A Zinc Binding Site in Viral Serine Proteinases

Raffaele De Francesco,* Andrea Urbani, Maria Chiara Nardi, Licia Tomei, Christian Steinkühler, and Anna Tramontano*

Istitituto di Ricerche di Biologia Molecolare (IRBM) P. Angeletti, Via Pontina Km 30.600, 00040 Pomezia (Rome), Italy

Received July 8, 1996[⊗]

ABSTRACT: The NS3 protein of hepatitis C virus contains a chymotrypsin-like serine proteinase domain. We built a homology model of this domain which predicts the presence of a tetradentate metal binding site formed by three cysteines and one histidine. These residues are strictly conserved in all known hepatitis C viral genotypes as well as in other recently discovered related hepatitis viruses. We show that the hepatitis C virus enzyme does indeed contain a Zn^{2+} ion with S_3N ligation and that the metal is required for structural integrity and activity of the enzyme. Strikingly, the residues forming the metal binding site are also conserved in the chymotrypsin-like 2A cysteine proteinases of picornaviruses. Remarkably, in these highly variable viral genomes the metal binding site is more conserved than the catalytic residues and thus allows us to define a novel class of zinc binding chymotrypsin-like proteinases and to identify a new attractive target for antiviral therapy.

Hepatitis C virus (HCV),[1] a member of the *Flaviviridae* family, is now recognized as the principal etiologic agent of parenterally transmitted non-A, non-B hepatitis (NANB-H) (Choo et al., 1989; Kuo et al., 1989). In addition, the development of cirrhosis and hepatocellular carcinoma in individuals with NANB-H suggests a causative linkage between HCV infection and hepatocarcinogenesis (Chien et al., 1992). Neither an effective therapy for HCV-associated chronic hepatitis nor a vaccine against viral infection has to date been developed.

The HCV virion has a positive-strand RNA genome of about 9.5 kilobases (kb) (Kato et al., 1990; Choo et al., 1991; Takamizawa et al., 1991). The genomic RNA contains a single open reading frame which encodes a polyprotein of about 3000 amino acids. The putative structural protein "core" and the envelope glycoproteins E1 and E2 are released from the N-terminal part of the HCV polyprotein by cellular peptidases (Hijikata et al., 1991), whereas a virus-encoded metalloprotease is responsible for the cleavage between NS2 and NS3 (Grakoui et al., 1993b; Hijikata et al., 1993). The proteolytic cleavage at the NS3/NS4A, NS4A/NS4B, NS4B/NS5A, and NS5A/NS5B junctions is mediated by a virus-encoded serine proteinase, contained within the N-terminal 180 amino acids of NS3, that is homologous to the small cellular proteinases of the chymotrypsin superfamily (Bartenschlager et al., 1993; Eckart et al., 1993; Grakoui et al., 1993a; Tomei et al., 1993; Manabe et al., 1994). In addition to the NS3 serine proteinase, the NS4A protein is required for cleavage at the NS3/NS4A and NS4B/NS5A sites, and it increases the efficiency of cleavage at the NS5A/NS5B and NS4A/NS4B junctions (Bartenschlager et al., 1994; Failla et al., 1994; Lin et al., 1994; Tanji et al., 1995). NS4A is a 54-residue membrane protein that has been demonstrated to bind to the N-terminus of NS3 via a central hydrophobic region, thereby enhancing the proteolytic activity of NS3 on all cleavage sites. It has recently been shown that all the effects of NS4A can be efficiently mimicked by peptides encompassing residues 21−34 of NS4A (Lin et al., 1995; Shimizu et al., 1996; Tomei et al., 1996).

The three-dimensional structure of the HCV NS3 proteinase has not yet been determined. In addition, its sequence identity with other serine proteinases of known structure is very low (less than 15%) (Pizzi et al., 1993), preventing the construction of an accurate sequence alignment and therefore of a detailed and complete homology model. However, we made use of the numerous available serine proteinase structures to align the sequences of their conserved regions and obtained a sequence profile to which we manually aligned the NS3 proteinase sequence (Pizzi et al., 1993; Failla et al., 1995). This allowed us to construct a partial model and to predict that the protein substrates recognized by the NS3 serine proteinase should have a cysteine residue in position P1. This hypothesis was subsequently proven by N-terminal sequencing of the proteinase products (Pizzi et al., 1993). More recently, we used the homology model as a guide in the redesign of the enzyme's specificity. We replaced the residues predicted to form its specificity pocket with the corresponding ones of *Streptomyces griseus* proteinase B (Failla et al., 1995) and obtained a mutant proteinase that, similarly to the bacterial protein and unlike the wild-type enzyme, was able to cleave a substrate containing a phenylalanine in the P1 position, thus increasing our confidence in the correctness of the model.

Here we show that our homology model predicts that the serine protease domain of NS3 contains a tetradentate Zn^{2+} binding site formed by three cysteines and one histidine, conserved in all HCV genotypes and in related hepatitis viruses, and demonstrate that the purified NS3 protease

* Corresponding authors. Tel: +39 691093207. Fax: +39 691093654. E-mail: Tramontano@IRBM.It.

[⊗] Abstract published in *Advance ACS Abstracts,* October 1, 1996.

