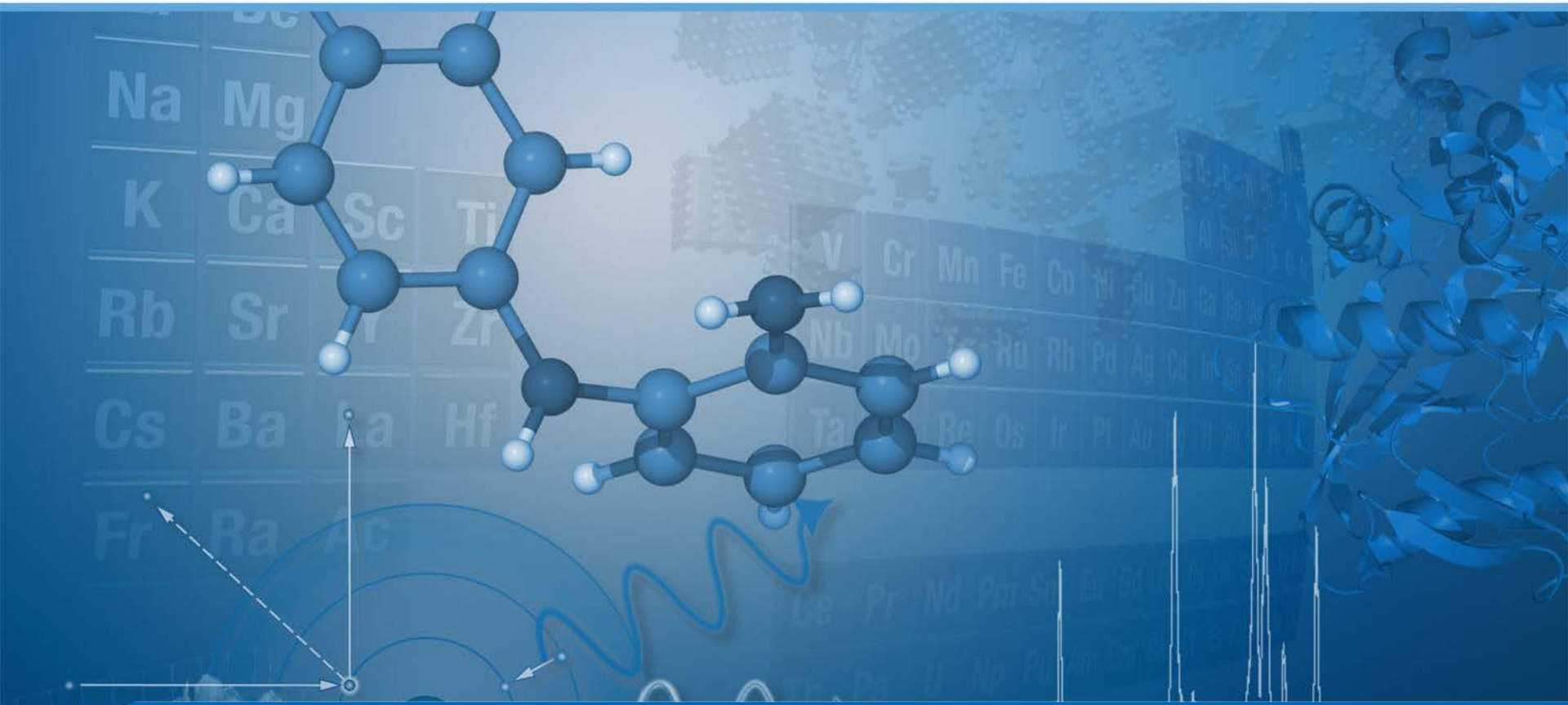
PROTEUM2 and Foreign Frame Formats



PROTEUM2 Software for Structural Biology

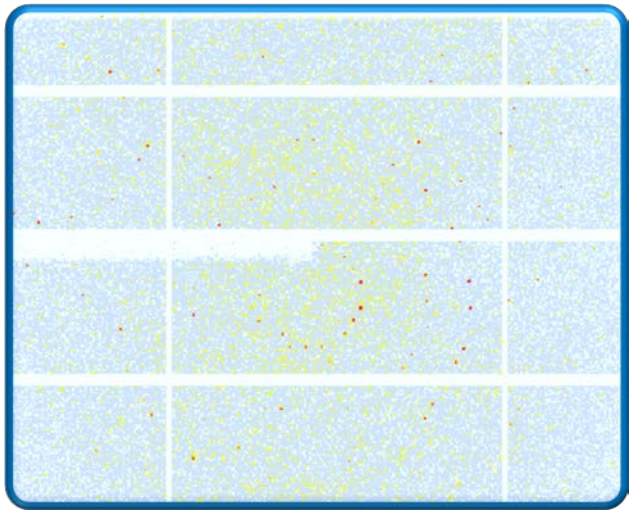
PROTEUM2

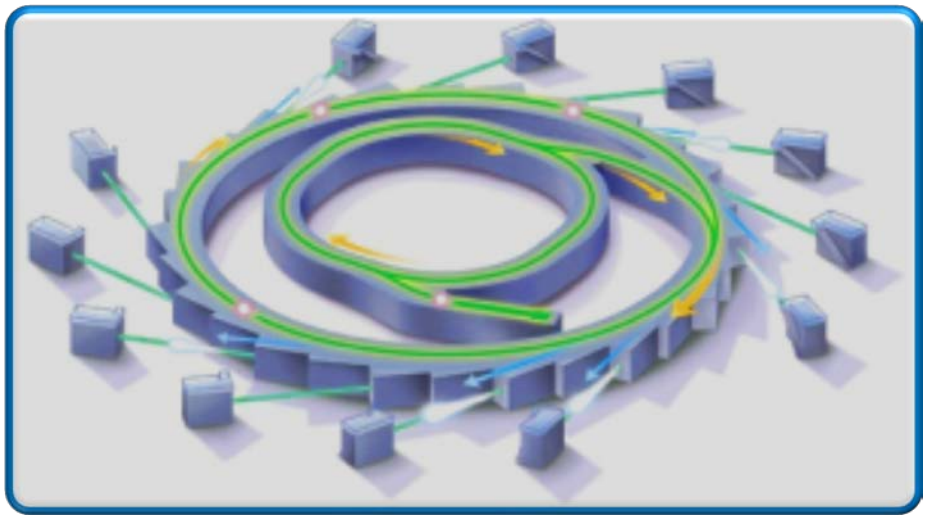
- Fully integrated software suite for structural biology
  - visualization, indexing, integration, scaling, spacegroup determination
  - Data consistency and integrity
- Handles multi-axis goniometer geometries correctly
  - Kappa geometry
  - Non perpendicular phi scans
- Optimized 3D profile data integration for finely sliced data
  - twinning
- Multi-run strategy,
  - Multi-run integration and scaling

PROTEUM2

- High intensity beamlines
- Beam lines with multi-axis goniometers
- Ultra low-noise and ultra fast detectors promote fine slicing data

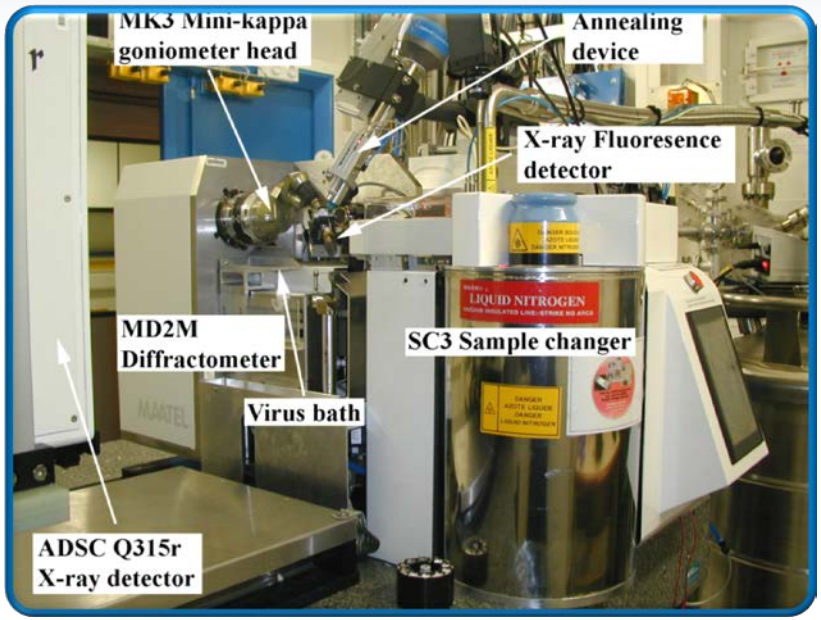
- **Highly redundant fine slicing data from Kappa goniometers becomes more prevalent**





Multi-Axis Goniometers

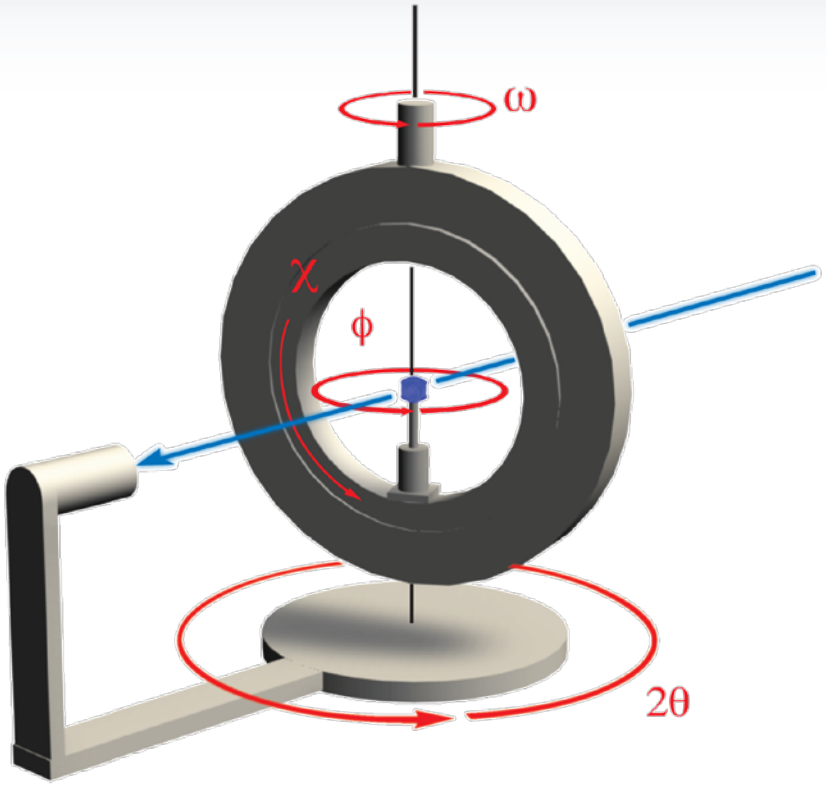
- Multi-Axis goniometers become more popular at beamlines
- Data requires proper processing
  - MK3 mini-kappa at ID14-4 at the ESRF
  - PRIGo: a novel multi-axis goniometer at beamline X06DA at the Swiss Light Source
  - MK3 mini-kappa BL14.1 at Bessy



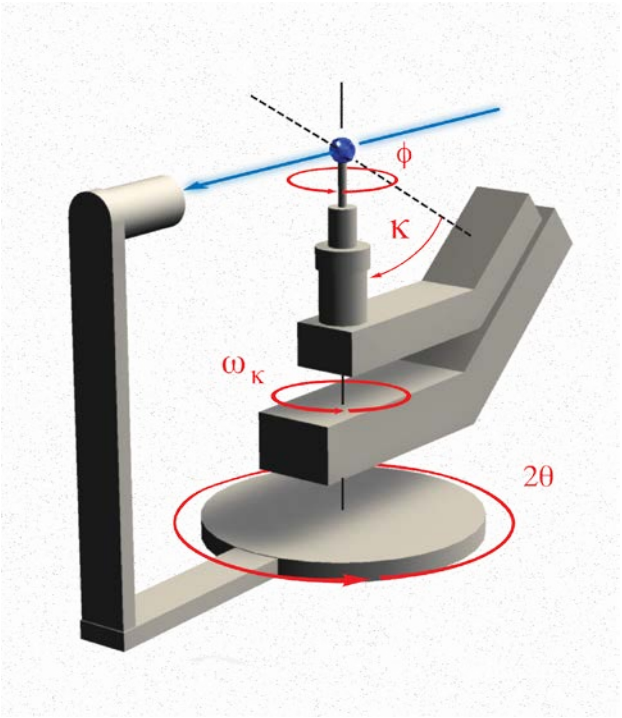
https://www.embl.fr/services/synchrotron_access/id14-4/

