BLEND

Complete data sets from multiple crystals

James Foadi, Pierre Aller, Gwyndaf Evans

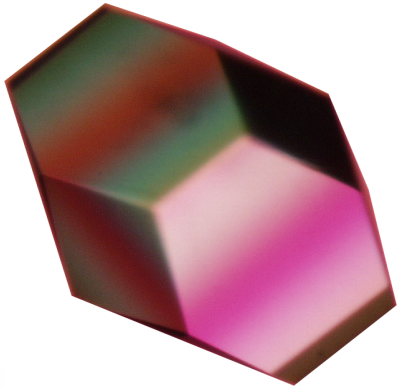
Membrane Protein Laboratory (MPL)
Imperial College London
Diamond Light Source Ltd

Imperial College London

MPL Membrane Protein Laboratory

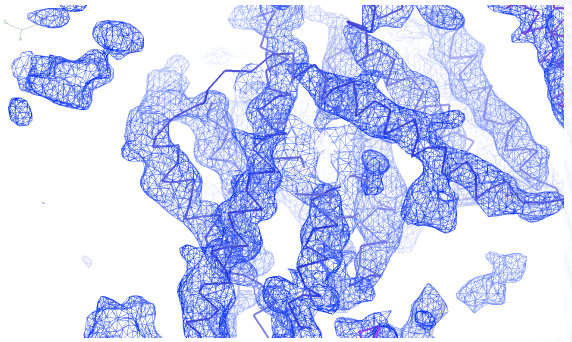
diamond

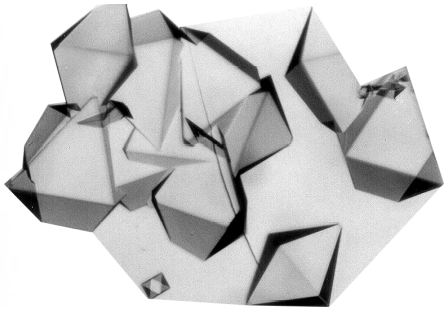
This talk in a nutshell



Single Crystal

Established Methods and Software

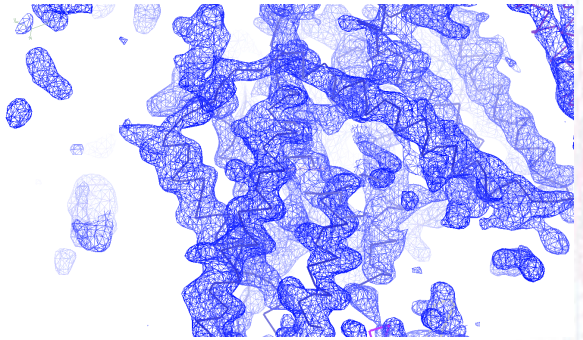




Multiple Crystals

New Methods and Software

BLEND



Imperial College London

MPL Membrane Protein Laboratory

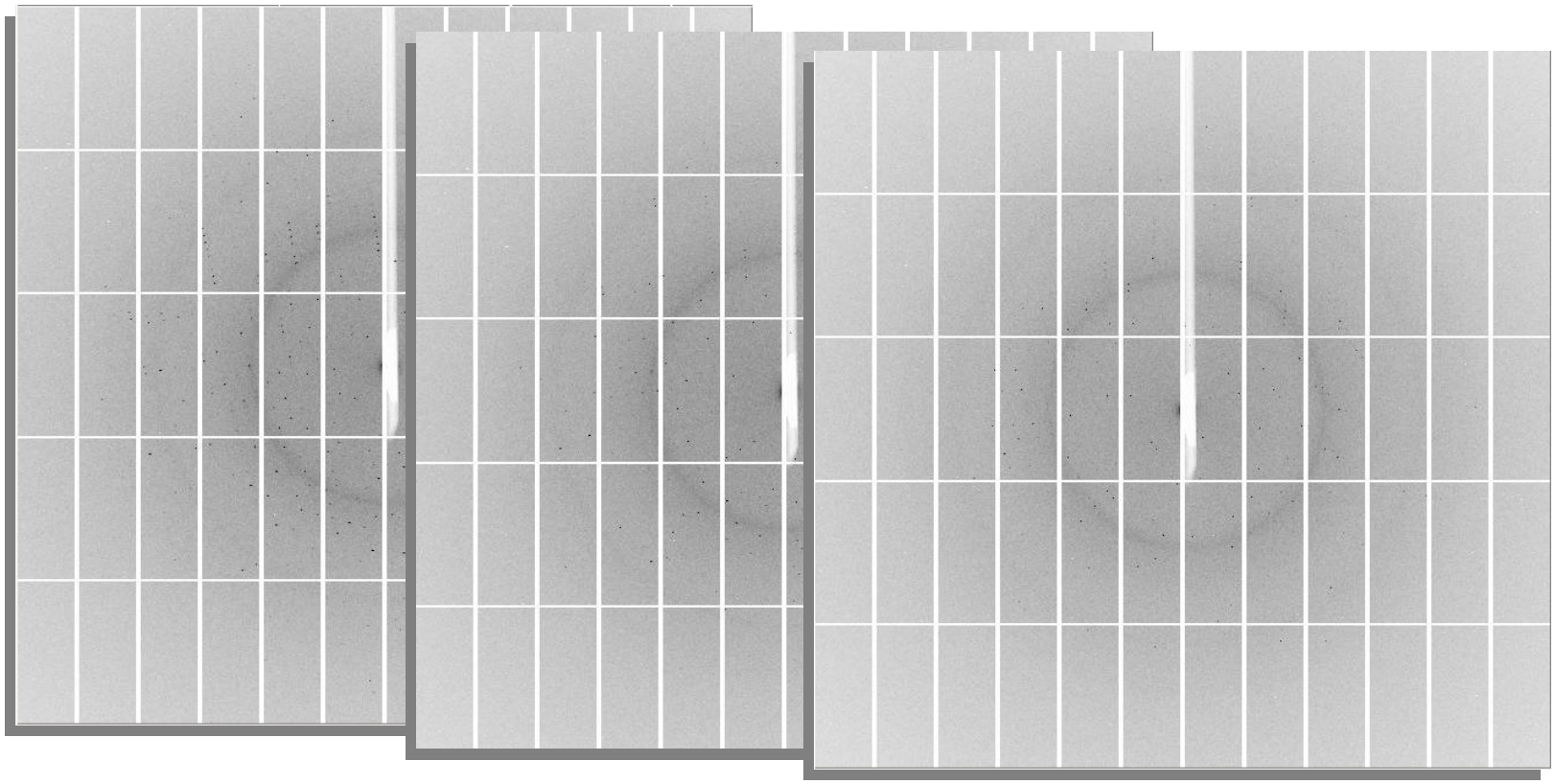
diamond

Outline

- (some) Reasons to use multiple crystals
- BLEND
- Established results
- Applications
- Slide on BLEND tutorials

