Surface topography of histidine residues: A facile probe by immobilized metal ion affinity chromatography
(transition metal/coordinations/Imidazole/protein surface)

ESMAT S. HEMDAN*†, YUN-JUAN ZHAO*†, EUGENE SULKOWSKI*§, AND JERKER POR ATH*∥

*Institute of Biochemistry, University of Uppsala, Biomedical Center, Box 576, S-751 23 Uppsala, Sweden; †Biochemistry Department, Ladies Section, King Abdullah University, Faculty of Science, Jeddah, Saudi Arabia; §Institute of Microbiology, Academia Sinica, Beijing 100080, China; ∥Department of Molecular and Cellular Biology, Roswell Park Memorial Institute, Buffalo, NY 14263; and ǂBiochemical Separation Centre, University of Uppsala, Box 577, S-751 23 Uppsala, Sweden

Communicated by Christian B. Anfinsen, December 9, 1988

ABSTRACT Immobilized metal ion affinity chromatography (IMAC) has been explored as a probe into the topography of histidyl residues of a protein molecule. An evaluation of the chromatographic behavior of selected model proteins—thioredoxin, ubiquitin, calmodulin, lysozyme, cytochrome c, and myoglobin on immobilized transition metal ions (Co²⁺, Ni²⁺, Cu²⁺, and Zn²⁺)—allows establishment of the following facets of the histidyl side chain distribution: (i) either interior or surface; (ii) when localized on the surface, accessible or unaccessible for coordination; (iii) single or multiple; (iv) when multiple, either distinct or vicinal. Moreover, proteins displaying single histidyl side chains on their surfaces may, in some instances, be resolved by IMAC; apparently, the microenvironments of histidyl residues are sufficiently diverse to result in different affinities for the immobilized metal ions. IMAC, previously introduced as an approach to the fractionation of proteins, has become also, upon closer examination, a facile probe into the topography of histidyl residues. This is possible because of the inherent versatility of IMAC; an appropriate metal ion (M²⁺) can be selected to suit the analytical purpose and a particular chromatographic protocol can be applied (isocratic pH, falling pH, and imidazole elution).

The underlying principle of immobilized metal ion affinity chromatography (IMAC) of proteins is the coordination between the electron donor groupings on a protein surface (histidine, tryptophan, cysteine) and chelated (imidodiacetate; IDA) transition metal ions [IDA-M(II)]. This principle of immobilized metal ion affinity (IMA) has been presented by now in some detail (1–4).

The practice of IMAC in the purification of proteins has had its empirical phase. There is now a need, from the body of data, to establish somewhat more detailed ground rules that would allow for the use of IMAC in a more predictive manner.

Our present experience (4) clearly indicates that histidine functions as the predominant ligand in the IMAC of proteins based on metal ions belonging to the latter part of the first series of transition metal elements (and zinc). Therefore, topography of the protein surface with respect to the location of histidine residues becomes critical for better understanding and exploitation of an IMA event (1, 4–6).

A facile probe into the topography of histidyl residues could be quite useful, especially if it could provide information about their surface disposition. This information is required for the rational exploitation of IMAC for protein isolation. In a simple case, one has only to seek information about the surface accessibility of a putative electron donor grouping (7). However, the evaluation of an IMA event in the case of proteins replete in histidine residues may require more comprehensive analysis (8).

We now report that IMAC may be exploited as an analytical tool in addition to its use as a protein purification technique (9, 10). IMAC can be used to ascertain several facets of the status of a histidyl residue(s) in a protein molecule: (i) localization (interior vs. surface), (ii) coordination potential as defined by the steric accessibility and the state of protonation, (iii) single vs. multiple, and (iv) surface density.

MATERIALS AND METHODS

Materials. Chelating Sepharose 6B and calmodulin (bovine testes) were purchased from Pharmacia. Thioredoxin, Escherichia coli (recombinant), was obtained from Calbiochem. Myoglobin, mollusc (Aplysia limacina), was a gift from E. Chiancone (University of Rome). All other proteins were purchased from Sigma: cytochromes c [horse heart (type VI, C7732), yeast, Candida krusei (type VII, C4381), tuna heart (type XI, C2111)], myoglobins [sperm whale skeletal muscle (type II, M0380), dog skeletal muscle (type V, M7382)], lysozyme [chicken egg white (grade 1 L6876)], ubiquitin [bovine erythrocytes (U6253)].

