

# International research networks in viral structural proteomics: Again, lessons from SARS

Bruno Canard<sup>a</sup>, Jeremiah S. Joseph<sup>b</sup>, Peter Kuhn<sup>b,\*</sup>

<sup>a</sup> *Architecture Et Fonction Des Macromolécules Biologiques UMR 6098, CNRS, Universités Aix-Marseille I & II,  
Case 932, 163 Avenue de Luminy, 13288 Marseille Cedex 9, France*

<sup>b</sup> *Department of Cell Biology, The Scripps Research Institute, 10550 N Torrey Pines Road,  
CB265 La Jolla, CA 92037, USA*

Received 10 August 2007; accepted 24 September 2007

## Abstract

Emerging and re-emerging pathogens and bioterror threats require an organized and coherent response from the worldwide research community to maximize available resources and competencies with the primary goals to understand the pathogen and enable intervention. In 2001, the Structural Proteomics In Europe (SPINE) project prototyped the pan-viral structural genomic approach, and the Severe Acute Respiratory Syndrome (SARS) outbreak in 2003 accelerated the concept of structural characterization of all proteins from a viral proteome and the interaction with their host partners. Following that approach, in 2004 the center for Functional and Structural Proteomics for SARS-CoV related proteins was initiated as part of the US NIH NIAID proteomics resource centers. Across worldwide efforts in Asia, Europe and America, the international research teams working on SARS-CoV have now determined experimental structural information for 45% of the SARS-CoV proteins and 53% of all its soluble proteins. This data is fully available to the scientific community and is providing an unprecedented level of insight to this class of RNA viruses. The efforts and results by the international scientific community to the SARS outbreak are serving as an example and roadmap of a rapid response using modern research methods.

© 2007 Elsevier B.V. All rights reserved.

**Keywords:** SARS-CoV; Infectious diseases; Structural genomics; FSPS; VIZIER

## 1. Introduction

### 1.1. The proof-of-concept: SARS-CoV and its proteome

Severe Acute Respiratory Syndrome (SARS) is the first severe and readily transmissible new disease to emerge in the 21st century. Although much about SARS remains poorly understood, it is clear that this disease has major implications, both for public health and the global economy. SARS emerged amongst the Chinese population in mid November 2002, but did not leave Guangdong Province until mid February 2003, when an infected individual brought the virus to a hotel in Hong Kong. From there, hotel guests rapidly disseminated the disease to distant places, such as Singapore, Vietnam and Canada. On August 7, 2003, 8422 cases were confirmed worldwide with overall case fatalities estimated at 14–15%; in persons over 50, the fatal-

ity exceeded 50%. SARS dramatically demonstrates the global havoc wreaked by a newly emerging infectious disease.

The World Health Organization coordinated an international investigation through its Global Outbreak Alert and Response Network (<http://www.who.int/csr/outbreaknetwork/en/>) and worked closely with health authorities in the affected countries to provide epidemiological, clinical and logistical support. Interestingly, the SARS outbreak came at a critical point when the scientific community was actively launching structural genomics projects (Lesley et al., 2002). The large size of the SARS-CoV RNA genome (~30 kb) justified a structural genomic approach in itself. A number of research groups around the world responded immediately to the challenge of SARS being the first pandemic of the new millennium. Four months after the outbreak, the virus was identified as a new class of coronavirus (Peiris et al., 2003), 2 weeks later the genome sequencing was completed (Marra et al., 2003; Rota et al., 2003; Zeng et al., 2003). The first structural model of a SARS-CoV protein was reported within another month (Anand et al., 2003) and the initial description of the proteome was reported a few months later

\* Corresponding author. Tel.: +1 858 784 9114; fax: +1 858 784 8996.  
E-mail address: [pkuhn@scripps.edu](mailto:pkuhn@scripps.edu) (P. Kuhn).

