European Symposium of Bio-Organic Chemistry 2003 (ESBOC): The Evolution of Catalysis

Many in England are convinced that strange things go on in Wales (the use of a language different from English being one of them). The theme of the latest Gregynog research conference—the evolution of catalysis—may seem strange, too, in the eyes of the synthetic and mechanistic chemists who normally flock every spring to the idyllic conference centre of *Prifysgol Cymru*, the University of Wales, in the Montgomeryshire countryside. However, the theme of the conference, Directed Evolution—that is, iterative cycles of selection from (or screening of) libraries of diverse molecules—has come of age, and its potential is recognised beyond the circle of pioneers of the area. Two such pioneers, Andreas Plückthun (Zürich University) and Don Hilvert (ETH Zürich), masterminded the conference and set out the challenge of research in directed evolution at the beginning of the meeting—to ultimately emulate enzyme function.

The first session defined why this is difficult: Richard Wolfenden (North Carolina) in collaboration with Nick Williams (Sheffield) has measured the exceedingly slow intrinsic rates of key biological reactions.[1] The half lives involved are in the order of 10^{12} years at 25 °C for phosphate monoester transfer. Too long to wait! So these exceedingly slow processes were studied at elevated temperatures, to allow extrapolation back to room temperature. The comparison with enzymes shows awe-inspiring rate accelerations of over 10^{20}. Next to efficiency, precision of biological machines is a fundamental challenge. Bob Stroud (UC San Francisco) described how, in a combination of site-directed mutagenesis, kinetics and crystallography, not only static active site interactions can be mapped out, but complex conformational changes can be tracked down for thymidylate synthase, a methyl transfer catalyst.[2, 3] Nenad Ban (ETH Zürich) described the ribosome structure:[4–7] a protein of this complexity is definitely no realistic target for directed evolution. Nonetheless we had a taste of a successful designed alteration of protein structure. Kevan Shokat (UC San Francisco) has demonstrated how tyrosine kinases with an engineered binding pocket in a conserved region of the ATP-binding site can be specifically inhibited in vivo, with synthetically prepared inhibitors found by screening. It greatly facilitated the study of the cell biology of these enzymes, which play a vital role in signal transduction and is being applied to to a growing list of key signal transduction kinases (e.g., src kinases).[8–12]

But what *is* realistic? It is definitely impossible to scan the *entirety* of sequence space—even a fairly small enzyme with about 50 amino acids would, if completely randomised, create a library of more molecules that there are in the whole universe. A good rationale could enable more targeted sampling of sequence space. Dagmar Ringe (Brandeis) introduced enzyme superfamilies based on active site chemistry as evolutionary neighbours—possible shortcuts for directed evolution.[13] For example, eukaryotic muconate lactonising enzyme has the same signature motif and chemical functionality as its bacterial analogue (and superfamily member), but a very different fold (β propeller instead of TIM barrel). Thus chemistry rather than structure becomes the guide for exploring new activities and understanding an enzyme's "catalytic phylogeny", an idea that differs from previous concepts of evolution by maintaining and re-evolving a given fold such as the TIM barrel.

Ultimately it is crucial that the desired solution sequence actually is in one's library, but often good library members are diluted by sequences that do not even give rise to a folded protein. Frances Arnold (Caltech) has developed an approach that would generate libraries that are relatively small, yet high in quality, thus increasing the chance of a hit. By using synthetic shuffling, fragments are chosen for being quasiautonomous, that is, amino acid stretches that share maximal contact, though do not (optimally) contact any other residue. These fragments are calculated from available structural data by the SCHEMA shareware.[14]

Occasionally it might even be possible to get by without experiment. Steve Benner (University of Florida) introduced the Master Catalog,[29] a genetic sequence database organized by evolutionary families. Every documented protein was grouped into a family of related sequences, postulating in essence their evolutionary history and reconstructing their ancestral common sequence.[30] In addition, this strategy can be useful for tracing the origins and function of new sequences with great speed, for example, of the SARS virus, thus identifying its primary source and suggest possible therapeutic strategies.

