Correspondence

Alternative splicing of ACE2 possibly generates variants that may limit the entry of SARS-CoV-2: a potential therapeutic approach using SSOs

Sayeed ur Rehman1 and Mohammad Tabish2

1Department of Biochemistry, School of Chemical and Life Sciences, Jamia Hamdard, New Delhi 110062, India; 2Department of Biochemistry, Faculty of Life Sciences, A.M.U, Aligarh, U.P. 202002, India

Correspondence: Sayeed ur Rehman (sayeed125@gmail.com)

Angiotensin-converting enzyme 2 (ACE2) plays an essential role in maintaining the balance of the renin–angiotensin system and also serves as a receptor for the SARS-CoV-2, SARS-CoV, and HCoV-NL63. Following the recent outbreak of SARS-CoV-2 infection, there has been an urgent need to develop therapeutic interventions. ACE2 is a potential target for many treatment approaches for the SARS-CoV-2. With the help of bioinformatics, we have predicted several novel exons of the human ACE2 gene. The inclusion of novel exons located in the 5′UTR/intronic region in the mature transcript may remove the critical ACE2 residues responsible for the interaction with the receptor-binding domain (RBD) of SARS-CoV-2, thus preventing their binding and entry into the cell. Additionally, inclusion of a novel predicted exons located in the 3′UTR by alternative splicing may remove the C-terminal transmembrane domain of ACE2 and generate soluble ACE2 isoforms. Splice-switching antisense oligonucleotides (SSOs) have been employed effectively as a therapeutic strategy in several disease conditions. Alternative splicing of the ACE2 gene could similarly be modulated using SSOs to exclude critical domains required for the entry of SARS-CoV-2. Strategies can also be designed to deliver these SSOs directly to the lungs in order to minimize the damage caused by SARS-CoV-2 pathogenesis.

The emergence of the highly pathogenic human coronavirus (SARS-CoV-2) is a major challenge faced by today’s world. SARS-CoV-2 has infected more than 5 million people across the globe leading to more than 330,000 deaths (as on 22/05/2020). It is related to the earlier reported Middle East respiratory syndrome coronavirus (MERS-CoV) and severe acute respiratory syndrome coronavirus (SARS-CoV). Coronaviruses (CoVs) belong to the Coronavirinae subfamily that are named for the presence of crown-like spikes on their surface. Phylogenetic clustering further classifies coronaviruses into four groups (α, β, γ, and δ CoVs). These viruses are known to cause mild-to-severe respiratory or intestinal infections in animals and humans [1]. Human CoVs like HCoV-229E, HCoV-NL63, HCoV-OC43, and HCoV-HKU1 are among the seven different strains that infect humans, causing mild and self-resolving infection [2]. However, SARS-CoV, MERS-CoV, and SARS-CoV-2 cause severe respiratory infections in humans [3,4]. Coronaviruses have four structural proteins, namely envelope (E), membrane (M), nucleocapsid (N), and spike (S) proteins. S proteins play a critical role in assisting the entry of the virus into the host cell by interacting with the receptors present on the host cell [5]. Recently, angiotensin-converting enzyme 2 (ACE2) was recognized as the host cell receptor (Figure 1) for the novel coronavirus (2019-nCoV/ SARS-CoV-2) [6]. ACE2 is also the primary receptor for SARS-CoV and HCoV-NL63 [7,8]. Notably, the S protein of SARS-CoV-2 binds human ACE2 more efficiently than SARS-CoV, reflecting its increased ability to transmit from person to person [9,10]. The critical residues of ACE2 that interact with the receptor-binding domain (RBD) of SARS-CoV-2 are Gln 24, Asp 30, His 34, Tyr 41, Gln 42, Met 82, Lys 353, and Arg 357 [11].
Figure 1. Structure and domains of human ACE2

