

Welcome!

**7th DLS-CCP4 Data Collection and
Structure Solution Workshop**

Welcome to the Virtual Space!

In order to make the Workshop most efficient, help tutors and each other, please:

- **Use your full name** when on Zoom/Slack, so that others know who you are
- **During a talk/presentation:**
 - ~ *Ensure your video/audio are off*
 - ~ *Post questions in the chat*
- **In any other moment:**
 - ~ *Feel free/you are encouraged to keep your video and audio on*
 - ~ *Just speak up to ask a question*
- **What we expect – our code of conduct:**
 - ~ The Diamond-CCP4 Data Collection and Structure Solution Workshop is an inclusive event where people should feel comfortable sharing their work, opinions, and perspectives. All of us commit to engaging with each other mindfully to ensure an environment that promotes shared learning and collaboration.

How to get the most out of the workshop

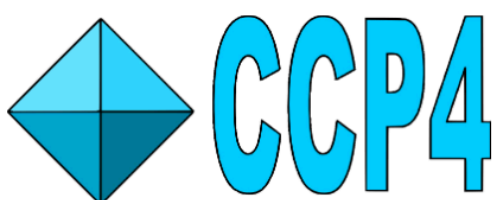
- **Ask questions!** The tutors are experts and developers of the software you will use. They are not all available for the full week, please approach them if you need help.
- **Talk to your each other.** The attendees at the course have various levels of expertise and scientific backgrounds. You may learn from each other and not just the tutors.
- **Take your time.** You will get the most out of the tutorials if you work through each step carefully and think about what you are doing
- **Course materials** will be uploaded to the website <http://www.ccp4.ac.uk/schools/DLS-2020>, so don't feel you have to copy every slide. You will have them available at the end of the workshop.

A few more points of the moment

- **If you still have problems with software setup and data upload:** post a message to the # setup channel on the Slack and we will help you out.
- **Your posters:** please upload them to the # 11-postersession channel. Each upload will start a thread, in which the poster discussion will then take place asynchronously by Slack chat

If progress you make during the workshop leads to published results we ask that you please acknowledge this in your papers. We will link to them on our website. This helps us to keep delivering similar workshops in the future.

We wish you every success with your projects!



Collaborative Computational Project No. 4
Software for Macromolecular X-Ray Crystallography



[2019 Home](#)

[2018 Home](#)

[2017 Home](#)

[2016 Home](#)

[2015 Home](#)

[2014 Home](#)

DLS-CCP4 Data Collection and Structure Solution Workshop



This page lists some of the publications where the research involved benefited from the assistance given at the DLS-CCP4 workshop. We are grateful to each of the authors for acknowledging the help they received at the workshop.

- [2017 Publications](#)
- [2015 Publications](#)
- [2014 Publications](#)

2017 Publications

The authors of the following publications have expressed their gratitude for the help they received in solving their structures at the DLS-CCP4 school in 2017.

- ["Crystal structure of human Acinus RNA recognition motif domain"](#) – Humberto Fernandes, Honorata Czapinska, Katarzyna Grudziak, Janusz M. Bujnicki and Martyna Nowacka. *PeerJ* (2018).
- ["Crystal Structure and Directed Evolution of Specificity of NlaIV Restriction Endonuclease"](#) –

DLS-CCP4 Data Collection and Structure Solution Workshop
November 30 - December 10, 2020, Diamond Light Source Ltd., UK

Introduction to CCP4: Project and Software

Eugene Krissinel & Ronan Keegan

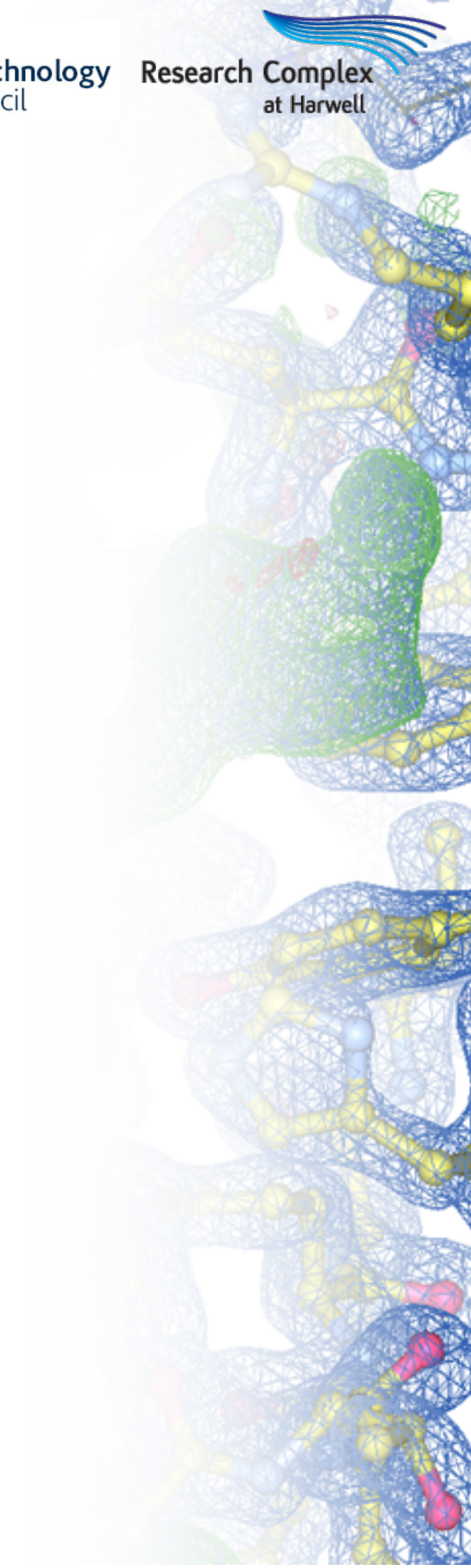
CCP4, Research Complex at Harwell
Didcot, Oxfordshire, UK



Science & Technology
Facilities Council

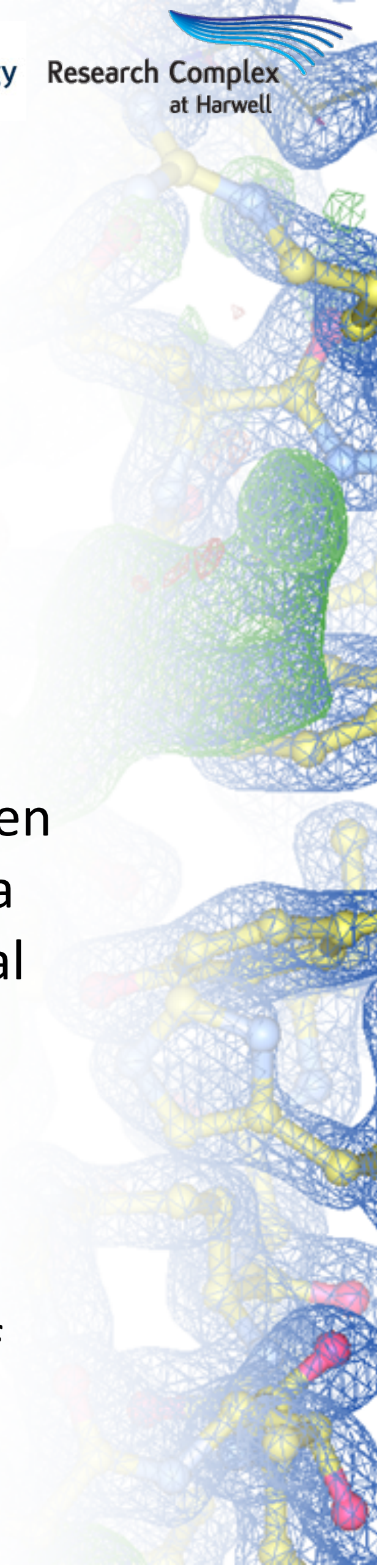


A Project with History



What is CCP4?

- CCP4 stands for “*Collaborative Computational Project No. 4 in Protein Crystallography*”
- One of several CCPs set up in the UK to advance and support scientific software developments
- CCP4 was formally set up in **1979** to support collaboration between researchers working on MX software in the UK, and to assemble a comprehensive collection of software to satisfy the computational requirements of the relevant UK groups
- Funded by the Biotechnology and Biological Sciences Research Council (BBSRC) and industrial licensing
- The project is coordinated at Scientific Computing Department of the Science and Technology Facilities Council (STFC)



Main Activities

- Today CCP4's main activities are:
 - ❖ Supporting the development and distribution of a comprehensive suite of software for performing structure solution in macromolecular crystallography
 - ❖ Providing educational and training events to users of the software including schools/workshops where users can get access to expert assistance with structure solution problems
 - ❖ Providing other resources to the user community including the CCP4BB mailing list, Wiki's and access to web/cloud-based computational resources

Main CCP4 Groups

- *Oxford*
 - ~ CCP4 core team – software development, maintenance, distribution, educational outreach and CCP4 online resources
 - ~ Diamond – data processing developments (DIALS)
- *Cambridge*
 - ~ MRC/LMB – software development for refinement (REFMAC) and data processing (MOSLFM, Aimless, etc.)
 - ~ Uni. of Cambridge – Phaser development team for phasing software
- *University of York*
 - ~ Software developments (CCP4mg, Buccaneer, Privateer)
 - ~ CCP4 GUI-2 developments



Main CCP4 Groups

● *University of Liverpool*

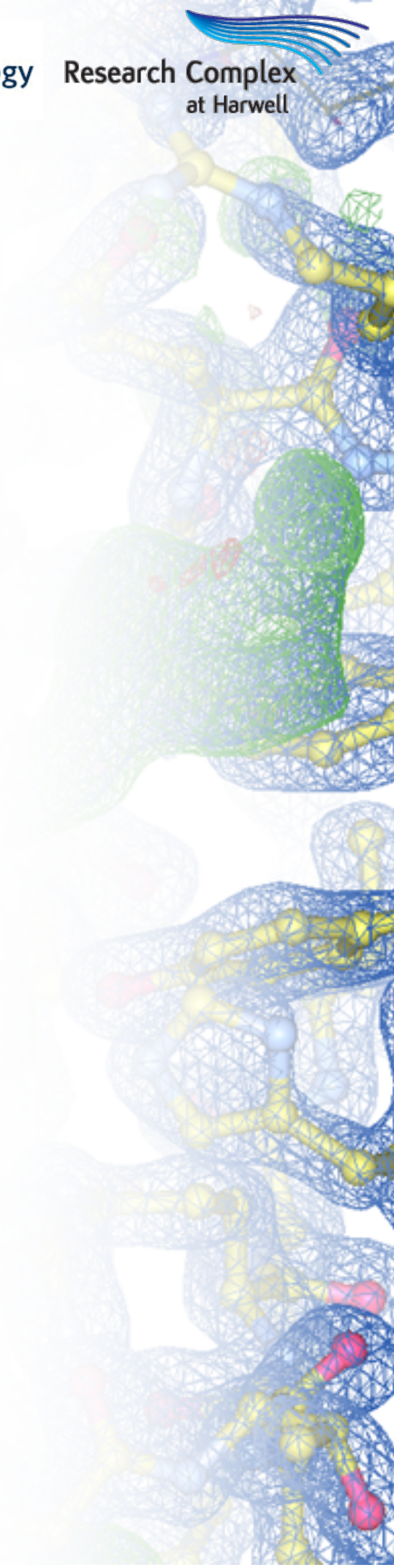
- ~ AMPLE (ab-initio MR)
- ~ SIMBAD (sequence-less MR)

● *Associated projects*

- ~ ARP/wARP (Model building, EMBL-Hamburg)
- ~ PDB redo (Validation and completion, NKI, the Netherlands)
- ~ SHELX (Experimental phasing, Goettingen and Barcelona)
- ~ Arcimboldo (Fragment-based MR, Barcelona)
- ~ Crank-2 (Automated experimental phasing, University of Leiden, The Netherlands)



Educational Outreach



CCP4 Main Event: the Study Weekend

- Takes place every January in the UK
- Each year tackles a different topic:
 - ~ Data processing
 - ~ Experimental phasing
 - ~ Molecular replacement
 - ~ Model building and refinement
 - ~ Related topics: Complementary methods (EM, NMR, SAX etc.), Structure analysis, Complexes, Ligands, Databases
- Lectures streamed live on the internet