But the field is still driven by technologies that are in themselves an area of active research. The crucial technological requirement in all library approaches is how one's selected molecule can be decoded: selection screens for phenotype, but the genotype holds the key, and has to be linked phenotype. This is more straightforward in selections for nucleic acids as genotype and phenotype *are* identical. Ron Breaker (Yale) showcased a variety of allosteric ribozymes that were evolved and used as molecular switches, governed by light, by the presence of metal ions or small organic molecules. Strikingly, such "riboswitches" (Figure 1) were identified as part of natural mRNAs, regulating their translation by recognising their cognate metabolite. Rather than the scientist mimicking nature, a novel discovery by directed evolution has shed light on a natural regulation mechanism.[15–17]

But where selection of protein catalysts is attempted, different genotype – pheno-

 DOI: 10.1002/cbic.200300796

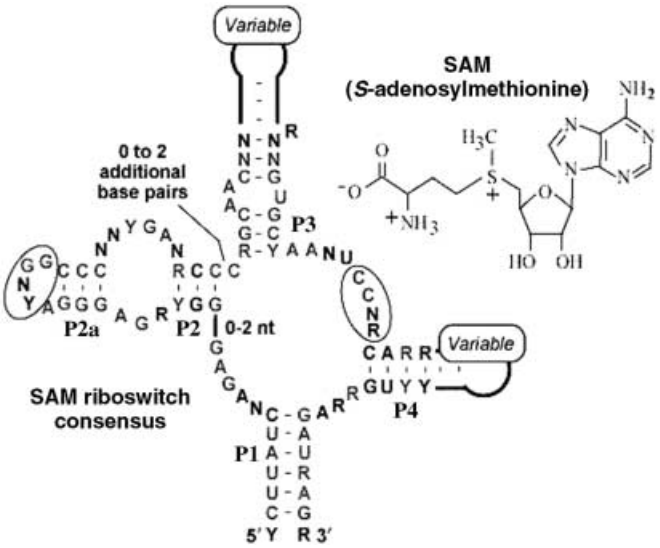

Figure 1. *The consensus sequence and secondary structure of the aptamer domain from S-adenosylmethionine (SAM) riboswitches. The grey and black notations depict nucleotides that are conserved in at least 90% and 80%, respectively, of the SAM riboswitches known to date.*

type unions are in use, as was exemplified in several key presentations. Alan Berry and Adam Nelson (Leeds) have solved a real-world problem in synthetic chemistry. From a library of randomly mutagenised class II aldolases they selected (by colony screening of cell lysates in 96-well format) enzymes that catalyse C–C bond formation in an aldol reaction, a key step in many natural-product syntheses. After rounds of DNA shuffling, the resulting enzyme reversed the stereochemistry of tagatose 100-fold, now accepting the unnatural substrate preferentially.[18] Presenting another approach to the synthesis and cleavage of C–C bonds, Veronique Gouverneur (Oxford) has explored the substrate versatility of catalytic aldolase antibodies for a variety of substrates.[19]

Tuning the classic phage-display selection for binding into selection for catalysis, Jacques Fastrez (UCL, Louvain-la-Neuve) presented two strategies of enriching enzyme-displayed phages for catalytic activity, either indirectly by panning on suicide inhibitors, or directly by "catalytic elution", selecting for allosteric sites.[20] Sandro Cesaro-Tadic (ETH Zürich) elaborated on a turnover-based selection of phage-displayed antibodies whose strength was exemplified by the selection of an efficient phosphatase antibody ($k_{cat}/k_{uncat} = 2 \times 10^5$) starting from a naive library and a tenfold further improvement after randomisation.[21] Phage display was also the tool used by Kai Johnsson (ETH Lausanne), who reported how directed evolution can provide reagents that can perform tasks that open the door to

