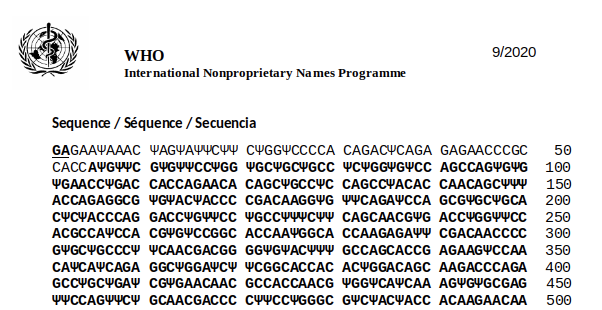
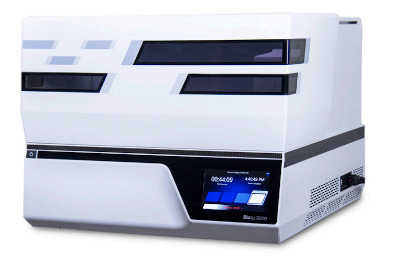
در این نوشتار می‌خواهیم کد منبع واکسن mRNA ویروس SARS-CoV-2، تولید BioNTech/Pfizer را حرف‌به‌حرف بشکافیم. این ادّعا ممکن است کمی عجیب و غریب به‌نظر برسد! بالاخره واکسن مایعی است که به بازوی ما تزریق می‌شود، چطور ممکن است که در مورد کدِ منبعِ آن صحبت کرد؟ این سوال خوبی است، بگذارید با قسمت کوچکی از کدِ منبعِ واکسنِ BioNTech/Pfizer یا همان BNT162b2 یا [Tozinameran](https://en.wikipedia.org/wiki/Tozinameran) یا [Comirnaty](https://twitter.com/PowerDNS_Bert/status/1342109138965422083) شروع کنیم.



۵۰۰ کاراکتر از کدِ منبعِ BNT162b2 mRNA،‌ منبع: [سازمان جهانی بهداشت](https://mednet-communities.net/inn/db/media/docs/11889.doc)

در قلب واکسن BNT162b2 mRNA این کد دیجیتال وجود دارد. این کد ۴۲۸۴ کاراکتر طول دارد و درنتیجه در تعدادی توییت هم می‌گنجد. در مراحلِ اوّلیه‌ی تولیدِ واکسن، یک‌نفر این کد را با یک چاپگر DNA (بله، چنین دستگاهی وجود دارد) از بایت‌های روی دیسک به مولکول‌های واقعی‌ِ DNA تبدیل کرده‌است.



یک چاپگر دی‌ان‌ای، [Codex DNA BioXP 3200](https://codexdna.com/products/bioxp-system/)

خروجی چنین دستگاهی، مقدار بسیار کمی DNA است که بعد از کلّی فرآیندِ شیمیایی و بیولوژیکی در نهایت تبدیل به RNA (جلوتر، بیشتر در این مورد صحبت می‌کنیم) در واکسنِ ما می‌شود. یک دوز ۳۰ میکروگرمی، در واقع ۳۰ میکروگرم RNA دارد. علاوه بر آن، یک بسته‌بندیِ هوشمندانه‌ی لیپیدی (چربی) در واکسن، باعث ورود این mRNA به سلول‌های می‌شود.

RNA یک نسخه‌ی ناپایدار و در واقع نسخه‌ای از DNA در حافظه‌ی موقّت سلول است. DNA معادل فلش‌درایو در بیولوژی است. DNA مانایی زیادی دارد و با مکانیزم‌های مختلفی ساختار پایدار و قابل اعتمادی را حفظ می‌کند. با این‌حال، همان‌طور که کامپیوترها برنامه‌ها را مستقیم از روی فلش‌درایو اجرا نمی‌کنند و قبل از اجرا، آن‌ها را به یک محیط سریع‌تر و منعطف‌تر کپی می‌کنند و از آن‌جا برنامه را اجرا می‌کنند، DNA هم به‌طور مستقیم عامل انفعالات منجر به تغییر نیست.

برای کامپیوترها، آن محیط سریع‌تر و انعطاف‌پذیرتر، حافظه‌ی RAM است، برای بیولوژی RNA. شباهت‌های این دو چشمگیر است. برخلافِ فلش‌درایو، در صورت عدم مراقبت و رسیدگی دایم، محتویاتِ حافظه‌ی RAM به سرعت از بین می‌روند. علّت این‌که واکسن mRNA ساخت Pfizer/BioNTech هم نیاز به محیط بسیار خنک برای نگهداری دارد هم همین است، RNA گُلِ بسیار حساسی است.

هر کاراکتر در RNA حدود ۰.۵۳x۱۰-۲۳ گرم وزن دارد، یعنی در یک دوز تکّیِ ۳۰ میکروگرمی واکسن ۶x۱۰۱۶ کاراکتر وجود دارد. در فرمتِ بایتی، این حدود ۲۵ پتابایت اطلاعات است. البته لازم به یادآوری است که در واقع معادل ۲۰۰ میلیارد تکرار یک ترکیب از ۴۲۸۴ کاراکتر است. محتویاتِ اصلیِ واکسن کمی بیشتر از یک کیلوبایت هستند. خود [SARS-CoV-2](https://www.ncbi.nlm.nih.gov/projects/sviewer/?id=NC_045512&tracks=%5Bkey:sequence_track,name:Sequence,display_name:Sequence,id:STD649220238,annots:Sequence,ShowLabel:false,ColorGaps:false,shown:true,order:1%5D%5Bkey:gene_model_track,name:Genes,display_name:Genes,id:STD3194982005,annots:Unnamed,Options:ShowAllButGenes,CDSProductFeats:true,NtRuler:true,AaRuler:true,HighlightMode:2,ShowLabel:true,shown:true,order:9%5D&v=1:29903&c=null&select=null&slim=0) حدود ۷.۵ کیلوبایت حجم دارد.

