parameters used to describe their structure (root-mean-square deviation or distances between specific points within the protein) as descriptors of the activation mechanism. This allows them to define a series of discrete conformational states by clustering the raw simulation data along these four regions, and also to compute the probabilities of transitions between these states.

Now comes the 'trick'. The authors show that, at a certain timescale, these discrete states have Markovian behaviour, that is, the properties of one state depend only on the properties of the previous state. This allows them to 'stitch' these states, coming from thousands of independent simulations, into a single model, where the transition probabilities between states represent the pathways that the receptor could follow during its activation process. This 'full' Markov model can then be further simplified into a 'human readable' activation model composed of only ten intermediate states (Fig. 1). Finally, by sampling these intermediate states using Monte Carlo techniques and the calculated transition probabilities, the authors condense the activation process into 'summary' trajectories of 150 μs. The trajectories represent condensed movies of the spontaneous fluctuations at equilibrium and provide a dynamic picture of GPCR activation, allowing the discovery of transition pathways hidden in the vastness of the raw data.

The authors also performed computational docking calculations — using a set of known ligands and structurally similar decoys — to test whether the intermediate states that they had identified were suitable for *in silico* drug screening. For comparison, they also performed the same docking study using active and inactive receptor structures obtained by X-ray crystallography, and

randomly selected structures extracted from long-timescale 'classical' molecular-dynamics simulations. They found that using the intermediate states determined using their methods improved the discrimination between true ligands and decoys, and interestingly, it also led to an enrichment of the chemical diversity of the docked ligands. These findings support an intuitive postulate in the GPCR field, that is, that intermediate states in the activation pathway are distinct enough to recognize different chemotypes. This represents an obvious advantage in the search for new drugs, by allowing the exploration of new types of chemical compound for therapeutic use. And also a not so obvious one: by stabilizing specific intermediate conformations, such drugs could result in the activation of specific subsets of cellular signalling pathways, acting as 'biased drugs', a field of intense current research in the GPCR field[5].

The most interesting aspect of this work is the depiction of the protein activation pathways. The existence of multiple discrete activation intermediates has been observed experimentally[6], but these simulations suggest how they may look and connect to each other. In other aspects, this work recapitulates previously described aspects of GPCR activation[7]. There are, however, some discrepancies to what we know from GPCR activation, in particular in how different ligands affect the dynamics of the receptor. There are many possible sources for these differences: simulations have been performed not on the full protein, but on models that lack flexible regions that may have a role in receptor dynamics. Also there are other factors (such as the presence of regulatory proteins, the influence of membrane composition and so on) that are not considered in the simulations,

despite their role in protein dynamics[8]. Such approximations are common in molecular dynamics simulations of GPCR function, and, therefore, this type of simulation should still be regarded as qualitative, as they are not yet able to reproduce every small detail of protein activation.

This research represents a remarkable example of the application of cloud computing, big data analysis and Markov state models to large-scale molecular dynamics simulations of biomolecules. Such cloud-computing methodologies, which complement standard simulations in parallel supercomputers, are already accessible to computational chemical biologists, but they are still in their infancy. However, with studies such as this from Kohlhoff, Altman, Pande and co-workers leading the way, they will certainly allow the length of simulations of protein dynamics to be extended to truly biologically relevant timescales in the near future. ❑

Xavier Deupi is in the Department of Research with Neutrons and Muons and in the Department of Biology and Chemistry at the Paul Scherrer Institute, 5232 Villigen PSI, Switzerland.
e-mail: xavier.deupi@psi.ch

References
1. Kohlhoff, K. J. *et al. Nature Chem.* **6,** 15–21 (2014).
2. Dror, R. O., Dirks, R. M., Grossman, J. P., Xu, H. & Shaw, D. E. *Ann. Rev. Biophys.* **41,** 429–452 (2012).
3. Johnston, J. M. & Filizola, M. *Curr. Opin. Struct. Biol.* **21,** 552–558 (2011).
4. Lohse, M. J. *et al. Brit. J. Pharmacol.* **153,** S125–S132 (2008).
5. Whalen, E. J., Rajagopal, S. & Lefkowitz, R. J. *Trends Molecular Med.* **17,** 126–139 (2011).
6. Bockenhauer, S., Fürstenberg, A., Yao, X. J., Kobilka, B. K. & Moerner, W. E. *J. Phys. Chem. B* **115,** 13328–13338 (2011).
7. Dror, R. O. *et al. Proc. Natl Acad. Sci. USA* **108,** 18684–18689 (2011).
8. Deupi, X. & Kobilka, B. K. *Physiology* **25,** 293–303 (2010).

