I, robot
I, the future

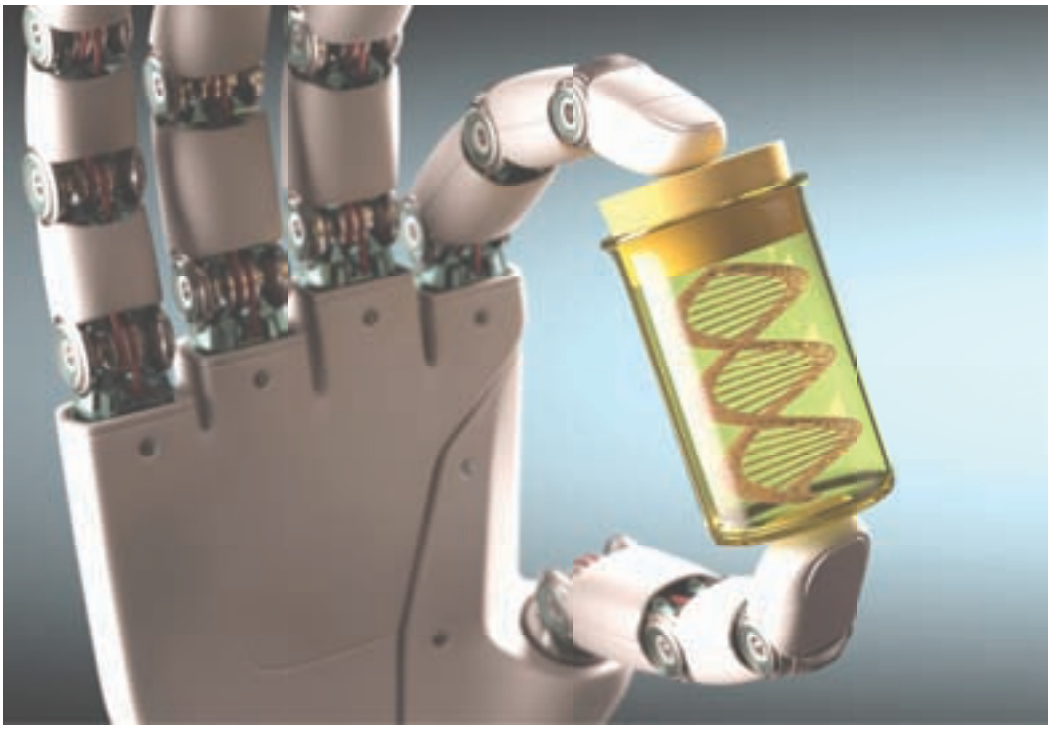
The Pistoia Alliance, whose membership includes most of the world's major pharmaceutical companies, technology vendors and publishers active in the life science R&D sector, is keenly interested in discovering the ways in which innovation through pre-competitive collaboration can be used to explore and exploit the constant stream of new technologies that become available. Recent interest in areas such as the Internet of Things, smart glasses and machine learning have all led to initial exploration through webinars and conference panels, but none seem to have generated quite so much interest as that of robotics both within and beyond the lab.

By Richard Holland

At the Pistoia Alliance conference in London in April 2016 we asked the audience a simple question: what is a robot? Unsurprisingly, most of the answers suggested that robots were machines, usually computer controlled, for automating repetitive tasks. However, we are already witnessing the first robots that are constructed not of cogs and levers but of proteins and chemical compounds, and have their behaviour programmed with encoded strings of DNA in place of computer software. These biological robots are changing the face of life science from early stage R&D right through to applications in clinical medicine, and, like it or not, as researchers and patients we must accept that they are here to stay.

We tend to think of robots as physical machines in the lab that automate repetitive work; tasks that can be done more efficiently or effectively by an automated mechanism than by a human, or to do things that lab staff otherwise simply could not do at all due to safety issues or the natural constraints of the human body. The repetitive nature of this kind of work is well suited to a mechanical robot that does not tire as a human does, although downtime cannot be entirely avoided as they still have to be turned off occasionally for maintenance. Most importantly, such robots allow the operating scientist to focus on cognitive work such as their research rather than spending hours performing laborious manual experimental procedures.

