1995

Proteolytic Processing of the Murine Coronavirus Polymerase Polyprotein

Shanghong Dong

Loyola University Chicago

Follow this and additional works at: https://ecommons.luc.edu/luc_diss

Part of the Molecular Biology Commons

Recommended Citation
https://ecommons.luc.edu/luc_diss/3387

This Dissertation is brought to you for free and open access by the Theses and Dissertations at Loyola eCommons. It has been accepted for inclusion in Dissertations by an authorized administrator of Loyola eCommons. For more information, please contact ecommons@luc.edu.

This work is licensed under a Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 License.
Copyright © 1995 Shanghong Dong
LOYOLA UNIVERSITY OF CHICAGO

PROTEOLYTIC PROCESSING OF THE MURINE CORONAVIRUS POLYMERASE POLYPROTEIN

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE GRADUATE SCHOOL
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

PROGRAM IN MOLECULAR BIOLOGY
STRITCH SCHOOL OF MEDICINE

BY

SHANGHONG DONG

MAYWOOD, ILLINOIS
JANUARY 1995
Copyright by Shanghong Dong, 1994.
All Rights Reserved.
ACKNOWLEDGMENTS

I would like to thank my advisor, Dr. Susan Baker for her patience, encouragement and guidance; my committee members: Drs. Leonard Erickson, Allen Frankfater, Mark Peeples, and Kenneth Thompson for their commitment, technical help and constructive suggestions. I would also like to thank the members of the Program in Molecular Biology and faculty in the Department of Microbiology and Immunology for their help and interest in my research. Special thanks to people in Dr. Baker's Laboratory including: Ricarda Carlisle, Juliet Glomb, Hongqiang Gao, Jim George, Cindy Rowe, Jenny Schiller, Dr. Cecilia Adami and Dr. Fabeha Fazal, for their assistance and friendship. Finally, I would like to dedicate the dissertation to my parents, my husband and my sisters, whose extraordinary love and support made it possible.
ABSTRACT

The goal of this dissertation research is to understand proteolytic processing of the polymerase polyprotein in murine coronavirus mouse hepatitis virus (MHV). MHV polymerase is encoded by gene 1 as a polyprotein, which is proteolytically processed into mature proteins by three proposed proteinase domains. The first papain-like cysteine proteinase domain (PCP-1) has been shown to autocleave the N-terminal peptide, p28 from the putative MHV polymerase polyprotein. This study was designed to further understand the p28 processing event, focusing on the catalytic sites of PCP-1 and the cleavage site recognized by PCP-1 to release p28.

The predicted PCP-1 domain at the 5'-end 3.5-4.3 kb of gene 1 was confirmed by in-frame deletion analysis. In addition, the proteolytic activity of PCP-1 was shown to be partially inhibited by cysteine proteinase inhibitors but not by serine, aspartic, or metallo-proteinases inhibitors, supporting the proposal that this proteinase is a papain-like cysteine proteinase. By site-specific mutagenesis, Cys-1137 and His-1288 were identified to be the catalytic sites of PCP-1. Thus, MHV PCP-1 is indeed a new member of the family of papain-like cysteine proteinases.

To determine the p28 cleavage site recognized by MHV PCP-1, the radiolabeled in vitro cleaved product adjacent to p28 was subjected to N-
terminal sequencing, and the result indicated that the cleavage occurred between Gly-247 and Val-248. By extensive site-specific mutagenesis of amino acid residues surrounding this cleavage site, Gly-247(P1) and Arg-246 (P2) were characterized as major determinants for cleavage site recognition by MHV PCP-1. In addition, some surrounding residues such as Arg-243 (P5) are probably required to maintain an important conformation of the cleavage site.

This work is an important step towards the understanding of the complete proteolytic processing pathway of MHV polymerase polyprotein. Future studies can be directed toward characterizing additional putative cleavage sites for MHV PCP-1 and determining what effect blocking the cleavage of p28 will have on coronavirus RNA replication.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ACKNOWLEDGMENTS</strong></td>
<td>iii</td>
</tr>
<tr>
<td><strong>ABSTRACT</strong></td>
<td>iv</td>
</tr>
<tr>
<td><strong>LIST OF FIGURES</strong></td>
<td>viii</td>
</tr>
<tr>
<td><strong>LIST OF TABLES</strong></td>
<td>x</td>
</tr>
<tr>
<td><strong>LIST OF ABBREVIATIONS</strong></td>
<td>xi</td>
</tr>
<tr>
<td><strong>Chapter</strong></td>
<td></td>
</tr>
<tr>
<td><strong>I. INTRODUCTION</strong></td>
<td></td>
</tr>
<tr>
<td>Coronavirus and their replication</td>
<td>1</td>
</tr>
<tr>
<td>Murine Coronavirus polymerase polyprotein</td>
<td>6</td>
</tr>
<tr>
<td>Proteolytic processing of the murine coronavirus polymerase polyprotein</td>
<td>10</td>
</tr>
<tr>
<td><strong>II. MATERIALS AND METHODS</strong></td>
<td></td>
</tr>
<tr>
<td>Plasmid DNA</td>
<td>14</td>
</tr>
<tr>
<td>Plasmid DNA preparation</td>
<td>14</td>
</tr>
<tr>
<td>Purification of linearized plasmid DNA</td>
<td>19</td>
</tr>
<tr>
<td>In vitro T7 RNA polymerase transcription</td>
<td>20</td>
</tr>
<tr>
<td>In vitro translation in rabbit reticulocyte lysates</td>
<td>21</td>
</tr>
<tr>
<td>Coupled in vitro transcription and translation</td>
<td>24</td>
</tr>
<tr>
<td>TCA precipitation assay to determine incorporation of radioactivity in the in vitro translated protein product</td>
<td>25</td>
</tr>
<tr>
<td>Anti-p28 serum and radioimmunoprecipitation</td>
<td>26</td>
</tr>
<tr>
<td>Polyacrylamide gel electrophoresis and autoradiograph</td>
<td>27</td>
</tr>
<tr>
<td>Proteinase inhibition assay</td>
<td>28</td>
</tr>
</tbody>
</table>
Chapter

Purification of DNA from low-melting agarose gels .............. 29
Construction of plasmid pS-N27 ............................... 30
Single-strand phagemid DNA preparation .................... 33
Oligonucleotide synthesis and phosphorylation ............... 35
Site-directed mutagenesis reaction ........................... 41
Plasmid DNA sequencing ....................................... 44
Preparation of radiolabeled protein sample
and N-terminal microsequencing .............................. 45

III. RESULTS ................................................................. 49
Effect of deletion mutation on proteinase activity ............ 49
Effect of proteinase inhibitors on the autoproteolytic
activity of MHV PCP-1 ........................................ 53
Effects of site-specific mutations of the putative MHV
PCP-1 catalytic residues Cys-1137 and His-1288
on proteolytic processing of p28 ............................ 60
Determining the amino-terminal amino acid sequence
of the protein adjacent to p28 .............................. 69
Effect of site-specific mutations surrounding the
cleavage site on proteolytic processing of p28 ............ 73

IV. DISCUSSION ............................................................ 107
LITERATURE CITED .................................................. 119
VITA ................................................................. 128
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Model of &quot;leader-primed&quot; transcription of coronavirus mRNAs</td>
<td>4</td>
</tr>
<tr>
<td>2.</td>
<td>Schematic diagram of murine coronavirus gene 1 and its putative functional domains</td>
<td>8</td>
</tr>
<tr>
<td>3.</td>
<td>Schematic diagram of corresponding gene 1 sequence in the plasmids of pT7-NBgl family</td>
<td>15</td>
</tr>
<tr>
<td>4.</td>
<td>Agarose gel electrophoresis analysis of the in vitro transcribed pT7-N27 RNA</td>
<td>22</td>
</tr>
<tr>
<td>5.</td>
<td>Schematic diagram of MHV-JHM gene 1 cDNA in plasmid pS-N27 and its encoded polypeptide and autoproteolytically processed proteins</td>
<td>31</td>
</tr>
<tr>
<td>7.</td>
<td>Schematic diagram of Altered Sites in vitro mutagenesis procedure</td>
<td>42</td>
</tr>
<tr>
<td>8.</td>
<td>Autoradiogram of representative sequencing result from screening plasmid DNA encoding mutations surrounding the p28 cleavage site</td>
<td>46</td>
</tr>
<tr>
<td>9.</td>
<td>Effect of deletion mutation on proteolytic processing of p28</td>
<td>51</td>
</tr>
<tr>
<td>10.</td>
<td>Inhibition of the proteolytic activity of MHV PCP-1 by proteinase inhibitors</td>
<td>55</td>
</tr>
<tr>
<td>11.</td>
<td>The amino acid sequence of the proposed first papain-like cysteine proteinase domain (PCP-1) of MHV-JHM gene 1</td>
<td>61</td>
</tr>
<tr>
<td>12.</td>
<td>Effect of site-specific mutation at the putative MHV PCP-1 catalytic residue Cys-1137 on proteolytic processing of p28</td>
<td>65</td>
</tr>
<tr>
<td>13.</td>
<td>Effect of site-specific mutation at the putative MHV PCP-1 catalytic residue His-1288 on proteolytic processing of p28</td>
<td>67</td>
</tr>
<tr>
<td>14.</td>
<td>Strategy of N-terminal protein sequencing to identify the p28 cleavage site</td>
<td>71</td>
</tr>
</tbody>
</table>
15. Schematic diagram of the autoradiogram of radiolabeled protein electrotransferred to the PVDF membrane .......... 74
16. N-terminal sequence analysis of radiolabeled p100 ............. 76
17. Schematic diagram of the putative p28 cleavage site .......... 78
18. Nomenclature of the amino acid residues surrounding a cleavage site for a proteinase ......................... 80
19. *In vitro* translation protein products of wild type pS-N27 and mutants of Gly-247 (P1) .................................. 83
20. *In vitro* translation protein products of wild type pS-N27 and mutants of Arg-246 (P2) .................................. 86
21. *In vitro* translation protein products of wild type pS-N27 and mutants of Tyr-245 (P3) .................................. 88
22. *In vitro* translation protein products of wild type pS-N27 and mutants of Gly-244 (P4) .................................. 90
23. *In vitro* translation protein products of wild type pS-N27 and mutants of Val-248 (P1') .................................. 92
24. *In vitro* translation protein products of wild type pS-N27 and mutants of Lys-249 (P2') .................................. 94
25. *In vitro* translation protein products of wild type pS-N27 and mutants of Pro-250 (P3') .................................. 96
27. Immunoprecipitation analysis of p28 from *in vitro* translation products of wild type pS-N27 and mutants at positions from P5 to P3' of the cleavage site ..................................... 101
28. Effect of site-specific mutation at the previously predicted p28 cleavage site Tyr-257 and Gly-258 on proteolytic processing of pS-N27 translation products ........................................... 104
29. Comparison of known cleavage sites for viral papain-like cysteine proteinases .............................................. 112
LIST OF TABLES

Table

1. Oligonucleotides used to generate site-specific mutations in the MHV PCP-1 proteinase domain ........................................ 38

2. Oligonucleotides used for site-directed mutagenesis of MHV-JHM p28 cleavage site and flanking amino acid sequence ........................................ 39

3. Effect of proteinase inhibitors on the autoproteolytic activity of the MHV PCP-1 proteinase ............................. 58
LIST OF ABBREVIATIONS

3C-Pro  poliovirus 3C like serine proteinase domain
οC  degree of centigrade
approxim.  approximately
DEPC  diethyl pyrocarbonate
dH2O  deionized water
DNA  deoxyribonucleic acid
dNTP  deoxynucleotide triphosphate
DTT  dithiothreitol
EDTA  disodium ethylenediaminetetra-acetate
g  gram
IBV  infectious bronchitis virus
kb  kilobase
kDa  kilodalton
l  liter
LB  Luria broth
M  molar
MHV  mouse hepatitis virus
μCi  micro Curie
μg  microgram
μl  microliter
μM  micromolar
xg  multiplication of the force of gravity

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Three-Letter Code</th>
<th>One-Letter Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>alanine</td>
<td>Ala(A)</td>
<td>A</td>
</tr>
<tr>
<td>arginine</td>
<td>Arg(R)</td>
<td>R</td>
</tr>
<tr>
<td>asparagine</td>
<td>Asn(N)</td>
<td>N</td>
</tr>
<tr>
<td>aspartic acid</td>
<td>Asp(D)</td>
<td>D</td>
</tr>
<tr>
<td>cysteine</td>
<td>Cys(C)</td>
<td>C</td>
</tr>
<tr>
<td>glutamine</td>
<td>Gln(Q)</td>
<td>Q</td>
</tr>
<tr>
<td>glutamic acid</td>
<td>Glu(E)</td>
<td>E</td>
</tr>
<tr>
<td>glycine</td>
<td>Gly(G)</td>
<td>G</td>
</tr>
<tr>
<td>histidine</td>
<td>His(H)</td>
<td>H</td>
</tr>
<tr>
<td>isoleucine</td>
<td>Ile(I)</td>
<td>I</td>
</tr>
<tr>
<td>leucine</td>
<td>Leu(L)</td>
<td>L</td>
</tr>
<tr>
<td>lysine</td>
<td>Lys(K)</td>
<td>K</td>
</tr>
<tr>
<td>methionine</td>
<td>Met(M)</td>
<td>M</td>
</tr>
<tr>
<td>phenylalanine</td>
<td>Phe(F)</td>
<td>F</td>
</tr>
<tr>
<td>proline</td>
<td>Pro(P)</td>
<td>P</td>
</tr>
<tr>
<td>serine</td>
<td>Ser(S)</td>
<td>S</td>
</tr>
<tr>
<td>threonine</td>
<td>Thr(T)</td>
<td>T</td>
</tr>
<tr>
<td>tryptophan</td>
<td>Trp(W)</td>
<td>W</td>
</tr>
<tr>
<td>tyrosine</td>
<td>Tyr(Y)</td>
<td>Y</td>
</tr>
<tr>
<td>valine</td>
<td>Val(V)</td>
<td>V</td>
</tr>
</tbody>
</table>
Coronaviruses and their replication

Coronaviruses are a family of large, enveloped positive-strand (+) RNA viruses which are widespread in nature (Wege et al., 1982; Siddell et al., 1983). This family of viruses consists of 11 members, including avian infectious bronchitis virus (IBV), mouse hepatitis virus (MHV), bovine coronavirus (BCV), and human coronavirus (HCV). Coronaviruses are of clinical and economical importance because they cause a variety of diseases, mainly in respiratory, gastrointestinal and neurological systems, such as common cold in humans; enteritis in cows, pigs, turkeys, cats and mice; and encephalitis in mice and cats. Mouse hepatitis virus (MHV), a prototype coronavirus, has been extensively studied to understand the molecular biology of coronaviruses (Spaan et al., 1988; Lai, 1990). The neurotropic strain of mouse hepatitis virus (MHV-JHM) can induce demyelination after injection into the brains of mice and serves as a model system for studying demyelination diseases such as multiple sclerosis (Weiner, 1973; Knobler et al., 1982; Fleming et al., 1987).

The coronavirion is a crown-like particle with a lipid-bilayer membrane outside and a helical encapsidated RNA genome inside (Siddell et al., 1983). These enveloped viruses are composed of three or four
structural proteins: the petal-shaped spike protein (S) projecting from the surface of the envelope giving coronavirus its distinctive morphology, the membrane embedded matrix protein (M), the hemagglutinin esterase (HE, expressed only in some virus strains) and the nucleocapsid protein (N) which surrounds the virion RNA. The viral genomic RNA is a single-strand positive-sense RNA molecule which is capped at the 5' end and polyadenylated at the 3' end (Yogo et al., 1977; Lai and Stohlman, 1978). In MHV, the viral RNA contains 8 genes and is about 31 kilobases in length, thus making it the largest viral RNA genome identified to date (Pachuk et al., 1989; Lee et al., 1991). The virion RNA itself carries all the genetic information for coronavirus replication and is infectious when it is transfected into eukaryotic cells (Wege et al., 1978).

Coronavirus replication occurs exclusively in the cytoplasm of the infected cells (Brayton et al., 1981; Wilhelmsen et al., 1981). Upon viral infection, virions use the spike proteins to attach to the receptors on the surface of the target cell and penetrate the cell by immediate membrane fusion or endocytosis. When viral genomic RNA is released into the host cell cytoplasm, the first event is the translation of the viral specific RNA-dependent RNA polymerase from the incoming genome. Using the genome as the template, this polymerase transcribes a genomic-length negative-strand RNA, which in turn serves as the template for the transcription of genomic RNA and seven subgenomic viral messenger RNAs. These mRNAs have the same 3' end, extend in the 5' direction for various lengths and contain an identical leader sequence (approximately 72 nucleotides) at
the 5' end (Lai and Stohlman, 1978; 1981; Lai et al., 1981; 1982a; 1983; Spaan et al., 1983). For each of these nested-set mRNAs, only the 5' most gene is translated into viral proteins (Siddell et al., 1983). For instance, although the genomic mRNA contains a total of 8 genes, only gene 1 at its 5' end is translated which results in the synthesis of the polymerase (Leibowitz et al., 1982b; Denison and Perlman, 1986). Following protein synthesis and viral replication, virions are assembled in the rough endoplasmic reticulum of the cytoplasm and bud out of the cell. Thus, like many positive-strand RNA viruses, coronavirus RNA is infectious and the viral replication cycle takes place in the cytoplasm of the infected cell.

