Acta Cryst. (1997). D53, 78–92

High-Pressure Krypton Gas and Statistical Heavy-Atom Refinement: a Successful Combination of Tools for Macromolecular Structure Determination

MARC SCHILTZ,[a] WILLIAM SHEPARD,[a] ROGER FOURME,[a] THIERRY PRANGÉ,[a,b] ERIC DE LA FORTELLE[c] AND GÉRARD BRICOGNE[a,c]

[a]*LURE, Bâtiment 209d, Université Paris Sud, 91405 Orsay CEDEX, France,* [b]*Chimie Structurale et Spectroscopie Biomoléculaire (URA 1430 CNRS) UFR-Biomédicale, 74, rue M. Cachin, 93012 Bobigny CEDEX, France, and* [c]*MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, England. E-mail: schiltz@lure.u-psud.fr*

(*Received 16 May 1996; accepted 17 July 1996*)

Abstract

The noble gas krypton is shown to bind to crystallized proteins in a similar way to xenon [Schiltz, Prangé & Fourme (1994). *J. Appl. Cryst.* **27**, 950–960]. Preliminary tests show that the major krypton binding sites are essentially identical to those of xenon. Noticeable substitution is achieved only at substantially higher pressures (above 50×10^5 Pa). As is the case for xenon, the protein complexes with krypton are highly isomorphous with the native structure so that these complexes can be used for phase determination in protein crystallography. Krypton is not as heavy as xenon, but its *K*-absorption edge is situated at a wavelength (0.86 Å) that is readily accessible on synchrotron radiation sources. As a test case, X-ray diffraction data at the high-energy side of the *K* edge were collected on a crystal of porcine pancreatic elastase (molecular weight of 25.9 kDa) put under a krypton gas pressure of 56×10^5 Pa. The occupancy of the single Kr atom is approximately 0.5, giving isomorphous and anomalous scattering strengths of 15.2 and 1.9 e, respectively. This derivative could be used successfully for phase determination with the SIRAS method (single isomorphous replacement with anomalous scattering). After phase improvement by solvent flattening, the resulting electron-density map is of exceptionally high quality, and has a correlation coefficient of 0.85 with a map calculated from the refined native structure. Careful data collection and processing, as well as the correct statistical treatment of isomorphous and anomalous signals have proven to be crucial in the determination of this electron-density map. Heavy-atom refinement and phasing were carried out with the program *SHARP*, which is a fully fledged implementation of the maximum-likelihood theory for heavy-atom refinement [Bricogne (1991). *Crystallographic Computing* 5, edited by D. Moras, A. D. Podjarny & J. C. Thierry, pp. 257–297. Oxford: Clarendon Press]. It is concluded that the use of xenon and krypton derivatives, when they can be obtained, associated with statistical heavy-atom refinement will allow one to overcome the two major limitations of the isomorphous replacement method *i.e.* non-isomorphism and the problem of optimal estimation of heavy-atom parameters.

1. Introduction

Following the pioneering work of Schoenborn and co-workers (Schoenborn, Watson & Kendrew, 1965; Schoenborn & Nobbs, 1966; Schoenborn & Featherstone, 1967; Schoenborn, 1969), and Tilton and co-workers (Tilton & Kuntz, 1982; Tilton, Kuntz & Petsko, 1984; Tilton *et al.*, 1986; Vitali, Robbins, Almo & Tilton, 1991) on myoglobin–xenon complexes, the use of xenon as a heavy atom for phase determination in protein crystallography has recently been generalized in our laboratory (Schiltz, Prangé & Fourme, 1994). Since xenon derivatives present a high degree of isomorphism (Tilton, Kuntz & Petsko, 1984; Schiltz, Fourme, Broutin & Prangé, 1995), particular attention has been paid to their phasing power in the multiple isomorphous replacement (MIR) method. Xenon derivatives have now been used successfully in the structure determinations of the human nuclear retinoid-X receptor RXR-α ligand-binding domain (Bourguet, Ruff, Chambon, Gronemeyer & Moras, 1995), the enzyme urate-oxidase (Colloc'h *et al.*, 1997) and the lipoamide dehydrogenase domain of the outer membrane protein P64k from *Neisseria meningitidis* (Li de la Sierra *et al.*, 1997).

The general principles of xenon binding to crystallized proteins have been presented in an earlier communication (Schiltz *et al.*, 1994) but an outline is given below. Xenon complexes with proteins can be obtained by subjecting a native protein crystal to a xenon-gas atmosphere of a relatively low pressure ($10–20 \times 10^5$ Pa). The appreciable solubility of xenon gas in water allows the rapid diffusion of atoms towards the interaction sites *via* the solvent channels that are present in protein crystals. Xenon binds to pre-existing atomic sized cavities that may be bound by hydrophobic residues, as is the case for myoglobin (Tilton *et al.*, 1984), or that display a more polar character, as is the case for serine proteinases (Schiltz *et al.*, 1995). The number and the occupancies of xenon binding sites varies with the applied pressure. The interaction of xenon with proteins is the result of non-covalent, weak-energy van der Waals forces (Ewing & Maestas, 1970; Tilton *et al.*, 1986), and, therefore, xenon derivatives are highly isomorphous with the native proteins (Tilton *et al.*, 1984; Schiltz *et al.*,

ISSN 0907-4449

1995). As the process of xenon binding is completely reversible (Miller *et al.*, 1981, Tilton & Kuntz, 1982), the gas pressure must be maintained during X-ray data collection. At pressures above 10×10^5 Pa, successful data collection is hampered by the absorption of X-rays from pressurized xenon gas and by the formation of xenon hydrate. Xenon hydrate is a solid clathrate whose formation is favoured by low temperatures and high xenon gas pressures. In some cases, the formation of xenon hydrate was observed at room temperature and prevented the collection of X-ray data under high xenon gas pressure (Schiltz *et al.*, 1994).

In the present communication, we describe the use of krypton as an alternative to xenon for isomorphous replacement. Krypton has a number of potential advantages over xenon even though it is 'lighter' (*i.e.* it has only 36 electrons). In particular, krypton hydrate is less stable than xenon hydrate (Forcrand, 1923, 1925) so that, with krypton, higher gas pressures can be applied (Fig. 1). In addition, krypton displays interesting anomalous scattering properties. More precisely, its *K*-absorption edge is situated at a wavelength (0.866 Å) that is readily accessible on synchrotron radiation sources and that is near the selenium and bromine edges, two atoms already exploited in anomalous scattering techniques (Hendrickson, Horton & LeMaster, 1990; Ogata, Hendrickson, Gao & Patel, 1989). In contrast, the xenon absorption edges are situated either at long or very short wavelengths (2.6 Å for the L_{III} edge and 0.36 Å for the *K* edge), both difficult to harness technically.

In the test case described below, we have been able to obtain a high-quality electron-density map for a 26 kDa protein from a single krypton derivative. Careful data collection and processing as well as new techniques for heavy-atom refinement and phasing have been essential to obtain this result. Beyond this particular case, more general conclusions can be drawn concerning data-collection strategies, the importance of eliminating systematic errors and the power of a complete implementation of statistical phasing methods.

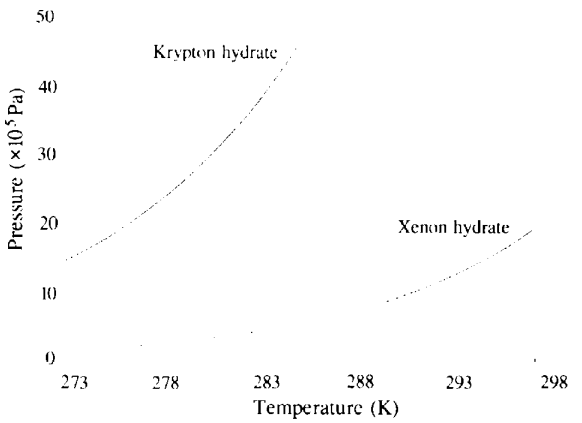


Fig. 1. Dissociation curves of xenon and krypton hydrates. These curves have been computed from data reported in Braun (1938). These curves pertain to pure water and the dissociation pressures should, therefore, be regarded as lower limits, below which one can safely assume that hydrate formation will not take place. The salts and/or organic solvents that are often present in crystal mother liquors may considerably inhibit hydrate formation and, consequently, increase the dissociation pressures (see also Schiltz *et al.*, 1994).

2. Binding of krypton to proteins

2.1. *Preliminary studies*

Initial krypton binding studies were carried out on proteins that have been shown to bind Xe atoms (Schiltz *et al.*, 1994, 1995): porcine pancreatic elastase (PPE), tetragonal hen egg-white lysozyme (HEL) and *Fusarium solani pisi* cutinase. The three-dimensional structures of these proteins have been solved and refined at resolutions up to 1.65 Å for PPE (Meyer, Cole, Radhakrishnan & Epp, 1988), 1.7 Å for HEL (Kurinov & Harrison, 1995) and 1.25 Å for cutinase (Martinez, De Geus, Lauwereys, Mathyssens & Cambillau, 1992). The experimental approach in these studies was similar to the one adopted in our previous work on protein–xenon complexes. For all of these structures, complete X-ray data sets at 2.2 Å resolution were recorded on crystals, first in the native state, and then under a krypton gas atmosphere of pressures ranging between 20 and 50×10^5 Pa. Kr atoms were located by difference-Fourier calculations using terms $|F_{\text{derivative}}| - |F_{\text{native}}|$ as amplitudes and phases computed from the refined native structures, using known coordinates deposited with the Protein Data Bank (PDB) (Bernstein *et al.*, 1977). In the cases of PPE and cutinase, a single Kr atom was found to be bound into the active site of each enzyme molecule, at exactly the same locations where xenon is known to bind. We have shown that xenon binding is a fairly general feature for the serine-hydrolase super-class of enzymes (Schiltz *et al.*, 1995), and this conclusion seems to hold for krypton as well. The signal strength of the krypton peaks in the difference-Fourier maps is rather weak, compared with peaks obtained from xenon derivatives. Although krypton is lighter than xenon, the comparatively weak signal also indicates that Kr atoms are not fully occupied. Binding of noble gases to proteins is essentially due to non-covalent van der Waals interactions, and the binding constant depends therefore on the electronic polarizability of the atom. As the polarizability of xenon is about twice as large as that of krypton, it seems plausible that krypton binding is less favoured. A more detailed study of the pressure-dependence of xenon and krypton binding was carried out for PPE and is presented in §2.3.

