

A Putative Role of de-Mono-ADP-Ribosylation of STAT1 by the SARS-CoV-2 Nsp3 Protein in the Cytokine Storm Syndrome of COVID-19

Jean-Michel Claverie

► To cite this version:

Jean-Michel Claverie. A Putative Role of de-Mono-ADP-Ribosylation of STAT1 by the SARS-CoV-2 Nsp3 Protein in the Cytokine Storm Syndrome of COVID-19. Viruses, 2020, 12 (6), pp.646. 10.3390/v12060646 . hal-02966045

HAL Id: hal-02966045 https://amu.hal.science/hal-02966045

Submitted on 13 Oct 2020 $\,$

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



Opinion

MDPI

A Putative Role of de-Mono-ADP-Ribosylation of STAT1 by the SARS-CoV-2 Nsp3 Protein in the Cytokine Storm Syndrome of COVID-19

Jean-Michel Claverie

Structural & Genomic Information Laboratory (IGS, UMR 7256), Mediterranean Institute of Microbiology (FR3479), Aix-Marseille University and CNRS, 13288 Marseille, France; claverie@igs.cnrs-mrs.fr

Received: 8 May 2020; Accepted: 11 June 2020; Published: 15 June 2020



Abstract: As more cases of COVID-19 are studied and treated worldwide, it had become apparent that the lethal and most severe cases of pneumonia are due to an out-of-control inflammatory response to the SARS-CoV-2 virus. I explored the putative causes of this specific feature through a detailed genomic comparison with the closest SARS-CoV-2 relatives isolated from bats, as well as previous coronavirus strains responsible for the previous epidemics (SARS-CoV and MERS-CoV). The high variability region of the nsp3 protein was confirmed to exhibit the most variations between closest strains. It was then studied in the context of physiological and molecular data available in the literature. A number of convergent findings suggest de-mono-ADP-ribosylation (de-MARylation) of STAT1 by the SARS-CoV-2 nsp3 as a putative cause of the cytokine storm observed in the most severe cases of COVID-19. This may suggest new therapeutic approaches and help in designing assays to predict the virulence of naturally circulating SARS-like animal coronaviruses.

Keywords: de-MARylation; interferon-stimulated gene; inflammation; PARP14; ACE2

1. Introduction

Probably started in December 2019 from a zoonotic transmission from traded wild animals in a Wuhan market (Hubei Province, China) [1,2], the coronavirus disease 2019 (COVID-19) quickly reached a planetary dimension forcing billions of people into lockdown, grinding to a halt most of the world economies, and killing in excess of 375,000 of the 6.1 million people known to be infected (as of 2 June 2020). Initially catalogued as a respiratory disease (severe acute respiratory syndrome, SARS), the investigation of severe COVID-19 cases progressively revealed their association with a dysfunctional pro-inflammatory response leading to severe lung damages [3], but also cardiovascular complications [4] and multi-organ failures [5]. The unexpected severity and dimension of the COVID-19 pandemic prompted national research organizations to ask their scientists, even without prior knowledge on coronaviruses, to use their expertise to contribute to the solution of this formidable challenge. This work is part of this approach, that of a novice researcher in the field, but competent in viral genomics. In this paper, I present a bioinformatic investigation of publicly available sequence data, in an attempt to identify which molecular processes might be responsible of the specific clinical picture of COVID-19. Several coincidental findings examined in the light of the rich pre-existing bibliographical information strongly suggest that virus-induced de-mono-ADP-ribosylation (de-MARylation) of STAT1 might be one of the key steps leading to the inflammatory cytokine storm seen in the most severe COVID-19 cases.

2. Materials and Methods

Due to the lockdown constraints imposed during the COVID-19 epidemic, all the bioinformatics analyses presented in this article were performed on publicly accessible bioinformatic platforms. Databases searches were performed at NCBI using the online BlastP on various subsets of NR [6], structural analyses were performed using the PSIPRED Workbench at University College London [7], and sequence alignments at the EMBL-European Bioinformatic Institute with the Clustal Omega and Water services [8].

3. Results

As a simple, straightforward attempt to delineate genomic changes that might correlate with the crossing of the xenographic barrier to human, and with the clinical features of COVID-19, I compared the whole predicted proteome of SARS-CoV-2 with its closest animal-infecting relatives available in the NCBI public database. In a preliminary step, a comparison of the proteome sequences of 250 human-infecting SARS-CoV-2 isolates did not reveal any region of significant variability, as they were quasi-identical. For instance, out of 250 nsp3 protein sequences (1945 residues long), 214 were identical and 36 exhibited single amino-acid changes (99.95% identity), 11 of which were conservative (data not shown). One sequence from the group of 214 identical sequences (Accession # YP_009725299, Wu et al., unpublished) was randomly chosen and used for subsequent comparative analyses.

