Z. Kristallogr. **217** (2002) 694–702



# Direct way to anomalous scatterers

Z. Dauter*,[I] and R. A. P. Nagem[I,II]

[I] Synchrotron Radiation Research Section, National Cancer Institute, Brookhaven National Laboratory, Bldg. 725A-X9, Upton, NY 11973, USA
[II] CBME, Laboratório Nacional de Luz Síncrotron and Instituto de Física Gleb Wataghin, UNICAMP, Caixa Postal 6192, CEP 13084-971, Campinas, SP, Brazil

Received July 5, 2002; accepted August 29, 2002

**Abstract.** The first step in solving macromolecular crystal structures by multi- or single-wavelength anomalous diffraction methods is the location of the anomalous scatterers. This can be done by direct methods, using either Bijvoet differences within the single data set, or anomalous scattering amplitudes estimated from measurements at several wavelengths. The calculations suggest that Bijvoet differences are equally successful for this purpose as anomalous amplitudes, $F_A$, which theoretically should be more suitable. The calculation of $F_A$ values is susceptible to the accumulation of errors contained in the individual intensity measurements at several wavelengths and in the inaccurate estimation of the anomalous atomic scattering corrections, $f'$ and $f''$. Direct methods often give better results at resolution lower than the full extent of the diffraction data limit. This may be attributed to the enhanced accuracy of measurements of the strong, low resolution reflections and to more effective phase refinement and propagation through $\sum_2$ relations.

## Introduction

The classic way of solving macromolecular crystal structures, the Multiple Isomorphous Replacement (MIR) method, relies on differences between structure amplitudes measured from the native protein crystal and from derivative crystals, containing a number of heavy atoms, usually incorporated by soaking native crystals in solutions of appropriate reagents (Blundell and Johnson, 1976; Drenth, 1997). The classic MIR procedure assumes that derivatives are isomorphous with the native crystals, which means that the lattice and the scattering contribution of the protein molecules in all crystals are unchanged, and the observed differences result only from the additional scattering of the heavy atoms. In practice, derivative atoms always perturb to some extent the protein structure, and non-isomorphism remains a problem in MIR phasing. As a consequence the lattice and the scattering contribution of the protein atoms in the native and derivative crystals are not unchanged, and the measured intensity differences do not result only from the additional scattering of the heavy atoms.

* Correspondence author (e-mail: dauter@bnl.gov)

The potential of anomalous scattering for phasing protein crystals was realized early (Rossmann, 1961; North, 1965; Matthews, 1966a, 1966b). However, at first the anomalous scattering effect of heavy atoms was used only as an auxiliary source of phase information, since the Bijvoet differences, usually smaller than isomorphous differences, were difficult to estimate accurately. The introduction of crystal freezing, availability of tunable synchrotron beam lines and fast and accurate X-ray detectors enhanced the use of anomalous scattering for phasing macromolecular diffraction data. The current most widely used method of phasing novel structures, the multi-wavelength anomalous diffraction (MAD) approach (Hendrickson, 1991, 1999), relies entirely on the anomalous scattering effect. Since the data are usually collected from the same specimen at various wavelengths, the non-isomorphism is almost totally alleviated in MAD, and may result only from the crystal radiation damage. If diffraction data contain accurately measured anomalous scattering signal, it is possible to obtain a structure solution by single-wavelength anomalous diffraction (SAD) approach (Hendrickson and Teeter, 1981; Wang, 1985; Dauter, Dauter and Dodson, 2002).

The first step in MAD and SAD phasing involves the solution of the partial structure of anomalous scatterers, since knowledge of the calculated anomalous atoms diffraction contribution, $F_A$, is necessary for estimation of the protein phases. The number of anomalously scattering atoms in the macromolecule is usually small and they are located at mutual distances longer than chemical bonds in a macromolecule. The partial structure of the anomalous scatterers consists therefore of non-overlapping atoms even at resolution of the diffraction data much lower than 1 Å. The condition of "atomicity", required by direct methods, is fulfilled at resolutions practically achieved in macromolecular crystallography. Since first proposed for this purpose (Mukherjee, Helliwell and Main, 1989), direct methods proved to be extremely useful for locating anomalous scatterers in macromolecular crystals.

## Anomalous scattering contributions and Bijvoet differences

Figure 1 illustrates the vector diagram for a pair of Friedel-related structure factors, $F^+_T$ and $F^-_T$, containing a

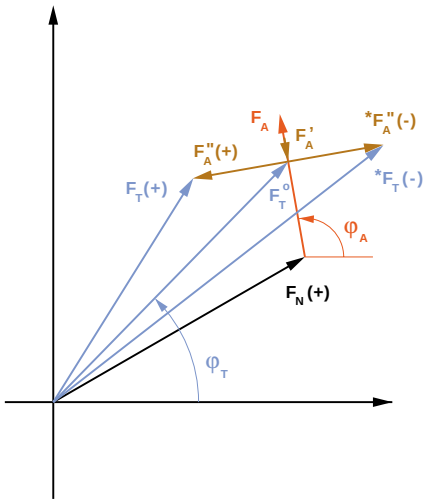



**Fig. 1.** Vector diagram showing the relations between the normal and anomalous scattering contributions to the entire scattering factor, $F_T = F_N + F_A + F_A' + iF_A''$.

contribution of normally scattering atoms, $F_N$, and of anomalous scatterers, $F_A$. It is assumed that all anomalous scatterers are of the same kind; hence, the imaginary contribution of the anomalous scatterers, $F''_A$, is perpendicular to their normal scattering vector, $F_A$. Since the normal scattering component of anomalous scatterers is represented by $F_A$, these values should be used for partial structure solution. However, the amplitudes $|F_A|$ cannot be estimated from the single-wavelength data, because the relation between the experimentally measured Bijvoet difference, $\Delta F^{\pm} = |F^+{}_T| - |F^-{}_T|$, and $|F_A|$, depends on the unknown difference of phases, $\varphi_T$ and $\varphi_A$, according to the relation:

$$\begin{aligned}\Delta F^{\pm} &= 2F''_A \sin(\varphi_T - \varphi_A) \\ &= 2(f''/f^0)\, F_A \sin(\varphi_T - \varphi_A)\,.\end{aligned}$$

Only if Bijvoet differences measured at more than one wavelength are available, as in the MAD experiment, it is possible to estimate $F_A$, assuming that the atomic scattering factors $f^0$, $f'$ and $f''$ are known. Several methods have been proposed to estimate values of $F_A$ from multi-wavelength data.

