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.

Hsu et al., 2004), an A family DNA polymerase, the molecular interactions and events leading to the bypass of a mutagenic DNA lesion are emerging.

In the DNA binary complexes of A family polymerases, the template base is typically positioned outside of the DNA helix with an aromatic side chain stacking with the template base immediately upstream of the coding base (Beard and Wilson, 2000). In the *Bacillus* fragment structure, 8oG is in a *syn* conformation in this extrahelical position (Hsu et al., 2004). A binary complex with T7 DNA polymerase shows 8oG to be extrahelical as well, but disordered suggesting conformational flexibility (Brieba et al., 2004).

In the newly solved ternary complex, the K536A mutant of T7 DNA polymerase facilitates closing of the fingers subdomain with 8oG in the *syn* conformation, which is further stabilized by dATP binding. Mechanistically, insertion of dATP and subdomain opening promotes translocation to position the (*syn*)8oG-(*anti*)A base pair at the primer terminus. The *syn* conformation of 8oG positions O8 in the minor groove at an equivalent position expected for O2 of thymidine thereby preserving the minor groove hydrogen bonding interactions expected for a Watson-Crick base pair. This permits escape of the mutagenic base pair from the intrinsic proofreading exonuclease activities of these polymerases (Brieba et al., 2004; Hsu et al., 2004).

The structural/kinetic/biochemical approach utilized by Brieba et al. (2005) has provided new insights into the molecular interactions that modulate the *anti-syn* and open-closed equilibriums of 8oG and the fingers subdomain, respectively. The results underscore the many dynamic aspects of DNA synthesis, and the importance of precise template base positioning for efficient and faithful replication. It remains an exciting challenge to decipher the general as well as the specific strategies utilized by DNA polymerases to make the right decisions about substrate selection and insertion. The structure of the mutagenic base pair in the confines of a closed polymerase complex exposes some of those strategies.

William A. Beard and Samuel H. Wilson
Laboratory of Structural Biology
National Institute of Environmental Health Sciences
National Institutes of Health
Research Triangle Park, North Carolina 27709

Selected Reading

Beard, W.A., and Wilson, S.H. (2000). Mutat. Res. *460*, 231–244.

Beard, W.A., Shock, D.D., Vande Berg, B.J., and Wilson, S.H. (2002). J. Biol. Chem. *277*, 47393–47398.

Brieba, L.G., Eichman, B.F., Kokoska, R.J., Doublie, S., Kunkel, T.A., and Ellenberger, T. (2004). EMBO J. *23*, 3452–3461.

Brieba, L.G., Kokoska, R.J., Bebenek, K., Kunkel, T.A., and Ellenberger, T. (2005). Structure *13*, this issue, 1653–1659.

Collins, A.R. (1999). Bioessays *21*, 238–246.

Freisinger, E., Grollman, A.P., Miller, H., and Kisker, C. (2004). EMBO J. *23*, 1494–1505.

Furge, L.L., and Guengerich, F.P. (1997). Biochemistry *36*, 6475–6487.

Grollman, A.P., and Moriya, M. (1993). Trends Gen. *9*, 246–249.

Hsu, G.W., Ober, M., Carell, T., and Beese, L.S. (2004). Nature *431*, 217–221.

Krahn, J.M., Beard, W.A., Miller, H., Grollman, A.P., and Wilson, S.H. (2003). Structure *11*, 121–127.

Shibutani, S., Takeshita, M., and Grollman, A.P. (1991). Nature *349*, 431–434.

Structure, Vol. 13, November, 2005, DOI 10.1016/j.str.2005.10.004

SARS Proteomics Reveals Viral Secrets

Worldwide cooperative efforts to understand the biology of the SARS coronavirus have already born significant fruit. In a further advance, the X-ray structure of a domain of nonstructural protein 3 is reported by Saikatendu et al. (2005) in this issue of *Structure*.

The SARS coronavirus, a positive sense single-stranded RNA virus which emerged from China in November 2002, caused 916 deaths out of 8422 reported cases. Its transmission by one "superspreader" in a Hong Kong hotel to wider Hong Kong (195 cases), Singapore (71), Vietnam (58), Canada (29), the United States (1), and Ireland (1) was a stark global warning of the potentially devastating effects of a surprise pandemic (Monaghan, 2004). The lessons learned from the averted SARS threat are of particular relevance in November 2005, as we face the increasing likelihood of a global pandemic if the H5N1 avian influenza virus, currently spreading from the bird population of Asia, becomes adapted for human to human transmissibility.

In a worldwide cooperative research effort involving a multidisciplinary approach, structural and functional characterization of the SARS virus and its host interactions has been swiftly pursued. Four months after the first case of SARS, the virus was identified, and determination of its complete sequence was finished only 2 weeks later. The virus genome was found to code for 28 proteins, and a month after its publication, the structure of the first SARS protein, the 3CL protease, was placed in the Protein Data Bank (Anand et al., 2003). This protease represented a prime target for design of anti-SARS drugs (Yang et al., 2003) and was followed by the structures of other SARS proteins, including parts of the SARS CoV replicase polyprotein: nsP7 (Peti et al., 2005) by NMR and nsP9 (Egloff et al., 2004) by crystallography. Among other published SARS crystal structures are the N-terminal RNA binding domain of the SARS CoV nucleocapsid protein (Huang et al., 2004), the sars2 (ORF2) spike receptor binding domain

complexed with the receptor (Li et al., 2005), and the sars7a (ORF7a) transmembrane protein of so far unknown function (Nelson et al., 2005).