Variable regions became only detectable when comparing the polyprotein 1ab sequence of SARS-CoV-2 with that of its closest relatives in animal hosts. As previously noticed, the most similar isolate is bat coronavirus RaTG13 with 96.7% similarity (Accession # QHR63299) [1], followed by two other bat isolates (# AVP78030, # AVP78041, 93.5% and 92.5% identical, respectively) [9], and several similar pangolin isolates (e.g., # QIQ54046, 85.4% identical) (Cao et al., unpublished).

Looking for amino-acid changes occurring in clusters, my attention was drawn to a segment located near the N-terminus of the nsp3 protein (Figure 1). While 82.3% of its 1945 positions exhibit strictly identical residues between the human and bat homologs, this segment falls down to 44.5% identical residues. In SARS-CoV-2 nsp3, this segment is strongly enriched in acidic amino-acids and coincides with a paucity of predicted secondary structure elements (Figure 1). On the contrary, it is predicted as an intrinsically disordered protein region. Such domains form the basis of the dynamic "on–off" switch-type of interactions commonly found in signalling networks (reviewed in [10]). This stretch of amino-acids roughly maps to the hypervariable region (HVR) previously recognized upon the comparison of nsp3 homologs from more distant coronaviruses (reviewed in [11]), and to which no function has been assigned or predicted. I then focused further analyses on this well-delineated island of changes within the otherwise highly similar nsp3 from human, bat, and pangolin SARS-CoV-2 closest relatives.

The nsp3 segment, delineated as highly variable in the above comparison between the human SARS-CoV-2 and its closest animal homologs, is also the most divergent upon alignment of the human SARS-CoV-2 with the previous SARS-CoV-1 (responsible of the 2002 epidemics) [12], and even more dissimilar in the MERS-CoV (agent of the Middle East respiratory syndrome) [13] (Figure 2). In particular, multiple deletions and insertions make it unlikely that these homologous regions could adopt identical folds and perform similar functions.

The above analyses suggested that the N-terminal variable region (approximately from positions 111 to 237 in SARS-CoV-2) of the nsp3 protein, bearing the greatest variability between human-infecting strains, might be linked to the specific clinical pictures of the associated diseases.

Nsp3 is a large multifunctional protein, several domains of which have been mapped and characterized (reviewed in [11]). Starting from the N-terminus, there are in succession: an ubiquitin-like domain (positions: 20–106) known to bind to single stranded RNA, an ADP-ribose-binding module (positions: 238–337), a single-stranded poly(A) binding domain (positions: 678–743), a papain-like viral protease (positions: 746–1064), a nucleic acid-binding domain (positions: 1089–1201), and a seven-transmembrane G protein-coupled receptor C (positions: 1494–1563).

HUMAN COV-2	APT <mark>KVTF</mark> GDD <mark>TVIEV</mark> QGY <mark>KSVNITFEL</mark> DERIDKVLNEKCSA <mark>YTVE</mark> LGT <mark>EVNEFACVVADA</mark> 60
QHR63299.1	APTKVTFGDDTVIEVQGYKSVNITFELDERIDKVLNEKCSTYTVELGTEVNEFACVVADA 60
AVP78030.1	$\verb+APTKVTFGDDTVIEVQGYKSVNITFELDERIDKVLNEKCSNYTVELGTEVNEFACVVADA 60$
AVP78041.1	$\verb+APTKVTFGDDTVIEVQGYKSVNITFELDERIDKVLNEKCSSYTVELGTDVNEFACVVADA 60$

HUMAN COV-2	VIKTLOPVSELLTPLGIDLDEWSMATYYLFDESGEFKLASHMYCSFYPPDEDEEEGDCEE 120
OHR63299.1	VIKTLOPVSELLIPLGIDLDEWGMATYYLFDESGEFKLASHMYCSFYPPDEDEEEGDCEE 120
AVP78030.1	VIKTLOPVSELIIPLGIDLDEWSMATYYLFDESGEFKLSSHMYCSFYPPED-EGEDDCEE 119
AVP78041.1	VIKTLOPVSELLTPLGIDLDEWSMATYYLFDESGEFKLSSHMYCSFYPPED-EGEDDCEE 119

