Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Quantitive evaluation of macromolecular crystallization experiments using 1,8-ANS fluorescence

David Watts, Jochen Müller-Dieckmann, Gohar Tsakanova,‡ Victor S. Lamzin and Matthew R. Groves*

EMBL Hamburg, c/o DESY, Building 25a, Notkestrasse 85, Hamburg, 22603, Germany

‡ Current address: Institute of Molecular Biology, National Academy of Sciences of the Republic of Armenia, 7 Hasratyan Str., 0014 Yerevan, Armenia.

Correspondence e-mail: groves@embl-hamburg.de

Received 20 April 2010
Accepted 29 May 2010

Modern X-ray structure analysis and advances in high-throughput robotics have allowed a significant increase in the number of conditions screened for a given sample volume. An efficient evaluation of the increased amount of crystallization trials in order to identify successful experiments is now urgently required. A novel approach is presented for the visualization of crystallization experiments using fluorescence from trace amounts of a nonspecific dye. The fluorescence images obtained strongly contrast protein crystals against other phenomena, such as precipitation and phase separation. Novel software has been developed to quantitatively evaluate the crystallization outcome based on a biophysical metric correlated with voxel protein concentration. In >1500 trials, 85.6% of the successful crystallization experiments were correctly identified, yielding a 50% reduction in the number of 'missed hits' compared with current automated approaches. The use of the method in the crystallization of three previously uncharacterized proteins from the malarial parasite *Plasmodium falciparum* is further demonstrated.

1. Introduction

One bottleneck in current crystallographic structural biological studies is the identification of conditions under which crystals are formed. Modern liquid-handling robotics allow the screening of large numbers of chemical conditions using low (nanolitre to picolitre) sample volumes (Mueller-Dieckmann, 2006; Heinemann *et al.*, 2003), resulting in a large number of crystallization trial images. Advances in image processing have demonstrated that computer algorithms may be applied to the detection and classification of individual objects within the image (Cumbaa *et al.*, 2003; Wilson, 2002; Watts *et al.*, 2008; Walker *et al.*, 2007; Pothineni *et al.*, 2006; Berry *et al.*, 2006; Spraggon *et al.*, 2002). Visible-light image-analysis techniques generally detect crystals based on contrast differences caused by refraction and absorption of light by the protein crystal. However, detection of the crystal interior by visible-light absorption is difficult and current methods mostly make use of light refraction from crystal surfaces. It is often the case that only certain faces refract and only partial identification of the entire protein crystal is achieved. The resolution of the camera, focal plane depth and image-compression rate (Berry *et al.*, 2004) are also limiting factors for successful crystal identification. This is particularly true for smaller crystals, which require a sufficient number of pixels at high resolution for reliable edge or texture detection (Supplementary Fig. 1*d*[1]). Image analysis is complicated by the varying

[1] Supplementary material has been deposited in the IUCr electronic archive (Reference: BW5354). Services for accessing this material are described at the back of the journal.

© 2010 International Union of Crystallography
Printed in Singapore – all rights reserved

morphology of crystals, which ranges from microcrystals, needles or crystal clusters to large single crystals, and the requirement to classify them into distinct crystal classes (Supplementary Fig. 1*a*). The presence of additional phenomena within the drop (*e.g.* precipitate or skin; Supplementary Figs. 1*b* and 1*c*) further complicates the identification of smaller protein crystals and results in 'missed hits' when using automated image analysis. In essence, visible-light imaging of crystallization drops is not specific for the biophysical properties of proteins, but is dependent on local illumination conditions and the alignment of crystal faces rather than biophysical characterization. Initial hits (*e.g.* Supplementary Fig. 1*a*) are often far from the single crystals required for diffraction studies. Imaging that records the spatial variation of protein concentration within a single crystallization drop allows quantification of the results based upon a meaningful biophysical parameter. An additional benefit of such an evaluation is the establishment of a metric for precipitation within a crystallization experiment rather than the commonly used nonspecific terms 'light' and 'heavy' precipitation.

We and other researchers (Forsythe *et al.*, 2006; Groves *et al.*, 2007) have previously shown that low concentrations (millimolar to picomolar) of specific fluorescence markers provide a significant increase in the quality of the signal obtained under fluorescent conditions. The improved quantum yield of the fluorescence probes used (>0.8 for carboxyrhodamine and 0.7 for 1,8-ANS; Molecular Probes, Invitrogen) results in a stronger fluorescence signal. There is little or no interaction of the excitation wavelength with the biomolecules, resulting in little or no radiation-induced damage to samples that are still undergoing crystallogenesis. Covalently attached dyes are physically linked to the proteins under examination and the freely diffusing dyes used in the experiments possess stronger fluorescence in close proximity to protein molecules, with weak residual fluorescence in the absence of protein. Thus, both dye-based methods provide fluorescence that is correlated with protein concentration in each region of the crystallization drop.

We have previously shown that light precipitate and microcrystals can be distinguished based purely upon the strength of the fluorescence signal from the corresponding region of the image (Groves *et al.*, 2007), even at dye concentrations of 9 p*M*. In this manuscript, we describe the construction of an imaging stage that is compatible with currently used 96-well high-throughput crystallization plates. We also describe the development of software that is able to score the results of crystallization experiments on a physically meaningful scale, whilst retaining spatial information.

2. Materials and methods

2.1. Detection equipment

The essential design of our visualization system is based on a 'straight-through' geometry (Supplementary Fig. 2). An array of 96 individual light-emitting diodes (LEDs; http://www.pur-led.de; catalogue No. 41437505010) is arranged on a 9 × 9 mm grid corresponding to the spacing between crystallization experiments in standard 96-well plate formats. The grid pattern ensures that each crystallization experiment in the standard high-throughput (HTP) wells is exposed to equal amounts of fluorescence excitation. The LEDs used have a peak emission wavelength at 365 nm, an intensity of ∼8 mW sr^{-1} and a viewing angle of 25°.

