

**PODCAST INTERVIEW** with Markus Gershater,  
Chief Scientific Officer at Synthace



“As an industry, we need to raise awareness that these capabilities are out there and also support people on their journey towards realizing the value from them.”

## Introducing flexibility to automation to unleash the power of biology

Markus co-founded Synthace after working as a Research Associate in Synthetic Biology at University College London where he developed novel biosynthesis methods using pathway engineering. Prior to UCL, he was a Biotransformation Scientist at Novacta Biosystems working as part of the industrial biotechnology group that conducted more than 90 contract research projects for over 20 clients. Markus has a PhD in Plant Biochemistry from Durham.

*Cell & Gene Therapy Insights* 2019; 5(12), 1691–1696

DOI: 10.18609/cgti.2019.177

**Q** Can you give us some brief background on what Synthace does, and why it has relevance to current trends and challenges in cell and gene therapy?

**MG:** Synthace is fundamentally about automation: automatic lab processes and also the data processes associated with those experiments.

The fundamental issue is that automation is inflexible. Once you have a process programmed then it can do that one thing, but as soon as you want to change something it becomes an issue. We make software that allows the flexible reprogramming of automation on both the lab and the subsequent data processing sides. This enables an escape from low throughput manual processes, which despite being arduous, variable and error-prone, are still the norm in biological R&D today.

This includes cell and gene therapy, of course, but it is not limited to it – biological R&D across the board is a lot more manual than we would all like it to be, and this is purely down to the inherent inflexibility of traditional automation. With a more flexible approach, we can bring automation into this space a lot more than was previously possible – that’s the potential we are seeking to unlock.

The interesting thing is that we started out as a bioprocess development company, not as a software company. We were looking to do ever more sophisticated experiments, for which we ultimately needed automation, and so we built the software to program that automation in the lab. Our chief focus was on addressing the complexity of biology for our own ends, but we came to realize that this software could enable not just our own lab work but everyone else’s, too – so it became a product its own right and now it is at the core of our business model.

**Q** What differentiates Synthace’s data message from those of other solution providers in the cell and gene therapy space?

**MG:** For me, this comes down to the fundamental way in which people are thinking about digitization as a whole. They tend to think, “OK, the digital output of what we’re doing is data, therefore all of our digitization problems have to do with what we want to do with that data.”

We see things differently. We want the digital world to reach much further into what we’re doing – to play an active role in helping us to define a much more sophisticated experiment in the first place, which then produces these data.

Because the digital world is already integrated with the experiments, we can better understand the context and structure of the data that is produced. We kind of flip the whole problem on its head: it’s not a question of, “we have this huge mass of data from our experiment – what can we do with it?” Instead, it’s about how we are generating the data in the first place. This approach gives us much more direct insight into the biology than we would get if we were trying to piece everything back together from all the various data received from different pieces of lab equipment.

**Q** Tell us about some of the key common misconceptions you encounter regarding data strategy in this relatively immature sector.

**MG:** Firstly, I’d mention that I don’t think this is limited to cell and gene therapy. Actually, the entire biopharma industry seems to be thinking about this in the reactive sense of what to do with all this data it’s generated.

Currently, there are people throughout the industry stating that they want to employ AI to transform their sector. We fully agree that AI will transform the industry, but it will only do so once the fundamental basis that is the routine production of structured, beautiful data sets is in place. The data are there solely for us to gain a better understanding of biology, whether or not that’s through efforts helped or augmented by AI. An issue I see at the moment is that people are seeing

data and experiments in isolation, as their own things, instead of seeing the data as the natural product of experiments, which are themselves there purely for us to gain insight into the biology.

So we see the need for a much deeper digitization strategy, which boils down to removing all of the obstacles between us and the biology. These obstacles are all the manual data structurings and other pieces of manual work that are typically required. We’re basically trying to disintermediate between the biologists and the stuff they’re trying to work with.

This idea of automated sophisticated experiments naturally producing sophisticated data sets feeds into the future potential for things like machine

“ Our chief focus was on addressing the complexity of biology for our own ends, but we came to realize that this software could enable not just our own lab work but everyone else’s, too. ”

“Cell and Gene Therapy is a field where the therapeutic modalities are more complex by orders of magnitude than anything we’ve had to deal with before in the biopharma industry.”

learning, which in turn will enable the design of even more sophisticated experiments which can be carried out in an automated way. We call this overall ecosystem of tools and capabilities computer-aided biology (CAB).

