ELSEVIER

Molecular biology of severe acute respiratory syndrome coronavirus
John Ziebuhr

The worldwide epidemic of severe acute respiratory syndrome (SARS) in 2003 was caused by a novel coronavirus called SARS-CoV. Coronaviruses and their closest relatives possess extremely large plus-strand RNA genomes and employ unique mechanisms and enzymes in RNA synthesis that separate them from all other RNA viruses. The SARS epidemic prompted a variety of studies on multiple aspects of the coronavirus replication cycle, yielding both rapid identification of the entry mechanisms of SARS-CoV into host cells and valuable structural and functional information on SARS-CoV proteins. These recent advances in coronavirus research have important implications for the development of anti-SARS drugs and vaccines.

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

Abbreviations

2′-O-MT	2′-O-ribose methyltransferase
3CL^{pro}	3C-like main protease
ADRP	ADP-ribose 1″-phosphatase
CPD	cyclic phosphodiesterase
ExoN	3′-to-5′ exoribonuclease
HCoV-229E	human coronavirus 229E
HR	heptad repeat
NendoU	nidoviral uridylate-specific endoribonuclease
ORF	open reading frame
PL2^{pro}	papain-like protease 2
RdRp	RNA-dependent RNA polymerase
SARS	severe acute respiratory syndrome
SARS-CoV	severe acute respiratory syndrome coronavirus
sg mRNA	subgenomic mRNA
TRS	transcription-regulating sequence

Introduction

In November 2002, an atypical pneumonia, characterized by progressive respiratory failure and death in approximately 10% of cases, emerged in Guangdong Province, Southern China [1$^{••}$,2,3]. Carlo Urbani, a WHO specialist in infectious diseases, was the first to recognize that this disease was unusual and that it represented a major threat to public health. Together with the Vietnamese authorities and WHO, he immediately introduced effective infection control measures that eventually stopped the further spread of the disease in Vietnam. Sadly, he would not survive to see this success. He contracted the disease and died as a result on March 29, 2003. The disease was named the severe acute respiratory syndrome (SARS) and a novel coronavirus, termed SARS-CoV, was rapidly identified as the etiological agent [4–6,7$^{•}$]. The rapid spread of the disease to neighboring regions and other countries prompted the WHO to issue a global alert on March 13, 2003. By the end of the epidemic, in July 2003, more than 8400 SARS cases and around 800 deaths due to SARS had been recorded worldwide. SARS-CoV is only distantly related to other human coronaviruses, such as 229E (coronavirus group 1) and OC43 (coronavirus group 2), which are known to cause the common cold and, in only a few cases, are associated with lower respiratory tract illness and diarrhea [8]. There is a large body of seroepidemiological data suggesting that SARS-CoV had not previously been endemic in humans [6,7$^{•}$]. Conversely, there is some initial evidence that at least a small proportion of healthy people in Hong Kong may have been exposed to SARS-CoV-related viruses up to two years before the SARS outbreak [9]. It is generally believed that SARS-CoV evolved from an animal coronavirus that recently crossed the species barrier. This hypothesis is supported by data published by Guan and co-workers [10$^{••}$], who isolated SARS-CoV-like viruses from Himalayan palm civets and a raccoon dog that were sold on a life-animal market in Guangdong, China. Significantly, an increased seroprevalence of SARS-CoV was observed in people trading with these animals. The potential for interspecies transmission of SARS-CoV is also illustrated by the fact that a whole range of animals, including cats, ferrets, mice and macaques can be infected with SARS-CoV [4,11,12]. The animal reservoir of SARS-CoV in nature remains to be identified.

Coronaviruses are enveloped, plus-strand RNA viruses that feature the largest RNA genomes currently known. In terms of genome structure and expression, the *Coronaviridae* (genera *Coronavirus* and *Torovirus*) and their distant relatives from the *Arteriviridae* and *Roniviridae* families, which together form the virus order *Nidovirales*, differ significantly from other positive-strand RNA viruses. In this review, current understanding of the crucial steps of the SARS-CoV life cycle will be summarized, focusing on genome organization, gene expression

and enzymes that are involved in genome replication and discontinuous synthesis of subgenomic mRNAs.

Genome organization

The SARS-CoV genome encompasses 29 727 nucleotides [excluding the 3′ poly(A) tail], of which 265 and 342 nucleotides, respectively, are located in the 5′ and 3′-nontranslated regions (Figure 1) [13,14]. The genome is predicted to contain 14 functional open reading frames (ORFs) (Figure 1) [15$^{\bullet\bullet}$]. Two large, 5′-terminal ORFs, 1a and 1b, constitute the replicase gene, which encodes the proteins that are required for viral RNA synthesis (and probably has other functions). The remaining twelve ORFs encode the four structural proteins, S, M, N and E, and eight accessory proteins that are not likely to be essential in tissue culture but may provide a selective advantage in the infected host. On the basis of unrooted phylogenetic trees, SARS-CoV was initially proposed to represent a new group ('group 4') within the genus *Coronavirus* [13,14]; however, rooted trees using torovirus and arterivirus sequences as outgroups convincingly placed SARS-CoV as a sister-lineage to the group 2 coronaviruses [16$^{\bullet\bullet}$,17].