Two unique features have been observed during coronavirus replication: (1) discontinuous transcription and (2) high frequency RNA recombination. The identical leader sequence (approximately 72 nucleotides) at the 5' end of the seven subgenomic RNAs is derived from the 5' end of the genomic RNA by a discontinuous transcription mechanism (Baric et al., 1983; Spaan et al., 1983; Lai et al., 1983; 1984; Makino et al., 1986b; 1988; 1989; Baker and Lai, 1990). For this unique discontinuous transcription, several models have been proposed, among which "leader-primed" transcription is the most favored and has been supported by a large body of evidence (Baric et al., 1983; Spaan et al., 1983; Lai et al., 1983; 1984; Makino et al., 1986b; 1988; 1989; Baker and Lai, 1990). According to this model (Fig. 1), free leader RNAs (approximately 72 nucleotides) are synthesized from the negative RNA template during coronavirus replication. These leader RNAs disassociate from the template and
Fig. 1. Model of "leader-primed" transcription of coronavirus mRNAs. A. "Leader-primed" transcription in the coronavirus RNA replication cycle. The viral genomic RNA (+) released in the infected cell is first used to synthesize the viral specific RNA-dependent RNA polymerase. This polymerase transcribes a genomic-length negative RNA (-) from the viral genomic RNA (+) and uses it as the template for the replication of a nested-set genomic and subgenomic mRNAs (+) shown below. A nested-set subgenomic negative RNAs (-) are also indicated as the potential templates for the transcription of the subgenomic mRNAs (+). The solid squares represent leader RNA. B. An example of the intergenic sequences. Partial sequences of a free leader RNA is denoted above the template of a genomic-length negative RNA. The intergenic sequence preceding gene 6 is shown to be hybridized to the free leader RNA by homologous sequences. The arrow indicates the consensus sequence UCUAAC in the leader RNA. (The figure was modified from Lai, 1990).
ORF 1a  PCP-1  PCP-2  ORF 1b
papain-like cysteine proteinases polio 3C-like serine proteinase polymerase motif helicase domain
Cys-rich transmembrane zinc-finger domains domains motif
rehybridize to the homologous sequence (several repeats of AGAUU) in the region between each gene. The leader RNA serves as a primer for the transcription of each of the seven subgenomic mRNAs. Transcription of these subgenomic mRNAs initiates at different intergenic sites and extends to the 3' end along the negative RNA template. This results in a nested-set structure of these genomic and subgenomic mRNAs.

The second unique feature, high frequency RNA recombination, was observed during mixed infection of different coronaviruses (Keck et al., 1987; 1988a; 1988b; Makino et al., 1986b; 1987). In the mixed infection, viral RNA transcription intermediates may dissociate from one template to bind to another viral template at the homologous sequence and be incorporated into genomic RNA in the subsequent RNA replication.

The complex viral transcription and replication strategy of coronavirus indicates a viral RNA polymerase with unusual properties.

**Murine coronavirus polymerase polyprotein**

The coronavirus encoded RNA polymerase is known to be responsible for replication of the viral negative RNA, synthesis of genomic and subgenomic RNAs, and replication of virion RNA (Brayton et al., 1982; 1984; Lai et al., 1982b; Leibowitz et al., 1982a). In addition, coronavirus polymerase is also known to transcribe the small leader RNAs which have been detected in the coronavirus infected cells (Lai et al., 1983; 1984; Baric et
al., 1983). To better understand the functional activities of coronavirus polymerase, we must elucidate the mechanism of MHV RNA polymerase expression and processing.

The murine coronavirus RNA polymerase is encoded by the 5' most gene of the virion RNA, termed gene 1. Gene 1 has been completely sequenced for coronavirus IBV (Bournell et al., 1987), MHV-JHM (Lee et al., 1991), MHV-A59 (Pachuk et al., 1989; Bredenbeek et al., 1990; Bonilla et al., 1994) and HCV (Herold et al., 1993). In the MHV-JHM virus strain, gene 1 is 22 kb in length and encompasses about two thirds of the entire viral genome (Lee et al., 1991). This gene is known to contain two overlapping open reading frames designated ORF1a (13 kb) and ORF1b (9 kb) which can be translated into a single polyprotein greater than 750 kDa via a ribosomal frameshifting mechanism (Brierly et al., 1987; 1989; Bredenbeek et al., 1990; Lee et al., 1991). Computer analysis of the deduced amino acid sequence of this large polymerase polyprotein suggests diverse functional domains, such as a polymerase motif, a helicase domain, a zinc-finger motif in ORF1b, transmembrane regions, cysteine-rich domains, and three proteinase domains in ORF1a (Gorbalenya et al., 1989; 1991; Bredenbeek et al., 1990; Lee et al., 1991) (see Fig. 2). The three proposed proteinase domains, two papain-like cysteine proteinases (PCP-1 and PCP-2) and a poliovirus 3C-like serine proteinase (3C-Pro), are approximately 3.5-4.2 kb, 5.3-6.1 kb, and 10.2-10.9 kb from the 5' end of the gene 1 respectively (Lee et al., 1991). The large size of the polymerase polyprotein and the presence of potentially active proteinase domains suggest that the
Fig. 2. Schematic diagram of murine coronavirus gene 1 and its putative functional domains. (The figure was modified from Lee et al., 1991).
A

Leader RNA-Primed Transcription

(+)

5' \[\text{Leader RNA}

3' \[\text{Primed Transcription}

(-)

3' \[\text{Leader RNA}

5' \[\text{Primed Transcription}

B

(-) 3' UACUAAUJAAGAUUGUUGUUAAC 5'

1 2 3 4 5 6 7 5' (-)
putative polymerase polyprotein undergoes extensive proteolytic processing to release multiple mature and functional protein products.

**Proteolytic processing of the murine coronavirus polymerase polyprotein**

Since gene 1 is expressed at a very low level and coronavirus does not inhibit host cell protein synthesis efficiently, it is very difficult to investigate viral gene expression *in vivo* without specific antibodies. Early studies of gene 1 translation products were mainly done in the rabbit reticulocyte lysate (RRL) cell free protein-synthesizing system (Leibowitz *et al.*, 1982b; Denison and Perlman, 1986). The first protein to be identified as a proteolytic processing product of the polymerase polyprotein was a 28 kDa protein, designated p28. Denison and Perlman (1986) demonstrated that *in vitro* translation of MHV-A59 genomic RNA isolated from virus-infected cells produced two prominent proteins, p28 and p220 (Denison and Perlman, 1986). The appearance of p28 could be inhibited by the addition of the proteinase inhibitor ZnCl$_2$, suggesting that p28 was processed from a precursor polyprotein (Denison and Perlman, 1986). p28 was shown to be the N-terminal peptide processed from the gene 1 polyprotein and also detected in MHV infected cells (Denison and Perlman, 1986; 1987). By *in vitro* T7 RNA polymerase transcription and *in vitro* translation in rabbit reticulocyte lysates of a plasmid containing the 5'-end 2 kb cDNA of gene 1, p28 was confirmed as the amino-terminal portion of the putative polymerase polyprotein (Soe *et al.*, 1987). Furthermore, by translating RNAs which represented up to 5.3 kb of the 5'-end of the MHV genome,
Baker et al. showed that the processing of p28 was an autoproteolytic event in which a downstream sequence about 4 kb from the 5' end of gene 1 was required for the cleavage of p28 (Baker et al., 1989). Significantly, this region overlaps with the first papain-like cysteine proteinase (PCP-1) domain predicted from sequence analysis (Gorbalenya et al., 1991, Lee et al., 1991), suggesting that PCP-1 is the proteinase which cleaves the p28 protein.

In addition to p28, several other protein products ranging from 50 kDa to 290 kDa in size have been detected from the in vitro translation of MHV genomic RNA or in the infected cells using antiserum directed against viral peptides or fusion proteins (Denison et al., 1991; 1992). However, the proteinases responsible for their processing are largely unknown. Our knowledge of how these proteins are processed from the polymerase polyprotein and the relationships between them is also incomplete. Currently, the first papain-like cysteine proteinase (PCP-1) is the only proteinase domain among the three proposed proteinases whose activity has been directly demonstrated (Baker et al., 1989). Although the only known function of PCP-1 is to mediate a cis-cleavage of p28, it is highly possible that PCP-1 cleaves at additional sites in the coronavirus polymerase polyprotein. The study of p28 processing will enrich our knowledge of proteolytic processing of the polymerase polyprotein and help to predict additional processing events.
The proteolytic processing of the polymerase polyprotein may play an important role in regulating viral replication, as has been shown in several other viruses such as poliovirus and Sindbis virus (Dewalt and Semler, 1987; Lemm and Rice, 1993; Shirako and Strauss, 1994). In addition, proteolytic processing may activate diverse functions of the polymerase polyprotein by inducing conformational changes. In terms of the understanding of how coronavirus polymerase works, it is possible that the polymerase polyprotein undergoes proteolytic cleavages at several sites to release multiple functional units. The functional subunits may still associate together as a working complex for coronaviral RNA replication. Understanding the proteolytic processing of gene 1 may ultimately help us to identify individual polymerase subunits and to understand their functions. Additionally, viral proteinases are potential targets for specific antiviral agents since viral replication can be interrupted by inhibiting the essential proteolytic processing events to inactivate fundamental enzymes for viral replication. Such work has already begun for human immunodeficiency virus (HIV), and results suggest that intervention with proteinase inhibitors may become a successful treatment for Acquired Immune Deficiency Syndrome (AIDS) (McQuade et al., 1990).

The goal of this Ph.D. dissertation research is to study proteolytic processing of the murine coronavirus polymerase polyprotein, especially focusing on the p28 processing event. Previous studies lead to the hypothesis that the MHV PCP-1 proteinase domain, which acts to cleave at the amino terminal region of polymerase polyprotein to release p28,
probably belongs to the family of cysteine proteinases (with plant proteinase papain as the prototype). The following questions were addressed in this research: 1) What are the catalytic residues of MHV PCP-1? 2) What is the cleavage site recognized by MHV PCP-1 to release p28?

To address these questions, I have used the well developed T7 RNA polymerase transcription/RRL translation system to express a variety of mutant polyproteins, and I have assayed the effect of these mutations on proteolytic processing of p28. I have characterized the proteinase activity and the catalytic sites, showing that PCP-1 is a new member of the family of papain-like cysteine proteinases. Additionally, I have identified the cleavage site by partial N-terminal protein sequencing, and defined the recognition motif of the cleavage site by site-directed mutagenesis.
CHAPTER II
MATERIALS AND METHODS

Plasmid DNA

A T7 RNA polymerase transcription plasmid pT7-NBgl, which contains the MHV-JHM genomic cDNA sequence from nucleotide 187 (Nar I site) to 5273 (Bgl II site), and its deletion plasmids pT7-NAc, pT7-N27 and pT7-N29 have been previously described (see Fig. 3, Baker et al., 1989; 1993). pT7-NAc is derived from pT7-NBgl with a deletion from nucleotide 4434 (Acc I site) to 5273 (Bgl II site). pT7-N27 is derived from pT7-NAc with a deletion from nucleotide 2068 (Nsi I site) to 2811 (Kpn I site). T7-N29 is derived from pT7-NAc with deletion from nucleotide 1167 (SnaB I site) to 2811 (Kpn I site).

Plasmid DNA preparation

Large scale preparations and minipreparations of plasmid DNAs were isolated by the alkaline lysis method basically as described by Maniatis et al. (1982). For large scale plasmid preparation, a single colony was inoculated into 5 ml Luria broth (LB, 10g/l bacto-trypotone, 5g/l bacto-yeast extract and 10g/l NaCl, pH 7.0) with 50 mg/ml ampicillin (Amp) and grown for 5-6 hours at 37°C with vigorous shaking (250 rpm). Two and one half ml of this small culture was inoculated into 250 ml LB/Amp (50 µg/ml) and incubated with shaking for 14-16 hours at 37°C. The bacteria were
Fig 3. Schematic diagram of corresponding gene 1 sequence in the plasmids of pT7-NBgl family. A partial restriction map of the 5'-end 5.3 kb of MHV-JHM genomic cDNA is shown to scale (Nar I:187, SnaB I:1167, Nsi I:2068, Kpn I:2811, Acc I:4434, Bgl II:5273). The translation initiation codon (AUG) is at nucleotide 215. The p28 region, which is released by the autoproteinase activity, is denoted by the stippled box. The putative first papain-like cysteine proteinase (PCP-1) domain is indicated by the hatched box. The putative p28 cleavage site at approximately 1 kb from the 5' end is indicated by an arrow. The sizes of encoded precursor polyprotein and potential processed products are indicated in the parenthesis.
harvested by centrifugation at 7 k rpm for 10 min at 4°C in a Sorvall GSA rotor. The pellet was resuspended in 4 ml of solution I (50 mM glucose, 100 mM EDTA, 28 mM Tris-HCl, pH 8.0) and incubated for 15 min at room temperature. Bacteria were lysed by adding 7.5 ml of solution II (0.2 N NaOH, 1% SDS, freshly made from 10 N NaOH and 10% SDS), and incubated for 10 min on ice. The cell lysates were neutralized by adding 6 ml of solution III (5 M KOAc, pH 4.8) and inverting the tubes and incubated for 10 min on ice. The cell debris was pelleted by centrifugation for 30 min at 12 k rpm at 4°C in a Sorvall SS34 rotor. The supernatant (about 16 ml) was transferred to a Corex tube (30 ml capacity), mixed with 0.6 volume (11 ml) of isopropanol and incubated for 15 min at room temperature. The nucleic acids were pelleted by centrifugation for 10 min at 10 k rpm in a Sorvall SS34 rotor and the pellet was dried in a Savant Speed Vac. This nucleic acid pellet was resuspended in 8 ml of TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0) and treated with 40 µl RNase A (10 mg/ml stock solution) for 30 min at 42°C to degrade tRNA (transfer RNA). The supercoiled plasmid DNA was purified by equilibrium centrifugation in cesium chloride (CsCl) ethidium bromide gradients. For each ml of DNA solution, 1 g of solid CsCl was added and the final density of the solution should be 1.55 g/ml (refractive index = 1.3860). When the salt is completely dissolved, 250 µl ethidium bromide (10 mg/ml stock solution, final concentration 312.5 µg/ml) was added. The mixed sample was centrifuged for 14-16 hours at 55 k rpm in a Beckman VTi 65 rotor. The band of the supercoiled plasmid DNA was collected by side puncture with a 16 gauge needle. The ethidium bromide was removed from the DNA by repeated
extraction with an equal volume of deionized water-saturated butanol until the pink color disappeared. The CsCl was removed by overnight dialysis at 4°C with at least two changes of 0.5 X TE buffer. Finally, the DNA solution was collected and stored at 4°C. The yield and purity of plasmid DNA was determined by measuring OD at 260 nm and OD at 280 nm in a UV spectrophotometer. The plasmid DNA was analyzed by electrophoresis in a 1% agarose gel. The DNA (0.2 µg) was diluted in 6 X loading buffer (0.25% bromophenol blue, 0.25% xylene cyanol FF and 30% glycerol) and loaded onto a 1% mini-agarose gel prepared in 1 X TBE buffer (89 mM Tris-HCl, 89 mM boric acid, 2 mM EDTA, pH 8.0) containing 0.5 µg/ml ethidium bromide. The DNA was analyzed by electrophoresis at 100 Volts (V) for 30 min. Lambda DNA digested with Hind III (Gibco BRL) was loaded as the size marker.

Minipreparations of plasmid DNA were carried out using the Magic™ Minipreps DNA Purification System (Promega) following the manufacturer's instructions. Briefly, 3-5 ml of the overnight cultured bacteria cells were pelleted in an Eppendorf microcentrifuge (1 min at high speed) and the pellet was resuspended in 200 µl of Cell Resuspension Solution (5m mM Tris-HCl, pH 7.5, 10 mM EDTA, pH 8.0, 100 µg/ml RNase A). The cells were lysed by adding 200 µl of Cell Lysis Solution (0.2 M NaOH, 1% SDS) and inverting the tube several times. To the cell lysates, 200 µl of Neutralization Solution (2.55 M KOAc, pH 4.8) was added and mixed by inverting the tubes. After centrifugation for 5 min at 12,000 X g, the supernatant was transferred to another tube and mixed with 1 ml of the
Magic Minipreps DNA Purification Resin. The DNA bound resin was transferred into a Magic Minicolumn-attached 3 ml luer-lock syringe on a vacuum manifold and concentrated in the column by applying the vacuum. The column was washed with 2 ml Column Wash Solution (200 mM NaCl, 20 mM Tris-HCl, 5 mM EDTA, pH 8.0) prediluted with 1.4 volume of 95% ethanol. The residual Column Wash Solution was removed by applying continuous vacuum for 2-5 min, followed by spinning the column for 20 sec in the microcentrifuge. The DNA product was eluted into a sterile microcentrifuge tube by adding preheated (65-70°C) dH2O to the column and spinning for 20 sec. Two μl of eluted DNA was analyzed by electrophoresis in a 1% mini-agarose gel and quantitated by UV spectrophotometry as described above.

