

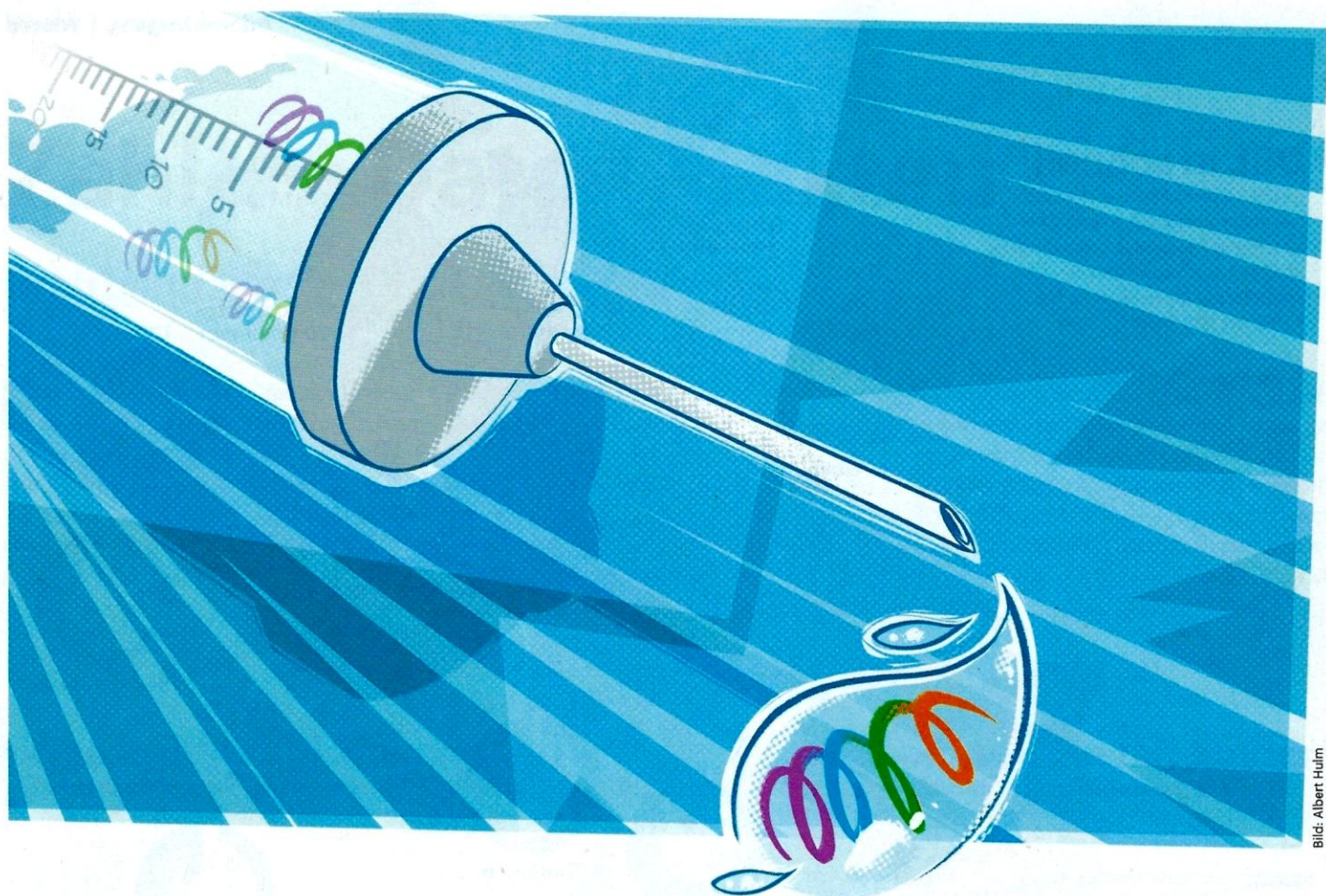


Bild: Albert Hulm

Zielwasser

Mit programmierbarem Impfstoff gegen SARS-CoV-2

Ende 2020 hat die Weltgesundheitsorganisation den Code des Corona-Impfstoffs von BioNTech und Pfizer publiziert. Er steckt voller Raffinessen und belegt, warum Impfstoffe, die nur das Viren-Erbgut enthalten, als das „nächste große Ding“ der Molekularbiologie gelten.

