Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

THE JOURNAL OF BIOLOGICAL CHEMISTRY Vol. 274, No. 21, Issue of May 21, pp. 14918–14925, 1999
Printed in U.S.A.

A Human RNA Viral Cysteine Proteinase That Depends upon a Unique Zn^{2+}-binding Finger Connecting the Two Domains of a Papain-like Fold*

(Received for publication, January 5, 1999, and in revised form, February 25, 1999)

Jens Herold‡§, Stuart G. Siddell‡, and Alexander E. Gorbalenya¶‖‡‡**

*From the ‡Institute of Virology and Immunology, University of Würzburg, Versbacher Strasse 7, 97078 Würzburg, Germany, the ¶Advanced Biomedical Computing Center, SAIC/NCI-Frederick Cancer Research and Development Center, National Institutes of Health, Frederick, Maryland 21702-1201, ‖Leiden University Medical Center, AZL P4–22, P.O. BOX 9600, 2300RC Leiden, The Netherlands, and the **M. P. Chumakov Institute of Poliomyelitis and Viral Encephalitides, Russian Academy of Medical Sciences, 142782 Moscow Region, Russia*

A cysteine proteinase, papain-like proteinase (PL1pro), of the human coronavirus 229E (HCoV) regulates the expression of the replicase polyproteins, pp1a and ppa1ab, by cleavage between Gly^{111} and Asn^{112}, far upstream of its own catalytic residue Cys^{1054}. In this report, using bioinformatics tools, we predict that, unlike its distant cellular homologues, HCoV PL1pro and its coronaviral relatives have a poorly conserved Zn^{2+} finger connecting the left and right hand domains of a papain-like fold. Optical emission spectrometry has been used to confirm the presence of Zn^{2+} in a purified and proteolytically active form of the HCoV PL1pro fused with the *Escherichia coli* maltose-binding protein. In denaturation/renaturation experiments using the recombinant protein, its activity was shown to be strongly dependent upon Zn^{2+}, which could be partly substituted by Co^{2+} during renaturation. The reconstituted, Zn^{2+}-containing PL1pro was not sensitive to 1,10-phenanthroline, and the Zn^{2+}-depleted protein was not reactivated by adding Zn^{2+} after renaturation. Consistent with the proposed essential structural role of Zn^{2+}, PL1pro was selectively inactivated by mutations in the Zn^{2+} finger, including replacements of any of four conserved Cys residues predicted to co-ordinate Zn^{2+}. The unique domain organization of HCoV PL1pro provides a potential framework for regulatory processes and may be indicative of a nonproteolytic activity of this enzyme.

Proteolytic enzymes control a large variety of processes in cellular organisms and viruses. Some proteases can completely digest most proteins, while others are highly selective, cleaving only one bond in a protein or a set of proteins (1). Mechanistically, four classes of proteases have been recognized, and they are named in accordance with the chemical nature of their catalytic site, namely cysteine proteases, serine proteases, aspartate proteases, and metalloproteases (Zn^{2+}) (2). Each protease class can be further divided into protein families that are united by a common origin and a common structural fold (3). Recently, however, it has become evident that more complex relationships exist among proteases. Thus, proteolytic enzymes employing different catalytic mechanisms within the framework of the same structural fold have been identified (showing that protease families can extend across the borders separating classes (4–6)) and proteases that are decorated with determinants of nonproteolytic activity, for example, polynucleotide-binding cysteine proteases (7–9) and Zn^{2+}-binding nonmetalloproteases (see below) have also been described. The proteases that exhibit these unusual features are very often encoded by positive sense RNA viruses. Many of these viruses rely upon proteolytic enzymes to regulate viral and cellular gene expression and virion morphogenesis (10–12).

The majority of RNA viral proteinases are recognized as being very distantly related to either cellular chymotrypsin-like or papain-like proteases (CLpro and PLpro, respectively).[1] RNA viral CLpro enzymes have been found that exhibit three different catalytic triads, the canonical Ser-His-Asp and two variations not found in cellular proteins, Cys-His-Asp and Cys-His-Glu (5, 13–16). Some RNA viral CLpro enzymes, with the canonical or Cys-His-Asp catalytic triads (*e.g.* the hepatitis C virus NS3 and picornavirus 2A proteases, respectively) employ a structural Zn^{2+} bound by two loops of the two β-barrel fold (17–22). Until very recently, the relationship between cellular and RNA viral PLpros had been primarily based on the results of sequence (23–26) and predicted secondary structure (27) analyses. A recent x-ray analysis of the foot-and-mouth virus (FMDV) Lpro (leader protease), one of the most well characterized viral PLpros (28), has proven the distant relationship between viral and cellular PLpros by showing that the FMDV Lpro adopts a compact version of the PLpro fold (29). Like its cellular homologues, the FMDV Lpro employs the catalytic Cys-His dyad residues, which are assisted by unique Asn and Asp residues replacing, respectively, Gln forming the oxyanion hole as well as Asn hydrogen-bonded to the catalytic His in cellular enzymes. Most probably, similar or other replacements could be found in the catalytic center of the other RNA viral PLpros (26) (see also below). It is reasonable to think that the structural diversity of RNA viral proteinases is related to the extremely high mutation rate of RNA viruses (30). As a result,

* This project has been funded in part by Deutsche Forschungsgemeinschaft Grant SFB 165/B1 (to S. G. S.), Russian Fund for Basic Research Grant 96-04-49562, a fellowship from the Netherlands Organization for Scientific Research (to A. E. G.), and federal funds from NCI, National Institutes of Health, under Contract NO1-CO-56000. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ Present address: Dept. of Microbiology and Immunology, Box 0414, University of California, San Francisco, CA 94143-0414.

‡‡ To whom correspondence should be addressed: SAIC/NCI-FCRDC, 430 Miller Dr., Rm. 235, Frederick, MD 21702-1201. Tel.: 301-846-1991; Fax: 1-301-846-5762; E-mail: gorbalen@ncifcrf.gov.

[1] The abbreviations used are: CLpro, chymotrypsin-like protease; PLpro, papain-like proteinase; aa, amino acid(s); ICP-OES, inductively coupled plasma optical emission spectrometry; MBP, maltose-binding protein; FMDV, foot-and-mouth disease virus; HCoV, human coronavirus 229E; MHV, murine hepatitis virus; TGEV, porcine transmissible gastroenteritis virus; IBV, avian infectious bronchitis coronavirus; CoOAc, cobalt acetate; ZnOAc, zinc acetate.

This paper is available on line at http://www.jbc.org

This is an Open Access article under the CC BY license.

many RNA viral proteinases have diverged to the point where sequence similarity between homologues can barely be detected.

One of the groups of RNA viral PLpro enzymes includes seven ortologous and paralogous coronaviral proteinases. Three coronaviruses (murine hepatitis virus (MHV) (31, 32), human coronavirus 229E (HCoV) (33), and porcine transmissible gastroenteritis virus (TGEV) (34)) encode two divergent copies of PLpro (PL1pro and PL2pro), while avian infectious bronchitis virus (IBV) (23, 26, 35) encodes only one PLpro. These enzymes are produced as part of large replicative precursor proteins (pp1a and pp1ab), with pp1ab being expressed by a mechanism involving (−1) ribosomal frameshifting (36). The PL1pros of MHV (37–40) and HCoV (41) and the PLpro of IBV (42) cleave one or two sites upstream of the protease domain. HCoV PL1pro, which is the focus of the work reported here, was previously mapped between the Gly^{861} and Gln^{1285} residues of pp1a/pp1ab. This protease releases an N-terminal protein, p9, from the pp1a/pp1ab precursors via the cleavage of the Gly^{111}–Asn^{112} bond. The conserved residues Cys^{1054} and His^{1205}, which were predicted to be the catalytic dyad, were found to be indispensable for proteolytic activity of the HCoV PL1pro (41). In this paper, we predict and provide experimental evidence to show that coronaviral PLpros comprise a unique group of enzymes that rely upon a structural Zn^{2+} within the framework of a papain-like fold.