- View of the highly automated ID14-4 experimental hutch at the ESRF Grenoble

Goniometer Geometry



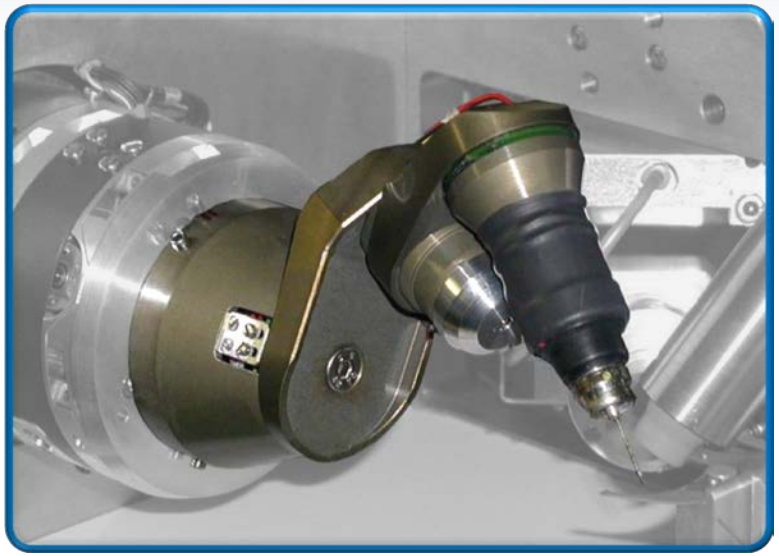
Eulerian Geometry



Kappa Geometry

Multi-Axis Goniometers

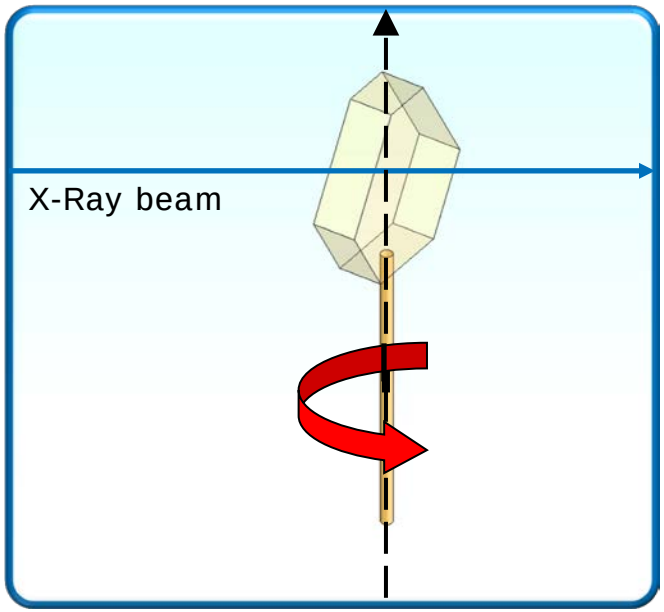
- Bruker KAPPA goniometer



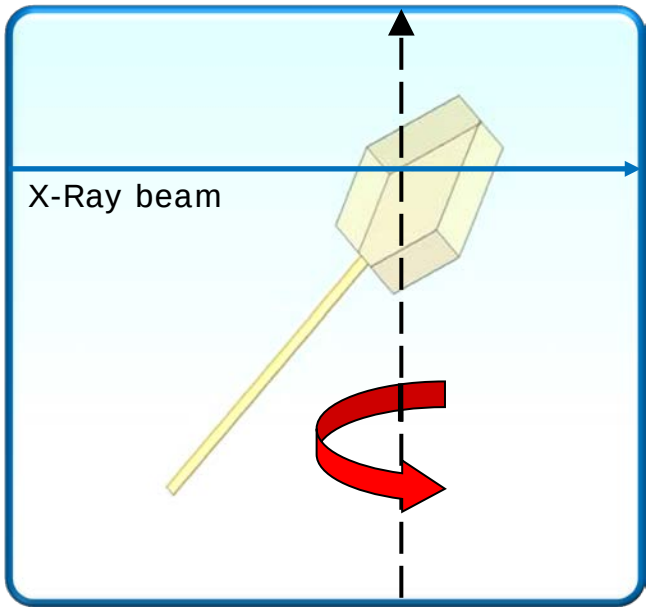
- EMBLEM, MAATEL Bruker corporation on the MK3 Mini-Kappa goniometer

Multi-Axis Goniometers

- A Kappa goniometer allows high multiplicity data collection
- Reflections are collected with multiple diffraction geometries
- Systematic errors are minimized
- Data provide a better model for scaling
- High multiplicity data is needed for S-SAD phasing



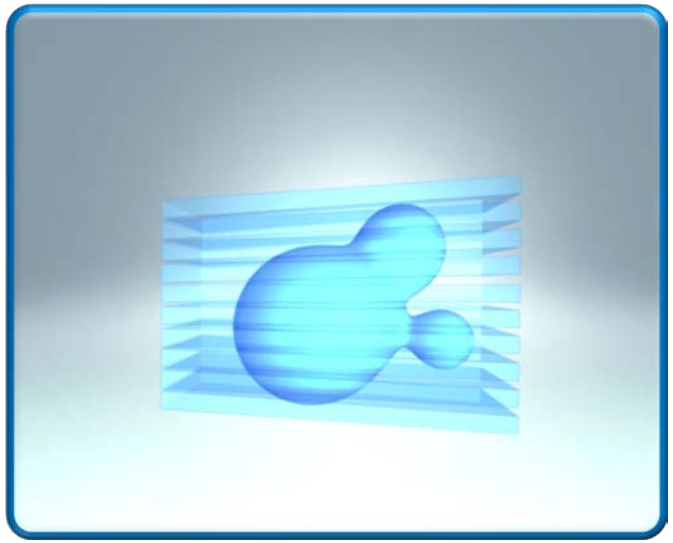
Static multiplicity

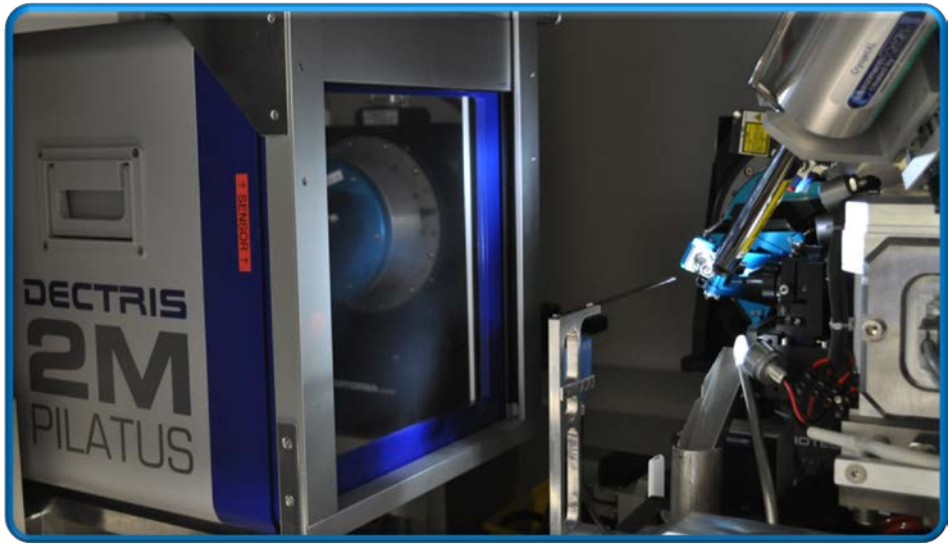


True multiplicity

Fast and Finely Sliced Data at Beamlines

- Ultra low-noise and ultra fast detectors promote fine data slicing





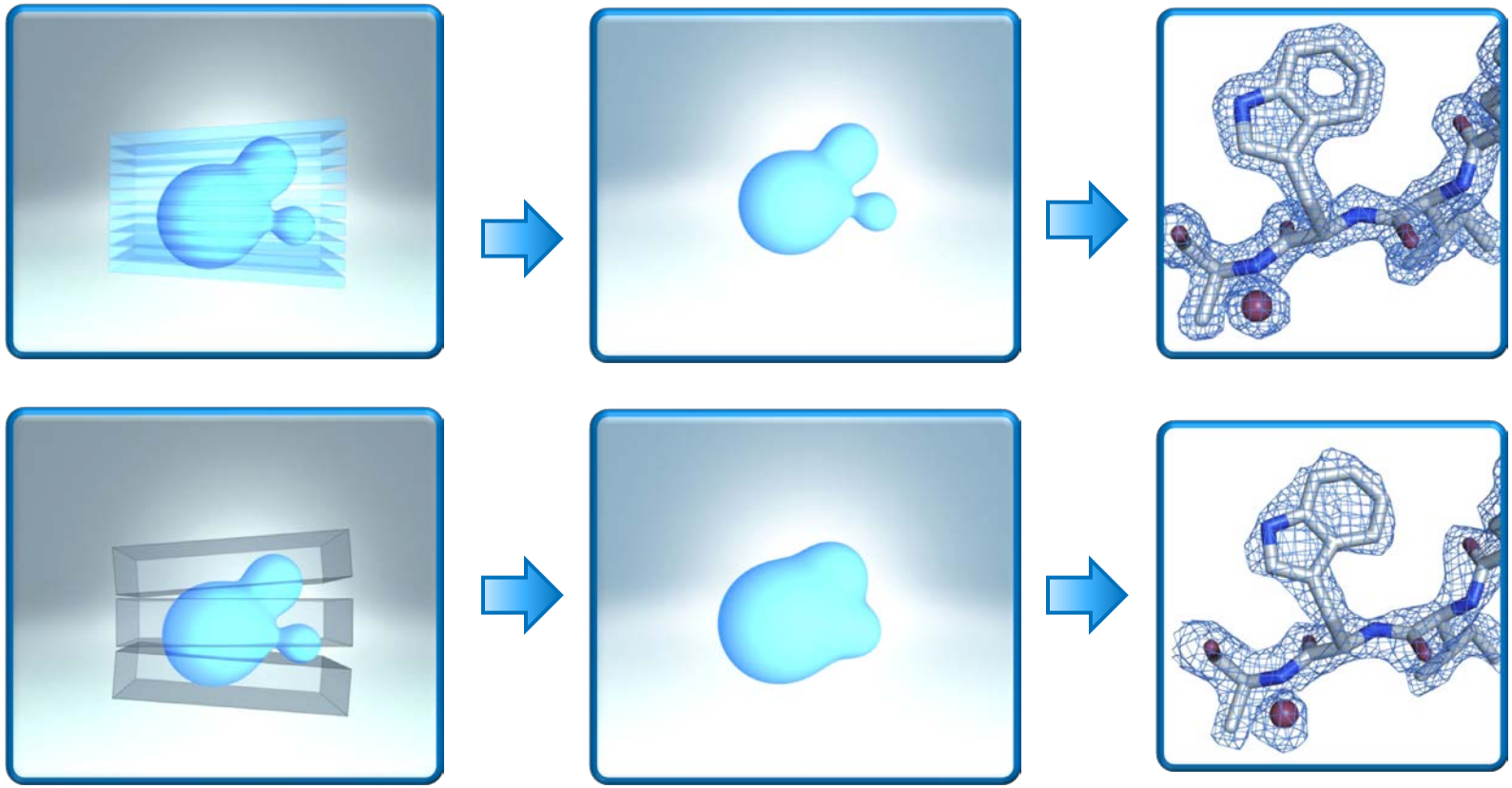
Rotation 0.025°
Exposure 0.02 sec

http://www.psi.ch/sls/pxiii/pxiii

X06DA endstation with the PILATUS 2M detector and the PRIGo goniometer

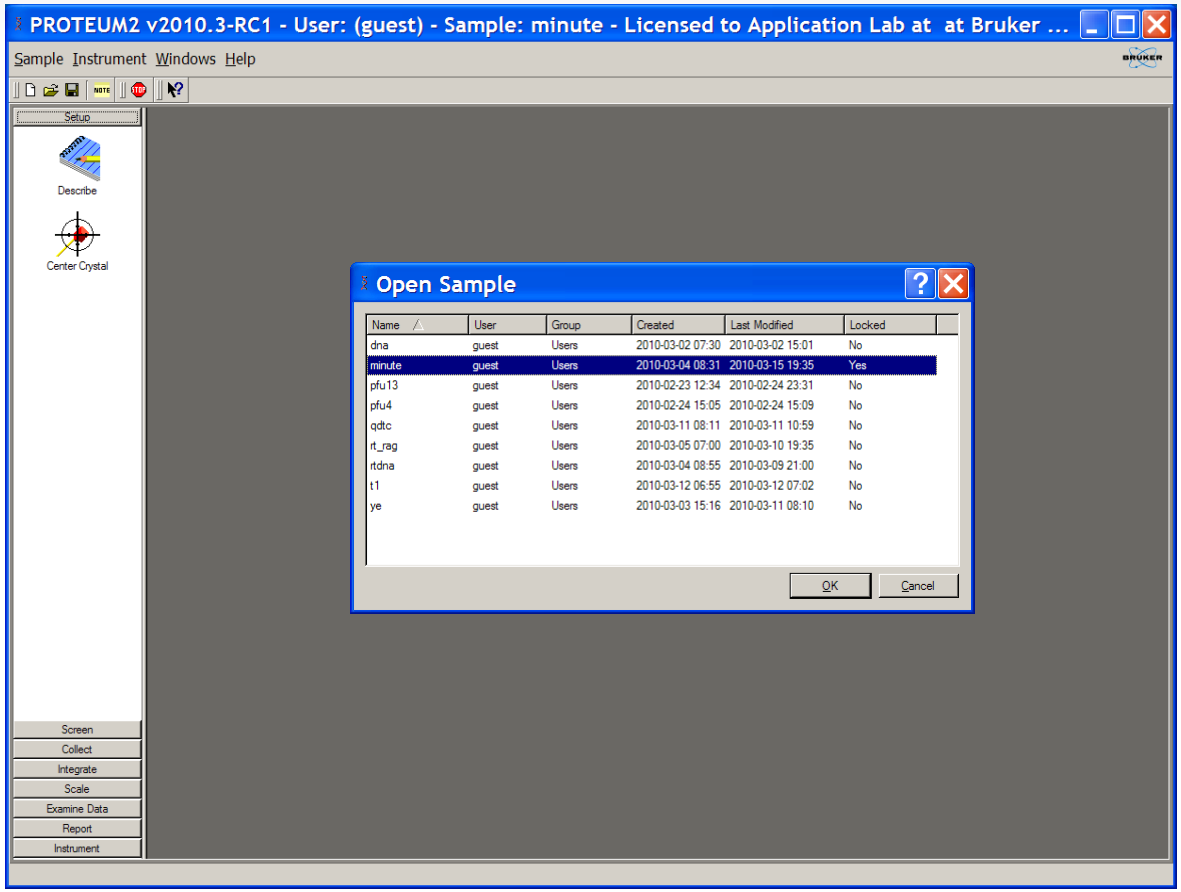
SAINT 3D Profile Data Integration

- Finely sliced data and 3D-profile integration improves data quality and map detail



