

Gene ID / UniProt ID / EC **Target Nominator** SGC Authors

Collaborating Authors Target PI Therapeutic Area(s) **Disease Relevance**

Date Approved by TEP Evaluation Group Document version Document version date Citation

Affiliations

SARS-Cov-2 NSP13



NSP13 / PODTD1 / -



Internal Nomination Joseph Newman¹, Yuliana Yosaatmadja¹, Alice Douangamath², Frank Von Delft^{1,2} Opher Gileadi¹ Opher Gileadi Infectious diseases NSP13 is a DNA/RNA helicase that is essential for SARS-Cov-2 replication, inhibitors to NSP13 could be developed as antiviral drugs. November 27th 2020 1.0 November 2020 Joseph Newman, Yuliana Yosaatmadja, Alice Douangamath, Frank Von Delft, & Opher Gileadi. (2020). SARS-CoV-2 NSP13; A Target Enabling Package [Data set]. Zenodo. http://doi.org/10.5281/zenodo.4449925 1. Centre for Medicines Discovery, Old Road Campus Research Building, University of Oxford, Roosevelt Dr, Headington,, Oxford, OX3 7DQ. 2. Diamond Light Source, Harwell Science and Innovation Campus, Fermi Avenue, Didcot, OX11 0DE.

USEFUL LINKS



SUMMARY OF PROJECT

To contribute towards the development of novel anti-viral therapeutics targeting the current and future emerging coronavirus threats, the Gileadi lab at the University of Oxford, together with the XChem team at Diamond Light Source, have teamed up to perform a crystallographic fragment screen against SARS-CoV-2 NSP13 helicase. NSP13 is believed to act in concert with the replication-transcription complex (NSP7/NSP8₂/NSP12), possibly being involved in either disrupting downstream RNA secondary structures or template switching, and plays an essential role in the life cycle of SARS-CoV-2.

This TEP includes expression clones and methods for producing the full length NSP13, and fluorescencebased activity assays suitable for compound screening. We provide a crystallisation system that produces reproducible crystals that diffract to high resolution, and have performed a crystallographic fragment screen revealing 63 fragment hits across 51 datasets. The fragment hits include several hits in pockets predicted to be of functional importance, including the nucleotide and nucleic acid binding sites, opening the way to development of novel antiviral agents.

SCIENTIFIC BACKGROUND

SARS-CoV-2 is the causative agent of the current global coronavirus (COVID-19) pandemic, a severe respiratory disease that emerged in the Chinese city of Wuhan in late 2019(1). Coronaviruses belong to the order Nidovirales which have a positive strand RNA genome that is amongst the largest know RNA genomes (approximately 30 KB in length). The main SARS-CoV genome encodes two open reading frames ORF1a and 1b, that when translated produce polyproteins that are processed by proteases into 16 non-structural proteins (NSP1-16)(2) that collectively form the machinery for viral replication and transcription. The NSP13 helicase is believed to act in concert with the replication-transcription complex (NSP7/NSP82/NSP12)(3) and is a critical component for viral replication(4). NSP13 is among the most conserved of the non-structural proteins, sharing only a single amino acid (V570I) difference with SARS-CoV-1. Thus, compounds targeting SARS-CoV-2 NSP13 would likely be effective against SARS-CoV-1 and potentially other future emerging coronaviruses, making it an ideal target for the development of new antiviral therapeutics.

NSP13 is a large 67 kDa protein that belongs to the helicase superfamily 1B. It utilises the energy of nucleotide triphosphate hydrolysis to catalyse the unwinding of double stranded DNA or RNA in a 5' to 3' direction. NSP13 contains 5 domains, a N-terminal zinc binding domain that coordinates 3 structural zinc ions, a helical "stalk" domain, a beta-barrel 1B domain and two "RecA like" helicase subdomains 1A and 2A that contain the residues responsible for nucleotide binding and hydrolysis (**Fig 1A**). Possible sites of inhibition of NSP13 include the nucleotide binding pocket, the DNA/RNA binding pocket and potentially allosteric pockets that may block domain movements that are required as part of the NSP13 catalytic cycle. Initial druggability analysis indicates both nucleotide and DNA/RNA binding pockets as being druggable and amongst the most well conserved pockets in the entire SARS-CoV-2 genome.