[1] Abbreviations: CHAPS, 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate; DMSO, dimethyl sulfoxide; DTT, dithiothreitol; GBV-A, GB virus A; GBV-B, GB virus B; GBV-C/HGV, GB virus C or hepatitis G virus; HCV, hepatitis C virus; HPLC, high-pressure liquid chromatography; IPTG, isopropyl 1-thio-β-D-galactopyranoside; LB, Luria−Bertani medium; NANB-H, non-A non-B hepatitis; NS, nonstructural protein; RPC, reverse-phase chromatography; TFA, trifluoroacetic acid.

S0006-2960(96)01645-5 CCC: $12.00

domain does indeed bind equimolar amounts of Zn^{2+}. We also show that, although Zn^{2+} is not the catalytic center of the enzyme, this metal ion is required for the generation of a functional enzyme, *in vitro* as well as *in vivo*, possibly through the stabilization of the protein fold.

MATERIALS AND METHODS

Model Building and Data Base Search. All analyses were performed using the GCG sequence analysis package (Devereux et al., 1984) and the graphics program Insight (Dayringer et al., 1986). Figure 1b was produced using Molscript (Kraulis, 1991).

Purification of the NS3 Proteinase Domain. Escherichia coli BL21(DE3) cells were transformed with a plasmid containing the cDNA coding for the serine proteinase domain of NS3 (amino acids 1−180) under the control of the bacteriophage T7 gene 10 promoter. Details of plasmid construction will be published elsewhere. Expression was achieved by incubating cultures at an A_{600nm} of 0.7−0.9 with 0.4 mM isopropyl 1-thio-β-D-galactopyranoside (IPTG) for 3 h at 20 °C in LB medium. The harvested cells were resuspended in 25 mM sodium phosphate, pH 7.5, 0.5% CHAPS, 10% glycerol, 10 mM DTT, and 500 mM NaCl (10 mL for each liter of original cell suspension) and disrupted in a French pressure cell. Homogenates were clarified by centrifugation at 100000*g* for 1 h, and nucleic acids were removed by precipitation with 0.1% poly(ethylenimine). Supernatants were loaded on a HiLoad 16/10 SP Sepharose high-performance column (Pharmacia), equilibrated with 50 mM sodium phosphate, pH 7.5, 5% glycerol, 0.1% CHAPS, and 3 mM DTT (buffer A). The proteinase was eluted by applying a 0−0.6 M NaCl gradient in buffer A. NS3-containing fractions were pooled and concentrated up to a protein concentration of 6−10 mg/mL, using an Amicon-stirred ultrafiltration cell equipped with a YM-10 membrane, and loaded on a HR 26/60 HiLoad Superdex 75 column (Pharmacia) equilibrated with buffer A. NS3-containing fractions were pooled and further purified on a HR 5/5 Mono S column (Pharmacia) equilibrated with buffer A and operating at a flow rate of 1 mL/min. The proteinase was eluted in a pure form from this column by applying a linear 0−0.6 M NaCl gradient in buffer A. Protein concentration was estimated by quantitative amino acid analysis. Purity of the enzyme was checked on silver-stained SDS−polyacrylamide gels and by HPLC using a reversed-phase Vydac C4 column (4.6 × 250 mm, 5 μm, 300 Å). Eluents were H_2O/0.1% TFA (A) and acetonitrile/0.1% TFA (B). A linear gradient from 3% to 95% B in 60 min was used.

Synthetic Peptides and HPLC Assay. Our standard substrate was a 13 amino acid synthetic peptide derived from the cleavage sequence of the NS4A−NS4B junction (DEEMEC-ASHLPYK). As a proteinase cofactor, we used a 14-mer peptide corresponding to the central hydrophobic core of the NS4A protein spanning residues 21−34 (pep4A$_{21-34}$: GSVVIVGRIILSGR).

Peptides were synthesized by solid-phase synthesis based on Fmoc chemistry. After cleavage and deprotection the crude peptides were purified by HPLC to >98% purity. The identity of peptides was checked by mass spectrometry. Concentration of stock solutions of peptides, prepared in DMSO and kept at −80 °C until use, was determined by quantitative amino acid analysis performed on HCl-hydrolyzed samples.

Cleavage assays were performed using 300 nM enzyme in 30 μL of 50 mM Tris-HCl, pH 7.5, 50% glycerol, 2% CHAPS, 3 μM NS4A peptide, and appropriate amounts of substrate and at 22 °C. The activity buffers used throughout this work were treated with Chelex-100 resin (2.5 g/L) in order to minimize the concentration of adventitious metal ions. Cleavage of peptide substrates was determined by HPLC using a Merck-Hitachi chromatograph. Samples were injected on a Lichrosphere C18 reverse-phase cartridge column (4 × 125 mm, 5 μm, Merck), and fragments were separated using a 3−100% acetonitrile gradient at 2%/min. Peak detection was accomplished by monitoring both the absorbance at 220 nm and tyrosine fluorescence (λ_{ex} = 260 nm, λ_{em} = 305 nm).