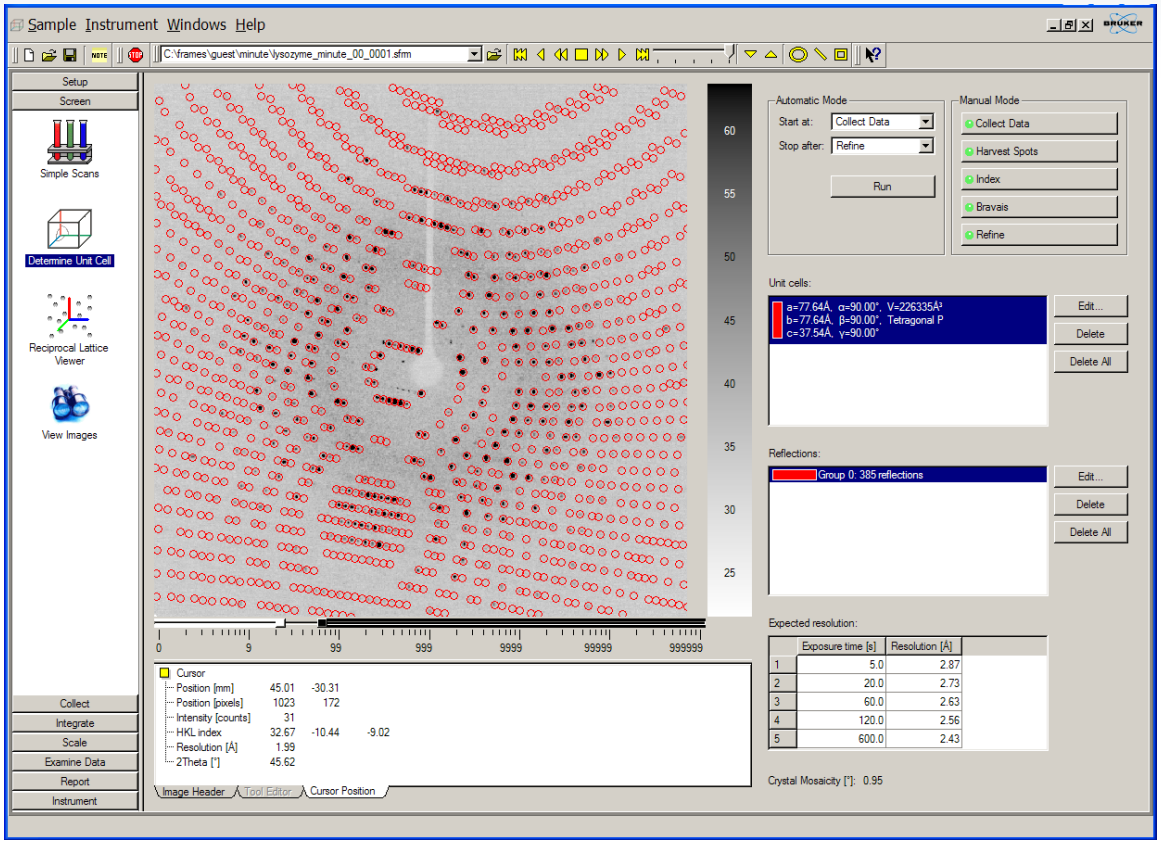
PROTEUM2 Software

- All tools necessary for initial screening to data reduction
- Easy to navigate interface
- Modular design
- Experimental results are stored in a SQL database
- Server/Client platform
- SHELXTL software for structure solution and phase determination



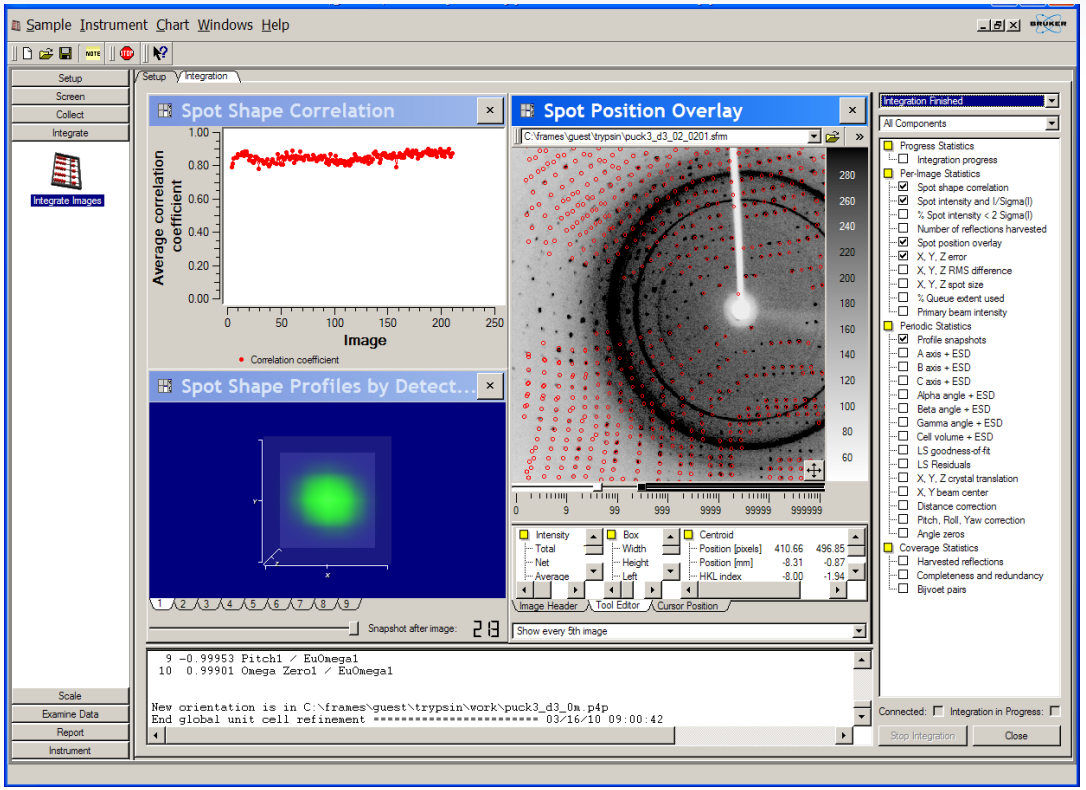
PROTEUM2 Software
Indexing

- Auto and manual modes
- Ability to store multiple reflection arrays and orientation matrices
- Uses different indexing algorithms
  - Difference vector
  - FFT
  - Cell_now



PROTEUM2 Software
Data Integration

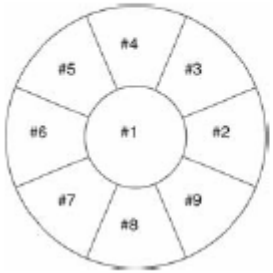
- True 3D profile fitting
- Extended Graphical feedback
  - 3D profile display
  - Spot overlays
- SAINT integration engine
- One button integration using intuitive defaults
- Handles twinned data



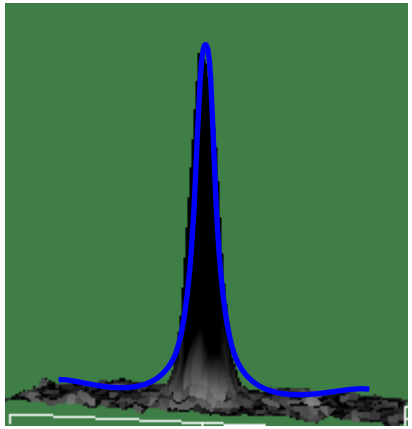
PROTEUM2 Software
Data Integration

Detector region

During the integration, the model profile shape is determined separately for nine regions of the detector (see picture). Blending the model profiles results in less variation across the detector area, and may provide better statistics for regions where the reflections are very weak.



