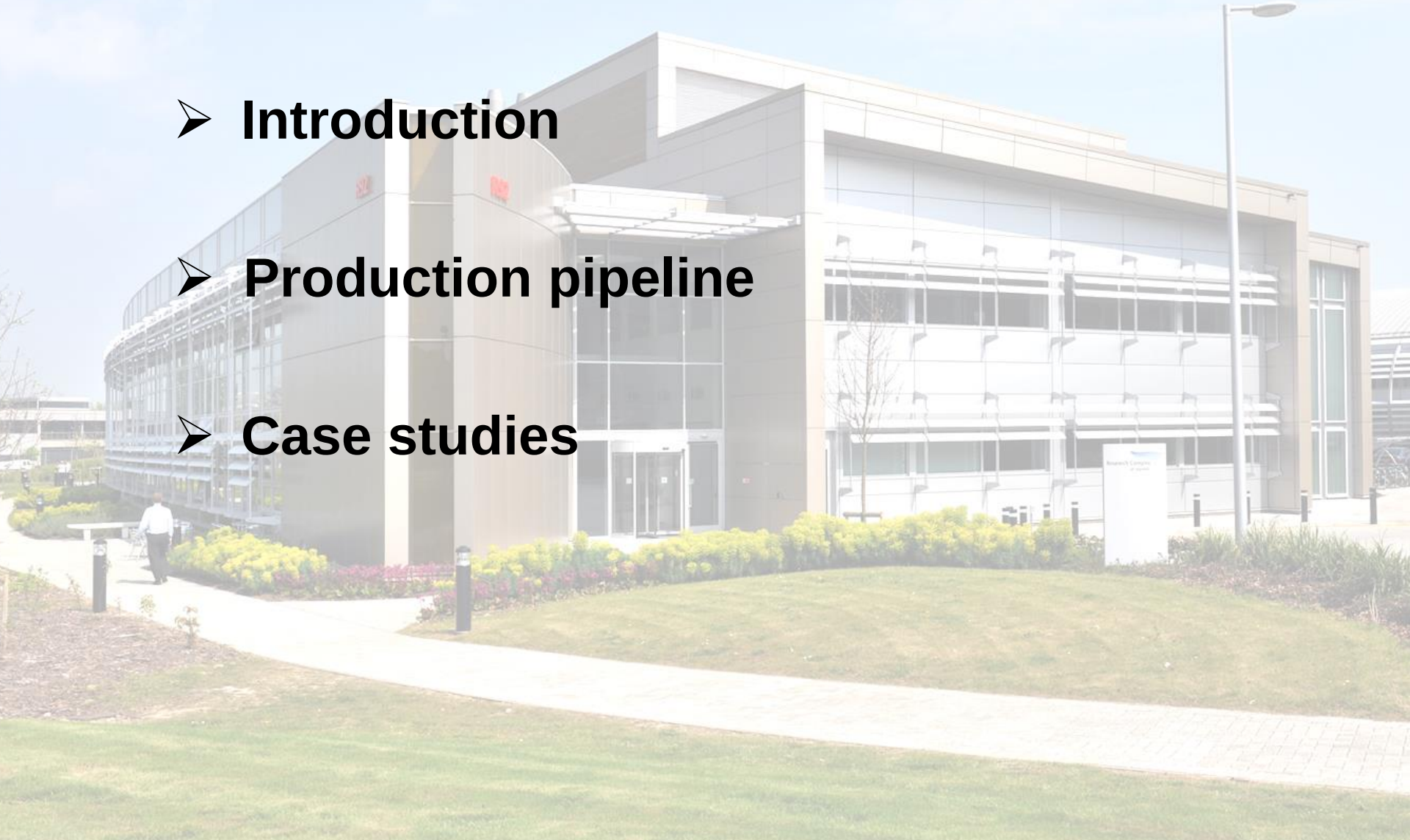
High throughput protein production for structural biology

Ray Owens
Oxford Protein Production Facility-UK
University of Oxford

OPPF-UK

UNIVERSITY OF OXFORD

Outline of presentation



- Introduction
- Production pipeline
- Case studies

The Oxford Protein Production Facility-UK



- Infrastructure for protein production and crystallization:
  - Construct design.
  - Parallel multi-vector construction & expression screening.
  - Scale-up & purification.
  - Crystallization (& testing in collaboration with Diamond).

- Active Research & development programme:
  - Technical improvement (vectors, protocols).
  - New methods (e.g. In situ diffraction testing).
  - Collaborative structural biology projects.

OPPF-UK

MRC Medical Research Council

BBSRC bioscience for the future

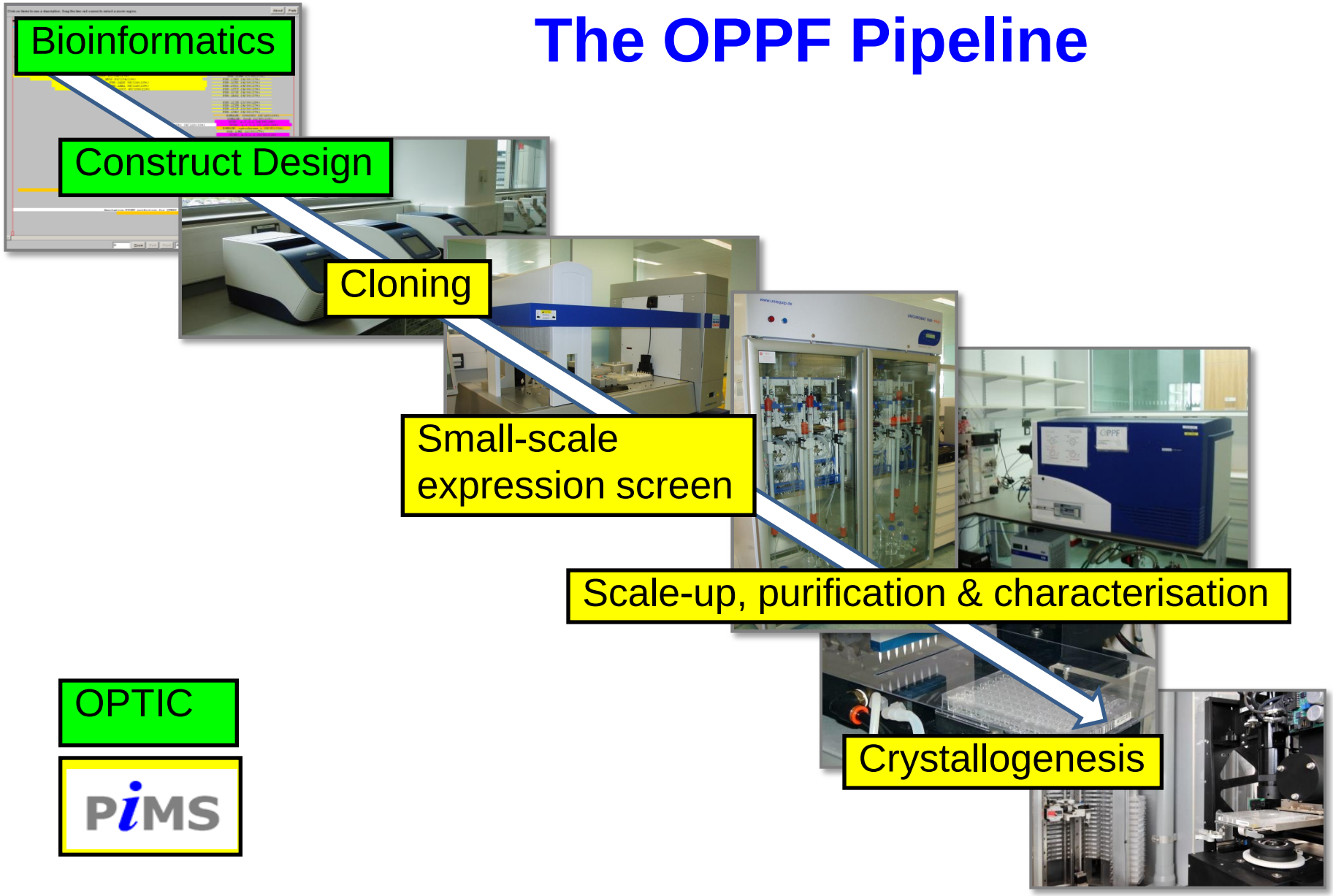
UNIVERSITY OF OXFORD

The concept of a pipeline

- A pipeline can be defined as a set of tasks connected in series, so that the output of one task is the input of the next one.
- Pipelining does not decrease the time for a single task; it only increases the **throughput** of the system by overlapping tasks.
- One key aspect of pipeline design is balancing the length of each pipeline stages to avoid creating bottlenecks.
- Another design consideration is the provision of adequate buffering between the pipeline stages.