In this issue of *Structure*, Kuhn and coworkers (Saikatendu et al., 2005) describe the X-ray crystal structure of a domain of the nsP3 (nonstructural protein 3: the virus has 16 nsPs), the first example of this particular domain to be solved from a ssRNA virus. Intact nsP3 is a 1922 amino acid multidomain protein with an N-terminal Glu-rich domain, an ADP-ribose-1″-phosphate phosphatase (ADRP) domain, a protease active domain, the papain-like protease (PLP) domain, and finally two or three domains of unknown function. The structure of the ADRP domain, consisting of 182 amino acids, was solved using the selenium-SAD (single-wavelength anomalous dispersion) method of phase determination to 1.4 Å, a resolution that allowed detailed analysis of the molecular interactions.

The structure was determined as part of an integrated structural and functional proteomics program aimed at understanding emerging infectious diseases. Previously, little was known about the function of the SARS ADRP domain, or of its role in virulence, although it is conserved across the four known groups of coronaviridae. The tremendous impact that knowledge of such structures can have on the functional characterization of unknown but conserved "hypothetical" proteins was exemplified by this study.

The overall fold of the ADRP domain was found to be that of the three-layered $\alpha/\beta/\alpha$ core of a macro-H2A-like domain, a unique class of phosphatases that possesses the ability to dephosphorylate ADRP. By sequence alignment and structural comparison with all known H2A domains, as well as examination of functional data, the authors conjecture that proteins from this superfamily form an emerging group of nucleotide phosphatases, all with similar functionality. The putative active site of ARDP has been identified by sequence alignment with its three structural homologs; yeast Ymx7, *E. coli* hypothetical protein Er58, and AF1521 from *Archeoglobus fulgidus.* The phosphatase activity of ADRP was subsequently confirmed by in vitro experiments on enzymatically synthesized substrate (provided by E. Phizicky, University of Rochester).

This functional discovery was surprising, because the presence of such a phosphatase in SARS had not been anticipated, and also, the original identification of a phosphatase with this substrate specificity was in an analogous tRNA maturation pathway in higher eukaryotes. It raises the interesting possibility that the SARS virus has a fairly sophisticated mechanism, which perhaps involves multiple components of the replicase machinery, to process newly synthesized genomic and subgenomic RNA. If this theory is validated, it will provide another avenue for intervention by targeted small-molecule ligands that could have therapeutic potential.

The integration of structural information with insights into the virology of SARS CoV is critical for an understanding of viral infection, its lifecycle, and the potential for intervention. A complete structural picture of the proteome of this virus would impact the understanding of many other viruses of this type.

However, there are still substantial challenges ahead, not least because structural characterization of viral proteins presents a number of severe obstacles. First, there is a paucity of genetic information available to assist with the identification of the putative protein function, since rather few viruses have been fully characterized. Second, the protocols for identifying domains, a prerequisite for expressing and purifying crystallization-grade protein, are often unsuccessful. Third, for a virus such as SARS with only 28 proteins, the success rate for structural genomics must be high for understanding of the viral biology to advance, unlike for a bacterial system where there are a many more proteins to investigate and a higher attrition rate in the pipeline can be tolerated. Last but not least, viruses are finely tuned biological systems with a large number of interdependent molecular interactions and little redundancy. This has two pivotal consequences for understanding the biology of the virus: A systematic approach is essential, and, even more importantly, a deeper structural and functional knowledge of the many complexes that the SARS CoV proteins form with one another and with proteins of the host organisms will be required—research that is still in its infancy.

Following the development of suitable antiviral drugs and vaccines, a global strategy for their administration must be agreed upon and implemented. A similar public health challenge is currently being faced for H5N1 influenza pandemic planning by governments worldwide.

Elspeth Garman
Department of Biochemistry
University of Oxford
South Parks Road
Oxford OX1 3QU
United Kingdom

Selected Reading

Anand, K., Ziebuhr, J., Wadhwani, P., Mesters, J.R., and Hilgenfeld, R. (2003). Science *300*, 1763–1767.

Egloff, M.P., Ferron, F., Campanacci, V., Longhi, S., Rancurel, C., Dutartre, H., Snijder, E.J., Gorbalenya, A.E., Cambillau, C., and Canard, B. (2004). Proc. Natl. Acad. Sci. USA *101*, 3792–3796.

Huang, Q., Yu, L., Petros, A.M., Gunasekera, A., Liu, Z., Xu, N., Hajduk, P., Mack, J., Fesik, S.W., and Olejniczak, E.T. (2004). Biochemistry *43*, 6059–6063.

Li, F., Li, W., Farzan, M., and Harrison, S.C. (2005). Science *309*, 1864–1868.

Monaghan, K. (2004). In Learning from SARS. S. Knobler, A. Mahmoud, S. Lemon, A. Mack, L. Sivitz, and K. Oberholtzer, eds. (Washington, DC: The National Academies Press), pp. 246–276.

Nelson, C.A., Pekosz, A., Lee, C.A., Diamond, M.S., and Fremont, D.H. (2005). Structure *13*, 75–85.

Peti, W., Johnson, M.A., Herrmann, T., Neuman, B.W., Buchmeister, M.J., Nelson, M., Joseph, J., Page, R., Stevens, R.C., Kuhn, P., and Wüthrich, K. (2005). J. Virol. *79*, 12905–12913.

Saikatendu, K.S., Joseph, J.S., Subramanian, M.S., Clayton, T., Griffith, M., Moy, K., Velasquez, J., Neuman, B.W., Buchmeier, M.J., Stevens, R.C., and Kuhn, P.K. (2005). Structure *13*, this issue, 1665–1675.

Yang, H., Yang, M., Ding, Y., Liu, Y., Lou, Z., Zhou, Z., Sun, L., Ye, S., Pang, H., Gao, G.F., et al. (2003). Proc. Natl. Acad. Sci. USA *100*, 13190–13195.