SWISS-Model was used to prepare the 3D model of human ACE2 using PDB: 6m1d.1.B as template. (A) Spike protein of SARS-CoV-2 binds to an ACE2 receptor present on the host cell. The interaction between the viral protein and human ACE2 is mediated by amino acid residues marked red (Gln 24, Asp 30, His 34, Tyr 41, Gln 42, Met 82, Lys 353, and Arg 357). The first coding exon, E1, encodes most of these residues. (B) There are different domains in ACE2. Amino acid 1-17 makes the signal peptide, while the amino acid 18-740 constitutes the ectodomain. The transmembrane domain is encoded by E17 and includes amino acid 741-763. Rest of the amino acids at C terminal code for the cytoplasmic domain (764-805).

The expression and distribution of the ACE2 receptor in the human body may indicate the potential entry route and infection risk for SARS-CoV-2. ACE2 is expressed in several tissues; however, high expression is reported in cardio-renal and gastrointestinal tissues [12]. More recently, single-cell RNA seq data revealed high expression of ACE2 in type II alveolar cells of lungs, cells of kidney proximal tubule, and bladder urothelial cells. Additionally, absorptive enterocytes from colon and ileum, esophagus upper and stratified epithelial cells, cholangiocytes, and myocardial cells also show high expression of ACE2 [13–16]. Epithelial cells of the oral cavity were also found to be highly enriched in ACE2, making them a high-risk route for SARS-CoV-2 infection [17].

ACE2 is a monacarboxypeptidase that plays an essential role in maintaining the balance of the renin–angiotensin system [18]. ACE2, a homolog of ACE, is a type I transmembrane protein and is known to play a beneficial role in diabetes, hypertension, and cardiovascular diseases [19,20]. ACE2 is involved in the cleavage of Angiotensin I (Ang-I) to Ang 1-9 and Angiotensin II (Ang-II) to Ang 1-7 [21]. ACE2 can also cleave several other peptides that are unrelated to the renin–angiotensin system [22]. Ang 1-7 exerts anti-inflammatory and antifibrotic action, stimulates the release of nitric oxide, and reduces blood pressure. In the absence of ACE2, Ang-II causes an increase in blood pressure, oxidative stress, inflammation, and fibrosis [23]. The deficiency of ACE2 causes impaired cardiac contractility and leads to the up-regulation of hypoxia-induced genes, suggesting a link with myocardial ischemia [21]. Since ACE2 is the primary route of Ang-II removal and generation of Ang (1-7) in the heart, during SARS-CoV-2 infection, virus-induced internalization and loss of ACE2 are expected to compromise the cardiac function and worsen the condition of patients with underlying CVD [23,24].
Figure 2. Prediction of novel exons that may cause variation at N-terminal due to alternative splicing
(A) The Human ACE2 gene contains 18 coding exons that are marked as E1-E18. With the help of computational tools, the 5′ UTR, 3′ UTR, and intronic regions of ACE2 gene (±50 kb) were analyzed for the putative novel exons that may participate in alternative splicing. With the help of FGENESH, FEX, and manual curation, five novel exons (N1–N5) were predicted. These exons can participate in alternative splicing and serve as an alternate first exon. (B) Prediction of novel exons that may cause variation at C-terminal due to alternative splicing. Exons were also predicted in the 3′ UTR. Exon N6, N7, and N8 were predicted and can serve as an alternate last exon. Blue rectangular box represents reported exon (E1–E18) while the colored box represents predicted exon (N1–N8). Solid line represents introns and the dashed lines represent the splicing pattern. Size of introns and exons are mentioned in nucleotides.

ACE2 is expressed in pancreatic acinar and islet cells [25]. ACE2, both dependent and independent of Ang (1-7), is linked to β-cell function and glucose control [26,27]. Loss of ACE2 causes a reduction in GLUT4 and increased insulin resistance, while the introduction of recombinant ACE2 reduces the progression of diabetic nephropathy [28,29]. The anti-inflammatory and antioxidant role of ACE2/Ang (1-7) system protects lung tissues [30] and a decrease in the expression of ACE2 in diabetes mellitus patients predisposes them to severe lung injury on the infection [20,31]. Indeed, people with diabetes and obesity are more prone to hospitalization due to SARS-CoV-2 infection [32].