EXPERIMENTAL PROCEDURES

Computer Sequence and Structure Analyses—Amino acid sequences were derived from Swiss-Prot and Genbank™ data bases, and protein structures were derived from the Brookhaven data base (Protein Data Bank, Brookhaven National Laboratory, Upton, NY). Sequence alignments were produced using a family of Clustal programs (43, 44) and the Macaw workbench (45). Nonredundant sequence data bases were searched with single sequences (46), blocks (47), and Hidden Markov Models trained on multiple sequence alignments (48). These alignments were also sent as input for the PhD (49) or DSC (50) programs to predict secondary structure. The 123D program (51) was used for the fold prediction by threading. The structure superimposition and three-dimensional protein homology modeling (52) were done and evaluated with the help of the Quanta 97 and Whatif 4.99 (53) packages. Analysis of the structure-based alignments was assisted by the CORE package (54, 55). Ribbons 2.81 (56) was used to display protein structures, and DSSP (57) was used to calculate secondary structure.

Plasmids—Plasmid pT7-IRES-Pap, encoding amino acids (aa) 1–1315 of pp1a/pp1ab, has been described previously (41). DNA was amplified from pT7-IRES-Pap by standard polymerase chain reaction procedures using oligonucleotides I (5′-ACTGCCATGGGTGGTATTTT-GGCAGTAATA-3′) and II (5′-TTATCACTTGGTAGAAAGCTACATT-GTC-3′). The polymerase chain reaction product was cloned into the *Xmn*I site of pMal-c2 (New England Biolabs, Schwalbach, Germany), resulting in pMal-PL1, and the construction was verified by sequencing. Plasmid pMal-PL1 encodes an isopropyl-1-thio-β-D-galactopyranoside-inducible fusion protein, MBP-PL1, consisting of three parts: the *Escherichia coli* maltose-binding protein (MBP), a linker of three artificial aa, and the HCoV PL1pro domain consisting of aa 861–1285 of pp1a/pp1ab (the so-called "minimal" trans-active protease) (41).

Site-directed Mutagenesis—A modified protocol of site-directed mutagenesis (58) was used. Two partially complementary oligonucleotides carrying the desired mutation were used in a polymerase chain reaction with plasmid DNA as template. The parental DNA was digested with *Dpn*I, and the polymerase chain reaction products were used to transform *E. coli* Top 10F′ cells. *In vivo* recombined plasmid DNA was isolated from individual clones and sequenced. Then DNA fragments from the parental plasmids were replaced with appropriate DNA fragments from the mutated plasmids.

Bacterial Expression and Purification of Recombinant Proteins—Plasmids pMal-c2 and pMal-PL1 were used to transform *E. coli* TB1 cells. Single colonies were inoculated into LB medium supplemented with 100 μM ZnOAc and incubated at 37 °C until an A_{600} of 0.5 was reached. Expression of the recombinant protein was induced with 1 mM isopropyl-1-thio-β-D-galactopyranoside. After incubation at 25 °C for 10 h, the cells were harvested and resuspended in 3 ml of buffer A (20 mM Tris-HCl, pH 7.5, 200 mM NaCl, 1 mM β-mercaptoethanol)/g of cell pellet. Cells were broken with a French pressure cell (SLM Instruments Inc.), and the lysate was clarified by centrifugation at 5000 × *g* for 30 min. The clarified lysate was loaded on a column containing amylose and washed with 10 column volumes of buffer A, and MBP-containing proteins were eluted using buffer A containing 10 mM maltose. The protein concentration was determined and adjusted to 1 or 2 mg/ml. The bacterially expressed recombinant proteins were analyzed by SDS-polyacrylamide gel electrophoresis. Protein preparations were stored at −70 °C.

In Vitro Trans-cleavage Assay—The proteolytic activity of *in vitro* synthesized proteins expressed from pT7-IRES-Pap and its derivatives or the activity of bacterially expressed MBP-PL1 was assayed using *in vitro* synthesized, [^{35}S]Met-labeled substrate, representing aa 1–956 of pp1a/pp1ab (41). The proteolytic reaction was done at 30 °C by incubation of 7 μl of the reticulocyte lysate substrate with either 20 μl of *in vitro* synthesized enzyme for 3 h (41) or 1 μg (0.5 μl) of recombinant protein for 1 h. The reaction products were immunoprecipitated with the polyclonal rabbit antiserum IS1720 (specific for aa 41–250 of pp1a/pp1ab) and analyzed by 10–17.5% gradient SDS-polyacrylamide gel electrophoresis (41). Radioactively labeled polypeptides were visualized by autoradiography.

Denaturation and Renaturation of MBP-PL1—Two 1-ml aliquots (I and II) of purified MBP-PL1 (1 mg/ml) were mixed with 9 volumes of buffer A containing 8.0 M urea. After incubation at 20 °C for 4 h, aliquot I was adjusted with EDTA to a final concentration of 10 mM, and aliquot II was adjusted with ZnOAc to a final concentration of 300 μM. The aliquots were subsequently dialyzed twice against 500 volumes of buffer A containing 1 mM EDTA (aliquot I) or 100 μM ZnOAc (aliquot II) at 4 °C for 16 h. The dialyzed material was then concentrated 10-fold using Centricon-10 membranes (Amicon). The EDTA-treated sample (aliquot I) was subjected to a second round of denaturation as described above and divided into three aliquots (Ia, Ib, and Ic). Aliquot Ia was adjusted with EDTA to a final concentration of 10 mM, aliquot Ib was adjusted with ZnOAc to a final concentration of 300 μM, and aliquot Ic was adjusted with CoOAc to a final concentration of 300 μM. The aliquots were dialyzed against 500 volumes of buffer A containing 1 mM EDTA (aliquot Ia), 100 μM ZnOAc (aliquot Ib), or 100 μM CoOAc (aliquot Ic).

Determination of Metal Content—The recombinant protein solutions were analyzed for the presence of Ca^{2+}, Co^{2+}, Cu^{2+}, Fe^{2+}, Mg^{2+}, Mn^{2+}, Ni^{2+}, Pb^{2+}, and Zn^{2+} by ICP-OES on a JY30 Plus spectrometer (Jobin Yvon, France). Prior to analysis, all samples were treated twice with Chelex-100 resin by mixing in suspension at 4 °C for 30 min.

RESULTS

Coronaviral PL Proteases Are Predicted to Have the α + β Structural Organization Found in Cellular Papain-like Proteinases—In this study, we have employed bioinformatics tools to gain insights into the structure and function of coronaviral PLpros. The predictions derived from these analyses have then been tested experimentally. Previously, this same strategy has proven very efficient in identifying and characterizing numerous RNA viral enzymes (for a review, see Ref. 59), including coronaviral PLpros (37, 40–42).

First, we extended previous analyses (26, 31, 33, 34) of the primary structure of the coronavirus PL protein family using multiple sequence comparison tools (see "Experimental Procedures"). A set of multiple alignments of the conserved PL domains, consisting of aa 202–216, was generated, and an overall alignment was constructed (Fig. 1, *A* and *C*). A number of different approaches were then used to search the data bases for similarities (see "Experimental Procedures"). Only several matches of marginal statistical significance were detected with cysteine proteases, namely between the coronavirus PLpros and the Lpro of different strains of FMDV (26) and also proteases of the ubiquitin isopeptidase T family (60). These weak similarities were structurally reasonable, since in the matches (putative) catalytic Cys or His residues were aligned (data not shown). This failure to identify a pronounced similarity between coronaviral and the other PLpros was, perhaps, not surprising given that pairwise similarity within the coronavirus PLpro family itself is very low (13–32% identical residues).

