on timescales that are, again, in excellent agreement with predictions from black-hole-scaling relations.

Many studies have explored whether there is a similarity between the accretion-state transitions in stellar-mass and supermassive black holes[9–12]. However, these studies have been plagued by uncertainties because of the extremely different timescales involved — days to months for transitions in stellar-mass black holes, compared with tens of thousands of years for those expected in active galactic nuclei. Kara and colleagues' detection of milli-second-timescale reverberation lags during an outburst from MAXI J1820 + 070 adds a valuable piece to the puzzle, but this event represents only a single data point. The authors' findings could be strengthened in two ways.

First, other stellar-mass black holes must be studied as they transition from outburst to quiescence, using NICER and future high-time-resolution X-ray instruments that have even better sensitivity. A compiled sample of data from these outbursts, more than one of which is associated with millisecond-timescale reverberation lags, would strengthen the current findings. Second, studies of highly variable active galactic nuclei, in particular those whose supermassive black holes transition rapidly from bright to faint states, or vice versa[13–16], would provide valuable tests of whether the structures of accretion disks are similar for stellar-mass and supermassive black holes. Together, these complementary studies could provide an unprecedented look at changes that take place in the immediate surroundings of all accreting black holes. ∎

Daryl Haggard *is in the Department of Physics and at the McGill Space Institute, McGill University, Montreal, Quebec H3A 2A7, Canada.*
e-mail: daryl.haggard@mcgill.ca

1. Kara, E. *et al. Nature* **565**, 198–201 (2019).
2. Shidatsu, M. *et al. Astrophys. J.* **868**, 54 (2018).
3. Tucker, M. A. *et al. Astrophys. J. Lett.* **867**, L9 (2018).
4. Uttley, P., Cackett, E. M., Fabian, A. C., Kara, E. & Wilkins, D. R. *Astron. Astrophys. Rev.* **22**, 72 (2014).
5. Kara, E. *et al. Mon. Not. R. Astron. Soc.* **462**, 511–531 (2016).
6. Uttley, P. *et al. Mon. Not. R. Astron. Soc. Lett.* **414**, L60–L64 (2011).
7. De Marco, B., Ponti, G., Muñoz-Darias, T. & Nandra, K. *Astrophys. J.* **814**, 50 (2015).
8. Gendreau, K. C. *et al. Proc. SPIE* **9905**, 99051H (2016).
9. Merloni, A., Heinz, S. & Di Matteo, T. *Mon. Not. R. Astron. Soc.* **345**, 1057–1076 (2003).
10. Falcke, H., Körding, E. & Markoff, S. *Astron. Astrophys.* **414**, 895–903 (2004).
11. McHardy, I. M., Koerding, E., Knigge, C., Uttley, P. & Fender, R. P. *Nature* **444**, 730–732 (2006).
12. Körding, E. G. *et al. Mon. Not. R. Astron. Soc.* **380**, 301–310 (2007).
13. Lamassa, S. M. *et al. Astrophys. J.* **800**, 144 (2015).
14. Ruan, J. J. *et al. Astrophys. J.* **826**, 188 (2016).
15. Gezari, S. *et al. Astrophys. J.* **835**, 144 (2017).
16. Noda, H. & Done, C. *Mon. Not. R. Astron. Soc.* **480**, 3898–3906 (2018).

COMPUTATIONAL BIOCHEMISTRY

Designer protein delivers signal of choice

A computational strategy has delivered a redesigned, more stable version of a cytokine protein that mimics the natural protein's interactions with receptors, opening the way for designer cytokine-based therapeutics. SEE ARTICLE P.186

E. YVONNE JONES

Messenger proteins called cytokines are secreted from cells and travel throughout the body. In the immune system, a cytokine called interleukin-2 (IL-2) delivers signals to receptors that regulate the activities of white blood cells. Cytokines such as IL-2 could therefore be used as therapeutics to modulate the strength of immune responses. However, the clinical uses of cytokines are often severely limited by properties that are inherent in their structures. On page 186, Silva *et al.*[1] describe how they have engineered a protein structure from scratch to replicate the beneficial receptor-binding properties of IL-2 without the drawbacks of the original cytokine.

IL-2 delivers its message by simultaneously binding to two receptor subunits known as IL-2 receptor β and IL-2 receptor γ (IL-2Rβ and IL-2Rγ). This binding pairs up the subunits to form a heterodimeric signalling protein called IL-2Rβγ$_c$. A third, non-signalling, receptor called IL-2Rα (also known as CD25) contributes to the formation of the signalling complex, strengthening the binding between IL-2 and IL-2Rβγ$_c$ roughly 100-fold. IL-2 is used to treat some people with cancer, in whom it stimulates the cells of the immune system to work more effectively to eliminate tumour cells[2]. However, IL-2 has toxic side effects that limit its use[3,4]. The mechanisms responsible for the toxicity are not fully understood, but studies[5] in mice indicate that the adverse effects are associated with the binding of CD25.

