Screening for phasing atoms in protein crystallography

Titus J Boggon and Lawrence Shapiro*

Address: Structural Biology Program, Department of Physiology and Biophysics, Mount Sinai School of Medicine of New York University, 1425 Madison Avenue, New York, NY 10029, USA.

*Corresponding author.
E-mail: shapiro@anguilla.physbio.mssm.edu

Structure 2000, **8**:R143–R149

0969-2126/00/$ – see front matter

Introduction

Determining crystallographic phases remains a key step in the solution of macromolecular structures. For most novel structures, phases must be determined experimentally using methods that rely on finding the positions of a few special (either electron-dense 'heavy', or anomalously scattering) atoms. The scattering contribution of these atoms, in combination with knowledge of their positions, can be used to calculate phases for the total crystal contents. Whether these calculations rely on multiple wavelength anomalous dispersion (MAD), multiple isomorphous replacement (MIR), single isomorphous replacement with anomalous scattering (SIRAS) or other methods will depend entirely on the ordered presence of these special atoms. Thus, the ability to efficiently prepare samples containing such sites is of fundamental importance in macromolecular crystallography.

Incorporation of selenium atoms in proteins via selenomethionine (SeMet) has become a staple method for protein crystallography [1]. However, in some cases, SeMet MAD is not an appropriate method for phase determination [2]. Thus screening for phasing atoms is still an indispensable arrow in the crystallographer's quiver. Although the screening process has often been approached as a somewhat arbitrary collection of trial-and-error experiments, recent advances have provided the basis for a more systematic experimental approach. Here we review methods of screening for phasing atoms, with an emphasis on the newer more systematic approaches that have streamlined the process. In particular, we describe a novel method of screening using native gel electrophoresis that enables dozens of compounds to be analyzed in only a few hours.

A little forethought goes a long way

There is no *a priori* way to predict with certainty compounds and conditions that will lead to the successful derivatization of a protein crystal. For example, a protein that contains free cysteine thiols is likely to be derivatized by mercury compounds but will not always be. Derivatization depends on the exposure of functional groups, local chemical environment, ionization state, and other parameters that can not be accurately foreseen, especially in the absence of a high-resolution structure.

Nonetheless, screening for derivatives can be directed toward a higher likelihood of success by simple conjectures based on known properties of the protein and the particulars of the crystallization conditions. An additional consideration concerns the X-ray diffraction properties of the phasing atom to be incorporated. Different types of atom are best suited to different types of diffraction experiments, and derivative screening can be designed from the inception to focus on the optimum scatterer for the chosen experiment.

The classic text describing the chemistry of derivatization compounds remains Blundell and Johnson [3]. More recent reviews (for example [4,5]) have expanded on this early work. Blundell and Johnson designated compounds as 'hard' or 'soft'. Hard ligands bind to their protein targets without covalent modification. For example, the lanthanides typically bind ionically via oxygen-containing sidechains and carbonyls. Soft ligands — which often include platinum, mercury and gold atoms — tend to bind covalently to sulfhydryl, imidazole and thiol groups. Some of these compounds have high specificity for particular sidechains, for example, K_2PtCl_4 preferentially reacts with methionine and many mercury compounds specifically target cysteine thiols. A small subset of the available derivatization reagents are responsible for a substantial portion of the structures determined using derivatized crystals. Statistics on derivatization can be evaluated using the Heavy Atom Databank (HAD) [6] (http://www.bmm.icnet.uk/had/heavyatom.html). Historically, the seven most successful compounds — occasionally referred to as the 'magic seven' — have been K_2PtCl_4, $KAu(CN)_2$, K_2HgI_4, $UO_2(C_2H_3O_2)$, $HgCl_2$, para-chloro mercury benzoic acid sulfonate (PCMBS) and $K_3UO_2F_5$. Whether these compounds are in fact better by nature, or are simply statistically advantaged because of their early arrival in protein crystallography, is not entirely clear.

The natural ligand-binding activities of some proteins can be used for the incorporation of phasing atoms. Natural metal centers containing iron, copper and zinc, (K edges at 1.74, 1.38 and 1.28 Å, respectively) have all been used for the determination of protein structures using MAD. Ca^{2+} and Mg^{2+} binding are also common but their X-ray properties exclude their use as phasing atoms (K edges at 3.07 and 9.50 Å, respectively). Divalent cations, however, can often be replaced with trivalent lanthanides (^{57}La through to ^{71}Lu). The lanthanides have among the best X-ray anomalous scattering properties of all atoms, often with f'' peaks at the L_{III} absorption edge as high as 30 electrons

[7]. Lanthanide ions, however, form insoluble colloids above neutral pH, and this excludes their use in some cases. Bound phosphate ions can often be replaced by tungstate [8,9], which has excellent MAD diffraction properties. More complex biological ligands have also been used for the introduction of phasing atoms. For example, the natural iodine of thyroid hormone was used to determine its receptor complex structure [10]. Phasing atoms can also be incorporated into small-molecule ligands by chemical or biological synthesis. This method was used in the first selenium MAD experiment; seleno–biotin was used as a phasing vehicle for the complex structure with streptavidin [11]. Similarly, it has now become common practice to use iodinated (for MIR) or brominated (for MAD) nucleotides in the determination of structures of nucleic acids and their complexes with proteins.