Cellular receptor

Entry of coronaviruses into target cells is initiated by binding of the viral S protein to receptor molecules. The S protein forms typical petal-shaped spikes on the surface of the virion. It is heavily glycosylated and consists of three domains, the external N-terminal domain with its conserved S1 and S2 subdomains, a transmembrane domain, and a short cytoplasmic domain at the C-terminus. The cellular receptor of several group 1 coronaviruses is aminopeptidase N, a zinc metalloprotease [18], whereas mouse hepatitis virus (MHV) (group 2) uses carcinoembryonic antigen-related cell adhesion molecules as a cellular receptor [19]. Recently, the angiotensin-converting enzyme 2 was demonstrated to be a functional cellular receptor of SARS-CoV [20$^{\bullet\bullet}$]. The minimal binding domain of the SARS-CoV S protein was delimited to the S1 residues 318–510 [21] and antibodies specific for the S1 subunit of the SARS-CoV S protein were shown to neutralize SARS-CoV infection [22$^{\bullet}$].

Genome expression

Following S protein-mediated fusion of the viral envelope with the host cell membrane (see Update) and release of the viral genome RNA into the cytoplasm of the infected cell [23], SARS-CoV genome expression begins with the (cap-dependent) translation of the genomic RNA (mRNA 1; Figure 1). The translation product that is encoded by ORF1a is a protein of 4382 amino acid residues and is called polyprotein 1a (pp1a). Due to ribosomal

Figure 1
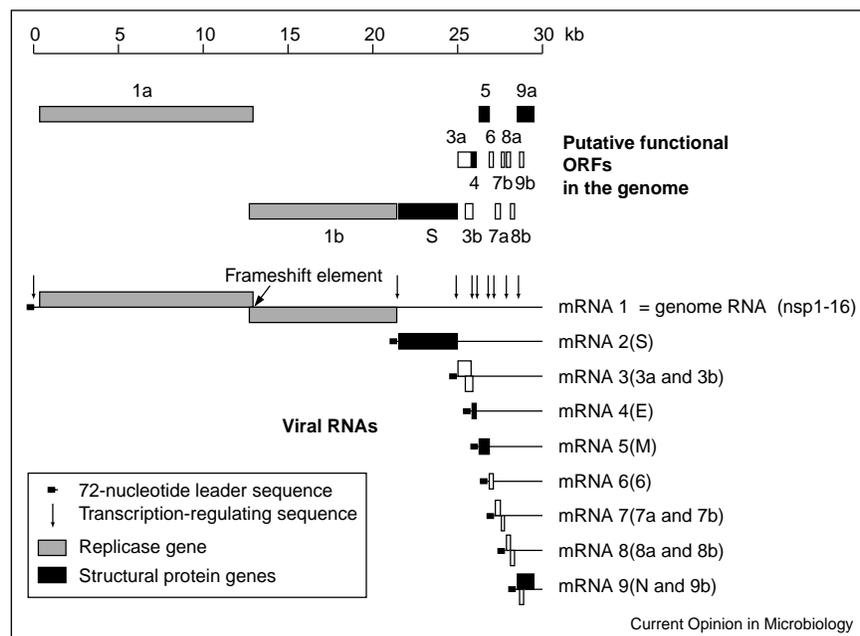
Genome organization and RNA synthesis of SARS-CoV. The putative functional ORFs in the genome of SARS-CoV are indicated. The 14 ORFs are expressed from the genome RNA (mRNA 1) and a nested set of sg mRNAs (mRNAs 2–9) that all have a common leader sequence derived from the 5′ end of the genome. The complement of this leader sequence ('antileader') is fused to the 3′ ends of nascent minus-strands by discontinuous RNA synthesis, which involves transcription-regulating sequences, the positions of which in the genome RNA are indicated here (see main text for details). The key functions that are required for the replication of the viral genome RNA and the synthesis of sg RNAs are encoded by the SARS-CoV replicase gene, comprising ORFs 1a and 1b. Expression of ORF1b sequences requires a programmed ribosomal frameshift into the −1 reading frame during translation of the genome RNA, which occurs just upstream of the ORF1a translation stop codon.

frameshifting into the –1 reading frame, occurring just upstream of the ORF1a translation stop codon, pp1a can be extended with ORF1b-encoded sequences to yield the 7073-residue polyprotein 1ab (pp1ab). The signal mediating the frameshift consists of the 'slippery' sequence, ^{13392}UUUAAAC^{13398}, and a downstream RNA pseudoknot structure [15••]. Polyproteins pp1a and pp1ab are extensively processed by viral proteinases [15••,24,25•] to yield a huge multi-subunit protein complex called 'viral replicase-transcriptase'. Together with a number of cellular factors, this protein complex mediates the replication of the viral genome and the transcription of a nested set of eight subgenomic (sg) mRNAs [15••]. Each of the sg mRNAs carries a 72-nucleotide, 5′-terminal leader sequence that is derived from the 5′-end of the genome [15••]. The leader sequence is acquired by a unique mechanism that involves discontinuous synthesis of sg minus strands and is dependent on *cis*-active RNA elements, known as 'transcription-regulating sequences' (TRSs) [26,27]. The TRSs of SARS-CoV have a common core sequence, ACGAAC [15••], that, by complementary base-pairing, assists in the transfer of the nascent minus strand to the TRS (leader TRS), located downstream of the 5′-leader sequence on the genomic RNA [26]. Besides complementary base-pairing, the transfer of the nascent minus strand to the 5′-leader TRS is thought to involve protein–protein interactions that keep the 5′-end of the genome in close proximity to the site of ongoing minus-strand synthesis. The current model of coronavirus sg RNA synthesis further suggests that, if the minus strand polymerase encounters attenuation signals that cause it to stall, the genome's 5′-end would provide an alternative template, allowing minus-strand synthesis to be continued and completed [27]. The resulting antileader-containing minus-strand RNAs are subsequently used as templates for (continuous) plus-strand synthesis of sg mRNAs. Analysis of SARS-CoV intracellular RNA synthesis, along with sequence analysis of the 5′-ends of SARS-CoV-specific RNAs confirmed the joining of noncontiguous genomic sequences in all the sg mRNAs and allowed reliable predictions on functional ORFs in the SARS-CoV genome [15••]. Thus, the SARS-CoV RNAs 2 to 9 are predicted to encode the four structural proteins S, M, N and E, as well as eight SARS-CoV-specific proteins with currently unknown functions. The sg RNAs are either functionally monocistronic (mRNAs 2, 4, 5 and 6) or bicistronic (mRNAs 3, 7, 8 and 9) (Figure 1) [15••,16••].