A reduction in the background fluorescence of the plate is achieved by the use of a mask that allows illumination of the experimental areas only. We have produced masks that are compatible with the experimental layout of the Greiner CQ PS and Innovadyne SD-2 crystallization plates that are currently supported on the EMBL-HTPX facility. The LEDs emit a small amount of visible light in the purple region of the spectrum. This was blocked by the use of a Hoya U360 band-pass filter which has 0% transmission in the visible region and ∼69% transmission at the peak wavelength of the LEDs.

An Olympus E420 Digital SLR camera was fitted with a Sigma 24 mm *f*/1.8 macro lens. The wide aperture of the low *f*-number lens enables faster shutter speeds and lower ISO numbers to be used, with a consequent reduction of noise in the image. The Sigma lens also has the advantage of accurate manual focusing with a lockable dual focusing system. The excitation radiation was filtered using a Tiffen Haze 1 UV-blocking filter attached to a Sigma lens. The camera is mounted vertically using a 30 cm steel pole and a Thorlabs swivel cross-clamp. All components were mounted on a 150 × 150 mm breadboard platform and placed within a lightproof box.

Since the experiment is enclosed within a lightproof box, external control of the camera is provided by *OLYMPUS Studio* 2 software. Control of the camera within a software application is possible using the Mocha Application Programming Interface (API) and image alignment or pincushion/barrel distortion can be corrected with the *OLYMPUS Studio* 2 software.

All images were taken with the camera in manual mode with a shutter speed of 1/3 s, an aperture of *f*/1.8 and an ISO value of 100; the white balance was set to 5500° and the image size was 3648 × 2736 pixels, giving a resolution of 35 µm per pixel.

2.2. Crystallization experiments

Crystallization experiments were performed using well characterized samples (lysozyme and proteinase K) as described previously (Groves *et al.*, 2007) unless indicated otherwise. The screens used in the course of this study included the Classics, Classics II, PEGs, PEGs II, pHClear, pHClear II, AmSO4 and PACT Suites from Qiagen and Index Screens from Hampton Research with an additional 20 µ*M* 1,8-ANS in each condition. The concentration of 1,8-ANS was chosen as a compromise between maximizing the fluorescence signal obtained and minimizing any potential crystallogenesis effects (Groves *et al.*, 2007).

2.3. Availability of *FREC* software

FREC is available as part of the EMBL Hamburg Outstation crystal-recognition suite *XREC* and will be freely available for academic use from http://www.embl-hamburg.de/XREC/.

3. Results and discussion

3.1. Analysis software

A typical image collected using an unmasked imaging stage is shown in Fig. 1(*a*), in which wells that contain crystals are clearly identifiable from the fluorescence signal. In order to provide quantitive ranking of the experimental results, software (*FREC*) has been developed to perform image processing. Accurate positioning of the plate (to better than 0.1 mm) is possible using the attached clamps and guide channels engraved into the mask surface (Supplementary Fig. 2) for the two most commonly used crystallization supports at the EMBL Hamburg HTP facility (Greiner CQ PS and Innovadyne SD-2), allowing the use of reasonably accurate preset definitions for cell and experimental area positions (Fig. 1*b*). User modification is also possible to account for minor differences in the tray position during imaging.

Fluorescence obtained from crystals is in the cyan region of the visible spectrum, in contrast to the darker blue fluorescence of the crystallization plate. To remove the fluorescence signal arising from the plate, a hue, saturation and lightness (HSL) filtering is performed. While preset values are available for commonly used plate types, user-defined HSL filter values may also be applied to the images. To aid in the choice of filter, the pixel HSL values are displayed for a given mouse position in both image and subimage regions. An alternative method involves a mouse click on potential crystals and filtering the image based on the Euclidean distance of the potential crystal HSL value and the HSL value of the pixels in the image.

Once a suitable initial HSL filter has been determined, the areas of each crystallization experiment that are above this

