# The Metal Binding Site of the Hepatitis C Virus NS3 Protease

## A SPECTROSCOPIC INVESTIGATION*

**Andrea Urbani, Renzo Bazzo, Maria Chiara Nardi, Daniel Oscar Cicero, Raffaele De Francesco, Christian Steinkühler‡, and Gaetano Barbato**

*From the Istituto di Ricerche di Biologia Molecolare "P. Angeletti," Via Pontina Km 30.600, 00040 Pomezia, Rome, Italy*

The NS3 region of the hepatitis C virus encodes for a serine protease activity, which is necessary for the processing of the nonstructural region of the viral polyprotein. The minimal domain with proteolytic activity resides in the N terminus, where a structural tetradentate zinc binding site is located. The ligands being been identified by x-ray crystallography as being three cysteines (Cys[97], Cys[99], and Cys[145]) and one histidine residue (His[149]), which is postulated to coordinate the metal through a water molecule. In this article, we present an analysis of the role of metal coordination with respect to enzyme activity and folding. Using NMR spectroscopy, the resonances of His[149] were assigned based on their isotropic shift in a Co(II)-substituted protein. Data obtained with ^{15}N-labeled NS3 protease were compatible with the involvement of the $\delta$-N of His[149] in metal coordination. pH titration experiments showed that the cooperative association of at least two protons is required in the protonation process of His[149]. Changes in the NMR signals of this residue between pH 7 and 5 are interpreted as evidence for a structural change at the metal binding site, which switches from a "closed" to an "open" conformation. Site-directed mutagenesis of His[149] has shown the importance of this residue in the metal incorporation pathway and for achieving an active fold. The metal coordination of the protease was also investigated by circular dichroism and electronic absorption spectroscopies using a Co(II)-substituted enzyme. We show evidence for rearrangements of the metal coordination geometry induced by complex formation with an NS4A peptide cofactor. No such changes were observed upon binding to a substrate peptide. Also, $CN^-$ and $N_3^-$ induced Co(II) ligand field perturbations, which went along with an 1.5-fold enhancement of protease activity.

The hepatitis C Virus (HCV)[1] has been identified as the major etiologic agent of parenterally transmitted non-A non-B hepatitis (1, 2). HCV is an enveloped virus with a positive-stranded RNA genome of 9.4 kb, which is translated into a precursor polyprotein of about 3010 amino acids (3). Both cellular and virally encoded proteases are involved in the maturational proteolytic processing of this precursor. Whereas the structural viral proteins arise from signal peptidase-catalyzed cleavages (4), two different proteolytic activities encoded by the HCV NS2 and NS3 proteins are responsible for the processing of the nonstructural region of the polyprotein. The NS2-NS3 precursor is cleaved intramolecularly by an autoprotease, the activity of which was shown to be zinc-dependent (5). The N-terminal part of the NS3 protein, furthermore, contains a 20-kDa serine protease domain that accomplishes all cleavage events downstream of NS3, including the generation of the mature viral polymerase (6). In order to perform its physiological task, the NS3 serine protease has to bind to the viral protein NS4A (7, 8). This binding event leads to an enhancement of the protease activity and to a stabilization of NS3. *In vitro*, activation of NS3 can be achieved by addition of peptides harboring residues 21–34 of NS4A (9–12).

Based on a homology model, we were able to predict the presence, in the NS3 protease domain, of a tetradentate metal binding site formed by three cysteines (Cys[97], Cys[99], and Cys[145]) and one histidine residue (His[149]) (13). Biochemical characterization has confirmed this prediction and demonstrated the presence of a zinc ion in a tetrahedral environment (13, 14). This zinc ion was shown to be essential for the structural integrity of the protein; its removal leads to unfolding and aggregation of the enzyme. Mutagenesis experiments have shown that mutations affecting any of the three cysteine residues resulted in an impaired NS3 protease activity as judged from *in vitro* translation experiments (14, 15). On the other hand, mutagenesis of His[149] into alanine had only minor effects on the autoprocessing of NS3-containing precursor polyproteins (14, 15). It is presently not clear whether the zinc ion of the NS3 protease is identical to the zinc that has been shown to be essential for the NS2-NS3 autoprotease activity.

Crystal structures of the NS3 protease domain (16) and of the complex with an NS4A cofactor peptide (17) have been published. Both structures confirm the prediction of the metal binding site and precisely locate it on the surface of the protein, well exposed to the solvent. In the structure of the NS3-cofactor complex the zinc ion is coordinated to the three predicted cysteine ligands and, through a water molecule, to His[149]. The indirect interaction between the metal and the histidine ligand is consistent with the weak effects of mutations in this position. On the other hand, coordination by His[149] through a water molecule is observed only in two of the three monomers in the asymmetric unit of the crystals obtained in the absence of the NS4A peptide (16). In the third monomer, the His[149]-N$\delta$ moves away from the zinc and thus does not participate in the coordination. The precise function of water and the influence of the cofactor on the coordination sphere of the zinc atom are still open issues because the metal binding site is located on the protein surface, and crystal packing is likely to play a role.

[1] The abbreviations used are: HCV, hepatitis C virus; U, $\alpha$-aminobutyric acid; CHAPS, 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate; DABCYL, 4-[[4′-(dimethylamino)phenyl]azo]benzoic acid; DTT, dithiothreitol; EDANS, 5-[(2′-aminoethyl)amino]naphthalenesulfonic acid; NOE, nuclear Overhauser enhancement; NOESY, nuclear Overhauser enhancement spectroscopy; NS, nonstructural; SW, sweep width; TOCSY, total correlated spectroscopy; MES, 4-morpholineethanesulfonic acid; ROESY, rotating frame Overhauser spectroscopy; HSQC, heteronuclear single quantum coherence.

Hence, the interest in further investigations of the solution state.

In the present study, we have investigated zinc- and cobalt-substituted NS3 protease using NMR, CD, and visible spectroscopy. We present evidence for pH-dependent changes in the metal coordination by His[149] that are compatible with the presence of a bridging hydroxyl group as metal-ligand. We furthermore show that the coordination geometry of the metal may undergo different conformational rearrangements that influence enzymatic activity.

## MATERIALS AND METHODS

*Purification of NS3 Protease Domain*—A plasmid containing the serine protease domain of NS3 (amino acids 1–180, from HCV Bk strain, followed by the sequence ASKKKK) cDNA under the control of the bacteriophage T7 gene 10 promoter was used to transform *Escherichia coli* BL21(DE3) cells (18). Protein expression and purification was carried out as described previously (13). $^{15}$N-Labeled NS3 was prepared using $(^{15}NH_4)_2SO_4$ in a minimal medium containing 100 mM potassium phosphate, pH 7.0, 0.5 mM MgSO$_4$, 0.5 mM CaCl$_2$, 13 $\mu$M FeSO$_4$, 50 $\mu$M ZnCl$_2$, 7 $\mu$M thiamine, 6 $\mu$M biotin, and glucose (4 g/liter). Purity of the enzyme was evaluated to be >95% by silver-stained SDS-polyacrylamide gels and by reversed-phase high performance liquid chromatography using a Vydac C4 column (4.6 × 250 mm, 5 $\mu$m, 300 Å). In the latter case, eluents were H$_2$O, 0.1% trifluoroacetic acid (eluent A) and acetonitrile, 0.1% trifluoroacetic acid (eluent B). A linear gradient from 3 to 95% eluent B in 60 min was used. The concentration of protein stocks was estimated by quantitative amino acid analysis.

*Synthesis of Co(II) NS3 Proteases*—Co(II)-containing recombinant wild type and S139A NS3 proteases were biosynthetically prepared as described previously (13). Contents of cobalt and adventitious zinc were determined on nitric acid-hydrolyzed proteins by atomic absorption spectroscopy on a Perkin-Elmer model 2100 atomic absorption spectrometer equipped with a graphite furnace. Standardized Zn$^{2+}$ and Co$^{2+}$ solutions were purchased from Merck. Glassware used for metal analysis was washed with 50% nitric acid and thoroughly rinsed with Chelex-100-treated deionized water. Wild type and S139A Co(II) proteins had a protein:metal stoichiometry of 1:0.91 ± 0.02 and 1:0.95 ± 0.03, respectively.

*NMR Spectroscopy*—NMR spectra of Zn NS3 were measured with Bruker AMX 400 (one-dimensional experiments) and AMX 500 (two-dimensional experiments) instruments. Samples (0.4–0.6 mM) were brought into the $^2$H$_2$O solvent system by extensive dialysis against 99% $^2$H$_2$O buffer 4% $^2$H$_5$-glycerol, 0.1% CHAPS, 3 mM DTT, 10 mM sodium phosphate, 10 mM MES, pH 6.5, at 4 °C under N$_2$. Protein samples were let exchange protons for 5 days before any measurements were acquired. The presence of glycerol in solution greatly enhances the viscosity, so that 20 °C was established to be the lower limit. On the other hand, the enzyme activity is greatly enhanced above 30 °C, so the working temperature of 25 °C was set by a compromise between these two restricting factors.