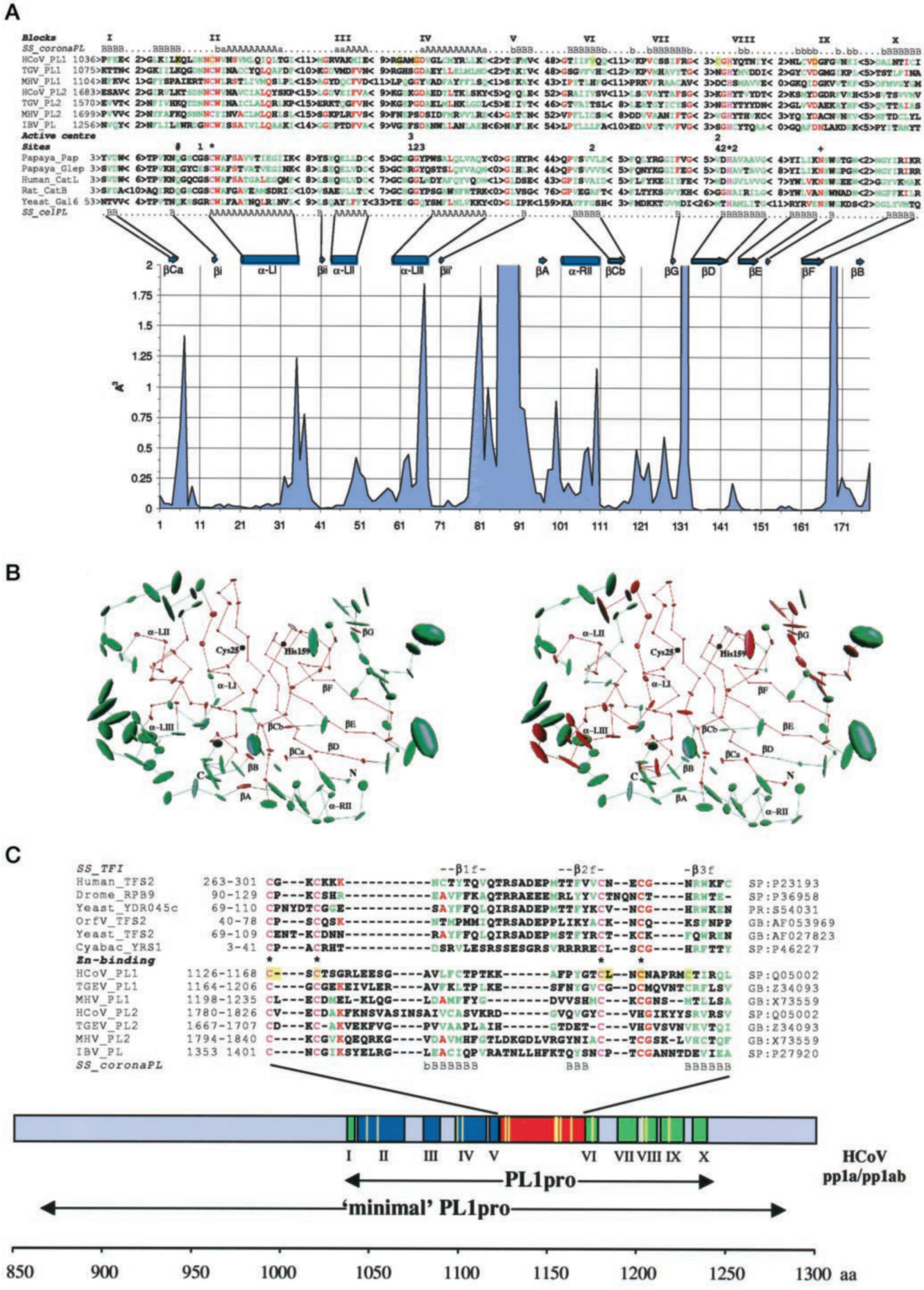
A
B
C
SS_TFI
SS_coronaPL
Zn-binding
HCoV_PL1
TGEV_PL1
MHV_PL1
HCoV_PL2
TGEV_PL2
MHV_PL2
IBV_PL
I II III IV V VI VII VIII IX X
PL1pro
'minimal' PL1pro
HCoV pp1a/pp1ab
850 900 950 1000 1050 1100 1150 1200 1250 1300 aa

FIG. 1

Thus, to check for the existence of more remote structural similarities, which were not evident at the amino acid sequence level (61), the predicted secondary structure of the coronavirus PLpros was threaded (62, 63) through the Protein Data Bank. Even then, no high scores were returned. However, consistent with the above observations and a previous analysis (26), papain was among the three top hits (Z score was 3.16) with the catalytic cysteines of papain and the coronavirus PLpros being matched (data not shown). It is also noteworthy that coronaviral PLpros were predicted to belong to the $\alpha + \beta$ structural class, which also includes cellular PLpros. Thus, we were encouraged to elaborate our model by producing a secondary structure-based alignment of seven coronaviral and 11 cellular PLpros, whose tertiary structures have been solved. This alignment includes 10 ungapped blocks encompassing the majority of the (predicted) structural elements of the two protein sets and has interblock spacing that is similar in the proteases of both cellular and viral origin (Fig. 1*A*). The interfamily amino acid conservation is most evident in blocks II, IV, VI, and VIII, which contain both catalytic and substrate-binding pocket residues of the cellular enzymes and four out of eight absolutely conserved residues of the viral enzymes (three in block II and one in block VIII). Three other invariant residues are located between blocks V and VI and described below. The eighth coronaviral invariant residue, Asp, is located in block IX. In an alignment slightly different from that shown in Fig. 1*A*, this residue would correspond to a highly conserved Asn, which is hydrogen-bonded to catalytic His residue of cellular enzymes. However, this alternative alignment was not favored by the MACAW-assisted analysis or by the results of a site-directed mutagenesis study (see below). In this respect, it is worth noting that the conserved Asn of cellular PLpros does not play the catalytic role (64) and is substituted by Lys in stem bromelain, one of the plant PLpros (65). Importantly, in the alignment shown in Fig. 1*A*, the coronavirus PL residues match the majority (77 out of 91) of cellular PL residues with a low space variation (the so-called core residues (55)) and only a minority (32 out of 87) of structurally less conserved residues (Fig. 1, *A* and *B*). Taken together, all of the above observations strongly imply that the coronaviral PLpros may adopt a variant of the $\alpha + \beta$ fold conserved in cellular PLpros.

Coronaviral PL Proteases Contain a Unique Zn^{2+} Finger Connecting the Two Domains of a Papain-like Fold—In the alignment displayed in Fig. 1*A*, three structural elements of the cellular PLpros, βA, βB, and α-RII, which are spatially juxtaposed in front of the substrate-binding pocket and comprise a peripheral part of the right-hand domain (Fig. 1*B*), were not aligned with coronaviral sequences. Two of these elements, βA and α-RII, form part of a long, poorly conserved structure found in the middle of the cellular PLpros (66). The corresponding region between blocks V and VI in the coronaviral PLpros was highly diverged, but, most conspicuously, it was found to contain a unique conserved sequence pattern characteristic of Zn^{2+} fingers, $CX_{1-2}CX_{22-31}CX_{1-2}$-[CH], where *X* is any aa (Fig. 1*C*).