Proteolytic processing of the replicative polyproteins

The production of a complex and diverse set of RNA molecules by SARS-CoV and other coronaviruses (and nidoviruses) is linked to an unparalleled complexity of the replicative polyproteins, which are anchored to intracellular membranes and contain a variety of enzymatic activities [16••,28]. Coronaviruses control the activities of their replicative proteins by co- and post-translational processing of the nonstructural polyproteins [24] and ribosomal frameshifting [29] – thus ensuring a specific molar ratio between ORF1a- and ORF1b-encoded proteins. Generally, coronaviruses employ two papain-like proteases, PL1^{pro} and PL2^{pro}, to process the N-proximal regions of the replicative polyproteins at three sites. By contrast, SARS-CoV encodes only one papain-like protease, the activity of which has been established recently [15••]. The SARS-CoV enzyme is a PL2^{pro} orthologue, which, in contrast to most other coronavirus papain-like proteases, features a narrow substrate specificity. This might improve the potential for identifying selective inhibitors [15••]. In common with other coronavirus papain-like proteases [30], the SARS-CoV PL2^{pro} contains a putative Zn-finger structure, connecting the α- and β-domains of a papain-like fold. Based on HCoV-229E (human coronavirus 229E) and EAV (equine arteritis virus) data, the Zn-finger is predicted to be required for the proteolytic activity of PL2^{pro} and may have distinct functions in coronavirus sg RNA synthesis [30,31]. The central and C-terminal regions of the replicative polyproteins, pp1a and pp1ab, are cleaved by a chymotrypsin-like protease that, because of its distant relationship with the 3C proteases of picornaviruses, is named 3C-like protease, 3CL^{pro} [24]. The 3CL^{pro} plays a pivotal role in coronavirus polyprotein processing and also releases the key replicative functions of the virus, such as RdRp and helicase; therefore, it is also called the coronavirus main protease, M^{pro} [24,28]. Both in terms of function and structure, it represents the best-characterized coronavirus enzyme to date. In common with the 3CL^{pro}s of group 1 coronaviruses [25•,32–34], the SARS-CoV 3CL^{pro} employs a catalytic Cys–His dyad and has a three-domain structure, in which the N-terminal, chymotrypsin-like two-β-barrel fold (domains I and II) is connected by a 16-residue loop to the C-terminal domain III, consisting of five α-helices [25•,35•]. Biochemical data, as well as crystal structure information and NMR data, consistently implicate the 16-residue loop in substrate-binding [25•,34,35•,36]. In the polyprotein, the 3CL^{pro} is flanked by hydrophobic, probably membrane-spanning domains. At present, it is not clear whether the 3CL^{pro} cleaves itself in *cis* or *trans* from the replicase polyprotein precursor; however, once released, the *trans*-cleavage activity seems to depend on 3CL^{pro} dimerization that mainly involves the enzyme's N-terminus, domain II and, in particular, the α-helical domain III [25•,35•,36,37]. Several intermolecular and intramolecular interactions appear to be tailor-made to keep the enzyme in a conformation that is capable of cleaving substrates in *trans* and preventing self-inactivation by backfolding of the chain termini.

The SARS-CoV 3CL^{pro} cleaves pp1a and pp1ab at 11 sites and has a substrate specificity [(A,V,T,P)-X-(L,I,F,V,M)-Q↓(S,A,G,N)] that is very similar to previously characterized coronavirus 3CL^{pro}s [15••,37,38].

Despite conservation of the P4, P2, P1, and P1′ positions among coronavirus 3CL^{pro} substrates, there is preliminary evidence to suggest a significant structural flexibility for the SARS-CoV 3CL^{pro} active site, which may even lead to differential binding modes of specific peptidyl inhibitors to group 1 [(porcine) transmissible gastroenteritis virus; TGEV] versus group 2 (SARS-CoV) coronavirus 3CL^{pro}s [25•,35•]. Both the flexibility of the active site and the data from another study, which recently suggested that the P′ residues (despite little conservation) also contribute significantly to the substrate-binding by SARS-CoV 3CL^{pro}, might have important implications for the design of protease inhibitors [36].

Proteins involved in RNA synthesis and processing

Due to its pivotal role in viral RNA synthesis, the ~106-kDa SARS-CoV RdRp (RNA-dependent RNA polymerase; nsp12) represents an attractive target for anti-SARS therapy. However, there is a lack of structural and biochemical information on any coronavirus RdRp and structural predictions are complicated by the fact that the coronavirus RdRps are significantly diverged from cellular and viral RNA polymerases. Recently, a structure model was built for the catalytic domain of the SARS-CoV RdRp [39•]. The model provides first insights into the active site of the protein and also enables conclusions to be drawn about the properties of potential nucleoside analog inhibitors of coronavirus RdRps. Thus, it was proposed that potential nucleoside analog inhibitors should contain groups at their 2′ and 3′ positions that are capable of making hydrogen-bonding interactions with RdRp residues 623 and 691. Furthermore, to avoid steric conflicts in the binding to the 2′ and 3′ positions, the potential nucleoside inhibitors should possess the C3′ *endo* sugar puckering conformation. Clearly, direct structural information is highly desirable for the development of effective inhibitors of this key enzyme.

