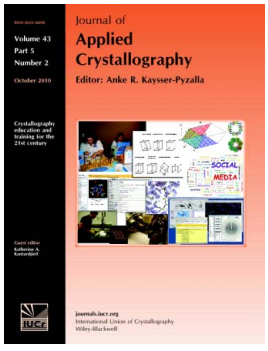
Journal of
Applied Crystallography
ISSN 0021-8898
Editor: **Anke R. Kaysser-Pyzalla**

Update on the tutorial for learning and teaching macromolecular crystallography

Annette Faust, Sandra Puehringer, Nora Darowski, Santosh Panjikar, Kay Diederichs, Uwe Mueller and Manfred S. Weiss

J. Appl. Cryst. (2010). **43**, 1230–1237



Journal of
Applied
Crystallography
ISSN 0021-8898

Received 14 March 2010
Accepted 14 July 2010

Update on the tutorial for learning and teaching macromolecular crystallography

Annette Faust,[a] Sandra Puehringer,[b] Nora Darowski,[b] Santosh Panjikar,[c] Kay Diederichs,[d] Uwe Mueller[b] and Manfred S. Weiss[b]*

[a]Centre for Biochemistry and Molecular Biology (ZBM), Christian-Albrechts-University of Kiel, Am Botanischen Garten 1–9, D-24118 Kiel, Germany, [b]Helmholtz-Zentrum Berlin für Materialien und Energie GmbH, Macromolecular Crystallography (BESSY-MX), Albert-Einstein-Strasse 15, D-12489 Berlin, Germany, [c]EMBL Hamburg Outstation, c/o DESY, Notkestrasse 85, D-22603 Hamburg, Germany, and [d]Department of Biology, University of Konstanz, Box M647, D-78457 Konstanz, Germany. Correspondence e-mail: msweiss@helmholtz-berlin.de

Two new experiments (single isomorphous replacement including anomalous-scattering effects and radiation damage-induced phasing) have been designed to complement the five experiments (sulfur single-wavelength anomalous diffraction, multiple-wavelength anomalous diffraction, molecular replacement, ion binding and ligand binding) of the first edition of the previously described tutorial for learning and teaching macromolecular crystallography [Faust, Panjikar, Mueller, Parthasarathy, Schmidt, Lamzin & Weiss (2008). *J. Appl. Cryst.* **41**, 1161–1172]. Furthermore, the tutorial has been re-organized and in part re-written to reflect the comments and suggestions of users. The most significant overhaul was applied to the data-processing part of the tutorial. Nevertheless, the convenient features that all of the utilized proteins used are commercially available and that they can be easily and reproducibly crystallized and mounted for diffraction data collection have been retained. Also, for all of the seven experiments the raw diffraction images and the processed data are provided for illustrating and teaching the steps of data processing and structure determination.

1. Introduction

Two years ago, we began to assemble a number of experiments that can be utilized to satisfy the growing needs for teaching various aspects and approaches of macromolecular crystallography (MX), mainly to a biological community (Jaskólski, 2001; Faust *et al.*, 2008*a*). Our efforts were motivated by the increased importance of MX in the greater context of biochemical, biological and biomedical research, a fact that is convincingly illustrated by the growth statistics of the Protein Data Bank (Berman *et al.*, 2000). Despite the enormously prominent role of MX in the biosciences, there is a striking scarcity of training tools that can be employed for teaching a generation of researchers with a mainly biological background. Needless to say there are courses, schools and workshops that are offered by people dedicated to teaching, such as the M2M course organized by the EMBL Hamburg, the RapiData course in Brookhaven, the EMBO course on anomalous scattering at the ESRF in Grenoble, the Cold Spring Harbor Crystallography School and certainly many more. However, because of their limited capacity for accepting students, these courses are in no way able to satisfy the needs of the community for high-level MX teaching. Furthermore, very few organizers make their course material publicly available, and even fewer go as far as to make diffraction data available to the community for self-study purposes. In order to address this shortage of well documented and described material, we set out to design a tutorial consisting of typical MX experiments. These experiments were originally developed for and utilized in the German Society of Crystallography (DGK) workshop on '*Diffraction Data Collection Using Synchrotron Radiation*', which was held at the BESSY II synchrotron radiation source in Berlin-Adlershof in the summer of 2007 (see also http://www.embl-hamburg.de/workshops/2007/diffraction/). The principles behind the design of the first edition of the tutorial, with five experiments, were that all material should be readily available, that all experiments should be easily reproducible at either a home or a synchrotron source, and that all data accumulated during the course of a structure determination should be available for download so that any interested user can enter the process at various entry points. The five experiments of the original tutorial were (i) a sulfur single-wavelength anomalous diffraction (S-SAD) structure determination of cubic insulin, (ii) a bromide multiple-wavelength anomalous diffraction (MAD) structure determination of thaumatin, (iii) a structure determination by molecular replacement using monoclinic hen egg-white lysozyme (HEWL), (iv) the identification of

bound solvent ions in tetragonal crystals of HEWL using a longer-wavelength data set and (v) the identification of a weakly bound ligand in the active site of HEWL. With these five experiments some of the most common scenarios encountered in structural biology using macromolecular crystallographic techniques were represented in the tutorial. However, despite its successful use, it became quickly apparent that the list of experiments was not complete. In order to remedy that, we have now complemented the first edition of the tutorial by adding two further experiments, (vi) a structure determination by single isomorphous replacement including anomalous scattering (SIRAS) using a gold derivative of tetragonal HEWL and (vii) a structure determination by ultraviolet radiation damage-induced phasing (UV-RIP) on the sweet protein thaumatin. The second edition of the tutorial now consists of seven experiments, which cover the whole breadth of approaches for structure determination using MX (Fig. 1).

2. The new projects

2.1. SIRAS on tetragonal lysozme

Since a detailed introduction on lysozyme has been given in the first version of the tutorial (Faust *et al.*, 2008*a*), only the most important features of HEWL will be repeated here. With 129 amino acids, HEWL is a relatively small protein. It was the first enzyme whose three-dimensional structure was determined (Blake *et al.*, 1965), and since then it has become a model protein for many systematic studies in MX and for teaching purposes. A very interesting feature of HEWL is that it can be crystallized in many different crystal forms (for an overview see Brinkmann *et al.*, 2006).

