Quarterly Reviews of Biophysics **47**, 1 (2014), pp. 49–93. © Cambridge University Press 2014
doi:10.1017/S0033583514000018 Printed in the United States of America

Anomalous diffraction in crystallographic phase evaluation

Wayne A. Hendrickson[1,2]*

[1] Department of Biochemistry and Molecular Biophysics, and Department of Physiology and Cellular Biophysics, Columbia University, New York, NY 10032, USA
[2] New York Structural Biology Center, 89 Convent Avenue, New York, NY 10027, USA

Abstract. X-ray diffraction patterns from crystals of biological macromolecules contain sufficient information to define atomic structures, but atomic positions are inextricable without having electron-density images. Diffraction measurements provide amplitudes, but the computation of electron density also requires phases for the diffracted waves. The resonance phenomenon known as anomalous scattering offers a powerful solution to this phase problem. Exploiting scattering resonances from diverse elements, the methods of MAD (multiwavelength anomalous diffraction) and SAD (single-wavelength anomalous diffraction) now predominate for *de novo* determinations of atomic-level biological structures. This review describes the physical underpinnings of anomalous diffraction methods, the evolution of these methods to their current maturity, the elements, procedures and instrumentation used for effective implementation, and the realm of applications.

1. Introduction 50

2. Physical basis for anomalous scattering 51

3. Historical background 53
3.1. Discovery and description of anomalous scattering 53
3.2. Antecedents of MAD and SAD phasing for macromolecules 54
3.3. Synchrotron radiation 55

4. Theoretical foundation for phase evaluation from anomalous diffraction 55

5. Development of MAD 58
5.1. Beginnings of MAD phasing for macromolecules 58
5.2. Experimental procedures 60
5.3. Synchrotron beamlines 62
5.4. Analytical and computational procedures 63

6. Evolution of SAD 67
6.1. Crambin and phase ambiguity 67
6.2. Ambiguity resolution from density modification 69

* Author for correspondence: W. A. Hendrickson, Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY 10032, USA.
Tel.: 1 212 305 3456; Email: wayne@xtl.cumc.columbia.edu

https://doi.org/10.1017/S0033583514000018 Published online by Cambridge University Press

6.3. SAD becoming the method of choice 70
6.4. Multi-crystal enhancement of signal-to-noise ratios 71

7. Applications and changing practice 74
7.1. Realms of application 74
7.2. Changing practice in *de novo* phase evaluation 76
7.3. Future prospects 80

8. Supplementary material 84

9. Acknowledgements 84

10. References 84

1. Introduction

X-ray diffraction patterns from crystals of biological macromolecules contain decisive information for determining structures in atomic detail. Provided that the diffraction extends to Bragg spacings corresponding to ~3 Å resolution, a typical pattern will have a number of observable, symmetry-independent reflections that exceeds the number of x, y, z parameters needed to specify atomic positions for the non-hydrogen atoms of the structure. At even lower resolution (~5 Å), the number will still suffice to specify conformational torsion angles. Thus, in principle, diffraction data define the structure; however, the structural implication is through a large set of transcendental equations, each of which depends on all of the atomic parameters. This problem is intractable without having a model within the radius of convergence for refinement. There are no lenses for the hard x-rays needed for crystallography, as there are for visible light (optical lenses), electrons (magnetic lenses) or even soft x-rays (Fresnel zone plates); so the starting model cannot come from direct imaging. Fortunately, by the theory of diffraction, computational imaging by Fourier transformation is possible once phases as well as amplitudes are known for the diffracted waves. Hence, we face the phase problem – the need for ways to evaluate phases.

Crystallographic structure determination for biological macromolecules is influenced greatly by the small molecule tradition, which began with the solution of salt and mineral structures by trial-and-error approaches (Bragg, 1913b, 1914; Bragg & Bragg, 1913) and later exploited images of interatomic-vector distributions (Patterson, 1934, 1935). The first approaches that really entailed phase evaluation, made relevant when Fourier syntheses were introduced (Bragg, 1929), were by isomorphous replacement from a centrosymmetric series of phthalocyanines (Robertson, 1936) and by the heavy-atom method as first used for the structure of platinum phthalocyanine (Robertson & Woodward, 1940). In the heavy-atom method, phases were approximated by the contribution from the heavy atom and used to produce images of the lighter atoms. This approach was followed by the more general isomorphous replacement method for non-centrosymmetric structures (Bijvoet *et al.* 1948, 1949) as introduced in solving the structure of strychnine (Bokhoven *et al.* 1951). In this first application of isomorphous replacement, diffraction differences from substitution at a specific atomic site were exploited to obtain definitive phase information, albeit ambiguous in this case. Remarkably presaging the coming importance of selenium in biological crystallography, the substitution here was of selenate for sulfate. Finally, direct methods were derived for extracting phase information from relationships among crystallographic structure factors (Hauptman & Karle, 1953; Karle & Hauptman, 1950, 1956), and computer implementations of direct methods came to dominate in small molecule crystallography.

https://doi.org/10.1017/S0033583514000018 Published online by Cambridge University Press

Only isomorphous replacement was carried forward very directly from small molecule crystallography to macromolecular crystallography (Blow, 1958; Green *et al.* 1954). As adapted to include multiple heavy-atom derivatives for resolving inherent ambiguities, the method of multiple isomorphous replacement (MIR) dominated in the first decades of protein crystallography. Anomalous scattering was realized to provide significant ancillary phase information (Blundell & Johnson, 1976), which was used to augment MIR phasing into MIR with anomalous scattering (MIRAS) and to resolve ambiguities from single isomorphous replacement (SIR) as the SIRAS method (Blow & Rossmann, 1961; Herriott *et al.* 1970). Anomalous scattering information was also shown to suffice on its own for phase evaluation in the solution of the structure of crambin (Hendrickson & Teeter, 1981), which we now see as the first application of method of single-wavelength anomalous diffraction (SAD). Ultimately, with the availability of tunable x-rays from synchrotron sources, practical implementation of the method of multiwavelength anomalous diffraction (MAD) was realized (Hendrickson, 1985; Kahn *et al.* 1985). Apart from isomorphous replacement, the small molecule methods are too weak to be effective for most macromolecular problems; however, the small molecule tradition influenced biological crystallography in many other ways. Patterson methods and direct methods are crucial for solving the substructures of phasing elements needed in MIR/SIR and MAD/SAD applications. Most importantly, the method of molecular replacement of knowns into unknowns (Lattman *et al.* 1971; Rossmann, 1972), which now dominates in macromolecular crystallography, is in essence a phase borrowing closely related to that of the heavy-atom method. The crucial innovation here was the rotation function by Rossmann & Blow (1962).

Anomalous diffraction follows from the foundations of crystallography and the early development of protein crystallography, but it also has important roots in spectroscopy and, of course, an essential connection to synchrotron radiation. Synchrotron radiation, which was brought to the attention of biological crystallographers by Rosenbaum *et al.* (1971), provides a bright continuum of x-rays; and fine tuning to the sharp resonant transitions responsible for anomalous scattering is critical for MAD phasing experiments and for optimized SAD phasing experiments. Such tuning was brought into practice for x-ray absorption spectroscopy (Kincaid *et al.* 1975), and then adapted for anomalous x-ray diffraction experiments (Phillips *et al.* 1978).

The development and characteristics of MAD and SAD phasing have been described in several previous reviews (Blow, 2003; Fourme & Hendrickson, 1990; Hendrickson, 1991, 1999; Hendrickson & Ogata, 1997; Ogata & Hendrickson, 1995; Smith & Hendrickson, 2001; Walsh *et al.* 1999). It took some time for this methodology to mature, however. Meanwhile, in the past decade and after most of these reviews, MAD and SAD have overtaken MIR and SIR and they now dominate dramatically as methods for the *de novo* crystal structures of biological macromolecules. In this review, I describe the overall course of development and document the impact of anomalous diffraction methods.

2. Physical basis for anomalous scattering

The interaction of x-rays with matter results in diverse physical phenomena including absorption, fluorescence, refraction and scattering (Als-Nielsen & McMorrow, 2011). These physical properties can be applied to problems in chemistry, biology and medicine including radiographic diagnosis and crystallographic structure determination. Our concern here is with the latter, which is based on elastic scattering whereby the incident x-ray energy (and hence wavelength) remains unchanged. X-rays are scattered from electrons, and the classical description of elastic scattering

https://doi.org/10.1017/S0033583514000018 Published online by Cambridge University Press

worked out by J. J. Thomson (1906) for a free electron applies to good approximation for electrons in atoms as well. An incident electromagnetic wave induces sympathetic vibrations in the electron and the accelerated electron emits radiation in accord with Maxwell's equations. When integrated over the probability density for all electrons in an atom, as specified by the quantum-mechanical wave function, a total atomic scattering profile can be defined relative to the theoretical scattering from a free electron. This is the 'normal' atomic scattering factor, f°, which decreases with scattering angle as a function of $S = |\mathbf{S}| = 2\ \sin(\theta)/\lambda$, where $\mathbf{S} = h\mathbf{a}^* + k\mathbf{b}^* + l\mathbf{c}^*$ is the scattering vector for reflection $\mathbf{h}(h,k,l)$ from a crystal with reciprocal cell dimensions $\mathbf{a}^*,\mathbf{b}^*,\mathbf{c}^*$, θ is the Bragg scattering angle, and λ is the x-ray wavelength.

In reality, electrons are not passively free to respond to incident radiation; they are tied up into electronic orbitals at characteristic frequencies. As the x-ray oscillatory frequency approaches an atomic orbital frequency, the induced electron vibrations can resonate with the natural oscillations of the bound electrons, or more precisely with transitions from bound states to other accessible orbitals. This perturbs the atomic scattering (Als-Nielsen & McMorrow, 2011; James, 1948). Such interactions also lead to photoelectric absorption whenever the x-ray energy exceeds the orbital binding energy, the absorption edge. This generates an ionized atom and a photoelectron; and such photoelectrons may recombine with the resulting ions to produce fluorescent x-rays with an energy characteristic of the particular electronic transition. The scattering perturbations due to orbital interactions add to the 'normal' Thomson scattering with increments, f^Δ, that have both amplitude and phase shift components, $|f^\Delta|$ and δ, and corresponding real and imaginary components, respectively f' and f''. Thus, the true atomic scattering factor, f, is complex:

$$f = f^\circ + f^\Delta = f^\circ + |f^\Delta| e^{i\delta} = f^\circ + f' + if''. \quad (1)$$

The f^Δ scattering perturbations are known as 'anomalous' scattering even though all electrons of interest are in electronic orbitals and truly are the norm. Physically, anomalous scattering is intimately connected to resonance and to absorption, and it is dispersive since the effect depends strongly on x-ray energy (E). Since anomalous scattering derives from core electrons, it is essentially independent of scattering angle; however, unlike normal scattering, it is a function of wavelength. Thus, $f^\Delta = f(\lambda) \neq f(S)$ whereas $f^\circ = f(S) \neq f(\lambda)$.

The spectra $f'(\lambda)$ and $f''(\lambda)$ can be evaluated for isolated atoms from quantum mechanical calculations (Cromer & Liberman, 1970), but these do not apply for many resonant transitions which are from core atomic orbitals to unoccupied molecular orbitals. Fortunately, however, these factors can be evaluated experimentally from the connection between scattering, absorption and fluorescence. By Fresnel diffraction theory, the f° scattered wave is $\frac{\pi}{2}$ out of phase with the incident wave (James, 1948), and f'' is out by another $\frac{\pi}{2}$; therefore, in the forward direction, the f'' component interferes destructively with the incident beam. This is the physical basis for absorption; $f''(E) = K\,\mu(E)\cdot E$, where K is a function only of physical constants and μ is the atomic absorption coefficient. The direct proportionality of fluorescence to absorption then makes the $f''(\lambda)$ spectrum accessible [$f''(\lambda)$ and $f''(E)$ are equivalent since $E = hc/\lambda$; $E(\text{keV}) = 12.3984/\lambda(\text{Å})$], and $f'(\lambda)$ follows by Kramers–Kronig transformation (Hendrickson *et al.* 1988; Lye *et al.* 1980). The full range of x-rays for diffraction experiments (~3.5–35 keV or 0.35–3.5 Å) includes K and/or L absorption edges for all elements of $Z \geqslant 20$ (Ca and upwards) (Hendrickson & Ogata, 1997). The light atoms of macromolecules (H, C, N and O) have negligible anomalous scattering in this range, but the f'' values for S and P are sufficiently substantial to be useful. Resonant enhancements can be pronounced even at K-edges as in selenoethers (Fig. 1*a*), and they can be dramatic at the M-edges of actinides and at the L_{III}-edges of lanthanides (Fig. 1*b*).

https://doi.org/10.1017/S0033583514000018 Published online by Cambridge University Press

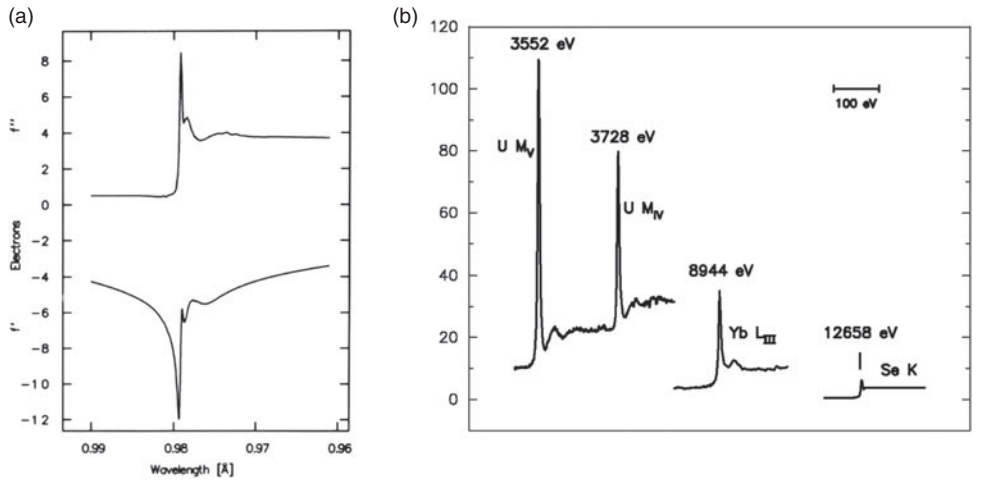


Fig. 1. Anomalous scattering factor profiles. (*a*) Se K-edge anomalous scattering factors. These data derive from an x-ray absorption spectrum of selenomethionyl thioredoxin (Hendrickson *et al.* 1990). (*b*) Imaginary components of anomalous scattering at different absorption edges. The U M_{IV} and U M_V edges are from uranyl nitrate (Liu *et al.* 2001), the Yb L_{III} edge is from a ytterbium-derivatized crystal of N-cadherin (Shapiro *et al.* 1995), and the Se K edge is from a crystal of selenomethionyl human choriogonadotropin (Wu *et al.* 1994). Reproduced from Hendrickson *et al.* (1990) for (*a*) with permission from the *EMBO Journal*, © 1990 EMBO, and from Liu *et al.* (2001) for (*b*) with permission from the *Proceedings of the National Academy of Sciences of the United States of America*, © 2001 National Academy of Sciences.

3. Historical background

3.1 Discovery and description of anomalous scattering

The first recorded evidence for anomalous x-ray scattering was from Mark & Szilard (1925) who showed that diffraction from RbBr crystals differed selectively depending on the wavelength of irradiating x-rays. The RbBr structure is like that of rock salt (NaCl) and, as Rb^+ and Br^- ions are isoelectronic and have equivalent f° scattering factors, only even orders of [111] diffraction were observed with x-rays from Fe, Co, Cu or Zn anodes, which have x-ray energies (7.1–9.7 keV) remote from the absorption edges of Rb or Br. With Sr Kα radiation, however, both the (1,1,1) and (3,3,3) diffraction lines were also observed; here the stimulating x-rays (14.1 keV) fall between the K absorption edges of Br (13.5 keV) and Rb (15.2 keV) whereby the anomalous scattering appreciably and differentially affects the net Br and Rb scattering.

An especially celebrated demonstration of anomalous scattering came from an elegant experiment by Coster *et al.* (1930) showing the breakdown of Friedel's law, $I(h,k,l) \neq I(-h,-k,-l)$, in diffraction from zincblende (ZnS). Using a gold anode to produce x-rays, these investigators found dramatic differences in the ratio of intensities in Au Lα_1 to Au Lα_2 lines seen for (1,1,1) as compared with (−1,−1,−1) reflections and also for (3,3,3) *versus* (−3,−3,−3) reflections. The Zn K absorption edge (9.66 keV) falls between the two Au L lines, with Au Lα_2 (9.63 keV) being below the edge and Au Lα_1 (9.71 keV) being above the edge; thus, since f'' (Zn) is then large at Lα_1, the differences in intensities between Friedel mates are also large. According to Ramaseshan (1993), this influential study was inspired by similar zincblende experiments, albeit more preliminary, based on tungsten Lβ lines (Nishikawa & Matukawa, 1928).

https://doi.org/10.1017/S0033583514000018 Published online by Cambridge University Press

Quantification of the dispersion of x-rays came from accurate measurements of atomic scattering factors as a function of x-ray wavelength in the neighborhood of an absorption edge. Observations by Wyckoff (1930) on nickel oxide with MoKα, CuKα and NiKα radiations were especially notable in this context, showing dispersion for Ni scattering factors but not for O scattering factors.

The observation of anomalous dispersion of x-rays prompted theoretical investigations of the phenomena (Kallman & Mark, 1927); and, to an extent, theoretical studies (Prins, 1928) spurred the experiments. The quantum dispersion theory (Kramers & Heisenberg, 1925) was extended to analysis of the refraction, absorption and dispersion of x-rays (Kronig, 1926; Kronig & Kramers, 1928). The most complete quantum elaboration came in the theory of Hönl (1933), which used atomic wave functions to obtain oscillator strengths and from them the photoelectric absorption cross-sections were needed to calculate scattering factor corrections.

3.2 Antecedents of MAD and SAD phasing for macromolecules

Phenomena associated with anomalous scattering fascinated physicists as being manifestations of the quantum nature of atoms and for developing x-ray diffraction theory. Even though phase shifts were understood to accompany the resonances responsible for the anomalies (Coster *et al.* 1930), there seems to have been no motivation for physicists at the time to exploit these effects for structure determination. Missing such opportunity is understandable since phase evaluation came only later to be employed in structure determination (Robertson, 1936) once Fourier approaches had been introduced (Bragg, 1929; see discussion by Hendrickson, 2013); however, it is surprising that it took so long for this inspiration to strike (Bijvoet, 1949). The lag is all the more remarkable since Bijvoet himself had been instrumental in prompting Ewald to reconsider his 'proof' of Friedel's Law (Ewald & Hermann, 1927).

Reduction to practice for the idea of phase evaluation from anomalous scattering was slow to come, although Bijvoet did develop the method of isomorphous replacement in the same time frame (Bijvoet *et al.* 1948, 1949). Anomalous differences tend to be smaller than isomorphous differences and, to be general, there are issues of phase ambiguity; however, the deviations from Friedel symmetry in diffraction from rubidium tartrate did suffice to determine absolute configuration for the first time (Bijvoet *et al.* 1951). Bijvoet later reviewed these influential and pioneering uses of anomalous diffraction (Bijvoet, 1954).

The phase problem was a preoccupying concern in the intellectual ferment of 1950s crystallography, and anomalous scattering stimulated several new formulations for phase determination (Herzenberg & Lau, 1967; Karle, 1967; Mitchell, 1957; Okaya & Pepinsky, 1956; Okaya *et al.* 1955; Ramachandran & Raman, 1956; Raman, 1959; Ramaseshan & Venkatesan, 1957; Singh & Ramaseshan, 1968). Practical implementations also followed; however, at the time, these were limited to the characteristic wavelengths available from the anode coatings of Roentgen-like x-ray tubes. Nevertheless, pioneering multi-wavelength tests were performed to resolve the phase ambiguity presented by non-centrosymmetric structures. Notably, Ramaseshan *et al.* (1957) bracketed the Mn K absorption edge with FeKα and CoKα radiations and solved the crystal structure of $KMnO_4$; and Hoppe & Jakubowski (1971) carried out two-wavelength tests on a fly hemoglobin using CoKα and NiKα x-ray, which similarly lie on either side of the Fe K edge of heme.

https://doi.org/10.1017/S0033583514000018 Published online by Cambridge University Press

Ramaseshan (1993) has given an engaging history of early work leading to multiwavelength phasing for biological molecules.

3.3 Synchrotron radiation

The bremsstrahlung emission from x-ray tubes does afford access to a continuous x-ray spectrum, indeed, this continuum proved to be the origin of the first x-ray patterns (Bragg, 1913a; Friedrich *et al.* 1912); however, the far more intense characteristic x-ray emissions from electron-bombarded anodes became the mainstay for structure determinations. By contrast, the continuum of synchrotron radiation is exceptionally bright across a wide spectrum determined by the critical energy of the source. Soon after the realization that synchrotron radiation could be useful for biological diffraction (Rosenbaum *et al.* 1971), experiments also followed to extract structural information about metal coordination in metalloproteins by x-ray absorption spectroscopy (Cramer & Hodgson, 1979; Kincaid *et al.* 1975) and to extract anomalous scattering factors from the finely tuned near-edge variations in x-ray absorption (Phillips *et al.* 1978; Templeton *et al.* 1980, 1982).

That synchrotron measurements of anomalous diffraction from appropriately derivatized crystals could be used effectively in structure determination was appreciated immediately (Lye *et al.* 1980; Phillips & Hodgson, 1980; Templeton *et al.* 1980) and a first application came with the test structure of Tb^{3+}-derivatized parvalbumin (Kahn *et al.* 1985). We had also learned from our experience in determining the structure of crambin from the weak anomalous scattering of CuKα x-rays by its intrinsic sulfur atoms that native anomalous scattering can suffice (Hendrickson & Teeter, 1981), and our experience in x-ray absorption spectroscopy of hemerythrin (Hendrickson *et al.* 1982) indicated that iron proteins might be excellent candidates. Thus, we successfully used lamprey hemoglobin as a first test subject (Hendrickson, 1985; Hendrickson *et al.* 1988).

We named the new method multiwavelength anomalous diffraction, giving the acronym MAD for its parallel with MIR, then the dominant phasing method. In both cases, information from multiple sets of measurements (at differing wavelengths for MAD and from different isomorphs for MIR) is used to break the trigonometric phase ambiguity that occurs in general when only a single set of measurements is used. The first use of the term MAD, to the best of my knowledge, was in a main lecture at the IUCr Congress in Hamburg in 1984 (Hendrickson, 1984), and this term was then adopted in the first publications describing structure determinations by the method (Harada *et al.* 1986; Hendrickson, 1985; Kahn *et al.* 1985).

4. Theoretical foundation for phase evaluation from anomalous diffraction

Our analysis for deriving phase information from measurements of anomalous diffraction is based on the fundamental additivity of scattered x-ray waves, whereby one here exploits the interference between scattering from anomalous centers and that from all other atoms. We distinguish the relatively few anomalous scatterers (A), for which the anomalous component (f^{Δ}) is substantial, from the many 'normal' scatterers (N), for which f^{Δ} is considered negligible, and the sum of the two gives the total from all atoms (T) in the structure. Since deviations from normal diffraction arise only from the anomalous scatterers, it is possible to isolate the anomalous contribution to the total diffraction and from it to deduce positions for the anomalous scatterers. With these scattering components in hand, both in phase and amplitude, then by the principles of interference it becomes possible to deduce the otherwise unknown phases for diffracted waves

https://doi.org/10.1017/S0033583514000018 Published online by Cambridge University Press

from the total structure. We perform this analysis in a unified treatment of all observations, using an algebraic approach developed in principle by Karle (1980) and reduced to practice after reformulation (Hendrickson, 1985, 1991; Hendrickson *et al.* 1988).

