Designing Protein Separations Based on Metal-Affinity Interactions

Thesis by

Robert J. Todd

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1993

(Submitted September 25,1992)

Abstract

Immobilized metal-affinity chromatography (IMAC) provides a technique to separate proteins based on their histidine multiplicity and microenvironment. A set of *S. cerevisiae* iso-1-cytochrome c variants which differed only in their histidine content and placement was used to study equilibrium binding and chromatographic retention on matrices of immobilized Cu(II) (TSK macroporous support). The structural integrity of the variants was verified by UV-visible absorbance and ^{1}H NMR spectroscopy. Equilibrium binding studies indicate that a surface histidine is required in iso-1-cytochrome for interaction with a Cu(II)IDA (iminodiacetic acid) matrix. The apparent binding affinity of variants depends on the accessibility of the histidine side chains and on the total number of surface histidines. The overall equilibrium binding behavior is consistent with a model in which bound protein covers multiple copper sites on the matrix surface. Chromatographic studies on a Cu(II)IDA TSK column showed that a surface histidine is required for significant protein retention. It was also shown that proteins can be separated based on the microenvironment (accessibility or pK_a) of a single surface accessible histidine or on the multiplicity of surface histidines and the order of retention in all cases correlated with the initial slope of the equilibrium binding isotherm. The effect of proton and imidazole competition on retention was also studied, and it was seen that simple equilibrium models, describing the interaction of a single-histidine protein with independent copper binding sites, are insufficient to account for the strong competitive behavior that is observed experimentally. The ability of IMAC to discriminate between proteins based on their histidine multiplicity and microenvironment provides a powerful dimension in protein separations.

The concept of engineering di-histidine metal-chelating sites into proteins for applications in purification and protein stabilization was investigated. A His-X_3-His site was engineered into the *N*-terminal α-helix of iso-1-cytochrome c at positions 4 and 8. A

di-histidine site was also engineered across a small section of β-sheet structure at positions 39 and 58. Cu(II)IDA titrations of ^{1}H NMR spectra show selective paramagnetic broadening of the C2 resonances of the chelating histidines. The His-X_3-His site at positions 4 an 8 binds Cu(II)IDA-PEG in solution 24 times stronger than a single histidine site and both metal-chelating variants have increased retention in a Cu(II) IMAC column relative to a non-chelating control protein. The engineered metal-chelating sites were also shown to stabilize the folded form of iso-1-cytochrome c in the presence of Cu(II)IDA by as much as 2.5 kcal/mol.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Background

Biotechnology has been developed to the point where we can now clone and manipulate the gene for essentially any protein found in nature.[1] This enables us to produce proteins for therapeutic and catalytic applications by fermentation in bacterial systems designed specifically for high levels of expression. But, at the same time, our ability to purify proteins has not kept pace with recent advances in their production. The problem of purifying a protein of interest from the complex mixture of proteins and other biological molecules produced during fermentation is a very difficult one. The number of different proteins present in a given fermentation may exceed one thousand, and because all proteins are comprised of the same twenty amino acids, the macroscopic properties on which separations may be based (size, charge, *etc.*) are very similar.

The standard chromatographic techniques in protein purification are based on specific biological interactions (bioaffinity chromatography) or on differences in the proteins' size (size-exclusion chromatography), charge (ion exchange chromatography), or hydrophobicity (hydrophobic interaction chromatography). A combination of several of these methods is often needed to purify a protein to homogeneity, and the costs involved in the purification of a protein are often the largest expense involved in its production. The replacement of a typical multi-step purification process with more effective novel separation methods represents one means to make the production of proteins more cost effective, and in addition, reduce the time and effort involved in the development of new therapeutic and catalytic proteins.

A relatively new concept in protein purification, first introduced by Porath in 1975, is that proteins can be separated based on their affinity for immobilized metal ions.[2] The protein's affinity for immobilized metal ions is derived from amino acid side chains on its surface which are able to donate electrons to the metal ion complex. Immobilized metal affinity chromatography (IMAC) has many properties which are advantageous for protein

purification. IMAC can be thought of as a good compromise between relatively non-specific protein purification techniques, such as ion exchange, and very specific techniques that involve biospecific interactions. The metal-affinity interaction is strong and selective for a small subset of ligands on the protein surface (mainly surface histidines). Although the metal-affinity interaction may not rival many biospecific interactions in terms of strength and specificity, the metal-affinity ligand does has advantages over many biospecific affinity ligands. These advantages are derived from the nature of the immobilized metal ion complex, which is inexpensive, extremely stable, and can be derivatized on a support to provide very high capacity. In addition, the metal-affinity interaction is kinetically labile,[3] thus providing fast kinetics and good resolution in chromatographic separations, and can be disrupted under gentle elution conditions.[4] In short, the metal-affinity interaction provides an inexpensive and selective means to separate proteins without the difficulties involved in biospecific affinity.

Figure 1.1 illustrates the metal-affinity interaction in the chromatographic separation of proteins. The metal ion (Me) is typically a divalent transition metal from the second half of the first transition series. The most commonly encountered metals in IMAC are those originally proposed by Porath, Cu(II), Ni(II), Zn(II), Co(II),[2] and more recently Fe(III).[10,11] The transition metal is immobilized via a spacer arm to the matrix using a multidentate chelating group such as iminodiacetic acid (IDA). Proteins in the mobile phase can interact with the immobilized metal ion through coordinating ligands on their surfaces. Once bound, proteins can be eluted under gentle conditions by a number of techniques: reduction in pH resulting in protonation of the coordinating ligand, addition of a mobile phase competitor, or removal of the metal ion with a strong chelating agent.[4,7]

The immobilized metal ion complex is made up of two components, the metal ion and the immobilizing ligand, each of which may impact the nature of a given IMAC separation. The primary component of the complex is the metal ion. The metal ion dictates

both the strength and ligand specificity of an IMAC separation. As originally proposed by Porath, Cu(II), Ni(II), Zn(II),and Co(II) have primary affinities for histidine and cysteine residues on the surface of proteins.[2] In contrast, Fe(III) prefers oxygen ligands (carboxyl and phosphate groups) and thus separates proteins based on a different mechanism.[5,6] The strength of the metal-affinity interaction for the divalent transition metals used in IMAC generally follows the series Cu(II) > Ni(II) > Zn(II) ≈ Co(II).[4,7] Thus, both the specificity and strength of the interaction are dictated by the metal ion.

The chelating group used to immobilize the metal ion may also affect the strength and specificity of the metal-affinity interaction.[8] The most common immobilizing group in IMAC is iminodiacetic acid (IDA), shown in Figure 1.2. This tridentate chelate immobilizes the metal ion by way of a single amine and two carboxyl ligands, and the resulting complex is neutral for divalent metals. Numerous other immobilizing agents have been reported in the IMAC literature, most of which contain some combination of amine and/or carboxyl ligands.[8-10] For example, nitrilotriacetic acid (NTA) uses a single nitrogen and three carboxyl groups to immobilize the metal ion, as shown in Figure 1.2.[10] The number of chelating ligands on the immobilizing group affects both the strength and selectivity of the metal-affinity interaction. The immobilizing group must bind the metal ion with sufficient strength so that it remains bound to the matrix, and yet it must also leave free coordination sites on the metal ion for interaction with the protein. In general, for a given metal ion, the more coordination sites that are occupied by the immobilizing ligand, the weaker the interaction with the protein.[8] Other aspects of the metal ion complex, such as overall charge and coordination geometry, are also influenced by the immobilizing ligand and may affect the selectivity of the metal-affinity interaction.[9]

Additional factors that affect an IMAC separation are the properties of the protein surface and mobile phase competitors. The number and positioning of ligands on the protein surface is a primary determinant of its affinity for a given IMAC matrix.[4,11-13]

This affinity can be competitively reduced by either protonation of the interacting ligands or by the addition of competitors to the mobile phase. Both of these methods provide convenient and non-destructive means to elute bound proteins from an IMAC column.[4,7]

Because surface histidines (and cysteines) are relatively rare,[14] IMAC is an effective means of purification for proteins which contain multiple surface histidines or have naturally high-affinity for metal ions.[4,15-18] Recent advances in protein engineering have proven that the structure of a protein can be manipulated in a controlled manner for increased stability or novel catalytic activities.[1] Protein engineering may also be used to study and develop novel purification schemes. One example is the engineering of a high-affinity metal-binding site into a protein that can be used for purification by IMAC.

Previous IMAC Studies

Proteins interact with immobilized metal ions through surface amino acid side chains which are able to donate electrons to the immobilized metal ion complex. The original supposition by Porath was that the primary contributors to IMAC retention at neutral pH would be surface accessible cysteine, histidine, and tryptophan residues.[2] The estimated metal ion binding constants of the amino acid side chains which are potential ligands in IMAC are given in Table 1.1.[19] The most informative of these values is the apparent binding affinity for Cu(II) (K'), which represents the affinity of each side chain at pH 7.0, typically used in IMAC separations. The table suggests that the side chain of histidine (imidazole), which has the strongest affinity for Cu(II) at neutral pH, is expected to dominate the metal-affinity interaction (surface cysteines are oxidized by Cu(II) and are therefore not included in this comparison). The contribution of different amino acids has also been studied by investigating the binding of amino acids and small peptides to metal affinity columns.[20-23] The most informative of these papers studied the retention of 67 biologically active peptides under a number of experimental conditions.[23] The results

showed that retention on Cu(II), Ni(II) and Zn(II) is determined predominantly by the number of histidines in the peptide.

The information available on the interaction of proteins with IMAC supports was, at the time we began our study, very qualitative in nature. The bulk of the data indicated that proteins are retained on IMAC columns according to their histidine multiplicity.[4] For example, a set of cytochrome c homologues containing zero (tuna), one (horse), and two (candida) surface histidines are retained on a Cu(II)IDA column according to their histidine content. Similar behavior was observed with lysozyme homologues.[4] Peking duck lysozyme does not contain any surface histidines and is not retained on a Cu(II)IDA column, while hen egg-white lysozyme contains a single histidine and is retained. These studies provided supporting evidence that binding to a Cu(II)IDA column correlates with the multiplicity of histidines and that proteins could be separated based on their surface histidine content.[4,24]

Recent IMAC Studies

The interaction of proteins with immobilized metal-affinity supports has been extensively studied in the last few years and has recently been reviewed by Wong *et al.*[15] The first direct evidence that IMAC retention could be due to a single histidine was provided by Mrabet *et al.,* who used site-directed mutagenesis to show that histidine 41 of D-xylose isomerase is solely responsible for retention on a Cu(II)IDA column.[13] Hemdan *et al.* showed that three different proteins, each containing a single surface histidine, can be resolved on a Cu(II)IDA column by a stepwise pH gradient and the order of their elution correlates inversely with the pK_a of the surface histidines involved.[11] This indicated that proteins could be separated based on differences in the pK_a's of their surface histidines. The elution behavior of eight different homologues of lysozyme was studied by Zhao *et al.* using an imidazole gradient.[12] This group was able to separate the proteins into groups

based on their histidine content, confirming that differences in the multiplicity of histidines can be used to separate proteins by IMAC.

The first attempts to quantitate the interaction of proteins with immobilized metal ions was published by Belew *et al.* in 1987.[25] Frontal elution chromatography was used to measure adsorption isotherms of four model proteins: lysozyme, ovalbumin, bovine serum albumin (BSA), and pig serum albumin (PSA). The equilibrium isotherms were described using a Langmuir-type model, and thus were characterized by a single binding constant and a maximum capacity. Binding constants ranged from 1.8×10^4 M^{-1} for lysozyme to 3.7×10^5 M^{-1} for PSA. Proteins which contained more surface histidines had larger apparent binding constants. One experimental difficulty with using frontal elution to measure binding isotherms is its lack of sensitivity. In order to accurately determine Langmuir binding constants, equilibrium binding data must be obtained in the linear portion of the binding isotherm (low protein concentrations). This was not accomplished by Belew *et al.* in all cases, and as a result, the uncertainty in the predicted binding constants for ovalbumin, BSA, and PSA is quite large. In an effort to overcome the experimental limitations and difficulties of using frontal analysis as a method of measuring binding isotherms, Hutchens *et al.* developed a batch equilibrium binding method.[26] This method could be used to determine isotherms much more quickly and with less protein than frontal elution. It also provided the sensitivity necessary for the accurate determination of binding constants. Using this technique, the binding behavior of eight commercially available proteins was analyzed,[27] and although it was difficult to interpret differences in binding constants and capacities in terms of protein properties, this technique proved to be a useful experimental approach to studying the variables involved in IMAC.

Project Goals

The design and application of IMAC separations depends on an understanding of how proteins interact with and are recognized by metal ion complexes. Our goal in this study is to devise a theoretical framework for the study of IMAC and then to quantitatively investigate how the metal-affinity interaction is influenced by the number and position of ligands on the protein surface. These can be varied in a controlled manner by protein engineering. In addition, we will address the role that mobile phase competitors play in IMAC. A fundamental understanding of the protein interaction and the competitive equilibrium processes in IMAC will allow for more efficient and informed designs of IMAC separations.

While a complete understanding of IMAC has not yet been obtained, it is clear that IMAC is a very effective purification method for proteins with a naturally high affinity for metal ions (those with many histidines or natural metal-binding sites on their surface).[4,15-18] This is the case for only a very small subset of natural proteins (e.g., human fibroblast interferon).[15-18] Through protein engineering, we propose that simple modifications can be made to the surface of a protein which would confer unusually strong metal-binding properties. The goal is to demonstrate a general technique by which a protein can be engineered to effect its purification by IMAC, without altering catalytic or functional activity.

Approach

Protein engineering is used to create a set of iso-1-cytochrome c variants that differ only in the number and location of their surface histidines, including two variants which contain specifically designed di-histidine metal-chelating sites. The availability of a well-characterized expression system[28] and a high-resolution crystal structure[29,30] made iso-1-cytochrome c from *S. cerevisiae* an ideal model protein for this study. Controlled

manipulation of the protein surface properties enabled us to create a set of otherwise identical proteins with which to investigate the effect of the multiplicity and microenvironment of histidine binding sites on metal affinity. The cytochrome c variants are characterized by UV-Visible absorbance and ^{1}H NMR to ensure their correct conformation, and the binding affinity of the variants is then studied on matrices of Cu(II)IDA by measuring both equilibrium binding isotherms and chromatographic retention.

Table 1.1 Estimated equilibrium binding constants for amino acid side chains expected to participate as ligands in metal ion binding.[19] K represents equilibrium binding constant for deprotonated species (K = [ML]/[M][L]). K' represents equilibrium binding constant at pH 7.0 (K' = [ML]/([M][L + HL])).

Amino Acid		**Log K**		**Log K'**	
Sidechain	**Cu(II)**	**Ni(II)**	**Zn(II)**	**Cu(II)**	**pKa[a]**
Imidazole (His)	4.2	3.0	2.6	4.2	6.7
RS^- (Cys)	_ b	3.0	4.0	_ b	8.5
RCO_2^- (Asp, Glu)	1.7	0.9	0.9	1.7	3.9
p-$RC_6H_4O^-$ (Tyr)	4.5	-	-	1.6	9.9
RNH_2 (Lys)	4.1	-	-	0.9	10.2
$RSCH_3$ (Met)	0.5	-	-	0.5	-7
e NH $RNHCNH_2$ (Arg)	4.0	-	-	< 0	12.5
e O $RCNH_2$ (Asn, Gln)	<0	-	-	< 0	-0.5
Indole (Trp)	<0	-	-	< 0	-4
O $CH_3OCCH_2CO_2^-$ (*N*-terminus)	3.9	-	-	3.2	7.7
e O $CH_3CNHCH_2CO_2^-$ (*C*-terminus)	1.3	-	-	1.3	3.4

a Estimated values for sidechains in *N*-acetyl amino acid amides.

b Cu(II) is reduced by free thiols.

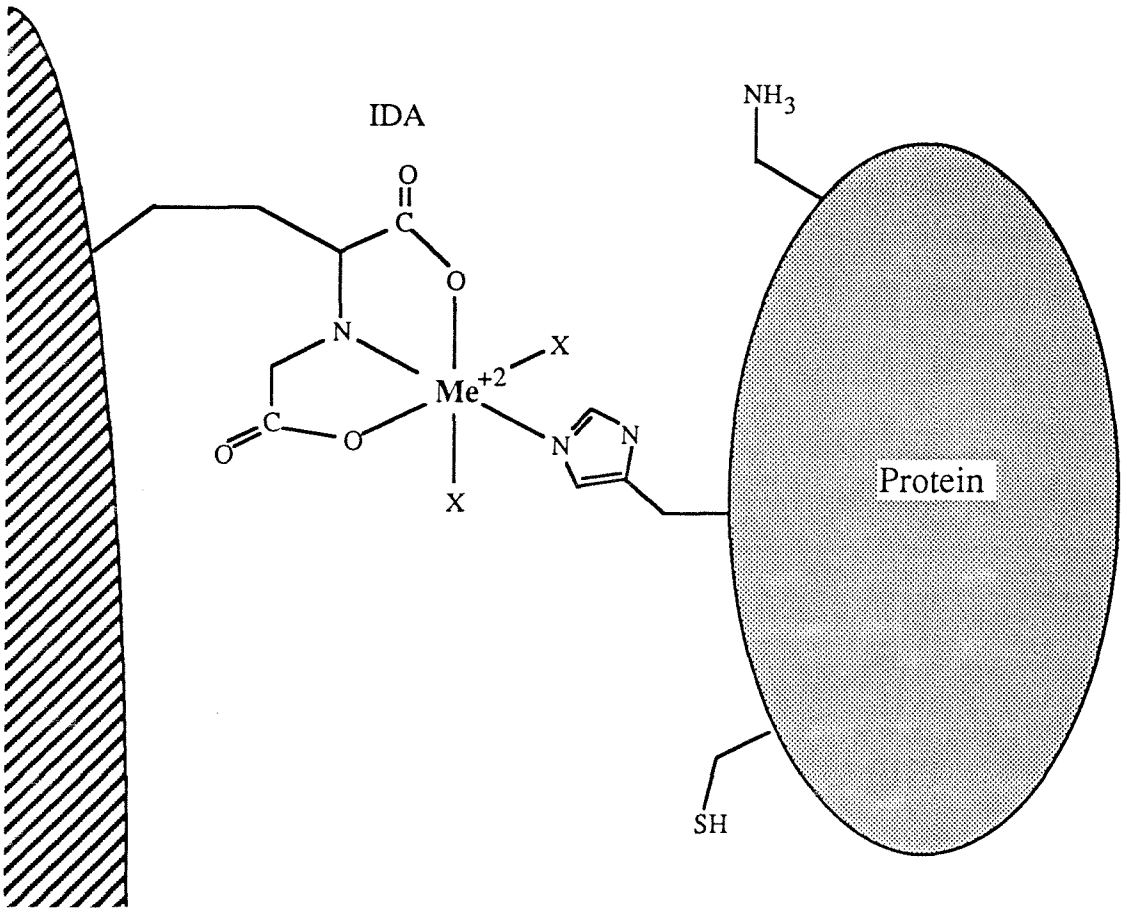


Figure 1.1 Illustration of interaction responsible for retention in metal-affinity chromatography. The metal ion (Me) is immobilized with a chelating ligand (iminodiacetic acid, IDA) which is attached to the chromatographic support via a spacer arm. Proteins interact with the immobilized metal ion through electron donating groups on their surface (imidazoles, thiols, *N*-terminus, *etc.*).

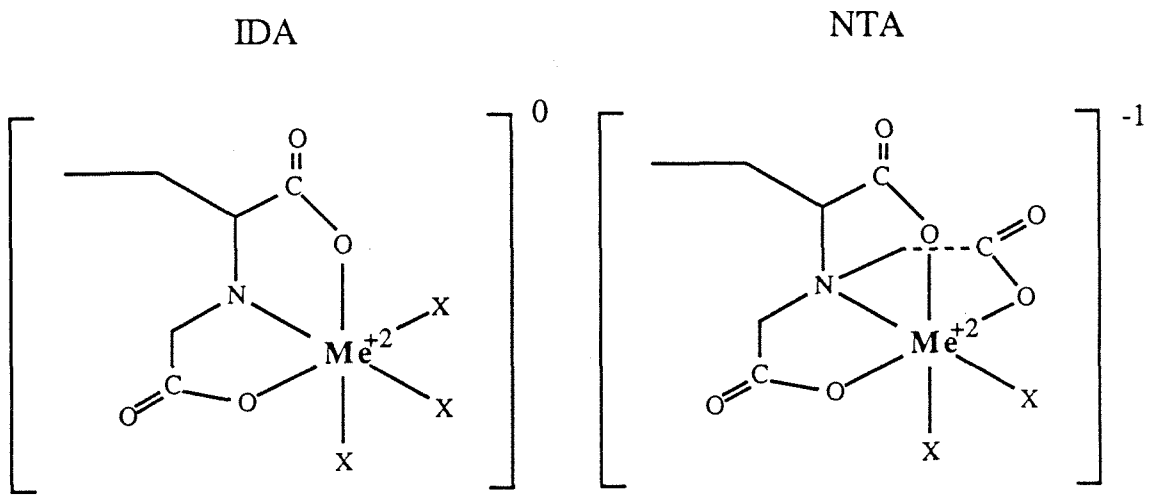


Figure 1.2 Selected immobilizing ligands used in IMAC. Imiodiacetic acid (IDA) chelates the metal ion using a single tertiary amine and two carboxyl groups. The resulting metal ion complex (metal-affinity ligand) is neutral and has three free coordination sites (X) with octahedral divalent transition metals. Nitrilotriacetic acid (NTA) chelates the metal ion via a single tertiary amine and three carboxyl groups resulting in a metal ion complex which has an overall charge of minus one and only two vacant coordination sites with octahedral divalent transition metals.

References

1. Ulmer, K. M., *Science* **219**, 666-671 (1983).
2. Porath, J., Carlsson, J., Olsson, I. and Belfrage, G., *Nature* **258**, 598-599 (1975).
3. Purcell, K. F. and Kotz, J. C., in *An Introduction to Inorganic Chemistry*, (Saunders, Philadelphia, 1980).
4. Sulkowski, E., *Trends in Biotechnology* **3**, 1-7 (1985).
5. Andersson, L. and Porath, J., *Analytical Biochemistry* **154**, 250-254 (1986).
6. Porath, J. and Olin,B., *Biochemistry* **22**, 1621-1630 (1983).
7. Fatiadi, A. J., *CRC Critical Reviews in Analytical Chemistry* **18**,1-44 (1987).
8. Porath, J., *Trends in Analytical Chemistry* **7,** 254-259 (1988).
9. Wuenschell, G. E., Wen, E, Todd, R. J., Shnek, D and Arnold, F. H., *Journal of Chromatography* **543**, 345-354 (1991).
10. Hochuli, E., Dobeli, H. and Schacher, A., *Journal of Chromatography* **411**, 177-184 (1987).
11. Hemdan, E. S., Zhao, Y., Sulkowski, E. and Porath, J., *Proceedings of the National Academy of Sciences USA* **86**, 1811-1815 (1989).
12. Zhao, Y.,Sulkowski, E. and Porath, J., *European Journal of Biochemistry* **202**, 1115-1119 (1991).
13. Mrabet, N. T. *Biochemistry* **31**, 2690-2702 (1992).
14. Klapper, M. H., *Biochemical and Biophysical Research Communications* **78**, 1018-1024 (1977).
15. Wong, J. W., Albright, R. L. and Wang, N. L., *Separation and Purification Methods* **20**, 49-106 (1991).
16. Matsuo, O., Tanbara, Y., Okada, K., Fukao, H., Bando, H and Sakai, T., *Journal of Chromatography* **369**, 391-397 (1986).
17. Kurecki, T., Kress, L. F. and Laskowski, M., *Analytical Biochemistry* **99**, 415-420 (1979).
18. Edy, V. G., Billiau, A. and De Somer, P., *Journal of Biological Chemistry* **252**, 5934-5935 (1977).
19. Haymore, B. L., unpublished results.
20. Hemdan, E.S. and Porath, J., *Journal of Chromatography* **323**, 255-264 (1985).
21. Hemdan, E.S. and Porath, J., *Journal of Chromatography* **323**, 265-272 (1985).
22. Nakagawa, Y., Yip, T., Belew, M. and Porath, J., *Analytical Biochemistry* **168**, 75-81 (1988).

23. Yip, T., Nakagawa, Y. and Porath, J., *Analytical Biochemistry* **183**, 159-171 (1989).

24. Figueroa, A., Corradini, C., Feibush, B. and Karger, B. L., *Journal of Chromatography* **371**, 335-352 (1986).

25. Belew, M., Yip, T.,Andersson, L. and Ehrnstron, R., *Analytical Biochemistry* **164**, 457-465 (1987).

26. Hutchens, T. W., Yip, T. and Porath, J., *Analytical Biochemistry* **170**, 168-182 (1988).

27. Hutchens, T. W., and Yip, T., *Analytical Biochemistry* **191**, 160-168 (1990).

28. Cutler, R. L., Pielak, G. J., Mauk, A. G. and Smith, M., *Protein Engineering* **1**, 95-99 (1987).

29. Louie, G. V., Hutcheon, L. B. and Brayer, G. D., *Journal of Molecular Biology* **199**, 295-314 (1988).

30. Louie, G. V. and Brayer, G. D., *Journal of Molecular Biology* **214**, 527-555 (1990).

Chapter 2

Protein Engineering

Introduction

The design and production of protein variants that contain histidines at various positions on the surface was carried out using iso-1-cytochrome c from *Saccharomyces cerevisiae*. Cytochrome c is the penultimate electron transfer protein of the respiratory chain whose function is to transfer electrons from cytochrome c reductase to cytochrome c oxidase. It is a comparatively small (103-113 amino acids), stable, water-soluble protein containing a prosthetic heme group. The heme group is covalently bound by two thio-ether linkages and cycles between the oxidized (Fe(III)) and reduced (Fe(II)) states during electron transfer.[1] X-ray crystallographic structures of cytochromes c from a number of species have been solved,[2-7] and the solution structure and dynamics have been studied by NMR.[8-16] The availability of a high-resolution crystal structure[2,3] and a well-characterized expression system[17] for iso-1-cytochrome c from *S. cerevisiae* made it an ideal choice for a model protein in this study.

Figure 2.1 is a graphical representation of the crystal structure of iso-1-cytochrome c from *Saccharomyces cerevisiae* that has been solved by Louie *et al.*[2,3]. Using the standard numbering for the alignment of mitochondrial cytochromes c, native iso-1-cytochrome c contains histidines at positions 18, 26, 33 and 39.[18] The histidine at position 18 is ligated to the heme iron at the center of the protein and is therefore unable to participate in binding to immobilized metal ions. The remaining native histidines at positions 26, 33 and 39 are at least partially accessible on the surface and are available to coordinate immobilized metal ions.

The sites chosen for histidines in this study were the native postitions at 26, 33, and 39, in addition to positions 4, 8 and 58, which were selected based on their high degree of surface accessibility. Surface histidines in various combinations at these six positions in iso-1-cytochrome c provided a powerful basis to address fundamental questions concerning the retention of proteins by IMAC. Moreover, mutations were introduced into cytochrome

c so as to minimize their effect upon the structure and function of the protein. Thus, potential mutation sites were referenced against known destructive cytochrome c mutations[19] and compared to conserved residues in cytochrome c homologues.[20] Site-directed mutagenesis was used to replace the native histidines at positions 26, 33, and 39 with either glutamine or asparagine, depending on sequence homology,[20] and the native residues at positions 4, 8, and 58 with histidines. In addition, the single surface cysteine at position 102 has been replaced by a serine in these studies to prevent protein dimerization.[17] Engineering of the protein surface allowed us to create a set of otherwise identical proteins which can be used to directly investigate how the multiplicity and environment of the histidine binding sites affects immobilized metal ion affinity.

Materials and Methods

Mutagenesis and Expression

A 2.5 kb BamH1-HindIII fragment (CYC1) encoding iso-1-cytochrome c cloned into M13mp8 was the generous gift of Professor M. Smith (University of British Columbia). Cytochrome c genes containing the mutations (C102S), (C102S;H39Q), (C102S;H39Q;H33Q;K4H), (C102S;H39Q;H33Q;T8H), and (C102S;H39Q;L58H) were kindly provided by J. H. Richards (California Institute of Technology). Site-directed mutagenesis on the 2.5 kb CYC1 fragment was performed by either the Amersham[21] or Bio-Rad[22] methods. The mutagenized M13 was transformed into *E. coli* TG1, and plaques were screened for the correct mutation using the dideoxy sequencing method.[23] The complete sequence of each mutant was confirmed by dideoxy sequencing, and the 2.5 kb CYC1 fragment was cloned into the yeast shuttle vector YEp213[24] and used to transform E. coli HB101. (The YEp213 shuttle vector contains both ampicillin and tetracycline resistance along with the Leu^{2+} gene from *S. cerevisiae*.) Upon insertion of the 2.5 kb CYC1 fragment, tetracycline resistance is lost, and plasmid from $amp^{+}tet^{-}$

colonies confirmed to contain the 2.5 kb fragment were used to transform yeast strain GM-3C-2[25] by the LiCl method.[26] The recipient strain contains a deletion mutation (CYC1-1) and a point mutation (CYC3-1) which eliminate and inactivate the genes for iso-1- and iso-2-cytochrome c, respectively. Thus the recipient strain cannot grow on glycerol as the sole carbon source.[25] Transformants were selected by growth on leu$^-$ plates to contain the plasmid, and further selected for production of functional iso-1-cytochrome c by growth on a non-fermentable carbon source (YPG plates: 1% yeast extract, 2% peptone, and 4% glycerol).

Cytochrome c Purification

Cytochrome c variants were purified according to a modified version of the procedure developed by Cutler and coworkers.[17] Yeast were grown at 30°C with moderate aeration for 4-5 days in a 15 L pyrex carboy containing 12 L of YPG media. The fermentation broth was concentrated to approximately 1 L in an Amicon DC 10 L cross-flow filtration device using a 0.1 μm cutoff hollow fiber cartridge (H5MP01). The cells were then washed twice with 5 L of 0.9% NaCl and reconcentrated to 1 L. The washed cells were harvested by centrifugation, and the cytochrome c was extracted by stirring the cell paste with 0.5 volumes of 1M NaCl, 0.1M sodium phosphate, pH 7.2, and 0.25 volumes ethyl acetate for 16-24 hours. Cell debris was removed by centrifugation, and the resulting supernatant dialyzed overnight against 0.1 M sodium phosphate, pH 7.2. The dialyzed cell extract was applied to a CM-Sepharose CL-6B cation exchange column (40 cm x 2.5 cm) equilibrated with 0.1 M sodium phosphate, pH 7.2. After the extract was loaded onto the column, the cytochrome c was oxidized by running 25 mL of a 5 mM solution of potassium ferricyanide in the loading buffer over the column. The cytochrome c was eluted using an 800 mL linear gradient of NaCl from 0 to 0.5 M in the running buffer. Fractions containing cytochrome c were collected, concentrated by ultrafiltration in

an Amicon 8050 stirred ultrafiltration cell using YM3 membranes and then further purified and desalted by gel filtration on a Sephadex G-75 column (45 cm x 1 cm). The final product was reconcentrated by ultrafiltration to approximately 1 mM and stored at -70° C.

Molecular Modeling Studies

Cytochrome c variants were designed with the aid of the computer graphics package Biograf (Molecular Simulations Inc., Pasadena Ca.). Using the coordinates of the refined structure of *S. cerevisiae* iso-1-cytochrome c,[2,3] potential mutation sites on the surface of the protein were located, and the native amino acid replaced by histidine. If warranted, the imidazole side chain was rotated about the α–β and β–γ carbon-carbon bonds to increase solvent exposure. The accessibility of surface histidines was estimated from the surface area of the ε-nitrogen available to a probe the size of Cu(II)IDA (3Å diameter).[27-29] The accessibility is reported as the percent surface area available relative to that of an unhindered histidine.

UV-Visible Spectroscopy

UV-Visible spectra were recorded at room temperature on a Milton Roy Spectronic 3000 diode-array spectrophotometer using a 1 cm path length semi-micro quartz cuvette. Protein samples were diluted in 50 mM phosphate buffer, pH 7.0, to a concentration of approximately 6 μM. Absorption spectra (250-600 nm) were taken in both the oxidized and the reduced state (prepared by the addition of a small amount of sodium dithionite). A concentrated solution of protein (50-100 μM) was used to observe the charge transfer band at 697 nm in the oxidized state.

NMR Spectroscopy

NMR data were collected at 37° C using AM 500-MHz and AMX 500-MHz Bruker spectrophotometers. Samples were prepared using a modified version of the method of

Pielak.[15] Purified cytochrome c was diluted 20-fold into buffer, 0.1 M sodium phosphate, 0.1M NaCl, pH 4.0 and then reconcentrated to approximately 0.5 mM using an Amicon 8050 stirred ultrafiltration cell with a PM3 membrane. Samples were then lyophilized, resuspended in D_2O and exchanged overnight at 37° C. Exchanged samples were lyophilized again and finally resuspended to a concentration of 0.5 mM in 100 percent D_2O under nitrogen. The volume of each sample was 400 μL, and the final pH was adjusted to approximately 4.5 (actual reading) using 0.1 M sodium phosphate, 0.1M NaCl, pH 3.0, in D_2O. Titration curves were prepared by taking NMR spectra at intervals of approximately 0.3 pH units from pH 4.5 to 8.0. The pH was adjusted by the addition of 0.1 M sodium phosphate, 0.1 M NaCl, pH 12 in D_2O. All samples contained approximately 100 μM of TSP (Na-3-trimethylsilylpropionate [2,2,3,3-d_4]), and chemical shifts are reported in ppm against this standard.

Results and Discussion

Cytochrome c Variants

The estimated solvent accessibility of the histidine positions selected for this study are given in Table 2.1. The crystal structure indicates that the native histidine at position 26 is involved in hydrogen bonding to the main chain amide of Asn 31 and the main chain carbonyl of Glu 44 and is almost completely inaccessible to a probe the size of Cu(II)IDA (3Å).[3] The native histidines at positions 33 and 39 are more accessible. The alternate sites for histidine placement (positions 4, 8, and 58) were chosen for their high degree of accessibility at the surface. These six sites all lie on the same face of the protein, and thus are not structurally precluded from forming simultaneous interactions at a solid surface.

Table 2.2 lists the cytochrome c variants used in this study. The label for each variant indicates its surface histidines. The mutations made to the wild-type protein and the total number of surface histidines is also given. All variants contain a serine at position 102

instead of the native cysteine, thereby preventing the dimerization that occurs in the wild-type protein.[17] The control protein for this study, H(-), was used to investigate binding in the absence of surface histidines. A set of three proteins containing one surface histidine at different locations (positions 4, 8, and 26) was used to study the effects of histidine accessibility and pK_a on binding affinity. The effect of multiple surface histidines on binding affinity was investigated using variants containing either two (positions 26 and 4, 8, or 33) or three (positions 26, 33 and 4, 8, 39 or 58) surface histidines. The final two variants, $H_{26}H_{33}H_4H_8$ and $H_{26}H_{33}H_{39}H_{58}$, were designed specifically to contain high-affinity binding sites on their surfaces, as will be discussed in detail in chapter 5.

The structural integrity and histidine content of each of the mutants was confirmed by several independent techniques. The production system in yeast is designed to guarantee that all variants are functional and thus able to carry out electron transfer.[17,30] This alone ensures that all variants are folded into essentially the native cytochrome c conformation. The conformation of the variants was further verified by UV-Vis and NMR spectroscopy as described below. Variants were purified in the oxidized form and used in this state throughout the study.

UV-Visible Spectroscopy

The structural integrity of the protein was investigated by UV-Vis spectroscopy. Figure 2.2 shows the absorbance spectrum of a typical member of the cytochrome c family.[31,32] Absorbance in the visible region is dominated by the heme group.[33] This produces bands at 416 nm (Soret band) and at 521 and 550 nm (β and α bands, respectively) for iso-1-cytochrome c in the reduced state. In the oxidized state the Soret band shifts to 410 nm and is reduced in intensity, while the α and β bands become a single broad peak. These bands have all been attributed to $\pi - \pi^*$ transitions occurring primarily

on the porphyrin ring of the heme.[33] Also present in the oxidized protein spectrum, although much less intense, is a charge transfer (CT) band at 697 nm.[33]

The conformation of cytochrome c can be monitored by the presence of the CT band at 697 nm.[34] This band is diagnostic of methionine coordination to the iron and is sensitive to the conformational state of the protein.[33] Minor structural perturbations or a change in the polarity of the heme environment can be seen from changes in the maximum and shape of the α band at 550 nm[33] and has been used to classify similar types of cytochromes. The existence and position of these two bands have been used to characterize the cytochrome c variants.

The Soret region of the spectra for each of the variants showed the characteristic peaks at 410 nm in the oxidized state and 416 nm in the reduced state. Coordination of the methionine ligand was confirmed by the presence of the charge transfer band at 697 nm in all variants. The α band maximum at 549.6 nm was identical to the wild-type cytochrome c and invariant among all mutants. The surface mutations did not result in any discernible alterations in the visible absorption spectrum compared to the wild-type, suggesting that the structures of the variants are unaffected by the histidine substitutions.

The purity of the cytochrome preparations was also examined by the ratio of absorbances at 280 nm and 410 nm in the oxidized state. All preparations were of equal purity or greater than that of the wild-type protein provided by Sigma (greater than 95 percent). Each preparation was also confirmed to be in the oxidized state from the placement of the Soret band at 410 nm and the absence of the α band at 550 nm.

NMR Spectroscopy

^{1}H NMR spectroscopy was used to investigate the surface histidines of the iso-1-cytochrome c variants. The ring proton NMR signals of cytochrome c surface histidines are readily observed in the downfield region of the spectrum where the C2 proton of the

imidazole ring (between the two nitrogens) usually resonates at a higher frequency than the C4 proton.[33] The pK_a of each histidine can be determined by monitoring the change in its chemical shift with pH.[33,35] In this study we used the titration behavior of the ring protons to both confirm the existence of mutations in the variants and to determine their pK_a's.

The NMR spectra for the pH titration of the $H_{26}H_{33}H_{39}$ variant are shown in Figure 2.3. The assignments for the titrating C2 protons of histidines 26, 33, and 39 were made by comparison to the published NMR data on iso-1-cytochrome c.[16,36,37] The C2 titration data for histidines 33 and 39 of $H_{26}H_{33}H_{39}$ are highlighted in Figure 2.4. The data fit standard pH titration curves and yield pK_a's of 6.37 and 6.54 for histidines 33 and 39, respectively.

Similar studies were performed on all cytochrome c variants, and the results are given in Appendix A. Comparing the titration behavior of each of the variants provided a method for confirming the identity of specific histidines in each variant. The presence of histidine 26 was confirmed from the non-titrating resonance at 7.71 ppm. Representative titration behavior of the C2 resonances for each of the other histidines are compared in Figure 2.5. With the exception of the metal-chelating variants, where two histidines are close to one another in space, the behavior of each of these resonances was similar in the different variants. The C2 resonances of histidines 33, 39 and 58 all exhibited standard pH titration behavior and were easily identified at low pH, where their terminal resonances are 8.81, 8.71, and 8.60 ppm, respectively. The resonances of histidines 4 and 8 were easily identified by their non-standard pH transitions.

Each histidine exhibits a unique C2 chemical shift at pH 5.6. Figure 2.6 displays the downfield region of the ^1H NMR spectra for selected cytochrome c variants at pH 5.6. It is clear from this figure that NMR at this pH provides a simple means to identify the histidines present in each of the variants.

Table 2.3 contains a listing of the mutants and the pK_a's of their surface histidines. Due to its involvement in hydrogen bonding, histidine 26 does not titrate appreciably in the pH range studied. This was not unexpected since the pK_a of histidine 26 has been reported to be less than 3.0.[33] The pK_a's for histidine 33 ranged from 6.37 to 6.53 among the different variants studied. The presence of histidines at positions 39 and 58 seemed to reduce the pK_a of histidine 33, while the introduction of histidines in the N-terminal helix at positions 4 and 8 seemed to increase its pK_a. The consistency of the titration behavior of histidine 33 indicates that its environment remains essentially unchanged among the variants where it is present.