HUMAN COV-2	EEFEPSTQYEYGTEDDYQGKPLEFGATSAAL-QPEEEQEEDWLDDDSQQTVGQQDGSEDN 179
OHR63299.1	EDFEPPTQYEYGTEDDYQGKSLEFGATSVTP-QPEEELEEDWLDDDSQQTVVQEDDSEVN 179
AVP78030.1	GQCEPSTQYEYGTEDDYQGKPLEFGATSFSSSSQEEEQEEDWLESDSQDGQETA 173
AVP78041.1	GQFEPSTQYEYGTEDDYQGKPLEFGATSFSSSSQEEEQEEDWLESDSQDGQETA 173
	: ** **********************************
HUMAN COV-2	QTTTIQ <mark>TIVE</mark> VQPQLEMELTPVV-QTIEVNSFSGYLKLTDN <mark>VYIKN</mark> ADIV <mark>EEAKK</mark> VKPT <mark>V</mark> 238
OHR63299.1	QTTITQSIAEVQPQLEMEPTPVV-QT-EVNSFSGYLKLTDNVYIKNADIVEEAKKVKPTI 237
AVP78030.1	VEENKIPSVEVPPVLQVESTPVVTETSEQNNFTGYLKLTDNVFIKNADIVEEAKKVKPTV 233
AVP78041.1	VVKTSEQNNFTGYLKLTDNVFIKNADIVEEAKKVKPTV 211
	· · · · · · · · · · · · · · · · · · ·
HUMAN_COV-2	<mark>VVN</mark> AANVYLKHGG <mark>GVAGALNKA</mark> TN <mark>NAMQVESDDYIAT</mark> NGPLKVGG <mark>SCVL</mark> SGHNLAK <mark>HCLH</mark> 298
QHR63299.1	VVNAANVYLKHGGGVAGALNKATNNAMQVESDHYIATNGPLKVGGGCVLSGHNLAKNCLH 297
AVP78030.1	VVNAANVYLKHGGGVAGALNKATNNAMQVESDKYITTNGPLIVGGGCVLSGHNLAKNCLH 293
AVP78041.1	VVNAANIYLKHGGGVAGALNKATNNAMQLNLISNITTNGPLIVGGSCVLSGHNLAKNCLH 271
	****** ********************************
HUMAN COV-2	VVGPNVNKGED <mark>IQLLKSAYENF</mark> NQHEV <mark>LLA</mark> PLLSAGIFGAD <mark>PIHSLRVCVDTV</mark> RT <mark>NVYLA</mark> 358
QHR63299.1	VVGPNVNRGEDIQLLKSAYENFNQHDVLLAPLLSAGIFGADPVHSLRVCVDTVRTNVYLA 357
AVP78030.1	VVGPNVNRGEDIQLLKNAYENFNQHEILLAPLLSAGIFGADPVHSLRVCVETVHTNVYLV 353
AVP78041.1	VVGPNVNRGEDIQLLKNAYENFNQHEILLAPLLSAGIFGVDPVHSLRVCVETVHTNVYLV 331
	*******:********:***********::*********
HUMAN COV-2	VFDKNLYDKLVSSFLEMKSEKQVEQKIAEIPKEEVKPFITESKPSVEQRKQD-DKKIKAC 417
QHR63299.1	VFDKNLYDKLVSSFLEMKSEKQVEQKTAEIPKEEVKSSITESKLSVEQRQQV-DKKIKAC 416
AVP78030.1	VFDKNLYDKLVSSFLEMKSEKQVEQKVAENPKEEVKPFFTENKPSVEQRQQAEEKKIKAS 413
AVP78041.1	VFDKNLYDKLVSSFLEMKSGKQVEQKVAENPKEEVKPFFTENKPSVEQRQQAEEKKIKAS 391

HUMAN COV-2	VEEVTTLEETKFLTENLLLYIDINGNLHPDSATLVSDIDITFLKKDAPYIVGDVVQEGV 477
QHR63299.1	VEEVTTTLEETKFLTENLLLYIDINGNLHPDSATLVNDIDITFLKKDAPYIVGDVVQEGV 476
AVP78030.1	IEEVTTTLEETKFLTENLLLYIDINGNLHPDSATLVKDIDTTFLKKDVPYIVGDVIKEGA 473
AVP78041.1	IEEVTTTLEETKFLTENLLLYIDINGNLHPDSATLVKDIDITFLKKDVPYIVGDVIKEGA 451
American American Statistics (SASARA SHIPPING)	:*************************************
HUMAN COV-2	LTAVVIPTKKAGG 490
OHR63299.1	LTAVVIPTKKAGG 489
AVP78030.1	LTAVVIPTKKAGG 486
AVP78041.1	LTAVVIPTKKAGG 464

Figure 1. Multiple alignment of the human SARS-CoV-2 nsp3 N-terminal sequence and that of its 3 closest homologs in non-human-infecting coronaviruses. Sequences with accession #: QHR63299.1, AVP78030.1, AVP78041.1 are from 3 different isolates of bat-associated viruses. Their overall similarity over the entire SARS-CoV-2 nsp3 sequence are 93%, 83.3%, 80.9% identical residues, respectively. It goes down to 44% identity in the "variable region" highlighted in green. Predicted secondary structure elements are highlighted in yellow (extended) and purple (helix). The segment highlighted in green is predicted as intrinsically disordered. HUMAN_COV-2 corresponds to Genbank accession #: YP_009725299.