CCP4 Study Weekend

● Proceedings published in special edition of **Acta Cryst. D**

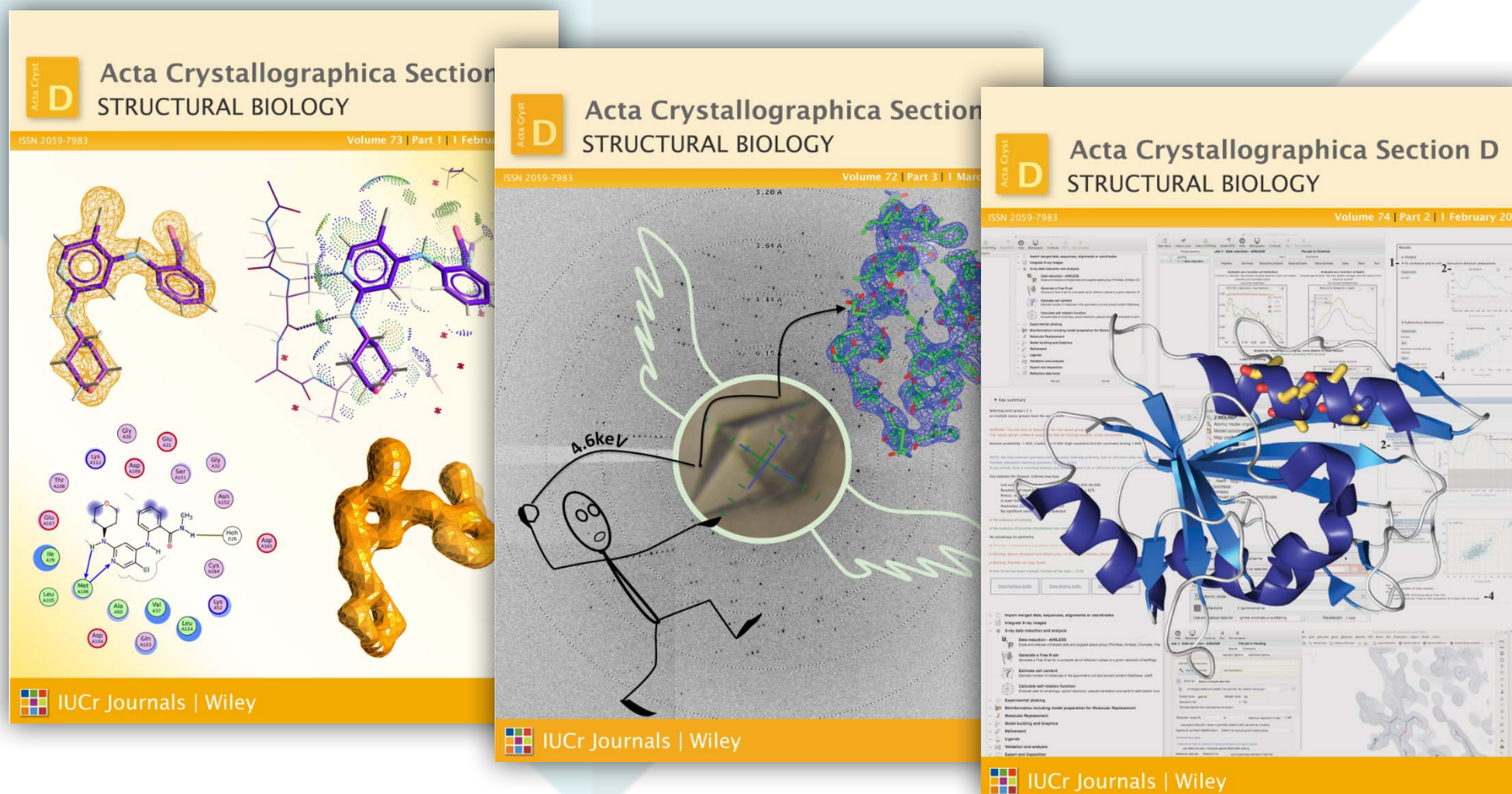
2017: *“From crystals to structure with CCP4”*

2018: *“Multi and Serial Crystal Data Collection and Processing”*

2019: *“Molecular Replacement”*

2020: *“Model Building”*

2021: *“Integrating Structural Biology”*



CCP4 Schools and Workshops

- Main aim is to give education and expert help to researchers working with collected crystal data
- Some schools also provide expert help with collecting crystal data (APS, Hamburg, Diamond)
- Applicants with challenging crystals and/or data are given strong consideration but this is not mandatory
- Software developers from CCP4, Phenix (APS Workshop), SHELX, ARP/wARP, XDS and HKL as well as others give lectures, tutorials and are also available to help with data processing and data solution throughout the meetings

<https://www.ccp4.ac.uk> – Courses & Events

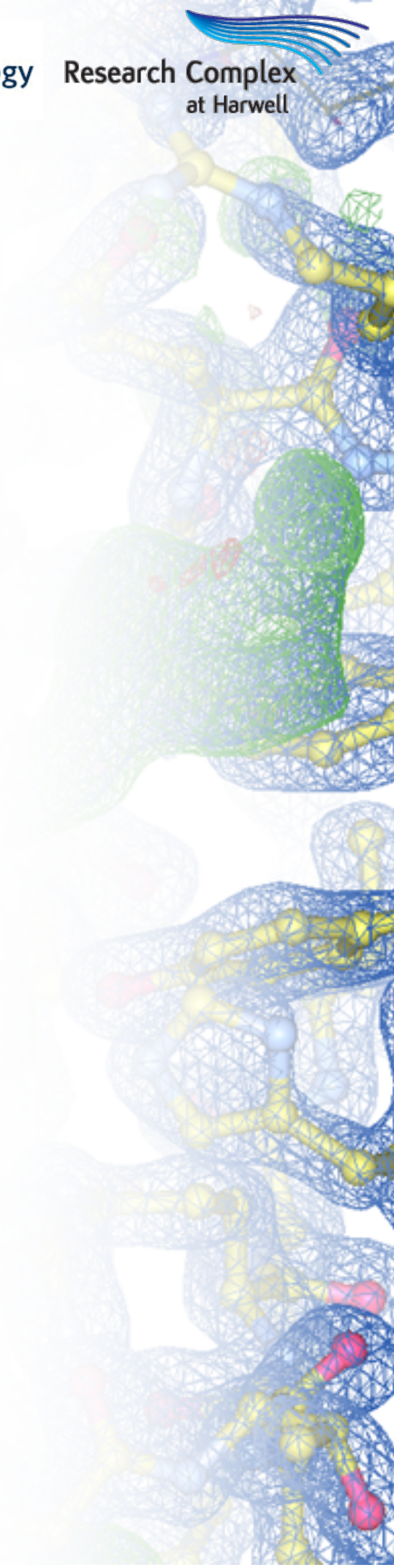
CCP4 users around the world



CCP4 Resources

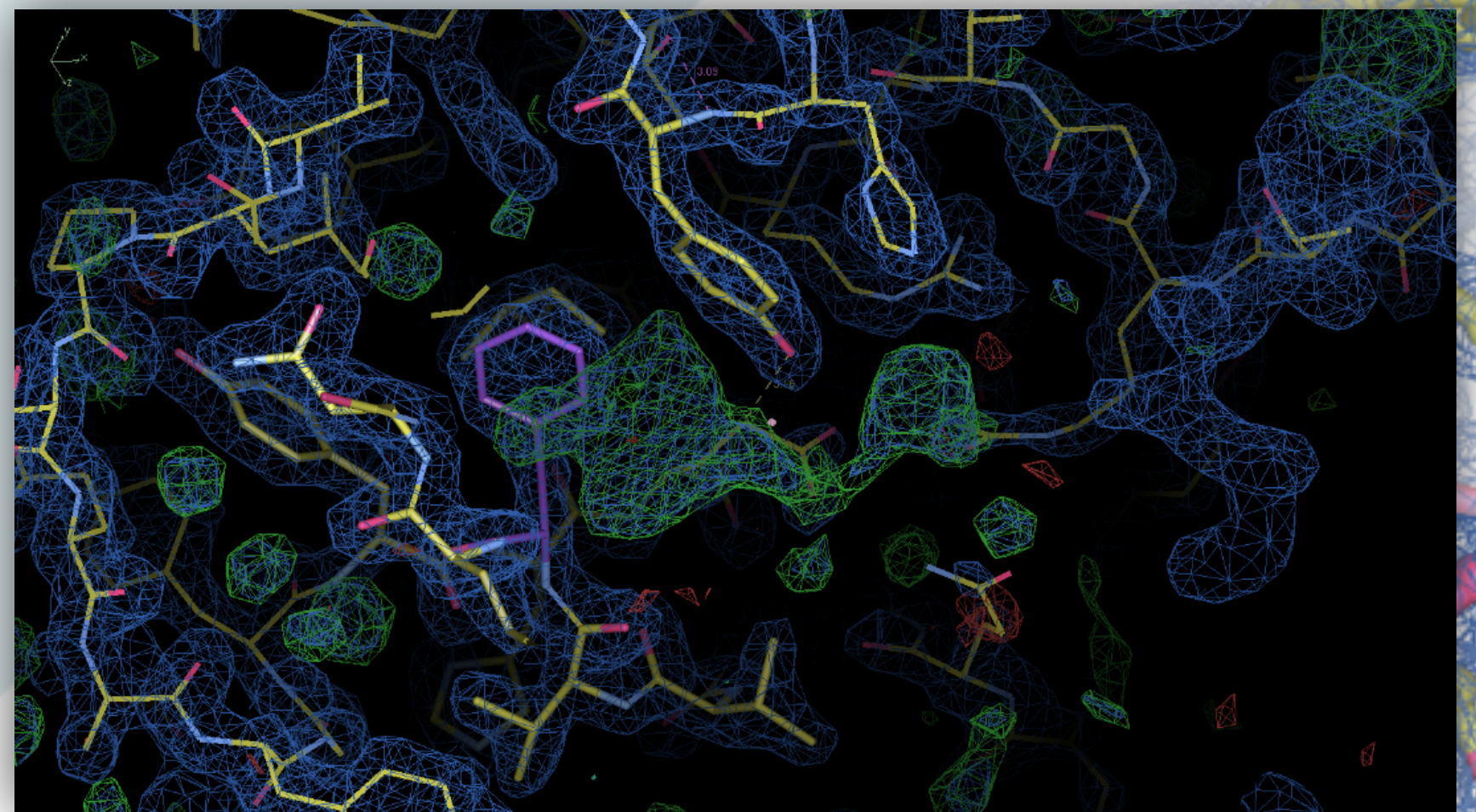
- CCP4BB mailing list — bulletin board, news, expert help and advice
- CCP4 help desk — ccp4@ccp4.ac.uk — bug reports, problems with software
- CCP4 wiki @ Konstanz — <http://strucbio.biologie.uni-konstanz.de/ccp4wiki/index.php/Crystallography>
(thanks to Kay Diederichs)
- CCP4 web site — <https://www.ccp4.ac.uk>
(downloads, documentation, news and more)

Crystallography in two words



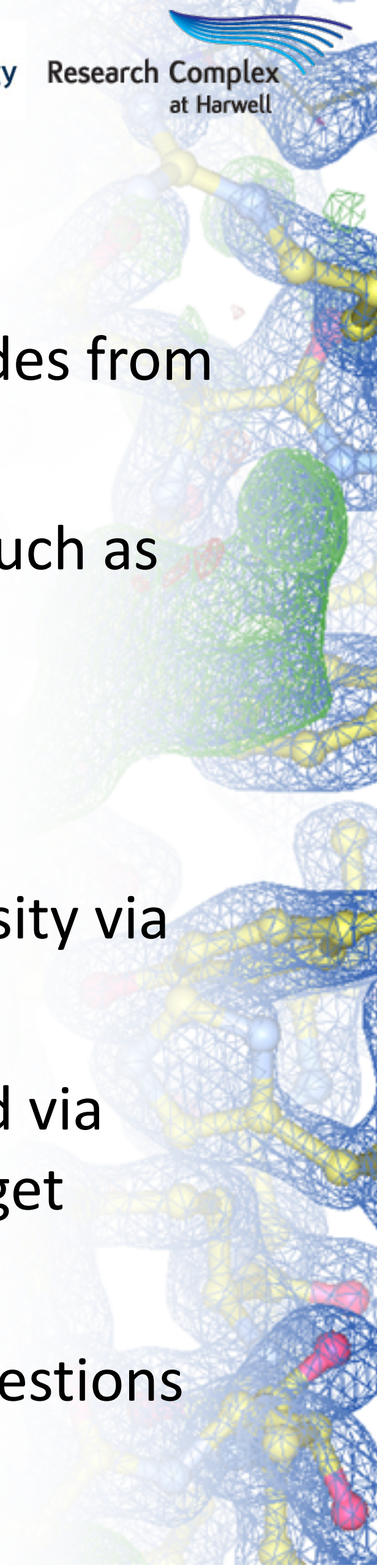
Structure Solution in Essence

- As an ultimate goal, one would like to generate an atomic model for their target
- This could be achieved by applying the inverse Fourier transform to experimental structure factors (which are Fourier transform of the structure's atomic coordinates) if they were known or could be indeed obtained in an experiment
- In reality, the X-ray diffraction experiment gives only moduli, ***but not phases***, of structure factors
- Therefore, phases need to be elicited by other means