Synthesis of Co^{2+}- and Cd^{2+}-NS3 Proteinase. Co^{2+}- and Cd^{2+}-containing recombinant NS3 proteinases were biosynthetically prepared in minimal medium containing 100 mM potassium phosphate, pH 7.0, 0.5 mM $MgSO_4$, 0.5 mM $CaCl_2$, 13 μM $FeSO_4$, 7 μM thiamin, and 6 μM biotin. Glucose (4 g/L) and $(NH_4)_2SO_4$ (1 g/L) were used as sources of carbon and nitrogen, respectively. To reduce the amount of Zn^{2+} in the medium, the phosphate buffer was passed through a Chelex-100 column. Zn^{2+} supplementation was achieved by the addition of $ZnCl_2$ to a final concentration of 50 μM to the Zn-depleted medium. Co^{2+}- and Cd^{2+}-substituted NS3 proteinases were synthesized using the same procedure, except that either 50 μM $CoCl_2$ or 50 μM $CdCl_2$ was added 20 min before the addition of IPTG. Co^{2+}- and Cd^{2+}-substituted NS3 proteinases were purified as described above for the Zn^{2+}-containing enzyme, except that all the buffers were treated with Chelex-100 resin (2.5 g/L) and DTT was omitted. In addition, the O_2 concentration in all buffers used was reduced by purging with N_2.

Spectroscopic and Metal Analysis. Electronic spectra of Zn^{2+}-, Co^{2+}-, and Cd^{2+}-NS3 proteinases were acquired on a Varian Cary 3E dual-beam spectrophotometer.

The concentrations of Zn^{2+}, Co^{2+}, and Cd^{2+} in the metal-substituted proteins were determined by atomic absorption spectroscopy on a Perkin-Elmer Instrument Model 2100 atomic absorption spectrometer equipped with a graphite furnace. Glassware used for metal analysis was washed with 30% nitric acid and thoroughly washed with deionized water. NS3 proteinase (4 mg/mL) was dialyzed at least 16 h against a buffer containing 50 mM Tris-HCl, pH 7.5, 3 mM DTT, 10% glycerol, and 0.1% CHAPS. Chelex-100 resin (2.5 g/L) was kept in suspension in the dialysis buffer in order to prevent contamination by adventitious metals. The protein was subsequently hydrolyzed with nitric acid and used for metal determination. Standardized Zn^{2+}, Co^{2+}, and Cd^{2+} solutions were purchased from Merck.

Preparation of Apo-NS3 Proteinase. Purified NS3 serine proteinase (1.7 mg) was denatured by the addition of TFA to a final concentration of 1%. The denatured protein was subsequently purified on a Resource RPC 3 mL column (Pharmacia) using a 0−85% acetonitrile gradient in the presence of 0.1% TFA. The flow rate was 2 mL/min, and the gradient volume was 45 mL.

RESULTS AND DISCUSSION

Prediction of a Metal Binding Site in the NS3 Proteinase. Three novel flavivirus-like hepatitis agents have recently been