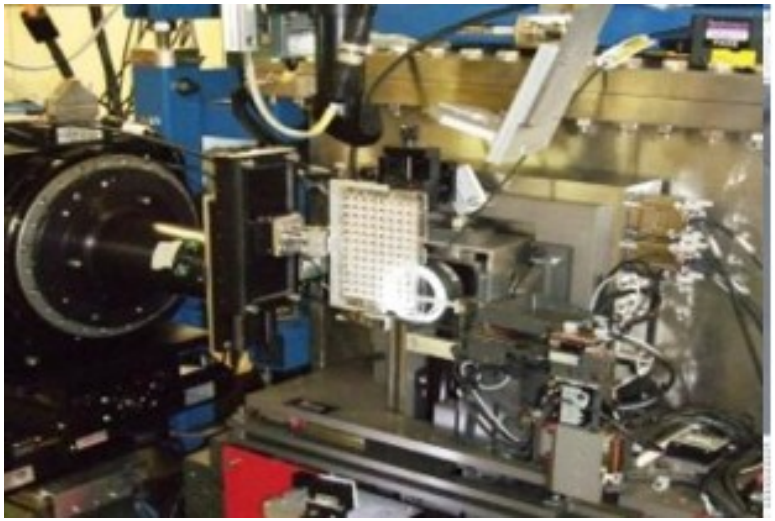
(some) Reasons to use multiple crystals

To reduce radiation damage



(some) Reasons to use multiple crystals

Constraints in data collection (*in situ*)

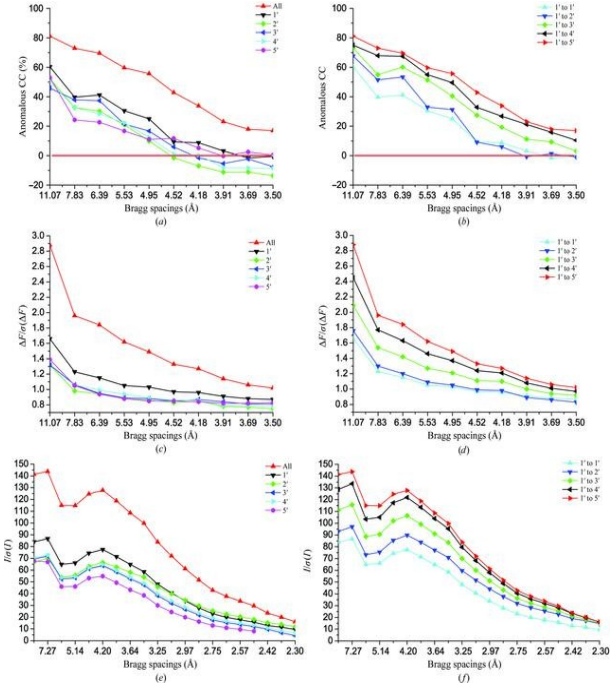


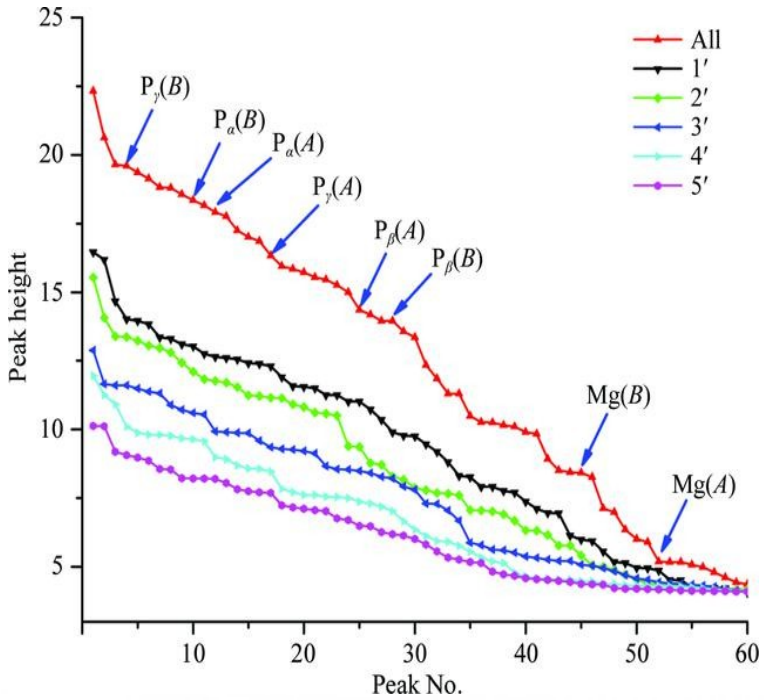
Diamond
I24 beamline

(some) Reasons to use multiple crystals

To increase the anomalous signal

Q. Liu, Q. Liu and W. A. Hendrickson (2013), Acta Cryst. D69, 1314-1332





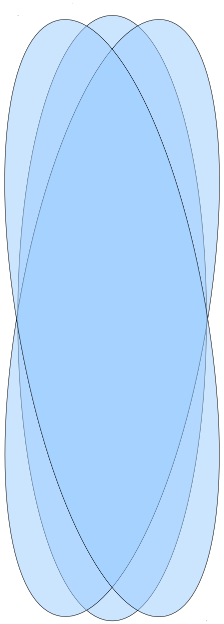
Imperial College London

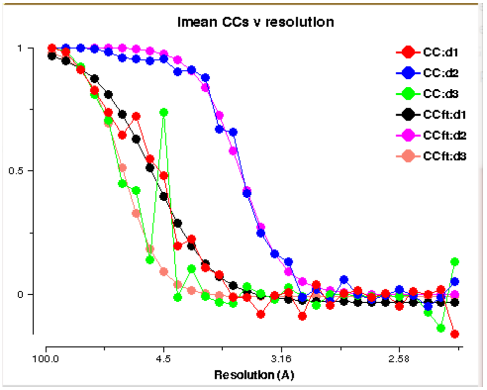
MPL Membrane Protein Laboratory

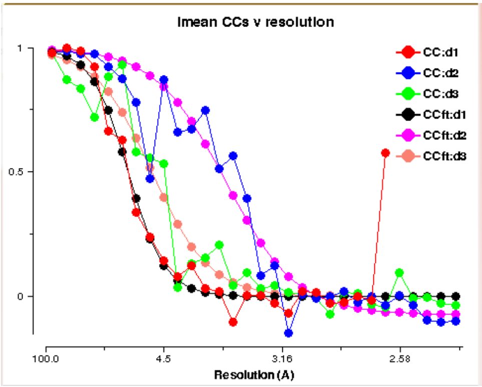
diamond

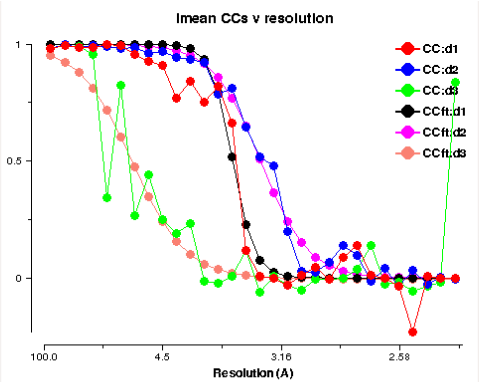
(some) Reasons to use multiple crystals

Slight improvement of anisotropic data









BLEND

Main reference:

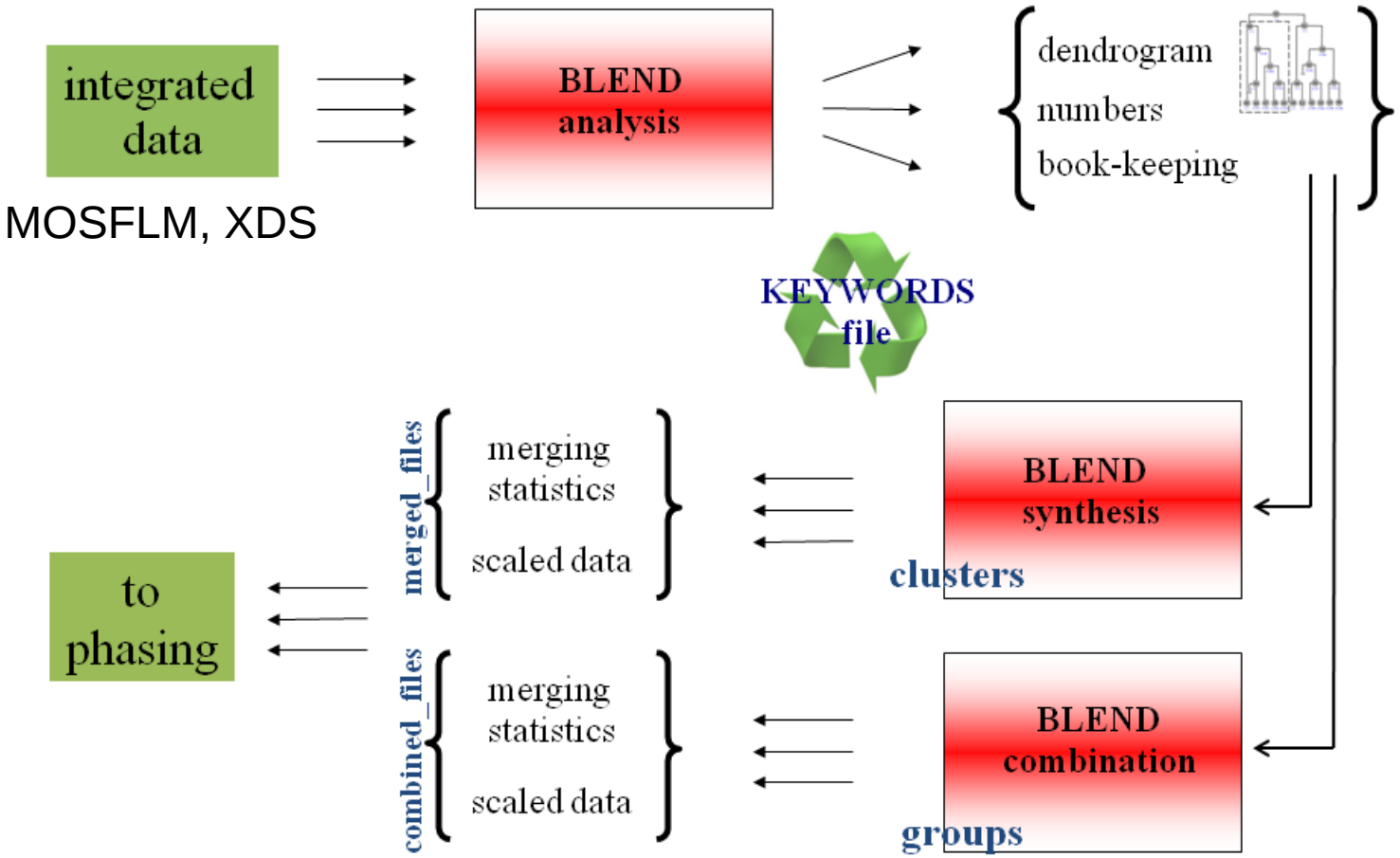
J. Foadi, P. Aller, Y. Alguel, A. Cameron, D. Axford,
R. L. Owen, W. Armour, D. G. Waterman, S. Iwata and G. Evans
Clustering procedures for the optimal selection of data sets from multiple crystals in macromolecular crystallography
Acta Cryst. (2013), D**69**, 1617-1632

See also:

Slides from various presentations
(web-search them)

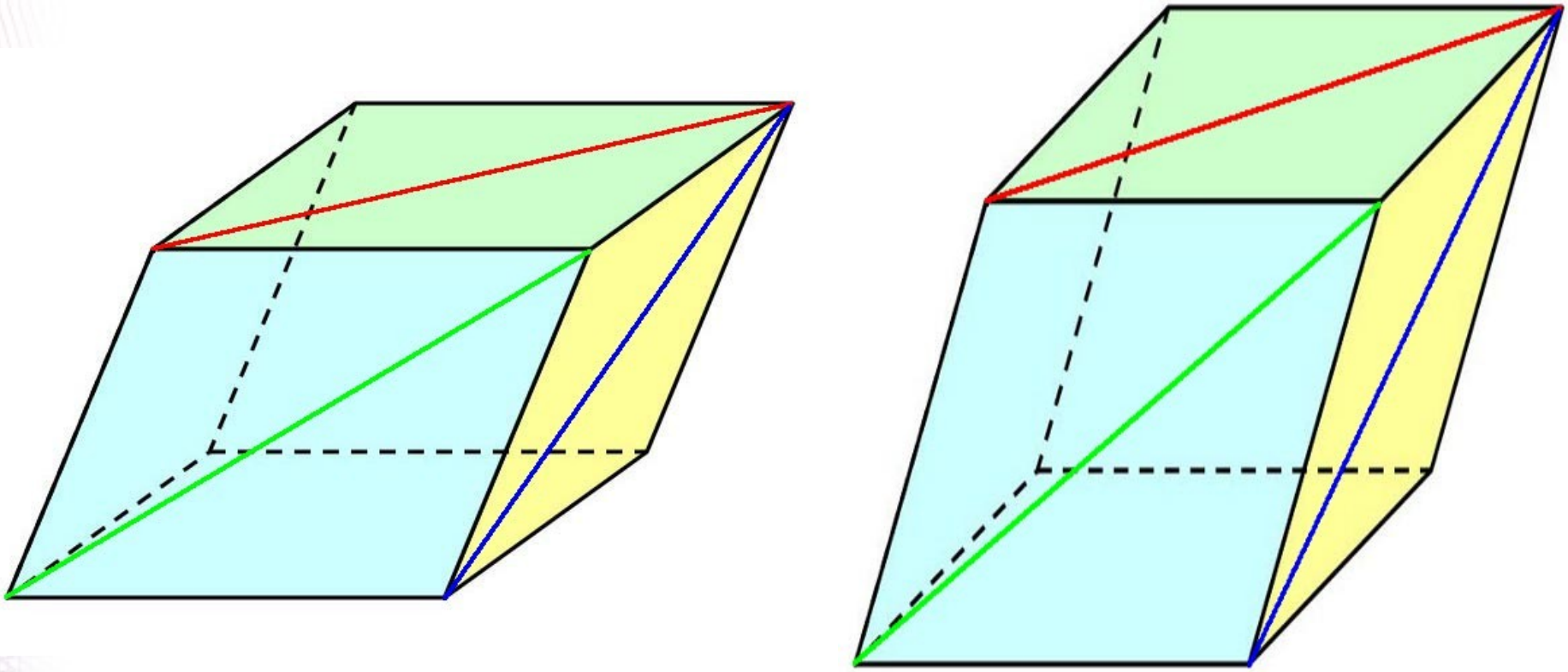
CCP4 html documentation
http://www.ccp4.ac.uk/dist/html/blend.html
→ also on personal web site:
http://www.jfoadi.me.uk/BLEND_documentation/CCP4_blend.html
Tutorials
On CCP4 web site
http://www.ccp4.ac.uk/tutorials/tutorial_files/blend_tutorial/BLEND_tutorial.html
→ also on personal web site:
http://www.jfoadi.me.uk/BLEND_documentation/BLEND_tutorial.html

BLEND


integrated data
MOSFLM, XDS
BLEND analysis
dendrogram
numbers
book-keeping
KEYWORDS file
merged_files
merging statistics
scaled data
BLEND synthesis
clusters
to phasing
combined_files
merging statistics
scaled data
BLEND combination
groups