Robots are not new to the life sciences, with examples such as DNA sequencing machines, automated liquid volume dispensers and plate movers all in widespread use in labs around the world. It would be unusual to visit a modern lab that does not have at least once piece of robotic equipment in action. All are electrical appliances, some with electric motors to allow movement, others with electrically-driven compressors to power hydraulic or pneumatic actuators on robotic arms. The more complex examples include on-board computers



that constantly perform calculations to monitor and adjust their behaviour mid-task, and the most complex of all allow scientists to programme them through the configuration of specialised software, or even train them by demonstrating tasks to artificial intelligence algorithms so that they can learn by example.

Software, defined by the mathematical concept of a Turing machine – machine being the operative word – is itself a robot, powered by, but independent from, the computer hardware it runs on, in much the same way that an electrical robot is independent from the electricity network. Performing calculations and making decisions rather than carrying out mechanical actions, software automates a scientist's work just as a physical machine can, taking away the drudgery of computation and interpretation of many common data-related tasks, accelerating the production of new data and permitting the scientist more time to concentrate on the finer detail of their research.

If software is a robot, then any autonomous entity built or programmed by humans to perform a defined task could also be considered so. Machines, electrical or not, do not have to be made of metal and plastic, and software does not have to run on digital computer hardware. The original concept of a Turing machine defined software only as a mathematical model, a predefined sequence of logical decisions made over a set of inputs, and that model can be implemented in any medium that permits the logic to be represented and executed in a predictable manner. Thus we can extend the idea of robotics in life sciences into living technologies and nanotechnologies where biological systems are engineered to do particular tasks in a reliable way. The biggest difference between the mechanical or software robots we are familiar with and the biological or nanotech robots in living technology is that the latter are built bottom up, often self-assembling, and they are generally encoded using the programming language of DNA.

Taking a simple organism with desirable traits, such as a virus which in its naturally occurring form is able to navigate its way through the human body to locate an optimal environment to inhabit and interact with, the behaviour of that virus can be adjusted by manipulating its DNA. While preserving its ability to move freely around the body, the precise part or parts of the body which it will target can be adjusted by modifying or entirely replacing the genes which control that behaviour. The way it behaves once it reaches its target can also be modified in a likewise manner, allowing these viral robots to perform such tasks as delivering drugs directly to diseased cells within an organ. Essentially a modified virus behaves like a USB stick in that it is capable of docking with and loading

customised content into individual cells. The concept can even be recursive; it is possible for the viral payload itself to be a portion of DNA that in turn modifies the DNA of the targeted cell, which has particularly interesting potential applications in the field of cancer research. With developments in Precise Genome Editing (PGE) through the rapid understanding of CRISPR technology, it will become possible to develop systems that will allow the tissue targeted modification of the sequence or level of expression of any gene, theoretically in any cell type.

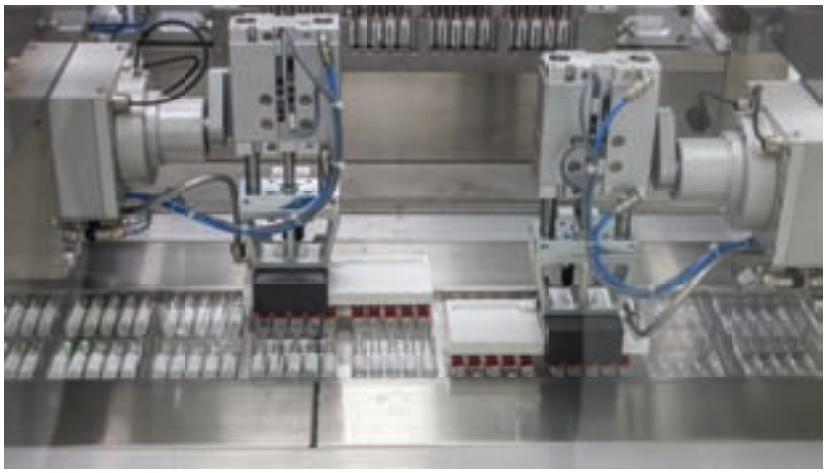
Viral robots are present-day reality and conceptually not far from being introduced into clinical settings. The use of viruses in research and clinical therapeutics has been increasing in recent years and we are already capable of making fully synthetic viruses where the entire genome of the virus is human-designed, replacing all of the natural virus' DNA. The whole viral robot is engineered from scratch based on a template provided by nature that makes it completely programmable right down to the atomic level. Yet, the abilities of a viral robot are still limited by the necessary constraints of biological, chemical and physical morphology that are required to make it identifiable as a virus and have the human body recognise and interact with it as such. Even more therapeutic potential could be unleashed if these constraints could be removed.

