An Improved Regularization Technique for Analysis of Anomalous X-ray Scattering Data; Platinum Uridine Blue Sulfate as an Example

R. Serimaa, a V. Eteläniemi, a Olli Serimaa, b T. Laitalainen c and A. Bienenstock d

aDepartment of Physics, POB 9, FIN-00014 University of Helsinki, Finland, bCenter for Scientific Computing, POB 405, FIN-02101 Espoo, Finland, cDepartment of Chemistry, POB 6, FIN-00014 University of Helsinki, Finland, and dStanford Synchrotron Radiation Laboratory, Slac Bin 69, PO Box 4349, Stanford, CA 94305, USA

(Received 25 September 1995; accepted 28 February 1996)

Abstract

A new iterative way to apply the Tikhonov regularization method for solving partial structure factors from anomalous X-ray scattering data is presented. It is used to obtain the regularized Pt–Pt partial structure factor in a complex amorphous material, platinum uridine blue, containing many atomic species, utilizing experimental anomalous wide-angle X-ray scattering data. Since only platinum has a suitable absorption edge for the measurement, the problem is very much ill-posed, and without regularization only unreliable solutions are obtained. The average Pt–Pt distances, determined from the Pt–Pt partial distribution function, are 3.0 and 5.7 Å and the coordination number at 3.0 Å is about 1. Accordingly, the dominating Pt-structures of Pt-uridine blue are dinuclear. The success of this method indicates that we may use anomalous scattering to obtain important, detailed information about the structure of very complex amorphous materials.

1. Introduction

Obtaining partial structure factors (PSFs) from anomalous X-ray scattering data is not an easy task because the scattering factors do not change much even close to adsorption edges. As a result, the mathematical equations are ill-conditioned and extremely sensitive to small errors of any sort. The problem is even more difficult if the sample contains only one element with a suitable absorption edge for the measurement. It is also quite difficult if the number of atomic species n in the material is more than two, since the number of independent PSFs is \(N = n(n + 1)/2\), making the equations even more sensitive to error. Often, however, the most important structural information can be obtained from the partial structure factor \(S_{xx}\), which corresponds to the single element x whose absorption edge is accessible.

In the following, we present a new iterative way to apply the Tikhonov regularization method (Tikhonov & Arsenin, 1977) to obtain the PSFs. The Tikhonov regularization has been applied earlier for analysis of anomalous X-ray scattering data by Licheri, Navarra & Seatzu (1990), Babanov et al. (1990) and Svergun (1991). The main difference between our procedures and these earlier procedures is that we include an initial estimate in the constraint (Tikhonov & Arsenin, 1977) to diminish the regularization error and present an iterative way to find it.

The regularization methods are well established methods to solve ill-posed problems and have been applied in various fields of physics (e.g. Glasko, 1988). The applications of Tikhonov regularization include, for instance, procedures for correcting instrumental broadening effects where the Fredholm integral equations of the first kind are solved (e.g. Paatero, Manninen & Paakkari, 1974; Glatter, 1977, Provencher, 1982; Serimaa, Serimaa, Paatero & Paakkari, 1984; te Riele, 1985; Svergun & Semenyuk, 1987; Trubin & Szaz, 1991). Another regularization method, the maximum-entropy method (Gull & Daniel, 1978), which also includes an initial estimate, has been applied in crystallography (Collins, 1982) and in small-angle scattering (Steenstrup & Hansen, 1994; Hansen & Pedersen, 1991; Müller & Hansen, 1994).

The regularization method is employed to yield a specific (Pt–Pt) partial structure factor in a complex amorphous material containing many different atomic species. This is the first time, to our knowledge, that the approach, which is analogous to the second difference method employed in neutron scattering (see, e.g., Neilson & Enderby, 1983) has been effectively implemented with anomalous wide-angle X-ray scattering (AWAXS). The success of these combined approaches indicates that we may use anomalous scattering to obtain important and detailed information about the structure of very complex amorphous materials. As will become apparent, it indicates that emphasis should be placed on the elimination of systematic errors in the experiment, since statistical errors can be treated with the regularization.

We apply the method to anomalous X-ray scattering data from Pt-uridine blue, which belongs to a class of...
amorphous platinum complexes known by the generic name platinum blues (Hofman & Bugge, 1908). Platinum blues are of interest owing to their antitumour activity and rich phenomenology (Rosenberg, 1969; Rosenberg et al. 1983; Davidson et al. 1975; Flynn, Viswanathan & Bruce Martin, 1977). The shortest Pt–Pt distance of 2.94 Å of Pt-uridine blue (Teo, Kijima & Bau, 1978) was determined for the first time by means of the extended X-ray absorption fine structure method (EXAFS). Results of our WAXS (Serimaa, Vahvasekkä, Paakkari, Laitalainen & Oksanen, 1993) and EXAFS (Eteläniemi, Serimaa, Paakkari & Laitalainen, 1996) studies of Pt-uridine blue sulfate synthesized by a newer method (Okuno, Tonosaki, Inoue, Yonemitsu & Sasaki, 1986; Serimaa et al., 1993) are comparable to those of Teo et al. (1978) and suggest Pt-structures analogous to crystalline models (Barton, Rabinowitz, Szalda & Lippard, 1987; O’Halloran, Mascharak, Williams, Roberts & Lippard, 1987; Matsumoto, Takahashi & Fuwa, 1984, and references therein). The length of the Pt-chain cannot be determined unambiguously from the total radial distribution function (RDF). The average coordination number at about 3 Å gives information on the length of the Pt chain, but the estimated Pt–Pt coordination number may be too large because of similar distances also between other elements. Moreover, EXAFS may give too small a Pt–Pt coordination number, since the distances between adjacent Pt atoms in the chain may vary and cause cancellation effects. Our aim was to obtain more precise information about the Pt-unit structure by solving the Pt–Pt partial structure factor. The differential anomalous scattering (DAS) technique (Shevchik, 1977; Fuoss, Eisenberg, Warburton & Bienenstock, 1981) was also applied for obtaining more accurate Pt–Pt distances by comparing the total and differential radial distribution functions (RDF, DDF).

2. Experimental

2.1. Synthesis

The synthesis of platinum uridine blue was carried out basically by the procedure given by Okuno et al. (1986) with slight modifications. Since, unlike Okuno, we did not observe the formation of Pt-red or Pt-green species during the isolation, the experiment is described in detail. The reagents were used as obtained.