Each kind of anomalous scatterer (k), having distinctive scattering factors, contributes to the total diffraction as the sum of a wavelength-independent component due to its normal scattering (f°) and a wavelength-variant component due to its anomalous scattering ($f^\Delta = f' + \mathrm{i}\, f''$). Thereby,

$$ {}^\lambda F_{A_k} = {}^\circ F_{A_k} + {}^\lambda F'_{A_k} + i\, {}^\lambda F''_{A_k} = \left[1 + \left(f'/f^\circ\right) + i\left(f''/f^\circ\right)\right] {}^\circ F_{A_k}, \tag{2} $$

where ${}^\lambda F_{A_k}$ denotes the complete wavelength-dependent contribution of atoms A_k, which is composed from ${}^\circ F_{A_k} = \mathrm{f}(f^\circ_k)$, ${}^\lambda F'_{A_k} = \mathrm{f}\left(f'_k\right)$ and ${}^\lambda F''_{A_k} = \mathrm{f}(f''_k)$. Furthermore, since the normal scattering component of each anomalous scatterer contributes along with all normal scatterers to the total scattering, ${}^\circ F_T = {}^\circ F_N + \sum_k {}^\circ F_{A_k}$. Thus, diffraction measurements associated with a specified Bragg reflection $\mathbf{h}(h,k,l)$ made at a particular wavelength λ are given by

$$ {}^\lambda F(\mathbf{h}) = {}^\circ F_T(\mathbf{h}) + \sum\nolimits_k \left[\left(f'/f^\circ\right) + i\left(f''/f^\circ\right)\right] {}^\circ F_{A_k}(\mathbf{h}), \tag{3} $$

where ${}^\lambda F(\mathbf{h})$ is the wavelength-dependent total structure factor for reflection $\mathbf{h}$, and ${}^\circ F_T(\mathbf{h}) = |{}^\circ F_T| \exp(i\, {}^\circ\phi_T)$ and ${}^\circ F_{A_k}(\mathbf{h}) = \left|{}^\circ F_{A_k}\right| \exp\left(i\, {}^\circ\phi_{A_k}\right)$ are corresponding wavelength-invariant components.

Equation (3) is general as the kinds of anomalous scatterers can be arbitrarily broad; however, in most cases the number of effective kinds is limited, often just to one. Usually, the light atoms (H, C, N and O) of macromolecules are treated as 'normal' scatterers having scattering factors $f = f^\circ$ even though an anomalous scattering component ($f^\Delta = f' + i\, f''$) is never strictly absent. Moreover, for measurements made at higher energy (>7 keV) the components from S and P atoms can be considered negligible relative to those at resonance from heavier elements such as Fe, Se, Hg or Yb. Even when multiple kinds of anomalous scatterers are present, one kind usually dominates; and to a useful initial approximation the equation can be so limited.

The observable quantity in a diffraction experiment is the diffracted intensity $I(\mathbf{h}) = K\, |F(\mathbf{h})|^2$; and $|F|^2$ expands by Eq. (2) for the case of a single kind of anomalous scatterer to

$$ \begin{aligned} \left|{}^\lambda F(\pm\mathbf{h})\right|^2 = |{}^\circ F_T(\mathbf{h})|^2 + a(\lambda)|{}^\circ F_A(\mathbf{h})|^2 + b(\lambda)|{}^\circ F_T(\mathbf{h})||{}^\circ F_A(\mathbf{h})|\cos\left({}^\circ\phi_T - {}^\circ\phi_A\right) \\ \pm c(\lambda)|{}^\circ F_T(\mathbf{h})||{}^\circ F_A(\mathbf{h})|\sin\left({}^\circ\phi_T - {}^\circ\phi_A\right), \end{aligned} \tag{4} $$

where

$$ a(\lambda) = \left(\left|f^\Delta\right|/f^\circ\right)^2; \quad b(\lambda) = 2\left(f'/f^\circ\right); \quad \text{and} \quad c(\lambda) = 2\left(f''/f^\circ\right). $$

Only the wavelength-invariant normal scattering factors f° contribute to ${}^\circ F_T$ and ${}^\circ F_A$, amplitudes and phases, including all atoms for ${}^\circ F_T$ but just one kind of anomalous scatterer for ${}^\circ F_A$. All wavelength dependence is separated into factors $a(\lambda)$, $b(\lambda)$ and $c(\lambda)$ defined by the spectral dependence of anomalous scattering.

Equations (4) encode an elegant and unified theory that can afford a definitive solution of the phase problem, which is implemented in the MADSYS system of programs (Hendrickson, 1985, 1991). Normal scattering factors f° are known from quantum mechanical calculation, and the anomalous scattering factors f' and f'' can be deduced from a combination of x-ray absorption spectra and theory (Hendrickson *et al.* 1988), whereby wavelength-dependent factors a(λ), b(λ) and c(λ) are determined. Alternatively, these factors can be refined from rough approximations

https://doi.org/10.1017/S0033583514000018 Published online by Cambridge University Press

(Fanchon & Hendrickson, 1990; Templeton & Templeton, 1982; Weis *et al.* 1991). The set of observations, $\{|{}^{\lambda}F(\mathbf{h})|^2\}$, of Bijvoet mates at various wavelengths constitute, for each reflection $\mathbf{h}$, a set of simultaneous equations (Eq. 4) that can be solved subject to trigonometric identities to yield the wavelength-invariant structure-factor quantities: $|{}^{\circ}F_{\mathrm{T}}(\mathbf{h})|$, $|{}^{\circ}F_{\mathrm{A}}(\mathbf{h})|$ and $\Delta\phi = ({}^{\circ}\phi_{\mathrm{T}} - {}^{\circ}\phi_{\mathrm{A}})$. The set of $|{}^{\circ}F_{\mathrm{A}}(\mathbf{h})|$ from all reflections can then be used to determine and refine the atomic substructure of anomalous scattering centers, $\{\mathbf{r_A}\}$, by relatively straightforward Patterson or direct methods. Phases calculated from this substructure solve the phase problem, ${}^{\circ}\phi_{\mathrm{T}} = \Delta\phi + {}^{\circ}\phi_{\mathrm{A}}^{\mathrm{calc}}$, provided that the correct enantiomorph, $\{\pm\mathbf{r_A}\}$, has been determined. In practice, the solution is represented as a phase probability distribution (Pähler *et al.* 1990).

In addition to providing a possible algebraic solution, Eq. (4) also provides useful, approximation-free descriptions of the deviations from normal diffraction due to anomalous scattering. It is evident that $|{}^{\lambda}F(\mathbf{h})|$ is independent of wavelength in the absence of anomalous scattering ($f^{\Delta} = 0$); and then structure factor values are independent of wavelength and identical for Friedel mates $|{}^{\lambda}F(\pm\mathbf{h})|$ or their symmetry equivalents (called Bijvoet mates). Deviations of $|{}^{\lambda}F(\pm\mathbf{h})|$ from $|{}^{\circ}F_{\mathrm{T}}(\mathbf{h})|$ are often small and best captured in salient differences. The Bijvoet difference is directly proportional to f'' and to $\sin(\Delta\phi)$:

$$\left|{}^{\lambda}F(\mathbf{h})\right|^2 - \left|{}^{\lambda}F(-\mathbf{h})\right|^2 = 2\,c(\lambda)|{}^{\circ}F_{\mathrm{T}}(\mathbf{h})||{}^{\circ}F_{\mathrm{A}}(\mathbf{h})|\sin\left({}^{\circ}\phi_{\mathrm{T}} - {}^{\circ}\phi_{\mathrm{A}}\right). \tag{5}$$

If anomalous scattering effects are relatively small, as commonly happens, then $|{}^{\circ}F_{\mathrm{T}}(\mathbf{h})| \approx (|{}^{\lambda}F(\mathbf{h})| + |{}^{\lambda}F(-\mathbf{h})|)/2$ and one obtains the Bijvoet-difference equation used for the SAD phasing analysis of the crambin structure (Hendrickson & Teeter, 1981):

$${}^{\lambda}\Delta F_{\pm\mathbf{h}} \approx \left|{}^{\lambda}F(\mathbf{h})\right| - \left|{}^{\lambda}F(-\mathbf{h})\right| = c(\lambda)|{}^{\circ}F_{\mathrm{A}}(\mathbf{h})|\sin\left({}^{\circ}\phi_{\mathrm{T}} - {}^{\circ}\phi_{\mathrm{A}}\right), \tag{6}$$

from which there are two possible values of ${}^{\circ}\phi_{\mathrm{T}}$ when the atomic substructure of anomalous scattering centers, $\{\mathbf{r_A}\}$, is known. Absent data from multiple wavelengths, other added information such as from density-modification methods is needed for resolution of this phase ambiguity.

Dispersive differences between two wavelengths, λ_{i} and λ_{j}, are more complicated, but they can be described as the differences between the respective intensity means, $\langle|{}^{\lambda}F(\mathbf{h})|^2\rangle = (|{}^{\lambda}F(\mathbf{h})|^2 + |{}^{\lambda}F(-\mathbf{h})|^2)/2$:

$$\begin{aligned}\left\langle\left|{}^{\lambda_{\mathrm{i}}}F(\mathbf{h})\right|^2\right\rangle - \left\langle\left|{}^{\lambda_{\mathrm{j}}}F(\mathbf{h})\right|^2\right\rangle = {} & \left[a(\lambda_{\mathrm{i}}) - a(\lambda_{\mathrm{j}})\right]|{}^{\circ}F_{\mathrm{A}}(\mathbf{h})|^2 \\ & + \left[b(\lambda_{\mathrm{i}}) - b(\lambda_{\mathrm{j}})\right]|{}^{\circ}F_{\mathrm{T}}(\mathbf{h})||{}^{\circ}F_{\mathrm{A}}(\mathbf{h})|\cos\left({}^{\circ}\phi_{\mathrm{T}} - {}^{\circ}\phi_{\mathrm{A}}\right).\end{aligned} \tag{7}$$

Unless $|{}^{\circ}F_{\mathrm{A}}(\mathbf{h})|$ is on the order of $|{}^{\circ}F_{\mathrm{T}}(\mathbf{h})|$, which only happens for very strong anomalous scattering, the second term dominates and $|{}^{\circ}F_{\mathrm{T}}(\mathbf{h})| \approx (\langle|{}^{\lambda_{\mathrm{i}}}F(\mathbf{h})|\rangle + \langle|{}^{\lambda_{\mathrm{j}}}F(\mathbf{h})|\rangle)/2$. In such typical cases, the dispersive difference is thereby:

$$\Delta F_{\Delta\lambda} \approx \left\langle\left|{}^{\lambda_{\mathrm{i}}}F(\mathbf{h})\right|\right\rangle - \left\langle\left|{}^{\lambda_{\mathrm{j}}}F(\mathbf{h})\right|\right\rangle = \tfrac{1}{2}\left[b(\lambda_{\mathrm{i}}) - b(\lambda_{\mathrm{j}})\right]|{}^{\circ}F_{\mathrm{A}}(\mathbf{h})|\cos\left({}^{\circ}\phi_{\mathrm{T}} - {}^{\circ}\phi_{\mathrm{A}}\right), \tag{8}$$

whereas Bijvoet differences are proportional to $f''(\lambda_{\mathrm{i}})$ and $\sin(\Delta\phi)$, dispersive differences are proportional to $|f'(\lambda_{\mathrm{i}}) - f'(\lambda_{\mathrm{j}})|$ and to $\cos(\Delta\phi)$, which nicely demonstrates the orthogonality of information needed for definitive phase evaluation.

It is helpful to have estimates of the strength of signal expected in a MAD or SAD experiment, and the approximations of Eqs. (6) and (8) for Bijvoet and dispersive differences make such estimates possible. We do this by obtaining the ratios of relevant averaged anomalous differences relative to the averaged total diffraction (Hendrickson, 1991), where averages are taken

https://doi.org/10.1017/S0033583514000018 Published online by Cambridge University Press

as root-mean-squared (rms) values over all data with extrapolation to zero scattering angle. Since the expected value for $\lim(\mathbf{h}\rightarrow 0)\ \langle |{}^{\lambda}F(\pm\mathbf{h})|^2\rangle \approx \langle |{}^{\circ}F_{\mathrm{T}}(\mathbf{h})|^2\rangle = \sum f^{\circ}(0)^2 \approx N_{\mathrm{T}}Z_{\mathrm{eff}}^2$ where N_{T} is the number of non-hydrogen atoms and Z_{eff} is the effective atomic number for an average non-hydrogen atom. Thus, rms $(F_{\mathrm{T}}) \approx \sqrt{N_{\mathrm{T}}}\, Z_{\mathrm{eff}}$. $Z_{\mathrm{eff}} = 6.7$ for proteins. A similar averaging pertains to differences from the anomalous substructure, but here the averaging must include the trigonometric factors, where $\langle \sin(\theta)\rangle = \langle \cos(\theta)\rangle = \frac{1}{2}$. Taken all together, we obtain the ratios

$$\mathrm{Rms}\left({}^{\lambda}\Delta F_{\pm\mathrm{h}}\right)/\mathrm{Rms}\,(F_{\mathrm{T}}) \approx \left(N_{\mathrm{A}}/2\,N_{\mathrm{T}}\right)^{1/2} \times \left(2f''(\lambda)/Z_{\mathrm{eff}}\right) \tag{9}$$

and

$$\mathrm{Rms}\,(\Delta F_{\Delta\lambda})/\mathrm{Rms}\,(F_{\mathrm{T}}) \approx \left(N_{\mathrm{A}}/2\,N_{\mathrm{T}}\right)^{1/2} \times \left(\left|f'(\lambda_{\mathrm{i}}) - f'(\lambda_{\mathrm{j}})\right|/Z_{\mathrm{eff}}\right). \tag{10}$$

MAD experiments can be thought of as *in situ* MIR experiments, carried out with physical instead of chemical replacements, and they can be analyzed accordingly (Phillips & Hodgson, 1980) or by earlier alternative methods (Okaya & Pepinsky, 1956). The dispersive differences defined by Eq. (7) are like those in isomorphous replacement and the Bijvoet differences defined by Eq. (6) are like the anomalous differences added to give SIRAS and MIRAS.

5. Development of MAD

The antecedents of MAD analyses for macromolecules did include the remarkable two-wavelength test on *Chrionomus* hemoglobin by Hoppe & Jakubowski (1971) and the single-wavelength structure of crambin (Hendrickson & Teeter, 1981), both using characteristic Kα lines from metal anodes; however, the true beginnings of MAD came with developments from synchrotron radiation.

5.1 Beginnings of MAD phasing for macromolecules

The first published MAD analysis of a protein structure was that for a parvalbumin from the toadfish *Opsanus tau*. Kahn *et al.* (1985) performed a three-wavelength experiment at LURE, taking advantage of the strong white-line resonance of Tb^{3+}, which was substituted into Ca^{2+} sites. Measurements were made at the absorption maximum ($\lambda = 1.649$ Å), at the inflection point ($\Delta\lambda = 0.0011$ Å) and at a shorter wavelength ($\Delta\lambda = -0.0021$ Å). These wavelengths corresponded to an energy of 7.5137 keV at the L_{III}-edge, an energy 5 eV beyond the edge at the absorption peak, and a slightly remote energy just 15 eV above the edge. Scattering factor values were first estimated from the scan of differences from a particularly strong Bijvoet pair and then refined against the complete data set. An electron-density map computed at 2.3 Å resolution compared favorably with the structure solved independently by molecular replacement.

In the same timeframe, experiments were conducted at SSRL using lamprey hemoglobin as a test problem, and it was results from these experiments that were used to describe MAD phasing at the IUCr Congress in 1984 (Hendrickson, 1984). The lamprey hemoglobin analysis was first described at 5.5 Å resolution (Hendrickson, 1985) and then at 3.0 Å resolution (Hendrickson *et al.* 1988). This was a four-wavelength experiment that included the inflection (edge) and peak positions at the Fe K-edge as well as a remote low- and high-energy positions, respectively 245 and 1140 eV below and above the edge. Fluorescence measurements of x-ray absorption were scaled to fit theoretical f'' values outside the edge region (6.68–7.40 keV) and then used to derive f' values by Kramers–Kronig transformation (Kronig & Kramers, 1928). The extended span of energies here was chosen, as illustrated in Fig. 2, to enhance the dispersive signals which depend

https://doi.org/10.1017/S0033583514000018 Published online by Cambridge University Press

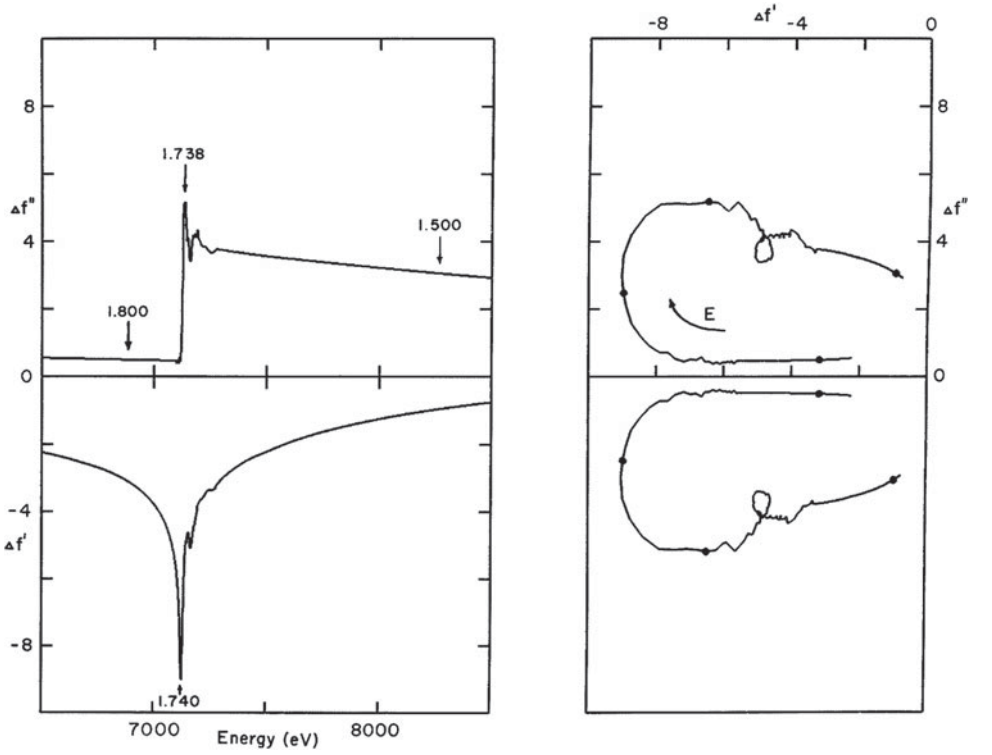


Fig. 2. Experimental anomalous scattering factors for lamprey hemoglobin. (Left) Imaginary and real components, above and below respectively, of Fe scattering factors in type *D*2 lamprey hemoglobin, measured with **E** parallel **b**. Experimental fluorescence data were fitted into theoretical atomic scattering factor spectra, retaining experimental values in the near-edge region. Wavelengths of the diffraction measurements are indicated by arrows. (Right) Herzenberg–Lau scattering factor profiles (Herzenberg & Lau, 1967). Wavelengths of diffraction measurements are indicated by solid circles. Reproduced from Hendrickson *et al.* (1988) with permission from Proteins, © 1988 Alan R. Liss, Inc.

on $\Delta f' = f'(\lambda_i) - f'(\lambda_j)$. Crystals were mounted in capillaries and maintained at ~10 °C, and diffraction data were measured with a multi-wire electronic area detector.

Other early MAD experiments were also conducted at low resolution or on test systems (Harada *et al.* 1986; Korzun, 1987), but the focus of attention soon moved to true applications. The first of these to be reported was that for cucumber basic blue protein by Guss *et al.* (1988), a structure that had eluded attempts by other methods. This study used the same experimental approach and instrumentation as that employed in the lamprey hemoglobin work, but in this case copper atoms produced the anomalous scattering. A four-wavelength experiment was conducted as for lamprey hemoglobin, here ranging from 949 eV below the edge to 1040 eV above the edge. The MAD data were measured to 2.5 Å spacings, but then used to compute electron-density maps at 3.0 Å resolution from which the structure was interpreted and subsequently refined at 1.8 Å resolution.

Our own first application to a novel structure was that to streptavidin, which we solved as a complex with selenobiotin (Hendrickson *et al.* 1989). Experiments were conducted both at SSRL and at a diffractometer facility at Photon Factory beamline 14A. Whereas high-energy resolution at SSRL permitted a characterization of anisotropy in anomalous scattering from seleonbiotinyl streptavidin, resulting diffraction data measured there were of insufficient quality for an

https://doi.org/10.1017/S0033583514000018 Published online by Cambridge University Press

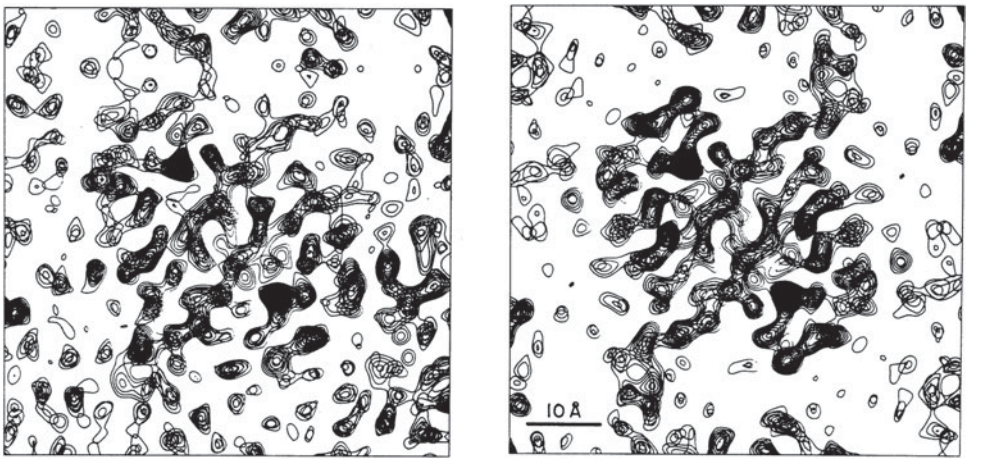


Fig. 3. MAD-phased electron-density distributions for selenobiotinyl streptavidin. (Left) Map at 3.3 Å resolution based exclusively on MAD phasing ($m = 0.78$). (Right) Map at 3.3 Å resolution after twofold molecular averaging refinement of initial phases ($m = 0.90$). Reproduced from Hendrickson *et al.* (1989) with permission from the *Proceedings of the National Academy of Sciences of the United States of America*, © 1989 National Academy of Sciences.

interpretable result. Despite poor energy resolution at PF 14A, a three-wavelength MAD experiment (inflection, peak and a remote point 1118 eV above the edge) gave a readily interpretable map that was further improved by averaging of non-crystallographic symmetry (Fig. 3). The excellence of this result based on a single Se atom from selenobiotin also inspired us to develop selenomethionyl proteins, which proved to provide great generality for MAD analyses (Hendrickson *et al.* 1990).

After these initial successes, we undertook additional MAD phasing experiments using beamlines at various synchrotrons to solve other novel structures, sometimes in collaborations and often through the use of selenomethionine. The resulting structures included interleukin 1-α (Graves *et al.* 1990), ribonuclease H (Yang *et al.* 1990), mannose-binding C-type lectin (Weis *et al.* 1991), urechis hemoglobin (Kolatkar *et al.* 1992) and tenascin (Leahy *et al.* 1992). During this period, we also used MAD data measured from gold-anode bremsstrahlung for our structure of CD4 (Ryu *et al.* 1990), and others reported a MAD-phased structure of cytochrome c553 (Nakagawa *et al.* 1990). All of these early results were obtained before cryoprotection was introduced to reduce radiation damage. Subsequently, with additional advances in instrumentation and analytic procedures, MAD usage expanded rapidly as described below in Section 7.

5.2 Experimental procedures

Experimental procedures introduced in early applications of MAD phase evaluation had considerable influence on the practices that followed as MAD blossomed. On the other hand, technical developments in available instrumentation and procedures as well as advances in data processing and analytical procedures have led to substantial evolution in standard practices, in many ways simplifying the process. Nevertheless, MAD presents several experimental alternatives to consider.

A primary consideration concerns the choice of atomic element and electronic resonance for the MAD experiment. In an earlier review (Hendrickson & Ogata, 1997), we tabulated all of the

https://doi.org/10.1017/S0033583514000018 Published online by Cambridge University Press

absorption edges within what we defined to be an accessible energy range of 3.5–35 keV (3.5–0.35 Å in wavelength) showing that all elements from atomic number 20 (Ca) to 92 (U) have K or L_{III} edges in this range, and sometimes both. Moreover, elements having edges within a more convenient central range (7–17 keV, 1.7–0.7 Å) include many transition elements and other metals (e.g. Fe and Mo) found in metalloproteins, elements such as Se and Br that can be introduced covalently into biological molecules, many of the lanthanides used to substitute for Mg^{2+} or Ca^{2+}, and nearly all 'heavy atoms' conventionally used for MIR and SIR. Whereas MIR analyses had focused on conventional heavy atoms with $Z \geqslant 73$ (Ta), the earliest MAD experiments included successes with much lighter elements such as Fe, Cu and Se ($Z = 26$, 29 & 34, respectively). Later MAD and SAD experiments have covered a broad range of elements, but systematic incorporations of Se in selenomethionyl proteins and Br in brominated nucleic acids have dominated for these respective macromolecules.