The classic MAD approach (Karle, 1980; Hendrickson, 1991) provides the value of $F_A$ (as well as $F_T$ and $\varphi_T - \varphi_A$) from the solution of the system of algebraic equations. In the REVISE procedure (Fan, Woolfson and Yao, 1993) the Bijvoet differences at different wavelengths are checked for consistency and corrected, to make the ratio $(|F^+|^2 - |F^-|^2)/f''$ wavelength-independent, and simultaneously the $F_A$ values are optimally evaluated. The recently introduced (Burla, Carrozzini, Cascarano, Giacovazzo, Polidori & Siliqi, 2002) rigorous method of evaluating $F_A$, based on the joint probability distribution functions, may potentially provide better estimates than the traditional techniques.

In the SAD case, the only available quantities are Bijvoet differences. From the sine relation between $\Delta F^{\pm}$ and $F''_A$ it can be inferred that for reflections with the largest anomalous differences, $(\varphi_T - \varphi_A) = \pm\pi/2$, and the magnitude of Bijvoet difference is proportional to $F_A$. Those reflections are quite appropriate for use for a direct methods solution of the anomalous substructure, since these methods use only a subset of the largest normalized structure amplitudes.

## Diffraction data

Almost all diffraction data presented in Table 1 were collected at the beam line X9B of NSLS (Brookhaven National Laboratory) using the Quantum-4 ADSC CCD detector and were processed with HKL2000 (Otwinowski and Minor, 1997). The only exception is ferredoxin, with data collected at EMBL Hamburg. The following data sets were used previously for structure solution and the results were published: ferredoxin (Dauter et al., 1997), native lysozyme (Dauter et al., 1999), *E. coli* thioesterase (Li et al., 2000), human thioesterase + Br (Devedjiev et al., 2001), and *Pseudomonas* serine-carboxyl proteinase + Br (PSCP, Dauter et al., 2001).

Data included in Table 1 contain various amounts of an anomalous signal from different anomalous scatterers. In some cases, the X-ray wavelength was optimized to maximize the anomalous diffraction signal of elements such as Br, Se, Ta and Lu, which are often used for MAD phasing; in other data sets, the anomalous signal originates from S, Cl, Fe and Mn, at wavelengths far remote from the absorption edges of these elements. In two MAD cases (subtilisin + Lu and *E. coli* thioesterase), the typical four-wavelengths MAD data were collected, whereas for lysozyme soaked in 1M solution of NaBr and glucose isomerase soaked in 2 mM solution of $Ta_6Br_{12}{}^{-2}$ cluster data were collected at several wavelengths through the absorption edge of bromine or tantalum. The latter multi-MAD data were acquired especially to estimate the $F_A$ values as accurately as possible and to investigate the variation of the anomalous scattering contributions in the vicinity of the absorption edge. The values of $f'$ and $f''$ were estimated with XPREP (Bruker Analytical X-ray Systems) for MAD data and are plotted in Fig. 2. In all MAD cases, estimation of the anomalous scattering factors $f'$ and $f''$ by XPREP gave quite reasonable results. Whereas this procedure gives the absolute values of $f''$, the $f'$ values cannot be estimated absolutely and only the relative values at different wavelengths are significant. For lysozyme crystal soaked in NaBr, the synchrotron ring injection occurred between collection of the fifth and sixth data sets, and the X-ray wavelength drifted by a small amount, which is evidenced by the perturbation in the curve of the estimated $f''$ values.

## Comparison of the use of $\Delta F^{\pm}$ and $F_A$

The $F_A$ values can only be estimated approximately, and their accuracy depends on the precision of intensity measurements at various wavelengths, as well as on the stability of wavelength during the whole data collection session. The effect of crystal deterioration may also play an important role. The large value of $F_A$ may result from a very small Bijvoet difference if the $F_T$ and $F_A$ vectors are nearly parallel, and in such cases the estimation may be

**Table 1.** Diffraction data with anomalous scattering contribution. Values of $f'$ and $f''$ for MAD data were estimated with XPREP (Bruker Analytical X-ray Systems).