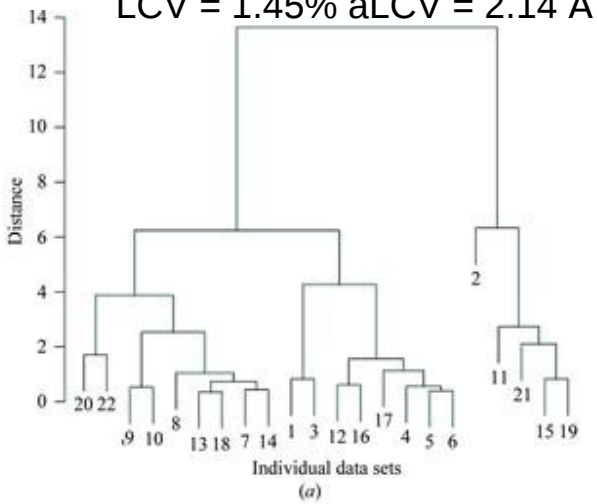
BLEND

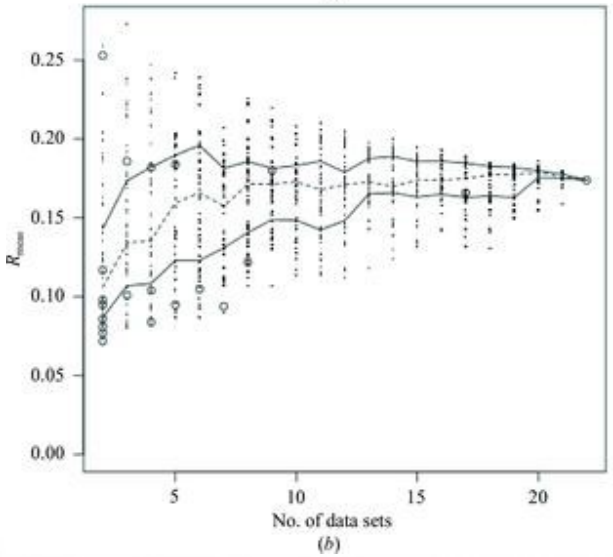
Meaning of LCV and aLCV

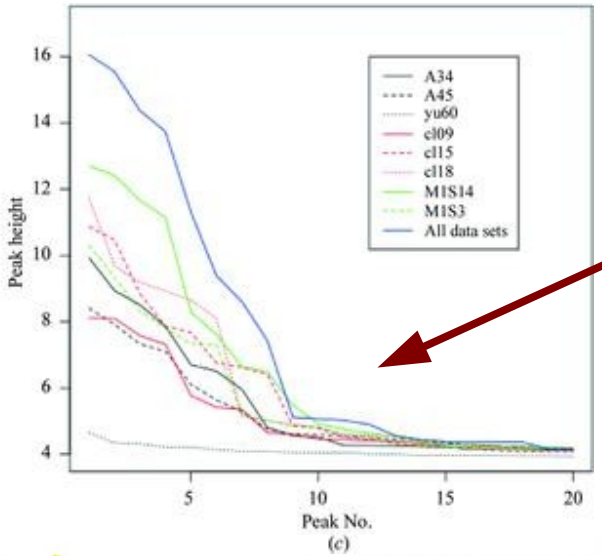


Established Results

LCV = 1.45% aLCV = 2.14 Å







<u>Cryo-cooled memPROT</u>
<u>22 data sets, 6 crystals – 4.21 Å)</u>

(Yilmaz Alguel, Alexander Cameron, So Iwata et al)

Confirms Hendrickson's findings

Imperial College London

MPL Membrane Protein Laboratory

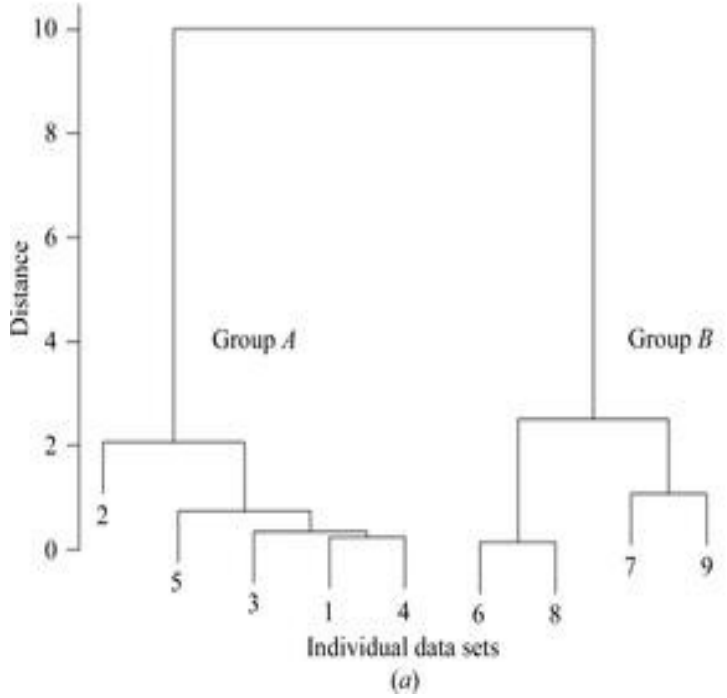
diamond

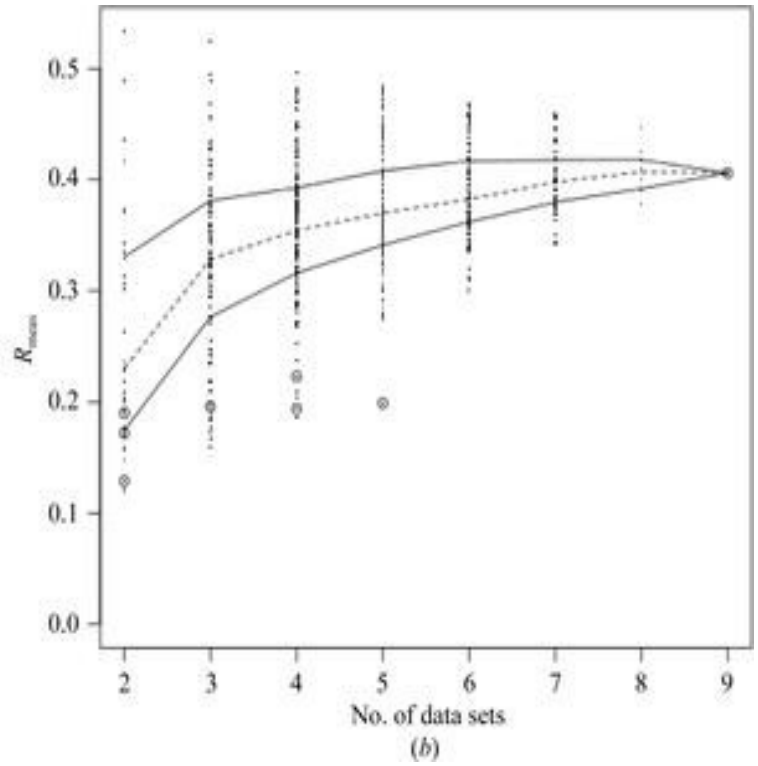
Established Results

Cryo-cooled ultralente insulin (9 crystals)

(courtesy of Armin Wagner)