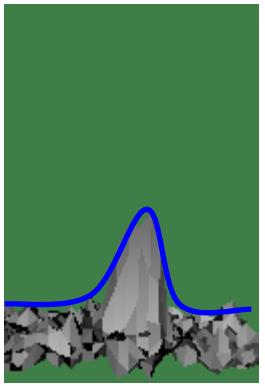
LS profile fitting



Strong reflections:
I/σI > 6

Calculated profile

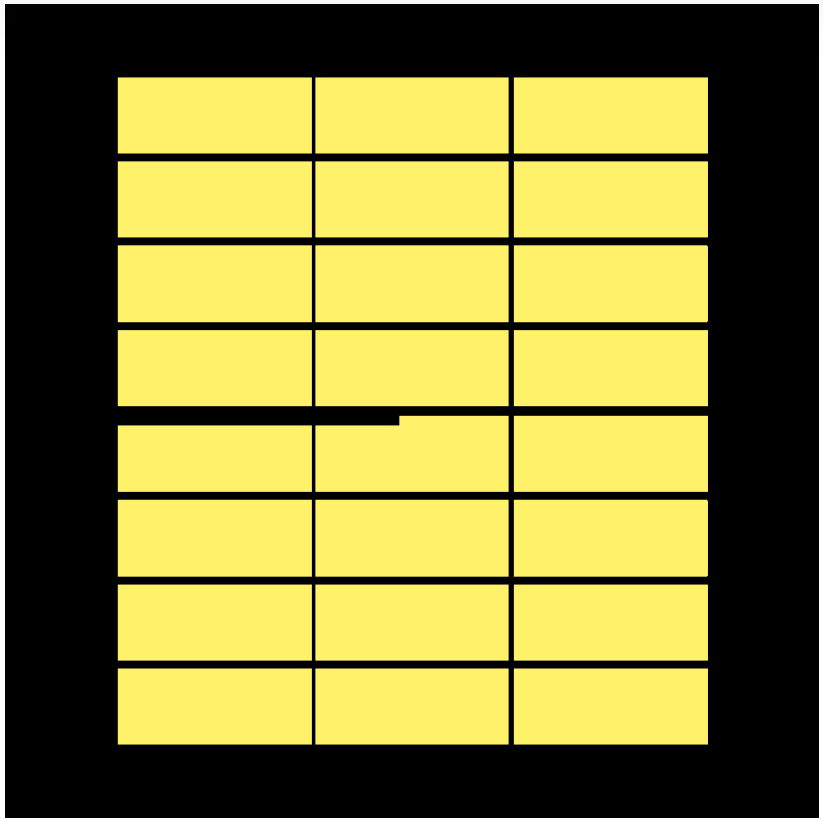
is applied to



Weak reflections:
I/σI < 4

PROTEUM2 Software
SAINT Active Mask

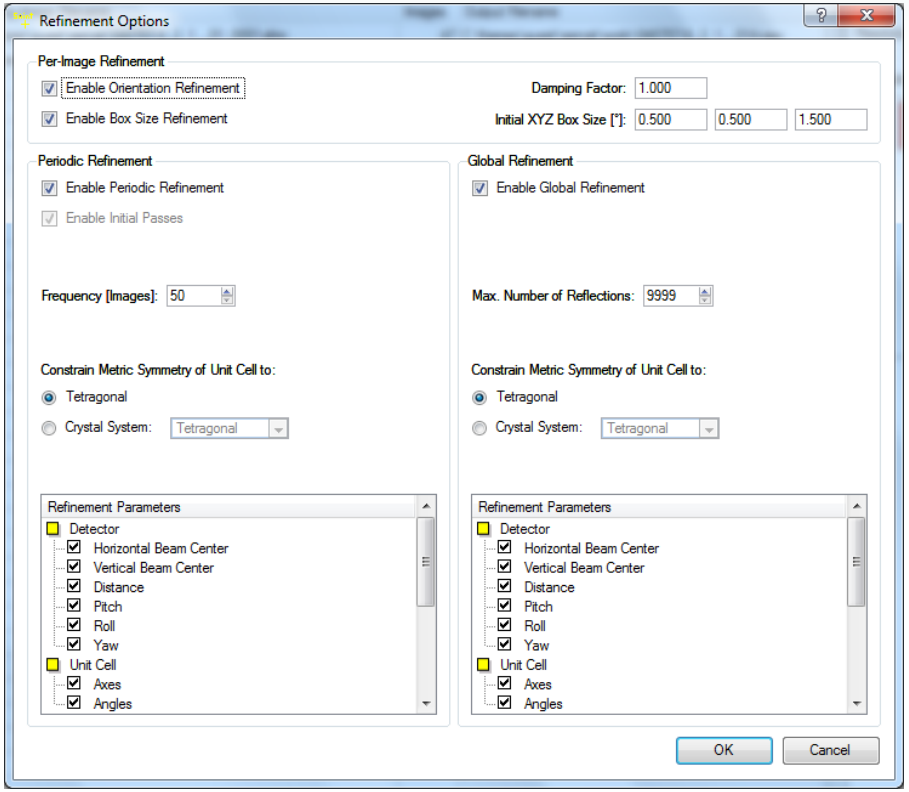
- SAINT uses an active mask to discard reflections that “touch” inactive areas of the detector
- Calculates inactive area based on median pixel values or pre-defined mask
- In this representation yellow areas are active
- The Bruker frame format uses a $2^n \times 2^n$ frame size to increase data processing efficiency (in this case frames are padded to 2048 × 2048 pixels)



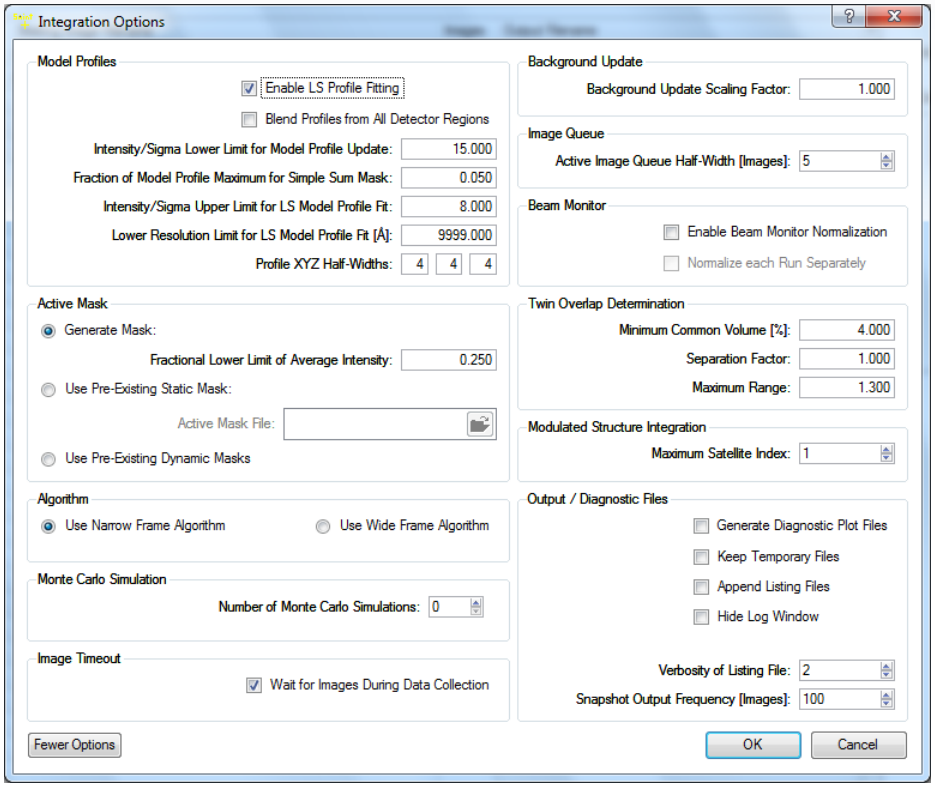
Mask for the 24 module PILATUS 2M detector at SLS X06DA endstation

PROTEUM2 Software
Data Integration

Refinement options for SAINT



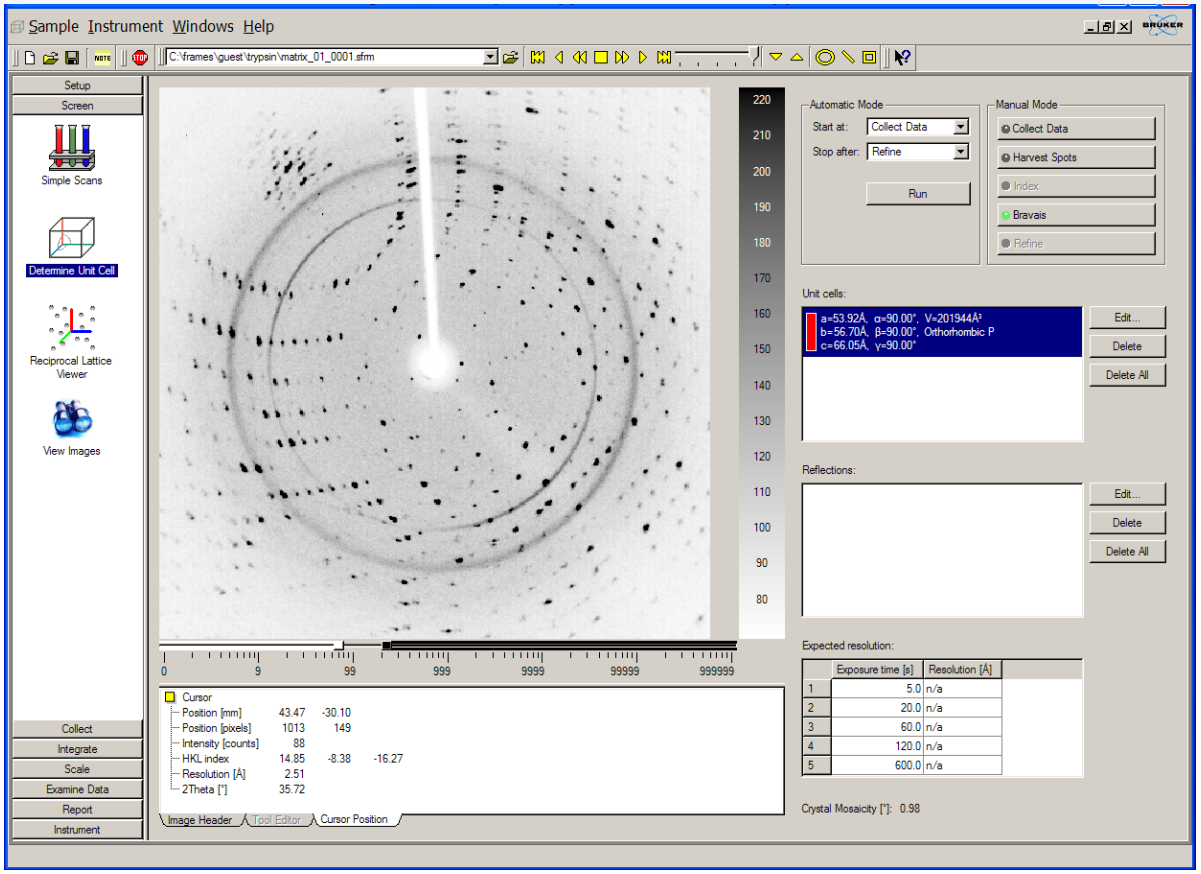
Integration options for SAINT



- Box size refinement
- Profile determination parameters
- Integration algorithm

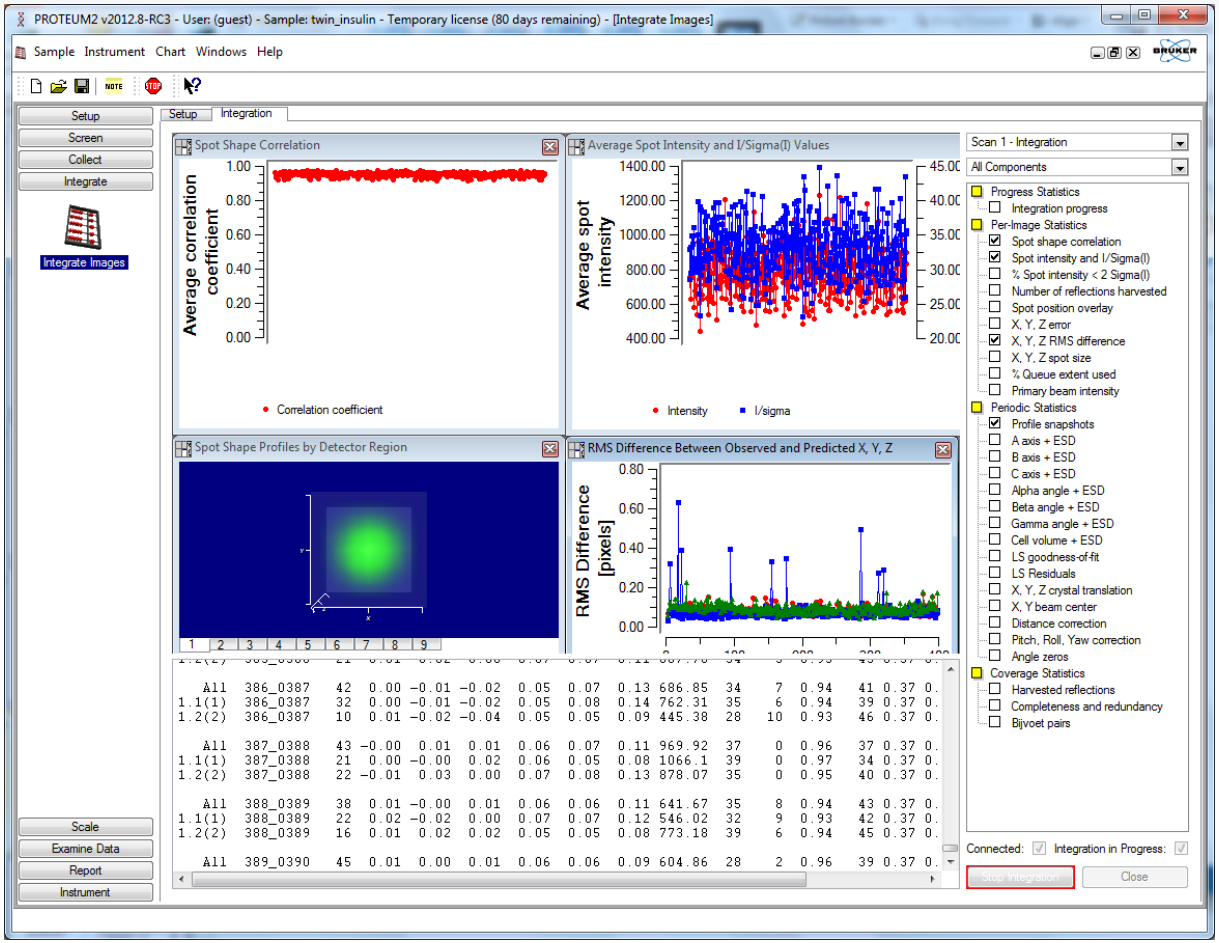
PROTEUM2 Software
RLATT

- Displays reflections in reciprocal space
- Provides an easy way of modifying the harvested reflection array
- Remove secondary scatter for indexing
- Manual separate twined lattices
- Good way to index crystals which are twinned, split or have satellites attached