For HEL, we have found earlier that xenon binds into a buried hydrophobic cavity that is essentially formed by the residues M12, I55, L56, I88 and V92 (Schiltz *et al.*, unpublished results). At a pressure of 25×10^5 Pa and a temperature of 297 K, the occupancy of the Xe

atom is rather low – approximately 0.35. Additional minor sites were also found in cavities formed by the contact of symmetry-related molecules. Krypton is found to bind into the major xenon binding site (Fig. 2), but it does not bind at the other minor sites. At a pressure of 46×10^5 Pa and a temperature of 281 K, the occupancy of the Kr atom is approximately 0.5. In conclusion, it can be said that the major krypton binding sites are identical to those of xenon but that with krypton, higher pressures have to be applied to achieve noticeable substitution.

2.2. *Time-resolved studies*

We have previously demonstrated that binding of xenon to crystallized proteins takes place on a time-scale of 0 to 20 min (Schiltz *et al.*, 1994). A similar experiment, using the DW32 set-up for protein crystallography at the wiggler beamline of the LURE-DCI synchrotron (Fourme *et al.*, 1992), was carried out with krypton. X-ray data were first collected on a native crystal of HEL and the crystal was then put under krypton gas pressure of 52×10^5 Pa. Six consecutive X-ray data sets were collected on the pressurized crystal. The first data collection started immediately after pressurization. Each data collection consisted of six frames, each frame covering a 3° rotation with an exposure time of 30 s. The spindle axis was reset to zero at the beginning of each new collection so that all data sets cover the same rotation range and, therefore, comprise essentially the same reflexions. Each data collection took approximately 15 min. After data reduction, the overall completeness of each set is 66% at 2.2 Å resolution. Difference-Fourier maps with the native amplitudes were computed for all data sets. Each map revealed a prominent peak at the krypton binding site. The peak height does not change significantly between the maps obtained from the various data sets. Also, a difference-Fourier map calculated with amplitudes from the first and the last derivative data sets is completely featureless. We conclude that krypton binding was essentially complete at the start of the first derivative data collection, thus taking less than a few minutes.

The krypton gas pressure was then released and a new data collection was started immediately afterwards. A difference-Fourier map with the native data is completely featureless, thus indicating that krypton was no longer bound to the protein. In conclusion, krypton binding (and unbinding) to crystallized proteins seems to be rather fast, taking place on a sub-minute timescale, and it is completely reversible upon release of the gas pressure.

2.3. *Pressure-dependent studies*

To get a more precise idea of the pressure dependence of xenon and krypton binding, a series of diffraction data at varying pressures were collected on a crystal of PPE. Native and derivative data were collected on the same crystal, in the same orientation (*i.e.* the spindle was reset to zero at the beginning of each collection). Each data set consists of ten frames, covering a total rotation range of 20°. Data were collected to a maximum resolution

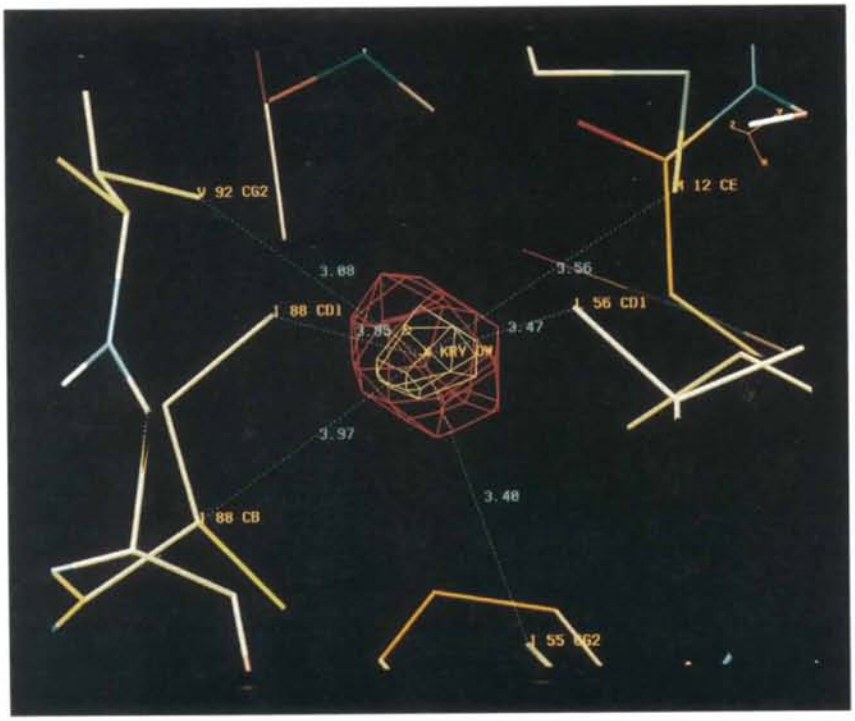


Fig. 2. Krypton binding site in a derivative of HEL (krypton pressure = 46×10^5 Pa). Displayed is a difference-Fourier map computed with terms $|F_{\text{derivative}}| - |F_{\text{native}}|$ as amplitudes and phases calculated from the refined native structure. Contours are at 10σ and 15σ above the mean density (where σ is the root-mean square of the map density). Superimposed is the refined native structure of the HEL molecule. The krypton binding site is located in a buried hydrophobic cavity that is essentially formed by the residues M12, I55, L56, I88 and V92. Also displayed are the distances (in Å) of the Kr atom to its closest non-H atom neighbours in the protein.

of 2.2 Å and after reduction are about 33% complete. The experiment was carried out with xenon and with krypton gas. An FTS-cooling system was used to keep the sample at a constant temperature of 293 K during all data collections. Difference-Fourier maps with the native data were computed for all derivative data. The value of the maximum density in each map, corresponding to the heavy-atom peak, was recorded and plotted against the value of the gas pressure (Fig. 3). On theoretical grounds, it can be shown (Schiltz *et al.*, in preparation) that, at equilibrium, the occupancy of the gas atom (Occ) is related to the applied pressure (P) by the Langmuir isotherm,

$$\mathrm{Occ} = \lambda P/(1 + \lambda P),$$

where λ is the equilibrium constant of the gas binding reaction. On the other hand, the maximum density in the difference-Fourier map is proportional to the heavy-atom occupancy. The experimental data were fitted by least-squares to the Langmuir isotherm. This allowed us to normalize the densities to unity. The experimental points together with the curve fittings are shown in Fig. 3. For a given occupancy, it is clear that with krypton gas a substantially higher pressure is required for binding, as compared to xenon. It should also be emphasized that the equilibrium constants are strongly temperature dependent (Schiltz *et al.*, unpublished results).

3. A case-study: SIRAS phasing of porcine pancreatic elastase

In order to evaluate the potential phasing power of isomorphous krypton derivatives, a SIRAS (single isomorphous replacement with anomalous scattering) test

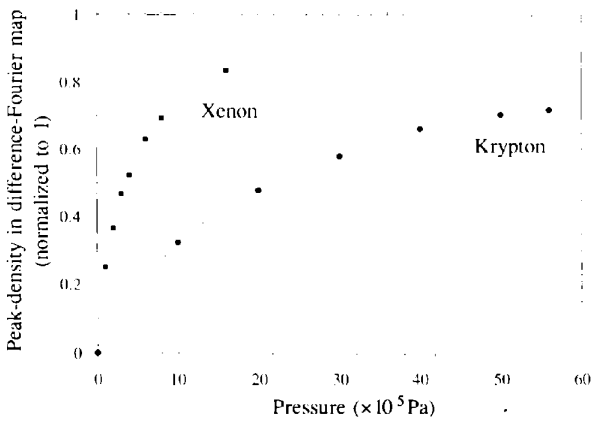


Fig. 3. Isotherms (293 K) of xenon and krypton binding to PPE. The quantity ρ/k, which is equivalent to the occupancy of the heavy atom, is plotted *versus* the gas pressure P. ρ is the maximum density (corresponding to the bound heavy atom) in a difference-Fourier map and k is a normalization factor that was obtained for each curve by least-squares fitting of the experimental data points (ρ,P) to the function, $\rho/k = \lambda P/(1 + \lambda P)$ The fitted curves are shown as dashed lines.

experiment at the high-energy side of the krypton K edge was performed on a crystal of PPE.

3.1. *Sample preparation*

Lyophilized PPE from SERVA Chemicals was used without further purification. Single crystals of PPE were grown at 291 K in 0.05 M acetate buffer, pH = 5.5, by the batch method using Na_2SO_4 as crystallizing agent (Meyer, Radhakrishnan, Cole & Presta, 1986). One prismatic shaped crystal of size 0.8 × 0.6 × 1.3 mm was used for the experiment. The crystallographic b direction is coincident with the long axis of the crystals. The crystal was mounted in a quartz capillary (diameter 1.0 mm, wall thickness 3/100 mm) with its long axis being approximately parallel to the axis of the capillary. Crystal mounting and pressurization under krypton gas proceeded as described in Schiltz *et al.* (1994). However, the pressure cell was a modified version of the one described in that paper: (i) the device is mounted on a standard Huber® goniometer head, so that the rotation arcs can be used to pre-orient the crystal; (ii) quartz capillaries with a wall thickness of 3/100 mm (Wolfgang Müller, Glas Technik & Konstruktion, Berlin, Germany) had to be used since the usual capillaries (with a wall thickness of 1/100 mm) do not resist to pressures above 20–30 × 10^5 Pa. (iii) For the same reason, 1/16″ PEEK® tubing (nominal pressure ratings 350 × 10^5 Pa) were used. The designs of the modified pressure cell will be published in a future laboratory note (Schiltz *et al.*, in preparation).

3.2. *X-ray data collection*

X-ray data collection was carried out at the DW21 station of the wiggler beamline at the LURE-DCI synchrotron, Orsay, France, with the storage ring operated at an energy of 1.85 GeV and with circulating positron currents up to 350 mA. The beamline optical elements feature an iridium-coated cylindrical mirror and a double crystal monochromator installed with Si (111) crystals (Renaud, Villette, Boeuf, Chandesris & Magnan, 1992; Renaud *et al.*, 1992). This combination focuses the X-ray beam vertically and horizontally, rejects X-rays with energies higher than 22 keV and provides for wavelength tunability in the range of 0.75–1.95 Å. The bandpass of the monochromated beam is better than 2 × 10^{-4} for X-ray energies near the krypton K edge, with an essentially complete rejection of higher order harmonics. The two-dimensional detector is a prototype of the 180 mm diameter imaging-plate system, built by J. Hendrix and A. Lentfer at the EMBL Outstation in Hamburg.

Calibration and determination of the X-ray wavelength for anomalous dispersion experiments is carried out using a silicon photodiode which has been installed to record X-ray fluorescence spectra of the sample and reference elements. This detector is mounted in the plane of the X-ray polarization to avoid any interference to the fluorescence spectrum by Bragg reflexions

from the sample or its holder. The synchronization of the monochromator crystals and the photodiode for X-ray fluorescence measurements are controlled by the same software and personal computer which pilots the monochromator. This allows X-ray fluorescence spectra to be recorded automatically and at regular intervals of energy.

