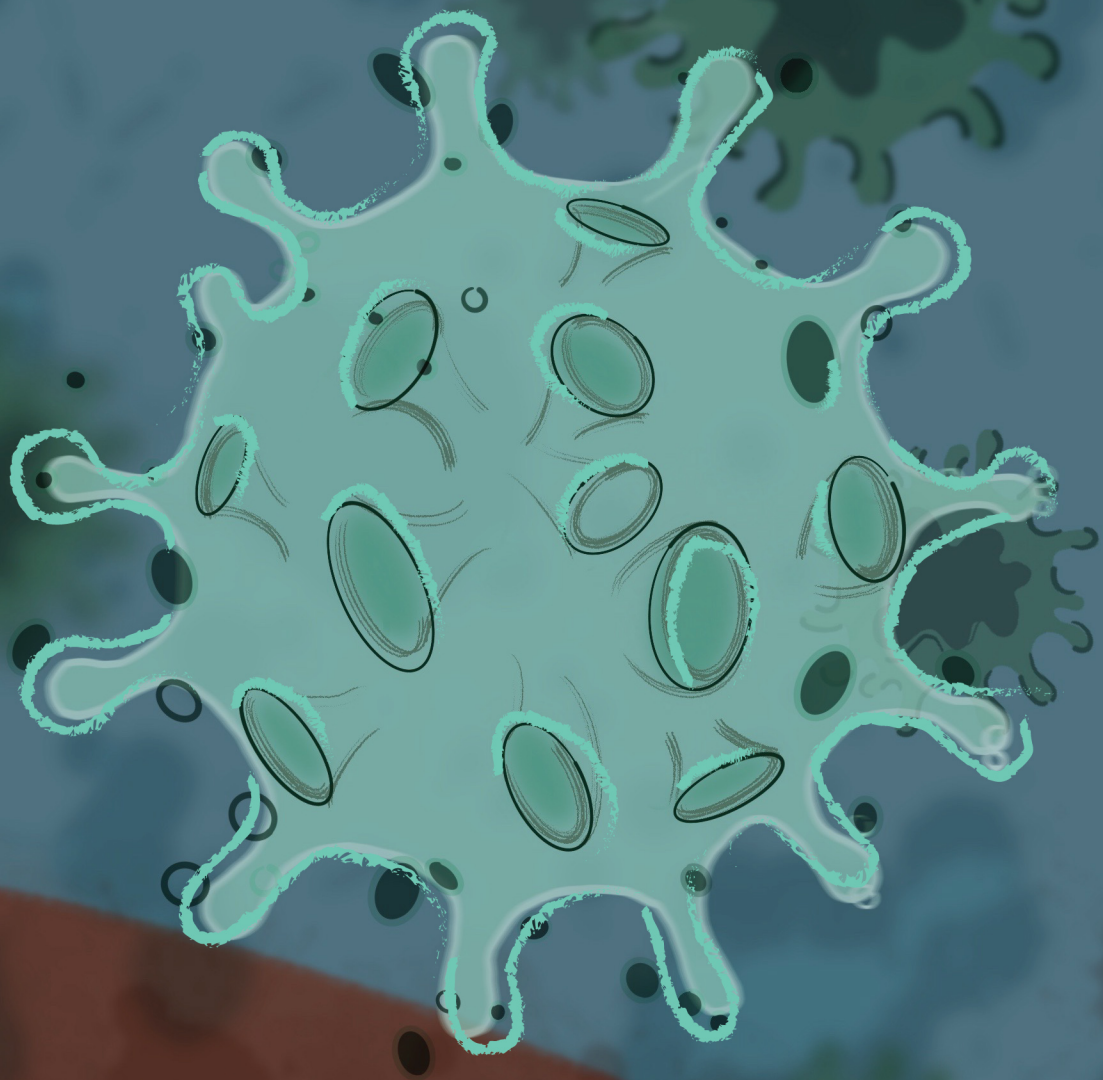
STEMedicine

Vol 1 • Issue 2 • Apr 2020

ISSN 2705-1188



Potential inhibitors targeting the SARS-CoV-2 coronavirus

STEMedicine

stemedicine.org

REVIEW ARTICLE

Functional studies of the coronavirus nonstructural proteins

Yanglin QIU and Kai XU*

Jiangsu Key Laboratory for Microbes and Functional Genomics, College of Life Sciences, Nanjing Normal University, Nanjing 210023, P. R. China.

*Correspondence: xukai@njnu.edu.cn
https://doi.org/10.37175/stemedicine.v1i2.39

ABSTRACT

Coronaviruses, including SARS-CoV, SARS-CoV-2, and MERS-CoV, have caused contagious and fatal respiratory diseases in humans worldwide. Notably, the coronavirus disease 19 (COVID-19) caused by SARS-CoV-2 spread rapidly in early 2020 and became a global pandemic. The nonstructural proteins of coronaviruses are critical components of the viral replication machinery. They function in viral RNA transcription and replication, as well as counteracting the host innate immunity. Studies of these proteins not only revealed their essential role during viral infection but also help the design of novel drugs targeting the viral replication and immune evasion machinery. In this review, we summarize the functional studies of each nonstructural proteins and compare the similarities and differences between nonstructural proteins from different coronaviruses.

Keywords: Coronavirus · SARS-CoV-2 · COVID-19 · Nonstructural proteins · Drug discovery

Introduction

The coronavirus (CoV) outbreaks among human populations have caused three major epidemics worldwide, since the beginning of the 21st century. These are the epidemics of the severe acute respiratory syndrome (SARS) in 2003 (1, 2), the Middle East respiratory syndrome (MERS) in 2012 (3, 4), and the coronavirus disease 19 (COVID-19) in 2020 (5-9).

Coronaviruses belong to the family *Coronaviridae* in the order *Nidovirales* (10). The *Coronaviridae* family has four genera, *Alphacoronavirus*, *Betacoronavirus*, *Gammacoronavirus*, and *Deltacoronavirus*. Currently, seven coronaviruses are known to infect humans. These include two alphacoronaviruses, human coronavirus 229E (HCoV-229E) (11) and HCoV-NL63 (12), as well as five betacoronaviruses, HCoV-OC43 (13), HCoV-HKU1 (14), SARS-CoV (15), MERS-CoV (4), and the novel coronavirus that causes COVID-19 (2019-nCoV) (5, 7, 9). The 2019-nCoV was later renamed to SARS-CoV-2 by the International Committee on Taxonomy of Viruses (16), due to high sequence homology and same receptor requirements shared with SARS-CoV (9).

The genome of coronaviruses is a nonsegmented, positive-sense, single-strand RNA of about 28 - 32 nt in length (17), which encodes for six major open reading frames (ORFs) and a various number of accessory genes (18). The first two major ORFs (ORF1a, ORF1ab) are the replicase genes, and the other four encode viral structural proteins that comprise the essential protein components of the coronavirus virions, including the spike surface glycoprotein (S), envelope protein (E), matrix protein (M), and nucleocapsid protein (N) (14, 19, 20).

In the case of *betacoronavirus*, ORF1a translates into a polyprotein. It is subsequently cleaved into 11 nonstructural proteins (nsp1-11) by itself or by nsp3 and nsp5, which contains the papain-like (21) and Chymotrypsin-like protease activities (22), respectively. ORF1ab is a polyprotein produced by translational readthrough of -1 ribosomal frameshift, which occurred at the coding region of nsp11. As a result, ORF1ab polyprotein carries the nsp1-10 at its N-terminus and five more nonstructural proteins (nsp12-16) on the C-terminus. These nonstructural proteins synergistically form the cytoplasmic membranous organelle-like viral replication/ transcription complexes (TRCs) (23). Besides the essential roles in viral RNA replication and transcription (18, 24-33), nonstructural proteins also determine the viral pathogenesis (18, 29, 34) as well as counteracting the cellular immune responses (21, 34-41). The coronavirus TRCs are among the most complex viral replication compartments (VRCs) known for positive-strand RNA viruses (23, 42, 43).

Besides its complex nature, structures of most of the nonstructural proteins have been determined (26, 44-58). These structures not only provide the functional insights of nonstructural proteins but also help the design of novel drugs targeting the viral replication and immune evasion machinery (47, 59-62). These studies facilitate our

Received: Mar 13, 2020; Accepted: Mar 21, 2020.

understanding of how viruses establish their infection and provide the foundations for the development of efficient antiviral strategies. Here we summarize current knowledge on each nonstructural proteins of coronaviruses.

Nsp1: the immune repressor

Nsp1 is the very 5' proximal nonstructural protein of β-CoVs. It is released by PLpro at conserved proteolytic sites from polyprotein encoded by ORF1ab (63-65) (Reviewed in **Table 1**). Due to the severe outbreak of SARS in 2003, SARS-CoV nsp1 is among the most extensively studied. SARS-CoV nsp1 is 20-kDa in size and distributed in the cytoplasm when transiently expressed in 293 cells (66). It was first observed that SARS-CoV nsp1 transient expression strongly inhibited IFN-β mRNA accumulation during Sendai virus infection, and promoted degradation of overexpressed exogenous mRNA and host endogenous mRNA, leading to an overall decrease in protein synthesis (66). Similar results were then obtained from studies on other β-CoVs including mouse hepatitis virus (MHV), bat coronavirus strains (Rm1, 133, and HKU9-1), and on a β-CoV human coronavirus 229E (HCoV-229E) (67, 68). These findings suggest that host translation arresting is a common feature during coronavirus infections.

The nsp1-mediated translation inhibition can be reproduced in the cell-free translation system (69). SARS-CoV nsp1 was demonstrated to bind and inactivate the 40S ribosomal subunits, resulting in translational inhibition. Meanwhile, the nsp1 protein, in its 40S ribosomal subunit binding form, could recruit a cellular endonuclease to mediate mRNA cleavage in the 5' untranslated region (5'-UTR) (70). Subsequently, 5'-truncated host mRNAs were degraded by host 5' - 3' exonuclease Xrn1 (71). Interestingly, protein translation under the control of internal ribosome entry site (IRES) from hepatitis C or cricket paralysis viruses, but not encephalomyocarditis virus, could escape from the SARS-CoV nsp1-mediated RNA cleavage, possibly due to different requirement for translational initiation factors in forming 48S initiation complex with the 40S subunit (69).

A SARS-CoV nsp1 mutant with two positively charged amino acid substitutions (R124A/K125A) loses the target mRNA/viral RNA binding and mRNA cleavage function but remains the ability to inhibit translation (72). MERS-CoV nsp1 contains the same RK motif. Nevertheless, this RK motif is not involved in binding to mRNAs but is required for the RNA cleavage. Instead, the R13 on the first alpha-helix of MERS-CoV nsp1 that is missing from SARS-CoV nsp1 is essential for mRNA binding (73).

Another SARS-CoV nsp1 mutant K164A/H165A was unable to bind to the 40S subunit and lost the ability to interfere with host gene expression (69). When the same mutations (K164A/ H165A) were introduced to SARS-CoV infectious clone, the recovered virus was replication-incompetent and unable to suppress innate immune responses or degrade host mRNA (66). A similar study was performed with MHV. MHV-nsp1-Δ99, which lacked 99 nucleotides in the nsp1 coding region essential for host translation arresting, could not reproduce well in wild-type mice (67). Moreover, the MHV-nsp1-Δ99 mutant restored its replication to the wt virus level when infecting mice defected in the type I interferons (IFN-I) recognition, highlighting the role of nsp1 in counteracting the IFN-I (72). Notably, the SARS-CoV-2 nsp1 shares a high protein sequence identity of 84.44% with SARS-CoV nsp1, including R124/K125 for mRNA binding (red-colored) and cleavage (asterisked, **Figure 1**), K164/H165 for translation shutoff (green colored, **Figure 1**), suggesting a likely same role of SARS-CoV-2 nsp1 in counteracting host immune responses.

SARS-CoV genomic RNA and subgenomic RNAs are somehow resistant to the nsp1 induced RNA cleavage (70). The resistant element was mapped to the 5'-end leader sequence of SARS-CoV RNAs, which contains two important nucleotides AU at the very 5' terminal, followed by AUUA. Coincidently, the SARS-CoV-2 genome also starts with the same two nucleotides AU, followed by UAAA. Whether SARS-CoV-2 avoids nsp1 mediated RNA cleavage by the same RNA coding remains to be answered.

Nsp2: fine-tuner of replication

Nsp2 is the most variable nonstructural protein among

Table 1 Proteolysis sites of the ORF1a/ORF1ab polyproteins. The polyproteins were cleaved by the nsp3 papain-like protease and the nsp5 3C-like protease, resulting in the release of sixteen nonstructural proteins. The position of the scissors represents the cleavage site.

papain-like protease clevage sites

	nsp1	nsp2	nsp3	nsp4	nsp5
SARS-CoV-2	···ELNGG	AYTR···TLKGG	APTK···ALKGG	KIVN···AVLQ	SGFR···VTFQ
SARS-CoV	···ELNGG	AVTR···RLKGG	APIK···SLKGG	KIVS···AVLQ	SGFR···VTFQ
MERS-CoV	···KLIGG	DVTP···RLKGG	APVK···KIVGG	APTW···GVLQ	SGLV···VVMQ

3C-like protease clevage sites

	nsp5	nsp6	nsp7	nsp8	nsp9	nsp10	nsp11	nsp12	nsp13	nsp14	nsp15	nsp16
SARS-CoV-2	SGFR···VTFQ	SAVK···ATVQ	SKMS···ATLQ	AIAS···VKLQ	NNEL···VRLQ	AGNA···PMLQ	SADA···	···TVLQ	AVGA···ATLQ	AENV···TRLQ	SLEN···PKLQ	SSQA···
SARS-CoV	SGFR···VTFQ	GKFK···ATVQ	SKMS···ATLQ	AIAS···VKLQ	NNEL···VRLQ	AGNA···PLMQ	SADA···	···TVLQ	AVGA···ATLQ	AENV···TRLQ	SLEN···PKLQ	ASQA···
MERS-CoV	SGLV···VVMQ	SGVR···AAMQ	SKLT···SVLQ	ATLS···VKLQ	NNEI···VRLQ	AGSN···AALP	QSKD···	···TVLQ	AVGS···YKLQ	SQIV···TKVQ	GLEN···PRLQ	ASAD···

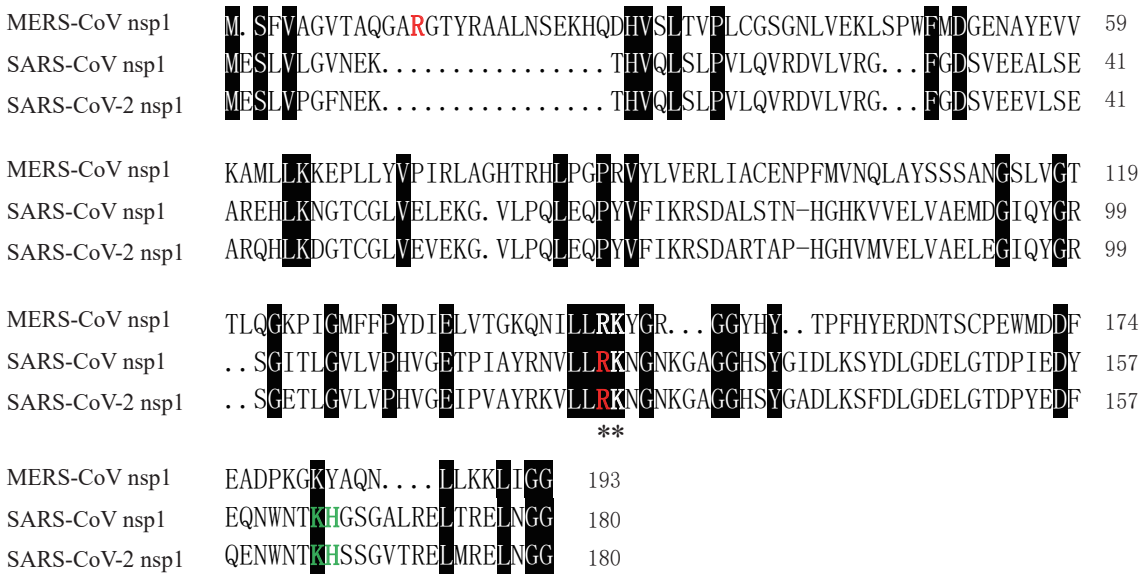


Figure 1. Alignment of nsp1 protein sequences of MERS-CoV, SARS-CoV, and SARS-CoV-2. The sequences of nsp1 from MERS-CoV (accession no. JX869059), SARS-CoV (accession no. AY278741), and SARS-CoV-2 (accession no. NC_045512) were aligned, and identical amino acids are highlighted. Amino acids essential for mRNA binding are colored in red, while those responsible for host translation shutoff are colored in green. Asterisked are the sites important for RNA cleavage.

SARS CoV-related viruses (74). A comparative analysis of the protein sequences of SARS-CoV-2 with SARS-CoV showed 61 amino acid substitutions in nsp2 between these two viruses (20). Due to the sequence variability of nsp2 among coronaviruses, it was speculated that nsp2 protein coevolved with the hosts to acquire host-specific functions, and modulating infection (75). A study utilizing nsp2-deleted MHV or SARS-CoV recombinant clones showed that nsp2 is required for optimal viral replication (30). When the nsp2 was deleted from the viruses, the viral titer, and viral RNA synthesis rate moderately reduced compared to the wild type virus. Nsp2 is shown to localize to the viral TRCs (43). However, MHV mutant with nsp2 deletion did not affect the morphology and subcellular localization of the TRCs (30). Importantly, Nsp2 expressed from other genomic loci still could not rescue the replication deficiency, pointing out that the function of nsp2 in viral growth depends on its correct genomic loci between nsp1 and nsp3 (76). Nsp2 was known to from nsp2-nsp3 proteolytic intermediate (30). Nsp2 may regulate protease cleavage in the form of nsp2-nsp3, thus fine-tuning the viral replication.

Nsp3: the scaffold protein and protease

Nsp3 is the largest among all the nonstructural proteins of β-CoV. It is cleaved off from ORF1a/ORF1ab by the papain-like protease domain or $PL2^{pro}$ domain that is within nsp3 itself. The function of nsp3 is to mediate the genome replication/transcription (77-80) and pathogenesis (81). Due to its large size and complex domain organization, nsp3 interacts with other nonstructural proteins (77, 78), structural proteins (79, 80), and host proteins (62) as a scaffold during viral infection (33). The SARS-CoV or SARS-CoV-2 nsp3 contains 16 domains (**Figure 2**), namely Ubiquitin-like domain 1 (Ubl1), Hypervariable region (HVR) or acidic domain, Macrodomain I/II/III (MacI/II/III), Domain Preceding Ubl2 and $PL2^{pro}$ (DPUP), Ubiquitin-like domain 2 (Ubl2), Papain-like protease two domain ($PL2^{pro}$), Nucleic acid-binding domain (NAB), betacoronavirus-specific marker domain (βSM), Transmembrane domain 1 (TM1), Nsp3 ectodomain (3Ecto), Transmembrane domain 2 (TM2), Amphipathic helix region (AH1), domains specific to *Nidovirales* and *Coronaviridae* (Y1 & CoV-Y) (82). These domains function synergistically to regulate viral infection.

Ubl1 is the first domain that locates on the N-terminus of β-CoV nsp3. The Nuclear magnetic resonance (NMR) structure of Ubl1 showed that the Ubl1 domain is structurally similar to ubiquitin-like proteins, albeit two additional helices (3_{10} helix and α helix) make the core structural more oval other than globular, comparing to human Ubiquitin or ISG15 (58, 82). Ubiquitination and ISGylation are associated with host regulation of innate antiviral responses (83-86), but the role of mimicry of ubiquitin by ubl1 (as well as ubl2) is currently unknown. It is reasonable to speculate that the ubiquitin-like domains of nsp3 could bridge the protease function of nsp3 to ubiquitination machinery in the cell and interfere with host antiviral immunity.

Ubl1 domain of nsp3 was shown to predominantly bind to single-stranded trinucleotide RNA sequence AUA, as mass-spectrometry analysis of the purified recombinant SARS-CoV Ubl1 from *E. Coli.* revealed co-purified unique RNA fragments (58). It is noted that both SARS-CoV and SARS-CoV-2 have AU-rich 5'-UTR or even 5' terminus in their genomic or subgenomic RNAs (20, 70). Whether this coincidence has a functional role remains to be tested. In addition to the studies on SARS-CoV, the Ubl1 domain of MHV nsp3 was found to bind to the viral.

N protein (80). This interaction is essential for N protein-mediated enhancement of viral infectivity (79).

The HVR domain locates at the C-terminus of the Ubl1 domain. HVR, also known as acidic domain, is rich in negatively charged amino acids aspartic acid (Asp/D) and glutamic acid (Glu/E). As its name indicates, it is the most variable region found in nsp3. Amino acids sequence identity between SARS-CoV HVR and SARS-CoV-2 HVR is 47.14%, much lower than the 76.6% overall nsp3 amino acid sequence identity. The HVR region is

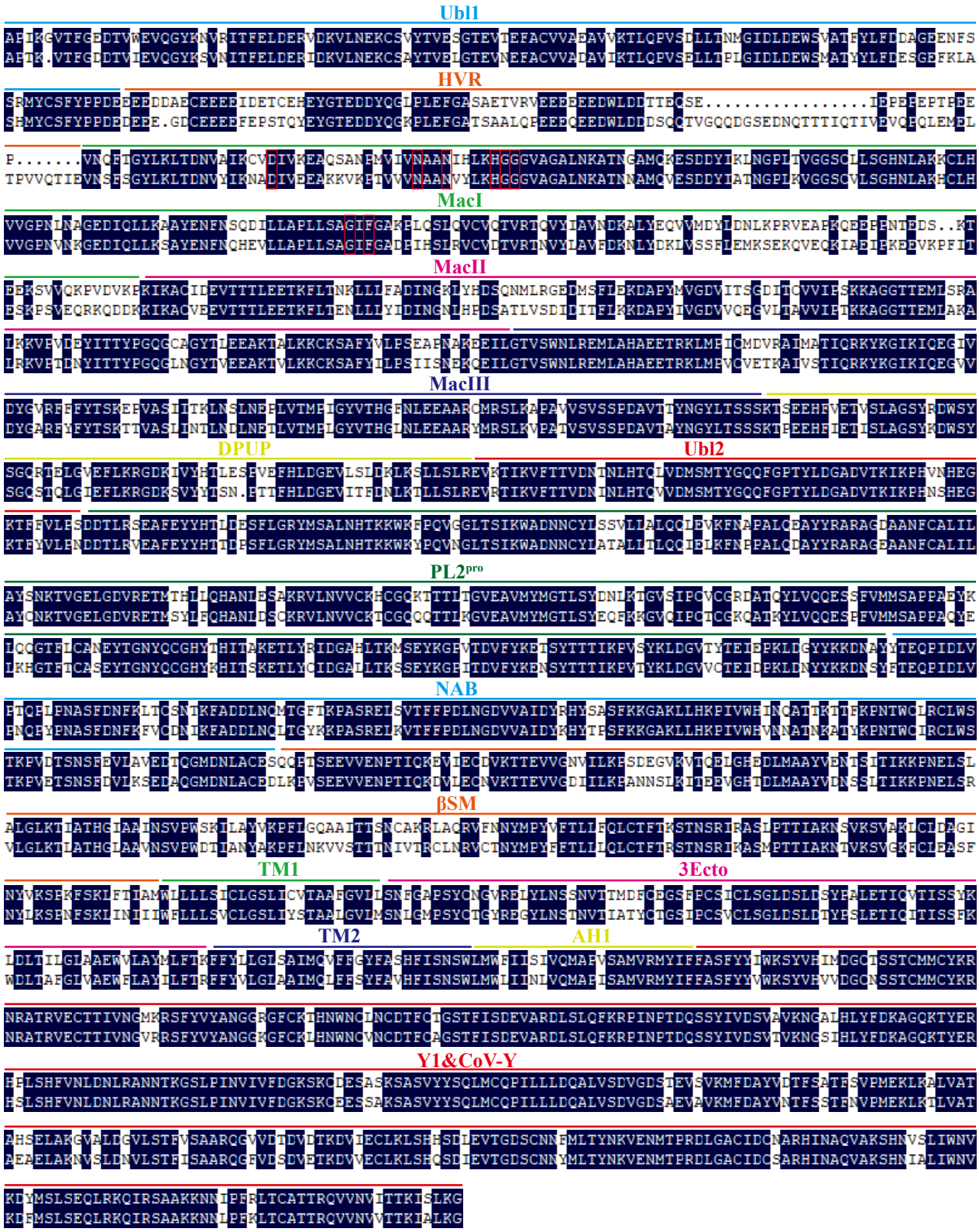


Figure 2. Sequence alignment of SARS-CoV nsp3 vs. SARS-CoV-2 nsp3. The domain organization of nsp3 is lined in different colors.

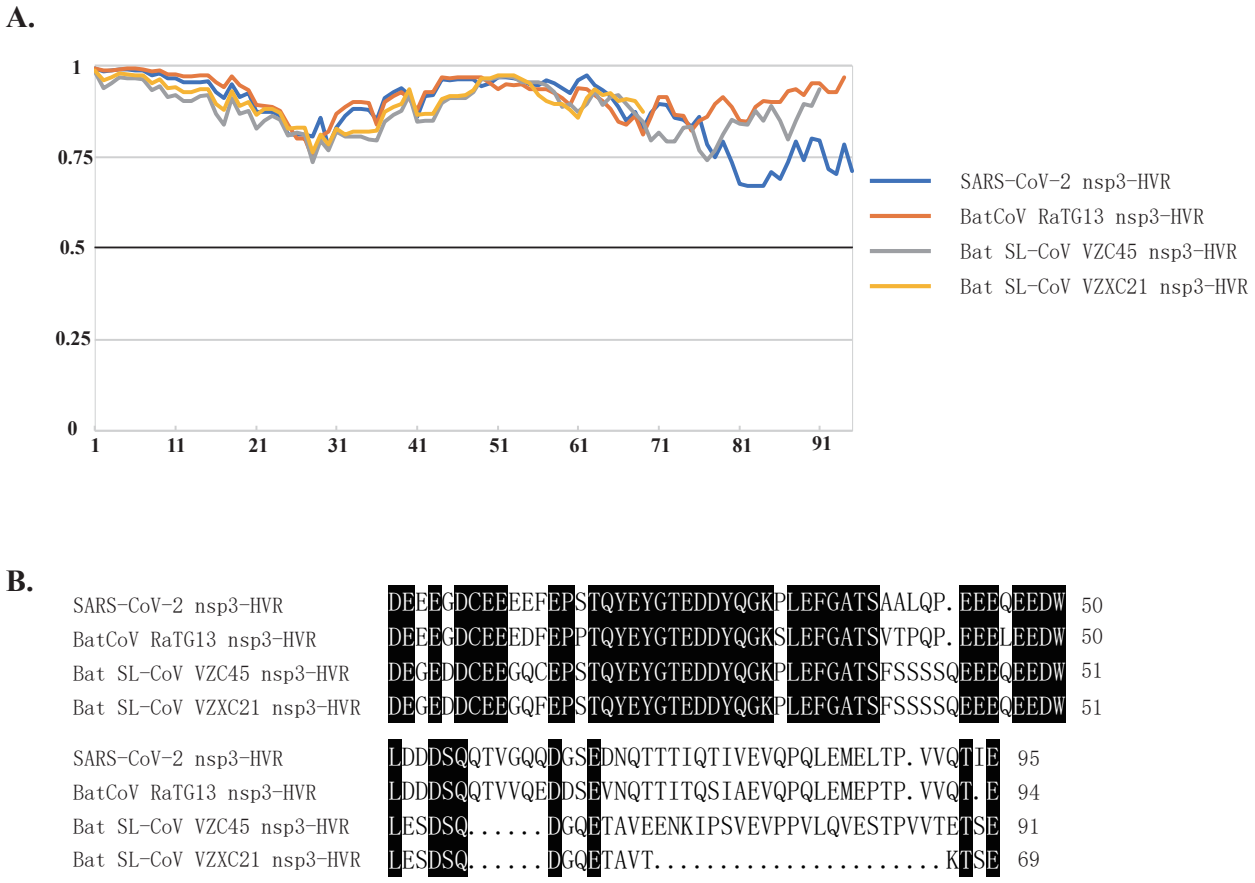


Figure 3. The HVR region of SARS-CoV-2 and its hypothetical ancestors is intrinsically disordered. A. The degree of disorder is shown graphically based on the analysis of IUPred2A (226). A score of more than 0.5 is considered disordered. B. The alignment of HVR region shows a high degree of negatively charged amino acid percentage in conserved amino acids.

intrinsically disordered in SARS-CoV (58) and MHV (57). The same feature is also observed in SARS-CoV-2 as well as in three highly similar bat coronavirus isolates BatCoV RaTG13 (accession no. MN996532) (9), Bat SL-CoV VZC45 (accession no. MG772933), and Bat SL-CoV VZXC21 (accession no. MG772934) (63) which show high nucleotide identity and protein identity to SARS-CoV-2 (**Figure 3A**). There are 45 consensus amino acids in the HVR among SARS-CoV-2 and these three bat viruses (**Figure 3B**), of which 48.9% are Asp/Glu. In the nonconsensus regions, the Asp/Glu percentage is 20%, 20.4%, 15.2%, and 8.3% for SARS-CoV-2, BatCoV RaTG13, Bat SL-CoV VZC45, and Bat SL-CoV VZXC21, respectively, and is much lower than that of the consensus sequence. These differences of HVR Asp/Glu percentage between consensus and nonconsensus regions indicate a possible function of negatively charged amino acids in viral replication that was selected during viral evolution. But currently, the exact role of HVR in the viral life cycle is unknown, studies on MHV suggest HVR is dispensable for viral infection *in vitro* (79).

Following the HVR is the Macrodomain I (MacI, previously known as X domain). Macrodomains are evolutionarily conserved domains that are ubiquitously existing in prokaryotes and eukaryotes. Three decades ago, bioinformatic analyses identified that members in *Coronaviridae*, *Togaviridae*, and *Hepeviridae* families encode this conserved domain of an unknown function, to which the name X domain was given (29, 59, 87-89). Protein crystallography studies on macrodomains of SARS-CoV (90, 91), MERS-CoV (54), and other coronaviruses (55, 56, 90) showed a three-layered alpha/beta/alpha core fold similar to the C-terminal nonhistone region of MacroH2A, a variant of human histone H2A (92). Macrodomains of SARS-CoV and some other coronaviruses contain in vitro ADP-ribose-1″-phosphate phosphatase (ADRP) activity (90, 91), de-mono-ADP-ribosylation (deMARylation) activity (36), and de-poly-ADP-ribosylation (dePARylation) activity (93). Studies by using a series of mutations on SARS-CoV and MHV showed that ADRP, deMARylation, and dePARylation activities of MacI are essential to viral virulence in vivo by suppressing the innate immune responses (36, 81, 91). These sites are conserved between SARS-CoV and SARS-CoV-2 (29) (**Figure 2**, boxed in red).

The SARS-CoV MacII+MacIII+DPUP forms a previously recognized SARS-unique domain (SUD), although more reports on betacoronavirus genome sequences suggest that this domain is not unique to SARS-CoV (94). MacII is the second marcodomain locating at the C-terminal side of MacI, and is dispensable for SARS-CoV replicon replication, while the third macrodomain

(MacIII) is required for SARS-CoV replication (32). MacIII binds to the G-quadruplexes formed by the Quadruplex forming G-Rich Sequences (QGRS) located in the nsp2 and nsp12 coding region (95). The MacII-III region also preferentially binds to oligo(G)-strings, which are present in the 3'-UTR of human mRNAs encoding defense-related genes (96). These RNA binding features are possibly essential for viral replication. DPUP is the domain that follows the MacIII. SARS-CoV lacking this domain displays reduced viral RNA even though the virus is still viable (32). Although the DPUP of SARS-CoV and MHV resembles a frataxin-like structure (95), which may involve controlling cellular oxidative stress (97, 98), the exact role of DPUP in viral infection is currently unknown.

Ubl2+PL2^{pro} locates immediately downstream of DPUP. Ubl2+PL2^{pro} of SARS-CoV and MERS-CoV were shown to modulate host innate immune responses (35, 99), albeit such immune suppression activity was separately reported to be only restricted in PL2^{pro} (41). Papain-like protease 2 (PL2^{pro}) is the protease domain within nsp3, which recognizes conserved motifs (**Table 1**) in the conjunction part of nsp1/nsp2, nsp2/nsp3, and nsp3/nsp4. In SARS-CoV and SARS-CoV-2, these motifs share XLXGG↓pattern, in which ↓ is the cleavage site. MERS-CoV contains the same pattern except for the cleavage site between nsp3 and nsp4 is KIVGG↓. These protease processing sites resemble the very C-terminal ends of ubiquitin and ubiquitin-like protein ISG15 (53), which is RLRGG and can be removed by deubiquitinating enzymes and deISGylating enzymes. Interestingly, the PL2^{pro} also possesses DUB and deISGylating activities besides its proteolytic activity (21, 41, 100-102). Ubiquitination and ISGylation play a central role in immune responses against viruses (85, 86, 103). Membrane-anchored SARS-CoV PL2^{pro} domain (PL2^{pro}-TM) could physically interact with STING (stimulator of IFN genes) - TRAF3 (The tumor necrosis factor receptor-associated factor 3) -TBK1 (TANK-binding kinase 1) complex and deubiquitinate intracellular dsRNA sensor RIG-I, STING, TRAF3, TBK1, and the transcription factor interferon regulatory factor 3 (IRF3) which cascadingly activate the IFN-β production pathway (104). SARS-CoV PL2^{pro} also removes Lys63-linked ubiquitin chains of TRAF3 and TRAF6, leading to suppression of the TLR7 signaling pathway (105). Human coronavirus NL63 (HCoV-NL63) deubiquitinates and stabilizes the E3 ubiquitin ligase MDM2 to increase the degradation of tumor suppressor p53 (40). Besides its DUB and deISGylating activities, coronavirus papain-like protease also depends on its interaction with host proteins to counteract host antiviral responses (106).

The NAB domain, which only exists in betacoronaviruses (94), forms flexibly extended linkers between the PL2^{pro} domain and the following domains of nsp3 (52). NAB domain could bind to RNA, especially repeats of GGGs (52), similar to the RNA recognition pattern of MacIII (95). Betacoronavirus-specific marker (βSM) domain follows NAB within nsp3. SARS-CoV βSM is intrinsically disordered (82), and its role in the viral life cycle is currently unknown.

Downstream of βSM domain is the transmembrane region that contains two transmembrane domains (TM1/2) and one luminal loop domain (3Ecto) (94). Subcellular localization analysis of SARS-CoV nsp3 truncated mutants revealed that the TM1/2 and the luminal 3Ecto domain are essential for the recruitment of nsp4 to discrete ER loci (78). The luminal 3Ecto domain of nsp3, possibly forming a disulfide bond, was proposed to interact with the luminal domains of nsp4 to "zipper" the ER membrane and induce discrete membrane formation (78). This membrane modification was recognized as the first step in forming the ER-origin viral replication organelles (28). Nsp3-nsp4 interaction of MERS-CoV also leads to the zippering of ER membranes and subsequent formation of Double-Membrane Vesicles (DVMs) (107).

AH1+Y1 & CoV-Y domain is the C-terminal portion of nsp3 that is facing towards the cytosol. AH1 encodes a predicted transmembrane domain that was shown to be a cytosolic region in SARS-CoV and MHV (42). Currently, the functions of these domains are less well understood as the N-terminal nsp3.

Nsp4: the DVM builder

Coronavirus nsp4 is an integral membrane protein with four transmembrane domains (42). In partnership with nsp3, it plays an essential role in the formation of the membranous structure of TRCs (27, 107, 108). SARS-CoV or MERS-CoV nsp4 /nsp3 localized to the reticular ER membrane when expressed separately (78, 107). However, when nsp4 and nsp3 were coexpressed, the formation of distinct perinuclear loci representing stacked double ER membranes was observed (78, 107). Such membrane rearrangements represent the critical step in TRCs formation. The N-terminus nsp4, including the first transmembrane domain and the first luminal loop region between TM1 and TM2, is required for this membrane rearrangement (78). The C-terminal TM4 and cytosolic part of nsp4 are dispensable for either formation of SARS-CoV induced aggregated ER loci (78), or efficient viral growth of MHV (27).

This first luminal region of SARS-CoV nsp4 could interact with the nsp3 luminal 3Ecto domain to bring two ER membranes in close proximity (109). This region was predicted to be glycosylated for various of coronaviruses (108). When glycosylated sites of this region were mutated in MVH, viral growth was reduced along with deficient DVM formation (108). Two-amino acid changes (H120N/F121L) near the SARS-CoV nsp4 glycosylation site (N131) abolished the nsp4-nsp3 interaction and also led to reduced genome replication and viral production (109). Comparing with SARS-CoV, the emerging SARS-CoV-2 contains same sites including both nsp3-interacting H120/F121 and the N131 glycosylation site (**Figure 4**).