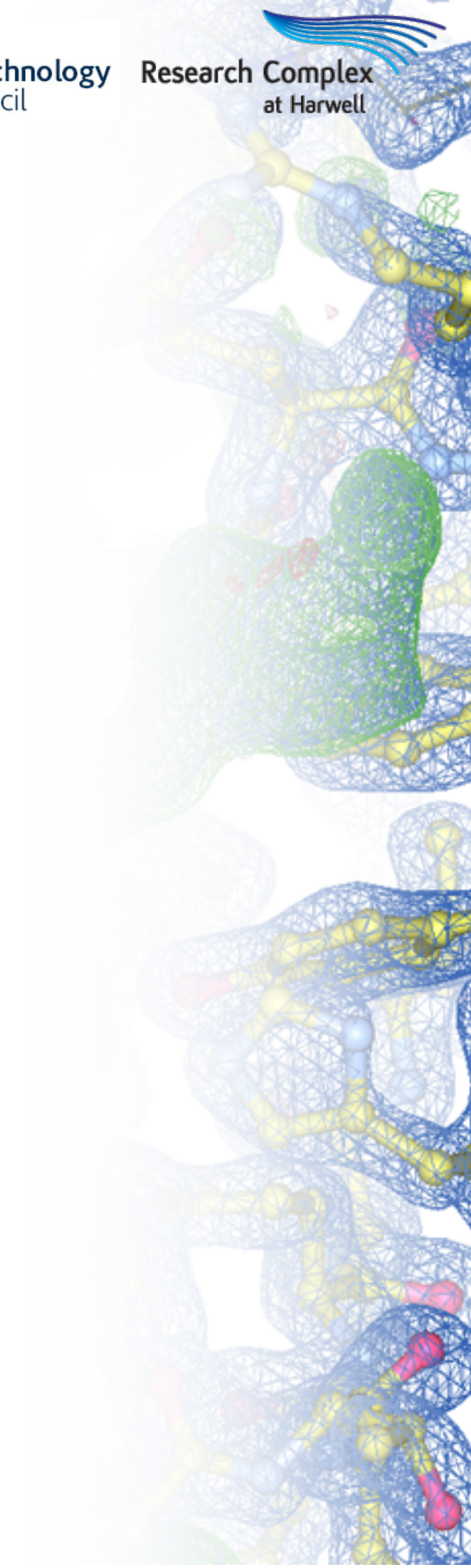


Structure Solution Pathway

- As a first step, one need to derive structure factor amplitudes from the intensities of spots in X-ray images
- Next, phases need to be approximated by using methods such as
 - ~ Experimental Phasing: SAD, MAD, SIRAS etc.
 - ~ Molecular Replacement
 - ~ Combination of the above
- Then, structure factors can be converted into electron density via *inverse Fourier transform*
- The obtained electron density map needs to be interpreted via model building and refinement to produce a model for target structure
- The model needs to be analysed to answer biochemical questions and deposited in the PDB

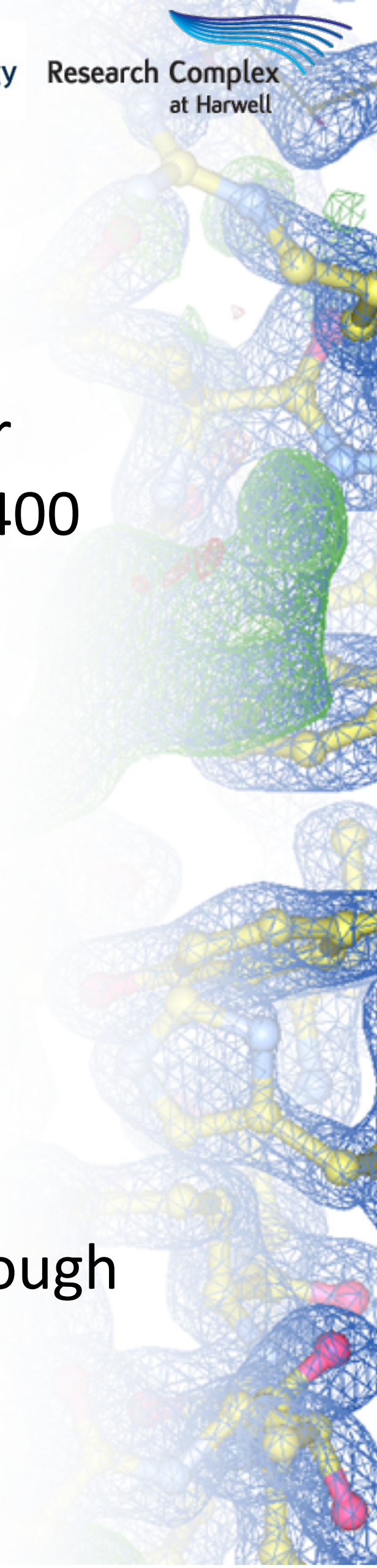


The Software



The CCP4 Software Suite

- The CCP4 Suite is a comprehensive suite of software for macromolecular crystallography and it contains about 400 program components
- New versions of the suite are released about every 12 months
 - ~ New programs
 - ~ Major updates to existing programs
 - ~ Other new features such as changes to CCP4i interface
- Revisions to the current release are made available through the CCP4 updates manager



CCP4 Scope

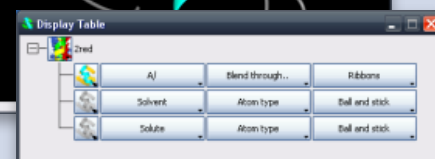
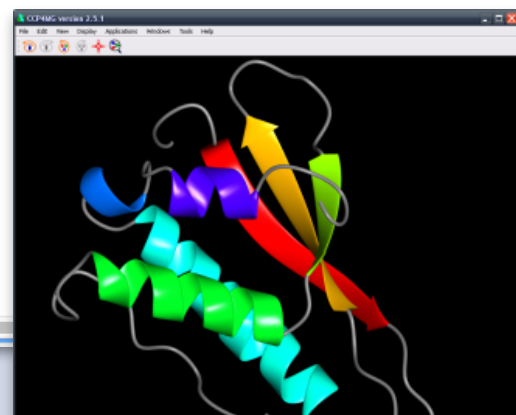
CCP4 Software

Task menu View in Coot View in CCP4mg Export MTZ Help Bibliography Clone job

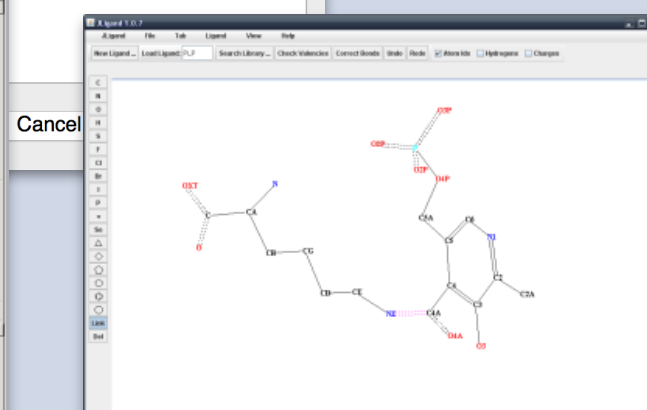
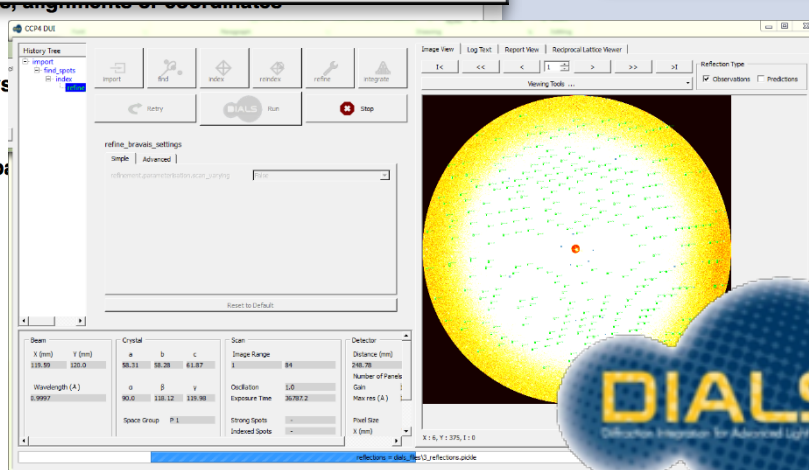
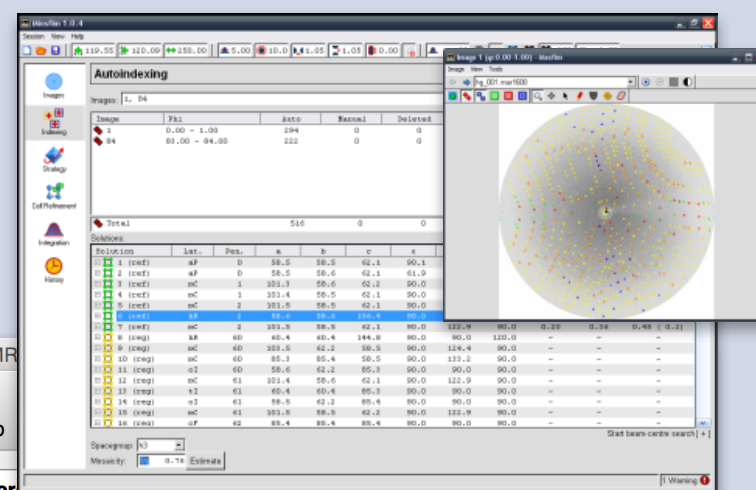
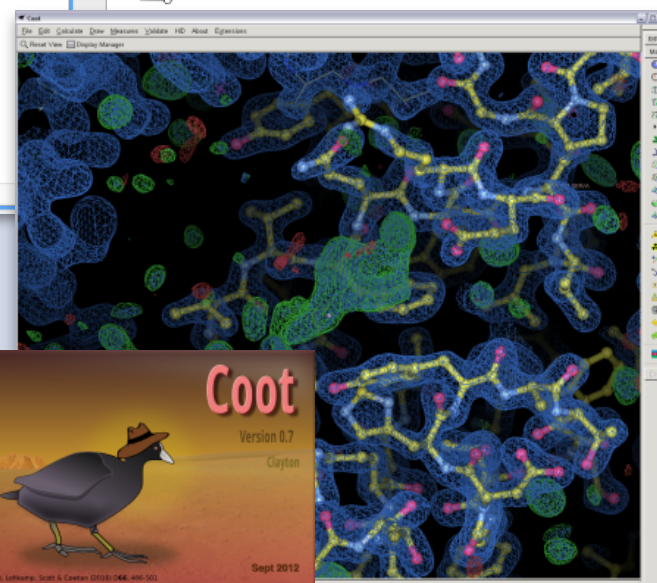
Job list Project directory

Job/File Evaluation

- 5 Molrep
- 4 Refinement - REFMAC5 $R=0.49$ $R_{free}=0.58$
- 3 Molrep
- 2 Import a coordinate set - optional se... $nRes=168$
- 1 Import merged



- Import merged data, sequences, alignment of coordinates
- Integrate X-ray images
- X-ray data reduction and analysis
- Experimental phasing
- Bioinformatics and model preparation
- Molecular Replacement
- Model building and Graphics
- Refinement
- Ligands
- Validation and analysis
- Export and Deposition
- Reflection data tools