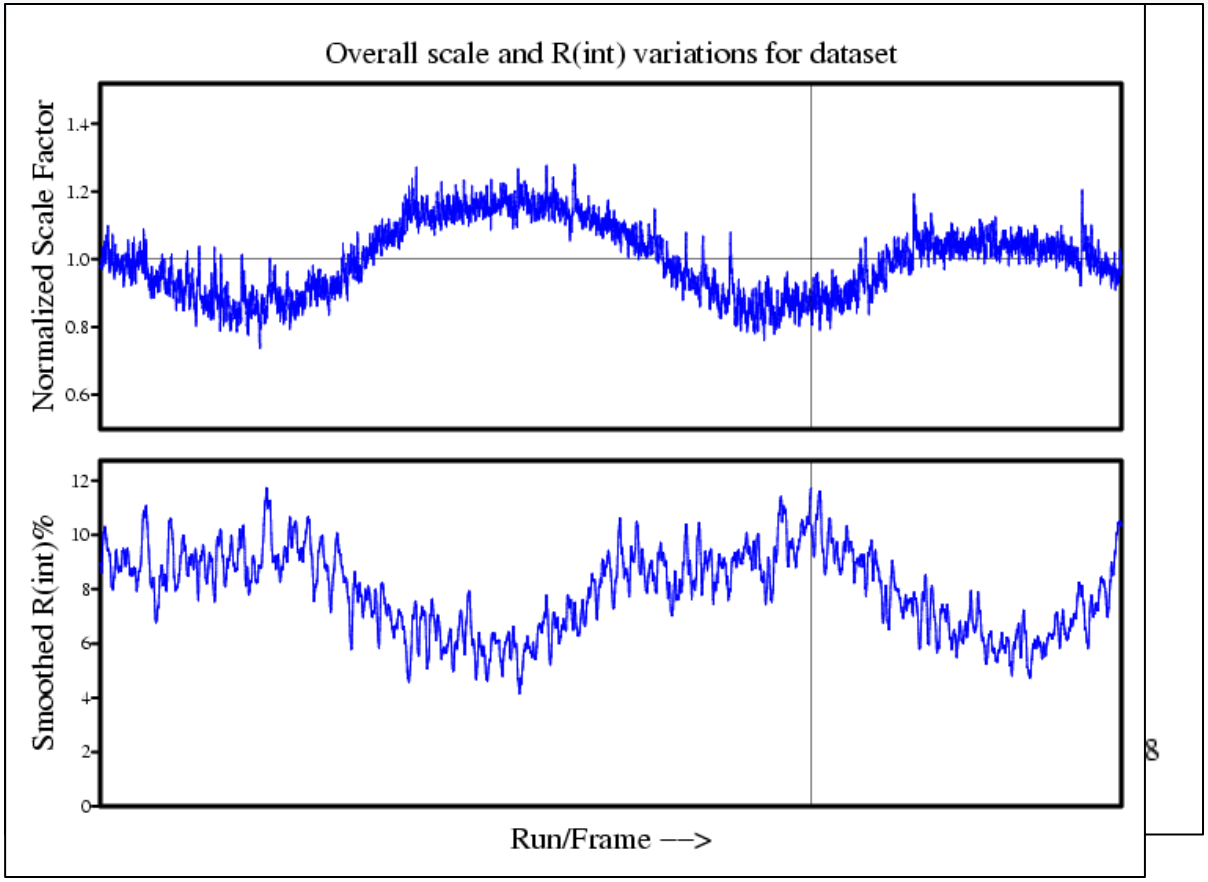
PROTEUM2 Software
Twinned Data Integration

- Twinned insulin crystal
- Manual separate domains
- Index all orientation matrices
- Pass information to SAINT
- TWINABS
  - HKL file from one domain
  - HKl file with all domains
- Cell_now
  - Auto-detwinning



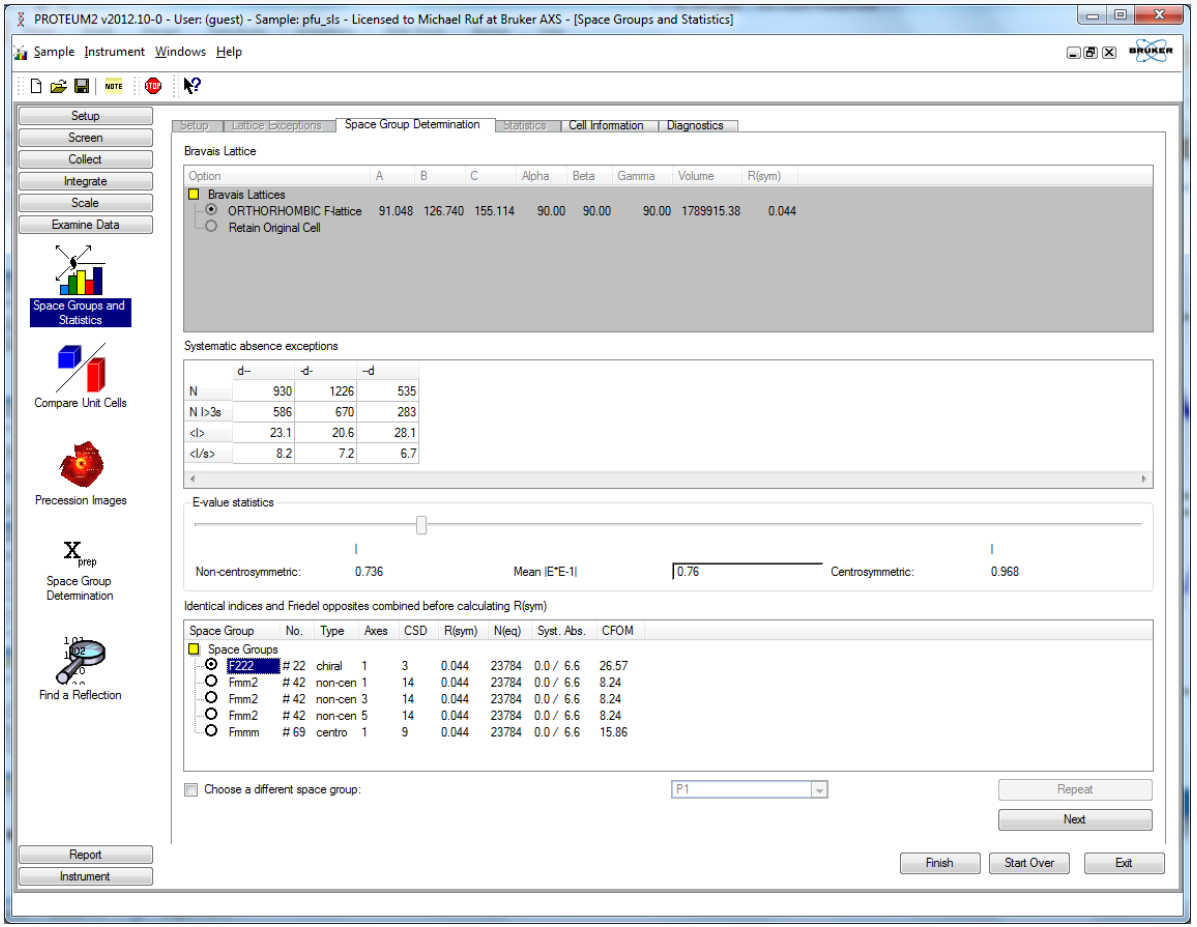
PROTEUM Software
Scaling

- Multiscan method
- SADABS engine
- Scaling and absorption correction
- Error model determination and adjustment of standard uncertainties
- Fast accurate and reliable
- Graphical feedback



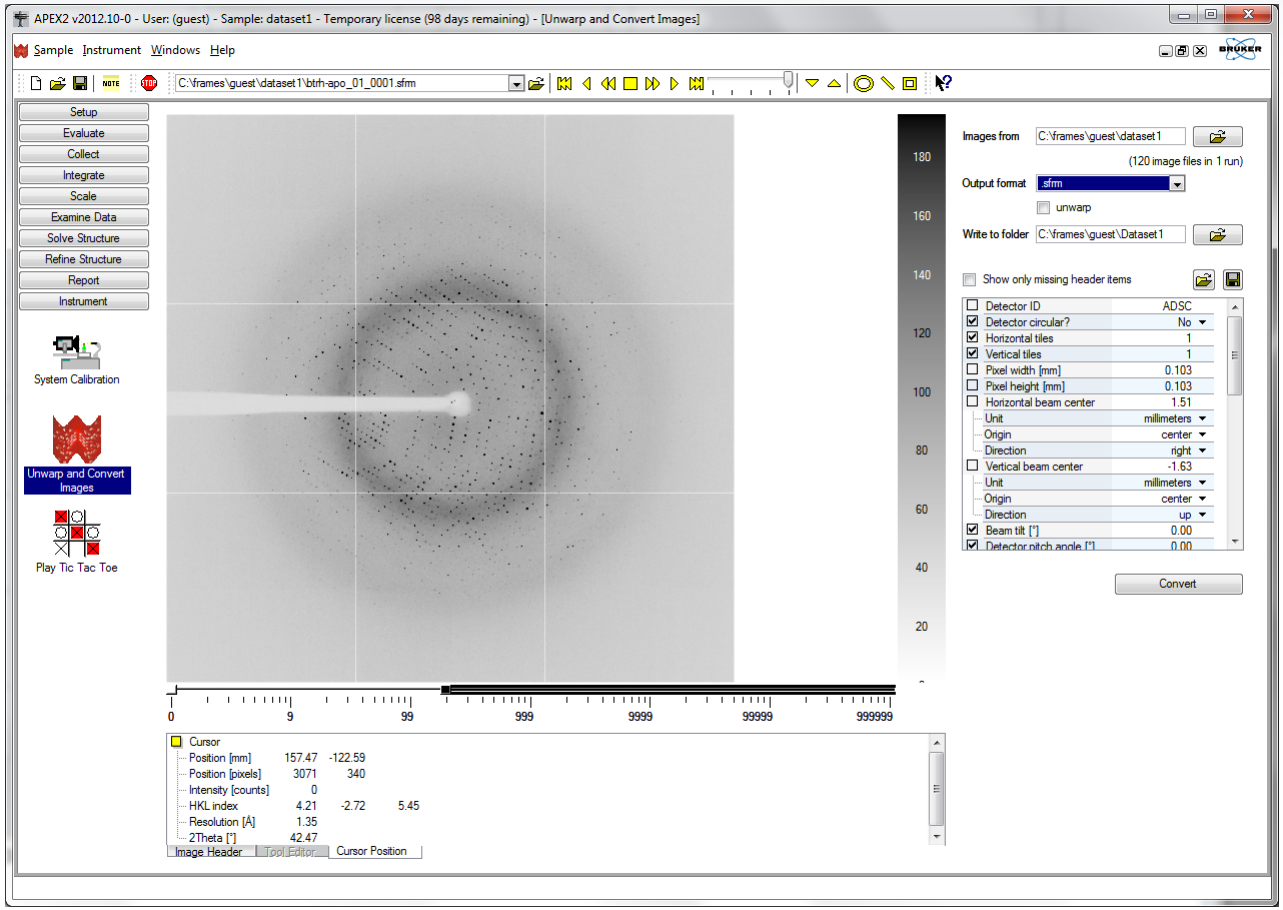
PROTEUM Software
Space Group and Statistics (XPREP)

- Automated spacegroup determination
- Intensity statistics
- Uses XPREP engines
- More functionality from XPREP
  - Setup of MAD, SAD experiments
  - Tests for twinning
  - Patterson maps
  - Self-rotation maps



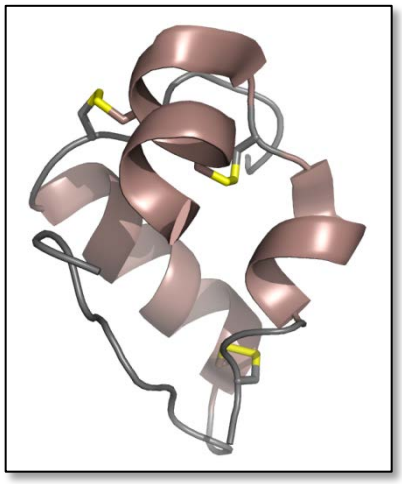
PROTEUM Software
Unwarp and Convert Images

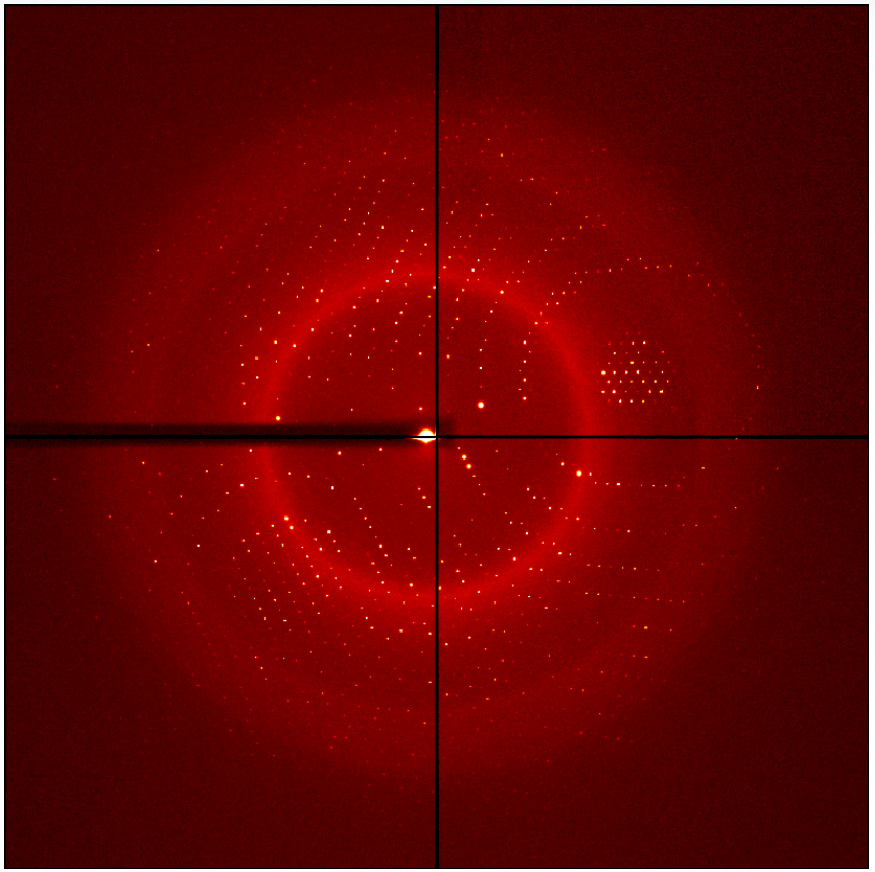
- Converts between different frame formats
- Detectors supported
  - MAR, Rayonix
  - ADSC
  - DECTRIS
- Create a hardware profile for each beamline
- Allows the frame header information to be updated
  - Beam center
  - Detector characteristics