Interestingly, when the whole nsp3 sequence was compared to the entire human proteome using BlastP, the only reported significant matches (E-value $< 10^{-6}$) were found to precisely encompass the ADP-ribose-binding module. Moreover, the top hits are human protein with a known function: all of them are mono-ADP-ribosyltransferases. This result strongly suggests that the ADP-ribose binding site predicted in the nsp3 protein is functional, as it has been experimentally verified in a number of other viruses, including SARS-CoV (reviewed in [14]). This is the first brick on which I will

build my hypothesis: the N terminal part of the SARS-CoV-2 nsp3 protein is constituted of a variable protein-binding domain immediately adjacent to the highly conserved active site of an ADP-ribose (ADPR) processing enzyme.

HUMAN COV2	APTK-VTFGDDTVIEVQGYKSVNITFELDERIDKVLNEK-CSAYTVELGTEVNEFACVVA	58
HUMAN COV1	APIKGVTFGEDTVWEVQGYKNVRITFELDERVDKVLNEK-CSVYTVESGTEVTEFACVVA	59
HUMAN MERS	APVKKVAFGGDOVHEVAAVRSVTVEYNIHAVLDTLLASSSLRTFVVDKSLSIEEFADVVK	60
HOMAN_MERS		00
	.* *: * * ** . :.* : :::. :*.:*:.*:: *** **	
HUMAN COV2	DAVIKTLOPVSELLTPLGIDLDEWSMATYYLFDESGEFKLASHMYCSFYPPDEDEE-EGD	117
HUMAN COV1	EAVVKTLOPVSDLLTNMGIDLDEWSVATFYLFDDAGEENFSSRMYCSFYPPDEEEEDDAE	119
And a second		100001-000
HUMAN_MERS	EQVSDLLVKLLRGMPIPDFDLDDFIDAPCYCFNAEGDASWSSTMIFSLHPVECDEECS-E	119
	. * . **** * * *. ** * **	
HUMAN COV2	CEEEEFEPSTQYEYGTEDDYQGKPLEFGATSAALQPEEEQEEDWLDDDSQ	167
		2201000000
HUMAN_COV1	CEEEEIDETCEHEYGTEDDYQGLPLEFGASAETVRVEEEEEEDWLDDTTE	169
HUMAN_MERS	VEASDLEEGESECISETSTEQVDVSHEVSDDEWAAAVDEAFPLDEA	165
	* . : : * . * . * . : . * . : : :	
HUMAN COV2	QTVGQQDGSEDNQTTTIQTIVEVQPQL	194
HUMAN COV1	QIUGQDGSEDNQIIIIQIIVEVQPQL	177
—		A 1 4 1 4 1 4 1 4 1
HUMAN_MERS	EDVTESVQEEAQPVEVPVEDIAQVVIADTLQETPVVSDTVEVPPQVVKLPSEPQTIQPEV	225
	* : *	
UUMAN COVO	EMELTPVVQTIEVNSFSGYLKLTDNVYIKNADIVEEAKKVK	235
HUMAN_COV2		1.000
HUMAN_COV1	EEPTPEEPVNQFTGYLKLTDNVAIKCVDIVKEAQSAN	213
HUMAN_MERS	${\tt KEVAPVYEADTEQTQSVTVKPKRLRKKRNVDPLSNFEHKVITECVTIVLGDAIQVAKCYG$	285
	: ::.: :: * * * :: *:	
UUBANI COVO	DEPENDENT STREET FUCACITY AS INVESTIGATION OF STREETS AND A DEPENDENT STREETS AND A DEPENDENT AND A DEPENDENT	295
HUMAN_COV2	PTVVVNAANVYLKHGGGVAGALNKATNNAMQVESDDYIATNGPLKVGGSCVLSGHNLAKH	
HUMAN_COV1	PMVIVNAANIHLKHGGGVAGALNKATNGAMQKESDDYIKLNGPLTVGGSCLLSGHNLAKK	273
HUMAN_MERS	${\tt ESVLV}$ NAANTHLKHGGGIAGAINAASKGAVQKESDEYILAKGPLQVGDSVLLQGHSLAKN	345
	*:**** :******:***:* *::.*:* ***:** :*** :*** :*:*	
WIRKIN CONO	OF WERE DEPENDENT OF FRAME AND THE FACTOR AND THE PARTY OF THE PARTY	355
HUMAN_COV2	CLHVVGPNVNKGEDIQLLKSAYENFNQHEVLLAPLLSAGIFGADPIHSLRVCVDTVRTNV	5.543
HUMAN_COV1	CLHVVGPNLNAGEDIQLLKAAYENFNSQDILLAPLLSAGIFGAKPLQSLQVCVQTVRTQV	333
HUMAN_MERS	ILHVVGPDARAKQDVSLLSKCYKAMNAYPLVVTPLVSAGIFGVKPAVSFDYLIREAKTRV	405
	******: . :*:.***: :* ::::**:******* *: : .:*.*	
WIRKIN COMO		415
HUMAN_COV2	YLAVFDKNLY <mark>DKLVSSFLEMKSEKQVEQKIAEIPKEEVKPFITESKPSVEQRKQDD</mark> KKIK	415
HUMAN_COV1	YIAVNDKALYEQVVMDYLDNLKPRVEAPKQEEPPNTEDSKTEEKSVVQKPVDVKPKIK	391
HUMAN_MERS	LVVVNSQDVYKSLTIVDIPQSLTFSYDGLRGAIRKAKDYGFTVF	449
	··* ·· · ·*···· · · · · · · · · · · · ·	
		467
HUMAN_COV2	ACVEEVTTTLEETKFLTENLLLYIDINGNLHPDSATLVSDIDITFLKKDAPY	467
HUMAN_COV1	ACIDEVTTTLEETKFLTNKLLLFADINGKLYHDSQNMLRGEDMSFLEKDAPY	443
HUMAN_MERS	VCTDNSANTKVLRNKGVDYTKKFLTVDGVQYYCYTSKDTLDDILQQANKSVGIISMPLGY	509
	.* :: :.* .**** . : :: *: ::. *	
TURNAN CONO		
HUMAN_COV2	IVGDVVQEGVLTAVVIPTKKAGG 490	
HUMAN_COV1	MVGDVITSGDITCVVIPSKKAGG 466	
HUMAN_MERS	VSHGLDLIQAGSVV 523	
	: *:: * :.	

