**March 15, 2021**

**General User proposal to access NCCAT-EM facility at SEMC, New York**

**Title: Structure and mechanism of large protein-nucleic acid complexes in cancer and viral pathogenesis**

**PI: Yogesh K. Gupta, Ph.D,** Assistant Professor, University of Texas Health Science Center, San Antonio, TX

**Project #1 Structures and mechanism of SARS-CoV-2 nonstructural protein complexes**

The massive global pandemic with high morbidity and mortality makes severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) one of the deadliest viruses in recent history[1]. To develop effective therapies, we need a better understanding of the mechanisms that permit the virus to invade cells and evade host immune restriction. SARS-CoV-2 is an enveloped, positive-sense single-stranded β-coronavirus with a large, complex RNA genome[2]. The RNA and RNA Cap synthesis complex (**RCS**) of CoVs comprises of eight nonstructural proteins (nsp) (**Fig. 1A**). Among them, nsp7, nsp8, and nsp12 processively replicate the RNA genome and accomplish sub-genomic transcription of SARS-CoV-2 genes[3-6] To evade the immune restriction and hijack the host translation machinery for propagation, enzymes encoded by the genome of coronaviruses (CoVs) modify the 5'-end of virally encoded mRNAs by creating a RNA cap. RNA capping and modification in CoVs involves activities of several nsp: nsp12, an RNA dependent RNA polymerase (RdRP), nsp13, a bifunctional RNA/NTP triphosphatase and helicase; nsp14, a bifunctional 3'→5' mismatch exonuclease and mRNA cap guanine-N7 methyltransferase; nsp16, a ribose 2'-O methyltransferase; and an elusive guanylyl transferase[7, 8]. The nsp16 and nsp10 covalently attach a methyl group to 2'-OH of ribose sugar of the first transcribing nucleotide (N1) of SARS-CoV-2 mRNA. This modification converts the Cap-0 (m7GppA) to Cap-1 (m7GpppAm), blocking the stimulation of interferon response and thereby helps the virus to evade immune restriction[9]. Nsp10 acts as an allosteric stimulator for nsp14 and nsp16[10] (**Fig. 1A**). We purified all major components of the RCS machinery of SARS-CoV-2 (**Fig. 1B**) and elucidated the first structure of nsp16/nsp10 complex in presence of an RNA cap ($^{m7}$GpppA) and a methyl donor *S*-adenosyl-L-methionine (SAM) to ~2.4Å resolution by X-ray crystallography (**Fig. 1D-E**). These results were recently published in *Nature Communications*[11]. Our long-term goal is to better understand the specific steps and interplay of nsp proteins during RNA synthesis and cap formation and their targeting by novel small molecules. By employing a cutting-edge machine learning method, we screened several million small molecules, which could potentially inhibit the activity of nsp proteins. We identified several lead compounds for further evaluation of their mode of action. In the next phase of drug development, we will combine the structural biology methods (X-ray crystallography and Cryo-EM) and classical biochemistry and mass spectrometry-based assays to better understand their mechanism of action. We have established a collaboration with **Dr. Doug Frantz**, an accomplished medicinal chemist with >20 years of industrial and academic experience in drug discovery program at the University of Texas San Antonio. The antiviral activity of the compounds will be tested in collaboration with **Dr. Luis Martinez-Sobrido,** an accomplished virologist at the Texas Biomedical Research Institute, San Antonio.

