Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Heavy-atom derivatization

Elspeth Garman* and James W. Murray

Laboratory of Molecular Biophysics, Department of Biochemistry, Oxford University, Rex Richards Building, South Parks Road, Oxford OX1 3QU, England

Correspondence e-mail: elspeth@biop.ox.ac.uk

Received 6 March 2003
Accepted 9 June 2003

Most of the standard methods of solving macromolecular structures involve producing a protein crystal that is derivatized by an anomalous scatterer or heavy atom (MIR, SIRAS, MAD, SAD *etc.*). The theoretical methodology which underpins the extraction of phase information from such derivatives is widely available in the literature. In addition, there are comprehensive sources of information on the chemistry of heavy-atom compounds and the ligands with which they are known to interact, as well as the Heavy Atom Databank accessible on the World Wide Web. This contribution therefore aims to provide some information on the less well documented practical problems of firstly deciding on an overall strategy for derivatization and secondly performing the physical manipulations involved in producing heavy-atom derivatives from native protein crystals and then cryocooling them. Ways to optimize the chances of isomorphous unit cells are suggested. Methods of determining whether or not the heavy atom is bound are outlined, including the powerful technique of PIXE (particle-induced X-ray emission).

1. Introduction

The incorporation of heavy atoms into protein crystals for the purposes of phasing was pioneered by Green *et al.* (1954). Although alternative methods of phasing have been developed, most still rely on the presence of either an endogenous heavy atom, selenomethionine residues in the protein or the derivatization of an existing crystal to include atoms of larger scattering power than the light elements from which proteins are built. Thus, the methods of multiple and single isomorphous replacement (MIR, SIR), multiple and single isomorphous replacement with anomalous scattering (MIRAS, SIRAS) and multi- and single-wavelength anomalous dispersion (MAD, SAD) all require a heavy atom to be present in enough sites in the molecule and at high enough occupancy to give a clear signal.

The essential foundations of techniques of heavy-atom derivatization and its theoretical basis can be found in Blundell & Johnson (1976) and Drenth (1999) and will not be repeated here. The basic theory has been summarized in this volume by Taylor (2003). Additional useful sources of information are the *International Tables for Crystallography Volume F* (Rossmann & Arnold, 2001) and a contemporary summary by Boggon & Shapiro (2000).

This paper will concentrate on the practical aspects of heavy-atom derivatization of protein crystals: which heavy atoms should be tried, how should they be incorporated into the crystals and how can success or failure be assessed? Some potential problems which might be encountered are also described, with possible remedies.

2. Which heavy atom?

2.1. Expected size of signal from heavy atoms

Before deciding which heavy atom to try first, it is useful to calculate the size of the expected signal to ensure that a heavy enough atom is used for the particular protein under study. Crick & Magdoff (1956) showed that for acentric reflections in a diffraction pattern, the average fractional expected intensity change, $\langle\Delta I\rangle/|I|$, resulting from the isomorphous addition of heavy atoms can be estimated as

$$\frac{\langle\Delta I\rangle}{|I|} = \left(\frac{2N_E}{N_P}\right)^{1/2}\frac{Z_h}{Z_{\rm eff}},$$

where N_E and N_P are the number of heavy and non-H protein atoms, respectively, and Z_h and $Z_{\rm eff}$ are the atomic numbers of the heavy atom and the average of protein atoms ($Z_{\rm eff} \simeq 6.7$), respectively.

Applying this formula to a 100 kDa protein with, for example, one fully occupied uranium (Z = 92) site, the expected MIR signal is a 16% average intensity change, whereas for a fully occupied copper (Z = 28) site this reduces to 5%. Thus, the higher the atomic number of the heavy atom, the larger the signal and the easier it will be to measure. For small signal changes, it can become very important to maximize the signal to noise in the diffraction images in order to optimize the data quality (Garman, 1999*a*). If the predicted signal is smaller than the $R_{\rm meas}$ (Diederichs & Karplus, 1997) of the highest resolution shell of the data, the derivative is unlikely to give reliable phasing information. Thus, the aim should be to use an atom heavy enough to give a signal larger than the experimental errors in the data.

For anomalous scattering and for MAD structure solution (using the change with incident wavelength of the dispersive part of the scattering), the signals are smaller than for MIR. Calculation of the anomalous signal expected for Bijvoet differences (Hendrickson & Ogata, 1997) and of the size of the dispersive differences between wavelengths requires knowledge of the imaginary (f''_A) and dispersive [$f'_A(\lambda_i)$] components of the scattering amplitude for the anomalous scatterer, as well as knowledge of how many atoms of it are present. A very useful World Wide Web site for experimenters wishing to check the expected signal size before carrying out an experiment is http://www.bmsc.washington.edu/scatter/AS_index.html. The site requires input of the mass and the number of anomalous scatterers as well as the protein size. It then calculates the approximate percentage contribution of these scatterers to the total diffraction intensity as a function of resolution.

The size of both the anomalous and dispersive signal is proportional to the square root of the proportion of atoms which are anomalous scatterers. It is worth noting that unlike the MIR signal, the fractional change in the total intensity owing to the anomalous signal [$f'_A(\lambda_i)$ and f''_A] increases with resolution.

As a general rule of thumb, for a successful MAD experiment on selenomethionine-containing proteins, at least one selenomethionine per 100 amino acids is required (Hendrickson & Ogata, 1997). However, under favourable conditions it may be possible to solve the structure with a proportionally lower selenuim concentration than this.

2.2. Classes of heavy atom (HA) available

There are broadly seven classes of heavy-atom derivative available for phasing:

(i) single metal ions which can be bound electrostatically to the protein,

(ii) endogenous metals which can be used directly for SAD/MAD or substituted by heavier metals with similar valency to obtain a larger signal, *e.g.* strontium for calcium,

(iii) selenium in selenomethionine,

(iv) metal compounds which require a chemical reaction to take place,

(v) multi-metal cluster complexes used for large proteins and multi-protein assemblies (Thygesen *et al.*, 1996),

(vi) the noble gases xenon (Scheonberg *et al.*, 1965) and krypton (Cohen *et al.*, 2001), and

(vii) halides (Dauter *et al.*, 2000) and triiodide (Evans & Bricogne, 2002), including brominated nucleic acids (*e.g.* Ennifar *et al.*, 2002).

Historically successful have been the so-called 'magic seven' (Boggon & Shapiro, 2000) of K_2PtCl_4 (platinum potassium chloride), $KAu(CN)_2$ (aurous potassium cyanide), K_2HgI_4 (mercuric potassium iodide), $UO_2(C_2H_3O_2)_2$ [uranium (VI) oxyacetate], $HgCl_2$ (mercuric chloride), $K_3UO_2F_5$ (potassium uranyl fluoride) and *para*-chloromercurybenzoic sulfate (PCMBS).

The success of the magic seven as derivatives may be just because they have been tried more often than other compounds. New more soluble and specific heavy-atom compounds currently available are now displacing them from popularity. In particular, trimethyllead acetate (Holden & Rayment, 1991) has been found to have wide applicability for lead derivatization and a buffered mercury compound, ethylmercury thiosalicylate (EMTS), known as thiomersal, has become the mercury compound of choice. It is a mild reagent, as the thiolsalicylate ligand chelates the mercury ion and thus competes for the mercury with the protein. Mercurial compounds tend to bind to cysteine sulfur (free sulfhydryls) or histidine nitrogen, whereas platinum compounds bind to cysteine, histidine and methionine residues.