**Purification of linearized plasmid DNA**

Plasmid DNA linearized by EcoR I digestion was used as the DNA template for an *in vitro* T7 RNA polymerase transcription reaction or a coupled *in vitro* transcription/translation reaction (TNT lysate, Promega). One to four μg of plasmid DNA was digested by 5-10 μl of EcoR I (Gibco BRL) in 50 mM Tris-HCl (pH 8.0), 10 mM MgCl₂, and 100 mM NaCl in a total volume of 100 μl at 37°C for 2-4 hours. (Long incubation times or high enzyme concentrations were avoided because of the potential star activity of EcoR I). Ten μl of the reaction was removed to check the completion of the digestion by electrophoresis in a 1% mini-agarose gel. If a single band of plasmid DNA of the expected size was detected, the reaction was subsequently treated with protease K (final concentration 50 μg/ml, Sigma)
in 5 mM EDTA and 0.5% SDS for 30 min at 56°C to digest the restriction enzymes. The final volume was brought up to 200 µl with dH2O and the linearized plasmid DNA was extracted with an equal volume of phenol:chloroform (1:1) once and then an equal volume of chloroform:isoamyl alcohol (24:1) once. Eight µl of 5 M NaCl (final concentration 200 mM) and 2.5 volume of ethanol were added to the supernatant and mixed by vortexing. The mixture was frozen at -70°C for 20 min and centrifuged for 10 min at 12,000 X g to precipitate DNA. The DNA pellet was gently rinsed with 1 ml of cold 70% ethanol and dried in a Savant Speed Vac. The final DNA pellet was resuspended in 20 µl of DEPC-dH2O (diethyl pyrocarbonate treated nuclease-free water) and 2 µl was removed for analysis by electrophoresis on a 1% mini-agarose gel as described above.

**In vitro T7 RNA polymerase transcription**

Plasmid DNA was transcribed *in vitro* with a T7 RNA polymerase (Promega) specific for a T7 promoter (Tabor and Richardson, 1985). The transcription reaction was performed in a total volume of 50 µl with 1 µg of linearized plasmid DNA in 40 mM Tris-HCl (pH 7.9), 6 mM MgCl2, 10 mM DTT (dithiothreitol), 2 mM spermidine, 10 mM NaCl, 0.5 mM of ATP, CTP, and UTP, 0.05 mM GTP, 0.25 mM m7G(5')ppp(5')G cap analog (NEB), 10 u of T7 RNA polymerase (Promega), and 40 u of RNasin ribonuclease inhibitor (Promega) and DEPC-dH2O at 37°C for 1 hour (Baker *et al.* 1989; 1993). The reaction mixture was treated with 2 u of RQI DNase (Promega) at 37°C for 30 min to degrade the DNA template. The final volume was
brought up to 200 µl with DEPC-H₂O and extracted with an equal volume of phenol:chloroform (1:1) and then an equal volume of chloroform:isoamyl alcohol (24:1). The upper aqueous phase was transferred to a sterile microcentrifuge tube and mixed with 0.1 volume of 3M NaOAc and 2.5 volume of ethanol and cooled to -70°C for 20 min. The mixture was centrifuged at 12,000 X g for 10 min. The pellet (normally milky white) was gently rinsed with 1 ml of cold 70% ethanol, dried and resuspended in 100 µl DEPC-H₂O. The resuspension solution was mixed with 100 µl 5M NH₄OAc (final concentration 2.5 M) and 2.5 volumes of ethanol and frozen for 20 min at -70°C. This second precipitation removes free nucleotides. The final pellet was gently rinsed in cold 70% ethanol, dried in a Savant Speed Vac and resuspended in 20 µl of DEPC-dH₂O. The yield was quantitated by measuring OD at 260 nm in a UV spectrophotometer. Two µl of RNA was diluted in 6 X loading buffer and analyzed in a 1.2% agarose gel (prepared in filtered 1 X TBE buffer containing 0.1-0.5 µg/ml ethidium bromide) at 60 V for 1-2 hour(s) (See Fig. 4). Lambda DNA Hind III fragments were loaded as the size marker. The remaining RNA was stored at -70°C until future use.

In vitro translation in rabbit reticulocyte lysates

Translation of in vitro transcribed RNA by T7 RNA polymerase was performed in the nuclease treated rabbit reticulocyte lysates following manufacturer's suggestion (Promega). Two hundred ng of purified RNA transcripts were mixed with 10 µl of the nuclease treated rabbit reticulocyte lysates, 1 µl of 1 mM amino acid mixture (minus Met) and 2 µl of 35S-Met
Fig. 4. Agarose gel electrophoresis analysis of the *in vitro* transcribed pT7-N27 RNA. Plasmid pT7-N27 linearized with EcoR I was transcribed by a T7 RNA polymerase from a specific T7 promoter as described in the Materials and Methods. RNA from four standard reactions was combined and 2 µl RNA (approximately 200 ng) was analyzed by electrophoresis in a 1.2% agarose gel and visualized by staining with ethidium bromide. Lambda DNA digested with Hind III was used as the size marker.
(1,000 μCi/ml, Amersham) or Trans 35S-label (1,000 μCi/ml with 70% 35S-Met and 15% 35S-Cys, ICN) and DEPC-dH2O in a final volume of 25 μl and incubated in 10 μl of SDS, 0.1% momlin, 3% alcohol, 3% containing 50% formaldehyde, 40% glycerol, and 10% TEMED. N7 RNA coupled to the membrane was fixed by crosslinking the membrane with 0.5 μg of puromycin. 20 μl of acid-ethanol, and 1.5% of BSA-labeled polyA tail, 1.5% of the polyA tail of the sample. The membrane was then washed with 70% ethanol and air-dried. The membrane was then exposed to autoradiography for 1-2 days.
Coupled in vitro transcription and translation

Linearized plasmid DNA template was transcribed with T7 RNA polymerase and translated in rabbit reticulocyte lysates in a coupled transcription translation system (TNT lysates, Promega) following the manufacturer's protocol. For a standard reaction, approximately 0.5 µg of linearized plasmid DNA was incubated with 12.5 µl TNT rabbit reticulocyte lysate, 1 µl TNT reaction buffer, 0.5 µl T7 RNA polymerase (Promega), 20 u RNasin ribonuclease inhibitor (Promega), 1 µl of 1 mM Met-free amino acid mixture, and 2 µl of 35-S-Met (1,000 µCi/ml, Amersham) or Trans 35S-label (1,000 µCi/ml, 70% 35S-Met and 15% 35S-Cys, ICN) and DEPC-dH2O in a final volume of 25 µl for 90 min at 30øC. Two µl of the reaction were removed and precipitated by trichloroacetic acid (TCA) to measure incorporation of radioactivity. One to two µl of the reaction was analyzed by electrophoresis in a 10% SDS polyacrylamide gel and the remainder was stored at -70øC.
TCA precipitation assay to determine incorporation of radioactivity in the in vitro translated protein product

The incorporation of radioactivity into the protein products of an in vitro translation was measured by a trichloroacetic acid (TCA) precipitation assay (Sambrook et al., 1989). The simplified protocol for TCA protein precipitation assay was from Dr. Tom Gallagher (Loyola University of Chicago). Individual Whatman 3MM filter paper disks (2-2.3 cm in diameter) were spotted with 100 µl of Solution A [3% CAA (casamino acids, as a protein carrier), 1% SDS] and air-dried for 0.5-1 hour. Two µl of the translation reaction was loaded at the center of the filter disk and air-dried for 10 min. Two µl from an in vitro translation reaction containing no mRNA was spotted onto a filter paper disc as a control to determine the background incorporation. The sample-loaded filters were transferred into a capped plastic tube (50 ml capacity, 4-6 filters per tube) and washed in 45 ml of Solution B (10% TCA, 3% CAA, 1% SDS) by slow shaking in an orbital shaker for 30 min at room temperature. Solution B and all other used solutions containing radioactivity were carefully discarded into a radioactive liquid disposal container. Forty-five ml of Solution C (5% TCA, 1.5% CAA) was added and incubated in an 85°C shaking water bath for 20 min. The filters were washed twice for 5 min with 45 ml of 5% TCA and then 45 ml of 95% ethanol. The filters were rinsed twice with dH2O and air-dried for 0.5-1 hour. Each dried filter was transferred to a 5 ml plastic scintillation vial with 4 ml of biodegradable counting cocktail Econo-Safe™ (Research Products International Corp.) and counted in a Beckman LS 5801 scintillation counter using the USER 2 program. The net incorporation of
radioactivity in the synthesized protein of each translation reaction was calculated by subtracting background from the cpm (count per min).

**Anti-p28 serum and radioimmunoprecipitation**

Polyclonal anti-p28 serum was generated in a rabbit against a synthetic peptide representing amino acids 78 to 93 (NH3-R-D-I-F-V-D-E-D-P-Q-K-V-E-A-S-T-COOH) of the p28 protein (previously prepared in our laboratory, Baker et al., 1989). One and one half g of Protein A-Sepharose CL-4B beads (Pharmacia) were swollen in 15 ml of dH2O and washed 2-3 times in PBS buffer (8 g/l of NaCl, 0.2 g/l KCl, 1.44 g/l Na2HPO4 and 0.24 g/l KH2PO4) to remove any additives and stored at 4°C for future use. To avoid the loss of beads during washing, the bead suspension can be briefly centrifuged after each step. The desired amount of beads were removed prior to each use. The beads were spun down, washed once in RIPA buffer (0.5% Triton X-100, 0.1% SDS, 300 mM NaCl, 4 mM EDTA and 50 mM Tris-HCl) and resuspended 1:1 in RIPA buffer. Samples of the *in vitro* translation reaction containing equivalent amount of TCA-precipitable radioactivity (10-15 µl, 0.5-5 X 10⁵ cpm) were incubated with 3 µl of anti-p28 serum in 1 ml of RIPA buffer in a safe-lock microcentrifuge tube with head-to-tail rotation at 4°C. Thirty µl of a 1:1 suspension of Protein A-Sepharose beads in RIPA buffer was added and the incubation reaction was continued for an additional 4 hours at 4°C. The beads were pelleted by a brief (1-2 min) microcentrifugation and washed three times with 1 ml of RIPA buffer. The washed beads were suspended in 50 µl of 2 x Laemmlili sample buffer, heated at 100°C for 2 min, and microcentrifuged for 5 min. The
supernatant was loaded on a 10% polyacrylamide gel containing 0.1% SDS (SDS-PAGE) for electrophoresis.

Polyacrylamide gel electrophoresis and autoradiography

Direct \textit{in vitro} translation products or the immunoprecipitated proteins of \textit{in vitro} translation were analyzed by polyacrylamide gel electrophoresis as described by Maizel, 1971. A 10% acrylamide, 0.1% SDS discontinuous gel was prepared in a Hoefer SE 600 Vertical Slab Unit. Twenty ml of 10% resolving gel was prepared by mixing 8.4 ml of dH$_2$O, 5.0 ml of 4 X Lower Buffer (181.7 g/l Tris-base, 0.4% SDS, pH 8.8), 6.6 ml of 30:0.8% Acrylamide:N, N'-methylene-bisacrylamide, 60 µl of 10% APS (ammonium persulfate) and 7.5 µl of TEMED (N, N, N', N'-tetramethylethyl-enediamine). For 10 ml of 5% stacking gel, 6.65 ml of dH$_2$O, 1 ml of 50% glycerol, 1.30 ml of 4 X Upper Buffer (60.6 g/l, 0.4% SDS, pH 6.8), 1.10 ml of 30:0.8% acrylamide:N, N'-methylenebisacrylamide, 75 µl 10% APS, and 20 µl of TEMED were mixed together. The gel was run in the 1 X Running Buffer (12.0 g/l, Tris-base, 57.6 g/l glycine, 0.1% SDS) at constant volts (50-200 V) until the bromophenol blue dye reached the bottom of the gel. Following electrophoresis, the gel was removed and fixed for 30-60 min in 40% methanol-7% acetic acid solution with gentle shaking in an orbital shaker. After fixing, the gel was enhanced for 1.5 hours with Du Pont Entensify, dried and exposed to Kodak X-ray film at -70°C.
Proteinase inhibition assay

The proteinase inhibition assay was performed using pT7-N27 RNA (see Fig. 4) in rabbit reticulocyte lysates in the presence of a variety of proteinase inhibitors. The proteinase inhibitors tested include a general inhibitor ZnCl₂, four cysteine proteinase inhibitors: cystatin (Boehringer Mannheim Biochemica, BMB), Iodoacetamide (Sigma), N-ethyl-maleimide (Sigma) and E-64 (BMB), a cysteine and serine proteinase inhibitor leupeptin (BMB), a serine proteinase inhibitor APMSF [94-(amidinophenyl)methanesulfonfonylfluoride, BMB], an aspartic proteinase inhibitor pepstatin A (BMB), and a metallo-proteinase inhibitor EGTA ([ethylenebis(oxyethylenenitrilo)]-tetraacetic acid, BMB). Large quantities of RNA were synthesized in vitro as described in the in vitro transcription reaction. Four standard in vitro transcription reactions (50 µl) were set up and the final products of each reaction were resuspended in 20 µl of dH₂O and combined together to quantitate the concentration and yield of the RNA synthesized. The same amount of homogenized RNA (approx. 200 ng) was used in each translation reaction as described in the in vitro translation reaction in rabbit reticulocyte lysates. The proteinase inhibitor was added to the translation reaction (25 µl, final volume) prior to the incubation at 30°C and the translation reaction was carried out for 90 min. The translation products were analyzed by electrophoresis in a 10% SDS-polyacrylamide gel and autoradiography. The radioactivity of p28 (with 7 Met) and the precursor polyprotein (with 37 Met) was quantitated in a Betagen scanner. The efficiency of p28 processing was calculated after normalizing the differences in Met contents in the following formula: 37 X
The percentage of the inhibition was measured by comparing the decrease of p28 processing efficiency in the translation reaction with proteinase inhibitors to that without proteinase inhibitors. A wide range of concentrations of each inhibitor was tested to achieve an optimal working concentration. The optimal concentration produced significant inhibition of p28 processing but had no effect on the overall efficiency of translation reactions.

**Purification of DNA from low-melting agarose gels**

Low-melting agarose gel (SeaKem, FMC) with 1 cm wide wells were prepared in 1 X TAE (40 mM Tris-acetate, 1 mM EDTA) containing 0.1 µg/ml ethidium bromide. Restriction enzyme digested plasmid DNA was loaded onto the gel, fragments were separated by electrophoresis at 60 V for 3-4 hours and visualized on an ultraviolet light transilluminator. The specific DNA bands of interest were identified according the DNA marker (lambda DNA digested with Hind III fragments, Gibco BRL) and excised from the gel with a razor blade. The agarose gel slice containing DNA was submerged in 3 volume of TE buffer in one (or several) microcentrifuge tube(s) and heated at 70ºC for 2-5 min in a heating block. The melted agarose solution was homogenized by briefly vortexing and quickly frozen in a dry-ice ethanol bath (-70ºC) for 5-10 min. The mixture was thawed by tapping the tube vigorously against the lab bench. The solidified agarose beads were removed by centrifugation for 1-5 min. The agarose-free DNA in the supernatant was extracted with equal volumes of phenol:chloroform and chloroform:isoamyl alcohol and concentrated by ethanol precipitation.
The DNA pellet was washed with cold 70% ethanol and resuspended in 20 µl of TE buffer.

