The suitability of multi-metal clusters for phasing in crystallography of large macromolecular assemblies

Jesper Thygesen[1], Shulamit Weinstein[2], Francois Franceschi[3] and Ada Yonath[1,2]*

Addresses: [1]Max-Planck Unit for Structural Molecular Biology, c/o DESY, Notkestrasse 85, D-22603, Hamburg, Germany, [2]Department of Structural Biology, Weizmann Institute, Rehovot 76100, Israel and [3]Max-Planck-Institute for Molecular Genetics, Ihne strasse 73, 14195 Berlin, Germany.

*Corresponding author. E-mail: csyonath@weizmann.weizmann.ac.il

Structure 15 May 1996, 4:513–518

 ISSN 0969-2126

The assignment of phases to the observed structure factor amplitude is the most crucial, albeit most complicated, step in structure determination. As the phases cannot be directly measured, their elucidation remains the least predictable task, even for average-sized proteins. Clearly, for large macromolecular assemblies the magnitude and the complexity of phasing is greatly enhanced.

The methods commonly used for phase determination in biological crystallography are single and multiple isomorphous replacement (SIR or MIR). Both require the preparation of derivatives, usually by introducing electron-dense compounds into the crystalline lattice at a limited number of distinct locations while keeping the crystal parameters isomorphous with those of the native molecule. As these replacement methods exploit the changes in the structure factor amplitudes resulting from the addition of the heavy atoms, the derivatization reagents are chosen according to their potential ability to induce measurable signals. For proteins of average size, useful heavy-atom derivatives consist of one or a few heavy-metal atoms that usually have an atomic number (Z)>70. However, for producing measurable signals from crystals of very large macromolecules that cannot be subdivided by non-crystallographic symmetry, numerous atoms (approximately 35 per 10^6 Da) are needed. Such multiple-site derivatives are extremely difficult to locate in the unit cell, and therefore, practically, this approach is not feasible.

Alternatively, advantage can be taken of compact and dense compounds containing a large number of heavy atoms linked directly to each other or arranged in close proximity. However, in contrast to the availability of numerous single-atom derivatizing agents, there are only a few stable water-soluble polymetallic compounds that may be suitable for derivatization of large biological compounds. These include heteropolyanions and multicoordination compounds, suitable for soaking experiments, and monofunctional reagents of dense metal clusters, designed for covalent binding at specific sites before crystallization.

Independent of which method is chosen, the phasing capacity of the clusters is limited by the usable resolution of the X-ray data, which is dictated not only by the quality of the crystals but also by the mode of binding of the clusters. As clusters offer more than one atom capable of coordination with the macromolecule, and as not all clusters possess internal symmetry, they may bind in several modes. If this occurs, they should be treated as group scatterers and the individual positions of the heavy atoms are not resolved. Under these conditions the clusters are expected to phase, in most cases, to 6–8 Å (about 66% of their average diameter). However, for clusters with perfect symmetry, or when the attachment to the macromolecule is created by a specific atom of the cluster, the individual positions of the cluster atoms may be resolved, and the phasing power may extend to almost atomic resolution.

Soaking and co-crystallization

Derivatization of crystals of biological macromolecules is routinely obtained by soaking in solutions containing millimolar concentrations of heavy-atom compounds. This method assumes that the high affinity of the metal compound for specific exposed side chains will lead to selective attachment of the heavy atom with a high yield. Using this equilibrium procedure, productive derivatization in one or a few sites is largely a matter of chance, but the probability of obtaining a useful derivative is sufficiently high that more sophisticated techniques are rarely needed. The large size of the dense clusters required for the derivatization of large biological macromolecular assemblies may be advantageous as it should eliminate multiple-site binding. However, at the same time, the cluster's large size may prohibit the diffusion of the clusters into the crystal. Despite this legitimate doubt, conditions can be established for the derivatization of crystals of very large particles, including ribosomes, by soaking in solutions containing multi-heavy-atom clusters.

The heavy metal cluster most frequently used for this type of study, TAMM (tetrakis [acetoxymercuri] methane), is rather small. It contains four mercury atoms and proved suitable for phasing data from crystals of rather large particles such as the photosynthetic reaction center [1], the nucleosome core histone octamer [2], an iodotype–anti–iodotype complex [3] and glutathione transferase [4].

Table 1

Heteropolyanion clusters.

