

# “Vaccine-Induced Covid-19 Mimicry” Syndrome: Splice reactions within the SARS-CoV-2 Spike open reading frame result in Spike protein variants that may cause thromboembolic events in patients immunized with vector-based vaccines

## Eric Kowarz

Institute of Pharmaceutical Biology/DCAL, Goethe-University of Frankfurt, Biocenter, Max-von-Laue-Str. 9, 60438 Frankfurt/Main, Germany

## Lea Krutzke

Department of Gene Therapy, Ulm University, Helmholtz Str. 8/1, 89081 Ulm, Germany

## Jenny Reis

Institute of Pharmaceutical Biology/DCAL, Goethe-University of Frankfurt, Biocenter, Max-von-Laue-Str. 9, 60438 Frankfurt/Main, Germany

## Silvia Bracharz

Institute of Pharmaceutical Biology/DCAL, Goethe-University of Frankfurt, Biocenter, Max-von-Laue-Str. 9, 60438 Frankfurt/Main, Germany

## Stefan Kochanek

Department of Gene Therapy, Ulm University, Helmholtz Str. 8/1, 89081 Ulm, Germany

## Rolf Marschalek (✉ [rolf.marschalek@em.uni-frankfurt.de](mailto:rolf.marschalek@em.uni-frankfurt.de))

Institute of Pharmaceutical Biology/DCAL, Goethe-University of Frankfurt, Biocenter, Max-von-Laue-Str. 9, 60438 Frankfurt/Main, Germany

---

## Research Article

**Keywords:** Vector-based vaccines, CVST, SVT, thrombosis, splicing, Spike protein

**DOI:** <https://doi.org/10.21203/rs.3.rs-558954/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

During the last months many countries have started the immunization of millions of people by using vector-based vaccines. Unfortunately, severe side effects became overt during these vaccination campaigns: cerebral venous sinus thromboses (CVST), absolutely rare under normal life conditions, were found as a severe side effect that occurred 4-14 days after first vaccinations. Besides CVST, Splanchnic Vein Thrombosis (SVT) was also observed. This type of adverse event has not been observed in the clinical studies of AstraZeneca, and therefore led immediately to a halt in vaccinations in several European countries. These events were mostly associated with thrombocytopenia, and thus, similar to the well-known Heparin-induced thrombocytopenia (HIT). Meanwhile, scientists have proposed a mechanism to explain this vaccine-induced thrombocytopenia. However, they do not provide a satisfactory explanation for the late thromboembolic events. Here, we present data that may explain these severe side effects which have been attributed to adenoviral vaccines. According to our results, transcription of wildtype and codon-optimized Spike open reading frames enables alternative splice events that lead to C-terminal truncated, soluble Spike protein variants. These soluble Spike variants may initiate severe side effects when binding to ACE2-expressing endothelial cells in blood vessels. In analogy to the thromboembolic events caused by Spike protein encoded by the SARS-CoV-2 virus, we termed the underlying disease mechanism the “Vaccine-Induced Covid-19 Mimicry” syndrome (VIC19M syndrome).

## Introduction

This world-wide pandemic, starting in the last months of 2019 in Wuhan (China), is caused by SARS-CoV-2, the most recent member of the large corona virus family. Until now (25th May 2021) we have more than 167 Mio infected people and nearly 3.5 Mio deaths that could be associated with a SARS-CoV-2 infection. We have observed that thousands of physicians and scientists focused their attention on this virus in order to address the pathobiology of the virus, to quickly find out how to treat acute Covid-19 patients, but also to deal with the consequences of the infection (Long Covid). Most importantly, the world has faced the fastest development and production of vaccines ever, resulting in 4 vaccines already approved by the EMA and national regulatory authorities in Germany and other European countries: the mRNA vaccines from BionTech and Moderna, as well as the recently authorized adenoviral vector-based vaccines from AstraZeneca and Johnson & Johnson.

While the national vaccination programs with Biontech and Moderna appear to cause only mild and typical immediate vaccination effects, severe side effects were first observed with AZD1222/Vaxzevria in Europe (~ 150 cases after ~ 17 mio doses) and more recently with the Johnson & Johnson vaccine in the United States (~ 6 cases after 6,8 mio doses). The first cases of cerebral venous sinus thrombosis led to a rapid suspension of the AZD1222/Vaxzevria vaccination campaign in several European countries (11-14th March 2021: Denmark, several northern European countries, Thailand, Ireland; 15th March 2021: Germany, Italy, France and Spain).

Most of these cases displayed thrombocytopenia combined with a rare cerebral venous sinus thromboembolic event that caused the death in one third of affected vaccinated persons in Germany. Many cases of CVST and some cases of splanchnic vein thrombosis (SVT) were reported in the EU drug safety database and in the UK. However, after several days of suspension, the EMA did not see any problem with the vaccine (18th March), and later also the German STIKO suggested administering the AZ vaccine solely to people over 60 years of age, since most cases in Germany affected women between the age of 20 and 60 years. Meanwhile (10th May 2021), the EMA has suggested to also use the Johnson & Johnson vaccine only for people over 60 years.

However, there remained a question regarding the mechanism behind these thromboembolic events. Thromboembolic events combined with a thrombocytopenia have many similarities with a known phenomenon termed "Heparin-induced thrombocytopenia" (HIT), where these combined effects are induced by the infusion of heparin. The mechanism involves Fc $\gamma$ -receptor-mediated activation of platelets by autoantibodies against the platelet derived factor PF4, in complex with heparin, and subsequent coagulation events. However, scientists in Greifswald together with several other international scientists were quickly able to demonstrate that the novel vaccine induced thrombosis and thrombocytopenia (VITT) mechanism is not provoked by antibodies directed against the Spike antigen (1,2), which was an important finding, because it reaffirmed the safety of all the other vaccines on the market. Although the proposed VITT mechanism explained the thrombocytopenia, the precise trigger behind the thromboembolic events has yet to be resolved.

Therefore, we investigated another possible mechanism. As already mentioned above, Covid-19 is caused by SARS-CoV-2, a single-stranded RNA virus of the  $\beta$ -Corona virus family, encoding 16 non-structural proteins (NSP's 1–16), 8 accessory proteins (ORF3a, 6, 7a, 7b, 8, 9b, 9c and 10) and 4 structural proteins (S, E, M and N) (3). The spike glycoprotein is responsible both for recognition of host cell membrane receptors ACE2 and TMPRSS2, for mediating fusion with the membrane of a host cell (4). Most importantly, the RNA of this virus is translated and replicated only in the cytosol of infected cells, and thus, evolution of coronaviruses has always taken place in the cytosol of permissive cells, and in the absence of processes that are necessarily taking place when nuclear-encoded genes are being transcribed (capping, splicing and poly-adenylation).

