

Ian M. Berry,^a O. Dym,^b
R. M. Esnouf,^a K. Harlos,^a
R. Meged,^b A. Perrakis,^c
J. L. Sussman,^b T. S. Walter,^a
J. Wilson^d and Albrecht
Messerschmidt^{e*}

^aDivision of Structural Biology and Oxford Protein Production Facility (OPPF), Wellcome Trust Centre for Human Genetics, Roosevelt Drive, Oxford OX3 7BN, England, ^bIsrael Structural Proteomics Center (ISPC), The Department of Structural Biology, Weizmann Institute of Science, Rehovot 76100, Israel, ^cDivision of Molecular Carcinogenesis, The Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands, ^dYork Structural Biology Laboratory, Department of Chemistry, University of York, York YO10 5YW, England, and ^eMax-Planck-Institute of Biochemistry, Department of Proteomics and Signal Transduction, Am Klopferspitz 18, 82152 Martinsried, Germany

Correspondence e-mail:
messersch@biochem.mpg.de

SPINE high-throughput crystallization, crystal imaging and recognition techniques: current state, performance analysis, new technologies and future aspects

This paper reviews the developments in high-throughput and nanolitre-scale protein crystallography technologies within the remit of workpackage 4 of the Structural Proteomics In Europe (SPINE) project since the project's inception in October 2002. By surveying the uptake, use and experience of new technologies by SPINE partners across Europe, a picture emerges of highly successful adoption of novel working methods revolutionizing this area of structural biology. Finally, a forward view is taken of how crystallization methodologies may develop in the future.

Received 5 December 2005
Accepted 28 July 2006

1. Introduction

The automation and miniaturization of high-throughput protein crystallization, *i.e.* the stages from soluble protein to crystal diffraction, is revolutionizing the process of protein structure determination. Workpackage 4 of the SPINE (Structural Proteomics In Europe) project was created to drive the development and uptake of hardware for robotic handling of nanolitre quantities of protein, to trial new strategies for crystallization and to automate the processes involved in crystal imaging and recognition.

The key developments for this area are

- (i) the implementation of small-volume crystallization methods;
- (ii) the development of novel methods for automated crystal image recognition;
- (iii) the development of systematic methods for optimizing crystal growth and quality;
- (iv) the testing of novel procedures for membrane-protein crystallization.

At the start of the project, very few laboratories in Europe had facilities for nanolitre-scale crystallization: of the larger laboratories, only the Protein Structure Factory (PSF), Berlin, Germany had any technologies in place, while both the Oxford Protein Production Facility (OPPF), Oxford, England and the Architecture et Fonction des Macromolécules Biologiques (AFMB), Marseille, France had placed orders. In contrast, the well established NIH-funded PSI (Protein Structure Initiative) structural genomics initiatives in the USA had invested heavily in robotics, including nanolitre-scale robotics in some of the PSI centres (see, for example, Lesley *et al.*, 2002; Stevens, 2004 and references therein). Similar levels of investment in robotics had been made in Japan as well as in the commercial sector. The PSI pilot phase has been reviewed as successful with more than 1200 structures solved and the demonstration that structural genomics pipelines can be

Table 1

Survey of crystallization robotics used across the SPINE consortium and elsewhere in Europe.

Laboratory	Master block	Reservoir	Protein drop	Reservoir drop	Optimization	Additives
Amsterdam	Tecan	Hydra	TTP Mosquito	TTP Mosquito and Tecan	TTP Mosquito, Tecan (and manual)	TTP Mosquito
Oxford	Qiagen Biorobot 8000	Hydra	Cartesian MicroSys 4000	Cartesian MicroSys 4000	MWG Robogo and Cartesian MicroSys 4000	Cartesian MicroSys 4000
Berlin	Zinsser Lissy XXL	Hydra Plus One	Hydra Plus One		Zinsser Lissy XXL and Hydra Plus One	
Weizmann	Tecan EVO	Tecan EVO	Douglas ORYX and TTP Mosquito		Douglas ORYX	Douglas ORYX
Strasbourg Hamburg	Zinsser Lissy 2002	Tecan Genesis Hydra II Plus One	Tecan Genesis Hydra II Plus One	Tecan Miniprep Zinsser Lissy 2002	Tecan Genesis Zinsser Lissy 2002 and Robodesign	Tecan Genesis Hydra II Plus One
Grenoble Marseille	Tecan Tecan Genesis	Tecan Tecan Genesis	Cartesian PixSys 4200 Cartesian MicroSys 4000	Tecan Cartesian MicroSys 4000	Tecan Tecan Genesis	Tecan Genesis
Munich	Cartesian Microsys 4000XL	Cartesian Microsys 4000XL	Cartesian Microsys 4000XL	(Manual)	(Manual)	(Manual)
GlaxoSmithKline, UK	Hamilton Star	Hamilton Star	Cartesian, TTP Mosquito	Hamilton Star	Cartesian, TTP Mosquito	Cartesian
CCLRC Laboratory, Warrington, UK	Hamilton Microlab Starlet	Innovadyne 96 plus 8 (Manual)	Innovadyne 96 plus 8		Hamilton Microlab Starlet	
University of Glasgow, Glasgow, UK						
SSPF, St Andrews, UK	Hamilton Microlab Star	Hamilton Microlab Star	Cartesian Honeybee	Hamilton Microlab Star	Cartesian Honeybee	Cartesian Honeybee
University of Zurich, Switzerland	Zinsser Lissy	Tecan Aquarius	Cartesian MicroSys 4000	Cartesian MicroSys 4000	Zinsser Lissy	Tecan Aquarius
Schering AG, Germany	(Manual)	(Manual)	Cartesian		Cartesian (and manual)	Cartesian (and manual)

Table 2

Protein quality-assurance (QA) techniques used by different laboratories.

QA method	No. of laboratories
Mass spectrometry	11
Gel filtration	10
UV spectroscopy	9
Concentration estimation	7
Dynamic light scattering (DLS)	5
Hampton Pre-Crystallization Test (PCT)	5
In-house protein concentration test	3
Electron microscopy	1
Biorad protein stain	1
Static light scattering (SLS)	1
IR spectroscopy	1
Trial-and-error protein concentration test	1
Other protein concentration test	1
No protein concentration test	1

constructed and scaled up and that HTP operation worked for many proteins (Novell & Berg, 2005).

The outstanding success of nanolitre-scale crystallization in obtaining more crystals from less protein (Santarsiero *et al.*, 2002; Brown *et al.*, 2003; Hosfield *et al.*, 2003), thereby leading to more structures and better science, has driven the adoption of these technologies across Europe, including in smaller laboratories. In order to facilitate dissemination of information and technology transfer from the larger laboratories, three workshops were organized to study some of the different technologies available: the first looked at the Fluidigm microfluidic crystallization technique (Hansen & Quake, 2003), the second looked at microbatch protein crystallization under oil (Chayen *et al.*, 1992) and the third looked at auto-

mated imaging of crystals (Wilson, 2002; Spraggon *et al.*, 2002). The lessons learned from these are presented in this review covering nanolitre-scale crystallization, plate storage and automated imaging. As part of the preparatory work for the workshop on automated imaging held in Oxford in June 2005, a survey was made of SPINE partners and other European laboratories in order to establish the nature and extent of technology transfer. The results of this survey into modern crystallization techniques form the basis of the following report.

2. Current state, performance analysis and new technologies

Nanolitre-scale crystallization technologies that were scarce at the start of the project are now common in the laboratories of the SPINE partners (Table 1). Many of the choices of equipment were influenced by cooperation and advice from those SPINE laboratories that had already invested in equipment and/or had contributed to hardware, software and methodological developments during the lifetime of SPINE. The developments and the improvements can be split into a number of categories, which are now considered separately.

2.1. Quality assurance

Checking the concentration, quality and homogeneity of each protein solution is an important first step in the process of crystallization. The usage of different quality-assurance

Table 3

Summary of liquid-handling robots used by different laboratories.