| Protein | Size (amino acids) | Resolution (Å) | Anomalous atoms | Wavelength (Å) | $f'$ | $f''$ |
|---|---|---|---|---|---|---|
| MAD data | | | | | | |
| Lysozyme + Br | 127 | 1.9 | 5Br[a] | 0.92357 | −0.68 | 0.58 |
| | | | | 0.92042 | −2.92 | 1.23 |
| | | | | 0.92028 | −3.25 | 1.89 |
| | | | | 0.92021 | −3.56 | 1.99 |
| | | | | 0.92014 | −3.51 | 3.18 |
| | | | | 0.92007 | −3.86 | 3.15 |
| | | | | 0.92001 | −2.99 | 4.80 |
| | | | | 0.91987 | −2.09 | 5.20 |
| | | | | 0.91674 | −0.24 | 3.98 |
| Gluc. Isom. + $Ta_6Br_{12}$ | 388 | 1.7 | 30Ta[a] | 1.28065 | −1.34 | 1.05 |
| | | | | 1.25549 | −6.40 | 1.75 |
| | | | | 1.25511 | −8.22 | 2.44 |
| | | | | 1.25486 | −9.00 | 3.82 |
| | | | | 1.25460 | −6.79 | 5.77 |
| | | | | 1.25435 | −3.32 | 6.05 |
| | | | | 1.25410 | −0.55 | 4.92 |
| | | | | 1.25384 | −0.31 | 3.95 |
| | | | | 1.25360 | −0.56 | 3.48 |
| | | | | 1.25336 | −0.32 | 3.24 |
| *E. coli* Thioesterase | 590 | 2.5 | 8Se | 0.98008 | −1.82 | 1.14 |
| | | | | 0.97931 | −5.99 | 3.98 |
| | | | | 0.97870 | −3.32 | 6.40 |
| | | | | 0.97469 | −1.17 | 5.12 |
| Subtilisin + Lu | 275 | 1.75 | 4Lu[a] | 1.37758 | −1.45 | 1.32 |
| | | | | 1.34120 | −4.29 | 4.87 |
| | | | | 1.34080 | −4.39 | 7.54 |
| | | | | 1.31280 | −1.80 | 2.99 |
| SAD data | | | | | | |
| Human Thioesterase | 418 | 1.8 | 22Br[a] | 0.92 | ∼ −3.0 | ∼5.0 |
| PSCP | 375 | 1.8 | 20Br[a] | 0.92 | ∼ −3.0 | ∼5.0 |
| Ferredoxin | 55 | 0.94 | 8Fe | 0.88 | 0.28 | 1.25 |
| Thermolysin | 316 | 1.45 | 1Zn + 4Ca | 1.28 | ∼ −6.0 + 0.35 | ∼4.0 + 0.92 |
| Gluc. Isom. II form | 776 | 1.50 | 2Mn + 18S | 1.54 | −0.57 + 0.32 | 2.80 + 0.56 |
| Gluc. Isom. | 388 | 1.50 | 1Mn + 9S | 1.34 | −0.06 + 0.28 | 2.23 + 0.43 |
| Gluc. Isom. | 388 | 1.60 | 1Mn + 9S | 1.08 | 0.24 + 0.21 | 1.54 + 0.28 |
| Gluc. Isom. | 388 | 1.45 | 1Mn + 9S | 0.98 | 0.29 + 0.18 | 1.30 + 0.23 |
| Lysozyme | 127 | 1.55 | 10S + 8Cl[a] | 1.54 | 0.32 + 0.35 | 0.56 + 0.70 |
| Thaumatin | 207 | 1.7 | 17S | 1.54 | 0.32 | 0.56 |

a: These atoms are partially occupied in the crystal structure.

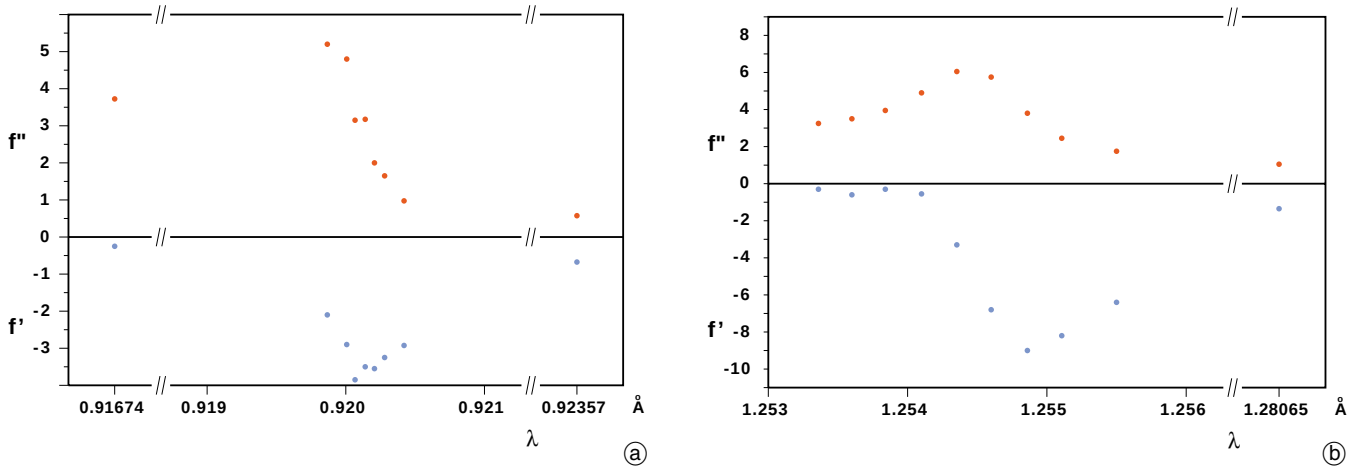



**Fig. 2.** Values of $f'$ and $f''$ estimated by XPREP (Bruker Analytical X-ray Systems) for MAD data sets: **a**) lysozyme + Br, **b**) glucose isomerase + $Ta_6Br_{12}$.

very inaccurate. It may be safer to use Bijvoet differences within a single wavelength data set, rather than $F_A$ set erroneously estimated from multi-wavelength data. The $F_A$ data set may contain centrosymmetric reflections, but those reflections do not occur in the $\Delta F^{\pm}$ set, since Friedel-related centrosymmetric reflections are equivalent.

The MAD data sets, shown in Table 1, collected from crystals containing various anomalous scatterers, were used for solution of anomalous scatterers' positions with program SHELXD (Sheldrick, 1998), comparing two approaches, using either Bijvoet differences, $\Delta F^{\pm}$, within the data set containing the highest $f''$ contribution (SAD data at peak wavelength) or amplitudes $F_A$ estimated by XPREP from all available data sets collected with various wavelengths (MAD data). One hundred multisolution trials were run, with the same parameters used in both approaches, including the resolution and number of highest normalized amplitudes (about 1500), although the reflections selected automatically were not the same. The SAD and MAD data sets are compared in Figure 3, which shows the fraction of reflections common for both sets in batches (multiples of 50) of the strongest normalized amplitudes. The number of centrosymmetric reflections in the $F_A$ data is also shown.

The number of centrosymmetric reflections in the $F_A$ data obviously depends on the crystal symmetry and it may be expected that in higher symmetry space groups their percentage is larger than in lower symmetry space groups. In the examples shown in Fig. 3, the centrosymmetric reflections amount to 10–20%, up to about 30% among the largest group of reflections for lysozyme in $P4_32_12$ symmetry.