The 'white line' of an absorption edge is a sharp peak in f'' at the absorption edge. Some elements give larger anomalous signals as they have intense 'white lines' at the absorption edge. The $L_{\rm III}$ edges of the lanthanides have particularly intense and sharp white lines. In contrast, the mercury $L_{\rm III}$ edge has a rather rounded spectral shape.

2.3. Strategy for a decision

The bewildering array of possibilities outlined above begs the question of how to decide which class/compound to try first, especially if the supply of crystals and/or time are limited. The first line of attack is to try to exploit what is already known about the protein: for example, if there is an unpaired

cysteine, the covalent attachment of an Hg atom would be an obvious start. Alternatively, there may be calcium or magnesium ions that could be exchanged for heavier metal ions, in particular lanthanides. If there is an endogenous metal, such as zinc or iron, MAD over an appropriate wavelength range would be worthwhile.

Inspection of the primary sequence of the protein can provide clues to possible strategy. The properties of the different amino acids relevant for HA derivatization, including the optimimum pH for binding and a discussion of pK_a values, have been comprehensively documented by Petsko (1985).

The biological function of the protein can also be relevant to the choice of heavy atom. For example, phosphate- or phospholipid-binding proteins may often be derivatized by tungstate.

The composition and pH of the mother liquor can adversely affect the metal compound used for derivatization and must be considered when deciding on a strategy. For instance, ammonium sulfate exists in equilibrium with ammonia. At high pH values there can be appreciable amounts of ammonia present, which can compete with the protein as a ligand for the heavy ion. Ammonium sulfate has a high ionic strength, which weakens ionic interactions, and so could lower the affinity of the protein for the heavy ion (Drenth, 1999). The heavy-atom compound should thus be chosen with care or the ammonium sulfate should be replaced with another salt (*e.g.* lithium, sodium or potassium sulfate) or PEG (polyethylene glycol). A second example of this problem is the incompatibility of mercury with DTT-containing mother liquors.

If there are no methionine or free cysteine residues in the protein, variant protein can be produced with cysteine (Sun *et al.*, 1987; Nagai *et al.*, 1990) or methionine (Skinner *et al.*, 1994; Leahy *et al.*, 1994) residues added to aid phasing by derivatization or by using selenomethionine.

The usual starting conditions for trying heavy-atom derivatization are 0.1, 1.0 and 10 m*M* concentrations of soaking solution for between 10 min and several days at the highest concentration that the crystal will tolerate.

2.4. Heavy Atom Databank

A useful source of information on heavy-atom use in protein crystallography can be found at http://www.bmm.icnet.uk/had/heavyatom.html. This data bank was compiled and made accessible to aid decision-making in heavy-atom selection and to allow crystallographers to disseminate their experience by adding to the data bank (Islam *et al.*, 1998).

The data bank enables the researcher to find out how different heavy atoms have been used in the past and for each heavy-atom compound it lists the buffer conditions, pH range, type of protein and the details of the interactions formed with the proteins for which it has been a useful derivative. Individual elements in the periodic table can be selected to provide details of all the useful compounds of that element, including their formulae, chemical structure, solution chemistry, the numbers of structures reported which have used that compound and references to previous successful use.

It is worth noting that this data bank will only develop as a resource if experimenters have a contribution conscience.

3. Crystal manipulation and treatment

3.1. Safety

Before any manipulation of heavy-atom compounds is undertaken, the safety aspects must be properly considered. The heavy-atom compounds used in protein crystallography are selected for their strong affinity for biological molecules. Experimenters are made of just such molecules, so the compounds are of real danger to users if they are careless or are uninformed of the risks.

In the UK it is a safety requirement that the chemicals are kept in a locked cabinet with restricted key access and that there should be a full and comprehensive inventory of the contents of the cabinet as well as a COSSH (Control of Substances Hazardous to Health) form for every compound. Worldwide, each synchrotron has particular local rules for heavy-atom usage and disposal, and for handling pressurized gases. These should be checked *before* you arrive on the site.

For weighing out the compounds, a dedicated balance is necessary, complete with spatulas and gloves. An Eppendorf tube can be placed on the balance, the balance zeroed and the required mass of heavy-atom compound placed in the tube. This can then be used to make stock solutions. Only small quantities should be used and the area should be cleaned up after heavy-atom preparation and use.

Proper disposal is essential and each laboratory is required to have an established system for this. Uranium compounds are disposed of in a separate waste stream, since they are radioactive (the half-life of ^{238}U is 4.5×10^9 y). For instance, 50 µl of 10 m*M* uranium acetate undergoes 90 disintegrations min^{-1}. Most synchrotrons have a requirement that they must be informed in advance if any uranium is being brought onto the site.

The majority of heavy-atom compounds are highly toxic and they should all be treated with great respect. Sensible practice should be established in the use of the compounds (*e.g.* absolutely NO mouth pipetting).

3.2. Heavy-atom solution preparation

Once a sealed vapour-diffusion crystallization experiment is opened, the equilibrium is disturbed. The protein crystals may no longer be stable and can sometimes dissolve or disintegrate. This situation may be exacerbated by the transfer of the crystal to a heavy-atom solution. Thus, it is important to find a solution in which the protein crystals will remain stable and from which other conditions may then be derived. Typically, a stabilizing solution contains the same components as the mother liquor from which the crystal was grown, but with a higher concentration of precipitant. For example, ^{4}F1^{5}F1 human fibronectin crystals were grown in nominal 7–8%(*w*/*v*) PEG 8K, 0.2 *M* ammonium sulfate but were stabilized into

24%(*w*/*v*) PEG 8K, 0.2 *M* ammonium sulfate. The cryosolution for these crystals is the stabilizing solution with 30%(*v*/*v*) glycerol, the glycerol replacing water in the solution (*i.e.* original solution is not diluted by the glycerol). Note that if the cryoprotection soak is longer than a few minutes, the heavy atom should be included in the cryosolution, unless back-soaking is the objective (see §3.6).

Heavy-atom compounds are usually best handled as ten times concentrated stock solutions in water. If 10 µl of 1.1 times concentrated stabilizing solution are added to 1 µl of the heavy-atom stock solution, this is usually enough for one soak. A final concentration in the soak in the range 1–50 m*M* of heavy atom is usual. Some heavy-atom compounds are quite insoluble and may be used as saturated solutions; others are only soluble in acidic or basic solutions. It is best if freshly prepared solutions are used for each experiment as the compounds tend to be unstable. Many heavy-atom compounds, especially platinum and iridium compounds, are also photosensitive and thus soaks should be carried out in the dark.

3.3. Crystal manipulation

Since protein crystals typically contain much disordered solvent and the intermolecular interactions which hold the protein molecules together are in general very weak, they can easily become dehydrated and disordered. How they are manipulated can seriously affect the final diffraction data quality. Handling procedures are thus worth some thought and forward planning in order to avoid unnecessary degradation and increased mosaicity. For instance, investment in a 'cold' microscope light source (*i.e.* a light source displaced from the microscope stage and directed onto the sample by means of an optic fibre) is highly recommended, as crystals can literally dissolve before the eyes if heated by the microscope lamp. This feature is especially important for crystals grown and harvested at 277 K.