For the pH titration experiments, protein samples were dialyzed for 2 h against 20 volumes of the same buffer with a slightly different pH. The pH was measured by a PHM84 Radiometer-pH meter equipped with an Aldrich calomel microelectrode. The one-dimensional titration experiments were acquired with a SW of 4800 Hz and 8K complex data points. No water presaturation was necessary. The data were analyzed with the SWAN-MR software (19). The p$K_a$ deviation due to the deuterium isotope effect was estimated to be +0.05 pH units, comparing the pH and p$^2$H profile of the imidazole $\epsilon$1H in the same protein buffer system. Data reported below are already corrected for this difference. Data were fitted using the following equation,

$$\text{ppm} = a + b*10^{(\text{pH}-pK_a)}/(10^{(\text{pH}-pK_a)} + 1) \quad \text{(Eq. 1)}$$

or, with allowance for cooperativity,

$$\text{ppm} = a - (K^n/(K^n + 10^{-n(\text{pH})}))\,(a - b) \quad \text{(Eq. 2)}$$

where ppm is the observed chemical shift, $a$ and $b$ are the asymptotic values, and $n$ is the Hill coefficient.

A line width analysis of peak C, identified as His[149], in the pH range 5.0–8.0 was carried out assuming the following minimal model,

$$C-HO^- + 2H^+ \underset{k_{-1}}{\overset{k_{+1}}{\rightleftarrows}} C^+ + H_2O \quad \text{(Eq. 3)}$$
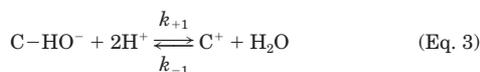
where C and C$^+$ represent the imidazole moiety in the unprotonated and protonated forms, respectively, and $k_{+1}$ and $k_{-1}$ are the kinetic constants of the process. The rate constant for the deprotonation reaction, $k_{-1}$, was estimated from the dependence of the apparent spin-spin relaxation times on the molar fraction of the two His species described by the following equation (20),

$$(1/T_{2C})_{\text{obs}} = 1/T_{2C} - k_{-1}(f\text{C}^+/f\text{C}) \quad \text{(Eq. 4)}$$

where $(1/T_{2C})_{\text{obs}}$ and $(1/T_{2C^+})_{\text{obs}}$ are the observed line widths at different pH values, $1/T_{2C}$ and $1/T_{2C^+}$ are the spin-spin relaxation times in absence of exchange, and $(f\text{C}^+/f\text{C})$ is the ratio of the respective molar fractions.

Two TOCSY experiments ($\tau_m$ = 28 and 48 ms), at several different pH values, were acquired, with a SW of 6000 Hz (t2) × 6000 Hz (t1) and 1K × 256 complex data points. No water presaturation was necessary. The time dependence of the exchange process was further investigated at pH 6.0 with two-dimensional NOESY spectroscopy. A set of experiments was acquired, with mixing times of 10, 30, 60, 100, and 200 ms. The SW values were 6000 Hz (t2) × 6000 Hz (t1) and 1K × 256 complex data points. All the experimental matrices were identically transformed with a Gauss to Lorentzian apodization function (t2) and a 70 degree shifted sine bell function (t1); the final dimensions after transform were 2K (t2) × 1K (t1). All the two-dimensional experiments were analyzed using the nmrPipe (21) and NMRView (22) programs on an Indigo2 SGI workstation. The kinetic process was analyzed following the procedure outlined by Clore *et al.* (23) that will be briefly summarized here. The time development of magnetization in the system described by Equation 3 is given by the following equation.

$$dM_C/dt = -(M_C - M_C^e)R_C - k_{+1}M_C + k_{-1}M_{C^+}$$

$$dM_{C^+}/dt = -(M_{C^+} - M_{C^+}^e)R_{C^+} + k_{+1}M_C - k_{-1}M_{C^+} \quad \text{(Eq. 5)}$$

The equilibrium constant of the process, at pH 6.0, is defined as $K = k_{+1}/k_{-1} = [C^+]/[C]$, where the concentration of hydrogen ion has been incorporated in the constant; $R_C$ and $R_{C^+}$ are the total spin lattice relaxation rates of the relevant protons. At time 0, the intensity of the diagonal peak of species C is set to 1 and the cross-peak to 0, whereas the equilibrium magnetization ($M_C^e$ and $M_{C^+}^e$) is expressed as a function of the constant rates by a trivial rearrangement of the mass conservation law. The build-up curves were fitted simultaneously to the differential equations (Equation 5), under the control of a nonlinear least squares optimization routine varying the values of the dissipative rates and of the kinetic constants, using the program MLAB (Civilized Software, Bethesda, MD).

The Co(II)-substituted NS3 samples in $^2$H$_2$O, for the NMR measurements, were prepared by H→$^2$H exchange of protein samples loading the purified protein (10 mg) on an Amersham Pharmacia Biotech HR 5/5 Mono S column equilibrated with H$_2$O buffer (4% glycerol, 0.1% CHAPS, 10 mM sodium phosphate, pH 7.5). The protein was then washed with the same buffer in 99% $^2$H$_2$O (buffers were flushed with dry nitrogen) and eluted by a steep NaCl gradient. The protein was further concentrated to 0.4 mM and dialyzed twice against 20 volumes of $^2$H$_2$O, 4% $^2$H$_5$-glycerol, 0.1% CHAPS, 10 mM sodium phosphate, pH 7.5, at 4 °C under argon. After storage for 2 days at 4 °C in $^2$H$_2$O buffer, the protein samples (0.3–0.4 mM) were analyzed by one-dimensional NMR spectroscopy.

$^{15}$N-filtered NOESY (60 and 100 ms) and ROESY (20 ms) were acquired in water using the $^{15}$N-labeled sample, at pH 6.0 and 6.6. The $^{15}$N labeling was used to filter out of the spectra all of the resonances originating from the amide groups of the protein, so that in the region 7.5–9.0 ppm, only the resonances originating from the $\epsilon$1H protons of the histidines were visible. The SW and data treatment were the same as described previously. The experiments were acquired using water flip back pulses combined with z-gradients to avoid saturation of the water signal and to allow the use of a convenient receiver gain (512 in our experiments). $^1$H-$^{15}$N HSQC experiments were acquired at pH 6.3 and 6.6, using as transfer delays $\tau_m$ = (1/(4J)) = 11.0 ms, to obtain the coherence transfer from the $\epsilon$1H and $\delta$2H protons to the $\epsilon$2N and $\delta$1N through the $^2$J coupling constant. The SW values were 6000 Hz and carrier at 7.8 ppm (t2) and 4700 Hz and carrier at 200 ppm (t1). The experimental matrix was 1K × 80 complex data points; it was transformed using a Gauss to Lorentzian apodization function (t2) and a 62°-shifted sine bell function to yield a final matrix of 2K × 512 data points.

*H149A Mutant and Apo NS3 Protease*—The H149A mutation was inserted by polymerase chain reaction site-directed mutagenesis using suitable primers. Full sequencing of the mutated cDNA confirmed the

presence of a single mutation. Transformed *E. coli* BL21(DE3) cells were grown at 37 °C and induced at an $A_{600\ nm}$ of 0.7–0.9 with 0.4 mM isopropyl-1-thio-β-D-galactopyranoside for 3 h at 37 °C in LB medium. Under these conditions, the protease was found in the insoluble fraction. The harvested cells were resuspended in 100 mM NaCl, 50 mM Tris-HCl, pH 8.0 (10% v/v), and 1 mg/ml lysozyme. After 20 min of incubation at room temperature, pellets were disrupted in a French pressure cell. The insoluble material was collected by centrifuging the homogenate at 5000 × g for 10 min at 4 °C. It was then resuspended in ice-cold 100 mM NaCl, 1 mM EDTA, 0.1% sodium deoxycolate, and 50 mM Tris-HCl, pH 8.0. After 10 min of mixing, 8 mM MgCl₂ was added, together with 10 μg/ml DNase I. Digestion was carried out for 45 min, followed by 10 min of centrifugation at 10,000 × g. Pellets were resuspended and centrifuged in 1% Nonidet P-40, 100 mM NaCl, 1 mM EDTA, 50 mM Tris-HCl, pH 8.0, and washed in 1 M urea, 50 mM Tris (pH 8.0). This procedure was repeated twice before dissolving the pellets in 8 M urea and 60 mM DTT overnight at 4 °C. The dissolved material contained protease with a purity >80%. Further purification to >90% purity was done on a Amersham Pharmacia Biotech Superdex 75 Hi-Load 26/60 gel filtration column, equilibrated in 6 M urea, 30 mM DTT and 100 mM sodium phosphate, pH 7.5. The gel-filtered protein (1 μM) was dialyzed against 1% trifluoroacetic acid, 500 mM urea for 16 h and against 0.1% trifluoroacetic acid for a further 18 h before being concentrated to 25 μM, nitrogen shock frozen, and stored at −80 °C. Protein stocks were quantified by amino acid analysis, and the purity grade was assessed as described previously.

Apo wild type NS3 was obtained by denaturing purified zinc protein in 8 M urea, 30 mM DTT, and 10 mM EDTA. Two steps of dialysis against trifluoroacetic acid as described previously were applied to remove urea. Then zinc content of the acid denatured enzymes has been estimated to be less then 1% of the total protein concentration. Proteins were refolded before being assayed in 50 mM Tris, pH 7.5, 2% CHAPS, 50% glycerol, 10 mM DTT (standard activity buffer) and treated with Chelex-100 resin before the addition of 10 μM ZnCl₂.