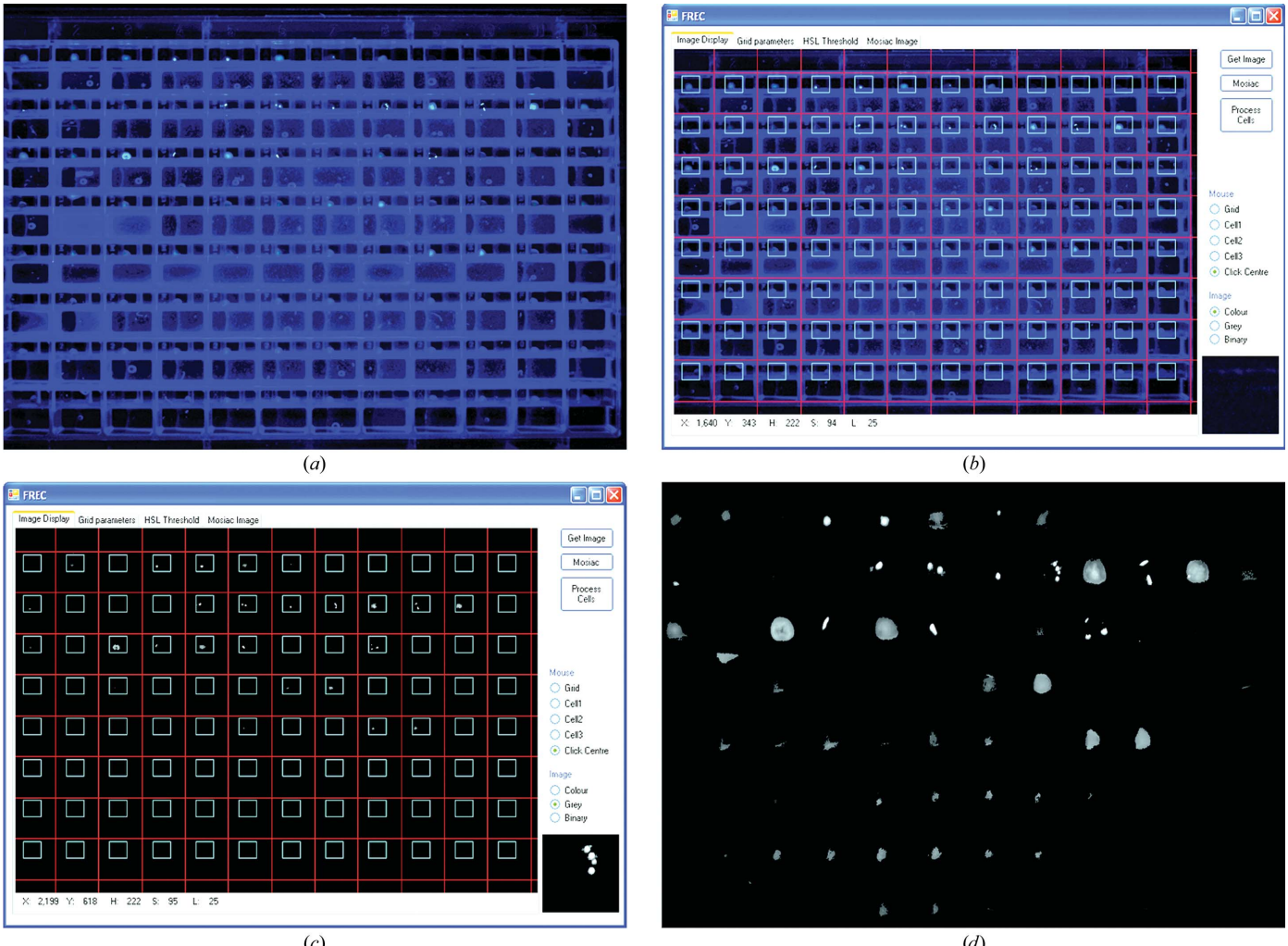


Figure 1
An example of *FREC*-based analysis of a crystallization experiment. The raw image captured by the camera is shown in (*a*). The user may choose to alter the preset definitions of cell and experimental area position through a mouse-driven interface (*b*). The results of hue–saturation–lightness (HSL) filtering are indicated in (*c*), with possibilities for the user to examine the effect of fine-tuning the HSL filter across the entire plate. The results of the filtering are shown in a 'mosaic' image (*d*), with each tile corresponding to the experimental area of the cell as defined in (*a*).

value provide a mask for downstream analysis and classification (*i.e.* only those fluorescence intensities that satisfy the HSL filter are used). Following filtering, the experimental areas defined in the previous stages are extracted from the image into an array and are re-combined into an image mosaic (Fig. 1*d*). At this stage, the spatial organization of the crystallization screen is preserved (*i.e.* the image corresponding to drop A1 is given at the top left). Visual inspection of the mosaic image enables the fluorescence intensities of the different cell areas to be easily compared and allows users to identify conditions that have resulted in clear drops, precipitation or crystalline material simply from the greyscale signal, producing similar results to those reported previously (Groves *et al.*, 2007).

Owing to the spatial resolution limit of the camera, heavy precipitate can have comparable greyscale brightness to some of the smaller crystals (Fig. 1*d*). This could have the effect of incorrectly classifying the precipitate as crystalline if the drops were to be scored solely on the basis of maximum intensity. However, in drops that contain only heavy precipitate the fluorescence signal is dispersed amorphously. Thus, the fluorescence is more uniform compared with drops that contain crystals, where images have high spatial variation in intensity. The *FREC* image-analysis algorithm obtains a measure of the statistical distribution of intensities within the image that is used for a ranking of sample crystallinity.

In an approach similar to that previously used in visible-light images (Watts *et al.*, 2008), a Haar wavelet transform (Haar, 1910; Mallat, 1989) is used to capture the intensity differences between neighbouring pixels at successive resolution levels. In many cases, the Haar wavelet has been found to outperform other wavelet bases in texture-recognition

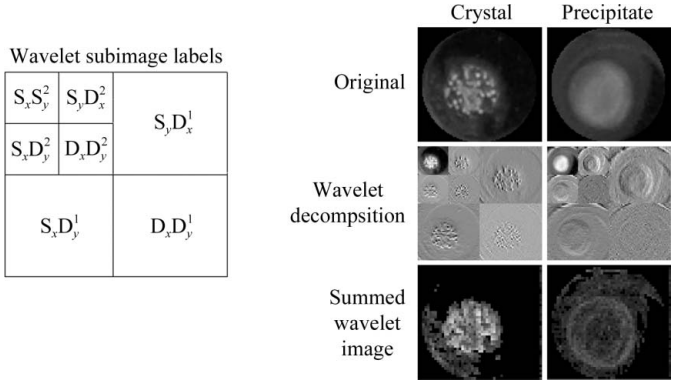


Figure 2
Demonstration of the effect of the Haar image transformation and subsequent processing of images containing crystals and heavy precipitation. A crystalline drop image (*cf.* B7 in Fig. 3) is compared with that of a heavy precipitate drop image (*cf.* H6 in Fig. 3). A two-level Haar wavelet decomposition is performed on both of the drop images. The labelling corresponding to subimages of a two-level wavelet transform is also given in the figure. The level 2 detail subimages ($S_xD_y^2$, $S_yD_x^2$ and $D_xD_y^2$) are up-sampled and summed with the level 1 detail subimages ($S_xS_y^1$, $S_xD_y^1$, $S_yD_x^1$ and $D_xD_y^1$) to produce a single image that contains the two-level detail information from the wavelet transform. This image is multiplied by the binary mask image of the experimental area as defined by *FREC* (0 for tray region, 1 otherwise) and the wavelet energy is calculated.

performance (Ohanian & Dubes, 1992; Mojsilović *et al.*, 2000) and edge detection (Singh *et al.*, 1997). Using our current system, the plate is imaged at a resolution of 35 µm per pixel and at this spatial resolution the image pixels belonging to crystal and drop become merged to a certain degree. The compact filter size of the Haar wavelet means that only derivative information between adjacent pixels is analyzed and spatial resolution is increased over the higher order wavelet functions.