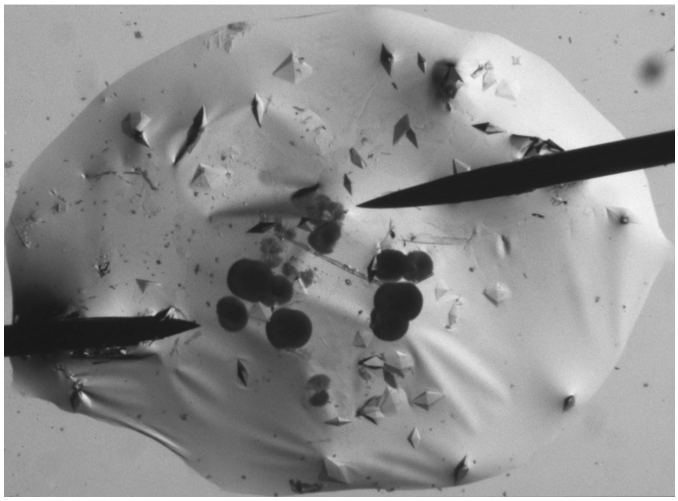

Figure 1
Acupuncture needles can conveniently be used to remove skin from a hanging- or sitting-drop crystallization experiment.

For heavy-atom soaking, the object is to extract one (if it is a new soaking condition) or a few crystals (if it is an established condition) from the growth drop (see §3.4) and transfer them into the heavy-atom-containing solution for soaking. The soaking solution and crystal can conveniently be incubated as a hanging or sitting drop in a labelled Linbro tray, with mother liquor in the well. In the situation where there are only a few crystals in the growth drop and the crystal supply is not too limited, the heavy-atom solution can be pipetted straight into the growth drop, making harvesting and transfer unnecessary.

Sometimes protein crystals grow in clusters or huddle together and require 'surgery' or separation. A particularly troublesome scenario is when a thick skin has grown over the surface of the crystallization drop, with the crystals firmly stuck on the underside of the skin. Experimenters have used a variety of tools for dislodging them in the past, including cats' whiskers, glass fibres, syringe needles and commercially available microtools. A convenient all-purpose aid to manipulation and surgery are acupuncture needles.[1] These are advantageous for several reasons: there is no loss of mother liquor (*cf.* syringe needle), they are slightly flexible, they have fine points and there are various sizes available suitable for a range of experimenter hand sizes. The needles are particularly useful for operations on protein-drop 'skins' where gentle surgery is possible to first chop round the edge of the drop to free the skin from the base, and then to chop around the outside of a single crystal using a needle in each hand: one to tether the crystal down and the other to free it from its neighbours (see Fig. 1). Since the skin does not diffract it can be left on the crystal; often, trying to remove it all will be disastrous. The presence of the skin will result in an increase in the diffuse X-ray background scatter.

3.4. Crystal harvesting

There are several ways of harvesting crystals from their growth drop for transfer to heavy-atom soaking solutions. The extraction should be performed as gently as possible, so that the crystals remain intact and undamaged. Before starting any crystal manipulation, it is very useful to establish a harvesting or stabilizing buffer.

Four possible transfer methods are described here; they are not exhaustive and there may also be others worth trying.

(i) A cryoloop mounted on a pin in a top hat can be used to catch the crystal in a film of liquid held across the loop by surface tension. The top hat can either be held magnetically on a straight or angled cryowand or grasped in self-opposing tweezers. Holding the top hat between finger and thumb is not recommended, since this method generally restricts the angles of approach, especially if a high-powered microscope is being used which has a small space between crystallization tray and lens. Different tools for holding the loop assembly are available which suit a variety of purpose. Experimenters should use whichever one is most convenient for their needs. The crystal in the cryoloop should be immersed in the receiving solution

[1] Obtainable from The AcuMedic Centre, 101–105 Camden High Street, London NW1 7JN; tel. 44(0)207 388 5783, fax 44(0)207 387 5766.

as quickly as possible in order to avoid dehydration by exposure to air, which will usually result in increased mosaicity.

(ii) A very effective tool giving exquisite suction control can be made from a Pasteur pipette that has been drawn out over a Bunsen flame and then broken into two parts. The part which has the original smaller diameter end is inserted into a piece of flexible hose attached to a small (1 ml) syringe body (see Fig. 2). The crystal can then be sucked up in a little of its liquid and transferred to its new location without drying out at all on the way. Note that a normal metal syringe needle tends to remove too much liquid by capillary action and the crystal is invisible under the microscope when inside it.

(iii) A Gilson pipette and plastic tip can be used to suck up the target crystal. This method is not ideal, as crystals often stick to the inside of the opaque plastic pipette and are hard to rescue since they cannot be seen under a microscope. This method can also result in too much liquid being removed from the crystallization drop, necessitating the addition of more liquid to avoid the crystals becoming 'beached'.

(iv) A glass Pasteur pipette with a compressible rubber teat can be used. This is harder to control than other options. It can also result in the removal of too much liquid.

3.5. Soaking in the heavy-atom solution

Once the crystal has been successfully manipulated out of its growth drop, there are several options for ways to soak the crystal in the heavy-atom solution.

The simplest of these is to immerse the crystal straight into a drop containing the heavy atom at the final concentration for a soak lasting anything from 10 min up to several days. Recent systematic work has provided evidence that it may be advantageous to soak for a short time (10 min) at near-saturation conditions. Crystals were found to be more isomorphous and better diffracting with short rather than longer (12 h) soaks (Sun *et al.*, 2002; Sun & Radaev, 2002) and the occupancy of heavy-atom sites did not increase with soak time for the longer soaked crystals.

A method which is gentler to the crystal is to transfer it from its growth drop into a drop of mother liquor or stabilizing solution and then to add the heavy-atom solution by serial addition of increasing concentrations (*e.g.* if a 10 m*M* soak is required, add 10 µl of 3 m*M* heavy-atom solution to 10 µl stabilizing solution and the crystal, mix with a pipette tip, remove 10 µl of solution, add 10 µl of 6 m*M* heavy-atom solution, mix, remove 10 µl of solution and then repeat for 10 m*M* twice).

The osmotic shock of immersion into the HA solution can be minimized by surrounding the crystal, on a cover slip in a drop of mother liquor, by up to eight 2 µl droplets of heavy-atom solution for 24–48 h (David & Burley, 1991). When there is no change in the size of drops, they are gently merged with the main drop one at a time.

Dialysis of the heavy-atom solution into the crystal is another possibility for crystals which are particularly sensitive.

Protein crystals can also be co-crystallized with the heavy-atom compound. However, this strategy might affect the crystal packing, unit cell and/or space group and the resulting derivative would then not be isomorphous with the native protein crystal.

3.6. Back-soaking

Back-soaking involves soaking the crystal in heavy-atom solution for a known time and then removing it back into its heavy-atom-free stabilizing buffer for anything from a few minutes to a period of days. The purpose of this procedure is to remove any heavy atoms which are non-specifically bound and also to reduce the likelihood of additional partially

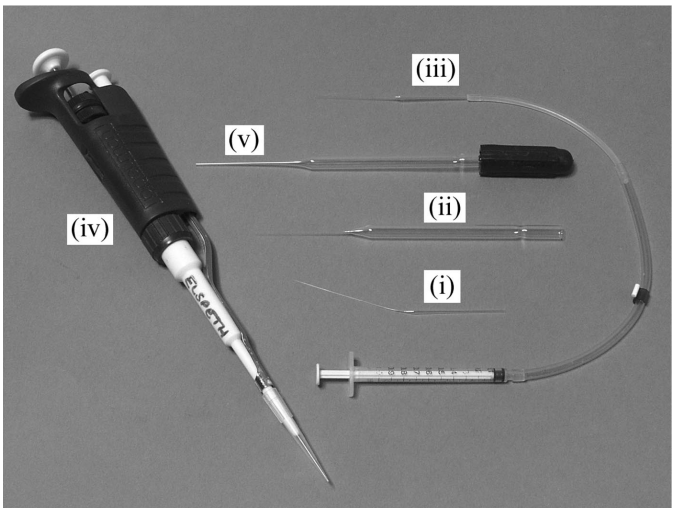

Figure 2
Crystal manipulation tools. (i) Small end of drawn Pasteur pipette. (ii) Larger end of drawn Pasteur pipette. (iii) Syringe attached using flexible tubing to smaller drawn end. (iv) Gilson pipette with tip. (v) Pasteur pipette with teat end.

