Published in final edited form as:

Nat Protoc. 2018 May ; 13(5): 1062–1090. doi:10.1038/nprot.2018.018.

Characterizing metal binding sites in proteins with X-ray crystallography

Katarzyna B. Handing^{#1,2}, Ewa Niedzialkowska^{#1,2,3}, Ivan G. Shabalin^{#1,2}, Misty L. Kuhn⁴, Heping Zheng^{1,2}, and Wladek Minor^{1,2}

¹Department of Molecular Physiology and Biological Physics, University of Virginia, Charlottesville, Virginia, USA. ²Center for Structural Genomics of Infectious Diseases (CSGID), University of Virginia, Charlottesville, Virginia, USA. ³Jerzy Haber Institute of Catalysis and Surface Chemistry, Polish Academy of Sciences, Krakow, Poland. ⁴Department of Chemistry and Biochemistry, San Francisco State University, San Francisco, California, USA.

These authors contributed equally to this work.

Abstract

Metals have vital roles in many physiological, pathological, toxicological, pharmaceutical, and diagnostic processes. Proper handling of metal-containing macromolecule samples for structural studies is not trivial, and failure to handle them properly is often the source of irreproducibility caused by issues such as pH changes, incorporating unexpected metals, or oxidizing/reducing the metal. This protocol outlines the guidelines and best practices for characterizing metal binding sites in protein structures and alerts the experimenter to potential pitfalls during preparation and handling of metal-containing protein samples for X-ray crystallography studies. The protocol features strategies for controlling the sample pH and the metal oxidation state, recording X-ray fluorescence spectra, and collecting diffraction datasets above and below the corresponding metal absorption edges. This protocol should allow the experimenter to gather sufficient evidence to unambiguously determine the identity and location of the metal of interest, as well as to accurately characterize the coordinating ligands in the metal binding environment within the protein. Meticulous handling of metal-containing macromolecular samples as described in our protocol will enhance experimental reproducibility in biomedical sciences, especially in X-ray macromolecular crystallography. For most samples, the protocol can be completed within 7–190 days, in which most of the time (2–180 days) is devoted to growing the crystal. The protocol should be readily understandable for structural biologists, in particular protein crystallographers with an intermediate level of experience.

Correspondence should be addressed to H.Z. (dust@iwonka.med.virginia.edu) or W.M. (wladek@iwonka.med.virginia.edu).

CONTRIBUTIONS

K.B.H. performed essential experiments to identify the pH-dependent conformational changes of residues coordinating the transition metal in the albumin; E.N. provided important data and experience on ICP-OES and TSA experiments; K.B.H., I.G.S., E.N., M.L.K. and H.Z. provided critical experience in metal-binding protein production and purification; K.B.H., I.G.S., E.N. and W.M. provided critical experience in metal-binding protein crystallization and data collection; H.Z. and I.G.S. provided extensive experience in characterization of metal binding sites in protein structures; H.Z. and I.G.S. laid out the structural framework of the manuscript; M.L.K. raised many constructive comments and extensively edited the manuscript; K.B.H., E.N., I.G.S., M.L.K., H.Z. and W.M. wrote and approved the manuscript; and H.Z. and W.M. supervised the project.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Keywords

metal binding site; pH; transition metal; metalloprotein; X-ray fluorescence spectrum; X-ray fluorescence scan; anomalous electron density map; X-ray crystallography

INTRODUCTION

Metal ions play important roles in a broad range of cellular processes^{1,2}. Typically, alkali and alkaline earth metals, such as Na^+ , K^+ , Mg^{2+} , and Ca^{2+} , play a structural role^{3,4}, while transition metals are often involved in both structure stabilization^{5,6} and catalysis^{6,7}. For example, some transition metals, such as Fe^{2+} or Cu^{2+} , have the ability to easily change oxidation states and are critical in catalyzing oxidation-reduction (redox) reactions⁸, while others, such as Zn^{2+} , act as Lewis acids in enzyme active sites or have structural roles in proteins and nucleic acids⁹. The concentration of metals required for physiological functions of living organisms varies from millimolar concentrations (e.g. Na^+ , Mg^{2+} , Ca^{2+} , K^+) to trace amounts (e.g. Co^{2+} , Ni^{2+} , V^{2+} , Mn^{2+})¹⁰. Metals are also important in medicine in the form of metal-based drugs¹¹, including the anti-cancer drugs cisplatin and carboplatin¹². Regardless of whether the presence of metal is essential, toxic, or therapeutic, the reactivity of metals is a consequence of their unique characteristics and interactions with macromolecules such as proteins and nucleic acids. To unveil the molecular details of these protein-metal interactions and gain a greater understanding of the role of metals in the cell, X-ray crystallography is often the method of choice.

While preparing metal binding proteins for structural studies, we observed that the formation of protein-metal complexes can be easily affected by sample preparation, contamination with trace amounts of other metals, and variations of environmental conditions such as pH or temperature. For example, the pH of the crystallization cocktail can differ significantly from the pH of the buffer used to make the cocktail after adding other components (e.g., metals). Chemicals used for buffer preparation may be insufficiently pure and can therefore be a source of trace amounts of metal. Also, when a protein is purified using metal affinity chromatography a metal may be stripped from the resin and bind to the protein. These observations prompted us to formulate a routine procedure for the preparation, crystallization, data collection, and structure determination of metal binding proteins. This protocol is designed for the most frequently encountered and biologically relevant transition metals (Mn^{2+} , Fe^{2+} , Fe^{3+} , Co^{2+} , Ni^{2+} , Cu^{2+} , Zn^{2+}). However, this protocol only describes the general guidelines regarding handling of redox-active metals, since some redox-active metal-protein complexes may require special handling^{13–15}. Most parts of the protocol can also be applied to the most frequently encountered and biologically relevant alkali and alkaline earth metals (Na^+ , Mg^{2+} , Ca^{2+} , K^+). With modification and careful consideration from other works^{16–20}, the protocol should also prove valuable for metals present only in the event of toxicity or when introduced as a part of a drug's active compound (Cd^{2+} , Au^{1+} , Pt^{2+} , Hg^{2+} , Pb^{2+}).

DEVELOPMENT OF THE PROTOCOL

Researchers have been interested in metalloprotein structures since the early-days of crystallography. The crystal structures of iron-containing hemoglobin (PDB ID: 1FDH)²¹ and manganese- and calcium-containing concanavalin A (PDB ID: 2CNA)²² are among the earliest structures deposited in the Protein Data Bank (PDB)²³. Thereafter, many manuscripts and book chapters regarding the expression and purification of metal-containing proteins, handling of metals with more than one oxidation state, crystallization, data collection of metal-containing crystals, and validation of protein-metal binding environments were published^{15,20,24–37}. To our knowledge, however, there are currently no publications that gather this information into a clear, easy to use, step-wise protocol.

The protocol presented herein was developed based on the structural and biochemical studies of several families of proteins^{38–48}, which resulted in the deposition of 45 crystal structures into the Protein Data Bank (PDB), data mining of metal binding environments in protein⁴⁹ and RNA structures⁵⁰, and extensive experience with tools for validation of metal binding sites in macromolecular structures³³. For example, we have successfully applied this protocol to the crystallization of serum albumin, the major plasma transport protein, to obtain complexes with a variety of metals⁵¹. This protein proved to be a convenient case for testing and optimizing techniques for sample preparation, crystallization, and data analysis of protein-metal complexes. Some steps of the protocol were similar to the strategies developed in parallel by different groups^{14,15}. For example, in studying interactions of metal-based drugs with proteins and consequent protein-metal adduct formation^{20,52,53}.

OVERVIEW OF THE PROCEDURE

The protocol summarizes common techniques used in various steps of structure determination of protein-metal complexes, with a particular focus on transition metals. It is intended to guide the user through the process of obtaining reliable chemically and biologically relevant data from the moment of sample preparation to the stages of crystallization, data collection, structure refinement, and validation of the metal identity and metal binding environment. The protocol is accompanied by potential pitfalls that may affect final results and is organized in a step-by-step manner. Additionally, we provide troubleshooting for problems and present potential caveats that may be encountered during sample preparation and data collection for protein-metal complexes.

The protocol is divided into five stages (Figure 1). The first stage involves the analysis of potential metal binding sites in the protein (Steps 1–2). The second stage is the preparation of a protein sample with unwanted metals removed (Steps 3–9). The third stage involves the preparation and flash-cooling of metal-containing protein crystals (Steps 10–12). The fourth stage is data collection (Steps 13–18). The fifth and final stage is modeling of metal binding sites (Steps 19–24). An overview of the steps is presented as a flow chart in Figure 1. For each step we provide additional background information and guidance in the Introduction. Steps 3–9 in the procedure are designed for a protein sample that can be produced without bound metals, and the metal of interest can be added during the crystallization process (step 10). For proteins that are produced with metals bound with high affinity, such as zinc finger proteins or superoxide dismutase, the procedures may be adjusted accordingly. For example,

the steps of removing (step 4) and/or adding metals (procedure step 10(A)i-vi) may be skipped.

CHALLENGES WITH CRYSTALLIZING PROTEIN-METAL COMPLEXES

There are several challenges associated with crystallizing protein-metal complexes. These challenges include preserving or incorporating the metal in the protein during sample preparation, monitoring and controlling pH, oxidation states, and metal concentration during the sample preparation process, and validating the metal content and metal binding environment during both sample preparation and model building processes. Considerations of these challenges are described in detail in the following sections.

Influence of pH on metal-binding sites and factors that affect it.—Crystallizing protein-metal complexes still poses challenges, even though the process of protein-metal binding has been studied for decades. One of the most important factors affecting protein-metal interaction and crystallization is the pH of the crystallization solution. The pH influences the binding properties of many amino acids in the metal binding environment, which is especially important to remember when working with proteins that have low (μM - mM) affinities for the metal of interest. The protonation state and conformation of amino acids that coordinate the metal ion (especially histidine and cysteine) can be significantly impacted by the variation of the environmental pH in solutions used during crystallization experiments, which in turn affect the ability to coordinate metals⁵⁴. For instance, this behavior was observed for zinc binding in the human S100B protein⁵⁵, where at pH 6.5 and pH 10.0 one zinc ion was coordinated by one glutamic acid and three histidine residues, while at pH 9.0 the glutamic acid residue was replaced by a fourth histidine residue. Changes in histidine protonation state and conformation induced by pH were also observed in a study of zinc binding to serum albumin⁵¹. The geometry of the main zinc binding site, which is responsible for the transport of 98% of the exchangeable fraction of Zn^{2+} in mammalian blood plasma⁵⁶, was affected by the proton concentration in the crystallization solution and resulted in different orientations of histidine residues coordinating zinc (Figure 2). Changes of the environmental pH may have a substantial effect on the fraction of a metal that is bound to the protein. For instance, a mobilferrin iron binding protein is 50% saturated with iron at pH 6.0 and 7.0, whereas at pH 5.0 it only contains 30%. Similar effects of environmental pH on the metal saturation of a protein were observed for bovine erythrocuprein, a protein that binds zinc and copper ions. At pH 3.6, 95% of zinc and 5% of copper dissociated from the protein, while at pH 5.8 there was a negligible loss of metal ions that originally bound to the protein⁵⁷.

The final pH of the environment in which the crystal grows (crystallization drop) is an important factor in protein crystallization experiments⁵⁸ and is often crucial for metal coordination⁵¹ and thus for functional analysis. The actual pH of the crystallization drop may be unknown and can be influenced by several factors. Firstly, information about the pH is not provided for approximately one-third of all crystallization solutions in commercially available screens. Even if information about the pH is included, it is often not pH of the final solution, but rather the pH of the buffering agent used as a solution component. The different pH levels of additional components, dilution factors, and source of water used for

preparation of the solutions⁵⁹ can affect the pH of final solution (Supplementary Table 1,⁶⁰). The buffering component is generally a minor part (0.05–0.1 M) of a crystallization solution, but the solution can also contain very high concentrations of salts (1–3 M) or polymers [up to 50% (wt./vol.)] that can swamp the effect of the buffering species. In particular, ammonium sulfate, a widely used salt, is known to alter the pH of any crystallization solution through the equilibrium of the ammonium ion with the volatile ammonia species: ammonia gas leaves and the solution that remains becomes more acidic⁶¹. Additionally, time and storage temperature can also affect the final pH of a crystallization solution. For example, during the aging of polyethylene glycols (PEGs), the final pH of a solution may decrease by as much as 3 points when solutions are stored at room temperature (20°C - 22°C) over a period of 16 months and exposed to light⁶². Furthermore, adding transition metal salts to a protein for co-crystallization and/or soaking native crystals may also significantly lower the final pH in the crystallization drop if the pH of the metal solution is not adjusted properly prior to the experiment. To illustrate the prominence of this issue, we have examined the influence of metal concentration on the pH of 50 mM and 100 mM Tris solutions by incrementally adding powdered ZnCl₂. Even in the presence of 100 mM Tris, the pH of the solution decreased by 0.5 pH units after adding 15 mM ZnCl₂ and further decreased with increasing concentrations of ZnCl₂ (Supplementary Figure 1). The addition of protein is yet another factor that affects the pH of the crystallization drop, especially when the protein buffer is at a high concentration and at a different pH than that of the crystallization solution. The pH of the crystallization drop may also change during the process of evaporation and may also be affected by the crystallization temperature, whereby it exhibits a change in pH during the crystallization experiment compared to its value measured at room temperature.

Inaccurately reporting pH in crystallization conditions may result in irreproducibility in the interpretation of structure- function relation as well as attempts of crystallizations by other laboratories. Unlike the modeling of metal binding sites, which can be validated later on in the refinement or re-refinement stages, the failure to report the correct pH can cause a permanent loss of data that is usually impossible to recover. Careful consideration of the many potential factors that can alter the pH of crystallization conditions during the preparation of metal-binding protein crystals is likely to enhance experimental reproducibility in protein crystallography.

Effective metal concentration may not be as expected.—Crystallization solutions contain a spectrum of chemicals that act as protein precipitants, buffers, and/or reagents to increase protein stability. The effective metal concentration in solution may be decreased by the formation of metal complexes with these chemicals. Many transition metal ions will form insoluble hydroxides in neutral to basic pH, which can significantly limit metal solubility and substantially decrease the effective metal concentration⁶³.

Besides the formation of insoluble hydroxides in solution, some crystallization reagents can act as chelators by creating coordinate bonds with metal ions in solution. This decreases the effective concentration of the metal and must be compensated for by increasing metal concentration added to solution. For example, Tris—a buffer commonly used in crystallization experiments—possesses a weak Zn²⁺-binding ability (log K_a = 1.94), which

prevents formation and precipitation of $\text{Zn}(\text{OH})_2$ and thus ensures that Zn^{2+} remains in solution⁶⁴. However, if the concentration of Tris is high (e.g. 50 mM Tris and 0.1 mM protein, which represents a 500-fold excess over protein), the effective concentration of free zinc is significantly decreased by Tris chelation and may be too low for the incorporation of Zn^{2+} into weaker Zn^{2+} binding sites in proteins⁶⁴. Similarly, citrate is a common crystallization precipitate that can decrease the effective metal ion concentration via chelation. The concentration of citrate in some of the commercial crystallization screens exceeds 1.5 M (15000-fold excess over protein), which significantly decreases the effective concentration of some metals.

Identity of the metal may not be as expected and requires verification.—The identity of the metal in a protein-metal complex structure may not be as expected for several reasons. During heterologous protein expression, physiological metals may bind to the protein of interest. Even if the experimenter adds exogenous metals, the physiological metal may remain in the macromolecule, especially if it is strongly bound. In a recent example, the structure of the copper- transport protein Atox-1 (PDB ID: 3IWX) contained copper^{38,65} which originated from the natural source during protein purification, instead of the cisplatin used for co-crystallization. Additionally, for proteins containing a His-tag that were purified using Ni^{2+} affinity chromatography (most commonly used nickel-nitrilotriacetic acid Ni-NTA), the Ni^{2+} can be sequestered from an affinity column during purification and “contaminate” the sample. During this process, Ni^{2+} can replace the weakly-bound physiological metal or bind to additional locations on the protein. For example, Ni^{2+} was found in the supposedly native structure of the cadmium stress protein YodA during the study of Zn^{2+} and Cd^{2+} binding⁶⁶. It is important to keep in mind that bioinformatics tools used to annotate protein function may also misidentify the metal predicted to interact with the protein in a physiological environment. And lastly, the solution used for co-crystallization or crystal soaking may contain unwanted metals which could be incorporated into the protein of interest, leading to confusion in interpreting results. For these reasons, it is necessary to control the presence and identity of a metal in the sample.

The metal identity within a protein can be confirmed using atomic absorption spectroscopy (AAS), inductively coupled plasma (ICP) methods, extended X-ray absorption fine structure (EXAFS)⁶⁷, X-ray fluorescence (XRF), and other methods. Unfortunately, the results of these experiments are not stored in the PDB or any other publicly available database and can sometimes only be found in manuscripts describing the results. Moreover, they only prove the presence of a metal in a sample and do not give information about its location in the macromolecule of interest. We suggest that the ultimate solution to confirm both the identity and location of the transition metal in the protein structure is to collect X-ray diffraction data above and below the metal anomalous absorption edge⁶⁸. This approach is fully applicable if the absorption edge of the metal is within the spectrum accessible by synchrotron beamlines (6 – 17.5 keV), which is the case for the most frequently encountered and biologically relevant transition metals (Mn^{2+} , Fe^{2+} , Fe^{3+} , Co^{2+} , Ni^{2+} , Cu^{2+} , Zn^{2+}), as well as some other metals (Cd^{2+} , Au^{1+} , Pt^{2+} , Hg^{2+} , Pb^{2+}). Regrettably, this practice has not been widely adopted by modern crystallography. There are two main reasons why researchers often skip transition metal identification with this method: 1) the beamline setup may not allow for

easy wavelength change and 2) the researcher assumes that a protein always binds a certain metal and therefore sees no reason to scrutinize the crystals for the presence of other metals during subsequent structures determinations. In addition, an incorrect wavelength may be used for data collection if the metal bound to the protein is different from what was anticipated. The absence of the anomalous signal should always raise red flag. Our own survey showed that only 1–8% of structures deposited in the PDB in complex with Zn, Cu, Ni, Mn, and Fe were collected at the wavelength near the metal absorption edge (–100eV and +25eV) (Supplementary Table 2).

Oxidation state of the metal or free cysteine residue(s) can change during crystallization or data collection.—Maintaining the anticipated oxidation states of the metal ion and free cysteine residues in the protein (if present) prior to and during a crystallization experiment is essential. The proper metal oxidation state is crucial for a metal-binding protein's function, as the oxidation state influences the metal's interactions with amino acid residues in the protein. The oxidation state of some metals (e.g., Fe, Cu, Co, Cr) may change upon exposure to atmospheric oxygen or when the pH of the solution is altered. For example, Fe²⁺ is stable in an acidic environment, but oxidizes to Fe³⁺ under neutral and basic conditions. Similarly, if free cysteine residues are present in the protein, their oxidation state can significantly affect protein conformation, physiological activity, and metal binding (if a cysteine residue is involved). The reagents used for crystallization may also change the redox environment and subsequently change the metal oxidation state even inside the crystal, such as in the case of residual peroxide contamination from PEG autooxidation⁶⁹.