The 0.1 M cis-diamminediaquaplatinum(II) sulfate, [Pt(NH$_3$)$_2$(H$_2$O)$_2$]SO$_4$ (pH = 3.2) was prepared from K$_2$PtCl$_4$ (Dhara, 1970) (Aldrich) via PtI$_2$(NH$_3$)$_2$ in a reaction with Ag$_2$SO$_4$ as described below. The reaction mixture of platinum uridine blue was gel filtered [Merck, Fractogel TSK HW-50(S)] by eluting with 0.012 M H$_2$SO$_4$. The product was stored in a brown desiccator over P$_2$O$_5$. Microanalysis is by Dornis und Kolbe, Mikroanalytisches Laboratorium, Germany. Visible (VIS) spectra were taken from aqueous solutions by means of a Schimadzu UV-200 double-beam spectrometer and infrared (IR) spectra with a Biorad FTS-7 Fourier-transform infrared spectrometer on KBr pellets pressed at 0.9 kN m$^{-2}$.

cis-Diamminediaquaplatinum(II) sulfate, [Pt(NH$_3$)$_2$(H$_2$O)$_2$]SO$_4$. A mixture of K$_2$PtCl$_4$ (5.00 g, 0.012 M) and KI (8.05 g, 0.0485 M) dissolved in 50 ml of deionized water by ultrasound was warmed in a brown flask at 313 K in an oil bath and stirred for 20 min. Into the warm solution 6.0 ml of aqueous NH$_3$ (14%) was added during 10 min and the mixture further vigorously stirred at 313 K for 15 min. The mixture was then kept in an ice bath for 30 min and the product was filtered off. The product was washed with hot water, cold ethanol and cold diethyl ether and dried and dried under vacuum (1.33 Pa) at 12 h to give 5.37 g (92% of theory) of PtI$_2$(NH$_3$)$_2$. A mixture of PtI$_2$(NH$_3$)$_2$ (5.3290 g, 0.011 M) and Ag$_2$SO$_4$ (3.4403 g, 0.011 M) in 110 ml of deionized water was stirred vigorously under light protection at room temperature for 5 h. Then, the mixture was cooled in an ice bath for 1 h and the precipitate was filtered off with a sintered glass funnel (G4) to give a faintly yellowish solution. This solution (pH = 3.2) was used immediately for the synthesis of platinum uridine blue.

Platinum uridine blue. A mixture of [Pt(NH$_3$)$_2$(H$_2$O)$_2$]SO$_4$ (72 ml, 0.0072 M) and uridine (1.7582 g, 0.0072 M) was heated in an oil-bath at 340–343 K in a transparent flask provided with a reflux condenser for 16 h and then stirred at room temperature for 7 h. The deep blue mixture was gel filtered, and the blue zone, was taken directly into a twentyfold amount of acetone. Precipitation at 281 K for 12 h gave a loose and lumpy product. The product was filtered with a sintered glass funnel (G4) and dried under vacuum (1.33 Pa) at 1 h and then stirred at room temperature for 7 h. The deep blue mixture was gel filtered, and the blue zone, was taken directly into a twentyfold amount of acetone. Precipitation at 281 K for 12 h gave a loose and lumpy product. The product was filtered with a sintered glass funnel (G4) and dried under vacuum (1.33 Pa) for 10 h to give 2.94 g (73.7% of theory when based on Pt) of a blue powder. The spectra were taken on the product described here. The sample for the elemental analysis was synthesized in exactly the same way and the results are given here. Elemental analysis: found [values in parentheses from Okuno et al. (1986)]: C, 18.90 (20.55); H, 3.53 (3.02); N, 9.20 (8.20); S, 3.88 (2.64); Pt, 35.22 (37.96). VIS(H$_2$O): 572 nm (broad, asymmetric); IR(KBr): 1638 cm$^{-1}$ (C = O).

2.2. WAXS measurements

The AWAXS measurements were carried out at the Stanford Synchrotron Radiation Laboratory on the wiggler beamline 4-3 utilizing an Si(111) double-crystal incident-beam monochromator and a two-circle diffractometer in symmetric reflection mode. The sample was a 3 mm thick pressed pellet with an area of 10 × 40 mm. The parasitic scattering was reduced using He paths but the sample itself was at normal atmosphere. The scattered intensity was measured in step-scan mode with a high-purity germanium solid-state detector. The primary beam
intensity was monitored by a scintillation crystal (NaI) and a photomultiplier detector. The data were measured in three parts, with the accumulation time increased to collect at least 100,000 counts at each point. The range of \( k \) was 0.3-10 \( \text{Å}^{-1} \) and the steps were 0.05 \( \text{Å}^{-1} \). (\( k = (4\pi/\lambda) \sin \theta \) is the magnitude of the scattering vector, \( \theta \) the scattering angle and \( \lambda \) the wavelength). Each scan took about 7 h. The employed photon energies of 10.6, 11.2, 11.4 and 11.5 keV were below the Pt L\( _{III} \) absorption edge (11.564 keV). One measurement was performed without a sample to take parasitic scattering into account. The air scattering was quite prominent at \( k < 2 \text{ Å}^{-1} \) but diminished rapidly at larger \( k \).

The experimental intensity curves were corrected for absorption and normalized to absolute scale with the large angle method. Normalization involved determination and subtraction of fluorescence in the data and subtraction of the calculated Compton intensity (Krogh-Moe, 1956; Norman, 1957; Fuoss et al., 1981). The energy-independent atomic form factors and incoherent scattering functions were taken from International Tables for X-ray Crystallography (1974). Anomalous scattering factors (Table 1), which were evaluated from experimental X-ray absorption coefficients using the Kramers-Kronig dispersion relation, were used in the data analysis.

### 3. Theory

#### 3.1. Total, differential and partial structure factors

Let us consider an amorphous material consisting of several types of atoms. The atomic fraction of the component \( \alpha \) is denoted by \( w_\alpha \), and its scattering factor is \( f_\alpha(E, k) = f_0\alpha(k) + f_\alpha(E) + if''_\alpha(E) \), where the functions \( f' \) and \( f'' \) are its anomalous scattering factors. The structure of the amorphous material is described in terms of partial radial distribution functions (PDRFs) \( P_{\alpha\beta}(r) \).