Nsp5: the main protease

Coronavirus nsp5 encodes an enterovirus 3C-like protease (3CL^{pro}) conserved across the Nidovirales (19).

Figure 4. SARS-CoV-2 and SARS-CoV nsp4 share identical amino acids responsible for DVM formation.

It is a cysteine protease with a chymotrypsin-like fold and is often referred to as the main protease (Mpro). Similar to the $PL2^{pro}$, $3CL^{pro}$ is essential to the nonstructural protein processing by cleavage at 11 sites downstream of the nsp4 coding region (**Table 1**). SARS-CoV nsp5 consists of an N-terminal domain with proteolytic activity, as well as a C-terminal domain that contains five alpha-helices (22, 51, 110, 111). SARS-CoV $3CL^{pro}$ has at least three formats, an inactive monomer (22, 112), an active homodimer (22, 51, 111, 112), and a highly active homooctamer (110). Besides its proteolytic activity, porcine deltacoronavirus (PDCoV) nsp5 cleaves Signal transducer and activator of transcription 2 (STAT2) at two locations with glutamine (Q) residue at the P1 position, leading to the inhibition on the transcription of IFN-stimulated genes (39). Deltacoronavirus nsp5 also targets the NF-κB essential modulator (NEMO) for degradation and also suppresses type I IFN production (37, 38). Thus, coronavirus nsp5 assists viral infection by proteolytically releasing nsp4-16, and suppresses innate immune responses by digesting essential enzymes in the immune signal transduction pathway.

Nsp6: forming DVM and activating autophagy

Coronavirus nsp6 is a transmembrane protein with six transmembrane domains (77). When expressed alone, it localized to ER and induced the generation of DFCP1 (Double FYVE domain-containing protein 1) - positive early autophagosomes, or omegasomes. Such a structure can mature into autophagosomes that are capable of delivering LC3 for lysosomal degradation (113). Autophagy is not required for either coronavirus replication or antiviral responses *in vitro*. The knockout of ATG5 or ATG7, essential genes in the autophagy pathway, does not affect betacoronavirus MHV replication (114, 115). ATG5 silencing in Vero cells or treatment with wortmannin, the class3 PI3K inhibitor, also does not affect replication of the Infectious bronchitis virus (IBV), a gammacoronavirus, (113). Although induction of autophagy is not required for coronavirus genome replication, the nsp6 plays a vital role in the viral life cycle. Coronavirus encodes three nonstructural proteins with transmembrane domains, nsp3, nsp4, and nsp6. While nsp3+nsp4 only produces aggregated zippered ER structures or maze-like body, nsp6 expression in addition to nsp3+nsp4 leads to the DMVs formation (77), resembling the authentic membranous structures of TRCs (77). Two HCoV-229E mutants, both contain single amino acid mutation on nsp6, confer antiviral drug K22 resistant, and result in partial recovery of drug-related DVMs loss. These mutations affected progeny infectivities, suggesting that nsp6 is critical for the viral life cycle (61). Although nsp6 expression induces autophagosomes-like DVMs, coronaviruses do not require autophagy for viral replication. Nsp6 or coronavirus likely recruits some host proteins shared with the autophagy pathway for viral DVMs production. However, such speculation needs further investigation.

Nsp7+Nsp8: the RdRp cofactor

Coronavirus nsp7 and nsp8 are indispensable and essential for viral survival (116). The crystal structure of SARS-CoV nsp7 with nsp8 is a hollow cylinder-like supercomplex, formed by two asymmetric units. Each unit includes four nsp7 and four nsp8 (26). In the nsp7+nsp8 supercomplex, a channel structure is apparent (26). The channel is mainly formed by the bridging of the four long helices of N-terminus nsp8, of which the structure resembles the "shaft" of a "golf-club" (26). Mutations on the positive-charged amino acids on this "shaft" region significantly reduce dsRNA binding ability of the supercomplex, while the mutations of positive-charged amino acids on nsp7 near the channel structure do not (26).

Unlike nsp12, nsp8 is a non-canonical RNA-dependent RNA polymerase (RdRp) that does not encode the conserved RdRp motif (117). SARS-CoV nsp8 could initiate short oligonucleotide (< 6 nt) synthesis at an internal template cytidine with a distance of at least two nucleotides from the 3'-end (117). A later study misinterpreted this internal initiating primer synthesis ability from nsp8 as the *de novo* initiation (118). In this study, the authors also reported that nsp8 has primer extension activities (118). The association of nsp8 with nsp7 was shown to enhance thermal stability (117) and primer extension activity of nsp8 (118). Thus, SARS-CoV nsp8, together with nsp7, provides RNA primer internally complementary to the viral genomic RNA for viral replication, which also requires the "main" RdRp nsp12 (119).

Nsp7+nsp8 complex of feline coronavirus (FCoV), an alphacoronavirus, is a 2:1 heterotrimer containing two conformational different nsp7 molecules and one nsp8 molecule. Two copies of heterotrimers could bind to each other through nsp8-nsp8 interaction and form a heterohexamer (120). This nsp7+nsp8 complex is also capable of synthesizing short oligonucleotides (120). Similar to FCoV, the alphacoronavirus HCoV-299E nsp7-10 polyprotein has this noncanonical RdRp activity as well (120).

However, due to the internal initiation nature of the nsp8 primase, the model including primer synthesis of nsp8 plus the primer-dependent RdRp activity of nsp12 still could not provide mechanism insights for the viral RNA

synthesis of the 5'-end. On the other hand, a later study on the nsp7+nsp8+nsp12 complex showed that this protein complex possesses both *de novo* initiation and primer extension RdRp activities. In this work, the nsp7+nsp8 complex indeed misses the *de novo* initiation activity, suggesting that this activity is mediated by nsp12 (119). Furthermore, the primase activity of nsp8 was not observed (119). Recently a single particle cryo-electron microscopic structure of nsp12 with nsp7 and nsp8 shows a heterodimer of nsp7+nsp8 as well as second nsp8 subunit binding to the N-terminal region of nsp12 (44). This structure favors biochemically established *de novo* initiation activity of nsp7+nsp8+nsp12 complex (119), where nsp7+nsp8 does not mediate RNA primer synthesis or form the higher-order oligomer (26, 117).

In SARS-CoV infected Vero cells, nsp8 can be detected as two forms, a 22 kDa full-length protein and a ~15 kDa version (65). The later was confirmed to be N-terminally truncated version (nsp8C) by western blotting analysis using an antibody only recognizing the C-terminal part (121). Nsp7+nsp8C forms a structure that displays the ability to fuse into the nsp7+nsp8 hexadecamer and was proposed to help the virus switch the replication to genome assembly (121).

Nsp9: the dimer forming RNA binding protein

Nsp9 is a ~12kDa proteolytic cleavage product of pp1a that has the nucleic acid binding activity (45, 122). It preferentially binds to single-stranded RNA (45, 122, 123). Biotin pull-down assay showed that IBV nsp9 preferentially interacts with the 3'-UTR region of the positive-strand viral RNA (124, 125). Nsp9 can interact with itself as well as the non-canonical RdRp nsp8 (124-127). Like most of the coronavirus nonstructural proteins, it locates in the viral TRCs (65, 128).

The crystal structure of the nsp9 monomer revealed a cone-shaped N-terminal β-barrel composed of seven β-strands and a C-terminal α-helix that is conserved among alpha-, beta, and gamma-CoVs (45, 122, 123, 129, 130). However, its dimerization varies among different coronaviruses. SARS-CoV, IBV, and porcine delta coronavirus (PDCoV) nsp9 were reported to form the "parallel helix-dimer" structure that stabilized by the hydrophobic interactions between two C-terminal α-helices. PDCoV nsp9 dimer is slightly different in that it also requires the N-terminal extended finger motif to stabilize the dimer structure. Besides, SARS-CoV nsp9 can form the "sheet-dimer" structure formed by interactions between β-strand five from both subunits (45). HCoV-229E nsp9 forms an "anti-parallel helix dimer" that requires interaction between two α-helices in the opposite direction with disulfide-bond from two nsp9 subunits (123). The porcine epidemic diarrhea virus (PEDV) nsp9 forms two possible dimer structure resembles the "parallel helix-dimer" and "sheet-dimer" stabilized by a disulfide-linkage (129).

Albeit various of dimerization structures, the dimer formation could enhance the nucleic acid binding and viral replication (123, 129-131). Mutations of the protein-protein interaction motif GXXXG on the C-terminal α-helix of SARS-CoV nsp9 disrupted dimer formation, and significantly decreased RNA binding of nsp9. The corresponding mutations in the SARS-CoV genome were either lethal to the viral growth or reverted to wt type amino acid coding (131). Similarly, the G98D mutant of IBV nsp9 significantly destabilized homodimer and also abolished the activity of RNA binding. The incorporated viral mutant was deficient in subgenomic RNAs transcription as well as viral growth. Interestingly, IBV nsp9 mutation I95N showed almost no effect on the RNA binding activity but moderately destabilized dimer formation, while the virus with this mutation has severe growth defects(132). Nsp9 dimerization may have essential roles in replication beyond the RNA binding.

Nsp10: cofactor in viral replication

Nsp10 is a zinc finger protein that contains two zinc finger domains conserved among coronaviruses (49, 50, 133). Several oligomer forms were reported for nsp10. MHV nsp10 shows the monomeric form in reducing SDS-PAGE and gel filtration analysis. At the same time, it also forms ~80 kDa and ~19600 kDa protein complexes when supplemented with zinc ions in dynamic light scattering assay (133). SARS-CoV nsp10 was reported to form a dimer in solution analyzed by gel filtration (50). A simultaneous report on SARS-CoV nsp10 also revealed a dodecameric structure (49). Currently, no evidence confirmed the biological relevance of nsp10 oligomerization in viral replication.

Nsp10 is essential to coronavirus infection. A temperature-sensitive mutation of nsp10 (Q65E) significantly inhibited MHV RNA synthesis at the nonpermissive temperature (134). Furthermore, the reverse-genetics study identified 16 nsp10 mutants of MHV clone, of which eight were viable but displayed attenuated viral growth, while the other eight clones were inviable (135). One of the nsp10 mutant (D47A/H48A) was studied into depth and had subtle effects on nsp4-10/11 polyprotein processing.

During viral replication, nsp10 enhances the enzyme activities of other replication proteins (136-138). SARS-CoV nsp10 interacts with the exoribonuclease domain of nsp14, resulting in significantly increased exoribonuclease activity (136). Mutations on MHV nsp10 (R80A/E82A) led to increased sensitivity of the virus towards RNA mutagen treatments (25), suggestive of the involvement of nsp10 in coronavirus proof-reading function, which relies on the exoribonuclease activity of nsp14. A heterodimer complex structure was also identified for nsp10/nsp16 (137). Nsp10 interacts with the nsp16 S-adenosyl-L-methionine (SAM)-binding pocket and stimulates the association of both the methyl donor SAM and capped RNA acceptor to nsp16 (137), thus activates nsp16 to methylate coronaviral mRNA cap at the 2'O-site (139).

In addition to the role as a part of viral TRCs, nsp10 is involved in the development of viral cytopathic effects. SARS-CoV nsp10 interacts with human NADH 4L subunit and cytochrome oxidase II and alters the activity of the NADH-cytochrome (140). Through these interactions,

nsp10 caused an impaired oxidoreductase system and induced the depolarization of the mitochondria inner membrane (140).

Nsp11: small peptide with unknown function

Nsp11 is a small peptide located at the C-terminus of ORF1a. A three-stemmed mRNA pseudoknot containing a typical hepta-nucleotide sequence UUUAAAC is situated in the nsp11 coding region (141). This RNA structure results in a programmed -1 ribosomal frameshift, which leads to the production of ORF1ab (141, 142).

A proximity-labeling experiment identified the existence of nearly all nonstructural proteins in the microenvironment of MHV replication complexes, except for nsp11 (128). SARS-CoV nsp11 does not interact with other nonstructural proteins in the mammalian two-hybrid assay (124). Currently, no known function was discovered for nsp11. According to these data, nsp11 is likely not a member of the replication complexes. However, the exact role of nsp11 remains to be explored.

Nsp12: the main RdRp

Nsp12 is the first nonstructural protein encoded by ORF1b and functions as the primary RNA-dependent RNA polymerase of coronaviruses. Nsp12 is at the center of the viral TRCs, which participate in both the synthesis of new full-length genomic RNA and the discontinuous transcription of subgenomic RNAs (18, 143-146).

Coronavirus nsp12 mainly contains two functional domains. The C-terminal portion of nsp12 is the canonical RdRp domain that resembles a cupped right hand with fingers, palm, and the thumb holding the template RNA (44, 147). The palm subdomain is the catalytic core that contains a conserved SDD motif in the active site. Like other positive-strand RNA viruses, mutations on the SDD motif abolished its RdRp activity (148). Asides from the C-terminal polymerase domain, nsp12 also contains a Nidovirus RdRp-associated nucleotidyltransferase (NiRAN) domain, which is unique to the *Nidovirales* (44, 147). The arterivirus equine arteritis virus (EAV) nsp9 is the homolog of coronaviruses nsp12. The NiRAN domain of EAV nsp9 can be nucleotidylated, as a phosphoamide bond can be formed in between the protein and a GTP or UTP molecule (149). Single-particle cryo-EM imaging shows a structure where the NiRAN domain of nsp12 interacts with an nsp7+nsp8 heterodimer as well as a second nsp8 (44). The interaction with nsp7 and nsp8 cofactors seems to help stabilize the nsp12 RNA binding region as well as extending the RNA-binding surface (44). Genetic studies also supported the essential role of this nucleotidylating activity of nsp12 in EAV and SARS-CoV replication (149); however, the exact function of this activity is still unknown.

The SARS-CoV nsp12 RdRp activity was investigated after the outbreak of the SARS epidemic in 2002-2003 by several groups. Early studies using recombinant nsp12 showed a primer-dependent RNA polymerase activity (118, 119, 150). This primer-dependent RdRp activity of nsp12 was proposed to work with the nsp8 primase for the viral genome synthesis (118, 150). However, biochemical data of nsp8 lacking the *de novo* initiation RNA synthesis activity could not fill the gap in virus replication cycles, as to how the virus maintains its 5'-end can not be explained. The biochemical assay by using recombinant nsp12 and nsp7/nsp8 cofactors showed *de novo* initiation ability of nsp12 (119), while the nsp8 primase activity could not be detected. Cryo-EM structure of this complex supported the *de novo* initiation activity for the nsp12 polymerase. This structure clearly showed that the active site of the nsp8 primase could not fit into the nsp12 RNA synthesis pocket (44), further confirmed the biochemical finding that nsp8 does not have primase activity.

Nsp13: the helicase

Helicases are the enzymes that unwind double-stranded DNAs or RNAs (151, 152). RNA viruses encode RNA helicases (153) or recruit host alternatives (154-156) to promote their genome replication and viral gene expression. Apart from the "pro-viral" functions, RNA helicases are also involved in host antiviral responses (157). Explicitly, both animal and plant host innate immune systems encode RNA helicases to recognize and respond to foreign double-stranded RNA in the cytoplasm (157-160).

Coronavirus nsp13 encodes a C-terminal helicase domain that belongs to the superfamily one helicases (152). While at the N-terminus of nsp13 is a zinc-binding domain (ZBD), which is conserved among the members in *Nidovirales* (142, 161, 162). Nsp13 exhibits both RNA and DNA duplex unwinding activities in vitro shown by the biochemical study of recombinant nsp13 from HCoV-229E (163, 164) and SARS-CoV (165, 166). Nsp13 unwinds its substrates in a 5'- 3' direction by using the energy generated from NTPs and dNTPs hydrolysis, with the most effectivity towards ATP, dATP, and GTP. Transient kinetic analysis showed that SARS-CoV nsp13 unwinds nucleic acid in discrete steps of 9.3 bp each, with a catalytic rate of 30 steps per second (167). Moreover, the unwinding activity can be enhanced 2-fold by nsp12 through nsp12-nsp13 interaction (167). Nsp13 preferentially binds to 5'-overhang and processes the double-strand with higher activity (168). ZBD is essential for helicase activity, and replacement of conserved ZBD Cys and His residues disrupted ATPase activities of HCoV-229E nsp13 (161). In addition to NTPase and dNTPase activities, RNA 5'-triphosphatase activity was discovered for HCoV-229E and SARS-CoV nsp13, which may catalyze the first step in the formation of the 5'-cap structure of viral RNAs (164, 165). The NTPase activity of arterivirus nsp10, the homolog of coronavirus nsp13, is essential to the viral survival (169). While mutation (A335V) on the RNA binding channel of MHV nsp13 conferred decreased viral replication both *in vitro* and *in vivo* (170).

Nsp14: dual-functional RNA modifier

The CoV nsp14 has dual functions in viral RNA processing (142). The N-terminus of nsp14 is a 3'-5' exonuclease (ExoN) (19). Nsp14 ExoN belongs to the

DEDD superfamily of exonucleases. The exonuclease activity that acts on both ssRNA and dsRNA, but can not hydrolyze DNA or ribose- 2'-O-methylated RNA substrates *in vitro* (171). The ExoN activity can be stimulated to >35 fold by interacting with nsp10, and this complex can release one mismatched nucleotide from the 3'-end of the newly synthesized RNA strand (172). Coronavirus nsp14 ExoN activity is involved in the RNA proofreading machinery during viral replication. Mutant MHV or SARS-CoV with deficient ExoN activity displays a higher mutation rate and lower replication fidelity (173, 174). Moreover, ExoN is related to the host innate immune response. MHV lacking ExoN activity showed increased sensitivity to cellular pretreatment with IFN-β (175). On the contrary, TGEV nsp14 probably is responsible for the induction of IFN-β production through interaction with cellular RNA helicase DDX1 (176).

The C-terminus of nsp14 encodes for the guanosine N7-methyltransferase activity (177). When SARS-CoV or TGEV nsp14 was overexpressed in yeast null mutant of mRNA guanine 7-methyltransferase *abd1*, the growth-deficient phenotype was restored (177). The nsp14 specifically methylates GTP, dGTP, or the inverted guanosine molecule attached to the 5'-end of RNA (178). When a point mutation at the MTase domain (D331A) was incorporated into SARS-CoV replicon carrying a luciferase reporter, the luciferase activity dropped to 10%, and the subgenomic RNA accumulation dropped to 19% of the wt level (177). The defect in viral transcription and gene expression is likely due to the unstable viral RNA produced by the virus.

Nsp15: endoribonuclease

Nsp15 is a uridine-specific endoribonuclease (EndoU) that processes RNA in the viral replication/transcription complex (RTC) (31). It is a signature molecular marker for vertebrate *Nidovirales* (179, 180), as insect nidoviruses and viruses outside the *Nidovirales* do not encode such endoribonuclease (19, 181-183).

Nsp15 forms a hexamer (46, 48, 184) that depends on manganese as a cofactor for its ribonuclease activity (185-187). Nsp15 can process both ssRNA and dsRNA, but not DNA (185). Blocking the 5'- or the 3'- ends of substrate RNAs did not prevent the RNA degradation, suggestive of the activity is mainly towards the middle (endo) portion of the RNA substrate (185, 188). A mass spectrometry analysis of SARS-CoV nsp15 digested RNA products revealed that the major RNA cleavage site is the 3' of the uridylate. The 3' of cytidylate in favored sequence contexts can also be the site for cleavage (189). The hexamer form of SARS-CoV nsp15 was found to be responsible for RNA binding (189). Other studies on MHV nsp15 showed a higher binding affinity to RNA and similar ribonuclease activity of the monomer form (47).

Early genetic attempts using mutations of the vital amino acids in the MHV nsp15 catalytic pocket found that nsp15 deficient caused decreased RNA replication and viral growth in cell culture (190). Breakthrough in understanding its unique function was made a decade later by testing viral replication in mouse macrophages and *in vivo* (24, 191). The nsp15 EndoU-deficient mutant MHV stimulated an early induction of cytosolic dsRNA during infection, led to robust induction of IFN-I and PKR-mediated apoptosis, and exhibited impaired viral growth (24, 191). Moreover, the infection of the mutant virus was restricted in primary cells *in vivo* and could not efficiently spread (24). Thus the endoribonuclease activity of nsp15 promotes digestion of excessive viral dsRNAs at the replication sites and mediates viral evasion of host dsRNA-mediated innate immunity at the early stage of infection (186).

Nsp16: 2′-O-methyltransferase

Nsp16 encodes 2′-O-methyltransferase, which can methylate the RNA cap at ribose 2'-O positions resulting in a cap-1 structure (192). The 2'-O-MTase activity was predicted for SARS-CoV nsp16 by 3D modeling soon after the SARS epidemic in 2003 (192). However, the 2'-O-MTase activity was first demonstrated for FCoV nsp16 (193). On the contrary, SARS-CoV nsp16 along is inactive. It requires nsp10 as a cofactor to bind to SAM and m7GpppA-RNA substrate (137) and depends on nsp10 for its 2'-O-MTase activity (139). The crystal structure of SARS-CoV nsp16 posts a non-canonical SAM-dependent 2'-O-MTase structure lacking two out of seven α-helices, which may explain the requirements of nsp10 for its activity (60, 137).

SARS-CoV or MERS-CoV nsp16 mutants that contain mutations on conserved KDKE motif strongly attenuated viral infection *in vitro* and *in vivo* (34, 194). As common strategies among various RNA viruses to counteract innate immunity (192), the cap-1 type of modification help coronavirus evade the RNA recognition machinery and the antiviral responses mediated by IFN-I (34, 194). Viruses defective in the 2'-O-MTase activity showed increased sensitivity to IFN-I treatment comparing to the wt virus (34, 194). Host cytoplasmic RNA sensor Mda5 was shown to recognize those viral transcripts produced by the nsp16 mutant virus, as in the absence of Mda5, the replication and virulence of the mutant virus was restored (34, 194).

Nonstructural proteins are useful drug targets

Phylogenetic studies and serological evidence shows that the human-infecting betacoronaviruses, including highly pathogenic SARS-CoV, SARS-CoV-2, and MERS-CoV, have animal origins (9, 63, 195-197). Bats are identified as the natural reservoirs for human coronaviruses (198-200). Bats usually do not display signs of disease when infected with coronaviruses and have evolved an immune system that allows virus propagation (201, 202). Due to the increasing human activities and global warming that result in the changing of bat habitats, the emergency of new zoonotic coronaviral diseases are very likely to occur (199-201). This consensus demands the development of novel anticoronaviral medicines. The nonstructural proteins or the viral replication processes of coronaviruses have been shown as potential antiviral drug targets (203).

During the epidemic of SARS in 2003, structure simulation of SARS-CoV nsp5 $3CL^{pro}$ suggested that the nsp5 substrate-binding sites are conserved with rhinovirus $3C^{pro}$ and can be targeted with the antirhinoviral drug AG7088 (204). Although AG7088 was soon demonstrated to be ineffective in inhibiting SARS-CoV replication (205), the HIV protease inhibitor lopinavir-ritonavir showed broad-spectrum antiviral activity targeting SARS-CoV, MERS-CoV $3CL^{pro}$, HCoV-229E, and HCoV-NL63 (64, 206-212). Immediately after the COVID-19 outbreak, several clinical studies for the use of protease inhibitors, such as lopinavir-ritonavir, darunavir-cobicistat, and ASC09F were started (213, 214). Unfortunately, one recently finished clinical trial showed that lopinavir-ritonavir treatment provides no therapeutic benefit to hospitalized severe COVID-19 adult patients (215). Similar to nsp5 $3CL^{pro}$, the PL^{pro} domain of nsp3 is also a drug target (216-218). Among the PL^{pro} inhibitors, disulfiram is clinically available for chronic alcohol dependence relief (216).

The coronaviral nsp12 RdRp serves as an important drug target. Nucleotide analogs can directly compete with nucleotide substrates of RdRp, resulting in the halt of the reaction as well as disruption of the viral replication (147). A recent report showed that remdesivir, an adenosine analog, can efficiently inhibit viral infection in SARS-CoV-2 sensitive Huh-7 cell (219). The first SARS-CoV-2 patient in the United States administered remdesivir under the protocol of "compassionate use" and showed improved clinical conditions in about 24 hours and finally discharged from hospital, suggestive of a possible efficacy of remdesivir against coronavirus in this individual case (8). The uses of remdesivir and several other RdRp inhibitors for the treatment of COVID-19 are currently in clinical trials (213, 214).

Helicase domain of nsp13 also showed promise as a potential target. SSYA10-001, a 1,2,4-triazole derivative, can block the unwinding activity of nsp13 in a non-competitive manner (220), while myricetin and scutellarein suppress the ATPase activity of nsp13 (221). Furthermore, the adamantane-derived bananins can inhibit both ATPase and helicase activity of nsp13 and cause decreased viral infection in cell culture (222). However, none of these drug candidates went to clinical trials by the end of February 2020 (214).

Summary

Significant progress has been made in the understanding of the coronavirus nonstructural proteins, especially after the SARS epidemic in 2003. Most of the studies utilized cultured cells to investigate coronaviral infection, but the transgenic humanized mouse model also played critical roles in dissecting viral pathogenesis as well as aiding drug discoveries (223-225). These studies of coronaviral nonstructural proteins provided in-depth knowledge of how the viruses establish their infection and will continue to aid the discovery of new drugs effectively against coronaviruses.

Acknowledgements

The authors thank Dr. Zhenghe Li (Zhejiang University, China), Ms. Jiaojiao Zhang, Mr. Wang Zhang, Ms. Lingyun Zhou, and Ms. Ruoning Jin for the critical reading of the manuscript. This work was supported by the National Natural Science Foundation of China (31770164), Natural Science Foundation of Jiangsu Province (BK20180039), the Program for Jiangsu Excellent Scientific and Technological Innovation Team (17CXTD00014), and the Priority Academic Program Development (PAPD) of Jiangsu Higher Education Institutions.

Conflict of interest

The authors declare no conflict of interest concerning this review.

References

1. Cyranoski D. China joins investigation of mystery pneumonia. Nature. 2003;422(6931):459.
2. Pearson H. Mystery virus slow to yield its identity as patient numbers rise. Nature. 2003;422(6930):364.
3. Assiri A, McGeer A, Perl TM, Price CS, Al Rabeeah AA, Cummings DA, et al. Hospital outbreak of Middle East respiratory syndrome coronavirus. N Engl J Med. 2013;369(5):407-16.
4. Zaki AM, van Boheemen S, Bestebroer TM, Osterhaus AD, Fouchier RA. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. N Engl J Med. 2012;367(19):1814-20.
5. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. N Engl J Med. 2020;382(8):727-33.
6. Wang C, Horby PW, Hayden FG, Gao GF. A novel coronavirus outbreak of global health concern. Lancet. 2020;395(10223):470-3.
7. Lu R, Zhao X, Li J, Niu P, Yang B, Wu H, et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. Lancet. 2020;395(10224):565-74.
8. Holshue ML, DeBolt C, Lindquist S, Lofy KH, Wiesman J, Bruce H, et al. First Case of 2019 Novel Coronavirus in the United States. N Engl J Med. 2020;382(10):929-36.
9. Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature. 2020.
10. International Committee on Taxonomy of Viruses., King AMQ. Virus taxonomy : classification and nomenclature of viruses : ninth report of the International Committee on Taxonomy of Viruses. London ; Waltham, MA: Academic Press; 2012. x, 1327 p. p.
11. Pene F, Merlat A, Vabret A, Rozenberg F, Buzyn A, Dreyfus F, et al. Coronavirus 229E-related pneumonia in immunocompromised patients. Clin Infect Dis. 2003;37(7):929-32.
12. van der Hoek L, Pyrc K, Jebbink MF, Vermeulen-Oost W, Berkhout RJ, Wolthers KC, et al. Identification of a new human coronavirus. Nat Med. 2004;10(4):368-73.
13. Lau SK, Lee P, Tsang AK, Yip CC, Tse H, Lee RA, et al. Molecular epidemiology of human coronavirus OC43 reveals evolution of different genotypes over time and recent emergence of a novel genotype due to natural recombination. J Virol. 2011;85(21):11325-37.
14. Woo PC, Lau SK, Chu CM, Chan KH, Tsoi HW, Huang Y, et al. Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. J Virol. 2005;79(2):884-95.
15. Marra MA, Jones SJ, Astell CR, Holt RA, Brooks-Wilson A, Butterfield YS, et al. The Genome sequence of the SARS-associated coronavirus. Science. 2003;300(5624):1399-404.
16. Coronaviridae Study Group of the International Committee on Taxonomy of V. The species Severe acute respiratory

syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. Nat Microbiol. 2020 Mar 2. [Epub ahead of print]
17. Masters PS. The molecular biology of coronaviruses. Adv Virus Res. 2006;66:193-292.
18. Fehr AR, Perlman S. Coronaviruses: an overview of their replication and pathogenesis. Methods Mol Biol. 2015;1282:1-23.
19. Lauber C, Goeman JJ, Parquet Mdel C, Nga PT, Snijder EJ, Morita K, et al. The footprint of genome architecture in the largest genome expansion in RNA viruses. PLoS Pathog. 2013;9(7):e1003500.
20. Wu A, Peng Y, Huang B, Ding X, Wang X, Niu P, et al.Genome composition and divergence of the novel coronavirus (2019-nCoV) originating in China. Cell Host Microbe. 2020 Mar 11;27(3):325-28.
21. Yang X, Chen X, Bian G, Tu J, Xing Y, Wang Y, et al. Proteolytic processing, deubiquitinase and interferon antagonist activities of Middle East respiratory syndrome coronavirus papain-like protease. J Gen Virol. 2014;95(Pt 3):614-26.
22. Lin PY, Chou CY, Chang HC, Hsu WC, Chang GG. Correlation between dissociation and catalysis of SARS-CoV main protease. Arch Biochem Biophys. 2008;472(1):34-42.
23. van Hemert MJ, van den Worm SH, Knoops K, Mommaas AM, Gorbalenya AE, Snijder EJ. SARS-coronavirus replication/transcription complexes are membrane-protected and need a host factor for activity in vitro. PLoS Pathog. 2008;4(5):e1000054.
24. Kindler E, Gil-Cruz C, Spanier J, Li Y, Wilhelm J, Rabouw HH, et al. Early endonuclease-mediated evasion of RNA sensing ensures efficient coronavirus replication. PLoS Pathog. 2017;13(2):e1006195.
25. Smith EC, Case JB, Blanc H, Isakov O, Shomron N, Vignuzzi M, et al. Mutations in coronavirus nonstructural protein 10 decrease virus replication fidelity. J Virol. 2015;89(12):6418-26.
26. Zhai Y, Sun F, Li X, Pang H, Xu X, Bartlam M, et al. Insights into SARS-CoV transcription and replication from the structure of the nsp7-nsp8 hexadecamer. Nat Struct Mol Biol. 2005;12(11):980-6.
27. Sparks JS, Lu X, Denison MR. Genetic analysis of Murine hepatitis virus nsp4 in virus replication. J Virol. 2007;81(22):12554-63.
28. Knoops K, Kikkert M, Worm SH, Zevenhoven-Dobbe JC, van der Meer Y, Koster AJ, et al. SARS-coronavirus replication is supported by a reticulovesicular network of modified endoplasmic reticulum. PLoS Biol. 2008;6(9):e226.
29. Fehr AR, Jankevicius G, Ahel I, Perlman S. Viral Macrodomains: Unique Mediators of Viral Replication and Pathogenesis. Trends Microbiol. 2018;26(7):598-610.
30. Graham RL, Sims AC, Brockway SM, Baric RS, Denison MR. The nsp2 replicase proteins of murine hepatitis virus and severe acute respiratory syndrome coronavirus are dispensable for viral replication. J Virol. 2005;79(21):13399-411.
31. Athmer J, Fehr AR, Grunewald M, Smith EC, Denison MR, Perlman S. In Situ Tagged nsp15 Reveals Interactions with Coronavirus Replication/Transcription Complex-Associated Proteins. mBio. 2017;8(1).
32. Kusov Y, Tan J, Alvarez E, Enjuanes L, Hilgenfeld R. A G-quadruplex-binding macrodomain within the "SARS-unique domain" is essential for the activity of the SARS-coronavirus replication-transcription complex. Virology. 2015;484:313-22.
33. Imbert I, Snijder EJ, Dimitrova M, Guillemot JC, Lecine P, Canard B. The SARS-Coronavirus PLnc domain of nsp3 as a replication/transcription scaffolding protein. Virus Res. 2008;133(2):136-48.
34. Menachery VD, Gralinski LE, Mitchell HD, Dinnon KH, 3rd, Leist SR, Yount BL, Jr., et al. Middle East Respiratory Syndrome Coronavirus Nonstructural Protein 16 Is Necessary for Interferon Resistance and Viral Pathogenesis. mSphere. 2017;2(6).
35. Alfuwaires M, Altaher A, Kandeel M. Molecular Dynamic Studies of Interferon and Innate Immunity Resistance in MERS CoV Non-Structural Protein 3. Biol Pharm Bull. 2017;40(3):345-51.
36. Fehr AR, Channappanavar R, Jankevicius G, Fett C, Zhao J, Athmer J, et al. The Conserved Coronavirus Macrodomain Promotes Virulence and Suppresses the Innate Immune Response during Severe Acute Respiratory Syndrome Coronavirus Infection. mBio. 2016;7(6).
37. Zhu X, Fang L, Wang D, Yang Y, Chen J, Ye X, et al. Porcine deltacoronavirus nsp5 inhibits interferon-beta production through the cleavage of NEMO. Virology. 2017;502:33-8.
38. Chen S, Tian J, Li Z, Kang H, Zhang J, Huang J, et al. Feline Infectious Peritonitis Virus Nsp5 Inhibits Type I Interferon Production by Cleaving NEMO at Multiple Sites. Viruses. 2019;12(1).
39. Zhu X, Wang D, Zhou J, Pan T, Chen J, Yang Y, et al. Porcine Deltacoronavirus nsp5 Antagonizes Type I Interferon Signaling by Cleaving STAT2. J Virol. 2017;91(10).
40. Yuan L, Chen Z, Song S, Wang S, Tian C, Xing G, et al. p53 degradation by a coronavirus papain-like protease suppresses type I interferon signaling. J Biol Chem. 2015;290(5):3172-82.
41. Clementz MA, Chen Z, Banach BS, Wang Y, Sun L, Ratia K, et al. Deubiquitinating and interferon antagonism activities of coronavirus papain-like proteases. J Virol. 2010;84(9):4619-29.
42. Oostra M, Hagemeijer MC, van Gent M, Bekker CP, te Lintelo EG, Rottier PJ, et al. Topology and membrane anchoring of the coronavirus replication complex: not all hydrophobic domains of nsp3 and nsp6 are membrane spanning. J Virol. 2008;82(24):12392-405.
43. Hagemeijer MC, Verheije MH, Ulasli M, Shaltiel IA, de Vries LA, Reggiori F, et al. Dynamics of coronavirus replication-transcription complexes. J Virol. 2010;84(4):2134-49.
44. Kirchdoerfer RN, Ward AB. Structure of the SARS-CoV nsp12 polymerase bound to nsp7 and nsp8 co-factors. Nat Commun. 2019;10(1):2342.
45. Sutton G, Fry E, Carter L, Sainsbury S, Walter T, Nettleship J, et al. The nsp9 replicase protein of SARS-coronavirus, structure and functional insights. Structure. 2004;12(2):341-53.
46. Joseph JS, Saikatendu KS, Subramanian V, Neuman BW, Buchmeier MJ, Stevens RC, et al. 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. 2007;81(12):6700-8.
47. Xu X, Zhai Y, Sun F, Lou Z, Su D, Xu Y, et al. New antiviral target revealed by the hexameric structure of mouse hepatitis virus nonstructural protein nsp15. J Virol. 2006;80(16):7909-17.
48. Ricagno S, Egloff MP, Ulferts R, Coutard B, Nurizzo D, Campanacci V, et al. Crystal structure and mechanistic determinants of SARS coronavirus nonstructural protein 15 define an endoribonuclease family. Proc Natl Acad Sci U S A. 2006;103(32):11892-7.
49. Su D, Lou Z, Sun F, Zhai Y, Yang H, Zhang R, et al. Dodecamer structure of severe acute respiratory syndrome coronavirus nonstructural protein nsp10. J Virol. 2006;80(16):7902-8.
50. Joseph JS, Saikatendu KS, Subramanian V, Neuman BW, Brooun A, Griffith M, et al. Crystal structure of nonstructural protein 10 from the severe acute respiratory syndrome coronavirus reveals a novel fold with two zinc-binding motifs. J Virol. 2006;80(16):7894-901.
51. Hu T, Zhang Y, Li L, Wang K, Chen S, Chen J, et al. Two adjacent mutations on the dimer interface of SARS