Von Dušan Živadinović

Der von BioNTech und Pfizer entwickelte Impfstoff BNT162b2 ist die erste Präzisionswaffe zur offensiven Bekämpfung der Corona-Pandemie. Es ist aber auch der erste Vertreter einer neuen Impfstoffklasse, die man buchstäblich programmieren kann. Mit genau 4284 Zeichen ist der Code zudem noch gut überschaubar, sodass er sich prima als Anschauungsmaterial eignet; er würde sogar locker auf eine c't-Seite passen. Sie finden ihn wie alle übrigen Literaturstellen zu diesem Beitrag über ct.de/ywa8. Leider gibt es unseres Wissens keine korrekte populärwissenschaftliche Beschreibung des BNT162b2-Impfstoffs (auch Tozinameran oder Comirnaty bezeichnet). Das holen wir hier nach.

Gute und böse Attrappen

Zunächst zwei Absätze zu Impfkonzepthen: Die Impfung gilt als eine der elegantesten Methoden der Bekämpfung von übertragbaren Krankheiten. Anstatt Krankheitserreger mit chemischen Keulen zu bekämpfen, wird der lernfähige Anteil des menschlichen Abwehrsystems mittels Attrappen trainiert. In vielen Impfstoffen dienen daher abgetötete oder abgeschwächte Erreger als Attrappen.

Spritzt man sie ins Muskelgewebe, reagiert das Abwehrsystem mit diversen Immunzellen darauf. Unter anderem bildet es Antikörper, die komplementär zu markanten Teilen der Attrappe sind, sodass sie daran anhaften. Wenn später der tatsächliche Erreger einwandert, haften sie ebenfalls daran an und markieren ihn dadurch, sodass er vom Immunsystem erkannt und „aufgefressen“ wird.

Die herkömmlichen Methoden zur Impfstoffherstellung sind aber aufwendig. Auch können abgeschwächte Erreger rückmutieren und dann wieder Krankheiten auslösen.

Den genetischen RNA-Impfstoffen wie BNT162b2 und mRNA-1273 vom US-Unternehmen Moderna liegt zwar dasselbe Trainingsprinzip mittels Attrappen zugrun-

de. Doch dabei werden weder abgetötete noch abgeschwächte Erreger gespritzt, sondern nur ausgesuchte Gene (Programmcodes) für den Bau von Oberflächenelementen. Die damit behandelten Zellen lesen die Gene und bauen die Attrappen selbst. Daher sind anders als bei Impfstoffen mit abgeschwächten Erregern keine Rückmutationen möglich. Und weil solche Impfstoffe ausschließlich Steuermodule und Gene für Oberflächenproteine enthalten, ist anders als etwa bei vollständigen HIV-Viren kein Erbguteinbau in menschliche Chromosomen möglich; den Impfstoffen fehlt dafür jegliche Ausrüstung. Wobei sich das SARS-CoV-2 ohnehin nicht in Chromosomen einnistet. Besonders elegant erscheint, dass die Programme am Ende ihre eigene Zersetzung einleiten.

Alle für die Biosynthese erforderlichen Anweisungen muss ein genbasierter Impfstoff mitbringen; man konstruiert ihn im Prinzip am Schreibtisch und synthetisiert ihn im Labor, indem man je nach Problem oder Anforderung Module aus dem Baukasten der Mikrobiologie zusammensetzt.

In der Pandemie kommt den Herstellern der genetischen Impfstoffe das Human Genome Writing Program zugute, das 2016 gestartet wurde und zum Ziel hat, alle menschlichen Gene innerhalb von 10 Jahren zu synthetisieren. Im Zuge dessen haben Biotech-Firmen wie SynBio Technologies schnelle „Printer“ entwickelt, für die die Synthese einzelner Virusgene ein Leichtes ist. So sanken in den letzten zehn Jahren die Kosten für die Synthese eines Gen-Bausteins (DNA-Basenpaar) von rund 1 US-Dollar auf weniger als 10 Cent.

Verpackte Botschaften

Weil genetische Impfstoffe thermisch labil sind (z. B. anfällig für Oxidation oder Spaltung durch interne Reaktionen), steckt man sie zum Transportschutz zumeist in Kügelchen, die zugleich wasser- und fettlöslich sind (Nano-Partikel aus mehreren Lipiden) und kühlt sie weit unter den Gefrierpunkt. Die Kügelchen sind in physiologischer wässriger Lösung aufgenommen, die sich nach dem Spritzen im Gewebe verteilt.

Außen an Zellmembranen angekommen, holt sie die Zelle über gängige Transportmechanismen herein (Endocytose). Die Zelle ist in Hohlräume (Kavernen) mit spezifischen Funktionen etwa zur Informationsspeicherung oder Energiegewinnung unterteilt. Die Kavernen sind über

Transportsysteme verbunden. BNT162b2 gelangt darüber weit ins Zellinnere bis zu den Nano-Maschinen (Ribosomen), welche anhand seiner Gensequenz ein Viren-Protein synthetisieren (Translation).

