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

Available online at www.sciencedirect.com

SCIENCE DIRECT

Virology 335 (2005) 165 – 176

VIROLOGY

www.elsevier.com/locate/yviro

Expression, purification, and characterization of SARS coronavirus RNA polymerase

Ao Cheng[a,1], Wei Zhang[b,1], Youhua Xie[a], Weihong Jiang[b], Eddy Arnold[c], Stefan G. Sarafianos[c], Jianping Ding[a,*]

[a]*Key Laboratory of Proteomics, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, 320 Yue-Yang Road, Shanghai 200031, China*

[b]*Institute of Plant Physiology and Ecology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, 300 Feng-Lin Road, Shanghai 200031, China*

[c]*Center for Advanced Biotechnology and Medicine (CABM) and Rutgers University Department of Chemistry and Chemical Biology, 679 Hoes Lane, Piscataway, NJ 08854-5638, USA*

Received 29 December 2004; returned to author for revision 20 January 2005; accepted 23 February 2005
Available online 24 March 2005

Abstract

The RNA-dependent RNA polymerase (RdRp) of SARS coronavirus (SARS-CoV) is essential for viral replication and a potential target for anti-SARS drugs. We report here the cloning, expression, and purification of the N-terminal GST-fused SARS-CoV RdRp and its polymerase catalytic domain in *Escherichia coli*. During purification, the full-length GST-RdRp was found to cleave into three main fragments: an N-terminal p12 fragment, a middle p30 fragment, and a C-terminal p64 fragment comprising the catalytic domain, presumably due to bacterial proteases. Biochemical assays show that the full-length GST-RdRp has RdRp activity and the p64 and p12 fragments form a complex that exhibits comparable RdRp activity, whereas the GST-p64 protein has no activity, suggesting that the p12 domain is required for polymerase activity possibly via involvement in template-primer binding. Nonnucleoside HIV-1 RT inhibitors are shown to have no evident inhibitory effect on SARS-CoV RdRp activity. This work provides a basis for biochemical and structural studies of SARS-CoV RdRp and for development of anti-SARS drugs.

Keywords: SARS; Coronavirus; RNA-dependent RNA polymerase; Polymerization

Introduction

Severe acute respiratory syndrome (SARS) is a new acute respiratory infectious disease and the outbreak of SARS in late 2002 in southeast China spread rapidly to over 30 countries and resulted in more than 800 deaths (Poutanen et al., 2003; Tsang et al., 2003). The causative agent of SARS is a previously unidentified positive-strand RNA virus that belongs to the Coronaviridae family, namely SARS coronavirus (SARS-CoV) (Drosten et al., 2003; Ksiazek et al., 2003; Peiris et al., 2003; Snijder et al., 2003). Currently, there is neither vaccine nor effective therapeutic treatments against this virus and a future resurgence of SARS is possible. So far, the very limited knowledge about SARS-CoV is mainly based on studies of other coronaviruses, in particular mouse hepatitis virus (MHV) which is very closely related to SARS-CoV (Navas-Martin and Weiss, 2003). Coronaviruses are enveloped RNA viruses with a single, positive-strand RNA genome (Lai and Holmes, 2001). The viral genome of SARS-CoV consists of about 29,727 nucleotides and encodes two large replicase polyproteins expressed by two open reading

Abbreviations: GST, glutathione S-transferase; MS, mass spectrometry; MHV, mouse hepatitis virus; RdRp, RNA-dependent RNA polymerase; SARS, severe acute respiratory syndrome; SARS-CoV, SARS coronavirus.

* Corresponding author. Fax: +1 86 21 54921116.
E-mail address: jpding@sibs.ac.cn (J. Ding).

[1] These authors contributed equally to this work.

0042-6822/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.virol.2005.02.017

frames (ORF1a and ORF1b) that are linked together by a ribosomal frameshift (Marra et al., 2003; Rota et al., 2003). These polyproteins undergo co-translational proteolytic processing by internal viral proteases into a set of mature non-structural proteins that carry out multiple important enzymatic functions during viral replication, including an RNA-dependent RNA polymerase (RdRp), a 3C-like serine proteinase (3CLpro), a papain-like proteinase (PL2pro), and a superfamily 1-like helicase (HEL1) (Snijder et al., 2003). In addition, the SARS-CoV genome also encodes a number of structural proteins characteristic to coronaviruses, including spike (S), envelope (E), membrane (M), nucleocapsid (N), and short untranslated regions at both termini.

By analogy to other positive-strand RNA viruses, SARS-CoV RdRp is predicted to be the central enzyme that, together with other viral and cellular proteins, constitutes a replication complex that is responsible for replicating the viral RNA genome (Bost et al., 2000; Brockway et al., 2003). The primary functions of the replication complex are to transcribe the full-length negative- and positive-strand RNAs, a 3′-coterminal set of nested subgenomic mRNAs that have a common 5′ "leader" sequence derived from the 5′ end of the genome, and the subgenomic negative-strand RNAs with common 5′ ends and leader complementary sequences at their 3′ ends (Lai and Holmes, 2001; Thiel et al., 2003). It has been shown that the RNA replication activity takes place at double-membrane vesicles in the host cell cytoplasm (Gosert et al., 2002; Pedersen et al., 1999; Prentice et al., 2004). Given its vital role in viral replication and the success obtained with polymerase inhibitors in the treatment of viral infections, SARS-CoV RdRp is an attractive target for anti-SARS agents. However, currently, our understanding about the biological functions of the RdRps of SARS-CoV and other coronaviruses is very meager because biochemical and structural studies of these enzymes are hampered by the problem of expressing and purifying a soluble and active protein. A modeling study of the polymerase catalytic domain of SARS-CoV RdRp was carried out based on sequence homology between SARS-CoV RdRp and other viral polymerases and has identified the conserved sequence motifs that are likely involved in polymerization and predicted the typical right-hand topology of the polymerase catalytic domain that consists of fingers, palm, and thumb subdomains (Xu et al., 2003). Most recently, immunoblotting and immunofluorescence analyses using an antibody directed against a fragment of SARS-CoV RdRp have detected a single protein with an observed mass of 106 kDa in SARS-CoV infected Vero cells, suggesting that a full-length SARS-CoV RdRp exists in the life cycle of viral replication (Prentice et al., 2004). In addition, an antibody generated against an MHV RdRp fragment can also identify a full-length SARS-CoV RdRp and several small proteins in SARS-CoV infected cell lysates, demonstrating an epitope conservation between MHV RdRp and SARS-CoV RdRp (Prentice et al., 2004).

We report here the cloning, expression, and purification of the full-length SARS-CoV RdRp and its polymerase catalytic domain as glutathione S-transferase (GST) fusion proteins. These recombinant proteins are characterized using Western blot, N-terminal sequencing, mass spectrometry (MS), and in vitro polymerase activity assay. The full-length GST-RdRp exhibits good RdRp activity and weak RNA-dependent DNA polymerase activity. During purification, the full-length enzyme is found to be hydrolytically cleaved into three main fragments: an N-terminal p12 fragment, a middle p30 fragment, and a C-terminal p64 fragment which comprises the polymerase catalytic domain. The cleavage pattern was consistent and was presumably due to bacterial proteases. The p64 and p12 fragments associate together to form a tightly bound complex that possesses a comparable RdRp activity, whereas the GST-p64 protein by itself has no detectable activity, suggesting that the p12 domain is required for the polymerase activity. Activity assays also show that nonnucleoside inhibitors of HIV-1 RT cannot inhibit the polymerase activity of SARS-CoV RdRp, confirming our previous prediction that SARS-CoV RdRp does not contain a hydrophobic pocket near the catalytic active site (Xu et al., 2003). This work provides a basis for further biochemical and structural studies of SARS-CoV RdRp and for development of anti-SARS drugs.

Results

Cloning, expression, and purification of SARS-CoV RdRp

SARS-CoV strain BJ101 was used to infect Vero cells and the total viral RNA was extracted from the infected cells. The cDNA complementary to the coding sequence of SARS-CoV RdRp was obtained using reverse transcription with random primers. Two overlapping DNA fragments (R1, 1428 nucleotides and R2, 1458 nucleotides) that cover the full-length SARS-CoV *RdRp* gene (2796 nucleotides) were cloned separately and then ligated together. The full-length RdRp (residues 1–932) has been successfully expressed in *Escherichia coli.* as a GST-fusion protein with a molecular mass of 132.8 kDa. Most of the recombinant GST-RdRp protein was expressed in inclusion bodies and found at the precipitated pellet of the cell lysate; a small portion remained as soluble form in the supernatant. To increase the quantity of the soluble protein, induction experiments at different IPTG concentrations and temperatures were carried out. Change of the induction conditions had very marginal effect (Fig. 1). High concentration of IPTG appeared to yield a slightly higher expression level of the soluble protein. It was also found that the expression level of GST-RdRp in the Origami (DE3) cells was slightly better than in the BL21 (DE3) (pLysS) cells (data not shown). Therefore, the final protein expression experiments were performed in the Origami (DE3) cells at the induction condition of 1 mM IPTG and 28 °C for 3 h.