LCV = 1.96% aLCV = 1.56 Å





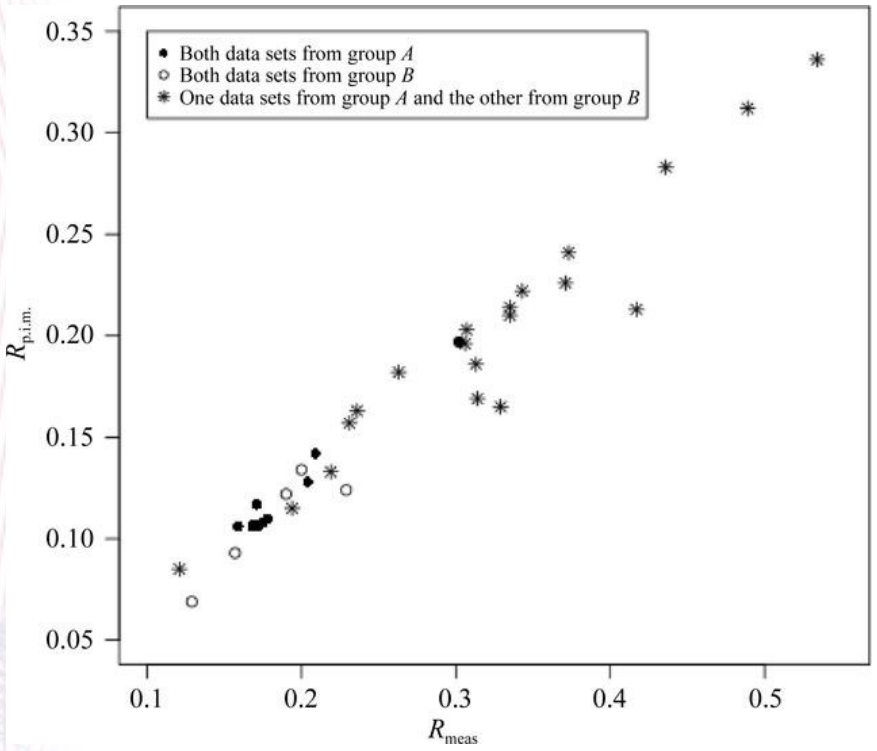
Imperial College London

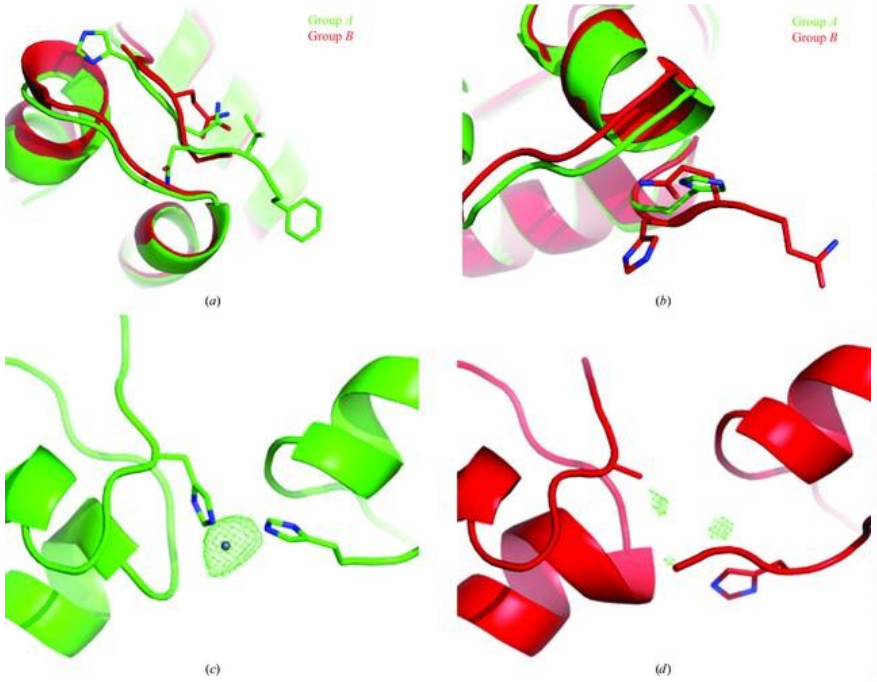
MPL Membrane Protein Laboratory

diamond

Established Results

Cryo-cooled ultralente insulin (9 crystals)

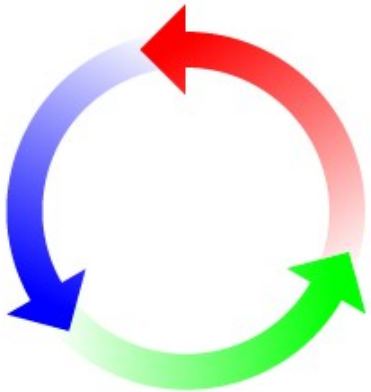




Applications

1) in situ data collection of a tellurite resistance protein homolog – **tehA**
67 data sets, 56 crystals (2.3 Å)

2) SAD data collection on crystals of lysozyme soaked with bromine
12 crystals (1.6 Å)

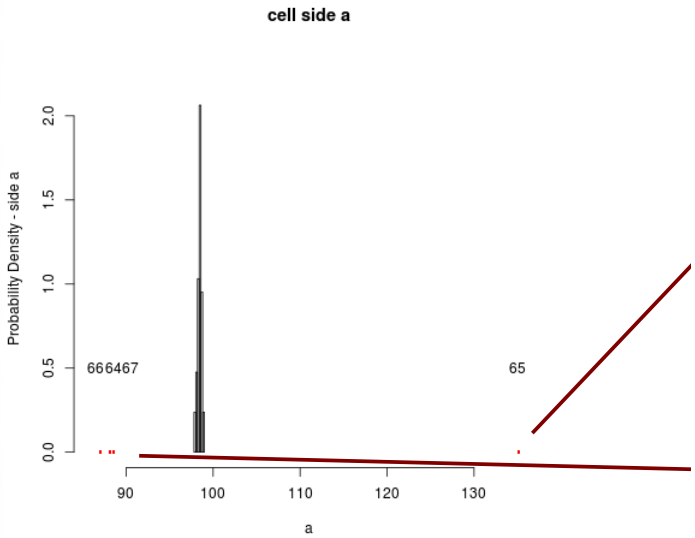


selection
combination
filtering

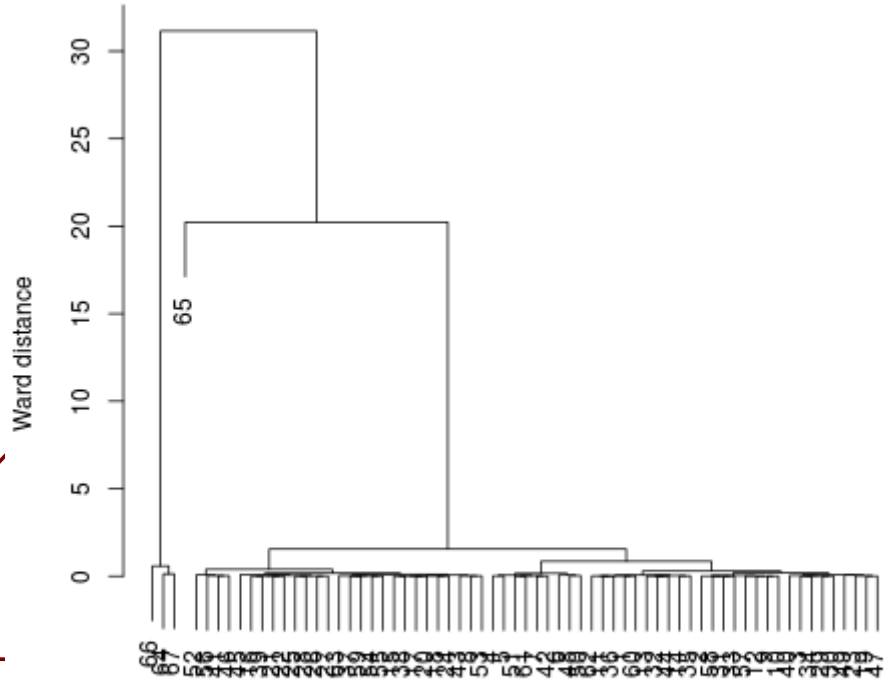
Applications

in situ Tellurite resistance protein homolog

huge value!