Molecular biology methods have also been useful for introducing phasing atoms. In addition to the standard method of selenomethionine substitution for natively encoded methionines [1], it is also possible to insert additional methionines by mutagenesis in order to enhance the phasing signal [12]. Some studies, relying on the avidity of mercury atoms for free thiols, have used ectopic cysteine residues — engineered for predicted surface accessibility — as covalent acceptors for phasing atoms [13,14].

Choice of atom type

Other than chemical reactivity properties, the best choice of atom type will depend on diffraction properties. For both anomalous scattering and isomorphous difference experiments, the primary criterion will be whether the change in total scattering brought about by derivatization is large enough to be observed well experimentally. For isomorphous differences, this is simply related to the number of phasing atoms incorporated and their atomic number. In 1956 Crick and Magdoff [15] showed that the average expected intensity change from isomorphously adding heavy atoms can be estimated shown in equation 1.

$$\frac{\langle \Delta F \rangle}{\langle |F| \rangle} = \sqrt{\frac{2N_H}{N_P}} \cdot \frac{Z_H}{Z_{eff}} \qquad (1)$$

N_H and N_P refer to the number of heavy and protein atoms respectively, Z_H is the atomic number for the heavy atom, and Z_{eff} is the effective average atomic number for a protein atom, about 6.7. This expression can be evaluated prior to derivative screening to limit the screen to atom choices likely to provide sufficient signal. But what is sufficient signal? An effect about the value of R_{sym} will be observable but larger signals are desirable. Because R_{sym} increases with resolution, the signal estimate is often best made for the highest resolution shell for which phasing is likely, given the quality of the data. For example, a single ordered 100% occupancy mercury atom (Z = 80) bound to an 80 kDa protein should provide about 21% isomorphous differences, whereas an yttrium atom (Z = 39) would provide only about 10% differences. Although an yttrium derivative could certainly be of value, it is probable that data quality would limit its usefulness to lower resolution shells. It is often more prudent to initially steer clear of such potential derivative atoms.

For anomalous scattering, the situation is similar [16]. The total scattering factor f for an anomalously scattering atom can be expressed as the sum of a wavelength-independent part f^0 and the real and imaginary components of the anomalous scattering (equation 2).

$$f = f^0 + f'(\lambda) + i f''(\lambda) \qquad (2)$$

f' reflects changes in the 'real' scattering component (as do isomorphous differences), whereas the f'' value at a particular wavelength determines the strength of 'imaginary' anomalous scattering. The root mean square expected signal from Bijvoet differences at a given wavelength λ can be expressed as in equation 3.

$$\frac{\langle \Delta F_{\pm h} \rangle}{\langle |F| \rangle} = \sqrt{\frac{2N_A}{N_T}} \cdot \frac{f''_\Delta(\lambda)}{Z_{eff}} \qquad (3)$$

N_A is the number of anomalous scatterers. The dispersive diffraction ratio between two wavelengths can be expressed as equation 4.

$$\frac{\langle \Delta F_{\Delta\lambda} \rangle}{\langle |F| \rangle} = \sqrt{\frac{N_A}{2N_T}} \cdot \frac{\left| f'_A(\lambda_i) - f'_A(\lambda_j) \right|}{Z_{eff}} \qquad (4)$$

N_T is the total number of atoms, including anomalous scatterers. Anomalous differences, in general, are smaller than isomorphous differences, and anomalous scattering behavior varies broadly among the elements. Hence great care must be taken in the choice of element for anomalous diffraction experiments.

Because anomalous data is most often measured from a single frozen crystal, non-isomorphism is usually near zero. Furthermore, special geometries, such as inverse-beam and mirror-alignment can minimize errors in collecting anomalous data so that errors on anomalous difference measurements are substantially below R_{sym}. It is for these reasons that a substantially lower signal can be tolerated in MAD experiments as compared with isomorphous experiments. For example, a rough rule of thumb has arisen that ~1 selenomethionine per 100 amino acids is sufficient for a SeMet MAD experiment [16]. Using this equation, a signal of only ~4% (for Bijvoet differences at the $f'' = 6e^-$ Se peak)

is predicted. Although this would certainly be inadequate for most isomorphous experiments, it is well within the realm of possibility for MAD data.