Isomorphous replacement is the traditional method of phase determination in MX. It is based on comparing the structure factor amplitudes of a native protein crystal ($|\mathbf{F}_{\rm P}|$) with those of a protein crystal that has been derivatized by co-crystallization or soaking with a heavy-atom-containing compound ($|\mathbf{F}_{\rm PH}|$). Under the assumption that both crystals are isomorphous, the differences ($|\mathbf{F}_{\rm PH}| - |\mathbf{F}_{\rm P}|$) can be used to identify the positions of the heavy atoms. The resulting heavy-atom structure factors ($\mathbf{F}_{\rm H}$) can then be used for phase determination. This process is shown graphically in the form of the so-called Harker construction (Fig. 2). It is based on the validity of the vector equation $\mathbf{F}_{\rm P} + \mathbf{F}_{\rm H} = \mathbf{F}_{\rm PH}$. If the magnitudes of $\mathbf{F}_{\rm P}$ and $\mathbf{F}_{\rm PH}$ are known (these are the measured structure factor amplitudes $|\mathbf{F}_{\rm P}|$ and $|\mathbf{F}_{\rm PH}|$ of the native and the derivative data sets) and if $\mathbf{F}_{\rm H}$ is known as vector (this means that the heavy-atom structure is known), the phase α can be determined. In the case of single isomorphous replacement (SIR; Fig. 2*a*), however, the phase determination yields two values (phase ambiguity), while in the case of multiple isomorphous replacement (not shown) and SIRAS (Fig. 2*b*) an unambiguous phase determination can be achieved.

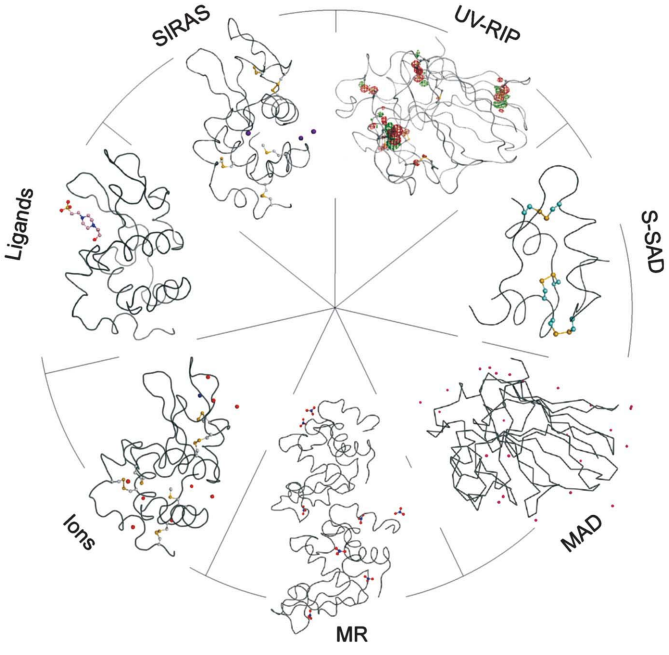


Figure 1
The seven projects of the second edition of the MX tutorial. S-SAD on cubic insulin, MAD on thaumatin, molecular replacement (MR) on monoclinic lysozyme, solvent ion identification in tetragonal lysozyme, ligand identification in the MPD form of tetragonal lysozyme, SIRAS on tetragonal lysozyme and UV-RIP on thaumatin.

2.2. UV-RIP on thaumatin

As for lysozyme, a detailed introduction on thaumatin has been presented in the previous version of the tutorial (Faust *et al.*, 2008*a*). Thaumatin is a naturally occurring sweet protein consisting of 207 amino acids. It contains a total of 16 Cys residues, which form eight disulfide bridges.

Structure determination by RIP is a very recent phasing technique, which is based on the X-ray- or UV-induced specific radiation damage to protein crystals (Ravelli *et al.*, 2005; Nanao & Ravelli, 2006; Schoenfeld *et al.*, 2008). In a typical RIP experiment, an initial diffraction data set with minimal exposure dose is collected, then the crystal is exposed to high doses of X-rays or UV in the so-called burn phase and a second diffraction data set is collected, again with minimal exposure dose. The two data sets, typically termed 'before' and 'after', now represent two states of the protein: an undamaged state and a damaged state. If the two states are sufficiently different and if the structural differences can be identified and modeled easily, then the 'before' and 'after' data sets can be used in a SIR-type approach for phase determination. The method works particularly well when the protein contains disulfide bridges.

3. Crystallization and heavy-atom derivatization

All crystallization experiments were carried out using the hanging-drop method at room temperature (293 K) in 24-well EasyXtal Tool screw-cap crystallization plates (Qiagen), which unfortunately are no longer commercially available. As a slightly less convenient replacement, greased Linbro plates with glass cover slides may be used instead. Tetragonal lyso-

zyme crystals and thaumatin crystals were produced as described in the first version of the tutorial (Faust *et al.*, 2008*a*). All chemicals used for the two experiments were from Sigma, unless specified otherwise.

3.1. Tetragonal lysozyme

Tetragonal crystals of HEWL were grown as described by Weiss *et al.* (2000) and by Faust *et al.* (2008*a*). Protein solution (4 µl, 30 mg ml^{-1} in water) was mixed with 4 µl of reservoir solution containing 50 m*M* sodium acetate pH 4.6 and 5%(*w*/*v*) sodium chloride and equilibrated against 1 ml of reservoir. Crystals belonging to space group $P4_32_12$ and exhibiting the usual unit-cell parameters of $a = 78.6$, $c = 36.8$ Å appeared within a few days. For heavy-atom derivatization, a 10 m*M* solution of $KAuCl_4$ in reservoir solution was freshly prepared and one crystal was soaked in this solution for 1 min (Sun *et al.*, 2002). While the native crystals diffracted to beyond 1.6 Å resolution, the diffraction properties of the derivatized crystals were slightly worse. Nevertheless, they still diffracted to about 1.8 Å resolution.