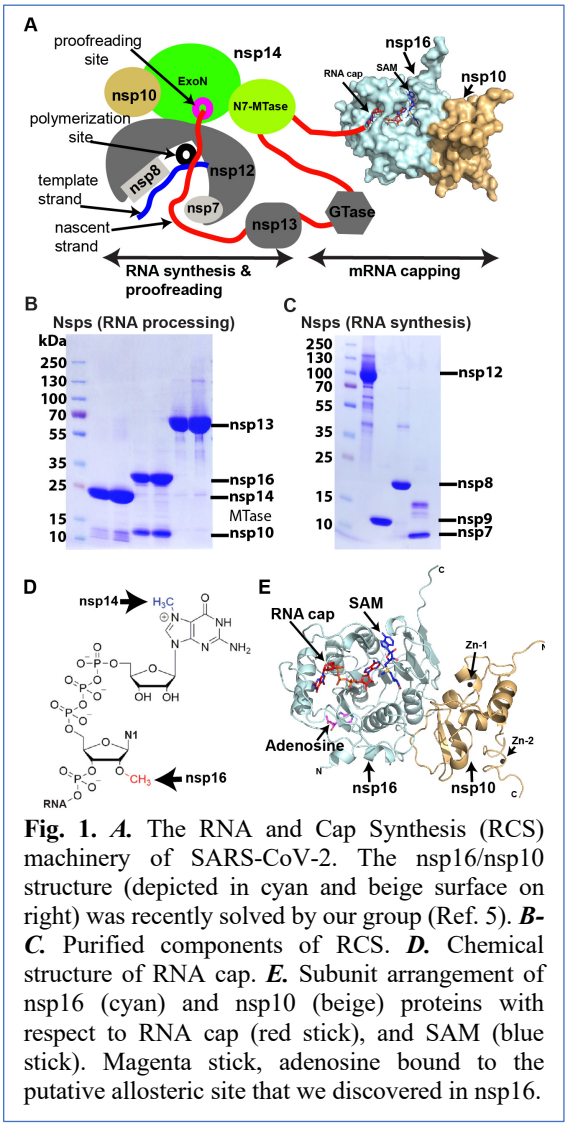


**Fig. 1.** ***A.*** The RNA and Cap Synthesis (RCS) machinery of SARS-CoV-2. The nsp16/nsp10 structure (depicted in cyan and beige surface on right) was recently solved by our group (Ref. 5). ***B-C.*** Purified components of RCS. ***D.*** Chemical structure of RNA cap. ***E.*** Subunit arrangement of nsp16 (cyan) and nsp10 (beige) proteins with respect to RNA cap (red stick), and SAM (blue stick). Magenta stick, adenosine bound to the putative allosteric site that we discovered in nsp16.

The successful outcome of our research will also further our understanding of the mechanisms by which SARS-CoV-2 evade immune restriction in human cells. This new knowledge will be combined with the structure-guided rational design of new class of antiviral agents for treatment of COVID-19 and emerging coronavirus illnesses.

**Project #2 Structure and Mechanism of human BAF complexes in cancer**

The chromatin remodeler BAF (BRG1 or BRM-associated factors) and PBAF (polybromo-associated BAF) complexes represent combinatorial assembly of up to 15 different proteins, including an essential SWI/SNF ATP motor subunit BRG1