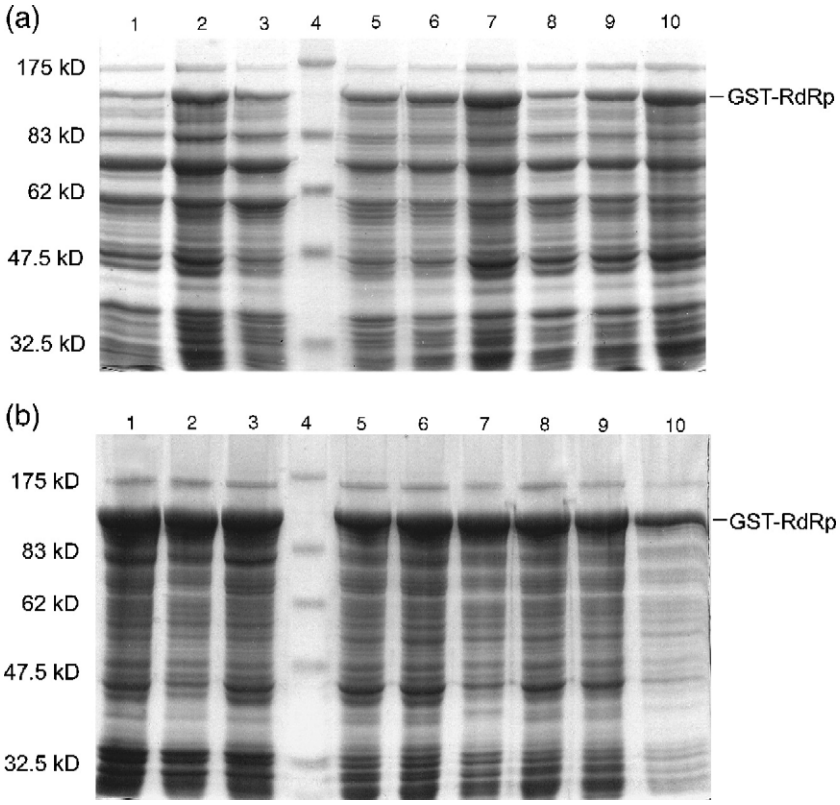

Fig. 1. Expression of SARS-CoV RdRp. (a) Supernatant of the cell lysates. (b) Pellet of the cell lysates. Protein expression was carried out in the Origami (DE3) cells and induced at different IPTG concentrations and temperatures. Lanes 1–3: different IPTG concentrations of 0.1, 0.5, and 1 mM, respectively, at 37 °C for 3 h; lane 4: molecular mass standards; lanes 5–7: different IPTG concentrations of 0.1, 0.5, and 1 mM, respectively, at 28 °C for 3 h; and lanes 8–10: different IPTG concentrations of 0.1, 0.5, and 1 mM, respectively, at 24 °C for 3 h. The protein samples were analyzed in a 15% SDS-PAGE gel stained with Coomassie blue.

The soluble GST-RdRp protein could bind to the glutathione Sepharose 4B column. The purification yield was low because most of the protein precipitated in inclusion bodies and the protein had a poor binding affinity with the column. Optimization of the purification procedure slightly improved the yield which was typically about 0.2 mg GST-RdRp protein with about 80% purity from 1 L of cell culture (Fig. 2a). During purification, it was observed that the full-length GST-RdRp protein was unstable and cleaved gradually when stored at both room temperature and 4 °C. Three fragments with molecular masses of approximately 64 kDa (p64), 39 kDa (p39), and 30 kDa (p30), respectively, are the main products of the cleavage (Fig. 2b). Addition of the protease inhibitor PMSF in the washing buffer did not prevent the cleavage. To characterize the cleavage fragments, the cleavage mixture of the full-length GST-RdRp protein was further purified with the polyA Sepharose 4B column (Fig. 2b). It is very interesting to find that the full-length GST-RdRp protein bound poorly to the polyA column and most of the protein went through the column. On the other hand, majority of the p64 and p39 fragments bound to the column and were co-purified at a molar ratio of about 1:1; only a small portion of them were in the flow-through fractions. The p30 fragment could not bind to the column and was found in the flow-through fractions. The GST could not be completely removed by the polyA column purification due to its abundance and existed as the main impurity in the purified sample. Treatment of the protein sample with thrombin cleaved the p39 fragment into two small fragments with molecular masses of 12 kDa (p12) and 26 kDa, respectively, while the p64 and p30 fragments remained intact. The 26-kDa fragment was shown to be GST by Western blot analysis. Purification of the thrombin processed protein sample with the polyA column showed that the p12 fragment also bound to the column and was co-purified with the p64 fragment in about 1:1 molar ratio (Fig. 3a). These results indicate that the p64 and p12 (or p39) fragments appear to form a complex that binds to polyA and the full-length GST-RdRp has a weaker ability to bind to polyA than the p64/p12 complex. Formation of a stable p64/p12 complex is also supported by results from native PAGE and isoelectric focusing (IEF) gel analyses. Native PAGE of the final protein sample exhibited a single band corresponding to the p64/p12 complex and two bands for GST (Fig. 3b). IEF gel of the protein sample also showed that the p64/p12 complex migrated in a pH gradient as a single band with a pI of about 5.2; GST had dual bands at pI of 5.3 and 5.4 (Fig. 3c).

The C-terminal p64 fragment of SARS-CoV RdRp comprises the polymerase catalytic domain (see results later). Attempts to chromatographically separate the p64 and

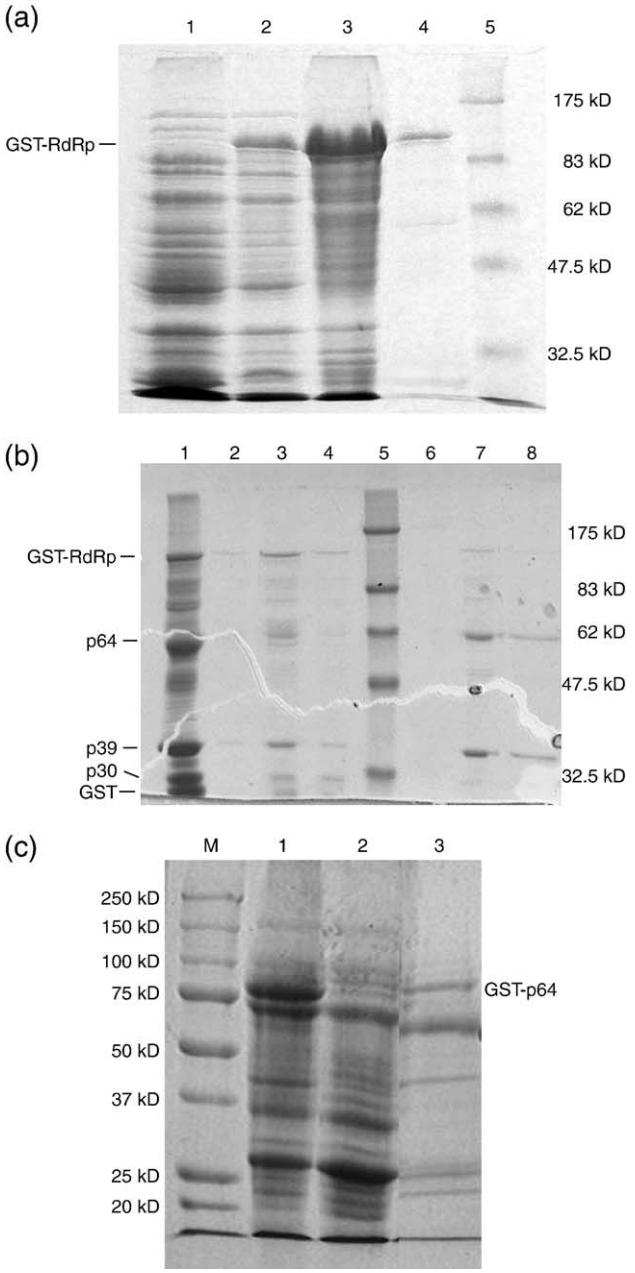

Fig. 2. Purification of SARS-CoV RdRp. (a) Purification of GST-RdRp with the glutathione Sepharose 4B column. Lane 1: whole cell lysate without the expression vector; lane 2: supernatant; lane 3: pellet; lane 4: elution fraction; and lane 5: molecular mass standards. (b) Purification of GST-RdRp with the polyA Sepharose 4B column. Lane 1: the cleavage mixture of the protein sample after purification with the glutathione Sepharose 4B column in 5 days; lanes 2–4: flow-through fractions; lane 5: molecular mass standards; and lanes 6–8: elution fractions. (c) Purification of GST-p64 with the glutathione Sepharose 4B column. M: molecular mass standards; lane 1: pellet; lane 2: flow-through fraction; and lane 3: elution fraction.