Non-standard pH transitions were observed for histidines 4 and 8. The values reported in Table 2.3 for histidines 4 and 8 are apparent pK_a's, representing the point at which the histidine is 50 percent protonated. The apparent pK_a's of histidines 4 and 8 were 5.9 and 5.4, respectively, when only one was present in a variant. The unusual titration behavior of histidines 4 and 8 has not been explained, but may be related to their location in the N-terminal α-helix of cytochrome c.

The pK_a's for histidines 39 and 58 ranged from 6.51 to 6.55. Residues 39 and 58 are close to each other in the folded stucture, and the titration behavior was significantly altered in the metal-chelating variant which contained both histidines, $H_{26}H_{33}H_{39}H_{58}$. Similar alterations were observed in the titration behavior of histidines 4 and 8 in the $H_{26}H_{33}H_4H_8$ variant. The altered titration behavior of variants with adjacent histidines could be a result of hydrogen bonding between the two residues or simply due to the proximity of the two titrating groups.

The measured pK_a's of histidines 33 and 39 are approximately 0.3 pH units lower than those reported in the literature for the native *S. cerevisiae* iso-1-cytochrome c.[33,36,37] This is likely due to the difference in NMR acquisition temperatures. Although previous studies have been performed at 25° C, we found that it was necessary to perform our

experiments at 37° C to resolve the C2 proton resonances for each of the histidines. The influence of temperature on the pK_a may be determined from the Clausius-Clapeyron equation,

$$\ln (K_a / K_{ao}) = - \Delta H/R \, (1/T - 1/T_o)$$

where K_a and K_{ao} are the proton association constants at the absolute temperatures T and T_o, respectively. The ΔH for the protonation of imidazole is -8.8 kcal/mol (25° C, 0.2M ionic strength) and can be used to approximate effect of temperature on pK_a. This equation predicts that a 12 degree increase in temperature results in a 0.25 unit decrease in the pK_a of imidazole. This is consistent with the 0.3 unit drop in pH that was observed for the histidines at positions 33 and 39 at 37° C.

Finally, the titration of histidine 33 in horse cytochrome c was performed in both D_2O and H_2O. The resulting pK_a's were 6.19 and 6.20, respectively (Appendix A). This indicates that the presence of D_2O had little effect on the measured pK_a's in this set of experiments.

Conclusions

A set of iso-1-cytochrome c variants which contain histidines at 6 different surface positions has been designed, produced, and characterized. The ability of these variants to support growth on a non-fermentable carbon source suggests that the protein conformation is not substantially altered by these mutations. The structural integrity was investigated by UV-visible absorbance, and the histidine content was confirmed using ^1H NMR spectroscopy. All data suggest that the critical structural characteristics of the native cytochrome c are retained in the variants.

Table 2.1 Calculated accessible surface area of selected histidines on the surface of iso-1-cytochrome c. Calculations were performed using coordinates from the high-resolution crystal structure of *S. cerevisiae* iso-1-cytcochrome c (1.23 Å).[3] The surface area is reported as the area, at the e-nitrogen, accessible to a probe the size of Cu(II)IDA (3 Å). The percent accessibility is the accessible surface area relative to an unhindered imidazole ring.

Histidine Position	**Accessible Surface Area (Å^2)**	**Percent Accessibility**
26	0.50	7
33	2.25	33
39	4.75	70
58	5.50	82
4	6.75	100
8	6.75	100

Table 2.2 Engineered variants of *S. cerevisiae* iso-1-cytochrome c. All variants contain the replacement of cysteine by serine at position 102 to prevent dimerization. Remaining mutations alter histidine content on the surface of cytochrome c. The total number of surface accessible histidines is also given (includes histidine 26).

Label	**Mutations**	**Number of Surface Histidines**
H(-)	C102S, H39Q, H33N, H26N	0
H_{26}	C102S, H39Q, H33N	1
H_4	C102S, H39Q, H33N, H26N,K4H	1
H_8	C102S, H39Q, H33N, H26N,T8H	1
$H_{26}H_4$[a]	C102S, H39Q, H33Q, K4H	2
$H_{26}H_8$[a]	C102S, H39Q, H33Q, T8H	2
$H_{26}H_{33}$[a]	C102S, H39Q	2
$H_{26}H_{33}H_{39}$[a]	C102S	3
$H_{26}H_{33}H_4$	C102S, H39Q, K4H	3
$H_{26}H_{33}H_8$	C102S, H39Q, T8H	3
$H_{26}H_{33}H_{58}$[a]	C102S, H39Q, L58H	3
$H_{26}H_{33}H_{39}H_8$	C102S, T8H	4
$H_{26}H_{33}H_{39}H_{58}$[b,c]	C102S, L58H	4
$H_{26}H_{33}H_4H_8$[b]	C102S, H39Q, K4H, T8H	4

a Genes containing these mutations provided by J. H. Richards, California Institute of Technology

b Engineered high-affinity metal-binding variants

c Genes containing these mutations provided by J. T. Kellis, California Institute of Technology

Table 2.3 pK_a's of surface histidines in *S. cerevisiae* iso-1-cytochrome c variants measured by ^1H NMR (tabulated data from Appendix A). Mutations were confirmed by comparing titration behavior of downfield imidazole proton resonances. Histidine at position 26 was verified by the non-titrating proton resonance at 7.71 ppm.

Variant	**His 26**	**His 33**	**His 39**	**His 4**	**His 8**	**His 58**
H(-)	—	—	—	—	—	—
H_{26}	+	—	—	—	—	—
H_4	—	—	—	5.88[a]	—	—
H_8	—	—	—	—	5.39[a]	—
$H_{26}H_4$	+	—	—	5.86[a]	—	—
$H_{26}H_8$	+	—	—	—	5.40[a]	—
$H_{26}H_{33}$	+	6.48	—	—	—	—
$H_{26}H_{33}H_4$	+	6.52	—	5.90[a]	—	—
$H_{26}H_{33}H_8$	+	6.50	—	—	5.43[a]	—
$H_{26}H_{33}H_{39}$	+	6.37	6.54	—	—	—
$H_{26}H_{33}H_{58}$	+	6.33	—	—	—	6.55
$H_{26}H_{33}H_{39}H_8$	+	6.44	6.53	—	5.48[a]	—
$H_{26}H_{33}H_{39}H_{58}$	+	6.36	6.52[a,b]	—	—	6.51[a,b]
$H_{26}H_{33}H_4H_8$	+	6.48	—	5.91[a,b]	5.73[a,b]	—

\+ Resonance at 7.71 ppm confirmed

a Non-standard pH transitions (apparent pKa)

b Close proximity of histidines disturbs titration behavior

Figure 2.1 Illustration of the crystal structure of *S. cerevisiae* iso-1-cytochrome c.[2,3] The protein backbone is displayed as a green ribbon and the prosthetic heme group is shown in red. The histidine at position 18 (yellow) is ligated to the heme iron. Native histidines at positions 26, 33 and 39 (purple) are surface accessible.

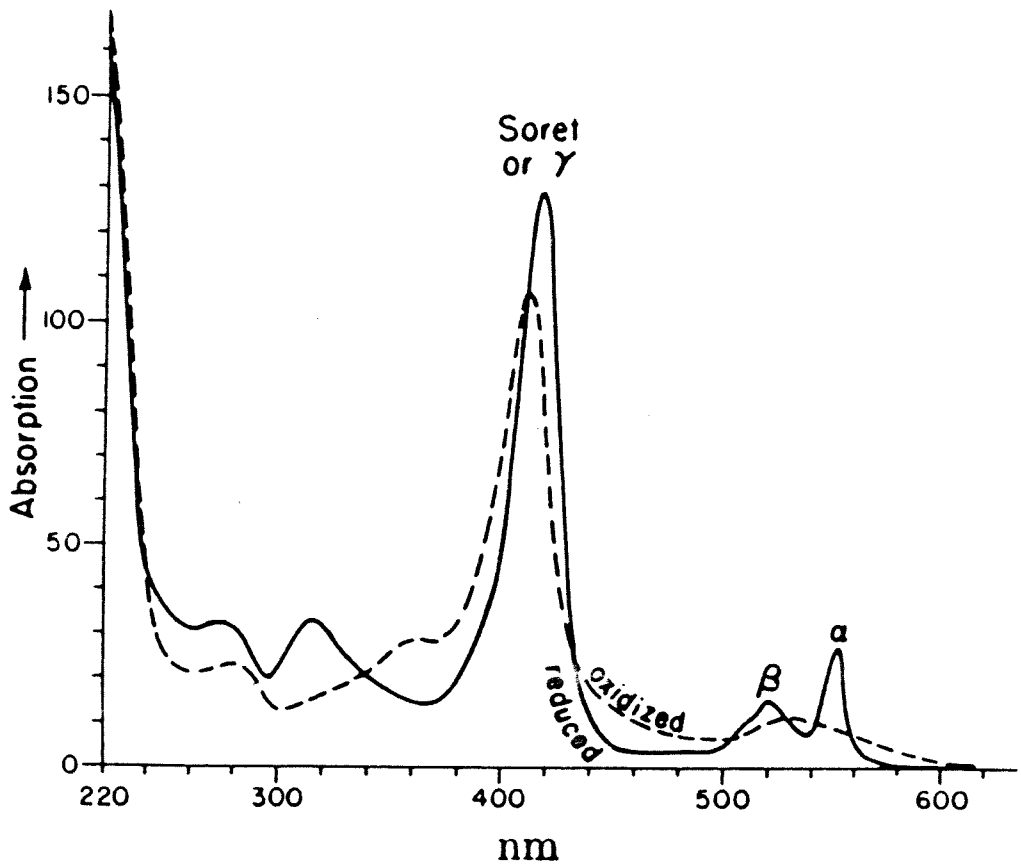


Figure 2.2 UV-Visible absorption spectrum of reduced and oxidized cytochrome c.[32,33] Absorbance in the visible region is dominated by the heme group. Cytochrome c shows characteristic three-band absorption spectrum in the reduced state with an α band at 550 nm, a β band at 521 nm, and a γ or Soret band at 516 nm. Oxidized cytochrome c has a single characteristic Soret band at 410 nm.

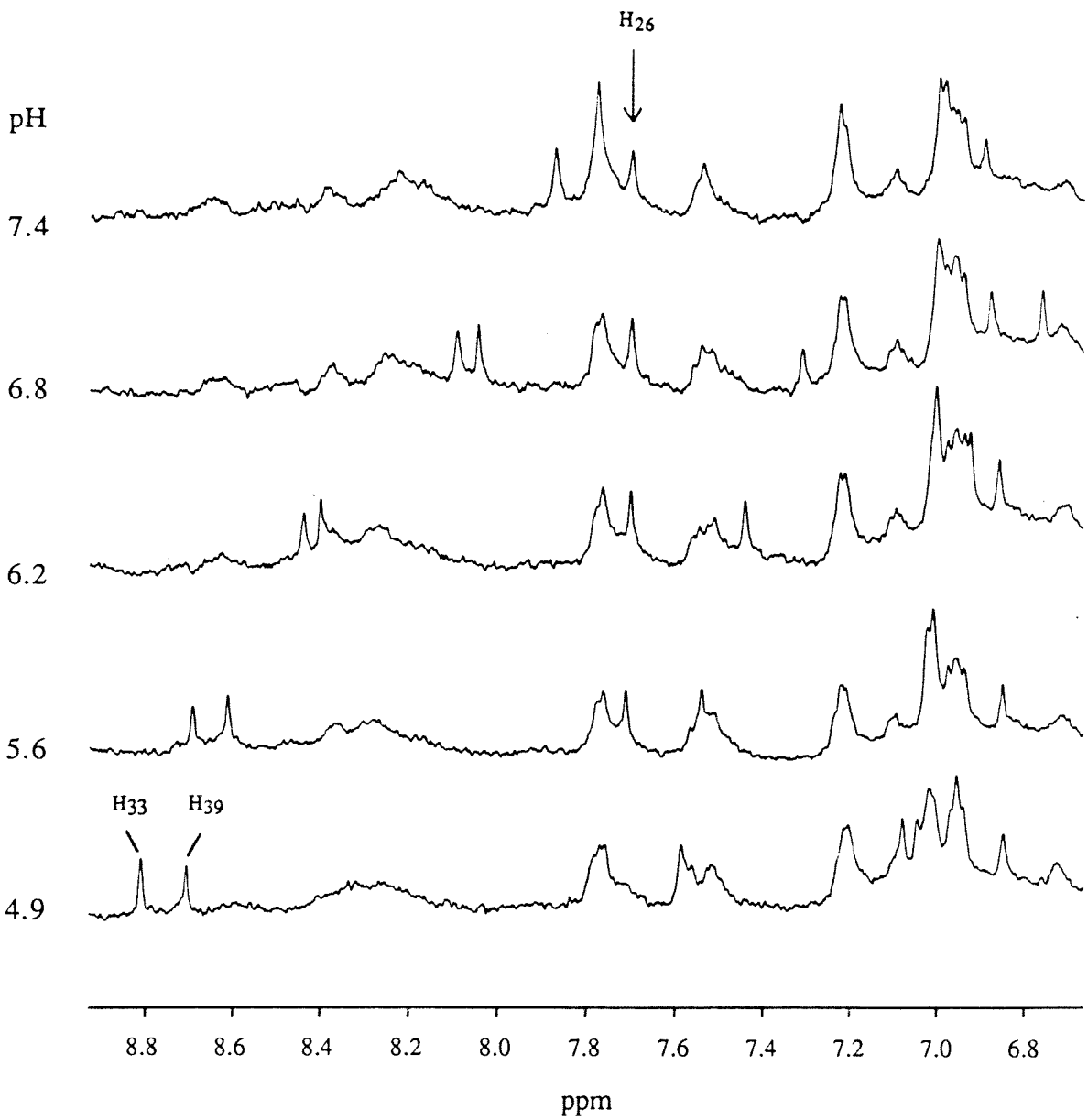


Figure 2.3 ^{1}H NMR spectra from the pH titration of iso-1-cytochrome c ($H_{26}H_{33}H_{39}$). The C2 protons of histidines 33 and 39 can be identified at 8.81 and 8.71 ppm at low pH, respectively. These resonances shift upfield as the pH is increased and terminate at 7.81 and 7.71 ppm. The C2 proton of histidine 26 can be observed at 7.71 ppm above pH 5.2.

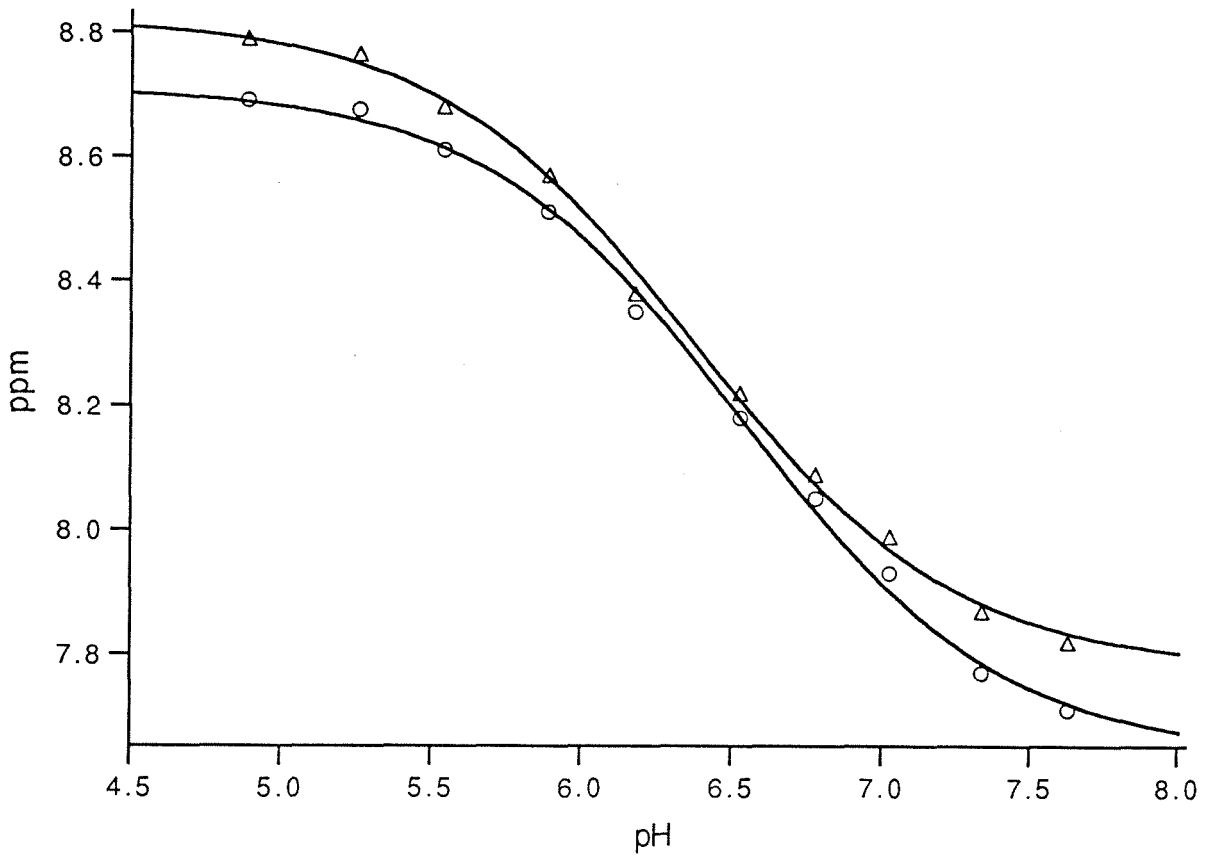


Figure 2.4 pH titration of *S. cerevisiae* iso-1-cytochrome c ($H_{26}H_{33}H_{39}$). The chemical shift in ppm of the C2 protons of histidines 33 and 39 are plotted versus the measured pH. The data are fit with the standard pH titration curve such that ppm = (ppm1 + ppm2 * $10^{(pH-pK_a)}$ / (1+ $10^{(pH-pK_a)}$). The measured pK_a's of histidines 33 and 39 are 6.39 and 6.54, respectively: (Δ) histidine 33, (O) histidine 39.

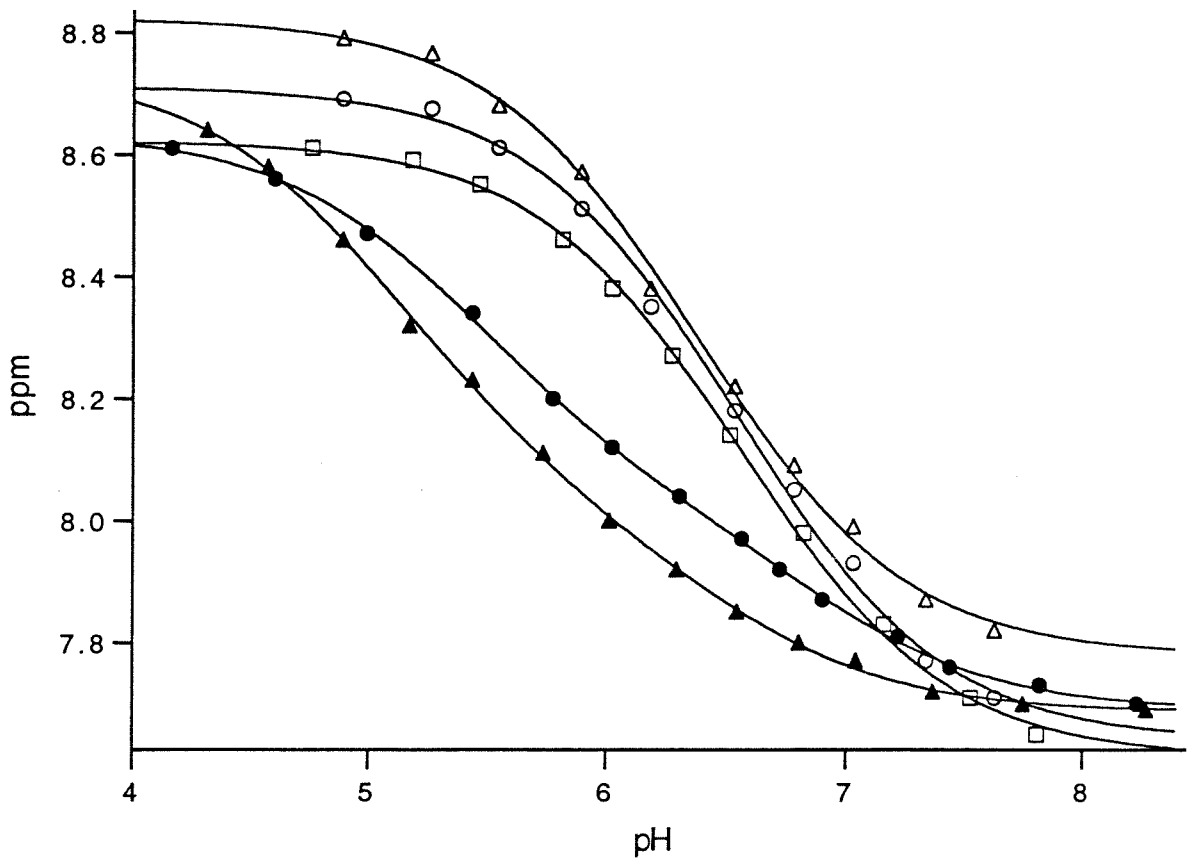


Figure 2.5 pH titration behavior of histidine at selected positions in iso-1-cytochrome c. The chemical shifts in ppm of the C2 protons of histidines at positions 4, 8, 33, 39 and 58 are shown. Histidines 33, 39, and 58 are fit with standard pH titration curves and have low pH resonances of 8.81, 8.71, and 8.60 ppm, respectively. Histidines 4 and 8 can be identified by their non-standard titration behavior. The titration data are obtained from the following variants: (Δ) histidines 33 and (O) histidine 39 from $H_{26}H_{33}H_{39}$, ( ) histidine 58 from $H_{26}H_{33}H_{58}$, (●) histidine 4 from $H_{26}H_{33}H_4$, and (▲) histidine 8 from $H_{26}H_{33}H_8$.

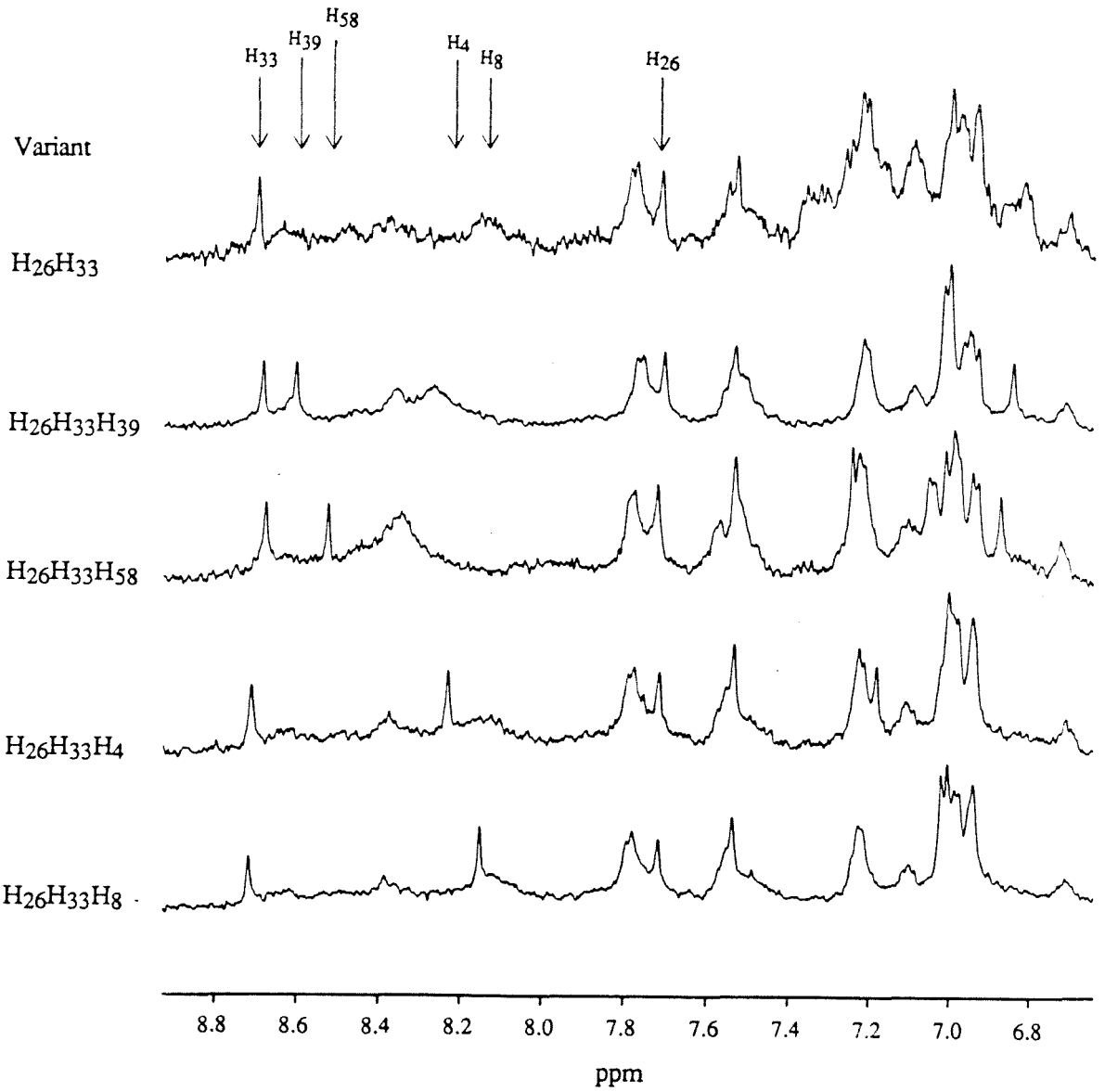


Figure 2.6 ^{1}H NMR spectra of selected iso-1-cytochrome c variants at pH 5.6. NMR spectroscopy of iso-1-cytochrome c variants at pH 5.6 provides a method for verification of histidine residues. The C2 protons of titrating histidines have distinct resonances at pH 5.6: histidine 33 (8.70 ppm), histidine 39 (8.60 ppm), histidine 58 (8.53 ppm), histidine 4 (8.23 ppm), and histidine 8 (8.16 ppm). The non-titrating C2 resonance of histidine 26 can be identified at 7.71 ppm.

References

1. Williams, G., Moore, G. R. and Williams, R. J. P., *Comments on Inorganic Chemistry* **4**, 55-98 (1985).
2. Louie, G. V., Hutcheon, L. B. and Brayer, G. D., *Journal of Molecular Biology* **199**, 295-314 (1988).
3. Louie, G. V. and Brayer, G. D., *Journal of Molecular Biology* **214**, 527-555 (1990).
4. Takano, T. and Dickerson, R. E., *Journal of Molecular Biology* **153**, 79-94 (1981).
5. Takano, T. and Dickerson, R. E., *Journal of Molecular Biology* **153**, 95-115 (1981).
6. Ochi, H., Hata, Y., Takana, N., Nakudo, M., Sakurai, T., Aihara, S. and Morita, Y., *Journal of Molecular Biology* **166**, 407-418 (1983).
7. Dickerson, R. E., Takano, T., Eisenberg, D., Kallai, O. B., Samson, L., Cooper, A. and Margoliash, E., *Journal of Biological Chemistry* **246**, 1511-1535 (1971).
8. Wand, A. J. and Englander, S. W., *Biochemistry* **24**, 5290-5294 (1985).
9. Wand, A. J. and Englander, S. W., *Biochemistry* **25**, 1100-1106 (1986).
10. Wand, A. J., Roder, H. and Englander, S. W., *Biochemistry* **25**, 1107-1114 (1986).
11. Williams, G., Clayden, N. J., Moore, G. R. and Williams, R. J. P., *Journal of Molecular Biology* **193**, 447-460 (1985).
12. Moore, G. R. and Williams, R. J. P., *European Journal of Biochemistry* **103**, 523-532 (1980).
13. Moore, G. R. and Williams, R. J. P., *European Journal of Biochemistry* **103**, 533-541 (1980).
14. Boswell, A. P., Moore, G. R., Williams, R. J. P., Chien, J. C. W. and Dickenson, L. C., *Journal of Inorganic Biochemistry* **13**, 347-352 (1980).
15. Pielak, G. J., Boyd, J., Moore, G. R. and Williams, R. J. P., *European Journal of Biochemistry* **177**, 167-177 (1988).
16. Gao, Y., Boyd, J., Williams, R. J. P. and Pielak, G. J., *Biochemistry* **29**, 6994-7003 (1990).
17. Cutler, R. L., Pielak, G. J., Mauk, A. G. and Smith, M., *Protein Engineering* **1**, 95-99 (1987).

18. Smith, M., Leung, D. W., Gillam, S., Astell, C. R., Montgomery, D. L. and Hall, B. D., *Cell* **16**, 753-761 (1979).
19. Hampsey, D. M., Das, G. and Sherman, F., *FEBS Letters* 231, 275-283 (1988).
20. Smith, E. L., in *The Enzymes*, P. D. Boyer, Ed. (Academic Press, London, 1970), vol. 1, pp. 267-339.
21. Sayers, J. R., Schmidt, W. and Eckstein, F., *Nucliec Acids Research* **16**, 791-802 (1988).
22. Kunkel, T. A., Roberts, J. D. and Zakour, R. A., *Methods in Enzymology* **154**, 367-382 (1987).
23. Sanger, F., Nicklen, S. and Coulsen, A. R., *Proceedings of the National Acadamy of Sciences U.S.A.* **74**, 5463-5467 (1977).
24. Hicks, J. B., Strathern, J. N., Klar, A. J. S. and Dellaporta, S. L., in *Genetic Engineering:Principles and Methods*, J. K. Setlow and A. Hollaender, Ed. (Plenum, New York, 1982) vol. 4, pp. 219-248.
25. Faye, G., Leung, D. W., Tatchell, K., Hall, B. D. and Smith, M., *Proceedings of the National Acadamy of Sciences U.S.A.* **78**, 2258-2262 (1981).
26. Ito, H., Fukuda, Y., Murata, K. and Kimura, A., *Journal of Bacteriology* **153**, 163-168 (1983).
27. Dung, N., Viossat, B., Busnot, A., Perez, J. M. G., Gutierrez, J. N. and Gardette, F., *Inorganica Chimica Acta* **174**, 145-148 (1990).
28. Dung, N., Viossat, B., Busnot, A., Zafra, A. G. S., Perez, J. M. G. and Gutierrez, J. N., *Inorganica Cimica Acta* **169**, 9-12 (1990).
29. Wuenschell, G. E., Naranjo, E. and Arnold, F. H., *Bioprocess Engineering* **5**, 199-202 (1990).
30. Sherman, F., Stewart, J. W., Jackson, M., Gilmore, R. A. and Parker, J. H., *Genetics* **77**, 255-284 (1974).
31. Dickerson, R. E. and Timkovich, R., in *The Enzymes*, P. D. Boyer, Ed. (Academic Press, London, 1975), vol. 11, pp. 397-547.
32. Margoliash, E. and Frohwirt, N., *Biochemical Journal* **71**, 570-572 (1959).
33. Moore, G. R. and Pettigrew, G. W., in *Cytochrome c: Evolutionary, Structural and Physicochemical Aspects* (Springer-Verlag, Berlin, 1990).
34. Schejter, A. and George, P., *Biochemistry* **3**, 1045-1049 (1964).
35. Cohen,J. S. and Hayes, M. B., *Journal of Biological Chemistry* 249, 5472-5477 (1974).

36. Cutler, R. L., Davies, A. M., Creighton, S., Warshel, A., Moore, G. R., Smith, M. and Mauk, A. G., *Biochemistry* **28**, 3188-3197 (1989).

37. Robinson, M. N., Boswell, A. P., Huang, Z., Eley, C. G. S. and Moore, G. R., *Biochemical Journal* **213**, 687-700 (1983).

Chapter 3

IMAC Studies I : Equilibrium Binding Isotherms

Introduction

Immobilized metal affinity chromatography (IMAC) is a separation technique which utilizes the interaction of proteins with immobilized metal ions as the basis for protein purification. IMAC has been shown to separate proteins based on the multiplicity and microenvironment of surface-accessible histidines,[1,2] although the precise molecular mechanisms by which this occurs are not well understood. In this study, a set of iso-1-cytochrome c variants will be used to evaluate how the number and placement of surface histidines affects the equilibrium binding isotherms on matrices containing immobilized Cu(II)IDA. The equilibrium binding isotherm contains fundamental information concerning the binding affinity and maximum capacity of a solute on a chromatographic matrix and can used evaluate surface interactions.[3] Equilibrium binding isotherms in this study are measured in the absence of any significant competitors at a pH where the histidines are effectively deprotonated (greater than 80 percent), and therefore, represent the inherent affinity of the cytochrome c variants for the Cu(II)IDA matrix.

Cu(II)IDA is the most commonly used metal ion complex in IMAC and provides a simple model system with which the effect of protein surface properties on IMAC retention can be studied. The metal-affinity ligand (Cu(II)IDA) is formed by the very strong complexation of Cu(II) by IDA (log K = 11.3),[4] in which the positive charge of the copper ion is neutralized by the acid groups of IDA and the resultant Cu(II)IDA complex has three coordination sites open for interaction with the protein. The Cu(II)IDA is attached to a hydrophilic macroporous polymer support matrix which is amenable to both batch equilibrium isotherm measurements and to HPLC chromatographic analysis. Batch equilibrium binding isotherms will be used to provide fundamental quantitative information about the protein-Cu(II)IDA interaction in this chapter. The translation of this equilibrium binding information into chromatographic retention behavior will be addressed using HPLC in Chapter 4.

Materials and Methods

Metal-Affinity Matrix

TSK guardgel Chelate-5PW (17 μm macroporous beads) was purchased from Tosohaas (#08648). The TSK matrix is polymer based and hydrophilic in nature. Although the surface area in the swollen state is not known, the manufacturer reports that the TSK beads have a pore size of 1000 Å. Iminodiacetic acid (IDA) is the immobilizing ligand and is attached to the TSK matrix via a long spacer arm.

Regeneration of TSK Cu(II) Matrix for Binding Isotherms

The TSK guardgel Chelate-5PW was packed into a column (1 cm x 10 cm) and washed extensively with a 50 mM solution of EDTA, pH 8.0. The column was loaded with copper by equilibrating with a solution of 50 mM $CuSO_4$ in water. Excess copper was removed by washing with 0.1 M sodium acetate buffer, pH 4.0. The matrix was finally equilibrated in 50 mM sodium phosphate, 0.5 M NaCl at pH 7.0 for use in the binding isotherm studies.

Equilibrium Binding Isotherms

The equilibrium binding isotherms were measured using a modified version of the procedure developed by Hutchens *et al.*[5] The TSK gel was regenerated as given above and then resuspended in a 25 percent solution (packed volume/total volume) of the equilibration buffer (50 mM sodium phosphate, 0.5 M NaCl at pH 7.0). Dilutions of cytochrome c ranging from 0.4 mM-.004 mM were prepared in the same buffer, and 800 μL of each protein solution was added to 200 μL of the gel suspension in a microfuge tube. This suspension was equilibrated for 30 minutes by inversion and then spun at 10,000 rpm in a microfuge for 5 minutes. The supernatant was transferred to a fresh tube, and the protein concentration determined. All cytochrome c concentrations were determined in

duplicate by first reducing the protein with sodium dithionite, and then measuring the visible absorbance at 550 nm (E_{550} = 27.6 cm^{-1} mM^{-1}).[6] The total volume of packed gel in each tube was 50 µL (200 µL of a 25 percent solution) of which approximately 60 percent was accessible to the protein solution (data not shown). Therefore, the final liquid volume was 980 µL (950 µL + 0.6 x 50 µL). The amount of bound protein in each sample (50 µL gel) is the difference between the amount input (800 µL x C_O) and the amount recovered in the supernatant (980 µL x C_L), where C_O is the concentration of the protein added and C_L is the concentration in the final supernatant. The amount of protein bound per mL of packed gel, Q_p, is then given by,

$$Q_p = 20\ (0.8\ C_O - 0.98\ C_L)$$

Measurement of Copper Loading

The copper loading of the TSK guardgel Chelate-5PW matrix was measured using the visible absorbance of EDTA-Cu at pH 7.0 (E_{800} = 73.6 cm^{-1} M^{-1}). Copper was removed by repeated washing with 50 mM EDTA at pH 7.0. The EDTA solutions were collected and the visible absorbance at 800 nm was determined. The copper loading of the TSK Chelate-5PW was 18.5 µmol copper per mL of packed gel.

Mathematical Modeling

The equilibrium isotherm plays a crucial role in predictive modeling and interpretation of adsorption systems.[3,7] The classical Langmuir theory for gas adsorption can be applied to the binding of dilute protein solutions at a surface.[7] The following assumptions are implicit in the Langmuir model;

1. Only one molecule can be adsorbed per site (monolayer assumption)
2. Only one type of site is present (homogeneous surface)

3. No lateral interactions or cooperativity

4. Reversible adsorption

5. Only one species present (no competitive adsorption)

Application of the Langmuir theory results in a two-parameter expression for the isotherm,

$$Q_p = \frac{Q_{max} K_p C_L}{1 + K_p C_L} \quad (1)$$

where Q_p is the amount of adsorbed protein per mL of packed gel and C_L is the liquid phase protein concentration at equilibrium. The two parameters that describe the isotherm are the binding constant, K_p, and the maximum capacity of adsorbed protein, Q_{max}, as seen in Figure 3.1. This model can be used to interpret chromatographic behavior in a quantitative manner,[8] and accurately describes the interaction of a single-histidine protein with an IMAC support.

The simple Langmuir model is insufficient to describe the equilibrium isotherms of multiple-histidine variants. Multiple histidines give rise to equilibrium isotherms which indicate heterogeneous surface binding. The heterogeneous binding is presumably due to the formation of two-site interactions at the matrix surface. These isotherms can be modeled using a bi-Langmuir isotherm, [7,9]

$$Q_p = \frac{Q_{max1} K_{p1} C_L}{1 + K_{p1} C_L} + \frac{Q_{max2} K_{p2} C_L}{1 + K_{p2} C_L} \quad (2)$$

where K_{p1} and K_{p2} represent the binding constants of the strong and weak binding sites, respectively. The corresponding maximum capacities are then given by Q_{max1} and Q_{max2}. The bi-Langmuir model was used in a similar situation by Huang *et al.* to

describe the adsorption isotherms of conalbumin, myoglobin, and γ-globulin on TSK DEAE 5PW anion exchange matrix where multi-point attachment was expected.[10]

Results

Single-Histidine Variants

Figure 3.2 shows the equilibrium isotherms for the H(-) control, H_{26} and tuna cytochrome c. The H(-) variant has no surface-accessible histidines and exhibits no measurable affinity for the Cu(II)IDA matrix. The H_{26} yeast cytochrome c variant and tuna cytochrome c have only histidine 26 on their surfaces and have similar binding isotherms with binding constants of 3.5 x 10^3 M^{-1} and 2.4 x 10^3 M^{-1}, respectively. The estimated maximum capacities for tuna cytochrome c and the H_{26} variant are between 1.1 and 1.3 μmols per mL of gel. By comparing the isotherm of the H(-) variant to the H_{26} variant, the presence of a surface histidine is shown to be a prerequisite for interaction with the Cu(II)IDA matrix. In the absence of surface histidines, the contribution of the remaining surface residues in yeast cytochrome c does not result in adsorption onto the matrix. The equilibrium binding behavior of tuna cytochrome c compared to the H_{26} variant provides further insight into the binding mechanism of IMAC. These two proteins have similar structures and identical histidine contents, yet contain many surface residues which differ, including a tryptophan at position 33.[11,12] The similarity in their binding behavior suggests that histidine content and position dominates the interaction with Cu(II)IDA, and that the contribution of the surface tryptophan at position 33 is insignificant (see Chapter 4 for a detailed discussion of tryptophan participation in IMAC).