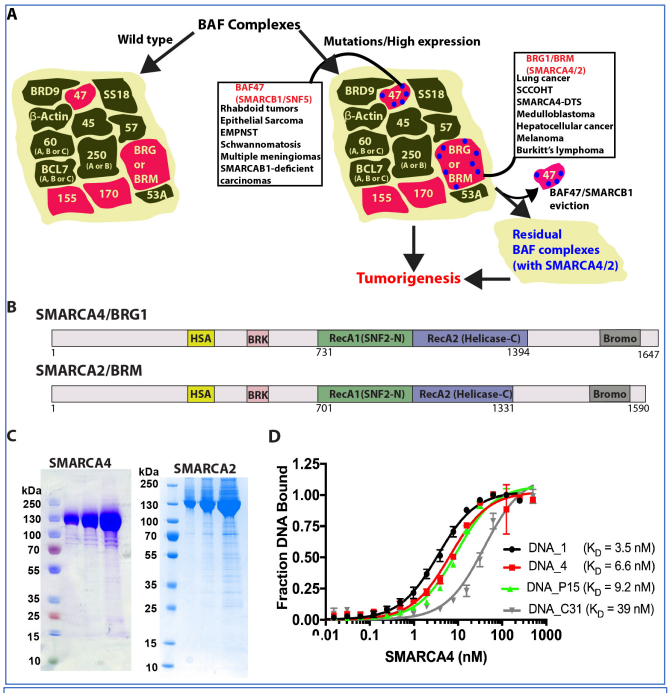


**Fig. 2. A.** Composition of human BAF. The four BAF subunits (SMARCA4/2, BAF47 or SMARCAB1, BAF155, BAF170; in red) constitute the functional chromatin remodeling core and are frequently mutated in different cancers (blue spheres). **B.** Domain architecture of the core ATPase subunit of BAFs, SMARCA4 (BRG1) and SMARCA2 (BRM). **C.** Commassie-stained gels of purified SMARCA4 and 2 enzymes after final size-exclusion column. **D.** FP-based assay identified a high affinity DNA substrate (DNA_1) for SMARCA4. Kd, equilibrium dissociation constant.

(or SMARCA4) or its biochemically redundant paralogue BRM (or SMARCA2)[12]. The SMARCA4/2-mediated chromatin remodeling activity facilitates recruitment of transcription factors, enhancer activation, and eviction of other BAF component(s) drive oncogenesis in various cancers[12, 13]. The core module of the BAF complex with only four subunits (SMARCA4/2, BAF47, BAF155, and BAF170) is capable of restoring efficient nucleosome remodeling activity *in vitro* with SMARCA4 serving as a minimal ATPase core required for this function[14] (**Fig. 2**). About 20% of all cancers harbor defective chromatin remodelers[15, 16]. The oncogenic role of BAF is also attributed to its perturbed function due to activity of residual BAF complexes[17, 18]. The mode of assembly of these 'aberrant' or 'residual' BAF complexes, however, remains unknown, partly due to lack of systematic biochemical/biophysical characterization. **Our studies may** provide a firm biophysical understanding of the ATPase core module and effects of cancer associated mutations on BAF assembly and activity. We have successfully purified the human SMARCA2 and SMARCA4 ATP motors from insect cells, identified high affinity DNA targets for highest ATPase activity. We propose to study these complexes by cryoEM. This study, if successful, will also systematically explore the architecture of the residual/aberrant BAFs as novel therapeutic regimen(s) for pediatric sarcomas. We have reconstituted the core module of BAF in vitro for EM studies. We will first optimize the sample conditions and measure cryo-EM data on most stable samples containing SMARC4 and/or SMARCA2 containing BAFs.

## References


[1] Wu, F., Zhao, S., Yu, B., Chen, Y. M., Wang, W., Song, Z. G., Hu, Y., Tao, Z. W., Tian, J. H., Pei, Y. Y., Yuan, M. L., Zhang, Y. L., Dai, F. H., Liu, Y., Wang, Q. M., Zheng, J. J., Xu, L., Holmes, E. C., and Zhang, Y. Z. (2020) A new coronavirus associated with human respiratory disease in China, *Nature 579*, 265-269.

[2] Zhou, P., Yang, X. L., Wang, X. G., Hu, B., Zhang, L., Zhang, W., Si, H. R., Zhu, Y., Li, B., Huang, C. L., Chen, H. D., Chen, J., Luo, Y., Guo, H., Jiang, R. D., Liu, M. Q., Chen, Y., Shen, X. R., Wang, X., Zheng, X. S., Zhao, K., Chen, Q. J., Deng, F., Liu, L. L., Yan, B., Zhan, F. X., Wang, Y. Y., Xiao, G. F., and Shi, Z. L. (2020) A pneumonia outbreak associated with a new coronavirus of probable bat origin, *Nature 579*, 270-273.

[3] Posthuma, C. C., Te Velthuis, A. J. W., and Snijder, E. J. (2017) Nidovirus RNA polymerases: Complex enzymes handling exceptional RNA genomes, *Virus Res 234*, 58-73.

[4] Snijder, E. J., Decroly, E., and Ziebuhr, J. (2016) The Nonstructural Proteins Directing Coronavirus RNA Synthesis and Processing, *Adv Virus Res 96*, 59-126.

[5] Sola, I., Almazan, F., Zuniga, S., and Enjuanes, L. (2015) Continuous and Discontinuous RNA Synthesis in Coronaviruses, *Annu Rev Virol 2*, 265-288.