p12 fragments were unsuccessful. In order to evaluate the biological property of the catalytic domain, we cloned the recombinant p64 fragment separately that includes 563 residues from the C-terminus of the enzyme (residues 369–932). Initial attempts of cloning this protein fragment (termed *p64*) as a hexahistidine-tagged fusion protein (in the pET22b+ or pET28b+ expression vector) resulted in insoluble protein. Hence, *p64* was subcloned in the pGEX-4T1 expression vector and expressed in *E. coli*. strain BL21

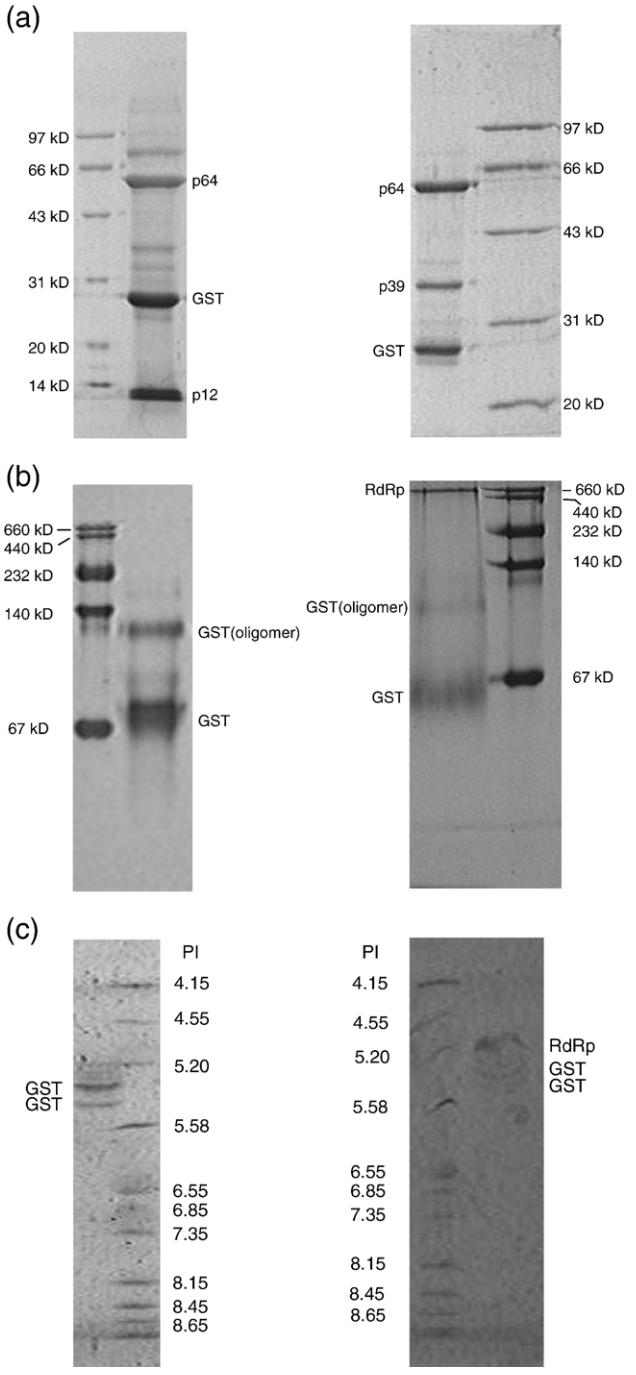

Fig. 3. Characterization of the p64/p12 complex. After purification with the glutathione Sepharose 4B column, the full-length GST-RdRp protein sample was stored at 4 °C for 5 days. One portion was treated with thrombin for the cleavage of GST and another was not treated with thrombin. Both samples were further purified with the polyA Sepharose 4B column. (a) SDS-PAGE of the purified protein samples without (left panel) and with thrombin treatment (right panel). (b) Native PAGE of GST (left panel) and the purified protein sample with thrombin treatment (right panel). (c) IEF gel of GST (left panel) and the purified protein sample with thrombin treatment (right panel).

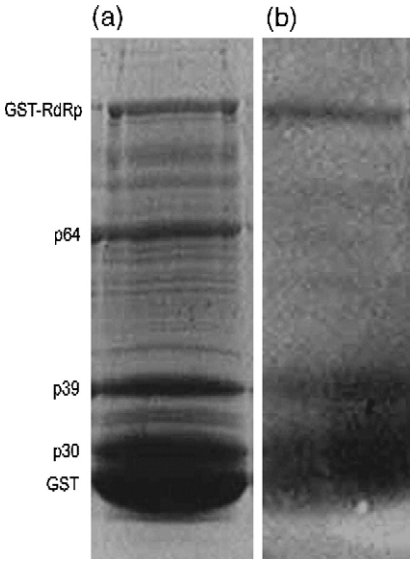

Fig. 4. Western blot analysis of the cleavage mixture of the full-length GST-RdRp. After purification with the glutathione Sepharose 4B column the protein sample was stored at 4 °C for 5 days. (a) SDS-PAGE of the protein sample. (b) Western blot of the protein sample. The gel shows that the full-length GST-RdRp and the p39 fragment contain a GST tag while the p64 fragment does not.

(DE3) (pLysS) as a GST-fusion protein. Similar to the full-length GST-RdRp, most of the expressed GST-p64 protein (molecular mass of 90.5 kDa) was insoluble (Fig. 2c). The GST-p64 protein could not bind to the polyA Sepharose 4B column. Therefore, purification of the GST-p64 protein was carried out only with the glutathione Sepharose 4B column. The purification effect was poor due to the weak binding of the protein with the column and the purified GST-p64 protein was mixed with other small protein impurities. One band with a molecular mass of 64 kD appears to be the GST-cleaved p64 protein. This GST-p64 protein sample was used in the polymerase activity assays.

Characterization of SARS-CoV RdRp

Characterization of the full-length GST-RdRp and of the three cleavage fragments was carried out using Western blot, mass spectrometry, and N-terminal sequencing analyses. Western blot analysis showed that the full-length GST-RdRp and the p39 fragment can be detected by anti-GST antibody, while the p64 fragment cannot (Fig. 4). The p30 fragment could not be detected by anti-GST antibody; however, since this fragment was very close to GST (26 kDa) and the Western blot band corresponding to GST was very broad due to the excessive amount, this result is less certain. These results indicate that both full-length GST-RdRp and the p39 fragment contain GST tag while the p64 fragment does not. It can be deduced that the p39 fragment (and hence the p12 fragment) is located at the N-terminus of SARS-CoV RdRp, while the p64 fragment is located in the middle or C-terminus of the protein. To determine the locations of the p64 and p30 fragments in SARS-CoV RdRp, we carried out MS analyses of these two fragments (data not shown). MS spectra showed that the sequences of the peptides corresponding to the base peaks of the p64 fragment can be mapped to the C-terminal region of SARS-CoV RdRp and those corresponding to the base peaks of the p30 fragment match the middle part of SARS-CoV RdRp, suggesting that p64 is located at the C-terminus and p30 in the middle of the enzyme. However, it is noteworthy that some peptides belonging to GST were also found in the base peaks with relatively low abundance in the MS spectra of the p30 fragment probably due to contamination by the nearby GST in sample preparation. To further identify the cleavage sites of the protein fragments, we performed N-terminal sequencing of the p64 and p30 fragments. The first 5 amino acids of the p64 fragment were determined to be KELLV which gives a calculated molecular mass of 64.2 kDa for p64, consistent with that determined by SDS-PAGE (64 kDa). The N-terminal sequence of the p30 fragment was determined to be VPHIS which gives a calculated mass of 29.9 kDa for p30, in agreement with that determined by SDS-PAGE (30 kDa).

