CTMI (2005) 287:57–94

The Coronavirus Replicase

J. Ziebuhr

Institute of Virology and Immunology, University of Würzburg, Versbacher Str. 7, 97078 Würzburg, Germany
j.ziebuhr@mail.uni-wuerzburg.de

Abstract Coronavirus genome replication and transcription take place at cytoplasmic membranes and involve coordinated processes of both continuous and discontinuous RNA synthesis that are mediated by the viral replicase, a huge protein complex encoded by the 20-kb replicase gene. The replicase complex is believed to be comprised of up to 16 viral subunits and a number of cellular proteins. Besides RNA-dependent RNA polymerase, RNA helicase, and protease activities, which are common to RNA viruses, the coronavirus replicase was recently predicted to employ a variety of RNA processing enzymes that are not (or extremely rarely) found in other RNA viruses and include putative sequence-specific endoribonuclease, 3′-to-5′ exoribonuclease, 2′-*O*-ribose methyltransferase, ADP ribose 1″-phosphatase and, in a subset of group 2 coronaviruses, cyclic phosphodiesterase activities. This chapter reviews (1) the organization of the coronavirus replicase gene, (2) the proteolytic processing of the replicase by viral proteases, (3) the available functional and structural information on individual subunits of the replicase, such as proteases, RNA helicase, and the RNA-dependent RNA polymerase, and (4) the subcellular localization of coronavirus proteins involved in RNA synthesis. Although many molecular details of the coronavirus life cycle remain to be investigated, the available information suggests that these viruses and their distant nidovirus relatives employ a unique collection of enzymatic activities and other protein functions to synthesize a set of 5′-leader-containing subgenomic mRNAs and to replicate the largest RNA virus genomes currently known.

1 Introduction

Plus-strand (+) RNA viruses exhibit an enormous genetic diversity that also applies to their RNA synthesis machinery. The RNA-dependent RNA polymerase (RdRp) is the only enzyme to be absolutely conserved, whereas other replicative and accessory protein domains vary considerably, in terms of both number and arrangement in the polyprotein (Koonin and Dolja 1993). Despite this diversity, phylogenetic relationships have been identified and used to group +RNA viruses into large superfamilies (or classes) (Goldbach 1987; Strauss and Strauss 1988; Koonin and Dolja 1993). As few as three superfamilies, the picornavirus-like, flavivirus-like and alphavirus-like viruses, were proposed to accommodate the vast majority of +RNA viruses infecting animals, plants, and microorganisms (Koonin and Dolja 1993). Interestingly, coronaviruses were among the few exceptions that did not easily fit into one of the established superfamilies; and the sequence analysis and characterization of arteri-, toro-, and roniviruses suggested that coronaviruses and their relatives may indeed exemplify a viral life form that, in several fundamental aspects, differs from that of other +RNA viruses (Gorbalenya et al. 1989c; Snijder et al. 1990a; den Boon et al. 1991; Snijder and Horzinek 1993; de Vries et al. 1997; Lai and Cavanagh 1997; Snijder and Meulenberg 1998; Cowley et al. 2000). Thus coronaviruses (and all their relatives) (1) produce a nested set of 3′-coterminal mRNAs (Lai et al. 1983; Spaan et al. 1983), (2) use ribosomal frameshifting into the –1 frame to express their key replicative functions (Brierley et al. 1987, 1989), (3) have a unique set of conserved functional domains that are arranged in the viral polyproteins in the following order: chymotrypsin-like proteinase, RdRp, helicase, and endoribonuclease (from N- to C-terminus) (Gorbalenya et al. 1989c; Gorbalenya 2001; Snijder et al. 2003), and (4) use RdRp and helicase activities that, based on the conservation of signature motifs, have been classified as belonging to the RdRp and helicase superfamilies 1, respectively (Koonin and Dolja 1993). Both the combination of two superfamily 1 domains and their sequential order in the polyprotein, with RdRp preceding the helicase, is extremely unusual (if not unique) among +RNA viruses. On the basis of these and other common properties, a new virus order, the *Nidovirales*, was introduced several years ago (Cavanagh 1997). At present, there is only little information on the toro- and ronivirus replicases, whereas information on the replicases of corona- and arteriviruses is accumulating rapidly. On the basis of both serological relationships and sequence sim-

ilarity, coronaviruses have been classified into three groups (Siddell 1995), with human coronavirus 229E (HCoV-229E, group 1), porcine transmissible gastroenteritis virus (TGEV, group 1), mouse hepatitis virus (MHV, group 2), and avian infectious bronchitis virus (IBV, group 3) being the best-studied coronaviruses to date. Because of its medical importance, SARS coronavirus (SARS-CoV) (tentatively classified as belonging to group 2) (Snijder et al. 2003) is currently becoming a major topic of coronavirus research.

2
Organization and Expression of the Replicase Gene

Complete genome sequences are currently available for seven species of coronaviruses, IBV (Boursnell et al. 1987), MHV (Bredenbeek et al. 1990; Lee et al. 1991; Bonilla et al. 1994), HCoV-229E (Herold et al. 1993), TGEV (Eleouet et al. 1995; Penzes et al. 2001), porcine epidemic diarrhea virus (PEDV) (Kocherhans et al. 2001), bovine coronavirus (Chouljenko et al. 2001), and SARS-CoV (Marra et al. 2003; Rota et al. 2003). In some cases (for example, SARS-CoV) complete genome sequences are available for several or even multiple isolates (Ruan et al. 2003). The genome sizes of coronaviruses range between 27.3 (HCoV-229E) and 31.3 (MHV) kb, making coronaviruses the largest RNA viruses currently known. About two-thirds of the coronavirus genome (~20,000 bases) are devoted to encoding the viral replicase that mediates viral RNA synthesis (Thiel et al. 2001b) and, possibly, other functions. The replicase gene is comprised of two large open reading frames, designated ORF1a and ORF1b, that are located at the 5′ end of the genome. The upstream ORF1a encodes a polyprotein of 450–500 kDa, termed polyprotein (pp)1a, whereas ORF1a and ORF1b together encode pp1ab (750–800 kDa) (Fig. 1). Expression of the C-terminal, ORF1b-encoded half of pp1ab requires a (–1) ribosomal frameshift during translation. It is generally accepted that frameshifting depends on two critical elements, the "slippery" sequence, UUUAAAC, at which the ribosome shifts into the (–1) reading frame and a tripartite RNA pseudoknot structure located more downstream, near the ORF1a/1b junction (Brierley et al. 1987, 1989; Herold and Siddell 1993). In vitro experiments using reticulocyte lysates indicate that frameshifting occurs in about 20%–30% of the translation events, but it is not known whether this reflects the situation in vivo. The fact that the core replicative functions, RdRp and helicase, are encoded by ORF1b implies that their expression critically depends

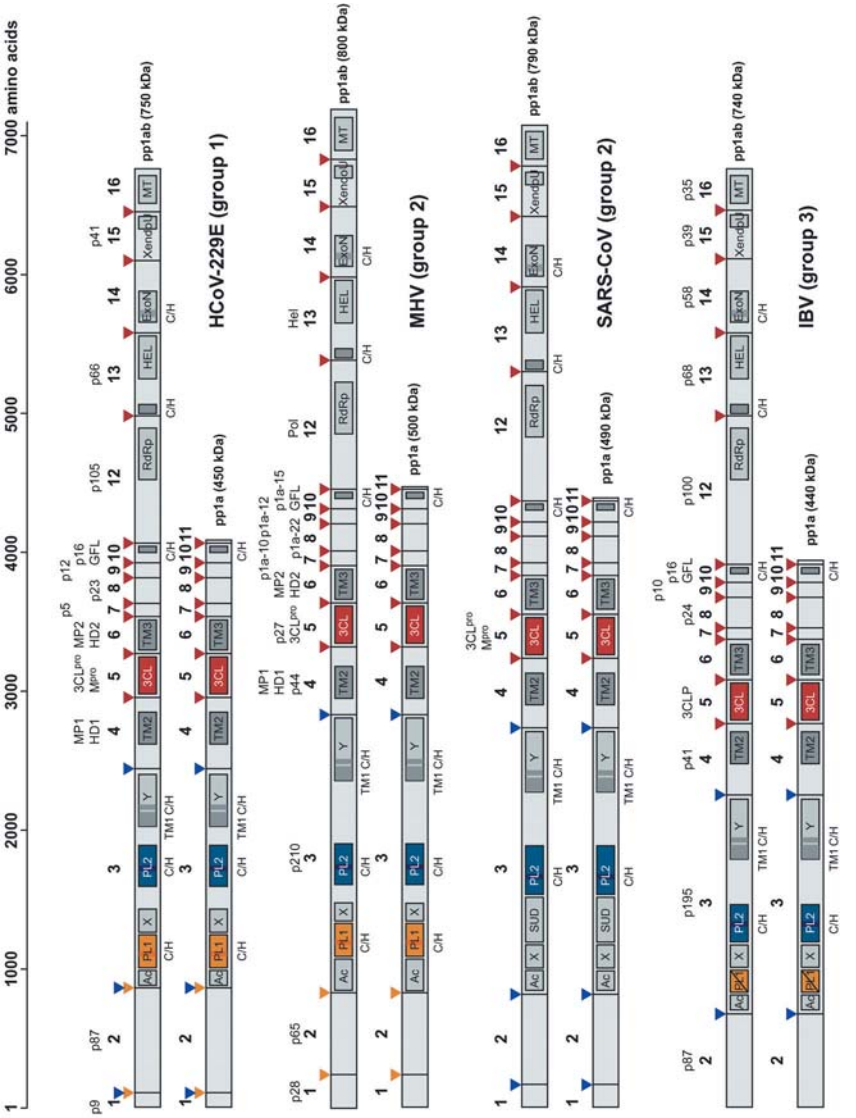

Fig. 1. Overview of the domain organization and proteolytic processing of coronavirus replicase polyproteins. Shown are the replicase polyproteins pp1a and pp1ab of human coronavirus 229E (*HCoV-229E*), mouse hepatitis virus (*MHV*), SARS coronavirus (*SARS-CoV*), and avian infectious bronchitis virus (*IBV*). The processing end-products of pp1a are designated nonstructural proteins (nsp) 1 to nsp11, and those

on ribosomal frameshifting, suggesting a requirement for a specific molar ratio between ORF1a- and ORF1b-encoded protein functions.

3 Replicase Polyproteins

3.1 Functional Domains

Initial sequence analyses in the late 1980s suggested a large divergence of the coronavirus replicase from the replicative machinery of other +RNA viruses. Accordingly, at this time, only very few functional predictions could be made for the ~800-kDa replicative polyproteins of coronaviruses (Boursnell et al. 1987). In 1989, a detailed comparative sequence analysis of the IBV replicase gene (Gorbalenya et al. 1989c) was pub-

of pp1ab are designated nsp1 to nsp10 and nsp12 to nsp16. Note that nsp1 to nsp10 may be released by proteolytic processing of either pp1a or pp1ab, whereas nsp11 is processed from pp1a and nsp12 to nsp16 are processed from pp1ab. nsp11 and nsp12 share a number of residues at the N-terminus. Alternative names that have been used in the past to designate specific processing products are given. Cleavage sites that are processed by the viral main proteinase are indicated by *red arrowheads*, and sites that are processed by the accessory papainlike proteinases 1 and 2 are indicated by *orange* and *blue arrowheads*, respectively. *Ac*, acidic domain (Ziebuhr et al. 2001); *PL1*, accessory papainlike cysteine proteinase 1 (Baker et al. 1989, 1993; Gorbalenya et al. 1991; Herold et al. 1998); *X*, X domain (Gorbalenya et al. 1991), which is predicted to have adenosine diphosphate-ribose 1″-phosphatase activity (Snijder et al. 2003); *SUD*, SARS-CoV unique domain (Snijder et al. 2003); *PL2*, accessory papainlike cysteine proteinase 2 (Gorbalenya et al. 1991; Liu et al. 1995; Kanjanahaluethai and Baker 2000; Ziebuhr et al. 2001); *Y*, Y domain containing a transmembrane domain and a putative Cys/His-rich metal-binding domain; *TM1*, *TM2*, and *TM3*, putative transmembrane domains 1 to 3; *3CL*, 3C-like main proteinase (Gorbalenya et al. 1989c; Liu and Brown 1995; Ziebuhr et al. 1995; Lu et al. 1995); *RdRp*, putative RNA-dependent RNA polymerase domain (Gorbalenya et al. 1989c); *HEL*, helicase domain (Seybert et al. 2000a); *ExoN*, putative 3′-to-5′ exonuclease (Snijder et al. 2003); *XendoU*, putative poly(U)-specific endoribonuclease (Snijder et al. 2003); *MT*, putative *S*-adenosylmethionine-dependent ribose 2′-*O*-methyltransferase (Snijder et al. 2003); *C/H*, Cys/His-rich domains predicted to bind metal ions. Note that IBV pp1a and pp1ab do not have a counterpart of nsp1 of other coronaviruses. The papainlike cysteine proteinase 1 of IBV is *crossed out* to indicate that the domain is proteolytically inactive

lished in which the RdRp and NTPase/helicase domains were predicted to be encoded by the 5′ region of ORF1b. Furthermore, a putative chymotrypsin-like (picornavirus 3C-like) cysteine proteinase domain ($3CL^{pro}$) was identified in ORF1a and predictions on putative cleavage sites in the C-terminal regions of pp1a and pp1ab were made. The proteinase was found to be flanked by membrane domains on both sides. The coronavirus replicative proteins were proposed to be only extremely distantly related to the corresponding homologs of other +RNA viruses, and many of the pp1a/pp1ab-encoded enzymes appeared to have unique structural properties. Thus, for example, the helicase was proposed to be linked at its N-terminus to a complex zinc-binding domain (ZBD) consisting of 12 Cys/His residues (see below). In several cases, mutations in otherwise strictly conserved signature sequences were found. Thus the typical G–D–D signature of the conserved RdRp motif VI (Koonin 1991) was found to be replaced by S–D–D in the coronavirus homolog and the G(A)–X–H motif conserved in the S1 subsite of the substrate-binding pocket of picornavirus 3C proteinases (Gorbalenya et al. 1989a, 1989c) was substituted with Y–M–H. The predictions on functional domains, putative active-site residues, and proteinase cleavage sites were continuously elaborated and extended when more coronavirus replicase sequences became available (Gorbalenya et al. 1991; Lee et al. 1991; Herold et al. 1993; Eleouet et al. 1995; Chouljenko et al. 2001; Kocherhans et al. 2001; Penzes et al. 2001; Ziebuhr et al. 2001; Snijder et al. 2003). In these studies, papainlike cysteine proteinase (PL^{pro}) domains (Gorbalenya et al. 1991), a conserved domain of corona-, alpha-, and rubiviruses, termed X[1] (Gorbalenya et al. 1991), an acidic domain (Ac) of unknown function, and a domain (termed Y) with putative metal-binding and membrane-targeting functions (Ziebuhr et al. 2001) were identified in the coronavirus ORF1a sequence (Fig. 1). Overall, the sequence similarities between the replicase genes of prototypic viruses from the three coronavirus groups corresponded well to those of the structural protein regions, providing support for the traditional classification of coronaviruses into three groups, which previously was based on structural protein sequence relationships and serological cross-reactivities (Siddell 1995).