**Construction of plasmid pS-N27**

For site-directed mutagenesis, the MHV-JHM gene 1 sequence in pT7-N27 was inserted into plasmid pSELECT-1 (Altered Sites *in vitro* Mutagenesis System, Promega) to create a new recombinant plasmid pS-N27 (see Fig. 5). pSELECT-1 (Amp<sup>+</sup> and Tec<sup>+</sup>), a hybrid of pBR322 and pGEM-3Zf{(+), carries modified ampicillin gene (4 base frameshifting), a tetracycline resistance gene, a f1 replication origin and a polylinker region (Lewis and Thompson, 1990). Plasmid pT7-N27 contains MHV-JHM genomic cDNA sequence from nucleotide 187 to nucleotide 4434 with a 747 nucleotide deletion of nucleotide 2064-2810 under the control of a T7 promoter. This 3.5 kb region of gene 1 (with a translation start codon at nucleotide 215) has a capacity to encode a 128 kDa precursor polyprotein which is expected to autoproteolytically release p28 and a 100 kDa C-terminal protein (Baker *et al.*, 1993). Four µg of pT7-N27 was digested by Pvu II (upstream of the T7 promoter) and EcoR I (just downstream of the MHV-JHM region) in a volume of 200 µl. Five µg of pSELECT-1 was digested by Sma I and EcoR I (Both sites are located within the polylinker region) in a reaction of 200 µl. Completion of each digestion was checked by electrophoresis in a 1% agarose gel. The double digested pSELECT-1 DNA was extracted with equal volumes of phenol:chloroform (1:1) and chloroform:isoamyl alcohol (24:1) and concentrated by ethanol precipitation. The resuspended pSELECT-1/Sma I/EcoR I DNA was
**Fig. 5.** Schematic diagram of MHV-JHM gene 1 cDNA in plasmid pS-N27 and its encoded polypeptide and autoproteolytically processed proteins. The MHV-JHM gene 1 cDNA from an in-frame deletion construct, pT7-N27, was inserted into plasmid pSELECT-1 as described in the Materials and Methods. The new recombinant plasmid was named as pS-N27 (A). This 3.5 kb gene 1 clone, containing nucleotides from 187 (Nar I site) to 2063 (Nsi I site) and nucleotides from 2811 (Kpn I site) to 4431 (Acc I site), is the smallest deletion clone which still maintains active proteolytic processing of p28. The two overlapping reading frames (ORF) are shown at the top with only ORF1a drawn to scale. The three proposed protease domains (Lee et al., 1991) are indicated, among which PCP-1 (hatched box, between 3.5-4.2 kb) has been demonstrated to be active (Baker et al., 1989). Proteolytic processing of the polypeptide translated from RNA derived from plasmid pS-N27 (also pT7-N27) is shown at the bottom (B). The precursor polypeptide (128 kDa) synthesized from linearized pS-N27 (also pT7-N27) in the *in vitro* T7 RNA polymerase transcription and rabbit reticulocyte lysate translation system is autoproteolytically processed to two products: N-terminal p28 and C-terminal peptide (100 kDa) (Baker et al., 1993). The p28 region is indicated by the stippled box and PCP-1 domain by the hatched box. The arrow represents the putative cleavage site of p28.
A

MHV-JHM Gene 1 ORF

scale (kb)

5' 7' 7 8 10 12 14 16 18 20 22

ORF 1a

ORF 1b

T7

(nt 187-2063)

(nt 2811-4431)

plasmid pT7-N27
or pS-N27

B

precursor polypeptide

p128

autoproteolytically processed peptide

p28

p100
further de-phosphorylated by calf intestinal alkaline phosphatase (CIP, BRL) in 1 mM ZnCl₂, 1mM MgCl₂, and 10 mM Tris-HCl, pH 8.3. CIP was removed by protease K treatment in 5 mM EDTA and 0.5 % SDS, and extracted with equal volumes of phenol:chloroform (1:1) and chloroform:isoamyl alcohol (24:1). The double digested pT7-N27 and pSELECT-1 (de-phosphorylated) were respectively separated by electrophoresis in a 1.2% or 0.8% low-melting agarose gel. The insert (a 3.5 kb DNA fragment of pT-N27) and the vector (a 5.6 kb DNA fragment of pSELECT-1) were purified from the low-melting agarose gels by the freeze and thaw method as described above. The gel purified fragments were ligated in a volume of 10-20 µl with 1 u of T4 DNA ligase (BRL), 50 mM Tris-HCl pH 7.6, 10 mM MgCl₂, 1 mM ATP and 1 mM DTT at 14°C for 8-16 hours. The control for the ligation reaction contained only CIP-treated vector but no insert. An E. coli DH5α strain was transformed by the ligation reaction and spread on a LB plate containing tetracycline (15 µg/ml), X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactoside, 40 µl of 20 mg/ml per plate) and IPTG (isopropylthio-β-D-galactoside, 4 µl of 200 mg/ml per plate). Only the white colonies were isolated to screen the recombinant plasmid, and the recombinant plasmid pS-N27 was confirmed by restriction enzyme digestion.

**Single-strand phagemid DNA preparation**

Phagemid pS-N27 containing a f1 origin (which is derived from a filamentous bacteriophage) can replicate in a single strand form in E. coli strain JM109 after infection with helper phage R408 (Vieira and Messing,
1987; Maniatis et al., 1982). JM109 were transformed with pS-N27 and spread on a minimal plate (3 g/l KH₂PO₄, 7 g/l K₂HPO₄, 2 g/l (NH₄)₂SO₄, 0.5 mg/l FeSO₄, 10 mM MgSO₄, 0.05 mg/l thiamine, and 4% glucose) containing tetracycline (15 μg/ml) to ensure the presence of bacterial F1 pili for bacteriophage infection. A single tetracycline-resistant colony of JM109 containing pS-N27 plasmid was inoculated in 5 ml of TYP broth (16 g/l Bacto-tryptone, 16 g/l Bacto-yeast extract, 5 g/l NaCl and 2.5 g/l K₂HPO₄) with 15 μg/ml tetracycline in a 50 ml tube to prepare an overnight culture. Two hundred μl of the overnight culture was inoculated into 5 ml of TYP broth containing 15 μg/ml tetracycline in a 125 ml flask with vigorous shaking for 30 min at 37°C. The culture was infected with helper phage R408 at an m.o.i. (multiplicity of infection) of 10 (40 μl of the stock from Promega, >1 X 10¹¹ plaque forming unit per ml or pfu/ml) and shaken for additional 6-8 hours. The bacteria were pelleted by centrifugation twice at 10 k rpm for 5 min at 4°C in a Sorvall SS34 rotor. The supernatant containing the secreted phage particles was transferred to a Corex tube (15 ml), mixed gently with 1/4 volume of phage precipitation solution (3.75 M NH₄OAc and 20% PEG) and incubated on ice for 15-30 min. The mixture was centrifuged at 10 k rpm for 15 min at 4°C. The pellet was drained thoroughly and then resuspended in 400 μl of TE buffer. The suspended solution was transferred to a microcentrifuge tube and extracted by equal volumes of chloroform (first, to lyse the phage particles), phenol:chloroform (1:1, twice), and then chloroform:isoamyl alcohol (24:1). The upper aqueous phase was transferred to sterile microfuge tubes and mixed with 0.5 volume of 7.5 M NH₄OAc, 2 volumes of ethanol and frozen at -70°C for 20
min. Single-strand DNA (phage R408 DNA and single-stranded pS-N27) was precipitated by centrifugation at 10 k rpm for 15 min at 4°C. The pellet was rinsed with cold 70% ethanol, dried in a Savant Speed Vac and resuspended in 20 µl of dH2O. Two µl of the final product was checked by electrophoresis in a 0.7% mini-agarose gel prepared in 0.5 X TBE containing 0.25 µg/ml ethidium bromide at 20 V overnight or at 50 V for 2-4 hours (see Fig. 5). M13mp18 DNA (7.25 kb, from the USB sequencing kit) and Lambda DNA Hind III fragments were loaded as the marker. The molar concentration of the single-stranded DNA was roughly estimated by measuring OD at 260 nm in a UV spectrophotometer.

**Oligonucleotide synthesis and phosphorylation**

Mutagenic oligonucleotides were synthesized by Dr. Basam Wakim at Loyola University Medical Center. Degenerate oligonucleotides listed in Table 1 were used for site-specific mutagenesis of the putative catalytic residues Cys-1137 and His-1288, and two control residues Cys-1172 and His-1317. Degenerate oligonucleotides listed in Table 2 were used for site specific mutagenesis at the identified cleavage sites and its flanking sequences (according to the nomenclature of Schechter and Berger, 1967) including residues Arg-243 (P5), Gly-244 (P4), Tyr-245 (P3), Arg-246 (P2), Gly-247 (P1) Val-248 (P1'), Lys-249 (P2') and Pro-250 (P3'), and also at the previously suggested cleavage site Tyr-257 and Gly-258 (Soe *et al.*, 1987). These oligonucleotides contained degenerate sequences at positions which would result in amino acid substitutions at specific residues of the MHV-JHM polymerase polyprotein. A forward primer B67 (5'-
Fig. 6. Agarose gel electrophoresis analysis of single-strand pS-N27 DNA. Single-strand pS-N27 DNA (lanes 1 and 2) was prepared as described in the Materials and Methods and analyzed by electrophoresis in a 0.7% agarose gel. Lambda DNA HindIII fragments (lane M) and wild type bacteriophage M13mp18 (lane M13, 7.25 kb) were used as the size markers. Single-stranded pS-N27 DNA (9.3 kb) and helper phage R408 DNA (6.4 kb) are indicated.
Table 1. Oligonucleotides used to generate site specific mutations in the MHV PCP-1 protease domain.

<table>
<thead>
<tr>
<th>Wild-type amino acid</th>
<th>Mutagenic oligonucleotide (5' to 3')*</th>
<th>Mutation Codon</th>
<th>Amino Acid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cys-1137</td>
<td>CGCAGCCAXXXATTAGTACGC</td>
<td>TAT</td>
<td>Tyr</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CGT</td>
<td>Arg</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AGT</td>
<td>Ser</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GGT</td>
<td>Gly</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TGG</td>
<td>Trp</td>
</tr>
<tr>
<td>His-1288</td>
<td>GCCATAGAXXXGCAATCATT</td>
<td>ACA</td>
<td>Thr</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CGT</td>
<td>Arg</td>
</tr>
<tr>
<td>His-1317</td>
<td>GTCATACCXXACCCACCAT</td>
<td>CAG</td>
<td>Gln</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GTG</td>
<td>Val</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TCT</td>
<td>Ser</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CCG</td>
<td>Pro</td>
</tr>
<tr>
<td>Cys-1172</td>
<td>CCACAAAGXXXTTGATCATAAGCC</td>
<td>AAC</td>
<td>Asn</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TCC</td>
<td>Ser</td>
</tr>
</tbody>
</table>

* X denotes the positions of degenerate nucleotides.
Table 2. Oligonucleotides used for site-directed mutagenesis of MHV-JHM p28 cleavage site and flanking amino acid sequence. (to be continued on the next page)

<table>
<thead>
<tr>
<th>Amino acid (position)</th>
<th>Substitution (single-letter code)</th>
<th>Mutant codon</th>
<th>Mutagenic oligonucleotide (5' to 3') *</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gly-247 (P1)</td>
<td>Ala (A)</td>
<td>GCT</td>
<td>AGGGGCTATCGCATTTGTTAAGGA</td>
</tr>
<tr>
<td></td>
<td>Asp (D)</td>
<td>GAT</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td>Val (V)</td>
<td>GTT</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Asn (N)</td>
<td>AAT</td>
<td></td>
</tr>
<tr>
<td>Arg-246 (P2)</td>
<td>Pro (P)</td>
<td>CCC</td>
<td>CTTAGGGGCTATCCGGTGTAAAGTT</td>
</tr>
<tr>
<td></td>
<td>Leu (L)</td>
<td>CTC</td>
<td>T</td>
</tr>
<tr>
<td></td>
<td>His (H)</td>
<td>CAC</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>Ser (S)</td>
<td>TCC</td>
<td>CTTAGGGGCTATTCGGTGTAAAGCT</td>
</tr>
<tr>
<td></td>
<td>Tyr (Y)</td>
<td>TAC</td>
<td>T</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tyr-245 (P3)</td>
<td>Arg (R)</td>
<td>CGT</td>
<td>CTTAGGGGGCCCTCGCGGTGTTT</td>
</tr>
<tr>
<td></td>
<td>Ser (S)</td>
<td>TCT</td>
<td>TT</td>
</tr>
<tr>
<td></td>
<td>Phe (F)</td>
<td>TTT</td>
<td>G</td>
</tr>
<tr>
<td></td>
<td>Cys (C)</td>
<td>TGT</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pro (P)</td>
<td>CCT</td>
<td></td>
</tr>
<tr>
<td>Gly-244 (P4)</td>
<td>Ala (A)</td>
<td>GCC</td>
<td>GCTCTTCTTAGGACCTATCGCGGTGT</td>
</tr>
<tr>
<td></td>
<td>Val (V)</td>
<td>GTC</td>
<td>GT</td>
</tr>
<tr>
<td></td>
<td>Asp (D)</td>
<td>GAC</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>Tyr (Y)</td>
<td>TAC</td>
<td>GCTCTTCTTAGGTCTATCGCGGTGT</td>
</tr>
<tr>
<td></td>
<td>Phe (F)</td>
<td>TTC</td>
<td>CT</td>
</tr>
<tr>
<td></td>
<td>Ser (S)</td>
<td>TCC</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>Leu (L)</td>
<td>CTC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>His (H)</td>
<td>CAC</td>
<td></td>
</tr>
<tr>
<td>Arg-243 (P5)</td>
<td>Lys (K)</td>
<td>AAA</td>
<td>GCTCTTCTTAAAGGCTATCGCGGA</td>
</tr>
<tr>
<td></td>
<td>Ile (I)</td>
<td>ATA</td>
<td>GT</td>
</tr>
<tr>
<td></td>
<td>Thr (T)</td>
<td>ACA</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td>Ala (A)</td>
<td>GCA</td>
<td></td>
</tr>
</tbody>
</table>
Table 2. Oligonucleotides used for site-directed mutagenesis of MHV-JHM p28 cleavage site and flanking amino acid sequence. (continued)

<table>
<thead>
<tr>
<th>Amino acid (position)</th>
<th>Substitution (single-letter code)</th>
<th>Mutant codon</th>
<th>Mutagenic oligonucleotide (5' to 3') *</th>
</tr>
</thead>
<tbody>
<tr>
<td>Val-248 (P1')</td>
<td>Ala (A)</td>
<td>GCT</td>
<td>GGCTATCGCGGTAAATAAGCCC</td>
</tr>
<tr>
<td></td>
<td>Asp (D)</td>
<td>GAT</td>
<td>GC</td>
</tr>
<tr>
<td></td>
<td>Gly (G)</td>
<td>GGT</td>
<td>G</td>
</tr>
<tr>
<td></td>
<td>Ser (S)</td>
<td>AGT</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Thr (T)</td>
<td>ACT</td>
<td></td>
</tr>
<tr>
<td></td>
<td>His (H)</td>
<td>CAT</td>
<td>GGCTATCGCGGTCAATAAGCCC</td>
</tr>
<tr>
<td></td>
<td>Arg (R)</td>
<td>CGT</td>
<td>TG</td>
</tr>
<tr>
<td></td>
<td>Leu (L)</td>
<td>CTT</td>
<td>T</td>
</tr>
<tr>
<td></td>
<td>Phe (F)</td>
<td>TTT</td>
<td></td>
</tr>
<tr>
<td>Lys-249 (P2')</td>
<td>Ala (A)</td>
<td>GCG</td>
<td>TCGCGGTGTTGGGCCCATCTCTG</td>
</tr>
<tr>
<td></td>
<td>Arg (R)</td>
<td>AGG</td>
<td>AT</td>
</tr>
<tr>
<td></td>
<td>Met (M)</td>
<td>ATG</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td>Thr (T)</td>
<td>ACG</td>
<td></td>
</tr>
<tr>
<td>Pro-250 (P3')</td>
<td>Arg (R)</td>
<td>CGC</td>
<td>TCGCGGTGTTAAGTGCATCTCTT</td>
</tr>
<tr>
<td></td>
<td>His (H)</td>
<td>CAC</td>
<td>CA</td>
</tr>
<tr>
<td></td>
<td>Leu (L)</td>
<td>CTC</td>
<td>T</td>
</tr>
<tr>
<td>Tyr-257 (*)</td>
<td>Leu (L)</td>
<td>CTT</td>
<td>GTGGACCAGCTTTGTTGCGAC</td>
</tr>
<tr>
<td></td>
<td>Cys (C)</td>
<td>TGT</td>
<td>TG</td>
</tr>
<tr>
<td></td>
<td>Arg (R)</td>
<td>CGT</td>
<td>G</td>
</tr>
<tr>
<td></td>
<td>Phe (F)</td>
<td>TTT</td>
<td></td>
</tr>
<tr>
<td>Gly-258 (*)</td>
<td>Ala (A)</td>
<td>GCT</td>
<td>GGACCAGTATGATTGCGACTATAC</td>
</tr>
<tr>
<td></td>
<td>Asp (D)</td>
<td>GAT</td>
<td>TC</td>
</tr>
</tbody>
</table>

* Tyr-257 and Gly-258 dipeptide bond was previously predicted to be the p28 cleavage site (Soe et al, 1987).
* The complimentary sequences of mutagenic oligonucleotides are shown with degenerate sequences in bold.
GGTAACAAAGGGTCTGTG-3', nucleotide 799 to 817 of MHV-JHM) was synthesized for DNA sequencing at the putative cleavage site. These synthesized oligonucleotides (approx. 0.2 µmole, as dry powder) were resuspended in 500 µl of dH2O and quantitated by measuring OD at 260 nm in a UV spectrophotometer. The mutagenic oligonucleotides are phosphorylated at the 5' end by T4 polynucleotide kinase to increase the ligation efficiency. One hundred picomole of the diluted mutagenic oligonucleotide was incubated with 5 u of T4 kinase in 50 mM Tris-HCl, pH 7.5, 10 mM MgCl2, 5 mM DTT, 0.1 mM spermidine and 1 mM ATP in a total volume of 25 µl at 37°C for 30 min. The reaction was heated at 70°C for 10 min to inactivate the kinase and stored at -20°C until use.