Liquid-handling system	Description	Manufacturer URL
Cartesian	Solenoid ink-jet dispensing system	http://www.genomicsolutions.com
Douglas	Drop dispenser for crystallization under oil and/or sitting drops	http://www.douglas.co.uk
Hamilton	Syringe-based liquid-dispensing platform	http://www.hamiltoncomp.com
Hydra	96-tip syringe-based liquid dispensing	http://www.matrixtechcorp.com
Hydra Plus One	96 + 1 tip syringe-based liquid dispensing and solenoid	http://www.matrixtechcorp.com
Innovadyne	96-way syringe dispenser plus eight solenoid dispensers	http://www.innovadyne.com
Mosquito	Positive-displacement nanolitre contact dispenser	http://www.ttplabtech.com
MWG Robogo	Single-channel syringe-based dispenser	No longer available
Qiagen Biorobot 8000	Syringe-based liquid-dispensing platform	http://www.qiagen.com
RoboDesign	Bird-feeder technology	http://www.robodesign.com
Tecan	Syringe-based liquid-dispensing platform	http://www.tecan.com
Zinsser Lissy	Syringe-based liquid-dispensing platform	http://www.zinsser-analytic.com
Manual	Multi-channel pipettes	

Table 4

Number of laboratories using different crystallization methods.

Crystallization method	No. of laboratories
Vapour diffusion (hanging drop)	9
Vapour diffusion (sitting drop, 1 drop per well)	14
Vapour diffusion (sitting drop, 2 drops per well)	5
Vapour diffusion (sitting drop, 3 drops per well)	9
Microbatch	4
Fluidigm	4

Table 5

Summary of plate suppliers.

Tray type	No. of laboratories	Supplier URL
Greiner	21	http://www.gbo.com
Linbro	3	http://www.jenabioscience.com
Corning	3	http://www.corning.com
Intelliplate	2	http://www.artrobbinsinstruments.com
Nunc HLA	1	http://www.douglas.co.uk
Douglas Vapor Batch Plate	1	http://www.douglas.co.uk
Hampton VDX	1	http://www.hamptonresearch.com

Table 6

Summary of sealing-tape products.

Sealing tape	No. of laboratories
Greiner (3M)	9
Hampton ClearSeal Film	3
Henkel Crystal-Clear	1
Zymark	1
Corning 6575	1
Hamilton Star	1
HJ Bioanalytik 100 m rolls (3M)	1

(QA) techniques across the laboratories in the survey is shown in Table 2 (see also Geerlof *et al.*, 2006).

Mass spectrometry is the most favoured QA measure as it gives confirmation of the protein mass in solution, providing reassurance that it is a uniform species, highlighting any unexpected proteolysis and quantifying selenomethionine incorporation. In contrast, testing for monodispersity by dynamic light scattering (DSL; for this particular application

of DSL, see Ferré-D'Amaré & Burley, 1997) is less common. Testing to determine the optimum protein concentration by trial experiments with a few carefully selected conditions has also become more common, with almost all laboratories performing some sort of test prior to large-scale crystallization screening.

2.2. Nanolitre-dispensing robots

The hardware systems used for liquid handling by the different laboratories participating in the survey are shown in Table 3. The survey revealed that there are a wide variety of models on the market, with no manufacturer having market dominance. There are different underlying technologies which may vary in their suitability to different experimental requirements and throughputs. Accurate dispensing of small drops is challenging, not least because of problems with sample cross-contamination and drop evaporation, which led laboratories and manufacturers to come up with different solutions. Accurate dispensing is complicated further by the diversity of screening conditions in general use which differ greatly in viscosity, volatility and surface tension.

The Cartesian Microsys and the TTP Mosquito are two instruments that are widely used to dispense nanolitre droplets. Both systems work well and require only about 15 µl of protein material for a full 96-well plate. The Cartesian is based on ink-jet technology using pressure and fast solenoid valves to eject the drops from special ceramic tips (Rose, 1999). As the tips need to go through a washing cycle between pipetting steps, setting up a 96-well plate is relatively slow and takes approximately 15 min. The Mosquito uses small disposable pipettes which are arranged on a belt and are then automatically fed into the instrument from a spool. The liquid is aspirated and dispensed in a positive-displacement mode by moving a small steel pin within a plastic capillary-type tip. A new tip is used for each pipetting step which eliminates any possible carry-over, but also increases the running costs. Setting up a 96-well plate is fast and only takes 1–2 min. Dispensing into hydrophobic plates with the Mosquito can be problematic as with this method the drop may not be able to form sufficient contact with the plate surface so that it can be released from the tip.

2.3. Crystallization methods

Across the survey, all common methods of crystallization are practiced, although sitting-drop vapour-diffusion experi-

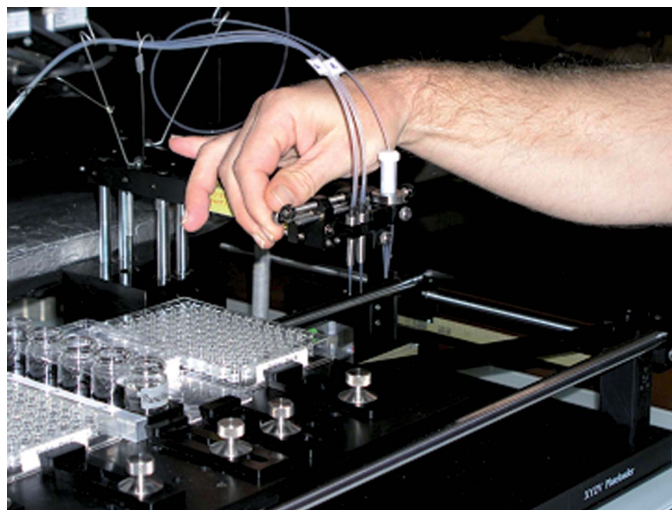


Figure 1
Douglas Oryx-6 robot used for microbatch crystallization under oil and sitting-drop crystallization by the Israel Structural Proteomics Center. The robot shown was modified 'in-house' by the addition of a double-pipette apparatus which approximately doubles the rate of drop delivery in screening experiments. The hand is adjusting the two-dimensional register on the dispenser.

ments are most common (Table 4). This may partly be a consequence of this method being the simplest to automate. In vapour-diffusion experiments liquid transfers, through the vapour phase, from a small drop of protein to a large volume of reservoir solution (or *vice versa*), which thereby increases the concentration of precipitant and/or protein and induces crystallization. It can be used with protein droplets placed as sitting or hanging drops (see, for example, Weber, 1997).

Microbatch crystallization under oil (Chayen *et al.*, 1992) is most frequently used in situations where other techniques have not been successful, although it is also used as a primary method and has been automated at the Weizmann Institute (Fig. 1). This technique relies on slow diffusion from the crystallization drop through a protective oil layer. A potential drawback of the method is that the experiment has no defined end-point, which may in turn limit crystal lifetime. Conversely, when the optimum protein concentration is not established this may be a useful feature and may in part explain the result from the survey which suggests that some proteins only crystallize in microbatch experiments (some others only crystallize *via* vapour diffusion), even when using the same screens.

The Fluidigm microfluidic technique (Hansen & Quake, 2003) uses free-interface diffusion within chips made specifically for screening 96 crystallization conditions whilst requiring only 1.5 μ l of protein solution in total (Fig. 2). The process consists of two phases: the first phase is free-interface

diffusion, while the second phase resembles microbatch crystallization since the experiment reaches no end-point. One advantage is that results are produced rapidly, with experiments typically lasting about one week, whereas one drawback is that it can be difficult to translate crystallization hits to higher volume solutions in order to grow usable crystals for diffraction experiments. Furthermore, the cost per 96-condition chip is more than tenfold higher than other setups and a cost/benefit analysis is necessary. This is being performed SPINE-wide as a way to maximize the diversity of samples and opinions and thus the validity of the conclusions, with the NKI, Amsterdam coordinating the effort.