For MAD experiments the best derivative elements will have large f'' values at an absorption edge and large differences between f' at different nearby wavelengths. These conditions are met by elements with strong and sharp (white line) X-ray absorption features. It is also necessary that these edges fall within the appropriate energy range, from ~0.5 Å (near the high-end of efficient X-ray production for most synchrotron sources) to ~1.7 Å (above which absorption effects become problematic). Among the best anomalous scatterers are the actinides and lanthanides. Of the actinides, all of which have predominant isotopes that are radioactive, only ^{92}U is commonly used in crystallography, and can provide extraordinary L_{III} edge anomalous signal at 0.7Å [7]. The lanthanides, ^{57}La to ^{71}Lu, have strong white line absorption features at their L_{III} edges, which increase in energy monotonically from 2.6 Å for ^{57}La to 1.3Å for ^{71}Lu. Lanthanide Ln^{3+} ions typically bind ionically to proteins. They can often substitute for divalent cations at specific protein-binding sites. Ln^{3+} ions across the lanthanide series often exhibit remarkably similar chemical behavior [17]. Thus, if one type of lanthanide ion can substitute at a particular binding site, other types will usually work as well. In this way, experiments can be tuned to achieve optimal X-ray characteristics by employing a particular lanthanide. Samarium, for ex ample, provides the strongest Bijvoet differences (f''~12e^-) with CuKα radiation; for synchrotron experiments lutetium provides the highest energy L_{III} transition, thus providing an easy means to minimize absorption difficulties.

Covalent heavy-atom derivatives of proteins can often be obtained by reacting with compounds containing elements ^{76}Os through ^{82}Pb. Although it is possible to perform MAD experiments that utilize L edges of each of these elements, their anomalous properties differ substantially. Elements ^{76}Os through ^{79}Au have substantial L_{III} white lines, but ^{80}Hg has a relatively flat spectrum, somewhat diminishing the value of dispersive differences in MAD experiments. Nonetheless, it should be noted that ^{80}Hg is the second most-used element for MAD, trailing only selenium. This underscores the point that experiments utilizing L edge transitions, in general, are preferable to K-edge experiments because of the larger absolute signal at the L edge. Although the Se K edge provides a well-defined spectrum for choosing wavelengths to maximize dispersive differences, the peak f'' value for Se is on the order of only 6 e^-. As a general rule, L_{III} edges have superior transition and white line intensities than the corresponding L_I and L_{II} edges. However, the L_I and L_{II} edges fall at higher energies. Thus an experiment out of range because of air absorption problems at L_{III} might in some cases be feasible for the same crystal using the L_I or L_{II} edge [16].

The inglorious history of heavy-atom screening

Crystallography has sometimes been likened (by non-crystallographers) to a modern day alchemy, in which strange and devious incantations are uttered over protein solutions to induce crystallization, and obscure witchcraft is employed to coax derivatization reactions. This reputation was certainly enhanced by the mostly anecdotal method of early derivative screening. In general, properties of crystals incubated with a potential derivatization compound were observed, and a particular behavior would suggest whether to proceed to the ultimate step of X-ray analysis. The problem is that the properties observed often correlated only poorly with the likelihood of derivatization.

For example, many phasing compounds are colored, and crystal color change has therefore been used as an indicator of derivatization. However, it has been found that many heavy atom compounds bind non-specifically to the surface of proteins in the crystal (or to hydrophobic areas on denatured protein that often forms the 'skin' that covers many crystals). This non-specific binding does not produce ordered sites, and is of no value for crystallographic purposes. These problems severely limit the usefulness of color change to evaluate derivatization.

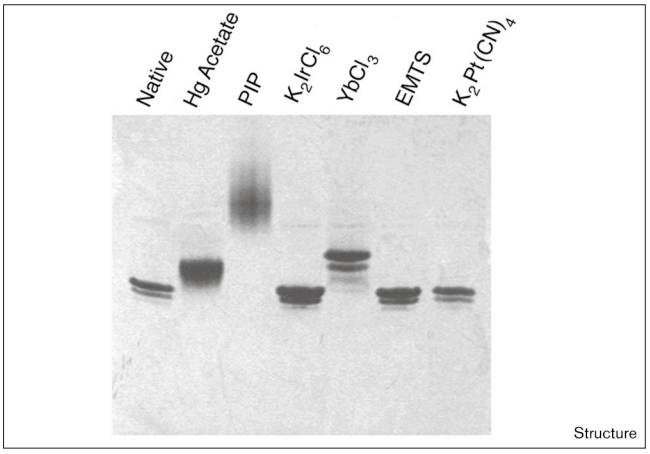
Another method relied on estimating the density of a crystal by the subjective speed with which it falls in a liquid-filled capillary. Heavy phasing atoms, it was reasoned, would increase the protein crystal density, thus causing faster sinking [18] —— an experiment Gallileo might have been proud of, but hardly rigorous or amenable to high-throughput application.

One of the most striking 'negative' results of a derivatization experiment is crystal cracking or even complete dissolution. Crystal cracking is often indicative of derivatization at a lattice contact, denaturation of the crystallized protein, or induction of a conformational change that cannot be accommodated within the crystal lattice. Although this can be an unhappy result, it can frequently provide a starting point for improvement. Lower concentrations of derivatization compound can be used, or less-reactive compounds from the same series. For example $K_2Pt(CN)_4$ is generally less reactive than its counterpart K_2PtBr_4. The character of leaving groups in common heavy-atom compounds can often dictate activity levels [3], with activity descending in the following order: F > Cl > Br > I > CN. Covalent cross-linking before soaking by gentle vapor diffusion of glutaraldehyde [19] will often prevent crystal cracking, but does not always result in a derivatized crystal with the desired diffraction properties.