Apo wild type (21 μM) and His¹⁴⁹ proteins (25 μM) were diluted to 0.1–1 μM in a cuvette containing standard activity buffer with different amounts of ZnCl₂ (1.5–100 μM) or 100 μM EDTA. Upon dilution, fluorescence traces ($\lambda_{EX}$ = 280 nm, 2.5 nm slits; $\lambda_{EM}$ = 340 nm, 8 nm slits) were recorded under continuos stirring on a Perkin-Elmer LS 50B fluorescence spectrometer equipped with a cuvette holder, thermostatted at 22 °C.

Experiments were repeated in triplicate and analyzed with the help of a Grafit software (Erithacus).

*High Performance Liquid Chromatography Activity Assay and Active Site Titration*—Peptide synthesis was performed by Fmoc (*N*-(9-fluorenyl)methyloxycarbonyl)/t-Bu chemistry as described previously on a NovaSyn Gem flow synthesizer (18, 24). Concentrations of stock solutions of peptides, which were prepared in Me₂SO or in buffered aqueous solutions and kept at −80 °C until use, were determined by quantitative amino acid analysis performed on HCl-hydrolyzed samples.

Standard cleavage assays, unless otherwise stated, were performed in 57 μl of 50 mM Tris, pH 7.5, 2% CHAPS, 50% glycerol, 10 mM DTT to which 3 μl of a synthetic peptide, derived from the cleavage sequence of the NS4A-NS4B junction ($S_1$ = DEMEEC-ASHLPYK), were added. DTT was omitted from buffers when assaying Co(II)-NS3 and wild type enzyme for KCN perturbation. As a protease cofactor, we used a 17-mer peptide corresponding to the central hydrophobic core of the NS4A protein spanning residues 21–34 with an N-terminal lysine tag (4A peptide, KKKGSVVIVGRIILSGR). Enzyme concentrations (10 nM to 1 μM) and incubation times were chosen in order to obtain <10% substrate conversion. Reactions were quenched by addition of 40 μl 1% trifluoroacetic acid and analyzed by high performance liquid chromatography on a Lichrospher C18 reversed-phase cartridge column (4 × 125 mm, 5 μm, Merck) using a 10–50% acetonitrile gradient at 5%/min. Peak detection was accomplished by monitoring both the absorbance at 220 nm and tyrosine fluorescence ($\lambda_{ex}$ = 260 nm, $\lambda_{em}$ = 305 nm).

Active site titrations were performed on an SX-MV18 Applied Photophysics stopped flow apparatus in standard activity buffer. Under these conditions, we measured a dead time of 10 ms with a standard reference reaction (reduction of 2,6-dichlorophenolindophenol by L-acorbic acid) (25). We used a fluorescence resonance energy transfer based ester substrate (26, 24) having the sequence $S_2$ = Ac-DED(EDANS)EE U Ψ[COO]ASK(DABCYL)-NH₂ to measure the initial burst phase of ester hydrolysis. This is directly proportional to the active site concentration (27). Fluorescence standard curves were calculated either by measuring the fluorescence of known concentrations of 100% hydrolyzed substrate or by reference to fully active wild type protein (24). Wild type and H149A mutant proteases at concentrations of 50, 100,
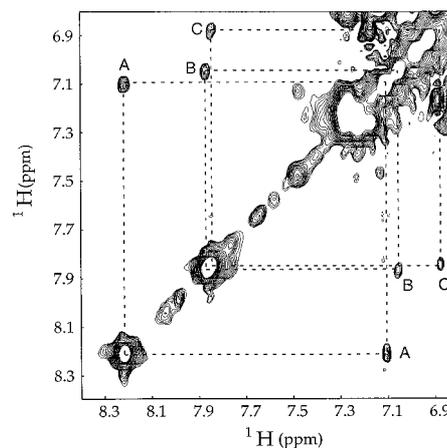


FIG. 1. **Two-dimensional NMR TOCSY spectra of Zn NS3 protease aromatic region.** Spectra were recorded in ²H₂O buffer containing 4% ²H₅-glycerol, 0.1% CHAPS, 1 mM DTT, 10 mM sodium phosphate, pH 6.4. Histidine ε₁H-δ₂H cross-peaks are labeled with letters; *A* corresponds to His⁵⁷, *B* to His¹¹⁰, and *C* to His¹⁴⁹.

and 200 nM were incubated with 80 μM 4A peptide and rapidly mixed with an equal volume of 50 μM $S_2$. Fluorescence traces (ex = 355 nm) were recorded for 5 s with a logarithmic time scale acquisition, using a 400-nm-cutoff filter. The averages of eight single determinations were analyzed by extrapolating the linear phase to zero time.

*Electronic Absorption and Circular Dichroism Spectroscopy*—80 μM of wild type Co(II) NS3 or 100 μM of the S139A mutant were incubated with different amounts of 4A peptide or $S_1$ in 50% glycerol, 2% CHAPS, 30 mM NaCl, and 20 mM sodium phosphate, pH 7.5. Sample concentrations were determined by quantitative amino acid analysis. KCN and NaN₃ additions were carried out using 40 μM Co(II) NS3 in presence of 150 μM 4A peptide. Electronic spectra of proteases were acquired on a Varian Cary 3E dual-beam spectrophotometer. Spectra were recorded at 15 °C at 60 nm/min scan speed with 1-cm path length quartz cuvettes. Circular dichroism measurements were performed using a Jasco 710 spectropolarimeter equipped with a cell holder thermostatically controlled by a circulating water bath. Spectra were collected with a 8-s time constant and a 5 nm/min scan speed at 15 °C by using rectangular quartz cells of 1 cm path length and a protein concentration of 40 μM. The mean residue ellipticity Ø was calculated referring to the protein residues concentration. Data were analyzed with the help of Kaleidagraph software (Abelbeck).

## RESULTS

*NMR Spectroscopy*—The frequencies arising from the ε1H and δ2H nonexchangeable protons of the three histidine residues were identified in the two-dimensional TOCSY experiment through their characteristic cross-peak correlation. In Fig. 1 the aromatic region of the TOCSY experiment at pH 6.4 is shown. The His ε1H→δ2H cross-peaks are labeled *A* (ε1H = 8.22; δ2H = 7.10), *B* (ε1H = 7.87; δ2H = 7.05), and *C* (ε1H = 7.84; δ2H = 6.88). The same lettering is held also in the one-dimensional set of the experiments shown in Fig. 3*a*, where the pH behavior of the His signals is explored.

To assign the His¹⁴⁹ ε1H signal, we have used a Co(II)-substituted protein. Co(II) has been used as a substitute for zinc in many metalloproteins due to its similar coordination behavior, leading to very little alteration of the coordination geometry (28). The paramagnetic nature of Co(II) isotropically shifts resonances of the surrounding atoms (29–32). This effect permits the direct identification of the protons in proximity to the paramagnetic metal ion. The imidazole ε1H signal of histidine residues involved in metal coordination is generally shifted to the downfield region of the spectrum (29).

One-dimensional spectra of Co(II) NS3 were obtained at pH 6.85 (Fig. 2). Experiments were also repeated at different pH values using both the Zn- and the Co(II)-substituted enzymes (Fig. 3*a*). In Fig. 2, it can be seen that the signal at 7.84 ppm (peak C) was not present in the Co(II) NS3 spectra and that a

FIG. 2. **NMR spectrum of Co(II) NS3 protease.** The spectrum was recorded in $^2H_2O$ buffer containing 4% $^2H_5$-glycerol, 0.1% CHAPS, 10 mM sodium phosphate, pH 6.85. The exponential line broadening of the aromatic region was set to 1 Hz, and the downfield region of Co(II) enzyme was processed with a 30 Hz exponential line broadening function. Baselines were corrected by a fourth-order polynomial function. The *arrow* indicates the position were peak C is missing.
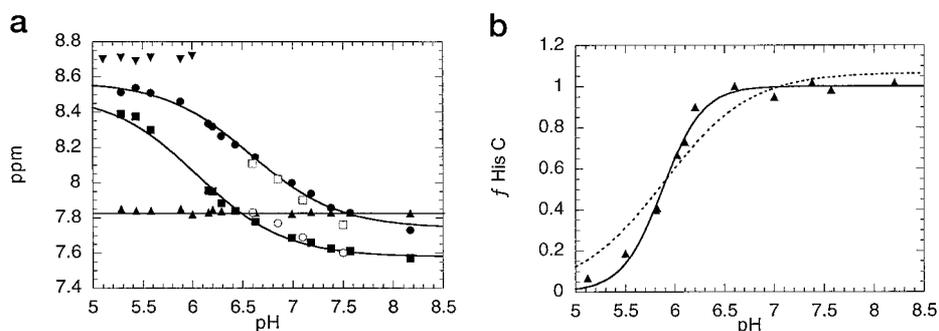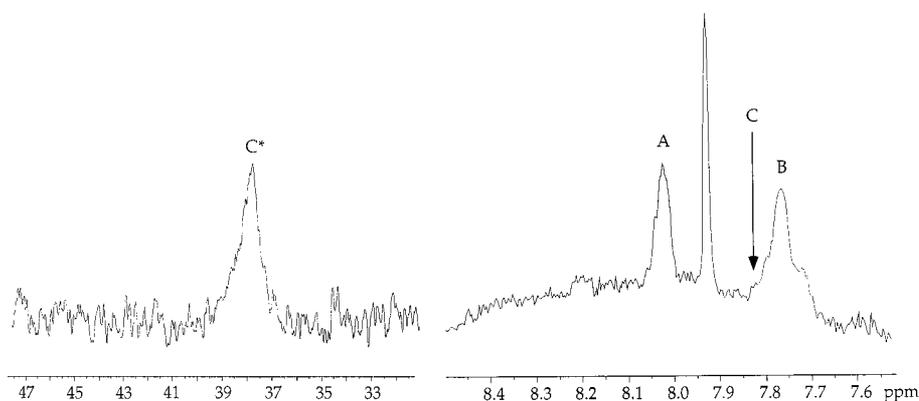