Here, we want to draw attention to the process of RNA splicing (5). All genes encoded by our genome have evolved in the presence of intronic sequences which may have played an important role during evolution to safe-guard coding information, but also to shuffle genetic information in order to create new cellular functions. With very few exceptions, nuclear encoded genes exhibit intronic sequences and primary transcripts are subject to splice reactions to eliminate intronic sequences. For this purpose, consensus splice donor and acceptor sites have evolved to drive this process that occurs co-transcriptionally in our cell nuclei. The RNA consensus sequences are recognized by RNA/protein complexes (spliceosome) with consists of nearly 100 proteins and specific U-RNA's (U1, U2, U4, U5 and U6). These RNA/protein complexes bind via specific RNAs in a complementary fashion to the consensus sequences. The perfect 5'-intronic site is characterized by "G•GUNNGU", while the 3'-end of an intron is

characterized by a branch-A nucleotide and the sequence "YYYYYYYCAG•**G** (the nucleotide in bold and the dots mark the last and first nucleotide of exonic sequences, respectively). Noteworthy, splice events are influenced by exonic splice enhancer (6 nucleotides; ESE) and exonic splice silencer sequences (4–18 nucleotides; ESS) which bind splice-activating or -suppressing proteins.

What is the fundamental difference between mRNA and vector-based vaccines? The mRNA vaccines are delivered by a lipid nanoparticle containing the appropriate mRNA molecule - coding for the spike protein of SARS-CoV-2 - to muscle cells surrounding the injection site. Cells that have successfully taken up these nanoparticles will release their cargo mRNA into the cytosol, where it will be translated into Spike protein in the rough endoplasmic reticulum (ER). Subsequently, the translated and folded Spike proteins will be post-translationally modified in the ER and Golgi apparatus and transported to the outer membrane - as membrane-anchored proteins. This way, the immune system is able to recognize the viral antigen, which in turn triggers the initial events for all subsequent immunological processes to produce specific B- and T-effector cells.

What happens to the same Spike gene when delivered via an adenoviral system? The adenovirus life cycle includes the infection of cells, uncoating of the virus in the cytosol, entry of the adenoviral DNA into the nucleus, and subsequently gene transcription by the host transcription machinery (6). All adenoviral systems follow exactly these steps (Ad5, Ad26 and chimp Ad). Thus, the SARS-CoV-2 Spike gene will be transcribed inside of the nucleus and subsequently exported as mRNA out of the nucleus. Arriving in the cytosol, the mRNA will again be translated into the Spike protein (see above).

And exactly here lies the problem: the viral piece of DNA - deriving from an RNA virus - is not optimized to be transcribed inside of the nucleus. Solely this 3,822 nucleotide long open reading frame, coding for a primary product of 1274 amino acid long Spike protein, contains 6 predicted splice donor and 5 predicted acceptor sites. This problem becomes even more severe when using codon-optimized Spike reading frames (depending on the company: up to 13 splice donor and 11 acceptor sites; see Fig. 1A). Thus, it could well be that the Spike open reading frame of SARS-CoV-2 is potentially disrupted by arbitrary splice events when transcribed inside the nucleus. Most, if not all, of these undesirable splice events would produce shorter protein variants, disrupting the Spike protein upstream of the C-terminally located membrane anchor, and thus, leading to soluble Spike protein variants.

To investigate this hypothesis, we examined a nuclear transcribed Spike open reading frame in a "splice reporter system" in order to qualitatively and quantitatively measure these splice events. In addition, we investigated a Human adenovirus C type 5 (HAdV-C5) encoding for a codon-optimized SARS-CoV2 Spike protein and the AZD1222 vaccine after infecting Hela cells to identify potential splicing events. Here, we present our data and discuss the consequences.

## Material And Methods

### Splice site prediction

The open reading frame of the SARS-CoV2 Spike protein was analyzed by using the SpliceRover online tool ([bioit2.irc.ugent.be/rover/splicerover](http://bioit2.irc.ugent.be/rover/splicerover)) which predicts splice donor and acceptor sites by using convolutional neural networks to achieve state-of-the-art accuracy. Here, we excluded all predicted sites below a score of 0.15 to identify only highly significant sites (7). In addition, we used the Spike sequences together with the Alternative Splice Site Predictor (ASSP) online tool, just to have another algorithm that identifies potential splice sites (8).

### Cloning of the open reading frame of SARS-CoV-2 Spike protein

cDNA prepared from the RNA of the Wuhan SARS-CoV-2 strain was kindly provided by colleagues from the Institute of Medical Virology, University Hospital Frankfurt am Main, Goethe-University, Frankfurt am Main, Germany (Tuna Toptan, Marek Widera). The cDNA was used for PCR experiments and to clone the open reading frame of the Spike gene by using the following primers: CoV-2\_S.Flag.Sfi.F 5'-aGGCCTCT-GAGGCCacatggattacaaggatgacgacgataagatgtttgttttctgttttattgccactagtctct-3', CoV-2\_S.Sfi.R 5'-aGGC-CTGACAGGCCttatgtgtaattgtaattgactcctttgagcac-3'. Resulting amplimers were cloned into the Topo II vector system and appropriate clones were validated by sequence analyses. The final clone was digested by Sfi1 and cloned into our well-established Sleeping Beauty transposon vector system pSBTet-GP (GFP/Puromycin; 9). The cloned Spike open reading frame exhibits at its very N-terminal portion a Flag-tag.

### Development of a Spike/splice reporter system

The cloned Spike open reading frame was used to establish an additional splice reporter system. The Spike/splice reporter system consists of the full-length Spike gene including the stop codon, followed by an intronic sequence that contains the necessary branch-A nucleotide, as well as a perfect splice acceptor site that has been fused in three different reading frames (0, +1, +2) to a near full-length Luciferase gene (without start codon). To this end, Luciferase activity can only be measured if splicing occurs. Thus, 3 independent Spike/splice reporter systems were constructed and cloned again into our Sleeping Beauty transposon vector pSBbi-GP (9).

### Stable transfection and transduction experiments

All 3 designed splice reporter constructs (pSBbi::Spike-Luc-0/+1/+2-GP; see above) were stably transfected into HEK293T cells and selected for 3-5 days with 2 $\mu$ g/ml Puromycin. Cells were used for Luciferase experiments or for the isolation of total RNA (Qiagen) in order to investigate potential splice events by RT-PCR experiments.

