Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Simultaneous X-ray diffraction from multiple single crystals of macromolecules

Karthik S. Paithankar,[a]‡ Henning O. Sørensen,[b]§ Jonathan P. Wright,[c] Søren Schmidt,[b] Henning F. Poulsen[b] and Elspeth F. Garman[a]*

[a]Laboratory of Molecular Biophysics, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, England, [b]Center for Fundamental Research: Metals in Four Dimensions, Materials Research Division, Risø National Laboratory for Sustainable Energy, Technical University of Denmark, Frederiksborgvej 399, PO Box 49, DK-4000 Roskilde, Denmark, and [c]European Synchrotron Radiation Facility, 6 Rue Jules Horowitz, BP 220, 38043 Grenoble CEDEX, France

‡ Current address: Macromolecular Crystallography (BESSY-MX), Institute F-I2, Helmholtz-Zentrum Berlin für Materialien und Energie, Elektronenspeicherring BESSY II, Albert-Einstein-Strasse 15, D-12489 Berlin, Germany.
§ Current address: Nano-Science Center, Department of Chemistry, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen Ø, Denmark.

Correspondence e-mail: elspeth.garman@bioch.ox.ac.uk

Received 24 November 2010
Accepted 25 April 2011

The potential in macromolecular crystallography for using multiple crystals to collect X-ray diffraction data simultaneously from assemblies of up to seven crystals is explored. The basic features of the algorithms used to extract data and their practical implementation are described. The procedure could be useful both in relation to diffraction data obtained from intergrown crystals and to alleviate the problem of rapid diffraction decay arising from the effects of radiation damage.

1. Introduction

Macromolecular crystallography (MX) has so far primarily been based on single-crystal methods. In this study, we explore a multicrystal approach as an alternative, in which a number of crystals are simultaneously illuminated during the diffraction experiment. We present the application of a novel indexing scheme which enables the use of conventional single-crystal diffraction pattern integration programs for data analysis of each of the crystals separately.

The motivation for employing this approach in macromolecular crystallography was twofold. The ever-increasing size of molecules in biological crystallography (such as ribosomal molecular machines) leads to lower intermolecular *versus* intramolecular interactions and this tends to result in small crystals (<10 µm), commonly known as microcrystals, which exhibit very weak diffraction. In such cases, data collected from a few exposures of each of a large number of single microcrystals can sometimes be merged to obtain a complete data set (Riekel *et al.*, 2005). In addition, crystals are often inter-grown and inseparable even in the skilful hands of experienced crystallographers. These give rise to more than one distinct diffraction pattern upon X-ray irradiation. Unlike merohedral twinning, in which the constituent lattices overlap with an exact symmetry, the different orientations from a multicrystal lead to a complex diffraction pattern that is often uninterpretable. In such cases, manual picking of lattices or elimination of unwanted overlapping spots is sometimes a possibility (Buts *et al.*, 2004; Sauter & Poon, 2010). Various software approaches have been employed in attempts to improve the treatment of weak and/or spatially overlapped reflections (Bourgeois *et al.*, 1998; Bourgeois, 1999), sometimes based upon the *ab initio* prediction of reflection profiles (Schreurs *et al.*, 2010). However, the programs freely available so far handle data sets from single-crystal diffraction only and do not allow the processing of multiple crystals. In the last

decade it has become possible to collect data by mounting the crystallization plate in which crystals are grown directly in front of the collimator and exposing the crystals to X-rays *in situ* at room temperature (Watanabe *et al.*, 2002; Jacquamet *et al.*, 2004, 2009; Yadav *et al.*, 2005; Ng *et al.*, 2008). These diffraction patterns may contain multiple lattices (arising from the different crystals inside the plates or the capillaries). Hence, additional programs to perform proper data reduction are needed.

Additionally, modern synchrotron beamlines offer high photon flux densities that lead to significant radiation damage to macromolecular crystals maintained at 100 K, which is manifested in terms of both specific structural damage and global damage (Ravelli & Garman, 2006). However, the high photon fluxes of third-generation synchrotrons are essential in order to obtain useful diffraction from small and weakly diffracting crystals, resulting in high absorbed doses that inflict significant radiation damage. In such cases, the development of a method that could extract single crystal diffraction patterns from multiple single crystals would be useful. Collecting data from multiple single crystals simultaneously in a single pass and merging the reflections detected from each presents a viable opportunity to obtain quality data before the crystals have decayed owing to radiation damage.

One of the challenges of analysing data collected from multiple crystals is that of indexing the individual crystal lattices. Programs have previously been developed to index diffraction patterns consisting of hundreds of crystalline grains in polycrystalline metallic samples (Lauridsen *et al.*, 2001; Jensen *et al.*, 2006). This multicrystal approach has also been applied to solving structures that lie between the extremes of single crystal and homogenous powders and has been successfully demonstrated in chemical crystallography for samples of Al_2O_3 (Schmidt *et al.*, 2003) and cupric acetate monohydrate [$Cu(C_2O_2H_3)_2.H_2O$; hoganite] (Vaughan *et al.*, 2004).

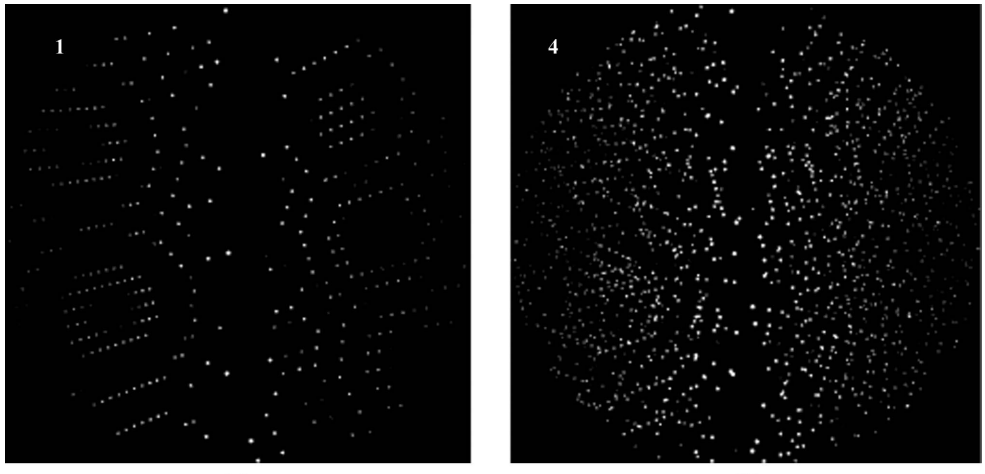
In this article, we explore the possibilities for generalizing the above work to multicrystals of macromolecules. By means of simulations as well as actual diffraction experiments, we investigate the degree of spot overlap and the resulting quality of structure refinement as a function of the number of crystals simultaneously irradiated. Two proteins, namely chicken egg-white lysozyme (HEWL) and cubic insulin, were studied in detail in order to estimate the effect of the size of the unit cell on spot overlap. We demonstrate that the indexing software developed as part of this study can be interfaced with the existing single macromolecular crystal data-integration program *MOSFLM* (Leslie, 1992).

2. Methods

2.1. Simulations

The influence of various parameters on a multicrystal X-ray diffraction experiment was analysed using simulated diffraction patterns calculated using the program *PolyXSim* (Sørensen, 2008; https://sourceforge.net/apps/trac/fable/wiki/PolyXSim). The program generates diffraction images of the polycrystalline sample, similar to those obtained experimentally from the detectors used in protein crystallography, and a file containing structure factors for the reflections for the crystal ensemble is produced. The inputs required by the program are values for the different experimental variables, namely the incident-beam characteristics (energy and flux), the oscillation angle, the number and size distribution of the crystals, their mosaicity, peak shape and reflection profile (square or Gaussian), a background level, a Boolean flag for the inclusion of Poisson noise (or not) and the detector configuration, along with a file containing the crystal symmetry, the unit-cell information and the atomic coordinates of the protein in PDB (Protein Data Bank) or CIF (crystallographic information file) format.

