The SARS-Unique Domain (SUD) of SARS Coronavirus Contains Two Macrodomains That Bind G-Quadruplexes

Jinzhi Tan[1], Clemens Vonrhein[2], Oliver S. Smart[2], Gerard Bricogne[2], Michela Bollati[1¶], Yuri Kusov[1], Guido Hansen[1], Jeroen R. Mesters[1], Christian L. Schmidt[1], Rolf Hilgenfeld[1,3]*

1 Institute of Biochemistry, Center for Structural and Cell Biology in Medicine, University of Lübeck, Lübeck, Germany, 2 Global Phasing Ltd., Sheraton House, Castle Park, Cambridge, United Kingdom, 3 Laboratory for Structural Biology of Infection and Inflammation, c/o DESY, Hamburg, Germany

Abstract

Since the outbreak of severe acute respiratory syndrome (SARS) in 2003, the three-dimensional structures of several of the replicase/transcriptase components of SARS coronavirus (SARS-CoV), the non-structural proteins (Nsps), have been determined. However, within the large Nsp3 (1922 amino-acid residues), the structure and function of the so-called SARS-unique domain (SUD) have remained elusive. SUD occurs only in SARS-CoV and the highly related viruses found in certain bats, but is absent from all other coronaviruses. Therefore, it has been speculated that it may be involved in the extreme pathogenicity of SARS-CoV, compared to other coronaviruses, most of which cause only mild infections in humans. In order to help elucidate the function of the SUD, we have determined crystal structures of fragment 389–652 ("SUD$_{core}$") of Nsp3, which comprises 264 of the 338 residues of the domain. Both the monoclinic and triclinic crystal forms (2.2 and 2.8 Å resolution, respectively) revealed that SUD$_{core}$ forms a homodimer. Each monomer consists of two subdomains, SUD-N and SUD-M, with a macrodomain fold similar to the SARS-CoV X-domain. However, in contrast to the latter, SUD fails to bind ADP-ribose, as determined by zone-interference gel electrophoresis. Instead, the entire SUD$_{core}$ as well as its individual subdomains interact with oligonucleotides known to form G-quadruplexes. This includes oligodeoxy- as well as oligoribonucleotides. Mutations of selected lysine residues on the surface of the SUD-N subdomain lead to reduction of G-quadruplex binding, whereas mutations in the SUD-M subdomain abolish it. As there is no evidence for Nsp3 entering the nucleus of the host cell, the SARS-CoV genomic RNA or host-cell mRNA containing long G-stretches may be targets of SUD. The SARS-CoV genome is devoid of G-stretches longer than 5–6 nucleotides, but more extended G-stretches are found in the 3′-nontranslated regions of mRNAs coding for certain host-cell proteins involved in apoptosis or signal transduction, and have been shown to bind to SUD in vitro. Therefore, SUD may be involved in controlling the host cell's response to the viral infection. Possible interference with poly(ADP-ribose) polymerase-like domains is also discussed.

Introduction

The SARS coronavirus (SARS-CoV) is much more pathogenic for humans than any other coronavirus. Therefore, protein domains encoded by the SARS-CoV genome that are absent in other coronaviruses are of particular interest, because they may be responsible for the extraordinary virulence. The most prominent such domain has been identified by bioinformatics as part of non-structural protein 3 (Nsp3) of the virus and appropriately named the "SARS-unique domain" (SUD) [1]. With a molecular mass of 213 kDa, Nsp3 is the largest of the non-structural proteins of SARS coronavirus (see Figure 1). Comprising 1922 amino-acid residues (polyprotein 1a/1ab residues Ala819 to Gly2740), SARS-CoV Nsp3 is larger than the entire replicase of Picornaviridae. It contains at least seven subdomains [2]: An N-terminal acidic domain (Ac, also called Nsp3a); an X-domain (also designated as

ADRP, or Nsp3b); the SUD (Nsp3c); a papain-like proteinase, PL2^{pro} (also called Nsp3d); and additional domains (Nsp3e–g) that include a transmembrane (TM) region.

At present, it is completely unclear whether and how the individual domains of Nsp3 interact with one another or with other components of the coronaviral replicase complex. Also, some of them possibly recognize proteins of the infected host cell [2]. In the absence of functional data on these domains, attempts have been made to derive their possible biological role from their three-dimensional structures (see [3] for a review). The NMR structure of an N-terminal fragment of the acidic domain (Nsp3a) has revealed a ubiquitin-like fold complemented by two additional short α-helices ([4], PDB code 2IDY). NMR chemical-shift analysis suggested that these non-canonical structural elements might bind single-stranded RNA with some specificity for AUA-containing sequences, although the K$_D$ values observed are

Author Summary

The genome of the SARS coronavirus codes for 16 non-structural proteins that are involved in replicating this huge RNA (approximately 29 kilobases). The roles of many of these in replication (and/or transcription) are unknown. We attempt to derive conclusions concerning the possible functions of these proteins from their three-dimensional structures, which we determine by X-ray crystallography. Non-structural protein 3 contains at least seven different functional modules within its 1922-amino-acid polypeptide chain. One of these is the so-called SARS-unique domain, a stretch of about 338 residues that is completely absent from any other coronavirus. It may thus be responsible for the extraordinarily high pathogenicity of the SARS coronavirus, compared to other viruses of this family. We describe here the three-dimensional structure of the SARS-unique domain and show that it consists of two modules with a known fold, the so-called macrodomain. Furthermore, we demonstrate that these domains bind unusual nucleic-acid structures formed by consecutive guanosine nucleotides, where four strands of nucleic acid are forming a superhelix (so-called G-quadruplexes). SUD may be involved in binding to viral or host-cell RNA bearing this peculiar structure and thereby regulate viral replication or fight the immune response of the infected host cell.

relatively high ($\sim$20 µM). Interestingly, a second ubiquitin-like domain occurs in Nsp3, as part of the papain-like proteinase (PL2^{pro}, Nsp3d, [5]; PDB code 2FE8). The PL2^{pro} cleaves the viral polyprotein after two consecutive glycine residues to release Nsp1, Nsp2, and Nsp3, respectively (The remaining cleavage reactions are performed by the coronaviral main proteinase (M^{pro}; [6–8])). In addition to its proteolytic activities on the N-terminal third of the polyproteins, the SARS-CoV PL2^{pro} has also been shown to be a deubiquitinating enzyme [9–12]. Lindner *et al.* [13] have shown that in addition to its proteolytic and deubiquitinating activity, the SARS-CoV PL2^{pro} acts as a de-ISGylating enzyme. Induction of ISG15 and its subsequent conjugation to proteins protects cells from the effects of viral infection [14,15]. Since the ISG15 gene is induced by interferon as part of the antiviral response of the innate immune system, the de-ISGylation activity of Nsp3d could explain

the suppression of the interferon response by the papain-like protease, in addition to a possible direct interaction between the PL2^{pro} and IRF3 [16].

Among the subdomains of the Nsp3 multidomain protein, there is also the so-called "X- domain" (Nsp3b), which shows structural homology to macrodomains. The latter name refers to the non-histone-like domain of the histone macro2A [17–19]. In animal cells, such domains are occasionally physically associated with enzymes involved in ADP-ribosylation or ADP-ribose metabolism. Because of this linkage and on the basis of sequence similarity to Poa1p, a yeast protein involved in the removal of the 1″-phosphate group from ADP-ribose 1″-phosphate (a late step in tRNA splicing; [20]), it has been proposed that the coronaviral X-domains may have the function of ADP-ribose-1″-phosphatases (ADRPs; [21]). The crystal structures of X-domains of SARS-CoV [22,23] as well as of HCoV 229E and Infectious Bronchitis Virus (IBV) [24] show that the protein has the three-layer $\alpha/\beta/\alpha$ fold characteristic of the macrodomains.

Embedded between the X-domain (Nsp3b) and the PL2^{pro} (Nsp3d), the SARS-unique domain (SUD; Nsp3c) fails to show sequence relationship to any other protein in the databases [1]. We have produced full-length SUD (residues 389 to 726 of Nsp3), and a more stable, shortened 264-residue version (residues 389 to 652; henceforth called SUD$_{core}$), by expression in *Escherichia coli*. This definition of the boundaries of the SUD is based on the structural results described here. We report crystallization of SUD$_{core}$ and its X-ray structure in two crystal forms, at 2.2 and 2.8 Å resolution, respectively. The structure turns out to consist of two further copies of the macrodomain, in spite of the complete absence of sequence similarity. In addition, we demonstrate that each of the subdomains binds G-quadruplexes, both in DNA and RNA fragments, and that selected mutations of lysine residues in the first subdomain, SUD-N, lead to reduced nucleic-acid binding, whereas those in the second subdomain, SUD-M, abolish it.

