CCP-EM Workshop

CCP-EM

8-12 Morning Session: SPA Reconstruction

8:00	Lecture - Introduction to CCP-EM and Doppio (Tom Burnley)
8:20	Lecture - Overview of Relion SPA workflow (Matt Iadanza)
9:00	Practical - Relion part 1 (getting started, review 2D classes, initial model generation, 3D classification)
10:00	*Coffee*
10:45	Practical - Relion part 2 (Select 3D classes, extract, Refine 3D, Mask create, Post Process)
12-1	*Lunch*

1-4 Afternoon Session: Model Building and Validation

1:00	Lecture - Overview of Model Building and Refinement tools in CCP-EM (Tom)
1:30	Practical - Refinement using Molrep & Refmac
2:00	Break
2:30	Lecture - Model Validation in CCP-EM (Agnel Joseph) [10am Chennai, India]

Better Wifi (limited to 20):

CCP4
Workshop1978

STFC Cloud VM Signup:

tom.burnley@stfc.ac.uk

UKRI Science and Technology Facilities Council


Tom


Matt


Agnel

Science and Technology Facilities Council

CCP-EM Software Suite

Tools for CryoEM

Tom Burnley

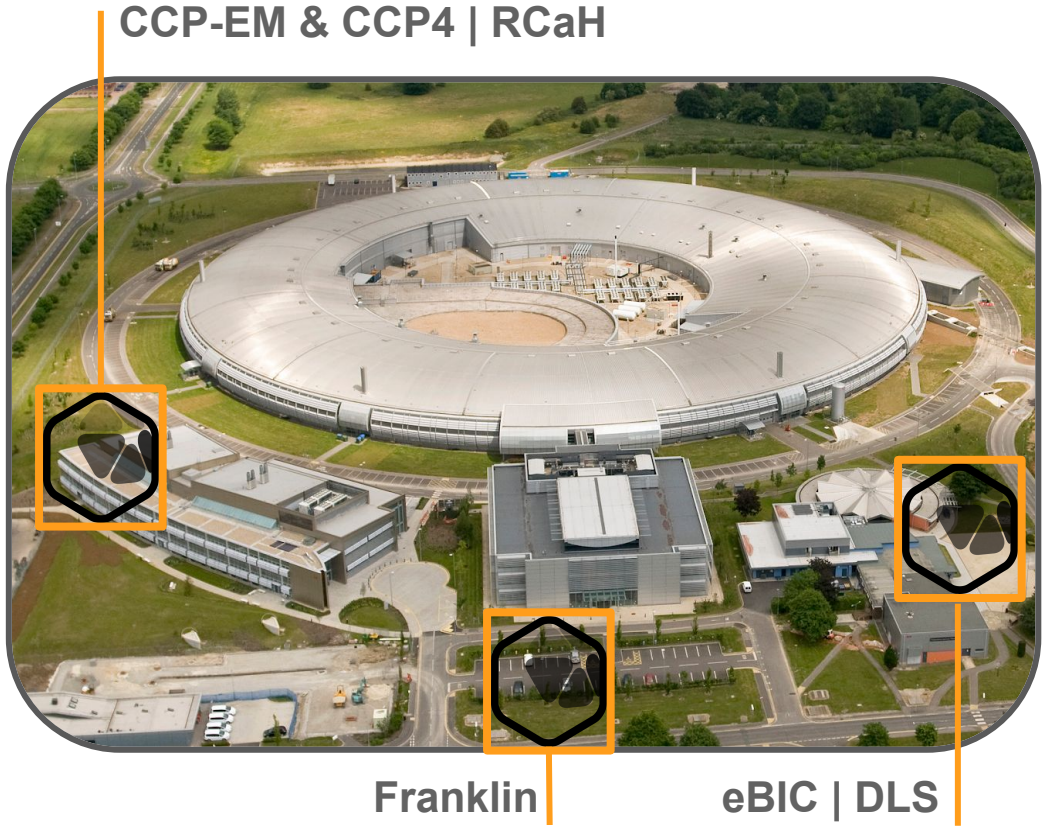
22 August 2023

IUCr CCP-EM Workshop

CCP-EM

What is CCP-EM?

- Support users and developers in computational aspects of biological EM
- Based at STFC RAL national laboratory
- CCP-EM software suite
- EM community (>3.5K subscribers)
- Software users (>1200 downloads, >30 industrial licences)
- Support developers (>10 external groups)



Science and Technology Facilities Council

Research Complex at Harwell

CCP-EM core team

CCP4 core team

STFC Business & Innovation

CCP-EM Commercial License Holders

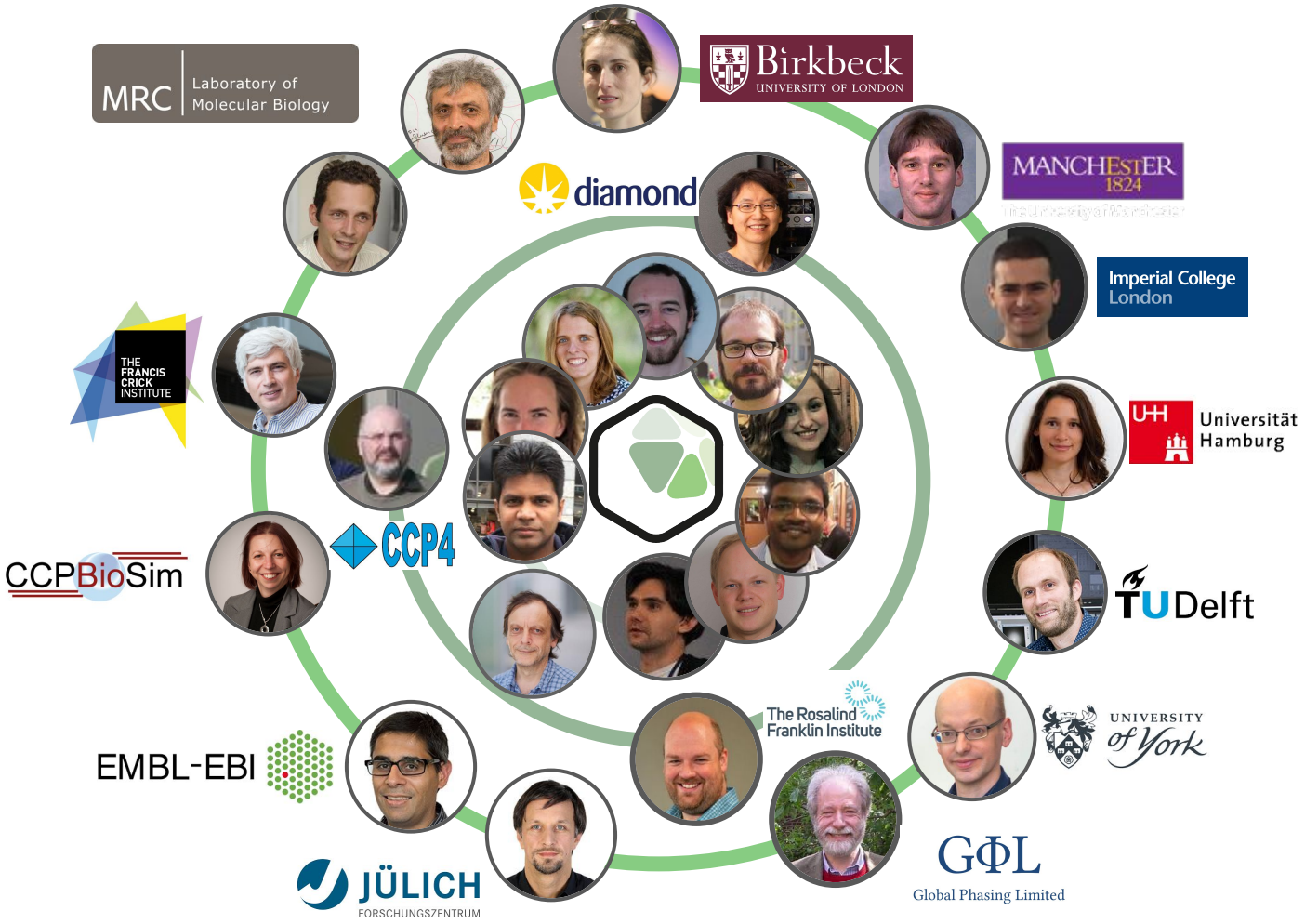
CCP-EM Collaborators

Website: www.ccpem.ac.uk

Mailing list: www.jiscmail.ac.uk/ccpem

Twitter: @ccp_em

Email: ccpem@stfc.ac.uk



Science and Technology Facilities Council

Medical Research Council

wellcome

Department for Business, Energy & Industrial Strategy

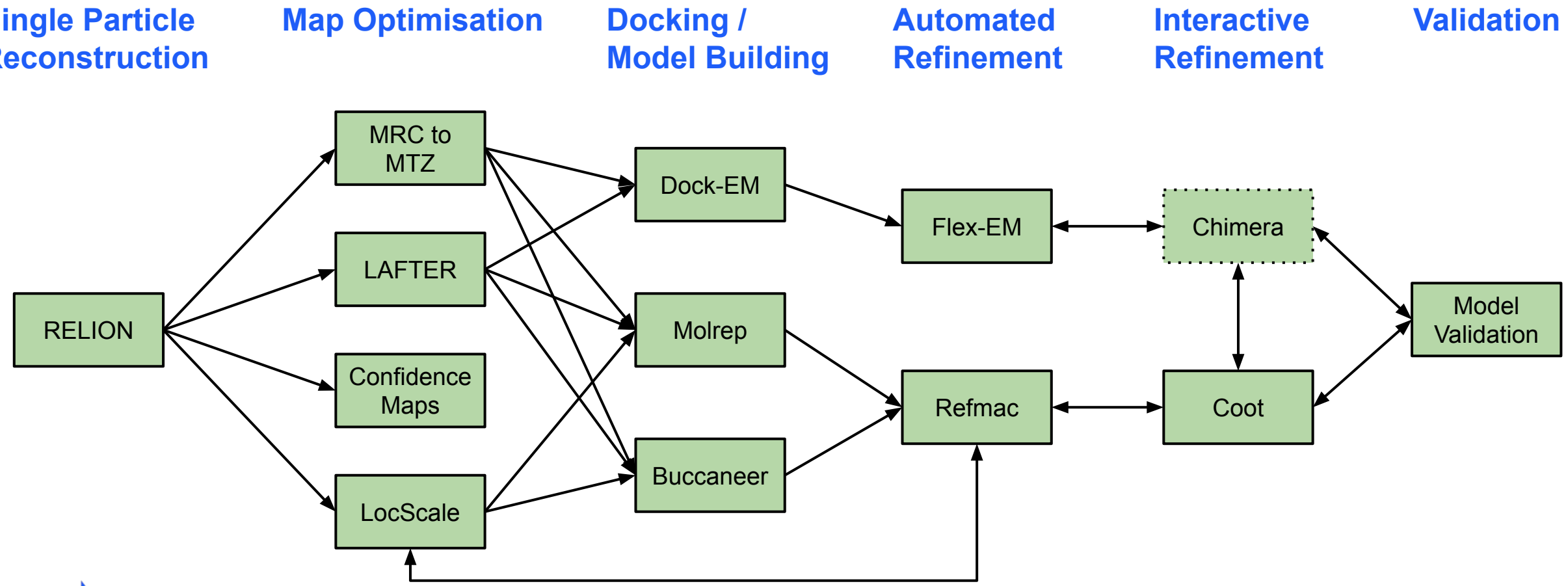
Research Complex at Harwell

CCP-EM

CCP-EM Software Suite

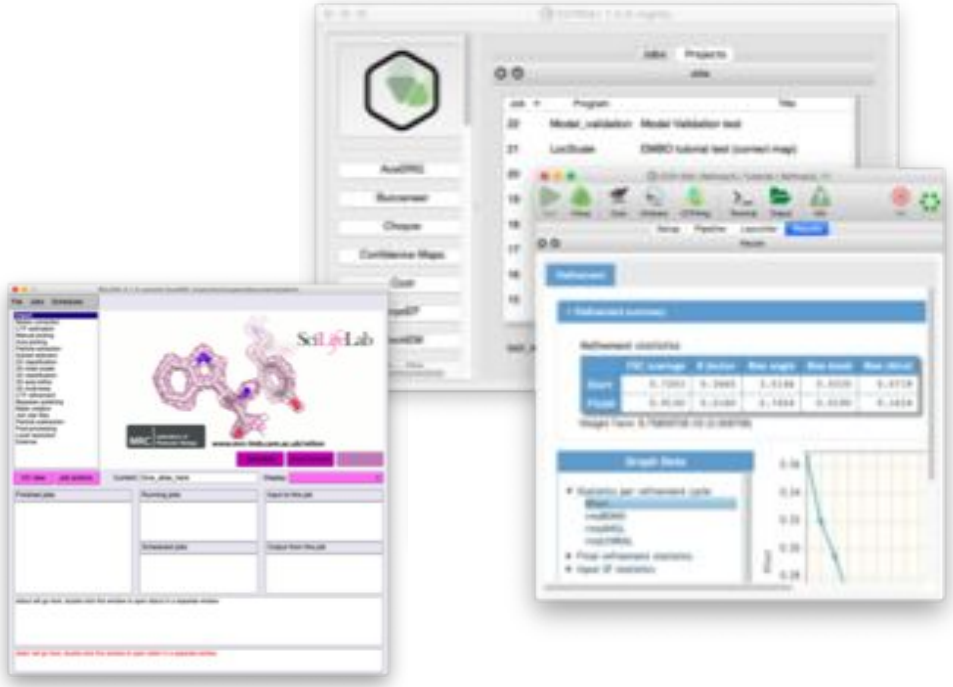
CCP-EM 2.0

CCP-EM workflow



Science and Technology Facilities Council

CCP-EM version 1.x



CCP-EM

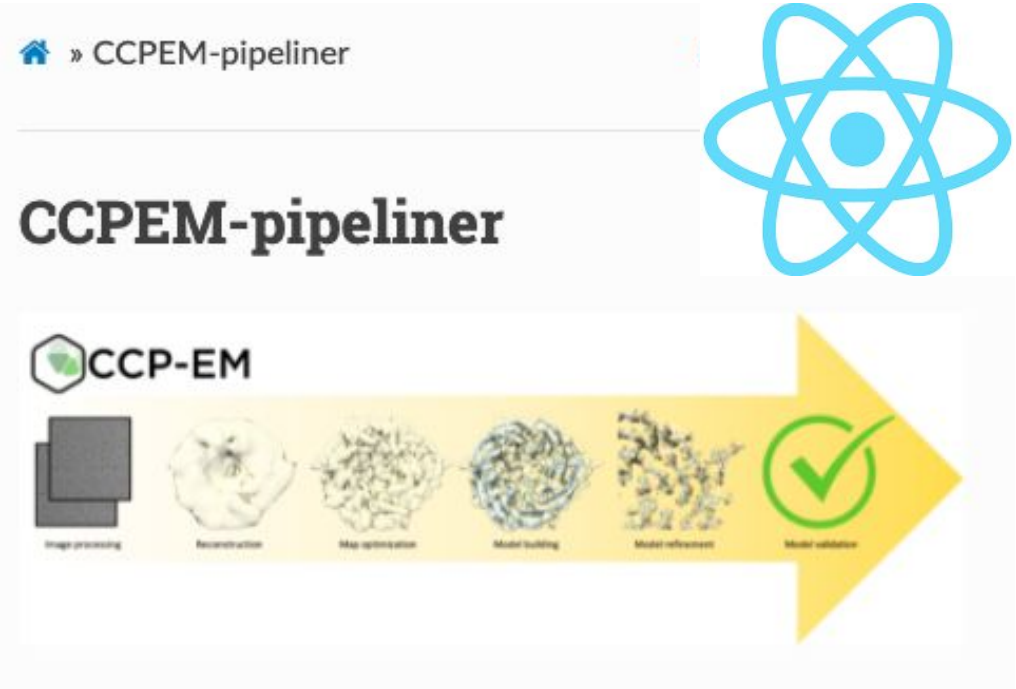
- Map post processing
- Atomic modelling
- PyQt5 GUI
- No data model
- Basic project management

RELION

- Single particle
- FLTK GUI
- Data model
- DAG project management

Science and Technology Facilities Council

CCP-EM version 2.x



Images to structures

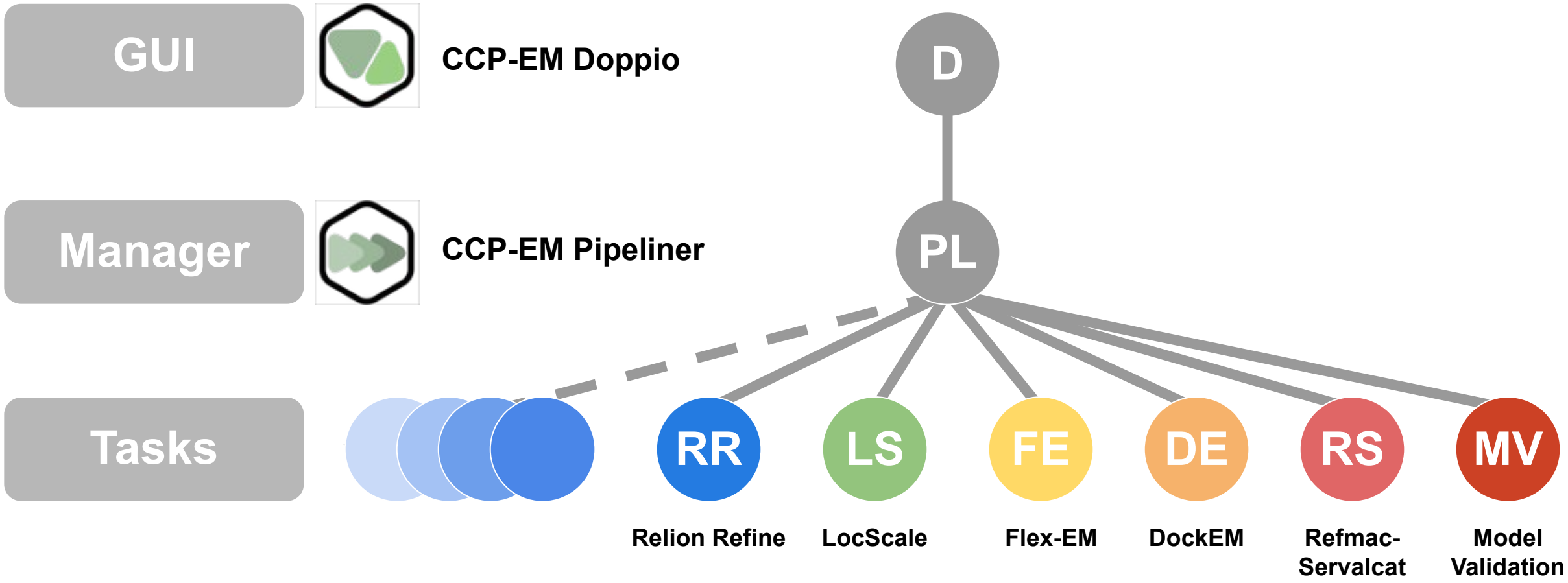
- Python pipeliner python API
- JS react GUI

- Updated data model
- Project management

- Single particle
- Atomic modelling
- Tomography (coming soon)

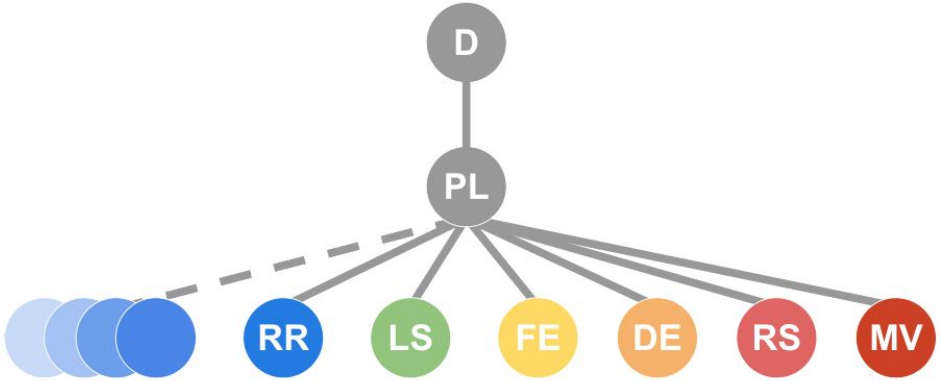
Science and Technology Facilities Council

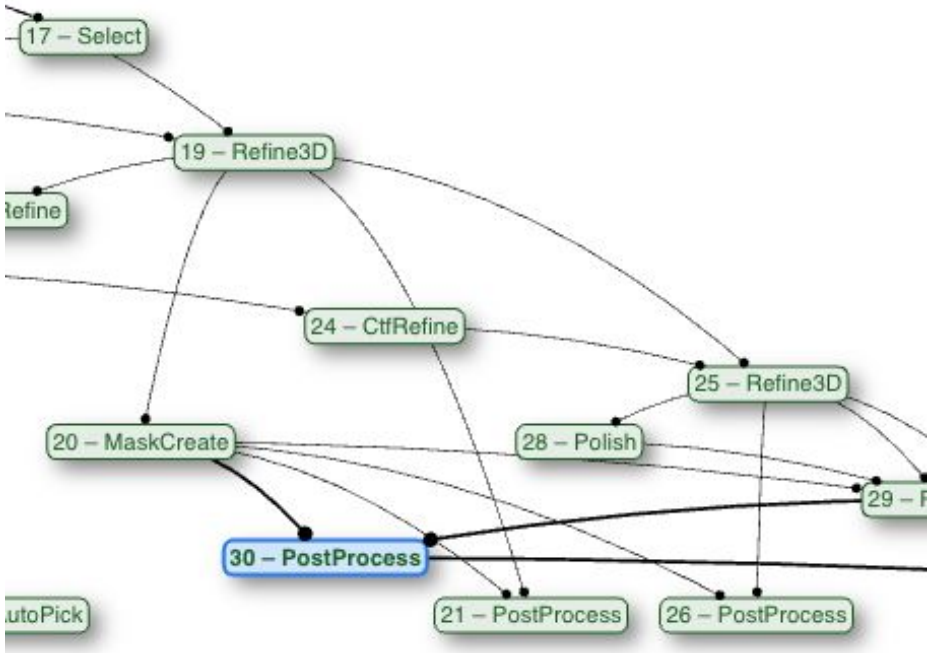
CCP-EM software plan



ccpem-pipeliner

- ***Business logic layer***
- ***Decoupled from tasks and UI***
- Python 3
- Plugin architecture
  - 100-500 lines per app + tests
- Directly Acyclic Graph dataflow
  - Metadata tracking
  - Archiving
  - Jobs have input/output nodes
- CL / UI / scripting APIs
- https://gitlab.com/ccpem/ccpem-pipeliner
- MPL 2.0 license
- Funding: STFC BID / Global Phasing / CCP-EM



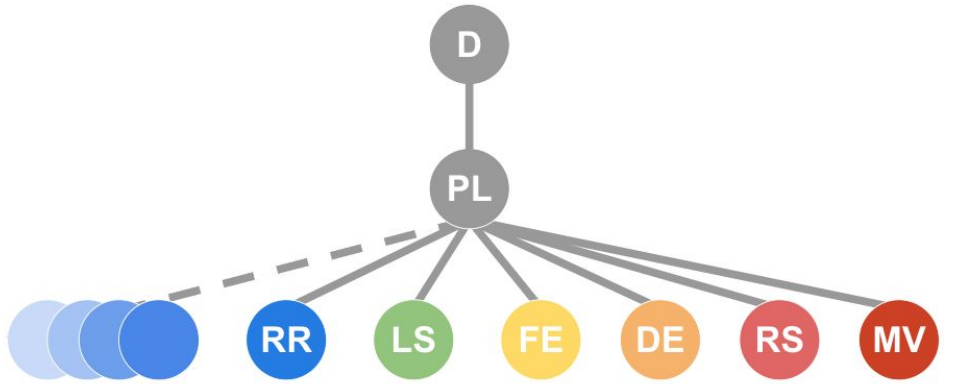


Science and Technology Facilities Council

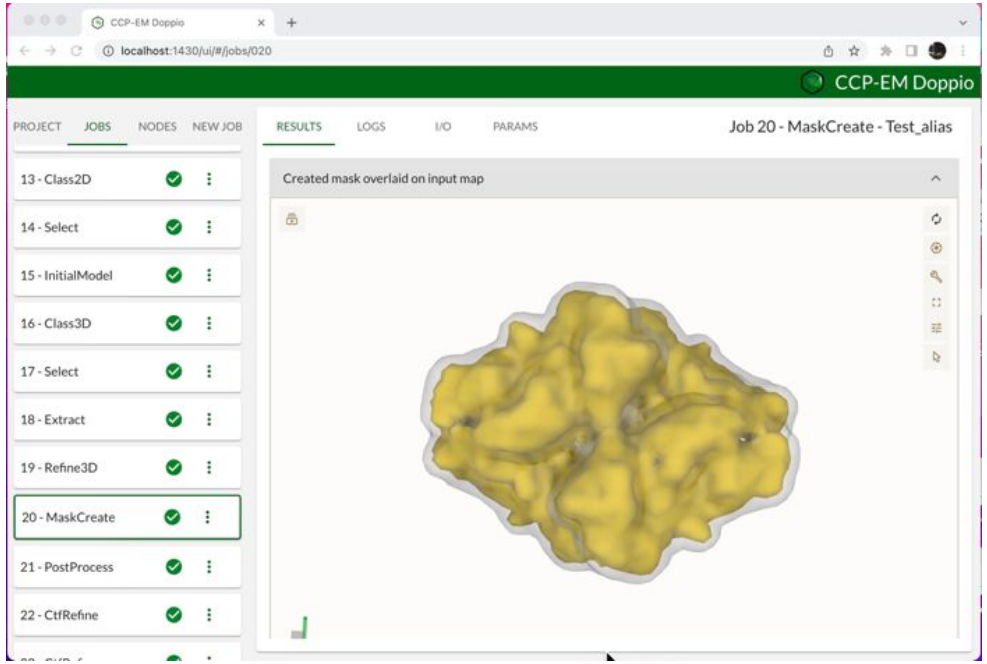


CCP-EM *Doppio*

- ***JavaScript UI***
- React.js / FastAPI / Material UI
- Remote browser app or local Electron app
- Task UIs auto generated from Python plugins