Simulations of the diffraction to be expected from between two and four crystals of the proteins HEWL (PDB entry 2blx; Nanao *et al.*, 2005) and cubic insulin (PDB entries 2ceu and 2bn3; Nanao *et al.*, 2005; Whittingham *et al.*, 2006) were carried out (Fig. 1). For the sake of simplicity (and time constraints) all of the crystals in the ensemble were assumed to be of the same size and their orientations relative to the beam were generated randomly. The incident beam was assumed to be a top-hat shape and have an energy of 13.2 keV (0.939 Å) [simulating the conditions at ESRF (European Synchrotron Radiation Facility) beamline ID14-4]. Since this work, the programs *MLFSOM* (J. Holton, http://bl831.als.lbl.gov/~jamesh) and *SIM_MX* (Diederichs, 2009) have been made available to simulate diffraction images. These simulation programs can be used as a tool to investigate the overlap fraction and number of crystals that it might be feasible to



(*a*) (*b*)

Figure 1
Simulations of diffraction patterns for one (*a*) and four (*b*) crystals of HEWL (PDB entry 2blx) to a resolution of 2 Å with an oscillation angle of 0.25° and a mosaicity of 0.5°.

Table 1
Data-collection statistics for multiple crystals of HEWL and insulin at 100 K.

Protein crystal	HEWL	Insulin
Total No. of crystals exposed in the loop	2, 3, 4, 7	2, 3, 4
Wavelength (Å)	0.939	0.939
Space group	$P4_32_12$	$I2_13$
Unit-cell parameters (Å)	$a = b = 78.5$, $c = 36.9$	$a = b = c = 78.1$
No. of exposures	360, 360, 360, 360	720, 720, 360
Oscillation angle (°)	0.5, 0.5, 0.5, 0.5	0.25, 0.25, 0.5
Maximum resolution at the detector edge circle (Å)	1.8	1.8
Exposure time (s)	1	1

Experiment	No. of lattices detected	Data integration possible
HEWL		
2	2	1, 2
3	3	1
4	4	1, 2, 3
7	7	4, 5, 6, 7
Insulin		
2	2	1, 2
3	3	1, 2, 3
4	5	1, 2, 3, 4

irradiate simultaneously for a specific system and experimental setup.

2.2. X-ray diffraction experiments

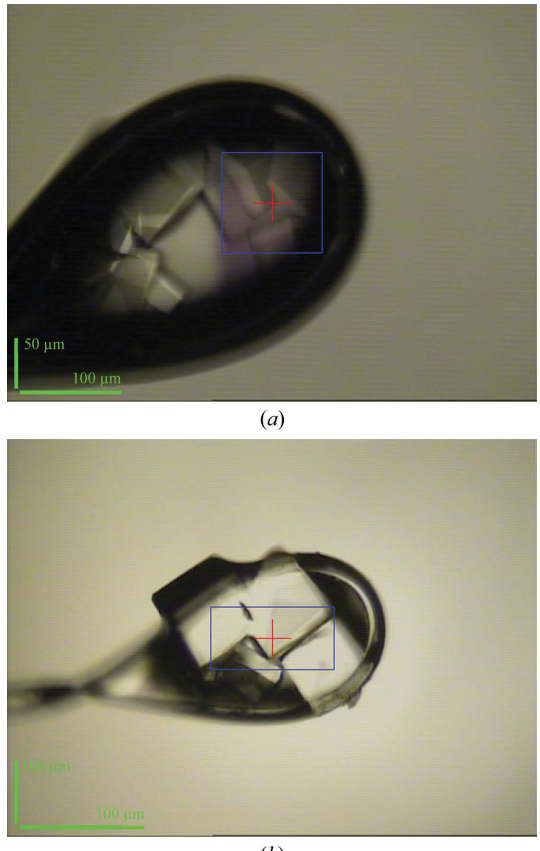
Single crystals of both chicken egg-white lysozyme and insulin were grown by the hanging-drop vapour-diffusion method.

For crystallization, 2 µl HEWL (30 mg ml^{-1}; Merck Biosciences, catalogue No. L6876) was mixed with an equal volume of crystallization buffer consisting of 4–8%(*w*/*v*) NaCl in 100 m*M* sodium acetate pH 4.5. For bovine insulin (Sigma, catalogue No. I5500), 2 µl protein solution (20 mg ml^{-1}) consisting of 0.02 *M* Na_2HPO_4 and 0.01 *M* Na_3EDTA was mixed with an equal volume of precipitant consisting of 0.5 *M* Na_2HPO_4/Na_3PO_4 pH 10.0–10.6.

Multiple single crystals of dimensions of up to 100 × 100 × 100 µm were soaked (for 30 s) in cryoprotectant solutions containing 35 and 30%(*v*/*v*) glycerol for HEWL and insulin, respectively, which were prepared by replacing some of the water in the crystallization buffer. These ensembles were exposed to X-radiation on ESRF beamline ID14-4 equipped with an X-ray detector (ADSC Quantum Q315r). All crystals were flash-cooled in a gaseous nitrogen stream and maintained at 100 K during data collection. A total of up to seven crystals were captured in a rayon-fibre cryo-loop and a controlled number were exposed to X-rays by suitable centring operations. The crystals were centred in such a way that a maximum number of crystals (visible on the on-axis microscope) could be illuminated throughout the experiment. In a normal diffraction experiment on a single crystal it is always recommended that there be minimal liquid (cryoprotectant) surrounding the crystal and a beam that matches the size of the crystal (or ensemble in this case) to optimize data quality. In the experiments reported here it should be emphasized that no special effort was made to pick up crystals of the same size and in almost all cases the crystals had substantially different sizes. Images of multiple crystals in a single rayon-fibre cryo-loop are shown in Fig. 2. Diffraction experiments were performed on ensembles comprising between two and seven multiple crystals for HEWL (designated H2, H3, H4 and H7, respectively) and two to four multiple crystals for insulin (designated I2, I3 and I4, respectively). In all experiments, two different data sets (Table 1) were collected: one at a high resolution (1.9 Å) and the other at lower resolution (3.0 Å).

2.3. Data processing

Data from the different ensembles were subjected to the same data-reduction pipeline. As illustrated in Fig. 3, initially novel multicrystal analysis programs were used to identify and index the reflection data. These programs are all part of the program suite *FABLE* (available at http://sourceforge.net/apps/trac/fable/ under an open-source licence). Having obtained orientation matrices for the individual lattices of the ensemble, the remainder of the analysis was performed with existing MX programs (*e.g.* within the *CCP*4 suite; Winn *et al.*, 2011). In the present study, the program *MOSFLM* was used for data integration, while *SCALA* (Evans, 2006) followed by



(*a*)

(*b*)

Figure 2
Multiple crystals of (*a*) HEWL in a cryo-loop viewed by an on-axis camera and (*b*) insulin. The blue box represents the position and size of the incident X-ray beam.

other utilities from the *CCP*4 suite were used for the merging and scaling of data. The program *MOLREP* (Vagin & Teplyakov, 2010) was used for molecular replacement, and model refinement was performed using *REFMAC*5 (Murshudov *et al.*, 2011) within the *CCP*4 suite. The low-resolution data sets were collected in the event that our methodology was unsuccessful on the higher resolution sets. In the event, since the latter were successfully processed, there was no necessity to analyse the low-resolution sets in detail.

2.3.1. Peak search. Peaks from the X-ray diffraction images were harvested with the *peaksearch* program from the *ImageD*11 suite (J. P. Wright, unpublished work). *peaksearch* finds three-dimensional peaks using a threshold procedure. For the experimental data, the *peaksearch* program was typically run to harvest peaks containing more than 100 counts and peaks very close to the boundaries between the individual tiles of the CCD (charge-coupled device) detector were filtered out. The remaining peak positions were transformed to scattering vectors in reciprocal space.

2.3.2. Indexing the lattices. Indexing of the multiple crystal diffraction patterns was performed with the program *GrainSpotter* (S. Schmidt, unpublished work), leading to a number of crystal orientation matrices. Given the unit-cell parameters and space-group symmetry, *GrainSpotter* performs searches for matches between the theoretical reciprocal lattice and the experimental pattern of scattering vectors. The search is performed in orientation space and a critical factor in optimizing its performance has been the choice of orientation representation, which is a hybrid of using quaternion space (Meister & Schaeben, 2005) and Rodrigues space (Frank, 1988). The search is carried out in two steps. Firstly, a subspace of orientation space is chosen either by random selection in quaternion space or by scanning through the full quaternion space. This orientation subspace is then parameterized in Rodrigues space because the geodesics of every scattering vector is a straight line in Rodrigues space, and if a set of scattering vectors originates from the same crystal orientation then these lines will have a common vertex in this space. The