In order to locate precisely the K edge of krypton (theoretically at 14.324 keV), the monochromator crystals were moved into the 14 keV region and calibrated using the $L_{\rm I}$ and $L_{\rm II}$ edges of mercury which are situated at 14.842 and 14.214 keV, respectively. A dried quartz capillary, 0.7 mm diameter, was filled with krypton gas and pressurized to 20×10^5 Pa. The krypton-filled capillary was then aligned into a non-collimated X-ray beam which was focused into a circular spot of approximately 2 mm in diameter. Behind the sample, the image of the direct beam was monitored with a CCD X-ray camera to follow the course of the spectra. An energy scan from 14.270 to 14.470 keV with 2 eV steps was completed without the krypton-filled capillary as a dry run, prior to the recording of the krypton fluorescence spectra. This dry run combined with the CCD monitor ensured that the beam position and intensity had remained stable. The X-ray fluorescence intensity was measured for 30 s at each 2 eV step, the monochromator crystals being kept stationary. During the recording of the fluorescence spectra of krypton, a dark shadow of the capillary could be seen on the CCD monitor when the threshold of the absorption edge was crossed. The form of the krypton K-edge spectrum is a simple smoothed step function showing no near-edge structure (see Fig. 4). Such a form is expected since krypton is a monatomic gas, and as such there is an absence of bonded interactions to the core electrons.

X-ray diffraction data were recorded at an energy of 14.361 keV, just past the krypton K edge, to optimize the anomalous signal (*i.e.* to maximize f''). Two diffraction images of the crystal, 90° apart in φ (spindle) were taken in order to determine the crystal missetting angles. The crystal was then oriented so that the b axis was exactly coincident with the spindle axis. This allowed for the simultaneous recording of Bijvoet pairs on the same frame. The crystal was then pressurized with krypton gas at 56×10^5 Pa as described in Schiltz *et al.* (1994). This value was chosen because it is the maximum pressure that we can reach with our current pressurization device. Data were collected at a room temperature of 296–297 K. Details of the data collections are given in Table 1. After 125° of data had been collected, the gas pressure was released and residual krypton gas was purged out of the system with the aid of a vacuum pump (Schiltz *et al.*, 1994). The spindle was brought back to its initial position and data collection on the native crystal was started. In this way, derivative and native data could be collected on the same crystal, in approximately the same orientation. Later, at the data-processing stage, we found that the crystal orientations were slightly different (by no more than 1.5°), thus indicating that the crystal had slipped somewhat during depressurization.

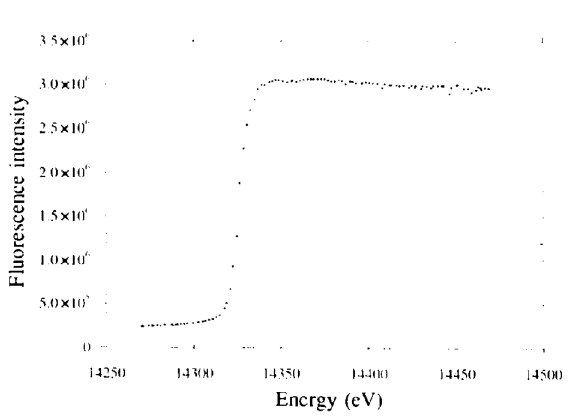


Fig. 4. Fluorescence spectrum of krypton gas.

Table 1. *Data collection and processing of native and a krypton derivative of PPE*

	Native PPE	PPE–krypton (56 bar)
Data collection		
Exposure time per frame (s)	50	60
Angular increment per frame (°)	2.0	1.0
Total rotation range (°)	110	125
Crystal-to-detector distance (mm)	180	180
Cell-parameter refinement*		
Unit-cell parameters a, b, c (Å)	51.64, 57.97, 75.32	51.65, 57.95, 75.26
Estimated standard deviations (Å)	0.03, 0.03, 0.04	0.04, 0.03, 0.05
Data reduction		
No. of measured reflexions	74579	90081
No. of unique reflexions	17752	17697
Resolution limits (Å)	30.4–1.87	30.4–1.87
Data completeness (%)	92.9	93.7
Data completeness (%) in the last resolution shell (1.97–1.87 Å)	78.6	86.2
Reflexions with $I > 3\sigma(I)$ (%)	96.7	93.8
Reflexions with $I > 3\sigma(I)$ (%) for data in the last resolution shell (1.97–1.87 Å)	94.2	83.9
Overall $R_{\rm merge}$†	0.022	0.059
$R_{\rm merge}$ for data up to 2.5 Å resolution	0.018	0.047
$R_{\rm merge}$ for data in the last resolution shell (1.97–1.87 Å)	0.050	0.161

*Space group $P2_12_12_1$. †The $R_{\rm merge}$ is defined as $\sum_{h,k,l}\sum_i |\langle I(h,k,l)\rangle - I_i(h,k,l)| / \sum_{h,k,l}\sum_i I_i(h,k,l)$ where $I_i(h,k,l)$ is the ith observation of reflexion h,k,l and $\langle I(h,k,l)\rangle$ is the weighted mean of all observations (after rejection of outliers).

With the crystal being aligned along a crystallographic axis, there is inevitably a so-called blind region of reciprocal space that can not be recorded. This was not too severe a problem as, at a wavelength of 0.863 Å, the curvature of the Ewald sphere is far less pronounced than is the case with Cu $K\alpha$ radiation. The data completeness is reasonably good, even in the highest resolution shell

(see Table 1). The blind region could have been recorded easily in a second data-collection pass by misaligning the crystal but because of beamtime constraints, this was not done here. Care was taken to record correctly low-resolution data. The imaging-plate scanner was specially modified so as to scan also the region close to the centre of the circular imaging plate and the beam-stop was placed far away from the crystal. In this way, reflexions at a resolution as low as 30 Å could be recorded. Data-collection parameters were chosen in such a way that these strong low-resolution reflexions did not saturate the detector.

For the derivative data collection, the pressurized krypton gas in the capillary caused two substantial problems: fluorescence background and absorption of X-rays. The background on the diffraction images caused by fluorescence of the pressurized krypton gas was about three to five times more intense than that observed for the native data. In order to enhance the signal-to-background ratio, the angular width per frame was chosen to be 1° whereas for the native crystal, an angular width of 2° could be applied without difficulties. Absorption of X-rays by the pressurized krypton gas in the capillary was another problem that we had already encountered in our work with xenon gas (Schiltz *et al.*, 1994). Care was taken to orient the crystal so that it was the incident X-ray beam, and not the diffracted beams, that travelled through the krypton gas atmosphere (see Fig. 5). The intensity loss of the incident beam was compensated for by increasing the exposure times. Residual systematic errors resulting from absorption effects could be eliminated by parameterized local scaling of derivative to native data as described below.

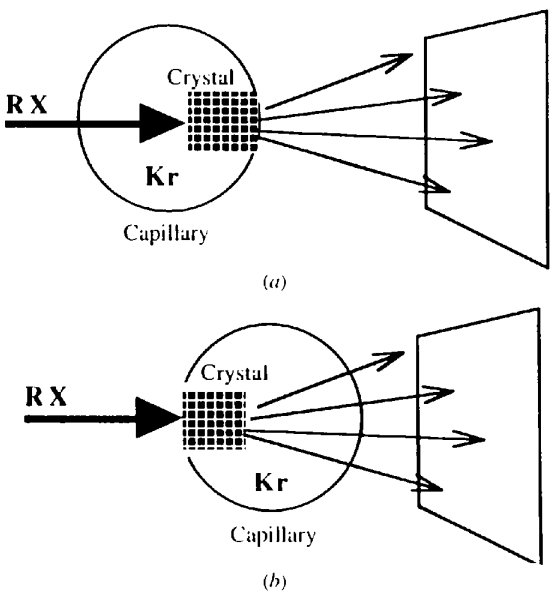


Fig. 5. Absorption of X-rays by pressurized krypton gas in the capillary. (*a*) If the crystal is oriented in such a way that it is the incident X-ray beam that travels through the krypton gas atmosphere, then only that beam will be absorbed and the overall effect will be a reduction in intensity that is the same for all diffracted beams. (*b*) If the crystal is oriented in such a way that the diffracted beams travel through the krypton gas atmosphere, then the absorption will affect the intensity of each beam differently and the quality of the data may severely deteriorate.

3.3. *Data processing*

Data processing was carried out with the *MOSFLM* program (Leslie, 1992). Data reduction and merging were carried out with programs of the *CCP*4 (Colaborative Computational Project, Number 4, 1994) suite. Results are summarized in Table 1. Cell parameters were refined independently for the native and the derivative data by the post-refinement technique (Winkler, Schutt & Harrison, 1979; Rossmann, Leslie, Abdel-Meguid & Tsukihara, 1979) implemented in *MOSFLM*. A few low-resolution reflexions did contain saturated pixels. These reflexions were not rejected but their intensity was evaluated *via* profile fitting of the non-saturated parts of the reflexion spots (Leslie, 1991). Lorentz-polarization corrections (Arndt & Wonacott, 1977; Kahn, Fourme, Gadet, Janin & André, 1982) were applied to all intensity measurements. For each data set, the various frames were brought on a common scale by applying overall and resolution-dependent scale factors (Hamilton, Rollet & Sparks, 1965) that were evaluated by the method of Fox & Holmes (1966). The derivative data were scaled to the native data by the local scaling method (Matthews & Czerwinski, 1975) of Kabsch (1988), as implemented in the *CCP*4 (Collaborative Computational Project, Number 4, 1994) program *SCALA* (Evans, 1993). A continuous scaling function was evaluated, every 10° of rotation at 4 × 4 equally spaced points on the detector surface

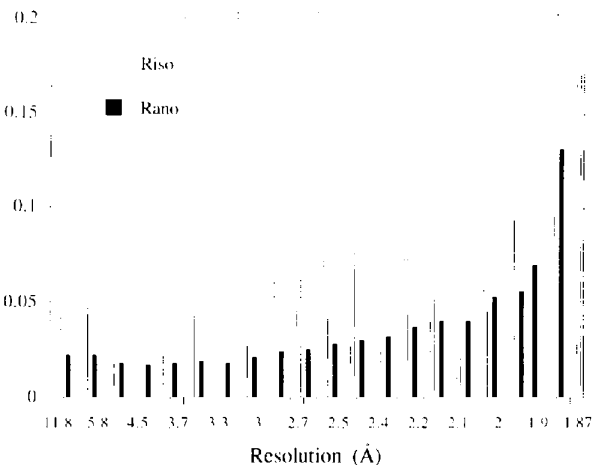


Fig. 6. Experimental isomorphous and anomalous differences. The R_{iso} is defined as $R_{iso} = (\sum_{h,k,l}|\mathbf{I}_P - \mathbf{I}_{PH}|)/\sum_{h,k,l}|\mathbf{I}_P|$, where $\mathbf{I}_P$ and $\mathbf{I}_{PH}$ pertain to the native and derivative intensity measurements of reflexion *h,k,l*, respectively. The R_{ano} is defined as, $R_{ano} = (\sum_{h,k,l}|\mathbf{I}_+ - \mathbf{I}_-|)/\sum_{h,k,l}|\mathbf{I}_+ + \mathbf{I}_-|$, where $\mathbf{I}_+$ and $\mathbf{I}_-$ are intensity measurements of the anomalous pairs.

