# The Crystal Structure of Urease from *Klebsiella aerogenes*

Evelyn Jabri, Mary Beth Carr, Robert P. Hausinger, P. Andrew Karplus*

The crystal structure of urease from *Klebsiella aerogenes* has been determined at 2.2 Å resolution and refined to an *R* factor of 18.2 percent. The enzyme contains four structural domains: three with novel folds playing structural roles, and an $(\alpha\beta)_8$ barrel domain, which contains the bi-nickel center. The two active site nickels are 3.5 Å apart. One nickel ion is coordinated by three ligands (with low occupancy of a fourth ligand) and the second is coordinated by five ligands. A carbamylated lysine provides an oxygen ligand to each nickel, explaining why carbon dioxide is required for the activation of urease apoenzyme. The structure is compatible with a catalytic mechanism whereby urea ligates Ni-1 to complete its tetrahedral coordination and a hydroxide ligand of Ni-2 attacks the carbonyl carbon. A surprisingly high structural similarity between the urease catalytic domain and that of the zinc-dependent adenosine deaminase reveals a remarkable example of active site divergence.

Urease (urea amidohydrolase; E.C. 3.5.1.5), a nickel-dependent metalloenzyme, catalyzes the hydrolysis of urea to form ammonia and carbon dioxide (*1*) with a rate approximately $10^{14}$ times the rate of the uncatalyzed reaction. In 1926, urease was isolated from seeds of the jack bean plant (*Canavalia ensiformis*) as a pure, crystalline enzyme by Sumner (*2*). These crystals, the first obtained for a known enzyme, played a decisive role in proving the proteinaceous nature of enzymes. Approximately 50 years later, jack bean urease was identified as the first nickel metalloenzyme (*3*). Nickel has since been found to be a component of hydrogenases for which a crystal structure was recently reported (*4*), methyl coenzyme M reductases, and carbon monoxide dehydrogenases (*5*).

Nickel-dependent ureases have been isolated from various bacteria, fungi, and higher plants (*1*). Their primary environmental role is to allow the organism to use external and internally generated urea as a nitrogen source (6) and, in plants, urease probably also participates in systemic nitrogen transport pathways and possibly acts as a toxic defense protein (*7*). In agricultural settings, rapid hydrolysis of fertilizer urea by soil bacterial ureases results in unproductive volatilization of nitrogen and in ammonia toxicity or alkaline-induced plant damage. Agricultural trials have shown that urease inhibitors can be combined with fertilizer to increase the overall efficiency of nitrogen utilization (6). Medically, bacterial ureases are important virulence factors. They are implicated in the formation of infection-induced urinary stones (accounting for 15 to 20 percent of all urinary stones), catheter encrustation, pyelonephritis, and hepatic encephalopathy (6). The urealytic *Helicobacter pylori* has also been implicated in peptic ulceration and possibly stomach cancer formation (8). Although some inhibitors of urease are used in treatment, more than half of the patients experience adverse side effects. The commonly used inhibitor, acetohydroxamic acid, depresses bone marrow biosynthesis, inhibits DNA synthesis, and is teratogenic in high doses (9).

The best characterized bacterial urease is that from *Klebsiella aerogenes*. The native enzyme has three subunits, α (60.3 kD, UreC), β (11.7 kD, UreB), and γ (11.1 kD, UreA), reportedly associating with $(\alpha\beta_2\gamma_2)_2$ stoichiometry (*10*). Jack bean urease, the benchmark for comparison, exists as a trimer or hexamer of identical 91-kD subunits (*11*). Despite the apparent variation in quaternary structure, amino acid sequence comparison shows that ureases are homologous, sharing more than 50 percent sequence identity (*11, 12*). The presence of multiple distinct gene products related to portions of the jack bean sequence has also been observed for other bacterial ureases. This clear correspondence between the sequence of the single subunit plant urease and the two or three subunit bacterial ureases indicates the occurrence of a gene fusion or disruption events during the evolution of this enzyme (*1*). Urease does not show significant sequence similarity with other proteins.

The stoichiometry of nickel and its role in the catalytic activity of the *K. aerogenes* and jack bean ureases have been extensively studied (*1*). Stoichiometric analysis of inhibitor binding to urease (*13*) and spectroscopic analysis of thiol inhibited urease (*14, 15*) have shown that both ureases contain a bi-nickel center per active site. This metallocenter is directly involved in binding of substrates and inhibitors (*16*). Recent evidence has been obtained showing that nickel binding to urease requires $CO_2$ incorporation probably involving a protein nucleophile ($pK_a > 9.0$) (*17*).

Despite the availability of crystals for nearly 70 years, a structure of jack bean urease has not been determined. We have reproduced the octahedral crystals of jack bean urease and obtained crystals of *K. aerogenes* urease (*18*). Whereas the jack bean urease crystals diffract only to ~3.0 Å, those of *K. aerogenes* urease diffract beyond 2.0 Å resolution and can provide a detailed view of the structure and active site. We describe here the crystal structure of this nickel metalloenzyme (urease, from the bacterium *K. aerogenes*) at 2.2 Å resolution.

**Structure determination.** Crystals of *K. aerogenes* urease, selected site-directed mutants (*19*), urease apoenzyme (nickel-free) (*20*), and selenomethionine (Se-Met) urease (*21*) were grown under the same conditions (*18*). The phases were determined by multiple isomorphous replacement (MIR) at 3.0 Å resolution with five heavy atom derivatives and inclusion of the anomalous signal (AS) and solvent flattening (SF) (Table 1). Although connectivities and side chain densities were ambiguous in the initial electron density map, the positions of most of the β strands and α helices were clear.

An initial Cα trace (720 atoms) from the MIR-AS-SF minimap was used to build a polyalanine chain (*22*) that served as the starting point for interactive model building (*23*). Difference maps for the urease apoenzyme and three histidine to alanine mutants at the active site ($His^{\alpha 134}$, $His^{\alpha 219}$, and $His^{\alpha 320}$) (*19*) provided unambiguous starting points for the insertion of sequence. Later, the positions of Se-Met difference peaks (*21*) and the heavy atom binding sites served as additional guides for placing sequence. Seven rounds of model building, refinement (*24*), and phase combination at 3.0 Å resolution (*25, 26*) led to a map in which the complete polypeptide chains, except the last five residues of the β subunit, could be traced unambiguously. Refinement of this model and phase extension to 2.0 Å resolution were completed by means of the simulated annealing and positional refinement protocols of X-PLOR (*27*). As refinement progressed, excess density at the Nζ atom of $Lys^{\alpha 217}$ became stronger and took on a well-defined branched shape connecting the lysine to both nickel ions (see below). In view of the recently documented

E. Jabri and P. A. Karplus are in the Section of Biochemistry, Molecular and Cell Biology, Cornell University, Ithaca, NY, 14853, USA. M. B. Carr and R. P. Hausinger are in the Departments of Microbiology and Biochemistry, Michigan State University, East Lansing, MI 48824–1101, USA.