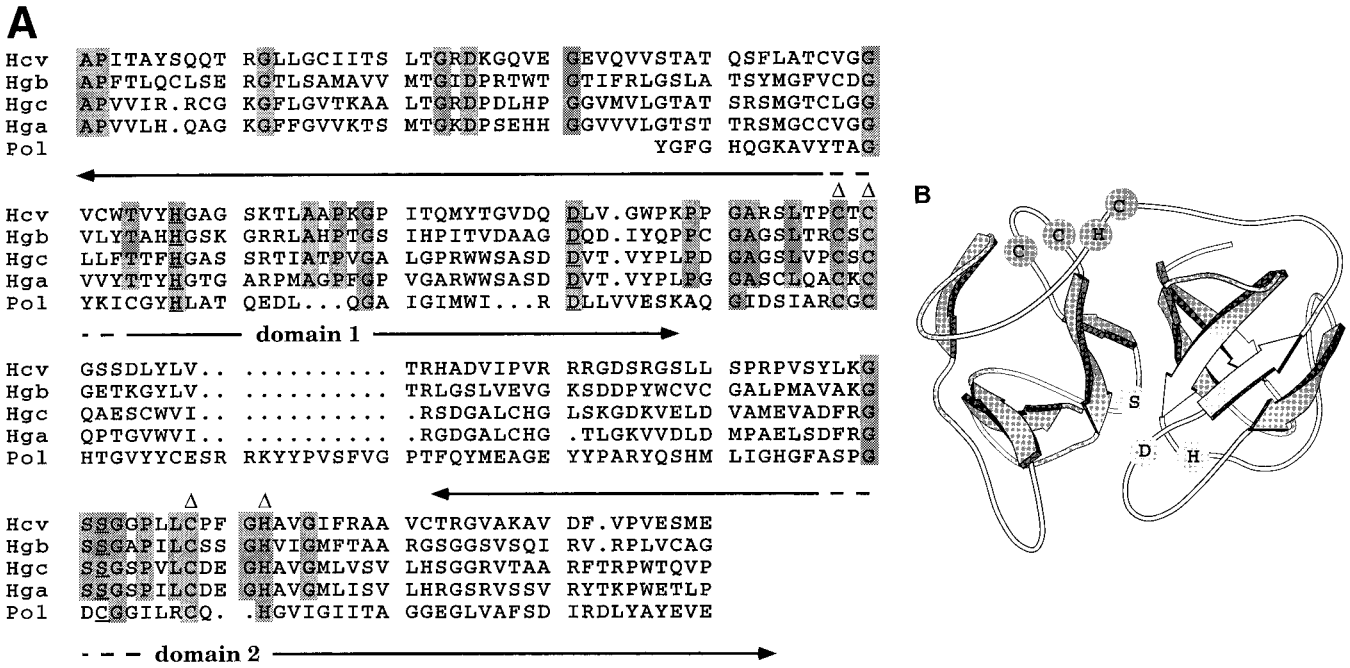

FIGURE 1: (A) Amino acid sequence comparison of HCV and HCV-related NS3 proteinase domains (Hcv, Hgb, Hgc, Hga) with the sequence of poliovirus 2A proteinase (Pol). Conserved residues in hepatitis C and related viruses are shaded, catalytic residues underlined, and metal binding residues indicated by △. (B) Schematic model of the NS3 proteinase domain. The positions in the structure of the residues involved in metal binding (dark gray) and catalytic residues (light gray) are shown.

described (Zuckerman, 1996). These new viruses, termed GB virus A (GBV-A), GB virus B (GBV-B), and GB virus C or hepatitis G virus (GBV-C/HGV), are only distantly related to HCV, with which they share a common polyprotein organization (Muerhoff et al., 1995; Leary et al., 1996). We have now identified strictly conserved residues of the NS3 serine proteinase domain from the alignment of different HCV genotypes and of GBV-A, GBV-B, and GBV-C/HGV and mapped their position in our model. Conserved residues include, besides the residues of the active site, several glycines and prolines most likely involved in the stabilization of the protein fold and three cysteines and one histidine (Figure 1A). In our model, these latter four residues are located on the side of the molecule opposite from the active site, cluster in space, and are likely to form a tetrahedral metal binding site (Figure 1B). Notably, previous mutagenesis work has indicated that the three conserved cysteines cannot be mutated without compromising the proteolytic activity of NS3 (Hijikata et al., 1993).

A data base search for similar sequence patterns in proteinases revealed that a similar arrangement of ligands exists in the 2A proteinases of picornaviruses (a chymotrypsin-like cysteine proteinase) where it is also strictly conserved (Yu & Lloyd, 1992). In these enzymes the metal is required for the formation of an active enzyme (Sommerburger et al., 1994) and is an essential component of the native structure (Voss et al., 1995). The metal binding residues of the picornavirus proteinase are located in topological positions similar to those in the HCV NS3 (Figure 1A); the sequence patterns (Cys-X-Cys/Cys-X-His in poliovirus and Cys-X-Cys/Cys-X-X-X-His in HCV NS3) are similar and could not be found in any other cellular or viral serine proteinase, including those of flavi- and pestiviruses, which are viruses closely related to HCV.

Zinc Is Required for Enzymatic Activity and Folding of the NS3 Serine Proteinase. We expressed the serine proteinase domain of NS3, from amino acid 1 to 180 in *E. coli.* The recombinant protein was recovered in the soluble fraction and purified to homogeneity as described in Materials and Methods.

Table 1: Metal Content of Zn^{2+}-, Apo-, Co^{2+}-, and Cd^{2+}-NS3 Proteinases Determined by Atomic Absorption Spectroscopy[a]

protein	Zn (mol/mol)	Co (mol/mol)	Cd (mol/mol)
Zn^{2+}-NS3	1.09	nd	nd
apo-NS3	0.02	nd	nd
Co^{2+}-NS3	0.19	0.90	nd
Cd^{2+}-NS3	0.09	nd	1.15

[a] The concentrations of Zn^{2+}, Co^{2+}, and Cd^{2+} in the metal-substituted protein and in the apoprotein were determined by atomic absorption spectroscopy on a atomic absorption spectrometer equipped with a graphite furnace (see Materials and Methods for details). NS3 proteinase (4 mg/mL) was dialyzed at least 16 h against a buffer containing 50 mM Tris-HCl pH 7.5, 3 mM DTT, 10% glycerol, and 0.1% CHAPS. Chelex-100 resin (2.5 g/L) was kept in suspension in the dialysis buffer in order to prevent contamination by adventitious metals. The protein was subsequently hydrolyzed with nitric acid and used for metal determination; nd = not determined.

The purified NS3 proteinase domain was extensively dialyzed against buffers containing a chelating agent in order to remove any adventitiously bound metal ion and then analyzed for its metal content by atomic absorption spectroscopy. The content of zinc was found to be approximately 1 mol/mol of enzyme, suggesting the presence of a tightly bound zinc ion (Table 1).

To ascertain whether zinc is required for the activity of the NS3 serine proteinase, we first measured its proteolytic activity on a synthetic peptide substrate in the presence of increasing concentrations of either EDTA or 1,10-phenanthroline. These two compounds were found not to inhibit proteolysis by NS3 at concentrations ≤1 mM. Above this concentration, both EDTA and 1,10-phenanthroline could inhibit NS3 activity. However, a similar pattern of inhibition was obtained in control experiments using structural analogs