Site-directed mutagenesis reaction

Site directed mutagenesis of pS-N27 was based on a coupled annealing of the desired mutagenic oligonucleotide and another AmpR oligonucleotide to the single-strand DNA templates (see Fig. 7) (Hutchison et al., 1978; Lewis and Thompson, 1990). The single-strand pS-N27 DNA served as the template for the site specific mutagenesis reaction. One and a quarter picomole of one of the phosphorylated synthetic mutagenic oligonucleotides (listed in Table 1 and Table 2) was annealed to 0.05 picomole of the single-stranded pS-N27 in conjunction with 0.25 picomole of an AmpR oligonucleotide in 20 mM Tris-HCl, pH 7.5, 10 mM MgCl2 and 50 mM NaCl in a 20 µl reaction at 70°C for 5 min and then cooled slowly. To the annealing reaction, 1 µl of T7 DNA polymerase (10 u/µl, Promega), 1 µl of T4 ligase (2 u/µl, Promega), 3 µl of the 10 X Synthesis Buffer (100 mM
Fig. 7. Schematic diagram of Altered Sites *in vitro* mutagenesis procedure (from the Promega protocol book). pS-N27 constructed as described in the Materials and Methods is the recombinant plasmid here, in which the insert (about 3.5 kb) is the MHV-JHM gene 1 cDNA from the plasmid pT7-N27. Single-stranded pS-N27 DNA was prepared as described in the Materials and Methods and used as the template for a mutagenic reaction. A mutagenic oligonucleotide, containing mutations at the MHV gene 1 region (see Table 1 and 2), was annealed to the single-stranded pS-N27 DNA in conjunction with Amp\(^\text{R}\)-oligonucleotide which can restore ampicillin resistance to the mutated plasmid. The plasmid DNA was synthesized by T4 DNA polymerase and ligated by T4 ligase. The ligated DNA was transformed into *E. coli* repair-deficient strain BMH 71-18 mut S and then JM109 which were subjected to ampicillin selection. The amplified mutant plasmids were screened by direct DNA sequencing of the desired region.
1. Anneal ampicillin repair oligo and mutagenic oligo.

2. Synthesize mutant strand with T4 DNA polymerase and ligate.

3. Transform BMH 71-18 mut S. Grow in media + ampicillin.

4. Prepare mini-prep DNA.


6. Screen for mutants by direct sequencing.
Tris-HCl, pH 7.5, 5 mM deoxynucleotide mixture with equal dATP, dGTP, dCTP, and dTTP and 5 µl of dH2O was added and incubated at 37°C for 90 min for the minus-strand plasmid DNA synthesis. The ampicillin repair oligonucleotide restores ampicillin resistance to the vector and the linking of it to MHV-JHM gene 1 mutagenic oligonucleotides allows powerful selection of mutants. A repair deficient E. coli strain (BHM71-18 mutS, Promega) was transformed by the ligated DNA in a mutation reaction (30 µl total). The bacteria were shaken in 0.5 ml of SOC medium (20 g/l Bacto-tryptone, 5 g/l Bacto-yeast extract, 0.5 g/l NaCl, 2.5 mM KCl, 10 mM MgCl₂ and 20 mM glucose) for one hour at 37°C. Five ml of TYP broth containing 125 µg/ml ampicillin was added for an additional 8-14 hours of shaking to amplify bacteria harboring mutant plasmids. Plasmid DNA was isolated from BHM71-18mutS cells and an E. coli strain JM109 was transformed by these mutated plasmid DNA. Individual ampicillin resistant colonies were isolated to inoculate a 5 ml overnight culture with TYP broth containing 125 µg/ml ampicillin for minipreparations of plasmid DNA. The prepared plasmid DNA was sequenced across the region encompassing the degenerate oligonucleotide binding site by double-stranded DNA sequencing (USB Sequenase Version 2.0) to screen the desired mutants.

**Plasmid DNA sequencing**

Plasmid DNA sequencing was performed on the basis of Sanger's dideoxynucleotide-mediated chain termination reaction using T7 DNA polymerase (Sanger et al., 1977). Briefly, 2-3 µg of the double-stranded plasmid DNA was denatured in 0.2 M NaOH, concentrated by ethanol
precipitation and resuspended in 7 µl of dH2O. One half pmol of the
sequencing primer was annealed to the single-strand pS-N27 DNA in a 20
µl reaction containing 40 mM Tris-HCl, 20 mM MgCl2, 50 mM NaCl, at 65-
70°C for 5 min and slowly cooled to 30°C. The DNA synthesis was carried
out using 2 u of T7 DNA polymerase (freshly diluted to 1 u/µl, USB
Sequenase Version 2.0) for 2-5 min at room temperature after adding 1 µl of
0.1 M DTT, 2 µl of labeling mixture (freshly diluted to 1.5 µM of dGTP, 1.5
µM of dCTP, 1.5 µM of dTTP), 0.5 µl of 35S-dATP (10 µCi/µl, NEN). The
extended DNA product was labeled with 35S-dATP. Two and one half µl of
the chain extension reaction was randomly terminated by adding 3.5 µl of
each dideoxynucleotide (8 µM of ddGTP, or ddATP, or ddCTP, or ddTTP, in
80 µM of four dNTP and 50 mM NaCl) and incubated at 37°C for 5 min. The
sequencing reaction was stopped by adding 4 µl of Stop Buffer (95 %
formamide, 20 mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanol
FF). The reaction was analyzed by electrophoresis in a 5% polyacrylamide-
7M urea gel (Long Ranger, J. T. Baker). The sequencing gel was
transferred to a Whatman 3MM paper, dried and exposed to a X-ray film at
room temperature. (see Fig. 8, a typical sequencing result for screening the
mutant plasmid)

Preparation of radiolabeled protein sample and N-terminal
microsequencing

To determine the p28 cleavage site, the radiolabeled C-terminal
cleaved protein product (100 kDa, or p100) downstream of p28 processed
from the polyprotein encoded by pT7-N27 was selected for partial N-
Fig. 8. Autoradiogram of representative sequencing result from screening plasmid DNA encoding mutations surrounding the p28 cleavage site. The sequencing result of two mutants: M1 (Lys) and M2 (Thr), and the wild type: Wt (Arg) of Arg-243 at P5 position are shown. The codons for wild type and mutants at position 243 are denoted at the left, and sequences of nucleotide 932-952 of MHV-JHM gene 1 from the wild type are indicated at the right.
terminal protein sequence analysis which is based on Edman degradation reaction (Matsudaira, 1987; LeGendre and Matsudaira, 1988). To obtain significant, large volume of 100 or 40 μl in the present protocol and small product size, the reaction was carried out at 50 °C. The precipitated PVDF membrane band was eluted with 2 M urea and labeled with 3H-leucine and 3H-lysine and subjected to gel electrophoresis in a Beckman model 35 gel apparatus. The gel was analyzed by scanning of radioactivity.
Terminal protein sequence analysis which is based on Edman degradation reaction (Matsudaira, 1987; LeGendre and Matsudaira, 1988). To obtain significant amounts of in vitro translation products for protein sequencing, large scale in vitro translation reactions were performed in a total volume of 100 µl using 4.0 µg of pT7-N27 DNA linearized with EcoR I in the presence of 80 µCi of 35S-Met (Amersham), or 40 µCi 3H-Leu (ICN Biomed.), or 40 µCi 3H-Val (ICN Biomed.) for 90 min at 30°C. The in vitro translation products were mixed with equal volumes of 2 X Laemmli Sample Buffer and separated by electrophoresis in a 7.5% SDS-PAGE gel for 24 hours at 50 V. Following electrophoresis, the translation products were electro-transferred to Immobilon polyvinylidene difluoride (PVDF) membrane (Millipore) at 40 V, 4°C, for 8-16 hours in buffer containing 25 mM tris-base, 192 mM glycine, 20% (v/v) methanol and 0.01% (w/v) SDS. The PVDF membrane was air-dried and exposed to a Kodak X-ray film at -70°C. The band of target protein labeled with 35S-Met labeled proteins was identified by autoradiography and the bands of corresponding proteins labeled with 3H-Leu or 3H-Val (not visible) were co-localized in parallel to 35S-Met labeled proteins. These labeled proteins were excised from the membrane and subjected to an Applied Biosystem gas-phase sequencer. Partial N-terminal amino acid sequence analysis was performed by Dr. Basam Wakim at Loyola University Medical Center. Radioactivity released from each cycle was quantitated by scintillation counting for 10 min in a Beckman scintillation counter and then plotted to identify peaks of radioactivity.
CHAPTER III
RESULTS

Coronavirus polymerase is encoded by gene 1 located at the 5' end of the viral genome. This gene (22 kb) is translated as a polyprotein with multiple functional domains, including two papain-like cysteine proteinase domains and one poliovirus 3C-like serine proteinase domain (see Fig. 2) (Lee et al., 1991). These proteinases are proposed to proteolytically process the polymerase polyprotein into mature and functional protein products. An autoproteinase in the MHV-JHM polymerase polyprotein was shown to be responsible for the cleavage of the amino-terminal peptide, p28 (Baker et al., 1989). Computer-assisted amino acid sequence searches revealed that this autoproteinase overlapped the putative first proteinase domain (PCP-1), which has some sequence homology to the cellular cysteine proteinases (Gorbalenya et al., 1991). The objective of the research reported in this dissertation was to characterize the first papain-like cysteine proteinase domain (PCP-1) and to identify the p28 cleavage site and its amino acid sequence motif recognized by this proteinase.

Effect of deletion mutation on proteinase activity

The first proteinase domain was predicted to be at 3.5 to 4.2 kb from the 5' end of the MHV-JHM gene 1 (Lee et al., 1991). To define more
precisely this putative proteinase domain, a deletion analysis was performed. Plasmid pT7-NBgl containing nucleotides 187 to 5273 of gene 1, was used as the parental construct and generated three in-frame deletion plasmids pT7-NAc, pT7-N27, and pT7-N29 (see Fig. 3). The plasmids were transcribed by T7 RNA polymerase and translated in rabbit reticulocyte lysates in the presence of \(^{35}\text{S-Met}\), and the protein products were analyzed directly or after immunoprecipitation with anti-p28 serum by SDS-PAGE (Fig. 9A and B). The pT7-NBgl translation yielded proteins of approximately 188 kDa, 160 kDa, and 28 kDa (Fig 9A, lane 2). This result is consistent with previous results which showed that the 188 kDa protein is the primary translation product and the 160 kDa and 28 kDa proteins are the products of an autoproteolytic cleavage (Baker et al., 1989). p28, the N-terminal cleavage product, is immunoprecipitated with anti-p28 antibody (Figure 9B, lane 2). The p28 cleavage efficiency for pT7-NBgl translation products was approximately 70%, as determined by quantitation of p28 and the precursor polyprotein using a Betagen scanner. Truncation of a 0.9 kb region at the 3' end of the PCP-1 domain in plasmid pT7-NAc had little effect on proteinase activity. It was demonstrated that p28 cleavage efficiency of the polypeptide encoded by pT7-NAc was still approximately 70% (Fig. 9A and B, lane 3). Furthermore, deletion of an internal 0.74 kb region between p28 and the proteinase domain had only a minor effect on the ability of the proteinase to cleave p28 from the precursor polyprotein (60-65% cleavage efficiency), as shown by translation of pT7-N27 RNA (Fig. 9A and B, lane 4). These results are consistent with the prediction that the proteinase domain is located in the region from 3.5 to 4.2 kb from the 5'-end
Fig. 9. Effect of deletion mutation on proteolytic processing of p28. Linearized plasmid DNA was transcribed and translated in rabbit reticulocyte lysates in the presence of T7 RNA polymerase and 35S-Met as described in Material and Methods. Equal TCA-precipitable radioactive counts of translation products were analyzed by PAGE in 10% polyacrylamide gels directly (A) and following immunoprecipitated with p28-specific antiserum (B). Lanes: 1) no RNA, 2) pT7-NBgl RNA, 3) pT7-NAc RNA, 4) pT7-N27 RNA, and 5) pT7-N29 RNA. Lane M contained 14C-labeled marker polypeptides; molecular weights are given in kilodaltons.
and that some spacing sequences between the proteinase and p28 can be deleted without abolishing p28 cleavage. However, a larger internal deletion of 1.6 kb (pT7-N28 RNA) resulted in an inability to cleave the p35 protein, and the appearance of two protein products of approximately 95

There are four classes of proteinases which have been proposed from the studies of cellular proteinases (Barrett and Salvesen, 1986). These four proteinases are serine proteinases, cysteine proteinases, aspartic proteinases and metallo-proteinases. Representatives of each of these proteinase classes have been found in viral systems, except for the metalloproteinases (Kremslich and Wimmer, 1988; Dougherty and Samler,
and that some spacing sequences between the proteinase and p28 can be deleted without abolishing p28 cleavage. However, a larger internal deletion of 1.6 kb (pT7-N29 RNA) resulted in an inability to cleave the p28 protein, and the appearance of two protein products of approximately 95 kDa and 90 kDa precipitable by p28 antiserum (Fig. 9B, lane 5). The 95 kDa band was likely to be the primary translation product from the pT7-N29, and the 90 kDa protein probably resulted from premature termination of translation, or from an inefficiently proteolytic cleavage at an alternative site. This result indicates that the internal region between the proteinase domain and the p28 cleavage site may act as a spacing region to provide a correct conformation essential for autoproteolytic activity. Alternatively, the protein product encoded by this region may play a supporting role in the processing of the polyprotein. Overall, these deletion studies supported the prediction of a proteinase domain at the 5' end 3.5 to 4.2 kb region of gene 1 and also suggested an important role for the spacing region between the proteinase domain and the cleavage site (Baker et al., 1998).

Effect of proteinase inhibitors on the autoproteolytic activity of MHV PCP-1

There are four classes of proteinases which have been proposed from the studies of cellular proteinases (Barrett and Salvesen, 1986). These four proteinases are serine proteinases, cysteine proteinases, aspartic proteinases and metallo-proteinases. Representatives of each of these proteinase classes have been found in viral systems, except for the metalloproteinases (Krausslich and Wimmer, 1988; Dougherty and Semler,
Computer-assisted sequence analysis suggests that the first proteinase of murine coronavirus gene 1 is probably a papain-like cysteine proteinase with the conserved Cys-1137 and His-1288 as the catalytic residues (Gorbalenya et al., 1991). In order to determine if MHV PCP-1 was indeed a cysteine proteinase, we tested the sensitivity of the proteolytic activity to a variety of proteinase inhibitors in rabbit reticulocyte lysate translation reactions of in vitro transcribed pT7-N27 RNA (see Fig. 10 and Table 3). The proteinase inhibitors tested included a general proteinase inhibitor ZnCl₂, four cysteine proteinase inhibitors: cystatin, iodoacetamide, and N-ethyl-maleimide and E-64, a cysteine and serine proteinase inhibitor leupeptin, a serine proteinase inhibitor APMSF [94-(amidinophenyl)-methanesulfonylfluoride], an aspartic proteinase inhibitor pepstatin A, and a metallo-proteinase inhibitor EGTA [(ethylenebis(oxyethlenenitrilo)]-tetraacetic acid]. The effect of a wide range of concentrations of each inhibitor on p28 processing was tested to achieve the optimal concentration for each of these inhibitors. The optimal concentration had no effect on the overall efficiency of the translation reaction but did inhibit specific proteinase activity. The optimal amounts of these selective proteinase inhibitors were added to the in vitro translation reaction of pT7-N27 RNA before the incubation at 30°C. The 35S-Met labeled translation products were analyzed by 10% SDS-PAGE gel and autoradiography. The radioactivity of the precursor polyprotein and the cleavage product p28 were quantitated by Betagen scanning to determine the p28 cleavage efficiency (or the proteolytic activity of MHV PCP-1). The cleavage efficiencies were normalized after setting the cleavage efficiency of
Fig. 10. Inhibition of the proteolytic activity of MHV PCP-1 by proteinase inhibitors. Approximately 200 ng of pT7-N27 RNA were translated in rabbit reticulocyte lysates in the presence of $^{35}$S-Met as described in the Materials and Methods. The optimal amount of the selective proteinase inhibitor was added to the \textit{in vitro} translation reaction of pT7-N27 RNA prior to incubation at 30°C for 90 min. The $^{35}$S-Met labeled protein product was analyzed by electrophoresis in a 10% acrylamide -- 0.1% SDS gel and autoradiography. A. Lanes: 1) no inhibitors, 2) 0.1 mM ZnCl$_2$, 3) 0.14 mM cystatin. B. Lanes: 1) no inhibitors, 2) 0.1 mM ZnCl$_2$, 3) 0.2 mM iodoacetamide, 4) 2.5 mM N-ethyl-maleimide, 5) 0.5 mM E64, 6) 0.8 mM leupeptin, 7) 1.00 mM APMSF, 8) 1.0 mg/ml pepstatin A and 9) 8.0 mM EGTA.
Table 3. Effect of proteinase inhibitors on the autoproteolytic activity of the MHV PCP-1 proteinase.

