

Accelerating towards the future of drug discovery

Using AI-led design and automated synthesis to speed up discovery of potential new leads

How can you accelerate the development of therapeutic drug candidates while also reducing their cost? The global AI-driven precision medicine company Exscientia, headquartered in Oxford, UK, thinks it has the answer – and big pharma is paying attention. By integrating AI and machine learning in its discovery process, Exscientia has been able to generate new molecules far faster than is possible with conventional methods, in an approach that has already delivered compounds into clinical development. We talked to Exscientia's scientists to understand the company's approach, and how they've recently collaborated with Shimadzu to further automate the discovery process with an aim to synthesise and purify new drug candidates in super-fast time.

Changing the status quo in drug discovery

Conventional drug development is a slow, high risk process, resulting in few new drugs reaching the market, at an expensive cost. For example, traditional drug discovery commonly requires between 2500 to 5000 molecules to be synthesised in order to generate just a single compound suitable for clinical development, in a process that often takes five years or more.

However, Exscientia is on a mission to transform how the biopharma industry designs and develops medical therapies, by using an AI-led and increasingly automated approach to make drug discovery much more efficient and productive. The company wants to “turn drugs from a sparse resource into an abundant resource”, and so ultimately help patients live better and healthier lives.

Founded in 2012, Exscientia has already made waves with its transformative approach, with the first AI-designed drug candidate to enter clinical trials announced in early 2020, developed for Sumitomo Pharma. Following this initial groundbreaking success and five other compounds that have since entered into clinical development, the company is typically able to deliver development candidates in around 12 to 15 months, with 150 to 250 molecules being synthesised in the process – a marked contrast with traditional industry averages.



Part of Exscientia's new building at Milton Park, south of Oxford, showing its state-of-the-art facility for robot-enabled synthesis and testing of pharmaceutical drug leads.

As well as upending the status quo in drug discovery, the company itself has also undergone a dramatic transformation, employing a substantial number of tech specialists who work hand-in-hand with savvy 'drug hunters' to integrate the best science with the most advanced technologies. Some of those staff are based at Milton Park to the south of Oxford, UK, where the company has recently completed a state-of-the-art automation facility (about which more later).

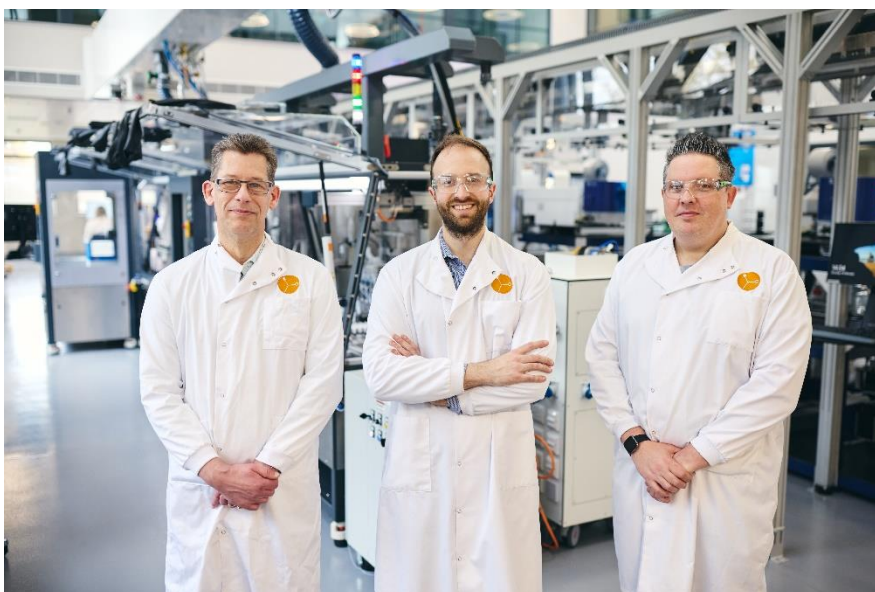
Encoding information and the role of AI

But first, let's examine why Exscientia's ideas have become such a compelling proposition. To get an insight, we first talked to Dr Paul Cox, Director of Chemistry Automation at Exscientia, who knows first-hand the pitfalls of conventional drug discovery. He explains what makes Exscientia different.

"It all started with our founders' deep frustration with how long it takes to discover new medicines, and the cost of getting them to patients. Of course, they understood very well that many processes in the drug discovery pipeline centre on handling complex sets of information – for example, about the binding affinities of molecular structures or the yields of synthetic processes". But one of the

difficulties, he says, has always been deciding what data needs to be looked at: “The sheer quantity and variety of data involved means that scientists spend a vast amount of time sifting through it to home in on the insights that may lead to new medicines”.