3.2. Thaumatin

Thaumatin crystals were grown as described by Mueller-Dieckmann *et al.* (2005). Protein solution [2 µl, 15 mg ml^{-1} in 0.1 *M* *N*-(2-acetamido)iminodiacetic acid (ADA) pH 6.5] was mixed with 2 µl of reservoir solution (0.1 *M* ADA pH 6.5, 1 *M* sodium/potassium tartrate) and equilibrated against 1 ml of reservoir. Tetragonal crystals of space group $P4_12_12$ and with unit-cell parameters $a = 57.9$, $c = 150.2$ Å appeared after a few days. They diffracted X-rays to at least 1.5 Å resolution.

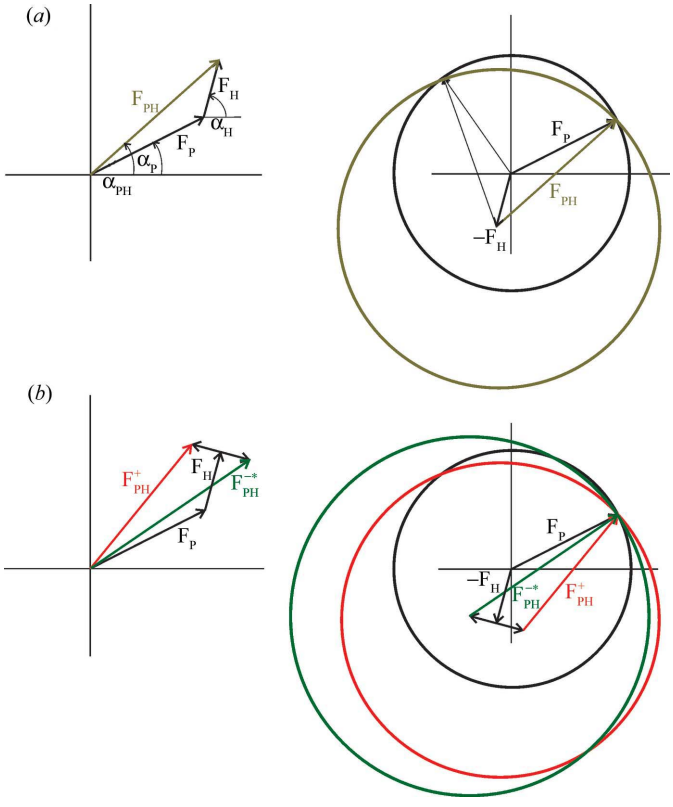
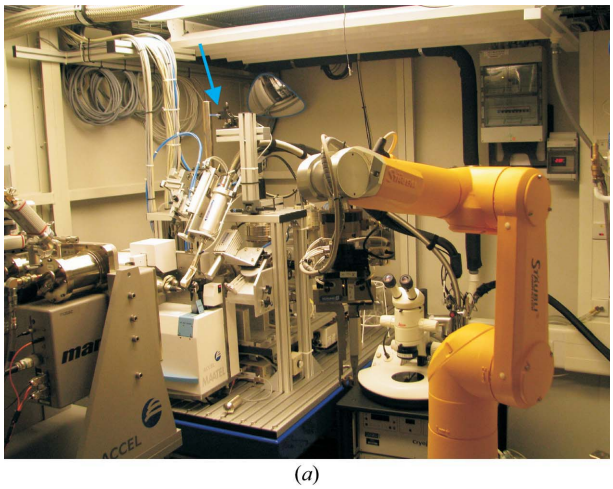


Figure 2
Harker construction. (*a*) SIR case. (*b*) SIRAS case. In (*a*), the two circles intersect at two different points, giving rise to the so-called phase ambiguity. In (*b*), the red, green and black circles intersect at one point, giving rise to an unambiguous phase determination.

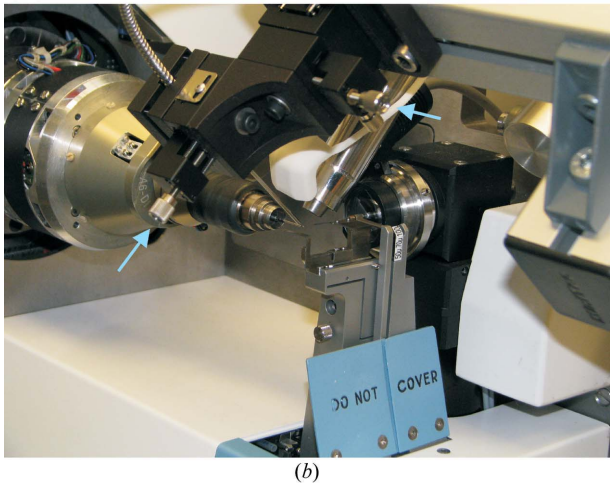
4. Diffraction data collection and processing

For the diffraction experiment the crystals were mounted in Nylon loops of suitable size. Subsequently, they were cryo-protected and shock-cooled in a cold nitrogen gas stream directly on the beamline. Both native and derivatized tetragonal HEWL crystals were cryo-protected by transferring them to a solution containing 25%(*v*/*v*) ethylene glycol, 10%(*w*/*v*) NaCl and 100 m*M* Na acetate pH 4.5. Thaumatin crystals were cryo-protected by soaking them for a few seconds in 25%(*v*/*v*) glycerol in water. For the RIP experiment described here it is particularly important to choose a thaumatin crystal that is not too large (100–150 µm in the longest dimension) and to make sure that before shock-cooling not too much mother liquor is left surrounding the crystal. Diffraction data were then collected on the beamlines BL14.1 and BL14.2 (Freie Universität Berlin and Helmholtz-Zentrum Berlin für Materialien und Energie, BESSY II synchrotron storage ring, Berlin-Adlershof, Germany) (Heinemann *et al.*, 2003). Both beamlines are equipped with a Rayonics MX-225, 3 × 3 tiled CCD detector. On BL14.2 the detector is mounted on a MARdtb goniometer system (MarResearch, Norderstedt, Germany), while beamline BL14.1 is equipped with an MD2 microdiffractometer (Maatel, Voreppe Cedex, France). On BL14.1, a 266 nm pulsed microchip UV laser (TEEM Photonics; http://www.teemphotonics.com) is permanently installed for performing UV-RIP experiments (U. Mueller *et al.*, in preparation). The BL14.1 setup is shown in Fig. 3.