<table>
<thead>
<tr>
<th>Inhibitor</th>
<th>Proteinase Family</th>
<th>Concentration</th>
<th>Cleavage Efficiency</th>
<th>Inhibition(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No inhibitor</td>
<td>------</td>
<td>100%</td>
<td></td>
<td>18%</td>
</tr>
<tr>
<td>ZnCl₂</td>
<td>general</td>
<td>0.10 mM</td>
<td>82%</td>
<td>18%</td>
</tr>
<tr>
<td>cystatin</td>
<td>cysteine</td>
<td>0.14 mM</td>
<td>82%</td>
<td>18%</td>
</tr>
<tr>
<td>iodoacetamide</td>
<td>cysteine</td>
<td>0.20 mM</td>
<td>82%</td>
<td>18%</td>
</tr>
<tr>
<td>N-ethyl-maleimide</td>
<td>cysteine</td>
<td>2.50 mM</td>
<td>77%</td>
<td>23%</td>
</tr>
<tr>
<td>E-64</td>
<td>cysteine</td>
<td>0.50 mM</td>
<td>82%</td>
<td>18%</td>
</tr>
<tr>
<td>leupeptin</td>
<td>serine/cysteine</td>
<td>0.80 mM</td>
<td>91%</td>
<td>9%</td>
</tr>
<tr>
<td>APMSF</td>
<td>serine</td>
<td>1.00 mM</td>
<td>100%</td>
<td>18%</td>
</tr>
<tr>
<td>pepstatin A</td>
<td>aspartic</td>
<td>1.0 ug/ml</td>
<td>100%</td>
<td>18%</td>
</tr>
<tr>
<td>EGTA</td>
<td>metallo-</td>
<td>8.00 mM</td>
<td>100%</td>
<td>18%</td>
</tr>
</tbody>
</table>
in vitro translation reaction without proteinase inhibitors as 100%. The proteinase inhibitory effect was estimated roughly by comparing radioactivity in p28 and precursor polyproteins and finally determined by calculating the decrease of the cleavage efficiency in the presence of proteinase inhibitors.

ZnCl₂ was shown to inhibit proteolytic activity in previous studies and used as the positive control (Denison and Perlman, 1986; Denison et al., 1992; Baker et al., 1989). In the presence of ZnCl₂ (Fig. 10 and B, lanes 2), the intensity of p28 was decreased (as well as the intensity of the large cleaved product p100) as compared to the cleavage products of in vitro translation without proteinase inhibitors (Fig. 10A and B, lanes 1). Correspondingly, the intensity of the precursor polyprotein was increased as compared to that from the negative control without the proteinase inhibitor (Fig 10A and B, lanes 1 and 2), confirm that ZnCl₂ inhibits the proteolytic activity of MHV PCP-1.

In the presence of four cysteine specific proteinase inhibitors: cystatin, Iodoacetamide, N-ethyl-maleimide and E64, the relative ratios of p28 to the precursor polyprotein were also decreased (Fig. 10A, lane 3; Fig. 10B, lanes 3 to 5). And these specific cysteine proteinase inhibitors were shown to inhibit proteolytic activity approximately to the same level as ZnCl₂. An inhibitor to both cysteine and serine proteinases, leupeptin, had some inhibitory effect on the proteolytic activity (Fig. 10B, lane 6), consistent with the published data (Denison et al., 1992). In contrast, the
representatives of the serine proteinase inhibitor--APMSF, the aspartic proteinase inhibitor--pepstatin A, and the metallo-proteinase inhibitor--EGTA, showed little or no inhibition of the proteolytic activity (Fig. 10B, lanes 7, 8, 9 respectively). The inhibition of the proteolytic activity of MHV PCP-1 by inhibitors of cysteine proteinases but not by inhibitors of serine proteinases, aspartic proteinases, and metalloproteinases supports the hypothesis that MHV PCP-1 is probably a papain-like cysteine proteinase.

**Effect of site-specific mutations of the putative MHV PCP-1 catalytic residues Cys-1137 and His-1288 on proteolytic processing of p28**

Amino acid sequence homology searches and the proteinase inhibition studies described above suggested that MHV PCP-1 is a papain-like cysteine proteinase with characteristic catalytic residues of Cys and His. Among the Cys and His residues within the predicted proteinase domain, Cys-1137 and His-1288 are the only two residues conserved among all the cellular and viral papain-like cysteine proteinases and thus were predicted to be the catalytic amino acids (Gorbalenya *et al.*, 1991). To determine if the predicted residues were indeed essential for catalytic activity, these putative catalytic amino acids were altered by site-directed mutagenesis (see Fig. 11). His-1317 and Cys-1172 in the proximity were also altered by site-directed mutagenesis for comparison.

Site specific mutagenesis was performed by the method of Hutchinson and co-workers (1978) using single-strand pS-N27 DNA and
Fig. 11. The amino acid sequence of the proposed first papain-like cysteine proteinase domain (PCP-1) of MHV-JHM gene 1. Cys-1137 and His-1288 marked by black triangles were predicted to be the catalytic residues of MHV PCP-1 (Gorbalenya et al., 1989). The putative catalytic residues and two control residues (marked by white triangles) were subjected to site-specific mutagenesis analysis.
oligonucleotides containing degenerate sequences at the specific nucleotides selected for mutation (see Table 2). The procedures were described in the Promega protocol of the Altered Sites *in vitro* mutagenesis system (Fig. 7, in Materials and Methods), which uses a phagemid vector pSELECT-1 (Lewis and Thompson, 1990). The gene 1 fragment from plasmid pT7-N27 (3.5 kb Pvu II-EcoR I fragment) was inserted into the polylinker region of pSELECT-1 (between the Sma I and BamH I sites), and the recombinant plasmid was designated pS-N27. Single-strand pS-N27 DNA was prepared by standard methods and serves as the template for mutagenesis reactions. Plasmid DNA replication was primed by two synthetic oligonucleotides: one restores ampicillin resistance to the plasmid and allows the selection of the mutant plasmids, and the other was designed to encode mutations at specific codons of the MHV gene 1 polyprotein. The two primers were added in excess so that they will be simultaneously incorporated into the same DNA strand during replication and produce double mutants selectable by ampicillin. Individual mutants at each site were identified by DNA sequencing and listed in Table 1. Plasmid DNAs encoding amino acid substitutions of each position were linearized by digestion with EcoR I, transcribed and translated *in vitro*. The $^{35}$S-Met labeled protein products with equivalent TCA-precipitable radioactivity are immunoprecipitated by p28 specific antiserum and analyzed by electrophoresis in a 10% polyacrylamide gel.

The putative catalytic residue, Cys-1137 was substituted to Tyr, Arg, Ser, Gly, or Trp. The polypeptides containing these mutations all resulted
in a total loss of p28 and corresponding increase of an uncleaved precursor polypeptide (Fig. 12A). In contrast, mutations of Cys-1172 to Asn or Ser revealed no alteration in the cleavage of p28 (Fig. 12B). The loss of the proteinase activity by mutating Cys-1137 to a variety of amino acids of different properties confirmed the importance of this specific residue. Thus, Cys-1137 is indeed essential for the proteinase function.

Mutation of the putative catalytic residue His-1288 to Thr or Arg also inactivated proteinase activity (Fig. 13A). Therefore, His-1288 as well as Cys-1137 is an essential residue for the proteinase activity. Interestingly, substitutions of His-1317 to Gln and Ser (Fig. 13B), also resulted in a 10-20% reduction in the production of p28, as determined by quantitation of cleavage products using a Betagen scanner. Substitution of His-1317 to Val or Pro (Fig. 13B) resulted in significant reductions in proteolytic processing to release p28. In particular, the substitution of Pro at position 1317 completely eliminated the production of p28 and resulted in a primary translation product which migrated slightly faster on the gel (Fig 13B). This is not surprising since the addition of a Pro residue may cause significant alteration in the protein folding. These data are consistent with the idea that His-1317, considered as a non-catalytic residue of the proteinase, may play a role in maintaining conformation of the proteinase domain to allow accurate and efficient cleavage of p28.

The above site-mutagenesis analysis of Cys-1137, His-1288, Cys-1172 and His-1317 demonstrated that the catalytic residues of this proteinase are
Fig. 12. Effect of site-specific mutation at the putative MHV PCP-1 catalytic residue Cys-1137 on proteolytic processing of p28. A series of mutants at Cys-1137 and control Cys residue at position 1172 were generated by oligonucleotide-directed single-strand DNA mutagenesis as described in the Materials and Methods. Wild type or mutant plasmid DNA pS-N27 was linearized by EcoR I digestion and used as the template in a coupled in vitro transcription/translation reaction which was performed as described in the Materials and Methods. Equal TCA counts of in vitro translation products were immunoprecipitated by p28-specific antiserum and analyzed by 10% SDS-PAGE. A. Lanes show products encoding the following amino acids at the position 1137: no RNA, wild type (Cys), Tyr, Arg, Ser, Gly, and Trp; B. Lanes show products encoding the following amino acids at position 1172: no RNA, wild type (Cys), Asn, and Ser.
### Cys-1137

<table>
<thead>
<tr>
<th>M</th>
<th>N</th>
<th>Wt</th>
<th>Tyr</th>
<th>Arg</th>
<th>Ser</th>
<th>Gly</th>
<th>Trp</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>97</td>
<td>68</td>
<td>43</td>
<td>29</td>
<td></td>
</tr>
</tbody>
</table>

### Cys-1172

<table>
<thead>
<tr>
<th>M</th>
<th>N</th>
<th>Wt</th>
<th>Asn</th>
<th>Ser</th>
</tr>
</thead>
<tbody>
<tr>
<td>p28</td>
<td>200</td>
<td>97</td>
<td>68</td>
<td>43</td>
</tr>
</tbody>
</table>

- **A** and **B** show the molecular weight ranges for the proteins. The arrows indicate the position of p28.
Fig. 13. Effect of site-specific mutation at the putative MHV PVP-1 catalytic residue His-1288 on proteolytic processing of p28. Various His-1288 mutants and control His-1317 mutants in pS-N27 were analyzed as in Fig. 12. A. Lanes show products encoding the following amino acids at position 1288: no RNA, wild type, Thr and Arg; B. Lanes show products encoding the following amino acids at position 1317: no RNA, wild type, Gln, Val, Ser, and Pro.
Cys-1137 and His-1288, and that it belongs to the family of papain-like cysteine proteinases (Baker at al., 1993).

**Determining the amino-terminal amino acid sequence of the protein adjacent to p28**

Studies on viral proteinases have shown that these proteinases have a high degree of specificity for their specific cleavage sites (Dougherty and Semler, 1993). For example, in picornaviruses, proteinase 2A cleaves predominantly at the Tyr(P1)-Gly(P1') dipeptide bond and proteinase 3C recognizes Gln/Glu(P1)-Gly/Ser/Ala(P1') dipeptide bond (Palmenberg, 1990). The specificity for cleavage is not only shown in the amino acids forming the dipeptide bond to be cleaved or the scissile bond, but also in some surrounding residues which may be important to maintain the appropriate conformation for cleavage site presentation (Krausslich and Wimmer, 1988; Palmenberg, 1990; Dougherty and Semler, 1993). According to molecular weight analysis, the p28 cleavage site is expected to be at approximately 1 kb from 5' end of MHV gene 1. This region contains several conventional cleavage sites (Tyr-Gly, Gly-Cys, Gly-Leu, etc.) similar to the cleavage sites of poliovirus proteinases 2A, and Sindbis virus proteinase nsP2. Previously, Soe et al. (1987) predicted Tyr-257--Gly-258 and Tyr-272-Gly-273 as the potential candidates for the p28 cleavage site.

To identify the actual cleavage site, a direct approach was used to isolate the cleavage product downstream of p28 and sequence its N
terminus. The most critical point for conventional protein sequencing is to prepare a sufficient quantity of protein sample (10-1,000 picomole, 0.5-50 µg of a 50 kDa molecule). Since *E. coli* and yeast expression systems have been developed, they are frequently used to overexpress eukaryotic proteins for purification purposes. Previous studies have shown that viral proteinases can be highly expressed in these systems and still retain their proteolytic activities (Ivanoff *et al.* 1986; Nicklin *et al.* 1988; Alvey *et al.* 1991). We have tried an *E. coli* protein overexpression system (PET system, Novagen) for expressing the gene 1 protein encoded in pT7-N27. However, it seems that MHV PCP-1 is highly toxic to bacteria and we were unable to express sufficient amounts of the protein product using this system.

To get around the limit of the requirement for microgram quantities of the pure protein, we chose to isolate a radiolabeled *in vitro* cleaved protein adjacent to p28 and subject the protein to amino-terminal microsequencing (see Fig. 14). The protein was generated by *in vitro* transcription and translation of our smallest in-frame deletion construct pT7-N27 which still maintains active proteolytic processing of p28 (Baker *et al.*, 1993). As shown in previous deletion analysis of the PCP-1 domain, *in vitro* translation of pT7-N27 generates a precursor polyprotein of 128 kDa which is autoproteolytically cleaved by MHV PCP-1 to yield p28 and a 100 kDa protein (p100) (Baker *et al.*, 1993). We found that pT7-N27 serves as an excellent template for an efficient transcription and subsequent translation reaction which can produces a large quantities of protein products. In addition, the 128 kDa precursor and 100 kDa C-terminal cleavage product were easily
Fig. 14. Strategy of N-terminal protein sequencing to identify the p28 cleavage site. The radiolabeled C-terminal cleaved 100 kDa peptide (p100) downstream of p28 was generated by \textit{in vitro} transcription/translation reaction of pT7-N27 and targeted for N-terminal microsequencing to determine the p28 cleavage site.
Translation Product from Plasmid pT7-N27

precursor polypeptide

autoproteolytically processed peptide

p28          p100

N-terminal sequencing
resolved on a 7.5% SDS-polyacrylamide gel which allowed us to isolate the target 100 kDa protein. Thus, *in vitro* translation products of pT7-N27, labeled with either $^3$H-Leu, $^3$H-Val or $^{35}$S-Met were synthesized, separated by electrophoresis in a polyacrylamide gel, and transferred to PVDF membrane (Fig. 15). The 100 kDa protein was identified by autoradiography and subjected to microsequencing as described in Materials and Methods. The sequencing profile generated from samples labeled with $^3$H-Leu indicated that Leu occupied amino acid positions 5, 18 and 22 from the cleavage site (Figure 16). Microsequencing of the 100 kilodalton protein labeled with $^3$H-Val revealed Val at positions 1 and 7 (Fig. 16). Microsequencing of the 100 kDa protein labeled with $^{35}$S-Met indicated that there was no Met at the first 25 positions from the cleavage site (data not shown). By aligning this profile with the amino acid sequence deduced from the MHV-JHM nucleotide sequence (Soe *et al.*, 1987; Lee *et al.*, 1991), we identified the cleavage site for the generation of p28 as the Gly-247 (P1)-Val-248 (P1') dipeptide bond (see Fig. 17).

**Effect of site-specific mutations surrounding the cleavage site on proteolytic processing of p28**

To determine the amino acid sequence requirements for the PCP-1 recognition of the p28 cleavage site, we have targeted 8 amino acid residues including P5 to P3' surrounding p28 cleavage site for site-directed mutagenesis (see Fig. 18). A variety of site-specific mutants (total 42) at
Fig. 15. Schematic diagram of the autoradiogram of radiolabeled protein electrotransferred to the PVDF membrane. 100 µl of pT7-N27 in vitro translation products were separated in a 7.5% SDS-PAGE gel and transferred to Immobilon polyvinylidene difluoride (PVDF) membrane (Millipore) as described in the Materials and Methods. The electrotransferred membrane was exposed to a X-ray film and autoradiogramed. The autoradiogram was matched with the PVDF membrane to identify p100 labeled with $^{35}$S-Met in the membrane. $^3$H-Leu or $^3$H-Val labeled p100 were not visible in the autoradiogram and were localized in parallel to $^{35}$S-Met labeled p100. All radiolabeled p100 was cut out and subjected to automated N-terminal protein sequence analysis to determine the p28 cleavage site.
Precursor peptide

C-terminal cleaved peptide (100kd)

35S-Met labeling

3H-Leu or 3H-Val labeling

35S-Met labeling

Precursor peptide

C-terminal cleaved peptide (100kd)
Fig. 16. N-terminal sequence analysis of radiolabeled p100. p100 was derived from the \textit{in vitro} transcription/translation reaction of plasmid pT7-N27 and were labeled with either $^{35}$S-Met, $^{3}$H-Leu, or $^{3}$H-Val. The radiolabeled samples on the PVDF membrane were identified by autoradiography and subjected to microsequencing analysis based on the Edman-degradation reaction. Radioactivity released at each cycle was quantitated using a Beckman scintillation counter and counts per minute (cpm) were plotted. The amino acid sequence in single letter codes shown at the top of the graph corresponds to amino acid 248 to 272 of the deduced amino acid sequence of MHV-JHM gene 1.
Fig. 17. Schematic diagram of the putative p28 cleavage site. Partial amino acid sequence of the potential cleavage site region is shown, in which the cleavage site Gly-247-Val-248 predicted from protein microsequencing analysis is indicated by a black triangle and the previously proposed cleavage site (Soe et al., 1987) is marked by an asterisk (*). The amino acids underlined were targeted for site-specific mutagenesis analysis to further characterize the sequence motif recognized by MHV PCP-1
Translation Product from Plasmid pT7-N27

precursor polypeptide

autoproteolytically processed peptide

N-terminal sequencing
Fig. 18. Nomenclature of the amino acid residues surrounding a cleavage site for a proteinase. The figure is modified from Dougherty and Semler, 1993, and the nomenclature was according to Schechter and Berger, 1967. The peptide bond to be cleaved is indicated by an arrow, and amino acids upstream of the scissile bond are designed as P1, P2, P3, P4, P5, etc. and amino acids downstream as P1', P2', P3', P4', P5', etc. A schematic proteinase domain is drawn at the bottom with the subsites (S5 through S5') of the substrate binding pocket indicated. The complexity of the substrate-binding pocket of an individual proteinase determines the requirement of its cleavage site, or its specificity.
positions from P5 to P3' (see Fig. 18) were created by single-strand DNA mutagenesis and isolated by direct plasmid DNA sequencing as previously described (see Materials and Methods, Table 2). Individual mutant pS-N27 DNAs at each position were linearized by digestion with EcoR I, transcribed and translated in vitro as described above. The 35S-Met labeled protein products with equal TCA-precipitable radioactivity were directly analyzed by 10% SDS-PAGE to determine the effect of amino acid substitutions on p28 cleavage (Fig. 19 to 26). Samples of in vitro translated proteins with equivalent radioactivity were further analyzed by immunoprecipitation with p28 specific antiserum, and the autoradiogram showed the amount of p28 cleaved from the wild-type pS-N27 and mutant polypeptides at these positions (Fig. 27).