The SARS-CoV superfamily 1 helicase resides in nsp13. The enzyme's catalytic domain is linked at its N-terminus to a complex zinc-binding domain (Figure 2) [16••,40]. Data obtained for an arterivirus homolog indicate that coronavirus helicases might have distinct functions in replication and transcription and, possibly, even in virion biogenesis [41]. The SARS-CoV helicase is a multifunctional protein. It has been shown to have: (i) single-stranded and double-stranded RNA and DNA binding activities; (ii) nucleic acid-stimulated NTPase and dNTPase activities; (iii) RNA and DNA duplex-unwinding activities and; (iv) RNA 5′-triphosphatase activity [42,43] (which is proposed to mediate the first step of 5′-cap synthesis on coronavirus RNAs). The coronavirus helicase acts processively in a 5′-to-3′ direction on partial-duplex RNA and DNA substrates and, consistent with its presumed replicative function, is capable of unwinding long stretches of double-stranded nucleic acids [42,44,45]. The 5′-to-3′ polarity of nidovirus

Figure 2
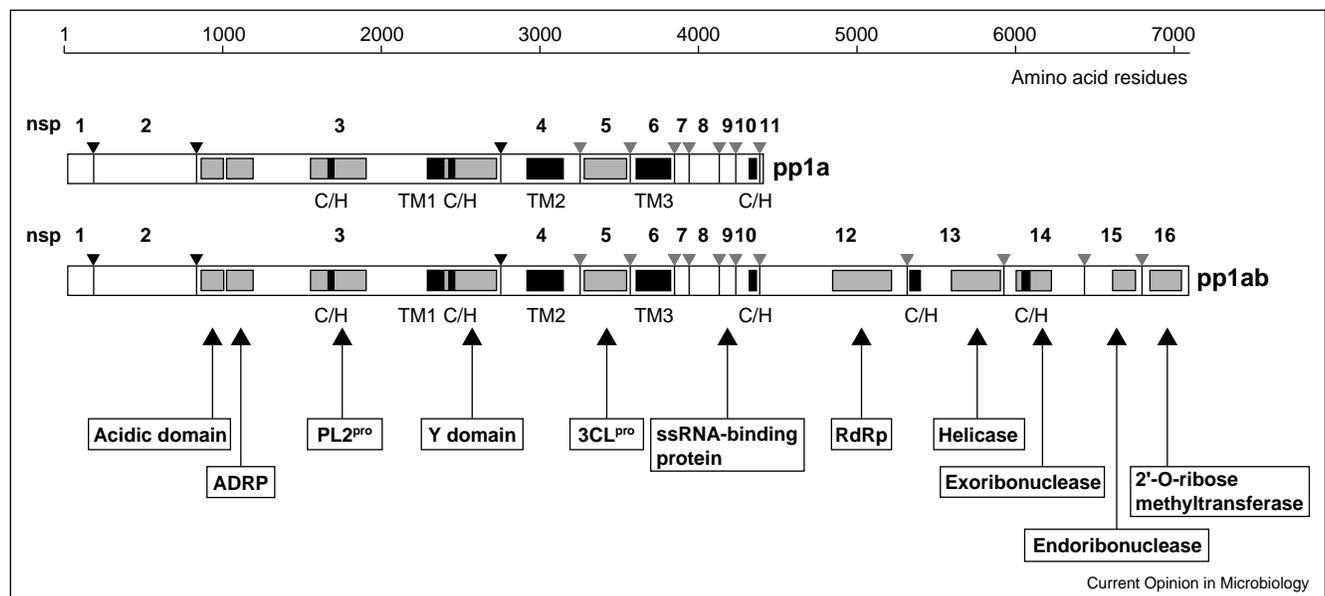
Overview of the domain organization and proteolytic processing of SARS-CoV replicase polyproteins, pp1a (486 kDa) and pp1ab (790 kDa). The processing end-products of pp1a are designated nonstructural proteins (nsp) 1 to nsp11 and those of pp1ab are designated nsp1 to nsp10 and nsp12 to nsp16. Cleavage sites that are predicted to be processed by the viral main protease, 3CL^{pro}, are indicated by grey arrowheads, and sites that are processed by the papain-like protease, PL2^{pro}, are indicated by black arrowheads. For further details on SARS-CoV replicative proteins, see Table 1. ADRP, ADP-ribose 1″-phosphatase; PL2^{pro}, papain-like protease 2; 3CL^{pro}, 3C-like main protease; RdRp, RNA-dependent RNA polymerase; TM1, TM2, TM3, transmembrane domains 1, 2 and 3; C/H, domains containing conserved Cys and His residues.

Table 1

The replicase gene products of severe acute respiratory syndrome coronavirus.

