Tapping into the drug discovery potential of AI

Plentiful financing and multiple pharma partnerships illustrate the burgeoning interest in applying artificial intelligence tools to drug research and development.

Credit: Westend61

Neil Savage

In April of this year, the German biotechnology company Evotec announced a phase 1 clinical trial on a new anticancer molecule. The candidate was created in partnership with Exscientia, a company based in Oxford, UK, that applies artificial intelligence (AI) techniques to small-molecule drug discovery.

Where it might have taken the traditional discovery process 4–5 years to come up with the drug candidate—an A2 receptor antagonist designed to help T cells fight solid tumors—it was found in 8 months by harnessing Exscientia's 'Centaur Chemist' AI design platform. This system can computationally sort through and compare various properties of millions of potential small molecules, looking for 10 or 20 to synthesize, test and optimize in lab experiments before selecting the eventual drug candidate for clinical trials.

Within 3 weeks, Exscientia, a 2012 spinoff of the University of Dundee, UK, announced a $225 million Series D financing round (Table 1), along with a $300 million equity commitment that it can draw on at its discretion. In 2020, the company did much the same, announcing a drug candidate with Sumitomo Dainippon Pharma in Osaka, Japan, then later raising $100 million in Series C funding. That drug, a selective serotonin reuptake inhibitor (SSRI) designed to treat obsessive compulsive disorder (OCD), and the oncology drug are the first two molecules designed with the help of AI to enter clinical trials, Exscientia claims. The company has also formed drug discovery partnerships with Bristol Myers Squibb (BMS), Sanofi, Bayer, GlaxoSmithKline, Roche and the University of Oxford (Table 2), and is building its own pipeline.

Exscientia is just one of many companies founded in the past decade around AI-based strategies for drug discovery and development, several of which have raised substantial funding recently (Table 1). Some of these are also developing tools to accelerate the identification of small-molecule drug candidates. Others such as Recursion Pharmaceuticals, which recently raised $436 million in

Table 1 | Selected recent financings of companies applying AI in drug discovery

Company	Date	Headline
Schrödinger	February 2020	Drug discovery software company closes $232 million IPO backed by Bill Gates and David Shaw.
Insitro	May 2020	Insitro raises $143 million in Series B funding, to help drive its machine learning-based drug discovery approaches further.
AbCellera	May 2020	AbCellera raises $105 million in Series B funding round to expand its antibody drug discovery platform.
Relay Therapeutics	July 2020	Relay Therapeutics, which focuses on understanding protein motion to design drug candidates, closes $400 million IPO.
Atomwise	August 2020	Sanabil Investments co-leads $123 million Series B funding round for Atomwise to support the development of its molecule identification software.
Recursion Pharmaceuticals	September 2020	Recursion Pharmaceuticals, which is applying machine learning to cellular imaging data, raises $239 million in Series D financing round led by Bayer's investment department Leaps. Other investors include Casdin Capital, Samsara BioCapital, Baillie Gifford and Lux Capital.
XtalPi	September 2020	More than a dozen investment companies raise $318 million in Series C round for start-up XtalPi, which is applying quantum physics with AI to discover drug candidates.
AbCellera	December 2020	AbCellera closes its IPO at $556 million.
Cellarity	February 2021	Cellarity raises $123 million in Series B funding for its drug discovery approach based on modulating cellular behaviors.
Valo Health	March 2021	Valo Health, which is developing its Opal computational drug discovery and development platform, raises $110 million to add to its $190 million raised in January 2021 for its Series B funding round.
Insitro	March 2021	Insitro raises $400 million in Series C financing led by Canada Pension Plan Investment Board.
Exscientia	March 2021	Exscientia completes $100 million Series C financing, with investors including Evotec, Bristol Myers Squibb and GT Healthcare.
Recursion Pharmaceuticals	April 2021	Recursion completes $436 million IPO.
Exscientia	April 2021	Exscientia secures additional $225 million in a series D round led by SoftBank Vision Fund 2.

IPO, initial public offering.