Scanning for derivatives by analysis of X-ray data

The most definitive way to assess the formation of useful derivatives is through X-ray analysis. Classically, a precession photograph of a particular zone, *hk0* for example, would be compared for each potential derivative crystal.

Figure 1



Native PAGE analysis of derivatization reactions. 2 µl of the protein Acrp30 (2 mg/ml) was mixed with 2 µl of derivatization solution (10 mM phasing atom compound in 10 mM bis-tris, pH 6.0) and left on ice for 10 min. Next, 1 µl of the reaction solution was loaded onto a Phastgel Homogenous 20 (Amersham Pharmacia Biotech) and run to 250Avh using native buffer strips.

These photographs could then be evaluated for lattice change and intensity changes — usually gauged visually for 'intensity reversals' of adjacent spots. This was a time consuming and often inexact endeavor.

Modern detectors have improved this process, enabling more rigorous comparison of data acquired from native and potential derivative crystals. This does not necessarily entail the collection of entire data sets; often the process can be streamlined by collecting only small 'swaths' of data, and analyzing these statistically. Description of this analysis is beyond the scope of this review, however it is worthy of note here that the main complication in these analyses comes from non-isomorphism in analysis between native and potential derivative crystals. Non-isomorphism can easily be mistaken for isomorphous differences. To obviate this problem, it is often best to analyze anomalous signals by collecting 5–10° of data and their corresponding Bijvoet mates using inverse-beam geometry.

Systematic experiments prior to X-ray exposure

Although the most definitive assay for phasing atom derivatization is the examination of X-ray data, this is a time-consuming process. Two simple chemical methods can be used to pre-screen dozens of potential heavy-atom derivatives in only a few hours, with a relatively high degree of accuracy and using no crystals. Firstly, native polyacrylamide gel electrophoresis (PAGE) of protein and heavy-atom mixtures can be used to search for gel shifts upon derivatization. This technique can show which reagents cause protein denaturation and are therefore less likely to be useful as heavy-atom derivatives. This method is valuable for both covalent and non-covalent derivatives. Secondly, mass spectrometry can be used to assay the formation of covalent complexes between protein and heavy atom reagent. These experiments can be performed in a highly parallel and efficient way.

Native gel electrophoresis experiments

Native PAGE separates proteins on the basis of net charge, size and conformation [20]. To first approximation, mobility in a native gel will be proportional to the quotient (net charge)/(frictional coefficient), where the frictional coefficient is closely correlated to molecular size. Single-charge differences are often clearly resolved; for example the phosphorylation of a protein often results in increased mobility because of the addition of a negatively charged phosphate group (e.g., [21]). The frictional coefficient resulting from addition of a small molecule is usually imperceptibly small, such that gel shifts originate primarily from charge differences. 'Laddering' patterns often indicate the presence of multiple discretely charged states.

Molecular changes caused by phasing atom derivatization can often be clearly observed on native gels. Native PAGE is also an easy method for determining whether or not a protein is properly folded, as a denatured protein will not usually enter the gel. Interactions with heavy-atom compounds commonly cause protein denaturation, and this often correlates with a loss of diffraction in similarly treated crystals. The concentration of derivatization reagent can be important to the behavior of these reactions. In the event that denaturation is observed, lower concentrations can be screened by gel analysis. Conversely, higher concentrations can be tried where no effect or incomplete derivatization is observed. Native PAGE is rapid and easy to perform (~1 hour). Literally dozens of heavy-atom conditions can be screened in a short time by application of this technique. Although proof of useful derivatization (the presence of ordered heavy-atom sites) still relies on diffraction analysis, this initial pre-screening process can provide a great increase in the efficiency of finding derivatives.

An example application of native gel screening is shown in Figure 1. Derivatization of the trimeric protein Acrp30 [22] was analyzed using native PHAST gels (Amersham Pharmacia Biotech). The data shown here correspond to about one third of the compounds screened in an afternoon. Clear band shifts are observed for Hg acetate, PIP, and $YbCl_3$. Note that the band shifts observed in these cases result from a retardation of mobility, as would be expected from the complexation of positively charged ions (e.g. Yb^{3+}) with the protein. No band shift is observed for $K_2Pt(CN)_4$, EMTS or K_2IrCl_6. X-ray analysis shows that all of the compounds eliciting band shifts did indeed form specific derivatives (see difference Fourier maps, Figure 2). The compounds that did not cause a band-shift did not derivatize the protein, with the exception of K_2IrCl_6. Although

the iridium compound bound specifically, it probably did not induce a change in net charge, thus accounting for the lack of a gel shift. Data not shown include experiments using gold compounds; we found that most of these caused aggregation of the protein, such that it could not enter the native gel. When Acrp30 crystals were treated with these compounds, their diffraction properties were typically destroyed. This is a correlated pattern that we have observed on many occasions: a compound that denatures a protein in solution often destroys the diffraction properties of its crystals. We commonly use this observation in selection of derivatization compounds. In our experience, native gel shift experiments provide one of the fastest and most reliable experiments for phasing atom screening.