The two-dimensional wavelet transform of the image is implemented (Mallat, 1989). At each frequency level of decomposition, the wavelet transform uses low-pass and high-pass filters to produce smoothed (S) and detailed (D) wavelet images, respectively (Fig. 2). The wavelet decomposition is implemented as a convolution of the image with the analyzing wavelet filter and down-sampling by 2 for each lower frequency level. Combining the results of horizontal (x) convolution followed by the vertical (y) convolution of the image produces a low-frequency (smoothed) $S_xS_y^1$ subimage in the top-left corner of the high-frequency image (detail) and $S_xD_y^1$, $S_yD_x^1$ and $D_xD_y^1$ subimages in the top-right, bottom-left and bottom-right corners of the image, respectively (Fig. 2*a*). Subsequent decomposition may then be performed by using the same algorithm on the $S_xS_y^1$ subimage, resulting in a two-level wavelet transform of the fluorescence images (Fig. 2*a*: $S_xS_y^2$, $S_xD_y^2$, $S_yD_x^2$ and $D_xD_y^2$). The result of a two-level wavelet transform of crystalline and precipitate drop images is shown in Fig. 2(*a*). Experiments performed using one-level and two-level wavelet decomposition demonstrated that the most accurate crystal ranking was achieved using the detail coefficients from level 1 of the wavelet transform alone.

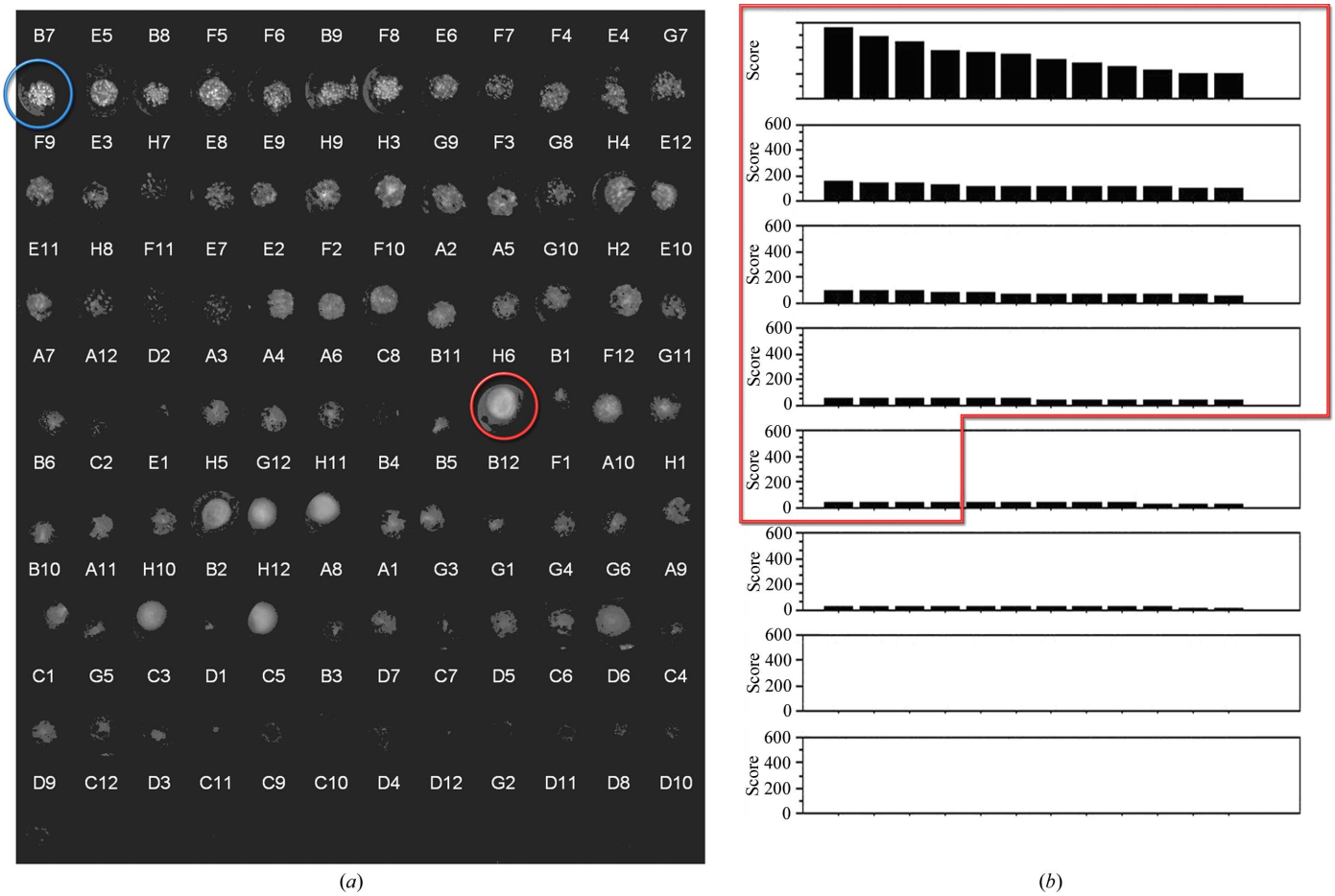
The detail wavelet images are combined into a single directionally invariant image by performing a summation of all detail wavelet coefficients at each pixel position in the image (Fig. 2*b*). The wavelet image is masked using the HSL-filtered binary mask image determined above and the squared magnitude (or energy) of each wavelet coefficient is calculated. The 70th percentile of the wavelet energies of the masked combined wavelet image is used for an overall score of the crystallization image to bias the scoring towards the detection of high-intensity single crystals rather than multiple crystalline objects as would be the case if all wavelet energy values were summed to produce an image score. An example of this process is shown in Fig. 2. The individual drop images are then sorted on the percentile score to produce a final 'score' image (Figs. 2 and 3). The entire image-processing protocol to this point requires ∼5 s on a modern desktop computer, whereas a reasonably thorough manual examination of each tray (4 s per drop) would need almost 6 min per plate.