Table 2 | Selected recent deals involving companies with a focus on applying AI in drug discovery

Companies	Date	Headline
BenevolentAI, AstraZeneca	April 2019	AstraZeneca partners with BenevolentAI to identify new drug candidates for chronic kidney disease and idiopathic pulmonary fibrosis.
Insitro, Gilead	April 2019	Gilead becomes first big pharma company to sign 3-year deal with Insitro to apply its Insitro Human platform to identify new drug targets for non-alcoholic steatohepatitis by creating experimental models of the disease. Insitro will be paid $15 million upfront in a deal potentially worth $1 billion.
Exscientia, Rallybio	July 2019	Exscientia brings its AI drug discovery platform Centaur Chemist to a partnership with Rallybio to discover small-molecule drugs for rare diseases.
Microsoft, Novartis	October 2019	Novartis announces strategic alliance with Microsoft to apply their AI algorithms to its large datasets to identify and develop therapeutics.
ZebiAI Therapeutics, Google Accelerated Science	January 2020	ZebiAI Therapeutics collaborates with Google Accelerated Science to discover small-molecule drug candidates using its machine learning and DNA-encoded library technologies.
Exscientia, Bayer	January 2020	Bayer collaborates with Exscientia to discover cardiovascular and oncology drug candidates using their artificial intelligence drug discovery platform Centaur Chemist.
BioSymetrics, Sema4 Johnson & Johnson	August 2020	J&J's Janssen partners with BioSymetrics and Sema4 in a collaboration that aims to apply machine learning to predict onset and severity of COVID-19 among different populations, with a goal of developing new treatments and vaccines.
Recursion Pharmaceuticals, Bayer	September 2020	In addition to its Series D financing round, Bayer partners with Recursion Pharmaceuticals in drug development deal for new small-molecule therapies to treat fibrotic diseases using Recursion's AI-guided drug discovery platform.
Insitro, Bristol Myers Squibb	October 2020	Insitro will use its machine-learning technology Insitro Human platform to identify potential drug targets by developing predictive models of amyotrophic lateral sclerosis and frontotemporal dementia. Bristol Myers Squibb will then choose candidates to develop further.
Genesis Therapeutics, Genentech	October 2020	Genesis Therapeutics partners with Genentech in multi-target drug development deal using Genesis' graph machine learning capabilities to identify drug candidates for a range of disorders.
Roivant, Silicon Therapeutics	February 2021	Roivant buys Silicon for $450 million, including its physics-based platform for in silico small-molecule drug design, to be integrated with Roivant's machine learning approaches.
Exscientia, University of Oxford	February 2021	Exscientia and the University of Oxford collaborate to develop treatments for Alzheimer disease.
Iktos, Pfizer	March 2021	Iktos will apply its AI-driven de novo design software to a number of Pfizer's small-molecule programs.
ZebiAI Therapeutics, Relay Therapeutics	April 2021	Relay buys ZebiAI for $85 million upfront and a further $185 million in potential milestone payments.

its initial public offering, are generating vast amounts of bespoke data on cellular behavior in the hope that these can be mined using AI to reveal biological insights that could inform the discovery of innovative drugs. More established, less health-care-focused tech companies, such as IBM, Microsoft and Google, are getting involved too.

Ultimately, the diverse range of AI applications being explored could help tackle the fundamental challenge that developing new drugs, from target identification through clinical trials, requires years of time and billions of dollars. Recognizing this potential, and hoping that the new technology can also help them develop more-effective and better-targeted drugs, pharmaceutical companies are building up their own in-house AI teams, as well as investing in and collaborating with these companies (Table 2).