BNT162b2 besteht hauptsächlich aus dem Gen für das Protein, das wie Spieße auf den kugelförmigen SARS-CoV-2-Viren aufgepflanzt ist (Spikes). Das Gen ist Teil eines strangförmigen Polymers aus vier Ribonukleinsäuren (hier Messenger-RNA, mRNA): Adenin, Cytosin, Guanin und Uracil, abgekürzt A, C, G und U. Je drei RNA-Bausteine (Basen) sind zu einem „Byte“ zusammengefasst (Basentriplett, Codon, z. B. AGA oder ACG). Wenn ein Ribosom eine solche RNA einfädelt, knüpft es entsprechend den Codons eine Aminosäurekette beziehungsweise ein Protein.

Jede Aminosäure steckt an einem individuellen Träger (Transport-RNA, tRNA), der am entgegengesetzten Ende komplementär zum Codon zusammengesetzt ist (Anticodon). Deshalb rastet zum Beispiel eine Arginin-beladene tRNA passgenau am Arginin-Codon (AGA) einer mRNA ein. Hat das Ribosom die Aminosäure abgetrennt und an das wachsende Protein angehängt, springt die leere tRNA ab und das Ribosomen-Getriebe dreht die mRNA um ein Codon weiter. Der Vorgang wiederholt sich mit der nächsten Aminosäure, bis das Protein fertig ist.

Optimierter Quellcode

Das gesamte Virengenom umfasst je nach Variante 27.400 bis 30.100 Kilobasen. Das Spike-Gen allein besteht aus nur rund 1000 Basen. Davor und dahinter haben

die Konstrukteure den Großteil der Steuerelemente eingefügt, die BNT162b2 zu einem sehr guten Impfstoff mit 95 Prozent Wirksamkeit machen. BNT162b2 ist daher 4282 Basen lang.

Die Steuerelemente kann man auch als Verwaltungs-Overhead ansehen, der bestimmt, wo und wie das Gen abgelesen wird. Ein Problem der Impfstoffkonstrukteure war nämlich, dass Ribosomen tief in der Zelle stecken und die mRNA erst mal dorthin muss, damit sie Ribosomen zum Spike-Protein translatieren. Die übliche Verarbeitungsrichtung der Zelle verläuft aber entgegengesetzt: Im Zellkern wird von der DNA eine mRNA-Abschrift erzeugt und zur Proteinherstellung in äußere Kavernen gebracht. Die künstliche mRNA des BNT162b2 soll aber von außen nach innen.

Außerdem soll die Zelle möglichst umgehend mit der Produktion starten, möglichst schnell viele Spike-Proteine erzeugen, diese an die Zelloberfläche befördern und anschließend jedes Impfstoffmolekül in Einzelteile zerlegen. Damit all das klappt, nutzen die BioNTech-Ingenieure einige Tricks und Steuerkommandos, auf die wir nun im Detail eingehen.

Sensoren austricksen

Wenn Sie das WHO-Dokument mit der Sequenz des BNT162b2-Wirkstoffs öffnen, dürfte Ihnen schnell auffallen, dass die Sequenz anders als übliche mRNAs nicht den Baustein Uracil enthält (U). Stattdessen findet man zahlreiche Ψ (griechisch Psi). Dabei handelt es sich nicht um einen Darstellungsfehler Ihrer Office-Software und die Auflösung des Rät-