coronavirus 3C-like protease cause different conformational changes in crystal structure. Virology. 2009;388(2):324-34.
52. Serrano P, Johnson MA, Chatterjee A, Neuman BW, Joseph JS, Buchmeier MJ, et al. Nuclear magnetic resonance structure of the nucleic acid-binding domain of severe acute respiratory syndrome coronavirus nonstructural protein 3. J Virol. 2009;83(24):12998-3008.
53. Baez-Santos YM, St John SE, Mesecar AD. The SARS-coronavirus papain-like protease: structure, function and inhibition by designed antiviral compounds. Antiviral Res. 2015;115:21-38.
54. Cho CC, Lin MH, Chuang CY, Hsu CH. Macro domain from middle east respiratory syndrome coronavirus (MERS-CoV)is an efficient ADP-ribose binding module: crystal structure and biochemical studies. J Biol Chem. 2016;291(10):4894-902.
55. Xu Y, Cong L, Chen C, Wei L, Zhao Q, Xu X, et al. 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. 2009;83(2):1083-92.
56. Piotrowski Y, Hansen G, Boomaars-van der Zanden AL, Snijder EJ, Gorbalenya AE, Hilgenfeld R. 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. 2009;18(1):6-16.
57. Keane SC, Giedroc DP. Solution structure of mouse hepatitis virus (MHV) nsp3a and determinants of the interaction with MHV nucleocapsid (N) protein. J Virol. 2013;87(6):3502-15.
58. Serrano P, Johnson MA, Almeida MS, Horst R, Herrmann T, Joseph JS, et al. Nuclear magnetic resonance structure of the N-terminal domain of nonstructural protein 3 from the severe acute respiratory syndrome coronavirus. J Virol. 2007;81(21):12049-60.
59. Han W, Li X, Fu X. The macro domain protein family: structure, functions, and their potential therapeutic implications. Mutat Res. 2011;727(3):86-103.
60. Subissi L, Imbert I, Ferron F, Collet A, Coutard B, Decroly E, et al. SARS-CoV ORF1b-encoded nonstructural proteins 12-16: replicative enzymes as antiviral targets. Antiviral Res. 2014;101:122-30.
61. Lundin A, Dijkman R, Bergstrom T, Kann N, Adamiak B, Hannoun C, et al. Targeting membrane-bound viral RNA synthesis reveals potent inhibition of diverse coronaviruses including the middle East respiratory syndrome virus. PLoS Pathog. 2014;10(5):e1004166.
62. Pfefferle S, Schopf J, Kogl M, Friedel CC, Muller MA, Carbajo-Lozoya J, et al. The SARS-coronavirus-host interactome: identification of cyclophilins as target for pan-coronavirus inhibitors. PLoS Pathog. 2011;7(10):e1002331.
63. Chan JF, Kok KH, Zhu Z, Chu H, To KK, Yuan S, et al. Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan. Emerg Microbes Infect. 2020;9(1):221-36.
64. Kilianski A, Mielech AM, Deng X, Baker SC. Assessing activity and inhibition of Middle East respiratory syndrome coronavirus papain-like and 3C-like proteases using luciferase-based biosensors. J Virol. 2013;87(21):11955-62.
65. Prentice E, McAuliffe J, Lu X, Subbarao K, Denison MR. Identification and characterization of severe acute respiratory syndrome coronavirus replicase proteins. J Virol. 2004;78(18):9977-86.
66. Kamitani W, Narayanan K, Huang C, Lokugamage K, Ikegami T, Ito N, et al. Severe acute respiratory syndrome coronavirus nsp1 protein suppresses host gene expression by promoting host mRNA degradation. Proc Natl Acad Sci U S A. 2006;103(34):12885-90.
67. Zust R, Cervantes-Barragan L, Kuri T, Blakqori G, Weber F, Ludewig B, et al. Coronavirus non-structural protein 1 is a major pathogenicity factor: implications for the rational design of coronavirus vaccines. PLoS Pathog. 2007;3(8):e109.
68. Tohya Y, Narayanan K, Kamitani W, Huang C, Lokugamage K, Makino S. Suppression of host gene expression by nsp1 proteins of group 2 bat coronaviruses. J Virol. 2009;83(10):5282-8.
69. Kamitani W, Huang C, Narayanan K, Lokugamage KG, Makino S. A two-pronged strategy to suppress host protein synthesis by SARS coronavirus Nsp1 protein. Nat Struct Mol Biol. 2009;16(11):1134-40.
70. Huang C, Lokugamage KG, Rozovics JM, Narayanan K, Semler BL, Makino S. SARS coronavirus nsp1 protein induces template-dependent endonucleolytic cleavage of mRNAs: viral mRNAs are resistant to nsp1-induced RNA cleavage. PLoS Pathog. 2011;7(12):e1002433.
71. Gaglia MM, Covarrubias S, Wong W, Glaunsinger BA. A common strategy for host RNA degradation by divergent viruses. J Virol. 2012;86(17):9527-30.
72. Tanaka T, Kamitani W, DeDiego ML, Enjuanes L, Matsuura Y. Severe acute respiratory syndrome coronavirus nsp1 facilitates efficient propagation in cells through a specific translational shutoff of host mRNA. J Virol. 2012;86(20):11128-37.
73. Terada Y, Kawachi K, Matsuura Y, Kamitani W. MERS coronavirus nsp1 participates in an efficient propagation through a specific interaction with viral RNA. Virology. 2017;511:95-105.
74. Menachery VD, Graham RL, Baric RS. Jumping species-a mechanism for coronavirus persistence and survival. Curr Opin Virol. 2017;23:1-7.
75. Graham RL, Sims AC, Baric RS, Denison MR. The nsp2 proteins of mouse hepatitis virus and SARS coronavirus are dispensable for viral replication. Adv Exp Med Biol. 2006;581:67-72.
76. Gadlage MJ, Graham RL, Denison MR. Murine coronaviruses encoding nsp2 at different genomic loci have altered replication, protein expression, and localization. J Virol. 2008;82(23):11964-9.
77. Angelini MM, Akhlaghpour M, Neuman BW, Buchmeier MJ. Severe acute respiratory syndrome coronavirus nonstructural proteins 3, 4, and 6 induce double-membrane vesicles. mBio. 2013;4(4).
78. Hagemeijer MC, Monastyrska I, Griffith J, van der Sluijs P, Voortman J, van Bergen en Henegouwen PM, et al. Membrane rearrangements mediated by coronavirus nonstructural proteins 3 and 4. Virology. 2014;458-459:125-35.
79. Hurst KR, Koetzner CA, Masters PS. Characterization of a critical interaction between the coronavirus nucleocapsid protein and nonstructural protein 3 of the viral replicase-transcriptase complex. J Virol. 2013;87(16):9159-72.
80. Hurst KR, Ye R, Goebel SJ, Jayaraman P, Masters PS. An interaction between the nucleocapsid protein and a component of the replicase-transcriptase complex is crucial for the infectivity of coronavirus genomic RNA. J Virol. 2010;84(19):10276-88.
81. Fehr AR, Athmer J, Channappanavar R, Phillips JM, Meyerholz DK, Perlman S. The nsp3 macrodomain promotes virulence in mice with coronavirus-induced encephalitis. J Virol. 2015;89(3):1523-36.
82. Lei J, Kusov Y, Hilgenfeld R. Nsp3 of coronaviruses: Structures and functions of a large multi-domain protein. Antiviral Res. 2018;149:58-74.
83. Davis ME, Gack MU. Ubiquitination in the antiviral immune response. Virology. 2015;479-480:52-65.
84. Heaton SM, Borg NA, Dixit VM. Ubiquitin in the activation and attenuation of innate antiviral immunity. J Exp Med. 2016;213(1):1-13.
85. Morales DJ, Lenschow DJ. The antiviral activities of ISG15. J Mol Biol. 2013;425(24):4995-5008.
86. Perng YC, Lenschow DJ. ISG15 in antiviral immunity and beyond. Nat Rev Microbiol. 2018;16(7):423-39.

87. Koonin EV, Gorbalenya AE, Purdy MA, Rozanov MN, Reyes GR, Bradley DW. 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. 1992;89(17):8259-63.
88. Gorbalenya AE, Koonin EV, Lai MM. Putative papain-related thiol proteases of positive-strand RNA viruses. Identification of rubi- and aphthovirus proteases and delineation of a novel conserved domain associated with proteases of rubi-, alpha- and coronaviruses. FEBS Lett. 1991;288(1-2):201-5.
89. Dominguez G, Wang CY, Frey TK. Sequence of the genome RNA of rubella virus: evidence for genetic rearrangement during togavirus evolution. Virology. 1990;177(1):225-38.
90. Saikatendu KS, Joseph JS, Subramanian V, Clayton T, Griffith M, Moy K, et al. Structural basis of severe acute respiratory syndrome coronavirus ADP-ribose-1"-phosphate dephosphorylation by a conserved domain of nsP3. Structure. 2005;13(11):1665-75.
91. Egloff MP, Malet H, Putics A, Heinonen M, Dutartre H, Frangeul A, et al. Structural and functional basis for ADP-ribose and poly(ADP-ribose) binding by viral macro domains. J Virol. 2006;80(17):8493-502.
92. Costanzi C, Pehrson JR. Histone macroH2A1 is concentrated in the inactive X chromosome of female mammals. Nature. 1998;393(6685):599-601.
93. Li C, Debing Y, Jankevicius G, Neyts J, Ahel I, Coutard B, et al. Viral Macro Domains Reverse Protein ADP-Ribosylation. J Virol. 2016;90(19):8478-86.
94. Neuman BW. Bioinformatics and functional analyses of coronavirus nonstructural proteins involved in the formation of replicative organelles. Antiviral Res. 2016;135:97-107.
95. Johnson MA, Chatterjee A, Neuman BW, Wuthrich K. SARS coronavirus unique domain: three-domain molecular architecture in solution and RNA binding. J Mol Biol. 2010;400(4):724-42.
96. Tan J, Kusov Y, Mutschall D, Tech S, Nagarajan K, Hilgenfeld R, et al. The "SARS-unique domain" (SUD) of SARS coronavirus is an oligo(G)-binding protein. Biochem Biophys Res Commun. 2007;364(4):877-82.
97. Bencze KZ, Kondapalli KC, Cook JD, McMahon S, Millan-Pacheco C, Pastor N, et al. The structure and function of frataxin. Crit Rev Biochem Mol Biol. 2006;41(5):269-91.
98. Chen Y, Savinov SN, Mielech AM, Cao T, Baker SC, Mesecar AD. X-ray Structural and Functional Studies of the Three Tandemly Linked Domains of Non-structural Protein 3 (nsp3) from Murine Hepatitis Virus Reveal Conserved Functions. J Biol Chem. 2015;290(42):25293-306.
99. Frieman M, Ratia K, Johnston RE, Mesecar AD, Baric RS. Severe acute respiratory syndrome coronavirus papain-like protease ubiquitin-like domain and catalytic domain regulate antagonism of IRF3 and NF-kappaB signaling. J Virol. 2009;83(13):6689-705.
100. Mielech AM, Kilianski A, Baez-Santos YM, Mesecar AD, Baker SC. MERS-CoV papain-like protease has deISGylating and deubiquitinating activities. Virology. 2014;450-451:64-70.
101. Lindner HA, Fotouhi-Ardakani N, Lytvyn V, Lachance P, Sulea T, Menard R. The papain-like protease from the severe acute respiratory syndrome coronavirus is a deubiquitinating enzyme. J Virol. 2005;79(24):15199-208.
102. Barretto N, Jukneliene D, Ratia K, Chen Z, Mesecar AD, Baker SC. The papain-like protease of severe acute respiratory syndrome coronavirus has deubiquitinating activity. J Virol. 2005;79(24):15189-98.
103. Malynn BA, Ma A. Ubiquitin makes its mark on immune regulation. Immunity. 2010;33(6):843-52.
104. Tanaka Y, Chen ZJ. STING specifies IRF3 phosphorylation by TBK1 in the cytosolic DNA signaling pathway. Sci Signal. 2012;5(214):ra20.
105. Li SW, Wang CY, Jou YJ, Huang SH, Hsiao LH, Wan L, et al. SARS Coronavirus Papain-Like Protease Inhibits the TLR7 Signaling Pathway through Removing Lys63-Linked Polyubiquitination of TRAF3 and TRAF6. Int J Mol Sci. 2016;17(5).
106. Ma-Lauer Y, Carbajo-Lozoya J, Hein MY, Muller MA, Deng W, Lei J, et al. p53 down-regulates SARS coronavirus replication and is targeted by the SARS-unique domain and PLpro via E3 ubiquitin ligase RCHY1. Proc Natl Acad Sci U S A. 2016;113(35):E5192-201.
107. Oudshoorn D, Rijs K, Limpens R, Groen K, Koster AJ, Snijder EJ, et al. Expression and Cleavage of Middle East Respiratory Syndrome Coronavirus nsp3-4 Polyprotein Induce the Formation of Double-Membrane Vesicles That Mimic Those Associated with Coronaviral RNA Replication. mBio. 2017;8(6).
108. Gadlage MJ, Sparks JS, Beachboard DC, Cox RG, Doyle JD, Stobart CC, et al. Murine hepatitis virus nonstructural protein 4 regulates virus-induced membrane modifications and replication complex function. J Virol. 2010;84(1):280-90.
109. Sakai Y, Kawachi K, Terada Y, Omori H, Matsuura Y, Kamitani W. Two-amino acids change in the nsp4 of SARS coronavirus abolishes viral replication. Virology. 2017;510:165-74.
110. Zhang S, Zhong N, Xue F, Kang X, Ren X, Chen J, et al. Three-dimensional domain swapping as a mechanism to lock the active conformation in a super-active octamer of SARS-CoV main protease. Protein Cell. 2010;1(4):371-83.
111. Shi J, Sivaraman J, Song J. Mechanism for controlling the dimer-monomer switch and coupling dimerization to catalysis of the severe acute respiratory syndrome coronavirus 3C-like protease. J Virol. 2008;82(9):4620-9.
112. Fan K, Wei P, Feng Q, Chen S, Huang C, Ma L, et al. Biosynthesis, purification, and substrate specificity of severe acute respiratory syndrome coronavirus 3C-like proteinase. J Biol Chem. 2004;279(3):1637-42.
113. Cottam EM, Maier HJ, Manifava M, Vaux LC, Chandra-Schoenfelder P, Gerner W, et al. Coronavirus nsp6 proteins generate autophagosomes from the endoplasmic reticulum via an omegasome intermediate. Autophagy. 2011;7(11):1335-47.
114. Reggiori F, Monastyrska I, Verheije MH, Cali T, Ulasli M, Bianchi S, et al. Coronaviruses Hijack the LC3-I-positive EDEMosomes, ER-derived vesicles exporting short-lived ERAD regulators, for replication. Cell Host Microbe. 2010;7(6):500-8.
115. Zhao Z, Thackray LB, Miller BC, Lynn TM, Becker MM, Ward E, et al. Coronavirus replication does not require the autophagy gene ATG5. Autophagy. 2007;3(6):581-5.
116. Deming DJ, Graham RL, Denison MR, Baric RS. Processing of open reading frame 1a replicase proteins nsp7 to nsp10 in murine hepatitis virus strain A59 replication. J Virol. 2007;81(19):10280-91.
117. Imbert I, Guillemot JC, Bourhis JM, Bussetta C, Coutard B, Egloff MP, et al. A second, non-canonical RNA-dependent RNA polymerase in SARS coronavirus. EMBO J. 2006;25(20):4933-42.
118. te Velthuis AJ, van den Worm SH, Snijder EJ. The SARS-coronavirus nsp7+nsp8 complex is a unique multimeric RNA polymerase capable of both de novo initiation and primer extension. Nucleic Acids Res. 2012;40(4):1737-47.
119. Subissi L, Posthuma CC, Collet A, Zevenhoven-Dobbe JC, Gorbalenya AE, Decroly E, et al. One severe acute respiratory syndrome coronavirus protein complex integrates processive RNA polymerase and exonuclease activities. Proc Natl Acad Sci U S A. 2014;111(37):E3900-9.
120. Xiao Y, Ma Q, Restle T, Shang W, Svergun DI, Ponnusamy R, et al. Nonstructural proteins 7 and 8 of feline coronavirus form a 2:1 heterotrimer that exhibits primer-independent RNA polymerase activity. J Virol. 2012;86(8):4444-54.
121. Li S, Zhao Q, Zhang Y, Zhang Y, Bartlam M, Li X, et al. New nsp8 isoform suggests mechanism for tuning viral RNA

synthesis. Protein Cell. 2010;1(2):198-204.
122. Egloff MP, Ferron F, Campanacci V, Longhi S, Rancurel C, Dutartre H, et al. The severe acute respiratory syndrome-coronavirus replicative protein nsp9 is a single-stranded RNA-binding subunit unique in the RNA virus world. Proc Natl Acad Sci U S A. 2004;101(11):3792-6.
123. Ponnusamy R, Moll R, Weimar T, Mesters JR, Hilgenfeld R. Variable oligomerization modes in coronavirus non-structural protein 9. J Mol Biol. 2008;383(5):1081-96.
124. Pan J, Peng X, Gao Y, Li Z, Lu X, Chen Y, et al. Genome-wide analysis of protein-protein interactions and involvement of viral proteins in SARS-CoV replication. PLoS One. 2008;3(10):e3299.
125. Tan YW, Fung TS, Shen H, Huang M, Liu DX. Coronavirus infectious bronchitis virus non-structural proteins 8 and 12 form stable complex independent of the non-translated regions of viral RNA and other viral proteins. Virology. 2018;513:75-84.
126. von Brunn A, Teepe C, Simpson JC, Pepperkok R, Friedel CC, Zimmer R, et al. Analysis of intraviral protein-protein interactions of the SARS coronavirus ORFeome. PLoS One. 2007;2(5):e459.
127. Bost AG, Carnahan RH, Lu XT, Denison MR. Four proteins processed from the replicase gene polyprotein of mouse hepatitis virus colocalize in the cell periphery and adjacent to sites of virion assembly. J Virol. 2000;74(7):3379-87.
128. V'Kovski P, Gerber M, Kelly J, Pfaender S, Ebert N, Braga Lagache S, et al. Determination of host proteins composing the microenvironment of coronavirus replicase complexes by proximity-labeling. Elife. 2019;8.
129. Zeng Z, Deng F, Shi K, Ye G, Wang G, Fang L, et al. Dimerization of Coronavirus nsp9 with Diverse Modes Enhances Its Nucleic Acid Binding Affinity. J Virol. 2018;92(17).
130. Hu T, Chen C, Li H, Dou Y, Zhou M, Lu D, et al. Structural basis for dimerization and RNA binding of avian infectious bronchitis virus nsp9. Protein Sci. 2017;26(5):1037-48.
131. Miknis ZJ, Donaldson EF, Umland TC, Rimmer RA, Baric RS, Schultz LW. Severe acute respiratory syndrome coronavirus nsp9 dimerization is essential for efficient viral growth. J Virol. 2009;83(7):3007-18.
132. Chen B, Fang S, Tam JP, Liu DX. Formation of stable homodimer via the C-terminal alpha-helical domain of coronavirus nonstructural protein 9 is critical for its function in viral replication. Virology. 2009;383(2):328-37.
133. Matthes N, Mesters JR, Coutard B, Canard B, Snijder EJ, Moll R, et al. The non-structural protein Nsp10 of mouse hepatitis virus binds zinc ions and nucleic acids. FEBS Lett. 2006;580(17):4143-9.
134. Sawicki SG, Sawicki DL, Younker D, Meyer Y, Thiel V, Stokes H, et al. Functional and genetic analysis of coronavirus replicase-transcriptase proteins. PLoS Pathog. 2005;1(4):e39.
135. Donaldson EF, Sims AC, Graham RL, Denison MR, Baric RS. Murine hepatitis virus replicase protein nsp10 is a critical regulator of viral RNA synthesis. J Virol. 2007;81(12):6356-68.
136. Ma Y, Wu L, Shaw N, Gao Y, Wang J, Sun Y, et al. Structural basis and functional analysis of the SARS coronavirus nsp14-nsp10 complex. Proc Natl Acad Sci U S A. 2015;112(30):9436-41.
137. Chen Y, Su C, Ke M, Jin X, Xu L, Zhang Z, et al. Biochemical and structural insights into the mechanisms of SARS coronavirus RNA ribose 2'-O-methylation by nsp16/nsp10 protein complex. PLoS Pathog. 2011;7(10):e1002294.
138. Bouvet M, Lugari A, Posthuma CC, Zevenhoven JC, Bernard S, Betzi S, et al. Coronavirus Nsp10, a critical co-factor for activation of multiple replicative enzymes. J Biol Chem. 2014;289(37):25783-96.
139. Bouvet M, Debarnot C, Imbert I, Selisko B, Snijder EJ, Canard B, et al. In vitro reconstitution of SARS-coronavirus mRNA cap methylation. PLoS Pathog. 2010;6(4):e1000863.
140. Li Q, Wang L, Dong C, Che Y, Jiang L, Liu L, et al. The interaction of the SARS coronavirus non-structural protein 10 with the cellular oxido-reductase system causes an extensive cytopathic effect. J Clin Virol. 2005;34(2):133-9.
141. Plant EP, Perez-Alvarado GC, Jacobs JL, Mukhopadhyay B, Hennig M, Dinman JD. A three-stemmed mRNA pseudoknot in the SARS coronavirus frameshift signal. PLoS Biol. 2005;3(6):e172.
142. Sola I, Almazan F, Zuniga S, Enjuanes L. Continuous and Discontinuous RNA Synthesis in Coronaviruses. Annu Rev Virol. 2015;2(1):265-88.
143. Ziebuhr J, Snijder EJ, Gorbalenya AE. Virus-encoded proteinases and proteolytic processing in the Nidovirales. J Gen Virol. 2000;81(Pt 4):853-79.
144. Ziebuhr J. The coronavirus replicase. Curr Top Microbiol Immunol. 2005;287:57-94.
145. Snijder EJ, Bredenbeek PJ, Dobbe JC, Thiel V, Ziebuhr J, Poon LL, et al. Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. J Mol Biol. 2003;331(5):991-1004.
146. Miller WA, Koev G. Synthesis of subgenomic RNAs by positive-strand RNA viruses. Virology. 2000;273(1):1-8.
147. Xu X, Liu Y, Weiss S, Arnold E, Sarafianos SG, Ding J. Molecular model of SARS coronavirus polymerase: implications for biochemical functions and drug design. Nucleic Acids Res. 2003;31(24):7117-30.
148. Ahn DG, Choi JK, Taylor DR, Oh JW. Biochemical characterization of a recombinant SARS coronavirus nsp12 RNA-dependent RNA polymerase capable of copying viral RNA templates. Arch Virol. 2012;157(11):2095-104.
149. Lehmann KC, Gulyaeva A, Zevenhoven-Dobbe JC, Janssen GM, Ruben M, Overkleeft HS, et al. Discovery of an essential nucleotidylating activity associated with a newly delineated conserved domain in the RNA polymerase-containing protein of all nidoviruses. Nucleic Acids Res. 2015;43(17):8416-34.
150. te Velthuis AJ, Arnold JJ, Cameron CE, van den Worm SH, Snijder EJ. The RNA polymerase activity of SARS-coronavirus nsp12 is primer dependent. Nucleic Acids Res. 2010;38(1):203-14.
151. Ranji A, Boris-Lawrie K. RNA helicases: emerging roles in viral replication and the host innate response. RNA Biol. 2010;7(6):775-87.
152. Singleton MR, Dillingham MS, Wigley DB. Structure and mechanism of helicases and nucleic acid translocases. Annu Rev Biochem. 2007;76:23-50.
153. Kadare G, Haenni AL. Virus-encoded RNA helicases. J Virol. 1997;71(4):2583-90.
154. Wang Y, Chen X, Xie J, Zhou S, Huang Y, Li YP, et al. RNA Helicase A Is an Important Host Factor Involved in Dengue Virus Replication. J Virol. 2019;93(4).
155. Matkovic R, Bernard E, Fontanel S, Eldin P, Chazal N, Hassan Hersi D, et al. The Host DHX9 DExH-Box Helicase Is Recruited to Chikungunya Virus Replication Complexes for Optimal Genomic RNA Translation. J Virol. 2019;93(4).
156. Kovalev N, Pogany J, Nagy PD. A Co-Opted DEAD-Box RNA helicase enhances tombusvirus plus-strand synthesis. PLoS Pathog. 2012;8(2):e1002537.
157. Wu CY, Nagy PD. Blocking tombusvirus replication through the antiviral functions of DDX17-like RH30 DEAD-box helicase. PLoS Pathog. 2019;15(5):e1007771.
158. Kato H, Takeuchi O, Sato S, Yoneyama M, Yamamoto M, Matsui K, et al. Differential roles of MDA5 and RIG-I helicases in the recognition of RNA viruses. Nature. 2006;441(7089):101-5.
159. Dalmay T, Horsefield R, Braunstein TH, Baulcombe DC. SDE3 encodes an RNA helicase required for post-transcriptional gene silencing in Arabidopsis. EMBO J. 2001;20(8):2069-78.
160. Carmell MA, Hannon GJ. RNase III enzymes and the

initiation of gene silencing. Nat Struct Mol Biol. 2004;11(3):214-8.
161. Seybert A, Posthuma CC, van Dinten LC, Snijder EJ, Gorbalenya AE, Ziebuhr J. A complex zinc finger controls the enzymatic activities of nidovirus helicases. J Virol. 2005;79(2):696-704.
162. Hao W, Wojdyla JA, Zhao R, Han R, Das R, Zlatev I, et al. Crystal structure of Middle East respiratory syndrome coronavirus helicase. PLoS Pathog. 2017;13(6):e1006474.
163. Seybert A, Hegyi A, Siddell SG, Ziebuhr J. The human coronavirus 229E superfamily 1 helicase has RNA and DNA duplex-unwinding activities with 5'-to-3' polarity. RNA. 2000;6(7):1056-68.
164. Ivanov KA, Ziebuhr J. Human coronavirus 229E nonstructural protein 13: characterization of duplex-unwinding, nucleoside triphosphatase, and RNA 5'-triphosphatase activities. J Virol. 2004;78(14):7833-8.
165. Ivanov KA, Thiel V, Dobbe JC, van der Meer Y, Snijder EJ, Ziebuhr J. Multiple enzymatic activities associated with severe acute respiratory syndrome coronavirus helicase. J Virol. 2004;78(11):5619-32.
166. Tanner JA, Watt RM, Chai YB, Lu LY, Lin MC, Peiris JS, et al. The severe acute respiratory syndrome (SARS) coronavirus NTPase/helicase belongs to a distinct class of 5' to 3' viral helicases. J Biol Chem. 2003;278(41):39578-82.
167. Adedeji AO, Marchand B, Te Velthuis AJ, Snijder EJ, Weiss S, Eoff RL, et al. Mechanism of nucleic acid unwinding by SARS-CoV helicase. PLoS One. 2012;7(5):e36521.
168. Lee NR, Kwon HM, Park K, Oh S, Jeong YJ, Kim DE. Cooperative translocation enhances the unwinding of duplex DNA by SARS coronavirus helicase nsP13. Nucleic Acids Res. 2010;38(21):7626-36.
169. Seybert A, van Dinten LC, Snijder EJ, Ziebuhr J. Biochemical characterization of the equine arteritis virus helicase suggests a close functional relationship between arterivirus and coronavirus helicases. J Virol. 2000;74(20):9586-93.
170. Zhang R, Li Y, Cowley TJ, Steinbrenner AD, Phillips JM, Yount BL, et al. The nsp1, nsp13, and M proteins contribute to the hepatotropism of murine coronavirus JHM.WU. J Virol. 2015;89(7):3598-609.
171. Minskaia E, Hertzig T, Gorbalenya AE, Campanacci V, Cambillau C, Canard B, et al. Discovery of an RNA virus 3'->5' exoribonuclease that is critically involved in coronavirus RNA synthesis. Proc Natl Acad Sci U S A. 2006;103(13):5108-13.
172. Bouvet M, Imbert I, Subissi L, Gluais L, Canard B, Decroly E. RNA 3'-end mismatch excision by the severe acute respiratory syndrome coronavirus nonstructural protein nsp10/nsp14 exoribonuclease complex. Proc Natl Acad Sci U S A. 2012;109(24):9372-7.
173. Eckerle LD, Lu X, Sperry SM, Choi L, Denison MR. High fidelity of murine hepatitis virus replication is decreased in nsp14 exoribonuclease mutants. J Virol. 2007;81(22):12135-44.
174. Eckerle LD, Becker MM, Halpin RA, Li K, Venter E, Lu X, et al. Infidelity of SARS-CoV Nsp14-exonuclease mutant virus replication is revealed by complete genome sequencing. PLoS Pathog. 2010;6(5):e1000896.
175. Case JB, Li Y, Elliott R, Lu X, Graepel KW, Sexton NR, et al. Murine Hepatitis Virus nsp14 Exoribonuclease Activity Is Required for Resistance to Innate Immunity. J Virol. 2018;92(1).
176. Zhou Y, Wu W, Xie L, Wang D, Ke Q, Hou Z, et al. Cellular RNA helicase DDX1 is involved in transmissible gastroenteritis virus nsp14-induced interferon-beta production. Front Immunol. 2017;8:940.
177. Chen Y, Cai H, Pan J, Xiang N, Tien P, Ahola T, et al. Functional screen reveals SARS coronavirus nonstructural protein nsp14 as a novel cap N7 methyltransferase. Proc Natl Acad Sci U S A. 2009;106(9):3484-9.
178. Jin X, Chen Y, Sun Y, Zeng C, Wang Y, Tao J, et al. Characterization of the guanine-N7 methyltransferase activity of coronavirus nsp14 on nucleotide GTP. Virus Res. 2013;176(1-2):45-52.
179. Snijder EJ, Horzinek MC, Spaan WJ. The coronaviruslike superfamily. Adv Exp Med Biol. 1993;342:235-44.
180. Cavanagh D, Horzinek MC. Genus Torovirus assigned to the Coronaviridae. Arch Virol. 1993;128(3-4):395-6.
181. Zirkel F, Roth H, Kurth A, Drosten C, Ziebuhr J, Junglen S. Identification and characterization of genetically divergent members of the newly established family Mesoniviridae. J Virol. 2013;87(11):6346-58.
182. Nga PT, Parquet Mdel C, Lauber C, Parida M, Nabeshima T, Yu F, et al. Discovery of the first insect nidovirus, a missing evolutionary link in the emergence of the largest RNA virus genomes. PLoS Pathog. 2011;7(9):e1002215.
183. Lauber C, Ziebuhr J, Junglen S, Drosten C, Zirkel F, Nga PT, et al. Mesoniviridae: a proposed new family in the order Nidovirales formed by a single species of mosquito-borne viruses. Arch Virol. 2012;157(8):1623-8.
184. Guarino LA, Bhardwaj K, Dong W, Sun J, Holzenburg A, Kao C. Mutational analysis of the SARS virus Nsp15 endoribonuclease: identification of residues affecting hexamer formation. J Mol Biol. 2005;353(5):1106-17.
185. Bhardwaj K, Guarino L, Kao CC. The severe acute respiratory syndrome coronavirus Nsp15 protein is an endoribonuclease that prefers manganese as a cofactor. J Virol. 2004;78(22):12218-24.
186. Deng X, Baker SC. An "Old" protein with a new story: Coronavirus endoribonuclease is important for evading host antiviral defenses. Virology. 2018;517:157-63.
187. Zhang L, Li L, Yan L, Ming Z, Jia Z, Lou Z, et al. Structural and Biochemical Characterization of Endoribonuclease Nsp15 Encoded by Middle East Respiratory Syndrome Coronavirus. J Virol. 2018;92(22).
188. Bhardwaj K, Palaninathan S, Alcantara JM, Yi LL, Guarino L, Sacchettini JC, et al. Structural and functional analyses of the severe acute respiratory syndrome coronavirus endoribonuclease Nsp15. J Biol Chem. 2008;283(6):3655-64.
189. Bhardwaj K, Sun J, Holzenburg A, Guarino LA, Kao CC. RNA recognition and cleavage by the SARS coronavirus endoribonuclease. J Mol Biol. 2006;361(2):243-56.
190. Kang H, Bhardwaj K, Li Y, Palaninathan S, Sacchettini J, Guarino L, et al. Biochemical and genetic analyses of murine hepatitis virus Nsp15 endoribonuclease. J Virol. 2007;81(24):13587-97.
191. Deng X, Hackbart M, Mettelman RC, O'Brien A, Mielech AM, Yi G, et al. Coronavirus nonstructural protein 15 mediates evasion of dsRNA sensors and limits apoptosis in macrophages. Proc Natl Acad Sci U S A. 2017;114(21):E4251-E60.
192. Decroly E, Ferron F, Lescar J, Canard B. Conventional and unconventional mechanisms for capping viral mRNA. Nat Rev Microbiol. 2011;10(1):51-65.
193. Decroly E, Imbert I, Coutard B, Bouvet M, Selisko B, Alvarez K, et al. Coronavirus nonstructural protein 16 is a cap-0 binding enzyme possessing (nucleoside-2'O)-methyltransferase activity. J Virol. 2008;82(16):8071-84.
194. Menachery VD, Yount BL, Jr., Josset L, Gralinski LE, Scobey T, Agnihothram S, et al. Attenuation and restoration of severe acute respiratory syndrome coronavirus mutant lacking 2'-o-methyltransferase activity. J Virol. 2014;88(8):4251-64.
195. Zhang W, Zheng XS, Agwanda B, Ommeh S, Zhao K, Lichoti J, et al. Serological evidence of MERS-CoV and HKU8-related CoV co-infection in Kenyan camels. Emerg Microbes Infect. 2019;8(1):1528-34.
196. Reusken CB, Haagmans BL, Muller MA, Gutierrez C, Godeke GJ, Meyer B, et al. Middle East respiratory syndrome coronavirus neutralising serum antibodies in dromedary camels: a comparative serological study. Lancet Infect Dis. 2013;13(10):859-66.
197. Ge XY, Li JL, Yang XL, Chmura AA, Zhu G, Epstein JH, et al. Isolation and characterization of a bat SARS-like

coronavirus that uses the ACE2 receptor. Nature. 2013;503(7477):535-8.
198. Li W, Shi Z, Yu M, Ren W, Smith C, Epstein JH, et al. Bats are natural reservoirs of SARS-like coronaviruses. Science. 2005;310(5748):676-9.
199. Hu B, Ge X, Wang LF, Shi Z. Bat origin of human coronaviruses. Virol J. 2015;12:221.
200. Banerjee A, Kulcsar K, Misra V, Frieman M, Mossman K. Bats and Coronaviruses. Viruses. 2019;11(1).
201. Brook CE, Boots M, Chandran K, Dobson AP, Drosten C, Graham AL, et al. Accelerated viral dynamics in bat cell lines, with implications for zoonotic emergence. Elife. 2020;9.
202. Banerjee A, Baker ML, Kulcsar K, Misra V, Plowright R, Mossman K. Novel insights into immune systems of bats. Front Immunol. 2020;11:26.
203. Zumla A, Chan JF, Azhar EI, Hui DS, Yuen KY. Coronaviruses - drug discovery and therapeutic options. Nat Rev Drug Discov. 2016;15(5):327-47.
204. Anand K, Ziebuhr J, Wadhwani P, Mesters JR, Hilgenfeld R. Coronavirus main proteinase (3CLpro) structure: basis for design of anti-SARS drugs. Science. 2003;300(5626):1763-7.
205. David A. Matthews, Amy K. Patick, Robert O. Baker, Mary A. Brothers, Peter S. Dragovich, Chris J. Hartmann, et al. In vitro antiviral activivty of human rhinovirus 3C protease inhibitors against the SARS coronavirus. In: Knobler S, Mahmoud A, Lemon S, Mack A, Sivitz L, Oberholtzer K, editors. Learning from SARS: Preparing for the Next Disease Outbreak: Workshop Summary. The National Academies Collection: Reports funded by National Institutes of Health. Washington (DC): National Academies Press (US); 2004.
206. Yang H, Xie W, Xue X, Yang K, Ma J, Liang W, et al. Design of wide-spectrum inhibitors targeting coronavirus main proteases. PLoS Biol. 2005;3(10):e324.
207. Ren Z, Yan L, Zhang N, Guo Y, Yang C, Lou Z, et al. The newly emerged SARS-like coronavirus HCoV-EMC also has an "Achilles' heel": current effective inhibitor targeting a 3C-like protease. Protein Cell. 2013;4(4):248-50.
208. de Wilde AH, Jochmans D, Posthuma CC, Zevenhoven-Dobbe JC, van Nieuwkoop S, Bestebroer TM, et al. Screening of an FDA-approved compound library identifies four small-molecule inhibitors of Middle East respiratory syndrome coronavirus replication in cell culture. Antimicrob Agents Chemother. 2014;58(8):4875-84.
209. Chu CM, Cheng VC, Hung IF, Wong MM, Chan KH, Chan KS, et al. Role of lopinavir/ritonavir in the treatment of SARS: initial virological and clinical findings. Thorax. 2004;59(3):252-6.
210. Chan KS, Lai ST, Chu CM, Tsui E, Tam CY, Wong MM, et al. Treatment of severe acute respiratory syndrome with lopinavir/ritonavir: a multicentre retrospective matched cohort study. Hong Kong Med J. 2003;9(6):399-406.
211. Chan JF, Yao Y, Yeung ML, Deng W, Bao L, Jia L, et al. Treatment With Lopinavir/Ritonavir or Interferon-beta1b Improves Outcome of MERS-CoV Infection in a Nonhuman Primate Model of Common Marmoset. J Infect Dis. 2015;212(12):1904-13.
212. Agnihothram S, Yount BL, Jr., Donaldson EF, Huynh J, Menachery VD, Gralinski LE, et al. A mouse model for Betacoronavirus subgroup 2c using a bat coronavirus strain HKU5 variant. mBio. 2014;5(2):e00047-14.
213. Zhang Q, Wang Y, Qi C, Shen L, Li J. Clinical trial analysis of 2019-nCoV therapy registered in China. J Med Virol. 2020 Feb 28. [Epub ahead of print]
214. Li G, De Clercq E. Therapeutic options for the 2019 novel coronavirus (2019-nCoV). Nat Rev Drug Discov. 2020;19(3):149-50.
215. Cao B, Wang Y, Wen D, Liu W, Wang J, Fan G, et al. A trial of lopinavir-ritonavir in adults hospitalized with severe Covid-19. N Engl J Med. 2020 Mar 18. [Epub ahead of print]
216. Lin MH, Moses DC, Hsieh CH, Cheng SC, Chen YH, Sun CY, et al. Disulfiram can inhibit MERS and SARS coronavirus papain-like proteases via different modes. Antiviral Res. 2018;150:155-63.
217. Lee H, Ren J, Pesavento RP, Ojeda I, Rice AJ, Lv H, et al. Identification and design of novel small molecule inhibitors against MERS-CoV papain-like protease via high-throughput screening and molecular modeling. Bioorg Med Chem. 2019;27(10):1981-9.
218. Cheng KW, Cheng SC, Chen WY, Lin MH, Chuang SJ, Cheng IH, et al. Thiopurine analogs and mycophenolic acid synergistically inhibit the papain-like protease of Middle East respiratory syndrome coronavirus. Antiviral Res. 2015;115:9-16.
219. Wang M, Cao R, Zhang L, Yang X, Liu J, Xu M, et al. Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro. Cell Res. 2020;30(3):269-71.
220. Adedeji AO, Singh K, Calcaterra NE, DeDiego ML, Enjuanes L, Weiss S, et al. Severe acute respiratory syndrome coronavirus replication inhibitor that interferes with the nucleic acid unwinding of the viral helicase. Antimicrob Agents Chemother. 2012;56(9):4718-28.
221. Yu MS, Lee J, Lee JM, Kim Y, Chin YW, Jee JG, et al. Identification of myricetin and scutellarein as novel chemical inhibitors of the SARS coronavirus helicase, nsP13. Bioorg Med Chem Lett. 2012;22(12):4049-54.
222. Tanner JA, Zheng BJ, Zhou J, Watt RM, Jiang JQ, Wong KL, et al. The adamantane-derived bananins are potent inhibitors of the helicase activities and replication of SARS coronavirus. Chem Biol. 2005;12(3):303-11.
223. Yang XH, Deng W, Tong Z, Liu YX, Zhang LF, Zhu H, et al. Mice transgenic for human angiotensin-converting enzyme 2 provide a model for SARS coronavirus infection. Comp Med. 2007;57(5):450-9.
224. Li K, Wohlford-Lenane CL, Channappanavar R, Park JE, Earnest JT, Bair TB, et al. Mouse-adapted MERS coronavirus causes lethal lung disease in human DPP4 knockin mice. Proc Natl Acad Sci U S A. 2017;114(15):E3119-E28.
225. Baric RS, Sims AC. Humanized mice develop coronavirus respiratory disease. Proc Natl Acad Sci U S A. 2005;102(23):8073-4.
226. Meszaros B, Erdos G, Dosztanyi Z. IUPred2A: context-dependent prediction of protein disorder as a function of redox state and protein binding. Nucleic Acids Res. 2018;46(W1):W329-W37.

