

Derek Lowe's commentary on drug discovery and the pharma industry. An editorially independent blog from the publishers of *Science Translational Medicine*. All content is Derek's own, and he does not in any way speak for his employer.



[By Derek Lowe](#)

COVID-19

RNA Vaccines And Their Lipids

By [Derek Lowe](#) 11 January, 2021

So now that people (not enough of them!) are getting vaccinated in the US with the Pfizer/BioNTech and Moderna mRNA vaccines, let's talk about some more details of what are in those injections and what happens once the shot is given. The workings of an mRNA vaccine touch on a lot of different cellular processes and a lot of drug-delivery issues, so we can Talk Corona while also talking drug discovery, biology, and chemistry at the same time. I want to start off by [recommending this piece](#) by Bert Hubert on the workings of the Pfizer/BioNTech vaccine – Bert goes into a lot of detail that I'm going to run through rather quickly in the next few paragraphs, and you're probably going to have a better shot understanding it from him than you do from me!)

One theme that will show up many times in this post is that these vaccines were [not invented](#) from scratch. There's a long list of things that had to be worked on in order for the field to be in the shape it was in at the beginning of 2020, and that's why things ran so quickly. "RNA as a therapeutic agent" is an idea that has had billions of dollars of work poured into it over the last twenty or thirty years, so when you hear about these vaccines as something new, remember that's only for certain definitions of "new".

As all the world knows, these vaccines are based on messenger RNA (mRNA). That, of course, is the type that's produced in a living cell by reading off a given stretch of DNA and assembling the matching RNA, after which it goes off on its own to be fed into a ribosome which will assemble proteins based on its code, reading off by three "codon" letters at a time. So messenger RNA has its feet in both worlds, if it had feet: it's down there in the nucleus being put together next to an exposed and unwound strand of DNA, but afterwards it's also present right in the middle of the ribosome machinery, as amino acids get brought in and spliced together into a growing protein strand. Genetic information gets turned into proteins (there's the Central Dogma of molecular biology for you), and mRNA is how that happens.

Now, the specialists in the room will appreciate the *huge* number of details that go into both those processes. The concepts are pretty straightforward (read off DNA to make mRNA, read off mRNA to make protein), but the execution is something else again. It's worth going into those in a little detail to explain why the mRNAs in the vaccines look the way that they do, and why designing a good one is a lot harder than it looks.

As a new mRNA strand is generated by the action of the RNA polymerase II machinery on a stretch of DNA, it gets a "cap" attached to the end that's coming out from the DNA (the "5-prime" end), a special nucleotide (7-methylguanosine) that's used just for that purpose. But don't get the idea that the new mRNA strand is just waving in the nucleoplasmic breeze – at all points, the developing mRNA is associated with a whole mound of

specialized RNA-binding proteins that keep it from balling up on itself like a long strand of packing tape, which is what it would certainly end up doing otherwise.

So the 5-prime end is capped, and then the other one (the “3-prime” end) undergoes some processing of its own. It has a certain number of residues scissored right back off, and then a stretch of “poly-A” (one adenosine residue after another) is added on – these processes are done by another big complex of enzymatic and scaffolding proteins working on that end of the molecule. By the time that’s finished, an mRNA can have a couple of hundred A residues tailing off its 3-prime end. This doesn’t get turned into protein, though – otherwise every protein that gets made would come out of the ribosome with a long tail of lysines on it, since the “AAA” codon under other circumstances means “Lys” to the translation machinery.

Then there’s another key step. In most organisms, the DNA doesn’t just read off the uninterrupted code for a whole protein. It has interruptions of other stretches of code (“introns”), and at this point those are clipped out and the actual mRNAs are spliced together by assembling their pieces (the “exons”) into their final form. That may seem like a rather weird process if you haven’t run into it, and it certainly was a surprise when it was discovered back in the late 1970s. This is done by yet another Death-Star-sized mass of proteins, the “spliceosome”, and it provides opportunities for “splice variants” along the way that will produce different proteins when a ribosome gets ahold of them. And that’s a big reason why we have a lot more different proteins in our bodies than we have different genes: many of them can be mixed-and-matched into these different variants back at the mRNA level.