Protein	Position in polyproteins pp1a and pp1ab[a], respectively		Protein size (amino acid residues)[a]	Features[a,b]
nsp1	Met^1	–Gly^{180}	180	ND
nsp2	Ala^{181}	–Gly^{818}	638	ND
nsp3	Ala^{819}	–Gly^{2740}	1922	Acidic domain, ADP-ribose 1″-phosphatase, papain-like protease 2 (C/H), Y domain (TM1, C/H)
nsp4	Lys^{2741}	–Gln^{3240}	500	TM2
nsp5	Ser^{3241}	–Gln^{3546}	306	3C-like protease
nsp6	Gly^{3547}	–Gln^{3836}	290	TM3
nsp7	Ser^{3837}	–Gln^{3919}	83	ND
nsp8	Ala^{3920}	–Gln^{4117}	198	ND
nsp9	Asn^{4118}	–Gln^{4230}	113	ssRNA-binding protein^c
nsp10	Ala^{4231}	–Gln^{4369}	139	C/H
nsp11	Ser^{4370}	–Val^{4382}	13	ND
nsp12	Ser^{4370}	–Gln^{5301}	932	RNA polymerase
nsp13	Ala^{5302}	–Gln^{5902}	601	C/H, NTPase, dNTPase, 5′-to-3′ RNA helicase and DNA helicase, RNA 5′-triphosphatase^d
nsp14	Ala^{5903}	–Gln^{6429}	527	3′-to-5′ exoribonuclease, C/H
nsp15	Ser^{6430}	–Gln^{6775}	346	uridylate-specific endoribonuclease
nsp16	Ala^{6776}	–Asn^{7073}	298	2′-O-ribose methyltransferase

C/H, domain with conserved Cys/His residues; ND, no data; TM, transmembrane domain;. [a]Data from Snijder et al. [16••]. [b]Data from Ziebuhr [28]. [c]Data from Sutton et al. and Egloff et al. [53,54]. [d]Data from Ivanov et al. [42].

helicases [42–44,46] contrasts with an opposite (3′-to-5′) polarity of helicases from the *Flaviviridae*, indicating differential functions of helicases in the life cycle of the respective viruses.

In the context of a comprehensive sequence analysis of the SARS-CoV genome, as many as five novel coronaviral RNA processing activities were predicted recently [16••] (Figure 2 and Table 1). These include a 3′-to-5′ exonuclease (ExoN), a uridylate-specific endoribonuclease (NendoU), an S-adenosylmethionine-dependent 2′-O-ribose methyltransferase (2′-O-MT), an ADP-ribose 1″-phosphatase (ADRP) [47], and a cyclic phosphodiesterase (CPD). Four of the activities are conserved in all coronaviruses, including SARS-CoV, supporting their essential role in the coronaviral life cycle [16••]. The fact that ExoN (nsp14), NendoU (nsp15) and 2′-O-MT (nsp16) are arranged in pp1ab as a single protein block downstream of the RdRp and helicase domains (Figure 2) suggests a cooperation of these activities in the same metabolic pathway. As an initial clue to the potential functions of the nsp14–nsp16 proteins, an interesting parallel to cellular RNA processing pathways was noted by Snijder and co-workers [16••]. Thus, cellular homologs of coronavirus nsp14–16 cleave mRNAs to produce small nucleolar RNAs that guide specific 2′-O-ribose methylations on ribosomal RNA [48,49]. The activities of the predicted coronavirus enzymes and their viral and/or cellular substrates remain to be established. ADRP and CPD, this being conserved only in a subset of group 2 coronaviruses, excluding SARS-CoV, were also proposed to cooperate in a common (but currently unknown) pathway. This hypothesis is based on the fact that cellular homologs of CPD and ADRP mediate two consecutive steps in the downstream processing of ADP-ribose 1″,2″ cyclic phosphate (Appr>p), a side product of tRNA splicing [50].

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 that are conserved in all coronaviruses (ExoN, NendoU, 2′-O-MT and ADRP), only one activity (NendoU) is conserved in arteriviruses [16••]. The differential conservation pattern of RNA processing activities among the nidovirus families and genera suggests a functional hierarchy for these enzymes, with NendoU playing a major role. It might also reflect subtle differences in the RNA synthesis mechanisms used by various nidovirus families and/or differential interactions of nidovirus nonstructural proteins with host cell functions. Alternatively, the extra functions that are encoded by coronaviruses and toroviruses (and, to a lesser extent, roniviruses) might be required to replicate the extremely large (~30 kb) RNA genomes of these viruses. Thus, the predicted 3′-to-5′ exonuclease, ExoN, has been speculated to be involved in recombination or repair mechanisms that may be required for the life cycle of corona-, toro-, and roniviruses but may be dispensable for the much smaller arteriviruses [16••].

Conclusions
The SARS outbreak has inspired a myriad of studies into virtually every aspect of SARS-CoV biology, including viral pathogenesis, tissue tropism, genome structure,

expression and replication, as well as SARS-CoV structural and nonstructural proteins. Within a remarkably short period of time, these studies have produced a wealth of functional and structural information that might be used for the development of SARS-CoV-specific drugs, as well as vaccines. Already, initial candidate vaccines are currently being tested [51,52], crystal structures of SARS-CoV proteins have been determined [35•,53,54], a full-length infectious clone has been constructed, allowing reverse genetics with SARS-CoV [55], a functional receptor of the virus has been identified [20••], and both interferons [56,57,58•] and antibodies [11,22•] have been successfully used to block SARS-CoV infections in model systems. The rapidly increasing information on SARS and its etiological agent, together with sensitive diagnostic tests and improved surveillance by public health authorities, should provide a good basis for the control of SARS-CoV infections, should the virus be reintroduced into the human population in the future.

Update
Recent work has demonstrated that the heptad repeat (HR) regions, HR1 and HR2, present in the S2 subunit of the SARS-CoV S protein, assemble into an antiparallel six-helix bundle, consisting of HR1 as a central triple-stranded coiled-coil structure and three HR2 α-helices [59•,60•,61]. Analogous to other type-1 fusion glycoproteins, the formation of the six-helix bundle has been suggested to contribute to a conformational change that occurs in the S protein, following receptor binding, to form a fusion-active core that brings the viral and host cell membranes into close proximity, leading to the fusion between these membranes. Importantly, peptides derived from the HR2 sequence were shown to block SARS-CoV infection of Vero cells, suggesting a potential approach to the development of drugs for treatment or prophylaxis of SARS-CoV infections [59•,60•].

