MASSIF-1: Player piano dystopia or brave new world?

Matthew W. Bowler and Didier Nurizzo
• Fully autonomous – no user control
• Arinax MD2s
• EMBL Flex SC – 368 sample capacity
• Pilatus3 2M
Changing projects

MASSIF1 post EBS

Size: 265 x 65 µm² FWHM
Flux: $3.1 \times 10^{12}$ ph/s

Size: 130 x 29 µm² FWHM
Flux: $8 \times 10^{12}$ ph/s
MASSIF1 post EBS

Size: 265 x 65 µm² FWHM
Flux: 3.1 x 10^{12} ph/s

Size: 130 x 29 µm² FWHM
Flux: 8 x 10^{12} ph/s
Fully automatic (unattended) data collection – what do we gain?

• 40,000 person hours for data collection and travel

• €1,070,400 in expenses and

• 625,845 kg of CO$_2$

(data from @EU_Eurostat)
Thanks to Tony Warne

β₁ adrenergic GPCR
<table>
<thead>
<tr>
<th>Diffraction Plan entry</th>
<th>Definition</th>
<th>Default value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protein acronym</td>
<td>Defines the protein that is registered with the ESRF safety group</td>
<td>Required field</td>
</tr>
<tr>
<td>Sample name</td>
<td>User defined unique identifier</td>
<td>Required field</td>
</tr>
<tr>
<td>Pin barcode</td>
<td>Barcode identifier</td>
<td>none</td>
</tr>
<tr>
<td>Experiment type</td>
<td>Define MXPressE / O / SAD / Score</td>
<td>MXPressE</td>
</tr>
<tr>
<td>Space Group</td>
<td>If present used for strategy calculation and autoprocessing</td>
<td>none</td>
</tr>
<tr>
<td>Pre-observed resolution</td>
<td>Resolution that the detector will be set to for mesh scans, characterisation images and default data collection</td>
<td>2.0 Å</td>
</tr>
<tr>
<td>Required resolution</td>
<td>Threshold resolution, samples below cutoff will not be collected</td>
<td>none</td>
</tr>
<tr>
<td>Radiation sensitivity</td>
<td>BEST input in case of highly radiation sensitive crystals</td>
<td>1</td>
</tr>
<tr>
<td>Required completeness</td>
<td>-</td>
<td>99%</td>
</tr>
<tr>
<td>Required multiplicity</td>
<td>-</td>
<td>4</td>
</tr>
<tr>
<td>Number of positions</td>
<td>For multiple crystals</td>
<td>1</td>
</tr>
<tr>
<td>Preferred beamsize</td>
<td>Select appropriate beamsize for crystals</td>
<td>50 µm</td>
</tr>
<tr>
<td>Diffraction Plan entry</td>
<td>Definition</td>
<td>Default value</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>---------------</td>
</tr>
<tr>
<td>Protein acronym</td>
<td>Defines the protein that is registered with the ESRF safety group</td>
<td>Required field</td>
</tr>
<tr>
<td>Sample name</td>
<td>User defined unique identifier</td>
<td>Required field</td>
</tr>
<tr>
<td>Pin barcode</td>
<td>Barcode identifier</td>
<td>none</td>
</tr>
</tbody>
</table>

**Required completeness** - 99%

**Required multiplicity** - 4

**Number of positions** For multiple crystals

**Preferred beamsize** Select appropriate beamsize for crystals 50 µm

---

**Svensson et al.**

Volume 71 | Part 8 | August 2015 | Pages 1757–1767
Software routines locate crystals and centre to best volume.

Characteristics such as beam size and flux as well as crystal volume lead to highly optimised data collection.
Svensson et al.
Volume 71 | Part 8 | August 2015 | Pages 1757–1767
COG determination
COG determination
Diffraction signal

Crystal map

Melnikov et al.

Volume 74 | Part 4 | April 2018 | Pages 355–365 | 10.1107/S2059798318002735
Pseudo-helical data collection – first automated helical workflow that accounts for crystal variability

- Full ‘normal’ data set the subsequent helical
- Auto-peak selection or user defined
- Stringent thresholding – regions within 30%
- SAD option available
Flu polymerase

2 Å
Flu polymerase

3 Å
“Fata Morgana”
The difference measurements make......