W30	$K_{14}(NaP_5W_{30}O_{110})31H_2O$	[8]
W12	$K_5H(PW_{12}O_{40})nH_2O$	[9]
W18	$(NH_4)_6(P_2W_{18}O_{62})14H_2O$	[10]
W17Co or CoWLi17	$Cs_7(P_2W_{17}O_{61}Co(NC_5H_5))nH_2O$	*
BuSnW17	$K_7[(BuSn)(P_2W_{17}O_{61})]nH_2O$	[11]
PhSnW15	$K_5H_4[(PhSn)_3(P_2W_{15}O_{59})]nH_2O$	[11]
BuSnW15	$K_5H_4((BuSn)_3(P_2W_{15}O_{59})]nH_2O$	[11]
	$Na_{16}[(O_3PCH_2PO_3)_4W_{12}O_{36}]nH_2O$	*

Abbreviations: Bu, butyl; Ph, phenyl. *MT Pope, personal communication.

The first attempts to use somewhat larger compounds, Ta_6Cl_{14} and Nb_6Cl_{14}, were made over 30 years ago. Both compounds were co-crystallized with lysozyme [5] and the difference in the electron densities of Ta and Nb was sufficient to determine their exact location in the unit cells. Nevertheless, these compounds were not useful for the structure determination of lysozyme as the co-crystals and native crystals were not isomorphous. Recently, Ta_6Br_{14} was shown to be suitable for phasing over a wide pH range at different resolution limits. Thus, it was used for structure determinations of ribulose-1,5-phosphate carboxylase/oxygenase (rubisco) and transketolase at 5.5 Å [6], and a proteosome from *Thermus acidophilum* at 3.4 Å [7].

Several multi-tungsten compounds (Table 1; [8–11]; MT Pope, personal communication) are being used for derivatization, some of which allow phasing at relatively high resolution: 2.6 Å for fumarase C [12] and 3.3 Å for riboflavin synthase [13]. These large heteropolytungstate anions are of exceptional stability over a wide range of pH and redox states. They possess a high degree of internal symmetry and a correlation between this and their binding sites was detected. Thus, a tungsten compound $[(W_3O_2(O_2CCH_3)_6]^{++}$, of a trigonal symmetry (bipyramid of W_3O_2), binds on the threefold axis of riboflavin synthase, whereas the W30 ion (Table 1), which is a pentagonal hollow disk formed by cyclic assembly of five PW_6O_{22} units [8], binds on the fivefold axis [13].

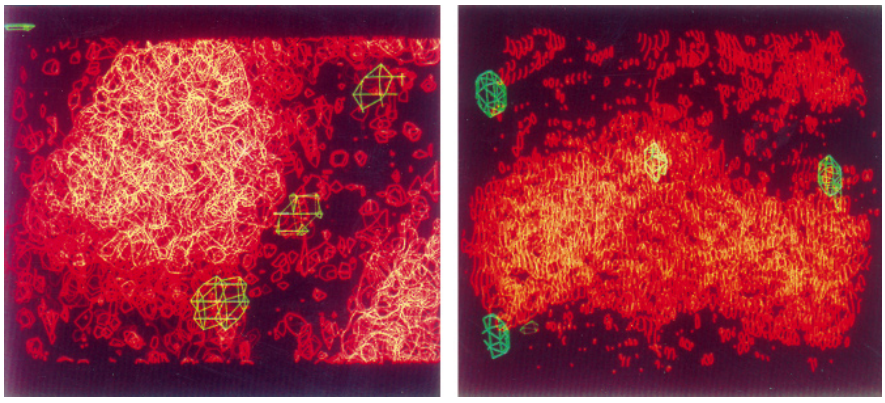
The high level of symmetry of these cluster compounds does not exclude them from binding to particles with no internal symmetry, but the mode of this binding is unpredictable. Crystals of the large ribosomal subunits from *Haloarcula marismortui* (H50S), soaked in either W30, W18, W17Co or W12 (Table 1) diffract best to 4–7 Å [14,15], compared with 2.9 Å for the native subunits [16]. All four clusters were found to bind at the same major sites and of these four clusters, only the locations of W30 could be directly determined from a 12 Å difference Patterson map. This was done by two parallel approaches. In one, the cluster was treated as a group scatterer. In the second, all significant peaks in the difference Patterson map were assigned as single W atoms. Interestingly, during the refinement process, the so-assigned W atoms merged to form clusters at the same positions found by the other approaches (HAS Hansen, unpublished data). These sites (Fig. 1) were also confirmed by *ab initio* calculations using the differences between the structure factors of the native and derivative crystals [14,15].