The OPPF Pipeline

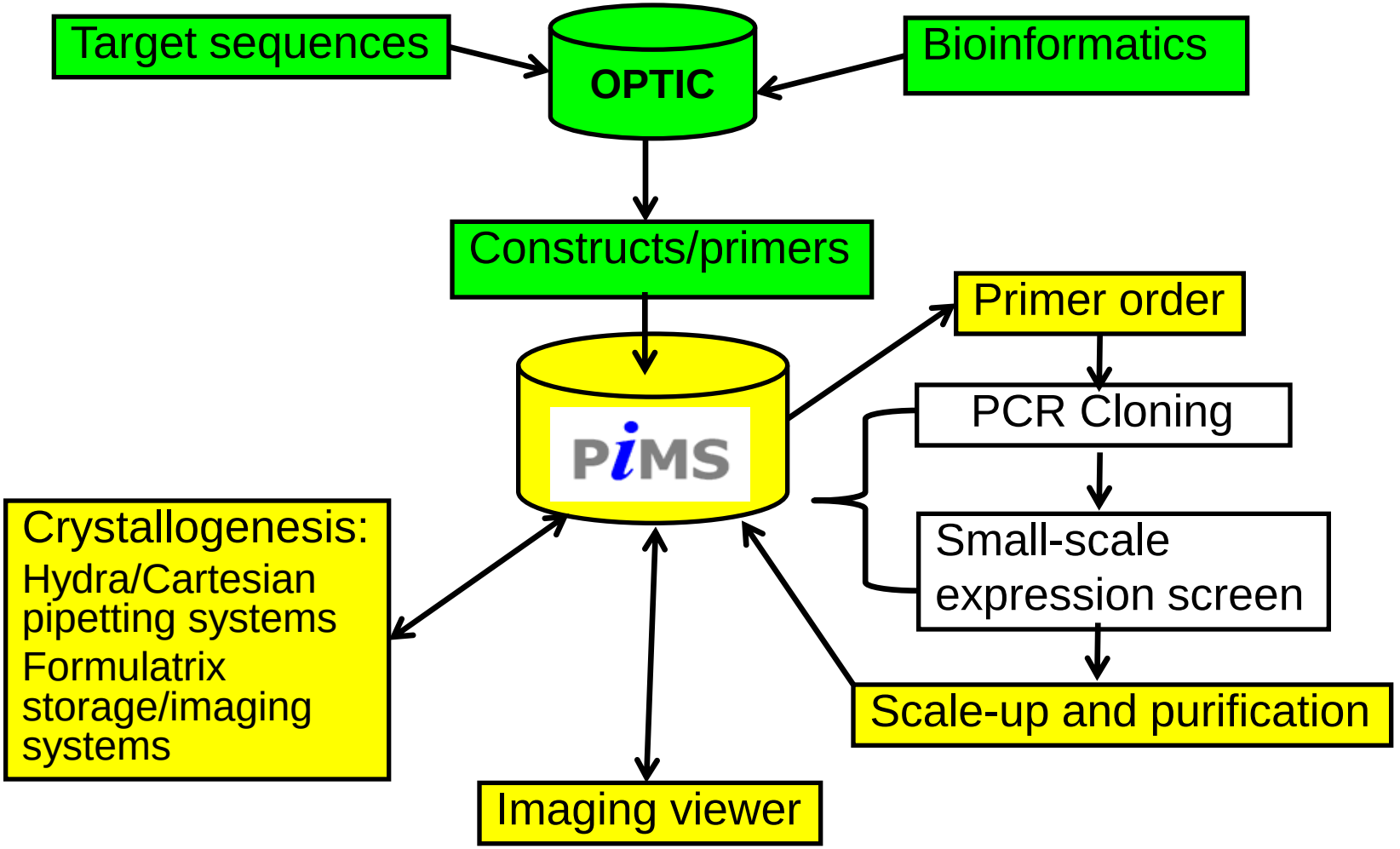


OPTIC

PiMS

OPPF-UK

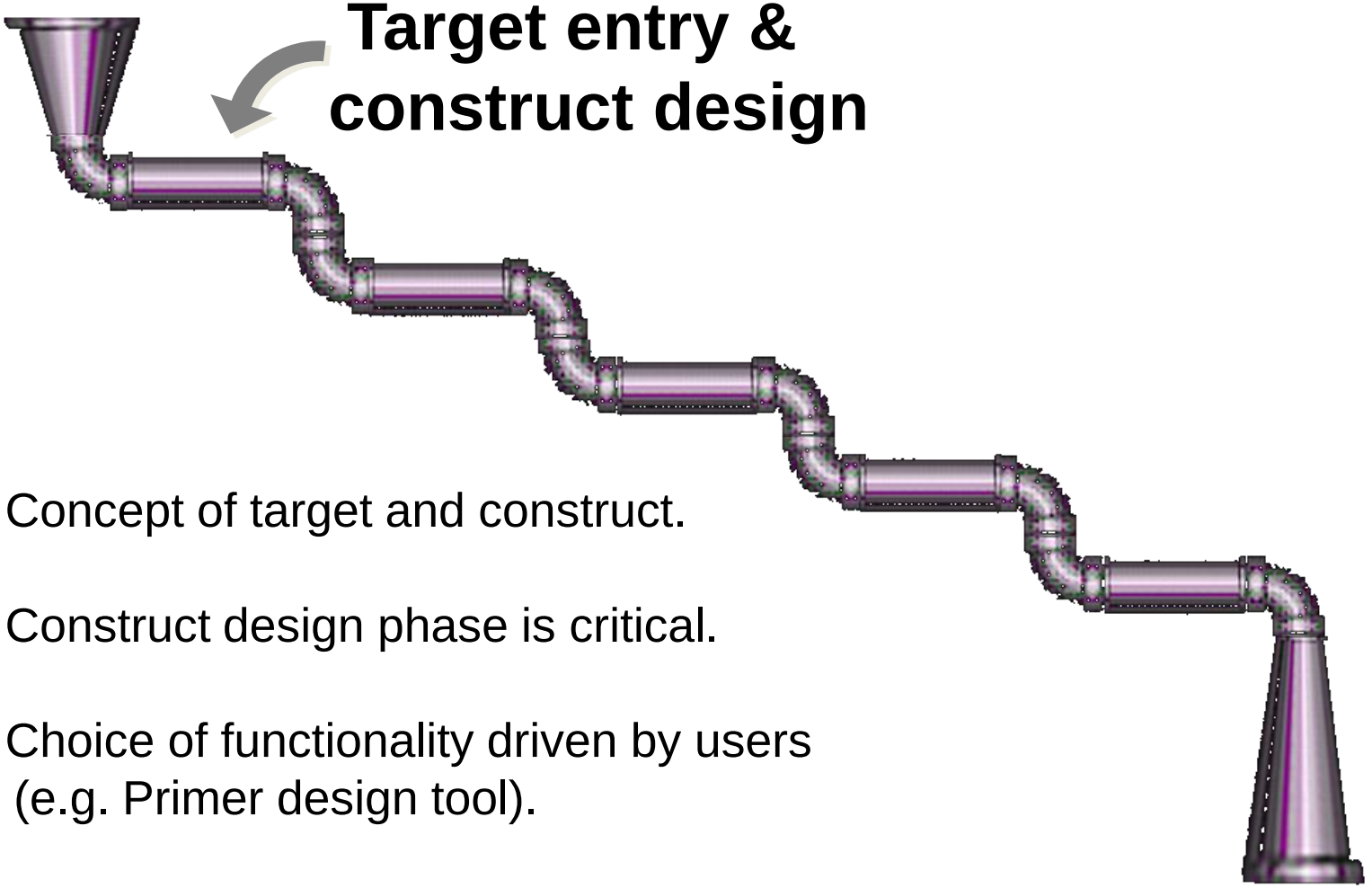
Laboratory Information Management



PPF-UK

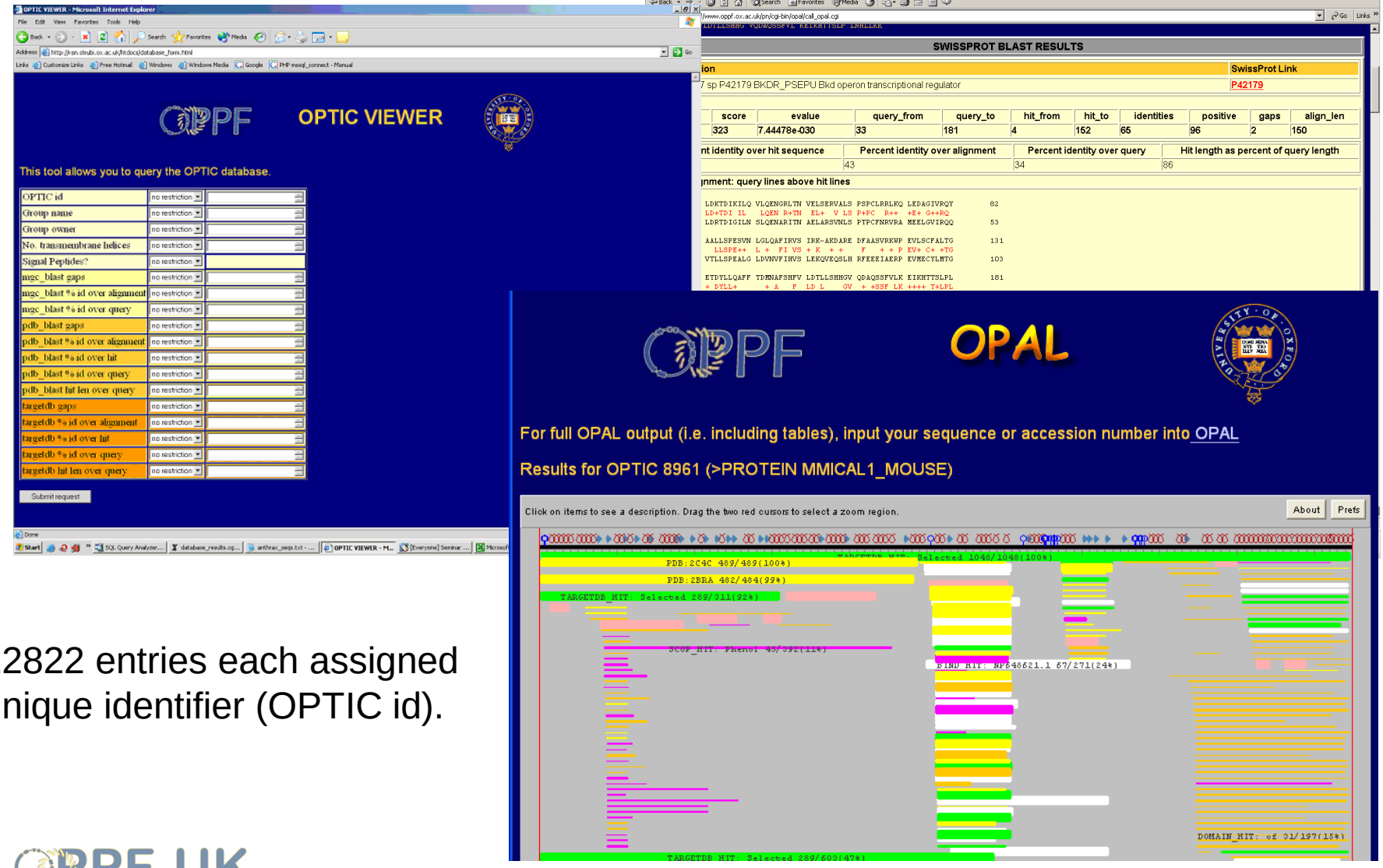
Morris et al. Acta Cryst. (2011) D67:249-260

Stage 1: bioinformatics



- Concept of target and construct.
- Construct design phase is critical.
- Choice of functionality driven by users (e.g. Primer design tool).

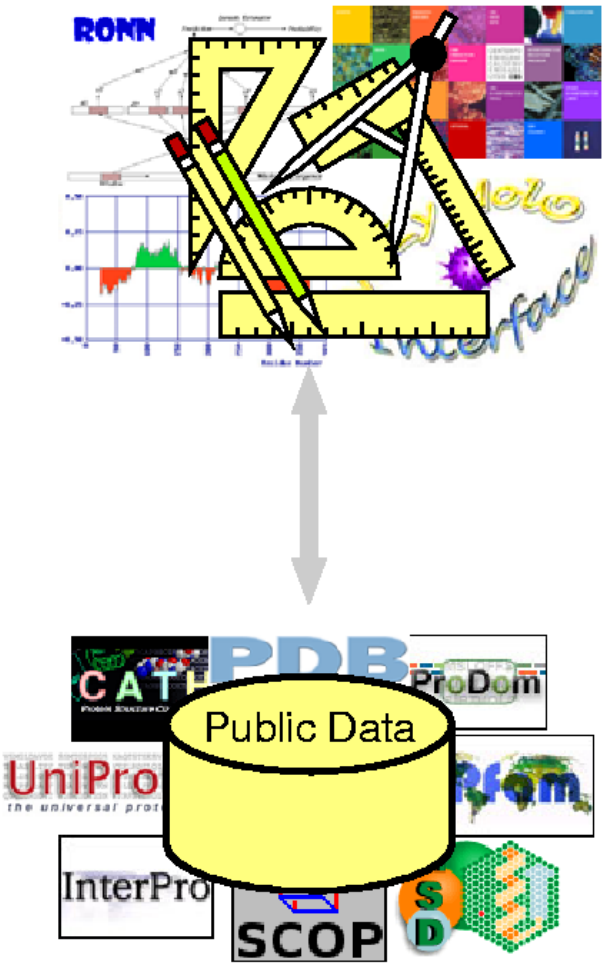
OPTIC database



12822 entries each assigned unique identifier (OPTIC id).