Several groups[6] have therefore sought to engineer a variant of IL-2 that no longer binds CD25 but still promotes the formation of IL-2Rβγ$_c$. The standard engineering approach is to start with the natural protein and find a combination of mutations that results in a variant protein with the desired binding properties. But this approach has foundered when applied to IL-2. This is partly because the unmutated protein is not very stable, and, as Silva *et al.* demonstrate, the mutated proteins are typically even less stable, which is problematic for the manufacture and storage of a therapeutic agent. In addition, the mutated cytokines have reduced potency
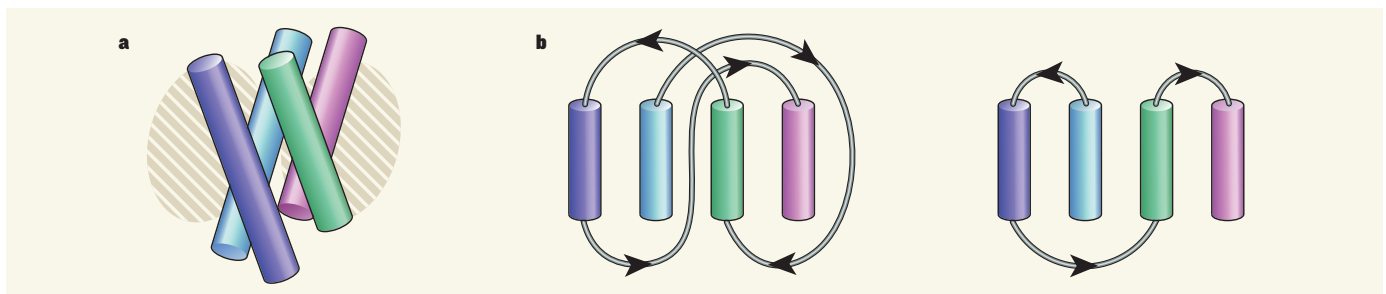

Figure 1 | Redesign for a cytokine protein. a, The cytokine protein IL-2 has potential therapeutic applications, but its clinical effectiveness is limited in part by its instability. The protein has a core of four α-helices (shown as cylinders) connected by loops (not shown), and exerts its biological effects by forming interactions with target receptors at two regions (shaded areas) on its surface. b, Silva *et al.*[1] used a computational strategy to redesign IL-2. They fixed the relative positions of the receptor-binding surfaces and then designed a stable four-helix bundle to present these surfaces to receptors. The designer protein is radically different in structure from the natural cytokine. For example, the connectivity of the four helices in the natural cytokine (left) is completely different from that of the analogous helices in the designer protein (right).

and exhibit residual binding to CD25.

Silva *et al.* therefore began afresh, and set out to design a protein structure from scratch that would provide a stable scaffold onto which they could add the structural elements required to produce the specific protein surfaces that bind to IL-2Rβ and IL-2Rγ. Crucially, these binding surfaces must be correctly positioned relative to each other in space to ensure that the designer cytokine engages the IL-2Rβγ$_c$ heterodimer and triggers signalling.

The authors obtained information about the structural and spatial requirements of their designer cytokine by analysing the crystal structures of naturally occurring cytokine–receptor complexes[7,8]. IL-2 is one of a large family of cytokines that have at their core a bundle of four structural elements termed α-helices (Fig. 1a). These four α-helices are linked, in a defined order, by a series of short or long connecting loops. Instead of keeping this particular arrangement of α-helices and re-engineering the binding surfaces, Silva *et al.* reversed the process. They started by defining the positions of the all-important binding surfaces, and then used computational methods to design an arrangement of α-helices that not only links these surfaces but is also predicted to be stable.

The proof of the pudding is in the eating, however. When the authors prepared and characterized the best candidates from the first round of design, the proteins showed promise in terms of IL-2Rβγ$_c$ binding, but had fairly poor thermal stability. Clearly, the recipe required some improvement. Silva *et al.* went back to the drawing board, taking the best arrangement of α-helices from the first round and substantially extending the computational search for optimal loops to link them together. This second round of design-generated candidates had improved stability and exhibited excellent binding to IL-2Rβγ$_c$.

Silva *et al.* then carried out an additional, experimentally driven round of mutagenesis — a fine-tuning process in which single amino-acid residues are changed — to enhance the binding properties of the best candidate proteins, and then fully characterized the cytokine that had the highest overall binding affinity for IL-2Rβγ$_c$. The results are impressive. The final designer cytokine is highly stable and binds strongly to IL-2Rβγ$_c$, but not at all to CD25. Excitingly, this new protein is effective as a therapy in mouse models of skin and colon cancer, delivering the immunotherapeutic effects characteristic of natural IL-2, but with lower toxicity. The authors named their designer protein Neoleukin-2/15 (Neo-2/15), because it is a new cytokine that mimics natural interleukins 2 and 15.

