Minutes

Meeting title	CCP4 Working Group 2 meeting (chaired by Jon Agirre, University of York)		
Date	Tuesday 19 March 2024	**Time**	10.00 – 12.00
Location	Virtual only	**Zoom**	Zoom link
Circulation	ccp4wg2@stfc.ac.uk		
Attending	Airlie McCoy, Arnaud Basle, Jon Agirre, Ben Bax, Charles Ballard, Keitaro Yamashita, Marcin Wojdyr, Eugene Krissinel, Garib Murshudov, Roberto Steiner, Alan Roseman, Briony Yorke, Clemens Vornhein, Dave Briggs, David Waterman, Deborah Harrus, Elke de Zitter, Filipe Menezes, Gustavo Lima, Hanna Kwon, Ivo Tews, Jacques-Philippe Colletier, Johan Turkunburg, Jordan Dialpuri, Judit Debreczeni, Karen McIntyre, Lou Holland, Lucy Schofield, Luiso Fuentes-Montero, Martin Maly, Nicola Coquelle, Paul Bond, Peter Moody, Rob Nicholls, Robbie P. Joosten, Ronan Keegan, Stuart McNicholas, Thao Pham, Jools Wills, Nick Pearce, Adrian Lapthorn, Dan Rigden, Ville Uski.		
Apologies	Mike Hough, Dave Brown, Tony Oliver, Robert Esnouf, Randy Read, Maria Fando, Mark Roe, Jon Grimes, Paul Emsley.		

10.00 **Welcome**

Jon gives the agenda.

10.05 **WG2 Chair's update (Jon Agirre)**

- Approval of February WG2 minutes
- Today's agenda

10:10 **Update on next release of CCP4 (Charles Ballard)**

Issues putting updates together for 8.0. Advise that any new developments are targeted at 9.0. Bundles available for 9.0. Going for 0.9 Coot for scripting and the current version of CCP4i2. There will be patch updates and minor (9.1, 9.2…) updates. DIALS, DUI-2 servalcat, monomer all going into 9.0. Also up to date versions of PDB-REDO tools (e.g. Tortoize).

At some point, Martin Noble's DJango backend for CCP4i2 will be released (9.0, but will need testing with those bundles).

Coot headless API (CHAPI) going in 9.0.

Charles had 9.0 for a year now. Coot 1.1?

9.0 caught up with the latest PDB-REDO. PDB needs that updated software.

CCP4 9.0 bundles: https://ccp4serv6.rc-harwell.ac.uk/9.0/one-page.html

Will probably update to Python 3.11 or 3.12 when 9.0 is shipped. Supported systems will go up. Red Hat 8.0 minimum. Mac OS X will go up too, but not sure by how much.

Charles encourages people to use WSL, Windows Subsystem for Linux (supports Ubuntu, Debian, etc.). But graphics drivers can be a bit problematic.

New Gemmi is included. Servalcat depends on that.

10.15 **Northern meeting debrief (Briony Yorke)**

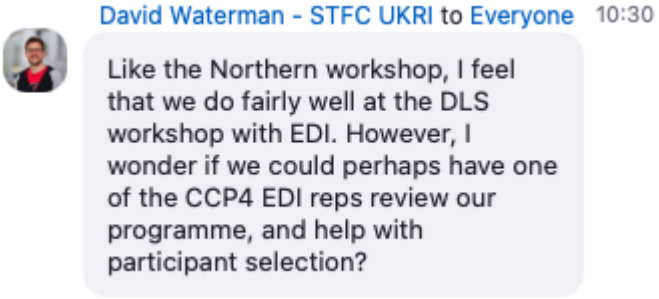
90 attendees. No delegation from Scotland (one from Glasgow). Northwest also missing (Liverpool And Manchester), which was also visible at York. ¾ ECR. Advertisers found they got more interaction from ECRs, so were very happy they sponsored the meeting.