PDB_REDO

Crystallisation

Data Collection

Data Processing
and Reduction

Experimental
Phasing

Molecular
Replacement

Density
Modification

Model Building

Refinement

Structure Analysis

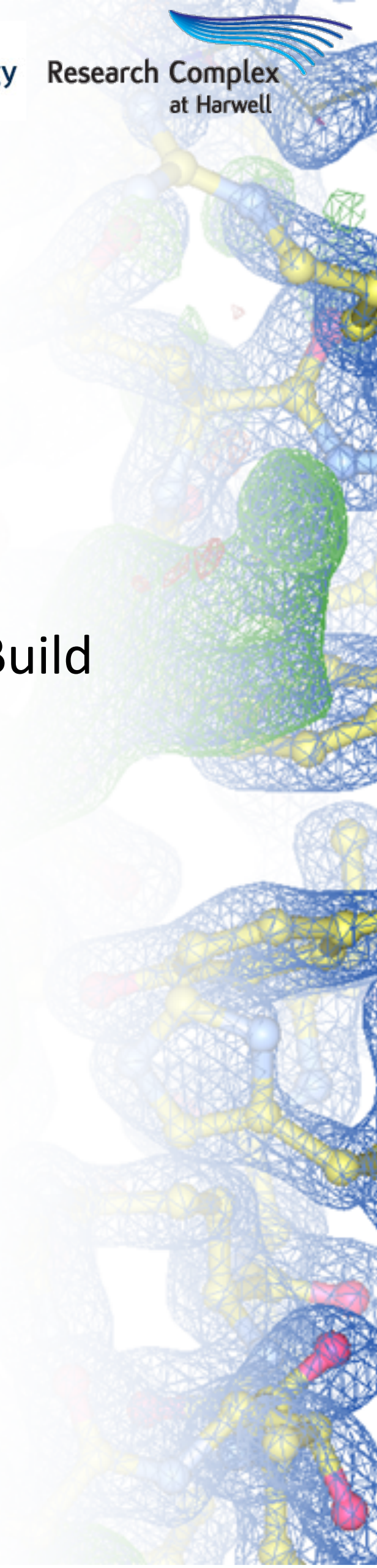
Deposition

SHELX

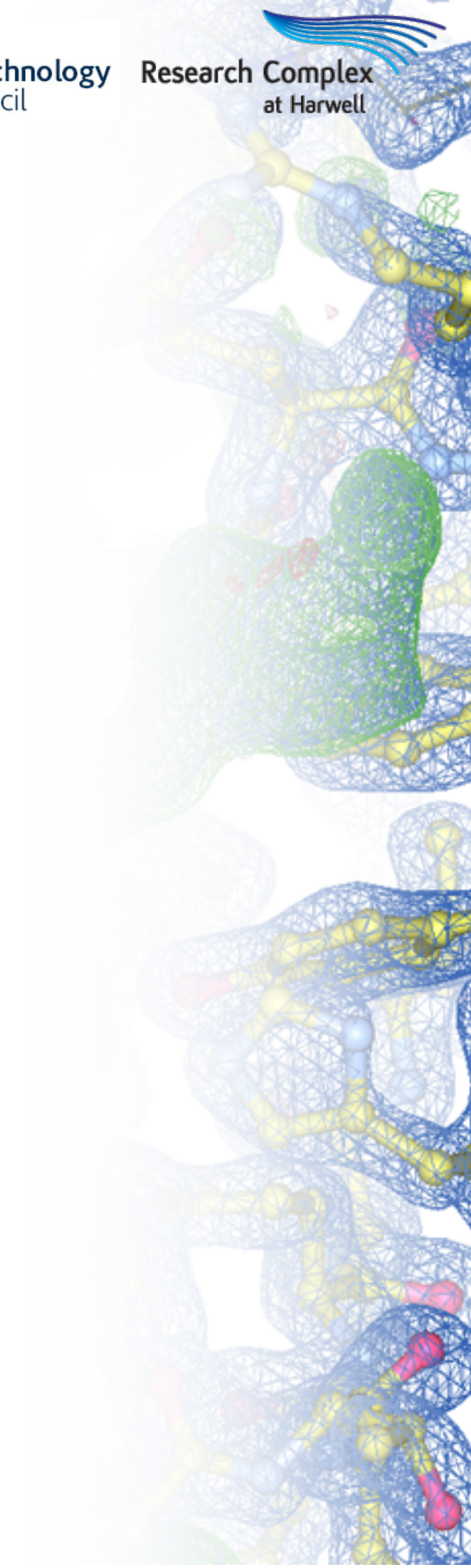
7.3
ARP
wARP

CCP4 Releases

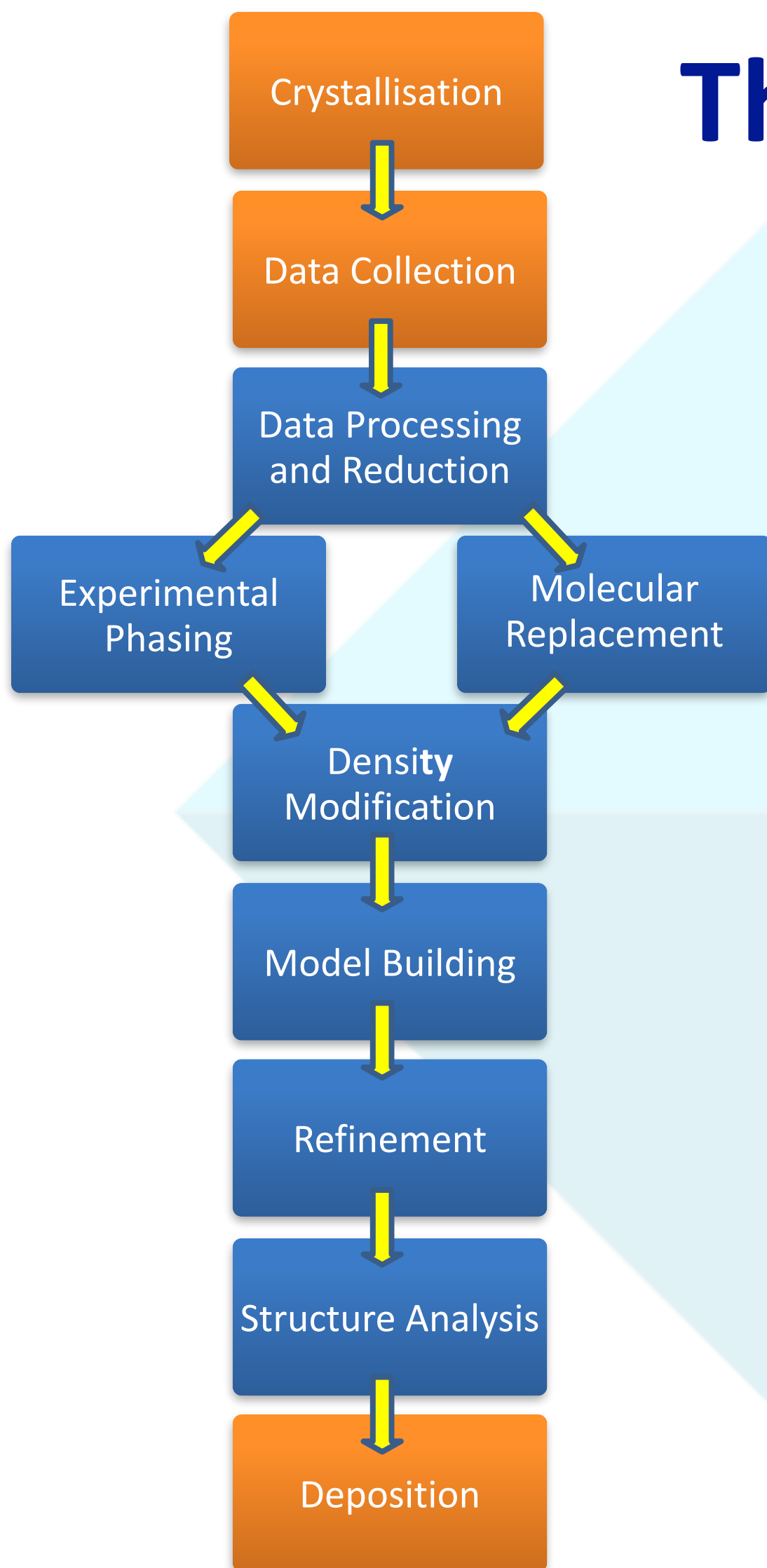
- Current release version — CCP4 7.1.008
- Recent programs:
 - ~ SHELX, DIALS, MorDA, Arcimboldo, CCP4i2, CCP4 Cloud, CCP4Build
 - ~ QtPISA, Aimless, Dimple, AMPLE, SIMBAD, Zanuda, ProSMART, Nautilus, AceDRG, Privateer, Feckless and other
 - ~ In collaboration with the EMBL-Hamburg, ARP/wARP model building software is now rolled with CCP4 Suite
- Continuing development of major components:
 - ~ Phaser, Refmac, Mosflm/iMosflm, Crank-2, MrBUMP
- Newer infrastructure
 - ~ Build system, Setup Manager and Updates Manager



Using CCP4



The Structure Solution Stages



Automated*

Xia-2

MrBump, Balbes, MorDA,
Crank-2, Simbad, Ample

Buccaneer, CCP4Build,
Arp/wArp

Lorestr, Dimple

PDB Deposition module

Fundamental*

XDS, DIALS, Aimless,
DUI, XDSCui, iMosflm

Phaser, Molrep, ShelxCD,
Sculptor, Chainsaw

Parrot, Acorn, ShelxE

Coot, AceDrg, jLigand

Refmac, Prosmart

Zanuda, PISA, Gesamt

* only selected representative programs are shown

Programs' use and Projects

- Crystallographic projects may be very complex
 - ~ over 100 programs may need to be used
 - ~ sometimes, 100s jobs may need to be run
- All CCP4 programs can be used from command line
 - ~ [+] gives full control of programs and access to all parameters
 - ~ [-] difficult to memorise all options and parameters
 - ~ [-] difficult to keep 100s and 1000s files organised
 - ~ [-] may be difficult to analyse the structure solution pathway *post factum*
- Most frequently used programs are available through graphical interfaces
 - ~ [+] greatly simplifies manipulation with data and program parameters
 - ~ [+] arranges work in **Projects**, easier to navigate, analyse and make decisions
 - ~ [+] take care of many petty jobs such as routine format conversions, map preparations and other
 - ~ [-] may give access only to a subset of most frequently used program parameters
 - ~ [-] may not exist for a particular program of interest

Command Line Operation

Example:

refmac5 HKLIN my.mtz HKLOUT output.mtz XYZIN my.pdb XYZOUT output.pdb

Keywords:

- Most CCP4 programs take “keywords” as command options through standard input to the program
- Gives access to all options of the program and the most fine-grained control
- Can be provided through ccp4i using “**Run & view com file**”

Scripting:

- Repetitive tasks can be scripted
- Pipelines of programs can be created in a script

Command
line

```
f2mtz hklin ${1} hklout shelxe-output.mtz << eof
```

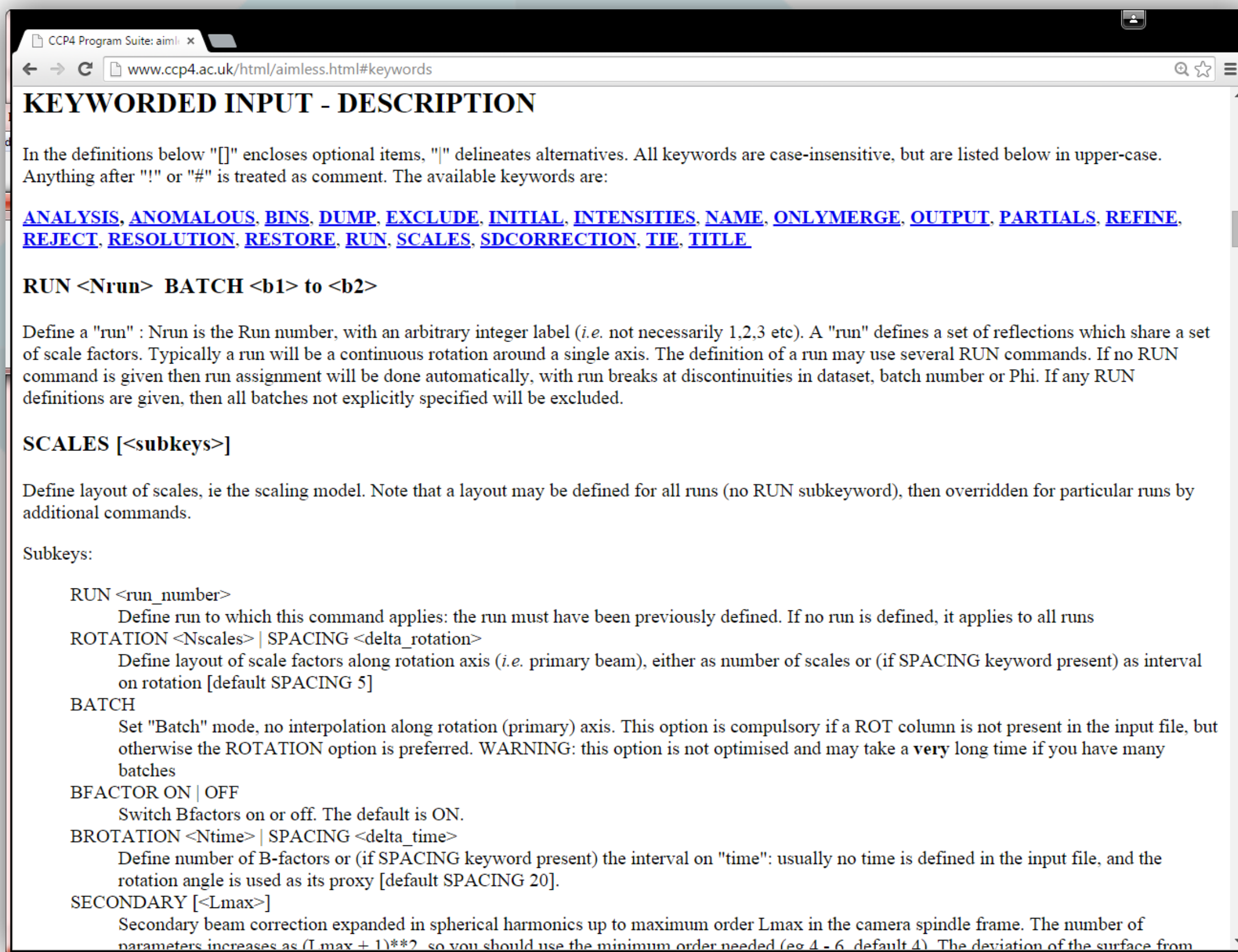
```

TITLE      ample
cell       ${cell}
symm       ${symm}
labout     H K L F FOM PHI SIGF
CTYPOUT    H H H F W P Q
pname      ample
dname      ample
END
```