When the predicted HCoV PL1pro Zn^{2+} finger was compared with the structural data base of Zn^{2+} fingers (67, 68), the nucleic acid binding domain of human transcriptional elongation factor TFIIS, which has a three-stranded, antiparallel, β-sheet fold (zinc ribbon) (69), was selected as being the most

Fig. 1. **Coronaviral and cellular PLpros: structural similarities and unique features.** *A*, secondary structure-based sequence alignment of coronaviral and cellular PLpros. The primary structures of HCoV PL1pro and its coronaviral relatives (for accession numbers see *C*) were aligned using ClustalW program (43) in a stepwise manner and manually corrected with the ClustalX program (44) and the MACAW workbench (45). The main portion of this alignment is presented as 10 ungapped blocks. Only blocks II and III were statistically significant ($p < 10^{-20}$ and $p = 1.5^{-13}$, respectively), and blocks IV and VII, excluding MHV PL1pro, were conditionally significant, using a searching space between blocks III and VIII and between blocks IV and VIII, respectively. The validity of block VIII was previously confirmed by site-directed mutagenesis of conserved His for MHV and HCoV PL1pros and IBV PLpro (37, 41, 42). The secondary structures predicted by the PhD program are shown at the *top* (SS_coronaPL; *A* and *a* represent α-helix, and *B* and *b* represent β-strand, predictions in *capital letters* have a reliability >5 and predictions in *lowercase letters* have a reliability of 5 and less (49)). The validity of this prediction was confirmed when similar secondary structure profiles were also returned for (i) the same alignment using the DSC program (50) and (ii) two automatically generated alignments containing either PL1pros or PL2pros encoded by HCoV and TGEV (not shown). The secondary structure profile of coronaviral PLpros was aligned with secondary structure elements conserved in the tertiary structure of 11 cellular PLpros (SS_celPL) (Protein Data Bank accession numbers: 1ppn, Papaya_Pap, papain (77); 1gec, Papaya_Glep, glycyl endopeptidase (78); 1ppo, caricain (79); 1yac, chymopapain (80); 1mem, cathepsin K (81); 1cjl, Human_CatL, cathepsin L (82); 1cte, Rat_Catb, cathepsin B (83); 2aim, trypanosoma cruzain (84); 2act, actinidin (85); 1gcb, yeast Gal6/bleomycin hydrolase (86)). The secondary structure alignment guided a sequence alignment of coronaviral and cellular proteases. A register of the alignment within each block was (arbitrarily) selected to maximize interfamily sequence similarity, although two or more poorly discriminated alignments were produced for all blocks except blocks II and VIII. When the three-dimensional structures of coronaviral PLpros become available, this alignment may need to be locally adjusted. For cellular PLpros, only a representative set of five sequences is shown. Coloring of the alignment of 12 sequences indicates the following: *pink*, invariant residues; *red*, residues conserved in >50% of the sequences; *green*, group of similar residues. The alignments of coronaviral and cellular PLpros highlight the active site residues of cellular proteases (66, 78). *, principal catalytic; +, "accessory" catalytic; *1*, *2*, *3*, and *4*, substrate-binding pocket subsites S1, S2, S3, and S4, respectively; #, oxyanion hole-forming residue. *Beneath* the alignments, a plot displaying the positional structural variability (55) of cellular PLpros is shown. *Above* the plot, the positions of conserved secondary structure elements of cellular PLpros (66) as well as four conserved hydrogen-forming elements consisting of one residue (not marked) in the primary structure are displayed. *Vertical axis*, space variability at a position of the alignment; *horizontal axis*, numeration in the structural alignment containing only aligned residues. *B*, core structural residues of the cellular PLpros and residues conserved in cellular and coronaviral PLpros. Using the CORE package (55), a structural alignment of 11 cellular PLpros was converted into an average PL structure. It is characterized by the mean position of each C-α atom common in the family. The size of the ellipsoid around each of these atoms is proportional to the volume of atom variance. The two identical average PL structures, consisting of 178 atoms, are displayed in the "standard" papain orientation (66) featuring left-hand and right-hand domains as well as the interdomain active site cleft with the two catalytic residues of papain, Cys^{25} and His^{159}. Conserved secondary structure elements of cellular PLpros are also marked. These structures are colored in *green* and *red* as follows. The left structure, the half of C-α atoms plus two atoms having the lowest space variance (91 atoms) are colored in *red* (core), and the remaining atoms are in *green* (noncore). The right structure, 109 atoms, whose residues were aligned with coronaviral PL residues in Fig. 1*A*, are shown in *red* (interfamily conserved residues), and the remaining atoms are in *green*. Note that the cellular PL core residues and the interfamily conserved residues are mainly from the same pool. *C*, A unique Zn^{2+} finger connects the two domains of the PL fold of coronaviral PLpros. A region of the coronaviral PLpros between blocks V and VI was aligned as specified in Fig. 1*A*. Using the secondary structures predicted for the PLpros (SS_coronaPL) (50) and derived from the NMR structure (69) of the TFIIS Zn^{2+} ribbon (SS_TFI), an alignment of Zn^{2+} fingers of coronaviral PLpros and TFIIS was generated. The positions of these sequences in the corresponding proteins are given on the *left*, and accession numbers in the sequence data bases are shown on the *right*. *Coloring* of the alignment is as detailed for *A*. Residues involved in Zn^{2+} binding in TFIIS (69) are marked. A *bar* depicts the region of HCoV pp1a/pp1ab characterized in this study with the conserved blocks (Fig. 1*A*) shown. These blocks are organized in three groups colored differently. *Blue*, left-hand α-helix domain; *green*, right-hand β-sheet domain without counterparts of βA- and βB-strands; *red*, Zn^{2+} finger domain. *Beneath* the *bar*, the positions of the PL1pro domain, which is conserved among coronaviruses, and the HCoV minimal PL1pro domain determined by deletion analysis (41) are shown. The positions of mutations (Ref. 41 and Table II) are depicted with *yellow vertical lines* in the *bar* and *yellow* amino acid background in the alignments in *A* and *C*.

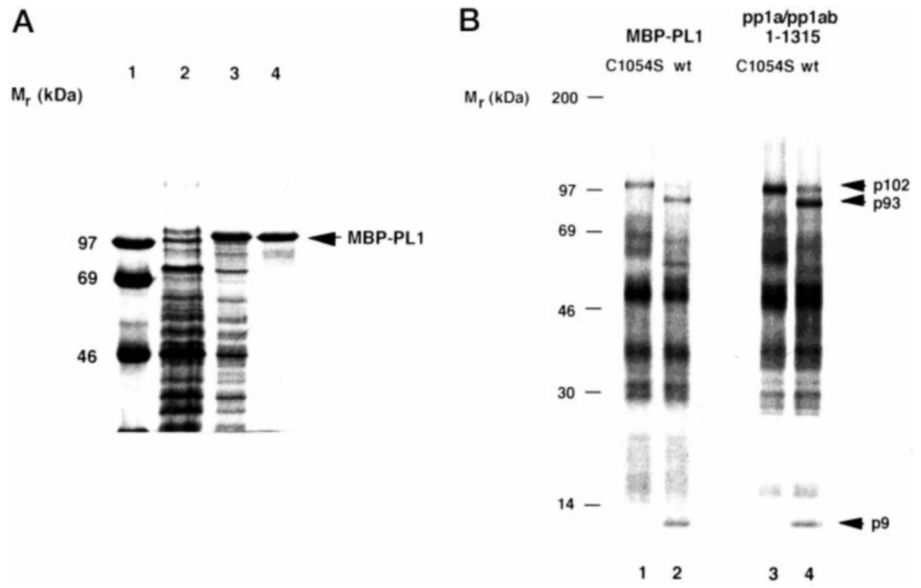

FIG. 2. **Expression and purification of proteolytically active HCoV PL1pro fused with the *E. coli* maltose-binding protein.** *A*, purification of the fusion protein. MBP-PL1 was purified by affinity chromatography on amylose column from lysates of *E. coli* transformed with TB1[pMal-PL1] as described under "Experimental Procedures." Aliquots taken from different stages of the purification were analyzed by 12.5% SDS-polyacrylamide gel electrophoresis. *Lane 1*, molecular mass markers; *lane 2*, noninduced bacterial lysate; *lane 3*, isopropyl-1-thio-β-D-galactopyranoside-induced bacterial lysate; *lane 4*, protein after amylose affinity chromatography. The position of MBP-PL1 is indicated. *B*, proteolytic activity of MBP-PL1. The trans-cleavage assay using *in vitro* generated [^{35}S]Met-labeled substrate was used to monitor proteolytic activity of purified MBP-PL1 (*lanes 1* and *2*) and *in vitro* generated, nonlabeled polypeptide containing PL1pro (pp1a/pp1ab-(1–1315)) and its mutated derivative (*lanes 3* and *4*). After immunoprecipitation of the cleavage reaction with IS 1720, proteins were separated by 10–17.5% gradient SDS-polyacrylamide gel electrophoresis, and labeled polypeptides were visualized by autoradiography. The positions of molecular mass markers, the substrate (p102), and cleavage products (p93 and p9) are indicated. The source of enzyme was as follows: MBP-PL1 C1054S (*lane 1*), MBP-PL1 (*lane 2*), *in vitro* produced PL1pro C1054S (*lane 3*), *in vitro* produced PL1pro (*lane* 4).