(in the x and y directions) by iterative least-squares minimization, assuming that 'true' intensities correspond to the native intensity measurements. Restraints were put on neighbouring scale factors in order to minimize large variations. Thus, the scaling function applied to derivative measurements is a smoothly varying, interpolated function of the three-dimensional coordinate of the observation. It should be noted that local scaling was performed on the unmerged data.

For data merging, intensity observations deviating by more than 3.0 (for the native) or 4.0 (for the derivative) standard deviations from the weighted mean were rejected as outliers. These constituted only small fractions of the total amount of measured data (3.2% for the native and 2.1% for the derivative). The merged intensity data were converted to structure-factor amplitudes by the method of French & Wilson (1978) and put on an quasi-absolute scale by a Wilson (1949) plot computed with data in the resolution range of 3.5–1.87 Å.

Isomorphous and anomalous differences plotted *versus* resolution are shown in Fig. 6. As can be seen, the isomorphous differences are significantly greater than experimental errors (as judged by merging R factors; see Table 1), whereas the anomalous differences are quite small.

3.4. *Phase determination and refinement*

Isomorphous and anomalous difference Patterson maps were computed. The three Harker sections of each Patterson map are shown in Figs. 7 and 8. In both maps, the Harker vectors corresponding to the single Kr atom are the signals that display maximum intensity. Apart from these peaks, both maps are featureless.

Heavy-atom refinement and phasing were carried out with the program *SHARP* (La Fortelle & Bricogne, 1997). *SHARP*, which stands for statistical heavy-atom refinement and phasing, effectively implements the full theory of maximum-likelihood heavy-atom refinement as described in Bricogne (1991*a*,*b*). The version of *SHARP* used in this work had the following capabilities. (1) Refinement of starting parameters. (2) Update of the substitution model (*i.e.* addition of minor sites, anisotropic thermal parameter refinement of individual heavy atoms) through the use of a new kind of residual map based on the gradient of the log-likelihood (LLG) function. (3) Final calculation, for each reflexion, of phase probability distributions, encoded in Hendrickson & Lattman (1970) coefficients, and of a centroid structure factor for the Fourier synthesis of an electron-density map.

The refinement proceeded as follows: we started refining a simple heavy-atom model, as suggested by the analysis of the Patterson maps (one site of occupancy 1.0 and with an isotropic temperature factor of 30 Å^2). Refinement was carried out until convergence, and Fourier coefficients were then calculated for residual LLG-gradient maps. Careful examination of the 'isomorphous' residual map around the krypton site showed fea-

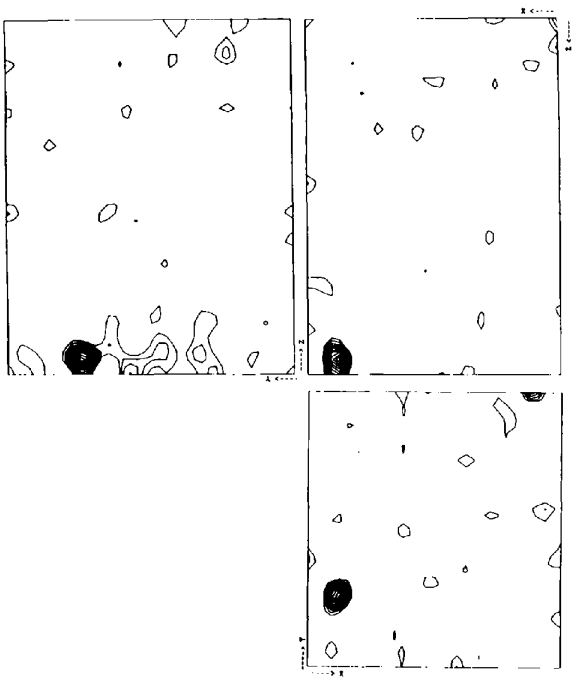


Fig. 7. Isomorphous-difference Patterson map. The asymmetric part ($0 \le X \le 0.5$, $0 \le Y \le 0.5$, $0 \le Z \le 0.5$) of the three Harker sections $X = 0.5$ (top left), $Y = 0.5$ (top right) and $Z = 0.5$ (bottom right) are shown. Contours are at intervals of 1σ, starting 2σ above the mean density (where σ is the root mean square of the map density).

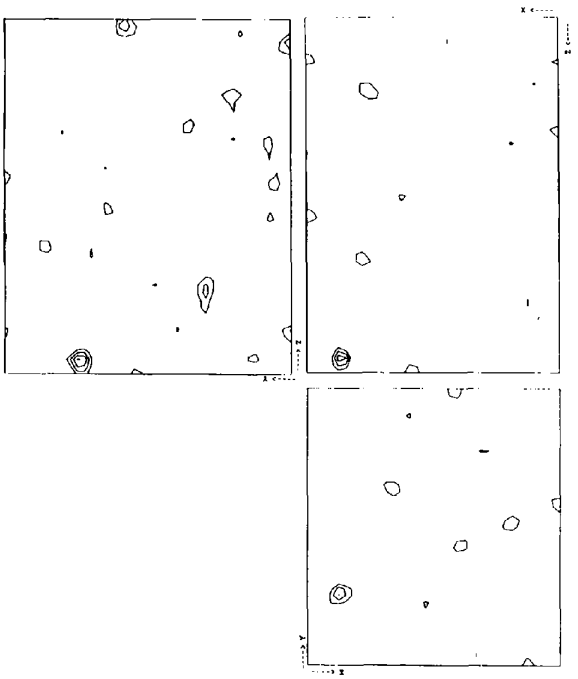
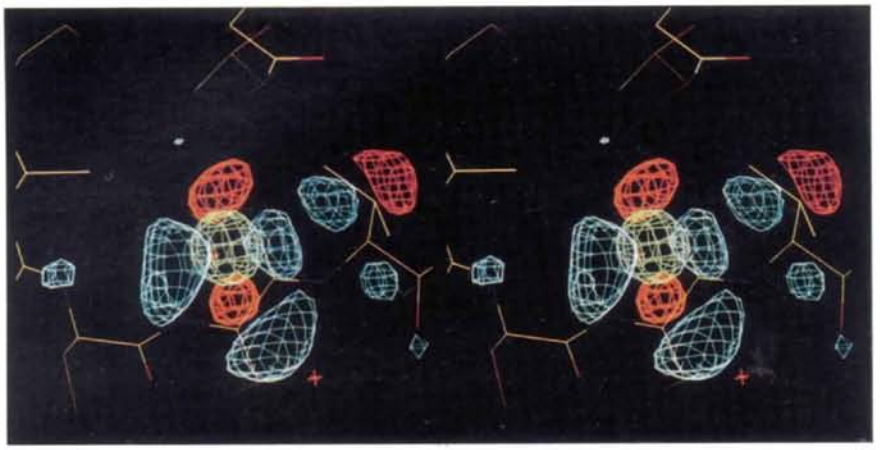


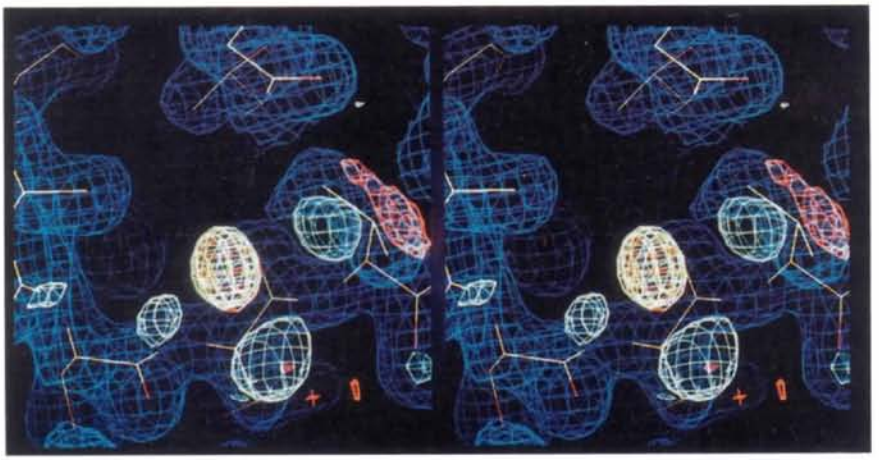
Fig. 8. Anomalous-difference Patterson map. The asymmetric part ($0 \le X \le 0.5$, $0 \le Y \le 0.5$, $0 \le Z \le 0.5$) of the three Harker sections $X = 0.5$ (top left), $Y = 0.5$ (top right) and $Z = 0.5$ (bottom right) are shown. Contours are at intervals of 1σ, starting 2σ above the mean density (where σ is the root mean square of the map density).

tures that strongly suggested anisotropic thermal motion of the Kr atom (Fig. 9*a*). As a consequence, an additional anisotropic perturbation to the temperature factor for that atom was refined. At this point, the isomorphous residual map (Fig. 9*b*) showed no feature above 4 standard deviations, except for two negative peaks close to the krypton site (that could correspond to departing water molecules) and a side-chain displacement. The anomalous residual map showed no significant feature. Lack-of-isomorphism parameters were refined along with the other parameters, and converged to small but non-zero values for the isomorphous signal, and to larger values for the anomalous signal, thus probably indicating that the standard deviations pertaining to anomalous differences were under-estimated in the measurement file. The final model for the heavy-atom substitution is summarized in Tables 2 and 3.