Given an element and its resonance, the next questions for experimental design are about how many and which wavelengths to use. The typical answer for MAD has been three or four. Since f'' factors are proportional to atomic absorption, fluorescence spectra of atomic absorption are important for decisions on choice of wavelengths. To be optimal, one wavelength should be at the peak of f'' since by Eq. (9) this specifies the greatest Bijvoet differences. Similarly, by Eq. (10) for dispersive differences, one seeks to maximize the f' difference between two wavelengths. One of the two is obviously at the minimum in the f' profile, which is identified at the ascending inflection point in the f'' profile (Figs 1*a* and 2), but there are multiple options for the second. For lanthanide L_{III} or uranium M edges (Fig. 1*b*), the descending inflections of the white-line transitions mark advantageous local maxima; for K-edge experiments, the near-edge local maxima are less pronounced (Fig. 1) and other wavelengths can be chosen as low or high-energy remotes (Fig. 2). Since Bijvoet and dispersive differences are orthogonal in dependence on phase angles (Eqs. 6 & 8), having strong differences for both is desirable for definitive phase determination. That said, as described below, the increasing success of SAD shows that very effective phase evaluations can also be made at a single wavelength.

Since the anomalous diffraction signals in many early MAD applications were expected to be relatively small, careful strategies were adopted to measure these small differences as precisely as possible. In particular, concerns about the effects of absorption and radiation damage led to procedures by which all data pertaining to a given phase (Bijvoet mates at different wavelengths) would be measured close together in time and with similar experimental conditions. To control for systematic error in Bijvoet differences, crystals were often aligned for mirror-symmetric simultaneity in diffraction (Guss *et al.* 1988; Hendrickson *et al.* 1988) or, more generally for when mirrors are absent, by inverse-beam geometry. The 'inverse-beam' method, which was introduced for streptavidin measurements at SSRL (Hendrickson *et al.* 1989), follows each diffracted image with one made after rotating by 180° about an axis perpendicular to the x-ray beam. If Friedel's law holds and alignments are perfect, the two images will be as if related by inverted beams, simply related by a mirror perpendicular to the axis of 180° rotation; if anomalous scattering is present, the Friedel mates will show differences but many common systematic errors will cancel. To control for systematic errors at different wavelengths, such as from progressive radiation damage, data would be recorded wedge-by-wedge or block-by-block, iterating among wavelengths and through inverse beams.

Early MAD experiments were conducted at room temperature or with limited cooling (e.g. at 4 °C); however, flash freezing to ~100 K became the norm after cryoprotection procedures were perfected (Rodgers, 1994). The greatly reduced radiation damage with cryoprotection was a

https://doi.org/10.1017/S0033583514000018 Published online by Cambridge University Press

special boon to MAD given its need for multiple data sets. The first MAD structure from a cryo-protected crystal was of human chorionic gonadotropin (Wu *et al.* 1994), and subsequent MAD results came almost exclusively from flash-frozen crystals. Both crystal freezing and the typically larger signals from the multiple sites of selenomethionyl proteins provided access to more challenging systems, and these advances also relaxed the need for special care in data collection. As MAD proved to be convincingly robust, a wide range of data measurement procedures came into use. Although the rates of radiation damage are reduced ∼50-fold on going from 300 to 100 K (Warkentin & Thorne, 2010), nevertheless, crystals at 100 K are not immortal and care in measurement strategy remains advisable.

5.3 Synchrotron beamlines

At the time of the beginnings of MAD phasing experiments, synchrotron radiation was already being used in macromolecular crystallography (MX) with a premium on high flux. Many early MX beamlines (e.g. CHESS A1, SSRL 7.1 and DESY X11) used horizontally focusing bent single-crystal monochromators, which were usually operated at a fixed wavelength and these used photographic film as the detector. Such instruments were ill suited for the tuning needed by MAD and for the precision envisioned for data collection. Other beamlines equipped with double-crystal monochromators (e.g. SSRL 1.5, PF 14A and LURE D1.5) and electronic detectors did anticipate multiwavelength experiments on protein crystals. Matching count-rate limitations of the detectors, beam focusing on these latter beamlines was quite limited, however, and data collections took several days. Nevertheless, these beamlines proved effective in early MAD experiments, as described above.

Even the single-crystal monochromator systems could be adapted for MAD, and this was done after imaging-plate detectors were introduced (Amemiya *et al.* 1988) in early MAD structure determinations at PF 6A2 for cytochrome c553 (Nakagawa *et al.* 1990) and CHESS F1 for tenascin (Leahy *et al.* 1992). Later, procedures were developed to achieve high-energy resolution from a bent, asymmetrically cut monochromator (Lidestri & Hendrickson, 2009), and the structure of a relatively large selenomethionyl protein complex was determined at horizontally focused NSLS X4C (Moore & Hendrickson, 2012). Many other horizontally focused, single-crystal monochromator systems are now advertised and used as SAD beamlines (http://biosync.sbkb.org/), but typically with limited energy resolution.

As successes mounted from MAD phasing experiments, often performed with special beamline adaptations or suboptimal characteristics, the need for beamlines optimized for MAD phasing became evident. Thus, we set out to develop such a beamline at Brookhaven National Laboratory as NSLS X4A (Staudenmann *et al.* 1989). X4A was designed to produce monochromatic x-rays, to be tuned conveniently over a wide energy range to accommodate the many absorption-edge opportunities for the method, to have high-energy resolution for optimizing experiments at white-line features (Fig. 1*b*), and to be focused for maximizing fluxes and minimizing data collection times. The first published structures from X4A were of the bacterial redox protein DsbA (Martin *et al.* 1993) and of human chorionic gonadotropin (Wu *et al.* 1994), and X4 has been highly productive ever since. MAD beamlines with characteristics similar to X4A were built at other synchrotrons, and among the first of these were ALS 5.0.2, APS 19ID and ESRF BM14. Many others have followed (http://biosync.sbkb.org/).

Advances in x-ray detectors have been critically important for the development of MAD. Multiwire proportional counters (MWPCs) were used for many of the first experiments at

https://doi.org/10.1017/S0033583514000018 Published online by Cambridge University Press

LURE D1 and SSRL 1-5, and a single-reflection scintillation-activated proportional counter was used initially at PF 14A. An MWPC detector was also used at LURE D23, notably for mannose-binding protein (Weis *et al.* 1991); however, count-rate limitations led to a push toward integrating area detectors for MAD phasing. It was with the introduction of imaging plates that MAD phasing really took off, first as read by off-line readers (Amemiya *et al.* 1988) and later as incorporated into commercial systems as in the Rigaku *R*-axis IV and the MAR345 detectors. Ever brighter beamlines led a push to fast detectors, first by way of charge-coupled detectors (CCDs, such as the ADSC Q315) as supplied by several vendors and more recently in the move to pixel array detectors (PADs, such as the Pilatus 6M).

5.4 Analytical and computational procedures

The analysis for MAD or SAD experiments begins with data reduction whereby the structure-dependent components, the $|F(\mathbf{h})|$ amplitudes of structure factors, are extracted from the observed intensities, $I(\mathbf{h})$. This is done, as in the *modus operandi* for crystallographic structure determinations generally, by determining and applying the various scaling and correction factors that relate intensities to structure factors: $I(\mathbf{h}) = K\, C_1\, C_2 \ldots C_N\, |F(\mathbf{h})|^2$. That is, rather than maintaining the actual observations against which models can be tested, the factors related to diverse experimental effects, including ones due to diffraction geometry (Lorentz factors), polarization, absorption, and even radiation damage, are evaluated and used to isolate the structure-dependent quantities. Multiple observations of each unique reflection can be merged together into a unique data set for the Friedel pairs at each wavelength $\{|{}^{\lambda}F(\mathbf{h})|, |{}^{\lambda}F(-\mathbf{h})|\}$, and often this is done with standard procedures encoded by programs such as HKL2000 (Otwinowski & Minor, 1997), SCALA (Evans, 2011) or XDS (Kabsch, 2010), which do not see distinctions in the mode of measurement. Such blind mergings have the defect, however, that they will obscure the precision that can be facilitated by practices of inverse-beam and mirror-simultaneity discussed above. Using our MADSYS program suite (Hendrickson, 1991), we have included crystal numbers, orientation codes and symmetry codes to preserve individual identities through parameterized local scaling and into phase evaluation by a 'phase first, merge later' approach (Hendrickson, 1985). MAD phasing has proved to be robust, however, and many structures have been determined even despite blind mergings of the data. Improvements in accuracy can be expected, however, when attention is paid to the control of systematic errors in anomalous differences.

In addition to the diffraction data, MAD analyses also require knowledge of the anomalous scattering factors, f' and f'', at each wavelength. These can be obtained from measurements of atomic absorption spectra and Kramers–Kronig conversion by programs such as KRAMIG (Hendrickson *et al.* 1988) or CHOOCH (Evans & Pettifer, 2001). Alternatively, and commonly in current practice, these values can be refined in the course of phase evaluation (Weis *et al.* 1991) even from rather crude initial estimates. Anisotropy of anomalous diffraction is seen in atomic absorption for simple cases such as for single-site selenoethers (Hendrickson *et al.* 1989, 1990; Templeton & Templeton, 1988), and it will also be present even when multiple orientations scramble the macroscopic effects. The intrinsic anisotropy is represented by an atomic scattering tensor, $\mathbf{f}^{\Delta}$, and the effective scattering factor, f^{Δ}_{eff}, depends on the orientation of that atom with respect to polarization directions of both the incident beam, $\hat{\mathbf{e}}$, and the diffracted beam, $\hat{\mathbf{e}}'$. Thus,

$$f^{\Delta}_{\text{eff}} = {}^{\text{t}}\hat{\mathbf{e}}\, \mathbf{f}^{\Delta} \hat{\mathbf{e}}'. \tag{11}$$

https://doi.org/10.1017/S0033583514000018 Published online by Cambridge University Press

Effects of anisotropy can be taken into account by refining each scattering tensor (Fanchon & Hendrickson, 1990), which requires unmerged diffraction data and the preservation of crystal orientations at each diffracting event. Our simulations showed that uncorrected anisotropy for selenomethionyl proteins does not spoil MAD phase evaluation; indeed, using a modified version of SHARP, the inclusion of anisotropy was shown to enhance SAD phasing power (Schiltz & Bricogne, 2008).

The process of structure determination by MAD or SAD requires first an initial determination of the positions of anomalous scatterers, just as MIR or SIR phasing requires that the heavy-atom positions be determined before phase evaluations can proceed. Once the substructure is known, the amplitudes and phases from its associated diffracted waves can be used as interference references to find phases for the diffraction from the complete structure.

The information in appropriately conducted MAD experiments provides for a definitive analysis of the phase problem, as is embodied in Eq. (4). We reduced this unified approach to practice in our initial application to lamprey hemoglobin (Hendrickson, 1985; Hendrickson *et al.* 1988) and then used it in our other early analyses. An instructive example is taken from the structure of the substrate-binding domain of selenomethionyl Hsp70 DnaK (Zhu *et al.* 1996). Here, the diffraction data from a four-wavelength MAD experiment were measured on imaging plates in successive iterations of wavelengths by wedges with mirror simultaneity of Bijvoet differences; and the data were analyzed by MADSYS in a 'phase first, merge later' strategy. The structure was refined at 2.0 Å resolution after model building based on a 2.3 Å resolution electron-density map, which in turn derived from a substructure determined at 3.0 Å resolution.

The multi-determination, 'phase first' approach for this DnaK structure gave MADLSQ statistics after mergers at 2.3 Å resolution of $R\,(|{}^{\circ}F_{\mathbf{T}}|) = 0.051$, $R\,(|{}^{\circ}F_{\mathbf{A}}|) = 0.356$, $\langle \Delta\varphi \rangle = 36.5°$ and $\langle \sigma(\Delta\varphi) \rangle = 17.2°$, where $\Delta\varphi = {}^{\circ}\phi_{\mathrm{T}} - {}^{\circ}\phi_{\mathrm{A}}$. Before phase information given by the $\Delta\varphi$ values can be used, the positions of anomalous scatterers, $\{\mathbf{r}_{\mathrm{A}}\}$, are needed to derive the reference phases, ${}^{\circ}\varphi_{\mathrm{A}}$. The accurately determined $|{}^{\circ}F_{\mathbf{A}}|$ coefficients from MADLSQ gave a Patterson map for the selenium positions (Fig. 4*a*) that was readily interpreted by Patterson superposition program HASSP (Terwilliger & Berendzen, 1999) and validated by the corresponding $|F_{\mathrm{calc}}|$ Patterson map from the deduced Se positions (Fig. 4*b*). The substructure thus deduced is unavoidably ambiguous, however: if a set of positions $\{\mathbf{r}_{\mathrm{A}}\}$ satisfies the $|{}^{\circ}F_{\mathbf{A}}|$ coefficients, its enantiomoph $\{-\mathbf{r}_{\mathrm{A}}\}$ also does so equally well. This ambiguity must be resolved, for if $\{\mathbf{r}_{\mathrm{A}}\}$ gives phases ${}^{\circ}\phi_{\mathrm{A}}$, then $\{-\mathbf{r}_{\mathrm{A}}\}$ gives phases $-{}^{\circ}\phi_{\mathrm{A}}$; and only with the correct choice will the phase solutions, ${}^{\circ}\phi_{\mathrm{T}} = \Delta\varphi + {}^{\circ}\phi_{\mathrm{A}}$, be useful. A statistical basis for resolving the enantiomorph ambiguity has proved elusive, but electron-density maps derived from the alternatives are decisive when phase evaluations are accurate as they were here ($\langle \sigma(\Delta\varphi) \rangle = 17.2°$). When the hand is correct (Fig. 5*b*), the density map shows clearcut molecular boundaries and structural features; whereas the map from the wrong hand (Fig. 5*a*) is without recognizable features.

Our unified analysis of MAD data, whereby $|{}^{\circ}F_{\mathbf{A}}|$ coefficients are extracted as in the DnaK analysis (Fig. 4), can provide a rigorous solution of the substructure problem. Such rigor is not essential, however; even the Bijvoet differences at a single wavelength suffice in many cases, and this is crucial for SAD phasing. The problem of finding anomalous scatterer positions was first solved by Rossmann (1961), who showed that positions for the anomalous scatterers could be obtained from a Patterson map based on $(|F(\mathbf{h})| - |F(-\mathbf{h})|)^2$ coefficients. This approximation can be appreciated from Eq. (6), which shows $\Delta F_{\pm\mathbf{h}}$ to be directly proportional to $|{}^{\circ}F_{\mathbf{A}}|$ but reduced by $\sin(\Delta\varphi)$. Terwilliger (1994) showed that $|{}^{\circ}F_{\mathbf{A}}|$ estimates can be improved by the use of prior expectations, going toward correction for the trigonometric factor. Nevertheless,

https://doi.org/10.1017/S0033583514000018 Published online by Cambridge University Press

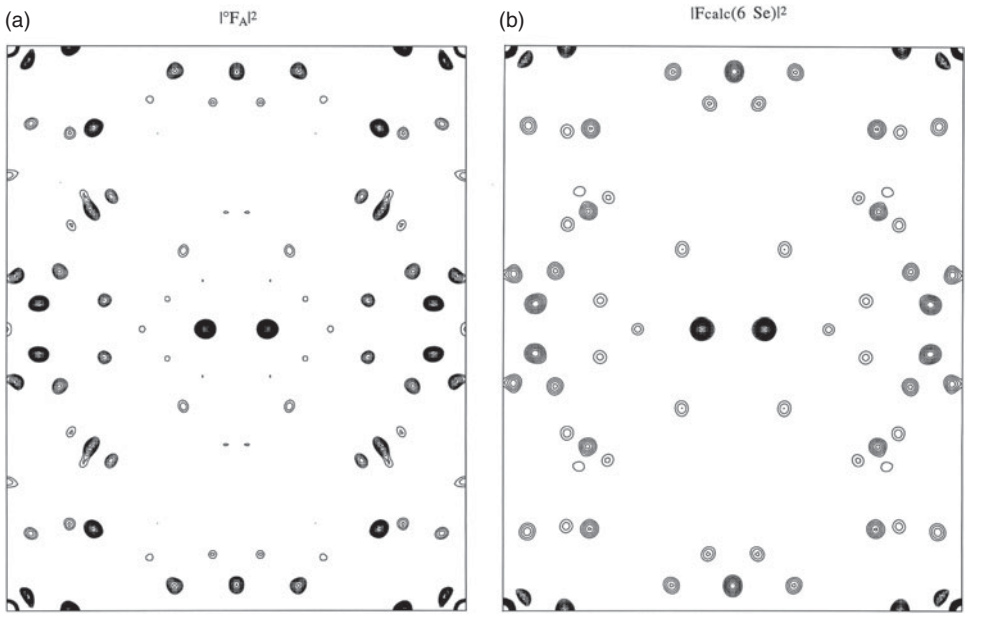


Fig. 4. Patterson maps for Se positions of DnaK substrate-binding domain (SBD). (*a*) Map based on experimentally derived $|^{\circ}F_{\mathrm{A}}|^{2}$ coefficients. (*b*) Map based on the Se model derived from $|F_{\mathrm{calc}}(6\ \mathrm{Se})|^{2}$ coefficients. In each case, the Harker section at $w = ½$ is shown with density at 3.5 Å resolution.

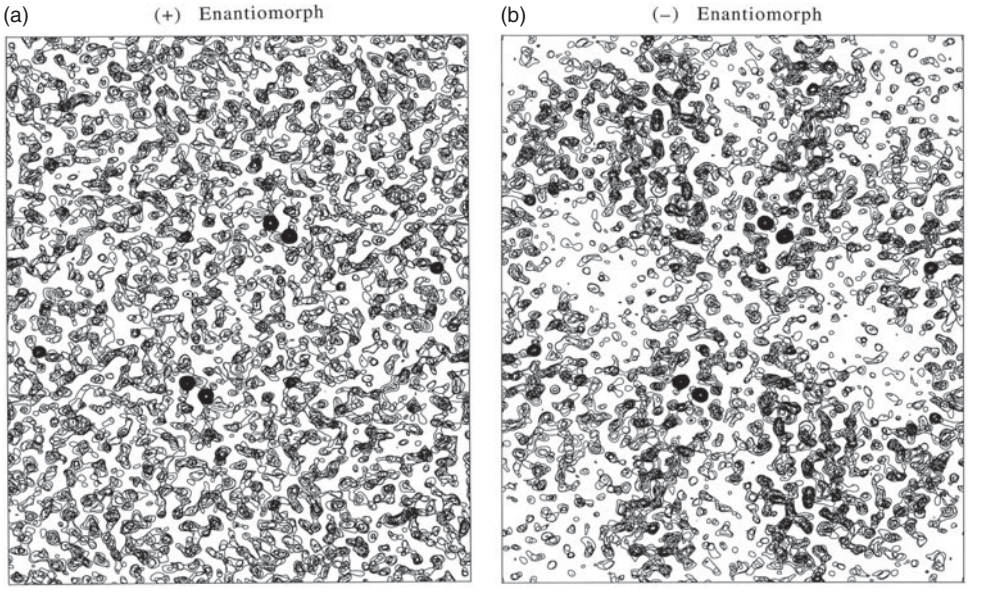


Fig. 5. Density distributions for DnaK SBD based on enantiomeric Se alternatives. (*a*) Map based on phases derived from the Se substructure as determined by Patterson superposition methods, the (+) enantiomorph. (*b*) Map based on phases derived from the alternative Se substructure having each z coordinate changed to $-z$, the (−) enantiomoph. Each distribution was produced at 2.3 Å resolution, and each image includes equivalent slabs of sections at $+z$ and $-z$, respectively.

https://doi.org/10.1017/S0033583514000018 Published online by Cambridge University Press

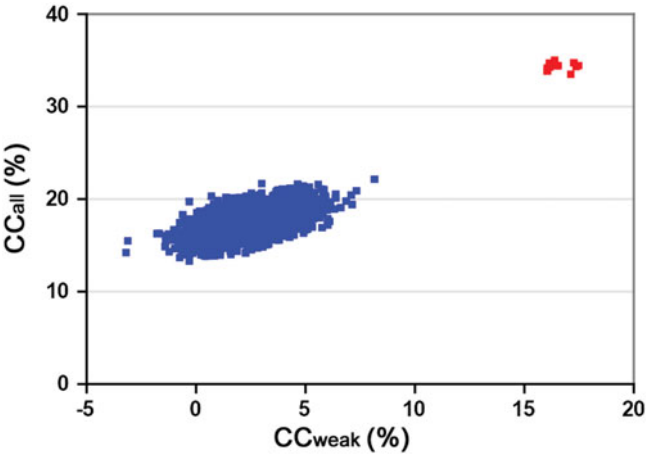


Fig. 6. Profile of correlation coefficients (CCs) from a substructure determination. The comparisons here are between observed Bijvoet differences from a TorT/TorS crystal and those calculated in SnB trials. The distribution of CCall and CCweak values is from 2000 SHELXD solution attempts. Successful solutions (14) are colored red and random solutions (1886) are colored blue. Reproduced from Fig. 2G of Liu *et al.* (2012) with permission from *Acta Crystallographica D*, © 2012 by the International Union of Crystallography.

the vast majority of MAD and SAD substructure determinations simply use the Rossmann approximation. An important development for substructure determination has come from the direct methods approach, first applied in the analysis of selenomethionyl S-adenosylhomocysteine hydrolase (Turner *et al.* 1998) where the Shake-and-Bake (SnB) algorithm as applied to Bijvoet differences (Smith *et al.* 1998) found all 30 Se atoms. The SHELXD implementation (Schneider & Sheldrick, 2002) of the SnB algorithm (Hauptman, 1997) has been especially effective. In this approach, random starting structures are refined by iterations of tangent-formula refinement and real-space model adjustments against Bijvoet-difference estimates of $|^{\circ}F_{\mathbf{A}}|$, using correlation-coefficient tests to find 'solutions' that stand out from the noise. Substructure determination from a native SAD application to intact DnaK–ATP (Liu *et al.* 2013) is shown in Fig. 6. Direct methods procedures were instrumental in extending the reach of MAD to large selenomethionyl proteins (Deacon & Ealick, 1999).

Phase evaluation itself can proceed once the anomalous substructure is known. Two distinct approaches are taken. In one, the unified formalism of Eq. (4) or (3) more generally, is used. The structure of anomalous scatterers serves to specify the associated $^{\circ}\phi_{\mathrm{A}}$ phases; and, subject to the enantiomorph ambiguity discussed above, an algebraic definition of the desired $^{\circ}\phi_{\mathrm{T}}$ phase can be obtained. In practice, especially since data sets are often incomplete, a more comprehensive description of the phase information is given by a probabilistic treatment (Pähler *et al.* 1990). Having determined the anomalous substructure, often from Bijvoet differences, $|^{\circ}F_{\mathrm{A}}|$ and $^{\circ}\phi_{\mathrm{A}}$ can be calculated and $|^{\circ}F_{\mathrm{T}}| \approx \langle |^{\lambda_i}F(\mathbf{h})|^2 \rangle^{1/2}$ can be estimated; thereupon the joint probability distributions, $P(|^{\circ}F_{\mathrm{T}}|, {}^{\circ}\phi_{\mathrm{T}})$, can be produced and integrated to yield the best Fourier coefficients.

Alternatively, MAD experiments can be analyzed as *in situ* MIR experiments (Phillips & Hodgson, 1980), based on differences from physics instead of from chemistry. In this approach, data from one of the wavelengths are chosen for the reference data set, and differences are taken relative to these data. The first structure to be analyzed in a MAD-as-MIR approach was that of histone H5 (Ramakrishnan *et al.* 1993), and this method of analysis quickly took hold (Ramakrishnan & Biou, 1997). Various computer programs are used in MAD-as-MIR analyses,

https://doi.org/10.1017/S0033583514000018 Published online by Cambridge University Press

including MLPHARE (Otwinowski, 1991), SOLVE (Terwilliger & Berendzen, 1999), SHARP (de la Fortelle & Bricogne, 1997) and PHASES (Furey & Swaminathan, 1997). Although the unified and MAD-as-MIR approaches do often give similar solutions, they are distinguished, in principle, in that the structure factors at a particular wavelength, ${}^{\lambda}F(\mathbf{h})$, correspond to a general density function (Hendrickson & Sheriff, 1987), which is complex due to anomalous scattering contributions, $f^{\Delta} = f' + \mathrm{i}\, f''$. The general density function may depart significantly from the true electron-density function, which is purely real as determined by the structure factors due to normal scattering factors, f°. While often just a nuance, the distinction is not ignorable when anomalous scattering is large fraction of the total (e.g. Rudenko *et al.* 2002; Weis *et al.* 1991).