Taken together, we are able to map the location and cleavage site of each fragment in SARS-CoV RdRp (Fig. 5). The cleavage site between the p12 and p30 fragments is at residues M110-V111 and the cleavage site between the p30 and p64 fragments is at residues F368–K369. The full-length SARS-CoV RdRp consists of 932 amino acid residues and has a molecular mass of 106.5 kDa and a calculated pI of 5.995. p12 is located at the N-terminus and consists of residues 1 to 110 (molecular mass 12.4 kDa and calculated pI 5.835). p30 is located in the middle and spans from residues 111 to 368 (molecular mass 29.9 kDa and calculated pI 5.005). p64 is located at the C-terminus and consists of residues 369 to 932 (molecular mass 64.2 kDa and calculated pI 6.635) which comprise the polymerase catalytic domain, as predicted in related molecular modeling studies (Xu et al., 2003).

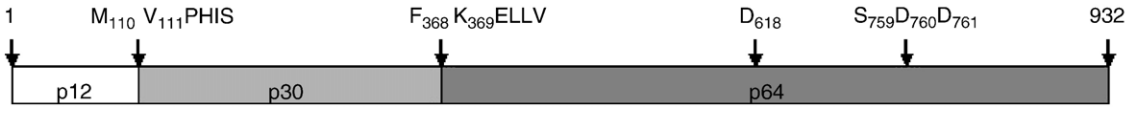

Fig. 5. Schematic diagram showing the cleavage sites and compositions of the full-length SARS-CoV RdRp. SARS-CoV RdRp consists of 932 residues and comprises three domains. The N-terminal p12 domain consists of residues 1 to 110; the middle p30 domain comprises residues 111 to 368, and the C-terminal p64 domain contains residues 369 to 932. The cleavage site between the p12 and p30 domains is at M110–V111 and the cleavage site between the p30 and p64 domains is at F368–K369. p64 comprises the polymerase catalytic domain which contains three strictly conserved aspartates (Asp618, Asp760, and Asp761) that form the polymerase active site.

Polymerase activity of the recombinant SARS-CoV RdRp

The RdRp activity of the full-length SARS-CoV RdRp and of the cleaved fragments was examined using the filter-binding polymerase assay. Specifically, we measured the incorporation of [α-^{32}P] UTP using polyA/oligoU$_{16}$ as the template/primer. The biochemical assays showed that the RdRp activity of the cell lysate supernatant is weak (0.3 pmol/μg/h) (Table 1). The full-length GST-RdRp purified with the glutathione column had an easily measurable RdRp activity (4 pmol/μg/h or 13 folds of that of the lysate supernatant). The full-length GST-RdRp was cleaved gradually into three major fragments, p30, p39, and p64. Proteolytic processing of this cleavage mixture with thrombin led to further proteolysis of p39 into GST and p12. The RdRp activity of this mixture was 7.5 pmol/μg/h or 24 folds of that of the lysate supernatant. Purification of the thrombin-treated protein sample with the polyA column yielded the p64/p12 complex. This complex had a slightly higher RdRp activity (10.9 pmol/μg/h or 35 folds of that of the lysate supernatant). The RdRp activity of the protein samples processed with or without thrombin was not inhibited by either rifampicin (20 μg/ml) or actinomycin D (50 μg/ml), indicating that the measured RdRp activity is not caused by contamination of bacterial RNA polymerase or DNA polymerase. Compared to HCV RdRp, SARS-CoV RdRp has a comparable RdRp activity (Yamashita et al., 1998). Moreover, the full-length GST-RdRp protein also showed a weak RNA-dependent DNA polymerase activity when using polyrA/oligo(dT)$_{12-18}$ as the template-primer (data not shown). However, the recombinant polymerase catalytic domain (GST-p64) by itself had no measurable enzymatic activity in the preliminary experiments (Table 1).

Table 1
Summary of the RdRp activity of SARS-CoV RdRp

	Concentration (μg/ml)	Incorporation (cpm)	Specific activity (pmol/μg/h)
GST-RdRp			
Sonication supernatant	54.6	962	0.3
Glutathione column elution	120.6	27,357	4.0
Thrombin cleaved	149.8	64,519	7.5
PolyA column elution	138.8	86,487	10.9
+Rifampicin (20 μg/ml)	138.8	59,394	7.5
+Actinomycin D (50 μg/ml)	138.8	83,091	10.5
GST-p64			
Glutathione column elution	194.4	369	0.03
Reaction buffer control		284	

Incorporation of [α-^{32}P]UTP into the reaction with polyA/oligoU$_{16}$ as the template/primer. For the GSTP-RdRp, the protein samples were the sonication supernatant, the full-length GST-RdRp after purification with the glutathione Sepharose 4B column, the cleavage mixture of the full-length GST-RdRp that was stored at 4 °C for 5 days and then processed with thrombin for GST cleavage, and the p64/p12 complex after purification with the polyA Sepharose 4B column. For the GST-p64, the protein sample was only purified with the glutathione Sepharose 4B column.

Table 2
Effect of nonnucleoside HIV-1 RT inhibitors on activity of SARS-CoV RdRp

Compound	Incorporation (cpm)	Relative activity (%)
GST-RdRp (46.0 μg/ml)	23,390	100
+α-APA R90384 (10 μM)	14,486	61.9
+HBY 097 (10 μM)	19,396	82.9
Reaction buffer control	310	

Incorporation of [α-^{32}P]UTP into the reaction with polyA/oligoU$_{16}$ as the template/primer. The GST-RdRp protein was purified with the glutathione Sepharose 4B column.

Molecular modeling studies of the polymerase catalytic domain of SARS-CoV RdRp based on sequence comparison of SARS-CoV RdRp with other viral polymerases suggest that SARS-CoV RdRp lacks a hydrophobic pocket near the polymerase active site that is observed in HIV-1 RT and is the binding site of nonnucleoside inhibitors (Xu et al., 2003). To evaluate this prediction, we performed polymerase activity assays of SARS-CoV RdRp in the presence of two potent nonnucleoside HIV-1 RT inhibitors (HBY 097 and α-APA R90384) (Kleim et al., 1999; Miller et al., 1998). The results showed that these inhibitors cannot inhibit the polymerase activity of SARS-CoV RdRp which confirms our prediction based on the modeling study (Table 2).

Discussion

Like other positive-strand RNA viruses, SARS-CoV RdRp is predicted to be part of a replication complex that is responsible for the replication of viral RNA genome. Currently, there is very limited knowledge about the biological function(s) of this enzyme and other coronavirus RdRps because of the lack of a soluble and active enzyme. To pursue functional and structural studies of SARS-CoV RdRp and understand the molecular basis of polymerization and potential drug susceptibility of the enzyme, we have cloned and expressed the full-length SARS-CoV RdRp and a C-terminal fragment that contains the predicted polymerase catalytic domain as fusion proteins with GST at the N-terminus in *E. coli*. The expression and purification of the proteins encountered difficulties. Most of the recombinant full-length GST-RdRp was present in inclusion bodies and a small portion existed as soluble form in supernatant. Furthermore, the purified full-length GST-RdRp was unstable and hydrolytically cleaved to three main fragments, presumably by bacterial proteases. The recombinant p64 domain was even more insoluble than the full-length GST-RdRp with the vast majority of the protein expressed in inclusion bodies. The full-length GST-RdRp was purified to about 80% purity with an RdRp activity comparable to that

of HCV RdRp. However, due to the poor binding affinity of the protein with the glutathione column, the purification yield is low. The p64 protein was purified to a less satisfactory purity and had no measurable polymerase activity. This is the first report of successful expression and purification of a soluble and active recombinant coronavirus polymerase. The quality and quantity of the purified proteins were sufficient good for preliminary biochemical analysis which showed that SARS-CoV RdRp is catalytically active in the absence of host factors. However, further work will be needed to optimize the expression and purification conditions to obtain large quantity of homogeneous protein that is more suitable for extensive biochemical and structural studies.

SARS-CoV RdRp has high contents of hydrophobic residues (43.0% nonpolar and hydrophobic residues, 32.6% of neutral and polar residues, 11.5% acidic residues, and 12.8% alkaline residues) and Cys residue (31 Cys or 3.32% in the full-length RdRp and 16 Cys or 2.83% in the polymerase catalytic domain), rendering it a hydrophobic protein which may contribute in part to its poor solubility. The poor solubility of SARS-CoV RdRp may be correlated with its biological function(s) in infected cells. As a characteristic feature of positive-strand RNA viruses, the replication complex is associated with membranes in the cytoplasm of host cells. Biochemical and biological data have shown that MHV RdRp is associated with membranes during cell lysis, centrifugation, and fractionation and that MHV RdRp not only forms an important part of the replication complex but is also involved in mediating its efficient association with membrane, proteins, and RNA (Brockway et al., 2003; Gosert et al., 2002; Pedersen et al., 1999; Sims et al., 2000). This association would require RdRp to interact with membrane proteins possibly and preferably via hydrophobic interactions. Immunoblotting and immunofluorescence analyses of SARS-CoV infected cell lysates also showed that the replicase proteins of SARS-CoV are co-localized to cytoplasmic complexes containing markers for autophagosome membranes (Prentice et al., 2004). Thus, it is possible that SARS-CoV RdRp may play a similar role in the replication complex and participates in interactions with membrane, proteins, and RNA.