FIG. 3. **pH dependence of the histidine resonances of Zn and Co(II) NS3 protease.** *a*, resonance shift of the three histidine signals as a function of pH. Zn NS3: His A (●), His B (■), His C (▲), and His C$^+$ (▼); Co(II) NS3: His A (○) and His B (□). His A and B *data points* are fitted with Equation 1, and His C is fitted with a linear equation. *b*, His C (▲) peak intensities as a function of pH. Data were fitted either with Equation 1, assuming a simple equilibrium (*dashed line*), or with Equation 2, taking into account cooperativity (*solid line*).

new broad peak was observed at 38.5 ppm (peak C*). This shift in the downfield region is consistent with the previously reported shifts of the histidine $\epsilon 1$ proton of Co(II)-substituted enzymes (30, 32) and with the imidazole spectra of model complexes, such as Co(1Me-Imi)2Cl2 and Co(Imi)2Cl2 (29, 31). The broad line width is characteristic of fast relaxing protons influenced by distorted tetrahedral coordinations of Co(II) complexes (29, 32). Enzyme preparations with different Co/Zn ratios were obtained by controlled dialysis against zinc-containing buffer. The Co/Zn ratio determined by atomic absorption spectroscopy correlated with the ratio of the intensities between the peaks at 38.5 and 7.84 ppm (data not shown). These findings assign the resonance of the $\epsilon H$ of the zinc-bound His[149] to that of peak C.

We next investigated the pH dependence of the histidine resonances in the zinc-containing protease. The three $\epsilon 1H$ signals exhibit a reversible pH profile (Fig. 3), with peaks A and B shifting downfield (Fig. 3*a*) as pH moves from basic to acidic. These signals therefore correspond to a protonation equilibrium, which is in fast exchange regime on the NMR time scale. The chemical shifts of peaks A and B were fitted using Equation 1, and the pK values obtained were 6.8 (peak A) and 6.2 (peak B). The pK of 6.8 compares well to the value of 6.9 found in activity titration experiments[2] and thus reflects the pK of the histidine residue of the catalytic triad (His[57]). His B has a pK of 6.2, which reflects the pK of a protein imidazole moiety exposed to the solvent (33), and was thus assigned to His[110]. Peak C (7.84 ppm) behaves differently (Fig. 3*b*): in the pH range 8.3–6.2, the chemical shift and the intensity of the peak is invariant. When the pH was lowered below 6.2, its intensity decreased, and a new peak arose at 8.71 ppm. The pK of this reversible transition has been estimated at 5.9. The best data

fit was achieved by Equation 2, which yields a Hill coefficient of $n = 2.1 \pm 0.2$ (Fig. 3*b*), indicating a cooperative association of at least two protons. We could exclude the possibility that this steep titration curve resulted from irreversible processes because in the same pH range, the titration curves of the other two histidine residues could be best described by a one-proton association, and the same titration curves were obtained proceeding either from high or from low pH. The line widths of species C and C$^+$ are 13 and 6 Hz, respectively. The line width analysis carried out on peak C, assuming the minimal model reaction described by Equation 3, resulted in a $k_{-1}$ of $\sim 9.4$ s$^{-1}$.

The protonated and unprotonated forms of C are in slow chemical exchange, as demonstrated by the two-dimensional NOESY experiments (Fig. 4*a*). The kinetic analysis of the slow exchange process was carried out assuming the model described in Equation 3, at pH 6.0. After the fitting procedure (Fig. 4*b*), we obtained the following macroscopic rates: $k_{+1}$ and $k_{-1}$ values of $5.3 \pm 0.9$ and $9.1 \pm 3.1$ s$^{-1}$ and $R_C$ and $R_{C^+}$ values of $0.5 \pm 0.4$ and $7.3 \pm 3.5$ s$^{-1}$, respectively.

In the $^{15}N$ filtered NOESY and ROESY (34) experiments, all three His $\epsilon 1H$ resonances exhibit NOEs, with a hydrogen resonating at the water frequency (not shown). Because the cross-peak disappears when the NOESY spectra are acquired in $D_2O$, we can exclude that the frequency observed arises from an aliphatic proton of the protein. However, it is still difficult to assign it to an NOE arising from water because, as was pointed out by Otting and Wüthrich (35), in order to assign an NOE to an interaction between an OH-hydroxyl and a protein hydrogen, one must rule out the possibility that the observed effect arises from chemically exchanging protons of the protein itself. In our case, an NOE involving a hydroxyl hydrogen and the $\epsilon 1H$ proton is ambiguous because within less than 3 Å, there are the exchangeable protons of the $\epsilon 2N$ and $\delta 1N$ nitrogens.

It is convenient to use the nomenclature introduced by Witanowski *et al.* (36) to classify the protonation and the tauto-

**a**



**b**
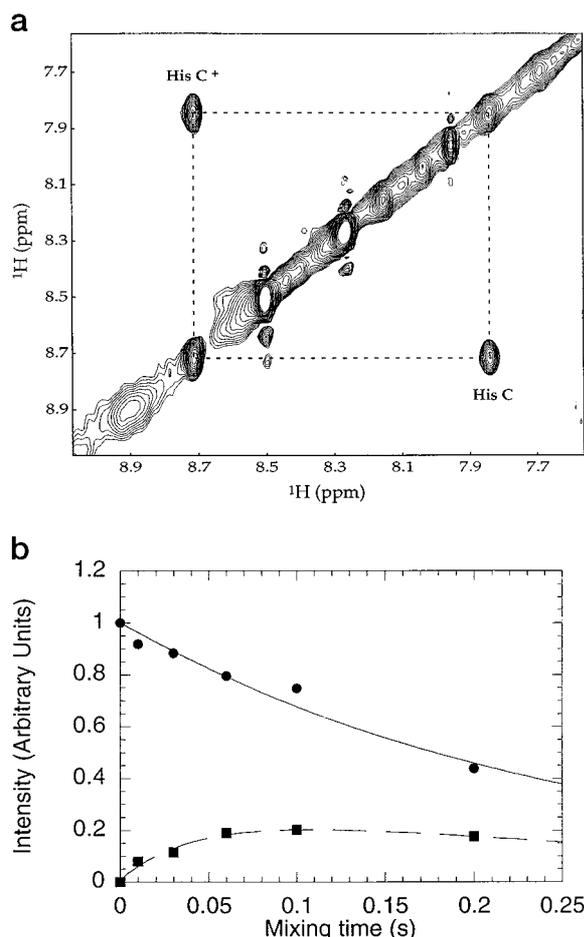


FIG. 4. **Magnetization exchange of His[149] species.** *a*, two-dimensional NMR NOESY spectra of the aromatic region of Zn NS3 protease. The spectrum was recorded in $^2H_2O$ buffer containing 4% $^2H_5$-glycerol, 0.1% CHAPS, 1 mM DTT, 10 mM sodium phosphate, pH 6.0, with a mixing time of 100 ms. The spin system C in slow exchange (7.8 ppm) and the protonated C+ (8.71 ppm) are highlighted with *letters* and *dashed lines*. Peak C was identified with His[149]. *b*, time development of magnetization of peak C. The intensity values measured for the diagonal (●) and the cross-peak (■) of the C-C+ process in the series of two-dimensional NOESY experiments were fitted simultaneously to the system of Equation 5 with a nonlinear least squares optimization routine, using the program MLAB.

TABLE I
*Reference values for His nitrogen states with respect to liquid $NH_3$*

| His States | N: β | N—H α | ⊕N—H α+ |
|---|---|---|---|
| | ppm | ppm | ppm |
| Solvent Exposed | 249.5 | 167.5 | 176.5 |
| H-Bonded | 239.5 | 177.5 | 186.5 |
| Zn Binding | 199.5–220.0 | 169.5–177.0 | - |

tion, and its interpretation is given under "Discussion."

*Characterization of the H149A Mutant*—In order to assess the functional role of His[149] in zinc coordination, we have mutagenized this residue into alanine by polymerase chain reaction site-directed mutagenesis (H149A). The H149A mutation has been already described as leading to a less active enzyme, which is accumulated in the insoluble fraction when expressed in bacteria (14, 15).

The H149A mutant was obtained pure to >90% by purification in a metal-free form from the insoluble fraction (see "Materials and Methods"). The amount of refolded H149A mutant enzyme was estimated by active site titration. The number of active sites for H149A was found to be 60% (±2%) of the total protein. Under the same conditions, 95% (±5%) of the wild type protein was refolded into an enzymatically active species. The specific activity of the refolded wild type protein was found to be indistinguishable from that of a preparation purified under native conditions.

