PORTLAND PRESS

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 Rehman[1] and Mohammad Tabish[2]

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

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

OPEN ACCESS

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].
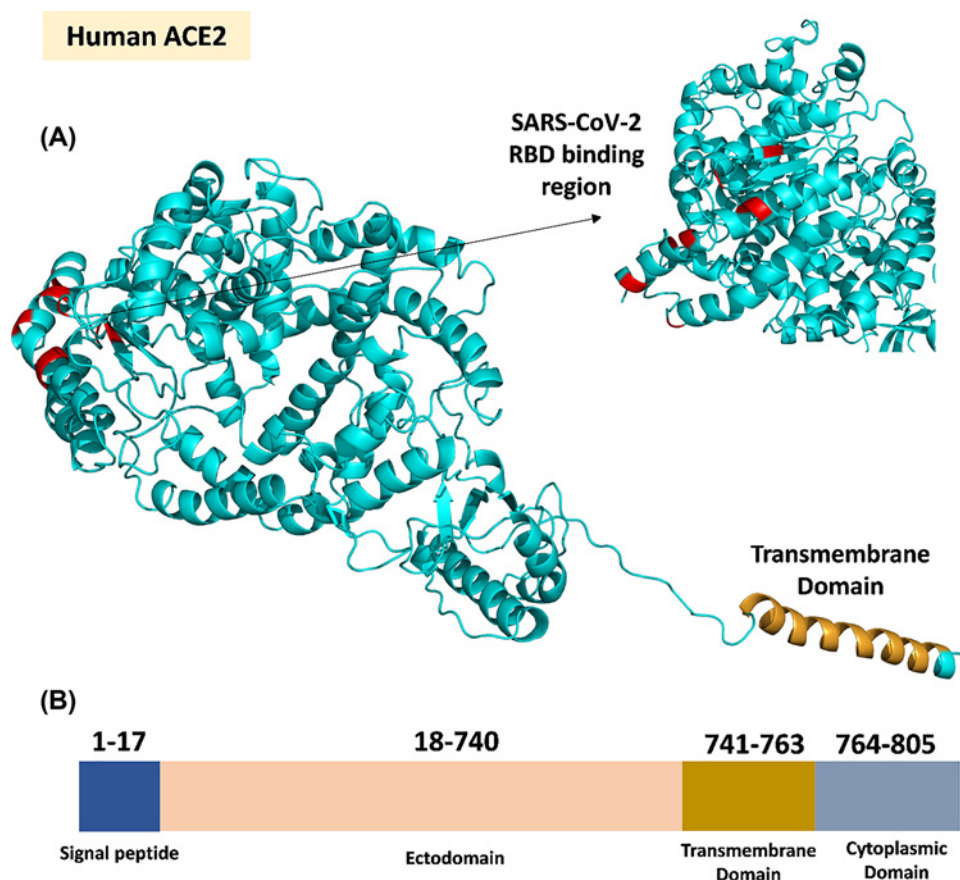
Figure 1. Structure and domains of human ACE2
SWISS-Model was used to prepare the 3D model of human ACE2 using PDB: 6m1d.1.B as template. (**A**) Spike protein of SARS–CoV-2 binds to an ACE2 receptor present on the host cell. The interaction between the viral protein and human ACE2 is mediated by amino 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 monocarboxypeptidase 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].
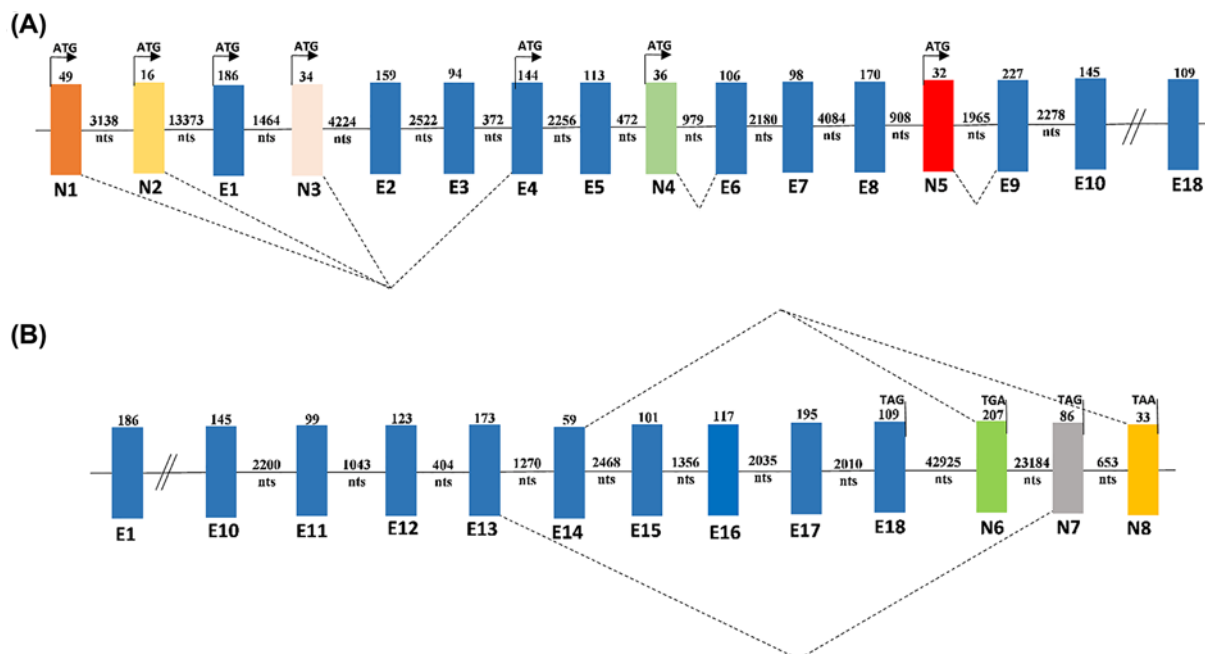
Figure 2. Prediction of novel exons that may cause variation at N-terminal due to alternative splicing
(**A**) The Human ACE2 gene contains 18 coding exons that are marked as E1-E18. With the help of computational tools, the 5′UTR, 3′UTR, and intronic regions of ACE2 gene ($\pm$50 kb) were analyzed for 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