Results

Quality of the structural models

Out of the many SUD constructs designed and tested by us, SUD$_{core}$ (Nsp3 residues 389–652) turned out to be relatively stable and could be crystallized (Table 1). Two crystal forms were observed under identical crystallization conditions: Form-1 crystals
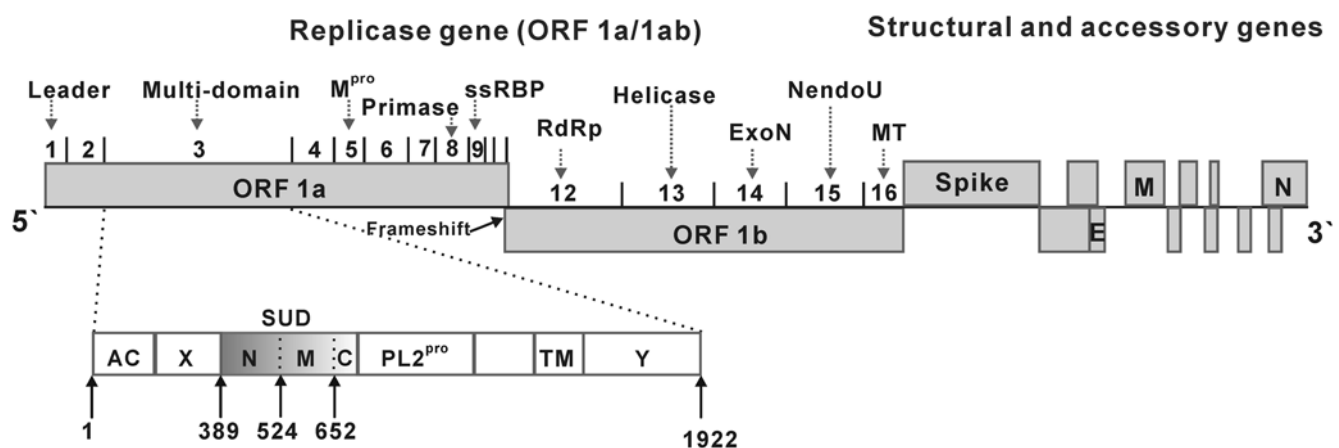
Figure 1. Genome organisation of SARS-CoV. Nsp3 and full-length SUD with subdomains N, M, and C are highlighted. M^{pro}, main (or 3CL) protease; ssRBP, single-stranded RNA-binding protein; RdRp, RNA-dependent RNA polymerase; ExoN, exonuclease; NendoU, uridine-specific endoribonuclease; MT, methyltransferase; Spike, spike protein; E, envelope protein; M, membrane (matrix) protein; N, nucleocapsid protein; Ac, acidic domain; X, X-domain; SUD, SARS-unique domain; PL2^{pro}, papain-like protease; TM, transmembrane region; Y, Y-domain.
doi:10.1371/journal.ppat.1000428.g001

Table 1. Data collection and refinement statistics.

	Monoclinic crystal form	Triclinic crystal form
Data collection		
Wavelength (Å)	1.25485	1.04123
Resolution (Å)	28.25–2.22 (2.34–2.22)	33.33–2.80 (2.96–2.80)
Space group	P2$_1$	P1
Unit-cell parameters		
a (Å)	46.36	68.68
b (Å)	68.55	75.52
c (Å)	94.21	80.54
α (°)	90.00	77.16
β (°)	99.17	75.61
γ (°)	90.00	74.48
Solvent content (%, v/v)	51	63
Overall reflections	166585 (7062)	101963 (9416)
Unique reflections	26598 (2508)	34003 (4086)
Multiplicity	6.3 (2.8)	3.0 (2.3)
Completeness (%)	92.1 (61.0)	93.1 (76.6)
R_{merge}[a]	0.055 (0.373)	0.075 (0.252)
I/σ(I)	16.9 (2.0)	8.8 (2.0)
R_{pim}[b]	0.025 (0.316)	0.056 (0.252)
Refinement		
R_{cryst}[c]	0.211	0.223
R_{free}[c]	0.268	0.240
r.m.s.d. from ideal geometry		
bonds (Å)	0.009	0.008
angles (°)	1.295	1.188
Ramachandran plot regions		
Most favored (%)	94.8	94.7
Additionally allowed (%)	4.6	4.6
Outlier (%)	0.6	0.7

Values in parentheses are for the highest resolution shell.
[a] $R_{merge} = \sum_{hkl} \sum_{i} |I(hkl)_i - \langle I(hkl) \rangle| / \sum_{hkl} \sum_{i} I(hkl)_i$, where $I(hkl)$ is the intensity of reflection hkl and $\langle I(hkl) \rangle$ is the average intensity over all equivalent reflections.
[b] R_{pim} is the precision-indicating merging R factor [78].
[c] $R_{cryst} = \sum_{hkl} |F_o(hkl) - F_c(hkl)| / \sum_{hkl} F_o(hkl)$. R_{free} was calculated for a test set of reflections (5%) omitted from the refinement.
doi:10.1371/journal.ppat.1000428.t001

were monoclinic (space group P2$_1$, two SUD$_{core}$ molecules per asymmetric unit) and diffracted X-rays to 2.2 Å resolution; form-2 crystals were triclinic (space group P1, four SUD$_{core}$ molecules per asymmetric unit) and diffracted to 2.8 Å. Both structures were determined by molecular replacement (see Materials and Methods). The r.m.s. deviations (on Cα atoms) between the models derived from the two different crystal structures are around 0.7 Å.

The models have good stereochemistry (Table 1). 94.7% of the amino-acid residues are in the favoured regions of the Ramachandran plot and 4.6% are in allowed regions. 0.6% are outliers. In all six independent copies of the SUD$_{core}$ monomer, residue Val611 adopts forbidden conformational angles. This residue is located in a turn described by the polypeptide chain where it leaves the subdomain interface (see below) and reaches the surface of the molecule. The side chain makes a hydrophobic contact across the subdomain interface and is also contacting the side chain of Phe406 of a symmetry-related SUD$_{core}$ dimer in the crystal lattice in the monoclinic crystal form (this also applies to two of the four monomers in the triclinic form).

Overall structure

SUD$_{core}$ exhibits a two-domain architecture (Figure 2A). The N-terminal subdomain (SUD-N) comprises Nsp3 residues 389–517, and the C-terminal subdomain of SUD$_{core}$ contains residues 525–652. We call the latter the "middle SUD subdomain", or SUD-M, because full-length SUD has a C-terminal extension of 74 residues compared to SUD$_{core}$. The SUD-N and SUD-M subdomains have a similar fold and can be superimposed with an r.m.s.d. of 3.3–3.4 Å (based on Cα positions); they share 11% sequence identity (see Figure 2C for a structural alignment). Of the 14 amino-acid residues identical between the two subdomains, four form a conserved Leu-Glu-Glu-Ala motif at the N-terminus of helix α4. The linker between the two subdomains (residues 518–524) has no visible electron density. This is due to elevated mobility of the linker, rather than proteolytic cleavage, since we showed by SDS-PAGE of dissolved crystals that the SUD$_{core}$ polypeptide (in the presence of β-mercaptoethanol) has the apparent molecular mass to be expected (~29 kDa; not shown). In addition to the linker, SUD-N and SUD-M are connected by a disulfide bond between cysteines 492 and 623 (Figure 2B). Disulfide bonds are rare in cytosolic proteins, but in coronaviral Nsps, examples of such bonds have been reported [25,26].

The fold of each SUD subdomain is that of a macrodomain (Figure 2A). Macrodomains consist of a largely parallel central β-sheet surrounded by 4–6 α-helices. The order of regular secondary-structure elements in SUD-N is βN1-αN1-βN2-αN2-βN3-βN4-αN3-βN5-αN4-βN6, and in SUD-M βM1-αM1-βM2-αM2-βM3-βM4-αM3-βM5-αM4-βM6-αM5. The topology of the β-strands is β1–β6–β5–β2–β4–β3, all of which are parallel except β3 (Figure 2A). Between the two subdomains, most of the secondary-structure elements are conserved with respect to their position in the three-dimensional structure, although they often differ in length. This is particularly obvious for α-helix 1, which comprises just four residues in the N-terminal subdomain but eleven in the M subdomain. Similarly, α-helix 2 has 5 vs. 10 amino-acid residues in the two subdomains. In general, the strands of the central β-sheet appear to align better between the two subdomains than do the α-helices.

Each of the SUD$_{core}$ subdomains is related to the macrodomain of the histone macro2A ([18]; PDB code 1ZR3, molecule C; for SUD-N: Z-score 9.8, r.m.s.d. 2.5 Å for 112 out of 184 Cα atoms, 12% sequence identity; for SUD-M: Z-score 8.6, r.m.s.d. 2.8 Å for 115 out of 184 Cα atoms, 19% sequence identity). Called "X-domains", single macrodomains are also found in alphaviruses, in hepatitis E virus, and in rubella virus, in addition to coronaviruses [27,28]. The SARS-CoV X-domain (Nsp3b), the domain immediately preceding the SUD in Nsp3, shares no recognizable sequence identity with SUD-N (12%) or SUD-M (7%) (Figure 2C), but its three-dimensional structure [22,23] (PDB code 2ACF, chain A) can be superimposed onto each of the two SUD subdomains with an r.m.s.d. (based on Cα atoms) of 2.7 Å and 2.3 Å, respectively (Figure 2D). Thus, within Nsp3, SARS-CoV has three macrodomains aligned one after the other.