The H149A mutant was found to have both impaired proteolytic activity, resulting in decreased $k_{cat}$ values, and an impaired affinity for the NS4A cofactor peptide (Table II). To establish the stability of the zinc coordination in the H149A mutant, we performed an EDTA titration. EDTA has previously been shown to inhibit Zn NS3 protease activity at high concentrations (13). In order to rule out differences due to the refolding protocol, we also compared the susceptibility to EDTA of the refolded and the native wild type proteins. The mutated protein was slightly more susceptible to inhibition by EDTA: we determined an $IC_{50}$ of 10.1 ± 3.4 mM for the H149A mutant compared with 21.4 ± 4 mM for the wild type enzyme (not shown). No measurable difference in susceptibility to EDTA was observed for the two wild type enzyme preparations (refolded and native). The relatively small difference in the inhibitory potency of EDTA between the H149A and the wild type enzymes suggests that the zinc atom is still strongly coordinated, even in absence of His 149. Therefore, this residue does not seem to play a pivotal role in the stabilization of coordination once the zinc ion is bound.

Refolding of acid-denatured NS3 protease has been shown to be a zinc-dependent process, accompanied by a decrease in tryptophan fluorescence.[3] We took advantage of these findings to determine the role of His[149] in the incorporation of zinc into the metal binding site. The H149A and wild type refolding kinetics were performed at plateau of zinc concentration with respect to the refolding efficiencies (Fig. 6).

meric states of His residues. In this notation, each nitrogen of the histidine imidazole ring is assigned to one of the following three categories: $\alpha$ if it is an N-H, $\alpha+$ if it is an N-H of a positively charged ring, and $\beta$ if it is a nonprotonated N. The reference values reported in Table I for the different states of the His nitrogens and for their hydrogen bonding state originate from several different systematic studies on model systems and on proteins (33, 36–45). On the other hand, there are very few literature data on the chemical shifts of histidine nitrogen atoms involved in zinc coordination (45–47), and so far, no attempt has been made to rationalize the observed values. It is possible for a His residue to assess both the tautomeric and the protonation state, at a given pH, from the disposition of the peaks in an $^1H$-$^{15}N$ HSQC-type experiment (44, 45). In Fig. 5, the HSQC at pH 6.6 is shown. At this pH value, we find the following chemical shifts: His[57] $\epsilon 2N$ = 184.0 ppm, $\delta 1N$ = 193.9 ppm; His[110] $\epsilon 2N$ = 223.6 ppm, $\delta 1N$ = 169.9 ppm; and His[149] $\epsilon 2N$ = 211.9 ppm, $\delta 1N$ = 180.0 ppm. The chemical shift of His[57] and His[110] can be explained by an equilibrium between the cationic and the amphionic forms with equivalent populations for His[57] and about 30% of cationic form for His[110]. The chemical shift of His[149] deserves more atten-

---

The fluorescence trace in the presence of EDTA shows a linear decrease in time both for the wild type and for the mutant protein. This could be ascribed to the formation, in the absence of accessible zinc ions, of a misfolded inactive enzyme. In fact, the resulting proteins were shown to be devoid of any appreciable catalytic activity.

In the presence of zinc, the wild type and the H149A mutant protein refold with significantly different velocities (Fig. 6). Furthermore, the refolding process led to only 60% of recovery of total active sites for the H149A mutant, whereas 95% of active enzyme molecules were recovered upon refolding of the wild type enzyme. These findings suggest that the H149A refolding kinetics could be ascribed to the sum of two competing processes, one that drives the formation of an active protein and a second one that leads to a misfolded, inactive enzyme. Increasing zinc concentrations did not augment the amount of refolded mutant protein, indicating that the mutation does not affect metal binding itself but rather leads to an incapacity to promptly assume an active conformation upon interaction with zinc ions.

*Perturbation of the Metal Coordination Sphere by NS3 Ligands*—We next addressed the question of whether conforma-

tional changes may occur in the metal coordination sphere. To this purpose, the effects of complex formation with the NS4A cofactor peptide and with a substrate peptide have been studied. Both substrate and co-factor binding have been shown to be glycerol-dependent (11). This precluded observation by NMR due to the excessive line-broadening caused by the high viscosity of glycerol-containing buffer solutions. We therefore chose to introduce Co(II) into the metal binding site, taking advantage of the properties of Co(II) as a spectroscopic probe. The high structural similarity between the Co(II) and the Zn NS3 is highlighted by the identical position in the NMR spectrum, recorded at a low glycerol concentration, of the His signals A and B, which also show the same pH dependence in the two proteins, as well as by the similar far UV circular dichroism spectra (data not shown). This evidence and the comparable kinetic data (Table II) let us consider the Co(II)-NS3 as a good structural substitute of the wild type zinc protein.

UV/Visible spectroscopy and circular dichroism were chosen as techniques, because the Co(II)-substituted protein failed to yield good quality low-temperature electron paramagnetic resonance spectra, possibly due to fast relaxation rates and resulting line broadening (data not shown).

Co(II) NS3 shows complex ligand field spectra with two major bands at 640 ($\epsilon = 460$ M cm$^{-1}$) and 685 nm ($\epsilon = 400$ M cm$^{-1}$) and two minor shoulders on either side at 585 and 740 nm. This transition envelope is characteristic of distorted tetracoordinated high-spin Co(II) complexes (28, 29). A strong charge transfer S→Co(II) band, belonging to the three Cys residues in coordination, was observed at 365 nm with a shoulder at 320 nm. Changes in the Co(II) ligand field spectrum became evident at pH values below 6, consistent with a change in the coordination geometry upon protonation of His$^{149}$ (not shown). At this acidic pH, the Co(II) NS3 achieved an unstable conformation, which is particularly prone to losing the metal ion. Therefore, in order to test the effect of protonation of His$^{149}$ on the affinity of the enzyme for the NS4A cofactor, the dissociation constants for the Zn NS3-NS4A peptide complex were determined at pH 7.5 and 5.1. We obtained values of $K_d = 5.3$ $\mu$M (pH 7.5) and $K_d = 70$ $\mu$M (pH 5.1), suggesting that the presence of His$^{149}$ in the metal coordination sphere is required to allow efficient co-factor binding. This loss in complex stability is comparable with the one observed for the H149A mutant protease (at pH 7.5 $K_d = 41.3$ $\mu$M), when the metal is no longer ligated by the imidazoyl moiety.

Upon addition of the NS4A peptide, we recorded the following changes in the optical spectra: in the envelope of the d-d transitions, the 685 nm band decreased, whereas the shoulder at 740 nm became more intense (Fig. 7a). Furthermore, the intensity of the charge transfer bands decreased (data not shown) and the intensity in the protein aromatic UV region, around 280 nm, increased (data not shown). This latter effect on the protein UV absorbance is not a specific rearrangement of the Co(II) protein because we also observed it with the zinc protein. These changes in the spectroscopic properties of Co(II) NS3 all reached a plateau at a 1:1 stoichiometric ratio of protein: cofactor (Fig. 7b).

The CD spectrum of the region between 300−420 nm (Fig.
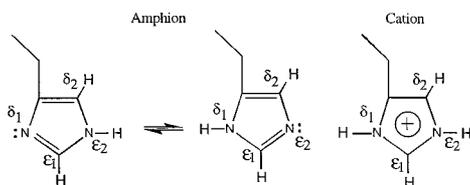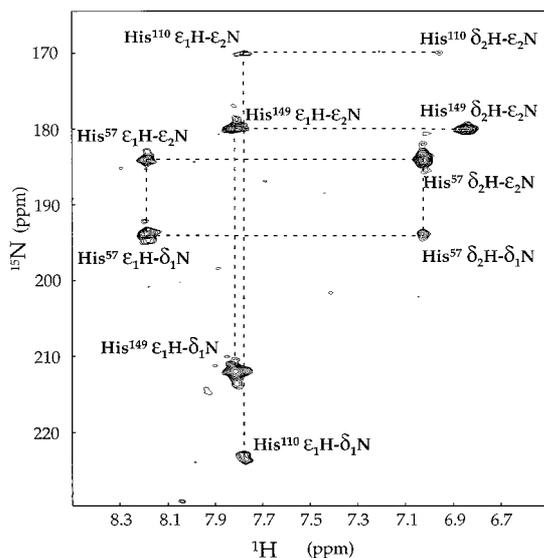


FIG. 5. **15N HSQC.** A portion of the $^1$H-$^{15}$N HSQC spectra of the NS3 protease domain at pH 6.6. The cross-peak pattern and the assignment of peaks is evidenced for all the three histidine residues. The delay during which $^1$H and $^{15}$N signals become antiphase was set to 22 ms $(1/(2J_{NH}))$ to refocus single-bond correlations. The schematic diagram of the nomenclature used through the article is shown at the *bottom* of the figure.

