discovery, a team of biophysicists used light traps called optical tweezers to manipulate a single molecule of RNA, which interacts with itself to form structures resembling hairpins. By repeatedly opening and closing an RNA hairpin, the researchers showed that the work fluctuations accurately obeyed the Jarzynski relation[2].

The nanometre scale is also the realm of molecular motors: protein complexes that act as mechanical engines in the cell. Molecular motors can perform work by stretching biopolymers or by transporting cargo against a drag force. In doing so, they drive mechanical processes — from cell division to muscle contraction — occurring at larger scales. A single molecular motor is fuelled by molecules that each provide around 20 times the thermal energy at room temperature. This means that the operation of biomolecular engines is, in principle, constrained by Jarzynski's equality, making his research the first step in revealing how the nanoscale thermodynamics of molecular motors[3] goes beyond the macroscopic thermodynamics of steam engines.

Jarzynski's discovery — an equality hidden in plain sight behind the second law's inequality — also had a major impact on the physics of non-equilibrium processes. Strikingly, the relationship makes no assumptions about how fast a thermodynamic transformation takes place. It thus applies, for example, if the system is driven far from thermal equilibrium during the experiment. This in turn implies that the work fluctuations of a non-equilibrium process are directly related to equilibrium free-energy differences — a surprising result, suggesting that such processes obey general laws that can be derived exactly.

Such laws are known as fluctuation theorems, and they had already started to appear in the years before Jarzynski's result, but their impact had been appreciated by only a narrow group of researchers[4,5]. The most popular fluctuation theorem, Gavin Crooks' refinement of Jarzynski's equality, offers a statistical description of the arrow of time of a process — a way of calculating the probability that a video of an experiment is running forwards or backwards[6]. This surge of fundamental discoveries sparked the development of modern stochastic thermodynamics, which is a mathematical framework describing the thermodynamic properties of microscopic systems ranging from nanometre-scale molecular motors to micrometre-scale colloids[7].

Although Jarzynski's equality remains essential to our understanding of modern statistical physics and has been verified in many contexts[2,8], its direct application has so far proven underwhelming, and it has yet to result in any major technological advances. For instance, although it can be used as a tool to determine free-energy differences in molecular dynamics simulations[9], this application is limited by the fact that the relationship relies on very rare events. A promising class of thermodynamical uncertainty relationship has emerged to describe non-equilibrium fluctuations, which are more practical because they do not rely on such rare events[10,11].

All in all, Jarzynski's article remains a physicist's favourite, and a must-read for students, for its simplicity and elegance, as well as its impact on our understanding of thermodynamics. Rarely has a breakthrough discovery been so clearly evident at the time of its publication.

Chase P. Broedersz is in the Department of Physics and Astronomy, Vrije Universiteit Amsterdam, 1081 HV Amsterdam, the Netherlands; and the Arnold Sommerfeld Center for Theoretical Physics, Ludwig Maximilian University of Munich, 80333 Munich, Germany. **Pierre Ronceray** is in the Interdisciplinary Center of Nanoscience of Marseille, the French National Centre for Scientific Research and the Turing Centre for Living Systems, Aix-Marseille Université, 13009 Marseille, France.
e-mails: pierre.ronceray@univ-amu.fr; c.p.broedersz@vu.nl

1. Jarzynski, C. *Phys. Rev. Lett.* **78**, 2690–2693 (1997).
2. Liphardt, J., Dumont, S., Smith, S. B., Tinoco, I. Jr, & Bustamante, C. *Science*, **296**, 1832–1835 (2002).
3. Pietzonka, P., Barato, A. C. & Seifert, U. *J. Stat. Mech.* **2016**, 124004 (2016).
4. Evans, D. J., Cohen, E. G. D. & Morriss, G. P. *Phys. Rev. Lett.* **71**, 2401–2404 (1993).
5. Gallavotti, G. & Cohen, E. G. D. *Phys. Rev. Lett.* **74**, 2694–2697 (1995).
6. Crooks, G. E. *Phys. Rev. E* **60**, 2721–2726 (1999).
7. Seifert, U. *Rep. Prog. Phys.* **75**, 126001 (2012).
8. Saira, O.-P. *et al. Phys. Rev. Lett.* **109**, 180601 (2012).
9. Park, S. & Schulten, K. *J. Chem. Phys.* **120**, 5946–5961 (2004).
10. Barato, A. C. & Seifert, U. *Phys. Rev. Lett.* **114**, 158101 (2015).
11. Gingrich, T. R., Horowitz, J. M., Perunov, N. & England, J. L. *Phys. Rev. Lett.* **116**, 120601 (2016).

The authors declare no competing interests.