- Full support for ***Image Processing*** through to ***Model Validation***
- https://gitlab.com/ccpem/doppio

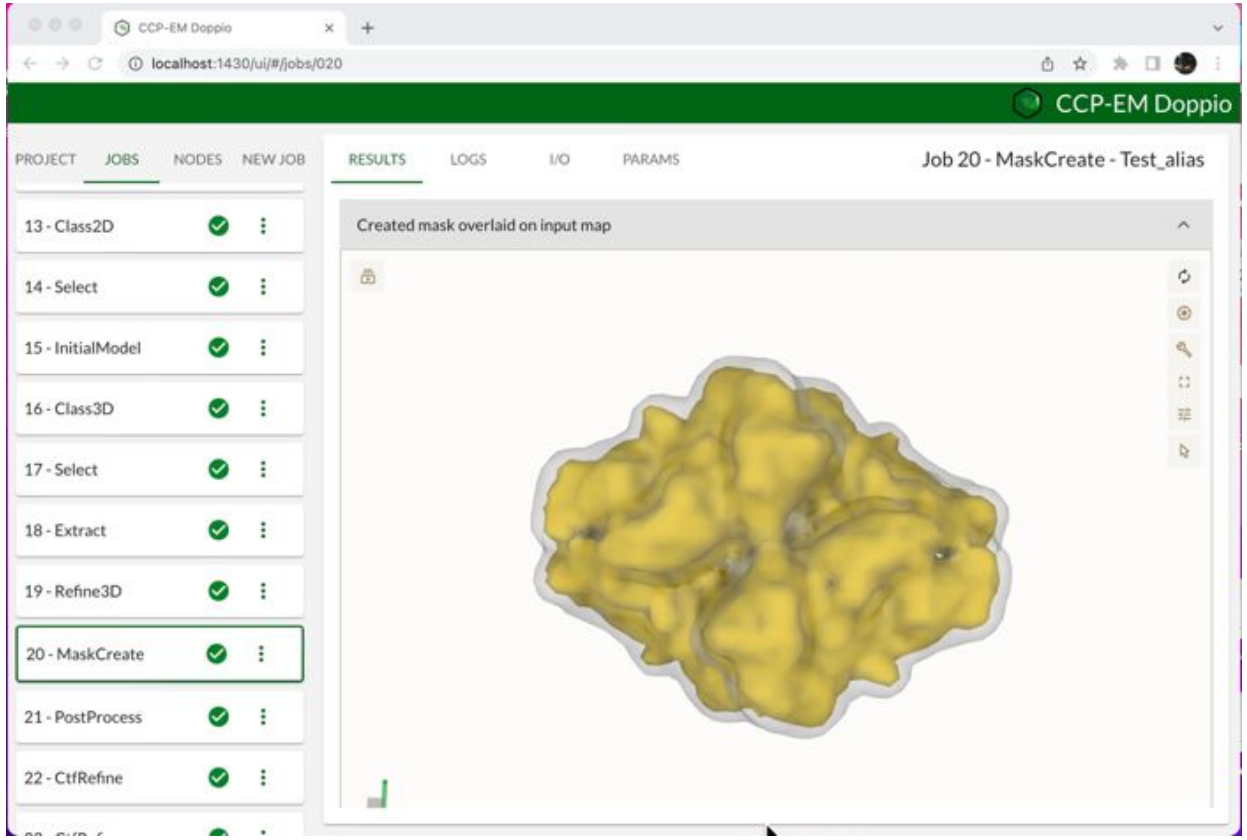


- STFC dual license
- Funding: BEIS / STFC BID / CCP-EM

Science and Technology Facilities Council



Doppio demo



Science and Technology Facilities Council



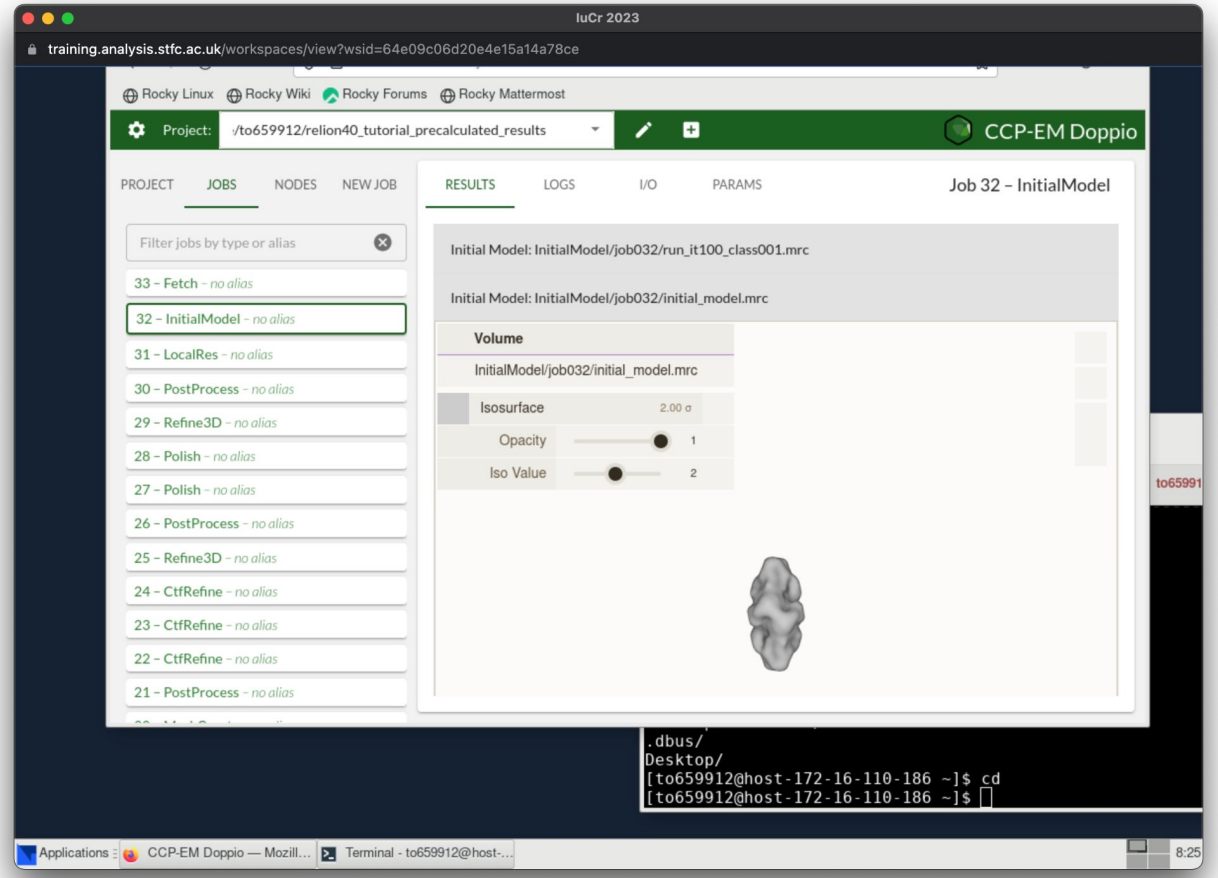
Installing Doppio Beta 3

- ***https://www.ccpem.ac.uk/download.php***
  - Linux & mac only
  - Requires CCP-EM 1.0 nightly
  - Can use local version of Relion 4
  - Full model building support requires CCP4 8.x

- Links to user guide and release notes also found from our downloads page

Running Doppio

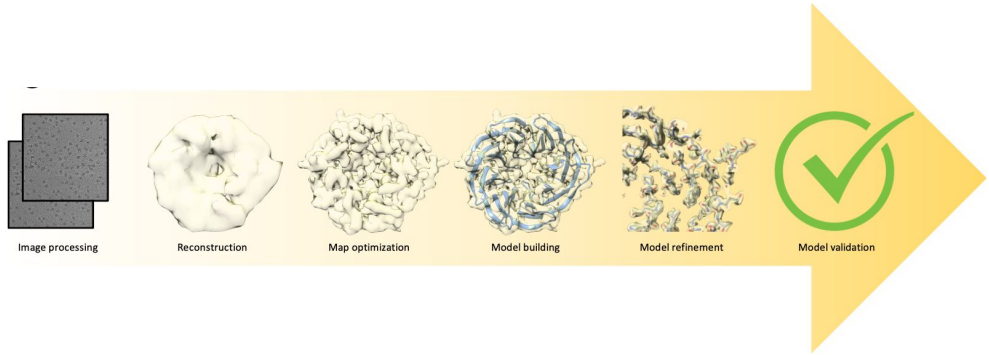
- **IUCr Workshop STFC Cloud Virtual Machines**
  - Rocky Linux
  - 12 core CPU
  - RTX 4000 GPU
  - Preinstalled CCP-EM & Relion, training data
  - *Requires decent internet connection!!!*
  - *Register for use and will be available for duration of conference*

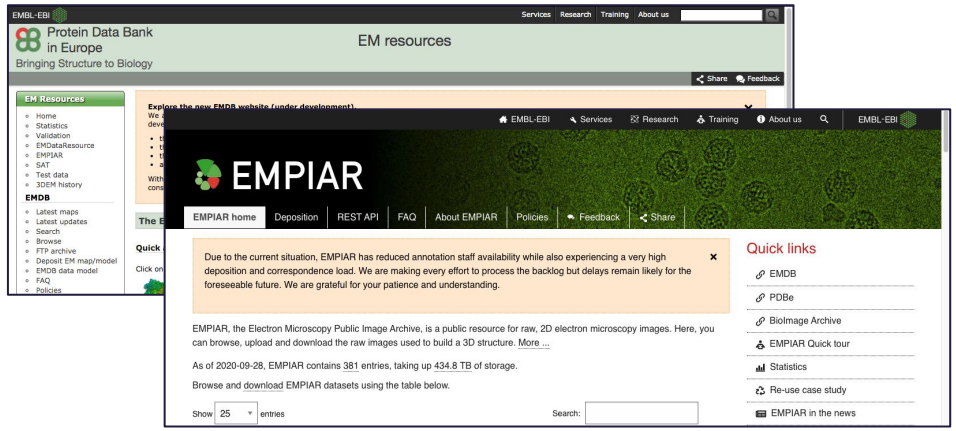


Science and Technology Facilities Council

CCP-EM

An integrated processing pipeline





RELION
Ctffind4
Motioncorr2
GCTF

CCP-EM
Buccaneer
Refmac-Servalcat
MolRep

New Tools

Pre-processing
Reconstruction

→

Model building
Model refinement
Model validation

→

Metadata capture
Data management

3rd Party Plugins

1) **Modern GUI:**
- User-friendly
- Platform independent

Pipeliner

Doppio

2) **Metadata capture and data management**
- Exhaustive metadata capture
- Processing traceability

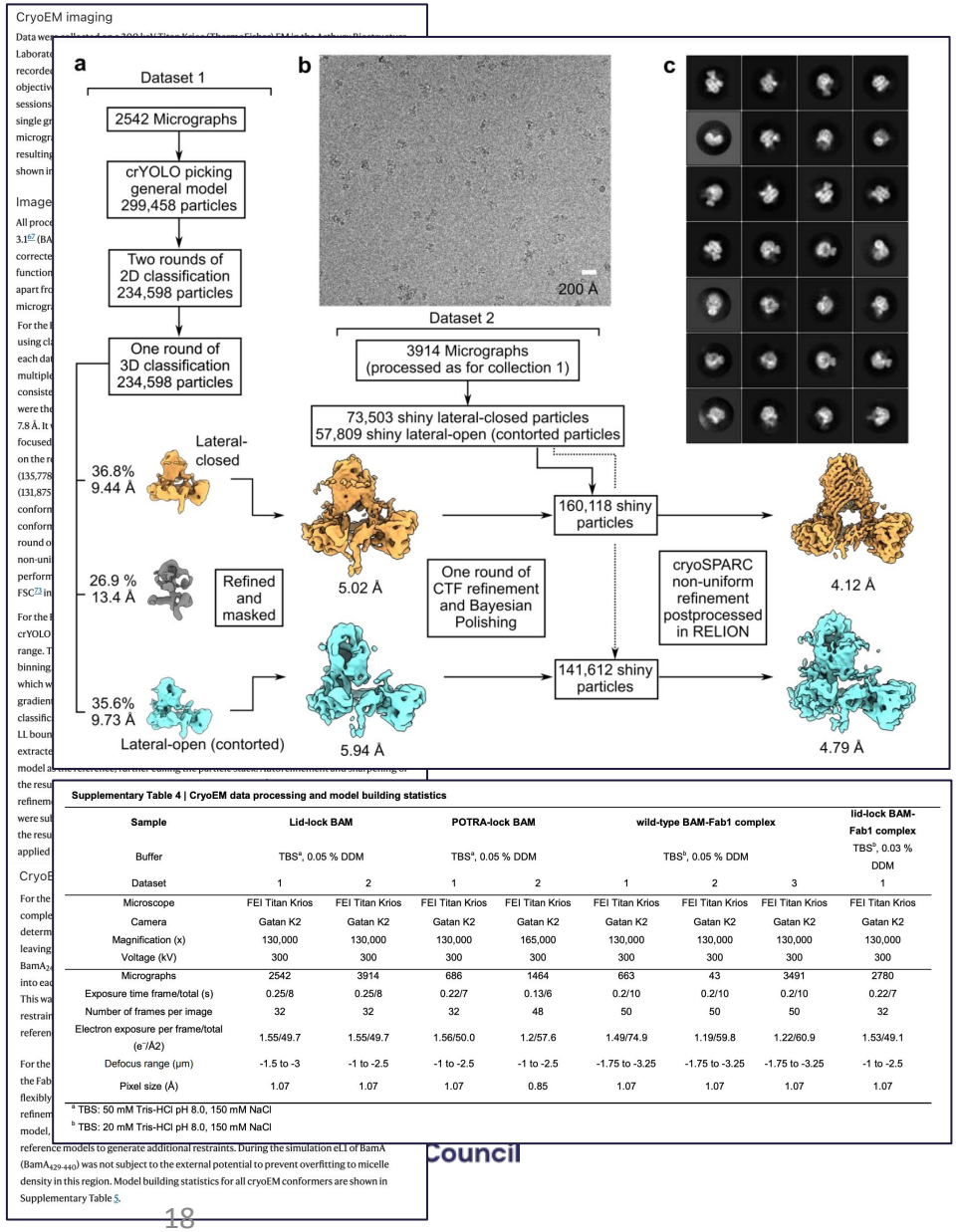
3) **Powerful scripting tools**
- Python and UNIX command line
- Basis for new applications

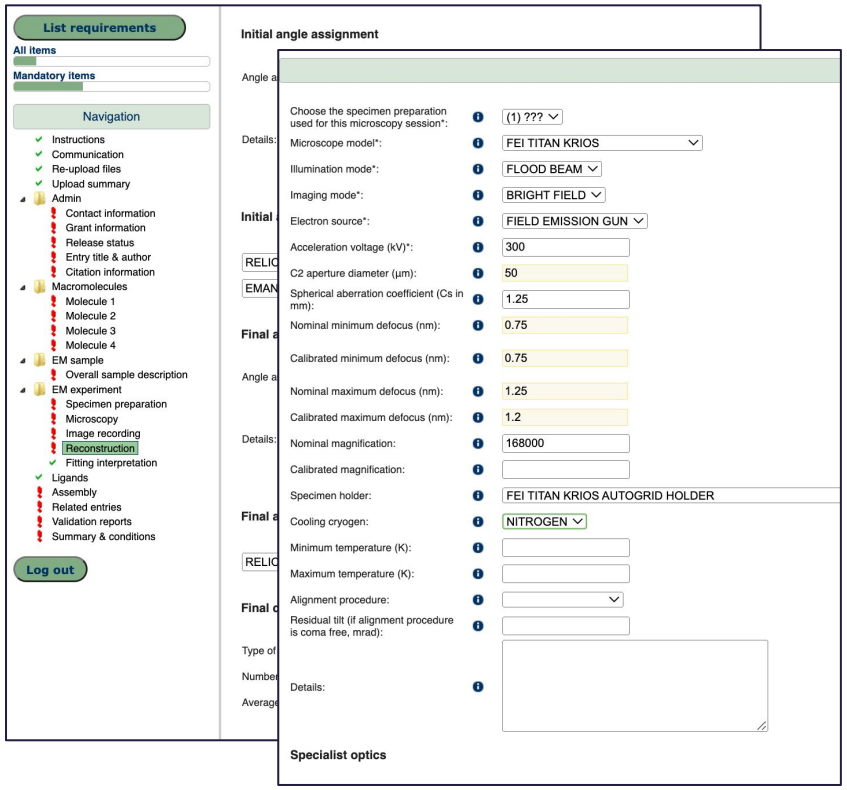
Sjors Scheres

Gerard Bricogne

Science and Technology Facilities Council

2) Managing Metadata





Supplementary Table 4 | CryoEM data processing and model building statistics

Sample	Lid-lock BAM		POTRA-lock BAM		wild-type BAM-Fab1 complex			lid-lock BAM-Fab1 complex
Buffer	TBS[a], 0.05 % DDM		TBS[a], 0.05 % DDM		TBS[b], 0.05 % DDM			TBS[a], 0.03 % DDM
Dataset	1	2	1	2	1	2	3	1
Microscope	FEI Titan Krios	FEI Titan Krios	FEI Titan Krios	FEI Titan Krios	FEI Titan Krios	FEI Titan Krios	FEI Titan Krios	FEI Titan Krios
Camera	Gatan K2	Gatan K2	Gatan K2	Gatan K2	Gatan K2	Gatan K2	Gatan K2	Gatan K2
Magnification (x)	130,000	130,000	130,000	165,000	130,000	130,000	130,000	130,000
Voltage (kV)	300	300	300	300	300	300	300	300
Micrographs	2542	3914	686	1464	663	43	3491	2780
Exposure time frame/total (s)	0.25/8	0.25/8	0.22/7	0.13/6	0.2/10	0.2/10	0.2/10	0.22/7
Number of frames per image	32	32	32	48	50	50	50	32
Electron exposure per frame/total (e^-/Å2)	1.55/49.7	1.55/49.7	1.56/50.0	1.2/57.6	1.49/74.9	1.19/59.8	1.22/60.9	1.53/49.1
Defocus range (µm)	-1.5 to -3	-1 to -2.5	-1 to -2.5	-1 to -2.5	-1.75 to -3.25	-1.75 to -3.25	-1.75 to -3.25	-1 to -2.5
Pixel size (Å)	1.07	1.07	1.07	0.85	1.07	1.07	1.07	1.07

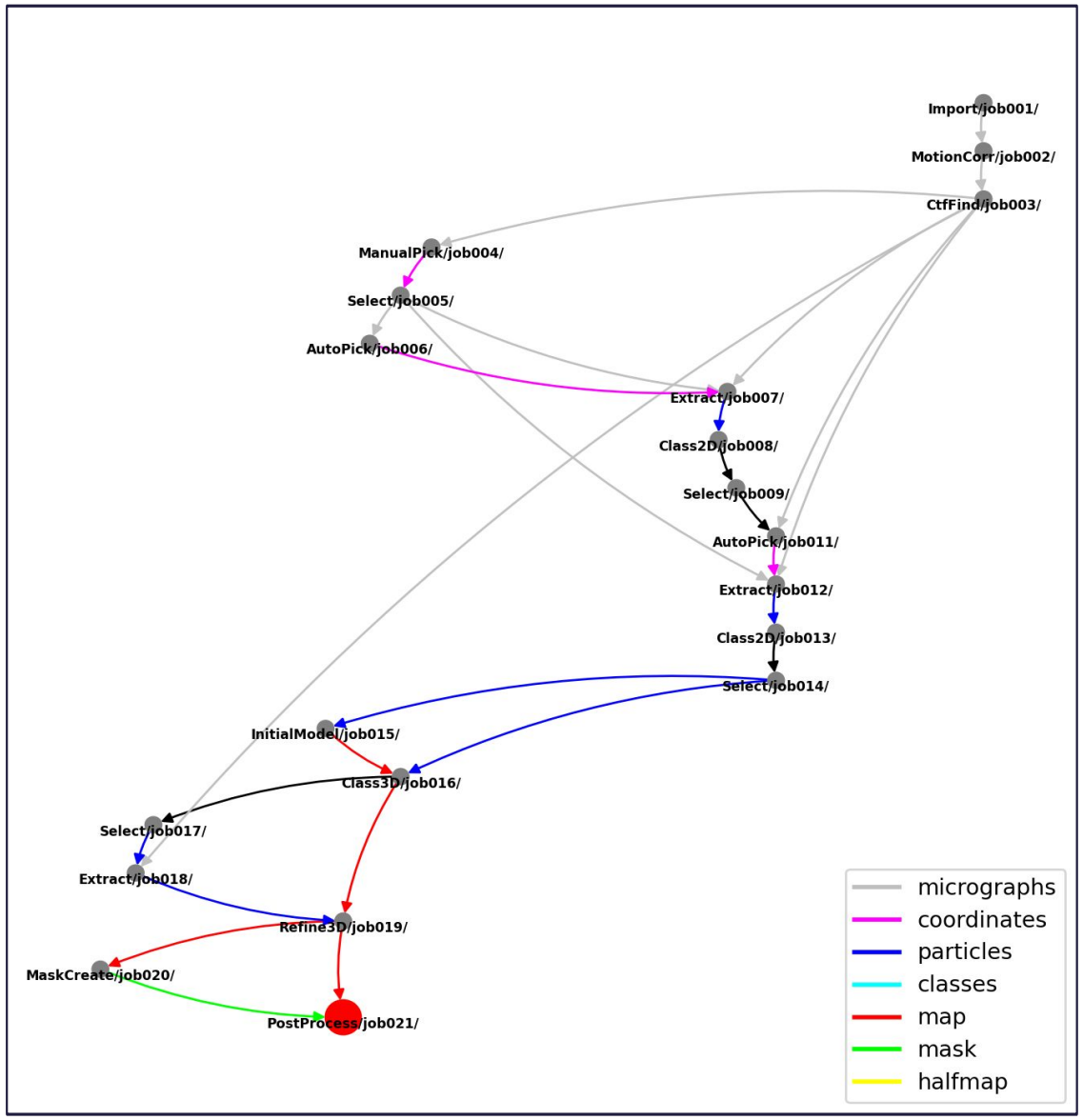
[a] TBS: 50 mM Tris-HCl pH 8.0, 150 mM NaCl
[b] TBS: 20 mM Tris-HCl pH 8.0, 150 mM NaCl

Inaccuracies:
- Transcription errors
- Complicated workflows
- Misunderstandings

Incompleteness:
- Unable to find specific data
- Nomenclature differences
- Space constraints
- ~~Laziness~~ Time constraints