Recently, the list of putative enzymes involved in coronavirus RNA synthesis was extended considerably. Thus, in the context of a bioinformatics study of the SARS-CoV genome, as many as five (putative) coro-

[1] The X domain has recently been predicted to be an adenosine diphosphate-ribose 1″-phosphatase (ADRP).

naviral RNA processing activities were identified (Snijder et al. 2003) (Fig. 1). These include (1) a 3′-to-5′ exonuclease (ExoN) of the DEDD superfamily (Zuo and Deutscher 2001), (2) a poly(U)-specific endoribonuclease (XendoU) (Laneve et al. 2003), (3) an *S*-adenosylmethionine-dependent ribose 2′-*O*-methyltransferase (2′-*O*-MT) of the RrmJ family (Bügl et al. 2000), (4) an ADRP (Martzen et al. 1999), and (5) a cyclic phosphodiesterase (CPD) (Martzen et al. 1999; Nasr and Filipowicz 2000). Four of the activities are conserved in all coronaviruses, indicating their essential role in the coronaviral life cycle. In fact, the number of enzymes predicted to be involved in coronavirus RNA synthesis and modification is unique in RNA viruses and indicates a remarkable functional complexity, which approaches that of DNA replication. Three of the newly identified activities, ExoN (nsp14), XendoU (nsp15), and 2′-*O*-MT (nsp16), are arranged in pp1ab as a single protein block downstream of the RdRp (nsp12) and helicase (nsp13) domains (Fig. 1), suggesting that their activities cooperate in the same metabolic pathway(s). This conclusion is supported by the identification of a stable processing intermediate in IBV-infected cells that exactly comprises these three domains (Xu et al. 2001). It is also supported by the fact that nsp14–16 expression involves common regulatory mechanisms, (1) ribosomal frameshifting and (2) $3CL^{pro}$-mediated proteolysis. As a first clue to possible functions encoded by this gene block in ORF1b, an exciting parallel to cellular RNA processing pathways was found by Snijder et al. (2003). Thus homologs of the coronavirus nsp14–16 processing products cleave and process mRNAs to produce small nucleolar (sno) RNAs that, in turn, guide specific 2′-*O*-ribose methylations of rRNA (Kiss 2001; Filipowicz and Pogacic 2002).

Two other coronavirus domains, CPD and ADRP, both of which do not require ribosomal frameshifting for expression, were speculated to cooperate in a pathway that again has parallels in the cell. Thus two cellular homologs are known to mediate two consecutive steps in the downstream processing of tRNA splicing products. In this pathway, CPD converts adenosine diphosphate ribose 1″-2″ cyclic phosphate (Appr>p) to adenosine diphosphate ribose 1″-phosphate (Appr-1″-p) (Culver et al. 1994) that, in a second reaction, is further processed (probably dephosphorylated) by an ADRP homolog (Martzen et al. 1999).

Obviously, the characterization of the substrate specificities of the newly identified enzymes will now be of major interest and may allow predictions or even conclusions on the functions of these proteins. Both (reverse) genetic and biochemical data will be required to answer the question of whether the RNA processing enzymes are directly involved

in the synthesis and/or processing of viral RNA or rather interfere with (and thereby reprogram) cellular pathways for the benefit of viral replication (or even have other functions).

The observed pattern of conservation in different nidovirus families suggests a functional hierarchy for the five RNA processing activities, with XendoU playing a central role. This enzyme is universally conserved in nidoviruses and was previously referred to as "nidovirus-specific conserved domain" (Snijder et al. 1990b; den Boon et al. 1991; de Vries et al. 1997). In contrast, CPD is only encoded by toroviruses and a subset of group 2 coronaviruses (excluding SARS-CoV) (Snijder et al. 2003). Given that coronaviruses and arteriviruses are generally believed to use very similar replication and transcription strategies, it is intriguing that, out of the four activities conserved in all coronaviruses (ExoN, XendoU, 2′-O-MT, and ADRP), only one activity (XendoU) is conserved in arteriviruses. One may therefore speculate that (1) arterivirus and coronavirus RNA synthesis mechanisms differ in several molecular details or (2) the viruses interact differentially with RNA processing pathways of the host cell. Alternatively, the extra functions encoded by corona- and toroviruses (and, to a lesser extent, roniviruses) may be required to synthesize and maintain the extremely large (~30 kb) RNA genomes of these viruses. Thus, on the basis of its sequence similarity with cellular 3′-to-5′ exonucleases involved in proofreading, repair, and/or recombination, ExoN has been speculated to be involved in related mechanisms that may be required for the life cycle of corona-, toro-, and roniviruses but may be dispensable for the much smaller arteriviruses (Snijder et al. 2003). The significance of the observation that overexpression of nsp14 induces apoptotic changes in the host cell (Liu et al. 2001) remains to be further investigated.

3.2 Proteolytic Processing by Viral Cysteine Proteinases

In common with many other +RNA viruses (Kräusslich and Wimmer 1988; Dougherty and Semler 1993), coronaviruses employ proteolytic processing as a key regulatory mechanism in the expression of their replicative protein functions (Ziebuhr et al. 2000). Proteinase inhibitors that block proteolytic processing also obviate coronavirus replication, illustrating the essential role of pp1a/pp1ab processing for viral RNA synthesis (Kim et al. 1995). On the basis of their physiological role, coronavirus proteinases can be classified into *accessory* proteinases, which are

responsible for cleaving the more divergent N-proximal pp1a/pp1ab regions at two or three sites, and *main* proteinases, which cleave the major part of the polyproteins at 11 conserved sites and also release the conserved key replicative functions, such as RdRp, helicase, and three of the RNA processing domains (Ziebuhr et al. 2000; Snijder et al. 2003). All coronaviruses encode one main proteinase and, depending on the virus (see below and Fig. 1), one or two accessory proteinases. The accessory proteinases are papainlike cysteine proteinases that are designated PL^{pro} ($PL1^{pro}$ and $PL2^{pro}$). The main proteinase is a cysteine proteinase with a serine proteinase-like structure (Anand et al. 2002). In previous publications, two alternative designations have been used for this protein. The name *main proteinase*, M^{pro}, is generally used to stress the dominant physiological role of this proteinase in coronavirus gene expression, whereas the name *3C-like proteinase* is used to stress the (distant) relationship with picornavirus 3C proteinases, which is based on a common chymotrypsin-like two-β-barrel structure and similar substrate specificities (Gorbalenya et al. 1989a,c; Ziebuhr et al. 2000). Despite this relationship, there are also important structural differences between picornavirus and coronavirus chymotrypsin-like proteinases (see below).

Peptide cleavage data obtained for several coronavirus main proteinases revealed differential processing kinetics for specific sites. The order of cleavages was found to be conserved among coronaviruses and appears to depend on the accessibility of specific sites in the context of the polyprotein (Piñon et al. 1999) as well as the primary and secondary structures of a given cleavage site. Thus deviation from the $3CL^{pro}$ cleavage site consensus sequence, L–Q|(A,S,G), resulted in most cases in significantly reduced cleavage efficiencies (Ziebuhr and Siddell 1999; Hegyi and Ziebuhr 2002; Fan et al. 2003). Furthermore, substrate peptides adopting extended β-strand structures appear to be favored by $3CL^{pro}$ over α-helical or disordered structures (Fan et al. 2003). On the basis of these data, it is reasonable to postulate that coronavirus polyprotein processing occurs in a temporally coordinated manner, which might lead to activation and inactivation of specific functions in the course of the viral life cycle, as has been demonstrated for other +RNA viruses (Lemm et al. 1994; Vasiljeva et al. 2003).

The combined data of numerous studies published in the past 15 years provide a (nearly) complete picture of the pp1a/pp1ab processing pathways of prototypic viruses from all three coronavirus groups (Fig. 1). Throughout this chapter, the replicase processing end products will be continuously numbered from nonstructural protein (nsp) 1 to

nsp16 (from N- to C-terminus[2]) to facilitate their comparison with homologs from other coronaviruses.

3.2.1 Accessory Proteinases

The N-proximal regions of the MHV and HCoV-229E replicase polyproteins are processed by two PL^{pro}s at three sites to produce nsp1–4, with the C-terminus of nsp4 being cleaved by the main proteinase (Fig. 1). The proteolytic activities of the MHV and HCoV-229E $PL1^{pro}$ and $PL2^{pro}$ domains and the IBV $PL2^{pro}$, which all reside in nsp3, have been characterized in detail (Ziebuhr et al. 2000). Briefly, the MHV $PL1^{pro}$ cleaves the nsp1|nsp2 and nsp2|3 sites, while $PL2^{pro}$ processes the third site, nsp3|nsp4 (Baker et al. 1989, 1993; Dong and Baker 1994; Denison et al. 1995; Hughes et al. 1995; Bonilla et al. 1997; Teng et al. 1999; Kanjanahaluethai and Baker 2000; Kanjanahaluethai et al. 2003). Also in HCoV-229E, $PL1^{pro}$ was shown to cleave the nsp1|nsp2 and nsp2|nsp3 sites (Herold et al. 1998; Ziebuhr et al. 2001). However, in the case of HCoV-229E, the regulation of proteolytic processing was shown to be more complex than previously thought. Thus $PL2^{pro}$ (originally believed to process only the nsp3|nsp4 site) was demonstrated also to process the nsp2|nsp3 site. The nsp2|nsp3 cleavages mediated by $PL1^{pro}$ and $PL2^{pro}$, respectively, were shown to occur at exactly the same scissile bond (Herold et al. 1998; Ziebuhr et al. 2001). Whereas the $PL1^{pro}$-mediated cleavage proved to be slow and incomplete in vitro, $PL2^{pro}$ cleaved this site efficiently under the same experimental conditions. Furthermore, evidence was obtained to suggest that the proteolytic activity of $PL1^{pro}$ at the nsp2|nsp3 site is downregulated by $PL2^{pro}$ by a noncompetitive mechanism (Ziebuhr et al. 2001). It was concluded that the activities of the two proteinase domains present in nsp3 are tightly regulated in HCoV-229E and, probably, also other coronaviruses, with $PL2^{pro}$ playing a major role and dominating over the activity of $PL1^{pro}$. This conclusion is also supported by the conservation of $PL2^{pro}$ in all coronaviruses (Ziebuhr et al. 2001; Snijder et al. 2003).

IBV encodes only one proteolytically active PL^{pro}, which is $PL2^{pro}$. The IBV $PL1^{pro}$ domain, although being conserved, has lost its proteolytic activity in the course of evolution because of the accumulation of active site mutations (Ziebuhr et al. 2001). Apparently, IBV does not en-

[2] Note that similar designations (nsp or ns) are occasionally used for some of the group-specific *nonstructural* proteins encoded in the 3′-structural protein regions of coronaviruses (Brown and Brierley, 1995).

code a counterpart of the nsp1 protein of other coronaviruses. Thus there are only two cleavage sites in this region of pp1a/pp1ab, nsp2|nsp3 and nsp3|nsp4, which are both processed by $PL2^{pro}$ (Lim and Liu 1998; Lim et al. 2000). In SARS-CoV, only one PL^{pro} is conserved (Marra et al. 2003; Rota et al. 2003). The domain occupies a position in pp1a/pp1ab that corresponds to that of the $PL2^{pro}$ domains of other coronaviruses and therefore is considered an ortholog of coronavirus $PL2^{pro}$s (Snijder et al. 2003). Obviously, the SARS-CoV $PL2^{pro}$ must be responsible for the processing of all three sites identified in this region and, indeed, the activity of $PL2^{pro}$ at the nsp2|nsp3 site was demonstrated recently (Thiel et al. 2003). The arrangement of the N-terminal domains of SARS-CoV nsp3 differs from that of other coronaviruses (Ziebuhr et al. 2001; Snijder et al. 2003). Thus, the conserved ADRP domain ("X" in Fig. 1) resides immediately downstream of the acidic domain (Ac) in nsp3, a position that is occupied by $PL1^{pro}$ in other coronaviruses. Further downstream, another domain of unknown function has been identified in the region separating the ADRP and $PL2^{pro}$ domains. It has been termed "SARS-CoV unique domain" (SUD) (Snijder et al. 2003) (Fig. 1).