I mentioned the poly-A tail, but there are also key regions at both the 5-prime and the 3-prime ends of an mRNA strand that also don’t get translated into protein. These contain important regulatory information for how that translation should go. There are “start” and “stop” codons that aren’t always associated with any particular amino acid (the start codon can also code for methionine, depending on the context), but rather convey those instructions to the ribosomes. The “leader” sequence at the beginning of the mRNA and sections at the other end as well can have profound effects on how readily it gets taken up by any given ribosome and how efficiently it moves through. Ribosomes themselves have at least two ways to feed an mRNA into their protein-making machinery: the normal way, which requires a “capped” mRNA and an “internal ribosome entry site” (IRES) that doesn’t care, and the use of these is also mediated by the untranslated RNA regions. It goes on and on! The last 30 or 40 years of biology have seen these details brought to light through vast amounts of effort in the lab (and similarly vast amounts of staring out windows trying to sort out mentally what’s going on), and that process is nowhere near at an end.

I’ve rambled on about all this to bring us back to the mRNA vaccines. You can see from that quick tour of the machinery that it would be a bit too hopeful just to produce a plain stretch of RNA that codes for the viral Spike protein and expect that to work right off the bat. No, you’re going to have to optimize both ends of it so that ribosomes are enthusiastic about it and zip right down the strand producing that Spike for you. (And remember, the vaccines we have are also producing a variation of the Spike that keeps it stable in its final active shape, the better to have antibodies recognizing that, so you’re not even coding for the “native” Spike from the very beginning).

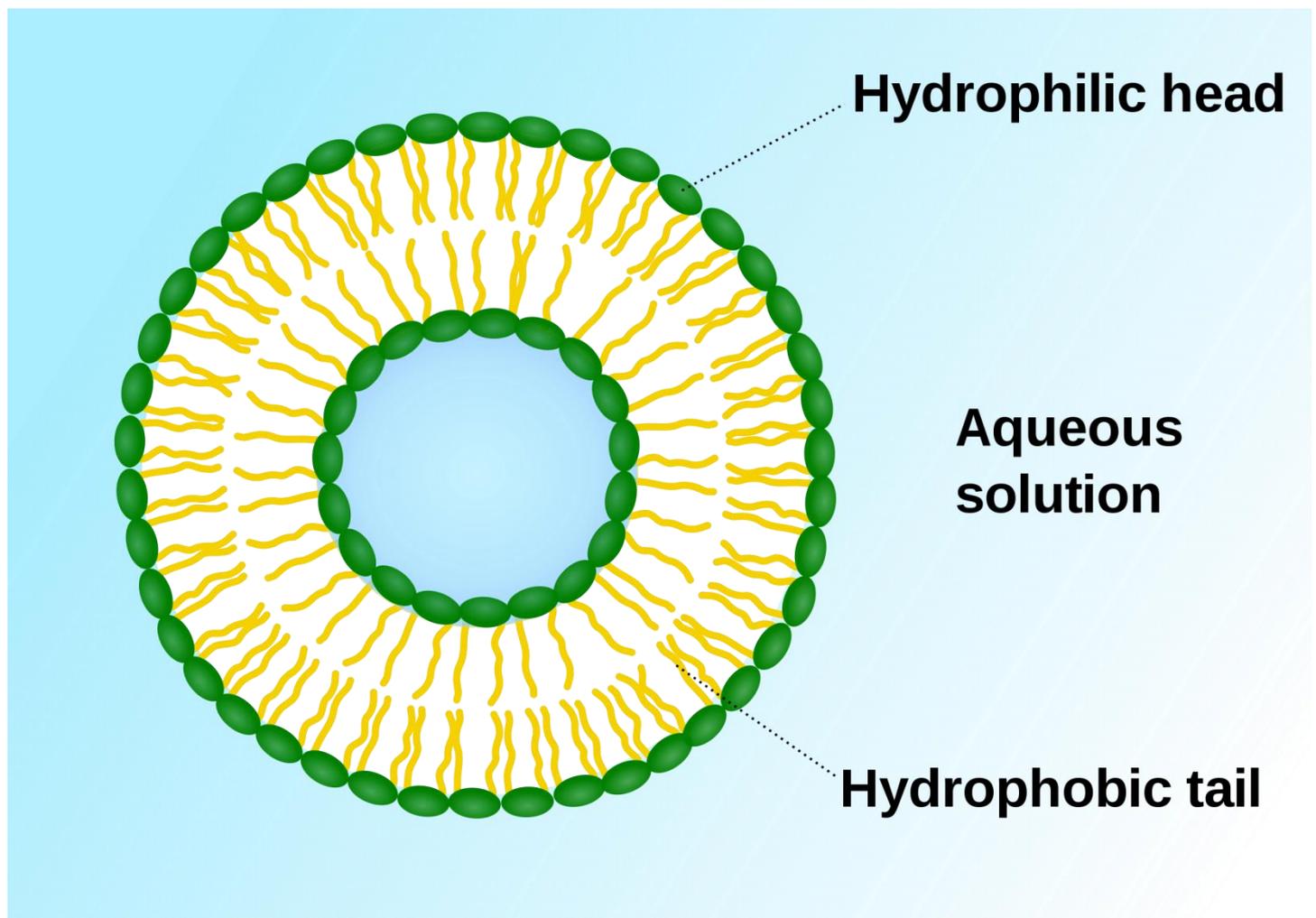
And as you’ll know if you’ve read that article from Bert Hubert that I linked to at the beginning, the mRNA vaccines also feature a good deal more such engineering. The three-letter codons for amino acids have some redundancies in them, but not all of those are processed with the same alacrity. Ones that are heavier in C and G residues seem to be run through more efficiently, so the sequences are biased that way. There are also the modified bases like pseudouridine/1-methylpseudouridine that get read off at the ribosome like their native cousins (in this case, good ol’ uridine, U) but make the mRNA strand both more stable and less likely to set off an immune response against itself. So the sequences in the vaccines have human fingerprints all over them – see Bert’s article for more.