6. Evolution of SAD

Compared to the course of development for MAD, SAD emerged slowly and then accelerated remarkably after a gestation nurtured by MAD. An appealing characteristic of MAD phasing is its prospect for unambiguously definitive phase evaluations. In practice, this appeal of MAD for giving decisive answers was realized right from the first applications, even when the anomalous scattering signals were relatively weak. SAD carries much the same potential as MAD for decisive phasing, but this comes with a trigonometric ambiguity intrinsic to the basic phasing relationship of Eq. (6). SAD is like SIR in this respect, where incorporation of the orthogonal information from anomalous scattering into SIRAS resolves the phase ambiguities. MAD is the obvious resolving counterpart for SAD, giving the complementary phase information of Eq. (8). When such orthogonality is not feasible or ineffective, as when measurements are remote from absorption resonances, the resolution of phase ambiguity becomes a critical issue.

SAD experiments are a subset of MAD experiments. Although wavelength changes are not involved, tuning to a resonance edge or to exceptionally low energy is often used for signal optimization. Thus, synchrotron radiation is important even if not essential for SAD. Nearly all aspects of data measurement and analysis described above for MAD also carry over to SAD. What differs most is the need to add information beyond anomalous scattering to resolve the phase ambiguity.

6.1 Crambin and phase ambiguity

Crambin was the first macromolecular structure to be determined *de novo* from the diffraction data of a single crystal and it was the first to be analyzed without incorporating exogenous heavy atoms (Hendrickson & Teeter, 1981). Crambin is a very small protein (46 residues) and it diffracts exceptionally well ($d_{\min} < 0.48$ Å; Schmidt *et al.* 2011); nevertheless, for its time it presented a substantial challenge. Sulfur SAD phasing was used to determine the structure, although the procedure was not identified then as SAD.

The diffraction data from a monoclinic crystal of crambin were measured at room temperature with CuKα radiation (sulfur $f'' = 0.56$e) by single-counter diffractometry, one reflection at a time. Bijvoet differences were measured to 1.5 Å spacings. Care was taken to reduce systematic errors by measuring data in blocks of 25 reflections at $(\varphi, \psi, 2\theta)$ followed by the corresponding Friedel mates at $(\varphi, \psi, -2\theta)$ and by applying a parameterized local scaling before differencing. Care was also taken to use counting times sufficient to assure adequate counting statistics. The error level for the data set, taking into account a term for instrumental instability, was rms $(\sigma_{\Delta F}) = 1.6$e as compared to rms $(\Delta F_{\pm \mathbf{h}}) = 2.3$e, which was 2.1% of rms $(|F|)$. Bijvoet-difference Patterson

https://doi.org/10.1017/S0033583514000018 Published online by Cambridge University Press

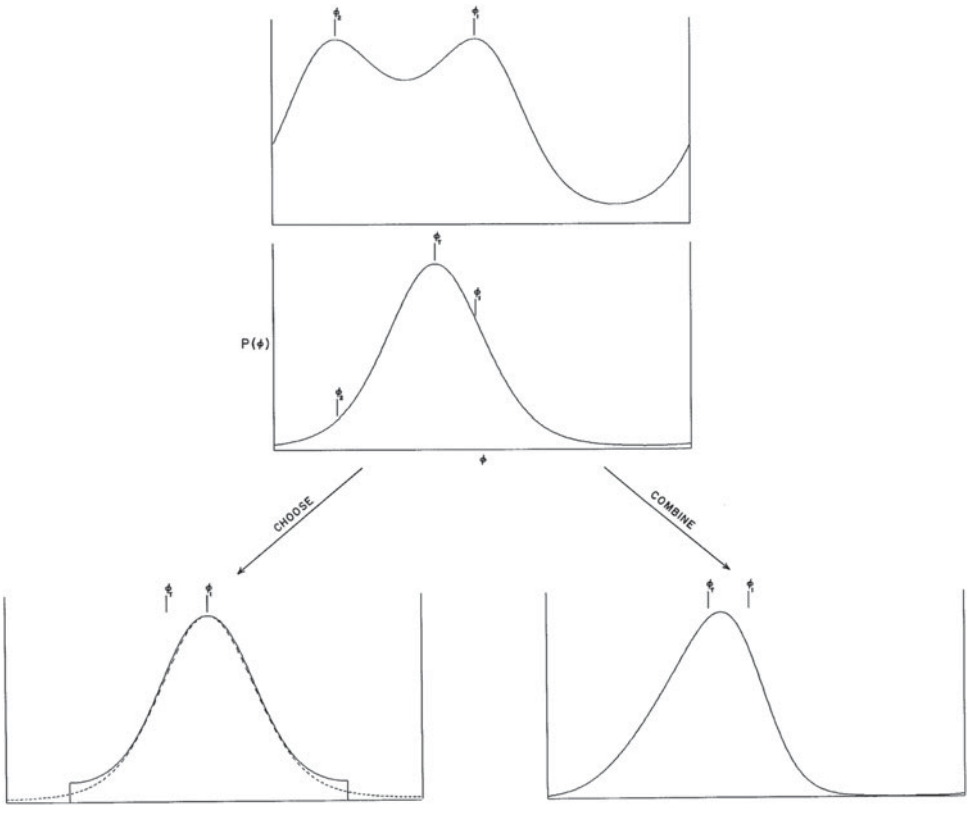


Fig. 7. Scheme of options used for resolving phase ambiguity with crambin. (Above) Bimodal 'SAD' probability distribution based on the Bijvoet difference for a particular reflection (Hendrickson, 1979). (Middle) Partial structure probability distribution based on the sulfur substructure. (Below) Options of multiplicative phase combination of the top two profiles ('COMBINE') or choice of the symmetrized modal distribution (Hendrickson, 1971) corresponding to the 'SAD' alternative, φ_1, that has partial-structure probability significantly higher than for the lesser alternative, φ_2 ('CHOOSE').

maps were used to solve for the sulfur positions, first at 3 Å resolution where disulfide-bonded atoms are not resolved and then at 1.5 Å resolution to deduce the six-atom sulfur structure. Only the subset of 38% of reflections with strongest Bijvoet differences, where $\sin(\Delta\varphi)\approx 1$; so that $|\Delta F_{\pm\mathbf{h}}|$ gives $|{}^{\circ}F_{\mathrm{A}}(\mathbf{h})|$ directly, were used to refine the sulfur substructure.

From the Bijvoet-difference relationship used for the crambin analysis, Eq. (6), the phase ambiguity of SAD is evident. Once the general enantiomorph ambiguity is resolved, there still exists for each acentric reflection the ambiguity that ${}^{\circ}\phi_{\mathrm{T}} = {}^{\circ}\phi_{\mathrm{A}} \pm \sin^{-1}[2({f''}_{\mathrm{A}}/{f^{\circ}}_{\mathrm{A}})/|{}^{\circ}F_{\mathbf{A}}|]$. In terms of phase probability (Hendrickson, 1979), this in general presents a bimodal probability distribution (Fig. 7) although the distribution may be unimodal if the alternative phases are close to one another. A figure-of-merit weighted map with centroid phases will give the least error (Blow & Crick, 1959); such a map will still be inaccurate, however it can be improved greatly with appropriate resolution of the phase ambiguity. For the crambin analysis, partial-structure phase information coming from the sulfur atoms was used for ambiguity resolution in a scheme for making choices between alternatives when the discriminative probability is high and for using multiplicative phase combination otherwise (Hendrickson, 1971; Fig. 7*s*). The sulfur partial structure here is relatively weak ($\langle|{}^{\circ}F_{\mathbf{A}}|\rangle/\langle|{}^{\circ}F_{\mathbf{T}}|\rangle = 2.9\%$), however, and the resulting density map (Fig. 8) was poor

https://doi.org/10.1017/S0033583514000018 Published online by Cambridge University Press

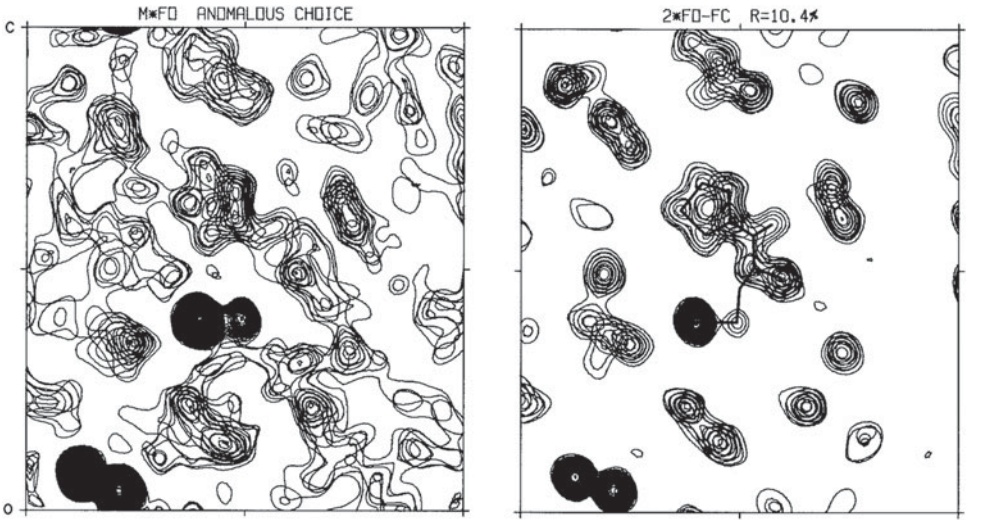


Fig. 8. Electron-density maps from the structure analysis of crambin. (Left) Experimental map with 'SAD' phases resolved by the partial-structure anomalous-choice procedure. (Right) Model-phased ($2F_{obs} - F_{calc}$) map from the final model refined at 1.5 Å resolution. The portion shown in these frames includes the disulfide-bonded sulfur atoms from Cys16–Cys26 at lower left, Sγ of Cys32 near the center and disulfide-bonded to Sγ of Cys4 (mostly out of the section), and the five-membered ring of Pro5 just above center. Reproduced from Fig. 2 of Hendrickson & Teeter (1981) with permission from *Nature*, © 1981 Macmillan Publishers Ltd.

compared with those from later MAD experiments (Figs 3 and 5). This map did nevertheless prove to be interpretable into a well determined structure, thanks largely to the high resolution.

The partial-structure resolved anomalous phasing method was used for a few other analyses (Conner *et al.* 1982; Sheriff *et al.* 1987; Smith *et al.* 1983), but such approaches were quickly overwhelmed by the growing successes of MAD phasing.

6.2 Ambiguity resolution from density modification

If only by comparison with the decisiveness of MAD, it became apparent that anomalous substructures alone are not generally adequate for effective resolution of phase ambiguity in single-wavelength anomalous phasing experiments. A striking example came with the application to a trimeric hemerythrin (Smith *et al.* 1983) where the di-iron sites of this metalloprotein were used in resolved anomalous phasing to obtain an initial map, which was inadequate for interpretation on its own but was improved dramatically by threefold molecular averaging and 'solvent leveling', as it was then described (Fig. 9). Similarly dramatic enhancements from symmetry averaging had been seen in the structure analysis of influenza virus hemagglutinin (Wilson *et al.* 1981). Non-crystallographic symmetry clearly provides a powerful means form ambiguity resolution, but it is not a general solution.

A general and powerful solution came with the innovation of procedures for molecular envelope definition and solvent flattening (Wang, 1985). As applied to anomalous scattering problems, this procedure was called the iterative single-wavelength anomalous scattering (ISAS) method, and as such its first application to an unknown was to a neurophysin–dipeptide complex using an iodinated derivative (Chen *et al.* 1991). A second application of this kind was made in the corrected analysis of Cd, Zn metallothionein (Robbins *et al.* 1991). A direct-methods approach to

https://doi.org/10.1017/S0033583514000018 Published online by Cambridge University Press

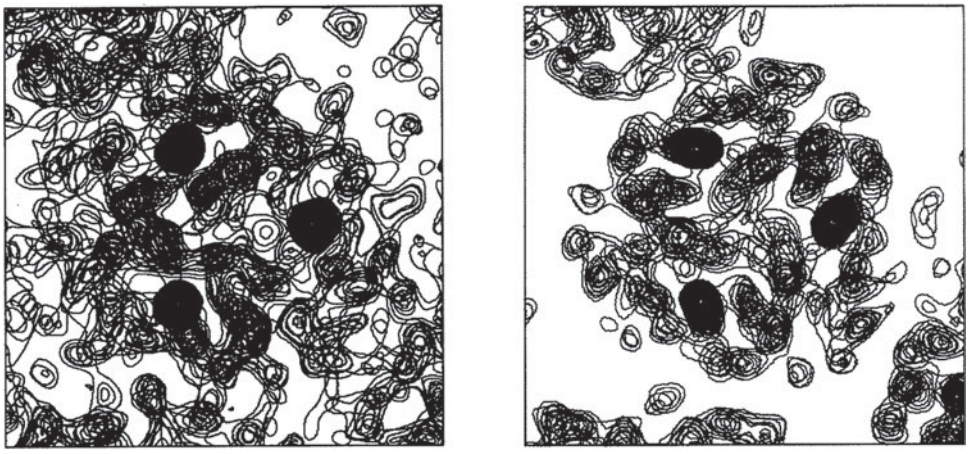


Fig. 9. Electron-density maps from the structure analysis of trimeric hemerythrin. (Left) Experimental map with 'SAD' phases resolved by the iron partial structure. (Right) Experimental map after refinement by iterative threefold molecular averaging and solvent leveling.

ambiguity resolution, called one-wavelength anomalous scattering (OAS), was also devised and tested (Fan & Gu, 1985); and this procedure was used in an application to the copper protein rusticyanin (Harvey *et al.* 1998). Phases for rusticyanin from the OAS procedure were not sufficiently accurate to produce a readily interpretable map (Harvey *et al.* 1998); however, refinement with density modification as implemented by Cowtan & Main (1993) did produce the structure. Direct methods by way of Sayre's equation had also been ineffective in the crambin study (Hendrickson & Teeter, 1981).

Density modification continued to develop after Wang's introduction of solvent flattening (1985) with the incorporations of histogram matching (Zhang & Main, 1990), the systematic inclusion of molecular averaging (Cowtan & Main, 1993), the innovation of solvent flipping (Abrahams & Leslie, 1996), the introduction of maximum-likelihood (Terwilliger, 2000), and with varied implementations as reviewed by Cowtan (2010). Another noteworthy addition to crystallographic practice with relevance to the interpretability of density maps was the introduction of automated chain tracing with the program ARP/wARP (Perraskis *et al.* 1999) and successors such as RESOLVE (Terwilliger, 2003a, b) and BUCCANEER (Cowtan, 2006). With these advances in computational procedures, the trigonometric phase ambiguity of SAD could be resolved to extract the intrinsic phase definition of careful experiments.

6.3 SAD becoming the method of choice

As density-modification procedures were implemented and the resulting power for effective ambiguity resolution was recognized, new opportunities became apparent and additional single-wavelength anomalous analyses were performed (Biou *et al.* 1995; Turner *et al.* 1998). Then, using hen egg-white lysozyme as a test system, Dauter and co-workers showed that density-modification resolved phasing could be effective when based on atoms intrinsic to the structure (Dauter *et al.* 1999) or based on introduced bromide atoms (Dauter & Dauter, 1999). The intrinsic-atom study (sulfur and chlorine) was performed at the wavelength of CuKα to emphasize that crambin-like analyses well away from absorption edges would now be straightforward,

https://doi.org/10.1017/S0033583514000018 Published online by Cambridge University Press

and this analysis used a combination of programs SHELXM (Sheldrick, 1998), SHARP (de la Fortelle & Bricogne, 1997) and SOLOMON (Abrahams & Leslie, 1996). The bromide ion studies were performed at or near K-edge resonance, and this analysis used a combination of SHELXM (Sheldrick, 1998), DM (Cowtan, 1994) and MLPHARE (Otwinowski, 1991). A key validation of the potential for sulfur SAD phasing came in the structure determination of obelin (Liu *et al.* 2000), which was performed at a wavelength of 1.74 Å and used a combination of SOLVE (Terwilliger & Berendzen, 1999) and PHASES (Furey & Swaminathan, 1997) to find the sulfur sites and the ISAS solvent flattening (Wang, 1985) for phase evaluation.

Even earlier, the PDB deposition for a green fluorescent protein (GFP) structure (Ormö *et al.* 1996) provided, to my knowledge, the first recorded use of the SAD acronym in citing 'MIR/SAD' as the method of determination to describe an analysis of selenomethionyl GFP. 'SAD' was also used in the Dauter papers on using sulfur and bromine as anomalous scatterers, but this was done off-handedly as if the term were already commonplace. It was used again in reporting the structure of psoriasin (Brodersen *et al.* 2000) and in a re-evaluation of seven structural analyses of selenomethionyl proteins (Rice *et al.* 2000). In the latter study, two structures were SAD-phasing rescues for MAD-phasing failures, spoiled by radiation damage; for the other five cases, the maps from density-modified SAD phases were nearly as good as those from MAD phases. With this background, Dauter *et al.* (2002) then presented a defining mandate for SAD that reviewed the method and described fifteen tests and novel applications.

The growing successes of SAD spurred a renewed interest in home-source experiments (Nagem *et al.* 2005; Yang & Pflugrath, 2001), notably employing x-rays generated from a chromium anode to better exploit sulfur anomalous scattering (Kitago *et al.* 2005; Yang *et al.* 2003) or to take advantage of the anomalous scattering from iodine as its L_I edge is approached (Evans & Bricogne, 2003). Nevertheless, synchrotron radiation provides compelling advantages; fine tuning of x-ray energy makes it possible to optimize the strength of f'' at the Se K-edge and other resonance peaks, and access to lower x-ray energies permits the enhancement of anomalous signals from low-Z elements such as sulfur since f'' values increase as these K-edges are approached (Weiss *et al.* 2001). The synchrotron advantage for selenium is responsible for most of the surge in selenomethionyl SAD structures documented below and the low-energy advantage is beginning to have an impact for sulfur SAD structures as multi-crystal experiments are developed.

6.4 Multi-crystal enhancement of signal-to-noise ratios

Although SAD has become the method of choice for *de novo* phase evaluation in crystallographic structure determination, it has remained relatively ineffective for challenges where diffraction is poor or signals are too weak. A recent analysis of SAD-phased structures in the Protein Data Bank (PDB) found only a small fraction to be at low resolution or to be based on native, low-Z elements (Liu *et al.* 2013). Among the 5286 identified SAD entries as of April 2012, only 32 structures were at low resolution ($d_{\min} \geqslant 3.5$ Å) and only 58 structures were based only on light atoms ($Z \leqslant 20$). Moreover, the only-light-atom (sulfur SAD) structures were at relatively high resolution: 74% at higher than 2 Å resolution; all at higher than 2.3 Å resolution. While the limited output of structures from low-resolution and only-light-atom SAD experiments might be due to a scarcity of such applications, it seemed more likely to have reflected complications from feeble anomalous signals embedded in noisy diffraction data.

We hypothesized that success with challenging SAD experiments could be improved by increasing the multiplicity of measurements; thereby, the enhancement in signal-to-noise over

https://doi.org/10.1017/S0033583514000018 Published online by Cambridge University Press

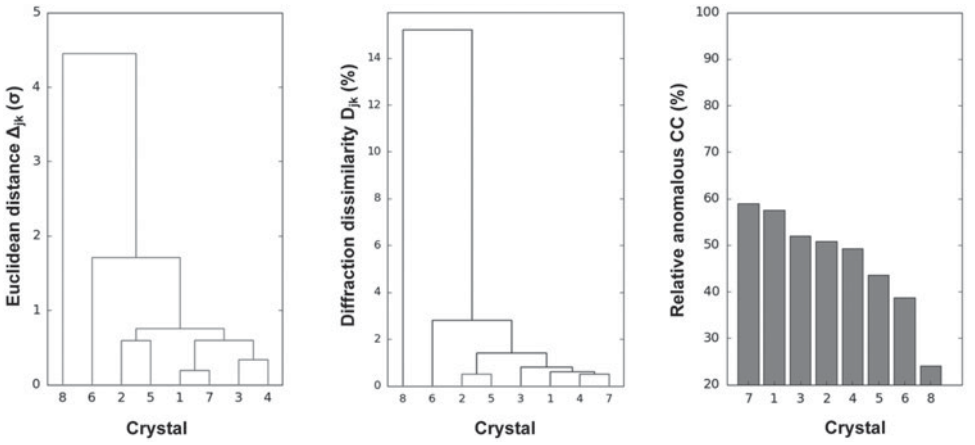


Fig. 10. Variations among crystals from a multi-crystal data set. (Left) Pairwise cluster analysis of unit cell variations. (Center) Pairwise cluster analysis of overall intensity variations. (Right) Relative anomalous correlation coefficient comparing Bijvoet differences from each crystal to those from the average of all together. Crystal 8 was eliminated as an outlier from the distribution, but the others were judged to be statistically equivalent and included for structure determination. Adapted from Fig. 1 of Liu *et al.* (2012) with permission from *Science*, © 2012 American Association for the Advancement of Science.

that from a single measurement is expected to be $\sqrt{M}$, where M is the multiplicity. Such improvements had been demonstrated for phosphorous SAD from a relatively large crystal of a DNA oligomer (Dauter & Adamiak, 2001); however, more typical macromolecular crystals cannot sustain multiple cycles of measurement without unacceptable radiation damage. Averaging from multiple crystals is an obvious solution; however, with cryoprotection, macromolecular crystallography had devolved into the search for one good crystal. Usually, one frozen crystal sufficed, and this was seen as preferable to multi-crystal averaging because crystal variation was expected due to stochastic lattice changes with flash freezing. For multi-crystal SAD experiments, it was thus important to assure that only crystals with statistically equivalent diffraction would be included. Clustering tests for such assurance were devised based on lattice dimensions, overall intensity patterns and anomalous diffraction profiles (Liu *et al.* 2012). Figure 10 exemplifies such clustering.

The idea of multi-crystal SAD phasing was tested first with selenomethionyl protein crystals at low resolution. We used the large extracellular domain of a histidine kinase receptor, which formed crystals that diffracted only to 3.5 Å Bragg spacings. With two copies in the asymmetric unit, this structure proved to be challenging in size as well as resolution; there proved to be 1456 ordered residues and 20 Se sites. Both the success in Se-substructure determination and the accuracy of resulting phase evaluations improved with the increasing multiplicity in diffraction from multiple crystals (Liu *et al.* 2011). A particularly common instance of low-resolution selenomethionyl problems is found with crystals of detergent-solubilized membrane proteins, and the multi-crystal approach has proved instrumental in such low-resolution analyses for a dicarboxylate transporter (Mancusso *et al.* 2012) and for a polyisoprenyl-phosphate glycosyltransferase (submitted for publication).

The challenge of only-light-atom native SAD analyses provided the second motivation for multi-crystal SAD experiments, and the methodology for such applications has also been described recently (Liu *et al.* 2012, 2013). Five structures were determined in these tests, which

https://doi.org/10.1017/S0033583514000018 Published online by Cambridge University Press

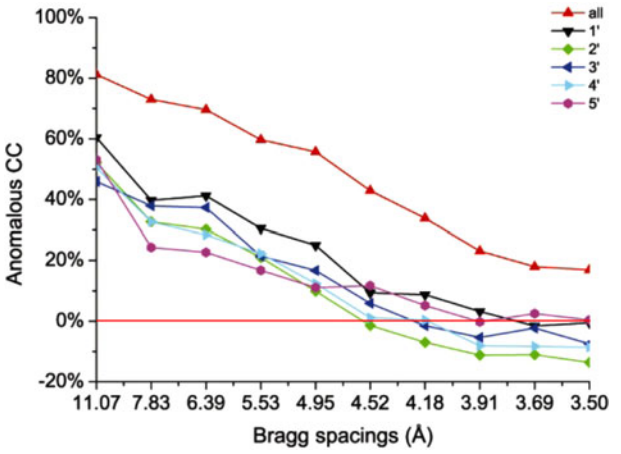


Fig. 11. Enhancement of anomalous signal-to-noise by multi-crystal averaging. Anomalous correlation coefficient (Anomalous CC) profiles as a function of Bragg spacings, comparing Bijvoet differences from five individual data sets with the data set compiled by averaging from all five. For individual DnaK-ATP data sets, the signal is essentially undetectable beyond 4.5 Å spacings and is at the 20% level down to ~5 Å spacings; for the averaged data set, the signal is retained at the 20% level out to 3.5 Å spacings. Reproduced from Fig. 4*a* of Liu *et al.* (2013) with permission from *Acta Crystallographica D*, © 2013 by the International Union of Crystallography.

used 5–13 crystals, included 4–52 anomalous scatterers, comprised 127–1200 ordered protein residues per asymmetric unit, and ranged in resolution from 2.3 to 2.8 Å.