similar. Both proteins are enriched with β-strands, belong to the C4 class, have a similar spacing between the pair of the C2 halves, and have a variable spacing separating cysteines within the C2 elements (or cysteine and histidine in the CH element of related proteins) (Fig. 1*C*). The zinc ribbon architecture is apparently conserved in a wide variety of Zn^{2+} fingers (69, 70) and was shown to tolerate a large size difference in the loop structure (71). These observations strongly suggest that the specific features of the coronaviral PLpro Zn^{2+} fingers, which are evident from the alignment shown in Fig. 1*C*, may also be compatible with this architecture.

From the comparative sequence analyses outlined above, we predicted that coronaviral PLpros are composed of the two domains of the papain-like fold which, unlike their known cellular relatives, are connected by a Zn^{2+} finger.

Purified PL1pro Fused with E. coli Maltose-binding Protein Is Proteolytically Active—To address the enzymatic and predicted Zn^{2+} binding properties of coronaviral PLpros experimentally, we decided to study a recombinant, purified HCoV PL1pro that was amenable to genetic and biochemical manipulations. The minimal HCoV PL1pro domain, which had been shown to be active when synthesized in a reticulocyte lysate (41), was expressed in *E. coli* as the fusion protein MBP-PL1 (Fig. 2*A*, *lanes 2* and *3*). MBP-PL1 was partly purified by amylose affinity chromatography (Fig. 2*A*, *lane 4*) and assayed for proteolytic activity. The recombinant MBP-PL1, but not a derivative carrying a replacement of the catalytic Cys1054 by serine, was proteolytically active (Fig. 2*B*, *lanes 1* and *2*). The proteolytic activity of recombinant MBP-PL1 was identical to that of *in vitro* synthesized HCoV PL1pro (Fig. 2*B*, *lanes 3* and *4*; Ref. 41).

HCoV PL1pro Is a Zn^{2+}-binding Enzyme—The partly purified MBP-PL1 was analyzed for eight metals by ICP-OES. No other metals except Zn^{2+} and Fe^{2+} were detected in significant amounts. The Zn^{2+}/Fe^{2+} ratio was approximately 3–4.5/1 (Table I). MBP, expressed and purified under identical conditions, bound only background amounts of the metal ions, confirming that it is the coronaviral portion of the fusion protein that binds Zn^{2+}. It was calculated that HCoV PL1pro binds Zn^{2+} in equimolar amounts.

TABLE I
Metal content in MBP-PL1

Sample	Batch	Zinc	Iron
		μg/ml	*μg/ml*
Buffer		ND[a]	ND
MBP (2 mg/ml)[b]	1	25	26
	2	ND	ND
MBP-PL1 (2 mg/ml)[b]	1	546	120
	2	554	184

[a] ND, not detectable.
[b] Total protein concentration, using gel densitometry.

Zn^{2+} Is an Essential Structural Cofactor of Proteolytically Active HCoV PL1pro—In a series of preliminary experiments, we observed that if ZnOAc was not added to the bacterial growth media, purified MBP-PL1 was proteolytically less active and there was a decreased amount of Zn^{2+} and increased amount of Fe^{2+} bound to the protein (data not shown). This observation indicated that Zn^{2+} might be an essential co-factor of PL1pro. To test this hypothesis, MBP-PL1 was denatured in the presence of 8 M urea and then renatured by dialysis against buffer A supplemented with either ZnOAc or EDTA (72). MBP-PL1 renatured in the presence of Zn^{2+}, but not EDTA, was proteolytically active (Fig. 3, *lanes 4* and *5*). Furthermore, the Zn^{2+}-depleted MBP-PL1 (apoenzyme) could be reactivated after a second round of denaturation and renaturation in the presence of ZnOAc or, to a lesser degree, CoOAc (Fig. 3, *lanes 6–8*). Importantly, Fe^{2+} could not substitute for Zn^{2+}, the addition of 100 μM ZnOAc to the apoenzyme did not restore activity, and the reconstituted Zn^{2+}-PL1pro was not inhibited

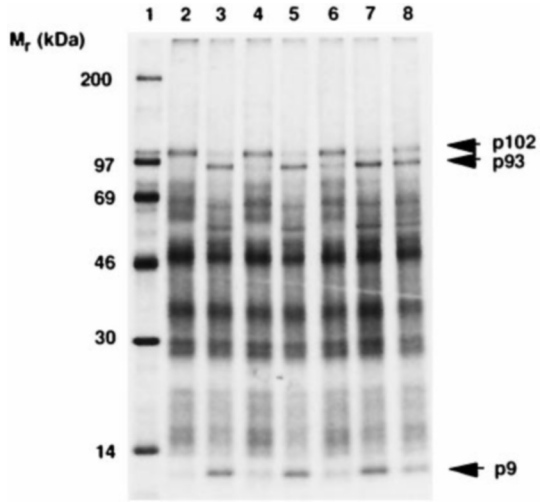

FIG. 3. **Effect of Zn^{2+} on the proteolytic activity of MBP-PL1.** The trans-cleavage assay was used to monitor proteolytic activity of MBP-PL1 subjected to one or two cycles of denaturation in the presence of 8 M urea and renaturation in the presence of EDTA, ZnOAc, or CoOAc. The substrate and cleavage products are indicated as in Fig. 2. *Lane 1*, molecular markers. The source of enzyme was as follows: buffer A (*lane 2*), MBP-PL1 (not treated) (*lane 3*), denatured MBP-PL1 renatured in the presence of EDTA (apoenzyme) (*lane 4*), denatured MBP-PL1 renatured in the presence of ZnOAc (*lane 5*), denatured apoenzyme MBP-PL1 renatured in the presence of EDTA (*lane 6*), denatured apoenzyme MBP-PL1 renatured in the presence of ZnOAc (*lane 7*), denatured apoenzyme MBP-PL1 renatured in the presence of CoOAc (*lane 8*).

by 10 mM 1,10-phenanthroline (data not shown). These results strongly suggest that Zn^{2+} is tightly bound to the PL1pro and that coordination of the Zn^{2+} has to occur during folding of the protein. This also implies that Zn^{2+} plays an essential structural rather than catalytic role.