Acknowledgements

References and recommended reading
Papers of particular interest, published within the annual period of review, have been highlighted as:

- of special interest
- of outstanding interest

1. Peiris JS, Yuen KY, Osterhaus AD, Stöhr K: **The severe acute**
•• **respiratory syndrome**. *N Engl J Med* 2003, **349**:2431-2441.
This paper reviews the cause, epidemiology and clinical features of SARS and illustrates the impact of globalization and international air travel on the dissemination of emerging infectious diseases.

2. Zhao Z, Zhang F, Xu M, Huang K, Zhong W, Cai W, Yin Z, Huang S, Deng Z, Wei M *et al.*: **Description and clinical treatment of an early outbreak of severe acute respiratory syndrome (SARS) in Guangzhou, PR China**. *J Med Microbiol* 2003, **52**:715-720.

3. Zhong NS, Zheng BJ, Li YM, Poon, Xie ZH, Chan KH, Li PH, Tan SY, Chang Q, Xie JP *et al.*: **Epidemiology and cause of severe acute respiratory syndrome (SARS) in Guangdong, People's Republic of China, in February, 2003**. *Lancet* 2003, **362**:1353-1358.

4. Kuiken T, Fouchier RA, Schutten M, Rimmelzwaan GF, van Amerongen G, van Riel D, Laman JD, de Jong T, van Doornum G, Lim W *et al.*: **Newly discovered coronavirus as the primary cause of severe acute respiratory syndrome**. *Lancet* 2003, **362**:263-270.

5. Drosten C, Günther S, Preiser W, van der Werf S, Brodt HR, Becker S, Rabenau H, Panning M, Kolesnikova L, Fouchier RA *et al.*: **Identification of a novel coronavirus in patients with severe acute respiratory syndrome**. *N Engl J Med* 2003, **348**:1967-1976.

6. Ksiazek TG, Erdman D, Goldsmith CS, Zaki SR, Peret T, Emery S, Tong S, Urbani C, Comer JA, Lim W *et al.*: **A novel coronavirus associated with severe acute respiratory syndrome**. *N Engl J Med* 2003, **348**:1953-1966.

7. Peiris JS, Lai ST, Poon LL, Guan Y, Yam LY, Lim W, Nicholls J,
•   Yee WK, Yan WW, Cheung MT *et al.*: **Coronavirus as a possible cause of severe acute respiratory syndrome**. *Lancet* 2003, **361**:1319-1325.
This paper reports the identification and preliminary characterization of a novel coronavirus from SARS patients.

8. Myint SH: **Human coronavirus infections**. In *The Coronaviridae*. Edited by Siddell SG. Plenum Press; 1995:389-401.

9. Zheng BJ, Wong KH, Zhou J, Wong KL, Young BW, Lu LW, Lee SS: **SARS-related virus predating SARS outbreak, Hong Kong**. *Emerg Infect Dis* 2004, **10**:176-178.

10. Guan Y, Zheng BJ, He YQ, Liu XL, Zhuang ZX, Cheung CL,
•• Luo SW, Li PH, Zhang LJ, Guan YJ *et al.*: **Isolation and characterization of viruses related to the SARS coronavirus from animals in southern China**. *Science* 2003, **302**:276-278.
The authors report the isolation of SARS-CoV-like viruses from palm civets and raccoon dogs, sold on a life-animal market in Guangdong, China. The data suggest interspecies transmission and implicate animals as potential reservoirs of SARS-CoV or closely related coronaviruses.

11. Subbarao K, McAuliffe J, Vogel L, Fahle G, Fischer S, Tatti K, Packard M, Shieh WJ, Zaki S, Murphy B: **Prior infection and passive transfer of neutralizing antibody prevent replication of severe acute respiratory syndrome coronavirus in the respiratory tract of mice**. *J Virol* 2004, **78**:3572-3577.

12. Martina BE, Haagmans BL, Kuiken T, Fouchier RA, Rimmelzwaan GF, Van Amerongen G, Peiris JS, Lim W, Osterhaus AD: **Virology: SARS virus infection of cats and ferrets**. *Nature* 2003, **425**:915.

13. Marra MA, Jones SJ, Astell CR, Holt RA, Brooks-Wilson A, Butterfield YS, Khattra J, Asano JK, Barber SA, Chan SY *et al.*: **The Genome sequence of the SARS-associated coronavirus**. *Science* 2003, **300**:1399-1404.

14. Rota PA, Oberste MS, Monroe SS, Nix WA, Campagnoli R, Icenogle JP, Penaranda S, Bankamp B, Maher K, Chen MH *et al.*: **Characterization of a novel coronavirus associated with severe acute respiratory syndrome**. *Science* 2003, **300**:1394-1399.

15. Thiel V, Ivanov KA, Putics A, Hertzig T, Schelle B, Bayer S,
•• Weissbrich B, Snijder EJ, Rabenau H, Doerr HW *et al.*: **Mechanisms and enzymes involved in SARS coronavirus genome expression**. *J Gen Virol* 2003, **84**:2305-2315.
This paper describes the transcriptional and (post)translational strategies that are used by the virus. It provides the first detailed analysis of SARS-CoV RNA synthesis, identifies the frameshifting signal and establishes the enzymatic activities of the SARS-CoV NTPase/helicase, and the papain-like and 3C-like proteases. The data, together with the results from a bioinformatics analysis by Snijder and co-workers [16••], allow reliable predictions on SARS-CoV gene products.