In general, the number of common reflections in the SAD and MAD data is rather low, especially among the largest amplitudes, less than 20%, and not exceeding 50% overall. Nevertheless, a comparison of the SHELXD results, Figure 4 (a–h), shows that the use of SAD and MAD data leads to similar results. The success rate of solutions with a high $E_o/E_c$ correlation coefficient (CC) in both cases is comparably high in various resolution ranges.

## Effect of resolution

In contrast to the normal atomic scattering factors, $f^0$, which diminish with increasing diffraction angle but do not depend on wavelength, the anomalous corrections, $f'$ and $f''$, do not depend on the diffraction angle but vary with wavelength of X-rays: $f(\theta,\lambda) = f^0(\theta) + f'(\lambda) + if''(\lambda)$. It may therefore be expected that the relative amount of anomalous signal should be larger at a higher resolution. Unfortunately, in macromolecular data this tendency is frustrated by the fact that the high-resolution re-

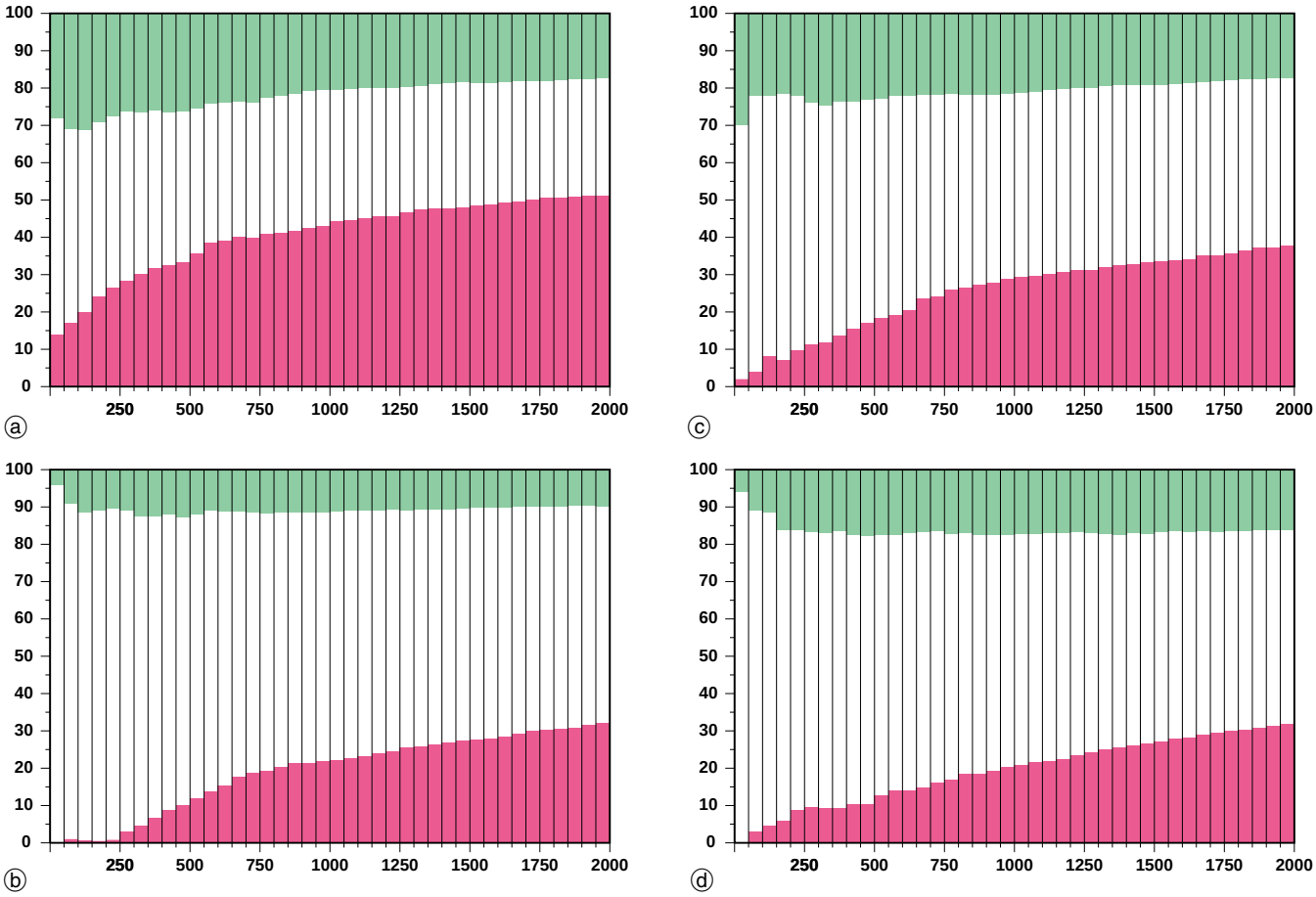



**Fig. 3.** The percentage of common reflections in the $F_A$ and $\Delta F^{\pm}$ data sets among a number (in batches of multiples of 50) of the largest normalized amplitudes in these sets, in: (**a**) lysozyme + Br, (**b**) glucose isomerase + $Ta_6Br_{12}$, (**c**) subtilisin + Lu, (**d**) *E. coli* thioesterase. The fraction of centrosymmetric reflections in the $F_A$ data is shown in green at the top of the histogram.