The crystals of the small ribosomal subunits from *Thermus thermophilus* (T30S), which diffract best to 7.3 Å [17], behave differently. Soaking them in a solution of W30 resulted in poor diffraction, whereas W12 and W18 (Table 1) led to a preliminary MIR electron density map at 19 Å. The map showed features similar to those obtained by image reconstruction [18] and the packing observed was in accord with the motif seen in thin sections of embedded crystals as well as with a map constructed from independent *ab initio* phase sets. For the interpretation of the difference Patterson maps, the W12 cluster was approximated by a sphere, whereas the ellipsoid W18 cluster was represented by two adjacent scatterers, each composed of nine W atoms [14,15].

Specific covalent binding of heavy atom clusters

Quantitative covalent binding of a heavy atom compound at predetermined sites prior to crystallization is an alternative for derivatization by soaking. This approach requires

Figure 1



Two views (in orange) of portions of the SIR 9–12 Å electron density map of H50S on which the positions of the W30 cluster are shown (in green). Crystals of H50S grow as extremely thin plates and diffract best to 2.9 A [16]. They are typically $0.3\times0.3\times0.01$ mm^3 in size; symmetry, $C222_1$; cell dimensions, 212 Å, 302 Å and 567 Å. Crystals were soaked for up to 34 days in solutions containing 1–2 mM of W30, W12, W18 or W17Co. For all sets the $R_{merge}(I)$=8.4–10.3% and the completeness was 74–91%. The number of major sites are 2, 2, 1 and 4, respectively, with occupancies between 0.2 and 0.6. Phasing power was between 0.7 (for the W17Co cluster) and 1.36 (for the W30 cluster). The overall figure of merit, after two cycles of solvent flattening was 0.65 [14,15].

complicated and time-consuming procedures, but is bound to yield indispensable information not only for phasing but also in the later stages of structure determination.

The feasibility of this approach for phasing data collected from weakly diffracting unstable crystals of large macromolecular assemblies at low resolution has been demonstrated in the construction of a 20 Å SIR electron density map of B50S [19]. Derivatization was performed by an undecagold cluster which contains a core of 8.2 Å diameter, consisting of eleven gold atoms linked directly to each other. These gold atoms are surrounded by hydrophilic organic groups, designed to increase the solubility and stability of the cluster [20], which causes the overall diameter to be 22 Å. As the metal components of this cluster are masked by a chemical envelope, they are not available for direct attachment with the macromolecule. For this reason, this cluster is not suitable for derivatization by soaking.

This undecagold cluster was prepared as a monofunctional reagent, specific to free sulfhydryls, by attaching to it a short aliphatic chain with a maleimido or an iodoacetyl group at its end [21,22]. This group was then bound to a sulfhydryl group of a purified ribosomal protein (BL11), after which the protein was reconstituted with core particles of a mutant lacking BL11. Although the cluster binding was carried out under denaturation conditions (as the isolated protein looses its *in situ* conformation so that its sulfhydryl group is no longer exposed as it is on the surface of the ribosome), and although the molecular weight of the cluster is almost half that of the protein (6.2 kDa versus 15.5 kDa), upon incorporation of the modified protein into the mutated ribosomal core quantitatively modified particles were formed (Fig. 2 and [18–22]). The fact that the bacteria can grow without BL11, in addition to the ability to crystallize cores lacking this protein, indicates that the removal of BL11 does not cause gross conformational changes in the ribosome and that BL11 is not likely to be involved in crystal packing. Indeed, the native, mutated and derivatized B50S crystallize under the same conditions and in the same space group and show the same level of isomorphism as found among typical preparations of native crystals.