Figure 2. Comparison of the N-term regions of the SARS-CoV-2, SARS-CoV-1, and MERS-CoV nsp3 polyproteins. The glutamine-rich hypervariable region (HVR) is highlighted in green. The conserved ADP-ribose binding site is highlighted in purple. A region of lower similarity between SARS-CoV-2 and SARS-CoV-1 is also seen between positions 365 and 411 (highlighted in yellow). The rest of the SARS-CoV-2 and SARS-CoV nsp3 polyproteins are more than 80% identical, but less than 31% identical with that of MERS-CoV. Strictly conserved positions are indicated by "*", conservative variations by ":" or ".", insertions by "-". Genbank accession numbers: CoV2: YP_009725299, CoV1: AAP13439, MERS: AGV08404.

Further inspections of the above BlastP output provided an exciting guess as to the putative target of the above ADPR enzyme. All the top matches of the nsp3 putative ADPR-binding site corresponded to that of various isoforms of the PARP14 protein (Figure 3). It turns out that this protein plays a significant role in the regulation of the inflammatory response that is out of control in severe cases of COVID-19.

```
>PARP14 [Homo sapiens] mono-ADP-ribosyltransferase
Sequence ID: XP_011511231.1 Length: 1209
Range 1: 818 to 937, Score:58.2 bits, Expect:7e-07,
Identities: 40/124(32%), Positives: 61/124(49%)
        238
        VVVNAANVYLKHGGGVAGALNKATNNAMQVESDDYIATNGPLKVGGSCVLSGHNLA-KHCLHVVG
NSP3
        VVVNA+N LKH GG+A AL+KA
                                 +Q + D +
                                             GL G++
                                                            т.
                                                                H +H VG
       VVVNASNEDLKHYGGLAAALSKAAGPELQADCDQIVKREGRLLPGNATISKAGKLPYHHVIHAVG
PARP14
        818
NSP3
        PNVNKGED---IQLLKSAYE-----NFNQHEVLLAPLLSAGIFGADPIHSLRVCVDTV 351
        P + E
                  + LL+ A +
                                    ++ + P +S+G+FG
                                                         L CV+T+
       PRWSGYEAPRCVYLLRRAVOLSLCLAEKYKYRSIAIPAISSGVFG----FPLGRCVETI 937
PARP14
```

Figure 3. Best match from the BlastP similarity search output of the entire SARS-CoV-2 nsp3 protein (1935 residues) against the human proteome subset of the NR database. This unique match precisely corresponds to the ADP-ribose binding site, central to the viral macrodomain defined by previous structural studies [11].

PARP14 was shown to bind to a group of IFN-stimulated gene (ISG)-encoded proteins, most with an unknown function, and is required for their nuclear accumulation [15]. However, its best characterized function is to regulate macrophage activation via STAT1 mono-ADP-ribosylation (i.e., MARylation) [16].

STAT1 (for "signal transducer and activator of transcription 1") is a member of the STAT protein family (i.e., STAT1-6), all members of which play some role in innate immune signalling. In response to cytokines and growth factors, these proteins are phosphorylated by the receptor-associated kinases, form homo- or heterodimers, and translocate to the cell nucleus. However, when it comes to signalling through any of the interferon (IFN) receptors, STAT1 is the most important [17].

