for the hydrogen-producing reaction in a protective shell of a chromium compound.

This combination of complex mitigation strategies proved highly successful: the authors reported EQEs of up to 96% when their photocatalysts were irradiated with light in the wavelength range of 350–360 nanometres. This is excellent news, because it means they have designed an almost perfect photocatalyst – the IQE must be between 96% and 100%.

This is a spectacular result for several reasons, even though strontium titanate is 'just' a model system for visible-light photocatalysts. First, it demonstrates that experiments can be designed in which EQEs come close to IQEs within an acceptable error margin of less than 4%. Improved experimental set-ups in which measured EQEs are very near to IQEs should facilitate the comparison of photocatalysts and therefore accelerate progress in this field.

Second, it proves that the combination of design strategies used by the authors can indeed eliminate efficiency losses associated with recombination. It is to be expected that the strategies used to improve the efficiency of strontium titanate will also apply to photocatalysts driven by visible light – and could therefore enable the conversion of solar energy to hydrogen with efficiencies of about 10%.

Finally, and most importantly, Takata and colleagues' findings will inspire and encourage other researchers to continue their work on photocatalysts. One of the authors of the work, Kazunari Domen, published his first paper[9] on the use of strontium titanate as a photocatalyst in 1980. This shows the timescale needed for success in this area. Although we do not yet have a route for the sustainable and economically viable production of hydrogen, we stand a good chance of finding one in the next few decades. This paper vouches for it.

Simone Pokrant is in the Department of Chemistry and Physics of Materials, University of Salzburg, Salzburg A-5020, Austria.
e-mail: simone.pokrant@sbg.ac.at

1. Takata, T. *et al. Nature* **581**, 411–414 (2020).
2. Abbott, D. *Proc. IEEE* **98**, 42–66 (2010).
3. Ochoa Robles, J., De-León Almaraz, S. & Azzaro-Pantel, C. in *Hydrogen Supply Chains* 8–11 (Academic, 2018).
4. Pinaud, B. A. *et al. Energy Environ. Sci.* **6**, 1983–2002 (2013).
5. Wrighton, M. S., Wolczanski, P. T. & Ellis, A. B. *J. Solid State Chem.* **22**, 17–29 (1977).
6. Ham, Y. *et al. J. Mater. Chem. A* **4**, 3027–3033 (2016).
7. Giocondi, J. L. & Rohrer, G. S. *J. Am. Ceram. Soc.* **86**, 1182–1189 (2003).
8. Mu, L. *et al. Energy Environ. Sci.* **9**, 2463–2469 (2016).
9. Domen, K., Naito, S., Soma, M., Onishi, T. & Tamaru, K. *J. Chem. Soc. Chem. Commun.* 543–544 (1980).

Molecular biology

Evolution of a molecular machine

Michael Berenbrink

The multi-subunit protein haemoglobin relies on complex interactions between its components to function properly. Analysis of ancient precursors suggests that its evolution from a simple monomer involved only a few steps. **See p.480**

The oxygen-transporting protein haemoglobin has undergone repeated adaptations as animals evolved to conquer new environments – from the depths of the oceans[1] to high mountain ranges[2]. These adaptations relied on changes in the long-range interactions between oxygen-binding sites buried in the protein's subunits, and between these regions and binding sites for a multitude of small effector molecules on the protein's surface[3]. How did this complex molecular machine, which can respond so exquisitely to available levels of both oxygen and several other effector molecules, come into being? On page 480, Pillai *et al.*[4] reconstruct the stepwise evolution of haemoglobin from precursors that existed more than 400 million years ago.

Almost nothing was previously known about how the four-subunit (tetrameric) form of haemoglobin that is found in modern-day jawed vertebrates evolved from ancient monomers. Tetrameric haemoglobin consists of two α- and two β-subunits. Pillai *et al.* computationally reconstructed an evolutionary tree to chart the protein's ancient history, using the amino-acid sequences of a large collection of the closely related vertebrate globin proteins, which exist as either monomers or tetramers.

"Pillai and colleagues' work serves as one of the clearest examples so far of how such complexity can arise."

The authors' tree was constructed taking into account that amino-acid substitutions a given protein shares with close relatives tend to have originated in more-recent common ancestors than have those it shares with more-distant relatives. The reconstructed evolutionary tree indicates that multiple rounds of gene duplication and subsequent divergence gave rise to the globin family and, by way of several ancestral proteins, to tetrameric haemoglobin (Fig. 1).

What is special about the study is that Pillai and colleagues went on to resurrect several of these extinct ancestral proteins, generating them from the amino-acid sequences predicted by the tree. The group then tested these proteins' functions.