**یک پیش‌زمینه‌ی خیلی خلاصه:**

DNA یک کد دیجیتال است. بر خلاف کامپیوترها که از ۰ و ۱ استفاده می‌کنند، ماشینِ زندگی، از A، C، G و U/T برای ساخت [نوکلئوتید](https://fa.wikipedia.org/wiki/%D9%86%D9%88%DA%A9%D9%84%D8%A6%D9%88%D8%AA%DB%8C%D8%AF) و [نوکلئوزید](https://fa.wikipedia.org/wiki/%D9%86%D9%88%DA%A9%D9%84%D8%A6%D9%88%D8%B2%DB%8C%D8%AF) یا [بازها](https://fa.wikipedia.org/wiki/%D8%A8%D8%A7%D8%B2_%D9%86%D9%88%DA%A9%D9%84%D8%A6%D9%88%D8%AA%DB%8C%D8%AF%DB%8C) استفاده می‌کند.

در کامپیوترها، ما ۰ و ۱ را به عنوان علامت حضور یا غیابِ یک بار، جریان الکتریکی، تغییر جهت آهن‌ربایی یا ولتاژ یا مدولاسیونِ یک سیگنال یا تغییر وضعیت انعکاس نور در نظر می‌گیریم. یعنی ۰ و ۱، مفاهیم انتزاعی‌ای نیستند که وجود خارجی نداشته‌باشند، آن‌ها در الکترون‌ها یا گونه‌های دیگری از مادّیات، نمود فیزیکی دارند.

در طبیعت، A، C، G و U/T مولکولهایی هستند که به‌صورتِ زنجیروار در DNA یا RNA ذخیره‌شده‌اند.

در کامپیوترها، ما ۸ بیت را یک بایت می‌گیریم و بایت واحد رایج پردازش داده توسط کامپیوتر است.

در طبیعت، ۳ نوکلئوتید (از این به بعد اشاره به نوکلئوتید، اشاره به یکی از چهار مولکول نام‌برده شده، یعنی C، G، A و U است، که با نام T هم شناخته می‌شود) یک رمز‌ژنتیکی را می‌سازند و این رمزِ ژنتیکی، واحدِ رایجِ پردازش است. یک رمز ژنتیکی معادل ۶ بیت اطلاعات دارد (۲ بیت به‌ازای هر کاراکتر DNA و ثبت ۳ کاراکتر در یک رمز، می‌کند به‌عبارتی ۶ بیت. این یعنی ۲ به توان ۶، یا درواقع ۶۴ رمز ژنتیکی مختلف.)

تا این‌جا که خیلی دیجیتال بود. اگر تردید دارید، [مستند سازمانِ جهانی بهداشت با کدِ دیجیتال](https://mednet-communities.net/inn/db/media/docs/11889.doc) را ببینید تا با چشمانِ خودتان این شباهت را باور کنید.

نوشتارهای بیشتری [این‌جا موجود است، این لینک، زندگی چیست](https://berthub.eu/articles/posts/what-is-life/)، ممکن است به درکِ بهترِ بقیه‌ی این نوشتار کمک کند، یا اگر ویدیو را ترجیح می‌دهید، [دو ساعت ویدیو این‌جا](https://berthub.eu/dna) هست.

**حالا این کد چه می‌کند؟**

ایده‌ی واکسن بر این اساس است که به سیستم‌ایمنی بدنِ ما می‌آموزد که چگونه با یک [پاتوژن](https://fa.wikipedia.org/wiki/%D8%A8%DB%8C%D9%85%D8%A7%D8%B1%DB%8C%E2%80%8C%D8%B2%D8%A7) (بیماری‌زا) بجنگد، بدونِ این‌که ما واقعاً درگیر آن بیماری شویم. به‌صورتِ کلاسیک، این‌کار با تزریق یک ویروسِ ضعیف‌شده یا فلج‌شده (محدود شده و دست‌پا و بسته و یا فقط قسمتی از ظاهر ویروس اصلی) به‌همراه یک مادّه‌ی کمکی یا مساند (adjuvant) به‌عنوانِ عاملی که سیستم ایمنی را طوری تحریک می‌کند که دست به عمل بزند و فعّال شود، به‌بدن انجام می‌شد. این یک شیوه‌ی «آنالوگ» شامل میلیاردها تخمک (و یا حشره) بود. در ضمن، موفقیتِ این روش شانسِ خیلی زیاد و زمانِ زیادی لازم داشت. گاهی اوقات هم از یک ویروس بی‌ربط و نامرتبط استفاده می‌شد.

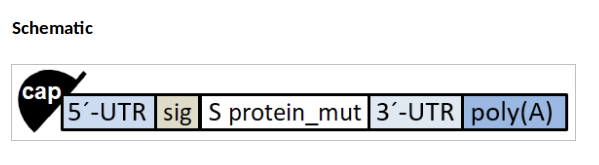
یک واکسن mRNA هم کار مشابهی می‌کند: به سیستم ایمنی بدنِ ما آموزش می‌دهد تا با دیدن نشانه‌های مشخصی وارد عمل شود و عامل آن نشانه‌ها را از بین ببرد. ولی با روشی بسیار دقیق هدف‌گیری شده و قدرتمند.

داستان از این قرار است: محلول تزریقیِ واکسن، شامل مواد ناایستای ژنتیکی است که شاخکِ معروفِ پوشش پروتئینی SARS-CoV-2 را توصیف می‌کند. با شیوه‌های هوشمندانه‌ی شیمیایی، واکسن راه انتقال قسمتی از محتویات ژنتیکی خودش به بعضی از سلول‌های ما را پیدا می‌کند.

سپس این سلول‌های بدنِ ما، خیلی مسوولانه شاخک مذکور را در تعداد زیاد تکثیر می‌کنند، تا جایی که سیستم ایمنی بدنِ ما نگران شده و وارد عمل می‌شود. رودررو با پروتئین‌های شاخک، و مهم‌تر از آن شواهد و علایمی که سلول‌های ما تسخیر شده‌اند، سیستمِ ایمنی بدنِ ما پاسخِ قوی‌ای برای مقابله با زوایای مختلفِ شاخک و همچنین سازوکارِ تکثیر آن می‌کند، و این همان چیزی است که به ما یک واکسن با موفقیت ۹۵ درصد می‌دهد. با این‌که در واکسنِ ما اثری از خودِ ویروس نیست و ما فقط سیستم‌ایمنی را برای مقابله با تکثیر شاخک‌های ویروس تربیت کرده‌ایم، ولی همین برای جلوگیری از امکان فعّالیت ویروس در درونِ بدنِ ما کافی‌است.