REVIEW ARTICLE

Practical applications of atomic force microscopy in biomedicine

Nicola GALVANETTO [a,b*]

[a] International School for Advanced Studies (SISSA), via Bonomea 265, Trieste 34136, Italy.
[b] Department of Biochemistry and Department of Physics, University of Zurich, 8057 Zurich, Switzerland.

*Correspondence: nicola.galvanetto@sissa.it
https://doi.org/10.37175/stemedicine.v1i2.15

ABSTRACT

The last thirty years of progress of atomic force microscopy (AFM) applied to living matter is reviewed with a focus on potential uses in drug discovery or screening of patient samples. AFM-based technologies are still at Proof of Concept level - or below, however, they are particularly promising for i) live imaging of unlabeled membrane proteins and ii) nanomechanical screening of biological samples, e.g. cancer biopsies.

Keywords: Atomic force microscopy · Cell membranes · Membrane proteins · Unroofing · Applications

1. Background

The progress of science is usually a graded process, but once in a while it shows discontinuities due to some disruptive revolutions. In physics, there have been mostly conceptual revolutions: theories that allow to link experimental results that were previously uncorrelated; in modern biology, on the contrary, there have been mostly technological revolutions: techniques that enable (previously) unthinkable experiments.

The invention of the Atomic Force Microscope (AFM) (1) (**Figure 1**) is one of these revolutions.

2. AFM is an enabling technology still under development

2.1 AFM is a special kind of microscope

The word microscope derives from the Greek words mikros - “small” - and skopein - “to look”. If we took literally this definition, it wouldn't be right to call the AFM really a ‘microscope', because AFMs are metaphorically much closer to touch than to sight. Indeed, the basic working principle of the AFM is a tip that touches and moves on top of a surface (**Figure 1**).

The working principle is simple, but the necessary technology to bring it to sub-nanometric precision is all but not trivial. Broadly speaking, the goal is to raster scan the surface of the sample with a tiny probe (cantilever tip), to record the scanned information as X-Y-Z coordinates, and to reconstruct this map into a topographical image true to the original sample. A modern AFM (2) is constituted by a micrometric cantilever with a sharp tip on one side, and connected to a cantilever holder on the other (**Figure 1**). The cantilever holder is fixed to a piezo tube (a piezoelectric crystal fused in a tubular volume) necessary to move the cantilever tip in the X-Y-Z directions. The piezo tube is a key element because it uses the special property of piezo crystals to deform in a very accurate manner under high electric voltages, thus to move the cantilever with sub-nanometric precision. The forces to which the cantilever is subjected are constantly monitored through a laser beam that points to the cantilever and reflects into a photodetector (**Figure 1**): in this way, even the smallest deflections of the cantilever are detected by the photodetector as movements of the laser spot. Moreover, in the regime of small deflections, the cantilever behaves like a spring, therefore the AFM can be used not only to raster scan a surface, but also to measure forces applied to an extremely small area (i.e. the apex of the tip). The power of the AFM comes from the fine intercombination of these elements and from the miniaturization that started in the 1960s, on the wave of the silicon revolution.

To summarize, an AFM is the combination of three advanced technologies: piezoelectric crystals, microfabrication for the cantilever tip, laser and electronics for the control device. In the following, the advancements of microfabrication and electronics that represent the domains of major gain of AFM performance of the last few years are briefly reviewed.

Received: Nov 7, 2019; Accepted: Dec 29, 2019.

2.2 'Make it smaller and faster' (microfabrication & electronics)

The resolution of the AFM depends on the size of the probe that is used to scan the surface of interest. In order to follow nanometer features, tips with an apex with nanometer radius of curvature are essential. Moreover, if the sample presents densely packed hills and holes, the tip necessitates to enter the holes and follow every detail to generate a faithful topographical representation of the sample. For these reasons, scientists developed methods to grow on top of the cantilever - or on top of the cantilever tip itself (**Figure 2a**) - an additional ultra-sharp tip made of carbon atoms. This additional tip can be several µm high and can have a radius of curvature of 1-2 nm (3), in order to follow the finest details of the sample.

But it is not only a matter of tips, indeed the quality of the topographical reconstruction depends also on the size of the cantilever. As a general principle, the smaller the better, because a small lever is softer (less invasive) and lighter (less inertia). The challenges were a) to manufacture small and precise levers and b) to fit the AFM accordingly. The advancements in microelectronics allowed to have the right technology to build small probes, and these small probes were also modified (sculptured) with Focused Ion Beams to make them even softer and lighter (**Figure 2b**) (4). Once small probes were made, a standard AFM became inadequate, in particular in the laser-photodetector coupling. In order to focus the laser beam onto these new small probes, the AFMs were customized with an optical objective necessary to focus the light onto the smaller reflection area of the new levers.

Now that all the ingredients for better AFMs were ready, the last step was to make them faster. An AFM image used to require several minutes to be recorded. Therefore, AFMs were equipped with special ultra-fast electronics giving birth to what is now called High-speed AFM (HS-AFM), which can be considered a second revolution in the field (5).

2.3 AFM is used to generate images or to probe the nanomechanics

There are two kinds of output that an AFM can generate: images or force curves.

AFM can be used to scan a surface, thus to record the surface topology of a sample (**Figure 4a**). These images can be collected in the so called 'contact mode' i.e. when the tip literally touches the surface moving laterally like in **Figure 1**; but the images can be generated also in the so called 'non-contact' mode or in 'intermittent contact' mode. In these cases, the tip, in addition to a lateral movement, is forced to oscillate in the vertical direction. These last modes of operation are particularly advantageous with soft samples, to prevent their damage.

But AFM is also a force probe, therefore the tip can also be used in one single location to probe the nanomechanical properties of the material. The tip is forced to move downwards and then upwards, while collecting the deflection of the cantilever. In this way it is possible to collect indentation curves or unfolding curves of a protein (**Figure 6c**).

In Section 3 I will present the applications in biomedicine of the AFM image generation, while in Section 4 and 5 I will discuss the use of AFM as a force probe of biological matter.

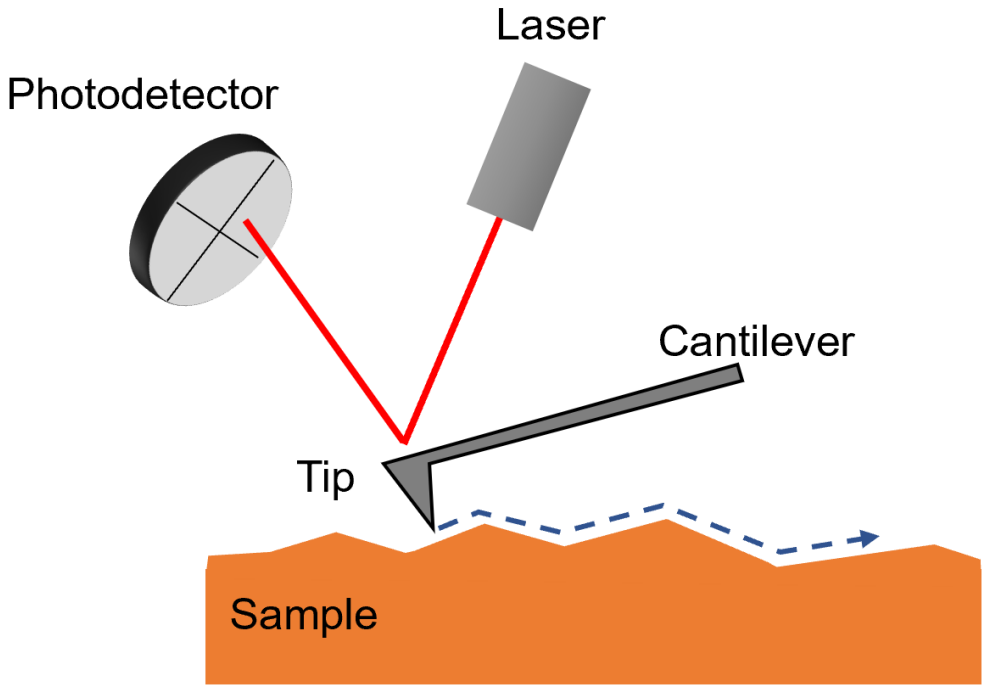


Figure 1. Sketch of an AFM probe moving on top of the imaged surface and basic scheme of an AFM device.

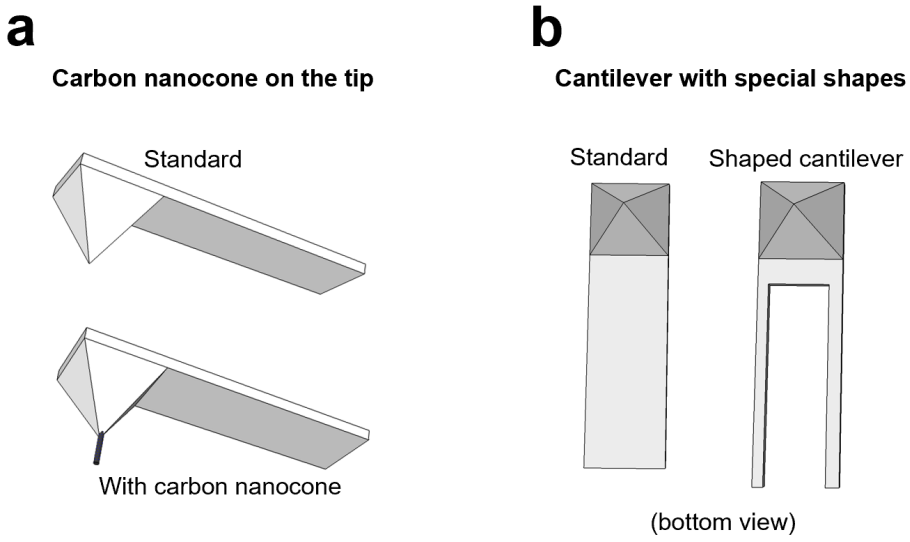


Figure 2. a: ultra-sharp tip (with carbon nanotube grown on top of a silicon tip; like in ref. (5)). **b: bottom views of modified cantilevers for enhanced temporal resolution** (like in ref. (4)).

3. AFM is a microscope for unlabeled molecules

Biology and modern medicine make extensive use of imaging techniques, and in particular of fluorescence microscopy (6,7) (**Figure 3c**). The general idea underlying fluorescence microscopy is to obtain signals from the specific biomolecule of interest, and to associate it with its spatial distribution within the sample, sometimes also for quantitative assessments. The biomolecule of interest is in almost all cases not fluorescent - i.e. invisible - therefore it needs to be labeled with another fluorescent molecule to be detected. In any case, with the exception of Forster Resonance Energy Transfer, no structural information of the biomolecules can be obtained with fluorescence microscopy.

To obtain structural information, structural biology and pharmacology use X-ray crystallography, nuclear magnetic resonance, and more recently cryogenic electron microscopy (cryo-EM) to investigate the internal structure of the biomolecule of interest (8,9) (**Figure 3a**). With these techniques it is possible to reach sub-nanometer resolution of the 3D arrangements of the amino acids in the case of proteins, which is very relevant in drug design. To gather this structural information there are some experimental steps which include the purification of the biomolecule, or its freezing, that necessarily does not allow to keep the molecules in a physiological (not even almost-physiological) environment.

There are situations, or scientific prerequisites, where either the label cannot be applied, or the project necessitate structural and dynamical information of the biomolecule in an almost-physiological environment. For these kinds of situations, the AFM is the ideal microscopy technique because it sits exactly in between fluorescence microscopy and crystallography/cryo-EM (**Figure 3b**).

3.1 Purified proteins reconstituted in membranes

Despite of being the target of about 50% of modern drugs (10), structural studies have been particularly complicated, with membrane proteins that are not prone to crystallization because they have a biphasic surface (hydrophilic at the poles and hydrophobic at the equator to accommodate the membrane), therefore it is difficult to measure the structural effect of a ligand.

In the late 1980s, the newborn AFM that was invented for applications on semiconductors or insulators also entered the biology departments, and the first pioneering studies on membrane proteins began. The first membrane protein for which it was developed a protocol for AFM imaging was bacteriorhodopsin (11), a membrane protein that form 2D crystals called purple membranes. Some years later it was possible to reach sub-molecular resolution and compare the AFM topographs with the atomic model derived from electron microscopy (12). The real AFM advantages over electron microscopy started to be evident at this point: indeed, AFM could be operated in liquid, i.e. emulating the physiological conditions of the membrane proteins environment. For instance, Muller et al. (12) showed that by changing the pH of the imaging solution, the purple membrane displayed conformational changes in its structure. Later studies demonstrated also that the application of forces induced changes in the bacteriorhodopsin structure (13), opening the way for “multidimensional" molecular imaging. In this regard, a more recent study of bacteriorhodopsin mechanical response (14) reached such a high level of precision that allowed the authors to determine the stiffest and the softest domains within a single protein. But the class of membrane proteins that is the most natural subject of investigation for a technique involving forces are the mechanosensitive proteins, e.g. PIEZO channels. In

mechanosensitive proteins, the application of forces is thought to be responsible for key structural changes (like channel gating). This is exactly what a recent study searched for in a combination of cryo-EM and AFM experiments, where the authors were able to propose a mechanism that translates the forces applied to the membrane to channel gating (15).

What was lacking in these experiments was a dynamical point of view in the time-scale of seconds. The slow operation mode of conventional AFMs was boosted starting from the year 2010 when Toshio Ando and his group introduced the technical improvements described in section 2.2 and built the first HS-AFM, which could acquire several frames per second. The fast frame rate enabled the visualization of fast molecular mechanisms, for instance, the first direct imaging of the movements of myosin (16) and the diffusion of membrane proteins in membranes (17).

I believe that HS-AFM will become increasingly prominent for practical applications in drug discovery because it allows to directly probe the structural effects of ligands on the membrane protein of interest with a relatively simple machine (18) (**Figure 4a and b**). In the next section I will describe how AFM has been applied to real cell membranes.

3.2 Imaging native cell membranes (from electron microscopy to AFM)

The question “how does a real cell membrane look like?" started to be answered only from the 1970s when the first rudimentary technique to break the cells and to create membrane-only samples was developed (19). Given the length-scale of the membrane features, the only microscopy technique available for their investigation was electron microscopy. This microscopy technique was extensively adopted in the following years, in particular thanks to a radical change that made the sample preparation more reliable: the use of sonic waves to break the cells to obtain the so called “unroofed" membranes (20). From these investigations we have learned how the membrane skeleton is arranged by looking directly into it from the cytoplasmic side of the membrane. For instance, it was possible to shed light on how the actin filament arrangements partition the plasma membrane, decreasing the lateral diffusion of lipids and proteins (21).

The versatility of AFM discussed also in the previous section comprises a relatively simpler instrument compared to the electron microscope and, more importantly, the possibility to work at room temperature with the sample under the physiological medium. Despite of these theoretical advantages on paper, the AFM was applied to native membranes relatively late, and starting with special compartments of the cell constituted just by membranes. An example of these applications is the rod outer segment disc, key for vision and rich in Rhodopsin (22). AFM was able to resolve the macromolecular structures of Rhodopsin in discs that form nano-domains with an average size of 500 nm^2 (23).

Only more recently, a group in Japan revisited the unroofing technique developed for electron microscopy (24)

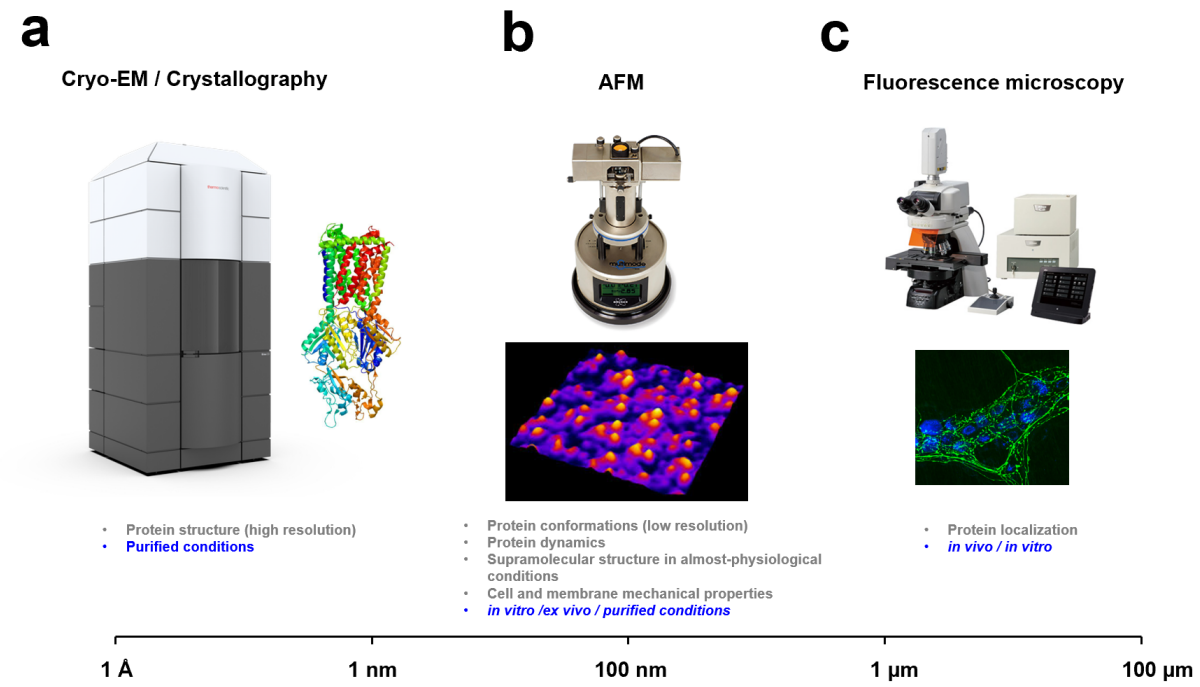


Figure 3. Comparison of imaging techniques for biomolecules and their domains of application. a: for structural, high-resolution, imaging. b: for intermediate molecular imaging with dynamics (AFM image obtained from JPK). c: for cellular imaging of biomolecules localization.

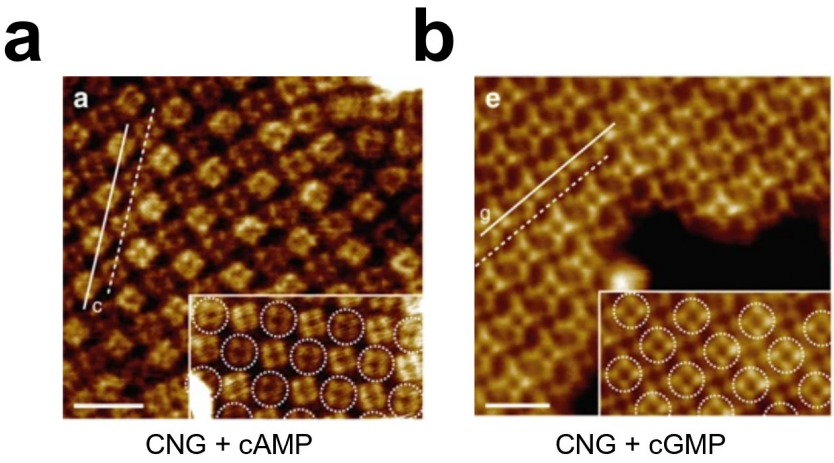


Figure 4. a and b: HS-AFM of conformational changes of an homolog of CNG channel in presence of cAMP and cGMP (adapted from (18)).

and imaged these samples after fixation with AFM, obtaining molecular resolution at room temperature (**Figure 5a and b**).

After having tested many of these methods, one of the major disadvantages of using AFM with native cell membranes is the fact that, the position of the isolated membrane itself is difficult to be found and requires time consuming scanning of the sample that also affects the quality of the tip. For this reason, inspired by an earlier study (25), I have developed an unroofing method that breaks only one single cell (or a few) on the very corner of a triangular coverslip (26) (**Figure 5c and d**). In this way, the identification of the position of the membrane is straightforward, and the membrane seems to be isolated without the membrane skeleton that was observed with other techniques. The very last goal in this case would be to image single membrane proteins with sub-molecular resolution in their complex macromolecular organization, potentially to recognize the signatures of diseases that affect directly the cell membrane.

4. AFM as a force probe for cells and cell membranes

In the previous sections I discussed the imaging capabilities of the AFM, but we cannot forget that the AFM is also an instrument able to measure forces at the micro- and nano-scale level. In the next sections I will present the major results of the past two decades obtained with the application of AFM-based force spectroscopy to biological materials.

4.1 Single-cell force spectroscopy

Cells are soft objects that range from 1 to 50 microns in size. As opposed to inanimate objects, cells are active entities constituted by a complex internal structure, they can duplicate and mix with other cells forming tissues, and finally organisms. The biochemical signals that the cells use to communicate and function have been the subject of molecular biology investigations which have made an enormous progress in the last 70 years. Different cell types are characterized by different biochemical signatures, and a question that one may ask is: are different cells characterized by different mechanical properties?

AFM can answer this question as the AFM tip can be used to indent a cell while recording the cantilever deflection. The recorded curve is informative of the cell stiffness and from the same curve it is possible to extrapolate conventional mechanical properties like the Young modulus (27). There is a vast scientific literature on this topic which focused very early on the attempt to find biomedical applications, in particular for cancer cell detection (28-30). The general conclusion of these studies is that cancer cells [lung, breast, pancreas (31), ovarian (29)] are softer and with a narrower standard deviation of the Young modulus compared to non-malignant cells of the same body cavity. Being softer may help cancer cells to be more dynamic and therefore successful in infiltrating tissues, but a definitive explanation of the problem is still to come.

More recently, people developed also some alternative operating mode to expand the range of mechanical properties that AFM can investigate. Yango el al. (32), for instance, introduced cycles of indentation and relaxation to measure both the Young modulus and the internal viscosity of the cell. An interesting advancement is the one introduced by Rigato et al. (33), where they measured the cell viscoelastic properties at different time scales, stimulating the cells from 1 Hz to 100 kHz. This is important because the cell is an active object whose constituents (e.g. cytoskeleton filaments) can rearrange at different time scales, therefore defining the ‘mechanical fingerprint' can help to sort cells more accurately.

Being able to uniquely identify cancerous cells (or specimens) just by probing their mechanics, without the use of biochemical markers, seems to be a challenge that may be completed in relatively few years.

4.2 Nanomechanics of lipid bilayers: from synthetic to native membranes

The mechanical properties of the lipid bilayer - i.e. the component that separates the cell interior from the external environment - were not accessible with the standard biochemical techniques available until the invention of the AFM. AFM opened the way for the direct investigation of the strength of molecular interactions among lipids in supported membranes. In these experiments, synthetic lipids (e.g. dipalmitoylphosphatidylcholine and dioleoylphosphatidylcholine) are mixed in solution and deposited on atomically flat surfaces (e.g. cleaved mica), then they are imaged with AFM to confirm their correct assembly, and finally indented with the AFM tip to probe the nanomechanical properties of the membrane. The prototypical indentation curve recorded by the AFM shows an initial elastic behavior of the membrane (deformation of the bilayer), until the compression force of the AFM tip reaches a critical value that causes the breakage of the bilayer. After the breakage, the tip is retracted while the membrane self-assembles back, closing the hole.

To my knowledge, the first experiments on lipids nanomechanics were performed in 1999 by Dufrene et al. (34): they demonstrated the versatility of AFM to probe the chemical and physical properties of lipid membranes. After these pioneering studies, many labs around the world started to apply AFM to membranes, characterizing their behavior in many different conditions of physiological relevance. For instance, it was shown that the stiffness of the membrane increases by reducing its temperature (35) or by increasing the percentage of cholesterol (36), with important consequences for our understanding of the dynamic of processes at cellular and sub-cellular levels. It is surprising how, from relatively crude indentations and enough statistics, it is possible to extrapolate very detailed physical quantities of the bilayer like the line tension and the spreading pressure, quantities that are essential in coarse grain simulations of the cell.

I would like to mention two noteworthy recent experimental advancements in the field of membrane biophysics. Goncalves et al. (37), instead of studying supported membranes, developed a substrate with sub-micrometric holes where the membrane could remain free-standing, emulating a real cell surface. In this way, they were able to precisely measure the breaking point of a membrane that could bend under the AFM tip push. While Al-Rekabi et al. (38), similarly to Rigato et al. (33) as described in the previous section, introduced a multifrequency AFM mode that allows to capture the viscoelastic properties of the membrane and demonstrates how cholesterol concentration is a key ingredient in changing membrane properties from viscous to elastic.

The goal of the recently developed single-cell unroofing (26) is to transfer the application of the powerful methods described above (and benchmarked with synthetic membranes) to native cellular membranes, in order to gather information of the real nanomechanics generated by the native complexity. Indeed, with single-cell unroofing it is possible to tear the cell membrane of cells (**Figure 5c**) and directly probe their indentation force, deriving elasticity and viscosity of cells of patients.

5. AFM as a force probe for single proteins

Measuring the mechanical properties of cells or cell membranes is definitely a technological leap, but it does not represent a conceptual breakthrough, in the sense that these results could be somehow expected. What really represents a practical revolution for an unsolved problem in biology is the possibility to record the unfolding of single proteins - one at a time - with AFM-based single-molecule force spectroscopy (SMFS).

In fact, the physico-chemical process that guides a linear sequence of amino acids into the three-dimensional precise structures that (almost) every protein shows may

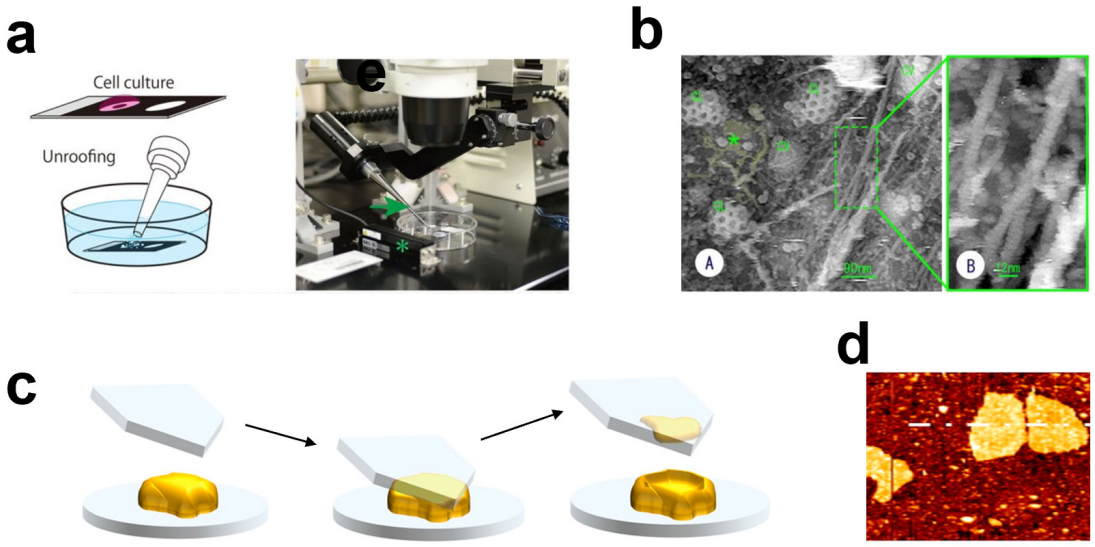


Figure 5. a & b: unroofing preparation through sonic waves and AFM imaging of actin filaments and clathrin (24). c & d: single cell unroofing through the gentle squeezing of a single cell and the resulting isolated membrane like (26).

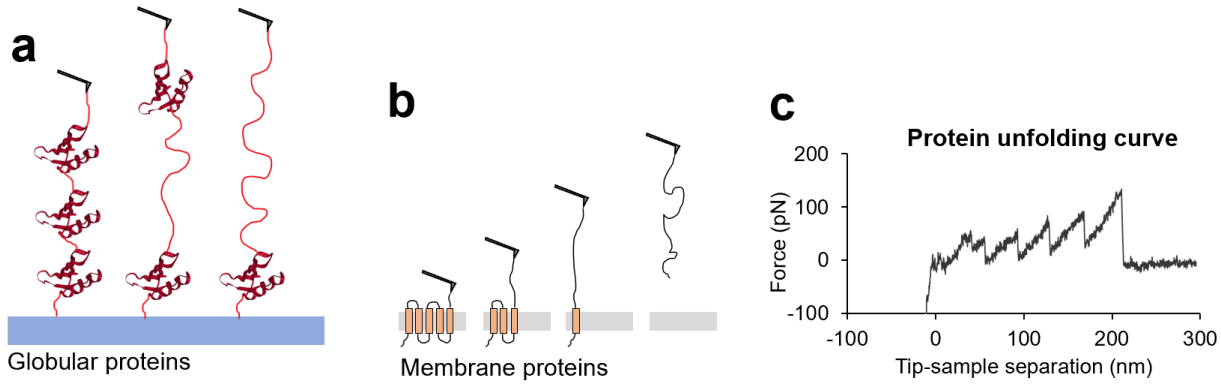


Figure 6. a: unfolding of a tandem of globular protein. b: unfolding of a trans-membrane protein. c: prototypical unfolding curve od a protein.

be taken for granted, but it has been a puzzle for physicists since the 1950s (39). The number of variables that should be taken into account to simulate with atomic resolution the folding of a protein is simply intractable for modern computational power, and therefore, in the past years many computational “shortcuts" have been invented to tackle the problem in an approximate manner (40). On the other hand, from the experimental point of view, the only way to study protein folding and unfolding was by chemical denaturation (41), thus at the ensemble level and not at the single-molecule level. There was no certain mechanistic understanding of how the dynamics of protein folding should work.

This is the context where SMFS, starting from the late 1990s, revolutionized the way people study protein folding. The experimental setup consists in an AFM cantilever tip that, instead of being used for imaging as discussed in section 3.1, is approached and retracted form a flat sample onto which many copies of the same protein of interest are deposited. The AFM tip pushes the sample and binds one of the terminus ends of the underlying protein. Then the tip is retracted while recording the force experienced by the cantilever, so that it is possible to form a force vs distance spectrum (F-D curve). If the other terminus of the protein is bounded to the sample surface, the tip will literally stretch the protein, forcing denaturation. The prototypical unfolding curve of a sequence of immunoglobulin domains was observed for the first time by Rief et al. in 1997 (42) (**Figure 6a**) showing the so called saw-tooth like shape, i.e. a sequence of rising phases followed by vertical jumps to the baseline at zero-force.

How can we describe this behavior? The rising phases can be fitted with the worm-like chain model, i.e. a model developed in polymer physics that describes the stretching of an ideal semi-flexible chain (43). The sudden jump to zero-force is instead representative of the abrupt unfolding of a complete protein domain, which is unexpected. The unfolding of the protein does not happen in a continuous way, but rather in an all-or-none fashion. This two-state behavior (folded vs unfolded) is typical for the majority of globular proteins investigated so far, e.g. GFP, Xylanase, Fibronectin and many others (44). Moreover, these globular proteins - when unfolded - tends to fold back into their original 3D structure even *in vitro* (42).

Membrane proteins, and in particular trans-membrane proteins, are topologically very different from globular proteins because, instead of being fully immersed in the cytosol, they perform their function embedded in the cell membrane. They are also more difficult to be studied since they are laborious to be purified and they don't crystallize well (it is often necessary to study homologous proteins that are easier to be handled). The fact that they operate across the membrane introduces an additional complex ingredient to the problem of folding, indeed it was discovered that trans-membrane proteins need the help of other proteins (translocon) to be correctly folded into the membrane (45).

SMFS has what it takes to be very powerful when applied to membrane proteins, possibly being informative also of membrane protein structure. The first experimental observation dates back to the year 2000 when Oesterhelt et al. (46) used AFM-based SMFS to unfold bacteriorhodopsin. The unfolding curve of bacteriorhodopsin (see **Figure 6b and c**) resembles the unfolding of a tandem of globular proteins (**Figure 6a**) but it has a slightly different interpretation. As opposed to what happens to the majority of globular proteins that unfold in an all-or-none fashion, in bacteriorhodopsin the different domains of the same molecule unfold separately and consecutively, one after the other. A general empirical observation is that these domains correspond to pairs of alpha helices (or beta hairpin). This observation is actually quite general, since further studies of other membrane proteins tend to support this empirical “rule", also in beta-barrel membrane proteins. E.g. aquaporin (47), sodium-proton antiporter (48), FhuA (49), etc. The unfolding behavior of membrane proteins is somehow reminiscent of the mechanistic intuition of how these proteins should exit the membrane when pulled.