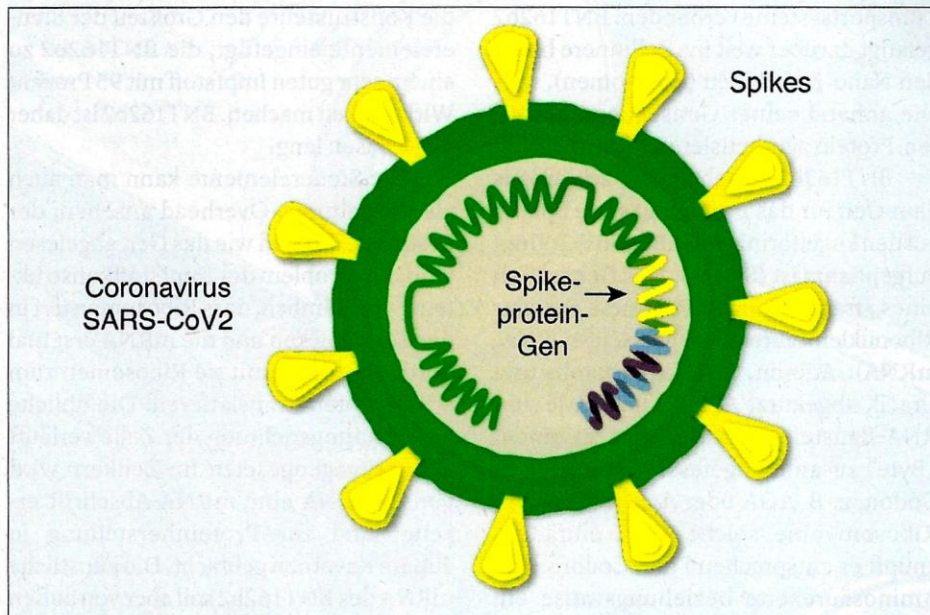
Wirkstoff gegen Zellpiraten

Mit dem genetischen Impfstoff BNT162b2 lässt sich das Corona-Virus auf faszinierende Weise bekämpfen. Cool daran ist zum Beispiel, dass er die Schleusen zum Zellinneren passiert, ohne das angeborene Abwehrsystem der Zelle zu alarmieren – so soll es sein, sonst hätte der Stoff keine Impfwirkung. Dieser und diverse weitere Tricks der Bio-Ingenieure sind Paradebeispiele angewandter Forschung. Um sie zu verstehen, braucht man aber zum Glück keine Spezialkenntnisse der Molekularbiologie.

Stattdessen genügt eine grobe Vorstellung davon, dass viele der kleinsten

Elemente des Lebens fadenartige Polymere sind (Proteine, Enzyme), die sich nach der Synthese individuell zu biologisch aktiven Formen zusammenfalten.

Manche Abschnitte solcher Polymere können zu Spiralen angeordnet sein (Alpha-Helix), andere sehen wie Faltblätter aus, wieder andere erscheinen wirr verknäuel (Random Coil). Ihre genaue Funktion ergibt sich aus ihrer 3D-Struktur. Beispielsweise faltet sich das Sauerstofftransportmolekül Hämoglobin so, dass es Taschen zur Aufnahme von Sauerstoff bildet.



Dockt das Coronavirus mit einem Spike an eine Zelle an, nimmt das Verhängnis seinen Lauf, die Zelle wird zur Virus-Fabrik ...

sels steht implizit am Ende der Sequenz: Ψ steht für 1-Methyl-3'-Pseudouridylyl, kurz Pseudo-Uridin. Pseudo-Uridin kommt zwar in der Natur vor, aber nicht in mRNAs.

In BNT162b2 setzt man es ein, um Teile des angeborenen Abwehrsystems zu umgehen. Dazu gehören intrazelluläre Sensoren, die einzelsträngige RNA buchstäblich abtasten und anhand ihrer Struktur identifizieren (Toll-like-Rezeptoren, TLR7 und TLR8). Dann geben sie Alarm, sodass die Zelle die Proteinbiosynthese abschaltet; die Sensoren können sogar den programmierten Zelltod auslösen. Beides ist schlecht für einen Impfstoff, denn ohne Proteinsynthese keine Spikes, ohne Spikes keine Immunantwort.

Aber mRNA-Stränge, die Ψ anstatt U enthalten (modifizierte RNA, modRNA), weisen eine leicht veränderte Struktur auf, sodass die Sensoren nicht anschlagen. Die Proteinsynthese der Ribosomen beeinträchtigt diese Strukturänderung nicht.

Ausführungs-Bit

Eine ähnliche Aufgabe wie die Ψ-Bausteine haben die beiden ersten GA-Bausteine ganz am Anfang von BNT162b2 (Cap-Sequenz): Sie sind chemisch leicht modifiziert, sodass der Strang für einen anderen Teil des angeborenen Abwehrsystems, die Nukleasen, unverdaulich ist; Nukleasen bauen von außen eintretende RNA oder DNA normalerweise ab. Außerdem täuscht man mit der GA-Markierung der

Zelle vor, dass es sich um eine mRNA aus dem Zellkern handelt. Und schließlich kennzeichnet man mit GA den Strang für die Ribosomen als „ausführbar“, etwa so wie bei Unix-Systemen mit dem Executable-Bit auf Dateisebene.

Weiter stromabwärts der Sequenz folgt ein längeres Element, mittels dessen mRNA und Ribosom aufeinander einrasten (Untranslated Region, UTR). Diese Sequenz reicht von Position 1 bis 54 (GAA...ACC).

Von der genauen Basenabfolge des UTR-Stücks hängt aber auch ab, wie schnell die Ribosomen synthetisieren. Inzwischen kennt man etliche Gene, deren Proteine schnell und in großer Zahl produziert werden. BioNTech setzt hier eine optimierte Sequenz ein, die vom Gen des häufig produzierten Alpha-Globin abgeleitet ist. Das ist mit der Bemerkung „optimized Kozak sequence“ gemeint, die im WHO-Dokument steht.