CAB is quite a specific route towards unblocking or unleashing all of the powers of the digital world within biology – to harness them to help us really grapple with biological complexities, especially in the area of cell and gene therapy. This is a field where the therapeutic modalities are more complex by orders of magnitude than anything we’ve had to deal with before in the biopharma industry.

**Q** What changes in both culture and enabling tools are needed to facilitate a shift to a stronger data strategy in cell and gene therapy companies?

**MG:** We envisage an ecosystem of tools that enable automation in the lab and data structuring – that’s where we see ourselves playing. But it’s important to be clear that we’re not trying to do everything ourselves.

When the structured data are produced, there are all sorts of ways in which they might then be analyzed. Anything from basic statistics all the way through to very sophisticated, deep learning methodologies could be used, according to what is most pragmatic and appropriate at the time.

There will of course be lots of companies out there developing those kinds of tools. When I talk about automation in the lab, we are not actually making the hardware. We’re relying on a fantastic ecosystem of physical tool providers – of automation manufacturers and analytics manufacturers. What it’s really about is the seamless integration of these fantastic physical tools that are already out there, and of new ones that might come through in the future.

At the moment, where this ecosystem exists, it is very fragmented. We believe that as an industry of technology providers to the cell and gene

“...as an industry of technology providers to the cell and gene therapy space, we need to be thinking about the best ways of integrating not just in terms of the physical devices in the lab, but also between the digital tools.”

therapy space, we need to be thinking about the best ways of integrating not just in terms of the physical devices in the lab, but also between the digital tools. This will allow a cell and gene therapy scientist to be able to use whatever is most powerful for the task they need to carry out – for example, if they have a favorite bit of data visualization or exploration software that they really want to use, then they should be able to use it. There shouldn’t be walls up between these different things that will slow everything down.

Regarding culture, I don’t expect scientists to have to change dramatically in order to use the tools. We see scientists as being very much within the loop. You sometimes hear people talking about AI as though everything is automated and machine learning drives it all. But that would work only if the machine learning somehow had all the expertise of all the different scientists coded into it somehow. To me, the optimum strategy would be to take all the power of machine learning and all the expertise of the scientists and bring it together.

This means you are creating tools that augment the capabilities of the scientists, they don’t replace them. The cultural shift is therefore really one that means we’re looking for people to start using these tools and start realizing what that means for how they can go about their science. It isn’t some sort of fundamental thing where everyone has to learn to code, for instance (although that is another misconception one sometimes hears). While there will certainly be some parts of the industry where that will be very helpful, without a doubt, the main thing for a scientist to think about is, “OK, now I can now do these much more sophisticated experiments within the lab, what does that allow me to do as a scientist? What problems does that allow me to address?”

The real shift is one to a much more systematic way of doing experiments – moving away from a kind of stepwise exploration of a space to a much more comprehensive characterization of a particular biological system. That’s what’s being enabled by these kind of tools – that’s the power they offer.

**Q** Let’s talk about Design of Experiments (DOE) and its utility in the cell and gene therapy space. Firstly, how is DOE being implemented in the wider biopharma space today, and with what impact?

**MG:** DOE is a somewhat unhelpful term in that it refers to something very specific, although the name doesn't really suggest that.

What it actually relates to is multi-factorial experimentation. The traditional approach to experimentation is to look at one factor at a time as you go along. For example, you might firstly look at the impact of temperature, then move on to the effect of a particular cytokine, and so on. DOE takes the alternative approach of asking what are all the different things that might affect our process, and how can you prioritize a subset in order to investigate all of them simultaneously?

For people who are less familiar with the mathematics involved, this sounds very unlikely to succeed. But it's actually a very well-developed branch of maths and it's been used for decades, although unfortunately, not nearly as much as it should have been in biology.

Biology is fundamentally an interconnected system, where you have lots of things coming together and then phenomena emerging out of the combination of lots of different simultaneous factors. What DOE allows is an unpicking of all of those different interactions in order to get to the underlying cause or causes. Getting to those causes enables you to really address the complexities inherent in biology.

In industry at the moment, that power to understand something more holistically is being applied only where it's absolutely critical that biological complexity is properly nailed down. For example, when we're producing a therapeutic, we want to know we're producing something that's going to help people and not hurt them. That process must obviously be exceptionally well understood and so the FDA demands that DOE is used as one of the methods of characterizing a biological process that's going to make a product.