LCV: 57.38% (absolute LCV: 48.10 Å)



64, 65, 66 and 67 can be filtered out

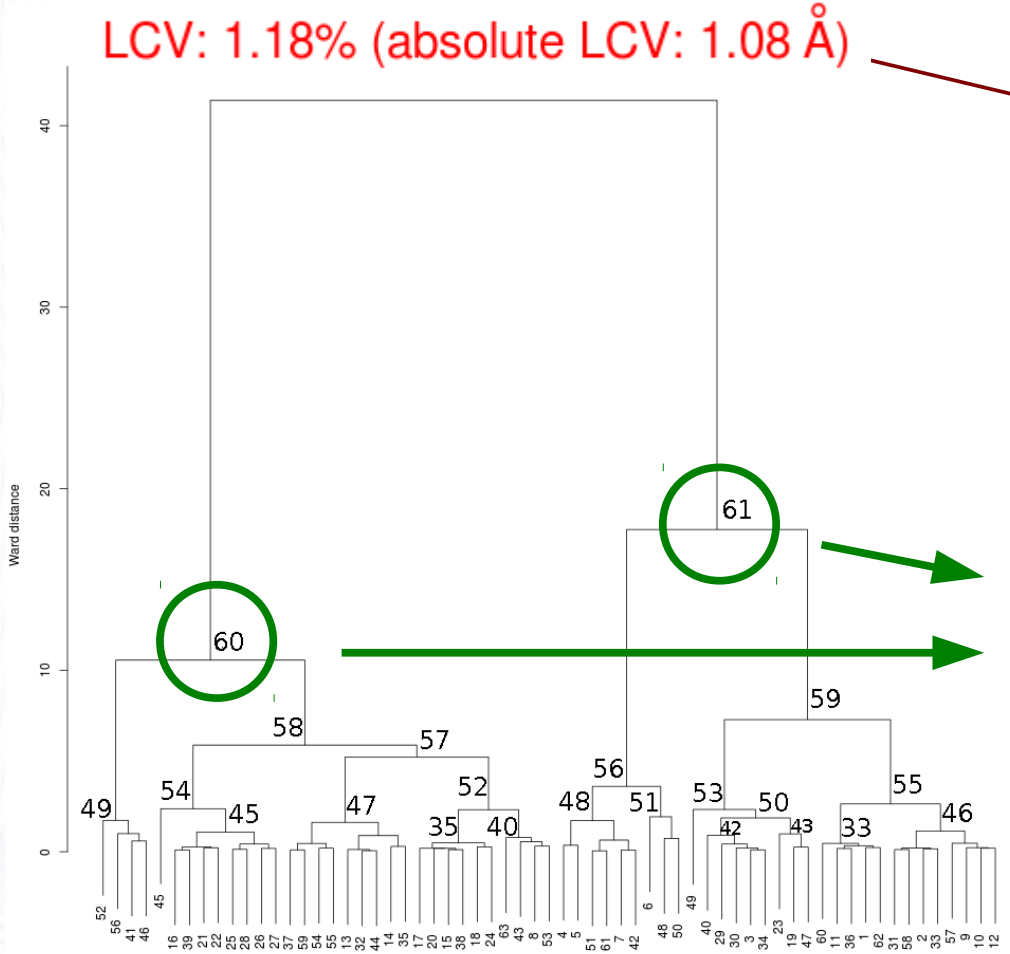
Imperial College London

MPL Membrane Protein Laboratory

diamond

Applications

in situ Tellurite resistance protein homolog



LCV = 1.18%
a reasonable value

The two different branches might point to some mild non-isomorphism.
But …
Cluster 60 is 89.7% complete
Cluster 61 is 70.7% complete

We select cluster 60

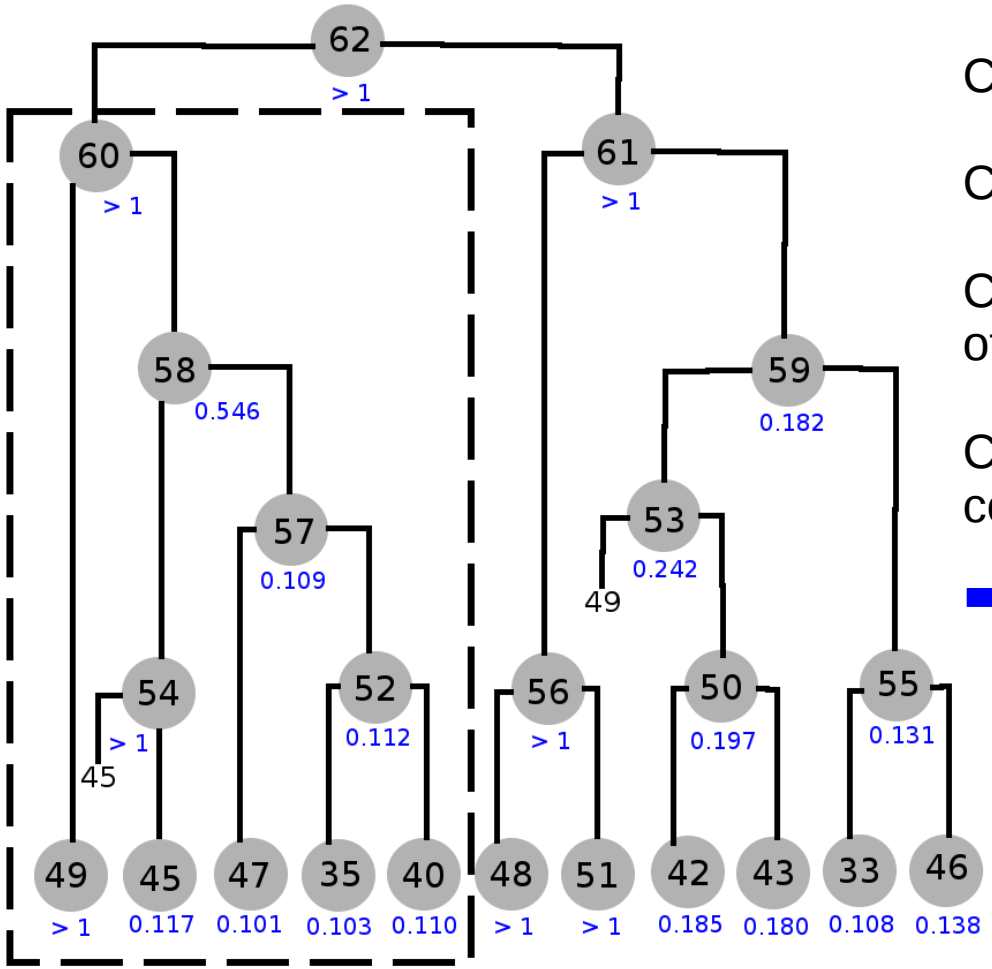
Imperial College London

MPL Membrane Protein Laboratory

diamond

Applications

in situ Tellurite resistance protein homolog



Cluster 57 shows a good value of Rmeas.

Cluster 58 shows a bad value of Rmeas.