The limited phasing resolution of this system stems not only from the rather poor quality of the B50S crystals (native crystals yield usable diffraction to about 18 Å) but also from the mode of binding of the undecagold cluster. As there are 21 possible non-identical positions for the attachment of the monofunctional arm, the position of the cluster in the unit cells should be treated as an average of many possibilities. Indeed, simulation studies have been carried out relating to this property as well as to the mean displacement that may occur due to the flexibility in the hydrophilic envelope and of the binding arm. They show that the phasing power of the cluster extends between 8 Å and 20 Å resolution [19], which is similar to that of the crystals of native B50S. Along these lines, it is conceivable that the undecagold cluster is suitable for phasing data from crystals of complexes mimicking defined states in the process of protein biosynthesis which diffract to comparable resolution. Examples are those of whole ribosomes with attached mRNA and tRNA, diffracting to 12–14 Å [22,23]. Therefore, the undecagold cluster was covalently attached to $tRNA^{phe}$ and $tRNA^{ile}$. It was found that the cluster binding did not impair the tRNA aminoacylation and its binding to the ribosome, and crystals of modified complexes could be grown and are currently being investigated [24].

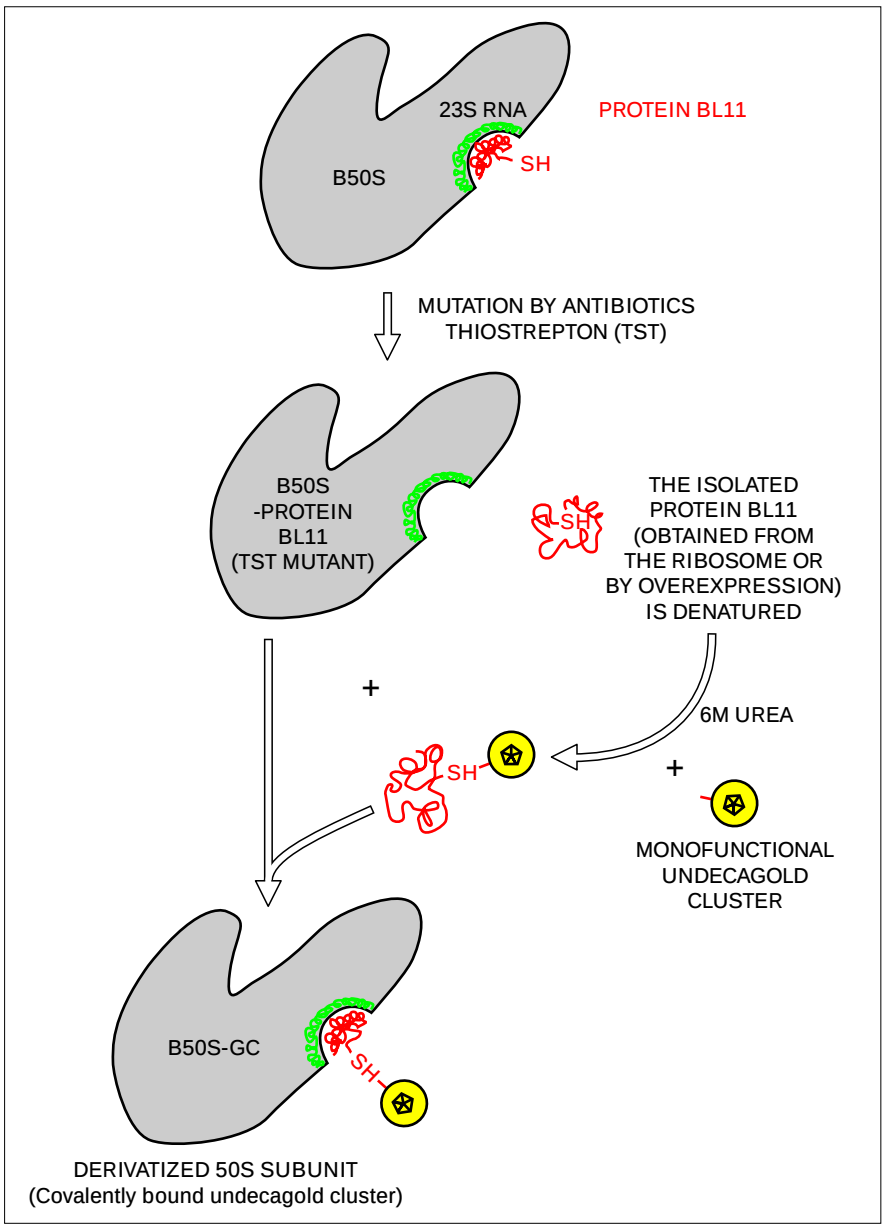
The demonstrated feasibility of using covalently bound clusters for low resolution phasing encouraged further effort relating to the construction of specifically derivatized crystals of large assemblies that diffract to higher resolution. At the same time it indicates the need for monofunctional reagents of either clusters of perfect symmetry, or those in which one specific atom can be 'flagged' and used for binding. An example of the latter is the tungsten clusters in which one or three structurally equivalent W atoms have been replaced by Sn, to which a chemically reactive arm can be bound (Table 1; [11]), or by metal, such as Rh or Ru, which has a high affinity to –SH groups allowing direct binding in a fashion similar to that of Hg. Such compounds have been prepared and their binding potential is currently being assessed (MT Pope, personal communication).