SHARP was then used to calculate the complex centroids of the two-dimensional probability distributions in the plane of the Harker construction for each reflexion. These complex numbers are the best coefficients for a Fourier synthesis to obtain an electron-density map



(*a*)



(*b*)

Fig. 9. Residual maps after heavy-atom parameter refinement with the program *SHARP* (stereoviews). (*a*) The heavy-atom parameter refinement used an isotropic thermal factor for the Kr atom. Displayed are the current heavy-atom model (in green – a single isotropic Kr atom) and the isomorphous residual map (red for postive density, light blue for negative density). This map is based on the gradient of the log-likelihood function and indicates discrepancies between the current heavy-atom model and the experimental data. The two pairs of positive and negative residual density peaks that are situated in orthogonal directions relative to each other around the heavy atom strongly suggest anisotropic thermal movement of the krypton atom. (*b*) The heavy-atom parameter refinement used an anisotropic thermal factor for the Kr atom. The positive and negative density peaks around the heavy atom are no longer present. The residual map only shows two negative peaks close to the krypton site (towards the bottom of the figure) that could correspond to replaced water molecules as well as a pair of positive and negative density peaks (near the side chain of Val209) that could correspond to a side-chain displacement. Also displayed (in yellow) is a standard difference-Fourier map computed with terms $|F_{\text{derivative}}| - |F_{\text{native}}|$ as amplitudes and phases calculated from the refined native structure. This difference-Fourier map displays a peak at the krypton binding site whose shape is in perfect agreement with the current heavy-atom model (in green), thus validating the use of an anisotropic thermal parameter refinement. The final SIRAS electron-density map after density modification is shown in blue. Superimposed to these maps is the refined native structure of the PPE molecule.

(Blow & Crick, 1959). During the same calculation, the Hendrickson–Lattman (1970) coefficients describing the marginal phase probability distribution were also computed and written to the output file.

The next step of the analysis was to improve the quality of the electron-density map thus obtained. We used the procedure of density modification as implemented in the program *SOLOMON* (Abrahams & Leslie, 1996). This program makes use of the phase probability distributions as encoded in the Hendrickson–Lattman (1970) coefficients, and embodies a set of new ideas such as the use of the map's local standard deviation to determine the solvent boundaries, as well as 'solvent flipping', where the electron density in the solvent region is inverted rather than set to a constant value. This procedure improved the phases to a great extent (Table 4).

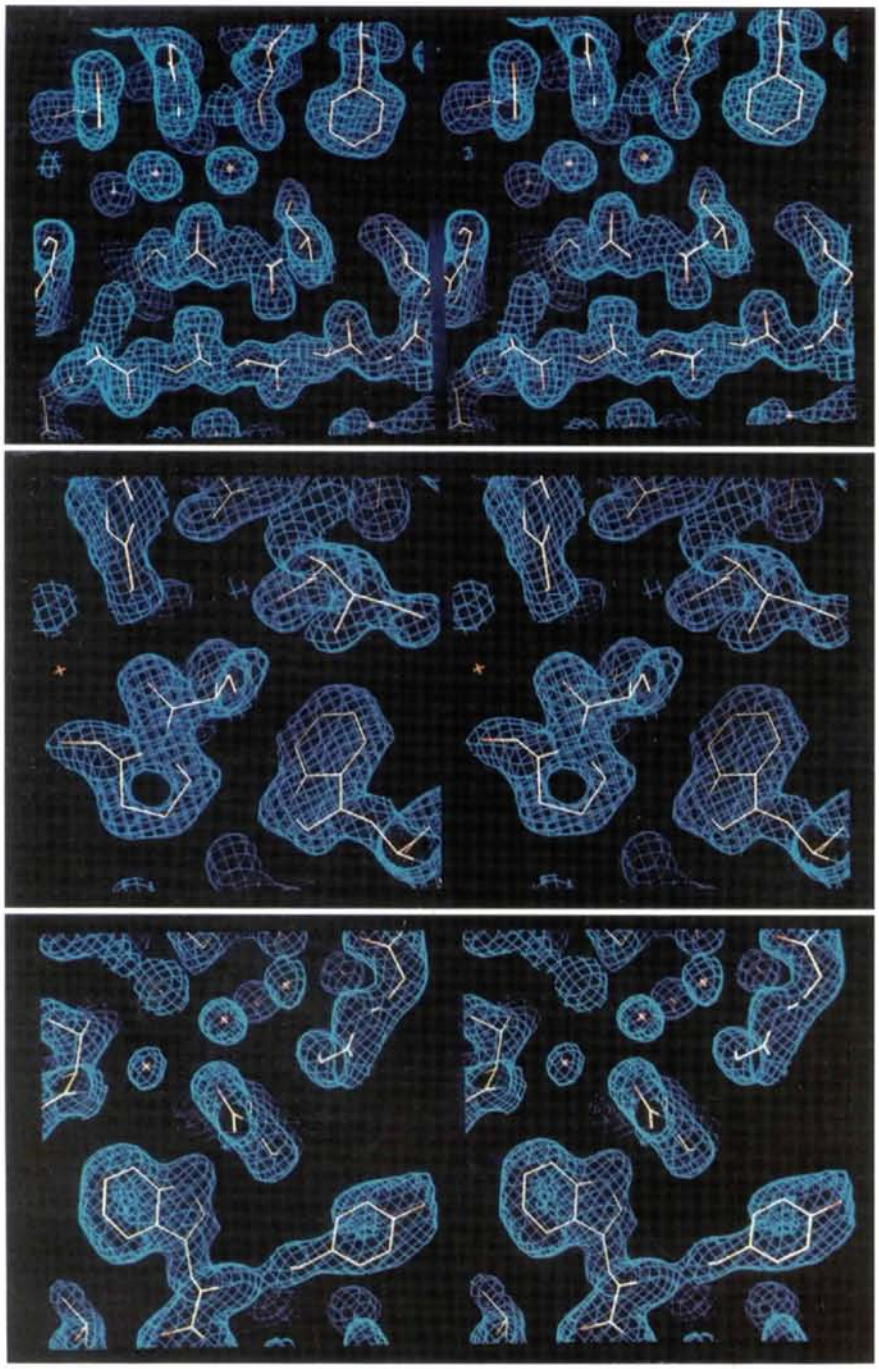


Fig. 10. Representative stereoviews of the SIRAS electron-density map after density modification with *SOLOMON*. Superimposed on this map is the refined native structure of the PPE molecule as obtained from the PDB (entry 3EST). It should be noted that this electron-density map was obtained from purely experimental data, *i.e.* prior to any model building. As can be seen, the agreement with the model is very good. The exceptional quality of the electron-density map is illustrated by the fact that internal water molecules show up as spherical density peaks.

3.5. *Results*

Examples of the final, solvent-flattened SIRAS map are shown in Fig. 10. The map is of exceptionally high quality and in almost perfect agreement with the known protein structure: the map correlation coefficient* between this final map and a map computed from the refined high-resolution native structure (PDB entry 3EST) is 0.85. Of the 129 refined water molecules that are present in the 3EST model, 70 are clearly visible in the solvent-flattened SIRAS map. Almost all of these molecules are internal waters. The electron density corresponding to water molecules at the surface of the protein may have been removed by the solvent-flattening procedure.

As the crystallization protocol used here was different from the one used in the high-resolution structure determination of PPE – the authors used crystals that were soaked in a 70% methanol solution prior to data collection (Meyer *et al.*, 1988) – we decided to refine the PPE structure against our native data set. Refinement started from the 3EST structure and was carried out by restrained least-squares minimization with the *PROLSQ* program (Konnert & Hendrickson, 1980) alternating with manual intervention on a graphics display. The final R and free R factors† for data in the resolution range of 10–1.87 Å are 0.175 and 0.224, respectively. The correlation coefficient between the 1.87 Å solvent-flattened SIRAS map and a map calculated from our refined native structure is 0.88. The same refinement procedure was also applied to the PPE–krypton complex. Details and results will be presented elsewhere (Schiltz *et al.*, in preparation). The coordinates of the refined structures and the measured structure-factor amplitudes have been deposited in the Protein Data Bank (Bernstein *et al.*, 1977).‡

4. Discussion

That the equivalent of half a Kr atom (about 15 and 2 e for the isomorphous and anomalous signals, respectively) is sufficient for the structure determination of a 26 kDa protein is in itself a remarkable result. The near-perfect isomorphism between the krypton derivative and the native structure, careful data collection and processing, and a proper statistical heavy-atom refinement and phasing were essential for the successful production of a high-quality electron-density map.

4.1. *Obtaining high-quality data*

The expected mean fractional changes of intensity upon heavy-atom binding can be evaluated by the Crick & Magdoff (1956) formula.* For the case of a half-occupied Kr atom in PPE, one obtains isomorphous differences of 0.073 (for acentric reflexions) and 0.103 (for centric reflexions) and anomalous differences of 0.018. These values should be compared to the experimental intensity differences reported in Fig. 6. Therefore, we expect the isomorphous signal to be weak, and a minute anomalous information. Nevertheless, if these signals are to be used for phasing, the experimental noise has to be reduced to the minimum, implying that the primary X-ray data have to be of the highest quality. The merging R factors reported in Table 1 are at the lower limit of what is usually achieved in protein crystallography (it should be noted that, for the derivative data, the $R_{\rm merge}$ not only measures the precision of the intensity data but also contains genuine intensity differences which result from the anomalous scattering by the Kr atom). As the whole process of phase determination relies on the accuracy of intensity differences, it is of greatest importance that Bijvoet mates on the one hand and corresponding native and derivative reflexions on the other hand are recorded under almost identical conditions. This strategy will allow for many systematic errors to cancel out in the process of phasing. The importance of simultaneous recording of Bijvoet pairs on the same (or adjacent) data frames has been emphasized in the development of the MAD (multiwavelength anomalous dispersion) method (Hendrickson, Smith, Phizackerley & Merritt, 1988). Here, this same procedure was used by orienting the crystal so that one of the crystallographic twofold axes was exactly coincident with the spindle axis.

One of the major limitations of the method of isomorphous replacement is that, in general, the native and derivative data measurements are performed on two different samples, thus giving rise to lack-of-isomorphism

* The map correlation coefficient between two maps (with densities ρ_1 and ρ_2, respectively), sampled at discrete intervals in x, y and z, is defined as $(\langle\rho_1\rho_2\rangle - \langle\rho_1\rangle\langle\rho_2\rangle)/[(\langle\rho_1^2\rangle - \langle\rho_1\rangle^2)^{1/2}(\langle\rho_2^2\rangle - \langle\rho_2\rangle^2)^{1/2}] = \text{Covariance}(\rho_1\rho_2)/[\text{Variance}(\rho_1)]^{1/2}[\text{Variance}(\rho_1)]^{1/2}$ where the mean values are taken over all sampling points.