SINGLE-MOLECULE CHEMISTRY

Knowing your neighbours

Quantitatively studying how the rate of a chemical reaction is affected by a reactant's atomic-scale environment is extremely challenging. This has now been achieved at the single-molecule level using scanning tunnelling microscopy to monitor tautomerization in an atomically well-defined environment.

Peter Liljeroth

Does the rate of a chemical reaction depend on the atomic-scale details of its environment? It is easy to reply "yes, of course" and to provide systems demonstrating this effect: for example, heterogeneous catalysis on different crystal

facets and photoisomerization in solution versus on a surface. It is much more difficult to answer the following, more detailed questions: how exactly does the rate of the chemical reaction depend on the atomic environment, and what happens if I move

one of the nearby atoms by, say, 5 Å? These are precisely the types of question that, as they report in *Nature Chemistry,* Leonhard Grill and co-workers have quantitatively answered using scanning tunnelling microscopy (STM) to study

porphycene molecules adsorbed on a copper surface[1].

They focus on the hydrogen tautomerization reaction in porphycene (Fig. 1a). This reaction occurs spontaneously at room-temperature, but can be frozen by carrying out the experiments at low temperatures — allowing the reaction to be controlled using the STM tip. As in earlier experiments on naphthalocyanine and porphyrin molecules[2,3], when the bias voltage between the STM tip and the substrate is increased sufficiently, the tunnelling electrons can drive the tautomerization reaction, causing the molecule to flip between the two tautomers. The reaction has a voltage threshold of 150 mV and is triggered by electron-tunnelling-induced vibrational excitation of the molecule.

The *cis* tautomers are stabilized by the interaction of the non-hydrogenated nitrogen atoms with the underlying atoms of the Cu(110) surface, making them more stable than the *trans* tautomers, which is in contrast to the gas phase. The reaction observed therefore corresponds to tautomerization from one *cis* form to another (see Fig. 1a), and is formally equivalent to a rotation of the whole molecule by 180 degrees. It affects the orientation of the frontier molecular orbitals, and hence the tunnelling current flowing between the STM tip and the substrate. This means that the reaction can be monitored (as well as induced) by STM because every tautomerization is associated with an abrupt change in the tunnelling current. By measuring the current over a period of time, the overall rate and residence times in the two tautomeric forms can be measured (Fig. 1a).

The tip is normally thought to be an innocent bystander in STM measurements, simply there to add or remove electrons at an atomically well-defined position on the sample surface. However, Grill and colleagues observed that bringing the tip too close had an effect on the rate of the tautomerization reaction and the equilibrium between the two tautomers. Motivated by this, they used the ability to move individual metal atoms on the surface using the STM tip — a process called lateral atomic manipulation in the STM community — to place individual copper atoms at specific locations around the porphycene molecule and then studied how their positions affected the hydrogen tautomerization reaction.