Four Cysteine Residues Implicated in Binding Zn^{2+} Are Crucial for Proteolytic Activity of HCoV PL1pro—In our model of the HCoV PL1pro, the Zn^{2+} is predicted to be tetrahedrally coordinated by Cys^{1126}, Cys^{1128}, Cys^{1154}, and Cys^{1157} (Fig. 1*C*). If this is correct, and in view of the strong dependence of MBP-PL1 activity on Zn^{2+} (Fig. 3), these four cysteine residues should be indispensable for proteolytic activity. We tested this prediction by assaying the proteolytic activity of *in vitro* synthesized PL1pro mutants. Consistent with our prediction, mutants with a replacement of any of the putative Zn^{2+}-coordinating cysteine residues were proteolytically inactive (Table II). Furthermore, PL1pro was found to be inactivated by two other mutations in the vicinity of the Zn^{2+}-coordinating Cys residues, an insertion of Val between aa 1126 and 1127 or a deletion of Leu^{1155} (Table II and Fig. 1*A*). The effect of the above mutations was highly selective, since two different mutations at Cys^{1163}, a nonconserved residue of the Zn^{2+} finger, did not inactivate the enzyme (Table II). Also, six other highly conserved residues, Lys^{1048}, Gly^{1099}, Gly^{1102}, Val^{1175}, Cys^{1203}, and Asp^{1218}, were probed outside the Zn^{2+} finger. The majority of these residues were predicted to be involved in the formation of the active site, as can be seen from the alignment of the coronaviral and cellular PLpros (Fig. 1*A*). Each of these residues (with the exception of Lys^{1048}) was replaced by a number of amino acids found in the same position in other PLpro(s) or by other amino acids (Fig. 1*A* and Table II). In contrast to the four conserved Cys residues of the Zn^{2+} finger, the majority of these residues tolerated at least one replacement. The exception, Lys^{1048}, which was aligned with the oxyanion hole-forming Gln^{15} of papain, did not tolerate a Glu substitution, the only tested mutation. These results are fully compatible with our alignment of coronaviral and cellular PLpros (Fig. 1) and indicate that the four conserved Cys residues of the Zn^{2+} finger are as equally important for the proteolytic activity of the HCoV PL1pro as the previously identified catalytic Cys^{1054} and His^{1205} residues (41).

DISCUSSION

In this report, we describe a unique cysteine proteinase, dependent on Zn^{2+} and encoded by an RNA virus. A combination of bioinformatics, biochemical, and molecular-genetic techniques has been used to obtain evidence that strongly supports a structural, rather than a catalytic, role of Zn^{2+} in the HCoV PL1pro. To characterize the HCoV PL1pro, we have expressed the enzyme in *E. coli* as a fusion protein that was shown to be proteolytically active in a trans-cleavage assay. The ability of PL1pro to cleave a cognate substrate seems to be completely dependent on additional (viral) cofactor(s), since we have been unable to develop a peptide-based assay for this protease to date. These putative cofactor(s) may reside in the approximately 900-aa region that separates the catalytic domain of PL1pro and its cleavage site. This region is known to modulate PL1pro-mediated processing in MHV (37, 73).

The most intriguing aspect of the HCoV PL1pro structure is, without doubt, the zinc finger. We have, first of all, delineated this structure on theoretical grounds and subsequently used a recombinant PL1pro to characterize the Zn^{2+} binding properties by spectroscopic analysis as well as the structure-function relationships by mutational analysis. The experimental results have validated the functional importance of the HCoV PL1pro Zn^{2+} finger, although we acknowledge that it remains to be shown directly that its four conserved Cys residues bind Zn^{2+}. In the first analysis, the Zn^{2+} finger was predicted to adopt a β-sheet topology, most probably of the zinc ribbon type, and be incorporated into a papain-like fold. As an extension of this analysis, a possible tertiary organization of PL1pro is presented in Fig. 4 as a ribbon model that was built by homology modeling (52) using the alignments shown in Fig. 1. We believe that this model is useful for rationalizing the results presented here (see below) and for the development of more precise mutagenesis analyses. At the same time, we recognize that the model most probably deviates significantly from the coordinates of the actual structure. This lack of robustness is due to the fact that none of the current methods of homology modeling (52) can rigorously predict the structure of proteins, when modeling is based upon two separate templates of a marginal similarity, as was the case for PL1pro. Despite these reservations about the model, the compatibility of the Zn^{2+} ribbon and the PL fold is supported by two observations: (i) the distance between the N and C termini of the Zn^{2+} ribbon fits between the two "acceptor" residues of the core PL domains in papain and (ii) the Zn^{2+} ribbon spatially replaces the C-terminal βB-strand in papain, and, accordingly, a counterpart of this strand is not conserved in the coronaviral PLpros (Fig. 1*A*). The interdomain junction role of the Zn^{2+} finger is also compatible with the observation that many RNA viral PLpros, although not all (26, 74), have considerably smaller sizes than coronaviral PLpros and may contain only a short interdomain loop in place of the Zn^{2+} finger.[2] A unique β-sheet topology of the interdomain junction has most recently been also described for the FMDV Lpro, the only RNA viral PLpro whose three-dimensional structure was solved (29).

The Zn^{2+} finger occupies one of the most pronouncedly diverged regions of coronaviral PLpros and seems to form a separate domain (Figs. 1, *A* and *C*, and Fig. 4). Surprisingly, the proteolytic activity of HCoV PL1pro was abolished by any of three Zn^{2+} finger mutations that mimicked the wild-type se-

[2] A. E. Gorbalenya, unpublished observations.

TABLE II
Proteolytic activity of HCoV PL1pro mutants

Region of HCoV PL1pro	Replacements in pp1a/pp1ab (from → to)[a]	PLpro whose sequence at this position matches the HCoV PL1 mutant	Proteolytic activity[b]	Reference
Zinc finger	C1126D, H		No	This work
	V1126/27ins	HCoV PL2	No	This work
	C1128A, D, P		No	This work
	C1154A, H, D		No	This work
	L1155del	MHV PL1/HCoV PL2	No	This work
	C1157A, D, P		No	This work
	H	HCoV PL2/TGV PL2	No	
	C1163A	IBV PL	Yes	This work
	D		Yes	
Outside of zinc finger	K1048E	IBV PL	No	This work
	C1054S, A, G		No	41
	G1099P	MHV PL1	Yes	This work
	A		No	
	G1102S		Yes	This work
	A	MHV PL2/IBV PL	Yes	
	V1175H	MHV PL1	No	This work
	P		No	
	N		Yes	
	T	TGV PL2	Yes	
	C1203D	MHV PL1	Yes	This work
	A		No	
	H1205A, G		No	41
	D1218N, Q, E, H, K, A		Yes	This work
			Yes	

[a] V1126/27ins, insertion of Val between aa 1126 and 1127; L1155del, deletion of L1155.
[b] Proteolytic activity was monitored using SDS-polyacrylamide gel electrophoresis of the products of reactions catalyzed by an *in vitro* produced PL1pro or its derivatives in the trans-cleavage and monomolecular assays (41).

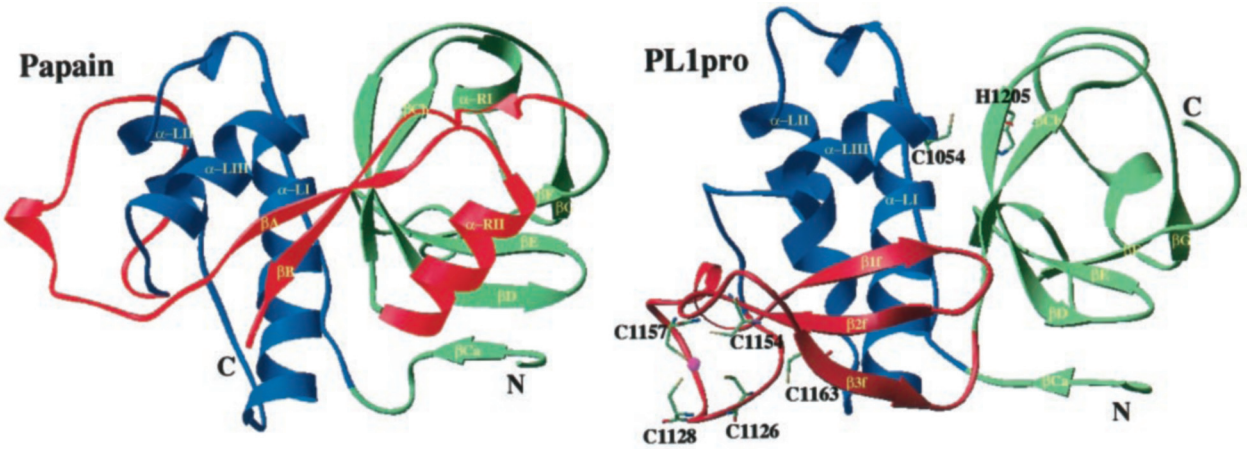

FIG. 4. **A crude structural model of HCoV PL1pro.** The PL1pro model (*right*; pp1a/1ab residues 1033–1242) was generated using the structures of papain (1ppn; shown *left*) and human TFIIS (1tfi) as templates for building the core and a loop library for constructing the variable regions by using Quanta 97 and Whatif 4.99. The model was partly refined but not minimized. Secondary structure was calculated according to Ref. 57 and, for some positions in PL1pro, from the predicted secondary structure. The structures of both proteases are displayed (56) in the standard papain orientation (66) and split into three domains. These are *colored* according to a scheme given in Fig. 1*C* as follows. *Blue*, left-hand α-helix domain; *green*, right-hand β-sheet domain without counterparts of βA- and βB-strands; *red*, Zn^{2+} finger domain in PL1pro and the interdomain loop along with βA- and βB-strands and α-RII-helix in papain. Cysteine residues of the Zn^{2+} finger as well as the catalytic dyad residues of PL1pro that have been probed by site-directed mutagenesis (Table II) are shown in the *ball-and-stick model*.