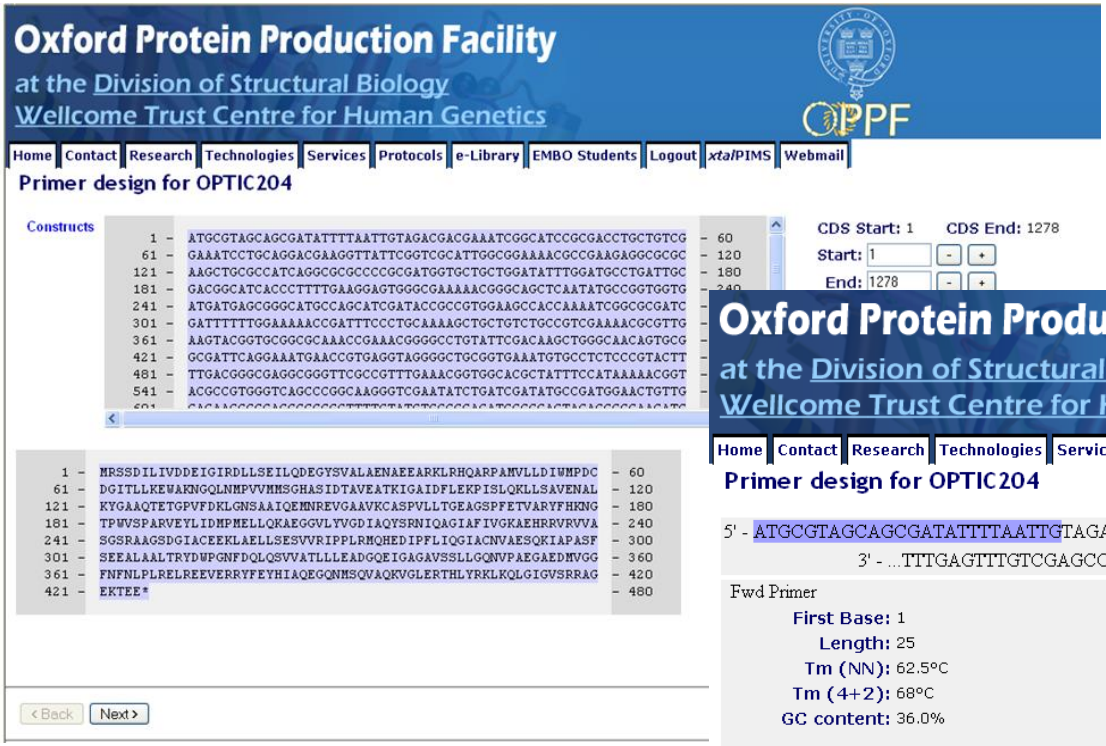
OPPF-UK

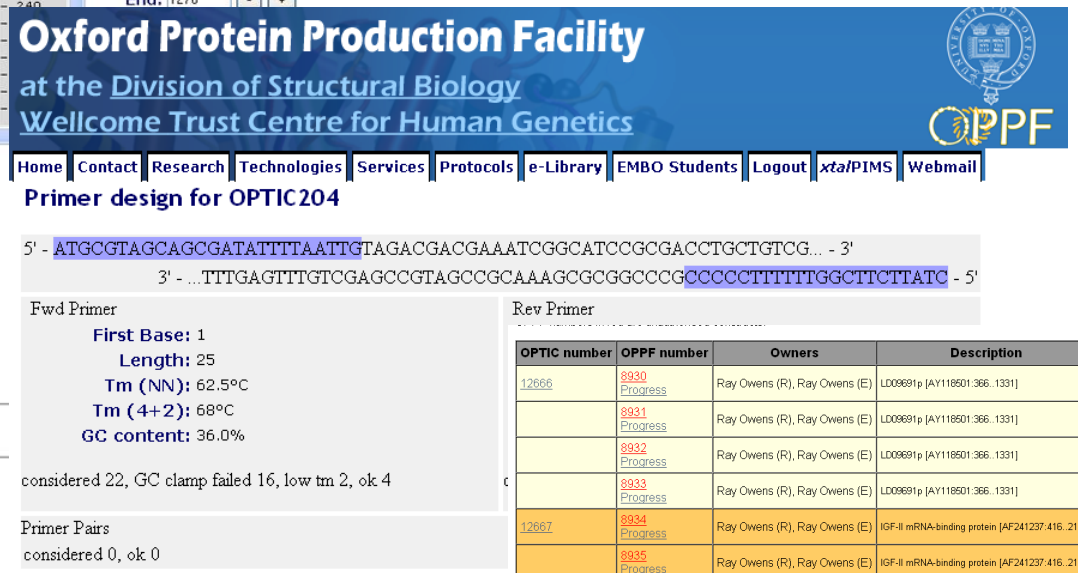
OPPF standard target annotation



- Tools
  - Protein calculator, rare codons
  - NetNGlyc, NetOGlyc
  - TMHMM2, PSORT, SignalP, TargetP
  - pI and GRAVY
  - RONN / PONDR
- Databases
  - PDB, TargetDB, MGC Clone, OPTIC
  - Genbank NR
  - Pfam, SMART, COGS, LOAD
  - Superfamily (SCOP HMM)
  - BIND
  - SwissProt

Primer design





OPTIC number	OPPF number	Owners	Description
12666	8930 Progress	Ray Owens (R), Ray Owens (E)	LD09691p [AY118501:366..1331]
	8931 Progress	Ray Owens (R), Ray Owens (E)	LD09691p [AY118501:366..1331]
	8932 Progress	Ray Owens (R), Ray Owens (E)	LD09691p [AY118501:366..1331]
	8933 Progress	Ray Owens (R), Ray Owens (E)	LD09691p [AY118501:366..1331]
12667	8934 Progress	Ray Owens (R), Ray Owens (E)	IGF-II mRNA-binding protein [AF241237:416..2116]
	8935 Progress	Ray Owens (R), Ray Owens (E)	IGF-II mRNA-binding protein [AF241237:416..2116]
	8936 Progress	Ray Owens (R), Ray Owens (E)	IGF-II mRNA-binding protein [AF241237:416..2116]
	8937 Progress	Ray Owens (R), Ray Owens (E)	IGF-II mRNA-binding protein [AF241237:416..2116]
12668	8938 Progress	Ray Owens (R), Ray Owens (E)	LD17549p [BT023837:801..3518]
	8939 Progress	Ray Owens (R), Ray Owens (E)	LD17549p [BT023837:801..3518]
	8940 Progress	Ray Owens (R), Ray Owens (E)	LD17549p [BT023837:801..3518]
	8941 Progress	Ray Owens (R), Ray Owens (E)	LD17549p [BT023837:801..3518]
12669	8942 Progress	Ray Owens (R), Ray Owens (E)	GH26816p [AY128430:568..1823]
	8943 Progress	Ray Owens (R), Ray Owens (E)	GH26816p [AY128430:568..1823]
	8944 Progress	Ray Owens (R), Ray Owens (E)	GH26816p [AY128430:568..1823]

Automatic/manual primer design to create constructs each with unique identifier (OPPF no.).

OPPF-UK

Vector design

Tag	Size (aa)	Use	Matrix/elution
Polyhistidine	6-10	Purification	Metal chelate elution with Divalent cation/imidazole/low pH
Glutathione S-transferase (GST)	396	Purification & enhance solubility	Glutathione agarose elution With glutathione
Strep-tag	8	Purification	Streptavidin (derivative) elution with desthiobiotin
FLAG ® tag	8	Purification	Anti-FLAG elution with peptide or low pH
Maltose binding protein	387	Enhanced expression & solubility & purification	Amylose elution with maltose
NusA	495	Enhanced expression & solubility	N/A
Thioredoxin (Trx)	109	Enhanced expression & solubility	N/A
DsbA	208	Enhanced expression & solubility	N/A
SUMO (smt3)	101	Enhanced expression & solubility	N/A
GFP	239	Detection	N/A

Experiments are organised into plates

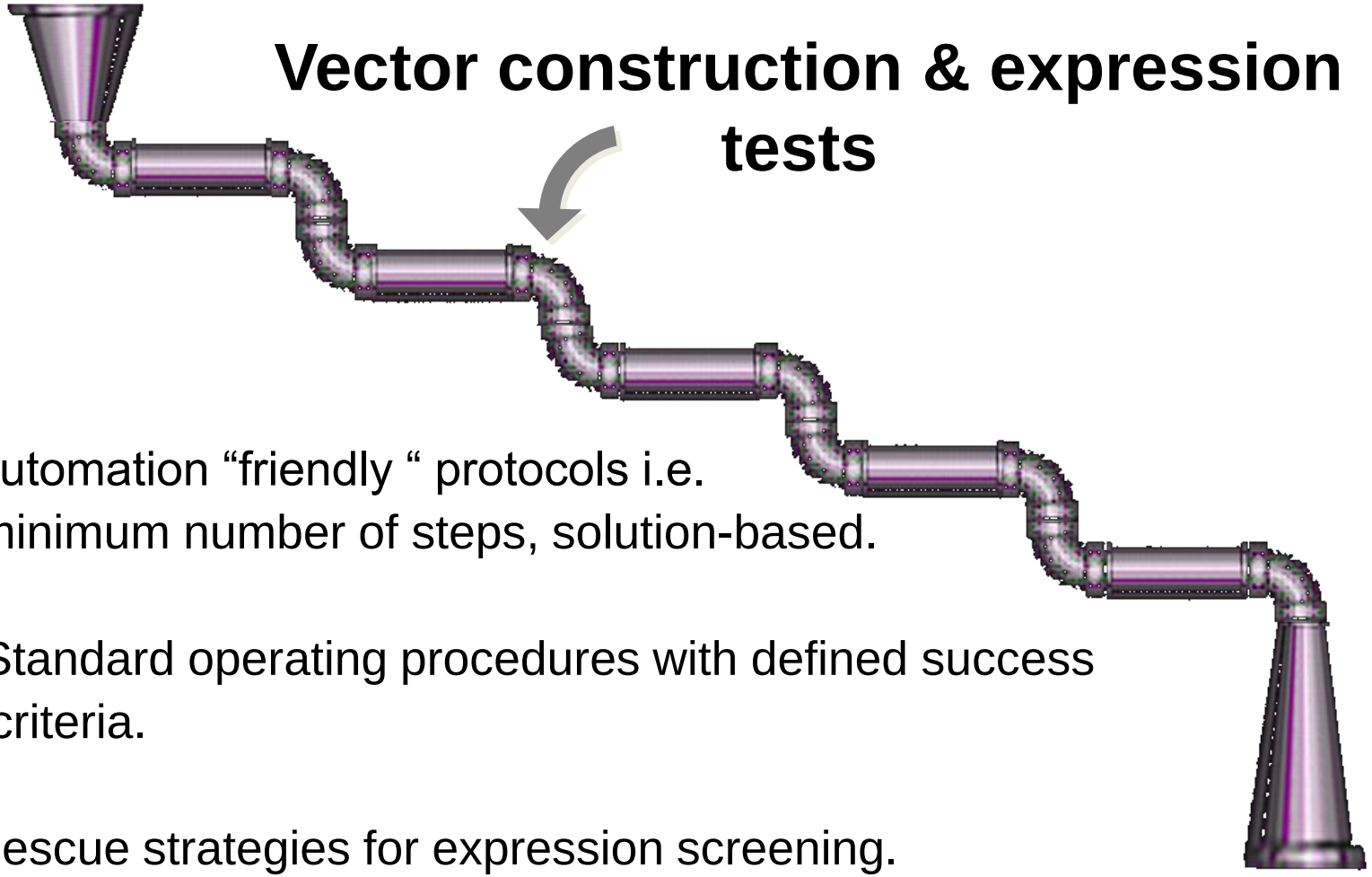
- Each project is organised into a 96 well plate format for parallel processing.

- Depending upon the project a plate may comprise:-
  - a few constructs of many different targets.
  - many constructs of a few targets.
  - full length and domain constructs.
  - multiple fusion tags.

- Plates are not re-formatted and progress to the next step in process depends upon achieving pre-set success criteria e.g. > 98 % PCR amplification of targets.

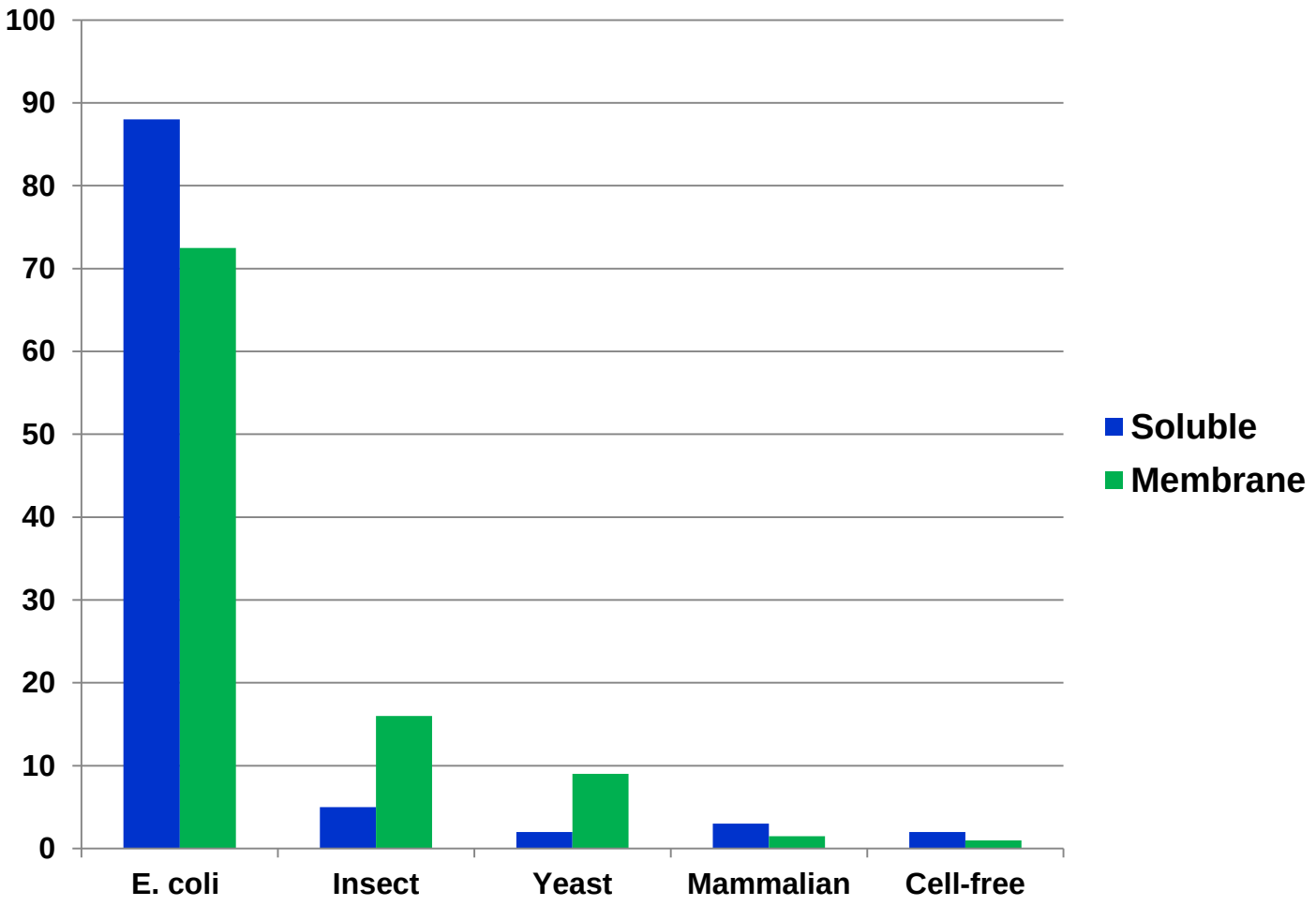
- Each 96 well plate experiment produces approximately 1000 samples (primers, PCR products, vectors, proteins from expression screens).

