Engineered Metal-Binding Proteins: Purification to Protein Folding

FRANCES H. ARNOLD* AND BARRY L. HAYMORE

PROTEINS CAN MAKE USE OF METAL IONS TO BIND SUBstrates, to maintain structure, to effect catalysis, and for allosteric control and regulation. In order to hold a particular metal ion with high affinity and specificity, proteins form multidentate binding pockets designed to fulfill both the chemical and geometric bonding requirements of that metal. Metal recognition can be engineered into proteins for applications such as protein purification.

Suitable metal complexes with vacant coordination sites [such as copper(II)-iminodiacetate; Cu(II)IDA] bind to ligating atoms exposed on protein surfaces. The relatively strong yet kinetically labile interaction between surface coordinating residues and metal complexes is ideal for selective purification processes. Metal ions attached to solid supports have been used in metal-affinity chromatography to purify a wide variety of proteins. Metal complexes have also been used to alter the selectivities of liquid-liquid extraction and precipitation processes (*1*). These inexpensive, stable metal complexes can be chemically modified and recycled without any loss of protein binding activity.

Separation in a metal-affinity process is dictated by the content and distribution of metal-coordinating residues on the protein surface. Exposed histidines and, when they exist, cysteines are the most important metal-coordinating ligands at neutral pH. If a sufficiently large number of such ligands is accessible, the protein can be recovered efficiently from crude mixtures. However, because histidines are relatively uncommon, and surface cysteines are very rare, most proteins do not exhibit high affinities for the metal complexes. If specific high-affinity metal-binding sites exist, they are usually inaccessible to the complexed metals. Proteins produced through recombinant DNA techniques can be engineered to have high affinities for metal complexes and thus be easily purified from other proteins in the host organism (*2*).

We have used both the metal-coordinating ability of histidine and the scaffolding afforded by proteins to create synthetic metal-binding sites in proteins. A high-affinity site can be formed from as few as two properly positioned metal-coordinating ligands. Two surface-accessible histidines separated by three intervening residues in an α helix can form a ternary complex with a metal ion such as copper and an anchoring ligand attached to a polymer or solid support (Fig. 1A). The metal-binding affinity conferred by a surface His-X_3-His site makes possible a single-step purification of the engineered protein from a crude cell lysate. This has been demonstrated for His-X_3-His variants of bovine somatotropin (bST) (*3*) and other proteins, including cytochrome c and insulin-like growth factor–1 (IGF1). bST has an accessible His^{19} residue in an α helix; the addition of a single His residue results in a His-X_3-His variant (His^{15}-bST) with a high affinity for Cu(II) IDA. His^{15}-bST is efficiently recovered (up to 95%) in a relatively pure and biologically active form (97 to 98%) from a crude cell lysate on a high-capacity Cu(II)IDA matrix (Fig. 1B). The ability of the metal columns to concentrate and purify the engineered protein in such a selective manner reflects the strong binding of the engineered protein as well as the paucity of natural proteins with exposed chelating sites.

The ability of these engineered proteins to bind copper complexes can also be exploited in extraction in aqueous polymer-polymer and polymer-salt two-phase systems. His-X_3-His variants of bST (*4*) and *Saccharomyces cerevisiae* iso-1-cytochrome c (*5*) partition strongly into the polyethylene glycol (PEG) phase of PEG-dextran two-phase systems containing small amounts of PEG derivatized with Cu(II)IDA, while cell debris and contaminating proteins largely partition to the bottom, dextran-rich phase. An engineered metal-binding site on the protein surface often does not interfere with biological function or expression of the functional variant (*5*).

The basis for the high metal affinity of the α-helical His-X_3-His configuration is the chelate effect. When two or more donor atoms of a multidentate ligand simultaneously bind a single metal, the strength and selectivity of the interaction are greater owing to the formation of multiple coordinate-covalent bonds. Modeling calculations for first-row transition metals such as Cu(II) show that selected short amino acid sequences coupled with appropriate elements of secondary structure can fulfill the requirements for metal chelation (Table 1). Di-histidine configurations His-X_3-His in an α helix, His-X_2-His in a reverse β turn, and His-X-His in a β strand form chelating sites. The geometric requirements of chelation are highly specific: while two histidines separated by three amino acids in an undistorted α helix can chelate a metal ion, histidines separated by two or four residues do not (*6*). Replacement of one of the two histidines by another metal-coordinating residue such as aspartate results in weaker chelating sites that still have significant affinity for free Cu(II) and cationic copper complexes. Chelating configurations can also be obtained with Cys and mixed Cys-His residues; however, Zn(II) is preferred because free thiols are oxidized by Cu(II). The number of practical metal-chelating configurations in common secondary structure elements is very limited. Although additional chelating configurations can be found in other types of structured regions of proteins, the simple motifs in Table 1 are generally applicable for engineering metal-binding sites into proteins. These motifs are usually buried and inaccessible natural metalloproteins.