However, thanks to a recent improvement in resolution of the AFM probes, it has been shown that the single and

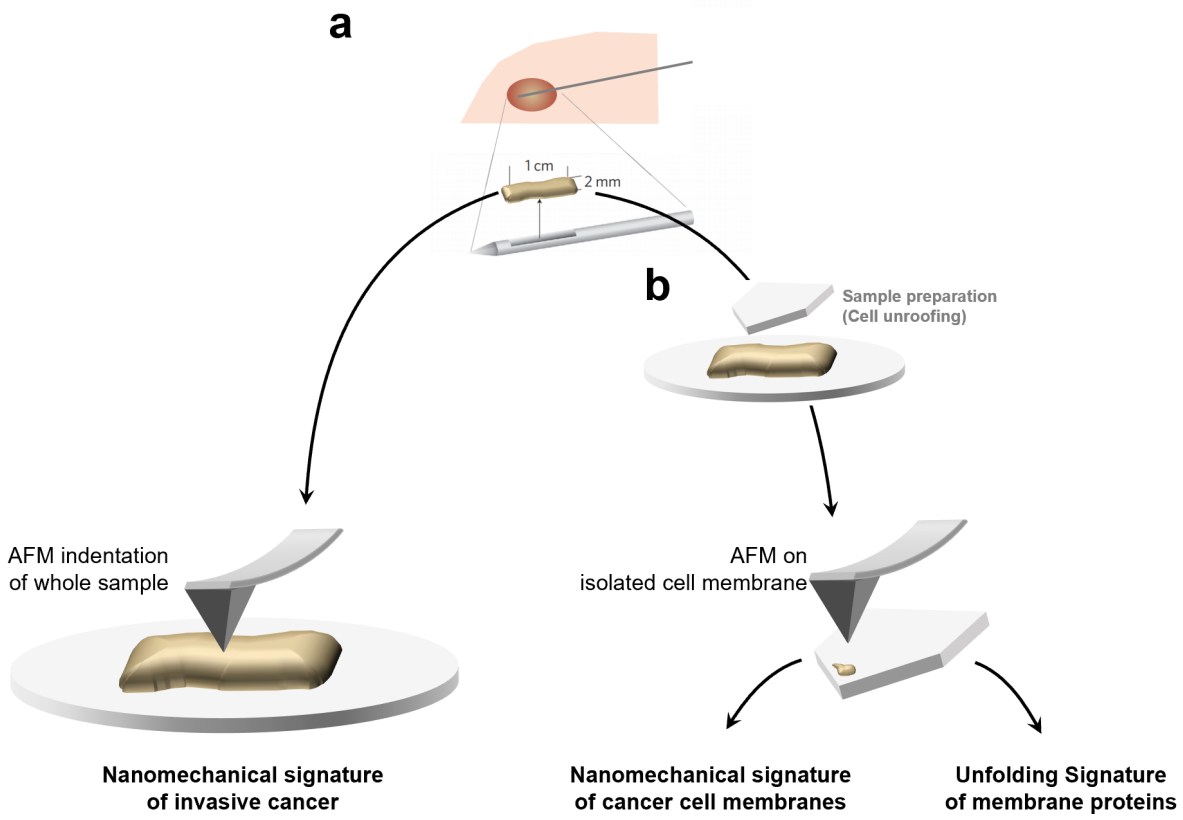


Figure 7. a: schematic of a biopsy of a patient. b: sample preparation: cell membrane isolation from biopsy cells. AFM force measurements performed on the biopsy or on the cell membranes reveal signatures that are characteristic of cancerous samples.

separated peaks of unfolding that was reported in the past, that correspond to the unfolding of pairs of alpha helices, are actually decorated by a rich series of minor unfolding events, indicating that the complexity of the process is still highly unresolved (50).

But with SMFS it is possible to test not only the detailed physics of folding, but also to obtain biological insights on the structure-function of specific proteins. An example is the work by Serdiuk et al., where they directly observed the action of chaperon proteins acting on membrane proteins that were previously unfolded, showing that without the help of these chaperons, a membrane protein cannot autonomously fold back into the membrane (51).

In the previous paragraphs I described how SMFS can be used to explore the problem of protein folding, but I strongly believe that the major applications of SMFS are to problems at the intersection with biomedicine. SMFS can be exploited to investigate protein behaviors where other techniques cannot be applied. I will now report some applications in this direction. One way to use SMFS to tackle a biomedical problem is to study the differences of a membrane protein and a mutated one that may be involved in a disease. Kawamura et al. (52) took a similar direction studying the differences between rhodopsin and opsin, which is the same protein not bounded to the chromophore that enables vision. In this case they detected slightly differences in the unfolding behavior, showing that SMFS can resolve these changes. Another approach is to study the mechanical differences induced by the binding of a ligand to a membrane protein, and how it affects the structure (53).

What seems a highly unexplored area of research is the one touched by Otten et al. (44), i.e. trying to change the scale and the throughput of the SMFS experiments. They developed an interesting platform that can express proteins *in vitro*, and then screen them in order to create a SMFS resource to test the most various hypothesis related to protein mechanics or folding, for instance malfunction due to misfolding.

6 Future perspectives

Much of the research presented in this review deals with new knowledge generated by AFM applied to biological samples. This knowledge has direct implications in real life problems.

It is estimated that the amount of funding that has been spent in basic and applied research involving bio-AFM is in the order of some billion dollars. Given the tremendous collaborative efforts of the last 30 years, it is reasonable to believe that a certain amount of organizations will eventually succeed in commercializing an AFM-based device, sold to hospitals. To my knowledge, the practical implications that nowadays are at the proof-of-concept level, and that could help the diagnosis of diseases are the

ones depicted in **Figure 7**:

i. direct indentation measurements of biopsies for automated tissue diagnostics (30) (technology readiness level 7-8, e.g. *ARDITIS AG (Swiss company)*);

ii. indentation measurements of the cell membranes isolated from biopsies (26) (technology readiness level 3-4);

iii. single-molecule force spectroscopy the native cell membranes isolated from biopsies (54) (technology readiness level 3-4).

The direct indentation of biopsies and the correlation of these data with the state of the patient is with no doubt the technology closest to a commercial device. This tool seems to help the physicians to identify the malignancy of different cancer biopsies from patients. There are clear substantial advancements also in hybrid systems of AFM plus Raman spectroscopy (55). In these cases, the benefit manifests in terms of chemical recognition of the material with sub-micrometer resolution; but these techniques are at the moment still difficult to be operated and still unmatured, but with high potential.

It is improbable that only one of these techniques will revolutionize the diagnostics by its own, but it is most likely that the intercombination of the mechanical signatures of each of these approaches, together with parallel diagnostic tools, will substantially help the decisions of physicians, even in a near future.

Conflict of interest

The authors declare that they have no conflict of interest.

References

1. Binnig G, Quate CF, Gerber Ch. Atomic force microscope. Phys Rev Lett. 1986 Mar 3;56(9):930–3.
2. McClelland GM, Erlandsson R, Chiang S. Atomic force microscopy: general principles and a new implementation. In: Thompson D.O., Chimenti D.E. (eds) Review of Progress in Quantitative Nondestructive Evaluation. 1987, vol 6 A. Springer, Boston, MA.
3. Chen I-C, Chen L-H, Ye X-R, Daraio C, Jin S, Orme CA, et al. Extremely sharp carbon nanocone probes for atomic force microscopy imaging. Appl Phys Lett. 2006 Apr 10;88(15):153102.
4. Edwards DT, Faulk JK, LeBlanc M-A, Perkins TT. Force spectroscopy with 9-μs resolution and sub-pN stability by tailoring AFM cantilever geometry. Biophys J. 2017 Dec 19;113(12):2595–600.
5. Ando T, Uchihashi T, Kodera N. High-speed AFM and applications to biomolecular systems. Annu Rev Biophys. 2013;42(1):393–414.
6. Kobayashi H, Ogawa M, Alford R, Choyke PL, Urano Y. New strategies for fluorescent probe design in medical diagnostic imaging. Chem Rev. 2010 May 12;110(5):2620–40.
7. Yguerabide J, Yguerabide EE. Fluorescence spectroscopy in biological and medical research. Int J Radiat Appl Instrum C Radiat Phys Chem. 1988 Jan 1;32(3):457–64.
8. Liao M, Cao E, Julius D, Cheng Y. Single particle electron cryo-microscopy of a mammalian ion channel. Curr Opin Struct Biol. 2014 Aug 1;27:1–7.
9. Williams DH, Mitchell T. Latest developments in crystallography and structure-based design of protein kinase inhibitors as drug candidates. Curr Opin Pharmacol. 2002 Oct 1;2(5):567–73.
10. Yin H, Flynn AD. Drugging membrane protein interactions. Annu Rev Biomed Eng. 2016;18(1):51–76.
11. Worcester DL, Miller RG, Bryant PJ. Atomic force microscopy of purple membranes. J Microsc.1988;152(3):817–21.
12. Müller DJ, Schabert FA, Büldt G, Engel A. Imaging purple membranes in aqueous solutions at sub-nanometer resolution by atomic force microscopy. Biophys J.1995 May 1;68(5):1681–6.
13. Müller DJ, Büldt G, Engel A. Force-induced conformational change of bacteriorhodopsin. J Mol Biol.1995 Jan 1;249(2):239–43.
14. Rico F, Su C, Scheuring S. Mechanical mapping of single membrane proteins at submolecular resolution. Nano Lett. 2011 Sep 14;11(9):3983–6.
15. Lin Y-C, Guo YR, Miyagi A, Levring J, MacKinnon R, Scheuring S. Force-induced conformational changes in PIEZO1. Nature. 2019 Sep;573(7773):230–4.
16. Kodera N, Yamamoto D, Ishikawa R, Ando T. Video imaging of walking myosin V by high-speed atomic force microscopy. Nature. 2010 Nov;468(7320):72–6.
17. Casuso I, Khao J, Chami M, Paul-Gilloteaux P, Husain M, Duneau J-P, et al. Characterization of the motion of membrane proteins using high-speed atomic force microscopy. Nat Nanotechnol. 2012 Aug;7(8):525–9.
18. Marchesi A, Gao X, Adaixo R, Rheinberger J, Stahlberg H, Nimigean C, et al. An iris diaphragm mechanism to gate a cyclic nucleotide-gated ion channel. Nat Commun. 2018 Sep 28;9(1):1–11.
19. Clarke M, Schatten G, Mazia D, Spudich JA. Visualization of actin fibers associated with the cell membrane in amoebae of Dictyostelium discoideum. PNAS. 1975 May 1;72(5):1758–62.
20. Heuser J. The Production of 'cell cortices' for light and electron microscopy. Traffic. 2000 Jul 1;1(7):545–52.
21. Morone N, Fujiwara T, Murase K, Kasai RS, Ike H, Yuasa S, et al. Three-dimensional reconstruction of the membrane skeleton at the plasma membrane interface by electron tomography. J Cell Biol. 2006 Sep 11;174(6):851–62.
22. Tanuj Sapra K, Park PS-H, Filipek S, Engel A, Müller DJ, Palczewski K. Detecting molecular interactions that stabilize native bovine rhodopsin. J Mol Biol. 2006 Apr;358(1):255–69.
23. Whited AM, Park PS-H. Nanodomain organization of rhodopsin in native human and murine rod outer segment disc membranes. Biochim Biophys Acta Biomembr. 2015 Jan 1;1848(1, Part A):26–34.
24. Usukura E, Narita A, Yagi A, Ito S, Usukura J. An unroofing method to observe the cytoskeleton directly at molecular resolution using atomic force microscopy. Sci Rep. 2016 Jun 7;6:srep27472.
25. Nermut MV, Williams LD. Freeze-fracturing of monolayers (capillary layers) of cells, membranes and viruses: some technical considerations. J Microsc. 1977 Jul 1;110(2):121–32.
26. Galvanetto N. Single-cell unroofing: probing topology and nanomechanics of native membranes. Biochim Biophys Acta Biomembr. 2018 Dec;1860(12):2532–8.
27. Faria EC, Ma N, Gazi E, Gardner P, Brown M, Clarke NW, et al. Measurement of elastic properties of prostate cancer cells using AFM. Analyst. 2008 Oct 20;133(11):1498–500.
28. Lulevich V, Zink T, Chen H-Y, Liu F-T, Liu G. Cell mechanics using atomic force microscopy-based single-cell compression. Langmuir. 2006 Sep 1;22(19):8151–5.
29. Xu W, Mezencev R, Kim B, Wang L, McDonald J, Sulchek T. Cell stiffness is a biomarker of the metastatic potential of ovarian cancer cells. PLoS One. 2012 Oct;7(10):e46609.
30. Plodinec M, Loparic M, Monnier CA, Obermann EC, Zanetti-Dallenbach R, Oertle P, et al. The nanomechanical signature of breast cancer. Nat Nanotechnol. 2012 Nov;7(11):757–65.
31. Cross SE, Jin Y-S, Rao J, Gimzewski JK. Nanomechanical analysis of cells from cancer patients. Nat Nanotechnol. 2007 Dec;2(12):780–3.

32. Yango A, Schäpe J, Rianna C, Doschke H, Radmacher M. Measuring the viscoelastic creep of soft samples by step response AFM. Soft Matter. 2016 Oct 12;12(40):8297–306.
33. Rigato A, Miyagi A, Scheuring S, Rico F. High-frequency microrheology reveals cytoskeleton dynamics in living cells. Nat Phys. 2017 Aug;13(8):771–5.
34. Dufrêne YF, Boland T, Schneider JW, Barger WR, Lee GU. Characterization of the physical properties of model biomembranes at the nanometer scale with the atomic force microscope. Faraday Discuss. 1999;111(0):79–94.
35. Leonenko ZV, Finot E, Ma H, Dahms TES, Cramb DT. Investigation of temperature-induced phase transitions in DOPC and DPPC phospholipid bilayers using temperature-controlled scanning force microscopy. Biophysical Journal. 2004 Jun 1;86(6):3783–93.
36. Sullan RMA, Li JK, Hao C, Walker GC, Zou S. Cholesterol-Dependent Nanomechanical Stability of Phase-Segregated Multicomponent Lipid Bilayers. Biophys J. 2010 Jul 21;99(2):507–16.
37. Gonçalves RP, Agnus G, Sens P, Houssin C, Bartenlian B, Scheuring S. Two-chamber AFM: probing membrane proteins separating two aqueous compartments. Nat Methods. 2006 Oct 22;3(12):nmeth965.
38. Al-Rekabi Z, Contera S. Multifrequency AFM reveals lipid membrane mechanical properties and the effect of cholesterol in modulating viscoelasticity. PNAS. 2018 Mar 13;115(11):2658–63.
39. Anfinsen CB, Redfield RR, Choate WL, Page J, Carroll WR. Studies on the gross structure, cross-linkages, and terminal sequences in ribonuclease. J Biol Chem. 1954;207(1):201–10.
40. Laio A, Parrinello M. Escaping free-energy minima. PNAS. 2002 Oct 1;99(20):12562–6.
41. Fong S, Hamill SJ, Proctor M, Freund SMV, Benian GM, Chothia C, et al. Structure and stability of an immunoglobulin superfamily domain from twitchin, a muscle protein of the nematode *Caenorhabditis elegans*. J Mol Biol. 1996 Dec 6;264(3):624–39.
42. Rief M, Gautel M, Oesterhelt F, Fernandez JM, Gaub HE. Reversible unfolding of individual titin immunoglobulin domains by AFM. Science. 1997;276(5315):1109–1112.
43. Doi M, Edwards SF. The theory of polymer dynamics. Clarendon Press; 1988. 420 p.
44. Otten M, Ott W, Jobst MA, Milles LF, Verdorfer T, Pippig DA, et al. From genes to protein mechanics on a chip. Nat Methods. 2014 Nov;11(11):1127–30.
45. Johnson AE, van Waes MA. The translocon: a dynamic gateway at the ER membrane. Annu Rev Cell Dev Biol. 1999;15:799-842.
46. Oesterhelt F, Oesterhelt D, Pfeiffer M, Engel A, Gaub HE, Müller DJ. Unfolding pathways of individual bacteriorhodopsins. Science. 2000 Apr 7;288(5463):143–6.
47. Möller C, Fotiadis D, Suda K, Engel A, Kessler M, Müller DJ. Determining molecular forces that stabilize human aquaporin-1. J Struct Biol. 2003 Jun 1;142(3):369–78.
48. Kedrov A, Ziegler C, Janovjak H, Kühlbrandt W, Müller DJ. Controlled unfolding and refolding of a single sodium-proton antiporter using atomic force microscopy. J Mol Biol. 2004 Jul 23;340(5):1143–52.
49. Thoma J, Bosshart P, Pfreundschuh M, Müller DJ. Out but not in: the large transmembrane β-barrel protein FhuA unfolds but cannot refold via β-hairpins. Structure. 2012 Dec 5;20(12):2185–90.
50. Yu H, Siewny MGW, Edwards DT, Sanders AW, Perkins TT. Hidden dynamics in the unfolding of individual bacteriorhodopsin proteins. Science. 2017 Mar 3;355(6328):945–50.
51. Serdiuk T, Balasubramaniam D, Sugihara J, Mari SA, Kaback HR, Müller DJ. YidC assists the stepwise and stochastic folding of membrane proteins. Nat Chem Biol. 2016 Nov;12(11):911–7.
52. Kawamura S, Gerstung M, Colozo AT, Helenius J, Maeda A, Beerenwinkel N, et al. Kinetic, energetic, and mechanical differences between dark-state rhodopsin and opsin. Structure. 2013 Mar 5;21(3):426–37.
53. Spoerri PM, Sapra KT, Zhang C, Mari SA, Kato HE, Kobilka BK, et al. Conformational plasticity of human protease-activated receptor 1 upon antagonist- and agonist-binding. Structure. 2019 Oct;27(10):1517–1526.
54. Galvanetto N, Maity S, Ilieva N, Ye Z, Laio A, Torre V. Unfolding and identification of membrane proteins from native cell membranes. bioRxiv. 2019 Sep;26;732933.
55. Fu W, Zhang W. Hybrid AFM for nanoscale physicochemical characterization: recent development and emerging applications.Small. 2017, 425;1603525.

REVIEW ARTICLE

Evaluation of *in vitro* neuronal networks for the study of spontaneous activity

Diletta POZZI[a,b*]

[a] *Neurobiology Sector, International School for Advanced Studies (SISSA), via Bonomea 265, 34136 Trieste, Italy.*
[b] *Department of Otolaryngology, Head and Neck Surgery, Stanford University School of Medicine, 300 Pasteur Drive, 94305 Stanford, California.*

*Correspondence: dpozzi@stanford.edu
https://doi.org/10.37175/stemedicine.v1i2.35

ABSTRACT

In the absence of external stimuli, the nervous system exhibits a spontaneous electrical activity whose functions are not fully understood, and that represents the background noise of brain operations. Spontaneous activity has been proven to arise not only *in vivo*, but in *in vitro* neuronal networks as well, following some stereotypical patterns that reproduce the time course of development of the mammalian nervous system. This review provides an overview of *in vitro* models for the study of spontaneous network activity, discussing their ability to reproduce *in vivo* - like dynamics and the main findings obtained with each particular model. While explanted brain slices are able to reproduce the neuronal oscillations typically observed in anaesthetized animals, dissociated cultures allow the use of patient-derived neurons and limit the number of animals used for sample preparation. Moreover, dissociated neurons can be cultured on three-dimensional (3D) substrates that mimic the extra-cellular matrix of the brain. Depending on the material used, 3D substrates are able to increase neuronal connectivity and synaptic strength.

Keywords: Cortical oscillations · Up states · Calcium transients · Neuronal cultures · Brain slices

1. Introduction

Spontaneous neural oscillations have been observed by researchers as early as the 1930s, when the electroencephalography (EEG) machine was invented (1). Specifically, slow oscillations in the frequency of *delta* waves (0.4 - 5 Hz) have first been observed in the cortex of sleeping patients by W Gray Walter (2). Subsequently, slow oscillations in the same frequency range were recorded in anesthetized mammals as well (3). Nevertheless, their function remained one of the most mysterious questions in neuroscience.

A first characterization of neuronal spontaneous oscillations was provided by Steriade *et al*. in 1993. Measuring the electrical activity of cortical and thalamic neurons in anesthetized cats, they observed oscillations composed of a slow depolarizing phase (0.8 - 1.5 s) followed by a long-lasting hyperpolarization. These oscillations had an even lower frequency than the one of *delta* waves, i.e. less than 1 Hz. Therefore, they were referred to as slow oscillations (4). The duration of the depolarizing phase was reduced by administration of the N-methyl-D-aspartate (NMDA) blocker ketamine.

In the hippocampus of anesthetized rats, instead, neuronal oscillations have been observed mainly in the higher, *theta* frequency range (4 - 8 Hz). Hippocampal *theta* waves are generated by the entorhinal input and by CA3 collaterals, through the activation of voltage-dependent Ca^{2+} currents in pyramidal cells. GABAergic and cholinergic neurons contribute as well to the generation of *theta* waves by feed-forward disinhibition of CA1 pyramidal cells by the interneurons, and cholinergic activation of CA3 (5).

Subsequent studies revealed the mechanisms of neuronal oscillations at the cellular level. In particular, some *in vivo* intracellular recordings of neostriatal spiny neurons showed fluctuations between a depolarized and an hyperpolarized membrane potential that were referred to as Up and Down states, respectively (6). Subsequently, Constantinople and Bruno demonstrated that this two-state behavior observed in anesthetized rats was eliminated during wakefulness. In addition, they showed that the dynamics of the awake state and under anesthesia were unaltered even after elimination of afferent thalamic input, suggesting an internal cortical neuromodulation (7). In fact, it was subsequently proven that - in both humans and rodents - the large majority of excitatory synapses onto cortical excitatory neurons come from other cortical excitatory neurons, giving the cortex a vast recurrent connectivity (8).

Received: Jan 25, 2020; Accepted: Mar 4, 2020.

Synchronized Up and Down states are the cellular equivalent of the slow waves observed - at the network level - from human EEG data (9) to anesthetized animals, and even in *in vitro* brain slices isolated from the rest of the brain (10) (**Figure 1**). Since they arise in such different systems, Up and Down states have been proposed to be a default activity pattern resulting from the basic architecture of the network. Nevertheless, they may still have a functional role. From experimental observations of their pathways, Up and Down states are supposed to transfer information across brain regions and to consolidate memories. A current hypothesis is that cortical Up states travel through the entorhinal and visual cortex to the hippocampus phasing the hippocampal rhythm, which in turn transfers its activity pattern - acquired during learning - back to the cortical networks for memory consolidation (5,11). Moreover, the hippocampal *theta* rhythm is believed to be critical for temporal coding/decoding of active neuronal ensembles and the modification of synaptic weights. Down states, as periods of low synaptic inputs, are thought to be necessary for metabolic restoration (see for review 12).

Spontaneous oscillations have been observed also in developing neuronal networks. An example is represented by synchronous plateau assemblies (SPAs) and giant depolarizing potentials (GDPs) in the immature hippocampus. Driven by synchronized GABAergic neurons, GDPs are thought to shape synaptic currents in the immature network (13,14).

In the regulation of this rhythmic activity, a key role is played by inhibitory interneurons. In fact, interneurons were proven to synchronize the spiking of hippocampal pyramidal cells, thus maintaining homeostatic levels of activity in the brain (15). Moreover, spike-triggered local field potentials measured in humans and monkeys were proven to be composed by inhibitory neurons for the largest part, suggesting that the onset of Up states is more driven by a decrease in inhibition rather than an increase in excitation (16).

Not only Up and Down states, but also spontaneous calcium signals were frequently observed in neuronal networks. In fact, intracellular calcium ions increase as a consequence of action potentials and membrane depolarization, entering the cytosol through voltage-gated calcium channels. Spontaneous calcium transients can be considered as indicators of Up and Down states since they follow parallel dynamics, despite being slower (17) (see **Figure 1** for a representation). Calcium imaging techniques have become a diffused method for studying network dynamics, since they allow the simultaneous recording of a large number of neurons.

Many of these findings were obtained *in vitro*. In this review, I discuss the commonly used methods for preparing neuronal samples, the patterns of spontaneous activity that typically arise from each one of these networks and their similarity with *in vivo* neuronal oscillations. Network disorders are at the basis of several pathologies starting from epilepsy, where an increased synchronization of excitatory neurons is causing seizures. Altered network dynamics are also reported in sleep disorders such as narcolepsy, where a specific neuronal population (hypocretin neurons) is reduced in number, and the EEG rhythms normally observed during sleep are modified (18). Moreover, an impaired connectivity of resting-state networks was observed in schizophrenia by functional magnetic resonance imaging (19). *In vitro* experimental models can be a powerful tool for a deeper understanding of network disorders, and for drug screening of new neuroactive compounds.

2. "Traditional" dissociated cultures – Methods, applications and spontaneous activity

Neurons preserve the ability to generate spontaneous electrical pulses even when disrupted from their original connectivity and cultured *in vitro*. In fact, primary dissociated cultures, obtained from embryonic or postnatal

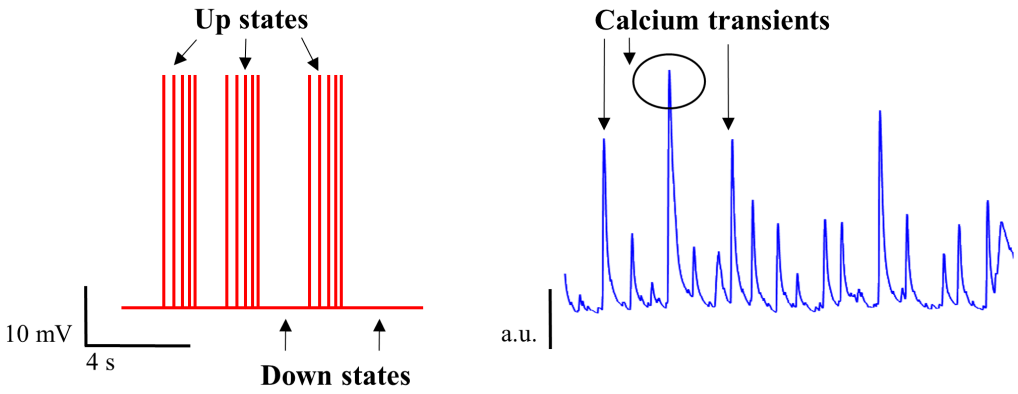


Figure 1. Schematic representation of spontaneous Up and Down states and calcium transients. The red trace on the left represents an example of a typical intracellular recording from a neuron in spontaneous conditions. Up and Down states are visible as rhythmic oscillations between membrane depolarization and hyperpolarization. The blue trace on the right depicts some typical calcium transients recorded from a single neuron in spontaneous condition. The time scale is the same for both traces. Note the increased duration of a single calcium transient comparing to a single Up state event.

rodent brain regions such as the hippocampus or the cortex, represent a well-established *in vitro* model for studying neuronal networks, both on a short (network formation) and a long term (differentiation, maturation) scale (20,21). When cells are enzymatically dissociated, plated on suitable substrates and cultured in appropriate media, they grow processes and form *ex-novo* a functional network. Dissociated cell cultures are traditionally plated on glass-made flat surfaces, allowing a fine space resolution for morphological investigation of small cellular structures, like neurites and growth cones. Neurons attach to surfaces through their anionic phospholipids (e.g. phosphatidylserine) (22), therefore the surface needs to be positively charged. Previous coating with polyaminoacids such as polylysine and polyornithine have since long been used for this purpose, together with the addition of components mimicking the extracellular matrix *in situ* (i.e. Matrigel® (23)). The morphology and development of dissociated neurons as well as their network activity are strongly influenced by the culturing conditions and, in particular, the culturing medium used (24). Different chemical compositions have been tested and are commercially available for neuronal cell cultures, all based on a physiological solution added with a particular supplier of nutrients and growth factors. Fetal bovine serum, for example, provides neurons with sufficient factors for short-term cultures. As any animal-derived product, however, it is prone to batch-to-batch variability (25) and, most of all, it results in a low differentiation of neuronal types (26). For this reason, serum-free media were tested with different additives. A very commonly used one is B27 Supplement ®, aimed at obtaining almost pure neuronal cultures (27). Another option is to use glial cells, specifically astrocytes, as the intrinsic source of growth factors in the neuronal culture. For this purpose, it is possible to plate neurons on a glial feeder layer (28) or, in order to avoid the spatial interference of astrocytes while maintaining the soluble factors released by them, to grow neurons in an astrocyte-conditioned medium (ACM). ACM is supposed to contain lipids, thrombospondins and Brain Derived Neurotrophic Factor (BDNF), that are essential components for neuronal growth and survival (29). In fact, it was demonstrated to increase the life span of dissociated cultures up to several months (24,30).

Once established the optimal plating and growing conditions for the specific experimental purposes, it is possible to explore the properties of dissociated networks with a fine, single cell resolution. A common application is to analyze different neuronal populations optically labeled. A comparison of spike parameters between the GABAergic and glutamatergic neurons has been possible, in fact, because of the labelling of the former with a green fluorescent protein (31,32).

Network (i.e. synchronized) activity in dissociated neurons emerges at 3 - 7 days in-vitro (DIV) and matures over the following several weeks in culture, as shown by electrophysiological and imaging measurements (33,34). Synchronized neuronal activity has been proven to be mediated by the NMDA receptor, by serotonin and acetylcholine (35). Network activity is also strongly dependent on the initial plating density of neurons: neurons plated at high density present higher rates of less synchronized activity comparing to sparse networks (28). In fact, there is an inverse relationship between the density of the network and its synaptic strength, such as sparse neurons establish less but stronger synaptic connections, as shown also by the longer duration of their bursts. Interestingly, dense neuronal cultures present a peak in synchronicity at DIV15-16, followed by a decrease (36) that resembles the one observed in the intact cortex during development (37).

Recent studies showed some differences in spontaneous, single-unit spike activity between dissociated neurons and awake animals: *in vitro*, neurons were more likely to fire in synchronized bursts; while *in vivo* the firing events were mainly composed by single spikes. Moreover, a greater variability was observed *in vitro* when quantifying the frequency of firing (38). However, as previously mentioned, the variability in spontaneous activity of dissociated neurons can be minimized by adopting a consistent culturing method and, most of all, the same plating density among different preparations.

Dissociated neurons served as a model for studying the progression of neuronal disorders. In a recent study (39), neurons dissociated from a mouse model of Fragile X syndrome (Fmr1 KO) were recorded by means of microelectrode arrays (MEAs) from 7 to 35 DIV, and exhibited an abnormally increased synchronization of spontaneous firing events at mature stages. Moreover, they showed oscillations in the *beta* frequency range, which were not observed in controls. These results are in line with previous studies showing network hyperactivity and neuronal hyperconnectivity in the intact cortex of Fmr1 KO mice (40,41).

Finally, dissociated cultures represent the only method for growing human–induced pluripotent stem cells (hiPSCs). These precursors can be differentiated into neurons, allowing the investigation of their molecular and electrophysiological properties (42,43). Most of all, they can be derived from patients for the purpose of specific drug development or cell therapies. Using appropriate differentiation methods, hiPSCs can generate three-dimensional (3D) brain organoids composed by neurons and glial cell types (44).

3. 3D neuronal cultures – methods, applications and spontaneous activity

Traditional, 2D neuronal cultures have been replaced in many studies by more recent 3D architectures for several reasons. First, neurons cultured on glass coverslips are prone to artifacts due to the flat condition in which they grow. Glass coverslips are in fact distant from the physiological extra-cellular environment of the brain, which is 3D and soft (less than 1 kPa in Young modulus, while the stiffness of the glass is in the order of GPa). Neurites forced to grow in a planar region present indeed an abnormal excitability (45). Moreover, using plastic materials instead of glass allows for directing the

differentiation of stem cells towards bone, muscle and neuronal cell types by progressively decreasing the stiffness of the substrate (46). Therefore, 3D substrates with appropriate stiffness were developed for neuronal cultures, and called neuroscaffolds.

An interesting example is represented by the carbon-based neuroscaffolds, engineered for interfacing the cell membrane structures. Carbon-based materials have since long raised the interest of neuroscientists because of the electrical conductivity of the material, that might find its most obvious application in restoring/improving neuronal connectivity (47,48). These materials can be assembled in the form of wrapped graphite sheets, and therefore called carbon nanotubes (CNTs), or in the form of graphene foams. Both cases were proven to be biocompatible and conductive scaffolds not only for culturing primary neurons, but also for neural stem cell differentiation (49,50).

Interestingly, primary hippocampal neurons plated on 3D graphene showed a more frequent and synchronized spontaneous activity comparing to both 2D graphene and traditional 2D glass substrates (51). These results suggest a specific role of dimensionality in network dynamics that is independent from the material where neurons grow. However, a recent study demonstrated that a single layer of graphene modifies the ionic concentration at the interface with neurons so that they will exhibit stronger potassium currents and will shift from adapting to tonically firing. This effect was not observed when neurons were plated on different conductive materials (52). Taken together, these results suggest that both the 3D structure and the electrical properties of the substrate can increase neuronal firing and connectivity. In fact, adding CNTs as a coating to a 3D substrate was proven to boost the synaptic activity of dissociated neurons, as shown by calcium imaging experiments (53). For these reasons, carbon-based materials have been tested in *in vivo* implants for restoring nerve mechanical or chemical injury (54,55) as well as for deep brain stimulation and monitoring (56,57).

Polydimethylsiloxane (PDMS) and other plastic materials represent another interesting category for neuronal studies: while not electrically conductive, these materials have the advantage of being highly ductile and with a Young modulus closer to the one of brain tissues. The fabrication of these types of scaffolds usually requires photolithography techniques combined with micro/nano-patterning (58). Some preliminary treatments - such as plasma cleaning and polylysine/polyornithine coating - are necessary to reduce the hydrophobicity of the plastic and ensure neuronal attachment. Plastic micropillars have often been used as substrates for cell cultures in order to decrease the stiffness of the original material: while a flat surface of PDMS has a Young modulus of around 1 MPa, PDMS micropillars have a stiffness of 5 - 20 kPa depending on their diameter and height (59). For these reasons, micropillars are useful for determining the mechanical responses and forces exerted by cells (60). Plastic materials can also be molded at the nanoscale: the use of nanopillars with different dimensions and stiffness has been proven to control the adhesion and differentiation of neuronal precursors (61). Moreover, the addition of plastic nanofibers made of poly (lactic-co-glycolic acid) to PDMS micropillars was proven to increase the frequency and amplitude of spontaneous calcium transients in primary hippocampal neurons (62). Due to their biocompatibility, PDMS-based microchannel electrodes have successfully been implanted in rodents for monitoring neural activity over time and for providing a mechanical support for peripheral nerve regeneration (55).

In addition, scaffold-free 3D cultures can be derived from hiPSCs to generate brain organoids. HiPSC-derived organoids have been used for drug screening and modeling of central nervous system diseases (63).

In summary, 3D neuronal cultures recapitulate several aspects of *in vivo* brain tissue, from synaptic properties of individual neurons to network synchronization. Moreover, comparing to 2D cultures, they offer the unique advantage of recreating organoids with neuronal and glial types interconnected in a cyto-architecture that is able to reproduce the one of the brain.

4. Explanted brain slices – methods, applications and spontaneous activity

Slices of the brain or spinal cord can be isolated from the region of interest and directly used for electrophysiological measurements and imaging (17,61). These preparations are referred to as acute slices and, in contrast to dissociated cell cultures, they preserve the native cyto-architecture of the tissue. Acute cortical slices are able to reproduce the most characteristic feature of the cerebral cortex, i.e. slow oscillations between Up and Down states, as shown by the work of Sanchez-Vives starting from the 2000s (65,66). In this work, intracellular recordings from ferret visual and prefrontal cortical slices were performed using microelectrode arrays. While no spontaneous activity was recorded in the "traditional" slice bathing medium (2 mM Ca^{2+}, 1 mM Mg^{2+}, 2.5 mM K^{+}), some rhythmic oscillations were observed after changing the bath solution to a composition more similar to the brain interstitial fluid (1.0 or 1.2 mM Ca^{2+}, 1 mM Mg^{2+} and 3.5 mM K^{+}). These oscillations were initiated by pyramidal neurons of layer V, had a periodicity of once every 3.5 seconds and were stable for the life of the slice (12 hours). The cortical slices were prepared from 3 to 8 month-old ferrets and showed a close similarity to those observed in adult, anesthetized cats. In fact, the spontaneous activity of slices was studied as the most likely *in vitro* equivalent of the activity of the brain in the absence of external stimuli, i.e. during quiescent sleep or under anesthesia. In a more recent study from the same group, variable propagation modes of Up states were observed, suggesting a memoryless generation of slow oscillations (10).

