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


Copyright © International Union of Crystallography

Author(s) of this paper may load this reprint on their own web site or institutional repository provided that this cover page is retained. Republication of this article or its storage in electronic databases other than as specified above is not permitted without prior permission in writing from the IUCr.

For further information see [http://journals.iucr.org/services/authorrights.html](http://journals.iucr.org/services/authorrights.html)

Many research topics in condensed matter research, materials science and the life sciences make use of crystallographic methods to study crystalline and non-crystalline matter with neutrons, X-rays and electrons. Articles published in the *Journal of Applied Crystallography* focus on these methods and their use in identifying structural and diffusion-controlled phase transformations, structure–property relationships, structural changes of defects, interfaces and surfaces, etc. Developments of instrumentation and crystallographic apparatus, theory and interpretation, numerical analysis and other related subjects are also covered. The journal is the primary place where crystallographic computer program information is published.

Crystallography Journals Online is available from [journals.iucr.org](http://journals.iucr.org)
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 Muellerb and Manfred S. Weissb*

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., 2008a). 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 lysozyme

Since a detailed introduction on lysozyme has been given in the first version of the tutorial (Faust et al., 2008a), 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 (|F_p|) with those of a protein crystal that has been derivatized by co-crystallization or soaking with a heavy-atom-containing compound (|F_H|). Under the assumption that both crystals are isomorphous, the differences (|F_H| − |F_p|) can be used to identify the positions of the heavy atoms. The resulting heavy-atom structure factors (F_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 F_p + F_H = F_{PH}. If the magnitudes of F_p and F_{PH} are known (these are the measured structure factor amplitudes |F_p| and |F_{PH}| of the native and the derivative data sets) and if F_H is known as vector (this means that the heavy-atom structure is known), the phase ϕ can be determined. In the case of single isomorphous replacement (SIR; Fig. 2a), however, the phase determination yields two values (phase ambiguity), while in the case of multiple isomorphous replacement (not shown) and SIRAS (Fig. 2b) an unambiguous phase determination can be achieved.

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., 2008a). 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-
enzyme crystals and thaumatin crystals were produced as described in the first version of the tutorial (Faust et al., 2008a). 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. (2008a). Protein solution (4 µl, 30 mg ml⁻¹ in water) was mixed with 4 µl of reservoir solution containing 50 mM sodium acetate pH 4.6 and 5% (w/v) sodium chloride and equilibrated against 1 ml of reservoir. Crystals belonging to space group P4₁2₁2 and exhibiting the usual unit-cell parameters of a = 78.6 Å, c = 36.8 Å appeared within a few days. For heavy-atom derivatization, a solution of KAuCl₄ 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 about 1.8 Å resolution, the diffraction properties of the derivatized crystals were slightly worse. Nevertheless, they still diffracted X-rays to at least 1.5 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 P4₁2₁2 and with unit-cell parameters a = 57.9 Å, c = 150.2 Å appeared after a few days. They diffracted X-rays to at least 1.5 Å resolution.

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 mM 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 Rayonix MX-225, 3 × 3 tiled CCD detector. On BL14.2 the detector is mounted on a MARdbt 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.teemphotronics.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 XSCALEx (Kabsch, 1993, 2010a,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., 2008a). Even though the new version of the tutorial contains a description of the data processing that has been completely re-written in
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 \( \frac{F}{C} \) 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 1

<table>
<thead>
<tr>
<th></th>
<th>Lysozyme (Experiment 6)</th>
<th>Thaumatin (Experiment 7)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Native</td>
<td>KAuCl₄ derivative</td>
</tr>
<tr>
<td>No. of crystals</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Beamline</td>
<td>BL14.2</td>
<td>BL14.2</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Crystal-detector distance (mm)</td>
<td>180</td>
<td>180</td>
</tr>
<tr>
<td>Rotation range per image (°)</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Total rotation range (°)</td>
<td>180</td>
<td>180</td>
</tr>
<tr>
<td>Exposure time per image (s)</td>
<td>2.50</td>
<td>5.0</td>
</tr>
<tr>
<td>Attenuation (mm Al)</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

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.
teaching and education

Table 2
Data-processing statistics for the two projects.
The numbers in parentheses define the outermost resolution shell used.