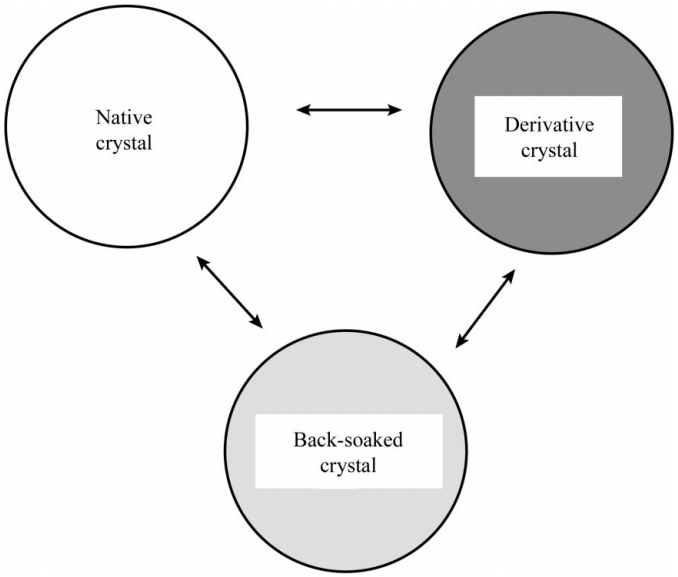

Figure 3
Diagram to show the three states of a crystal which are available if a derivative is back-soaked.

occupied sites caused by weak heavy-atom binding. Note, however, that low-occupancy derivatives can sometimes be used to solve structures if the data are collected at high resolution with high redundancy.

There are several advantages in carrying out back-soaking. It results in a third state of the crystal (the three being native, heavy-atom soaked and back-soaked; see Fig. 3) and thus gives the possibility of three different pair combinations for obtaining an isomorphous signal. Theoretically, back-soaking should also give less background and reduced radiation damage (see §5.5). It does involve an extra manipulative step, but the back-soaking can be carried out by exchanging the liquid around the crystal rather than physically moving the crystal to a new location, thus cutting down the possibility of damaging it.

4. Are there heavy-atom sites in the crystal?

An inherent problem with heavy-atom derivatization is that without collecting a diffraction data set and looking for the sites (see Dall'Antonia *et al.*, 2003), it is often not obvious whether or not the heavy atom has bound to the protein. This can be time-consuming and frustrating and it would be more efficient if there was a faster and unambiguous way to find out whether the soak had been successful or not. There are some indicators that can aid the search and several analytical approaches are being explored some of which are briefly described below. None is yet in widespread routine use.

4.1. Crystal changes colour during soaking

A colour change during soaking (see Fig. 4) does not necessarily imply that the heavy atom has bound at specific sites, since non-specific binding may be the cause of the colour change. Back-soaking (see §3.6) can be used to minimize the non-specific binding. Crystals which undergo a colour change can also become less ordered, *e.g.* the crystals shown in Fig. 4 normally diffract to 3 Å but after this soak only diffracted to 8 Å. A shorter soak time and/or a lower concentration of heavy-atom solution can then be tried.

4.2. Crystal undergoes physical changes

During soaking, the crystal sometimes cracks or, more spectacularly, explodes. This can indicate derivatization of the crystal, interference with the crystal contacts, denaturation of the protein or the induction of a major conformational change in the protein molecule. The heavy atom is clearly having an effect and another soak should be tried with the heavy-atom compound at a lower concentration or with a less reactive compound of the same heavy atom.

4.3. Mass spectrometry

The application of recent advances in mass spectrometry (MS) to proteins allow it to be used to determine the stoichiometry of heavy atoms both covalently and noncovalently bound to a liquid protein sample (Loo, 1997; Cohen *et al.*, 2000; Cohen & Chait, 2001). That the results have a correlation with the number of crystallographic binding sites has been shown for solutions of the human immunoglobulin receptor FcγRIII (Sun & Hammer, 2000) using electrospray ionization mass spectrometry (ESI–MS). For the four heavy-atom compounds compared, the number of crystallographic sites was greater (in three cases) or the same (in one case) as the number of heavy atoms detected per protein molecule by ESI–MS. Two other compounds were not detected in MS as they failed to react with the protein. This technique thus clearly has potential for the identification of possible heavy-atom derivatives. Before the measurement, samples studied by ESI–MS must be thoroughly cleaned of salts, buffers and any other additives which may be present. Not all proteins can survive this procedure, so an an alternative MS method, matrix-associated laser desorption ionization (MALDI), can be used instead, since the presence of the salts *etc.* do not disrupt it as much. Recently improved matrix materials are overcoming previous problems caused by the heavy-atom interfering with the desorption of the sample from the MALDI chip. This affected the spectrum quality, which was often rather poor (Boggon & Shapiro, 2000). For a study of the binding of mercurial compounds to proteins, ESI–MS has been shown to give more accurate results than MALDI (Cohen *et al.*, 2000). Despite this, MALDI can more easily be used for quick screening of heavy-atom binding, giving an indication of binding useful to the crystallographer rather than an accurate stoichiometric ratio.

Both MS methods have been used to check the incorporation of selenium into proteins grown in the presence of selenomethionine, the ~47 Da mass difference between sulfur and selenium being easily distinguishable.

4.4. Gel electrophoresis on liquid protein

As described by Boggon & Shapiro (2000), many compounds can be screened relatively quickly using native polyacrylamide gel electrophoresis (PAGE) on liquid protein

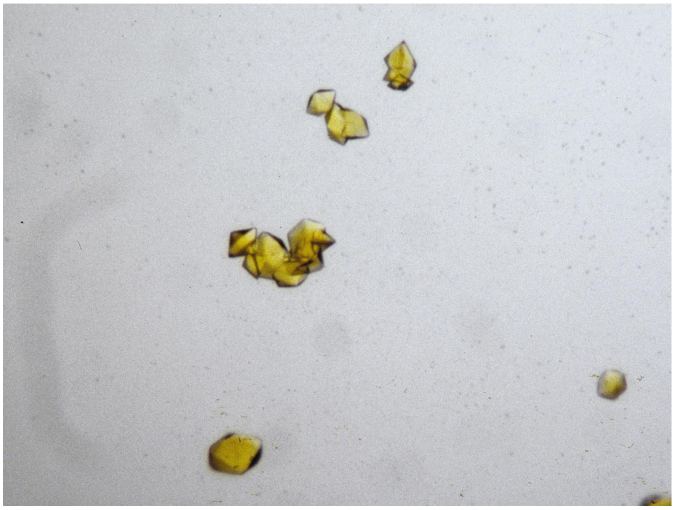

Figure 4
Crystals of ^{4}F1^{5}F1 human fibronectin module pair after soaking overnight in 10 m*M* platinum (II) terpyridine chloride dihydrate. The crystals have changed from colourless to orange.

samples with and without heavy-atom solution added. Binding of heavy atom will induce gel shifts owing to the change in mobility of the protein with the heavy atom in tow, or cause denaturation of protein so that it does not load properly onto the gel. Thus, unpromising derivatives can be abandoned before any crystals have been sacrificed. The concentration of the heavy-atom solution can be increased or decreased according to the results of the PAGE.

