

PFビームラインでのXDS自動 処理とその活用

山田悠介

KEK IMSS PF/SBRC

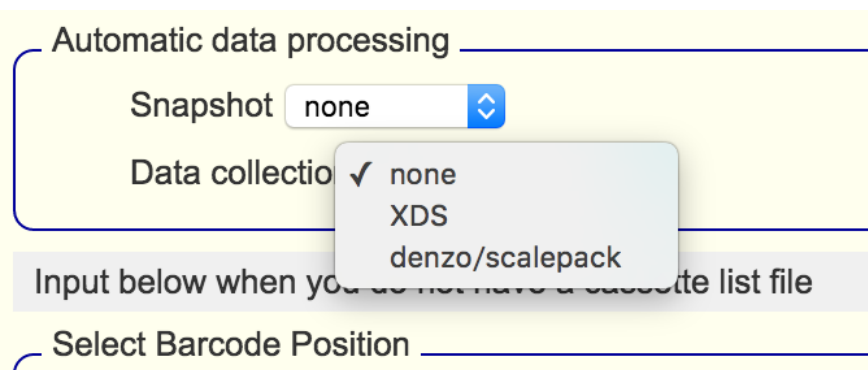
自動回折データ処理

回折実験終了後に自動的に処理開始

2つのソフトウェアを利用した独自スクリプトが利用可能

- DENZO/SCALEPACK(HKL2000)
 - 空間群情報が必要
(サンプルリストファイル)