10.20 **Update on ED&I plans (Briony Yorke & Lucy Schofield)**

Lucy Schofield presents a summary of the latest discussions within the ED&I team, now integrated by Lucy herself, Briony Yorke, Mads Gabrielsen, Karen McIntyre and Jon Agirre.
Team can become a leader in the field compared to other groups.

Change of culture needed in CCP4 for example grant-writing demanding hours out of working time, thus primary caregivers cannot contribute equally.

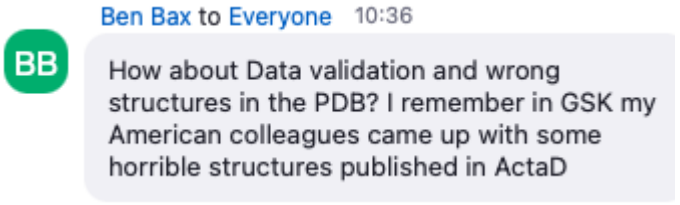
Briony and Arwen Pearson have started a discussion group for female and non binary coders.



ED&I team are happy to help.

Briony Yorke introduces a new idea to address the leaky pipeline problem in STEM, which is clearly affecting us – lots of diversity at ECR level, gradually lost throughout career progression: she has created a discussion group for female and non-binary coders. Another idea for which Briony will seek funds with CCP4 is a boot camp, where people can learn crystallographic coding skills with experts in the field. Supported by Garib, Ivo and Jon. Briony to submit an application to CCP4 exec for funds.

10.25 **Study Weekend 2025: topic and organisers (Jon Agirre & Arnaud Basle)**



Some support for Validation at every step of the process. Roberto: add QM for validation, additional focus on ligands (interesting point – last talk today is on QM - in-pocket).

Garib: model is much bigger than x,y,z.

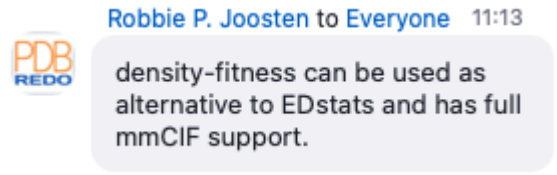
The discussion converges on constructive validation/verification of modelling.
We nominate a few potential organisers, to be contacted by WG2 chairs.

11.00 **The road to full mmCIF support, five-letter monomer codes in CCP4 (Eugene Krissinel)**
Eugene outlines the many benefits of transitioning CCP4 to full mmCIF support. For example having mmCIF instead of MTZ files – mmCIF are plain text.

No effort to go into mmCIF-ication of CCP4i, despite it being used possibly more than CCP4i2 and cloud together.

Support for mmCIF fades as we progress from phasing to validation and deposition. One exception is Modelcraft.

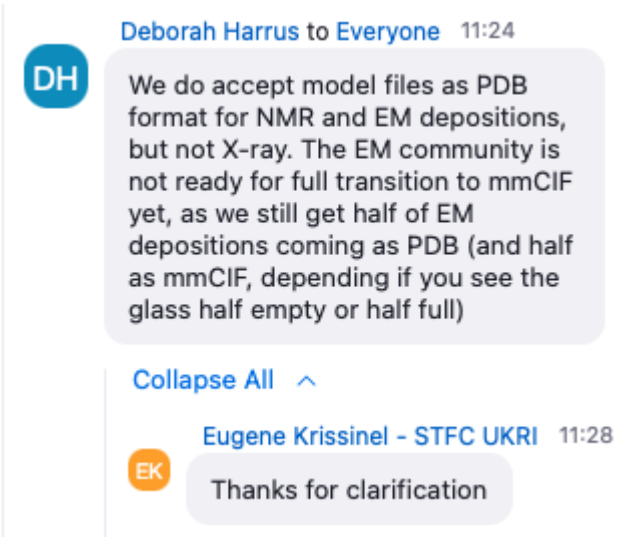
Eugene has a list of programs including supported, unsupported and deprecated ones. 51 programs most commonly used in CCP4i2 and Cloud, 10 of those are mmCIF ready, others awaiting assessment. 7 programs exposed in CCP4i2 are not mmCIF ready. 32 programs are in a low priority list. 6 additional programs are unsupported, 6 are deprecated. Total of 103 programs.