Some metalloproteins are extremely sensitive to oxygen and require special handling from purification to crystallization. When working with oxygen-sensitive metalloproteins, manipulation of both protein and crystal samples should be performed under an anaerobic atmosphere. The general guidelines for maintaining an anaerobic environment are described in the section “PRELIMINARY ANALYSIS OF THE PROTEIN AND METAL BINDING SITE(S)”, while detailed procedures for handling redox-active metals are not described in this protocol.

Metal reduction due to radiation damage during X-ray data collection is also possible. The inelastic interactions of X- rays with the protein in the crystal during data collection can lead to crystal radiation damage and different chemical modifications of the protein⁷⁰. One specific effect of radiation damage to the protein is the reduction of some metal ions^{71–73}. During the crystallographic experiment, it is important to consider and address the possibility of the metal reduction, especially if the metal is expected to be in a higher oxidation state when multiple oxidation states are possible in a protein-metal complex.

Metal site modeling and refinement can be performed incorrectly.—Once the identity and location of the metal in the protein is verified using the various techniques mentioned above, the next step in determining the metal binding mode in the structure is to refine and validate the metal binding site. This step in the process is of ultimate importance because it finalizes the interpretation of the experimental data. It is critical to ensure that the coordination geometry, type of ligands, and bond distances are consistent with both the

experimental electron density and known coordination chemistry. Unfortunately, incomplete or even incorrect coordination geometry and coordination bond distances are commonly found in the PDB. We demonstrated previously that many functionally critical metal binding sites in the PDB have been refined incorrectly, even for structures determined at resolutions better than 2.0 Å³². Some structures show incorrect metal assignments or modeled binding environments. For example, the structure of the mature and fully active Der p 1 allergen (PDB code: 2AS8) has a magnesium ion incorrectly modeled with all six metal-oxygen distances in the range of 2.31 Å - 2.49 Å. This range is much longer than the typical Mg-O distance of 2.08 Å. A calcium ion would provide a much better fit in this case, as its typical metal-oxygen distance is 2.41 Å. Many additional examples of incorrectly modeled metals and/or binding environments have been discussed previously³³. Therefore, structures of protein-metal complexes must be analyzed with extreme caution, even when they have already been deposited into the PDB.

APPLICATIONS OF THE APPROACH

The protocol described herein is designed for studying metal complexes with soluble globular proteins. It may also be applied (with respective modification of sample preparation steps) to metal complexes with membrane proteins, nucleic acids, and proteins in complex with nucleic acids. Some steps of this protocol may be valuable for handling metal-binding proteins investigated by other structural methods (e.g. NMR and Cryo-EM), as well as by biophysical and biochemical methods (e.g., binding assays and kinetics studies). ...

The protocol can be applied to metal ion-macromolecular complexes that display a broad range of binding affinities. Oxygen-sensitive metal-protein complexes require special handling as noted in the “LIMITATIONS” and “PRELIMINARY ANALYSIS OF THE PROTEIN AND METAL BINDING SITE(S)” sections. With proper handling of oxygen-sensitive metalloproteins, the protocol is fully applicable to the most frequently encountered and biologically relevant transition metals (Mn²⁺, Fe²⁺, Fe³⁺, Co²⁺, Ni²⁺, Cu²⁺, Zn²⁺), as well as other metals that have an absorption edge within the spectrum accessible by most synchrotron beamlines with a tunable monochromator (6 – 17.5 keV, or 0.7 – 2.1 Å). Most parts of the protocol can be applied to any metal, except steps 13–18, which are not fully applicable to most alkali, alkaline earth, and some transition metals that do not have an absorption edge within the spectrum accessible by synchrotron beamlines. Therefore, a dataset below the absorption edge cannot be collected, and the disappearance of the corresponding anomalous signal cannot be verified. In most cases, however, these other metals still produce noticeable anomalous signals to identify their location. For example, potassium (K-edge at 3.607 keV) and calcium (K-edge at 4.038 keV) produce noticeable anomalous signals at the copper K-edge (8.979 keV) and lower energies, hence their location in the structure can still be identified by the anomalous signal. In combination with the metal content validation by ICP and X-ray fluorescence, one dataset with anomalous signal present may be enough to prove the identity and location of these metals that produce significant anomalous map peaks but do not have an absorption edge within the accessible spectrum. The same applies if the comparison of the anomalous maps does not show a significant difference between the peaks in datasets collected at the energies above and below the absorption edge (which may happen for transition metals from the 5th and 6th periods).

LIMITATIONS

It is essential for any X-ray crystallography study that the protein sample is soluble, homogenous, and in a sufficiently high concentration for crystallization (preferably free of affinity tags). Other methods, such as EXAFS or mass spectrometry, may have less strict sample requirements. However, structural information obtained from these techniques is usually complementary to X-ray crystallography. It is important to note that if the metal binds with a very high affinity, it may be difficult to replace the metal using the suggested protocol. One should be cautious when attempting to crystallize a protein in the presence of a metal ion that is different from its physiological metal, especially when the metal plays a structural or catalytic role in the protein. In such cases, the very high-affinity interactions between the protein and the metal ion may make it difficult to completely remove and replace the physiological metal ion from the protein, even when dialyzing the protein in the presence of a chelating agent or with more aggressive methods, such as protein denaturation and subsequent refolding.

It is critical for oxygen-sensitive metal-protein complexes to stay under an anaerobic environment during all steps of the procedure from protein purification to data collection. Some steps of the protocol may be difficult to perform inside an anaerobic chamber. For example, a glovebox equipped with instruments for measuring the binding affinity of the metal to the protein, or influence of metal on the stability of redox-active metal-protein (Steps 2 and 3) may be necessary. Additionally, protein production, purification and sample preparation (including affinity tag removal) must be performed under anaerobic conditions, which may require additional equipment (e.g. system for degassing buffers) (Step 3).

Some steps of the protocol presented here are suitable only for proteins interacting with metals with an already determined or estimated dissociation constant. Prior knowledge about the strength of the interaction between metal and protein is helpful in assessing which concentration of metal should be used for crystallization experiments. Finally, the metal identity may not be confidently verified by the diffraction experiments (steps 13–18) if the metal absorption edge falls outside of the tunable wavelength range of the synchrotron and if it is not possible to tune the wavelength both above and below the absorption edge.

ALTERNATIVE METHODS

The methods described in this protocol are routinely employed by the authors and are not intended to represent an exhaustive list of available methods for the analysis of metals in macromolecular samples. In the sections below, we provide references to the alternative methods to our best knowledge:

Alternative approaches to characterize metal binding affinity (Step 2): When metals bind with picomolar range affinity, it is non-trivial to measure these affinities by isothermal titration calorimetry (ITC), and ITC assays can yield incorrect metal affinity measurements. In these circumstances, competition assays using chelators with known metal affinities may be used to measure metal affinity such as in the case of zinc binding site in human calprotectin⁷⁴.

Alternative approaches to validate metal content in a protein sample (Steps 5, 13–15): In addition to the methods described in this protocol, other methods used to verify the metal content include electrospray ionization mass spectrometry (ESI-MS)^{53,75,76}, Raman microspectroscopy⁷⁷, and infrared spectroscopy (IR)⁷⁸. Additionally, the presence of metals can be studied even directly on crystals of metalloproteins or of protein-metal adducts using non-destructive methods such as *in crystallo* UV-visible spectroscopy⁷⁹, Raman microspectroscopy^{80–82} or IR⁸³. Using one of these non-destructive methods directly on the crystal offers the benefit that the data to verify the metal content can be collected on the same crystal right before the subsequent collection of a full X-ray diffraction data set at the synchrotron facility. Additionally, if the protein of interest is an metallo-enzyme, enzyme activity assays can be used to identify the physiological metal at the catalytic site in the protein, as well as to confirm the presence of the correct metal in the sample⁸⁴.

Alternative approaches to determine the location of the metal binding site (Step 18): An alternative method of metal assignment in crystal structures is calculation and display of dispersive $[F(\lambda_{\text{above}}) - F(\lambda_{\text{below}})]$ Fourier difference map that should give a positive peak at the metal position¹⁴. This method may be especially useful for transition metals from the 5th and 6th periods.

Alternative approaches to characterize protein-bound metal-based drugs: Similar methods have been developed to identify metal binding sites and metal ligands in the reaction between proteins and metal-based drugs containing Re, Ru, Pt or Au. A similar two-wavelength procedure has been used to unambiguously identify a Re ion in the structure of the lysozyme-Re compound adduct¹⁶ and to discriminate between Cd²⁺ and Pt²⁺ in the adducts formed in the reaction between cisplatin or carboplatin and ferritin¹⁷. Moreover, including data with a high R_{merge} but significant $CC_{1/2}$, can improve the electron density map to allow for the definition of metals^{18,19}.

PRELIMINARY ANALYSIS OF THE PROTEIN AND METAL BINDING SITE(S)

Predicting metal binding sites (Step 1).—Amino acid residues that interact with a certain metal tend to be conserved among homologous proteins. The most common residues that bind metal ions are histidine, cysteine, aspartate, and glutamate because the polar or charged atoms of their side chains can coordinate metals by providing a lone pair of electrons^{49,85}. Metal sites can also be coordinated by backbone atoms, and are sometimes coordinated by backbone atoms only (e.g. for K, Na, and Ca sites). Analyzing a structure of a close homolog of the protein of interest is helpful for predicting a metal-binding site in the protein being studied. Moreover, if a crystal structure of the protein of interest has been determined, it can be used to predict potential metal binding sites in the protein. Several web servers predict metal binding sites based on a structural model, including FEATURE⁸⁶, TEMSP⁸⁷, MetSite⁸⁸, MetalDetector⁸⁹, DiANNA⁹⁰, or FINDSITE-metal⁹¹. When inspecting the putative metal binding site, one should consider the potential change in amino acid conformation upon metal binding⁸⁶. If a structure is not available, the information about a putative metal binding site in a protein can be provided by resources that annotate or predict protein function, such as UniProt⁹² or ESG⁹³. Otherwise, it is suggested to search for homologs using BLAST⁹⁴. It is also possible to use more comprehensive machine-learning

predictions, such as MetalDetector⁸⁹, to infer the existence of a metal-binding site in the protein being investigated. In this protocol we describe procedures for identifying metal-bound homologous structures and using them to obtain information about the amino acids forming the putative metal binding sites in the investigated protein.

Working with oxygen-sensitive proteins (Steps 2–12).—It is critical that redox-active proteins are kept under an anaerobic atmosphere at all times. Therefore, we recommend performing steps 2–12 of the protocol inside an anaerobic chamber (e.g., a glove box) where the sample is exposed to an inert atmosphere (e.g., a mix of nitrogen and hydrogen) to eliminate the presence of oxygen^{95,96}. Procedures for special handling of redox-active metals are not described in detail in this protocol.

Estimating or experimentally determining the metal binding affinity (Step 2).—To better assess particular conditions for metal-protein crystallization experiments, biophysical and biochemical characterization of the metal binding features of the protein is of particular importance. For example, knowledge of stoichiometry and metal binding affinities ensures a sufficient metal concentration is used to saturate the metal-binding sites. Therefore, it is worthwhile to estimate the binding affinity between the protein and metal ion based on known data of a highly-homologous protein or by using quick screening techniques such as thermal shift assays (TSA) (Figure 3, Box 1). To fully characterize the binding affinity between the protein and metal ion, use ITC (Box 2). ITC experiments may show if there are multiple binding sites in the same protein or whether more than one type of metal ion can bind to the same binding site. Competition assays⁷⁴ may also be used as an alternative for measuring affinity as discussed in the alternative methods section. Other popular binding techniques, such as surface plasmon resonance (SPR) or fluorescence polarization (FP), may not be suitable to analyze very small ligands like metal ions. Additionally, constructing mutant proteins in combination with binding experiments may help pinpoint residues involved in metal binding. If the mutant proteins fold correctly and binding experiments show protein interaction with metal is abolished, this may indicate that the metal binding region in the protein has been predicted successfully. Nevertheless, the biophysical and biochemical characterization of the metal binding protein does not unambiguously identify amino acid residues responsible for metal binding. Therefore, initial clues regarding metal ion binding sites in a protein should always be confirmed by collecting and analyzing X-ray diffraction data.

CONSIDERATIONS FOR PROTEIN SAMPLE PREPARATION

Removing affinity tags from recombinant proteins (Step 3).—When working with recombinant metal-binding proteins, we highly recommend using affinity tags that allow for affinity purification using metal-free resins, such as maltose binding protein (MBP) or glutathione S-transferase (GST) tags. These tags may be left uncut because they are not likely to interfere with metal binding. The rationale for keeping the tags present on the protein may be imposed by factors such as protein stability, folding, or tendency to aggregate. Conversely, if a metal affinity purification tag, such as His-tag or calmodulin-binding domain/peptide⁹⁷, is present on a protein, it should be removed, as the tag may interact with metals. In addition, a His tag can bind in the active site and consequently affect

protein activity⁹⁸ or cause precipitation upon the addition of divalent metals due to aggregate formation through metal chelation⁹⁹. If removal of metal-affinity tags is impossible, extended dialysis or purification using chelating ion-exchange resin should be performed to deplete the metal ions chelated by the affinity tag. His-tags are especially likely to be problematic for proteins with weaker metal binding sites, but some enzymes with a tightly bound metal may be used in combination with a His-tag without complications. In summary, if any affinity purification tag is not removed from the protein, the effect of its presence should be carefully considered during data analysis.

Removing unwanted metals from the protein sample (Step 4).—It is crucial to remove any competing metal ions from the protein sample, especially in cases in which literature search, protein sequence, structure analyses, and/or experimental data reveal that a protein exhibits a proclivity for binding different metal ions. In extreme cases, the presence of different ions may influence the structure and characteristics of a protein, as has been shown for Fe and Ni ions in an acireductone dioxygenase enzyme¹⁰⁰. This step greatly depends on the properties of the protein of interest and the role that the metal plays in the protein-metal complex. For the removal of weakly bound metals, we recommend performing dialysis with chelators present in buffer, such as ethylenediaminetetraacetic acid (EDTA), ethylene glycol- bis(β-aminoethyl ether)-N,N,N',N'-tetraacetic acid (EGTA), 1,10-phenanthroline, or N,N,N',N'-tetrakis(2-pyridinylmethyl)-1,2-ethanediamine (TPEN), to trap metals. The selection of a proper chelator will depend upon the metal that needs to be removed. Some chelators have a higher affinity for a particular metal, while some can interact with a broad spectrum of metals. EDTA is generally a good choice to chelate most metals, citrate chelates iron and is less aggressive than EDTA, 1,10-phenanthroline chelates iron, and TPEN is a zinc chelator¹⁰¹. Alternatively, a Chelex 100 chelating ion-exchange resin (Bio-Rad) with a strong selectivity for Fe, Cu, and other polyvalent metal ions can be used to purify protein samples, buffers, or other reagents. If the metal is tightly bound and dialysis is not sufficient to remove it, more aggressive methods, such as protein denaturation and subsequent refolding, can be used. It should be noted, however, that the protein may not be stable after metal removal and may precipitate if the metal has a crucial structural role or constitutes the active site of the enzyme.

Validating the metal content in the protein sample (Step 5).—Establishing the metal content of the sample is an important step when working with proteins that bind metals. Techniques such as atomic absorption spectroscopy (AAS), extended X-ray absorption fine structure (EXAFS)⁶⁷, X-ray fluorescence (XRF), inductively coupled plasma-mass spectrometry (ICP-MS), or inductively coupled plasma-optical emission spectrometry (ICP-OES) are commonly used to determine the metal presence in a sample. AAS is most likely a more common instrument found across a variety of labs due to its lower cost compared to ICP-MS and ICP-OES (Box 3). These methods are useful in determining if a metal of interest was successfully introduced into or removed from a protein or for measuring metal content in prepared buffers. For other methods, see the ALTERNATIVE METHODS section. We recommend checking the metal content in the protein sample just after purification but before the removal of unwanted metals. If satisfied with the metal content of the protein sample, one may not need to proceed with the removal

of unwanted metals. After the removal of unwanted metals, the final metal content should be verified again. We suggest validating metal content for each protein sample because the metal content may change depending upon the chemicals used during sample preparation (different batches of salts and buffers). The result of the metal content verification should normally show that only the metal of interest is present in the protein sample or the protein is free of metals, in which case the metal of interest could be incorporated into the protein during the crystallization procedure outlined below. Alternatively, crystals can also be dissolved after X-ray data collection to determine the presence of a metal with a method of choice. For example, ICP-MS has been used to verify the presence of Ru in the crystals of a human serum albumin (HAS)-KP1019-like molecule adduct¹⁰².

Selecting the pH and a metal-compatible buffer (Steps 6–7).—Buffers are not innocuous, especially the so called Good's buffers¹⁰³. Some form soluble or insoluble metal complexes, which can alter the pH of the solution through the addition of protons to solution while forming the complex or through removing the buffer from solution via precipitation. Therefore, it is important to select a buffer for both sample preparation and crystallization that is compatible with the metal(s) being investigated. Selecting the best metal-compatible buffer requires knowledge of certain characteristics of the protein being studied. For instance, one should identify the optimal pH for the protein of interest (ideally close to the physiological pH) and determine if certain buffers are inhibitory to its function. Once the ideal pH of the experiment is known, compile a list of biological buffers with pKa values that are within one pH unit of the desired pH and do not inhibit the protein. Next, assess which of these buffers form metal-complexes with the metal(s) selected for analysis by looking at scientific literature investigating the properties of biological buffers and their stability constants with different metals^{64,104,105}, especially a recent review in which much of this information is compiled¹⁰⁶. If there is a buffer that does not form metal-complexes, select this one for further experiments. For instance, HEPES and MES are buffers that have low metal binding constants. If all of the buffers form metal complexes with the selected metal(s), choose the one(s) with the lowest stability constant(s) and proceed; remember that the equilibrium concentration of the metal in the solution can be lower than expected.