Alternatively, clusters of smaller size, such as, TAMM or a tetra iridium compound [25], may be used. The reduced size of the clusters may necessitate quantitative binding at several sites (4–6 for ribosomal particles). As it is unlikely that so many natural potential binding sites are available on the surface of the macromolecule, binding sites should be inserted either by mutagenesis or by chemical modifications of surface amino acids. (See, for example, the studies on the large halophilic ribosomal subunits [26].) For choosing appropriate locations for these insertions, procedures for specific quantitative detachment and reconstitution of selected ribosomal proteins have been developed, and the surfaces of the ribosomal particle have been mapped by chemical and enzymatic methods [26].

Clusters for phasing by multiwavelength anomalous diffraction (MAD)

MAD phasing is advantageous among crystals of large and unstable assemblies as these usually exhibit a low level of isomorphism. In contrast to MIR, which requires collecting data from a native and several heavy atom derivatized crystals, for MAD one crystal should be sufficient, provided that there is no radiation decay while measuring a few datasets at different wavelengths. MAD has been used to a lesser extent compared with MIR because its

Figure 2

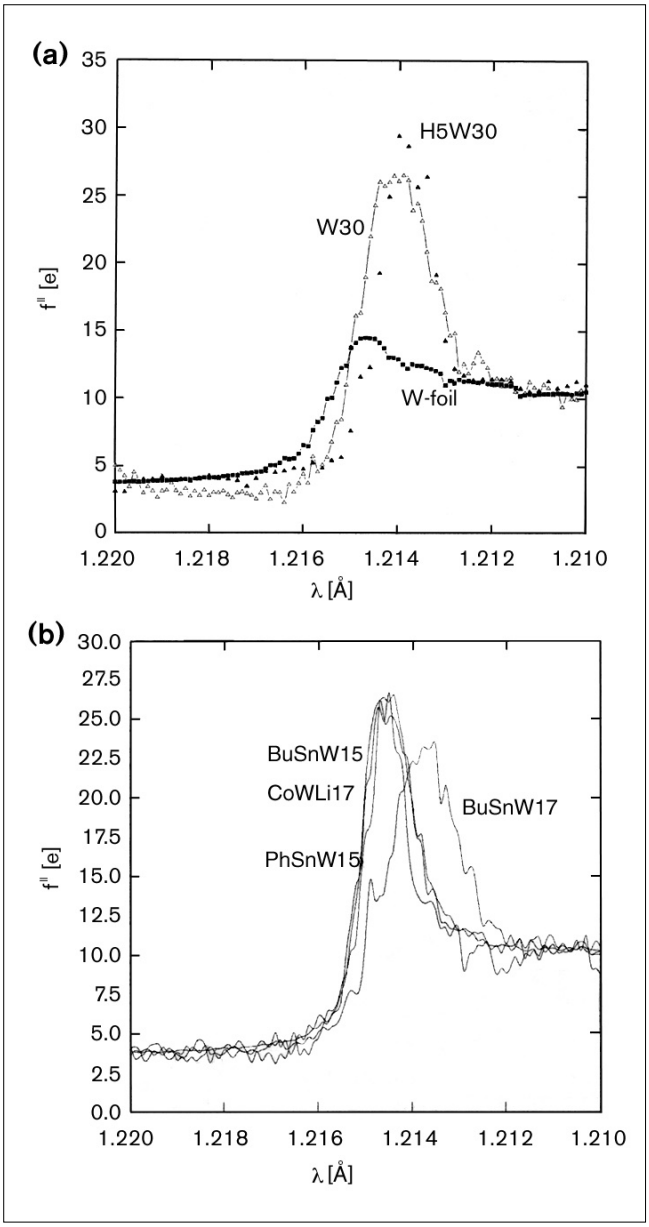


The procedure used for the quantitative and specific derivatization of B50S by binding a monofunctional undecagold cluster [19–22] to an exposed SH moiety on the ribosomes.

experimental requirements are very demanding [27]; the anomalous signal is 5–10 times lower than that obtained per heavy atom with MIR. Nevertheless, due to the recent vast progress in techniques for data collection, this method is currently gaining power and is being used to determine the structures of an increasing number of proteins of average size.