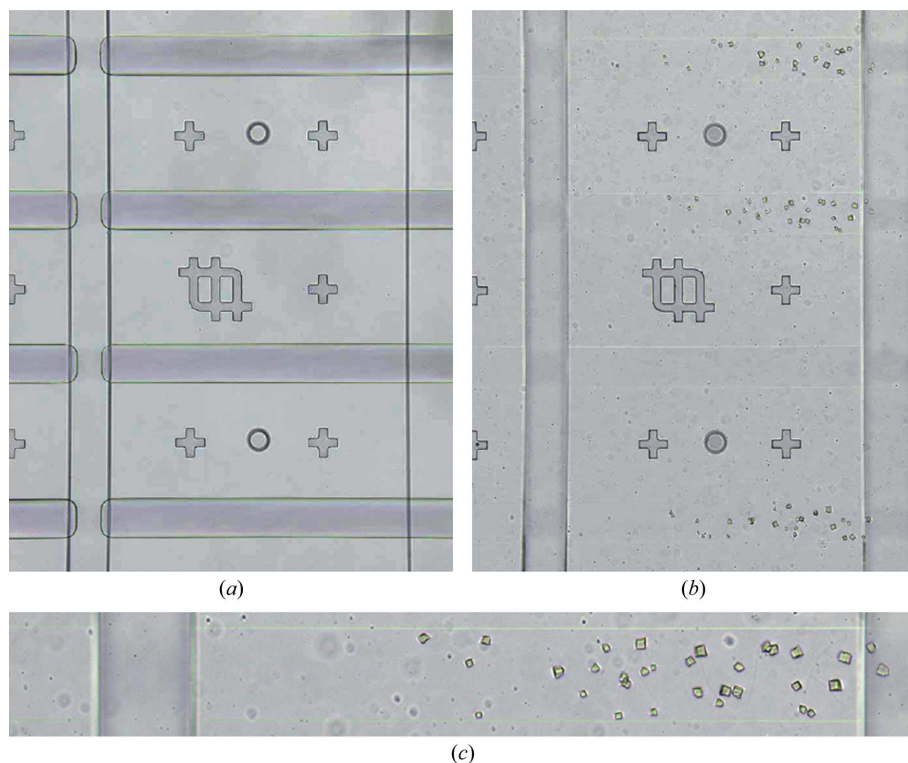


Figure 2
Images of protein channels, as marked by horizontal lines, of Fluidigm chips. (a) Four channels before loading, (b) the same channels after the experiment and (c) an enlarged view of one of the channels showing well formed crystals. During an experiment, precipitant solutions diffused into the channel from the left; note the gradient in crystal quality towards the right, presumably arising from the variation in protein concentration. Images were taken with the Fluidigm 'AutoInspeX Workstation'. Channel dimensions are 0.1 \times 0.7 mm.

2.4. Crystallization plasticware

A variety of tray types and sealants are used to set up vapour-diffusion experiments and these are summarized in Tables 5 and 6. There are currently problems with several brands of sealing tapes in that they become cloudy owing to volatile reagents and hence prevent imaging. Furthermore, those tapes which use encapsulated sealant appear

to have a limited shelf-life. However, SPINE partners are collaborating with the manufacturers to test improved products.

A wide variety of drop volumes are in use, ranging from 50 nl through to 1 μ l. Most laboratories did not report problems with excessive evaporation once plates had been sealed, indicating that the tapes were effective in sealing the wells. In general, most laboratories only use a single concentration of protein per plate; however, some use each shelf in any given well for a different protein. Most plate types were found to have good optical quality for the imaging methods and systems employed. However, work is still required on the surface and optical properties of trays to optimize their use

with all automated systems and with various illumination modes.

Oxford have collaborated with Greiner Bio-One GmbH to produce crystallization plates suitable for membrane proteins (and indeed for any crystallization experiments where the liquid has a low surface tension). Experiments with a variety of Greiner plates were performed and the best one is now commercially available from Greiner Bio-One GmbH (96-well CrystalQuik™ Plus). The differences between a normal plate and the new hydrophobic plate are demonstrated in Fig. 3. In both cases, 100 nl drops of 50 mM *n*-octylglucoside solution were dispensed and the plates were subjected to the normal OPPF storage and imaging procedure. The diameter of the

drops in the normal plate (Fig. 3*a*) is much larger than that of drops in the new plate (Fig. 3*b*); indeed, it is difficult to see the actual drop with the normal plate because the drop is so shallow. With both plates there were some drops which wicked into the walls, but this is typical of all such plate setups. The new plates can be used in standard crystallization procedures without any modifications and are suitable for crystallizations in the presence of detergents.

2.5. Crystallization screening kits and layouts

The Hampton Research screens are the most widely used, followed by the Molecular Dimensions and Emerald BioSystems screens. Half the survey respondents used commercially available optimization kits, while the remainder had developed bespoke approaches to optimization (see below). There is no universal way in which these kits are combined to form crystallization screens, but most involve setting up so-called 'Master Blocks'. These contain 1–2 ml of 96 different screening solutions, sufficient to set up 10–20 plates. Some laboratories adopt the approach of having 'standard' blocks which combine a unique set of screen conditions from different manufacturers' screen sets. For instance, Oxford use seven standard blocks made up from commercial screens (*e.g.* 'Block 1' comprises 96 solutions selected from Hampton Research Crystal Screen and Crystal Screen 2; Walter *et al.*, 2003). Table 7 shows the variety and frequency of screen usage amongst the SPINE partners.

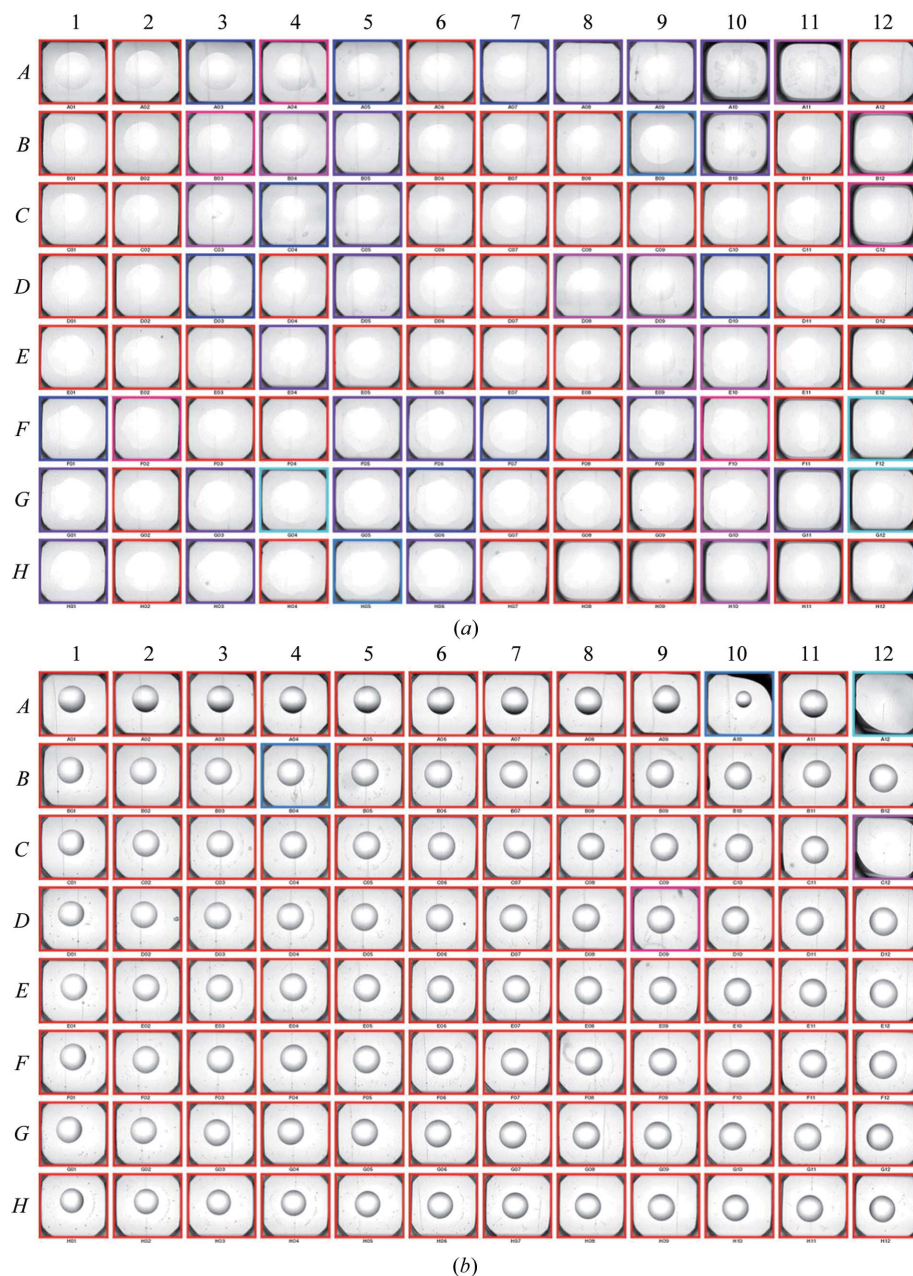


Figure 3

Comparison of 100 nl droplets containing 50 mM *n*-octyl-glucoside dispensed (*a*) into a standard plate and (*b*) into the new CrystalQuik™ Plus plate for membrane proteins from Greiner. Images were recorded at the Oxford node using the Veeco Oasis 1700 imaging system.

