DIALS

Gwyndaf Evans

DIALS
Diffraction Integration for Advanced Light Sources

diamond

Graeme Winter

David Waterman

James Parkhurst

Luis Fuentes Montero

Richard Gildea

Sauter group, Berkley

DIALS 6 workshop, Berkeley, CA.

and …
Phil Evans
Andrew Leslie
Garib Murshudov
Harry Powell
MRC-LMB

DIALS
Diffraction Integration for Advanced Light Sources

diamond

DIALS: what is it?
(Diffraction Integration for Advanced Light Sources)

- integration software for SR and FEL MX
- a development environment for new algorithms and approaches to integration
- correct treatment of physical processes in detection
  - to produce better error estimates for downstream analysis
- keep pace with data collection rates
  - when we started 100 fps was fairly exciting – now we are facing kHz frame rates!
  - even higher with FEL sources
- modular design to permit easy developmental upgrades
  - e.g. dealing rapidly with new detectors and other advances, CSPAD, curved Pilatus 12M
- collaborative development with open source license
  - built on cctbx and using Python/C++
- Major collaboration with Nick Sauter at LBNL, Berkley CA.

http://dials.diamond.ac.uk
http://dials.lbl.gov

Why did/do we need a new data analysis package?

- MOSFLM, XDS, HKL2000, d*trek, EVAL15 …
- Difficult to code into, one-man show, proprietary, alternative algorithm, key people at retirement age!!!!!!
  - need to create a new generation of experts!
- Desire to develop and test new ideas and algorithms
- Felt the time was right to take a fresh look at data analysis given many of the programs are > 20 yrs old (even though they are constantly developed)
- Unify the analysis of SR and XFEL data!
  - Structural biology community would benefit from an single-point of entry into such analysis
  - they have lots in common!

Motivation

Journal of
Applied Crystallography

ISSN 0021-8898

Received 11 December 2009
Accepted 18 August 2010

Estimation of errors in diffraction data measured by CCD area detectors

David Waterman and Gwyndaf Evans*

Diamond Light Source, Harwell Science and Innovation Campus, Didcot, Oxfordshire OX11 0DE, England. Correspondence e-mail: gwyndaf.evans@diamond.ac.uk

Current methods for diffraction-spot integration from CCD area detectors typically underestimate the errors in the measured intensities. In an attempt to understand fully and identify correctly the sources of all contributions to these errors, a simulation of a CCD-based area-detector module has been produced to address the problem of correct handling of data from such detectors. Using this simulation, it has been shown how, and by how much, measurement errors are underestimated. A model of the detector statistics is presented and an adapted summation integration routine that takes this into account is shown to result in more realistic error estimates. In addition, the effect of correlations between pixels on two-dimensional profile fitting is demonstrated and the problems surrounding improvements to profile-fitting algorithms are discussed. In practice, this requires knowledge of the expected correlation between pixels