It's not just ligand codes, but also chain names. Deborah adds the following on that:

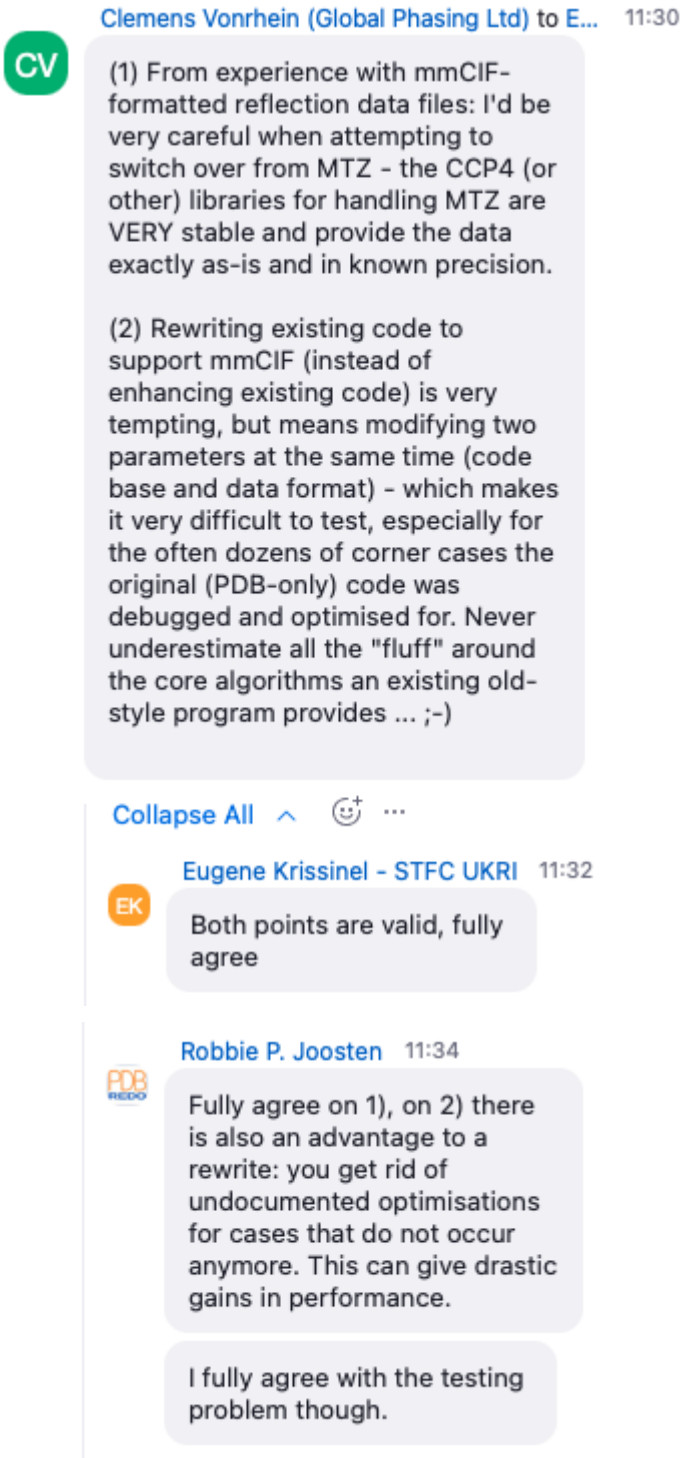
Ben Bax proposes to have converters.