It is very intriguing to observe that during purification the full-length SARS-CoV RdRp was unstable and was hydrolytically cleaved into three fragments, namely p12, p30, and p64. The p64/p12 complex has comparable RdRp activity as the full-length enzyme. The cause of the cleavage is unclear. Whether these cleavages occur in the life cycle of viral replication and whether the observed proteolytic cleavages of the full-length enzyme have any biological implications are also not clear. We explored the possibility of self cleavage; however, sequence comparison of SARS-CoV RdRp with known proteases did not reveal any potential protease domain in SARS-CoV RdRp. At this point, it appears that the observed proteolytic cleavages of the full-length SARS-CoV RdRp are not specific and are likely caused by bacterial proteases, such as thermolysin and subtilisin which have relatively broad cleavage specificity. Since these cleavage sites are not typical proteolytic sites for 3CLpro (usually a conserved Gln at the P1 position) and PL2pro (usually a conserved Gly at the P1 position) of coronaviruses (Anand et al., 2003; Harcourt et al., 2004; Kanjanahaluethai et al., 2003; Yang et al., 2003; Ziebuhr et al., 1995) and there is no conserved residue motif in the cleavage regions that is potential proteolytic site for 3CLpro or PL2Pro, these two viral proteases are less likely to be involved in the cleavages during the viral replication. However, there are some cellular proteases such as elastase, chymotrypsin, and trypsin that have relatively broad substrate specificity. These cellular proteases might play some roles in the cleavage of the full-length enzyme in the virus-infected cells.

A full-length MHV RdRp (100 kDa) was detected as a mature product in MHV infected cells by immonoprecipitation experiments using antiserum against an N-terminal peptide of MHV RdRp (Brockway et al., 2003). Immunoblotting and immunofluorescence analyses of SARS-CoV infected cell lysates using an antibody directed against a SARS-CoV RdRp fragment (which corresponds to approximately residues 326-637 of SARS-CoV RdRp and covers the C-terminal part of the p30 domain and the N-terminal part of the p64 domain) detected a single protein with a molecular mass of 106 kDa, corresponding to the full-length SARS-CoV RdRp, in SARS-CoV infected Vero cells (within 24 h after viral infection) (Prentice et al., 2004). These results suggest that SARS-CoV RdRp appears to exist as the full-length enzyme in the life cycle of viral replication. However, immunoblotting experiments using the antibody directed against the N-terminal of MHV RdRp identified the full-length SARS-CoV RdRp as well as several small proteins in SARS-CoV infected Vero cells (Prentice et al., 2004). Moreover, since the sequences of the cleavage sites are well conserved in RdRps of SARS and group II coronaviruses (including MHV), it is possible that these cleavages may also take place in other group II coronavirus RdRps. Thus, we could not completely exclude the possibility that the cleavage of the full-length SARS-CoV RdRp may also occur in the viral life cycle and the cleaved form of SARS-CoV RdRp may play some functional role in the viral replication. Additional experiments will be required to unequivocally determine the cause of the cleavages and the functional form of the enzyme in the virus infected cells.

The unusual high susceptibility of the full-length SARS-CoV RdRp to proteolysis by bacterial proteases suggests that the enzyme consists of multiple domains connected by flexible regions that contain surface-exposed cleavage sites to bacterial proteases. The flexibility of the enzyme might be reduced when it interacts with other viral proteins and/or factors from the infected cells. Our data suggest that SARS-CoV RdRp appears to consist of three domains. The C-

terminal p64 domain contains the polymerase catalytic domain which is comparable in size to RdRps of other positive-strand RNA viruses and is predicted to consist of the fingers, palm, and thumb subdomains (Xu et al., 2003). The N-terminal region of SARS-CoV and other coronavirus RdRps is significantly large compared to RdRps of other positive-strand RNA viruses and appears to consist of two domains, p12 and p30, that have no equivalents in other viral RdRps with known structures. The N-terminal p12 domain forms a complex with the p64 domain that bound tightly to the polyA column. The p64/p12 complex is catalytically active and has an RdRp activity comparable to that of the full-length RdRp. The recombinant GST-p64 protein did not bind to the polyA column and had no measurable polymerase activity. These results suggest that the N-terminal p12 domain is required for optimal polymerase activity of SARS-CoV RdRp, whereas the p30 domain is dispensable for the catalytic reaction and has an unknown function. Since both p64 and p12 domains contain numerous Cys residues, the tight interaction between these two domains suggests that disulfide bond(s) might be involved in inter-domain interaction in addition to hydrophobic and hydrophilic contacts. The requirement of the p12 domain for the binding of the p64 domain to the polyA column and for the polymerase activity of the enzyme suggests that the p12 domain is likely involved in interaction with template-primer and/or stabilization of the template-primer binding region of the enzyme. The RdRp of reovirus (a double-strand RNA virus) comprises an N-terminal domain (residues 1–380) that is comparable in size to the N-terminal region of SARS-CoV RdRp. Crystal structure of reovirus RdRp indicates that the N-terminal domain flanks on one side of the nucleotide binding cleft and forms part of a channel through which the incoming nucleotide enters into the catalytic active site during polymerization (Tao et al., 2002). The N-terminal region of SARS-CoV RdRp might play a similar functional role as that of reovirus RdRp. In addition, it is also possible that the N-terminal domain(s) of SARS-CoV RdRp may be involved in interactions with either other proteins from the virus itself or host factors from the infected cells that regulate or optimize the replication functions of SARS-CoV RdRp.

Viral polymerases are essential for viral genome replication and are attractive targets for antiviral drug development. Nonnucleoside inhibitors are a class of small organic compounds of hydrophobic nature that have been known to be potent and effective therapeutics with great specificity against HIV-1. Similar inhibitors targeting HCV RdRp are currently under development (Chan et al., 2003; Dhanak et al., 2002; Love et al., 2003). These inhibitors act kinetically in a non-competitive manner with respect to dNTP or rNTP substrates. They either bind to a hydrophobic pocket close to the polymerase active site in HIV-1 RT causing conformational change of the polymerase active site and restricting the flexibility of the nucleotide binding cleft (Das et al., 1996; Ding et al., 1995; Esnouf et al., 1995; Kohlstaedt et al., 1992; Ren et al., 1995), or bind to a hydrophobic pocket on the surface of the thumb subdomain in HCV RdRp, interfering allosterically with conformational change of the thumb (Love et al., 2003). Modeling studies indicate that SARS-CoV RdRp contains neither a hydrophobic pocket near the polymerase active site nor an inhibitor-binding pocket in the thumb subdomain. Thus, it was predicted that nonnucleoside inhibitors which can inhibit HIV-1 or HCV polymerase would not work for SARS-CoV RdRp (Xu et al., 2003). Here, we showed that two potent nonnucleoside HIV-1 RT inhibitors (α-APA R90384 and HBY 097) exhibited no evident inhibitory effect on the SARS-CoV RdRp activity. This information is valuable in the development of anti-SARS drugs. We should exclude these compounds that have similar chemical structures and properties as nonnucleoside HIV-1 RT inhibitors in the drug screening test. Nevertheless, alternative allosteric sites or surface pockets may exist in SARS-CoV RdRp that could be potential targets for antiviral agents. Detailed biochemical and structural studies of the enzyme will likely reveal new inhibitor binding sites. Large-scale drug screening might also identify new inhibitors of SARS-CoV RdRp that could lead to the discovery of potential inhibitor binding sites.

In summary, we present here a simple system for expressing and purifying soluble and active SARS-CoV RdRp in GST-fused form in *E. coli*. The availability of an active recombinant SARS-CoV RdRp protein will not only provide a tool for biochemical and structural studies of the enzyme, but also facilitate efforts of antiviral drug development. In addition, the active enzyme can also be used to prepare both monoclonal and polyclonal antibodies against SARS-CoV.