Anhängerkupplung und Seilbahn

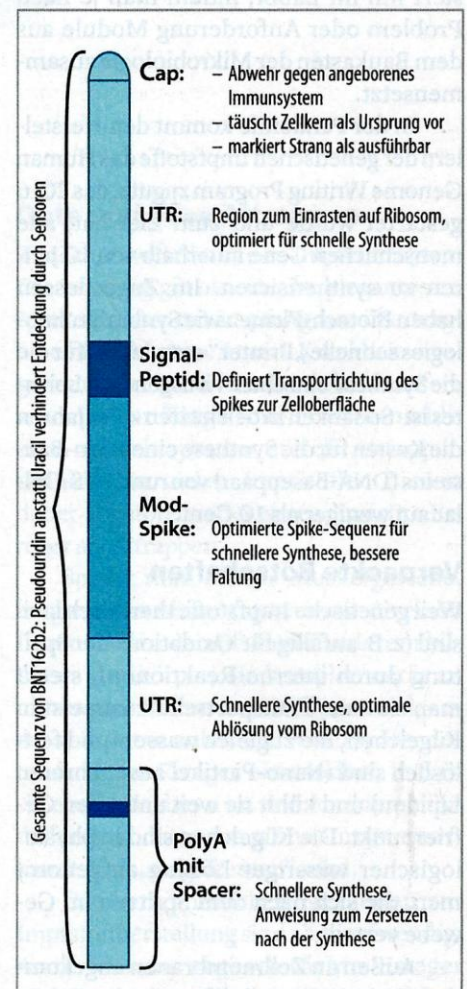
Auf das UTR-Element folgen die ersten Basen-Triplets, die für ein Protein kodieren (Position 55 bis 102). Dabei handelt es sich um die Sequenz für das S-Glycoprotein. Auf den Positionen 103 bis 3879 folgt das Spike-Gen. Beide sind in sehr ähnlicher Form auch in SARS-CoV-2 enthalten.

Das S-Glycoprotein kann man als spezifische Kupplung auffassen, denn alle Proteine, an denen ein solches Signalpeptid hängt, docken damit an spezi-

fische Transportproteine an, die sie in Lipid-Kapseln verpacken (Vesikel), damit sie die zelleigene Seilbahn an die Zelloberfläche befördert (Exocytose). Proteine mit anderem Bestimmungsort tragen andere Signalpeptide und gelangen so auf andere Transportwege.

Am S-Glycoprotein ist auch das erste Basen-Triplett, AYG, von besonderer Bedeutung. Übersetzt in den normalen mRNA-Code lautet es AUG und stellt das Start-Codon dar. Es stellt sicher, dass das Ribosom auf dem richtigen von drei möglichen Leserastern einhakt (Reading Frame). Denn da je 3 Basen gemeinsam gelesen und in ein Protein übersetzt werden, ist es wichtig hier das Leseraster mit der richtigen Base zu beginnen.

Würde das Ribosom um ein Frame versetzt mit dem Triplett UGU starten, müsste es Cystein anstatt Methionin ein-



Genetische Impfstoffe lassen sich am Reißbrett konstruieren. Je nach Anforderung oder Problem flanscht man Steuermodule aus dem Baukasten der Mikrobiologie an Gene eines Erregers an.

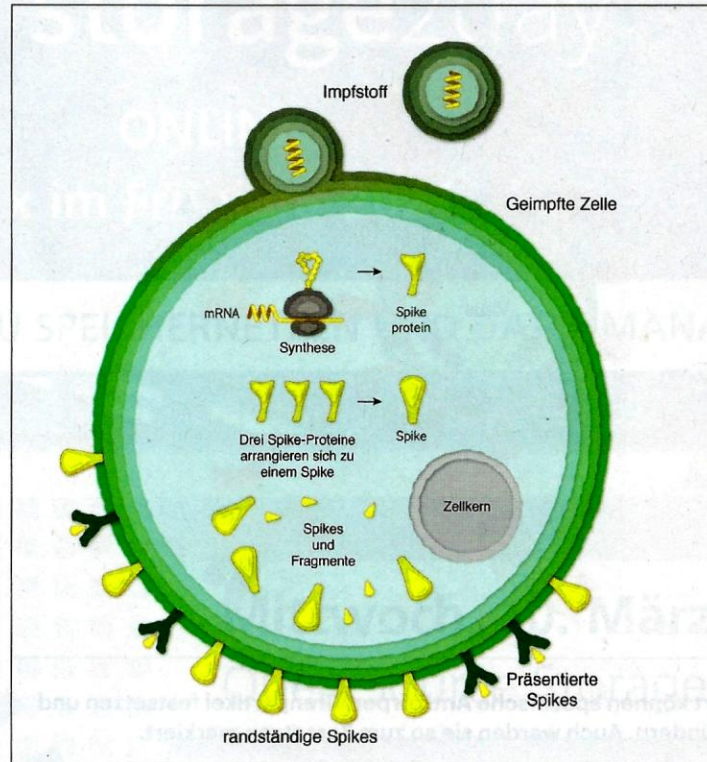
bauen und so weiter. Das so gebildete Protein hätte eine gänzlich andere Sequenz und würde sich nicht zum Spike falten. Das Abwehrsystem würde darauf reagieren, wäre aber nicht gegen das SARS-CoV-2-Virus scharfgeschaltet.