TABLE II
*Steady state kinetic parameters*

| | NS3 | | | NS3/4A | | | $K_d$ NS3/4A |
|---|---|---|---|---|---|---|---|
| | $K_{cat}$ | $K_m$ | $K_{cat}/K_m$ | $K_{cat}$ | $K_m$ | $K_{cat}/K_m$ | |
| | $s^{-1}$ | $\mu$M | $M^{-1} s^{-1}$ | $s^{-1}$ | $\mu$M | $M^{-1} s^{-1}$ | $\mu$M |
| Zn-WT | $8.3 \times 10^{-3}$ | 80.1 | 104 | $58.3 \times 10^{-3}$ | 40.5 | 1460 | 5.3 |
| Zn-H149A | $4.72 \times 10^{-5}$ | 622 | 0.07 | $6.9 \times 10^{-3}$ | 74.4 | 93 | 41.3 |
| Co(II)-WT | $4.3 \times 10^{-3}$ | 103 | 42 | $37.3 \times 10^{-3}$ | 40 | 932 | 7.9 |

8a) shows a positive band at 365 nm that matches with an intense peak in the absorption spectrum and a negative shoulder at 350 nm that is followed by a negative band with the minimum at 322 nm. This latter band corresponds to a shoulder recorded in the absorption spectrum. These bands can be assigned to a complex S→Co(II) charge transfer system, because the number of expected transitions ($S^{\sigma;\pi} \rightarrow dx^2, dx^2\text{-}y^2$) for three cysteines is 12 (49). Upon binding of the NS4A peptide to Co(II) NS3, the three transitions in the CD spectrum underwent a decrease in molar ellipticity, reaching a plateau at a protein:cofactor ratio of 1:1 (Fig. 8b). Therefore, S-Co(II) bonds are structurally reorganized upon binding of the peptide cofactor.

The reported changes in both CD and visible spectra are indicative of rearrangements in the metal coordination geometry occurring upon formation of the NS3-NS4A peptide complex. We next explored whether conformational changes affect-
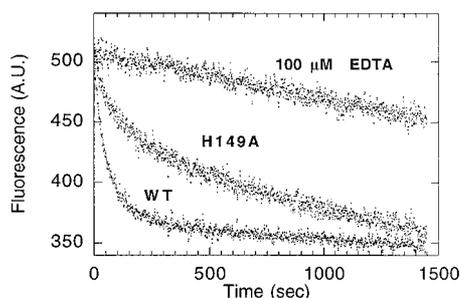


FIG. 6. **Zinc-dependent refolding kinetics of acid-denatured apo NS3 wild type and H149A.** The fluorescence (340 nm) traces *versus* the time of wild type (*WT*) and H149A apo protein diluted 1:100 (200 nM) in a refolding buffer containing 50 mM Tris, pH 7.5, 2% CHAPS, 50% glycerol, 10 mM DTT to which 10 $\mu$M ZnCl$_2$ or 100 $\mu$M EDTA were added. The data in presence of EDTA were obtained using the H149A protein. The wild type protein gave the same results.

ing the metal binding site also occur upon substrate binding. To this purpose, we used a Co(II)-substituted catalytically inactive mutant protein in which the active site serine residue was mutagenized into alanine (S139A). The protein showed $^1$H-$^{15}$N NMR spectra that were virtually superimposable on those obtained with the wild type enzyme (not shown). Furthermore, visible and CD spectra and their perturbation by the addition of the NS4A peptide were indistinguishable from those of the wild type Co(II) protein. Addition of a substrate peptide harboring the sequence of the polyprotein NS4A/NS4B junction to S139A Co(II)-NS3 only resulted in a very small change of the ligand field spectrum characterized by a slight decrease of the two major bands at 640 and 685 nm (data not shown). Furthermore, addition of NS4A peptide to samples at substrate saturation gave rise to the same spectroscopic transitions recorded for the enzyme in the absence of substrate (data not shown). Thus, we can conclude that the substrate does not significantly rearrange the metal coordination geometry, whereas in the ternary complex, the metal experiences a structural environment similar to that in the binary enzyme-cofactor complex.

*Activation of NS3 Protease by Metal Ligands*—Because we have shown that the NS3 protease activator NS4A causes a conformational rearrangement of the metal binding site, we wanted to test whether externally added metal ligands that perturb the metal coordination were also capable of modulating the enzymatic activity. As a proof of this principle, we chose CN$^-$, due to its high affinity for both zinc and cobalt (49), as well as N$_3^-$. Addition of KCN or NaN3 to Co(II)-substituted NS3 protease caused dramatic changes in the visible spectrum of the ligand field (Fig. 9). These changes titrated with a $K_d$ of about 7 mM for N$_3^-$ and with a $K_d = 106 \pm 51$ $\mu$M for CN$^-$. Both N$_3^-$ and CN$^-$ also enhanced Co(II) NS3 activity by a maximum of 1.5-fold. The CN$^-$ activation of the Co(II)-substituted enzyme titrated with $K_d = 176 \pm 96$ $\mu$M, which is in good agree-



FIG. 7. **Perturbation of the Co(II) NS3 ligand field transitions by the NS4A peptide.** *a*, to 80 $\mu$M Co(II) NS3, increasing amounts of NS4A peptide were added in 50% glycerol, 2% CHAPS, 30 mM NaCl, and 20 mM sodium phosphate, pH 7.5, and electronic absorption spectra were recorded. *Arrows* indicate the direction of the intensity changes upon cofactor binding. *b*, the differences in the extinction coefficient (mM) at 690 nm are plotted *versus* the protein cofactor equivalents. Data from three different experiments are shown.
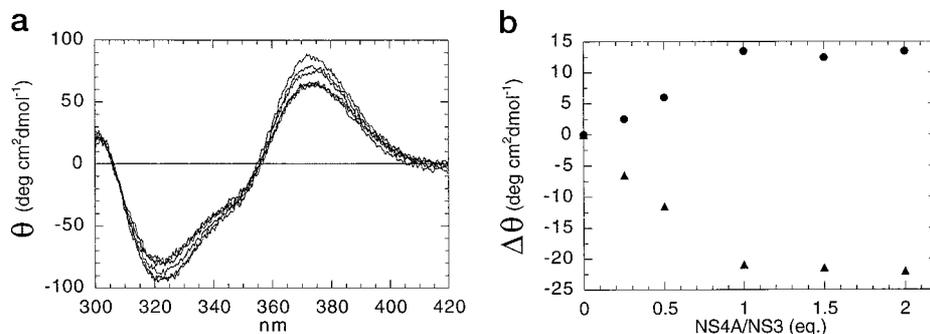


FIG. 8. **Circular dichroism spectra of the Co(II) NS3 S→Co charge transfers upon cofactor binding.** *a*, 40 $\mu$M Co(II) NS3 was incubated with increasing amounts of NS4A peptide in 50% glycerol, 2% CHAPS, 30 mM NaCl, and 20 mM sodium phosphate, pH 7.5, and CD spectra were recorded. *b*, the differences in molar ellipticities at 322 nm (●) and 374 nm (▲) are plotted as a function of added NS4A peptide equivalents.

FIG. 9. **Perturbation of the Co(II) NS3 ligand field transitions by cyanide and azide.** To 40 $\mu$M Co(II) NS3 and 150 $\mu$M 4A peptide in 50% glycerol, 2% CHAPS, 30 mM NaCl, and 20 mM sodium phosphate, pH 7.5 (●), 0.5 mM KCN (■) or 50 mM NaN$_3$ (▲) was added, and spectra were recorded.

ment with the $K_d$ value obtained from the titration of the visible spectral bands (data not shown). $CN^-$ activation was also observed with the native, zinc-containing protein. In this case, a maximum 1.4-fold activation was observed that titrated with a $K_d = 111 \pm 51$ $\mu$M (data not shown). The activatory effects of $CN^-$ and $N_3^-$ were specific and not shown by other anions such as $Cl^-$, $I^-$, or $PO_4^{3-}$, which did not perturb the Co(II) coordination sphere. In fact, both $I^-$ and $PO_4^{3-}$ were competitive inhibitors of the enzyme and thus behaved similarly to $Cl^-$, which has previously been reported to be a competitive inhibitor of the NS3 protease (18).

### DISCUSSION

The crystallographic structures (16, 17) of the NS3 protease have shown that the protein folds in a chymotrypsin-like fold consisting of two $\beta$-barrel-like domains. The topology of these domains is crucial for the correct orientation of the residues of the catalytic triad, which are distributed between the domains. Most of the chymotrypsin-like proteases have disulfide bridges that are believed to maintain the relative orientations of the residues involved in catalysis (50). Disulfide bridges present in these extracellular serine proteases are unlikely to be stable in the reducing intracellular milieu. In a series of viral proteases, which accomplish their physiological role intracellularly, such as the NS3 protease domains of HCV, GB viruses A and B, and hepatitis G virus, as well as in the picornavirus 2A proteases, zinc binding sites probably play an analogous role of structural stabilization. In contrast to disulfide bridges, these metal binding sites are stable in the reducing intracellular milieu. It is remarkable that the conservation of the three cysteines and of the histidine residue in these enzymes is even stronger than the one of the catalytic triad because the 2A proteases belong to the chymotrypsin-like cysteine protease family (13).

In the present work, we have undertaken a spectroscopic investigation of the metal binding site of the HCV NS3 protease. We tried to asses, in solution, the coordination role of His[149]. To address this issue, we have investigated the NMR spectra of Zn- and Co(II)-substituted proteins. The nonexchangeable protons of the three histidine residues were assigned and investigated by their pH dependence and by the Co(II)-induced isotropic shift of the signals. The signals arising from His[57] and His[110] show a noncooperative pH profile and a peak resonance shift characteristic of a fast exchange process.

