HKL-3000

Toward the future of protein crystallography
Long and winding road (with traps)

Phasing

Data collection

Data reduction

PDB

Protein production crystallization

Project start
SG dynamics

In every PSI Large Center

Roughly

new structure every 48 hours

Argonne Users Meeting, May 2008
W0308

Teaching Elves to Collect Data: An Analysis of the Last Million Diffraction Images from ALS 8.3.1.
James Holton, Physical Biosciences, Lawrence Berkeley National Laboratory, Berkeley, CA 94720.

Most X-ray data sets collected at synchrotron sources do not produce usable results. An analysis of data collected in 2003 at the ALS beamline 8.3.1 shows that 2346 datasets were collected and 41 structures were deposited in the PDB. Although it is understandable that not every dataset leads to a published structure, it is troubling that ~98% of them do not. This large gap between collected data and useful results is not unique to 8.3.1. The 28 operating American PX beamlines collect ~100,000 datasets/year. This suggests that a great deal of improvement in scientific productivity can be attained if the reasons for failed projects are better understood.

57 datasets/deposit !!!!
Worldwide < 40 hours/deposit
~ 40-150 crystals/structure

Argonne Users Meeting, May 2008
Priceless crystals

Star of Africa

Crystallography lab

How to define high throughput at the beamline?

Argonne Users Meeting, May 2008
Integration

HKL-3000 is a pipeline component

Pepdb
expression & purification

Xtaldb
crystallization

HKL-3000
Data collection, diffraction & structure solution

Other systems provide additional contextual up- and downstream information

Wetlab
chemicals & solutions

Argonne Users Meeting, May 2008
Project
Check against DB during data collection

HKL-2000
SHELXD, SHELXE
CCP4
SOLVE, RESOLVE
ARP/WARP
O, COOT, CCP4
Completeness – overlaps
Multi-crystal strategy
Space Group Determination

Argonne Users Meeting, May 2008
P2?2?2?
Scaling

Signal from Sulfur

Argonne Users Meeting, May 2008
Data management in HKL-3000 - substructure solution
NCS handling

Argonne Users Meeting, May 2008
Model building and preliminary refinement (10 minutes, 190aa)

Secondary structure after each cycle

Unidentified density -> Xtaldb
MCSG 712 [HKL-3000]

HKL-2000/HKL-3000
as reported for MCSG structures
Low resolution SAD phasing

difficult structure – no

1ze0

Argonne Users Meeting, May 2008
Structures solved in wrong target

APC81521 (ChangSoo)

To figure out the right target at the resolution of 2.5Å

1. Manually build helixes and stands;
2. Submit this model to dali to find out several similar structures in PDB;
3. Search targetDB(MCSG) using the sequence of these structures to find out the right target.

Argonne Users Meeting, May 2008
Twinning

Argonne Users Meeting, May 2008
Xtaldb – crystallization database
TM1086 (APC4579)
TM1086 (APC4579)

10 Se-Met/282 aa
35.1kDa
I222.

a = 290Å, b = 300Å, c = 316Å

~81% solvent content = 17 molecules in AU
~38% solvent content = 57 molecules in AU
Argonne Users Meeting, May 2008

TM1086 (APC4579)

282 aa
35.1 kDa
P3221
a = 165 Å, b = 165 Å, c = 98 Å

~78% solvent content = 3 molecules in AU
~40% solvent content = 8 molecules in AU
I222 crystal form

~210 Å

Argonne Users Meeting, May 2008
I222 crystal form

30 chains in AU
8460 aa
1.05MDa

~210 Å

Argonne Users Meeting, May 2008
Fast refinement (full control)

Argonne Users Meeting, May 2008
Unidentified densities?