In both crystal forms, SUD$_{core}$ displays the same head-to-tail dimer, with the SUD-N subdomain of monomer A interacting with the SUD-M subdomain of monomer B, and *vice versa*. Approximately 1130 Å^2 of solvent-accessible surface per monomer is buried upon dimerization (Figure 3). Due to the two-domain architecture of each monomer, the resulting four lobes give the
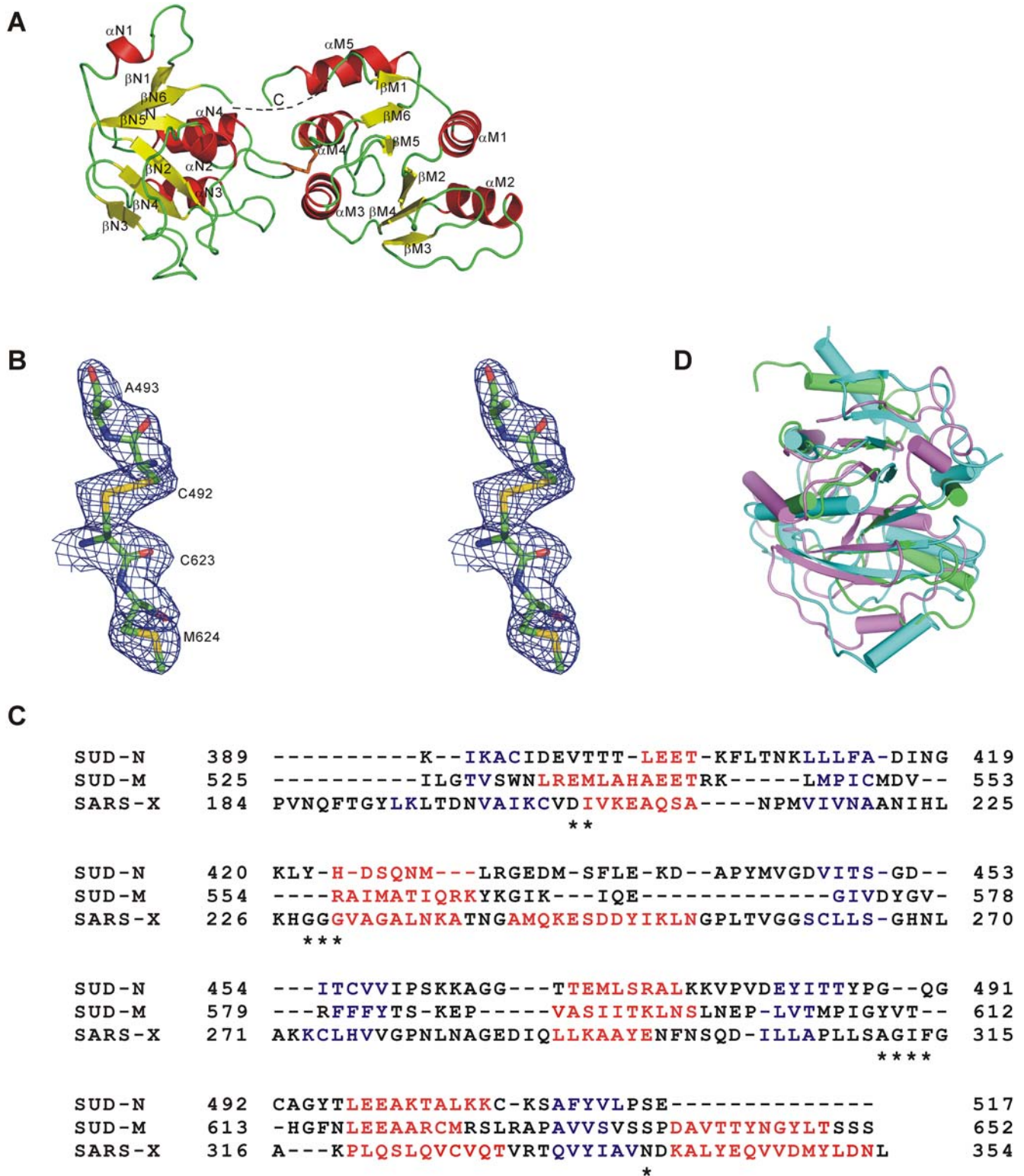
Figure 2. Structure of the SUD_core monomer and comparison with the SARS-CoV X-domain. (A) Ribbon representation of the SUD_core structure (residues 389–652 of Nsp3). The flexible linker connecting the two macrodomains is indicated by a dotted line. The disulfide bond between cysteines 492 of SUD-N and 623 of SUD-M is shown in orange. (B) Stereo image of the 2F_o–F_c electron-density map (1σ above the mean) around the disulfide bond connecting the SUD-N and SUD-M subdomains. (C) Structure-based sequence alignment of the SUD_core subdomains N (SUD-N) and M (SUD-M), and the SARS-CoV X-domain (SARS-X). α-Helices and β-strands are marked red and blue, respectively. Residues 518–524 form the linker between the two SUD subdomains and have not been included in the alignment. Asterisks mark SARS-X residues involved in binding of ADP-ribose. (D) Superimposition of the structures of the SUD-N (violet) and SUD-M (green) subdomains with the SARS-CoV X-domain (cyan).
doi:10.1371/journal.ppat.1000428.g002

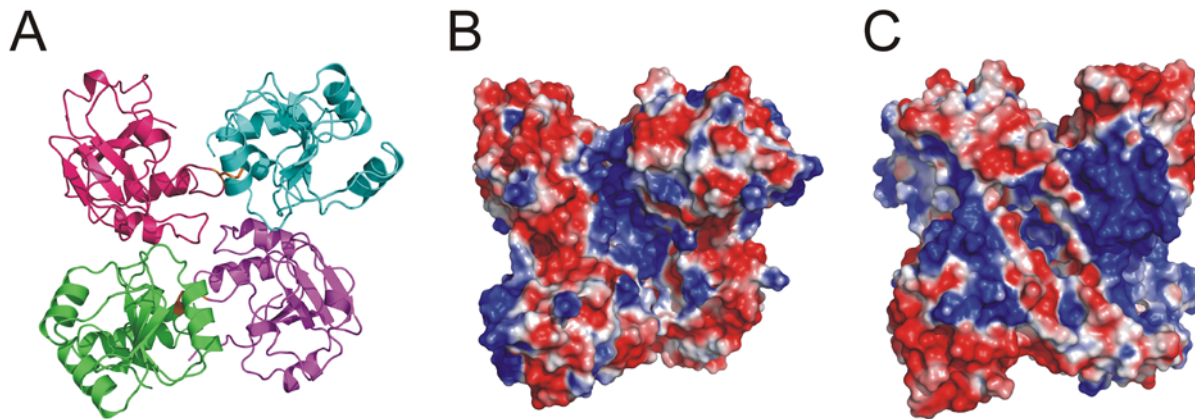**Figure 3. Structure of the SUD$_{core}$ dimer. (A)** SUD$_{core}$ forms a head-to-tail dimer. SUD-N and SUD-M of monomer A are colored violet and cyan, respectively, and SUD-N and SUD-M of monomer B are colored magenta and green, respectively. **(B)** Surface of the SUD$_{core}$ homodimer colored according to electrostatic potential (blue, positive potential; red, negative potential). Orientation is the same as in the cartoon representation in (A). The extended patches of positive potential (blue) are possible binding sites for G-quadruplexes or other nucleic acids. **(C)** As (B), but rotated by 180°. The narrow cleft running across the dimer surface (with a ~45° orientation relative to the monomer-monomer interface, which runs horizontal in this illustration) could be a potential protein-binding site. The monomer–monomer interface is largely hydrophilic and buries ~1130 Å^2 of exposed surface per monomer.
doi:10.1371/journal.ppat.1000428.g003

dimer a quasi-tetrahedral shape (Figure 3A). Involving ~10 hydrogen bonds and four well-defined salt-bridges (Asp-B440…ArgA554, ArgB473…GluA619, ArgB554…AspA440, and GluB619…ArgA473), interactions between the monomers are largely hydrophilic. As to be expected, the structures of the monomers are very similar to one another, with r.m.s.d. values (for Cα atoms) of 0.58 Å between monomers A and B of the monoclinic crystal form, and 0.11–0.37 Å between monomers A–D of the triclinic form. The structure of SUD-M alone is even better conserved between the individual copies of SUD$_{core}$. Also, the fold of the SUD-M subdomain is similar to the model of the SUD fragment 527–651 derived from NMR measurements, which was published very recently (r.m.s.d. ~0.9 Å) [29].

The SUD$_{core}$ macrodomains fail to bind ADP-ribose

The function of the coronaviral X-domain is still unclear; for some coronaviruses such as HCoV 229E and SARS-CoV, it has been shown to exhibit a low ADP-ribose-1″-phosphate phosphatase (Appr-1″-pase, occasionally also called "ADRP") activity and to bind the product of the reaction, ADP-ribose [21–23,30]. However, the two subdomains of SUD$_{core}$ do not bind ADP-ribose, as we have demonstrated by zone-interference gel electrophoresis (Figure S1).

SUD$_{core}$ and its individual subdomains bind G-quadruplexes

When we investigated possible interactions between SUD and nucleic acids by zone-interference gel electrophoresis, we found that the domain binds oligo(G) and oligo(dG) stretches with a K$_D$ of ~1 μM, but not oligo(dA), (dC), or (dT) [31]. Single-stranded nucleotides of random sequence are only bound if they are longer than ~15 nucleotides. Here we demonstrate that each of the two individual SUD subdomains also binds oligo(dG) (Figure 4A). With oligo(dH), where H stands for A, C, or T, but not G, only very small gel shifts, if at all, were observed. As oligo(G) stretches are known to form G-quadruplexes, i.e. four-stranded nucleic-acid structures formed by contiguous guanines [32], we also examined the binding to the oligodeoxynucleotide 5′-GGGCGCGGGAG-GAATTGGGCGGG-3′, a G-rich sequence present in the *bcl-2*

promoter region. This oligonucleotide has been shown by NMR spectroscopy to form a G-quadruplex ([33]; PDB code 2F8U). We found that both full-length SUD and SUD$_{core}$ do indeed bind this oligodeoxynucleotide and that this process is enhanced by the addition of K$^+$ ions, which are known to stabilize G-quadruplex structures (Figure 4B). In agreement with the ability of SUD to non-specifically bind to oligonucleotides of >15 bases [31], both SUD and SUD$_{core}$ were found to bind the reverse-complementary sequence, but with low affinity and, more importantly, independent of K$^+$ ions.