As crystals or crystalline objects result in higher fluorescence-intensity variations across drop images, we propose that the position of the experiment within the 'score' image reflects the likelihood of a crystal being present in that particular drop. Based upon the outcome of the 1536 experiments, a current score of 30 has been chosen as a cutoff for identifying the presence of crystals based on a compromise between the

number of crystals correctly identified and the false-positive rate (Fig. 4, inset). However, the user may manually increase or decrease this filter as desired and the effect of modifying the wavelet energy score cutoff is shown in Fig. 4 (inset). We have defined the false-positive rate (FP) as the number of false positives identified at a given wavelet energy score threshold divided by the total number of conditions within the 16 screens that do not produce crystals; *i.e.* all available false positives (red line in the inset in Fig. 4). As expected, increasing the wavelet energy threshold results in a reduction in the number of false positives. However, an increase in the wavelet energy threshold also results in a decrease in the number of crystals detected by *FREC* (expressed as a fraction of all positives; blue line). We have chosen to define the success rate of *FREC* as 1 − FP and the choice of higher wavelet energy thresholds results in an increase in the fidelity of crystal identification by *FREC*. Thus, the use of the wavelet energy threshold may be tuned to the requirements of the individual crystallization experiment. The analysis of sparse-matrix screening experiments should be performed at a lower wavelet energy threshold in order to identify as many potential initial hits as possible. However, subsequent optimization experiments should be analyzed using a high wavelet energy threshold in order to identify conditions that contain larger crystals.

3.2. Demonstration of *FREC* on two trial proteins

An experiment was performed with lysozyme and proteinase K using eight crystallization screens available at EMBL Hamburg, resulting in a total of 768 different crystallization conditions for each protein. The experimental images were classified into crystal and noncrystal classes based on a preset filter value and the results of automated analysis are shown in Fig. 3. Visible-light imaging was used to manually identify all of the crystals in the experiments for comparison purposes. Automated classification of the plates resulted in an overall success rate of 85.6%, with 12.7% of the crystallization experiments falsely classified as crystals (Fig. 4). We believe that the relative performance between lysozyme and proteinase K (*i.e.* the difference in false positives) is a consequence of the difference in the number of successful



(*a*) (*b*)

Figure 3
An example of the use of *FREC* to rank the outcome of a crystallization experiment. (*a*) The individual fluorescence images resulting from a proteinase K crystallization experiment are presented ranked based on the 70th percentile wavelet energy scores. The conditions indicated by blue and red circles resulted in crystals and heavy precipitate, respectively. The wavelet energy scores of these samples were 552 and 42, respectively, and these images are used in the Haar transformation demonstration shown in Fig. 2. (*b*) The wavelet energies of each condition shown in (*a*) are indicated based upon the same arrangement; conditions that are predicted by *FREC* to result in crystals are boxed in red.

crystallization trials. In both cases the absolute number of false positives is very similar, as is the overall success rate (*i.e.* those correctly identified).

3.3. Use of *FREC* to guide the crystallization of novel proteins

In order to demonstrate the viability of ANS-based fluorescence to identify conditions that result in diffraction-quality crystals, we performed crystallization screens with three previously uncharacterized proteins (PF1, PF2 and PF3, with molecular masses ranging from 27 to 120 kDa) from *P. falciparum*. Initial screening was performed with an EMBL-HTPX robot using the Qiagen Classics Suite and PEGs Suite and Hampton Research Index Screens supplemented with 20 µ*M* 1,8-ANS. The classification results of the initial screening were analyzed using *FREC* and initial hits that were correctly identified for two of the proteins (PF1 and PF2) are shown in Supplementary Fig. 3. While well shaped initial crystals could be identified for PF3 (Jain *et al.*, 2010), further optimization of the conditions identified by *FREC* as containing crystalline material resulted in the identification of conditions that produced diffraction-quality crystals. The purification and structure determination of the proteins PF1 and PF2 will be described elsewhere. It should be noted that the crystallization conditions identified by *FREC* covered a range of crystal morphologies (well formed crystals, small rods, thin plates and spherulites; Supplementary Figs. 3*a*–3*d*, respectively).

4. Discussion

Recent advances in examination of crystallization experiments have clearly shown that a large amount of valuable data can be fed back into the crystallization experiments from such biophysical data (D'Arcy *et al.*, 2003). Examination of crystallization experiments using dynamic light scattering (DLS), cross-polarization (CP), dual polarization interferometry (DPI), absorption of ultraviolet light (UV), native or dye-based fluorescence (FL_{nat} and FL_{ext}) and X-ray diffraction (X-ray) have been demonstrated to have great potential to provide a biophysical characterization of crystallization drop phenomena.

