
Parkhurst JM, Winter G, Waterman DG, Fuentes-Montero L, Gildea RJ, Murshudov GN, Evans G.
J Appl Crystallogr 49, 1912-1921 (01 Dec 2016). [PMID:27980508] [PMC reprint: PMC5139990]

Diffraction-geometry refinement in the DIALS framework
Waterman DG, Winter G, Gildea RJ, Parkhurst JM, Brewster AS, Sauter NK, Evans G.
Acta Crystallogr D Struct Biol 72, 558-75 (1 Apr 2016).

2016年4月　DIALS1.1リリース

New methods for indexing multi-lattice diffraction data
Gildea RJ, Waterman DG, Parkhurst JM, Axford D, Sutton G, Stuart DI, Sauter NK, Evans G, Winter G.
Acta Crystallogr D Biol Crystallogr 70, 2652-66 (1 Oct 2014)

dxtbx: the diffraction experiment toolbox
Parkhurst JM, Brewster AS, Fuentes-Montero L, Waterman DG, Hattne J, Ashton AW, Echols N, Evans G, Sauter NK, Winter G.
J Appl Crystallogr 47, 1459-1465 (01 Aug 2014). [PMID:25242914] [PMC reprint: PMC4119952]

Diffuse X-ray scattering to model protein motions
Wall ME, Adams PD, Fraser JS, Sauter NK.
Structure 22, 182-4 (04 Feb 2014). [PMID:24507780] [PMC reprint: PMC4070675] (Reprint)

New Python-based methods for data processing
Sauter NK, Hattne J, Grosse-Kunstleve RW, Echols N.
Acta Crystallogr D Biol Crystallogr 69, 1274-82 (1 Jul 2013).

Automatic Fortran to C++ conversion with FABLE
Grosse-Kunstleve RW, Terwilliger TC, Sauter NK, Adams PD.
Source Code Biol Med 7, 5 (28 May 2012). [PMID:22640868] [PMC reprint: PMC3448510]

2012年ミーティング開始

DIALS References

最近の文献が多く，活発に開発が進められていることが伺える。

xia2.multiplex: a multi-crystal data-analysis pipeline
Gildea RJ, Beilsten-Edmands J, Axford D, Horrell S, Aller P, Sandy J, Sanchez-Weatherby J, Owen CD, Lukacik P, Strain-Damerell C, Owen RL, Walsh MA, Winter G.
Acta Crystallogr D Struct Biol 78, 752-769 (01 Jun 2022)

DIALS as a toolkit
Winter G, Beilsten-Edmands J, Devenish N, Gerstel M, Gildea RJ, McDonagh D, Pascal E, Waterman DG, Williams BH, Evans G.
Protein Sci 31, 232-250 (1 Jan 2022). [PMID:34747533] [PMC reprint: PMC8740827]

Scaling diffraction data in the DIALS software package: algorithms and new approaches for multi-crystal scaling
Beilsten-Edmands J, Winter G, Gildea R, Parkhurst J, Waterman D, Evans G.
Acta Crystallogr D Struct Biol 76, 385-399 (01 Apr 2020).

How best to use photons
Winter G, Gildea RJ, Paterson NG, Beale J, Gerstel M, Axford D, Vollmar M, McAuley KE, Owen RL, Flaig R, Ashton AW, Hall DR.
Acta Crystallogr D Struct Biol 75, 242-261 (01 Mar 2019).

Improving signal strength in serial crystallography with DIALS geometry refinement
Brewster AS, Waterman DG, Parkhurst JM, Gildea RJ, Young ID, O'Riordan LJ, Yano J, Winter G, Evans G, Sauter NK.
Acta Crystallogr D Struct Biol 74, 877-894 (01 Sep 2018).