Schnelle Codons

Bei genauem Vergleich der Virus-Ab schnitte mit BNT162b2 fällt auf, dass BioNTech an vielen Stellen die dritte Position von Codons ausgetauscht hat; wo im Virus-Gen ein A oder U steht, setzt BioNTech meist G oder C ein. Aber trotz etlicher geänderter Codons liefert BNT162b2 fast die gleiche Aminosäuresequenz wie das Spike-Gen.

Der Zusammenhang ist ziemlich trickreich. Die kurze Erklärung: Die Modifikationen beschleunigen die Synthese des Spikes. Beispielsweise fand man in älteren Experimenten, dass mRNAs mit hohem GC-Gehalt eine schnellere Proteinproduktion bewirken als solche mit hohem AU-Gehalt.

Die lange Erklärung: Zunächst sind solche stillen Änderungen überhaupt nur wegen der Kodier-Unschärfe von biologischen Systemen möglich. Sie nutzen zwar alle 64 möglichen Codons, müssen aber nur rund 20 Aminosäuren kodieren (proteinogene AS). Die dritte Position ist bei den meisten Codons unspezifisch (wobble), sodass oft mehr als ein Codon für dieselbe Aminosäure kodiert. Beispielsweise sind CGG und CGA beides Synonyme für die Aminosäure Arginin.



Bei Impfungen genügt es, das Abwehrsystem mittels Attrappen gegen einzelne Oberflächenelemente von Erregern zu trainieren. Das Spike eignet sich prima als Attrappe.

Dass die Codon-Wahl überhaupt die Synthesegeschwindigkeit beeinflusst, liegt vermutlich daran, dass die Zelle manche Codon-Synonyme seltener verwendet als andere, sodass auch die zugehörigen Träger, die die Aminosäuren zu den Ribosomen bringen (tRNA), unterschiedlich häufig sind. Einige Experimente deuten darauf hin, dass häufig benötigte Träger eher Mangelware sind, während die Zelle die selten benötigten immer auf Lager hat.

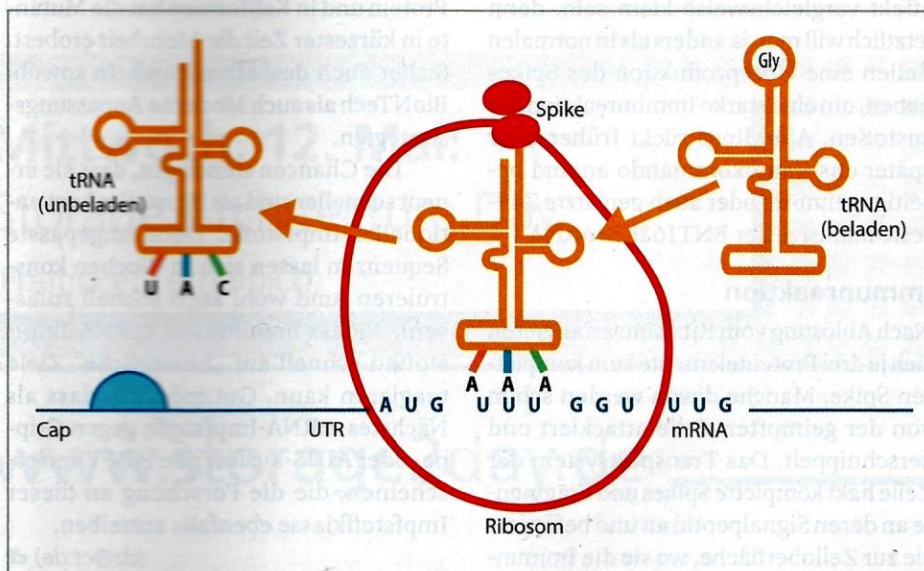
Pflanzenzellen nutzen das, um in der Not Stressproteine schnell herstellen zu können, zum Beispiel Hitzeschockproteine zum eigenen Schutz.