The partial radial distribution functions are defined by \( P_{\alpha\beta}(r) = 4\pi r^2 \rho_{\alpha\beta}(r) \), where \( \rho_{\alpha\beta}(r) \) represents the average atomic density of \( \beta \)-type atoms at the distance of \( r \) from an \( \alpha \)-type atom. By definition, \( P_{\alpha\beta} \) is related to the partial structure factor \( S_{\alpha\beta}(k) \) by the Fourier transform:

\[
P_{\alpha\beta}(r) = 4\pi r^2 \rho_{\alpha\beta}(r) = 4\pi r^2 w_\beta \rho_0 + (2r/\pi) \int_0^\infty k S_{\alpha\beta}(k) \sin kr \, dk,
\]

where \( \rho_0 \) is the average atomic density.

The coherently scattered intensity \( I(k) \) can be written in terms of the partial structure factors \( S_{\alpha\beta}(k) \) as follows (Keating, 1963):

\[
I(E, k) = \sum_\alpha w_\alpha f_\alpha^*(E, k) f_\alpha(E, k) + \sum_{\alpha,\beta} w_\alpha f_{\alpha\alpha}^*(E, k) f_{\beta\beta}(E, k) S_{\alpha\beta}(k).
\]

#### 3.1. Total, differential and partial structure factors

The photon energies \( E_i, i = 1, 2, \ldots \), are chosen close to the absorption edge of one of the elements, \( \alpha \), but it is favourable that the energies \( E_i \) are far away from absorption edges of the other elements. Hence, the scattering factor \( f_\alpha(E, k) \) of the chosen element \( \alpha \) chan-
ges considerably more than the other scattering factors, and $\Delta I(E_i, E_j, k)$ is approximately a weighted sum of only those partial structure factors $S_{\alpha \beta}(k)$ that involve the chosen element.

The differential distribution function $DDF(r)$, which is calculated as a Fourier transform of the differential structure factor

$$\text{DSF}_r(k) = [(I(E_i, k) - I(E_j, k) - (f(E_i, k)^2) - (f(E_j, k)^2)]/[I(f(E_i, k)^2) - (f(E_j, k)^2)],$$

shows the average environment of the chosen element.

For a two-component system, the weights of the PDFs in the total RDF and DDF are approximately (Ludwig, Warburton, Wilson & Bienenstock, 1987)

$$\text{RDF}(r) \approx \omega_1 K_1^2(E) P_{a}(r) + 2\omega_1^2 K_1^2(E) P_{b}(r) + \omega_2 K_2^2(E) P_{a}(r),$$

$$\text{DDF}(r) \approx 2\omega_1 \text{Re}(K_1(E_1 - K_2(E_2)) P_{a}(r)) \times [(K_1(E_1)^2 - (K_2(E_2))^2)^{-1}.$$

All the partial structure factors may be solved from the set of equations (2), but if the number of atomic species in the sample is large, the number of independent intensity curves must be large as well.

However, one of the PSFs of a multicomponent system can be solved from a set of only a few intensity differences. Then, the minimum number of independent intensity curves (measured at different energies) is three.

The partial structure factor $[S_{2a}(k)]$ is solved from a set of these intensity differences presented in the following form:

$$\Delta I(E_i, E_j, k) = [f_a(E_i, k)^2] - [f_a(E_j, k)^2]w_2 S_{2a}(k) + [f_b(E_i, k)^2]w_2 S_{2b}(k) \times \sum_{\beta \neq a} [f_\beta(k) + f_\beta^*(k)]w_2 S_{\beta \beta}(k).$$

Then, the following (unregularized) solution is obtained from two intensity differences $\Delta I(E_1, E_2, k)$ and $\Delta I(E_1, E_3, k)$ by elimination of the common sum

$$w_2 S_{2a}(k) \approx [f_a^2(E_2) - f_a^2(E_3)]^{-1} \times [(\Delta I(E_1, E_2, k)/[f_a(E_1) - f_a(E_2)] - \Delta I(E_1, E_3, k)/[f_a(E_1) - f_a(E_3))].$$

This solution is analogous to that employed earlier in neutron scattering (second difference method; Neilson & Enderby, 1983).

3.2. Linear equation

For each fixed $k$, (2) and (10) are both linear equations from which the PSFs can be solved. The matrix form of the equations is $y = A(k)x(k)$. The matrix $A(k)$ is the unknown coefficient matrix, of size $M \times N$, which depends on the scattering factors $f_a(E, k)$, atomic fractions $w_a$, and the magnitude of the scattering vector $k$. The column matrix $y(k)$, whose components are indexed by the energies $E_i, i = 1, 2, \ldots, M$, is the data of the equations; in this case it is the experimental intensities from which the term $(f^2)$ is subtracted. In the case of a two-component system, the number of unknowns $N$ in (2) is three and the column matrix $x(k)$ represents the PSFs to be determined: $x_1(k) = S_{a2}(k), x_2(k) = S_{a b}(k)$ and $x_3(k) = S_{b b}(k)$.

In a more abstract setting, the matrices in the equation $y = Ax$ can be interpreted as operators and vectors: $A$ is a known linear operator from a vector space $V$ of dimension $N$ to a space $W$ of dimension $M$, $y \in W$ is an (unprecisely) known vector in the target space and $x \in V$ is the solution vector.

Because the scattering factors depend weakly on the photon energy $E$, the rows representing the equations are almost identical (i.e. the matrix $A$ is nearly singular).

Hence, the problem of solving the PSFs is unfortunately very much ill-posed, and statistical inaccuracy and systematic errors in the data (and to a certain extent also the inevitable systematic regularization bias) severely reduce the goodness of the solution.

3.3. Random variable

Because of the errors inherent in the measurements, we furthermore consider the vector $y$ as a vector-valued random variable. As a result, the solution $x$ will be a vector-valued random variable as well. Note that for simplicity we consider here the matrix $A$ as exactly known (nonrandom).

In the following theoretical considerations, the (unknown) expectation value of the variable $y$ is denoted by $y_T = Ey$. Its Hermitian and positive covariance matrix is denoted by $\text{cov}(y)$ and estimated assuming a Poisson distribution for $y$. Its diagonal elements, the variances of the components $y_i$, are denoted by $\text{var}(y_i)$. We also assume that there is a unique exact solution $x_T$ (the unknown 'true solution') satisfying the equation $y_T = Ax_T$ exactly.