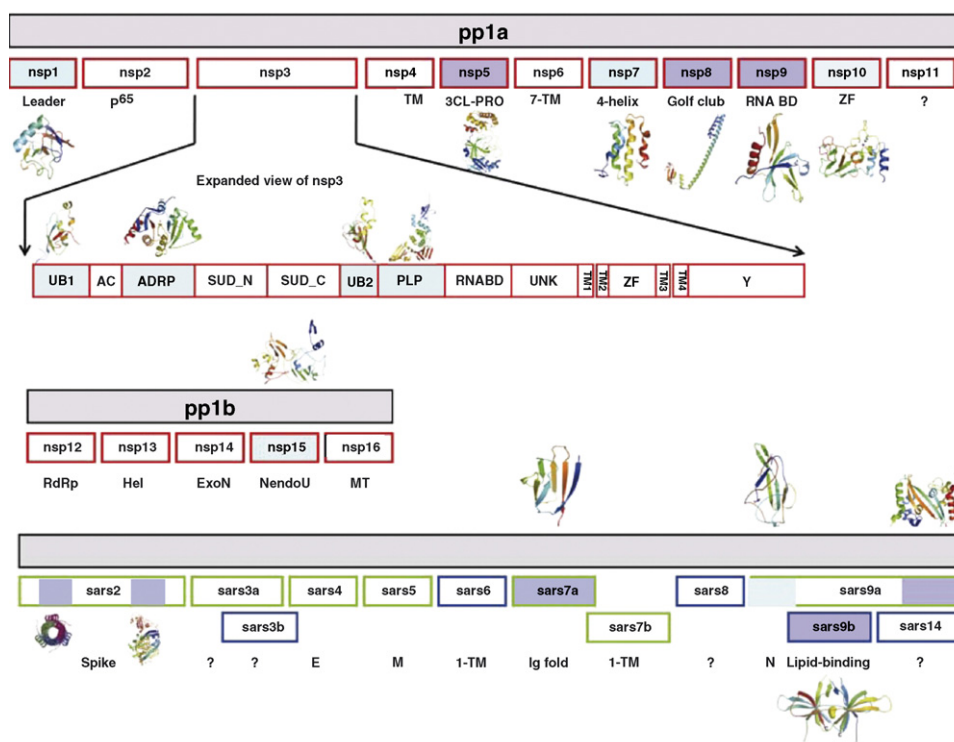


Fig. 1. A schematic depicting the progress made in structural characterization of the SARS-CoV proteome. The top two panels show the mature proteins that emerge from the replicase polyprotein while the bottom panel shows the ORFs for structural and accessory proteins. The individual domains of nsp3 derived by *in silico* analysis are shown in the inset. Those proteins whose structures have been described in the literature are shown as ribbon diagrams that are color coded from the N- (red) to the C-terminus (blue).

(Snijder et al., 2003; Thiel et al., 2003). Since then, a number of groups around the world have worked towards the complete structural description of the proteome and all related functions (Almeida et al., 2007; Campanacci et al., 2003; Imbert et al., 2006; Joseph et al., 2006, 2007; Li et al., 2005; Meier et al., 2006; Nelson et al., 2005; Peti et al., 2005; Ratia et al., 2006; Ricagno et al., 2006; Saikatendu et al., 2005, 2007; Serrano et al., 2006; Sutton et al., 2004; Su et al., 2006; Zhai et al., 2005). While a number of important contributions were made by individual groups, a very systematic approach was taken by the European, US and Asian structural proteomics teams. Fig. 1 shows the current state of the structural biology of SARS-CoV related proteins.