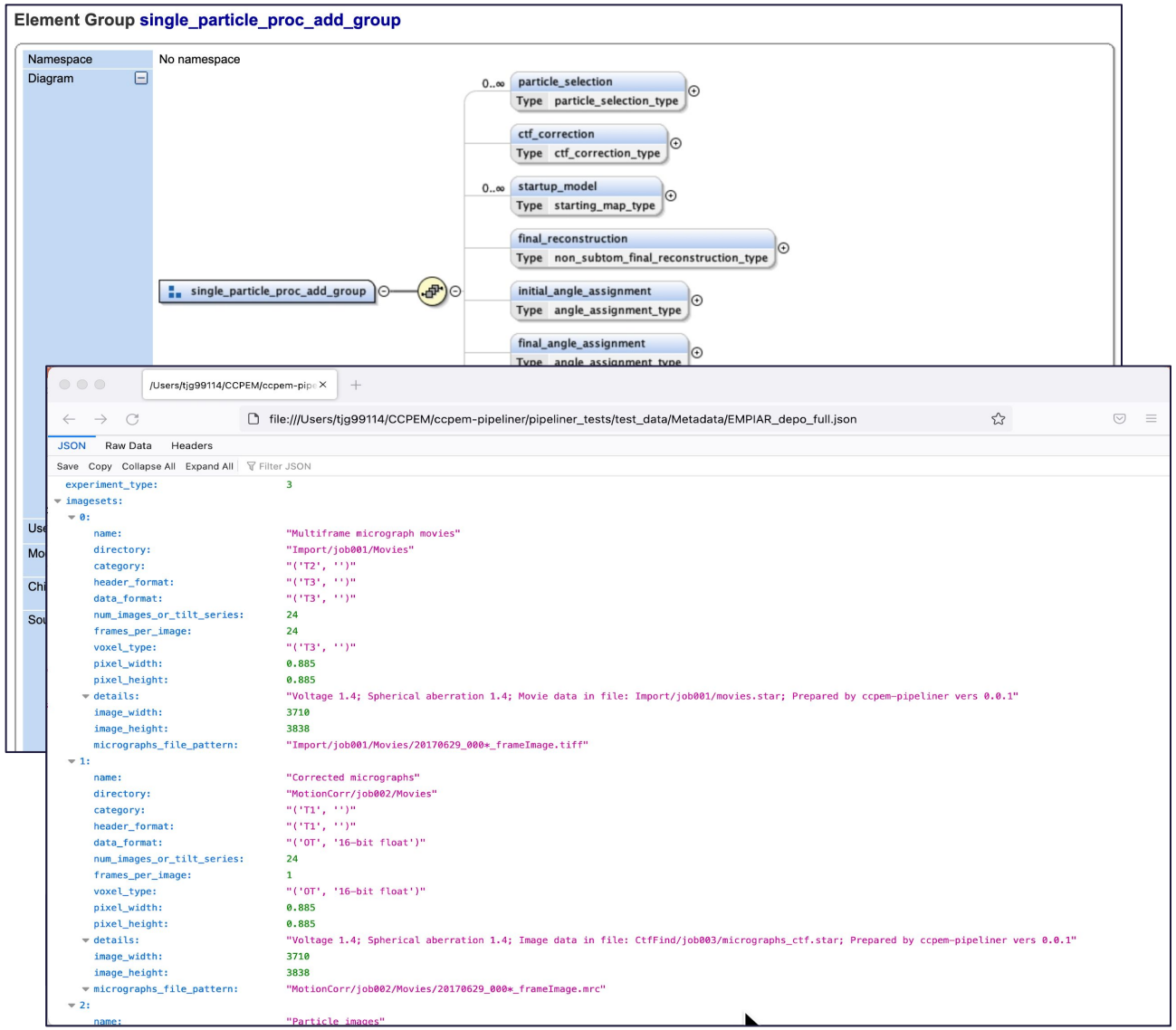
As **machine learning** becomes more prominent in CryoEM the published methods, EMDB, PDB, and EMPIAR will be important sources of training data

Metadata Collection and Analysis



- Every job records the complete history of all operations
- Every job defines the metadata it returns
- Generate a metadata report for an entire projector upstream/downstream from specific jobs

Automated Deposition Preparation



EMPIAR Deposition:

- Launched from any job with Raw movies, Corrected micrographs, Particles or Corrected (polished) particles
- Prepares upload directory
- Creates deposition data file

EMDB Deposition

- Launched from a refined map or model

PDB deposition

- Launched from a refined model

Data management

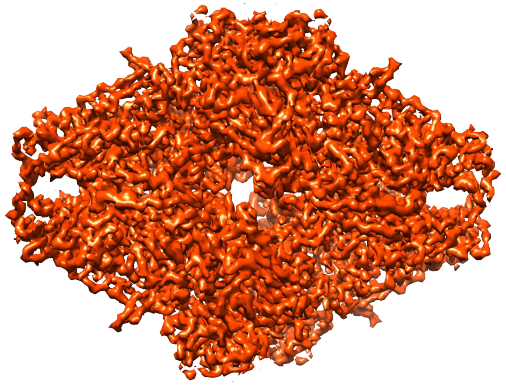
Tools to reduce data clutter and meet data retention needs

```
$> CL_pipeline --cleanup job002
$> CL_pipeline --cleanup job002 -harsh
$> CL_pipeline --cleanup ALL
$> CL_pipeline --cleanup ALL --harsh
```

- Deletes non-essential intermediate processing files
- Each job defines what type of files should be removed
- Files in the network of input/output nodes are protected

```
$> CL_pipeline --full_archive job045

Wrote 20210623_pipeliner_project.tar.gz

$> CL_pipeline --simple_archive job045

Wrote 20210623_pipeliner_project.tar.gz
```

- Full archive contains all the files used to generate the output of the terminal job
- Simple archive contains all the parameter files and a script to automatically re-run the project



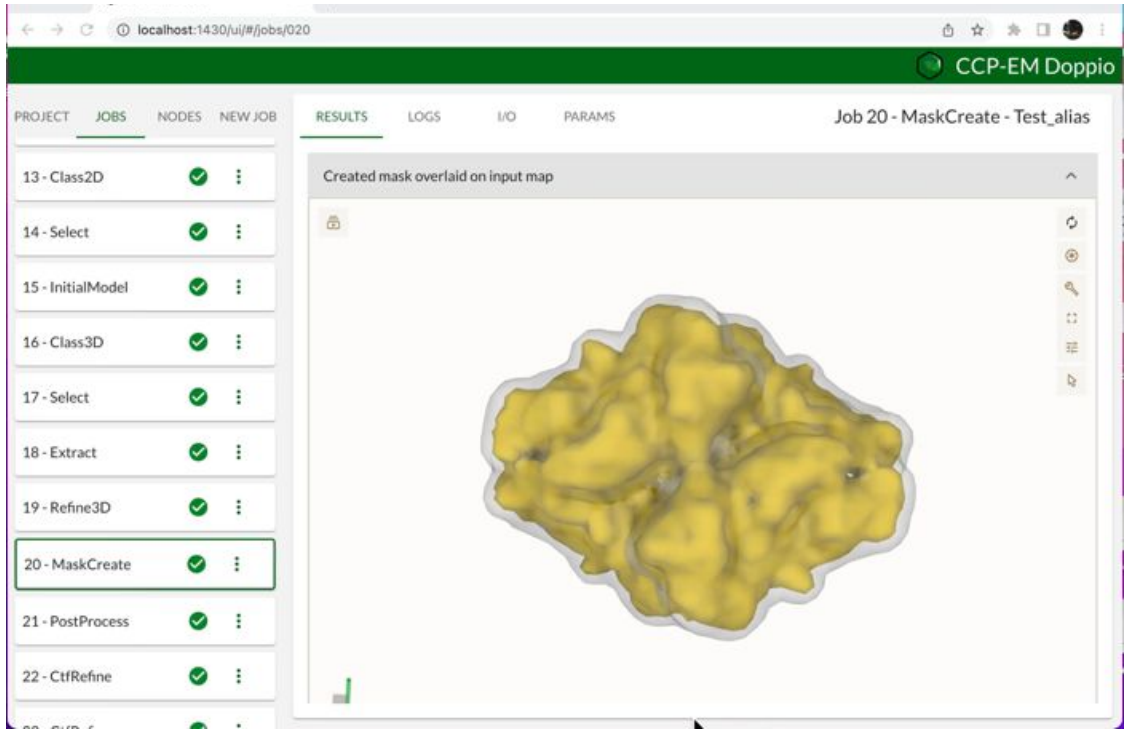
β-gal
2x binned
2179 particles
3.6 Å (3.54 Å Nyquist)

Required 20 jobs – 5.2 G of data produced

	Files	Size
Project	5487	5.2 G
Gentle clean	1482	2.9 G
Harsh clean	820	2.7 G

New Software Features Coming Soon

- Full implementation of model refinement and validation tools
- Tomography and STA support
- EMDB and PDB deposition tools
- Additional GUI functionality
- Alpha released to collaborators now
- Public beta release at Symposium



Participate!

- Source code: gitlab.com/ccpem/ccpem-pipeliner
- Documentation: https://ccpem-pipeliner.readthedocs.io
- Guide to writing plugins is available

3) Scripting the Pipeliner

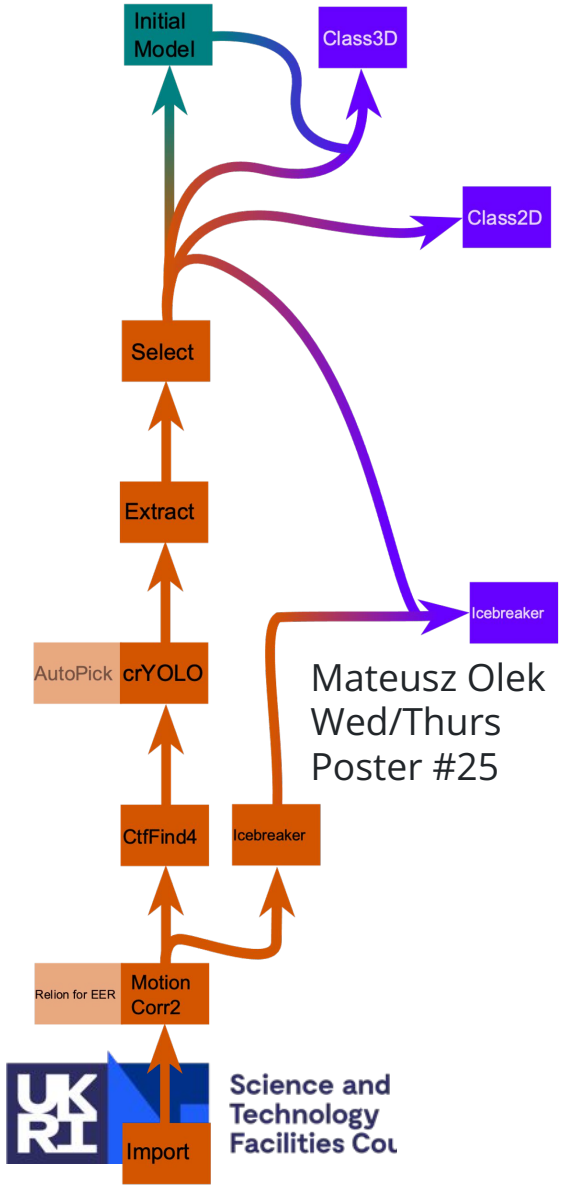
```
#!/user/bin/bash
CL_pipeline --start_new_project
CL_pipeline --schedule_job Import_job.star
CL_pipeline --schedule_job MotionCorr_job.star
CL_pipeline --schedule_job CtfFind_job.star

CL_pipeline --run_schedule --name Schedule1 --jobs Import/job001/ job002 CtfFind/new_alias --min_between 15
    --nr_repeats 3 --min_wait_before 2 -sec_wait_after 15

CL_pipeline --metadata_report job003
```

```
from pipeliner.api.manage_project import PipelinerProject

my_project = PipelinerProject()

my_project.schedule_job("Import_job.star")          # adds Import/job001/ to the pipeline
my_project.schedule_job("MotionCorr_job.star")          # adds MotionCorr/job002/ to the pipeline
my_project.schedule_job("CtfFind_job.star")         # adds AutoPick/job003/ to the pipeline

my_project.run_schedule(
     name="Schedule1",
     job_ids=["Import/job001/", "MotionCorr/job002/", "CtfFind/job003/"],
     nr_repeat=3,
     minutes_wait=15,
     minutes_wait_before=2,
     seconds_wait_after=15,
     )
my_project.get_network_metadata("CtfFind/job003/", "my_metadata.json")      # returns metadata for listed job
                                    # and all upstream jobs
```

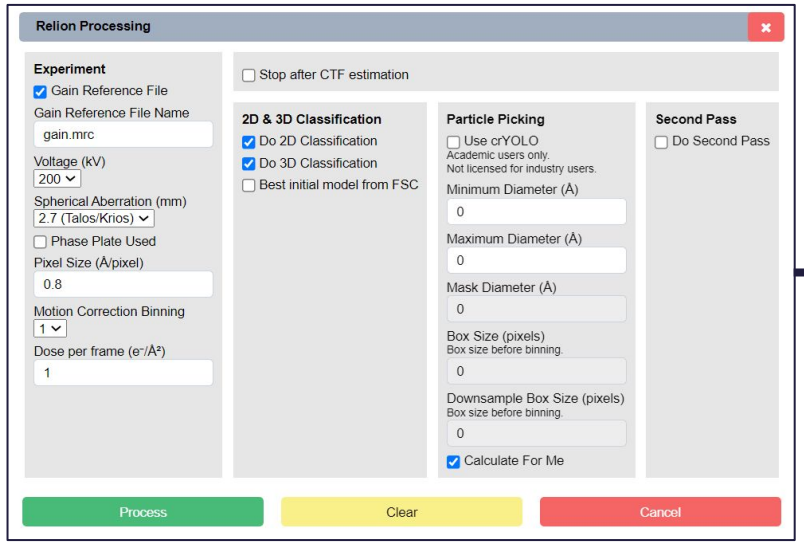
Scripting Example: RelionIt

on-the-fly batched preprocessing and evaluation

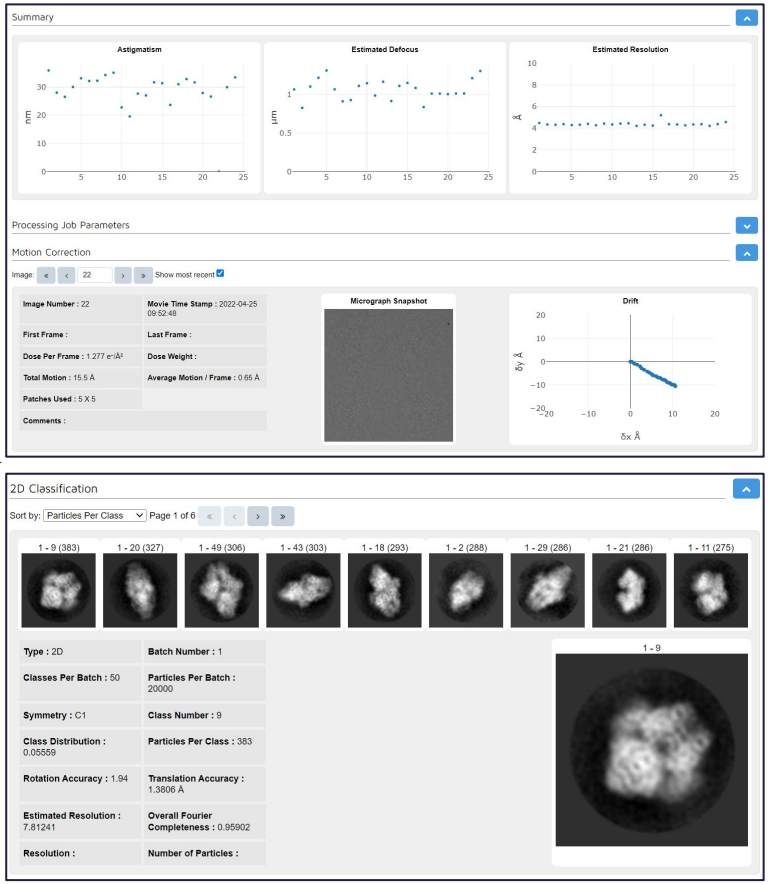
(Dan Hatton, Diamond Light source)



Mateusz Olek
Wed/Thurs
Poster #25



ISPyB



CCP-EM Doppio Beta Release

- ***Beta 1 - April 25 Spring Symposium:***
  - Requires CCP-EM 1.x, CCP4 8.x
  - Linux only
- ***Beta 2 - July:***
  - Asynchronous job runner
  - Relion image display compatibility
- ***Beta 3 - August:***
  - Improved support for Refmac
  - Metadata & reference reports
  - Basic archiving functionality
- ***Full Release:***
  - Conda based distribution
- ***Future plans:***
  - Moorhen integration
  - Relion 5 integration
  - Full metadata deposition (EMPIAR/EMDB/PDB)
  - JS 2D image viewer
  - Tomography / Relion 5.1 support
- https://www.ccpem.ac.uk/download.php

UKRI Science and Technology Facilities Council

CCP-EM

Acknowledgements

CCP-EM core team

CCP4 core team

STFC Business & Innovation Development

CCP-EM Commercial License Holders

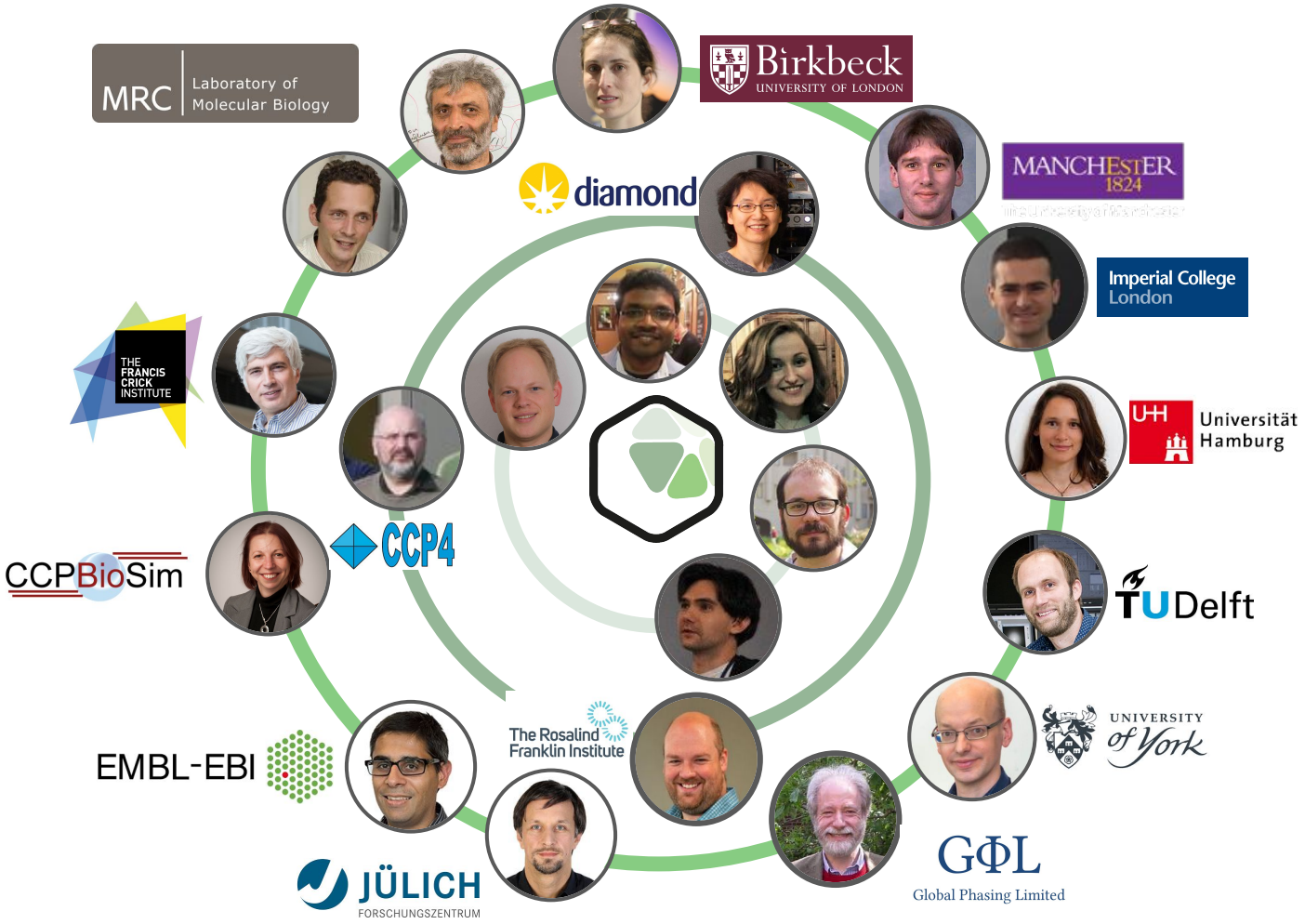
CCP-EM Collaborators

Website: www.ccpem.ac.uk

Mailing list: www.jiscmail.ac.uk/ccpem

Twitter: @ccp_em

Email: ccpem@stfc.ac.uk



Science and Technology Facilities Council

Medical Research Council

wellcome

Department for Business, Energy & Industrial Strategy

Research Complex at Harwell

GΦL Global Phasing Limited

CCP-EM

Science and Technology Facilities Council

CCP-EM 1.0

CCP-EM software suite

30 tasks in a common Python framework

Uses some CCP4 programs

Download from ccpem.ac.uk

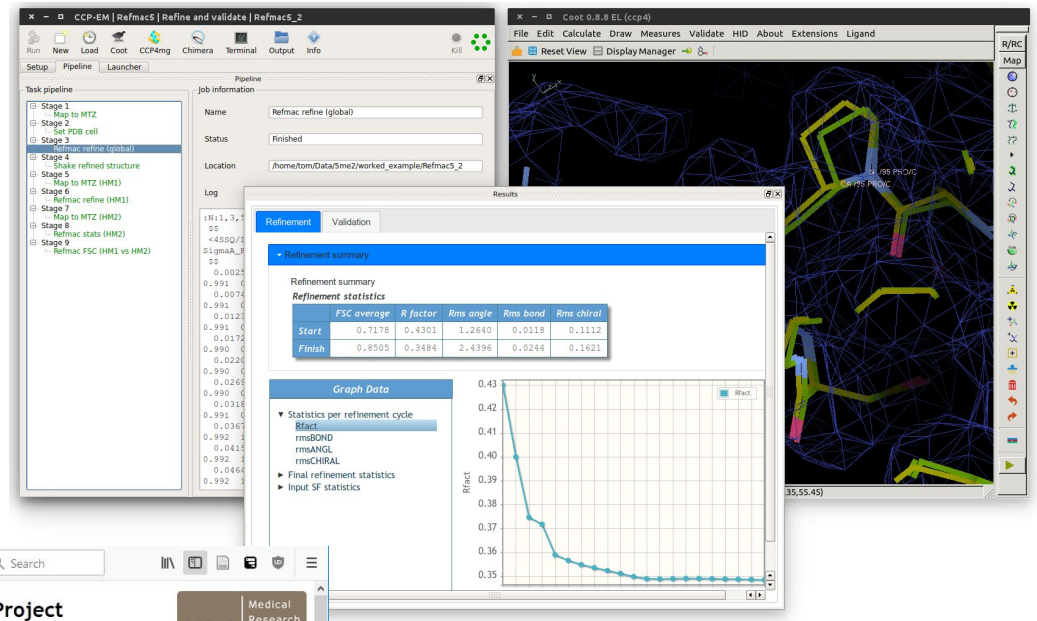
Linux & Mac

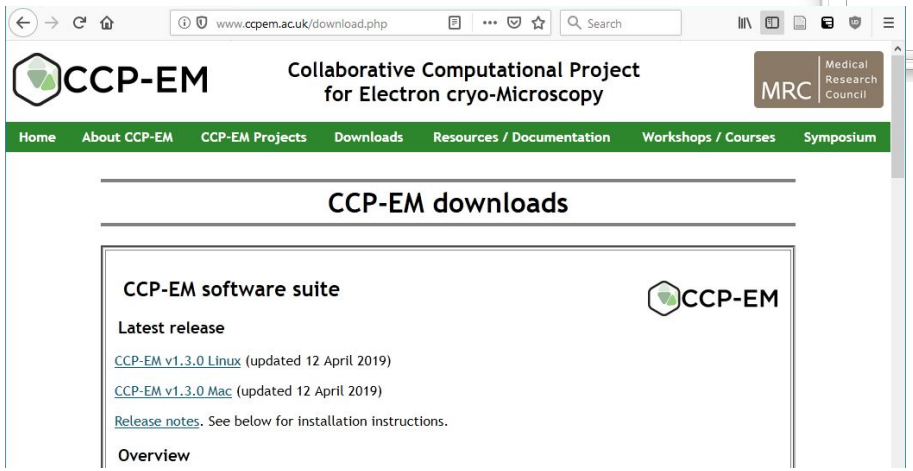
Free for academic use, fee for commercial