Preparation of IDA-M(II) Columns. Chelating Sepharose 6B was thoroughly washed with water, degassed, and equilibrated with 0.1 M sodium acetate/1 M NaCl, pH 4.0. The gel was poured into a chromatographic column (0.9 × 8 cm) (bed volume, 5 ml) and charged with 50 ml of metal salt (5 mg/ml) dissolved in 0.1 M sodium acetate/1 M NaCl, pH 4.0. Unbound metal ion (M²⁺, Ni²⁺, Cu²⁺, or Zn²⁺) was washed out of the column with 25 ml of metal-free buffer (0.1 M sodium acetate/1 M NaCl, pH 4.0). The charging of the column with Co²⁺ and the subsequent washing were done at pH 6.0. Finally, each IDA-M(II) column was equilibrated with 100 ml of 0.02 M sodium phosphate/1 M NaCl, pH 7.0. The IDA-M(II) columns, prepared as described above, were used for the adsorption of protein samples at pH 7.0 and their subsequent isocratic elution at pH 7.0 or by a pH gradient, either stepwise or continuous. When the IDA-M(II) column was to be used in an imidazole elution protocol, it was charged with 50 ml of 10 mM imidazole/0.02 M sodium phosphate/1 M NaCl, pH 7.0, and finally equilibrated with 1 mM imidazole/0.02 M sodium phosphate/1 M NaCl, pH 7.0.

Chromatographic Procedure. All IDA-M(II) columns were appropriately equilibrated and all protein samples (5–20 mg) were applied in the equilibrating solvents. The columns were developed at room temperature. The flow rates were main-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: IMAC, immobilized metal ion affinity chromatography; IDA, imidodiacetate; IDA-M(II), chelate of transition metal ion (M²⁺).

∥To whom reprint requests should be addressed.
tained at ≈1 ml per 5 min by means of peristaltic pump. Fractions of 1 ml were collected and, when appropriate, the pH was measured. The protein content of the fractions was measured spectrophotometrically at 280 nm. All other details were given in the figure legends.

RESULTS

A study of the contribution of histidine side chains to an IMA event could be obscured by a contribution of other putative ligands—tryptophan and cysteine (1). Thus, an appropriate selection of the protein models that will permit a study of the contribution of a histidine residue(s) is in order. This can be accomplished only to a limited degree and is ultimately determined by the availability of the histidine groups. Our set of "model proteins" has been selected on the basis of their suitable chemistry (composition, sequence) and stereochemistry (tertiary structure): thioredoxin (11), ubiquitin (12), calmodulin (13), lysozyme (14), cytochrome c (15), and myoglobin (16).

A Single Histidine Residue May Give Rise to the Affinity of a Protein for an IDA-Cu(II). Fig. 1 illustrates the chromatography of a set of model proteins on an IDA-Cu(II) column: thioredoxin (E. coli), cytochrome c (horse), calmodulin (bovine), ubiquitin (bovine), and lysozyme (chicken). All these proteins, except for cytochrome c, contain single histidine residues in their compositions and those histidines are accessible to the solvent. Cytochrome c (horse) contains three histidines in its composition; however, only one histidine is available for coordination. Ubiquitin and calmodulin lack tryptophan and cysteine in their molecules; no protein of this set has a free cysteinyl residue. All proteins, when applied to IDA-Cu(II) columns at pH 7, were retained; their elution was accomplished by a sequential stepwise decrease in the pH of the eluant. Thioredoxin and cytochrome c were eluted upon a stepwise decrease in the pH of the eluant, down to pH 6; calmodulin, ubiquitin, and lysozyme were eluted at a still lower pH (pH 5). Clearly, this elution sequence does not correlate with the pI values of the proteins: thioredoxin, pI 4.5; cytochrome c, pI 10.3; calmodulin, pI 4.1; ubiquitin, pI 6.7; lysozyme, pI 11.3. This chromatographic outcome is anticipated since the electrostatic interaction of a protein with an IDA-M(II) is effectively quenched by the sodium chloride (1 M) present in the eluants (2, 6). There seems to be, however, an inverse correlation between the pK values of the histidyl residues and the avidity of the protein retention: cytochrome c (His-33, pK 6.5), calmodulin (His-107, pK 6.1), and lysozyme (His-15, pK 5.2). The latter correlation is also anticipated since the histidyl side chains coordinate to an IDA-M(II) in their unprotonated forms (1, 6).