The sequence similarity between coronaviral PL^{pro}s and the prototypic cellular proteinases is very low. A closer relationship seems to exist between the active sites of coronavirus PL^{pro}s and the leader proteinase (L^{pro}) of the picornavirus foot-and-mouth-disease virus (FMDV) (Gorbalenya et al. 1991). Crystal structure analysis revealed that the active site of L^{pro} also diverged profoundly from its cellular homologs, which explains some of the unique biochemical properties of this enzyme, such as salt sensitivity and narrow pH optimum (Guarné et al. 1998, 2000). It remains to be studied whether the sequence affinity between L^{pro} and coronavirus PL^{pro}s is associated with common structural and functional features.

Only very few amino acids are absolutely conserved among coronavirus PL^{pro}s (Herold et al. 1999). Furthermore, there are only very few $PL1^{pro}$ versus $PL2^{pro}$ lineage-specific residues, which do not provide sufficient evidence for clustering the $PL1^{pro}$ and $PL2^{pro}$ domains into two separate groups. Despite this divergency at the sequence level, coronavirus PL^{pro}s share a number of common properties. Thus they all (1) process sites that are located in the N-terminal half of the replicase polyproteins, far upstream of the conserved ORF1b-encoded domains (Fig. 1), (2) cleave sites that have at least one small residue (Gly, Ala) at the scissile bond (Dong and Baker 1994; Hughes et al. 1995; Bonilla et al. 1997; Herold et al. 1998; Lim and Liu 1998; Lim et al. 2000; Ziebuhr et al. 2001; Kanjanahaluethai et al. 2003), (3) have a catalytic dyad consisting of Cys

(followed by Trp or Tyr) and a downstream His (Baker et al. 1993; Herold et al. 1998; Lim and Liu 1998), and (4) employ variants of the papainlike $\alpha+\beta$ fold (Gorbalenya et al. 1991; Herold et al. 1999). Molecular modeling suggests that the α and β domains are connected by a transcription factor-like domain that includes a zinc-binding domain (ZBD) essential for proteolytic activity (Herold et al. 1999) (Fig. 1). It seems likely that the domain also has other functions, for example, in sg mRNA transcription. This hypothesis is based on (1) the sequence similarity with cellular transcription factors (Herold et al. 1999) and (2) the fact that the related ZBD-containing EAV nsp1 papainlike proteinase has a clearly established role in arterivirus sg mRNA synthesis (Tijms et al. 2001).

The presence of two PL^{pro}s in most coronavirus replicases suggests that these enzymes originated from the duplication of a PL^{pro} domain in one of the ancestors of the contemporary coronaviruses. Surprisingly, however, phylogenetic trees inferred from multiple sequence comparisons of coronavirus PL^{pro}s revealed that only the $PL1^{pro}$ and $PL2^{pro}$ domains of the most closely related coronaviruses were clustered together (Ziebuhr et al. 2001). Therefore, multiple independent gene duplications in different coronaviruses cannot be excluded entirely. Alternatively and much more probably, the above result can be interpreted to reflect homoplasy events that, subsequent to the initial gene duplication, have driven a parallel evolution of the two coronavirus PL^{pro} paralogs, while other regions of the replicase diverged much more profoundly (Ziebuhr et al. 2001). Often, such homoplasy events are driven by common substrates. Thus the identification of a common cleavage site that is processed by both $PL1^{pro}$ and $PL2^{pro}$ in HCoV-229E may indicate that, in this virus and probably also other coronaviruses, the conservation of overlapping substrate specificities was an important driving force of evolution. The underlying selective advantage that led to the conservation of such a partial redundancy of two proteinase domains in most coronaviruses remains to be investigated. Conservation of overlapping substrate specificities also appears to affect the cleavage site structures. Thus a comparison of PL^{pro} cleavage sites of SARS-CoV and IBV, which both employ only one PL^{pro} activity, with the corresponding cleavage sites of HCoV-229E, which employs two PL^{pro} domains, revealed a much better conservation of the IBV/SARS-CoV $PL2^{pro}$ sites compared with the HCoV $PL1^{pro}$/$PL2^{pro}$ sites (Thiel et al. 2003).

3.2.2 Main Proteinase

The coronavirus main proteinase, $3CL^{pro}$, is encoded by ORF1a and resides in nsp5 (Fig. 1). In the polyprotein, it is flanked by hydrophobic domains. The ~33-kDa proteinase releases itself from pp1a/pp1ab at flanking sites and directs the proteolytic processing of all downstream domains of pp1a/pp1ab (Fig. 1). In total, $3CL^{pro}$ cleaves at 11 conserved sites to produce 13 processing end products and, probably, multiple intermediates. Because of the central role in the expression of the major replicative proteins, $3CL^{pro}$ is also called "main" proteinase (M^{pro}).

Coronavirus $3CL^{pro}$s represent a highly diverged branch of two-β-barrel proteinases (Gorbalenya et al. 1989a,c). In contrast to what the name suggests, coronavirus $3CL^{pro}$s also deviate significantly from the picornavirus 3C and other +RNA viral 3C-like proteinases. Characterization of a roniviral $3CL^{pro}$ has indicated that the 3C-like proteinases of potyviruses may represent the closest relatives of coronavirus $3CL^{pro}$s (outside the *Nidovirales* order) (Cowley et al. 2000; Gorbalenya 2001; Ziebuhr et al. 2003). In common with the prototypic picornavirus 3C proteinases (Allaire et al. 1994; Matthews et al. 1994; Mosimann et al. 1997), coronavirus 3C-like proteinases have a chymotrypsin-like, two-β-barrel fold that is formed by 12 antiparallel β-strands (Allaire et al. 1994; Matthews et al. 1994; Mosimann et al. 1997; Anand et al. 2002, 2003). However, both the size and orientation of secondary structure elements vary considerably between the two groups of enzymes, making reliable structural alignments difficult, if not impossible. Furthermore, in contrast to 3C proteinases but in common with other nidovirus 3C-like proteinases (Barrette-Ng et al. 2002; Ziebuhr et al. 2003), coronavirus $3CL^{pro}$s have a C-terminal extension, which is called domain III to distinguish it from the β-barrel domains I and II. Domain III of the TGEV $3CL^{pro}$ comprises 103 amino acids and consists of 5 α-helices that adopt a unique structure that currently has no homologs in the database (Anand et al. 2002) (Figs. 2 and 3). The structure of the coronavirus $3CL^{pro}$ domain III differs from the corresponding domain of the arterivirus nsp4 proteinase, which comprises only 49 residues and consists of 2 short pairs of β-strands and 2 α-helices (Barrette-Ng et al. 2002).

The differences between picornavirus and coronavirus chymotrypsin-like proteinases also extend to the catalytic residues. Thus, whereas the vast majority of picornavirus enzymes employ a catalytic triad, Cys-

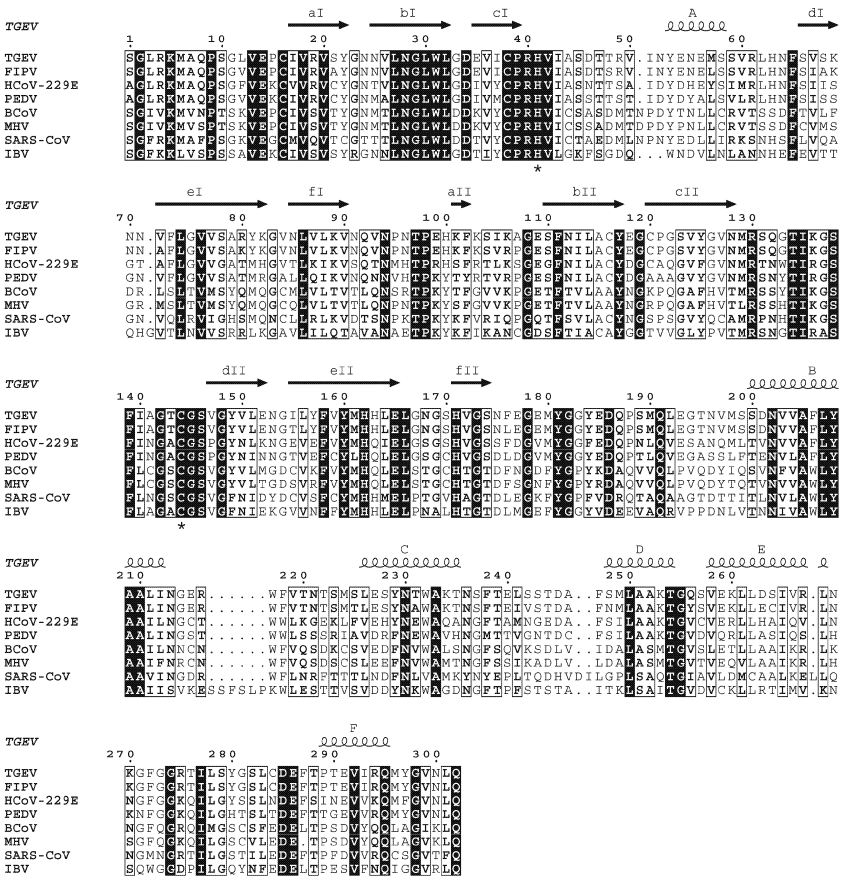

Fig. 2. Sequence comparison of coronavirus 3C-like main proteinases. The alignment was generated with the ClustalW program (version 1.82) (http://www.ebi.ac.uk/clustalw/) and used as input for the ESPript program (version 2.1) (http://prodes.toulouse.inra.fr/ESPript/cgi-bin/ESPript.cgi). The $3CL^{pro}$ sequences of transmissible gastroenteritis virus (*TGEV*, strain Purdue 46), feline infectious peritonitis virus (*FIPV*, strain 79-1146), human coronavirus 229E (*HCoV-229E*), porcine epidemic diarrhea virus (PEDV, strain CV777) bovine coronavirus (*BCoV*, isolate LUN), mouse hepatitis virus (*MHV*, strain A59), avian infectious peritonitis virus (IBV, strain Beaudette), and SARS coronavirus (*SARS-CoV*, isolate Frankfurt 1) were derived from the replicative polyproteins of the respective viruses whose sequences are deposited at the DDBJ/EMBL/GenBank database (accession numbers: TGEV, AJ271965; FIPV, AF326575; HCoV, X69721; PEDV, AF353511; BCoV, AF391542; MHV, NC 001846; IBV, M95169; SARS-CoV, AY291315). The β-strands and α-helices as revealed by the TGEV $3CL^{pro}$ crystal structure (Anand et al. 2002; PDB 1LVO) are shown *above* the sequence alignment. Catalytic Cys and His residues are indicated by *asterisks*

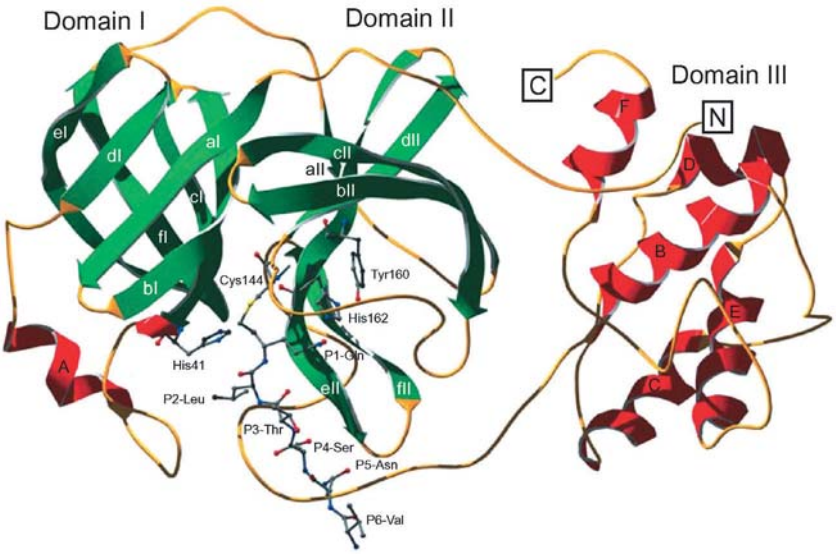

Fig. 3. Structure of monomer B of TGEV 3CL^{pro} with a hexapeptidyl chloromethyl ketone inhibitor bound to the active site (Anand et al. 2002, 2003). 3CL^{pro} domains I, II, and III are indicated. α-Helices are shown in *red* and are *labeled A* to *F*. β-Strands are shown in *green* and are *labeled a* to *f*, followed by an indication of the domain to which they belong. Shown in ball-and-stick representation are the substrate analog inhibitor (residues P1 to P6), the catalytic residues (Cys144 and His41), and the S1 subsite His162 residue interacting with Tyr160 and the P1 Gln side chain of the substrate (see text for details). N- and C termini are labeled *N* and *C*

His-Asp(Glu) (Allaire et al. 1994; Matthews et al. 1994; Mosimann et al. 1997; Seipelt et al. 1999), which is reminiscent of the charge-relay system of chymotrypsin-like serine proteinases, the coronavirus 3CL^{pro}s use a catalytic dyad consisting of Cys (nucleophile) and His (general base) (Figs. 2 and 3). Mutation analyses performed with recombinant enzymes from different coronavirus species had consistently failed to identify a third catalytic residue, suggesting that coronavirus 3CL^{pro}s may lack a counterpart to the catalytic Asp(Glu) of other chymotrypsin-like proteinases (Liu and Brown 1995; Lu and Denison 1997; Ziebuhr et al. 1997). This hypothesis was confirmed by crystal structure analyses of the TGEV (Anand et al. 2002), HCoV-229E (Anand et al. 2003), and SARS-CoV 3CL^{pro} enzymes (PDB acc: 1Q2W). Thus, for example, in the TGEV 3CL^{pro} structure, a buried water molecule was found in the place that is normally occupied by the third member of the triad (Asp or Glu).