Exons	Nature of exon	Nucleotide sequence (5′-3′)	Amino acid sequence
N1	First	ATGGAGTTTCACCATGTTGGCCGGGCTGGTCTTGAACTCCCGACCTCAG	MEFHHVGRAGLELPTS
N2	First	ATGAGAGTTCCTCCAG	MRVPP
N3	First	ATGCCCCTTAATACACATGAACACCTACACACAG	MPLNTHEHLHT
N4	First	ATGCCTCCCCTTTACGTGAACCTTCAAACTTCTGAT	MPPLYVNLQTSD
N5	First	ATGAGGGAAGCAGGCTGGGACAAAGGAGGGAG	MREAGWDKGG
N6	Last	GCCAACTCCACTCTTGGGAAAAAGTTGGCTGACAGCCATCTTGAAAGATTGAGGGCTGAA AATCCAAGAACTGAGGATCAAGATCTCTCCCCTGTCATAAAACTACATATGGATCTGCCCTT CAGTAGGAAATTCCTAAAAGTCTCCCATGAGATAAAGAATCAGTGCTGGAAAACTCACTC CGATACCACCACCACCACCAAATCATGA	ANSTLGKKLADSHLERLRAENP RTEDQDLSPVIKLHMDLPFSRK FLKVSHEIKNQCWKTHSDTTTT KS-
N7	Last	GCTGGAGTGCAATGGCACGACCTCGGCTCACGGCAACCTCCGTCTCCTGGGTTCAAGC AATTCTCCTGCCTCAGTCTCCCGAGTAG	WSAMARPRLTATSVSWVQAILLPQSPE-
N8	Last	GGGCCTGTAAATGGAATTCCTGCATTGCTCTAA	GPVNGIPALL-

FGENESH: http://www.softberry.com/berry.phtml?topic=fgenesh&group=programs&subgroup=gfind
FEX: http://www.softberry.com/berry.phtml?topic=fex&group=programs&subgroup=gfind

spike protein. The transmembrane domain of ACE2 includes amino acid 741-764 (encoded by exon E17) necessary for membrane attachment [38]. So far, there are two reported isoforms of the human ACE2 that differ in the coding region (Uniprot ID: Q9BYF1-1 and Q9BYF1-2). The larger variant with 18 coding exons encodes a full-length protein of 805 amino acids while the smaller variant gets terminated after coding exon 12 and encodes a soluble isoform of 555 amino acids. An alternative 5′-untranslated exon was also identified in an earlier study [39]. Keeping in mind the importance of alternative splicing, we studied the alternative splicing patterns of the ACE2 gene using computational tools.