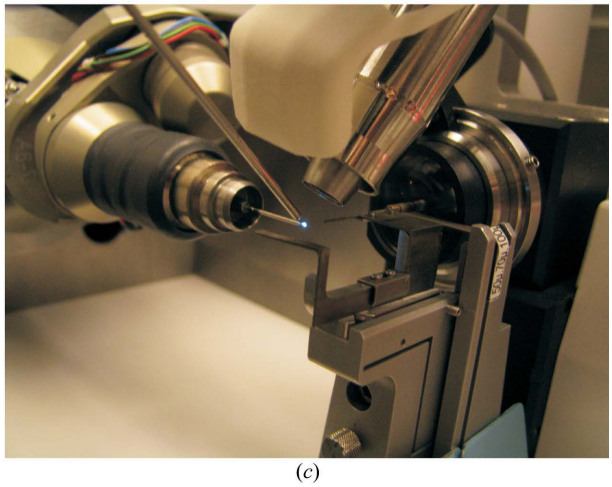
For the SIRAS experiment on HEWL, diffraction data were collected over a range of 180° at a wavelength of 1.00 Å. The starting angle was chosen arbitrarily. For the UV-RIP experiment on thaumatin, the data-collection wavelength was also 1.00 Å and the data-collection sequence was as follows: at first a minimal-exposure data set was collected, then the crystal was exposed for 15 min to a beam from the UV laser and a second minimal-exposure data set was collected over the same rotation range as the first one. The starting angle for the first data set was determined using the strategy option of *iMOSFLM* (Leslie, 1992). The second data set was collected covering the same wedge as the first one. The relevant data-collection parameters are summarized in Table 1. During the UV exposure the crystal was completely bathed in the laser beam (beam diameter approximately 600 µm) and rotated. Indexing, integration and scaling of the data sets were carried out using the programs *XDS* and *XSCALE* (Kabsch, 1993, 2010*a*,*b*). All respective data-processing statistics for the various data sets are given in Table 2. The whole procedure of how the data were processed has been described extensively in the first edition of the tutorial (Faust *et al.*, 2008*a*). Even though the new version of the tutorial contains a description of the data processing that has been completely re-written in



(*a*)



(*b*)



(*c*)

Figure 3
Setup of beamline BL 14.1 (BESSY II, Berlin) (U. Mueller *et al.*, in preparation). (*a*) Global view of the experimental hutch. Clearly discernible are the CCD detector (left), the MD2 diffractometer with the mounted cryo-system (center) and the robot arm of the CATS sample changer (right) (Irelec, Saint-Martin-d'Hères, France). The UV laser source (indicated by a blue arrow) is mounted on a movable arm. The current view is in the resting state. (*b*) Close-up view of the sample position with the UV laser moved down into the experiment state. The two screws (indicated by the light-blue arrows) on the *xy* table (black) are used for aligning the UV laser onto the crystalline sample. (*c*) Further close-up view of the sample environment. The crystal is fluorescing in the UV beam.

Table 1
Data collection parameters for the two projects.

	Lysozyme (Experiment 6)		Thaumatin (Experiment 7)	
	Native	$KAuCl_4$ derivative	'Before'	'After'
No. of crystals	1	1	1	
Beamline	BL14.2	BL14.2	BL14.1	BL14.1
Wavelength (Å)	1.00	1.00	1.00	1.00
Crystal–detector distance (mm)	180	180	180	180
Rotation range per image (°)	1.0	1.0	1.0	1.0
Total rotation range (°)	180	180	90	90
Exposure time per image (s)	2.5	5	2.0	2.0
Attenuation (mm Al)	–	–	0.29	0.29

order to incorporate the many comments we have received from users over the past two years, the basic principles and the steps of data processing have remained the same. The revised detailed description of the data processing including the relevant lines of the input files is provided in the tutorials, which can be downloaded from http://www.helmholtz-berlin.de/bessy-mx-tutorial or obtained from the authors upon request.

5. Structure determination

The focus of the workshop for which this tutorial was originally designed was on crystal mounting, data collection and processing. The subsequent structure determination was used merely as validation of the success of the diffraction experiment. Because of its ease of use, all structure-determination steps were carried out automatically using the respective protocols of *Auto-Rickshaw* (Panjikar *et al.*, 2005, 2009). *Auto-Rickshaw* can be accessed at http://www.embl-hamburg.de/Auto-Rickshaw/ (registration may be required, which is free for academic users). Needless to say that the tutorial may also be used with the emphasis placed on teaching the successive steps of structure determination. In such cases, the structure determination may also be performed the conventional manual way. The numbers provided here may then simply serve as a rough guide of what can be achieved given a certain data set.

For the SIRAS experiment on HEWL, phases need to be derived from the measured isomorphous and anomalous differences, while for the RIP experiment on thaumatin, the isomorphous differences are the only source of phase information. The first step in the phase-determination process is the calculation of ΔF values. This can be done, for instance, with the program *SHELXC* (Sheldrick *et al.*, 2001; Sheldrick, 2008). *SHELXC* will also write an input file for the subsequent program *SHELXD* (Schneider & Sheldrick, 2002), which is widely used to determine isomorphous or anomalous substructures. Since a *SHELXD* calculation is started based on randomly placed atoms, it can and should be run many times (a typical number of *SHELXD* cycles is 100, but it is often worthwhile to run 1000 or even 10 000 cycles). A correct solution can then be identified by looking at the two correlation coefficients CC(All) and CC(weak). Values of 30.0 and

Table 2
Data-processing statistics for the two projects.

The numbers in parentheses define the outermost resolution shell used.