The equilibrium binding isotherms of three single-histidine cytochrome c variants are shown in Figure 3.3. The binding affinity of H_{26} is weak compared to the variants with fully accessible histidines at positions 4 and 8. The H_4 and H_8 variants have binding constants of 5.6 and 4.9 x 10^4 M^{-1}, respectively, as compared to 3.5 x 10^3 M^{-1} for the

H_{26} variant. The single-histidine isotherms are adequately described by the Langmuir model with maximum capacities of approximately 1.1 μmol per mL of gel. The reduced binding affinity of the H_{26} variant, by more than a factor of 10, is undoubtedly due to the steric hindrance of histidine at position 26, which has less than 10 percent the accessibility of histidines 4 and 8 as measured by molecular modeling (Table 2.1). This suggests that the accessibility of surface histidine residues plays an important role in the affinity of a single-histidine protein for immobilized metal-affinity matrices.

Figure 3.4 compares the equilibrium isotherms of imidazole and the H_8 cytochrome c variant. The imidazole isotherm has been modeled using equation 3 shown below,

$$Q_p = \frac{Q_{max} K_{I1} C_I \left(1 + 2 K_{I2} C_I\right)}{1 + K_{I1} C_I \left(1 + K_{I2} C_I\right)} \qquad (3)$$

where C_I is the free imidazole concentration in solution and K_{I1} and K_{I2} represent the binding of the first and second imidazoles to each copper site, respectively. The maximum capacity (Q_{max}) has been fixed as the maximum loading of copper in the matrix (18.5 μmol/mL gel). Using this model the binding constants measured for imidazole to Cu(II)IDA were 5.4 x 10^3 M^{-1} (K_{I1}) and 30 M^{-1} (K_{I2}). The first binding constant corresponds closely to the solution binding constant of imidazole for Cu(II)IDA measured potentiometrically (2.2 x 10^3 M^{-1} at 35° C).[13] The binding of the H8 variant is dominated by a single histidine interaction, as evidenced by the H(-) variant which has no measurable affinity for the matrix, yet the equilibrium binding isotherm of the H8 variant is significantly different than that of imidazole. The most notable difference occurs at maximum capacity, where cytochrome c occupies only 5-10 percent of the copper sites available to imidazole. In addition, the binding constant for a single-histidine protein interacting with immobilized copper (4.9 x 10^4 M^{-1}) is almost ten times larger than that of imidazole (5.4 x 10^3 M^{-1}).

Multiple-Histidine Variants

The equilibrium binding isotherms for the $H_{26}H_4$ and $H_{26}H_8$ cytochrome c variants are compared to the isotherms of variants that contain the single histidines individually in Figures 3.5 and 3.6. In both Figures 3.5 and 3.6, the initial slopes of the two-histidine isotherms are dramatically larger than the single-histidine variants. Binding behavior at intermediate and higher protein concentrations is similar, although the maximum capacity of the two-histidine variants is approximately 30 percent greater than the single-histidine variants. The shape of the two-histidine isotherms suggest that more than one type of binding is taking place on the surface of the matrix (heterogeneous sites).[7] This is emphasized using a Scatchard plot as seen in Figure 3.7, where the $H_{26}H_4$ and $H_{26}H_4$ isotherms are concave upwards, indicating at least two modes of binding behavior.[7] This behavior cannot be fit to a simple Langmuir model, and hence, the isotherms for the $H_{26}H_4$ and $H_{26}H_8$ variants were modeled using bi-Langmuir isotherms.

The addition of a second histidine to the surface of cytochrome c has two effects: the initial slope of the binding isotherm is dramatically increased (by more than a factor of 10), and the isotherm takes on a bimodal shape. These results, taken together, suggest that the two histidines are able to simultaneously coordinate different copper sites at the surface. The first binding constant may represent the formation of strong two-site attachments to the matrix, while the second binding constant represents single-site attachment. Although histidine 26 is a relatively weak binding site, its addition to the single-histidine variants (H_4 and H_8) significantly influences the binding isotherm presumably by allowing the formation of strong two-site attachments.

A direct comparison of the isotherms for the variants which contain two surface histidines is given in Figure 3.8. The first binding constants for the two-histidine variants are approximately 1 x 10^6 M^{-1}, while the second binding constants are similar to the binding constants of the single-histidine variants (1-5 x 10^4 M^{-1}). The total capacity for

each two-histidine variant is approximately 1.4 μmol per mL of gel. The isotherms for the three-histidine variants are given in Figure 3.9, where each variant contains histidines at positions 26 and 33 in addition to a third surface histidine. These isotherms also show at least two modes of binding behavior, one corresponding to a very strong binding site. The binding constants observed for strong binding sites in the three-histidine variants ranged from 5-10 x 10^6 M^{-1}, and the total capacity is approximately twice that of the single-histidine variants.

A summary of the binding constants and maximum capacities for the complete set of cytochrome c variants is given in Table 3.1. Table 3.1 clearly shows the strong dependence of the first binding constant (K_{p1}) on the number of surface histidines. In addition, for the multiple-histidine variants, the data indicate that a large portion of the surface sites are able to accommodate strong two-site binding (Q_{max1}/Q_{tot} > 30 percent in all cases).

Discussion

Modeling of Surface Interactions

If a better understanding of protein interactions with IMAC surfaces is going to be advanced, the underlying molecular mechanism must be elucidated. Toward this end, the equilibrium binding isotherms for the set of cytochrome c variants interacting with a Cu(II)IDA matrix have given us the following information.

1. Surface histidines are required for interaction with a Cu(II)IDA matrix.
2. Cytochrome c variants occupy only 5-10 percent of the available copper sites at maximum capacity.
3. The binding constants measured for variants containing a single fully-accessible histidine are much stronger than for imidazole.

4. Variants containing more than one surface histidine appear to bind via two-site binding.
5. A large fraction of the surface sites can accommodate two-site binding.

These results are all consistent with the simple model shown in Figure 3.10, which describes the interaction of a protein with an IMAC matrix. The fundamental assumption of this model is that the binding of a single protein can block access to a number of copper sites on the surface of the matrix. Therefore, maximum protein capacity is reached when a monolayer of protein is deposited on the surface and all the copper sites are either occupied or blocked. This model will be used in the following discussion in an effort to interpret the results presented in this chapter.

Histidine Interaction

As demonstrated by the data in Table 1.1,[14] Cu(II) prefers to coordinate the imidazole side chain of histidine over other potential ligands on a protein surface. Thus, it is expected that surface-accessible histidine side chains will be the primary ligand involved in binding immobilized copper. This has been verified by removing the surface-accessible histidines from *S. cerevisiae* iso-1-cytochrome c (H(-) variant) and thereby eliminating any measurable interaction with immobilized Cu(II)IDA. A similar result was recently published by Mrabet in which the only surface histidine of D-xylose isomerase from *Actinoplanes missourienis* was replaced by a lysine.[15] Removal of this histidine completely abolished retention on a Cu(II)IDA matrix. Zhao *et al.* also found that a surface histidine was required for chromatographic retention by comparing lysozymes from different species.[2] Peking duck lysozyme, which has no surface-accessible histidines, was eluted in the wash buffer from a Cu(II)IDA column. These findings provide direct

evidence that histidine is the predominant ligand in IMAC and suggest that histidine may be required for any significant interaction with an IMAC matrix.

Single Histidine Interactions

The binding isotherms of the two variants with single fully-accessible histidines (H_4 and H_8) gave binding constants of approximately 5 x 10^4 M^{-1} and maximum capacities that were 5-10 percent of the total copper loading. A similar result was obtained by Hutchens *et al.* using chicken egg white lysozyme, which also contains a single histidine on its surface.[5,16] Hutchens measured the equilibrium binding isotherm of chicken egg white lysozyme on a Cu(II)IDA TSK matrix and reported a binding constant of approximately 3 x 10^4 M^{-1} with a maximum capacity that was 15 percent of the total copper loading.[5,16] The similarity in the binding isotherms of single-histidine cytochrome c and single-histidine lysozyme proteins suggest a common mechanism of protein interaction in IMAC.

It has been clearly demonstrated that the interaction of a single-histidine protein with a Cu(II)IDA matrix is dominated by the imidazole side chain of its lone histidine.[2,15] Therefore, we would expect the binding affinity of a single-histidine variant to be similar to that of imidazole. If one examines the initial slopes of the binding isotherms of the H_8 variant and imidazole in Figure 3.4, they indicate that at low surface coverage, their affinity for the matrix is similar. Yet, the overall binding isotherm of the H_8 variant is very different than the isotherm of imidazole. The H_8 isotherm has approximately 10 percent of the maximum capacity and an apparent binding constant that is approximately 10-fold greater than imidazole. We have shown using chromatographic evidence, obtained by measuring the void volumes for cytochrome c and *N*-acetyl tryptophan, that essentially all of the pore space in the TSK matrix is accessible to cytochrome c (data not shown), and therefore, the

difference in isotherms is not a result of copper sites being located in pores that are inaccessible to the protein.

A possible explanation for the limited capacity and apparently high binding constants observed for the single-histidine proteins relative to imidazole can be obtained using the model illustrated in Figure 3.10. For a single-histidine protein, the interaction seen in Figure 3.10 can be modeled in the following manner,

$$P + nCu \Leftrightarrow PCu + \{(n-1)Cu\} \qquad K_1 = \frac{[PCu]}{[P]\,[Cu]}$$

where the protein (P) coordinates only a single copper (Cu) upon binding to the surface, but also blocks n-1 copper sites from further interaction (blocked sites designated by brackets). This model results in a Langmuir-type binding isotherm,

$$Q_P = \frac{Cu_{tot} K_1 C_L}{1 + n K_1 C_L} \qquad (4)$$

where Cu_{tot} is the total copper loading on the matrix.

Using equation 4 to fit the single-histidine data in Figure 3.3 ($Cu_{tot} \approx 18.5$ mmol Cu/mL of packed gel), the predicted binding constants (K_1) for the H_{26}, H_4 and H_8 variants are 0.2 x 10^3 M^{-1}, 3.5 x 10^3 M^{-1} and 2.7 x 10^3 M^{-1}, respectively (Table 3.2). The binding constants for the fully-accessible single-histidine variants (H_4 and H_8) are in line with the expected affinity of a single-histidine interaction with Cu(II) and are similar to the measured binding constant of imidazole (5.4 x 10^3 M^{-1}). The model also predicts that approximately 16 sites are blocked by a single cytochrome c binding to the matrix. This simple model provides an explanation for both the limited use of copper sites and the increased apparent binding affinities for the single-histidine proteins relative to imidazole.

Another possible explanation for the increased binding affinity of the single-histidine proteins relative to imidazole is that increased binding affinity could result from

additional weak coordination events occurring through other ligands on the protein surface (*e.g.*, *N*-terminus). A weak secondary interaction resulting in 1 kcal/mol of free energy would be sufficient to explain the observed 10-fold increase in apparent binding affinity for the single-histidine proteins relative to imidazole.

Multiple-Histidine Interactions

The introduction of multiple histidines into the surface of a protein can have an impact on its binding affinity in various ways, depending on the type of interaction with the IMAC matrix. If two histidines in a protein bind independently to the IMAC matrix, the apparent binding constant of the two-histidine protein would be approximately twice the binding constant of a single-histidine protein. This is simply due to the statistical increase in binding sites available on the protein. If, on the other hand, the two histidines are able to bind simultaneously to different copper sites on the matrix (two-site binding), the binding constant of the two-histidine protein could be greatly increased. This type of two-site binding has been termed "macro-chelation" by Hutchens *et al.*,[16] and can be analyzed in a manner analogous to chelation as described for inorganic complexes. The increase in the binding constant of the two-histidine protein will depend on how much enthalpy is gained from the second histidine interaction relative to entropic losses. The gain in enthalpy results from the favorable formation of the Cu(II)IDA-imidazole interaction while the loss in entropy is a result of the protein being further constrained on the matrix. If entropic losses were negligible, the apparent binding constant of the two-histidine protein would be approximately the square of the binding constants for a single-histidine protein. Of course, it is expected that entropic losses will reduce this value significantly.

Macro-chelation requires that the copper sites be densely packed on the matrix surface so that two-site interactions can form (Figure 3.11). This is consistent with the

proposed protein interaction model where protein binding blocks multiple copper sites on the surface. If the protein has two potential binding sites (surface histidines), then equation 3 can be modified to include the second histidine binding as follows,

$$P + nCu \Leftrightarrow PCu + \{(n-1)Cu\} \qquad K_1 = \frac{[PCu]}{[P]\,[Cu]}$$

$$PCu + \{(n-1)Cu\} \Leftrightarrow PCu_2 + \{(n-2)Cu\} \qquad K_2 = \frac{[PCu_2]}{[PCu]}$$

where PCu_2 represents a protein bound at two separate copper sites. The second binding constant (K_2) represents the apparent binding affinity of the second histidine side chain for a nearby copper site on the surface (zero-order reaction). Analysis of these equilibria result in the modified isotherm seen below.

$$Q_P = \frac{Cu_{tot} K_1 (1 + K_2) C_L}{1 + n K_1 (1 + K_2) C_L} \tag{5}$$

By assuming that only a portion the surface sites will accommodate strong two-site binding ($K_2 >> 1$), while the remainder of the sites will only accommodate single-site binding ($K_2 \approx 0$) (*i.e.,* heterogeneous binding), equation 5 can be modified to fit the bi-Langmuir model (equation 2), as seen below,

$$Q_P = \frac{\theta\, Cu_{tot} K_1 K_2 C_L}{1 + n K_1 K_2 C_L} + \frac{(1 - \theta)\, Cu_{tot} K_1 C_L}{1 + n K_1 C_L} \tag{6}$$

where θ represents the fraction of sites that can accommodate two-site binding. In this equation, K_1K_2 represents the apparent binding constant of the protein binding via two-site binding.

The binding isotherms of the $H_{26}H_4$ and $H_{26}H_8$ variants can be analyzed using equation 6 where the total copper loading ($Cu_{tot} \approx 18.5$ μmol Cu/mL of packed gel) and the binding constant due to the strong histidine ($K_1 = 3.5 \times 10^3$ M^{-1} for H_4 and 2.7×10^3 M^{-1} for H8) are already known. The values of K_2, θ, and n obtained from this analysis are given in Table 3.2. The addition of histidine 26 to the variants containing either histidine 4 or histidine 8 results in a greater than 20-fold increase in the apparent binding constant relative to single-site binding ($K_2 \geq 20$) for greater than 60 percent of surface sites ($\theta \geq 0.6$). Additionally, the values for n are approximately the same as for the single-histidine variants, with approximately 15 copper sites being blocked by each protein.

Direct support for the macro-chelation model could be obtained by preparing a set of IMAC supports with various copper densities. If the binding constant decreased with the density of copper sites (loss of two-site binding), while the capacity remained the same (limited only by the surface area), the macro-chelation hypothesis would be supported. This type of experiment may emphasize the important role that the copper density plays in metal-affinity chromatography. For example, if one wishes to separate a single-histidine protein from a two-histidine protein chromatographically, the separation will be governed by the ratio of the initial slopes of their respective binding isotherms (see Chapter 4).[8] From equations 4 and 6, it can be shown that this ratio is given by the quantity θK_2. Both these quantities are dependent on the density of copper sites on the surface. If the density of copper sites decreases, the fraction of sites able to accommodate two-site binding (θ) will be reduced. In addition, if the copper sites are spaced too far apart, the entropic losses associated with forming the second histidine binding site will be increased, resulting in a reduction in the apparent binding constant, K_2. The apparently high density of copper sites on this matrix provides for large values of θ and K_2 which translate into efficient separations, as will be demonstrated in Chapter 4.

Variations in Maximum Capacity

The complete set of data suggest that increasing either the binding strength or the number of surface histidines results in greater surface coverage. The total capacity for the two-histidine variants is about 30 percent greater than that of the single-histidine variants and the total capacity for the three-histidine variants is more than double. One explanation for this behavior is that the increased binding affinity resulting from additional surface histidine can overcome steric forces at the surface, thereby allowing tighter packing.

Contribution of Histidine 26 to Metal Binding

Previous studies have suggested that the inaccessibility of histidine 26 precluded its contribution to cytochrome c retention in immobilized metal-affinity chromatography.[1,17] Although the crystal structure of cytochrome c suggests that the imidazole ring nitrogens of histidine 26 are involved in hydrogen bonding and are inaccessible, we have shown that histidine 26 does contribute weakly to retention. By directly examining the effects of its removal by site-directed mutagenesis, we are able to conclusively show the binding observed for the H_{26} variant is due solely to histidine 26 and not to other specific or non-specific interactions. The contribution of histidine 26, despite its inaccessibility, can be justified by realizing that proteins are flexible, and the crystal structure may not adequately represent their solution behavior or their behavior at a solid surface. Thus, as cytochrome c undergoes dynamic motions in solution, histidine 26 may become more accessible to the protein surface. In addition, the partial unfolding of cytochrome c upon interaction with the surface cannot be excluded.[4] Recently, the destabilization of cytochrome c was observed by Muga *et al.* upon binding to an acidic phospholipid membrane.[18] Adsorption may increase the accessibility of partially buried residues and facilitate their interaction with the surface.

The results obtained for cytochrome c binding to a solid phase Cu(II)IDA matrix may be compared to the interactions of soluble Cu(II)IDA with cytochrome c (*Candida krusei*) as analyzed by Johnson and Arnold.[19,20] Johnson and Arnold measured the copper-induced paramagnetic line broadening of protein ^{1}H NMR spectra at low concentrations of Cu(II)IDA. Cu(II)IDA binding sites can be identified by the selective relaxation of proton resonances near those sites. In cytochrome c, the imidazole ring protons of histidines 26, 33, and 39 are all specifically relaxed in the presence of Cu(II)IDA, with the NMR signals for histidines 33 and 39 being broadened to a much greater extent than those of histidine 26. These results indicate that histidine 26 is accessible to Cu(II)IDA in solution, albeit with a much lower affinity than the more accessible histidines at positions 33 and 39. These results support our findings that histidine 26 can interact with Cu(II)IDA, although the interaction is weakened by its lack of accessibility.

Conclusions

Equilibrium binding isotherms for a set of cytochrome c variants have been used to investigate protein interactions with a Cu(II)IDA chromatographic matrix. The presence of a surface histidine is a prerequisite for interaction with the Cu(II)IDA matrix. Variants that contained a single surface histidine bind the matrix with affinity that is greater than imidazole and dependent on the histidine's accessibility. Two-histidine variants had bimodal binding isotherms and 10-fold greater binding affinity than single-histidine variants. The bimodal isotherm and the strong binding of the two-histidine variants is attributed to the formation of two-site binding interactions with the IMAC matrix. Finally, the overall binding behavior of the cytochrome c variants could be explained using a model in which bound protein covered multiple copper sites on the surface and multiple-histidine variants were involved in two-site interactions (macro-chelation).

Table 3.1 Langmuir binding parameters for the complete set of iso-1-cytochrome c variants and imidazole. Multi-histidine variant isotherms have been fit to bi-Langmuir model as described in the Mathematical Modeling section.

	K_{p1} (mM^{-1})	Q_{max1} (mM)	K_{p2} (mM^{-1})	Q_{max2} (mM)
Imidazole	5.4 ± 0.3	18.5 ± 0.5	0.03 ± .003	18.5 ± 0.05
Tuna cytochrome c	2.4 ± 0.1	1.25 ± 0.3		
H_{26}	3.5 ± 0.6	1.12 ± 0.13		
H_4	56 ± 3.	1.15 ± 0.02		
H_8	49 ± 5.	1.03 ± 0.04		
$H_{26}H_4$	900 ± 100	1.10 ± 0.10	19 ± 10	0.29 ± 0.10
$H_{26}H_8$	1000 ± 100	0.78 ± 0.06	18 ± 10	0.56 ± 0.09
$H_{26}H_{33}$	740 ± 70	0.52 ± 0.03	24 ± 5	0.90 ± 0.02
$H_{26}H_{33}H_4$	5600 ± 800	1.33 ± 0.13	48 ± 20	1.08 ± 0.13
$H_{26}H_{33}H_8$	3600 ± 700	1.32 ± 0.10	15 ± 8	0.95 ± 0.11
$H_{26}H_{33}H_{39}$	4300 ± 400	1.27 ± 0.04	29 ± 5	0.96 ± 0.04

Table 3.2 Estimated binding parameters for selected iso-1-cytochrome c variants. Single-histidine isotherms have been fit using the Langmuir-type model described by equation 4. Two-histidine isotherms have been fit using equation 6. K_1 is the binding constant for a single histidine interaction with the matrix, n is the number of copper sites that are occupied by a single bound protein, K_2 is the apparent binding constant of the second histidine to nearby copper sites, and θ is the fraction of surface sites that can accommodate two-site binding. (Cu_{tot} is fixed at 18.5 μmol Cu/mL of packed gel.)

Variant	**K_1 (mM^{-1})**	**n**	**K_2**	θ
H_{26}	0.2	16		
H_4	3.5	16		
H_8	2.7	18		
$H_{26}H_4$	3.5[a]	14	20	0.8
$H_{26}H_8$	2.7[a]	15	27	0.6

[a] Value is fixed during parameter fit

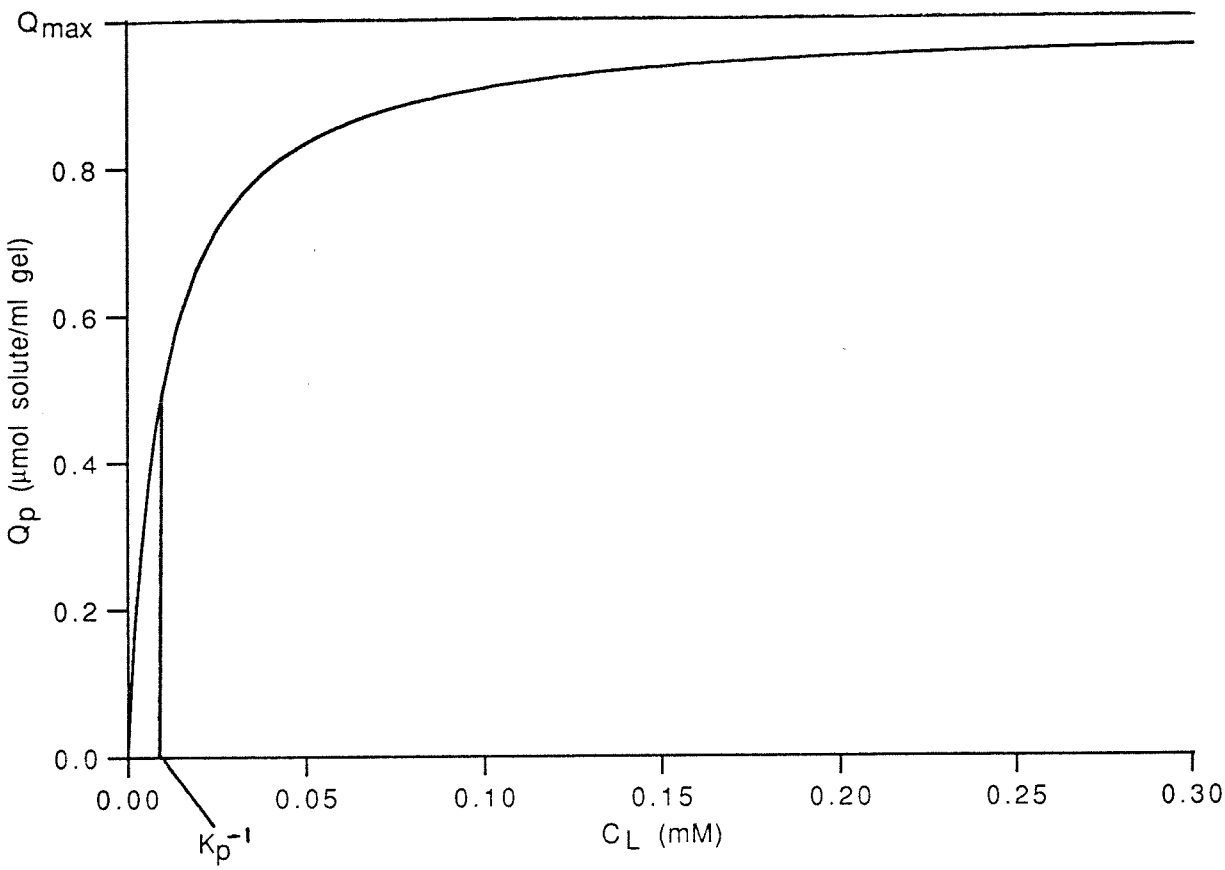


Figure 3.1 Langmuir representation of equilibrium binding isotherm. The amount of solute bound per mL of gel, Q_p, is plotted versus the concentration of solute at equilibrium in the liquid phase, C_L. Q_{max} is the maximum solute that can be adsorbed at infinite solute concentration, and K_p is the apparent binding constant between the solute and sites on the gel surface.

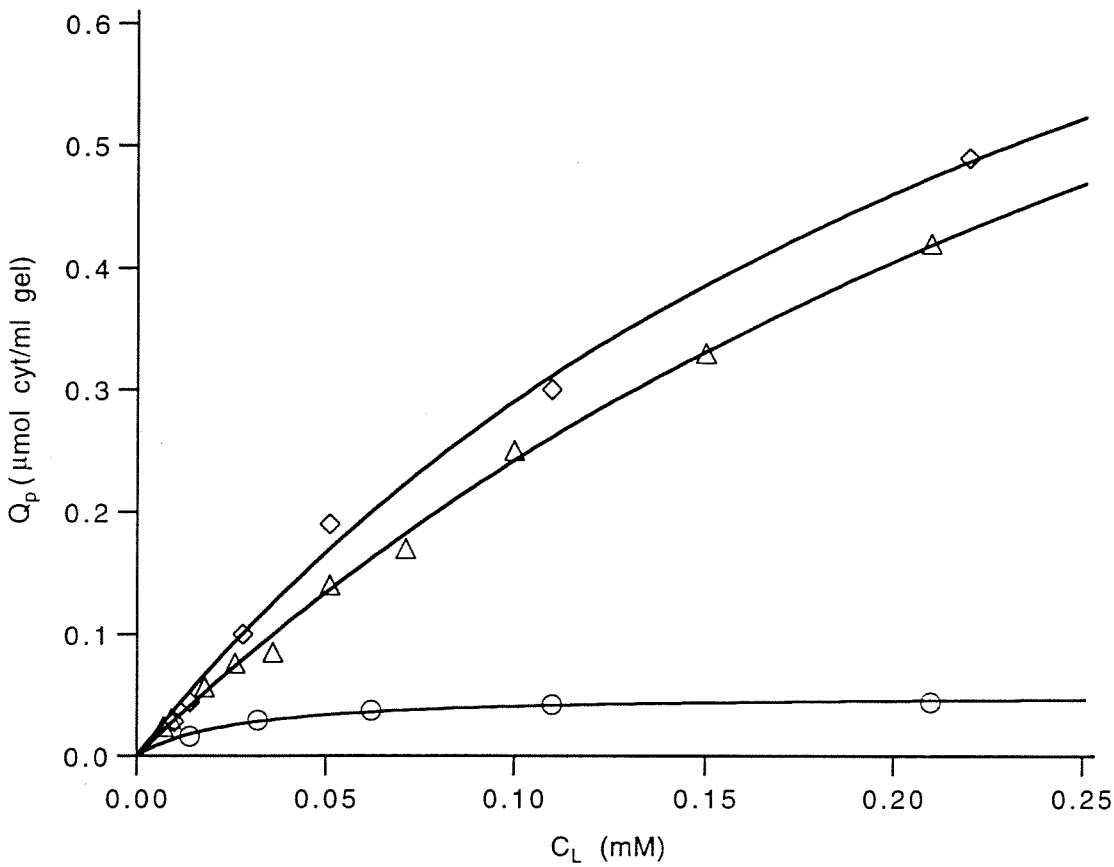


Figure 3.2 Equilibrium binding isotherms of weakly binding cytochromes c. Solid lines represent best fit Langmuir isotherms using parameters given in Table 3.1. H(-) shows negligible adsorption onto matrix. H_{26} and tuna cytochrome c have similar adsorption isotherms. (◇) H_{26}, (Δ) Tuna, (O) H(-).

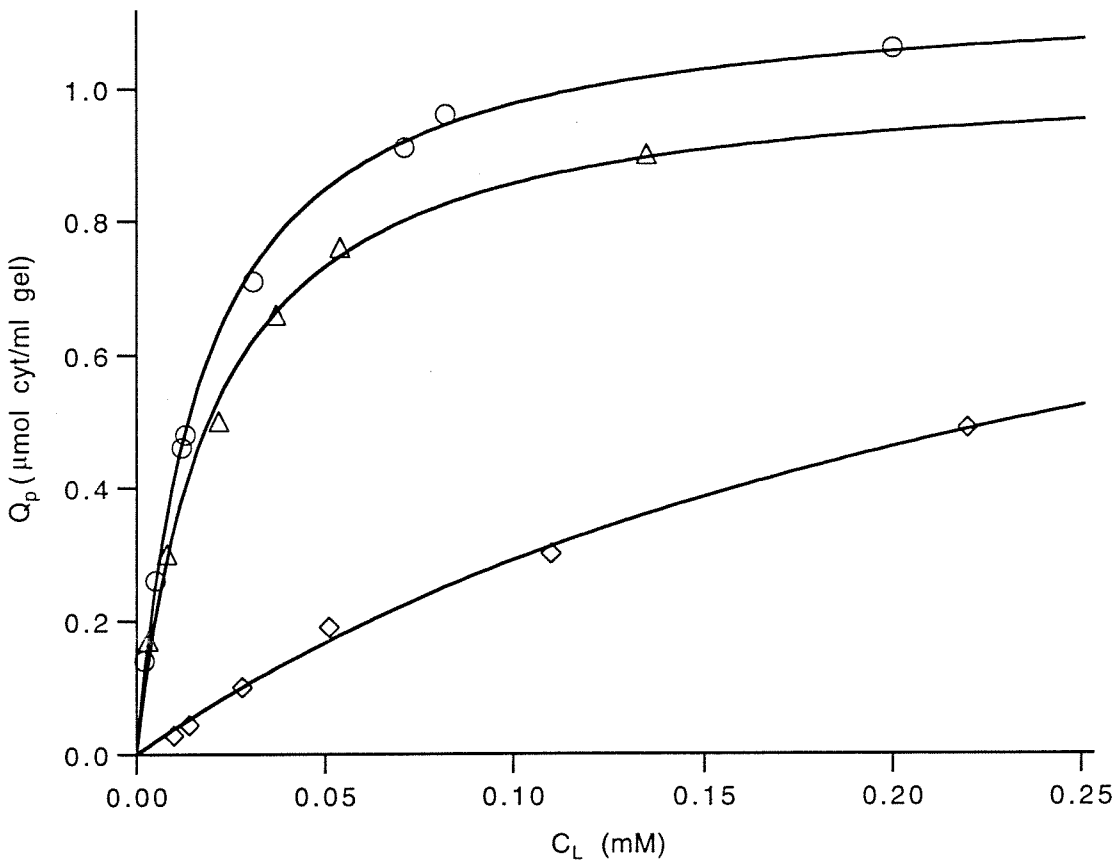


Figure 3.3 Equilibrium binding isotherms of single-histidine iso-1-cytochrome c variants. Solid lines represent best fit Langmuir isotherms using parameters given in Table 3.1. H_4 and H_8 show increased binding over H_{26} due to greater surface accessibility of their single surface histidines. (O) H_4, (Δ) H_8, (◇) H_{26}.

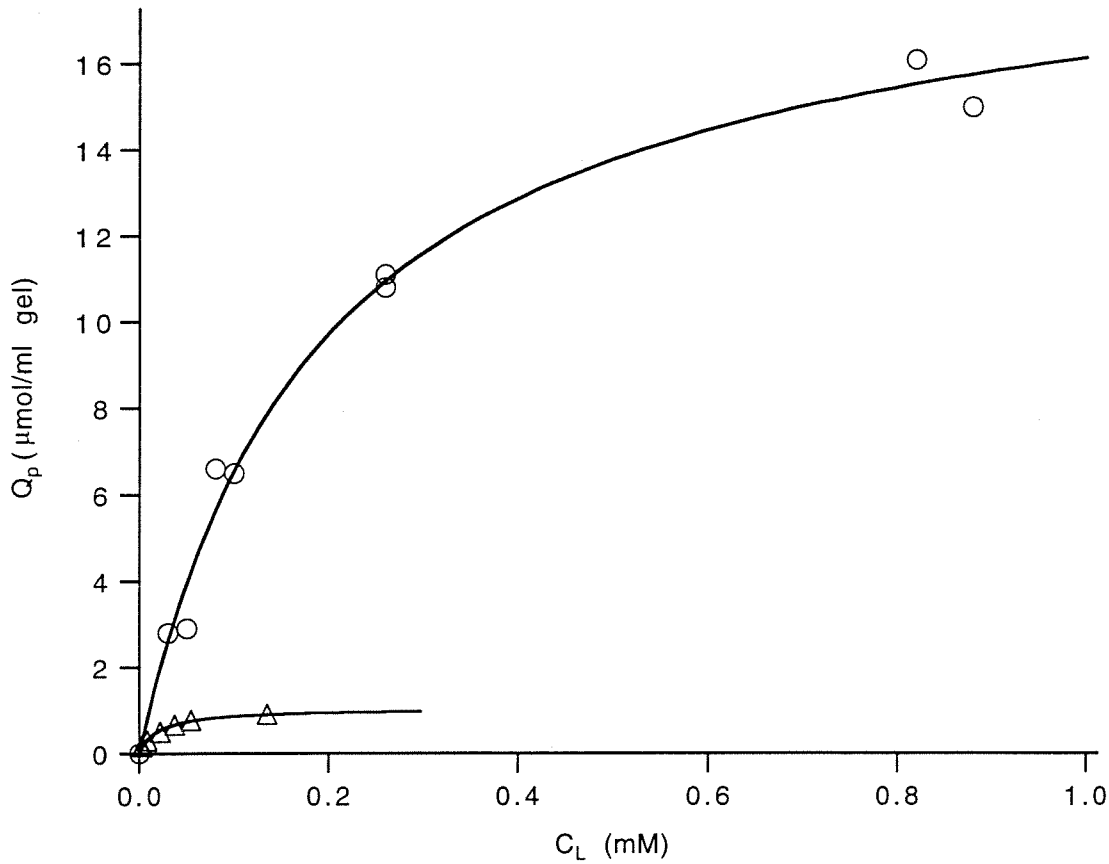


Figure 3.4 Comparison of imidazole and iso-1-cytochrome c (H_8 variant) equilibrium binding isotherms. Solid lines represent best fit isotherms for imidazole (bi-Langmuir) and H_8 (Langmuir) using parameters given in Table 3.1. H_8 has reduced capacity and increased binding affinity relative to imidazole. (O) imidazole, (Δ) H_8.

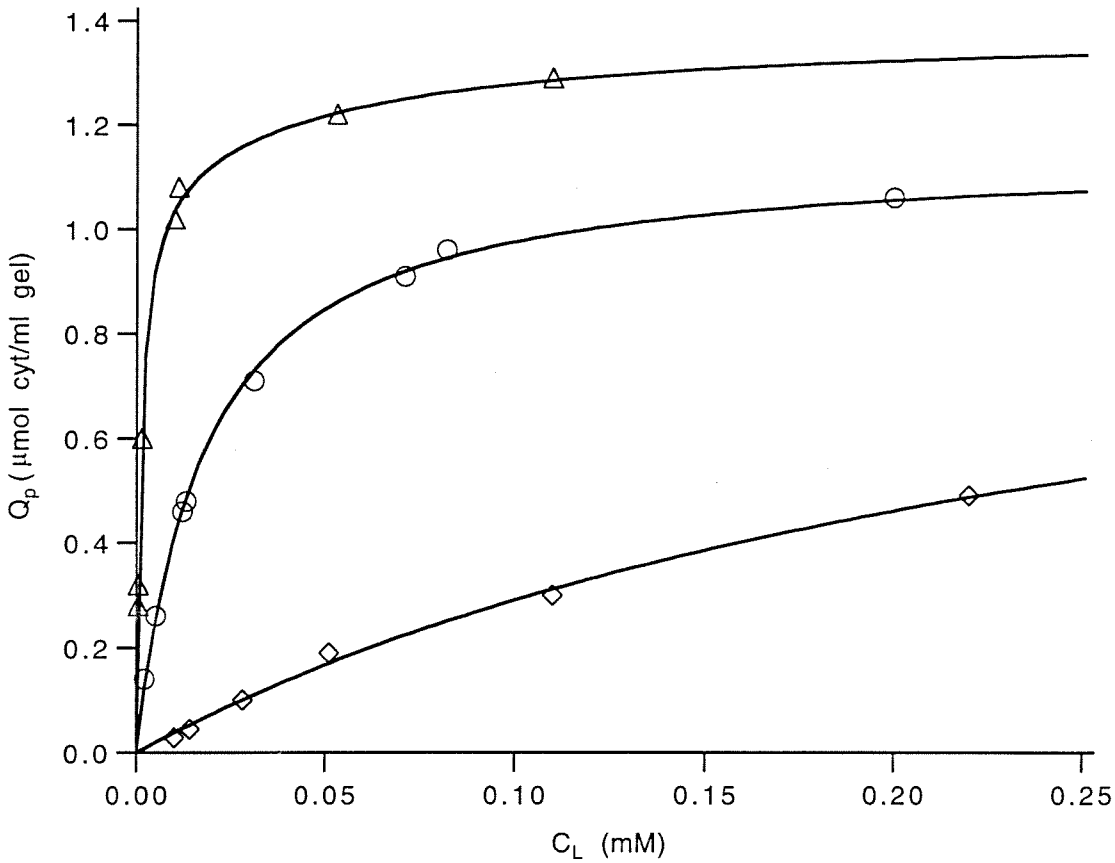


Figure 3.5 Evidence of two-site binding in $H_{26}H_4$ iso-1-cytochrome c. Comparison of two-histidine variant ($H_{26}H_4$) equilibrium isotherm to the individual one-histidine variant (H_{26} and H_4) isotherms . Solid line through $H_{26}H_4$ variant data represents the best fit isotherms to a bi-Langmuir model using the parameters given in Table 3.1. Initial slope of two-histidine variant isotherm is more than a factor of 10 greater than either single histidine isotherm. (Δ) $H_{26}H_4$, (O) H_4, (◇) H_{26}.

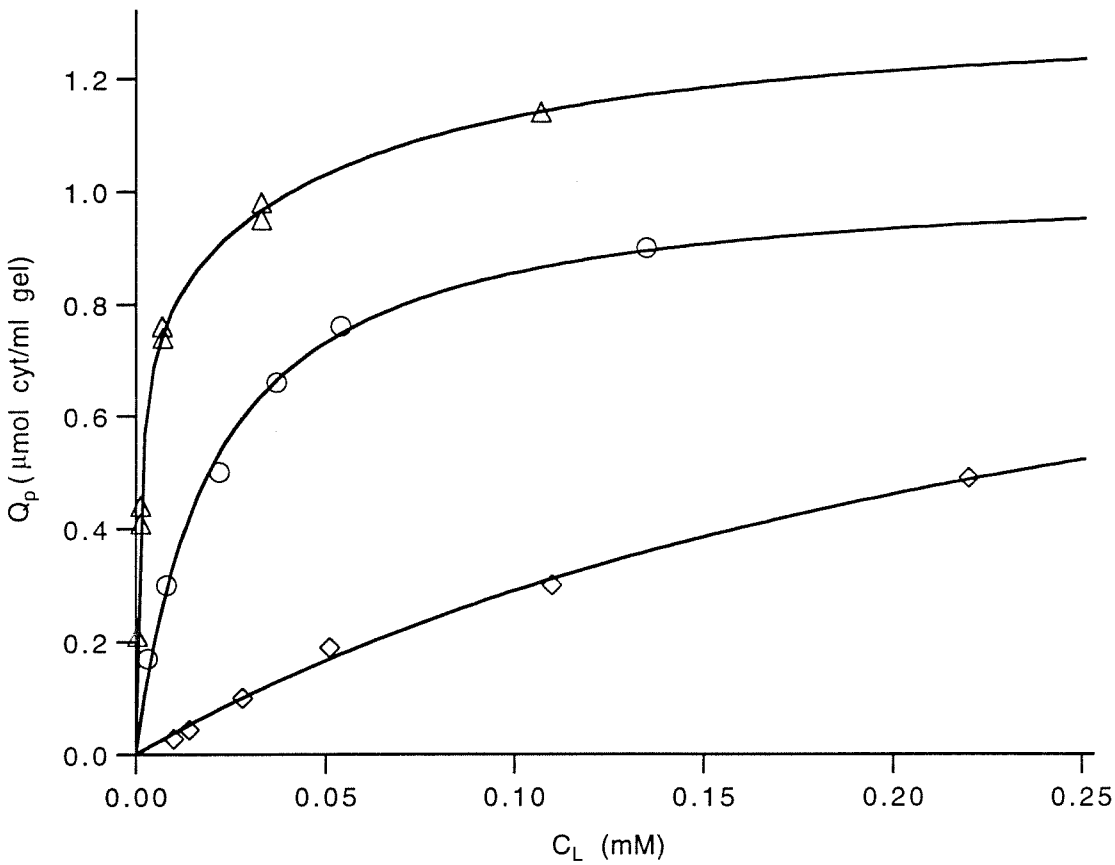


Figure 3.6 Evidence of two-site binding in $H_{26}H_8$ iso-1-cytochrome c. Comparison of two-histidine variant ($H_{26}H_8$) equilibrium isotherm to the individual one-histidine variant (H_{26} and H_8) isotherms . Solid line through $H_{26}H_8$ variant data represents the best fit isotherms to a bi-Langmuir model using the parameters given in Table 3.1. Initial slope of two-histidine variant isotherm is more than a factor of 10 greater than either single histidine isotherm. (Δ) $H_{26}H_8$, (O) H_8, (◇) H_{26}.