ACE2 is an extensively studied target for developing a potential treatment for the SARS-CoV-2. Zhang and group compiled a series of approaches used to target ACE2 and its related pathways [10]. One of the strategies is to block the ACE2 receptor using antibody or small molecules. Also, ACE2 receptor-mediated entry of the virus involves the role of transmembrane protease, serine 2 (TMPRSS2). Inhibitors of TMPRSS2 may serve as a promising candidate for targeting SARC-CoV-2 [33]. The soluble form of ACE2 is also being evaluated as an approach to target SARS-CoV-2 [34]. The soluble ACE2 lacks the transmembrane domain and can compete with the membrane-bound ACE2, limiting the binding of viral particles to the surface-bound, full-length ACE2. This may restrict the entry of SARS-CoV-2 and other coronaviruses using ACE2 as the host receptor. Short ACE2 variants are also reported to have improved pharmacokinetic properties and can undergo glomerular filtration and tubular uptake, thus serving as a potential therapeutic approach in the treatment of acute kidney injury [35].

Alternative splicing plays an essential role in human development and is also linked to several diseases [36]. The isoforms arising due to alternative splicing may have different structure and functions [37]. Alternative splicing is regulated by RNA binding proteins that dictate exon inclusion or skipping. Functional characterization of alternatively spliced isoforms helps in understanding the diverse roles of target protein in the cell. The human ACE2 gene contains 18 coding exons that encode a protein containing 805 amino acids. The amino acids encoded by the first exon includes the signal peptide as well as several critical residues responsible for the binding of the SARS-CoV-2
Table 1 Nucleotide sequence of the predicted novel exons with the conceptually translated sequence

<table>
<thead>
<tr>
<th>Exons</th>
<th>Nature of exon</th>
<th>Nucleotide sequence (5'-3')</th>
<th>Amino acid sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>N1</td>
<td>First</td>
<td>ATGGAGTTTCCACATGGGCTGGGCTGGTCTGAAACCCGACCCAG</td>
<td>MEFHHVGRAGLELPTS</td>
</tr>
<tr>
<td>N2</td>
<td>First</td>
<td>ATGAGGAGTTCCTCCAGGATGAGCCG</td>
<td>MRVPP</td>
</tr>
<tr>
<td>N3</td>
<td>First</td>
<td>ATGCCCTCTCTTTAATACACATGAACACCTACACACAG</td>
<td>MPPLYVNQTS</td>
</tr>
<tr>
<td>N4</td>
<td>First</td>
<td>ATGCCCTCTCTTTTACGTGAACCTTCAAACTTCTGAT</td>
<td>MPFLYVNQTS</td>
</tr>
<tr>
<td>N5</td>
<td>First</td>
<td>ATGCCCTCTCTTTTACGTGAACCTTCAAACTTCTGAT</td>
<td>MPPLYVNQTS</td>
</tr>
<tr>
<td>N6</td>
<td>Last</td>
<td>GCCAACTCCACTCTGGCAAAATTTGCGGCTAGGAGGTATCCATGGAAGAGGGCTGAA</td>
<td>ANSTLGGKDAHSLERLRAENP</td>
</tr>
<tr>
<td>N7</td>
<td>Last</td>
<td>GCTGGAGTGACAGGACGCCCTCCGGCTCGCTCCGGTGAACCCGACCCAG</td>
<td>WSAMARPRLTATSVSWVQAILLPQS</td>
</tr>
<tr>
<td>N8</td>
<td>Last</td>
<td>GGCGGCTGCAATGGGATCCCTCTCTCCGGCATGAGATAAAGAATCAGTGCTGGAAAACTCACTC</td>
<td>KSA-</td>
</tr>
</tbody>
</table>