- XDS
 - 空間群の情報は必要ない



Automatic data processing

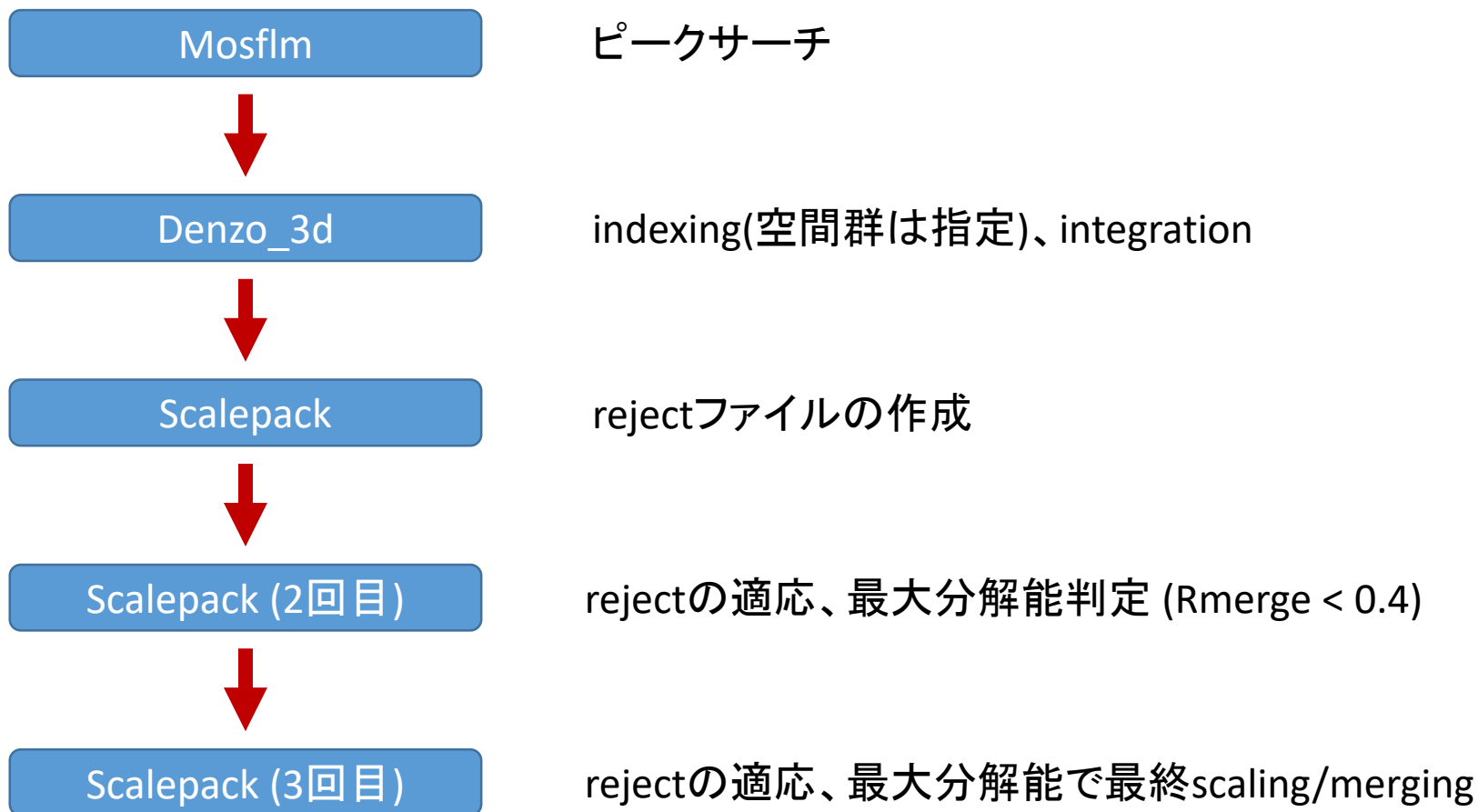
Snapshot

Data collection none
 XDS
 denzo/scalepack

Input below when you do not have a cassette list file

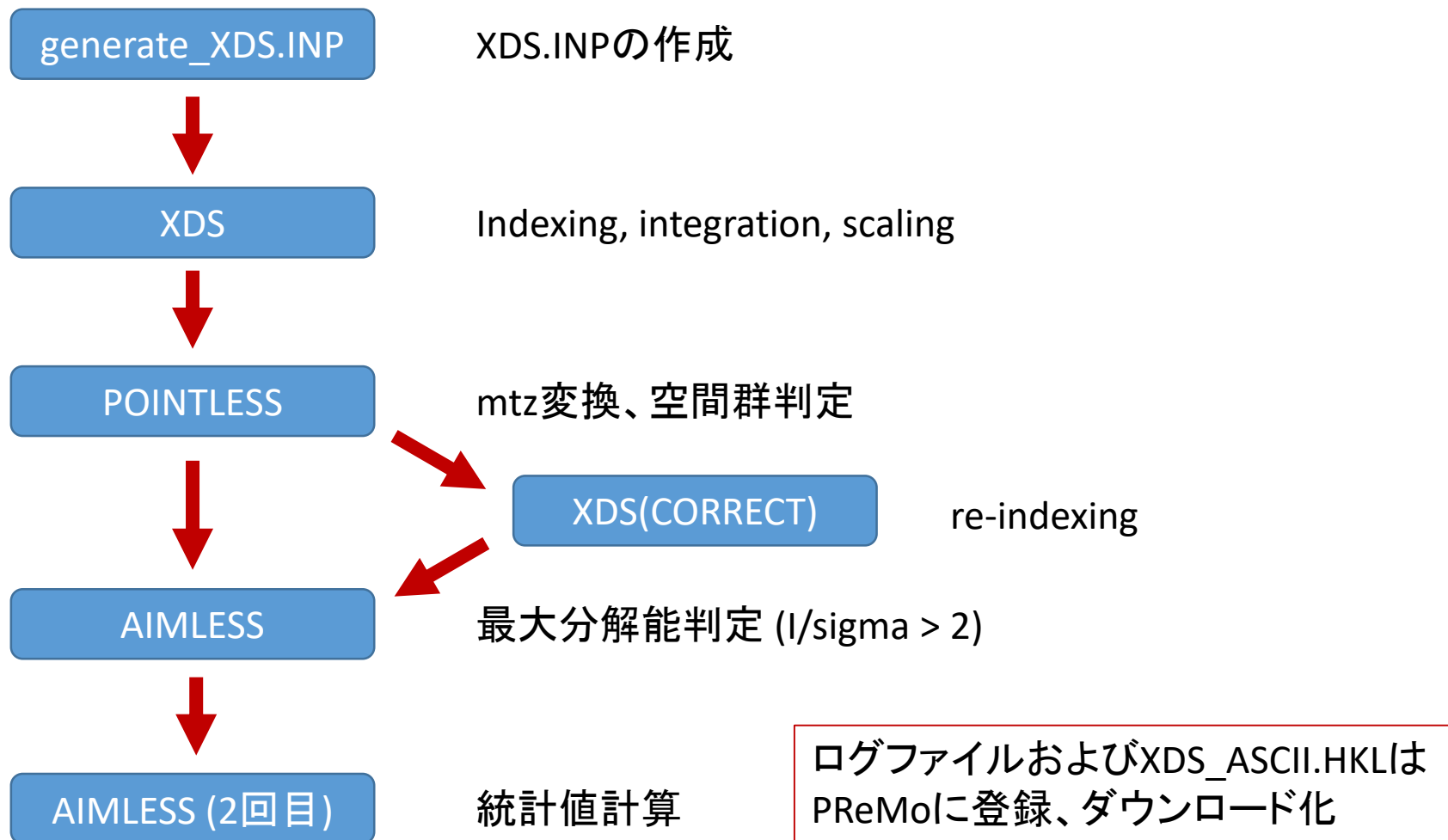
Select Barcode Position

DENZO/SCALEPACK



ログファイルおよび.scaファイルは
PReMoに登録、ダウンロード化

XDSを用いた自動処理



PReMo: 実験データ管理データベース

- ビームラインでの回折実験に関する情報をデータベースに格納
 - 試料情報
 - 実験条件
 - 実験結果(回折イメージ等)
 - データ処理結果
- Webで閲覧
 - <https://premo.kek.jp/RCM-Web/>



PReMo: 実験データ管理データベース

実験リスト一覧




Year: 2015

Beamtime: AR-NE3A
Date: 2015-12-04
Capture: 75
Snapshot: 72
Runs: 7
Mca: 0
Xaxis: 0
Oxiscan: 0

Time	Experiment	A	ω	Osc.	Exp.	Max.Resp.	Directory	Filename
15:28:00	Exchange							Dismounted: PUCK-2 / 8
15:24:41	Runs	1.0000	0.0-180.0	1.0	1.0	1.63	2015-12-04_AR-NE3Arun/PUCK-2/08/osc	collect_01
15:24:29	Capture		90.0				2015-12-04_AR-NE3Arun/PUCK-2/08/crystal	90.bmp
15:24:27	Capture		45.0				2015-12-04_AR-NE3Arun/PUCK-2/08/crystal	45.bmp
15:24:26	Capture		0.0				2015-12-04_AR-NE3Arun/PUCK-2/08/crystal	0.bmp
15:23:28	Exchange							Mounted: PUCK-2 / 8 Dismounted: PUCK-2 / 4