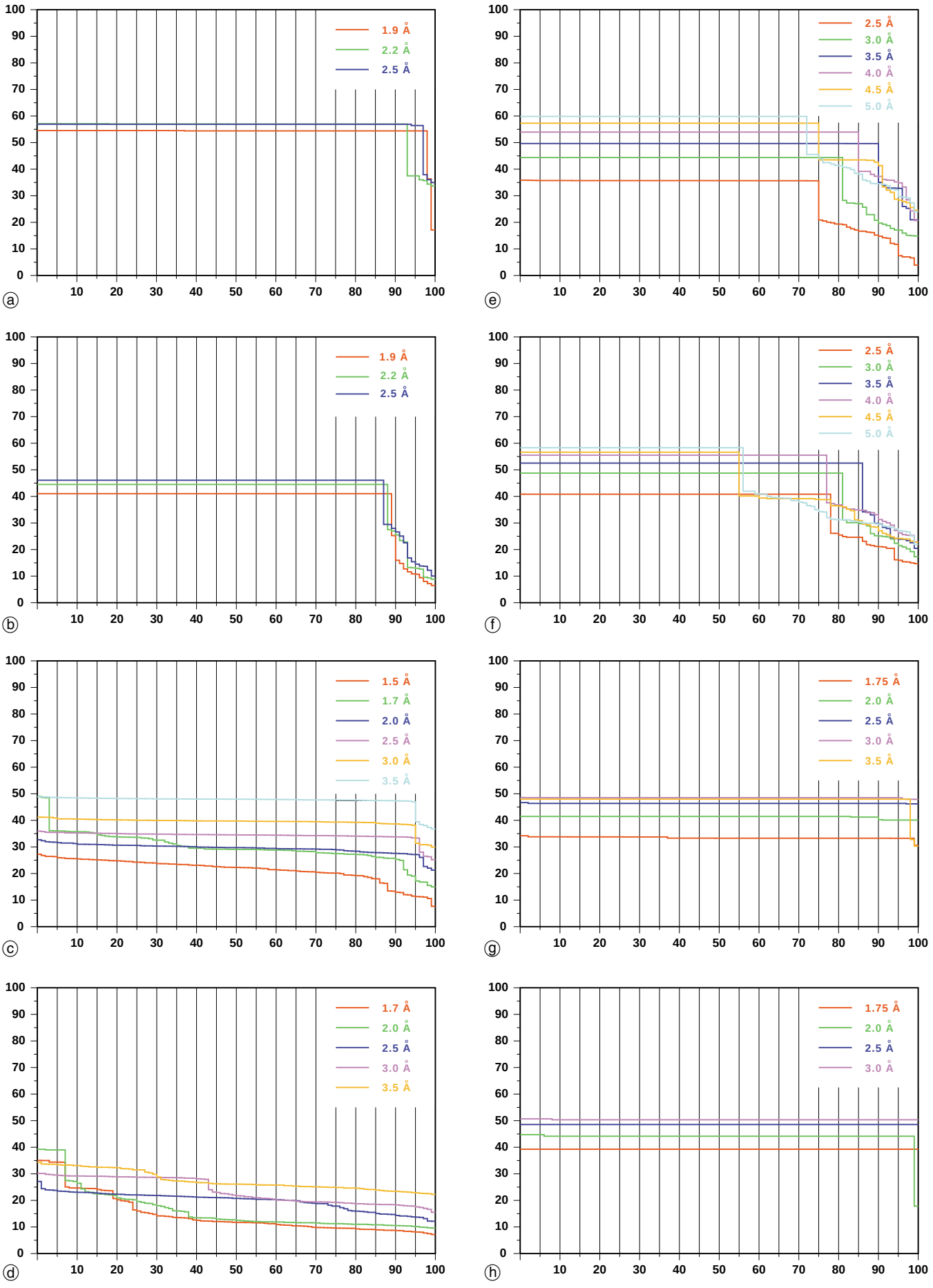

a
1.9 Å
2.2 Å
2.5 Å
e
2.5 Å
3.0 Å
3.5 Å
4.0 Å
4.5 Å
5.0 Å
b
1.9 Å
2.2 Å
2.5 Å
f
2.5 Å
3.0 Å
3.5 Å
4.0 Å
4.5 Å
5.0 Å
c
1.5 Å
1.7 Å
2.0 Å
2.5 Å
3.0 Å
3.5 Å
g
1.75 Å
2.0 Å
2.5 Å
3.0 Å
3.5 Å
d
1.7 Å
2.0 Å
2.5 Å
3.0 Å
3.5 Å
h
1.75 Å
2.0 Å
2.5 Å
3.0 Å