Identification of novel exons using bioinformatics

The genomic sequence of the human ACE2 gene was analyzed using gene (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.
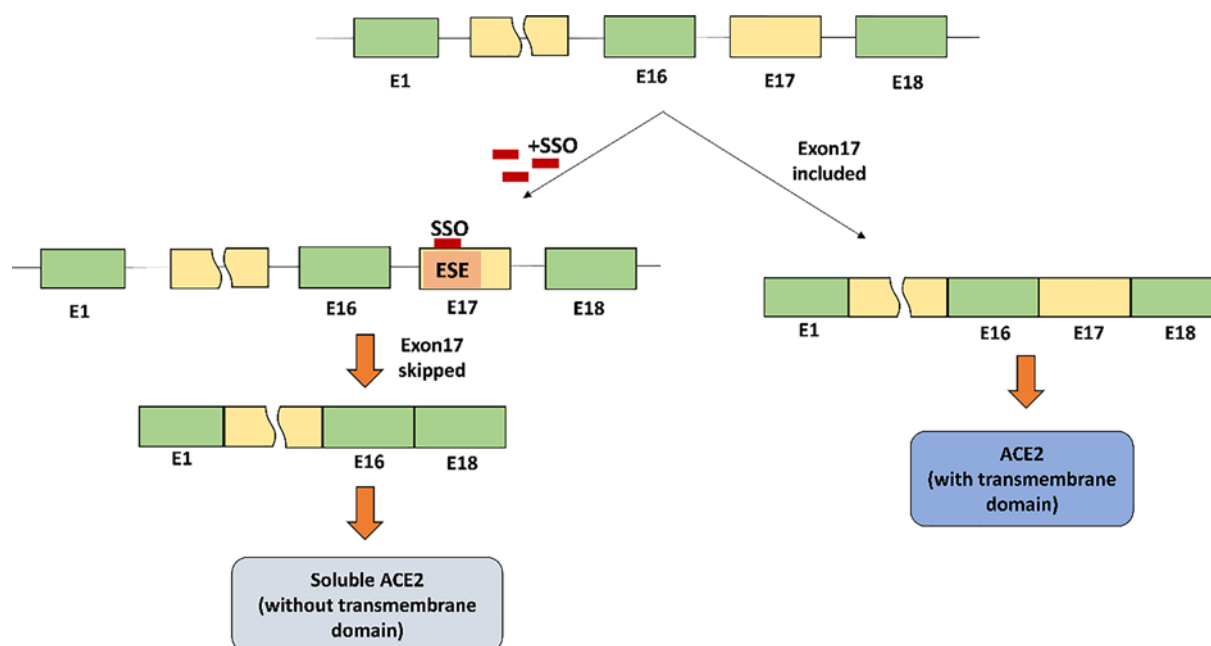
Figure 3. Use of splice-switching antisense oligonucleotides to target SARS-CoV-2
SSOs can be designed to target the exonic splicing enhancer (ESE) present in the exon E17. This will lead to the exon skipping and generation of the transcript without E17. The protein encoded by the modulated variant will 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

1　Cheng, V.C.C., Lau, S.K.P., Woo, P.C.Y. and Yuen, K.Y. (2007) Severe acute respiratory syndrome coronavirus as an agent of emerging and reemerging infection. *Clin. Microbiol. Rev.* **20**, 660–694, https://doi.org/10.1128/CMR.00023-07

2　Madjid, M., Safavi-Naeini, P., Solomon, S.D. and Vardeny, O. (2020) Potential Effects of Coronaviruses on the Cardiovascular System: A Review. *JAMA Cardiol.*, https://doi.org/10.1001/jamacardio.2020.1286

3　Li, W., Hulswit, R.J.G., Kenney, S.P., Widjaja, I., Jung, K., Alhamo, M.A. et al. (2018) Broad receptor engagement of an emerging global coronavirus may potentiate its diverse cross-species transmissibility. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E5135–E5143, https://doi.org/10.1073/pnas.1802879115

4　Chen, N., Zhou, M., Dong, X., Qu, J., Gong, F., Han, Y. et al. (2020) Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet* **395**, 507–513, https://doi.org/10.1016/S0140-6736(20)30211-7

5　Tai, W., He, L., Zhang, X., Pu, J., Voronin, D., Jiang, S. et al. (2020) Characterization of the receptor-binding domain (RBD) of 2019 novel coronavirus: implication for development of RBD protein as a viral attachment inhibitor and vaccine. *Cell Mol. Immunol.*, https://doi.org/10.1038/s41423-020-0400-4