DNA nanomachines remove the limitations of virus morphology by creating completely new autonomous structures that have no historical biological context whatsoever and bear no resemblance to any natural biological construct. These microscopic mechanical robots are constructed at the atomic level from chemical components that are very different from their benchtop ancestors, yet they behave in much the same way. DNA nanomachines use DNA-encoded software to passively react when they reach and detect their intended context, or actively control their physical motion by exploiting the chemical properties of their constituent parts. They are constructed using a technique known as DNA origami[1] combining proteins and chemicals with a DNA backbone or container that doubles up as the control logic. With this technology, programmable biological nanomotors become possible, driving nanorobotic arms to manipulate individual cells or propelling the nanomachine on its journey to locate its target. Once the target is reached, the chemical payload can be released (most likely a highly specific drug), physical manipulations can be made, or diagnostics can be performed. Since the morphology of the DNA origami nanomachine can be designed from the ground up to be perfectly complementary to its intended purpose without needing to comply with the limitations of viral mimicry the possible applications are endless.

DNA origami may still be futuristic, although simplistic early examples have already been demonstrated, but along with viral robotics it already has profound implications for life science R&D. Historically, the delivery mechanism for a drug was an important, but secondary, consideration compared to the design of the drug itself. Tablets, capsules, inhalers, injections, suppositories and other traditional delivery routes were and still are the only regular way of getting the medicine to where it is needed, and the only assessment required was which of these methods would get sufficient active component to its intended target. With viral robots and nanomachines, the design of the virus or machine is inextricably interlinked with the design of the chemical compound that it will be delivering, since the structure of that compound will affect the biological and chemical characteristics of the robot required to deliver them.

Certain generic categories or classifications for viral robots will doubtless emerge, for instance a general-purpose virus capable of delivering any compound that belongs to a class defined by its physical structure, but the detailed design of the delivery mechanism will have to take place in parallel with and on an equal footing to research into the compounds themselves. It may not be too far-fetched to suggest that the basic underlying scientific discipline driving future life science R&D will be biomedical engineering, not biochemistry alone, just as modern research lab design is more about ergonomics, process engineering and workflow optimisation, rather than simply allocating bench space.

Modern research labs usually feature benchtop robots, if not larger standalone machines and in some cases completely automated rooms or labs-in-containers that are optimally engineered for a single process or workflow. These industrial-scale robots have gained widespread acceptance and are in everyday use, particularly in fields such as high throughput screening (HTS) where many thousands if not millions of compounds have to be analysed in a short period of time to find an optimal candidate. Machines are tasked with setting up plates, creating samples, running standardised tests against those samples, automating a wide array of ingredient selection and formulation, assessing



their likely effectiveness and toxicity and making decisions regarding which experiments to do next based on predetermined logic provided by the supervising scientist. Such is the level of automation possible today that commercial services already exist that allow a selection of experiments to be commissioned online, carried out in a fully automated lab-in-a-box and results uploaded online, all without the intervention of a single human being.

An increasing level of automation in laboratories can significantly reduce the cost of science. As well as the obvious benefits for larger companies running their own labs, it also lowers barriers to access for smaller players, just as cloud computing has enabled even the smallest of companies to work with massively powerful computational resources for an affordable fee. A growing number of experiment types can now be commissioned as low-cost services from research partners or commercial third parties. The reduced cost is a result of the robot's ability to perform experiments at scale, delivering time savings, consistency, reproducibility and operational efficiency both in terms of reagent usage and scientist time. Since robots are able to work 24/7, the health and safety of the scientists is improved as well through reduction of overtime and night-time working. Further, the automation of sample and compound storage and retrieval can itself improve the reliability and integrity of the materials being handled, helping to avoid experimental failures due to degraded material and the costs associated with restocking or replacing from source.

Even with the increasing level of lab automation there has been no recent significant advance in the underlying technology. Robots have become more compact, more efficient and more communicative with each other, but the basic concepts have largely remained the same in terms of equipment that is already available in the lab today. One of the most transformative advances that has occurred has been the development of acoustic dispensing systems which have revolutionised the ability to move liquids around plates. Detection technologies have also evolved, as have microfluidics, the speed and capacity of high throughput assay systems, the ability to create cellular organoids on chips and even 3D print cells, and complete analysis platforms, but none have yet created a widespread step change in the way science is done.

The major exception is nanorobotics. The development of nanomaterials and nanodevice design has made huge leaps forward and has enabled the field of living technology to be explored in much greater depth. Without advances in ultra-miniaturisation it would simply not be possible to construct this new class of robots, but the development of these and their viral robot relatives has depended just as much on equivalent advances in the understanding of genomics and DNA technology.