Selecting a reducing agent (Step 7).—Another aspect that may significantly influence the metal's binding to protein is the oxidation of the metal and/or the protein side chains, especially cysteines. We recommend ensuring a reducing environment in the sample by keeping all solutions in an anaerobic environment (e.g. anaerobic chambers) and/or adding reducing agents such as 2-mercaptoethanol (BME), dithiothreitol (DTT), tris(2-carboxyethyl)phosphine (TCEP), or sodium dithionite. The mentioned reagents differ in reducing capability. The strongest, sodium dithionite, can be used when working with metal ions of the lower oxidation state, such as Fe²⁺. TCEP is a stronger reducing agent compared to BME and DTT. It does not reduce metals commonly used in metal affinity chromatography (Ni, Co), but it is possible that TCEP could reduce some other metals (for example, Cu(II) to Cu(I))¹⁰⁷. TCEP is a commonly used reducing agent for protein crystallization since it is stable at room temperature for about 2 to 3 weeks in a typical crystallization experiment, compared to 1–2 days for sodium dithionite, 2–3 days for BME and 3–7 days for DTT.

CONSIDERATIONS FOR CRYSTALLIZATION

Choosing a crystallization strategy (Step 10).—Obtaining crystals of protein-metal complexes can pose significant challenges. There are three common strategies to approach this problem: Option A describes introducing metal to the protein before the crystallization experiment; Option B describes mixing the metal-free protein sample with a metal-containing crystallization condition; Option C describes soaking crystals of a metal-free protein with a metal solution. In Option A, the protein is combined with the metal prior to the crystallization experiment and the mixture is used to screen crystallization conditions. In some cases, it may not be necessary to add the metal if the protein has a very high affinity for the metal and if it was retained during purification as shown by the metal detection methods described above. Nevertheless, maintaining sufficient metal concentration in the solution is highly recommended, otherwise the bound metal may dissociate from the protein prior to or during the crystallization experiment. If the crystallization conditions are already known for the protein, Option B, in which the metal is added to the crystallization cocktail, is suitable. When the presence of metal provides a proper protein conformation prior to crystal formation, co-crystallization options A and B are advisable strategies. If co-crystallization does not produce satisfactory results, Option C can be tested by crystallizing the metal-free protein and then soaking the crystals with a metal solution.

Ensuring sufficient metal concentration (Step 10).—The concentration of the metal in the crystallization drop must be high enough to visualize the metal in the crystal structure by occupying the binding site(s) in every protein molecule in the crystal. If the protein-metal affinity is determined or estimated as suggested in the preliminary analysis (Step 2), we recommend having the metal concentration in the crystallization drop at least ten times higher than the metal-protein K_d . If the affinity is unknown, one can begin with 1–10 mM metal concentration. In some cases, mixing the protein with certain metals in high concentrations may cause immediate protein precipitation¹⁰⁸. In this case, we suggest adding chelators or preparing several protein-metal mixtures with increasing metal concentrations and screening all of them for crystal formation and presence of the metal in the structure. If crystals form in the presence of the metal, but the metal is not visible in the crystal structure or there is a partial occupancy of the metal-ion binding site, it is advisable to try co-crystallization with a higher metal concentration (even if some protein precipitation is observed upon metal addition) or additional soaking of protein-metal crystals (Step 10 Option C). During crystal soaking of the protein-metal complexes or metal-free protein crystals, the high local concentration of metal may cause crystal cracking or even dissolve the crystal. Therefore, we recommend beginning with a lower metal concentration and gradually increasing the metal concentration after the metal diffuses into the drop. If there are several drops with crystals on the same cover slip, we suggest adding a high metal concentration to one of the drops (this will most likely damage crystals in this particular drop) and connecting the remaining drops with channels made between the drops to allow for slow diffusion of highly concentrated metal to other drops. Keep in mind that if the concentration of the metal is relatively high, sites with low metal affinity and marginal physiological importance (or even non-specific sites) may be present in the resulting crystal structure¹⁰⁹. Therefore, it is important to take into account the information collected on the

preliminary analysis steps (e.g. binding affinity and residues involved in binding) when analyzing the structure to distinguish between sites of interest and crystallization artifacts.

Controlling the reducing environment (Step 10).—Many transition metals have more than one oxidation state available in physiological conditions. When working with a metal that may exist in multiple oxidation states, such as iron or copper, the metal can be kept in the reduced state by adding 1–10 mM reducing agent (such as TCEP) to the dialysis buffer. If a stronger reducing agent is needed, 10 mM sodium dithionite may be used as an oxygen scavenger. In addition to the type of metal, selecting a reducing agent also requires knowledge of the particular protein's redox potential since the protein can change the redox properties of a metal dramatically. For oxidation-sensitive metals and proteins, it is recommended to perform all work with solutions and crystallization inside anaerobic chambers, in which the oxygen concentration can be kept at different levels and should be adjusted to the requirements of a particular metal and protein^{95,96}. In a crystallized protein sample, measurements of the oxidation state of the metalloprotein should also be considered because changes in a redox state of the metal can impact the interpretation of metalloprotein crystal structures. The oxidation state can be assessed by single-crystal spectroscopy methods⁷⁹ such as EPR^{15,110}, UV-vis absorption microspectrophotometry, microspectrofluorimetry, microRaman, Resonance Raman, and Infrared Spectroscopy (IR)⁷⁹.

Controlling the pH (Steps 10).—The pH of the crystallization drop significantly influences metal binding to proteins regardless of the crystallization method used. pH may also be temperature dependent. Consequently, the pH must be monitored throughout the entire process of sample preparation and crystallization. We recommend adjusting the pH of the protein buffer, metal solution(s) used for co-crystallization and/or soaking, and the crystallization cocktail once the successful crystallization condition is found. It is important to remember that the pH of the crystallization cocktail may change upon addition of metal to the solution. Ideally, the optimal crystallization pH would be the one closest to the native pH of the *in vivo* protein environment. If the protein initially crystallizes at a pH away from the optimal level, one may consider optimizing the initial hits towards physiological pH. Protein-metal crystals at pH values than physiological pH should be interpreted with extra caution to ensure that the biological significance of the results obtained is relevant.

Flash-cooling crystals (Steps 11–12).—The crystallization condition is considered cryo-ready if flash-cooling does not lead to ice crystals around the protein crystal and does not damage the crystal in any other way. If the condition is not cryo-ready, slow dehydration by drying the harvested crystal against a relatively high concentration of salt (mother liquor or 1.0–1.5 M NaCl) or by careful removal of solvent around the crystal (crystal dehydrates during the manipulations) (Step 11 Option A) can be used^{20,111,112}. When available, high-pressure cryo-cooling is an excellent alternative¹¹³. If these methods do not work and an additional cryoprotectant is required to flash-cool the crystals, it is wise to use a cryoprotectant in which the metal solubilizes poorly (such as paratone-N or mineral oil) or has the same concentration of metal (Step 11 Option B). Flash-cooling should be performed in an anaerobic chamber for oxygen-sensitive proteins¹¹⁴.

DATA COLLECTION

Collecting the first fluorescence spectrum (at a high excitation energy) (Step 13).—X-ray fluorescence is the emission of characteristic secondary X-rays from the sample that was excited with X-rays of a single given energy. X-rays are ionizing radiation and by definition, may have sufficient energy to eject a core electron (located on an inner orbital) from an atom. Because the resulting electron structure is unstable, the hole that is created is filled by an electron from a higher orbital. The electron falling is accompanied by the release of a photon with an energy equal to the energy difference of the two orbitals involved, resulting in secondary radiation with a characteristic energy. Recording the fluorescence emission spectrum allows for the quick detection of particular elements in the crystal based on the presence of peaks with characteristic energy of the secondary X-rays. Each element has a clearly distinguishable signature of peaks; usually, the highest peak is used for identification. Since elements require different minimal energies to excite the atoms, the number of elements that can be detected by the spectrum depends on the excitation energy used. Hence, the first spectrum should be collected with a high excitation energy to allow for the detection of as many elements as possible. The typical energy on the selenium absorption K-edge (12,664 eV) is sufficient to detect most of the transition metals as well as Ca and K. Therefore, this spectrum can often be collected even before changing the energy of the synchrotron beam. The fluorescence spectrum is usually represented as a graph, plotting the number of detected photons of emitted fluorescence X-ray (y-axis) for each emission energy (x-axis). An example of collecting a fluorescence spectrum at the selenium absorption K-edge is shown in Figure 4a. A strategy and examples of collecting fluorescence spectra at high excitation energy other than selenium K-edge are described in the next paragraph (Step 14). Since the emitted energy is characteristic for each element, it is possible to use a standard table of X-ray emission lines¹¹⁵ to assign emissions to a particular element that was present in the sample and excited by the incident beam. Data collection programs on modern synchrotron stations can provide the user with spectra that have automatically assigned peaks for convenience. The detection limit depends in part on the instrument used to generate the spectrum, the detector used to collect the data, and the element that is being evaluated. Modern synchrotron stations permit trace-element analysis in the picogram range^{116–118}. Collecting fluorescence spectra is a quick experiment to perform, which can confirm the presence of the desired metal in the sample, and also may show the presence of contaminant metals in addition to the metal of interest.

Collecting the second and third fluorescence spectra (above and below the absorption edge) (Step 14).—If the first spectrum (as described in the section above) shows the presence of the metal of interest, it is recommended to proceed to the next step and change the energy of the beam. The second fluorescence spectrum, collected with the excitation energy on or slightly above (+10–20 eV) the metal absorption edge will show a similar graph to the first spectrum, although fewer elements may manifest on the second spectrum compared to the first one (Figure 4b, Supplementary Figure 2, top panels). For theoretical values of the absorption edges of the elements and plots for the scattering factor coefficients f and f' , the reader is referred to *The International Tables for Crystallography Volume C*¹¹⁹, online resources such as <http://skuld.bmsc.washington.edu/scatter/ASindex.htm>, and programs for data collection such as HKL-3000 that have these values

presented in the form of the periodic table. If the second fluorescence spectrum is collected at a high excitation energy close to the selenium absorption K-edge, it can serve the dual purpose of both the first and the second spectra (Supplementary Figure 2b–d, top panels). Under such circumstances, collecting the first fluorescence spectrum described in Step 13 is redundant and unnecessary. The third fluorescence spectrum, collected at the energy slightly below (–30–50 eV) the metal absorption edge will produce a graph in which the peak for the fluorescence of the metal of interest is non-existent (Figure 4b, Supplementary Figure 2, middle panels). The difference between the second and the third spectra will prove the presence of the metal in the sample and confirm that excitation energies were selected accurately (to be used in the following steps). Importantly, this method does not provide any insight regarding whether the metal directly interacts with the protein (i.e. the metal can be present in the solution in the crystal channels but not be bound to the protein).

Collecting the fluorescence scan (Step 15).—An X-ray fluorescence scan is the total emission photon count of characteristic secondary X-rays from the sample that was excited with X-rays of multiple varying steps of energy. Usually, the data collection software will present the raw photon count graph and the graphs for both f and f' versus excitation energy (see Figure 4c and Supplementary Figure 2, bottom panels for the examples). The scan is usually collected with excitation energy ~30 eV below and above the tabulated value of the metal absorption edge, which is usually enough to cover the whole absorption edge. The sufficient range of excitation energy depends on the absorption edge width of the metal in the sample (the energy difference between the f' inflection point and the f' maximum), which depends on the type of metal, its oxidation state, and its chemical environment^{120,121}. For instance, the absorption scan for Zn is usually performed in the energy range of 9630 eV to 9690 eV (Figure 4c). If the collected scan does not cover the whole edge (e.g. Hg L-III edge, Supplementary Figure 2c, bottom panel), the range of the excitation energy could be increased. The fluorescence scan can be used for additional verification of the metal identity because the absorption edge value (approximated as the inflection f' point; may differ from the theoretical value by up to 20 eV depending on the metal oxidation state and chemical environment)^{120,121} and the graphs for f and f' are characteristic for each element. The scan can help further optimize the energy for collecting X-ray diffraction data above the absorption edge (the maximum of f' , located at the top of the fluorescence scan), as well as the optimal energy for collecting X-ray diffraction data below the absorption edge (the highest energy below the absorption edge that gives only background fluorescence signal).

Collecting the X-ray diffraction data above and below the metal absorption edge (Step 16).—To determine and verify the identity and position of the metal in the crystal structure, it is important to collect X-ray diffraction data above and below the metal absorption edge. Collecting X-ray data on or slightly above (10–20 eV) the absorption edge of a metal will allow the characteristic anomalous signal to emerge on the anomalous electron density map, pinpointing the location(s) of the atom(s). Using the energy determined by the fluorescence scan in step 15 will help maximize the anomalous signal. Collecting X-ray data below the absorption edge of the same metal using the energy determined by the scan will show an absence (or significant decrease) of the anomalous map peak(s), proving both the identity and location of the metal in the structure. Because the

height of the anomalous peaks depends on the data quality and redundancy, it is important to have two datasets collected under the same conditions, including starting angle, frame width, expose dose, and number of frames. This is especially important for transition metals from the 5th and 6th periods, which have a significant value of f'' at exciting energies below the absorption edge (as a result, a decrease, but not disappearance of the anomalous peaks, is expected in the dataset collected below the absorption edge).

Radiation damage during X-ray data collection (Step 16).—Macromolecules containing transition metals are more susceptible to radiation damage than ‘native’ samples, due to the higher X-ray absorption by these metals. Therefore, pay special attention to avoid significant radiation damage because it can affect the quality of the datasets, which is important for comparing the anomalous signal between the datasets above and below the absorption edge. The radiation damage may also include reduction of the metal, especially if the metal is expected to be in a higher oxidation state when multiple oxidation states are possible in a protein-metal complex (e.g., Fe^{3+} , Co^{3+} , Cu^{2+}). Protein structures that may be affected by the X-ray-induced reduction of metals should be analyzed with great care. When metal reduction is possible, low radiation dose data collection and monitoring of the metal oxidation stage by single-crystal microspectrophotometry⁷¹ or by X-ray absorption spectroscopy¹²² are recommended.

Identifying the location(s) of the metal (Step 18).—Location(s) of the metal in the protein structure can be detected by the presence of a strong peak(s) on the anomalous electron density map. Disappearance (or significant decrease) of the peak(s) on the anomalous map calculated from the data collected below the metal absorption edge is the ultimate proof for metal identity. Alternatively, one can use dispersive [$F(\lambda_{\text{above}}) - F(\lambda_{\text{below}})$] Fourier difference map for metal assignment¹⁴ see the ALTERNATIVE METHODS section. If there is no anomalous data available or if the metal does not produce visible anomalous signal at available energies of the X-ray source, the location of the transition metal can be pinpointed by the presence of a high positive peak on the difference electron map and by a characteristic coordination pattern. Regardless of the availability of the anomalous data, it is important to make sure that the identity of the placed metal corresponds to the observed geometry and coordination bond distances. Table 1 presents the summarized features of ligands, distances with respective estimated standard deviations (ESDs), and geometries for commonly encountered metal binding sites.

REFINEMENT AND VALIDATION OF METAL-CONTAINING PROTEIN STRUCTURES

Modeling the metal-binding environment and specifying the metal-ligand restraints (Steps 19–23).—After the metal ion is located, the coordination ligands should be inspected to verify that they are chemically capable of coordinating the metal and the metal-ligand distances are within the acceptable range (Table 1). Coordination ligands other than protein atoms (e.g. water molecules, Cl^- , NH_2) should be modelled to complete the coordination sphere of the metal. Crystallographic programs used for model building and refinement of macromolecular structures are well optimized for handling macromolecules, whereas refinement and validation of small molecules, especially metal ions, usually require extra effort. During restrained refinement of organic compounds, it is usually sufficient to

use intra-residue bond lengths and bond angle definitions in the form of a CIF file in the CCP4 monomer library¹²³, although corrections have to be introduced in some cases³⁸. The restrained refinement of metal ions involves mostly inter-residue coordination bonds that need to be defined between the metal ion and the ligands in its binding environment. Otherwise, values from the CCP4 monomer library¹²³ are used by default in refinement programs. Unfortunately, these values are sometimes wrong^{49,50}, the default ESDs are often insufficient to properly restrain the distances, and changing the ESDs is problematic. Restraints used for refinement should be designed based on the residues and atom types of the coordinating ligands, the arrangement of ligands around the ion, the ion-ligand coordination bond distances, inter-coordination bond angles, and correct ESD values. This can be performed by downloading a REFMAC5¹²⁴ restraint file in the CIF format and the LINK records in the PDB format from the CMM server with the “REFMAC” option selected as the output format³³. For lower resolution refinement, applying restraints is imperative but should be treated with extreme caution because heavy restraints can enforce incorrect distances between the metal and ligand that may not be reflected in the electron density map. Small changes in metal-ligand distances may not be observed after adding restraints on the metal, especially for diffraction data collected at low resolution. General guidelines on applying restraints in metal binding site modeling and validation, together with some examples can be found in a recent publication¹²⁵.

Refining the metal B factor and occupancy (Steps 19–23).—The B-factor of the metal should be approximately equal ($\pm 14\%$) to the B-factors of the environment, defined as protein atoms located within 2–4 Å from the metal. The B-factor of a metal may appear higher if the metal is highly anisotropic but is refined isotopically. Often this situation is accompanied by characteristic positive and negative peaks on the F_o-F_c difference map (for an example, see <https://www.phenix-online.org/presentations/faq.pdf>). Therefore, anisotropic refinement should always be applied to transition metals or other heavy scatters in structures at around 2.5 Å resolution or better. Commonly used refinement programs such as REFMAC5 and PHENIX allow for the anisotropic refinement of a particular atom or a type of atom while refining the rest of the model isotopically or with Translation/Libration/Screw (TLS) parameterization. If the B-factor of the metal is higher than the B-factors of the environment ($>14\%$), even after the application of the anisotropic refinement, the occupancy of the metal should be decreased gradually until the B-factor of the metal is within the acceptable range. Some refinement programs can refine occupancies automatically, but we recommend manual adjustment because these automated protocols may not be optimized for the evaluation of the occupancy of metals. Big discrepancies between the B-factor of the metal and the environment may also be a sign of metal misidentification. Importantly, if the B-factors are heavily restrained, the B-factor of the metal may stay within the acceptable range even if the occupancy or the identity of the metal has been corrected. Often, this discrepancy will manifest on the F_o-F_c difference map with either a negative peak (if the occupancy of the metal has to be reduced or the identity of the metal should be corrected for a lighter element) or a positive peak (if the identity of the metal should be corrected for a heavier element) around the metal.

Validating the refined metal sites (Step 24).—If the output of CMM³³, Molprobity¹²⁶, or the PDB validation report¹²⁷ indicates the presence of sub-optimally modeled metal binding sites, it is recommended to correct the problem either by reexamining the metal identity or by remodeling the types and arrangement of ligands coordinating the metal. When reexamining the metal identity, it is still possible to make an educated guess based on the geometry and type of ligands around the metal-binding environment, even if the aforementioned experimental procedures failed to unambiguously identify the type and location of the metal. This can be achieved using the “alternative metal” section in the CMM tool, followed by manual examination. When remodeling the type and arrangement of ligands coordinating the metal, it is important to pay additional attention to metal binding sites with potentially incomplete coordination spheres, especially in structures determined at a relatively low resolution. This can be achieved using the “low resolution adjustment” option in the CMM to ensure that the effects of lower resolution have been taken into consideration during validation. Importantly, the validation procedure in CMM focuses on mononuclear metal-binding sites. Metal clusters with two or more metal centers are handled as individual metal-binding sites; their validation as a whole cluster has not been parameterized by CMM. In addition, CMM is not yet fully optimized for some metals that are rarely found in protein structures.