Metal-ligand binding can be far stronger than the sum of the binding strengths of the individual monodentate ligands. This "excess" binding strength, known as the chelate effect, is ~2.5 kcal mol^{-1} for a good bidentate chelate and primarily results from favorable conformational entropy changes associated with ligand binding, although conformational enthalpy and solvation-desolvation energetics can also be important. The sum of the strength of the second coordinate-covalent bond plus the chelate effect equals the energy difference between monodentate binding and chelating bidentate binding ($\Delta\Delta G°_{chelate}$). Thus $\Delta\Delta G°_{chelate}$ represents the

Table 1. Possible metal-chelation sites in α helix, β strand, and reverse β turn (type I′ and II′): (+) Chelation is possible; and (−) chelation cannot occur. Bonding parameters are M–N = 2.01 ± 0.02 Å and N–M–N = 75° to 115°.

Sequence	α Helix	β Strand	Reverse β Turn
HH	−	−	−
HXH	−	+	−
HX_2H	−	−	+
HX_3H	+	−	−
HX_4H	−	−	−

F. H. Arnold, Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125.
B. L. Haymore, Chemical Sciences Department, Central Research Laboratories, Monsanto Company, St. Louis, MO 63167.

*To whom correspondence should be addressed.

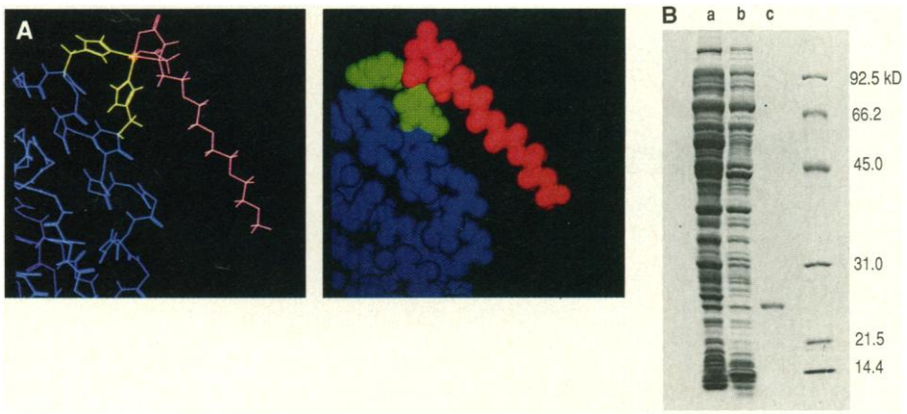

Fig. 1. (**A**) Two histidines (green) positioned His-X_3-His in NH_2-terminal α helix (blue) of *S. cerevisiae* iso-1-cytochrome c can chelate a metal ion. Increased partitioning or retention in metal-affinity separations occur when surface-accessible chelating sites form ternary complexes with the metal and an anchoring ligand (IDA) attached to a polymer (red) or solid support. (**B**) Purification of (refolded) His^{15}-bST (2.1 mg) from crude *Escherichia coli* cell lysate (1.1 g of protein) by chromatography on Cu(II)IDA-trisacryl (8-ml column) with a linear imidazole gradient. Polyacrylamide gel electrophoresis of (a) crude lysate; (b) column wash through under adsorption conditions (proteins not bound by the Cu(II)IDA matrix); and (c) eluate after two adsorption-elution steps.

positive contribution of the protein "scaffolding" to metal binding; it can be as great as 4 kcal mol^{-1} for Cu(II)IDA binding to a di-histidine site and can be significantly greater for binding to free copper (II).

Cu(II)IDA binding constants measured for a number of engineered His-X_3-His sites in exposed α helices vary over two orders of magnitude, from 2×10^4 to 2×10^6 M^{-1}, and depend on accessibility and rigidity of the particular site and the stereochemical match to the metal (*4*). For a good chelating di-histidine site, binding constants are two orders of magnitude greater than those for the same copper complex binding to a similar protein with two exposed, nonchelating His residues. Observed values of $\Delta\Delta G^\circ_{chelate}$ range from 0.9 to 3.5 kcal mol^{-1} for different His-X_3-His sites and reflect both the high sensitivity of metal binding to local structure and the differences in the structures and stabilities of helices in proteins. Without a stable structural framework, chelation does not occur. Thus the metal-binding strength of a chelating site is also highly sensitive to environmental conditions that alter its structure or stability. For example, the chelating His^4-His^8 site in the NH_2-terminal helix of cytochrome c loses its high affinity for soluble Cu(II)IDA-PEG above pH 8.0, even though the protein remains folded up to pH 12. Furthermore, this His-X_3-His site in the NH_2-terminal cytochrome c helix exhibits a relatively low affinity for Cu(II)IDA immobilized on a solid support. The loss of the protein scaffolding contribution at high pH or upon adsorption to a surface arises from small changes in the NH_2-terminal helix. This high sensitivity to environmental conditions is potentially useful in designing readily reversible metal-binding agents.