6　Zhou, P., Yang, X.L., Wang, X.G., Hu, B., Zhang, L., Zhang, W. et al. (2020) A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270–273, https://doi.org/10.1038/s41586-020-2012-7

7　Lu, R., Zhao, X., Li, J., Niu, P., Yang, B., Wu, H. et al. (2020) Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet North Am. Ed.* **395**, 565–574, https://doi.org/10.1016/S0140-6736(20)30251-8

8　Li, W., Moore, M.J., Vasilieva, N., Sui, J., Wong, S.K., Berne, M.A. et al. (2003) Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature* **426**, 450–454, https://doi.org/10.1038/nature02145

9　Wan, Y., Shang, J., Graham, R., Baric, R.S. and Li, F. (2020) Receptor recognition by the novel coronavirus from Wuhan: an analysis based on decade-long structural studies of SARS coronavirus. *J. Virol.* **94**, pii: e00127–20, https://doi.org/10.1128/JVI.00127-20

10　Zhang, H., Penninger, J.M., Li, Y., Zhong, N. and Slutsky, A.S. (2020) Angiotensin-converting enzyme 2 (ACE2) as a SARS-CoV-2 receptor: molecular mechanisms and potential therapeutic target. *Intensive Care Med.* **46**, 586–590, https://doi.org/10.1007/s00134-020-05985-9

11　Yan, R., Zhang, Y., Li, Y., Xia, L., Guo, Y. and Zhou, Q. (2020) Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science* **367**, 1444–1448, https://doi.org/10.1126/science.abb2762

12　Harmer, D., Gilbert, M., Borman, R. and Clark, K.L. (2002) Quantitative mRNA expression profiling of ACE 2, a novel homologue of angiotensin converting enzyme. *FEBS Lett.* **532**, 107–110, https://doi.org/10.1016/S0014-5793(02)03640-2

13　Zou, X., Chen, K., Zou, J., Han, P., Hao, J. and Han, Z. (2020) The single-cell RNA-seq data analysis on the receptor ACE2 expression reveals the potential risk of different human organs vulnerable to Wuhan 2019-nCoV infection. *Front. Med.*, https://doi.org/10.1007/s11684-020-0754-0

PORTLAND PRESS

14 Zhang, H., Kang, Z., Gong, H., Xu, D., Wang, J., Li, Z. et al. (2020) Digestive system is a potential route of 2019-nCov infection: an analysis of single-cell coexpression pattern of key proteins in viral entry process. *Gut* **69**, 1010–1018

15 Zhao, Y., Zhao, Z., Wang, Y., Zhou, Y., Ma, Y. and Zuo, W. (2020) Single-cell RNA expression profiling of ACE2, the receptor of SARS-CoV-2. *BioRxiv.*, https://doi.org/10.1101/2020.01.26.919985

16 Chai, X., Hu, L., Zhang, Y., Han, W., Lu, Z., Ke, A. et al. (2020) Specific ACE2 expression in cholangiocytes may cause liver damage after 2019-nCoV infection. *BioRxiv.*, https://doi.org/10.1101/2020.02.03.931766

17 Xu, H., Zhong, L., Deng, J., Peng, J., Dan, H., Zeng, X. et al. (2020) High expression of ACE2 receptor of 2019-nCoV on the epithelial cells of oral mucosa. *Int. J. Oral. Sci.* **12**, 8, https://doi.org/10.1038/s41368-020-0074-x

18 Lambert, D.W., Hooper, N.M. and Turner, A.J. (2008) Angiotensin-converting enzyme 2 and new insights into the renin–angiotensin system. *Biochem. Pharmacol.* **75**, 781–786, https://doi.org/10.1016/j.bcp.2007.08.012

19 Tipnis, S.R., Hooper, N.M., Hyde, R., Karran, E., Christie, G. and Turner, A.J. (2000) A human homolog of angiotensin-converting enzyme cloning and functional expression as a captopril-insensitive carboxypeptidase. *J. Biol. Chem.* **275**, 33238–33243, https://doi.org/10.1074/jbc.M002615200

20 Tikellis, C. and Thomas, M.C. (2012) Angiotensin-converting enzyme 2 (ACE2) is a key modulator of the renin angiotensin system in health and disease. *Int. J. Pept.* **2012**, 1–8, https://doi.org/10.1155/2012/256294

21 Burrell, L.M., Risvanis, J., Kubota, E., Dean, R.G., MacDonald, P.S., Lu, S. et al. (2005) Myocardial infarction increases ACE2 expression in rat and humans. *Eur. Heart J.* **26**, 369–375, https://doi.org/10.1093/eurheartj/ehi114