The water was hydrogen-bonded to His41[3] $N^{\delta 1}$, His163 $N^{\delta 1}$, and Asp186 $O^{\delta 1}$. An equivalent water molecule is also found in the HCoV $3CL^{pro}$ structure. Here, it is stabilized by His41 $N^{\delta 1}$, Gln163 $N^{\delta 1}$, and Asp186 $O^{\delta 1}$. The TGEV $3CL^{pro}$ structure also suggested that, after the attack of the active-site Cys144 nucleophile on the carbonyl carbon of the scissile bond, the developing oxyanion is stabilized by hydrogen bonds donated by the main chain amides of Gly142, Thr143, and Cys144, which together form the "oxyanion hole."

The substrate specificity of coronavirus $3CL^{pro}$s resembles that of many other 3C and 3C-like proteinases (Blom et al. 1996; Ryan and Flint 1997) in so far as all the coronavirus $3CL^{pro}$ sites share a Gln residue at the P1 position, whereas small residues (Ala, Ser, and Gly) are conserved at the P1′ position (Ziebuhr et al. 2000). Larger residues, such as Asn (which is found at the P1′ position of all coronavirus nsp8|nsp9 sites), result in significantly reduced cleavage efficiencies (Ziebuhr and Siddell 1999; Hegyi and Ziebuhr 2002; Fan et al. 2003). Leu is strongly preferred at the P2 position of coronavirus $3CL^{pro}$ substrates, although other hydrophobic residues, such as Ile, Val, Phe, and Met, are occasionally also found at this position. At the P4 position, small residues, Val, Thr, Ser, Pro, and Ala, are favored. The structural basis for the pronounced specificity of coronavirus $3CL^{pro}$s was elucidated recently by structure analysis of a hexapeptidyl chloromethyl ketone inhibitor bound to the active site of the TGEV $3CL^{pro}$ (Anand et al. 2003). Because the sequence of the inhibitor was derived from the P6–P1 region of a natural cleavage site (Val-Asn-Ser-Thr-Leu-Gln) of TGEV $3CL^{pro}$, the structure most likely represents the binding mode of coronavirus $3CL^{pro}$ substrates in general. It was found that the P region of $3CL^{pro}$ substrates binds in a shallow groove at the surface of the proteinase, between domains I and II (Fig. 3). Residues P5 to P3 form an antiparallel β-sheet with residues 164–167 of strand eII and residues 189–191 of the loop linking domains II and III. Deletion of the loop region abolishes the proteolytic activity of $3CL^{pro}$, supporting the functional significance of the interaction between the substrate and this loop region (Anand et al. 2002).

The conserved Gln side chain at the P1 position of $3CL^{pro}$ substrates interacts with the imidazole of His162 (Fig. 3), at the bottom of the S1 subsite, which is formed by the main-chain atoms of Ile51, Leu164, Glu165, and His171 (Anand et al. 2003). The neutral state of His162 over a broad pH range appears to be maintained by (1) stacking onto the

[3] Amino acid residues of coronavirus $3CL^{pro}$s are numbered from Ser(Ala)1 to Gln302.

phenyl ring of Phe139 and (2) accepting a hydrogen bond from the hydroxyl group of the buried Tyr160. This interpretation is supported by mutagenesis data obtained for bacterially expressed HCoV-229E and feline infectious peritonitis (FIPV) $3CL^{pro}$s (Ziebuhr et al. 1997; Hegyi et al. 2002). Tyr160 is part of the conserved coronavirus $3CL^{pro}$ signature, Tyr-X-His, whereas Gly(Ala)-X-His is found at the equivalent sequence position in most 3C and 3C-like proteinases (Gorbalenya et al. 1989a; Gorbalenya and Snijder 1996). Accordingly, stabilization of histidine in the neutral tautomeric state needs to be ensured by other residues (Bergmann et al. 1997; Mosimann et al. 1997).

The hydrophobic S2 subsite of the proteinase, which accommodates the conserved Leu residue and, in few cases, other hydrophobic residues, is formed by the side chains of Leu164, Ile51, Thr47, His41, and Tyr53 (Anand et al. 2003). The fact that, in the structure, the P3 side chain of the substrate analog was oriented toward bulk solvent explains why there is no specificity for any particular side chain at the P3 position of coronavirus $3CL^{pro}$ cleavage sites (Ziebuhr et al. 2000). The S4 site is rather congested (Anand et al. 2003), explaining the conservation of small residues, such as Ser, Thr, Val, or Pro, at this position of coronavirus $3CL^{pro}$ substrates. On the basis of the TGEV $3CL^{pro}$–inhibitor structure, it has been proposed that the relatively small P1′ residues (Ser, Ala, or Gly) may be accommodated by a S1′ subsite that involves Leu27, His41, and Thr47 (Anand et al. 2003).

It is generally believed that most of the pp1a/pp1ab cleavages are mediated *in trans* by the fully processed form of $3CL^{pro}$ (nsp5). The *trans* activity of $3CL^{pro}$ has been well characterized, both biochemically and structurally (Ziebuhr et al. 1995; Grötzinger et al. 1996; Lu et al. 1996; Heusipp et al. 1997a,b; Tibbles et al. 1999; Ziebuhr and Siddell 1999; Anand et al. 2002, 2003; Hegyi and Ziebuhr 2002; Fan et al. 2003). However, it is not clear whether $3CL^{pro}$ cleaves itself from pp1a/pp1ab *in cis* or *in trans*. Also, it is not clear whether $3CL^{pro}$ can cleave downstream pp1a/pp1ab sites *in cis*. Thus, on the one hand, there is biochemical and structural evidence to suggest that $3CL^{pro}$ self-processing occurs *in trans* (Lu et al. 1996; Anand et al. 2002). Furthermore, in MHV-infected cells, $3CL^{pro}$ was found to be part of a rather stable 150-kDa processing intermediate (nsp4–10 or nsp4–11), which also argues against a rapid, cotranslational release of $3CL^{pro}$ *in cis* (Schiller et al. 1998). On the other hand, a number of MHV and IBV $3CL^{pro}$-containing precursors were shown to require microsomal membranes for efficient autocatalytic release of $3CL^{pro}$ from the flanking TM2 (nsp4) and TM3 (nsp6) domains (Tibbles et al. 1996; Piñon et al. 1997), indicating that the flanking do-

mains (when properly folded) affect the activity of $3CL^{pro}$. In other words, interdomain interactions in pp1ab may modulate the structure (and activity) of the enzyme, for example, to render $3CL^{pro}$ competent for *cis* cleavages at flanking sites or even further downstream sites. In fact, one might expect that at least some of the pp1a/pp1ab cleavages need to occur *in cis* early in infection, when the concentration of $3CL^{pro}$ is low and intermolecular reactions are less likely to occur. Otherwise, if there were no *cis* cleavages at all, pp1a/pp1ab should operate initially as an extremely large polyprotein that is only processed at its N-terminus by PL^{pro} cleavages. Structure information for larger $3CL^{pro}$ precursors will be required to answer the question of whether or not $3CL^{pro}$ adopts alternative conformations in its fully processed form and larger precursor molecules. Notably, reorientation of secondary structure elements after intramolecular release is believed to occur in picornavirus 3C proteinases (Khan et al. 1999), illustrating the significance of this question.

At present, structure information is only available for the fully processed coronavirus $3CL^{pro}$ (Anand et al. 2002, 2003). Both the crystal structures and dynamic light scattering data show that $3CL^{pro}$ forms dimers (Anand et al. 2002, 2003). The two molecules in the dimer are oriented perpendicular to one another (Fig. 4). The contact interface mainly involves conserved residues of the N-terminus of one molecule and domain II of the other molecule (and vice versa). The N-terminal amino acid residues are squeezed in between domains II and III of the parent

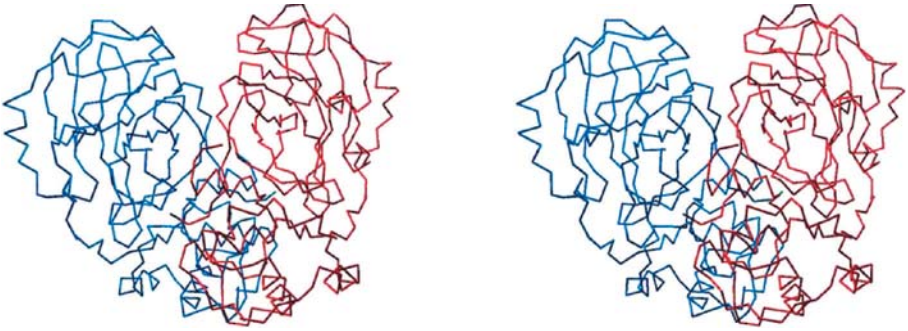

Fig. 4. Coronavirus main proteinases form dimers (Anand et al. 2002). Stereo representation of a Cα plot of a TGEV $3CL^{pro}$ dimer (PDB accession number: 1LVO). Monomers A and B are shown in *blue* and *red*, respectively. The monomers are oriented perpendicular to one another. Dimerization mainly involves interactions of the N terminus with domain II of the other dimer (see text for details). The N termini of monomers A and B are shown in *green* and *brown*, respectively

monomer and domain II of the other monomer, where they make a number of very specific interactions that appear tailor-made to bind this segment with high affinity. Apparently, this mechanism allows the active site to remain competent for binding and cleaving other sites in the polyprotein after autocleavage of $3CL^{pro}$. In addition, the exact placement of the N-terminus seems to have a structural role for the mature $3CL^{pro}$, because deletion of residues 1 to 5 leads to a dramatic decrease in proteolytic activity (Anand et al. 2003). It has been speculated that the tight interaction of the N-terminus with domains II and III may help to maintain the loop connecting domains II and III in the orientation required to bind the P3–P5 residues of the substrate (Anand et al. 2002, 2003). The presumed indirect role of domain III in proteolysis may explain the results from previous mutagenesis studies that consistently reported a dramatic loss of *trans*-cleavage activity with C-terminally truncated forms of HCoV-229E, TGEV, MHV, and IBV $3CL^{pro}$s (Lu and Denison 1997; Ziebuhr et al. 1997; Ng and Liu 2000; Anand et al. 2002).

Genetic data also point to a (direct or indirect) role of domain III in RNA synthesis. Thus characterization of temperature-sensitive (*ts*) MHV mutants revealed that substitution of the MHV $3CL^{pro}$ Phe219 residue, which is part of the loop connecting α-helices B and C in domain III (Fig. 2), with Leu causes an RNA-minus phenotype at the restrictive temperature (Siddell et al. 2001). Further characterization of the *ts* mutant, Alb *ts*16, showed that both plus- and minus-strand synthesis was not greatly affected when the temperature was shifted late in infection. However, when the temperature was shifted to the nonpermissive temperature early, at a time when the rate of MHV RNA synthesis increases rapidly, no increase of plus-strand synthesis was observed with Alb *ts*16. Furthermore, inhibition of minus-strand synthesis (by inhibition of protein synthesis) was found to cause a decline of plus-strand synthesis after 30–60 min. The data can be interpreted to indicate that the defect in $3CL^{pro}$ activity interferes with minus-strand synthesis and reduces it to a low level that merely ensures the replenishment of minus strands being lost because of turnover. Alternatively, the mutation may cause a defect in the activity of $3CL^{pro}$ that blocks the formation of plus-strand polymerase activity (or prevents its conversion from the minus strand-synthesizing precursor). It remains to be determined whether the observed *ts* phenotype is caused by specific defects in the proteolytic activity of $3CL^{pro}$ or whether another, nonproteolytic function of domain III is affected. Thus, for example, protein-protein interactions involving domain III—as proposed to be mediated by the C-terminal domain of the EAV nsp4 proteinase (Barrette-Ng et al. 2002)—may be affected.

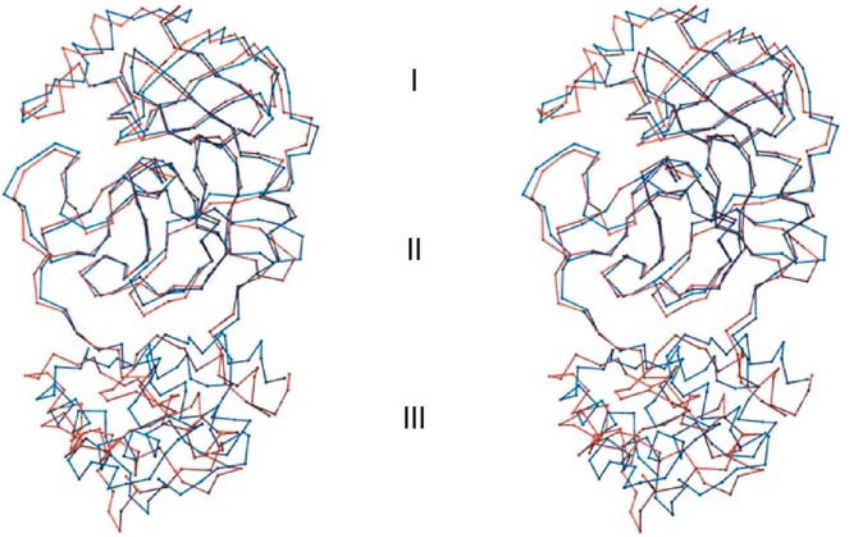

Fig. 5. Differential orientation of the C-terminal domains III of TGEV and SARS-CoV 3C-like main proteinases (PDB 1LVO and 1Q2 W). Superimposition (stereo image) of TGEV (*orange*) and SARS-CoV (*blue*) $3CL^{pro}$s shows little variation between the structures of the N-terminal β-barrel domains I and II. The orientation (rather than the structure) of the respective C-terminal domains of TGEV and SARS-CoV $3CL^{pro}$ differs slightly in the two proteins, resulting in less perfect superimposition

Comparison of coronavirus main proteinase structures shows that domains I and II superimpose much better than the C-terminal domains III (Fig. 5). This is mainly due to a slightly different orientation of domain III in relation to domains I and II rather than differences in the domain III structures themselves.