A. limacina myoglobin does not contain cysteine in its composition, but there are two tryptophans present. A single histidine (His-95) corresponds to the heme-binding proximal histidine and is not accessible (17, 18). When A. limacina myoglobin was applied on IDA-Cu(II) column at pH 7, it did not bind (data not shown). Thus, the lack of a single histidine residue on the surface of this protein correlates with the lack of an IMA event on an IDA-Cu(II) column. This same result was also obtained with tuna cytochrome c and Peking duck lysozyme, both of which lack a single histidine residue available for coordination (see below).

Correlation Between the Multiplicity of Histidine Residues and the Affinity of a Protein for IDA-Cu(II). Fig. 2 illustrates the chromatography of cytochrome c (tuna, horse, and yeast). The cytochrome c of tuna is devoid of an external histidine ("zero"-histidine protein), the cytochrome c of horse has one external histidine available for coordination ("one"-histidine protein), and the cytochrome c of yeast (C. kruzei) has two histidines available ("two"-histidine protein). Cytochrome c of tuna was found in the breakthrough fractions from an IDA-Cu(II) column when applied at pH 7, whereas cytochromes c of horse and yeast were retained. Both bound cytochromes c were subsequently eluted from IDA-Cu(II) columns in a falling pH gradient: first horse cytochrome c and then yeast cytochrome c.

The affinity of a cytochrome c molecule for the IDA-Cu(II) column seems to be determined by the presence and the multiplicity of a histidyl side chain. The contribution of other amino acid side chains resident on the molecular surface of a cytochrome c molecule to the chromatographic event on an IDA-Cu(II) column, if any, seems to be marginal.

Fig. 3 illustrates the chromatography of two cytochrome c variants: horse (one-histidine protein) and yeast (two-histidine protein) on IDA-Cu(II) in the presence of imidazole in the mobile phase. Both cytochromes c were applied in 1 mM imidazole on IDA-Cu(II) columns, which were appropriately saturated and equilibrated with imidazole. Horse

![Figure 1](image1.png)  
**Fig. 1.** Affinity of model proteins (single histidine) for IDA-Cu(II). A sample of proteins (5 ml) was applied and the columns were rinsed with 20 ml of the equilibrating buffer (pH 7.0). The column was then developed with a pH gradient (stepwise): 1. 0.1 M sodium acetate/1 M NaCl, pH 6.0; 2. 0.1 M sodium acetate/1 M NaCl, pH 6.0; 3. Thioredoxin (E. coli); 4. cytochrome c (horse); 5. calmodulin (bovine); 6. ubiquitin (bovine); 7. lysozyme (chicken).

![Figure 2](image2.png)  
**Fig. 2.** Resolution of cytochromes c on IDA-Cu(II) column. A sample of protein (5 ml) was applied and the column was rinsed as follows: 1. with 20 ml of the equilibrating buffer (pH 7.0); 2. with 25 ml of 0.1 M sodium acetate/1 M NaCl, pH 6.0; 3. the column was then developed with a pH gradient formed by mixing 25 ml of 0.1 M sodium acetate/1 M NaCl, pH 4.0, into 25 ml of 0.1 M sodium acetate/1 M NaCl, pH 6.0; 4. finally, the column was rinsed with 0.1 M sodium acetate/1 M NaCl, pH 4.0. Tuna heart; horse heart; C. kruzei.
cytochrome c was found in the breakthrough fractions, whereas yeast cytochrome c was retained; the latter protein was subsequently recovered from the column in a linear concentration gradient of imidazole at pH 7 to pH 7 mM imidazole.