And even if you’re not concerned about timescales, he says, there’s another problem. “Conventional drug discovery processes may exhibit human bias. Consider the role of molecular structure, where medicinal chemists, drawing from their expertise, may lean towards a specific type of molecule for targeting a particular receptor. In their pursuit, they may synthesise many variants in search of the optimal candidate. However, what if this approach merely leads to a ‘local minimum’, overlooking a potentially superior alternative structure? Striking the right balance between exploring diverse options and leveraging existing data becomes crucial in this context”.



(Left to right) Craig White, Paul Cox and Andrew Ledgard in the lab at Exscientia.

The big insight at Exscientia, Cox explains, was that if you could encode all available information on biological targets and molecular structures, the analytical power of AI (or strictly speaking, machine learning) could be used to process it and come up with a small set of diverse molecules that most effectively test the widest range of hypotheses. As a result, to get one compound into clinical trials, you would only need to test a few hundred, rather than thousands or tens of thousands. And

importantly, because computer processing would be taking care of the time-consuming work, you'd also dispense with much of the tedious trial-and-error aspect of drug discovery.

But the team at Exscientia isn't just applying these methods to molecule design, it's looking at every aspect of the drug discovery and early development processes, and how they could be improved, says Cox.

In many tasks they apply AI to, Exscientia staff use a paradigm they call 'model-driven adaptive design', involving repeated rounds of feedback to refine the models. Cox says: "Of course, the AI models need ongoing improvement, and by feeding back all the test data and every single human decision, we're constantly making them better. So for compound discovery, we'll typically go through about 10 to 15 'design cycles' each involving perhaps 25 compounds, to evolve the compound set and answer a project's hypothesis".

Streamlining synthesis with robotics

Although Exscientia has attracted most attention for its use of AI, the company is also pushing forward on another labour- and time-saving innovation: laboratory automation. Cox explains the thinking behind this: "Traditional drug discovery is not just held back by the speed of information processing, hypothesis testing and decision-making, but also by the practical aspects of making, purifying and testing drug candidates. So we want to address that, too".

But once again, conventional thinking needs to be re-assessed. He explains that automation in chemistry has typically focused on 'parallel synthesis' – basically, making lots of rather similar molecules. But this often isn't the most effective way of testing hypotheses. That's why Exscientia is increasingly looking at more bespoke workflows that enable a wider range of molecules to be synthesised. To move towards this goal, the team designed and commissioned a series of automated

platforms geared towards enabling the synthesis of a more diverse range of chemical structures through a wider array of reaction classes and workflows.

But true automation of these syntheses will require more than just hardware, says Cox. This is because reaction protocols, whether generated in-house or taken from the literature, must also be encoded to turn them into automation-friendly procedures that can be reliably executed. However, experimental descriptions can vary substantially between chemists. Therefore, a big challenge is standardising this information to make it machine-interpretable. That's why, he adds, the company is developing its own reaction datasets in-house – “we absolutely have to ensure that the data we feed our AI models is properly structured, unbiased and reliable”.

A new facility for a new era

This combined philosophy of integrating AI with automation – led by highly skilled human scientists and tech experts – is coming to life in Exscientia's new facility at Milton Park, UK. The building was an empty shell when acquired in 2021, but has since been rapidly kitted out to house the company's new automation equipment.

And nobody knows this end of the company's drug pipeline better than Andrew Ledgard, Senior Automation Chemist at Exscientia. He describes the basic workflow: “Our AI-driven process will come up with a target molecule and a number of synthetic routes to it. It's then our job to select the ones that are most likely to work, set up the dosing sequences ready for automated synthesis, and ultimately ensure that we get a sample of pure product ready for testing”.

This all needs to happen at speed, he says: “For each step in the synthesis, we're looking to replicate what a super-efficient lab chemist would do – basically run a reaction overnight, then work up, purify and analyse the next day, ready for the next reaction”. He says the team is working at moderate to

small scales: "For a multi-step synthesis with unoptimised yields, we might be starting at the gram scale, in order to end up with a few milligrams for testing".

Versatile instruments for novel compounds

Ledgard's colleague Craig White, who leads Analysis and Purification, describes the four 'analytical touchpoints' across the whole process. These involve the application of Shimadzu analytical and preparative liquid chromatography (LC) and supercritical fluid chromatography (SFC) coupled with mass spectrometry (MS).