But all that engineering availeth one not if the mRNA doesn’t make it to the cells and inside the cells. And that takes us to the formulations, which are another essential part of the whole mRNA vaccine story. Cell and molecular biologists tend to think of RNA molecules in general as pretty fragile things, and that reputation has been earned. They’re intrinsically less stable than the corresponding DNA molecules, and the odds are further stacked against them in the body by our own immune system’s defenses against foreign RNAs from pathogens like the current coronavirus. Just for starters, there are plenty of “RNA-ase” enzymes out there ready to tear any wandering RNAs to bits – the body can use circulating RNA molecules as signals, but these things are under

tight control. So if you just inject a naked RNA sequence into someone's blood, it'll get stripped down to nothing before it's traveled very far.

What are your alternatives for a more suitably clothed RNA? Well, as mentioned earlier, mRNA vaccines are not a new idea, nor is the idea of therapeutic RNA in general (remember siRNA?). So there's been a lot of work over the years to find suitable carriers (see [this 2016 review](#) for an overview). It was not obvious which of these possibilities (lipids, carrier proteins, synthetic polymers, and more) would work out, of course. The only way to find out was (and is) to spend the time, spend the money, and go run the experiments. One thing that many of these ideas have in common is the carrier molecules having numerous positive charges on them, though, because RNA (and DNA) have lots of negatively-charged phosphate groups, and these would match up together to form a stable complex. Results from those experiments have [tended to elevate](#) the idea of lipid nanoparticles as a carrier, because they can help out in two ways simultaneously: they protect the mRNA construct itself as it travels through the bloodstream, and they seem to help it cross cell membranes and get from the blood into its destination. That's not something you can just assume is [going to happen](#) on its own.

That point deserves a quick elaboration, because one thing that you have likely noticed is that there's been a lot more work during this pandemic on RNA vaccines as opposed to DNA ones, even though DNA has that stability advantage mentioned above. There are several reasons for that, but one big one is that an RNA payload just has to get into the cell to encounter its site of action (the ribosomes, which are all over the place). A DNA therapeutic, though, has to get into the nucleus to do anything, and that's yet another membrane to cross (and one with its own set of properties and gatekeepers). There's also the possibility for a DNA species to get mistakenly incorporated into a cell's own genome, which for a vaccine you don't want (as opposed to a gene therapy), and using RNA completely takes that off the table, but the "just get into the cytosol" advantage is a real one, too.