Data Processing
ADSC Quantum 4

- Cubic Insulin
  - Rotation angle 1°
  - *a*=78.23 Å
  - 0.5 s/°
  - 100° of data collected in a single phi scan

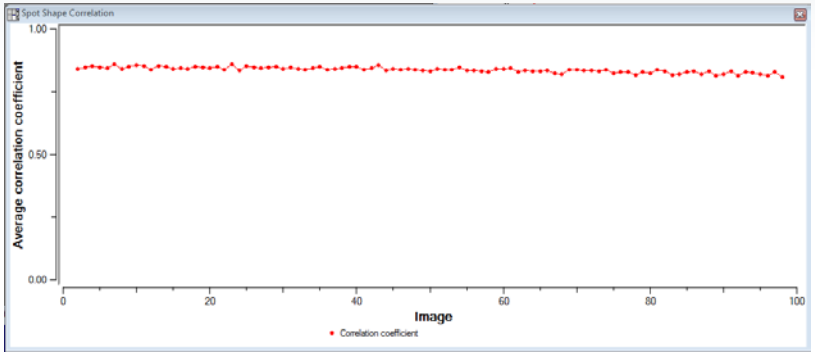


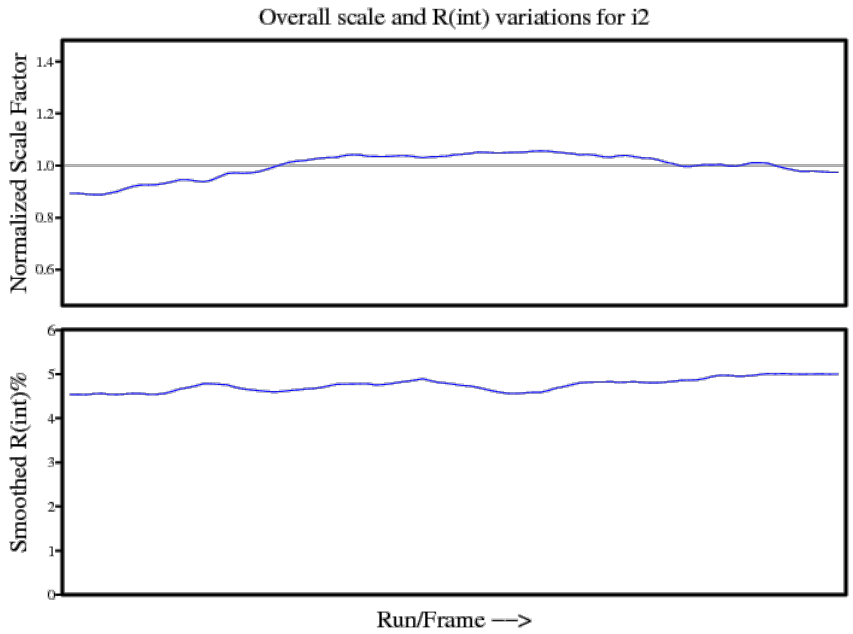


ID14-2 Macromolecular Crystallography Beamline with ADSC Quantum 4

Data Processing
ADSC Quantum 4

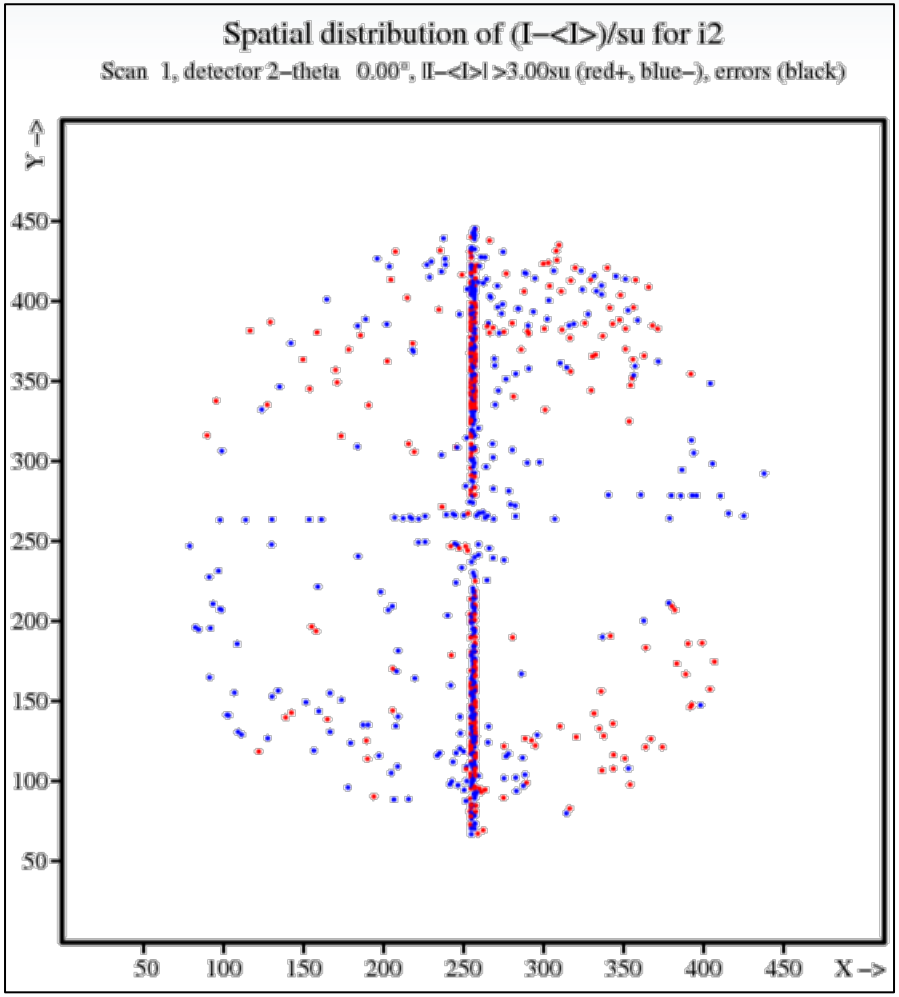
- High > 90% spot shape correlation
- Good model profiles
- Smooth and small overall scale variations
- Small and consistent internal consistency of data (small R_{merge})





Data Processing
Proper Active Mask

- Data where processed without an active mask
- Spatial distribution of the deviation from the mean Intensity shows the gaps between the individual modules
- About 8% of the reflections were rejected during the error model determination and outlier rejection step of scaling



Data Processing
ADSC Quantum 4

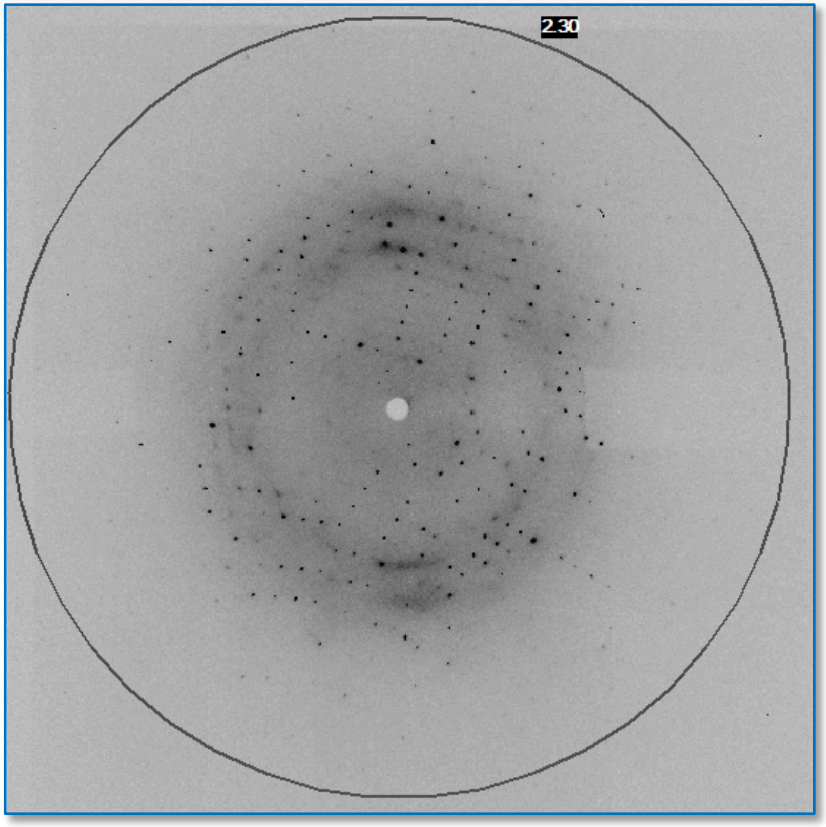
Intensity statistics

Resolution	%Comp	Redundancy	Mean I/σI	R(int)
Inf - 5.08	99.7	10.44	67.38	0.0307
5.08 - 3.97	100	11.39	73.33	0.0282
3.97 - 3.45	100	11.59	66.78	0.0284
3.45 - 3.13	100	11.64	62.32	0.0331
3.13 - 2.89	100	11.78	54.52	0.0378
2.89 - 2.71	100	11.72	46.9	0.0441
2.71 - 2.58	100	11.83	44.34	0.0483
2.58 - 2.46	100	11.74	39.19	0.0535
2.46 - 2.36	100	11.84	34.79	0.0626
2.36 - 2.27	100	11.82	30.08	0.0714
2.27 - 2.20	100	11.8	29.13	0.0768
2.20 - 2.13	100	11.77	27.13	0.0829
2.13 - 2.07	100	11.86	22.73	0.0988
2.07 - 2.02	100	11.67	20.2	0.1120
2.02 - 1.97	100	11.81	16.4	0.1370
1.97 - 1.93	100	11.8	14.52	0.1559
1.93 - 1.89	100	11.72	13.13	0.1638
1.89 - 1.85	100	11.75	10.81	0.1968
1.85 - 1.81	100	11.67	8.92	0.2246
Inf - 1.81	100	11.67	35.5	0.0468

Data Processing
MAR 225

- Beamline SERCAT 22-ID
- Structural genomics: AF1382
- Wavelength 1.9 Å
- Space group $P4_2$
- Cell: $a=b=53.61$, $c=41.18$ Å

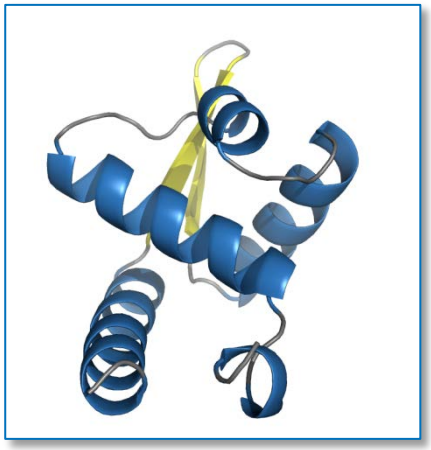
Resolution	%Comp	Redundancy	Mean I/σI	R(int)
Inf - 4.94	99.7	13.72	62.73	0.0314
4.94 - 3.92	100	14.32	61.76	0.0335
3.92 - 3.42	100	14.18	52.57	0.0461
3.42 - 3.11	100	14.13	40.61	0.058
3.11 - 2.89	100	14.04	29.09	0.0816
2.89 - 2.72	100	13.91	21.02	0.1203
2.72 - 2.58	100	13.77	15.21	0.1635
2.58 - 2.47	100	13.65	11.59	0.2067
2.47 - 2.38	100	13.62	8.45	0.2659
2.38 - 2.30	100	13.39	5.38	0.3655
Inf - 2.30	100	13.89	31.71	0.0542