**کُدِ منبع!**

ابتدای کُد، جای خیلی خوبی برای شروع است. [مستند سازمانِ جهانیِ بهداشت این تصویر راه‌گشا را دارد](https://youtu.be/jp0opnxQ4rY?t=8):



این درواقع یک‌جور فهرست است. با «cap» که در تصویر با یک کلاه نشان‌داده شده‌است شروع می‌کنیم.

همان‌طوری که شما نمی‌توانید یک‌سری دستورالعمل بنویسید و انتظار داشته‌باشید که کامپیوتر مستقیم آن‌ها را اجرا کند، سیستم‌عامل بیولوژیکی ما هم قبل از شروع به اجرای دستورالعمل‌ها نیازمند دریافتِ یک سرآیند یا Header است که درآن چیزهایی مثلِ [پیوند‌دهنده‌ها](https://fa.wikipedia.org/wiki/%D9%BE%DB%8C%D9%88%D9%86%D8%AF%D8%AF%D9%87%D9%86%D8%AF%D9%87) یاهمان linkerها و [قراردادهای صدازدن](https://fa.wikipedia.org/wiki/%D9%82%D8%B1%D8%A7%D8%B1%D8%AF%D8%A7%D8%AF_%D8%B5%D8%AF%D8%A7%D8%B2%D8%AF%D9%86) یا calling conventionها تعریف‌شده باشند. کد واکسن با دو نوکلئوتید شروع می‌شود.

|  |
| --- |
| GA |

این کدِ شروع، خیلی شبیه کدِ شروع فایل‌های اجرایی [داس و ویندوز است که با MZ](https://en.wikipedia.org/wiki/DOS_MZ_executable) شروع می‌شوند و یا اسکریپت‌های یونیکس که با که [!#](https://en.wikipedia.org/wiki/Shebang_(Unix)) شروع می‌شوند. هم در ماشینِ زندگی و هم در سیستم‌های عامل، این کاراکترها در واقع اجرا نمی‌شوند و دستوری برای انجامِ عملی نیستند، امّا باید در شروع ورودی ما باشند وگرنه هیچ عملی اتّفاق نمی‌افتد.

این کلاه در mRNA [چند کارکرد](https://en.wikipedia.org/wiki/Five-prime_cap#Function) دارد. یکی از آن‌ها این است که منبع کد را هسته‌ی سلولی معرفی می‌کند. در مورد واکسن‌ِ ما، واضح است که منبع این واکسن هسته‌ی سلول نیست ولی در عین‌حال هیچ لزومی هم ندارد که ما این موضوع را به سلول اطلاع‌دهیم! این کلاه، ظاهر کد ما را کاملاً موجّه نشان می‌دهد و جلوی شناسایی آن به‌عنوان یک کدِ خارجی و درنتیجه نابود کردنش توسط سلول را می‌گیرد.

در ضمن، این دو نوکلئوتید GA اولیه از نظرِ شیمیایی کمی متفاوت از بقیه‌ی RNA هستند. به‌همین جهت GA شامل پیام‌های دیگری به‌جز پیامِ محضِ ژنتیکی هم هست.

**ترتیب رهبرِ RNA یا قسمت‌ِ ترجمه‌نشو‌ی پنج ‌پریم:**

این‌جا با کمی زبانِ فنّی طرف هستیم، مولکول‌های RNA فقط می‌توانند از یک جهت خوانده‌شوند. اسمِ جایی که خواندنِ RNA شروع می‌شود، پنج‌ پریم (’۵) است. اسمِ جایی که خواندن متوقّف می‌شود هم سه‌ پریم (’۳) است که کمی گیج‌کننده است!

مبنای زندگی یا پروتئین‌ها هستند و یا چیزهایی که توسط پروتئین‌ها ساخته‌می‌شوند. پروتیئن‌ها در RNA توصیف می‌شوند و توسط ماشین‌آلاتی که در همه‌ی سلول‌های زنده‌ی دنیا وجود دارد، تبدیل به پروتئین می‌شوند. عمل تبدیل RNA به پروتئین را ترجمه می‌گویند.

این‌جا ما قسمتِ ترجمه‌نشو‌ی پنج‌پریم را داریم و درنتیجه این قسمت از کد تبدیل به پروتئین نمی‌شود:

|  |
| --- |
| GAAΨAAACΨAGΨAΨΨCΨΨCΨGGΨCCCCACAGACΨCAGAGAGAACCCGCCACC |

حالا اوّلین سورپرایزمان را می‌بینیم. کاراکترهای معمولی A، C، G و U بودند، U در DNA به‌نامِ T هم شناخته می‌شود. ولی ما در این‌ ترکیب علامت Ψ (پسی، انگلیسی: سای) هم می‌بینیم. داستان چیست؟

این یکی از نوآوری‌های هوشمندانه در مورد واکسن است. بدنِ ما یک آنتی‌ویروس خیلی قوی به اسم «نسخه‌ی اصلی» دارد. به‌همین دلیل هم، سلول‌ها هیچ علاقه‌ای به RNAهای خارجی ندارند و تمام تلاش خودشان را می‌کنند که قبل از این‌که هر کاری از این RNAهای خارجی سر بزند، آن‌ها را نابود کنند.

این برای واکسن ما مساله‌ساز است، واکسن لازم است که بتواند از سیستم ایمنی ما عبور کند. بعد از سال‌ها آزمایش، بشر متوجّه‌شد که اگر U در RNA را با مولکول کمی متفاوتی جایگزین کند، سیستم ایمنی ما دیگر توجّهی به‌آن نمی‌کند و به آن اجازه‌ی عبور می‌دهد. جالب‌تر این‌که در هنگام ترجمه‌شدن، این مولکول متفاوت می‌تواند درست مثلِ U ترجمه‌شود و درنتیجه، پروتئین خروجی، کاملاً مشابه پروتئنی خواهد بود که از یک U آمده!