As there is no evidence for SARS-CoV Nsp3 entering the nucleus and binding to DNA, we examined whether SUD would bind to an RNA known to form a quadruplex structure. Indeed, zone-interference gel shift experiments revealed major shifts for both SUD and SUD$_{core}$ in the presence of the oligoribonucleotide 5′-UGGGGGGGGAGGGAGGGAGGGA-3′, which is a protein-binding element in the 3′-nontranslated region of chicken elastin mRNA [34] and forms G-quadruplexes [35] (Figure 4C). Furthermore, we observed a significant gel shift for SUD$_{core}$ when we added the short oligonucleotide UGGGGU, which has also been shown to form a G-quadruplex ([36]; PDB code 1J8G). This shift was also enhanced by the addition of K$^+$ (Figure 4D). Thus, SUD binds RNA (rG)-quadruplexes and DNA (dG)-quadruplexes with comparable affinity.

Effect of lysine mutations on G-quadruplex binding

Inspection of the structure of the SUD dimer reveals a central narrow cleft running across the dimer surface, but distinct from the monomer-monomer interface (Figure 3C), which could be a binding site for another protein. In addition, there are several positively charged patches in the center of the dimer (Figure 3B), and on its backside (Figure 3C), which could be involved in binding to G-quadruplexes. We have prepared four sets of mutations by replacing lysine residues (and one glutamate) in these patches by alanines. The first two pairs of mutations, K505A+K506A (M1, at the end of helix αN4) and K476A+K477A (M2, in the loop between αN3 and βN5), are located on the surface of the SUD-N subdomain and lead to reduced shifts with G-quadruplexes in the zone-interference gel electrophoresis experiment, both with the G-quadruplex from the
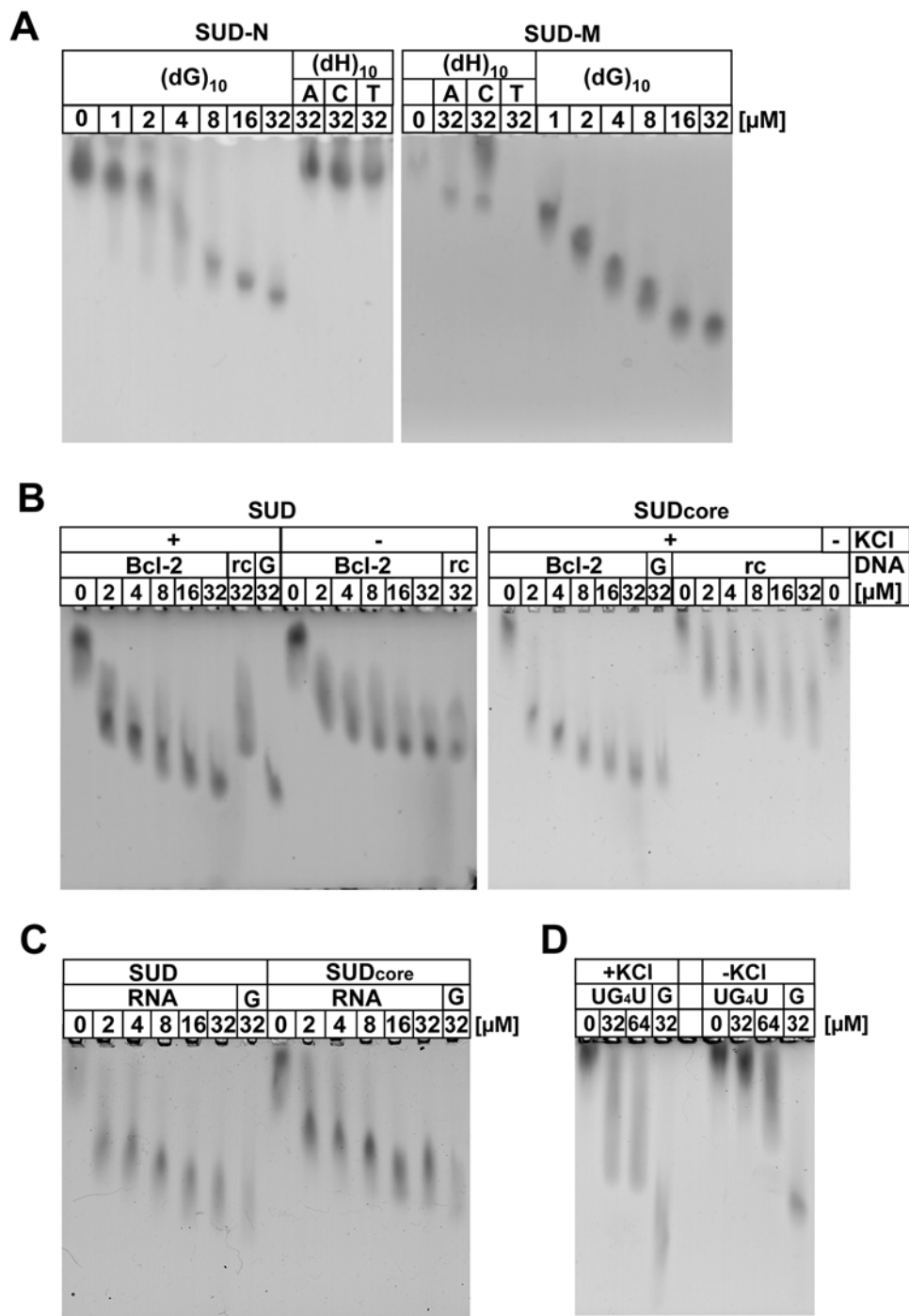
Figure 4. Binding of oligonucleotides to SUD as demonstrated by zone-interference gel electrophoresis. Protein concentration was 10 µM in all experiments. (A) Binding of increasing concentrations (indicated above the lanes) of $(dG)_{10}$ to the SUD-N and SUD-M subdomains of SUD_{core} (left and right panel, resp.). Comparison with 32 µM $(dA)_{10}$, $(dC)_{10}$, or $(dT)_{10}$ shows that the binding is specific for $(dG)_{10}$. "H" stands for A, C, or T. (B) Binding of increasing concentrations (indicated above the lanes) of the quadruplex-forming oligodeoxynucleotide 5'-GGGCGCGGGAG-GAATTGGGCGGG-3' (labeled "Bcl-2") as occurring within the *bcl-2* promoter region, in the presence and absence of 100 mM KCl, which is known to promote quadruplex formation. Left panel, full-length SUD; right panel, SUD_{core}. The reverse-complementary oligodeoxynucleotide (labeled "rc"), which fails to form a quadruplex but exceeds the minimum length of ~15 nucleotides for non-quadruplex interaction with SUD, is also bound, but with reduced affinity and independently of KCl. $(dG)_{10}$ (labeled "G") has been included as a positive control. (C) Binding of increasing concentrations (indicated above the lanes) of the quadruplex-forming oligoribonucleotide 5'-UGGGGGGGAGGGAGGGAGGGA-3' (labeled "RNA") as occurring in the 3'-nontranslated region of chicken elastin mRNA. Left panel: interaction with full-length SUD; right panel: interaction with SUD_{core}. Binding of $(dG)_{10}$ (labeled "G") is shown for comparison. 100 mM KCl was present in all lanes. (D) Binding to SUD_{core} of the quadruplex-forming oligonucleotide 5'-UGGGGU-3' (labeled "UG4U") in the presence (left) and absence (right) of 100 mM KCl. $(dG)_{10}$ (labeled "G") has been included as a positive control.
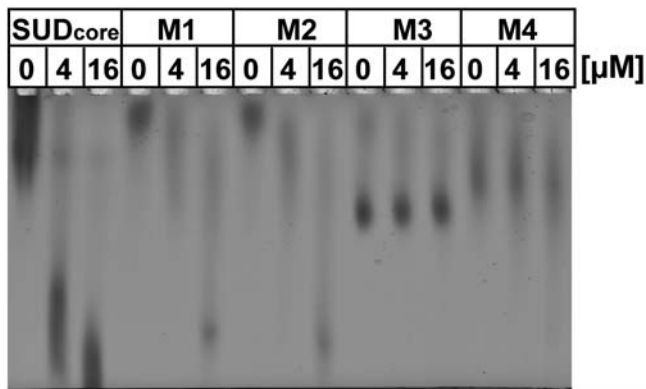doi:10.1371/journal.ppat.1000428.g004

Figure 5. G-quadruplex binding is affected by mutations of lysine residues on the surface of SUD$_{core}$. Binding of double and triple mutants of SUD$_{core}$ to the quadruplex-forming oligodeoxynucleotide 5′-GGGCGCGGGAGGAATTGGGCGGG-3′ as occurring within the *bcl-2* promoter region, in the presence of 100 mM KCl, as demonstrated by zone-interference gel electrophoresis. Protein concentration was 10 μM in all experiments. Oligonucleotide at two concentrations (4 and 16 μM) was added to wild-type SUD$_{core}$ and mutants M1 (K505A+K506A), M2 (K476A+K477A), M3 (K563A+K565A+K568A), and M4 (K565A+K568A+E571A). Mutants M1 and M2 show reduced shifts, in particular at 4 μM nucleotide, whereas mutants M3 and M4 abolish binding. Note that in the absence of nucleotide, mutant proteins M3 and M4 behave differently on the gel because of different charges. doi:10.1371/journal.ppat.1000428.g005

bcl-2 promoter region (Figure 5) and with (dG)$_{10}$ (not shown). The second set of mutations, K563A+K565A+K568A (M3) and K565A+K568A+E571A (M4) are located in the loop connecting αM2 and βM3 of the SUD-M subdomain and abolish G-quadruplex binding altogether (Figure 5), again with both oligonucleotides.