サンプルリスト一覧



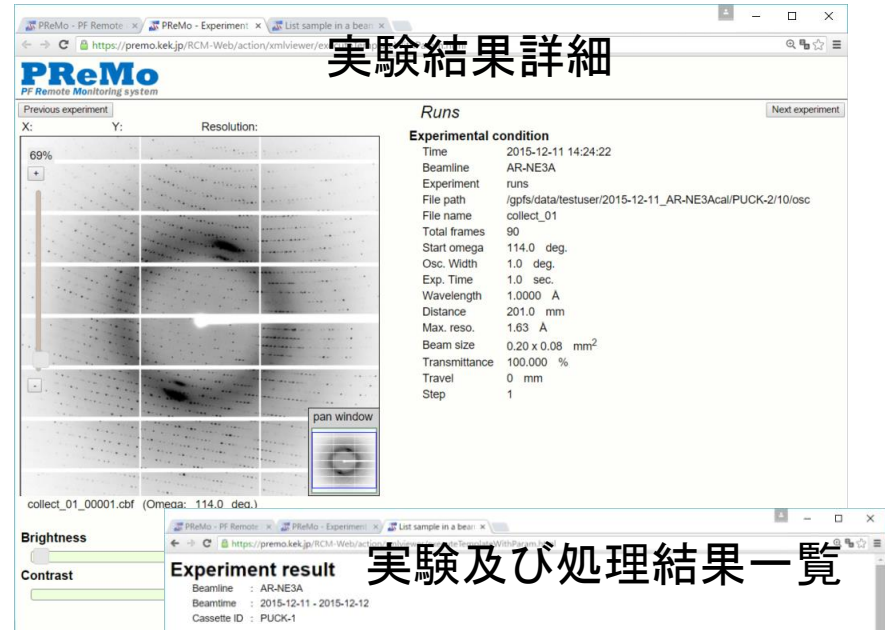
Year: 2015

Beamtime: AR-NE3A
Start date: 2015-12-11
End date: 2015-12-12
Status: Finished

Containers: PUCK-1, PUCK-2

No	ContainerID	Port	Barcode	Protein	CrystalID	Snapshot	Runs	Comment
1	PUCK-1	1		Thamatin	Sample1-01	Done		
2	PUCK-1	2		Thamatin	Sample1-02	Done		
3	PUCK-1	3		Thamatin	Sample1-03	Done		
4	PUCK-1	4		Thamatin	Sample1-04	Done		
5	PUCK-1	5		Thamatin	Sample1-05	Done		
6	PUCK-1	6		Lysozyme	Sample1-06	Done		
7	PUCK-1	7		Lysozyme	Sample1-07	Done		
8	PUCK-1	8		Lysozyme	Sample1-08	Done		
9	PUCK-1	9		Lysozyme	Sample1-09	Done		
10	PUCK-1	10		Lysozyme	Sample1-10	Done		
11	PUCK-2	1		Thamatin	Sample1-01	Done		
12	PUCK-2	2		Thamatin	Sample1-02	Done		
13	PUCK-2	3		Thamatin	Sample1-03	Done		

実験結果詳細



Resolution: X: Y: Resolution:

69%

collect_01_00001.cbf (Omega: 114.0 deg.)

Brightness

Contrast

Experiment result

Beamline : AR-NE3A
Beamtime : 2015-12-11 - 2015-12-12
Cassette ID : PUCK-1

Runs

Experimental condition

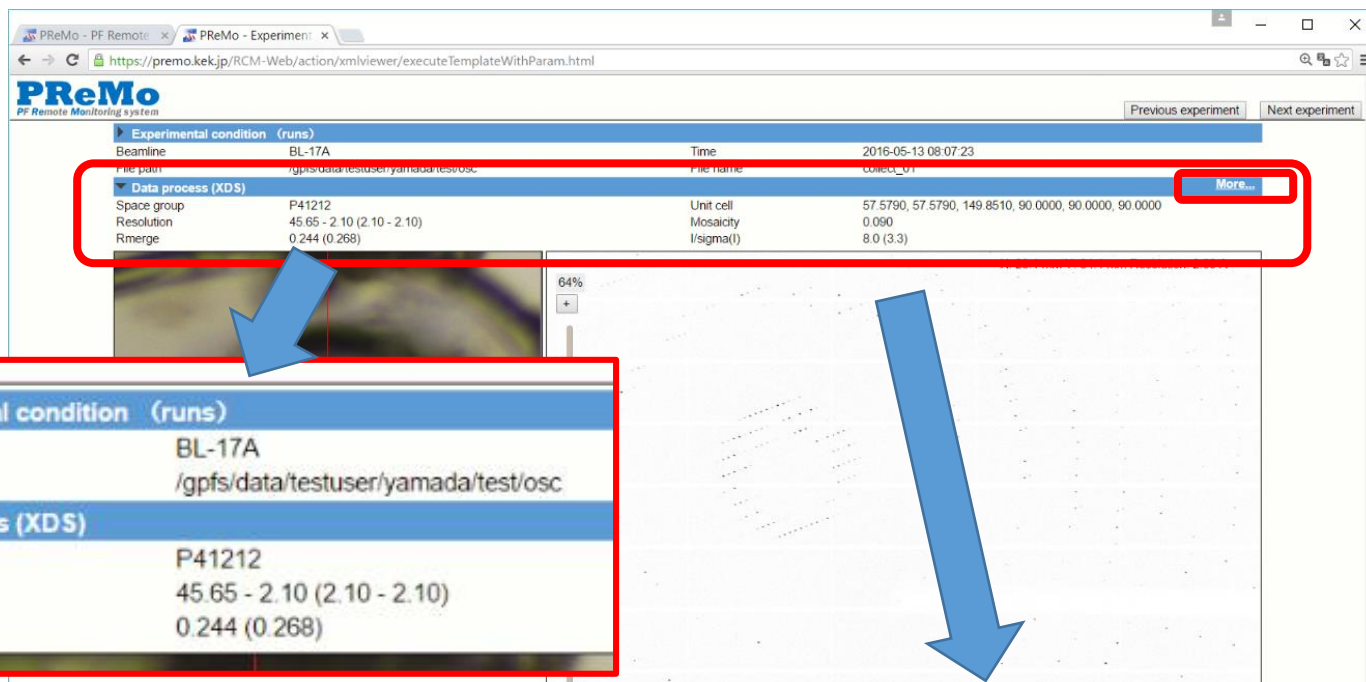
Time: 2015-12-11 14:24:22
Beamline: AR-NE3A
Experiment: runs
File path: /gpts/data/testuser/2015-12-11_AR-NE3Acal/PUCK-2/10/osc
File name: collect_01
Total frames: 90
Start omega: 114.0 deg.
Exp. Width: 1.0 deg.
Exp. Time: 1.0 sec.
Wavelength: 1.0000 A
Distance: 201.0 mm
Max. reso: 1.63 A
Beam size: 0.20 x 0.08 mm²
Transmittance: 100.000 %
Travel: 0 mm
Step: 1