The three different classes of IFN signal through distinct cell surface receptors. Binding of type I and II IFN to their respective receptors leads to the phosphorylation of STAT1 and STAT2 which then associate in heterodimers, recruit the IFN-regulatory factor 9 (IRF9) to form the IFN-stimulated gene factor 3 (ISGF3). ISGF3 then translocate to the nucleus were it to induce genes regulated by IFN-stimulated response element (ISRE). Alternatively, binding of type II IFN (IFN- γ) to its receptor complex lead to the sole phosphorylation of STAT1 that then associates in homodimers to form the IFN- γ activation factor (GAF) that translocate to the nucleus and activate genes regulated by the gamma-activated sequence (GAS) promoter elements. Through these distinct activation pathways STAT1 is involved in the regulation of different subsets of ISGs [18]. The STAT1 phosphorylation/de-phosphorylation process is thus central to the various immune responses triggered by distinct IFN sensing: antiviral defenses through the STAT1/STAT2 heterodimer formation (IFN I and III), or macrophage activation and proinflammatory responses through STAT1 homodimer formation (IFN - γ) (reviewed in [19]).

In the context of this complex interaction network, PARP14, hinders the phosphorylation of STAT1 at "Tyr 701" through the MARylation at "Glu-657" and "Glu-705". This precludes the formation of the STAT1/STAT2 heterodimer or the STAT1 homodimers and their translocation to the nucleus. PARP14 activity is thus down-regulating the various IFN-triggered responses, in particular, the pro-inflammatory cytokine production in macrophages responding to IFN- γ stimulation [16].

This well-documented consequence of STAT1 MARylation on the inflammatory response, together with the presence of a PARP14-like ADP-ribose binding site flanking the variable region of nsp3, is at the basis of my hypothesis on the origin of the cytokine storm associated to the severe cases of COVID-19, detailed below.

4. Discussion

The ADPR-binding site that was identified in the SARS-CoV-2 nsp3 sequence corresponds to a conserved macrodomain (also called "X domain") that follows the HVR in all coronaviruses [11].

The macrodomain is a ubiquitous structural domain that removes mono-ADP-ribose from proteins, reversing the activity of ADP-ribosyltransferases. This domain has been functionally characterized [14,20–22]. It was shown that the macrodomain from SARS-CoV (as well as that of other single-stranded positive-sense, single-stranded RNA viruses) can remove mono-ADP-ribose in vitro from a variety of targets, including PARP1, PARP10, and PARP15 [22], although no specific protein targets for viral macrodomains have been identified in vivo. Here I propose that STAT1 is the in vivo target of the SARS-CoV-2 macrodomain, thus precisely counteracting its mono-ADP-ribosylation by human PARP14.

The fact that the ADP-ribose binding site of the SARS-CoV-2 nsp3 macrodomain is most resembling that of PARP14 (Figure 3) suggests that they have a similar conformation, and are both compatible with the addition (for PAPR14) or removal (for nsp3) of an ADP-ribose moiety at the same site in the context of the STAT1 3-D structure.

However, the variable disordered region flanking the conserved macrodomain, a predicted protein-protein interaction domain, would exhibit a differential (or no) affinity for STAT1 in different SARS-CoV strains, hence modulating the efficiency of its de-MARylation by nsp3 and its consequence on the various IFN-triggered responses, including the proinflammatory activation of macrophages. Alternatively, the nsp3 variable region of SARS-CoV-1 or MERS-CoV could interact with different MARyled targets altogether, leading to different behavior vis-à-vis the host innate immune response [14,20,21].

According to this molecular scenario, the expression of nsp3 in IFN- γ -activated macrophages would indirectly promote a prolonged pro-inflammatory STAT1-dependent expression of interferon stimulated genes by inhibiting the downregulation of STAT1 through its MARylation and increasing the pool of STAT1 molecules available for phosphorylation. This would participate in the cytokine storm characteristic of severe COVID-19 cases.

Moreover, a recent unexpected finding makes the above hypothesis even more appealing. Following a monumental work, Ziegler et al. [23] just showed that the SARS-CoV-2 receptor ACE2 (Angiotensin-converting enzyme 2) is an interferon-stimulated gene in cell types found in the in the upper airway (epithelial secretory goblet cells), in the lung (type II pneumocytes), and in the small intestine (absorptive enterocytes), all locations and organs linked to clinical symptoms and viral shedding [4]. Accordingly, they found evidence for STAT1 binding sites in the promoter region of ACE2. In consequence, the higher phosphorylated/unphosphorylated ratio of STAT1 promoted by nsp3 in IFN- γ -stimulated macrophages might lead to more SARS-CoV-2 receptors being expressed at the surface of bystander epithelial cells, enhancing the infection process. Once started, a positive feedback loop will proceed, with more infections generating more IFN, activating more STAT1, in turn generating more inflammation and more virus receptors, with nsp3 permanently counteracting the PARP14-dependent MARylation of STAT1, thus precluding its downregulation and the termination of the pro-inflammatory phase.

Alternatively, MARylated (hence unphosphorylated) forms of STAT1 might have a protective role against viral pathogenesis that is counterbalanced by the SARS-CoV-2 nsp3 activity. Such a hypothesis is consistent with previous studies showing in a mouse model that SARS-CoV pathogenesis is regulated by STAT1 through an IFN-receptor-independent mechanism (reviewed in [24]). Unphosphorylated STAT1 (in the context of ISGF3) was shown to drive the constitutive expression of genes inhibiting viral replication [25]. In all cases, STAT1 remains the predicted target of SARS-CoV-2 nsp3 activity.