<table>
<thead>
<tr>
<th></th>
<th>Crystal larger than beam</th>
<th>Crystal smaller than beam</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Default Crystal Size</td>
<td>Measured Crystal Size</td>
</tr>
<tr>
<td><strong>Crystal dimensions (µm)</strong></td>
<td>100 x 100 x 100</td>
<td>603 x 238 x 397</td>
</tr>
<tr>
<td><strong>Beam diameter (µm)</strong></td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td><strong>Space group</strong></td>
<td>C2_1</td>
<td>C2_1</td>
</tr>
<tr>
<td><strong>Unit cell dimensions</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a, b, c (Å)</td>
<td>129.00, 208.38, 117.50, 90, 109.3, 90</td>
<td>129.00, 208.38, 117.50, 90, 109.3, 90</td>
</tr>
<tr>
<td>α, β, γ (°)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Flux (ph/s)</strong></td>
<td>1.2 x 10^{12}</td>
<td>1.2 x 10^{12}</td>
</tr>
<tr>
<td><strong>Transmiision (%)</strong></td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td><strong>Dose (MGy)</strong></td>
<td><strong>13.23</strong></td>
<td><strong>36.38</strong></td>
</tr>
<tr>
<td><strong>Total exposure time (s)</strong></td>
<td>178.9</td>
<td>507.1</td>
</tr>
<tr>
<td><strong>Detector resolution (Å)</strong></td>
<td>2.16</td>
<td>2.05</td>
</tr>
</tbody>
</table>
β₁ adrenergic GPCR

Thanks to Tony Warne
### β₁ adrenergic GPCR

<table>
<thead>
<tr>
<th>Crystal</th>
<th>Crystal dimensions (x, y, z, mm)</th>
<th>Fixed beam diameter</th>
<th>Adaptable beam diameter</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Resolution limit (Å)</td>
<td>&lt;I/σ(I)&gt;</td>
</tr>
<tr>
<td>adrcpt-For41</td>
<td>0.109 x 0.053 x 0.025</td>
<td>3.77</td>
<td>6.7</td>
</tr>
<tr>
<td>adrcpt-For42</td>
<td>0.084 x 0.025 x 0.025</td>
<td>4.22</td>
<td>4.3</td>
</tr>
<tr>
<td>adrcpt-For45</td>
<td>0.035 x 0.045 x 0.051</td>
<td>3.95</td>
<td>6.2</td>
</tr>
<tr>
<td>adrcpt-For47</td>
<td>0.105 x 0.061 x 0.051</td>
<td>3.74</td>
<td>4.7</td>
</tr>
<tr>
<td>adrcpt-For48</td>
<td>0.105 x 0.039 x 0.064</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>adrcpt-For58</td>
<td>0.169 x 0.050 x 0.061</td>
<td>3.88</td>
<td>6.6</td>
</tr>
<tr>
<td>adrcpt-For59</td>
<td>0.042 x 0.024 x 0.025</td>
<td>3.25</td>
<td>9.2</td>
</tr>
<tr>
<td>adrcpt-For67</td>
<td>0.064 x 0.026 x 0.031</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
## β₁ adrenergic GPCR

<table>
<thead>
<tr>
<th>Crystal</th>
<th>Crystal dimensions (x, y, z, mm)</th>
<th>Fixed beam diameter</th>
<th>Adaptable beam diameter</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Resolution limit (Å)</td>
<td>&lt;I/σ(I)&gt;</td>
</tr>
<tr>
<td>adrcpt-For41</td>
<td>0.109 x 0.053 x 0.025</td>
<td>3.77</td>
<td>6.7</td>
</tr>
<tr>
<td>adrcpt-For42</td>
<td>0.084 x 0.025 x 0.025</td>
<td>4.22</td>
<td>4.3</td>
</tr>
<tr>
<td>adrcpt-For45</td>
<td>0.035 x 0.045 x 0.051</td>
<td>3.95</td>
<td>6.2</td>
</tr>
<tr>
<td>adrcpt-For47</td>
<td>0.105 x 0.061 x 0.051</td>
<td>3.74</td>
<td>4.7</td>
</tr>
<tr>
<td>adrcpt-For48</td>
<td>0.105 x 0.039 x 0.064</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>adrcpt-For58</td>
<td>0.169 x 0.050 x 0.061</td>
<td>3.88</td>
<td>6.6</td>
</tr>
<tr>
<td>adrcpt-For59</td>
<td>0.042 x 0.024 x 0.025</td>
<td>3.25</td>
<td>9.2</td>
</tr>
<tr>
<td>adrcpt-For67</td>
<td>0.064 x 0.026 x 0.031</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
## B<sub>1</sub> adrenergic GPCR