Table 7

Usage of commercially available crystallization screens.

Manufacturer	Screen name	No. of conditions	No. of laboratories
Emerald BioSystems, http://www.emeraldbiosystems.com	Cryo 1	48	1
	Cryo 2	48	1
	Wizard 1	48	11
	Wizard 2	48	11
Fluka, http://www.sigmaaldrich.com	Automatic Screening Kit	96	1
	Crystallization Extension Kit	50	1
	Crystallization Kit for Membrane Proteins	50	1
	Crystallization Cryo Kit	50	1
	Crystallization Basic Kit	50	1
	Crystallization Low Ionic Strength Kit	50	1
	Crystal Screen	50	13
	Crystal Screen 2	50	13
Hampton Research, http://www.hamptonresearch.com	Grid Screen Ammonium Sulfate	24	5
	Crystal Screen Cryo	50	7
	Memfac	48	8
	Grid Screen MPD	24	5
	Natrix	48	9
	PEG/Ion Screen	48	17
	Grid Screen PEG 6000	24	7
	Quik Screen	24	4
	Grid Screen Sodium Chloride	24	4
	Salt RX	96	7
	Index	96	11
	JCSG screen	96	1
	1–10	240	3
	HTS1	96	1
	HTS2	96	1
Molecular Dimensions, http://www.moleculardimensions.com	3D Structure Screen	48	2
	Clear Strategy Screen 1	24	4
	Clear Strategy Screen 2	24	4
	MacroSol	48	1
	MemStart	48	4
	MemSys	48	4
	NR-LBD	48	3
	NR-LBD Extension	48	3
	PACT Premier	96	4
	Structure Screen 1	50	3
	Structure Screen 2	50	3
	Stura Footprint	96	5
	Zeta Sol	98	1
	Anions, PEGS, Ammonium Sulfate, MPD	96	2
Nextal, http://www.nextalbiotech.com	Classics	96	2
	The Cryos	96	2
Bespoke solution	Own creation	96	3

2.6. Scale-up and optimization

Once crystals or crystallizable precipitates have been found in a screening condition then, unless the crystals are particularly good and immediately usable, researchers need to either (i) scale up the volume of the trial in an attempt to obtain larger crystals or (ii) optimize the crystallization condition around the 'hits' in order to create better crystals. Which approach is taken is very much laboratory- and researcher-dependent. However, the survey shows that in many cases the drop-dispensing (but not necessarily the making up of the reservoir solution) employs robotics and this has led to a higher success rate in terms of both crystal quality and reproducibility.

2.7. Tray storage and imaging systems

Since the inception of SPINE, tray storage has increasingly become an issue as more experiments are created through the

use of high-throughput robotic systems. A key feature for the usability of such systems is integration with an image-acquisition system that allows automated and scheduled imaging of plates. During the period of the project, such systems have rapidly developed and become more cost-effective. The majority of laboratories perform most crystallization experiments at room temperature. However, increasingly laboratories are also investigating the effect of different temperatures such as 277, 285 and 310 K, especially for high-value targets. Table 8 shows details of the crystal-imaging systems that are currently in use in the laboratories that responded to the survey. There has been a large increase in the variety of commercial systems available and also the degree of automation has increased, particularly related to imaging capabilities. It should be noted that Veeco are no longer selling imaging systems. At present there are a number of commercial systems available and it seems that several of these perform at least adequately. Unfortunately, there has as yet been no systematic comparison of them. The imaging schedule defined in each laboratory is determined by several factors: the amount of storage available (how

long the tray can be kept in the storage system before the space is needed for newer plates), the importance of the tray's contents (laboratories may wish to keep high-value samples for longer periods and image trays more regularly in the hope of obtaining a crystal) and the speed of the imaging device. The result is that the time period during which trays are imaged at regular intervals varies between one month and a year. For those sites that image for longer periods, it has been found that useful crystals can still start to appear even late into an imaging schedule.

Viewing images produced by imaging systems is achieved in three main ways: through a manufacturer-supplied interface specific to each machine, through a bespoke interface designed by the laboratory (*e.g.* Mayo *et al.*, 2005) or through Microsoft Windows Explorer. As experience with these systems has grown, a clearer idea has emerged of their requirements in terms of accessibility, ease of use and sources of information. These are now being adopted for a new

Table 8

Crystal-imaging systems described in the survey.

Imaging system	No. of laboratories	Image formats	Image size	Resolution (μm per pixel)	Type	Image depth (bpp)	z-slicing	Birefringence	Focus	Schedule	Additional imaging	Manual imaging
Veeco Oasis 1700†	1	BMP	700 × 750	1.0	Greyscale	8	None	None	Fixed	Automated	Yes	Yes
Veeco Oasis 1750†	1	BMP	700 × 750	1.0	Greyscale	8	None	None	Fixed	Automated	Yes	Yes
Veeco Oasis LS3†	1	BMP	700 × 750	1.0	Greyscale	8	Available	None	Fixed	Automated	Yes	Yes
BioTOM	2	JPG	1024 × 1024	Unknown	Colour	8	Available	Available	Automatic	Automated	Yes	Yes
BioSTORE‡												
Bruker Crystal Farm§	2	JPG	1024 × 1024	Unknown	Colour	8	Available	None	Automatic	Automated	Yes	Yes
TriTek	1	BMP, JPG	1600 × 1200	3.0	Colour	8	Available	Available	Manual	Manual	No	Yes
CrystalPro¶												
Homemade (PSF)	1	JPG, TIFF	1000 × 1000	Unknown	Greyscale	8	None	None	Automatic	Automated	Yes	Yes
Homemade (IGBMC)	1	JPG	1392 × 1040	4.0	Colour	8	Available	Available	Manual	Manual	No	Yes
Homemade (Schering AG)	1	JPG	2000 × 1500	Unknown	Colour	8	None	Available	Manual	Manual	Yes	Yes
RoboDesign	1	JPG	2560 × 1920	1.0	Colour	8	Available	Available	Automatic	Automated	Yes	No
Minstrel I††												
RoboDesign	2	JPG	2560 × 1920	1.0	Colour	24	Available	Available	Automatic	Automated	Yes	Yes
Minstrel III††												
RoboDesign	1	BMP, JPG, PNG	2560 × 1920	1.0	Colour	16	Available	Available	Automatic	Automated	Yes	No
Minstrel IV††												
Thermo Rhombix‡‡	2	BMP, JPG, TIFF, PNG	1360 × 1036	0.6	Colour	24	Available	Available	Automatic	Automated	Yes	Yes

† <http://www.veeco.com>. ‡ <http://www.biotom.net>. § <http://www.bruker-axs.de>. ¶ <http://www.proteincrystalimaging.com>. †† <http://www.robodesign.com>. ‡‡ <http://www.rhombix.com>.

generation of viewers, for example as part of the PIMS (Protein Information Management System) project funded by the UK BBSRC and the European Commission (<http://www.pims-lims.org>). One important consideration for viewing systems is integration with image-analysis and human annotation tools. This integration creates the prospect of useful data mining from the huge databases of images that are being generated (the largest single database amongst the SPINE partners now comprises over 36 million images).

A variety of imaging techniques are now available in imaging systems, but not all are actively being used. For example, where z-slicing (taking many images of a drop at different focal planes) is available it is only used in 60% of laboratories; equally, where birefringence imaging is available it is only used in 40% of laboratories. It is likely that the use of simplified protocols reflects, at least in part, the time overheads incurred, but does suggest that these protocols are of marginal value.

3. Image analysis

Automation has led to a dramatic increase in the number of crystallization experiments performed in most laboratories, especially structural genomics or proteomics centres, where many thousands of experiments are produced on a daily basis. Images of these experiments, also often acquired automatically, must be assessed repeatedly over a period of time.

Inspection of these many thousands of images by eye is becoming increasingly impractical (especially with multiple lighting effects, z-slicing and dark-field images), hence driving the development of tools to automate analysis.