## Adenoviral vectors

HAdV-C5-Spike vector (shortly HAd5::Spike) particles produced in N52.E6 cells (10), had a physical titer of  $3.35 \times 10^8$  VP/ $\mu$ l, are dissolved in 50 mM HEPES, 150 mM NaCl, 10% glycerol, pH 7.4. According to the manufacturer ChAdOx1 nCov-19 vaccine from AstraZeneca (lot number ABV9317) has a physical titer of  $1 \times 10^8$  VP/ $\mu$ l and is dissolved in 10 mM histidine, 7.5% sucrose (w/v), 35 mM NaCl, 1 mM MgCl<sub>2</sub>, 0.1% polysorbate 80 (w/v), 0.1 mM EDTA and 0.5% EtOH (w/v). According to the manufacturer Ad26.COVS-2 vaccine from Janssen (lot number XD955) has a infectious titer of  $1.78 \times 10^8$  VP/ $\mu$ l and is dissolved in citric acid monohydrate (0.14 mg), trisodium citrate dihydrate (2.02 mg), ethanol (2.04 mg), 2-hydroxypropyl- $\beta$ -cyclodextrin (HBCD) (25.50 mg), polysorbate-80 (0.16 mg), sodium chloride (2.19 mg).

## Production and purification of HAdv-C5-Spike vectors

HAd5::Spike vector particles used in this study are *E1*- and *E3*-deleted replication-incompetent vector particles based on human adenovirus species C type 5 (based on GenBank AY339865.1,  $\Delta$ 441–3522 and  $\Delta$ 28123-30813). The CMV-promoter controlled SARS-CoV2 Spike expression cassette with SV40 poly-A is based on the GenBank sequence YP\_009724390.1, codon-optimized for *Homo sapiens* and was inserted in the *E1* region. Particles were produced in *E1*-complementing N52.E6 cells (10) and purified by CsCl<sub>2</sub> gradient ultracentrifugation as described earlier (11). In brief:  $4 \times 10^8$  cells were transduced with MOI 300 from stock solution. 48 h post transduction cells were harvested, resuspended in 3 ml 150 mM NaCl, 50 mM HEPES, pH 7.4 and lysed by freeze/thaw cycles. Cell debris was separated by centrifugation at 2,000 x g for 10 min, and supernatants were layered on a CsCl step gradient (density bottom: 1.41 g/ml; density top: 1.27 g/ml, 50 mM HEPES, 150 mM NaCl, pH 7.4) and centrifuged at 176,000 x g for 2 h at 4°C. Vector particles were aspirated and further purified by a consecutive continuous CsCl gradient (density: 1.34 g/ml, 50 mM HEPES, 150 mM NaCl, pH 7.4) and centrifuged at 176,000 x g for 20 h at 4°C. Vector particles were aspirated and desalted by size exclusion chromatography using PD10 columns (GE Healthcare). Physical vector titers were determined by optical density measurement at OD<sub>260nm</sub> (12).

## Determination of cell transduction efficiencies by different adenovirus strains

A total of  $1 \times 10^5$  Hela cells were transduced with MOI 500, 1000, 3000, 5000, or 10000 of HAdV-C5-Spike, ChAdOx1 or Ad26.COVS-2. Cells were incubated with the vector for 2 h at 37°C, before cells were harvested and DNA isolated using GenElute Mammalian Genomic DNA Miniprep Kit (Sigma) according to the manufacturer instructions. DNA concentrations were determined by optical density measurement at 260<sub>nm</sub>. Twenty ng total DNA was analyzed for the adenoviral genome DNA content by quantitative real time PCR and compared to HAd5::Spike genome copy numbers. Primers used for HAdV-C5-Spike and

ChAdOx1 samples: fw.: 5'-TAGACGATCCCTACTGTACG-3'; rv.: 5'-GGAAATATGACTACGTCCGG-3'. Primers used for HAdV-C5-Spike and Ad26.COVID-2-S samples: fw.: 5'-CAGGACGCCTCGGAGTACCTGAG-3'; rv.: 5'-GGGGCCACCGTGGGGTT-3'. DNA was dissolved in 2 µl and mixed with 8 µl SYBR Green (Kapa Biosystems), 0.4 µl 10 pmol/µl of each forward and reverse primer in a total volume of 20 µl. Thermocycles: first cycle: 10 min at 95 °C; 40 cycles: 30 sec 95 °C, 30 sec 60 °C, 8 sec 72 °C; followed by a final cycle of 10 min at 72 °C.

## Analysis of SARS-CoV2 spike sequence splicing upon adenoviral cell transduction

Adenoviral cell transduction assays were performed using HeLa cells, cultivated in MEM medium with FBS and antibiotics. To achieve equivalent cell transduction efficiencies with adenovirus vectors of different strains, HeLa cells were transduced with MOI 1000 for HAd5::Spike, MOI 3000 for AZD1222 ChAdOx1-nCov19 and MOI 6000 for Ad26.COVID-2-S. Cells were incubated with the vectors for 48 h at 37 °C, before RNA was isolated using RNeasy Plus Mini Kit (Qiagen) followed by cDNA generation using Maxima H Minus cDNA Synthesis (ThermoScientific) according to the respective manufacturer instructions. PCR products were separated on agarose gels, bands were excised, purified with gel extraction kit (Qiagen) and sequenced. PCR samples were cloned using the pCR 2.1-TOPO TA cloning kit (ThermoScientific) and sequenced.

## RT-PCR experiments

For RT-PCR experiments a series of primers were used. These were the following: S.0233.F 5'-GGTTTGATAACCCTGTCTACCA-3', S.1027.F 5'-AACGCCACCAGATTTGCATC-3', S.1712.F 5'-ACACTACTGATGCTGTCCGTG-3', S.2482.F 5'-CTTGCAGATGCTGGCTTCAT-3', S.2982.F 5'-TAGGTTGATCACAGGCAGACT-3', S.3723.R 5'-ACAGCAACTGGTCATACAGCA-3', S.3168.R 5'-TGCTGACTGAGGGAAGGACA-3', S.2587.R 5'-GTGGCAAACAGTAAGGCCG-3', S.2019.R 5'-ACTAGCGCATATACCTGCACC-3'. RT-PCR experiments were carried out under stringent conditions, and PCR fragments deviating from expected PCR bands were analyzed by sanger sequencing.