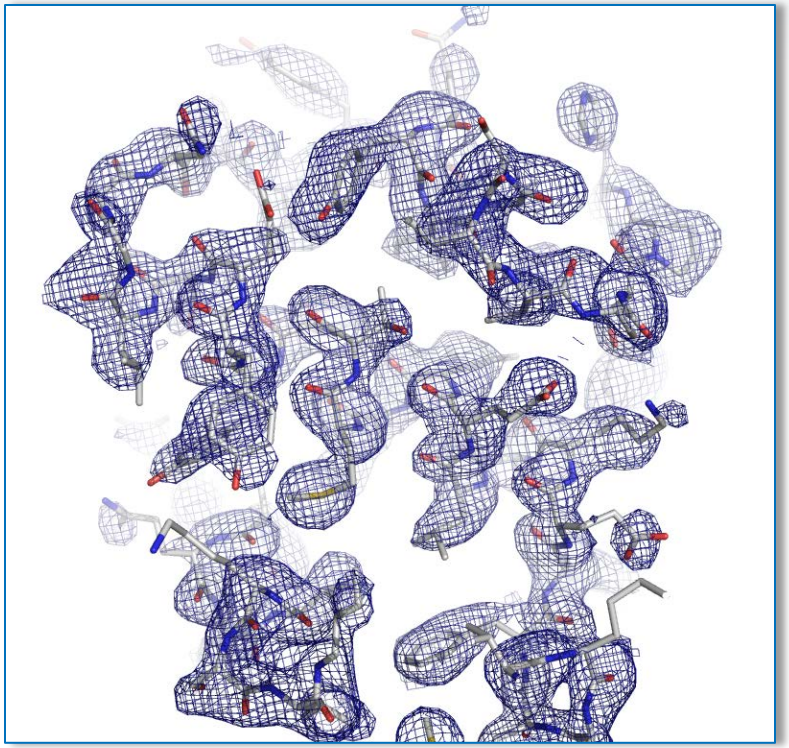


B.C. Wang, John Rose
University of Georgia

PROTEUM2
Mar Mosaic

- Solved using Sulfur SAD
- 95 aa, 4 Met, 1Cys
- 4 sulfur positions were found, all but n-terminal Met
- Sub-structure determined using SHELXD, initial phases produced from SHELXE



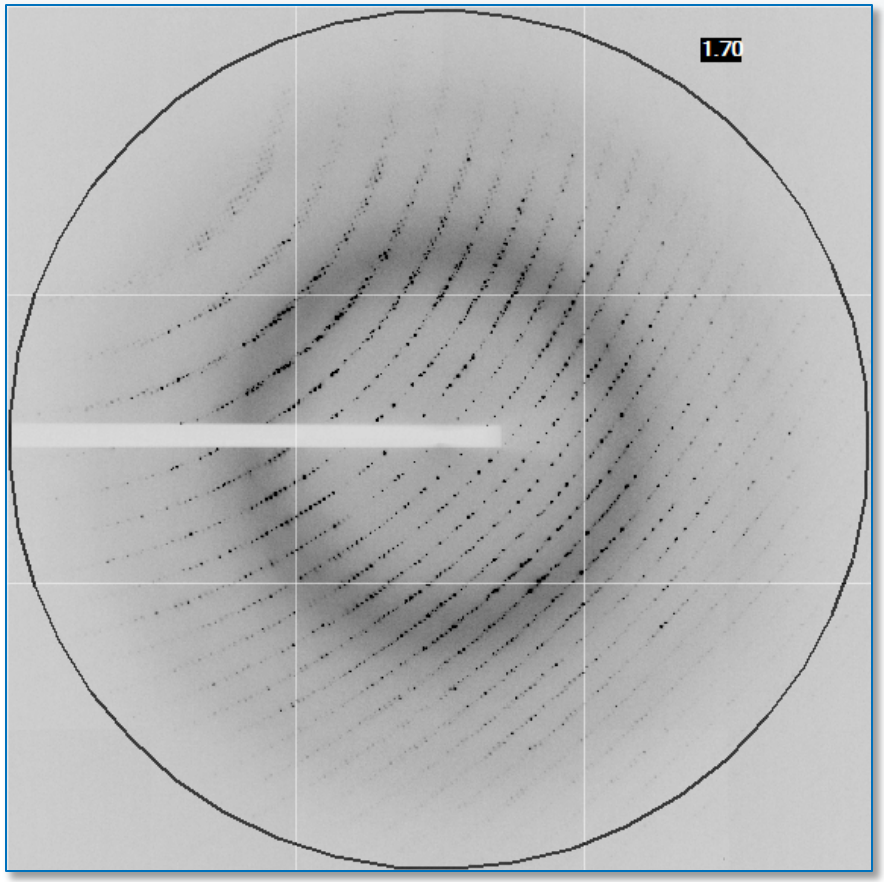


Experimental map contoured at 1 σ

Data Processing
ADSC Quantum 315

- Beamline: Diamond IO2
- Space group: $P4_12_12$
  $a=b=121.73$, $c=33.7$ Å
- Rotation angle 1°
- Exposure time 0.5 s/°

Resolution	%Comp	Redundancy	Mean I/σI	R(int)
Inf - 3.64	99.6	12.22	36.51	0.0451
3.64 - 2.89	100	12.88	33.22	0.0711
2.89 - 2.52	100	13.48	29.35	0.0782
2.52 - 2.29	100	13.71	24.99	0.0973
2.29 - 2.12	100	13.68	20.56	0.1218
2.12 - 2.00	100	13.63	16.08	0.1587
2.00 - 1.90	100	13.51	11.04	0.2246
1.90 - 1.82	100	13.35	8.42	0.2812
1.82 - 1.75	100	13.26	5.75	0.3825
1.75 - 1.70	100	13.16	4.58	0.4497
Inf - 1.70	100	13.28	19.85	0.0788



Dimitri Chirgadze
University of Cambridge

Data Processing
ADSC Quantum 315

- Comparison between SAINT and MOSFLM

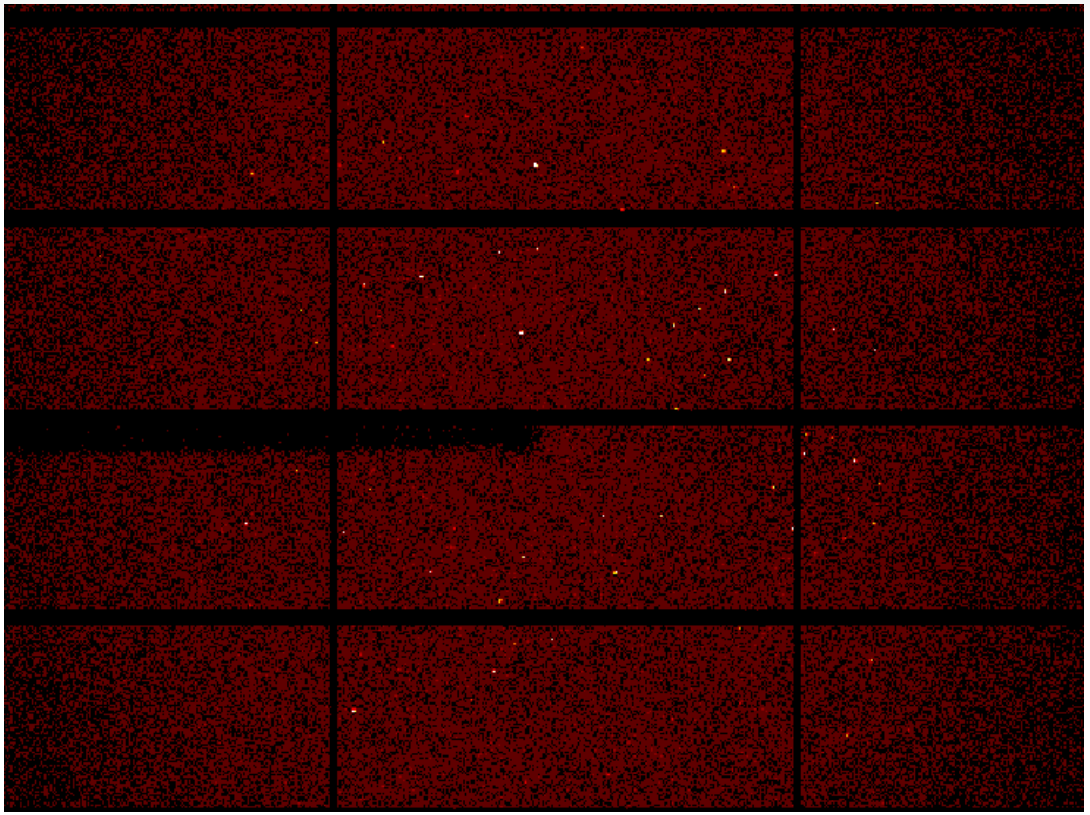
Software	Resolution	Unique Reflections	I/σI	Rmerge (%)	% Comp	Multiplicity
SAINT	43 - 1.69 (1.78 - 1.70)	28,693	19.9	7.9 (42.7)	100 (100)	13.3 (13.2)
MOSFLM	43 - 1.69 (1.78 - 1.70)	29,025	17.3	9.0 (61.0)	100 (100)	13.3 (13.4)

PROTEUM with PILATUS 2M Detector and the PRIGo Goniometer Data

- PILATUS provides mini cbf that needs to be converted to cbf
  - for n in *.cbf; do convert_minicbf –i $n –o converted/$n –p template; done
- PROTEUM reads cbf and can index cbf
- PILATUS 2M data “difficult” to visualize
  - Sharp small spots are “difficult” to see in 2D
- Need to convert data to Bruker frame format for integration
- Data statistics look respectable
- Not sure whether 0.02s/0.025 deg are appropriate
  - Wider scans and longer exposure times for better counting statistics
- Not sure whether goniometer and detector are properly synchronized
  - Large apparent goniometer error during integration

PROTEUM with PILATUS 2M Detector and the PRIGo Goniometer Data

- Ultra-fast data collection in a few seconds or minutes
- 14000 images in 4.5 min

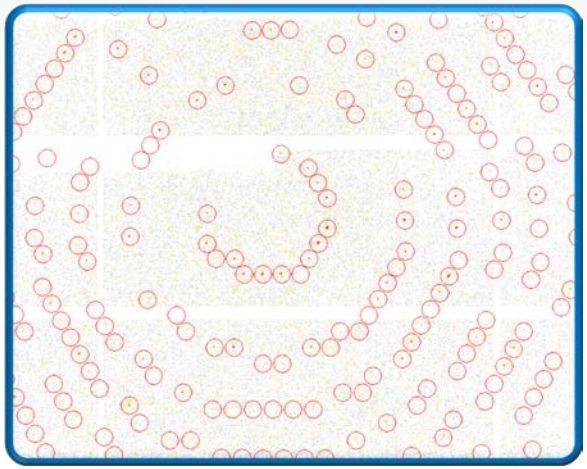


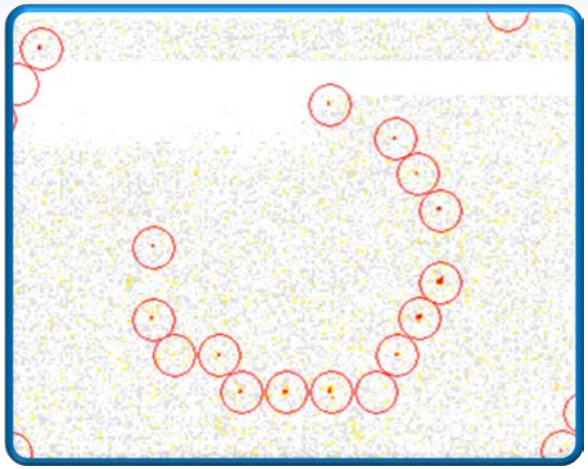
Data Statistics
X06DA Endstation with the PILATUS 2M Detector and the PRIGo Goniometer

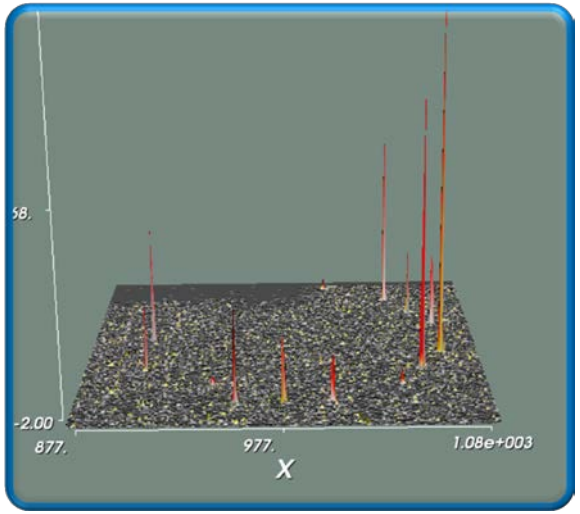
- Exposure 0.02 sec
- Rotation 0.025°
- 200 times shorter exposure time (per deg) and 10 times narrower frames compared to lab data