Bugs & requests:

ccpem@stfc.ac.uk

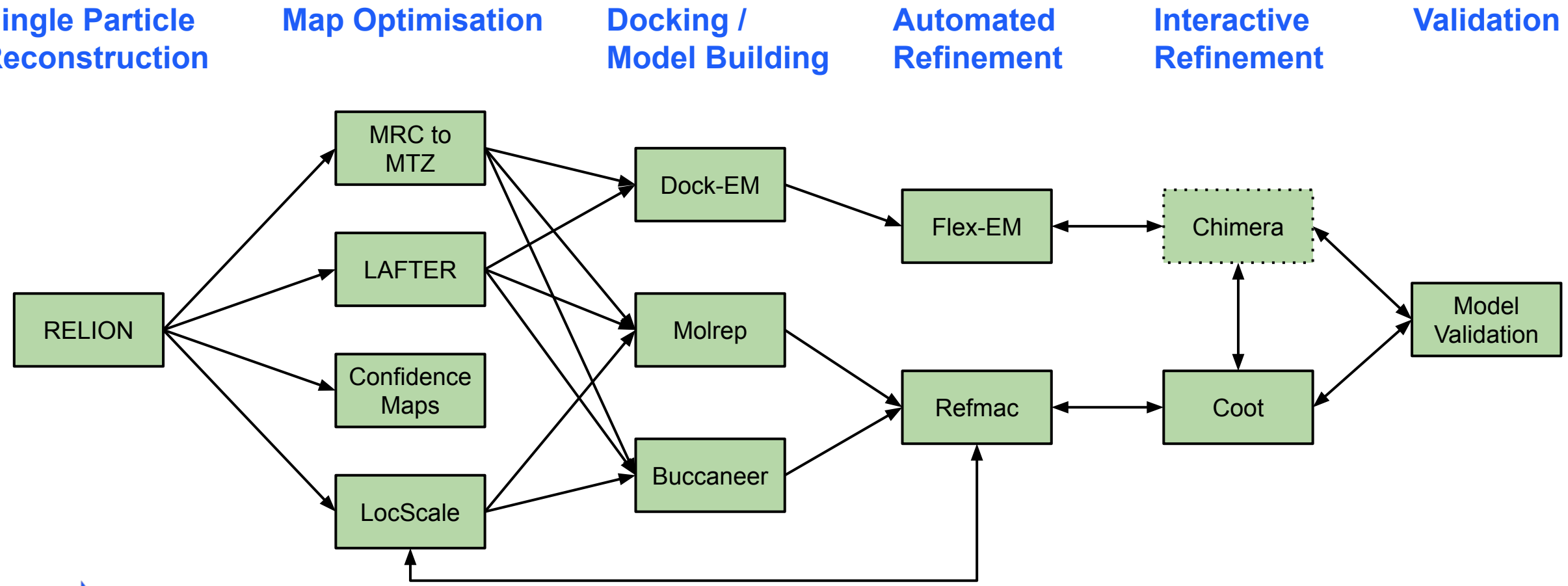
Recommend using latest nightly





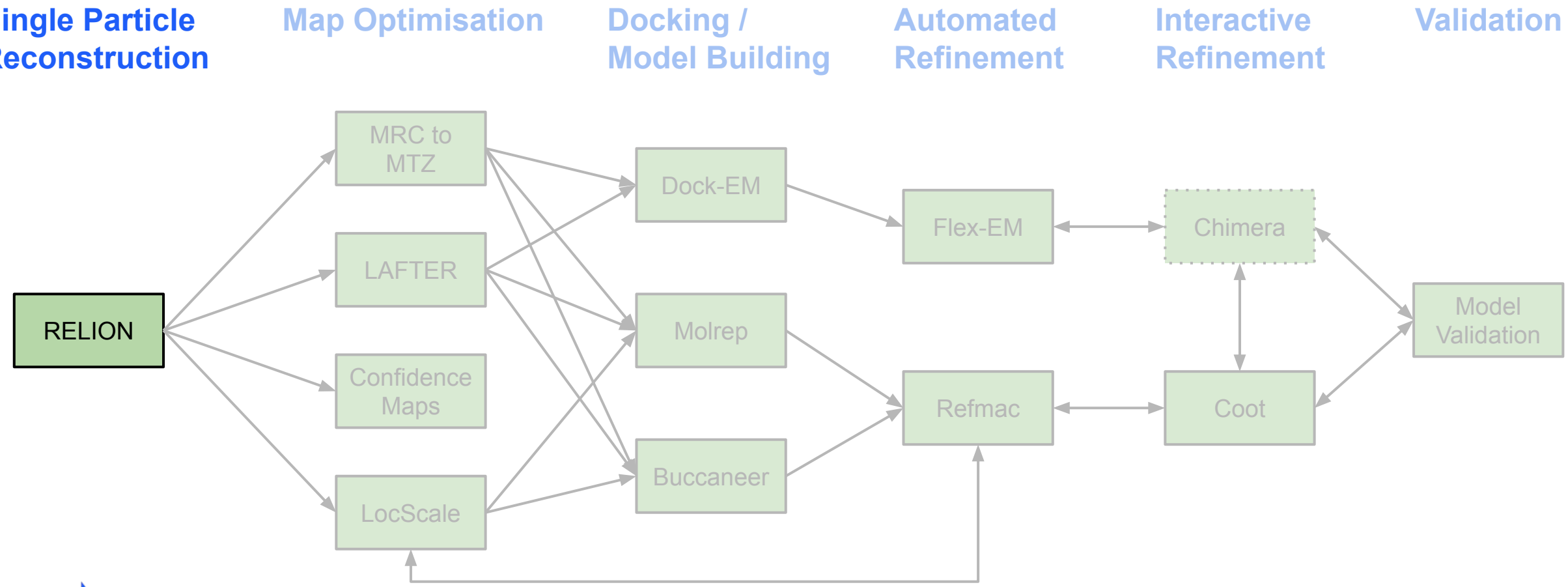
UKRI Science and Technology Facilities Council

CCP-EM workflow



Science and Technology Facilities Council

CCP-EM workflow



Science and Technology Facilities Council

RELION

Software suite for Single Particle Reconstruction and Subtomogram Averaging

CCP-EM includes pre-compiled RELION binaries for Linux and Mac

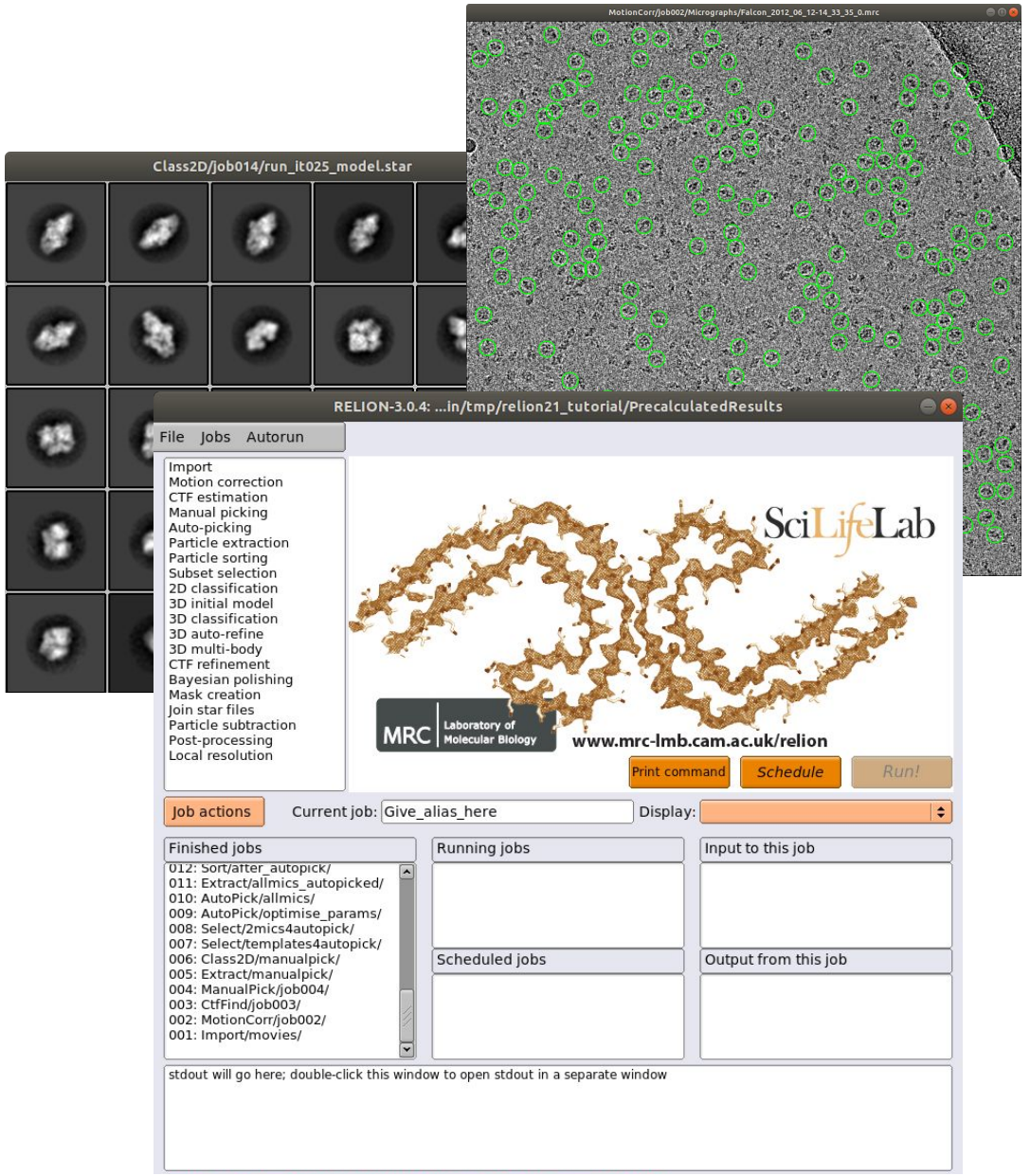
CUDA GPU support on Linux

v3.1 in CCP-EM 1.6

RELION will be integrated more closely with the rest of the suite in future

Sjors Scheres

Science and Technology Facilities Council

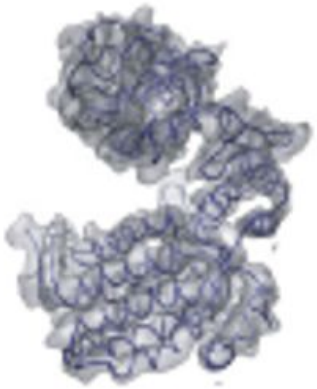


Science and Technology Facilities Council

Atomic modelling tools

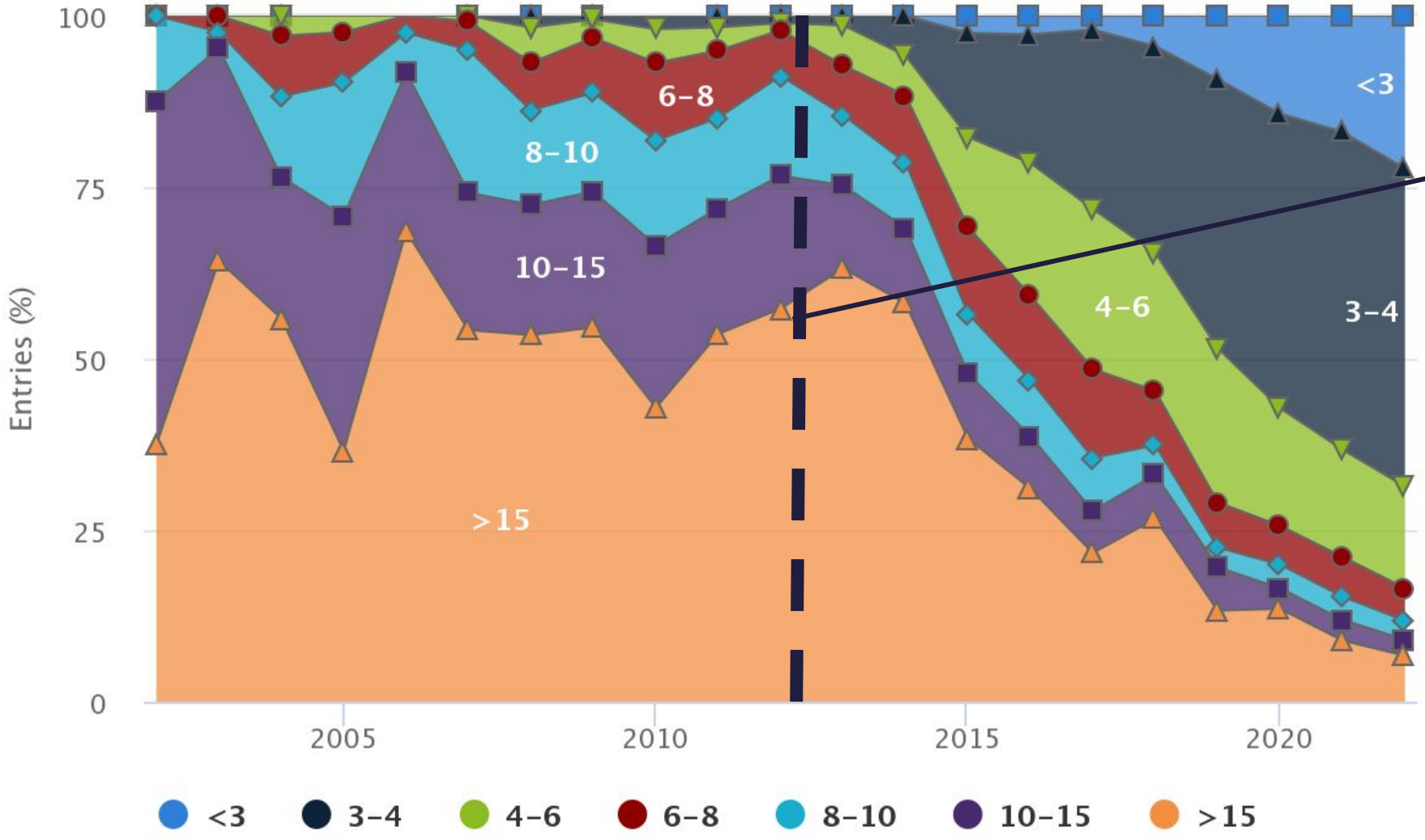
Resolution of interpretable features






Up to 20 Å	Up to 9 Å	Up to 6 Å	Up to 4 Å
Conformational changes	Conformational changes	Conformational changes	Conformational changes
Domain boundaries	Domain boundaries	Domain boundaries	Domain boundaries
		Beta sheets	Beta sheets
			Individual beta strands
	Alpha helices	Alpha helices	Alpha helices
			Pitch of alpha helices
		Pitch of RNA helices	Pitch of RNA helices
			Phosphate “bumps”
			Side chains

EMDB resolution per year



Direct electron detectors

Global resolution

2022 (Sep):

22% <3Å

70% <4Å

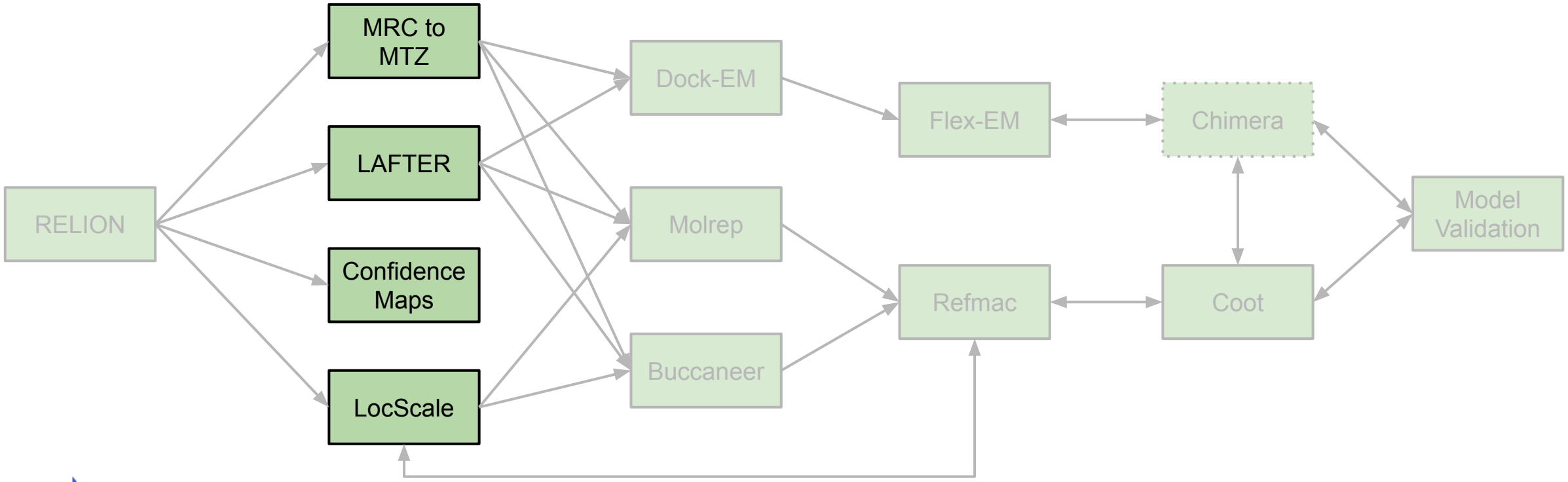
30% >4Å

Science and Technology Facilities Council

Map optimisation

CCP-EM workflow

Single Particle Reconstruction | **Map Optimisation** | Docking / Model Building | Automated Refinement | Interactive Refinement | Validation

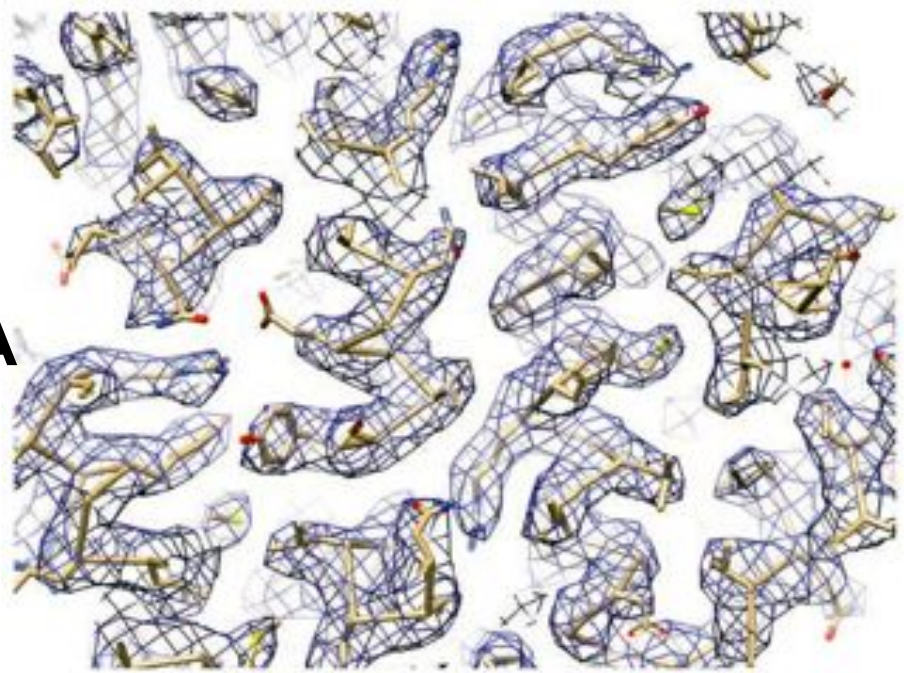


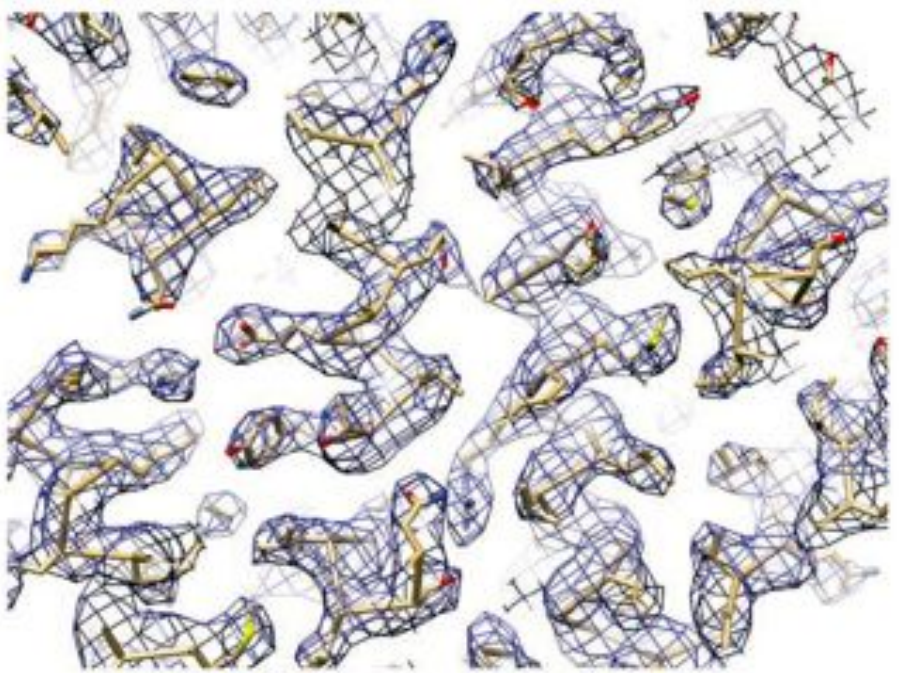
Science and Technology Facilities Council

EM and MX maps are different

Maps from cryo-EM look like maps from X-ray crystallography...

EM 3.3Å





MX 3.4Å

20S proteasome

Science and Technology Facilities Council

Li et al, Nature Methods, 2013

EM and MX maps are different

Maps from cryo-EM look like maps from X-ray crystallography...

...But some important differences:

- Map is phased!
  - Better quality initially
  - Doesn’t change during model building
- Densities represent ***Coulomb potential***, not ***electron density***
- Box size is arbitrary
- Local resolution can be highly variable
- Map sharpening needs to be optimised

Crystallography software can often handle cryo-EM maps, but needs care

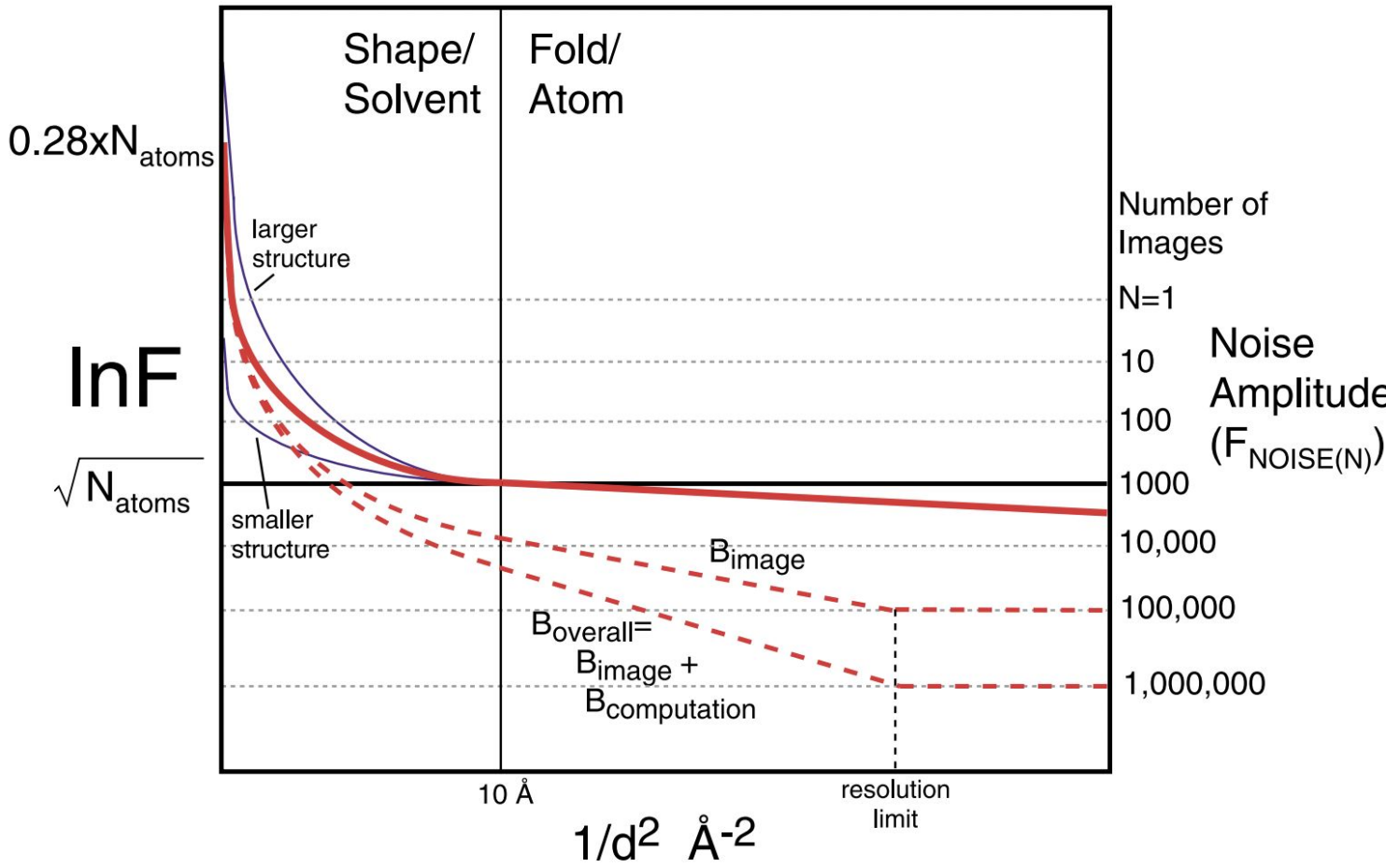
Science and Technology Facilities Council

Contrast loss in cryoEM

Combination of several effects leads to a Gaussian blurring (contrast loss):

- Alignment errors
- classification errors
- interpolation errors…
- … all other imperfections