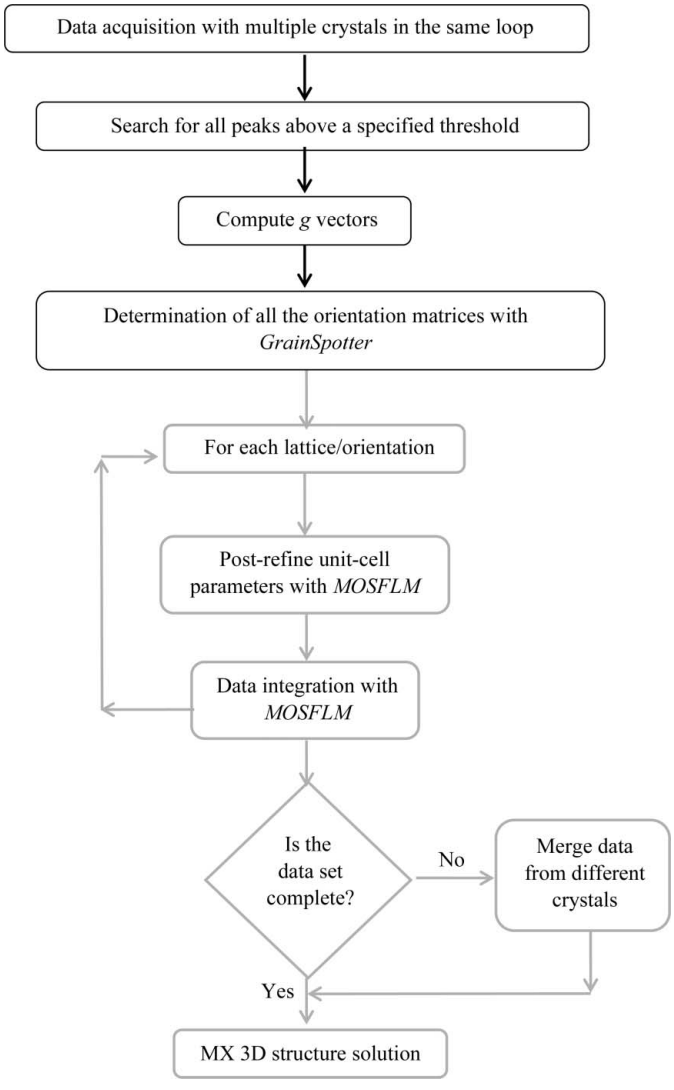


Figure 3
A flow diagram of the different steps involved in structure solution using multiple crystal diffraction data. The components that are already routine in MX are shown in grey.

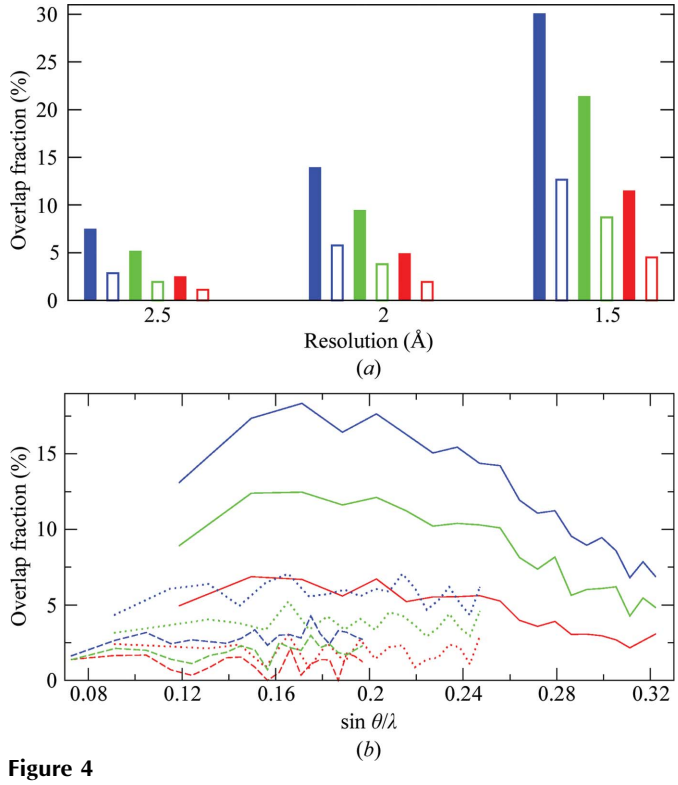


Figure 4
(*a*) Spot overlap as a function of resolution for two different crystal forms of insulin: cubic $I2_13$ (PDB entry 2bn3) and orthorhombic *I*222 (PDB entry 2ceu) with crystal volumes of 472 729 and 185 374 Å^3, respectively. Filled bars show results for cubic insulin and hollow bars show results for orthorhombic insulin. Simulated data were produced for two (red), three (green) and four (blue) simultaneously diffracting crystals. (*b*) The calculated fraction of spot overlaps for orthorhombic insulin [hollow bars in (*a*)] are plotted as a function of $\sin\theta/\lambda$ in equal-volume reciprocal-space shells. Red, green and blue colours again denote simulations performed with two, three and four crystals, respectively. Solid, dotted and broken lines indicate data-set resolutions of 1.5, 2.0 and 2.5 Å, respectively, at the detector edge.

Table 2
Data-reduction statistics for the two-crystal HEWL ensemble and the four-crystal HEWL ensemble for the oscillation range 0–180°.

Values in parentheses are for the outer resolution shell.

Crystal identifier	H2-1	H2-2	H4-1	H4-2	H4-3
Resolution range (Å)	39.3–1.8 (1.9–1.8)	39.3–1.8 (1.9–1.8)	36.9–1.9 (2.0–1.9)	36.9–1.9 (2.0–1.9)	36.9–1.9 (2.0–1.9)
No. of reflections: total/unique	157753/11753	157954/11746	65399/9171	65579/9647	65832/9680
Completeness (%)	100 (100)	100 (100)	95.6 (97.3)	99.9 (100)	100 (100)
Multiplicity	13.4 (13.4)	13.4 (13.4)	7.1 (7.3)	6.8 (7)	6.8 (7.1)
$R_{\rm meas}$†	0.07 (0.24)	0.07 (0.2)	0.13 (0.4)	0.07 (0.15)	0.1 (0.3)
$R_{\rm p.i.m.}$‡	0.02 (0.06)	0.02 (0.05)	0.05 (0.15)	0.03 (0.06)	0.04 (0.1)
PCV§	0.1 (0.3)	0.09 (0.23)	0.16 (0.52)	0.08 (0.19)	0.13 (0.4)
$\langle I\rangle/\langle\sigma(I)\rangle$	25.0 (10.5)	25 (12.2)	12.6 (4.9)	22.5 (11.2)	21.4 (12)
Wilson B factor (Å^2)	15.6	15.9	12.3	12.4	11.5

† $R_{\rm meas} = \sum_{hkl}[N/(N-1)]^{1/2}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$. hkl is a particular reflection, N is the multiplicity of reflection hkl, $I_i(hkl)$ is the ith intensity measurement of reflection hkl and $\langle I(hkl)\rangle$ is the average intensity of reflection hkl. $R_{\rm meas}$ accurately reflects the reliability of individual measurements independent of multiplicity (Diederichs & Karplus, 1997), while $R_{\rm p.i.m.}$ describes the precision of the average measurement (Weiss & Hilgenfeld, 1997). The $[N/(N-1)]^{1/2}$ factor allows the contributions from the individual reflections to be appropriately weighted according to their multiplicity. ‡ $R_{\rm p.i.m.} = \sum_{hkl}[1/(N-1)]^{1/2}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$. § PCV (pooled coefficient of variance) = $\sum_{hkl}[1/(N-1)]^{1/2}\sum_i[I_i(hkl) - \langle I(hkl)\rangle]^2/\sum_{hkl}\langle I(hkl)\rangle$. For the PCV (a multiplicity-weighted measure of the data quality), the pooled standard deviation (the statistically valid measure of the noise level) is divided by the sum of the intensities (the signal level).

orientations can now be found by searching for vertices in the Rodrigues subspace. In cases where information on the unit cell and symmetry is unavailable (an issue which will be the subject of an upcoming publication), two or three diffraction patterns from a single crystal can be collected widely spaced in φ and indexed with existing MX software to obtain the crystal habit.

Having determined a crystal orientation within the uncertainty limits provided by the user, this orientation (and if desired the crystal position) is fitted to reciprocal-lattice vectors. In the scenario where some crystals show substantially weaker diffraction than others and their orientations are close, it can be difficult to index the weaker lattice. The user can remove the indexed scattering vectors using the program *filtergrain* from *ImageD*11 and repeat the *GrainSpotter* search using only the non-indexed scattering vectors. The resulting orientations are output as Busing–Levy orientation matrices (Busing & Levy, 1967), which are individually input into *MOSFLM* (Leslie, 1992) using the keyword MATRIX so that the data can integrated crystal by crystal. Alternatively, the orientation can be suitably input to programs such as *XDS* (Kabsch, 2010), *d*TREK* (Pflugrath, 1999) or the *HKL* package (Otwinowski & Minor, 1997).