Discussion

Although homology of both PHI and NR5A proteins are found extensively in all three kingdoms of life, the function of NR5A and its homologs that contain both PHI-like and NR5-like domains remain completely unknown. The fusion of the two proteins in one peptide chain indicates a functional coupling of these two proteins. The structural features revealed in the present study, such as the caps-like hexameric unit structure with its NR5 domains as solubilized and the two PHI-like domain trimers as helices, the flexible tail stretch, and the centrally bound ligand, may provide certain clues of their functional roles. This fusion role of the estrogen regulatory PHI protein is to integrate various intracellular proteins and signaling signals by regulating expression involving estrogen metabolism [10]. PHI exerts its regulatory effects by undergoing different post-translational modifications, such as acetylation [11] and phosphorylation [12] by various modifying enzymes in response to the primary cell regulatory signals. The PHI protein and PHI-like proteins have been shown to bind growth factors [13, 14], such as transforming growth factor beta (TGF-β), interleukin 1 (IL-1), and interleukin 6 (IL-6). It has been shown that some inflammatory cytokines, such as interleukin 1 (IL-1), stimulate the expression of PHI in the liver [15], and the nuclear factor kappa B (NF-κB) is involved in the regulation of PHI expression. The PHI protein has been shown to bind growth factors [16], such as transforming growth factor beta (TGF-β), interleukin 1 (IL-1), and interleukin 6 (IL-6). It has been shown that some inflammatory cytokines, such as interleukin 1 (IL-1), stimulate the expression of PHI in the liver [15], and the nuclear factor kappa B (NF-κB) is involved in the regulation of PHI expression. 

Unlike PHI domains, the NR5-like domains have only recently been identified to play a regulatory role in the regulation of expression of the PHI protein. The NR5-like domains are ubiquitously found in all three kingdoms of life, whereas the PHI-like domains are exclusively found in eukaryotes. This suggests that the NR5-like domains are more ancient in evolution and may play a regulatory role in the expression of the PHI protein. 

Argonne Users Meeting, May 2008
Unidentified densities
Fast refinement
Unidentified density
Integration

**HKL-3000** is a pipeline component

*Pepdb*
expression & purification

*Xtaldb*
crystallization

**HKL-3000**
Data collection, diffraction & structure solution

*Wetlab*
chemicals & solutions

Other systems provide additional contextual up- and downstream information

Argonne Users Meeting, May 2008
Ligand libraries

Argonne Users Meeting, May 2008
Automatic ligand building and refinement
Unidentified densities
Automatic building
Fully refined ligand
Structure validation

Argonne Users Meeting, May 2008
Ca-O distances

A. $\text{Ca}^{2+}$ - O (CSD)

B. $\text{Ca}^{2+}$ - O$_{\text{water}}$ (CSD)

C. $\text{Ca}^{2+}$ - O (PDB - HR)

D. $\text{Ca}^{2+}$ - O$_{\text{water}}$ (PDB - HR)

E. $\text{Ca}^{2+}$ - O (PDB - MR)

F. $\text{Ca}^{2+}$ - O$_{\text{water}}$ (PDB - MR)

Number of PDB deposits vs resolution

Argonne Users Meeting, May 2008
Metal ion in Der p 1

Mg$^{2+}$

Ca$^{2+}$

Argonne Users Meeting, May 2008
Incorrect assignments 1kyt
Correct assignment
suboptimal refinement
B factors

![Graph showing the relationship between Ca B-factor (Å²) and environment B-factor (Å²).]
Absolute configuration and small molecule agents
HKL-3000SM

Argonne Users Meeting, May 2008
People involved

Wladek Minor
- Maks Chruszcz
- Marcin Cymborowski
- Matt Zimmerman
- Heping Zheng
- Marek Grabowski
- Haijun Qiu
- Olga Kirillova
- John Raynor

Andrzej Joachimiak
- Rongguang Zhan
- Marianne Cuff
- Chris Lazarski
- SBC staff

Zbyszek Dauter

Zbyszek Otwinowski
- Dominika Borek
- Andrzej Kudlcki
- Roger Ljungberg

Grants:
NIH GM53163, GM 62414,
DOE, NCI
HKL Research

Argonne Users Meeting, May 2008