Note that the statistical parameters $y_T$ and $\text{cov}(y)$ and the distribution of the random variable $y$ are unknown. The only thing we know is the experimental data, which can be regarded as one realization of the random variable $y$.

3.4. Regularization method

A possibly ill-posed set of linear equations $y = Ax$ can be solved approximately using the Tikhonov regulariza-
Here, a quadratic functional of the form
\[ \chi^2(x) = \|y - Ax\|_S^2 + \|x - x_M\|_B^2 \] (12)
\[ = \|(y - y_M) - A(x - x_M)\|_S^2 + \lambda \|x - x_M\|_B^2 \] (13)
is used, where \( \lambda \) is the regularization factor, \( x_M \) is the centre vector (initial estimate), \( y_M = Ax_M \) is the image vector corresponding to the centre vector, \( B \) is the regularizing term (or matrix) and \( S \) is the weighting matrix. These parameters defining the regularization procedure are to be chosen carefully to obtain good approximative solutions.

Here, the notation \( \|z\|_C^2 \) means the norm of the vector \( z \) with respect to the Hermitian positive operator \( C = D^*D \), the so-called \( C \) norm:
\[ \|z\|_C^2 = (z^*Cz) = \|Dz\|^2 = z^*Cz = z^*D^*Dz, \]
where the last two forms are in the matrix notation. Note that any positive Hermitian operator \( C \) can always be presented in a (nonunique) form \( C = D^*D \) (Cholesky decomposition).

The use of the centre vector \( x_M \) in the regularized functional is emphasized, since it was found to improve considerably the results. Below, an iterative technique to find a proper centre vector will be applied.

The regularization term \( \|x - x_M\|_B^2 \) including the centre vector represents a constraint to minimize the difference of the solution \( x \) from \( x_M \). The closer \( x_M \) is to the true solution \( x_T \), the smaller the effect of this regularizing term \( \|x - x_M\|_B^2 \) will be. Hence, the idea of the centre vector \( x_M \) is to be an initial guess for the solution, i.e., to choose a known vector \( x_M \) (loosely) approximating the true solution \( x_T \). If no centre vector is used, a too 'flattened' solution may be obtained, since the usually used regularization term \( \|x\|_S^2 \), without the centre vector, is a constraint to minimize the norm of the solution vector.

The usual choice for the weighting matrix \( S \) defining the \( y \) norm is essentially the inverse of the covariance matrix of the data \( y \):
\[ S = (1/M) \text{cov}(y)^{-1}. \]
The dimension \( M = \dim W \) of the target space (number of equations) is introduced to the operator \( S \) for normalization. With this choice for \( S \), the expectation value of the square of the norm of the difference \( y - y_T \) is exactly one, \( E(\|y - y_T\|_S^2) = 1 \). This choice for the weighting matrix \( S \) ensures that the optimal residual \( \|y - Ax\|_S^2 \) is on average about 1 (Tikhonov, 1977).

The regularization matrix \( B \) can be interpreted to correspond to smoothness constraints being imposed on the regularized solution. There are two possible cases here. In the first one regularizes the 'small' set of equations \( y(k) = A(k)x(k) \) with \( M \) equations for \( N \) unknowns for each fixed value of the scattering vector \( k \) separately (separate case). In this case, the constraints do not couple different \( k \) values. This can lead to a solution \( x(k) \) that is not behaving smoothly as a function of the variable \( k \). Alternatively, one could choose the norms so that they do couple the different equations for different values \( k_i, i = 1, 2, \ldots, K \) of the scattering vector (coupled case). These norms are function norms with respect to the variable \( k \). In this case, all the equations have to be solved together, and we have one 'huge' regularized equation \( Y = AX \) with \( MK \) equations for \( NK \) unknowns. Here, the matrix \( A \) is now a large block-diagonal \( MK \times NK \) matrix with the small \( M \times N \) matrices \( A(k_i) \) on its diagonal. The matrix \( B \) is now an \( NK \times NK \) matrix that, in general, is not block diagonal. In this case, the regularized solution tends to be a smoother function of the scattering vector \( k \).

Technically the separate case can be formulated as a coupled case with a block-diagonal \( B \) matrix, with \( N \times N \) blocks on its diagonal. Then, the large regularized equation separates into many small equations that are regularized versions of the equations \( y(k) = A(k)x(k) \).

These types of constraints have been imposed to control the norm of the solution vector \( x \) and its first and second derivatives with respect to \( k \) (coupled case; Babanov et al., 1990). Similarly, in the separate case, one can control the norm of the solution vector \( x \) and its first and second 'derivatives' or differences with respect to its component indices (Licheri et al., 1990), even though in this case the term derivative has no meaning. Here, a combination of both types of constraints are used. The regularization operator employed here is described in detail in the Appendix.

The regularization factor \( \lambda > 0 \) corresponds to a Lagrange multiplier and controls the 'strength' of the regularization. For more universality it is scaled as follows:
\[ \lambda = \tilde{\lambda}(\|A\|_S^2/\|B\|). \] (14)
The operator norm \( \|A\|_S \) of \( A \) with respect to \( S \) can be calculated as the square root of the usual operator norm of the operator \( A^*SA \): \( \|A\|_S^2 = \|A^*SA\| \). This scaling reduces the \( k \) dependence of the regularization parameter, in the separate case where the set of equations is solved independently at each \( k \).

The solution depends on the form of \( B \) and on the parameters of the regularization. This will be discussed in more detail below.