Identification of novel exons using bioinformatics

The genomic sequence of the human ACE2 gene was analyzed using gene finding tool (FGENESH) and exon finding tool (FEX) [40–42]. After manual curation, several novel exons in the sequence were predicted that could participate in alternative splicing with the reported exons. As seen in Figure 2, novel exons were predicted in the 5'UTR, intronic regions as well as in the 3'UTR. The splicing pattern of predicted novel exons N1–N5 is shown in Figure 2A. Each novel exon, N1, N2, and N3, is capable of replacing exons E1, E2, and E3 from the mature transcript and splicing directly with exon E4. Similarly, novel exon N4 may splice with E6 and N5 with exon E9, replacing all upstream exons, resulting in a truncated ACE2. Alternatively spliced transcripts containing these novel exons (N1–N5) would, in theory, code for isoforms lacking the N-termini required to bind the spike protein of the SARS-CoV-2 and thus may limit the entry of the virus in the host cell. The N-terminal amino acid sequences encoded by these novel predicted isoforms were aligned with the reported isoform using Clustal Omega (Supplementary Figure S1).

Novel exons were also predicted in the 3'UTR of the ACE2 gene (Figure 2B). These exons can function as an alternate last exon, and their inclusion in the mature transcript would lead to the termination of the protein with an alternate C-terminus. Exon E14 can splice directly with predicted exon N6 or N8 and hence lead to the removal of exons E15–E18 from the mature transcript. Similarly, exon E13 can splice with exon N7, causing the removal of E14–E18 from the mature transcript. Inclusion of exon N6, N7, and N8 would replace several coding exons from the 3’ end of the transcript, resulting in the loss of the majority of C-terminal residues, more importantly, exon E17 that encodes the transmembrane domain required for ACE2 membrane-association. Loss of this domain would potentially render the ACE2 isoform soluble. The amino acid sequences encoded by these predicted alternative last exons are shown in Supplementary Figure S2. Complete sequences of the predicted exons are given in Table 1, along with their amino acid sequences.

Targeting SARS-CoV-2 by splice-switching antisense oligonucleotides

Splice-switching oligonucleotides (SSOs) are short synthetic antisense sequences that are allowed to base-pair with the target sequence of the pre-mRNA in the nucleus, thus altering the splicing pattern [43]. SSOs are designed to base-pair with the target intronic or exonic regions to limit access to splicing factors, thus modulating the splicing pattern. There are several examples where splicing therapies have been approved by the FDA [44]. The treatment of
spinal muscular atrophy (SMA) using SSO-induced retention of SMN2 exon 7 and SSO-induced skipping of exon 51 of the Dmd gene are some of the most highlighted cases of splice altering therapies [43,45]. Another oligonucleotide drug, golodirsen, which modulates alternative splicing, was recently approved for Duchenne Muscular Dystrophy treatment [46]. Antisense oligonucleotides to treat Huntington’s disease have also entered clinical trials owing to the advances in the ASO chemical stability and delivery methods [47]. Therapies using antisense oligos for the treatment of Parkinson’s disease, amyotrophic lateral sclerosis (ALS), and Alzheimer’s disease are under clinical trials [47]. The use of SSOs as therapeutic approaches in other disease conditions are discussed by Montes et al. [44].

The use of SSOs in targeting the SARS-CoV-2 can offer a novel approach for therapeutic intervention. SSOs can be designed from regions that may lead to the exclusion or inclusion of exons from the mature transcript. With the help of bioinformatics, existence of novel putative exons were predicted in the 5’UTR, 3’ UTR, and in the intronic regions of the ACE2 gene. These novel exons may participate in the alternative splicing and replace the critically important ACE2 domains required for the entry of SARS-CoV-2. The critical amino acid residues required for viral binding are encoded by different exons of the ACE2 gene. Exon E1 encodes Gln 24, Asp 30, His 34, Tyr 41, and Gln 42, critical for RBD binding. Another important amino acid, Met 82, is encoded by Exon E2, while E8 encodes Lys 353. The Arg 357 residue is encoded by the nucleotides present at the junction of exon E8 and E9.