The coordination amongst the principal investigators of the research teams involved has led to efficient communication of progress and exchange of reagents. Updates on progress and joint publication strategies have effectively leveraged the competencies in each of the teams. In the US, the participation in the systematic investigation of the structures of all proteins in the proteome was primarily through the NIH NIAID Proteomics Resource Center for Biodefense and Emerging Infectious Diseases program and its support of the center for Functional and Structural Proteomics of SARS-CoV (FSPS) related proteins (<http://visp.scripps.edu/sites/sars>; <http://www.proteomicsresource.org>). The aim was to characterize the complete proteome of SARS-CoV related proteins both structurally and functionally to provide initial validation to targets for drug discovery and vaccine development. This sys-

tematic approach is providing the coronavirus community with structural insight at an unprecedented rate leading constantly to new scientific findings.

Critical to this success in structural elucidation was the worldwide participation of exceptional competencies and effective collaboration amongst these experts in the field. In Europe, the research groups participating through Structural Proteomics In Europe (SPINE) (<http://www.spineurope.org/>) and VIZIER (<http://www.vizier-europe.org/>) included Professors Canard, Hilgenfeld, and Stuart. SPINE was one of the earliest funded integrated research projects (October 2003–2006) bringing together top European structural biology institutions in an unprecedented collaborative effort to develop new methods and technologies for high-throughput structural biology. SPINE was funded within Framework V of the European Commission and coordinated by Professor Dave Stuart of the Division of Structural Biology, University of Oxford. The project ran for 3 years from October 1, 2003 and involved more than 20 partners in 11 European countries working together to build a standard pipeline for identifying protein structures important for new drug discovery. VIZIER was funded through Framework VI of the European Commission and focuses on viral enzymes involved in replication. In addition, the European Commission supported SARS-CoV research through several Specific Targeted Research Projects. Two of these also made important contributions to SARS-CoV structural genomics: SEPSDA (<http://www.sepsda.biochem.uni-luebeck.de>), which was coordinated by Professor Rolf Hilgenfeld at the University of

Lübeck, and SARS-DTV (<http://www.sars-dtv.nl>), coordinated by Professor Willy Spaan and Eric Snijder at Leiden University Medical Center. In the US, the FSPS program of the NIAID Proteomics Resource Center involves the laboratories of Professors Kuhn, Buchmeier, Stevens, Wüthrich, and Wilson.

This consortia network is then expanded with individual research teams including Professor Zihe Rao and colleagues of Tsinghua University in Beijing and Professor Tai-huang Huang and colleagues of Taiwan's Academia Sinica and Professor Andrew Mesecar of the University of Chicago, USA and Professor Richard Kuhn of Purdue University, USA.

The coordination amongst these groups has formal and informal aspects. For example, the structure and function of the PLP Protease was jointly determined by the FSPS team and the Mesecar group combining expertise in nanovolume crystallization and protease enzymology. Publications on the non-structural protein 10 were coordinated amongst the Rao and FSPS teams and following the initial work on nsp15 by the VIZIER team, the reagents were exchanged with the FSPS and Rao teams to allow the structural elucidation and interpretation in a different space group and functional state. While the senior investigators frequently converse via electronic mail and at meetings, the students in the labs are also frequently exchanging information and working jointly on projects. Two recent graduates from the Rao laboratory started in the Kuhn laboratory at The Scripps Research Institute and lab members from the Mesecar lab just presented their recent findings at the bi-weekly FSPS meeting.

Formal meetings on the topic of emerging and re-emerging infectious diseases such as the workshop on Discovery of Antiviral compounds in Lübeck, 2006 are used, where members from EU, US and China present and share their progress. The annual meeting of the American Society of Microbiology has also served as a forum for exchange of both, published and unpublished data. As a result of the efforts by the NIH NIAID, reagents and procedures are being deposited to and made available by the BEI Resources [<http://www.beiresources.org>].