22 Vickers, C., Hales, P., Kaushik, V., Dick, L., Gavin, J., Tang, J. et al. (2002) Hydrolysis of biological peptides by human angiotensin-converting enzyme-related carboxypeptidase. *J. Biol. Chem.* **277**, 14838–14843, https://doi.org/10.1074/jbc.M200581200

23 South, A.M., Diz, D. and Chappell, M.C. (2020) COVID-19, ACE2 and the cardiovascular consequences. *Am. J. Physiol. Heart Circ. Physiol.* **318**, H1084–H1090, https://doi.org/10.1152/ajpheart.00217.2020

24 Yousif, M.H., Dhaunsi, G.S., Makki, B.M., Qabazard, B.A., Akhtar, S. and Benter, I.F. (2012) Characterization of Angiotensin-(1–7) effects on the cardiovascular system in an experimental model of type-1 diabetes. *Pharmacol. Res.* **66**, 269–75, https://doi.org/10.1016/j.phrs.2012.05.001

25 Liu, F., Long, X., Zhang, B., Zhang, W., Chen, X. and Zhang, Z. (2020) ACE2 Expression in Pancreas May Cause Pancreatic Damage After SARS-CoV-2 Infection. *Clin. Gastroenterol. Hepatol.* pii: S1542–3565(20)30537-1, https://doi.org/10.1016/j.cgh.2020.04.040

26 Shoemaker, R., Yiannikouris, F., Thatcher, S. and Cassis, L. (2015) ACE2 deficiency reduces beta-cell mass and impairs beta-cell proliferation in obese C57BL/6 mice. *Am. J. Physiol. Endocrinol. Metab.* **309**, E621–E631, https://doi.org/10.1152/ajpendo.00054.2015

27 Bindom, S.M., Hans, C.P., Xia, H., Boulares, A.H. and Lazartigues, E. (2010) Angiotensin I-converting enzyme type 2 (ACE2) gene therapy improves glycemic control in diabetic mice. *Diabetes* **59**, 2540–2548, https://doi.org/10.2337/db09-0782

28 Takeda, M., Yamamoto, K., Takemura, Y., Takeshita, H., Hongyo, K., Kawai, T. et al. (2013) Loss of ACE2 exaggerates high-calorie diet-induced insulin resistance by reduction of GLUT4 in mice. *Diabetes* **62**, 223–233, https://doi.org/10.2337/db12-0177

29 Oudit, G.Y., Liu, G.C., Zhong, J., Basu, R., Chow, F.L., Zhou, J. et al. (2010) Human recombinant ACE2 reduces the progression of diabetic nephropathy. *Diabetes* **59**, 529–538, https://doi.org/10.2337/db09-1218

30 Zou, Z., Yan, Y., Shu, Y., Gao, R., Sun, Y., Li, X. et al. (2014) Angiotensin-converting enzyme 2 protects from lethal avian influenza A H5N1 infections. *Nat. Commun.* **5**, 1–7, https://doi.org/10.1038/ncomms4594

31 Wu, C., Chen, X., Cai, Y., Xia, J., Zhou, X., Xu, S. et al. (2020) Risk factors associated with acute respiratory distress syndrome and death in patients with coronavirus disease 2019 pneumonia in Wuhan, China. *JAMA Intern. Med.*, https://doi.org/10.1001/jamainternmed.2020.0994

32 Drucker, D.J. (2020) Coronavirus infections and type 2 diabetes-shared pathways with therapeutic implications. *Endocr. Rev.* **41**, pii: bnaa011, https://doi.org/10.1210/endrev/bnaa011

33 Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S. et al. (2020) SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* **181**, 271–280, https://doi.org/10.1016/j.cell.2020.02.052

34 Batlle, D., Wysocki, J. and Satchell, K. (2020) Soluble angiotensin-converting enzyme 2: a potential approach for coronavirus infection therapy? *Clin. Sci. (Lond.)* **134**, 543–545, https://doi.org/10.1042/CS20200163

35 Wysocki, J., Schulze, A. and Batlle, D. (2019) Novel Variants of Angiotensin Converting Enzyme-2 of Shorter Molecular Size to Target the Kidney Renin Angiotensin System. *Biomolecules* **9**, 886, https://doi.org/10.3390/biom9120886

36 Chen, M. and Manley, J.L. (2009) Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. *Nat. Rev. Mol. Cell Biol.* **10**, 741–754, https://doi.org/10.1038/nrm2777