These methods can be broadly divided into two groups: 'global' (DLS and X-ray) and 'local' (DPI, CP, UV, FL_{nat} and FL_{ext}). Dynamic light-scattering experiments performed *in situ* on crystallization drops (Kadima *et al.*, 1990; Thibault *et al.*, 1992; Saridakis *et al.*, 2002; Dierks *et al.*, 2008) provide highly

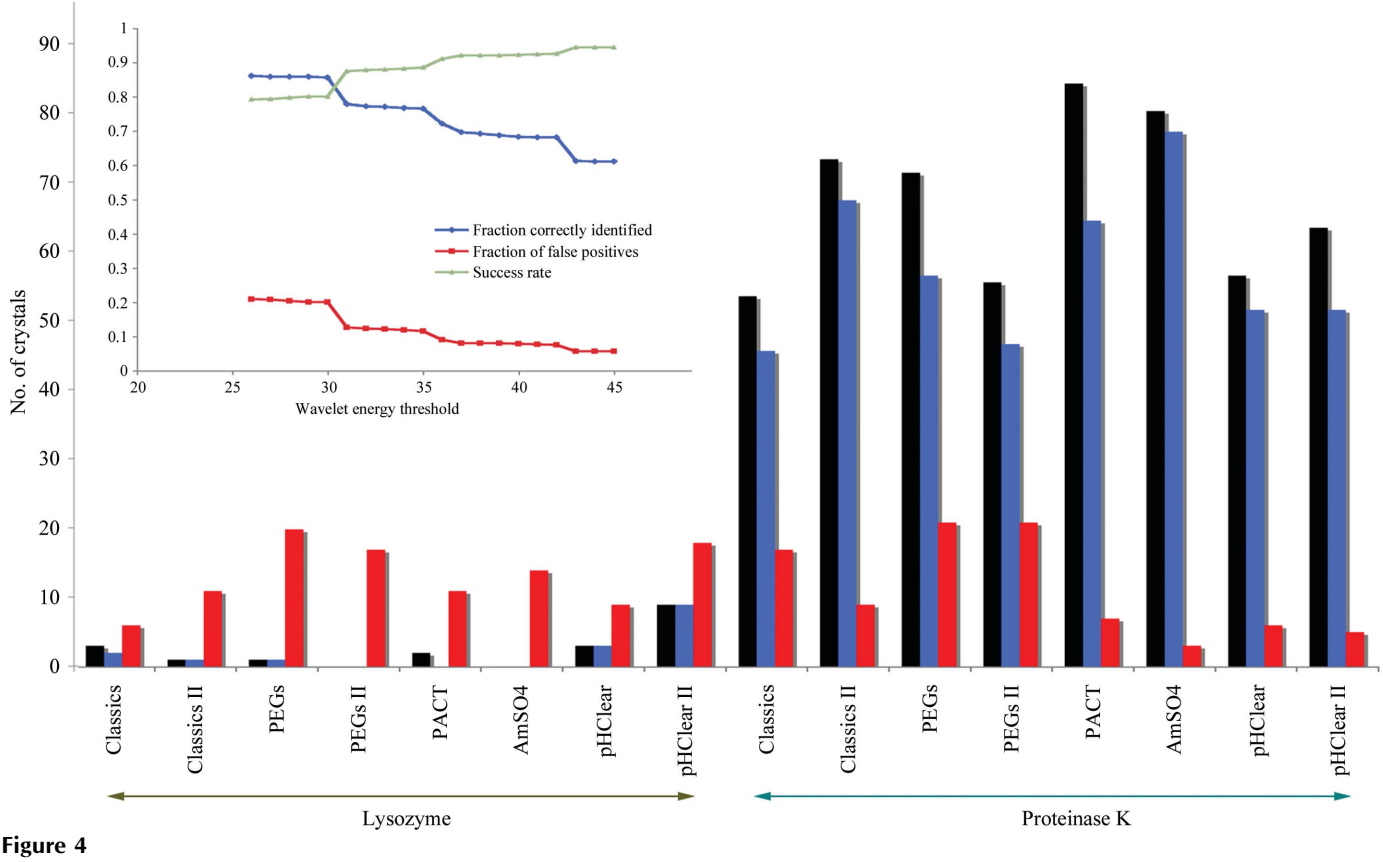


Figure 4
Eight independent screens (from Qiagen) were performed at the EMBL Hamburg HTPX facility using lysozyme and proteinase K, resulting in a total of 1536 crystallization experiments. The results of *FREC* automated crystal identification compared with the presence of crystals determined by manual examination are shown. The experiments were classified as 'crystalline' or 'noncrystalline' by *FREC* using a wavelet energy threshold value of 30. The black bars indicate the number of crystals observed manually in each screen, the blue bars represent those that were successfully classified by the *FREC* algorithm and the red bars represent false positives. The percentage of all crystals correctly identified is 85.6%, with a success rate (*i.e.* crystals correctly identified) of 80%. The percentage of false positives of the *FREC*-identified crystals is 20%. The inset shows the effect of changes in the wavelet energy score threshold for crystal identification on the percentage of crystals identified, the percentage of false positives and the percentage success rate.

valuable information about the solution state of the biomolecules and their oligomerization/aggregation state, indicating potential crystallization conditions. The presence of protein crystals can also be screened/confirmed by exposure of the entire crystallization experiment to X-ray radiation (Roth *et al.*, 2002; Jacquamet *et al.*, 2004).

Direct X-ray diffraction screening of conditions still has limitations that are imposed by the need to identify and centre a potential crystal or potential degradation of the signal-to-noise ratio when the contents of an entire drop, which may contain many small crystals, is exposed in a single 'snapshot'. In addition to the requirement for well diffracting crystals [*e.g.* those shown in Supplementary Fig. 1(*a*) did not give ordered diffraction], there is also the question of radiation damage to samples still undergoing crystallogenesis.