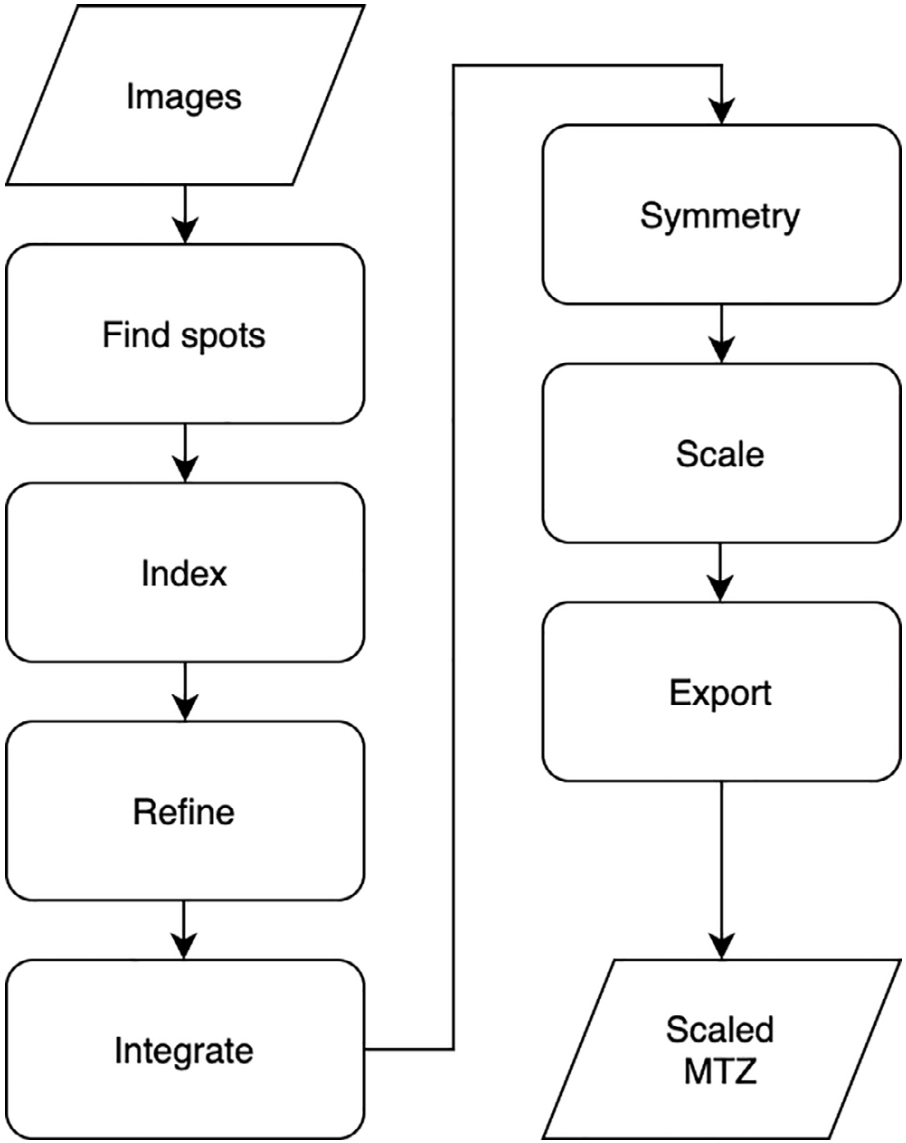
Electron diffraction data processing with DIALS
Clabbers MTB, Gruene T, Parkhurst JM, Abrahams JP, Waterman DG.
Acta Crystallogr D Struct Biol 74, 506-518 (01 Jun 2018)

Determination of Patterson group symmetry from sparse multi-crystal data sets in the presence of an indexing ambiguity
Gildea RJ, Winter G.
Acta Crystallogr D Struct Biol 74, 405-410 (01 May 2018).

DIALS: implementation and evaluation of a new integration package
Winter G, Waterman DG, Parkhurst JM, Brewster AS, Gildea RJ, Gerstel M, Fuentes-Montero L, Vollmar M, Michels-Clark T, Young ID, Sauter NK, Evans G.
Acta Crystallogr D Struct Biol 74, 85-97 (01 Feb 2018).

DIALSで行うこと

XDS, d*TREKなどの主流なデータ処理ソフトのワークフローと同じ。



Winter G, et al. “DIALS as a toolkit” *Protein Sci* 31, 232-250 (2022)