37 Rehman, S.U., Husain, M.A., Sarwar, T., Ishqi, H.M. and Tabish, M. (2015) Modulation of alternative splicing by anticancer drugs. *Wiley Interdiscip. Rev. RNA* **6**, 369–379, https://doi.org/10.1002/wrna.1283

38 Chen, I.Y., Chang, S.C., Wu, H.Y., Yu, T.C., Wei, W.C., Lin, S. et al. (2010) Upregulation of the chemokine (CC motif) ligand 2 via a severe acute respiratory syndrome coronavirus spike-ACE2 signaling pathway. *J. Virol.* **84**, 7703–7712, https://doi.org/10.1128/JVI.02560-09

39 Itoyama, S., Keicho, N., Hijikata, M., Quy, T., Phi, N.C., Long, H.T. et al. (2005) Identification of an alternative 5′-untranslated exon and new polymorphisms of angiotensin-converting enzyme 2 gene: Lack of association with SARS in the Vietnamese population. *Am. J. Med. Genet. A* **136**, 52–57, https://doi.org/10.1002/ajmg.a.30779

40 Solovyev, V., Kosarev, P., Seledsov, I. and Vorobyev, D. (2006) Automatic annotation of eukaryotic genes, pseudogenes and promoters. *Genome Biol.* **7**, 10.1–10.12, https://doi.org/10.1186/gb-2006-7-s1-s10

41 Aartsma-Rus, A. and Corey, D.R. (2020) The 10th oligonucleotide therapy approved: golodirsen for duchenne muscular dystrophy. *Nucleic Acid Ther.* **30**, 67–70, https://doi.org/10.1089/nat.2020.0845

42 Rehman, S.U., Ashraf, S., Ahamad, S., Sarwar, T., Husain, M.A., Ahmad, P. et al. (2019) Identification of a novel alternatively spliced isoform of antithrombin containing an additional RCL-like loop. *Biochem. Biophys. Res. Commun.* **517**, 421–426, https://doi.org/10.1016/j.bbrc.2019.07.113

43 Ishqi, H.M., Husain, M.A., Rehman, S.U., Sarwar, T. and Tabish, M. (2018) Identification and expression of alternatively spliced novel isoforms of cancer associated MYD88 lacking death domain in mouse. *Mol. Biol. Rep.* **45**, 699–711, https://doi.org/10.1007/s11033-018-4209-5

44 Havens, M.A. and Hastings, M.L. (2016) Splice-switching antisense oligonucleotides as therapeutic drugs. *Nucleic. Acids. Res.* **44**, 6549–6563, https://doi.org/10.1093/nar/gkw533

45 Montes, M., Sanford, B.L., Comiskey, D.F. and Chandler, D.S. (2019) RNA splicing and disease: animal models to therapies. *Trends Genet.* **35**, 68–87, https://doi.org/10.1016/j.tig.2018.10.002

46 Havens, M.A., Duelli, D.M. and Hastings, M.L. (2013) Targeting RNA splicing for disease therapy. *Wiley Interdiscip Rev. RNA* **4**, 247–266, https://doi.org/10.1002/wrna.1158

47 Leavitt, B.R. and Tabrizi, S.J. (2020) Antisense oligonucleotides for neurodegeneration. *Science* **367**, 1428–1429, https://doi.org/10.1126/science.aba4624

48 Le, T.T., Andreadakis, Z., Kumar, A., Román, R.G., Tollefsen, S., Saville, M. et al. (2020) The COVID-19 vaccine development landscape. *Nat. Rev. Drug Discov.* **19**, 305–306, https://doi.org/10.1038/d41573-020-00073-5

49 Lythgoe, M.P. and Middleton, P. (2020) Ongoing clinical trials for the management of the COVID-19 pandemic. *Trends Pharmacol. Sci.* **S0165-6147**, 30070–30075

50 Lurie, N., Saville, M., Hatchett, R. and Halton, J. (2020) Developing Covid-19 vaccines at pandemic speed. *N. Engl. J. Med.* **382**, 1969–1973, https://doi.org/10.1056/NEJMp2005630