RESULTS – THE TEP

Proteins purified

We have expressed and purified full length NSP13 (1-601) in E. coli cells.

Structural data

Previously crystal structures of NSP13 have been solved for MERS-CoV and the highly related SARS-CoV to 3.0 Å and 2.8 Å respectively(4,5). We used these structures to define domain boundaries and identified a new crystal form for the full-length protein that is reproducible and diffracts routinely to around 2.0 Å.

6ZSL: Crystal structure of the SARS-CoV-2 helicase at 1.94 Angstrom resolution:

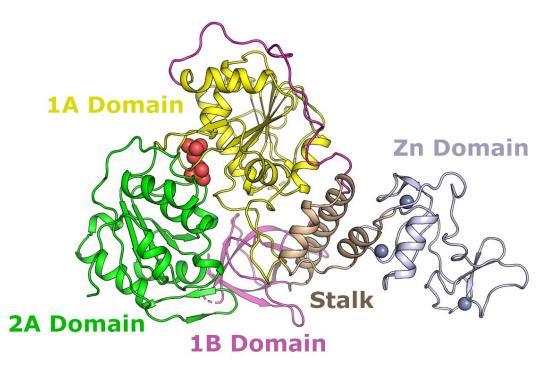


Figure 1. Overall structure of the SARS-CoV-2 NSP13 helicase with domains labelled and coloured individually (PDB: 6ZSL).

These crystals were used in X-ray fragment screening to identify fragment binding sites and initial chemical matter. Based on previous structures of related proteins(6), the nucleotide is believed to bind in a cleft between the 1A and 2A domains, whilst the RNA passes through a cavity formed between the IB domain and the 1A and 2A domains (**Fig 2**). Fragment hot spots were identified in the nucleotide and RNA binding sites as well as potential allosteric sites between domains (**Fig 2**).

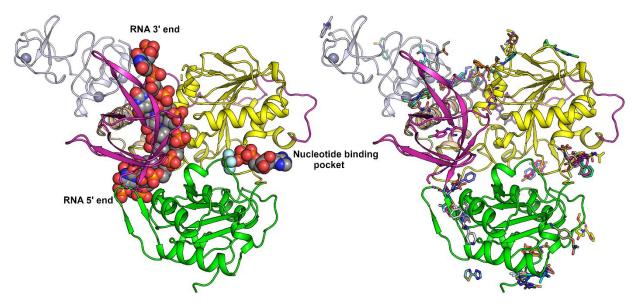


Figure 2. Structural model of NSP13 with the RNA and nucleotide bound in their expected positions based on the crystal structure of Upf-1 RNA complex **2XZL** (6). The right-hand panel shows the NSP13 fragment hits viewed from the same orientation.

Assays

DNA helicase assay: We have adapted a fluorescence dequenching DNA strand separation assay to measure the helicase activity of NSP13 (**Fig 3**). Whilst both DNA and RNA helicase activities have been observed for

NSP13, the activity on DNA appears to be more robust(7). This assay has been previously used in a high throughput screen to identify helicase inhibitors of Blooms syndrome helicase(8).

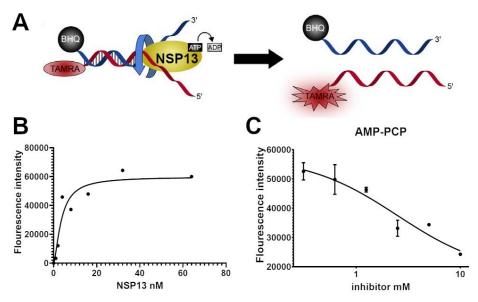


Figure 3. Fluorescence dequenching-based helicase unwinding assay. (A) Schematic diagram showing the principal of the helicase unwinding assay (adapted from(8)). (B) NSP13 enzyme titration measured after 10 minutes incubation. (C) Dose response of a low affinity ATP competitive inhibitor AMP-PCP.