Discussion

When the SARS-unique domain was first predicted [1], the boundaries of the domain were set approximately at Nsp3 residues 352 and 726. We made major efforts to produce this protein in a stable form, but with little success. Only when we used *in-vitro* protein synthesis, were we able to obtain small amounts of a relatively stable preparation comprising Nsp3 residues 349–726 [31]. At the N-terminus of this construct, up to eleven residues actually correspond to the C-terminus of the preceding X-domain (Nsp3b). When we expressed a gene construct coding for SUD (349–726) in *E. coli*, we observed rapid proteolytic degradation of the N-terminal segment. The relatively stable intermediate obtained had its N-terminus at Nsp3 residue 389. The N-terminal segment ~359–388 is predicted to be intrinsically unfolded by several prediction programs (not shown). Therefore, we assume segment 359–388 to be merely a linker between Nsp3b and SUD, and 389 to be the first residue of the latter. This assignment is justified by the observation that in our crystal structures reported here, the SUD-N subdomain is a complete macrodomain without any residues lacking at the N-terminus. Therefore, the protein corresponding to Nsp3 residues 389–726 is called "full-length SUD" here.

In this communication, we describe the crystal structures at 2.2 Å and 2.8 Å resolution (monoclinic and triclinic form, respectively) of the core of the SARS-unique domain (SUD$_{core}$, Nsp3 residues 389–652). SUD$_{core}$ turns out to consist of two subdomains, SUD-N (Nsp3 residues 389–517) and SUD-M (525–652), each exhibiting the fold of a macrodomain. The two subdomains are connected by a flexible linker (residues 518–524)

and a disulfide bond. Even though coronavirus replication occurs in the cytosol, where the environment is reductive, it is unlikely that the formation of this disulfide is an artifact owing to handling of the protein: As the linker between the SUD-N and SUD-M subdomains is very short (seven residues), and the mutual orientation of the subdomains is fixed due to the tight dimerization, cysteine residues no. 492 and 623 will be very close to one another irrespective of the exact conformation of the linker. In fact, disulfide bonds are not uncommon in coronaviral non-structural proteins (Nsps) involved in RNA replication or transcription. Among others, they have been observed in HCoV-229E Nsp9 [25] and turkey coronavirus Nsp15 [26], but in these cases, the disulfide bond connects two subunits of the homo-oligomeric proteins, whereas the occurrence in SUD$_{core}$ is the first case of an *intra*molecular disulfide bond described in a coronavirus Nsp.

Coronavirus replication in the perinuclear region of the cell is localized to double-membrane vesicles that have been hijacked from the endoplasmic reticulum or late endosomes [37–40]. These vesicles are around 200–350 nm in diameter and present alone or as clusters in the cytosol [38]. The milieu inside or at the surface of these vesicles is unknown, but it is well possible that it is partially oxidative. It has also been speculated [25] that formation of disulfide bonds may be a way for the coronaviral Nsps to function in the presence of the oxidative stress that is the consequence of the viral infection [41–43].

Our identification of two macrodomains in SUD$_{core}$ brings the number of these domains in SARS-CoV Nsp3 to three. What are the functions of these modules? The original SARS-CoV "X-domain" (Nsp3b) has been shown to have low ADP-ribose-1″-phosphate phosphatase (Appr-1″-pase or "ADRP") activity [21–23]. However, this assignment is controversial. A nuclear Appr-1″-pase (Poa1p in yeast, [20]) is an enzyme of a tRNA metabolic pathway, but there is no evidence for coronavirus Nsp3 ever being translocated to the nucleus, and the other enzymes involved in this pathway are missing in coronaviruses (with the exception of the cyclic 1″,2″-phosphodiesterase (CPDase) in group 2a viruses such as Mouse Hepatitis Virus, Bovine Coronavirus, and Human Coronavirus OC43). Therefore, it has been proposed that the X-domain may be involved in binding poly(ADP-ribose), a metabolic product of NAD$^+$ synthesized by the enzyme poly(ADP-ribose) polymerase (PARP; [23]). However, we have recently demonstrated that the X-domain of Infectious Bronchitis Virus (IBV) strain Beaudette, a group-3 coronavirus, does not have significant affinity to ADP-ribose [24]. This can be explained on the basis of crystal structures: In the X-domain (Nsp3b) of SARS-CoV [23], and in that of HCoV 229E [24], a stretch of three conserved glycine residues is involved in binding the pyrophosphate unit of ADP-ribose, whereas in the corresponding domain of IBV strain Beaudette (but not in all IBV strains, see [44]), the second glycine is replaced by serine, leading to steric interference with ADP-ribose binding [24]. In the two SUD subdomains, the triple-glycine sequence is not conserved (see Figure 2C), and hence, they do not bind ADP-ribose either.

Neuman *et al.* [2] reported that full-length SUD binds cobalt ions, whereas a domain called SUD-C by these authors, which is however almost identical (residues 513–651) to our SUD-M (525–652), does not. From this, they concluded that the metal-binding activity is associated with the cysteine residues in the N-terminal subdomain. We were also able to observe binding of cobalt ions to SUD$_{core}$ by following the occurrence of a peak at 310 nm in the UV spectrum, which, in contrast to the data presented by Neuman *et al.* [2], could be reverted by addition of zinc ions. However, when we removed the N-terminal His-tag, this phenomenon could

no longer be observed. Furthermore, we note that of the four cysteine residues in the SUD-N subdomain (residues 393, 456, 492, and 507), 456 and 507 are non-accessible in the interior of the subdomain, and 492 is involved in the buried disulfide bond to Cys623; therefore, Cys393 and perhaps the solvent-exposed His423 would remain the only potential ligands for cobalt ions in SUD-N. However, these residues are >12 Å apart and thus unlikely to chelate cobalt ions.

For SUD-M, a recent publication [29] reported binding to oligo(A). However, we fail to observe this (Figure 4A, lane labeled "A"). Instead, we have demonstrated that full-length SUD and SUD$_{core}$ bind oligodeoxynucleotides and oligoribonucleotides that form G-quadruplexes. For full-length SUD and SUD$_{core}$, we had previously shown binding to oligo(dG) and oligo(G) stretches [31], but the demonstration here of oligo(dG) binding to the individual SUD$_{core}$ subdomains, SUD-N and SUD-M, is unexpected because their overall electrostatic properties are very different from one another: SUD-N is acidic (pI = 5.3), whereas SUD-M is basic (pI = 9.0). However, even SUD-N shows surface patches with positive electrostatics that could bind nucleic acid (Figure 3B).

We have used automatic docking procedures to place the G-quadruplex found in the *bcl-2* promoter region ([33]; PDB code 2F8U) into our crystal structures. One potential binding site identified is in the cleft between the SUD-M and the SUD-N subdomains within the SUD$_{core}$ dimer (Figure S2A); this binding site is spatially close to the mutations M3 and M4, consistent with the observation that these mutations abolish binding completely. However, we have previously shown by Dynamic Light-Scattering that G-quadruplex binding leads to oligomerization of SUD$_{core}$ [31]. Consequently, we have also constructed models based on the packing modes of SUD$_{core}$ dimers observed in our crystal structures. One potential binding site for G-quadruplexes might be in a cleft between two consecutive SUD$_{core}$ dimers as they occur in both the monoclinic and triclinic crystal forms (Figure S2B), but for confirmation, any of these models will have to await crystallographic determination of the complex. In summary, our mutation experiments demonstrate an involvement of several of the many lysine residues of SUD in binding G-quadruplexes, but as it is probably extended surfaces of SUD$_{core}$ oligomers that participate in this process, it is not possible to pinpoint any single amino-acid residue.

The target of SUD binding could be G-quadruplexes in RNA of viral or/and cellular origin. The SARS-CoV genome contains three G$_6$-stretches (one on the plus-strand and two on the minus-strand) and an additional two G$_5$-sequences, which could perhaps form local G-quadruplexes. However, the G-stretch binding capabilities of SUD and SUD$_{core}$ seem to have been optimized for recognition of longer G-rich sequences. By systematic variation of the length of oligo(dG), we found that SUD$_{core}$ exhibits strongest affinity (K$_D$ ~0.45 µM) for (dG)$_{10}$ to (dG)$_{14}$ [31]. The 3′-nontranslated regions of several host-cell mRNAs coding for proteins involved in the regulation of apoptosis and in signaling pathways contain long G-stretches and could also be targets of SUD. Examples of such mRNAs are those coding for the pro-apoptotic protein Bbc3 [45], RAB6B (a member of the Ras oncogene family, [46]), MAP kinase 1 [47], and TAB3, a component of the NF-κB signaling pathway [48]. It is conceivable that these proteins might be targets for the virus when interfering with cellular signaling. Changes in the stability and/or translation efficiency of these mRNAs due to the binding of a viral regulatory factor could result in an altered reaction of the infected cell to apoptotic signals, or it could silence the antiviral response.