<table>
<thead>
<tr>
<th>Resolution range (Å)</th>
<th>Native</th>
<th>KAuCl₄ derivative</th>
<th>‘Before’</th>
<th>‘After’</th>
</tr>
</thead>
<tbody>
<tr>
<td>50.0–1.60 (1.70–1.60)</td>
<td>50.0–1.80</td>
<td>30.0–1.60 (1.70–1.60)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Space group</th>
<th>Unit-cell parameters (Å)</th>
<th>Mosaicity (%)</th>
<th>Total No. of reflections</th>
<th>Unique reflections</th>
</tr>
</thead>
<tbody>
<tr>
<td>P₄₁₂₂</td>
<td>a = 78.62, a = 78.73, a = 78.9, a = 78.9,</td>
<td>0.13</td>
<td>187,497</td>
<td>28,363</td>
</tr>
<tr>
<td></td>
<td>c = 36.81, c = 36.73, c = 150.15, c = 150.18</td>
<td>0.06</td>
<td>152,422</td>
<td>20,439</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Completeness (%)</th>
<th>Redundancy</th>
<th>I/σ(I)</th>
<th>R_meas (%)</th>
<th>Overall Wilson B factor (Å²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>97.4 (86.3)</td>
<td>6.6 (3.4)</td>
<td>27.4 (5.2)</td>
<td>4.8 (24.6)</td>
<td>20.3</td>
</tr>
<tr>
<td>99.5 (97.5)</td>
<td>7.5 (7.1)</td>
<td>15.6 (3.0)</td>
<td>9.9 (71.1)</td>
<td>28.3</td>
</tr>
<tr>
<td>98.3 (91.0)</td>
<td>3.2 (1.6)</td>
<td>12.5 (2.1)</td>
<td>8.4 (44.7)</td>
<td>17.1</td>
</tr>
<tr>
<td>99.4 (97.9)</td>
<td>3.3 (1.6)</td>
<td>12.0 (1.8)</td>
<td>8.9 (51.7)</td>
<td>17.8</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Au site</th>
<th>x</th>
<th>y</th>
<th>Z</th>
<th>Occupancy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>67.093</td>
<td>11.725</td>
<td>0.909</td>
<td>1.00</td>
</tr>
<tr>
<td>2</td>
<td>69.980</td>
<td>9.972</td>
<td>-4.114</td>
<td>0.69</td>
</tr>
<tr>
<td>3</td>
<td>67.297</td>
<td>11.323</td>
<td>-9.203</td>
<td>0.33</td>
</tr>
</tbody>
</table>

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 | |Fₐₐ₇ₙ₃| - |Fᵢₙ₃| | 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 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 KAuCl₄-derivative data using REFMACS (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). |Fₐₐ₇ₙ₃| - |Fᵢₙ₃| 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 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 KAuCl₄-derivative data using REFMACS (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.

Figure 4
Refined structure of the KAuCl₄ 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 4.0ε. 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 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 $R_{\text{free}} = 19.3\%$. 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). 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), 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.

#### 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 KAuCl₄, 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

![Figure 5](image)

**Figure 5**

Difference electron density ($F_{\text{after}} - F_{\text{before}}$, $\alpha_{\text{calc}}$) 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 $+5.0\sigma$ (green) and $-5.0\sigma$ (red). The positive peaks identify structural elements that appear after the damage, while the negative peaks identify sites that are destroyed by the damage.