	Lysozyme (Experiment 6)		Thaumatin (Experiment 7)	
	Native	$KAuCl_4$ derivative	'Before'	'After'
Resolution range (Å)	50.0–1.60 (1.70–1.60)	50.0–1.80 (1.91–1.80)	30.0–1.60 (1.70–1.60)	
Space group	$P4_32_12$	$P4_32_12$	$P4_12_12$	$P4_12_12$
Unit-cell parameters (Å, °)	a = 78.62, c = 36.81	a = 78.73, c = 36.73	a = 57.89, c = 150.15	a = 57.89, c = 150.18
Mosaicity (°)	0.13	0.17	0.06	0.05
Total No. of reflections	187 497	152 422	205 403	205 771
Unique reflections	28 363	20 439	63 215	63 237
Completeness (%)	97.4 (86.3)	99.5 (97.5)	98.3 (91.0)	99.4 (97.9)
Redundancy	6.6 (3.4)	7.5 (7.1)	3.2 (1.6)	3.3 (1.6)
$I/\sigma(I)$	27.4 (5.2)	15.6 (3.0)	12.5 (2.1)	12.0 (1.8)
$R_{r.i.m.}/R_{meas}$ (%)	4.8 (24.6)	9.9 (71.1)	8.4 (44.7)	8.9 (51.7)
Overall Wilson B factor (Å^2)	20.3	28.3	17.1	17.8

15.0 indicate a correct solution, although correct solutions may appear with even lower values. The next step is the actual phase calculation and the improvement of the phase by density modification. This can be achieved by using the program *SHELXE* (Sheldrick, 2002), but also programs such as *MLPHARE*, *BP3* (Pannu *et al.*, 2003), *SHARP* (Bricogne *et al.*, 2003) and *DM* (Collaborative Computational Project, Number 4, 1994). At this point, the correct hand of the substructure has also to be established. Once the phases have been determined an electron-density map can be calculated, for instance, using the program *COOT* (Emsley & Cowtan, 2004), and displayed and examined for parts that can be interpreted in terms of an amino acid chain or a secondary structure element. With high enough resolution at hand, however, it is also possible to attempt a completely automatic density interpretation and model building using a β version of *SHELXE* (Sheldrick, 2010) and *ARP*/*wARP* 7.0 (Perrakis *et al.*, 1999; Morris *et al.*, 2002).

5.1. SIRAS on tetragonal lysozyme

The structure was solved using the SIRAS protocol of *Auto-Rickshaw*. The input diffraction data files XDS_ASCII.HKL were uploaded and then prepared and converted using programs of the *CCP4* suite (Collaborative Computational Project, Number 4, 1994). ΔF and $|\,|\mathbf{F}_{PH}| - |\mathbf{F}_{P}|\,|$ values were calculated using the program *SHELXC*. Based on an initial analysis of the data, the maximum resolution for substructure determination and initial phase calculation was set to 2.4 Å. Three heavy-atom sites (Table 3) were found by using the program *SHELXD*, with correlation coefficients of 36.0 and 23.7% for all and for just the weak reflections, respectively. The correct hand for the substructure was determined using the programs *ABS* (Hao, 2004) and *SHELXE*. Initial phases were calculated after density modification using the program *SHELXE*. A poly-Ala model was automatically built using the

Table 3
List of heavy-atom peaks in orthogonal Å coordinates found using *SHELXD* and their correspondence to the three Au sites.

Au site	x	y	Z	Occupancy
1	67.093	11.725	0.909	1.00
2	69.980	9.972	−4.114	0.69
3	67.297	11.323	−9.203	0.33

programs *SHELXE* and *ARP*/*wARP*. The resulting model was 95% complete. In order to further complete the model and to enhance the phases, the MRSAD module of *Auto-Rickshaw* (Panjikar *et al.*, 2009) was used. This resulted in a model consisting of 127 residues, of which 125 were docked in the amino acid sequence. This model was subsequently refined against the $KAuCl_4$-derivative data using *REFMAC5* (Murshudov *et al.*, 1997) and further modified using *COOT*. The final R factor for the model was 21.8% and the free R factor was 27%. The refined structure is shown schematically in Fig. 4.

5.2. UV-RIP on thaumatin

The structure was solved using the RIP protocol of *Auto-Rickshaw*. The input diffraction data were prepared and converted for use in *Auto-Rickshaw* using programs of the *CCP4* suite. The scale factor applied to the 'after' data set was set to the default value of 0.966 (Nanao *et al.*, 2005; Schoenfeld *et al.*, 2008). $|\mathbf{F}_{after} - \mathbf{F}_{before}|$ values were calculated using the program *SHELXC*. Based on an initial analysis of the data,

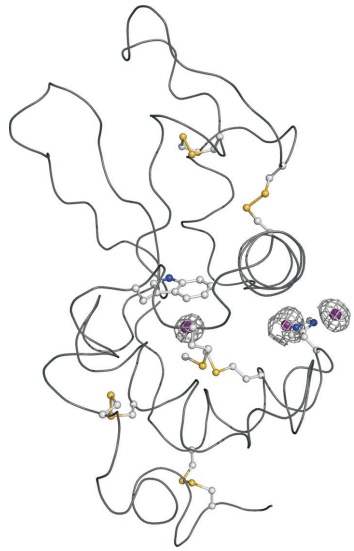


Figure 4
Refined structure of the $KAuCl_4$ derivative of tetragonal HEWL presented as a thin tube (in gray). Also shown are the Met and Cys side chains of HEWL (as ball-and-stick) and the bound Au atoms (in purple). Superimposed on the structure is the anomalous difference Fourier electron-density map contoured at 4.0σ. The first Au atom is located 3.6 Å from the SD atom of Met105 and 4.4 Å from the NE atom of Trp108. The second and the third Au atoms are located on both sides of the side chain of His15. One is located 2.1 Å from the NE2 atom, while the other is found 2.0 Å from the ND1 atom of His15.

the maximum resolution for substructure determination and initial phase calculation was set to 1.9 Å. All of the 20 damage sites requested were found using the program *SHELXD* with correlation coefficients of 25.7 and 11.2% for all and for just the weak reflections, respectively. The correct hand for the substructure was determined using the program *SHELXE*. Initial phases were calculated after density modification using the program *SHELXE*. Of the 217 amino acids, 197 (90.8%) were built automatically using the program *ARP/wARP*. The model was then extended and refined against the 'before' data set using *COOT* and *REFMAC5*. The *R* factors for the final model were $R = 16.6\%$ and $R_{\rm free} = 19.3\%$. The peaks in Fig. 5 and Table 4 identify the sites in the structure that were affected during the exposure to UV radiation.