Keywords

Command Line Operation

- Documentation for CCP4 program keywords provided by documentation pages online or in the suite (`$CCP4/html/INDEX.html`):



CCP4 Program Suite: aiml x

www.ccp4.ac.uk/html/aimless.html#keywords

KEYWORDED INPUT - DESCRIPTION

In the definitions below "[]" encloses optional items, "|" delineates alternatives. All keywords are case-insensitive, but are listed below in upper-case. Anything after "!" or "#" is treated as comment. The available keywords are:

[ANALYSIS](#), [ANOMALOUS](#), [BINS](#), [DUMP](#), [EXCLUDE](#), [INITIAL](#), [INTENSITIES](#), [NAME](#), [ONLYMERGE](#), [OUTPUT](#), [PARTIALS](#), [REFINE](#), [REJECT](#), [RESOLUTION](#), [RESTORE](#), [RUN](#), [SCALES](#), [SDCORRECTION](#), [TIE](#), [TITLE](#)

RUN <Nrun> BATCH <b1> to <b2>

Define a "run" : Nrun is the Run number, with an arbitrary integer label (*i.e.* not necessarily 1,2,3 etc). A "run" defines a set of reflections which share a set of scale factors. Typically a run will be a continuous rotation around a single axis. The definition of a run may use several RUN commands. If no RUN command is given then run assignment will be done automatically, with run breaks at discontinuities in dataset, batch number or Phi. If any RUN definitions are given, then all batches not explicitly specified will be excluded.

SCALES [<subkeys>]

Define layout of scales, ie the scaling model. Note that a layout may be defined for all runs (no RUN subkeyword), then overridden for particular runs by additional commands.

Subkeys:

RUN <run_number>
Define run to which this command applies: the run must have been previously defined. If no run is defined, it applies to all runs

ROTATION <Nscale> | SPACING <delta_rotation>
Define layout of scale factors along rotation axis (*i.e.* primary beam), either as number of scales or (if SPACING keyword present) as interval on rotation [default SPACING 5]

BATCH
Set "Batch" mode, no interpolation along rotation (primary) axis. This option is compulsory if a ROT column is not present in the input file, but otherwise the ROTATION option is preferred. WARNING: this option is not optimised and may take a very long time if you have many batches

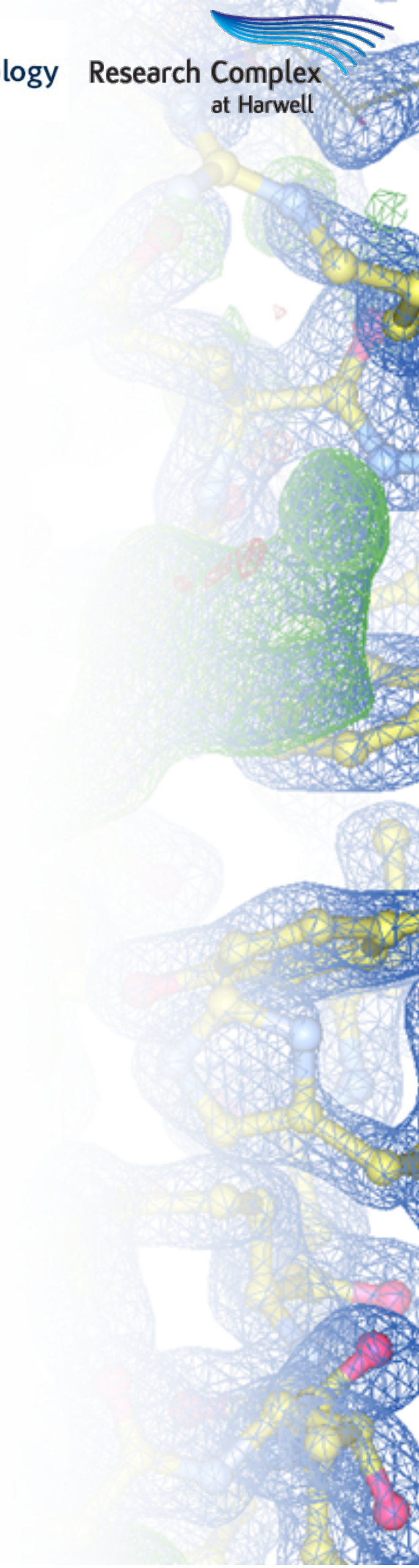
BFACTOR ON | OFF
Switch Bfactors on or off. The default is ON.

BROTATION <Ntime> | SPACING <delta_time>
Define number of B-factors or (if SPACING keyword present) the interval on "time": usually no time is defined in the input file, and the rotation angle is used as its proxy [default SPACING 20].

SECONDARY [<Lmax>]
Secondary beam correction expanded in spherical harmonics up to maximum order Lmax in the camera spindle frame. The number of parameters increases as $(L_{max} + 1) * 2$, so you should use the minimum order needed (eg 4 - 6, default 4). The deviation of the surface from

The CCP4 user interface(s)

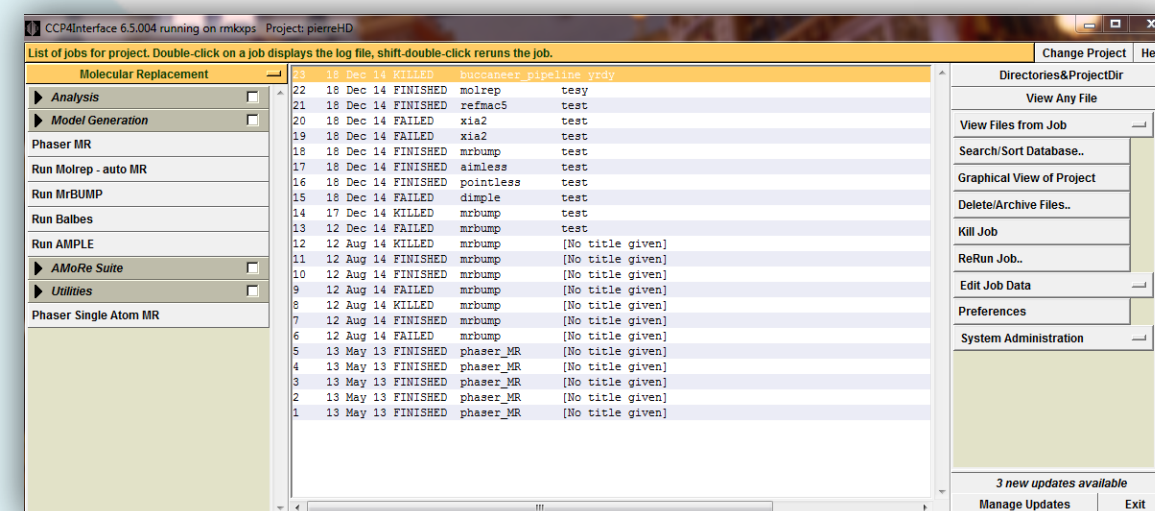
- Three main interfaces to the CCP4 suite



The CCP4 user interface(s)

- Three main interfaces to the CCP4 suite

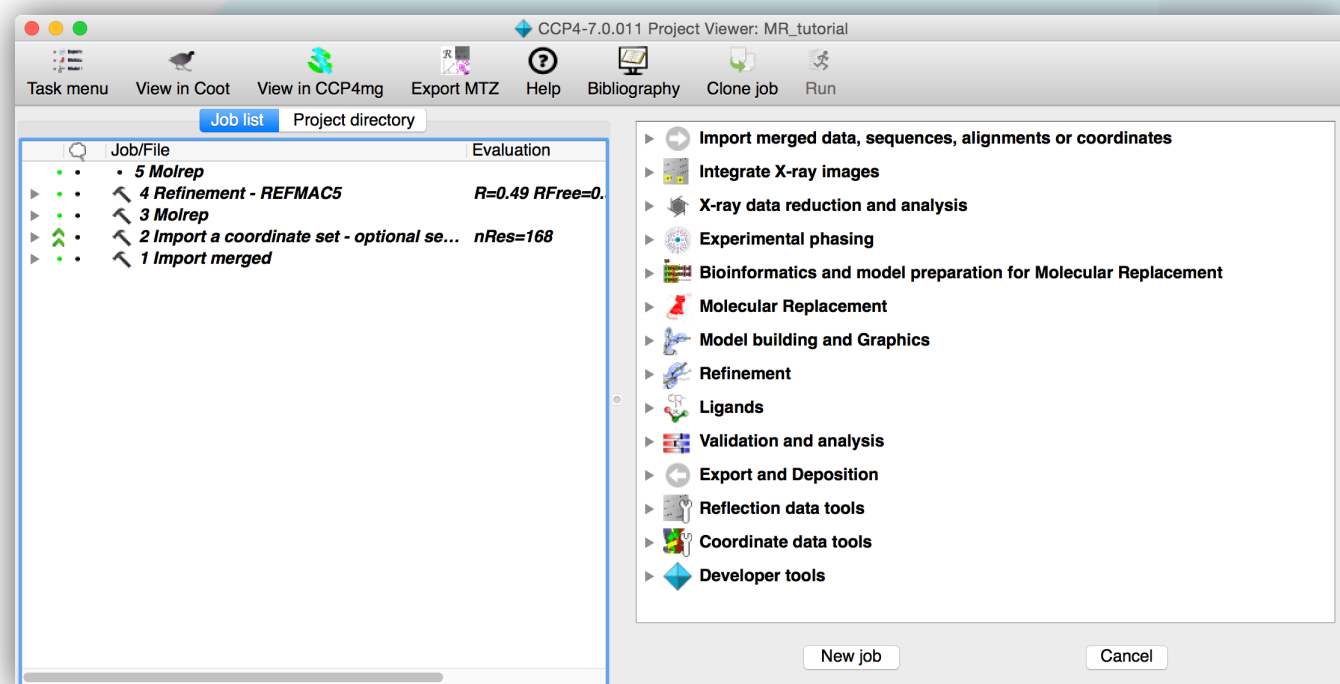
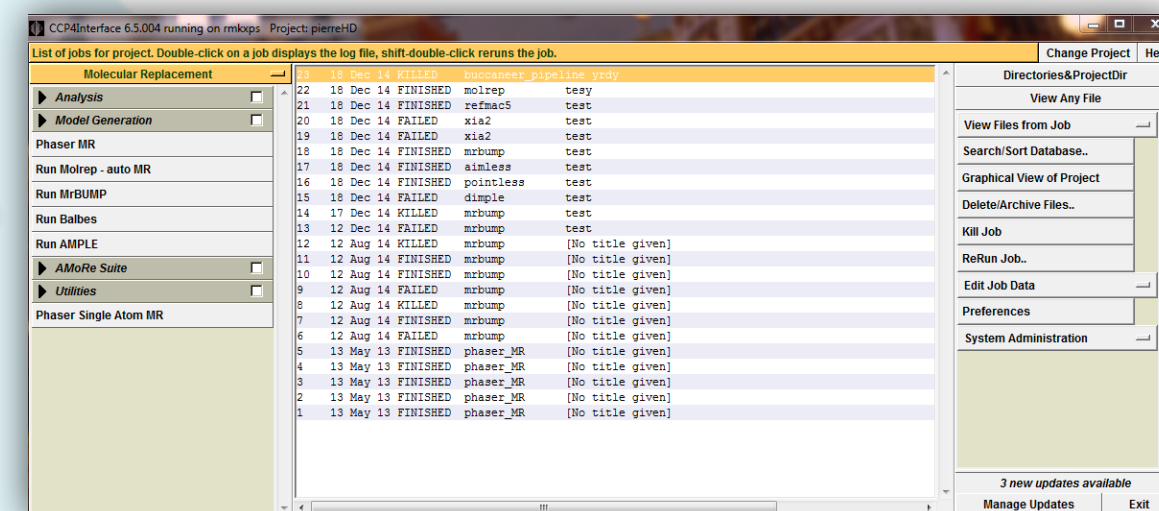
1. **CCP4i** – original interface developed around 2000