<table>
<thead>
<tr>
<th>Site No.</th>
<th>X</th>
<th>Y</th>
<th>Z</th>
<th>Peak height ($\sigma$)</th>
<th>Closest atom</th>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>18.74</td>
<td>31.68</td>
<td>22.47</td>
<td>-39.49</td>
<td>Cys56-SG</td>
<td>S bridge with Cys68</td>
</tr>
<tr>
<td>2</td>
<td>9.03</td>
<td>3.12</td>
<td>13.61</td>
<td>-34.17</td>
<td>Cys68-SG</td>
<td>S bridge with Cys56</td>
</tr>
<tr>
<td>3</td>
<td>7.21</td>
<td>8.81</td>
<td>34.45</td>
<td>-33.01</td>
<td>Cys193-SG</td>
<td>S bridge with Cys121</td>
</tr>
<tr>
<td>4</td>
<td>5.66</td>
<td>7.71</td>
<td>33.66</td>
<td>-24.67</td>
<td>Cys121-SG</td>
<td>S bridge with Cys193</td>
</tr>
<tr>
<td>5</td>
<td>20.77</td>
<td>32.20</td>
<td>25.12</td>
<td>+24.65</td>
<td>Cys68-SG</td>
<td>New position</td>
</tr>
<tr>
<td>6</td>
<td>31.47</td>
<td>28.43</td>
<td>49.95</td>
<td></td>
<td>Cys126-SG</td>
<td>S bridge with Cys177</td>
</tr>
<tr>
<td>7</td>
<td>25.05</td>
<td>39.60</td>
<td>51.69</td>
<td>+20.69</td>
<td>Cys56-SG</td>
<td>New position</td>
</tr>
<tr>
<td>8</td>
<td>5.70</td>
<td>38.06</td>
<td>35.92</td>
<td>-19.86</td>
<td>Cys134-SG</td>
<td>S bridge with Cys145</td>
</tr>
<tr>
<td>9</td>
<td>5.95</td>
<td>36.03</td>
<td>36.26</td>
<td>-18.09</td>
<td>Cys145-SG</td>
<td>S bridge with Cys134</td>
</tr>
<tr>
<td>10</td>
<td>5.66</td>
<td>8.40</td>
<td>34.99</td>
<td>+17.28</td>
<td>Cys121-SG</td>
<td>New position</td>
</tr>
<tr>
<td>11</td>
<td>1.37</td>
<td>26.54</td>
<td>34.91</td>
<td>-16.15</td>
<td>Cys177-SG</td>
<td>S bridge with Cys126</td>
</tr>
<tr>
<td>12</td>
<td>8.77</td>
<td>8.22</td>
<td>34.18</td>
<td>+16.12</td>
<td>Cys193-SG</td>
<td>New position</td>
</tr>
<tr>
<td>13</td>
<td>23.99</td>
<td>-0.81</td>
<td>14.53</td>
<td>-13.31</td>
<td>Cys204-SG</td>
<td>S bridge with Cys9</td>
</tr>
<tr>
<td>14</td>
<td>32.46</td>
<td>-4.63</td>
<td>40.89</td>
<td>+11.18</td>
<td>Cys126-SG</td>
<td>New position</td>
</tr>
<tr>
<td>15</td>
<td>10.87</td>
<td>27.78</td>
<td>51.76</td>
<td>-9.71</td>
<td>Cys164-SG</td>
<td>S bridge with Cys159</td>
</tr>
<tr>
<td>16</td>
<td>16.45</td>
<td>33.82</td>
<td>47.14</td>
<td>-9.65</td>
<td>Cys158-SG</td>
<td>S bridge with Cys149</td>
</tr>
<tr>
<td>17</td>
<td>6.86</td>
<td>27.55</td>
<td>22.12</td>
<td>-9.55</td>
<td>Cys9-SG</td>
<td>S bridge with Cys204</td>
</tr>
<tr>
<td>18</td>
<td>16.23</td>
<td>32.96</td>
<td>45.22</td>
<td>-9.12</td>
<td>Cys149-SG</td>
<td>S bridge with Cys158</td>
</tr>
<tr>
<td>19</td>
<td>19.70</td>
<td>32.72</td>
<td>20.79</td>
<td>-8.19</td>
<td>Cys56-CB</td>
<td>S bridge with Cys66</td>
</tr>
<tr>
<td>20</td>
<td>23.83</td>
<td>27.56</td>
<td>27.39</td>
<td>-8.10</td>
<td>Cys71-SG</td>
<td>S bridge with Cys77</td>
</tr>
<tr>
<td>21</td>
<td>18.24</td>