4.5. MicroPIXE (particle-induced X-ray emission)

The characteristic emission X-ray energy of an element can be used to identify it unambiguously. This X-ray emission can be induced by bombardment with X-rays (XRF), electrons (EPMA) or protons (PIXE). A scanning 1 µm diameter beam of 3 MeV protons has been used to analyse a variety of liquid and crystal protein samples (Garman, 1999*b*), identifying elements in them by detection of the emitted X-rays in a high-energy resolution lithium drifted silicon [Si(Li)] detector. The technique is sensitive to parts per million of an element in the sample dry weight. Thus, a single metal atom bound to a 100 kDa protein can be unambiguously identified. X-rays from elements lighter than neon (atomic mass number 20) do not penetrate the front detector window and thus are not detected.

MicroPIXE is an ideal technique for obtaining stoichiometric ratios of elements bound to proteins because the sulfur content of the protein molecule is known from the primary sequence (from the number of cysteine and methionine residues). Thus, quantitation is achieved by finding the ratios of the concentration of other elements to the concentration of sulfur. This internal calibration removes most of the systematic errors and gives stoichiometric ratios with ±10% accuracy.

For microPIXE on putative heavy-atom derivatized crystals, the soaked crystal (which can be down to 10 µm in size and not necessarily of diffraction quality) is thoroughly washed in MilliQ water and then mounted on a 2 µm thick Mylar film. The microbeam of protons is scanned across the crystal in *x* and *y* and elemental maps are built up by setting software windows round the X-ray energies of the elements of interest and sorting those events into separate two-dimensional maps for each element.

An example of such sulfur and heavy-atom maps collected from a crystal of EMR2 grown in the presence of 0.2 m*M* barium acetate is shown in Fig. 5(*a*). EMR2 is a three EGF domain-containing protein which by homology is expected to have at least two calcium-binding sites; calcium has been shown to be essential for its function (Lin *et al.*, 2000, 2001). It was hoped that the calcium would be replaced by barium during crystal growth. For the microPIXE analysis, X-ray spectra were collected at three points on the crystal and analysis of these with the *Dan*32 PIXE software (Grime & Dawson, 1995) gave a result of 2.44 ± 0.05 Ba atoms per protein molecule. The likely solvent content of the EMR2 crystals implies two molecules per asymmetric unit which, in conjunction with the microPIXE results, indicates the expected two sites and a third partially occupied site per molecule or a third site between monomers. Following an unoptimized SAD experiment on the crystals, the Patterson peak search could be informed by this knowledge of the number of sites. The structure solution of EMR2 is still in progress. Note that the microPIXE technique gives no information on the chemical environment of the metal atom, nor can it distinguish between one fully occupied site and two half-occupied metal sites.

The sulfur and gold maps in Fig. 5(*b*) resulted from a similar experiment carried out on DAF1234 crystals. DAF1234 protein consists of the four extracellular SCR (short concensus repeat) domains of the complement regulator CD55. The crystals, grown at pH 5.6, had been soaked at 277 K in 10 m*M* $NaAuCl_4$ for 24 h. Earlier attempts to bind gold to crystals

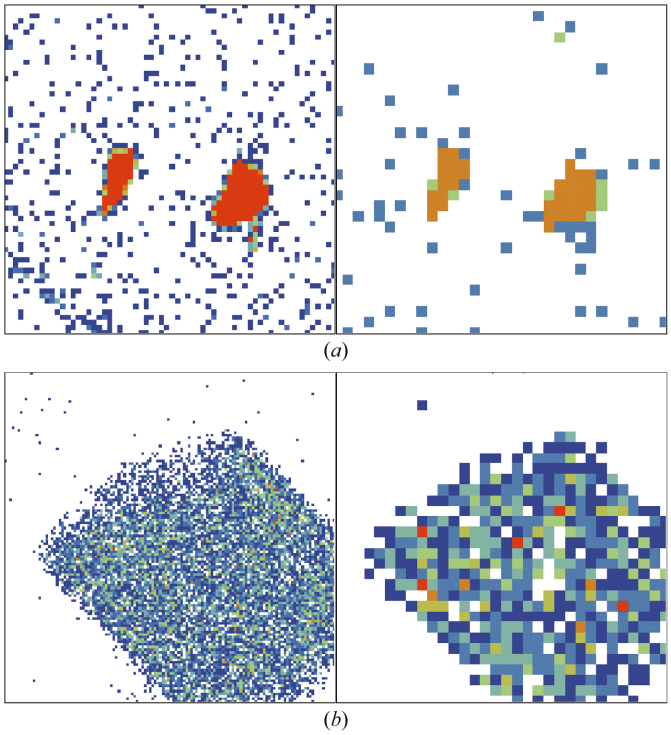

Figure 5
MicroPIXE elemental maps showing (*a*) sulfur (left) and barium (right) from a 0.5 × 0.5 mm scan over an EMR2 crystal and (*b*) sulfur (left) and gold (right) from a 0.25 × 0.25 mm scan over a DAF1234 crystal.

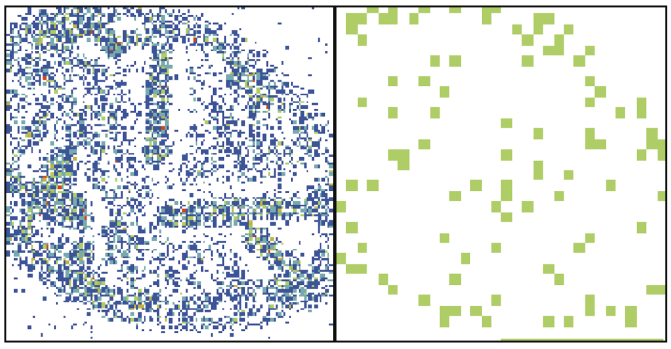

Figure 6
MicroPIXE sulfur and selenium elemental maps (1 × 1 mm scan) from a 0.3 µl drop of RsbW kinase evaporated protein solution.

grown at pH 4.5 had failed, but previous work on DAF34 had shown that at pH 6 gold binds to the protein and a gold derivative had contributed to the DAF34 structure solution (Williams *et al.*, 2002). The microPIXE results indicated that between 11 and 12 Au atoms were bound per DAF1234 protein molecule and therefore, although it is much harder to grow crystals at the higher pH of 5.6, efforts are now being concentrated in this area (Lea, 2003).

During the structure solution of DAF34, a crystal was analysed by microPIXE following a failed MAD experiment. The crystal had been soaked for 3 d in 2 m*M* HgI_2 and no mercury edge was detectable using the fluorescence detector at ID14-4 at the ESRF, Grenoble even though the crystals had changed colour. MicroPIXE on a washed crystal showed that iodine, not mercury, was bound. The structure of this protein was subsequently solved by Williams *et al.* (2002) using gold and platinum derivatives.