Chemical Matter

The NSP13 crystal form contains two protomers and the fragment screen revealed a total of 63 fragment hits across 51 datasets. These structures have been published in the PDB and are featured as a curated collection in the RCSB PDB SARS-CoV-2 resource. We show in the table below structures and binding sites of selected fragment hits binding to functionally important sites.

PDBID	Ligand	Binding Location	Binding Pocket	Resolution (Å)
5RL7	۲ <u>3</u> 64321922	Nucleotide pocket A & RNA 3' B	Rucleotide pocket A	1.89
5RLV	Z2467208649	Nucleotide pocket A & Other A & RNA 5' proximal B	Rucleotide pocket A	2.21
5RLY	z2027049478	Nucleotide pocket A & RNA 5' proximal B	Nucleotide pocket A	2.43

For more information regarding any aspect of TEPs and the TEP programme, please contact teps@thesgc.org

	_			
5RLS	Z59181945	Nucleotide pocket A	Nucleotide pocket A	2.28
5RLN	сзб4328788	Nucleotide pocket A	Nucleotide pocket A	2.15
5RL9	HN C C C C C C C C C C C C C C C C C C C	Nucleotide pocket B	Nucleotide pocket B	1.79
5RLI	Z45617795	Nucleotide pocket B & Other A	Nucleotide pocket B	2.26
5RLJ	ر بالمراجع المراجع الم 21407673036	Nucleotide pocket B	Nucleotide pocket B	1.88
5RLO	Z1454310449	Nucleotide pocket B	Nucleotide pocket B	2.10
5RLR	Z822382694	Nucleotide pocket B	Nucleotide pocket B	2.32

5RLW	کر آپ 245705015	Nucleotide pocket B & Other A	Nucleotide pocket B	1.97
5RM2	р но Z1741964527	Nucleotide pocket B & RNA 5' Proximal B	Nucleotide pocket B	1.82
5RM7	روب نوب ش Z69118333	Nucleotide pocket B	Nucleotide pocket B	1.84
5RLL	Z425387594	RNA 5' Proximal B	RNA 5' Proximal B	2.08
5RLE	Z1429867185	RNA 5' Proximal B	RNA 5' Proximal B	2.27
5RLP	ларания Ин Z166605480	RNA 5' Proximal B	RNA 5' Proximal B	2.56
5RMK	HN - 0 	RNA 5' Proximal B	RNA 5' Proximal B	2.08

5RLH	рание и страниции и страници	RNA 5' B	RNA 5' B	2.38
5RMM	POB0066	RNA 5' B	RNA 5' B	2.20
5RLZ	страната и с Страната и страната и с	RNA 5' B	RNA 5' B	1.97
5RL6	پر ۲198195770	RNA 3' B	RNA 3' B	1.92
5RLU	Z744754722	RNA 3' B & Other A	RNA 3' B	2.35
5RLK	HN % N Z1509882419	RNA central B	RNA central B	1.96

Table 2: Structures and binding sites of selected fragment hits binding to functionally important sites.

IMPORTANT: Please note that the existence of small molecules within this TEP indicates only that chemical matter might bind to the protein in potentially functionally relevant locations. The small molecule ligands are intended to be used as the basis for future chemistry optimisation to increase potency and selectivity and yield a chemical probe or lead series. As such, the molecules within this TEP should not be used as tools for functional studies of the protein, unless otherwise stated, as they are not sufficiently potent or well-characterised to be used in cellular studies.

Future Plans

We are currently preparing a manuscript describing the crystal structure and X-ray fragment screening for NSP13. Following the publication of our data in the PDB we have been contacted by academic groups and

pharma partners and have agreed to work together, providing structural support for follow up compound soaking and early stage drug discovery.

CONCLUSION

NSP13 is an important target for the development of new antiviral compounds targeting SARS-CoV-2 and other potential emerging viral threats. We describe a robust crystallisation system for NSP13 that produces crystals that routinely diffract to high resolution and have performed an X-ray fragment screen to identify initial binders. Of particular interest are the large number of hits in the nucleotide binding site and the DNA/RNA interface which may be developed into ATP or DNA competitive inhibitors. Alternatively, several pockets were discovered at interfaces between domains that could lead to allosteric inhibitors that prevent domain movements that occur as part of the helicase catalytic cycle.