Cluster 54 is “polluting” the good quality of cluster 57.

Cluster 60 is “polluted” by cluster 58, but could also be “polluted” by cluster 49.

→ **We need to combine cluster 57 with groups of data sets or individual data sets of cluster 54, in order to form a “cleaned” cluster 58. Then we combine with data sets from cluster 49, in ordert o form a “cleaned” cluster 60.**

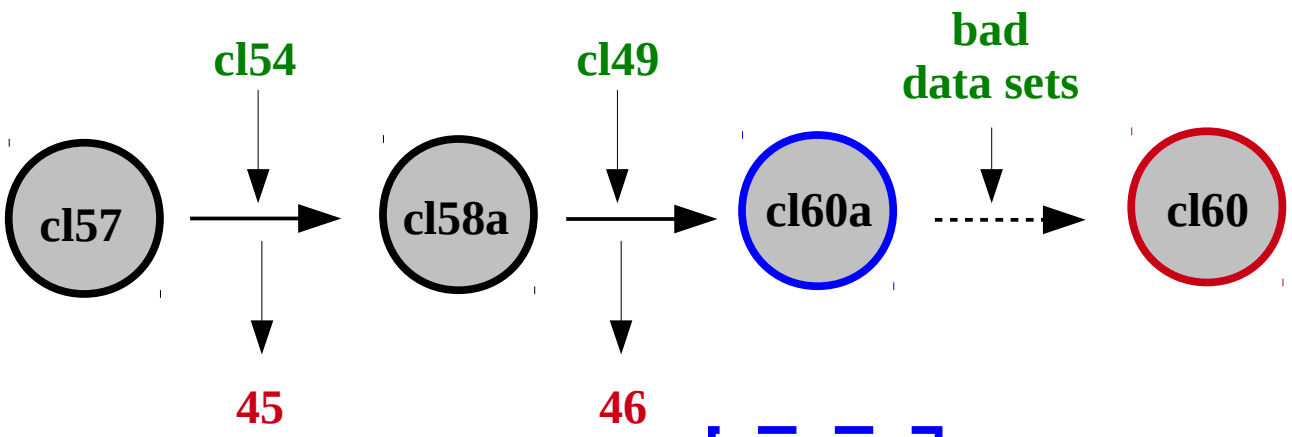
Imperial College London

MPL Membrane Protein Laboratory

diamond

Applications

in situ Tellurite resistance protein homolog



Rmeas	0.109	0.113	0.121	> 1
Rpim	0.055	0.050	0.052	> 1
Compl.	81.9%	84.1%	89.1%	89.7%