**Fig. 4.**

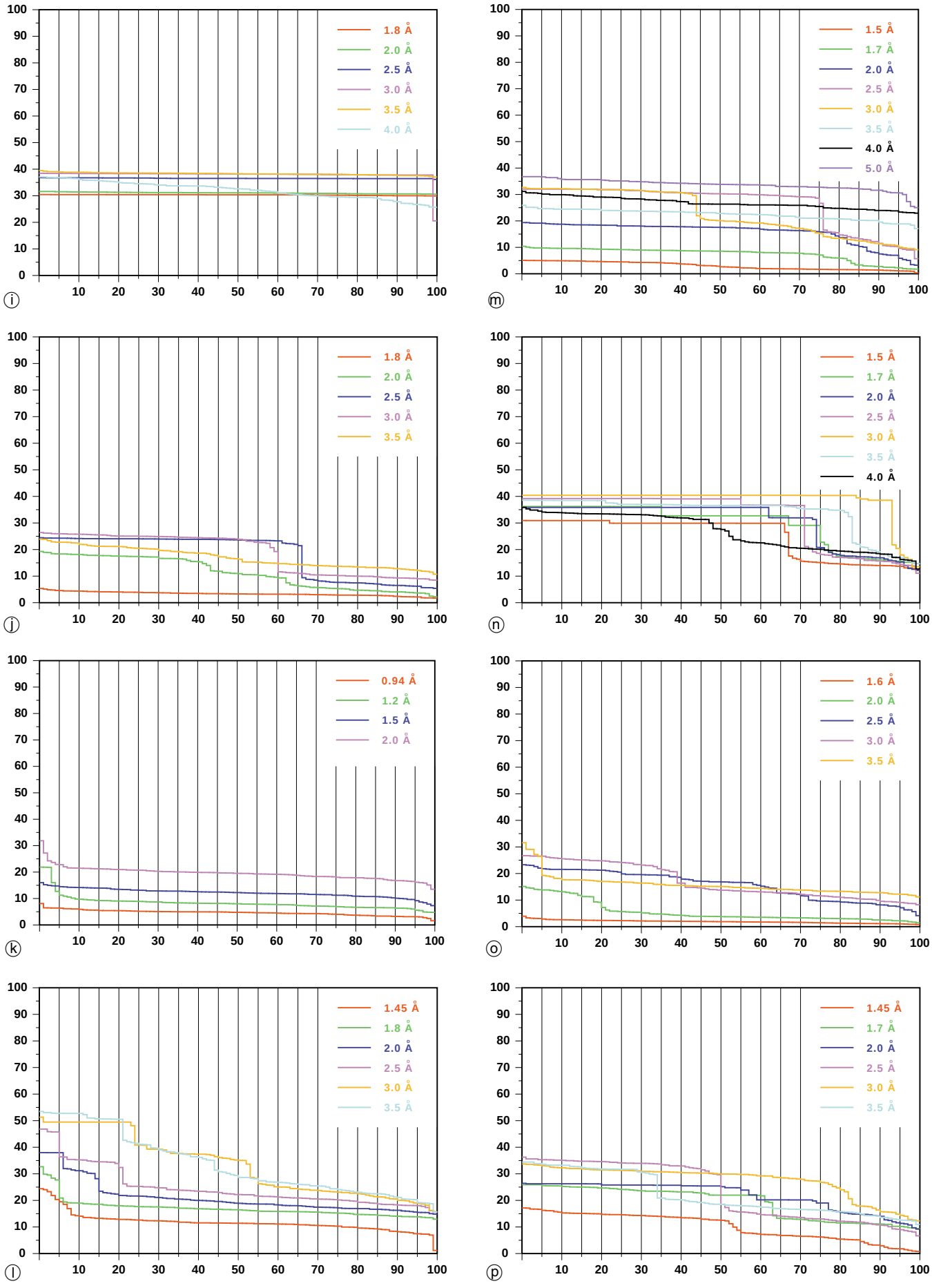
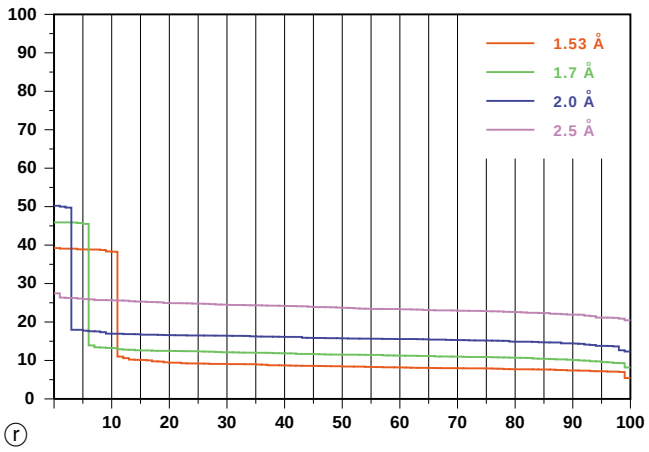

1.8 Å
2.0 Å
2.5 Å
3.0 Å
3.5 Å
4.0 Å
(i)
1.5 Å
1.7 Å
2.0 Å
2.5 Å
3.0 Å
3.5 Å
4.0 Å
5.0 Å
(m)
1.8 Å
2.0 Å
2.5 Å
3.0 Å
3.5 Å
(j)
1.5 Å
1.7 Å
2.0 Å
2.5 Å
3.0 Å
3.5 Å
4.0 Å
(n)
0.94 Å
1.2 Å
1.5 Å
2.0 Å
(k)
1.6 Å
2.0 Å
2.5 Å
3.0 Å
3.5 Å
(o)
1.45 Å
1.8 Å
2.0 Å
2.5 Å
3.0 Å
3.5 Å
(l)
1.45 Å
1.7 Å
2.0 Å
2.5 Å
3.0 Å
3.5 Å
(p)

**Fig. 4.**



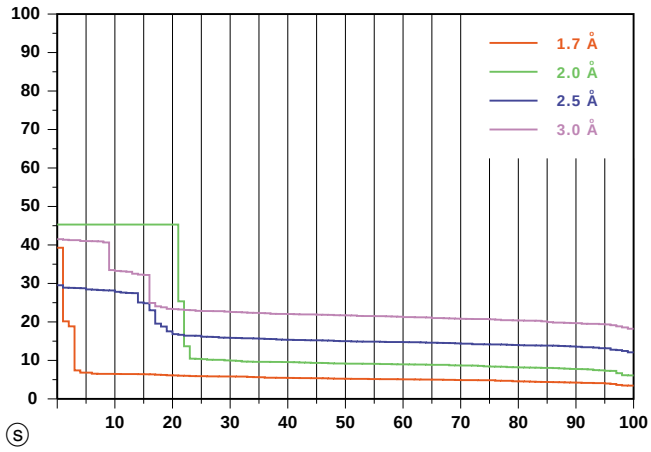




**Fig. 4.** Results of SHELXD for various data sets, showing the $E_o/E_c$ correlation coefficient for 100 multisolution trials run for each data. Various colours correspond to different data resolution cut-offs. (**a**) lysozyme + Br, $F_A$ data; (**b**) lysozyme + Br, $\Delta F^{\pm}$ data; (**c**) glucose isomerase + $Ta_6Br_{12}$, $F_A$ data; (**d**) glucose isomerase + $Ta_6Br_{12}$, $\Delta F^{\pm}$ data; (**e**) *E. coli* thioesterase, $F_A$ data; (**f**) *E. coli* thioesterase, $\Delta F^{\pm}$ data; (**g**) subtilisin + Lu, $F_A$ data; (**h**) subtilisin + Lu, $\Delta F^{\pm}$ data; (**i**) human thioesterase; (**j**) PSCP; (**k**) ferredoxin; (**l**) thermolysin; (**m**) glucose isomerase, II form, $\lambda = 1.54$ Å; (**n**) glucose isomerase, $\lambda = 1.34$ Å; (**o**) glucose isomerase, $\lambda = 1.08$ Å; (**p**) glucose isomerase, $\lambda = 0.98$ Å; (**r**) lysozyme; (**s**) thaumatin.

flections are much weaker and generally measured with less accuracy than the strong, low-resolution data. The anomalous differences usually amount to 1–5% of the total structure amplitude, and may be easily lost in the noise contained in the high resolution, weak intensities.