Materials and methods

Construction of plasmids

Cultured Vero cells were infected with SARS-CoV BJ01 isolate (NCBI accession code: AY278488). The total viral RNA was extracted from the infected cells with TRIZol reagent (Invitrogen). The cDNA complementary to the coding sequence of SARS-CoV RdRp (genome locations 13357...13383 and 13383...16151) was obtained by reverse transcription using random primers. The two overlapping DNA fragments of RdRp (R1 and R2) were amplified using primer sets RdRp1/RdRp1429r and RdRp1338/RdRp2796r, respectively. The primers used are RdRp1: 5′-GGGGCTCGAGCATGTCTGCGGATGCATCAACGTTTTTAAAccGGGTTTGCGGTG -3′ (the lowercase letters indicate the site where the frameshift occurs in the SARS-CoV genome); RdRp1429r: 5′-CAACAACTTCAACTACGAATAGGA-3′; RdRp1338: 5′-GGGGGGATCCAACGCTGCTATCAGTGATTATG-3′; RdRp2796r:

5′-GGGGAAGCTTCTGCAAGACTGTATGTGGTG-TGTA-3′; R11: 5′-GGATCCTCTGCGGATGCATCAACG-3′; and R11r: 5′-CTACAGATAGAGACACCAGCTACG-3′.

R1 was further amplified using primers R11 (containing a *Bam*HI site) and R11r using Pryobest PCR Taq enzyme. R2 was digested by *Eco*RV and *Hin*dIII and then cloned into pBluescript KS(+) (Stratagene) to get pKSR2. The amplified R1 was digested by *Eco*RV and ligated into the *Eco*RV site of pKSR2 to get pKSRdRp. pKSRdRp was then digested by *Bam*HI and *Xho*I and the *RdRp* gene was subcloned into the pGEX-4T1 vector (Amersham Biosciences) to form the pGEX-4T1-RdRp construct. At the 5′-end of the *RdRp* gene, a sequence encoding the GST protein was attached. The open reading frame of the final construct and the encoding of SARS-CoV RdRp (residues 1–932) were confirmed by DNA sequencing.

A 1.7-kb DNA fragment corresponding to the polymerase catalytic domain of SARS-CoV RdRp (p64, residues 369–932) was amplified from the pGEX-4T1-RdRp construct by using primers S1 (5′-ATCGGGATCCAAGGAACTTT-TAGTGTATGCTGC-3′) and S1r (5′-ATCGCTCGAGT-CACTGCAAGACTGTATGTGGTGT-3′). S1 contains a *Bam*HI site and S1r contains an *Xho*I site. The DNA fragment was digested by these two enzymes and the *p64* gene was subcloned into the pGEX-4T1 vector to form the pGEX-4T1-p64 construct. The sequence encoding the GST tag was inserted at the 5′-end of the *p64* gene. The pGEX-4T1-p64 construct was also verified by DNA sequencing.

Expression and purification of the full-length SARS-CoV RdRp

The pGEX-4T1-RdRp plasmid was transformed in two *E. coli* strains: Origami (DE3) (Novagen) and BL21 (DE3) (pLysS) (Novagen). A single colony was grown at 37 °C for 12 h in 15 mL of 2× YT medium supplemented with ampicillin (0.1 mg/ml). The cell culture was further grown at 37 °C in 1 L of 2 × YTG medium containing ampicillin (0.1 mg/ml) until OD_{600} reached 0.8–1.0. Protein expression was induced with 0.5 mM IPTG at 28 °C for 3 h and the cells were harvested by centrifugation at 7000 × *g* and 4 °C for 20 min and washed with a PBS buffer.

Protein purification was carried out using affinity chromatography with the glutathione Sepharose 4B and polyA Sepharose 4B columns (Amersham Biosciences). The cell pellet was resuspended in buffer A (4.3 mM Na_2HPO_4, 1.5 mM KH_2PO_4, pH 7.3, 0.14 M NaCl, and 2.7 mM KCl) containing 1 mM PMSF and lysed on ice by sonication. The lysate was centrifuged at 14,000 × *g* and 4 °C for 30 min. After being filtered through a 0.45 μm micron membrane, the supernatant was loaded onto a glutathione Sepharose 4B column equilibrated with the washing buffer. The column was washed with the buffer until no protein was detected in the flow-through solution. The bound protein was then eluted with 4 column volumes of buffer B (10 mM reduced glutathione, 50 mM Tris–HCl, and pH 8.0). The elution fractions were pooled and dialyzed against buffer C (20 mM Tris–HCl, pH 8.0, 10% glycerol, 1 mM EDTA, and 1 mM DTT) for 24 h.

The purified full-length GST-RdRp was found to be cleaved gradually into three main fragments with apparent molecular masses of 30 kDa (p30), 39 kDa (p39), and 64 kDa (p64), respectively. Later experiments showed that the p39 fragment can be further proteolyzed using thrombin into two small fragments: a 26-kDa GST and a 12-kDa fragment (p12). To separate and characterize these cleavage fragments, the protein solution was applied onto a polyA Sepharose 4B column equilibrated with buffer C. After washing with buffer C several times, the target proteins were eluted with a linear gradient of 0–1 M NaCl in buffer C. SDS-PAGE analyses were performed to check the purity and quality of the protein samples at every stage of purification. The purified proteins were characterized by Western blot, mass spectrometry, N-terminal sequencing, and polymerase activity assay.

Expression and purification of the catalytic domain of SARS-CoV RdRp

The expression and purification of the polymerase catalytic domain (p64) were similar to those for the full-length RdRp. The pGEX-4T1-p64 plasmid was expressed in *E. coli* strain BL21 (DE3) (pLysS). A single colony was grown overnight at 30 °C in 20 mL of 2 × YTG medium in the presence of ampicillin and chloramphenicol (0.1 mg/ml). The cell culture was used to inoculate 1 L of 2 × YTG medium containing ampicillin (0.1 mg/ml) and grown to an OD_{600} of 1.0 at 30 °C. Protein expression was induced overnight with 0.1 mM IPTG at 20 °C. The cells were harvested by centrifugation at 6000 × *g* and 4 °C for 10 min and washed with a PBS buffer.

The cell pellet was resuspended in buffer A containing 1 mM PMSF and lysed on ice by sonication. The lysate was centrifuged at 18,000 × *g* and 4 °C for 30 min. The supernatant was immediately applied onto a glutathione Sepharose 4B column equilibrated with buffer A and the column was washed with the same buffer until no protein was detected in the flow-through solution. The bound proteins were eluted with 4 column volumes of buffer B and the elution fractions were pooled and dialyzed against buffer C for 24 h. Further purification with a polyA Sepharose 4B column was unsuccessful because the p64 protein could not bind to the column. Therefore, the protein sample purified with the glutathione Sepharose 4B column was used in the polymerase activity assays.

Western blot analysis

During purification, the full-length GST-RdRp was found to be proteolytically cleaved into three main fragments, p30, p39, and p64. To identify which fragment is located at the N-terminus of GST-RdRp, we performed Western blot

analysis using anti-GST antibody. The protein sample purified with the glutathione Sepharose 4B column was stored at 4 °C for 5 days to allow proteolysis of sufficient portion of GST-RdRp. The cleaved protein fragments were separated with a 12% SDS-PAGE gel and transferred into a PVDF film (0.2 μm, Bio-Rad) using Bio-Rad Mini Trans-Blot Cell. The PVDF film was washed with TTBS, blocked with 5% milk-TTBS, and further washed with TTBS. The PVDF film was then incubated with an anti-GST antibody (Amersham Pharmacia, 1:200 dilution) at 4 °C for overnight. After another washing with TTBS, it was incubated with mouse anti-rabbit antibody coupled with AP (Amersham Pharmacia, 1:5000 dilution) at room temperature for 60 min and then washed again with TTBS. Bands were detected with a BCIP/NBT coloration kit (SABC). The host strain BL21 (DE3) (pLysS) was used as a negative control.

Mass spectrometry analysis

To identify the locations of the p64 and p30 fragments in the full-length GST-RdRp and their molecular masses, we carried out mass spectrometry analyses of these fragments. After purification with the glutathione Sepharose 4B column, the protein sample was allowed to store at 4 °C for 5 days and then separated with SDS-PAGE. The stained protein bands corresponding to the p64 and p30 fragments were excised from the gel and digested with trypsin (Zeng et al., 2003). Specifically, the gel slices were destained in a solution containing 30% acetonitrile and 100 mM NH_4HCO_3 for 20 min and then dried in vacuum. 10 μL of 40 mM NH_4HCO_3 containing trypsin (the ratio of sample to trypsin is about 40) (sequencing grade, Promega) was added in the gel and incubated at 4 °C for 1 h, and then additional 40 mM NH_4HCO_3 solution was added to cover gel pieces. Tube was sealed and incubated at 37 °C for 22 h. These samples were further prepared using a reverse phase HPLC and a C18 column (120 μm × 150 mm, Thermo Hypersil-Keystone) on a surveyor LC system (Thermo Finnigan). MS data were acquired on an LTQ linear ion trap mass spectrometer (Thermo Finnigan) equipped with an electrospray interface and operated in positive ion mode. The acquired MS spectra were automatically searched against protein databases for SARS-CoV, *Schistosoma*, and *E. coli* using the TurboSEQUEST program in the BioWorks™ 3.0 software suite. An accepted SEQUEST result had a ΔCn score of at least 0.1 (regardless of charge state). Peptides with a +1 charge state were accepted if they were fully tryptic and had a cross correlation (*X*corr) >1.9. Peptides with a +2 charge state were accepted if they had an *X*corr >2.5 and peptides with a +3 charge state were accepted if they had an *X*corr >3.7.