As shown in Figure 19, mutation of Gly-247 at the P1 position to Ala, Asp, Asn or Val resulted in a dramatic decrease in the yield of p28 generated by autoproteolytic processing of the mutant polyprotein. However, mutation at the P1 position did not completely abolish proteinase activity as is seen in a mutant of the catalytic site of PCP-1 in which the catalytic residue Cys-1137 was mutated to Ser (Figure 19, lane C) (Baker et al., 1993). When the products of the in vitro translation reaction were immunoprecipitated with the anti-serum to p28, we detected two p28 specific bands (Figure 27, P1 position), indicating that a low level of aberrant processing may also be taking place in the P1 mutants. Such aberrant autoprocessing of a polypeptide encoding mutations in the cleavage site has also been reported for the poliovirus 2A proteinase (Hellen
Fig. 19. *In vitro* translation protein products of wild type pS-N27 and mutants of Gly-247 (P1). Site-specific mutations were introduced by degenerate oligonucleotide-directed mutagenesis as described in the Materials and Methods. Linearized plasmids were translated in the T7 RNA polymerase coupled rabbit reticulocyte lysates in the presence of $^{35}$S-Met and translation products were analyzed by 10% SDS-PAGE. Specific mutations at the P1 position are indicated by amino acids A (Ala), D (Asp), N (Asn), and V (Val) at the top of the corresponding lanes. Additional lanes are: M, molecular weight marker; N, no RNA; Wt, wild type (Gly-247) pS-N27 translation products; C, control polypeptide with inactive PCP-1 in which the catalytic residue of the proteinase, Cys-1137, was mutated to Ser (Baker et al., 1993).
This aberrant processing may be due to an alteration in the presentation of the cleavage site to the proteinase domain. Clearly, Gly-247 at the position of p28 for Hyp contributed to the cleavage of p28. The results indicated an impaired cleavage site for Hyp and a smaller, and the wild type of p8. The wild type of p8, or f3, or f4, and f5, respectively. Most of these substitutions result in drastic changes in the side chains of the amino acids, but all the mutations had little or no effect on the cleavage.
et al., 1992). This aberrant processing may be due to an alteration in the presentation of the cleavage site to the proteinase domain. Clearly, Gly-247 at the P1 position is an important determinant of the cleavage efficiency and specificity of p28.

Like Gly-247 at the P1 position, mutation of Arg-246 at the P2 position to His, Leu, Pro, Ser, or Tyr also resulted in dramatic reduction in cleavage of p28 (Fig. 20 and 27). No p28 from all these mutant polyproteins was detectable even after a long exposure of the SDS-PAGE gel. Thus, Arg-246 at the P2 position is also critical for efficient cleavage of p28. These results indicate that the P2 residue (Arg-246) plays an important role for cleavage-site determination, which is consistent with the studies of the cleavage sites for other viral cysteine proteinases in tobacco etch virus, and Hypovirulence-associated virus (Carrington and Herndon, 1992; Shapira and Nuss, 1991).

In contrast, mutations at the surrounding positions including P3, P4, P1', P2', and P3' had little or no effect on cleavage efficiency. Polyproteins encoding mutations at the P1', P2', P3', P3 or P4 positions generally maintained equivalent cleavage of p28 as compared to the wild type polyprotein (Fig. 21 to 25, and Fig. 27). At the P3 or P4 position, the wild type has been selectively mutated from Tyr-245 to Cys, Phe, Pro, Arg, or Ser, or from Gly-244 to Ala, Asp, Phe, His, Leu, Ser, Val, or Tyr, respectively. Most of these substitutions result in drastic changes in the side chains of the amino acids, but all the mutations had little or no effect on the cleavage
Fig. 20. *In vitro* translation protein products of wild type pS-N27 and mutants of Arg-246 (P2). Site-specific mutagenesis analysis was performed as described in Fig. 19. Specific mutations at the P2 position are indicated by amino acids Y(Tyr), S(Ser), P(Pro), L(Leu), and H(His) at the top of the corresponding lanes and lane Wt represents wild type Arg-246 pS-N27 translation products.
Arg-246 (P2)

<table>
<thead>
<tr>
<th>Wt</th>
<th>Y</th>
<th>S</th>
<th>P</th>
<th>L</th>
<th>H</th>
</tr>
</thead>
</table>

p28

Fig. 2 shows the effect of mutations at position 246 on protein expression. The mutants were generated and analyzed as described in the methods section. The gel shows the expression levels of wild-type (Wt) and various mutants (Y, S, P, L, H). The p28 band is indicated by an arrow.
Fig. 21. *In vitro* translation protein products of wild type pS-N27 and mutants of Tyr-245 (P3). Site-specific mutagenesis analysis was performed as described in Fig. 19. Specific mutations at the P3 position are indicated by amino acids C (Cys), F (Phe), P (Pro), R (Arg), and S (Ser) at the top of the corresponding lanes and lane Wt represents wild type Tyr-245 pS-N27 translation products.
### Tyr-245 (P3)

<table>
<thead>
<tr>
<th>Wt</th>
<th>C</th>
<th>F</th>
<th>P</th>
<th>R</th>
<th>S</th>
</tr>
</thead>
</table>

![Image of gel electrophoresis](image_url)

- **p28**

Fig. 3: The table and gel electrophoresis image show the analysis of Tyr-245 (P3) in different conditions. The wild-type (Wt) and various mutant conditions (C, F, P, R, S) are compared. The gel highlights the p28 band, indicating specific expression or alteration in these conditions. Further analysis is performed as indicated by the experiment's methodology.
Fig. 22. *In vitro* translation protein products of wild type pS-N27 and mutants of Gly-244 (P4). Site-specific mutagenesis analysis was performed as described in Fig. 19. Specific mutations at the P4 position are indicated by amino acids A (Ala), D (Asp), F (Phe), H (His), L (Leu), S (Ser), V (Val), and Y (Tyr) at the top of the corresponding lanes and lane Wt represents wild type Gly-244 pS-N27 translation products.
<table>
<thead>
<tr>
<th>Gly-244 (P4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wt</td>
</tr>
</tbody>
</table>

![Image of gel electrophoresis with bands labeled p28]
**Fig. 23.** *In vitro* translation protein products of wild type pS-N27 and mutants of Val-248 (P1'). Site-specific mutagenesis analysis was performed as described in Fig. 19. Specific mutations at the P1' position are indicated by amino acids A (Ala), D (Asp), G (Gly), S (Ser), T (Thr), L (Leu), F (Phe), R (Arg), and H (His) at the top of the corresponding lanes and lane Wt represents wild type Val-248 pS-N27 translation products.
Fig. 3a: The wild-type and five mutants, as described in the legend, were expressed in bacteria and analyzed by an immunoblot in which corresponding polypeptides were visualized using a polyclonal antibody. The Western blots were performed as described in the legend. The positions of the wild-type (Wt), and mutants A, D, G, S, T, L, F, R, H, and pS, are indicated by arrows. The p8-N27 mutant is not detected by this antibody.
Fig. 24. In vitro translation protein products of wild type pS-N27 and mutants of Lys-249 (P2'). Site-specific mutagenesis analysis was performed as described in Fig. 19. Specific mutations at the P2' position are indicated by amino acids A (Ala), M (Met), R (Arg) and T (Thr) at the top of the corresponding lanes and lane Wt represents wild type Lys-249 pS-N27 translation products.
<table>
<thead>
<tr>
<th>Wt</th>
<th>A</th>
<th>M</th>
<th>R</th>
<th>T</th>
</tr>
</thead>
</table>

**Lys-249 (P2')**

- **p28**
Fig. 25. *In vitro* translation protein products of wild type pS-N27 and mutants of Pro-250 (P3'). Site-specific mutagenesis analysis was performed as described in Fig. 19. Specific mutations at the P3' position are indicated by amino acids H (His), L (Leu), and R (Arg) at the top of the corresponding lanes and lane Wt represents wild type Pro-250 pS-N27 translation products.
of p28 (Fig. 21 and 22). As this P2-gated system was shown to be important for some viral cleavage sites, Val.250 (with small increases in side chain length, e.g., Gly, Lys, or Thr in the pseudolipopeptide or Thr in the pseudodipeptide) may be gatekeeper for this cleavage site. In this regard, it should be noted that the mutant with Ala.250 (Leu, polypeptide) or Thr.250 (Leu, polypeptide) has no apparent impact in a similar cleavage reaction. The C-terminal Arg.343 also reduced the cleavage efficiency, whereas the N-terminal Lys.101 has little impact. As a result, Arg.343 may play an important role, presumably a structural role, maintaining an appropriate conformation of the cleavage site.
of p28 (Fig. 21 and 22). At the P1' position, which was shown to be important for some viral cleavage sites, Val-248 (with small branched side chain) was replaced by amino acids with similar small side chains (Ala, Gly, Ser or Thr), amino acids with a large side chain or an aromatic chain (Leu, or Phe), or charged amino acids (Asp, Arg, or His). All nine mutant polypeptides had about the same or even slightly higher p28 cleavage as compared to the wild-type polyprotein (Fig. 23 and 27). Thus, this result enabled us to rule out the importance of the P1' position for the cleavage site determination. At two other positions (P2' and P3') downstream of the peptide bond to be cleaved, substitutions of Lys-249 (P2') with Ala, Met, Arg, or Thr, and Pro-250 (P3') with His, Leu, or Arg had no effect on p28 cleavage (Fig. 24, 25 and 27). Thus, these five positions are less important for cleavage site determination.

Interestingly, mutation of the P5 position, Arg-243, did have an impact on the cleavage of p28. Mutation of Arg to Ala, Ile, or Thr resulted in a significant reduction in p28 (Fig. 26 and 27). Substitution of Arg with Lys, another similar positively charged amino acid, also reduced the cleavage of p28 (Fig. 26 and 27). This result is somewhat surprising because Lys is the wild-type amino acid at this position in the A59 strain of MHV. This observation raises a question as to the role of a different amino acid at the P5 position for MHV JHM versus A59 (see Discussion). The results of all the mutation analysis at the P5 position indicate that Arg-243 may play an important role, presumably a structural role, maintaining an appropriate conformation of the cleavage site.
Fig. 26. *In vitro* translation protein products of wild type pS-N27 and mutants of Arg-243 (P5). Site-specific mutagenesis analysis was performed as described in Fig. 19. Specific mutations at the P5 position are indicated by amino acids A (Ala), I (Ile), K (Lys) and T (Thr) at the top of the corresponding lanes and lane Wt represents wild type Arg-243 pS-N27 translation products.
Fig. 27. Immunoprecipitation analysis of p28 from *in vitro* translation products of wild type pS-N27 and mutants at positions from P5 to P3' of the cleavage site. Site-specific mutagenesis on the positions from P5 to P3' was performed as described in Materials and Methods. Translation products were generated by *in vitro* transcription/translation of various pS-N27 plasmid DNAs and equal TCA-precipitable counts from each reaction were immunoprecipitated using anti-sera specific to p28. Immunoprecipitation products were analyzed by 10% SDS-PAGE. The region corresponding to p28 is shown for each mutant. The amino acid at the site examined is shown above each lane.
Previous studies on proteolytic processing of p28 suggested that the cleavage site for p28 generation is the Tyr-257 and Gly-258 amenable bond (see above). The cleavage of the N-terminal number of sites (Palm to Ala) and Gly contribute to the mobility of the N-terminus. In addition, it is to have any effect on the KRSY site.

<table>
<thead>
<tr>
<th>Wt</th>
<th>Mutants</th>
</tr>
</thead>
<tbody>
<tr>
<td>R A I K T</td>
<td></td>
</tr>
<tr>
<td>P5</td>
<td></td>
</tr>
<tr>
<td>G A D F H L S V Y</td>
<td></td>
</tr>
<tr>
<td>P4</td>
<td></td>
</tr>
<tr>
<td>Y C F P R S</td>
<td></td>
</tr>
<tr>
<td>P3</td>
<td></td>
</tr>
<tr>
<td>P2</td>
<td></td>
</tr>
<tr>
<td>G A D N V</td>
<td></td>
</tr>
<tr>
<td>P1</td>
<td></td>
</tr>
<tr>
<td>P1'</td>
<td></td>
</tr>
<tr>
<td>V A D G S T L F R H</td>
<td></td>
</tr>
<tr>
<td>P2'</td>
<td></td>
</tr>
<tr>
<td>K A M R T</td>
<td></td>
</tr>
<tr>
<td>P3'</td>
<td></td>
</tr>
<tr>
<td>P H L R</td>
<td></td>
</tr>
</tbody>
</table>
Previous studies on proteolytic processing of p28 suggested that the cleavage site for p28 generation is the Tyr-257 and Gly-258 dipeptide bond (Soe et al., 1987). This suggestion was based primarily on a search of the N-terminal region of the polymerase polyprotein for homology to a limited number of known viral proteinase cleavage sites. Certain Tyr-Gly (Y-G) sites (among others) are recognized by enterovirus 2A proteinases (Palmenberg, 1990). Mutation of Tyr-257 to either Ala or Asp and of Gly-258 to Ala, Asp, Phe or Leu had no effect on p28 cleavage nor any effect on the mobility of the p28 protein (Fig. 28). These results indicated that Tyr-257 and Gly-258 are not essential for proteolytic processing of p28. In addition, the substitutions of residues at positions 257 and 258 do not appear to have any effect on the presentation of the nearby authentic cleavage site Gly-247-Val-248.

Overall, the results from our site-specific mutagenesis analysis at eight amino acids surrounding p28 cleavage site Gly-247-Val-248, indicate that Gly-247 at the P1 position and Arg-248 at the P2 position are the major determinants for efficient cleavage of p28 by MHV PCP-1. In addition, among the surrounding residues, Arg-243 at the P5 position may be important, perhaps in maintaining a special conformation for the cleavage site.

In summary, the studies of this dissertation showed that the first papain-like cysteine proteinase (PCP-1) domain responsible for autocleaving the N-terminal peptide p28 from the MHV-JHM polymerase
Fig. 28. Effect of site-specific mutation at the previously predicted p28 cleavage site Tyr-257 and Gly-258 on proteolytic processing of pS-N27 translation products. Site-specific mutagenesis was performed as described in the Materials and Methods. Specific mutations at Gly-258 and Tyr-257 are indicated by amino acids Ala, Asp, Phe, Leu, Lys, and Arg at the top of the corresponding lanes. Additional lanes are: M, 14C-labeled molecular weight polypeptide marker, masses are given in kilodaltons; N, no RNA; Wt, wild type pS-N27 translation products.
polyprotein is located at the 5' end 3.5-4.2 kb of gene 1. MHV PCP-1 is indeed a member of the family of papain-like cysteine proteinases which utilizes Cys-1137 and His-1288 as the catalytic residues. Additionally, MHV PCP-1 cleaves at Gly-247 and Val-248 dipeptide bond to release p28. For this p28 cleavage site, Gly-247 (P1) and Arg-246 (P2) are major determinants, and an important conformation is required for this cleavage site which might be determined by some surrounding residues including Arg-243 at the P5 position.
Proteolytic processing of a polyprotein precursor is a common and essential event in the replication of positive-strand RNA viruses including retroviruses. Viral polyprotein precursors are processed into functional subunits by viral and/or host cell proteinases (Krausslich and Wimmer, 1988; Dougherty and Semler, 1993). During replication of MHV, three viral-encoded proteinase domains, PCP-1, PCP-2, and 3C-Pro, are proposed to be responsible for separating the polymerase polyprotein (more than 750 kDa) into functional protein products. To date, the activity of only the first papain-like cysteine proteinase domain, PCP-1, has been demonstrated (Baker et al., 1989). PCP-1 acts to autocleave the amino-terminal portion of the MHV polymerase polyprotein, releasing p28 (Baker et al., 1989; 1993). The goal of this dissertation was to understand this particular processing event by characterizing the proteinase functional domain and the cleavage site specificity of the proteinase.