quences of HCoV PL1pro relatives (V1126/27ins, L1155del, and C1157H in Table II). To reconcile these findings, one must propose that, in the evolution of the HCoV PL1pro homologues, such mutations have been suppressed by other replacements in the Zn^{2+} finger or the catalytic domains. If the changes were limited to the Zn^{2+} finger, this would imply that the Zn^{2+} finger is either involved in substrate binding or controls movement of the catalytic domains. Both options seem possible given the interdomain position of the Zn^{2+} finger and its likely proximity to the active site (Fig. 4). If the catalytic domains accepted mutations as well, this would be indicative of an interaction between these domains and the Zn^{2+} finger. Again, the current model does not preclude an interaction between the Zn^{2+} finger and α-LI helix, which is capped by the catalytic Cys^{1054} (Fig. 4). It is conceivable that the PL1pro Zn^{2+} finger interacts with both the catalytic domains and substrate.

HCoV PL1pro and, by implication, the other coronaviral PLpros extend the list of nonmetal proteinases dependent on Zn^{2+} and encoded by RNA viruses. Recently, a Zn^{2+} has been identified as a structural component in the picornavirus 2A and hepatitis C virus NS3 proteases (17–22). In these enzymes, Zn^{2+} is uniquely coordinated by residues that have been naturally engineered in two loops of the CLpro fold. In both the chymotrypsin-like or papain-like Zn^{2+}-binding viral proteases, Zn^{2+} is located at the site opposite to the active center. The recurrent emergence of Zn^{2+} dependence in the two independent lineages of RNA viral nonmetal proteinases could be linked to their biology (12). For example, in the coronaviral PLpros,

the prominent position of the Zn^{2+} finger makes it an excellent candidate for mediating external signals that might modulate proteolytic activity inside cells. Through the Zn^{2+} finger, the PLpros might interact with proteins or polynucleotides and, in turn, affect processes regulated by the PLpro partner (75).

Among the positive sense RNA viruses, coronaviruses have the largest genomes (close to 30 kilobases) and probably use the most sophisticated mechanisms of RNA replication and transcription (76). As is the case for poliovirus (7), coronaviral PLpros might turn out to be important regulators at these levels. It may not be entirely coincidental that the Zn^{2+} finger used as the template in our HCoV PL1pro model employs a variant of the general architecture motif of RNA polymerases (70). The lack of an infectious DNA clone for any coronavirus remains a major technical obstacle in our attempts to elucidate the possible roles of PLpros in the viral replicative cycle. However, regardless of the role(s) played by PLpros, the integrity of its unusual domain organization could be targeted by specific drugs for the control of coronaviral infections.

Acknowledgments—We thank B. Schelle for excellent technical assistance, P. Schramel for ICP-OES analyses, M. Gerstein and R. Altman for help with the CORE package, K. Miaskiewicz for the administration of computer resources, and S. Burt and J. Maizel for commenting on the manuscript.