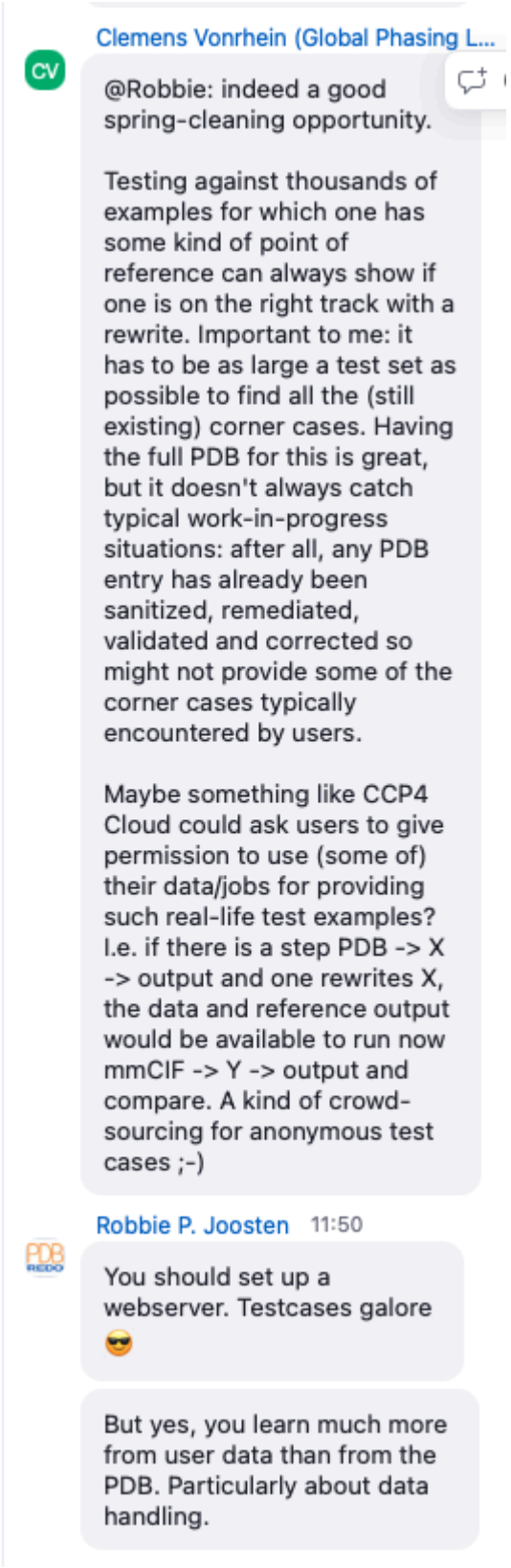
Judit says that there are lots and lots of tools in industry that use the PDB format at a low level. These could be rewritten, but would be a huge job.



Eugene to make the list public so that users can follow updates on this.

Paul Bond shares a link to the convert utility in the Gemmi module/package: https://gemmi.readthedocs.io/en/latest/utils.html#convert


Clemens Vonrhein (Global Phasing Ltd) to E... 11:30
CV
(1) From experience with mmCIF-formatted reflection data files: I'd be very careful when attempting to switch over from MTZ - the CCP4 (or other) libraries for handling MTZ are VERY stable and provide the data exactly as-is and in known precision.
(2) Rewriting existing code to support mmCIF (instead of enhancing existing code) is very tempting, but means modifying two parameters at the same time (code base and data format) - which makes it very difficult to test, especially for the often dozens of corner cases the original (PDB-only) code was debugged and optimised for. Never underestimate all the "fluff" around the core algorithms an existing old-style program provides ... ;-)
Collapse All
Eugene Krissinel - STFC UKRI 11:32
EK
Both points are valid, fully agree
Robbie P. Joosten 11:34
PDB REDO
Fully agree on 1), on 2) there is also an advantage to a rewrite: you get rid of undocumented optimisations for cases that do not occur anymore. This can give drastic gains in performance.
I fully agree with the testing problem though.