<table>
<thead>
<tr>
<th>Crystal</th>
<th>adrcpt-For42</th>
<th>adrcpt-For42</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beam diameter (μm)</td>
<td>50</td>
<td>30</td>
</tr>
<tr>
<td>Space group (no.)</td>
<td>P2&lt;sub&gt;1&lt;/sub&gt;,2&lt;sub&gt;1&lt;/sub&gt;, (19)</td>
<td>P2&lt;sub&gt;1&lt;/sub&gt;,2&lt;sub&gt;1&lt;/sub&gt;, (19)</td>
</tr>
<tr>
<td>Cell parameters (Å, °)</td>
<td>116.8, 121.1, 129.5, 90, 90, 90</td>
<td>116.5, 120.78, 128.7, 90, 90, 90</td>
</tr>
<tr>
<td>Flux (ph/s)</td>
<td>5.7 x 10&lt;sup&gt;11&lt;/sup&gt;</td>
<td>2.2 x 10&lt;sup&gt;11&lt;/sup&gt;</td>
</tr>
<tr>
<td>Rotation width (°)</td>
<td>0.15</td>
<td>0.05</td>
</tr>
<tr>
<td>Total oscillation Range (°)</td>
<td>149.1</td>
<td>124.0</td>
</tr>
<tr>
<td>Total dose (MGy)</td>
<td>5.25</td>
<td>5.92</td>
</tr>
<tr>
<td>Detector resolution</td>
<td>4.05</td>
<td>3.9</td>
</tr>
<tr>
<td>Wilson B-factor (Å&lt;sup&gt;2&lt;/sup&gt;)</td>
<td>94.0</td>
<td>97.0</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>48.2 – 3.95 (4.09 – 3.95)</td>
<td>48.7 – 3.53 (3.66-3.53)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>98.2 (90.9)</td>
<td>81.5 (33.0)</td>
</tr>
<tr>
<td>Observed reflections</td>
<td>15,940 (1413)</td>
<td>18,785 (721)</td>
</tr>
<tr>
<td>Average redundancy</td>
<td>4.1 (2.7)</td>
<td>4.4 (1.6)</td>
</tr>
<tr>
<td>&lt; I/σ(I)&gt;</td>
<td>4.3 (0.8)</td>
<td>10.6 (0.6)</td>
</tr>
<tr>
<td>R&lt;sub&gt;meas&lt;/sub&gt;</td>
<td>0.23 (1.45)</td>
<td>0.1 (1.64)</td>
</tr>
<tr>
<td>R&lt;sub&gt;merge&lt;/sub&gt;</td>
<td>0.18 (1.05)</td>
<td>0.077 (1.17)</td>
</tr>
<tr>
<td>CC&lt;sub&gt;1/2&lt;/sub&gt;</td>
<td>0.99 (0.43)</td>
<td>1 (0.33)</td>
</tr>
</tbody>
</table>
Molecular-weight dependence of the minimum required crystal size
Molecular-weight dependence of the minimum required crystal size

![Graph showing the relationship between molecular weight and crystal size for different resolutions.](image)
Change existing CRL

Add additional CRL
100 µm diameter beam

Flux: $5 \times 10^{12}$ ph/sec
75 µm diameter beam

Flux: $5 \times 10^{12}$ ph/sec
50 µm diameter beam

Flux: $5 \times 10^{12}$ ph/sec
10 µm diameter beam

Flux: \( 5 \times 10^{12} \text{ ph/sec} \)

Flux: \( 5 \times 10^{11} \text{ ph/sec} \)
Molecular-weight dependence of the minimum required crystal size

![Graph showing molecular-weight dependence](image-url)
HC1 and CrystalDirect allow automated RT
Automated humidity control experiments
Automated humidity control experiments
Not just for lysozyme!

Scc3/Scc1 complex with DNA: – 600 crystals final 3.8 Å data set

p52 and ETS1: 300 crystals final 3.0 Å data set

ARS2: 300 crystals Se-Met phasing at 4.0 Å
**Example** – Bayer discover a new anti-inflammatory targeting Tank Binding Kinase (TBK)

- Problem: Crystals diffract poorly (1 in 20 good) but many fragments need to be screened
- Automation allowed large number of molecules to be screened (20 per molecule, over 1000 total) without human presence
- Optimised data collection allowed high enough resolution to be obtained to develop the molecule

Acknowledgements:

Didier Nurizzo
Pascal Theveneau
Matias Guijarro
Olof Svensson
Debby Davison
Marcus Oscarsson
Marjolaine Bodin
Stephanie Monaco
Gordon Leonard
Christoph Mueller-Dieckmann
David Flot
Max Nanao
Andrew McCarthy

All EMBL and SB groups

User Office
Safety Group
ATF

Thierry Giraud
Fabien Dobias
Hugo Caserotto
Nicolas Guichard
John Surr
Mario Lentini

Stephen Cusack
Florent Cipriani
Andrew McCarthy
Josan Marquez
Gergely Papp
Serena Rocchio

Instrumentation and HTX teams