The 'local' methods share a single important feature: the physical measurement of a property associated with a biomolecule. Examination of crystallization drops using cross-polarization (Echalier *et al.*, 2004) has the potential to highlight the location of many protein crystals provided that they do not belong to cubic space groups (currently less than 5% of all structures in the PDB; Berman *et al.*, 2000). While the potential of the technique has yet to be fully explored, polarization occurs only from crystals and the technique is not suited to providing a direct biophysical measurement of protein concentration or degree of precipitation. The relatively weak polarization signal obtained from smaller protein crystals is also occluded or masked by visible light absorption from other phenomena, such as precipitation (Supplementary Figs. 1*b* and 1*c*), within crystallization experiments.

Native fluorescence of proteins when excited by ultraviolet radiation can be valuable in the analysis of crystallization experiments (Judge *et al.*, 2005), either through the detection of absorption or the resultant fluorescence. Indeed, native fluorescence has also been used to improve the visualization and centring of cryocooled protein crystals on synchrotron beamlines (Vernede *et al.*, 2006). As both UV absorption and fluorescence are intrinsic properties of a protein, these techniques may provide a biophysical measurement of the relative protein concentration across the drop. All methods based upon absorption spectroscopy have limitations, as precipitate or other phenomena also absorb and potentially mask the presence of smaller crystals. However, absorption spectroscopy is based upon the absence of a signal absorbed by the sample. Increasing the flux of UV radiation in order to increase the contrast will result in an increase in the dose absorbed by the biomolecules, potentially leading to increased radiation damage (Dose & Sena, 1968; Kehoe *et al.*, 2008).

The detection of fluorescence signals partially circumvents this last limitation (Judge *et al.*, 2005), as suitable filtering and an increase in counting/integration time may result in improved signal-to-noise ratios. However, the presence of some cofactors (*e.g.* iron) and the necessity for the presence of specific amino acids (tryptophan, tyrosine) are limiting factors, as is radiation damage from the excitation wavelength as discussed above. Additionally, native fluorescence has a relatively low quantum yield ($\sim$0.13 for a fully exposed tryptophan; Lakowicz, 1999) and re-absorption of the emitted fluorescence by the high local concentration of protein within the lattice may reduce the generated signal (Franzen *et al.*, 1972). This is reduced by covalent labelling and added free dye approaches, both of which limit the number of fluorescent molecules present and their final potential concentration.

While we have previously demonstrated that the fluorescence signal available from crystallization experiments performed in the presence of 1,8-ANS may be used to successfully identify protein crystals as small as 1 µm (Groves *et al.*, 2007), the capture and interpretation of these images required a significant amount of manual labour. In this manuscript, we have presented the construction and use of a simple and cost-effective device that allows the rapid imaging of crystallization screens performed in the presence of 1,8-ANS. The total cost of all materials used to construct the image stage was under 1000 euros. We also describe the developed software that can be used to aid identification and also quantitatively evaluate the results of the crystallization experiment based on the fluorescence intensity of pixels within the image. An analysis based on the physical behaviour of the proteins under investigation is likely to result in a reduction in the number of false negatives by the identification of successful conditions that would be otherwise occluded under visible-light visualization (Supplementary Fig. 1). Such a ranking can also be used to quantatively describe the physical behaviour of the sample under sparse-matrix screening conditions, rather than simply assigning the results as 'light' or 'heavy' precipitation. Such a quantitative analysis of the results could also be used to design subsequent optimization screens, even in the absence of clearly identifiable initial crystals. Other authors (D'Arcy *et al.*, 2003) have suggested that conditions resulting in crystalline precipitate can be used in subsequent seeding experiments. The method that we present can aid in the identification of such conditions through a physical characterization of the nature of the precipitate in terms of its local (voxel) concentration. However, further experimentation is under way to demonstrate the efficacy of this.

The use of wavelet energy to characterize the crystallization drops allows the biophysical behaviour of the protein within a crystallization drop to be used as a classification metric to determine the crystallinity of phenomena in each experiment. While the current wavelet energy preset is somewhat subjective, we believe that the ranking of the crystallization results is of more use than any cutoff value. The individual user may use the results of the 'score' image to concentrate attention in further examination with visible-light microscopy on drops that are most likely to contain crystals. While it is not unreasonable to expect that the researcher will wish to visually examine the results of all his/her experiments, the described software *FREC* provides a basis for the choice of which drops to examine first. It is also to be anticipated that not all screens will result in crystals, but a biophysical characterization of the protein behaviour may aid the design of further screens as described above. We can currently identify no general characteristics, in terms of morphology, size or position, that can easily be attributed to the 14% missed hits, although it is our

working hypothesis that the resolution of the camera used is the major limitation here. While crystal size and morphology certainly play a role in successful identification (see Groves *et al.*, 2007) the use of the wavelet energy is not dependent upon the absolute fluorescence and both needle-shaped and plate-shaped crystals gave similar results.

While care must be taken to address the concerns of experimenters who may not wish to add additional components to a crystallization experiment, it should be noted that a single additional screen using this technique would consume less than 35 µl of material (based on a drop size of 200 nl). As other researchers have also recently demonstrated the efficacy of adding small molecules as 'magic bullets' in crystallization screening (McPherson & Cudney, 2006), the use of 1,8-ANS could be seen as an aid to crystal identification rather than an aid to the crystallization process, as we believe that this method provides significant advantages in imaging and successfully identifying crystallization conditions that may otherwise be overlooked based upon visible-light imaging alone.

It should also be noted that once crystallization conditions have been identified using the described method, 1,8-ANS might be omitted from all downstream optimization crystallization experiments. Through repetitions of the experiments described in this manuscript, we have established that 1,8-ANS is stable in the crystallization screen for over six months (when stored at 277 K), with fluorescence-based measurements showing no significant difference over this time period. An initial trial of mixing a small volume (1 µl) of protein solution with 1,8-ANS and assaying for fluorescence is sufficient to establish whether a particular protein is suitable for this method. The presented analysis is clearly limited by the resolution of the camera and the imaging of an entire crystallization plate. Future developments will include the use of an *X*–*Y* stage and microscope to provide significantly increased resolution beyond the current 35 µm, although the current resolution is sufficient to identify crystals as small as ∼15–20 µm (Fig. 3). Such developments could reasonably be expected to improve our current low-cost prototype into something more suitable for high-throughput laboratories.