Procedures were devised to make multi-crystal native SAD phasing robust and routine. First, it was in these applications that clustering procedures were devised to assure statistical equivalence of crystals admitted into averaging, employing tests on unit-cell deviations, diffraction dissimilarity and relative anomalous correlation coefficient (CC) (Fig. 10). As addressed above and advocated previously (Weiss *et al.* 2001; Yang *et al.* 2003), enhanced f'' signals are expected at lower x-ray energy. We used the moderately low energy of 7 keV for our initial studies (Liu *et al.* 2013) since all of those crystals were relatively large; nevertheless, the use of a helium cone proved to be critical for the most challenging case (histidine kinase TorT/TorS$_S$ with 1148 residues and diffraction limited to 2.8 Å spacings). Standard data-reduction and scaling procedures were adopted, but with care to avoid including radiation-damaged frames. The enhancement in signal-to-noise was especially evident in the improvement of anomalous CC (from randomly split $\Delta F_{\pm \mathbf{h}}$ data) for data after merging compared to those from each individual crystal (Fig. 11). Substructures were determined using the SnB procedure of SHELXD (Hauptman, 1997; Schneider & Sheldrick, 2002). Initial phases were evaluated by Phaser (Read & McCoy, 2011), density modification was performed with DM (Cowtan & Zhang, 1999), and initial atomic models were built automatically with ARP/wARP (Langer *et al.* 2008).

Four of the five structures analyzed in our initial multi-crystal native SAD experiments were novel. Two of these novel structures are now published (Brasch *et al.* 2011; Qi *et al.* 2013) and papers are in preparation on the other two. Meanwhile, we and others have solved other structures by multi-crystal native SAD procedures. One on a flavivirus NS1 protein is recently published (Akey *et al.* 2014) and another on a calcium-leak channel has been submitted for publication.

https://doi.org/10.1017/S0033583514000018 Published online by Cambridge University Press

We anticipate that multi-crystal SAD procedures will have broad application to native macromolecules without recourse to heavy-atom incorporation (Liu *et al.* 2013). The initial set of applications included structures as large as 1200 amino-acid residues and at resolutions as limited as 2.8 Å. These are properties shared by 90% of current PDB holdings. Anticipated advances in synchrotron instrumentation and automation should facilitate making such experiments routine.

7. Applications and changing practice

At the outset, when MAD was instigated in the mid-1980s, such experiments were considered rather exotic and even heroic; moreover, the need for a new approach to *de novo* structure determination was then questionable in light of historic success with MIR and SIRAS. Ensuing applications showed convincingly that MAD was practical and that it could be the approach of choice for challenging structures. MAD was energized by technical innovations, notably cryopreservation and selenomethionyl proteins, and it flourished at beamlines designed expressly for MAD. MAD soon displaced MIR as the predominant *de novo* method; then, as density-modification procedures were perfected, SAD surpassed MAD. This section takes stock of current applications of anomalous diffraction in biological crystallography and changing practice in *de novo* structure determination.

7.1 Realms of application

The domain of MAD and SAD is nearly as broad as the periodic table itself. Each element has its distinctive resonant energies; and, as noted above in Section 5.2, all elements from Ca ($Z=20$) to U ($Z=92$) have K or L_{III} resonances in the energy range from 3.5 to 35 keV (3.5 to 0.35 Å in wavelength). Indeed, with specialized beamlines, the range can be extended down to 2.1 keV (5.8 Å) to include the K resonances from phosphorous and sulfur (Stuhrmann *et al.* 1997). Moreover, off-resonance f'' strength often suffices for SAD based on elements with resonances outside of accessible x-ray energies, which in practice may be restricted (e.g. 6 to 17 keV/2.1 to 0.7 Å). This opens significant opportunities for SAD with elements in the range $15 \leqslant Z \leqslant 24$, where K-edge strength can be approached, and $40 \leqslant Z \leqslant 59$ where L-edge strength can be approached. Success with sulfur SAD is a case in point. Indeed, even the anomalous scattering from oxygen can be used for structural characterization (Hope & de la Camp, 1972).

With the sensitivity of anomalous diffraction and the wide-open options for useful elements, versatility abounds. All of the conventional heavy-atom derivatives, such as those with U, Hg, Au, Pt and Os are also ideal for MAD and/or SAD; intrinsic metal centers, including many first-series transition elements such as Fe, Cu and Zn as well as Mo and W from higher series, make almost all metalloproteins natural candidates; many lighter ions have ready and potent replacements, e.g. PO_4^{2+} by WO_4^{2+}, SO_4^{2+} by SeO_4^{2+}, and Ca^{2+} or Mg^{2+} by a lanthanide such as Yb^{3+}; and there are several robust means for systematic incorporation. The most notable systematic approach is with selenomethionyl proteins (Doublié, 2007; Hendrickson *et al.* 1990); other possibilities include the synthesis of brominated or selenated nucleic acids (Sheng & Huang, 2010), the quick-soaked adsorption of solvent ions such as the halides Br^- and I^- (Dauter & Dauter, 2001) or cations Cs^+ or Gd^{3+} (Nagem *et al.* 2001), or the pressurized binding of noble gases Kr and Xe into hydrophobic cavities (Schiltz *et al.* 2003).

Endowed with this rich armoury of phasing elements, biological crystallographers have expeditiously solved thousands of *de novo* macromolecular structures. These applications are from the complete range of biochemical systems, having in common only the need for *de novo* structure

https://doi.org/10.1017/S0033583514000018 Published online by Cambridge University Press

determination. Because of this bias toward novelty, the resulting MAD and SAD structures have often broken new ground for biological discoveries and structural precedents. Instances of application include metabolic enzymes [e.g. NO synthase (Raman *et al.* 1998); HMG-CoA reductase (Istvan *et al.* 2000)], DNA replication and repair systems [e.g. clamp-loader complex (Jeruzalmi *et al.* 2001); RecA:DNA recombination complex (Chen *et al.* 2008); DNA polymerase η : DNA (Biertümpfel *et al.* 2010)], transcription factor complexes [e.g. TBP:TFIIA:DNA complex (Geiger *et al.* 1996); NFAT:Fos/Jun:DNA (Chen *et al.* 1998); nuclear receptor PPARγ–RXRα: DNA complex (Chandra *et al.* 2008)], RNA molecules [e.g. P4–P6 self-splicing intron (Cate *et al.* 1996), RNase P ribozyme:tRNA complex (Reiter *et al.* 2010)], signaling proteins [e.g. insulin receptor kinase (Hubbard *et al.* 1994); EGF receptor ectodomain (Cho & Leahy, 2002)], cytoskeletal proteins [Arp 2/3 complex (Robinson *et al.* 2001); centriolar hub protein SAS-6 (Kitagawa *et al.* 2011)], viral proteins [e.g. HIV integrase (Dyda *et al.* 1994); papilloma virus E2 protein (Harris & Botchan, 1999)], vesicle transport proteins [SNARE complex (Sutton *et al.* 1998); coatamer cage complex (Lee & Goldberg, 2010)], molecular chaperones [e.g. trigger factor (Martinez-Hackert & Hendrickson, 2009; Hsp70 DnaK (Qi *et al.* 2013)], membrane transporters [neurotransmitter transporter homolog LeuT (Yamashita *et al.* 2005); P-glycoprotein (Jin *et al.* 2012)], and ion channels [stomatal anion channel (Chen *et al.* 2010); voltage-gated sodium channel (Payandeh *et al.* 2011)]. There are many more instances for these and other categories of application. Notable among them are high-significance challenges such as the first crystal structures of ribosomes [70S (Cate *et al.* 1999); 30S (Wimberly *et al.* 2000)], transcriptional RNA polymerase (Cramer *et al.* 2001), and the G-protein-coupled receptor rhodopsin (Palczewski *et al.* 2000). From these precedents, other structures have followed by molecular replacement.

In addition to addressing the broad spectrum of structural biology associated with the projects of individual investigators, MAD and SAD also became instrumental in the development of the higher throughput approaches of structural genomics. The then newly realized efficacy of selenomethionyl MAD analysis of protein structure was a founding rationale for the structural genomics approach whereby structural representatives are sought for each sequence family in the protein universe (Shapiro & Lima, 1998). *De novo* structure determination is essential here since the focus is on sequences without known structures, and selenomethionyl MAD crystallography seemed most expeditious. Indeed, selenomethionyl proteins became a mainstay for the protein production pipelines of all four large-scale centers of the NIH Protein Structure Initiative (PSI): JCSG (van den Bedem *et al.* 2011), MCSG (Kim *et al.* 2011), NESG (Xiao *et al.* 2010) and NYSGXRC (Bonanno *et al.* 2005); however, as for the world more generally, SAD came to be preferred over MAD. Over the 10-year period of the PSI protein-universe effort, PSI-1 and PSI-2, 5097 atomic structures were deposited into the PDB, of which 4598 (90%) were crystal structures. These crystal structures included 1353 (29%) done by MAD and 2168 (47%) done by SAD.

In addition to the direct impact of anomalous diffraction in *de novo* structure determination, as exemplified by applications described above, ancillary benefits often accrue as well. First, the sites of anomalous scatterers can provide precise landmarks to assist in chain tracing at resolutions too low for clear residue identification for sequence alignment. This is particularly true for selenium sites marking the positions of methionine residues. There are many notable examples of such utility (e.g. Feng *et al.* 2010; Jin *et al.* 2012; Pomeranz Krummel *et al.* 2009; Sawaya *et al.* 2008; Wu *et al.* 1997), including cases where methionine sites are introduced strategically by mutation (Leahy *et al.* 1994). Second, information from anomalous diffraction measurements can be coupled usefully with molecular replacement. In favorable situations, the sites of anomalous

https://doi.org/10.1017/S0033583514000018 Published online by Cambridge University Press

scatterers can be used to specify the translational positioning, or even the orientation, of a known structure into the subject crystal [e.g. Bes *et al.* (1999)]. More typically, when at least some part of a structure can be approximated by a replacement model, SAD phasing may be incorporated to remove phase bias and to define unmodeled portions of the structure. Indeed, in cases where the anomalous substructure is not readily solved, a molecular replacement solution may produce a readily interpretable Bijvoet-difference Fourier depiction of anomalous-scatterer positions. Combinations of molecular replacement with SAD, often called MR-SAD (Panjikar *et al.* 2009; Read & McCoy, 2011; Schuermann & Tanner, 2003), have become increasingly common; since 2009, 88 SAD structures (3%) included an MR component. Third, the element specificity of anomalous scattering provides a basis for unambiguous identification of ion sites. Classically, this has been done by taking measurements at diagnostic energies above and/or below that for an element in question. When the element in question has no resonance within the accessible energy range, ionic replacements have sometimes been used, such as Rb^+ or Tl^+ for K^+ (Zhou & MacKinnon, 2003) or Os(III) hexamine for Mg^{2+} hydrates (Kazantsev *et al.* 2009). Alternatively, we have shown that robust identifications can also be made through the refinement of f'' values against Bijvoet mates (Liu *et al.* 2013), especially for data measured at a low x-ray energy. Fourth, the definitive character of MAD phasing can lead to experimental phases with exceptional accuracy, and this can be useful in providing model-free pictures of solvation, motion and conformational heterogeneity (Burling *et al.* 1996). Accurate experimental phases from MAD phasing are particularly useful for analyses at low resolution (Bass *et al.* 2002; Rudenko *et al.* 2002).

7.2 Changing practice in *de novo* phase evaluation

The early history of macromolecular crystallography related almost entirely to *de novo* structure determination done by MIR. All structures presented at a watershed Cold Spring Harbor Meeting in 1971 (Watson, 1972) either came directly from MIR (some reinforced with anomalous scattering, i.e. MIRAS or SIRAS) or were from difference-Fourier analyses of isomorphous ligand complexes. The first textbook on protein crystallography (Blundell & Johnson, 1976) treated MIR extensively; however, while it gave excellent discussions of anomalous scattering and molecular replacement, there was almost nothing to report in practice. The emphasis for molecular replacement, as in the monograph by Rossmann (1972) by this very name, was on solving for phases from symmetry transformations. Molecular replacement as we know it now, for repositioning a known structure into the crystal of a relative, was mentioned only in citing a test at 5 Å resolution on seal myoglobin (Tollin, 1969). As recounted above, it was not until 1981 that anomalous scattering was shown to suffice on its own for biological structure determination (Hendrickson & Teeter, 1981), introducing what we now call SAD. MAD followed in 1985 (Hendrickson, 1985; Kahn *et al.* 1985), but it took time for it to develop (Hendrickson, 1991).

Meanwhile, structure determination practices were changing rapidly. By the time that MAD was emerging, structures by *de novo* methods were already the minority fraction. We know this from statistics in our *Macromolecular Structures* (*MS*) series (Hendrickson & Wüthrich, 2000). Starting in 1990 and on for a decade, through 1999, we compiled annual *MS* books of abstracts for structures meeting criteria of: (1) having an atomic model, (2) being experimentally derived, (3) being novel and (4) being macromolecular. *MS 1991* (Hendrickson & Wüthrich, 1991), reporting structures published in 1990, described 135 structures meeting the *MS* criteria as compared to 236 PDB depositions in total (57%); the remainder were from follow-up studies that had

https://doi.org/10.1017/S0033583514000018 Published online by Cambridge University Press

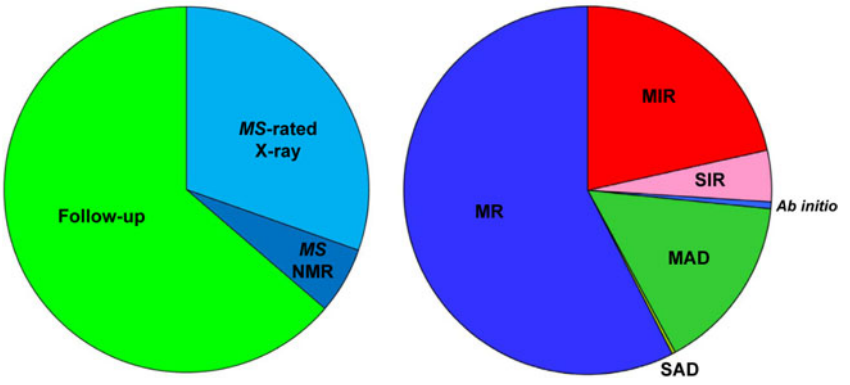


Fig. 12. Distribution of macromolecular structures of 1999 by determination method. (Left) Pie chart showing distribution among *MS*-rated (novel) and follow-up structures. (Right) Pie chart of *MS*-rated x-ray structures as divided among molecular replacement structures (MR) and *de novo* structures as determined by various methods.

failed the novelty test, mainly as being ligand complexes or mutant variants of previously deposited structures. Already by 1990, the majority of *MS*-qualified x-ray structures (57/111) were determined by molecular replacement; among the 51 *de novo* structures, there were 78% by MIR (40), 12% by SIR (6), 8% by MAD (4) and one virus by *ab initio* molecular averaging. By 1999 (*MS* 2000), the number of *MS*-rated structures had risen to 929 but as a fraction of all PDB entries it had fallen to 36%. Among *MS*-rated x-ray structures, again a clear majority were determined by molecular replacement (58%) but the breakdown of *de novo* strutures was evolving to an increasing fraction done by MAD (Fig. 12).

To chronicle the evolution of MAD phasing applications, I undertook to produce a record of all published MAD structures meeting the MS criteria. An account of the progression through 1998 showed a slow start (1–4 entries per year during 1988–1993) followed by an exponential rise to 58 in 1998 (Hendrickson, 1999). This compilation was continued through 2001 (Fig. 13; Supplemental Table 1), at which point the growing flood of MAD structures became overwhelming (260 recorded for 2001); indeed, some might have been missed. Moreover, SAD structures were becoming numerous (21 for 2001); and a separate chronicle was developed for SAD, retrospectively including earlier off-resonance anomalous diffraction structures (Supplemental Table 2). In compiling structural results from the publications, rather than PDB entries, it was possible to glean experimental details on the x-ray source and phasing element. The breakdown on x-ray beamlines (Supplemental Table 3; Hendrickson, 1999, 2000) showed that by 2001 novel MAD and SAD structures had been determined at 44 different beamlines from 13 synchrotrons and also from CuKα sources. The breakdown on phasing elements shows coverage across the periodic table (Fig. 14; Supplemental Table 4; Hendrickson, 1999). These MAD and SAD experiments reflect all of the categories of application described above in section 7.1; however, not surprisingly, the distribution is far from uniform. Most notably, Se predominates (66% overall); however, the conventional heavy atoms (Hg, Pt, Au, Os), the transition metals of metalloproteins (Fe, Cu, Zn), and bromide ions and brominated nucleic acids (Br) also showed appreciable fractions at 13, 10 and 5%, respectively. Besides the phasing elements used in MAD and SAD structures through 2001 (Supplemental Tables 1 & 2), a few others that have been used subsequently (identified in Supplemental Table 5) are also included in the periodic chart (Fig. 14) but not in the bar graph.

https://doi.org/10.1017/S0033583514000018 Published online by Cambridge University Press

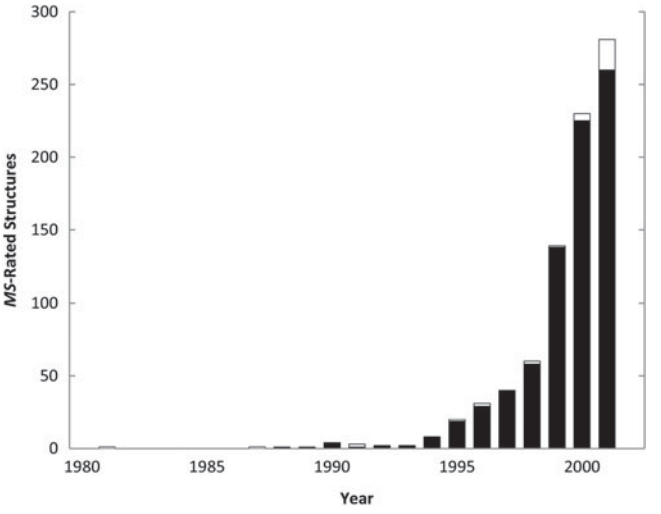


Fig. 13. Early growth in the production of anomalous diffraction structures. The number of structures based on MAD or SAD phase evaluation, published in a given calendar year, and meeting the criteria for inclusion in *Macromolecular Structures* (*MS*-rated; Hendrickson & Wüthrich, 1991) is plotted year by year. Structures included here are identified in Supplemental Tables 1 & 2 and the statistics are summarized in Supplemental Tables 3 & 4. MAD structures are in solid black and the SAD structures are in open boxes. Updated from Fig. 1 of Hendrickson (1999) with permission from the *Journal of Synchrotron Radiation*, © 1999 International Union of Crystallography.

An alternative approach to the tracking of trends in structure determination became available when, starting in 1996, the PDB began to include details of the kind that were being abstracted into MS entries. To retrieve information of interest from PDB header records, we developed a parser for extracting and categorizing the depositor declarations on 'method used to determine the structure.' We sorted those entries from *de novo* determinations into MIR, SIR, MAD, SAD and *ab initio* categories, including variant descriptors such as MIRAS, SIRAS, SE-MET MAD PHASING, SAS or DIRECT METHODS, respectively, into these groupings. The progression of changing practice in *de novo* structure determination is illustrated in Fig. 15. Three things are notable. First, over this period of record, the total number of *de novo* structures grew dramatically from an annual average of ~300 in the span from 1997 to 2000 to nearly 1400 annually ten years later in the span from 2007 to 2010. The drop to ~1100 *de novo* structures per year in the last two full years can be ascribed to the ending of the high-throughput phase of PSI; whereas annual PSI output of SAD plus MAD structures averaged 597 during 2007–2009, this dropped to 521 in 2010 (transition year between PSI-2 and PSI:Biology) and then to 230 on average during PSI:Biology years 2011–2012. Second, promoting the growth in total *de novo* structures, there was a dramatic shift in market share from analyses by isomorphous replacement (MIR and SIR, reddish bars) to analyses by anomalous diffraction (MAD and SAD, greenish bars) analyses. At the start of the period, 80% of *de novo* structures were done by MIR or SIR; whereas a decade later, nearly 90% were done by MAD or SAD. Already by 2000, more structures were done by MAD than by MIR and SIR combined. Third, SAD now predominates altogether. SAD overtook MAD in 2006, and now nearly 70% of all *de novo* structures are SAD structures.

I estimate that, over all of biological crystallography through 2012, 77% of the x-ray crystal structures determined by *de novo* methods were done by MAD or SAD. For the period of

https://doi.org/10.1017/S0033583514000018 Published online by Cambridge University Press

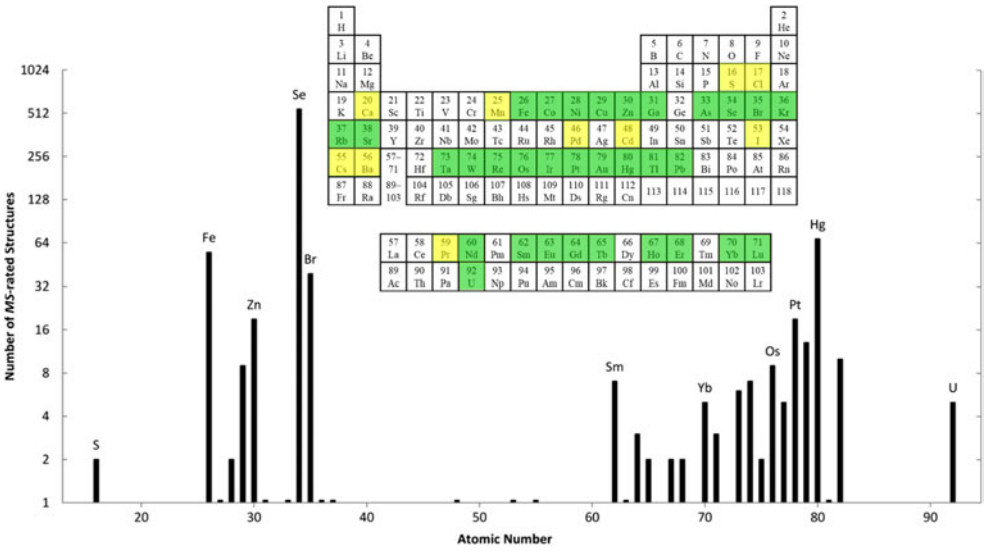


Fig. 14. Distribution of phasing elements used in anomalous diffraction analyses. The numbers of *MS*-rated (Hendrickson & Wüthrich, 1991) MAD or SAD structures published through 2001 (Supplemental Tables 1 & 2) are plotted as a function of the atomic numbers of elements used in those phase evaluations (Supplemental Table 4). If an analysis used more than one anomalous element, all are included. A logarithmic scale ($\log_2$) is used because of the great disparity between numbers for Se and other elements (Se:Hg≈8; Se:Fe≈10). Atomic numbers having $N=1$ are distinguished from those with $N=0$ (logN indefinite) by small bars. The periodic table shown as an inset identifies elements that have been used for novel structure determinations by MAD or SAD phasing. Elements are colored green if any reported analysis included measurements made at resonance, whether by MAD or SAD; elements are colored yellow if only off-resonance measurements were used in the reported analyses. Elements first used in studies published after 2001 (Cl, Ca, Mn, Sr, Pd, Ba, Nd and Pr) are identified in Supplemental Table 5.