*To whom correspondence should be directed.

requirement of $CO_2$ for nickel binding (*17*), we modeled this density as a carbamate derivative of $Lys^{\alpha 217}$, from here on referred to as $Lys^{\alpha 217*}$. In the course of refinement, changes in the density near the active site led us to suspect that the five crystals merged in the original native data set (Nat1) (Table 1) might not be perfectly isomorphous. Difference Fourier analysis with the unmerged data sets allowed us to group equivalent data sets on the basis of the absence (Nat2) or presence (Nat3) of excess active site density. Two structures were subsequently refined with gradient minimization in X-PLOR. The Nat2 structure, in which a water molecule is bound to Ni-2, was refined at 2.2 Å resolution with an *R* factor of 18.2 percent and an $R_{free}$ of 23.2 percent (Table 1). The Nat3 structure, in which an unidentified ligand bridges the nickel ions, was refined at 2.0 Å resolution with an *R* of 18.5 percent and $R_{free}$ of 22.5 percent. Aside from small differences at the active site, the Nat2 and Nat3 structures are equivalent. The uninterpreted density at the active site in the Nat3 structure is the appropriate size for a urea molecule or a $HCO_3^-$. The protein models had good geometry (*28*), and most heavy atom positions were near Cys, Arg, Gln, and Asp residues. The final electron density map was consistently of high quality (see below) except in the region of residues α316 to α336 where the density was sparse and the refined model had high temperature factors.

**Overall structure**. *Klebsiella aerogenes* urease is a tightly associated trimer of $(\alpha\beta\gamma)$-units in a triangular arrangement (Fig. 1A). This $(\alpha\beta\gamma)_3$ stoichiometry differs from the proposed $(\alpha\beta_2\gamma_2)_2$ structure that was based on the relative intensities of subunit bands in Coomassie blue stained gels coupled with gel filtration chromatographic and native gel electrophoretic analysis (*10*). All three subunits make extensive contacts to build the trimer: each α subunit packs between the two symmetry related α subunits, and contacts two β subunits and two γ subunits to form the sides of the triangle; each β subunit packs between two adjacent α subunits at the corners of the triangle; and each γ subunit interacts with two α subunits and tightly with two other γ subunits at the crystallographic threefold axis.

Because of the close interactions of the subunits, it is not obvious which α, β, and γ chains make up a primary $(\alpha\beta\gamma)$-unit. However, the $(\alpha\beta\gamma)$-unit discussed below was chosen such that the proximity of the various termini were compatible with the known homology of this urease with the two subunit urease from *H. pylori* (*29*) and the one subunit urease from jack bean (*11*). The carboxyl terminus of the γ subunit is 16 Å from the amino terminus of the β subunit, within the distance needed for the insertion of four residues in the *H. pylori* urease but requiring a looping out of 30 residues of the jack bean sequence. The carboxyl terminus of the β subunit is about 35 Å from the terminus of the α subunit, a distance that would readily allow the insertion of the extra 33 residues that exist in jack bean urease. The high conservation of sequence in all ureases combined with extensive interactions of the trimer suggest that all known ureases adopt a similar trimeric structure. Approximately 3300 $Å^2$

**Table 1.** Summary of crystallographic data and results. Intensity data were collected at room temperature with a San Diego Multiwire Systems Detector (hardware and software) on a Rigaku RU-200 rotating anode x-ray generator (*47*). Heavy atom derivatives were prepared by soaking crystals at room temperature in crystal storage buffer (100 mM Hepes, pH 7.5, 2.0 M $Li_2SO_4$) containing the respective heavy atom. Forty heavy atom compounds were screened and five derivatives were obtained. Multiple isomorphous replacement phases, at 3.0 Å including anomalous data (AS), were calculated by means of the program Darefi (*48*). This initial set of phases was mediocre because the heavy atom derivatives shared common sites and the phasing power beyond ~3.8 Å resolution was weak. The overall figure of merit to 3.0 Å resolution was 0.672. The phases were improved with three rounds of SF with zeroing of electron density at the heavy atom sites to remove heavy atom "ghost" peaks (*26*). Model building into a 3.0 Å MIR-AS-SF map was done with the programs O (*22*) and CHAIN (*23*). The phases were improved with rounds of partial model refinement by means of gradient minimization in the program TNT (*24*) (33,436 reflections, with weighting to the MIR-AS-SF phases) followed with phase combination by the program SIGMAA (*25*). Phase improvement was evidenced by improved separation of the density of the nickel ions and improved connectivity and side chain densities in regions not included in the model. Upon completion of the chain tracing, refinement was continued with X-PLOR (*27*). The models for Nat2, Nat3, and apoenzyme were also refined in X-PLOR. The final models for Nat2 and Nat3 include 767 residues and two $Ni^{2+}$ ions. The model for urease apoenzyme contains 767 residues but lacks the two $Ni^{2+}$ ions and the $CO_2$ modification of $Lys^{\alpha 217}$.

| Data set | Resolution | | Reflections | | | $R_{sym}$* |
|---|---|---|---|---|---|---|
| | (Å) | No. of crystals | Total | Unique | Complete (%) | |
| *Data collection statistics* | | | | | | |
| Nat1† | 2.0 | 5 | 663,385 | 58,463 | 93 | 9.5 |
| Nat2 | 2.2 | 2 | 331,399 | 42,583 | 92 | 11.0 |
| Nat3 | 2.0 | 3 | 386,731 | 58,334 | 93 | 8.9 |
| Apoenzyme | 2.8 | | 85,114 | 20,532 | 95 | 12.2 |
| $HOHgC_6H_4CO_2Na$ | 3.3 | | 18,660 | 11,027 | 89 | 7.5 |
| $EuCl_2$ | 3.3 | | 18,953 | 12,210 | 89 | 6.5 |
| $Hg_2(CH_3COO)_2$ | 2.5 | | 67,132 | 28,709 | 94 | 7.6 |
| $C(HgOOCCH_3)_4$ | 2.4 | | 69,997 | 29,672 | 93 | 8.5 |
| $(CH_3)_3Pb(CH_3COO)$ | 2.4 | | 106,954 | 23,486 | 93 | 7.9 |
| Se-Met | 3.0 | | 92,935 | 20,332 | 90 | 14.7 |

| Derivatives | Soaking conditions | | *R* factor versus native (%) | Heavy atom sites | Phasing power vs. resolution‡ | | |
|---|---|---|---|---|---|---|---|
| | Conc. (mM) | Time (days) | | | ∞–6 Å | 6–4 Å | 4–3 Å |
| *Phasing statistics* | | | | | | | |
| $HOHgC_6H_4CO_2Na$ | 1 | 5 | 19.5 | a,b | 1.30 | 1.18 | 0.79 |
| $EuCl_2$ | 10 | 7 | 9.5 | a,b | 1.48 | 1.31 | 0.74 |
| $Hg_2(CH_3COO)_2$ | Sat'd. | 2 | 15.3 | a,b,c,d,e | 1.30 | 1.18 | 0.93 |
| $C(HgOOCCH_3)_4$ | Sat'd. | 1 | 18.3 | a,b,c,f,g,h,i,j,k | 2.00 | 1.58 | 0.97 |
| $(CH_3)_3Pb(CH_3COO)$ | 10 | 8 | 12.5 | b,g,l,m,n | 1.75 | 1.49 | 0.97 |
| FOM§ | – | – | – | – | 0.825 | 0.789 | 0.612 |

| Data set | Resolution (Å) | Reflections (*N*) | Non-hydrogen atoms | Solvent molecules | *R*‖ | $R_{free}$¶ | rms deviations | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | Bonds (Å) | Angles (deg) |
| *Refinement statistics* | | | | | | | | |
| Nat2 | 10–2.2 | 41,809 | 5,964 | 177 | 18.2 | 23.2 | 0.008 | 1.88 |
| Nat3 | 10–2.0 | 55,572 | 6,002 | 215 | 18.5 | 22.5 | 0.008 | 1.98 |
| Apoenzyme | 10–2.8 | 20,184 | 5,944 | 157 | 18.4 | 25.7 | 0.009 | 1.88 |