The CCP4 user interface(s)

- Three main interfaces to the CCP4 suite

1. **CCP4i** – original interface developed around 2000

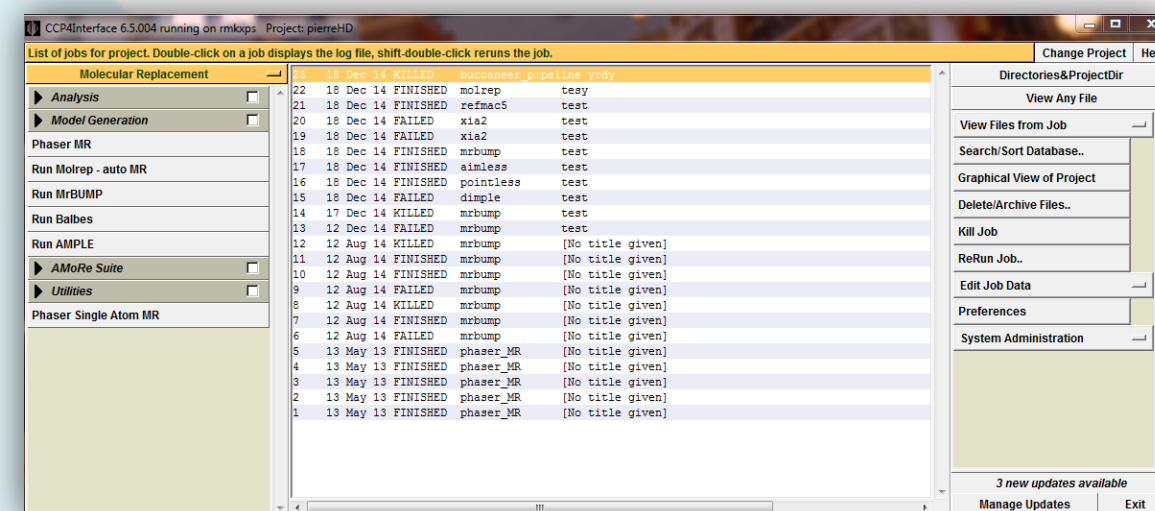


2. **CCP4i2** – new graphical desktop interface (2016)

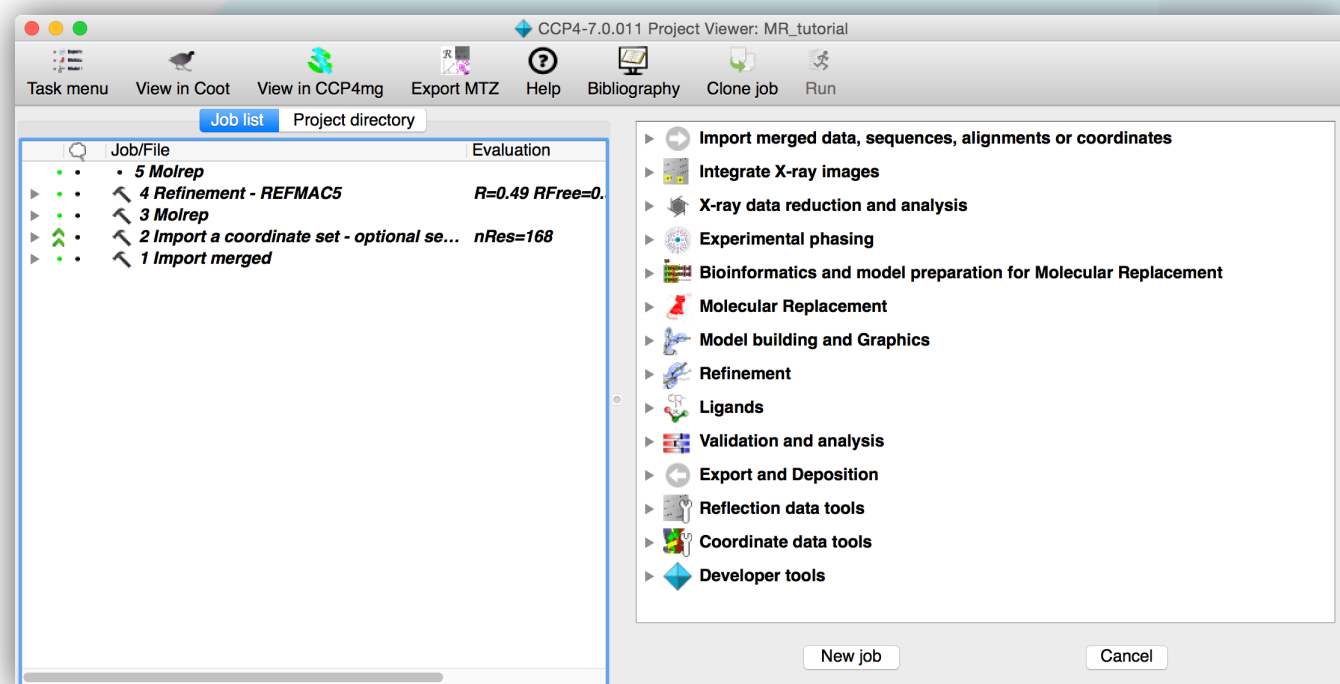
The CCP4 user interface(s)

Three main interfaces to the CCP4 suite

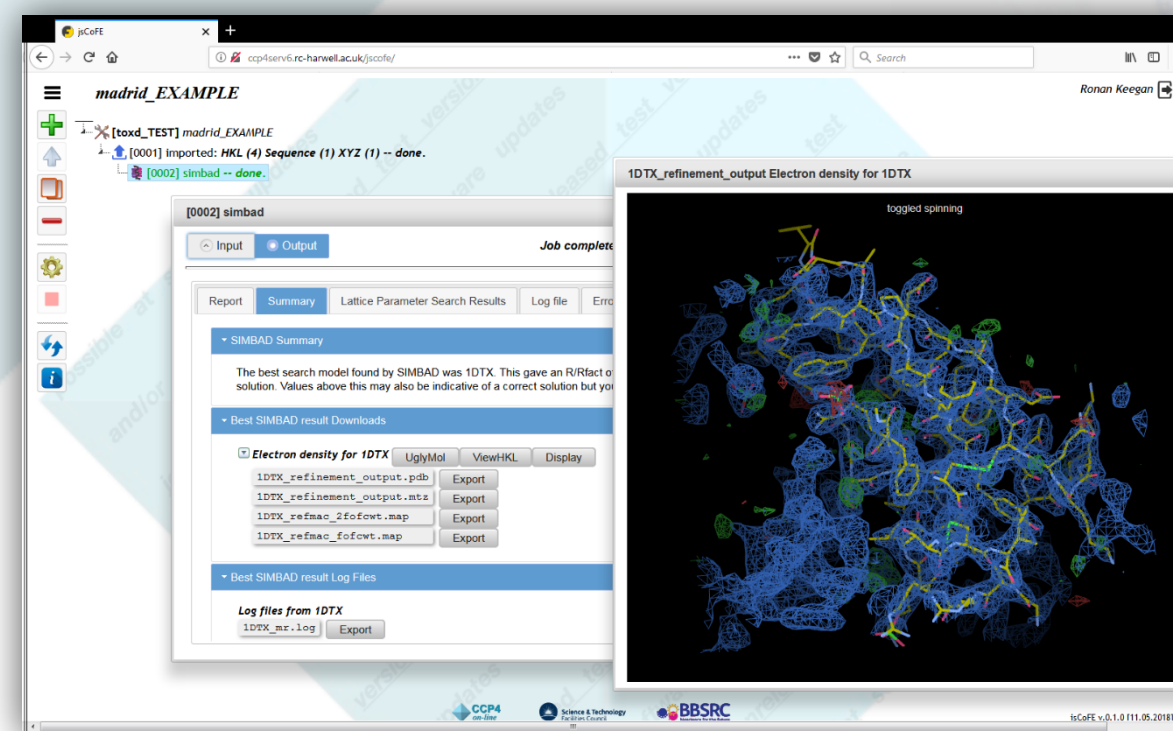
1. CCP4i – original interface developed around 2000



2. CCP4i2 – new graphical desktop interface (2016)



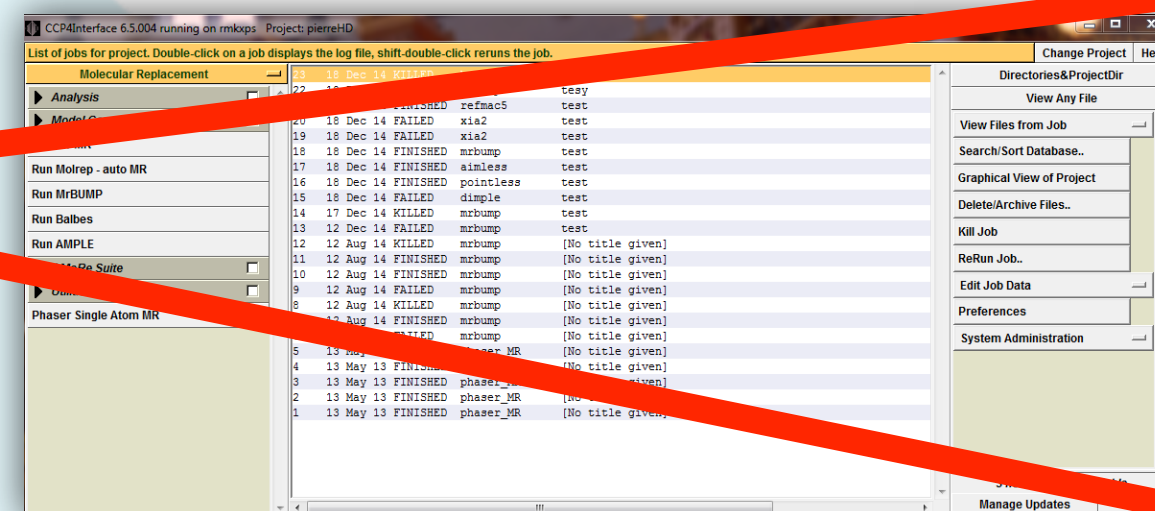
3. CCP4 Cloud, browser based (2019)



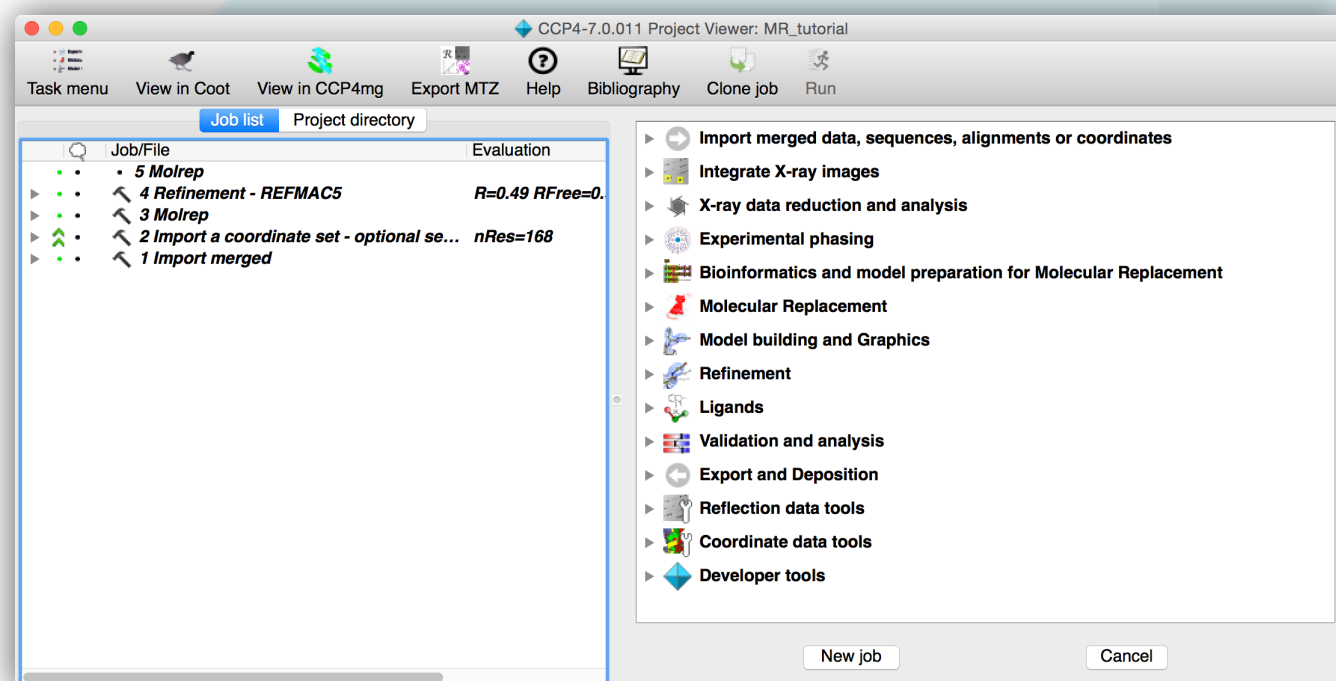
The CCP4 user interface(s)