Supplementary Figure S1.

```
N3-E4--E9   ------------------------------------------------------------   0
N1-E4--E9   ------------------------------------------------------------   0
E1--E9      MSSSSWLLLSLVAVTAAQSTIEEQAKTFLDKFNHEAEDLFYQSSLASWNYNTNITEENVQ   60
N2-E4--E9   ------------------------------------------------------------   0
N4-E6-E9    ------------------------------------------------------------   0
N5-E9       ------------------------------------------------------------   0


N3-E4--E9   ------------------------------------------------------------   0
N1-E4--E9   ------------------------------------------------------------   0
E1--E9      NMNNAGDKWSAFLKEQSTLAQMYPLQEIQNLTVKLQLQALQQNGSSVLSEDKSKRLNTIL   120
N2-E4--E9   ------------------------------------------------------------   0
N4-E6-E9    ------------------------------------------------------------   0
N5-E9       ------------------------------------------------------------   0

                          ↓E4
N3-E4--E9   ---------------MPLNTHEHLHTGLNEIMANSLDYNERLWAWESWRSEVGKQLRPLY   45
N1-E4--E9   ----------MEFHHVGRAGLELPTSGLNEIMANSLDYNERLWAWESWRSEVGKQLRPLY   50
E1--E9      NTMSTIYSTGKVCNPDNPQECLLLEPGLNEIMANSLDYNERLWAWESWRSEVGKQLRPLY   180
N2-E4--E9   --------------------MRVPPGLNEIMANSLDYNERLWAWESWRSEVGKQLRPLY   39
N4-E6-E9    ------------------------------------------------------------   0
N5-E9       ------------------------------------------------------------   0

                                              ↓E6
N3-E4--E9   EEYVVLKNEMARANHYEDYGDYWRGDYEVNGVDGYDYSRGQLIEDVEHTFEEIKPLYEHL   105
N1-E4--E9   EEYVVLKNEMARANHYEDYGDYWRGDYEVNGVDGYDYSRGQLIEDVEHTFEEIKPLYEHL   110
E1--E9      EEYVVLKNEMARANHYEDYGDYWRGDYEVNGVDGYDYSRGQLIEDVEHTFEEIKPLYEHL   240
N2-E4--E9   EEYVVLKNEMARANHYEDYGDYWRGDYEVNGVDGYDYSRGQLIEDVEHTFEEIKPLYEHL   99
N4-E6-E9    ---------------------------------------MPPLYVNLQTSDIKPLYEHL   20
N5-E9       ------------------------------------------------------------   0
```

```
N3-E4--E9   HAYVRAKLMNAYPSYISPIGCLPAHLLGDMWGRFWTNLYSLTVPFGQKPNIDVTDAMVDQ   165
N1-E4--E9   HAYVRAKLMNAYPSYISPIGCLPAHLLGDMWGRFWTNLYSLTVPFGQKPNIDVTDAMVDQ   170
E1--E9      HAYVRAKLMNAYPSYISPIGCLPAHLLGDMWGRFWTNLYSLTVPFGQKPNIDVTDAMVDQ   300
N2-E4--E9   HAYVRAKLMNAYPSYISPIGCLPAHLLGDMWGRFWTNLYSLTVPFGQKPNIDVTDAMVDQ   159
N4-E6-E9    HAYVRAKLMNAYPSYISPIGCLPAHLLGDMWGRFWTNLYSLTVPFGQKPNIDVTDAMVDQ   80
N5-E9       ------------------------------------------------------------   0

                                                        ↓E9
N3-E4--E9   AWDAQRIFKEAEKFFVSVGLPNMTQGFWENSMLTDPGNVQKAVCHPTAWDLGKGDFRILM   225
N1-E4--E9   AWDAQRIFKEAEKFFVSVGLPNMTQGFWENSMLTDPGNVQKAVCHPTAWDLGKGDFRILM   230
E1--E9      AWDAQRIFKEAEKFFVSVGLPNMTQGFWENSMLTDPGNVQKAVCHPTAWDLGKGDFRILM   360
N2-E4--E9   AWDAQRIFKEAEKFFVSVGLPNMTQGFWENSMLTDPGNVQKAVCHPTAWDLGKGDFRILM   219
N4-E6-E9    AWDAQRIFKEAEKFFVSVGLPNMTQGFWENSMLTDPGNVQKAVCHPTAWDLGKGDFRILM   140
N5-E9       ----------------------------------------MREAGWDKG---GRILM   14
                                                    : :.** *    ****

N3-E4--E9   CTKVTMDDFLTAHHEMGHIQYDMAYAAQPFLLRNGANEGFHEAVGEIMSLSAATPKHLKS   285
N1-E4--E9   CTKVTMDDFLTAHHEMGHIQYDMAYAAQPFLLRNGANEGFHEAVGEIMSLSAATPKHLKS   290
E1--E9      CTKVTMDDFLTAHHEMGHIQYDMAYAAQPFLLRNGANEGFHEAVGEIMSLSAATPKHLKS   420
N2-E4--E9   CTKVTMDDFLTAHHEMGHIQYDMAYAAQPFLLRNGANEGFHEAVGEIMSLSAATPKHLKS   279
N4-E6-E9    CTKVTMDDFLTAHHEMGHIQYDMAYAAQPFLLRNGANEGFHEAVGEIMSLSAATPKHLKS   200
N5-E9       CTKVTMDDFLTAHHEMGHIQYDMAYAAQPFLLRNGANEGFHEAVGEIMSLSAATPKHLKS   74
            ************************************************************

N3-E4--E9   IGLLSPDFQEDN   297
N1-E4--E9   IGLLSPDFQEDN   302
E1--E9      IGLLSPDFQEDN   432
N2-E4--E9   IGLLSPDFQEDN   291
N4-E6-E9    IGLLSPDFQEDN   212
N5-E9       IGLLSPDFQEDN   86
            ************
```