The human genome project opened the doors to low-cost sequencing which has now been used to sequence many different genomes. The resulting bioinformatics tools and genetic databases have given us a vast library of resources that are structured, annotated and in some cases validated, providing us with a solid foundation in the general principles of genes and DNA. This foundation has informed and driven the field of synthetic biology, producing software tools such as plasmid editors and sophisticated genome editors, modelling, simulation and visualisation technologies. Most importantly, the DNA synthesiser was developed with its ability to act as a software compiler, a bridge between the world of electron computing and the construction of the DNA instructions that run living molecular systems. Synthetic biology[2] thus drives the ability to construct entire genomes from scratch, making the viral robots and DNA origami nanomachines of the future entirely possible using technology that is already available today.

In the years to come DNA technology is set to advance even further. There is a need to accelerate the development of infrastructure and tools necessary to design and build genomes cost-effectively at scale, exemplified by the proposal of a new Human Genome Project-Write (HGP-write) that aims to synthesise entire genomes – including human genomes – from their individual chemical components and get them to function in living cells. The foundational tools to do this kind of biological

engineering, and improve the technology underpinning viral robots and DNA nanomachines as a result, should see a rapid evolution in the short- to mid-term along with the emergence of small technology-savvy groups bringing novel applications to market. It is likely that these first-to-market applications will initially appear in the field of cancer research, bringing together personalised treatments, detection and perhaps even genetic surgery, all combined effectively to make cancer a manageable illness.

DNA is not new technology, and it is not untested. DNA is a key component and driver of biology and it works very well. The changes we will see are not in the technology but in the tools we use to work with it. From the viewpoint of DNA as a programming language for biology, and with the development of tools that can translate high-level logical constructs into low-level physical synthesis of DNA sequences, we will eventually see a whole new generation of software developers emerge that are capable, have the tools, are empowered, and have the vision to engineer this new wave of biological robots and living systems. A whole new branch of software engineering will be established and democratised, not just restricted to the major players in life science R&D, enabling any individual with an innovative new DNA-based idea to rapidly prototype and publish it in the same way that hobby programmers bring us an endless stream of new apps and websites today.

This vision could be utopian, or dystopian, depending on your point of view, but it is very much long-term. Shorter-term we are more likely to see existing life science robotics refined to be able to operate safely and reliably across the healthcare spectrum, from research through to the clinic. Treatment robots, already found in some operating theatres for minimally invasive surgery, will become more common in a number of other clinical settings, in some cases assisting with procedures and in others carrying out the entire operation alone. These applications are not too distant from the tasks carried out by lab robots today, although with a living subject undergoing robotic treatment the level of safety, reliability and predictability naturally becomes of huge importance.

Lab robots will not be left behind in this march towards therapeutic progress. Scientists can already command robots to carry out chemistry in µl volumes in a variety of plate formats, test the chemistry immediately in a biological assay, feed data into a learning algorithm, ask it to predict which compounds to make next, then feed the results back in a complete loop to make the next compound and repeat the cycle. However, these stages, while automated, are still largely separated and require a scientist to move compounds between each machine at each stage. An increased ability to fully integrate these machines through software standards and communications protocols and allowing them to self-improve through development of machine learning techniques will be the biggest step forward we are likely to see in the mid-term, and is where much current research is focused.

For now, and until better methods of integration remove the need for a human interface between each stage, lab robotics continue to improve mostly along the lines of speed and efficiency. In labs currently under construction at AstraZeneca's new site in Cambridge, UK, acoustic dispensing from 75µl acoustic tubes will miniaturise compound storage by as much as 90%, and improve lab efficiency and reduce experimental contamination by allowing samples to be placed directly on an assay plate without manipulation or physical contact. Similar application of acoustic dispensing to mass spectrometry, where samples can be dispensed directly into the machine, will see throughput increase from an already impressive three samples per minute to an incredible six samples per second. Collaborative robotic arms will work beside researchers on their own benches in an open lab environment using advanced motion detection sensors to avoid collisions with humans. Additionally, freezer storage will incorporate robotics to locate, extract and deliver samples from freezers in an environment maintained at the same temperature as the freezer itself, removing the temperature fluctuation and related sample degradation that occurs whenever humans, who require normal room temperature to operate, open the lid of a freezer to search for and remove a sample themselves. These robots can also find the samples faster than humans through the use of barcodes and databases that map everything to precise physical locations within each freezer. Such cold storage systems are already on the market at up to -20˚C, but will soon be available at -80˚C and even -140˚C in the very near future.