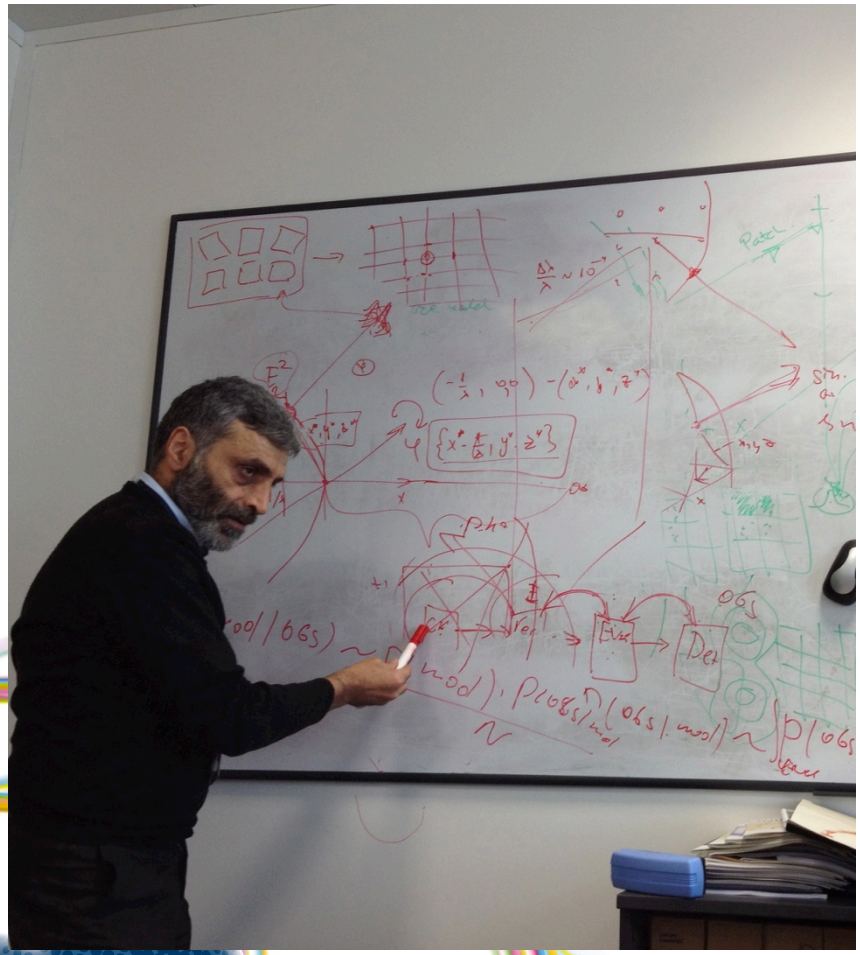
Fundamental problems needed to be addressed

- Error analysis and propagation
- Algorithm selection
  - 2D, 2.5D vs 3D case by case
  - there appears to be no ‘one-size-fits-all’ solution (yet)
- Time/dose dependence of crystal parameters
  - radiation damage was being managed at scaling (to some extent) but not at integration level
  - global approach to refinement
- Multicrystal analysis
  - these should not necessarily be treated as independent jobs
  - for refinement, indexing etc…..