We are grateful to Doris Jahn for mechanical engineering support and Bernd Robrahn for construction of the 96 UV LED array and power supply. We also gratefully acknowledge the help of Xandra Kreplin in the preparation of the crystallization experiments. We would also like to acknowledge Dr R. Jordanova and Dr A. Schmidt for helpful discussion during the writing of this manuscript.

References

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.

Berry, I., Wilson, J., Mayo, C., Diprose, J. & Esnouf, R. (2004). *Intelligent Data Engineering and Automated Learning – IDEAL 2004*, edited by Z. R. Yang, R. Everson & H. Yin, pp. 117–124. Berlin: Springer.

Berry, I. M., Dym, O., Esnouf, R. M., Harlos, K., Meged, R., Perrakis, A., Sussman, J. L., Walter, T. S., Wilson, J. & Messerschmidt, A. (2006). *Acta Cryst.* D**62**, 1137–1149.

Cumbaa, C. A., Lauricella, A., Fehrman, N., Veatch, C., Collins, R., Luft, J. R., DeTitta, G. & Jurisica, I. (2003). *Acta Cryst.* D**59**, 1619–1627.

D'Arcy, A., Mac Sweeney, A. & Haber, A. (2003). *Acta Cryst.* D**59**, 1343–1346.

Dierks, K., Meyer, A., Einspahr, H. & Betzel, C. (2008). *Cryst. Growth Des.* **8**, 1628–1634.

Dose, K. & Sena, L. (1968). *Photochem. Photobiol.* **7**, 59–64.

Echalier, A., Glazer, R. L., Fülöp, V. & Geday, M. A. (2004). *Acta Cryst.* D**60**, 696–702.

Forsythe, E., Achari, A. & Pusey, M. L. (2006). *Acta Cryst.* D**62**, 339–346.

Franzen, J. S., Kuo, I. & Chung, A. E. (1972). *Anal. Biochem.* **47**, 426–441.

Groves, M. R., Müller, I. B., Kreplin, X. & Müller-Dieckmann, J. (2007). *Acta Cryst.* D**63**, 526–535.

Haar, A. (1910). *Math. Ann.* **69**, 331–337.

Heinemann, U., Bussow, K., Mueller, U. & Umbach, P. (2003). *Acc. Chem. Res.* **36**, 157–163.

Jacquamet, L., Ohana, J., Joly, J., Legrand, P., Kahn, R., Borel, F., Pirocchi, M., Charrault, P., Carpentier, P. & Ferrer, J.-L. (2004). *Acta Cryst.* D**60**, 888–894.

Jain, R., Jordanova, R., Müller, I. B., Wrenger, C. & Groves, M. R. (2010). *Acta Cryst.* F**66**, 409–412.

Judge, R. A., Swift, K. & González, C. (2005). *Acta Cryst.* D**61**, 60–66.

Kadima, W., McPherson, A., Dunn, M. F. & Jurnak, F. A. (1990). *Biophys. J.* **57**, 125–132.

Kehoe, J. J., Remondetto, G. E., Subirade, M., Morris, E. R. & Brodkorb, A. (2008). *J. Agric. Food Chem.* **56**, 4720–4725.

Lakowicz, J. R. (1999). *Principles of Fluorescence Spectroscopy*, 2nd ed. New York: Springer.

Mallat, S. (1989). *IEEE Trans. Pattern Anal. Mach. Intell.* **11**, 674–693.

McPherson, A. & Cudney, B. (2006). *J. Struct Biol.* **156**, 387–406.

Mojsilović, A., Popović, M. V. & Rackov, D. M. (2000). *IEEE Trans. Image Processing*, **9**, 2043–2050.

Mueller-Dieckmann, J. (2006). *Acta Cryst.* D**62**, 1446–1452.

Ohanian, P. P. & Dubes, R. C. (1992). *Pattern Recognition*, **25**, 819–833.

Pothineni, S. B., Strutz, T. & Lamzin, V. S. (2006). *Acta Cryst.* D**62**, 1358–1368.

Roth, M., Carpentier, P., Kaïkati, O., Joly, J., Charrault, P., Pirocchi, M., Kahn, R., Fanchon, E., Jacquamet, L., Borel, F., Bertoni, A., Israel-Gouy, P. & Ferrer, J.-L. (2002). *Acta Cryst.* D**58**, 805–814.

Saridakis, E., Dierks, K., Moreno, A., Dieckmann, M. W. M. & Chayen, N. E. (2002). *Acta Cryst.* D**58**, 1597–1600.

Singh, R., Vasquez, R. E. & Singh, R. (1997). *Proc. SPIE*, **3074**, 151–159.

Spraggon, G., Lesley, S. A., Kreusch, A. & Priestle, J. P. (2002). *Acta Cryst.* D**58**, 1915–1923.

Thibault, F., Langowski, J. & Leberman, R. (1992). *J. Cryst. Growth*, **122**, 50–59.

Vernede, X., Lavault, B., Ohana, J., Nurizzo, D., Joly, J., Jacquamet, L., Felisaz, F., Cipriani, F. & Bourgeois, D. (2006). *Acta Cryst.* D**62**, 253–261.

Walker, C. G., Foadi, J. & Wilson, J. (2007). *J. Appl. Cryst.* **40**, 418–426.

Watts, D., Cowtan, K. & Wilson, J. (2008). *J. Appl. Cryst.* **41**, 8–17.

Wilson, J. (2002). *Acta Cryst.* D**58**, 1907–1914.