*$R_{sym} = \Sigma |I - \langle I \rangle| / \Sigma \langle I \rangle$, where *I* is the integrated intensity of a given reflection. †Nat1 = native data set for MIR analysis. ‡Phasing power is the ratio between the root mean square (rms) of the heavy atom scattering amplitude and the lack of closure error. §FOM is the mean figure of merit (cosine of the estimated phase error). ‖$R = \Sigma |F_{obs} - F_{calc}| / \Sigma |F_{obs}|$, where *F* is the structure factor. ¶$R_{free}$ is the cross-validation *R* factor computed for the test set of reflections (5 percent of the total), which are omitted in the refinement process.

(10 percent) of the α, β, and γ surfaces are buried to form the (αβγ)-unit, and 19,100 Å$^2$ (23 percent) of the (αβγ)-unit surface is buried when the trimer is in formation.

The (αβγ)-unit itself forms a T-shaped molecule with dimensions of 75 by 80 by 80 Å (Fig. 1B). The urease (αβγ)-unit consists of four structural domains, two in the α chain and one each in the β and γ chains (Fig. 1, B and C). Three of these domains appear to be unusual types of folds (*30*). They do not contribute any residues to the active site, but rather have structural roles. The α subunit has an $(\alpha\beta)_8$ barrel domain and a primarily β domain (Fig. 1). The $(\alpha\beta)_8$ barrel is rather elliptical with the long axis (about 20 Å) connecting strands 3 and 7. Two extra parallel strands (9 and 10) extend the lower portion of strands 1 to 3, accentuating the flattened appearance of the barrel. The active site is located at the carboxyl termini of the strands, and a helical excursion (H2-H4) between strand 7 and helix 7 forms a "flap" across the active