† The R factor is defined as being the normalized crystallographic residual $R = \sum_{h,k,l}||F_{\rm obs}(h,k,l)| - k|F_{\rm calc}(h,k,l)||/\sum_{h,k,l}|F_{\rm obs}(h,k,l)|$, where h,k,l are the reciprocal lattice points of the crystal, $|F_{\rm obs}(h,k,l)|$ and $|F_{\rm calc}(h,k,l)|$ are the observed and calculated structure-factor amplitudes, respectively, and k is a scale factor. The definition of the free R factor (Brünger, 1992, 1993) is the same, but the summations are carried out over a small subset of reflexion (10% in the present case) that have been excluded from the refinement.

‡ Atomic coordinates and structure factors have been deposited with the Protein Data Bank, Brookhaven National Laboratory (Reference: 1LVY, R1LVYSF). Free copies may be obtained through The Managing Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England (Reference: HA0149).

* The formula is given by $[\langle(\Delta\mathbf{I})^2\rangle]^{1/2}/\langle\mathbf{I}\rangle = \gamma(N_H/N_P)^{1/2}(f_H/f_P)$, where N_H is the number of heavy atoms, N_P is the number of protein atoms, f_P the average protein atomic scattering factor (at zero scattering angle), f_H is the heavy-atom scattering factor and γ equals 2 for centric reflexions and $2^{1/2}$ for acentric reflexions. For the case of a half-occupied Kr atom in PPE: $N_P = 1942$, $N_H = 1$, $f_P = 6.7$ and $f_H = 15.25$ (f_H for a half-occupied Kr atom, corrected by $f' = -5.5$ at 14.361 keV). The same formula can be used for the estimation of anomalous differences (with $\Delta I = I_+ - I_-$) by setting $f_H = 2f''$. For the PPE–krypton derivative, $2f''$ was taken to be 3.8 (to account for the occupancy of 0.5). It has to be remembered that, strictly speaking, this formula is only valid for reflexions at zero scattering angle.

Table 2. *Heavy-atom parameter refinement*

	Kr atom
Fractional coordinates (x, y, z)	$(0.2808 \pm 0.0002; 0.0649 \pm 0.0003; 0.4888 \pm 0.0002)$
Occupancy	0.496 ± 0.009
Isotropic temperature factor (B_{iso}) (Å^2)	(18.9 ± 0.7)
Anisotropic perturbations (tensor of b_{ij})*	$(-1.8 \pm 0.2; 1.4 \pm 0.3; 1.0 \pm 0.3; 2.3 \pm 0.4; 1.6 \pm 0.4; 0.3 \pm 0.3)$
Anomalous scattering factors	
f'	−5.5 (not refined)
f''	3.5 ± 0.2 (refined)†

* The temperature factor of the Kr atom is modelled as the sum of an isotropic term (a diagonal matrix with elements B_{iso}) and an anisotropic perturbation (a tensor with elements b_{ij}):

$$\exp\left[-B_{iso}\left(\frac{\sin\theta}{\lambda}\right)^2\right]\exp\left[-b_{11}\left(\frac{h}{a^*}\right)^2-2b_{12}\left(\frac{h}{a^*}\right)\left(\frac{k}{b^*}\right)-2b_{13}\left(\frac{h}{a^*}\right)\left(\frac{l}{c^*}\right)-b_{22}\left(\frac{k}{b^*}\right)^2-2b_{23}\left(\frac{k}{b^*}\right)\left(\frac{l}{c^*}\right)-b_{33}\left(\frac{l}{c^*}\right)^2\right].$$

† The occupancy of the heavy atom, and its imaginary and real dispersion corrections (f' and f'') cannot be refined independently as each parameter is linearly correlated with the two other. For this reason, f' was fixed at the theoretical value of −5.5 whereas the two other parameters were refined.

Table 3. *Lack-of-isomorphism parameters versus resolution*

Isomorphous lack-of-closure is defined as $||\mathbf{F}_{PH}| - |\mathbf{F}_P + \mathbf{F}_H||$ where $\mathbf{F}_H$ is the calculated heavy-atom structure factor. $\mathbf{F}_P$ is the protein structure factor with $\mathbf{F}_P = |\mathbf{F}_P|\exp(i\varphi_P)$ where $|\mathbf{F}_P|$ is the measured protein structure-factor amplitude and φ_P is the calculated centroid phase angle of the protein structure factor. $|\mathbf{F}_{PH}|$ is the measured heavy-atom derivative structure-factor amplitude. Anomalous lack-of-closure is defined as $\left|||\mathbf{F}_{PH^+}| - |\mathbf{F}_{PH^-}|| - 2\mathbf{F}_{H''}\sin(i\varphi_P)\right|$ where $|\mathbf{F}_{PH^+}| - |\mathbf{F}_{PH^-}|$ is the measured Bijvoet difference of the heavy-atom derivative structure-factor amplitude and $\mathbf{F}_{H''}$ is the imaginary part of the calculated heavy-atom structure factor.

Resolution (Å)								
50.0	5.26	3.73	3.05	2.64	2.36	2.16	2.00	1.87
R.m.s. isomorphous lack-of-closure								
8.2	14.3	12.6	10.2	9.3	9.2	8.4	7.3	
R.m.s. anomalous lack-of-closure								
3.0	3.5	4.1	4.6	5.3	5.9	6.6	7.3	

Table 4. *Phasing statistics in resolution shells*

Heavy-atom refinement and phasing were carried out with the *SHARP* program. Density modification was carried out with the *SOLOMON* program. See text for details. FOM is the mean figure of merit. $\langle\Delta\varphi\rangle$ is the figure-of-merit weighted mean phase error (in °) with respect to phases calculated from the refined native PPE structure. CORR is the mean map correlation coefficient with respect to a map computed from the refined native PPE structure.

	Resolution (Å)								
	50.0	5.26	3.73	3.05	2.64	2.36	2.16	2.00	1.87
SIRAS phasing									
FOM	0.56	0.48	0.47	0.43	0.38	0.33	0.27	0.18	
$\langle\Delta\varphi\rangle$	55.0	53.4	53.9	57.4	62.5	66.2	72.8	78.9	
CORR	0.47	0.52	0.49	0.47	0.41	0.35	0.28	0.19	
SIRAS phasing and density modification									
FOM	0.95	0.97	0.95	0.93	0.92	0.93	0.92	0.85	
$\langle\Delta\varphi\rangle$	20.2	16.0	21.4	24.0	23.9	23.0	24.6	26.9	
CORR	0.82	0.93	0.89	0.87	0.86	0.87	0.86	0.83	
SIR phasing									
FOM	0.45	0.42	0.42	0.39	0.34	0.31	0.26	0.17	
$\langle\Delta\varphi\rangle$	62.0	59.4	60.9	63.2	67.1	69.4	75.0	80.4	
CORR	0.39	0.46	0.43	0.41	0.36	0.31	0.24	0.17	
SIR phasing and density modification									
FOM	0.94	0.95	0.94	0.93	0.93	0.93	0.92	0.91	
$\langle\Delta\varphi\rangle$	32.7	32.6	36.9	39.5	39.9	39.6	41.0	44.0	
CORR	0.68	0.78	0.74	0.71	0.70	0.70	0.69	0.65	

and intercrystal scaling errors. The benefits of recording native and derivative data under nearly identical conditions have been illustrated in the structure determination of bovine mitochondrial F_1 ATPase where, for each derivative data set, a corresponding native data set was collected from crystals of the same batch which had been soaked under identical conditions in the stabilizing solution, but without the heavy-atom compound (Abrahams, Lutter, Leslie & Walker, 1995; Abrahams & Leslie, 1996). The use of xenon and krypton derivatives associated with fast data-collection schemes using synchrotron radiation offers the unique possibility of recording native and derivative data on the same crystal, in the same orientation. This strategy was followed in the present study, and the largest remaining source of systematic errors is the absorption of X-rays by the pressurized gas in the capillary. As indicated in the previous section, a judicious orientation of the crystal, increased exposure times and local scaling of derivative to native data allowed one to deal with this problem in a satisfactory manner. We have suggested earlier (Schiltz *et al.*, 1994) that the amount of gas in the X-ray beam path can be reduced by using capillaries of the smallest possible diameter allowed by the crystal size. It has also been proposed (Tilton, 1988) that the void volume should be filled with a medium that is transparent to X-rays, such as Mylar. These procedures will help to minimize the fluorescence background as well. Fluorescence by the pressurized gas is in itself not a source of systematic errors as the background on the detector surface is remarkably uniform. The main effect of the increased background level is to reduce the signal-to-noise ratio and thus impair the quality of the data. As indicated before, this can be counteracted by using small angular rotations. In this respect, multiwire proportional chamber detectors and fast-readout CCD detectors might have clear advantages over imaging-plate scanners, which have a rather low duty cycle. Another remedy against high fluorescence background would be to increase the crystal-to-detector distance (d). The fluorescence background, being uniformly distributed in space, falls off as d^2, whereas the intensity of the diffracted beams, as a direct result of the very narrow divergence of the synchrotron source, remains

approximately constant (Helliwell, Ealick, Doing, Irving & Szebenyi, 1993). Thus, by increasing the effective detector surface (which is possible with imaging plates) the detector can be moved further away for a given resolution limit and a significant improvement in signal-to-noise ratio is possible.

The single most important step in obtaining high-quality data and eliminating residual systematic errors in the isomorphous and anomalous differences was the local scaling of derivative to native data. Parametrized local scaling of reflexions from adjacent data sets is commonly used with the MAD method (Hendrickson, 1985) but has only occasionally been carried out in conjunction with the MIR technique (Matthews & Czerwinski, 1975). For the present study, local scaling was absolutely essential. Initial attempts to determine phases from derivative data that had been scaled to the native by Kraut's method (Kraut, Sieker, High & Freer, 1962) ran into difficulties. Also, no clear signal was detectable in the anomalous Patterson map computed from these data. After the local scaling procedure described in §3.3 had been carried out, the Harker peaks in the isomorphous Patterson map were twice as strong as before local scaling. Even more important, the Harker vectors in the anomalous Patterson map were now clearly visible as being the most prominent peaks. Thus, the local scaling procedure not only reduces the remaining systematic errors in isomorphous differences, but also enhances the accuracy of the anomalous differences.

A final remark is in order concerning the resolution range of the measured data. We have been especially careful not only to record the high-resolution data but also to accurately collect the reflexions at low angles. Thus, in the resolution range from 8 to 25 Å, the data display a completeness of better than 90%. It is an unfortunate but still common practice in protein crystallography not to measure the low-resolution reflexions accurately (*e.g.* by saturating the detector) or not to measure them at all (*e.g.* by placing the beam-stop too close to the crystal). The low-resolution terms contain the basic information about the solvent distribution in the crystal and, hence, about the molecular boundaries. In the present study, one reason why the solvent-flattening procedure was so successful in refining phases is to be found in the fact that accurate amplitudes for the low-resolution reflexions were available and that a substantial number of these reflexions were already correctly phased by the SIRAS procedure.