Drei spezielle Modifikationen gehören nicht zu dieser Kategorie. An der Position 79 verwendet BioNTech CCΨ anstatt CCA. Beide Triplets kodieren für die Aminosäure Prolin, sodass die Modifikation für das Signal-Peptid folgenlos bleibt und daher zunächst verwundert. Ihre Aufgabe besteht aber nur darin, Verschlingungen des mRNA-Strangs durch interne Basenanlagerungen zu verhindern (Stem-Loops). Stem-Loops bremsen das Ribosom, es kann nur entfaltete mRNA vorzugsfrei einlesen.

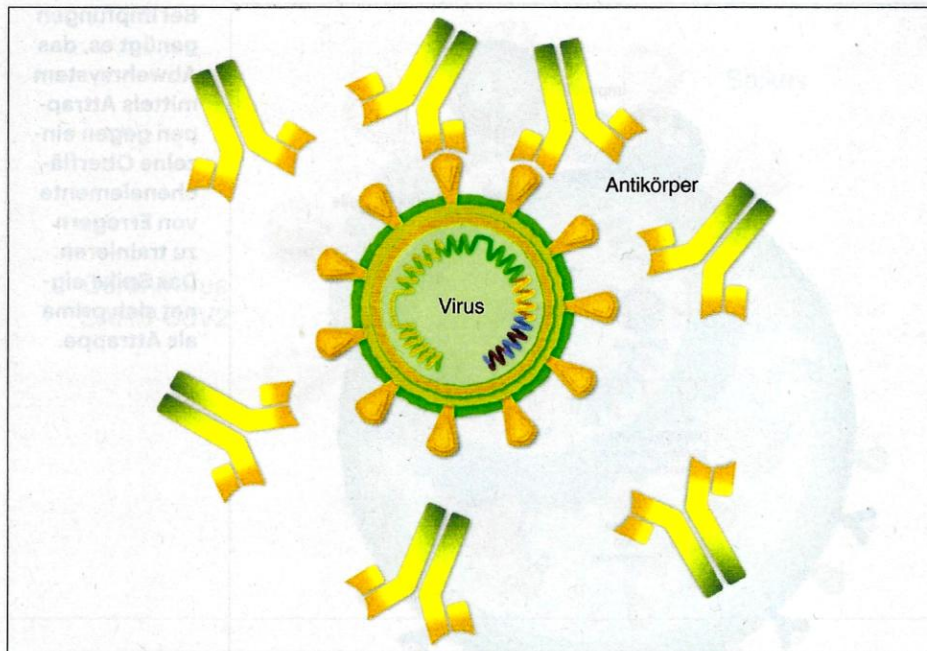
Mit und ohne Faltung

Die anderen beiden Codon-Änderungen von BNT162b2 gegenüber der ursprünglich in Wuhan isolierten (Spike-)Gensequenz hat man schon in der Natur als Mutationen gefunden. Sie sind daher in Gendatenbanken verzeichnet (Katalognummern K986P und V987P). Beide führen die Aminosäure Prolin neu ein, bei K986P anstatt Lysin und bei V987P anstatt Valin.

Diese beiden Änderungen sind nützlich, weil sie die Raumstruktur des Spikes stabilisieren. Die fertige Proteinsequenz faltet sich nämlich normalerweise nur durch Aggregation mit den übrigen Virus-



Bei der Proteinsynthese übersetzt ein Ribosom eine mRNA in eine Aminosäuresequenz (Protein); die Aminosäuren bringen die tRNA-Moleküle zum Ribosom.



Bei der Immunantwort können spezifische Antikörper Virenpartikel festsetzen und am Eintritt in Zellen hindern. Auch werden sie so zum Zersetzen markiert.

Proteinen zur Spike-Form. Wenn der Rest der Virus-Hülle fehlt, tendiert dieselbe Proteinsequenz zu einer anderen Raumstruktur. Damit wäre es aber keine geeignete Attrappe für das Abwehrsystem. Da Prolin vergleichbar einem Rohrknie aufgebaut ist, erzwingt es einen Knick in der Kette. Für die Spike-Form sind zwei solcher Knicke sehr förderlich.

Am Ende des Gens stehen gleich zwei Stopp-Codons (ΨGA, Position 3874), doppelt hält besser. Danach folgt eine zweite untranslatierte Sequenz. Sie dient dazu, dass sich das Ribosom und die mRNA geordnet trennen und verbessert ebenfalls die Syntheseeffizienz. Man kennt inzwischen diverse UTR-Enden der Tier-, Menschen- und Pflanzenwelt. BioNTech kombiniert hier zwei Elemente („Amino-terminal Enhancer of Split“, AES, und „mitochondriale 12S rRNA“). Dass der AES-Anteil einen starken Einfluss auf die Effizienz hat, haben Ugur Sahin, der Gründer von BioNTech und weitere Mitarbeiter erst 2019 herausgefunden.