3.3 Helicase

RNA helicases represent the second most conserved subunit of the RNA synthesis machinery of +RNA viruses and are involved in diverse steps of the viral life cycle (Buck 1996; Kadaré and Haenni 1997). They utilize the energy derived from hydrolysis of nucleoside triphosphates (NTPs) to unwind double-stranded (ds) RNA. Conservation of specific sequence motifs allows helicases to be classified into three large superfamilies (SFs), termed SF1, SF2, and SF3, as well as several small families (Gorbalenya et al. 1989b; Gorbalenya and Koonin 1993). The coronavi-

rus helicase resides in nsp13 and has been classified as belonging to SF1 (Gorbalenya et al. 1989b, c) (Fig. 1). Nsp13 and its homologs in other nidoviruses have a putative zinc-binding domain (ZBD) at their N-terminus (Gorbalenya et al. 1989c), which is known to be required for the enzymatic activities of coronavirus and arterivirus helicases (Seybert, van Dinten, Posthuma, Snijder, Gorbalenya, and Ziebuhr, unpublished data). EAV reverse genetics data have shown that the ZBD and a downstream segment ("hinge spacer") that links ZBD to the C-terminal helicase domain have distinct functions in arterivirus replication, sg mRNA transcription, and virion morphogenesis (van Dinten et al. 2000). It is tempting to suggest that coronavirus helicases may have similarly diverse functions. Biochemical characterization of a recombinant form of HCoV-229E nsp13 demonstrated both nucleic acid-stimulated NTPase and duplex-unwinding activities (Seybert et al. 2000a). Similar data have subsequently been obtained for two arterivirus nsp10 helicases and the SARS-CoV nsp13 helicase (Seybert et al. 2000b; Bautista et al. 2002; Tanner et al. 2003; Thiel et al. 2003).

Coronavirus (and arterivirus) helicases were shown to unwind their dsRNA substrates with 5′-to-3′ polarity, that is, they move in a 5′-to-3′ direction on the strand to which they initially bind (Seybert et al. 2000a, b). Obviously, this stands in contrast to the 3′-to-5′ polarity of the SF2 helicases of flavi-, pesti-, and hepaciviruses (Kadaré and Haenni 1997; Kwong et al. 2000) and may indicate fundamental differences in biological functions between the two groups of enzymes. For example, the 5′-to-3′ polarity of the coronavirus nsp13 helicase activity argues against a role in the separation of secondary structures in the RNA template during minus-strand synthesis (as has been suggested for RNA viral SF2 helicases), because this would require a helicase with 3′-to-5′ polarity.

Interestingly, coronavirus nsp13 is one of the few helicases that have no marked preference for RNA or DNA substrates. Thus they have been found to unwind partial-duplex DNA substrates with high efficacy (Seybert et al. 2000; Thiel et al. 2003). This property allows DNA-based assays to be used in the characterization of coronavirus helicases (for example, in mutagenesis studies and high-throughput tests of potential inhibitors). Because coronaviruses replicate in the cytoplasm and the helicase has not been found to localize to the nucleus (Sims et al. 2000; Bost et al. 2001), a biological significance of the DNA-unwinding activity of nsp13 seems unlikely, although it cannot be excluded entirely at the present stage. It should be mentioned in this context that the hepatitis C virus (HCV) NS3 helicase also has DNA duplex-unwinding activity,

which, however, has been proposed to affect the structure of host cell DNA (Pang et al. 2002).

Duplex unwinding by coronavirus helicases is an energy-dependent process that derives its energy from NTP hydrolysis (Seybert et al. 2000a; Seybert and Ziebuhr 2001). Coronavirus helicases appear to be highly promiscuous with respect to the NTP cofactor used. Thus all standard NTPs and dNTPs were found to be hydrolyzed by coronavirus helicases (Seybert et al. 2000a; Seybert and Ziebuhr 2001; Tanner et al. 2003). Finally, coronavirus helicases possess RNA 5′-triphosphatase activity that may be involved in the formation of the 5′ RNA cap structure of coronavirus plus-strand RNAs (Ivanov et al. 2004; Ivanov and Ziebuhr 2004).

3.4 RNA-Dependent RNA Polymerase

As discussed above for other coronavirus pp1a/pp1ab proteins, the RdRp domain also differs substantially from its homologs in other +RNA viruses. Coronavirus RdRps and their nidovirus relatives have been classified as an outgroup of SF1 RdRps (Koonin 1991). The coronavirus RdRp domain comprising the finger, palm, and thumb subdomains occupies the C-terminal two-thirds of nsp12 (Gorbalenya et al. 1989c). Recent data suggest that replication complex association of the RdRp may occur through interactions of the nsp12 segment 411–448 (located upstream of the RdRp core domain in nsp12) with ORF1a-encoded proteins, such as nsp5 ($3CL^{pro}$), nsp8, and nsp9 (Brockway et al. 2003). Consistent with the presumed RdRp activity of nsp12, a mutation in nsp12 (His868 to Arg) was found to cause an RNA-negative phenotype in an MHV *ts* mutant, Alb *ts*22 (Siddell et al. 2001). Thus, when infected cultures of Alb *ts*22 were shifted to the restrictive temperature at 40°C, both plus- and minus-strand RNA synthesis ceased immediately. Even at the permissive temperature, the *ts* mutant synthesized 4–5 times less RNA compared with revertants. The defect of this mutant in RNA synthesis can easily be explained by the fact that His868 is part of the predicted thumb subdomain of the MHV RdRp that, in other RNA polymerases, has been implicated in polymerase activity (Burns et al. 1989; Mills et al. 1989; Plotch et al. 1989; Hansen et al. 1997).

The Cys/His-rich nsp10 that immediately precedes RdRp in pp1ab (Fig. 1) has also been implicated in RNA synthesis. An MHV *ts* mutant, Alb *ts*6, encoding a mutant form of nsp10 (Gln65 to Glu), was shown to have a defect in minus-strand RNA synthesis (Siddell et al. 2001). Thus,

when the temperature was shifted to 40°C, minus-strand synthesis stopped immediately but plus-strand synthesis continued at the same level as was occurring at the time of temperature shift. Plus-strand RNA synthesis gradually declined over 3–4 h (starting at 30–60 min after the shift to 40°C) because the minus strands produced at the permissive temperature were turned over (Wang and Sawicki 2001) and, because of the defect in their synthesis, were not replenished at the restrictive temperature.

Nsp10 and nsp12 (RdRp) are adjacent domains in pp1ab (Fig. 1). Peptide cleavage data have shown that, most likely because of a replacement of the conserved P2 Leu residue, the nsp10|nsp12 cleavage site is less efficiently cleaved than other SARS-CoV $3CL^{pro}$ sites (Fan et al. 2003). Also, the nsp10|nsp12 sites of other coronaviruses have the P2 position occupied by noncanonical residues. It is thus tempting to speculate that the nsp10|nsp12 site has to be cleaved more slowly than other sites, probably to attain a specific activity mediated by an nsp10–nsp12-containing intermediate. The IBV nsp10 has been reported to form dimers. It localizes to membranes near the site of viral RNA synthesis (Ng and Liu 2002).

4 Subcellular Localization of the Coronavirus Replicase

Genome replication and transcription of virtually all +RNA viruses takes place at intracellular membranes that are derived from various cellular organelles including, for example, the endoplasmic reticulum, lysosomes and endosomes, intermediate compartment and *trans*-Golgi network, peroxisomes, mitochondria, and chloroplasts (Russo et al. 1983; Froshauer et al. 1988; Peränen and Kääriäinen 1991; De Graaff et al. 1993; Peränen et al. 1995; Restrepo-Hartwig and Ahlquist 1996; Schaad et al. 1997; van der Meer et al. 1998; Mackenzie et al. 1999; Restrepo-Hartwig and Ahlquist 1999; Miller et al. 2001). The viral replication complex, which consists of multiple viral but also cellular subunits (see the chapter by Shi and Lai, this volume), is associated with these membranes and, in many cases, also directs their synthesis and/or modification (Peränen and Kääriäinen 1991; Cho et al. 1994; Schlegel et al. 1996; Teterina et al. 1997; Snijder et al. 2001; Egger et al. 2002). Typically, multiple vesicles or membrane invaginations (spherules) on cellular organelles are induced to which the replication complex is attached by specific structural elements, such as hydrophobic domains (van Kuppeveld

et al. 1995; Snijder et al. 2001) amphipathic helices (Datta and Dasgupta 1994), palmitate side chains (Laakkonen et al. 1996), and C-terminal membrane insertion sequences (Schmidt-Mende et al. 2001). As a result, replication takes place in a membrane-protected (and, thus, nuclease resistant) microenvironment that contains (and sequesters) the protein functions required for viral RNA synthesis. This strategy is believed to improve template specificity by retaining negative strands for template use and to repress host defenses that may be induced by double-stranded RNA (Schwartz et al. 2002).

Association of the viral replication/transcription complex with intracellular membranes has also been established for coronaviruses (Sethna and Brian 1997). Thus TGEV genome- and subgenome-length minus strands, which are the templates for viral genome RNA replication and subgenomic mRNA transcription, respectively (Sethna et al. 1989; Sawicki and Sawicki 1990; Schaad and Baric 1994; Sawicki et al. 2001), were predominantly found in nuclease-resistant membranous complexes. In contrast, positive-strand RNAs proved to be much more susceptible to nuclease digestion, indicating that plus-strand RNAs, which also act as mRNAs, are mainly in solution or part of easily dissociable complexes in the cytosol (Sethna and Brian 1997).

Immunofluorescence (IF) studies provided clear evidence that the vast majority of coronavirus replicase subunits localize to perinuclear membrane compartments (Heusipp et al. 1997a; Bi et al. 1998; Denison et al. 1999; Shi et al. 1999; van der Meer et al. 1999; Ziebuhr and Siddell 1999; Bost et al. 2000; Sims et al. 2000; Bost et al. 2001; Xu et al. 2001; Ng and Liu 2002). Whereas most ORF1a-encoded replicase components remain tightly associated with membranes throughout the viral life cycle, at least some of the ORF1b-encoded subunits seem to be only temporarily present in the complex, probably when still part of the polyprotein. Thus, for example, partial detachment from the membrane-bound complexes was reported for MHV nsp12 and nsp13 later in infection (van der Meer et al. 1999; Bost et al. 2001; Xu et al. 2001). Also, the most C-terminal IBV pp1ab processing products show, in contrast to all other IBV pp1a/pp1ab proteins tested, a diffuse, cytoplasmic staining pattern in IF experiments (van der Meer et al. 1999; Bost et al. 2001; Xu et al. 2001). The membrane-bound replicase proteins overlap to a large extent with the site of viral RNA synthesis (Denison et al. 1999; Shi et al. 1999; van der Meer et al. 1999; Bost et al. 2001; Gosert et al. 2002; Ng and Liu 2002). There is some controversy regarding the intracellular compartment at which viral RNA synthesis takes place and, in particular, the cellular origin of the membranes employed. In a recent EM study (Gosert

et al. 2002), virus-induced double membrane vesicles (DMVs) were reported to be the site of MHV-A59 replication and transcription in HeLa-MHVR (Gallagher 1996) and 17CL-1 cells. These DMVs have a diameter of 200–350 nm and consist of a double membrane that, occasionally, is fused into a trilayer. At the time of maximum RNA synthesis, both genome- and subgenome-length positive-strand RNA was detected on DMVs by in situ hybridization, and also the results of BrUTP labeling suggest that DMVs are the site of viral RNA synthesis. The subcellular origin of the DMVs has not been determined to date. However, a previous IF study (Shi et al. 1999) using MHV-A59-infected 17CL-1 and HeLa-MHVR cells suggested that N-terminal pp1a/pp1ab proteins and newly synthesized RNA colocalize with ER- or Golgi-derived membranes, depending on the cell type studied.

In clear contrast to these results, another study revealed that, in MHV-A59-infected L cells at 5 h p.i., the C-terminal pp1a region (CT1a), $3CL^{pro}$ (nsp5), RdRp (nsp12), helicase (nsp13), and the N protein are associated with virus-induced, late endosomal/lysosomal membranes, which were confirmed to be the site of RNA synthesis (van der Meer et al. 1999). In IF experiments, the sites of maximum CT1a accumulation overlapped only partially with those of nsp5, nsp12, and nsp13. A thorough EM study suggested that the low (albeit significant) degree of colocalization of CT1a and nsp12 is probably due to the existence of two distinct types of membrane structures that are closely adjacent to each other but have different morphologies and protein compositions. Thus CT1a was found to be associated mainly with endosomes, whereas the majority of nsp12 was associated with multilayered membranes, probably originating from invaginations on continuous membrane sheets. The latter structures were morphologically reminiscent of endocytic carrier vesicles (ECVs) or multivesicular bodies (MVBs). However, the fact that many of these structures had membrane continuities to late endosomes argues against typical ECVs and rather favors the idea that both the multivesicular (carrying the bulk of CT1a) and multilayered (carrying the bulk of nsp12) structures represent different subdomains of the same endocytic compartment. Most intriguingly, it has also been found (van der Meer et al. 1999) that CT1a- and nsp12-positive membranes appear to be secreted. Similar observations have also been reported recently for endosome-derived cytoplasmic vacuoles carrying the alphavirus replication complex (Kujala et al. 2001). The functional significance of this phenomenon is currently unclear but may have parallels in the regulated lysosomal secretion systems employed by, for example, lymphocytes (Stinchcombe and Griffiths 1999).