REFERENCES

1. Neurath, H. (1984) *Science* **224,** 350–357
2. Hartley, B. S. (1960) *Annu. Rev. Biochem.* **29,** 45–72
3. Rawlings, N. D., and Barrett, A. J. (1993) *Biochem. J.* **290,** 205–218
4. Gorbalenya, A. E., Blinov, V. M., and Donchenko, A. P. (1986) *FEBS Lett.* **194,** 253–257
5. Bazan, J. F., and Fletterick, R. J. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85,** 7872–7876
6. Ollis, D. L., Cheah, E., Cygler, M., Dijkstra, B., Frolow, F., Franken, S. M., Harel, M., Remington, S. J., Silman, I., Schrag, J., Sussman, J. L., Verschueren, K. H. G., and Goldman, A. (1992) *Prot. Eng.* **5,** 197–211
7. Andino, R., Rieckhof, G. E., and Baltimore, D. (1990) *Cell* **63,** 369–380
8. Xu, H. E., and Johnson, S. A. (1994) *J. Biol. Chem.* **269,** 21177–21183
9. Mangel, W. F., Toledo, D. L., Ding, J., Sweet, R. M., and McGrath, W. J. (1997) *Trends Biochem. Sci.* **22,** 393–398
10. Kräusslich, H. G., and Wimmer, E. (1988) *Annu. Rev. Biochem.* **57,** 701–754
11. Dougherty, W. G., and Semler, B. L. (1993) *Microbiol. Rev.* **57,** 781–822
12. Gorbalenya, A. E., and Snijder, E. J. (1996) *Perspect. Drug Discov. Des.* **6,** 64–86
13. Gorbalenya, A. E., Donchenko, A. P., Blinov, V. M., and Koonin, E. V. (1989) *FEBS Lett.* **243,** 103–114
14. Choi, H.-K., Tong, L., Minor, W., Dumas, P., Boege, U., Rossmann, M. G., and Wengler, G. (1991) *Nature* **354,** 37–43
15. Allaire, M., Chernaia, M. M., Malcolm, B. A., and James, M. N. G. (1994) *Nature* **369,** 72–76
16. Matthews, D. A., Smith, W. W., Ferre, R. A., Condon, B., Budahazi, G., Sisson, W., Villafranca, J. E., Janson, C. A., McElroy, H. E., Gribskov, C. L., and Worland, S. (1994) *Cell* **77,** 761–771
17. Yu, S. F., and Lloyd, R. E. (1992) *Virology* **186,** 725–735
18. De Francesco, R., Urbani, A., Nardi, M. C., Tomei, L., Steinkuhler, C., and Tramontano, A. (1996) *Biochemistry* **35,** 13282–13287
19. 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
20. 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,** 343–55
21. Sommergruber, W., Seipelt, J., Fessl, F., Skern, T., Liebig, H. D., and Casari, G. (1997) *Virology* **234,** 203–214
22. Stempniak, M., Hostomska, Z., Nodes, B. R., and Hostomsky, Z. (1997) *J. Virol.* **71,** 2881–2886
23. Gorbalenya, A. E., Koonin, E. V., Donchenko, A. P., and Blinov, V. M. (1989) *Nucleic Acids Res.* **17,** 4847–4861
24. Hardy, W. R., and Strauss, J. H. (1989) *J. Virol.* **63,** 4653–4664
25. Oh, C.-S., and Carrington, J. C. (1989) *Virology* **173,** 692–699
26. Gorbalenya, A. E., Koonin, E. V., and Lai, M. M. C. (1991) *FEBS Lett.* **288,** 201–205
27. Skern, T., Fita, I., and Guarne, A. (1998) *J. Gen. Virol.* **79,** 301–307
28. Ryan, M. D., and Flint, M. (1997) *J. Gen. Virol.* **78,** 699–723
29. Guarne, A., Tormo, J., Kirchweger, R., Pfistermueller, D., Fita, I., and Skern, T. (1998) *EMBO J.* **17,** 7469–7479
30. Holland, J., and Domingo, E. (1998) *Virus Genes* **16,** 13–21
31. Lee, H. J., Shieh, C. K., Gorbalenya, A. E., Koonin, E. V., La Monica, N., Tuler, J., Bagdzhadzhyan, A., and Lai, M. M. (1991) *Virology* **180,** 567–582
32. Bonilla, P. J., Gorbalenya, A. E., and Weiss, S. R. (1994) *Virology* **198,** 736–740
33. Herold, J., Raabe, T., Schelle-Prinz, B., and Siddell, S. G. (1993) *Virology* **195,** 680–691
34. Eleouet, J. F., Rasschaert, D., Lambert, P., Levy, L., Vende, P., and Laude, H. (1995) *Virology* **206,** 817–822
35. Boursnell, M. E. G., Brown, T. D. K., Foulds, I. J., Green, P. F., Tomley, F. M., and Binns, M. M. (1987) *J. Gen. Virol.* **68,** 57–77
36. Brierley, I. (1995) *J. Gen. Virol.* **76,** 1885–1892
37. Baker, S. C., Yokomori, K., Dong, S., Carlisle, R., Gorbalenya, A. E., Koonin, E. V., and Lai, M. M. (1993) *J. Virol.* **67,** 6056–6063
38. Dong, S., and Baker, S. C. (1994) *Virology* **204,** 541–549
39. Hughes, S. A., Bonilla, P. J., and Weiss, S. R. (1995) *J. Virol.* **69,** 809–813
40. Bonilla, P. J., Hughes, S. A., and Weiss, S. R. (1997) *J. Virol.* **71,** 900–909
41. Herold, J., Gorbalenya, A. E., Thiel, V., Schelle, B., and Siddell, S. G. (1998) *J. Virol.* **72,** 910–918
42. Lim, K. P., and Liu, D. X. (1998) *Virology* **245,** 303–312
43. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res.* **22,** 4673–4680
44. Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., and Higgins, D. G. (1997) *Nucleic Acids Res.* **25,** 4876–4882
45. Schuler, G. D., Altschul, S. F., and Lipman, D. J. (1991) *Proteins* **9,** 180–190
46. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) *Nucleic Acids Res.* **25,** 3389–3402
47. Henikoff, S., Pietrokovski, S., and Henikoff, J. G. (1998) *Nucleic Acids Res.* **26,** 309–312
48. Eddy, S. R. (1996) *Curr. Opin. Struct. Biol.* **6,** 361–365
49. Rost, B. (1996) *Methods Enzymol.* **266,** 525–539
50. King, R. D., Saqi, M., Sayle, R., and Sternberg, M. J. (1997) *Comput. Appl. Biosci.* **13,** 473–474
51. Alexandrov, N. N., Nussinov, R., and Zimmer, R. M. (1996) *Pacific Symposium on Biocomputing, 1996* (Hunter, L., and Klein, T. E., eds) pp. 53–72, World Scientific, Singapore
52. Sanchez, R., and Sali, A. (1997) *Curr. Opin. Struct. Biol.* **7,** 206–214
53. Vriend, G., and Sander, C. (1993) *J. Appl. Crystallogr.* **26,** 47–60
54. Altman, R. B., Hughes, C., and Gerstein, M. B. (1995) *J. Mol. Graph.* **13,** 142–152
55. Gerstein, M., and Altman, R. B. (1995) *Comput. Appl. Biosci.* **11,** 633–644
56. Carson, M. (1997) *Methods Enzymol.* **277,** 493–505
57. Kabsch, W., and Sander, C. (1983) *Biopolymers* **22,** 2577–2637
58. Herold, J., Siddell, S., and Ziebuhr, J. (1996) *Methods Enzymol.* **275,** 68–89
59. Gorbalenya, A. E., and Koonin, E. V. (1993) *Sov. Sci. Rev. D. Physicochem. Biol.* **11,** 1–84
60. Wilkinson, K. D. (1995) *Annu. Rev. Nutr.* **15,** 161–189
61. Aurora, R., and Rose, G. D. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95,** 2818–2823
62. Alexandrov, N. N., and Fischer, D. (1996) *Proteins* **25,** 354–365
63. Bowie, J. U., Luthy, R., and Eisenberg, D. (1991) *Science* **253,** 164–170
64. Vernet, T., Tessier, D. C., Chatellier, J., Plouffe, C., Lee, T. S., Thomas, D. Y., Storer, A. C., and Menard, R. (1995) *J. Biol. Chem.* **270,** 16645–16652
65. Ritonja, A., Rowan, A. D., Buttle, D. J., Rawlings, N. D., Turk, V., and Barrett, A. J. (1989) *FEBS Lett.* **247,** 419–424
66. Kamphuis, I. G., Drenth, J., and Baker, E. N. (1985) *J. Mol. Biol.* **182,** 317–329
67. Schwabe, J. W., and Klug, A. (1994) *Nat. Struct. Biol.* **1,** 345–349
68. Hubbard, T. J. P., Murzin, A. G., Brenner, S. E., and Chothia, C. (1997) *Nucleic Acids Res.* **25,** 236–239
69. Qian, X., Gozani, S. N., Yoon, H., Jeon, C. J., Agarwal, K., and Weiss, M. A. (1993) *Biochemistry* **32,** 9944–9959
70. Wang, B., Jones, D. N., Kaine, B. P., and Weiss, M. A. (1998) *Structure* **6,** 555–569
71. Zhu, W., Zeng, Q., Colangelo, C. M., Lewis, L. M., Summers, M. F., and Scott, R. A. (1996) *Nat. Struct. Biol.* **3,** 122–124
72. Sommergruber, W., Casari, G., Fessl, F., Seipelt, J., and Skern, T. (1994) *Virology* **204,** 815–818
73. Bonilla, P. J., Hughes, S. A., Pinon, J. D., and Weiss, S. R. (1995) *Virology* **209,** 489–497
74. Liu, X., Ropp, S. L., Jackson, R. J., and Frey, T. K. (1998) *J. Virol.* **72,** 4463–4466
75. Berg, J. M., and Shi, Y. (1996) *Science* **271,** 1081–1085
76. Lai, M. M. C., and Cavanagh, D. (1997) *Adv. Virus Res.* **48,** 1–100
77. Harris, G., Pickersgill, R., Howlin, B., and Moss, D. (1992) *Acta Crystallogr. B* **48,** 67–75
78. O'Hara, B. P., Hemmings, A. M., Buttle, D. J., and Pearl, L. H. (1995) *Biochemistry* **34,** 13190–13195
79. Pickersgill, R. W., Rizkallah, P., Harris, G. W., and Goodenough, P. W. (1991) *Acta Crystallogr. Sec. B* **47,** 766–771
80. Maes, D., Bouckaert, J., Poortmans, F., Wyns, L., and Looze, Y. (1996) *Biochemistry* **35,** 16292–16298
81. McGrath, M. E., Klaus, J. L., Barnes, M. G., and Bromme, D. (1997) *Nat. Struct. Biol.* **4,** 105–109
82. Coulombe, R., Grochulski, P., Sivaraman, J., Menard, R., Mort, J. S., and Cygler, M. (1996) *EMBO J.* **15,** 5492–5503
83. Jia, Z., Hasnain, S., Hirama, T., Lee, X., Mort, J. S., To, R., and Huber, C. P. (1995) *J. Biol. Chem.* **270,** 5527–5533
84. McGrath, M. E., Eakin, A. E., Engel, J. C., McKerrow, J. H., Craik, C. S., and Fletterick, R. J. (1995) *J. Mol. Biol.* **247,** 251–259
85. Baker, E. N. (1980) *J. Mol. Biol.* **141,** 441–484
86. Joshua-Tor, L., Xu, H. E., Johnston, S. A., and Rees, D. C. (1995) *Science* **269,** 945–950