DIALSのワークフロー

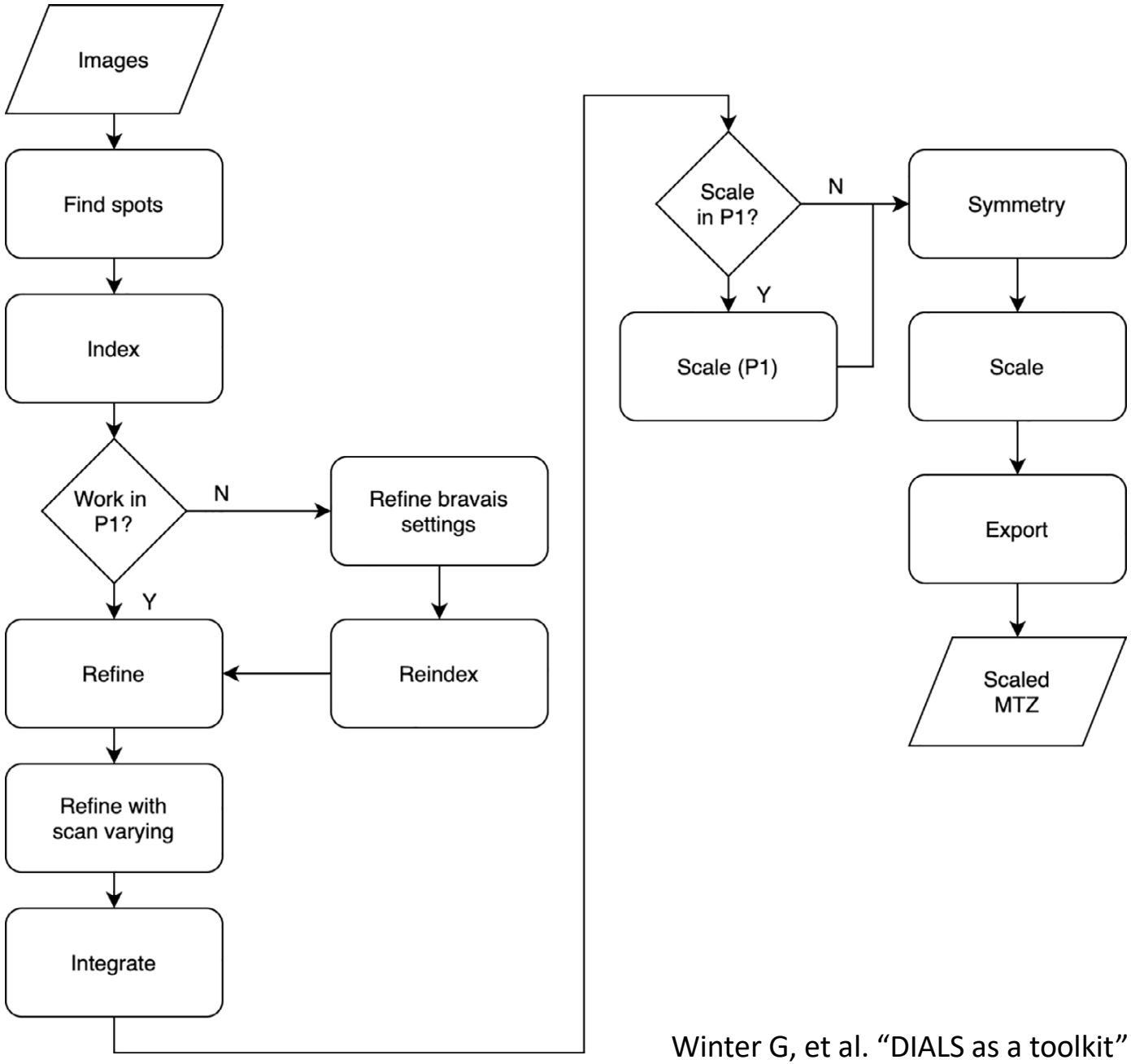
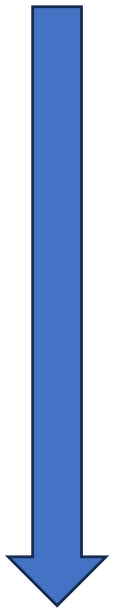


FIGURE 3 Simple workflow of data processing with DIALS, which deliberately follows the simple workflow shown in Figure 2

ワークフロー



Import
Find_spots
Index
Lattice
→Re-index
Refine
Integrate
Symmetry
Scale
Export

Winter G, et al. “DIALS as a toolkit” *Protein Sci* 31, 232-250 (2022)