Several systems for analyzing and classifying images are being developed (see, for example, Bern *et al.*, 2004; Cumbaa *et al.*, 2003; Spraggon *et al.*, 2002; Wilson, 2004). Whilst the fundamental aim is the detection of crystals, reliable classification of other outcomes potentially provides vital feedback to aid the development of optimization protocols. In the absence of crystals, the occurrence of microcrystals, spherulites or 'sea-urchins' indicate conditions that are likely to be very close to those required for crystal growth, while conditions resulting in phenomena such as phase separation or crystalline precipitate may also be optimized. On the other hand, experiments resulting in heavy amorphous precipitate or denatured protein suggest that the conditions are unlikely to lead to crystal growth.

Current image-analysis systems are generally specific to a particular imaging system. Within the SPINE project the aim was to produce software that could be adapted to multiple systems through retraining rather than reprogramming. A variety of images from different imaging systems is depicted in Fig. 4 and shows what a difficult aim this is. Lighting and resolution have the most dramatic implications, but there are many other aspects to consider including image formats and colour. Some differences can be dealt with by preprocessing

the images. At present, colour images are converted to grey-scale before analysis. The software must be trained to recognize objects using test image sets and will always require retraining for different imaging systems.

ALICE (automated analysis of images from crystallization experiments) is software developed at York within the SPINE project and will be continued in the EC-funded BIOXHIT project. Like many methods, it relies on edge-detection, just as we are able to recognize an object from a rough outline sketch. *ALICE* uses an edge detector (Sobel detector), in contrast to earlier work of Zuk & Ward (1991) which used the Hough transform (see, for example, Gonzales & Wintz, 1987; Spraggon *et al.*, 2002), which defines the drop boundary by Canny edge detection (Canny, 1986). Analysis proceeds in several stages. Firstly, unwanted artifacts that could cause problems in classification are removed (see Wilson & Berry,

2005). For example, some crystallization plates can have mould defects that appear as lines across the image and since straight lines are an obvious characteristic of crystals, these would produce many false positives in classification. Bubbles can also be identified and eliminated at this stage. Next, the boundary of the well is identified and the image cropped accordingly. A keyword indicating the plate type allows the most widely used crystallization plates to be dealt with efficiently, although other plates are available and shadowing arising from curved wells creates particular problems. Identification of the drop boundary then allows the analysis to be restricted to the appropriate area of the image, speeding up processing dramatically.

Having found the drop, the objects within it are detected and characteristic features that can be quantified are measured and used as variables in classification algorithms. As

for any pattern-recognition problem, the extracted features should have common values within a class and be discriminatory between classes. Unfortunately, this is rarely the case with features extracted from crystallization images. Whilst crystals may be expected to have straight edges and definite angles, there is no fixed pattern that describes every crystal; indeed, apparently round crystals have been found that diffract beautifully. Conversely, wrinkles in the protein skin on the drop often display similar angular features, while denatured protein can create an assortment of strange effects. In general, classification variables may be features of the crystallization drop as a whole or of individual objects within the drop. Most methods take the drop-based approach, quite possibly because the task of defining a training set for individual objects from each class is extremely tedious. In *ALICE*, individual objects are identified within the crystallization drop and evaluated separately (Wilson, 2002). However, this object-based method does not naturally take advantage of the overall environment or inter-relatedness of the objects and a method to combine the two approaches is being developed. *ALICE* uses a combination of self-organizing maps (see, for example, Kohonen & Somervuo, 2002) and learning-vector quantization to classify objects into groups and assign them to particular types. Results from both of these methods, as well as Bayesian probabilities calculated from the training data, are combined to give a

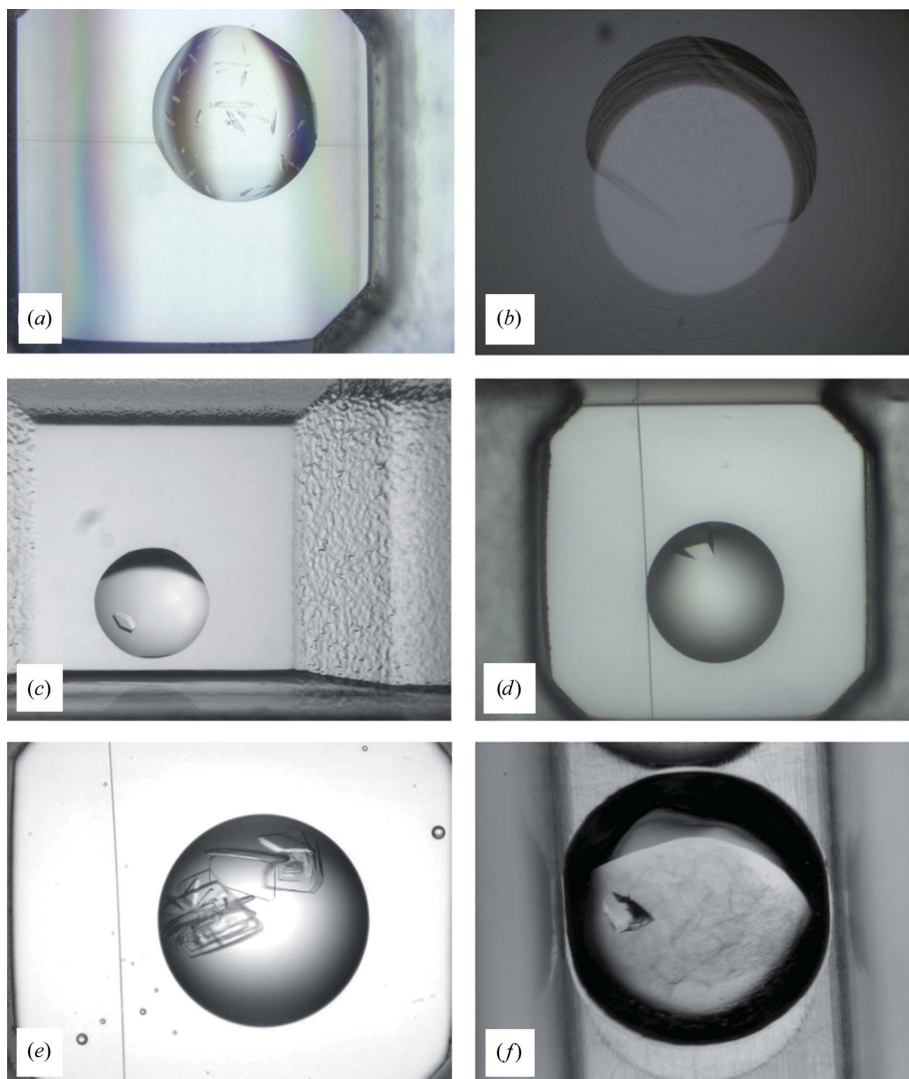


Figure 4

A variety of images from different sources. The image systems and laboratories are (a) the Rhombix imager in St Andrews, Scotland, (b) the Tritex Crystal Pro imaging system at the Weizmann node, Israel, (c) the BioTom imaging system in Amsterdam, the Netherlands, (d) the RoboDesign system at Grenoble, France, (e) the Veeco Oasis 1700 imaging system at Oxford, England and (f) the in-house imaging system at Berlin, Germany.

final class for each object identified within a drop. The final stage involves calculating an overall score for the drop from the classifications of individual objects within it.

ALICE was developed in collaboration with Oxford, where it is now used routinely to annotate images. The software was therefore initially tailored to the OPPF's Veeco Oasis 1700

and 1750 imaging systems, although adaptations to other systems are now under way and pre-processing software has already been supplied to the Amsterdam and Grenoble partners. This software allows images to be cropped to the size of the drop, thereby drastically reducing the storage space required for images. Whilst pre-processing reduces the variation between different imaging systems, the classification algorithms used within *ALICE* must be trained to recognize the variables expected from a particular imaging system. This requires a training set of images, composed of (at least) 200 from each class, to be classified by eye in terms of individual objects. This task is facilitated by a graphical user interface that allows objects within an image identified by *ALICE* to be classified by the user at the click of a mouse button. The creation of training sets is still extremely tedious, although it should not need to be performed more than once for each new system.