First, Pillai and colleagues analysed whether each ancestral protein could form dimers and tetramers of like or unlike subunits. The earliest protein – a common ancestor of haemoglobin and the monomeric globin protein myoglobin, named AncMH by the authors – exists only as a monomer. A later protein, named Ancα/β, which is the ancestor of all existing haemoglobin subunits, forms homodimers when expressed at high levels. The authors' tree indicates that Ancα/β underwent gene duplication to produce two proteins: the ancestors of all existing α- or β-subunits, which the group respectively named Ancα and Ancβ. These proteins also form homodimers, or even homotetramers, when expressed alone. However, when the two are expressed together in equal proportions, they can form heterodimers, which then further align to yield haemoglobin tetramers.

The group next investigated the oxygen-binding affinity of the ancestral proteins, along with their oxygen cooperativity (the ability of oxygen-binding subunits to interact with one another) and their 'allosteric' regulation by a potent, artificial effector molecule, inositol hexaphosphate (IHP). They found that only Ancα and Ancβ – when expressed together at high concentrations – show similar oxygen-binding affinity, cooperativity and allosteric regulation to today's haemoglobin protein. These features are shared by all living jawed vertebrates, but are absent or achieved in a different way in jawless vertebrates, whose haemoglobin proteins are of more ancient origin. This indicates that the basic functions of jawed-vertebrate haemoglobin had already evolved in a common ancestor of these animals but at some time after the split with jawless vertebrates.

Next, Pillai *et al.* modelled the stepwise changes in α- and β-subunit interfaces that might have allowed Ancα and Ancβ first to form heterodimers with one another, and later heterotetramers from pairs of such dimers. The modelling indicated that strikingly few amino-acid substitutions might have been needed to transform a simple monomeric

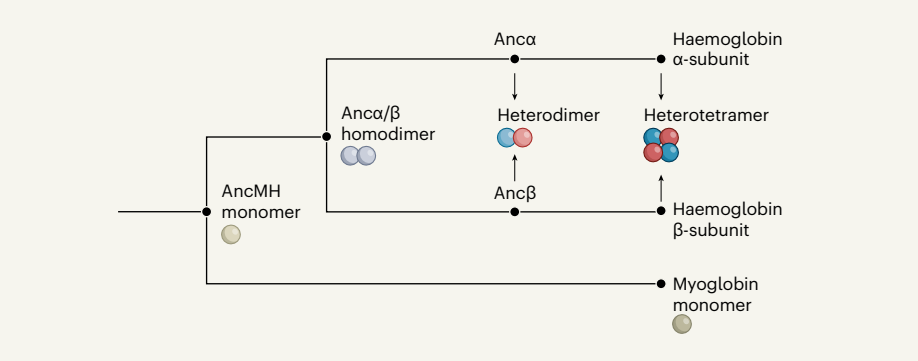

Figure 1 | Key steps in the evolution of the tetrameric haemoglobin protein. In jawed vertebrates, haemoglobin exists as a tetramer, formed from two α- and two β-subunits. Pillai *et al.*[4] resurrected extinct ancestors of haemoglobin using predicted amino-acid sequences to reconstruct the protein's evolution. The authors showed that the protein AncMH, the last common ancestor of haemoglobin and the related protein myoglobin, existed as a monomer. Duplication of the gene that encoded AncMH, and subsequent divergence into two genes, produced monomeric myoglobin and the ancestor of haemoglobin, Ancα/β, which forms a homodimer. Further gene duplication of Ancα/β and subsequent divergence yielded the ancestors of the α- and β-subunits, dubbed Ancα and Ancβ. These two subunits evolved an interface that allowed the formation of heterodimers. A few further changes in amino-acid residues generated a second interface that allowed the assembly of modern-day α- and β-subunits into haemoglobin heterotetramers.

oxygen-binding protein similar to myoglobin (whose oxygen binding is non-cooperative and almost totally unaffected by allosteric effector molecules[5]) into haemoglobin. Importantly, the researchers verified the results of their model by generating proteins that carried mutations of the amino-acid residues identified, and showing that heterotetramer formation was disrupted.

The authors' work shows how natural selection, acting on pre-existing biophysical protein properties, can, in just a few evolutionary steps, create multimeric structures that have complex functions. Most cellular processes involve the action of protein multimers, and Pillai and colleagues' work serves as one of the clearest examples so far of how such complexity can arise during protein evolution.