TEP IMPACT

This TEP was generated as part of a wider partnership between the University of Oxford and the XChem facility at Diamond Light Source to contribute to the global effort to combat COVID-19. In the short time following publication of the fragment screening results in the PDB we have been contacted by industrial collaborators to screen follow up compounds with the ultimate aim of developing orally available antiviral drugs within two years. The NSP13 work has also been included in an application to an internal University of Oxford Institutional Strategic Support Fund to fund cellular screening of small molecules in an antiviral assay.

FUNDING INFORMATION

The work performed at the SGC has been funded by a grant from the Wellcome [106169/ZZ14/Z].

ADDITIONAL INFORMATION

Structure Files

PDB ID	Structure Details	
6ZSL	1.94 Å structure of the SARS-CoV-2 helicase	

Materials and Methods

Molecular Biology

Clone Source: Synthetic Gene (Twist Biosciences) SGC Construct ID: CVNSP13A-c000 (https://www.addgene.org/159614/) Vector: pNIC-ZB (GenBank: GU452710.1) Tag: N-terminal 6HIS, Z-basic tag with TEV cleavage Host: BL21(DE3)-R3-pRARE2

Sequence with tag (underlined; *: TEV protease cleavage site):

MHHHHHHSSGVDNKFNKERRARREIRHLPNLNREQRRAFIRSLRDDPSQSANLLAEAKKLNDAQPKGTENLYFQ*SMAV GACVLCNSQTSLRCGACIRRPFLCCKCCYDHVISTSHKLVLSVNPYVCNAPGCDVTDVTQLYLGGMSYYCKSHKPPISFPLCAN GQVFGLYKNTCVGSDNVTDFNAIATCDWTNAGDYILANTCTERLKLFAAETLKATEETFKLSYGIATVREVLSDRELHLSWEV GKPRPPLNRNYVFTGYRVTKNSKVQIGEYTFEKGDYGDAVVYRGTTTYKLNVGDYFVLTSHTVMPLSAPTLVPQEHYVRITGL YPTLNISDEFSSNVANYQKVGMQKYSTLQGPPGTGKSHFAIGLALYYPSARIVYTACSHAAVDALCEKALKYLPIDKCSRIIPAR ARVECFDKFKVNSTLEQYVFCTVNALPETTADIVVFDEISMATNYDLSVVNARLRAKHYVYIGDPAQLPAPRTLLTKGTLEPEY FNSVCRLMKTIGPDMFLGTCRRCPAEIVDTVSALVYDNKLKAHKDKSAQCFKMFYKGVITHDVSSAINRPQIGVVREFLTRNP AWRKAVFISPYNSQNAVASKILGLPTQTVDSSQGSEYDYVIFTQTTETAHSCNVNRFNVAITRAKVGILCIMSDRDLYDKLQFT SLEIPRRNVATLQ

Sequence after tag cleavage:

SMAVGACVLCNSQTSLRCGACIRRPFLCCKCCYDHVISTSHKLVLSVNPYVCNAPGCDVTDVTQLYLGGMSYYCKSHKPPISF PLCANGQVFGLYKNTCVGSDNVTDFNAIATCDWTNAGDYILANTCTERLKLFAAETLKATEETFKLSYGIATVREVLSDRELHL SWEVGKPRPPLNRNYVFTGYRVTKNSKVQIGEYTFEKGDYGDAVVYRGTTTYKLNVGDYFVLTSHTVMPLSAPTLVPQEHYV RITGLYPTLNISDEFSSNVANYQKVGMQKYSTLQGPPGTGKSHFAIGLALYYPSARIVYTACSHAAVDALCEKALKYLPIDKCSR IIPARARVECFDKFKVNSTLEQYVFCTVNALPETTADIVVFDEISMATNYDLSVVNARLRAKHYVYIGDPAQLPAPRTLLTKGTL EPEYFNSVCRLMKTIGPDMFLGTCRRCPAEIVDTVSALVYDNKLKAHKDKSAQCFKMFYKGVITHDVSSAINRPQIGVVREFL TRNPAWRKAVFISPYNSQNAVASKILGLPTQTVDSSQGSEYDYVIFTQTTETAHSCNVNRFNVAITRAKVGILCIMSDRDLYD KLQFTSLEIPRRNVATLQ