MATERIALS

REAGENTS

Macromolecule sample

- Solubilized protein sample at a sufficient concentration for crystallization.

▲ **CRITICAL** To avoid sample oxidation, oxygen-sensitive samples should be kept in an anaerobic chamber during all steps of the protocol. Detailed procedures for the special handling of redox-active metals are not described here.

▲ **CRITICAL** To avoid sample contamination with metals or other compounds, all chemical reagents should be high-grade and meet the specifications defined by the Committee on Analytical Reagents of the American Chemical Society (ACS) when possible. The catalog number is only listed when the purity of a compound reaches this recommended level. It is not critical to use particular reagents from the indicated suppliers. It is important to use reagents of Trace*SELECT* grade if available, especially for ICP or AAS analysis. For chemicals of Trace*SELECT* grade the blank values for metal traces are typically below 0.01 mg/kg (0.01 ppm).

Sample preparation

- Buffering agents: e.g., Trizma® base *TraceSELECT* (#42616, Sigma-Aldrich), HEPES, BIS-TRIS, MES, or citric acid *TraceSELECT* (#94068, Sigma-Aldrich).
- Reducing agents: e.g. 2-Mercaptoethanol (#97622, Sigma-Aldrich), DTT (#D5545, Sigma-Aldrich), or Sodium dithionite ! **CAUTION** BME is hazardous.

It is harmful to eyes and can cause skin irritation. Use it in a fume hood and wear a mask, protective eyewear and gloves.

! CAUTION DTT is hazardous. It causes skin, eye, and respiratory irritation. Wear protective clothing and eyewear.

! CAUTION Sodium dithionite is hazardous. It may catch fire, is toxic to aquatic life, and harmful if swallowed. Wear protective clothing and eyewear.

- Protease to remove affinity purification tag: e.g. rTEV protease, precision protease, thrombin **! CAUTION** Thrombin is hazardous. It is an irritant, may cause allergic respiratory and skin reactions. Wear protective clothing and eyewear.
- Metal chelators: e.g. EDTA (#431788, Sigma-Aldrich), EGTA (# 03777, Sigma-Aldrich), 1,10- phenanthroline (#131377, Sigma-Aldrich), TPEN (#4413, Sigma-Aldrich), or citric acid *Trace*SELECT (#94068, Sigma-Aldrich) **! CAUTION** EDTA is hazardous. It causes skin, eye, and respiratory irritation. Wear protective clothing and eyewear.
- **! CAUTION** Citric acid is an irritant. It causes skin, eye, and lung irritation. Wear protective clothing and eyewear.
- Chelex 100 resin (#1422822, Bio-Rad)

Crystallization

- Crystallization screens available on the market (e.g. from Hampton Research, Jena Biosciences, MiTeGen, Molecular Dimensions, Qiagen, Anatrace) or homemade - prepared according to needs.
- Polymers needed for crystal optimization: various polyethylene glycols.
- Salts needed for crystal optimization: e.g. sulfates, phosphates, formates.
- Organic compounds needed for crystal optimization: e.g., citrates, isopropanol, or malic acid **! CAUTION** Isopropanol is flammable. **! CAUTION** Malic acid is hazardous. Wear protective clothing and eyewear.
- Metal ion solution for co-crystallization and soaking; **! CAUTION** Solutions of some metals are hazardous. Always check the MSDS for a metal ion solution before working with it.

Data collection

- Cryoprotectant solutions: Parabar 10312 (#HR-643, Hampton Research), glycerol (#G7893, Sigma Aldrich), LV CryoOil™ (#LVCO-1, Mitegen), ethylene glycol (#85978, Sigma Aldrich), or (+/-)-2-Methyl-2,4- pentanediol (#HR2-627, Hampton Research) **! CAUTION** Ethylene glycol is hazardous. Wear protective clothing and eyewear. **! CAUTION** (+/-)-2-Methyl-2,4-pentanediol is hazardous. Wear protective clothing and eyewear.

Acid and base for pH adjustment

- Hydrochloric acid, *TraceSELECT* (#96208, Sigma-Aldrich); ! **CAUTION** Hydrochloric acid causes severe burns. Wear full protective clothing and protective eyewear. Use it in a fume hood.
- Sodium hydroxide solution, *TraceSELECT* (#13171, Sigma-Aldrich); ! **CAUTION** Sodium hydroxide is an irritant and corrosive. Wear full protective clothing and protective eyewear. Use it in a fume hood.
- Acetic acid, *TraceSELECT* (#07692, Sigma-Aldrich); ! **CAUTION** Acetic acid is an irritant and corrosive. Wear full protective clothing and protective eyewear. Use it in a fume hood.

EQUIPMENT

- Equipment for carrying out binding assays, such as ITC (Malvern, MicroCal iTC200; TA Instruments, Nano ITC) and TSA (Bio-Rad, CFX Connect™ Real-Time PCR Detection System; Applied Biosystem, QuantStudio 6 Flex Real-Time PCR System).
- Equipment for protein buffer exchange: dialysis tubing, beakers, magnetic stirrer with stirring bars, centrifugal filters.
- pH meter (accumet™ AE150 pH Benchtop Meter Bio Set; Mettler Toledo™, FE20 FiveEasy™ Benchtop pH Meter); pH-Indicator strips graduated in 1pH or 0.5pH units (Whatmann #2613991; MilliporeSigma #109584).
- Anaerobic chamber: rigid or soft anaerobic chambers manufactured by companies such as Coy Laboratory, Belle Technology or Shel Lab (BACTRON chambers), or others.
- Inductively coupled plasma-mass spectrometer (ICP-MS), or inductively coupled plasma-optical emission spectrometer (ICP-OES), sometimes referred to as an ICP-atomic emission spectrometer (ICP-AES) (Perkin Elmer, Sciex ELAN 6100 ICP/MS), Atomic absorption spectrometer (ThermoFisher Scientific, iCE™ 3300 AAS Atomic Absorption Spectrometer).
- Equipment for setting up crystallization plates: hanging or sitting drop crystallization plates (Hampton Research, 3 Well Crystallization Plate, Swisisci; Qiagen, EasyXtal 15-Well Tool) (96-, 48-, 24-, or 15-well plates), plate seals, cover slides, grease or greaseless screw-in crystallization support (depending on the type of crystallization plate used).
- Light microscope for crystal observation and manipulation (ZEISS SteREO Discovery.V8; Olympus SZX16 Stereo Microscope).
- Equipment for protein crystal flash-cooling: loops with bases for crystal mounting, vials, storage canes, magnetic wands, foam dewars for crystal manipulation; dewars for storing samples at cryogenic temperatures, dewars for transporting samples at cryogenic temperatures, pucks for crystal shipping with

accessories (puck holder, puck transfer tools, tongs for sample or puck manipulation).

- Tunable X-ray source (Advance Photon Source, European Synchrotron Radiation Facility, etc.) with the possibility of moving the beam energy to the absorption edge of the metal of interest. A suitable beamline can be searched through <http://biosync.sbk.org/>.
- Computer with the appropriate crystallographic software installed.

Software

- TEMSP server (<http://netalign.ustc.edu.cn/temsp/>)⁸⁷ or other servers for metal binding site prediction.
- HKL-3000 version 715 or later (<http://www.hkl-xray.com/hkl-3000>)¹²⁸, JBluIce-EPICS (current version as implemented on APS GM/CA-CAT beamlines)¹²⁹, or other software for data collection.
- HKL-3000 version 715 or later, or XDS version June 1, 2017 or later (<http://xds.mpimf-heidelberg.mpg.de/>)¹³⁰ for diffraction data processing.
- HKL-3000 version 715 or later, CCP4 version 7.0 or later (<http://www.ccp4.ac.uk/>)¹²³, or PHENIX version 1.12 or later (<https://www.phenix-online.org/>)¹³¹ for initial structure determination.
- COOT version 0.8.7 or later (<https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/>)¹³² or other software for displaying and manipulating atomic models of macromolecules and electron density maps.
- REFMAC5 version 5.5 or later (<http://www.ccp4.ac.uk/html/refmac5.html>)¹²⁴, PHENIX version 1.12 or later, or other refinement programs that support definition of metal-ligand restraints.
- CMM server (http://csgid.org/csgid/metal_sites)³³, PDB Validation Server (<http://validate.wwpdb.org/>)¹²⁷, or other servers for metal binding site validation.

Data

- Protein sequence
- Optional: structure of the metal-free protein or a homologous protein

PROCEDURE

Analyzing (potential) metal binding site(s) in the protein • TIMING 4–12 hours

1. Determine the potential metal that will bind to the protein, and/or the potential coordinating amino acids for each metal binding site of interest. Always try Option A first to search for the potential presence of a metal-bound structure(s) of homologous proteins. If a metal-free structure of the same protein is available, try Option B, either in addition or as an alternative to Option A. Proceed with Option B with a metal-free structure(s) of homologous proteins only when a

metal-free structure of the same protein is unavailable and Option A fails to identify any metal-bound structures(s) of homologous proteins (Figure 1).

A. Begin with a structure of a homologous protein with metal bound.

- i.** Identify structures of the same protein or homologous proteins that are known to have metal-binding sites. For example, to identify these structures by amino acid sequence comparison search, use tools such as “protein blast” on the NCBI BLAST website (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>), applying the “Protein Data Bank proteins (pdb)” database search option or “search by sequences” on the PDB website (<http://www.rcsb.org>).

▲ CRITICAL STEP When considering the potential presence of a single point mutation, structures of the same protein typically share >98% sequence identity with the query sequence. The global sequence identities with the structures of homologous proteins can vary. The identification of a homologous metal binding site depends more on the local conservation of the residues coordinating the metal and not just the global sequence identity.

- ii.** Identify metal-containing structures from the search results by looking at the small-molecule section of the PDB entry. If no such structure is available, skip the rest of Option A and proceed to Option B.
- iii.** Inspect the metal binding site in the homologous structures with a metal ion bound. Use tools such as CheckMyMetal³³ to verify the metal identity. Use programs such as COOT¹³² to visualize the metal binding site and corresponding electron density. Refer to Table 1 for the characteristics of the metal binding environment.
- iv.** Align sequences obtained in step i and the sequence of the protein of interest to determine the (potential) coordinating amino acids for each metal ion-binding site of interest.
- v.** Compare the identities of the coordinating amino acids forming the metal binding sites in the homologous model to those in the protein of interest. If the coordinating amino acids are conserved for a site, the similar metal binding site can be expected in the structure of interest.

B. Begin with a metal-free structure of the same or homologous protein.

- i.** Obtain the structure of a metal-free protein from a previous structure determination experiment, or fetch the structure of a

metal-free protein of interest or homologous protein from the PDB by following the instructions in Step 1 Option A step i-iv.

- ii. Predict metal binding site(s) using servers for metal binding site prediction, such as FINDSITE-metal⁹¹, TEMSP⁸⁷ or FEATURE¹³³ for Zn²⁺ (Ref.⁸⁷).
 - iii. Inspect the potential metal binding environment of the metal-free protein structure using a program like COOT¹³² for structure visualization. Manually examine the orientation of amino acids and select sites with several residues in close proximity that are able to coordinate a given metal. For example, the most common ligands for Zn²⁺ are the side-chain nitrogen of histidine residue, side-chain sulfur of cysteine residue, and carboxyl oxygen of aspartic acid or glutamic acid residues. The most common coordinating ligands and geometries are summarized in Table 1.
2. Estimate or experimentally determine the binding affinity of the metal to the protein based on either known data of a highly-homologous protein or by performing binding assays such as ITC or TSA as described in the corresponding section in experimental design. For detailed protocols on ITC and TSA, see Box 1 and Box 2 respectively.

Preparing a protein sample with unwanted metals removed • TIMING 1–2 days

1. (Optional) Remove the purification tag. If a His-tag is used and a cleavage site is constructed between the His-tag and the target protein, remove the His-tag from the protein sample using the appropriate protease before any crystallization experiment. For proteins that are unstable after His-tag cleavage, or if the His-tag is non-cleavable from a protein, it is recommended to either perform extended dialysis in the presence of a metal chelator or apply the protein sample to a Chelex 100 chelating ion-exchange resin (see **Step 4**). Tags that allow for affinity purification using metal-free resins, namely maltose binding protein (MBP) or glutathione S-transferase (GST) tags, may be left uncut, although their cleavage is preferable.

▲ CRITICAL STEP For immobilized metal affinity chromatography, immobilized Co²⁺, Cu²⁺, Ni²⁺, Zn²⁺, Ca²⁺, and Fe³⁺ can all be used to purify His-tagged proteins¹³⁴. If a metal affinity tag is not removed, there is a high likelihood that some of these metal ions will be present in the sample for all subsequent steps of this protocol.
2. (Optional) Remove unwanted or contaminant metals. This step is recommended in most cases. An alternative strategy is to skip this step, proceed with ICP experiments (**Step 5**), and only perform this step if ICP experiments detect metal contaminants. Remove the unwanted or contaminant metals from the sample using chelators via dialysis or Chelex 100 resin (Bio-Rad). During dialysis, add EDTA (typically 0.5 mM), 1,10-phenanthroline (typically 1–10 mM), citric acid

(typically 5 mM), or another chelator to the protein sample. The optimal concentration and type of chelator should be determined empirically, since the quantity of metal contaminants in the sample varies, and different types of chelators sequester metals with different affinities. Choose a dialysis membrane with a molecular weight cut-off (MWCO) 3–5 times smaller than the MW of the protein. Dialyze the protein in an approximately 200:1 (vol:vol) ratio of dialysis buffer to sample in the selected membrane with an appropriate chelator in order to reach equilibrium in 3–4 h¹³⁵. Exchange the dialysis buffer and repeat the dialysis process three times¹³⁶. If a shorter time is desired, a fourth buffer change is required¹³⁶.

▲ **CRITICAL STEP** Adjust the dialysis protocol to account for several factors that can cause variation in the speed of dialysis (e.g., temperature, type of chemicals in the dialysis buffer, mixing speed, volume of the sample and dialysis buffer, or size of the pores in the membrane).

▲ **CRITICAL STEP** Dialysis should be performed under anaerobic conditions for oxygen sensitive proteins.

? TROUBLESHOOTING

3. Check the amount of metal contaminants still present in the sample. To assess the amount of metal in the protein sample, use AAS, ICP-MS, or ICP-OES. When preparing samples for ICP experiments, use a validated standard metal solution¹³⁷ prepared in a similar buffer as the protein to assess sample matrix effects and to compare the metal content in samples with a known reference. For detailed protocols on ICP-OES or ICP-MS, see Box 3. Sample validation with ICP is critical to monitor the level of metal contaminants that may affect later steps of the protocol.

▲ **CRITICAL STEP** The lack of metal content in ICP results indicates the metal content is below the detection level but will not necessarily guarantee the sample is free of certain metal contaminants.

4. Choose the pH at which you would like to carry out all experiments, including crystallization. Usually, physiological pH for the protein under study is the preferred choice; a different pH may be chosen depending upon the question being addressed. Choose the pH according to the guidelines in the “CONSIDERATIONS FOR PROTEIN SAMPLE PREPARATION” section in the Introduction.

▲ **CRITICAL STEP** When crystals grow at a non-physiological pH, use extra caution when performing subsequent steps in this protocol and when analyzing and interpreting the biological significance of the resulting structural model.

5. Select a metal-compatible buffer that has a good buffering capacity at the chosen pH according to the experimental design section. Select a reducing agent according to the guidelines in the experimental design section. Prepare the buffer at a typical concentration in the range of 50–200 mM with the reducing agent at

a concentration either recommended by the experimental design guidelines or determined experimentally.

▲ CRITICAL STEP If possible, use high-grade chemicals and reagents that meet the ACS specifications on Analytical Reagents and are specially designed and tested for suitability in inorganic trace analysis. Select products with purity over 99% and metal trace impurities below 1 µg/kg (1 ppb).

▲ CRITICAL STEP When working with oxygen sensitive proteins to maintain desired (reduced) oxidation states of metal and protein, buffers should be degassed to remove oxygen and kept under an anaerobic atmosphere (e.g., a mix of H₂ and nitrogen (N₂) (5/95%) or N₂/carbon dioxide (CO₂)/H₂ (85/10/5 %)).

6. Adjust the pH of the buffer at the temperature that will be used for all subsequent experiments.
7. Prepare a metal-free protein sample from step 5 in the buffer from step 8 by exchanging the buffer using either dialysis or a centrifugal filter. Oxygen-sensitive proteins should be maintained under an anaerobic atmosphere.

Preparing and flash cooling metal-containing protein crystals • TIMING 2–180 days, or longer

1. Perform crystallization experiments of protein-metal complexes. There are three generalized strategies to prepare metal-containing protein crystals for X-ray crystallography experiments as outlined below (Options A to C). Option A describes incorporating metal into the protein before the crystallization experiment; Option B describes mixing the metal-free protein sample with a metal-containing crystallization condition; Option C describes soaking crystals of a metal-free protein with a metal solution. Refer to the guidelines in the “CONSIDERATIONS FOR CRYSTALLIZATION” section of the Introduction for details. These three options are not meant to be mutually exclusive. Due to the unpredictable nature of the crystallization experiment, one should attempt all strategies one after another or in parallel until diffraction-quality crystals are obtained. For any of the strategies, make sure the metal concentration in the crystallization drops is at least 10 times higher than the metal-protein K_d. If the affinity is unknown, start with 1–10 mM metal concentration.

▲ CRITICAL Check for the presence of potential metal chelators in protein crystallization conditions. These may compete with and weaken protein metal binding. If the K_d between metal and protein is similar to the K_d between metal and chelators in the buffer (e.g., Tris, citrate), and/or the concentration of the chelator greatly exceeds the concentration of metal, be sure to use an excessive amount in terms of the molar ratio of metal compared to chelator. For the examples of buffers that do not interfere with metal, see the “CONSIDERATION FOR PROTEIN SAMPLE PREPARATION” section in the Introduction.

▲ CRITICAL Crystallization should be performed under an anaerobic atmosphere for oxygen sensitive proteins.

A. (A) Preparing protein-metal complexes and crystallizing the complex

▲ **CRITICAL** In order to incorporate a metal into the protein sample, either dialysis or direct mixing of the metal solution with protein can be used. For dialysis, continue with steps 10(A)i-iii. For direct mixing of the metal solution with protein, proceed to steps 10(A)iv-v. Direct mixing is much faster than dialysis, but a rapid, local increase in metal concentration can cause protein precipitation. This method is suitable for very stable proteins, such as albumins.