3. Results and discussion

3.1. Study of spot overlap using simulations

The fraction of spot overlaps was quantitatively analysed as a function of resolution and of the number of crystals with the help of diffraction-pattern simulations for crystals of the proteins HEWL (tetragonal form) and insulin (in orthorhombic and cubic forms) using *PolyXSim*. Spot-overlap analyses for the two forms of insulin are presented in Fig. 4. The simulations enabled the expected resolution limits for a specific number of crystals to be determined. The analyses were performed assuming a reflection spot size with a diameter of 11 pixels on the CCD detector mentioned above.

As expected, there is an overall increase in spot overlap when collecting data to high resolution (limited by the detector edges) compared with moving the detector further away and only collecting data to lower resolution. The increase in overlap fraction is mainly a consequence of the fact that at higher resolution more reflections have to be collected simultaneously on the same detector area. However, when the overlap fraction (number of overlapping reflections in a given volume/total number of reflections in the volume) is binned as a function of resolution within one data set, a close to constant value is found in all resolution bins (Fig. 4). Increasing the number of illuminated crystals also increases the overlap, but interestingly it rises almost linearly, at least up to four crystals, *e.g.* the spot overlap at 2 Å for insulin increases from ∼5% for two crystals to ∼10% for three crystals and ∼14% for four crystals. Comparing the simulation results for the two different unit cells of insulin (with volumes of 472 729 and 185 374 Å^3 for the cubic and orthorhombic forms, respectively), as expected the larger unit cell leads to an increased number of spot overlaps compared with the smaller unit cell (see Fig. 4*a*). Despite the increasing spot-overlap fraction observed as the number of crystals becomes larger, a value of only 14% was predicted for the four-crystal simulated data from cubic insulin at 2 Å resolution and at the same time the percentage of spot overlap is close to constant over all d^* within each data set (Fig. 4*b*). This is in sharp contrast to data collected by the Laue diffraction method, in which the overlap fraction increases dramatically with resolution as the accessible area of reciprocal space increases.

3.2. Indexing and integration of multiple-crystal experimental data

Once the diffraction patterns had been indexed individually as described above, data belonging to individual crystals from the various multiple-crystal data collections of HEWL and insulin were integrated individually in *MOSFLM* and the quality of the resulting merged data sets was analysed.

Strong diffuse scattering was observed from some crystals and as the crystal sizes differed in the ensembles, the obtainable resolution limit from each was also variable. This meant that although all crystals in the ensemble were successfully

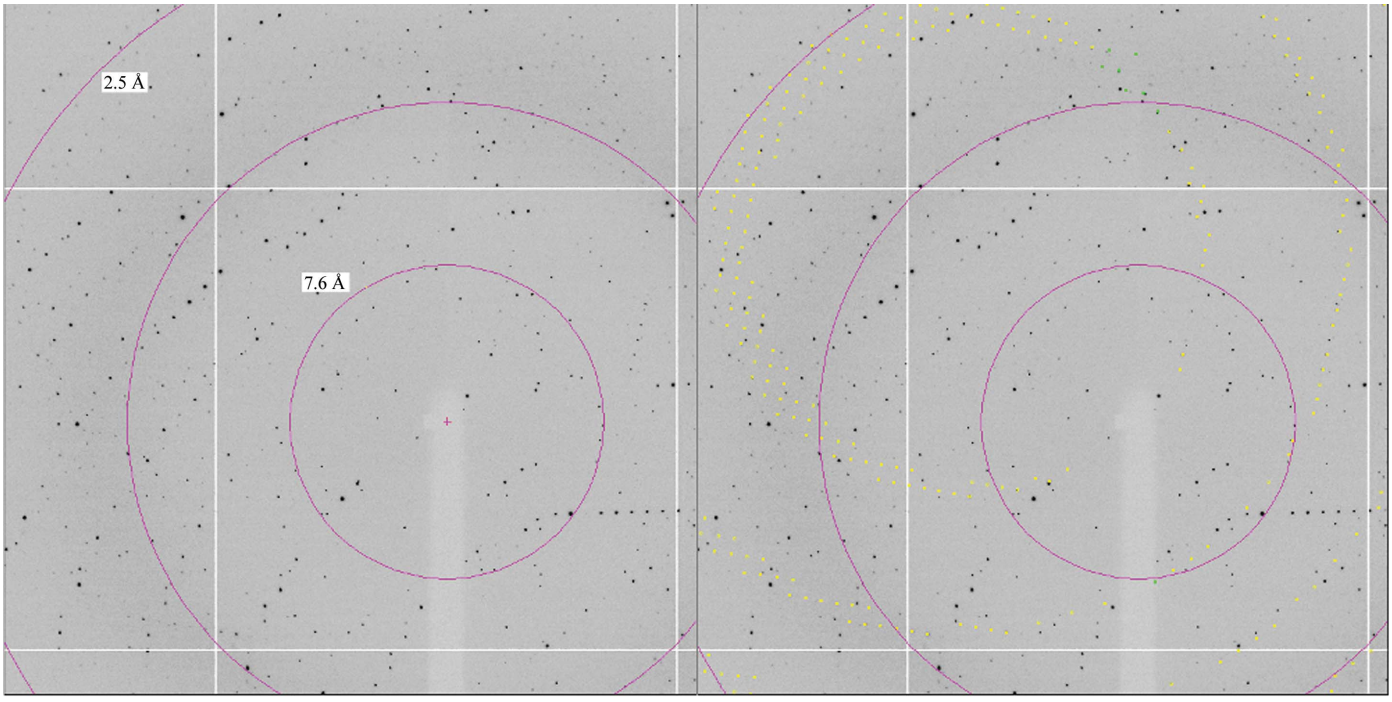


Figure 5
Diffraction data obtained from the simultaneous X-ray exposure of four crystals of HEWL (*a*). The yellow circles in (*b*) are reflection positions predicted by *MOSFLM* following the determination of the orientation matrix for the weakest crystal by *GrainSpotter*.

identified and indexed, data from some crystals were not carried through to the integration step. Finally, an attempt was made to perform structure solution with a complete data set obtained by merging the data obtained from individual crystals in the ensemble of multiple crystals (exposed simultaneously).

3.2.1. Data from two crystals of HEWL. Exposing two crystals of HEWL (H2) gave two distinct diffraction lattices that could be visually distinguished. Whereas neither of the two lattices could be successfully indexed with *MOSFLM*, *GrainSpotter* could identify and index both of them. The orientation matrices given by *GrainSpotter* were subsequently used to integrate the spots from the two crystals individually using *MOSFLM*. A summary of data-integration results for the multiple-crystal data collection of H2 (higher resolution data set 1.77 Å) is shown in Table 2. Data-reduction statistics of data for 18° wedges in φ collected for the two individual crystals (H2-1 and H2-2) are given separately in the supplementary material (Table S1[1]). It can be seen that in all φ rotation ranges for both crystals data completeness is around 63% with excellent $\langle I\rangle/\langle\sigma(I)\rangle$ values. The values of the redundancy-independent merging R factor ($R_{\rm meas}$) are uniform throughout the data collection at 7.7%, except for the rotation ranges 163–180° (H2-1) and 109–126° (H2-2). Not surprisingly, integration over the complete range of data collection (0–180°) for the two crystals (of tetragonal symmetry) separately leads to complete and good-quality data sets (Table 2). Since the crystals in the three-crystal HEWL experiment (H3) showed very weak diffraction, they were not subjected to a detailed data analysis.

3.2.2. Data from four crystals of HEWL. In the experiment with four HEWL crystals (H4) in the cryo-loop, all crystals but one showed good diffraction (Fig. 5). Data-reduction statistics are thus shown in Table 2 for three of the four crystals and those for individual crystals (composed of φ wedges of 18°) in Table S2[1]. All three crystals show reasonable mosaicity (0.5°) and low Wilson B values (∼13 Å^2). A completeness of above 50% is obtained for all crystals in all wedges except for H4-3 in the rotation range φ = 0–18°. There are at least two reasons for this behaviour. Firstly, this data-collection experiment was performed with a beam of size 100 × 100 µm and crystals with sizes between 40 and 100 µm. Thus, at some φ angles (rotation axis) not all of the crystals in the ensemble could be fully illuminated simultaneously by the X-ray beam. In addition, some crystals were plate-like and hence diffracted poorly in certain rotation ranges. Despite this, integrating the data set over 180° (for a single crystal) shows that all three indexed crystals show very high data quality and completeness.