ワークフローにおける各処理について

Import

EIGERの場合，******_master.h5　ファイルを読み込む。

Find spots

MOSFLMやHKL3000と違い，すべてのイメージに対してfind spotsがなされる。
→より完全なデータのモデル化を可能とする。

Indexing

まずP1でリファインして指数付け。

Lattice

空間群をリファイン。Re−indexウィンドウに候補が出る。
→最も対称性が高くmaxδが最大の空間群を選ぶ（best guess solution =**と出る）。

ワークフローにおける各処理について

Refine

最適な空間群でリファイン。格子定数，カメラ距離，結晶の方位などのパラメーターがリファインされる。

Integrate

デフォルトではUse profile fitting=True, Background algorithm=glm

profile fitting（高角の弱い反射データが解析可能）

Symmetry

より高い対称性を見つけるための計算。→Best solutionが示される。

Scale

scale.logをみてI/σ(I)やcc1/2 を指標にhigh res.を決める。High res.を入れてもう一度Scaleを行う。→2度目はDUI上での表示が赤くなりログファイルが見れなくなるが，処理はされている（バグ？）。ログファイルも格納されている。

Export

mtzファイルをつくる

一方で，XDS

- ✓ Wolfgang Kabsch (Max Planck Institute for Medical Research)が開発。
  1986年リリース

- ✓ Wolfgang Kabsch は2011年に退職している。

  SB Grid Consortium “Developper Tale”より
  https://sbgrid.org/software/tale/escape-from-the-darkroom

- ✓ Kay Diederichs (University of Konstanz)が後継者として2007年からXDSのメンテナンスを行っている。

- ✓ ソースコードは非公開

- ✓ アカデミア（非営利）は無償だが，企業の利用にはライセンスが必要。

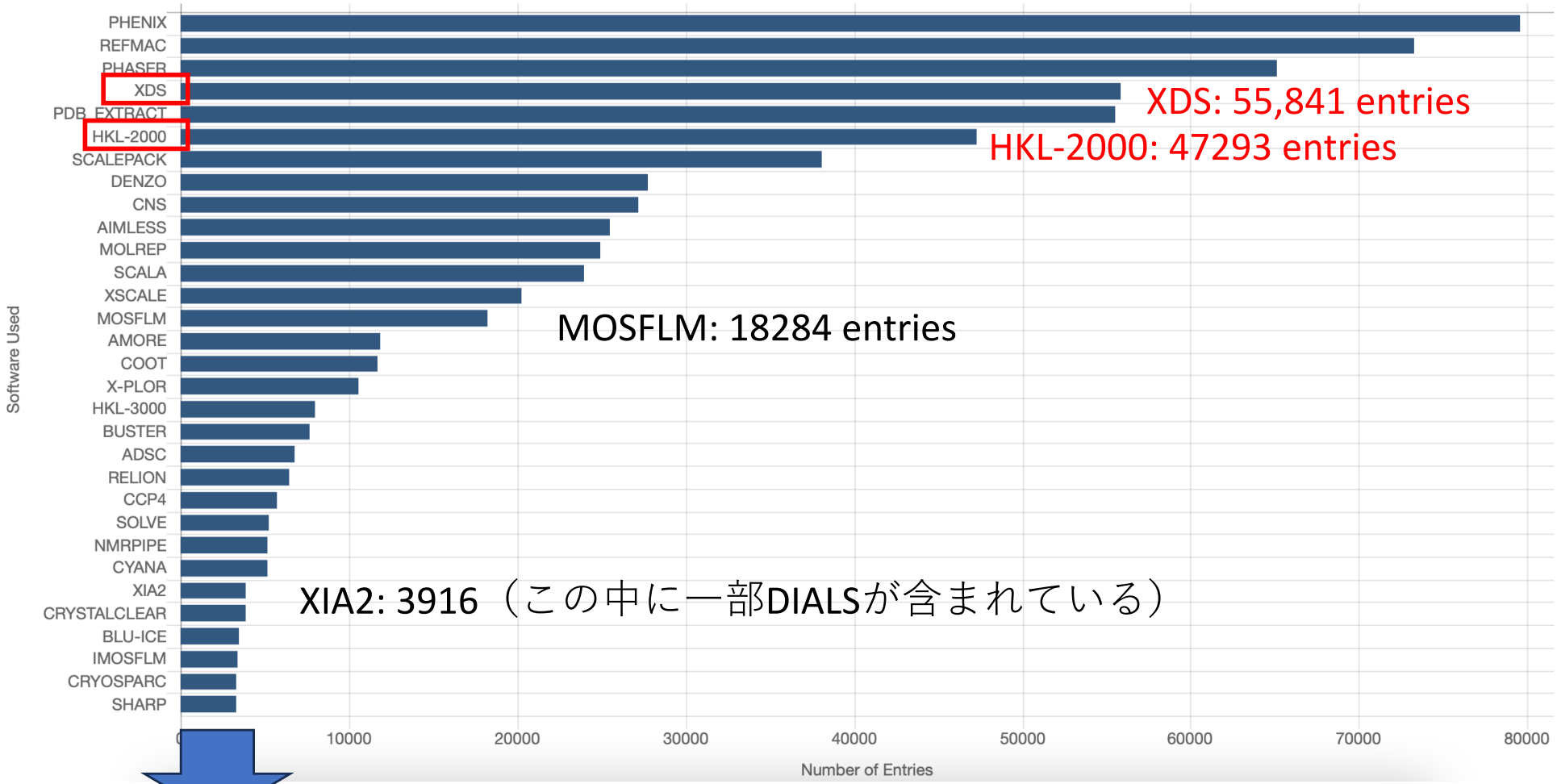
- ✓ 放射光施設で全自動測定のデータ処理プログラムに（標準的に？）
  採用されている。