MicroPIXE can also be used to find the degree of incorporation of selenium into proteins to check that MAD structure determination is feasible, in particular for mammalian expression systems where the degree of selenomethionine take-up is highly variable. Liquid protein samples can conveniently be used for these measurements; only 0.3 µl of protein solution is required, at a concentration which varies with the sulfur content and molecular weight of the protein, although a few mg ml^{-1} is typical (Yates *et al.*, 2003). The degree of selenium incorporation is determined using microPIXE in the following way from the measured selenium (D_{Se}) and sulfur areal densities (D_S). Let the total number of Se atoms per protein molecule = X and the total number of S atoms left per protein molecule = Y. The original number of S atoms (cysteine + methionine) in the sequence was $Z = X + Y$.

$$\frac{N_S}{N_{Se}} = \frac{D_S}{D_{Se}} \times \frac{M_{Se}}{M_S} = \frac{Y}{X} = A = \frac{Z - X}{X}.$$

Since Z is known and A is measured, X can be determined,

$$X = \frac{Z}{1 + A}.$$

An example of such a measurement on a 0.3 µl liquid drop of 1 mg ml^{-1} RsbW kinase in sodium chloride buffer is shown in Fig. 6 (Lewis, 2003). The results showed that selenium had been successfully incorporated into the protein and crystallization conditions which will give diffraction-quality crystals are currently being sought.

The microPIXE technique is available as a service at the University of Surrey Centre for Research in Ion Beam Applications, Guildford, UK. Inquiries should be directed to Dr Geoff Grime by e-mail at g.grime@surrey.ac.uk.

4.6. Beamline fluorescence detector

Tunable beamlines for MAD and SAD experiments are equipped with fluorescence detectors which are used to locate the absorption edge so that the incident X-ray wavelength can be correctly set to optimize the anomalous signal. The fluorescence detector shows that the heavy atom is present in the crystal, but not necessarily bound to defined sites in the protein. However, if the protein crystal has been back-soaked in the mother liquor and the fluorescence signal is still present, this is a good sign that the derivatization has been successful.

4.7. X-ray data statistics

The time-honoured method of detecting the presence of ordered heavy atoms in the crystal is through the effects they have on the reflection intensities.

These effects can be detected by inspection of various statistics calculated from the data. For isomorphous signals, the R_{merge} between the putative heavy-atom data set and an isomorphous native data set is a strong indicator of their differences.

For finding an anomalous signal, a rough indicator is the difference between R_{meas} with and without Friedel mates merged (R_{meas} and R_{meas0} in *SCALA*; Evans, 1993). However, an appropriate difference Patterson map should always be calculated to check for heavy-atom sites. Normal probablility analysis can also be used (Howell & Smith, 1992) to identify heavy-atom derivatives likely to give phasing information, although good estimates of standard errors are required.

As already mentioned, finding heavy-atom derivatives can require the collection and analysis of many data sets before sufficient phasing power is achieved. For instance, in the solution of *Vibrio cholerae* neuraminidase six mediocre heavy-atom derivatives were used to phase the data before a structure solution could be found, following data collection on over 50 different heavy-atom soaks (Crennell *et al.*, 1994).

5. Potential problems

5.1. Crystals disintegrate in HA solutions

After immersion in heavy-atom solution, crystals monitored under the microscope may visibly degrade. The time of the soak may be shortened, the concentration reduced and/or the temperature at which the soak is carried out can also be reduced, since most heavy-atom binding will be slower at 277 K than at room temperature. Unless there is a large supply of crystals, it is unwise to add more than one crystal at a time to a new heavy-atom test solution, since if it makes the crystals disintegrate they will be wasted.

A possible strategy for sensitive crystals is to cross-link them with glutaraldehyde (Quiocho & Richards, 1964; Lusty, 1999) prior to heavy-atom soaking. Using a very low concentration (0.001%) of glutaraldehyde for a very short time (seconds to minutes) is advisable in order to prevent the crystal turning into a tiny rubber ball.

5.2. Non-isomorphism

One of the biggest problems of heavy-atom derivatization is that incorporation of a heavy atom into the lattice often induces a change in the unit cell away from the native crystal values, *i.e.* the derivatized crystal is non-isomorphous to the native crystals. The heavy atom may perturb the arrangement of protein molecules in the crystal or distort the protein molecule, causing a change in unit-cell lengths. Note, however,

Table 1
The rotation range in degrees required to collect a complete dataset with optimal crystal alignment and data collection strategy compared with that in the worst possible starting φ angle.

In addition the crystal may need to be moved around a κ axis to collect the blind region. More data may be required if the detector has a 2θ offset.

Crystal system	Laue class	Optimal native data	Worst-case native data	Optimal anomalous data	Worst-case anomalous data
Triclinic	$\bar{1}$	180	180 + 2θ	180 + 2θ	180
Monoclinic	2/*m*	90	180	180	180
Orthorhombic	*mmm*	90	180	90	180
Tetragonal	4/*m*	90	180	90	180
	4/*mmm*	45	180	45	180
Trigonal	$\bar{3}$	60	180	60 + 2θ	180
	$\bar{3}m$	30	180	30 + 2θ	180
Hexagonal	6/*m*	60	180	60	180
	6/*mmm*	30	180	30	180
Cubic	$m\bar{3}$	45	90	45	90
	$m\bar{3}m$	30	90	30	90

that it is also possible for the protein to move within the original unit cell (resulting in a different sampling of the molecular transform). The same unit cell is thus a necessary but not sufficient condition for isomorphism.

Crick & Magdoff (1956) calculated that a 0.5 Å change in all three unit-cell edges of a 100 Å cubed unit cell would change the diffraction intensities by an average of 15% in a 3 Å resolution sphere. The predicted intensity changes induced by non-isomorphism increase at higher resolution. When faced with a non-isomorphous derivative, it is the absolute change in the cell which should be considered compared with the working resolution, rather than the relative change, *i.e.* a change of 1.0% in a 100 Å unit cell edge has a similar effect to that of a 0.5% change in a 200 Å unit cell edge, if compared at similar resolutions. As a general rule of thumb, a change in cell dimensions of $d_{\rm min}/4$ is tolerable, where $d_{\rm min}$ is the resolution limit (Drenth, 1999). For instance, for 2.5 Å data, a 0.6 Å change in the unit cell might be acceptable, whereas at 3.5 Å this could rise to 0.8 Å.

It can be seen that the changes in intensity arising from non-isomorphism can easily swamp the isomorphous signal from the heavy atom (see §2.1) and thus the success of the experiment can depend on finding an isomorphous derivative. To optimize the chances of this, it is vital to use exactly the same protocol for the preparation, handling and cryotreatment (cryosolution concentrations, cryosolution soak time, cryogen, flash-cooling technique) of all crystals used for data collection. It is particularly important to keep constant the length of time the crystal spends in its cryosolution, since cryoprotecting agents are often also dehydrating agents and thus induce unit-cell shrinkage. Another opportunity for inadvertantly inducing cell shrinkage and thus non-isomorphism in crystals occurs during transfer of the crystal from its cryosolution into liquid or gaseous cryogen. The crystal can suffer dehydration and thus cell shrinkage if this action is slow. Crystals also commonly shrink on being flash-cooled owing to the difference in volume between water and vitreous ice. Again, it is important in obtaining a reproducible cell to subject all crystals to the same regime.