This study demonstrated that the MHV-JHM PCP-1 proteinase domain, which is responsible for the proteolysis of p28 protein from the N-terminus of the MHV gene 1 polyprotein, is located at 3.5-4.2 kb from the 5' end of gene 1 (see Fig. 3 and Fig. 9). It was shown that some of the intervening sequence between the proteinase domain and the cleavage site...
can be deleted without a significant effect on the proteolytic cleavage (Baker et al., 1993). Also, the kinetics of p28 processing during in vitro translation of the MHV genomic RNA (data not shown) suggest that p28 is cleaved as soon as the proteinase domain is synthesized. Since the proteinase has been shown to work in cis to release p28, p28 processing is probably a cotranslational event (Baker et al., 1989). In addition, it is likely that the tertiary conformation of the polyprotein, particularly that formed by sequences neighboring the cleavage site or the proteinase domain, are critical for efficient proteolytic processing. This conclusion is supported by the failure of pT7-N29, which has an internal deletion of 1.6 kb, to generate p28.

To determine if MHV PCP-1 is a cysteine proteinase, we tested a variety of proteinase inhibitors and monitored the effect of these inhibitors on the proteolytic processing of p28. The level of p28 processing was shown to be specifically reduced by cysteine proteinase inhibitors, but not by inhibitors of serine proteinases, aspartic proteinases, and metalloproteinases (see Fig. 9). Furthermore, mutagenesis studies have identified Cys-1137 and His-1288 as critical residues for the proteinase activity of MHV-JHM PCP-1 (Baker et al., 1993). These two residues have been proposed to be the catalytic residues of this proteinase on the basis of the sequence similarity with cellular and viral papain-like cysteine proteinases (Gorbalenya et al., 1991). The substitutions of amino acids of diverse properties for Cys-1137 and His-1288 resulted in the total loss of proteinase activity which confirmed that these amino acids are indeed at the catalytic
center. In contrast, substitutions of the neighboring His-1317 or Cys-1172 residues did not inactivate proteinase activity, except for the Pro substitution at the His-1317 position, which is expected to alter global protein conformation. These studies thus strongly supported the computer prediction that MHV PCP-1 is a new member of the family of papain-like cysteine proteinases.

Coronavirus PCP-1 adds to the growing list of the viral papain-like proteinases. Previous studies have suggested that Sindbis virus (SV), equine arteritis virus (EAV), rubella virus, hepatitis E virus, plant virus tobacco etch virus (TEV) and hypovirulence-associated virus (HAV) all contain proteinase domains with homology to the family of cysteine proteinases and have conserved Cys and His residues essential for catalytic activity (Gorbalenya et al., 1991; Koonin et al., 1992; Dougherty and Semler, 1993). Among these viral cysteine proteinases, MHV PCP-1 appears to be related to PCP in another coronavirus superfamily member, equine arteritis virus (EAV) (den Boon et al., 1991; Snijder et al., 1992). EAV has a similar replication strategy and genomic organization to that of coronavirus MHV, but has only a 13-kb RNA genome. The product of the EAV papain-like cysteine proteinase is a 30 kDa protein, which is also derived from the N terminus of the ORF1a protein of the polymerase gene, but the cleavage site is located at the C-terminal end of the proteinase. Like MHV PCP-1, this proteinase also acts only in cis (Snijder et al., 1992). For both MHV and EAV, the function of the released amino terminal peptide (p28 or p30) is not yet clear.
The second question addressed in this research was the site recognized by MHV PCP1 for the cleavage release of p28. By N-terminal protein sequence analysis of the *in vitro* cleaved protein downstream of p28, the cleavage site for p28 was identified as the Gly-247 (P1) and Val-248 (P1') dipeptide bond (Fig. 16 and 17). Site-specific mutagenesis of eight residues surrounding the cleavage site from the P5 to P3' positions was performed to further characterize the sequence motif of the p28 cleavage site. A total of 42 mutants were generated, with 3 to 9 mutants at each position. Mutations at P1 Gly-247 and P2 Arg-246 dramatically decreased the amount of p28 cleaved from the mutant polyproteins compared to that from the wild type polyprotein. In contrast, the majority of the mutations at the five other positions (P3, P4, P1', P2' and P3') had little or no effect on p28 processing and only a few mutants led to significant p28 cleavage. Surprising results were seen from the mutational analysis at the P5 position (Arg-243). Mutation of Arg-243 to Lys, another positive charged amino acid also reduced p28 cleavage as did three other mutations (Ala, Ile and Thr). The P5 residue is the only residue that varies within positions from P5 to P3' between two p28 cleavage sites of MHV-JHM and -A59 strains (Soe *et al.*, 1987; Pachuk *et al.*, 1989). The observation that strain JHM prefers Arg at the P5 position instead of Lys as in strain A59, is interesting, as it might result from a minor difference between PCP-1 domains of A59 and JHM.

The extensive mutagenesis analysis of the cleavage site has shown that Gly-247 at the P1 position and Arg-246 at the P2 position are major
determinants of the p28 cleavage site recognized by MHV PCP-1. Since the Arg-243-Gly-244 dipeptide, appears only three residues upstream of the cleavage site (Arg-246--Gly-247), but is not recognized by MHV PCP-1, the conformation of the p28 cleavage site Gly-247-Val-248 (unknown at this point) or other surrounding amino acids must also be important for defining the cleavage site. Surrounding residues such as Arg-243 (P5), may be involved in maintaining the correct conformation of the cleavage site or for interacting with the substrate-binding pocket of the MHV PCP-1 proteinase domain.

The Gly-Val dipeptide bond at the p28 cleavage site appears to be similar to cleavage sites used by other viral papain-like cysteine proteinases (see Fig. 29). The cleavage site for the amino terminal product of the equine arteritis virus (EAV) polymerase polyprotein has been defined (Snijder et al., 1992). The EAV papain-like cysteine proteinase (PCP) cleaves at a Gly-Gly dipeptide bond to release an amino terminal protein of 30 kDa. Snijder and co-workers have demonstrated by site-directed mutagenesis that the Gly residue at the P1 position was important for efficient autoproteolytic processing whereas the Gly residue at the P1' position was more flexible to substitution (Snijder et al., 1992). In EAV, when the P1 Gly was replaced with Ala, there was a reduction in the amount of p30 detected; when the P1 Gly was replaced with Val, little or no processing of p30 was detected. However, the Gly at the P1' position was tolerant of mutation to Ala or Val and proteolytic processing of p30 was detected.
**Fig. 29.** Comparison of known cleavage sites for viral papain-like cysteine proteinases. Viral proteinases are indicated, and protein junctions cleaved by Sindbis virus nsP2 proteinase are denoted. Amino acid sequences are shown in single letter codes and the arrow indicates the dipeptide bond being cleaved. MHV stands for mouse hepatitis virus, EAV for equine arteritis virus, TEV for tobacco etch virus, HAV for hypovirulence-associated virus, and SV for Sindbis virus.
<table>
<thead>
<tr>
<th>Viral Proteinase</th>
<th>P5</th>
<th>P4</th>
<th>P3</th>
<th>P2</th>
<th>P1</th>
<th>P1'</th>
<th>P2'</th>
<th>P3'</th>
<th>P4'</th>
</tr>
</thead>
<tbody>
<tr>
<td>MHV-JHM PCP1</td>
<td>R</td>
<td>G</td>
<td>Y</td>
<td>R</td>
<td>G</td>
<td>V</td>
<td>K</td>
<td>P</td>
<td>I</td>
</tr>
<tr>
<td>EAV nsP1</td>
<td>A</td>
<td>G</td>
<td>N</td>
<td>Y</td>
<td>G</td>
<td>G</td>
<td>Y</td>
<td>N</td>
<td>P</td>
</tr>
<tr>
<td>TEV HC-Pro</td>
<td>T</td>
<td>Y</td>
<td>N</td>
<td>V</td>
<td>G</td>
<td>G</td>
<td>M</td>
<td>N</td>
<td>R</td>
</tr>
<tr>
<td>HAV p29</td>
<td>L</td>
<td>A</td>
<td>R</td>
<td>I</td>
<td>G</td>
<td>G</td>
<td>R</td>
<td>L</td>
<td>N</td>
</tr>
<tr>
<td>HAV p48</td>
<td>D</td>
<td>I</td>
<td>L</td>
<td>V</td>
<td>G</td>
<td>A</td>
<td>E</td>
<td>E</td>
<td>G</td>
</tr>
<tr>
<td>SV nsP2</td>
<td>(nsP1-nsP2) A</td>
<td>D</td>
<td>I</td>
<td>G</td>
<td>A</td>
<td>A</td>
<td>L</td>
<td>V</td>
<td>E</td>
</tr>
<tr>
<td></td>
<td>(nsP2-nsP3) D</td>
<td>G</td>
<td>V</td>
<td>G</td>
<td>A</td>
<td>A</td>
<td>P</td>
<td>S</td>
<td>Y</td>
</tr>
<tr>
<td></td>
<td>(nsP3-nsP4) T</td>
<td>G</td>
<td>V</td>
<td>G</td>
<td>G</td>
<td>Y</td>
<td>I</td>
<td>F</td>
<td>S</td>
</tr>
</tbody>
</table>
The best characterized cleavage sites for all identified viral cysteine proteinases are in two plant viruses: the tobacco etch virus helper component protease (TEV HC-Pro) and the proteinase p48 of hypovirulence-associated virus (HAV) of the chestnut blight fungus. The cleavage site of TEV HC-Pro is also at a Gly-Gly dipeptide bond and both residues (P1 and P1') are critical for efficient cleavage (Carrington et al., 1989; Carrington and Herndon, 1992). There are two PCP proteinases (p29 and p48) reported in HAV. HAV p29 and p48 utilize Gly-Gly and Gly-Ala as their cleavage sites respectively. The P1 Gly residue for both HAV cleavage sites is required for efficient cleavage (Choi et al., 1991a; 1991b; Shapira and Nuss, 1991). Overall, a Gly at the P1 position is required by MHV, EAV, TEV and HAV cysteine proteinases for efficient cleavage activity; but the P1' residue is somewhat flexible for MHV PCP-1, EAV PCP and HAV p48 recognition and processing.

The amino acid residues surrounding these cleavage sites, the P2 and P4 residues of the cleavage sites for both HC-Pro and p48 have been shown to be critical for the cleavage site determination. Similarly, the P2 residue of p28 cleavage site for MHV-JHM PCP-1 is a critical residue, but the P4 residue does not seem to be as important for MHV PCP-1 cleavage as it is for HC-Pro and HAV p48 cleavages. The difference in the sequence motifs of these cleavage sites might result from the differences in the substrate-binding pockets of these viral cysteine proteinases.
Viral cysteine proteinases such as EAV PCP, TEV HC-Pro, and HAV p29 described above have been termed "leader" viral papain-like proteinases (Gorbalenya et al., 1991). These leader proteinases are located near their N termini of the respective polyproteins and mediate a single cleavage event in cis (or intramolecularly) at their own C termini. In these polyproteins, the cleavage sites are located 50-100 amino acids downstream of the catalytic cysteine residues (Snijder et al., 1992; Carrington et al., 1989; Choi et al., 1991a; 1991b). In contrast to the leader proteinase, the MHV p28 cleavage site (Gly-247-Val-248) is upstream of MHV PCP-1 and is separated from the catalytic residue Cys-1137 of PCP-1 by almost 900 amino acids (approximately 3 kb). We believe that MHV PCP-1 may be more similar to Sindbis virus nsP2 (classified as a "main" proteinase by Gorbalenya et al., 1991) which cleaves the polyprotein precursor at multiple sites both in cis and in trans (Hardy and Strauss, 1989; de Groot et al., 1990, Shirako and Strauss, 1990). Although MHV PCP-1 is so far only known to release p28 from the N-terminus of the polymerase polyprotein, it is possible that the mature polymerase polyprotein may possess several sites which may be cleaved in trans or in cis by MHV PCP-1. Questions remain about what the mature form of the PCP-1 protein is and whether there are additional cleavage sites in the polyprotein.

Compared to cellular proteinases, viral proteinases have been shown to recognize specific amino acid sequence motifs (Dougherty and Semler, 1993). The characterization of the cis-cleavage site of p28 for MHV PCP-1 should help us predict additional cleavage sites in the polyprotein.
Indeed, there are 16 other Arg-Gly (RG) sites in MHV-JHM polymerase polyprotein which may serve as potential targets for trans-cleavage by the mature PCP-1 product. We plan to subject these RG sites (the major determinants of the putative cleavage site recognized by MHV PCP-1) to future investigation in order to determine which ones are in fact recognized by the MHV PCP-1. These studies will help us to identify potential new processed products of MHV polymerase polyprotein.

MHV PCP-1 is the first of three predicted proteinases in MHV demonstrated to be active. Interestingly, there is a second predicted papain-like cysteine proteinase domain approximately 2 kb downstream from the proteinase domain described here. MHV PCP-2 domain has a higher sequence homology to the only PCP domain suggested in the chicken coronavirus IBV ORF1a polyprotein than does PCP-1. The biological activity of MHV PCP-2 domain has not yet been demonstrated. Because of the finding that the first papain-like proteinase works only in cis and requires at least some spacing regions between the cleavage site and the proteinase domain, it seems unlikely that the downstream papain-like proteinase domain will also cleave p28. Rather, it is more likely that the second proteinase has other cleavage sites. Significantly, the IBV gene 1 polyprotein has only one PCP domain, a 3C-Pro domain (Gorbalenya et al., 1989), and the N-terminal portion of ORF1a polyprotein is substantially diverged from that of MHV. IBV also lacks regions corresponding to p28 and the PCP-1 domain. Therefore, MHV PCP-2 may cleave in a region of
the coronavirus ORF1a polyprotein which is more conserved between IBV and MHV.

The two papain-like cysteine proteinase domains are generally believed to mainly process the polyprotein portion encoded from the MHV ORF1a, so the 3C-Pro proteinase domain is probably responsible for the processing of the polyprotein portion from the ORF1b which is about 300 kDa. Preliminary evidence suggests that MHV 3C-Pro is active, and there are multiple potential cleavage sites which have been predicted in the ORF1b-encoded polyprotein by searching conventional poliovirus 3C cleavage sites (Q/E-G/S/A) (Lee et al., 1991). Further confirmation of the proteinase activity of MHV PCP-2 and 3C-Pro domains are the prerequisite for future studies of the complicated proteolytic processing of MHV polymerase polyprotein.

The long term goal of the studies of the proteolytic processing of MHV polymerase polyprotein is to define the complete proteolytic processing pathway of this polyprotein. Preliminary evidence has suggested that a plethora of cleavage events are likely to occur during or after translation of the coronavirus polymerase polyprotein (Denison et al., 1991; 1992). Viral proteinases in Sindbis virus and poliovirus have been shown to play a pivotal role in the regulation of processing and enzymatic activity of viral nonstructural proteins including the polymerase protein (Shirako and Strauss, 1990; Strauss and Strauss, 1990; de Groot et al., 1990, Lemm and Rice, 1993; Shirako and Strauss, 1994; Dewalt and Semler, 1987).
Conceivably, MHV proteinases may play similar roles in the regulation of coronavirus replication. Future experiments can be directed to study the functional importance of p28 processing in coronavirus replication by introducing mutations at the p28 cleavage site by the unique high-frequency RNA recombination observed in MHV. Ultimately, the functions of p28 and other processed protein products of the polymerase polyprotein will be pursued. By studying the functions of all of these processed protein products from coronavirus polymerase polyprotein, it is possible to understand the molecular basis of the complex replication in coronavirus replication. Hopefully in the future, the studies of the molecular mechanism of coronavirus proteolytic processing will provide useful information for designing antiviral drugs to interrupt coronavirus replication and block coronavirus infection and pathogenesis.
LITERATURE CITED


VITA

The author, Shanghong Dong, was born in Hubei, People's of Republic China on October 16, 1966 to Huijie Jiang and Yaping Dong.

In June, 1988, Ms. Dong received the degree of Bachelor of Sciences in Genetics from Wuhan University in Hubei, China. Ms. Dong worked as a research assistant in Wuhan Institute of Biologic Products, Ministry of Public Health, People's Republic of China for the next year. In August, 1989, Ms. Dong was awarded a Loyola University Basic Science Fellowship and entered the Program in Molecular Biology of Loyola University Medical Center for her graduate studies. In January, 1991, Ms. Dong joined Dr. Baker's Laboratory and started her dissertation project.

Ms. Dong will begin her postdoctoral research in Dr. Nussenzweig Laboratory in New York University Medical Center in September, 1994.
ABSTRACTS AND PRESENTATIONS


PUBLICATIONS


The dissertation submitted by Shanghong Dong has been read and approved by the following committee:

Dr. Susan Baker, Ph. D. (Director)
Assistant Professor
Department of Microbiology and Immunology
Loyola University of Chicago

Dr. Leonard Erickson, Ph. D.
Professor
Department of Medicine and Pharmacology
Loyola University of Chicago

Dr. Allen Frankfater, Ph. D.
Associate Professor
Department of Molecular and Cellular Biochemistry
Loyola University of Chicago

Dr. Mark Peeples, Ph. D.
Professor
Department of Immunology/Microbiology
Rush-Presbyterian-St. Luke's Medical Center

Dr. Kenneth Thompson, Ph. D.
Professor
Department of Microbiology
University of Chicago, Medical Center

The final copies have been examined by the director of the dissertation and the signature which appears below verifies the fact that any necessary changes have been incorporated and that the dissertation is now given final approval by the Committee with reference to the content and form. The dissertation is therefore accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

August 24, 1994

Susan Baker
Director's Signature