Although classification by an alert and expert human is always likely to be superior, the volume of images being generated means that any method of automated analysis that can reduce the number of images needing human inspection promises to increase the quality of classification. Equally important, however, is the potential for the development of optimization procedures. Automatic classification of images will mean that many more 'near-misses' will be annotated and so even with rather imperfect classification it may be possible to guide the researchers towards conditions which produce better crystals.

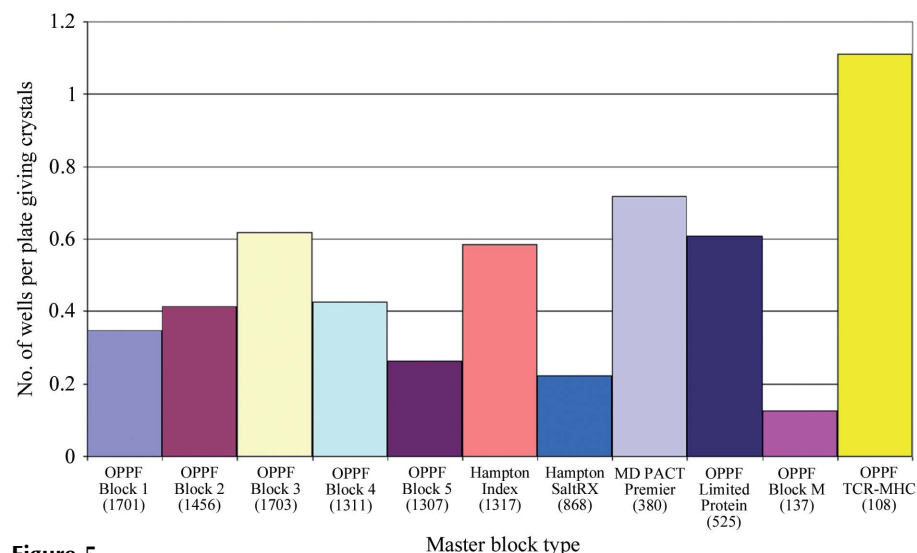


Figure 5

Effectiveness of different screening blocks at producing crystals at the Oxford node. Each block contains 96 reagents and drops containing crystals were annotated manually. The number of plates set up using each block is given in parentheses. The blocks are based on conditions from the following commercial screens: Block 1, Hampton Crystal Screens I and II; block 2, Emerald Wizard Screens I and II; block 3, Hampton PEG/Ion, Grid Screens PEG 6000 and AS; block 4, Hampton Natrix Screen and Crystal Screen Cryo; block 5, Hampton PEG/LiCl, NaCl, MPD and Quik Screen Phosphate. Block M consists of Molecular Dimensions MemStart and MemSys screens and is focused on screening membrane proteins (Walter *et al.*, 2003, 2005).

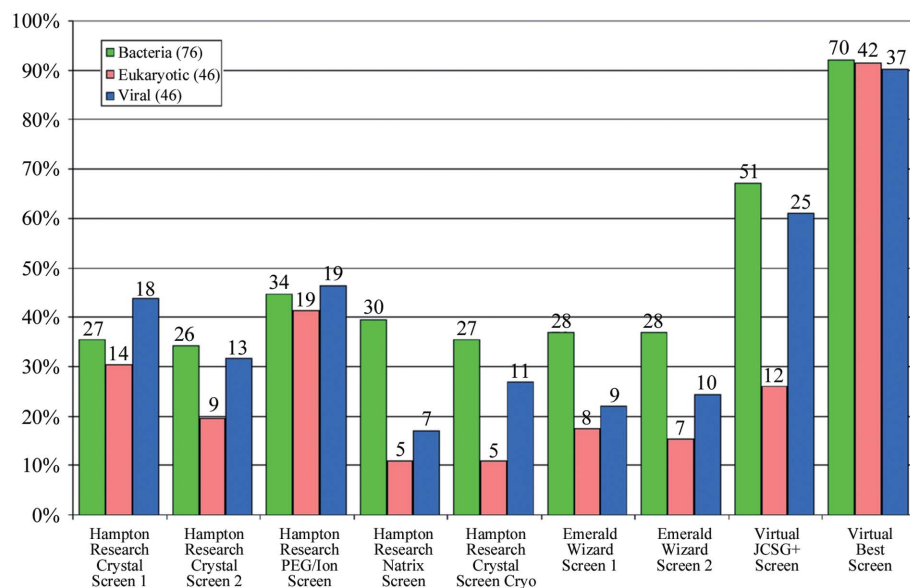


Figure 6

Coverage achieved by various commercial screens and by two 'virtual' screens in crystallizing proteins of different origins at the Oxford node. The coverage shown represents the percentage of proteins crystallized with each screen when compared with the total number of proteins crystallized for each group. The number of proteins crystallized for each protein origin is given in parentheses and above each column. The commercially available screens each contained 48 different conditions. Based on the conditions of these actual screens, 'virtual screens' can be constructed. A virtual JCSG+ screen comprising 86 of the 96 conditions would give good coverage of proteins with bacterial and virus origins, but a much reduced coverage for eukaryotic proteins. (Details of the JCSG+ screen are given in the text). Preliminary analysis of the data indicates that the 'Virtual Best Screen' of 96 conditions would cover more than 90% of all the proteins crystallized, irrespective of their origin.

4. Future development of crystallization technologies

4.1. Protein optimization

Technology improvements upstream of crystallization tend to have the potential to produce higher quality protein samples and purer samples produce more and better crystals

(discussed in Geerlof *et al.*, 2006). To realise these improvements, increasing emphasis is being placed on quality-assurance procedures and mass spectrometry has overtaken chromatography in most laboratories as the best way to confirm the precise nature of the sample and its homogeneity. A second series of tests confirm the uniformity of the protein species in solution (*e.g.* its homogeneity as a monomer, dimer *etc.*). Assessing the protein concentration is necessary to ensure that crystallization trials are set up at appropriate concentrations and methods such as gels and UV absorption are now routinely complemented by pre-screening tests. These tests use a small panel of crystallization screen solutions to see whether precipitation occurs in the expected fraction of conditions (PCT, Hampton Research). As large-scale crystallization trials accumulate more data on which properties of proteins affect their behaviour in different screen solutions, such tests will become increasingly important in allowing tailored screening where more trial conditions reach (super)-saturation gently and thereby encourage crystal growth. New QA technologies are also emerging, such as Thermofluor (a fluorescence-based thermal shift assay using an environmentally sensitive dye to monitor protein unfolding with respect to temperature, which can be used to investigate conditions favouring protein stability), and within SPINE this is now routine in three laboratories (see Geerlof *et al.*, 2006).

Not all proteins, however pure, can be persuaded to crystallize usefully. Experience with SPINE and from the Joint Center for Structural Genomics (JCSG) suggests that around one-third of proteins fall into this category (Brown *et al.*, 2003; Page *et al.*, 2003) and in these cases the only way forward is some sort of modification to the protein itself. One common approach is the production of multiple constructs for targets of high value, although this is expensive and poses challenges in construct design (see Banci *et al.*, 2006). A second approach is some sort of post-translational modification. Controlling glycosylation and other naturally occurring post-translational events is useful in some cases (see Aricescu *et al.*, 2006), while additions of specific cross-linking reagents or very dilute proteases have also been reported to be successful (Newman, 2006). A method that is becoming increasingly popular is methylation of exposed lysine residues (Rayment, 1997). The reaction goes essentially to completion, giving a homogeneous product where the surface properties have been modified to be less hydrophilic. Systematic studies in Oxford have given several notable successes in producing crystals from otherwise recalcitrant proteins (see Bahar *et al.*, 2006). A related approach is surface-entropy reduction by mutation of surface-exposed residues to replace large flexible side chains with smaller amino acids (Derewenda, 2004).

4.2. Crystallization hardware

Automation and miniaturization of crystallization screening has dramatically increased the number of crystallization trials that can be set up for any given amount of protein and their reproducibility. The increase in demand has drawn several manufacturers into the market and driven prices down,

making the use of nanolitre-dispensing robots feasible for all structural biology laboratories and their widespread adoption seems assured. Crystallization plasticware has also evolved with the developments in optical quality and suitability for membrane-protein crystallizations. The increase in scale does, however, demand better record keeping and several laboratories and collaborations are working on developing informatics solutions, particularly ones that are linked with automated storage and imaging of these trials. However, the adoption of such software is not guaranteed since it has to be flexible enough to cope with different working practices without being burdensome or prescriptive.