16. Snijder EJ, Bredenbeek PJ, Dobbe JC, Thiel V, Ziebuhr J, Poon LL,
•• Guan Y, Rozanov M, Spaan WJ, Gorbalenya AE: **Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage**. *J Mol Biol* 2003, **331**:991-1004.
This paper provides the most exhaustive analysis of the SARS-CoV genome that has been published to date. It predicts as many as five novel coronavirus RNA processing activities. Four of the enzymes, including two ribonucleases, are conserved in SARS-CoV. The paper provides an important theoretical basis for future studies aimed at elucidating the molecular mechanisms of coronavirus (nidovirus) RNA synthesis that, in several aspects, differs from that of other RNA viruses.

17. Gibbs AJ, Gibbs MJ, Armstrong JS: **The phylogeny of SARS coronavirus**. *Arch Virol* 2004, **149**:621-624.

18. Yeager CL, Ashmun RA, Williams RK, Cardellichio CB, Shapiro LH, Look AT, Holmes KV: **Human aminopeptidase N is a receptor for human coronavirus 229E**. *Nature* 1992, **357**:420-422.

19. Williams RK, Jiang GS, Holmes KV: **Receptor for mouse hepatitis virus is a member of the carcinoembryonic antigen family of glycoproteins**. *Proc Natl Acad Sci USA* 1991, **88**:5533-5536.

20. Li W, Moore MJ, Vasilieva N, Sui J, Wong SK, Berne MA,
•• Somasundaran M, Sullivan JL, Luzuriaga K, Greenough TC *et al.*: **Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus**. *Nature* 2003, **426**:450-454.
This study identifies angiotensin-converting enzyme 2, a zinc metallo-protease, as a functional receptor for SARS-CoV. Expression of this protein was confirmed for the lung and kidney and, at lower levels, also for the heart and gastrointestinal tract, which correlates with the observed tropism of SARS-CoV in infected patients and tissue culture.

21. Wong SK, Li W, Moore MJ, Choe H, Farzan M: **A 193-amino acid fragment of the SARS coronavirus S protein efficiently binds angiotensin-converting enzyme 2**. *J Biol Chem* 2004, **279**:3197-3201.

22. Sui J, Li W, Murakami A, Tamin A, Matthews LJ, Wong SK, Moore
• MJ, Tallarico AS, Olurinde M, Choe H *et al.*: **Potent neutralization of severe acute respiratory syndrome (SARS) coronavirus by a human mAb to S1 protein that blocks receptor association**. *Proc Natl Acad Sci USA* 2004, **101**:2536-2541.
The authors describe a recombinant human single-chain variable region fragment that binds with high affinity to the S1 subdomain of the SARS-CoV S protein and efficiently neutralizes SARS-CoV in tissue culture.

23. Ng ML, Tan SH, See EE, Ooi EE, Ling AE: **Early events of SARS coronavirus infection in Vero cells**. *J Med Virol* 2003, **71**:323-331.

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

25. Anand K, Ziebuhr J, Wadhwani P, Mesters JR, Hilgenfeld R:
• **Coronavirus main proteinase (3CL^{pro}) structure: basis for design of anti-SARS drugs**. *Science* 2003, **300**:1763-1767.
The paper presents the crystal structures of the HCoV-229E 3CL^{pro} and a peptidyl inhibitor bound to the active site of TGEV 3CL^{pro}. Based on the conservation of these structures and, in particular, their main specificity sites, a homology model was built for the SARS-CoV 3CL^{pro}.

26. Zúñiga S, Sola I, Alonso S, Enjuanes L: **Sequence motifs involved in the regulation of discontinuous coronavirus subgenomic RNA synthesis**. *J Virol* 2004, **78**:980-994.

27. Sawicki SG, Sawicki DL: **A new model for coronavirus transcription**. *Adv Exp Med Biol* 1998, **440**:215-219.

28. Ziebuhr J: **The coronavirus replicase**. *Curr Top Microbiol Immunol* 2004, **287** in press.

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

30. Herold J, Siddell SG, Gorbalenya AE: **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* 1999, **274**:14918-14925.

31. Tijms MA, van Dinten LC, Gorbalenya AE, Snijder EJ: **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* 2001, **98**:1889-1894.

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

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

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

35. Yang H, Yang M, Ding Y, Liu Y, Lou Z, Zhou Z, Sun L, Mo L, Ye S,
• Pang H *et al.*: **The crystal structures of severe acute respiratory syndrome virus main protease and its complex with an inhibitor**. *Proc Natl Acad Sci USA* 2003, **100**:13190-13195.
This paper describes crystal structures of SARS-CoV 3CL^{pro} at different pH values and in complex with an inhibitor. The crystal structures presented in this and two previous studies [25•,34] are anticipated to provide a solid basis for the design of selective 3CL^{pro} inhibitors that may be developed to anticoronaviral drugs.

36. Shi J, Wei Z, Song J: **Dissection study on the SARS 3C-like protease reveals the critical role of the extra domain in dimerization of the enzyme: Defining the extra domain as a new target for design of highly-specific protease inhibitors**. *J Biol Chem* 2004, **279**:24765-24773.

37. Fan K, Wei P, Feng Q, Chen S, Huang C, Ma L, Lai B, Pei J, Liu Y, Chen J *et al.*: **Biosynthesis, purification, and substrate specificity of severe acute respiratory syndrome coronavirus 3C-like proteinase**. *J Biol Chem* 2004, **279**:1637-1642.