To test for splice events in the 3 reporter cell lines, we used the following primer sequences: For analyzing splice events between the Spike and Luciferase ORF, we used the forward primer S.2482.F 5'-CTTGCAGATGCTGGCTTCAT-3', S.2982.F 5'-TAGGTTGATCACAGGCAGACT-3' in combination with the Luc.3912.R 5'-TTCGAGTGGGTAGAATGGCG-3' for all three reporter constructs to monitor the alternative splice events, that create the different Spike-Luciferase fusion proteins.

## Luciferase assays

The same cell lines were used to measure luciferase activity. Since most splicing events from the Spike open reading frame to the ATG-less Luciferase cassette will destroy the membrane anchor, we also investigated the supernatant of the 3 cell cultures for the presence of Luciferase. Briefly, we seeded 5,000 cells per well in black glass bottom 96-well plates (Greiner BioOne). After 2 days of culture, GFP fluorescence (RFU) was measured with the Varioskan Flash Plate-reader (Thermo Fischer Scientific). Subsequently, D-Luciferin (final concentration of 187.5ug/ml) was added with the automatic dispense system, and Luciferase activity was measured (RLU). A normalization procedure was used to express the RLU/RFU ration to account for intercellular differences. All measurements were carried with n=6, while luciferase activity in medium was measured once at two different time points (t=0 and t=48h). For his purpose 200µl medium was centrifugated for 5 minutes at 800 x g, and 100µl supernatant was used again to measure Luciferase activity.

## Results

### In silico analyses of potential splice sites in various Spike open reading frames

The Spike open reading frame of the Wuhan SARS-CoV-2 isolate was tested for potential splice sites by using the available SpliceRover online tool. In addition, we used the online Alternative Splice Site Predictor (ASSP) tool to confirm the predicted splice sites (data not shown). Both servers identified several dozens of splice sites in the 3,822 nucleotide long open reading frame of the Spike gene, and so we decided to filter these by their individual scores to find the ones which are highly likely to be used (wildtype Spike gene: 11, see **Table 1A**). As shown in **Figure 1A**, we performed the same *in silico* analysis also for codon-optimized Spike open reading frames. The AZ sequence was deduced from a publication which describes the use of the GeneArt codon optimizer (Fisher Scientific), while the Johnson & Johnson sequence was obtained directly from the Janssen company (kindly provided by Roland Zahn). Codon optimization resulted in both cases in a higher number of potential splice sites (AZ: 24, Janssen: 20, see **Table 1B** and **1C**).

### The SARS-CoV-2 Spike open reading frame enables arbitrary splice events that enables the secretion of a soluble Spike protein

First, we analyzed potential splice events within the open reading frame of wildtype Spike, where we had already predicted 6 potential splice donor sites. Such unsaturated splice donor sites may give rise to artificial splice events (to cryptic splice acceptors) or even generate trans-splicing events. When we tested all three splice reporters in classical Luciferase assay, we realized that the “+2 construct” exhibited the strongest Luciferase activity, while the “0 construct” revealed a much lower activity. The activity measured for the +1 construct was near the values measured non-transfected HEK293T cells which served as negative controls (see **Figure 1B**). More importantly, we also tried to analyze the medium of the cultured cells and could indeed show that tiny amounts of Luciferase activity could be measured also extracellularly. Due to the design of the 3 splice traps, Luciferase activity can only be measured when



splicing occurs. And if splicing occurs within the Spike open reading frame, the C-terminal located domain that serves as membrane anchor will be lost. To this end, these Luciferase experiment strongly indicated that splice events are indeed taking place.

To validate these findings, we used isolated RNA from all three cell lines and investigated potential splice events by sequencing RT-PCR bands that we obtained when we performed 4 independent RT-PCR experiments (4 Spike primers pointing 3'-prime against 1 Luciferase primer pointing 5'-prime). As exemplarily shown in **Figure 1B**, we observed several splice reactions from predicted splice sites (e.g. SD1-3), but also internal splice reactions that lead to frame-shift events. We also found 2 splice events downstream of the ACE2 binding domain that fused again to Luciferase. At this point, we have to conclude that splice reactions do indeed take place and all of them cause a C-terminal truncation of the Spike protein, exactly as predicted by our initial hypothesis.

In order to also analyze potential splice events in the vector-based vaccines, an HAd5::Spike constructs and the AZ vaccine were used to transduce Hela cells. After 2 days, RNA was isolated and analyzed again by RT-PCR experiments. This time, however, primers were designed not only for the Spike open reading frame (LK76, LK78, LK113), but also a primer for a gene that is located 3' of the Spike transgene in the backbone of the adenovirus (LK114). As shown in **Figure 1C** (left panel), some smaller PCR bands can be observed when analyzing the full-length open reading frame of the Spike transgene. Noteworthy, a strong splice donor site is predicted at position +506, and several splice acceptor sites further downstream may explain the observed pattern. The middle panel demonstrates the integrity of the second half of the Spike cassette. In the third panel, a full-length PCR band was observed, as well as 2 smaller bands. These smaller bands might be explained by splice events between the C-terminal portion of Spike with the further downstream localized pIX gene. The difference in size can be explained by a splice donor site at nt 3,614 within the 4,075 shortly upstream of the pIX gene (-14), which was subsequently confirmed by direct sequencing of the faster migrating PCR band. This artificially spliced transcript can only be explained by transcriptional read through into the pIX ORF, or by a trans-splicing event from the Spike transcript into the 5'NTR of the pIX transcript. Either way, it will lead to a fully functional Spike-pIX fusion protein that could be secreted into body fluids due to the missing membrane anchor of Spike.

## Discussion

Here, we present first molecular evidence that vector-based vaccines encoding the Spike protein exhibit a problem that is completely absent in mRNA-based vaccines. This is due to the fact that during the vaccination step, the adenoviral DNA enters the nucleus and use the host machinery to transcribe its (trans)genes inside the nucleus. However, RNA viruses have evolved in the absence of any post-transcriptional modification systems that are usually enabled to process the primary RNA transcripts of nuclear encoded genes.

In this study, we experimentally validated our assumption that potential splice events cause the production of Spike protein variants that have lost the important membrane anchor, resulting in secreted,

soluble Spike protein variants. We could show that (1) predicted splice donor sites are used to create Spike-Luciferase fusion proteins, that Luciferase activity can be measured intra- and extracellularly which was a clear indication that splice reactions are occurring, and (3) that codon-optimized Spike protein – that exhibits more and stronger splice donor sites – was able to fuse with the pIX gene either by transcriptional read-through and splicing, or by a trans-splicing event. All these results indicate that splice reactions occur and create soluble Spike protein. Soluble Spike protein has been described to cause adverse effects, e.g. a strong inflammatory response on endothelial cells [13–16]. Moreover, nearly all severe cases of SARS-CoV-2 infections (COVID-19) suffer from life-threatening thromboembolic events due to the many viruses with Spike surface protein in the blood stream. Even pseudoviruses with Spike protein on the surface cause strong inflammatory reactions in tissues and endothelial cells, indicating the danger of this protein when available in a systemic fashion [17].