Note also that if the original crystallization conditions and stabilizing solution contain a dehydrating agent such as PEG, the concentration of this reagent should be kept constant in different batches. The length of time the crystal spends in the stabilizing solution may also affect the unit-cell values.

Experience shows that for heavy-atom derivative searches, it is best to have the same person perform all the crystal manipulations, since then idiosyncrasies of technique are at least reproduced each time.

A possible solution to the non-isomorphism problem is to use a single crystal for phasing using SAD. Alternatively, the phases from severely non-isomorphous derivatives can be combined in real space using cross-crystal averaging.

Another successful strategy has been to use a single crystal which has been broken in two. Native and back-soaked heavy-atom data are collected from the two parts at room temperature in order to minimize any non-isomorphism which may have been induced by cryotechniques.

A question often asked is whether the problems of non-isomorphism are exacerbated by cryotechniques. Owing to the time involved and the tedium of repeating soaking and diffraction experiments at room temperature once a structure has been solved, there is currently no evidence other than anecdotal to answer this question.

5.3. Poor diffraction and/or incomplete anomalous data

Derivative crystals frequently diffract less well than native crystals, having lower order and larger mosaicity (although some derivative crystals diffract better than native crystals). However, the lower resolution diffraction limit is generally not a limiting problem, since initial phases to 3.0–3.5 Å are usually sufficient to start bootstrapping to a structure solution.

Far more important is the collection of complete and multiplicitous anomalous data that is isomorphous to the native. There are no good excuses for collecting data that are incomplete. There exist many programs which, following autoindexing of the first image, will predict the optimum φ range to collect and the φ value at which to start data collection so that the completeness is maximized using the minimum beamtime (Leslie, 1993; Kabsch, 1993; Noble, 1996; Otwinowski & Minor, 1997; Ravelli *et al.*, 1997).

Table 1 shows the contrast between the φ ranges for optimized data collection (Dauter, 1999) and that required in the worst-case scenario (crystal aligned with the lowest symmetry axis parallel to the beam and starting the φ rotation from the worst possible place).

5.4. Oxidation of selenium

Selenomethionines in SeMet proteins are usually kept reduced using DTT (dithiothreitol) or β-mercaptoethanol. Oxidation of selenomethionine causes the selenium *K* edge to move from 12.65 keV (0.9795 Å) to a higher energy and the 'white line' to be enhanced owing to changes in the chemical environment of the selenium on oxidation. If mixed species

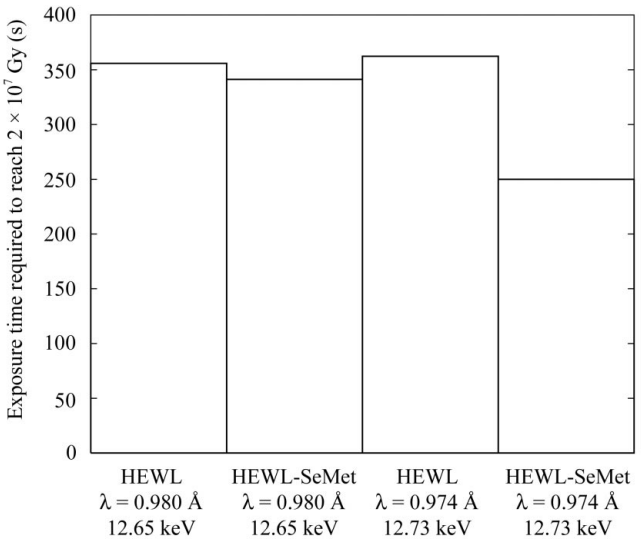

Figure 7
Histogram to show the calculated time taken for crystals of hen egg-white lysozyme, with and without two sulfur atoms replaced by selenium, to reach the 'Henderson dose limit' of 2×10^7 Gy at two different incident wavelengths, using typical conditions on beamline ID14-EH4 at the ESRF, Grenoble. Calculations were performed with *RADDOSE* (Murray *et al.*, 2003).

are present, the signal is degraded and thus partial oxidation should be avoided. However, it has been demonstrated that peroxide may be used to ensure that a single oxidized species is present (Sharff *et al.*, 2000).

5.5. Radiation damage

The large photon fluxes available from second- and third-generation synchrotrons and the push to use smaller crystals of ever larger protein complexes have contributed to bringing radiation damage to the forefront as an increasingly limiting problem in protein crystallography. Despite the widespread use of cryotechniques (Garman & Schneider, 1997; Rodgers, 1997), which significantly reduce the mobility of free radicals formed by interactions of the X-ray beam with the crystal and thus prolong the lifetime of crystals in the beam, radiation damage is commonly reported as a reason for the failure of MAD experiments.

As data collection proceeds, the resolution limit degrades, the $R_{\rm merge}$ becomes worse (Murray & Garman, 2002) and structural damage occurs at specific sites. Even worse, the unit cell expands in size (Burmeister, 2000; Ravelli & McSweeney, 2000) and so the problem of non-isomorphism is exacerbated. In the MAD technique, the signal is even smaller than for MIR, so non-isomorphism induced by damage during the three- or four-wavelength experiment can completely mask the MAD signal and thus cause the structure solution to fail.

The heavy atoms in a derivative have a higher primary photoelectric cross-section than all the light atoms in the protein molecule combined. Therefore, back-soaking is strongly recommended where necessary to remove non-specifically bound heavy atoms, which contribute to absorption and diffuse background scatter but not to the signal. A careful data-collection strategy is advised so that the necessary data are gathered with the lowest possible absorbed dose in the crystal. For example, a brominated RNA–protein complex will damage faster than a native crystal and will be debrominated during the experiment (Ennifar *et al.*, 2002). Specific radiation damage such as this has even been used for phasing, using the isomorphous differences caused by the specific sites of radiation damage (Ravelli *et al.*, 2003).

The effect on the time to reach the so-called 'Henderson limit' (the absorbed radiation dose which destroys approximately half the original diffraction intensity) of substituting two Se atoms ($Z = 34$) for two S atoms ($Z = 16$) in hen egg-white lysozyme (HEWL) crystals is illustrated in Fig. 7, calculated using *RADDOSE* (Murray *et al.*, 2003) for incident wavelengths below and above the selenium absorption edge. It can be seen that below the edge (the selenium K edge is at an energy of 12.66 keV, 0.9795 Å) the effect is not significant, but that above the edge the time available for data collection decreases by 30% owing to the presence of the Se atoms. For heavy atoms with higher Z values present in the crystal, the effect on the predicted time to reach the Henderson limit is even more marked.

6. Conclusions

The widespread use of the MAD/SAD methods of phasing, which can be carried out on either a selenomethionine-containing protein crystal, or on a crystal with an endogenous anolmalous scatterer, has significantly reduced the number of heavy-atom soaking experiments performed by macromolecular crystallographers. The increasing number of protein structures in the Protein Data Bank also allows molecular replacement to be more frequently used for phasing than in the past.

Despite these developments, however, there will continue to be a place for the use of heavy-atom derivatives in MIR and MAD/SAD phasing for some time to come. The soaking of these compounds into protein crystals is now a little more systematic than in the past, although it still involves a large element of trial and error and requires some luck and persistence. As ever, careful attention to detail in crystal treatment and in the application of the techniques can save time and effort in the long run.

We thank Airlie McCoy, Jim McDonnell, Martin Noble, Robin Owen, Gérard Bricogne and the CCP4 workshop delegates for helpful discussions. We are grateful to Geoff Grime and David Yates for their help with the microPIXE experiments, and to Susan Lea and Rick Lewis for permission to report results on their samples. We particularly thank Zbyszek Dauter for discussion regarding Table 1. JWM is supported by an MRC studentship.