38. Hegyi A, Ziebuhr J: **Conservation of substrate specificities among coronavirus main proteases**. *J Gen Virol* 2002, **83**:595-599.

39. Xu X, Liu Y, Weiss S, Arnold E, Sarafianos SG, Ding J: **Molecular
• model of SARS coronavirus polymerase: implications for biochemical functions and drug design**. *Nucleic Acids Res* 2003, **31**:7117-7130.
The viral RNA-dependent RNA polymerase represents a key target for the antiviral therapy of SARS. The paper is the first to describe a structural model for the catalytic domain of the coronavirus polymerase, an enzyme that has profoundly diverged from both cellular and viral RNA polymerases.

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

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

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

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

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

45. Ivanov KA, Ziebuhr J: **Human coronavirus 229E nonstructural protein 13: characterization of duplex-unwinding, nucleoside triphosphatase, and RNA 5′-triphosphatase activities**. *J Virol* 2004, **78** in press.

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

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

48. Filipowicz W, Pogacic V: **Biogenesis of small nucleolar ribonucleoproteins**. *Curr Opin Cell Biol* 2002, **14**:319-327.

49. Kiss T: **Small nucleolar RNA-guided post-transcriptional modification of cellular RNAs**. *EMBO J* 2001, **20**:3617-3622.

50. Culver GM, Consaul SA, Tycowski KT, Filipowicz W, Phizicky EM: **tRNA splicing in yeast and wheat germ. A cyclic phosphodiesterase implicated in the metabolism of**

ADP-ribose 1″,2″-cyclic phosphate. *J Biol Chem* 1994, **269**:24928-24934.

51. Marshall E, Enserink M: **Medicine. Caution urged on SARS vaccines**. *Science* 2004, **303**:944-946.

52. Gao W, Tamin A, Soloff A, D'Aiuto L, Nwanegbo E, Robbins PD, Bellini WJ, Barratt-Boyes S, Gambotto A: **Effects of a SARS-associated coronavirus vaccine in monkeys**. *Lancet* 2003, **362**:1895-1896.

53. Sutton G, Fry E, Carter L, Sainsbury S, Walter T, Nettleship J, Berrow N, Owens R, Gilbert R, Davidson A *et al.*: **The nsp9 replicase protein of SARS-coronavirus, structure and functional insights**. *Structure (Camb)* 2004, **12**:341-353.

54. Egloff MP, Ferron F, Campanacci V, Longhi S, Rancurel C, Dutartre H, Snijder EJ, Gorbalenya AE, Cambillau C, Canard B: **The severe acute respiratory syndrome-coronavirus replicative protein nsp9 is a single-stranded RNA-binding subunit unique in the RNA virus world**. *Proc Natl Acad Sci USA* 2004, **101**:3792-3796.

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

56. Cinatl J, Morgenstern B, Bauer G, Chandra P, Rabenau H, Doerr HW: **Treatment of SARS with human interferons**. *Lancet* 2003, **362**:293-294.

57. Haagmans BL, Kuiken T, Martina BE, Fouchier RA, Rimmelzwaan GF, Van Amerongen G, Van Riel D, De Jong T, Itamura S, Chan KH *et al.*: **Pegylated interferon-alpha protects type 1 pneumocytes against SARS coronavirus infection in macaques**. *Nat Med* 2004, **10**:290-293.

58. Hertzig T, Scandella E, Schelle B, Ziebuhr J, Siddell SG, Ludewig
•   B, Thiel V: **Rapid identification of coronavirus replicase inhibitors using a selectable replicon RNA**. *J Gen Virol* 2004, **85**:1717-1725.
This study describes a non-cytopathic, selectable replicon RNA (based on HCoV-229E) that expresses a reporter gene and that can be stably maintained in eukaryotic cells. The replicon-containing cell line can be used to assess the effects of potential coronavirus replicase inhibitors without the need to grow infectious virus. Once adapted to SARS-CoV, the replicon approach might allow high-throughput testing of compound libraries at a reduced biosafety level.

59. Liu S, Xiao G, Chen Y, He Y, Niu J, Escalante CR, Xiong H, Farmar
•   J, Debnath AK, Tien P *et al.*: **Interaction between heptad repeat 1 and 2 regions in spike protein of SARS-associated coronavirus: implications for virus fusogenic mechanism and identification of fusion inhibitors**. *Lancet* 2004, **363**:938-947.
See annotation for [60•].

60. Bosch BJ, Martina BE, van Der Zee R, Lepault J, Haijema BJ,
•   Versluis C, Heck AJ, de Groot R, Osterhaus AD, Rottier PJ: **Severe acute respiratory syndrome coronavirus (SARS-CoV) infection inhibition using spike protein heptad repeat-derived peptides**. *Proc Natl Acad Sci USA* 2004, **101**:8455-8460.
This paper, together with [59•,61], demonstrates that the HR sequences, HR1 and HR2, of the SARS-CoV S protein assemble into an antiparallel six-helix bundle. The formation of this structure probably contributes to a conformational change in the S protein that triggers the fusion between viral and host-cell membranes. Peptides derived from the HR2 sequence are shown to effectively inhibit SARS-CoV infection of Vero cells.

61. Ingallinella P, Bianchi E, Finotto M, Cantoni G, Eckert DM, Supekar VM, Bruckmann C, Carfi A, Pessi A: **Structural characterization of the fusion-active complex of severe acute respiratory syndrome (SARS) coronavirus**. *Proc Natl Acad Sci USA* 2004, **101**:8709-8714.