実験及び処理結果一覧



No	Port	Crystal ID	Protein Name	Directory	S	Capture	SS	Diffraction	Data statistics			
									Total	Index	Index	
1	1				S	Capture	SS	Diffraction	Resolution	43.02	43.02	1.48
									Range	0.000	0.040	0.100
									Integral	20.4	40.7	3.7
									Completeness	90.7	90.5	42.3
2	2				S	Capture	SS	Diffraction	Resolution	43.00	43.00	1.52
									Range	0.001	0.020	0.400
									Integral	20.0	40.0	3.0
									Completeness	97.0	90.5	83.7
3	3				S	Capture	SS	Diffraction	Resolution	43.00	43.00	1.48
									Range	0.001	0.020	0.400
									Integral	22.7	27.0	2.5
									Completeness	93.0	90.5	42.3
4	4				S	Capture	SS	Diffraction	Resolution	43.00	43.00	1.48
									Range	0.000	0.040	0.100
									Integral	22.7	27.0	2.5
									Completeness	93.0	90.5	42.3
5	5				S	Capture	SS	Diffraction	Resolution	43.00	43.00	1.48
									Range	0.001	0.020	0.400
									Integral	20.0	40.0	3.0
									Completeness	90.0	90.5	42.3
6	6				S	Capture	SS	Diffraction	Resolution	43.00	43.00	1.48
									Range	0.000	0.040	0.100
									Integral	20.0	40.0	3.0
									Completeness	90.0	90.5	42.3
7	7				S	Capture	SS	Diffraction	Resolution	43.00	43.00	1.48
									Range	0.001	0.020	0.400
									Integral	20.0	40.0	3.0
									Completeness	90.0	90.5	42.3

回折データ処理結果表示(I)



Experimental condition (runs)

Beamline	BL-17A	Time	2016-05-13 08:07:23
File path	/gpfs/data/testuser/yamada/test/osc	File name	collect_01

Data process (XDS)

Space group	P41212	Unit cell	57.5790, 57.5790, 149.8510, 90.0000, 90.0000, 90.0000
Resolution	45.65 - 2.10 (2.10 - 2.10)	Mosaicity	0.090
Rmerge	0.244 (0.268)	I/sigma(I)	8.0 (3.3)

ing system

Experimental condition (runs)

Beamline	BL-17A
File path	/gpfs/data/testuser/yamada/test/osc

Data process (XDS)

Space group	P41212
Resolution	45.65 - 2.10 (2.10 - 2.10)
Rmerge	0.244 (0.268)

Time	2016-05-13 08:07:23
File name	collect_01
More...	
Unit cell	57.5790, 57.5790, 149.8510, 90.0000, 90.0000, 90.0000
Mosaicity	0.090
I/sigma(I)	8.0 (3.3)

回折データ処理結果表示(II)

PReMo - PF Remote x PReMo - Experiment x PReMo - Processing x
 https://premo.kek.jp/RCM-Web/action/xmlviewer/executeTemplateWithParam.html