Soluble Spike variants together with newly built antibodies against Spike protein as well as the highly specific blood flow conditions in the central venous sinus of the brain may result in the rare but severe events after vaccination observed with ADZ1222/Vaxzevria. Noteworthy, the vaccine from Johnson & Johnson appears to carry fewer splice donor sequences, especially SD<sub>506</sub> and SD<sub>3614</sub> (see **Table 1B and 1C**), which are the strongest predicted splice donor sites in the AZD1222 sequence (see Fig. 1A). This may explain the ~ 10-fold lower incidence of severe side effects with the Johnson & Johnson vaccine when compared to the AZD1222 vaccine.

In principle, such thromboses may occur in any site of the human body where endothelial cells express ACE2. Soluble Spike proteins which still exhibit the important core portion of the S1 domain (R319-F551) will be able to bind these receptors. When the immune system starts to produce antibodies against the Spike protein, the endothelial cells will not only bind the soluble Spike protein variants, but would also be decorated with the newly formed antibodies. This will give rise to strong inflammatory reactions either by ADCC (antibody dependent cell-mediated cytotoxicity) or CDC (complement dependent cytotoxicity) occurring in these vessels at various sites where such soluble Spike protein variants accumulate. In ADCC, NK cells could be recruited via their surface receptors (CD16 or CD32) to increase inflammation. During CDC, the C1q protein may recognize the local accumulation of Fc receptors to attract C3b, and thus, will recruit erythrocytes that express the CR1 receptor on their surface or will cause cell damage and inflammation via the normal complement cascade (18).

Several questions are remaining, e.g. why does this happen preferentially in the sinus of the central nervous system? The involvement of central nervous system sinuses (besides other sites in the body) could be explained by the non-unidirectional blood flow (depending on body posture or when sleeping) in these venous vessels that are missing the usual venous valves. Thus, the residence time of soluble Spike protein in this area of our body and the possibility to bind to endothelial cells expressing ACE2 is *per se* increased. This simply increases the probability of adverse immune reactions at these special places of our venous vessel system.

Therefore, we propose a pathological disease mechanism that is depicted in Fig. 1D. On one hand, the recently described VITT mechanism is based on the artificial activation of PF4 by adenoviral proteins or DNA molecules, which can similarly to heparin, act as a poly-anion to mediate PF4 activation. In patients that exhibit a high load of auto-antibodies against PF4, this may cause the observed thrombocytopenia (Fig. 1D, left side). The other side of the pathological disease mechanism is depicted as well (Fig. 1D, right side). Based on our splicing data, membrane-anchored and soluble Spike protein variants are produced after the vaccination procedure. When the immune system now starts the production of anti-Spike antibodies (days 4–16), these antibodies will recognize the membrane-anchored as well as soluble Spike proteins. However, the soluble fraction of Spike protein variants is disseminated throughout the body and concentrates at various sites of those endothelial cells expressing the ACE2 surface protein. These ACE-bound Spike protein variants will become targets of the newly produced antibodies and will cause an ADCC/CDC-mediated inflammatory reaction involving several immune cells. It may also cause the already described NET reaction of neutrophils which has already been shown to induce thromboembolic events (19–21). It also may cause a strong inflammatory reaction by destroying endothelial cells at those sites, explaining the phenomenon of the rarely observed capillary leak syndrome (Clarkson-disease; 22).

If this mechanism is correct, we may also speculate about some other peculiar observations. What about the age-distribution of thromboembolic events and why this happens more in younger people than in older people? As a matter of fact, elderly people may use more frequently, or even on a daily basis, drugs that decrease blood coagulation and inflammatory events (ASS100 or Vit K inhibitors such as Macurmar). This could be a reason why elderly people suffer much less from these thromboembolic events. In addition, an older immune system displays more immune senescence which means that young people exhibit stronger immune reactions than elderly people, and women even stronger than men. All this would imply a higher incidence in young women when compared to men or elderly people (23,24).

Another speculation could potentially explain why this happens so rarely in first-vaccinated people. Neutralizing antibodies prevent binding of soluble Spike proteins to endothelial cells that express ACE2 on the cell surface. Moreover, only ACE2-bound soluble Spike protein will cause the above mentioned ADCC reactions. Therefore, it could well be that due to specific MHC combinations some vaccinated people are unable to produce neutralizing antibodies against Spike. Such an immunological situation in combination with the presence of auto-antibodies against PF4 could potentially explain the rather low incidence of central venous sinus thrombosis - or other thromboembolic events, although all vaccinated persons should face the same pro-inflammatory situation due to the presence of soluble Spike protein when vaccinated with a vector based vaccine.

Based on our findings, we strongly suggest that the Spike open reading frames – wildtype or codon-optimized - in vector-based vaccines has to be re-optimized to avoid unintended splice reactions and to increase the safety of these pharmaceutical products. *Vice versa*, all mRNA-based vaccines should represent safe products, because the delivered mRNA will only be translated into surface antigen, without having any possibility to participate in nuclear splice events.

# Declarations

**Conflict of interests:** The authors declare no conflict of interests

## Acknowledgement

I thank all my lab members for fruitful discussions and their help. I also want to thank Prof. Owen Williams for critical reading of the manuscript. I also thank Dr. Tuna Toptan and Dr. Marek Widera for providing us the cDNA of SARS-CoV-2 (Wuhan isolate). I am also grateful to Roland Zahn (Janssen) for providing the digital Spike gene sequence of the Johnson & Johnson vaccine. This work was supported by a small funding grant of the Corona Task Force of the Goethe-University to RM.