- Three main interfaces to the CCP4 suite

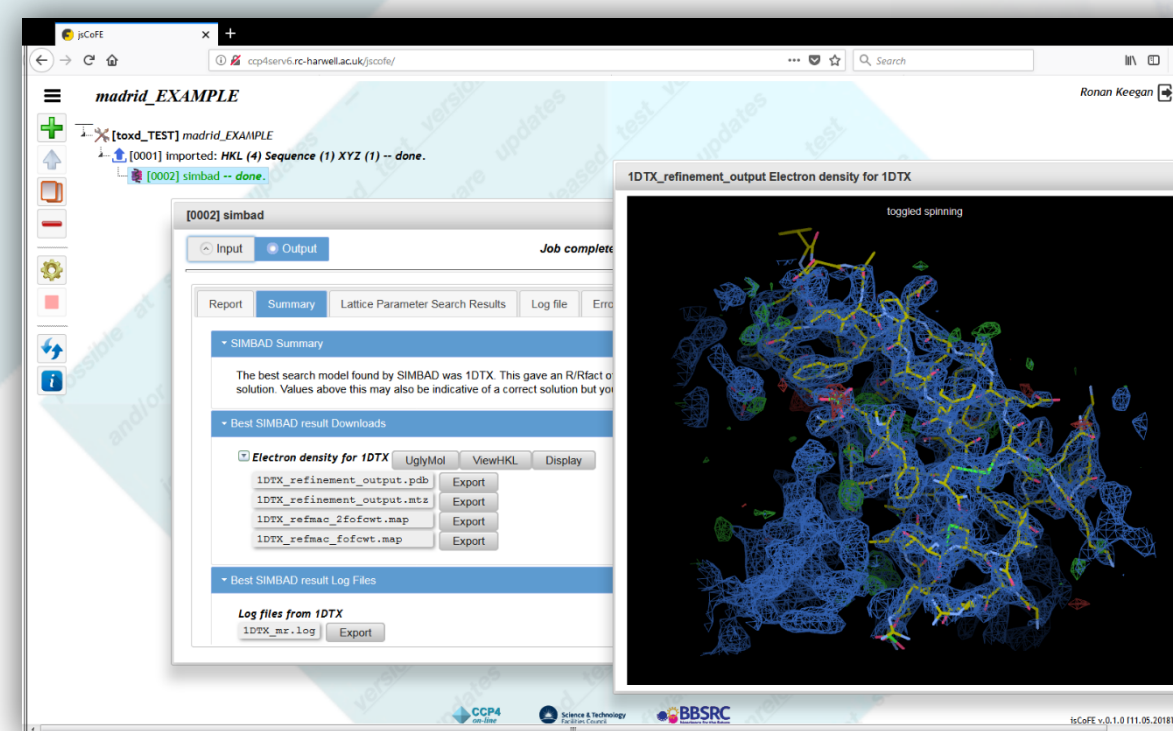
~~1. CCP4i – original interface
developed around 2000
deprecated~~



2. CCP4i2 – new graphical
desktop interface (2016)



3. CCP4 Cloud, browser
based (2019)



CCP4 on-line services

<http://www.ccp4.ac.uk/ccp4online>

Welcome to CCP4 online

Login

Other Options - [Register](#), [Forgot](#)

Runnable programs

The following programs and pipelines are available:

Balbes

An automated Molecular Replacement pipeline that searches for the components necessary for a successful search.

MrBUMP

An automated Molecular Replacement pipeline that searches for the structure factors, it will search for the template structure and restrained refinement.

Zanuda

Space group and crystallographic data.

jsPISA

Calculation and analysis of protein-protein interfaces.

AMPLE

Automated ab initio search.

SHELX

Automated SHELXC/D/E steps. Data in XDS, Scalepack, SHARP. Protein sequence is provided.

CRANK2

Automated structure solution.

BETA

CCP4 online

Collaborative Computational Project No. 4
Software for Macromolecular X-Ray Crystallography

Welcome to CCP4 online

Login

Other Options - [Register](#), [Forgot](#)

Runnable programs

The following programs and pipelines are available:

Balbes
An automated Molecular Replacement pipeline that searches for the components necessary for a successful search.

MrBUMP
An automated Molecular Replacement pipeline that searches for the structure factors, it will search for the template structure and restrained refinement.

Zanuda
Space group and crystallographic data.

jsPISA
Calculation and analysis of protein-protein interfaces.

AMPLE
Automated ab initio search.

SHELX
Automated SHELXC/D/E steps. Data in XDS, Scalepack, SHARP. Protein sequence is provided.

CRANK2
Automated structure solution.

BETA

CCP4 STFC Rutherford Appleton Laboratory, Harwell Campus

CCP4 online

Collaborative Computational Project No. 4
Software for Macromolecular X-Ray Crystallography

Welcome to CCP4 online

Login

Other Options - [Register](#), [Forgot](#)

Runnable programs

The following programs and pipelines are available:

Balbes
An automated Molecular Replacement pipeline that searches for the components necessary for a successful search.

MrBUMP
An automated Molecular Replacement pipeline that searches for the structure factors, it will search for the template structure and restrained refinement.

Zanuda
Space group and crystallographic data.

jsPISA
Calculation and analysis of protein-protein interfaces.

AMPLE
Automated ab initio search.

SHELX
Automated SHELXC/D/E steps. Data in XDS, Scalepack, SHARP. Protein sequence is provided.

CRANK2
Automated structure solution.

BETA

CCP4 STFC Rutherford Appleton Laboratory, Harwell Campus

jsPISA 2.0.3 [PDB 1e94]



Input Monomers Interfaces Stock Crystal Splits Log file

Interfaces

- List of interfaces
- 1. F || E
 - Bonds
 - Residues
 - 2. E || F
 - 3. E || E
 - 4. A || B
 - 5. B || A
 - 6. D || C
 - 7. D || C
 - 8. B || A
 - 9. D || C
 - 10. C || C
 - 11. A || A
 - 12. B || B
 - 13. D || D
 - 14. [ANP]E:500 || E
 - 15. [ANP]F:501 || F
 - 16. F || B
 - 17. E || A
 - 18. C || F
 - 19. E || A
 - 20. [ANP]E:500 || F
 - 21. E || [ANP]F:501
 - 22. F || E
 - 23. E || E
 - 24. F || F

Interface F || E

Summary

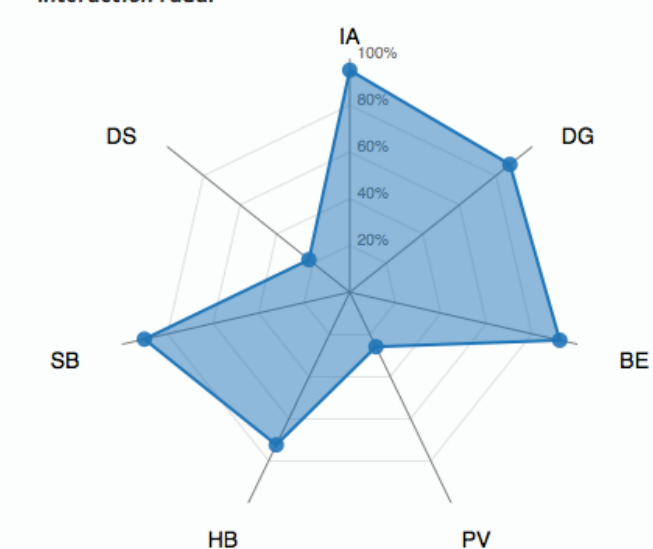
	Monomer 1		Monomer 2	
Monomer ID	F		E	
Class	Protein		Protein	
Symmetry operation	X, Y, Z		X, Y, Z	
Symmetry ID	1_555		1_555	
Interface atoms	232	7.3%	221	6.9%
Surface atoms	1925	60.5%	1921	60.4%
Total atoms	3184	100.0%	3183	100.0%
Interface residues	66	16.1%	65	15.9%
Surface residues	380	92.9%	385	94.4%
Total residues	409	100.0%	408	100.0%
BSA, Å ²	2146.6	9.6%	2345.5	10.6%
ASA, Å ²	22253.4	100.0%	22072.6	100.0%
Solvation energy, kcal/mol	-331.9		-334.3	
SE gain, kcal/mol	-7.5		-5.0	

Download PDB

Download PDBx

View in JSMol

Interaction radar



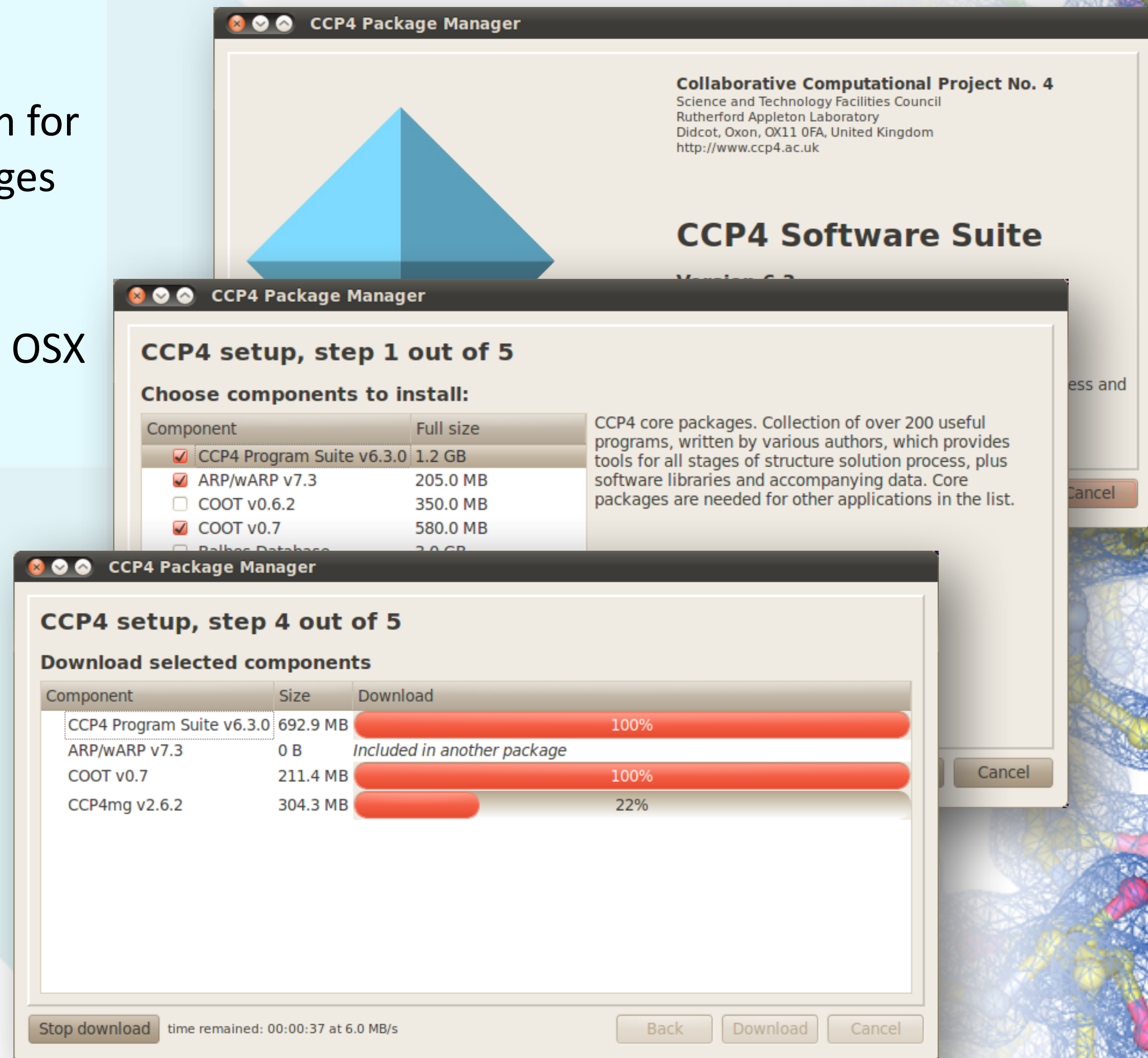
Interface parameters

IA : Interface area, Å ²	2246
DG : Solvation Energy, kcal/mol	-12.46
BE : Total Binding Energy, kcal/mol	-26.65
PV : Hydrophobic P-value	0.3542
HB : Number of Hydrogen Bonds	11
SB : Number of Salt Bridges	25
DS : Number of Disulphide Bonds	0