3.2.3. Data from seven crystals of HEWL. When the seven-crystal HEWL (H7) data set was processed with *GrainSpotter* all seven crystal lattices from the ensemble could be identified. This matches with an expectation based on visual inspection of the sample. Only four of the seven crystals diffracted to high resolution (<2 Å) and could be successfully integrated (see Table 3). For the other three crystals, very weak diffraction (around six spots per degree were observed in some rotation

[1] Supplementary material has been deposited in the IUCr electronic archive (Reference: EA5138). Services for accessing this material are described at the back of the journal.

Table 3
Data-reduction statistics for the seven-crystal HEWL ensemble.

Only data for crystals that showed good diffraction spots are shown in the table. The oscillation angle ($\Delta\varphi$) per exposure was 0.5° over a total of 180°. Values in parentheses are for the outer resolution shell. Definitions are as in the footnotes to Table 2.

Crystal identifier	H7-4	H7-5	H7-6	H7-7
Resolution range (Å)	36.9–1.9 (2.0–1.9)	56.0–2.6 (2.7–2.6)	78.8–2.0 (2.1–2.0)	38.9–3.0 (3.1–3.0)
No. of reflections: total/unique	64510/8837	24630/3943	56467/8332	14668/2510
Completeness (%)	93.6 (89.1)	99.6 (100)	99.9 (99.9)	99.9 (100)
Multiplicity	7.3 (6.2)	6.2 (6.5)	6.8 (7.0)	5.8 (6.2)
$R_{\rm meas}$	0.06 (0.13)	0.2 (0.4)	0.25 (0.77)	0.3 (0.5)
$R_{\rm p.i.m.}$	0.02 (0.05)	0.09 (0.17)	0.09 (0.28)	0.1 (0.2)
PCV	0.08 (0.17)	0.3 (0.67)	0.3 (1.0)	0.42 (0.68)
$\langle I\rangle/\langle\sigma(I)\rangle$	27.1 (14.6)	7.6 (4.8)	8.8 (4.2)	3.5 (2.5)
Wilson B factor (Å^2)	12.3	28.2	12.9	31.3

Table 4
Data-reduction statistics for the two-crystal ensemble of insulin.

Values in parentheses are for the outer resolution shell. Definitions are as in the footnotes to Table 2.

Crystal identifier	I2-1	I2-2
Resolution range (Å)	39–1.5 (1.6–1.5)	55.3–1.7 (1.8–1.7)
No. of reflections: total/unique	136734/12814	77131/8544
Completeness (%)	100 (100)	91.8 (44.2)
Multiplicity	10.7 (10.7)	9.0 (2.9)
$R_{\rm meas}$	0.08 (0.32)	0.1 (0.47)
$\langle I\rangle/\langle\sigma(I)\rangle$	20.4 (7.4)	14.6 (2.4)
Wilson B factor (Å^2)	15.3	15.5

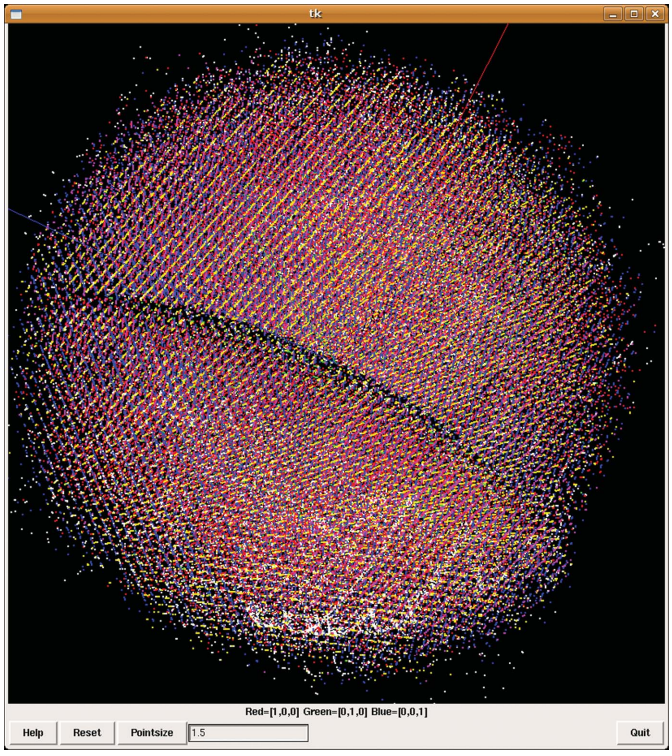


Figure 6
The experimental reciprocal-lattice points calculated from the peaks harvested with *peaksearch* from the diffraction images of sample I4 plotted with the *ImageD*11 program *plot3d*. The points are coloured red, green, blue, cyan and yellow for the five different crystal orientations assigned to the reflections during the *GrainSpotter* indexing procedure. The white spots denote unindexed reciprocal-lattice points and the red line (at the top right) indicates the x axis of reciprocal space.

ranges) was observed and a high positional r.m.s. residual (above 6 mm) for the spot prediction when compared with the actual spot position led to failure during reflection integration in *MOSFLM*. The crystals for which the data could be integrated, H7-4, H7-5, H7-6 and H7-7, have mosaic spreads of about 0.2, 0.5, 0.35 and 1.2°, respectively.

Crystal H7-4 provides the best data-reduction statistics, with good values for the signal-to-noise ratio over the entire data collection. For crystal H7-5, the positional r.m.s. residual for the average spot profile was very high (0.3–0.4 mm) throughout the range of integration. Crystals H7-5 and H7-6 possibly suffered from poor crystallinity, as the diffraction spots looked sharp in the early part of the rotation range (φ range 1–36°) but became increasingly diffuse as the sample was rotated to φ ranges of 54–90° (Table S3). This change cannot be ascribed to radiation damage because when more data were collected from the same sample in the same φ rotation ranges (with a different crystal-to-detector distance to collect low-resolution data after the higher resolution data collections), the exposures in the rotation range 1–36° of the second data set showed sharp reflections which become diffuse when the φ range 54–90° was reached. In the case of H7-7, the crystal diffracted to a worse resolution (3.0 Å and lower) throughout data collection except in the φ range 18–54°, where very few spots were observed as the crystal moved out of the beam. Analysis of the results from this sample illustrates one of the advantageous features of using more than one crystal in an X-ray diffraction experiment, since here we mounted seven crystals simultaneously and obtained one good single-crystal data set without having to test the seven crystals individually.

3.2.4. Data from two crystals of insulin. When data from the two crystals of insulin (I2) were treated with *GrainSpotter*, both lattices were identified. Since crystal I2-1 diffracted better than I2-2, data were integrated for I2-1 and I2-2 to 1.5 and 1.7 Å, respectively (Table 4). Data completeness for all wedges of I2-1 is above 87%, whereas I2-2 shows a slightly lower completeness of around 81% (Table 5). It is clear that crystal I2-1 not only diffracts to a higher resolution but shows significantly better values for $\langle I\rangle/\langle\sigma(I)\rangle$ across all of the φ rotation range. For both crystals, the spot profiles show very low values of positional r.m.s. residual (0.04 mm) over all exposures (data not shown).

3.2.5. Data from three crystals of insulin. All data obtained from exposure of the three crystals of insulin were integrated in the range 55–1.9 Å (Table 6). Owing to the high symmetry of the insulin lattice (space group $I2_13$), even data processed in 18° wedges (Table S4) showed very good completeness (>90%) and also good values of $\langle I\rangle/\langle\sigma(I)\rangle$, except for crystal I3-3. Scaling of I3-1 and I3-2 gave uniform values of the

Table 5
Data-reduction statistics for the two-crystal insulin ensemble for φ wedges of 18°.

Values in parentheses are for the outer resolution shell. Definitions are as in the footnotes to Table 2.