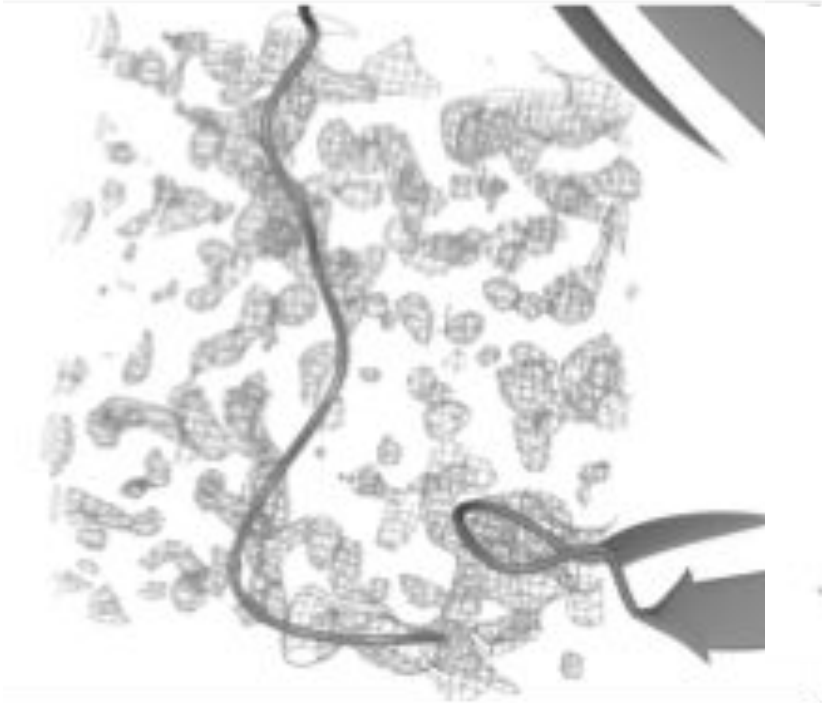
Unlike MX intensity fall off cannot be directly measured from the experiment



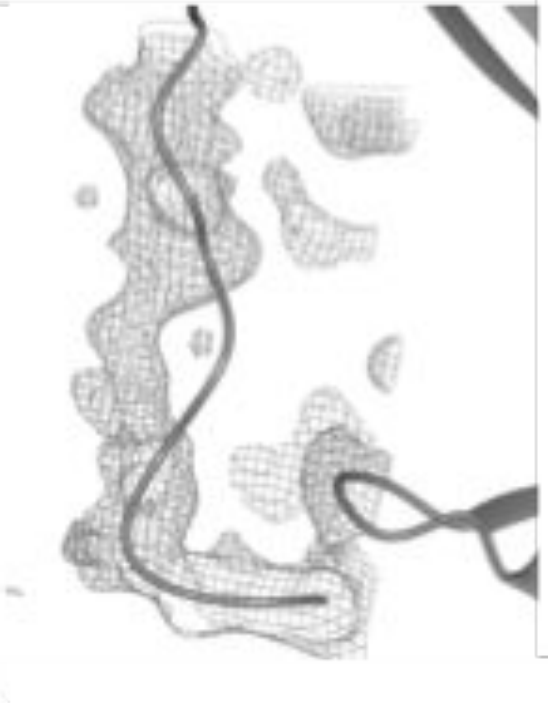
Guinier plot (Guinier & Fournet, 1955) from Rosenthal & Henderson 2003

Science and Technology Facilities Council

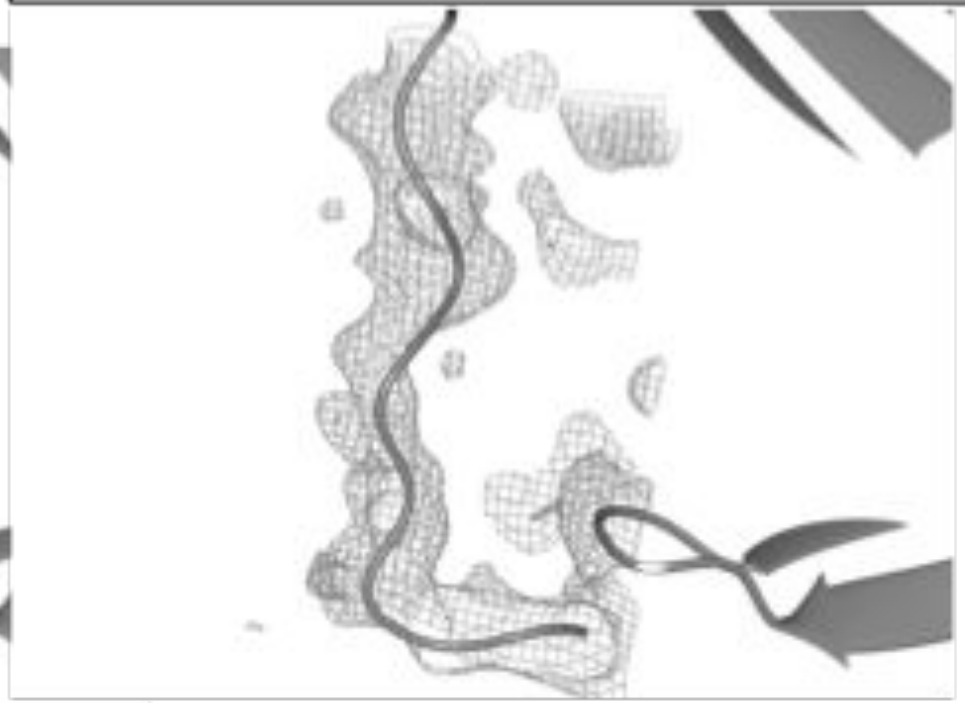
Map from PDB



Blurred: B = 60



Model refined against blurred map



Science and Technology Facilities Council

Garib Murshudov

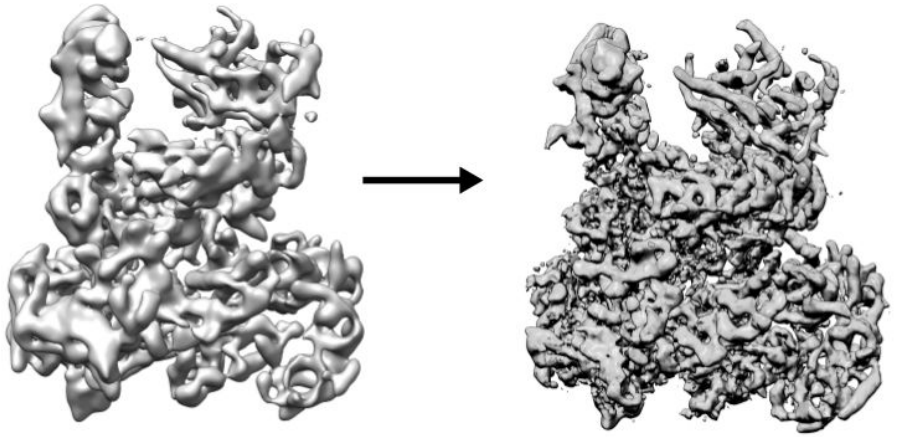
LocScale

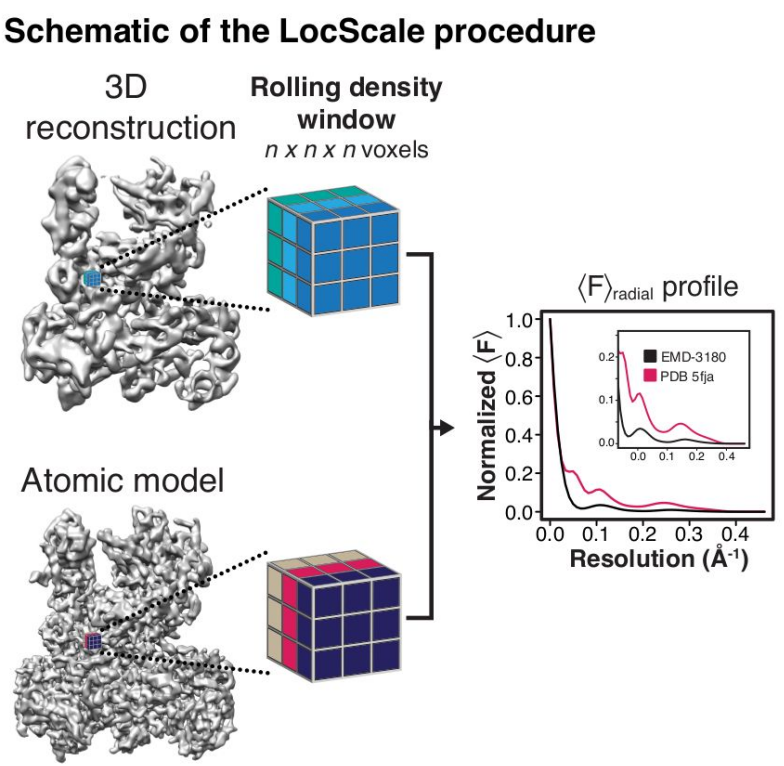
Locally adaptive map sharpening

Fits experimental map to local amplitude profile from atomic model B-factors

Requires a refined model (for now!)

Iterative process of model building and map improvement





Arjen Jakobi

Science and Technology Facilities Council

Jakobi et al. eLife (2017) 6:e27131

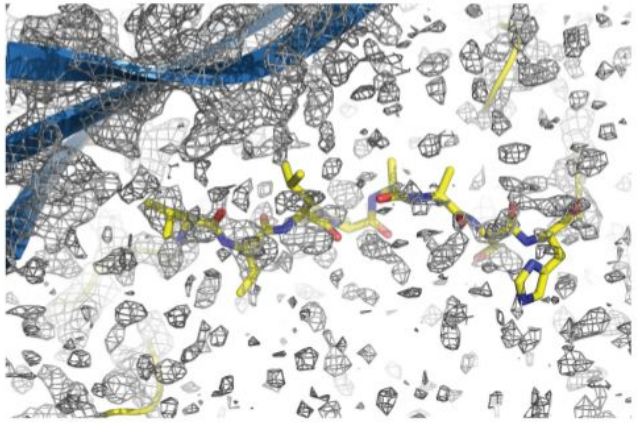
LocScale

New version “LocScale2” available now which has model-free mode https://gitlab.tudelft.nl/aj-lab/locscale

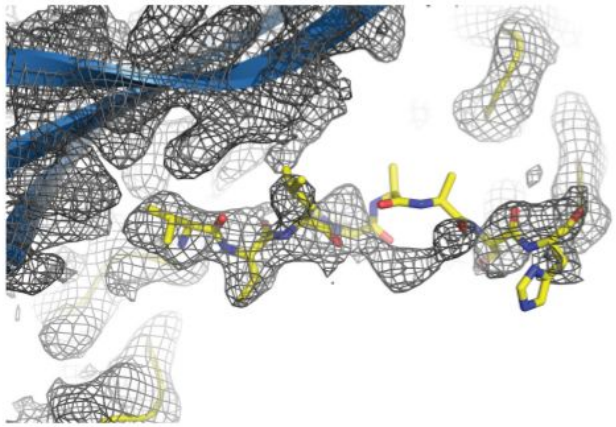
Will be added to CCP-EM in next major release

Other programs e.g. phenix.auto_sharpen, LocalDeblur, DeepEMhancer… etc

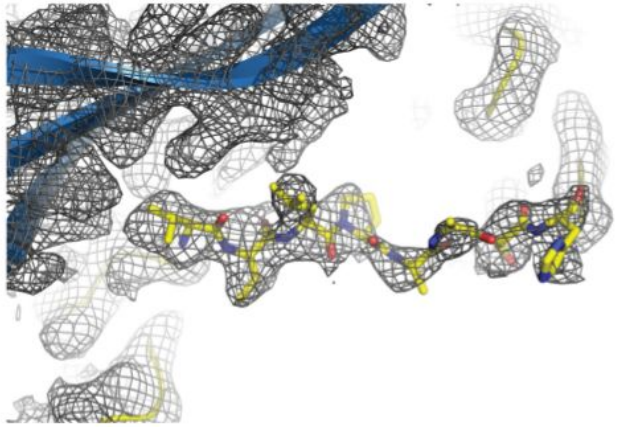
Use LocScale maps for visualisation
Do not use LocScale maps for refinement!



EMD-2984 / PDB 5a1a



LocScale / PDB 5a1a



LocScale / rebuilt

Arjen Jakobi

Science and Technology Facilities Council

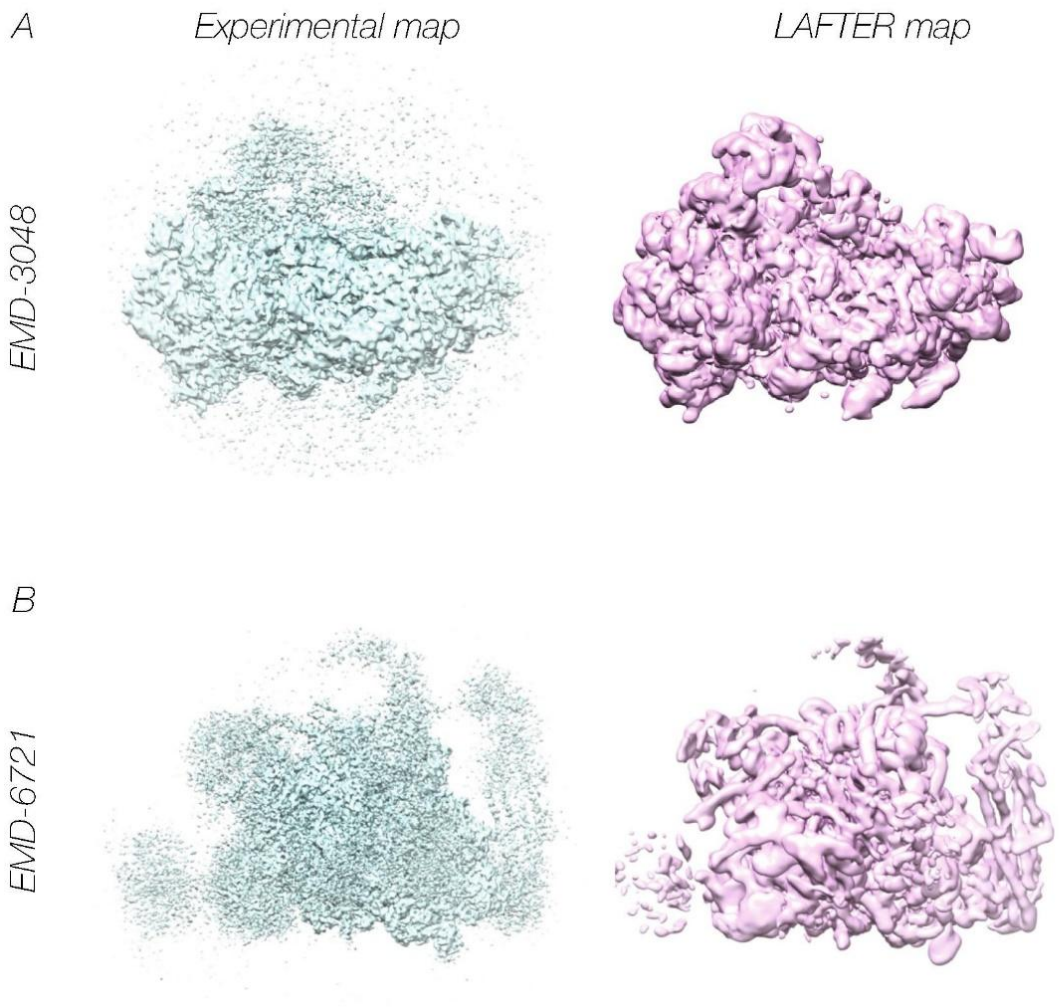
Jakobi et al. eLife (2017) 6:e27131

LAFTER

Local **A**greement **F**ilter for **T**ransmission
EM **R**econstructions

Compares band-passed half maps to identify locally-shared features

Preserves shared signal, suppresses noise



Chris Aylett

Science and Technology Facilities Council

Ramlaul, Palmer & Aylett (2019) J. Struct. Biol ***205****:30–40*

LAFTER

Local **A**greement **F**ilter for **T**ransmission **E**M **R**econstructions

Compares band-passed half maps to identify locally-shared features

Preserves shared signal, suppresses noise

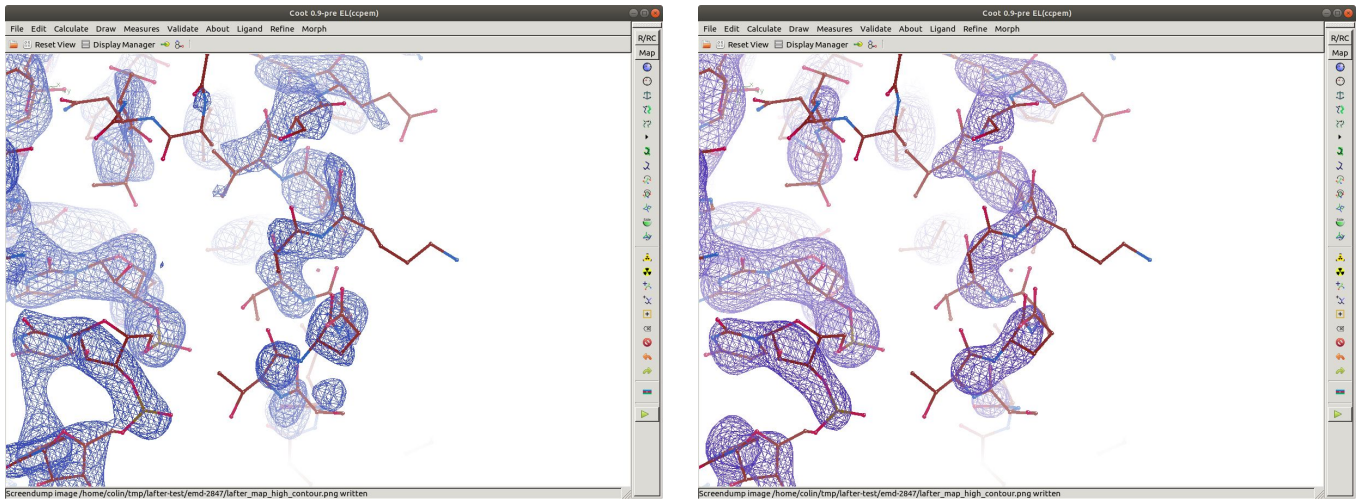
High contour: strong features remain similar

Low contour: weak noise features are removed

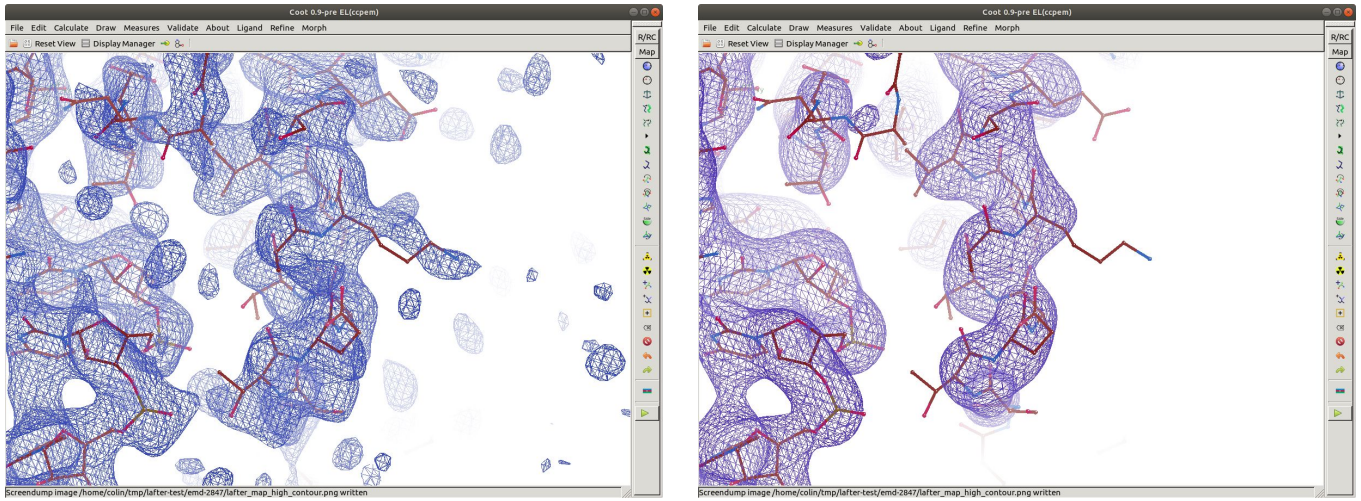
Chris Aylett

Science and Technology Facilities Council

Original EMD-2847 | LAFTER filtered



High contour



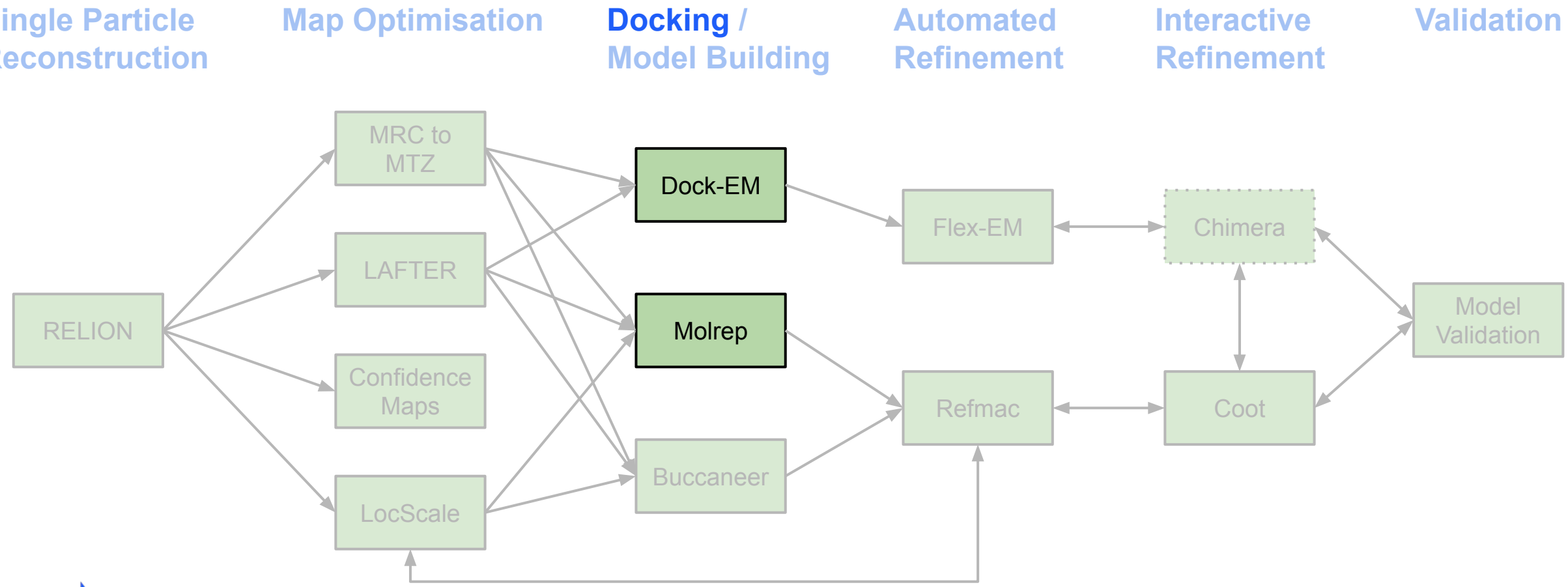
Low contour

Ramlaul, Palmer & Aylett (2019) J. Struct. Biol ***205****:30–40*

Science and Technology Facilities Council

Docking and model building

CCP-EM workflow



Science and Technology Facilities Council

Rigid-body fitting

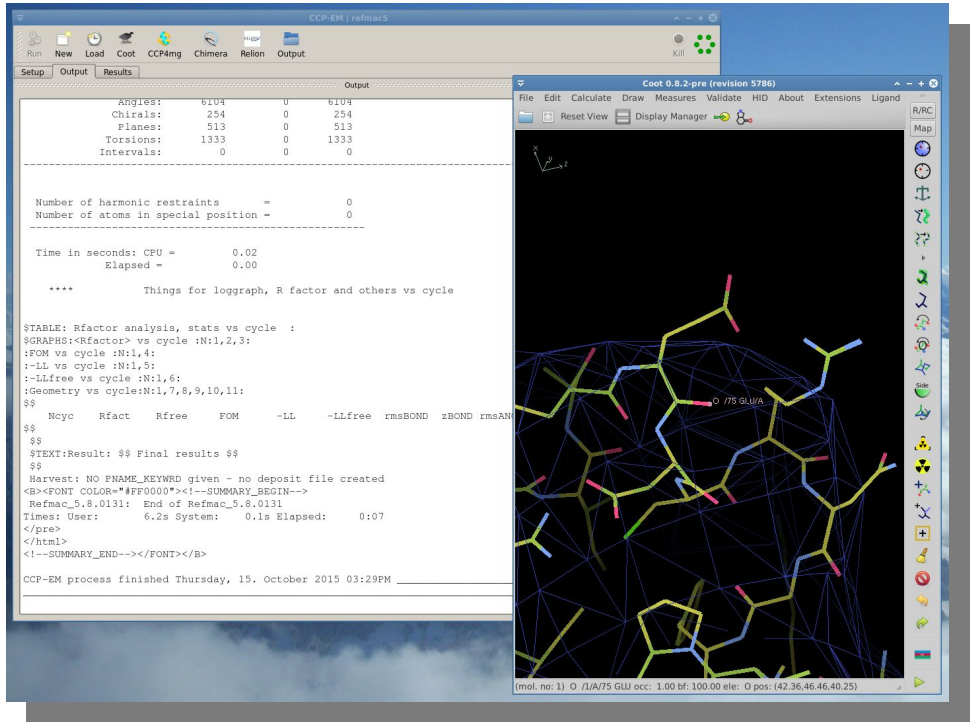
MOLREP

Fast docking of molecular models

High resolution EM maps

Find multiple copies

Reference to target sequence correction



Alexei Vagin

Science and Technology Facilities Council

CCP4

Rigid-body fitting

Dock-EM

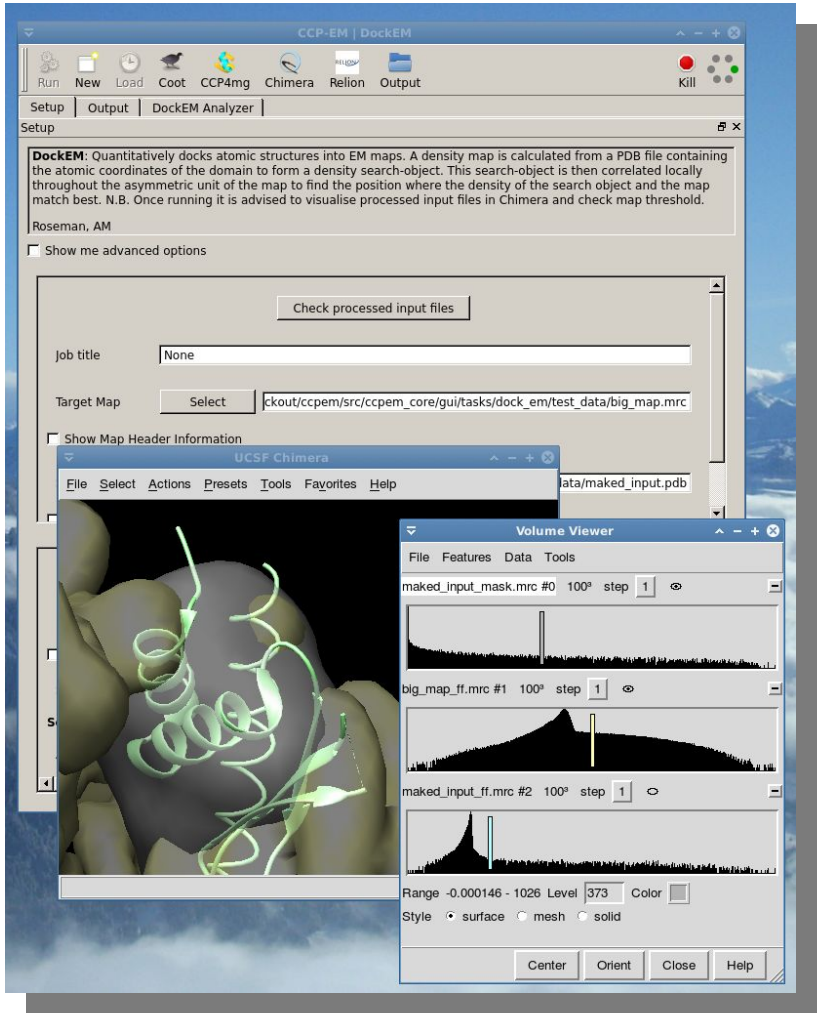
Docking atomic models at medium to low resolution