[6] Yin, W., Mao, C., Luan, X., Shen, D. D., Shen, Q., Su, H., Wang, X., Zhou, F., Zhao, W., Gao, M., Chang, S., Xie, Y. C., Tian, G., Jiang, H. W., Tao, S. C., Shen, J., Jiang, Y., Jiang, H., Xu, Y., Zhang, S., Zhang, Y., and Xu, H. E. (2020) Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV-2 by remdesivir, *Science*.

[7] Decroly, E., Ferron, F., Lescar, J., and Canard, B. (2011) Conventional and unconventional mechanisms for capping viral mRNA, *Nat Rev Microbiol 10*, 51-65.

[8] Yan, L., Ge, J., Zheng, L., Zhang, Y., Gao, Y., Wang, T., Huang, Y., Yang, Y., Gao, S., Li, M., Liu, Z., Wang, H., Li, Y., Chen, Y., Guddat, L. W., Wang, Q., Rao, Z., and Lou, Z. (2021) Cryo-EM Structure of an Extended SARS-CoV-2 Replication and Transcription Complex Reveals an Intermediate State in Cap Synthesis, *Cell 184*, 184-193 e110.

[9] Daffis, S., Szretter, K. J., Schriewer, J., Li, J., Youn, S., Errett, J., Lin, T. Y., Schneller, S., Zust, R., Dong, H., Thiel, V., Sen, G. C., Fensterl, V., Klimstra, W. B., Pierson, T. C., Buller, R. M., Gale, M., Jr., Shi, P. Y., and Diamond, M. S. (2010) 2'-O methylation of the viral mRNA cap evades host restriction by IFIT family members, *Nature 468*, 452-456.

[10] Bouvet, M., Debarnot, C., Imbert, I., Selisko, B., Snijder, E. J., Canard, B., and Decroly, E. (2010) In vitro reconstitution of SARS-coronavirus mRNA cap methylation, *PLoS Pathog 6*, e1000863.

[11] Viswanathan, T., Arya, S., Chan, S. H., Qi, S., Dai, N., Misra, A., Park, J. G., Oladunni, F., Kovalskyy, D., Hromas, R. A., Martinez-Sobrido, L., and Gupta, Y. K. (2020) Structural basis of RNA cap modification by SARS-CoV-2, *Nat Commun 11*, 3718.

[12] St Pierre, R., and Kadoch, C. (2017) Mammalian SWI/SNF complexes in cancer: emerging therapeutic opportunities, *Curr Opin Genet Dev 42*, 56-67.

[13] Wu, Q., Lian, J. B., Stein, J. L., Stein, G. S., Nickerson, J. A., and Imbalzano, A. N. (2017) The BRG1 ATPase of human SWI/SNF chromatin remodeling enzymes as a driver of cancer, *Epigenomics 9*, 919-931.

[14] Phelan, M. L., Sif, S., Narlikar, G. J., and Kingston, R. E. (1999) Reconstitution of a core chromatin remodeling complex from SWI/SNF subunits, *Mol Cell 3*, 247-253.

[15] Kadoch, C., Hargreaves, D. C., Hodges, C., Elias, L., Ho, L., Ranish, J., and Crabtree, G. R. (2013) Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy, *Nat Genet 45*, 592-601.

[16] Shain, A. H., and Pollack, J. R. (2013) The spectrum of SWI/SNF mutations, ubiquitous in human cancers, *PLoS One 8*, e55119.

[17] Helming, K. C., Wang, X., and Roberts, C. W. M. (2014) Vulnerabilities of mutant SWI/SNF complexes in cancer, *Cancer Cell 26*, 309-317.

[18] Helming, K. C., Wang, X., Wilson, B. G., Vazquez, F., Haswell, J. R., Manchester, H. E., Kim, Y., Kryukov, G. V., Ghandi, M., Aguirre, A. J., Jagani, Z., Wang, Z., Garraway, L. A., Hahn, W. C., and Roberts, C. W. (2014) ARID1B is a specific vulnerability in ARID1A-mutant cancers, *Nat Med 20*, 251-254.