Mass spectrometry

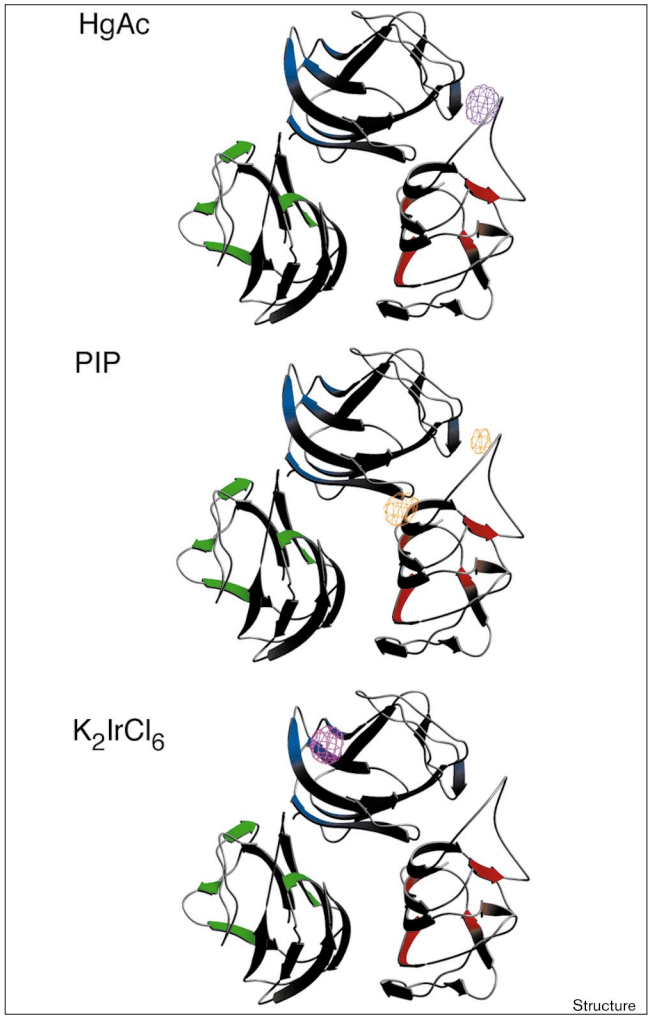
Mass spectrometry of proteins [23–25] is now commonly used to determine the degree of protein modification by heavy-atom compounds [26,27]. Although this technique is extraordinarily powerful for analyzing the stoichiometry of covalent complexes, it is not without problems. In particular, many heavy-metal compounds appear to interfere with the desorption of the sample from the matrix associated laser desorption ionization (MALDI) applicator chip. This can severely impact the quality of the spectra. The MALDI mass spectra in Figure 3 show a comparison of native and mercury acetate-reacted forms of Acrp30. Although the presence of the heavy atom degrades the signal, a mass shift of ~213 Da is clearly seen. This corresponds approximately to the 200 Da mass of a single mercury atom.

New kinds of derivatives

Since the time of Blundell and Johnson's description [3], many new derivatization reagents have been explored. Some of these are essentially classical heavy-atom compounds with improved properties. For example, trimethyllead acetate has far better solubility and specificity properties than its lead-based predecessors, and has become a prominent choice for protein derivatization [28]. Derivatization under pressure by the noble gas xenon — although first described in 1965 [29] — has recently enjoyed a resurgence, primarily because of better pressurization equipment [30,31]. Techniques for the derivatization of nucleic acids have also advanced substantially [32].

In cases where the contents of the asymmetric unit is very large, as for many multi-protein complexes, several phasing atoms might need to be incorporated to achieve sufficient anomalous or isomorphous signal. This need has led to the development of metal cluster compounds for use in phasing large structures [33]. The problem of non-isomorphism between crystals is especially common for crystals of large molecules. To alleviate this difficulty, clusters of metals with substantial anomalous white line features — particularly Tungsten clusters — have been developed for use in MAD phasing [33].