"The first stage is reaction monitoring, for which we use an analytical LC-MS configured with a flow-through vial. The next stage is solid-phase extraction (SPE), followed by purification using either preparative LC or SFC. And finally we need to perform quality control prior to biological screening", he explains.

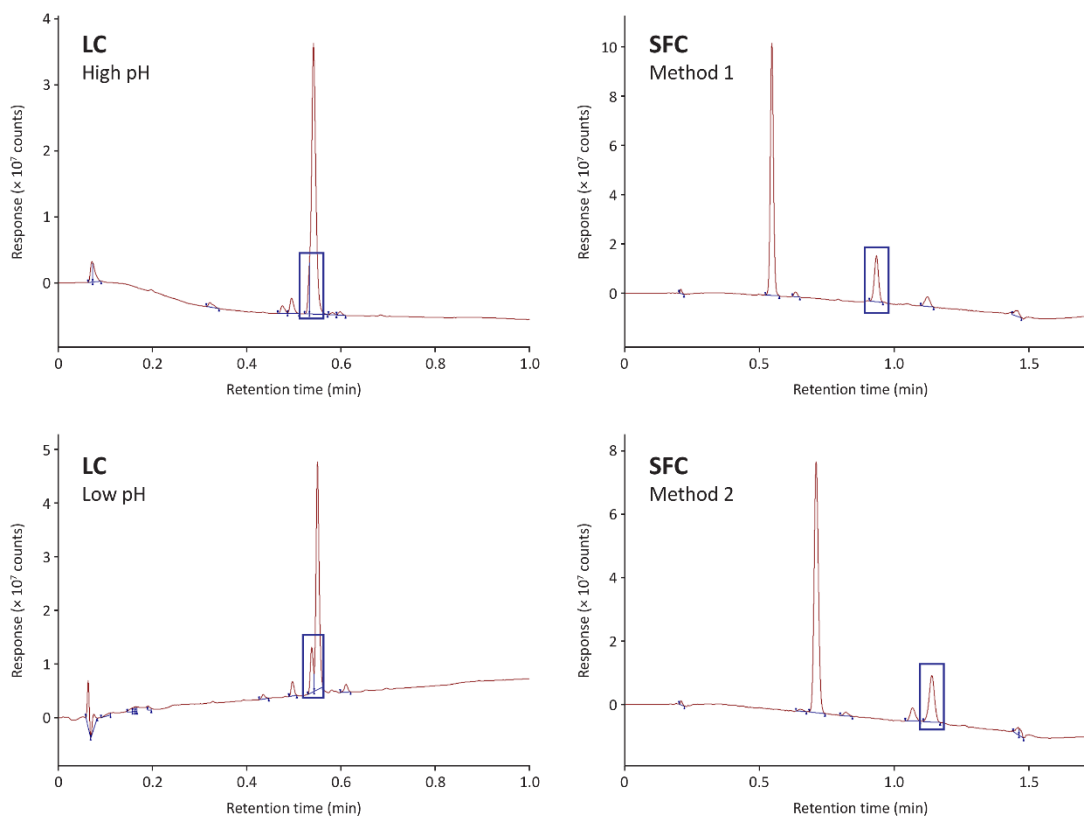
The entire analytical landscape is completely different from what would be encountered in conventional drug discovery, says White, because of the need to stream intermediates and final compounds with diverse polarity, across different automated platforms. It was therefore important to consider LC and SFC 'side by side' as complementary techniques, to successfully isolate pure material.

"And when we're dealing with multi-sample and multi-gram purifications overnight, there are big implications for solvent handling", he adds. "Specifically, we



Some of the Shimadzu equipment used at Exscientia for reaction monitoring: (top) the analytical LC-MS, and (bottom) the flow-through vial configuration, which physically connects to the chemistry platform.

needed to design novel solutions for solvent delivery and solvent waste removal – which are all now operational and integrated with the Shimadzu instruments”.



Example of exploratory purification work carried out at Exscientia on an impure sample, showing (left) two LC separations at different pH, and (right) two SFC separations using different methods. As can be seen, both SFC methods easily separate a significant impurity that is only partially resolved or completely unresolved using LC.