- i. ***Metal incorporation by dialysis (Steps i-iii)*** Prepare a dialysis buffer by dissolving the metal in the buffer prepared in step 8 at the desired concentration. Adjust the pH afterwards.

▲ **CRITICAL STEP** Maintain a proper reducing environment for the dialysis buffer according (see “CONSIDERATIONS FOR CRYSTALLIZATION” in the Introduction).

- ii. Perform dialysis of the protein sample against the metal-containing dialysis buffer. To avoid potential protein precipitation, begin dialysis with a smaller concentration of metal (for example 1:10 metal to protein) and gradually increase the concentration over time (transfer the sample to a buffer with higher concentration of metal for example 1:2 then 1:1, next 10:1 ligand to protein and higher once every 2–4 hours). Metal concentration in each buffer and duration of dialysis should be empirically determined based on protein precipitation and/or metal incorporation in the protein. Continue with step 10(A)iii to validate the presence of desired metal, or proceed to step 10(A)vi.
- iii. (Optional) Perform ICP-MS or ICP-OES as described in Box 3 to verify the presence of the desired metal in the protein sample and metal to protein molar ratio. Use the dialysis buffer as a reference sample in order to assess the amount of metal bound to the protein. Proceed to step 10(A)vi.
- iv. ***Metal incorporation by direct mixing (Steps iv-v)*** Prepare a concentrated metal stock solution in the buffer from step 8. Adjust the pH afterwards.

▲ **CRITICAL STEP** Maintain a proper reducing environment for the metal stock solution (see “CONSIDERATIONS FOR CRYSTALLIZATION” in the Introduction).
- v. Mix the metal solution with the protein.

- vi. **Crystallization and optimization (Steps vi-ix)** Set up crystallization screens using a metal-containing protein sample from step 10(A)ii/10(A)iii or 10(A)v. Use standard crystallization screening techniques for preliminary screening of protein-metal complexes¹³⁸. Set up the crystallization drops by mixing the mother liquor with the metal-containing protein of interest.
- vii. Examine the drops to identify crystal-containing conditions as initial hits. Try to identify conditions with a final pH close to physiological pH for the protein of interest.
- viii. Attempt to optimize initial hits. If the final pH of the crystallization condition is not close to physiological, optimize the initial condition by varying the pH across the plate, including the physiological pH of interest. If the final pH is close to physiological and you wish to increase the metal concentration in the crystallization drop, incorporate metal into mother liquor and increase metal concentration in mother liquors from consecutive wells. Adjust the pH in the mother liquor, and set up the crystallization drops by mixing mother liquor with metal-containing protein sample or soak protein-metal crystals as described in step 10 Option C.

▲ CRITICAL STEP When crystals are obtained at non-physiological pH, extra caution is needed to perform subsequent steps in this protocol and to interpret the biological significance of the resulting structural model.
- ix. Further optimize the crystallization conditions to obtain crystals that can diffract to a better resolution.

B. Incorporating the metal during crystallization experiment setup

▲ CRITICAL In case crystallization conditions for a metal-free protein are available from previous experiments, immediately proceed to step 10(B)vi. Otherwise, use one of the two alternative approaches to estimate potential crystallization conditions for a metal-free protein: In order to determine protein crystallization conditions for the protein of interest from the PDB, proceed to steps 10(B)i-iii, in order to screen crystallization conditions for initial hits and optimize crystals, proceed to steps 10(B)iv-v.

- i. **Determining protein crystallization conditions using the PDB (Steps i-iii)** Find the 4-letter PDB code of a protein with >98% sequence identity to the protein of interest by querying the “Protein Data Bank proteins (PDB)” database using the protein sequence in FASTA format on the NCBI BLAST website. Refer to step 1(A)i-ii for a detailed procedure.

- ii. Locate the crystallization method, temperature, and conditions under the “Crystallization” section in the “Experiment” tab by searching the 4-letter PDB code found on the PDB website at <http://www.rcsb.org>.
- iii. Verify that protein crystals can be produced using the crystallization method, temperature and conditions reported in the PDB. Proceed to step 10(B)vi.
- iv. ***Determining protein crystallization conditions by screening (Steps iv-v)*** Use standard crystallization screening techniques (sitting-drop, hanging-drop vapor diffusion) for preliminary screening of the metal-free protein.
- v. Evaluate initial hits. Try to find conditions with a pH close to physiological pH for the protein of interest in a particular organism or cellular environment. If crystals are found at physiological pH with the desired metal as a component in the crystallization condition formulation, proceed to crystal flash-cooling in step 11. If crystals are found in any other pH, attempt to optimize conditions using a procedure similar to step 10(A)viii.

? TROUBLESHOOTING

- vi. ***Crystallization and optimization (Steps vi-viii)*** Mix the metal with mother liquor from known and/or optimized crystallization conditions in the desired ratio. Adjust the pH as needed. The known and/or optimized crystallization conditions are obtained from experiments carried out prior to exercising this protocol, step 10(B)iii or step 10(B)v.

▲ **CRITICAL STEP** Maintain a proper reducing environment for the mother liquor (see “CONSIDERATIONS FOR CRYSTALLIZATION” in the Introduction)

- vii. Screen for different metal concentrations by increasing the metal concentration in the mother liquor across the plate. Adjust the pH as needed.
- viii. Set up new crystallization drops by mixing the mother liquor containing increasing concentrations of metal with either the native protein of interest or the protein in complex with 2–5-fold excess metal depending on the expected binding affinity of metal to the protein.

C. Incorporating the metal after crystallization setup

- i. Prepare the mother liquor according to known and/or optimized crystallization conditions obtained from step

10(B)iii and step 10(B)v respectively. Grow the protein crystal of interest using the mother liquor.

- ii. Prepare the metal stock solution in the mother liquor prepared in step 10(C)i or buffer that includes a specific concentration of metal, and adjust the pH. Make sure the metal concentration in the mix is at least 10 times higher than the metal-protein K_d . If the affinity is unknown, start with 1–10 mM metal concentration. This step can be used to soak crystals of *apo*-protein or crystals of metal-protein complexes in order to increase metal concentration. The metal soaking approach is recommended for metals displaying lower affinities (μM - mM) toward the protein.

▲ CRITICAL STEP Maintain a proper reducing environment of the metal stock solution (see “CONSIDERATIONS FOR CRYSTALLIZATION” in the Introduction).

- iii. Soak crystals with the metal solution prepared in step 10(C)ii. Only use metal solutions with the pH adjusted.

▲ CRITICAL STEP Soaking crystals by adding metal salt powder directly to the drop is not recommended since metal diffusion can significantly change the pH of the crystallization drop beyond the buffering range as mentioned in “CONSIDERATIONS FOR CRYSTALLIZATION” in the Introduction.

- iv. If the protein of interest does not crystallize in a pH close to physiological pH, soak the *apo*-protein crystals with metal solutions that have the pH adjusted to the pH of the crystallization condition.

▲ CRITICAL STEP If the final pH of the crystallization condition is outside the physiological range, caution must be taken when analyzing the binding environment.

2. (Optional) If the crystallization condition is cryo-ready, proceed with step 12. Otherwise, prepare crystals for flash-cooling. This can be done either by dehydration (Option A) or by preparing the crystal in cryoprotectant solution (Option B). Try both approaches to see which one produces better results.

A. Dehydrate the crystal.

- i. Slowly dry the harvested crystal (5–10 minutes) against either the mother liquor or relatively high concentration of salt such as 1.0–1.5M NaCl; alternatively, carefully remove the solvent around the crystal^{20,111,112}. In these methods, the final

concentration of metal will be slightly higher due to water evaporation.

▲ CRITICAL STEP To maintain a certain oxidation state of a metal in a crystal of an oxygen sensitive protein, crystal flash-cooling should be performed in an anaerobic glove box equipped with microscope and liquid nitrogen dewar assembly for the direct-freezing of samples.

B. Prepare the crystal in cryoprotectant solution.

- i. Prepare a cryoprotectant solution that exhibits a slow metal diffusion rate (e.g., 100% paratone-N oil or its mixture with other oils such as paraffin)¹³⁹. If the metal diffuses into the cryoprotectant, add the same concentration of metal that was used for crystallization of the protein-metal complex into the cryoprotectant solution prior to use. Adjust the pH of the cryoprotectant solution so it will be compatible with the pH of the crystallization drop.

▲ CRITICAL STEP It is crucial to use a cryoprotectant solution that solubilizes metals with a negligible metal diffusion rate, such as oils. Water-based cryoprotectant solutions (e.g. 25% glycerol/mother liquor (vol./vol.) or 30% ethylene glycol/mother liquor (vol./vol.)) may have fast metal diffusion rates. It is possible that manipulating the crystal over long periods of time may cause metal diffusion from the crystal. Add metal to the cryoprotectant solution to prevent such diffusion.

- ii. Hold the metal-containing protein crystal in the cryoprotectant for different amounts of time (typically 1–5 minutes), or sequentially transfer the metal-containing protein crystal into a series of solutions containing increasing concentrations of precipitant (various molecular weight PEGs, organic solvents, inorganic salts).

3. Flash-cool the metal-containing protein crystal. If crystals are not harvested at the synchrotron facility, use pucks or vials to store and transport the samples for data collection on a tunable X-ray source at the synchrotron facility.

PAUSE POINT Flash-cooled crystals can be stored in liquid nitrogen for several months.

Collecting diffraction data and determining the initial model • TIMING 15 minutes-4 hours

1. Collect the first fluorescence emission spectrum with a high excitation energy (i.e. around selenium absorption K-edge 12664 eV). Examine the spectrum to verify the presence of the desired metal and to detect the presence of any other metals. Proceed to step 14 if the spectrum indicates the presence of metal. If the

spectrum indicates the absence of metal, continue with other crystals potentially containing metal until the presence of the desired metal is confirmed. If none of the protein crystals contain the metal, check for potential factors that may affect the presence of metals in the protein crystal in all previous steps 1–12. Identify other metals present on the spectrum and see if any of them are unwanted (not the desired metal or a buffer component). If a contaminant metal is identified, try to explain the source of it by checking for potential factors that may introduce contaminant metal in the protein crystal in all previous steps 1–12 and repeat the necessary steps to avoid the presence of the unwanted metal.

▲ **CRITICAL STEP** The fluorescence emission spectrum should be analyzed carefully. For example, peaks coming from other metal ions in the solution can also be visible in the spectrum, as well as small peaks coming from the metals constituting the loop pin (for examples, see Figure 4a–b).

▲ **CRITICAL STEP** The presence of metal in the protein crystal should be verified using step 13 and/or step 14 before all subsequent steps are performed.

? TROUBLESHOOTING

2. Collect the second fluorescence spectrum with the excitation energy on or 10–20 eV above the energy (for examples, see Figure 4b and Supplementary Figure 2, top panels) of the metal absorption edge; confirm the presence of the fluorescence signal coming from the metal. Collect a third fluorescence emission spectrum 30–50 eV below the metal absorption edge (for examples, see Figure 4b and Supplementary Figure 2, middle panels); confirm the absence of the fluorescence signal coming from the metal. For example, if the theoretical value of K-edge absorption energy for Zn equals 9659 eV, collect the fluorescence spectrum above the metal absorption K-edge using 9668 eV and below the metal absorption K-edge using 9618 eV. For theoretical values of the absorption edges of the elements and plots for f' and f'' see *The International Tables for Crystallography Volume C*¹⁹, online resources such as <http://skuld.bmsc.washington.edu/scatter/ASindex.html>, and programs for data collection such as HKL- 3000.
3. Collect a narrow fluorescence absorption scan with an excitation energy ~30 below and above the approximate metal absorption edge depending on the type of metal (for examples, see Figure 4c and Supplementary Figure 2, lower panels). Determine the energy of the metal absorption edge, approximated as the inflection f'' point, and use this value together with f and f'' plots to additionally verify the metal's identity by comparing it to those in Table 1 and other published tables (see Step 14 for references). Record the energy above the absorption edge that gives the highest f'' signal to ensure the presence of the highest peak on the anomalous electron density map that will be calculated in the step 18. Record the highest energy below the absorption edge that gives only background fluorescence signal to ensure a significant decrease of the peak on the anomalous electron density map.

▲ CRITICAL STEP Although the energies of absorption edges are defined, exact values of the energy may vary slightly depending on the oxidation state and local chemical environment of the metal^{120,121}. In addition, the values determined by the scan can deviate from the theoretical values if the beamline is not calibrated properly.

? TROUBLESHOOTING

4. Collect two complete diffraction datasets using the energies above and below the identified metal of interest absorption edge as determined in step 15. Use identical settings (except the beam energy) for both datasets, including starting angle, frame width, expose dose, and number of frames. The energies used for the fluorescence spectrum in step 14 may also be used for data collection, but the signal in the anomalous electron density map may be weaker. Avoid significant radiation damage.

▲ CRITICAL STEP If it is important to obtain the best resolution possible, it is recommended to collect a third dataset at a high energy like the selenium edge at 12680 eV.

5. Determine initial structures from the data collected above and below the metal absorption edge in step 16. This includes indexing, integrating, and scaling the diffraction images, structure solution, and model building. During scaling, keep Bijvoet Pairs separate (e.g. use the option “Use auto corrections” or “anomalous” in HKL-3000). Calculate the initial anomalous maps using the phases from the initial structure and the experimental structure factors.
6. Compare the initial structures determined at the energies above and below the metal absorption edge to unambiguously determine the location of the metal binding site with the strongest anomalous signal. Verify that the anomalous signal for the metal in the structure determined from the data collected above the metal absorption edge indeed disappears (or significantly decreases) in the structure determined from the data collected below the metal absorption edge. The decrease in anomalous peak is related to the lower energy provided during data collection, which was chosen to be insufficient for exciting the metal of interest.

▲ CRITICAL STEP Anomalous signals from elements other than the metal of interest can be present. For example, if data is collected on the iron absorption edge, anomalous signal from elements such as manganese, chloride, and sulfur can be seen in the structure.

▲ CRITICAL STEP After establishing the presence of the desired metal in the crystal sample, a diffraction experiment above and below the metal absorption edge is critical for finding the exact position of the metal in the unit cell¹⁴⁰.

? TROUBLESHOOTING

Refining the structures and characterizing metal binding sites • TIMING 3–5 days

1. Refine the structure and obtain the optimized phase information from the refined model. Update the $2F_o - F_c$, $F_o - F_c$, and the anomalous maps. Identify all possible locations of the metal using the updated anomalous map. Model the metal-binding environment and refine the metal binding sites according to the guidelines in “REFINEMENT AND VALIDATION OF METAL-CONTAINING PROTEIN STRUCTURES” section in the Introduction.

▲ **CRITICAL STEP** For medium to high-resolution structures (resolution better than 2.5 Å), anisotropic refinement should be applied to heavy scatters such as transition metals.

? TROUBLESHOOTING

2. Examine each metal binding site and specify which residues and small molecules coordinate the metal. Verify that the coordinating ligands are chemically capable of coordinating the metal of interest⁴⁹.
3. Check if the geometry of metal coordination is allowed for the metal (Table 1) and all inner-sphere coordinating ligands are visible in the electron density maps and modelled to complete the coordination sphere for the particular geometry. If needed and possible, model additional ligands (e.g. water molecules, Cl^- , NH_2) to complete the coordination sphere.

▲ **CRITICAL STEP** The full metal coordination sphere may be difficult to determine for low-resolution structures, binding sites located in disordered parts of structures, metals with very elastic coordination (K, Na), or sites that bind metal with weak affinities in the micromolar range.

4. Inspect and verify that the metal-ligand distances are within the allowable range. For example, the coordination bond between Cu and O is around 2.0 Å, whereas that between K and O is around 2.8 Å (Table 1).

? TROUBLESHOOTING

5. Define restraints for refinement according to the guidelines in the experimental design and using the parameters from steps 20–22. Copy the REFMAC5 restraint file in the CIF format and the LINK records in the PDB format from the file generated by the CMM server with the “REFMAC” output format selected. Alternatively, create LINK records using COOT¹³² software and manually correct the distances in the LINK records in the resulting .pdb file if needed. Use restraints during refinement to keep the proper geometry of the metal binding site and reasonable distances between the metal and ligands. Apply restraints on distances and angles for refinement in programs such as COOT¹³², REFMAC5¹²⁴ or PHENIX¹³¹ until a satisfactory metal binding site in the structure is achieved.

▲ **CRITICAL STEP** Small changes in metal-ligand distances may not be observed after adding restraints on the metal, especially if the diffraction data

were collected at a low resolution. Applying restraints is obligatory for low-resolution data but should be treated with extreme caution.

? TROUBLESHOOTING

6. Validate the refined metal binding sites using tools tuned toward metal binding analysis in proteins, such as the CheckMyMetal Server (CMM)³³, or servers that allow for ligand validation, such as the PDB Validation Server¹²⁷. Proper validation may pinpoint potential issues with metal site modeling and analyses. Refer to the “REFINEMENT AND VALIDATION OF METAL-CONTAINING PROTEIN STRUCTURES” section in the Introduction for additional details on features and limitations of using CMM for metal binding site validation.

? TROUBLESHOOTING

See Table 2 for troubleshooting guidelines.

• TIMING

The timing used in this protocol assumes that the macromolecule sample has been prepared with a purity and concentration suitable for crystallization. Discussion of the timing for the preparation of macromolecule samples suitable for crystallization is beyond the scope of this protocol. In extreme cases, this preparation can take many years of efforts.

Steps 1–2, Analyze (potential) metal binding site(s) in the protein ~ 4–12 hours

Steps 3–9, Prepare a protein sample with unwanted metals removed ~ 1–2 days

Steps 10–12, Prepare and flash cool metal-containing protein crystals ~ 2–180 days, or longer

The timing used for the crystallization of a metal-containing crystal typically ranges from a few days to a few months. It varies depending on many factors, such as the knowledge of prior crystallization conditions for metal-free proteins. Obtaining diffraction-quality metal-containing crystals can take many years of efforts in extreme cases.

Steps 13–18, Collect diffraction data and determine the initial model ~ 15 minutes–4 hours

The timing used for data collection varies by many factors including the brightness of the beamline, the size of the crystal, facility to tune the wavelength, the hardware and software setup of the workstation at the beamline etc. In the ideal situation, it takes around 5 minutes to perform various fluorescence spectra and scans, 5 minutes to set up the crystal and collect the datasets, and 5 minutes to determine the initial model. However, this process may take much longer, for example, when multiple crystals need to be screened for the optimal result.

Steps 19–24, Refine the structures and characterize metal binding sites ~ 3–5 days

ANTICIPATED RESULTS

One should be able to unambiguously determine the identity and location of a transition metal in a protein crystal structure and, in many cases, the ligands that coordinate the metal by following the above protocol. This protocol was successfully applied to the structural

characterization of metal binding sites in several different proteins, including the recently published crystal structures of horse and human serum albumin in complex with zinc⁵¹. To illustrate what one can expect for selected steps of the protocol, we present our results for four different transition metal-binding proteins as examples in the sections below (see also Table 3).