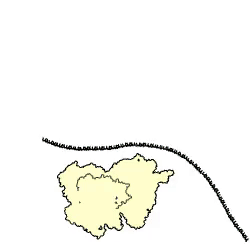
برهمین اساس هم، در واکسن BioNTech/Pfizer، همه‌ی Uها با مولکول «1-methyl-3’-pseudouridylyl» که با Ψ نمایش داده‌شده است، جایگزین شده‌اند. همان‌طور که خواندید، نکته‌ی خیلی هوشمندانه‌‌ی این روش این است که این جایگزینی U با Ψ سیستم ایمنی ما را آرام و رام می‌کند ولی توسط بقیه‌ی قسمت‌های مرتبط سلول، به‌سانِ یک U معمولی پذیرفته می‌شود.

در مبحث امنیت کامپیوترها ما این کلک را می‌شناسیم. گاهی این امکان وجود دارد که با ارسال پیام‌هایی که کمی دست‌کاری شده‌اند، از فایروال‌ها و دیوارهای امنیتی عبور کنیم ولی در مقصد، همچنان اجرا شویم و باعث نفوذ به سیستم مقصد شویم.

در این‌جا ما داریم نتیجه‌ی کاشتِ حاصل از تحقیقات علمی که در گذشته انجام شده‌است را برداشت می‌کنیم. [مکتشفین](https://twitter.com/PennMedicine/status/1341766354232365059) روش Ψ مجبور به کلّی جنگ بودند که بودجه‌ی لازم برای انجام [تحقیقات و قبول شدن کارشان](https://www.statnews.com/2020/11/10/the-story-of-mrna-how-a-once-dismissed-idea-became-a-leading-technology-in-the-covid-vaccine-race/) را بگیرند. ما همه باید قدردانِ آن‌ها باشیم و مطمئن هستم [به‌وقتش جایزه‌ی نوبل](https://twitter.com/PowerDNS_Bert/status/1329861047168225281) به آن‌ها تعلّق خواهد گرفت.

خیلی‌ها سوال کرده‌اند که آیا ویروس‌ها هم می‌توانند از روش Ψ استفاده‌کنند تا از سیستم ایمنی ما عبور کنند؟ به‌طورِ خلاصه، احتمالش فوق‌العاده کم است. خیلی ساده، زندگی ماشین‌آلاتِ تولیدِ نوکلئوتیسِ «*1-methyl-3’-pseudouridylyl*» را ندارد. ویروس‌ها برای بازتولید خود، متّکی به امکانات تولیدی سلول‌ها هستند و چنین امکانی در سلول‌های ما نیست. واکسن mRNA خیلی سریع در بدنِ انسان تجزیه و ترجمه می‌شود و امکان بازتولید RNA با *Ψ* به‌جای U از طریق RNAهای حاملِ *Ψ* وجود ندارد، بل‌که وقتی ماشین‌آلاتِ درون سلول ما مولکول‌های *Ψ در* واکسن را می‌بینند، دست به تکثیر RNA با U به‌جای *Ψ می‌زنند.* «[نه، واقعاً قرار نیست واکسن‌های mRNA باعث تغییر DNA شما شوند](https://www.deplatformdisease.com/blog/no-really-mrna-vaccines-are-not-going-to-affect-your-dna)» متن خوبی برای مطالعه است.

خوب، برگردیم سر منطقه‌ی ترجمه‌نشو‌ی پنج ‌پریم. این ۵۱ کاراکتر چه می‌کنند؟ مثل هر چیزی در طبیعت، هیچ چیزی تنها یک کاربری مشخص و شفاف ندارد و همه چیز چند کارکرد دارد.سلول‌های ما RNA را با ماشینی به‌اسم [ریبوزوم](https://fa.wikipedia.org/wiki/%D8%B1%DB%8C%D8%A8%D9%88%D8%B2%D9%88%D9%85) ترجمه می‌کنند، یعنی ریبوزوم، ماشینِ خواندنِ RNA و تولید پروتئین است. می‌شود گفت ریبوزوم مثل یک پرینتر سه‌بعدی برای پروتئین‌هاست. ریبوزوم یک نخ RNA را می‌گیرد و بر اساس آن یک رشته [اسید آمینه](https://fa.wikipedia.org/wiki/%D8%A7%D8%B3%DB%8C%D8%AF_%D8%A2%D9%85%DB%8C%D9%86%D9%87) تولید می‌کند که بعد از تا شدن، تبدیل به یک مولکول پروتئین می‌شوند.

  
منبع: کاربر ویکی‌پدیا [Bensaccount](https://commons.wikimedia.org/wiki/File:Protein_translation.gif)

در انیمیشن بالا کارکردِ ریبوزوم را می‌بینیم. نوار سیاه در پایین تصویر RNA است. نوار سیاهی که در قسمت سبز می‌بینیم و رو به بالا می‌رود، پروتئینی است که دارد ساخته می‌شود. چیزهایی که داخل و خارج می‌شوند هم اسیدهای آمینه و مبدّل‌هایی هستند که آن‌ها را با اتصال مناسب روی RNA سوار می‌کنند.

این ریبوزوم باید به‌صورت فیزیکی روی نوار RNA بنشیند که بتواند کار کند. بعد از نشستن، ریبوزوم می‌تواند بر اساس ادامه‌ی RNAای که دریافت می‌کند پروتئین بسازد. بر همین اساس می‌توان حدس زد که ریبوزوم نمی‌تواند ترجمه را از همان جایی که فرود آمده شروع کند، وگرنه بسته به زمانی که روی RNA نشسته‌بود و نقطه‌ی شروع، هر دفعه یک پروتئن جدید تولید می‌کرد. این یکی از کاربردهای منطقه‌ی ترجمه‌نشو‌ی پنج پریم است، ایجاد یک باند فرود برای ریبوزوم یا مشخص کردن یک نقطه‌ی شروع.