There are inevitable uncertainties in these kinds of reconstruction of the deep past of life, because the accuracy of such reconstructions relies on the proteins under consideration having several specific properties[6]. The proteins should, ideally, show small overall rates of amino-acid sequence divergence from one another, have thoroughly known and well-supported evolutionary relationships to each other, and exhibit a dense evolutionary branching pattern over the time period(s) of interest. Finally, a detailed knowledge of structure–function relationships is essential. It would be difficult to reconstruct with any confidence ancestral sequences and functions for proteins that do not fulfil all or any of these conditions. However, haemoglobin is well suited for this type of study for several reasons. For instance, a wealth of comparative data on globin function across vertebrates has accumulated over the past 100 years[7]. We have intimate knowledge of haemoglobin structure–function relationships[3,8,9]. In addition, there is an ever-expanding pool of globin sequence information, thanks to genome-sequencing projects in diverse organisms (these efforts will also benefit similar studies on other proteins).

Pillai and colleagues' study is sure to raise several follow-up questions and to spark further research. For instance, the authors used the artificial effector IHP in their experiments, but the binding sites for physiologically relevant effectors of haemoglobin oxygen affinity, such as hydrogen ions, only partly overlap with — and in some cases are quite different from — the IHP binding site[3,8,9]. Some evidence[1] suggests that mechanisms by which hydrogen ions modify haemoglobin oxygen affinity have evolved independently multiple times in vertebrates. This would make the picture much more complicated than can be assessed using IHP.

It will be interesting to probe the evolutionary origins of the regulation of haemoglobin oxygen binding by other effectors, such as carbon dioxide or physiologically relevant organic phosphates including ATP and 2,3-bisphosphoglycerate. In doing so, we could examine, for instance, how haemoglobin regulation changed as demands on the body's oxygen-transport system rose during the evolution of warm-blooded, active vertebrates, or how it was affected by changes over geological time in atmospheric oxygen levels, which have been proposed by some to have shaped vertebrate evolution[10].

The assembly of multimeric proteins depends on specific concentrations and thus expression levels of the protein's subunits. Natural selection presumably prevents imbalanced subunit production, both to limit the costly energy expenditure involved in protein synthesis and to prevent accumulation of potentially harmful spare globin subunits, as occurs in some hereditary human blood disorders[11]. But, at some point, an increase and balancing of expression levels of haemoglobin's subunits would have been needed to enable tetramer formation. When did this occur? It has been shown[12] that the net surface charge of myoglobin acts as a molecular 'signature' that can be used to assess the expression levels of ancestral myoglobin. However, such markers are largely unknown from other globin proteins, and we do not know the ancestral expression levels of any of the reconstructed haemoglobin precursors in Pillai and colleagues' study.

Finally, as previously noted[13], one of the most fascinating frontiers in this research field might be uncovering the evolutionary history of gene regulation. This remains an open question in the evolution of the genes that encode haemoglobin's subunits.

Michael Berenbrink is in the Department of Evolution, Ecology and Behaviour, University of Liverpool, Liverpool L69 7ZB, UK.
e-mail: michael.berenbrink@liverpool.ac.uk

This article was published online on 20 May 2020.

1. Berenbrink, M., Koldkjær, P., Kepp, O. & Cossins, A. R. *Science* **307**, 1752–1757 (2005).
2. Natarajan, C. *et al. Science* **354**, 336–339 (2016).
3. Perutz, M. F. *Mol. Biol. Evol.* **1**, 1–28 (1983).
4. Pillai, A. S. *et al. Nature* **581**, 480–485 (2020).
5. Antonini, E. & Brunori, M. *Hemoglobin and Myoglobin in Their Reactions with Ligands* 21 (North-Holland, 1971).
6. Benner, S. A. in *Ancestral Sequence Reconstruction* (ed. Liberles, D. A.) Ch. 1, 3–19 (Oxford Univ. Press, 2007).
7. Krogh, A. & Leitch, I. *J. Physiol. (Lond.)* **52**, 288–300 (1919).
8. Storz, J. F. *Hemoglobin: Insights into Protein Structure, Function, and Evolution* (Oxford Univ. Press, 2019).
9. Imai, K. *Allosteric Effects in Haemoglobin* (Cambridge Univ. Press, 1982).
10. Schachat, S. R. *et al. Proc. R. Soc. B* **285**, 20172631 (2018).
11. Rund, D. & Rachmilewitz, E. *N. Engl. J. Med.* **353**, 1135–1146 (2005).
12. Mirceta, S. *et al. Science* **340**, 1234192 (2013).
13. Zuckerkandl, E. in *Ancestral Sequence Reconstruction* (ed. Liberles, D. A.) x–xi (Oxford Univ. Press, 2007).