Another effect, characteristic for direct methods applied to macromolecular data at high resolution, is the sparse distribution of a relatively small subset of strong reflections (*i.e.*, with large $F_A$ or $\Delta F^{\pm}$) in the large reciprocal diffraction sphere, densely populated by weak reflections. As a consequence, the number of $\Sigma_2$ interactions between reflection triplets is small at high resolution, particularly in lower symmetry space groups. The only remedy is to increase the number of reflections at the cost of computing time. At low resolution the diffraction sphere is smaller, and the same number of strong reflections interrelate through more $\Sigma_2$ triplets. However, at very low resolution there may not be enough strong reflections above the minimal $E = 1.0$ limit.

Inspection of Fig. 4 shows that for most data sets the solution of the partial structure of anomalous scatterers can be obtained at high as well as low resolution, although at low resolution the chance of success is higher. For PSCP, ferredoxin, glucose isomerase at $\lambda = 1.54$ Å and 1.08 Å, attempts at full diffraction data resolution were unsuccessful with CC below 10%, but at resolution lower than 2.0 Å, in all these cases the correct solution was obtained. The higher rate of success at low resolution is particularly visible for data with a small amount of the anomalous signal resulting from weak anomalous scatterers, such as iron in ferredoxin at $\lambda = 0.88$ Å, manganese in glucose isomerase, zinc in thermolysin or sulfur in thaumatin or lysozyme. The apparent lower success for native lysozyme at 2.5 Å than at a higher resolution results from the fact that eight out of ten sulfurs in lysozyme form disulfide bridges with S–S distance of about 2.1 Å; and at a resolution lower than that value, those sulfurs cannot be successfully resolved.

## Data redundancy

Practically all diffraction data from macromolecular crystals are collected with the rotation method and two-dimensional detectors. Usually the minimalist approach is applied, with the rotation range selected to give a complete data set, with (almost) each reflection measured at least once. However, the two-dimensionality of the detector and the crystal symmetry lead to most reflections being measured several times. This makes it possible to merge and scale the individual intensities and to estimate their uncertainties. If the redundancy of measurements is low, the accuracy of estimated intensities is poor. The accuracy of estimated intensities can be enhanced by increasing the multiplicity of measurements of individual reflections, including their symmetry equivalents. Since the anomalous scattering signal is usually estimated as a small difference between large amplitudes, the data accuracy is especially important in MAD and SAD.

The effect of data redundancy on the success rate of solving the anomalous substructure was investigated for the native lysozyme. The original data were collected in four passes with various exposure times, each through 180° of total rotation (Dauter et al., 1999). These data were reprocessed in narrower batches of total rotation and the statistics is given in Table 2.

The increased redundancy has a clear effect on the data quality, as evidenced by the increasing $I/\sigma(I)$ ratio. The standard $R_{merge}$ increases with redundancy, which only underlines the known fact that it is a poor data quality criterion (Weiss and Hilgenfeld, 1997; Diederich and Karplus, 1997). Fig. 5 shows the Bijvoet ratio as a function of resolution for these data sets. For the most accurate data from 180° rotation the Bijvoet ratio is actually smallest, but closest to 1.4%, the value theoretically expected from the presence of ten sulfurs and eight chlorides among 1001 atoms of lysozyme.

As evidenced in Table 2, the success rate of SHELXD solutions critically depends on data quality and parallels

| Total rotation (°) | 45 | 60 | 90 | 135 | 180 |
|---|---|---|---|---|---|
| Redundancy[a] | 3.4 | 4.1 | 6.1 | 8.3 | 10.4 |
| Completeness (%)[b] | 86.1 (64.5) | 94.5 (68.5) | 96.6 (69.3) | 98.1 (83.5) | 100.0 (99.8) |
| $R_{merge}$ (%)[b] | 3.1 (11.4) | 3.2 (10.8) | 3.4 (12.9) | 3.8 (14.3) | 4.0 (15.7) |
| $I/\varrho(I)$ | 34.8 (5.8) | 39.1 (6.5) | 47.7 (7.8) | 50.2 (7.4) | 53.5 (6.8) |
| Solutions[c] | | | | 45 | 144 |

a: Redundancy and completeness refer to individual Friedel mates
b: In parentheses are values for the highest resolution shell, 1.56–1.53 Å
c: The number of correct solutions in 1000 SHELXD phasing trials

**Table 2.** Batches of data for native lysozyme processed with various redundancy.

the redundancy of measurements. When the diffraction data contains only a small amount of the anomalous signal, as in case of the native lysozyme, it is important to enhance the accuracy of intensity estimation by recording data with high redundancy of measurements.

## Conclusions

As evidenced in Fig. 4, it was possible to obtain by the direct methods program SHELXD the positions of anomalous scatterers from the anomalous signal contained in all the diffraction data quoted in Table 1, collected from various crystals at different resolutions. The amount of anomalous signal in these data varies from a few percent to about 0.5% for glucose isomerase at $\lambda = 0.98$ Å. For the successful use of very weak anomalous signal it is important to collect highly redundant diffraction data. In all of the cases discussed above, it was also possible not only to locate the anomalous scatterers, but also subsequently to solve the protein model by SAD phasing.

From the comparison of the performance of phasing based on $\Delta F^{\pm}$ and $F_A$, it seems that both types of data can lead to a partial structure solution with a similar chance of success. From the theoretical point of view, the phasing process based on Bijvoet differences may be expected to be inferior, since the $\Delta F^{\pm}$ data lack all centrosymmetric reflections and their values only approximate the scattering contribution of the anomalous atoms. However, the $F_A$ values, which theoretically are more appropriate for this purpose, may contain additional errors resulting from difficulties in their accurate estimation from the multi-wavelength data with uncertain values of $f'$ and $f''$. As a result, the large part of the theoretical advantage of using $F_A$ may be lost in practical applications. It may be expected that the full advantage of using $F_A$ values will result from the improved methods of their estimation, such as based on the appropriate probability distributions (Burla et al., 2002).