1960–1996 (from atomic-level myoglobin until PDB deposits came to include methods records), I estimate that 430 MS-rated structures were done by MIR and SIR. This number is deduced by starting from our estimate of 2346 total x-ray structures through 1996 (Hendrickson & Wüthrich, 2000), assuming a *de novo* fraction (21.6%) as found for 1990 (MS 1991), subtracting the 67 MAD and 7 SAD structures recorded from publications (Supplemental Tables 1 & 2), and neglecting *ab initio* structures. For the period of 1997–2012, based on PDB records, the integrated totals for MAD and SAD (11 395 for 79.6%) are much larger than those for MIR and SIR (2604 for 18%). By summing over the two periods, I arrive at the 77% estimate for MAD and SAD overall. A small but steady residual fraction (~3% overall) of *de novo* structures have come from *ab initio* methods. It should be noted that another small fraction of *de novo* determinations (<1%) have made use of multiple *de novo* methods (mainly MAD + MIR). Of course, nearly all current *de novo* structure determinations take advantage of density modification, and some also include molecular replacement information. That said, generally speaking, MAD and SAD are self-sufficient and reliably effective methods. Indeed, that effectiveness was clearly such as to motivate the community to adopt MAD and then SAD as methods of choice. Associated with this shift in behaviour, the increases from MAD and SAD after 2000 correlate with a concomitant growth in total *de novo* structure determinations (Fig. 15), which is consistent with efficiency of the anomalous diffraction approach.

https://doi.org/10.1017/S0033583514000018 Published online by Cambridge University Press

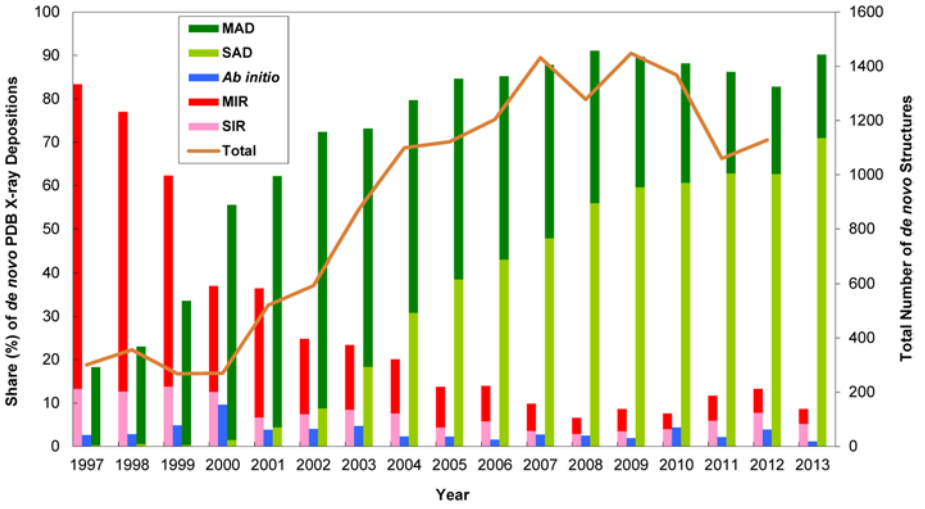


Fig. 15. Changing practice in *de novo* structure determination. Depositor declarations for the method of structure determination were parsed from PDB entries beginning in 1997, the first full year for such recordings. Declarations judged to be synonymous were reduced to five *de novo* options; thus, for example, MIR includes MIRAS determinations, SIR includes SIRAS, and *ab initio* includes 'direct methods'. Numbers were normalized into percentages of total *de novo* determinations in each year, and plotted in histograms identified by color coding. Multiple declarations (e.g. SIR and MAD) were counted for each method, but there were few such instances. The total number of *de novo* structures is plotted year by year as identified; 2013 is excluded because of incompleteness in releases. Results here are from the PDB holdings of 10 December 2013. Updated and elaborated from Fig. 3 of Hendrickson (2013) with permission from *Acta Crystallographica A*, © 2013 International Union of Crystallography.

7.3 Future prospects

Anomalous diffraction has moved from being ancillary to predominating for *de novo* structure determination, and this dominance seems likely to continue for some years. Indeed, it is difficult to imagine what might overtake SAD. An advance in *ab initio* methods could happen, but there is no evidence in the historical record (Fig. 15) for incipient ascendency; moreover, theoretical and practical limitations on size and resolution complicate broad applicability. Molecular replacement will doubtless continue to advance as more 'knowns' become available to serve as templates and improved methods permit the detection of ever smaller portions of unknown structures. Still, truly novel structures will demand *de novo* phase evaluation, and many of these structures are likely to present crystallographic challenges.

The frontier for MAD and SAD technology lies in efficaciously tackling small crystals of large, radiation-sensitive molecular systems that diffract no better than to moderate resolution. Progress with such challenges will likely need one or another improvement: sound analytical procedures to address extremes of both very weak and extraordinarily large signals, enhancement of anomalous signals such as from low-energy x-rays, possibly a new generation of phasing vehicles, enhancement of signal-to-noise ratios as from multiple crystals, improved computational procedures for handling data from multiple crystals, appropriate instrumentation for advanced x-ray sources and the presentation of small crystals, and appropriate applications to test the limits. Many applications will benefit from a coordinated implementation of multiple methodological improvements.

One challenging area where advanced methods might contribute concerns structure determination at low resolution, where atomic models are ill-specified. Solvent flattening approaches lose

https://doi.org/10.1017/S0033583514000018 Published online by Cambridge University Press

effectiveness at lower resolution, which makes accurate ambiguity resolution for low-resolution SAD structures problematic (although molecular averaging remains highly effective). On the other hand, properly performed MAD experiments can yield highly accurate phases (Burling *et al.* 1996) whenever the anomalous sites can be determined accurately. Situations are becoming commonplace where even best crystals diffract poorly (d_{min} ~4 Å); examples include membrane proteins (Bass *et al.* 2002; Nyblom *et al.* 2013), receptors (Rudenko *et al.* 2002), flexible multi-domain proteins (DeLaBarre & Brünger, 2003; Leonard *et al.* 2011) and multi-component complexes (Ataide *et al.* 2011). Accurate structures at much lower resolution (~6 Å) have proved to be highly informative, as was true for the globins (Kendrew *et al.* 1958; Perutz *et al.* 1960) and hemerythrins (Hendrickson *et al.* 1975; Smith *et al.* 1983; Ward *et al.* 1975), and similar studies are likely to become important once again where intrinsic resolution is very limited.

Many low-resolution structures are from rather large systems, which in turn require strong anomalous scatterers for accurate phasing. An effective approach to strong anomalous signals has come through heavy-metal cluster compounds (Knäblein *et al.* 1997; Thygesen *et al.* 1996), which scatter as super-atoms when viewed at low-enough resolution ($d_{min} > 4$ Å). This can be a great boon to structure determination; e.g. we estimate that the complex of a Au_{102}–gold cluster compound (Jadzinsky *et al.* 2007) attached 4:1 to a 2.3 MDa tetramer of 5000-residue protein chains will produce Bijvoet-diffraction ratios (Eq. 9) of 110%. Signals of such strength will overpower many conventional phasing programs; indeed, complications were already noted from a tungsten-cluster complex with the LDL receptor (Rudenko *et al.* 2003). On the other hand, the exact MAD treatment (Eqs. 4–7) can handle such colossal MAD signals readily, as was demonstrated for the very large (~27%) Ho L_{III} signals from the mannose-binding protein (Weis *et al.* 1991). Colossal cluster compounds can also be used at resolutions beyond the super-atom level; however, the cluster must then be resolved into its constituent atoms, and procedures have been described for doing so (Dahms *et al.* 2013). Moreover, the low-resolution phases can be used in Bijvoet-difference analyses to find the potential sites of anomalous scattering in a very large structure, as for selenomethionyl proteins, halide soaks or some heavy-atom derivatives.

Heavy-atom cluster compounds are important as new phasing vehicles, as noted above, but other directions are in progress as well. Selenated compounds alternative to the predominating selenomethionine have also been introduced by biological synthesis into proteins, using selenatryptophan analogs of tryptophan (Bae *et al.* 2001; Boles *et al.* 2002) or selenocysteine to generate cystine-like diselenides (Strub *et al.* 2003), and by chemical synthesis into nucleic acids (Sheng & Huang, 2010). Chelating frameworks for incorporating lanthanide ions (Girard *et al.* 2003; Purdy *et al.* 2002) provide potentially effective means for introducing these potent anomalous centers. Uranyl compounds can provide spectacular anomalous signals (Fig. 1*b*; Liu *et al.* 2001) and chelated delivery for such phasing vehicles might also be made possible. No doubt, other possibilities will emerge.

Another challenging category of application concerns native structures that are composed only from light atoms. Here the multi-crystal native SAD approach described above in Section 6.4 becomes important. The anomalous signals from the intrinsic sulfur and phosphorous atoms in biological macromolecules can be enhanced remarkably by using low-energy x-rays (Liu *et al.* 2013). This enhancement comes about as the K (1 s) resonances for S and P atoms at 2.472 keV (5.016 Å) and 2.1455 keV (5.779 Å), respectively, are approached. Thus, f'' signals increase with lower energy; however, so too does x-ray absorption and background scattering. Indeed, x-ray penetration depth decreases with energy such that the transmitted signals actually diminish for larger crystals (Fig. 16). By Bragg's law, scattering angles within diffraction patterns

https://doi.org/10.1017/S0033583514000018 Published online by Cambridge University Press

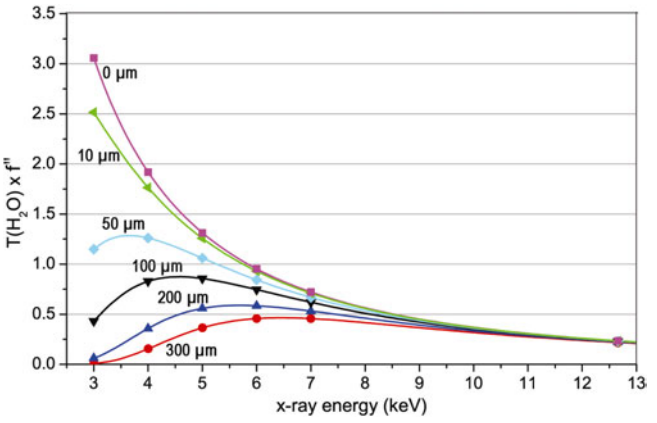


Fig. 16. Dependence of transmitted anomalous signals on x-ray energy. To understand the effects of x-ray absorption on the anomalous signal to be expected from sulfur (f' in electrons) in crystals of varying size, the x-ray transmission through various thicknesses of water was calculated as a function of x-ray energy. Water is used as having x-ray absorptivity similar to that of a protein crystal. Transmitted signals are plotted for various thicknesses; symbols identify the Se K-edge energy and selected relevant low-energy values. Reproduced from Fig. 2 of Liu *et al.* (2013) with permission from *Acta Crystallographica D*, © 2013 by the International Union of Crystallography.

also increase. These factors cause complications, but ones that can be addressed by appropriate instrumentation and experimental design (Liu *et al.* 2013). Notably, the use of small crystals becomes imperative below 4 keV. An additional complication at lower energy comes from the contaminating anomalous noise from other organic elements (C, N and O), which is negligible at higher energy but increases at lower energy. Taking as an example the structure of the TehA homolog of anion channel SLAC1 (Chen *et al.* 2010; PDB: 3M71), for which the ordered protein and water has the composition C_{1701} N_{350} O_{401} S_9, we find that the intrinsic signal-to-noise ratio [rms(f''_S)/rms($f''_{C,N,O}$)] drops from 1.98 at 7 keV to 1.49 at 3 keV. Such CNO noise should not become a problem, however, unless the sulfur/phosphorous content is low.

As discussed in Section 5.4 above, there are many existing computational options for analyzing MAD and SAD data. We can also expect continuing developments of direct and indirect relevance, notably from the PHENIX and CCP4 consortia. Surely there is room for improvements, for example in phase evaluation from very large anomalous signals and in the treatment of probability. The more open questions, however, concern procedures for the optimal assembly of complete data sets from the partial observations recorded from many crystals or multiple regions within crystals. One active area of study in this regard concerns the problem of crystal clustering and outlier identification, with new contributions appearing shortly after we reported our first results (Foadi *et al.* 2013; Giordano *et al.* 2012; Liu *et al.* 2011, 2012) and programs to treat the single-shot data from serial femtosecond crystallography (White *et al.* 2013). Another area of importance relates to the control and adjustment for systematic errors, for example by preserving the identity of inverse-beam pairings for local scaling (Hendrickson & Teeter, 1981; Matthews & Czerwinski, 1975) prior to globally collective scaling. Perhaps most importantly, there is an opportunity to improve resolution and data quality through improved monitoring and selection based on understanding of mechanisms of radiation damage (Hendrickson, 1976; Warkentin & Thorne, 2010).

On the instrumentation side, exciting developments are in store from new x-ray sources as well as at beamlines discussed above in Section 5.3. Most recently, entirely new approaches to

https://doi.org/10.1017/S0033583514000018 Published online by Cambridge University Press

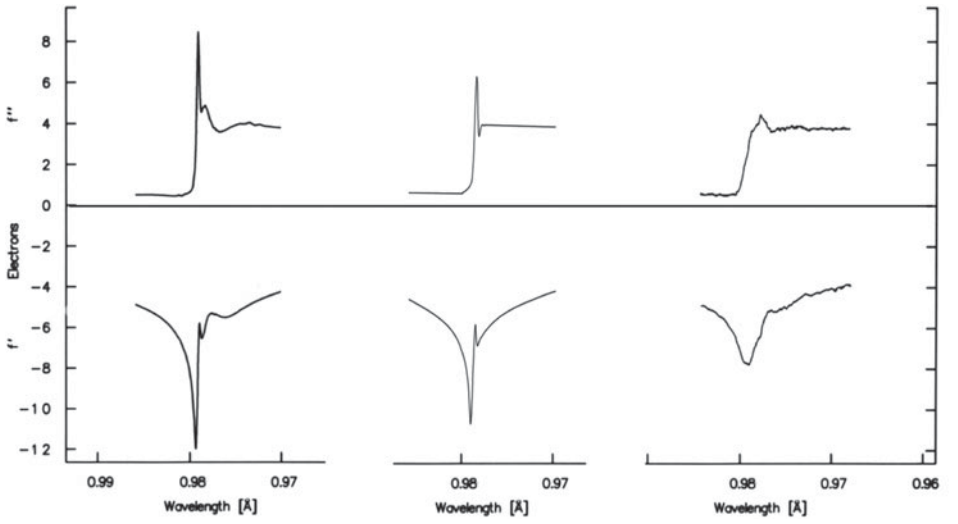


Fig. 17. Dependence of apparent resonance strength on energy resolution. (Left) Anomalous scattering spectra measured from an *E. coli* selenomethionyl thioredoxin crystal at SSRL beamline 1–5, without any focusing elements (Hendrickson *et al.* 1990). (Center) Spectra measured from CHO-produced selenomethionyl human chorionic gonadotropin at NSLS beamline X4A, with focusing (Wu *et al.* 1994). (Right) Spectra measured from selenobiotinyl streptavidin at horizontally focused PF beamline 14A, and analyzed subsequently relative to a comparable sample measured as above at SSRL 1–5 (Fanchon & Hendrickson, 1990). In each case, x-ray absorption spectra were measured from the selenoproteins by fluorescence, and then transformed into theoretical anomalous scattering for atomic selenium to yield experimental edge data on an absolute scale. The monochromator was of Si111 in all cases, and exactly the same analytical procedures and program were used in all cases (Hendrickson *et al.* 1988). Signal reductions are due to compromised energy resolution from beam divergences.

biological structure analysis are emerging from serial femtosecond crystallography (Chapman *et al.* 2011) at x-ray free electron lasers (FELs). This was first put into practice at Stanford's Linac Coherent Light Source (LCLS) and other sources are being developed worldwide. Radiation damage can be eliminated by using single-shot, diffraction-before-destruction procedures with nanocrystals, but challenges are faced in achieving the kind of accuracy required to measure anomalous diffraction signals. Nevertheless, results are beginning to emerge; tests with lysozyme show that SAD phasing of FEL data should become feasible (Barends *et al.* 2013, 2014). In addition, MAD-like methods are being tested based on FEL-induced stripping of electrons from heavy atoms (Son *et al.* 2011). New synchrotron sources are also under development. Most imminent are the National Synchrotron Light Source – II (NSLS-II) at Brookhaven, which is diffraction limited in the vertical, and MAX IV in Lund, which employs a novel multibend-achromat accelerator lattice to produce diffraction-limited emission horizontally as well as vertically. Upgrade implementations of such ultimate-storage-ring lattices are in planning for several existing machines.

The new synchrotron sources can readily produce very bright, low-divergence, micron-sized x-ray beams to match emerging needs generally. These qualities are especially important for the challenges of anomalous diffraction. For the optimization of MAD or SAD at resonance (Fig. 1), having high-energy resolution is critical. As shown in Fig. 17, poor energy resolution can spoil intrinsically sharp white-line features, such as at the K-edge resonance of selenomethionyl proteins. Signal strength is then reduced accordingly. We are developing microdiffraction

https://doi.org/10.1017/S0033583514000018 Published online by Cambridge University Press

beamline NYX at NSLS-II with a novel monochromator design that preserves intrinsic resonant signals. For the optimization of native SAD analyses, low-energy experiments are needed (see Section 6.4); here, when at the preferred lower reaches of energy, microcrystals and thereby microbeams are essential for optimization (Fig. 16). A new long-wavelength (low-energy) beamline (I23) is being built at Diamond, and a low-energy anomalous x-ray diffraction beamline (LAX) is approved for building at NSLS-II. For synchrotron beamlines as for FEL instruments, the appropriate delivery of small crystals into microbeams is a critical focus of instrument development. Detectors are another.

8. Supplementary material

Supplemental tables 1–5 can be found at http://dx.doi.org/10.1017/S0033583514000018.

9. Acknowledgements

I thank the many associates whose work over the past decades has contributed to our developments of MAD and SAD phase evaluation; most notably, I thank Janet Smith (our initial MAD phasing tests on lamprey hemoglobin), Arno Pähler (our first MAD application to an unknown, streptavidin), Craig Ogata (X4A synchrotron beamline) and Qun Liu (native SAD analyses). I also thank Arno Pähler for developing the PDB parser for analyzing the progress of *de novo* structure determinations, Jie Qin for help with figures, and Maggie Gabanyi and Helen Berman for PDB statistics on PSI structures. Finally, I am grateful to Qun Liu for his critical review and helpful comments. This work was supported in part by NIH grants GM034102 and GM095315, by NSF instrumentation grant DBI-0960034 (NYX beamline) and by the Howard Hughes Medical Institute.

10. References

Abrahams, J. P. & Leslie, A. G. W. (1996). Methods used in the structure determination of bovine mitochondrial F1 ATPase. *Acta Crystallographica D* **52**, 30–42.

Akey, D. L., Brown, W. C., Dutta, S., Konwerski, J., Jose, J., Jurkiw, T. J., DelProposto, J., Ogata, C. M., Skiniotis, G., Kuhn, R. J. & Smith, J. L. (2014). Flavivirus NS1 crystal structures reveal a surface for membrane association and regions of interaction with the immune system. *Science* **343**, 881–885.

Als-Nielsen, J. & McMorrow, D. (2011). *Elements of Modern X-ray Physics*, 2nd edn, Chichester: Wiley.

Amemiya, Y., Matsushita, T., Nakagawa, A., Satow, Y., Miyahara, M. & Chikawa, J.-I. (1988) Design and performance of an imaging plate system for X-ray diffraction study. *Nuclear Instruments and Methods A* **266**, 645–653.

Ataide, S. F., Schmitz, N., Shen, K., Ke, A., Shan, S. O., Doudna, J. A. & Ban, N. (2011). The crystal structure of the signal recognition particle in complex with its receptor. *Science* **331**, 881–886.

Bae, J. H., Alefelder, S., Kaiser, J. T., Friedrich, R., Moroder, L., Huber, R. & Budisa, N. (2001). Incorporation of β-selenolo[3,2-b]pyrrolyl-alanine into proteins for phase determination in protein X-ray crystallography. *Journal of Molecular Biology* **309**, 925–936.

Barends, T. R., Foucar, L., Botha, S., Doak, R. B., Shoeman, R. L., Nass, K., Koglin, J. E., Williams, G. J., Boutet, S., Messerschmidt, M. & Schlichting, I. (2014). *De novo* protein crystal structure determination from X-ray free-electron laser data. *Nature* **505**, 244–247.

Barends, T. R., Foucar, L., Shoeman, R. L., Bari, S., Epp, S. W., Hartmann, R., Hauser, G., Huth, M., Kieser, C., Lomb, L., Motomura, K., Nagaya, K., Schmidt, C., Strecker, R., Anielski, D., Boll, R., Erk, B., Fukuzawa, H., Hartmann, E., Hatsui, T., Holl, P., Inubushi, Y., Ishikawa, T., Kassemeyer, S., Kaiser, C., Koeck, F., Kunishima, N., Kurka, M., Rolles, D., Rudek, B., Rudenko, A., Sato, T., Schroeter, C. D., Soltau, H., Strueder, L., Tanaka, T., Togashi, T., Tono, K., Ullrich, J., Yase, S., Wada, S. I., Yao, M., Yabashi, M., Ueda, K. & Schlichting, I. (2013). Anomalous signal from S atoms in protein crystallographic data from an X-ray free-electron laser. *Acta Crystallographica D* **69**, 838–842.

Bass, R. B., Strop, P., Barclay, M. & Rees, D. C. (2002). Crystal structure of *Escherichia coli* MscS,

https://doi.org/10.1017/S0033583514000018 Published online by Cambridge University Press

a voltage-modulated and mechanosensitive channel. *Science* **298**, 1582–1587.

Bes, M. T., Parisini, E., Inda, L. A., Saraiva, L. M., Peleato, M. L. & Sheldrick, G. M. (1999). Crystal structure determination at 1.4 Å resolution of ferredoxin from the green alga *Chlorella fusca*. *Structure* **7**, 1201–1211.

Biertümpfel, C., Zhao, Y., Kondo, Y., Ramón-Maiques, S., Gregory, M., Lee, J. Y., Masutani, C., Lehmann, A. R., Hanaoka, F. & Yang, W. (2010). Structure and mechanism of human DNA polymerase η. *Nature* **465**, 1044–1048.

Bijvoet, J. M. (1949). Phase determination in direct fourier-synthesis of crystal structures. *Proceedings of the National Academy of Sciences Amsterdam B* **52**, 313–314.

Bijvoet, J. M. (1954). Structure of optically active compounds in the solid state. *Nature* **173**, 888–891.

Bijvoet, J. M., Bokhoven, C., & Schoone, J. C. (1948). On the crystal structure of strychnine sulfate and selenate. II. [010] projection and structure formula. *Koninklijke Nederlandse Akademie van Wetenschappen Proceedings* **51**, 990–990.

Bijvoet, J. M., Bokhoven, C. & Schoone, J. C. (1949). On the crystal structure of strychnine sulfate and selenate. III. [001] projection. *Koninklijke Nederlandse Akademie van Wetenschappen Proceedings* **52**, 120–121.

Bijvoet, J. M., Peerdeman, A. F. & van Bommel, A. J. (1951). Determination of the absolute configuration of optically active compounds by means of X-rays. *Nature* **168**, 271–272.

Biou, V., Shu, F. & Ramakrishnan, V. (1995). X-ray crystallography shows that translational initiation factor IF3 consists of two compact α/β domains linked by an α-helix. *EMBO Journal* **14**, 4056–4064.

Blow, D. M. (1958). The structure of haemoglobin. VII. Determination of phase angles in the non-centrosymmetric [100] zone. *Proceedings of the Royal Society A* **247**, 302–336.

Blow, D. M. (2003). How Bijvoet made the difference: the growing power of anomalous scattering. *Methods in Enzymology* **374**, 3–22.

Blow, D. M. & Crick, F. H. C. (1959). The treatment of errors in the isomorphous replacement method. *Acta Crystallographica* **12**, 794–802.

Blow, D. M. & Rossmann, M. G. (1961). The single isomorphous replacement method. *Acta Crystallographica* **14**, 1195–1202.

Blundell, T. L. & Johnson, L. N. (1976). *Protein Crystallography*. London: Academic Press.

Bokhoven, C., Schoone, J. C. & Bijvoet, J. M. (1951). The Fourier synthesis of the crystal structure of strychnine sulphate pentahydrate. *Acta Crystallographica* **4**, 275–280.

Boles, J. O., Henderson, J., Hatch, D. & Silks, L. A. (2002). Synthesis and incorporation of [6,7]-selenatryptophan into dihydrofolate reductase. *Biochemical and Biophysical Research Communication* **298**, 257–261.

Bonanno, J. B., Almo, S. C., Bresnick, A., Chance, M. R., Fiser, A., Swaminathan, S., Jiang, J., Studier, F. W., Shapiro, L., Lima, C. D., Gaasterland, T. M., Sali, A., Bain, K., Feil, I., Gao, X., Lorimer, D., Ramos, A., Sauder, J. M., Wasserman, S. R., Emtage, S., D'Amico, K. L. & Burley, S. K. (2005). New York-Structural GenomiX Research Consortium (NYSGXRC): a large scale center for the protein structure initiative. *Journal of Structural and Functional Genomics* **6**, 225–232.

Bragg, W. H. & Bragg, W. L. (1913). The structure of the diamond. *Proceedings of the Royal Society London A* **89**, 277–291.

Bragg, W. L. (1913a). The diffraction of short electromagnetic waves by a crystal. *Proceedings: Cambridge Philosophical Society* **17**, 43–57.

Bragg, W. L. (1913b). The structure of some crystals as indicated by their diffraction of x-rays. *Proceedings of the Royal Society London A* **89**, 248–277.