علاوه بر این، منطقه‌ی ترجمه‌نشو‌ شامل فراداده (metadata) هم هست. ترجمه در چه زمانی باید صورت بگیرد؟ و چقدر؟ برای واکسن، آن‌ها منطقه‌ی ترجمه‌نشده‌ای با حداکثرِ مقدارِ «همین» الآن را یافتند، ترکیبی که از [ژنِ آلفا گلوبین](https://www.tandfonline.com/doi/full/10.1080/15476286.2018.1450054) گرفته‌بودند. این ژن برای تولید قویِ مقدار زیادی پروتئین شناخته‌شده است. در سال‌های گذشته، دانشمندان راه‌هایی برای بهینه‌سازی حتّی بیشترِ همین منطقه‌ی ترجمه‌ی نشده‌ی ژنِ آلفا گلوبین یافته‌بودند (بر اساس مستند سازمان جهانی بهداشت) و بر همین اساس، این منطقه‌ی ترجمه‌شده، دقیقاً منطقه‌ی ترجمه‌شده‌ی ژنِ آلفاگلوبین نیست، بلکه‌ نسخه‌ای است که از آن هم موثرتر است.

**پپتیدِ سیگنال گلیکوپروتئینِ S:**

همان‌طور که گفته‌شد، هدفِ واکسن راضی کردنِ سلول به تولید حداکثری پروتئین شاخکِ ویروسِ SARS-CoV-2 است. تا این‌جا ما فقط با فراداده و قراردادهای صدازدن در کدِ منبعِ واکسن طرف بوده‌ایم. از این‌جا به‌بعد، واردِ محدوده‌ی واقعیِ پروتئین ویروس می‌شویم. البته، هنوز یک لایه فراداده پیش رو داریم. وقتی که ریبوزوم -همان‌طور که در انیمیشن دیدیم- یک پروتئین ساخت، این پروتئین همچنان باید به مقصدی برود. این اطلاعات در «پپتیدِ سیگنالِ گلیکوپروتئینِ S»، یا «ترتیبِ افزوده‌ی رهبر» نوشته‌شده‌ است.

راهِ دیدن این مقصد در یک نوع برچسبِ آدرس است که در شروعِ پروتئین است و در خودِ پروتئین نوشته‌شده‌است. در این نمونه‌ی خاص، پپتیدِ سیگنال می‌گوید که این پروتئین باید از طریق [شبکه‌ی آندوپلاسمی](https://fa.wikipedia.org/wiki/%D8%B4%D8%A8%DA%A9%D9%87_%D8%A2%D9%86%D8%AF%D9%88%D9%BE%D9%84%D8%A7%D8%B3%D9%85%DB%8C) از سلول خارج شود.

پپتیدِ سیگنال خیلی طولانی نیست ولی وقتی که به کد آن نگاه می‌کنیم، تفاوت‌هایی بین RNA واکسن و ویروس می‌بینیم:

(نکته، برای هدف مقایسه، من Ψ های جایگزین‌شده با U در RNAی ویروس را با U عوض کرده‌ام)

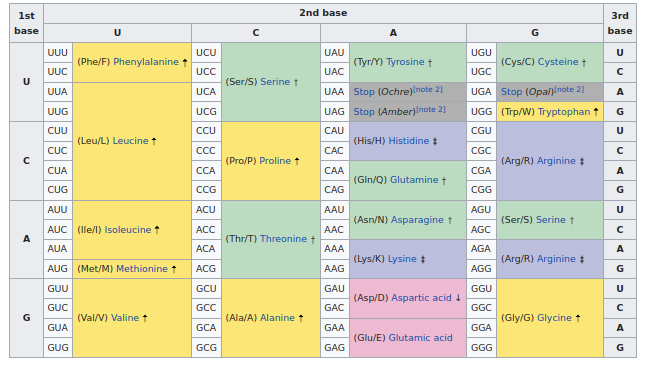
|  |
| --- |
| 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3  Virus: AUG UUU GUU UUU CUU GUU UUA UUG CCA CUA GUC UCU AGU CAG UGU GUU  Vaccine: AUG UUC GUG UUC CUG GUG CUG CUG CCU CUG GUG UCC AGC CAG UGU GUU  ! ! ! ! ! ! ! ! ! ! ! ! ! |

خوب؛ این‌جا چه خبر است؟ این‌که من RNA را در گروه‌های سه‌حرفی لیست‌کرده‌ام تصادفی نبوده‌است. سه کاراکتر در RNA یک رمز ژنتیک می‌سازند و هر رمز ژنتیک حاوی دستور یک اسیدآمینه‌ی خاص است. پپتیدِ سیگنال در واکسن حاوی دقیقاً همان آمینو اسیده‌هایی است که در ویروس هم هستند.

**پس چرا RNA متفاوت است؟**

از آن‌جایی که ۴ کاراکتر RNA داریم و سه‌تای آن‌ها در رمز ژنتیک هستند، معادل۴۳، یعنی ۶۴ رمز ژنتیک مختلف خواهیم داشت. با این‌حال در خروجیِ ترجمه‌ی ریبوزوم همیشه فقط ۲۰ اسید آمینه‌ی مختلف داریم. این یعنی مواردی هست که یک اسیدِ آمینه چند رمز ژنتیک متفاوت دارد که ترجمه‌ی هریک از آن رمزها منتج به تولید همان یک اسید می‌شود.