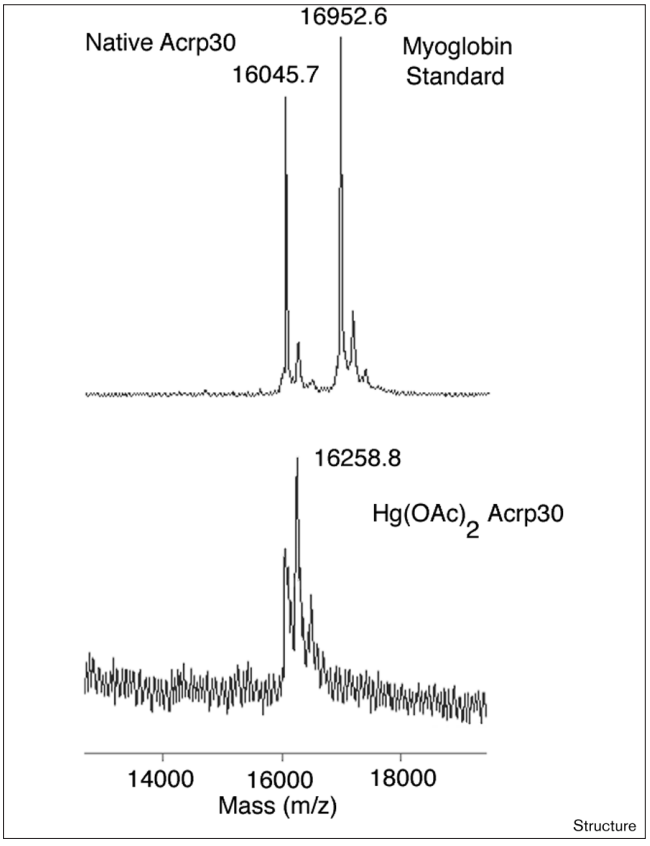
Figure 2



Bijvoet difference electron density maps (5σ) of **(a)** $Hg(C_2H_3O_2)_2$, **(b)** PIP, and **(c)** K_2IrCl_6 derivatized Acrp30 crystals superimposed on a ribbon diagram of the refined protein crystal structure (PDB entry 1C28). $YbCl_3$ also formed an excellent derivative, although electron density is not shown for this compound.

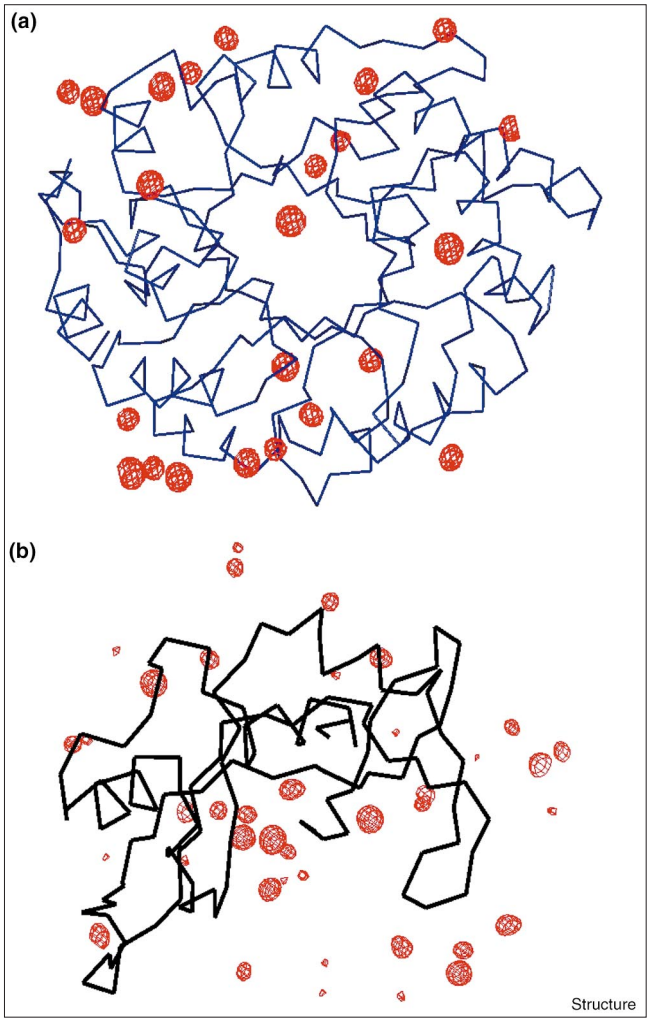
There have also been changes of more fundamental character in the usage of derivatization for protein crystallography. For example, derivatization prior to crystallization has been used to alter the surface properties of proteins recalcitrant to crystallization to promote the formation of useful crystals [34]. This technique has arisen because of the advent of MAD phasing, for which the availability of an isomorphous native crystal is irrelevant. Techniques that rely on derivatization prior to crystallization could benefit substantially by the application of native gel screening. In a few cases, the resolution limits of diffraction for some crystals have been found to improve upon derivatization. Heavy-atom screening with poorly diffracting crystals has been used on at least one occasion expressly for the purpose of attaining an improved diffraction limit [35].

Figure 3



Mass spectrometry analysis of native and $Hg(C_2H_3O_2)_2$ derivatized Acrp30. The native protein has a mass of 16045 Da and the $Hg(C_2H_3O_2)_2$ derivatized protein a mass of 16258 m/z, corresponding to modification of the 213 Da protein. This is approximately the mass of a single mercury atom (200 Da). The 13 Da difference must stem from either the addition of another chemical group, or an error in the measurement. The myoglobin standard had to be excluded from the derivative spectra so as not to overwhelm the small signal. Derivatization was carried out as described in Figure 2. The solution was then diluted with 200 µl of 10 mM bis-tris pH 6.0, and, filtered through a Microcon 10 (Amicon, Inc.) to remove unreacted $Hg(C_2H_3O_2)_2$, 5 µl of the 30 µl retentate was then lyophilized and redissolved in 6 µl matrix solution for MALDI analysis.