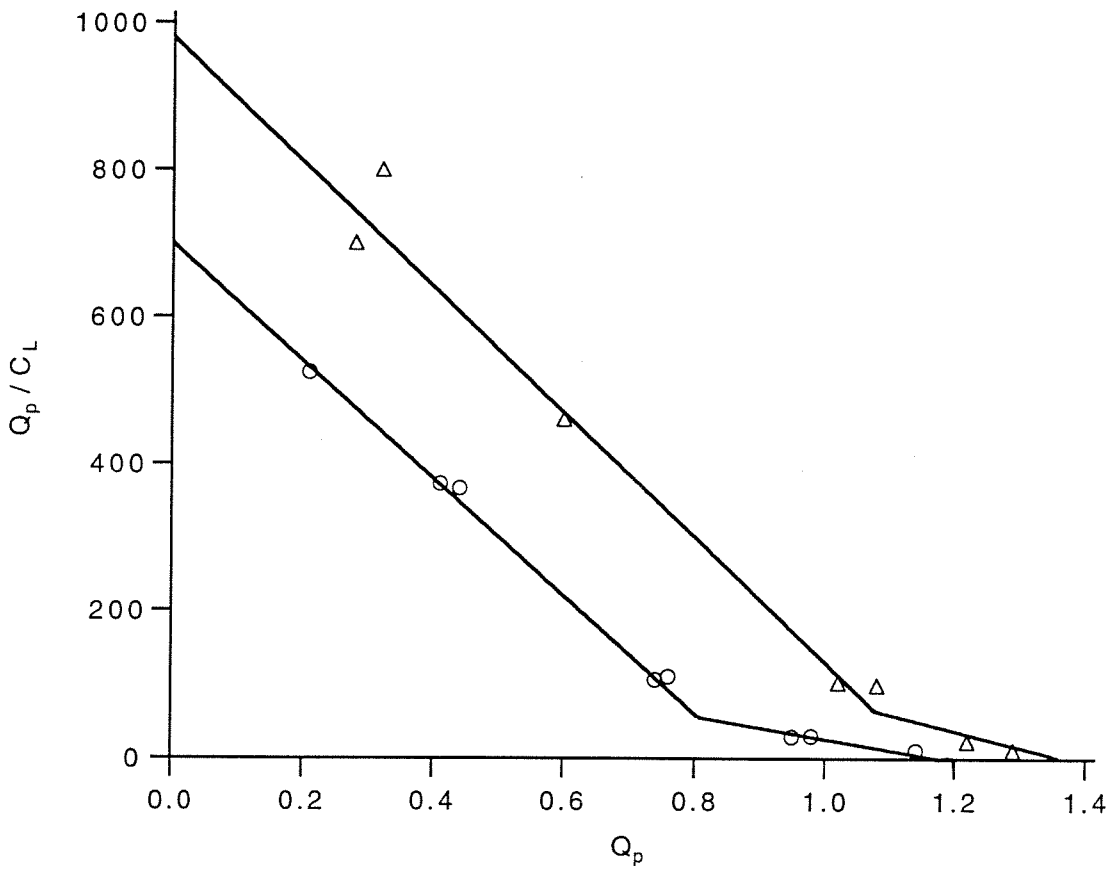


Figure 3.7 Scatchard plot of $H_{26}H_4$ and $H_{26}H_8$ isotherms fit to bimodal binding behavior. Q_p/C_L is plotted versus Q_p to give Scatchard representation. Slope of data represents $-K_p$ and the y-intercept represents K_pQ_{max}. A nonlinear Scatchard plot indicates heterogeneous binding to the surface. (Δ) $H_{26}H_4$, (O) $H_{26}H_8$.

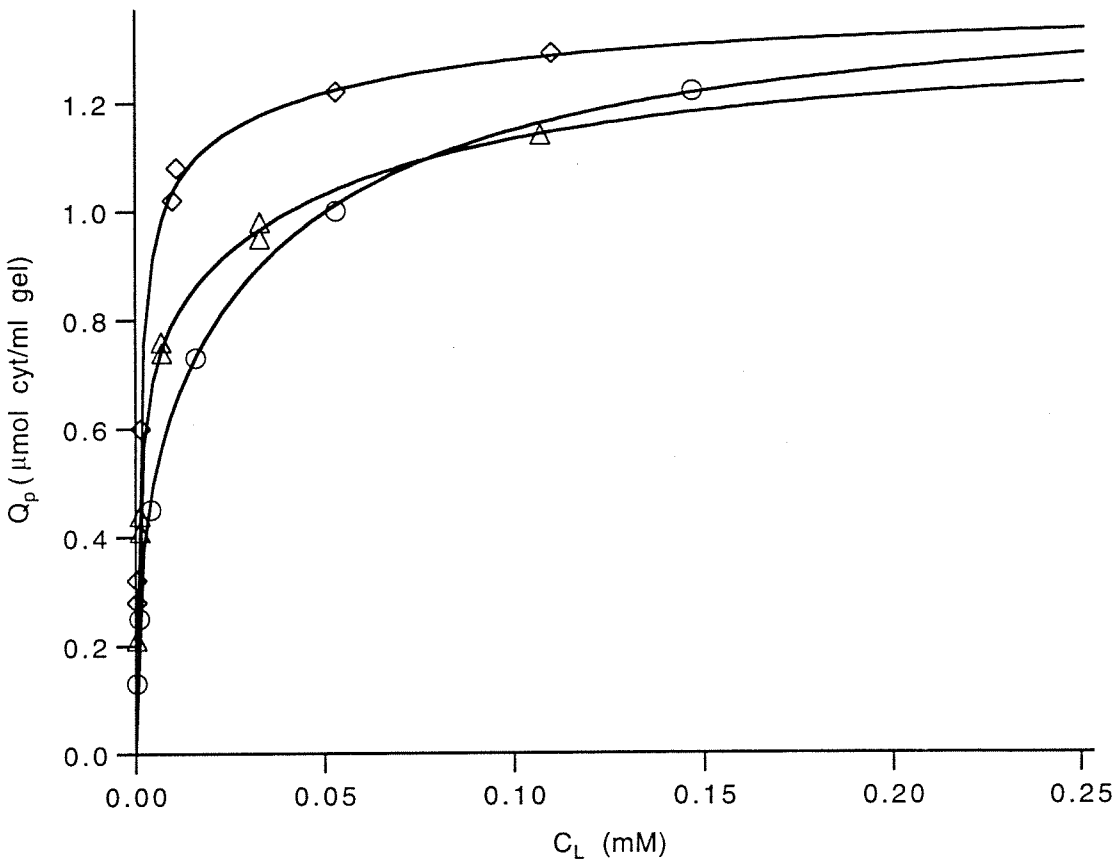


Figure 3.8 Equilibrium binding isotherms of two-histidine iso-1-cytochrome c variants. Solid lines represent best fit to bi-Langmuir model using parameters given in Table 3.1. All variants have similar maximum capacities (1.4 mmol/mL gel). The measured binding constants are approximately 10^6 M^{-1} for the strong sites and 2 x 10^4 M^{-1} for the weak sites. (◇) $H_{26}H_4$, (Δ) $H_{26}H_8$, (O) $H_{26}H_{33}$.

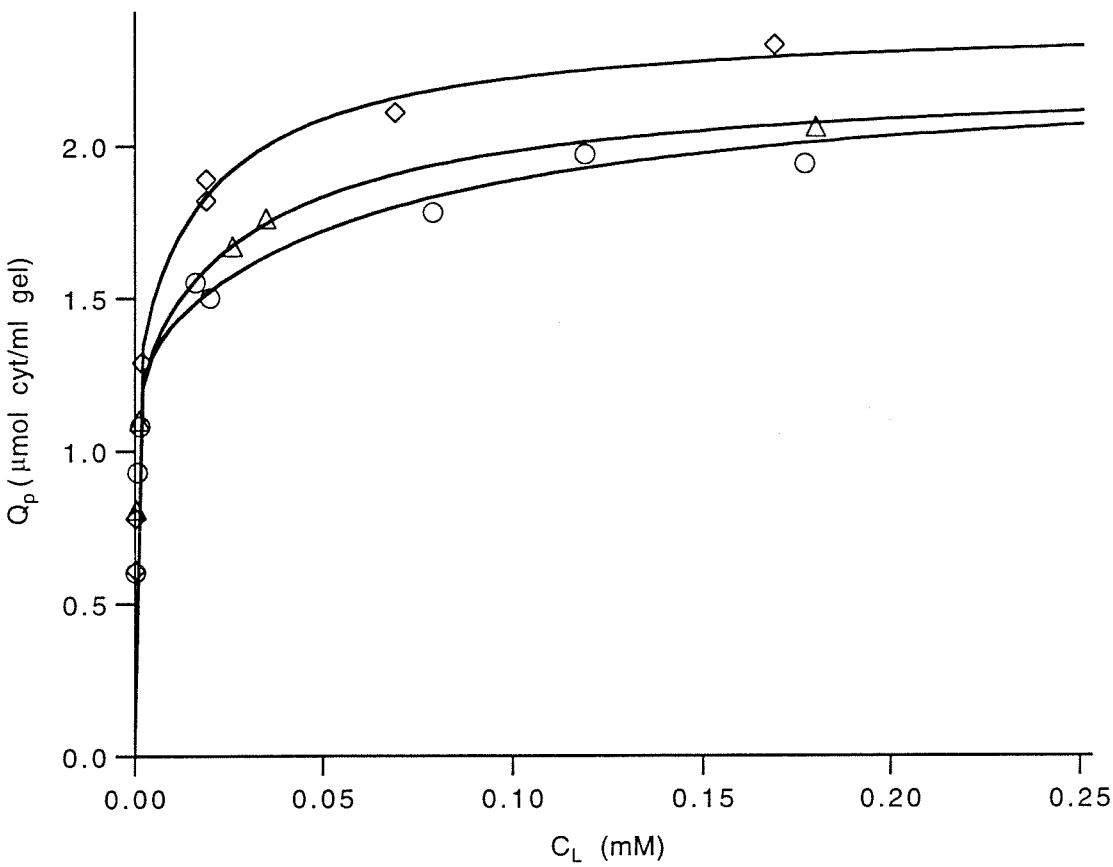


Figure 3.9 Equilibrium binding isotherms of three-histidine iso-1-cytochrome c variants. Solid lines represent best fit to bi-Langmuir model using parameters given in Table 3.1. (◊) $H_{26}H_{33}H_4$, (Δ) $H_{26}H_{33}H_{39}$, (O) $H_{26}H_{33}H_8$.

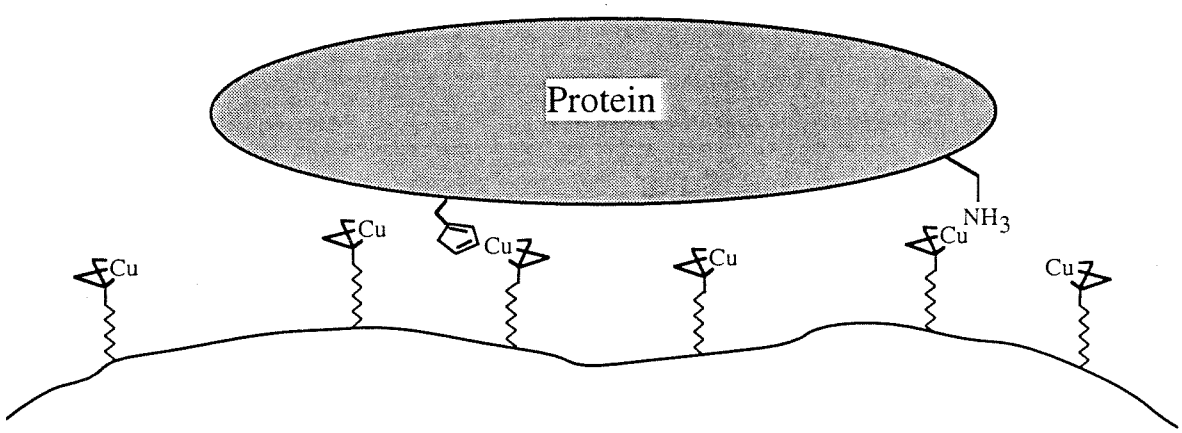


Figure 3.10 Illustration of proposed single-histidine binding interaction in IMAC. Protein may block access to a number of copper sites upon binding to the surface of the matrix. Additional weak coordination events may result from secondary ligands on the protein surface (*e.g.*, *N*-terminus).

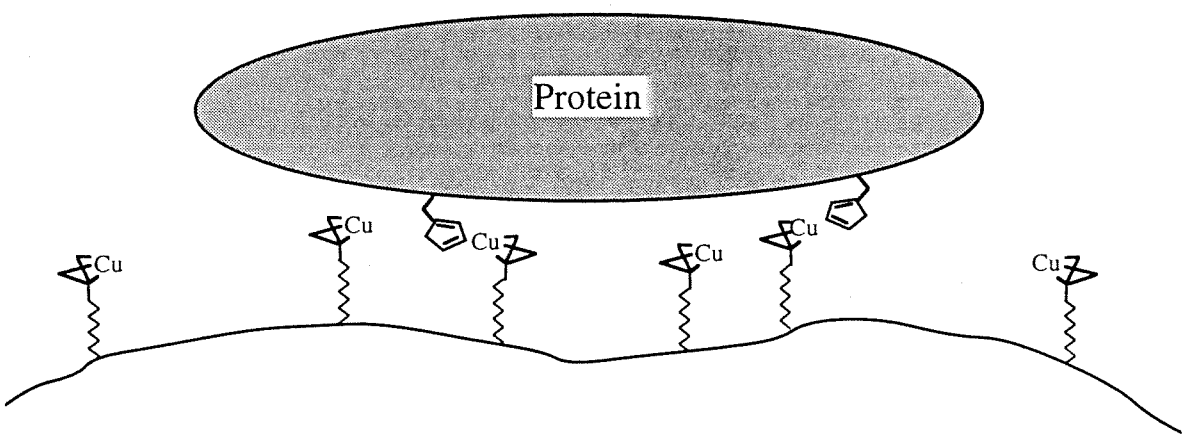


Figure 3.11 Illustration of "macro-chelation" by protein on IMAC surface. Two histidines are able to coordinate different copper sites at the surface resulting in enhanced affinity for the matrix.

References

1. Hemdan, E. S., Zhao, Y., Sulkowski, E. and Porath, J., *Proceedings of the National Academy of Sciences USA* **86**, 1811-1815 (1989).
2. Zhao, Y.,Sulkowski, E. and Porath, J., *European Journal of Biochemistry* **202**, 1115-1119 (1991).
3. Weber, W. J. Jr., in *Adsorption Technology: A Step-by-Step Approach to Process Evaluation and Application,* F. L. Slejko, Ed. (Tall Oaks Publishing, Voorhees, N. J., 1985) pp.1-36.
4. Martell, A. E. and Smith, R. M., in *Critical Stability Constants*, (Plenum Press, New York, 1974) vol.1-6.
5. Hutchens, T. W., Yip, T. and Porath, J., *Analytical Biochemistry* **170**, 168-182 (1988).
6. Margoliash, E. and Frohwirt, N., *Biochemical Journal* **71**, 570-572 (1959).
7. Andrade, J. D. and Hlady, V., in *Advances in Polymer Science*, (Springer-Verlag, Berlin, 1986) vol. 79, pp. 1-64.
8. Arnold, F. H., Schofield, S. A. and Blanch, H. W., *Journal of Chromatography* **355**, 1-12 (1986).
9. Laub, R. J., in *ACS Symposium Series: Chromatography and Separation Chemistry*, S. Ahuja, Ed. (American Chemical Society, Washington , D. C., 1986) vol. 297, pp. 1-33.
10. Huang, J., Schudel, J. and Guiochon, G., *Journal of Chromatographic Science* **29**, 122-126 (1991).
11. Takano, T. and Dickerson, R. E., *Journal of Molecular Biology* **153**, 95-115 (1981).
12. Louie, G. V. and Brayer, G. D., *Journal of Molecular Biology* **214**, 527-555 (1990).
13. Sinha, P. C., Saxena, V. K., Nigam, N. B. and Srivastava, M. N., *Indian Journal of Chemistry* **28A**, 335-336 (1989).
14. Haymore, B. L., unpublished results.
15. Mrabet, N. T. *Biochemistry* **31**, 2690-2702 (1992).
16. Hutchens, T. W. and Yip, T., *Analytical Biochemistry* **191**, 160-168 (1990).
17. Sulkowski, E., *Trends in Biotechnology* **3**, 1-7 (1985).
18. Muga, A., Mantsch, H. H. and Surewicz, W. K., *Biochemistry* **30**, 7219-7224 (1991).

19. Johnson, R. D. and Arnold, F. H., *Journal of Cellular Biochemistry* **15g**, 76 (1991).

20. Johnson, R. D., Masters Thesis, California Institute of Technology (1991).

Chapter 4

IMAC Studies II: Chromatography and Competitive Elution

Introduction

Immobilized metal-affinity chromatography (IMAC) has proven to be a useful technique for the separation of proteins. There have been over 70 different reports of IMAC being used as a primary step in protein purification,[1] yet the precise molecular mechanism by which these separations are taking place is not well understood. Current analysis of IMAC separations is based on the postulate that proteins interact with IMAC matrices predominantly via surface histidine residues. This has led to the proposal that IMAC may also be used as an analytical tool to probe the surface of proteins to determine their surface histidine content.[2-4] Previous studies have suggested that selectivity in IMAC separations can be derived from both the multiplicity and local environment of the surface histidine residues.[2-4] Selectivity of a single-histidine proteins may result from differences in the pK_a of the histidine or from differences in its accessibility. When more than a single histidine is present, selectivity may be derived from either the total number or even the placement of histidines on the surface of the protein. The current understanding of selectivity in IMAC is qualitative in nature and the effect of competitors on the interaction (protons and imidazole) has not been sufficiently addressed. Our goal is to set forth a quantitative understanding of the interaction of proteins and competitors in immobilized metal-affinity chromatography that will allow for more efficient designs of IMAC separations in the future.

The quantitative interpretation of metal-affinity chromatography in terms of the properties of a protein surface and mobile phase competitors is a complex task. The most common elution techniques in IMAC involve either pH or imidazole.[2,5] Therefore, immobilized metal-affinity chromatography in its simplest form (single surface histidine protein) involves at least four sets of equilibria, as seen below. These equilibria provide a theoretical framework in which differences among proteins and changes in the running

conditions (pH and imidazole concentration) can be dealt with separately and quantitatively.

$$P + M\text{-}IDA \Leftrightarrow P\text{-}M\text{-}IDA$$

$$P + H \Leftrightarrow P\text{-}H$$

$$I + M\text{-}IDA \Leftrightarrow I\text{-}M\text{-}IDA$$

$$I + H \Leftrightarrow I\text{-}H$$

The first equilibrium represents the affinity of the protein (P) for the immobilized metal ion (M-IDA) in the absence of competitors. This interaction can involve multiple histidines and metals and therefore can be more complicated than indicated here. The second equilibrium represents the competition by protons (H) for individual histidines on the protein surface. The third equilibrium involves the binding of the competitor molecule, which for the purposes of this study will be imidazole (I), to the metal-ion complex. The final equilibrium represents the protonation of imidazole, thereby making it a less effective competitor. It is the interplay of these equilibria that provides both the power and the complexity of IMAC.

In this chapter, the retention of a set of N-acetyl amino acids is studied, in order to confirm the underlying postulate that surface histidines residues are the predominant ligand responsible for protein retention in IMAC. The retention of the set of cytochrome c variants will then be used to study how the multiplicity and microenvironment of surface histidines effects IMAC retention. The effect of both proton and imidazole competition on retention will be studied using both gradient and isocratic elution and interpretation of the results will be done in light of the equilibrium binding information that was presented in Chapter 3. Experiments are performed under analytical conditions (local equilibrium and

low protein loading) so that retention is a function of only the initial slope of the equilibrium binding isotherm and the competitors present in the mobile phase.[6-9]

Materials and Methods

IMAC Experiments

A TSK Chelate-G6000XL (10 μm macroporous beads) HPLC column (7.5 cm x 7.5 mm) from Tosohaas (#08645) was used in all experiments. Experiments were performed at a flow rate of 1 mL/min using a Hitachi model L6200A high-pressure pump/low pressure gradient former. The eluant was monitored by UV-visible absorbance using a Kratos Spectraflow 783 variable wavelength detector. The column was regenerated by first washing with 15 mL of 50 mM ethylenediaminetetraacetic acid (EDTA), pH 8.0, and then reloading with 15 mL of 50 mM $CuSO_4$ in water. The column volume (v) is 3.31 mL and the void volume (V_O) is 2.3 mL. In all cases, experiments were performed under analytical conditions, with less than one-hundredth of the column capacity being utilized for each injection. Alterations in the flow rate had no effect on measured retention volumes (the kinetics of the imidazole-Cu(II)IDA interaction are very fast), and therefore, local equilibrium conditions were assumed throughout. In addition, peaks were generally symmetrical, allowing us to approximate the center of mass by the maximum absorbance.

Isocratic Elution of N-Acetyl Amino Acids

Isocratic experiments were performed in 0.05M sodium phosphate and 0.5 M NaCl at pH 7.6. 20 μL of a 10 mM solution of each *N*-acetyl amino acid was injected and the eluant absorbance monitored at 210 nm. The column was adequately equilibrated in running buffer before each injection and the reported retention time was that at which the eluant absorbance was maximum.

Isocratic Elution of Cytochrome c Variants

Following regeneration, the column was equilibrated with 30 mL of one of the following buffers: A) 50 mM sodium phosphate, 50 mM sodium acetate, 0.5 M NaCl at pH's between 4.0 and 7.6; or B) 50 mM sodium phosphate, 0.5 M NaCl, pH 7.0 at imidazole concentrations from 0 to 1 mM. 100 μL of a 50 μM solution of protein was injected and the retention is reported as the time at which the eluant absorbance at 410 nm was maximum.

pH Gradient Elution

The TSK chelate column was regenerated and then equilibrated with 10 mL of low pH buffer (50 mM sodium phosphate, 50 mM sodium acetate, 0.5M NaCl at pH 4.0) followed by 20 mL of the high pH loading buffer (50 mM sodium phosphate, 50 mM sodium acetate, 0.5M NaCl at pH 7.0) in between each injection. A 100 μL aliquot of a 50 μM protein solution was injected and the column washed for 6 minutes in the loading buffer. The protein was then eluted in a linear gradient from pH 7 buffer to pH 4.0 buffer in 60 minutes. The gradient was approximately linear in pH and retention will be reported as the pH at which the eluant absorbance at 410 nm was maximum.

Imidazole Gradient Elution

The TSK chelate column was regenerated and then equilibrated in between each experiment with 10 mL of 50 mM sodium phosphate, 0.5M NaCl, 10 mM imidazole at pH 7.0 followed by 20 mL of the same buffer containing 1 mM imidazole. 20 μL of a 100 μM protein solution was injected and the column washed for 6 minutes. The protein was eluted in a linear gradient from 1 mM imidazole to 10 mM imidazole in 54 minutes. The retention will be reported as the concentration of imidazole at which the eluant absorbance at 410 nm was maximum.

Measurement of Copper Loading

The copper loading of the TSK chelate column was measured using the visible absorbance of EDTA-Cu at pH 7.0 (E_{800} = 73.6 cm^{-1} M^{-1}). Copper was eluted using 50 mM EDTA at pH 7.0 immediately following the regeneration and equilibration protocols. The complete eluant was collected and the absorbance determined at 800 nm. The copper loading of the TSK column was 25.0 μmol copper per mL of column volume.

Theory

Prediction of Retention in Isocratic Experiments

Chromatographic retention under isocratic conditions can be predicted directly from the Langmuir isotherm parameters K_p (M^{-1}) and Q_{max} (mol solute per L of packed gel) as given by Arnold *et al.*,[6]

$$\frac{V_e - V_o}{v} = Q_{max} K_p \qquad (1)$$

where Kp is the Langmuir binding constant, Qmax is the maximum solute capacity on the matrix, V_e is the elution volume, V_o is the column void volume, and v is the column volume. This equation is valid for chromatographic experiments performed under local equilibrium conditions when the mobile phase protein concentration (C_L) is low ($K_pC_L<<1$). The quantity $Q_{max}K_p$ represents the initial slope of the Langmuir isotherm and is the relevant parameter used to predict elution volumes. $Q_{max}K_p$ has also been called the linear equilibrium distribution coefficient (K')[6] and represents the partitioning of solute to the solid phase at low surface coverage. For the bi-Langmuir isotherm model, the strongest binding site determines the initial slope of the isotherm, which is given by $Q_{max1}K_{p1}$.

pH Competition

The effect of pH on binding can be modeled for a single-histidine protein (P) as a simple proton competition as given below,

$$P + Cu \Leftrightarrow P\text{-}Cu \qquad K_p = \frac{[P\text{-}Cu]}{[P][Cu]}$$

$$P + H \Leftrightarrow P\text{-}H \qquad K_a = \frac{[P\text{-}H]}{[P][H]}$$

where K_p is the association constant for the immobilized copper (Cu) binding to the surface histidine, and K_a is the association constant for a proton (H) binding to the surface histidine. Using this equilibrium analysis, the following equation describes the effect of proton competition on the binding isotherm,

$$Q_p = \frac{Q_{max} K_p C_L}{1 + K_a[H] + K_p C_L}$$

where proton competition reduces the initial slope of the isotherm by $1/(1+K_a[H])$. Assuming local equilibrium and low mobile phase protein concentration ($K_pC_L << 1+ K_a[H]$), the following equation should describe isocratic retention as a function of pH.

$$\frac{V_e - V_o}{v} = \frac{Q_{max} K_p}{1 + K_a[H]} \qquad (2)$$

Imidazole Competition

A simple equilibrium model can also be formulated for competition by imidazole. Imidazole (I) competes with the protein for copper sites on the surface of the chromatographic matrix as described below,

$$P + Cu \Leftrightarrow P\text{-}Cu \qquad K_p = \frac{[P\text{-}Cu]}{[P][Cu]}$$

$$I + Cu \Leftrightarrow I\text{-}Cu \qquad K_I = \frac{[I\text{-}Cu]}{[I][Cu]}$$

where K_I is the association constant for imidazole binding to immobilized copper. Using this equilibrium analysis, the following equation describes the effect of imidazole on the binding isotherm,

$$Q_p = \frac{Q_{max} K_p C_L}{1 + K_I[I] + K_p C_L}$$

Imidazole competition reduces the initial slope of the isotherm by $1/(1 + K_I [I])$. Assuming local equilibrium and low mobile phase protein concentration ($K_p C_L << (1+K_I[I])$), the effect of imidazole concentration on isocratic elution volume can be described by,

$$\frac{V_e - V_o}{v} = \frac{Q_{max} K_p}{1 + K_I[I]} \qquad (3)$$

Results

Isocratic Chromatography of N-Acetyl Amino Acids

The relative contribution of amino acid side chains to retention in immobilized metal-affinity chromatography was studied using a set of *N*-acetyl amino acids. As shown

in Figure 4.1, the *N*-acetyl blocking group is used to eliminate chelation between the amino and carboxyl groups. The carboxyl group alone has an affinity for Cu(II) (K) of approximately 10^1-10^2 M^{-1} and provides only minimal binding in the absence of the amino group. For example, according to Martell, glycine has a log K of 8.15 for Cu(II) in solution, while the log K of *N*-acetyl glycine is only 1.30.[10] According to equation 1, any significant retention (greater than one column volume) of the *N*-acetyl amino acids will require a log K greater than 1.5. Therefore, *N*-acetyl amino acids provide a good model system because, as is the case in proteins, their primary affinity for immobilized metal ions will be due to their side chains.

The measured isocratic retention values of the set of *N*-acetyl amino acids on a Cu(II)IDA TSK column are given in Table 4.1. The retention of *N*-acetyl histidine is much greater than the other *N*-acetyl amino acids studied. *N*-acetyl tryptophan, *N*-acetyl tyrosine and *N*-acetyl phenylalanine showed slight retention on the column, while *N*-acetyl glycine eluted in the void volume. *N*-acetyl lysine and *N*-acetyl leucine also showed slight, but measurable, retention. The calculated linear equilibrium distribution coefficients (K') according to equation 1 for the set of *N*-acetyl amino acids are also given in Table 4.1. Since all the *N*-acetyl amino acids are small, they should have access to all of the copper sites, and therefore, have the same maximum capacity (Q_{max}). This being the case, the linear equilibrium distribution coefficient ($K' = Q_{max}K_p$) is proportional to the binding affinity for the Cu(II)IDA matrix (K_p). The linear equilibrium distribution coefficients indicate that *N*-acetyl histidine has a binding constant at least 100 times greater than *N*-acetyl tryptophan. Although some slight contribution is seen from the indole ring of tryptophan, the predominant role that histidine plays in IMAC is evident in these experiments.

Isocratic Chromatography of Single-Histidine Variants

Table 4.2 shows the elution behavior for a set of weakly binding cytochrome c variants using the Cu(II)IDA TSK HPLC matrix in the absence of any significant competitors. The H(-) variant is eluted from the column immediately following the void volume, and hence has almost no interaction with the matrix. The H_{26} variant was retained weakly and eluted in 6.1 column volumes, while the H_8 variant was strongly retained and did not elute until 115 column volumes. Although the dilution resulting from band spreading for the H_8 variant was significant, the three variants could be separated with baseline resolution using isocratic elution.

The retention of the H_{26} and H_8 variants can be predicted from the equilibrium isotherms measured in Chapter 3 using equation 1. The predicted retention values for the H_{26} and H_8 variants, based on the increased copper loading of the HPLC matrix (25 μmol Cu/mL gel) compared to the guardgel matrix (18.5 μmol Cu/mL gel), matches the predicted values fairly well. The predicted retention of H_8 variant (70 column volumes) was less than the actual value (115 column volumes), suggesting possible differences between the TSK matrix used for HPLC (10 μm beads) and the TSK guardgel matrix (17 μm beads) used for the equilibrium isotherm measurements.

pH Competition

A common mode of elution in IMAC is by proton competition. A decrease in the pH of the mobile phase results in protonation of surface histidines which are then unable to interact with Cu(II)IDA. The isocratic retention behavior of the single-histidine H_8 variant as a function of pH is shown in Figure 4.2. Also shown is the expected retention behavior based on the simple pH competition model developed in equation 2, using the independently measured value of approximately 5.7 as the pK_a of histidine 8 (given in Chapter 2 and corrected for temperature). Because $Q_{max}K_p$ is unknown in this buffer

system, it has been set at the experimentally determined value at pH 7.0 ($(Q_{max}K_p)_{max}$ = 45). The pH dependence of retention does not follow the behavior predicted by the simple model for single-histidine binding. As the pH is increased, the measured retention increases rapidly but does not reach a plateau once the histidine is fully deprotonated. In addition, the sharp rise in retention values above neutral pH does not correlate with the measured pK_a of histidine 8.

Imidazole Competition

The most common mode of elution in IMAC is by imidazole competition. The isocratic retention behavior of the H_8 variant as a function of imidazole concentration is shown in Figure 4.3. The predicted behavior using the imidazole competition model given by equation 3 is also shown, where K_I for imidazole is set at the independently measured value of the association constant given in Chapter 3 (5.4×10^3 M^{-1}). The experimental data suggest that imidazole is a far more effective competitor than would be predicted by its measured affinity for the matrix. The K_I for imidazole determined experimentally using this data is 6.8×10^4 M^{-1}. The discrepancy in the predicted and actual behavior of imidazole as a competitor will be discussed in the next section.

pH Gradient Elution

The retention behavior of a set of cytochrome c variants has been analyzed using a gradient in pH from 7 to 4 (a mixed phosphate-acetate buffer was used to approximate a linear gradient). Figure 4.4 shows the pH at which five different variants were eluted from the IMAC column versus their histidine contents. As in the non-competitive experiments, the H(-) variant was eluted in the void volume at pH 7.0. The H_{26} variant was eluted at pH 6.9, shortly after the gradient was begun. The H_4 and H_8 variants were retained until pH 5.0 and 5.3, respectively, while $H_{26}H_{33}$ was retained until pH 4.4. The remaining two-

histidine variants ($H_{26}H_4$ and $H_{26}H_8$) bound so tightly that they were not eluted in the 60 minute pH gradient (data not shown).

Imidazole Gradient Elution

The retention behavior of the engineered cytochrome c variants in an imidazole gradient is given in Figure 4.5 versus their histidine contents. The imidazole gradient is linear from 1 to 10 mM and the retention is reported as the concentration of imidazole at elution. The H(-) variant and all single histidine variants are eluted in the void volume, at 1 mM imidazole (data not shown). The two-histidine variants elute between 1.1 and 3.0 mM imidazole, depending on histidine location. The three-histidine variants are eluted as a group between 4.2 and 5.1 mM imidazole and the four-histidine variant is retained until 7.0 mM. The affinity for the Cu(II)IDA matrix is strongly dependent on the number of surface histidines and elution with an imidazole gradient allows the variants to be grouped in terms of their surface histidine multiplicity.

Discussion

N-Acetyl Amino Acid Studies

The role that histidine plays as the dominant ligand in IMAC has been confirmed by studying the behavior of a set of *N*-acetyl amino acids. *N*-acetyl amino acids have two possible coordinating ligands, the carboxyl group and the amino acid side chain. Because the affinity of the carboxyl group is relatively small (log K = 1.3),[10] only the side chain ligand is capable of providing any significant chromatographic retention.

In his original work on IMAC, Porath postulated that the side chains of histidine (imidazole), cysteine (thiol) and tryptophan (indole) will act as the primary metal-affinity ligands.[11] Our results indicate that *N*-acetyl histidine has a binding constant which is 100 times greater than other representative *N*-acetyl amino acids (for the Cu(II)IDA matrix) and

suggest that histidine is the primary contributor to IMAC retention (cysteine not included). The proposal by Porath that tryptophan will be a primary contributor to IMAC retention is inconsistent with its structure, in which the lone nitrogen is unable to donate its electrons to a metal ion complex because they are already involved in the π-bonding system of the indole ring. Furthermore, the pK_a of the indole nitrogen is so high that under conditions found in IMAC separations it will be protonated and unavailable to coordinate metal ions. Nevertheless, tryptophan pervades the literature as an explanation for increased protein retention in IMAC.[2,3,11,12]

Our study indicates that the retention of *N*-acetyl tryptophan is only slightly greater than that of *N*-acetyl glycine and therefore that the contribution of the indole ring to retention is minimal. This result is supported by two separate studies which have shown that *N*-acetyl tryptophan coordinates Cu(II)IDA via the carboxyl group and not the indole ring, and hence the interaction of *N*-acetyl tryptophan with Cu(II)IDA is very similar to that of *N*-acetyl glycine. The inability of the tryptophan side chain to directly coordinate Cu(II)IDA is evidenced first by X-ray crystal structures of copper complexes with *N*-acetyl tryptophan and *N*-acetyl glycine which indicate that Cu(II) coordinates both molecules via the carboxylate oxygens.[13,14] In addition, Johnson and coworkers have taken advantage of the paramagnetism of Cu(II) to localize the site of metal-amino acid interactions in aqueous solution .[15,16] A qualitative description of the structure of the ternary complexes (e.g., *N*-acetyl glycine-Cu(II)IDA) arises from the effect of the copper complex upon the ^{13}C NMR spectra of the *N*-acetyl amino acid. The spectra of *N*-acetyl tryptophan and *N*-acetyl glycine clearly indicate that Cu(II)IDA is coordinating at the carboxylic acid moiety in solution, in accord with crystallographic predictions. Similar studies on N-acetyl histidine support the direct interaction of the imidazole ring with Cu(II)IDA. X-ray crystal structures of Cu(II)IDA complexes predict that the imidazole nitrogens are directly coordinating ligands.[13,14] Additionally, Johnson *et al.* have shown that Cu(II)IDA is coordinated

exclusively by the imidazole nitrogens of *N*-acetyl histidine in aqueous solution.[15,16] Therefore, as expected, the indole ring of *N*-acetyl tryptophan does not directly interact with Cu(II)IDA, as is observed for the imidazole ring of *N*-acetyl histidine.

However, our chromatographic results do indicate a slight increase in the retention of *N*-acetyl tryptophan relative to *N*-acetyl glycine. This result, as well as many other examples of tryptophan enhancing an existing binding interaction,[17-19] can be explained by the proximity of the indole ring to the Cu(II) ion complex. Studies by Leporati on ternary complexes of Cu(II) with ethylenediamine-N-acetic acid and amino acids suggest that the increase in stability constants resulting from proximal aromatic side chains is a solvation effect, resulting from a decrease in the hydration of the copper ion.[20] Therefore, the presence of a tryptophan in close proximity to a direct binding site could provide some small contribution to retention in IMAC without directly coordinating the metal ion.

A secondary solvation interaction could explain the apparent contributions of tryptophan to IMAC retention that have been reported in the literature.[17-19] For example, Hemdan and Porath measured the retention of the naturally-occurring amino acids on a Ni(II)IDA column.[17] The authors reported that tryptophan, cysteine and histidine were all retained significantly longer than glycine or other non-aromatic residues. This experiment measured the retention of free amino acids, all of which contain a strong metal-chelating site involving the amino and carboxyl groups, and therefore, any contribution by the side chain is only secondary. The increased retention resulting from the indole ring of tryptophan in this study was most likely due to a solvation effect and not to direct tryptophan interaction as the authors imply.

Finally, the inability of tryptophan to provide retention in the absence of proximal binding sites can be seen in a study by Yip *et al.*, which showed that a peptide containing two tryptophans and no histidines (human gastrin I) was not retained on Cu(II) or Ni(II)

IMAC columns.[21] Therefore, the indole side chain of tryptophan is unable to contribute to IMAC retention without the direct interaction of another proximal ligand.

Non-Competitive Elution

Isocratic elution under non-competitive conditions has been used to analyze the retention behavior of weakly binding cytochrome c variants on a Cu(II)IDA HPLC matrix. The H(-) variant, which contains no surface histidines, has no significant interaction with the matrix and is eluted shortly after the void volume. This is the expected behavior based on the equilibrium isotherm measurements presented in Chapter 3 which also indicated negligible affinity of this variant for the Cu(II)IDA guardgel matrix. The elution of the H(-) variant shortly after the void volume reemphasizes the requirement of surface histidines for IMAC retention on Cu(II)IDA matrices as shown by Zhao *et al.* in a similar set of experiments with lysozymes from different species.[4] The presence of histidine 26 alone results in sufficient retention of the H_{26} variant that it can be easily separated from H(-) variant. The presence of a single fully-accessible histidine (H_8) results in strong binding to the matrix which allows this variant to be completely resolved from the H_{26} variant. Although the band spreading involved in this type of isocratic experiment can be extensive, it provides a simple method to separate the variants based on the existence and accessibility of a single surface histidine.