<td>36.21</td>
<td>22.05</td>
<td>-7.17</td>
<td>Asp55-O</td>
<td>–</td>
</tr>
<tr>
<td>22</td>
<td>13.01</td>
<td>34.79</td>
<td>24.26</td>
<td>+6.88</td>
<td>Thr54-O</td>
<td>New position</td>
</tr>
<tr>
<td>23</td>
<td>16.25</td>
<td>34.06</td>
<td>23.77</td>
<td>+6.88</td>
<td>Thr54-O</td>
<td>New position</td>
</tr>
<tr>
<td>24</td>
<td>12.13</td>
<td>28.43</td>
<td>49.95</td>
<td>-6.84</td>
<td>Cys159-SG(B)</td>
<td>S bridge with Cys164</td>
</tr>
<tr>
<td>25</td>
<td>5.22</td>
<td>38.52</td>
<td>34.54</td>
<td>+6.80</td>
<td>fCys134-SG</td>
<td>New position</td>
</tr>
<tr>
<td>26</td>
<td>4.61</td>
<td>9.56</td>
<td>30.37</td>
<td>-6.44</td>
<td>Pro116-CG</td>
<td>–</td>
</tr>
<tr>
<td>27</td>
<td>13.48</td>
<td>34.24</td>
<td>25.18</td>
<td>-6.37</td>
<td>Thr54-GO1</td>
<td>–</td>
</tr>
<tr>
<td>28</td>
<td>8.58</td>
<td>7.01</td>
<td>35.45</td>
<td>-6.37</td>
<td>Cys193-CB</td>
<td>S bridge with Cys121</td>
</tr>
<tr>
<td>29</td>
<td>22.44</td>
<td>26.04</td>
<td>27.17</td>
<td>-6.28</td>
<td>Cys77-SG</td>
<td>S bridge with Cys71</td>
</tr>
<tr>
<td>30</td>
<td>6.05</td>
<td>10.00</td>
<td>31.32</td>
<td>+6.21</td>
<td>Pro116-CG</td>
<td>New position</td>
</tr>
<tr>
<td>31</td>
<td>2.17</td>
<td>25.24</td>
<td>34.38</td>
<td>+6.39</td>
<td>Cys177-SG</td>
<td>New position</td>
</tr>
<tr>
<td>32</td>
<td>2.87</td>
<td>27.82</td>
<td>22.61</td>
<td>+6.08</td>
<td>Cys204-SG</td>
<td>New position</td>
</tr>
<tr>
<td>33</td>
<td>9.05</td>
<td>22.72</td>
<td>31.27</td>
<td>-6.05</td>
<td>Met112-SD</td>
<td>–</td>
</tr>
</tbody>
</table>
may be the case that older solutions of KAuCl₄ 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., 2008a) 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
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 diffrac-
tion 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
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., 2008b).

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.,
Res. 28, 235–242.

Blake, C. C. F., Koenig, D. F., Mair, G. A., North, A. C. T., Philipp,
Bricegno, G., Vonrhein, C., Flensburg, C., Schiltz, M. & Paciorek, W.
349–355.
D50, 760–763.
Faust, A., Panjikar, S., Mueller, U., Parthasarathy, V., Schmidt, A.,
1172.
Faust, A., Panjikar, S., Mueller, U., Parthasarathy, V., Schmidt, A.,
968–975.
Mueller-Dieckmann, C., Panjikar, S., Tucker, P. A. & Weiss, M. S.
D53, 240–255.
D61, 1227–1237.
326.
Panjikar, S., Parthasarathy, V., Lamzin, V. S., Weiss, M. S. & Tucker,
Panjikar, S., Parthasarathy, V., Lamzin, V. S., Weiss, M. S. & Tucker,
1801–1808.
Biol. 6, 458–463.
Ravelli, R. B. G., Nanao, M. H., Lovering, A., White, S. & McSweeney,
1779.
Sheldrick, G. M., Hauptman, H. A., Wecks, C. M., Miller, R. & Uson,
Academic Publishers.
1098.
952–958.