References

Blundell, T. L. & Johnson, L. N. (1976). *Protein Crystallography*. New York: Academic Press.
Boggon, T. J. & Shapiro, L. (2000). *Structure*, **8**, R143–R149.

Burmeister, W. P. (2000). *Acta Cryst.* D**56**, 328–341.
Cohen, A., Ellis, P., Kresge, N. & Soltis, S. M. (2001). *Acta Cryst.* D**57**, 233–238.
Cohen, S. L. & Chait, B. T. (2001). *Annu. Rev. Biophys. Biomol. Struct.* **30**, 67–85.
Cohen, S. L., Padovan, J. C. & Chait, B. T. (2000). *Anal. Chem.* **72**, 574–579.
Crennell, S., Garman, E., Laver, G., Vimr, E. & Taylor, G. (1994). *Structure*, **2**, 535–544.
Crick, F. H. C. & Magdoff, B. S. (1956). *Acta Cryst.* **9**, 901–908.
Dauter, Z. (1999). *Acta Cryst.* D**55**, 1703–1717.
Dauter, Z., Dauter, M. & Rajashankar, K. R. (2000). *Acta Cryst.* D**56**, 232–237.
David, P. R. & Burley, S. K. (1991). *J. Appl. Cryst.* **24**, 1073–1074.
Diederichs, K. & Karplus, A. P. (1997). *Nature Struct. Biol.* **4**, 269–275.
Drenth, J. (1999). *Principles of Protein X-ray Crystallography*, 2nd ed. Berlin: Springer–Verlag.
Ennifar, E., Carpentier, P., Ferrer, J.-L., Walter, P. & Dumas, P. (2002). *Acta Cryst.* D**58**, 1262–1268.
Evans, P. (1993). *Proceedings of the CCP4 Study Weekend*, pp. 114–122.
Evans, G. & Bricogne, G. (2002). *Acta Cryst.* D**58**, 976–991.
Garman, E. (1999*a*). *Acta Cryst.* D**55**, 1641–1653.
Garman, E. (1999*b*). *Structure*, **7**, R291–R299.
Garman, E. F. & Schneider, T. R. (1997). *J. Appl. Cryst.* **30**, 211–237.
Green, D. W., Ingram, V. M. & Perutz, M. F. (1954). *Proc. R. Soc. London Ser. A*, **225**, 287–307.
Grime, G. W. & Dawson, M. (1995). *Nucl. Instrum. Methods Phys. Res. B*, **104**, 107–113.
Hendrickson, W. A. & Ogata, C. M. (1997). *Methods Enzymol.* **276**, 494–523.
Holden, H. M. & Rayment, I. (1991). *Arch. Biochem. Biophys.* **291**, 187–194.
Howell, P. L. & Smith, G. D. (1992). *J. Appl. Cryst.* **25**, 81–86.
Islam, S. A., Carvin, D., Sternberg, M. J. & Blundell, T. L. (1998). *Acta Cryst.* D**54**, 1199–1206.
Kabsch, W. (1993). *J. Appl. Cryst.* **26**, 795–800.
Lea, S. (2003). Personal communication.
Leahy, D. J., Erickson, H., Aukhil, I., Joshi, P. & Hendrickson, W. (1994). *Proteins*, **19**, 48–54.
Leslie, A. G. W. (1993). *Jnt CCP4/ESF–EAMCB Newsl. Protein Crystallogr.* **26**.
Lewis, R. (2003). Personal communication.
Lin, H.-H., Stacey, M., Hamann, J., Gordon, S. & McKnight, A. J. (2000). *Genomics*, **67**, 188–200.
Lin, H.-H., Stacey, M., Saxby, C., Knott, V., Chaudhry, Y., Evans, D., Gordon, S., McKnight, A. J., Handford, P. & Lea, S. (2001). *J. Biol. Chem.* **276**, 24160–24169.
Loo, J. A. (1997). *Mass Spectrom. Rev.* **16**, 1–23.
Lusty, C. J. (1999). *J. Appl. Cryst.* **32**, 106–112.
Murray, J. & Garman, E. (2002). *J. Synchrotron Rad.* **9**, 347–354.
Murray, J., Garman, E. & Ravelli, R. (2003). *RADDOSE: A Program for the Calculation of Protein Crystal Absorption Cross-sections.* Unpublished work.
Nagai, K., Oubridge, C., Jessen, T. H., Li, J. & Evans, P. (1990). *Nature (London)*, **348**, 515–520.
Noble, M. (1996). *PREDICT Package.* http://biop.ox.ac.uk/www/distrib/predict.html.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Petsko, G. A. (1985). *Methods Enzymol.* **114,** 147–156.
Quiocho, F. A. & Richards, F. M. (1964). *Proc. Natl Acad. Sci. USA*, **52**, 833–839.
Ravelli, R. B. G., Sweet, R. M., Skinner, J. M., Duisenberg, A. J. M., & Kroon, J., (1997). *J. Appl. Cryst.* **30**, 551–554.
Ravelli, R. B. & McSweeney, S. M. (2000). *Structure*, **8**, 315–328.
Ravelli, R. B. G., Schroder Leiros, H. K., Pan, B., Caffrey, M. & McSweeney, S. (2003). *Structure*, **11**, 217–224.
Rodgers, D. W. (1997). *Methods Enzymol.* **276**, 183–203.
Rossmann, M. G. & Arnold, E. (2001). Editors. *International Tables For Crystallography*, Vol. *F.* Dordrecht: Kluwer Academic Publishers.
Scheonberg, B. P., Watson, H. C. & Kendrew, J. C. (1965). *Nature (London)*, **207**, 28–30.
Dall'Antonia, F., Baker, P. J. & Schneider, T. (2003). *Acta Cryst.* D**59**, 1987–1994.
Sharff, A. J., Koronakis, E., Luisi, B. & Koronakis, V. (2000). *Acta Cryst.* D**56**, 785–788.
Skinner, M. M., Zhang, H., Leschnitzer, D. H., Guan, Y., Bellamy, H., Sweet, R. M., Gray, C. W., Konings, R. N., Wang, A. H.-J. & Terwilliger, T. C. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 2071–2075.
Sun, D., Alber, T., Bell, J., Waever, L. & Matthews, B. (1987). *Protein Eng.* **1**, 115–123.
Sun, P. D. & Hammer, C. H. (2000). *Acta Cryst.* D**56**, 161–168.
Sun, P. D. & Radaev, S. (2002). *Acta Cryst.* D**58**, 1099–1103.
Sun, P. D., Radaev, S. & Kattah, M. (2002). *Acta Cryst.* D**58**, 1092–1098.
Taylor, G. L. (2003). *Acta Cryst.* D**59**, 1881–1890.
Thygesen, J., Weinstein, S., Franceschi, F. & Yonath, A. (1996). *Structure*, **4**, 513–518.
Williams, P., Chaudhry, Y., Goodfellow, I. G., Billington, J., Spiller, O. B., Evans, D. J. & Lea, S. (2002). *J. Biol. Chem.* **278**, 10691–10696.
Yates, D., Grime, G. & Garman, E. (2003). *Identification of Elements in Proteins by MicroPIXE.* http://biop.ox.ac.uk/www/lj2001/garman/garman_03.html.