The derivatization procedures for MAD are essentially the same as performed for MIR, that is, soaking has been the procedure of choice. For such experiments, Ta_6Br_{14} seems an attractive compound, as it contains two different moieties (Ta and Br) each with a significant anomalous signal (M Roth and E Pebay-Peyroula, personal communication). However, for large macromolecular assemblies, special efforts for enhancing the probability of obtaining measurable anomalous signals should be made, preferably by quantitative binding of rather large dense heavy atom clusters at several sites. Some of the clusters mentioned above, especially the monofunctional reagents prepared from W12 and W18 (Table 1), may be suitable for this purpose. The anomalous contribution of several metal clusters, including those of W, has been demonstrated by fluorescence measurements even for soaked ribosomal crystals (Fig. 3). In fact, recently it was shown, for the first time, that tungstate is indeed an appropriate anomalous scatterer [28].

Figure 3



f″ as a function of wavelength derived from fluorescence spectra measured from (a) a crystal soaked in a solution of W30 (closed triangles), the solution itself (open triangles) and a tungsten foil (squares); and (b) the solutions of different W clusters in the H50S stabilization liquid. (For abbreviations see Table 1.)

Acknowledgements

We owe exceptional gratitude to the late Professor HG Wittmann with whom we initiated these studies, to Drs M Pope and W Jahn who gave us generous gifts of heavy atom clusters and to the team working with us on this project. Data were collected at the EMBL and MPG beam lines at DESY; F1/CHESS, Cornell University; D2AM/ESRF, Grenoble; and PF/KEK, Japan. Support was provided by the Max-Planck Society, the US National Institute of Health (NIH GM 34360), the German Ministry for Science and Technology (BMFT 05-641EAC) and the Kimmelman Center for Macromolecular Assembly at the Weizmann Institute. AY holds the Martin S Kimmel Professorial Chair.