The pH dependence profile of His[149] has revealed some peculiarities: it titrates with pK 5.9, lower than the pK of a protein histidine side chain exposed to the bulk water (pK = ~6.2) (33) but still significantly higher than that of a directly coordinating imidazoyl moiety (pK < 5.3) (51). Furthermore, the pH titration profile is consistent with the cooperative association of at least two protons (Hill coefficient, $n = 2.1$) in the

histidine protonation process. At acidic pH, the dominant species is expected to be the protonated His[149], corresponding to peak $C^+$. This species is unlikely to bind the metal ion due to electrostatic repulsion of the positive charges of the metal and the protonated His residue. The difference in line width between the unprotonated and the protonated forms of His[149] points toward a more rigid conformation for the former, whereas most probably, the latter is freely rotating (52). The experimental kinetic exchange parameters for the process in equation 3 allow us to calculate a $k_{-1} = 9.1 \pm 3.1$ s$^{-1}$. This value is similar to one recently published for the conformational switch of the histidines ligating the zinc in the HIV-1 integrase N-terminal protein (53), although in this case both states involved were ordered. Our experimental evidence suggests that the imidazole moiety of His[149] modulates the accessibility of the zinc ion, allowing an "open" and a "closed" conformation in the protonated and unprotonated states, respectively. This switch mechanism also parallels the observation in one published crystallographic structure (16), in which His[149] is postulated to participate in metal coordination in only two of the three molecules in the asymmetric unit, whereas in the third, the imidazoyl side chain of His[149] moves away.

NOESY and ROESY spectra provided evidence for an NOE involving the $\epsilon$1H nonexchangeable hydrogen and either a water molecule or an exchangeable group of the protein resonating at water frequency (in this case, the only candidates are the $\epsilon$2NH and the $\delta$1NH of the same residue). The $^{15}$N peak positions in the HSQC experiment are compatible with the $\delta$1N being in a $\beta$ state and the $\epsilon$2N in an $\alpha$ state (Table I). This tautomeric state is coherent with what can be inferred from the x-ray structures (16, 17), in which the nitrogen involved in metal chelation is the $\delta$1N, which is therefore not protonated. In our case, the $\beta$ state nitrogen corresponds to the $\delta$1N and its chemical shift falls in the range described by literature data (Table I). In contrast, the $\alpha$ state nitrogen ($\epsilon$2NH) is shifted downfield with respect to reference values. This, according to Bachovchin (41), would be an indication of its involvement, as donor, in a strong hydrogen bond. However, because nitrogen is a very sensitive nucleus, in proteins, other factors can also alter its chemical shift to the same extent of an H-bond (44), *e. g.* an indirect influence of the involvement of the $\delta$1N in metal binding on the electronic distribution in the coordinated imidazol ring.

The simplest explanation for this experimental evidence is the assumption that His[149] is ligated to the metal using the $\delta$1N through an $OH^-$. This model accounts both for the intermediate pK value and for the stoichiometry necessary to fit the titration data. In fact, the simultaneous association of two protons is required to protonate both the histidyl imidazole and the bridging hydroxyl. We also observed the NOE expected if such an oxhydril group was actually present. However, we are well aware that by itself this observation is not unambiguous. In line with crystal structure data, our spectroscopic findings are compatible with the chelation of zinc occurring via the $\delta$1N. This is not the usually preferred situation for histidine residues, occurring in only 2 cases out of the 14 described for which the $^{15}$N chemical shifts are available (46–48).

An alternative model has to invoke the presence of a hydrogen bond acceptor in the vicinity of the $\epsilon$2NH, the pK of which, by chance, should match that of His[149]. In fact, only a cooperative protonation of the $\delta$1N and of the hydrogen bond acceptor group could justify the Hill coefficient of the pH titrations. The binding to zinc in this case should occur directly with the $\delta$1N. However, this second hypothesis appears less plausible. In fact, to account for the observed NOE we should admit an exchange

process of the $\epsilon 2NH$ with the bulk water. But if this hydrogen position is assumed to be in fast exchange with water (to the extent of not being directly observable) it is unlikely to be, at the same time, involved in an interaction with a hydrogen bond acceptor group of the enzyme. Another objection comes from the crystallographic structures. In fact, it is well established that the zinc binding site is on the surface of the protein and that in the vicinity of His[149] there are no candidate groups that could act as hydrogen bond acceptors or could be protonated with a pK reasonably close to the observed one (the carbonyl groups of the peptidic bonds are the only groups below 5 Å distance that could be involved).

It is presently not clear whether the zinc ion present in NS3 is the same zinc ion shown to be essential for the NS2-NS3 autoprotease activity. If so, it could be speculated that a facile movement of His[149], leading to exposure of the coordinated metal, might play some role in the so far not completely understood mechanism of proteolysis of the NS2-NS3 junction. In this context, it is interesting to notice that the sequence pattern of the metal binding residues of the picornavirus 2A proteases (Cys-*X*-Cys/Cys-*X*-His) differs from the pattern of HCV-related viruses (Cys-*X*-Cys/Cys-*X*-*X*-*X*-His). The insertion of two additional residues between Cys[145] and His[149] may well contribute to the observed conformational flexibility of His[149] and could be related to a putative role of the metal, in HCV-related viruses, in the processing of the NS2-NS3 site.

The differences in the spacing of the metal-coordinating residues in HCV NS3 and in picornavirus 2A proteases could also explain the different effects of mutagenesis of His[149] and its corresponding residue in 2A. In fact, biochemical characterization of the H149A mutation in the NS3 protease (14, 15) has pointed out that the removal of the imidazoyl moiety only leads to minor effects on enzyme activity but causes the protein to accumulate in the insoluble fraction when expressed in *E. coli* (15). On the other hand, mutagenesis studies on 2A proteases from rhinovirus (53) and poliovirus (54) have revealed the essential role of the residue corresponding to His[149] of HCV NS3 for the catalytic activity of these enzymes. These results, indeed, are rather qualitative because they have been obtained in *in vitro* translation experiments. We therefore decided to quantitatively characterize the purified H149A mutant protein. Our results indicate that the H149A mutation has a major impact on the productive incorporation of zinc ions into the metal binding site during folding of the protein. In the mutant, this process was shown to occur on a time scale similar to misfolding, leading to only partial recovery of active protein. In contrast, only a 2-fold-increased susceptibility to EDTA-inactivation was observed, suggesting that, once incorporated, the metal is bound with a similar affinity in wild type and H149A mutant proteins. The mutation further affected both specific activity and affinity of the protein for its co-factor. Experimental difficulties did not allow the introduction of Co(II) into the metal binding site or concentration of the mutant protein, thus impeding a spectroscopic characterization of the conformational consequences of the H149A mutation. The activity data, however, suggest that changes in the coordination geometry resulting from the H149A mutation may have significant effects on the protein structure.

The relationship between metal coordination geometry and enzymatic activity is further highlighted by the spectral changes that accompany the complexation of the NS3 protease with its NS4A co-factor peptide. In fact, electronic absorption and circular dichroism spectra demonstrated ligand-metal rearrangements of the Co(II)-substituted NS3 protease induced by binding of the co-factor, whereas no significant effects could be detected upon substrate binding to the active site. Move-

ments of cysteine ligands seemed to play a crucial role in rearrangements of the coordination geometry, as pointed out by the changes in the circular dichroism of the S→Co(II) transitions. Activation of the HCV NS3 protease by its co-factor NS4A involves both a structural reorganization of the N-terminal domain and a rearrangement of the catalytic triad (16, 17). Our data show that at least some of these conformational transitions involve also rearrangements of the metal binding site. In fact, it may be argued that the metal coordination sphere is endowed with an intrinsic flexibility, necessary to respond to the conformational rearrangements of the protease during its complexation with the NS4A cofactor. Indeed, His[149] seems to play a major role in this conformational flexibility, as in the open metal coordination, where the His[149] is not ligating, and in the H149A mutant protease, the NS3-NS4A complex stability is reduced.

NS4A binds to the N-terminal region of NS3 (55). Interestingly, Mori *et al.* have shown that deletions in this region lead to an enhanced susceptibility toward inactivation of the NS3 protease by EDTA (56), which is suggestive of an NS4A-dependent tightening of metal-ligand bonds and in line with our findings of cofactor-induced conformational changes of the metal coordination.

The relationship between coordination geometry and enzymatic activity suggests that compounds having the capability of perturbing the native zinc coordination in the NS3 protease might modulate the enzymatic activity. We have shown, as a proof of principle, that $CN^-$ has such an effect. $CN^-$ was shown to bind to the metal in the Co(II)-substituted enzyme, inducing spectroscopically detectable changes in the coordination geometry that titrated with the same apparent $K_d$ as the activation process. Furthermore, $CN^-$ was able to enhance also the activity of the native, zinc-containing protease.

Perturbations of the ligand field spectrum of Co(II) substituted proteins by externally provided ligands have been used to probe the presence of transient coordination moieties such as water molecules (57, 58). Our observations on $CN^-$ and $N_3^-$ effects are in line with these reports, suggesting a distorted tetrahedral site, at which the water ligand has been replaced by the added anions.

In the light of the described correlation between metal coordination geometry and enzymatic activity, it is tempting to speculate that compounds capable of perturbing the metal coordination geometry of the NS3 protease might either activate or inhibit the enzyme, depending on the nature of the induced perturbation, thus revealing a new potential mechanism for NS3 protease inhibitors.