The key to the success of these advanced systems is the human factor. To make the best use of a new system its operators must be fully trained and supported, and approach its application from the viewpoint of the biology and not the technology. Just as in other fields, technology for its own sake is rarely helpful, but a solution purchased and designed for a specific purpose can deliver real value. Adapting the biology to play to the strengths of the automation is the only successful approach, and removing

References

1 http://dx.doi.org/10.1038/464158a.

2 http://dx.doi.org/10.1073/pnas.1508521112.

the variety in the biology gives a greater return on investment by reducing the time spent customising and reprogramming between each experiment.

To make the most of the technology available to them and to help progress it, scientists will need to keep abreast of the developments in robotics and be aware of the likely impact from deploying even the smallest machines in their own labs. Inter-company collaboration to share ideas and experiences, and internal collaboration between research groups will be essential in order to ensure that all staff within an organisation are aware of what is available internally in their labs and externally on the market. Procedures that are habitually carried out by hand a few at a time can be scaled up to several thousand a day with only the simplest of robotic solutions if the investigator is able to identify the need for and seek out an appropriate piece of technology. Mindset and awareness will thus be the greatest influencing factor over the adoption of robotics in labs for at least the short-term, rather than anything to do with cost, installation requirements or technological capability of the machines themselves.

In the clinic, as with DNA engineering, the human factor plays a very different role in the expansion of robotic therapeutics. Patients may show resistance to the idea of robotic surgery or therapies, and, as we begin to merge robotics with advancements in precision medicine and translational medicine, ethics will also come into play. What does it mean for the individual and society when significant amounts of gene therapy and genetic modification become possible, even addressing prenatal or neonatal considerations that target maladies before they occur in the foetus? Society needs answers to the ethical questions raised by the types of treatment which become possible, as well as to consider who will bear the likely greater cost of developing and administering them.

Ultimately, as with all major technological advances, the ethical and cost concerns will eventually be addressed and the technology accepted in the minds of the public, and it is unlikely that any single issue will prevent robotic medicine from becoming mainstream. In our lifetime we have seen similar advances in information and communications technology that were resisted at first but eventually became mainstream, such as cloud computing and augmented reality, and there are also parallels with the industrial revolution and early advances in the mechanisation of agriculture. In computing we could not now stop the development of computing devices if we tried, and biotech is even more compelling as it has already proliferated, is acceptable, inexpensive, sustainable and has a single universal programming language – DNA – that we are becoming increasingly adept with.

As with the lab robots, agreeing on standards for communications protocols and physical integration would be a huge help in co-ordinating and optimising the development of DNA robotics. What is required now is to educate the next generation of scientists, and patients, so that the world becomes more biologically literate, just as we aim to make our populations computer-literate today. Not much is required to understand the fundamentals of biological technology and its building blocks, and as with computing we frequently find that the emerging masters of this technology are the younger generation with their extreme expertise in highly specific fields.

The Pistoia Alliance discussed the definition of robotics and the opportunities and challenges around this subject at its London conference because its member organisations see this is an area in which there is true potential to make a significant impact. To realise that potential as an industry, working across all therapeutic areas, we must support the creation of the investment, regulatory and academic systems that will recognise and nurture the younger generation of scientists in order to help accelerate the development of life science robotics. The Pistoia Alliance is uniquely well positioned to lead the development of the open, shared software, protocols and standards that will be needed. Since the challenges faced will be the same for all, then only through collaboration will it be possible to make optimal progress. We all need to work together to make sure that our robotic future comes to us in the best way possible, because come it will regardless.

This article is based on interviews with Andrew Hessel, Distinguished Research Scientist at Autodesk; Ted Pawela, Product Management Executive at Dassault Systèmes BIOVIA; and Steve Rees, Vice-President of Screening Sciences and Sample Management at AstraZeneca. Additional editing was carried out by Steve Arlington, President of the Pistoia Alliance. The author would like to thank all four for their extensive assistance. DDW

Richard Holland *is an independent consultant writing on behalf of the Pistoia Alliance. Richard was co-founder of Eagle Genomics, a Cambridge-based bioinformatics startup. He has held positions at major academic research institutes including the European Bioinformatics Institute (EMBL-EBI), with more than 15 years' experience in life science R&D IT.*