Steps 1–9: Analyzing (potential) metal binding site(s) in the protein and preparing a protein sample with unwanted metals removed.

Three different projects are used to illustrate the expected outcomes in steps 1, 2, and 6 (Table 3). We used equine serum albumin (ESA) to illustrate the expected outcome of preliminary analysis in step 1 (PDB ID: 5IJ5)⁵¹, STM1931 from *Salmonella typhimurium* to illustrate thermofluor shift assays (TSA) in step 2 (PDB ID: 4K2H), and a molybdenum cofactor containing chaperone for steroid C25 dehydrogenase from *Sterolibacterium denitrificans* to illustrate ICP-OES in step 6¹⁴¹ (Table 3).

In step 1, we expected to observe a series of specific amino acid residues that are known to interact with certain metal ions, such as those found in zinc fingers⁴⁴ or interacting with hemes¹⁴² if we had correctly analyzed the potential metal-binding site. The location of the metal binding site and residues involved in metal coordination must be confirmed via a crystallographic experiment. For example, the location of the zinc atom and the residues involved in zinc coordination in the major zinc binding site on HSA were deduced based on the interpretation of EXAFS data for both wild-type and mutant HSA¹⁴³. The proposed model correctly identifies the position of the major zinc binding site, but the details of zinc coordination were further revealed by X-ray crystallography. The crystal structure differs from the proposed model in the absence of the Asn-99 sidechain and the backbone carbonyl oxygen of His-247 in zinc coordination in the major zinc binding site on HSA. This difference could be explained by the ~ 2 Å shift of zinc position in the proposed model when compared to the zinc position in the crystal structure⁵¹.

In step 2, binding assays such as ITC or TSA are expected to characterize the formation of a macromolecular-metal complex using information about the strength of the interaction between a protein and metal ion, the influence of a metal on protein stability, and an estimate of the number of metal-binding sites present in the protein. For example, the protein STM1931 from *S. typhimurium* shows higher stability in the presence of Zn²⁺ in a TSA experiment (Figure 3).

In step 6, the chaperone protein responsible for molybdenum cofactor loading into molybdoenzymes was overexpressed under aerobic conditions and purified by immobilized metal affinity chromatography (on Ni-NTA resin). The presence of the molybdenum cofactor in the chaperone protein was analyzed by ICP-OES (Supplementary Table 3). The analysis shows that the molybdenum content in the sample is below the detection limit of ICP-OES, suggesting the absence of the molybdenum cofactor in the protein. The results are in accordance with current knowledge that the cofactor is produced only under anaerobic conditions¹⁴¹. Additionally, the analysis shows Ni²⁺ ions are present in a protein sample eluted from a Ni-NTA resin (most likely bound to the His-tag on the N-terminus of the protein)¹⁴¹.

Steps 10–12: Preparing and flash-cooling metal-containing protein crystals

The crystallization strategy described in procedure step 10 Option B was employed to crystallize the HSA-Zn²⁺ complex in the presence of 0.5 mM ZnCl₂ (Ref.⁵¹) (Supplementary Table 4). Crystallization conditions contained 23% (wt./vol.) PEG 1500; therefore, crystals were considered cryo-ready, and no additional cryoprotectant was needed before flash-cooling in liquid nitrogen. The crystallization strategies described in procedure step 10 Option B and Option C were combined to produce ESA crystals with different concentrations of zinc. Crystals of ESA-Zn²⁺ complexes were obtained by co-crystallization with 2.5 or 50 mM ZnCl₂ (Ref.⁵¹) (Supplementary Table 4) to obtain a final zinc concentration of 10, 15, and 30 mM. Crystals were harvested and placed in 100% paratone-N solution or a mixture of 50% vol./vol. paratone-N and 50% vol./vol. mineral oil for about 2 minutes. Manual manipulation of crystals in the cryoprotectant solution allowed for the removal of excess water before flash-cooling in liquid nitrogen. We successfully crystallized and flash-cooled both HSA and ESA in complexes with Zn²⁺ (Ref.⁵¹).

Steps 13–18: Collecting diffraction data and determining the initial model

The described procedure was applied to the data collection of serum albumin in complex with Zn²⁺. For a detailed description and overview of results, please refer to the original publication⁵¹. Here, we present another example of results obtained from an improved protocol of X-ray fluorescence analysis and data collection of dihydroorotase from *Yersinia pestis* CO92 (UniProt accession code: Q8ZFU4) (Table 3).

To check which metals were present in the sample, we collected a fluorescence spectrum around the selenium absorption K-edge at 12664 eV according to step 13 in the procedure section. The fluorescence spectrum verified that zinc was present in the sample with its characteristic K α emission energy around 8600 eV (Figure 4a). The peak at around 12600 eV corresponds to the energy of the incident X-ray beam. Additionally, peaks at 3300 eV and 6400 eV were visible on the fluorescence spectrum. The peak at 3300 eV may correspond to the K α emission energy of potassium, which was present at a high concentration in the buffer. The small peak at 6400 eV may correspond to the K α emission energy of iron present in the pin that supports the loop containing the protein crystal.

We collected X-ray fluorescence spectra at excitation energies below (9618 eV) and above (9668 eV) the theoretical values of zinc absorption K-edge according to step 14 in the procedure section. The peaks at around 9600 eV correspond to the energies of the incident X-ray beams at 9668 eV and 9618 eV in the two spectra (Figure 4b). The characteristic K α emission energy of zinc at around 8600 eV was visible in the X-ray fluorescence spectrum recorded at an energy of 9668 eV but absent in the X-ray fluorescence spectrum recorded at an energy of 9618 eV. The disappearance of the K α zinc emission peak in the fluorescence spectrum collected at the excitation energy below the zinc absorption edge (9618 eV) provides evidence that the zinc was responsible for the anomalous scattering. Additional peaks were also visible at around 3300 eV and 6400 eV, which were similar to those observed in the first fluorescence spectrum collected around the selenium absorption K-edge at 12664 eV.

X-ray fluorescence absorption scans were performed on the crystal according to step 15 in the procedure section. X-ray absorption scans were collected in the zinc's K-edge region with energy ranging between 9630 eV and 9690 eV (Figure 4c). The value of the observed absorption edge at 9663 eV (1.283 Å) verified the presence of Zn²⁺ in the sample (theoretical value 9659 eV). The energy above the absorption edge in the fluorescence scan that gives the highest f'' signal was located within the energy range of 9665–9671 eV. This confirmed that data collection at the energy used to collect the second fluorescence spectrum (9668 eV) would give the optimal anomalous signal above the absorption edge. The highest energy below the absorption edge in the fluorescence scan that gives only a background fluorescence signal is near 9650 eV. This confirmed that the data collection at the energy used to collect the third fluorescence spectrum (9618 eV) would ensure the absence of the peak on the anomalous electron density map. Therefore, an energy of 9618 eV was selected as being below the zinc absorption K-edge and 9668 eV as being above the zinc absorption K-edge for data collection.

The fluorescence spectrum unambiguously determines the identity of the metal but not its location. Unambiguous determination of both the identity and location of the metal requires the collection of diffraction datasets at energies above (9668 eV) and below (9618 eV) the zinc absorption K-edge for the same crystal. Examination of a zinc binding site in the initial model reveals the presence of the anomalous signal in the data collected above the absorption edge at 9668 eV and the absence of the anomalous signal in the data collected below the absorption edge at 9618 eV, proving that anomalous scattering was caused by the presence of Zn²⁺ (Figure 4d).

For other examples of fluorescence data collected for ESA complexed with different metals see Supplementary Figure 2.

Steps 19–24: Refining the structures and characterizing the metal binding sites

The use of restraints for metal binding site refinement is demonstrated by the major zinc binding site on ESA featuring three amino acid residues and one water molecule coordinating metal in a tetrahedral geometry (Figure 5, Table 1). We added Zn²⁺ in the middle of the anomalous peak and introduced LINKs between the metal and coordinating atoms using COOT software. The distance restraints between the metal and its ligands were further manually adjusted in the .pdb file by changing the distance in the LINK record (Figure 5b). After several rounds of refinement, correct distances in the range of 1.9 Å–2.2 Å for Zn–O_{ASP} and Zn–N_{HIS} and 2.1 Å for Zn–water were observed (Figure 5c). The use of the CheckMyMetal (CMM)³³ server for metal binding site validation is demonstrated by the crystal structure of zinc bound to ESA (PDB code: 5IIH)⁵¹. In terms of geometry, CMM reported zinc in a tetrahedral geometry with the CMM parameter *gRMSD* of 10.3 degrees and no missing vertices (both in a green background), which indicates an acceptable range for geometry in the CMM algorithm (Figure 6). In terms of ligands, zinc is coordinated by two histidine residues, aspartate with bidentate coordination, and by a water molecule, with Zn–O distances in the range of 1.98 Å–2.01 Å and Zn–N distances at 2.07 Å. These ligands are common for zinc, and the distances are within the acceptable distance range (Table 1). Moreover, all CMM parameters regarding valence, nVECSUM, Occupancy, and B-factor

agreement are in a green or yellow background, indicating the copper-binding site was modeled correctly (Figure 6).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

This work was supported with Federal funds from the National Institute of General Medical Sciences by the grant numbers GM117325 and GM117080 and NIH BD2K grant HG008424, as well as from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, Department of Health and Human Services under Contract Nos. HHSN272201200026C and HHSN272201700060C. E.N. was also supported by the Foundation for Polish Science (FNP) and received funding from the Marian Smoluchowski Krakow Research Consortium—a Leading National Research Centre KNOW supported by the Polish Ministry of Science and Higher Education. We thank Joanna Lipowska for providing the fluorescence and diffraction data for dihydroorotase from *Yersinia pestis* CO92. We thank Mateusz P. Czub for providing thermofluor shift data for STM1931 protein from *Salmonella typhimurium*. We thank Randy Alkire from the Structural Biology Center at Argonne National Laboratory for providing the fluorescence spectrum data for the zinc foil. We thank David R. Cooper and Marcin Cymborowski for help in X-ray fluorescence data collection and interpretation. We thank Marek Grabowski, Wen-Shyong Tzou, and Barat Venkataramany for valuable discussions and critically reading the manuscript.

REFERENCES

1. Röttschke O, Lau JM, Hofstätter M, Falk K & Strominger JL A pH-sensitive histidine residue as control element for ligand release from HLA-DR molecules. *Proc. Natl. Acad. Sci. U. S. A* 99, 16946–50 (2002). [PubMed: 12471156]
2. Holm RH, Kennepohl P & Solomon EI Structural and Functional Aspects of Metal Sites in Biology. *Chem. Rev* 96, 2239–2314 (1996). [PubMed: 11848828]
3. Pyle AM Metal ions in the structure and function of RNA. *J. Biol. Inorg. Chem* 7, 679–690 (2002). [PubMed: 12203005]
4. Potter JD, Sheng Z, Pan B & Zhao J A Structural Role for the Ca²⁺ - Mg²⁺ Sites on Troponin Regulation of Muscle Contraction. *J. Biol. Chem* 270, 2557–2562 (1995). [PubMed: 7852318]
5. Pace N & Weerapana E Zinc-Binding Cysteines: Diverse Functions and Structural Motifs. *Biomolecules* 4, 419–434 (2014). [PubMed: 24970223]
6. McCall KA, Huang C & Fierke CA Function and Mechanism of Zinc Metalloenzymes. *J Nutr* 130, 1455–1458 (2000).
7. Solomon EI et al. Copper active sites in biology. *Chem. Rev* 114, 3659–3853 (2014). [PubMed: 24588098]
8. MacPherson IS & Murphy MEP Type-2 copper-containing enzymes. *Cell. Mol. Life Sci* 64, 2887–2899 (2007). [PubMed: 17876515]
9. Mccall KA, Huang C-C & Fierke CA Zinc and Health: Current Status and Future Directions Function and Mechanism of Zinc Metalloenzymes 1. *J. Nutr* 130, 1437–1446 (2000).
10. Lyons T & Eide D Transport and storage of metal ions in biology. *Univ. Sci. Books* 2006, 57–78 (2007).
11. Farrell N Metal Complexes as Drugs and Chemotherapeutic Agents in *Comprehensive Coordination Chemistry II* (eds. McCleverty JA and Meyer TJ) 9, 809–840 (Elsevier Science, 2003)
12. Williams CJ, Whitehouse JMA & Medical B Cis-platinum: a new anticancer agent. *Br. Med. J* 1, 1689–1691 (1979). [PubMed: 380759]
13. Crack JC, Green J, Thomson AJ & Le Brun NE in (eds. Fontecilla-Camps JC & Nicolet Y) 33–48 (Humana Press, 2014). doi:10.1007/978-1-62703-794-5_4
14. Volbeda A in (eds. Fontecilla-Camps JC & Nicolet Y) 189–206 (Humana Press, 2014). doi: 10.1007/978-1-62703-794-5_13

15. Bowman SEJ, Bridwell-Rabb J & Drennan CL Metalloprotein Crystallography: More than a Structure. *Acc. Chem. Res* 49, 695–702 (2016). [PubMed: 26975689]
16. Brink A & Helliwell JR New leads for fragment-based design of rhenium/technetium radiopharmaceutical agents. *IUCrJ* 4, 283–290 (2017).
17. Helliwell JR New developments in crystallography: exploring its technology, methods and scope in the molecular biosciences. *Biosci. Rep* 37, BSR20170204 (2017).
18. Tanley SWM, Schreurs AMM, Kroon-Batenburg LMJ & Helliwell JR Re-refinement of 4g4a: room-temperature X-ray diffraction study of cisplatin and its binding to His15 of HEWL after 14 months chemical exposure in the presence of DMSO. *Acta Crystallogr. Sect. F, Struct. Biol. Commun* 72, 253–254 (2016). [PubMed: 26948967]
19. Tanley SWM, Schreurs AMM, Kroon-Batenburg LMJ & Helliwell JR Re-refinement of 4xan: hen egg-white lysozyme with carboplatin in sodium bromide solution. *Acta Crystallogr. Sect. F, Struct. Biol. Commun* 72, 251–252 (2016). [PubMed: 26919531]
20. Russo Krauss I et al. Principles and methods used to grow and optimize crystals of protein-metalloadducts, to determine metal binding sites and to assign metal ligands. *Metallomics* (2017). doi:10.1039/C7MT00219J
21. Frier JA & Perutz MF Structure of human foetal deoxyhaemoglobin. *J. Mol. Biol* 112, 97–112 (1977). [PubMed: 881729]
22. Reeke GNJ, Becker JW & Edelman GM The covalent and three-dimensional structure of concanavalin A. IV. Atomic coordinates, hydrogen bonding, and quaternary structure. *J Biol Chem* 250, 1525–1548 (1975). [PubMed: 1112816]
23. Berman HM et al. The Protein Data Bank. *Nucleic Acids Res* 28, 235–42 (2000). [PubMed: 10592235]
24. Gatzeva-topalova PZ, Warner LR, Pardi A & Carlos M Metalloproteomics: Forward and Reverse Approaches in Metalloprotein Structural and Functional Characterization. *Curr Opin Chem Biol* 18, 1492–1501 (2011).
25. Hall JF et al. Towards the high-throughput expression of metalloproteins from the *Mycobacterium tuberculosis* genome. *J. Synchrotron Radiat* 12, 4–7 (2005). [PubMed: 15616357]
26. Acton AQ Metalloproteins—Advances in Research and Application: 2013 Edition (ScholarlyEditions, 2013).
27. Nicolini C Molecular Manufacturing (Springer Science & Business Media, 2013).
28. Maret W Metalloproteomics, metalloproteomes, and the annotation of metalloproteins. *Metallomics* 2, 117–125 (2010). [PubMed: 21069142]
29. Shi W et al. Characterization of metalloproteins by high-throughput X-ray absorption spectroscopy. *Genome Res* 21, 898–907 (2011). [PubMed: 21482623]
30. Macedo S et al. Can soaked-in scavengers protect metalloprotein active sites from reduction during data collection? *J. Synchrotron Radiat* 16, 191–204 (2009). [PubMed: 19240331]
31. Metalloproteins: Theory, Calculations, and Experiments (CRC Press, 2015).
32. Domagalski MJ et al. in *Structural Genomics* 297–314 (Springer Protocols, 2013).
33. Zheng H et al. Validation of metal-binding sites in macromolecular structures with the CheckMyMetal web server. *Nat. Protoc* 9, 156–170 (2014). [PubMed: 24356774]
34. *Methods in Enzymology*, Volume 226: Metallobiochemistry, Part C (Academic Press, 1993).
35. *Methods in Enzymology*, Volume 158: Metallobiochemistry, Part A (Academic Press, 1988).
36. *Methods in Enzymology*, Volume 227: Metallobiochemistry, Part D (Academic Press, 1993).
37. Papageorgiou AC & Mattsson J Protein structure validation and analysis with X-ray crystallography. *Methods in Molecular Biology* 1129, (2014).
38. Shabalin I, Dauter Z, Jaskolski M, Minor W & Wlodawer A Crystallography and chemistry should always go together: A cautionary tale of protein complexes with cisplatin and carboplatin. *Acta Crystallogr. Sect. D Biol. Crystallogr* 71, 1965–1979 (2015). [PubMed: 26327386]
39. Zimmerman MD, Proudfoot M, Yakunin A & Minor W Structural insight into the mechanism of substrate specificity and catalytic activity of an HD-domain phosphohydrolase: the 5'-deoxyribonucleotidase YfbR from *Escherichia coli*. *J. Mol. Biol* 378, 215–26 (2008). [PubMed: 18353368]