The chromatographic resolution of these two proteins, cytochrome c of horse heart and yeast, was thus accomplished in two experimental protocols, one involving the protonation of the imidazole side chain of the protein molecule (Fig. 2) and the other involving coordination of free imidazole (mobile phase) to IDA-Cu(II), thus resulting in affinity elution of the bound protein. In both cases, the cytochrome c endowed with two histidyl residues available for coordination displayed a higher affinity for an IDA-Cu(II) column than did cytochrome c with a single histidyl. As a result, the two cytochromes c were chromatographically resolved on IDA-Cu(II) columns.

Presumably, other evolutionary variants of cytochrome c and of thioredoxin, calmodulin, and lysozyme can be readily segregated according to the multiplicity of histidine residues. The separation of other one- and two-histidine proteins, which are unrelated and grossly different with respect to size, charge, glycosylation, etc., is, in principle, feasible but remains to be demonstrated.

Affinity of a Protein for an IDA-Ni(II) Column. Fig. 4 illustrates the chromatography, under isocratic elution conditions at pH 7, of two cytochromes c: horse (one histidine) and yeast (two histidines). Horse heart cytochrome c was not retained on an IDA-Ni(II) column, whereas C. krusei cytochrome c was and was then recovered with 10 total column volumes of eluant. Clearly, the presence of only one histidine residue does not give rise to a sufficiently high affinity of a cytochrome c for an IDA-Ni(II) and ensuing retention on the column. By contrast, the presence of two histidine residues results in the retention of a cytochrome c on an IDA-Ni(II) column.

The affinity of dog myoglobin ("four- to five-histidine") protein for an IDA-Ni(II) column is considerably higher than that of C. krusei cytochrome c: when chromatographed isocratically at pH 7, the dog myoglobin emerges from the column in a broad elution profile with a peak tube at 300 ml (60 total volumes of the column) of the eluate (data not shown). However, in the presence of imidazole (20 mM) in the mobile phase, dog myoglobin can be readily eluted from an IDA-Ni(II) column (data not shown).

Affinity of a Protein for an IDA-Co(II) and an IDA-Zn(II) Column. Fig. 5 illustrates the chromatography of horse cytochrome c and yeast cytochrome c on IDA-Co(II) and IDA-Zn(II) columns. Clearly, the yeast cytochrome c (two-histidine protein) displays a somewhat stronger affinity at pH 7 for both sorbents than does horse cytochrome c (one-histidine protein). However, the availability of only two histidines (yeast cytochrome c) gives rise to a barely discernible retention on these IDA-M(II) columns.

Fig. 5 illustrates the chromatography of dog myoglobin and sperm whale myoglobin on IDA-Co(II) and IDA-Zn(II) columns. Dog myoglobin was only marginally retained when applied on IDA-Co(II) and IDA-Zn(II) columns at pH 7 despite the presence of four or five histidyls available for coordination; its binding was not much stronger than that of C. krusei cytochrome c. However, in sharp contrast, the sperm whale myoglobin emerged from both IDA-Co(II) and IDA-Zn(II) columns only after a prolonged elution. Yet, dog myoglobin and sperm whale myoglobin differ, conspicuously, by only two histidines, which are present in the sperm.
whale myoglobin (His-113 and His-116) and absent in the dog myoglobin. Apparently, it is the vicinal disposition (an α-helical localization) of those two histidines on the surface of myoglobin and not merely the multiplicity of histidine residues that has such a dramatic impact on the affinity of myoglobin toward IDA-Co(II) and IDA-Zn(II). The chromatographic behavior of human and horse myoglobins was similar to that of dog and sperm whale myoglobins, respectively (data not shown). Conspicuously, the human myoglobin lacks His-113 and His-116, which, however, are present in horse myoglobin.