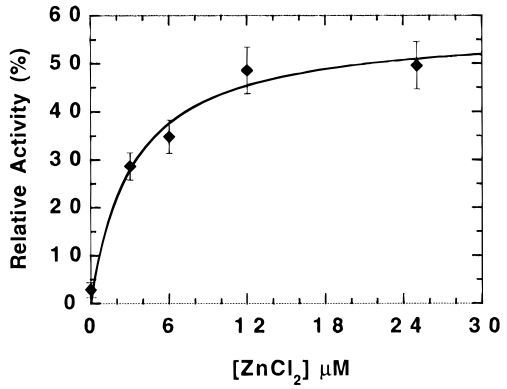

FIGURE 2: Effect of Zn^{2+} on the activity of the NS3 serine proteinase. NS3 proteinase apoenzyme was prepared by acid denaturation followed by reverse-phase chromatography as described in the Materials and Methods section. The apoprotein was diluted 200-fold to a final concentration of 600 nM in activity buffer containing the indicated concentration of $ZnCl_2$ and 10 mM DTT to prevent thiol oxidation. Following incubation at 22 °C for 1 h, the reaction was started by addition of the peptide substrate at a concentration of 40 μM. The reaction proceded for another hour prior to analysis.

of 1,10-phenanthroline that cannot chelate zinc ions (1,4-phenanthroline or 1,7-phenanthroline), and inhibition was not reverted by the presence of excess Zn^{2+} (data not shown). These results suggested that either zinc is not required for activity or it is tightly bound and could not be removed by treatment with chelating agents.

We set out to prepare a zinc-free protein and measure its biochemical activity in the absence and in the presence of metal. Again, the bound zinc could not be removed by dialysis against chelators at pH >7. Conversely, prolonged dialysis of the enzyme at pH <5 and in the presence of 10 mM EDTA caused a loss of zinc coupled with irreversible precipitation of the sample. Altogether, these observations suggest that zinc is tightly bound and that it may be essential for the structural integrity of the protein.

In order to facilitate the release of zinc, we denatured the enzyme by addition of 1% trifluoroacetic acid and then purified the apoprotein by reverse-phase chromatography. The zinc content of the resulting apoprotein was negligible (Table 1). We thus tested the enzymatic activity of the apoprotein after rapid refolding by dilution directly in the activity buffer in the presence and in the absence of zinc. As shown in Figure 2, reconstitution of an enzymatically active NS3 serine proteinase is dependent upon the concentration of Zn^{2+} in the activity buffer. Maximum reactivation was observed at a $ZnCl_2$ concentration of 25 μM, where recovery of the enzymatic activity was about 50% with respect to a zinc-containing proteinase diluted to the same final concentration in the same buffer. Less then 2% of the activity could instead be recovered in the absence of $ZnCl_2$ (Figure 2). The finding that the reconstitution of active enzyme is only partial, even in the presence of a molar excess of $ZnCl_2$, suggests that not all the enzyme is refolding properly during the dilution step.

The experiments described so far confirm that the NS3 serine proteinase domain does contain zinc and show that the zinc ion is required for the formation of an active enzyme.

The environment of the bound zinc within the NS3 proteinase can be investigated by replacing it with metals that, unlike zinc, are amenable to spectroscopic studies.

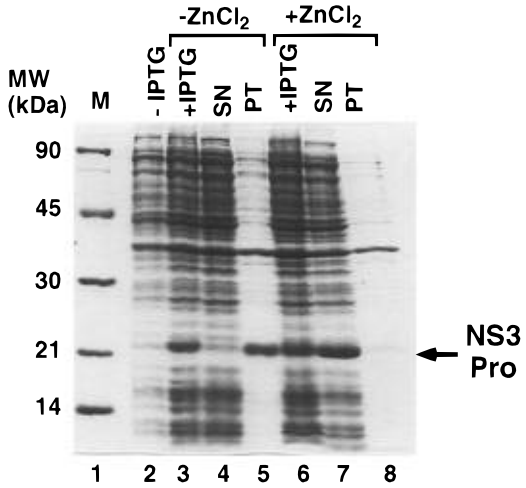

FIGURE 3: Effect of Zn^{2+} on the folding of the NS3 serine proteinase in *E. coli*. The production of NS3 proteinase was induced with IPTG. NS3 proteinase was obtained as an insoluble protein when produced by bacterial cells growing in zinc-depleted medium (lanes 3−5). Conversely, when the medium was supplemented with 50 μM $ZnCl_2$, the proteinase was completely soluble (lanes 6−8). SN: cytoplasmic fraction. PT: particulate fraction. Lane 1: molecular mass markers.

The tight binding of the constitutive zinc to the enzyme makes it difficult remove the metal and replace it with a different one *in vitro*. Therefore, we attempted to achieve metal substitution of Co^{2+} and Cd^{2+} for Zn^{2+} by *in vivo* incorporation. We grew *E. coli* transformed with the appropriate expression plasmid in Zn^{2+}-depleted minimal medium. The bacterial cells grew normally in this medium, and the expected amount of NS3 proteinase was produced after induction with IPTG (Figure 3, lane 2), but it accumulated in the insoluble fraction (Figure 3, lanes 3 and 4). Strikingly, when $ZnCl_2$ was added to the minimal medium, the NS3 proteinase remained in the cytoplasmic soluble fraction (Figure 3, lanes 5−7). This finding suggests that Zn^{2+} is required for proper folding of the NS3 proteinase in *E. coli* cells. Addition of $CoCl_2$ or $CdCl_2$ to the Zn-depleted media also resulted in the production of soluble enzyme in *E. coli* (data not shown), suggesting that, through biosynthetic incorporation, Co^{2+} or Cd^{2+} could replace Zn^{2+} in the metal binding site. NS3 proteinases purified from *E. coli* cells grown in the Co^{2+}- and Cd^{2+}-enriched media were analyzed by atomic absorption spectroscopy and found to contain stoichiometric amounts of Co^{2+} and Cd^{2+}, respectively (Table 1). The enzymatic activity of the Co^{2+}- and Cd^{2+}-substituted enzymes was comparable to that of the Zn^{2+}-containing enzyme (Table 2).