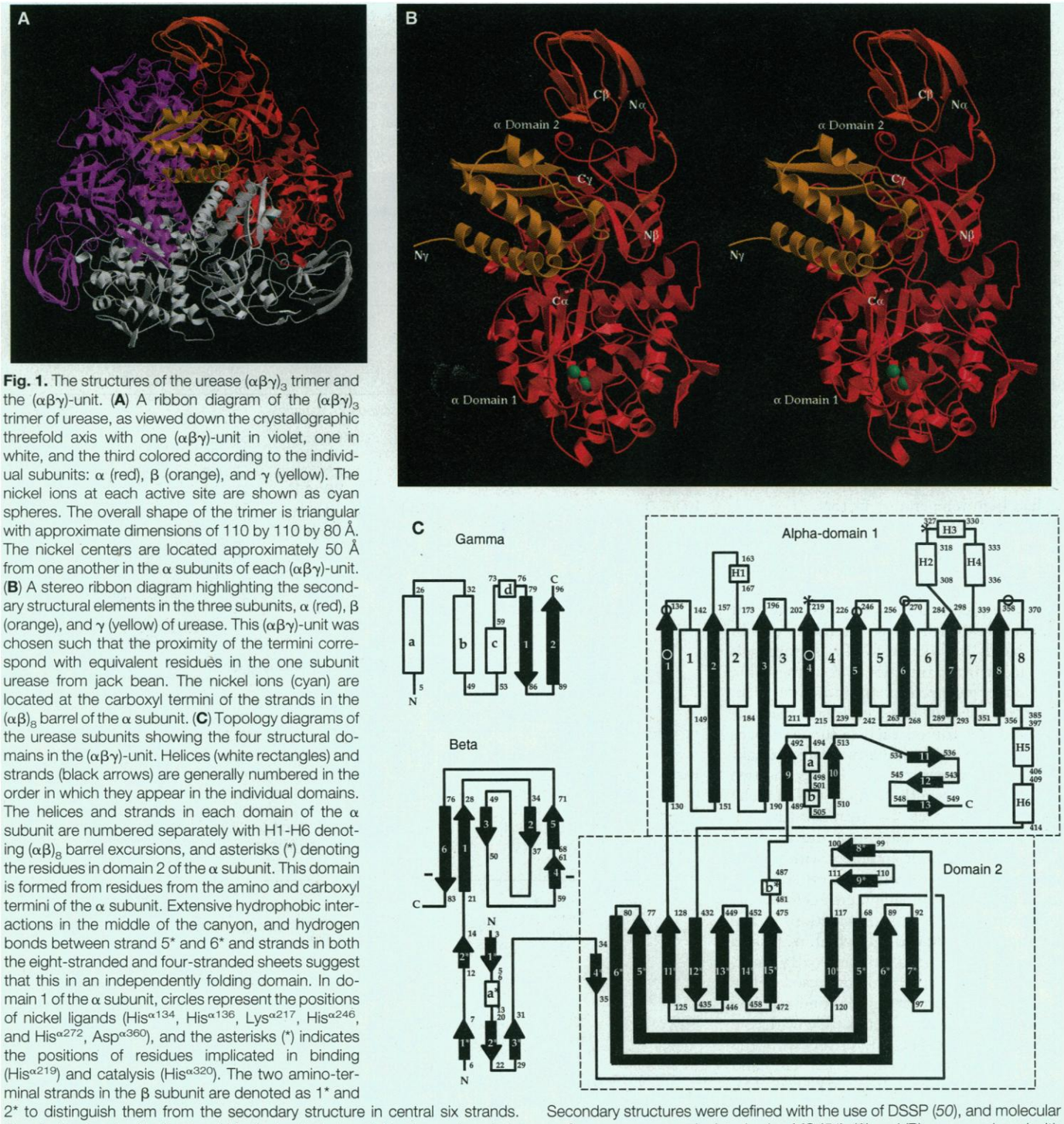


**Fig. 1.** The structures of the urease $(\alpha\beta\gamma)_3$ trimer and the (αβγ)-unit. (**A**) A ribbon diagram of the $(\alpha\beta\gamma)_3$ trimer of urease, as viewed down the crystallographic threefold axis with one (αβγ)-unit in violet, one in white, and the third colored according to the individual subunits: α (red), β (orange), and γ (yellow). The nickel ions at each active site are shown as cyan spheres. The overall shape of the trimer is triangular with approximate dimensions of 110 by 110 by 80 Å. The nickel centers are located approximately 50 Å from one another in the α subunits of each (αβγ)-unit. (**B**) A stereo ribbon diagram highlighting the secondary structural elements in the three subunits, α (red), β (orange), and γ (yellow) of urease. This (αβγ)-unit was chosen such that the proximity of the termini correspond with equivalent residues in the one subunit urease from jack bean. The nickel ions (cyan) are located at the carboxyl termini of the strands in the $(\alpha\beta)_8$ barrel of the α subunit. (**C**) Topology diagrams of the urease subunits showing the four structural domains in the (αβγ)-unit. Helices (white rectangles) and strands (black arrows) are generally numbered in the order in which they appear in the individual domains. The helices and strands in each domain of the α subunit are numbered separately with H1-H6 denoting $(\alpha\beta)_8$ barrel excursions, and asterisks (*) denoting the residues in domain 2 of the α subunit. This domain is formed from residues from the amino and carboxyl termini of the α subunit. Extensive hydrophobic interactions in the middle of the canyon, and hydrogen bonds between strand 5* and 6* and strands in both the eight-stranded and four-stranded sheets suggest that this in an independently folding domain. In domain 1 of the α subunit, circles represent the positions of nickel ligands ($His^{\alpha134}$, $His^{\alpha136}$, $Lys^{\alpha217}$, $His^{\alpha246}$, and $His^{\alpha272}$, $Asp^{\alpha360}$), and the asterisks (*) indicates the positions of residues implicated in binding ($His^{\alpha219}$) and catalysis ($His^{\alpha320}$). The two amino-terminal strands in the β subunit are denoted as 1* and 2* to distinguish them from the secondary structure in central six strands. Also, in the β subunit a hyphen (-) indicates that strand 4 hydrogen bonds to strand 6. The beginning and end residue numbers are given for each element. Secondary structures were defined with the use of DSSP (*50*), and molecular surface areas were calculated using MS (*51*). (A) and (B) were produced with MOLSCRIPT (*52*) and rendered with Raster3D (*53*).

site. Two helices (H5 and H6) and a long loop (residues α515 to α532) cover the amino-terminal end of the barrel. This loop is part of an unusual feature in which the carboxyl terminal 100 residues wrap completely around the $(\alpha\beta)_8$ barrel domain. The second β domain is formed by residues α33 to α129 and α430 to α488. Here, eight-stranded and four-stranded mixed β sheets form the walls of a U-shaped canyon. The walls are connected by the long strands 5* and 6* together with strands 10* and 11*, which go down one side and up the other.

The first 15 residues of the β subunit form two antiparallel β sheets with the amino-terminal residues of the α subunit (Fig. 1). The core of the β subunit adopts an imperfect six-stranded antiparallel β jellyroll. The jellyroll is left-handed and may be the first of this kind observed (*31*). Strand 6 is followed by a long loop that packs against the loops before and after strand 3. The last five residues (β102 to β106) are disordered. This subunit stabilizes the trimer by associating with domain 2 of its own α subunit and domain 1 of a symmetry related α subunit.

The γ subunit adopts a novel αβ domain with four helices and two antiparallel strands (Fig. 1). Two of the helices (b and c) and the two strands pack tightly like the helices of a four-helix bundle and have the commonly seen right-handed up-down-up-down topology for that family (*32*). Helix a is clearly peripheral and helix d is little more than a turn. This subunit facilitates trimer formation through association with the α subunit and the symmetry related γ subunits. The packing of γ subunits at the crystallographic threefold axis is dominated by three copies of helix a, one from each γ subunit, which pack against one another in an almost orthogonal manner.

**The nickel center**. Spectroscopic studies of jack bean and *K. aerogenes* urease have suggested that the two active site nickels are within approximately 3.5 Å, and that each metal ion is approximately pentacoordinate, with about two imidazoles and additional N or O atoms serving as ligands (*14*). Site-directed mutagenesis of *K. aerogenes* urease tentatively identified $His^{\alpha134}$, $His^{\alpha136}$, and $His^{\alpha246}$ as three of the nickel ligands (*19*).

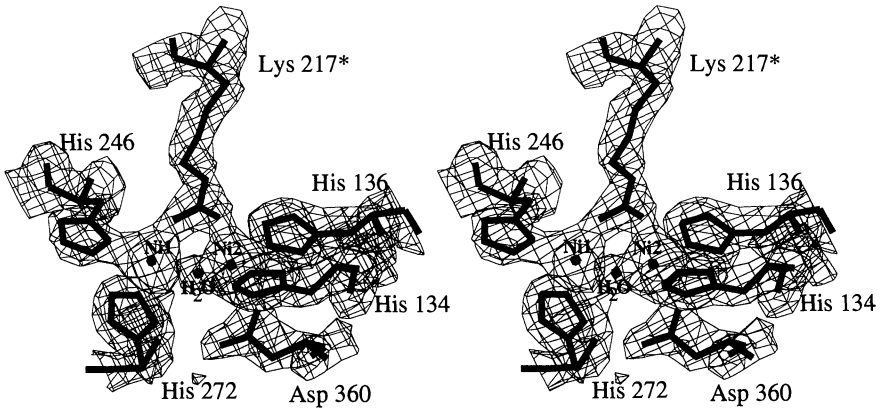


**Fig. 2.** Stereo diagram of the urease bi-nickel center showing the carbamylated lysine residue ($Lys^{\alpha217*}$). The electron density map was calculated with coefficients $2F_{(Nat2)} - F_C$, $\alpha_C$ and is shown contoured at 1.5 times the rms density. This figure was produced with CHAIN (*23*).

**Table 2.** Coordination geometry at the bi-nickel center. Coordinate accuracy is approximately 0.2 Å in this region of the structure, as estimated by a Luzzati plot (*49*).

| Metal-ligand | Distance (Å) | Ligand-metal-ligand | Angles (degrees) |
|---|---|---|---|
| Ni-1···$Lys^{\alpha217*}$Oθ1 | 2.0 | $Lys^{\alpha217*}$Oθ1···Ni-1···$His^{\alpha246}$Nδ | 92 |
| Ni-1···$His^{\alpha246}$Nδ | 2.1 | $Lys^{\alpha217*}$Oθ1···Ni-1···$His^{\alpha272}$Nε | 107 |
| Ni-1-$His^{\alpha272}$Nε | 2.2 | $His^{\alpha246}$Nδ···Ni-1···$His^{\alpha272}$Nε | 95 |
| Ni-2···$His^{\alpha134}$Nε | 2.3 | $His^{\alpha134}$Nε···Ni-2···$His^{\alpha136}$Nε | 110 |
| Ni-2···$His^{\alpha136}$Nε | 2.2 | $His^{\alpha134}$Nε···Ni-2···$Lys^{\alpha217*}$Oθ2 | 88 |
| Ni-2···$Lys^{\alpha217*}$Oθ2 | 2.1 | $His^{\alpha134}$Nε···Ni-2···$Asp^{\alpha360}$Oδ1 | 83 |
| Ni-2···$Asp^{\alpha360}$Oδ1 | 2.1 | $His^{\alpha134}$Nε···Ni-2···Wat-1 | 151 |
| Ni-2···Wat-1 | 2.0 | $His^{\alpha136}$Nε···Ni-2···$Lys^{\alpha217*}$Oθ2 | 88 |
| | | $His^{\alpha136}$Nε···Ni-2···$Asp^{\alpha360}$Oδ1 | 88 |
| Ni-1···Ni-2 | 3.5 | $His^{\alpha136}$Nε···Ni-2···Wat-1 | 98 |
| | | $Asp^{\alpha360}$Oδ1···Ni-2···$Lys^{\alpha217*}$Oθ2 | 168 |
| | | $Asp^{\alpha360}$Oδ1···Ni-2···Wat-1 | 92 |