Matching strengths

Partnering between pharma companies and AI companies is "definitely blossoming across the industry," said Jim Weatherall, VP of data science and AI R&D at AstraZeneca. In some cases, "we would look to them and say, you've spent the last few years developing all these algorithms and tools so that we don't have to. We'll partner with you because we understand our science, and share the risk that way," he said. Though there hasn't been much merger and acquisition activity in the area, Weatherall says that in some cases pharma companies are making a 'partial purchase' and gaining a seat on the board of an AI company to help steer its direction. "It's not necessarily that they then come up with some killer algorithm that we could not," he said of the results of such partnerships. "It's usually something more like they've hooked into a specific data set we don't have access to or they've created a nice tool, which is a nice software wrapper around what they've done."

As an example, Weatherall points to a collaboration launched in 2019 between AstraZeneca and BenevolentAI, to discover new drugs for chronic kidney disease and idiopathic pulmonary fibrosis, a serious lung disease. The pharma company was interested in using knowledge graphs, which pull together large amounts of data from disparate sources—some of it mined from the literature using natural language processing, some produced by chemical or pharmacological experiments, some arising from gene studies. A knowledge graph allows a computer to find connections among the various data types and present them graphically, producing insights no single set of experiments could lead to.

AstraZeneca decided BenevolentAI was the best company to turn to for knowledge graph work. In January 2021, the companies announced a new drug target for chronic kidney disease. BenevolentAI's computers predicted the target, and AstraZeneca's experiments validated it.

Deals can take a variety of forms, said Najat Khan, chief data science officer at Janssen, Johnson & Johnson's pharmaceutical arm. "We really like to invest in the external ecosystem, whether collaborations or equity," she said. "We use many different mechanisms to be part of that ecosystem." With some companies, the pharma's venture capital arm, Johnson & Johnson Innovation, makes an equity investment and takes a seat on the board. The group also makes investments through JLabs, its startup incubator, to collaborate with young companies. One such case is Insilico Medicine, of Hong Kong, which in 2019 joined JLabs and in November 2020 launched a collaboration with Janssen to design small-molecule candidates for several drug targets chosen by Janssen.

AstraZeneca is also working with Schrödinger, a company which has used physics-based computational approaches to predict properties of potential small-molecule drug candidates such as their binding affinity for protein targets since it was founded in 1990, but in the past decade has added more machine learning into its toolkit.

Driven by data

Computational approaches to small-molecule drug design go back at least to the 1970s, when efforts focused on understanding relationships between the structural characteristics of molecules and their biological activities to help suggest new candidates. Today, various computational approaches are routinely used to screen virtual libraries of large numbers of existing compounds looking for those that might match a newly discovered target to test in experimental assays. And in fact, small-molecule discovery is the focus of many companies applying AI to pharmaceutical questions.

Small molecules are attractive for AI approaches in part owing to the availability of appropriate data to learn from, thereby enabling good predictions about new molecules to be made. Small molecules are well described by their chemical structure, which can be rendered easily in a format that can be used by computers. And scientists have a good understanding of the physicochemical principles that underlie the behavior of small molecules, such as their interactions with protein targets, based on huge amounts of high-quality data that have been amassed in public and industry databases in the past few decades.

AI, with its ability to look at vast quantities of existing data and learn patterns that might be too subtle or complex for humans to recognize, can then predict new small molecules with desirable properties, taking the computational screening process to a new level. For instance, Exscientia is working with companies such as Dainippon Sumitomo on finding bispecific molecules, which can bind to more than one target. The company signed a deal with Sanofi in 2017, which was worth up to €250 million, to search for such a drug for metabolic diseases, and, in 2019, Sanofi exercised its option to license a molecule that binds to two distinct pathways related to inflammation and the progression of fibrosis. Andrew Hopkins, CEO at Exscientia, said AI provided the ability to search a much vaster chemical space than traditional processes could hope to handle. "You would never find these molecules by large-scale, high-throughput screening. It's not looking for a needle in a haystack. It's looking for a needle on the farm," he said. "I think you're going to open up now other possibilities where there'll be new types of small molecule starting to be discovered, which can only be done by the kind of scale and search space that an AI algorithm can do, compared to a traditional sort of brute force screening method."