Figure 4



Difference Fourier maps produced from fast soaks with halides. **(a)** Bijvoet difference Fourier map produced from a xylanase crystal soaked in 0.5 M sodium iodide cryo-solution for 10 s. **(b)** Bijvoet difference Fourier map produced from an RNAase A crystal soaked in 1 M sodium bromide cryo-solution for 40 s. The peaks are contoured at 5σ. This figure was kindly provided by K.R. Rajashankar of Brookhaven National Laboratory.

Fast incorporation of halides

A rapid soak of protein crystals (on the order of 1 minute) in a cryo-protectant solution containing up to 1 M bromide or iodide anions can often lead to incorporation of these anomalously scattering centers into the ordered solvent region around protein molecules (Figure 4) [36]. These ions bind through electrostatic interactions with hydrophilic regions of the protein surface. This is in marked contrast to xenon, which tends to bind in hydrophobic cavities.

Bromide and iodide ions have suitable anomalous and isomorphous scattering properties to provide phasing for protein structure solution. Bromine can be used for MAD because of its accessible K-absorption edge (0.92 Å). Although no absorption edge for iodine is within an accessible energy range for MAD, its anomalous signal can be useful for single wavelength anomalous diffraction (SAD) and SIR/AS methods.

Halides, being small monoatomic ions, are able to substitute for solvent water molecules around the protein surface and do not show strong preference for specific coordination geometry. In contrast, most metal ions show a preference for a particular coordination, and do not as easily find appropriate ordered sites. This approach requires very little preparative effort and might be particularly applicable for high-throughput crystallographic projects.

Application of the fast halide soak method can be used in conjunction with other heavy-atom and anomalous scattering experiments to produce suitable phases for protein structure determination. Because halide soaking might result in many poorly occupied sites, it might often be difficult to determine the position of these sites with standard Patterson or direct techniques. This approach can be used in combination with other phasing vehicles to augment the phase calculations, however, with the site positions determined using difference Fourier analysis. Although each site might be poorly occupied, the aggregate phasing power can be substantially increased over either method alone. Because of the close proximity in energies of the selenium and bromine K edges, these methods could conceivably be combined in a single MAD experiment using five (or more) wavelengths around these edges.

The bottom line

The famous American baseball player Yogi Berra coined the adage, 'It ain't over 'til it's over.' Unfortunately, this saying applies in full force to derivative screening. Although all non-crystallographic methods, and even crystallographic statistics might attest to the near certain presence of phasing atoms, only finding them by Patterson or other methods is real proof of their usefulness in carrying out a structure determination. The use of systematic screening methods, such as native PAGE, however, can rapidly provide useful leads in finding a phasing derivative, and so speed the process of structure solution.

Acknowledgements

The history of phasing atom methods is expansive; we therefore must apologize for our inability to reference many important contributors due to space limitations. We thank Mary Ann Gawinowicz for conducting mass spectrometry experiments, KR Rajashankar for providing figures, and Aneel Aggarwal, Peter Kwong, and Wayne Hendrickson for helpful suggestions. LS is the recipient of a Career Scientist Award from the Irma T Hirschl Foundation. TJB is a recipient of a Wellcome Trust International Prize Travelling Research Fellowship.