- ✓ これまで最もよく使われてきたデータ処理ソフト。

XDSは現在まで圧倒的に使われてきた！

PDB Statistics: PDB Data Distribution by Processing Software

Data shown include structures solved by any experiment methods.



DIALS: 2,181 entries

PDB statistics, 2023年11月6日時点

XDS References

1.Kabsch, W. (1988a). *Automatic indexing of rotation diffraction patterns. J. Appl. Cryst.* **21**, 67-71.
2.Kabsch, W. (1988b). *Evaluation of single-crystal X-ray diffraction data from a position-sensitive detector. J. Appl. Cryst.* **21**, 916-924.
3.Kabsch, W. (1993). *Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. J. Appl. Cryst.* **26**, 795-800.
4.Kabsch, W. (2001a) Chapter **11.3.** *Integration, scaling, space-group assignment and post refinement* in *International Tables for Crystallography*, Volume F. Crystallography of Biological Macromolecules, Rossmann, M.G. and Arnold, E. (2001). Editors. Dordrecht: Kluwer Academic Publishers.
5.Kabsch, W. (2001b) Chapter **25.2.9.** *XDS* in *International Tables for Crystallography*, Volume F. Crystallography of Biological Macromolecules, Rossmann, M.G. and Arnold, E. (2001). Editors. Dordrecht: Kluwer Academic Publishers.

6.Kabsch, W. (2010a). *XDS. Acta Cryst.* D**66**, 125-132. → Key reference
7.Kabsch, W. (2010b). *Integration, scaling, space-group assignment and post refinement. Acta Cryst.* D**66**, 133-144.
8.Abrahams, J. P. (1993). *Compression of X-ray images* in Joint CCP4 and ESF-EACBM Newsletter on protein crystallography, no. 28 (May 1993), 3-4.
9.Diederichs, K. and Karplus, P. A. (1997). *Improved R-factors for diffraction data analysis in macromolecular crystallography. Nature Struct. Biol.* **4**, 269-275.
10.French, S. and Wilson, K. (1978). *On the treatment of negative intensity observations. Acta Cryst.* **A34**, 517-525.
11.Diederichs, K., McSweeney, S. and Ravelli, R.B.G. (2003). *Zero-dose extrapolation as part of macromolecular synchrotron data reduction. Acta Cryst.* **D59**, 903-909.
12.Karplus, P. A. and Diederichs, K. (2012). *Linking Crystallographic Model and Data Quality. Science* **336**, 1030-1033.
13.Schwarzenbach, D. and Flack, H.D. (1989). *On the definition and practical use of crystal-based azimuthal angles. J. Appl. Cryst.* **22**, 601-605.

https://xds.mr.mpg.de/html_doc/references.html

XDS_docやXDSwikiに記載されている文献は比較的古い。

XDSも継続してアップデートされている。

Release notes

Version June 30, 2023

BUILT=20230630

- Avoidance of strange characters in standard output upon reading images using the DECTRIS neggia library. (Problem reported by Aaron Finke).
- Maximum length of file names accepted by XDSCONV and XSCALE increased to 256 characters.
- Automatic removal of unused memory space in XSCALE.