ماشینِ زندگی، تقریباً در همه‌ی موارد از جدول زیر برای تبدیل رمزِهای ژنتیک به اسیدهای آمینه استفاده می‌کند [کاوه: البته در اختاپوس و چند فامیل ژنتیکی آن، ماجرا متفاوت است و این موجودات می‌توانند روی RNA هم تغییر ایجاد کنند، به‌زبان کامپیوتری، یعنی اختاپوس می‌تواند در زمان Runtime بر اساس ورودی‌ها، کدِ اجرایی را تغییر دهد! امیدوارم روزی بتوانم چیزی در این مورد بنویسم یا ترجمه کنم، برای یک مقاله‌ی جالب: <https://www.sciencealert.com/octopus-and-squid-evolution-is-weirder-than-we-could-have-ever-imagined>]

جدول رمزِ ژنتیک RNA از [ویکی‌پدیا](https://fa.wikipedia.org/wiki/%D8%AC%D8%AF%D9%88%D9%84_%DA%A9%D8%AF%D9%88%D9%86_%D8%AF%DB%8C%E2%80%8C%D8%A7%D9%86%E2%80%8C%D8%A7%DB%8C)

در این جدول می‌بینیم که تفاوت‌هایی که در واکسن با ویروس وجود دارد (UUU -> UUC) همه هم معنی هستند، یعنی با وجود تفاوتِ RNA در واکسن، اسیدآمینه‌ و پروتئین مشابه‌ای به‌عنوانِ خروجیِ عملِ ریبوزوم تولید می‌شود.

اگر دقیق‌تر نگاه کنیم، می‌بینیم که بیشتر تغییرات در موقعیت رمزِ ژنتیکی سوّم هستند، که در کد با کاراکتر 3 در بالای آن‌ها را علامت‌گذاری کرده‌ام و اگر جدول رمزِ ژنتیک را نگاه کنیم، می‌بینینم که این رمزِ ژنتیکی سوّم در خیلی از موارد روی آمینو اسید تولید شده تاثیری ندارد.

پس این تغییرات هم‌ارز هستند، امّا اصلاً چرا وجود دارند؟ اگر بیشتر دقّت کنیم، می‌بینیم که همه‌ی این‌تغییرات به‌جز یکی از آن‌ها در جهت داشتنِ C و Gهای بیشتر بوده‌اند.

چرا باید چنین کاری بکنیم؟ همان‌طور که گفتیم، سیستم ایمنی بدنِ ما خیلی در موردَِ RNAی که از خارج از سلول وارد شده‌باشد سخت‌گیر است. برای جلوگیری از شناخته‌شدن ما پیشتر U را با Ψ جایگزین کرده‌بودیم.

علاوه‌بر این، ما می‌دانیم که RNA [با مقدار بیشتری G و C](https://www.nature.com/articles/nrd.2017.243) با بازدهی بالاتری به پروتئین ترجمه می‌شود و در RNA واکسن، با یافتن مشابه‌هایی با G و C بیشتر برای رسیدن به همان آمینواسیدهایی که ویروس هم ترجمه می‌کند، [بازدهی تولید ساختار پروتئنی شاخک را بالا برده‌ایم.](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1463026/)

**پروتئین شاخکِ واقعی:**

۳۷۷۷ کاراکتر بعدی از RNA واکسن به همان صورت برای رمزهای ژنتیکی بهینه‌سازی‌ شده‌اند تا از رمزهای ژنتیکی با حداکثر C و G ولی معادل با خروجیِ پروتئین ویروس را استفاده کنند. برای حفظ فضا من کلِّ کد را این‌جا لیست نمی‌کنم ولی یک قسمتِ خاص را بزرگ‌نمایی می‌کنم. این قسمتی است که باعث می‌شود که کد کار کند، قسمتی که باعث می‌شود ما به زندگی معمولی بازگردیم:

|  |
| --- |
| \* \*  L D K V E A E V Q I D R L I T G  Virus: CUU GAC AAA GUU GAG GCU GAA GUG CAA AUU GAU AGG UUG AUC ACA GGC  Vaccine: CUG GAC CCU CCU GAG GCC GAG GUG CAG AUC GAC AGA CUG AUC ACA GGC  L D P P E A E V Q I D R L I T G  ! !!! !! ! ! ! ! ! ! ! |

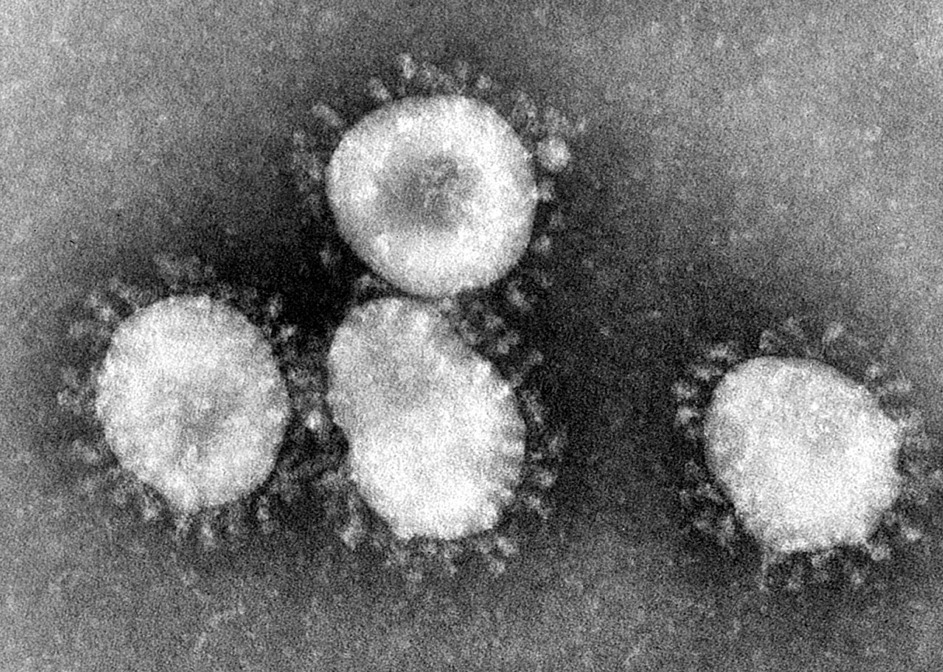
این‌جا هم تغییرات RNA با معادل‌های هم‌ارز را می‌بینیم. برای نمونه، در رمزِ ژنتیک اوّل، می‌بینیم که CUU به CUG تبدیل شده‌است، این یک G جدید به واکسن اضافه می‌کند که ما می‌دانیم که کارآیی تولید پروتئین را بالا می‌برد. هر دو CUU و CUG منتج به تولید اسیدآمینه‌ی L یا [لوسین](https://fa.wikipedia.org/wiki/%D9%84%D9%88%D8%B3%DB%8C%D9%86_(%D8%A7%D8%B3%DB%8C%D8%AF_%D8%A2%D9%85%DB%8C%D9%86%D9%87)) می‌شود، پس چیزی در پروتئین تولید‌شده توسط سلول که قرار است توسط سیستم ایمنی به‌عنوانِ یک تهدید شناسایی شود، تغییر نمی‌کند.