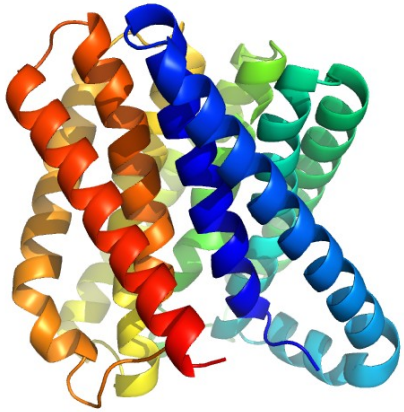
Imperial College London

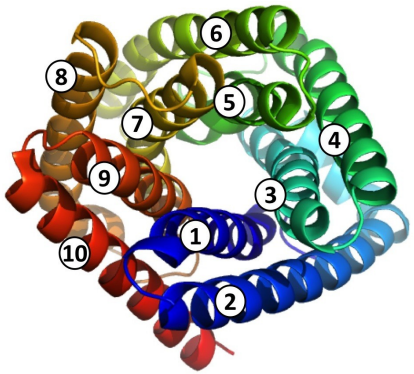
MPL Membrane Protein Laboratory

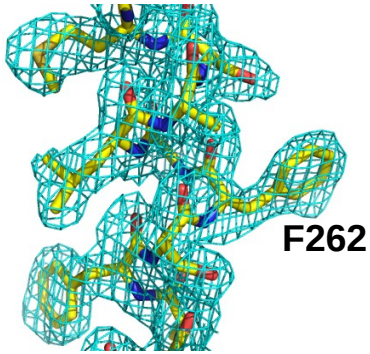
diamond

Applications

in situ Tellurite resistance protein homolog

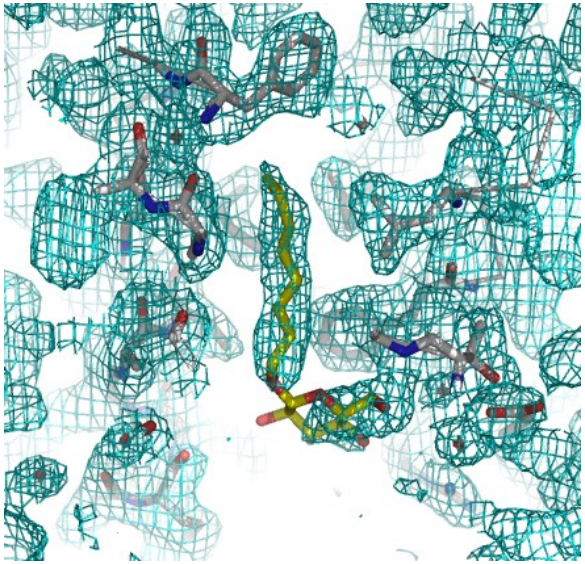


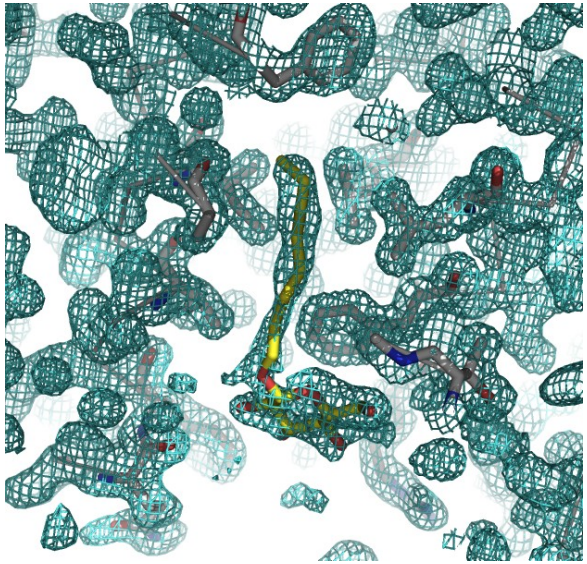




F262

2.3 Å





1.5 Å

Imperial College London

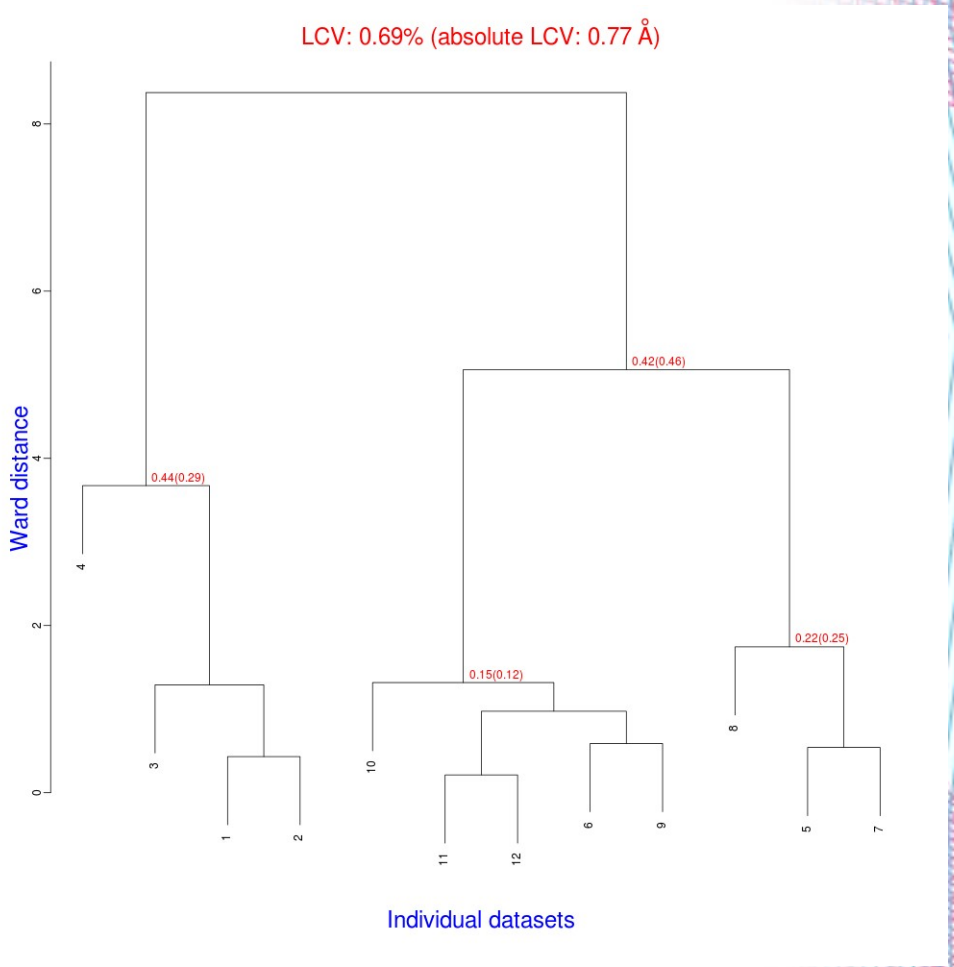
MPL Membrane Protein Laboratory

diamond

Applications

SAD lysozyme with bromine

Data Set	Completeness	Anomalous Completeness
1	70.90	30.3
2	73.30	65.3
3	94.70	80.1
4	75.20	36.0
5	84.30	22.2
6	85.50	52.6
7	71.10	51.0
8	60.90	42.3
9	82.30	40.5
10	79.30	20.2
11	81.30	51.0
12	90.40	76.2



Imperial College London

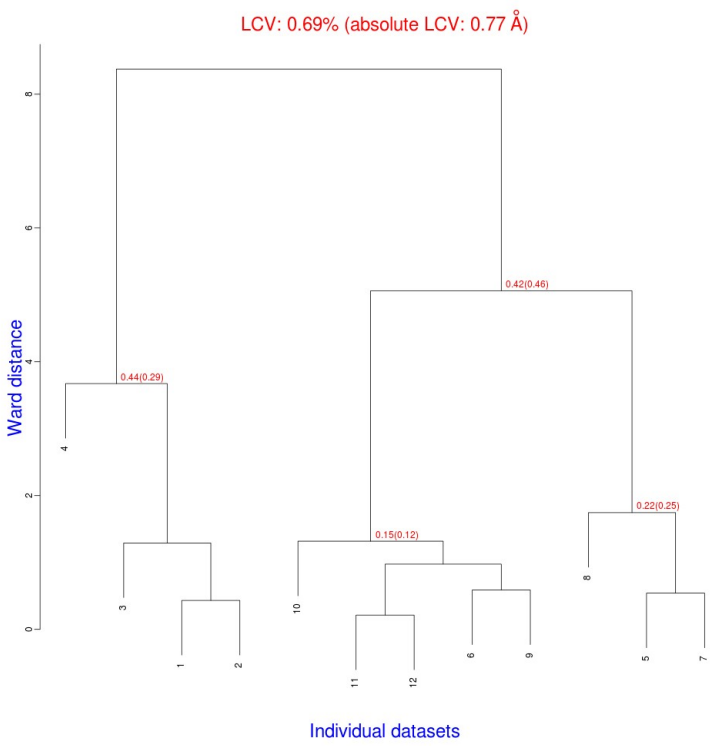
MPL Membrane Protein Laboratory

diamond

Applications

SAD lysozyme with bromine

Cluster	Completeness	Anomalous Completeness
1	88.3	73.7
2	93.6	70.4
3	92.9	70.9
4	96.0	69.5
5	98.0	92.6
6	99.8	98.4
7	98.7	94.8
8	96.3	83.9
9	99.9	98.9
10	99.4	98.2
11	99.9	99.9



Imperial College London

MPL Membrane Protein Laboratory

diamond

Applications

SAD lysozyme with bromine

SHELXC → SHELXD

Data set	Sufficient d"/sig ?	Solution ?	CCall	CCweak	Cfom
ind_003	yes (-)	yes (-)	33.68	15.34	49.02
ind_012	yes	yes	43.52	25.45	68.97
005	yes	yes	45.07	24.94	70.01
006	yes	yes	37.17	19.07	56.25
007	yes	yes	44.72	26.03	70.75
009	yes	yes	36.78	18.05	54.83
010	yes	yes	45.68	26.01	71.69
011	yes	yes	44.25	26.60	70.85

Future Work

- Good indicators for anomalous phasing
- Preparation for anomalous phasing (sca files, etc.)
- Mapping data sets reciprocal space completeness (GUI)
- Splitting of discontinuous sweeps (inverse beam)
- Improved procedure for radiation damage

Acknowledgments

Phil Evans

David Waterman

DIAMOND SOFTWARE SUPPORT
Alun Ashton
Karl Levik

MPL SUPPORT
Isabel DeMoraes
Momi Iwata
So Iwata

CCP4 SOFTWARE SUPPORT
Marcin Wojdyr
Ville Uski
Andrey Lebedev

FUNDING
Wellcome Trust
BBSRC

Imperial College London

MPL Membrane Protein Laboratory

diamond

Tutorials

http://www.jfoadi.me.uk/BLEND_documentation/BLEND_tutorial.html
http://www.jfoadi.me.uk/BLEND_documentation/CCP4_blend.html

BLEND: managing, scaling and merging multiple datasets

James Foadi[(a)] and Pierre Aller[(b)]

(a) Imperial College London (j.foadi@imperial.ac.uk) , Diamond Light Source Ltd (james.foadi@diamond.ac.uk)

(b) Diamond Light Source Ltd (pierre.aller@diamond.ac.uk

Main aim of this tutorials is to get users acquainted with the program BLEND (Foadi et al, 2013) and to show how it can be used effectively to obtain complete data sets out of several partial or complete ones. The reader interested in exploring pros and cons of using data from multiple crystals can refer, for instance, to Liu et al. (2011), Giordano et al. (2012), Axford et al. (2012), Hanson et al. (2012).

CONTENTS

DATA SETS INCLUDED WITH THIS TUTORIALS

For convenience both tutorials can be carried out inside a same directory and specific results can be collected in separate sub-directories. We will assume that "blend_tutorial" is the name of the overall directory and that it has been created in the $HOME area. Different directory names and locations can be used, but appropriate modifications will have to be applied to the paths displayed in this document. Data downloaded using the above links, when unpacked, yield a directory named "data". For the rest of the tutorials this directory can be moved inside "blend_tutorial". There are three groups of data in the sub-directory “data”. One group is formed by crystals of insulin, the second by crystals of lysozyme and the third still by lysozyme soked with bromine.

Imperial College London

MPL Membrane Protein Laboratory

diamond