The existence of two closely associated but physically distinct membrane compartments was also shown by iodixanol gradient centrifugation of intracellular membranes isolated from MHV-A59-infected DBT cells (Sims et al. 2000). The ORF1a-encoded proteins nsp2 (p65) and nsp8 (p22) cofractionated with membranes with a buoyant density of 1.05–1.09 g/ml. In contrast, nsp13, the N protein, nsp1 (p28), and newly synthesized RNA were detected in another membrane fraction of 1.12–1.13 g/ml. Both membrane fractions were LAMP-1 positive, confirming previous conclusions on the endosomal/lysosomal origin of the MHV replication compartment. Interestingly, later in infection, there appears to be a translocation of nsp13 and the N protein to the ER/*cis*-Golgi compartment, resulting in colocalization of these two proteins with the M protein at the site of virion assembly (Bost et al. 2001). The combined data suggest a multipartite structure of the coronavirus replication complex, with the N protein playing a specific role in RNA synthesis as suggested earlier (Compton et al. 1987; Baric et al. 1988). Apparently, the coronavirus replication complex undergoes structural rearrangements at the transition from maximum RNA synthesis to virion assembly at later time points (8–12 h p.i.). If this is confirmed, the localization of nsp13 at the site of assembly may correspond with a specific role of nsp13 in virion biogenesis. Such an activity has also been proposed for the related arterivirus nsp10 helicase (van Dinten et al. 1999, 2000; Seybert et al. 2000b).

To date, the mechanisms by which components of the coronavirus replication complex are integrated in or attached to intracellular membranes have not been elucidated in detail. However, it seems very likely that the strongly hydrophobic domains, TM1 to TM3 (see Fig. 1), that are present in nsp3, nsp4, and nsp6 (Gorbalenya et al. 1989c; Ziebuhr et al. 2001) play a major role in this process. This hypothesis is supported by arterivirus data showing that homologous hydrophobic domains present in EAV nsp2 and nsp3 are necessary and sufficient to trigger the synthesis of the membrane structures carrying the arterivirus replication complex (Pedersen et al. 1999; Snijder et al. 2001). The fact that several MHV pp1a/pp1ab processing products including nsp3 (Gosert et al. 2002) and nsp4–10(11) (Schiller et al. 1998), which contain TM1 and TM2/TM3, respectively, are integral membrane proteins strongly suggests a scaffold function for these proteins. There is also biochemical evidence indicating that the majority of ORF1a-encoded proteins and, to a lesser extent, ORF1b-encoded proteins are tightly bound in the complex (Gosert et al. 2002). The precise protein-protein and protein-RNA interactions stabilizing this complex remain to be characterized.

5 Concluding Remarks

Although much has been learned about coronavirus replicase organization, localization, proteolytic processing, and some of the viral replicative enzymes (e.g., proteinases and helicases), there are still major gaps in our knowledge. Given the availability of full-length clones of coronaviruses, directed genetic analysis is now possible (Almazán et al. 2000; Yount et al. 2000; Casais et al. 2001; Thiel et al. 2001a; Yount et al. 2002, 2003). In vivo studies as well as biochemical and structural information should yield important new information on the molecular details of coronaviral RNA synthesis. In this context, it will be of particular interest to define the proteins that are responsible for the unique features of coronavirus RNA synthesis, for example, the production of an extensive set of 5′- and 3′-coterminal subgenomic RNAs and the synthesis and maintenance of RNA genomes of this unique size. Studies on coronavirus replicases and their homologs on closely related viruses may also help to determine the structural and functional constraints that have driven the evolution of nidoviruses and enable them to infect a broad range of vertebrate and invertebrate hosts. Furthermore, the relationship of the recently identified coronavirus RNA processing activities with cellular proteins may reveal interesting insights into similarities and differences (or even an interplay) between coronaviral and cellular RNA metabolism pathways. In the long term, the unique structural properties of coronavirus replicative enzymes may allow the development of very selective enzyme inhibitors and possibly even drugs suitable to combat coronavirus infections.

References

Allaire M, Chernaia MM, Malcolm BA, James MN (1994) Picornaviral 3C cysteine proteinases have a fold similar to chymotrypsin-like serine proteinases. Nature 369:72–76

Almazán F, González JM, Pénzes Z, Izeta A, Calvo E, Plana-Durán J, Enjuanes L (2000) Engineering the largest RNA virus genome as an infectious bacterial artificial chromosome. Proc Natl Acad Sci USA 97:5516–5521

Anand K, Palm GJ, Mesters JR, Siddell SG, Ziebuhr J, Hilgenfeld R (2002) Structure of coronavirus main proteinase reveals combination of a chymotrypsin fold with an extra alpha-helical domain. EMBO J 21:3213–3224

Anand K, Ziebuhr J, Wadhwani P, Mesters JR, Hilgenfeld R (2003) Coronavirus main proteinase ($3CL^{pro}$) structure: basis for design of anti-SARS drugs. Science 300:1763–1767

Baker SC, Shieh CK, Soe LH, Chang MF, Vannier DM, Lai MM (1989) Identification of a domain required for autoproteolytic cleavage of murine coronavirus gene A polyprotein. J Virol 63:3693–3699

Baker SC, Yokomori K, Dong S, Carlisle R, Gorbalenya AE, Koonin EV, Lai MM (1993) Identification of the catalytic sites of a papain-like cysteine proteinase of murine coronavirus. J Virol 67:6056–6063

Baric RS, Nelson GW, Fleming JO, Deans RJ, Keck JG, Casteel N, Stohlman SA (1988) Interactions between coronavirus nucleocapsid protein and viral RNAs: implications for viral transcription. J Virol 62:4280–4287

Barrette-Ng IH, Ng KK, Mark BL, Van Aken D, Cherney MM, Garen C, Kolodenko Y, Gorbalenya AE, Snijder EJ, James MN (2002) Structure of arterivirus nsp4. The smallest chymotrypsin-like proteinase with an alpha/beta C-terminal extension and alternate conformations of the oxyanion hole. J Biol Chem 277:39960–39966

Bautista EM, Faaberg KS, Mickelson D, McGruder ED (2002) Functional properties of the predicted helicase of porcine reproductive and respiratory syndrome virus. Virology 298:258–270

Bergmann EM, Mosimann SC, Chernaia MM, Malcolm BA, James MN (1997) The refined crystal structure of the 3C gene product from hepatitis A virus: specific proteinase activity and RNA recognition. J Virol 71:2436–2448

Bi W, Piñon JD, Hughes S, Bonilla PJ, Holmes KV, Weiss SR, Leibowitz JL (1998) Localization of mouse hepatitis virus open reading frame 1A derived proteins. J Neurovirol 4:594–605

Blom N, Hansen J, Blaas D, Brunak S (1996) Cleavage site analysis in picornaviral polyproteins: discovering cellular targets by neural networks. Protein Sci 5:2203–2216

Bonilla PJ, Gorbalenya AE, Weiss SR (1994) Mouse hepatitis virus strain A59 RNA polymerase gene ORF 1a: heterogeneity among MHV strains. Virology 198:736–740

Bonilla PJ, Hughes SA, Weiss SR (1997) Characterization of a second cleavage site and demonstration of activity in trans by the papain-like proteinase of the murine coronavirus mouse hepatitis virus strain A59. J Virol 71:900–909

Bost AG, Carnahan RH, Lu XT, Denison MR (2000) Four proteins processed from the replicase gene polyprotein of mouse hepatitis virus colocalize in the cell periphery and adjacent to sites of virion assembly. J Virol 74:3379–3387

Bost AG, Prentice E, Denison MR (2001) Mouse hepatitis virus replicase protein complexes are translocated to sites of M protein accumulation in the ERGIC at late times of infection. Virology 285:21–29

Boursnell ME, Brown TD, Foulds IJ, Green PF, Tomley FM, Binns MM (1987) Completion of the sequence of the genome of the coronavirus avian infectious bronchitis virus. J Gen Virol 68:57–77

Bredenbeek PJ, Pachuk CJ, Noten AF, Charite J, Luytjes W, Weiss SR, Spaan WJ (1990) The primary structure and expression of the second open reading frame of the polymerase gene of the coronavirus MHV-A59; a highly conserved polymerase is expressed by an efficient ribosomal frameshifting mechanism. Nucleic Acids Res 18:1825–1832

Brierley I, Boursnell ME, Binns MM, Bilimoria B, Blok VC, Brown TD, Inglis SC (1987) An efficient ribosomal frame-shifting signal in the polymerase-encoding region of the coronavirus IBV. EMBO J 6:3779–3785

Brierley I, Digard P, Inglis SC (1989) Characterization of an efficient coronavirus ribosomal frameshifting signal: requirement for an RNA pseudoknot. Cell 57:537–547

Brockway SM, Clay CT, Lu XT, Denison MR (2003) Characterization of the expression, intracellular localization, and replication complex association of the putative mouse hepatitis virus RNA-dependent RNA polymerase. J Virol 77:10515–10527

Buck KW (1996) Comparison of the replication of positive-stranded RNA viruses of plants and animals. Adv Virus Res 47:159–251

Bügl H, Fauman EB, Staker BL, Zheng F, Kushner SR, Saper MA, Bardwell JC, Jakob U (2000) RNA methylation under heat shock control. Mol Cell 6:349–360

Burns CC, Lawson MA, Semler BL, Ehrenfeld E (1989) Effects of mutations in poliovirus 3Dpol on RNA polymerase activity and on polyprotein cleavage. J Virol 63:4866–4874

Casais R, Thiel V, Siddell SG, Cavanagh D, Britton P (2001) Reverse genetics system for the avian coronavirus infectious bronchitis virus. J Virol 75:12359–12369

Cavanagh D (1997) Nidovirales: a new order comprising Coronaviridae and Arteriviridae. Arch Virol 142:629–633

Cho MW, Teterina N, Egger D, Bienz K, Ehrenfeld E (1994) Membrane rearrangement and vesicle induction by recombinant poliovirus 2C and 2BC in human cells. Virology 202:129–145

Chouljenko VN, Lin XQ, Storz J, Kousoulas KG, Gorbalenya AE (2001) Comparison of genomic and predicted amino acid sequences of respiratory and enteric bovine coronaviruses isolated from the same animal with fatal shipping pneumonia. J Gen Virol 82:2927–2933

Compton SR, Rogers DB, Holmes KV, Fertsch D, Remenick J, McGowan JJ (1987) In vitro replication of mouse hepatitis virus strain A59. J Virol 61:1814–1820

Cowley JA, Dimmock CM, Spann KM, Walker PJ (2000) Gill-associated virus of *Penaeus monodon* prawns: an invertebrate virus with ORF1a and ORF1b genes related to arteri- and coronaviruses. J Gen Virol 81:1473–1484

Culver GM, Consaul SA, Tycowski KT, Filipowicz W, Phizicky EM (1994) tRNA splicing in yeast and wheat germ. A cyclic phosphodiesterase implicated in the metabolism of ADP-ribose 1″,2″-cyclic phosphate. J Biol Chem 269:24928–24934

Datta U, Dasgupta A (1994) Expression and subcellular localization of poliovirus VPg-precursor protein 3AB in eukaryotic cells: evidence for glycosylation in vitro. J Virol 68:4468–4477

De Graaff M, Coscoy L, Jaspars EM (1993) Localization and biochemical characterization of alfalfa mosaic virus replication complexes. Virology 194:878–881

de Vries AAF, Horzinek MC, Rottier PJM, de Groot RJ (1997) The genome organization of the Nidovirales: similarities and differences between arteri-, toro-, and coronaviruses. Sem Virol 8:33–47

den Boon JA, Snijder EJ, Chirnside ED, de Vries AA, Horzinek MC, Spaan WJ (1991) Equine arteritis virus is not a togavirus but belongs to the coronaviruslike superfamily. J Virol 65:2910–2920

Denison MR, Hughes SA, Weiss SR (1995) Identification and characterization of a 65-kDa protein processed from the gene 1 polyprotein of the murine coronavirus MHV-A59. Virology 207:316–320

Denison MR, Spaan WJ, van der Meer Y, Gibson CA, Sims AC, Prentice E, Lu XT (1999) The putative helicase of the coronavirus mouse hepatitis virus is processed from the replicase gene polyprotein and localizes in complexes that are active in viral RNA synthesis. J Virol 73:6862–6871

Dong S, Baker SC (1994) Determinants of the p28 cleavage site recognized by the first papain-like cysteine proteinase of murine coronavirus. Virology 204:541–549

Dougherty WG, Semler BL (1993) Expression of virus-encoded proteinases: functional and structural similarities with cellular enzymes. Microbiol Rev 57:781–822

Egger D, Wölk B, Gosert R, Bianchi L, Blum HE, Moradpour D, Bienz K (2002) Expression of hepatitis C virus proteins induces distinct membrane alterations including a candidate viral replication complex. J Virol 76:5974–5984

Eleouet JF, Rasschaert D, Lambert P, Levy L, Vende P, Laude H (1995) Complete sequence (20 kilobases) of the polyprotein-encoding gene 1 of transmissible gastroenteritis virus. Virology 206:817–822

Fan K, Wei P, Feng Q, Chen S, Huang C, Ma L, Lai B, Pei J, Liu Y, Chen J, Lai L (2003) Biosynthesis, purification, and substrate specificity of severe acute respiratory syndrome coronavirus 3C-like proteinase. J Biol Chem 279:1637–1642

Filipowicz W, Pogacic V (2002) Biogenesis of small nucleolar ribonucleoproteins. Curr Opin Cell Biol 14:319–327

Froshauer S, Kartenbeck J, Helenius A (1988) Alphavirus RNA replicase is located on the cytoplasmic surface of endosomes and lysosomes. J Cell Biol 107:2075–2086

Gallagher TM (1996) Murine coronavirus membrane fusion is blocked by modification of thiols buried within the spike protein. J Virol 70:4683–4690