Stage 2: screening



- Automation “friendly “ protocols i.e. minimum number of steps, solution-based.
- Standard operating procedures with defined success criteria.
- Rescue strategies for expression screening.

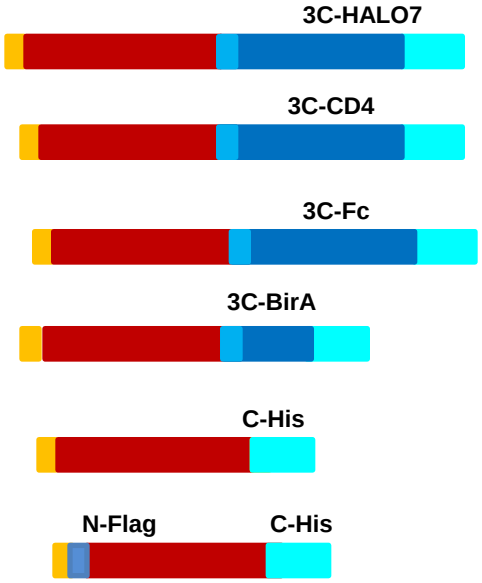
Choice of expression systems

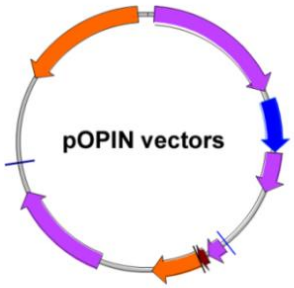


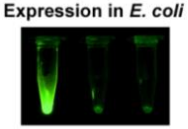
http://blanco.biomol.uci.edu
http://www.pdb.org/pdb

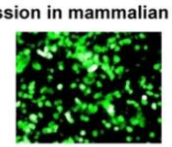
Choice of vector

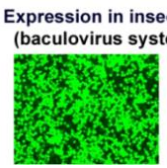
Mammalian expression vectors
(incl. NeoR)











Multi-host vectors

PCR products can be used to create a family of fusion proteins for expression in different hosts.

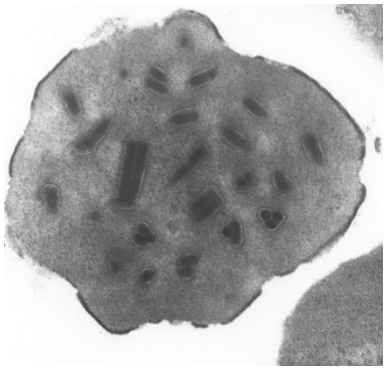
His_6 tag
3C cleavage site
Signal sequence
Protein of interest

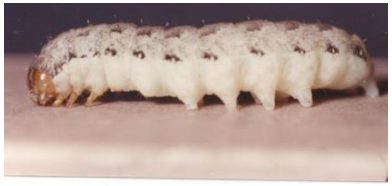
Berrow, et al. (2007) Nucleic Acids Res. 35(6):e45
Berrow, et al. (2009) Methods Mol Biol. (2009) 498:75-90
Bird (2011) Methods 55, 29-37
Bird et al. (2014) Methods Mol Biol. 1116:209-34.

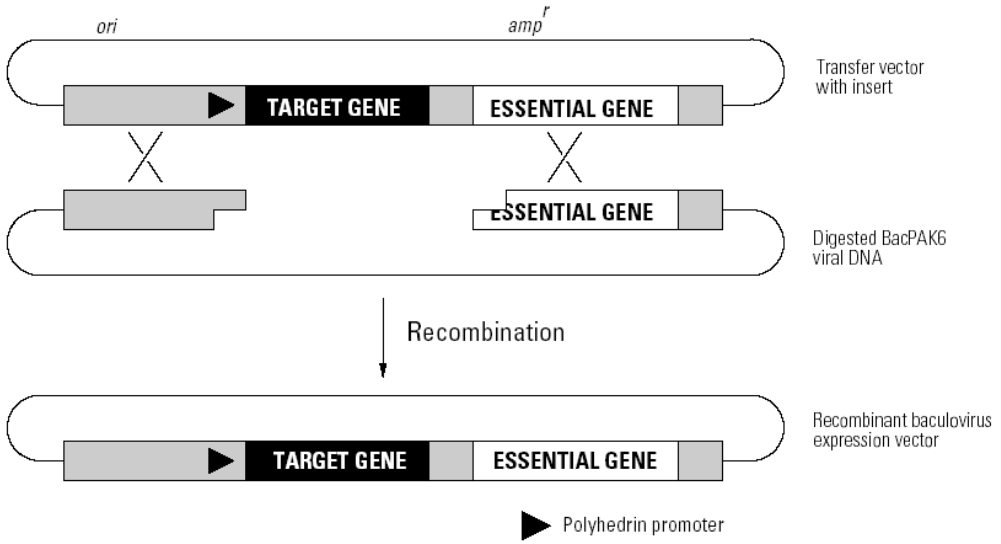
PPF-UK

Baculovirus construction

The polyhedrin gene is required for propagation of the virus in the natural habitat but in cell culture, the polyhedrin is not required and can be replaced with a gene of interest by homologous recombination in insect cells.



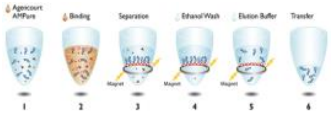


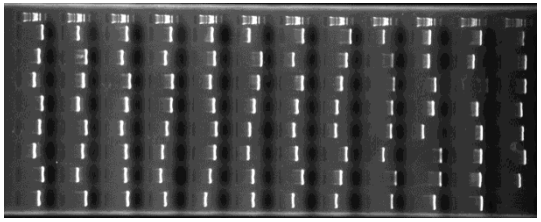


Hitchman et al.(2010) Biotechnol Appl Biochem.56(3):85-93.
Hitchman et al. (2010) Cell Biol Toxicol.26(1):57-68.
Possee et al (2008) Biotechnol Bioeng.101(6):1115-22.
Zhao et al.(2003)Nucleic Acids Res.31(2):E6-6.

96 Well PCR Cloning

PCR amplification (50ul) in 96 well plate: DpnI treat and purify using magnetic beads (Agencourt AMPure).

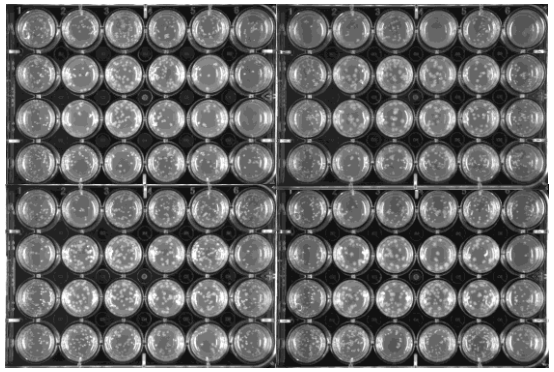




Quality assessment.

Miniprep and PCR verification.

Total elapsed time = 5 days



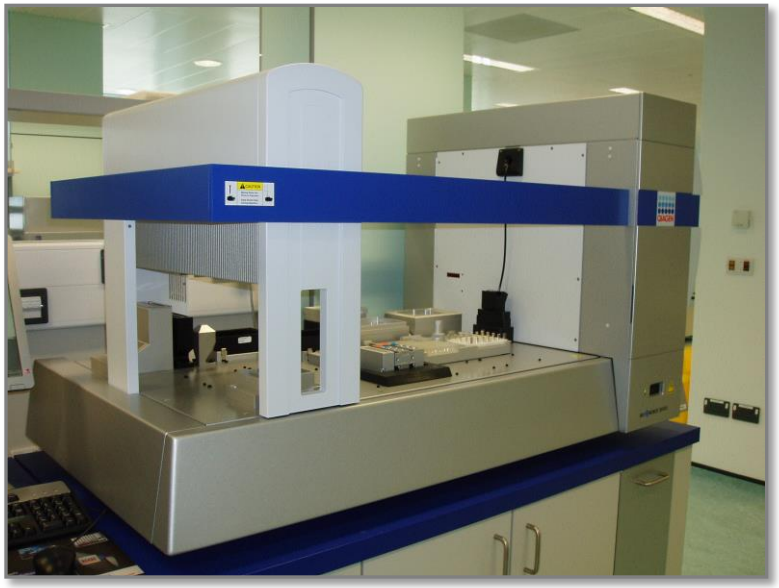
96 well plate ligation independent cloning into expression vector and transformation plating on 4 x 24 well plates.

OPPF-UK

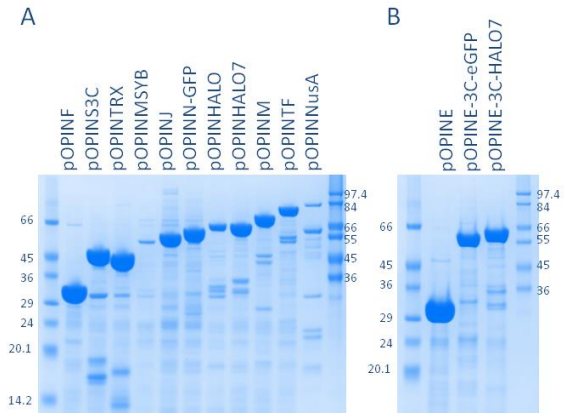
Expression screening in *E. coli* and Insect Cells

E. Coli
24 deep-well plates.
Two strains.
Two Induction methods.

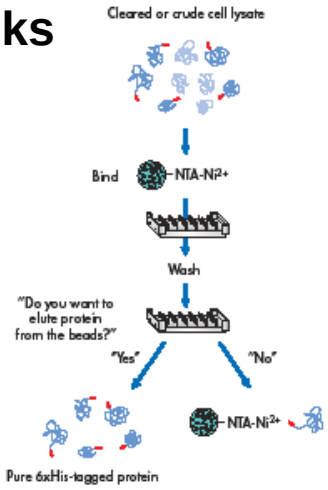
Insect cells (Sf9)
24 deep-well plates.
Two virus concentrations.
Two harvest points.



Total elapsed time
***E. coli* = 1 week**
Insect Cells = 3 weeks

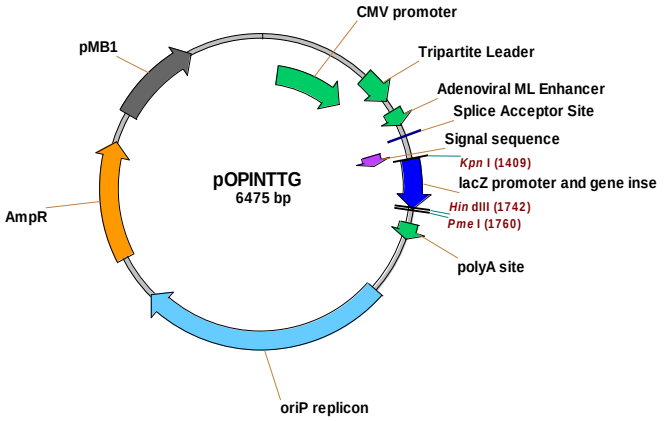


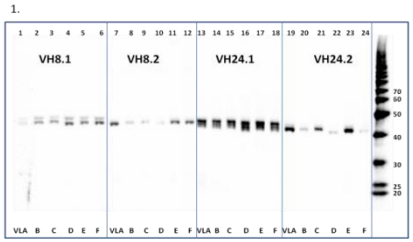
SDS-PAGE analysis

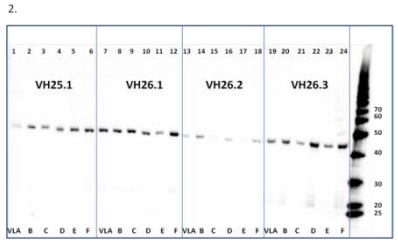


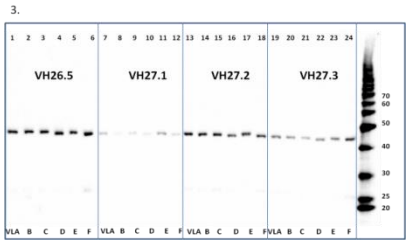
Expression screen: expression from 1ml of cells analysed Ni NTA-magnetic beads for soluble/detergent solubilised cells . Ni-NTA agarose on a vacuum manifold for insoluble proteins.