Properties of the Co^{2+}- and Cd^{2+}-Substituted NS3 Proteinases. The Co^{2+}- and Cd^{2+}-NS3 proteinases were characterized by electronic absorption spectroscopy. The Co^{2+}-NS3 proteinase showed a characteristic absorption spectrum in the visible region (Figure 4A) suggestive of a tetrahedral Co^{2+} complex (Maret & Vallee, 1993). Two major bands, with maxima at 640 and 685 nm and minor shoulders at 585 and 740 nm, are indicative of d−d transitions. The energy of these transitions and the molar extinction coefficients are typical of four-coordinate high-spin Co^{2+} complexes with distorted tetrahedral coordination geometry (Bertini & Luchinat, 1984). The energy of the d−d absorption envelope is consistent with mixed sulfur−nitrogen ligation. In addition, the centroid of the d−d envelope is suggestive of a Co^{2+}

Table 2: Enzymatic Activity of Zn^{2+}-, Co^{2+}-, and Cd^{2+}-NS3 Proteinases on a Synthetic Peptide Substrate[a]

metal	K_m (μM)	k_{cat} (min^{-1})	k_{cat}/K_m (M^{-1} s^{-1})
Zn^{2+}	58	2.4	690
Co^{2+}	98	2.0	340
Cd^{2+}	47	1.4	496

[a] Our standard substrate was a 13 amino acid synthetic peptide derived from the cleavage sequence of the NS4A−NS4B junction (DEEMEC-ASHLPYK). As a proteinase cofactor, we used a 14-mer peptide corresponding to the central hydrophobic core of the NS4A protein spanning residues 21−34 (pep4A_{21-34}: GSVVIVGRIILSGR). Cleavage assays were performed using 300 nM enzyme in a metal-free buffer containing 50 mM Tris-HCl, pH 7.5, 50% glycerol, 2% CHAPS, 3 μM NS4A peptide, and appropriate amounts of substrate and at 22 °C. Cleavage of peptide substrates was determined by HPLC as described in Materials and Methods. Data are mean values of two different experiments with standard errors <30%.

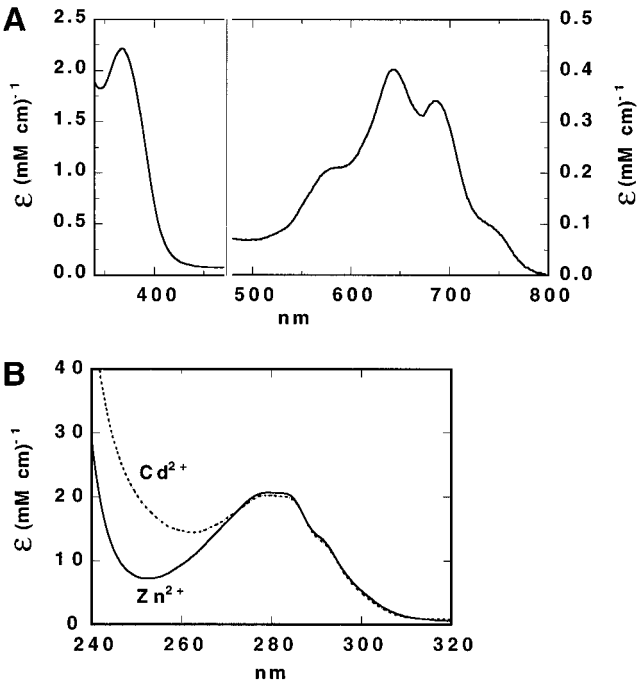

FIGURE 4: Electronic spectra of Co^{2+}-, Zn^{2+}-, and Cd^{2+}-substituted NS3 proteinases. Spectra were acquired at 15 °C. The buffer used consisted of 50 mM sodium phosphate, pH 7.5, 3 mM DTT, 10% glycerol, and 0.1% CHAPS. (A) Near-UV and visible absorption spectra of Co^{2+}-NS3 proteinase. The protein concentration was 125 μM. The d−d envelope centered near 650 nm is indicative of a Co^{2+} complex with S_3N ligation. A typical S → Co^{2+} charge-transfer band is present around 365 nm. (B) UV absorption spectra of Zn^{2+}- and Cd^{2+}-NS3 proteinases. The protein concentrations were 60 and 30 μM, respectively. The increased absorbance below 250 nm shown by the Cd^{2+}-NS3 proteinase is due to the S → Cd^{2+} charge-transfer band.

complex with S_3N ligation (Maret & Vallee, 1993). A typical S → Co^{2+} charge-transfer band was observed around 365 nm (Figure 4A), implying that the metal ion is in fact coordinated by thiolates. Consistently, the UV absorption spectrum of the Cd^{2+}-NS3 proteinase (Figure 4B) shows increased absorbance around 250 nm that is most likely due to the S → Cd^{2+} charge-transfer band (Fitzgerald & Coleman, 1991).