Goldbach R (1987) Genome similarities between plant and animal RNA viruses. Microbiol Sci 4:197–202

Gorbalenya AE, Donchenko AP, Blinov VM, Koonin EV (1989a) Cysteine proteases of positive strand RNA viruses and chymotrypsin-like serine proteases. A distinct protein superfamily with a common structural fold. FEBS Lett 243:103–114

Gorbalenya AE, Koonin EV, Donchenko AP, Blinov VM (1989b) Two related superfamilies of putative helicases involved in replication, recombination, repair and expression of DNA and RNA genomes. Nucleic Acids Res 17:4713–4730

Gorbalenya AE, Koonin EV, Donchenko AP, Blinov VM (1989c) Coronavirus genome: prediction of putative functional domains in the non-structural polyprotein by comparative amino acid sequence analysis. Nucleic Acids Res 17:4847–4861

Gorbalenya AE, Koonin EV, Lai MM (1991) Putative papain-related thiol proteases of positive-strand RNA viruses. Identification of rubi- and aphthovirus proteases and delineation of a novel conserved domain associated with proteases of rubi-, alpha- and coronaviruses. FEBS Lett 288:201–205

Gorbalenya AE, Koonin EV (1993) Helicases: amino acid sequence comparisons and structure-function relationships. Curr Opin Struct Biol 3:419–429

Gorbalenya AE, Snijder EJ (1996) Viral cysteine proteinases. Persp Drug Discov Des 6:64–86

Gorbalenya AE (2001) Big nidovirus genome. When count and order of domains matter. Adv Exp Med Biol 494:1-17

Gosert R, Kanjanahaluethai A, Egger D, Bienz K, Baker SC (2002) RNA replication of mouse hepatitis virus takes place at double-membrane vesicles. J Virol 76:3697–3708

Grötzinger C, Heusipp G, Ziebuhr J, Harms U, Süss J, Siddell SG (1996) Characterization of a 105-kDa polypeptide encoded in gene 1 of the human coronavirus HCV 229E. Virology 222:227–235

Guarné A, Tormo J, Kirchweger R, Pfistermueller D, Fita I, Skern T (1998) Structure of the foot-and-mouth disease virus leader protease: a papain-like fold adapted for self-processing and eIF4G recognition. EMBO J 17:7469–7479

Guarné A, Hampoelz B, Glaser W, Carpena X, Tormo J, Fita I, Skern T (2000) Structural and biochemical features distinguish the foot-and-mouth disease virus leader proteinase from other papain-like enzymes. J Mol Biol 302:1227–240

Hansen JL, Long AM, Schultz SC (1997) Structure of the RNA-dependent RNA polymerase of poliovirus. Structure 5:1109–1122

Hegyi A, Friebe A, Gorbalenya AE, Ziebuhr J (2002) Mutational analysis of the active centre of coronavirus 3C-like proteases. J Gen Virol 83:581–593

Hegyi A, Ziebuhr J (2002) Conservation of substrate specificities among coronavirus main proteases. J Gen Virol 83:595–599

Herold J, Raabe T, Schelle-Prinz B, Siddell SG (1993) Nucleotide sequence of the human coronavirus 229E RNA polymerase locus. Virology 195:680–691

Herold J, Siddell SG (1993) An 'elaborated' pseudoknot is required for high frequency frameshifting during translation of HCV 229E polymerase mRNA. Nucleic Acids Res 21:5838–5842

Herold J, Gorbalenya AE, Thiel V, Schelle B, Siddell SG (1998) Proteolytic processing at the amino terminus of human coronavirus 229E gene 1-encoded polyproteins: identification of a papain-like proteinase and its substrate. J Virol 72:910–918

Herold J, Siddell SG, Gorbalenya AE (1999) A human RNA viral cysteine proteinase that depends upon a unique Zn^{2+}-binding finger connecting the two domains of a papain-like fold. J Biol Chem 274:14918–14925

Heusipp G, Grötzinger C, Herold J, Siddell SG, Ziebuhr J (1997a) Identification and subcellular localization of a 41 kDa, polyprotein 1ab processing product in human coronavirus 229E-infected cells. J Gen Virol 78:2789–2794

Heusipp G, Harms U, Siddell SG, Ziebuhr J (1997b) Identification of an ATPase activity associated with a 71-kilodalton polypeptide encoded in gene 1 of the human coronavirus 229E. J Virol 71:5631–5634

Hughes SA, Bonilla PJ, Weiss SR (1995) Identification of the murine coronavirus p28 cleavage site. J Virol 69:809–813

Ivanov KA, Thiel V, Dobbe JC, van der Meer Y, Snijder EJ, Ziebuhr J (2004) Multiple enzymatic activities associated with severe acute respiratory syndrome coronavirus helicase. J Virol 78:5619–5632

Ivanov KA, Ziebuhr J (2004) Human coronavirus nonstructural protein 13: characterization of duplex-unwinding, (deoxy)nucleoside triphosphatase, and RNA 5′-triphosphatase activities. J Virol 78:7833–7838

Kadaré G, Haenni AL (1997) Virus-encoded RNA helicases. J Virol 71:2583–2590
Kanjanahaluethai A, Baker SC (2000) Identification of mouse hepatitis virus papain-like proteinase 2 activity. J Virol 74:7911–7921
Kanjanahaluethai A, Jukneliene D, Baker SC (2003) Identification of the murine coronavirus MP1 cleavage site recognized by papain-like proteinase 2. J Virol 77:7376–7382
Khan AR, Khazanovich-Bernstein N, Bergmann EM, James MN (1999) Structural aspects of activation pathways of aspartic protease zymogens and viral 3C protease precursors. Proc Natl Acad Sci USA 96:10968–10975
Kim JC, Spence RA, Currier PF, Lu X, Denison MR (1995) Coronavirus protein processing and RNA synthesis is inhibited by the cysteine proteinase inhibitor E64d. Virology 208:1-8
Kiss T (2001) Small nucleolar RNA-guided post-transcriptional modification of cellular RNAs. EMBO J 20:3617–3622
Kocherhans R, Bridgen A, Ackermann M, Tobler K (2001) Completion of the porcine epidemic diarrhoea coronavirus (PEDV) genome sequence. Virus Genes 23:137–144
Koonin EV (1991) The phylogeny of RNA-dependent RNA polymerases of positive-strand RNA viruses. J Gen Virol 72:2197–2206
Koonin EV, Dolja VV (1993) Evolution and taxonomy of positive-strand RNA viruses: implications of comparative analysis of amino acid sequences. Crit Rev Biochem Mol Biol 28:375–430
Kräusslich HG, Wimmer E (1988) Viral proteinases. Annu Rev Biochem 57:701–754
Kujala P, Ikäheimonen A, Ehsani N, Vihinen H, Auvinen P, Kääriäinen L (2001) Biogenesis of the Semliki Forest virus RNA replication complex. J Virol 75:3873–3884
Kwong AD, Kim JL, Lin C (2000) Structure and function of hepatitis C virus NS3 helicase. Curr Top Microbiol Immunol 242:171–196
Laakkonen P, Ahola T, Kääriäinen L (1996) The effects of palmitoylation on membrane association of Semliki forest virus RNA capping enzyme. J Biol Chem 271:28567–28571
Lai MM, Patton CD, Baric RS, Stohlman SA (1983) Presence of leader sequences in the mRNA of mouse hepatitis virus. J Virol 46:1027–1033
Lai MM, Cavanagh D (1997) The molecular biology of coronaviruses. Adv Virus Res 48:1-10
Laneve P, Altieri F, Fiori ME, Scaloni A, Bozzoni I, Caffarelli E (2003) Purification, cloning, and characterization of XendoU, a novel endoribonuclease involved in processing of intron-encoded small nucleolar RNAs in *Xenopus laevis*. J Biol Chem 278:13026–13032
Lee HJ, Shieh CK, Gorbalenya AE, Koonin EV, La Monica N, Tuler J, Bagdzhadzhyan A, Lai MM (1991) The complete sequence (22 kilobases) of murine coronavirus gene 1 encoding the putative proteases and RNA polymerase. Virology 180:567–582
Lemm JA, Rümenapf T, Strauss EG, Strauss JH, Rice CM (1994) Polypeptide requirements for assembly of functional Sindbis virus replication complexes: a model for the temporal regulation of minus- and plus-strand RNA synthesis. EMBO J 13:2925–2934

Lim KP, Liu DX (1998) Characterization of the two overlapping papain-like proteinase domains encoded in gene 1 of the coronavirus infectious bronchitis virus and determination of the C-terminal cleavage site of an 87-kDa protein. Virology 245:303–312

Lim KP, Ng LF, Liu DX (2000) Identification of a novel cleavage activity of the first papain-like proteinase domain encoded by open reading frame 1a of the coronavirus avian infectious bronchitis virus and characterization of the cleavage products. J Virol 74:1674–1685

Liu C, Xu HY, Liu DX (2001) Induction of caspase-dependent apoptosis in cultured cells by the avian coronavirus infectious bronchitis virus. J Virol 75:6402–6409

Liu DX, Brown TD (1995) Characterisation and mutational analysis of an ORF 1a-encoding proteinase domain responsible for proteolytic processing of the infectious bronchitis virus 1a/1b polyprotein. Virology 209:420–427

Lu X, Lu Y, Denison MR (1996) Intracellular and in vitro-translated 27-kDa proteins contain the 3C-like proteinase activity of the coronavirus MHV-A59. Virology 222:375–382

Lu Y, Denison MR (1997) Determinants of mouse hepatitis virus 3C-like proteinase activity. Virology 230:335–342

Mackenzie JM, Jones MK, Westaway EG (1999) Markers for trans-Golgi membranes and the intermediate compartment localize to induced membranes with distinct replication functions in flavivirus-infected cells. J Virol 73:9555–9567

Marra MA, Jones SJ, Astell CR, Holt RA, Brooks-Wilson A, Butterfield YS, Khattra J, Asano JK, Barber SA, Chan SY, Cloutier A, Coughlin SM, Freeman D, Girn N, Griffith OL, Leach SR, Mayo M, McDonald H, Montgomery SB, Pandoh PK, Petrescu AS, Robertson AG, Schein JE, Siddiqui A, Smailus DE, Stott JM, Yang GS, Plummer F, Andonov A, Artsob H, Bastien N, Bernard K, Booth TF, Bowness D, Czub M, Drebot M, Fernando L, Flick R, Garbutt M, Gray M, Grolla A, Jones S, Feldmann H, Meyers A, Kabani A, Li Y, Normand S, Stroher U, Tipples GA, Tyler S, Vogrig R, Ward D, Watson B, Brunham RC, Krajden M, Petric M, Skowronski DM, Upton C, Roper RL (2003) The genome sequence of the SARS-associated coronavirus. Science 300:1399–1404

Martzen MR, McCraith SM, Spinelli SL, Torres FM, Fields S, Grayhack EJ, Phizicky EM (1999) A biochemical genomics approach for identifying genes by the activity of their products. Science 286:1153–1155

Matthews DA, Smith WW, Ferre RA, Condon B, Budahazi G, Sisson W, Villafranca JE, Janson CA, McElroy HE, Gribskov CL, et al. (1994) Structure of human rhinovirus 3C protease reveals a trypsin-like polypeptide fold, RNA-binding site, and means for cleaving precursor polyprotein. Cell 77:761–771

Miller DJ, Schwartz MD, Ahlquist P (2001) Flock house virus RNA replicates on outer mitochondrial membranes in *Drosophila* cells. J Virol 75:11664–11676

Mills DR, Priano C, DiMauro P, Binderow BD (1989) Q beta replicase: mapping the functional domains of an RNA-dependent RNA polymerase. J Mol Biol 205:751–764

Mosimann SC, Cherney MM, Sia S, Plotch S, James MN (1997) Refined X-ray crystallographic structure of the poliovirus 3C gene product. J Mol Biol 273:1032–1047

Nasr F, Filipowicz W (2000) Characterization of the *Saccharomyces cerevisiae* cyclic nucleotide phosphodiesterase involved in the metabolism of ADP-ribose 1″,2″-cyclic phosphate. Nucleic Acids Res 28:1676–1683

Ng LF, Liu DX (2000) Further characterization of the coronavirus infectious bronchitis virus 3C-like proteinase and determination of a new cleavage site. Virology 272:27–39

Ng LF, Liu DX (2002) Membrane association and dimerization of a cysteine-rich, 16-kilodalton polypeptide released from the C-terminal region of the coronavirus infectious bronchitis virus 1a polyprotein. J Virol 76:6257–6267

Pang PS, Jankowsky E, Planet PJ, Pyle AM (2002) The hepatitis C viral NS3 protein is a processive DNA helicase with cofactor enhanced RNA unwinding. EMBO J 21:1168–1176

Pedersen KW, van der Meer Y, Roos N, Snijder EJ (1999) Open reading frame 1a-encoded subunits of the arterivirus replicase induce endoplasmic reticulum-derived double-membrane vesicles which carry the viral replication complex. J Virol 73:2016–2026

Penzes Z, González JM, Calvo E, Izeta A, Smerdou C, Mendez A, Sánchez CM, Sola I., Almazán F, Enjuanes L (2001) Complete genome sequence of transmissible gastroenteritis coronavirus PUR46-MAD clone and evolution of the Purdue virus cluster. Virus Genes 23:105–118

Peränen J, Kääriäinen L (1991) Biogenesis of type I cytopathic vacuoles in Semliki Forest virus-infected BHK cells. J Virol 65:1623–1627

Peränen J, Laakkonen P, Hyvönen M, Kääriäinen L (1995) The alphavirus replicase protein nsP1 is membrane-associated and has affinity to endocytic organelles. Virology 208:610–620

Piñon JD, Mayreddy RR, Turner JD, Khan FS, Bonilla PJ, Weiss SR (1997) Efficient autoproteolytic processing of the MHV-A59 3C-like proteinase from the flanking hydrophobic domains requires membranes. Virology 230:309–322