Burn After Reading

Den Abschluss von BNT162b2 bildet ein weiteres Stück untranslatierter Sequenz, die Poly-A-Region. Diese wird genauso wie die vordere UTR-Region für ein korrektes Aufsetzen des Ribosoms beim Translationsstart benötigt. Das Poly-A-Ende an BNT162b2 besteht aus zwei Segmenten, separiert durch einen Spacer (GCAU-

AUGACU). Aus Experimenten geht hervor, dass die Segmentierung und der kurze Spacer ebenfalls die Syntheseeffizienz verbessern (siehe ct.de/ywa8).

Der Clou am Poly-A-Teil ist aber, dass Ribosomen nach jedem Translationsdurchgang ein Stück davon abknapsen und, wenn alle As aufgebraucht sind, das betreffende mRNA-Molekül in seine Einzelteile zerlegt wird. So stellt die Zelle sicher, dass ein mRNA-Molekül nicht beliebig oft abgelesen wird (mRNA-Halbwertszeit), sondern nur bedarfsgerecht.

In geimpften Zellen dürfte der Abbau-effekt vergleichsweise klein sein, denn letztlich will man ja anders als in normalen Zellen eine Überproduktion des Spikes haben, um eine starke Immunreaktion anzustoßen. Allerdings rückt früher oder später das Räumkommando an und beseitigt geimpfte oder auch geplatzte Zellreste inklusive der BNT162b2-mRNA.

Immunreaktion

Nach Ablösung vom Ribosom arrangieren sich je drei Proteinelemente zum kompletten Spike. Manche davon werden schon von der geimpften Zelle attackiert und zerschnippelt. Das Transportsystem der Zelle hakt komplette Spikes und Fragmente an deren Signalpeptid an und befördert sie zur Zelloberfläche, wo sie die Immunantwort auslösen. Wegen der Überproduktion des Spikes platzen manche Zellen und rufen das Immunsystem ebenfalls auf

den Plan. Dann nehmen antigenpräsentierende Zellen Spikes und Fragmente auf und präsentieren sie an ihrer Oberfläche. Wenn dann andere Helfer-Zellen solche Fragmente detektieren, T-Zellen etwa, schlagen sie Alarm und holen weitere Immunzellen zur Hilfe, um den Eindringling zu bekämpfen.

Hinzu kommen B-Zellen, die ebenfalls auf herumliegende oder auf geimpften Zellen aufgepflanzte Spikes treffen. Wenn sie daran binden, können sie durch T-Zellen aktiviert werden. Daraufhin produzieren sie Spike-spezifische Antikörper und schütten sie aus. Wenn die Antikörper auf Virus-Spikes anhaften, können sie es für die Zerlegung markieren. Bei kompletten SARS-Cov-2-Viren verkleben sie die Spikes, sodass sie keine Zellen infizieren können.

Antigenpräsentierende Zellen können auch Killer-T-Zellen aktivieren. Diese suchen und zerstören mit SARS-Cov-2 infizierte Zellen, die auf der Oberfläche Spikes oder Spike-Fragmente aufweisen.

Aussichten

Bisherigen Analysen zufolge eignet sich BNT162b2 nicht nur gegen die in Wuhan isolierte ursprüngliche Corona-Virus-Variante sehr gut, sondern auch gegen neue Mutanten, die in England und Brasilien gefunden wurden (B.1.1.7 und P.1). Bei der in Südafrika entdeckten Mutante B.1.351 ist die Wirksamkeit zwar abgeschwächt, aber immer noch ausreichend. Über weitere Mutanten, gefunden etwa in Kalifornien oder Japan, weiß man bisher wenig. Sie tragen ein stärker verändertes Spike-Protein und in Kalifornien hat die Mutante in kürzester Zeit die Mehrheit erobert. Sicher auch deshalb entwickeln sowohl BioNTech als auch Moderna Anpassungsstrategien.

Die Chancen stehen gut, dass sie erneut schneller sind als Hersteller konventioneller Impfstoffe. Denn angepasste Sequenzen lassen sich in Wochen konstruieren (und wohl auch schnell zulassen), sodass man mittels mRNA-Impfstoffen schnell auf „bewegliche“ Ziele reagieren kann. Gut möglich, dass als Nächstes mRNA-Impfstoffe gegen Grippe- oder AIDS-auslösende HIV-Viren erscheinen, die die Forschung an dieser Impfstoffklasse ebenfalls antreiben.

(dz@ct.de) **ct**

Forschungspublikationen,
BNT162b2-Code: ct.de/ywa8