**DISCUSSION**

Histidine has been shown to be distributed more or less evenly between the interior and the exterior of proteins analyzed in this respect (19). The IMAC of an unknown protein could, therefore, serve as a prima facie probe for the presence of histidyl side chains on the surface of that protein. To substantiate that claim, one has to experimentally verify the following notions: (i) a protein devoid of a putative electron donor residue (histidine) on its molecular surface will not display affinity for a transition metal ion, IDA-M(II); (ii) a protein endowed with a putative electron donor residue (histidine) on its molecular surface will display an affinity for a transition metal ion, IDA-M(II); (iii) a protein with multiple electron donor residues (histidines) on its surface will have an enhanced affinity for an IDA-M(II) resulting in a prolonged residence time on a chromatographic column under standard experimental conditions.

The first notion has been tested with tuna cytochrome c and mollusc myoglobin (A. limacina). Tuna cytochrome c was found in the breakthrough fractions (Fig. 2) when applied on an IDA-Cu(II) column. A tally of the putative electron donor groupings (I) in tuna cytochrome c is as follows: there is no free cysteine (thiol) present; Trp-59 and Trp-33 are buried (20–22). Therefore, only histidine residues remain to be considered as possible electron donor groupings. There are two histidines present in tuna cytochrome c; His-18 is buried and the low pK 2.5 reflects its microenvironment; His-26, although classified as "S" (surface) residue (20), is not readily accessible to the solvent (pK 3.5) due to crowding by the neighboring hydrophobic residues (15, 20–22). Thus, His-18 and His-26 are ruled out a priori as electron donors for the coordination with IDA-Cu(II). In the absence of a thiol group and since imidazolyl and indolyl groups are not available as well, one can anticipate (I) that tuna cytochrome c will not bind to an IDA-Cu(II) column. The experimental outcome (Fig. 2) is consistent with that anticipation.

A. limacina myoglobin did not bind to the IDA-Cu(II) column. Again, as in the case of tuna cytochrome c, the absence of an accessible histidine residue determined the chromatographic behavior of that protein on the IDA-Cu(II) column.

Peking duck lysozyme, lacking a histidine residue in its composition, does not bind to an IDA-Cu(II) column (Y.-J.Z., E.S., and J.P., unpublished data). The thioredoxin of *Rhodobacter sphaeroides* Y, which lacks histidine in its composition (23), is yet to be tested.

The diversity of the molecular surfaces of the selected variants of cytochrome c, myoglobin, and lysozyme notwithstanding, their chromatographic behavior on IDA-Cu(II) columns seems to be determined by the absence of an accessible histidyl side chain, an observation consistent with the first notion.

The second notion has been tested with thioredoxin (E. coli), horse cytochrome c, ubiquitin, calmodulin, and chicken lysozyme. All these proteins were retained on IDA-Cu(II) columns (Fig. 1).

None of these protein models has free thiol groups. Ubiquitin and calmodulin do not contain cysteine and tryptophan. All these proteins have, however, single histidines, which are available for coordination. Thioredoxin (E. coli) has His-6, which is accessible to the solvent (11). Horse heart cytochrome c has His-33 (pK 6.5), which is classified as an "E" (external) residue (15, 20). Ubiquitin (bovine) has His-68, which is well exposed (24, 25). Calmodulin (bovine) displays His-107 (pK 6.1), which is readily carboxylated (26, 27). Finally, chicken egg white lysozyme has His-15 (pK 5.2), which is only partially shielded from the solvent (14); of six tryptophan residues present, Trp-62 and Trp-123 are the most exposed.

The retention of these proteins on an IDA-Cu(II) column (Fig. 1) occurs with somewhat different avidity, possibly reflecting the differences in the pK values of their histidyl residues; their steric accessibility and an additional contribution by tryptophan side chains to an IMAC event (lysozyme) may also play a role.

The third notion has been tested with cytochromes c and with myoglobins.

Yeast cytochrome c (His-33 and His-39) was better retained on an IDA-Cu(II) column than horse heart cytochrome c, both in decreasing pH (Fig. 2) and increasing imidazole concentration gradients (Fig. 3). While horse cytochrome c was not retained at pH 7 on an IDA-Ni(II) column, yeast cytochrome was eluted only after a prolonged wash (Fig. 4). The behavior of the horse cytochrome c and the yeast cytochrome c on IDA-Co(II) and IDA-Zn(II) columns correlated again with the multiplicity of available histidyl side chains. Although only a very weak interaction was observed (Fig. 5), it was the yeast cytochrome c that was better retained.