Isocratic Competition Experiments

The standard competitors used in IMAC experiments are pH and imidazole. There is a fundamental difference in the mode of interaction between these two species. Proton competition takes place in the mobile phase, while imidazole competition takes place at the surface of the matrix. Imidazole competes with the protein's surface histidines for copper sites on the matrix, while protons compete in the mobile phase by binding to surface

histidines. This basic difference in molecular mechanism may impact both how these competitors influence protein retention and how this influence is modeled.

The effect of pH on the retention of the single-histidine variant, H_8, was investigated by isocratic elution. The simple equilibrium model (equation 2) cannot describe the experimental results which indicated that an increased in proton concentration has a more pronounced competitive effect than expected. A possible explanation for this behavior is that at high pH additional weak coordination interactions are forming at the matrix surface. For example, the *N*-terminus of cytochrome c could potentially contribute to IMAC retention above neutral pH. Using glycine methyl ester as a model for the *N*-terminus (pKa 7.7), Table 1.1 suggests that the *N*-terminus would have a log K of 3.2 for Cu(II) at neutral pH, which would provide an interaction only slightly weaker than a surface histidine. The availability of additional ligands at high pH makes the effect of pH on retention very difficult to model.

The effect of imidazole competition on retention was investigated for the H_8 variant. The predicted effect of imidazole competition on protein retention for a single-histidine protein is given by equation 3, which assumes that the protein and imidazole are competing for isolated copper sites. According to equation 3, the data suggest that the K_I for imidazole in this competition experiment is 6.8×10^4 M^{-1}. Yet, the measured binding constant for imidazole to the matrix in the absence of protein was 5.4×10^3 M^{-1}, as given in Chapter 3. Thus, the competitive effect of imidazole is much stronger than would be predicted simply from its binding affinity. Belew *et al.* also observed the strong competitive nature of imidazole, although no mechanism was proposed.[12]

A possible explanation for the strong imidazole competition can be derived from the protein interaction model presented in Chapter 3. For example, equation 3 predicts that if half the copper sites are occupied by imidazole then the retention of the protein will be reduced by a factor of two. But if the protein needs 5 free copper sites in order to interact

at the surface, then an occupation of half of the copper sites by imidazole could completely eliminate the protein interaction. Thus, if the protein occupies multiple copper sites upon binding to the surface, then the presence of imidazole at only a limited percentage of the copper sites may eliminate the protein interaction, as shown in Figure 4.6. This hypothesis provides a means to explain the strong imidazole competition that is still consistent with imidazole's relatively weak binding affinity.

The results obtained for both pH and imidazole competition are consistent with the protein interaction model presented in Chapter 3. For both protons and imidazole, competition is much more effective than would be expected from the simple equilibrium models given by equations 2 and 3. If the protein is occupying multiple copper sites on the surface of the matrix, the proton competition may be enhanced due to the breakage of additional weak interactions at nearby copper sites, while imidazole may only need to occupy a small percentage of copper sites to be an effective competitor (Figure 4.6). The modeling of multiple surface interactions and the effect of competitors represents a significant challenge for future workers in this field.

Gradient IMAC

In all chromatographic studies, the cytochrome c variants were retained in the order predicted by the initial slope of their equilibrium binding isotherms (K'), as given in Table 4.3. In the absence of competitors, the H(-) variant showed no measurable binding in the equilibrium isotherm and no significant retention in chromatography, while the H_{26} and H_8 variants had increased values for K' and showed increased retention. Elution by pH gradient allowed for the separation of proteins which had K' values up to approximately 400 (single-histidine proteins), above which point the competition provided by protons was not sufficient for elution. Elution using an imidazole gradient from 1-10 mM provided for the separation of variants with K' values ranging from approximately 500 to 10,000

(multiple-histidine variants). Weakly binding variants (K'<500) are eluted in the void volume at 1 mM imidazole and increasing concentrations of imidazole provide the necessary competition to elute even strongly binding four-histidine proteins.

The separation of single-histidine proteins in a falling pH gradient may result from differences in either the inherent affinity of the protein (due to histidine accessibility) or the pK_a of the surface histidine involved. The difference in pK_a's between the fully-accessible single-histidine cytochrome c variants, H_4 and H_8 (5.9 and 5.4, respectively), was not sufficient to result in their separation in this pH gradient. In a similar experiment by Hemdan et al., an inverse correlation between histidine pK_a and retention was shown for a set of three single-histidine proteins; cytochrome c (horse), calmodulin (bovine), and lysozyme (chicken).[3] But as indicated in the isocratic pH competition experiments, the pK_a of the single histidine may only play a secondary role in the overall effect of pH on retention, and even though the correlation between histidine pK_a and retention is applicable to the three proteins that Hemdan studied, there are other factors that must be considered.

Competition by imidazole in the gradient mode has proved to be the most versatile of IMAC separation techniques. Imidazole competes effectively with proteins over a wide range of affinities. In the studies presented here, the imidazole gradient was started at 1 mM which is sufficient to elute weak-to-moderately bound variants in the void volume. In a gradient from 1 to 10 mM imidazole, the remaining variants separated into groups according to their histidine multiplicity. Zhao *et al.* also showed that multiple histidine lysozyme homologues were separated in an imidazole gradient according to their histidine multiplicity.[4] Although the exact mechanism of imidazole competition has not been elucidated in these studies, its ability to resolve proteins based on histidine multiplicity has been verified.

The similarity in chromatographic results for the cytochrome c variants used in this study and the lysozyme homologues used by Hemdan et al.[3] suggests a standard

methodology in the application of IMAC. Elution by a pH gradient is effective for separating proteins which contain single surface histidines, but does not provide enough competition to elute proteins bound with multiple histidines, whereas elution using an imidazole gradient provides a powerful method to elute proteins on the basis of their histidine multiplicity.

Conclusions

Using both cytochrome c variants and N-acetyl amino acids, this study has clearly shown that histidine is required for any significant chromatographic retention of proteins on matrices of Cu(II)IDA. Additionally, the effect of proton and imidazole competition on retention was studied, and it was seen that simple equilibrium models, describing the interaction of a single-histidine protein with the Cu(II)IDA matrix in the presence of these competitors, are insufficient to account for the strong competitive effect on retention. Finally, the ability of IMAC to discriminate between proteins based on their histidine multiplicity and microenvironment was investigated. Single-histidine cytochrome c variants are separated based on the microenvironment of their surface accessible histidine (due primarily to differences in the accessibility of the histidine) using isocratic or pH gradient elution. Multiple-histidine cytochrome c variants are separated into groups based on their histidine multiplicity by elution in an imidazole gradient. The order of retention in all cases correlates with the initial slope of the equilibrium binding isotherms as measured in Chapter 3.

Table 4.1 Isocratic elution of *N*-acetyl amino acids at pH 7.6 on a Cu(II)IDA-TSK column. Elution volume (V_e) of representative *N*-acetyl amino acids is given along with the linear equilibrium partition coefficient (K') as determined by equation 1. Histidine is the only *N*-acetyl amino acid that is significantly retained on the Cu(II)IDA column.

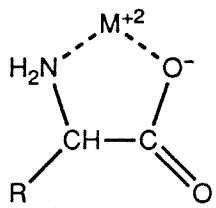
***N*-acetyl amino acid**	**V_e (mL)**	**$K' = (V_e - V_o)/v$**
Glycine	2.3	0.00
Lysine	2.5	0.06
Leucine	2.6	0.09
Phenylalanine	2.7	0.12
Tyrosine	2.8	0.15
Tryptophan	3.3	0.3
Histidine	123.0	36.3

Table 4.2 Elution volume (V_e) of selected cytochrome c variants on a Cu(II)IDA-TSK column at pH 7.0. The predicted retention values are based on the equilibrium binding isotherms presented in Chapter 3 and are calculated using equation 1. The H(-) variant shows no significant retention while the H_{26} variant is weakly retained. The H_8 variant contains a single fully-accessible histidine and is strongly retained.

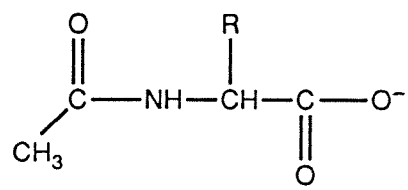
Variant	**Measured V_e (mL)**	**Predicted V_e (mL)**
H(-)	3.1	~ 2.3
H_{26}	22.5	19.8
H_8	380.0	230.0

Table 4.3 Linear equilibrium partition coefficients (K') of iso-1-cytochrome c variants. The equilibrium partition coeficients were calculated based on equilibrium binding isotherms measured in Chapter 3 from $K' = Q_{max}K_p$.

Variant	K'	Number of Surface Histidines
H(-)	~ 0.0	0
H_{26}	3.9	1
H_8	50.5	1
H_4	64.4	1
$H_{26}H_{33}$	385	2
$H_{26}H_8$	780	2
$H_{26}H_4$	990	2
$H_{26}H_{33}H_8$	4700	3
$H_{26}H_{33}H_{39}$	5500	3
$H_{26}H_{33}H_4$	7500	3
$H_{26}H_{33}H_{39}H_8$	N.A.	4



Amino Acid



N-acetyl Amino Acid

Figure 4.1 Illustration of *N*-acetyl blocking group in amino acids. *N*-acetyl blocking group eliminates the strong chelating site that exists between the carboxyl and amino groups in amino acids. In the absence of the amino group, the carboxyl group alone has a log K of only 1.3 for Cu(II). Therefore, any significant retention of the *N*-acetyl amino acids will be due to the side chain (R).

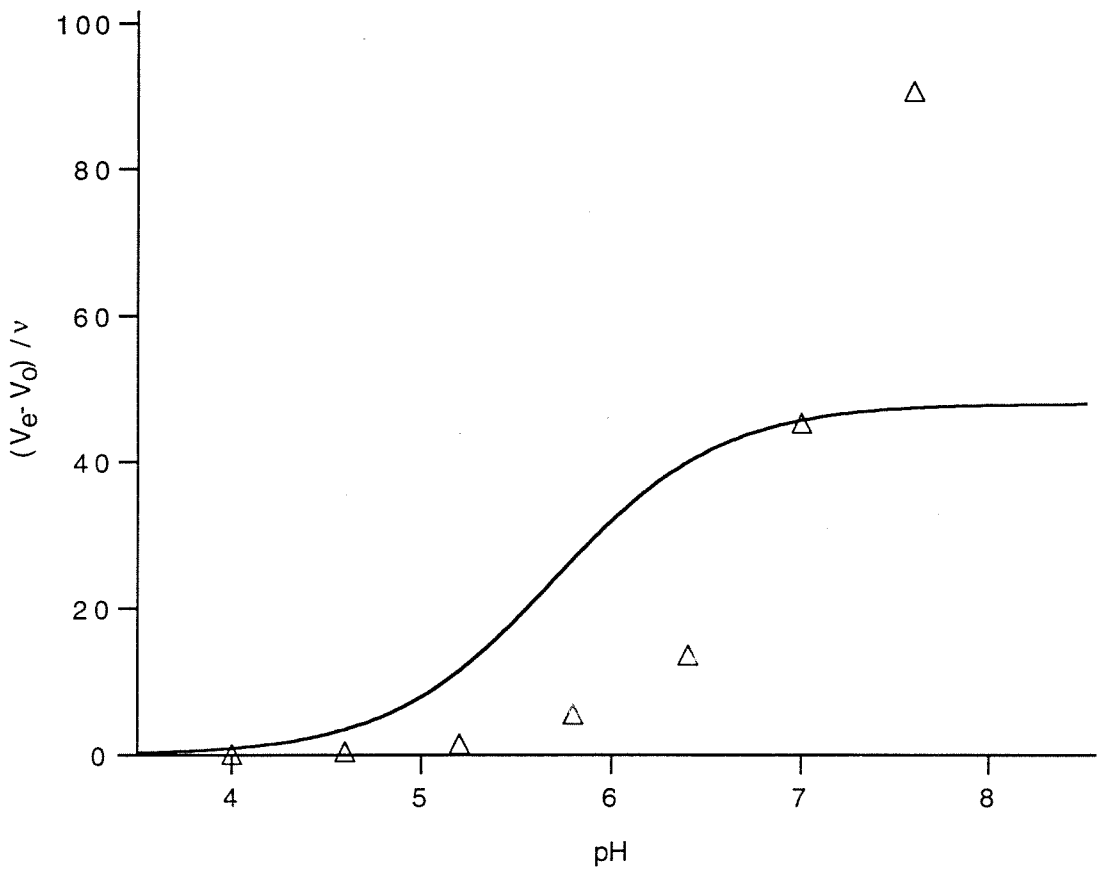


Figure 4.2 Effect of pH on retention behavior of iso-1-cytochrome c H_8 variant (see Materials and Methods). Solid line represents predicted retention behavior based on known pK_a of histidine 8 (5.7) as given by equation 2. The experimental behavior is not consistent with the simple equilibrium model developed based on the protonation of a single histidine (equation 2). (Δ) experimental retention of H_8.

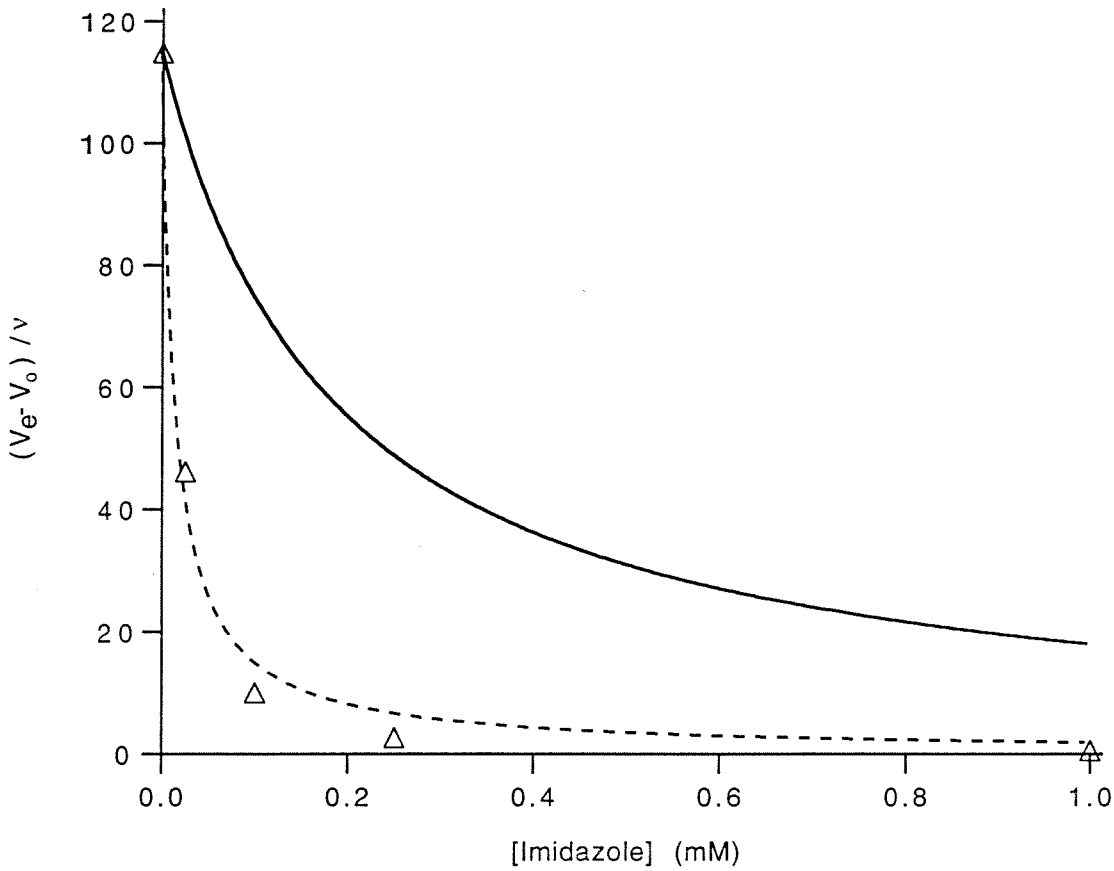


Figure 4.3 Effect of imidazole on the retention behavior of iso-1-cytochrome c H_8 variant (see Materials and Methods). Solid line represents predicted retention behavior based on known binding constant of imidazole (5.4×10^3 M^{-1}) as given by equation 3. Dashed line represents best fit of experimental data to equation 3, and yields a inhibition constant (K_I) of 6.8×10^4 M^{-1}. The experimental data suggest that imidazole is a more effective competitor than would be predicted by its binding constant. (Δ) experimental retention of H_8.

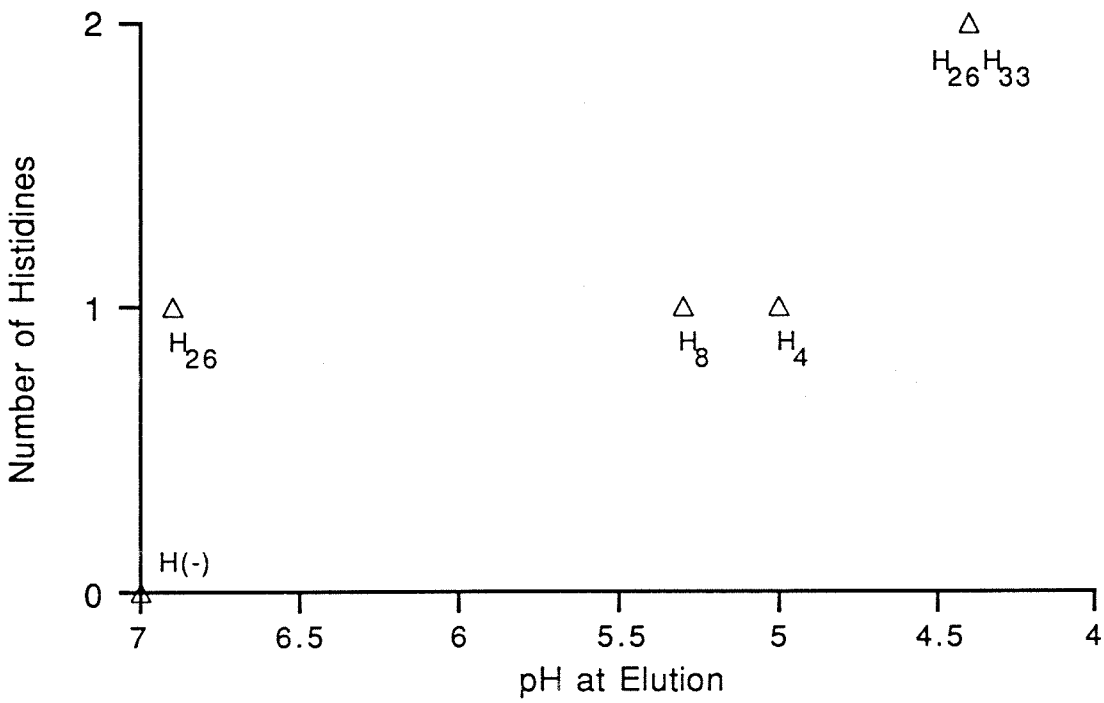


Figure 4.4 Retention of selected iso-1-cytochrome c variants eluted from a TSK Cu(II)IDA column by a linear gradient from pH 7 to pH 4 (see Materials and Methods). Variants are eluted as sharp peaks in the order predicted by their equilibrium binding isotherms (Table 4.3). (Δ) pH at maximum absorbance.

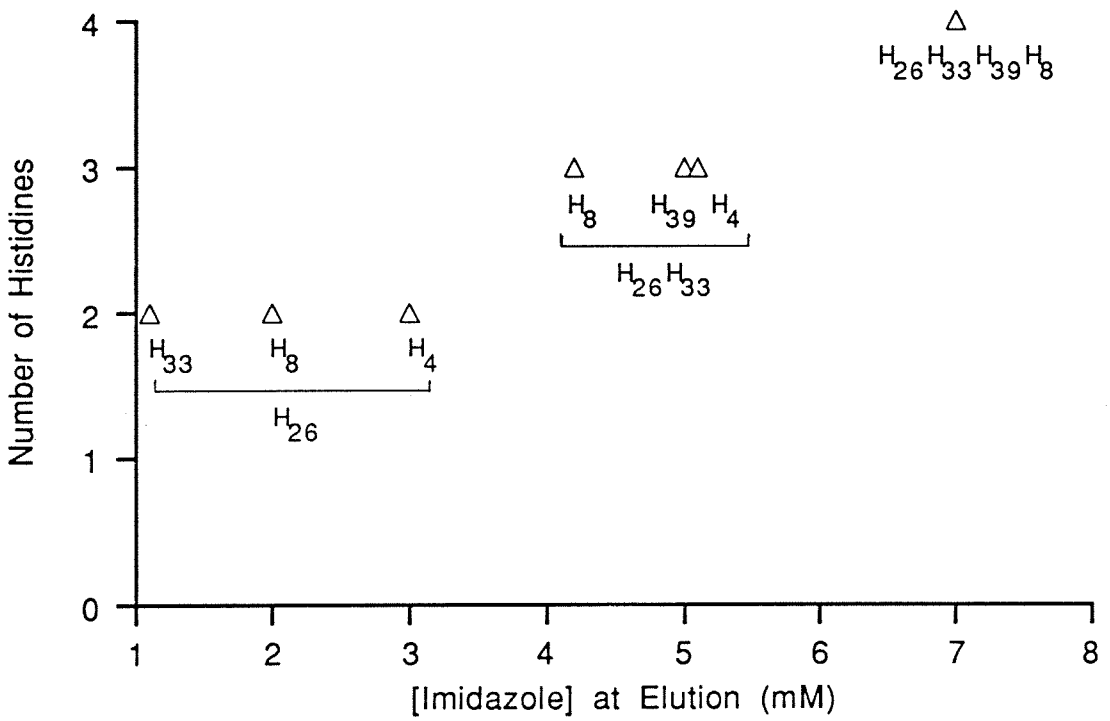


Figure 4.5 Retention of selected iso-1-cytochrome c variants eluted from a TSK Cu(II)IDA column by a linear gradient in imidazole from 1-10 mM (see Materials and Methods). Single-histidine variants are not retained in the loading buffer (1 mM imidazole). Multiple-histidine variants are eluted in groups according to their histidine multiplicity and in accordance with their equilibrium binding isotherms (Table 4.3). (Δ) concentration of imidazole at maximum absorbance.

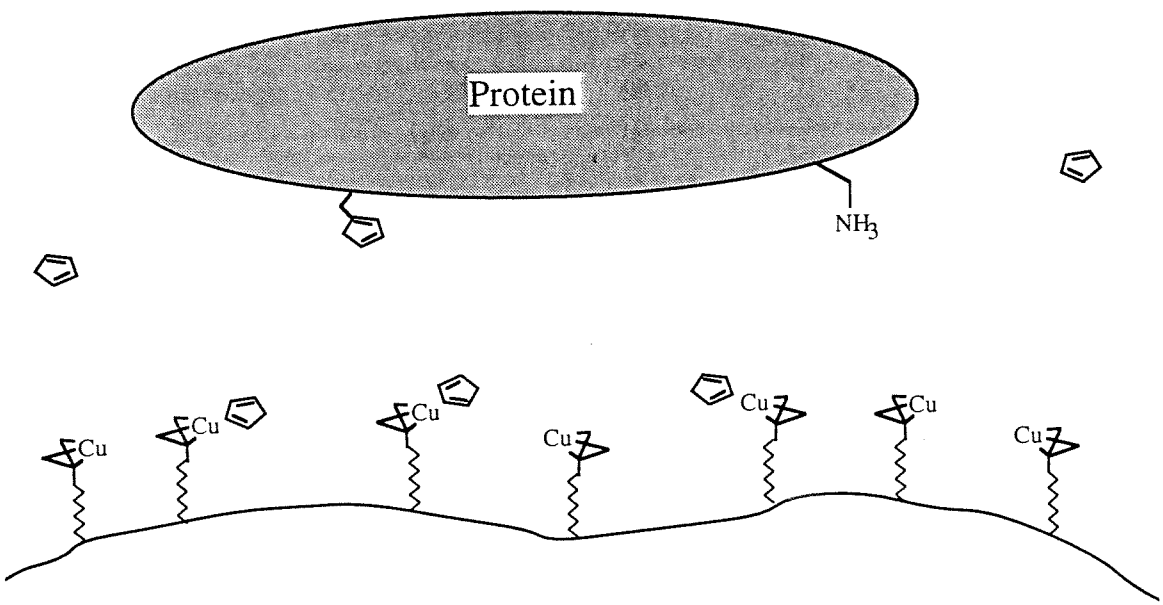


Figure 4.6 Illustration of competition by imidazole at the surface of the Cu(II) matrix. Binding of imidazole at a limited number of sites on the surface may be sufficient to block access of the protein to the surface.

References

1. Wong, J. W., Albright, R. L. and Wang, N. L., *Separation and Purification Methods* **20**, 49-106 (1991).
2. Sulkowski, E., *Trends in Biotechnology* **3**, 1-7 (1985).
3. Hemdan, E. S., Zhao, Y., Sulkowski, E. and Porath, J., *Proceedings of the National Academy of Sciences USA* **86**, 1811-1815 (1989).
4. Zhao, Y.,Sulkowski, E. and Porath, J., *European Journal of Biochemistry* **202**, 1115-1119 (1991).
5. Porath, J., *Trends in Analtytical Chemistry* **7**, 254-259 (1988).
6. Arnold, F. H., Schofield, S. A. and Blanch, H. W., *Journal of Chromatography* **355**, 1-12 (1986).
7. Fassina, G. and Chaiken, I. M., *Advances in Chromatography* **27**, 247-297 (1987).
8. Nichol, L. W., Ogston, A. G., Winzor, D. J. and Sawyer, W. H., *Biochemical Journal* **143**, 435-443 (1974).
9. Dunn, B. M. and Chaiken, I. M., *Proceedings of the National Academy of Sciences USA* **271**, 2382-2385 (1974).
10. Martell, A. E. and Smith, R. M., in *Critical Stability Constants*, (Plenum Press, New York, 1974) vol.1-3.
11. Porath, J., Carlsson, J., Olsson, I. and Belfrage, G., *Nature* **258**, 598-599 (1975).
12. Belew, M., Yip, T., Andersson, L. and Porath, J., *Journal of Chromatography* **403**, 197-206 (1987).
13. Battaglia, L., Corradi, A., Marcotrigiano, G., Menabue, L. and Pellacani, G., *Journal of the American Chemical Society* **102**, 2663-2669 (1980).
14. Udupa, M. and Krebs, B., *Inorganica Chimica Acta* **31**, 251-255 (1978).
15. Johnson, R. D. and Arnold, F. H., *Journal of Cellular Biochemistry* **15g**, 76 (1991).
16. Johnson, R. D., Masters Thesis, California Institute of Technology (1991).
17. Hemdan, E.S. and Porath, J., *Journal of Chromatography* **323**, 255-264 (1985).
18. Hemdan, E.S. and Porath, J., *Journal of Chromatography* **323**, 265-272 (1985).
19. Smith, M. C., Furman, T. C., Ingolia, D. T. and Pidgeon, C., *The Journal of Biological Chemistry* **263**, 7211-7215 (1988).
20. Leporati, E., *Journal of the Chemical Society: Dalton Transactions* **1**, 199-203 (1986).

21. Yip, T., Nakagawa, Y. and Porath, J., *Analytical Biochemistry* **183**, 159-171 (1989).

Chapter 5

High-Affinity Metal-Binding Variants of Iso-1-cytochrome c

Introduction

The ability of proteins to bind metal ions is widely exploited in biological systems to effect catalysis, stabilize folded conformations, and to contribute to the binding of other molecules. There is a growing interest in engineering metal-binding sites into proteins for catalysis,[1,2] regulation of enzyme activity,[3,4] and protein recognition and purification.[5-7] Natural metal-binding proteins use a variety of metal-coordinating groups and elements of secondary structure to create specific binding sites for metal ions. Since these basic elements are common to nearly all proteins, proteins provide a convenient framework for the design of synthetic metal-binding sites. As part of an effort to more fully understand mechanisms of metal recognition in proteins and to explore the design of synthetic metal-binding sites, we have engineered high-affinity metal-binding sites into the surface of *S. cerevisiae* iso-1-cytochrome c.

In general, retention in immobilized metal-affinity chromatography depends upon the number of surface histidines that a protein possesses.[8,9] Since histidine residues are relatively rare in proteins (~1 in 50),[10] the number of strongly-retained naturally-occurring proteins is relatively small. If a protein can be engineered to have high affinity for an IMAC matrix, it should be easily purified by IMAC. The introduction of a metal-chelating site into the surface of a protein is one means to create a very high affinity and hence selective interaction with an IMAC support. The di-histidine metal-chelating site also provides a means to stabilize the protein by providing preferential binding of the metal ion to the folded form of a protein.[11] As depicted in Figure 5.1 for a His-X_3-His metal-chelation site, the secondary structure of the α-helix provides a rigid framework from which two histidines can chelate a single metal ion in the folded state. In the unfolded state, the rigid framework is no longer present, and the entropic loss associated with chelating a single metal ion precludes the formation of a stable di-histidine complex. The difference in binding energies between the chelating interaction in the folded state and the

individual single-histidine interactions in the unfolded state forms the basis for protein stabilization.

The purpose of this study is to assess the affinity of engineered di-histidine sites for metal ions and to evaluate their usefulness in protein purification and stabilization. Metal-binding affinity in solution has been investigated using differential partitioning in PEG/dextran aqueous two-phase systems containing small amounts of Cu(II)IDA-PEG and immobilized metal-affinity chromatography (IMAC) has been used to evaluate binding affinity to a solid support. We have also studied the unfolding (by guanidinium chloride as well as thermal denaturation) of engineered high-affinity metal-binding cytochrome c variants in the presence of Cu(II)IDA by to evaluate metal chelation as a means of protein stabilization.

Materials and Methods

NMR Spectroscopy of iso-1-cytochrome c

^{1}H NMR data were collected at 37° C using AM 500-MHz and AMX 500-MHz Bruker spectrometers. Samples were prepared as given in Chapter 2, adjusted to pH 7.0 (actual reading), and then titrated with Cu(II)IDA.

Two-Phase Partitioning

Two-phase partitioning experiments were used to measure the solution phase binding constants for a number of cytochrome c variants to Cu(II)IDA. The two-phase systems contained 7% MPEG 5000, 4.4% Dextran T500 (Pharmacia), 0.023M sodium phosphate, 0.023 M NaCl at pH 7.6 (4.0 g total weight). For copper-containing systems, Cu(II)IDA-PEG varied from 0.5% (0.8 x 10^{-4} M) to 4% of the total PEG. The Cu(II)IDA-PEG was prepared from MPEG 5000 as described previously.[12] The partitioning experiments were performed at room temperature in the following manner. Cytochrome c

was diluted in buffer (0.5 M NaCl, 50 mM sodium phosphate, pH 7.6) to a concentration of approximately 10 μM. 2.0 grams of the cytochrome c solution was added to 1.3 grams of 13.54 % dextran and 0.7 grams of a 40% PEG solution containing the appropriate amount of Cu(II)IDA-PEG. The system was made up directly in a disposable 4 mL cuvette, equilibrated for 15 minutes by repeated inversion and then centrifuged at 2500 rpm for 15 minutes. The top (PEG) phase was then transferred to a fresh cuvette and the absorbance at 410 nm of both phases was measured directly. The partition coefficient (K) was determined as the ratio of the absorbance in the top phase to that in the bottom. In all cases the partition coefficient in the absence of copper was determined as reference (K_O).

The ratio of the partition coefficients K/K_O can be described by the following model for n independent and identical histidine binding sites,[6,13]

$$\frac{K}{K_O} = \left[\frac{1 + K_p'[M_{tot}]\left[\frac{R+1}{R+1/K_M}\right] + K_a[H]}{1 + K_p''[M_{tot}]\left[\frac{R+1}{RK_M+1}\right] + K_a[H]} \right]^n \qquad (1)$$

where K_p' and K_p'' are the association constants for Cu(II)IDA-PEG interacting with individual histidines on the protein surface. $[M_{tot}]$ is the total metal concentration in the system, R is the phase ratio, and K_M is the partition coefficient of the metal-chelating polymer. The volume ratio (R = 2.52) and the partitioning of the Cu(II)IDA-PEG (K_M = 3.61) were determined previously for this system.[13] In addition, at pH 7.6, the term consisting of the proton association constant (K_a) times the proton concentration ([H]) is much less than 1 for each histidine. Thus, the following simplified equation can be used to determine the binding constants in the top (K_p') and bottom (K_p'') phases.

$$\frac{K}{K_o} = \left[\frac{1 + 1.26\, K_p'[M_{tot}]}{1 + 0.35\, K_p''[M_{tot}]} \right]^n \qquad (2)$$

In a study with natural histidine-containing proteins, values of $K'_a = 2.2 \times 10^3\ M^{-1}$ and $K''_a = 4.5 \times 10^3\ M^{-1}$ were found for the association of Cu(II)IDA-PEG with a single surface histidine using identical experimental conditions.[13]

If two histidines are positioned to chelate a single metal ion, they are no longer independent binding sites. For chelating proteins, a variation of the above equation can be used to determine the association constant of the di-histidine chelating site in the top and bottom phases (K_c' and K_c'', respectively),

$$\frac{K}{K_o} = \left[\frac{1 + 1.26\, K_p'[M_{tot}]}{1 + 0.35\, K_p''[M_{tot}]} \right]^m \left[\frac{1 + 1.26\, K_c'[M_{tot}]}{1 + 0.35\, K_c''[M_{tot}]} \right] \qquad (3)$$

where m is the number of accessible histidines not participating in the high-affinity chelation site.[6]

Gradient IMAC

Experiments were performed as described in Chapter 4. The gradient was extended to 20 mM imidazole in order to elute the high-affinity cytochrome variants. The slope of the gradient remained constant, from 1-20 mM imidazole over 120 minutes.

Guanidinium Chloride Unfolding

Unfolding of cytochrome c variants by guanidinium chloride (Gdm.Cl) was carried out at 25° C using 2 μM protein in 50 mM sodium phosphate, pH 7.6, in the presence of 0-0.8 mM Cu(II)IDA. Unfolding was measured by fluorescence spectroscopy, monitoring

the increase in tryptophan fluorescence that accompanies cytochrome c unfolding.[14,15] The tryptophan was excited at 280 nm and the emission monitored at 346 nm. The fraction of unfolded protein was calculated based on a two-state approximation.[16] The free energy of unfolding is given by,

$$\Delta G_{unf} = - RT \ln K_{unf}$$

where K_{unf} = fraction unfolded /fraction folded.[16]

Thermal Unfolding

Thermal unfolding experiments were carried out with 2 μM protein in 50 mM sodium phosphate, pH 7.6, with 0 and 0.8 mM Cu(II)IDA. Unfolding was monitored by fluorescence spectroscopy. The temperature was increased at a rate of approximately 0.5-1.0° C per minute and was measured directly by a thermocouple in the cuvette.[17] The fraction unfolded was calculated based on a two-state model as given for Gdm.Cl unfolding.

Protein Design

Multidentate protein-metal interactions in rigid elements of secondary structure can potentially provide high-affinity metal-binding sites and increase protein stability. Simple modeling calculations have been carried out on common secondary structural motifs.[18] Chelating di-histidine interactions were found for α-helices (His-X_3-His), β-strands (His-X-His), and reverse β-turns (His-X_2-His). A subsequent survey of structural data showed similar modes of metal binding have been observed in natural metalloproteins. Thermolysin contains a Zn-binding His-X_3-His site in an α-helix,[19] superoxide dismutase a His-X-His site Cu-binding in a β-strand,[20] carbonic anhydrase a His-X-His Zn-binding site in a β-strand,[21] among others. The chelating residues in these naturally occurring

binding sites are buried and therefore unable to interact with metal complexes in solution. To bind metals ions for purification purposes, the binding site must be located on the surface of the protein.

Molecular modeling was used to locate possible sites which would be surface accessible and fulfill the structural requirements for di-histidine chelation. In an effort to retain the native structure of the protein, potential sites were referenced against known destructive mutations and conserved positions among cytochrome c homologous to insure their mutability.[15,22] Residues 4 and 8 in the N-terminal α-helix were chosen as a representative His-X_3-His site in cytochrome c ($H_{26}H_{33}H_4H_8$). In this configuration, the histidines are located on adjacent turns of the α-helix, and the ε2 nitrogens of the imidazole side chains can coordinate a single metal ion (Figure 5.2). Cytochrome c contains primarily α-helical structure,[23,24] but a potential site consisting of histidines at positions 39 and 58 located across from each other in a small section of β-sheet appeared to have the correct geometry needed to coordinate a metal ion ($H_{26}H_{33}H_{39}H_{58}$). In this configuration, the β-sheet structure could provide the rigid scaffolding that enables chelation of the two histidines (Figure 5.3).

Results and Discussion

NMR Spectroscopy

Evidence that the di-histidine sites in iso-1-cytochrome c are able to chelate Cu(II)IDA in solution is provided by ^{1}H NMR studies. Paramagnetic NMR relation effects can be used to identify the sites of Cu(II)IDA binding on a protein surface. When Cu(II)IDA is added at very low concentrations to a protein solution, proton resonances in the NMR spectrum which are close to the binding site of the paramagnetic copper ion are selectively broadened.[25-27] Figures 5.4 and 5.5 show the effect of Cu(II)IDA on the downfield region of the ^{1}H NMR spectrum of the $H_{26}H_{33}H_4H_8$ and $H_{26}H_{33}H_{39}H_{58}$

variants, respectively. The high affinity of the His-X_3-His site for Cu(II)IDA in the $H_{26}H_{33}H_4H_8$ variant can be seen in Figure 5.4. At very low copper concentrations, the paramagnetic copper ion selectively and simultaneously broadens the C2 proton resonances corresponding to histidines 4 and 8. A similar result is obtained for histidines 39 and 58 in the $H_{26}H_{33}H_{39}H_{58}$ variant (Figure 5.5).[28] These experiments suggest that the two histidines involved in the metal-chelation sites are acting cooperatively to bind Cu(II)IDA in solution.

Metal-Affinity Two-Phase Partitioning

Metal-affinity two-phase partitioning was developed as a method for the purification and characterization of proteins.[6,13,29] By attaching an affinity ligand to one of the phase-forming polymers of an aqueous two-phase system, the partitioning of proteins that interact with the ligand can be preferentially altered.[30] This concept has been successfully applied using dyes,[31] ATP,[32] and fatty acids[33] as affinity ligands, among others. Addition of a small amount of Cu(II)IDA-PEG to a two-phase system formed from aqueous solutions of PEG and dextran alters the partitioning of proteins that contain surface histidines.[13] Partition experiments can be used to obtain fundamental information on complexes formed between proteins and metal complexes: binding constants, number of binding sites, and pK_a's of binding sites.[6,13]

Partition coefficients of cytochrome c variants and native proteins containing exposed histidines are presented in Table 5.1. In the absence of Cu(II)IDA-PEG, the partition coefficients (K_O) do not vary greatly among the cytochrome c variants. Addition of Cu(II)IDA-PEG increases the partitioning of proteins that contain surface-accessible histidines into the top (PEG-rich) phase. When surface histidines independently bind polymer-bound metal ions, the difference between the logarithms of the partition coefficients in the absence and presence of Cu(II)IDA-PEG ($\ln K/K_O$) is proportional to the

number of surface histidines (equations 1 and 2). The contribution of histidine 26 to partitioning is negligible compared to the other, more accessible histidines, and therefore, histidine 26 has not been included as a surface-accessible histidine in these results. The $\ln(K/K_O)$ for natural histidine-containing proteins and for the cytochrome c variants has been plotted versus the number of surface-accessible histidines in Figure 5.6. These proteins define a calibration curve for the influence of non-interacting, accessible histidines on partitioning.