Bragg, W. L. (1914). The analysis of crystals by the x-ray spectrometer. *Proceedings of the Royal Society London A* **89**, 468–489.

Bragg, W. L. (1929). The determination of parameters in crystal structures by means of Fourier series. *Proceedings of the Royal Society London A* **123**, 537–559.

Brasch, J., Harrison, O. J., Ahlsen, G., Liu, G. & Shapiro, L. (2011). Crystal structure of the ligand binding domain of netrin G2. *Journal of Molecular Biology* **414**, 723–734.

Brodersen, D. E., de La Fortelle, E., Vonrhein, C., Bricogne, G., Nyborg, J. & Kjeldgaard, M. (2000). Applications of single-wavelength anomalous dispersion at high and atomic resolution. *Acta Crystallographica D* **56**, 431–441.

Burling, F. T., Weis, W. I., Flaherty, K. M. & Brunger, A. T. (1996). Direct observation of protein solvation and discrete disorder with experimental crystallographic phases. *Science* **271**, 72–77.

Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Kundrot, C. E., Cech, T. R. & Doudna, J. A. (1996). Crystal structure of a group I ribozyme domain: principles of RNA packing. *Science* **273**, 1678–1685.

Cate, J. H., Yusupov, M. M., Yusupova, G. Z., Earnest, T. N. & Noller, H. F. (1999). X-ray crystal structures of 70S ribosome functional complexes. *Science* **285**, 2095–2104.

Chandra, V., Huang, P., Hamuro, Y., Raghuram, S., Wang, Y., Burris, T. P. & Rastinejad, F. (2008). Structure of the intact PPAR-gamma-RXR- nuclear receptor complex on DNA. *Nature* **456**, 350–356.

Chapman, H. N., Fromme, P., Barty, A., White, T. A., Kirian, R. A., Aquila, A., Hunter, M. S., Schulz, J., DePonte, D. P., Weierstall, U., Doak, R. B.,

https://doi.org/10.1017/S0033583514000018 Published online by Cambridge University Press

Maia, F. R., Martin, A. V., Schlichting, I., Lomb, L., Coppola, N., Shoeman, R. L., Epp, S. W., Hartmann, R., Rolles, D., Rudenko, A., Foucar, L., Kimmel, N., Weidenspointner, G., Holl, P., Liang, M., Barthelmess, M., Caleman, C., Boutet, S., Bogan, M. J., Krzywinski, J., Bostedt, C., Bajt, S., Gumprecht, L., Rudek, B., Erk, B., Schmidt, C., Hömke, A., Reich, C., Pietschner, D., Strüder, L., Hauser, G., Gorke, H., Ullrich, J., Herrmann, S., Schaller, G., Schopper, F., Soltau, H., Kühnel, K. U., Messerschmidt, M., Bozek, J. D., Hau-Riege, S. P., Frank, M., Hampton, C. Y., Sierra, R. G., Starodub, D., Williams, G. J., Hajdu, J., Timneanu, N., Seibert, M. M., Andreasson, J., Rocker, A., Jönsson, O., Svenda, M., Stern, S., Nass, K., Andritschke, R., Schröter, C. D., Krasniqi, F., Bott, M., Schmidt, K. E., Wang, X., Grotjohann, I., Holton, J. M., Barends, T. R., Neutze, R., Marchesini, S., Fromme, R., Schorb, S., Rupp, D., Adolph, M., Gorkhover, T., Andersson, I., Hirsemann, H., Potdevin, G., Graafsma, H., Nilsson, B. & Spence, J. C. (2011). Femtosecond X-ray protein nanocrystallography. *Nature* **470**, 73–77.

Chen, L., Glover, J. N., Hogan, P. G., Rao, A. & Harrison, S. C. (1998). Structure of the DNA-binding domains from NFAT, Fos and Jun bound specifically to DNA. *Nature* **392**, 42–48.

Chen, L., Rose, J. P., Breslow, E., Yang, D., Chang, W. R., Furey, W. F. Jr., Sax, M. & Wang, B.-C. (1991). Crystal structure of a bovine neurophysin II dipeptide complex at 2.8 Å determined from the single-wavelength anomalous scattering signal of an incorporated iodine atom. *Proceedings of the National Academy of Sciences of the United States of America* **88**, 4240–4244.

Chen, Y. H., Hu, L., Punta, M., Bruni, R., Hillerich, B., Kloss, B., Rost, B., Love, J., Siegelbaum, S. A. & Hendrickson, W. A. (2010). Homologue structure of the SLAC1 anion channel for closing stomata in leaves. *Nature* **467**, 1074–1080.

Chen, Z., Yang, H. & Pavletich, N. P. (2008). Mechanism of homologous recombination from the RecA-ssDNA/dsDNA structures. *Nature* **453**, 489–494.

Cho, H. S. & Leahy, D. J. (2002). Structure of the extracellular region of HER3 reveals an interdomain tether. *Science* **297**, 1330–1333.

Conner, B. N., Takano, T., Tanaka, S., Itakura, K. & Dickerson, R. E. (1982). The molecular structure of d(^{I}CpCpGpG), a fragment of right-handed double helical A-DNA. *Nature* **295**, 294–299.

Coster, D., Knol, K. S. & Prins, J. A. (1930). Unterschiede in der Intensität der Röntgenstrahlenreflexion an den beiden 111-Flächen der Zinkblende. *Zeitschrift für Physik* **63**, 345–369.

Cowtan, K. (1994). *DM, an Automated Procedure for Phase Improvement by Density Modification*. Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography no 31, pp. 34–48. Warrington, UK: SERC Daresbury Laboratory.

Cowtan, K. (2006). The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallographica D* **62**, 1002–1011.

Cowtan, K. (2010). Recent developments in classical density modification. *Acta Crystallographica D* **66**, 470–478.

Cowtan, K. D. & Main, P. (1993). Improvement of macromolecular electron-density maps by the simultaneous application of real and reciprocal space constraints. *Acta Crystallographica D* **49**, 148–157.

Cowtan, K .D. & Zhang, K. Y. J. (1999). Density modification for macromolecular phase improvement. *Progress in Biophysics and Molecular Biology* **72**, 245–270.

Cramer, P., Bushnell, D. A. & Kornberg, R. D. (2001). Structural basis of transcription: RNA polymerase II at 2.8 Ångstrom resolution. *Science* **292**, 1863–1876.

Cramer, S. P. & Hodgson, K. O. (1979). X-ray absorption spectroscopy: a new structural method and its applications to bioinorganic chemistry. *Progress in Inorganic Chemistry* **25**, 1–39.

Cromer, D. T. & Liberman, D. (1970). Relativistic calculation of anomalous scattering factors for X-rays. *Journal of Chemical Physics* **53**, 1891–1898.

Dahms, S. O., Kuester, M., Streb, C., Roth, C., Sträter, N. & Than, M. E. (2013). Localization and orientation of heavy-atom cluster compounds in protein crystals using molecular replacement. *Acta Crystallographica D* **69**, 284–297.

Dauter, Z. & Adamiak, D. A. (2001). Anomalous signal of phosphorus used for phasing DNA oligomer: importance of data redundancy. *Acta Crystallographica D* **57**, 990–995.

Dauter, Z. & Dauter, M. (1999). Anomalous signal of solvent bromides used for phasing of lysozyme. *Journal of Molecular Biology* **289**, 93–101.

Dauter, Z. & Dauter, M. (2001). Entering a new phase: using solvent halide ions in protein structure determination. *Structure* **9**, R21–R26.

Dauter, Z., Dauter, M. & Dodson, E. (2002). Jolly SAD. *Acta Crystallographica D* **58**, 494–506.

Dauter, Z., Dauter, M., de La Fortelle, E., Bricogne, G. & Sheldrick, G. M. (1999). Can anomalous signal of sulfur become a tool for solving protein crystal structures? *Journal of Molecular Biology* **289**, 83–92.

Deacon, A. M. & Ealick, S. E. (1999). Selenium-based MAD phasing: setting the sites on larger structures. *Structure* **15**, R161–166.

DeLaBarre, B. & Brünger, A. T. (2003). Complete structure of p97/valosin-containing protein reveals communication between nucleotide domains. *Nature Structural Biology* **10**, 856–863.

https://doi.org/10.1017/S0033583514000018 Published online by Cambridge University Press

de la Fortelle, E. & Bricogne, G. (1997). Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. *Methods in Enzymology* **276**, 472–494.

Doublié, S. (2007). Production of selenomethionyl proteins in prokaryotic and eukaryotic expression systems. *Methods in Molecular Biology* **363**, 91–108.

Dyda, F., Hickman, A. B., Jenkins, T. M., Engelman, A., Craigie, R. & Davies, D. R. (1994). Crystal structure of the catalytic domain of HIV-1 integrase: similarity to other polynucleotidyl transferases. *Science* **266**, 1981–1986.

Evans, G. & Bricogne, G. (2003). Triiodide derivatization in protein crystallography. *Acta Crystallographica D* **59**, 1923–1929.

Evans, G. & Pettifer, R. (2001). *CHOOCH*: a program for deriving anomalous-scattering factors from X-ray fluorescence spectra. *Journal of Applied Crystallography* **34**, 82–86.

Evans, P. R. (2011). An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallographica D* **67**, 282–292.

Ewald, P. P. & Hermann, C. (1927). Gilt der Friedelsche Satz über die Symmetrie der Röntgensinterferenzen? *Zeitschrift für Kristallographie* **65**, 251–260.

Fan, H.-F. & Gu, Y.-X. (1985). Combining direct methods with isomorphous replacement or anomalous scattering data. III. The incorporation of partial structure information. *Acta Crystallographica A* **41**, 280–284.

Fanchon, E. & Hendrickson, W. A. (1990). Effect of anisotropy in anomalous scattering on the MAD phasing method. *Acta Crystallographica A* **46**, 809–820.

Feng, L., Campbell, E. B., Hsiung, Y. & MacKinnon, R. (2010). Structure of a eukaryotic CLC transporter defines an intermediate state in the transport cycle. *Science* **330**, 635–641.

Foadi, J., Aller, P., Alguel, Y., Cameron, A., Axford, D., Owen, R. L., Armour, W., Waterman, D. G., Iwata, S. & Evans, G. (2013). Clustering procedures for the optimal selection of data sets from multiple crystals in macromolecular crystallography. *Acta Crystallographica D* **69**, 1617–1632.

Fourme, R. & Hendrickson, W. A. (1990). Analysis of macromolecular structures by the method of multiwavelength anomalous diffraction. In *Synchrotron Radiation and Biophysics* (ed. S. S. Hasnain), pp. 156–175. Chichester: Ellis Horwood Limited.

Friedrich, W., Knipping, P. & Laue, M. (1912). Interferenz-Erscheinungen bei Röntgenstrahlen. *Sitzungsberichte Der Akademie Der Wissenschaften* 303–322.

Furey, W. & Swaminathan, S. (1997). PHASES-95: a program package for processing and analyzing diffraction data from macromolecules. *Methods in Enzymology* **276**, 590–620.

Geiger, J. H., Hahn, S., Lee, S. & Sigler, P. B. (1996). Crystal structure of the yeast TFIIA/TBP/DNA complex. *Science* **272**, 830–836.

Giordano, R., Leal, R. M., Bourenkov, G. P., McSweeney, S. & Popov, A. N. (2012). The application of hierarchical cluster analysis to the selection of isomorphous crystals. *Acta Crystallographica D* **68**, 649–658.

Girard, E., Stelter, M., Vicat, J. & Kahn, R. (2003). A new class of lanthanide complexes to obtain high-phasing-power heavy-atom derivatives for macromolecular crystallography. *Acta Crystallographica D* **59**, 1914–1922.

Graves, B. J., Hatada, M. H., Hendrickson, W. A., Miller, J. K., Madison, V. S. & Satow, Y. (1990). Structure of interleukin-1α at 2.7 Å resolution. *Biochemistry* **29**, 2679–2684.

Green, D. W., Ingram, V. M. & Perutz, M. F. (1954). The structure of haemoglobin IV. Sign determination by the isomorphous replacement method. *Proceedings of the Royal Society London A* **225**, 287–307.

Guss, J. M., Merritt, E. A., Phizackerley, R. P., Hedman, B., Murata, M., Hodgson, K. O. & Freeman, H. C. (1988). Phase determination by multiple-wavelength x-ray diffraction: crystal structure of a basic "blue" copper protein from cucumbers. *Science* **241**, 806–811.

Harada, S., Yasui, M., Murakawa, K., Kasai, N. & Satow, Y. (1986). Crystal structure analysis of cytochrome c′ by the multiwavelength anomalous diffraction method using synchrotron radiation. *Journal of Applied Crystallography* **19**, 448–452.

Harris, S. F. & Botchan, M. R. (1999). Crystal structure of the human papillomavirus type 18 E2 activation domain. *Science* **284**, 673–677.

Harvey, I., Hao, Q., Duke, E. M., Ingledew, J. & Hasnain, S. S. (1998). Structure determination of a 16.8 kDa copper protein at 2.1 Å resolution using anomalous scattering data with direct methods. *Acta Crystallographica D* **54**, 629–635.

Hauptman, H. A. (1997). Shake-and-bake: an algorithm for automatic solution *ab initio* of crystal structures. *Methods in Enzymology* **277**, 3–13.

Hauptman, H. & Karle, J. (1953). Solution of the Phase Problem I. The Centrosymmetric Crystal. ACA Monograph No. 3. Polycrystal Book Service.

Hendrickson, W. A. (1971). Some aids for breaking the phase ambiguity in the single isomorphous replacement methods. *Acta Crystallographica B* **27**, 1474–1475.

Hendrickson, W. A. (1976). Radiation damage in protein crystallography. *Journal of Molecular Biology* **106**, 889–893.

Hendrickson, W. A. (1979). Phase information from anomalous-scattering measurements. *Acta Crystallographica A* **35**, 245–247.

Hendrickson, W. A. (1984). Measurement and use of anomalous x-ray scattering. *Acta Crystallographica A* **40**, C3.

https://doi.org/10.1017/S0033583514000018 Published online by Cambridge University Press

Hendrickson, W. A. (1985). Analysis of protein structure from diffraction measurements at multiple wavelengths. *Transactions of the American Crystallographic Association* **21**, 11–21.

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

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

Hendrickson, W. A. (2000). Synchrotron crystallography. *Trends in Biochemical Sciences* **12**, 637–643.

Hendrickson, W. A. (2013). Evolution of diffraction methods for solving crystal structures. *Acta Crystallographica A* **69**, 51–59.

Hendrickson, W. A., Co, M. S., Smith, J. L., Hodgson, K. O. & Klippenstein, G. L. (1982). X-ray absorption spectroscopy of the dimeric iron site in azidomethemerythrin from *Phascolopsis gouldii*. *Proceedings of the National Academy of Sciences of the United States of America* **79**, 6255–6259.

Hendrickson, W. A., Horton, J. R. & LeMaster, D. M. (1990). Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. *EMBO Journal* **9**, 1665–1672.

Hendrickson, W. A., Klippenstein, G. L. & Ward, K. B. (1975). Tertiary structure of myohemerythrin at low resolution. *Proceedings of the National Academy of Sciences of the United States of America* **72**, 2160–2164.

Hendrickson, W. A. & Ogata, C. M. (1997). Phase Determination from Multiwavelength Anomalous Diffraction Measurements. *Methods in Enzymology* **276**, 494–523.

Hendrickson, W. A., Pähler, A., Smith, J. L., Satow, Y., Merritt, E. A. & Phizackerley, R. P. (1989). Crystal structure of core streptavidin determined from multiwavelength anomalous diffraction of synchrotron radiation. *Proceedings of the National Academy of Sciences of the United States of America* **86**, 2190–2194.

Hendrickson, W. A. & Sheriff, S. (1987). General density function corresponding to X-ray diffraction with anomalous scattering included. *Acta Crystallographica A* **43**, 121–125.

Hendrickson, W. A., Smith, J. L., Phizackerley, R. P. & Merritt, E. A. (1988). Crystallographic structure analysis of lamprey hemoglobin from anomalous dispersion of synchrotron radiation. *Proteins* **4**, 77–88.

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

Hendrickson, W. A. & Wüthrich, K. (eds.) (1991). *Macromolecular Structures 1991*. London: Current Biology.

Hendrickson, W. A. & Wüthrich, K. (eds.) (2000). *Macromolecular Structures 2000*. London: Current Biology.

Herriott, J. R., Sieker, L. C., Jensen, L. H. & Lovenberg, W. (1970). Structure of rubredoxin: an X-ray study to 2.5 Å resolution. *Journal of Molecular Biology* **50**, 391–406.

Herzenberg, A. & Lau, H. S. M. (1967). Anomalous scattering and the phase problem. *Acta Crystallographica* **22**, 24–28.

Höhl, H. (1933). Zur Dispersionstheorie der Röntgenstrahlen. *Zeitschrift für Physik* **84**, 1–16.

Hope, H. & de la Camp, U. (1972). Anomalous scattering by oxygen: measurements on (+)-tartaric acid. *Acta Crystallographica A* **28**, 201–207.

Hoppe, W. & Jakubowski, V. (1971). The determination of phases of erythrocruorin using the two-wavelength method with iron as anomalous scatterer. In *Anomalous Scattering* (eds. S. Ramaseshan, & S. Abrahams), pp. 437–461. Copenhagen: Munksgård.

Hubbard, S. R., Wei, L., Ellis, L. & Hendrickson, W. A. (1994). Crystal structure of the tyrosine kinase domain of the human insulin receptor. *Nature* **372**, 746–754.

Istvan, E. S., Palnitkar, M., Buchanan, S. K. & Deisenhofer, J. (2000). Crystal structure of the catalytic portion of human HMG-CoA reductase: insights into regulation of activity and catalysis. *EMBO Journal* **19**, 819–830.

Jadzinsky, P. D., Calero, G., Ackerson, C. J., Bushnell, D. A. & Kornberg, R. D. (2007). Structure of a thiol monolayer-protected gold nanoparticle at 1.1 Å resolution. *Science* **318**, 430–433.

James, R. W. (1948). *The Optical Principles of the Diffraction of X-rays*. London: Bell.

Jeruzalmi, D., O'Donnell, M. & Kuriyan, J. (2001). Crystal structure of the processivity clamp loader gamma (gamma) complex of E. coli DNA polymerase III. *Cell* **106**, 429–441.

Jin, M. S., Oldham, M. L., Zhang, Q. & Chen, J. (2012). Crystal structure of the multidrug transporter P-glycoprotein from *Caenorhabditis elegans*. *Nature* **490**, 566–569.

Kabsch, W. (2010). XDS. *Acta Crystallographica D* **66**, 125–132.

Kahn, R., Fourme, R., Bosshard, R., Chiadmi, M., Risler, J. L., Dideberg, O. & Wery, J. P. (1985). Crystal structure study of *Opsanus tau* Parvalbumin by multiwavelength anomalous diffraction. *FEBS Letters* **179**, 133–137.

Kallman, H. & Mark, H. (1927). Über die Dispersion und Streuung von Röntgenstrahlen. *Annalen der Physik* **82**, 585–604.

Karle, J. (1967). Anomalous scatterers in x-ray diffraction and the use of several wavelengths. *Applied Optics* **6**, 2132–2135.

Karle, J. (1980). Some developments in anomalous dispersion for the structure investigation of macromolecular systems in biology. *International Journal of Quantum Chemistry: Symposium* **7**, 357–367.

https://doi.org/10.1017/S0033583514000018 Published online by Cambridge University Press

KARLE, J. & HAUPTMAN, H. (1950). The phases and magnitudes of the structure factors. *Acta Crystallographica* **3**, 181–187.

KARLE, J. & HAUPTMAN, H. (1956). A theory of phase determination for the four types of non-centrosymmetric space groups 1*P*222, 2*P*22, 3$P_1$2, 3$P_2$2. *Acta Crystallographica* **9**, 635–651.

KAZANTSEV, A. V., KRIVENKO, A. A. & PACE, N. R. (2009). Mapping metal-binding sites in the catalytic domain of bacterial RNase P RNA. *RNA* **15**, 266–276.

KENDREW, J. C., BODO, G., DINTZIS, H. M., PARRISH, R. G., WYCKOFF, H. & PHILLIPS, D. C. (1958). A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. *Nature* **181**, 662–666.

KIM, Y., BABNIGG, G., JEDRZEJCZAK, R., ESCHENFELDT, W. H., LI, H., MALTSEVA, N., HATZOS-SKINTGES, C., GU, M., MAKOWSKA-GRZYSKA, M., WU, R., AN, H., CHHOR, G. & JOACHIMIAK, A. (2011). High-throughput protein purification and quality assessment for crystallization. *Methods* **55**, 12–28.

KINCAID, B. M., EISENBERGER, P., HODGSON, K. O. & DONIACH, S. (1975). X-ray absorption spectroscopy using synchrotron radiation for structural investigation of organometallic molecules of biological interest. *Proceedings of the National Academy of Sciences of the United States of America* **72**, 2340–2342.

KITAGAWA, D., VAKONAKIS, I., OLIERIC, N., HILBERT, M., KELLER, D., OLIERIC, V., BORTFELD, M., ERAT, M. C., FLÜCKIGER, I., GÖNCZY, P. & STEINMETZ, M. O. (2011). Structural basis of the 9-fold symmetry of centrioles. *Cell* **144**, 364–375.

KITAGO, Y., WATANABE, N. & TANAKA, I. (2005). Structure determination of a novel protein by sulfur SAD using chromium radiation in combination with a new crystal-mounting method. *Acta Crystallographica D* **61**, 1013–1021.

KNÄBLEIN, J., NEUEFEIND, T., SCHNEIDER, F., BERGNER, A., MESSERSCHMIDT, A., LÖWE, J., STEIPE, B. & HUBER, R. (1997). $Ta_6Br_{12}^{2+}$, a tool for phase determination of large biological assemblies by X-ray crystallography. *Journal of Molecular Biology* **270**, 1–7.

KOLATKAR, P. R., ERNST, S. R., HACKERT, M. L., OGATA, C. M., HENDRICKSON, W. A., MERRITT, E. A. & PHIZACKERLEY, R. P. (1992). The structure determination and refinement of homotetrameric hemoglobin from *Urechis caupo* at 2.5 Å resolution. *Acta Crystallographica B* **48**, 191–199.

KORZUN, Z. R. (1987). The tertiary structure of azurin from *Pseudomonas denitrificans* as determined by Cu resonant diffraction using synchrotron radiation. *Journal of Molecular Biology* **196**, 413–419.

KRAMERS, H. A. & HEISENBERG, W. (1925). Über die Streuung von Strahlung durch Atome. *Zeitschrift für Physik* **31**, 681–708.

DE KRONIG, R. L. (1926). On the theory of dispersion of x-rays. *Journal of the Optical Society of America* **12**, 547–557.

DE KRONIG, R. L. & KRAMERS, H. A. (1928). Zur Theorie der Absorption und Dispersion in den Röntgenspektren. *Zeitschrift für Physik* **48**, 174–179.

LANGER, G., COHEN, S. X., LAMZIN, V. S. & PERRAKIS, A. (2008). Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nature Protocols* **3**, 1171–1179.

LATTMAN, E. E., NOCKOLDS, C. E., KRETSINGER, R. H. & LOVE, W. E. (1971). Structure of yellow fin tuna metmyoglobin at 6 Å resolution. *Journal of Molecular Biology* **60**, 271–277.

LEAHY, D. J., ERICKSON, H. P., AUKHIL, I., JOSHI, P. & HENDRICKSON, W. A. (1994). Crystallization of a fragment of human fibronectin: introduction of methionine by site-directed mutagenesis to allow phasing via selenomethionine. *Proteins* **19**, 48–54.

LEAHY, D. J., HENDRICKSON, W. A., AUKHIL, I. & ERICKSON, H. P. (1992). Structure of a fibronectin type III domain from tenascin phased by MAD analysis of the selenomethionyl protein. *Science* **258**, 987–991.

LEE, C. & GOLDBERG, J. (2010). Structure of coatomer cage proteins and the relationship among COPI, COPII, and clathrin vesicle coats. *Cell* **142**, 123–132.

LEONARD, T. A., RÓŻYCKI, B., SAIDI, L. F., HUMMER, G. & HURLEY, J. H. (2011). Crystal structure and allosteric activation of protein kinase C βII. *Cell* **144**, 55–66.

LIDESTRI, J. P. & HENDRICKSON, W. A. (2009). Optimization of X-ray energy resolution from a horizontally focused single-crystal monochromator. *Nuclear Instruments and Methods A* **599**, 289–300.