## References

1. Greinacher A, Thiele T, Warkentin TE, Weisser K, Kyrle P, Eichinger S. **A Prothrombotic Thrombocytopenia disorder resembling Heparin-induced Thrombocytopenia following Coronavirus-19 vaccination. Researchsquare doi:** 10.21203/rs.3.rs-362354/v2
2. Greinacher A, Selleng K, Mayerle J, Palankar R, Wesche J, Reiche S, Aebischer A, Warkentin TE, Muenchhoff M, Hellmuth JC, Keppler OT, Duerschmied D, Lothar A, Rieg S, Gawaz MP, Mueller KAL, Scheer Cr, Napp M, Hahnenkamp K, Lucchese G, Vogelgesang A, Flöel A, Lovreglio P, Stufano A, Marschalek R, Thiele T. **Anti-SARS-CoV-2 Spike Protein and Anti-Platelet Factor 4 Antibody Response Induced by COVID-19 Disease and ChAdOx1 nCov-19 vaccination. Researchsquare doi:** 10.21203/rs.3.rs-404769/v1
3. Huang Y, Yang C, Xu XF, Xu W, Liu SW. Structural and functional properties of SARS-CoV-2 spike protein: potential antiviral drug development for COVID-19 *Acta Pharmacol Sin* 2020;41(9):1141–1149.
4. Hoffmann M, Kleine-Weber H, Schroeder S, et al. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor *Cell* 2020;181(2):271–280 e8.
5. Lee Y, Rio DC. **Mechanisms and Regulation of Alternative Pre-mRNA Splicing.** *Annu Rev Biochem.* 2015;**84**:291–323.
6. W. Doerfler, P. Böhm, **Adenoviruses: Model and Vectors in Virus-Host interactions.** Springer-Verlag, Berlin Heidelberg, 2003.
7. Zuallaert J, Godin F, Kim M, Soete A, Saeys Y, De Neve W. 2018. **SpliceRover: interpretable convolutional neural networks for improved splice site prediction,** *Bioinformatics* **34**, 4180–4188.
8. Wang M. and Marín A. 2006. **Characterization and Prediction of Alternative Splice Sites.** *Gene* **366**: 219–227.
9. Kowarz E, Löscher D, Marschalek R. Optimized Sleeping Beauty transposons rapidly generate stable transgenic cell lines. *Biotechnol J.* 2015 Apr;10(4):647–53.

10. Schiedner G, Hertel S, Kochanek S. **Efficient transformation of primary human amniocytes by E1 functions of Ad5: generation of new cell lines for adenoviral vector production.** Hum Gene Ther. 2000 Oct 10;11(15):2105-16.
11. Krutzke L, Allmendinger E, Hirt K, Kochanek S. **Chorioallantoic Membrane Tumor Model for Evaluating Oncolytic Viruses.** Hum Gene Ther. 2020 Oct;31(19–20):1100–1113.
12. Mittereder N, March KL, Trapnell BC. **Evaluation of the concentration and bioactivity of adenovirus vectors for gene therapy.** J Virol. 1996 Nov;70(11):7498 – 509.
13. Lei Y, Zhang J, Schiavon CR, He M, Chen L, Shen H, Zhang Y, Yin Q, Cho Y, Andrade L, Shadel GS, Hepokoski M, Lei T, Wang H, Zhang J, Yuan JX, Malhotra A, Manor U, Wang S, Yuan ZY, Shyy JY. **SARS-CoV-2 Spike Protein Impairs Endothelial Function via Downregulation of ACE2.** bioRxiv [Preprint]. 2020 Dec 4:2020.12.04.409144. doi: 10.1101/2020.12.04.409144. Update in: Circ Res. 2021 Mar 31;: PMID: 33300001; PMCID: PMC7724674.
14. Nuovo GJ, Magro C, Shaffer T, Awad H, Suster D, Mikhail S, He B, Michaille JJ, Liechty B, Tili E. Endothelial cell damage is the central part of COVID-19 and a mouse model induced by injection of the S1 subunit of the spike protein. Ann Diagn Pathol. 2021 Apr;51:151682.
15. Patra T, Meyer K, Geerling L, Isbell TS, Hoft DF, Brien J, Pinto AK, Ray RB, Ray R. SARS-CoV-2 spike protein promotes IL-6 trans-signaling by activation of angiotensin II receptor signaling in epithelial cells. PLoS Pathog. 2020 Dec 7;16(12):e1009128.
16. Amraei R, Rahimi N. COVID-19, Renin-Angiotensin System and Endothelial Dysfunction. Cells. 2020 Jul 9;9(7):1652.
17. Lei Y, Zhang J, Schiavon CR, He M, Chen L, Shen H, Zhang Y, Yin Q, Cho Y, Andrade L, Shadel GS, Hepokoski M, Lei T, Wang H, Zhang J, Yuan JX, Malhotra A, Manor U, Wang S, Yuan ZY, Shyy JY. SARS-CoV-2 Spike Protein Impairs Endothelial Function via Downregulation of ACE 2. Circ Res. 2021 Apr 30;128(9):1323–1326.
18. C. Janeway, Paul Travers, Mark Walport, Mark Shlomchik: Immunologie. 5. Auflage, Spektrum Akademischer Verlag, Heidelberg 2002.
19. Günther Schönrich, Martin J. Raftery, Yvonne Samstag. Devilishly radical NETwork in COVID-19: Oxidative stress, neutrophil extracellular traps (NETs), and T cell suppression. Adv Biol Regul. 2020 Aug; 77: 100741.
20. Veras FP, Pontelli MC, Silva CM, Toller-Kawahisa JE, de Lima M, Nascimento DC, Schneider AH, Caetité D, Tavares LA, Paiva IM, Rosales R, Colón D, Martins R, Castro IA, Almeida GM, Lopes MIF, Benatti MN, Bonjorno LP, Giannini MC, Luppino-Assad R, Almeida SL, Vilar F, Santana R, Bollela VR, Auxiliadora-Martins M, Borges M, Miranda CH, Pazin-Filho A, da Silva LLP, Cunha LD, Zamboni DS, Dal-Pizzol F, Leiria LO, Siyuan L, Batah S, Fabro A, Mauad T, Dolhnikoff M, Duarte-Neto A, Saldiva P, Cunha TM, Alves-Filho JC, Arruda E, Louzada-Junior P, Oliveira RD, Cunha FQ. SARS-CoV-2-triggered neutrophil extracellular traps mediate COVID-19 pathology. J Exp Med. 2020 Dec 7;217(12):e20201129.