DNA Sequence (codon-optimized for E. coli expression):

Protein Expression and Purification

Medium: Terrific Broth (TB) Merck with 4 ml of glycerol

Antibiotics: Kanamycin, 50 µg/ml

From the glycerol stock, bacteria were inoculated in 15 ml of 1 x TB in a 50 ml tube with Kanamycin 0.05 mg/ml and 0.034 mg/ml of chloramphenicol and grown overnight in a shaker at 37°C, 250rpm. The following day, 40 ml of the overnight culture were inoculated in 4L of TB. The bacteria grew in an incubator at 37°C, shaking 180 rpm. Once the OD reached 2-3, IPTG (300uM) was added to the media and left overnight at 18°C, shaking 180 rpm. The pellets were harvested the next following day.

Protein Purification

The pellet (45g cell mass) was re-suspended in 200 ml lysis buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 5% Gol, 10 mM Imidazole, 0.5 mM TCEP) with protease inhibitors (500 Merck set III). Cells were disrupted by sonication for 15 mins, 10sec on 5sec off, spin in JA25.5 24500 RPM for 30 mins. The supernatant was incubated for 40 mins with 5ml of Ni resin (IAMC sepharose) for batch binding. The tubes containing the lysate were centrifuged at 700 x g at 4°C for 5 minutes and the supernatant discarded.

Beads were loaded on a gravity flow column and washed with 40 ml lysis buffer, 25 ml wash buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 5% Gol, 45 mM imidazole, 0.5 mM TCEP). A further wash with 10 ml Hi-salt buffer (50 mM HEPES pH 7.5, 1 M NaCl, 5% Gol, 0.5 mM TCEP) and again with another 10 ml of wash buffer. Proteins were eluted with addition of 15ml of elution buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 5% Gol, 300 mM imidazole, 0.5 mM TCEP).

Elution fraction was immediately applied to a 5ml HItrap SP column using a syringe, collecting the flow through. The SP column was washed with 10ml elution buffer and proteins were eluted with 15ml Hi-salt buffer. The NSP13 protein was found to be present in flow-through, wash and elution fractions and these fractions were pooled and treated separately from this point onward.

For further purification both proteins were incubated overnight with TEV protease (1:40 mass ratio) and loaded onto gel filtration systems using a superdex 200 16/60 column equilibrated in 50 mM Hepes, 500 mM NaCl, 0.5 mM TCEP.

Both pools were found to crystallise with the majority of the crystals coming from the SP flow through and wash which had greater yield although slightly dirtier.

Protein was concentrated to 20 mg /ml and diluted in half with water for initial crystallization trials at 10 mg/ml using a combination of screens and nucleotide combinations. Total yield was around 6 mg. The protein was confirmed by ESI-TOF intact mass spectrometry.

Crystallization

Initial diffracting crystals were found in the Morpheus screen from Molecular dimensions for the APO form of the protein at 10 mg/ml. Initial crystals grew at 20 degrees from conditions containing 20 % Ethylene Glycol, 10 % PEG 8K, 0.05 M HEPES, 0.05 M MOPS, 0.03 M sodium nitrate, 0.03 M sodium phosphate, 0.03 M ammonium sulphate. For crystal optimization seeding was performed, around 5-10 crystals were crushed with a glass probe and transferred to 25 μ l of well solution. A seed bead was added, and the mixture was sonicated for around 30-60 seconds with pulsing. Final seeding was performed with a 1 in 400 dilution of seed stock. Final plates were set up with protein at 5 mg/ml (diluted 4 fold in water from 20 mg/ml stock) with a slightly reduced precipitant concentration (16 % ethylene glycol, 8 % PEG 8K, 0.05 M HEPES, 0.05 M MOPS, 0.03 M sodium nitrate, 0.03 M sodium phosphate, 0.03 M ammonium sulphate), using 300 nl drops (1:1 ratio) with 20 nl seeds (added last).