A relatively straightforward test of my hypothesis would involve the investigation of the ADP-ribosylation status of STAT1 during SARS-CoV-2 in vitro infection (using wild-type and macrodomain-mutated strains) and the experimental demonstration of a physical interaction between STAT1 and the relevant N-terminal regions of nsp3. As the decrease of the MARylated STAT1 pool could indirectly arise from the inhibition of the MARylation function of PARP14, the eventual interaction of the latter with nsp3 should also be investigated.

Finally, the discovery that ACE2 transcription is under the positive control of STAT1 [23] provides an elegant explanation for the conservation of a functional de-MARylation macrodomain, even though it was shown to be dispensable for the in vitro replication of the virus [21]. By following the natural incentive of all viruses to maximize their infectivity, hereby increasing the density of its receptor, SARS-CoV-2 triggers a cytokine storm as an unfortunate side-effect of COVID-19. Beyond the fundamental interest of establishing a link between nsp3 and the inflammatory response associated to the disease, an in vitro measurement of the de-MARylation efficiency of STAT1 by nsp3 synthetic constructs designed after new SARS-like coronavirus genomic sequences could help predict the virulence of wild animal strains before they start circulating within human populations. The identification of STAT1 as a target might also suggest some therapeutic avenues, such as the use of anti-IFN-y drugs (e.g., Emapalumab), or Jack1/Jack2 inhibitors (e.g., Baricitinib, Ruxolinitinib) (reviewed in [26]).

Funding: This work supported by the recurrent CNRS funding to the IGS laboratory (UMR CNRS-AMU 7256).

Acknowledgments: I thank Chantal Abergel for her support, for correcting earlier versions of the manuscript, and for helpful discussions.

Conflicts of Interest: The author declares no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

- Zhou, P.; Yang, X.L.; Wang, X.G.; Hu, B.; Zhang, L.; Zhang, W.; Si, H.R.; Zhu, Y.; Li, B.; Huang, C.L.; et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* 2020, 579, 270–273. [CrossRef]
- 2. Lam, T.T.; Shum, M.H.; Zhu, H.C.; Tong, Y.G.; Ni, X.B.; Liao, Y.S.; Wei, W.; Cheung, W.Y.; Li, W.J.; Li, L.F.; et al. Identifying SARS-CoV-2 related coronaviruses in Malayan pangolins. *Nature* **2020**. [CrossRef] [PubMed]
- 3. Tay, M.Z.; Poh, C.M.; Rénia, L.; MacAry, P.A.; Ng, L.F.P. The trinity of COVID-19: Immunity, inflammation and intervention. *Nat. Rev. Immunol.* **2020**. [CrossRef] [PubMed]
- 4. Atri, D.; Siddiqi, H.K.; Lang, J.; Nauffal, V.; Morrow, D.A.; Bohula, E.A. COVID-19 for the Cardiologist: A Current Review of the Virology, Clinical Epidemiology, Cardiac and Other Clinical Manifestations and Potential Therapeutic Strategies. *JACC Basic Transl. Sci.* **2020**. [CrossRef] [PubMed]
- 5. Liu, P.P.; Blet, A.; Smyth, D.; Li, H. The Science Underlying COVID-19: Implications for the Cardiovascular System. *Circulation* **2020**. [CrossRef]
- Sayers, E.W.; Beck, J.; Brister, J.R.; Bolton, E.E.; Canese, K.; Comeau, D.C.; Funk, K.; Ketter, A.; Kim, S.; Kimchi, A.; et al. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* 2020, 48, D9–D16. [CrossRef]
- 7. PSIPRED. Available online: http://bioinf.cs.ucl.ac.uk/psipred/ (accessed on 1 May 2020).
- 8. Multiple Sequence Alignment. Available online: www.ebi.ac.uk/Tools/msa/ (accessed on 1 May 2020).
- Hu, D.; Zhu, C.; Ai, L.; He, T.; Wang, Y.; Ye, F.; Yang, L.; Ding, C.; Zhu, X.; Lv, R.; et al. Genomic characterization and infectivity of a novel SARS-like coronavirus in Chinese bats. *Emerg. Microbes Infect.* 2018, 7, 154. [CrossRef]
- Uversky, V.N. Intrinsically Disordered Proteins and Their "Mysterious" (Meta) Physics. *Front. Phys.* 2019, 7, 10. [CrossRef]
- 11. Lei, J.; Kusov, Y.; Hilgenfeld, R. Nsp3 of coronaviruses: Structures and functions of a large multi-domain protein. *Antivir. Res.* **2018**, *149*, 58–74. [CrossRef]
- 12. Rota, P.A.; Oberste, M.S.; Monroe, S.S.; Nix, W.A.; Campagnoli, R.; Icenogle, J.P.; Peñaranda, S.; Bankamp, B.; Maher, K.; Chen, M.H.; et al. Characterization of a novel coronavirus associated with severe acute respiratory syndrome. *Science* **2003**, *300*, 1394–1399. [CrossRef]
- 13. Moreno, A.; Lelli, D.; de Sabato, L.; Zaccaria, G.; Boni, A.; Sozzi, E.; Prosperi, A.; Lavazza, A.; Cella, E.; Castrucci, M.R.; et al. Detection and full genome characterization of two beta CoV viruses related to Middle East respiratory syndrome from bats in Italy. *Virol. J.* **2017**, *14*, 239. [CrossRef] [PubMed]
- 14. Alhammad, Y.M.O.; Fehr, A.R. The Viral Macrodomain Counters Host Antiviral ADP-Ribosylation. *Viruses* **2020**, *12*, 384. [CrossRef] [PubMed]