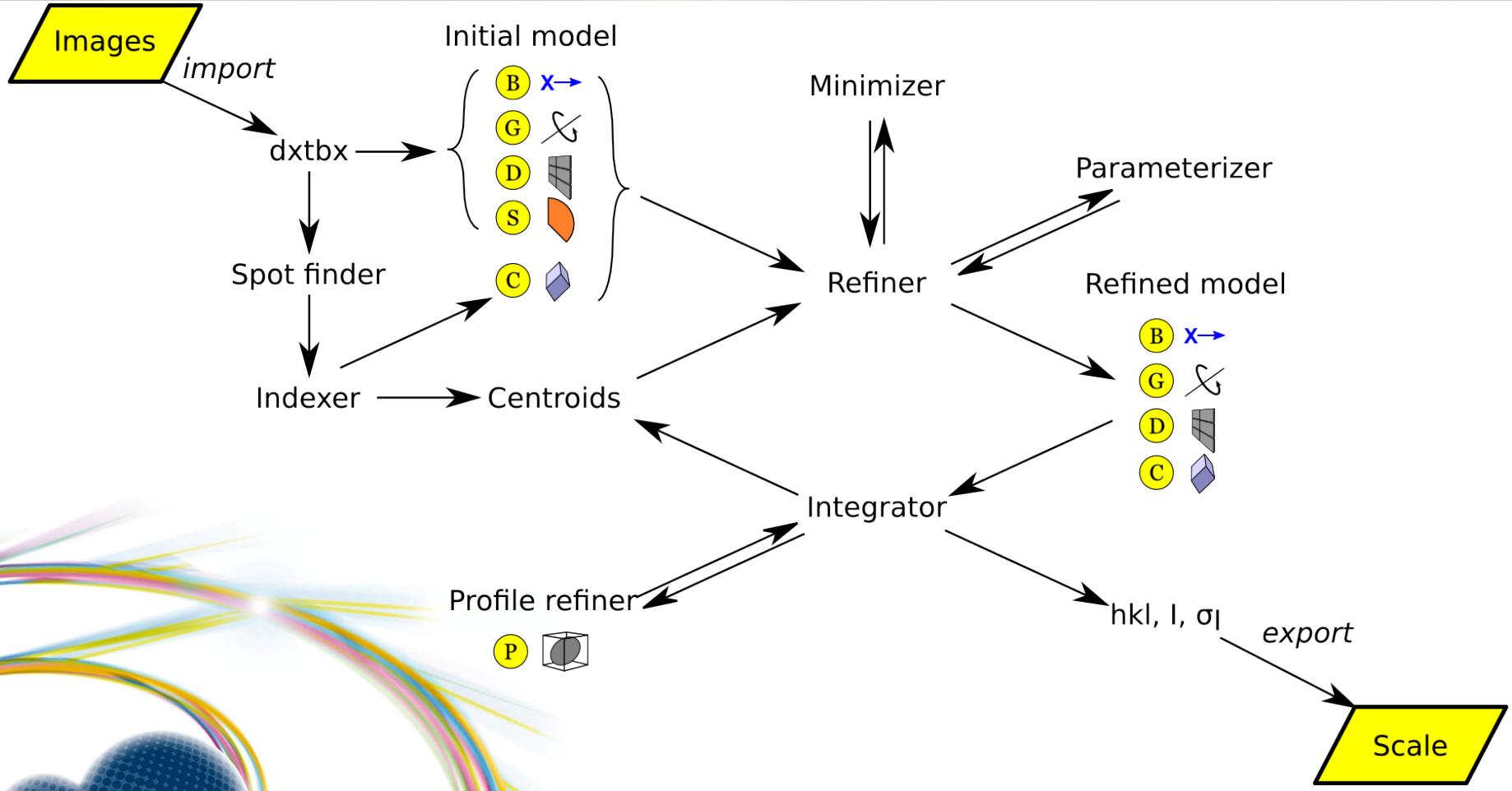
Support in Europe

- Funded by EU Framework 7 as BioStruct-X (Work Package 6 – Evans/Chapman)
  - 2011 – 2015
  - my third attempt at getting funding for a new data analysis package!!
- CCP4 support in the form of David Waterman
  - DIALS will replace Mosflm as main data analysis package in CCP4
- Diamond Light Source matched the EU funding giving us one extra post

Early planning meetings



DIALS


Images
import
dxtbx
Initial model
B
G
D
S
C
Spot finder
Indexer
Centroids
Minimizer
Refiner
Parameterizer
Refined model
Integrator
Profile refiner
P
hkl, I, σI
export
Scale

Impact to date

- DIALS is now embedded in the Diamond data analysis scheme
  - is the default for Xia2
- Raised the bar for other integration programs
- New collaborations
  - US and European SR sources
  - Global Phasing
- Keen interest from Neutron and electron diffraction communities

Which of our goals have been achieved?

- Modular design and open source
  - major opportunity for training a younger generation of developers (proving the point already in-house)
- Relatively easy to adapt to new detectors
  - e.g. I23 12M detector could be supported now, CSPAD also
- Advanced refinement
  - time/dose dependent variation of crystal parameters
  - better unit cell dimension estimation (critical for BLEND, multicrytsal refinement in REFMAC etc.)
  - can refine multiple lattices/crystals with same beamline/detector geometry (or other combinations of this)
- Uses both 2D, 3D or other algorithms
- Can be used to integrate XFEL data
  - sort of… but parts of DIALS are being used by cctbx.xfel
  - this will be made more standard in the coming years and form part of the Diamond XFEL Hub delivery of data analysis

What's missing

- User friendly GUI and visualization
  - CCP4 have proposed funding me to develop a GUI
  - We'll start this at Diamond and see how it develops
- Algorithms
  - we're just scratching the surface of the ideas we had initially
  - this is now starting
- More collaborators, contributors and users!
- ……