Crystallization methods themselves have not changed dramatically in recent years and most effort has gone into automating and miniaturizing the existing approaches: sitting-drop and hanging-drop vapour diffusion and (micro)batch. However, one emergent technology is that of the Fluidigm TOPAZ crystallization chip (Thorsen *et al.*, 2002), a free-interface diffusion method. Its key feature is the very small amount of protein solution required to perform a screen (1.5 µl for 96 trials) and it has been shown to be effective in producing many crystallization hits, although the crystals are usually far too small to be useful for diffraction studies. However, at present there are trade-offs: the chips are expensive and it can be difficult to translate hits into completed diffraction-competent crystals by scale-up in 'traditional' crystallization trials.

4.3. Crystallization screens

Screening kits have been commercially available for many years and are widely used for their convenience and increased repeatability owing to rigorous quality assurance by the manufacturers. However, mining data from the vast number of crystallization trials now performed has created scope for re-evaluation of the screen make-up. The goal for initial screening is to find some approximate crystallization conditions as starting points for optimization. For example, while PEG/Ion screens produce many crystals (see Figs. 5 and 6), the component solutions are chemically quite similar and so proteins tend either to crystallize in many conditions or in none. A better screen would sample chemical space more widely so as to give the maximum chance of finding a hit and ideally would also suggest a strategy for optimization. Some re-evaluations of screen conditions have already been reported, *e.g.* the JCSG has derived a minimal core set of 66 conditions based on results for its study of *Thermotoga maritima* proteins (Page *et al.*, 2003; Page & Stevens, 2004) which has been expanded to 96 conditions by the SPINE Amsterdam partners (JCSG+; Newman, 2006). Further analysis has confirmed the effectiveness of such a minimal screen, at least when used on proteins from the same host that the screen was designed for (Page *et al.*, 2005).

The Oxford and Strasbourg laboratories have both shown the value of targeting screens at a particular family of proteins (see, for example, Walter *et al.*, 2005). Furthermore, the initial data mining of the Oxford crystallization results for protein

targets spread across different species has shown up some unexpected but surprisingly strong correlations between screen success and source organism (Fig. 6). Thus, while the Hampton Research PEG/Ion screen gives similar success (in terms of coverage) for different host kingdoms, most other screens are markedly less successful for proteins from eukaryotic sources. We are not able to identify the physico-chemical basis for the difference in success rate, but note that a significant number of the mammalian proteins are extracellular and, when expressed in eukaryotic cells, are often glycosylated. As might have been expected from this initial observation, the selection of conditions in the JCSG+ screen, optimized against proteins from a bacterial source, shows very good success rates with other bacterial proteins, but does not show a similar increase in success with eukaryotic proteins (Fig. 6). First-pass data mining has led to the creation of a 'Virtual Best Screen' of the 96 conditions that provide the best coverage of protein families. These observations are being analyzed in more detail and will be reported elsewhere, but already demonstrate that data mining on the systematic results from a large number of crystallization experiments may significantly improve the crystallization method. In the future, analysis of proteins based on sequence and physical properties may allow custom screens to be defined or at least allow a rational selection from standard screens to be made that makes best use of the available material.

Crystallization of protein–protein complexes is made more difficult because the crystallization conditions are further limited to those which maximize complex solubility and stability. An analysis of known protein–protein complex structures in the Protein Data Bank (Berman *et al.*, 2000) and the Biological Macromolecule Crystallization Database (Gilliland & Bickham, 1990) has shown that compared with non-complex structures, the observed crystallization conditions are much less diverse and strongly favour polyethylene glycols (PEGs; 71% *versus* 27%) rather than ammonium sulfate and other high-salt conditions (Radaev & Sun, 2002). As part of this analysis a set of sparse-matrix conditions was established targeted at nuclear receptor complexes (Busso *et al.*, 2005).

4.4. Crystal optimization

The implementation of systematic optimization protocols has already made an impact; however, a single approach has not yet emerged suggesting that improvements are also possible in this area. Systematic variation of protein concentration (or relative drop volumes) and pH is easily to automate and such a protocol has proven successful in Oxford (Walter *et al.*, 2005), whilst an automated 8×8 matrix of conditions has been used effectively by Marseille (Lartigue *et al.*, 2003). Other approaches may be devised once the important parameters in 'crystallization space' are better understood. However, in common with other analyses this may require pooling crystallization results from many laboratories.

4.5. Crystallization experiments, imaging and management

To handle effectively the ever-increasing rate of creation of crystallization trials requires the use of automated storage, retrieval and imaging robots. At the start of the SPINE project there were no commercial offerings in this area, but now several manufacturers are competing to produce increasingly cost-effective robotics combining compact, environmentally controlled storage with high-quality imaging (Table 8). Although likely to remain out of reach of smaller structural biology laboratories, these storage systems are now finding their way into many larger laboratories. However, their usefulness is dependent on a software platform for integrating crystallization trial setup, imaging-session scheduling and image presentation. Within Europe, the crystallization management in Oxford, developed as part of the BBSRC-funded e-HTPX (e-science for High-Throughput Protein Crystallography) project has set the benchmark for high throughput and availability over the web (Mayo *et al.*, 2005), while manufacturers have tended to focus on lower throughput platforms with imagery licensed to a small number of local workstations. Encouraged by SPINE, a collaborative effort has been established to provide a freely available academic Laboratory Information Management System (LIMS) covering protein production and crystallization (PIMS; Protein Information Management System), which is now primarily supported by grants from the EC and UK BBSRC. As part of the project, a freely available web-based crystallization-management system is now under development (Fig. 7).

Imaging of crystallization trials tends to fall into two categories: large-scale imaging of screens and more detailed higher-quality imaging (perhaps with different lighting modes). The vast preponderance of failed uninteresting trials means that taking a single, standardized, quick image of a drop is the major activity, while more detailed imaging could be performed manually. An alternative is to fit imaging systems with two sets of optics, although this would obviously have cost implications. Despite the vast number of crystallization images produced in automated systems, almost all image analysis is still performed by eye. This burden is driving the development of image-analysis software, of which the SPINE-supported development of *ALICE* (Wilson, 2004) is the most ambitious, aiming to provide a general framework for classification that can be trained for use with images from different systems. Amongst other sites, this software is currently in routine use by the Oxford partners where it is relied on to prioritize human analysis, but the value and scarcity of crystals and the subtlety of differences between crystalline and non-crystalline objects makes it difficult to place complete trust in automated systems.

4.6. Links to downstream activities

Looking downstream from crystallization, crystals are grown to allow X-ray diffraction experiments. These experiments entail (i) manipulation of crystals to mount them from their crystallization drops, (ii) (potentially) treatment with

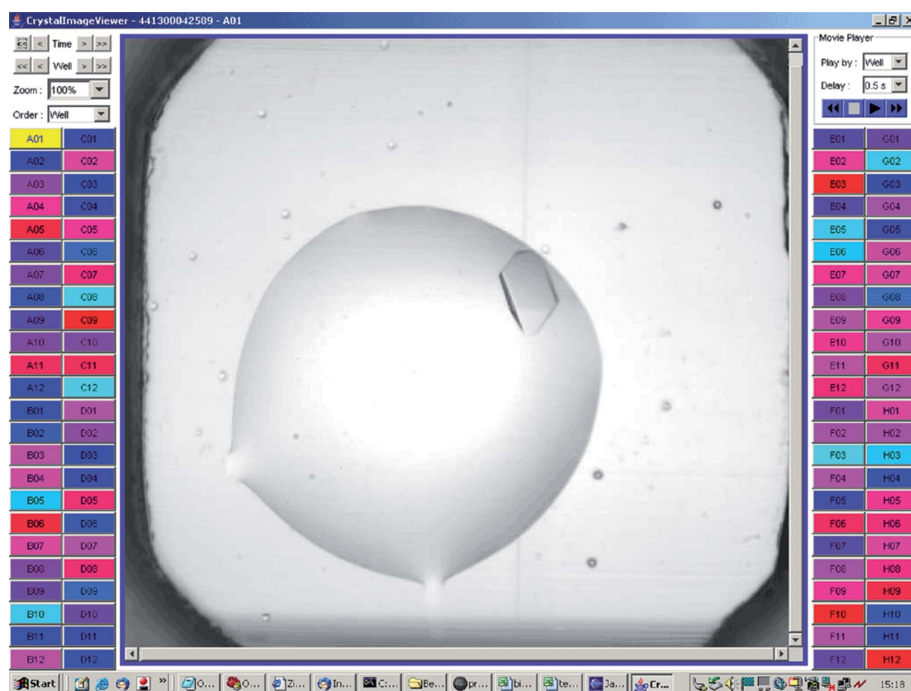


Figure 7

An image of a crystal trial shown in a development version of the PIMS crystallization management system. Figure courtesy of Dr J. M. Diprose (unpublished work).