40. Minor W et al. Crystal structure of soybean lipoxygenase L-1 at 1.4 Å resolution. *Biochemistry* 35, 10687–701 (1996). [PubMed: 8718858]
41. Tomchick DR, Phan P, Cymborowski M, Minor W & Holman TR Structural and Functional Characterization of Second-Coordination Sphere Mutants of Soybean Lipoxygenase-1. *Biochemistry* 40, 7509–7517 (2001). [PubMed: 11412104]
42. Chruszcz M, Wlodawer A & Minor W Determination of protein structures—a series of fortunate events. *Biophys. J* 95, 1–9 (2008). [PubMed: 18441029]
43. Saikatendu KS et al. Structure of a conserved hypothetical protein SA1388 from *S. aureus* reveals a capped hexameric toroid with two PII domain lids and a dinuclear metal center. *BMC Struct. Biol* 6, 27 (2006). [PubMed: 17187687]
44. Niedzialkowska E et al. Molecular basis for phosphospecific recognition of histone H3 tails by Survivin paralogues at inner centromeres. *Mol. Biol. Cell* 23, 1457–66 (2012). [PubMed: 22357620]
45. Majorek KA, Kuhn ML, Chruszcz M, Anderson WF & Minor W Structural, functional, and inhibition studies of a Gcn5-related N-acetyltransferase (GNAT) superfamily protein PA4794: A new C-terminal lysine protein acetyltransferase from *Pseudomonas aeruginosa*. *J. Biol. Chem* 288, 30223–30235 (2013). [PubMed: 24003232]
46. Luo H-B et al. Crystal structure and molecular modeling study of N-carbamoylsarcosine amidase Ta0454 from *Thermoplasma acidophilum*. *J. Struct. Biol* 169, 304–11 (2010). [PubMed: 19932181]
47. Murphy TA et al. Crystal structure of *Pseudomonas aeruginosa* SPM-1 provides insights into variable zinc affinity of metallo-beta-lactamases. *J. Mol. Biol* 357, 890–903 (2006). [PubMed: 16460758]
48. Segreaves EN et al. Kinetic, spectroscopic, and structural investigations of the soybean lipoxygenase-1 first-coordination sphere mutant, Asn694Gly. *Biochemistry* 45, 10233–42 (2006). [PubMed: 16922498]
49. Zheng H, Chruszcz M, Lasota P, Lebioda L & Minor W Data mining of metal ion environments present in protein structures. *J. Inorg. Biochem* 102, 1765–76 (2008). [PubMed: 18614239]
50. Zheng H, Shabalin IG, Handing KB, Bujnicki JM & Minor W Magnesium-binding architectures in RNA crystal structures: Validation, binding preferences, classification and motif detection. *Nucleic Acids Res* 43, 3789–3801 (2015). [PubMed: 25800744]
51. Handing KB et al. Circulatory zinc transport is controlled by distinct interdomain sites on mammalian albumins. *Chem. Sci* 7, 6635–6648 (2016). [PubMed: 28567254]
52. Gabbiani C et al. Protein metalation by metal-based drugs: reactions of cytotoxic gold compounds with cytochrome c and lysozyme. *J Biol Inorg Chem* 17, 1293 (2012). [PubMed: 23132507]
53. Merlino A, Marzo T & Messori L Protein Metalation by Anticancer Metallodrugs: A Joint ESI MS and XRD Investigative Strategy. *Chem. - A Eur. J* 23, 6942–6947 (2017).
54. Allen JP *Biophysical Chemistry* (Wiley-Blackwell, 2008).
55. Ostendorp T, Diez J, Heizmann CW & Fritz G The crystal structures of human S100B in the zinc- and calcium-loaded state at three pH values reveal zinc ligand swapping. *Biochim. Biophys. Acta - Mol. Cell Res* 1813, 1083–1091 (2011).
56. Giroux EL & Henkin RI Competition for zinc among serum albumin and amino acids. *Biochim. Biophys. Acta* 273, 64–72 (1972). [PubMed: 5038289]
57. Pantoliano MW, Valentine JS, Mammone RJ & Scholler DM The pH dependence of metal ion binding to the native zinc site of bovine erythrocyte (superoxide dismutase). *J. Am. Chem. Soc* 104, 1717–1723 (1982).
58. McPherson A Current approaches to macromolecular crystallization. *Eur. J. Biochem* 189, 1–23 (1990). [PubMed: 2185018]
59. Newman J, Sayle R. a & Fazio VJ A universal indicator dye pH assay for crystallization solutions and other high-throughput applications. *Acta Crystallogr. D. Biol. Crystallogr* 68, 1003–9 (2012). [PubMed: 22868766]
60. Newman J, Fazio VJ, Lawson B & Peat TS The C6 Web Tool: A Resource for the Rational Selection of Crystallization Conditions. *Cryst. Growth Des* 10, 2785–2792 (2010).

61. Mikol V, Rodeau JL & Giegé R Changes of pH during biomacromolecule crystallization by vapor diffusion using ammonium sulfate as the precipitant. *J. Appl. Crystallogr* 22, 155–161 (1989).
62. Hampton Research Corporation. PEG Stability: A Look at pH and Conductivity Changes over Time in Polyethylene Glycols (2012).
63. Orlov YF, Maslov EI & Belkina EI Solubilities of Metal Hydroxides 58, 1458–1466 (2013).
64. Fischer BE, Häring UK, Tribolet R & Sigel H Metal ion/buffer interactions: Stability of binary and ternary complexes containing 2-amino-2(hydroxymethyl)-1,3-propanediol (Tris) and adenosine 5'-triphosphate (ATP). *Eur. J. Biochem* 94, 523–530 (1979). [PubMed: 428398]
65. Belviso BD et al. Oxaliplatin Binding to Human Copper Chaperone Atox1 and Protein Dimerization. *Inorg. Chem* 55, 6563–6573 (2016). [PubMed: 27305454]
66. David G, Blondeau K, Schiltz M, Penel S & Lewit-Bentley A YodA from *Escherichia coli* is a Metal-binding, Lipocalin-like Protein. *J. Biol. Chem* 278, 43728–43735 (2003). [PubMed: 12909634]
67. Cotelesage JJH, Pushie MJ, Grochulski P, Pickering IJ & George GN Metalloprotein active site structure determination: Synergy between X-ray absorption spectroscopy and X-ray crystallography. *J. Inorg. Biochem* 115, 127–137 (2012). [PubMed: 22824156]
68. Ascone I & Strange R Biological X-ray absorption spectroscopy and metalloproteomics. *J. Synchrotron Radiat* 16, 413–421 (2009). [PubMed: 19395808]
69. Yukl ET et al. Diradical intermediate within the context of tryptophan tryptophylquinone biosynthesis. *Proc. Natl. Acad. Sci* (2013). doi:10.1073/pnas.1215011110
70. Garman EF Radiation damage in macromolecular crystallography: what is it and why should we care? *Acta Crystallogr. Sect. D Biol. Crystallogr* 66, 339–351 (2010). [PubMed: 20382986]
71. Seitlich T, Kühnel K, Schulze-Briesse C, Shoeman RL & Schlichting I Cryoradiolytic reduction of crystalline heme proteins: Analysis by UV-Vis spectroscopy and X-ray crystallography. *J. Synchrotron Radiat* 14, 11–23 (2007). [PubMed: 17211068]
72. Frankaer CG, Mossin S, Ståhl K & Harris P Towards accurate structural characterization of metal centres in protein crystals: the structures of Ni and Cu T6 bovine insulin derivatives. *Acta Crystallogr. Sect. D Biol. Crystallogr* 70, 110–122 (2014). [PubMed: 24419384]
73. Trofimov AA et al. Structural study of the X-ray-induced enzymatic reaction of octahaem cytochrome C nitrite reductase. *Acta Crystallogr. Sect. D* 71, 1087–1094 (2015). [PubMed: 25945574]
74. Brophy MB, Hayden JA & Nolan EM Calcium Ion Gradients Modulate the Zinc Affinity and Antibacterial Activity of Human Calprotectin. *J. Am. Chem. Soc* 134, 18089–18100 (2012). [PubMed: 23082970]
75. Shirran S, Garnaud P, Daff S, McMillan D & Barran P The formation of a complex between calmodulin and neuronal nitric oxide synthase is determined by ESI-MS. *J. R. Soc. Interface* 2, 465–476 (2005). [PubMed: 16849206]
76. Kaltashov Igor A., Zhang Mingxuan, Eyles Stephen J. & Abzalimov Rinat R.. Investigation of structure, dynamics and function of metalloproteins with electrospray ionization mass spectrometry. *Anal Bioanal Chem* 386, 472–481 (2006). [PubMed: 16932945]
77. Czernusiewicz RS in *Spectroscopic Methods and Analyses: NMR, Mass Spectrometry, and Metalloprotein Techniques* (eds. Jones C, Mulloy B & Thomas AH) 345–374 (Humana Press, 1993). doi:10.1385/0-89603-215-9:345
78. Vincent K Triggered infrared spectroscopy for investigating metalloprotein chemistry. *Philos Trans A Math Phys Eng Sci* 368, 3713–3731 (2010). [PubMed: 20603378]
79. Ronda L, Bruno S, Bettati S, Storici P & Mozzarelli A From protein structure to function via single crystal optical spectroscopy. *Front. Mol. Biosci* 2, 12 (2015). [PubMed: 25988179]
80. Stoner-Ma D et al. Single-crystal Raman spectroscopy and X-ray crystallography at beamline X26-C of the NSLS. *J. Synchrotron Radiat* 18, 37–40 (2011). [PubMed: 21169688]
81. Merlino A et al. Crystallization, preliminary X-ray diffraction studies and Raman microscopy of the major haemoglobin from the sub-Antarctic fish *Eleginops maclovinus* in the carbomonoxy form. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun* 66, 1536–1540 (2010).

82. Vergara A, Merlino A, Pizzo E, D'Alessio G & Mazzarella L A novel method for detection of selenomethionine incorporation in protein crystals {it via} Raman microscopy. *Acta Crystallogr. Sect. D* 64, 167–171 (2008). [PubMed: 18219116]
83. Hall JP et al. Monitoring one-electron photo-oxidation of guanine in DNA crystals using ultrafast infrared spectroscopy. *Nat Chem* 7, 961–967 (2015). [PubMed: 26587711]
84. Knape MJ et al. Divalent Metal Ions Mg 2+ and Ca 2+ Have Distinct Effects on Protein Kinase A Activity and Regulation. *ACS Chem. Biol* 10, 2303–2315 (2015). [PubMed: 26200257]
85. Lu CH, Lin YF, Lin JJ & Yu CS Prediction of metal ion-binding sites in proteins using the fragment transformation method. *PLoS One* 7, 1–12 (2012).
86. Ebert JJC & Altman RRB Robust recognition of zinc binding sites in proteins. *Protein Sci* 17, 54–65 (2008). [PubMed: 18042678]
87. Zhao W et al. Structure-based de novo prediction of zinc-binding sites in proteins of unknown function. *Bioinformatics* 27, 1262–1268 (2011). [PubMed: 21414989]
88. Sodhi JS et al. Predicting Metal-binding Site Residues in Low-resolution Structural Models. *J. Mol. Biol* 342, 307–320 (2004). [PubMed: 15313626]
89. Passerini A, Lippi M & Frasconi P MetalDetector v2.0: Predicting the geometry of metal binding sites from protein sequence. *Nucleic Acids Res* 39, 288–292 (2011).
90. Ferrè F & Clote P DiANNA 1.1: an extension of the DiANNA web server for ternary cysteine classification. *Nucleic Acids Res* 34, W182–5 (2006). [PubMed: 16844987]
91. Brylinski M & Skolnick J FINDSITE-metal: Integrating evolutionary information and machine learning for structure-based metal binding site prediction at the proteome level. *Proteins* 79, 735–751 (2011). [PubMed: 21287609]
92. Alpi E et al. Analysis of the tryptic search space in UniProt databases. *Proteomics* 15, 48–57 (2015). [PubMed: 25307260]
93. Chitale M, Hawkins T, Park C & Kihara D ESG: extended similarity group method for automated protein function prediction 25, 1739–1745 (2009).
94. Altschul SF, Gish W, Miller W, Myers EW & Lipman DJ Basic local alignment search tool. *J. Mol. Biol* 215, 403–410 (1990). [PubMed: 2231712]
95. Leuthner B et al. Biochemical and genetic characterization of benzylsuccinate synthase from *Thauera aromatica*: a new glycyl radical enzyme catalysing the first step in anaerobic toluene metabolism. *Mol. Microbiol* 28, 615–628 (1998). [PubMed: 9632263]
96. Echavarri-Erasun C, Arragain S & Rubio LM in *Metalloproteins: Methods and Protocols* (eds. Fontecilla-Camps JC & Nicolet Y) 5–18 (Humana Press, 2014). doi: 10.1007/978-1-62703-794-5_2
97. Pestov NB & Rydstro J Purification of recombinant membrane proteins tagged with calmodulin-binding domains by affinity chromatography on calmodulin-agarose: example of nicotinamide nucleotide transhydrogenase. *Nat Protoc* 2, 198–202 (2007). [PubMed: 17401354]
98. Majorek KA, Kuhn ML, Chruszcz M, Anderson WF & Minor W Double trouble-Buffer selection and His-tag presence may be responsible for nonreproducibility of biomedical experiments. *Protein Sci* 23, 1359–68 (2014). [PubMed: 25044180]
99. Van Dam ME, Wuenschell GE & Arnold FH Metal affinity precipitation of proteins. *Biotechnol Appl Biochem* 11, 492–502 (1989). [PubMed: 2508699]
100. Ju T et al. One protein, two enzymes revisited: a structural entropy switch interconverts the two isoforms of acireductone dioxygenase. *J. Mol. Biol* 4, 823–834 (2011).
101. Smith RM, Martell AE & Motekaitis RJ NIST critically selected stability constants of metal complexes database (2004).
102. Bijelic A, Theiner S, Keppler BK & Rompel A X-ray Structure Analysis of Indazolium trans-[Tetrachlorobis(1H-indazole)ruthenate(III)] (KP1019) Bound to Human Serum Albumin Reveals Two Ruthenium Binding Sites and Provides Insights into the Drug Binding Mechanism. *J. Med. Chem* 59, 5894–5903 (2016). [PubMed: 27196130]
103. Yu Q, Kandegedara a, Xu Y & Rorabacher DB Avoiding interferences from Good's buffers: A contiguous series of noncomplexing tertiary amine buffers covering the entire range of pH 3–11. *Anal. Biochem* 253, 50–56 (1997). [PubMed: 9356141]

104. Nakon R & Krishnamoorthy CR Free-metal ion depletion by 'Good's' buffers. *Science* (80-.). 221, 749–750 (1983).
105. Sokolowska M & Bal W Cu(II) complexation by 'non-coordinating' N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid (HEPES buffer). *J. Inorg. Biochem* 99, 1653–1660 (2005). [PubMed: 15993944]
106. Ferreira CMH, Pinto ISS, Soares EV & Soares HMVM (Un)suitability of the use of pH buffers in biological, biochemical and environmental studies and their interaction with metal ions - a review. *RSCAdv* 5, 30989–31003 (2015).
107. Lundblad RL *Chemical Reagents for Protein Modification* (CRC Press, 2014).
108. Satyanarayana U *Biochemistry* (Elsevier, 2014).
109. Phillips CM, Schreiter ER, Stultz CM & Drennan CL Structural Basis of Low-Affinity Nickel Binding to the Nickel-Responsive Transcription Factor NikR from *Escherichia coli*. *Biochemistry* 49, 7830–75838 (2010). [PubMed: 20704276]
110. Champloy F, Gruber K, Jogl G, Kratky C & IUCr. XAS spectroscopy reveals X-ray-induced photoreduction of free and protein-bound B 12 cofactors. *J. Synchrotron Radiat* 7, 267–273 (2000). [PubMed: 16609206]
111. Huang Q & Szebenyi DME Improving diffraction resolution using a new dehydration method. *Acta Crystallogr. Sect. F, Struct. Biol. Commun* 72, 152–159 (2016). [PubMed: 26841767]
112. Newman J A review of techniques for maximizing diffraction from a protein crystal in stilla. *Acta Crystallogr. Sect. D* 62, 27–31 (2006). [PubMed: 16369091]
113. Kim CU, Wierman JL, Gillilan R, Lima E & Gruner SM A high-pressure cryocooling method for protein crystals and biological samples with reduced backgrounda X-ray scatter. *J. Appl. Crystallogr* 46, 234–241 (2013). [PubMed: 23396891]
114. Vernède X, Fontecilla-Camps JC & IUCr. A method to stabilize reduced and/or gas-treated protein crystals by flash-cooling under a controlled atmosphere. *J. Appl. Crystallogr* 32, 505–509 (1999).
115. Bearden JA X-Ray Wavelengths. *Rev. Mod. Phys* 39, 78–124 (1967).
116. Lower limits of detection of synchrotron radiation high-energy X-ray fluorescence spectrometry and its possibility for the forensic application for discrimination of glass fragments. *Forensic Sci. Int* 175, 227–234 (2008). [PubMed: 17764863]
117. Awaji N et al. Detection Limits of Trace Elements for Wavelength Dispersive Total X-Ray Fluorescence under High Flux Synchrotron Radiation. *Jpn. J. Appl. Phys* 43, 1644–1648 (2004).
118. Sutton SR, Rivers ML & Smith JV Synchrotron x-ray fluorescence: diffraction interference. *Anal. Chem* 58, 2167–2171 (1986).
119. *The International Tables for Crystallography Volume C* (1999).
120. Srivastava UC & Nigam HL X-ray absorption edge spectrometry (xaes) as applied to coordination chemistry. *Coord. Chem. Rev* 9, 275–310 (1973).
121. Sathyanarayana DN *Electronic Absorption Spectroscopy and Related Techniques* (Universities Press, 2001).
122. Cotelesage JHH, Grochulski P, Pickering IJ, George GN & Fodje MN X-ray absorption spectroscopy at a protein crystallography facility: the Canadian Light Source beamline 08B1–1. *J. Synchrotron Radiat* 19, 887–891 (2012). [PubMed: 23093746]
123. Winn MD et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr. D. Biol. Crystallogr* 67, 235–242 (2011). [PubMed: 21460441]
124. Murshudov GN et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D. Biol. Crystallogr* 67, 355–367 (2011). [PubMed: 21460454]
125. Zheng H et al. CheckMyMetal: Macromolecular metal binding validation tool. *Acta Crystallogr. Sect. D* 73, (2017).
126. Chen VB et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D. Biol. Crystallogr* 66, 12–21 (2010). [PubMed: 20057044]
127. Gore S, Velankar S & Kleywegt GJ Implementing an X-ray validation pipeline for the Protein Data Bank. *Acta Crystallogr. Sect. D Biol. Crystallogr* 68, 478–483 (2012). [PubMed: 22505268]