## 1.2. Technologies

The last decade has seen a rapid development of miniaturized, automated and integrated methods and instrumentation development for structural biology (Lesley et al., 2002). While initially only applied to simple prokaryotic organisms (Lesley et al., 2002), it is now clear that these approaches are critical to the rapid understanding of complete viral organisms at the structural and functional level in support of developing means of therapeutic intervention. Clearly, these methods are being pushed to their limits and continued development is required that will ultimately enable individual laboratories and large international teams to tackle viral proteomes of both, emerging and re-emerging diseases. Many of the participating research groups are constantly developing, improving and sharing new technologies (Brooun et al., 2007; Stuart et al., 2006) to more reliably and faster clone, express, purify, crystallize, and functionally characterize viral proteins. A significant advancement of what was accomplished over the past 2 years through the biology-guided

approach, which demonstrated that many of these methods and instrumentation had reached a level of maturity so that they could be applied rapidly to new problems. Intellectual drivers for these experiments have, in many cases, been the virologists. This development will likely prevail over the coming decade as more and more of the traditionally specialized resource intensive experiments such as mammalian or baculoviral expression systems, automated high-throughput crystallization and the use of international and local synchrotron sources have been streamlined and implemented for widespread laboratory use. Examples like miniaturized and automated crystallization and imaging systems as well as next generation crystallization geometries have demonstrated how new technologies can accelerate discovery on challenging targets. The next wave of these technology developments is now being driven by new roadblocks in the quest for high content structural biology on challenging targets such as viral proteins and their host partners.

## 2. Conclusion—future prospect

Pathogenic viruses do not know any frontiers. Perhaps the greatest boost to viral structural genomics was the availability of the SARS-CoV sequence on the internet, as well as the openness of the small but dynamic coronavirus research world. Strains, sequences, clones, databases and all sorts of viral resources should become more easily available to appropriate labs, and large-scale efforts such as those ongoing in VIZIER and FSPS should be coordinated and bridged to other groups worldwide. (Investigators interested in joining this collaborative network are invited to contact the authors.) Structural proteomics has the great advantage of delivering atomic coordinates that can be quickly disseminated and not only used to considerably speed-up crystallographic problems but also promote cooperation. Given the importance of these pathogens and the possible worldwide impact of an outbreak, it is critical that the scientific community takes responsibility for addressing the needs for identifying and validating targets of therapeutic intervention. In this respect, the short response time of such big consortia has been truly remarkable in providing, for example, key structural data on SARS-CoV enzymes a prerequisite for expediting state-of-the-art drug design. This has been clearly demonstrated by the complementary work of the Canard, Rao and Kuhn labs on nsp15, a critical coronavirus endoribonuclease (Joseph et al., 2006; Ricagno et al., 2006; Xu et al., 2006).

Programs like VIZIER and FSPS are laying the foundation for any emerging virus and are delivering first line research products for the discovery of vaccines and therapeutics. The availability of reagents and methods for the production of the viral proteins, the high-resolution structural information, the functional and high-throughput assays, and small molecule tool compounds provide the starting point for drug discovery efforts in both, academic and pharmaceutical settings.

## Acknowledgements

The authors thank their many collaborators and colleagues who contributed to the advances in functional and structural

proteomics of the SARS-CoV proteins and their host interactions. We thank Elvia Nunez for editorial assistance. This short summary is by no means comprehensive and there are many more research groups working in various aspects of developing a comprehensive understanding of the SARS-CoV proteome and validating individual drug targets. This is TSRI manuscript number 19099. JJ and PK are supported by through NIH-NIAID Contract HHSN266200400058C. Work in the laboratory of BC was supported (in part) by the European Union 6th Framework Programme (FP6) in the context of the activities of the Euro-Asian SARS-DTV Network (SP22-CT-2004-511064) and the VIZIER integrated project (LSHG-CT-2004-511960).