3.5. Regularized solution

The minimum point(s) satisfy the normal equation

[\text{Euler–Lagrange equation of the functional } \chi^2(x)]

\[ (A^*SA + \lambda B)(x - x_M) = W_S(x - x_M) = A^*S(y - y_M). \] (15)
For positive $\lambda > 0$, the operator $W_\lambda = A^\dagger SA + \lambda B$ is a Hermitian, strictly positive and invertible operator, which ensures that the regularized normal equation has the unique solution $x_\lambda$ given by

$$x_\lambda = x_M + G_\lambda (y - y_M) = G_\lambda y + x_M - G_\lambda Ax_M = G_\lambda y + P_\lambda x_M.$$  

(16)

The operators

$$G_\lambda = W_\lambda^{-1} A^\dagger S = (A^\dagger SA + \lambda B)^{-1} A^\dagger S,$$

$$P_\lambda = id_y - G_\lambda A = W_\lambda^{-1} \lambda B$$

and

$$R_\lambda = id_w - AG_\lambda$$

are called the resolvent operator, the regularization bias operator [for motivation, see (18)] and the residual operator, respectively. The resolvent operator $G_\lambda$ is a regularized pseudoinverse of the operator $A$. The minimum value of the functional $\chi^2$ is

$$\min [\chi^2(x)] = \chi^2(x_\lambda) = \langle y - y_M | SR_\lambda (y - y_M) \rangle.$$  

Both the regularization and the statistical errors of the data $y$ affect the solution $x_\lambda$. Because the solution $x_\lambda$ depends linearly on $y$, the covariance matrix of $x$ can be written as a function of the covariance matrix of $y$ and the resolvent operator $G$ as follows:

$$\text{cov} (x_\lambda) = G_\lambda \text{cov} (y) G_\lambda^\dagger = (id_y - P_\lambda) W_\lambda^{-1} = W_\lambda^{-1} - W_\lambda^{-1} (\lambda B) W_\lambda^{-1}.$$  

(17)

The expectation value of $x_\lambda$ is

$$E x_\lambda = G_\lambda E y + P_\lambda x_M = G_\lambda Ax_T + P_\lambda x_M = x_T + P_\lambda (x_M - x_T).$$  

(18)

where the quantity $P_\lambda (x_M - x_T)$ represents the regularization bias, i.e. the regularized solution is a biased estimate for the true solution $x_T$. The regularization bias depends of course on the choice of the regularization matrix $B$ and the regularization parameter $\lambda$. Since the true solution $x_T$ is unknown, the regularization bias can be estimated reliably only for simulated data.

Several criteria for choosing the optimum regularization parameter have been proposed (e.g. Tikhonov, 1977; Glatter, 1977; Svergun, 1991). To find a suitable regularization parameter $\lambda$, the solutions $x_\lambda$, the regularization bias and statistical errors, the values of the residual norm $\| y - Ax_\lambda \|_S$ and the stability $\| x_\lambda \| / \| \lambda \| \| x_\lambda \| = \| P_\lambda \| / \| \lambda \|$ (the prime denotes the derivative with respect to $\lambda$), were examined. In the vicinity of the optimal regularization parameter, the residual norm is close to 1 (Tikhonov, 1977).

The determination of the regularization parameter is not straightforward when the errors in the data are not known (Bertero & Pike, 1992; Svergun, 1992). In the measurements of anomalous scattering, the information is usually available. Good statistical precision is essential for the success of the data analysis and the experimentalist should estimate the statistical precision of the data already during the experiment when choosing the measuring strategy.

4. Simulated partial structure factors

The Pt-complex is considered as a two-component system of platinum (denoted below by component $x$) and other elements hydrogen, carbon, nitrogen, oxygen and sulfur lumped together (component $\beta$). For our data, the condition number of the original coefficient matrix $A$ varies between 10 000 and 1000 as a function of the magnitude of the scattering vector $k$. The singular values of the original coefficient matrix $A$ at $k = 0.3 \text{Å}^{-1}$, for example, are 0.068, 0.001 and 0.00001, which indicates that one of the partial structure factors dominates the scattering.

Calculations with simulated data with several regularization operators and parameters gave insight on the quality of the regularized solution. In all these calculations, the data $y$ included statistical errors but the coefficient matrix $A$ was exact. (In practice, also the matrix $A$ is not known accurately since there are potential errors in $f$ and $f''$.) In most simulations, the data matrix $y$ consisted of simulated intensity curves of an equidistant linear four-atom Pt chain of average spacing 3 Å. Each platinum was fourfold-coordinated by O atoms at 2 Å. In addition, in order to have the same atomic density of Pt as in the sample studied, 124 randomly distributed O atoms were included with the minimum mutual distance of 1.5 Å. The statistical precision of the simulated data was about same as that of the experimental data.

Because of the ill-posedness of the coefficient matrix $A$, the statistical errors of the data $y$ may contaminate significantly the unregularized solution, giving rise to kind of 'mirror' effects. These have been previously discussed by Ludwig et al. (1987). The statistical precision of the regularized solution is much better than that of the unregularized solution. However, the regularized solution will also be somewhat biased even when $y$ is exact. By choosing the centre vector properly, one can try to minimize the regularization bias.

The simulations were carried out for six different regularization schemes, i.e. for the representative choices 1–6 of the regularization operator $B$ described below. The regularized solution $x_\lambda$ depends to some extent on the regularization operator $B$. The regularization parameter $\lambda = 0.0001$ was found to be in the optimum range using the residual and stability criteria for all $B$ matrices (1–6).

The tested regularization matrices $B$ in the separate case represented (see the Appendix)

(i) the norm of the solution vector $x_\lambda$,

(ii) the first differences of $x_\lambda$ and

(iii) the second differences of $x_\lambda$ with respect to its component indices. In the coupled case, the matrices $B$ represented
(iv) the norm of the solution vector \( x_2 \),
(v) the first derivatives of \( x_2 \), and
(vi) the second derivatives of \( x_2 \) 
with respect to \( k \).

Note that in cases (ii), (iii), (v) and (vi) the matrix \( B \) is singular. The advantage of the \( k \)-dependent regularization matrices (iv)--(vi) was that they produced smoother solutions with respect to \( k \) than the \( k \)-independent ones (i)--(iii). The obvious drawback is that in the cases (iv)--(vi) one has to solve one large linear system of equations instead of several small ones needed in cases (i)--(iii).

4.1. Regularization bias with different regularization matrices

First, the regularization errors caused by the matrices (i)--(vi) were compared. No centre vector \( x_M \) was used in these calculations. In all cases (i)--(vi) the bias operator \( P_3 \) differs from the unity matrix. Then, all components of the expectation value of the obtained solution \( E x_2 = G_{ij} A x_T = x_T - P_3 x_T \) will be linear combinations of the components of the true solution \( x_T \).

The regularization matrices (i) and (iv) are constraints to minimize the norm of \( x \) and have a tendency to produce too 'flat' solutions when \( \lambda \) is too large. The derivative regularization matrices (ii) and (iii) do not restrict the mean value of solution, but (ii) has a tendency to bring the components of the solution vector equal and (iii) to linear combinations of each other: \( x_1 - 2x_2 + x_3 = 0 \). Similarly, the regularization matrix (v) tends to bring all the components of \( x \) as constants with respect to \( k \) and (vi) to linear functions of \( k \).