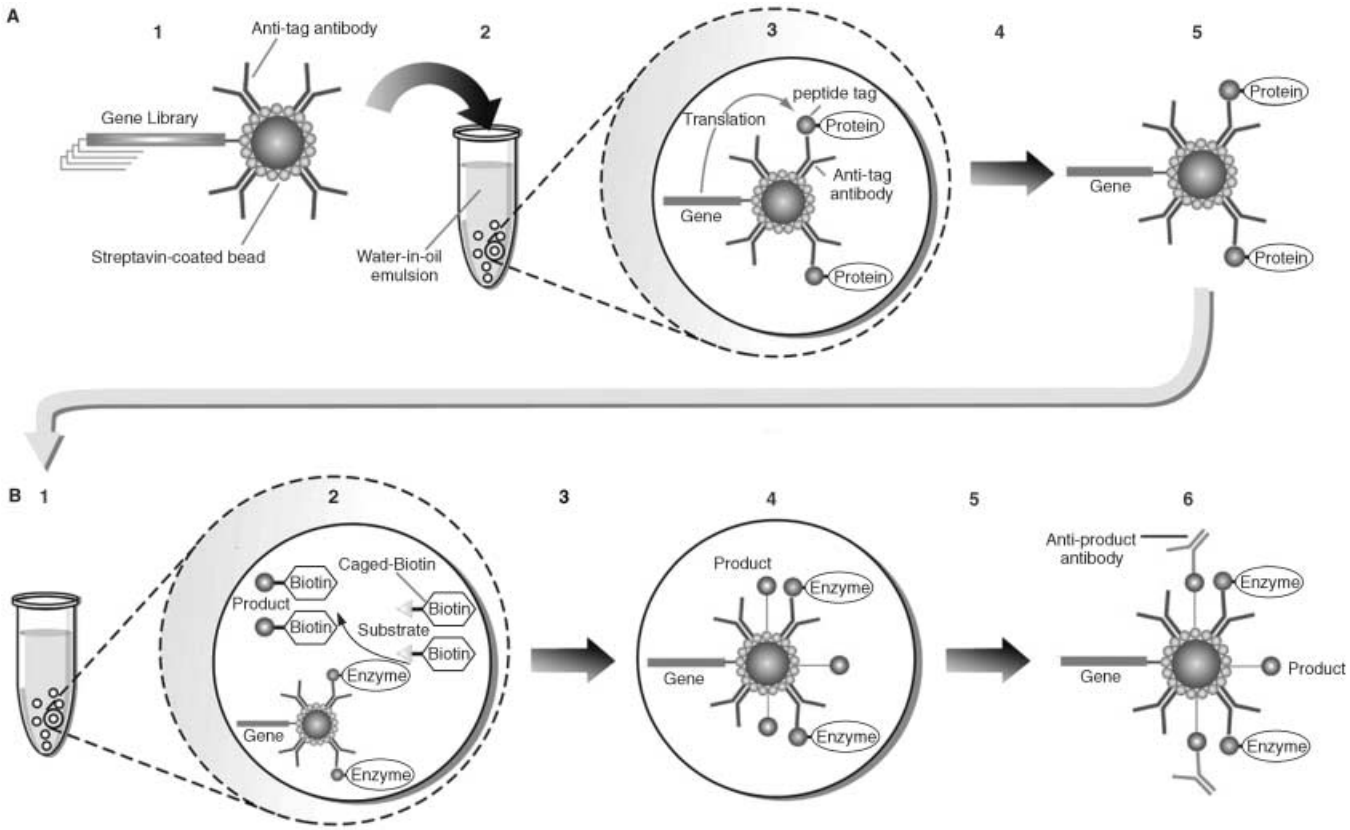

Figure 2. *A new on-bead display system that allows selection of a diffusion-controlled phosphotriesterase in water-in-oil microcompartments.*[27]

solving a biological problem, namely specific protein labelling to trace the cellular fate of individual proteins.[22] Johnsson evolved the methyl transferase hAGT that normally repairs alkylated guanine in DNA to react with the cell-permeable nucleobase O6-benzylguanine (substituted at the 4-position with fluorescein).[23] This leaves hAGT covalently labelled with a fluorescent marker. Indeed, N or C terminus protein fusions of the evolved hAGT were shown to be specifically labelled in vivo; this enables the localisation of specific proteins to be studied.[24]

Andrew Griffiths (MRC, Cambridge, UK) presented cutting-edge technology of in vitro selection by compartmentalisation, using $\sim 1\,\mu m$ droplets of water-in-oil emulsions. This universal in vitro system, developed jointly with Dan Tawfik (Weizmann Institute, Rehovot),[25, 26] may now be practically applied to a wide range of reactions under various conditions (e.g. at 95 °C). It is based on decoupling the transcription/translation step within the emulsion (by using gene-coupled beads (Figure 2), from the enzymatic reaction (in a second emulsion), followed by FACS selection of positive beads that carry the reaction product. As proof of principle, a phosphotriesterase with a very fast k_{cat} of $> 10^5\ s^{-1}$ (60 times higher than the wild-type enzyme) was selected.[27] Pim Stemmer of MAXYGEN has written about this technology: "This is how enzymes will be made a decade from now".[28]

Whatever technique will be used in the future, this meeting poignantly made the point that directed evolution is here to stay and becoming more and more important as well as more and more popular. The conference ended with the angelic tones of Welsh folk tunes, accompanied by harp music. At last something familiar! The same singer appears every year with much success, albeit no less exciting than the novel insights into one aspect of the future of chemistry and biology.