Metal complexes can act as sensitive probes of local protein conformational states. Two folding isomers of a His^8-His^{12} variant of IGF1 are completely separated from each other as well as from the unfolded protein by Cu(II)IDA chromatography (*7*). IGF1 contains six Cys residues in three disulfide bonds: 6-48, 18-61, and 47-52. When subjected to conditions that allow disulfide interchange, a biologically inactive isomer with different disulfide pairings (6-47, 18-61, and 48-52) forms in appreciable quantity. The properly folded His-X_3-His IGF1 variant is strongly retained on the copper column, owing to a well-formed chelating site, while the misfolded variant is less tightly bound. Although both isomers contain chelating His-X_3-His sites, the site in the misfolded protein is somehow distorted by the interchange of disulfide bonds.

Engineered metal-chelating sites are also effective in stabilizing folded proteins. Metal ions can shift the thermodynamic folding-unfolding equilibrium for a peptide or protein by binding preferentially to the folded form. Thus metal ions can stabilize existing or incipient structures: metal chelation by a His-X_3-His site can lock in a marginally stable helix or even add to the global stability of a folded protein. For proteins that obey a simple two-state folding mechanism, $\Delta\Delta G^\circ_{chelate}$ is equal to $\Delta\Delta G^\circ_{unf}$, the difference in the free energy of unfolding in the presence and absence of the metal complex. This quantity is the degree to which metal chelation stabilizes the folded protein. In fact, Cu(II)IDA stabilizes the His^4-His^8 cytochrome c variant to denaturation by guanidinium hydrochloride by the same free energy as its (independently measured) chelating contribution to metal binding (*8*). The addition of Cu(II)IDA to a variant containing a single His-X_3-His site can realize up to 4 kcal mol^{-1} of stabilization; many folded proteins are more stable than their unfolded forms by only 5 to 15 kcal mol^{-1}. Provided that the substitution of surface residues by histidine does not drastically reduce stability, engineered chelating sites can generally enhance protein stability at elevated temperatures or in the presence of denaturants such as organic solvents.

The ability to engineer strong metal binding into proteins has a number of technological applications (*9*). Metal chelation by engineered proteins can make metal ions available for catalysis and provide a rich source of structural information. Additional uses include development of metal-ion sensors and synthetic enzymes, regulation of protein function, probing protein conformations and folding pathways, and attachment of prosthetic metals for radiolabeling, spectroscopic labeling, or x-ray crystallography.

REFERENCES AND NOTES

1. For a review of metal-affinity protein separations, see F. H. Arnold, *Bio/Technology* 9, 151 (1991).
2. Smith and co-workers at Eli Lilly [M. C. Smith, T. C. Furman, T. D. Ingolia, C. Pidgeon, *J. Biol. Chem.* **263**, 7211 (1988)] and Hochuli and co-workers at Hoffmann-La Roche [E. Hochuli, W. Bannwarth, H. Dobeli, R. Gentz, D. Stuber, *Bio/Technology* **6**, 1321 (1988)] have attached histidine-containing peptides to recombinant proteins in order to confer a high affinity for metal complexes in metal-affinity chromatography.
3. S. J. Brewer, B. L. Haymore, T. P. Hopp, H. M. Sassenfeld, in *Purification and Analysis of Recombinant Proteins*, R. Seethram and S. Sharma, Eds. (Dekker, New York, 1991), pp. 239–265.
4. S.-S. Suh, B. L. Haymore, F. H. Arnold, *Protein Eng.* **4**, 301 (1991).
5. R. Todd, M. Van Dam, D. Casimiro, B. L. Haymore, F. H. Arnold, *Proteins* **10**, 156 (1991).
6. Similar geometric requirements are seen for the stabilization of α-helical peptides by ion pairs [S. Marqusee and R. L. Baldwin, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 8898 (1987)].
7. B. L. Haymore and A. M. Easton, unpublished results. To maintain the unfolded state, the protein is sulfonated at its six Cys residues.
8. J. T. Kellis, Jr., and F. H. Arnold, unpublished results.
9. For examples of other engineered metal-binding proteins, see: M. R. Ghadiri and C. Choi, *J. Am. Chem. Soc.* **112**, 1630 (1990); J. N. Higaki, B. L. Haymore, S. Chen, R. J. Fletterick, C. S. Craik, *Biochemistry* **29**, 8582 (1990); B. L. Iverson *et al.*, *Science* **249**, 659 (1990); T. Handel and W. F. DeGrado, *J. Am. Chem. Soc.* **112**, 6710 (1990).

10. We thank G. Wuenschell, R. Todd, S.-S. Suh, J. Kellis, R. Leimgruber, G. Bild, and G. Krivi for their contributions to this work. Support from Monsanto Company, the National Science Foundation, and the Office of Naval Research and a fellowship from the David and Lucile Packard Foundation are gratefully acknowledged (F.H.A.).