When the regularization parameter \( \lambda \) was in the optimum range, good solutions \( x_{i1} \) were obtained with any of the regularization matrices (i)--(iv). The regularization bias \( P_3 (x_M - x_T) \) was smallest when the regularization matrix (iii) was used, but the differences were small. The matrices (v) and (vi) (and their linear combinations) gave considerably more biased solutions than the matrices (i)--(iv). The other two components of the solution vector \( x_{i2} \) and \( x_{i3} \) were too biased with all the matrices (i)--(vi).

The regularization bias increased with \( \lambda \) when the regularization matrices (i), (ii), (iv), (vi) were used. That the second derivative regularization matrix (iii) was an exception is a special feature of this particular problem: with this singular \( B \) the true solution \( x_T \) solves also quite well the equation \( B x = 0 \). For this reason, this matrix (iii) did not give as smooth a solution \( x \) with respect to \( k \) as the other regularization matrices.

Good solutions were obtained also with various linear combinations of matrices (i)--(vi). The inclusion of (i) or (iv) in \( B \) is preferable since it ensures that \( B \) is invertible. The regularization matrix \( B \) that was finally chosen for analysis of experimental data included the norm of \( x - x_M \) with respect to \( k \) and the second derivative with respect to the component indices and the scattering vector \( k \). The obtained solution was reasonably smooth and the share of \( x_{T1} \) in the solved \( x_{i1} \) was about 0.8. This indicates that, when no initial estimate is used, the Pt–Pt PDF has maxima in correct positions, but the coordination numbers are roughly 20% smaller than the correct values.

4.2. The role of the centre vector

A proper centre vector was found iteratively. At first, all the components of the centre vector \( x_M \) were set to 0 to obtain an approximation to its first component \( x_{M1} \). Since the first component of the solution vector, PSF \( S_{az} \), is the most interesting one, we did not want to use any \textit{a priori} information to obtain it. The calculated \( x_{i1} \) was used as the first component of the centre vector \( x_{M1} \) at the next iteration step. At subsequent iteration steps, the new obtained \( x_{i1} = S_{az} \) was always used as the first component of the centre vector \( x_{M1} \). After three iterations, the result did not change any more.

In this case, the PSFs \( S_{gb} \) and \( S_{bb} \) of a unit containing one Pt atom and four O atoms at the distance of 2 Å from the Pt atom are used as the centre-vector components \( x_{M2} \) and \( x_{M3} \). The centre-vector component \( x_{M2} \) may also be iterated, but the result for the component \( x_2 = S_{gb} \) is only slightly better than that obtained by iterating only \( x_{M1} \). Even dividing the centre vector \( x_{M2} \) by two, which means reducing the Pt–O coordination number from four to two, gives nearly identical results, also. However, a too large centre vector may give erroneous results.

Since the component \( x_3 = S_{bb} \) is small compared to the other two components of \( x \), it was not possible to obtain any reliable solution for the component \( x_3 \) irrespective of the choice for the centre-vector component \( x_{M3} \), and hence the choice \( x_{M3} = 0 \) was used. For

![Fig. 1. The partial structure factor \( S_{az} \) of a four-atom Pt chain as a function of \( k (\text{Å}^{-1}) \). The iterated \( S_{az} \) is denoted by filled circles and the exact PSF by the solid line. The regularized \( S_{az} \) solved without a centre vector is denoted by open circles.](image-url)
$x_3 = S_{BB}$, only a mirror image of the strongest component, $x_1 = S_{22}$, is obtained.

Fig. 1 presents the exact $S_{22}$ and solutions that were obtained with and without a centre vector. It shows that, with a properly chosen centre vector, an improved solution with less regularization bias is obtained. The improvement is considerable at small k (especially in the small-angle scattering range not shown in the figure). The Pt–Pt coordination number obtained from the $S_{22}$ of the third step was correct with the accuracy of about 1%.

Fig. 2 presents the exact $S_{AB}$ and solutions obtained with and without a centre vector. The iterated $S_{AB}$ is better than that calculated without a centre vector, but both of them are considerably less accurate than $S_{22}$. The first maxima of the PDF $x_2 = S_{AB}$ are in correct positions but the precision of the coordination numbers is not good enough.

In Fig. 3, the difference of the exact and solved component $x_1 = S_{22}$ is compared with the standard deviation of the regularized solution $\sqrt{\text{var}(x)}$ and with the approximate regularization bias $P_A(x_M - x_T)$. The statistical errors of the solution $x_{11}$ are reasonably small. In the case of $x_{23}$, the statistical errors remain larger than the function values. Fig. 3 shows that the estimation of the precision of the result in the case of experimental data, when the true solution is unknown, is a problem. The use of $x_I$ instead of $x_T$ does not give a reliable estimate for the regularization bias $P_A(x_M - x_T)$. Computer simulations are thus very helpful for clarifying the effects of regularization bias on the solution.

5. Experimental results

The experimental intensity curve shown in Fig. 4 has broad maxima at about 0.5, 1.4, and 2.6 Å$^{-1}$. First, the total (TSF) and the differential (DSF) structure factors at different energies were calculated (Fig. 5) and from them the total and differential RDFs (Fig. 6). Then, the experimental Pt–Pt partial structure factor was calculated iteratively in the same way as the simulated one. The centre-vector components $x_{M1}$ and $x_{M3}$ for $S_{22}$ and $S_{BB}$, respectively, were set equal to zero, while $x_{M2}$ for $S_{AB}$ was a simulated PSF $S_{AB}$ of a unit containing one Pt atom, and four O atoms at the distance of 2 Å from the platinum, on the basis of the interpretation of the first maximum of the RDF (Serimaa et al., 1993). Reducing $x_{M2}$ by a factor of two did not change the result.

The Pt–Pt partial structure factor, which was solved from the intensity differences only [(10)], was essentially the same as that obtained from the original set of equations (2). The intensity differences were also used together with the intensities when the PSFs were solved (Munro, 1982; Fuoss et al., 1981). In the cases of both experimental and simulated data (not containing systematic errors), this set of equations gave essentially the same PSFs as the intensities alone.