Iktos, meanwhile, announced in March that it was applying its technology to a number of small-molecule discovery programs at Pfizer. It also has collaborations in place with Merck KGaA, Janssen and Almirall, which is a Barcelona-based company working on treatments for skin diseases. Often, pharmaceutical companies have identified potentially useful compounds, then they turn to Iktos to learn the parameters of those compounds and find more candidates. "They basically see the AI-designed compounds as additional ideas on top of the original ones that they have," said Yann Gaston-Mathé, president and CEO of Iktos. "We really feed them with new ideas and then they are in the driver's seat deciding which compounds they will make." Such collaborations usually last about a year.

Iktos also licenses machine learning modules it has created to pharma companies, which can apply them with their data and in-house expertise. Other deals take the form of a joint technology development agreement, in which the pharma company helps fund the development of new algorithms.

Clinical trial design is another area where the power of AI is being brought to bear on various existing data sets, including electronic health records, patient demographics, the results of previous clinical trials, and information gleaned from omics fields. For instance, in March, Janssen inked a deal with Komodo Health, which combines stores of claims codes and other patient data with predictive AI algorithms to match patients to clinical trials, potentially making those trials more statistically powerful. Similarly, Janssen partnered in November 2020 with Tempus, a company that find biomarkers through predictive algorithms, focused on oncology. In cases like these, the main driver for the collaboration is less about the algorithms than the data. "We partner with them for access to those data sets," Janssen's Kahn said.

For some applications though, there has been a surge of interest in approaches based around generating data specifically with AI applications in mind from the outset, as existing data may have limitations related to quantity, quality and suitability. For example, Insitro was founded in 2018 to rapidly generate high-quality biological data sets suitable for machine learning in drug discovery. Now it is using its technology to create predictive models of amyotrophic lateral sclerosis and frontotemporal dementia for target identification in a deal with BMS. That 5-year development deal, signed last October, provides a $50 million payment upfront and another $20 million for hitting certain milestones, with a total of up to $2 billion for future work, plus royalties on any drug sales. Once Insitro identifies targets, BMS will choose which of those to pursue (Table 2).

Similarly, Recursion uses an array of robots to treat millions of cell samples with drugs and genetic perturbations, stain them, and image them. It then applies machine learning algorithms to search for informative relationships between the perturbations and the morphological features of the cells. Along with hit screening, Recursion CEO Chris Gibson told *Nature Reviews Drug Discovery* that its creation of well-curated image data could also be useful across a wide array of problems in drug discovery, including target identification, target deconvolution, library enrichment, lead optimization and toxicity testing (*Nat. Rev. Drug Discov.* **18**, 653–655; 2019).

Bayer found this approach promising enough that it signed an agreement with Recursion in September 2020 to work on fibrotic diseases. The deal included a $30 million upfront payment, plus $100 million each for reaching milestones in up to ten drug discovery programs, making the deal potentially worth more than $1 billion. At the same time, Bayer's investment arm, Leaps, contributed $50 million to Recursion's $239 million Series D financing.

Big tech is getting into the game too

Novartis launched a multi-year strategic alliance in 2019 with Microsoft to apply the computer company's AI algorithms to pharma's large datasets. The companies said they planned to use image analysis and generative approaches to develop personalized medicine and optimize cell and gene therapy. Nvidia, maker of graphics processing units, which has been expanding its AI work, in April partnered with Schrödinger in the hopes of increasing the speed and accuracy of their molecule prediction software.

Those involved in this research say the role of AI in drug discovery and development is only likely to grow. "It opens up the prospect of having some kind software that could automatically design molecules 100% in silico, without intervention of the chemists," said Gaston-Mathé. "You would need to make only one molecule and test it and it would be the good one." He acknowledges that such a vision is unlikely to come completely true any time soon. "That's the science fiction objective."

Hopkins, meanwhile, predicts that by the end of the decade AI will have a part in the design of every new drug. "What we are seeing is that the competitive advantage and the capital efficiency of these approaches is so superior that actually they will win out as ways to discover drugs," he said.

Neil Savage is a science and technology writer in Lowell, Massachusetts.