4.2. *Heavy-atom parameter refinement and phasing*

The use of the program *SHARP* for heavy-atom refinement and phasing has played an essential role in the successful determination of the high-quality electron-density map. Our initial approach was to use the *CCP*4 (Collaborative Computational Project, Number 4, 1994) program *MLPHARE* (Otwinowski, 1991). However, with this algorithm we had difficulties in dealing correctly with the rather small anomalous differences. After less than five cycles of refinement, the anomalous occupancy of the Kr atom dropped to zero. Also, refinement of an individual atomic *B* factor was unstable and never converged. Simultaneous refinement of individual *B* factors and occupancies for the heavy atoms may be difficult as both parameters are not completely uncorrelated if the resolution of the data is limited. In the present case however, accurate high-resolution data were available (up to 1.87 Å), and only one single heavy atom was refined. It was, therefore, unexpected that the refinement should have run into trouble. Difficulties in refining weak anomalous signals with *MLPHARE* have already been reported elsewhere (Baker, Anderson, Dobbs & Dodson, 1995). Not surprisingly, the SIRAS map which was calculated from *MLPHARE* phases is of rather poor quality and has a correlation coefficient of 0.39 with a map computed from the refined native structure of PPE. After phase refinement with the density-modification procedure *SOLOMON*, the correlation coefficient rose to 0.69.

Heavy-atom parameter refinement in Patterson space with the program *VECREF* (Tickle, 1991) was somewhat more successful. The refinement behaves well and gives an overall occupancy and *B* factor for the Kr atom that are similar to those obtained with *SHARP*. However, this program only refines heavy-atom parameters so that estimation of lack-of-isomorphism and phasing still have to be carried out with *MLPHARE*. In principle, parameters pertaining to anomalous scattering (*i.e.* anomalous occupancies, dispersion correction f'') can also be refined with *VECREF* on an anomalous Patterson map with an appropriate form factor but this was not tested here.

The new approach of maximum-likelihood refinement implemented in *SHARP* was prompted by an analysis of the pathologies associated with least-squares refinement methods which stem from two major pitfalls (Bricogne, 1991*a*,*b*). (1) The least-squares residual needs to refer to a 'perfect value' for the modulus and phase of the native structure factor $\mathbf{F}_P$. These perfect values are then replaced by estimates, a process which introduces unpredictable amounts of bias in the result. Phase-integrated least-squares programs, such as *MLPHARE*, only bring a partial solution to this problem because the native modulus is still considered error-free, leading to over-weighting the native contribution with respect to that of the derivatives. (2) Least-squares refinement algorithms can only cyclically estimate parameters for the lack-of-isomorphism error. This results in poor convergence properties, the lack-of-isomorphism variance being correlated with heavy-atom occupancies and temperature factors.

The maximum-likelihood formalism used in *SHARP* deals with these two problems as follows (La Fortelle & Bricogne, 1997). (1) The bias problem related to the 'status of $\mathbf{F}_P$' is solved by recognising that the complex

native structure factor is not known. Instead of using estimated values for $\mathbf{F}_P$ in setting up the observational equations, *SHARP* uses a trial value for it, and eventually integrates the likelihood function over all possible trial values for $\mathbf{F}_P$. (2) The randomness caused by lack-of-isomorphism is treated as part of the substitution model. As a result, it does not attempt to match a calculated value to an observed value for the structure-factor amplitudes (as is the case in the least-squares method), but to match the probability distributions of these two quantities. This approach has proved, through tests on calculated data, that it can effectively remove bias on the values of the refined parameters and on the phase-probability distributions derived from them.

4.3. *The contribution from the anomalous signal*

Given the fact that the anomalous differences are rather weak, as compared to the uncertainties of experimental measurements, the question arises whether these differences contain any significant phase information. The clearly interpretable Harker peaks in the anomalous difference Patterson map already indicate that the measured anomalous differences do not simply arise from random experimental noise.

A more conclusive method to quantify the amount of information provided by the anomalous signal is to refine and phase the data based on the isomorphous differences only – in effect an SIR (single isomorphous replacement) experiment. Refinement did not present difficulties: the anisotropy of the Kr atom was again clearly visible in the residual maps and could be refined. The map obtained after phasing by *SHARP* was less informative than the SIRAS map, because of the inherent bimodality of the SIR phase distributions. Nonetheless, we observed that the information added to the phase distributions in the process of density modification enabled most of this bimodality to be removed. The final electron-density map is of lower quality than the equivalent SIRAS map, but is nevertheless clearly interpretable and has a correlation coefficient of 0.72 with the map computed from the refined native structure.

It is worth noting that the SIR figures of merit after solvent flattening (Table 4) are roughly similar to the equivalent SIRAS figures of merit, even though the electron-density map is of lower quality. This fact underlines the main role of the solvent-flattening procedure, namely that of choosing the right mode in bimodal phase distributions, under the constraint of an envelope hypothesis. The phase distributions after solvent flattening are made unimodal in the SIR case as in the SIRAS case, but since the first SIR envelope was less precisely defined than the first SIRAS envelope, the incorrect mode was favoured for a number of reflexions, therefore introducing self-consistent bias in the iterative solvent-flattening procedure. The important role of the weak anomalous signal is then, when exploited with proper statistical methods, to produce a sufficient number of unimodal reflexions (mostly at low resolution) to minimize envelope bias.

4.4. *Krypton versus xenon*

Initially, the driving force behind our experiments with krypton was the fact that the applicability of xenon derivatives may be limited by the formation of xenon hydrate. This becomes a serious problem at temperatures below 277 K and/or when the crystal mother liquor does not contain salts or organic solvents (which would normally inhibit hydrate formation). Krypton, on the other hand, does not present such problems as its corresponding hydrate is far less stable (Fig. 1). As an example, for HEL at a temperature of 285 K xenon hydrate formed at a gas pressure of 25×10^5 Pa, whereas with krypton gas, pressures above 50×10^5 Pa could be applied without hydrate formation. In a number of cases, the use of krypton might therefore prove to be more advantageous than xenon.

The price to pay for this advantage is the smaller isomorphous signal of krypton with respect to xenon as well as the fact that with krypton gas, higher pressures need to be applied in order to achieve the same degree of substitution. As far as the anomalous scattering properties are concerned, it should be emphasized that SIRAS experiments can easily be performed with xenon as well. In fact, for the SIRAS method it is not necessarily required to collect data at a wavelength close to an absorption edge (as is the case with MAD). The important point is to have an imaginary anomalous scattering factor (f'') of a significant magnitude. For xenon, the wavelength range between 1.0 and 1.7 Å (where f'' increases monotonically from 3.5 to 8.6) is most useful. Indeed, SIRAS experiments with xenon have been carried out successfully using Cu $K\alpha$ radiation from a rotating anode (Vitali *et al.*, 1991). The conclusions of the present study, regarding data collection and processing and heavy-atom refinement and phasing remain valid, of course, for xenon derivatives as well.

There remain problems with both krypton and xenon. The absorption of X-rays and the fluorescence background from pressurized gas in the capillary are approximately proportional to the imaginary anomalous scattering factor (f'') of the gas atoms. Thus, these problems are inevitably linked to the exploitation of anomalous scattering effects. We have already described in this article a number of ways to overcome these limitations but we have also explored a far more radical approach to avoid these problems. The procedure consists of freezing krypton or xenon derivatives at cryogenic temperatures (113 K) and then releasing the gas pressure. Cryogenic techniques are now widely used in protein crystallography (for a recent review, see Rodgers, 1994), basically because they enable one to drastically reduce radiation damage to the crystal (Dewan & Tilton, 1987; Young, Dewan, Nave & Tilton, 1993). We performed a test experiment on HEL crystals,

mounted in a quartz capillary and put under high krypton gas pressure. The crystals were soaked in a cryoprotective mother liquor (Petsko, 1975) consisting of a mixture of 10%(*v*/*v*) glycerol and 90%(*v*/*v*) aqueous NaCl (1 *M*) solution. Under krypton gas pressure, crystals cannot normally be cryocooled to 113 K as the boiling point of krypton is situated at a higher temperature (119.7 K at 1×10^5 Pa and rapidly increasing with pressure). Initial attempts to flash-freeze crystals under krypton gas pressure led to the instantaneous condensation of the krypton gas in the capillary. We found two possible ways out of this quandary. The crystals can be frozen at gas pressures above the critical point of krypton (*i.e.* above 55.0×10^5 Pa). In this case, the phase transition is continuous, taking place on a timescale of tens of seconds, enough to flash-freeze the crystal and release the gas pressure. Another method is to release the pressure just before flash-freezing the crystal. Although the binding (and unbinding) of krypton is fast (see §2.2), the pressurization and freezing steps can be separated. We have been able to flash-freeze a HEL crystal at 113 K, less than one second after the krypton gas pressure (50×10^5 Pa) had been released. X-ray data were collected on this crystal and a Kr atom could be located in a difference-Fourier map.

Clearly, these initial studies must be followed by further investigations. In particular, the means to replace the method of cryo-cooling crystals in a capillary tube by the more efficient technique of flash-freezing crystals on a free-standing film (Teng, 1990) or any other method must be found. The potential advantages of frozen krypton or xenon derivatives are, however, considerable. Not only are there no longer problems with absorption of X-rays and fluorescence background (as there is no pressurized gas), but also one would benefit from the usual elimination of radiation damage in the crystal at low temperatures. In addition, hydrate formation (which typically develops slowly during data collection) can be avoided. The major drawback is that native and derivative data can no longer be collected on the same sample, thus annihilating one of the unique advantages of noble gas derivatives. This problem is aggravated by the fact that cryo-cooling often leads to unreproducible cell-parameter changes, thus being a source of possible non-isomorphism between native and derivative crystals. A solution would be to make use of the MAD method, which only requires one single crystal and is, therefore, essentially free from non-isomorphism problems.

5. Conclusions

The method of isomorphous replacement has played a central role in the development of protein crystallography. In spite of its successes, the effectiveness of the method has been limited by a number of factors which can be grouped into three distinct classes.

(i) Non-isomorphism between native and derivative structures has always been considered to be the major problem of MIR. This view is supported by the results obtained with the MAD method, where perfect isomorphism has allowed the determination of protein structures in spite of the much smaller anomalous and dispersive signal strengths (Hendrickson *et al.*, 1988).

(ii) The difficulties associated with the construction and refinement of the heavy-atom substitution model have long been a matter of debate. An optimal and definitive answer was proposed (Bricogne, 1991*a*,*b*) and implemented in the program *SHARP*, which has been shown to be superior to previous approaches (La Fortelle & Bricogne, 1997).