References

1. Deisenhofer, J., Epp, O., Miki, K., Huber, R. & Michel, H. (1984). X-ray structure analysis of a membrane protein complex. Electron density map at 3 Å resolution and a model of the chromophores of the photosynthetic reaction center from *Rhodopseudomonos viridis*. *J. Mol. Biol.* 180, 385–398.
2. O'Halloran, T.V., Lippard, S.J., Richmond, T.J. & Klug, A. (1987). Multiple heavy-atom reagents for macromolecular X-ray structure determination. Application to the nucleosome core particle. *J. Mol. Biol.* 194, 705–712.
3. Bentley, G.A., Boulot, G., Riottot, M.M. & Poljak, R.J. (1990). Three-dimensional structure of an idiotype anti-idiotype complex. *Nature* 348, 254–257.
4. Reinemer, P., Dirr, H.W., Ladenstein, R., Schaeffer, J., Gallay, O. & Huber, R. (1991). The three dimensional structure of class II glutathione S-transferase in complex with glutathione sulfonate at 2.3 Å resolution. *EMBO J.* 10, 1997–2005.
5. Corey, R.B., Stanford, R.H., Marsh, R.E., Leung, Y.C. & Kay, L.M. (1962). An X-ray investigation of wet lysozyme chloride crystals. Preliminary report on crystals containing complex ions of niobium and tantalum. *Acta Cryst.* 15, 1157–1163.
6. Schneider, G. & Lindquist, Y. (1994). Ta_6Cl_{14} is a useful cluster compound for isomorphous replacement in protein crystallography. *Acta Cryst. D* 50, 186–191.
7. Löwe, J., Stock, D., Jap, B., Zwickl, P., Baumeister, W. & Huber, R. (1995). Crystal structure of the 20S proteasome from the Archaeon *T. acidophilum* at 3.4 Å resolution. *Science* 268, 533–539.
8. Alizadeh, M.H., Harmalker, S.P., Jeannin, Y., Martin-Frere, J. & Pope, M.T. (1985). A heteropolyanion with fivefold molecular symmetry that contains a nonlabile encapsulated sodium ion. The structure and the chemistry of $(NaP_5W_{30}O_{110})^{-14}$. *J. Am. Chem. Soc.* 107, 2662–2669.
9. Brown, G.M., Noe-Spirlet, M.R., Busig W.R. & Levy, H.A. (1977). Dodecatungtophosphoric acid hexahydrate $(H_5O_2+)_3(PW_{12}O_{40}{}^{-3})$. The true structure of Keggin's 'pentahydrate' from single-crystal X-ray and neutron diffraction data. *Acta Cryst. B* 33, 1038–1046.
10. Dawson, B. (1953). The structure of the 9(18)-heteropoly anion in potassium 9(18)-tungstophosphate, $K_6(P_2W_{18}O_{62})14H_2O$. *Acta Cryst.* 6, 113–126.
11. Xin, F. & Pope, M.T. (1994). Polyoxometalate derivatives with multiple organic groups. 1. Synthesis and structure of tris(organotin) beta-keggin and alpha-Dawson tungstophosphates. *Organometallics* 13, 4881–4886.
12. Weaver, T.M., Levitt, D.G., Donnelly, M.I., Wilkens Stevens, P.P. & Banaszak, L.J. (1994). The multisubunit active site of fumarase C from *E. coli*. *Nat. Struct. Biol.* 2, 654–662.
13. Ladenstein, R., Bacher, A. & Huber, R. (1987). Some observations of a correlation between the symmetry of large heavy atom complexes and their binding sites on proteins. *J. Mol. Biol.* 195, 751–753.
14. Schlünzen, F., *et al.*, & Yonath, A. (1996). A milestone in ribosomal crystallography: the construction of preliminary electron density maps at intermediate resolution. *J. Biochem. Cell Biol.*, in press.
15. Thygesen, J., *et al.*, & Yonath, A. (1996). Ribosomal crystallography: From crystal growth to initial phasing. *J. Cryst. Growth*, in press.
16. von Böhlen, K., *et al.*, & Yonath, A. (1991). Characterization and preliminary attempts for derivatization of crystals of large ribosomal subunits from *Haloarcula marismortui*, diffracting to 3 Å resolution. *J. Mol. Biol.* 222, 11–15.
17. Yonath, A., *et al.*, & Wittmann, H.G. (1988). Characterization of crystals of the small ribosomal subunits. *J. Mol. Biol.* 203, 831–833.
18. Berkovitch-Yellin, Z., Bennett, W.S. & Yonath, A. (1992). Aspects in structural studies on ribosomes. *Crit. Rev. Biochem. Mol. Biol.* 27, 403–439.
19. Bartels, H., *et al.*, & Yonath, A. (1995). The suitability of a mono functional reagent of an undecagold cluster for phasing data collected from the large ribosomal subunit from *Bacillus stearothermophilus*. *J. Peptide Sci.* 37, 411–419.
20. Jahn, W. (1989). Synthesis of water soluble undecagold cluster for specific labelling of proteins. *Z. Naturforsch.* 44b, 1313–1322.
21. Weinstein S., Jahn W., Hansen H.A.S., Wittmann, H.G. & Yonath, A. (1989). Novel procedures of derivatization of ribosomes for crystallographic studies. *J. Biol. Chem.* 264, 19138–19142.
22. Franceschi F., *et al.*, & Yonath A. (1993). Towards Atomic Resolution of Prokaryotic Ribosomes: Crystallographic, Genetic and Biochemical Studies. In *The Translational Apparatus.* (Nierhaus, K., ed), pp. 397–409, Plenum Press, NY.
23. Hansen, H.A.S., *et al.*, & Yonath, A. (1990). Crystals of complexes mimicking protein biosynthesis are suitable for crystallographic studies. *Biochem. Biophys. Acta* 1050, 1–5.

24. Weinstein, S., *et al.*, & Yonath, A. (1992). Derivatization of ribosomes and of tRNA with an undecagold cluster: crystallographic and functional studies. *J. Cryst. Growth* **122**, 286–292.
25. Jahn, W. (1989). Synthesis of water soluble tetrairidium cluster for specific labelling of proteins. *Z. Naturforsch.* **44b**, 79–82.
26. Sagi, I., *et al.*, & Yonath, A. (1995). Crystallography of ribosomes: Attempts at decorating the ribosomal surface. *Biophys. J.* **55**, 31–41.
27. Hendrickson, W.A. (1991). Determination of molecular structures from anomalous diffraction of synchrotron radiation. *Science* **254**, 51–58.
28. Egloff, M.P., Cohen, P.T.W., 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.