cryoprotectants to allow cryocooling, (iii) management of the transport to X-ray facilities and (iv) linking protein production data with diffraction data. Strategies for successful cryoprotection are still poorly understood and a significant fraction of presumably good crystals are wasted in establishing protection protocols. Just like screening and optimization, accurate recording of this information on a large scale may well be able to inform choice of strategy for individual cases. This pipeline between crystallization and (automated) data collection has been addressed by collaborative work between SPINE, the e-HTPX project, synchrotrons and others to produce systems such as *DNA* (automated collection of data) and ISPyB (information system for protein crystallography beamlines) that exploit recent developments in automated beamline sample changers (see Beteva *et al.*, 2006; Cipriani *et al.*, 2006). Test pipelines are now in place, with one immediate benefit being the ability to view diffraction images remotely over the web as JPEG thumbnails and full-size images within a few seconds of collection.

5. Conclusion

In the past 3 y, and especially in Europe, nanolitre crystallography and crystal imaging and recognition have come from a state where few laboratories had anything to a state where these high-throughput technologies are considered standard practice for good science and are in regular use. This may be judged by the estimate, from a single manufacturer alone, that they have installed nanolitre dispensers in 30 laboratories around Europe. Such nanolitre technologies allow the crea-

tion of diffraction-quality crystals sufficient for structure determination, although scale-up experiments are still common in many laboratories. In general, the introduction of automation and miniaturization into crystallization workflows has resulted in the growth of a greater number of high-quality crystals for a wider variety of proteins in a range of conditions using experimental protocols that are more reproducible. Combined with the increased availability and decreased costs of automation, the economic and scientific case for adopting these technologies is compelling.

This work formed part of SPINE (Structural Proteomics In Europe), contract No. QL2-CT-2002-00988, funded by the European Commission under the Integrated Programme 'Quality of Life and Management of Living Resources'.

References

- Aricescu, A. R. *et al.* (2006). *Acta Cryst.* **D62**, 1114–1124.
- Bahar, M. *et al.* (2006). *Acta Cryst.* **D62**, 1170–1183.
- Banci, L. *et al.* (2006). *Acta Cryst.* **D62**, 1208–1217.
- Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). *Nucleic Acids Res.* **28**, 235–242.
- Bern, M., Goldberg, D., Stevens, R. C. & Kuhn, P. (2004). *J. Appl. Cryst.* **37**, 279–287.
- Beteve, A. *et al.* (2006). *Acta Cryst.* **D62**, 1162–1169.
- Brown, J. *et al.* (2003). *J. Appl. Cryst.* **36**, 315–318.
- Busso, D., Poussin-Courmontagne, P., David, R., Ripp, R., Litt, A., Thierry, J.-C. & Moras, D. (2005). *J. Struct. Funct. Genomics*, **6**, 81–88.
- Canny, J. (1986). *IEEE Trans. Pattern Anal. Mach. Intell.* **8**, 679–698.
- Chayen, N. E., Shaw Stewart, P. D. & Blow, D. M. (1992). *J. Cryst. Growth*, **122**, 176–180.
- Cipriani, F. *et al.* (2006). *Acta Cryst.* **D62**, 1251–1259.
- Cumbaa, C. A., Lauricella, A., Fehrman, N., Veatch, C., Collins, R., Luft, J., DeTitta, G. & Jurisica, I. (2003). *Acta Cryst.* **D59**, 1619–1627.
- Derewenda, Z. S. (2004). *Structure*, **12**, 529–535.
- Ferré-D'Amaré, A. R. & Burley, S. K. (1997). *Methods Enzymol.* **276**, 157–166.
- Geerlof, A. *et al.* (2006). *Acta Cryst.* **D62**, 1125–1136.
- Gilliland, G. L. & Bickham, D. (1990). *Methods*, **1**, 6–11.
- Gonzales, R. & Wintz, P. (1987). *Digital Image Processing*, 2nd ed. Reading, MA, USA: Addison-Wesley.
- Hansen, C. & Quake, S. R. (2003). *Curr. Opin. Struct. Biol.* **13**, 538–544.
- Hosfield, D., Palan, J., Hilgers, M., Scheibe, D., McRee, D. E. & Stevens, R. C. (2003). *J. Struct. Biol.* **142**, 207–217.
- Kohonen, T. & Somervuo, P. (2002). *Neural Netw.* **15**, 945–952.
- Lartigue, A., Gruez, A., Briand, L., Pernollet, J.-C., Spinelli, S., Tegoni, M. & Cambillau, C. (2003). *Acta Cryst.* **D59**, 919–921.
- Lesley, S. A. *et al.* (2002). *Proc. Natl Acad. Sci. USA*, **99**, 11664–11669.

- Mayo, C. J., Diprose, J. M., Walter, T. S., Berry, I. M., Wilson, J., Owens, R. J., Jones, E. Y., Harlos, K., Stuart, D. I. & Esnouf, R. M. (2005). *Structure*, **13**, 175–182.
- Newman, J. (2006). *Acta Cryst.* **D62**, 27–31.
- Novell, J. & Berg, J. M. (2005). *The Scientist*, **19**, 30–30.
- Page, R., Deacon, A. M., Lesley, S. A. & Stevens, R. C. (2005). *J. Struct. Funct. Genomics*, **6**, 209–217.
- Page, R., Grzechnik, S. K., Canaves, J. M., Spraggon, G., Kreusch, A., Kuhn, P., Stevens, R. C. & Lesley, S. A. (2003). *Acta Cryst.* **D59**, 1028–1037.
- Page, R. & Stevens, R. C. (2004). *Methods*, **34**, 373–389.
- Radaev, S. & Sun, P. D. (2002). *J. Appl. Cryst.* **35**, 674–676.
- Rayment, I. (1997). *Methods Enzymol.* **276**, 171–179.
- Rose, D. (1999). *Drug Discov. Today*, **4**, 411–419.
- Santarsiero, B. D., Yegian, D. T., Lee, C. C., Spraggon, G., Gu, J., Scheibe, D., Uber, D. C., Cornell, E. W., Nordmeyer, R. A., Kolbe, W. F., Jin, J., Jones, A. L., Jaklevic, J. M., Schultz, P. G. & Stevens, R. C. (2002). *J. Appl. Cryst.* **35**, 278–281.
- Spraggon, G., Lesley, S. A., Kreusch, A. & Priestle, J. P. (2002). *Acta Cryst.* **D58**, 1915–1923.
- Stevens, R. C. (2004). *Nature Struct. Mol. Biol.* **11**, 293–295.
- Thorsen, T., Maerkl, S. J. & Quake, S. R. (2002). *Science*, **298**, 580–584.
- Walter, T. S., Diprose, J., Brown, J., Pickford, M., Owens, R. J., Stuart, D. I. & Harlos, K. (2003). *J. Appl. Cryst.* **36**, 308–314.
- Walter, T. S., Diprose, J. M., Mayo, C. J., Siebold, C., Pickford, M. G., Carter, L., Sutton, G. C., Berrow, N. S., Brown, J., Berry, I. M., Stewart-Jones, G. B. E., Grimes, J. M., Stammers, D. K., Esnouf, R. M., Jones, E. Y., Owens, R. J., Stuart, D. I. & Harlos, K. (2005). *Acta Cryst.* **D61**, 651–657.
- Weber, P. C. (1997). *Methods Enzymol.* **276**, 13–23.
- Wilson, J. (2002). *Acta Cryst.* **D58**, 1907–1914.
- Wilson, J. (2004). *Crystallogr. Rev.* **10**, 73–84.
- Wilson, J. & Berry, I. (2005). *J. Appl. Cryst.* **38**, 493–500.
- Zuk, W. M. & Ward, K. B. (1991). *J. Cryst. Growth*, **110**, 148–155.