(iii) Finally, the quality of MIR phases may be limited by the inherent accuracy of the primary X-ray data. With careful data collection and processing schemes, experimental errors can be reduced significantly.

In the present study, it has been demonstrated that the use of highly isomorphous xenon and krypton derivatives and the use of statistical heavy-atom refinement constitute a powerful association of tools for macromolecular structure determination with the SIRAS method. With respect to MAD, this method has additional advantages: (i) the isomorphous signal is in general much stronger than the anomalous signal and (ii) experiments are much easier to perform as only two data collections at a fixed wavelength are required (with xenon, SIRAS experiments can even be carried out on a rotating anode with Cu $K\alpha$ radiation). It was stated by Schoenborn & Featherstone (1967) that '*. . . xenon is a little lighter than desirable for a heavy atom, but this is counteracted by the fact that xenon protein complexes show a high degree of isomorphism with the native crystals . . .*'. The results presented in this communication would suggest that the unique potential provided by almost perfectly isomorphous derivatives relegates considerations on the 'lightness' of the atom to the status of a minor disadvantage.

We thank C. Martinez, S. Longhi and C. Cambillau for a gift of *Fusarium solani pisi* Cutinase crystals and J. M. Dubuisson and J. Walocha for the machining of the pressure cell. Financial support was provided by a NATO grant (to MS).

References

Abrahams, J. P. & Leslie, A. G. W. (1996). *Acta Cryst.* D**52**, 30–42.

Abrahams, J. P., Lutter, R., Leslie, A. G. W. & Walker, J. E. (1995). *Nature (London)*, **370**, 621–628.

Arndt, U. W. & Wonacott, A. J. (1977). *The Rotation Method in Crystallography*. Amsterdam: North-Holland.

Baker, E. N., Anderson, B. F., Dobbs, A. J. & Dodson, E. J. (1995). *Acta Cryst.* D**51**, 282–289.

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977). *J. Mol. Biol.* **112**, 535–542.

Blow, D. M. & Crick, F. H. C. (1959). *Acta Cryst.* **12**, 794–802.

Bourguet, W., Ruff, M., Chambon, P., Gronemeyer, H. & Moras, D. (1995). *Nature (London)*, **375**, 377–382.
Braun, B. (1938). *Diss. Bonn.* pp. 10–21.
Bricogne, G. (1991*a*). *Crystallographic Computing* 5, edited by D. Moras, A. D. Podjarny & J. C. Thierry, pp. 257–297. Oxford: Clarendon Press.
Bricogne, G. (1991*b*). *Isomorphous Replacement and Anomalous Scattering. Proceedings of the CCP4 Study Weekend 25–26 January 1991*, edited by W. Wolf, P. R. Evans & A. G. W. Leslie, pp. 60–68. Warrington: Daresbury Laboratory.
Brünger, A. T. (1992). *Nature (London)*, **355**, 472–475.
Brünger, A. T. (1993). *Acta Cryst.* D**49**, 24–36.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Colloc'h, N., Bachet, B., L'Hermite, G., Mornon, J. P., Castro, B., El Hajji, M., Prangé, T. & Schiltz, M. (1997). In preparation.
Crick, F. H. C. & Magdoff, B. S. (1956). *Acta Cryst.* **9**, 901–908.
Dewan, J. C. & Tilton, R. F. (1987). *J. Appl. Cryst.* **20**, 130–132.
Evans, P. R. (1993). In *Proceedings of the CCP4 Study Weekend on Data Collection and Processing*, pp. 114–122. Warrington: Daresbury Laboratory.
Ewing, J. G. & Maestas, S. (1970). *J. Phys. Chem.* **74**, 2341–2344.
French, S. & Wilson, K. (1978). *Acta Cryst.* A**34**, 517–525.
Forcrand, M. de (1923). *C. R. Acad. Sci.* **176**, 355–358
Forcrand, M. de (1925). *C. R. Acad. Sci.* **181**, 15–17
Fourme, R., Dhez, P., Benoit, J.-P., Kahn, R., Dubuisson, J.-M., Besson, P. & Frouin, J. (1992). *Rev. Sci. Instrum.* **63**, 982–987.
Fox, G. C. & Holmes, K. C. (1966). *Acta Cryst.* **20**, 886–891.
Hamilton, W. C., Rollet, J. S. & Sparks, R. A. (1965). *Acta Cryst.* **18**, 129–130.
Helliwell, J. R., Ealick, S., Doing, P., Irving, T. & Szebenyi, M. (1993). *Acta Cryst.* D**49**, 120–128.
Hendrickson, W. A. (1985). *Crystallographic Computing* 3, edited by G. M. Sheldrick, C. Kruger & R. Goddard, pp. 277–285. Oxford University Press.
Hendrickson, W. A., Horton, J. R. & LeMaster, D. M. (1990). *EMBO J.* **9**, 1665–1672.
Hendrickson, W. A. & Lattman, E. E. (1970). *Acta Cryst.* B**26**, 136–143.
Hendrickson, W. A., Smith, J. L., Phizackerley, R. P. & Merritt, E. A. (1988). *Proteins Struct. Funct. Genet.* **4**, 77–88.
Kabsch, W. (1988). *J. Appl. Cryst.* **21**, 916–924.
Kahn, R., Fourme, R., Gadet, A., Janin, J. & André, D. (1982). *J. Appl. Cryst.* **15**, 330–337.
Konnert, J. H. & Hendrickson, W. A. (1980). *Acta Cryst.* A**36**, 344–350.
Kraut, J., Sieker, L. C., High, D. F. & Freer, S. T. (1962). *Proc. Natl Acad. Sci. USA*, **48**, 1417–1424.
Kurinov, I. V. & Harrison, R. W. (1995). *Acta Cryst.* D**51**, 98–109.
La Fortelle, E. de & Bricogne, G. (1997). *Methods in Enzymology. Macromolecular Crystallography*, edited by R. M. Sweet & C. W. Carter Jr. New-York: Academic Press. In the press.
Leslie, A. G. W. (1991). *Crystallographic Computing* 5, edited by D. Moras, A. D. Podjarny & J. C. Thierry, pp. 50–61. Oxford: Clarendon Press.
Leslie, A. G. W. (1992). *Jnt CCP4 ESF-EACMB Newslett. Protein Crystallogr.* p. 26.
Li de la Sierra, I. M., Pernot, L., Prangé, T., Saludjian, P., Schiltz, M., Fourme, R. & Padrón, G. (1997). Submitted.
Martinez, C., De Geus, P., Lauwereys, M., Mathyssens, G. & Cambillau, C. (1992). *Nature (London)*, **356**, 615–618.
Matthews, B. W. & Czerwinski, E. W. (1975). *Acta Cryst.* A**31**, 480–487.
Meyer, E., Cole, G., Radhakrishnan, R. & Epp, O. (1988). *Acta Cryst.* B**44**, 26–38.
Meyer, E. F., Radhakrishnan, R., Cole, G. M. & Presta, L. G. (1986). *J. Mol. Biol.* **189**, 553–559.
Miller, K. W., Reo, N. V., Uiterkamp, A. J. M. S., Stengle, D. P., Stengle, T. R. & Williamson, K. R. (1981). *Proc. Natl Acad. Sci. USA*, **78**, 4946–4949.
Ogata, C. M., Hendrickson, W. A., Gao, X. & Patel, D. J. (1989). *Abstr. Am. Crystallogr. Assoc. Ser.* 2, **17**, 53.
Otwinowski, Z. (1991). *Isomorphous Replacement and Anomalous Scattering. Proceedings of the CCP4 Study Weekend 25–26 January 1991*, edited by W. Wolf, P. R. Evans & A. G. W. Leslie, pp. 80–86. Warrington: Daresbury Laboratory.
Petsko, G. A. (1975). *J. Mol. Biol.* **96**, 381–392
Renaud, G., Villette, B., Boeuf, A., Chandesris, D. & Magnan, H. (1992). *LURE: Rapport général* 1989–1992, Vol. I, p. 134. Saclay: Edition SDEM CE (CNRS, CEA, MENC).
Renaud, G., Villette, B., Guenard, P., Ragonnet, D., Frouin, J., Chandesris, D. & Magnan, H. (1992). *LURE: Rapport général* 1989–1992, Vol. I, p. 135. Saclay: Edition SDEM CE (CNRS, CEA, MENC).
Rodgers, D. W. (1994). *Structure*, **2**, 1135–1140.
Rossmann, M. G., Leslie, A. G. W., Abdel-Meguid, S. S. & Tsukihara, T. (1979). *J. Appl. Cryst.* **12**, 570–581.
Schiltz, M., Fourme, R., Broutin, I. & Prangé, T. (1995). *Structure*, **3**, 309–316.
Schiltz, M., Prangé, T. & Fourme, R. (1994). *J. Appl. Cryst.* **27**, 950–960.
Schoenborn, B. P. (1969). *J. Mol. Biol.* **45**, 297–303
Schoenborn, B. P. & Featherstone, R. M. (1967). *Adv. Pharmacol.* **5**, 1–17.
Schoenborn, B. P. & Nobbs, C. L. (1966). *Mol. Pharmacol.* **2**, 491–498.
Schoenborn, B. P., Watson, H. C. & Kendrew, J. C. (1965). *Nature (London)*, **207**, 28–30.
Teng, T. (1990). *J. Appl. Cryst.* **23**, 387–391.
Tickle, I. J. (1991). *Isomorphous Replacement and Anomalous Scattering. Proceedings of the CCP4 Study Weekend 25–26 January 1991*, edited by W. Wolf, P. R. Evans & A. G. W. Leslie, pp. 87–95. Warrington: Daresbury Laboratory.
Tilton, R. F. (1988). *J. Appl. Cryst.* **21**, 4–9.
Tilton, R. F. & Kuntz, I. D. (1982). *Biochemistry*, **21**, 6850–6857.
Tilton, R. F., Kuntz, I. D. & Petsko, G. A. (1984). *Biochemistry*, **23**, 2849–2857.
Tilton, R. F., Singh, U. C., Weiner, S. J., Connolly, M. L., Kuntz, I. D., Kollmann, P. A., Max, N. & Case, D. A. (1986). *J. Mol. Biol.* **192**, 443–456.
Vitali, J., Robbins, A. H., Almo, S. C. & Tilton, R. F. (1991). *J. Appl. Cryst.* **24**, 931–935.
Wilson, A. J. C. (1949). *Acta Cryst.* **2**, 318–321.
Winkler, F. K., Schutt, C. E. & Harrison, S. C. (1979). *Acta Cryst.* A**35**, 901–911.
Young, A. C. M., Dewan, J. C., Nave, C. & Tilton, R. F. (1993). *J. Appl. Cryst.* **26**, 309–319.