6. Challenges and didactic value

All experiments have been designed such that they can easily be carried out by beginners in the field but are also challenging in some respect. All proteins used for crystallization are available commercially, and the crystallization experiments themselves usually work reliably and reproducibly. The experiments as they are described have been carried out by students, albeit under supervision of tutors, during two practical workshops in the years 2007 and 2009 (see §8 below). Some general points about the didactic value of the various experiments have already been made by Faust *et al.* (2008*a*). These points do still apply to the second edition. In response to the many comments and suggestions we received from users of the tutorial, the data-processing part was completely rewritten. It is still based on the program *XDS* (Kabsch, 1993, 2010*a*,*b*), but it is now organized in a step-by-step procedure, where for each step the relevant parameters that need to be looked at are described. Needless to say that other data-processing programs such as *MOSFLM* (Leslie, 1992), *DENZO* or *HKL2000* (Otwinowski & Minor, 1997) may be used just as well.

In the following, the two new experiments of the tutorial are discussed in more detail.

Table 4
The local maxima and minima of the difference electron-density ($\mathbf{F}_{\rm after} - \mathbf{F}_{\rm before}, \alpha_{\rm calc}$) map for the thaumatin UV-RIP data identifying the sites at which radiation damage occurs.

					Damage site	
Site No.	*X*	*Y*	*Z*	Peak height (σ)	Closest atom	Structure
1	18.74	31.68	22.47	−39.49	Cys56-SG	S bridge with Cys68
2	9.03	3.12	13.61	−34.17	Cys68-SG	S bridge with Cys56
3	7.21	8.81	34.45	−33.01	Cys193-SG	S bridge with Cys121
4	5.66	7.71	33.66	−24.67	Cys121-SG	S bridge with Cys193
5	20.77	32.20	25.12	+24.65	Cys68-SG	New position
6	31.47	−3.42	40.91	−22.40	Cys126-SG	S bridge with Cys177
7	25.05	39.60	51.69	+20.69	Cys56-SG	New position
8	5.70	38.06	35.92	−18.96	Cys134-SG	S bridge with Cys145
9	5.95	36.03	36.26	−18.09	Cys145-SG	S bridge with Cys134
10	5.66	8.40	34.99	+17.28	Cys121-SG	New position
11	1.37	26.54	34.91	−16.15	Cys177-SG	S bridge with Cys126
12	8.77	8.22	34.18	+16.12	Cys193-SG	New position
13	23.99	−0.81	14.53	−13.31	Cys204-SG	S bridge with Cys9
14	32.46	−4.63	40.89	+11.18	Cys126-SG	New position
15	10.87	27.78	51.76	−9.71	Cys164-SG	S bridge with Cys159
16	16.45	33.82	47.14	−9.65	Cys158-SG	S bridge with Cys149
17	6.86	27.55	22.12	−9.55	Cys9-SG	S bridge with Cys204
18	16.23	32.96	45.22	−9.12	Cys149-SG	S bridge with Cys158
19	19.70	32.72	20.79	-−8.19	Cys56-CB	S bridge with Cys66
20	23.83	27.56	27.39	−8.10	Cys71-SG	S bridge with Cys77
21	18.24	36.21	22.05	−7.17	Asp55-O	–
22	15.01	34.79	24.26	+6.88	Thr54-O	New position
23	16.25	34.06	23.77	−6.88	Thr54-O	–
24	12.13	28.43	49.95	−6.84	Cys159-SG(B)	S bridge with Cys164
25	5.22	38.52	34.54	+6.80	fCys134-SG	New position
26	4.61	9.56	30.37	−6.44	Pro116-CG	–
27	13.48	34.24	25.18	−6.37	Thr54-OG1	–
28	8.58	7.01	35.45	−6.37	Cys193-CB	S bridge with Cys121
29	22.44	26.04	27.17	−6.28	Cys77-SG	S bridge with Cys71
30	6.05	10.00	31.32	+6.21	Pro116-CG	New position
31	2.17	25.24	34.38	+6.19	Cys177-SG	New position
32	2.87	27.82	22.61	+6.08	Cys204-SG	New position
33	9.05	22.72	31.27	−6.05	Met112-SD	–

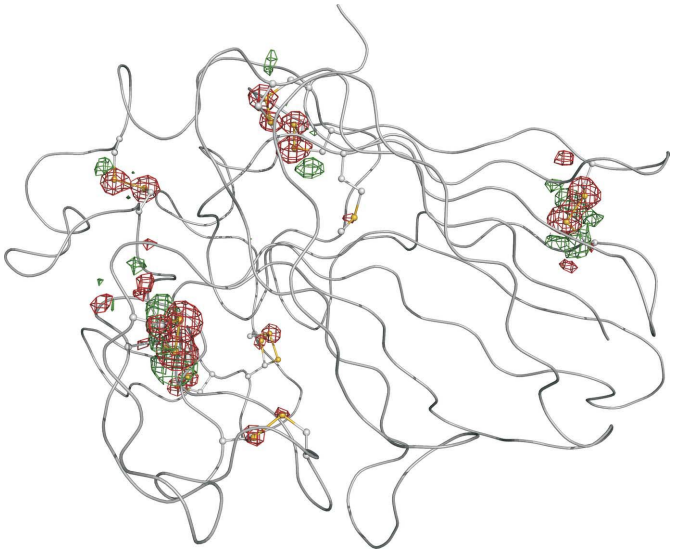


Figure 5
Difference electron density ($\mathbf{F}_{\rm after} - \mathbf{F}_{\rm before}, \alpha_{\rm calc}$) map superimposed onto a Cα trace of thaumatin showing all the sites that were affected as a result of the UV exposure. The map is contoured at +5.0σ (green) and −5.0σ (red). The positive peaks identify structural elements that appear after the damage, while the negative peaks identify sites that are destroyed by the damage.