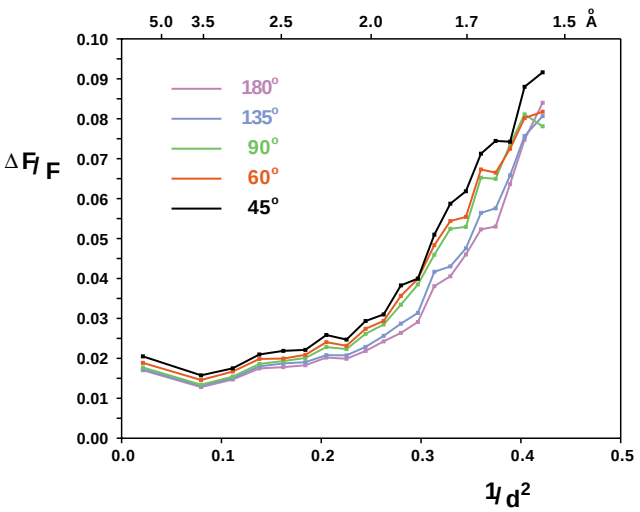



**Fig. 5.** The Bijvoet ratio $\langle \Delta F^{\pm} \rangle / \langle F \rangle$ as a function of resolution for native lysozyme data collected with various total rotation ranges. The theoretically expected value is 1.4% for ten sulfur and eight chlorine atoms among 1001 atoms of the lysozyme molecule.

The results of the direct methods trials using the same data at various resolutions suggest that it may be more productive to search for anomalous scatterers at low resolution, which is faster and seems to be more successful. This effect may be attributed to the fact that low-resolution intensities are stronger and can be measured with higher accuracy. Moreover, at low resolution the reflections with large normalized amplitudes are interconnected through more $\Sigma_2$ relationships, which may lead to more effective phase estimation and refinement.

All diffraction data are available upon request from ZD.

## References

Blundell, T. L.; Johnson, L. N.: Protein Crystallography. New York, Academic Press, (1976).

Burla, M. C.; Carrozzini, B.; Cascarano, G. L.; Giacovazzo, C.; Polidori, G.; Siliqi, D.: MAD phasing: probabilistic estimate of $|F_{oa}|$. Acta Crystallogr. **D58** (2002) 928–935.

Dauter, Z.; Dauter, M.; Dodson, E. J.: Jolly SAD. Acta Crystallogr. **D58** (2002) 494–506.

Dauter, Z.; Dauter, M.; de La Fortelle, E.; Bricogne, G.; Sheldrick, G. M.: Can anomalous signal of sulfur become a tool for solving protein crystal structures? J. Mol. Biol. **289** (1999) 83–92.

Dauter, Z.; Wilson, K. S.; Sieker, L. C.; Meyer, J.; Moulis, J. M.: Atomic resolution (0.94 Å) structure of *Clostridium acidurici* ferredoxin. Detailed geometry of [4Fe–4S] clusters in a protein. Biochemistry **36** (1997) 16065–16073.

Dauter, Z.; Li, M.; Wlodawer, A.: Practical experience with the use of halides for phasing macromolecular structures: a powerful tool for structural genomics. Acta Crystallogr. **D57** (2001) 239–249.

Devedjiev, Y.; Dauter, Z.; Kuznetsov, S. R.; Jones, T. L. Z.; Derewenda, Z. S.: Crystal structure of the human acyl protein thioesterase I from a single X-ray data set to 1.5 Å. Structure **8** (2000) 1137–1146.

Diederichs, K.; Karplus, P. A.: Improved R-factor for diffraction data analysis in macromolecular crystallography. Nature Struct. Biol. **4** (1997) 269–275.

Drenth, J.: Principles of protein X-ray crystallography. New York, Springer, (1994).

Fan, H. F.; Woolfson, M. M.; Yao, J. X.: New techniques of applying multi-wavelength anomalous scattering data. Proc. R. Soc. Lond. **A442** (1993) 13–32.

Hendrickson, W. A.: Determination of macromolecular structures from anomalous diffraction of synchrotron radiation. Science **254** (1991) 51–58.

Hendrickson, W. A.: Maturation of MAD phasing for the determination of macromolecular structures. J. Synchrotron Rad. **6** (1999) 845–851.

Hendrickson, W. A.; Teeter, M. M.: Structure of the hydrophobic protein crambin determined directly from the anomalous scattering of sulfur. Nature **290** (1981) 107–113.

Karle, J.: Some developments in anomalous dispersion for the structural investigation of macromolecular systems in biology. Int. J. Quant. Chem. **7** (1980) 357–367.

Li, J.; Derewenda, U.; Dauter, Z.; Smith, S.; Derewenda, Z. S.: Crystal structure of the *Eschericha coli* thioesterase II, a homolog of the human Nef binding enzyme. Nature Struct. Biol. **7** (2000) 555–559.

Matthews, B. W.: The extension of the isomorphous replacement method to include anomalous scattering measurements. Acta Crystallogr. **20** (1966a) 82–86.

Matthews, B. W.: The determination of the position of anomalously scattering heavy atom groups in protein crystals. Acta Crystallogr. **20** (1966b) 230–239.

Mukherjee, A. K.; Helliwell, J. R.; Main, P.: The use of *MULTAN* to locate the positions of anomalous scatterers. Acta Crystallogr. **A45** (1989) 715–718.

North, A. C. T.: The combination of isomorphous replacement and anomalous scattering data in phase determination of non-centrosymmetric crystals. Acta Crystallogr. **18** (1965) 212–216.

Otwinowski, Z.; Minor, W.: Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. **276** (1997) 307–326.

Rossmann, M. G.: The position of anomalous scatterers in protein crystals. Acta Crystallogr. **14** (1961) 383–388.

Sheldrick, G. M.: In: *Direct Methods for Solving Macromolecular Structures.* (Fortier S., ed.) Dordrecht, The Netherlands: Kluwer Academic Publishers, pp. 401–411 (1998).

Wang, B. C.: Resolution of phase ambiguity in macromolecular crystallography. Methods Enzymol. **115** (1985) 90–112.

Weiss, M. S.; Hilgenfeld, R.: On the use of merging R factor as a quality indicator for X-ray data. J. Appl. Cryst. **30** (1997) 203–205.