The researchers then determined the crystal structure of Neo-2/15 in complex with IL-2Rβγ$_c$. Gratifyingly, the binding surfaces are positioned as designed, and the four-helix bundle matches the computational blueprint with almost pinpoint accuracy. The redesign of the interleukin's four-helix bundle achieved

by Silva *et al.* is remarkably radical: the order in which the α-helices are linked has been rearranged (Fig. 1b), and the amino-acid sequence of the resulting 100-residue protein is very different from that of either mouse or human IL-2.

It remains to be seen whether Neo-2/15 will deliver on its initial promise in the clinic. Moreover, perhaps the four-helix bundle is a particularly favourable case for re-engineering — other cytokine families that have more-complex architectures might be harder to redesign. Nevertheless, Neo-2/15 excitingly demonstrates that bold *de novo* design, when combined with a deep knowledge of the structural determinants of receptor binding, can deliver designer cytokines that have bespoke binding properties. More broadly, Silva and colleagues' approach to protein design has the potential for re-engineering any of the myriad biological systems that involve interactions

between multiple proteins. In the meantime, the authors have opened up uncharted territory for therapeutics based on four-helix bundles, and there is plenty still to explore. ∎

E. Yvonne Jones is in the Division of Structural Biology, Wellcome Centre for Human Genetics, University of Oxford, Oxford OX3 7BN, UK.
e-mail: yvonne@strubi.ox.ac.uk

1. Silva, D.-A. *et al. Nature* **565**, 186–191 (2019).
2. Rosenberg, S. A. *J. Immunol.* **192**, 5451–5458 (2014).
3. Siegel, J. P. & Puri, R. K. *J. Clin. Oncol.* **9**, 694–704 (1991).
4. Sim, G. C. *et al. J. Clin. Invest.* **124**, 99–110 (2014).
5. Boyman, O. & Sprent, J. *Nature Rev. Immunol.* **12**, 180–190 (2012).
6. Kureshi, R., Bahri, M. & Spangler, J. B. *Curr. Opin. Chem. Eng.* **19**, 27–34 (2018).
7. Wang, X., Rickert, M. & Garcia, K. C. *Science* **310**, 1159–1163 (2005).
8. Stauber, D. J., Debler, E. W., Horton, P. A., Smith, K. A. & Wilson, I. A. *Proc. Natl Acad. Sci. USA* **103**, 2788–2793 (2006).

COMPUTER SCIENCE

Unprovability comes to machine learning

Scenarios have been discovered in which it is impossible to prove whether or not a machine–learning algorithm could solve a particular problem. This finding might have implications for both established and future learning algorithms.

LEV REYZIN

During the twentieth century, discoveries in mathematical logic revolutionized our understanding of the very foundations of mathematics. In 1931, the logician Kurt Gödel showed that, in any system of axioms that is expressive enough to model arithmetic, some true statements will be unprovable[1]. And in the following decades, it was demonstrated that the continuum hypothesis — which states that no set of distinct objects has a size larger than that of the integers but smaller than that of the real numbers — can be neither proved nor refuted using the standard axioms of mathematics[2–4]. Writing in *Nature Machine Intelligence*, Ben-David *et al.*[5] show that the field of machine learning, although seemingly distant from mathematical logic, shares this limitation. They identify a machine-learning problem whose fate depends on the continuum hypothesis, leaving its resolution forever beyond reach.

Machine learning is concerned with the design and analysis of algorithms that can learn and improve their performance as they are exposed to data. The power of this idea is illustrated by the following example: although it seems hopelessly difficult to explicitly

program a computer to determine what objects are in a picture, the Viola–Jones machine-learning system can detect human faces in real time after being trained on a labelled sample of photographs[6]. Today, we regularly interact with machine-learning algorithms, from virtual assistants on our phones to spam filters for our e-mail. But these modern real-world applications trace their origins to a subfield of machine learning that is concerned with the careful formalization and mathematical analysis of various machine-learning settings.

The goal of learning a predictor (a mathematical function that can be used to make predictions) from a database of random examples was formalized in the aptly named probably approximately correct (PAC) learning model[7]. In this model, the aim is to train the predictor to match some true function that labels the data. A different model, called online learning, has the learner making immediate predictions as data arrive — for example, capturing a trading system's task of executing transactions in an ever-changing market. And another model known as multi-armed bandits can simulate clinical trials, in which the medical outcomes that an experimenter observes depend on his or her own choices.

These are only a few examples of the many