6.1. SIRAS on tetragonal lysozyme

While isomorphous replacement experiments constitute the traditional approach to phase determination in MX, they are rarely used nowadays. Nevertheless, the method should still be taught, since the obtainable signal is much larger than what can be achieved in a typical MAD or SAD experiment. Consequently, obtaining initial phase information for large complexes such as ribosome still requires heavy-atom derivatization techniques. The most challenging part of the SIRAS experiment described here is the actual heavy-atom derivatization by soaking. When using a freshly prepared solution of $KAuCl_4$, a soaking time of 60 s is sufficient to yield a usable heavy-atom derivative. Soaking times of three minutes or longer severely reduce the diffraction power of the crystal. It

may be the case that older solutions of $KAuCl_4$ require different soaking times, but this was not investigated in this experiment. In contrast to the first edition of the tutorial, the cryo-protection of the crystals was achieved with ethylene glycol rather than paraffin oil, which seems to work better and in a more reproducible manner. The actual data collection is straightforward and poses no problems. The data-collection wavelength was chosen arbitrarily as 1.0 Å. One could, however, optimize the anomalous signal by performing an X-ray fluorescence scan in a similar fashion as for the original experiment No. 2 (Faust *et al.*, 2008*a*) and by choosing a wavelength that corresponds to one of the Au *L*-absorption edges. An important aspect of the data collection is that 180° of data appear to be necessary for successful substructure determination using *SHELXD*; just 90° of data were not sufficient. For teaching purposes, it may be useful to interpret the isomorphous difference Patterson and the anomalous Patterson maps (Fig. 6) by hand, determine the coordinates of the major gold site, and continue from this point onwards.

6.2. UV-RIP on thaumatin

This experiment serves two purposes. First of all, it constitutes a fairly recent addition to the arsenal of phase determination techniques in MX, and second, it demonstrates to students in a very convincing way how UV radiation (or X-rays for that matter) can damage protein molecules. Since radiation damage cannot be avoided in MX data-collection experiments and since it is always present, students should be aware of it at various stages: when they plan the experiment but also when they interpret their structural results. The most important aspect of this experiment is that the crystal for the experiment should not be too big and that the crystal should be mounted with as little solvent around it as possible. This requirement is a consequence of the limited penetration depth of the 266 nm UV laser radiation that is used for introducing the damage. Otherwise, this experiment is straightforward. Just note that when working with a laser, particularly an invisible class 4 UV laser, great care must be taken and suitable eye protection must be worn during the experiment. Furthermore, since it is important that most of the damage to the crystal is introduced during the UV exposure, one should also ensure that the 'before' and 'after' data sets are collected with minimal exposure to X-rays. This makes the use of a strategy program a must. The actual structure determination based on UV-RIP works best when the resolution of the data extends to about 2.0 Å. This limits the applicability of UV-RIP somewhat, but one has to keep in mind that UV-RIP is a fairly new method and that it may still require the development of new and better approaches in how to best deal with such data.

7. Conclusion

With the here presented second edition of the MX tutorial describing (i) a sulfur-SAD structure determination of cubic insulin, (ii) a bromide-MAD structure determination of thaumatin, (iii) a structure determination by molecular replacement using monoclinic lysozyme, (iv) the identification of bound solvent ions in lysozyme using a longer-wavelength data set, (v) the identification of a weakly bound ligand at the active site of lysozyme, (vi) a structure determination by SIRAS using tetragonal lysozyme and (vii) a structure determination *via* UV-RIP on thaumatin, we attempted to cover the whole breadth of MX experiments. To the best of our

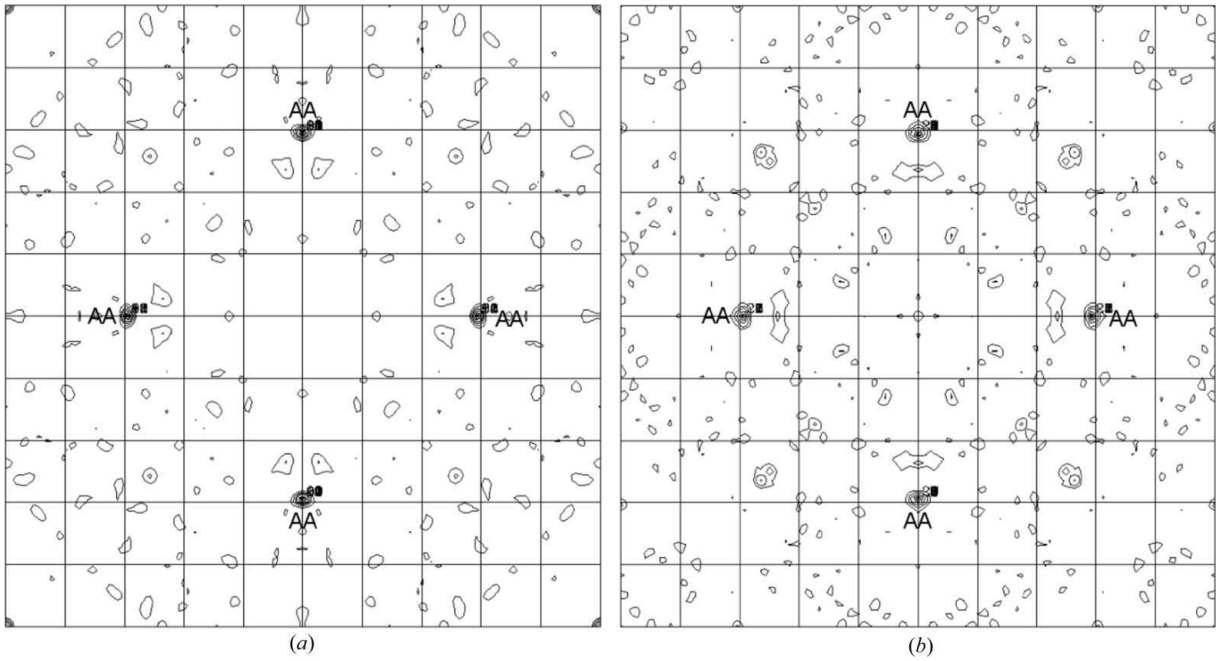


Figure 6
Difference (*a*) and anomalous difference (*b*) Patterson maps. Shown are the Harker sections w = 1/4 in both cases. The peak denoted AA and the corresponding symmetry-related peaks correspond to the Harker peaks of the Au site No. 1.

knowledge this tutorial is unique in that it provides a complete description of a crystallographic experiment covering all aspects from crystallization and crystal mounting, to diffraction data collection and processing as well as structure determination and interpretation. It details not only the crystallization recipes but also the sources of all relevant materials as well as instructions on how to perform the experiments. In addition, example diffraction data are provided for those who do not have easy access to a diffraction facility, together with information on the various steps from data processing to structure determination. All of the computer programs used and described in the tutorial are freely available for academic users.