وقتی که کلِّ پروتئین شاخک در واکسن را با شاخک خودِ ویروس مقایسه می‌کنیم، همه جا چنین تغییراتی می‌بینیم، به‌جز دو جا، و این چیزی است که در این قسمت از کد می‌بینیم.

رمزهای ژنتیکی سوّم و چهارم در بالا، نمایانگر تغییرات هستند. اسیدهای آمینه‌ی K و V هر دو با اسیدِ P یا [پرولین](https://fa.wikipedia.org/wiki/%D9%BE%D8%B1%D9%88%D9%84%DB%8C%D9%86) تعویض شده‌اند. برای K این نیاز به سه تغییر داشته (مشخص شده‌با !!!) و برای V نیاز به دو تغییر (!!).

مشخص شده که این دو تغییر، باعث افزایش بسیارزیادِ کارآمدی واکسن می‌شود.

حالا سعی کنیم ببینیم در این‌جا چه اتّفاقی می‌افتد؟ اگر به ذرات ویروس SARS-CoV-2 نگاه کنید، می‌توانید پروتئین شاخک را به شکل یک‌سری، درست حدس زدید، شاخک!، ببینید:



ذرات ویروس SARS-CoV-2، منبع [ویکی‌پدیا](https://en.wikipedia.org/wiki/Severe_acute_respiratory_syndrome_coronavirus)

این شاخک‌ها روی بدنه‌ی ویروس یا پروتئین [کپسید](https://fa.wikipedia.org/wiki/%DA%A9%D9%BE%D8%B3%DB%8C%D8%AF) هسته‌ای سوار شده‌اند. نکته‌این‌جاست که واکسن ما فقط کد ساختِ شاخک‌ها را دارد و ما آن‌ها را سوار هیچ بدنه‌ی ویروسی نمی‌کنیم.

مشخص‌شده که پروتئین‌های شاخکِ تنها، به‌شکل ساختارهای دیگری «تا» می‌شوند، یعنی اگر ما فقط دستور چاپ شاخک را به ریبوزوم بدهیم، بعد از تولید نوارِ پروتئین توسط ریبوزوم، این نوار به فرمی تا می‌شود که آن را تبدیل به مولکولی متفاوت از شاخک می‌کند. اگر این شاخک‌های تا شده به‌عنوان واکسن به بدن ما تزریق شوند، باعث می‌شوند تا سیستم ایمنی ما راهی برای ایمن کردن ما در مقابل آن‌ها پیدا کند، تنها نکته‌اش این است که این شاخک‌های تا شده، مشابه شاخک‌های ویروس نیستند و درنتیجه بدنِ ما در مقابل ویروس مقاوم نخواهد شد و فقط در مقابل شاخک‌های تا شده ایمنی خواهیم یافت.

پس چه باید کرد؟ [در سال ۲۰۱۷ مشخص شد که چگونه جاگزینی دو اسیدآمینه با پرولین، در جای مناسب در ساختار ویروس](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5584442/) باعث می‌شود شاخک‌های SARS-CoV-1 و MERS S حتی پس از تا شدن، فرمی مشابه با فرمِ شاخکِ سوار بر بدنه‌ی ویروس را بگیرند، حتی بدونِ این‌که ویروسی وجود داشته‌باشد که بدنه‌ای داشته‌باشد. علّت این ماجرا هم سختیِ پرولین به عنوانِ آمینو اسید است. پرولین مثلِ یک نبشی، باعث ثباتِ اسیدآمینه در وضعیتی می‌شود که ما می‌خواهیم به سیستم‌ایمنی نمایش بدهیم.

[مکتشفین](https://twitter.com/goodwish916) این ماجرا باید [خیلی خوشحال](https://twitter.com/KizzyPhD) باشند و باید بی‌نهایت از آن‌ها [قدردان](https://twitter.com/McLellan_Lab/status/1291077489566142464) بود.

**پایانِ پروتئین، قدم‌های بعدی:**

اگر بقیه‌ی کدِ منبع را نگاه کنیم، تغییرات کوچکی در پایانِ کدِ شاخک می‌بینیم:

|  |
| --- |
| V L K G V K L H Y T s  Virus: GUG CUC AAA GGA GUC AAA UUA CAU UAC ACA UAA  Vaccine: GUG CUG AAG GGC GUG AAA CUG CAC UAC ACA UGA UGA  V L K G V K L H Y T s s  ! ! ! ! ! ! ! ! |

در انتهای یک پروتئین، ما رمزِ ژنتیکی «ایست» را می‌بینیم که با یک s کوچک مشخص شده‌است. این روش مودبانه‌ی اعلام توقف ترجمه‌ی پروتئین به ریبوزوم است. ویروس اصلی از ترتیبِ UAA استفاده می‌کند واکسن ولی از دو UGA استفاده می‌کند. به‌احتمال زیاد برای قطعیتِ بیشتر از اتمامِ پروسه‌ی تولید.

**قسمتِ ترجمه‌نشو‌ی سه‌پریم:**

همان‌طور که ریبوزوم برای پیدا کردنِ نقطه‌ی شروعِ ترجمه به قسمتِ ترجمه‌نشو‌ی پنج‌پریم احتیاج داشت، در محلِ انتهای ترجمه هم یک قسمت ترجمه‌نشده با ساختاری مشابه نیاز داریم که آن را سه‌پریم می‌نامند.