The maxima arising from the Pt–Pt distances can be distinguished by comparison of the functions $\Delta PDF = P_{AB} - 4\pi r^2 w_k P_0$, $\Delta RDF$ and $\Delta DDF$. The first maximum of the RDF at 1.4 Å arises from C–C, C–N, C–O and N–O distances, but the features of the DDF and PDF below 2 Å are ‘Fourier ripples’. The maximum at 2 Å, which is not present in the $\Delta PDF$, arises mainly from Pt–N and Pt–O distances. The $\Delta PDF$ shows one broad maximum at 3 and another at 5–6 Å, both of which are present in the $\Delta RDF$ and $\Delta DDF$, too. Since they are more intense in the $\Delta DDF$ than in the $\Delta RDF$, it is concluded that they arise from the Pt–Pt distances. The maximum that exists in the $\Delta RDF$ at 4.3 Å but not in the $\Delta DDF$ and $\Delta PDF$ cannot arise from a Pt–Pt distance. Finally, the maximum at about 8 Å, which is present in the $\Delta RDF$ and $\Delta DDF$, may arise also from a Pt–Pt dis-
Fig. 4. Experimental intensity curve (solid line) with the coherent independent (open circles) and Compton intensities (filled circles) as a function of $k$ (Å$^{-1}$).

Fig. 5. Experimental total (solid line) and differential structure factors (open circles) and Pt–Pt partial structure factor (filled circles) as a function of $k$ (Å$^{-1}$). For comparison, the differential structure factor was divided by 2 and the partial structure factor by 6.

Fig. 6. The $\Delta$RDFs of Pt-uridine blue sulfate obtained with 11.5 keV (thick line) and 17.4 keV (thin line) and the differential distribution obtained with 11.2 and 11.5 keV (filled circles).

6. Discussion

The measurements were made using synchrotron light, whereas the previous measurements (Serimaa et al., 1993) were made using a conventional X-ray source (Mo $K_{\alpha}$, 17.4 keV). The statistical precision of the present data (10.6–11.5 keV) is better than in the previous study. In the previous work, the background scattering was reduced using a vacuum path, whereas now the sample is at normal atmosphere. The TSFs from the present and the previous study are in good agreement. We believe that the minor differences at small $k$ arise from the differences in the accuracies of the data, rather than from the fact that the samples were different, since the two samples were synthesized in the same way. The RDFs from the present and the previous study are compared in Fig. 6. The better spatial resolution of the RDF obtained with 17.4 keV is due to the larger $k$ range available. The most important difference is that in the RDF at 17.4 keV there are separate maxima at 2, 2.6 and 3 Å and in the RDF at 11.5 keV only two broad maxima at about 2 and 3 Å.
Table 2. The positions of the most prominent maxima (in Å) of the ΔRDF and ΔDDF compared with the estimated Pt–Pt distances of 2–9 Å determined from the Pt–Pt PDF and the average coordination numbers at about 2 and 3 Å

<table>
<thead>
<tr>
<th></th>
<th>Distances (Å)</th>
<th>Coordination number at about</th>
</tr>
</thead>
<tbody>
<tr>
<td>RDF</td>
<td>2.1</td>
<td>4.5</td>
</tr>
<tr>
<td>RDF(a)</td>
<td>2.08</td>
<td>3.8</td>
</tr>
<tr>
<td>DDF</td>
<td>2.0</td>
<td>5.2</td>
</tr>
<tr>
<td>PDF</td>
<td>3.03</td>
<td>1</td>
</tr>
</tbody>
</table>

Reference: (a) Serimaa et al. (1993).

In the previous study (Serimaa et al., 1993), it was concluded on the basis of the total RDF that the nuclearity of Pt chains in Pt-uridine blue does not exceed four. The Pt–Pt PDF determined in this work indicates that the dominating Pt structures in Pt-uridine blue are dinuclear. The maximum at 3 Å can be interpreted to arise from intradimeric Pt–Pt distances and that at 5–6 Å from the shortest interdimeric Pt–Pt distances. The broadness of these maxima can be explained partly by Fourier filtering and partly by variation in the distances.

The 2.6 Å maximum is an interesting detail in the ΔRDF obtained with 17.4 keV. We demonstrate with simulation that the maximum can arise from a Pt–Pt distance but not exist in the calculated 11.5 keV ΔRDF because of the k-space termination. Fig. 7(a) shows model k × TSF(k) for a mixture of two dinuclear Pt units with Pt–Pt distances of 2.6 and 3 Å, terminated as the experimental TSFs obtained with 11.5 and 17.4 keV, and Fig. 7(b) shows the corresponding calculated ΔRDFs. The 17.4 keV ΔRDF contains a maximum due to the Pt–Pt distance of 2.6 Å but the 11.5 keV ΔRDF a minimum at this distance. The simulation thus explains the difference between the 11.5 and 17.5 keV experimental ΔRDFs and indicates that the 2.6 Å maximum arises from atomic distances.

The maximum at 2.6 Å in the experimental 17.4 keV ΔRDF is intensive enough to arise from a Pt–Pt distance, but it might also arise from distances between Pt atoms and O atoms of the SO₄²⁻ counter ions. To study this, the experimental RDF (17.4 keV) and Pt–Pt PDF were compared with those based on the structures of the crystalline dinuclear model Pt complexes (O’Halloran et al., 1987; Faggiani, Lock, Pollock, Rosenberg & Turner, 1981). The model RDF and PDF were obtained by calculation first of the total and the Pt–Pt partial structure factor from atomic coordinates and then of the model RDF and the Pt–Pt PDF with the same procedure that was used for obtaining the experimental ones.

According to the calculations, a model of randomly distributed dinuclear Pt units and counter ions, where the Pt–Pt distance is 3 Å and the distance between Pt and O atoms of the SO₄²⁻ counter ion is about 2.6 Å, gave a too intensive maximum at 3 Å for both the model Pt–Pt PDF (Fig. 8) and the total RDF and a too low maximum of the total RDF at 2.6 Å. Agreement with the Pt–Pt PDF but not with the total RDF at 2.6 Å would be improved by the assumption that about 20% of Pt atoms are in monomers.

The mixture of two dinuclear Pt units with Pt–Pt distances of 3 and 2.58 Å gave a better agreement with the experimental data (Fig. 8). The dinuclear Pt unit with the Pt–Pt distance of 2.58 Å was constructed on the basis of...
the structure of the crystalline dinuclear Pt(III) complex by Hollis & Lippard (1982). At 5.7 Å, the agreement could be improved further by allowing more variation in the interdimeric Pt–Pt distances.

