ASTROCHEMISTRY

Extracts of meteorite

The Sutter's Mill meteorite exploded in a dazzling fireball over California last year (pictured). Writing in *Proceedings of the National Academy of Sciences*, Pizzarello *et al*. report that it contained organic molecules not found in any previously analysed meteorites (S. Pizzarello *et al. Proc. Natl Acad. Sci. USA* http://dx.doi.org/10.1073/pnas.1309113110; 2013).

The authors' analysis of extracts of organic matter from the meteorite revealed a complex mixture of oxygen-rich compounds, which probably resulted from oxidative processes that occurred in the parent body. Intriguingly, the authors suggest that such compounds could have been released to prebiotic Earth during or after meteorite bombardments, adding to the roster of organic molecules that might have contributed to the evolution of life. **Andrew Mitchinson**

LISA WARREN

COMPUTATIONAL BIOLOGY

A recipe for ligand–binding proteins

Cellular cross-talk, enzymatic catalysis and regulation of gene expression all depend on molecular recognition. A method that allows the design of proteins with desired recognition sites could thus be revolutionary. **SEE LETTER P.212**

GIOVANNA GHIRLANDA

To support cellular processes, natural proteins have evolved to recognize a relatively small set of ligand molecules with high affinity and specificity. Broadening this set of protein–ligand pairs with synthetic proteins that are specific for ligands of choice could transform the development of biosensors, protein-based drugs, artificial enzymes and tools for chemical biology. Current *in vitro* approaches to the design of protein–ligand pairs rely on laborious directed-evolution techniques to engineer binding sites in existing protein scaffolds, and cannot be generalized. In this issue, Tinberg *et al.*[1] (page 212) describe a computational method that promises to streamline and greatly facilitate the development of tailored ligand-binding proteins*.

One way to engineer synthetic ligand-binding proteins is repurposing existing proteins through directed evolution. This approach involves selecting a few amino acids within an existing protein, typically located in a cavity, and randomly mutating them to create libraries of millions of protein variants, which are then selected for their ability to bind the target. Depending on the ligand and the scaffold protein, successive rounds of directed evolution yield binding proteins with dissociation constants in the physiologically relevant micromolar to nanomolar ranges[2,3]. The process is arduous, and success is not guaranteed: current methods limit the size of combinatorial libraries to about 10^9, which is often insufficient to explore a large number of positions and so limits the fitness of the resulting binders. Furthermore, the process works best when the starting protein already binds weakly to the desired ligand, rather than when starting from scratch.

Computational methods hold great promise to bypass this lengthy process by using virtual selection on a large number of scaffolds. Such approaches have enjoyed much success lately, with the design of artificial enzymes for a variety of reactions[4–8]. Despite these feats, the design of high-affinity ligand-binding proteins has remained elusive[9].

Tinberg *et al.* now present a generalizable strategy for designing artificial recognition modules for any desired target molecule. As their target they chose digoxigenin — a natural steroid that forms part of the cardiac drug digoxin, and that has found a second life as a recognition tag in biotechnology applications. Consequently, a few digoxigenin-binding agents, including both antibodies and non-antibody proteins, have been developed through traditional protein-engineering methods[10–12].

Taking some clues from the previously developed binding agents, Tinberg and colleagues sketched out a couple of possible ways to bind ligands to digoxigenin using amino acids that could provide hydrogen bonds and hydrophobic interactions (Fig. 1a). They then transposed these minimalistic sets into three-dimensional models, which were used as probes to interrogate a selected set of 401 proteins, encompassing several classes of

*This article and the paper under discussion[1] were published online on 4 September 2013.