## References

- Almeida, M.S., Johnson, M.A., Herrmann, T., Geralt, M., Wuthrich, K., 2007. Novel beta-barrel fold in the nuclear magnetic resonance structure of the replicase nonstructural protein 1 from the severe acute respiratory syndrome coronavirus. *J. Virol.* 81, 3151–3161.
- 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.
- Brooun, A., Foster, S.A., Chrencik, J.E., Chien, E.Y., Kolatkar, A.R., Streiff, M., Ramage, P., Widmer, H., Weckbecker, G., Kuhn, P., 2007. Remedial strategies in structural proteomics: expression, purification, and crystallization of the Vav1/Rac1 complex. *Protein Expr. Purif.* 53, 51–62.
- Campanacci, V., Egloff, M.P., Longhi, S., Ferron, F., Rancurel, C., Salomoni, A., Durousseau, C., Tocque, F., Bremond, N., Dobbe, J.C., Snijder, E.J., Canard, B., Cambillau, C., 2003. Structural genomics of the SARS coronavirus: cloning, expression, crystallization and preliminary crystallographic study of the Nsp9 protein. *Acta Crystallogr. D. Biol. Crystallogr.* 59, 1628–1631.
- Imbert, I., Guillemot, J.C., Bourhis, J.M., Bussetta, C., Coutard, B., et al., 2006. A second, non-canonical RNA-dependent RNA polymerase in SARS coronavirus. *EMBO J.* 25, 4933–4942.
- Joseph, J.S., Saikatendu, K.S., Subramanian, V., Neuman, B.W., Brooun, A., Griffith, M., Moy, K., Yadav, M.K., Velasquez, J., Buchmeier, M.J., Stevens, R.C., Kuhn, P., 2006. Crystal structure of nonstructural protein 10 from the severe acute respiratory syndrome coronavirus reveals a novel fold with two zinc-binding motifs. *J. Virol.* 80, 7894–7901.
- Joseph, J.S., Saikatendu, K.S., Subramanian, V., Neuman, B.W., Buchmeier, M.J., Stevens, R.C., Kuhn, P., 2007. Crystal structure of a monomeric form of severe acute respiratory syndrome coronavirus endonuclease nsp15 suggests a role for hexamerization as an allosteric switch. *J. Virol.* 81, 6700–6708.
- Lesley, S.A., Kuhn, P., Godzik, A., Deacon, A.M., Mathews, I., Kreusch, A., Spraggon, G., Klock, H.E., McMullan, D., Shin, T., Vincent, J., Robb, A., Brinen, L.S., Miller, M.D., McPhillips, T.M., Miller, M.A., Scheibe, D., Canaves, J.M., Guda, C., Jaroszewski, L., Selby, T.L., Elsliger, M.A., Woolley, J., Taylor, S.S., Hodgson, K.O., Wilson, I.A., Schultz, P.G., Stevens, R.C., 2002. Structural genomics of the *Thermotoga maritima* proteome implemented in a high-throughput structure determination pipeline. *Proc. Natl. Acad. Sci. U.S.A.* 99, 11664–11669.
- Li, F., Li, W., Farzan, M., Harrison, S.C., 2005. Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. *Science* 309, 1864–1868.
- Marra, M.A., et al., 2003. The Genome sequence of the SARS-associated coronavirus. *Science* 300, 1399–1404.
- Meier, C., Aricescu, A.R., Assenberg, R., Aplin, R.T., Gilbert, R.J., Grimes, J.M., Stuart, D.I., 2006. The crystal structure of ORF-9b, a lipid binding protein from the SARS coronavirus. *Structure* 14, 1157–1165.
- Nelson, C.A., Pekosz, A., Lee, C.A., Diamond, M.S., Fremont, D.H., 2005. Structure and intracellular targeting of the SARS-coronavirus Orf7a accessory protein. *Structure* 13, 75–85.
- Peiris, J.S., Lai, S.T., Poon, L.L., Guan, Y., Yam, L.Y., Lim, W., Nicholls, J., Yee, W.K., Yan, W.W., Cheung, M.T., et al., 2003. Coronavirus as a possible cause of severe acute respiratory syndrome. *Lancet* 361, 1319–1325.