Our structural results agree reasonably well with these studies (Fig. 2). The nickel ions are well ordered (*B* factors of about 15.0 $Å^2$) and are 3.5 Å apart. Ni-1 is coordinated by three (two N and one O) ligands: $His^{\alpha246}$ through the Nδ atom, $His^{\alpha272}$ (*33*) through Nε, and $Lys^{\alpha217*}$ through Oθ1 of the carbamate (see below). Ni-2 is coordinated by five (two N and three O) ligands: $His^{\alpha134}$ and $His^{\alpha136}$ both through Nε, $Asp^{\alpha360}$ through Oδ1, Wat-1, and $Lys^{\alpha217*}$ through Oθ2 (Table 2). The geometry for Ni-1 can best be described as pseudotetrahedral with a weakly occupied fourth ligand (*34*). The geometry for Ni-2 can be described as distorted trigonal bipyramidal or distorted square pyramidal with $His^{\alpha136}$ as the apical ligand (Fig. 2 and Table 2).

The presence of a carbamate ligand to the nickel ions would have been surprising if it were not for the recent finding that activation of urease apoenzyme can be achieved in vitro only in the presence of $CO_2$ (*17*). The activation results and the structure are consistent with the requirement of $CO_2$ reacting with $Lys^{\alpha217*}$ prior to nickel binding. The structure of the urease apoenzyme (Table 1) shows no carbamate, suggesting that carbamylation of $Lys^{\alpha217}$ occurs readily in aqueous solution, but the binding of nickel ions is required to stabilize the carbamylated form. A similar carbamylated lysine has been observed in another enzyme, namely ribulose-1,5-bisphosphate carboxylase oxygenase (RUBISCO) (*35*). There, carbamylation is required for the binding of $Mg^{2+}$, which in turn coordinates the substrate. In contrast to urease which binds metal tightly (*36*), RUBISCO can be readily deactivated by metal chelators. Although urease can be activated in vitro by the addition of $CO_2$, its activation in vivo appears to be aided by specific proteins (*20, 37*).

The use of a carbamate ligand, rather than an Asp or Glu residue, appears to be necessitated by the position of the active site in the $(\alpha\beta)_8$ barrel in that shorter Asp or Glu residues would not reach the nickel center. Also, the ability of the carbamate to form resonance isomers that allow higher partial negative charges on the oxygens compared to carboxylic acids may be important in the ligation of the nickels or the modulation of their chemistry.

**Active site and catalytic mechanism.** Adjacent to the nickel, near Wat-1, is a pocket roughly the size of urea, which in a static structure is sequestered by the side chain of $Cys^{\alpha319}$ from a narrow channel leading to the enzyme surface (Fig. 3). However, the H2-H4 flap (residues α308 to α336), which forms one wall of the channel, is highly mobile suggesting that this flap may easily open to allow extensive

access to the active site. In fact, the facile modification of $Cys^{\alpha 319}$ at the base of the channel by bulky reagents (*38*) and, in our analysis, by large heavy atom compounds such as $C(HgOOCCH3)_4$ suggests that the channel can readily open.

The putative urea binding pocket is lined by residues from loops 1, 2, 4, 5, 6, 7, and 8 of the $(\alpha\beta)_8$ barrel (Fig. 3). These include $Ala^{\alpha 167}$, $Gly^{\alpha 277}$, $His^{\alpha 219}$, $Ala^{\alpha 363}$, $Met^{\alpha 364}$, $Cys^{\alpha 319}$, and $His^{\alpha 320}$. Three of these residues, $His^{\alpha 219}$, $Cys^{\alpha 319}$, and $His^{\alpha 320}$ have been implicated as important by chemical modification and mutagenesis studies. Site-directed mutagenesis of $His^{\alpha 219}$ to alanine produced an active enzyme with a $K_m$ of 1100 mM compared to 2.3 mM for wild type (*19*). This result suggests that $His^{\alpha 219}$ is involved in substrate binding, and in the structure it is positioned 3.1 Å from Ni-1 and 3.7 Å from Ni-2. Furthermore, $His^{\alpha 219}$ receives a hydrogen bond to Nδ from the main chain nitrogen of $Asp^{\alpha 221}$ and is therefore protonated on Nϵ.

Modification of $Cys^{\alpha 319}$ (or its equivalent residue in jack bean urease) by chemical reagents (*11*, *38*) blocks activity. Dixon and colleagues (*39*) proposed that a cysteine serves as a general acid in jack bean urease catalysis. Later, mutagenesis studies (*36*) showed that $Cys^{\alpha 319}$ is not essential, as an alanine mutation of this residue retained 48 percent of the wild-type activity while mutations to larger residues were more deleterious. The *Staphylococcus xylosus* enzyme does not have a cysteine at this position (*40*), further supporting the notion that $Cys^{\alpha 319}$ is not essential for activity. In the structure, $Cys^{\alpha 319}$ Sγ is 4.4 Å from Wat-1. The effect of larger $Cys^{\alpha 319}$ mutations could easily be caused directly by steric effects or indirectly by changes in the position of the general base, $His^{\alpha 320}$ (see below).

Diethylpyrocarbonate (DEP) modification of the bacterial urease indicated that a histidine with a $pK_a$ of 6.5 is essential for activity (*41*). Site-directed mutagenesis of $His^{\alpha 320}$ produces an enzyme that is not DEP sensitive and although the mutant has normal nickel content, it has very low (~0.003 percent) activity (*19*). These results are consistent with $His^{\alpha 320}$ acting as a catalytic base. In the structure, $His^{\alpha 320}$ Nϵ is 4.8 Å from Ni-1 and 4.3 Å from Wat-1. The $pK_a$ and orientation of $His^{\alpha 320}$ are influenced by hydrogen bonds to the side chains of $Asp^{\alpha 221}$ and $Arg^{\alpha 336}$ (Fig. 3).

From the results of biochemical studies, Dixon and colleagues (*39*) also proposed that a carboxyl group is present at the active site of jack bean urease. In addition, the efficacy of various thiols as inhibitors and the reactivity of thiol-specific modification reagents for inactivating *K. aerogenes* urease are consistent with the presence of a negatively charged group at the active site (*13*, *42*). The structure confirms the presence of three acidic residues, $Glu^{\alpha 220}$, $Asp^{\alpha 221}$, $Asp^{\alpha 360}$ and three main chain carbonyl oxygens from residues $Ala^{\alpha 167}$, $Gly^{\alpha 277}$, and $Ala^{\alpha 363}$. The carbonyl oxygens of $Gly^{\alpha 277}$ and $Ala^{\alpha 363}$ serve to hold a water molecule in place, which hydrogen bonds to both Wat-1 and $His^{\alpha 320}$. In a urea-bound enzyme, these carbonyl groups along with that of $Ala^{\alpha 167}$ and possibly Oδ1 of $Asp^{\alpha 360}$ may be involved in binding the $NH_2$ groups of urea.