On the basis of these results, it is suggested that Pt-uridine blue contains also structures where the Pt–Pt distance is 2.6 Å, but further studies are needed to prove this unambiguously. We attribute this distance to dinuclear Pt(III)–Pt(III) complexes. These valuable results encourage the application of the same methods to biologically more active green products for investigation of the structural origin of the anticancer activity.

The regularization was found to be essential for obtaining the dominating Pt–Pt structure factor. The simulations indicated that it is important to have a reasonable centre vector for all significant components of the solution and that the dominant partial structure factor can be found iteratively, starting from zero. It is notable that the applied iterative process gives good results also at small k and may thus be useful in the analysis of anomalous small-angle scattering (ASAXS) data as well, if an appropriate centre vector can be found.

Calculations with simulated data including statistical errors before the experiments are emphasized. The simulations help in choosing a proper measuring strategy for obtaining sufficiently accurate intensity curves from which the partial structure factors can be solved.

Regularization affects the behaviour of the covariances of the PSFs as a function of k. With the chosen measuring strategy, the regularization matrix and the regularization parameter, the statistical error of partial structure factors decreased slightly as a function of k. In general, because the calculation of the distribution function involves multiplication of the structure factors by k and a ‘sharpening factor’ \((f)\), it is important to measure the intensities more accurately at large k than at low k to obtain a well behaved partial distribution function.

7. Conclusions

Even though the anomalous scattering experiments were time consuming owing to the precision of the data needed, they gave valuable information on the average Pt structures in the sample. The Pt–Pt partial radial distribution function of an amorphous platinum pyrimidine complex, Pt-uridine blue, was solved for the first time in a new iterative way by using the Tikhonov regularization technique. The result indicates that the dominating Pt species are dinuclear. The study demonstrates that a specific partial structure factor of a multicomponent system can be solved from a carefully measured data set even if the sample contains only one element with a suitable absorption edge for the measurement. The use of the regularization technique minimizes problems related to ill-conditioning and statistical errors in the data.

This research was supported by the Academy of Finland, the Neste Foundation, the Technology Development Centre, the University of Helsinki and the Office of Basic Energy Sciences of the US Department of Energy. The synchrotron-radiation work was performed at SSRL, which is supported by the US Department of Energy, Office of Basic Energy Sciences.

APPENDIX

For the coupled case, a plausible regularization functional is the following general second-order derivative norm:

\[ \|x\|_B^2 = \|x\|_2^2 + \|x'\|_2^2 + \|x''\|_2^2, \]

which is a linear combination of the function norms of the vector \(x\) and its first and second derivatives with respect to the scattering vector \(k\). Note, however, that this ‘derivative type’ norm is not the most general choice for the norm (or the matrix \(B\)).

When the derivatives are discretized as differences, one obtains the norm

\[ \|x\|_B^2 = \sum_{i=1}^K (\alpha_i \|x_i\|_{b_i}^2 + \beta_i \|x_i - x_{i-1}\|_{c_i}^2 + \gamma_i \|x_{i+1} - 2x_i + x_{i-1}\|_{d_i}^2). \]

The summation is over the index \(i = 1, 2, \ldots, K\) representing the discrete \(k\) values \(k_i\).

The weights \(\alpha_i, \beta_i\) and \(\gamma_i\) are arbitrary constants to be chosen suitably. If \(\beta_i\) or \(\gamma_i\) differs from 0, the positive definite regularization matrix \(B\) couples different \(k\).

Fig. 8. The experiment Pt–Pt partial radial distribution function (filled circles) together with models of randomly distributed dinuclear Pt units with the Pt–Pt distance of 3 Å (thin line) and with Pt–Pt distances of 3 and 2.58 Å (thick line).
values. If \(\beta_i = 0 = \gamma_i\), the regularization matrix is block diagonal and the \(k\) values are not coupled (separate case).

The solution \(x = x(k)\) is a vector with components \(x_{ij}, j = 1, 2, \ldots, N\). The number of the components is equal to the number of independent PSFs, which depends on the number of atomic species in the sample. For a two-component sample, there are \(N = 3\) components in the vectors \(x_i\).

The norms in the above equation are norms in the \(N\)-dimensional space of \(x\) vectors, and in principle they might all be chosen to be different (or the same, for that matter). They too can be defined with positive Hermitian operators \(b_i, c_i\) and \(d_i\), as indicated in the above formula by subscripts. For instance,

\[
\|x_i\|_{b_i}^2 = \|x_i\|_{\gamma_i}^2 = \sum_{l,m=1}^{N} x_i^* b_{ilm} x_{im}
\]

(similar equations hold for other norms in the above definition). The (small) \(N \times N\) square matrices \(b_i, c_i\) and \(d_i\) are Hermitian and positive (the matrix \(b_i\) should be definite, too), but otherwise arbitrary; they usually even depend on the variable \(k_i\), as indicated by the index \(i\).

To reduce the number of undetermined coefficients, we have chosen \(c_i\) and \(d_i\) to be \(N \times N\) unit matrices, and the matrices \(b = b_i\) are chosen independent of the index \(i\).

In the case \(N = 3\), we choose the ‘derivative type’ norm

\[
\|x_i\|_b^2 = \beta_0(x_{i1}^2 + x_{i2}^2 + x_{i3}^2) + \beta_0(x_{i1} - x_{i2})^2 + \gamma_0(x_{i1} - 2x_{i2} + x_{i3})^2,
\]

which corresponds to the matrix

\[
b = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix} + \gamma_0 \begin{pmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{pmatrix},
\]

where \(\beta_0\) and \(\gamma_0\) are arbitrary constants to be chosen suitably. (The matrices are discretized difference versions of the squares of different orders of the derivatives, and can be obtained from Pascal’s triangle.) This can easily be generalized for other values of \(N\), and to norms corresponding to higher-order derivatives.

Hence, any of the matrices \(b_i, c_i\) or \(d_i\) can contain the norms for the zeroth-, first- and second-order derivatives or differences with respect to the component index for the vector \(x_i = (x_{i1}, x_{i2}, x_{i3})\) (when \(N = 3\); generally the norm can contain derivatives up to the order \(N - 1\)). Of course, here these ‘derivatives’ have no meaning whatsoever.

**References**