می‌توان خیلی در مورد قسمتِ سه‌پریم نوشت ولی من به [نقلِ آن‌چه که در ویکی‌پدیا آمده](https://en.wikipedia.org/wiki/Three_prime_untranslated_region)، اکتفا می‌کنم: «قسمتِ سه‌پریم، تاثیرگذاری حیاتی در کارآییِ محلیّت، ثبات، صدور و ترجمه‌ی mRNA دارد... علی‌رغمِ درکِ فعلی ما از قسمتِ سه‌پریم، به‌صورتِ نسبی آن‌ها هنوز برای ما یک معمّا هستند.»

چیزی که ما می‌دانیم این است که بعضی از قسمت‌های سه‌پریم خیلی موفق‌تر در ارایه‌ی توصیفی از پروتئین هستند. بر اساس مستندِ سازمانِ جهانیِ بهداشت، قسمتِ ترجمه‌نشوی سه‌پریم در واکسنِ BioNTech/Pfizer از amino-terminal enhancerهای میتوکندریایی ریبوزومی 12S تولید شده [کاوه: در هنگام نگارشِ این متن من هیچ اطلاعی از این که منابع سه‌پریم مورد اشاره چیست، ندارم] تا به ثابت و توصیف حداکثری پروتئین برسیم.

**و AAAAAAAAAAAAAAAAAAAAAA در پایان همه‌ی این‌ها:**

انتهای mRNA [چند آدننینی](https://fa.wikipedia.org/wiki/%DA%86%D9%86%D8%AF%D8%A2%D8%AF%D9%86%DB%8C%D9%86%DB%8C%E2%80%8C%D8%B4%D8%AF%D9%86) شده‌است. این مدل دیگری برای گفتنِ این است که این mRNA با کلّی AAAAAAAAAAAAAAAAAAA پایان می‌یابد. ظاهراً حتّی mRNA هم از سال ۲۰۲۰ سیر شده‌است!

mRNAها می‌توانند بارها توسط ریبوزوم و برای تولید پروتئین، استفاده‌ی مجدد شوند اما در هر استفاده تعدادی از Aهای انتها را از دست می‌دهند، وقتی که A ها تمام شوند، mRNA دیگر فایده‌ای برای ریبوزوم ندارد و به‌دور ریخته‌ می‌شود. در موردِ ما، این Aها محافظی هستند در مقابلِ بی‌مصرف‌شدنِ زودهنگامِ mRNA.

تحقیقاتی در مورد تعدادِ بهینه‌ی A در انتهای واکسن‌های mRNA انجام شده‌است. چیزی که من در مقالاتِ باز دیده‌ام این است که ظاهراً عدد ۱۲۰ تعداد مناسبی برای Aها در انتهای mRNA می‌باشد.

**واکسن BNT162b2 با این ترتیب پایان می‌یابد:**

|  |
| --- |
| \*\*\*\*\*\* \*\*\*\* UAGCAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAGCAUAU GACUAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAA |

این ۳۰ تا A است، بعد ۱۰ نوکلئوتید مرتبطِ ساز و به‌دنبال آن ۷۰ A دیگر. حدس می‌زنم، چیزی که این‌جا می‌بینیم، نتیجه‌ی بهینه‌سازی بیشتر و استفاده از روشی اختصاصی برای توصیف پروتئین با کارآیی حداکثری باشد.

**خلاصه:**

با توجّه‌به این‌ها، ما الآن می‌دانیم که دقیقاً محتوای mRNA در واکسن BNT162b2 چیست و برای بیشتر آن‌ها هم توضیح مناسبی داریم:

- استفاده از کلاه، تا RNA ما شبیه به یک mRNA معمولی باشد

- استفاده از یک قسمتِ ترجمه‌ی نشوی پنج‌پرایمِ موفق و بهینه‌سازی شده

- رمز‌ژنتیکی بهینه‌سازی شده برای ارسال پروتئین شاخک به‌جای درست که ۱۰۰٪ از روی ویروس اصلی کپی شده.

- رمز ژنتیکیِ بهینه‌سازی شده‌ی شاخک اصلی، با جایگزینی دو اسیدآمینه با پرولین، به‌عنوانِ نبشی برای حفظِ فرمِ شاخک بدونِ وجودِ بدنه‌ی ویروس

- استفاده از یک قسمتِ ترجمه‌نشوی سه‌پریمِ موفق و بهینه‌سازی شده

- یک دُمِ چند A کمی عجیب با یک ربط دهنده‌ی توضیح داده‌نشده در میان آن  
- بهینه‌سازی رمز‌های ژنتیکی تعداد زیادی C و G به mRNA می‌افزاید. در همین حال، استفاده از Ψ یا همان مولکول (1-methyl-3’-pseudouridylyl) به‌‌جای U به گول زدنِ سیستم‌ِ ایمنی بدنِ ما و وارد شدنِ ویروسِ تولیدی به سلول‌ها برای تولید پروتئین شاخک ویروس و در نتیجه آموزش سیستمِ ایمنی بدنِ ما برای مقابله با ویروس.

**منابع برای مطالعه‌ی بیشتر:**

من در سال ۲۰۱۷ یک ارایه‌ی دو ساعته در مورد DNA داشتم که می‌توانید از [این‌جا](https://berthub.eu/dna) ببینید. مشابه با همین نوشتار، مخاطبِ هدف من آدم‌های آشنا با سازوکارِ کامپیوتر بوده‌است.

در عینِ حال، از سال ۲۰۰۱ من صفحه‌ای به‌اسم DNA [برای برنامه‌نویسان](https://berthub.eu/amazing-dna) دارم.

شاید از این [معرفی سیستم ایمنیِ فوق‌العاده‌ی بدنِ ما](https://berthub.eu/articles/posts/immune-system/) هم لذّت ببرید.

درنهایت، [این لیست از نوشته‌های وبلاگیِ من](https://berthub.eu/articles)، مقالات خوبی در مورد DNA، SARS-CoV-2 و مسایل مرتبط با COVID دارد.