The spectroscopic analysis of the Co^{2+}- and Cd^{2+}-substituted NS3 proteinases is fully consistent with our homology model where the metal binding site within the NS3 proteinase is formed by three cysteine thiolates and one histidine side chain nitrogen.

CONCLUSIONS

In most Zn^{2+}-containing enzymes, the zinc acts as an electrophilic catalyst, often involved in the activation of water molecules, and is usually ligated by nitrogen and oxygen atoms and exposed to the solvent. However, a few enzymes bind zinc ions that have a purely structural role (Berg & Shi, 1996). In this latter case the metal ion is more often ligated by cysteine thiolates, as is the case for the NS3 proteinase domains of HCV, GBV-A, GBV-B, and HGV and for the 2A proteinase of picornaviruses (Voss et al., 1995). It is interesting to speculate on the role of this metal binding site in NS3 and 2A proteinases. The metal binding residues are located in the long loop connecting the two proteinase domains and in a hairpin loop in the second domain. Catalytic residues of chymotrypsin-like proteinases are also distributed between the two domains (Figure 1), and this imposes very stringent structural requirements on their relative orientation. Almost all chymotrypsin-like proteinases have disulfide bridges that involve cysteine residues adjacent to or nearby residues in the catalytic triad and that are believed to help in maintaining their relative position (Lesk & Fordham, 1996). The NS3 proteinase domains and the 2A proteinases do not have any disulfide bridges, and their metal binding sites could therefore be required to constrain the position of the two domains and, consequently, of the catalytic triad.

We also constructed four mutants of the NS3 proteinase domain, each bearing the substitution of one of the putative metal binding residues into alanine. In line with the notion that Zn^{2+} binding is required for proper folding of the NS3 proteinase (Figure 3), none of these mutants could be expressed in a soluble form in *E. coli* (data not shown).

The presence of a similar and functionally important metal binding site in the NS3 and 2A proteinases is striking from an evolutionary point of view, since 2A proteinases belong structurally to the serine proteinase family but utilize a cysteine and not a serine as an active site nucleophile (Bazan & Fletterick, 1988). Our observation consequently implies that the metal binding site is more conserved than the catalytic site (even more remarkable in the light of the hypervariability of viral genomes) and that the NS3 proteinase domains of hepatitis C-related viruses and the 2A proteinases of picornaviruses are members of a novel class of zinc binding chymotrypsin-like proteinases.

It is also important to stress that the strict conservation of this site in hepatitis C viral genotypes makes it a potentially very attractive target for antiviral therapy (Rice et al., 1993, 1995).

The experimental support for our structural prediction confirms that the general features of our model are sufficiently accurate to permit the prediction of relevant properties of the enzyme. This shows that, in certain instances, an approximate model can be sufficient to direct the experimental work to achieve a better understanding of the protein characteristics. The repeated success of our predictions in this system (Pizzi et al., 1993; Failla et al., 1995) emphasizes the importance of the synergy between theoretical analysis and experimental work especially in the frequent cases when only an approximate and often partial model of a protein can be constructed.

ACKNOWLEDGMENT

We thank Prof. Alessandro Desideri and Dr. Francesca Polizio for helpful discussions. We are particularly grateful to Dr. Angelo De Martino for performing atomic absorption spectroscopy and to Drs. Antonello Pessi and Elisabetta Bianchi for peptide synthesis.

REFERENCES

Bartenschlager, R., Ahlborn-Laake, L., Mous, J., & Jacobsen, H. (1993) *J. Virol. 68*, 1045−1055.

Bartenschlager, R., Ahlborn-Laake, L., Mous, J., & Jacobsen, H. (1994) *J. Virol. 68*, 5045−5055.

Bazan, F. J., & Fletterick, R. J. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 7872−7876.

Berg, J. M., & Shi, Y. (1996) *Science 271*, 1081−1085.

Bertini, I., & Luchinat, C. (1984) *Adv. Inorg. Biochem. 6*, 71−111.

Chien, D., Choo, Q.-L., Tabrizi, A., Kuo, C., McFarland, J., Berger, K., Lee, C., Shuster, J., Nguyen, T., Moyer, D., Tong, M. M., Furuta, S., Omata, M., Tegtmeyer, G., Alter, H., Schiff, E., Jeffers, L., Houghton, M., & Kuo, G. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 10011−10015.

Choo, Q.-L., Kuo, G., Weiner, A. J., Bradley, L. R. D. W., & Houghton, M. (1989) *Science 244*, 359−362.

Choo, Q.-L., Richman, K. H., Han, J. H., Berger, K., Lee, C., Dong, C., Gallegos, C., Coit, D., Medina-Selby, A., Barr, P. J., Weiner, A. J., Bradley, D. W., Kuo, G., & Houghton, M. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 2451−2455.

Dayringer, H. E., Tramontano, A., Sprang, S. R., & Fletterick, R. J. (1986) *J. Mol. Graphics 4*, 82−87.

Devereux, J., Haeberli, P., & Smithies, O. (1984) *Nucleic Acids Res. 12*, 387−395.