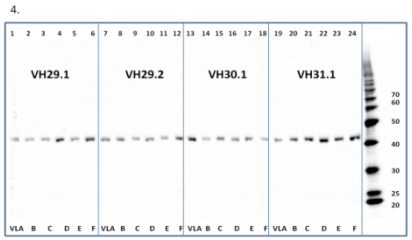
Expression screening in HEK 293T cells











Semi-automated transfection of HEK293T cells 24-well plates

SDS-PAGE : W. Blot media 72h post-transfection

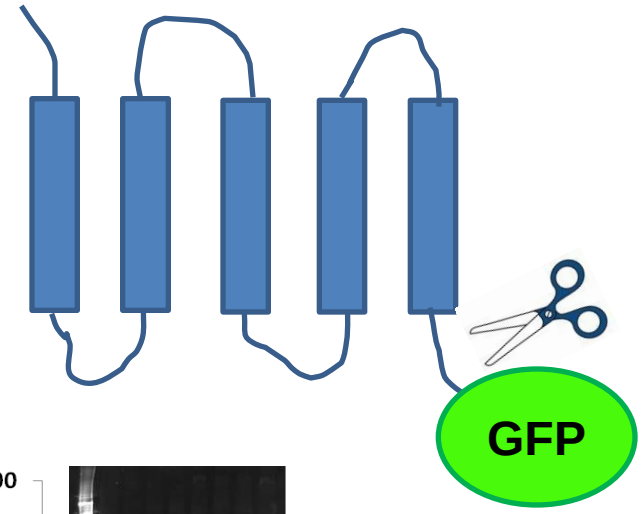
Selection for large-scale transient (1 L) and/or stable pool

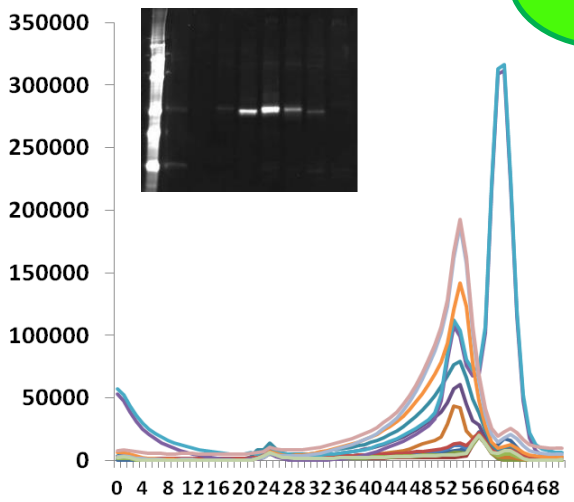
Total elapsed time = 1 week

PPF-UK

Nettleship et al. (2009) Methods Mol. Biol. 498: 245-263
pTT vector: Durocher et al. (2002) Nucleic Acids Res.30:e9

Fluorescence detection Size Exclusion Chromatography (FSEC)



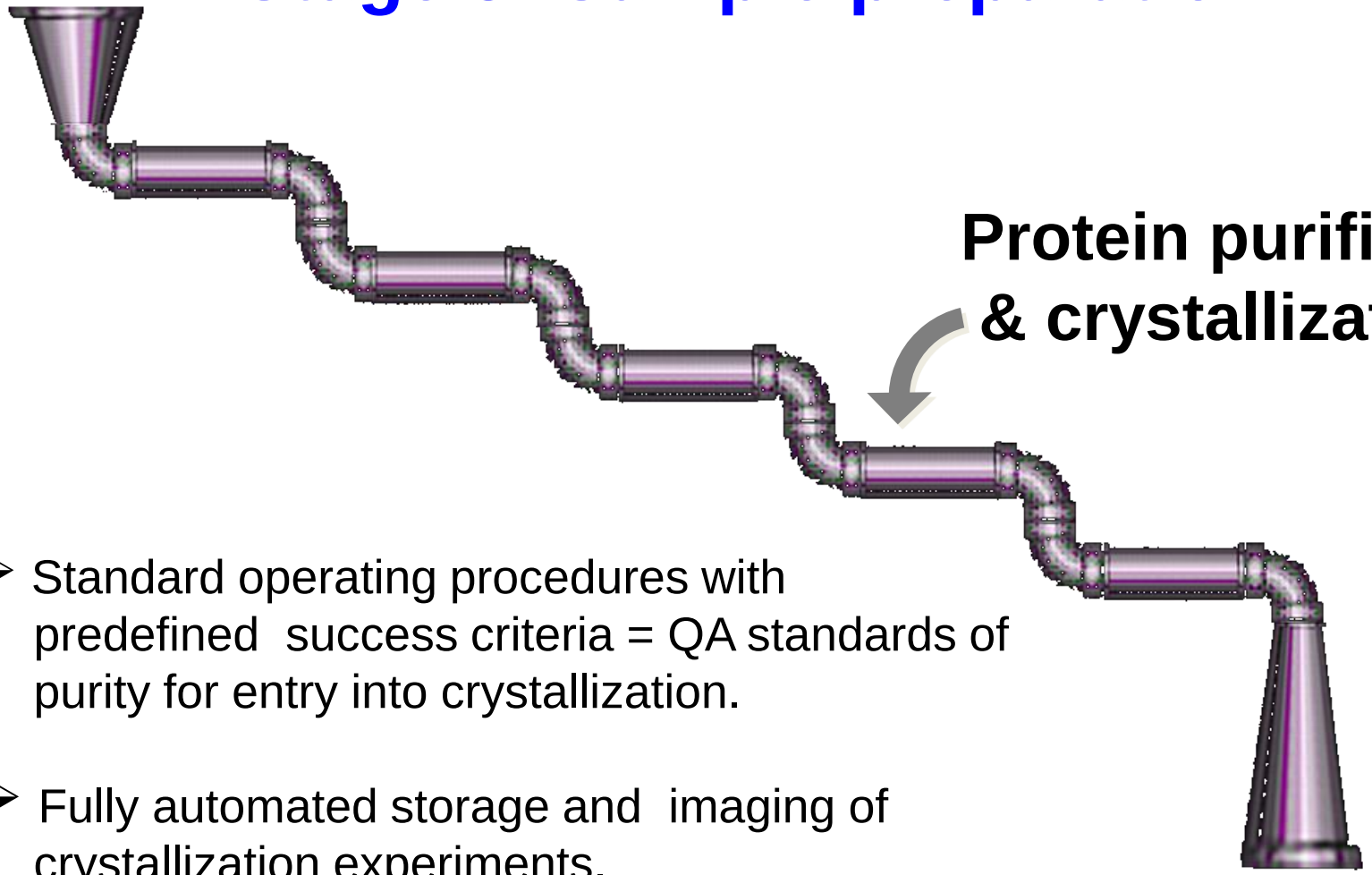


Screening membrane protein expression

Kawate and Gouaux (2006) Structure. 14(4):673-81
Drew et al. (2006) Nat Methods 3: 303-313.
Parcej et al. (2013) PLOSOne 8: e67112.

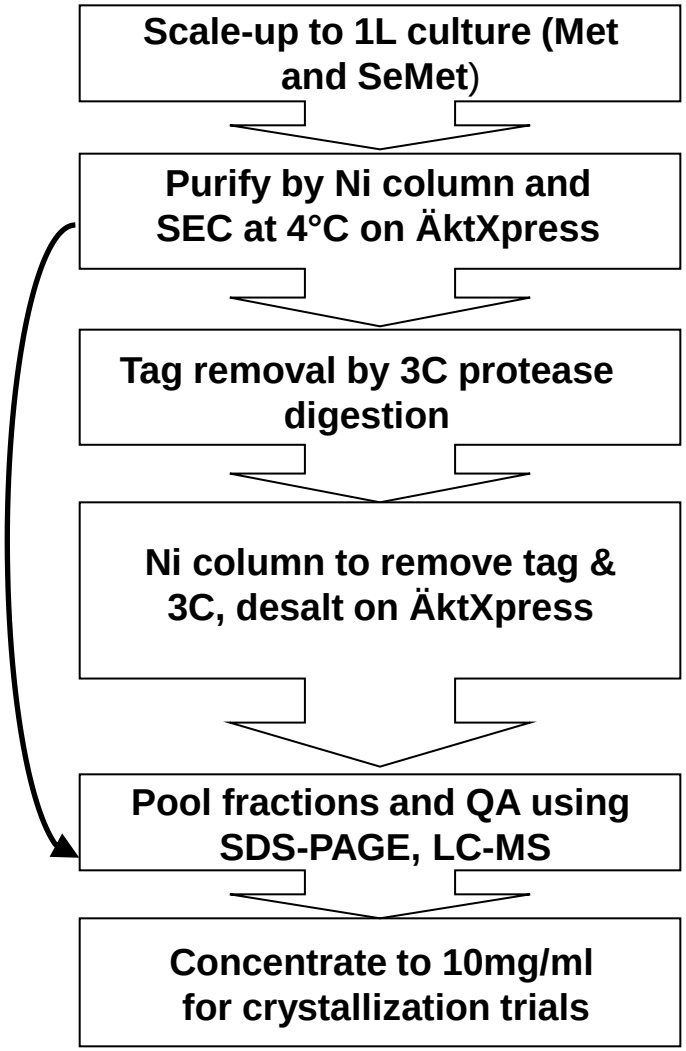
PPF-UK

Stage 3: Sample preparation



- Standard operating procedures with predefined success criteria = QA standards of purity for entry into crystallization.
- Fully automated storage and imaging of crystallization experiments.

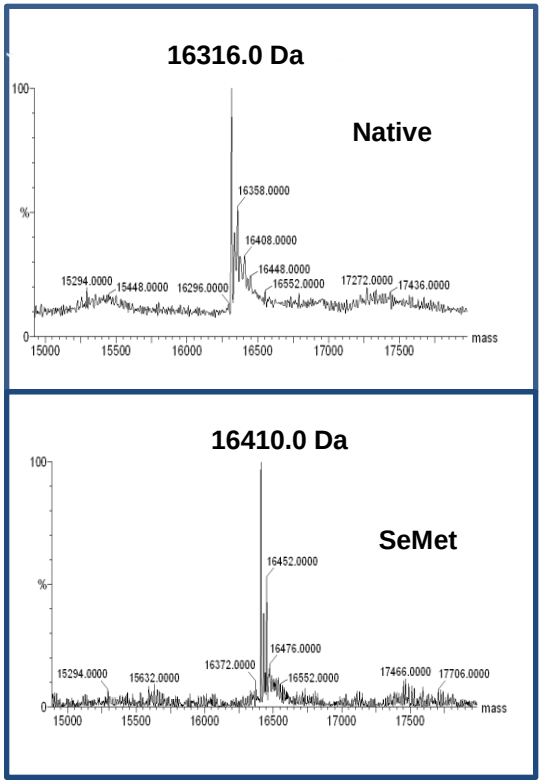
Scale-up and purification: intracellular (*E .coli*, insect cells)



PPF-UK

Mass Spectrometry – Intact Proteins

- Intact protein MS measures the accurate mass (to within 1Da) of a protein sample.
- Verifies the construct and measures purity.
- Assess SeMet incorporation.
- Disulphide redox state.

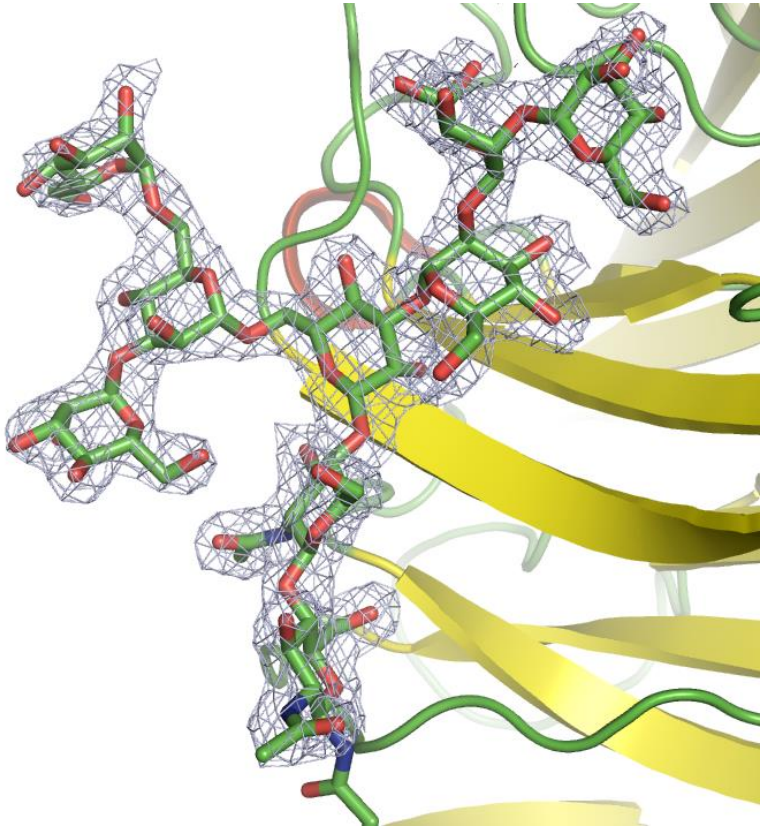


Molecular mass difference between Native and SeMet = 94 showing substitution by 2 seleniums (100% incorporation).