	15Kev	Chi40	Omega1	Omega2	Ori-ba
Resolution (Å)	2.8	2.8	2.8	2.8	2.9
Completeness (%)	98.50 (98.5)	100.00 (100)	97.3 (96.0)	98.1 (96.4)	94.0 (92.0)
Multiplicity	13.10 (12.3)	12.20 (11.1)	12.0 (10.8)	12.3 (11.0)	12.4 (12.2)
Mean I/sI	15.90 (3.3)	14.30 (3.4)	17.3 (3.5)	13.5 (2.67)	30.5(2.4)
R_{int}	6.82 (55.0)	6.25 (48.0)	6.32 (47.2)	6.31 (55.0)	6.3(44.4)
R_{pim}	1.98 (15.5)	1.84 (14.9)	1.85 (14.5)	1.83 (16.8)	1.8(12.4)

X06DA Data Visualization

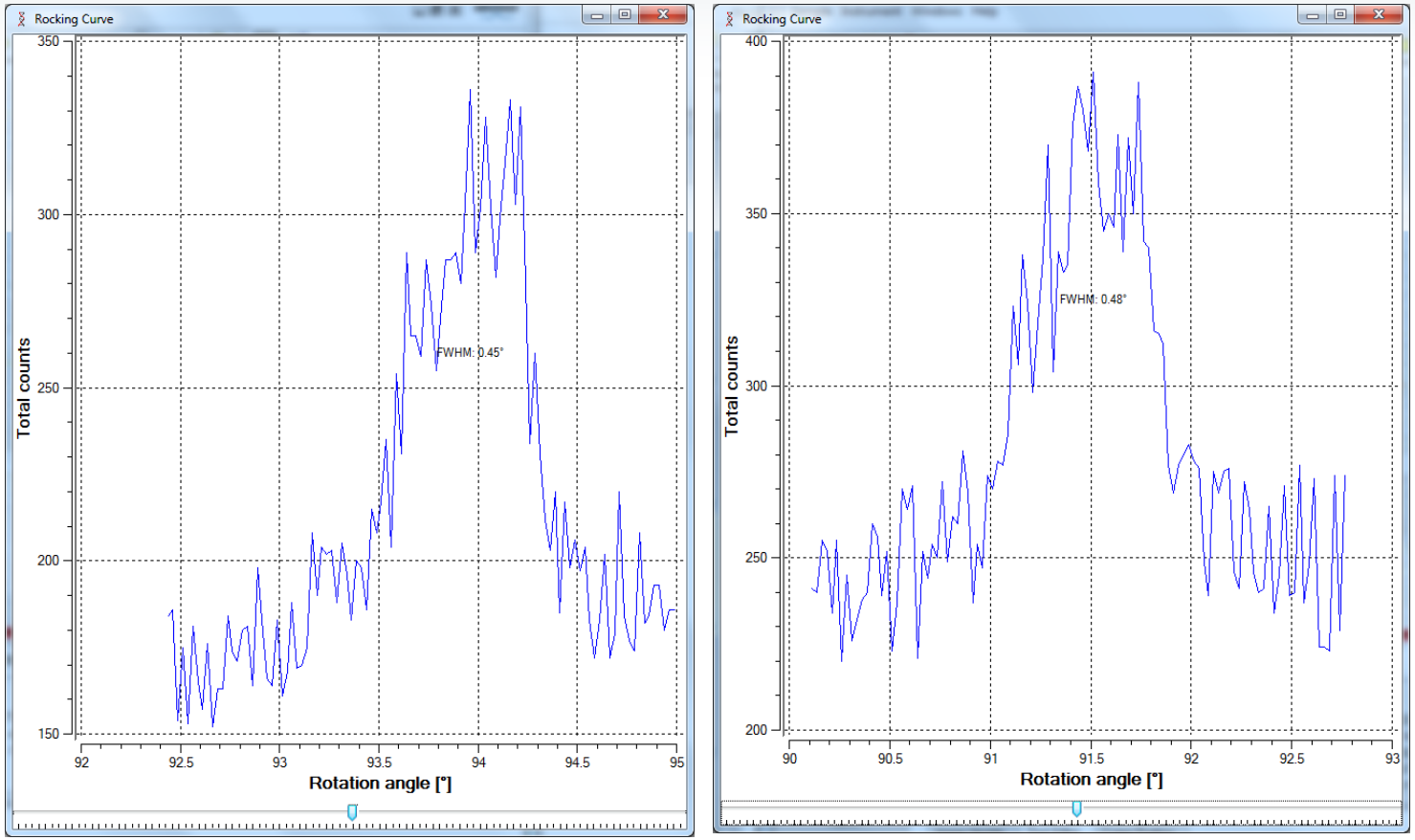






- The very small spots are very difficult to see
- 3D tools help to better visualize diffraction spots

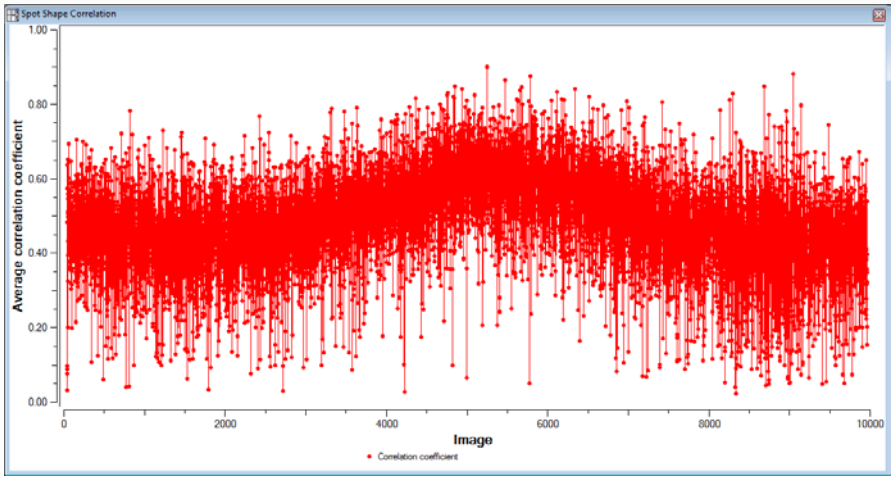
X06DA Data Rocking Curve



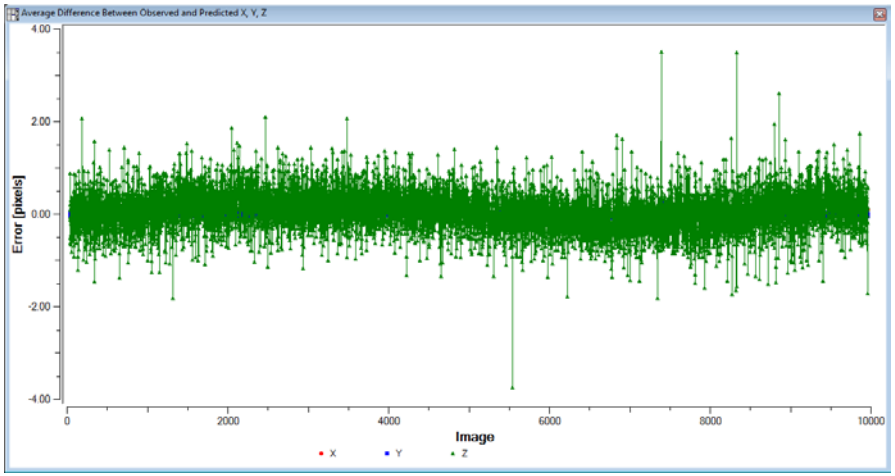
- Refection rocking curve of about 0.5°

X06DA Data Integration

- Spot shape correlations are pretty low
  - SAINT problem with super-fine slicing?
  - Oversampling?
  - Low intensity?
  - Poor 2-D spot info?

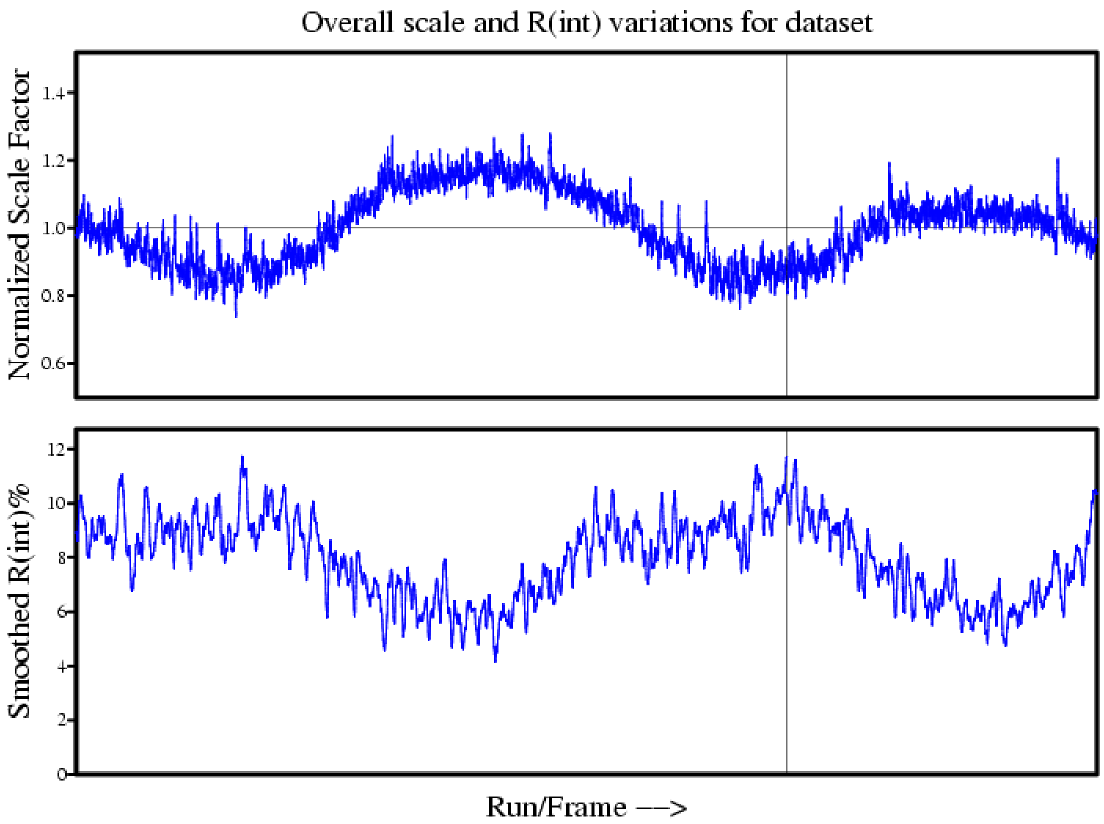


- Z-error in pixels are very high
  - Scan direction in Kabsch space
  - Goniometer problems?
  - Synchronization problems?
  - Loose pin?



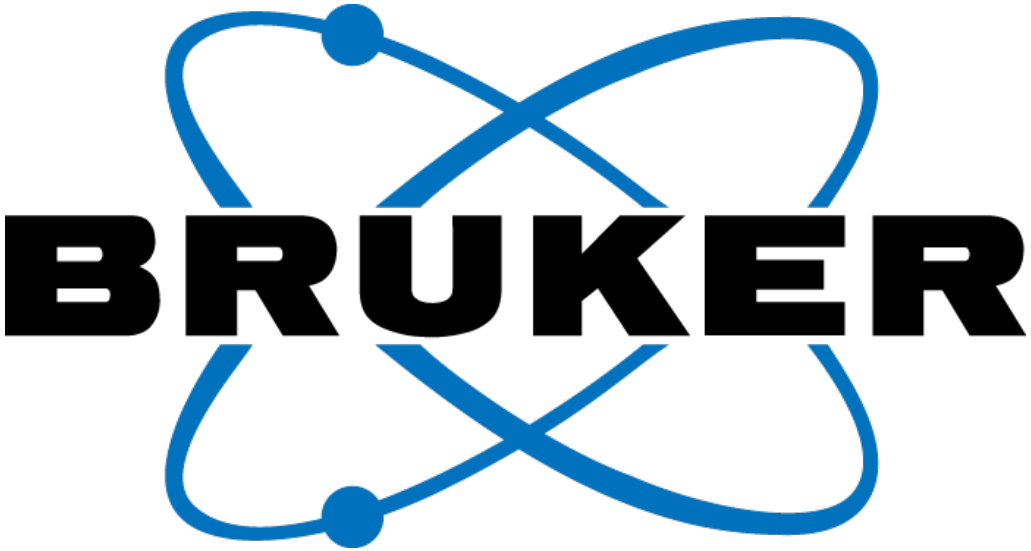
X06DA Data Scaling

- “too much” structure on normalized scaling factor versus frame number
- “too much” structure on smoothed Rint versus frame number



Summary

- Highly redundant fine slicing data from Kappa goniometers becomes more prevalent at synchrotron beam-lines
- PROTEUM2 provides a full suite of crystallographic tools to process these data
- PROTEUM2 has unique features and tools for data and result visualization
- PROTEUM2 is the only package that handles non-merohedrally twinned data appropriately



Innovation with Integrity