PReMo
 PF Remote Monitoring system

Summary Log Plot Files

COLSPOT LP
 CORRECT LP
 DEFPX LP
 IDXREF LP
 INTLP
 INTEGRATE LP
 XDSSTAT LP
 XYCORR LP
aimless.log
 aimless.log.1
 pointless.log
 process.log.o29075
 xds_auto.log

```

180 359 0.966 0.336 0.342 0.350 0.359 0.371 0.391 0.392 0.399 0.358
$$$
<!--SUMMARY_BEGIN--> $TEXT:Result: $$$
Summary data for Project: XDSproject Crystal: XDScrystal Dataset: XDSdataset
Overall InnerShell OuterShell
Low resolution limit 45.65 45.65 2.16
High resolution limit 2.10 8.91 2.10
Rmerge (within I+/I-) 0.235 0.218 0.241
Rmerge (all I+ and I-) 0.244 0.224 0.268
Rmeas (within I+/I-) 0.255 0.236 0.315
Rmeas (all I+ & I-) 0.256 0.236 0.319
Rpin (within I+/I-) 0.099 0.089 0.200
Rpin (all I+ & I-) 0.074 0.072 0.168
Rmerge in top intensity bin 0.227 - -
Total number of observations 156852 2562 2691
Total number unique 14969 269 901
Mean(I)/sd(I) 8.0 8.1 3.3
Mn(I) half-set correlation CC(1/2) 0.973 0.967 0.857
Completeness 96.6 99.6 75.1
Multiplicity 10.5 9.5 3.0
Anomalous completeness 94.4 100.0 60.6
Anomalous multiplicity 5.6 6.9 1.8
DelAnom correlation between half-sets 0.069 0.238 0.056
Mid-Slope of Anom Normal Probability 0.989 - -
    
```

	Overall	InnerShell	OuterShell
Low resolution limit	45.65	45.65	2.16
High resolution limit	2.10	8.91	2.10
Rmerge (within I+/I-)	0.235	0.218	0.241
Rmerge (all I+ and I-)	0.244	0.224	0.268
Rmeas (within I+/I-)	0.255	0.236	0.315
Rmeas (all I+ & I-)	0.256	0.236	0.319
Rpin (within I+/I-)	0.099	0.089	0.200
Rpin (all I+ & I-)	0.074	0.072	0.168
Rmerge in top intensity bin	0.227	-	-
Total number of observations	156852	2562	2691
Total number unique	14969	269	901
Mean(I)/sd(I)	8.0	8.1	3.3
Mn(I) half-set correlation CC(1/2)	0.973	0.967	0.857
Completeness	96.6	99.6	75.1
Multiplicity	10.5	9.5	3.0
Anomalous completeness	94.4	100.0	60.6
Anomalous multiplicity	5.6	6.9	1.8
DelAnom correlation between half-sets	0.069	0.238	0.056
Mid-Slope of Anom Normal Probability	0.989	-	-



Estimate of maximum resolution for significant anomalous
 Estimates of resolution limits: overall
 from half-dataset correlation CC(1/2) > 0.50: limit
 from Mn(I/sd) > 2.00: limit
 from Mn(I/sd) > 2.00: limit
 Estimates of resolution limits in reciprocal lattice dir
 Along h k plane
 from half-dataset correlation CC(1/2) > 0.50: limit
 from Mn(I/sd) > 2.00: limit
 Along l axis
 from half-dataset correlation CC(1/2) > 0.50: limit
 from Mn(I/sd) > 2.00: limit
 Anisotropic deltaB (i.e. range of principal components),
 Average unit cell: 57.58 57.58 149.85 90.00 90.
 Space group: P 41 21 2
 Average mosaicity: 0.09
 Minimum and maximum SD correction factors: Fulls 0.10
 The anomalous signal appears to be weak so anomalous file
 \$\$\$ <!--SUMMARY_END-->
 =====
 === Writing merged data for dataset XDSproject/XDScryst
 Number of reflections written 14969 maximum resolution

ビームラインでの自動処理の今後

これまでは以下の選択性

- XDS (自作スクリプト)
- Denzo_3d/scalepack (自作スクリプト)

今後は既存パイプラインの並行利用

- Fast_dp
– 高速性重視  ビームタイム中の実験
計画へのフィードバック
- XIA2 (XDS and DIALS)
– ある程度の精度、再現性を重視  ビームタイム後のその
後の解析

テストケース

Case 1.

- ソーマチン
- Pilatus3 S6M @ BL17A
- 0.98 Å
- 1度振り、180枚
- $P4_12_12$, $a=b=58$,
 $c=150$

Case 2.