PPF-UK

Crystallization of Glycoproteins

- Large sugar chains are disordered and can prevent crystallization.
- Mutate variably occupied N-glycosylation sites.
- Reduce and simplify glycosylation by modification of N-glycan processing.

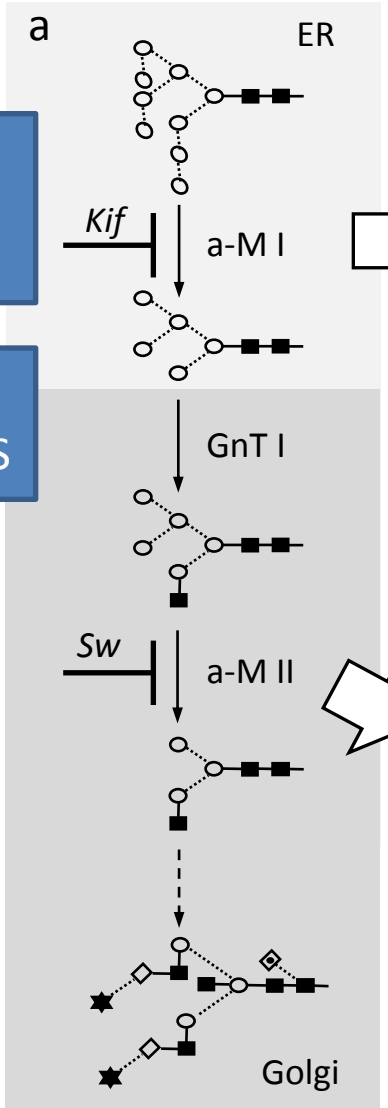


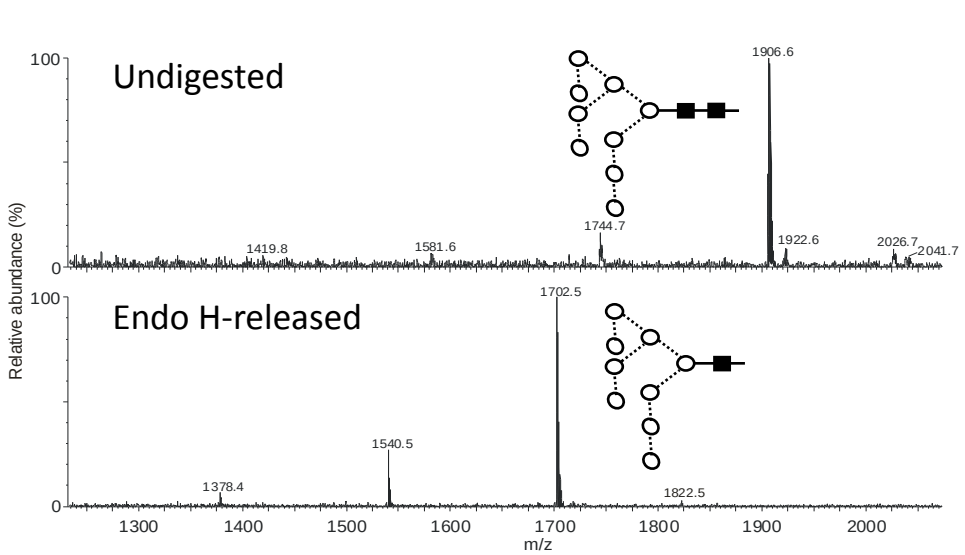
Graphic: Max Crispin

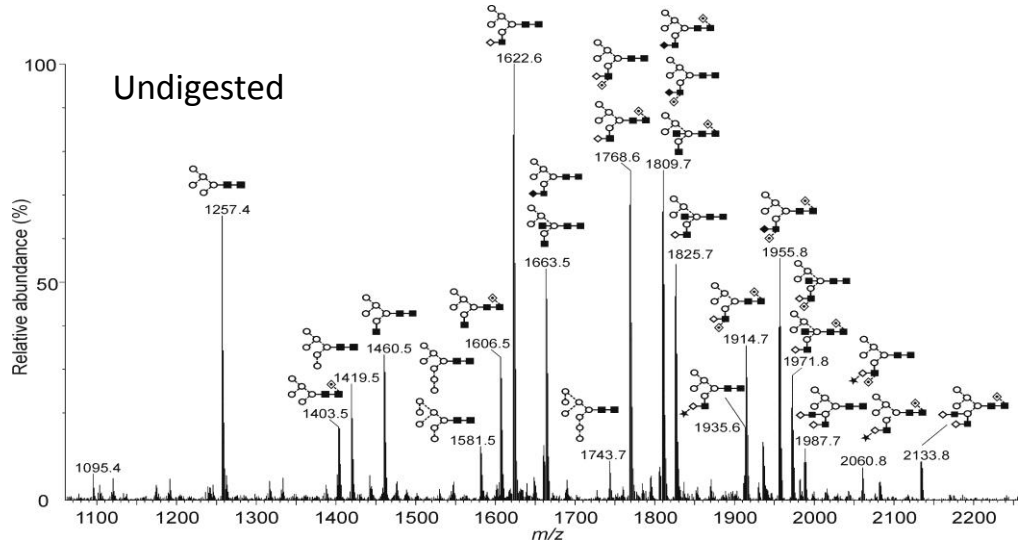
Modification of *N*-Glycan Processing

Use of the inhibitor kifunensine

Use of the GnTI$^-$ cell line HEK 293S



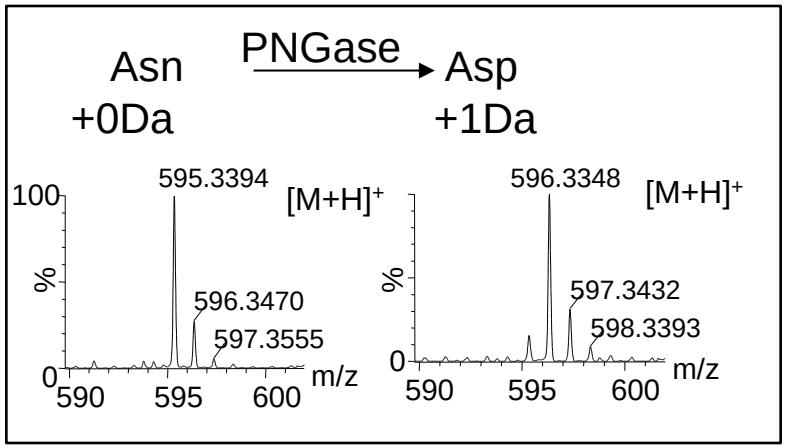


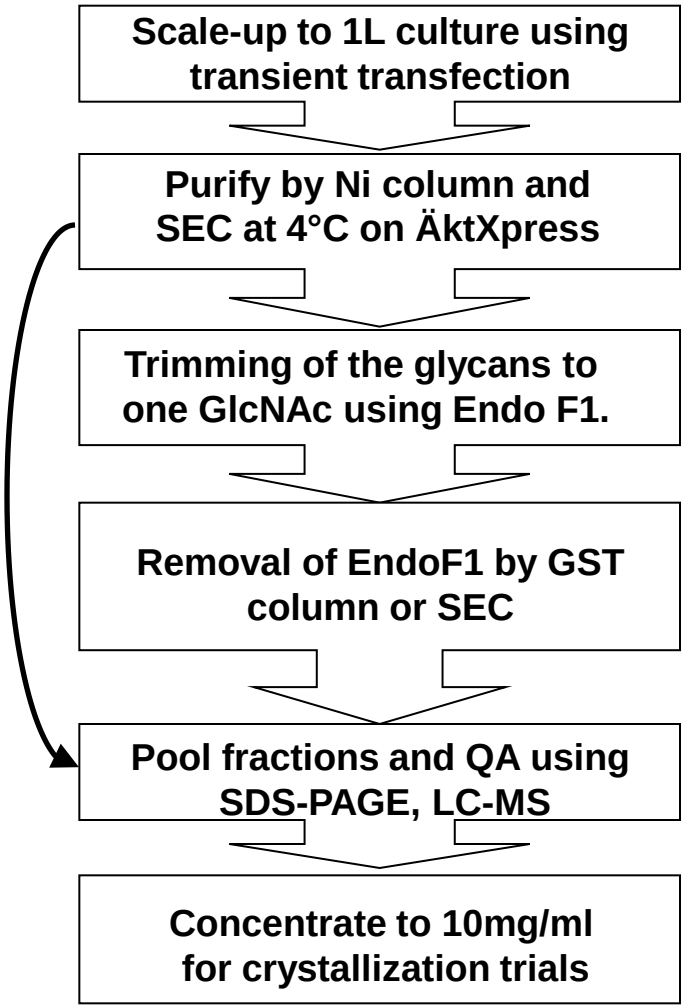


PPF-UK

V. Chang et al. (2007) Structure 15, 267-273

Scale up and Purification: secreted (mammalian cells)

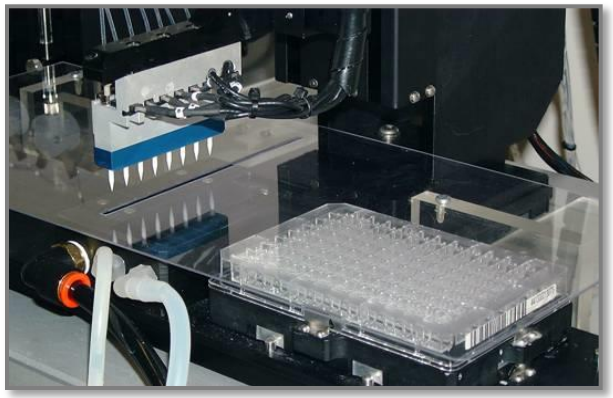




PPF-UK

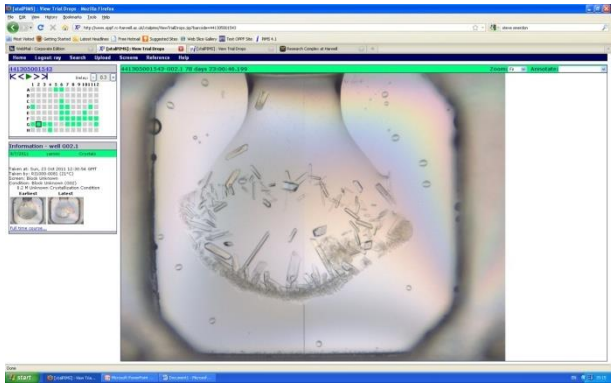
Nettleship et al. (2009) Methods Mol. Biol., 498, 245-263

Semi-automated protein crystallization



- Sitting drop experiments set-up using Cartesian dispenser (x 2).

- Formulatrix imaging systems (2 x 1000 plates @ RT; 1 x 1000 plates @ 4 ^{0} C): visible and UV.



Walter, et al. Acta Cryst. (2008) F64:14-8.
Walter, et al. Acta Cryst. (2005) D61:651-7.
Walter, et al. J. Appl. Cryst. (2003) 36, 308-314.
Brown, et al. J. Appl. Cryst. (2003) 36, 315-318.

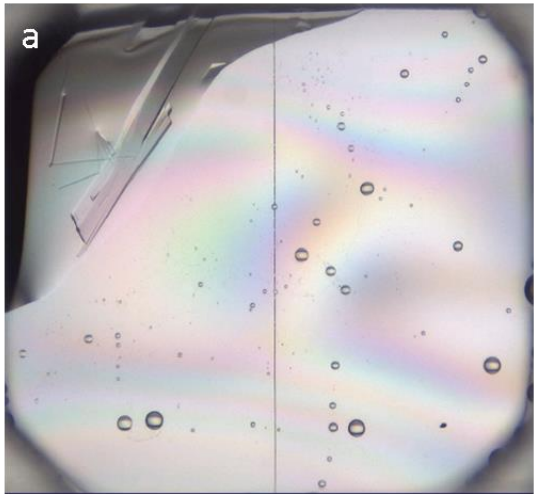
PPF-UK

Daniel, E., et al. (2011) J. Struct. Biol. 75(2):230-5

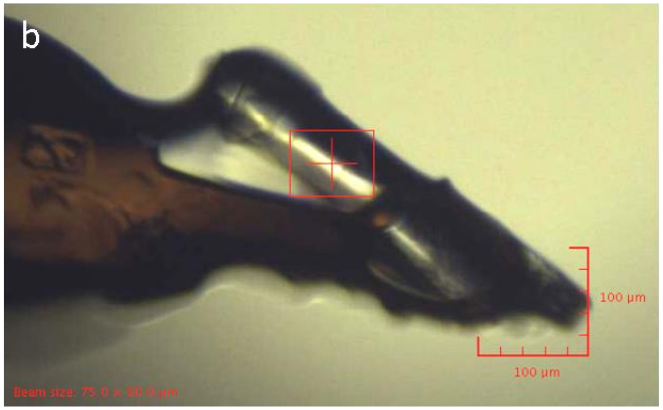
Screening crystallization plates *in situ*

(collaboration with Martin Walsh, Diamond Light Source)

- Routinely used to screen crystals in primary crystallization plates.
- Diffraction used to prioritise conditions for optimisation or for data collection.