So what are these lipid formulations like? They've been investigated for many years themselves, because these sorts of carrier properties could of course be useful for a lot of other therapeutic agents beside RNA. [Here's](#) a short article at *STAT* about them. There are a lot of variations on the lipid idea, and one kind involves a sort of spherical bubble of lipid (a liposome) – generally a bilayer, as with our own cell membranes, because lipid

molecule just naturally stack up like this, with greasy interior layers and the polar parts facing the solvent on the outside (see above, illustration by SuperManu via Wikipedia). In this case, the “hydrophilic head” will tend to incorporate some sort of positively charged group (as mentioned above). The payload will be in that little blue area in the middle, safe and secure as it drifts along. The lipid nanoparticles being used now are more of a solid lump, with the RNA and the lipids mixed together into tiny masses. The cell membrane is largely made of phospholipid bilayer, with the outside hydrophilic part being negatively charged, so these positively charged nanoparticles have all the more reason to stick to them.

When that happens, it appears that endocytosis kicks in, the general process of importing larger particles into a cell. There are several varieties of endocytosis, but they tend to end up with the external particle emerging on the other side of the cell membrane wrapped in a new endosomal vesicle of its own (can’t be too careful, from a cellular perspective). A well-chosen lipid nanoparticle formulation can actually help the RNA payload escape such an endosomal compartment and finally make it into the cytosol itself, ready for action.

Now we get into a forest of picky details. There is also no way to be sure from first principles which of the many, many, many possible lipid nanoformulations is going to work out the best for carrying therapeutic mRNAs. Small amounts of various other lipid species present in the bilayer can affect their properties a great deal, so you have a lot of experimentation to do and lessons to learn, and years of work have already been spent on just that sort of thing. For example, one broad lesson has been that nanoparticles formed from lipids that have permanently charged head groups (like quaternary amines) don’t seem to perform as well as ones made from amines that are charged by having ionizable H atoms on them. You don’t want to have to discover all this on your own at the same time you’re working out the details of the RNA construct, so therapeutic development has almost invariably been through partnerships.

The Pfizer/BioNTech vaccine uses lipid nanoparticles [developed by](#) the Canadian company Acuitas, who have (under one name or another!) been working in this area for [over a decade](#) now, [trying out](#) countless variations on various lipid combinations. Back then, it was mostly for siRNA delivery, but the lessons learned from that work have been invaluable for mRNA vaccine delivery. Meanwhile, Moderna has been involved in a [vigorous](#) and [long-running](#) patent dispute with a smaller company called Arbutus, who have also been investigating lipid nanoparticle formulations and whose technology Moderna once licensed. Arbutus has been claiming that Moderna’s research programs (and indeed their now-launched vaccine) avail themselves of Arbutus’ intellectual property, while Moderna (naturally) disputes this with equal vigor. I Am Not a Patent Attorney, and a damn good thing, too, so I have no useful opinion about who’s in the right. If Arbutus has a case, I would expect them to eventually get a judgement giving them some royalties off the Moderna vaccine, but my only solid prediction is that a number of lawyers will have steady employment thanks to this issue for some time to come.

A [closer look](#) at the Pfizer/BioNTech vaccine shows that it has four lipid components, two of which appear to be proprietary to Acuitas. One of these is [ALC-0315](#), and the other is [ALC-0159](#). You’ll note that both of those are tertiary amines (protonated to a positive charge under physiological conditions) and not quaternary charged ones, for the reasons mentioned above. The other two lipids are 1,2-distearoyl-*sn*-glycero-3-phosphocholine ([DPSC](#)), which is a well-known [phosphatidylcholine](#) lipid (as evidenced by the number of references in that link) and cholesterol, which is rather better-known still. These four components are of course present in a specific ratio, which I would rather not try to exfoliate out of the patent filings. But that should give you some idea of what’s in a formulation like this and what the lipids themselves look like. The physical process by which you reliably prepare such nanoparticles is another thing that needs experimentation, of course, but they’re cranking out the vials as we speak.

So that’s a look under the hood, and as promised, there’s a lot in there. It’s all the more remarkable that these therapeutics came together as quickly as they did, but if it had not been for the years of prep work in all of these areas, we would still be waiting!