But unfortunately, from my perspective, it is mainly being used as this kind of regulatory compliance tool, as opposed to a tool with enormous power not just to help you understand a biological system, but to enable you to engineer that system a lot more predictively than would otherwise be possible.

I think that cell and gene therapy has this opportunity to not just use DOE as a tool for regulatory compliance, but to wield it to help address the extremely high levels of complexity within the space. If we can get to that higher level of understanding, it will result in all of the things we are producing become that much more scalable, that much more tractable, that much more engineerable. We'll be able to roll these products out to all the patients who need them, as opposed to the few we've managed to treat as a relatively nascent therapeutic field to date.

“...things get really exciting when you don't just use Design of Experiments and automation in isolation, but in conjunction. At that point, you can do high throughput, sophisticated experiments.”

However, things get really exciting when you don't just use DOE and automation in isolation, but in conjunction. At that point, you can do high throughput, sophisticated experiments.

High throughput has been used before in the therapeutics industry, of course, for things like screening. Those are often pretty unsophisticated experiments, though – you're basically just posing the same hypothesis millions of times. DOE, on the other hand, can pose far more sophisticated hypotheses in a much more holistic way.

If we can take these sophisticated ways of experimenting and make them high throughput, then what can we achieve? Well, what's really exciting at the moment is we're just starting to see the impact within cell and gene therapy of exactly this kind of method. For example, Oxford Biomedica has been using Antha for a number of years now, and we released a joint case study where they used automated DOE to optimize transfection and transduction at the heart of their lentiviral vector production process. They got an order of magnitude increase in yield from properly addressing the different factors that might affect that transfection/transduction.

Beyond the very positive result, you could also see them starting to change the way they think about their science. So we come back to the cultural aspect: you give people new tools and new capabilities, and these things become transformative. It's quite remarkable when you first start using them to see just how powerful they can be. You often find people who haven't used DOE before become really evangelical about it, because of the step change in the amount of power it provides to address biological complexity. The next thing those individuals ask themselves is 'what other problems can I apply this to?' In this way, it becomes a part of their thinking, and without the need for a major cultural shift, these quite transformative tools become endemic within an organization.

**Q** Can you go deeper on where specifically you see DOE bringing benefits to the cell and gene therapy field? And what will be the key obstacles to overcome before its full potential can be realized in this space?

**MG:** DOE is a statistical tool. It's a general method of being able to pose lots of sophisticated hypotheses simultaneously. In that respect, it can be used extremely widely.

“We can use Design of Experiments to make sure that all the different components of a media are properly balanced, which in turn ensures we’re differentiating robustly to a particular cell type...”

In our own labs, we use DOE whenever we have a process to optimize. For instance, if we have an analytical process that has too much variance and we need to tighten up those error bars, then we can use DOE to make sure it’s just that much more robust – that the precision is really dialed in and we get high quality data from it.

You can use it on much more complex processes, too, as in the case of Oxford Biomedica. We’ve also used it for optimizing all sorts of molecular biology methods in the lab, as well as cell growth methods and media optimizations. We can use DOE to make sure that all the different components of a media are properly balanced, which in turn ensures we’re differentiating robustly to a particular cell type, for one example, or that we can make organoids in a robust and reproducible manner, for another.

Wherever there is any complexity, then we need to be using these more powerful tools in order to be able to ask more sophisticated questions. And in cell and gene therapy, complexity is everywhere!

We can also think about how DOE can be applied in the cell and gene therapy space in a similar way to how it’s already being applied within the broader biotherapeutics area – in antibody production, for example. That relates to how we can use DOE to really understand the production processes that are required to make our therapies. Once we understand the production processes that much better, then we know that even given the diverse inputs that we often encounter with cell therapies in particular, we’re always going to get to a high-quality product that is suitable for the patient.

The issue with DOE is that there is a learning curve. This is a different way of thinking about science to the way we are all taught through school and university. Indeed, I think it’s one of the major problems we face with the way science is being taught today: these much more powerful experimental techniques are just not ‘baked in’ from the start. This means there is a bit of a cultural shift to negotiate, a bit of a knowledge gap and also a trust gap. These tools sound really powerful, but I’m sure there are a lot of people reading who remain skeptical, and that’s entirely correct. As biologists, we should be skeptical about things – I was profoundly skeptical before I started using them. But when you start to see the data coming out, that’s when you get excited.

So there is this gap between first hearing about it and actually receiving those first data and becoming really excited about what they’re showing you. It is not something that is significant, but it does need to be addressed. As an industry, we need to raise awareness that these capabilities are out there and also support people on their journey towards realizing the value from them.