REFERENCES

1. Choo, Q.-L., Kuo, G., Weiner, A. J., Bradley, L. R. D. W., and Houghton, M. (1989) *Science* **244,** 359–362
2. Kuo, G., Choo, Q.-L., Alter, H. J., Gitnick, G. L., Redecker, A. G., Purcell, R. H., Myamura, T., Dienstag, J. L., Alter, M. J., Syevens, C. E., Tagtmeier, G. E., Bonino, F., Colombo, M., Lee, W.-S., Kuo, C., Berger, K., Shister, J. R., Overby, L. R., Bradley, D. W., and Houghton, M. (1989) *Science* **244,** 362–364
3. Kato, M., Hijikata, M., Ootsuyama, Y., Nakagawa, M., Ohkoshi, S., Sugimura, T., and Shimotohno, K. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87,** 9524–9528
4. Hijikata, M., Kato, N., Ootsuyama, Y., Nagakawa, M., and Shimotohno, K. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **90,** 5547–5551
5. Pieroni, L., Santolini, E., Fipaldini, C., Pacini, L., Migliaccio, G., and La Monica, N. (1997) *J. Virol.* **71,** 6373–6380

6. Tomei, L., Failla, C., Santolini, E., De Francesco, R., and La Monica, N. (1993) *J. Virol.* **67,** 4017–4026
7. Failla, C., Tomei, L., and De Francesco, R. (1994) *J. Virol.* **68,** 3753–3760
8. Bartenschlager, R., Ahlborn-Laake, L., Mous, J., and Jacobsen, H. (1994) *J. Virol.* **68,** 5045–5055
9. Lin, C., Thomson, J. A., and Rice, C. M. (1995) *J. Virol.* **69,** 4373–4380
10. Tomei, L., Failla, C., Vitale, R. L., Bianchi, E., and De Francesco, R. (1996) *J. Gen. Virol.* **77,** 1065–1070
11. Bianchi, E., Urbani, A., Biasiol, G., Brunetti, M., Pessi, A., De Francesco, R., and Steinkühler, C. (1997) *Biochemistry* **25,** 7890–7897
12. Landro, J. A., Raybuck, S. A., Luong, Y. P. C., O'Malley, E. T., Haberson, S., Morgenstern, K. A., Rao, G., and Livingston, D. (1997) *Biochemistry* **36,** 9340–9348
13. De Francesco, R., Urbani, A., Nardi, M. C., Tomei, L., Steinkühler, C., and Tramontano, A. (1996) *Biochemistry* **35,** 13282–13287
14. Stempniak, M., Hostomska, Z., Nodes, B. R., and Hostomsky, Z. (1997) *J. Virol.* **71,** 2881–2886
15. Hijikata, M., Mizushima, H., Akagi, T., Mori, S., Kakiuchi, N., Kato, N., Tanaka, T., Kimura, K., and Shimotohno, K. (1993) *J. Virol.* **67,** 4665–4675
16. Love, R. A., Parge, H. E., Wickersham, J. A., Hostomsky, Z., Habuka, N., Moomaw. E. W., Adachi, T., and Hostomska, Z. (1996) *Cell* **87,** 331–342
17. Kim, J. L., Morgenstern, K. A., Lin, C., Fox, T., Dwyer, M. D., Landro, J. A., Chambers, S. P., Markland, W., Lepre, C. A., O'Malley, E. T., Harbeson, S. L., Rice, C. M., Murcko, M. A., Caron, P. R., and Thomson, J. A. (1996) *Cell* **87,** 343–355
18. Steinkühler, C., Urbani, A., Tomei, L., Biasiol, G., Sardana, M., Bianchi, E., Pessi, A., and De Francesco, R. (1996a) *J. Virol.* **70,** 6694–6700
19. Balacco, G. (1994) *J. Chem. Inf. Comput. Sci.* **34,** 1235–1241
20. Lian, L. Y., and Roberts, G. C. K. (1993) in *NMR of Macromolecules* (Roberts, G. C. K., ed) pp.153–182, IRL Press, Oxford
21. Delaglio, F., Grzesiek, S.,Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR* **6,** 277–293
22. Johnson, B., and Blevins, R. A. (1994) *J. Biomol. NMR* **4,** 603–614
23. Clore, G. M., Omichinski, J. G., and Gronenborn, A. M. (1991) *J. Am. Chem. Soc.* **113,** 4350–4351
24. Urbani, A., Bianchi, E., Narjes, F., Tramontano, A., De Francesco, R., Steinkühler, C., and Pessi, A. (1997) *J. Biol. Chem.* **272,** 9204–9209
25. Hiromi, I. (1979) *Kinetics of Fast Enzyme Reactions*, pp. 99–101, John Wiley & Sons, Inc.
26. Taliani, M., Bianchi, E., Narjes, F., Fossatelli, M., Urbani, A., Steinkühler, C., De Francesco, R., and Pessi, A. (1996) *Anal. Biochem.* **240,** 60–67
27. Fersht, A. (1985) *Enzyme Structure and Mechanism*, 2nd Ed., pp. 143–147, W. H. Freeman and Co., New York
28. Maret, W., and Vallee, B. L. (1993) *Methods Enzymol.* **226,** 52–71
29. Bertini, I., and Luchinat, C. (1984) *Adv. Inorg. Biochem.* **6,** 71–111
30. Bertini, I., Turano, P., and Vila, A. J. (1993) *Chem. Rev.* **93,** 2833–2932
31. Bertini, I., Canti, G., Luchinat, C., and Mani, F. (1981) *J. Am. Chem. Soc.* **103,** 7784–7788
32. Vila, A. J., and Fernandez, C. O. (1996) *J. Am. Chem. Soc.* **118,** 7291–7298
33. Blomberg, F., Maurer, W., and Rüterjans, H. (1977) *J. Am. Chem. Soc.* **99,** 8149–8159
34. Grzesiek, S., and Bax, A. (1993) *J. Am. Chem. Soc.* **115,** 12593–12594
35. Otting, G., and Wüthrich, K. (1989) *J. Am. Chem. Soc.* **111,** 1871–1875
36. Witanowski, M., Stefaniak, L., Januszewski, H., Grabowski, Z., and Webb, G. A. (1972) *Tetrahedron* **28,** 637–653
37. Bachovchin, W. W., and Roberts, J. D. (1978) *J. Am. Chem. Soc.* **100,** 8041–8047
38. Schuster, I. I., and Roberts, J. D. (1979) *J. Org. Chem.* **44,** 3864–3867
39. Roberts, J. D., Yu, C., Flanagan, C., and Birdseye, T. R. (1982) *J. Am. Chem. Soc.* **104,** 3945–3949
40. Munowitz, M., Bachovchin, W. W., Herzfeld, J., Dobson, C. M., and Griffin, R. G. (1982) *J. Am. Chem. Soc.* **104,** 1192–1196
41. Bachovchin, W. W. (1986) *Biochemistry* **25,** 7751–7759.
42. Bachovchin, W. W., Wong, W. L. Farr-Jones, S., Shenvi, A. B., and Kettner, C. A. (1988) *Biochemistry* **27,** 7689–7697
43. van Dijk, A. A., de Lange, L. C. M., Bachovchin, W. W., and Robillard, G. T. (1990) *Biochemistry* **29,** 8164–8171
44. van Dijk, A. A., Scheek, R. M., Dijkstra, K., Wolters, G. K., and Robillard, G. T. (1992) *Biochemistry* **31,** 9063–9072
45. Pelton, J. G., Torchia, D. A., Meadow, N. D., and Roseman, S. (1993) *Protein Sci.* **2,** 543–558
46. Gooley, P. R., Johnson, B. A., Marcy, A. I., Cuca, G. C., Salowe, S. P., Hagmann, W. K., Esser, C. K., and Springer, J. P. (1993) *Biochemistry* **32,** 13098–13108
47. Schmiedeskamp, M., Rajagopal, P., and Klevit, R. E. (1997) *Protein Sci.* **6,** 1835–1848
48. Mengli, C., Zheng, M. C., Caffrey, M., Craigie, R., Clore, G. M., and Gronenborn, A. M. (1997) *Nat. Struct. Biol.* **7,** 567–577
49. Holmquist, B. (1986) *Methods Enzymol.* **130,** 270–290
50. Lesk, A., and Fordham, W. D. (1996) *J. Mol. Biol.* **258,** 501–537
51. Krizk, B. A., Aman, B. T., Kilofoil, V. J., Merckle, D. L., Berg, J. M. (1991) *J. Am. Chem. Soc.* **113,** 4518–4523
52. Clore, G. M., and Gronenborn, A. M. (1982) *Biochemistry* **21,** 4048–4055
53. Sommergruber, W., Zorn, M., Blaas, D., Fessl, F., Volkmann, P., Maurer-Fogy, I., Pallai, P., Merluzzi, V., Matteo, M., Skern, T., and Kuechler, E. (1989) *Virology* **169,** 68–77
54. Shuyuarn, F., and Lloyd, R. E. (1992) *Virology* **186,** 725–735
55. Failla, C., Tomei, L., and De Francesco, R. (1995) *J. Virol.* **69,** 1769–1777
56. Mori, A., Yuasa, S, Yamada, K., Nagami, Y., and Miyamura, T. (1997) *Biochem. Biophys. Res. Commun.* **231,** 738–742
57. Merkle, D. L., Schimidt, M. H., and Berg, J. M. (1991) *J. Am. Chem. Soc.* **113,** 5450–5451
58. Bertini, I., Canti, G., Luchinat, C., and Scozzafava, A. (1978) *J. Am. Chem. Soc.* **110,** 4873–4877