The introduction of the His-X_3-His metal-binding site at positions 4 and 8 in cytochrome c increases the partition coefficient dramatically, far more than expected for the addition of two non-interacting histidines. As seen in Figure 5.6, the partition behavior of the $H_{26}H_{33}H_4H_8$ variant at this copper concentration is equivalent to that of a hypothetical protein with nine or ten independent surface-accessible histidines. In contrast to what was observed for the His-X_3-His variant ($H_{26}H_{33}H_4H_8$), the metal-binding site introduced at positions 39 and 58 in the $H_{26}H_{33}H_{39}H_{58}$ variant had minimal effect on partitioning. The $H_{26}H_{33}H_{39}H_{58}$ variant had a partition coefficient similar to the non-chelating four-histidine control protein ($H_{26}H_{33}H_{39}H_8$ variant). The histidine at position 58 appears to be unable to form a chelate with histidine 39 and Cu(II)IDA-PEG in this system.

To determine stability constants for formation of the complexes between surface histidines and Cu(II)IDA-PEG, partition experiments were performed at different concentrations of Cu(II)IDA-PEG. Figure 5.7 shows the effect of Cu(II) concentration on partitioning for the $H_{26}H_{33}H_4H_8$ variant. The value of the stability constant derived from these experiments using equation 3 for the His-X_3-His site in $H_{26}H_{33}H_4H_8$ is 5.3×10^4 M^{-1} for the PEG-rich phase (3.2×10^4 M^{-1} for the dextran-rich phase). This stability constant is 24 times the binding constant for the interaction of Cu(II)IDA-PEG with a single histidine. The difference between the binding free energies in the PEG-rich phase for a single Cu(II)IDA-PEG interaction with the His-X_3-His site (ΔG_C) and two surface

histidines that are not interacting with one another (ΔG_p) is given by,

$$\Delta\Delta G_{bind} = \Delta G_c - \Delta G_p = -RT \ln (Kc') + RT \ln (2Kp')$$

and corresponds to the free energy contribution resulting from di-histidine chelation. Using the experimentally determined values for the binding constants of the His-X_3-His site ($K_c' = 5.3 \times 10^4\ M^{-1}$) and the single histidine interaction ($K_p' = 2.2 \times 10^3\ M^{-1}$) in the PEG-rich phase, the resulting contribution to the binding free energy from chelation ($\Delta\Delta G_{bind}$) is 1.5 kcal/mol.

The incorporation of His-X_3-His sites in α-helices on the surface of bovine somatotropin resulted in binding affinities ranging from 2.0×10^4 to $1.6 \times 10^6\ M^{-1}$ as measured by metal-affinity partitioning.[6] Suh *et al.* suggest that the differences in affinity between the His-X_3-His metal-binding sites are a function of the geometry and flexibility of the helices in which they are located.[6] Attempts to form a chelation interaction with incorrect geometry will result in enthalpic losses due to strain or distortion of the helix backbone and therefore a weaker binding site. Increased flexibility in the binding site will result in entropic losses upon chelation and thus weaker binding. The His-X_3-His site in $H_{26}H_{33}H_4H_8$ cytochrome c was engineered into an unconstrained α-helix at the N-terminus of the protein. The positioning of this metal-binding site in a flexible portion of the protein may be responsible for the relatively low affinity ($5.3 \times 10^4\ M^{-1}$) as compared to some of the bovine somatotropin sites.

Gradient IMAC

The affinity of the engineered metal-binding variants for immobilized Cu(II) was investigated using immobilized metal-affinity chromatography. The concentrations of imidazole required to elute the cytochrome c variants from a Cu(II)IDA-TSK G6000PW column are shown in Figure 5.8 as a function of histidine multiplicity. As with the metal-

affinity partitioning, the retention of the cytochrome c variants containing non-interacting histidines parallel their surface histidine content. The engineered metal-chelating variants require a substantially higher concentration of imidazole for elution than does the $H_{26}H_{33}H_{39}H_8$ variant, even though these proteins each contain four accessible histidines.

An engineered metal-binding site has been used in a successful purification by Brewer *et al.*, in which a high-affinity His-X_3-His site was engineered into bovine somatotropin.[34] This high-affinity metal-binding site made it possible to purify the protein from crude lysate in a single purification step by Cu(II)-IMAC. Genetic approaches have also been used to facilitate protein purification by IMAC.[35-37] In these studies, fusion proteins were prepared by linking the coding sequence of the protein of interest with that of a known high-affinity metal-binding peptide. In the original paper by Hochuli *et al.*, mouse dihydrofolate reductase was fused to a poly-histidine tail and purified to homogeneity on a Ni(II)NTA adsorbent.[35] In a similar manner, Smith *et al.* used the metal-binding dipeptide His-Trp attached to the N-terminus of proinsulin to effect a separation on a Ni(II)IDA matrix.[36] These fusion proteins are easily purified by IMAC, but the recovery of the homogeneous native protein has not been adequately demonstrated.

The two strategies used to facilitate the purification of recombinant proteins by IMAC, protein-engineered high-affinity metal-binding sites and genetically-introduced poly-histidine tails, each have their merits and drawbacks. The "affinity tail" approach offers a method which is generally applicable and does not involve alteration of the native protein. The difficulty, as in all affinity-tail approaches, is that it can be very troublesome to remove the fusion tail selectively to recover the native protein. Incorporating the metal-binding site directly into the native structure involves a larger initial design effort (the primary protein sequence should provide sufficient information) and risks disturbing the structure and/or function of the protein. But once a successful design is found, the engineered protein may be easily purified without the need for further processing.

Protein Stabilization

The ability of high-affinity metal-binding sites to stabilize the folded form of cytochrome c was investigated using guanidinium chloride (Gdm.Cl) and thermal denaturation studies. Cu(II)IDA stabilizes the $H_{26}H_{33}H_4H_8$ and $H_{26}H_{33}H_{39}H_{58}$ variants with respect to unfolding induced by Gdm.Cl, as seen in Figures 5.9 and 5.10, respectively. The fraction unfolded was measured by monitoring the increase in fluorescence of the single tryptophan at position 59 as cytochrome c unfolds. The presence of 0.8 mM Cu(II)IDA increases the amount of Gdm.Cl required to unfold both variants. The resulting stabilization can be quantified by plotting the free energy of unfolding versus the Gdm.Cl concentration, as seen in Figures 5.11 and 5.12. The stabilization resulting from the addition of 0.8 mM Cu(II)IDA ($\Delta\Delta G_{unf}$) can be determined from the difference in the midpoint of the unfolding transition ($\Delta[Gdm.Cl]_0$) and the average slope of the data $(Slope)_{avg}$,[38]

$$\Delta\Delta G_{unf} = \Delta[Gdm.Cl]_0 \, (Slope)_{avg}$$

The measured values for $\Delta\Delta G_{unf}$ are 1.5 kcal/mol for the $H_{26}H_{33}H_4H_8$ variant and 2.5 kcal/mol for the $H_{26}H_{33}H_{39}H_{58}$ variant.

$\Delta\Delta G_{unf}$ for the metal-chelating variants is plotted as a function of Cu(II)IDA concentration in Figure 5.13 (data given in Appendix C). These data display saturation behavior and can be used to determine the maximal stabilization that can be attained by the addition of Cu(II)IDA. The predicted maximum stabilization, $(\Delta\Delta Gunf)_{max}$, for the metal-chelating variants is 1.9 kcal/mol ($H_{26}H_{33}H_4H_8$) and 2.5 kcal/mol ($H_{26}H_{33}H_{39}H_{58}$).

The effect of Cu(II)IDA on thermal denaturation was studied by measuring the melting temperature of the metal-chelating variants in the presence and absence of 0.8 mM Cu(II)IDA (Figure 5.14). Thermal denaturation was monitored by the increase in

tryptophan fluorescence that accompanies unfolding and the melting temperature is the temperature at which the protein is half unfolded (data given in Appendix C). Cu(II)IDA stabilizes the $H_{26}H_{33}H_4H_8$ variant by 0.4° C and the $H_{26}H_{33}H_{39}H_{58}$ variant by 4.5° C.

One possible drawback to using metal-chelating sites to stabilize proteins is that the mutations needed to introduce the chelation site can reduce the stability of the native protein structure ($\Delta\Delta G_{mut}$). The $H_{26}H_{33}H_{39}$ variant contains only a single mutation of the wild-type iso-1-cytochrome c (C102S) and represents the starting point from which we attempted to engineer stability into the protein using metal chelation. In the absence of Cu(II)IDA, both metal-chelating cytochrome c variants had reduced stabilities ($\Delta\Delta G_{mut} \approx$ 1.0 kcal/mol) compared to the wild-type protein ($H_{26}H_{33}H_{39}$), as measured by Gdm.Cl unfolding. As seen in Figure 5.15, the stability gained by the introduction of the chelating site in the presence of 0.8 mM Cu(II)IDA for the $H_{26}H_{33}H_4H_8$ variant ($\Delta\Delta G_{unf}$ = 1.5 kcal/mol) offset the loss in protein stability resulting from mutation of the native protein structure by only 0.5 kcal/mol ($\Delta\Delta G_{stab} \approx$ 0.5 kcal/mol). The stabilization gained from the metal-chelating site in the $H_{26}H_{33}H_{39}H_{58}$ variant ($\Delta\Delta G_{unf}$ = 2.5 kcal/mol) offset the loss in stability of the native structure more substantially, resulting in a $\Delta\Delta G_{stab} \approx$ 1.5 kcal/mol), as seen in Figure 5.16. Similar results were obtained when comparing the melting temperatures of the metal-chelating variants to the $H_{26}H_{33}H_{39}$ cytochrome c variant (T_m = 53.5° C). The mutations required to create the chelating-site in both metal-chelating variants reduced their melting temperature in the absence of Cu(II)IDA by approximately 3° C. The addition of Cu(II)IDA had little effect on the melting temperature of the $H_{26}H_{33}H_4H_8$ variant (T_m = 51.5 ° C) but was sufficient to increase the $H_{26}H_{33}H_{39}H_{58}$ variant to slightly over the original melting temperature (T_m = 54.1° C). Thus, the success of introducing a metal-chelating site into a protein not only depends on the binding strength of the chelating site but also on the effect of the mutations on the native structure.

Figure 5.17 represents the thermodynamic cycle for a two-state folding-unfolding transition in the presence of ligand binding. The free energy of protein stabilization by metal binding ($\Delta\Delta G_{unf}$) is given by the difference in the folding free energy in the absence and presence of the metal ion ($\Delta G_{unf}^{*} - \Delta G_{unf}$). The difference in binding free energy of the metal ion to the folded ($\Delta G_{bind(N)}$) and unfolded states ($\Delta G_{bind(U)}$) is defined as $\Delta\Delta G_{bind}$. By inspection of the thermodynamic cycle, the protein stabilization resulting from metal binding ($\Delta\Delta G_{unf}$) should be equal to the preferential binding of the metal ion in the folded state over the unfolded state ($\Delta\Delta G_{bind}$).[11] In the presence of Cu(II)IDA, the $H_{26}H_{33}H_4H_8$ variants had a predicted maximum increase in stability $(\Delta\Delta G_{unf})_{max}$ of 1.9 kcal/mol (Figure 5.13), as measured by Gdm.Cl denaturation. In comparison the $H_{26}H_{33}H_4H_8$ variant had a $\Delta\Delta G_{bind}$ of 1.5 kcal/mol as measured by aqueous two-phase extraction. These data suggest that the binding energy of the $H_{26}H_{33}H_4H_8$ variant can be translated into protein stabilization in this example. Metal-affinity partitioning studies have shown that $\Delta\Delta G_{bind}$ for a single His-X_3-His site in bovine somatotropin can be as high as 3.5 kcal/mol.[6] The ability to translate this energy into protein stabilization, as was accomplished for the $H_{26}H_{33}H_4H_8$ variant, would have a significant impact on overall protein stability.

The general concept of stabilizing proteins by engineering metal-chelation sites has been applied to short peptides by Ghadiri *et al.* in which His-X_3-His sites were used to induce helicity.[39] In the presence of Cu(II) ions, a 17 amino acid peptide containing a His-X_3-His site displayed approximately 90 percent helicity at 0° C as measured by circular dichroism. The His-X_3-His motif was also used by Handel *et al.* to stabilize a *de novo* designed four-helix bundle protein.[40] This design incorporated a third histidine on an adjacent α-helix, positioned in a manner such that it would form a tridentate interaction with the His-X_3-His site. The three-histidine design resulted in a $\Delta\Delta G_{unf}$ of 2.8 kcal/mol

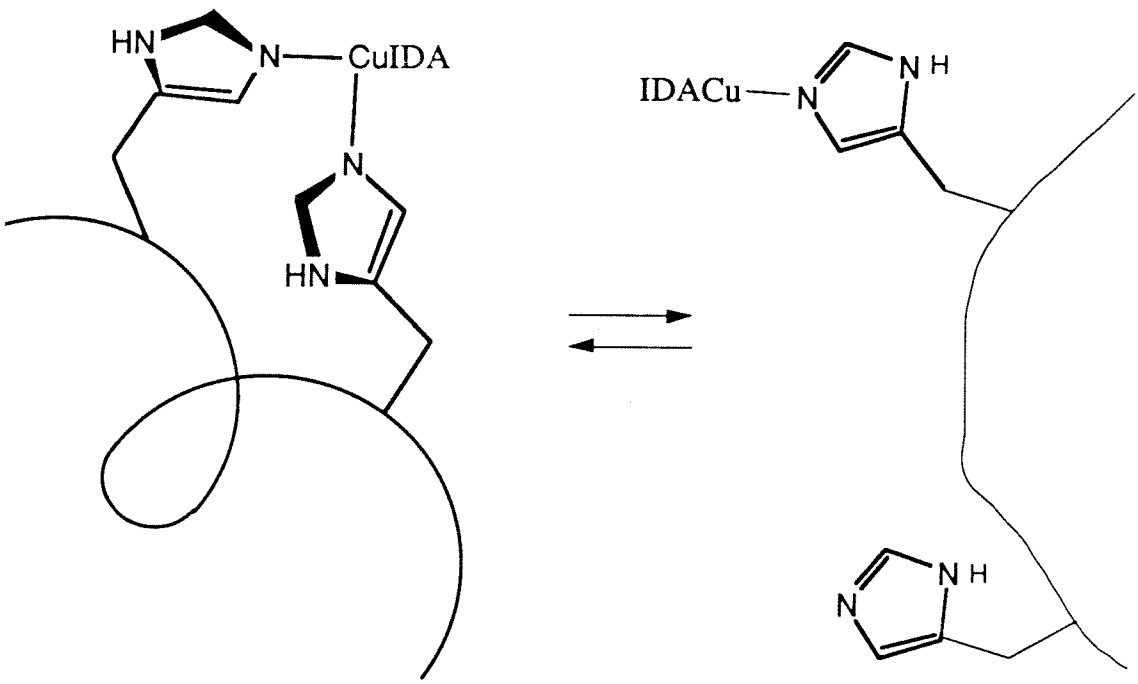
in the presence of 1 mM Zn(II). These studies support the general applicability of the His-X_3-His motif as a mechanism for protein and peptide stabilization.

Conclusions

All evidence collected to this point, with the exception of the two-phase partitioning of $H_{26}H_{33}H_{39}H_{58}$, indicates that the di-histidine metal-binding sites engineered in cytochrome c are able to chelate Cu(II)IDA. Cu(II)IDA titrations of the ^1H NMR spectra of the metal-binding variants show selective and simultaneous broadening of the C2 resonances of the chelating histidines. The His-X_3-His site at positions 4 and 8 in the N-terminal α-helix of cytochrome c binds Cu(II)IDA-PEG in solution 24 times stronger than a single histidine interaction and both variants have increased retention in a Cu(II)-IMAC column relative to a four-histidine control variant. The engineered metal-chelation sites were also shown to stabilize the folded form of cytochrome c in the presence of Cu(II)IDA by as much as 2.5 kcal/mol and may provide a general means by which proteins can be stabilized.

Table 5.1 Metal-affinity partitioning of native heme proteins and *S. cerevisiae* iso-1-cytochrome c variants in the absence (K_O) and presence (K) of 1.6 x 10^{-4} M Cu(II)IDA-PEG. Accessible histidine positions were determined from crystal structures. Histidine 26 of cytochrome c proteins is less than 10 percent accessible and is therefore not counted as an accessible histidine position. All experiments were carried out in duplicate at room temperature, pH 7.6.

Protein	**K_O**	**K**	**Accessible histidines**
Tuna cytochrome c	0.56 ± 0.01	0.57 ± 0.01	0
C. krusei cytochrome c	0.61 ± 0.01	0.83 ± 0.01	2
Horse heart myoglobin	0.41 ± 0.01	0.72 ± 0.01	4
Sheep myoglobin	0.48 ± 0.01	0.97 ± 0.01	5
S. cerevisiae cytochromes c			
$H_{26}H_{33}$	0.52 ± 0.01	0.63 ± 0.01	1
$H_{26}H_{33}H_{39}$	0.51 ± 0.01	0.74 ± 0.01	2
$H_{26}H_{33}H_4$	0.60 ± 0.01	0.84 ± 0.02	2
$H_{26}H_{33}H_8$	0.52 ± 0.01	0.67 ± 0.01	2
$H_{26}H_{33}H_{39}H_8$	0.52 ± 0.01	0.79 ± 0.01	3
$H_{26}H_{33}H_{39}H_{58}$	0.53 ± 0.01	0.83 ± 0.02	3
$H_{26}H_{33}H_4H_8$	0.57 ± 0.01	2.20 ± 0.05	3



Folded Protein **Unfolded Protein**

Figure 5.1 Diagram showing preferential binding of Cu(II)IDA to a folded His-X_3-His site. The α-helix in the folded state provides a rigid framework that positions the histidines such that single metal ion may be chelated. In the unfolded state, the large entropic loss associated with chelating a single metal ion favors independent Cu(II)IDA binding.

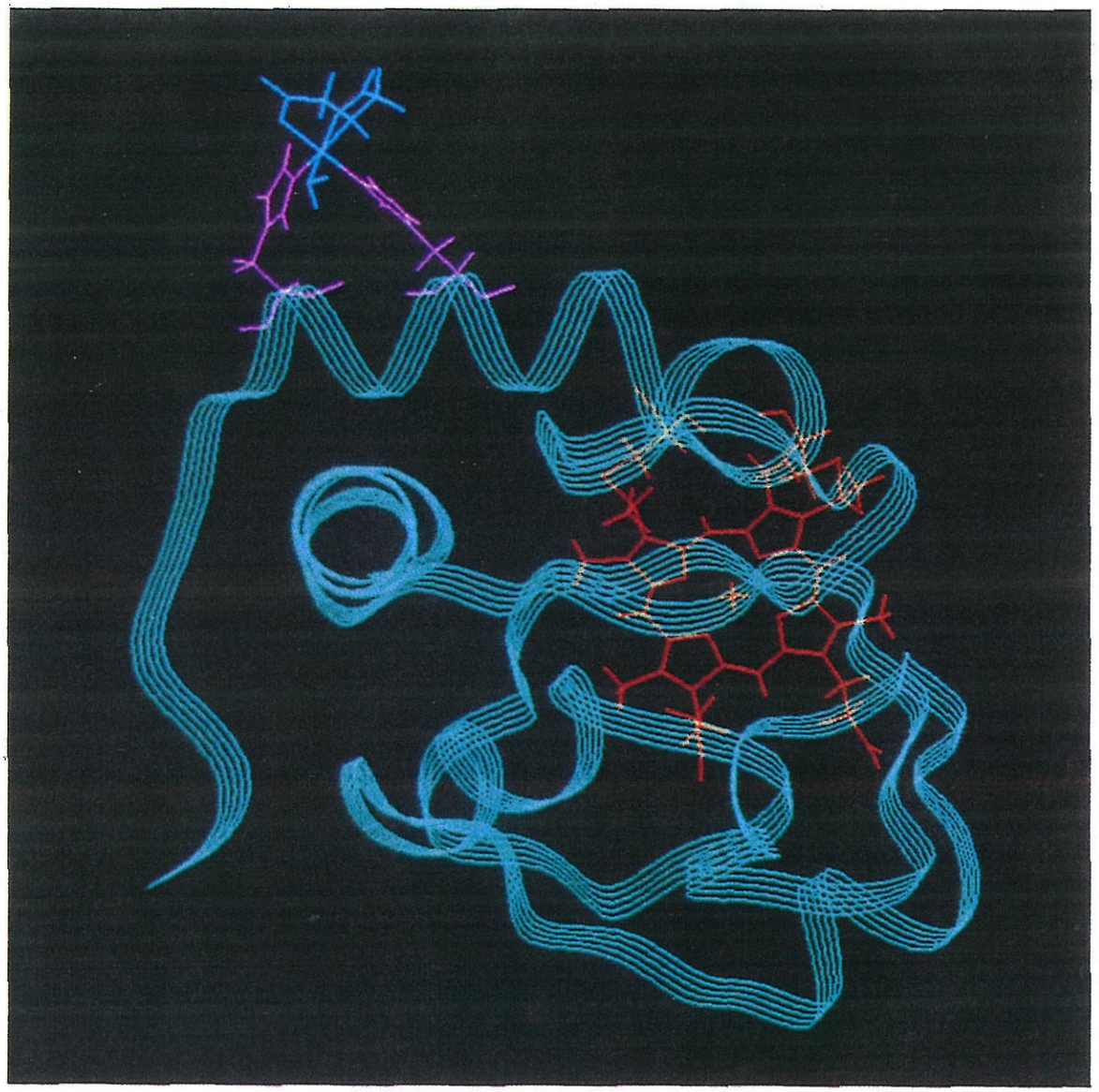


Figure 5.2 Model of His-X_3-His metal-chelating site in $H_{26}H_{33}H_4H_8$ iso-1-cytochrome c. Backbone is shown as a green ribbon and the heme group in red. Histidines at positions 4 and 8 (purple) in the *N*-terminal α-helix are able to coordinate a single Cu(II)IDA (blue).

Figure 5.3 Model of di-histidine metal-chelating site in $H_{26}H_{33}H_{39}H_{58}$ iso-1-cytochrome c. Backbone is shown as a green ribbon and the heme group in red. Histidines at positions 39 and 58 (purple) are able to bridge the anti-parallel β-sheet and coordinate a single Cu(II)IDA (blue).

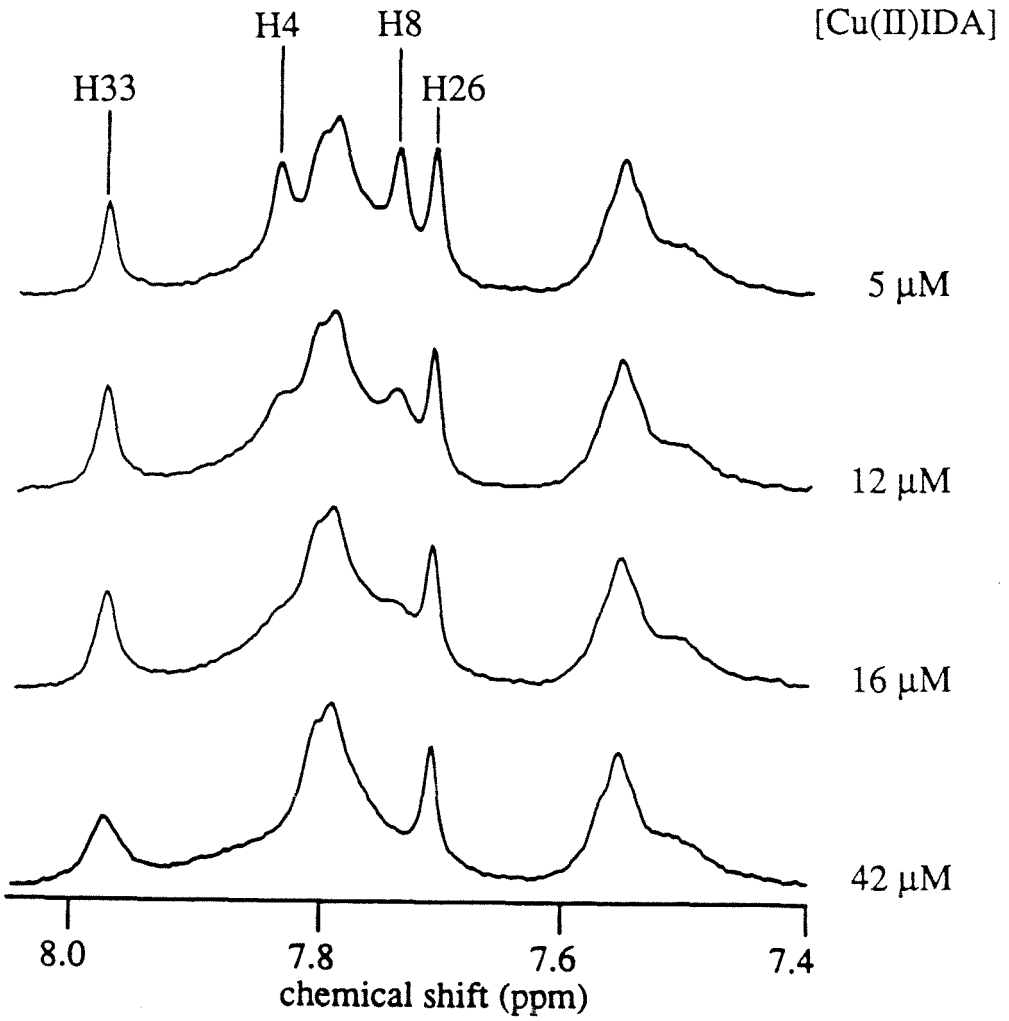


Figure 5.4 ^{1}H NMR spectra from the Cu(II)IDA titration of $H_{26}H_{33}H_4H_8$ iso-1-cytochrome c provides evidence for chelation. The downfield region of the spectra show the selective and simultaneous broadening of the C2 resonances assigned to histidines 4 and 8 at low Cu(II)IDA concentrations (11 μM). The C2 resonance associated with histidine 33 is not significantly broadened until the concentration of Cu(II)IDA is 42 μM.

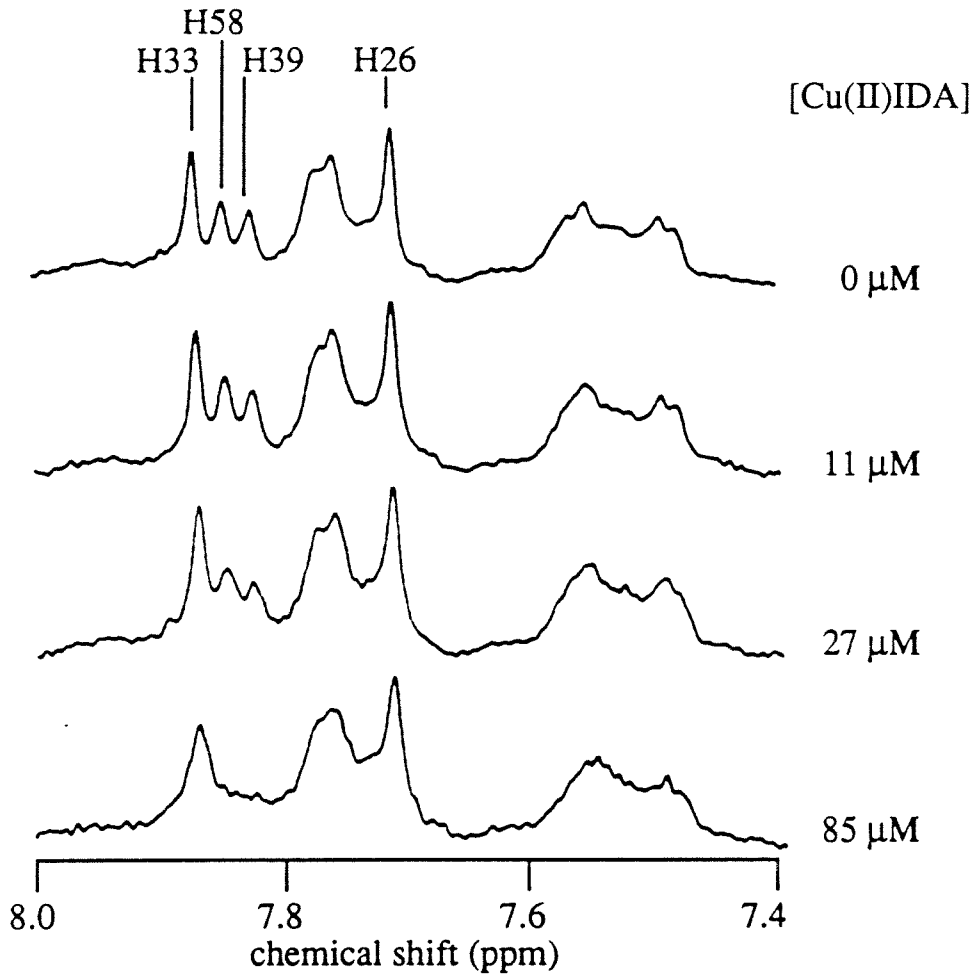


Figure 5.5 ^{1}H NMR spectra from the Cu(II)IDA titration of $H_{26}H_{33}H_{39}H_{58}$ iso-1-cytochrome c provides evidence for chelation. The downfield region of the spectra show the selective and simultaneous broadening of the C2 resonances assigned to histidines 39 and 58 at low Cu(II)IDA concentrations (27 μM). The C2 resonance associated with histidine 33 is not significantly broadened until the concentration of Cu(II)IDA is 85 μM.

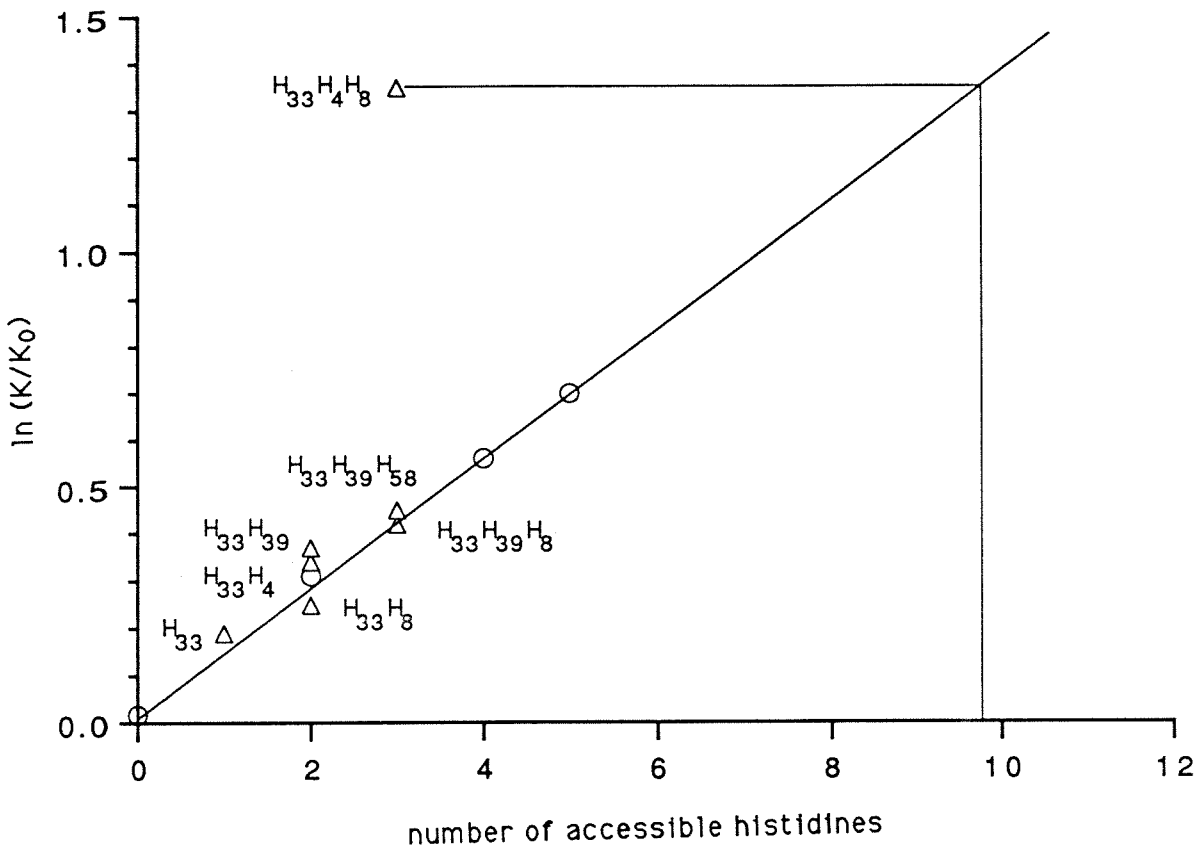


Figure 5.6 Metal-affinity partitioning of native (O) and variant (Δ) proteins in PEG/Dextran aqueous two-phase systems at pH 7.6. Partitioning of native proteins defines a linear calibration for the effect of independent, accessible histidines on $\ln(K/K_O)$. Addition of two histidines (His-X_3-His) in the $H_{26}H_{33}H_4H_8$ variant increases $\ln(K/K_O)$ to a value which corresponds to a hypothetical protein with 9-10 surface histidines. Addition of two histidines at positions 39 and 58 in the $H_{26}H_{33}H_{39}H_{58}$ variant have the expected effect for independent histidines on the partitioning.

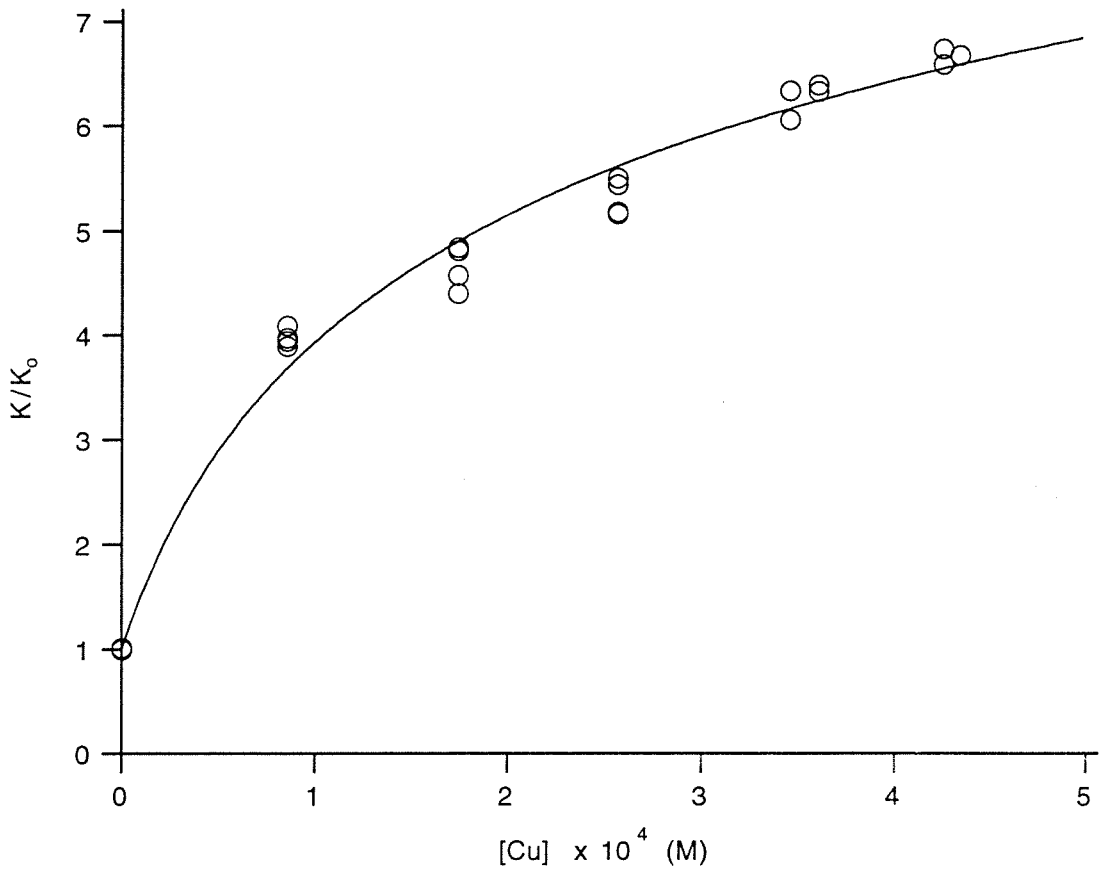


Figure 5.7 Effect of Cu(II)IDA-PEG concentration on partitioning (K/K_O) of *S. cerevisiae* iso-1-cytochrome c His-X_3-His variant ($H_{26}H_{33}H_4H_8$) in PEG/dextran two-phase systems, pH 7.6. The solid line represents the best fit of equation 3 to the partitioning data. $K_p' = 2.2 \times 10^3\ M^{-1}$, $K_p'' = 4.5 \times 10^3\ M^{-1}$, $K_c'=5.3 \times 10^4\ M^{-1}$, and $K_c'' = 3.2 \times 10^4\ M^{-1}$.

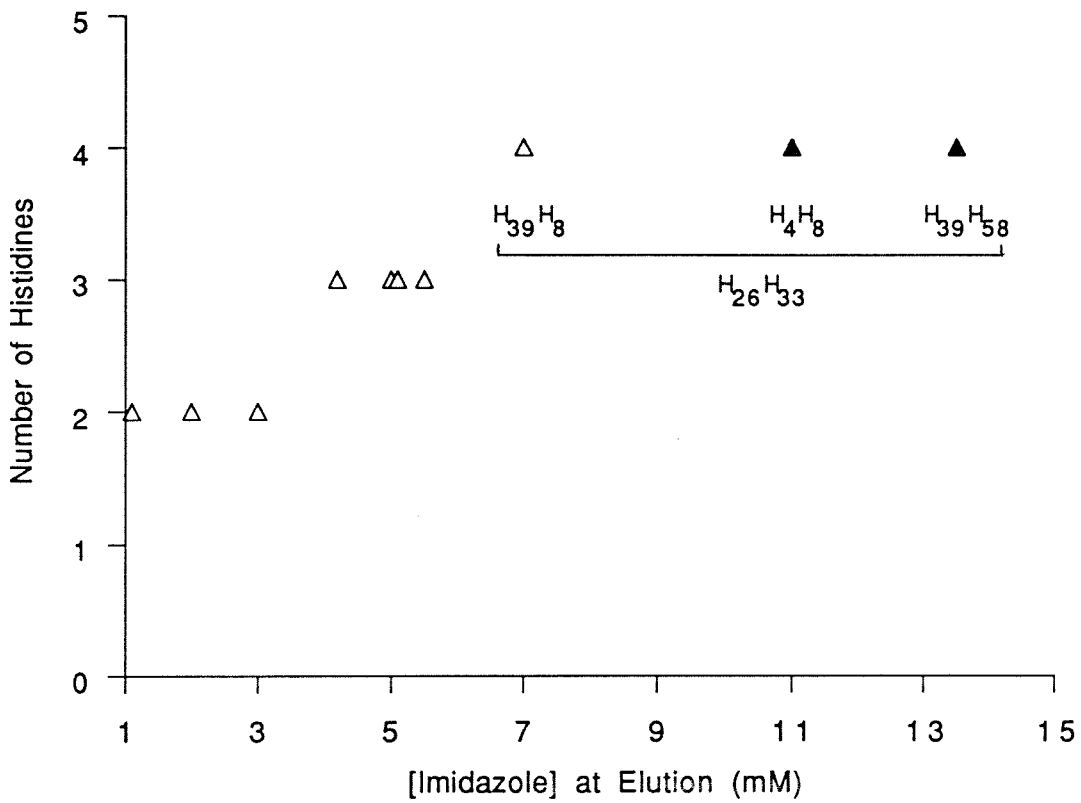


Figure 5.8 Retention of selected iso-1-cytochrome c variants eluted from a TSK Cu(II)IDA column by an imidazole gradient at pH 7.0. Symbols represent the concentration of imidazole at maximum absorbance. Multiple-histidine variants are eluted in groups according to their histidine content. Metal-chelating variants are retained longer than expected for a four-histidine protein. (Δ) non-chelating variants, (▲) metal-chelating variants.