Structure Determination (6ZSL)

Data Collection: Data were collected to 1.94Å resolution at Diamond light source beamline I04-1 and processed using XDS.

Data Processing: The structure was solved by molecular replacement using the program PHASER and the structure of SARS-CoV-1 NSP13 (GJYT) as a search model. Refinement was performed using PHENIX REFINE to a final Rfactor = 20.9%, Rfree = 25.3%.

X-ray Fragment Screening

Fragment soaking: Fragments from the DSI-Poised library were added to the crystallisation drops by acoustic dispensing using an ECHO acoustic liquid handler from a 500 mM stock concentration dissolved in DMSO to a final concentration of 10%. Soaking times varied from 1 to 3 hours.

Data Collection: Data were collected at Diamond light source beamline IO4-1 and processed using the XChem Explorer pipeline.

Data Processing: Structures were solved by difference Fourier synthesis using the XChem Explorer pipeline. Fragment hits were identified using the PanDDA program. Refinement was performed using REFMAC or BUSTER.

References

- Chan, J. F., Yuan, S., Kok, K. H., To, K. K., Chu, H., Yang, J., Xing, F., Liu, J., Yip, C. C., Poon, R. W., Tsoi, H. W., Lo, S. K., Chan, K. H., Poon, V. K., Chan, W. M., Ip, J. D., Cai, J. P., Cheng, V. C., Chen, H., Hui, C. K., and Yuen, K. Y. (2020) A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. *Lancet* **395**, 514-523
- Chan, J. F., Kok, K. H., Zhu, Z., Chu, H., To, K. K., Yuan, S., and Yuen, K. Y. (2020) Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan. *Emerg Microbes Infect* 9, 221-236
- Chen, J., Malone, B., Llewellyn, E., Grasso, M., Shelton, P. M. M., Olinares, P. D. B., Maruthi, K., Eng, E. T., Vatandaslar, H., Chait, B. T., Kapoor, T. M., Darst, S. A., and Campbell, E. A. (2020) Structural Basis for Helicase-Polymerase Coupling in the SARS-CoV-2 Replication-Transcription Complex. *Cell* 182, 1560-1573 e1513

- Jia, Z., Yan, L., Ren, Z., Wu, L., Wang, J., Guo, J., Zheng, L., Ming, Z., Zhang, L., Lou, Z., and Rao, Z.
 (2019) Delicate structural coordination of the Severe Acute Respiratory Syndrome coronavirus Nsp13 upon ATP hydrolysis. *Nucleic Acids Res* 47, 6538-6550
- Hao, W., Wojdyla, J. A., Zhao, R., Han, R., Das, R., Zlatev, I., Manoharan, M., Wang, M., and Cui, S. (2017) Crystal structure of Middle East respiratory syndrome coronavirus helicase. *PLoS Pathog* 13, e1006474
- 6. Chakrabarti, S., Jayachandran, U., Bonneau, F., Fiorini, F., Basquin, C., Domcke, S., Le Hir, H., and Conti, E. (2011) Molecular mechanisms for the RNA-dependent ATPase activity of Upf1 and its regulation by Upf2. *Mol Cell* **41**, 693-703
- 7. Mickolajczyk, K. J., Shelton, P. M. M., Grasso, M., Cao, X., Warrington, S. R., Aher, A., Liu, S., and Kapoor, T. M. (2020) Force-dependent stimulation of RNA unwinding by SARS-CoV-2 nsp13 helicase. *bioRxiv*
- Nguyen, G. H., Dexheimer, T. S., Rosenthal, A. S., Chu, W. K., Singh, D. K., Mosedale, G., Bachrati, C. Z., Schultz, L., Sakurai, M., Savitsky, P., Abu, M., McHugh, P. J., Bohr, V. A., Harris, C. C., Jadhav, A., Gileadi, O., Maloney, D. J., Simeonov, A., and Hickson, I. D. (2013) A small molecule inhibitor of the BLM helicase modulates chromosome stability in human cells. *Chem Biol* 20, 55-62

We respectfully request that this document is cited using the DOI value as given above if the content is used in your work.