Slices of tissue cultured *in vitro* for several days are instead referred to as organotypic slice cultures. These

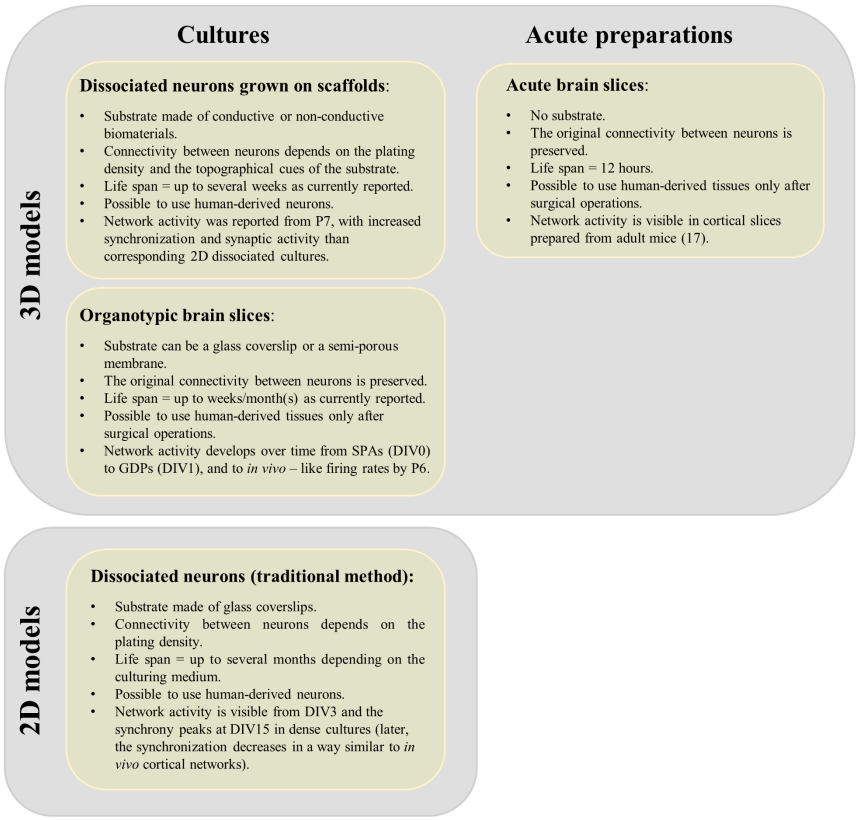


Figure 2. Summary and comparison of in vitro neuronal network models. 3D models include brain slices and dissociated neurons grown on scaffolds. 2D models are represented by dissociated neurons grown on glass (traditional method). The only preparation that does not require culturing is represented by acute brain slices.

preparations were proven to regrow synaptic processes and to restore the functions potentially impaired during the cutting procedure. There are two commonly used methods for preparing organotypic cultures from nervous tissues. In the Gähwiler method (67), slices are cut and attached on glass coverslips by coagulating chicken plasma with thrombin; the coverslips are then inserted into plastic tubes and kept in serum-added culturing medium on a roller drum (ensuring balanced time of oxygenation) at 37 °C. The Stoppini method (68) follows a similar procedure, but the slices are placed on sterile, transparent semiporous membranes in Petri or multiwell dishes, and kept at the interface between air and culturing medium at 37 °C. In this case, it is possible to culture the slices in Neurobasal medium added with B27 supplement, with no addition of animal serum. However, organotypic slices grown in the absence of serum have been reported to develop spontaneous epileptiform activity (69).

The use of organotypic cultures allows to perform chronic multineuron imaging. For example, daily morphological analyses and calcium imaging in hippocampal slices revealed that GABAergic interneurons are involved in the generation of SPAs at DIV0; starting from DIV1, the spontaneous activity of the slice switches to giant depolarizing potentials (70), as in line with previous findings on the development of hippocampal networks (71).

Interestingly, in 2014, Okamoto *et al*. demonstrated that organotypic hippocampal slices exhibit firing rates and excitatory synaptic activity more similar to those observed *in vivo* comparing to acute slices (72). In fact, cutting several neurites during the preparation might affect the network activity of acute slices, which do not have time to regrow these processes.

Without the technical issues presented by *in vivo* measurements (i.e. background noise, the interference of anesthetics and the need of complex surgeries), experiments on slices have been useful to address questions like how inhibitory neurons coordinate synchrony in the network. In a study of 1995 (73), GABAergic neurons were anatomically identified from hippocampal slices and depolarized to produce fast inhibitory postsynaptic potentials: the excitatory, pyramidal neurons synaptically connected to them exhibited some sub-threshold oscillations in the *theta* frequency range, demonstrating that these spontaneous oscillations in the hippocampus are tuned by the inhibitory activity of GABAergic neurons. Subsequently, a reconstruction of functional connectivity in acute hippocampal slices was performed by Bonifazi *et al*., combining electrophysiological recordings with calcium imaging: connections in between neurons were assumed

on the base of the temporal correlations of their activation. Stimulating some specific, highly connected neurons in the network, they observed a higher occurrence of spontaneous GDPs in the network, suggesting the existence of functional hubs that were identified as a subpopulation of GABAergic neurons (64). Results obtained in cortical slices also showed that networks of fast spiking, electrically coupled interneurons coordinate the spontaneous activity (74).

These results are in line with subsequent studies in behaving mice, characterizing the membrane potential dynamics of fast-spiking inhibitory neurons vs excitatory ones (75), and also in human and monkeys during sleep, when the local field potential of the neocortex was measured with MEAs (16). Taken together, these results confirm that brain slices are a suitable *in vitro* model for investigating network dynamics. They are prepared typically from animal models. However, the use of human brain slices is possible from resected or postmortem tissues (76,77).

5. Conclusions

Both dissociated and organotypic cultures exhibit a spontaneous neuronal firing with *in vivo*-like patterns in their course of development (quantified by days *in vitro*). Acute and organotypic slices were used for the most remarkable studies on spontaneous neuronal oscillations, given the fact that they preserve the original cyto-architecture of the region of interest. On the other side, 3D substrates made of carbon-based or plastic materials can compensate for the lack of architecture in traditional neuronal cultures and boost the synaptic activity of dissociated neurons, opening the way for neural prostheses and cell culture implants in the central and peripheral nervous system. However, there is no strong evidence that differentiated neurons plated on neuroscaffolds can reproduce the time course of development of intact neuronal networks. HiPSCs, as opposite, are able to reproduce some aspects of human brain development when differentiated into scaffold-free, 3D organoids. HiPSCs offer in fact a powerful model for drug screening and customized therapy. Finally, the use of dissociated neurons and hiPSCs can diminish the number of experimental animals used, making it a more ethical tool for scientific studies and even a pre-clinical test platform for screening of new neuroactive compounds.

For a summary and comparison of the properties of the *in vitro* models discussed within this review, see **Figure 2**.

Conflict of interest

The authors declare that they have no conflict of interest.

References

1. Berger H. Über das Elektrenkephalogramm des Menschen. Archiv f. Psychiatrie. 1929;87:527–570.
2. Freeman WJ. W.G. Walter: The Living Brain. In: Palm G, Aertsen A, editors. Brain Theory. Springer Berlin Heidelberg; 1986. p. 237–8.
3. Steriade M, Dossi RC, Nuñez A. Network modulation of a slow intrinsic oscillation of cat thalamocortical neurons implicated in sleep delta waves: cortically induced synchronization and brainstem cholinergic suppression. J Neurosci. 1991 Oct;11(10):3200–17.
4. Steriade M, Nuñez A, Amzica F. A novel slow (< 1 Hz) oscillation of neocortical neurons in vivo: depolarizing and hyperpolarizing components. J Neurosci. 1993 Aug;13(8):3252–65.
5. Sirota A, Buzsáki G. Interaction between neocortical and hippocampal networks via slow oscillations. Thalamus Relat Syst. 2005 Dec;3(04):245.
6. Wilson CJ, Kawaguchi Y. The origins of two-state spontaneous membrane potential fluctuations of neostriatal spiny neurons. J Neurosci. 1996 Apr 1;16(7):2397–410.
7. Constantinople CM, Bruno RM. Effects and mechanisms of wakefulness on local cortical networks. Neuron. 2011 Mar;69(6):1061–8.
8. Seeman SC, Campagnola L, Davoudian PA, Hoggarth A, Hage TA, Bosma-Moody A, et al. Sparse recurrent excitatory connectivity in the microcircuit of the adult mouse and human cortex. eLife. 2018 Sep 26;7:e37349.
9. Massimini M, Huber R, Ferrarelli F, Hill S, Tononi G. The sleep slow oscillation as a traveling wave. J Neurosci. 2004 Aug 4;24(31):6862–70.
10. Capone C, Rebollo B, Muñoz A, Illa X, Del Giudice P, Sanchez-Vives MV, et al. Slow waves in cortical slices: how spontaneous activity is shaped by laminar structure. Cereb Cortex. 2019 Jan 1;29(1):319-335.
11. Ji D, Wilson MA. Coordinated memory replay in the visual cortex and hippocampus during sleep. Nat Neurosci. 2007 Jan;10(1):100–7.
12. Buzsáki G. Theta Oscillations in the hippocampus. Neuron. 2002 Jan;33(3):325–40.
13. Mohajerani MH, Cherubini E. Role of giant depolarizing potentials in shaping synaptic currents in the developing hippocampus. Crit Rev Neurobiol. 2006;18(1–2):13–23.
14. Ben-Ari Y, Cherubini E, Corradetti R, Gaiarsa JL. Giant synaptic potentials in immature rat CA3 hippocampal neurones. J Physiol. 1989 Sep;416:303–25.
15. Somogyi P, Klausberger T. Defined types of cortical interneurone structure space and spike timing in the hippocampus. J Physiol. 2005 Jan 1;562(Pt 1):9–26.
16. Teleńczuk B, Dehghani N, Le Van Quyen M, Cash SS, Halgren E, Hatsopoulos NG, et al. Local field potentials primarily reflect inhibitory neuron activity in human and monkey cortex. Sci Rep. 2017 Jan 11;7:40211.
17. Cossart R, Aronov D, Yuste R. Attractor dynamics of network UP states in the neocortex. Nature. 2003 May 15;423(6937):283–8.
18. Thannickal TC1, Moore RY, Nienhuis R, Ramanathan L, Gulyani S, Aldrich M, Cornford M, Siegel JM. Reduced number of hypocretin neurons in human narcolepsy: Neuron. 2000 Sep;27(3):469-474.
19. Karbasforoushan H1, Woodward ND. Resting-state networks in schizophrenia. Curr Top Med Chem. 2012 Dec;12:2404-2414.
20. Dotti CG, Sullivan CA, Banker GA. The establishment of polarity by hippocampal neurons in culture. J Neurosci. 1988 Apr;8(4):1454–68.
21. da Silva JS, Dotti CG. Breaking the neuronal sphere: regulation of the actin cytoskeleton in neuritogenesis. Nat Rev Neurosci. 2002 Sep;3(9):694–704.
22. Eichberg J, Hauser G. The subcellular distribution of polyphosphoinositides in myelinated and unmyelinated rat brain. Biochim Biophys Acta BBA - Lipids Lipid Metab. 1973 Nov 29;326(2):210–23.
23. Hughes CS, Postovit LM, Lajoie GA. Matrigel: a complexprotein mixture required for optimal growth of cell culture. Proteomics. 2010 May;10(9):1886–90.
24. Pozzi D, Ban J, Iseppon F, Torre V. An improved method for growing neurons: comparison with standard protocols.

J Neurosci Methods. 2017 Mar 15;280:1-10.
25. Lafon-Cazal M, Adjali O, Galeotti N, Poncet J, Jouin P, Homburger V, et al. Proteomic analysis of astrocytic secretion in the mouse: comparison with the cerebrospinal fluid proteome. J Biol Chem. 2003 Jun 27;278(27):24438–48.
26. Landis DM, Weinstein LA, Skordeles CJ. Serum influences the differentiation of membrane structure in cultured astrocytes. Glia. 1990;3(3):212–21.
27. Brewer GJ, Torricelli JR, Evege EK, Price PJ. Optimized survival of hippocampal neurons in B27-supplemented Neurobasal, a new serum-free medium combination. J Neurosci Res. 1993 Aug 1;35(5):567–76.
28. Ivenshitz M, Segal M. Neuronal density determines network connectivity and spontaneous activity in cultured hippocampus. J Neurophysiol. 2010 Aug;104(2):1052–60.
29. Ebrahimi M, Yamamoto Y, Sharifi K, Kida H, Kagawa Y, Yasumoto Y, et al. Astrocyte-expressed FABP7 regulates dendritic morphology and excitatory synaptic function of cortical neurons. Glia. 2016 Jan;64(1):48–62.
30. Todd GK, Boosalis CA, Burzycki AA, Steinman MQ, Hester LD, Shuster PW, et al. Towards neuronal organoids: a method for long-term culturing of high-density hippocampal neurons. PloS One. 2013;8(4):e58996.
31. Weir K, Blanquie O, Kilb W, Luhmann HJ, Sinning A. Comparison of spike parameters from optically identified GABAergic and glutamatergic neurons in sparse cortical cultures. Front Cell Neurosci. 2014;8:460.
32. Tamamaki N, Yanagawa Y, Tomioka R, Miyazaki J-I, Obata K, Kaneko T. Green fluorescent protein expression and colocalization with calretinin, parvalbumin, and somatostatin in the GAD67-GFP knock-in mouse. J Comp Neurol. 2003 Dec 1;467(1):60–79.
33. Murphy TH, Blatter LA, Wier WG, Baraban JM. Spontaneous synchronous synaptic calcium transients in cultured cortical neurons. J Neurosci. 1992 Dec;12(12):4834–45.
34. Cohen E, Ivenshitz M, Amor-Baroukh V, Greenberger V, Segal M. Determinants of spontaneous activity in networks of cultured hippocampus. Brain Res. 2008 Oct 15;1235:21–30.
35. Björklund U, Persson M, Rönnbäck L, Hansson E. Primary cultures from cerebral cortex and hippocampus enriched in glutamatergic and GABAergic neurons. Neurochem Res. 2010 Nov;35(11):1733–42.
36. Biffi E, Regalia G, Menegon A, Ferrigno G, Pedrocchi A. The influence of neuronal density and maturation on network activity of hippocampal cell cultures: a methodological study. PLoS ONE. 2013 Dec 27;8(12):e83899.
37. Golshani P, Gonçalves JT, Khoshkhoo S, Mostany R, Smirnakis S, Portera-Cailliau C. Internally mediated developmental desynchronization of neocortical network activity. J Neurosci. 2009 Sep 2;29(35):10890–9.
38. Belle AM, Enright HA, Sales AP, Kulp K, Osburn J, Kuhn EA, et al. Evaluation of in vitro neuronal platforms as surrogates for in vivo whole brain systems. Sci Rep. 2018 Jul 17;8:10820.
39. Moskalyuk A, Kooy RF, Giugliano M. Single-cell and neuronal network alterations in an in vitro model of Fragile X syndrome. bioRxiv. 2018 Jul 10;366997.
40. Gonçalves JT, Anstey JE, Golshani P, Portera-Cailliau C. Circuit level defects in the developing neocortex of Fragile X mice. Nat Neurosci. 2013 Jul;16(7):903–9.
41. Testa-Silva G, Loebel A, Giugliano M, de Kock CPJ, Mansvelder HD, Meredith RM. Hyperconnectivity and slow synapses during early development of medial prefrontal cortex in a mouse model for mental retardation and autism. Cereb Cortex N Y N 1991. 2012 Jun;22(6):1333–42.
42. Prè D, Nestor MW, Sproul AA, Jacob S, Koppensteiner P, Chinchalongporn V, et al. A time course analysis of the electrophysiological properties of neurons differentiated from human induced pluripotent stem cells (iPSCs). PloS One. 2014;9(7):e103418.
43. Zhou R, Jiang G, Tian X, Wang X. Progress in the molecular mechanisms of genetic epilepsies using patient-induced pluripotent stem cells. Epilepsia Open. 2018 Jul 8;3(3):331–9.
44. Marton RM, Miura Y, Sloan SA, Li Q, Revah O, Levy RJ, et al. Differentiation and maturation of oligodendrocytes in human three-dimensional neural cultures. Nat Neurosci. 2019 Mar;22(3):484–91.
45. Irons HR, Cullen DK, Shapiro NP, Lambert NA, Lee RH, Laplaca MC. Three-dimensional neural constructs: a novel platform for neurophysiological investigation. J Neural Eng. 2008 Sep;5(3):333–41.
46. Engler AJ, Sen S, Sweeney HL, Discher DE. Matrix Elasticity Directs Stem Cell Lineage Specification. Cell. 2006 Aug 25;126(4):677–89.
47. Lovat V, Pantarotto D, Lagostena L, Cacciari B, Grandolfo M, Righi M, et al. Carbon nanotube substrates boost neuronal electrical signaling. Nano Lett. 2005 Jun;5(6):1107–10.
48. Cellot G, Toma FM, Varley ZK, Laishram J, Villari A, Quintana M, et al. Carbon nanotube scaffolds tune synaptic strength in cultured neural circuits: novel frontiers in nanomaterial-tissue interactions. J Neurosci. 2011 Sep 7;31(36):12945–53.
49. Li N, Zhang Q, Gao S, Song Q, Huang R, Wang L, et al. Three-dimensional graphene foam as a biocompatible and conductive scaffold for neural stem cells. Sci Rep. 2013;3:1604.
50. Posypanova GA, Gayduchenko IA, Moskaleva EY, Fedorov GE. Neuronal differentiation of PC12 cells and mouse neural stem cells on carbon nanotube films. Cell Tissue Biol. 2016 May 1;10(3):194–201.
51. Ulloa Severino FP, Ban J, Song Q, Tang M, Bianconi G, Cheng G, et al. The role of dimensionality in neuronal network dynamics. Sci Rep. 2016 Jul 11;6:29640.
52. Pampaloni NP, Lottner M, Giugliano M, Matruglio A, D'Amico F, Prato M, et al. Single-layer graphene modulates neuronal communication and augments membrane ion currents. Nat Nanotechnol. 2018;13(8):755–64.
53. Bosi S, Rauti R, Laishram J, Turco A, Lonardoni D, Nieus T, et al. From 2D to 3D: novel nanostructured scaffolds to investigate signalling in reconstructed neuronal networks. Sci Rep. 2015;5:9562.
54. Aurand ER, Usmani S, Medelin M, Scaini D, Bosi S, Rosselli FB, et al. Nanostructures to engineer 3D neural-interfaces: directing axonal navigation toward successful bridging of spinal segments. Adv Funct Mater. 28(12):1700550.
55. Musick KM, Rigosa J, Narasimhan S, Wurth S, Capogrosso M, Chew DJ, et al. Chronic multichannel neural recordings from soft regenerative microchannel electrodes during gait. Sci Rep. 2015 Sep 24;5:14363.
56. Park D-W, Ness JP, Brodnick SK, Esquibel C, Novello J, Atry F, et al. Electrical neural stimulation and simultaneous in vivo monitoring with transparent graphene electrode arrays implanted in GCaMP6f mice. ACS Nano. 2018 23;12(1):148–57.
57. Masvidal-Codina E, Illa X, Dasilva M, Calia AB, Dragojević T, Vidal-Rosas EE, et al. High-resolution mapping of infraslow cortical brain activity enabled by graphene microtransistors. Nat Mater. 2019 Mar;18(3):280–8.
58. Li S, Severino FPU, Ban J, Wang L, Pinato G, Torre V, et al. Improved neuron culture using scaffolds made of three-dimensional PDMS micro-lattices. Biomed Mater. 2018;13(3):034105.
59. Wei J, Shi J, Wang B, Tang Y, Tu X, Roy E, et al. Fabrication of adjacent micropillar arrays with different heights for cell studies. Microelectron Eng. 2016 Jun;158:22–5.
60. Nagayama K, Inoue T, Hamada Y, Matsumoto T. A novel patterned magnetic micropillar array substrate for analysis of cellular mechanical responses. J Biomech. 2017 Dec 8;65:194–202.
61. Migliorini E, Ban J, Grenci G, Andolfi L, Pozzato A, Tormen M,

et al. Nanomechanics controls neuronal precursors adhesion and differentiation. Biotechnol Bioeng. 2013;110(8):2301–10.
62. Wei J, Pozzi D, Ulloa Severino FP, Torre V, Chen Y. Fabrication of PLGA nanofibers on PDMS micropillars for neuron culture studies. Microelectron Eng. 2017 May;175(C):67–72.
63. D'Aiuto L, Naciri J, Radio N, Tekur S, Clayton D, Apodaca G, et al. GABAergic hub neurons orchestrate synchrony in developing hippocampal networks. Science. 2009 Dec 4;326(5958):1419–24.
65. Sanchez-Vives MV, McCormick DA. Cellular and network mechanisms of rhythmic recurrent activity in neocortex. Nat Neurosci. 2000 Oct;3(10):1027–34.
66. Compte A, Reig R, Descalzo VF, Harvey MA, Puccini GD, Sanchez-Vives MV. Spontaneous high-frequency (10-80 Hz) oscillations during up states in the cerebral cortex in vitro. J Neurosci. 2008 Dec 17;28(51):13828–44.
67. Gähwiler BH, Capogna M, Debanne D, McKinney RA, Thompson SM. Organotypic slice cultures: a technique has come of age. Trends Neurosci. 1997 Oct;20(10):471–7.
68. Stoppini L, Buchs PA, Muller D. A simple method for organotypic cultures of nervous tissue. J Neurosci Methods. 1991 Apr;37(2):173–82.
69. Magalhães DM, Pereira N, Rombo DM, Beltrão-Cavacas C, Sebastião AM, Valente CA. Ex vivo model of epilepsy in organotypic slices—a new tool for drug screening. J Neuroinflammation. 2018 Jul 11;15(1):203.
70. Allene C, Picardo MA, Becq H, Miyoshi G, Fishell G, Cossart R. Dynamic changes in Interneuron morphophysiological properties mark the maturation of hippocampal network activity. J Neurosci. 2012 May 9;32(19):6688–98.
71. Crépel V, Aronov D, Jorquera I, Represa A, Ben-Ari Y, Cossart R. A parturition-associated nonsynaptic coherent activity pattern in the developing hippocampus. Neuron. 2007 Apr 5;54(1):105–20.
72. Okamoto K, Ishikawa T, Abe R, Ishikawa D, Kobayashi C, Mizunuma M, et al. Ex vivo cultured neuronal networks emit in vivo-like spontaneous activity. J Physiol Sci. 2014 Nov;64(6):421–31.
73. Cobb SR, Buhl EH, Halasy K, Paulsen O, Somogyi P. Synchronization of neuronal activity in hippocampus by individual GABAergic interneurons. Nature. 1995 Nov 2;378(6552):75–8.
74. Galarreta M, Hestrin S. A network of fast-spiking cells in the neocortex connected by electrical synapses. Nature. 1999 Nov 4;402(6757):72–5.
75. Gentet LJ, Avermann M, Matyas F, Staiger JF, Petersen CCH. Membrane potential dynamics of GABAergic neurons in the barrel cortex of behaving mice. Neuron. 2010 Feb;65(3):422–35.
76. Wickham J, Brödjegård NG, Vighagen R, Pinborg LH, Bengzon J, Woldbye DPD, et al. Prolonged life of human acute hippocampal slices from temporal lobe epilepsy surgery. Sci Rep. 2018 Mar 7;8(1):1–13.
77. Qi X-R, Verwer RWH, Bao A-M, Balesar RA, Luchetti S, Zhou J-N, et al. Human brain slice culture: a useful tool to study brain disorders and potential therapeutic compounds. Neurosci Bull. 2019 Apr;35(2):244–52.

RESEARCH ARTICLE

Potential inhibitors for targeting Mpro and Spike of SARS-CoV-2 based on sequence and structural pharmacology analysis

Chuanjun SHU[1#], Xuan HUANG[2#], Ting HUANG[3], Li CHEN[2], Bing YAO[2*], Jianwei ZHOU[4*], Cheng DENG[5*]

[1] *Department of Bioinformatics, School of Biomedical Engineering and Informatics, Nanjing Medical University, Nanjing, 211166, China.*
[2] *Reproductive Medical Center, Jinling Hospital Affiliated to The Medical School of Nanjing University, Nanjing, 210002, China.*
[3] *Department of Pharmacy, Xuzhou Maternity and Child Health Care Hospital, Xuzhou, 221000, China.*
[4] *Department of Molecular Cell Biology & Toxicology, Center for Global Health, School of Public Health, Nanjing Medical University, Nanjing 211166, China.*
[5] *Jiangsu Key Laboratory for Biodiversity and Biotechnology, College of Life Sciences, Nanjing Normal University, Nanjing, 210023, China.*

[#]Contributed equally.
*Correspondence: dengcheng@njnu.edu.cn; jwzhou@njmu.edu.cn; yaobing@nju.edu.cn
https://doi.org/10.37175/stemedicine.v1i2.41

ABSTRACT

The SARS-CoV-2 outbreak has spread rapidly and widely since December 2019, and the effective drugs are urgently needed. The two key proteins, Mpro and Spike, are attractive therapy targets for developing drugs against SARS-CoV-2 infection. In this study, we searched for the potential inhibitors targeting Mpro and Spike based on protein sequences and structural pharmacological analysis. We found that both Mpro and Spike of SARS-CoV-2 were homologous with bat SARS-like-CoV. SARS-CoV-2 Mpro showed high conservation (sequence similarities >99%), and the existing few point mutants in different patients from diverse cities suggested that SARS-CoV-2 probably underwent adaptive evolution when the virus infection transmitted from Wuhan patients to other non-Wuhan patients. Moreover, some inhibitors for SARS-CoV Mpro could probably inhibit the activity of SARS-CoV-2 Mpro, because they do not target conserved mutated sites of SARS-CoV-2 Mpro, such as SDJ, ACE-THR-VAL-ALC-HIS-H, B4Z inhibitor, Beclabuvir, Saquinavir, and Lopinavir. In contrast, Spike of SARS-CoV-2 had more mutations and some mutant sites were distributed in the interaction domain between Spike and ACE2. A new peptide FRKSNLKPFERDISTEIYQAGSTPC, based on interactions between Spike and ACE2, could be a potential drug to treat SARS-CoV-2 patients. In summary, our study provided potential new inhibitors for targeting Mpro and Spike in SARS-CoV-2 virus-infected patients based on sequence and structural pharmacology analysis.

Keywords: SARS-CoV-2 · Mpro · Spike · Structural pharmacology analysis · New peptide inhibitor

Introduction

In December 2019, a pneumonia associated with the 2019 novel coronavirus (SARS-CoV-2) occurred in Wuhan, Hubei Province, China (1). And the new coronavirus pneumonia (NCP) named as COVID-19 has spread rapidly over many countries around the world (2, 3). As of 30 March 2020, >730,000 cases have been confirmed and the number of patients is still increasing, with an estimated mortality risk of ~ 4.6% (3, 4). However, the source of the virus, the effective drugs, and the pathogenesis are still not clear (5, 6).

SARS-CoV-2 has a single-stranded RNA, with distinct clade from the β-coronaviruses associated with the human severe acute respiratory syndrome (SARS) and was classified in the β-coronavirus 2b lineage (7, 8). Similar to SARS, SARS-CoV-2 genome encodes non-structural proteins, structural proteins and accessory proteins (7). Non-structural protein, such as 3-chymotrypsin-like protease (3CLpro, also known as main protease Mpro), is one of the key enzymes for the viral life cycle (9). Structural protein-Spike protein, responsible for viral entry, binds to the cellular receptor angiotensin-converting enzyme 2 (ACE2) and mediates the fusion between the viral and cellular membranes (10). There are two regions

Received: Mar 30, 2020; Accepted: Mar 31, 2020.

(S1 and S2) in the Spike protein (11). In the S1 region, there is a receptor binding domain (Spike-RBD) that interacts with ACE2 (12). The functional importance of the Mpro and Spike in the viral life cycle has attracted a lot of interests for developing drugs against SARS-CoV-2.

Many scientists have envisaged that vaccines, monoclonal antibodies, peptides, interferon therapies and small-molecule drugs might be used to control and prevent emerging infections (13-15). However, there is no evidence to support specific drug treatment against the NCP in suspected or confirmed cases. Hence, our objectives were to search for potential drugs based on sequence and structural pharmacological analysis from Mpro and Spike of SARS-CoV-2.

Materials and Methods

Data preparation

Whole genome sequences of SARS-CoV-2 were downloaded from CNCB/BIG (https://bigd.big.ac.cn/ncov) and NCBI. According to gene locations, the nucleotide sequences of Mpro, Spike and Spike-S1 of SARS-CoV-2 and their corresponding amino acids were acquired by BioEdit software (16). To explore the origin of proteins, each of Mpro, Spike and Spike-S1 in SARS-CoV-2 was employed in BLAST search. The 100 sequences with the highest similarity to Mpro/Spike/Spike-S1 of SARS-CoV-2 were downloaded from the BLAST results, redundant sequences were deleted (17). Subsequently, 93 Mpro sequences of SARS-CoV-2 and 87 Mpro sequences of SARS/SARS-like, and 82 Spike/Spike-S1 sequences of SARS-CoV-2 and 98 Spike/Spike-S1 sequences of SARS/SARS-like were obtained.

Three-dimensional (3D) spatial structure of Mpro of SARS-CoV-2 was downloaded from the PDB database (18). And 72 Mpro structures in the PDB database were download for comparing the Mpro structure of SARS-CoV-2 with other coronavirus. Meanwhile, structures of Mpro-ligand complex and Spike-ACE2 complex for SARS-CoV were downloaded from the PDB database (PDB ID: 5i08) (11).

Genetic and phylogenetic analysis

Weblogo was implemented to find the conserved sites/area in Mpro/Spike/Spike-S1 (19). The sequences of proteins and the similarities between sequences were aligned by the BioEdit software. Unrooted tree topology based on multiple alignments of amino acids was established with the neighbor-joining method in MEGA 6.06 (20). Consistency of branching was tested using a bootstrap analysis with 500 resamplings of the data in MEGA 6.06.

Structural pharmacology analysis

The I-TASSER (Iterative Threading Assembly Refinement) algorithm was utilized to predict the structures of Spike and Spike-S1 of SARS-CoV-2 (21). The RMSD (root-mean-square deviation) between two structures was computed by the Rosetta software. Physical and chemical parameters for a protein were predicted by the ProtParam tool (https://web.expasy.org/protparam/) (22).

To investigate differences in the electrostatic properties between proteins, adaptive Poisson-Boltzmann solver (APBS) and PDB2PQR were applied to each protein (http://nbcr-222.ucsd.edu/pdb2pqr_2.1.1/) (23). The pqr file of each structure was generated using the PDB2PQR program. The dx file of each structure was generated by utilizing APBS. The pqr file and dx file were then uploaded in VMD to render the molecular surface electrostatic potential map.

The largest possible binding pocket of these proteins, i.e., Mpro and Spike-S1, was predicted by Discovery Studio 3.0, respectively. These predicted pockets were utilized to construct an initial coarse model of the Mpro-ligand and Spike-S1-ACE2 complexes. Then, structures of complexes were refined by the Rosetta software (RosettaDock and FelxPepDock module), respectively (24). The final structure was obtained based on energy scores. The interactions between proteins and molecular ligands were calculated by Discovery Studio 3.0 (25). Meanwhile, interactions between proteins were computed based on distances between atoms and type of residues. Expression levels for ACE2 in human tissues were obtained from The Genotype-Tissue Expression (26). High quality 3D images of the proteins were drawn by PyMOL (27).

Results

Sequence analysis could be helpful to evaluate the repurposing of existing antiviral agents to treat SARS-CoV-2. Phylogenic trees were built (neighbor-joining, bootstrap = 500) for three proteins (Mpro, Spike, Spike-S1) based on selected 185 virus sequences obtained from CNCB/BIG (https://bigd.big.ac.cn/ncov) and NCBI blast results (**Figure 1A**, **B** and **C**, access date 7 March 2020). The results indicated that all of these three proteins of SARS-CoV-2 probably originated from bat SARS-like-CoV, and pangolin as a mammal is probably a potential intermediate host (**Figure 1A**, **B** and **C**). According to homologous analysis, Mpro sequences of SARS-CoV-2 had very high conservation (100% identify for 93 Mpro), and they were remarkably close to corresponding proteins of bat SARS-CoV and SARS-CoV (sequence similarities > 95%) (**Figure 1D**). However, Spike and Spike-S1 protein sequences of SARS-CoV-2 already displayed point mutants (SARS-CoV-2 to SARS-CoV-2 sequence similarities > 99%), and all their corresponding sequence similarity values for bat SARS-like-CoV and SARS-CoV are between 0.6 and 0.8 (**Figure 1D**, **F**). These results indicated that Mpro is much more conserved than Spike proteins in SARS-CoV-2. Compared to SARS-CoV, twelve conserved amino acids mutations, i.e. 35V, 46S, 65N, 86V, 88K, 94A, 134F, 180N, 202V, 267S, 285A, and 286L, were detected in the Mpro sequences of SARS-CoV-2 (**Figure 1E**, **Figure 2A** and **Figure 2B**). Meanwhile, six point mutations (46, 65, 86, 88, 134 and 180 sites)

were distributed in the inhibitor binding pocket of SARS-CoV Mpro (**Figure 1F**). It suggested that inhibitors of Mpro for SARS which target these sites probably could not inhibit the activity of Mpro for SARS-CoV-2.

On the other hand, Spike-S1 has a conserved domain which interacts with mammalian ACE2. Point mutation occurred in this domain may affect the interaction between Spike and ACE2, and then impact the capability of coronavirus entry into mammalian normal cells with ACE2. In 82 sequences of SARS-CoV-2 Spike-S1, there

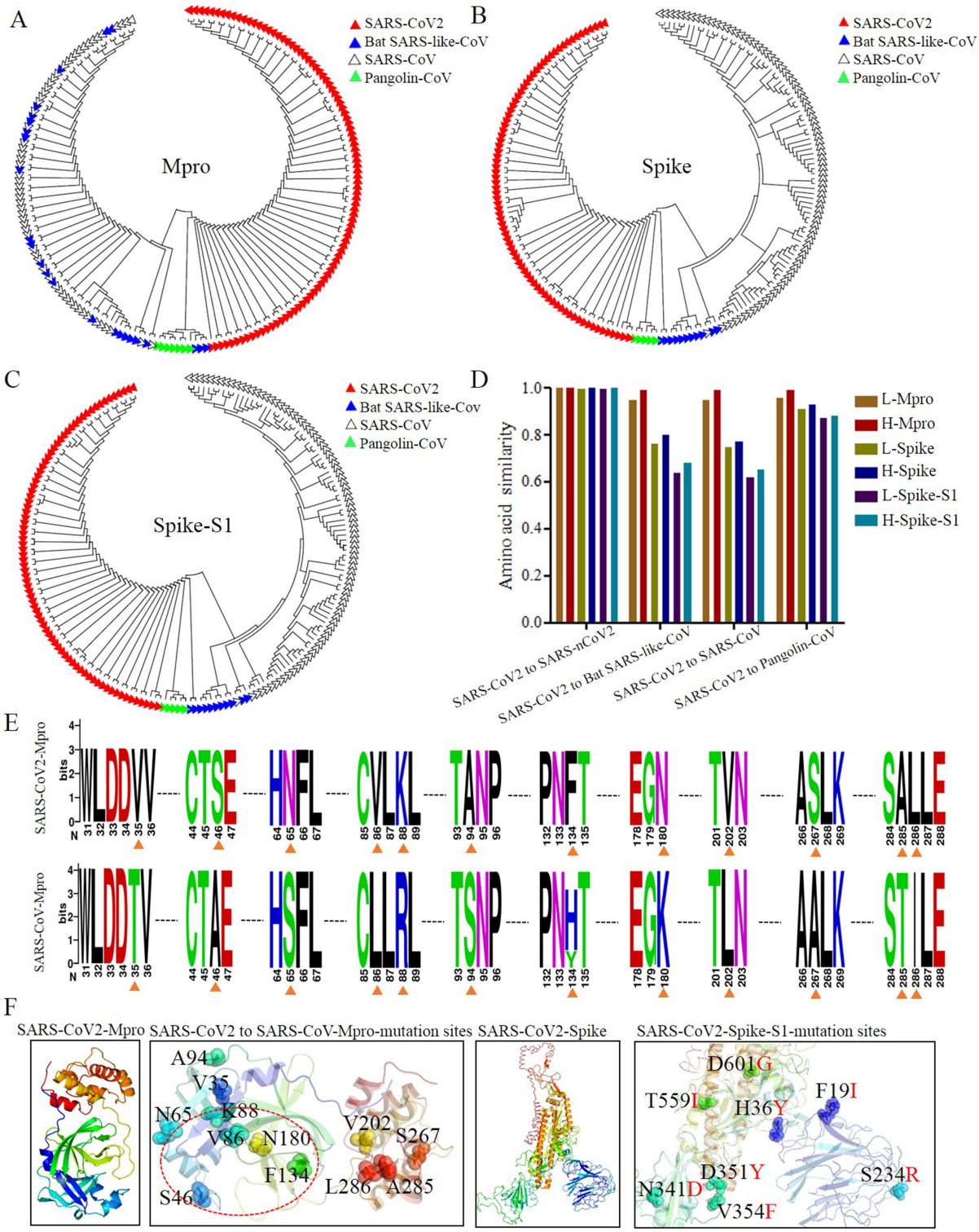


Figure 1. The evolution and conservation for Mpro and Spike of SARS-CoV-2. (A), (B) and (C) represent phylogenetic trees of Mpro, Spike, and Spike RBD. (D) Amino acid sequence similarity among SARS-CoV-2, bat SARS-like-CoV, and SARS-CoV. L- and H- represent the lowest and highest sequence similarity between sequences in two types of coronavirus, respectively. (E) Sequence logo of Mpro for SARS-CoV-2 and SARS-CoV. (F) Differential conserved amino acids of Mpro/Spike between SARS-CoV-2 and SARS-CoV.

were 11 sequences with amino acid point mutation (11/82, 13%) and up to three mutation sites for each mutation sequence (**Figure 1D**). Eight key mutation sites (F19I, H36Y, S234R, N341D, D351Y, V354F, T559I, and D601G) located in different regions (**Figure 1F**). F19I and T559I were only distributed in SARS-CoV-2 that were from Wuhan, China. N341D and D351Y were only presented in SARS-CoV-2 that were from Shenzhen, China. H36Y, S234R, V354F, and D601G only occurred in Spike-RDB domains of SARS-CoV-2 that were from Guangdong (China), Australia, France, and Germany, respectively. These results indicated SARS-CoV-2 probably underwent adaptive evolution in the human body.