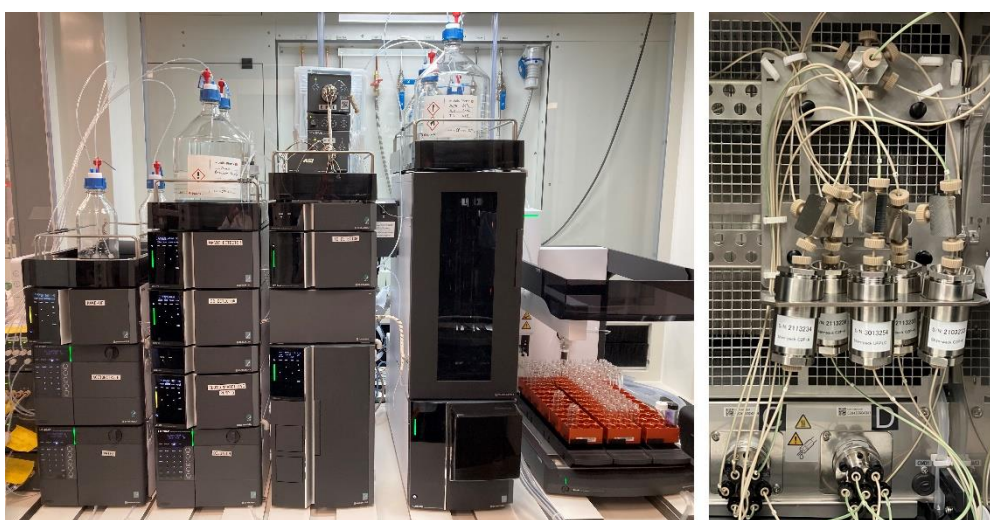
Innovation, collaboration, integration

Ledgard and White have been closely involved in system design at Exscientia’s automation lab throughout its development, and they’re not exaggerating when they say that these are just a small sample of the challenges they’ve faced.

But through it all, Shimadzu’s help with the analytical setup has been invaluable, says Ledgard: “We initially contacted Shimadzu in late 2021 because we were interested in the cutting-edge fraction-trapping capability on their ultra-fast preparative LC. This was vital because it enabled us to eliminate

the lengthy dry-down that is otherwise unavoidable when isolating a target compound from an aqueous mobile phase”.

But then they realised that other considerations also made Shimadzu a strong match. For instance, it offered both the LC and SFC techniques, simplifying the process of creating the integrated platform that Exscientia was looking to design. “At that time not many vendors offered that combination of assets”, White says.



The Shimadzu ultra-fast preparative LC system used at Exscientia, showing (left) the complete setup and (right) the fraction-trapping capability that was vital for avoiding lengthy dry-down procedures.

Another factor, he explains, was reliability: “Any 24/7 operation will need a robust platform, so the reported reliability of Shimadzu equipment and software was a big draw for us. It’s also been great to integrate their LabSolutions Sync software into our setup, as it provides flexibility to orchestrate sample submissions outside the environment of the Shimadzu chromatography data system, using our in-house informatics platform”.

And finally, there’s the collaborative mindset, he says. “Our project here in Oxfordshire has never been about simply optimising individual instruments. We’ve been on a mission to put together a completely novel, integrated setup, so we needed bespoke configurations and bespoke software.

Therefore, it was really important that we had an instrument vendor who really understood the need for collaboration, and were prepared to innovate quickly to provide solutions”.

A new vision for drug discovery

Although the new facility will need time to ramp up to peak performance, a lot of progress has been achieved in a short period. “From the analytical and purification angle, we still have a lot to do”, says White, “especially in automating the movement of physical samples and the flow of information. However we installed the complete lab infrastructure in just two years, and as part of that timeframe, it took only six months to get our analytical and purification platforms operational. We also have an internal team of engineers and software developers working on automating the input/output of purification tools, to further streamline the entire workflow”.

The joined-up approach to compound preparation, says Ledgard, is becoming more important, because vendors are beginning to realise that their customers don’t just want standalone instruments. “They want to be able to connect compound workup to purification, and purification to evaporation. Developing solutions for connectivity is mostly our responsibility currently, but through collaboration with vendors, I think we’ll see that become more mainstream in the near future”.

That sentiment – automating and linking up the process so it runs seamlessly from start to finish – fits neatly into the wider picture of drug discovery at Exscientia, concludes Paul Cox. “Our ultimate aim is to create a drug discovery pipeline that’s human-led in terms of overall strategy, target setting and other key decision making, yet could operate more-or-less without human interference from concept to reaction execution to delivery of the product for testing – and in a matter of weeks rather than years. It kind of blows my mind when I think that this might be possible, let alone that Exscientia seems well on track to achieving it!”