The idea that coronaviral X-domains might function as modules binding poly(ADP-ribose) [23] received support from the observation that some macrodomains are connected with domains showing poly(ADP-ribose) polymerase (PARP) activity, i.e. in the so-called macroPARPs (PARP-9 and PARP-14) [49]. There are 18 human genes for members of the PARP family; the prototype enzyme, PARP-1, catalyzes the post-translational modification of many substrate proteins, including itself, in a multitude of cellular processes (DNA repair, transcriptional regulation, energy metabolism, and apoptosis) [50–52]. Interestingly, SUD-M and the C-terminal 74-residue subdomain (SUD-C) that is missing in our SUD$_{core}$ construct together show a ~15% sequence identity (32% similarity) to the catalytic domain of PARP-1. However, the three-dimensional structures of SUD-M (this work) and the C-terminal domain of PARP-1 [53] are different and cannot be superimposed. Another feature common between SARS-CoV SUD and PARP-1 is that the latter has recently been shown to bind to G-quadruplexes [54], although it is generally assumed that this occurs through the DNA-binding domain rather than the catalytic domain of PARP-1.

PARP-1 and most of its family members are located to the nucleus, while PARP-4 and others predominantly act in the cytoplasm [50–52]. PARP-4 is incorporated into vaults, RNA-containing subcellular particles in the cytoplasm [55]. Furthermore, ZAP, a human antiviral protein comprising a C-terminal PARP-like domain devoid of catalytic activity, has been shown to exhibit antiviral activity on alphaviruses [56], which contain an X-domain similar to that of coronaviruses [23,27,28]. In addition, ZAP contains an N-terminal zinc-finger domain, a central TiPARP (2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD)-inducible PARP) domain, and a WWE domain (a protein-protein interaction module in ubiquitin and ADP-ribose conjugation proteins). In fact, ZAP appears to be part of the human innate immune system and to play a role comparable to APOBEC3G in HIV infection [57]. It is possible that this group of viruses has evolved macrodomains to counteract the antiviral activity of ZAP. Indeed, macrodomains can inhibit PARPs, as has been shown for the macrodomain of the histone mH2A1.1, which downregulates the catalytic activity of PARP-1 [58]. Having three macrodomains at its disposal, SARS-CoV may be much more efficient in knocking down the antiviral response of the host cell than other coronaviruses. Whether this involves a direct interaction between SUD and ZAP or another member of the PARP family, or competition for G-quadruplexes in viral or host-cell RNA, remains to be shown.

Materials and Methods

Recombinant protein production and purification

Full-length SUD (Nsp3 residues 389–726) and the fragment SUD$_{core}$ (Nsp3 residues 389–652, previously called "SUDc5b") of SARS-CoV strain TOR2 (acc. no. AY274119) were produced recombinantly in *E. coli* as described [31]. The coding regions for the SUD-N subdomain (Nsp3 residues 389–524) and the SUD-M subdomain (Nsp3 residues 525–652) were constructed by introducing an appropriate deletion into the previously described plasmid pQE30-Xa-c5b [31] using site-directed mutagenesis. Plasmids encoding SUD-N and SUD-M were prepared using primers listed in Table S1. The coding regions for four sets of mutations of SUD$_{core}$, M1 (K505A+K506A), M2 (K476A+K477A), M3 (K563A+K565A+K568A), and M4 (K565A+K568A+E571A), were constructed by introducing appropriate mutations into plasmid pQE30-Xa-c5b [31] using site-directed mutagenesis. Plasmids encoding these mutants were prepared using primers also listed in Table S1. All plasmids provided an N-terminal His-tag and a short linker sequence encoding a factor-Xa cleavage site. The

coding regions of the expression plasmids were verified by DNA sequencing. *E. coli* M15 (pRep4) was used as expression host for these constructs. SUD-N, SUD-M, and the mutated proteins were purified according to the same protocol as for SUD$_{core}$ [31].

Crystallization

SUD$_{core}$ displayed >95% purity in SDS-PAGE, and monodispersity in Dynamic Light- Scattering. Initial crystallization screening was performed using the sitting-drop vapor-diffusion method in 96-well Intelli-Plates (Dunn Laboratories). Several commercial kits (Sigma, Jena Bioscience) were used for the screening. The protein concentration was 6 mg/ml. Using a Phoenix robotic system (Art Robbins), drops were made of 260 nl protein and 260 nl precipitant solution. The optimized crystallization condition consisted of 20% polyethylene glycol monomethyl ether 5000 and 0.2 M ammonium sulfate in 0.1 M morpholinoethane sulfonic acid (pH 6.5). Plate-like crystals grew in 3–5 days, to maximum dimensions of $0.02 \times 0.02 \times 0.01$ mm^3.

Data collection and processing

Many SUD$_{core}$ crystals had to be tested for diffraction until one yielding data to 2.2 Å resolution was found. The best diffracting crystals belonged to space group P2$_1$. Under the same crystallization conditions, a second crystal form belonging to space group P1 was observed, diffracting to lower resolution of about 2.8 Å. Crystals were cryoprotected in reservoir solution that included 30% glycerol, and were harvested into a loop prior to flash-cooling in liquid nitrogen. All data were collected at 100 K from a single crystal each at beamline BL14.2, BESSY (Berlin, Germany), using an MX225 CCD detector (Rayonics), or at beamline I911-2 at MAX-lab (Lund, Sweden), using a Mar165 CCD detector (Marresearch). Data were processed with MOSFLM [59], and reduced and scaled using the SCALA [60] program from the CCP4 suite [61]. Crystals belonging to space group P2$_1$ had unit-cell parameters $a = 46.36$ Å, $b = 68.55$ Å, $c = 94.21$ Å, $\beta = 99.17°$, those belonging to space group P1 had unit-cell parameters $a = 68.68$ Å, $b = 75.52$ Å, $c = 80.54$ Å, $\alpha = 77.16°$, $\beta = 75.61°$, $\gamma = 74.48°$. Data-collection statistics for both crystal forms are shown in Table 1. The asymmetric unit of the P2$_1$ form contained two SUD$_{core}$ monomers, giving a Matthews coefficient [62] of 2.5 Å^3 Da^{-1} and a solvent content of 51%, whereas the P1 crystal form had four monomers per asymmetric unit, giving corresponding parameters of 3.2 Å^3 Da^{-1} and 63%.

Structure determination

We attempted to solve the structure by molecular replacement into the P2$_1$ form using the NMR coordinates of a subdomain comprising SARS-CoV Nsp3 residues 513–651; PDB code 2JWJ [29,63]), which is almost identical to the SUD-M subdomain of SARS-CoV Nsp3. Using the program Phaser [64,65], we could find two solutions, and the C-terminal part of SUD$_{core}$ was well defined in the electron-density maps. However, for the N-terminal half, only a few segments of poly(Ala) chain could be built into the maps. This starting model was then refined in BUSTER-TNT [66] using Local Structure Similarity Restraints (LSSR) [67] as non-crystallographic symmetry (NCS) restraints to give R and R$_{free}$ values of 0.453 and 0.479, respectively. The resulting 2mF$_o$-DF$_c$ electron density was subjected to density modification using solvent flattening, histogram matching, and 2-fold NCS-averaging using DM [68]. The averaging masks were calculated and updated using the auto-correlation procedure [69] as implemented in DM. Using the automatic building program BUCCANEER [70] together with REFMAC [71] (as implemented in the CCP4i

[72] interface for CCP4) in an iterative procedure for 20 cycles resulted in a model for 501 residues in 10 chains (the longest having 208 residues), in which 448 residues were assigned both a chemical identity and a sequential residue number, while the remaining 53 residues were modeled as poly(Ala) in 8 shorter chains. The R and R$_{free}$ values resulting from REFMAC were 0.374 and 0.414, respectively. This model was refined in BUSTER-TNT, again using LSSR as NCS restraints for the common parts in the already sequenced 448 residues of the dimer, to R and R$_{free}$ values of 0.269 and 0.316. The improved electron density was again subjected to density modification using DM as detailed above, but using a lower solvent content of 35% as well as anisotropically scaled observed amplitudes as output by BUSTER-TNT. The resulting density-modified and NCS-averaged map was then used for automatic model building using the iterative BUCCANEER/REFMAC procedure described above. This produced a model with 511 residues in 5 chains with 487 residues sequenced. The R and R$_{free}$ values from REFMAC for this model were 0.289 and 0.326, respectively.

Since the refinements in BUSTER-TNT at that point showed some problematic low correlations between F$_o$ and F$_c$ at low resolution, the original images collected from the P2$_1$ crystal were reprocessed using XDS [73] and SCALA, applying different high-resolution cutoffs for different segments of the collected images. Details for this dataset are given in Table 1. Subsequent refinement of the P2$_1$ form with REFMAC, under application of weak NCS restraints, yielded a model with R = 0.211, R$_{free}$ = 0.264. The advanced handling of NCS restraints through LSSR in BUSTER-TNT gave a final model R = 0.211 and R$_{free}$ = 0.268. The final model in the P2$_1$ form comprises 513 residues (A389–A516; A524–A652; B393–B519; B526–B652).

Chain A of the P2$_1$ form was used for molecular replacement with the program MOLREP [74] into the P1 form. There was an unambiguous solution for four molecules in the asymmetric unit. This model was refined with BUSTER-TNT (using LSSR for NCS restraints) and rebuilt in Coot [75] to final values of R = 0.223 and R$_{free}$ = 0.240. The final model of the P1 form comprises 1014 residues.

The figures were made with PyMOL [76].