It was well-known that protein sequence determines its structure, which in turn decides its biological function, such as pharmacological properties. According to sequence analysis, we tested the hypothesis that the Mpro structure of SARS-CoV-2 was very close to that of SARS-CoV. Then, Mpro of SARS-CoV-2 (PDB ID: 6LU7, https://www.rcsb.org/structure/6LU7) was utilized to compare with other Mpro proteins in the PDB database. We found that Mpro of SARS-CoV (PDB ID: 5c5o) was very close to Mpro of SARS-CoV-2 (RMSD = 0.41Å)

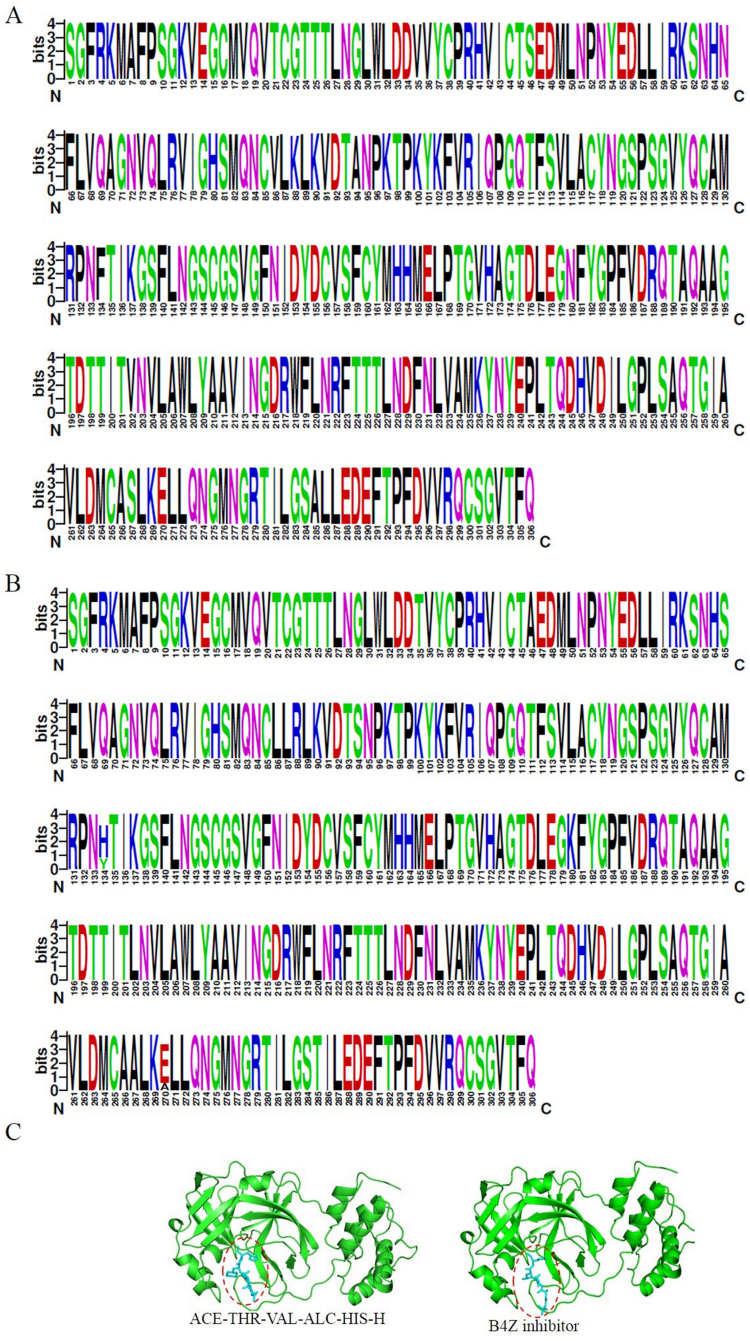


Figure 2. Sequence logos and inhibitors for Mpro in SARS-CoV-2 and SARS-CoV. (A) The sequence logo for Mpro in SARS-CoV-2; (B) The sequence logo for Mpro in SARS-CoV; (C) The complex for peptide and Mpro-SARS-CoV-2.

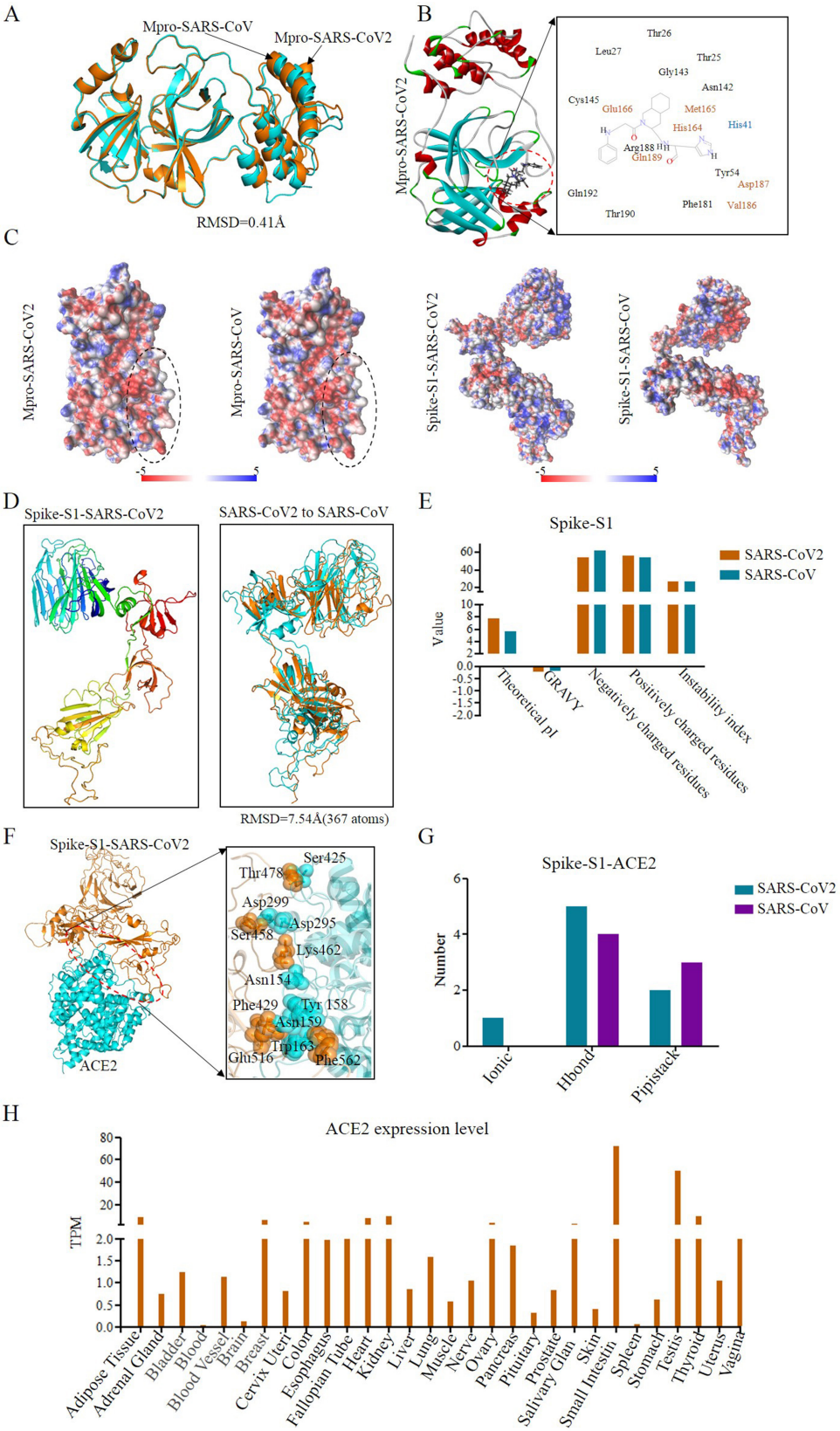


Figure 3. Structural pharmacology analysis of Mpro and Spike. (A) Structural alignment of SARS-CoV-2 Mpro and corresponding protein of SARS-CoV (PDB ID: 5c5o). (B) A ligand (SDJ) of SARS-CoV interacts with Mpro of SARS-CoV-2. Orange, cyan, and black represent h bond, Pipistack, and VDW. (C) Surface electrostatic potential values of Mpro and Spike-S1 for SARS-CoV-2 and SARS-CoV. (D) Structure comparison of SARS-CoV-2 Spike-S1 and SARS-CoV Spike-S1 (orange cartoon). (E) Physical and chemical parameters, including theoretical pI (isoelectric point), GRAVY (grand average of hydropathicity), negatively/ positively, and instability index of Spike-S1 for SARS-CoV-2 and SARS-CoV. (F) The complex of Spike-S1 for SARS-CoV-2 and ACE2, and the interaction sites between two proteins. (G) The interaction force type of Spike-S1 and ACE2 in SARS-CoV-2 and SARS-CoV. (H) ACE2 expression levels in different human organs.

(**Figure 3A**). This result indicated that the inhibitors of this protein probably also inhibit the activity of Mpro of SARS-CoV-2. After screening, a unique ligand for 5c5o, i.e. SDJ (phenyl-β-alanyl (S, R)-N-decalin type inhibitor: (2S)-3-(1H-imidazol-5-yl)–2-({[(3S, 4aR, 8aS)-2-(N-phenyl-β-alanyl) decahydroisoquinolin-3-yl] methyl} amino) propanal, was utilized to dock with Mpro of SARS-CoV-2 (**Figure 3B**). The interaction sites between the inhibitor and Mpro of SARS-CoV-2 were almost the same (**Figure 3B**, PDB ID: 5c5o) (9), and there is no conserved amino acid mutations for Mpro. Hence, this inhibitor of SARS-CoV Mpro could act as an effective inhibitor for Mpro of SARS-CoV-2. Furthermore, based on other proteins that were close to Mpro of SARS-CoV-2, we also found that peptide ACE-THR-VAL-ALC-HIS-H [Biologically Interesting Molecule Reference Dictionary (BIRD), ID: PRD_000815] and B4Z inhibitor (BIRD ID: PRD_000910) could probably inhibited SARS-CoV-2 infection (**Figure 2C**). Meanwhile, electrostatic potential values for the surface of Mpro proteins of SARS-CoV and SARS-CoV-2 were almost the same as well (**Figure 3C**). These results indicated that inhibitors (molecules and peptides, such as, Beclabuvir, Saquinavir, and Lopinavir) for Mpro of SARS-CoV probably could be utilized to inhibit the activity of Mpro in SARS-CoV-2, if these inhibitors do not target conserved mutation amino acids of Mpro SARS-CoV-2.

For drug design based on the interaction between Spike-S1 and ACE2, we obtained protein structures of Spike and Spike-S1 by utilizing the I-TASSER algorithm. Compared to SARS-CoV, five factors were different from Spike-S1 proteins of SARS-CoV-2: 1) surface electrostatic potential values (**Figure 3C**); 2) 3D spatial structure (RMSD = 7.54 Å, aligned based on 367 atoms) (**Figure 3D**, PDB ID: 5i08); 3) physical and chemical parameters (theoretical pI, GRAVY, negatively/ positively, and instability index) (**Figure 3E**); 4) interaction sites between Spike-S1 and ACE2 (**Figure 3F**, PDB ID: 5i08) (11); 5) interaction force type between Spike-S1 and ACE2 (**Figure 3G**, PDB ID: 5i08). These factors play important roles in designing inhibitors for ACE2. Therefore, some inhibitors of ACE2, effectively inhibit the entry of SARS-CoV into normal cells, might not perform well for SARS-CoV-2. However, peptides for SARS-CoV-2 could be designed based on predicted structure of the Spike-S1-ACE2 complex. For example, such as a predicted peptide, FRKSNLKPFERDISTEIYQAGSTPC, could interact with ACE2 with one ionic bond and four H bonds according to predicted Spike-S1-ACE2 complex (**Figure 3F**). Meanwhile, according to ACE2 expression level in human tissues (**Figure 3H**), we should pay more attention to the functional changes of intestinal tract, testis, liver, lung and kidney, during clinical diagnosis and treatment for SARS-CoV-2 patients.

Discussion

Our results showed that both Mpro and Spike of SARS-CoV-2 were homologous with those of bat SARS-like-CoV. Although protein sequence and structure for SARS-CoV-2 Mpro were conserved, Spike of SARS-CoV-2 had many mutations and some mutant sites were distributed in the Spike-S1. Based on sequence and structural pharmacology analysis, we found that some Mpro inhibitors for SARS-CoV probably also inhibit activity of Mpro if they do not target conserved mutated sites for Mpro of SARS-CoV-2 (**Figure 1**). In addition, we found that SARS-CoV-2 probably underwent adaptive evolution when the virus spread from Wuhan patients to other non-Wuhan patients, which could be helpful for discovering potential drugs for the treatment of NCP cases. Furthermore, a potential peptide, that was predicted based on interactions between Spike-S1 and ACE2, could probably serve as a potential drug, and further study can be conducted for its function test and peptide modification.

Conflict of interest

The authors declare that they have no competing interests.

Acknowledgement

Authors acknowledge the contributions for Mpro structure of Zihe Rao group.

References

1. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A novel coronavirus from patients with pneumonia in China, 2019. N Engl J Med. 2020 Feb 20;382(8):727-733.
2. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet. 2020 Feb 15;395(10223):497-506.
3. Guan W-j, Ni Z-y, Hu Y, Liang W-h, Ou C-q, He J-x, et al. Clinical characteristics of 2019 novel coronavirus infection in China. medRxiv. 2020:2020.02.06.20020974.
4. Wang D, Hu B, Hu C, Zhu F, Liu X, Zhang J, et al. Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan, China. JAMA. 2020 Feb 7. [Epub ahead of print]
5. Jin YH, Cai L, Cheng ZS, Cheng H, Deng T, Fan YP, et al. A rapid advice guideline for the diagnosis and treatment of 2019 novel coronavirus (2019-nCoV) infected pneumonia (standard version). Mil Med Res. 2020;7(1):4.
6. Zhang Z, Li X, Zhang W, Shi ZL, Zheng Z, Wang T. Clinical features and treatment of 2019-nCov pneumonia patients in Wuhan: report of a couple cases. Virol Sin. 2020 Feb 7. [Epub ahead of print]
7. Wu A, Peng Y, Huang B, Ding X, Wang X, Niu P, et al. Genome composition and divergence of the novel coronavirus (2019-nCoV) originating in China. Cell Host Microbe. 2020 Mar 11;27(3):325-328.
8. Woo PC, Lau SK, Chu CM, Chan KH, Tsoi HW, Huang Y, et al. Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. J Virol. 2005 Jan; 79(2): 884-895.
9. Shimamoto Y, Hattori Y, Kobayashi K, Teruya K, Sanjoh A, Nakagawa A, et al. Fused-ring structure of decahydroisoquinolin as a novel scaffold for SARS 3CL protease inhibitors. Bioorg Med Chem. 2015 Feb 15; 23(4):876-90.
10. Wong, K. S. A 193-amino acid fragment of the SARS coronavirus S protein efficiently binds angiotensin-converting enzyme 2. J Biol Chem. 2004 Jan 30;279(5):3197-201.
11. Kirchdoerfer RN, Cottrell CA, Wang N, Pallesen J, Yassine HM,

Turner HL, et al. Pre-fusion structure of a human coronavirus spike protein. Nature. 2016 Mar 3;531(7592):118-21.
12. Wang S, Guo F, Liu K, Wang H, Rao S, Yang P, et al. Endocytosis of the receptor-binding domain of SARS-CoV spike protein together with virus receptor ACE2. Virus Res. 2008 Sep;136(1-2):8-15.
13. Richardson P, Griffin I, Tucker C, Smith D, Oechsle O, Phelan A, et al. Baricitinib as potential treatment for 2019-nCoV acute respiratory disease. Lancet. 2020 Feb 15;395(10223):e30-e31.
14. Liu W, Morse JS, Lalonde T, Xu S. Learning from the past: possible urgent prevention and treatment options for severe acute respiratory infections caused by 2019-nCoV. Chembiochem. 2020 Mar 2;21(5):730-738.
15. Wang M, Cao R, Zhang L, Yang X, Liu J, Xu M, et al. Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro. Cell Res. 2020 Mar;30(3):269-271.
16. Helge-Friedrich T. Analysis for free: Comparing programs for sequence analysis. Brief Bioinform. 2004 Mar;5(1):82-7.
17. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool (BLAST). J Mol Biol. 1990 Oct 5;215(3):403-10.
18. Sussman JL, Lin D, Jiang J, Manning NO, Prilusky J, Ritter O, et al. Protein data bank (PDB): database of three-dimensional structural information of biological macromolecules . Acta Crystallographica. 2010;54(6-1):1078-84.
19. Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: a sequence logo generator. Genome Res. 2004 Jun; 14(6):1188–1190.
20. Lewis PO, Kumar S, Tamura K, Nei M. MEGA: molecular evolutionary genetics analysis, version 1.02. Systematic Biology. 1995;44(4).
21. Roy A, Kucukural A, Zhang Y. I-TASSER: a unified platform for automated protein structure and function prediction. Nat Protoc. 2010 Apr;5(4):725–738.
22. Garg VK, Avashthi H, Tiwari A. MFPPI – multi FASTA ProtParam interface. Bioinformation. 2016;12(2): 4–77.
23. Dolinsky TJ, Nielsen JE, Andrew MJ, Baker NA. PDB2PQR: an automated pipeline for the setup of Poisson–Boltzmann electrostatics calculations. Nucleic Acids Res. 2004 Jul 1; 32(Web Server issue):W665–W667.
24. Rohl CA, Strauss CE, Misura KM, Baker D. Protein structure prediction using Rosetta. Method Enzymol. 2004;383:66-93.
25. Gao YD, Huang JF. An extension strategy of Discovery Studio 2.0 for non-bonded interaction energy automatic calculation at the residue level. Dongwuxue Yanjiu. 2011 Jun;32(3):262-6. [Article in Chinese]
26. Lonsdale J, Thomas J, Salvatore M, Phillips R, Lo E, Shad S, et al. The genotype-tissue expression (GTEx) project. Nat Genet. 2013 Jun;45(6):580–585.
27. Ordog R. PyDeT, a PyMOL plug-in for visualizing geometric concepts around proteins. Bioinformation. 2008;2(8):346–347.

RESEARCH ARTICLE

Effects of abiotic elicitors on the production of bioactive flavonols in *Emilia sonchifolia*

Mohan RAJI, Zhong CHEN*

Natural Sciences and Science Education, National Institute of Education, Nanyang Technological University, Singapore 637616, Singapore

*Correspondence: zhong.chen@nie.edu.sg
https://doi.org/10.37175/stemedicine.v1i2.33

ABSTRACT

Background: The usage of plants as a source of medicine begins with the isolation of active bio-compounds which are responsible for their therapeutic action. Research on flavonols has flourished in the last decade, fuelled by the recognized importance of natural antioxidants on health.

Methods: Based on the meagre information reported on the effect of stress treatment on the production of flavonol metabolites, the current research attempted to investigate if heat (one-day for 6 hours at 45 °C) and drought stress (one week) could be used to enhance the production of three flavonols, including rutin, quercetin and kaempferol, in the tropical medicinal plant *Emilia sonchifolia*. The study focused on a novel elicitation approach to improve the therapeutic value of medicinal plants.

Results: The increment in the antioxidant levels of *Emilia sonchifolia* was attributed to the increased accumulation of the three flavonols and total flavonoid content with strong scavenging ability against free radicals.

Conclusion: Short-term heat stress and moderate water deficit proved as effective strategies to increase the rutin, quercetin and kaempferol contents in *Emilia sonchifolia*.

Keywords: Lilac tassel flower · Antioxidant activity · Rutin · Quercetin · Kaempferol

1. Introduction

Tropical countries are home to a variety of different medicinally important plants, due to the presence of bioactive compounds which are secondary plant metabolites. The latest interest in secondary plant metabolites lies in obtaining new sources of compounds that have promising potential to function in therapeutic systems (1, 2). Medicinal plants are usually assessed based on their therapeutic value. The therapeutic effects of these metabolites have been associated with their antioxidant activity, which could be the scientific basis of traditional herbal-based treatment of diseases (3). The isolation and characterization of drugs are based on the pharmacologically active plant secondary metabolites (4). Root, bark, stem and leaves are the primary sources of different and unique secondary metabolites in plants, of which the most widely distributed bioactive groups are flavonoids that exist both in free state and as glycosides (5). Flavonols belong to the most abundant group of flavonoids, with most prominent antioxidant activity by possessing the chemical structure of 3-hydroxyflavone skeleton (6). The presence of flavonols in the form of fruits (quercetin in blueberries, strawberries), leafy green vegetables, onions and beverages (tea and red wine) is inversely associated with the incidence of breast cancer, diabetes, heart disease and neurodegenerative diseases including Alzheimer's. Moreover, medicinal plants such as *Moringa oleifera*, *Azadirachta indica*, *Aloe barbadensis*, *Ginkgo biloba* and *Calendula officinalis* are reported to contain higher level of flavonols (7). Flavonols such as rutin, quercetin and kaempferol (**Figure 1**) are unique reactive oxygen species (ROS) scavengers with antiplatelet, anti-neural, cancer preventive, antioxidant and anti-inflammatory properties (7, 8). Consequently, they possess beneficial pharmaceutical, nutritional and therapeutic effects in humans. Rutin is usually recommended for treatment of haemorrhoids or internal bleeding, varicose veins and reducing the risk of arteriosclerosis (9). Quercetin could contribute significantly

Received: Dec 30, 2019; Accepted: Mar 8, 2020.

to the antioxidant defenses present in the blood and plasma as it is only slowly eliminated in humans (10). It is a lipid peroxidation inhibitor and reduces oxidative stress produced as a result of chronic diseases. Among the three flavonols, quercetin exhibits a wide range of biological activities and therapeutic applications (11, 12). Kaempferol is responsible for decreasing thrombin that prevents clot formation and inhibits blood clotting (13). Due to its anti-thrombotic property, it has the potential to serve as a therapeutic agent for cardiovascular patients. In plants, flavonol metabolism dates back during the evolution of terrestrial plants mostly in mosses and liverworts, and it has been reported to have relevant function in the responsive mechanism of plants under stressful conditions (14). They respond to abiotic stresses by increasing their production and detoxifying the free radicals through donation of hydrogen atom from the hydroxyl moiety thereby rendering tolerance (15).

Indeed, the exploitation of more medicinal plants is demanding due to the health hazards associated with the use of synthetic drugs. Even though the potential of medicinal plants are boundless, there are many challenges for the large scale utilization in pharmaceutics (16). Among them are the lack of reproducibility due to the variation in secondary metabolites and the limited production of metabolites (less than 1% dry weight). In medicinal plants, environmental conditions regulate the production of these complex active constituents, and their type and quantity of metabolites are often determined by the changing environmental conditions (17). When wild plants are procured and cultivated under optimal growing conditions without exposure to natural stressors, the content and therapeutic activity of secondary metabolites are reduced (18). The production of secondary metabolites should be increased to acquire maximum medicinal and nutritive values. Heat and drought are abiotic elicitors that stimulate stress responses in plants and increase the amount of secondary metabolites produced (19). This strategy helps to induce the desired chemical response by establishing a relationship between phytochemistry and plant stress. This technique is known as elicitation which not only increases the therapeutic activity of medicinal plants but also enhances the health benefits of edible plants that are consumed as food.

Comparatively, little information has been gained in recent years regarding the methods to improve the therapeutic activity in whole medicinal plants. There is also not much updated knowledge on the antioxidant metabolism under heat and drought stress in medicinal plants. It seems necessary to study the connection between medicinal plants and abiotic stresses. Therefore, we aimed to test the medicinal plant *Emilia sonchifolia* (*E. sonchifolia*) for its antioxidant activity and study its abiotic stress response on flavonol metabolite production through the elicitation process. Under thermal stress flavonols could be either increased, unaffected or decreased depending upon the degree/intensity, duration of the temperature, flavonoid structure and finally the plant species. The accumulation of flavonols under water deficit conditions could be attributed to the severity of drought, growth stage of the plant and the stress tolerant level inherent in each species. To date, there hasn't been sufficient data to conclude on the impact of temperature or drought on flavonols. The present work also aimed to understand the adaptability of flavonols in response to environmental stresses and find out effective methods for their natural production. The choice of investigated medicinal plant depends on four reasons: firstly, availability as it is a locally available and easily grown weed plant. Secondly, it is a good source of flavonols (20) and widely used in traditional medicine. Thirdly, it is tolerant to abiotic stresses like heat and drought. Fourthly, no elicitation studies have been reported in this plant species.

E. sonchifolia belongs to family Asteraceae and is known by the names 'Cupid's shaving brush' and 'Lilac tassel flower'. It is listed in the Indian system of Ayurveda and traditional Chinese medicine (TCM) (21). It is widely present in tropical regions of India, China and Southeast Asia as weeds in grassy fields, crop fields and along roadsides. The leaves of *E. sonchifolia* proficiently treat night blindness, wounds and abscesses, roundworm infestations, snake bites, burns, cough, sore ears, fever and bronchitis, while the roots are used for treating diarrhoea (22). Various studies also have documented the pharmacological activities of *E. sonchifolia* extracts against microbials, inflammation (23), cataract (24), oxidation (25), diabetes (26) and virus (27).

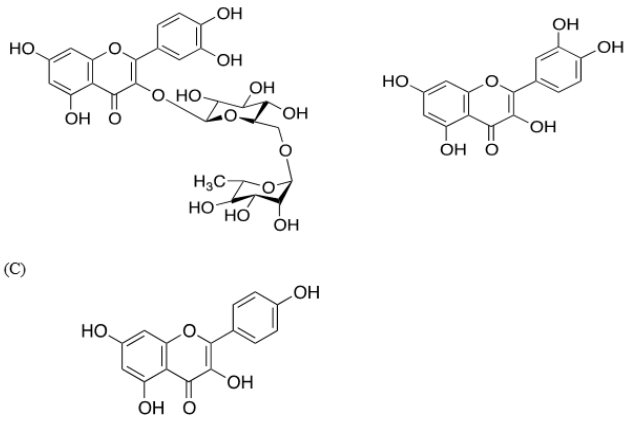


Figure 1: Chemical structures of (A) rutin, (B) quercetin, and (C) kaempferol.

2. Materials and Methods

2.1 Plant Materials

The tropical medicinal plant *E. sonchifolia* (3-month-old) was used for the study. The seeds of *E. sonchifolia* were obtained from Molecular Genetics Lab at National Institute of Education (NIE), Singapore. The seedlings of uniform size were transplanted to black polythene bags (21×12 cm) containing a mixture of peat moss and vermiculite (3:1). The plants were grown under natural light conditions of photosynthetic photon flux density (PPFD) 300 µmol $m^{-2}s^{-1}$ in the NIE greenhouse, watered once daily for three months.

2.2 Experimental conditions

The experiment was conducted at NIE. Plants grown in optimal temperature and watered regularly were used as reference or control for each treatment. For the heat treatment, five replicates of the plant were transferred to an incubator with white fluorescent light of PPFD 50 µmol $m^{-2}s^{-1}$ and exposed to heat stress for 6 hours continuously in a single day under temperature stress set at 45 °C. For the drought treatment, another batch of five replicates of the plant was subjected to a continuous water deficit for 7 days. The drought stress experiment was conducted in the NIE greenhouse. The leaf samples for analysis were harvested at time 0 (control) and at the end of respective treatments.

2.3 Determination of relative water content (RWC) in drought-stressed plants

The leaves of *E. sonchifolia* were harvested and immediately transferred to a sealed plastic bag to avoid moisture loss. The sample leaves were cut into small pieces and immediately weighed with an analytical balance to record the fresh weight (FW). The weighed leaves were then submerged in water for 24 hours in the dark to record their saturated weight (SW). Subsequently, the samples were wrapped in aluminum foil and dried in an oven at 80 °C for 4 days. The dry weight (DW) of the leaves was recorded. The RWC of samples was calculated using the formula: RWC = (FW-DW) / (SW-DW) × 100% (28).

2.4 DAB staining for the qualitative detection of ROS (H_2O_2) under the influence of heat and drought stress treatments

The DAB staining solution of concentration 1 mg/ml was prepared freshly in a 50 ml falcon tube according to the standard protocol (29). Harvested leaves of *E. sonchifolia* were placed inside the petri plates and covered with aluminium foil as DAB is light sensitive. 1 mg/ml DAB staining solution was added to keep the leaves fully immersed. They were then vacuum infiltrated for 15 minutes. The treated leaves were incubated inside a standard laboratory shaker for 8 hours until the leaves were thoroughly infiltrated with the staining solution at 100 rpm. Following the incubation, the DAB solution was replaced with 80% (v/v) ethanol. The plates were carefully placed in 80 °C water bath for 15 minutes to remove chlorophyll. DAB-stained leaves were visualized as brown colour. It was then stored in a solution of ethanol:acetic acid:glycerol in a ratio of 3:1:1 until photographed.

2.5 Preparation of Extracts

The air-dried leaves were subjected to freeze-drying or lyophilization to remove the moisture content. The thoroughly dried leaves were then ground into a fine powder using a pestle and mortar. Using 5 ml of extracting solvent 80% (v/v) methanol-water, approximately 200 mg of leaf powder was extracted using an ultrasonic cleaner (37 kHz, 150W, S70 Elmasonic Germany) for 30 minutes at room temperature. Following the extraction, the mixture was centrifuged at 3500 rpm, 25 °C for 5 minutes. The supernatant was recovered and filtered using 0.45 µm syringe filters, followed by solid phase extraction (SPE) using the Sep-Pak C18 cartridge. To start the process, the cartridge was saturated with methanol. Around 400 µl of the plant sample was then added into the cartridge with the help of vacuum pump, followed by eluting with 50 ml of methanol solvent and then rotary evaporated. The residue was then re-dissolved in 1 ml LC/MS grade methanol, before being injected to the column, Following the modified method from Neves, Stringheta (30).

2.6 LC-MS analysis

The detection of three flavonol compounds was performed using an Agilent 1100 series (Agilent Technologies, CA, USA) liquid chromatography coupled to ion trap mass spectrometer (LC/MSD Trap XCT) with ESI interface. The mobile phase was degassed ultrasonically following a gradient program using solvent A: 0.1% formic acid (FA) in water and solvent B: 0.1% FA in methanol. The flow rate used was 0.7 ml/min while 15 µl of sample was injected into the Phenomenex Kinetex C18 reverse phase column (ø 100 mm × 4.6 mm × 2.6 µm). The column temperature was maintained at 30 °C. The samples were passed through a 0.45 µm syringe filter before being injected into a column. The system was run with the following elution program: 18% to 95% B for 31 minutes, then hold at 95% B for 3 minutes before re-conditioning back to the initial composition in 1 minute. Mass spectra of analyses were performed with the scan range of *m/z* 200 - 650 in a negative ion mode. The mass spectrometric parameters including nebulizing gas pressure, capillary voltage, dry gas temperature and flow rate were set to 40 psi, 4.5 kV, 350 °C and 10 L/min, and kept constant throughout the analysis. LC/MS Chemstation software was used for data analysis. The detection of the eluted peaks of rutin, quercetin and kaempferol was performed at 360 nm. The retention time obtained based on triplicates for rutin, quercetin and kaempferol were 27.2, 29.2 and 30.2 minutes, respectively.

2.7 Quantification of total flavonoids using spectrophotometry

The total flavonoid content (TFC) was determined spectrophotometrically following the aluminium

chloride ($AlCl_3$) method (31). The flavonol standard rutin (1 mg/ml) prepared in 80% methanol was used to generate a standard curve with four concentrations of 0.0125, 0.025, 0.05, and 0.1 mg/ml, respectively. 0.5 ml of dried plant extract from SPE was mixed with 1.5 ml of 80% methanol, 0.1 ml of 10% aqueous dilution of $AlCl_3$, 0.1 mL of 1 M potassium acetate and finally with 2.8 ml of distilled water. The mixed contents in each tube were allowed to incubate at 22 °C for 30 minutes. The blank was prepared for each sample which consisted of all reagents except $AlCl_3$. The absorbance was measured at 415 nm. The total content of flavonoids in plant extracts in rutin equivalent was calculated by the following equation: C = (c × v) / m. C = Total content of flavonoid compounds, mgg^{-1} DW, in rutin equivalent (RE), c = Concentration of rutin established from the calibration curve, mgg^{-1} DW, v = Volume of extract, ml, m = Mass of extract, gm.

2.8 Measurement of antioxidant activity using DPPH method

Different concentrations of leaf extracts (1 mg/ml) of *E. sonchifolia* were prepared in methanol equivalent to 50, 100, 200 and 400 ppm. 80% methanol was used as a solvent. Ascorbic acid was used as the positive control. A stock solution of 0.06 mM methanolic solution of DPPH was prepared by weighing 5.91 mg of DPPH powder in an Eppendorf tube containing 1 ml of methanol. The tube was centrifuged at 3000 rpm and then transferred to 250 ml volumetric flask, with the volume of methanol adjusted accordingly. 5 ml of 0.06 mM solution of DPPH was mixed with 0.5 ml of plant samples and standard solutions, separately. These solution mixtures were kept in a dark place for 30 minutes, and the optical density was measured at 517 nm. The blank control was prepared for each sample which consisted of all reagents except the tested extracts. Each analysis was done in triplicate. The radical scavenging of the methanolic extract of *E. sonchifolia* leaves was calculated from the formula: Percentage radical-scavenging activity = (OD control - OD sample / OD control) × 100. OD control = Absorbance of the control, OD sample = Absorbance of the tested plant extract.

2.9 Statistical analysis

ANOVA (one-way analysis of variance) was performed to test for the significant difference between optimal and the treated conditions. The Dunnett's test was also used for multiple comparisons and to determine differences among treatment means at significance levels of $P < 0.05$. The Pearson's correlation coefficient (r) was calculated (between +1 and -1) using the SPSS software (version 16; SPSS Inc, Chicago, USA) that measures the strength of linear correlation between two variables. The Origin Pro software was used to plot the graph. The Graphpad Prism 8 was used to plot the graph for antioxidant studies.

3. Results and Discussion

3.1 RWC in drought-stressed plant leaves

The plant *E. sonchifolia* at optimal conditions recorded leaf RWC value of 97.5%. At the end of water deficit stress, RWC was reduced to 80% (**Figure 2**). The RWC of *E. sonchifolia* tended to decrease slightly with increased water stress level. However, the percentage drop of 18% suggested that the stressed leaves of the plant experienced a slow rate of water loss that developed acclimation to drought stress. The results showed that RWC was an indicator of water deficit stress, consistent with previous observation in *Stellaria dichotoma* (32).