Crystal identifier	Image range ($\Delta\varphi$ = 0.5°)	Completeness (%)	Multiplicity	R_{meas}	$\langle I\rangle/\langle\sigma(I)\rangle$	No. of reflections: total/unique
I2-1	1–72	87.4 (90.8)	2.5 (2.4)	0.06 (0.3)	10.5 (4.0)	27306/11040
	73–144	94.3 (94)	2.3 (2.3)	0.07 (0.3)	9.3 (3.3)	27497/11991
	145–216	97.1 (98.2)	2.1 (2.1)	0.07 (0.4)	8.9 (2.8)	26207/12510
	217–288	97.6 (98.9)	2.1 (2.1)	0.07 (0.4)	8.5 (2.6)	26879/12508
	289–360	92.8 (90.8)	2.3 (2.4)	0.08 (0.3)	8.4 (2.8)	27218/11825
I2-2	1–72	84.9 (85.5)	1.7 (1.7)	0.2 (0.8)	2.4 (1.0)	12466/7521
	73–144	90.3 (55)	2.1 (1.6)	0.1 (0.4)	5.9 (1.8)	16215/7788
	145–216	78.3 (20.8)	2.0 (1.1)	0.1 (0.5)	5.7 (1.4)	13625/6836
	217–288	78.2 (30.5)	1.9 (1.4)	0.1 (0.6)	5.8 (1.4)	12198/6350
	289–360	77.5 (43)	2.0 (1.5)	0.09 (0.45)	6.6 (1.9)	13601/6774

Table 6
Data-reduction statistics for the three crystal ensemble of insulin. The oscillation angle ($\Delta\varphi$) per exposure was 0.25°..

Values in parentheses are for the outer resolution shell. Definitions are as in the footnotes to Table 2.

Crystal identifier/No. of exposures	I3-1/1–360	I3-2/1–360	I3-3/1–360	I3-3/1–180
Resolution range (Å)	55.1–1.9 (2.0–1.9)	55.1–1.9 (2.0–1.9)	55.1–1.9 (2.0–1.9)	55.1–1.9 (2.0–1.9)
No. of reflections: total/unique	67939/6399	68291/6396	68368/6393	34010/6344
Completeness (%)	100 (100)	100 (100)	100 (100)	99.8 (100)
Multiplicity	10.6 (10.9)	10.7 (11)	10.7 (11)	5.4 (5.5)
R_{meas}	0.1 (0.3)	0.07 (0.12)	0.25 (1.1)	0.1 (0.16)
$\langle I\rangle/\langle\sigma(I)\rangle$	17.8 (7.1)	29.7 (19.4)	17.2 (11.9)	14.6 (8.9)
Wilson B factor (Å^2)	16.6	18.4	17.7	17.0

Table 7
Data-reduction statistics for the four-crystal ensemble of insulin.

The oscillation angle ($\Delta\varphi$) per exposure was 0.5° for a total of 180 images. Values in parentheses are for the last shell. Definitions are as in the footnotes to Table 2.

Crystal identifier	I4-1	I4-2	I4-3	I4-4
Resolution range (Å)	55.2–1.9 (2.0–1.9)	55.2–1.9 (2.0–1.9)	55.2–1.9 (2.0–1.9)	55.2–1.9 (2.0–1.9)
No. of reflections: total/unique	68052/6393	67753/6396	67266/6369	67562/6391
No. of rejected reflections	218	349	218	818†
Completeness (%)	100 (100)	100 (100)	100 (100)	100 (100)
Multiplicity	10.6 (10.9)	10.6 (10.9)	10.6 (10.8)	10.6 (10.9)
R_{meas}	0.10 (0.34)	0.08 (0.18)	0.08 (0.2)	0.14 (0.3)
$R_{p.i.m.}$	0.03 (0.10)	0.03 (0.05)	0.03 (0.06)	0.04 (0.09)
PCV	0.13 (0.48)	0.11 (0.24)	0.11 (0.27)	0.20 (0.43)
$\langle I\rangle/\langle\sigma(I)\rangle$	21.0 (8.2)	23.5 (12.5)	23.2 (10.6)	13.5 (7.2)
Wilson B factor (Å^2)	18.6	16.7	17.9	17.2

† Most of the rejected reflections were detected in the last 30 images.

Table 8
Statistics for merging of wedges of data from crystals I1, I2, I3 and I4 of the insulin four-crystal ensemble.

For this exercise data were integrated incrementally from the first exposure to n for each crystal separately. Each of these incremental wedges from the different crystals were then merged and scaled together. See §3.3 for details. Values in parentheses are for the last shell. Definitions are as in the footnotes to Table 2.

Image range from four crystal lattices	Completeness (%)	Multiplicity	R_{meas}	$\langle I\rangle/\langle\sigma(I)\rangle$	No. of reflections: total/unique	R value/ R_{free} (%)
1–5	62.8 (62.8)	1.5 (1.5)	0.12 (0.45)	7.5 (3.0)	6064/3969	16.1/20.1
1–10	90.3 (90.2)	2.4 (2.4)	0.11 (0.44)	9.1 (3.7)	13642/5733	16.2/19.5
1–15	97.5 (98.3)	3.4 (3.5)	0.10 (0.41)	11.7 (4.9)	21256/6212	15.6/18.9
1–20	99.4 (100)	4.6 (4.6)	0.10 (0.38)	13.8 (5.6)	28891/6344	15.4/18.8
1–25	99.8 (100)	5.7 (5.8)	0.10 (0.38)	15.8 (6.2)	36465/6376	15.3/18.8
1–45	99.8 (100)	10.5 (10.7)	0.10 (0.34)	21.1 (8.5)	66917/6379	15.2/18.8
1–60	99.9 (100)	14.0 (14.4)	0.09 (0.31)	24.4 (10.2)	89621/6386	15.1/18.9
1–90	100 (100)	21.1 (21.7)	0.09 (0.29)	29.5 (12.9)	135089/6393	15.1/18.9
1–180	100 (100)	42.3 (43.6)	0.10 (0.28)	41.2 (19.3)	270490/6393	15.1/18.9

average applied scale factor of between 0.25 and 1 for the entire rotation range of 90°. However, for crystal I3-3 the scale factors ranged from 1–176 over the data-collection range 0–90°. Again, we attribute this to movement of the crystal out of the beam as the crystal was rotated during data collection. For all of the crystals in the ensemble the Wilson B-factor estimates are low at around 17 Å^2.

3.2.6. Data from four crystals of insulin. When the data from four crystals of insulin (I4-1, I4-2, I4-3 and I4-4) were indexed with *GrainSpotter*, five crystal lattices were identified. This was one more than was apparent by visual inspection of the sample in the loop (both through the goniometer on-axis view and while fishing the ensemble). Using *peaksearch*, 166 845 peaks were harvested and 99% (165 263) of these peaks could be matched by the five crystal orientations (Fig. 6). The majority of the remaining unindexed peaks in this case (shown in white in Fig. 6) arose from detector defects. Three of the crystals (I4-1, I4-2 and I4-3) were in the diffracting condition throughout the entire data collection and displayed similar diffraction qualities (Table 7). Crystal I4-4 rotated partly out of the beam after a rotation of 60° (120 exposures) and thus had a slightly lower $\langle I\rangle/\langle\sigma(I)\rangle$ for the latter part of the data collection. I4-5 showed weak diffraction and its data could not be reliably integrated to a reasonable resolution. As the sample was not preserved after the experiment, it was not possible to verify whether I4-5 was a unique crystal or a satellite crystal. Over the complete rotation range (0–180°), three of the well diffracting crystals (I4-1, I4-2 and I4-3) showed a low R_{meas} of around 7–9% and a rather high $\langle I\rangle/\langle\sigma(I)\rangle$ value (>20, whereas I4-4 had a slightly higher R_{meas} of 13% and a lower $\langle I\rangle/\langle\sigma(I)\rangle$ (∼13). All crystals

showed similar Wilson B-factor values of about 17 Å^2. Statistics from integrating the data from the different crystals separately for subsets of rotation angles (exposures 1–20 and 1–60) are presented in Table S5.

3.3. Merging of the individual data sets to construct a complete data set

One of the goals of the technique development presented here was to reduce the data-collection time as well as to obtain complete data with minimum X-ray irradiation. The discussion below will be limited to the data collected with four insulin crystals in the sample, as the data with two and three crystals are solely less complex versions of the same case.

In Table 8, the data-reduction statistics (by merging data from all four insulin crystals) are shown for a number of different subsets of rotation angles. When merging data from more than one crystal, it must be kept in mind that the crystals can be indexed using twin-related orientations. Insulin is in space group $I2_13$ and there is one possible twin symmetry operation: h, k, l to k, h, $-l$ (leading to ambiguity in the direction of one of the axes). Therefore, the data were re-indexed before merging: if this step is inadvertently omitted the $R_{\rm meas}$ value becomes 40% instead of the true value of around 10%. Table 8 shows that after a rotation of just 12.5° (images 1–25), data with a completeness of 99.8%, a multiplicity of 5.7, an $R_{\rm meas}$ of 10% and an $\langle I\rangle/\langle\sigma(I)\rangle$ of 15.8 could be obtained. This multicrystal data set is thus comparable in quality to typical single-crystal data (with only one crystal in the loop).