Zone-interference gel electrophoresis (ZIGE)

The zone-interference gel electrophoresis (ZIGE) device was adapted from Abrahams *et al.* [77]. ZIGE assays were performed using a horizontal 1% agarose gel system in TBE buffer (20 mM Tris, 50 mM boric acid, 0.1 mM ethylenediaminetetraacetic acid (EDTA), pH 8.3). The protein was incubated at room temperature for 30 min with different concentrations of oligodeoxynucleotides, such as (dG)$_{10}$ and *bcl-2* promoter region (5′-GGGCGCGGGAG-GAATTGGGCGGG-3′), or oligoribonucleotides (5′-UGGGGG-GAGGGAGGGAGGGA-3′ and 5′-UGGGGU-3′). The samples were mixed with dimethylsulfoxide (DMSO; final concentration 10% (v/v)) and a trace of bromophenolblue (BPB). These protein-oligonucleotide samples were applied to the small slots. Oligonucleotide with the same concentration as in the small slots was also mixed with DMSO and BPB in 1xTBE buffer and applied to the long slots of the gel (total volume 100 µl). Electrophoresis was performed at 4°C for 1 h with a constant current of 100 mA. Staining was performed as outlined in [77].

Accession Codes

Protein Data Bank: Coordinates and structure factors have been deposited with accession code 2W2G (P2$_1$ crystal form) and 2WCT (P1 crystal form).

Supporting Information

Figure S1 Zone-interference gel electrophoresis experiment showing that SUD_{core} fails to bind NAD+ and ADP-ribose. SUD_{core} alone (label 0) and decreasing concentrations (1, 0.5, 0.1, 0.05 and 0.02 mM) of NAD^+, or decreasing concentrations (1, 0.5, 0.1, 0.05 and 0.02 mM) of ADP-ribose.
Found at: doi:10.1371/journal.ppat.1000428.s001 (0.70 MB DOC)

Figure S2 Alternative models of G-quadruplex binding to SUD_{core}, obtained by automated docking into the crystal structures. The SUD-N and SUD-M subdomains are in violet and cyan, respectively, the G-quadruplex as found in the *bcl-2* promoter region (PDB code: 2F8U) is in orange. The pairs of mutations in SUD-N are indicated by green (M1, K505A+K506A) and blue (M2, K476A+K477A) spheres. The M3 set of mutations in SUD-M is indicated by olive (K563A) and orange (K565A+K568A) spheres. The M4 set of mutations, also in SUD-M, is indicated by orange (K565A+K568A) and yellow (E571A) spheres. (A) A possible binding site is in a cleft between monomers in the SUD_{core} dimer. The binding site is close to the lysine residues replaced by the M3 and M4 mutations, compatible with the inability of these mutants to bind G-quadruplexes. (B) A second potential binding site is a cleft between two neighboring SUD_{core} dimers as found in both crystal packing arrangements (space groups $P2_1$ and P1). This binding mode is compatible with the observation of SUD_{core} oligomerization upon G-quadruplex binding.
Found at: doi:10.1371/journal.ppat.1000428.s002 (3.46 MB DOC)

Table S1 Primer sequences used for SUD-N, SUD-M, and four sets of mutants.
Found at: doi:10.1371/journal.ppat.1000428.s003 (0.01 MB PDF)

Author Contributions

Conceived and designed the experiments: JT YK CLS RH. Performed the experiments: JT YK. Analyzed the data: JT CV GB MB GH JRM CLS RH. Contributed reagents/materials/analysis tools: CV OSS GB. Wrote the paper: JT RH.

References

1. Snijder EJ, Bredenbeek PJ, Dobbe JC, Thiel V, Ziebuhr J, et al. (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.
2. Neuman BW, Joseph JS, Saikatendu KS, Serrano P, Chatterjee A, et al. (2008) Proteomics analysis unravels the functional repertoire of coronavirus nonstructural protein 3. J Virol 82: 5279–5294.
3. Hilgenfeld R, Tan J, Chen S, Shen X, Jiang H (2008) Structural proteomics of emerging viruses: The examples of SARS-CoV and other coronaviruses. In: Sussman JL, Silman I, eds (2008) Structural proteomics and its impact on the life sciences. Singapore: World Scientific. pp 361–433.
4. Serrano P, Johnson MA, Almeida MS, Horst R, Herrmann T, et al. (2007) Nuclear magnetic resonance structure of the N-terminal domain of nonstructural protein 3 from the severe acute respiratory syndrome coronavirus. J Virol 81: 12049–12060.
5. Ratia K, Saikatendu KS, Santarsiero BD, Barretto N, Backer SC, et al. (2006) Severe acute respiratory syndrome coronavirus papain-like protease: structure of a viral deubiquitinating enzyme. Proc Natl Acad Sci U S A 103: 5717–5722.
6. Anand K, Ziebuhr J, Wadhwani P, Mesters JR, Hilgenfeld R (2003) Coronavirus main proteinase (3CL^pro) structure: Basis for design of anti-SARS drugs. Science 300: 1763–1767.
7. Yang H, Yang M, Ding Y, Liu Y, Lou Z, et al. (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.
8. Tan J, Verschueren KHG, Anand K, Shen J, Yang M, et al. (2005) pH-dependent conformational flexibility of the SARS-CoV main proteinase (M^pro) dimer: Molecular dynamics simulations and multiple X-ray structure analyses. J Mol Biol 354: 25–40.
9. Sulea T, Lindner HA, Purisima EO, Ménard R (2005) Deubiquitination, a new function of the severe acute respiratory syndrome coronavirus papain-like protease? J Virol 79: 4550–4551.
10. Barretto N, Jukneliene D, Ratia K, Chen Z, Mesecar AD, et al. (2005) The papain-like protease of severe acute respiratory syndrome coronavirus has deubiquitinating activity. J Virol 79: 15189–15198.
11. Lindner HA, Fotouhi-Ardakani N, Lytvyn V, Lachance P, Sulea T, et al. (2005) The papain-like protease from the severe acute respiratory syndrome coronavirus is a deubiquitinating enzyme. J Virol 79: 15199–15208.
12. Barretto N, Jukneliene D, Ratia K, Chen Z, Mesecar AD, et al. (2006) Deubiquitinating activity of the SARS-CoV papain-like protease. Adv Exp Med Biol 581: 37–41.
13. Lindner HA, Lytvyn V, Qi H, Lachance P, Ziomek E, et al. (2007) Selectivity in ISG15 and ubiquitin recognition by the SARS coronavirus papain-like protease. Arch Biochem Biophys 466: 8–14.
14. Kim KI, Zhang DE (2003) ISG15, not just another ubiquitin-like protein. Biochem Biophys Res Commun 307: 431–434.
15. Ritchie KJ, Zhang DE (2004) ISG15: the immunological kin of ubiquitin. Semin Cell Dev Biol 15: 237–246.
16. Devaraj SG, Wang N, Chen Z, Chen Z, Tseng M, et al. (2007) Regulation of IRF-3-dependent innate immunity by the papain-like protease domain of the severe acute respiratory syndrome coronavirus. J Biol Chem 282: 32208–32221.
17. Pehrson JR, Fried VA (1992) MacroH2A, a core histone containing a large nonhistone region. Science 257: 1398–1400.
18. Kustatscher G, Hothorn M, Pugieux C, Scheffzek K, Ladurner AG (2005) Splicing regulates NAD metabolite binding to histone macroH2A. Nature Struct Mol Biol 12: 624–625.
19. Chakravarthy S, Gundimella SK, Caron C, Perche PY, Pehrson JR, et al. (2005) Structural characterization of the histone variant macroH2A. Mol Cell Biol 25: 7616–7624.
20. Shull NP, Spinelli SL, Phizicky EM (2005) A highly specific phosphatase that acts on ADP-ribose 1″-phosphate, a metabolite of tRNA splicing in *Saccharomyces cerevisiae*. Nucleic Acids Res 33: 650–660.
21. Putics Á, Filipowicz W, Hall J, Gorbalenya AE, Ziebuhr J (2005) ADP-ribose-1″-monophosphatases: a conserved coronavirus enzyme that is dispensable for viral replication in tissue culture. J Virol 79: 12721–12731.
22. Saikatendu KS, Joseph JS, Subramanian V, Clayton T, Griffith M, et al. (2005) Structural basis of severe acute respiratory syndrome coronavirus ADP-ribose-1″-phosphate dephosphorylation by a conserved domain of nsP3. Structure 13: 1665–1675.
23. Egloff MP, Malet H, Putics Á, Heinonen M, Dutartre H, et al. (2006) Structural and functional basis for ADP-ribose and poly(ADP-ribose) binding by the coronavirus macro domains. J Virol 80: 8493–8502.
24. Piotrowski Y, Hansen G, Boomaars-van der Zanden AL, Snijder EJ, Gorbalenya AE, et al. (2009) Crystal structures of the X-domains of a group-1 and a group-3 coronavirus reveal that ADP-ribose-binding may not be a conserved property. Protein Sci 18: 6–16.
25. Ponnusamy R, Moll R, Weimar T, Mesters JR, Hilgenfeld R (2008) Variable oligomerization modes in coronavirus non-structural protein 9. J Mol Biol 383: 1081–1096.
26. Cao J, Wu CC, Lin TL (2008) Turkey coronavirus non-structure protein NSP15 - an endoribonuclease. Intervirology 51: 342–351.
27. Koonin EV, Gorbalenya AE, Purdy MA, Rozanov MN, Reyes GR, et al. (1992) Computer-assisted assignment of functional domains in the nonstructural polyprotein of hepatitis E virus: Delineation of an additional group of positive-strand RNA plant and animal viruses. Proc Natl Acad Sci U S A 89: 8259–8263.
28. Neuvonen M, Ahola T (2008) Differential activities of cellular and viral macro domain proteins in binding of ADP-ribose metabolites. J Mol Biol 385: 212–225.
29. Chatterjee A, Johnson MA, Serrano P, Pedrini B, Joseph JJ, et al. (2009) Nuclear magnetic resonance structure shows that the SARS-unique domain contains a macrodomain fold. J Virol 83: 1823–1836.
30. Putics Á, Slaby J, Filipowicz W, Gorbalenya AE, Ziebuhr J (2006) ADP-ribose-1″-phosphatase activities of the human coronavirus 229E and SARS coronavirus X domains. Adv Exp Med Biol 581: 93–96.
31. Tan J, Kusov Y, Mutschall D, Tech S, Nagarajan K, et al. (2007) The "SARS-unique" domain (SUD) of SARS coronavirus is an oligo(G)-binding protein. Biochem Biophys Res Commun 364: 877–882.
32. Keniry MA (2001) Quadruplex structures in nucleic acids. Biopolymers 56: 123–146.
33. Dai J, Chen D, Jones RA, Hurley LH, Yang D (2006) NMR solution structure of the major G-quadruplex structure formed in the human BCL2 promoter region. Nucleic Acids Res 34: 5133–5144.