CCP4 Setup Manager

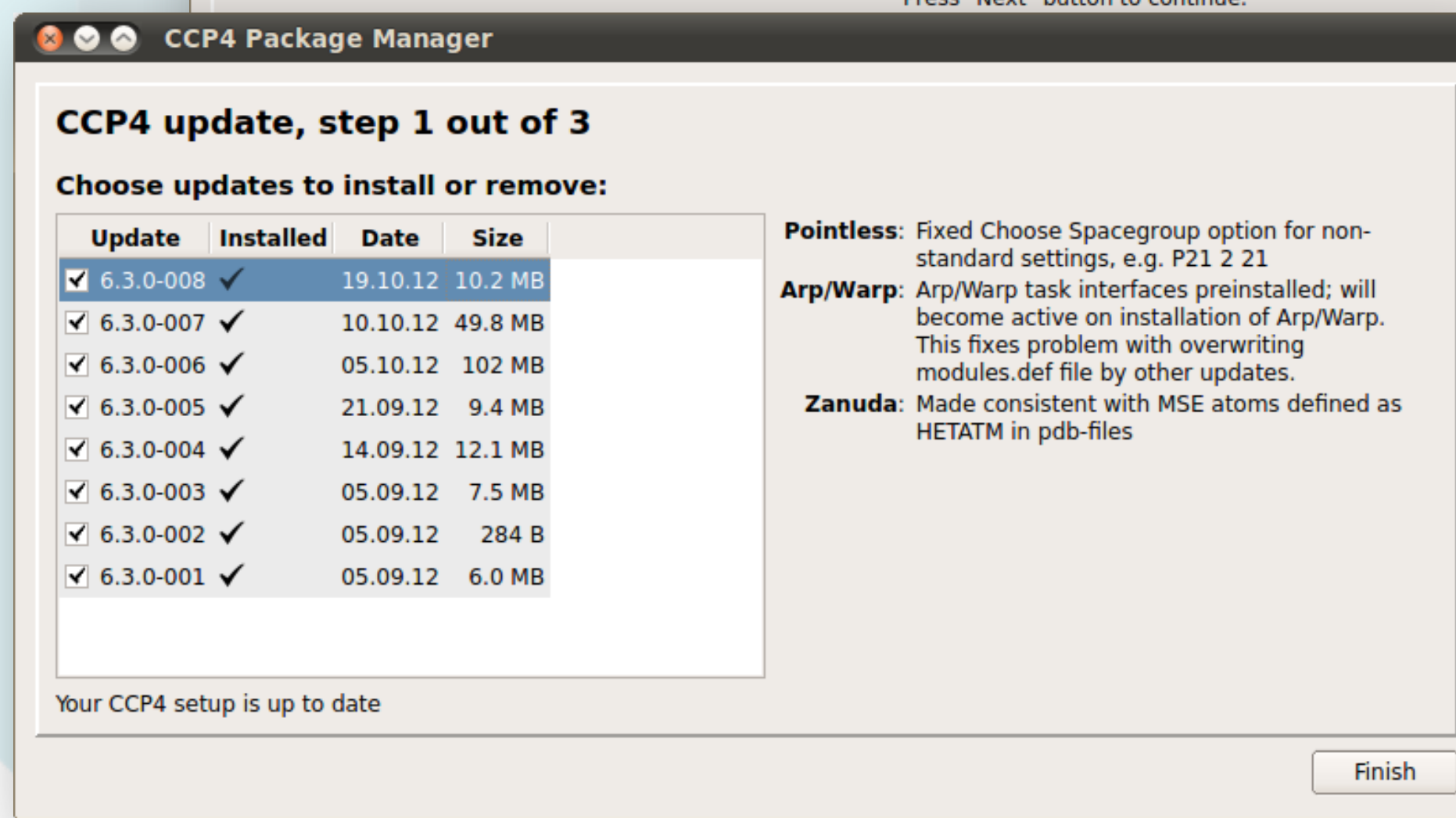
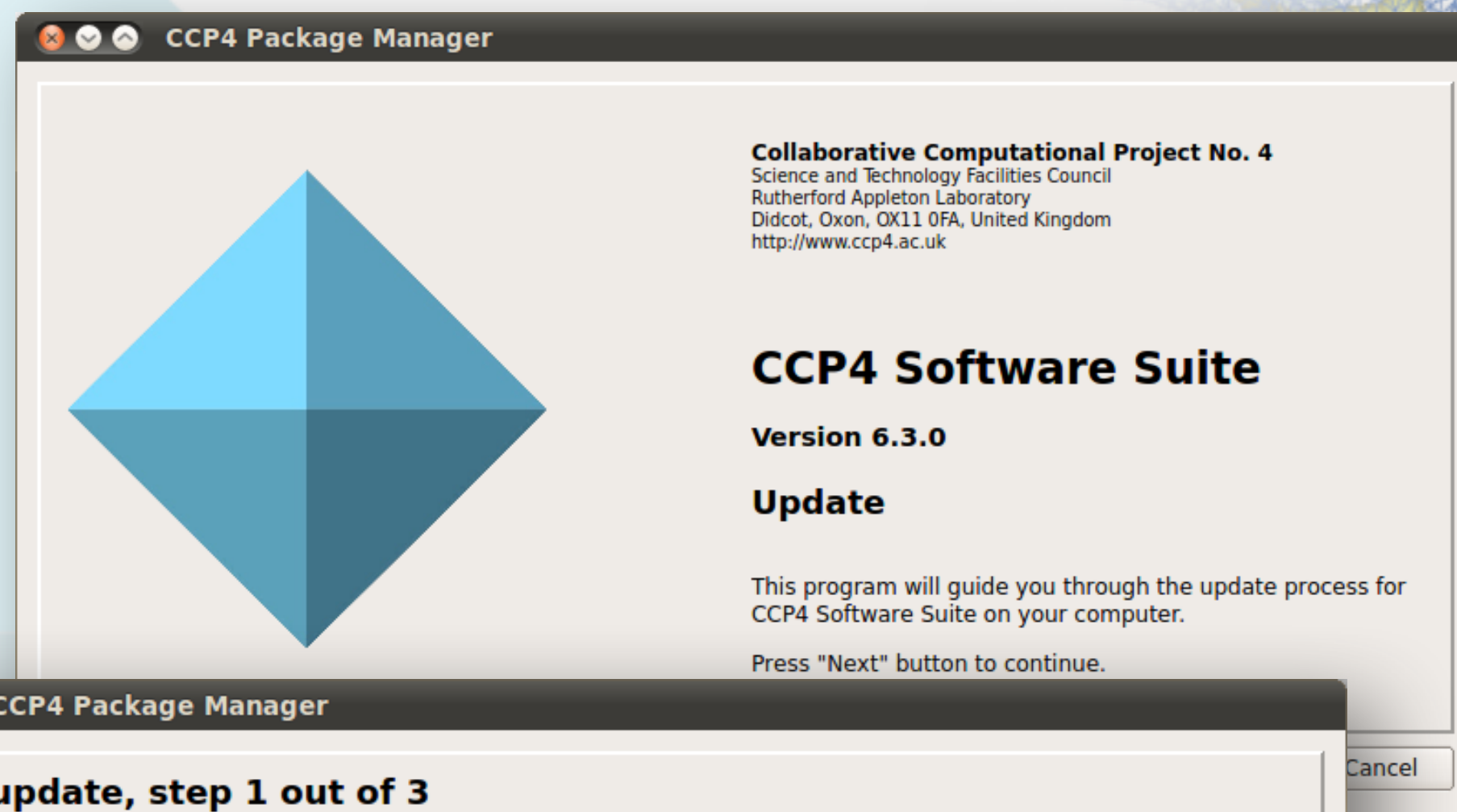
- Easy installation mechanism for CCP4 and associated packages (Coot, ARP/wARP, CCP4mg)
- Available on Linux and Mac OSX platforms
- Sets up system variables automatically if you wish





Update Manager

- Automatically install latest fixes and new features
- Allows for rolling back to previous versions
- Updates made available once every two weeks or as and when they are needed
- Available on Linux, Mac OSX and Windows



Acknowledgements

CCP4 Core group: Andrey Lebedev, Eugene Krissinel, Charles Ballard, David Waterman, Marcin Wojdyr, Ville Uski, Karen McIntyre, Oleg Kovalevskyi, Kyle Stevenson, Martyn Winn

LMB/MRC: Andrew Leslie, Phil Evans, Garib Murshudov, Rob Nicholls, Harry Powell, Owen Johnson, Fei Long, Paul Emsley, Andrea Thorn

Phaser Group: Airlie McCoy, Randy Read, Rob Oeffner, Gabor Bunkoczi

YSBL York: Keith Wilson, Kevin Cowtan, Liz Potterton, Stuart McNicholas, Eleanor Dodson

Diamond: Gwyndaf Evans, Graeme Winter, Luis Fuentes-Montero

Leiden: Raj Pannu, Pavol Skubak

Others: Bernhard Lohkamp, Clemens Vonrhein, Ruslan Sanishvili, Frank Von Delft, Martin Noble, Jaclyn Bibby, Daniel Rigden, Jens Thomas, Alun Ashton, David Brown, Arwen Pearson, Tim Gruene, George Sheldrick and many more ...

CCP4 People (core and associated developers)



Martyn
Winn



Charles
Ballard



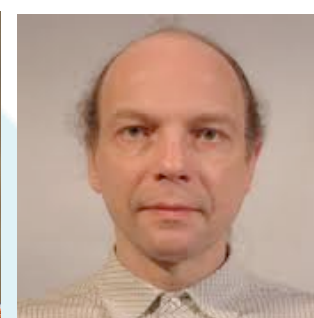
Eugene
Krissinel



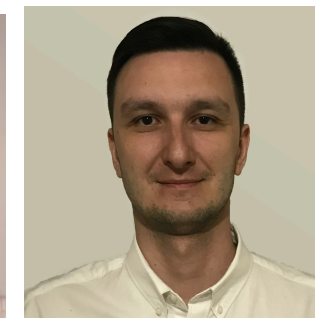
Ronan
Keegan



David
Waterman



Andrey
Lebedev



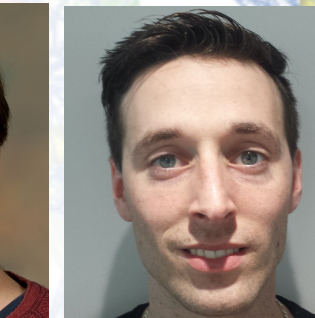
Oleg
Kovalevskyi



Kyle
Stevenson



Ville
Uski



Tarik
Drevon



Keith
Wilson



Eleanor
Dodson



Liz
Potterton



Stuart
McNicholas



Kevin
Cowtan



Garib
Murshudov



Alexei
Vagin



Fei
Long



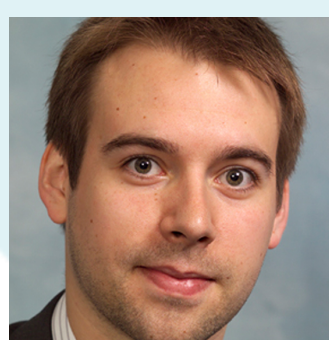
Hew
Jenkins



Adam
Simpkin



Pavol
Skubak



Robert
Nichols



Daniel
Rigden



Graeme
Winter



Gwyndaf
Evans



Phil
Evans



Andrew
Leslie



Harry
Powell



Owen
Johnson



Randy
Reed



Martin
Noble



Paul
Emsley



Raj
Pannu



Chris
Morris