- Peti, W., Johnson, M.A., Herrmann, T., Neuman, B.W., Buchmeier, M.J., Nelson, M., Joseph, J., Page, R., Stevens, R.C., Kuhn, P., Wuthrich, K., 2005. Structural genomics of the severe acute respiratory syndrome coronavirus: nuclear magnetic resonance structure of the protein nsP7. *J. Virol.* 79, 12905–12913.
- Ratia, K., Saikatendu, K.S., Santarsiero, B.D., Barretto, N., Baker, S.C., Stevens, R.C., Mesecar, A.D., 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.
- Ricagno, S., Egloff, M.P., Ulferts, R., Coutard, B., Nurizzo, D., Campanacci, V., Cambillau, C., Ziebuhr, J., Canard, B., 2006. Crystal structure and mechanistic determinants of SARS coronavirus nonstructural protein 15 define an endoribonuclease family. *Proc. Natl. Acad. Sci. U.S.A.* 103, 11892–11897.
- Rota, P.A., et al., 2003. Characterization of a novel coronavirus associated with severe acute respiratory syndrome. *Science* 300, 1394–1399.
- Saikatendu, K.S., Joseph, J.S., Subramanian, V., Clayton, T., Griffith, M., Moy, K., Velasquez, J., Neuman, B.W., Buchmeier, M.J., Stevens, R.C., Kuhn, P., 2005. Structural basis of severe acute respiratory syndrome coronavirus ADP-ribose-1''-phosphate dephosphorylation by a conserved domain of nsP3. *Structure* 13, 1665–1675.
- Saikatendu, K.S., Joseph, J.S., Subramanian, V., Neuman, B.W., Buchmeier, M.J., Stevens, R.C., Kuhn, P., 2007. Ribonucleocapsid formation of severe acute respiratory syndrome coronavirus through molecular action of the N-terminal domain of N protein. *J. Virol.* 81, 3913–3921.
- Serrano, P., Almeida, M.S., Johnson, M.A., Wuthrich, K., 2006. NMR assignment of the protein nsp3a from SARS-CoV. *J. Biomol. NMR* 36, 45.
- 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.
- Stuart, D., Jones, Y.E., Wilson, K.S., Daenke, S., 2006. Structural proteomics in Europe. *Acta D* 62 Pt. 10.
- Su, D., Lou, Z., Sun, F., Zhai, Y., Yang, H., Zhang, R., Joachimiak, A., Zhang, X.C., Bartlam, M., Rao, Z., 2006. Dodecamer structure of severe acute respiratory syndrome coronavirus nonstructural protein nsp10. *J. Virol.* 80, 7902–7908.
- Sutton, G., Fry, E., Carter, L., Sainsbury, S., Walter, T., Nettleship, J., Berrow, N., Owens, R., Gilbert, R., Davidson, A., Siddell, S., Poon, L.L., Diprose, J., Alderton, D., Walsh, M., Grimes, J.M., Stuart, D.I., 2004. The nsp9 replicase protein of SARS-coronavirus, structure and functional insights. *Structure* 12, 341–353.
- Thiel, V., Ivanov, K.A., Putics, A., Hertzog, 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.
- Xu, X., Zhai, Y., Sun, F., Lou, Z., Su, D., Xu, Y., Zhang, R., Joachimiak, A., Zhang, X.C., Bartlam, M., Rao, Z., 2006. New antiviral target revealed by the hexameric structure of mouse hepatitis virus nonstructural protein nsp15. *J. Virol.* 80, 7909–7917.
- Zhai, Y., Sun, F., Li, X., Pang, H., Xu, X., Bartlam, M., Rao, Z., 2005. Insights into SARS-CoV transcription and replication from the structure of the nsp7-nsp8 hexadecamer. *Nat. Struct. Mol. Biol.* 12, 980–986.
- Zeng, F.Y., et al., 2003. The complete genome sequence of severe acute respiratory syndrome coronavirus strain HKU-39849 (HK-39). *Exp. Biol. Med.* (Maywood) 228, 866–873.