LIU, Q., DAHMANE, T., ZHANG, Z., ASSUR, Z., BRASCH, J., SHAPIRO, L., MANCIA, F. & HENDRICKSON, W. A. (2012). Structures from anomalous diffraction data of native biological macromolecules. *Science* **336**, 1033–1037.

LIU, Q., LIU, Q. & HENDRICKSON, W. A. (2013). Robust structural analysis of native biological macromolecules from multi-crystal anomalous diffraction data. *Acta Crystallographica D* **69**, 1314–1332.

LIU, Q., ZHANG, Z. & HENDRICKSON, W. A. (2011). Multi-crystal anomalous diffraction for low resolution macromolecular phasing. *Acta Crystallographica D* **67**, 45–59.

LIU, Y., OGATA, C. M. & HENDRICKSON, W. A. (2001). Multiwavelength anomalous diffraction analysis at the M absorption edges of uranium. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 10648–10653.

LIU, Z.-J., VYSOTSKI, E. S., CHEN, C.-J., ROSE, J. P., LEE, J. & WANG, B.-C. (2000). Structure of the Ca2+-regulated photoprotein obelin at 1.7 A resolution determined directly from its sulfur substructure. *Protein Science* **9**, 2085–2093.

LYE, R. C., PHILLIPS, J. C., KAPLAN, D., DONIACH, S. & HODGSON, K. O. (1980). White lines in L-edge X-ray absorption spectra and their implications for anomalous

https://doi.org/10.1017/S0033583514000018 Published online by Cambridge University Press

diffraction studies of biological materials. *Proceedings of the National Academy of Sciences of the United States of America* **77**, 5884–5888.

Mancusso, R., Gregorio, G. G., Liu, Q. & Wang, D.-N. (2012). Structure and mechanism of a bacterial sodium-dependent dicarboxylate transporter. *Nature* **491**, 622–626.

Mark, H. & Szilard, L. (1925). Ein einfacher Versuch zu Auffindung eines selektiven Effektes bei der estreuung von Röntgenstrahlen. *Zeitschrift für Physik* **33**, 688–691.

Martin, J. L., Bardwell, J. C. & Kuriyan, J. (1993). Crystal structure of the DsbA protein required for disulphide bond formation *in vivo*. *Nature* **365**, 464–468.

Martinez-Hackert, E. & Hendrickson, W. A. (2009). Promiscuous substrate recognition in folding and assembly activities of the trigger factor chaperone. *Cell* **138**, 923–934.

Matthews, B. & Czerwinski, E. W. (1975). Local scaling: a method to reduce systematic errors in isomorphous replacement and anomalous scattering measurements. *Acta Crystallographica A* **31**, 480–487.

Mitchell, C. M. (1957). Phase determination by the two-wavelength method of Okaya and Pepinsky. *Acta Crystallographica* **10**, 475–476.

Moore, J. O. & Hendrickson, W. A. (2012). An asymmetry-to-symmetry switch in signal transmission by the histidine kinase receptor for TMAO. *Structure* **20**, 729–741.

Nagem, R. A., Ambrosio, A. L., Rojas, A. L., Navarro, M. V., Golubev, A. M., Garratt, R. C. & Polikarpov, I. (2005). Getting the most out of X-ray home sources. *Acta Crystallographica D* **61**, 1022–1030.

Nagem, R. A., Dauter, Z. & Polikarpov, I. (2001). Protein crystal structure solution by fast incorporation of negatively and positively charged anomalous scatterers. *Acta Crystallographica D* **57**, 996–1002.

Nakagawa, A., Higuchi, Y., Yasuoka, N., Katsube, Y. & Yagi, T. (1990). S-class cytochromes c have a variety of folding patterns: structure of cytochrome c-553 from *Desulfovibrio vulgaris* determined by the multi-wavelength anomalous dispersion method. *Journal of Biochemistry* **108**, 701–703.

Nishikawa, S. & Matukawa, K. (1928). Hemihedry of Zincblende and X-Ray Reflexion. *Proceedings of the Imperial Academy of Japan* **4**, 96–97.

Nyblom, M., Poulsen, H., Gourdon, P., Reinhard, L., Andersson, M., Lindahl, E., Fedosova, N. & Nissen, P. (2013). Crystal structure of Na+, K (+)-ATPase in the Na(+)-bound state. *Science* **342**, 123–127.

Ogata, C. M. & Hendrickson, W. A. (1995). Protein structures from MAD experiments at the howard hughes medical institute beam line X4A at NSLS. *Synchrotron Radiation News* **8**, 13–18.

Okaya, Y. & Pepinsky, R. (1956). New formulation and solution of the phase problem in x-ray analysis of non-centric crystals containing anomalous scatterers. *Physical Review* **103**, 1645–1647.

Okaya, Y., Saito, Y. & Pepinsky, R. (1955). New method in X-ray crystal structure determination involving the use of anomalous dispersion. *Physical Review* **98**, 1857–1858.

Ormö, M., Cubitt, A. B., Kallio, K., Gross, L. A., Tsien, R. Y. & Remington, S. J. (1996). Crystal structure of the Aequorea victoria green fluorescent protein. *Science* **273**, 1392–1395.

Otwinowski, Z. (1991). In *Isomorphous Replacement and Anomalous Scattering*, (eds. W. Wolf, P. R. Evans & A. G. W. Leslie), pp. 80–86. Warrington: SERC, Daresbury Laboratory.

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

Pähler, A., Smith, J. L. & Hendrickson, W. A. (1990). A probability representation for phase information from multiwavelength anomalous dispersion. *Acta Crystallographica A* **46**, 537–540.

Palczewski, K., Kumasaka, T., Hori, T., Behnke, C. A., Motoshima, H., Fox, B. A., Le Trong, I., Teller, D. C., Okada, T., Stenkamp, R. E., Yamamoto, M. & Miyano, M. (2000). Crystal structure of rhodopsin: a G protein-coupled receptor. *Science* **289**, 739–745.

Panjikar, S., Parthasarathy, V., Lamzin, V. S., Weiss, M. S. & Tucker, P. A. (2009). On the combination of molecular replacement and single-wavelength anomalous diffraction phasing for automated structure determination. *Acta Crystallograllographica D* **65**, 1089–1097.

Patterson, A. L. (1934). A fourier series method for the determination of the components of interatomic distances in crystals. *Physical Review* **46**, 372–376.

Patterson, A. L. (1935). A direct method for the determination of the components of interatomic distances in crystals. *Zeitschrift für Kristallographie* **90**, 517–542.

Payandeh, J., Scheuer, T., Zheng, N. & Catterall, W. A. (2011). The crystal structure of a voltage-gated sodium channel. *Nature* **475**, 353–358.

Perraskis, A., Morris, R. & Lamzin, V. S. (1999). Automated protein model building combined with iterative structure refinement. *Nature Structural Biology* **6**, 458–463.

Perutz, M. F., Rossmann, M. G., Cullis, A. F., Muirhead, H., Will, G. & North, A. C. (1960). Structure of haemoglobin: a three-dimensional Fourier synthesis at 5.5-Å. resolution, obtained by X-ray analysis. *Nature* **185**, 416–422.

Phillips, J. C. & Hodgson, K. O. (1980). The use of anomalous scattering effects to phase diffraction

https://doi.org/10.1017/S0033583514000018 Published online by Cambridge University Press

patterns from macromolecules. *Acta Crystallographica A* **36**, 856–864.

Phillips, J. C., Templeton, D. H., Templeton, L. K. & Hodgson, K. O. (1978). L_{III}-Edge anomalous X-ray scattering by cesium measured with synchrotron radiation. *Science* **201**, 257–259.

Pomeranz Krummel, D. A., Oubridge, C., Leung, A. K., Li, J. & Nagai, K. (2009). Crystal structure of human spliceosomal U1 snRNP at 5.5 Å resolution. *Nature* **458**, 475–480.

Prins, J. A. (1928). Über die Dispersion und Absorption von Röntgenstrahlen. *Zeitschrift für Physik* **48**, 479–498.

Purdy, M. D., Ge, P., Chen, J., Selvin, P. R. & Wiener, M. C. (2002). Thiol-reactive lanthanide chelates for phasing protein X-ray diffraction data. *Acta Crystallographica D* **58**, 1111–1117.

Qi, R., Sarbeng, E. B., Liu, Q., Le, K. Q., Xu, X., Xu, H., Yang, J., Wong, J. L., Vorvis, C., Hendrickson, W. A., Zhou, L. & Liu, Q. (2013). Allosteric opening of the polypeptide-binding site when an Hsp70 binds ATP. *Nature Structural and Molecular Biology* **20**, 900–907.

Ramachandran, G. N. & Raman, S. (1956). A new method for the structure analysis of non-centrosymmetric crystals. *Current Science* (India) **25**, 348–351.

Ramakrishnan, V. & Biou, V. (1997). Treatment of multi-wavelength anomalous diffraction data as a special case of multiple isomorphous replacement. *Methods in Enzymology* **276**, 538–557.

Ramakrishnan, V., Finch, J. T., Graziano, V., Lee, P. L. & Sweet, R. M. (1993). Crystal structure of globular domain of histone H5 and its implications for nucleosome binding. *Nature* **362**, 219–223.

Raman, C. S., Li, H., Martásek, P., Král, V., Masters, B. S. & Poulos, T. L. (1998). Crystal structure of constitutive endothelial nitric oxide synthase: a paradigm for pterin function involving a novel metal center. *Cell* **95**, 939–950.

Raman, S. (1959). Theory of the anomalous dispersion method of determining the structure and absolute configuration of non-centrosymmetric crystals. *Proceedings of the Indian Academy of Sciences A* **50**, 95–107.

Ramaseshan, S. (1993). Four decades in anomalous scattering – some reminiscences. *Current Science* **65**, 644–651.

Ramaseshan, S. & Venkatesan, K. (1957). The use of anomalous scattering without phase change in crystal structure analysis. *Current Science* **26**, 352–353.

Ramaseshan, S., Venkatesan, K. & Mani, N. V. (1957). The use of anomalous scattering for the determination of crystal structures – $KMnO_4$. *Proceedings of the Indian Academy of Sciences A* **46**, 95–111.

Read, R. J. & McCoy, A. J. (2011). Using SAD data in phaser. *Acta Crystallographica D* **67**, 338–344.

Reiter, N. J., Osterman, A., Torres-Larios, A., Swinger, K. K., Pan, T. & Mondragón, A. (2010). Structure of a bacterial ribonuclease P holoenzyme in complex with tRNA. *Nature* **468**, 784–789.

Rice, L. M., Earnest, T. N. & Brunger, A. T. (2000). Single-wavelength anomalous diffraction phasing revisited. *Acta Crystallographica D* **56**, 1413–1420.

Robbins, A. H., McRee, D. E., Williamson, M., Collett, S. A., Xuong, N. H., Furey, W. F., Wang, B. C. & Stout, C. D. (1991). Refined crystal structure of Cd, Zn metallothionein at 2.0 Å resolution. *Journal of Molecular Biology* **221**, 1269–1293.

Robertson, J. M. (1936). An x-ray study of the phthalocyanines. Part II. Quantitative structure determination of the metal-free compound. *Journal of the Chemical Society* 1195–1209.

Robertson, J. M. & Woodward, I. (1940). An x-ray study of the phthalocyanines. Part IV Direct quantitative analysis of the platinum compound. *Journal of the Chemical Society* 36–48.

Robinson, R. C., Turbedsky, K., Kaiser, D. A., Marchand, J. B., Higgs, H. N., Choe, S. & Pollard, T. D. (2001). Crystal structure of Arp2/3 complex. *Science* **294**, 1679–1684.

Rodgers, D. W. (1994). Cryocrystallography. *Structure* **15**, 1135–1140.

Rosenbaum, G., Holmes, K. C. & Witz, J. (1971). Synchrotron radiation as a source for X-ray diffraction. *Nature* **230**, 434–437.

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

Rossmann, M. G. (1972). *The Molecular Replacement Method*. New York: Gordon & Breach.

Rossmann, M. G. & Blow, D. M. (1962). The detection of sub-units within the crystallographic asymmetric unit. *Acta Crystallographica* **15**, 24–31.

Rudenko, G., Henry, L., Henderson, K., Ichtchenko, K., Brown, M. S., Goldstein, J. L. & Deisenhofer, J. (2002). Structure of the LDL receptor extracellular domain at endosomal pH. *Science* **298**, 2353–2358.

Rudenko, G., Henry, L., Vonrhein, C., Bricogne, G. & Deisenhofer, J. (2003). 'MAD'ly phasing the extracellular domain of the LDL receptor: a medium-sized protein, large tungsten clusters and multiple non-isomorphous crystals. *Acta Crystallographica* **59**, 1978–1986.

Ryu, S.-E., Kwong, P. D., Truneh, A., Porter, T. G., Arthos, J., Rosenberg, M., Dai, X., Xuong, Ng.-H., Axel, R., Sweet, R. W. & Hendrickson, W. A. (1990). Crystal structure of an HIV-binding recombinant fragment of human CD4. *Nature* **348**, 419–426.

Sawaya, M. R., Wojtowicz, W. M., Andre, I., Qian, B., Wu, W., Baker, D., Eisenberg, D. & Zipursky, S. L. (2008). A double S shape provides the structural basis for the extraordinary binding specificity of Dscam isoforms. *Cell* **134**, 1007–1018.

https://doi.org/10.1017/S0033583514000018 Published online by Cambridge University Press

Schuermann, J. P. & Tanner, J. J. (2003). MRSAD: using anomalous dispersion from S atoms collected at Cu K alpha wavelength in molecular-replacement structure determination. *Acta Crystallographica D* **59**, 1731–1736.

Schiltz, M. & Bricogne, G. (2008). Exploiting the anisotropy of anomalous scattering boosts the phasing power of SAD and MAD experiments. *Acta Crystallographica D* **64**, 711–729.

Schiltz, M., Fourme, R. & Prangé, T. (2003). Use of noble gases xenon and krypton as heavy atoms in protein structure determination. *Methods in Enzymology* **374**, 83–119.

Schneider, T. R. & Sheldrick, G. M. (2002). Substructure solution with SHELXD. *Acta Crystallographica D* **58**, 1772–1779.

Schmidt, A., Teeter, M., Weckert, E. & Lamzin, V. S. (2011). Crystal structure of small protein crambin at 0.48 Å resolution. *Acta Crystallographica F* **67**, 424–428.

Shapiro, L., Fannon, A. M., Kwong, P. D., Thompson, A., Lehmann, M. S., Grübel, G., Legrand, J. F., Als-Nielsen, J., Colman, D. R. & Hendrickson, W. A. (1995). Structural basis of cell–cell adhesion by cadherins. *Nature* **374**, 327–337.

Shapiro, L. & Lima, C. D. (1998). The Argonne Structural Genomics Workshop: Lamaze class for the birth of a new science. *Structure* **15**, 265–267.

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

Sheng, J. & Huang, Z. (2010). Selenium derivatization of nucleic acids for x-ray crystal structure and function studies. *Chemistry and Biodiversity* **7**, 753–785.

Singh, A. K. & Ramaseshan, S. (1968). The use of neutron scattering in crystal structure analysis. I. Noncentrosymmetric structures. *Acta Crystallographica B* **24**, 35–39.

Smith, G. D., Nagar, B., Rini, J. M., Hauptman, H. A. & Blessing, R. H. (1998). The use of SnB to determine an anomalous scattering substructure. *Acta Crystallographica D* **54**, 799–804.

Smith, J. L. & Hendrickson, W. A. (2001). Multiwavelength anomalous diffraction. In *International Tables for Crystallography*, vol. F (ed. M. G. Rossmann and E. Arnold), pp. 299–303. Chester: IUCr.

Smith, J. L., Hendrickson, W. A. & Addison, A. W. (1983). Structure of trimeric haemerythrin. *Nature* **303**, 86–88.

Sheriff, S., Hendrickson, W. A. & Smith, J. L. (1987). Structure of myohemerythrin in the azidomet state at 1.7/1.3 Å Resolution. *Journal of Molecular Biology* **197**, 273–296.

Son, S. K., Chapman, H. N. & Santra, R. (2011). Multiwavelength anomalous diffraction at high x-ray intensity. *Physics Review Letter* **107**, 218102.

Staudenmann, J.-L., Hendrickson, W. A. & Abramowitz, R. (1989). The synchrotron resource of the Howard Hughes Medical Institute. *Review of Scientific Instruments* **60**, 1939–1942.

Strub, M. P., Hoh, F., Sanchez, J. F., Strub, J. M., Böck, A., Aumelas, A. & Dumas, C. (2003). Selenomethionine and selenocysteine double labeling strategy for crystallographic phasing. *Structure* **11**, 1359–1367.

Stuhrmann, S., Bartels, K. S., Braunwarth, W., Doose, R., Dauvergne, F., Gabriel, A., Knöchel, A., Marmotti, M., Stuhrmann, H. B., Trame, C. & Lehmann, M. S. (1997). Anomalous dispersion with edges in the soft X-ray region: first results of diffraction from single crystals of trypsin near the K-absorption edge of sulfur. *Journal of Synchrotron Radiatation* **4**, 298–310.

Sutton, R. B., Fasshauer, D., Jahn, R. & Brunger, A. T. (1998). Crystal structure of a SNARE complex involved in synaptic exocytosis at 2.4 Å resolution. *Nature* **395**, 347–353.

Templeton, D. H. & Templeton, L. K. (1982). X-ray dichroism and polarized anomalous scattering of the uranyl ion. *Acta Crystallographica A* **38**, 62–67.

Templeton, L. K. & Templeton, D. H. (1988). Biaxial tensors for anomalous scattering of X-rays in selenolanthionine. *Acta Crystallographica A* **44**, 478–481.

Templeton, D. H., Templeton, L. K., Phillips, J. C. & Hodgson, K. O. (1980). Anomalous scattering of X-rays by cesium and cobalt measured with synchrotron radiation. *Acta Crystallographica A* **36**, 436–442.

Templeton, D. H., Templeton, L. K., Phizackerley, R. P. & Hodgson, K. O. (1982). L_3-edge anomalous scattering by gadolinium and samarium measured at high resolution with synchrotron radiation. *Acta Crystallographica A* **38**, 74–78.

Terwilliger, T. C. (1994). MAD phasing: Bayesian estimates of F_A. *Acta Crystallographica D* **50**, 11–16.

Terwilliger, T. C. (2000). Maximum-likelihood density modification. *Acta Crystallographica D* **56**, 965–972.

Terwilliger, T. C. (2003a). Automated main-chain model building by template matching and iterative fragment extension. *Acta Crystallographica D* **59**, 38–44.

Terwilliger, T. C. (2003b). Automated side-chain model building and sequence assignment by template matching. *Acta Crystallographica D* 45–49.

Terwilliger, T. C. & Berendzen, J. (1999). Automated MAD and MIR structure solution. *Acta Crystallographica D* **55**, 849–861.

Thomson, J. J. (1906). *Conduction of Electricity through Gases*. Cambridge: Cambridge University Press.

Thygesen, J., Weinstein, S., Franceschi, F. & Yonath, A. (1996). The suitability of multi-metal clusters for phasing in crystallography of large macromolecular assemblies. *Structure* **4**, 513–518.

https://doi.org/10.1017/S0033583514000018 Published online by Cambridge University Press

Tollin, P. (1969). Determination of the orientation and position of the myoglobin molecule in the crystal of seal myoglobin. *Journal of Molecular Biology* **45**, 481–490.

Turner, M. A., Yuan, C. S., Borchardt, R. T., Hershfield, M. S., Smith, G. D. & Howell, P. L. (1998). Structure determination of selenomethionyl S-adenosylhomocysteine hydrolase using data at a single wavelength. *Nature Structural Biology* **5**, 369–376.

van den Bedem, H., Wolf, G., Xu, Q. & Deacon, A. M. (2011). Distributed structure determination at the JCSG. *Acta Crystallographica D* **67**, 368–375.

Walsh, M. A., Evans, G., Sanishvili, R., Dementieva, I. & Joachimiak, A. (1999). MAD data collection – current trends. *Acta Crystallographica D* **55**, 1726–1732.

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

Ward, K. B., Hendrickson, W. A. & Klippenstein, G. L. (1975). Quaternary and tertiary structure of hemerythrin. *Nature* **257**, 818–821.

Warkentin, M. & Thorne, R. E. (2010). Glass transition in thaumatin crystals revealed through temperature-dependent radiation-sensitivity measurements. *Acta Crystallographica D* **66**, 1092–1100.

Watson, J. D. (1972). *Cold Spring Harbor Symposia on Quantitative Biology*, vol. XXXVI. Structure and Function of Proteins at the Three-Dimensional Level. Cold Spring Harbor: Cold Spring Harbor Laboratory.

Weis, W. I., Kahn, R., Fourme, R., Drickamer, K. & Hendrickson, W. A. (1991). Structure of the calcium-dependent lectin domain from a rat mannose-binding protein determined by MAD phasing. *Science* **254**, 1608–1615.

Weiss, M. S., Sicker, T., Djinovic-Carugo, K. & Hilgenfeld, R. (2001). On the routine use of soft X-rays in macromolecular crystallography. *Acta Crystallographica D* **57**, 689–695.

White, T. A., Barty, A., Stellato, F., Holton, J. M., Kirian, R. A., Zatsepin, N. A. & Chapman, H. N. (2013). Crystallographic data processing for free-electron laser sources. *Acta Crystallographica D* **69**, 1231–1240.

Wimberly, B. T., Brodersen, D. E., Clemons, W. M. Jr., Morgan-Warren, R. J., Carter, A. P., Vonrhein, C., Hartsch, T. & Ramakrishnan, V. (2000). Structure of the 30S ribosomal subunit. *Nature* **407**, 327–339.

Wilson, I. A., Skehel, J. J. & Wiley, D. C. (1981). Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. *Nature* **289**, 366–373.

Wu, H., Kwong, P. D. & Hendrickson, W. A. (1997). Dimeric association and segmental variability in the structure of human CD4. *Nature* **387**, 527–530.

Wu, H., Lustbader, J. W., Liu, Y., Canfield, R. E. & Hendrickson, W. A. (1994). Structure of human chorionic gonadotropin at 2.6 Å resolution from MAD analysis of the selenomethionyl protein. *Structure* **2**, 545–558.

Wyckoff, R. W. G. (1930). The X-ray scattering powers of nickel and oxygen in nickel oxide. *Physical Review* **35**, 583–587.

Xiao, R., Anderson, S., Aramini, J., Belote, R., Buchwald, W. A., Ciccosanti, C., Conover, K., Everett, J. K., Hamilton, K., Huang, Y. J., Janjua, H., Jiang, M., Kornhaber, G. J., Lee, D. Y., Locke, J. Y., Ma, L. C., Maglaqui, M., Mao, L., Mitra, S., Patel, D., Rossi, P., Sahdev, S., Sharma, S., Shastry, R., Swapna, G. V., Tong, S. N., Wang, D., Wang, H., Zhao, L., Montelione, G. T. & Acton, T. B. (2010). The high-throughput protein sample production platform of the Northeast Structural Genomics Consortium. *Journal of Structural Biology* **172**, 21–33.

Yamashita, A., Singh, S. K., Kawate, T., Jin, Y. & Gouaux, E. (2005). Crystal structure of a bacterial homologue of Na+/Cl– dependent neurotransmitter transporters. *Nature* **437**, 215–223.

Yang, C. & Pflugrath, J. W. (2001). Applications of anomalous scattering from S atoms for improved phasing of protein diffraction data collected at Cu Kα wavelength. *Acta Crystallographica D* **57**, 1480–1490.

Yang, C., Pflugrath, J. W., Courville, D. A., Stence, C. N. & Ferrara, J. D. (2003). Away from the edge: SAD phasing from the sulfur anomalous signal measured in-house with chromium radiation. *Acta Crystallographica D* **59**, 1943–1957.

Yang, W., Hendrickson, W. A., Crouch, R. J. & Satow, Y. (1990). Structure of ribonuclease H phased at 2 Å resolution by MAD analysis of the selenomethionyl protein. *Science* **249**, 1398–1405.

Zhang, K. Y. J. & Main, P. (1990). Histogram matching as a new density modification technique for phase refinement and extension of protein molecules. *Acta Crystallographica A* **46**, 41–46.

Zhou, Y. & MacKinnon, R. (2003). The occupancy of ions in the K+ selectivity filter: charge balance and coupling of ion binding to a protein conformational change underlie high conduction rates. *Journal of Molecular Biology* **333**, 965–975.

Zhu, X., Zhao, X., Burkholder, W. F., Gragerov, A., Ogata, C. M., Gottesman, M. E. & Hendrickson, W. A. (1996). Structural analysis of substrate binding by the molecular chaperone DnaK. *Science* **272**, 1606–1614.

https://doi.org/10.1017/S0033583514000018 Published online by Cambridge University Press