Figure S1. Alignment of the N-terminal amino acid sequences conceptually encoded by predicted exons. The nucleotide sequences were translated using the ExPaSy translate tool, and alignment was done using Clustal Omega.

Supplementary Figure S2.

```
E11--13-N7      REIVGVVEPVPHDETYCDPASLFHVSNDYSFIRYYTRTLYQFQFQEALCQAAKHEGPLHK   60      E11--13-N7      ------------------------------------------------------------   159
E11--E14-N6     REIVGVVEPVPHDETYCDPASLFHVSNDYSFIRYYTRTLYQFQFQEALCQAAKHEGPLHK   60      E11--E14-N6     -----NPRTEDQDLSPVIKLHMDLPFSRKF------------LKVSHEIKNQCWKTHSDT   214
E11--E18        REIVGVVEPVPHDETYCDPASLFHVSNDYSFIRYYTRTLYQFQFQEALCQAAKHEGPLHK   60      E11--E18        MILFGEEDVRVANLKPRISFNFFVTAPKNVSDIIPRTEVEKAIRMSRSRINDAFRLNDNS   240
E11--E14-N8     REIVGVVEPVPHDETYCDPASLFHVSNDYSFIRYYTRTLYQFQFQEALCQAAKHEGPLHK   60      E11--E14-N8     ------------------------------------------------------------   161
                ************************************************************                            

E11--13-N7      CDISNSTEAGQKLFNMLRLGKSEPWTLALENVVGAKNMNVRPLLNYFEPLFTWLKDQNKN   120     E11--13-N7      ------------------------------------------------------------   159
E11--E14-N6     CDISNSTEAGQKLFNMLRLGKSEPWTLALENVVGAKNMNVRPLLNYFEPLFTWLKDQNKN   120     E11--E14-N6     TTTKS-------------------------------------------------------   219
E11--E18        CDISNSTEAGQKLFNMLRLGKSEPWTLALENVVGAKNMNVRPLLNYFEPLFTWLKDQNKN   120     E11--E18        LEFLGIQPTLGPPNQPPVSIWLIVFGVVMGVIVVGIVILIFTGIRDRKKKNKARSGENPY   300
E11--E14-N8     CDISNSTEAGQKLFNMLRLGKSEPWTLALENVVGAKNMNVRPLLNYFEPLFTWLKDQNKN   120     E11--E14-N8     ------------------------------------------------------------   161
                ************************************************************                            

E11--13-N7      SFVGWSTDWSPCWSAMARPRLTATSV------SWVQAILLPQSPE---------------   159     E11--13-N7      -----------------------   159
E11--E14-N6     SFVGWSTDWSPYADQSIKVRISLKSALGDKAANSTLGKKLAD-SHL-ERLR-----AE--   171     E11--E14-N6     -----------------------   219
E11--E18        SFVGWSTDWSPYADQSIKVRISLKSALGDKAYEWNDNEMYLFRSSVAYAMRQYFLKVKNQ   180     E11--E18        ASIDISKGENNPGFQNTDDVQTSF   324
E11--E14-N8     SFVGWSTDWSPYADQSIKVRISLKSALGDKAGPVNGIPALL------------------   161     E11--E14-N8     -----------------------   161
                **********    .  : *::  .*
                       E13              E14
```

Figure S2. Alignment of C-terminal amino acid sequences encoded by predicted exons. The nucleotide sequences were translated using the ExPaSy translate tool, and alignment was done using Clustal Omega.