Fig. 7 shows $R_{\rm p.i.m.}$ and $\langle I\rangle/\langle\sigma(I)\rangle$ for the merged data sets as a function of φ. A comparison of completeness between data from only one of the four crystals in the ensemble (Table S5) with the merged data (Table 8) for an equivalent number of exposures shows clearly that the completeness is marginally lower in the case of the merged data. For example, whereas I4-1 (60 exposures, 30° rotation range) shows a completeness of 99.5%, merging I4-1, I4-2, I4-3 and I4-4 (15 exposures each, 7.5° rotation range) shows a completeness of 97.5%. $R_{\rm meas}$ values for the merged I4 data sets are similar to the values obtained from integrating the individual crystals separately; by merging the four data sets into one $\langle I\rangle/\langle\sigma(I)\rangle$ shows good improvement in most cases.

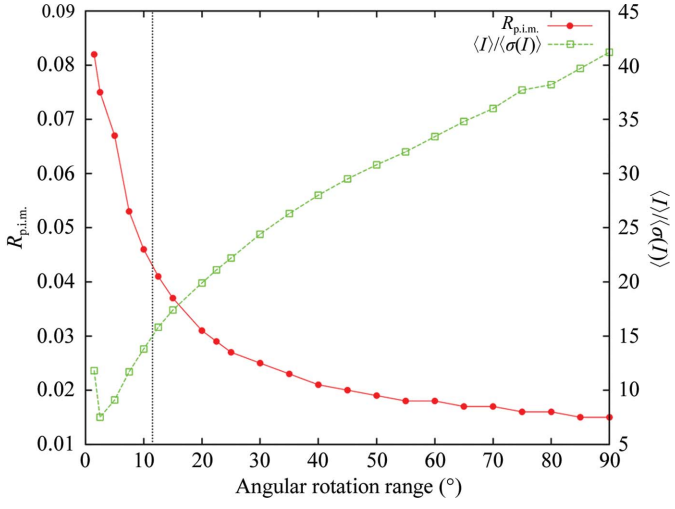


Figure 7
The residual $R_{\rm p.i.m.}$ and $\langle I\rangle/\langle\sigma(I)\rangle$ are plotted as a function of the number of degrees of the rotation range included in the data merging of the insulin four-crystal data into single data sets. The vertical dashed line marks the position at which the merged data set is of similar quality to the individual full rotation (90°) data sets from single crystals.

The merged reflection files were used for structure refinement to a resolution of 1.9 Å using *REFMAC*5, following molecular replacement with *MOLREP* using an insulin model obtained from the PDB (PDB entry 2bn3) containing only protein atoms (excluding waters). The R value and $R_{\rm free}$ values after just ten cycles of restrained refinement for wedges of X-ray exposures 1–30, 31–60, 61–90, 91–120, 121–150, 151–180 are 0.152 (0.191), 0.154 (0.189), 0.154 (0.187), 0.152 (0.188), 0.151 (0.187) and 0.153 (0.187), respectively (without any manual rebuilding of the model). The results shown in Table 8 illustrate that data with minimal exposures (φ rotation from 0 to 15°) from multicrystal ensembles are sufficient to acquire a complete data set and determine the three-dimensional molecular structure of the protein. No effects of radiation damage were observed in any of our data sets.

3.4. Overlapping reflections

A major challenge in this method is the presence of two types of reflection overlap: those falling close to one another arising from either the same crystal or a different one, and those that originate from different crystals but that fall at exactly the same position on the detector. In the former case, the default settings in *MOSFLM* are used to reject spots that are too close to one another. In the latter case, the reflections have unusual spot intensities and are rejected during the scaling and merging process within *SCALA* (classified as 'rogues').

The number of rejected reflections based on bad profiles are shown for the four-crystal insulin data in Table 7. It can be seen that the numbers are surprisingly low: around 300 for each of the crystals I4-1, I4-2 and I4-3 and around 800 for crystal I4-4 out of about 68 000 reflections. The majority of the rejected reflections for I4-4 are observed in the last quarter of the rotation range (135–180°), where the crystal perhaps moved partially out of the beam. Reducing the I4-4 data without the last 30 exposures only gave about 300 rejected reflections, similar to the numbers observed for the other three data sets (I4-1, I4-2 and I4-3). Hence, a fraction of less than 1% are rejected, which is far less than was expected from the simulation results, in which a fraction of 15–20% of the data were overlapped for similar experimental parameters (see §2.1). This reduction is a consequence of the use of conservative values for parameters in the estimation of the overlap fraction in the simulated data. Looking at the experimental reflection profiles in the output of the *MOSFLM* integration suggests that a cutoff of less than the 11 pixels that were specified in the simulation input might have been chosen as

the criterion for a neighbouring reflection to be classified as overlapped, and that a value as low as 8 or 9 pixels could have been used. If the overlapped reflections are inspected, it is found that generally only the very tail of a peak goes into the peak area of another. Thus, the effect is small and the reflection is not identified as having an odd spot shape. If these overlapping spots are not identified, one might suspect that the accuracy of the reflection intensities would be compromised. However, in the data sets presented here this does not seem be the case, *e.g.* all four data sets of I4 show low $R_{\rm merge}$ and high $\langle I\rangle/\langle\sigma(I)\rangle$. In the more extreme cases where there is severe overlap, *e.g.* caused by large crystal mosaicity, algorithms have been developed to overcome this problem (Hansen *et al.*, 2009; Kazantsev *et al.*, 2009), but these have yet to be tested with MX data sets.

4. Conclusions

New and automated approaches to structure solution are required as MX moves from being used on individual proteins and their complexes to molecular machines and large assemblies. The latter will not generally yield large well diffracting crystals (Dauter, 2006) but often give only small crystals that are susceptible to severe radiation damage. Here, using a suite of software tools developed to handle polycrystal data, multicrystal diffraction data collected from protein crystals have been successfully indexed. Subsequently, standard programs for the integration of single-crystal protein diffraction data can be used to successfully integrate the data sets one by one. Thus, these programs (which are available under an open-source licence) could easily be adopted and adapted by existing software pipelines available in MX.

It is clear that the data do not suffer a significant loss of completeness from the effects of spot overlap. The reflection profiles do not seem to be affected, as only a very few reflections are rejected because of the profile criterion and the internal residuals (*e.g.* $R_{\rm merge}$, $R_{\rm p.i.m.}$) are also low. High residuals are found, *e.g.* in the HEWL seven-crystal data set, but these are a consequence of bad crystal quality and not of spot overlap. Another implication of this multicrystal technique is the possible potential to reduce the problems of X-ray radiation damage to the crystal that are often encountered in structural biology (Ravelli & Garman, 2006). In the data collection of the four-crystal ensemble of cubic insulin (I4), a complete data set could be obtained after a rotation of just 12.5° with a quality comparable to that obtained with a conventional single-crystal diffraction experiment. More importantly, all of the crystals in the ensemble would have been in the early part of their radiation damage-induced decay during this 12.5° exposure, leading to a low-dose data set. The ensemble could be utilized in a data-collection experiment to give a series of data sets with increasing absorbed dose. Each of these could be individually analysed to give a snapshot of the protein state at that stage of the data collection. Thus, data collected in a multicrystal X-ray diffraction experiment could be used to analyse catalytic mechanisms, reduction of metal centres and photo-activated states, leading to better biological understanding. Such experiments are usually performed with single crystals exposed one at a time, leading to poor merging statistics. In a multicrystal experiment in which all the crystals are being cryocooled at the same time in a cryo-loop, better merging statistics would be expected.

The indexing programs used in this study could be used to analyse diffraction images from samples of known space group and unit-cell dimensions. At the time of writing this manuscript, a new indexing program, *index_unknown* (part of the *ImageD*11 program suite), that can process diffraction data from crystals of unknown space groups (for up to ten crystals) has been developed and has successfully indexed a two-crystal data set from an unknown protein structure (data not shown). Furthermore, an extension of the *GrainSpotter* algorithm to handle unknown space groups and a very high number of crystals is presently under development.

This method has significant potential as part of the process of the automation of X-ray diffraction experiments. Thanks to structural genomics initiatives, there has been a rapid acceleration in techniques of molecular biology for protein production and crystal growth, leading to an immense increase in the number of new three-dimensional structures from X-ray crystallography. The use of robots to speed up crystallization trials and the implementation of sophisticated computer-based image analysis even from remote locations (Mayo *et al.*, 2005) has reduced the problems associated with screening for crystallization conditions. However, the major problem in MX remains that of growing crystals of sizes suitable for diffraction, although this has been partially alleviated by the development of high-flux and microfocus beams (Cusack *et al.*, 1998; Sanishvili *et al.*, 2008; Riekel *et al.*, 2005; Yamamoto *et al.*, 2010). The whole structure-determination pipeline is also significantly slowed down by the necessary manual work involved in picking up crystals from the crystallization plates into a cryo-loop even during the initial crystal-screening process.