References

1. 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 J.* **9**, 1665-1672.
2. Smith, J.L. & Thompson, A. (1998). Reactivity of selenomethionine – dents in the magic bullet? *Structure* **6**, 815-819.
3. Blundell, T. & Johnson, L.N. (1976). *Preparation of Heavy Atom Derivatives.* Academic Press, London.
4. Petsko, G.A. (1985). Preparation of isomorphous heavy-atom derivatives. *Methods Enzymol.* **114,** 147-156.
5. Rould, M.A. (1997). Screening for heavy-atom derivatives and obtaining accurate isomorphous differences. *Meth. Enzymol.* **276**, 461-472.
6. Carvin, D., Islam, S.A., Sternberg , M.J.E. & Blundell, T.L. (1998). HAD, A databank of heavy-atom binding sites in protein crystals: a resource for use in multiple isomorphous replacement and anomalous scattering. *Acta. Crystallogr. D* **54**, 1199-1206.
7. Shapiro, L., *et al.*, & Hendrickson, W.A. (1995). Structural basis of cell–cell adhesion by cadherins. *Nature* **374**, 327-337.
8. Egloff, M.P., Cohen, P.T., Reinemer, P. & Barford, D. (1995). Crystal structure of the catalytic subunit of human protein phosphatase 1 and its complex with tungstate. *J. Mol. Biol.* **254**, 942-959.
9. Lima, C.D., Klein, M.G. & Hendrickson, W.A. (1997). Structure-based analysis of catalysis and substrate definition in the HIT protein family. *Science* **278**, 286-290.
10. Wagner, R.L., Apriletti, J.W., McGrath, M.E., West, B.L., Baxter, J.D. & Fletterick, R.J. (1995). A structural role for hormone in the thyroid hormone receptor. *Nature* **378**, 690-697.
11. Hendrickson, W.A., Pahler, 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. *Proc. Natl Acad. Sci. USA* **86**, 2190-2194.
12. 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.
13. Sun, D.P., Alber, T., Bell, J.A., Weaver, L.H. & Matthews, B.W. (1987). Use of site-directed mutagenesis to obtain isomorphous heavy-atom derivatives for protein crystallography: cysteine-containing mutants of phage T4 lysozyme. *Protein Eng.* **1**, 115-123.
14. Hatfull, G.F., Sanderson, M.R., Freemont, P.S., Raccuia, P.R., Grindley, N.D. & Steitz, T.A. (1989). Preparation of heavy-atom derivatives using site-directed mutagenesis. Introduction of cysteine residues into gamma delta resolvase. *J. Mol. Biol.* **208**, 661-667.
15. Crick, F.H.C. & Magdoff, B. (1956). The theory of the method of isomorphous replacement for protein crystals. *Acta Cryst.* **9**, 901-908.
16. Hendrickson, W.A. & Ogata, C.M. (1997). Phase determination from multiwavelength anomalous diffraction measurements. *Methods Enzymol.* **276**, 494-523.
17. 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.
18. Sigler, P.B. & Blow, D.M. (1965). A means of promoting heavy-atom binding in protein crystals. *J. Mol. Biol.* **14**, 640-644.
19. Lusty, C.J. (1999). A gentle vapor-diffusion technique for cross-linking of protein crystals for cryocrystallography. *J. Appl. Cryst.* **32**, 106-112.
20. Davis, B.J. (1964). Disc electrophoresis II: methods and application to human serum proteins. *Ann. NY Acad. Sci.* **121**, 404-427.
21. Cann, A.D., Bishop, S.M., Ablooglu, A.J. & Kohanski, R.A. (1998). Partial activation of the insulin receptor kinase domain by juxtamembrane autophosphorylation. *Biochemistry* **37**, 11289-11300.
22. Shapiro, L. & Scherer, P.E. (1998). The crystal structure of a complement-1q family protein suggests an evolutionary link to tumor necrosis factor. *Curr. Biol.* **8**, 335-338.
23. Beavis, R.C. & Chait, B.T. (1990). High-accuracy molecular mass determination of proteins using matrix-assisted laser desorption mass spectrometry. *Anal. Chem.* **62**, 1836-40.
24. Hillenkamp, F., Karas, M., Beavis, R.C. & Chait, B.T. (1991). Matrix-assisted laser desorption/ionization mass spectrometry of biopolymers. *Anal. Chem.* **63**, 1193A-1203A.
25. Chait, B.T. (1994). Mass spectrometry – a useful tool for the protein X-ray crystallographer and NMR spectroscopist. *Structure* **2**, 465-467.
26. Krishna, T.S., *et al.*, & Kuriyan, J. (1994). Crystallization of proliferating cell nuclear antigen (PCNA) from *Saccharomyces cerevisiae. J. Mol. Biol.* **241**, 265-268.
27. Budisa, N., Steipe, B., Demange, P., Eckerskorn, C., Kellermann, J. & Huber, R. (1995). High-level biosynthetic substitution of methionine in proteins by its analogs 2-aminohexanoic acid, selenomethionine, telluromethionine and ethionine in *Escherichia coli. Eur. J. Biochem.* **230**, 788-796.
28. Holden, H.M. & Rayment, I. (1991). Trimethyllead acetate: a first-choice heavy atom derivative for crystallography. *Arch. Biochem. Biophys.* **291,** 187-194.
29. Schoenborn, B.P., Watson, H.C. & Kendrew, J.C. (1965). Binding of xenon to sperm whale myoglobin. *Nature* **207,** 28-30.
30. Vitali, J., Robbins, A.H., Almo, S.C. & Tilton, R.F. (1991). Using xenon as a heavy atom for determining phases in sperm whale. *J. Appl. Cryst.* **24**, 931-935.
31. Schiltz, M., Prange, T. & Fourme, R. (1994). On the preparation and X-ray data collection of isomorphous xenon. *J. Appl. Cryst.* **27**, 950-960.
32. Cate, J.H. & Doudna, J.A. (2000). Solving large RNA structures by X-ray crystallography. *Methods Enzymol.* **317**, 169-180.
33. 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.
34. Krishna, T.S.R., Kong, X.P., Gary, S., Burgers, P.M. & Kuriyan, J. (1994). Crystal structure of eukaryotic DNA polymerase procesivity factor PCNA. *Cell* **79**, 1233-1243.
35. Chang, G., Spencer, R.H., Lee, A.T., Barclay, M.T. & Rees, D.C. (1998) Structure of the MscL homolog from mycobacterium tuberculosis: a gated mechanosensitive ion channel. *Science* **282**, 2220-2226.
36. Dauter, Z., Dauter, M. & Rajashankar, K.R. (2000). Novel approach to phasing proteins: derivatization by short cryo-soaking with halides. *Acta Crystallogr. D* **56**, 232-237.