Version January 10, 2022

BUILT=20220110

Apple Silicon processorへの適応

- Correction of a bug in the IDXREF step that could lead to an odd indexing failure (error reported by Clemens Vonrhein).
- Values for the input parameters TEST_RESOLUTION_RANGE=, MIN_RFL_Rmeas=, and MAX_FAC_Rmeas= as used in the CORRECT step of XDS are shown in CORRECT.LP. Default values for these parameters are defined in the documentation.
- An additional version of the XDS package, XDS-Apple_M1, is provided for running natively under macOS Big Sur (or newer) on an Apple Silicon processor.

BUILT=20220120

- Correction of a numerical problem in controlling the acceptance of reflections in the CORRECT step of XDS by the input parameter INCLUDE_RESOLUTION_RANGE=. (problem reported by Jan Wollenhaupt).

BUILT=20220220

- Correction of a type mismatch in the routine linking the input of data images to an external library routine.
- Correction of reported indices for the difference vector clusters in the special 2-dimensional case in the IDXREF step.
- Exchange of 3 instructions to make the source code of the XDS package strictly obeying the Fortran 90 language definition.

BUILT=20220820

- Correction of an error that might allocate processor threads to an empty set of images in the COLSPOT step.
- Optimized assignment of threads for a large multi-core processor system during the COLSPOT step.
- Change of the meaning of the correlation factors between data sets reported by XSCALE. Previously, these factors indicated the success of radiation damage corrections as correlation between mean intensities of data sets as function of resolution. The correlation factors now reported are calculated between intensities of equivalent reflections from each data set pair which is the more intuitive view (pointed out by Oliver Weiergräber and Kay Diederichs).

Version February 5, 2021

XDSも継続してアップデートされている。

2023年

JOURNAL OF APPLIED CRYSTALLOGRAPHY

ISSN 1600-5767

XDSGUI: a graphical user interface for *XDS*, *SHELX* and *ARCIMBOLDO*

Wolfgang Brehm,[a]‡ Josep Triviño,[b]‡ Juno M. Krahn,[c] Isabel Usón[b,d]* and Kay Diederichs[e]*

[a]Department of Physics, University of Hamburg, Hamburg 22761, Germany, [b]Instituto de Biologia Molecular de Barcelona (IBMB), Baldiri Reixach 15, Barcelona 08028, Spain, [c]Genome Integrity and Structural Biology Laboratory, National Institute of Environmental Health Sciences, Research Triangle Park, NC 27709, USA, [d]ICREA: Institució Catalana de Recerca i Estudis Avançats, Pg. Lluis Companys 23, Barcelona 08010, Spain, and [e]Department of Biology, University of Konstanz, Universitätsstrasse, Konstanz 78457, Germany. *Correspondence e-mail: uson@ibmb.csic.es, kay.diederichs@uni-konstanz.de

Received 5 June 2023
Accepted 8 August 2023

Edited by A. Barty, DESY, Hamburg, Germany

‡ These authors contributed equally.

Keywords: X-ray diffraction; neutron diffraction; electron diffraction; data processing; graphical user interfaces; phasing; *XDS*; *ARCIMBOLDO*; *SHELX*.

XDSGUI is a lightweight graphical user interface (GUI) for the *XDS*, *SHELX* and *ARCIMBOLDO* program packages that serves both novice and experienced users in obtaining optimal processing and phasing results for X-ray, neutron and electron diffraction data. The design of the program enables data processing and phasing without command line usage, and supports advanced command flows in a simple user-modifiable and user-extensible way. The GUI supplies graphical information based on the tabular log output of the programs, which is more intuitive, comprehensible and efficient than text output can be.

演習の目的

XDSは圧倒的に使われてきたソフトだが,
開発の将来について心配？

1つのソフト（XDS）に頼らずに,
DIALSを使ってみましょう。

DIALSの実行

・本来はコマンドラインベース。
・GUI（DUI），自動処理（xia2_DIALS）が便利。

本日の実習では,

１，DUI（DIALSのGUIバージョン）を使ったデータ処理。

→個別のプログラムで順次解析。

２，xia2_DIALSを使ったデータ処理。

→全自動処理。

（時間があれば）xia2_XDSで処理したデータと比較してみましょう。