[1] C. Lad, N. H. Williams, R. Wolfenden, *Proc. Natl. Acad. Sci. USA* **2003**.
[2] J. S. Finer-Moore, D. V. Santi, R. M. Stroud, *Biochemistry* **2003**, *42*, 248.
[3] D. L. Birdsall, J. Finer-Moore, R. M. Stroud, *Protein Eng.* **2003**, *16*, 229.
[4] J. L. Hansen, J. A. Ippolito, N. Ban, P. Nissen, P. B. Moore, T. A. Steitz, *Mol. Cell* **2002**, *10*, 117.
[5] P. Nissen, J. A. Ippolito, N. Ban, P. B. Moore, T. A. Steitz, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 4899.
[6] P. Nissen, J. Hansen, N. Ban, P. B. Moore, T. A. Steitz, *Science* **2000**, *289*, 920.
[7] N. Ban, P. Nissen, J. Hansen, P. B. Moore, T. A. Steitz, *Science* **2000**, *289*, 905.
[8] K. Shah, K. M. Shokat, *Methods Mol. Biol.* **2003**, *233*, 253.
[9] K. M. Specht, K. M. Shokat, *Curr. Opin. Cell Biol.* **2002**, *14*, 155.
[10] K. Shokat, M. Velleca, *Drug Discovery Today* **2002**, *7*, 872.
[11] S. M. Ulrich, D. M. Kenski, K. M. Shokat, *Biochemistry* **2003**, *42*, 7915.
[12] A. Denzel, K. J. Hare, C. Zhang, K. Shokat, E. J. Jenkinson, G. Anderson, A. Hayday, *J. Immunol.* **2003**, *171*, 519.
[13] M. J. Ondrechen, J. G. Clifton, D. Ringe, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 12473.
[14] M. M. Meyer, J. J. Silberg, C. A. Voigt, J. B. Endelman, S. L. Mayo, Z. G. Wang, F. H. Arnold, *Protein Sci.* **2003**, *12*, 1686.
[15] W. Winkler, A. Nahvi, R. R. Breaker, *Nature* **2002**, *419*, 952.
[16] W. C. Winkler, A. Nahvi, N. Sudarsan, J. E. Barrick, R. R. Breaker, *Nat. Struct. Biol.* **2003**, *10*, 701.
[17] W. C. Winkler, S. Cohen-Chalamish, R. R. Breaker, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 15908.
[18] G. J. Williams, S. Domann, A. Nelson, A. Berry, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 3143.
[19] V. Maggiotti, M. Resmini, V. Gouverneur, *Angew. Chem. Int.* **2002**, *114*, 1054; *Angew. Chem. Int. Ed. Engl.* **2002**, *41*, 1012.
[20] I. Ponsard, M. Galleni, P. Soumillion, J. Fastrez, *ChemBioChem* **2001**, *2*, 253.
[21] P. O'Brian, Ph.D. thesis, University of Cambridge (UK), **1995**.
[22] N. Johnsson, K. Johnsson, *ChemBioChem* **2003**, *4*, 803.
[23] A. Juillerat, T. Gronemeyer, A. Keppler, S. Gendreizig, H. Pick, H. Vogel, K. Johnsson, *Chem. Biol.* **2003**, *10*, 313.
[24] A. Keppler, S. Gendreizig, T. Gronemeyer, H. Pick, H. Vogel, K. Johnsson, *Nat. Biotechnol.* **2003**, *21*, 86.
[25] A. D. Griffiths, D. S. Tawfik, *Curr. Opin. Biotechnol.* **2000**, *11*, 338.
[26] D. S. Tawfik, A. D. Griffiths, *Nat. Biotechnol.* **1998**, *16*, 652.
[27] A. D. Griffiths, D. S. Tawfik, *EMBO J.* **2003**, *22*, 24.
[28] J. Minshull, W. P. Stemmer, *Curr. Opin. Chem. Biol.* **1999**, *3*, 284.
[29] D. A. Liberles, D. R. Schreiber, S. Govindarajan, S. G. Chamberlin, S. A. Brenner, *Genome Biology* **2001**, *2*, research0028.1-0028.6.
[30] E. A. Gaucher, J. M. Thomson, M. F. Burgan, S. A. Benner, *Nature*, **2003**, *425*, 285.

Ariel Lindner
Laboratoire de Génétique Moléculaire Evolutive et Médicale
Université René Descartes – Paris V (France)
Florian Hollfelder
Department of Biochemistry
University of Cambridge (UK)