It is intuitively clear that the atomic-scale environment of a chemical reaction has an effect on the reaction rate, and this effect has been demonstrated in the case of molecular motion[4] and catalytic reactions[5]. However,
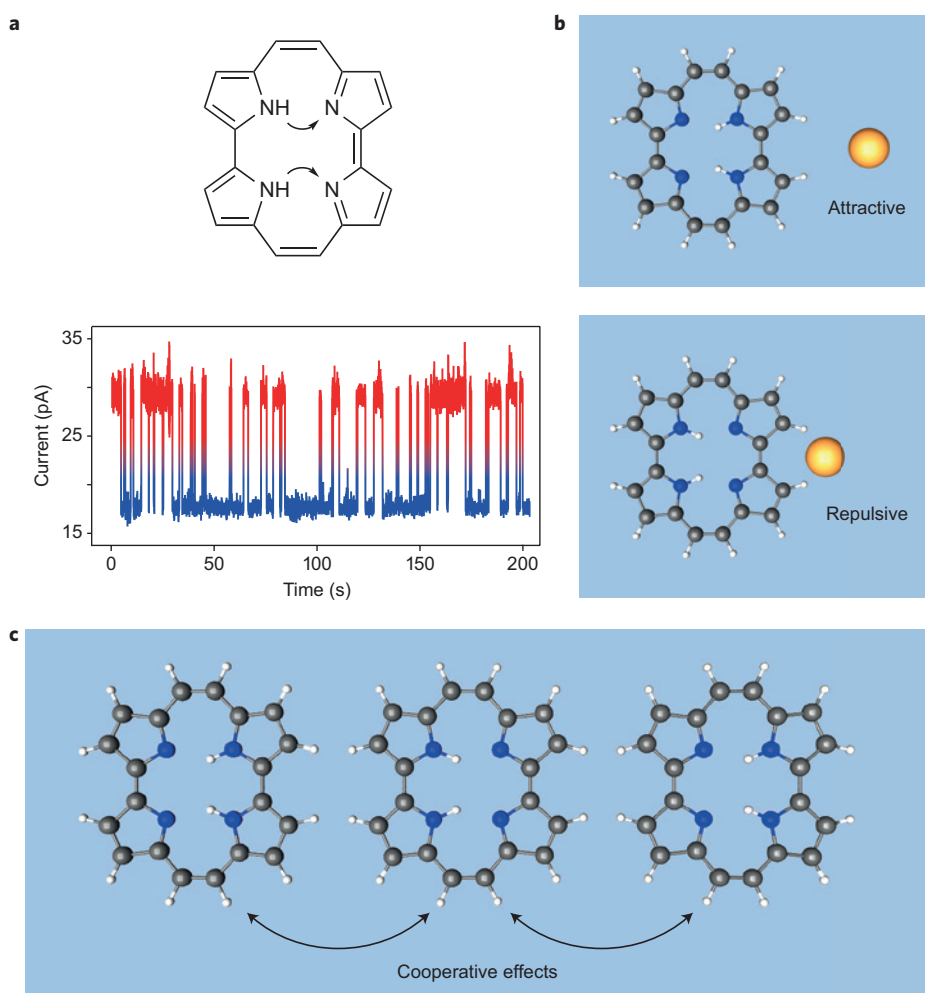
Figure 1 | Hydrogen tautomerization reaction in porphycene. **a**, Schematic of the porphycene molecule undergoing *cis*–*cis* isomerization and a current–time trace showing discrete jumps between two stable states corresponding to the two different tautomers[1]. **b**, Copper adatoms positioned with atomic precision have an effect on the equilibrium position of the hydrogen atoms showing both attractive (top) and repulsive (bottom) interactions. **c**, Multiple porphycene molecules in close-packed structures show cooperative effects, where the tautomerization state of a neighbouring molecule has an effect on the reaction on a given molecule.

up to now there have been no experimental results on the dependence of a chemical reaction rate on the actual position of neighbouring atoms on the single-molecule level. The experiments from Grill and co-workers show that individual copper atoms can modify the tautomerization reaction substantially —even when placed relatively far from the reaction centre. Effects were seen up to a distance of 0.9 nm, which is much larger than the van der Waals radius of porphycene (0.65 nm).

When there is no copper adatom present, the two *cis* foms are energetically degenerate and there is an equal chance of observing either form. The presence of the extra copper adatom, however, can shift the tautomerization equilibrium in either direction. At relatively large distances, the

tautomer for which the central hydrogen atoms are closer to the adatom is preferred (Fig. 1b, top). In contrast, when the adatom is closer, there are repulsive interactions between the hydrogen atoms and the adatom, and the tautomer for which the hydrogen atoms are further from the adatom is preferred. Both situations create energy penalties for tautomerization to the respective less stable form, but the effect is much larger in the latter (close) case and the reaction is essentially quenched. Experiments confirm that placing two adatoms symmetrically around the porphycene has no effect on the overall reaction rate, which can be rationalized based on simple symmetry arguments.

These experimental results were corroborated by extensive density functional

theory calculations, which are particularly valuable in examining specifically how the adatoms affect the reaction rate. It turns out that at large distances of up to around 1 nm, the effect is mostly due to attractive van der Waals interactions. This also explains the preference for the configuration in which the hydrogen atoms are closer to the adatom. On the other hand, at short distances, surface-related repulsive interactions become important and induce a shift in the tautomerization equilibrium.

The experiments undertaken are an extremely elegant, quantitative demonstration of basic chemical phenomena, and offer a platform for studying such effects further. This behaviour is highly unlikely to be limited to the tautomerization reaction of porphycene, and further studies will undoubtedly offer important insights into, for example, catalysis.

It is also worthy to note that this type of tautomerization reaction has also been proposed as a molecular-scale switch[2,3], and the present results bring both good and bad news for the achievement of this goal. The bad news is that when the active components of electronic circuitry shrink down to the molecular scale, the disrupting effects from disorder in the atomic environment become more relevant. This is probably unavoidable, and strategies for tolerating the inherent variability in the properties of the components (here chemical reaction rate and equilibrium) have to be developed instead of simply concentrating on trying to eliminate this variability.