Exhaustive 6D rigid body search

Target region of interest

Solutions ranked by cross-correlation coefficient (CCC)

View best hits in Chimera

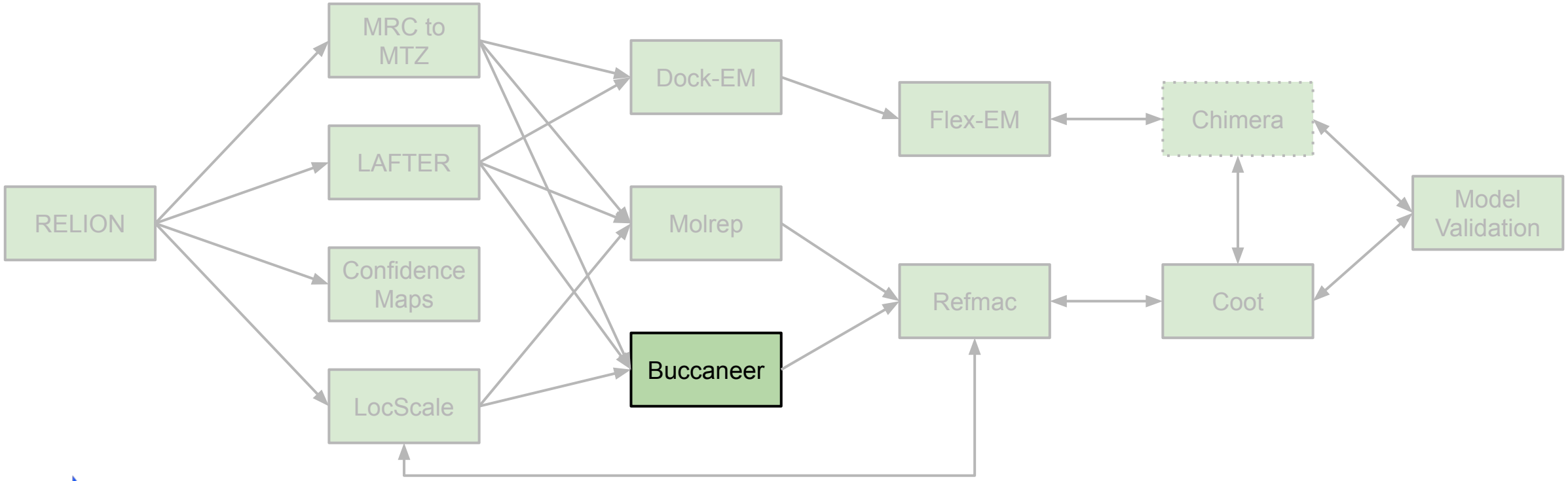


Alan Roseman

Science and Technology Facilities Council

CCP-EM workflow

Single Particle Reconstruction | Map Optimisation | **Docking / Model Building** | Automated Refinement | Interactive Refinement | Validation



Science and Technology Facilities Council

Buccaneer & Nautilus

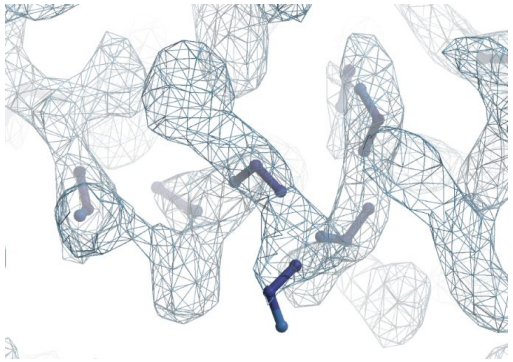
Automated model building in high-resolution maps (<4Å)

Buccaneer: amino acid
Nautilus: nucleic acid

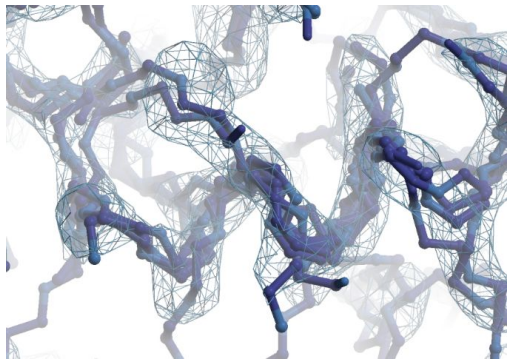
Requires map and sequence

Kevin Cowtan

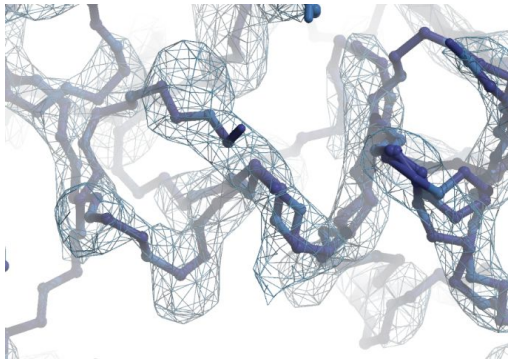
Scott Hoh



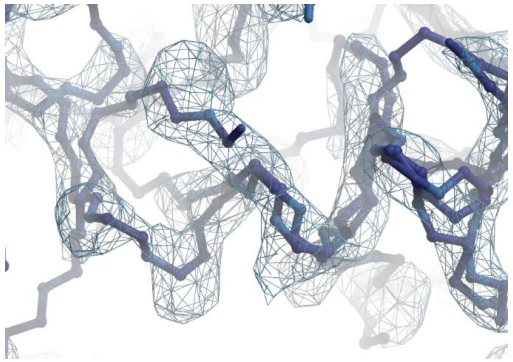
Find Cα seed positions



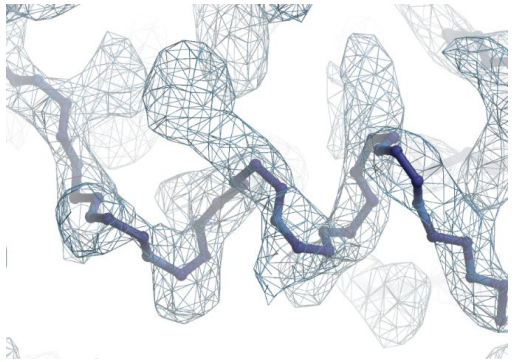
Grow chain fragments



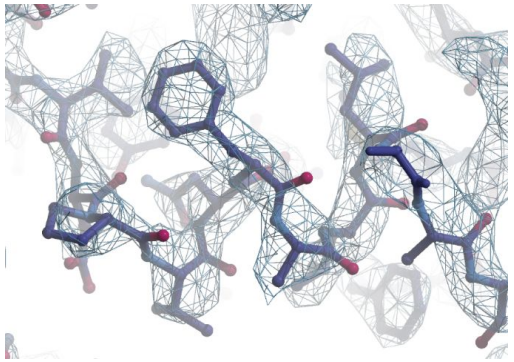
Join overlapping fragments



Link adjacent fragments, assign & correct **sequence**, **filter** poor-quality fragments



Prune inconsistent fragments



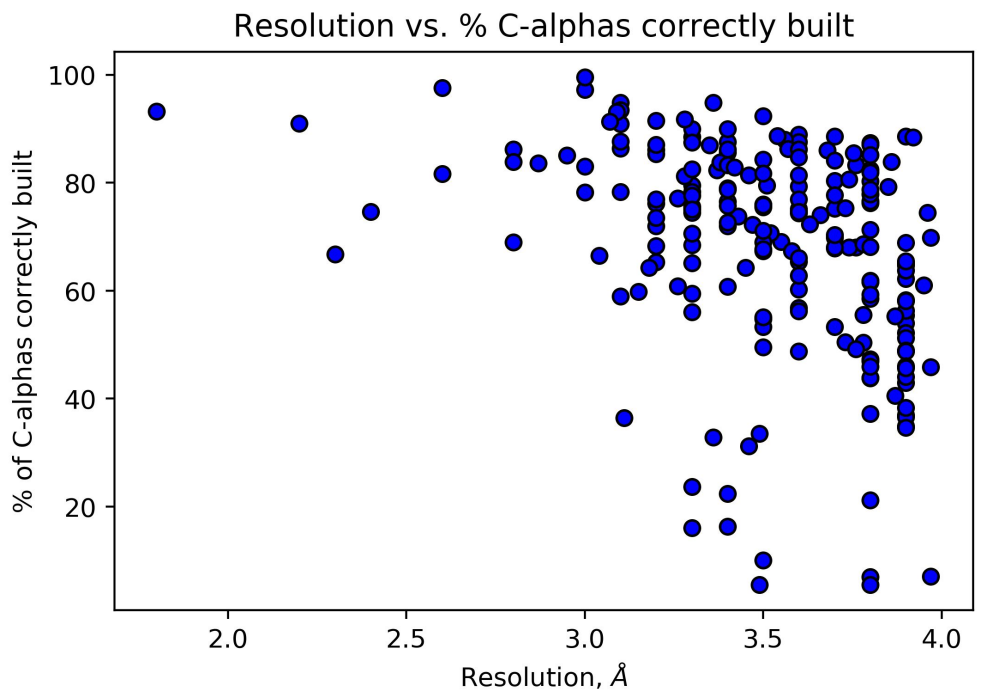
Rebuild side chains

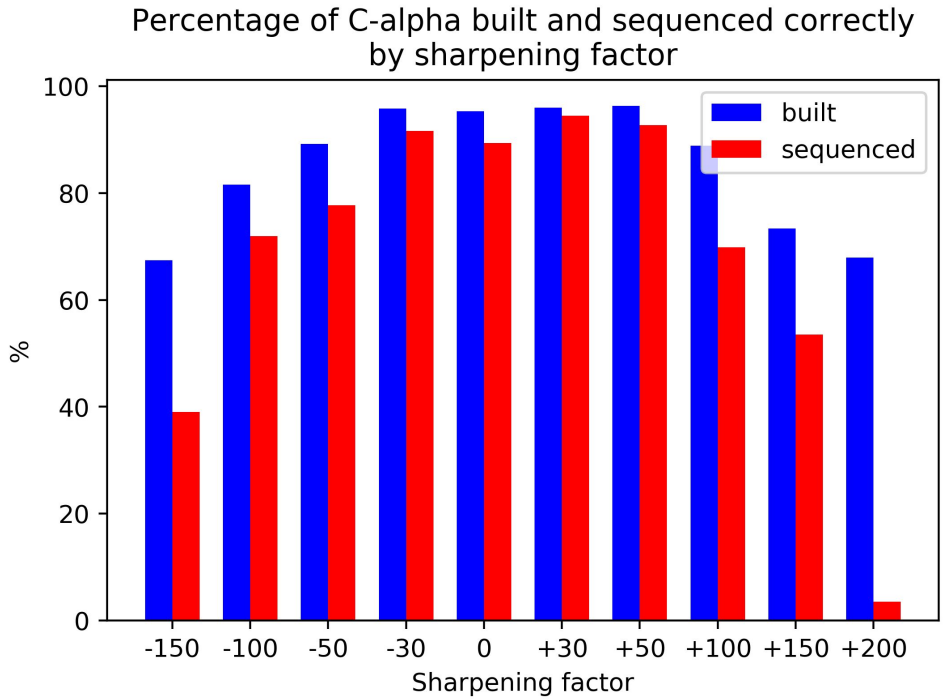
Science and Technology Facilities Council

CCP4

Buccaneer & Nautilus

Effect of resolution and map sharpening





Science and Technology Facilities Council

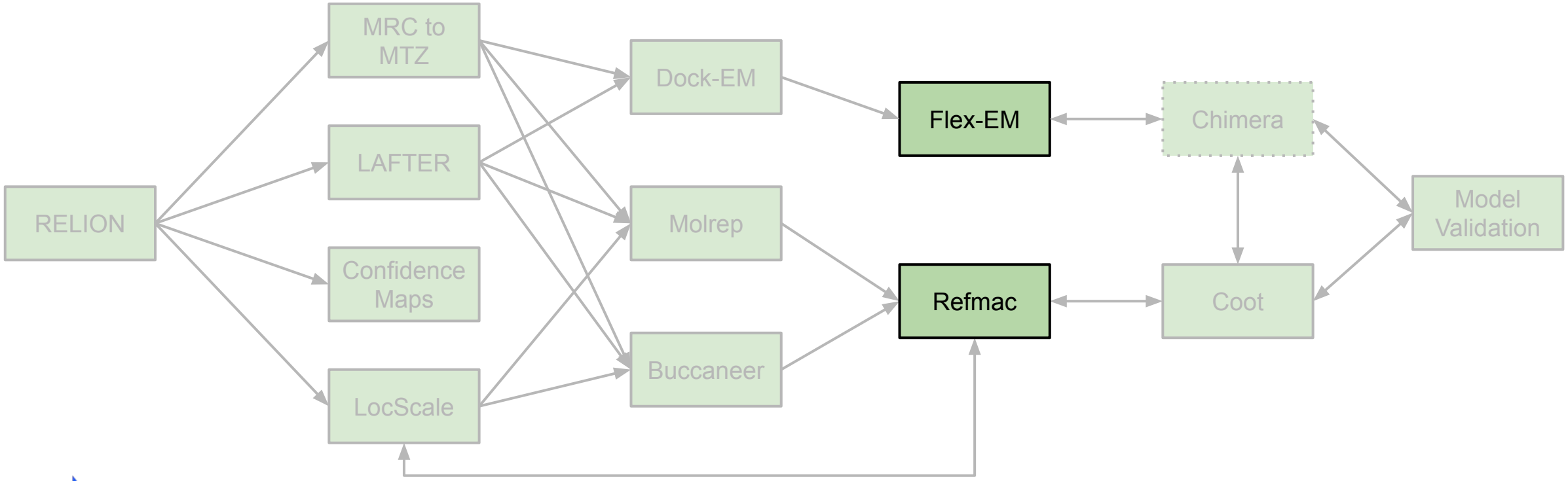
CCP4

Science and Technology Facilities Council

Model refinement

CCP-EM workflow

Single Particle Reconstruction | Map Optimisation | Docking / Model Building | **Automated Refinement** | Interactive Refinement | Validation



Science and Technology Facilities Council

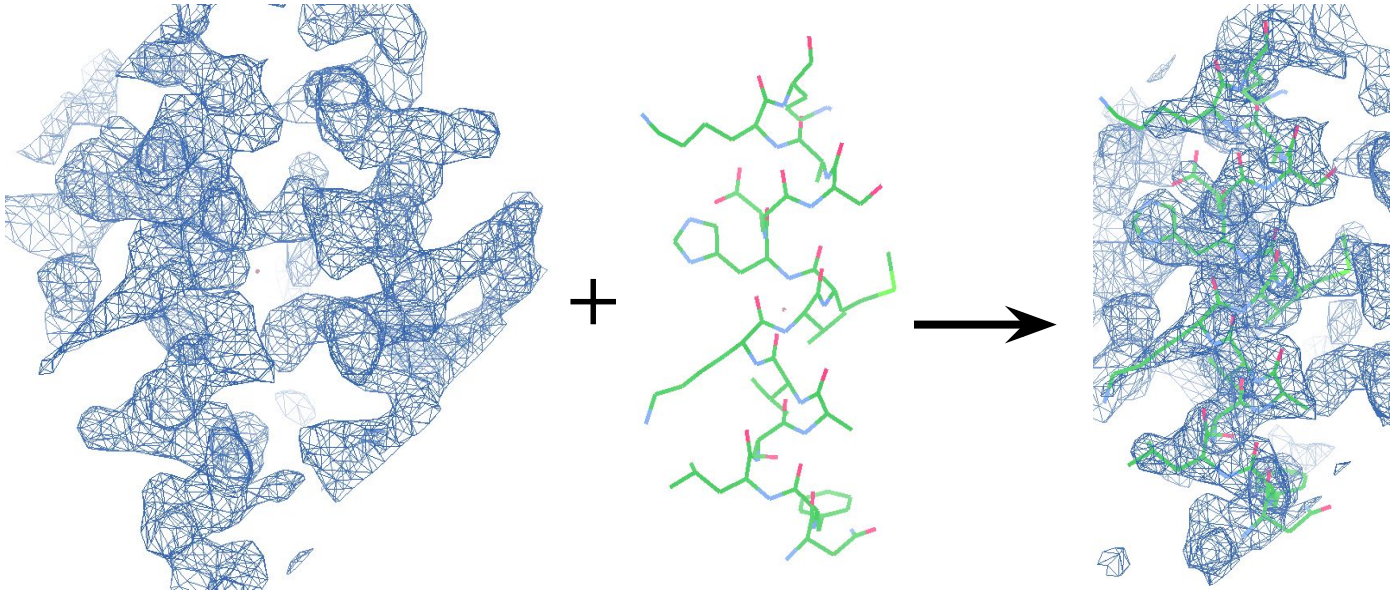
Purpose of Refinement

Aims:

- Model should agree with the observed data
- Model must be chemically and structurally sensible

Restraints:

- Bond lengths
- Angles
- Chirals
- Planes
- Some torsion angles
- B-values
- VDW repulsions



Data **Atomic model** **Fit and refine**

$$f_{total} = wf_{data} + f_{geometry}$$

Weight ***Likelihood of the data*** ***Probability of the model***

Science and Technology Facilities Council

REFMAC

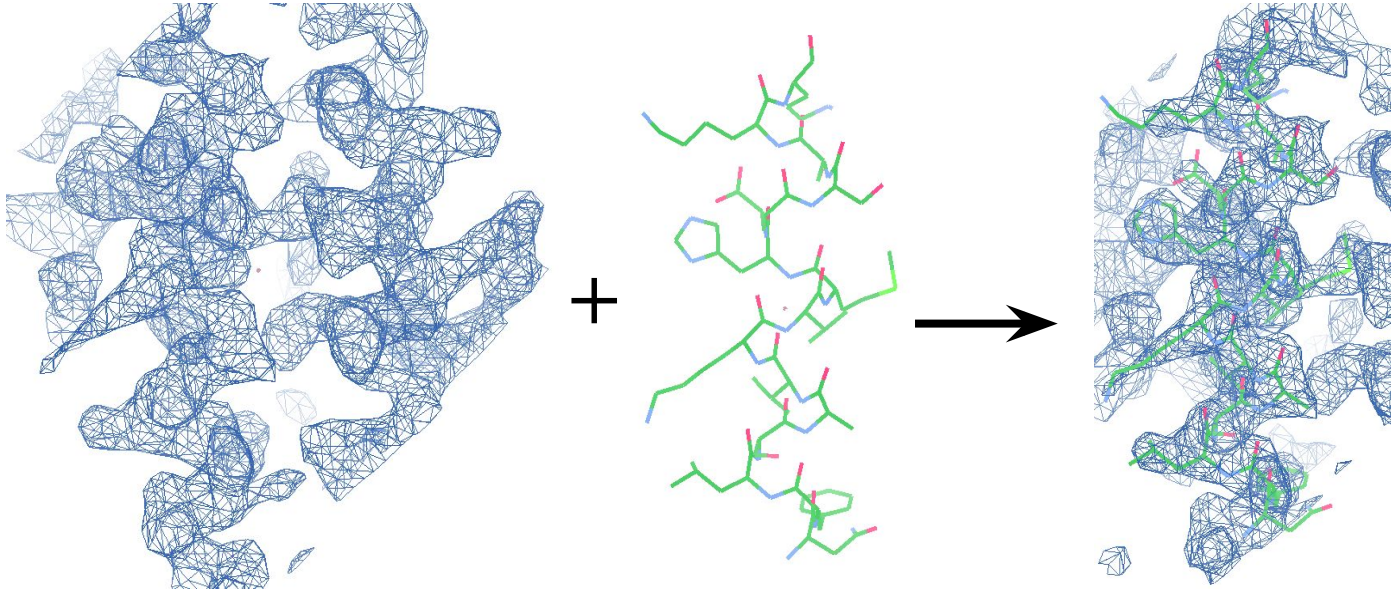
Model refinement at high resolution (<~5Å)

Vector difference refinement
- Amplitudes and phases

Automatic handling of EM maps:
- Map to MTZ conversion
- Electron structure factors
- Map sharpening