**Q** Why is it so important to push Quality by Design (QbD) further upstream in cell and gene therapy R&D?

**MG:** What is QbD? QbD is basically a system, a framework in which you can think about all of the process that you’re addressing, all the biology, and consider what are all the things that could contribute to that biology not working.

You start off with something called the root cause analysis. This is where you consider all the different inputs and ask which of them could vary, or what happens if the lab temperature is different from this day to that day, etc. There are lots of different things that could contribute to variability or failure within biology.

QbD therefore begins with really in-depth thinking, which I think is something we don’t pursue a lot of the time – we tend to think about the things that are more immediately in front of us, as opposed to all of the things that could potentially go wrong. I guess it’s quite a negative way of thinking!

But what it does give you is a list of all of these different things that could result in or contribute to problems further down the line. And we do have the tools to address a lot of these things. For example, you can then use DOE in order to explore the potential issues systematically and see which ones really matter, and which ones might not.

In our own labs, we use this kind of methodology just for routine lab tasks. For instance, you want a PCR to work every single time – well, if you actually do this kind of analysis and you do the experiments associated with it, you get that PCR nicely optimized and it will work every time. You don’t have to go back and redo things. Fundamentally, what we’re looking to do is build that foundation of quality, which then means we can proceed to the much more interesting and meaningful questions of how we can actually develop and produce these therapies in a really reliable and scalable manner.

**Q** What does the lab of the future look like to you, and what tangible steps and marginal gains can be achieved today to put cell and

gene therapy companies on the right path to realizing this vision?

**MG:** I think 'lab of the future' means a lot of things to a lot of different people.

However, we think that within whatever vision people might have, there needs to be this component of automated lab processes, and then the automated structuring of the data that comes from them to make really high quality, contextual data sets. There could also be some kind of machine learning, which is usually another component of most people's labs of the future.

It's really just a subset of the overall lab of the future – there will be other technologies that are needed as well – but this computer-aided biology vision is something we're looking to define quite clearly. So in contrast to a much more expansive vision of the future, if you like, we're saying 'look, this is something that is obviously very powerful and that could also build the foundation for something a lot more exciting going forward, and these are the steps we can take to get there'.

**Q** Finally, can you summarize what needs to happen over the decade ahead if cell and gene therapy is to fully capitalize on the promise of automation and machine learning by 2030?

**MG:** I think it's actually a reasonable timeframe. I don't think it's too unrealistic. That's because there are a lot of pressing issues right now, and a lot of the sensible ways of addressing them are through the technologies we've been talking about. It's not as though we're expecting people to make a huge leap – in terms of culture, for instance, as we've discussed.

Overall, I'm pretty optimistic. I think there can be some clear arguments made that are based purely on hard-headed things like return on investment from automation, and how we can get better data integrity – higher quality data, data that is actually put in the context of the experiment it comes from. These are all perfectly logical things we want to do. Again, I don't think there needs to be a massive leap forward. In fact, when you do see people try to make that direct leap towards an AI-augmented future, they tend to spend a lot of time later making up for the fact that the foundations weren't really there in the first place.

So it's just about get those foundation things in place: making sure we're building on really high quality, automated protocols, both for lab and for data. I believe we will then get there quite naturally through the curiosity

of scientists who are motivated to solve problems, because they're such important problems to solve.

#### AFFILIATIONS

**Markus Gershater**

Chief Scientific Officer, Synthace

#### AUTHORSHIP & CONFLICT OF INTEREST

Contributions: *All named authors take responsibility for the integrity of the work as a whole, and have given their approval for this version to be published.*

Acknowledgements: *None.*

Disclosure and potential conflicts of interest: *The authors declare that they have no conflicts of interest.*

Funding declaration: *The authors received no financial support for the research, authorship and/or publication of this article.*

#### ARTICLE & COPYRIGHT INFORMATION

Copyright: Published by Cell and Gene Therapy Insights under Creative Commons License Deed CC BY NC ND 4.0 which allows anyone to copy, distribute, and transmit the article provided it is properly attributed in the manner specified below. No commercial use without permission.

Attribution: Copyright © 2020 Gershater M. Published by Cell and Gene Therapy Insights under Creative Commons License Deed CC BY NC ND 4.0.

Article source: Invited.

**Revised manuscript received:** Dec 18 2019; **Publication date:** Jan 23 2020.

Synthace