The good news is that this sensitivity to the environment can be exploited and the first steps in this direction are already indicated by Grill and colleagues. They demonstrate that cooperative effects are important in assemblies of several porphycene molecules (Fig. 1c): the tautomerization reaction of a given molecule is affected by the tautomerization state of the neighbouring molecules. Realization of more complicated molecular information processing units based on assemblies of molecular switches requires precisely this kind of interaction between the switching units. Actual devices based on such concepts are some way off in the future, but the work of Grill and colleagues gives us a solid understanding of the underlying chemical phenomena of such systems. ❒

Peter Liljeroth is in the Department of Applied Physics, Aalto University School of Science, PO Box 15100, FI-00076 AALTO, Finland.
e-mail: peter.liljeroth@aalto.fi

References
1. Kumagai, T. et al. Nature Chem. 6, 41–46 (2014).
2. Liljeroth, P., Repp, J. & Meyer, G. Science 317, 1203–1206 (2007).
3. Auwärter, W. et al. Nature Nanotech. 7, 41–46 (2012).
4. Heinrich, A. J., Lutz, C. P., Gupta, J. A. & Eigler, D. M. Science 298, 1381–1387 (2002).
5. Wintterlin, J., Völkening, S., Janssens, T. V. W., Zambelli, T. & Ertl, G. Science 278, 1931–1934 (1997).

BIOSYNTHESIS

Non-stick natural products

Fluorine imparts many drugs with beneficial properties, however, the synthesis of fluorinated complex natural products is challenging. Biosynthetic strategies and recent experimental precedents have paved the way for bioengineered fluorinated polyketides.

Peter A. Jordan and Bradley S. Moore

Among the brave, early pioneers of fluorine chemistry, few would have guessed that this perilous element — which is rarely found in natural molecules — would hold such a privileged place in modern drug design. Yet, the current pharmacopoeia consists of an ever-expanding array of organofluorine drugs. At recent counts, over 20% of all pharmaceuticals contain at least one fluorine atom[1], and improvements in virtually all aspects of pharmacokinetics and pharmacodynamics have been demonstrated through the strategic placement of fluorine substituents[2]. Although significant progress has been made in developing safe and efficient methods of controlling organofluorine chemistry, many of the properties that make fluorine so attractive to medicinal chemists — such as electronegativity, sigma bond strength and high redox potential[3] — also contribute to the difficulties of working with fluorine. Nonetheless, great strides in organofluorine chemistry have been made, and many labs have set their sights on synthesizing fluorinated analogues of complex natural products. These non-fluorinated natural products are often discovered to be new lead compounds in drug-discovery programs[4], and a simple method of incorporating fluorine into these structures could unlock its beneficial properties.

Lab bench chemical synthesis can take advantage of transition metal catalysts, highly reactive fluorinating reagents and a water-free environment[4], whereas the biosynthetic strategies developed for this field so far depend on harnessing existing enzymatic reactions. With only a handful of fluorine-containing natural products and only one enzyme that converts inorganic fluoride to an organofluorine species, the careful consideration of appropriate fluoride sources and enzymes is essential[5]. A recent report in Science by Michelle C. Y. Chang and co-workers is set to change this by introducing a general strategy for biosynthesizing fluorinated natural products via the incorporation of fluorinated building blocks[6].

Acetate is one of the principal building blocks in biology that gives rise to a diverse range of metabolites — including fatty acids, polyketides (like the immunosuppressant rapamycin), and terpenoids (such as the anticancer agent taxol). In polyketide biosynthesis, acetate must first be loaded onto coenzyme A and carboxylated to form malonyl-CoA (Fig. 1). Simply substituting fluoroacetate for acetate at an early stage of biosynthesis could enable the design of fluorinated polyketide molecules. However, the extreme electronegativity of fluorine potentially compromises the stability of all biosynthetic intermediates. Despite these challenges, Chang and co-workers have described two distinct enzymatic approaches starting with fluoroacetate or fluoromalonate that converge to produce fluoromalonyl-CoA with high yields.

A biosynthetic route to fluoromalonyl-CoA is only the beginning. In its most basic form, polyketide assembly involves sequential Claisen condensation reactions. This requires generation of a highly nucleophilic carbanion intermediate from a malonyl substrate bound to an acyl carrier protein, with the concurrent loss of CO_2.