Myoglobins are known to have typically six or seven histidyl residues accessible to the solvent; for example, sperm whale myoglobin (28, 29). This number is diminished by two residues in the case of dog myoglobin: His-113 and His-116 are missing (30). The retention of both myoglobins on IDA-Cu(II) columns was extremely strong and was not further studied. The binding of dog myoglobin to an IDA-Ni(II) column was much stronger than that of yeast cytochrome c, in accordance with the multiplicity of their histidine residues available for coordination. Yeast protein was eluted with 55 ml of the eluant (Fig. 4), whereas dog
myoglobin was eluted with 300 ml of the eluant (data not shown).

The binding of yeast cytochrome c to IDA-Co(II) and IDA-Zn(II) is weak (Fig. 5). Even dog myoglobin displayed only a marginal affinity (Fig. 6). In sharp contrast, sperm whale myoglobin displayed considerable avidity in binding to both IDA-Co(II) and IDA-Zn(II) columns (Fig. 6). We interpret this strong binding as being due to two histidines—His-113 and His-116—present in sperm whale myoglobin and absent in dog myoglobin. Those residues are a part of an α-helix in the myoglobin molecule. Being separated by two other amino acid residues, they are at a favorable distance for a strong coordination to all IDA-M(II). Their binding to IDA-Co(II) and IDA-Zn(II) is particularly conspicuous because of the lack of binding of all the other tested proteins.

Essentially the same structural feature—i.e., two histidines resident in an α-helix and separated by three residues—occurs in “zinc fingers” (31).

All in all, it can be said that a protein with an increasing number of histidyl side chains displays an increasing affinity for an IDA-M(II). The increment in the avidity of retention varies with the particular metal ion: IDA-Cu(II), IDA-Ni(II), IDA-Co(II), or IDA-Zn(II). This relationship becomes more complex in the case of IDA-Co(II) and IDA-Zn(II) where it is the density of histidine residues (α-helix) on the protein surface (“vicinal histidines”) that is of critical significance.

General Comments. Our present study indicates that IMAC can be used as a probe of the surface localization, multiplicity, and density of the histidine residues.

Ultimately, whether or not IMAC will experience limitations as a histidine topography probe will depend on the relative abundance and availability for an IMA event of the other residues—tryptophan and cysteine (1).

It should be recalled that tryptophan is not an abundant amino acid in proteins and, when present, is usually sequestered in the interior of a protein molecule (19). Our data with avidin (Y.-j.Z., E.S., and J.P., unpublished observations) also indicate that tryptophan, when present on the surface of a protein (32, 33), is a weak ligand. In a typical case then, the outcome of an IMAC experiment should not be significantly influenced by a contribution of tryptophan.

Free thiol groups are quite frequent on the surface of intracellular proteins. This is where the use of IMAC as a probe for a histidine residue, might fail. It should be said, however, that in the realm of extracellular proteins, typically not endowed with free thios, this should not be a serious problem. Moreover, it is plausible, although not certain at this time, that a free thiol group may scavenge a transition metal ion from an IDA-M(II) and thus become irrelevant to the outcome of an IMAC experiment.

Another limitation, already experimentally recognized (34), is the scavenging of transition metal ions (Co, Ni, Cu, Zn) by a so-called “high-affinity binding site” present on the N terminus of some albumins. The existence of His-3, a constituent of that binding site, could be overlooked by the IMAC, although His-3 is readily accessible to the solvent (34).

There may also be some other limitations not recognized at this time.

Finally, it should be stressed that all the observed correlations between the binding and nonbinding, on the one hand, and the number and topography of the histidines, on the other, depend not only on the coordination properties of a particular transition metal ion but also on other characteristics of the sorbent itself. To wit, one can readily envisage a stronger retention of any protein carrying at least one histidine residue on its surface on any IDA-M(II) if the density of IDA functions is increased. The optimal chromato- graphic conditions will eventually be selected since there is vigorous investigation being carried out in several laboratories (35–37).