128. Minor W, Cymborowski M, Otwinowski Z & Chruszcz M HKL-3000: the integration of data reduction and structure solution--from diffraction images to an initial model in minutes. *Acta Crystallogr. D. Biol. Crystallogr* 62, 859–866 (2006). [PubMed: 16855301]
129. Stepanov S et al. JBluIce - EPICS control system for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* 67, 176–188 (2011). [PubMed: 21358048]
130. Kabsch W XDS. *Acta Crystallogr. Sect. D Biol. Crystallogr* 66, 125–132 (2010). [PubMed: 20124692]
131. Adams PD et al. PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D Biol. Crystallogr* 66, 213–221 (2010). [PubMed: 20124702]
132. Emsley P & Cowtan K Coot: model-building tools for molecular graphics. *Acta Crystallogr. D. Biol. Crystallogr* 60, 2126–2132 (2004). [PubMed: 15572765]
133. Bruvold WH A meta-analysis of the California school-based risk reduction program. *J Drug Educ* 20, 139–152 (1990). [PubMed: 2144590]
134. Kimple ME, Brill AL & Pasker RL Overview of Affinity Tags for Protein Purification. *Curr. Protoc. Protein Sci* 73, Unit-9.9 (2013).
135. Bonner PLR Protein Purification (Taylor & Francis Group, 2007).
136. Scopes RK Protein Purification: Principles and Practice (Springer, 1993).
137. Amman AA Inductively coupled plasma mass spectrometry (ICP MS): a versatile tool. *J Mass Spectrom* 42, 419–427 (2007). [PubMed: 17385793]
138. McPherson A Crystallization of Biological Macromolecules (Cold Spring Harbor Laboratory Press, 1999).
139. Pflugrath JW Macromolecular cryocrystallography—methods for cooling and mounting protein crystals at cryogenic temperatures. *Methods* 34, 415–423 (2004). [PubMed: 15325658]
140. Tereshko V et al. Detection of alkali metal ions in DNA crystals using state-of-the-art X-ray diffraction experiments. *Nucleic Acids Res* 29, 1208–15 (2001). [PubMed: 11222771]
141. Niedzialkowska E et al. Optimization of overexpression of a chaperone protein of steroid C25 dehydrogenase for biochemical and biophysical characterization. *Protein Expr. Purif* 134, 47–62 (2017). [PubMed: 28343996]
142. Lion T et al. The translocation t(1;22)(p13;q13) is a nonrandom marker specifically associated with acute megakaryocytic leukemia in young children. *Blood* 79, 3325–30 (1992). [PubMed: 1596573]
143. Blindauer CA et al. Structure, properties, and engineering of the major zinc binding site on human albumin. *J. Biol. Chem* 284, 23116–23124 (2009). [PubMed: 19520864]
144. Newman J Expanding screening space through the use of alternative reservoirs in vapor-diffusion experiments. *Acta Crystallogr. Sect. D Biol. Crystallogr* 61, 490–493 (2005). [PubMed: 15805606]
145. Till M et al. Improving the success rate of protein crystallization by random microseed matrix screening. *J. Vis. Exp* e50548 (2013). doi:10.3791/50548
146. Benvenuti M & Mangani S Crystallization of soluble proteins in vapor diffusion for x-ray crystallography. *Nat. Protoc* 2, 1633–51 (2007). [PubMed: 17641629]
147. Malawski GA et al. Identifying protein construct variants with increased crystallization propensity-- a case study. *Protein Sci* 15, 2718–28 (2006). [PubMed: 17132859]
148. Vedadi M et al. Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. *Proc. Natl. Acad. Sci. U. S. A* 103, 15835–40 (2006). [PubMed: 17035505]
149. McPherson A & Cudney B Optimization of crystallization conditions for biological macromolecules. *Acta Crystallogr. Sect. FStructural Biol. Commun* 70, 1445–1467 (2014).
150. Dong A, Xu X & Edwards A In situ proteolysis for protein crystallization and structure determination. *Nat Methods* 4, 1019–1021 (2007). [PubMed: 17982461]
151. Holdgate G Isothermal titration calorimetry and differential scanning calorimetry. *Methods Mol. Biol* 572, 101–133 (2009). [PubMed: 20694688]

Box 1.**Thermofluor shift assay (TSA)**

TSA measures the thermal stability of a protein under varying conditions such as pH, presence and concentration of different salts or other additives. Protein is mixed with a fluorescent dye (most often SYPRO Orange), which exhibits an increase of fluorescence signal in a hydrophobic environment, and the sample is gradually heated. The protein unfolds with increasing temperature, and its exposed hydrophobic regions bind the fluorescent dye. The melting temperature (T_m), which is the temperature at which 50% of the protein is unfolded, is determined. Buffering reagents and other components present such as metal ions affect the melting temperature - protein is stabilized (higher T_m) or destabilized (lower T_m) when compared to a control condition without metal. Fluorescence and temperature can be monitored using a real time PCR instrument. A detailed protocol for sample preparation, including protein and dye concentrations and run parameters is described here¹⁴¹.

Additional reagents

SYPRO Orange 5000x stock (#S6650, ThermoFisher Scientific)

Hard-Shell® 96-Well PCR Plates (# HSP9601, Bio-Rad)

PCR Plate Sealing Film (#MSB1001, Bio-Rad)

Additional equipment

Real-Time PCR instrument

Procedure

1. Mix protein (typically a final concentration is between 0.5 mg/mL and 4 mg/mL), SYPRO Orange (typically stock solution is diluted between 1:100 and 1:1000)) and additives, such as salts of metals and buffering reagents. The total volume of a TSA reaction is typically 25 μ L. A control reaction includes the protein and dye, in absence of the reagent being tested.
2. Run TSA experiment. Example protocol: incubate sample at 20°C for 5 min, increase temperature from 20°C to 95°C with heating at a rate of 1°C/15 seconds. Record the fluorescence signal every 1°C.
3. Determine melting temperature (T_m) values from the maximum of the first derivative of the fluorescence intensity. Protein and dye concentration, as well as parameters of the run should be tested empirically. An unfolding curve with an easily defined melting point is an indicator of well-adjusted parameters of the TSA experiment.

Box 2.**Isothermal titration calorimetry (ITC)**

ITC allows the affinity of the interaction between a protein and a ligand to be measured. Usually, the ligand is titrated into a protein sample over certain time spans and volumes, and the emitted or absorbed heat of each injection is recorded. Integrated heat values are plotted against a concentration ratio of ligand to protein. A binding isothermal fit is performed using various fit models depending on the nature of binding (i.e. one site, two sites etc). The stoichiometry, dissociation constant and ΔH are calculated from the fit. Typical protocols for ITC sample preparation are described here^{51,151}.

Additional equipment

ITC instrument

Procedure

1. Dialyze protein into a buffer that will be used for the ITC experiment.
2. Prepare protein at a concentration of 10–100 μM or higher, and a ligand solution at a concentration 5–20 times higher than the protein in the same dialysis buffer. Proper ligand and protein concentrations should be tested empirically.
3. Run ITC experiment. An example of run parameters: 16 injections of 2.0 μL over 4.0 s with an interval of 180 s between injections to allow for complete equilibration; stirring speed: 700 rpm, temperature 25°C.

▲ **CRITICAL STEP** Buffers used for sample preparation should be degassed before an ITC experiment.

Box 3.

Inductively coupled plasma optical emission spectrometry (ICP-OES) and inductively coupled plasma mass spectrometry (ICP-MS)

ICP-OES or ICP-MS provide quantitative composition of elements in a sample.

Additional reagents

70% HNO₃ (#225711 –475ML, Sigma)

! CAUTION HNO₃ may cause metal leaching from glass bottles. Use polypropylene tubes for sample preparation. Metal ICP standards for example: Nickel ICP standard (#1703820100, Merck); Iron standard (#1197810100, Merck)

Procedure

1. Determine the detection limits for ICP-OES or ICP-MS spectrometers for a particular metal and estimate the amount of protein required for detection. For instance, the detection limit of zinc can be in the range of 1.0×10^{-3} to 100 mg/L, nickel is 1.0×10^{-3} to 10 mg/L, and molybdenum is 0.3×10^{-3} to 1 mg/L for ICP-MS.
2. Prepare serial dilutions of a metal solution standard so that the amount of metal is above and slightly below the detection limit. The metal standard solution should be diluted in the same buffer as that used for the protein. A standard solution should be handled exactly the same way as the protein and control sample to take into account the effects of sample manipulation.
3. Prepare the protein solution and a control sample in which the protein sample is substituted with the same buffer used for the protein.
4. Add HNO₃ to a final concentration of 1% up to 3% (vol./vol.) to metal standard solutions, protein and control sample.
5. Run the experiment, analyze results and calculate the metal content in a sample considering dilutions that were used and metal content in the control sample.

Figure 1. Overview of the protocol depicted in a flow chart.
The anticipated timing is denoted on the top of each stage.

Figure 2. Representative structures of horse serum albumin (ESA)-Zn²⁺ complexes showing the dynamic behavior of His-247.

Both structures were crystallized in 100 mM Tris buffer, and the pH refers to the final pH in the crystallization drop. (a) ESA-Zn²⁺, 2.5 mM Zn²⁺, pH 7.4; PDB ID: 5IIH. (b) ESA-Zn²⁺, 15 mM Zn²⁺, pH 6.5; PDB ID: 5IIX. Residues are shown as sticks, zinc ion in gray, oxygen in red, nitrogen in dark blue, and carbon in green. Coordination bonds are marked with black dashed lines. Gray grid represents 2mF_o - DF_c map (σ - 1.0), orange - anomalous map (σ - 3.0). Data for this figure were taken from Handing K.B., et al.⁵¹

Figure 3. The influence of transition metals on the stability of STM1931 protein from *S. typhimurium* analyzed by thermal shift assay (TSA) experiments.

The STM1931 protein in the presence of 189 mM NaCl was used as a reference. The presence of 0.273 mM CdCl₂ decreased the stability of this protein, while the concentrations of 4.38 mM of ZnCl₂ and 2.35 mM of (NH₄)₂MoO₄ (ammonium molybdate) increased the stability of the protein. The diffraction experiments showed that the C103A mutant of STM1931 protein from *S. typhimurium* crystallized in the presence of zinc ions (PDB ID: 4K2H).

Figure 4. Fluorescence spectra, fluorescence scan, and electron density maps of a zinc-containing protein dihydroorotase from *Yersinia pestis* CO92.

(a) X-ray fluorescence spectrum of the sample excited at the selenium K-edge (12664 eV). The peak around 12600 eV corresponds to the energy of the incident X-ray beam. The peak around 8600 eV corresponds to the $K\alpha$ emission energy of zinc. The peak at 3300 eV may correspond to the $K\alpha$ emission energy of potassium. The small peak at 6400 eV may correspond to the $K\alpha$ emission energy of iron present in the loop pin. (b) Fluorescence spectra of the crystal at excitation energies below (9618 eV – orange) and above (9668 eV – blue) the theoretical value of zinc absorption K-edge (9659 eV). The peaks at 3300 eV and 6400 eV are similar to those in panel a. (c) X-ray fluorescence scan of the sample at zinc K-edge region. The f'' plot and the observed absorption edge (9663 eV, or 1.283 Å) correspond to the table values for zinc K-edge. (d) Models and the electron density maps based on the data collected above the zinc absorption edge (left, model in cyan) and below the edge (right, model in yellow). Residues are shown in sticks, zinc ion in gray, oxygen in red, nitrogen in dark blue, carbon in cyan or yellow. Gray grid represents $2mF_o - DF_c$ map ($\sigma = 1.5$), pink - anomalous map ($\sigma = 4.0$).

Figure 5. Use of restraints for a zinc binding site in horse serum albumin (ESA).
 (a) 2Fo-Fc map with model and anomalous map, no metal. (b) Fragment of restraint file with LINK field present. (c) Refined model with metal and distances between ligands and metal shown. Residues are shown in sticks, oxygen in red, nitrogen in dark blue, carbon in cyan, water shown as a red sphere, and zinc shown as a gray sphere. Coordination bonds are marked with black dashed lines. Gray grid represents $2mF_o - DF_c$ map ($\sigma - 1.0$), pink - anomalous map ($\sigma - 3.0$).

Figure 6.
An example of a validation report for a zinc binding site ESA.

Table 1. Summarized features of ligands, distances, and geometry for commonly-encountered metal binding sites.

Metal class	Metal	Prevalence and distance to oxygen ligands in Å				Prevalence and distance to other ligands in Å			Prevalent geometry			Absorption edge in eV
		Asp/Glu	Water	Ser/Thr/Tyr, main-chain O	His-N _δ /N _ε 2	Cys-S _γ	Met-S _δ	Octahed ral	Tetrahed ral	Others		
Alkali and alkaline earth metals	Na ⁺	Common 2.44(14)	Common 2.41(9)	Common 2.41(16)	--	--	--	Common	--	--	--	
	Mg ²⁺	Common 2.08(5)	Common 2.07(4)	Common 2.08(5)	Rare 2.19(4)	--	--	Common (strict)	--	--	--	
	K ⁺	Common 2.81(15)	Common 2.80(10)	Common 2.83(14)	--	--	--	Common	--	--	--	
	Ca ²⁺	Common 2.40(11)	Common 2.42(6)	Common 2.44(11)	--	--	--	Common	--	--	--	
Transition metals from the 4 th period	Mn ^{2+/3+}	Common 2.16(10)	Common 2.19(6)	--	Common 2.23(7)	Rare 2.39(16)	--	Common	Rare (Mn ²⁺)	--	6539 (K)	
	Fe ^{2+/3+}	Common 2.06(8)	Common 2.10(4)	--	Common 2.12(9)	Common 2.27(19)	--	Common	Fe-S cluster	--	7112 (K)	
	Co ^{2+/3+}	Common 2.07(10)	Common 2.10(5)	--	Common 2.06(8)	Common 2.26(25)	--	Common	Rare	--	7709 (K)	
	Ni ²⁺	Common 2.06(8)	Common 2.08(6)	--	Common 2.08(6)	Common 2.18(19)	--	Common	Rare	Square planar	8333 (K)	
	Cu ^{+/2+}	Common 1.97(13)	Common 1.97(3) 2.37(17)	--	Common 1.99(7)	Common 2.33(27)	Rare 2.34(29)	Common (Cu ²⁺)	Common	Square planar (Cu ²⁺)	8979 (K)	
	Zn ²⁺	Common 1.97(15)	Common 2.09(8)	--	Common 2.02(7)	Common 2.36(26)	--	Rare	Common	Trigonal bipyramid	9659 (K)	
Transition metals from the 5 th and 6 th periods	Mo ^{4+/6+}	--	--					Common (Mo ⁴⁺)	--	Trigonal bipyrim (Mo ⁶⁺)	--	
	Cd ²⁺	2.34(17)	2.32(7)					Common	Common	--	--	
	Pt ^{2+/4+}	2.03(38)	2.12(23)		--	--	--	Common (Pt ⁴⁺)	--	Square planar (Pt ²⁺)	11,564 (L-III)	
	Au ^{+/3+}	--	--	--				--	--	Linear (Au ⁺) Square planar(Au ³⁺)	11,919 (L-III)	
	Hg ²⁺	2.73(25)	2.68(28)					--	--	Linear (Hg ²⁺)	12,284 (L-III)	
	Pb ²⁺	2.63(18)	2.68(18)					--	Common	--	13,035 (L-III)	

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

For ligands, the prevalence is denoted by “Common”, “Rare”, or a double dash. “Rare” denotes rare metal-ligand pairs; a double dash means the presence of a certain ligand around a metal is either impossible or negligible. For prevalent geometry, a double dash means the presence of a certain geometry is either impossible or negligible. The metal-ligand distance and ESDs are shown for common and rare ligands. Adapted from Ref. 125. Absorption edge column presents energy values for absorption edges with highest Γ change within 6000 – 17500 eV range¹¹⁹.

Table 2.

Troubleshooting Table.

Step	Problem	Possible reason	Solution
4	The protein precipitates after metal removal.	Metal plays a structural role.	Try to produce a recombinant protein using a defined growth medium containing metal of interest.
10(B)iii	No initial hits from crystallization screens.	Have not explored enough crystallization conditions in the crystallization space.	Set up more screens, try to find additional ligands that may stabilize your protein, use additional methods such as in situ proteolysis, or change the genetic construct of the protein. Many additional techniques to troubleshoot obtaining initial crystallization hits have been presented elsewhere and are beyond the scope of this protocol ^{144–150} .
13	Signal from the metal is not visible in the fluorescence spectrum.	Too low metal concentration.	Increase the concentration of metal during co-crystallization or soaking.
15	Large difference (>5 eV) between the observed and the tabulated value of the absorption edge.	<ol style="list-style-type: none"> 1 The chemical environment of the metal causes the shift. 2 Wavelength of the beamline may need to be calibrated. 	Use metal foil of the relevant metal to verify the wavelength of the beamline.
18	Signal from the metal is not visible on the anomalous map.	<ol style="list-style-type: none"> 1 The metal does not bind to weak metal binding site. 2 Competitive binding of the metal ion to chemicals present in the protein buffer or crystallization condition (e.g. Tris). 3 The metal binding site is occupied by a different metal. 4 Poor quality of the dataset does not allow for extraction of the anomalous signal. 	<ol style="list-style-type: none"> 1 Increase the concentration of metal during co-crystallization or soaking. 2 Decrease the concentration of the chelating compound, use a different buffer that is not a metal chelator, or find different crystallization solution. 3 Use fluorescence spectra to identify the metal. Consider the possibility that the metal unexpectedly bound to the protein could be physiological. 4 Re-process the dataset paying special attention to the quality parameters, sufficient redundancy, and absence of significant radiation damage. Screen for better crystals.
19	Metal has too high of a B-factor compared to the environment, or has a negative peak on the difference map.	The metal binding site is not fully occupied, or the metal is wrongly identified.	Decrease the metal occupancy. Consider the possibility that a metal with noticeably smaller atomic number (and hence with less electrons) is bound.
21–22	Unexpected coordination geometry and/or metal-ligand distances.	<ol style="list-style-type: none"> 1 Incorrect metal is bound. 2 Reduction of the metal due to radiation damage. 3 Oxidation of the metal during crystallization. 	<ol style="list-style-type: none"> 1 Verify the metal content. 2 Try lower radiation dose during data collection. 3 Use anaerobic chamber and/or appropriate reducing agents.
22	Position of the metal does not correspond to the peak in the electron density.	Restraints are not tight enough to keep metal ion close to ligand.	Try to use tighter restraints.
23	Metal binding site possesses seemingly perfect geometry and distances for both metal and ligands, but the ligands exhibit poor agreement with experimental electron density maps.	Metal binding site is over-restrained in low-resolution data.	Try to loosen the restraints.

Table 3.

Proteins used in different steps of anticipated results

Anticipated results	Procedure	Experiment	Metal	Protein	Organism
Steps 1–9	Step 1	Preliminary analysis	Zn	Albumin	<i>Equus caballus</i>
	Step 2	TSA	Zn	STM1931	<i>Salmonella typhimurium</i>
	Step 6	ICP-OES	Mo	molybdenum cofactor-containing chaperone protein	<i>Sterolibacterium denitrificans</i>
Steps 10–12	Steps 10–12	Crystallization	Zn	Albumin	<i>Homo sapiens & Equus caballus</i>
Steps 13–18	Steps 13–18	Data collection	Zn	Dihydroorotase	<i>Yersinia pestis CO92</i>
Steps 19–24	Steps 19	Refinement	Zn	Albumin	<i>Equus caballus</i>
	Steps 20–24	Validation	Zn	Albumin	<i>Equus caballus</i>

Used proteins include human albumin (HSA) and horse albumin (ESA) (PDB IDs: 5IJF, 5IJ5)⁵¹, STM1931 from *Salmonella typhimurium* (PDB ID: 4K2H), molybdenum cofactor-containing chaperone protein from *Sterolibacterium denitrificans*¹⁴¹, dihydroorotase from *Yersinia pestis CO92* (PDB ID: 5V0G).