N-terminal sequencing

To determine the cleavage sites, we performed amino acid sequencing of the N-termini of the p64 and p30 fragments. After purification with a glutathione Sepharose 4B column the protein sample was stored at 4 °C for 5 days. To prepare the p30 fragment, a fraction of the protein mixture was separated with SDS-PAGE and the protein band corresponding to p30 was cut from the gel for further analysis. To prepare the p64 sample, a fraction of the protein mixture was first subjected to thrombin cleavage and then purified with the polyA Sepharose 4B column. After separation with SDS-PAGE, the protein band corresponding to p64 was cut from the gel for further analysis. These gel pieces were collected and purified using the electro-elution method (Model 422 Electro-Eluter, Bio-Rad). The protein sample was transferred to a PVDF film through Prosorb and washed with 1% TFA and H_2O, and the film was then dried at 40–60 °C. The N-terminal five amino acids were analyzed using an ABI 491A protein sequencer (Procise, Applied Biosystems).

RNA polymerase activity assays

The RdRp activity of the protein samples was examined using the filter-binding polymerase assay modified from that used in the polymerase activity assay of hepatitis C virus (HCV) RdRp (Ronald et al., 1998; Yamashita et al., 1998). Specifically, a total volume of 50 μl reaction mixture was prepared that contains 50 mM Tris–HCl, pH 8.0, 7.5 mM KCl, 8 mM $MgCl_2$, 10 mM DTT, 1% BSA, 3.5 μL of 1 mM UTP, 3.33 μCi of [α-^{32}P] UTP (Amersham Pharmacia), 3.125 μg/ml polyA (Fluka), 1 μg/ml oligoU_{16} (Takara), 20 units of RNAase inhibitor (SABC), and 50 μg/ml actinomycin D. The mixture was incubated at 37 °C for 1 h and the reaction was stopped by adding a cold buffer solution containing 20 mM sodium pyrophosphate and 5% trichloroacetic acid. 30 μl of the reaction solution was dotted on the GF/C glass microfibre filters (Whatman) and washed with the cold buffer 5 times and then with 75% ethanol once. The incorporated triphosphate was assayed by measuring ^{32}P using a liquid scintillation counter (LS6000, Beckman). To evaluate whether SARS-CoV RdRp has an RNA-dependent DNA polymerase activity, a reverse transcriptase activity assay was also carried out in which a polyrA/oligo$(dT)_{12-18}$ was used as template-primer and dTTP and [α-^{32}P]TTP (Amersham Pharmacia) were used as substrates, respectively. MMLV RT (Promega) was used as a standard reverse transcriptase for comparison.

Acknowledgments

We thank Qin Huang, Baozhen Peng, and Rong Zeng for assistance of the experiments. This work was supported by the NSFC grants (30125011, 30170223, and 30130080), the MOST grants (2002BA711A13, 2004AA235091, 2004CB720102, and 2004CB520801), and the CAS grant (KSCX1-SW-17).

References

Anand, K., Ziebuhr, J., Wadhwani, P., Mesters, J.R., Hilgenfeld, R., 2003. Coronavirus main proteinase (3CLpro) structure: basis for Design of anti-SARS drugs. Science 300, 1763–1767.

Bost, A.G., Carnahan, R.H., Lu, X.T., Denison, M.R., 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.

Brockway, S.M., Clay, C.T., Lu, X.T., Denison, M.R., 2003. Characterization of the expression, intracellular localization, and replication complex association of the putative mouse hepatitis virus RNA-dependent RNA polymerase. J. Virol. 77 (19), 10515–10527.

Chan, L., Reddy, T.J., Proulx, M., Das, S.K., Pereira, O., Wang, W., Siddiqui, A., Yannopoulos, C.G., Poisson, C., Turcotte, N., Drouin, A., Alaoui-Ismaili, M.H., Bethell, R., Hamel, M., L'Heureux, L., Bilimoria, D., Nguyen-Ba, N., 2003. Identification of N,N-disubstituted phenylalanines as a novel class of inhibitors of hepatitis C NS5B polymerase. J. Med. Chem. 46, 9489–9495.

Das, K., Ding, J., Hsiou, Y., Clark Jr., A.D., Moereels, H., Koymans, L., Andries, K., Pauwels, R., Janssen, P.A.J., Boyer, P.L., Clark, P., Smith Jr., R.H., Kroeger Smith, M.B., Michejda, C.J., Hughes, S.H., Arnold, E., 1996. Crystal structures of 8-Cl and 9-Cl TIBO complexed with wild-type HIV-1 RT and 8-Cl TIBO complexed with the Tyr181Cys HIV-1 RT drug-resistant mutant. J. Mol. Biol. 264, 1085–1100.

Dhanak, D., Duffy, K.J., Johnston, V.K., Lin-Goerke, J., Darcy, M., Shaw, A.N., Gu, B., Silverman, C., Gates, A.T., Nonnemacher, M.R., Earnshaw, D.L., Casper, D.J., Kaura, A., Baker, A., Greenwood, C., Gutshall, L.L., Maley, D., DelVecchio, A., Macarron, R., Hofmann, G.A., Alnoah, Z., Cheng, H.Y., Chan, G., Khandekar, S., Keenan, R.M., Sarisky, R.T., 2002. Identification and biological characterization of heterocyclic inhibitors of the hepatitis C virus RNA-dependent RNA polymerase. J. Biol. Chem. 277, 38322–38327.

Ding, J., Das, K., Tantillo, C., Zhang, W., Clark Jr., A.D., Jessen, S., Lu, X., Hsiou, Y., Jacobo-Molina, A., Andries, K., Pauwels, R., Moereels, H., Koymans, L., Janssen, P.A.J., Smith Jr., R.H., Koepke, M.K., Michejda, C., Hughes, S.H., Arnold, E., 1995. Structure of HIV-1 reverse transcriptase in a complex with the nonnucleoside inhibitor a-APA R 95845 at 2.8A resolution. Structure 3, 365–379.

Drosten, C., Gunther, S., Preiser, W., van der Werf, S., Brodt, H., Becker, S., Rabenau, H., Panning, M., Kolesnikova, L., Fouchier, R.A.M., Berger, A., Burguiere, A.M., Cinatl, J., Eickmann, M., Escriou, N., Grywna, K., Kramme, S., Manuguerra, J.C., Muller, S., Rickerts, V., Sturmer, M., Vieth, S., Klenk, H.D., Osterhaus, A.D.M.E., Schmitz, H., Doerr, H.W., 2003. Identification of a novel coronavirus in patients with severe acute respiratory syndrome. N. Engl. J. Med. 348, 1967–1976.

Esnouf, R., Ren, J., Ross, R., Jones, Y., Stammers, D., Stuart, D., 1995. Mechanism of inhibition of HIV-1 reverse transcriptase by non-nucleoside inhibitors. Nat. Struct. Biol. 2, 303–308.

Gosert, R., Kanjanahaluethai, A., Egger, D., Bienz, K., Baker, S.C., 2002. RNA replication of mouse hepatitis virus takes place at double-membrane vesicles. J. Virol. 76 (8), 3697–3708.

Harcourt, B.H., Jukneliene, D., Kanjanahaluethai, A., Bechill, J., Severson, K.M., Smith, C.M., Rota, P.A., Baker, S.C., 2004. Identification of severe acute respiratory syndrome coronavirus replicase products and characterization of papain-like protease activity. J. Virol. 78, 13600–13612.

Kanjanahaluethai, A., Jukneliene, D., Baker, A., 2003. Identification of the murine coronavirus MP1 cleavage site recognized by papain-like proteinase 2. J. Virol. 77, 7376–7382.

Kleim, J.-P., Winters, M., Dunkler, A., Suarez, J.-R., Riess, G., Winkler, R., Balzarini, J., Oette, D., Merigan, T.C., 1999. Antiviral activity of the human immunodeficiency virus type 1-specific nonnucleoside reverse transcriptase inhibitor HBY 097 alone and in combination with Zidovudine in a phase II study. J. Infect. Dis. 179, 709–713.