Piñon JD, Teng H, Weiss SR (1999) Further requirements for cleavage by the murine coronavirus 3C-like proteinase: identification of a cleavage site within ORF1b. Virology 263:471–484

Plotch SJ, Palant O, Gluzman Y (1989) Purification and properties of poliovirus RNA polymerase expressed in *Escherichia coli*. J Virol 63:216–225

Restrepo-Hartwig M, Ahlquist P (1999) Brome mosaic virus RNA replication proteins 1a and 2a colocalize and 1a independently localizes on the yeast endoplasmic reticulum. J Virol 73:10303–10309

Restrepo-Hartwig MA, Ahlquist P (1996) Brome mosaic virus helicase- and polymerase-like proteins colocalize on the endoplasmic reticulum at sites of viral RNA synthesis. J Virol 70:8908–8916

Rota PA, Oberste MS, Monroe SS, Nix WA, Campagnoli R, Icenogle JP, Penaranda S, Bankamp B, Maher K, Chen MH, Tong S, Tamin A, Lowe L, Frace M, DeRisi JL, Chen Q, Wang D, Erdman DD, Peret TC, Burns C, Ksiazek TG, Rollin PE, Sanchez A, Liffick S, Holloway B, Limor J, McCaustland K, Olsen-Rasmussen M, Fouchier R, Gunther S, Osterhaus AD, Drosten C, Pallansch MA, Anderson LJ, Bellini WJ (2003) Characterization of a novel coronavirus associated with severe acute respiratory syndrome. Science 300:1394–1399

Ruan YJ, Wei CL, Ee AL, Vega VB, Thoreau H, Su ST, Chia JM, Ng P, Chiu KP, Lim L, Zhang T, Peng CK, Lin EO, Lee NM, Yee SL, Ng LF, Chee RE, Stanton LW, Long PM, Liu ET (2003) Comparative full-length genome sequence analysis of 14 SARS coronavirus isolates and common mutations associated with putative origins of infection. Lancet 361:1779–1785

Russo M, Di Franco A, Martelli GP (1983) The fine structure of *Cymbidium* ringspot virus infections in host tissues. III. Role of peroxisomes in the genesis of multivesicular bodies. J Ultrastruct Res 82:52–63

Ryan MD, Flint M (1997) Virus-encoded proteinases of the picornavirus supergroup. J Gen Virol 78:699–723

Sawicki D, Wang T, Sawicki S (2001) The RNA structures engaged in replication and transcription of the A59 strain of mouse hepatitis virus. J Gen Virol 82:385–396

Sawicki SG, Sawicki DL (1990) Coronavirus transcription: subgenomic mouse hepatitis virus replicative intermediates function in RNA synthesis. J Virol 64:1050–1056

Schaad MC, Baric RS (1994) Genetics of mouse hepatitis virus transcription: evidence that subgenomic negative strands are functional templates. J Virol 68:8169–8179

Schaad MC, Jensen PE, Carrington JC (1997) Formation of plant RNA virus replication complexes on membranes: role of an endoplasmic reticulum-targeted viral protein. EMBO J 16:4049–4059

Schiller JJ, Kanjanahaluethai A, Baker SC (1998) Processing of the coronavirus MHV-JHM polymerase polyprotein: identification of precursors and proteolytic products spanning 400 kilodaltons of ORF1a. Virology 242:288–302

Schlegel A, Giddings TH, Jr., Ladinsky MS, Kirkegaard K (1996) Cellular origin and ultrastructure of membranes induced during poliovirus infection. J Virol 70:6576–6588

Schmidt-Mende J, Bieck E, Hügle T, Penin F, Rice CM, Blum HE, Moradpour D (2001) Determinants for membrane association of the hepatitis C virus RNA-dependent RNA polymerase. J Biol Chem 276:44052–44063

Schwartz M, Chen J, Janda M, Sullivan M, den Boon J, Ahlquist P (2002) A positive-strand RNA virus replication complex parallels form and function of retrovirus capsids. Mol Cell 9:505–514

Seipelt J, Guarne A, Bergmann E, James M, Sommergruber W, Fita I, Skern T (1999) The structures of picornaviral proteinases. Virus Res 62:159–168

Sethna PB, Hung SL, Brian DA (1989) Coronavirus subgenomic minus-strand RNAs and the potential for mRNA replicons. Proc Natl Acad Sci USA 86:5626–5630

Sethna PB, Brian DA (1997) Coronavirus genomic and subgenomic minus-strand RNAs copartition in membrane-protected replication complexes. J Virol 71:7744–7749

Seybert A, Hegyi A, Siddell SG, Ziebuhr J (2000a) The human coronavirus 229E superfamily 1 helicase has RNA and DNA duplex-unwinding activities with 5′-to-3′ polarity. RNA 6:1056–1068

Seybert A, van Dinten LC, Snijder EJ, Ziebuhr J (2000b) Biochemical characterization of the equine arteritis virus helicase suggests a close functional relationship between arterivirus and coronavirus helicases. J Virol 74:9586–9593

Seybert A, Ziebuhr J (2001) Guanosine triphosphatase activity of the human coronavirus helicase. Adv Exp Med Biol 494:255–260

Shi ST, Schiller JJ, Kanjanahaluethai A, Baker SC, Oh JW, Lai MM (1999) Colocalization and membrane association of murine hepatitis virus gene 1 products and de novo-synthesized viral RNA in infected cells. J Virol 73:5957–5969

Siddell S, Sawicki D, Meyer Y, Thiel V, Sawicki S (2001) Identification of the mutations responsible for the phenotype of three MHV RNA-negative *ts* mutants. Adv Exp Med Biol 494:453–458

Siddell SG. (1995). The Coronaviridae: an introduction. *In* "The Coronaviridae" (Siddell SG, ed.), pp. 1–10. Plenum Press, New York.

Sims AC, Ostermann J, Denison MR (2000) Mouse hepatitis virus replicase proteins associate with two distinct populations of intracellular membranes. J Virol 74:5647–5654

Snijder EJ, den Boon JA, Bredenbeek PJ, Horzinek MC, Rijnbrand R, Spaan WJ (1990a) The carboxyl-terminal part of the putative Berne virus polymerase is expressed by ribosomal frameshifting and contains sequence motifs which indicate that toro- and coronaviruses are evolutionarily related. Nucleic Acids Res 18:4535–4542

Snijder EJ, Horzinek MC (1993) Toroviruses: replication, evolution and comparison with other members of the coronavirus-like superfamily. J Gen Virol 74:2305–2316

Snijder EJ, Meulenberg JJ (1998) The molecular biology of arteriviruses. J Gen Virol 79:961–979

Snijder EJ, van Tol H, Roos N, Pedersen KW (2001) Non-structural proteins 2 and 3 interact to modify host cell membranes during the formation of the arterivirus replication complex. J Gen Virol 82:985–994

Snijder EJ, Bredenbeek PJ, Dobbe JC, Thiel V, Ziebuhr J, Poon LL, Guan Y, Rozanov M, Spaan WJ, Gorbalenya AE (2003) Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. J Mol Biol 331:991–1004

Spaan W, Delius H, Skinner M, Armstrong J, Rottier P, Smeekens S, van der Zeijst BA, Siddell SG (1983) Coronavirus mRNA synthesis involves fusion of non-contiguous sequences. EMBO J 2:1839–1844

Stinchcombe JC, Griffiths GM (1999) Regulated secretion from hemopoietic cells. J Cell Biol 147:1-6

Strauss JH, Strauss EG (1988) Evolution of RNA viruses. Annu Rev Microbiol 42:657–683

Tanner JA, Watt RM, Chai YB, Lu LY, Lin MC, Peiris JS, Poon LL, Kung HF, Huang JD (2003) The severe acute respiratory syndrome (SARS) coronavirus NTPase/helicase belongs to a distinct class of 5′ to 3′ viral helicases. J Biol Chem 278:39578–39582

Teng H, Piñon JD, Weiss SR (1999) Expression of murine coronavirus recombinant papain-like proteinase: efficient cleavage is dependent on the lengths of both the substrate and the proteinase polypeptides. J Virol 73:2658–2666

Teterina NL, Bienz K, Egger D, Gorbalenya AE, Ehrenfeld E (1997) Induction of intracellular membrane rearrangements by HAV proteins 2C and 2BC. Virology 237:66–77

Thiel V, Herold J, Schelle B, Siddell SG (2001a) Infectious RNA transcribed in vitro from a cDNA copy of the human coronavirus genome cloned in vaccinia virus. J Gen Virol 82:1273–1281

Thiel V, Herold J, Schelle B, Siddell SG (2001b) Viral replicase gene products suffice for coronavirus discontinuous transcription. J Virol 75:6676–6681

Thiel V, Ivanov KA, Putics A, Hertzig T, Schelle B, Bayer S, Weissbrich B, Snijder EJ, Rabenau H, Doerr HW, Gorbalenya AE, Ziebuhr J (2003) Mechanisms and enzymes involved in SARS coronavirus genome expression. J Gen Virol 84:2305–2315

Tibbles KW, Brierley I, Cavanagh D, Brown TD (1996) Characterization in vitro of an autocatalytic processing activity associated with the predicted 3C-like proteinase domain of the coronavirus avian infectious bronchitis virus. J Virol 70:1923–1930

Tibbles KW, Cavanagh D, Brown TD (1999) Activity of a purified His-tagged 3C-like proteinase from the coronavirus infectious bronchitis virus. Virus Res. 60:137–145

Tijms MA, van Dinten LC, Gorbalenya AE, Snijder EJ (2001) A zinc finger-containing papain-like protease couples subgenomic mRNA synthesis to genome translation in a positive-stranded RNA virus. Proc Natl Acad Sci USA 98:1889–1894

van der Meer Y, van Tol H, Krijnse Locker J, Snijder EJ (1998) ORF1a-encoded replicase subunits are involved in the membrane association of the arterivirus replication complex. J Virol 72:6689–6698

van der Meer Y, Snijder EJ, Dobbe JC, Schleich S, Denison MR, Spaan WJ, Krijnse Locker J (1999) Localization of mouse hepatitis virus nonstructural proteins and RNA synthesis indicates a role for late endosomes in viral replication. J Virol 73:7641–7657

van Dinten LC, Rensen S, Gorbalenya AE, Snijder EJ (1999) Proteolytic processing of the open reading frame 1b-encoded part of arterivirus replicase is mediated by nsp4 serine protease and Is essential for virus replication. J Virol 73:2027–2037

van Dinten LC, van Tol H, Gorbalenya AE, Snijder EJ (2000) The predicted metal-binding region of the arterivirus helicase protein is involved in subgenomic mRNA synthesis, genome replication, and virion biogenesis. J Virol 74:5213–5223

van Kuppeveld FJ, Galama JM, Zoll J, Melchers WJ (1995) Genetic analysis of a hydrophobic domain of coxsackie B3 virus protein 2B: a moderate degree of hydrophobicity is required for a *cis*-acting function in viral RNA synthesis. J Virol 69:7782–7790.

Vasiljeva L, Merits A, Golubtsov A, Sizemskaja V, Kaariainen L, Ahola T (2003) Regulation of the sequential processing of Semliki Forest virus replicase polyprotein. J Biol Chem 278:41636–41645

Wang T, Sawicki SG (2001) Mouse hepatitis virus minus-strand templates are unstable and turnover during viral replication. Adv Exp Med Biol 494:491–497

Xu HY, Lim KP, Shen S, Liu DX (2001) Further identification and characterization of novel intermediate and mature cleavage products released from the ORF 1b region of the avian coronavirus infectious bronchitis virus 1a/1b polyprotein. Virology 288:212–222

Yount B, Curtis KM, Baric RS (2000) Strategy for systematic assembly of large RNA and DNA genomes: transmissible gastroenteritis virus model. J Virol 74:10600–10611

Yount B, Denison MR, Weiss SR, Baric RS (2002) Systematic assembly of a full-length infectious cDNA of mouse hepatitis virus strain A59. J Virol 76:11065–11078

Yount B, Curtis KM, Fritz EA, Hensley LE, Jahrling PB, Prentice E, Denison MR, Geisbert TW, Baric RS (2003) Reverse genetics with a full-length infectious cDNA of severe acute respiratory syndrome coronavirus. Proc Natl Acad Sci USA 100:12995–13000

Ziebuhr J, Herold J, Siddell SG (1995) Characterization of a human coronavirus (strain 229E) 3C-like proteinase activity. J Virol 69:4331–4338

Ziebuhr J, Heusipp G, Siddell SG (1997) Biosynthesis, purification, and characterization of the human coronavirus 229E 3C-like proteinase. J Virol 71:3992–3997

Ziebuhr J, Siddell SG (1999) Processing of the human coronavirus 229E replicase polyproteins by the virus-encoded 3C-like proteinase: identification of proteolytic products and cleavage sites common to pp1a and pp1ab. J Virol 73:177–185

Ziebuhr J, Snijder EJ, Gorbalenya AE (2000) Virus-encoded proteinases and proteolytic processing in the *Nidovirales*. J Gen Virol 81:853–879

Ziebuhr J, Thiel V, Gorbalenya AE (2001) The autocatalytic release of a putative RNA virus transcription factor from its polyprotein precursor involves two paralogous papain-like proteases that cleave the same peptide bond. J Biol Chem 276:33220–33232

Ziebuhr J, Bayer S, Cowley JA, Gorbalenya AE (2003) The 3C-like proteinase of an invertebrate nidovirus links coronavirus and potyvirus homologs. J Virol 77:1415–1426

Zuo Y, Deutscher MP (2001) Exoribonuclease superfamilies: structural analysis and phylogenetic distribution. Nucleic Acids Res 29:1017–1026