Global or local refinement modes

Standard and additional restraints



Data | Atomic model | Fit and refine

Garib Murshudov

Oleg Kovalevskiy

Rob Nicholls

UKRI Science and Technology Facilities Council

CCP4

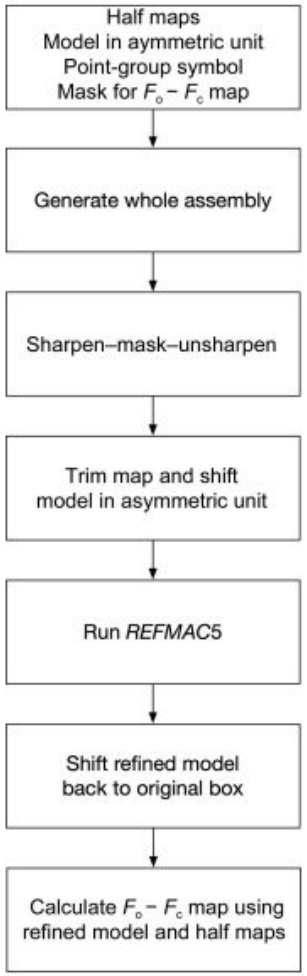
CCP-EM refinement

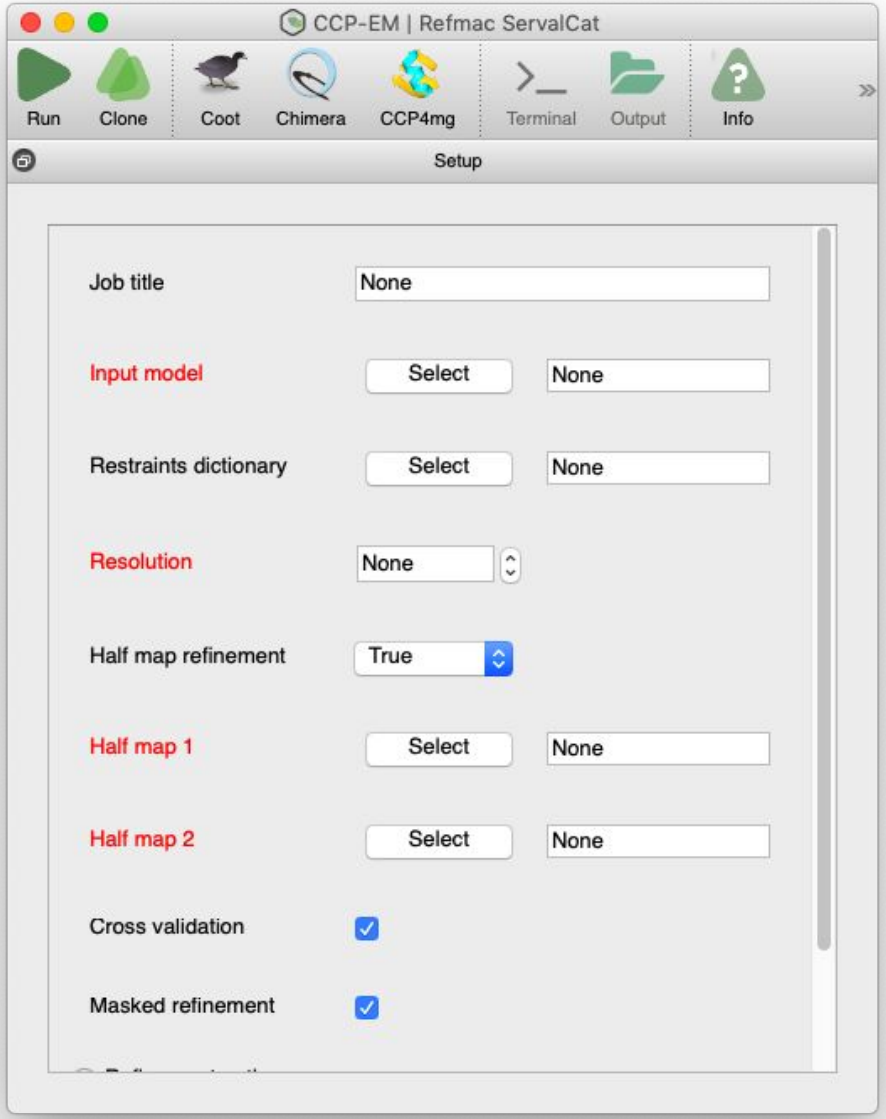
Refmac-Servalcat Pipeline

Servalcat Python wrapper for Refmac

Pre-/post-Refmac processing functionality

- Point group symmetry
- Half map input
- Auto Sharpening
- Subvoluming masking
- Weight optimisation
- Variance weighted difference map
- Half map cross validation



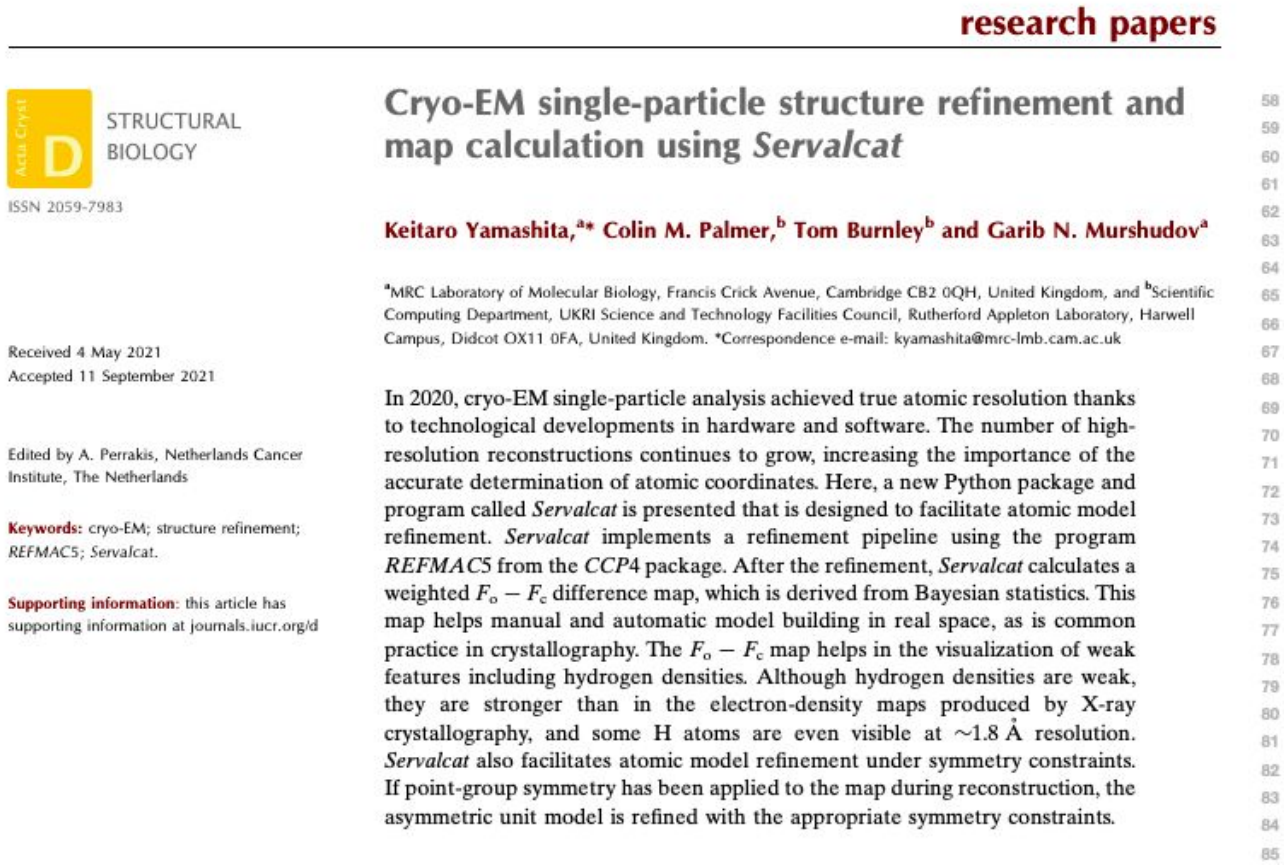


Science and Technology Facilities Council

Keitaro Yamashita

Servalcat use of symmetry

- ~50% EMDB entries have non-C1 symmetry
- Downstream analysis should be aware of this
- 2 modes: auto or strict
  - Point group Relion
- Output ASU model and symmetry expanded model
- Recommend refining / deposited ASU



research papers

STRUCTURAL BIOLOGY

ISSN 2059-7983

Received 4 May 2021
Accepted 11 September 2021

Edited by A. Perrakis, Netherlands Cancer Institute, The Netherlands

Keywords: cryo-EM; structure refinement; *REFMAC5*; *Servalcat*.

Supporting information: this article has supporting information at journals.iucr.org/d

Cryo-EM single-particle structure refinement and map calculation using *Servalcat*

Keitaro Yamashita,[a]* Colin M. Palmer,[b] Tom Burnley[b] and Garib N. Murshudov[a]

[a]MRC Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge CB2 0QH, United Kingdom, and [b]Scientific Computing Department, UKRI Science and Technology Facilities Council, Rutherford Appleton Laboratory, Harwell Campus, Didcot OX11 0FA, United Kingdom. *Correspondence e-mail: kyamashita@mrc-lmb.cam.ac.uk

In 2020, cryo-EM single-particle analysis achieved true atomic resolution thanks to technological developments in hardware and software. The number of high-resolution reconstructions continues to grow, increasing the importance of the accurate determination of atomic coordinates. Here, a new Python package and program called *Servalcat* is presented that is designed to facilitate atomic model refinement. *Servalcat* implements a refinement pipeline using the program *REFMAC5* from the *CCP*4 package. After the refinement, *Servalcat* calculates a weighted $F_o - F_c$ difference map, which is derived from Bayesian statistics. This map helps manual and automatic model building in real space, as is common practice in crystallography. The $F_o - F_c$ map helps in the visualization of weak features including hydrogen densities. Although hydrogen densities are weak, they are stronger than in the electron-density maps produced by X-ray crystallography, and some H atoms are even visible at ~1.8 Å resolution. *Servalcat* also facilitates atomic model refinement under symmetry constraints. If point-group symmetry has been applied to the map during reconstruction, the asymmetric unit model is refined with the appropriate symmetry constraints.

Yamashita *et al*, Acta D, Accepted
https://www.biorxiv.org/content/10.1101/2021.05.04.442493v2

Science and Technology Facilities Council

Weighted & sharpened difference map

Unsharpened and unweighted half maps for input

Sharpened maps can result in non-physical ADPs

Weighted = visualisation & model building

Unweighted = refinement

Calculate variance of noise (σ^2_n) in fourier frequency shells based on agreement of half maps

Global sharpening level on based on comparison of Fo vs Fc with atomic ADPs set to zero

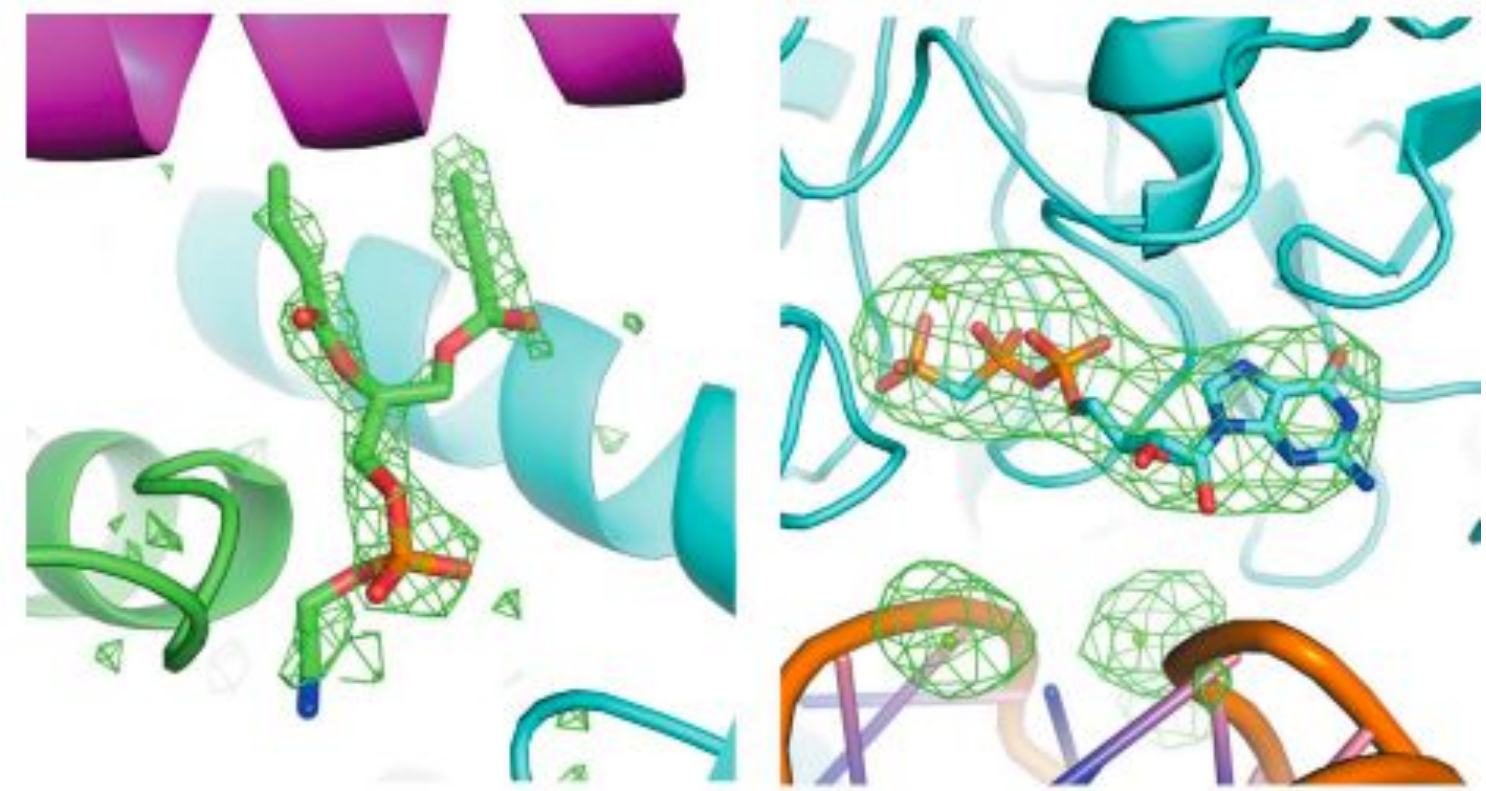
$$\sigma_n^2 = \frac{\mathrm{var}(F_{o1} - F_{o2})}{4}.$$

$$w = \frac{\sigma^2_{U,T}}{\sigma^2_{U,T} + \sigma^2_n}.$$

$$F_{diff} = \frac{w}{(\mathrm{FSC}_{full}\langle|F_o|^2\rangle)^{1/2}}(F_o - DF_c).$$

Science and Technology Facilities Council

Difference map: ligand inspection



Fo-Fc omit maps

(L) SARS-CoV-2 3a ion channel

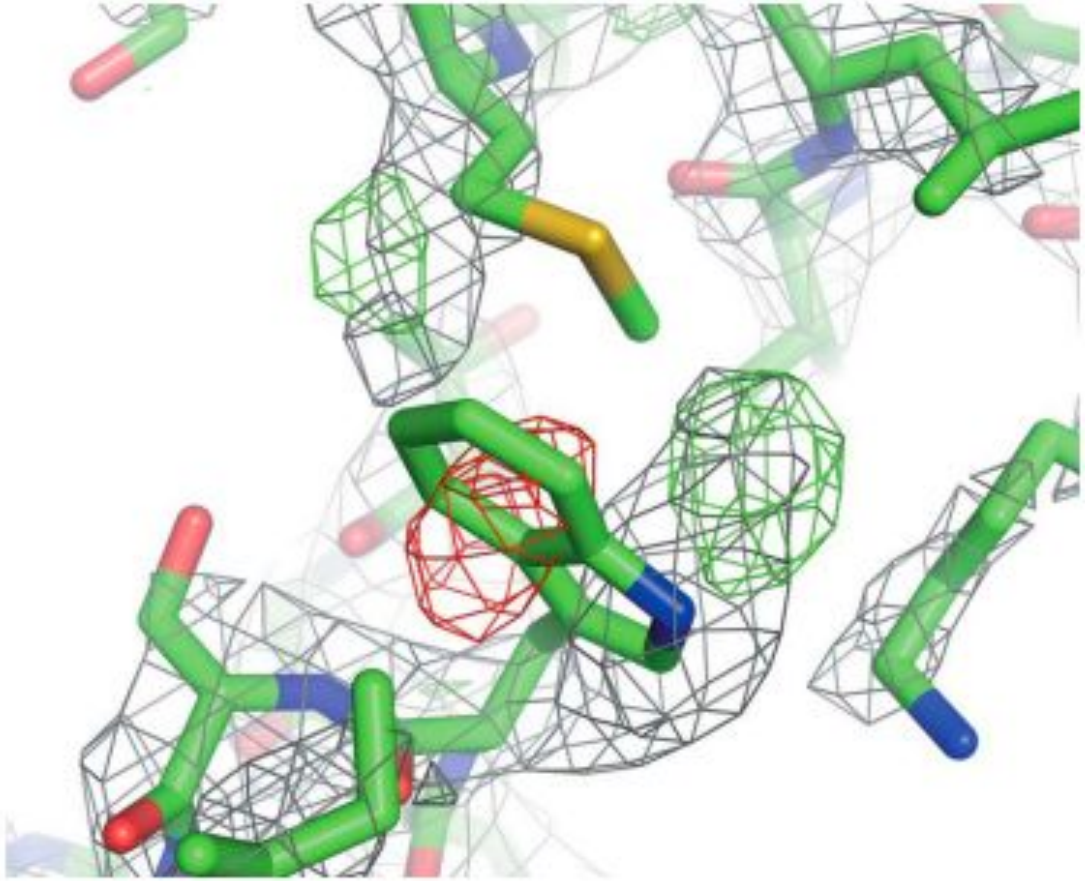
EMD-22898 (2.1Å)

(R) Kluyveromyces lactis 80S ribosome

EMD-8123 (3.6Å)

Science and Technology Facilities Council

Difference map: model improvement



Fo-Fc

Killifish Calcium homeostasis modulator

EMD-0919 (2.7Å)

Acta Crystallographica
Section D

search IUCr Journals

archive editors for authors for readers submit subscribe open ac

research papers

STRUCTURAL BIOLOGY

Volume 77 | Part 10 | October 2021 | Pages 1282-1291
https://doi.org/10.1107/S2059798321009475
OPEN ACCESS
Cited by 5

ISSN: 2059-7983

Cryo-EM single-particle structure refinement and map calculation using *Servalcat*

Keitaro Yamashita,[a]* Colin M. Palmer,[b] Tom Burnley[b] and Garib N. Murshudov[a]*

[a]MRC Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge CB2 0QH, United Kingdom, and [b]Scientific Computing Department, UKRI Science and Technology Facilities Council, Rutherford Appleton Laboratory, Harwell Campus, Didcot OX11 0FA, United Kingdom
*Correspondence e-mail: kyamashita@mrc-lmb.cam.ac.uk, garib@mrc-lmb.cam.ac.uk

Science and Technology Facilities Council

Masked refinement

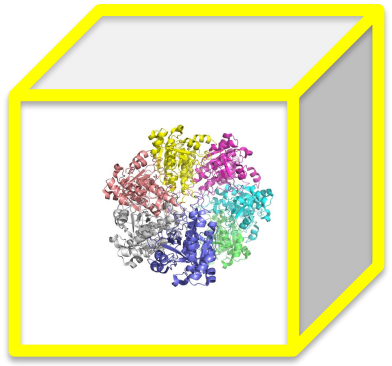
Automated subvolume refinement

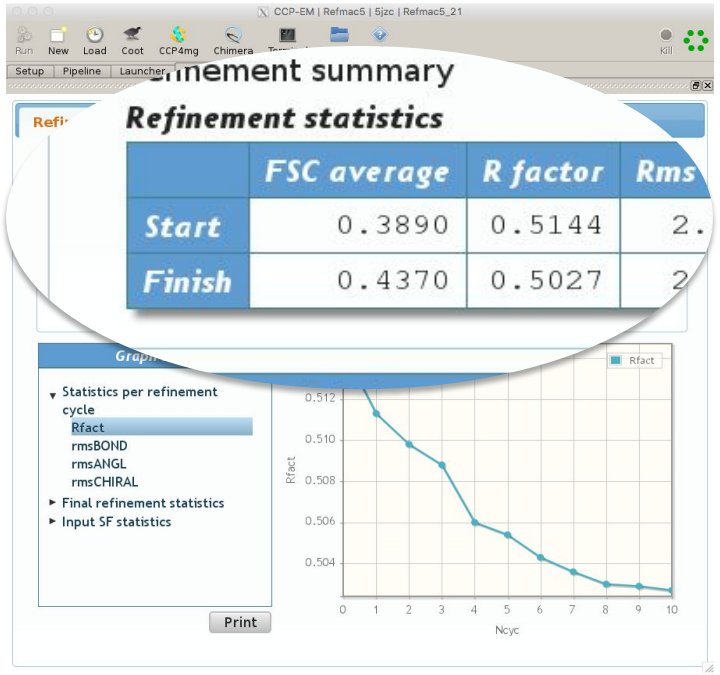
Mask map around input PDB

Sharpen-mask-unsharpen

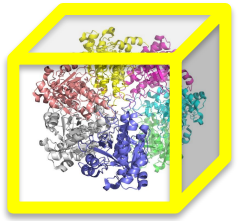
Refine and validated in subvolume

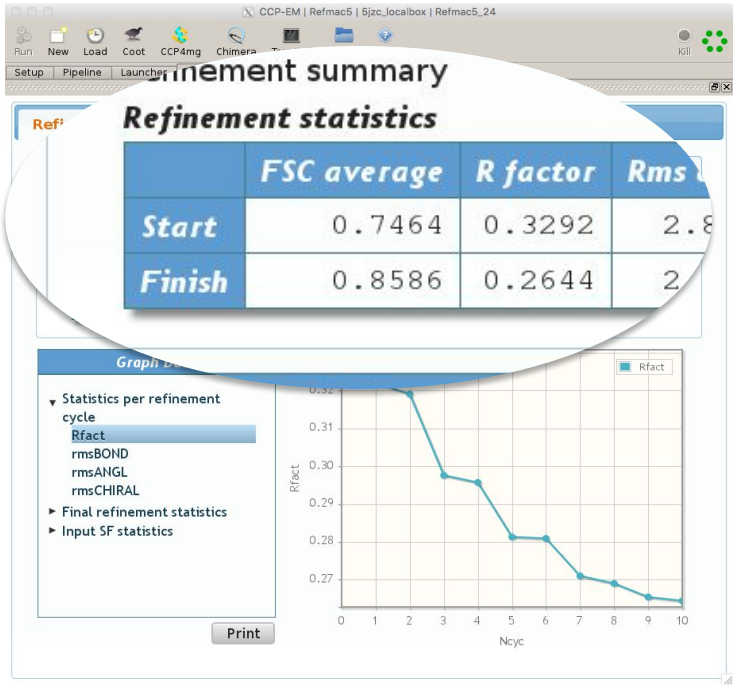
~~Local refinement~~ == Masked refinement





Helical filament 5jzc (4.2 Å)





Science and Technology Facilities Council

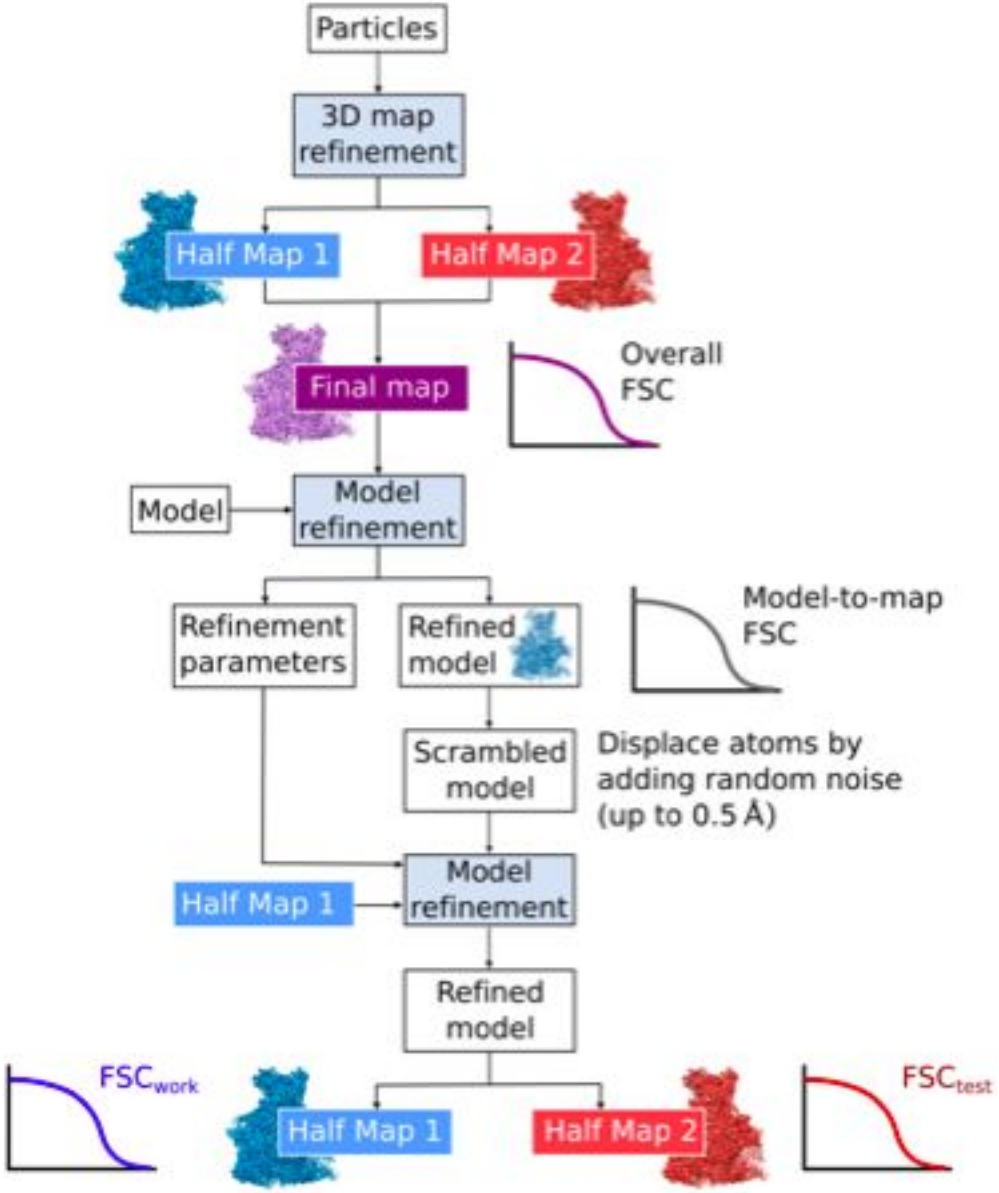
Weight optimisation and cross validation