Kohlstaedt, L.A., Wang, J., Friedman, J.M., Rice, P.A., Steitz, T.A., 1992. Crystal structure at 3.5 A resolution of HIV-1 reverse transcriptase complexed with an inhibitor. Science 256, 1783–1790.

Ksiazek, T.G., Erdman, D., Goldsmith, C.S., Zaki, S.R., Peret, T., Emery, S., Tong, S., Urbani, C., Comer, J.A., Lim, W., Rollin, P.E., Dowell, S.F., Ling, A.-E., Humphrey, C.D., Shieh, W.-J., Guarner, J., Paddock, C.D., Rota, P., Fields, B., DeRisi, J., Yang, J.-Y., Cox, N., Hughes, J.M., LeDuc, J.W., Bellini, W.J., Anderson, L.J., 2003. A novel coronavirus associated with severe acute respiratory syndrome. N. Engl. J. Med. 348, 1953–1966.

Lai, M.M.C., Holmes, K.V., 2001. Coronaviridae: the viruses and their replication. In: Knipe, D.M., Howley, P.M. (Eds.), Fields Virology. Lippincott Williams and Wilkins, Philadelphia, pp. 1163–1185.

Love, R.A., Parge, H.E., Yu, X., Hickey, M.J., Diehl, W., Gao, J., Wriggers, H., Ekker, A., Wang, L., Thomson, J.A., Dragovich, P.S., Fuhrman, S.A., 2003. Crystallographic identification of a noncompetitive inhibitor binding site on the hepatitis C virus NS5B RNA polymerase enzyme. J. Virol. 77, 7575–7581.

Marra, M.A., Jones, S.J.M., Astell, C.R., Holt, R.A., Brooks-Wilson, A., Butterfield, Y.S.N., Khattra, J., Asano, J.K., Barber, S.A., Chan, S.Y., Coughlin, S.M., Freeman, D., Girn, N., Griffith, O.L., Leach, S.R., Mayo, M., McDonald, H., Montgomery, S.B., Pandoh, P.K., Petrescu, A.S., Robertson, A.G., Schein, J.E., Siddiqui, A., Smailus, D.E., Stott, J.M., Yang, G.S., Plummer, F., Andonov, A., Artsob, H., Bastien, N., Bernard, K., Booth, T.F., 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, G.A., Tyler, S., Vogrig, R., Ward, D., Watson, B., Brunham, R.C., Krajden, M., Petric, M., Skowronski, D.M., Upton, C., Roper, R.L., 2003. The genome sequence of the SARS-associated coronavirus. Science 300, 1399–1404.

Miller, V., De Bethune, M.P., kober, A., Aturmer, M., Hertogs, K., Pauwels, R., Stoffels, P., Staszewski, S., 1998. Patterns of resistance and cross-resistance to human immunodeficiency virus type 1 reverse transcriptase inhibitors in patients treated with the nonnucleoside reverse transcriptase inhibitor loviride. Antimicrob. Agents Chemother. 42, 3123–3129.

Navas-Martin, S., Weiss, S., 2003. SARS: lessons learned from other coronaviruses. Viral Immunol. 16, 461–474.

Pedersen, K.W., van der Meer, Y., Roos, N., Snijder, E.J., 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.

Peiris, J.S.M., Lai, S.T., Poon, L.L.M., Guan, Y., Yam, L.Y.C., Lim, W., Nicholls, J., Yee, W.K.S., Yan, W.W., Cheung, M.T., Cheng, V.C.C., Chan, K.H., Tsang, D.N.C., Yung, R.W.H., Ng, T.K., Yuen, K.Y., 2003. Coronavirus as a possible cause of severe acute respiratory syndrome. Lancet 361, 1319–1325.

Poutanen, S.M., Low, D.E., Henry, B., Finkelstein, S., Rose, D., Green, K., Tellier, R., Draker, R., Adachi, D., Ayers, M., Chan, A.K., Skowronski, D.M., Salit, I., Simor, A.E., Slutsky, A.S., Doyle, P.W., Krajden, M., Petric, M., Brunham, R.C., McGeer, A.J., 2003. Identification of severe acute respiratory syndrome in Canada. N. Engl. J. Med. 348 (20), 1995–2005.

Prentice, E., McAuliffe, J., Lu, X., Subbarao, K., Denison, M.R., 2004. Identification and characterization of severe acute respiratory syndrome coronavirus replicase proteins. J. Virol. 78, 9977–9986.

Ren, J., Esnouf, R., Garman, E., Somers, D., Ross, C., Kirby, I., Keeling, J., Darby, G., Jones, Y., Stuart, D., Stammers, D., 1995. High resolution structures of HIV-1 RT from four RT-inhibitor complexes. Nat. Struct. Biol. 2 (4), 293–302.

Ronald, H.A., Xie, Y.P., Wang, Y.H., Hagedorn, C.H., 1998. Expression of recombinant hepatitis C virus non-structural protein 5B in *Escherichia coli*. Virus Res. 53, 141–149.

Rota, P.A., Oberste, M.S., Monroe, S.S., Nix, W.A., Campagnoli, R., Icenogle, J.P., Penaranda, S., Bankamp, B., Maher, K., Chen, M., Tong, S., Tamin, A., Lowe, L., Frace, M., DeRisi, J.L., Chen, Q., Wang, D., Erdman, D.D., Peret, T.C.T., Burns, C., Ksiazek, T.G., Rollin, P.E.,

Sanchez, A., Liffick, S., Holloway, B., Limor, J., McCaustland, K., Olsen-Rasmussen, M., Fouchier, R., Gunther, S., Osterhaus, A.D., Drosten, C., Pallansch, M.A., Anderson, L.J., Bellini, W.J., 2003. Characterization of a novel coronavirus associated with severe acute respiratory syndrome. Science 300, 1394–1399.

Sims, A.C., Ostermann, J., Denison, M.R., 2000. Mouse hepatitis virus replicase proteins associate with two distinct populations of intracellular membranes. J. Virol. 74 (12), 5647–5654.

Snijder, E.J., Bredenbeek, P.J., Dobbe, J.C., Thiel, V., Ziebuhr, J., Poon, L.L., Guan, Y., Rozanov, M., Spaan, W.J., Gorbalenya, A.E., 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.

Tao, Y., Farsetta, D.L., Nibert, M.L., Harrison, S.C., 2002. RNA synthesis in a cage—structural studies of reovirus polymerase λ3. Cell 111, 733–745.

Thiel, V., Ivanov, K.A., Putics, A., T. Hertzig, T., Schelle, B., Bayer, S., Weissbrich, B., Snijder, E.J., Rabenau, H., Doerr, H.W., Gorbalenya, A.E., Ziebuhr, J., 2003. Mechanisms and enzymes involved in SARS coronavirus genome expression. J. Gen. Virol. 84, 2305–2315.

Tsang, K.W., Ho, P.L., Ooi, G.C., Yee, W.K., Wang, T., Chan-Yeung, M., Lam, W.K., Seto, W.H., Yam, L.Y., Cheung, T.M., Wong, P.C., Lam, B., Ip, M.S., Chan, J., Yuen, K.Y., Lai, K.N., 2003. A cluster of cases of severe acute respiratory syndrome in Hong Kong. N. Engl. J. Med. 348, 1977–1985.

Xu, X., Liu, Y., Weiss, S., Arnold, E., Sarafianos, S.G., Ding, J., 2003. Molecular model of SARS coronavirus polymerase: implications for biochemical function and drug design. Nucleic Acids Res. 31 (24), 7117–7130.

Yamashita, T., Kaneko, S., Shirota, Y., Qin, W.P., Nomura, T., Kobayashi, K., Murakami, S., 1998. RNA-dependent RNA polymerase activity of the soluble recombinant hepatitis C virus NS5B protein truncated at the C-terminal region. J. Biol. Chem. 273, 15479–15486.

Yang, H., Yang, M., Ding, Y., Liu, Y., Lou, Z., Zhou, Z., Sun, L., Mo, L., Ye, S., Pang, H., Gao, G.F., Anand, K., Bartlam, M., Hilgenfeld, R., Rao, Z., 2003. The crystal structures of severe acute respiratory syndrome virus main protease and its complex with an inhibitor. Proc. Natl. Acad. Sci. U.S.A. 100, 13190–13195.

Zeng, R., Chen, Y.B., Shao, X.X., Miller, K., Tran, H., Xia, Q.C., 2003. Identification of proteins separated by one-dimensional sodium dodecyl sulfate/polyacrylamide gel electrophoresis with matrix-assisted laser desorption/ionization ion trap mass spectrometry; comparison with matrix-assisted laser desorption/ionization time-of-flight mass fingerprinting. Rapid Commun. Mass Spectrom. 17, 1995–2004.

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