21. Middleton EA, He XY, Denorme F, Campbell RA, Ng D, Salvatore SP, Mostyka M, Baxter-Stoltzfus A, Borczuk AC, Loda M, Cody MJ, Manne BK, Portier I, Harris ES, Petrey AC, Beswick EJ, Caulin AF, Iovino A, Abegglen LM, Weyrich AS, Rondina MT, Egeblad M, Schiffman JD, Yost CC. Neutrophil extracellular traps contribute to immunothrombosis in COVID-19 acute respiratory distress syndrome. *Blood*. 2020 Sep 3;136(10):1169–1179.
22. Kawabe S, Saeki T, Yamazaki H, Nagai M, Aoyagi R, Miyamura S. Systemic capillary leak syndrome. *Intern Med*. 2002 Mar;41(3):211–5.
23. Klein, S., Flanagan, K. Sex differences in immune responses. *Nat Rev Immunol* **16**, 626–638 (2016).
24. Ortona E, Pierdominici M, Rider V. Editorial: Sex Hormones and Gender Differences in Immune Responses. *Front Immunol*. 2019 May 9;10:1076.

## Tables

**Table 1A: splice donor site prediction in wildtype Spike ORF**

Position	pot. splice donor sites	Score	Type
454-473	TGGATGGAAA•GTGAGTTCAG	0,187	+1
541-560	GGAAAACAGG•GTAATTTCAA	0,151	+1
894-911	GAAACAAAAGT•GTACGTTGAA	0,565	+1
1323-1342	TGATTCTAAG•GTTGGTGGTA	0,357	0
1996-2015	ATTGGTGCAG•GTATATGCCG	0,707	+1
3296-3317	GCACACACTG•GTTTGTAAACA	0,160	+2

Consensus MAG•GTNNGTG

Position	pot. splice acceptor sites	Score	Type
760-779	TCTTCTTCAG•GTTGGACAGC	0,667	+2
1183-1202	GTCTATGCAG•ATTCATTTGT	0,772	+2
1830-1849	TCTTTATCAG•GATGTAACT	0,321	0
1906-1925	TATTCTACAG•GTTCTAATGT	0,226	+2
2815-2834	TCTTCCACAG•CAAGTGCAGT	0,395	+2

Consensus YYYYYYNCAG•GGTN

**Table 1B: splice donor site prediction in codon-optimized Spike ORF of AZ**

Position	pot. splice donor sites	Score	Type
497-516	GCACCTTCGA•GTACGTGTCC	0,990	+2
563-582	ACCTGCGCGA•GTTCTGTTT	0,200	+2
1175-1194	TCACAAACGT•GTACGCCGAC	0,187	+2
1209-1228	GGGAGATGAA•GTGCGGCAGA	0,162	0
1812-1831	CTCCAACCAG•GTGGCCGTGC	0,547	0
2331-2350	CACCCAAGAG•GTGTTCCGCC	0,202	+2
2949-2968	ACTGGACAAG•GTGGAAGCCG	0,378	0
2961-2980	GGAAGCCGAG•GTGCAGATCG	0,265	0
3083-3102	AGATGTCTGA•GTGTGTGCTG	0,516	+2
3296-3315	GCACCCATTG•GTTCTGTACC	0,357	+2
3452-3471	AACTGGATAA•GTAATTTAAG	0,296	+2
3555-3574	GCTGAACGAG•GTGGCCAAGA	0,309	0
3605-3624	AACTGGGGAA•GTACGAGCAG	0,813	+2

Position	pot. splice acceptor sites	Score	Type
170-189	CTTTCTTCAG•CAACGTGACC	0,299	+1
708-727	CCGGTTTCAG•ACACTGCTGG	0,201	0
804-823	CTACCTGCAG•CCTAGAACCT	0,276	0
1011-1030	CTTCGGCGAG•GTGTTCAATG	0,206	0
1561-1580	CCTGCCACAG•TGTGCGGCC	0,265	+2
1910-1929	CCACCGGCAG•CAATGTGTTT	0,198	+1
1923-1942	TGTGTTTCAG•ACCAGAGCCG	0,260	0
2153-2172	TCACCATCAG•CGTGACCACA	0,192	+1
2676-2695	CGCTCTGCAG•ATCCCTTTCG	0,175	0
2694-2713	TGCTATGCAG•ATGGCCTACC	0,782	0

3205-3224 CCCGCTCAAG•AGAAGAATTT 0,287 +2

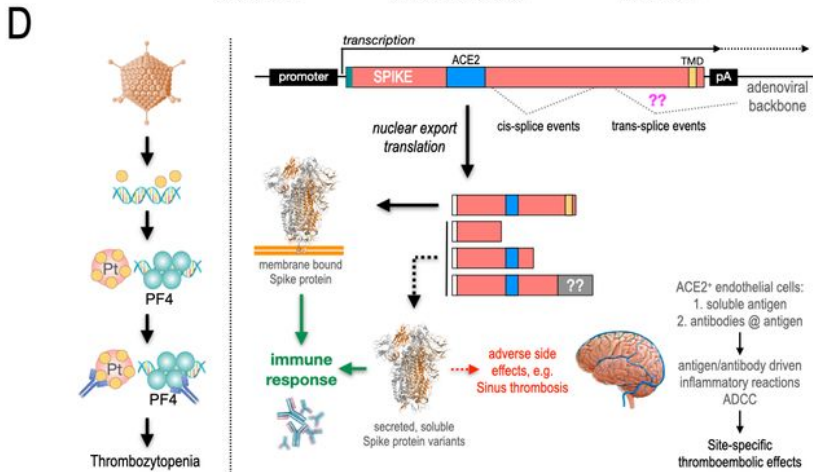
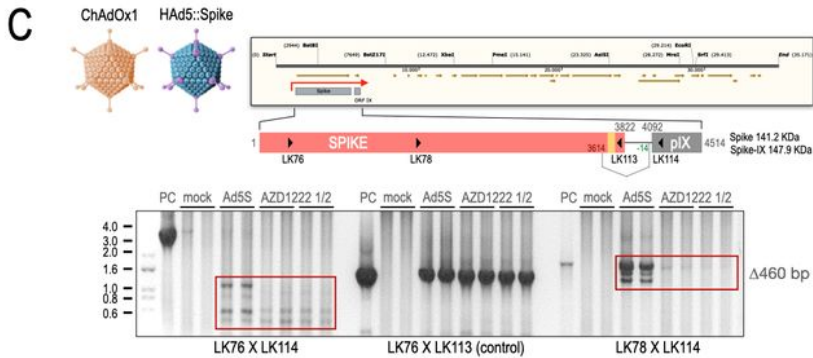
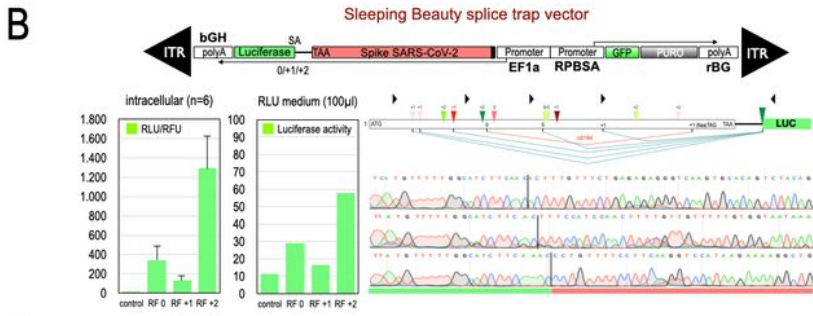
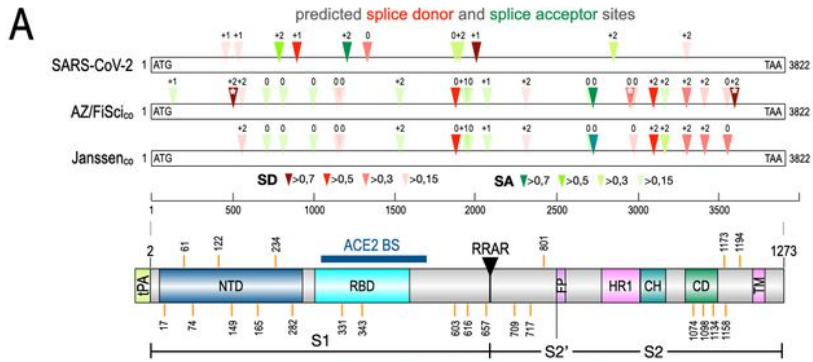
Table 1C: splice donor site prediction in codon-optimized Spike ORF of J&J