Improved auto weighting (data vs stereochemical) procedure

Empirically tested with cryoEM datasets based on resolution mask box ratio

Recommend running half map cross-validation:

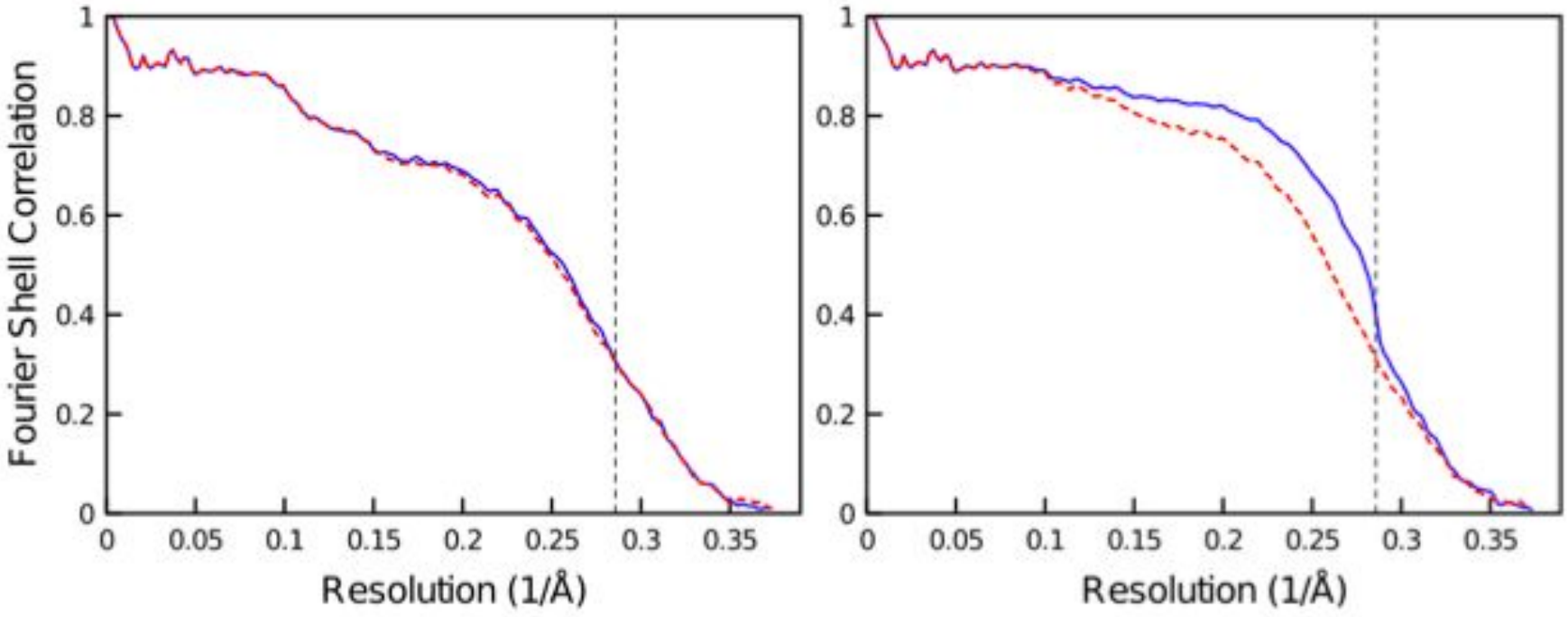
1) Initial refinement using full map
2) Second refinement using scrambled model and half map 1
3) Compare model vs map for half map 1 and half map 2



Science and Technology Facilities Council

Not over-fitted

Over-fitted



FSCwork: model refined against half map 1; compared to half map 1
FSCfree: model refined against half map 1; compared to half map 2

Science and Technology Facilities Council

Ligand refinement

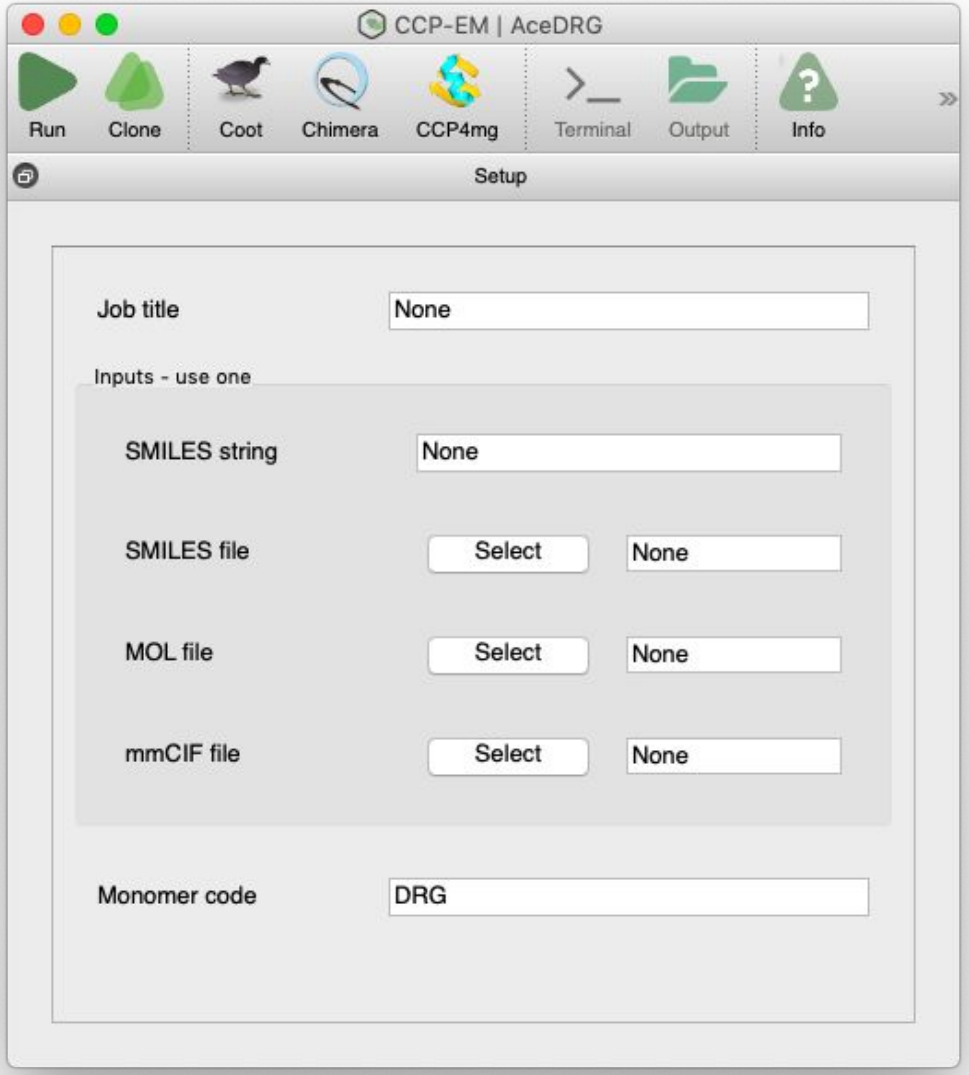
Novel ligands require custom restraints

AceDRG CCP-EM interface

Ligand cif dictionary generation

SMILES / MOL / mmCIF Input

Use as input to Refmac-Servalcat



Science and Technology Facilities Council

Garib Murshudov

Rob Nicholls

Flexible fitting

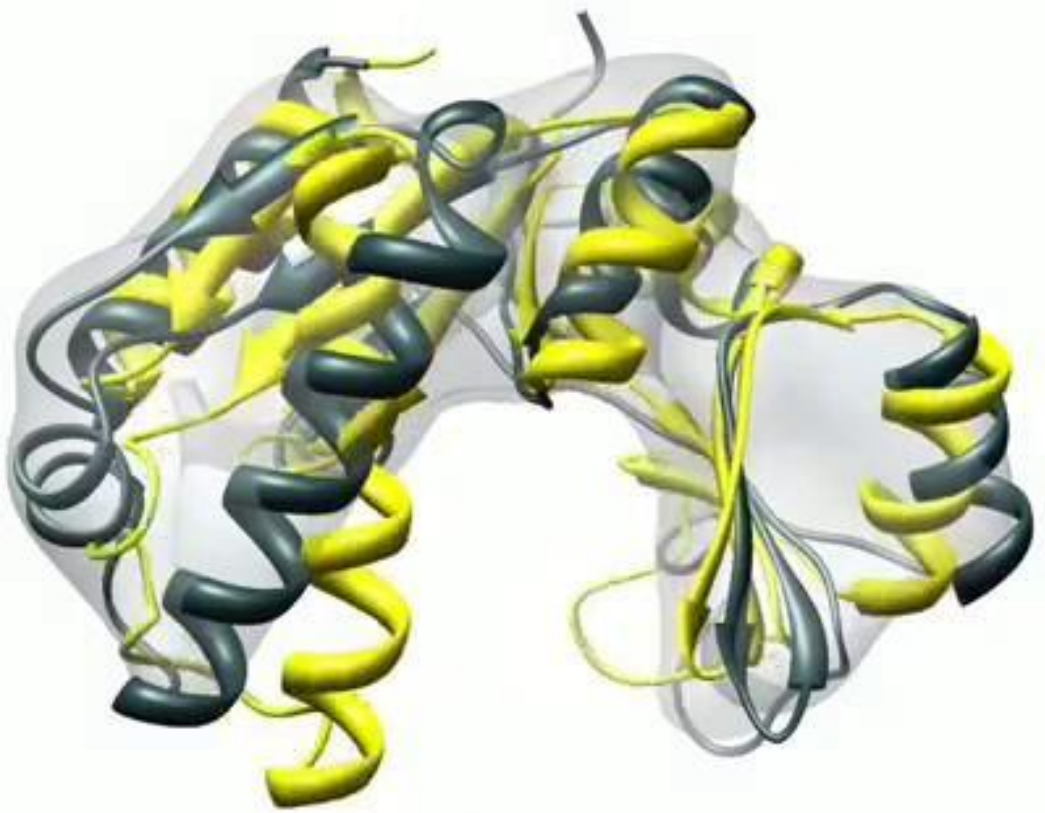
FlexEM

Model refinement at medium resolutions or early stage model building

Flexible fitting of rigid body domains to EM maps

Real space MD refinement

Rigid bodies detected based on clusters of secondary structure elements



Science and Technology Facilities Council

Maya Topf

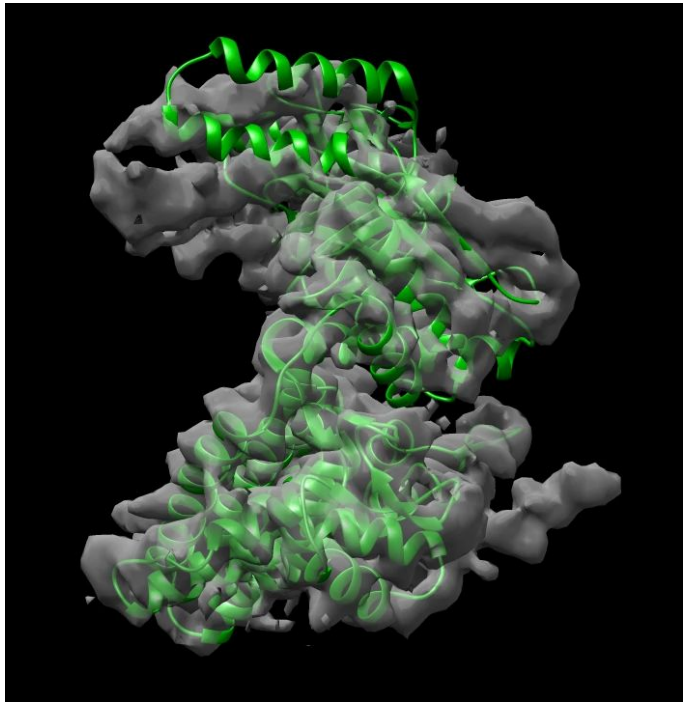
Agnel Joseph

FlexEM

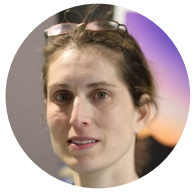
- Create series of rigid bodies - clustered secondary structure elements
- Simulated annealing MD
- In refinement rigid bodies are move to maximizes **cross-correlation with the cryoEM density map** (E^{CCC}) and minimize the violations of the stereochemical (E^{SC}) and non-bonded terms (E^{NB})

$$E=w1\times E^{CCC}+w2\times E^{SC}+w3\times E^{NB}$$

- Stereochemical (E^{SC}), non-bonded terms (E^{NB}) and rigid body restraints help to preserve the correct geometry in cases where local structures would otherwise be distorted during refinement
- Rigid body restraints based on inter-atomic or inter-residue contacts
- Atoms in a rigid body move together during the course of refinement



Science and Technology Facilities Council



Maya Topf

Agnel Joseph

Other refinement tools

Phenix: Gradient/Simulated annealing MD/exhaustive (Afonine et al. 2012, Afonine et al. 2018))

MDFF: Molecular Dynamics (Trabuco et al. 2008; Singharoy et al. 2016)

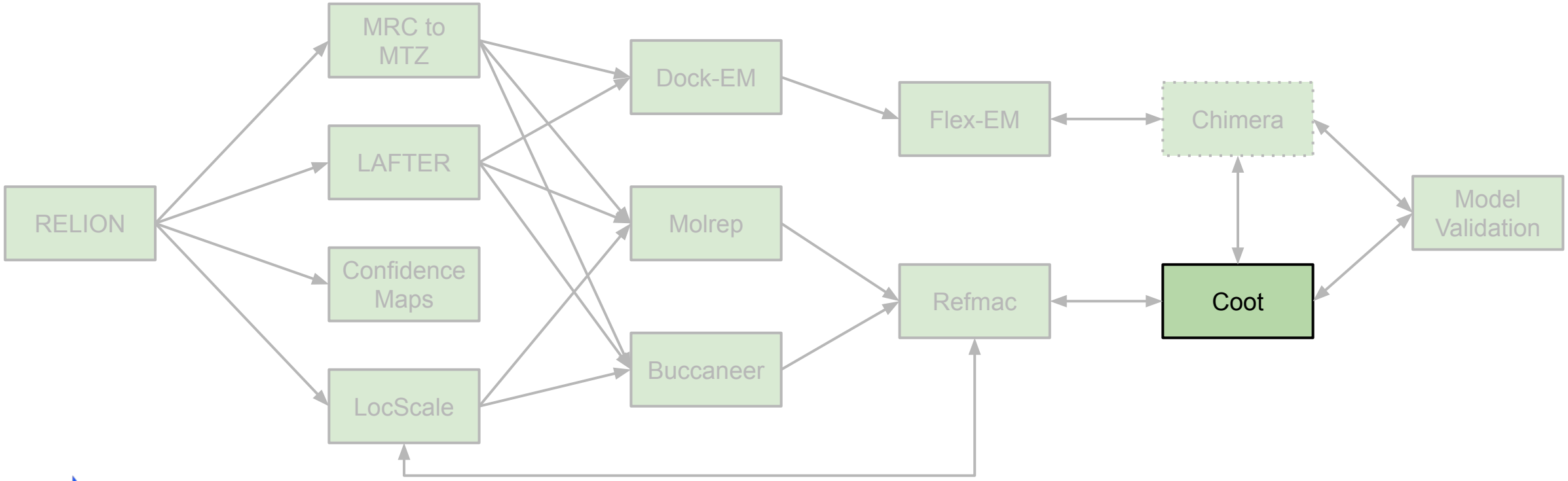
Direx, NMFF, iMODFIT: Normal modes (Wang and Schroder 2012; Tama et al. 2004; Blanco and Chacon 2013)

Rosetta, Direx: Monte-Carlo/stochastic (Wang et al 2016; DiMaio et al. 2015; Wang and Schroder 2012)

ISOLDE: Interactive steered MD refinement (Croll, 2018)

CCP-EM workflow

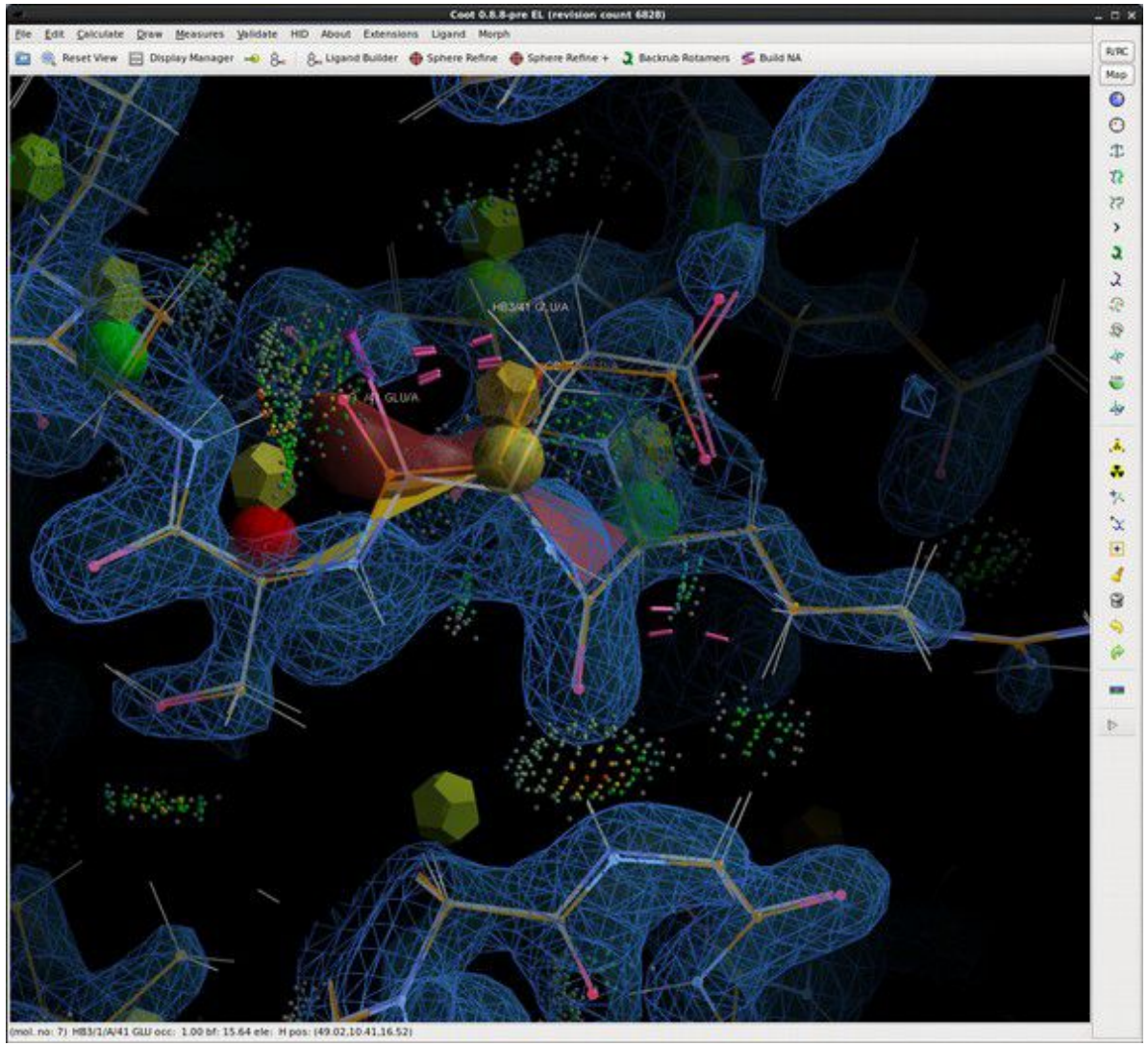
Single Particle Reconstruction | Map Optimisation | Docking / Model Building | Automated Refinement | **Interactive Refinement** | Validation



Science and Technology Facilities Council

Coot

Coot 0.9x packaged with CCP-EM at present

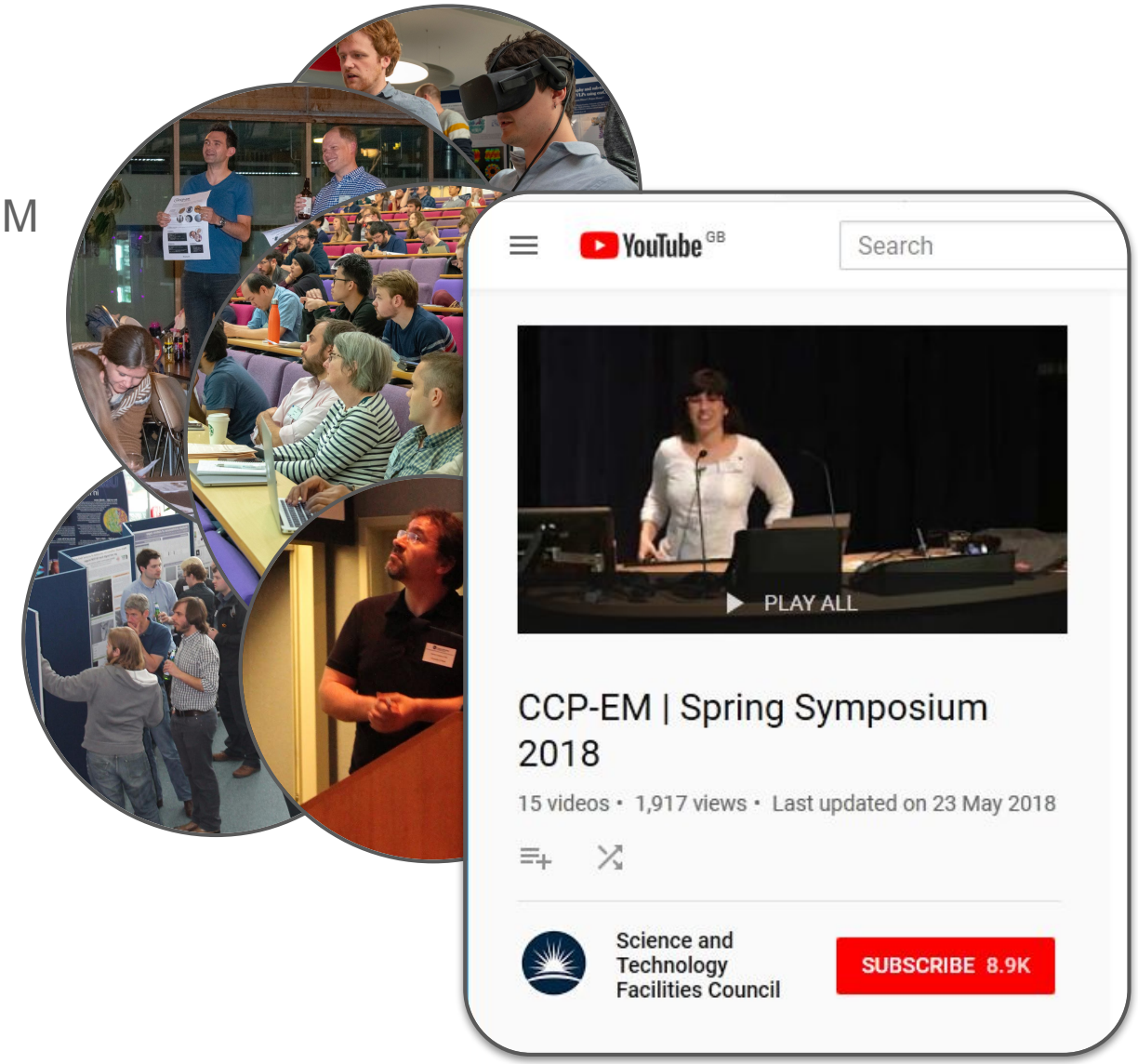


Paul Emsley

Science and Technology Facilities Council

CCP-EM Symposium

- **National UK cryoEM conference**
- **Celebrating** new method developments in cryo-EM
- Running since 2015
- **Last 8 years talks on YouTube**
  - www.ccpem.ac.uk/symposium.php



Science and Technology Facilities Council

9th CCP-EM Symposium - 25-27th April

Confirmed speakers include:

Alan Brown (Harvard)
Beatriz Costa Gomes (Turing)
Tristan Croll (Altos)
Ieva Drulyte (TFS)
Rene Frank (Leeds)
Michael Grange (Franklin)
Matt Iadanza (STFC)
Kiarash Jamali (MRC-LMB)
Gerard Kleywegt (EBI)
Sam Lacey (Human Technopole)
Bonnie Murphy (Max Planck)
Randy Read (Cambridge)
Ricardo Righetto (Basel)
Elizabeth Villa (UCSD)* remote
Janet Vonck (Max Planck Biophysics)

Bridget Carragher (Chan Zuckerberg Imaging Institute)
Victoria Garcia Giner (Imperial College London)
Peter Wing (Oxford)
Elizabeth Wright (UW-Madison)

Fully hybrid with free virtual attendance
See CCP-EM Mailing List / Website for registration details

Science and Technology Facilities Council