34. Hew Y, Lau C, Grzelczak Z, Keeley FW (2000) Identification of a GA-rich sequence as a protein-binding site in the 3′-untranslated region of chicken elastin mRNA with a potential role in the developmental regulation of elastin mRNA stability. J Biol Chem 275: 24857–24864.

35. Pan B, Xiong Y, Shi K, Deng J, Sundaralingam M (2003) Crystal structure of an RNA purine-rich tetraplex containing adenine tetrads: Implications for specific binding in RNA tetraplexes. Structure 11: 815–823.

36. Deng J, Xiong Y, Sundaralingam M (2001) X-ray analysis of an RNA tetraplex (UGGGGU)$_4$ with divalent Sr^{2+} ions at subatomic resolution (0.61 Å). Proc Natl Acad Sci U S A 98: 13665–13670.

37. Prentice E, Jerome WG, Yoshimori T, Mizushima N, Denison MR (2004) Coronavirus replication complex utilizes components of cellular autophagy. J Biol Chem 279: 10136–10141.

38. Snijder EJ, van der Meer Y, Zevenhoven-Dobbe J, Onderwater JJ, van der Meulen J, et al. (2006) Ultrastructure and origin of membrane vesicles associated with the severe acute respiratory syndrome coronavirus replication complex. J Virol 80: 5927–5940.

39. van Hemert MJ, van den Worm SHE, Knoops K, Mommaas AM, Gorbalenya AE, et al. (2008) SARS-coronavirus replication/transcription complexes are membrane-protected and need a host factor for activity *in vitro*. PLoS Pathog 4: e1000054. doi:10.1371/journal.ppat.1000054.

40. Knoops K, Kikkert M, van den Worm SHE, Zevenhoven-Dobbe JC, van der Meer Y, et al. (2008) SARS-coronavirus replication is supported by a reticulovesicular network of modified endoplasmic reticulum. PLoS Biol 6: e226. doi:10.1371/journal.pbio.0060226.

41. Eleouet JF, Chilmonczyk S, Besnardeau L, Laude H (1998) Transmissible gastroenteritis coronavirus induces programmed cell death in infected cells through a caspase-dependent pathway. J Virol 72: 4918–4924.

42. Wu YH, Tseng CP, Cheng ML, Ho HY, Shih SR, et al. (2008) Glucose-6-Phosphate dehydrogenase deficiency enhances human coronavirus 229E infection. J Infect Dis 197: 812–816.

43. Imai Y, Kuba K, Neely GG, Yaghubian-Malhami R, Perkmann T, et al. (2008) Identification of oxidative stress and Toll-like receptor 4 signaling as a key pathway of acute lung injury. Cell 133: 235–249.

44. Xu Y, Cong L, Chen C, Wei L, Zhao Q, et al. (2009) Crystal structures of two coronavirus ADP-ribose-1″-monophosphatases and their complexes with ADP-ribose: a systematic structural analysis of the viral ADRP domain. J Virol 83: 1083–1092.

45. Wie JS, Whiteford CC, Cenacchi N, Son CG, Khan J (2005) BBC3 mediates fenretinide-induced cell death in neuroblastoma. Oncogene 24: 7976–7983.

46. Garcia-Saez I, Tcherniuk S, Kozielski F (2006) The structure of human neuronal Rab6B in the active and inactive form. Acta Crystallogr D 62: 725–733.

47. Mizutani T (2007) Signal transduction in SARS-CoV-infected cells. Ann N Y Acad Sci 1102: 86–95.

48. Jin G, Klika A, Callahan M, Faga B, Danzig J, et al. (2004) Identification of a human NF-κB-activating protein, TAB3. Proc Natl Acad Sci U S A 101: 2028–2033.

49. Hakmé A, Huber A, Dollé P, Schreiber V (2008) The macroPARP genes *Parp-9* and *Parp-14* are developmentally and differentially regulated in mouse tissues. Dev Dyn 237: 209–215.

50. Bürkle A (2005) Poly(ADP-ribose): The most elaborate metabolite of NAD$^+$. FEBS J 272: 4576–4589.

51. Schreiber V, Dantzer F, Ame JC, de Murcia G (2006) Poly(ADP-ribose): novel functions for an old molecule. Nature Rev Mol Cell Biol 7: 517–528.

52. Hassa PO, Hottiger MO (2008) The diverse biological roles of mammalian PARPs, a small but powerful family of poly-ADP-ribose polymerases. Front Biosci 13: 3046–3082.

53. Ruf A, de Murcia JM, de Murcia G, Schulz GE (1996) Structure of the catalytic fragment of poly(ADP-ribose) polymerase from chicken. Proc Natl Acad Sci U S A 93: 7481–7485.

54. Soldatenkov VA, Vetcher AA, Duka T, Ladame S (2008) First evidence of a functional interaction between DNA quadruplexes and poly(ADP-ribose) polymerase-1. ACS Chem Biol 3: 214–219.

55. van Zon A, Mossink MH, Schoester M, Houtsmuller AB, Scheffer GL, et al. (2003) The formation of vault-tubes: a dynamic interaction between vaults and vault PARP. J Cell Sci 116: 4391–4400.

56. Kerns JA, Emerman M, Malik HS (2008) Positive selection and increased antiviral activity associated with the PARP-containing isoform of human zinc-finger antiviral protein. PLoS Genet 4: e21. doi:10.1371/journal.pgen.0040021.

57. Goila-Gaur R, Strebel K (2008) HIV-1 vif, APOBEC, and intrinsic immunity. Retrovirol 5: 51.

58. Ouararhni K, Hadj-Slimane R, Ait-Si-Ali S, Robin P, Mietton F, et al. (2006) The histone variant mH2A1.1 interferes with transcription by down-regulating PARP-1 enzymatic activity. Genes Dev 20: 3324–3336.

59. Leslie AGW (1992) Recent changes to the MOSFLM package for processing film and image plate data. Joint CCP4+ESF–EAMCB Newsletter on Protein Crystallography 26: 27–33.

60. Evans PR (1993) Data reduction. In: Sawyer L, Isaacs N, Bailey S, eds (1993) Proceedings of the CCP4 Study Weekend. Data Collection and Processing. Warrington: Daresbury Laboratory. pp 114–122.

61. Collaborative Computational Project, Number 4 (1994) The CCP4 suite: Programs for protein crystallography. Acta Crystallogr D 50: 760–763.

62. Matthews BW (1968) Solvent content of protein crystals. J Mol Biol 33: 491–497.

63. Chatterjee A, Johnson MA, Serrano P, Pedrini B, Wüthrich K (2007) NMR assignment of the domain 513-651 from the SARS-CoV nonstructural protein nsp3. Biomol NMR Assign 1: 191–194.

64. McCoy AJ, Grosse-Kunstleve RW, Storoni LC, Read RJ (2005) Likelihood-enhanced fast translation functions. Acta Crystallogr D 61: 458–464.

65. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, et al. (2007) Phaser crystallographic software. J Appl Cryst 40: 658–674.

66. Bricogne G, Blanc E, Brandl M, Flensburg C, Keller P, et al. (2008) BUSTER-TNT, version 2.5.1. Cambridge, United Kingdom: Global Phasing Ltd.

67. Smart OS, Brandl M, Flensburg C, Keller P, Paciorek W, et al. (2008) Annual Meeting of the American Crystallographic Association, Knoxville, TN. Abstract TP139. pp 137.

68. Cowtan K (1994) DM: an automated procedure for phase improvement by density modification. Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography 31: 34–38.

69. Vellieux FMD, Hunt JF, Roy S, Read RJ (1995) DEMON/ANGEL: a suite of programs to carry out density modification. J Appl Cryst 28: 347–351.

70. Cowtan K (2006) The Buccaneer software for automated model building. Acta Crystallogr D 62: 1002–1011.

71. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D 53: 240–255.

72. Potterton E, Briggs P, Turkenburg M, Dodson E (2003) A graphical user interface to the CCP4 program suite. Acta Crystallogr D 59: 1131–1137.

73. Kabsch W (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. J Appl Cryst 26: 795–800.

74. Vagin A, Teplyakov A (1997) MOLREP: an automated program for molecular replacement. J Appl Cryst 30: 1022–1025.

75. Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. Acta Crystallogr D 60: 2126–2132.

76. DeLano WL The PyMOL molecular graphics system. DeLano Scientific, Palo Alto, CA, USA.

77. Abrahams JP, Kraal B, Bosch L (1988) Zone-interference gel electrophoresis: a new method for studying weak protein-nucleic acid complexes under native equilibrium conditions. Nucleic Acids Res 16: 10099–10108.

78. Weiss MS, Hilgenfeld R (1997) On the use of the merging R-factor as a quality indicator for X-ray data. J Appl Cryst 30: 203–205.