The structure of *K. aerogenes* urease can accommodate the mechanistic model proposed by Zerner and colleagues for the jack bean enzyme (*39*). The essence of this proposal is that one nickel ion binds a hydroxide ion, and urea binds to the other nickel, via O-coordination, to polarize the substrate. Based on the structure, Wat-1 would be the hydroxide and the urea oxygen would bind to Ni-1 completing its tetrahedral coordination. Modeling urea in this manner places its oxygen within hydrogen bonding distance of Nϵ of $His^{\alpha 219}$ which may aid in orienting and further polarizing the carbonyl of urea and account for the importance of $His^{\alpha 219}$. The polarized carbonyl of urea is subsequently attacked by a hydroxide, which results in the formation of a tetrahedral intermediate. Zerner's model suggests that a base abstracts a proton from a water molecule to yield a hydroxide. Oδ1 of $Asp^{\alpha 360}$ is in a position to activate Wat-1 for attack, whereas the putative catalytic base, $His^{\alpha 320}$, is positioned too far from the attacking hydroxide. For $His^{\alpha 320}$ to act as the catalytic base, it must move upon binding of urea to a position favorable for proton abstraction from Wat-1. The resulting tetrahedral intermediate is thought to decompose, with the participation of a general acid, to ammonia and a Ni-bound carbamate, which subsequently dissociates. At this time, the identity of the general acid present in *K. aerogenes* urease (*10*) remains unclear.

Although the mechanism outlined above does account for most of the available data, the urease structure is also compatible with alternative models that may involve more structural changes at the active site. For example, urea may displace Wat-1 at Ni-2 and coordinate via the oxygen, or might even bridge both nickels through its carbonyl oxygen. In these cases, there are more possible orientations of bound urea and less can be said about the specific roles for active site residues.

**A Zn-metalloenzyme relative.** The $(\alpha\beta)_8$ barrel observed in the catalytic Ni-binding domain of the α subunit is a rather common topology for proteins, and active debate centers on whether the many $(\alpha\beta)_8$ barrels that do not share significant sequence similarity are a result of convergent or divergent evolution (*43*). Given the lack of recognizable sequence similarity of urease with any other protein, we were surprised

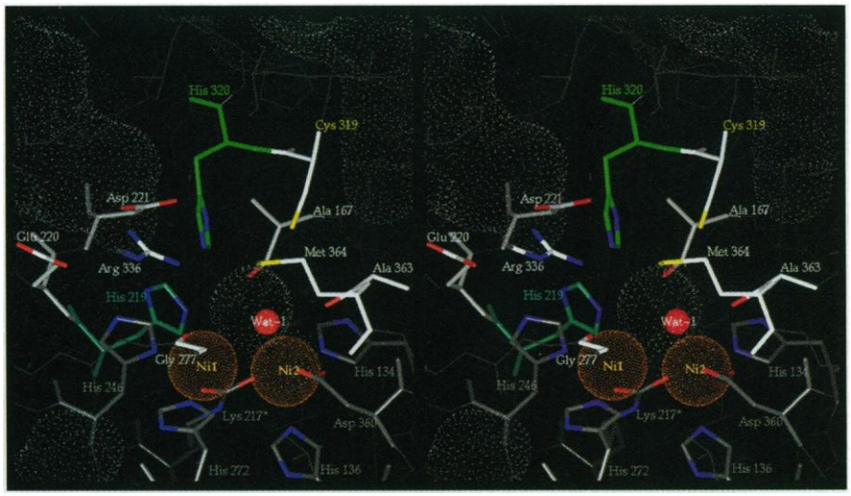


**Fig. 3.** Stereo diagram of the active site of urease. The nickel ions are sequestered in a pocket from a narrow channel by the side chain of $Cys^{\alpha 319}$. The channel is formed from residues in the α subunit ($His^{\alpha 166}$, $Leu^{\alpha 316}$, $Cys^{\alpha 319}$, $Ala^{\alpha 363}$, $Met^{\alpha 364}$) with some contributions from residues in the α subunit of a symmetry related molecule ($Ile^{\alpha 465'}$, $Pro^{\alpha 466'}$, $Thr^{\alpha 467'}$). $Cys^{\alpha 319}$ sits at the base of the channel positioned such that Cβ and Sγ are 3.6 Å from the plane of the ring of the catalytic base, $His^{\alpha 320}$ (green). $Asp^{\alpha 221}$ positions the nickel ligand $His^{\alpha 246}$ through less commonly observed hydrogen bonding involving the anti-orbital of the carboxylate. $His^{\alpha 219}$ (cyan) is positioned to donate a hydrogen bond to the substrate and the main chain oxygens of $Ala^{\alpha 167}$, $Gly^{\alpha 277}$, and $Ala^{\alpha 363}$ are positioned to accept hydrogen bonds from water or urea. Nitrogen atoms are blue, oxygen atoms are red, and carbon atoms are green ($His^{\alpha 320}$), cyan ($His^{\alpha 219}$), gray (nickel ligands) or white (other active site residues). Solvent accessible surface area was calculated with the use of a 1.4 Å radius probe. This figure and Fig. 4A were produced with InsightII version 2.3.5 (BIOSYM).

when a structural comparison with all known protein structures (*30*) revealed that the $(\alpha\beta)_8$ barrel of the α subunit is strikingly similar to that of adenosine deaminase (ADA), an enzyme containing one zinc per active site. This enzyme catalyzes the deamination of adenosine to yield inosine and ammonia (*44*), a reaction mechanistically related to that of urease. Both mechanisms involve the attack of a metal coordinated water on the amide carbon to form a tetrahedral intermediate that subsequently degrades to yield $NH_3$ as one of two products. The nucleophilic water molecule in ADA is stabilized through hydrogen bonds with an aspartic acid and with the catalytic base.

ADA is one of two members of a new class of $(\alpha\beta)_8$ barrels characterized by a unique elliptical barrel axis and a requirement for metal ions (*43*). The second is the Zn-dependent phosphotriesterase for which the apoenzyme crystal structure was recently determined (*45*). A superposition of urease and ADA shows that 72 atoms in the strands of the barrel overlay with 1.8 Å root-mean-square (rms) deviation (Fig. 4A). For comparison, the best overlay with structures of other representative members of the $(\alpha\beta)_8$ barrel family have rms deviations of approximately 3.0 Å (*46*). The helices in urease and ADA all have the same tilt but are somewhat shifted, so that 90 equivalent residues overlay with 2.8 Å rms deviation. The active sites of urease and ADA also show a high degree of structural similarity (Fig. 4B). The position of the Ni-2 and Wat-1 in urease are structurally equivalent to the zinc ion and the hydrolytic water molecule at the active site of ADA. Ni-2 ligands $His^{\alpha134}$, $His^{\alpha136}$, and $Asp^{\alpha360}$ of urease are structurally equivalent to Zn ligands $His^{15}$, $His^{17}$, and $Asp^{295}$ of ADA. $Lys^{\alpha217*}$ in urease is equivalent to $Asp^{181}$, a residue too short to coordinate the Zn ion. Rather, $His^{214}$ of ADA, structurally equivalent to Ni-1 ligand $His^{\alpha246}$ of urease, rotates 90° relative to $His^{\alpha246}$, to take the place of $Lys^{\alpha217*}$ as a metal ligand. $His^{\alpha272}$, a Ni-1 ligand in urease, is equivalent to $His^{238}$, the proposed catalytic base in ADA.