3.2 DAB staining for ROS (H_2O_2) under heat and drought stress treatments

The DAB-stained leaves of *E. sonchifolia* were visually detected (**Figure 3**). The brown colour developed on the leaves directly correlated with H_2O_2 accumulation, and it was evident that the leaves of stressed plants appeared more brownish and hence, accumulated more H_2O_2 than the leaves of unstressed plants. The appearance of dark spots was more visible in heat-stressed (**Figure 3C**) and drought-stressed (**Figure 3B**) plants, with no dark spots accumulated in control plants. The pattern observed in DAB staining suggested that the stressed plants accumulated more ROS. The data indicated that stress developed in all treated plants at the time of biochemical evaluations, and the visibility of ROS depended on the stress tolerance of the species.

3.3 Detection of flavonols rutin, quercetin and kaempferol using LC-MS

Using LC-MS method, the three flavonol compounds in E. sonchifolia were confirmed by comparing their respective characteristic fragmentation patterns with the corresponding flavonol standards at a specific retention time (**Figures 4-5**). The distinctive fragmentationpattern of each flavonol standard was studied using both alternating positive $[M+H]^+$ and negative $[M-H]^-$ ion ESI-MS. The negative ion mode was selected for the detection

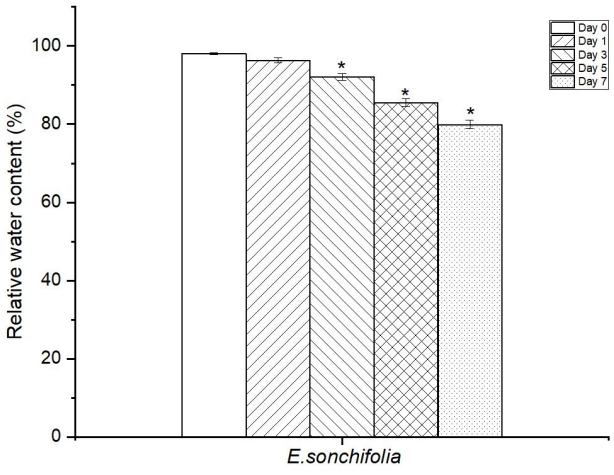


Figure 2. RWC of E. sonchifolia leaves during drought stress. The height of each bar and the error bar show the mean and standard deviation from 5 replicates. The column with single asterisk (*) adhered indicates significant differences ($P < 0.05$) as compared with day 0, determined by Dunnett's test.

of three flavonols as these compounds were more stable as deprotonated adducts. Furthermore, the deprotonated ions (precursor ions) of rutin, quercetin and kaempferol were detected from the methanolic extracts of leaves of each plant species that were subjected to optimal and stressed conditions (via heat and drought) at retention time 27.2, 29.2, and, 30.2, respectively. The deprotonated ions detected for rutin, quercetin and kaempferol were *m/z* 609.2 $[M-H]^-$, *m/z* 301 $[M-H]^-$, *m/z* 285 $[M-H]^-$, respectively (33). The *m/z* of each compound was compared with previous data analysis, including rutin (34, 35), quercetin (35, 36) and kaempferol (36, 37). The MS analyses of rutin, quercetin and kaempferol in *E. sonchifolia* were represented in **Figures 6-8**. The data of ESI-MS analysis were summarized in **Table 1**.

The LC-MS analysis was carried out in E. sonchifolia to evaluate the flavonol contents under normal and stressed conditions (heat and drought) and to validate the influence of stress level on the bioactive metabolites. In non-treated *E. sonchifolia* plants, the rutin content was observed in higher intensity than quercetin and kaempferol (**Table 1**). In heat-stressed and water deficit *E. sonchifolia* plants, the relative intensity of rutin was higher by 5.3-fold and 4-fold, respectively than control. Furthermore, the relative intensity of quercetin in heat-stressed ones was higher by 4-fold. Similarly, an increasing trend of quercetin was followed in drought-stressed plants with 3.3-fold higher than control. The relative intensity of kaempferol in heat, drought treated plants was higher by 3.4-fold and 2-fold, respectively than control (**Table 1**).

3.4 Total flavonoid content (TFC)

The TFC of *E. sonchifolia* plants under heat, drought and optimal conditions was shown in **Table 2**. Among the stressed and optimal *E. sonchifolia* plants, highest TFC was detected in heat-stressed samples (8.66 ± 0.73 mgREg-1DW), followed by drought-stressed

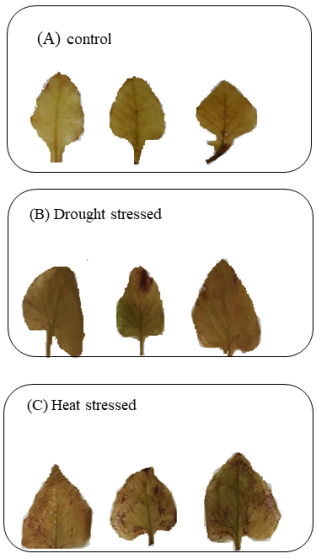


Figure 3. DAB-stained leaves of *E. sonchifolia* (3-month old). (A) control plants, (B) drought-stressed plants (1 week), and (C) heat-stressed plants (45 °C for 6 hours in a day).

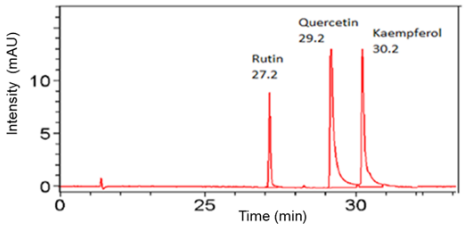
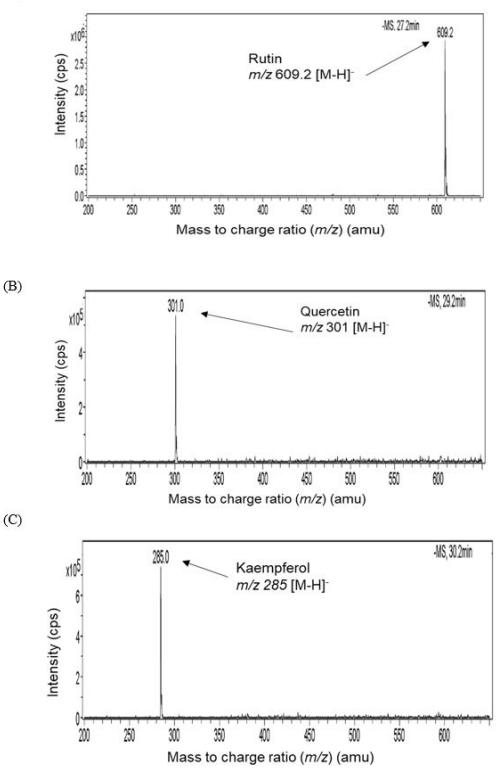


Figure 4. HPLC-UV chromatogram of flavonol standards rutin, quercetin, and kaempferol (10 ppm) with the respective retention times detected at 360 nm.



Figures 5 (A-C) Indicates the ESI-MS analysis of the mass spectra fragmentation patterns of the flavonol standards rutin, quercetin, and kaempferol with x-axis of mass to charge ratio (*m/z*) (amu) and y-axis indicated intensity or the ion abundance (cps). The deprotonated ions detected for rutin, quercetin and kaempferol were *m/z* 609.2 $[M-H]^-$, *m/z* 301 $[M-H]^-$, *m/z* 285 $[M-H]^-$, respectively at respective retention times.

(6.01 ± 0.85 mgREg-1DW) and then control (3.52 ± 0.84 mgREg-1DW). The data indicated that heat and drought stress enhanced the TFC in all treated plants (per gm dry weight of leaf material) compared to control plants. A notable fold change of TFC was observed in heat-stressed *E. sonchifolia* (1.5-fold increase). An investigation revealed that apart from the increase in flavonols, drought stress also stimulated the production of other flavonoids and hence influenced the TFC (38). Water stress is known to augment the TFC in *Chrysanthemum morifolium* as compared to normal (39). Xiaolu, Jie showed an increase in TFC in stems of *Dendrobium moniliforme* after 5 days of drought stress (40).

In conclusion, quercetin, rutin and kaempferol were

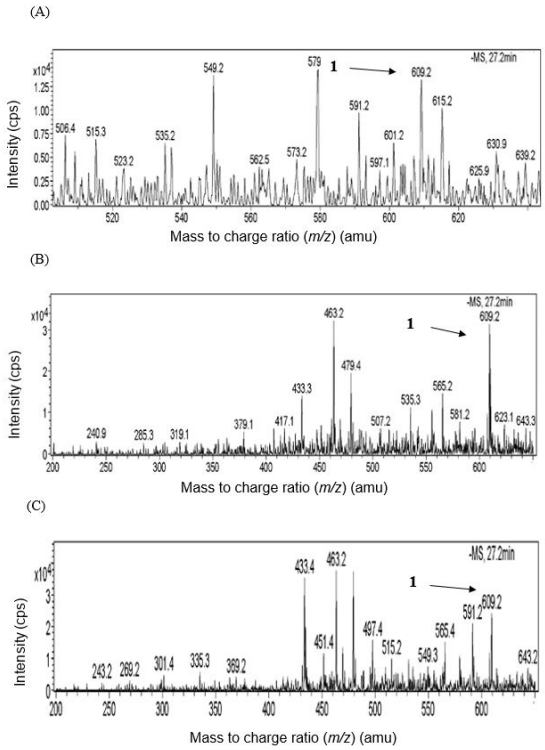


Figure 6. ESI-MS analysis of rutin in *E. sonchifolia*. 1. rutin, with *m/z* 609.2 $[M-H]^-$, indicated above at retention time 27.2. (A) control plants, (B) heat-stressed plants, and (C) drought-stressed plants.

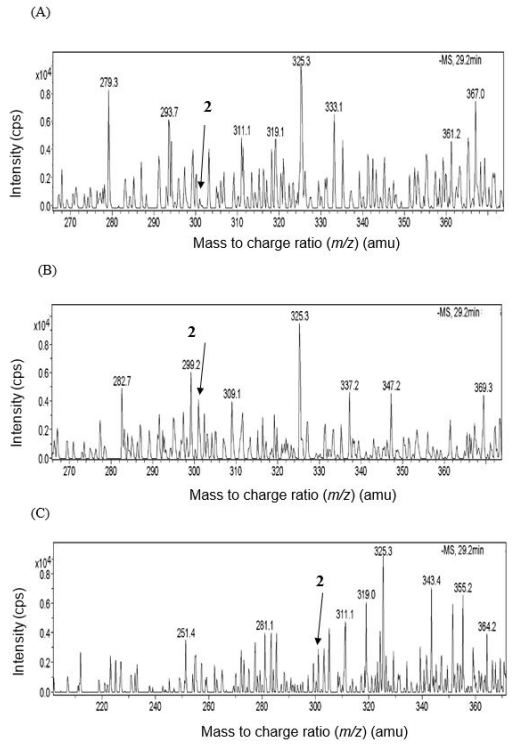


Figure 7. ESI-MS analysis of quercetin in E. sonchifolia. 2. quercetin, with *m/z* 301 $[M-H]^-$, indicated above at retention time 29.2. (A) control plants, (B) heat-stressed plants, and (C) drought-stressed plants.

found to increase in stressed *E. sonchifolia* plants compared to the control ones, and the abiotic stresses such as drought and heat had a significant impact on the three flavonol metabolites as evident from their increasing trends. The variation in metabolite contents is due to the growth conditions, the difference in genotypes and the plant stress tolerance (41). In plants grown under optimal growing conditions, the intensity of natural stressors on the plant is lower compared to the wild plants, and the content of secondary metabolites is therefore reduced, leading to decreased therapeutic activity (18). In the study, optimal *E. sonchifolia* plants grown under greenhouse condition had lesser TFC than stressed plants (**Table 2**). As stated by Agati, Azzarello, *A. thaliana* grown under low light irradiance (less than 200 µmol $m^{-2}s^{-1}$) was unable to stimulate the biosynthesis of quercetin, unless triggered by temperature and high light (42). In *E. sonchifolia* plants, rutin and quercetin increased under water deficit conditions (**Table 1**). The result was in agreement with the report of Nacif de Abreu and Mazzafera (43). Another study by Hodaei, Rahimmalek demonstrated that drought stress for 3 days stimulated the synthesis of rutin and quercetin in "Taraneh" cultivar of *Chrysanthemum morifoilum* plant (39). However, prolonged drought stress for 7 days was accompanied by a decrease in quercetin content while an increase in rutin. In another study of drought stress in two related species of hawthorn, Kirakosyan, Kaufman revealed that quercetin increased in *Crataegus monogyna* and decreased in *Crataegus laevigata* whereas rutin increased in both species (38). Rutin is a glycoside that is composed of aglycone quercetin and disaccharide rutinose (**Figure 1A**). Although rutin is derived from quercetin with the inclusion of rhamnoside residue in the third position,there are variations among the compounds even in related

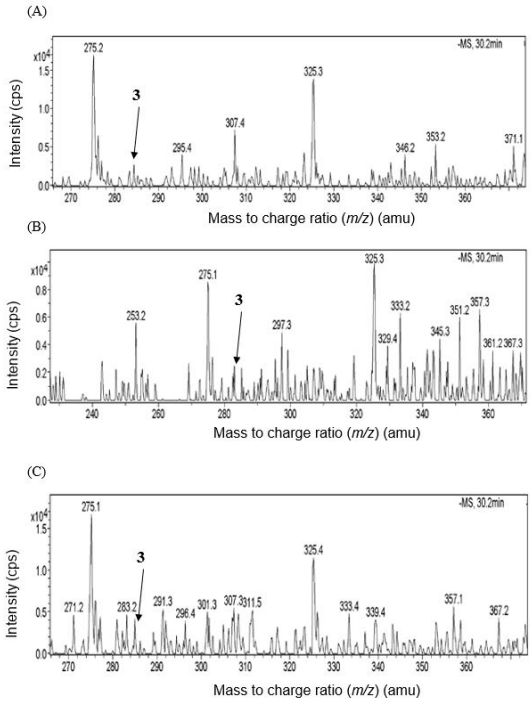


Figure 8. ESI-MS analysis of kaempferol in *E. sonchifolia*. 3. kaempferol, with *m/z* 285 $[M-H]^-$, indicated above at retention time 30.2. (A) control plants, (B) heat-stressed plants, and (C) drought-stressed plants.

species. Quercetin is O-glycosylated by rhamnose units at C_3 position to form quercetin-3-O-rutinoside (**Figure 1A** and **Figure 1B**). It is clear from the studies that flavonols are mostly glycosylated in the majority of tissues to favour their storage and transport and even to protect the OH group from oxidation (44). In the study, it is evident that during heat and drought stress, rutin was dominated likely due to the induced rhamnosylation process that resulted in the conjugation of rhamnose to aglycone quercetin forming rutin to attain stability and tolerance against stress conditions. According to Le Roy, Huss, the pile-up of flavonol glycosides acted as a reserve of flavonols that could be utilized at any time under stress conditions like UV (44). Hectors, Van Oevelen emphasized on the role of rhamnosylated quercetin and kaempferol glycosides against UV stress in *A. thaliana* (45). According to the study by Nacif de Abreu and Mazzafera, although temperature stresses negatively affected the plant growth including reduced photosynthetic rate, lower biomass and premature leaf senescence, these thermal stresses significantly increased the concentration of rutin by 3-fold at 17 °C, and 7-fold at 36 °C and also 4-fold during drought stress (43). Moreover, many authors considered the positive response of metabolites as a means to improve product quality in medicinal plants (46).

3.5 Antioxidant activity in *E. sonchifolia* leaf extracts

The antioxidant activity of E. sonchifolia plants was examined using the DPPH method. Figure 9 shows the radical scavenging activity of *E. sonchifolia* extracts at increasing concentration. The maximum percentage inhibition was observed in heat-stressed plants, followed by drought-stressed and finally the optimal ones, i.e., the stressed plant extracts exhibited higher antioxidant activity. Among the various concentrations of *E. sonchifolia* extracts tested, the concentration of 400 µg/ml showed the highest scavenging activity of 95.5%, and it was observed in extracts of heat-stressed *E. sonchifolia* samples. Meanwhile ascorbic acid at the same concentration showed 98.57% which showed close similarity to each other. The IC_{50} values, which correspond to the plant extract concentration required to scavenge 50% of the free radical, of *E. sonchifolia* were calculated from the percentage scavenging curve plotted against the concentration (**Table 3**). The IC_{50} value was inversely related to the antioxidant activity. The lesser IC_{50} value and higher radical scavenging activity indicated higher antioxidant activity. The plant extract that exhibited the highest reducing capacity was the heat-stressed *E. sonchifolia* extracts as determined by the neutralization of DPPH with the IC_{50} value of 40.65 µg/ml. Also, the highest TFC was found in heat-stressed *E. sonchifolia* extracts.

The results in **Figure 9** indicated that the 80% methanolic extracts of *E. sonchifolia* showed strong radical scavenging activity compared to vitamin C. The medicinal plant *E. sonchifolia* has been found to be promising in terms of antioxidant activity, hence the plant can be considered as a source of natural antioxidants for medicinal uses. The superiority demonstrated by *E. sonchifolia* extracts in scavenging free radicals therefore could be attributed to the presence of flavonoid metabolites. According to Nakabayashi, Yonekura-Sakakibara, quercetin and its glycosides containing more hydroxyl group exhibited more antioxidant activity than kaempferol (47). Therefore, the higher rutin and quercetin contents in *E. sonchifolia* plants might contribute to the enhanced antioxidant activity of *E. sonchifolia* plants, which counterbalanced the negative factors and protected plants against oxidative damage. The visible signs of stress were prominent in *E. sonchifolia* (**Figure 3**), indicating that these plants were more exposed to severe oxidative stress surpassing their defence mechanism during their initial encounter with ROS. Since ROS generation is an indication of stress, it has been postulated that biosynthesis of flavonoids is triggered as a supplementary defense mechanism, aimed at detoxifying ROS (48). Under severe stress conditions, secondary flavonoid metabolites function as a secondary or auxiliary scavenging system in plants, complementing the inefficient enzymatic defense mechanism in cells (42).

Table 1. LC-MS analysis of rutin, quercetin, and kaempferol in *E. sonchifolia* under heat-stressed and drought-stressed conditions.

Species	Condition	Compounds	Intensity(cps)	RT	Molecular formula	Calculated $[M-H]^-$	Observed $[M-H]^-$
E. sonchifolia	Control	Rutin	7.5×10^3	27.2	$C_{27}H_{30}O_{16}$	609.5	609.2
	Heat		4×10^4				
	Drought		3×10^4				
	Control	Quercetin	4.5×10^2	29.2	$C_{15}H_{10}O_7$	301.2	301
	Heat		2×10^3				
	Drought		1.5×10^3				
	Control	Kaempferol	1.2×10^3	30.2	$C_{15}H_{10}O_6$	285.2	285
	Heat		4.1×10^3				
	Drought		2.3×10^3				

Therefore, the stress conditions might trigger the production of more flavonoids, supporting the results obtained in this study. The results indicated that the three flavonols (**Table 1**), and the TFC (**Table 2**) were higher in stressed *E. sonchifolia* plants than other plants.

Flavonoids are excellent ROS scavengers, including quercetin and its glycosides. The stable glycosylated flavonol rutin responded positively to both heat and drought stresses regardless of the species, whereas the metabolites quercetin and kaempferol responded differently to two types of stresses, indicating the role

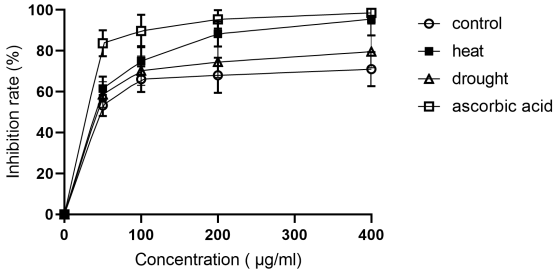


Figure 9. DPPH scavenging activity of methanolic extracts of *E. sonchifolia* with ascorbic acid. Data were represented as mean ± standard deviation of triplicate experiments.

of the three flavonols in plants as stress protectors. According to Nakabayashi, Yonekura-Sakakibara, the increased production of flavonoids reduced the accumulation of H_2O_2 and exhibited enhanced oxidative and drought tolerance (47). A study by Zhu, Li suggested that anthocyanins, a flavonoid class, acted as a secondary ROS scavenger in sugarcane leaves under chilling stress, that compensated the ineffective activity of antioxidant enzymes and contributed to chilling tolerance (49). These evidences suggest that flavonoids have a protective role during stress, and different species have specific molecular mechanisms to synthesize a variety of secondary metabolites to cope with the impact of stress.

Correlation coefficient (r) analysis between the TFC and the DPPH free radical scavenging activity (IC_{50} values) of *E. sonchifolia* extracts was determined using the data from triplicate experiments. According to data analysis using SPSS, it was found that there was a negative significant correlation between TFC and antioxidant activity of *E. sonchifolia* extracts (IC_{50}) ($r = -0.998$, $P<0.01$). The negative correlation coefficient indicated that the lower the value of DPPH (IC_{50}), the higher the TFC. The data suggested that flavonoids had a stronger influence on the antioxidant activity of *E. sonchifolia* extracts. In conclusion, heat and water stress has a higher impact on the radical scavenging activity of *E. sonchifolia* which eventually leads to the increased antioxidant activity. As the stress level increases, the oxidative stress in plant tissues is elevated causing peroxidation of lipids. The increment of the antioxidant levels is attributed to their higher radical scavenging ability, which can prevent oxidative stress due to the presence of bioactive flavonol metabolites.

Factors such as the severity of stress, the time duration of exposure and the type of species should be taken into consideration. In a study by Guo, He on *Scaevola aemula*, it was shown that with short time high-temperature stress of 46 °C for 6 hours in a day, the biosynthesis of photosynthetic pigments chlorophyll and carotenoid in the leaves increased (50). However, the prolonged high-temperature intensity in the subsequent two days resulted in the denaturation of the biosynthetic process of pigments and caused a further decrease in soluble protein content. In contrast, moderate and short-term deliberate stresses proved to be effective without much reduction in the biomass and thereby increasing product quality (51). The results from this study showed that following the severe short-term temperature stress and extended moderate drought stress (as the severity of stress was moderate in plants), serious leaf injuries were not visible in plants under study. Also, it is understood that the removal of elicitation treatment before the onset of the appearance of wilt would be advantageous for the increased production of flavonol metabolites without much injury to the plant.

Furthermore, the glycosylated flavonol production was prominent in analysed medicinal plant *E. sonchifolia*, which warrants future analysis looking for more glycosylated compounds in this medicinal plant. The LC-MS/MS could be a better option for identification and structural characterization of glycosylated compounds, as it can ionize polar compounds readily. Various pharmacological studies revealed the potent antioxidant and anticancer effects of flavonoids present in *E. sonchifolia* (52, 53). Another study suggests that the flavonoids present in *E. sonchifolia* extracts can modulate lens opacification and oxidative stress in selenite-induced cataract (54). Furthermore, Maikaeo, Chotigeat reported that *E. sonchifolia* leaf extracts, when fed to shrimp, acted as immunostimulants that protect them against virus (white spot syndrome virus and yellow head virus) induced diseases (55). The bioactive metabolites in *E. sonchifolia* had significantly high biological activities and in the future, the plant can be actively used in preclinical, clinical andtherapeutic trials investigating new treatments of diseases.

4. Conclusions

Elicitation through short-term heat stress and moderate

Table 2. Total flavonoid concentration in *E. sonchifolia* plants

Treatment	Total flavonoid Concentration (mgREg-1DW)
	E. sonchifolia
Control	3.52 ±0.84
Heat	8.66 ±0.73 *
Drought	6.01 ±0.85 *

Data was represented as mean ± standard deviation of triplicate experiments. The treated samples with single asterisk (*) adhered indicate $P < 0.05$ and were statistically significant from control as determined by Dunnett's test.

Table 3. IC_{50} values of *E. sonchifolia* extracts

Treatment	IC50 values (μgml^{-1})
	E. sonchifolia
Control	46.9
Heat	40.6 *
Drought	42.6 *

The IC_{50} values of the plant extract, as shown above, were calculated from Figure 9. The control and treated samples were statistically significant ($P < 0.05$) as determined by Dunnett's test, indicated by an asterisk (*).

water deficit stress proved to be an effective strategy to enhance the production of flavonols including rutin, quercetin and kaempferol in medicinal plants, well known for their potential benefits in therapeutic and nutraceutical industries. The results from this study inferred *E. sonchifolia* as a promising source of natural antioxidants. With reference to various pharmacological studies, it could be used for the development of natural product-based therapies against cataract as well as in the treatment of cancer. The water deficit results suggest that, non-severe or moderate stress applied before the harvest of medicinal plants can induce a significant increase in secondary metabolites without detriment to biomass accumulation. However, each plant species responds to stress differently. Consequently, further research is needed to optimise metabolic production in each species before applying these simple methods.

Acknowledgements
Authors thank the core funding from National Institute of Education, Nanyang Technological University Singapore to Z.C.

Conflict of interest
Authors declare that there are no conflicts of interest.

References

1. Davies KM, Deroles SC. Prospects for the use of plant cell cultures in food biotechnology. Curr Opin Biotechnol. 2014;26:133-40.
2. Sathuluri RR, Gokare R. Vanilla flavour: production by conventional and biotechnological routes. J. Sci. Food Agric. 2000;80:289-304.
3. Wink M. Modes of action of herbal medicines and plant secondary metabolites. Medicines (Basel, Switzerland). 2015;2(3):251-86.
4. Lee LK, Foo KY. Recent advances on the beneficial use and health implications of Pu-Erh tea. Food Research International. 2013;53(2):619-28.
5. Frederick Hill A. Economic botany: a textbook of useful plants and plant products. Science. 1939;89(2309):294.
6. Nguyen NH, Kim JH, Kwon J, Jeong CY, Lee W, Lee D, et al. Characterization of Arabidopsis thaliana FLAVONOL SYNTHASE 1 (FLS1) -overexpression plants in response to abiotic stress. Plant Physiol Biochem. 2016;103:133-42.
7. Sultana B, Anwar F. Flavonols (kaempeferol, quercetin, myricetin) contents of selected fruits, vegetables and medicinal plants. Food Chem. 2008;108(3):879-84.
8. Perez-Vizcaino F, Fraga CG. Research trends in flavonoids and health. Arch Biochem Biophys. 2018;646:107-12.
9. Javed H, Khan MM, Ahmad A, Vaibhav K, Ahmad ME, Khan A, et al. Rutin prevents cognitive impairments by ameliorating oxidative stress and neuroinflammation in rat model of sporadic dementia of Alzheimer type. Neuroscience. 2012;210:340-52.
10. Li, Yao J, Han C, Yang J, Chaudhry MT, Wang S, et al. Quercetin, inflammation and immunity. Nutrients. 2016;8(3):167.
11. Anand David AV, Arulmoli R, Parasuraman S. Overviews of biological importance of quercetin: a bioactive flavonoid. Pharmacogn Rev. 2016;10(20):84-9.
12. Kim JK, Park SU. Quercetin and its role in biological functions: an updated review. EXCLI J. 2018;17:856-63.
13. Choi JH, Park SE, Kim SJ, Kim S. Kaempferol inhibits thrombosis and platelet activation. Biochimie. 2015;115:177-86.
14. Winkel-Shirley B. Biosynthesis of flavonoids and effects of stress. Curr Opin Plant Biol. 2002;5(3):218-23.
15. Saito K, Yonekura-Sakakibara K, Nakabayashi R, Higashi Y, Yamazaki M, Tohge T, et al. The flavonoid biosynthetic pathway in Arabidopsis: structural and genetic diversity. Plant Physiol Biochem. 2013;72:21-34.
16. Oksman-Caldentey, Kirsi-Marja, Inzé D. Plant cell factories in the post-genomic era: new ways to produce designer secondary metabolites. Trends Plant Sci. 2004;9(9):433-40.
17. Yang L, Wen K-S, Ruan X, Zhao Y-X, Wei F, Wang Q. Response of plant secondary metabolites to environmental factors. Molecules (Basel, Switzerland). 2018;23(4):762.
18. Gorelick J, Bernstein N. Chapter Five - Elicitation: an underutilized tool in the development of medicinal plants as a source of therapeutic secondary metabolites. In: Sparks DL, editor. Advances in Agronomy. 124: Academic Press; 2014. p. 201-30.
19. Ramakrishna A, Ravishankar GA. Influence of abiotic stress signals on secondary metabolites in plants. Plant Signal Behav. 2011;6(11):1720-31.
20. Shen SM, Shen LG, Lei QF, Si JY, Liu CM, Lu H. Chemical constituents contained in aerial parts of Emilia sonchifolia. Zhongguo Zhong Yao Za Zhi. 2012;37(21):3249-51.
21. Dash G, Syafiq Abdullah M, Yahaya R. Traditional uses, phytochemical and pharmacological aspects of Emilia sonchifolia (L.) Int J Res Ayurveda Pharm. 2015;6(4):551-556.
22. Arun Raj G, Shailaja U, Rao PN, Sharanesh T, Gokul J. Review on the contribution of Dashapushpa: a traditional medicine in the management of cancer. Global J Res Med Plants & Indigen Med. 2013;2(9):656-63.
23. Muko KN, Ohiri FC. A preliminary study on the anti-inflammatory properties of Emilia sonchifolia leaf extracts. Fitoterapia. 2000;71(1):65-8.
24. Lija Y, Biju PG, Reeni A, Cibin TR, Sahasranamam V, Abraham A. Modulation of selenite cataract by the flavonoid fraction of Emilia sonchifolia in experimental animal models. Phytother Res. 2006;20(12):1091-5.
25. Sophia D, Ragavendran P, Arulraj C, Gopalakrishnan VK. In vitro antioxidant activity and HPTLC determination of n-hexane extract of Emilia sonchifolia (L.)DC. J Basic Clin Pharm. 2011;2(4):179-83.
26. Mamta R. Natural antioxidant (flavone glycoside) from Emilia sonchifolia DC. and its potential activity. Int J Pharm Pharm Sci. 2012;4:159-162.
27. Maikaeo L, Chotigeat W. Emilia sonchifolia extract activity against white spot syndrome virus and yellow head virus in shrimp cell cultures. Dis Aquat Organ. 2015;115(2):157-64.
28. Nasrollahi V, Mirzaie-asl A, Piri K, Nazeri S, Mehrabi R. The effect of drought stress on the expression of key genes involved in the biosynthesis of triterpenoid saponins in liquorice (Glycyrrhiza glabra). Phytochemistry. 2014;103:32-7.
29. Daudi A, O'Brien JA. Detection of hydrogen peroxide by DAB staining in arabidopsis leaves. Bio Protoc. 2012;2(18):e263.
30. Neves NdA, Stringheta PC, Gómez-Alonso S, Hermosín-Gutiérrez I. Flavonols and ellagic acid derivatives in peels of different species of jabuticaba (Plinia spp.) identified by HPLC-DAD-ESI/MSn. Food Chem. 2018;252:61-71.

31. Pękal A, Pyrzynska K. Evaluation of aluminium complexation reaction for flavonoid content assay. Food Anal Methods. 2014;7(9):1776-82.
32. Zhang, Cao Z, Xie Z, Lang D, Zhou L, Chu Y, et al. Effect of water stress on roots biomass and secondary metabolites in the medicinal plant Stellaria dichotoma L. var. lanceolata Bge. Sci Hortic. 2017;224:280-5.
33. Gai Q-Y, Jiao J, Luo M, Wei Z-F, Zu Y-G, Ma W, et al. Establishment of hairy root cultures by agrobacterium rhizogenes mediated transformation of Isatis Tinctoria L. for the efficient production of flavonoids and evaluation of antioxidant activities. PLoS One. 2015;10(3):e0119022.
34. Matteini P, Agati G, Pinelli P, Goti A. Modes of complexation of rutin with the flavonoid reagent diphenylborinic acid 2-aminoethyl ester. Monatsh Chem. 2011;142:885.
35. Cen M, Ruan J, Huang L, Zhang Z, Yu N, Zhang Y, et al. Simultaneous determination of thirteen flavonoids from Xiaobuxin-Tang extract using high-performance liquid chromatography coupled with electrospray ionization mass spectrometry. J Pharm Biomed Anal. 2015;115:214-24.
36. Li Z, Guo H, Xu WB, Ge J, Li X, Alimu M, et al. Rapid identification of flavonoid constituents directly from PTP1B inhibitive extract of raspberry (Rubus idaeus L.) leaves by HPLC-ESI-QTOF-MS-MS. J Chromatogr Sci. 2016;54(5):805-10.
37. Chen G, Li X, Saleri F, Guo M. Analysis of flavonoids in Rhamnus davurica and its antiproliferative activities. Molecules (Basel, Switzerland). 2016;21(10).
38. Kirakosyan A, Kaufman P, Warber S, Zick S, Aaronson K, Bolling S, et al. Applied environmental stresses to enhance the levels of polyphenolics in leaves of hawthorn plants. Physiol Plant. 2004;121(2):182-6.
39. Hodaei M, Rahimmalek M, Arzani A, Talebi M. The effect of water stress on phytochemical accumulation, bioactive compounds and expression of key genes involved in flavonoid biosynthesis in Chrysanthemum morifolium L. Ind Crops Prod. 2018;120:295-304.
40. Xiaolu W, Jie Y, Aoxue L, Yu C, Yijun F. Drought stress and re-watering increase secondary metabolites and enzyme activity in dendrobium moniliforme. Ind Crops Prod. 2016;94:385-93.
41. Barton KE, Koricheva J. The ontogeny of plant defense and herbivory: characterizing general patterns using meta-analysis. Am Nat. 2010;175(4):481-93.
42. Agati G, Azzarello E, Pollastri S, Tattini M. Flavonoids as antioxidants in plants: location and functional significance. Plant Sci. 2012;196:67-76.
43. Nacif de Abreu I, Mazzafera P. Effect of water and temperature stress on the content of active constituents of Hypericum brasiliense Choisy. Plant Physiol Biochem. 2005;43(3):241-8.
44. Le Roy J, Huss B, Creach A, Hawkins S, Neutelings G. Glycosylation is a major regulator of phenylpropanoid availability and biological activity in plants. Front Plant Sci. 2016;7(735).
45. Hectors K, Van Oevelen S, Geuns J, Guisez Y, Jansen MA, Prinsen E. Dynamic changes in plant secondary metabolites during UV acclimation in Arabidopsis thaliana. Physiol Plant. 2014;152(2):219-30.
46. Selmar D, Kleinwächter M. Influencing the product quality by deliberately applying drought stress during the cultivation of medicinal plants. Ind Crops Prod.2013;42:558-66.
47. Nakabayashi R, Yonekura-Sakakibara K, Urano K, Suzuki M, Yamada Y, Nishizawa T, et al. Enhancement of oxidative and drought tolerance in Arabidopsis by overaccumulation of antioxidant flavonoids. Plant J. 2014;77(3):367-79.
48. Fini A, Guidi L, Ferrini F, Brunetti C, Di Ferdinando M, Biricolti S, et al. Drought stress has contrasting effects on antioxidant enzymes activity and phenylpropanoid biosynthesis in Fraxinus ornus leaves: an excess light stress affair? J Plant Physiol. 2012;169(10):929-39.
49. Zhu J-J, Li Y-R, Liao J-X. Involvement of anthocyanins in the resistance to chilling-induced oxidative stress in Saccharum officinarum L. leaves. Plant Physiol Biochem. 2013;73:427-33.
50. Guo T, He B, Huang H, Xi W, Chen X. Physiological responses of Scaevola aemula seedlings under high temperature stress. S Afr J Bot. 2017;112:203-9.
51. Nowak M, Kleinwächter M, Manderscheid R, Weigel HJ, Selmar D. Drought stress increases the accumulation of monoterpenes in sage (Salvia officinalis), an effect that is compensated by elevated carbon dioxide concentration. J Appl Bot Food Qual. 2010;83(2):133-6.
52. Devi D G, Y L, Raghavan C, Biju P, Devi V, Annie A. Evaluation of the protective effects of Emilia sonchifolia Linn. (DC.) on perchlorate-induced oxidative damage. J Biol Sci. 2006;6.
53. Raghavan C, Gopala S, Devi D G, Srinivas P, Y L, Annie A. Antioxidant and antiproliferative effects of flavonoids from Emilia sonchifolia Linn on human cancer cells. International Journal of Pharmacology. 2006;2.
54. Lija Y, Biju P, Reeni A, Raghavan C, Sahasranamam V, Abraham A. Modulation of selenite cataract by the flavonoid fraction of Emilia sonchifolia in experimental animal models. Phytother Res. 2006;20:1091-5.
55. Maikaeo L, Chotigeat W, Mahabusarakam W. Emilia sonchifolia extract activity against white spot syndrome virus and yellow head virus in shrimp cell cultures. Dis Aquat Organ. 2015;115(2):157-64.