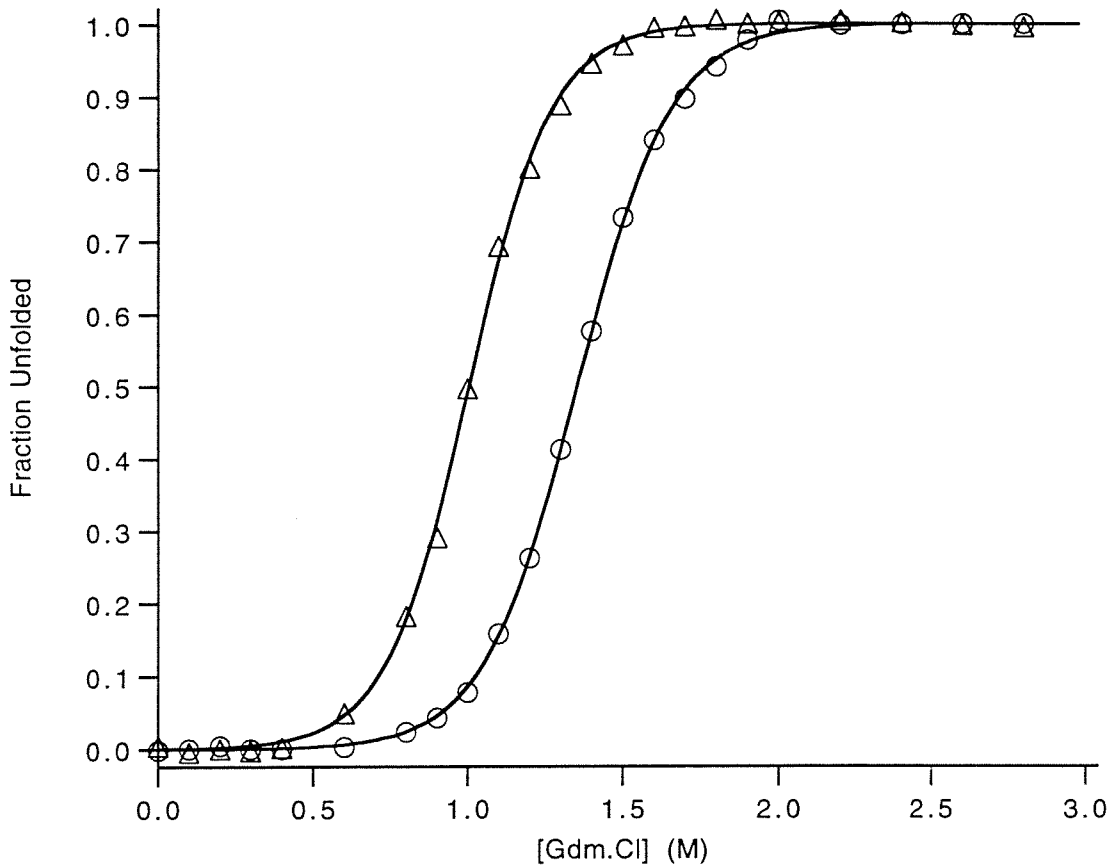


Figure 5.9 Stabilization of $H_{26}H_{33}H_4H_8$ iso-1-cytochrome c variant by Cu(II)IDA at 25° C, pH 7.6. The fraction unfolded is given as a function of guanidinium chloride (Gdm.Cl) concentration in the presence and absence of 0.8 mM Cu(II)IDA. The midpoint of the unfolding transition shifts from 1.0 M Gdm.Cl to 1.35 M Gdm.Cl in the presence of Cu(II)IDA. (Δ) 0.0 mM Cu(II)IDA, (O) 0.8 mM Cu(II)IDA.

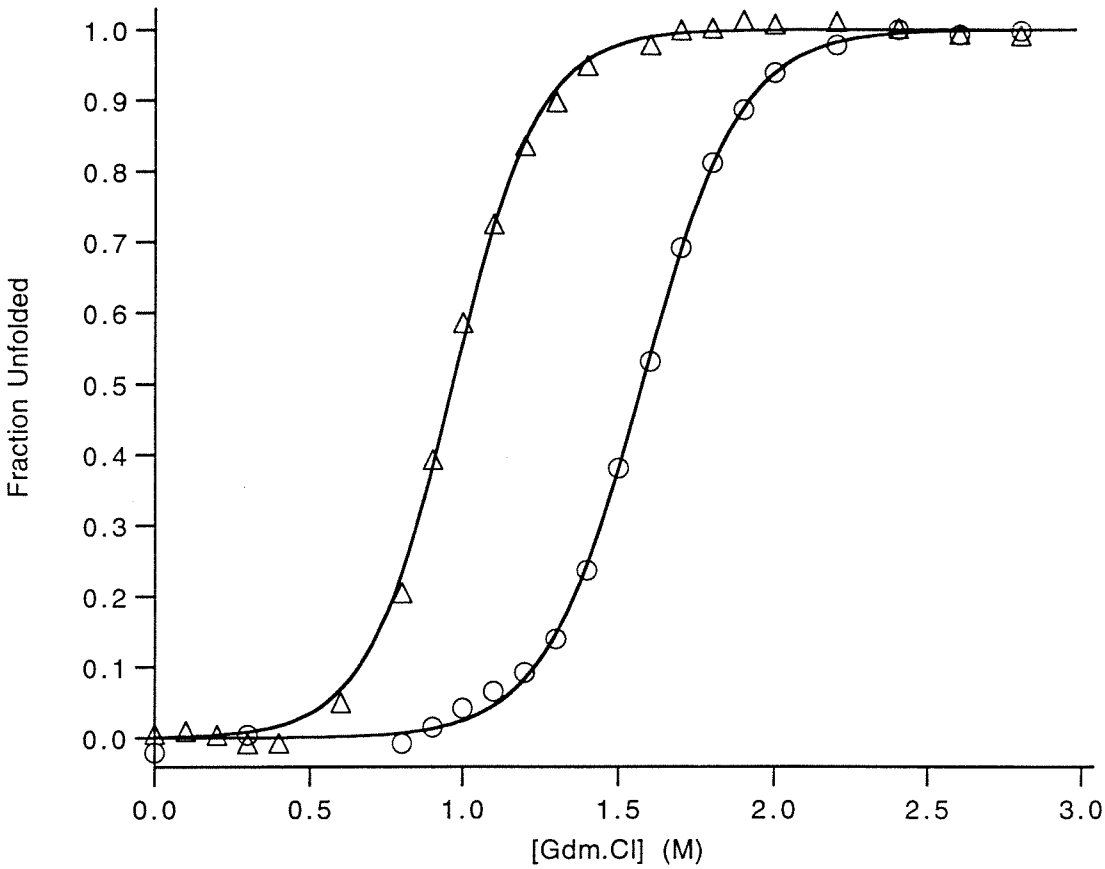


Figure 5.10 Stabilization of $H_{26}H_{33}H_{39}H_{58}$ iso-1-cytochrome c variant by Cu(II)IDA at 25° C, pH 7.6. The fraction unfolded is given as a function of guanidinium chloride (Gdm.Cl) concentration in the presence and absence of 0.8 mM Cu(II)IDA. The midpoint of the unfolding transition shifts from 1.0 M Gdm.Cl to 1.58 M Gdm.Cl in the presence of Cu(II)IDA. (Δ) 0.0 mM Cu(II)IDA, (O) 0.8 mM Cu(II)IDA.

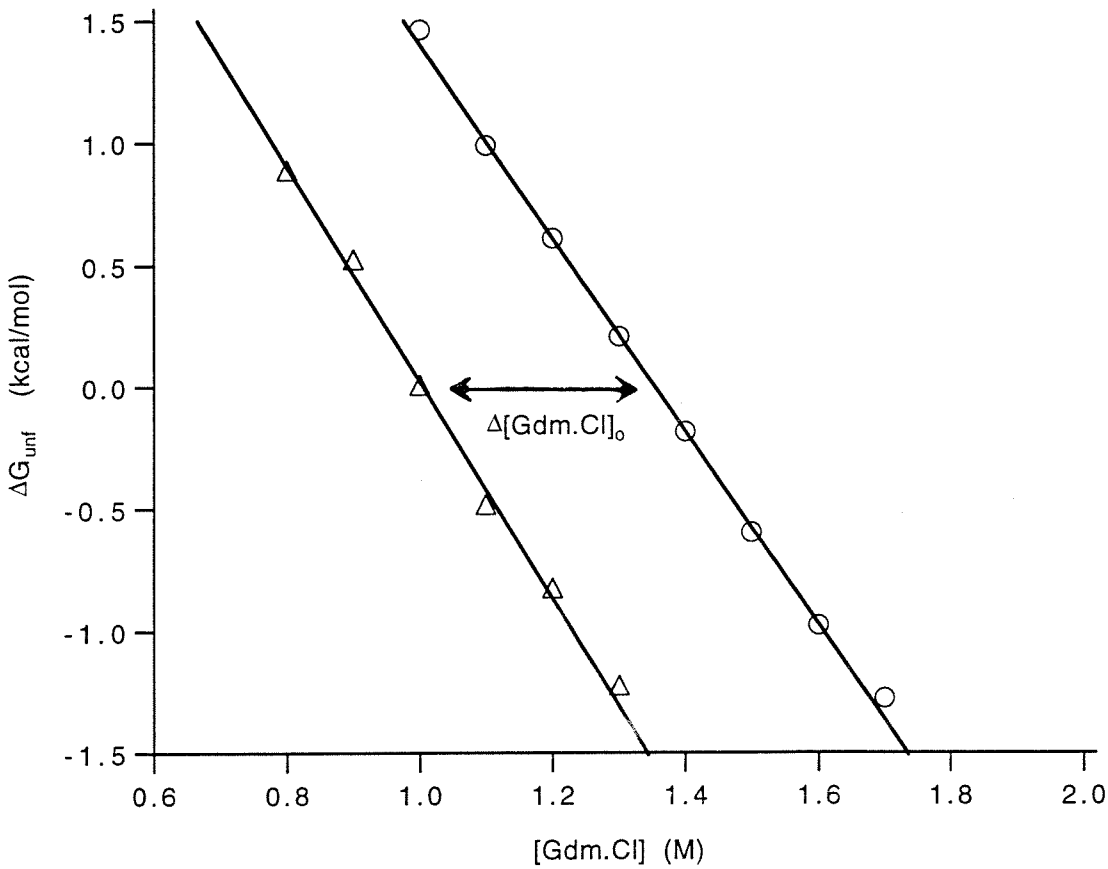


Figure 5.11 Effect of Gdm.Cl on the free energy of unfolding of the $H_{26}H_{33}H_4H_8$ variant in the absence and presence of Cu(II)IDA (25° C, pH 7.6). The stabilization that results from the addition of 0.8 mM Cu(II)IDA in the absence of Gdm.Cl ($\Delta\Delta G_{unf}$) is approximately 1.5 kcal/mol ($\Delta[Gdm.Cl]_0$ = 0.35 M, Slope avg = 4.2 kcal/mol/M).

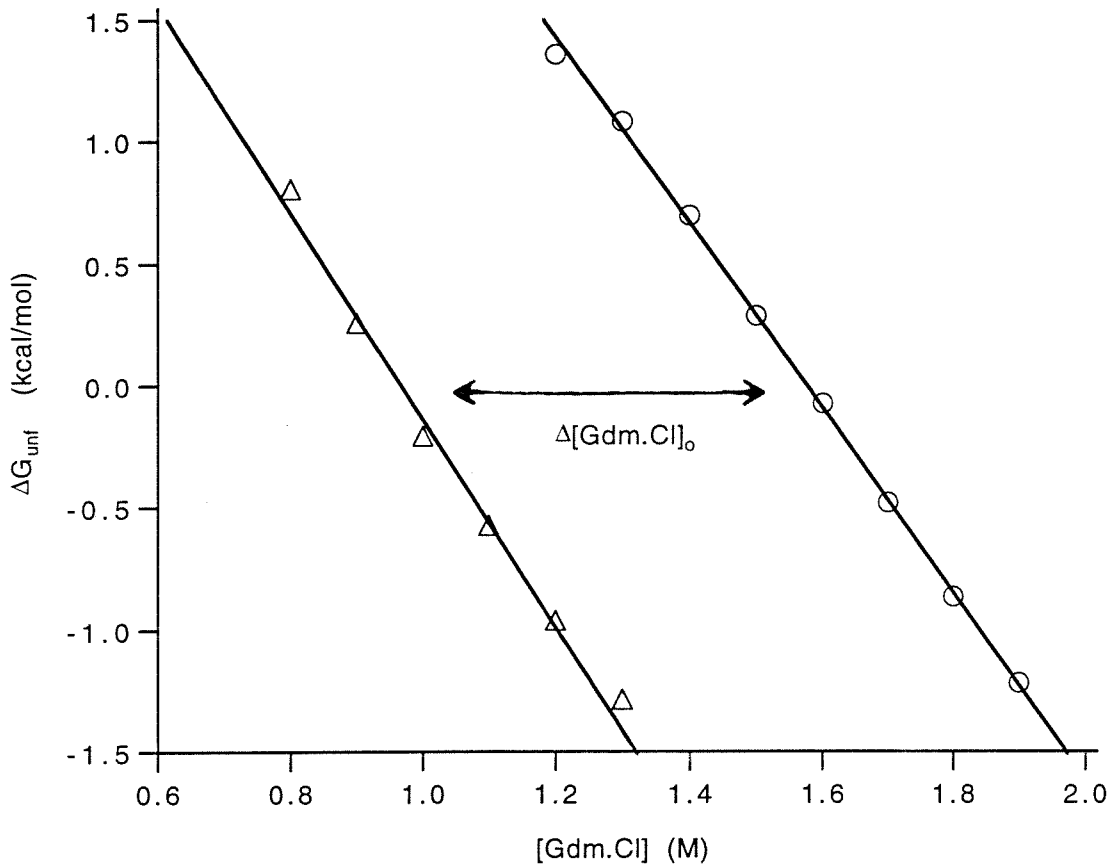


Figure 5.12 Effect of Gdm.Cl on the free energy of unfolding of the $H_{26}H_{33}H_{39}H_{58}$ variant in the absence and presence of Cu(II)IDA (25° C, pH 7.6). The stabilization that results from the addition of 0.8 mM Cu(II)IDA in the absence of Gdm.Cl ($\Delta\Delta G_{unf}$) is approximately 2.5 kcal/mol ($\Delta[Gdm.Cl]_0$ = 0.58 M, Slope avg = 4.2 kcal/mol/M).

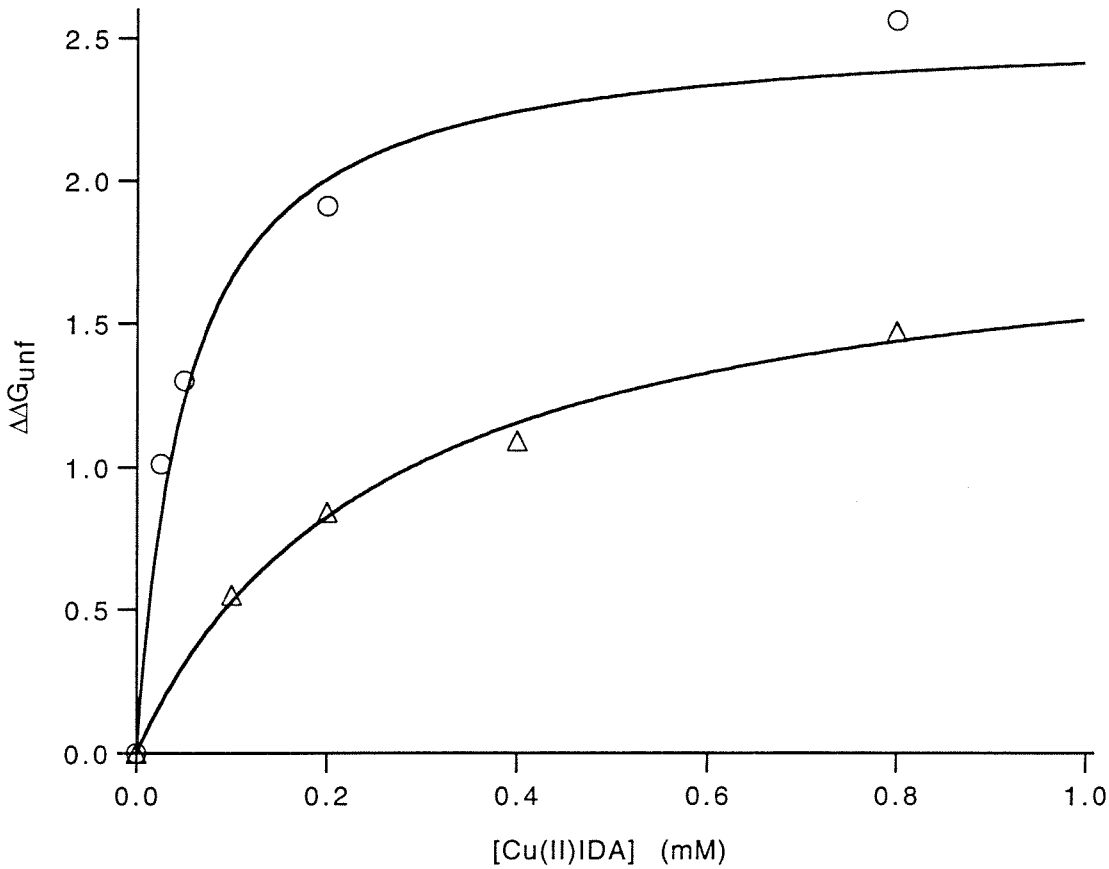


Figure 5.13 Stabilization of metal-chelating variants as a function of Cu(II)IDA as measured by Gdm.Cl unfolding (25° C, pH 7.6). The data display saturation behavior (solid lines represent best fit) and result in the following values for $(\Delta\Delta G_{unf})_{max}$; 2.5 kcal/mol ($H_{26}H_{33}H_{39}H_{58}$) and 1.9 kcal/mol ($H_{26}H_{33}H_4H_8$). (O) $H_{26}H_{33}H_4H_8$, (Δ) $H_{26}H_{33}H_{39}H_{58}$.

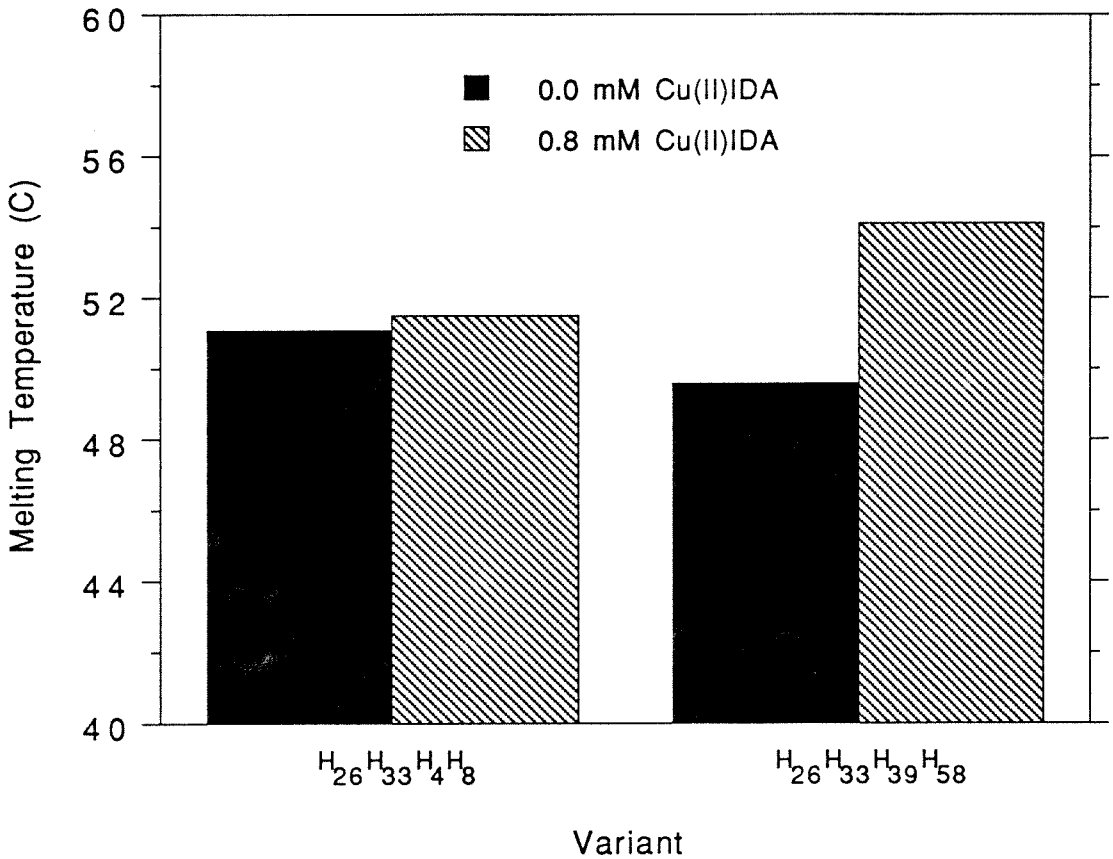


Figure 5.14 Melting temperatures of the metal-chelating variants in the presence and absence of Cu(II)IDA at pH 7.6. Cu(II)IDA increases the melting point of the $H_{26}H_{33}H_4H_8$ and $H_{26}H_{33}H_{39}H_{58}$ variants by 0.4° C and 4.5° C, respectively.

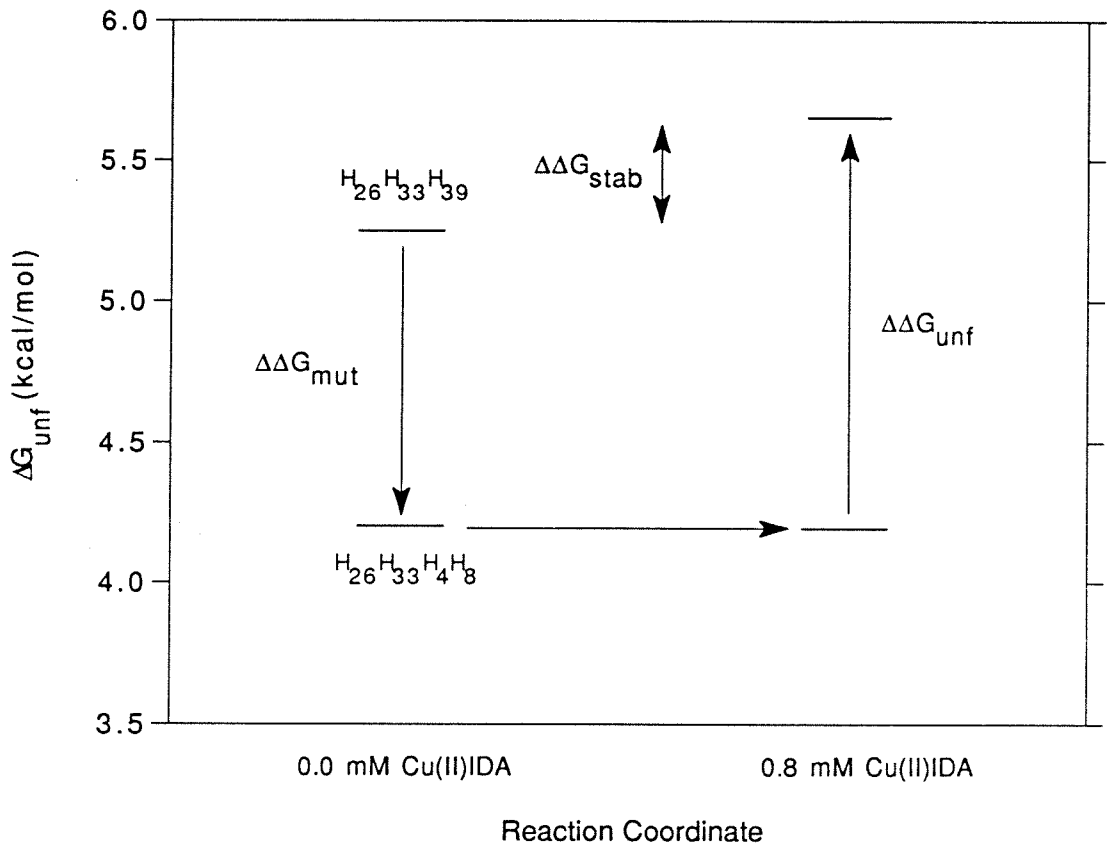


Figure 5.15 Overall stabilization of the $H_{26}H_{33}H_4H_8$ variant compared to the $H_{26}H_{33}H_{39}$ variant as measured by Gdm.Cl unfolding (pH 7.6). The stabilization gained by the introduction of the His-X_3-His metal-chelating site into cytochrome c was insufficient to overcome the destabilization resulting from the mutations ($\Delta\Delta G_{stab}$ = -0.7 kcal/mol).

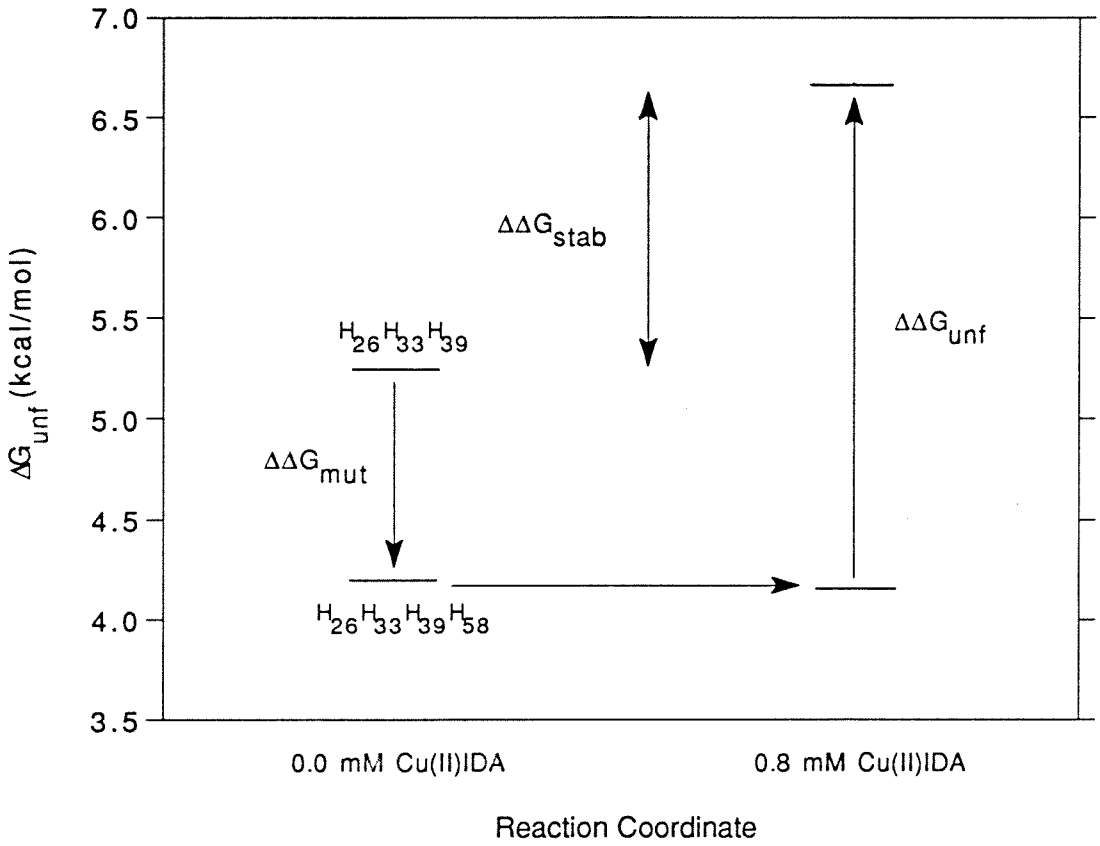


Figure 5.16 Overall stabilization of the $H_{26}H_{33}H_{39}H_{58}$ variant compared to the $H_{26}H_{33}H_{39}$ variant as measured by Gdm.Cl unfolding (pH 7.6). The stabilization gained by the introduction of the β-sheet bridge metal-chelating site into cytochrome c was sufficient to overcome the destabilization resulting from the mutations ($\Delta\Delta G_{stab}$ = 0.3 kcal/mol).

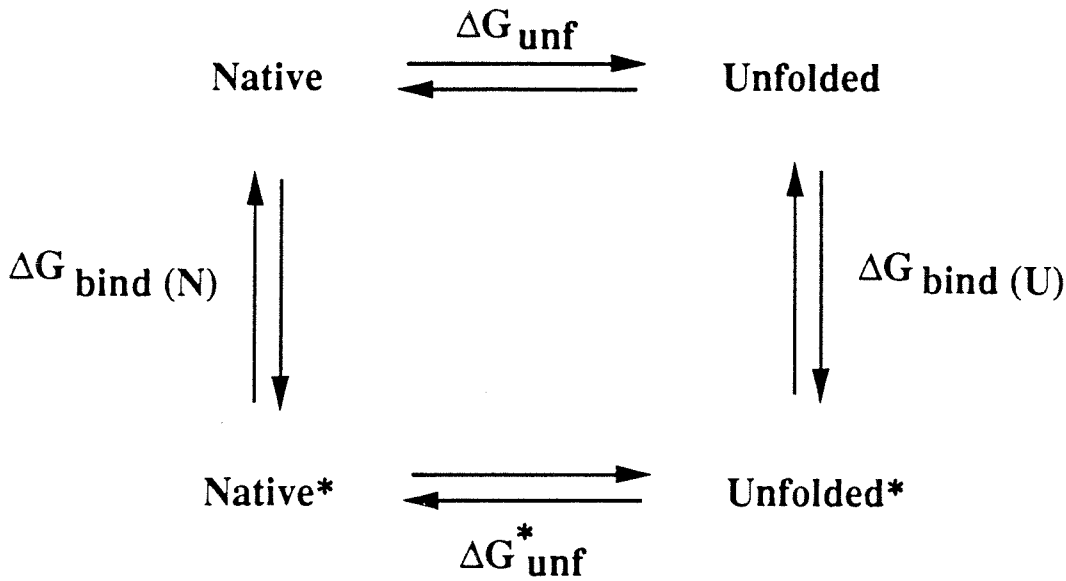


Figure 5.17 Thermodynamic cycle describing stabilization of the native state by preferential binding of a ligand (*e.g.*, a metal ion complex) to the native state. The asterisks denote protein in the ligand-bound state. $\Delta\Delta G_{unf}$ is defined as ΔG_{unf} - ΔG_{unf}^* and $\Delta\Delta G_{bind}$ is defined as $\Delta G_{bind(U)}$ - $\Delta G_{bind(N)}$; by inspection of the cycle, $\Delta\Delta G_{unf} = \Delta\Delta G_{bind}$. Thus, the difference in the free energy of protein unfolding in the presence and absence of the ligand ($\Delta\Delta G_{unf}$) is equal to the energy difference between ligand binding in the native and unfolded states ($\Delta\Delta G_{bind}$).

References

1. Mack, D. P. and Dervan, P. B., *Journal of the American Chemical Society* **112**, 4604-4606 (1990).
2. Iverson, B. L., Iverson, S. A., Roberts, V. A., Getzoff, E. D., Tainer, J. A., Benkovic, S. J. and Lerner, R. A., *Science* **249**, 659-662 (1990).
3. Higaki, J. N., Haymore, B. L., Chen, S., Flatterick, R. J. and Craik, C. S., *Biochemistry* **29**, 8582-8586 (1990).
4. Corey, D. R. and Schultz, P. G., *Journal of Biological Chemistry* **264**, 3666-3669 (1989).
5. Hochuli, E., Bannwarth, W., Dobeli, H., Gentz, R. and Stuber, D., *Bio/Technology* **6**, 1321-1325 (1988).
6. Suh, S.-S., Haymore, B. L. and Arnold, F. H., *Protein Engineering* **4**, 301-305 (1991).
7. Arnold, F. H., *Bio/Technology* **9**, 151-156 (1991).
8. Hemdan, E. S., Zhao, Y., Sulkowski, E. and Porath, J., *Proceedings of the National Academy of Sciences USA* **86**, 1811-1815 (1989).
9. Zhao, Y.,Sulkowski, E. and Porath, J., *European Journal of Biochemistry* **202**, 1115-1119 (1991).
10. Klapper, M. H., *Biochemical and Biophysical Research Communications* **78**, 1018-1024 (1977).
11. Kellis, J. T. Jr., Todd, R. J. and Arnold, F. H., *Bio/Technology* **9**, 994-995 (1991).
12. Van Dam, M., Wuenschell, G. E. and Arnold, F. H., *Biotechnology and Applied Biochemistry* **11**, 492-502 (1989).
13. Suh, S.-S. and Arnold, F. H., *Biotechnology and Bioengineering* **35**, 682-690 (1990).
14. Pryse, K. M., Bruckman, T. G., Maxfield, B. W. and Elson, E. L., *Biochemistry* **31**, 5127-5136 (1992).
15. Moore, G. R. and Pettigrew, G. W., in *Cytochrome c: Evolutionary, Structural and Physicochemical Aspects* (Springer-Verlag, Berlin, 1990).
16. Pace, C. N., *Methods in Enzymology* **131**, 266-280 (1986).
17. Hickey, D. R.,McLendon, G. and Sherman, F., *Journal of Biological Chemistry* **263**, 18298-18305 (1988).
18. Haymore, B. L., Clare, M. C., unpublished results.

19. Holmes, M. A. and Matthews, B. W., *Journal of Molecular Biology* **160**, 623-639 (1982).
20. Tainer, J. A., Getzoff, E. D., Richardson, J. S. and Richardson, D. C., *Nature* **306**, 284-287, (1983).
21. Eriksson, A. E., Jones, T. A. and Liljas, A., *Proteins: Structure, Function and Genetics* **4**, 274-282 (1988).
22. Hampsey, D. M., Das, G. and Sherman, F., *FEBS Letters* **231**, 275-283 (1988).
23. Louie, G. V., Hutcheon, L. B. and Brayer, G. D., *Journal of Molecular Biology* **199**, 295-314 (1988).
24. Louie, G. V. and Brayer, G. D., *Journal of Molecular Biology* **214**, 527-555 (1990).
25. Bertini, I and Scozzafava, A., in *Metal Ions in Biological Systems: Properties of Copper*, H. Sigel, Ed. (Dekker, New York, 1981), vol. 12, pp. 31-74.
26. Johnson, R. D. and Arnold, F. H., *Journal of Cellular Biochemistry* **15g**, 76 (1991).
27. Johnson, R. D., Masters Thesis, California Institute of Technology (1991).
28. Johnson, R. D. and Arnold, F. H. (unpublished results).
29. Wuenschell, G. E., Naranjo, E. and Arnold, F. H., *Bioprocess Engineering* **5**, 199-202 (1990).
30. Flanagan, S. D. and Barondes, S. H., *Journal of Biological Chemistry* **250**, 1484-1489 (1975).
31. Tjerneld, F., Johansson, G. and Joelsson, M., *Biotechnology and Bioengineering* **30**, 809-816 (1987).
32. Persson, L.-O. and Olde, B., *Journal of Chromatography* **457**, 183-193 (1988).
33. Johansson, G. and Shanbhag, V. P., *Jounal of Chromatography* **284**, 63-72 (1984).
34. Brewer, S. J., Haymore, B. L., Hopp, T. P. and Sassenfeld, H. M., in *Purification and Analysis of Recombinant Proteins*, R. Seethram and S. Sharma, Eds. (Dekker, New York, 1991), pp. 239-265.
35. Hochuli, E., Bannwarth, W., Dobeli, H., Gentz, R. and Stuber, D., *Bio/Technology* **6**, 1321-1325 (1988).
36. Smith, M. C., Furman, T. C., Ingolia, D. T. and Pidgeon, C., *The Journal of Biological Chemistry* **263**, 7211-7215 (1988).
37. Ljungquist, C., Breitholtz, A., Brink-Nilsson, H., Moks, T., Uhlen, M. and Nilsson, B., *European Journal of Biochemistry* **186**, 563-569 (1989).

38. Serrano, L., Kellis, J. T., Cann, P., Matouschek, A. and Fersht, A. R., *Journal of Molecular Biology* **224**, 783-804 (1992).

39. Ghadiri, M. R. and Choi, C., *Journal of the American Chemical Society* **112**, 1630-1632 (1990).

40. Handel, T. and DeGrado, W. F., *Journal of the American Chemical Society* **112**, 6710-6711 (1990).

Appendix A

pH Titration Data for Iso-1-cytochrome c Variants

Table A.1 ^{1}H NMR data for the pH titrations of selected iso-1-cytochrome c variants. The assigned histidine is given in parentheses following the ppm header. Data were obtained as described in Chapter 2.

H_4		H_8	
pH	**ppm (4)**	**pH**	**ppm (8)**
4.27	8.57	4.20	8.64
4.80	8.48	4.81	8.45
5.14	8.39	5.20	8.28
5.50	8.27	5.58	8.13
5.85	8.15	5.88	8.00
6.10	8.06	6.16	7.91
6.45	7.95	6.52	7.82
6.77	7.87	6.79	7.76
7.05	7.80	7.14	7.71
7.40	7.74	7.40	7.69
7.85	7.70	7.72	7.67

$H_{26}H_4$		$H_{26}H_8$		$H_{26}H_{33}$	
pH	**ppm (4)**	**pH**	**ppm (8)**	**pH**	**ppm (33)**
4.48	8.54	4.30	8.66	4.57	8.82
5.17	8.40	4.85	8.48	4.88	8.81
5.76	8.18	5.27	8.29	5.26	8.78
6.16	8.04	5.70	8.13	5.66	8.68
6.54	7.93	6.00	7.99	6.03	8.51
6.95	7.80	6.36	7.89	6.28	8.40
7.25	7.76	6.70	7.81	6.50	8.30
7.60	7.72	7.05	7.75	6.69	8.18
		7.51	7.69	6.86	8.10
				7.05	8.02
				7.25	7.92
				7.49	7.86
				7.84	7.80
				8.18	7.78

Table A.2 ^{1}H NMR data for the pH titrations of three-histidine iso-1-cytochrome c variants. The assigned histidine is given in parentheses following the ppm header. Data were obtained as described in Chapter 2.

	$H_{26}H_{33}H_4$			$H_{26}H_{33}H_8$	
pH	**ppm (33)**	**ppm (4)**	**pH**	**ppm (33)**	**ppm (8)**
4.16		8.61	3.43		8.72
4.60	8.80	8.56	3.89		8.70
4.99	8.79	8.47	4.31		8.64
5.43	8.74	8.34	4.57	8.82	8.58
5.77	8.64	8.20	4.89	8.81	8.46
6.02	8.54	8.12	5.17	8.77	8.32
6.30	8.41	8.04	5.43	8.73	8.23
6.56	8.26	7.97	5.73	8.66	8.11
6.72	8.17	7.92	6.01	8.54	8.00
6.90	8.06	7.87	6.29	8.40	7.92
7.22	7.93	7.81	6.54	8.25	7.85
7.44	7.85	7.76	6.80	8.13	7.80
7.82	7.79	7.73	7.04	8.01	7.77
8.23	7.76	7.70	7.37	7.87	
			7.75	7.79	7.70
			8.27	7.76	7.69

	$H_{26}H_{33}H_{39}$			$H_{26}H_{33}H_{58}$	
pH	**ppm (33)**	**ppm (39)**	**pH**	**ppm (33)**	**ppm (58)**
4.89	8.79	8.69	4.76	8.81	8.61
5.26	8.77	8.68	5.18	8.77	8.59
5.54	8.68	8.61	5.46	8.71	8.55
5.89	8.57	8.51	5.81	8.58	8.46
6.18	8.38	8.35	6.02	8.47	8.38
6.53	8.22	8.18	6.27	8.34	8.27
6.78	8.09	8.05	6.51	8.20	8.14
7.03	7.99	7.93	6.82	8.06	7.98
7.34	7.87	7.77	7.16	7.93	7.83
7.63	7.82	7.71	7.53	7.84	7.71
			7.81	7.80	7.65

Table A.3 ^{1}H NMR data for the pH titrations of the four-histidine iso-1-cytochrome c variants. The assigned histidine is given in parentheses following the ppm header. Data were obtained as described in Chapter 2.