Molecular biology

Accessible method maps mutations to biophysics

Debora S. Marks & Stephen W. Michnick

A general method that quantifies and disentangles the effects of a gene's mutations on the traits of its protein enables assessments of mutational effects on protein biophysics for many of the proteins of a living organism. **See p.175**

Proteins function through their interactions with other biomolecules, and can be modulated not only by changes to the amino-acid residues at the protein–biomolecule interface, but also by changes at distal sites — a phenomenon called allostery. Discovering which protein sites are allosteric has been difficult, because methods that simply and comprehensively quantify allosteric effects have been lacking. On page 175, Faure *et al.*[1] report the first integrated method to globally map, quantify and distinguish the effects of mutations on allostery and protein stability. Their approach is also easy to do, making it as accessible to geneticists as it is to biophysicists. This method will have broad applications in biotechnology and drug discovery.

Advances in experimental technologies, especially DNA sequencing, have enabled the simultaneous measurement of the effects of thousands of mutations on a single protein[2]. Many hundreds of these high-throughput experiments have been published[3], focusing on assays that can infer mutational effects on a wide range of protein properties and processes — including heat stability, catalytic activity, binding to other proteins and small molecules, cell growth and drug resistance. The COVID-19 pandemic has showcased the value of these technologies, which have been used to screen how mutations in the spike protein of the SARS-CoV-2 virus affect the protein's expression and binding to host receptors[4] and antibodies[5]. Computational methods have also matured, and can now make surprisingly accurate predictions of mutational effects[6] across thousands of proteins without using experimental data[7]. All of these successes have led to the emergence of an international consortium called the Atlas of Variant Effects Alliance (www.varianteffect.org), whose goals include the open sharing of experimental and computational methods among scientists.

However, in practice, neither the experimental nor computational approaches have been able to distinguish the effects of mutations on different protein traits. To predict and, ultimately, engineer proteins that have desired traits requires an understanding of how biophysical properties are encoded by protein sequences.

Furthermore, ambiguities in discerning

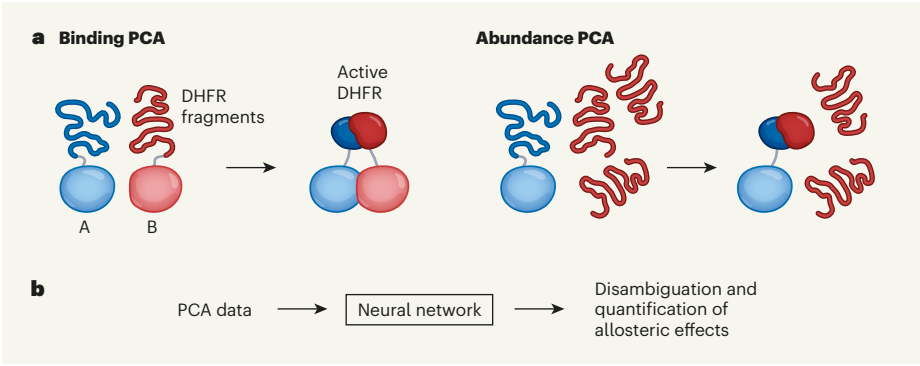

Figure 1 | A method for quantifying allosteric effects. a, Faure and colleagues' method[1] uses a combination of screening techniques known as protein-fragment complementation assays (PCAs). In the binding PCA, cells express protein A fused to a fragment of an enzyme (dihydrofolate reductase, DHFR) and protein B fused to a second DHFR fragment. If A binds to B, the DHFR fragments come together and form active DHFR. DHFR activity accelerates cell growth, which acts as a measure of the affinity of A for B. In the abundance PCA, the second DHFR fragment is not attached to B. Cell growth associated with DHFR activity correlates only with the abundance (and therefore stability) of A. **b**, Data from the binding and abundance PCA screening of thousands of mutants of A are fitted to thermodynamic models of the protein by an artificial neural network. The fitted models are used to calculate the free-energy differences specifically associated with the effects of mutations on the proteins' stability and allosteric effects on binding affinity.

which protein traits are affected by mutations need to be resolved. For instance, we might want to know how mutations affect the binding of an antibody to a viral protein; but mutations to a protein that affect its binding to an antibody in an experiment might also decrease the protein's stability or increase its abundance. A single experiment will therefore not be able to identify effects that are specifically attributable to changes in antibody binding. Furthermore, in genetics, different mutations (or combinations of mutations) can have the same effect on an organism, but arise from multiple biophysical causes. How can we distinguish between these different causes?