The high degree of global and active site similarity, combined with possible mechanistic similarity, strongly suggests that urease and ADA are homologs. This surprising homology raises some interesting questions: how did urease acquire the requirement for $CO_2$ activation, and why is the enzyme specific for nickel when zinc or other metals could serve similar functions? Furthermore, what mechanisms lie behind the emergence of four accessory proteins, UreD, UreE, UreF, and UreG (*20*, *37*), which are required for in vivo assembly of urease? Thus, in addition to the utility of this structure for inhibitor design, further study of this system should lead not only to insight into the mechanism and regulation of urease, but also insight into the mechanisms of protein evolution.

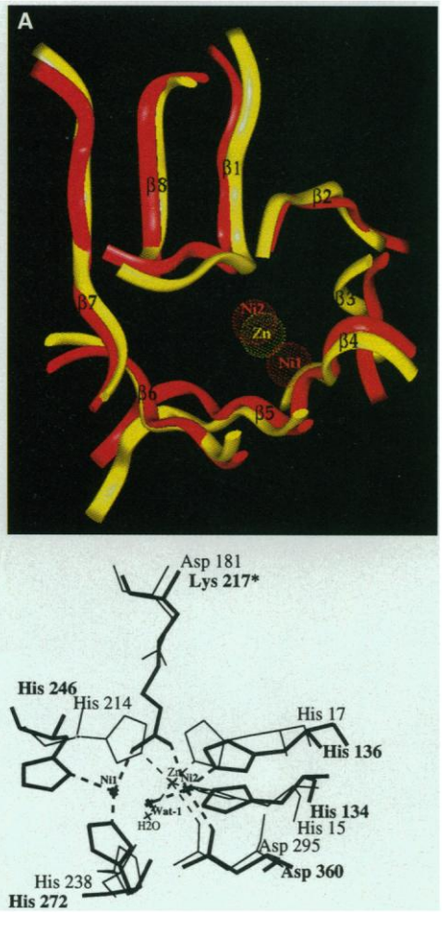


**Fig. 4.** Comparison of urease and adenosine deaminase (ADA). (**A**) An overlay the $(\alpha\beta)_8$ barrels, according to the strategy of Chothia and Lesk (*54*). The view is looking from the amino to carboxyl termini of the strands. (**B**) A stereo view of the urease active site (thick lines and bold lettering) overlaid onto the ADA active site. Structurally equivalent residues shown (urease:ADA) include $His^{\alpha134}$:$His^{15}$, $His^{\alpha136}$:$His^{17}$, $Asp^{\alpha360}$:$Asp^{295}$, $His^{\alpha272}$:$His^{238}$, $His^{\alpha246}$:$His^{214}$, and $Lys^{\alpha217*}$:$Asp^{181}$.

## REFERENCES AND NOTES


1. R. P. Hausinger, *Biochemistry of Nickel* (Plenum Press, New York, 1993), chap. 3.
2. J. B. Sumner, *J. Biol. Chem.* **69**, 435 (1926).
3. N. E. Dixon, C. Gazzola, R. L. Blakeley, B. Zerner, *J. Amer. Chem. Soc.* **97**, 4131 (1975).
4. A. Volbeda *et al.*, *Nature* **373**, 580 (1995).
5. R. P. Hausinger, *Biochemistry of Nickel* (Plenum Press, New York, 1993), chap. 4 to 6.
6. H. L. T. Mobley and R. P. Hausinger, *Microbiol. Rev.* **53**, 85 (1989); H. L. T. Mobley, M. O. Island, R. P. Hausinger, *ibid.*, in press.
7. J. C. Polacco and M. A. Holland, in *International Review of Cytology*, K. W. Leon and J. Jarvik, Eds. (Academic Press, San Diego, 1993), pp. 65.
8. A. Lee, J. Fox, S. Hazell, *Infect. Immun.* **61**, 1601 (1993); J. Alper, *Science* **260**, 159 (1993); M. Marchetti *et al.*, *ibid.* **267**, 1655 (1995).
9. D. P. Griffith, *Urol. Res.* **7**, 215 (1979).
10. M. J. Todd and R. P. Hausinger, *J. Biol. Chem.* **262**, 5963 (1987).
11. S. Wang *et al.*, *Inorg. Chem.* **33**, 1589 (1993).
12. S. B. Mulrooney and R. P. Hausinger, *J. Bacteriol.* **172**, 5837 (1990).
13. M. J. Todd and R. P. Hausinger, *J. Biol. Chem.* **264**, 15835 (1989).
14. K. Takishima, T. Suga, G. Mamiya, *Eur. J. Biochem.* **175**, 151 (1988).
15. M. G. Finnegan *et al.*, *J. Am. Chem. Soc.* **113**, 4030 (1991).
16. N. E. Dixon, R. L. Blakeley, B. Zerner, *Can. J. Biochem.* **58**, 481 (1980); R. L. Blakeley, N. E. Dixon, B. Zerner, *Biochim. Biophys. Acta* **744**, 219 (1983).
17. I.-S. Park and R. P. Hausinger, *Science* **267**, 1156 (1995).
18. E. Jabri, M. H. Lee, R. P. Hausinger, P. A. Karplus, *J. Mol. Biol.* **227**, 934 (1992). Crystals grew in hanging drops equilibrated against 100 mM Hepes (pH, 7.5) and 1.6 M $Li_2SO_4$. All crystals were isomorphous, having the cubic space group $I2_13$ with $a$ = 170.8 Å and one catalytic unit per asymmetric unit.
19. I. S. Park and R. P. Hausinger, *Protein Sci.* **2**, 1034 (1993). The data statistics for the three mutants are similar to those observed for Nat1. Details of the crystal structure are in preparation.
20. M. H. Lee, S. B. Mulrooney, M. J. Renner, Y. Markowicz, R. P. Hausinger, *J. Bacteriol.* **174**, 4324 (1992).
21. Se-Met urease was obtained from *K. aerogenes* carrying pKAU19 (*20*) grown in MOPS minimal medium plus chloramphenicol. When a 1-liter culture growing at 37°C reached $A_{600}$ = 0.6, the following compounds were added: 50 mg of Se-Met, 100 mg of lysine-HCl, 100 mg of threonine, 100 mg of phenylalanine, 50 mg of leucine, 50 mg of isoleucine, and 50 mg of valine. The culture was grown for an additional 22 hours at 30°C reaching $A_{600}$ of 3.4. Purification was the same as for the holoenzyme. The low specific activity observed for the Se-Met enzyme (207 U/mg) may be due to substitution-induced conformational changes in the region of the active site base, $His^{\alpha320}$, placing it in a catalytically unfavorable position. A retrospective analysis shows that 22 of the 23 Se-Met residues have difference peaks ($F_{(Se\text{-}Met)} - F_{(Nat2)}, \alpha_C$) greater than 4.0 times the rms value of the map. The difference peak $Met^{\alpha317}$ is low because of the high mobility of this stretch of residues.
22. T. A. Jones, J. Y. Zou, S. W. Cowan, S. W. Kjeldgaard, *Acta Cryst.* **A47**, 110 (1991).
23. J. S. Sack, *J. Mol. Graphics* **6**, 224 (1988).