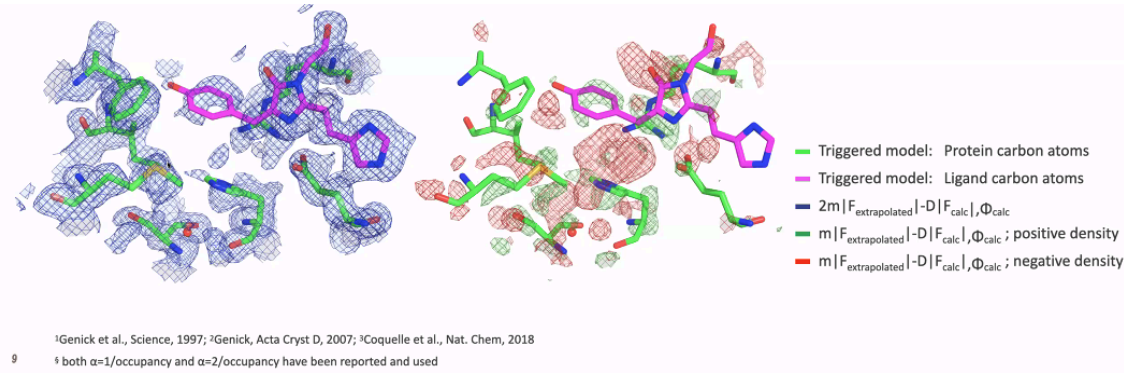


11.20 **Xtrapol8 (Elke de Zitter)**

Elke introduces the software, talks about low occupancy states.

Xtrapol8 uses Bayesian weighted Fourier difference electron density maps.
Calculates extrapolated structure factors to mimic the triggered state at full occupancy.

Low occupancy can be a consequence of a number of things: crystal preparation, time-resolved experiments, radiation damage, ligand binding…

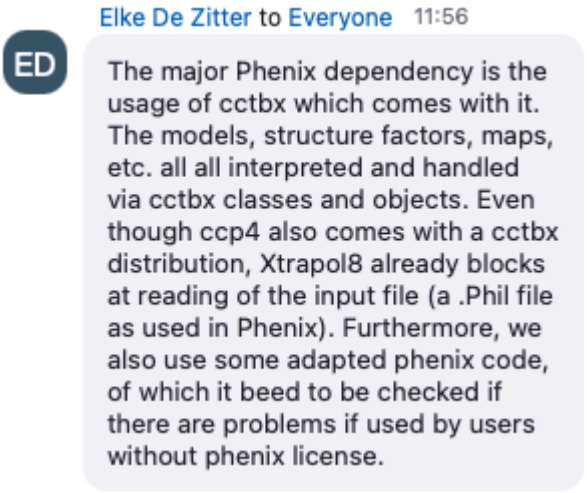


Different methods for extrapolation: based on difference maps, based on the output of refinement. Test different occupancies input by the user.

Next versions will explore relationship between alpha and occupancy of the triggered state, improvements in the refinement strategy, parallelisation, mmCIF generation for deposition, real space methods and compatibility with CCP4 – obviously critical for inclusion in the suite (discussed below).

GitHub: https://github.com/ElkeDeZitter/Xtrapol8
Need to have CCP4, Phenix and Coot installed to use it.
The package is Python 3 compatible, generates mmCIF for deposition.

Elke demos the PHENIX interface of the software.



11.40 **In-pocket (Filipe Menezes)**

The code requires the ULYSSES package, which is available in the repository: https://gitlab.com/siriius/ulysses
Incorporates a lot of methods already: PMx, RM1, PM3-BP, AM1, PM3-PDDG…
Open Babel required for in-pocket.

At the centre of the package is the calculation of deformation energy. This is the energy required to go from ligand free to ligand bound, as the enzyme deforms the ligand to accommodate it. Or, put into a formula, $E_{def} = E_{bound} - E_{free}$

Can get distances and bond orders after optimization. Calculated maps using QM look good.
In pocket can be useful in determining H-bond networks, but also for validation.
Software runs very quickly.

Brilliant 2D information on the interactions (Ligplot+ style with energies), before and after optimisation. Great example where a H-bond is gained with a side-chain flip.

This software could find lots of uses in CCP4, Eugene and Jon say.

ULYSSES is dependent on Eigen. Eugene says it's fine.
ULYSSES/in-Pocket licence is LGPL 2.0.

GitLab: https://gitlab.com/siriius/ulysses

12.00 **AOB**