8. Availability of the data and the tutorial

All data composing the tutorial, from the raw diffraction images for all seven experiments, and the integrated and scaled data, to the solved and refined structures, are available for download from http://www.helmholtz-berlin.de/bessy-mx-tutorial. Accompanying the experiments, detailed descriptions of the data processing and structure solution are also available for download or directly from the authors upon request. The whole tutorial comprises about 16 GB of data, which are conveniently distributed over 54 g'zipped tar-archive (.tgz) files of sizes of 700 MB or less. The tutorial has been successfully employed during two practical workshops in the years 2007 and 2009 (see http://www.embl-hamburg.de/workshops/2007/diffraction and http://www.embl-hamburg.de/workshops/2009/diffraction). Furthermore, it has been used and favorably received by many individual researchers and students to date, who have learned about it from colleagues participating in one of the workshops, by word of mouth, or from presentations at the 16th Annual Meeting of the German Society of Crystallography and the XXI Congress and General Assembly of the International Union of Crystallography (Faust *et al.*, 2008*b*).

We would like to acknowledge the support of the DGK for their financial contribution towards the two workshops mentioned in this paper. We would also like to thank the students and the tutors of the workshops as well as the many users of the tutorial who have provided us with constructive feedback.

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.
Blake, C. C. F., Koenig, D. F., Mair, G. A., North, A. C. T., Philipps, D. C. & Sarma, V. R. (1965). *Nature* (*London*), **206**, 757–761.
Bricogne, G., Vonrhein, C., Flensburg, C., Schiltz, M. & Paciorek, W. (2003). *Acta Cryst.* D**59**, 2023–2030.
Brinkmann, C., Weiss, M. S. & Weckert, E. (2006). *Acta Cryst.* D**62**, 349–355.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Faust, A., Panjikar, S., Mueller, U., Parthasarathy, V., Schmidt, A., Lamzin, V. S. & Weiss, M. S. (2008*a*). *J. Appl. Cryst.* **41**, 1161–1172.
Faust, A., Panjikar, S., Mueller, U., Parthasarathy, V., Schmidt, A., Lamzin, V. S. & Weiss, M. S. (2008*b*). *Acta Cryst.* A**64**, C80.
Hao, Q. (2004). *J. Appl. Cryst.* **37**, 498–499.
Heinemann, U., Buessow, K., Mueller, U. & Umbach, P. (2003). *Acc. Chem. Res.* **36**, 157–163.
Jaskólski, M. (2001). *J. Appl. Cryst.* **34**, 371–374.
Kabsch, W. (1993). *J. Appl. Cryst.* **26**, 795–800.
Kabsch, W. (2010*a*). *Acta Cryst.* D**66**, 125–132.
Kabsch, W. (2010*b*). *Acta Cryst.* D**66**, 133–144.
Leslie, A. G. W. (1992). *Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography*, No. 26.
Morris, R. J., Perrakis, A. & Lamzin, V. S. (2002). *Acta Cryst.* D**58**, 968–975.
Mueller-Dieckmann, C., Panjikar, S., Tucker, P. A. & Weiss, M. S. (2005). *Acta Cryst.* D**61**, 1263–1272.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Nanao, M. H. & Ravelli, R. B. (2006). *Structure*, **14**, 791–800.
Nanao, M. H., Sheldrick, G. M. & Ravelli, R. B. G. (2005). *Acta Cryst.* D**61**, 1227–1237.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Panjikar, S., Parthasarathy, V., Lamzin, V. S., Weiss, M. S. & Tucker, P. A. (2005). *Acta Cryst.* D**61**, 449–457.
Panjikar, S., Parthasarathy, V., Lamzin, V. S., Weiss, M. S. & Tucker, P. A. (2009). *Acta Cryst.* D**65**, 1089–1097.
Pannu, N. S., McCoy, A. J. & Read, R. J. (2003). *Acta Cryst.* D**59**, 1801–1808.
Perrakis, A., Morris, R. J. & Lamzin, V. S. (1999). *Nat. Struct. Biol.* **6**, 458–463.
Ravelli, R. B. G., Nanao, M. H., Lovering. A., White, S. & McSweeney, S. (2005). *J. Synchrotron Rad.* **12**, 276–284.
Schneider, T. R. & Sheldrick, G. M. (2002). *Acta Cryst.* D**58**, 1772–1779.
Schoenfeld, D. L., Ravelli, R. B. G., Mueller, U. & Skerra, A. (2008). *J. Mol. Biol.* **384**, 393–405.
Sheldrick, G. M. (2002). *Z. Kristallogr.* **217**, 644–650.
Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.
Sheldrick, G. M. (2010). *Acta Cryst.* D**66**, 479–485.
Sheldrick, G. M., Hauptman, H. A., Weeks, C. M., Miller, R. & Uson, I. (2001). *International Tables for Crystallography*, Vol. F, edited by M. G. Rossmann & E. Arnold, pp. 333–351. Dordrecht: Kluwer Academic Publishers.
Sun, P. D., Radaev, S. & Kattah, M. (2002). *Acta Cryst.* D**58**, 1092–1098.
Weiss, M. S., Palm, G. J. & Hilgenfeld, R. (2000). *Acta Cryst.* D**56**, 952–958.