Inclusion of novel exons N1, N2, N3, N4, or N5 in the mature transcript will lead to the exclusion of exons encoding the N-terminal of protein. The novel exon N1 (encoding 16 amino acids), N2 (encoding 5 amino acids), and N3 (encoding 11 amino acids) are predicted to splice with E4 replacing E1-E3 from the mature transcript. Removal of E1-E3 will replace 146 amino acids at the N-terminal with a short sequence encoded by either N1 or N2 or N3. The novel exon N4 (encoding 12 amino acids) is capable of splicing with E6 and may replace E1–E5 (encoding 232 amino acids) from the mature transcript. Exon N5 (encoding 10 amino acids) is capable of splicing directly with E9 and may remove exon E1-E8 (encoding 356 amino acids) from the mature transcript. Thus, depending on the inclusion of novel exon in the mature transcript, the N-terminal truncation of ACE2 protein will occur. This will remove the major amino acid residues responsible for the binding of the RBD of SARS-CoV-2 and hence may limit their entry into the host cell. SSOs can be designed to modulate the alternative splicing to favor the inclusion of the predicted exons (N1, N2, N3, N4, or N5) in the mature transcript, thus altering the N-terminal of the protein. Similarly, other putative exons, N6, N7, and N8, may act as an alternate last exon, and their SSOs mediated inclusion in the mature transcript will remove the amino acid sequences at C-terminal, including the transmembrane domain. This may lead to the generation of soluble ACE2 isoforms. SSOs can also be designed to exclude exon 17 from the transcript that would lead to the removal of the transmembrane domain from the ACE2 protein (Figure 3). The soluble ACE2 thus formed may bind the SARS-CoV-2 spike protein but may not assist their entry into the cell. Targeting SARS-CoV-2 using splice-switching antisense oligonucleotides could be tried as a stand-alone strategy or in conjunction with other approaches.

Conclusions and perspectives
The SARS-CoV-2 pandemic has caused significant health and economic burden in the world. So far, there are no treatments or therapies to prevent SARS-CoV-2 infection. Several approaches involving different platforms like the use of recombinant proteins, peptides, virus-like particles, live attenuated virus, inactivated virus, viral vectors, and nucleic acid-based therapies are being developed and evaluated [48,49]. There is an urgent need to develop a vaccine against SARS-CoV-2 supported by better scientific understanding in the field of genomics and structural biology and an intense global R&D activity [50].

We propose that using splice switch oligonucleotides, alternative splicing of the ACE2 gene can be modulated, and critical domains of ACE2 responsible for the binding and the entry of SARS-CoV-2 can be removed from the mature transcript. Novel exon predicted in the present study may also provide the nucleotide sequences for designing SSOs. With the recent advancement in the synthesis and delivery of SSOs, this may provide a novel and effective approach to target SARS-CoV-2 pathogenesis.

Competing Interests
The authors declare that there are no competing interests associated with the manuscript.

Abbreviations
ACE2, angiotensin-converting enzyme 2; CoV, Coronavirus; ESE, exonic splicing enhancer; MERS-CoV, Middle East respiratory syndrome coronavirus; RBD, receptor-binding domain; SARS-CoV, severe acute respiratory syndrome coronavirus; SSO, splice-switching antisense oligonucleotide; TMPRSS2, transmembrane protease, serine 2.
Figure 3. Use of splice-switching antisense oligonucleotides to target SARS-CoV-2

SSOs can be designed to target the exonic splicing enhancer (ESE) present in the exon E17. This will lead to the exon skipping and generation of the transcript without E17. The protein encoded by the modulated variant will lack the transmembrane domain and hence remain soluble limiting the entry of SARS-CoV-2 (and SARS-CoV) in the host cell. Similarly, SSOs can also be designed to modulate the splicing of novel exons.

References