Faure and colleagues address this fundamental challenge using a combination of experimental and computational strategies. First, the authors describe an *in vivo* screening method that amplifies the effects of individual and double protein mutations, and distinguishes those that alter the protein's stability from those that affect its binding affinity for other proteins. Their approach uses two versions of previously reported techniques, known as protein-fragment complementation assays (PCAs), that are based on the activity of the enzyme dihydrofolate reductase (DHFR)[8,9].

In the first type of PCA used by the authors, two proteins are expressed in cells (Fig. 1). The first protein (let's call it A) is fused to an amino-terminal fragment of DHFR; and a second protein (B), which can bind to protein A, is fused to a complementary carboxy-terminal fragment of DHFR. If A binds to B, then the two fragments are brought together to form a functional DHFR enzyme. This confers the cells with resistance to an antibiotic, and thereby allows them to grow. The higher the binding affinity is for A to B, the greater the number of DHFR molecules reconstituted from fragments, and the faster the cell growth. Libraries of mutants of protein A can thus be screened in cells to find out which mutations increase its binding affinity for protein B.

However, cell growth in that assay could also change if the mutations alter the stability (and therefore the abundance in cells) of protein A. To distinguish effects on protein stability from effects on binding affinity, Faure and co-workers performed another PCA screen in which protein A, still fused to the N-terminal fragment of DHFR, is expressed in cells with the C-terminal fragment of DHFR alone. In this case, reconstitution of DHFR and associated cell growth depends only on the abundance of protein A. The authors call their combination of PCAs 'double deep PCA' (ddPCA).

"This technique is much simpler than the technologies currently used to determine mutational effects on protein traits."

Faure *et al.* used ddPCA to screen thousands of variants of two proteins, with each variant containing one or two changes in the amino-acid sequence, and thereby captured a comprehensive sample of mutational effects on the growth of cells in the assays. They then inferred from these data how each single or double amino-acid substitution alters the free energy of the proteins' stability and the free energy of the binding of the proteins to their partner protein (that is, protein B in the assay), using an artificial neural network to fit thermodynamic models to the experimental data. Importantly, by quantifying the effects of double mutants and measuring the mutational effects on two molecular traits, the number of mutations needed to fit the data to the models is reduced.

From the fitted models, the authors could infer the free-energy differences specifically associated with stability and binding affinity for all single mutations of the two proteins and their binding partners, relative to the free energies of the non-mutated proteins and the other mutants. In other words, Faure *et al.* were able to disentangle and quantify the size of the allosteric effects of mutations on these two protein traits. The authors compared their inferred results with free-energy differences determined experimentally from *in vitro* measurements of protein binding and stability, and observed remarkably strong agreement. Faure and colleagues' technique is much simpler than the technologies currently used to determine mutational effects on these traits. The methods and resources used are similar to those available to most university or industry laboratories, so the approach could be widely adopted by other researchers.

So what's next for ddPCA? The generalizability of the technique suggests that it could be scaled up to investigate the tens of thousands of protein interactions that occur in an organism[10]. If so, then it could be used to identify the residues involved in a huge range of protein-mediated processes – such as enzyme catalytic activity, post-translational modifications (chemical modifications of proteins) and disease states, to name but a few. Much as genetics is increasingly becoming a computational discipline, ddPCA could now take it into the realm of biophysics.

Debora S. Marks is in the Department of Systems Biology, Harvard Medical School, Boston, Massachusetts 02115, USA, and at the Broad Institute of Harvard and MIT, Cambridge, Massachusetts. **Stephen W. Michnick** is in the Department of Biochemistry, University of Montreal, Montreal H3C 3J7, Canada.
e-mails: debbie@hms.harvard.edu; stephen.michnick@umontreal.ca

1. Faure, A. J. *et al. Nature* **604**, 175–183 (2022).
2. Fowler, D. M. & Fields, S. *Nature Methods* **11**, 801–807 (2014).
3. Rubin, A. F. *et al.* Preprint at bioRxiv https://doi.org/10.1101/2021.11.29.470445 (2021).
4. Starr, T. N. *et al. Cell* **182**, 1295–1310 (2020).
5. Greaney, A. J. *et al. Cell Host Microbe* **29**, 463–476 (2021).
6. Hopf, T. A. *et al. Nature Biotechnol.* **35**, 128–135 (2017).
7. Frazer, J. *et al. Nature* **599**, 91–95 (2021).
8. Pelletier, J. N., Campbell-Valois, F.-X. & Michnick, S. W. *Proc. Natl Acad. Sci. USA* **95**, 12141–12146 (1998).
9. Campbell-Valois, F.-X., Tarassov, K. & Michnick, S. W. *Proc. Natl Acad. Sci. USA* **102**, 14988–14993 (2005).
10. Hein, M. Y. *et al. Cell* **163**, 712–723 (2015).

The authors declare no competing interests.