Position	pot. splice donor sites	Score	Type
563-582	ACCTGCGCGA•GTTCGTGTTT	0,150	+2
1175-1194	TCACAAACGT•GTACGCCGAC	0,180	+2
1209-1228	GGGAGATGAA•GTGCGGCAGA	0,207	0
1812-1831	CTCCAACCAG•GTGGCCGTGC	0,547	0
2331-2350	CACCCAAGAG•GTGTTGCGCC	0,202	+2
2949-2968	ACTGGACAAG•GTGGAAGCCG	0,378	0
2961-2980	GGAAGCCGAG•GTGCAGATCG	0,194	0
3083-3102	AGATGTCTGA•GTGTGTGCTG	0,495	+2
3296-3315	GCACCCATTG•GTTCGTGACC	0,392	+2
3452-3471	AACTGGATAA•GTACTTTAAG	0,436	+2
3555-3574	GCTGAACGAG•GTGGCCAAGA	0,335	0

Position	pot. splice acceptor sites	Score	Type
708-727	CCGGTTTCAG•ACACTGCTGG	0,172	0
804-823	CTACCTGCAG•CCTAGAACCT	0,205	0
1011-1030	CTTCGGCGAG•GTGTTCAATG	0,207	0
1561-1580	CCTGCCACAG•TGTGCGGCC	0,265	+2
1910-1929	CCACCGGCAG•CAATGTGTTT	0,194	+1
1923-1942	TGTGTTTCAG•ACCAGAGCCG	0,267	0
2153-2172	TCACCATCAG•CGTGACCACA	0,163	+1
2676-2695	CGCTCTGCAG•ATCCCTTTC	0,175	0
2694-2713	TGCTATGCAG•ATGGCCTACC	0,782	0
3205-3224	CCCGCTCAAG•AGAAGAATTT	0,181	+2

ACE2 binding domain: encoded by nucleotides 998-1720 (R319-F551)

## Figures



**Figure 1**

1A: Splice site prediction within the Spike open reading frames Splice site prediction was carried out by using SpliceRoover. Splice donor sites are given in red, splice acceptor sites in green. SpliceRoover calculates splice sites with a score between 0 to 1, but only splice site with  $>0,15$  were displayed. All sites were displayed in 4 colors as indicated. All splice sites are numbered with "0", "+1" or "+2", to indicate how the open reading frame is disrupted. Therefore, all splice events between splice site with identical



numbers will be in-frame, while all splice reactions between unequal numbers will result in out-of frame fusions. Below, the protein domain structure of Spike is displayed to immediately understand what domains could be spliced-out or deleted. White asterisks are marking 3 splice donor sites that are present in the codon optimized Spike reading frame of AZ, but not of the J&J transgene.

1B: Splice trap and Luciferase experiments in HEK293T cells The 3 different splice traps were cloned into the pSBbi-GP vector. This vector encodes two polycistronic transcripts, one encoding GFP and Puromycin resistance (from left to right), while the other encodes the full-length Spike gene fused to an artificial intron and an ATG-deleted Luciferase gene. Splice events to the Luciferase will result in “in-frame” and “out-of-frame” Spike-Luciferase fusion transcripts. Since we used 3 different constructs, all successful in-frame splice events will result in Luciferase activity which could then be measured in- or outside of the cell. Several splice events have already been validated by sequencing and demonstrated that the predicted splice donor site were indeed used. Moreover, gene-internal out-of-frame splicing was observed, as well as splicing in a region of the Spike gene that occurred downstream of the encoded ACE2-binding domain.

1C: Ad5-Spike and AZD1222 transduction experiments in Hela cells Two adenoviral constructs were used to infect Hela cells (HAd5::Spike and AZD1222). After RNA isolation specific RT-PCR experiments were performed. The transgene in the HAd5::Spike vector is 99% identical to the AZ and J&J sequence and contains all splice donor sites of the AZ vector, and all splice acceptors of the J&J vector. The potential splice events between the Spike transcript and the pIX transcript can be explained either by transcriptional read-through, or by a trans-splice event between both transcripts. The size change were in the range of about 460 bp and would predict a splicing between a nearly perfect splice donor site in Spike (nt 3614) and the 5'-NTR of pIX (nt 4075). Direct Sequencing of the resulting PCR band confirmed this assumption.

1D: Predicted disease mechanism We predict for the VIC19M syndrome that rare side effects are caused by two independent mechanism: one is the postulated VITT mechanism that is accompanied by the thrombozytopenic situation, while the other mechanism is based on the availability of soluble Spike protein variants in a systemic fashion. We also postulate that soluble Spike protein variants are able to bind to ACE2 expressing endothelial cells which in turn may trigger an ADCC-like mechanisms that result in thromboembolic events (CVST and SVT). Thus, the presence of auto-antibodies against PF4 in combination with soluble Spike protein variants that bind to ACE2 expressing endothelial cells may explain the rare severe side effects (estimated 1 in 120,000 vaccinated persons in Europe).