Diffraction to 2.7Å in plate on beamline I04-1.



Cryo-protected in loop diffraction to 1.7 Å on beamline I04-1.

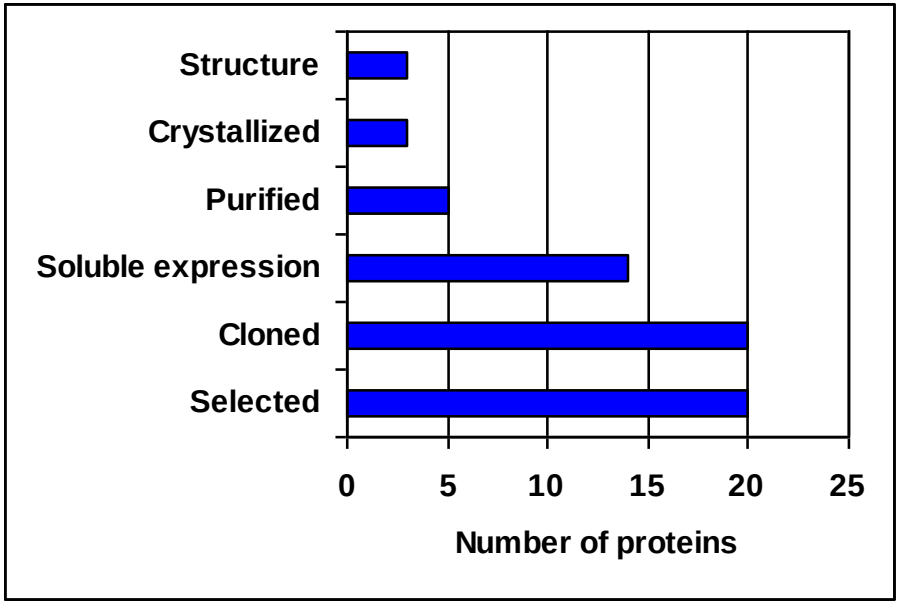
PPF-UK

Lobley et al. 2012, Acta Cryst. F 68, 1427- 1433.

diamond

Identification of novel anti-microbials targeted at gram-negative bacteria

(Collaboration with Bill Hunter. Dundee University)



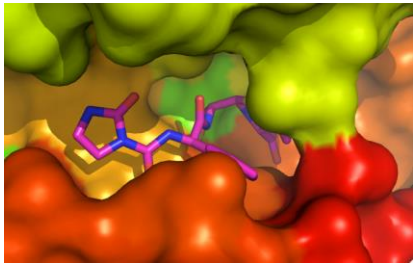
- PCR to crystal trials in 3 weeks.
- First structure in total of 8 weeks.
- 3 Structures (10 including complexes).

PPF-UK

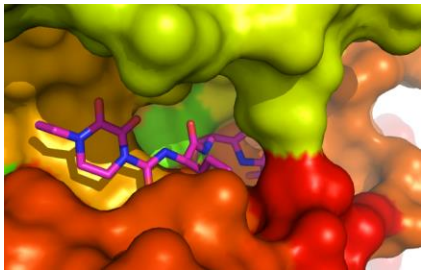
aeropath

Collaborative Project: FP7-HEALTH-2007-B Duration: 01.11.2008 > 31.10.2012

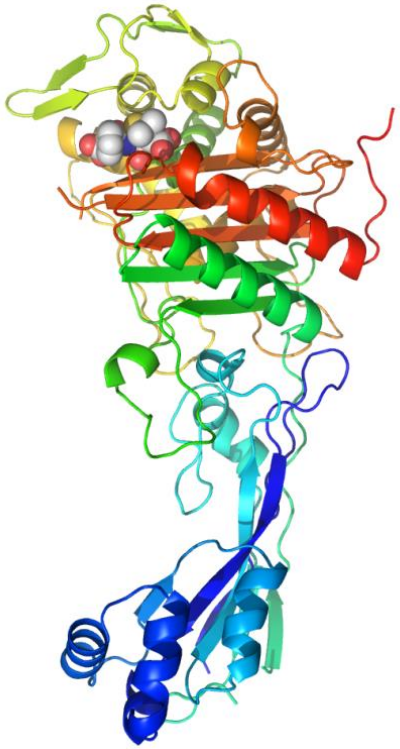
Penicillin-binding proteins



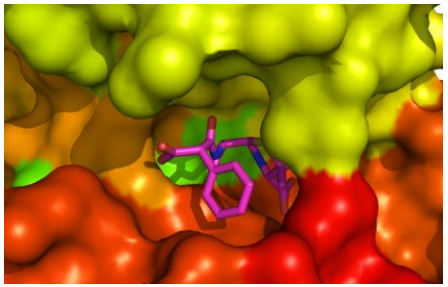
Azlocillin (+16.7°C ΔT_m)



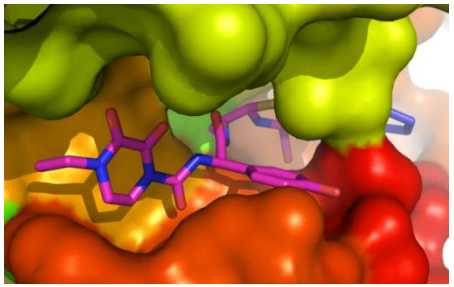
Piperacillin (+17.6°C ΔT_m)



PBP3: carbenicillin



Carbenicillin (+13.2°C ΔT_m)



Cefoperazone (+17.2°C ΔT_m)

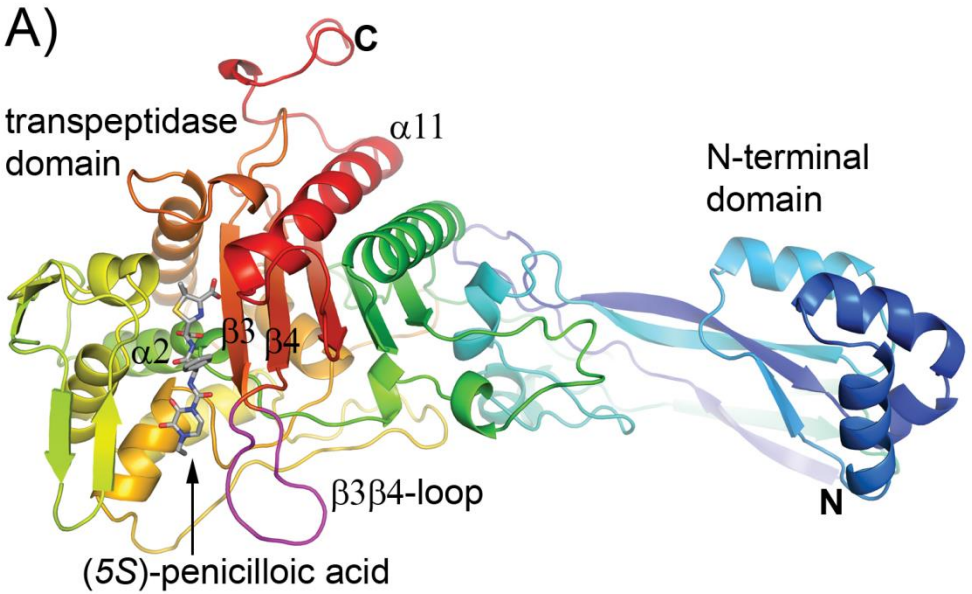
Binding of all β-lactams to PBP3 increases the thermal stability of the enzyme significantly and is associated with local conformational changes which lead to a narrowing of the substrate binding cleft.

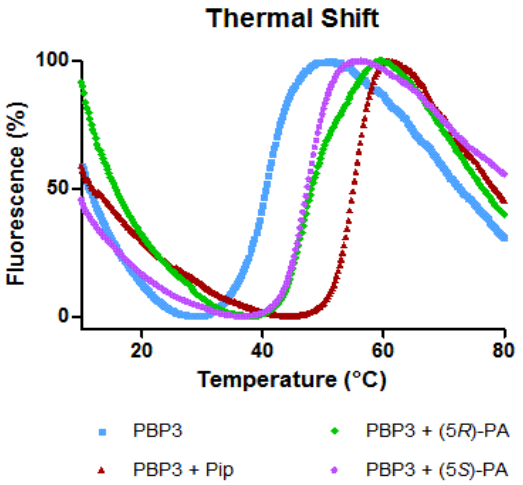
Sainsbury et al. (2011) J. Mol. Biol. 405(1):173-84.
Ren et al. (unpublished)

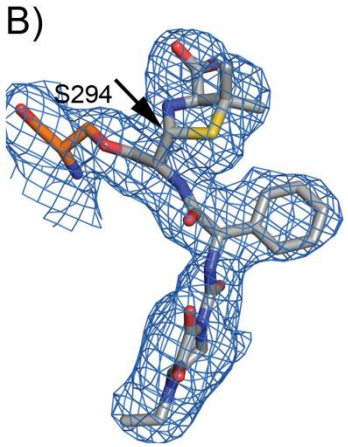
PPF-UK

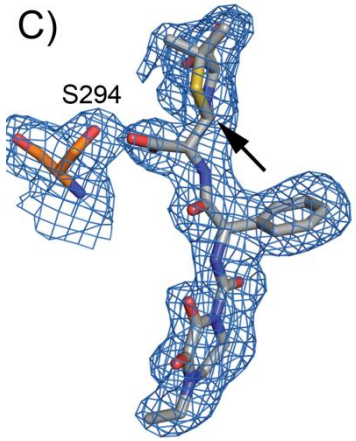
Binding of Penicilloic Acid to *P. aeruginosa* PBP3

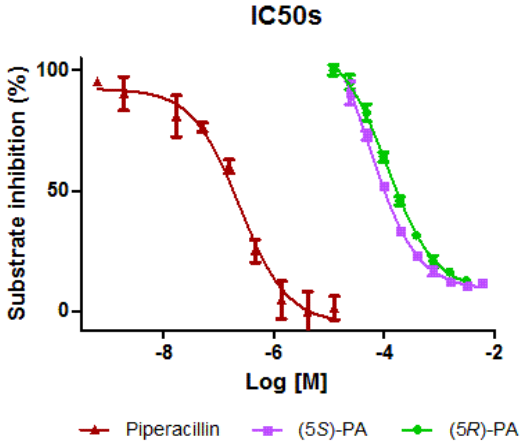
(collaboration with Chris Schofield, Dept. of Chemistry)











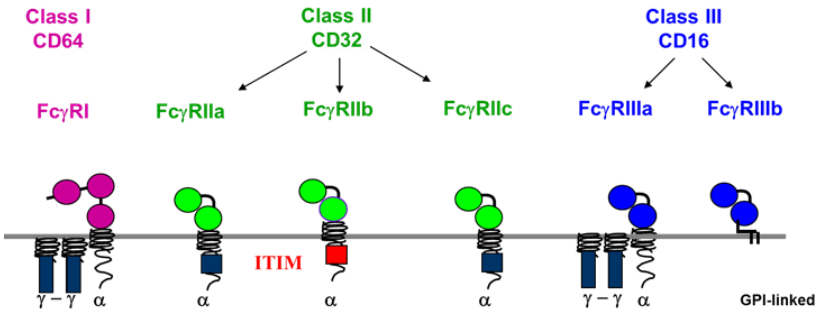
OPPF-UK

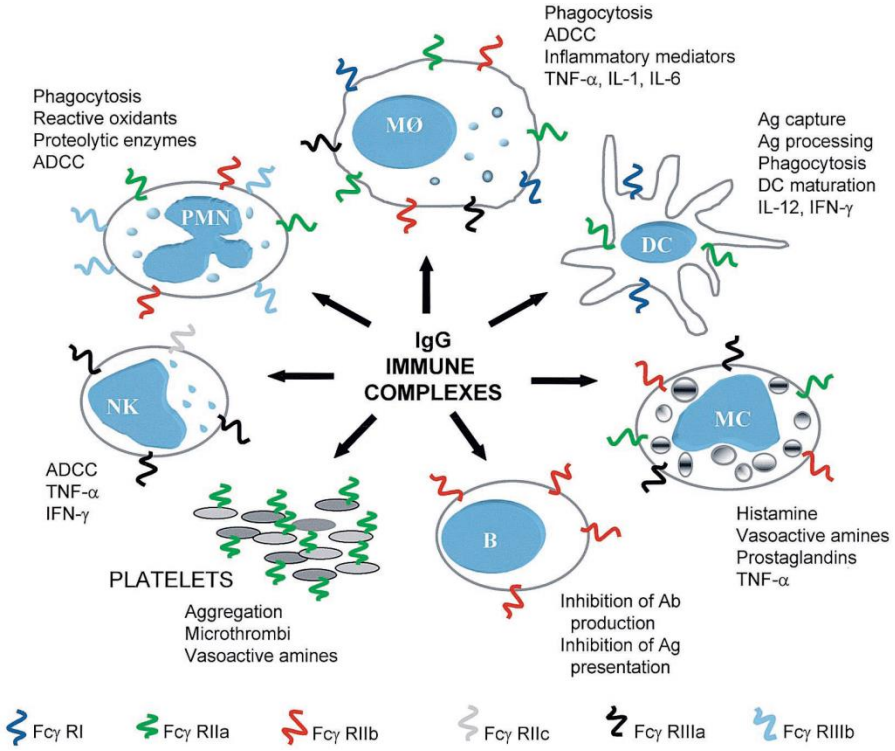
Sander et al. (2013) ACS Chemical Biology 2013 8, 2112-6

FcγRIIIa

(collaboration with Ann Morgan, Leeds Institute of Molecular Medicine Darren Tomlinson & Mike)

- IgG Receptor Fcγ Receptor IIIa
- Implicated in rheumatoid arthritis.
- Binds IgG, CRP, Pentraxins and other small molecules.
- Two polymorphisms studied: 158F and 158V.