For screening, this step is obviated by performing *in situ* data collection directly from crystals still within crystallization plates, although multiple crystals are often thus irradiated. New materials for fabricating the plates are being trialled to reduce the background inherent in these methods (Emamzadah *et al.*, 2009). Data sets obtained in this way (*i.e. in situ* at room temperature; 293 K) show very low mosaicity (0.1–0.2°; compared with cryocooled samples, which have a mosaicity of 0.5° and above) which would give sharp reflections on the detector, resulting in fewer spot overlaps in the multicrystal experiment. Developments of an *in situ* screening system coupled to a crystallization-tray imaging robot can be used to screen samples in an automated manner. It should be remembered that although the dose tolerance of crystals is lower at room temperature by around a factor of 70 (Nave & Garman, 2005), the counter-intuitive observation that room-temperature tolerance is higher at greater dose rates (Southworth-Davies *et al.*, 2007) allows reasonable amounts of data to be collected using intense synchrotron beams. These improvements could be enhanced by the use of radical scavengers, which have been found to improve the dose

tolerance of a room-temperature crystal severalfold (Barker *et al.*, 2009).

An extension of *GrainSpotter* could be used to determine the three-dimensional location of a selected crystal in a multicrystal environment; for example, at an automated crystal screening station. By using a defocused beam one can illuminate the multicrystal ensemble and then subsequently use a microbeam to collect data from the best diffracting crystal automatically. Alternatively, phase-contrast X-ray tomography could be used to accurately find and locate the number of crystals in the ensemble (Brockhauser *et al.*, 2008).

In conclusion, data collected from an ensemble of protein crystals in a single loop simultaneously can be used for structure determination. The presence of a potentially large number of overlapping spots does not appear to significantly affect the quality of the data obtained. The programming suite developed outputs data so that existing MX integration and structure-solution software can be used in a conventional way. This method would allow the collection of low-dose complete data sets in a short amount of time, with the possibility of achieving a completely automated data collection from an ensemble of crystals, and play an important role in methodological progress in MX.

KSP and HOS were supported by the EU FP6 program TotalCryst. HOS was further supported by the Danish Research Council FNU (*via* Danscatt). HFP and SS acknowledge support from the Danish National Research Foundation Council. We thank both the ESRF (Radiation Damage BAGs MX-666 and 812) and Diamond Light Source for beamtime. KSP and EFG thank Harry Powell for useful discussions and suggestions.

References

Barker, A. I., Southworth-Davies, R. J., Paithankar, K. S., Carmichael, I. & Garman, E. F. (2009). *J. Synchrotron Rad.* **16**, 205–216.
Bourgeois, D. (1999). *Acta Cryst.* D**55**, 1733–1741.
Bourgeois, D., Nurizzo, D., Kahn, R. & Cambillau, C. (1998). *J. Appl. Cryst.* **31**, 22–35.
Brockhauser, S., Di Michiel, M., McGeehan, J. E., McCarthy, A. A. & Ravelli, R. B. G. (2008). *J. Appl. Cryst.* **41**, 1057–1066.
Busing, W. R. & Levy, H. A. (1967). *Acta Cryst.* **22**, 457–464.
Buts, L., Dao-Thi, M.-H., Wyns, L. & Loris, R. (2004). *Acta Cryst.* D**60**, 983–984.
Cusack, S., Belrhali, H., Bram, A., Burghammer, M., Perrakis, A. & Riekel, C. (1998). *Nature Struct. Biol.* **5**, 634–637.
Dauter, Z. (2006). *Acta Cryst.* D**62**, 1–11.
Diederichs, K. (2009). *Acta Cryst.* D**65**, 535–542.
Diederichs, K. & Karplus, P. A. (1997). *Nature Struct. Biol.* **4**, 269–275.
Emamzadah, S., Petty, T. J., De Almeida, V., Nishimura, T., Joly, J., Ferrer, J.-L. & Halazonetis, T. D. (2009). *Acta Cryst.* D**65**, 913–920.
Evans, P. (2006). *Acta Cryst.* D**62**, 72–82.
Frank, F. C. (1988). *Metall. Mater. Trans A*, **19**, 403–408.
Hansen, P. C., Sørensen, H. O., Sükösd, Z. & Poulsen, H. F. (2009). *SIAM J. Imaging Sci.* **2**, 593–613.
Jacquamet, L., Joly, J., Bertoni, A., Charrault, P., Pirocchi, M., Vernede, X., Bouis, F., Borel, F., Périn, J.-P., Denis, T., Rechatin, J.-L. & Ferrer, J.-L. (2009). *J. Synchrotron Rad.* **16**, 14–21.
Jacquamet, L., Ohana, J., Joly, J., Borel, F., Pirocchi, M., Charrault, P., Bertoni, A., Israel-Gouy, P., Carpentier, P., Kozielski, F., Blot, D. & Ferrer, J.-L. (2004). *Structure*, **12**, 1219–1225.
Jensen, D. J., Lauridsen, E. M., Margulies, L., Poulsen, H. F., Schmidt, S., Sørensen, H. O. & Vaughan, G. B. M. (2006). *Mater. Today*, **9**, 18–25.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Kazantsev, I. G., Schmidt, S. & Poulsen, H. F. (2009). *Inverse Probl.* **25**, 105009.
Lauridsen, E. M., Schmidt, S., Suter, R. M. & Poulsen, H. F. (2001). *J. Appl. Cryst.* **34**, 744–750.
Leslie, A. G. W. (1992). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **26**.
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.
Meister, L. & Schaeben, H. (2005). *Math. Methods Appl. Sci.* **28**, 101–126.
Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.
Nanao, M. H., Sheldrick, G. M. & Ravelli, R. B. G. (2005). *Acta Cryst.* D**61**, 1227–1237.
Nave, C. & Garman, E. F. (2005). *J. Synchrotron Rad.* **12**, 257–260.
Ng, J. D., Clark, P. J., Stevens, R. C. & Kuhn, P. (2008). *Acta Cryst.* D**64**, 189–197.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pflugrath, J. W. (1999). *Acta Cryst.* D**55**, 1718–1725.
Ravelli, R. B. & Garman, E. F. (2006). *Curr. Opin. Struct. Biol.* **16**, 624–629.
Riekel, C., Burghammer, M. & Schertler, G. (2005). *Curr. Opin. Struct. Biol.* **15**, 556–562.
Sanishvili, R., Nagarajan, V., Yoder, D., Becker, M., Xu, S., Corcoran, S., Akey, D. L., Smith, J. L. & Fischetti, R. F. (2008). *Acta Cryst.* D**64**, 425–435.
Sauter, N. K. & Poon, B. K. (2010). *J. Appl. Cryst.* **43**, 611–616.
Schmidt, S., Poulsen, H. F. & Vaughan, G. B. M. (2003). *J. Appl. Cryst.* **36**, 326–332.
Schreurs, A. M. M., Xian, X. & Kroon-Batenburg, L. M. J. (2010). *J. Appl. Cryst.* **43**, 70–82.
Sørensen, H. O. (2008). *PolyXSim – 3DXRD Far-field Simulation of Polycrystals*. Risø National Laboratory for Sustainable Energy, Technical University of Denmark.
Southworth-Davies, R. J., Medina, M. A., Carmichael, I. & Garman, E. F. (2007). *Structure*, **15**, 1531–1541.
Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* D**66**, 22–25.
Vaughan, G. B. M., Schmidt, S. & Poulsen, H. F. (2004). *Z. Kristallogr.* **219**, 813–825.
Watanabe, N., Murai, H. & Tanaka, I. (2002). *Acta Cryst.* D**58**, 1527–1530.
Weiss, M. S. & Hilgenfeld, R. (1997). *J. Appl. Cryst.* **30**, 203–205.
Whittingham, J. L., Youshang, Z., Žáková, L., Dodson, E. J., Turkenburg, J. P., Brange, J. & Dodson, G. G. (2006). *Acta Cryst.* D**62**, 505–511.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.
Yadav, M. K., Gerdts, C. J., Sanishvili, R., Smith, W. W., Roach, L. S., Ismagilov, R. F., Kuhn, P. & Stevens, R. C. (2005). *J. Appl. Cryst.* **38**, 900–905.
Yamamoto, M., Hirata, K., Hikima, T., Kawano, Y. & Ueno, G. (2010). *Yakugaku Zasshi*, **130**, 641–648.