- Protein X
- Pilatus3 S6M @ BL17A
- 1.9 Å
- 0.1度振り、3600枚
- $P2_12_12_1$, $a=59$, $b=76$,
 $c=107$

Case1

Fast_dp

```

-----
Low resolution 28.63 28.63 1.46
High resolution 1.42 6.34 1.42
Rmerge 0.063 0.027 0.856
I/sigma 25.00 66.20 3.30
Completeness 99.9 99.0 98.5
Multiplicity 12.8 11.0 12.7
CC 1/2 1.000 1.000 0.931
Anom. Completeness 99.8 100.0 97.8
Anom. Multiplicity 6.6 7.1 6.4
Anom. Correlation 0.016 0.256 -0.098
Nrefl 615453 7192 43620
Nunique 48081 652 3448
Mid-slope 0.993
dF/F 0.044
dl/sig(dl) 0.804
-----
Merging point group: P 4 2 2
Unit cell: 57.26 57.26 149.40 90.00 90.00 90.00

```

02m 39s

Xeon E5-2670
2.6 GHz, 32 thread

Xia2 (XDS)

```

For AUTOMATIC/DEFAULT/NATIVE Overall Low High
High resolution limit 1.26 3.42 1.26
Low resolution limit 40.52 40.55 1.28
Completeness 99.6 100.0 95.9
Multiplicity 11.6 12.0 5.0
I/sigma 17.6 67.0 1.2
Rmerge(I) 0.073 0.030 1.105
Rmerge(I+/-) 0.070 0.029 0.997
Rmeas(I) 0.076 0.031 1.232
Rmeas(I+/-) 0.076 0.031 1.222
Rpim(I) 0.022 0.009 0.533
Rpim(I+/-) 0.030 0.012 0.693
CC half 0.999 0.999 0.508
Wilson B factor 9.884
Anomalous completeness 99.4 100.0 93.2
Anomalous multiplicity 6.1 7.0 2.7
Anomalous correlation 0.009 0.047 -0.023
Anomalous slope 1.009
Total observations 786585 45414 16205
Total unique 67994 3769 3227
Assuming spacegroup: P 41 21 2
Other likely alternatives are:
P 43 21 2
Unit cell:
57.310 57.310 149.500
90.000 90.000 90.000

```

09m 56s

Case 2.

Fast_dp

```

-----
Low resolution 29.31 29.31 2.45
High resolution 2.38 10.66 2.38
  Rmerge 0.090 0.043 0.639
  I/sigma 19.60 42.20 3.10
Completeness 98.6 95.2 82.1
Multiplicity 11.9 10.6 10.1
  CC 1/2 0.999 0.999 0.883
Anom. Completeness 98.3 97.9 77.5
Anom. Multiplicity 6.4 6.9 5.5
Anom. Correlation -0.090 0.183 -0.148
  Nrefl 231807 2685 11934
  Nunique 19417 254 1181
  Mid-slope 0.826
  dF/F 0.044
  dl/sig(dl) 0.717
-----
Merging point group: P 2 2 2
Unit cell: 58.63 75.59 107.28 90.00 90.00 90.00
  
```

04m 35s

Xia2 (XDS)

```

For AUTOMATIC/DEFAULT/NATIVE Overall Low High
High resolution limit 2.17 5.88 2.17
Low resolution limit 37.79 37.80 2.21
Completeness 96.2 100.0 68.7
Multiplicity 10.9 11.5 4.8
I/sigma 14.8 32.6 1.5
Rmerge(I) 0.100 0.051 0.978
Rmerge(I+/-) 0.097 0.050 0.861
Rmeas(I) 0.105 0.054 1.100
Rmeas(I+/-) 0.106 0.054 1.072
Rpim(I) 0.031 0.016 0.487
Rpim(I+/-) 0.042 0.021 0.623
CC half 0.998 0.999 0.350
Wilson B factor 31.192
Anomalous completeness 95.3 99.9 3.3
Anomalous multiplicity 5.8 6.7 2.6
Anomalous correlation -0.149 0.017 0.013
Anomalous slope 0.770
Total observations 271259 16435 4170
Total unique 24923 1433 860
Assuming spacegroup: P 21 21 21
Unit cell:
58.650 75.590 107.310
90.000 90.000 90.000
  
```

15m 07s

回折データ処理のインフラ

	2017年3月まで	2017年4月以降
ノード数	8	14
総CPUコア数	256	544
総メモリ容量 (GB)	256	1,024
ネットワーク	1 Gbps	10 Gbps

計算ノード14台

- 2台 Findspots
- 6台 Fast_dp (5ビームライン)
- 6台 Xia2 (5ビームライン)

} 即時実行
 順次実行

まとめ

- PFではデータ測定終了後にXDSによる自動処理が行われている。
- 自動処理結果は_process/xds_<N>に出力され、さらにPReMo(実験データベース)に記録され、Webから閲覧可能である。
- 2017年秋からは、Fast_dp, Xia2などの既存のパイプラインが複数走るようになる。