$H_{26}H_{33}H_{39}H_8$

pH	ppm (33)	ppm (39)	ppm (8)
3.56			8.71
3.94			8.69
4.29			8.66
4.55	8.82	8.72	8.59
4.92	8.80	8.70	8.45
5.21	8.77	8.67	8.34
5.57	8.69	8.61	8.18
5.81	8.61	8.54	8.08
6.04	8.52	8.47	8.00
6.32	8.35	8.32	7.91
6.64	8.18	8.14	7.83
6.94	8.03	7.97	7.78
7.21	7.93	7.85	7.74
7.61	7.83	7.78	7.71
8.04	7.77	7.70	
8.52	7.76	7.68	

$H_{26}H_{33}H_4H_8$

pH	ppm (33)	ppm (4)	ppm (8)
4.57	8.81	8.48	8.44
5.00	8.79	8.43	8.38
5.21	8.78		
5.50	8.72	8.27	8.18
5.78	8.62	8.16	8.04
6.03	8.52	8.08	7.95
6.23	8.43	8.02	7.90
6.48	8.29	7.96	7.84
6.69	8.17	7.91	7.79
6.92	8.04	7.85	7.75
7.16	7.94	7.81	7.72
7.42	7.86	7.77	7.69
7.71	7.79	7.74	7.67
8.55	7.75	7.69	7.64

Table A.4 ^{1}H NMR data for the pH titrations of the $H_{26}H_{33}H_{39}H_{58}$ iso-1-cytochrome c variant. The assigned histidine is given in parentheses following the ppm header. Data were obtained as described in Chapter 2.

	$H_{26}H_{33}H_{39}H_{58}$		
pH	**ppm (33)**	**ppm (39)**	**ppm (58)**
4.45	8.81	8.69	8.69
5.08	8.78	8.61	8.61
5.39	8.73	8.55	8.53
5.75	8.60	8.43	8.41
6.06	8.47	8.33	8.31
6.31	8.32	8.24	8.22
6.54	8.19	8.15	8.14
6.76	8.08	8.06	8.06
7.00	7.98	7.98	7.98
7.30	7.89	7.87	7.85
7.57	7.83	7.77	7.75
7.91	7.78	7.71	7.67

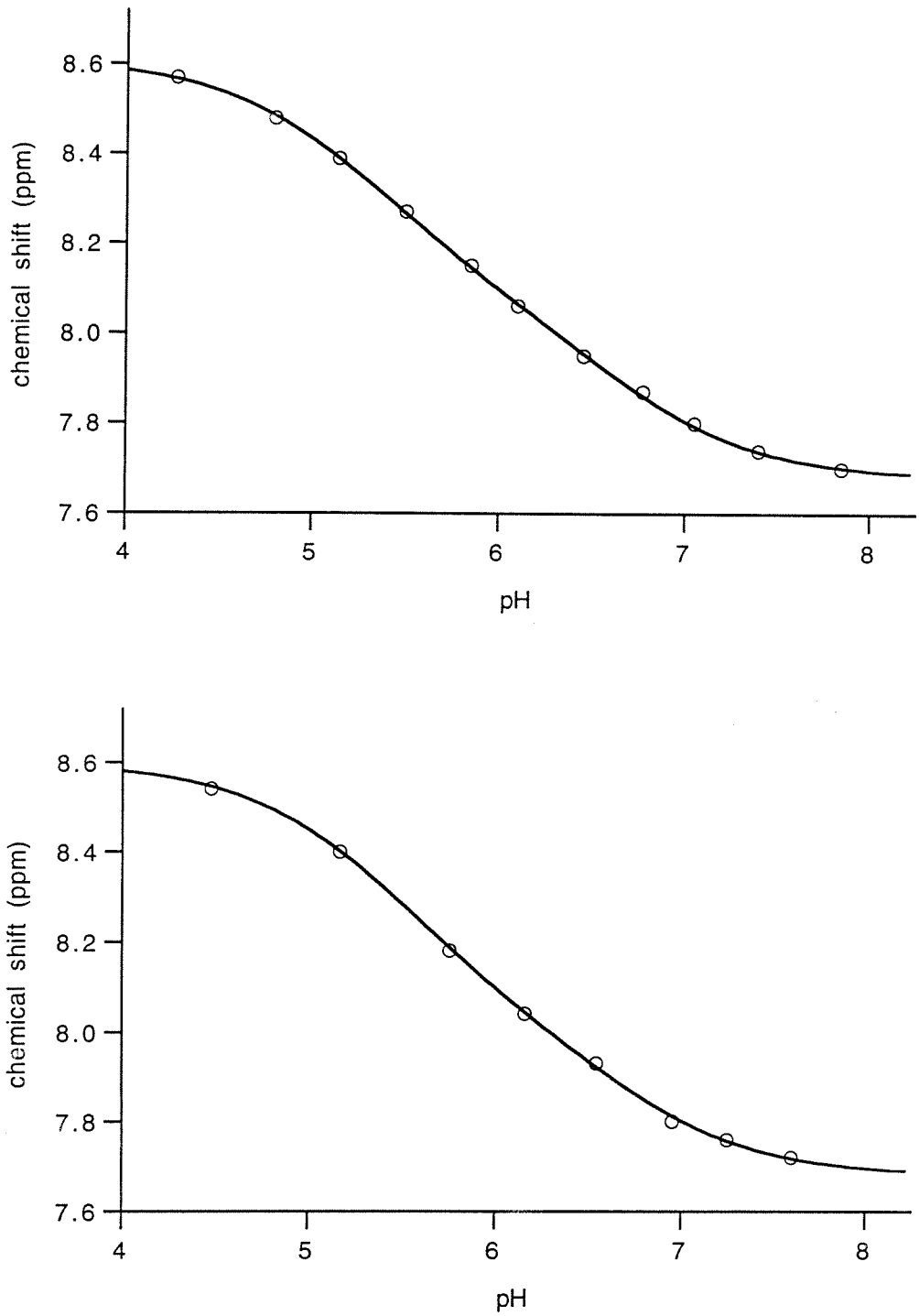


Figure A.1 pH titration behavior of H_4 (top) and $H_{26}H_4$ (bottom) iso-1-cytochrome c variants. The chemical shift in ppm of the C2 proton of histidine at position 4 is shown. The data is fit with a non-standard pH titration curve which yields an apparent pK_a of 5.88 (H_4) and 5.86 ($H_{26}H_4$).

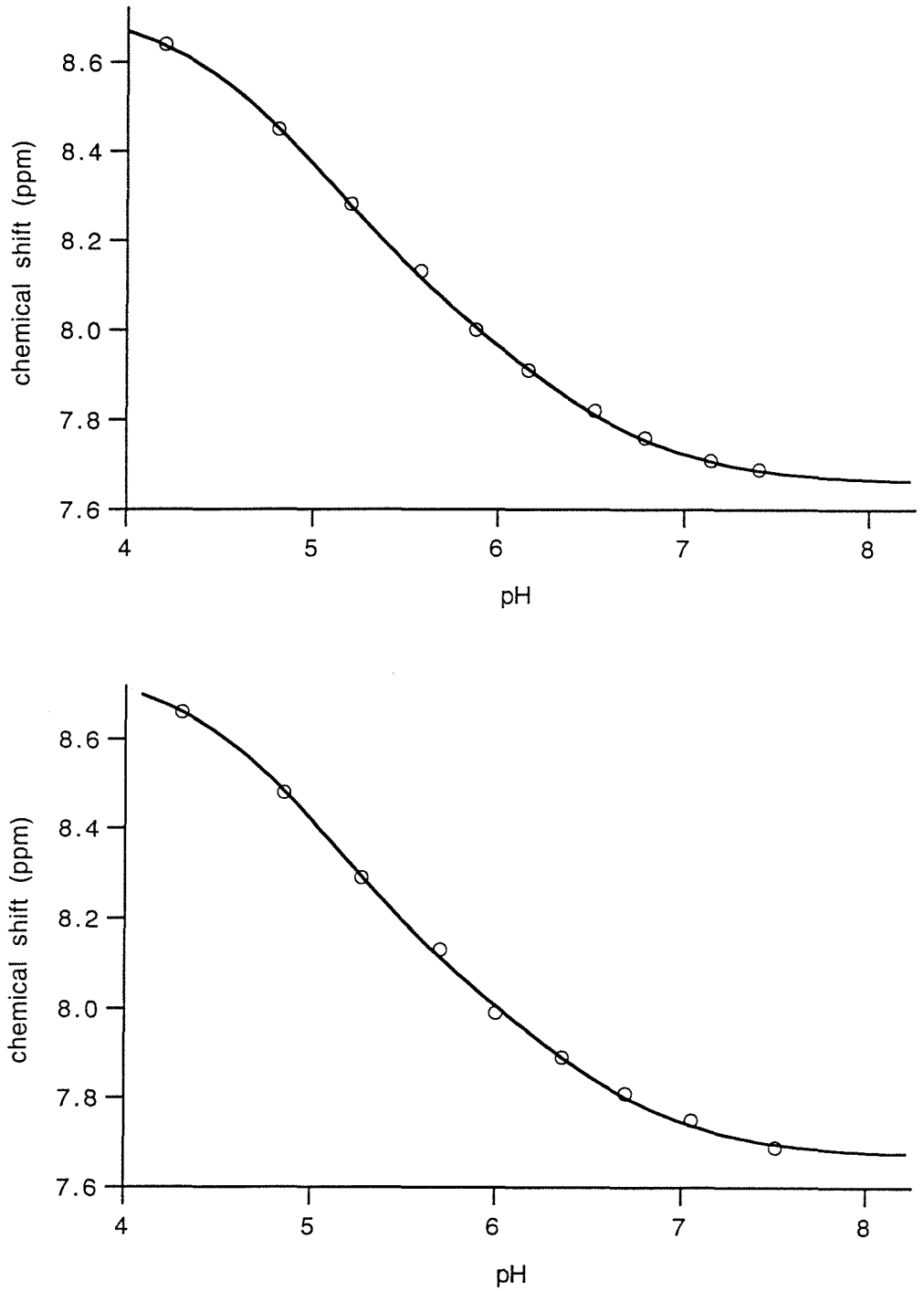


Figure A.2 pH titration behavior of H_8 (top) and $H_{26}H_8$ (bottom) iso-1-cytochrome c variants. The chemical shift in ppm of the C2 proton of histidine at position 8 is shown. The data is fit with a non-standard pH titration curve which yields an apparent pK_a of 5.39 (H_8) and 5.40 ($H_{26}H_8$).

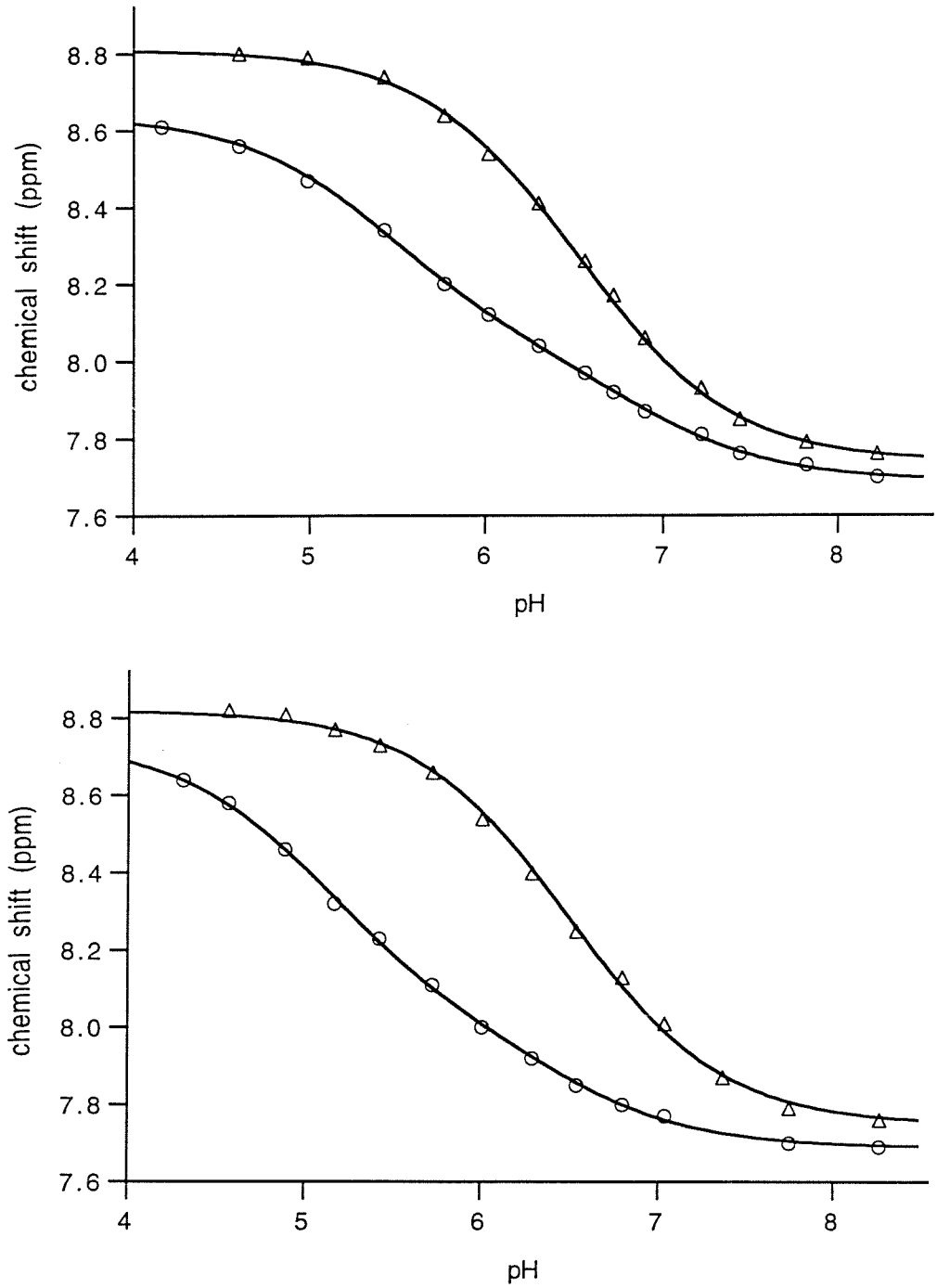


Figure A.3 pH titration behavior of $H_{26}H_{33}H_4$ (top) and $H_{26}H_{33}H_8$ (bottom) iso-1-cytochrome c variants. The chemical shifts in ppm of the C2 protons of histidines at positions 33 (Δ) and 4 or 8 (O) are shown. The data for histidine 33 is fit with a standard pH titration curve and yields pKa's of 6.52 ($H_{26}H_{33}H_4$) and 6.50 ($H_{26}H_{33}H_8$). The data for histidines 4 and 8 are fit with a non-standard pH titration curve which yields an apparent pK_a of 5.90 ($H_{26}H_{33}H_4$) and 5.43 ($H_{26}H_{33}H_8$).

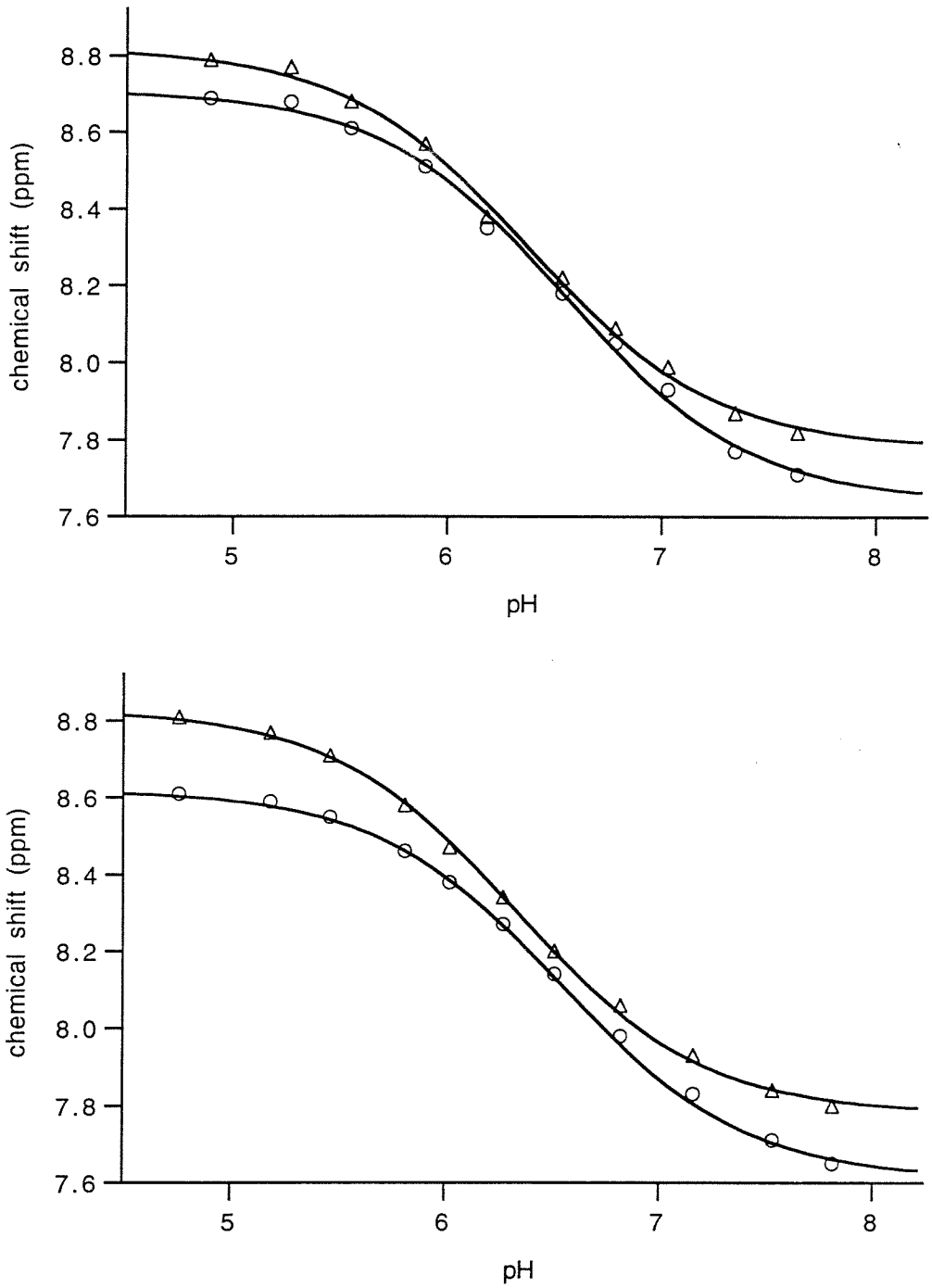


Figure A.4 pH titration behavior of $H_{26}H_{33}H_{39}$ (top) and $H_{26}H_{33}H_{58}$ (bottom) iso-1-cytochrome c variants. The chemical shifts in ppm of the C2 protons of histidines at position 33 (Δ) and 39 or 58 (O) are shown. The data for histidine 33 is fit with a standard pH titration curve and yields pKa's of 6.37 ($H_{26}H_{33}H_{39}$) and 6.33 ($H_{26}H_{33}H_{58}$). The data for histidines 39 and 58 are also fit with a standard pH titration curve and yields pK_a's of 6.54 ($H_{26}H_{33}H_{39}$) and 6.55 ($H_{26}H_{33}H_{58}$).

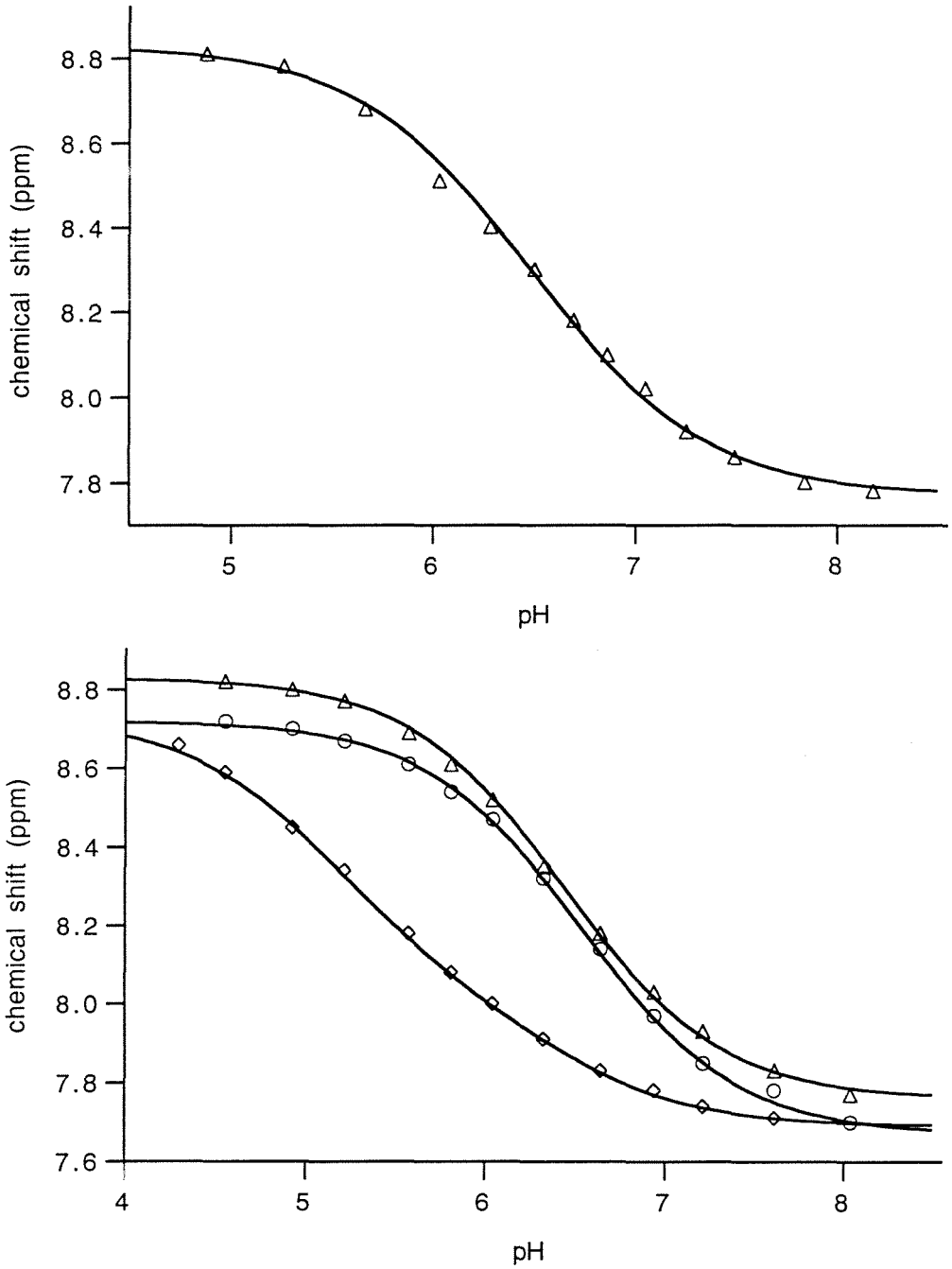


Figure A.5 pH titration behavior of $H_{26}H_{33}$ (top) and $H_{26}H_{33}H_{39}H_8$ (bottom) iso-1-cytochrome c variants. The chemical shifts in ppm of the C2 protons of histidines at position 33 (Δ), 39 (O), and 8 (◊) are shown. The data for histidine 33 is fit with a standard pH titration curve and yields pKa's of 6.48 ($H_{26}H_{33}$) and 6.44 ($H_{26}H_{33}H_{39}H_8$). The data for histidine 39 is also fit with a standard pH titration curves which yielda a pK_a of 6.53 ($H_{26}H_{33}H_{39}H_8$). The data for histidine 8 is fit with a non-standard pH titration curve which yield an apparent pK_a of 5.48 ($H_{26}H_{33}H_{39}H_8$).

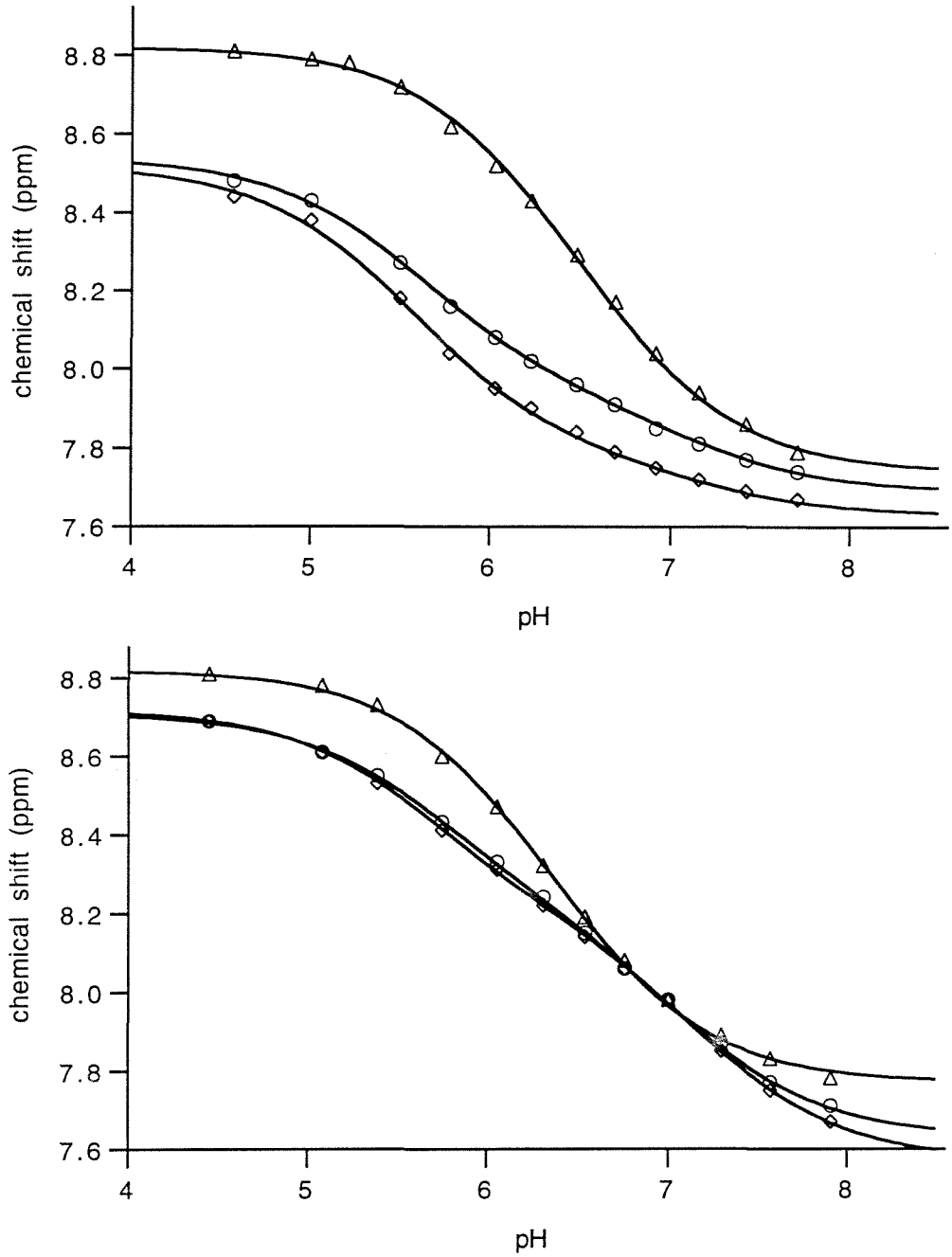


Figure A.6 pH titration behavior of metal-chelating iso-1-cytochrome c variants, $H_{26}H_{33}H_4H_8$ (top), $H_{26}H_{33}H_{39}H_{58}$ (bottom). The chemical shifts in ppm of the C2 protons of histidines at position 33 (Δ), 39 or 8 (O), and 59 or 4 (◊) are shown. The data for histidine 33 is fit with a standard pH titration curve and yields pKa's of 6.48 ($H_{26}H_{33}H_4H_8$) and 6.36 ($H_{26}H_{33}H_{39}H_{58}$). The data for histidines 4, 8, 39, and 58 are fit with non-standard pH titration curves which yield apparent pK_a's of 5.91 ($H_{26}H_{33}H_4H_8$), 5.73 ($H_{26}H_{33}H_4H_8$), 6.52 ($H_{26}H_{33}H_{39}H_{58}$), and 6.51 ($H_{26}H_{33}H_{39}H_{58}$), respectively.

Appendix B

Equilibrium Binding Isotherm Data

Table B-1 Equilibrium binding isotherm data for selected iso-1-cytochrome c variants, tuna cytochrome c and imidazoe. C_L is the solute concentrationin the liquid phase in mM and Q_p is the amount of solute bound at equilibrium (μmol solute/mL of packed gel). Data measured as described in Chapter 3.

H(-)		H_{26}		Tuna	
C_L	Q_p	C_L	Q_p	C_L	Q_p
0.21	0.044	0.22	0.49	0.21	0.42
0.11	0.042	0.11	0.30	0.15	0.33
0.062	0.037	0.051	0.19	0.10	0.25
0.032	0.029	0.028	0.10	0.071	0.17
0.014	0.016	0.014	0.044	0.051	0.14
		0.010	0.028	0.036	0.085
				0.026	0.076
				0.018	0.056
				0.009	0.030
				0.007	0.024

H_4		H_8		Imidazole	
C_L	Q_p	C_L	Q_p	C_L	Q_p
0.20	1.06	0.135	0.90	15.1	25.0
0.082	0.96	0.054	0.76	15.1	23.5
0.071	0.91	0.037	0.66	7.1	21.5
0.031	0.71	0.022	0.50	7.0	21.5
0.013	0.48	0.008	0.30	3.1	18.6
0.012	0.46	0.003	0.17	3.0	19.5
0.005	0.26			0.88	15.0
0.002	0.14			0.82	16.1
				0.26	11.1
				0.26	10.8
				0.10	6.5
				0.08	6.6
				0.05	2.9
				0.03	2.8

Table B.2 Equilibrium binding isotherm data for multiple-histidine iso-1-cytochrome c variants. C_L is the solute concentrationin the liquid phase in mM and Q_p is the amount of solute bound at equilibrium (μmol solute/mL of packed gel). Data measured as described in Chapter 3.

$H_{26}H_4$		$H_{26}H_8$		$H_{26}H_{33}$	
C_L	Q_p	C_L	Q_p	C_L	Q_p
0.110	1.29	0.107	1.14	0.147	1.22
0.053	1.22	0.033	0.95	0.053	1.00
0.011	1.08	0.033	0.98	0.016	0.73
0.010	1.02	0.0070	0.74	0.004	0.45
0.0013	0.60	0.0069	0.76	0.0010	0.25
0.0004	0.32	0.0012	0.44	0.0004	0.13
0.0004	0.28	0.0011	0.41		
		0.0004	0.21		

$H_{26}H_{33}H_4$		$H_{26}H_{33}H_8$		$H_{26}H_{33}H_{39}$	
C_L	Q_p	C_L	Q_p	C_L	Q_p
0.169	2.33	0.29	2.13	0.33	2.15
0.069	2.11	0.177	1.94	0.18	2.06
0.019	1.89	0.119	1.97	0.035	1.76
0.019	1.82	0.079	1.78	0.026	1.67
0.0018	1.29	0.020	1.50	0.0012	1.10
0.0002	0.78	0.016	1.55	0.0004	0.81
0.00017	0.61	0.0013	1.08		
		0.0007	0.93		
		0.0002	0.60		

Appendix C

Protein Stabilization Data

Table C.1 Parameters from guanidinium unfolding studies on $H_{26}H_{33}H_4H_8$ iso-1-cytochrome c. The slope and intercept refer to the best fit linear parameters to a plot of ΔGunf versus [Gdm.Cl]. $[Gdm.Cl]_o$ is the concentration of Gdm.Cl where ΔG_{unf} equals zero.

	[Cu(II)IDA] (mM)				
	0.0	**0.1**	**0.2**	**0.4**	**0.8**
Slope	-4.43	-3.86	-4.18	-4.17	-3.95
Intercept	4.45	4.36	5.03	5.24	5.35
$[Gdm.Cl]_o$	1.00	1.13	1.20	1.26	1.35
$\Delta\Delta G_{unf}$	0.00	0.55	0.84	1.09	1.47

Table C.2 Parameters from guanidinium unfolding studies on $H_{26}H_{33}H_{39}H_{58}$ iso-1-cytochrome c. The slope and intercept refer to the best fit linear parameters to a plot of ΔGunf versus [Gdm.Cl]. $[Gdm.Cl]_o$ is the concentration of Gdm.Cl where ΔG_{unf} equals zero.

	[Cu(II)IDA] (mM)				
	0.0	**0.025**	**0.05**	**0.2**	**0.8**
Slope	-4.24	-4.16	-4.33	-4.12	-3.80
Intercept	4.10	5.03	5.53	5.87	5.99
$[Gdm.Cl]_o$	0.97	1.21	1.28	1.42	1.58
$\Delta\Delta G_{unf}$	0.00	1.01	1.30	1.91	2.56

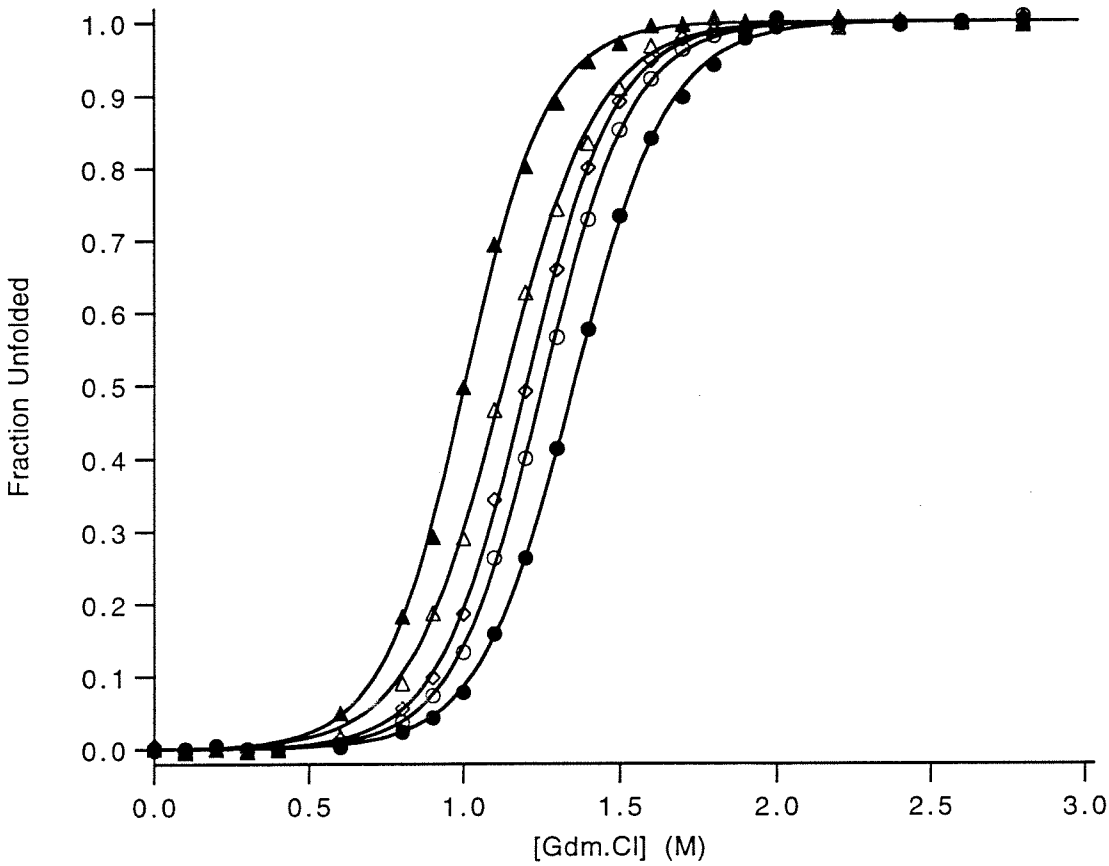


Figure C.1 Stabilization of $H_{26}H_{33}H_4H_8$ iso-1-cytochrome c variant by Cu(II)IDA at 25° C, pH 7.6. The fraction unfolded is given as a function of guanidinium chloride (Gdm.Cl) concentration in the presence of increasing concentrations of Cu(II)IDA. (▲) 0.0 mM, (Δ) 0.1 mM, (◊) 0.2 mM, (O) 0.4 mM, (●) 0.8 mM.

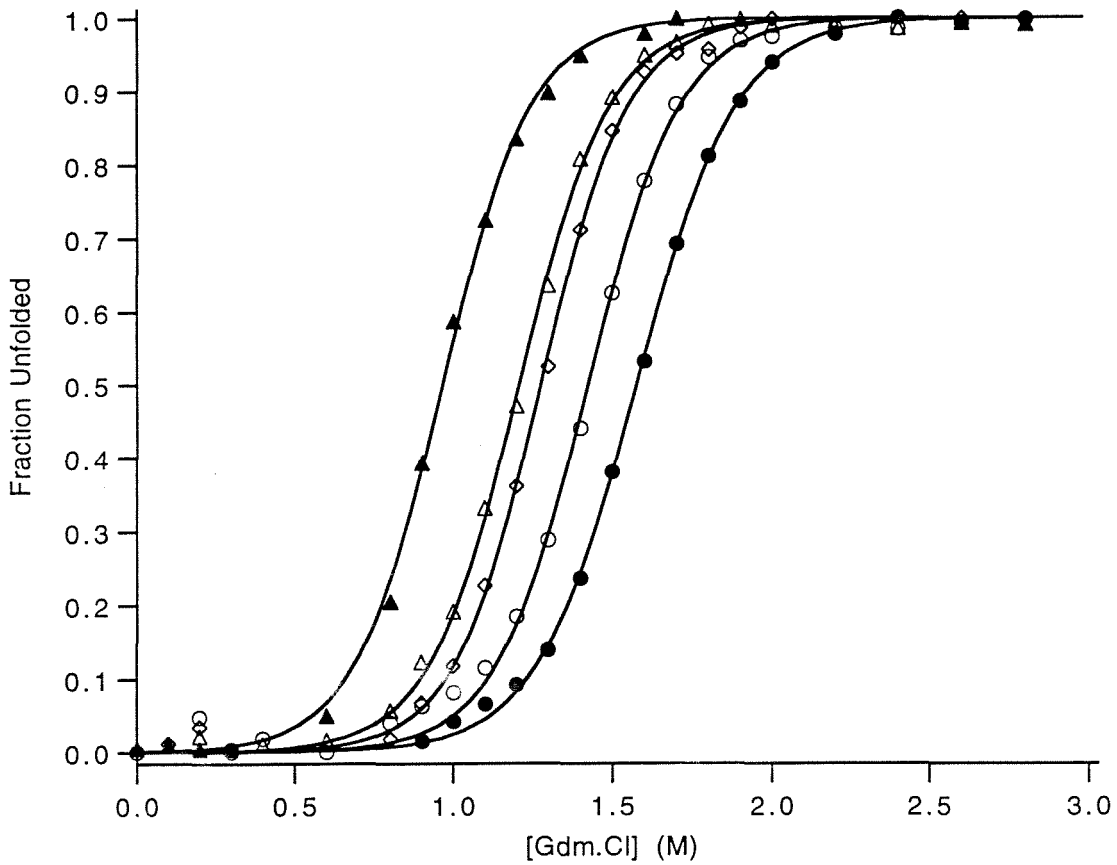


Figure C.2 Stabilization of $H_{26}H_{33}H_{39}H_{58}$ iso-1-cytochrome c variant by Cu(II)IDA at 25° C, pH 7.6. The fraction unfolded is given as a function of guanidinium chloride (Gdm.Cl) concentration in the presence of increasing concentrations of Cu(II)IDA. (▲) 0.0 mM, (Δ) 0.025 mM, (◊) 0.05 mM, (O) 0.2 mM, (●) 0.8 mM.