Eckart, M. R., Selby, M., Masiarz, F., Lee, C., Berger, K., Crawford, K., Kuo, C., Kuo, G., Houghton, M., & Choo, Q.-L. (1993) *Biochem. Biophys. Res. Commun. 192*, 399−406.

Failla, C., Tomei, L., & De Francesco, R. (1994) *J. Virol. 68*, 3753−3760.

Failla, C. M., Pizzi, E., De Francesco, R., & Tramontano, A. (1995) *Folding Des. 1*, 35−42.

Fitzgerald, D. W., & Coleman, J. E. (1991) *Biochemistry 30*, 5195−5201.

Grakoui, A., McCourt, D. W., Wychowski, C., Feinstone, S. M., & Rice, C. M. (1993a) *J. Virol. 67*, 2832−2843.

Grakoui, A., McCourt, D. W., Wychowski, C., Feinstone, S. M., & Rice, C. M. (1993b) *Proc. Natl. Acad. Sci. U.S.A. 90*, 10583−10587.

Hijikata, M., Kato, N., Ootsuyama, Y., Nagakawa, M., & Shimotomo, K. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 5547−5551.

Hijikata, M., Mizushima, H., Akagi, T., Mori, S., Kakiuchi, N., Kato, N., Tanaka, T., Kimura, K., & Shimotohno, K. (1993) *J.Virol. 67*, 4665−4675.

Kato, M., Hijikata, M., Ootsuyama, Y., Nakagawa, M., Ohkoshi, S., Sugimura, T., & Shimotohno, K. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 9524−9528.

Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.

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., & Houghton, M. (1989) *Science 244*, 362−364.

Leary, T. P., Muerhoff, A. S., Simons, J. N., Pilot-Matias, T. J., Erker, J. C., Chalmers, M. L., Schlauder, G. C., Dawson, G. J., Desai, S. M., & Mushahwar, I. K. (1996) *J. Med. Virol. 48*, 60−67.

Lesk, A. M., & Fordham, W. D. (1996) *J. Mol. Biol. 258*, 501−537.

Lin, C., Pragai, B. M., Grakoui, A., Xu, J., & Rice, C. M. (1994) *J. Virol. 68*, 8147−8157.

Lin, C., Thomson, J. A., & Rice, C. M. (1995) *J. Virol. 69*, 4373−4380.

Manabe, S., Fuke, I., Tanishita, O., Kaji, C., Gomi, Y., Yoshida, S., Mori, C., Takamizawa, A., Yohida, I., & Okayama, H. (1994) *Virology 198*, 636−644.

Maret, W., & Vallee, B. L. (1993) *Methods Enzymol. 226*, 52−71.

Muerhoff, A. S., Leary, T. P., Simons, J. N., Pilot-Matias, T. J., Dawson, G. J., Erker, J. C., Chalmers, M. L., Schlauder, G. C., Desai, S. M., & Mushahwar, I. K. (1995) *J. Virol. 69*, 5621−5630.

Pizzi, E., Tramontano, A., Tomei, L., La Monica, N., Failla, C., Sardana, M., Wood, T., & De Francesco, R. (1993) *Proc. Natl. Acad. Sci. U.S.A. 91*, 888−892.

Rice, W. G., Schaeffer, C. A., Harten, B., Villinger, F., South, T. L., Summers, M. F., Henderson, L. E., Bess, J. W. J., Arthur, L. O., McDougall, J. S., Orloff, S. L., Mendeleyev, J., & Kun, E. (1993) *Nature 361*, 473−475.

Rice, W. G., Supko, J. G., Malspeis, L., Buckheit, R. W. J., Clanton, D., Bu, M., Graham, L., Schaeffer, C. A., Turpin, J. A., Domogala, J., Gogliotti, R., Bader, J. P., Halliday, S. M., Coren, L., Sowder, R. C. I., Arthur, L. O., & Henderson, L. E. (1995) *Science 270*, 1194−1197.

Shimizu, Y., Yamanii, K., Masuho, Y., Yokota, T., Inoue, H., Sudo, K., Satoh, S., & Shimotono, K. (1996) *J. Virol. 70*, 127−132.

Sommerburger, W., Casari, G., Fessl, F., Seipelt, J., & Skern, T. (1994) *Virology 204*, 815−818.

Takamizawa, A., Mori, C., Fuke, I., Manabe, S., Murakami, S., Fujita, J., Onoshi, E., Andoh, T., Yoshida, I., & Okayama, H. (1991) *J. Virol. 65*, 1105−1113.

Tanji, Y., Hijikata, M., Satoh, S., Kaneko, T., & Shimotono, K. (1995) *J. Virol. 69*, 1575−1581.

Tomei, L., Failla, C., Santolini, E., De Francesco, R., & La Monica, N. (1993) *J. Virol 67*, 4017−4026.

Tomei, L., Failla, C., Vitale, R. L., Bianchi, E., & De Francesco, R. (1996) *J. Gen. Virol. 77*, 1065−1070.

Voss, T., Meyer, R., & Sommerburger, W. (1995) *Protein Sci. 4*, 2526−2531.

Yu, S. F., & Lloyd, R. E. (1992) *Virology 186*, 725−735.

Zuckerman, A. J. (1996) *Lancet 347*, 558−559.

BI9616458