- 15. Caprara, G.; Prosperini, E.; Piccolo, V.; Sigismondo, G.; Melacarne, A.; Cuomo, A.; Boothby, M.; Rescigno, M.; Bonaldi, T.; Natoli, G. PARP14 Controls the Nuclear Accumulation of a Subset of Type I IFN-Inducible Proteins. *J. Immunol.* **2018**, *200*, 2439–2454. [CrossRef] [PubMed]
- Iwata, H.; Goettsch, C.; Sharma, A.; Ricchiuto, P.; Goh, W.W.; Halu, A.; Yamada, I.; Yoshida, H.; Hara, T.; Wei, M.; et al. PARP9 and PARP14 cross-regulate macrophage activation via STAT1 ADP-ribosylation. *Nat. Commun.* 2016, *7*, 12849. [CrossRef] [PubMed]
- 17. Schneider, W.M.; Chevillotte, M.D.; Rice, C.M. Interferon-stimulated genes: A complex web of host defenses. *Annu. Rev. Immunol.* **2014**, *32*, 513–545. [CrossRef]
- 18. Rauch, I.; Müller, M.; Decker, T. The regulation of inflammation by interferons and their STATs. *JAKSTAT* **2013**, 2, e23820. [CrossRef]
- 19. Ivashkiv, L.B. IFNγ: Signalling, epigenetics and roles in immunity, metabolism, disease and cancer immunotherapy. *Nat. Rev. Immunol.* **2018**, *18*, 545–558. [CrossRef]
- 20. Fehr, A.R.; Jankevicius, G.; Ahel, I.; Perlman, S. Viral Macrodomains: Unique Mediators of Viral Replication and Pathogenesis. *Trends Microbiol.* **2018**, *26*, 598–610. [CrossRef]
- 21. Fehr, A.R.; Channappanavar, R.; Jankevicius, G.; Fett, C.; Zhao, J.; Athmer, J.; Meyerholz, D.K.; Ahel, I.; Perlman, S. The Conserved Coronavirus Macrodomain Promotes Virulence and Suppresses the Innate Immune Response during Severe Acute Respiratory Syndrome Coronavirus Infection. *mBio* **2016**, *7*, e01721-16. [CrossRef]
- 22. Li, C.; Debing, Y.; Jankevicius, G.; Neyts, J.; Ahel, I.; Coutard, B.; Canard, B. Viral Macro Domains Reverse Protein ADP-Ribosylation. *J. Virol.* **2016**, *90*, 8478–8486. [CrossRef]
- Ziegler, C.G.K.; Allon, S.J.; Nyquist, S.K.; Mbano, I.M.; Miao, V.N.; Tzouanas, C.N.; Cao, Y.; Yousif, A.S.; Bals, J.; Hauser, B.M.; et al. SARS-CoV-2 receptor ACE2 is an interferon-stimulated gene in human airway epithelial cells and is detected in specific cell subsets across tissues. *Cell* 2020, 181, 1016–1035.e19. [CrossRef] [PubMed]
- 24. Totura, A.L.; Baric, R.S. SARS coronavirus pathogenesis: Host innate immune responses and viral antagonism of interferon. *Curr. Opin. Virol.* **2012**, *2*, 264–275. [CrossRef] [PubMed]
- 25. Wang, W.; Yin, Y.; Xu, L.; Su, J.; Huang, F.; Wang, Y.; Boor, P.P.C.; Chen, K.; Wang, W.; Cao, W.; et al. Unphosphorylated ISGF3 drives constitutive expression of interferon-stimulated genes to protect against viral infections. *Sci. Signal.* **2017**, *10*, eaah4248. [CrossRef] [PubMed]
- 26. Merad, M.; Martin, J.C. Pathological inflammation in patients with COVID-19: A key role for monocytes and macrophages. *Nat. Rev. Immunol.* **2020**, *20*, 355–362. [CrossRef]



© 2020 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).