24. D. E. Tronrud, L. F. Ten Eyck, B. W. Matthews, *Acta Crystallogr.* **A43**, 489 (1987). The command file for TNT was modified to include an additional *R* factor ($R_{free}$) calculation before initiation of each cycle of refinement. The test data set, generated in X-PLOR, served as the data set for $R_{free}$ calculation while the working data set was used in refinement.
25. R. J. Read, *Acta Crystallogr.* **A46**, 900 (1990). SIGMAA weighted structure factors were used to generate $F_O$, $\alpha_{combined}$ electron density maps for use in fitting.
26. CCP4, "Collaborative Computational Project, Number 4," *ibid.* **D50**, 760 (1994).
27. A. T. Brünger, A. Krukowski, J. W. Erickson, *ibid.* **A46**, 585 (1990).
28. All residues are in the allowed regions of $\phi/\psi$ angles with seven residues (1.1 percent) falling in the generously-allowed regions. Two are in type 2′ turns ($Ala^{\alpha 24}$, $Phe^{\beta 93}$), one is in a type 2 turn ($His^{\alpha 527}$), two are metal ligands ($His^{\alpha 272}$, $Asp^{\alpha 360}$), and one has a high *B* factor ($Met^{\alpha 317}$). $Ala^{\alpha 561}$ has clear and well-defined density. The model contains three cis-prolines: α282, α303, and α470.
29. C. L. Clayton, M. Pallen, H. Kleanthous, B. W. Wren, S. Tabaqchali, *Nucleic Acid Res.* **18**, 362 (1990); A. Labigne, V. Cussac, P. Courcoux, *J. Bacteriol.* **173**, 1920 (1991).
30. L. Holm and C. Sander, *J. Mol. Biol.* **233**, 123 (1993); N. S. Boutonnet, M. J. Rooman, M.-E. Ochagavia, J. S. L. Wodak, in preparation; D. F. Fischer, H. Wolfson, S. L. Lin, R. Nussinov, *Protein Sci.* **3**, 769 (1994). A comprehensive search of the Protein Data Bank was done by the three groups cited. The search by L. Holm with the program Dali showed that domain 1 of the α subunit, an $(\alpha\beta)_8$ barrel, is similar to the $(\alpha\beta)_8$ barrel of adenosine deaminase. Domain 2 of the α subunit, the β subunit, and the γ subunit had no closely similar structure.
31. J. S. Richardson, *Adv. Protein Chem.* **34**, 167 (1981).
32. C. Cohen and D. A. D. Parry, *Proteins* **7**, 1 (1990).
33. $His^{\alpha 272}$ was not studied by site-directed mutagenesis because it was not conserved in the *Ureaplasma urealyticum* sequence. An analysis of the reported *U. urealyticum* sequence (A. Blanchard, personal communication) indicates that a sequencing error caused a reading frame shift so that about a dozen amino acids, including $His^{\alpha 272}$, reported to be distinct compared to other ureases, are in fact conserved.
34. The electron density map shows weak bridging density between Ni-1 and Wat-1. The density is ~1.8 Å from Wat-1 and therefore, cannot be a second water molecule at full occupancy. We estimate the position of Wat-1 is highly occupied as it has a reasonable **B** factor of 7.0 $Å^2$. During the other small portion of the time, Wat-1 may instead occupy a position bridging the nickel ions or solely as Ni-1 ligand.
35. F. C. Hartman and M. R. Harpel, *Annu. Rev. Biochem.* **63**, 197 (1993).
36. P. R. Martin and R. P. Hausinger, *J. Biol. Chem.* **267**, 20024 (1992).
37. M. B. Carr and R. P. Hausinger, in *Mechanisms of Metallocenter Assembly*, R. P. Hausinger, G. L. Eichhorn, L. G. Marzilli, Eds. (VCH, New York), in press.
38. M. J. Todd and R. P. Hausinger, *J. Biol. Chem.* **266**, 24327 (1991).
39. N. E. Dixon, P. W. Riddles, C. Gazzola, R. L. Blakeley, B. Zerner, *Can. J. Biochem.* **58**, 1335 (1980).
40. J. Jose, S. Christians, R. Rosenstein, F. Götz, H. Kaltwasser, *FEMS Microbiol. Lett.* **80**, 277 (1991).
41. I. S. Park and R. P. Hausinger, *J. Protein Chem.* **12**, 51 (1993).
42. M. J. Todd and R. P. Hausinger, *J. Biol. Chem.* **266**, 10260 (1991).
43. G. K. Farber and G. A. Petsko, *Trends Biochem. Sci.* **15**, 228 (1990), G. K. Farber, *Curr. Opin. Struct. Biol.* **3**, 409 (1993); D. Reardon and G. K. Farber, *FASEB J.*, in press.
44. D. K. Wilson, F. B. Rudolph, F. A. Quiocho, *Science* **252**, 1278 (1991).
45. M. M. Benning, J. M. Kuo, F. M. Raushel, H. M. Holden, *Biochemistry* **33**, 15001 (1994).
46. E. Jabri and P. A. Karplus, unpublished results.
47. R. C. Hamlin, *Methods Enzymol.* **114**, 416 (1985); N. H. Xuong, C. Nielson, R. Hamlin, D. Anderson, *J. Appl. Crystallogr.* **18**, 342 (1985).
48. R. E. Dickerson, J. E. Weinzierl, R. A. Palmer, *Acta Crystallogr.* **B24**, 997 (1968). Darefi is a modified version of a program used in the laboratory of G. E. Schulz.
49. P. V. Luzzati, *Acta Crystallogr.* **5**, 802 (1952).
50. W. Kabsch and C. Sander, *Biopolymers* **22**, 2577 (1983).
51. M. L. Connolly, *Science* **221**, 709 (1983).
52. P. Kraulis, *J. Appl. Crystallogr.* **24**, 946 (1991).
53. D. J. Bacon and W. F. Anderson, *J. Mol. Graph.* **6**, 219 (1988); E. A. Merritt and M. E. P. Murphy, *Acta Crystallogr.* **D50**, 869 (1994).
54. L. Chothia and A. Lesk, *EMBO J.* **4**, 823 (1986).
55. We thank I.-S. Park and M. Lee for selected mutant enzymes and apoenzyme, R. Shagita for TAMM and DMA heavy atom compounds, L. Holm, N. Boutonnet, M. Rooman, and D. Fischer for searching the PDB for homologs, S. Ealick for the use of his detector, and B. Carpenter and S. Lippard for helpful discussion of the metal center. Supported by USDA grant 9303870 (P.A.K. and R.P.H.) and by NIH training grant 5T32-GM08384-04 (E.J.). The coordinates and structure factors are deposited in the Protein Data Bank with entry codes 1KAU for Nat2, 2KAU for Nat3, and 3KAU for the apoenzyme.

1 January 1995; accepted 5 April 1995