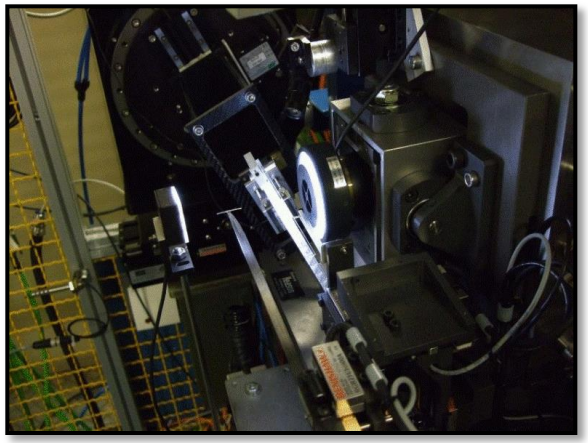




PPF-UK

UNIVERSITY OF LEEDS

FcγRIIIa Polymorphisms

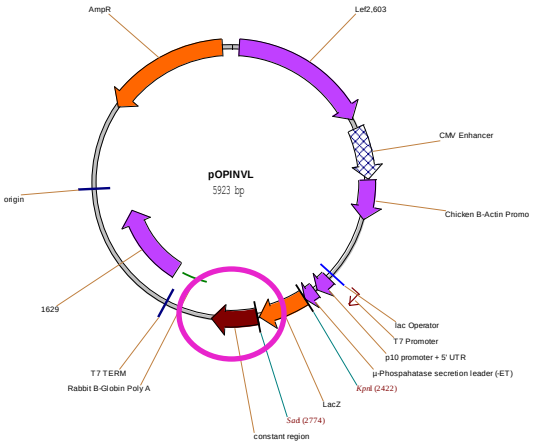


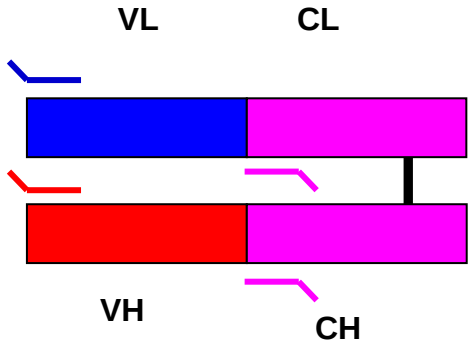
158V - Cyan: 2.4 Å
158F - Blue: 2.4 Å

PPF-UK

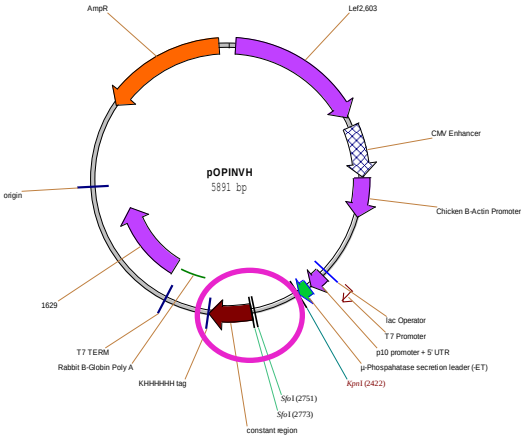
diamond

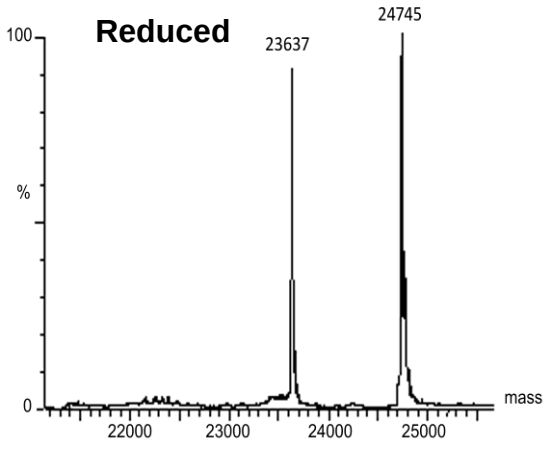
Production of recombinant Fabs

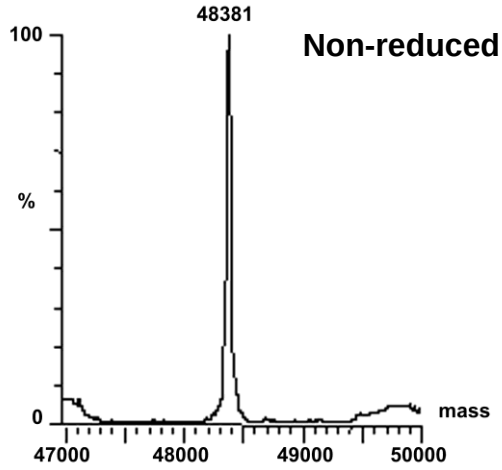




Vh and Vl genes cloned by PCR into vectors containing resident CH1 and CL constant regions.





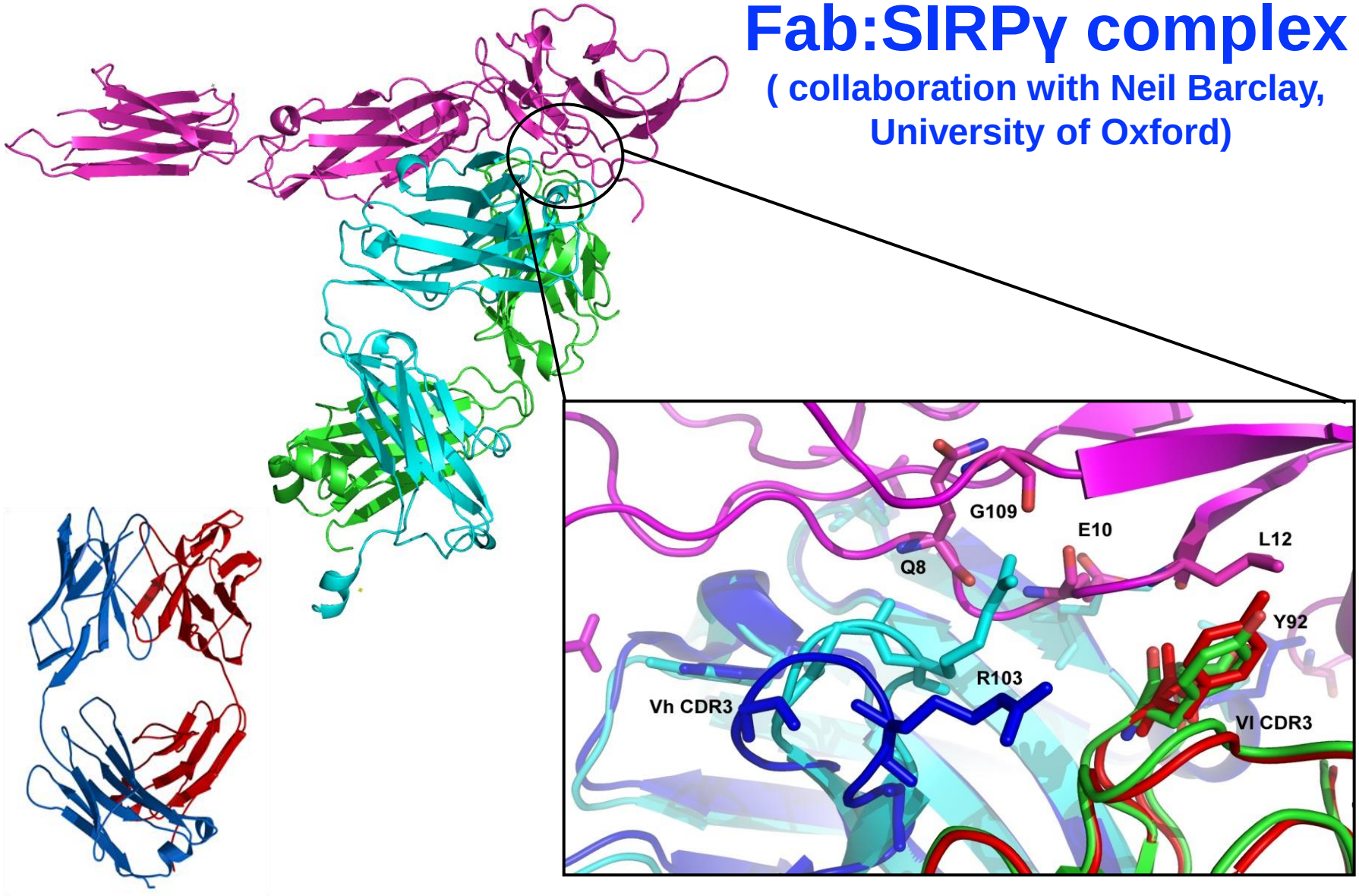


PPF-UK

Nettleship et al. Protein Expression and Purification(2008) 62:83-9
Nettleship et al Methods in Mol. Biol. (2012) 801:137-59.

Fab:SIRPγ complex

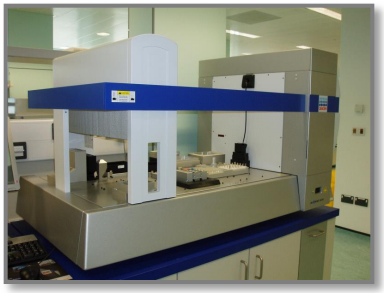
(collaboration with Neil Barclay, University of Oxford)



PPF-UK

Nettleship, et al. (2013)BMC Structural Biology, 4:13.

Summary of workflow



Target selection

Source of templates

Design/order primers

PCR cloning

Insect cell (co)-expression screen

Mammalian (HEK293T) cell (co)-expression screen

E.coli expression screen

Soluble

Intracellular (incl. membrane)

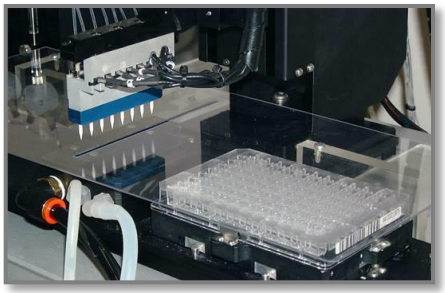
Soluble

Secreted

Scale-up + Purification, characterisation

Soluble

Crystallization



PPF-UK

Acknowledgements

Research Complex at Harwell

CPPF-UK

Ray Owens
Anil Verma
Jo Nettleship
Louise Bird
Nahid Rahman
Heather Rada
Yamini Reddivari
Vicky Arena

diamond

Gwyndaf Evans
Robin Owen
Danny Axford
James Foadi
Martin Walsh

University of Oxford

Neil Barclay
Deborah Hatherley
Chris Scofield
Jurgen Bremen

Leeds Institute of Molecular Medicine

Ann Morgan
Dawn Cooper
Jim Robinson

BioScreening Technology Group

Darren Tomlinson
Mike McPherson
Christian Tiede

The Pirbright Institute 100 Years

Bryan Charleston
Clare Grant
John Hammond

STRUBI

Dave Stuart
Liz Fry
Jingshan Ren
Abhay Kotecha

MRC Medical Research Council

BBSRC bioscience for the future