OBC
www.rsc.org/obc

Directed evolution of enzymes: new biocatalysts for asymmetric synthesis

Marina Alexeeva, Reuben Carr and Nicholas J. Turner *
School of Chemistry, University of Edinburgh, King's Buildings, West Mains Road, Edinburgh, UK EH9 3JJ. E-mail: n.j.turner@ed.ac.uk

Received 10th September 2003, Accepted 2nd October 2003
First published as an Advance Article on the web 21st October 2003

Directed evolution has been employed to generate new enzymes for the deracemisation of chiral amines.

Introduction

The use of combinatorial methods for lead discovery and optimisation has become a powerful and now firmly established technique within the pharmaceutical industry. More recently, analogous approaches of parallel synthesis and screening have been applied to the discovery and development of new catalysts for asymmetric synthesis.[1] The underlying principles are simple and very appealing, namely coupling systematic variations in catalyst structure, ligand composition, reaction conditions *etc*., with effective methods for high-throughput screening for a range of properties (*e.g.* turnover number, enantioselectivity) resulting in catalysts that are fine tuned and optimised for specific applications.[2] Not surprisingly, similar concepts of combinatorial optimisation have recently gained popularity in the discovery and development of biocatalysts. However, here the situation is somewhat different given the relative complexity of enzymes. A typical enzyme is composed of between 250–1000 amino acids and hence the notion of systematically varying even a small number of amino acids in an enzyme sequence, to explore the effect on activity or selectivity, is unrealistic since the number of possible permutations rapidly exceeds the capabilities of high-throughput screens, let alone conventional analytical HPLC or GC-MS methods. Even if only one amino acid in any part of the sequence is replaced by any other of the 19 other amino acids the number of permutations[3] is in the region of *ca.* 10^4 whereas for combinatorial mutagenesis, wherein two or more residues are simultaneously randomised, the library sizes can become truly vast (no. of permutations = 20^n where n is the number of amino acids that are simultaneously varied). Furthermore, it is generally not obvious which amino acids, or combinations thereof, are important for controlling specific characteristics (*e.g.* enantioselectivity, catalytic activity, stability). Although the increased availability of X-ray structures can help to guide identification of important residues, in practice it is mainly active-site residues that are targeted on the

Marina Alexeeva, graduated from the Moscow State Pirogov Medical Institute in 1980 with a degree in Biochemistry. Following a period as a Research Assistant in the Institute of Molecular Genetics at the Russian Academy of Sciences, she joined Edinburgh University in 1992 as a Technical Officer, a position she held for 7 years until beginning her PhD in 1999 with Nick Turner. Her thesis research involved the directed evolution of amine oxidases for deracemisation reactions. At the beginning of 2003 she joined Ingenza as a Senior Scientist.

Reuben Carr, graduated with a first class honours degree in Chemistry and Biochemistry from the University of Southampton in 2000. He moved to Edinburgh in 2001 to begin his PhD under the supervision of Nick Turner as a CASE student in collaboration with GlaxoSmithKline. His research involves the deracemisation of chiral amines using enzymes in combination with chemical reducing agents.

Nicholas Turner, obtained his D. Phil. in 1985 under the supervision of Professor Sir Jack Baldwin, followed by postdoctoral studies as a Royal Society Junior Research Fellow with Professor George Whitesides at Harvard University, USA. In 1987 he joined the Chemistry Department at Exeter University as a Lecturer before moving to Edinburgh University in 1995 as a Reader. In 1998 he was promoted to a Personal Chair in Biological Chemistry. Recently, along with his colleagues Ian Fotheringham, Robert Speight and Sunil Srivastava, he founded Ingenza, a new company based within the School of Chemistry at Edinburgh, that is developing novel bioprocesses for the pharmaceutical and fine chemical industry.

Marina Alexeeva

Reuben Carr

Nicholas Turner

DOI: 10.1039/b311055a

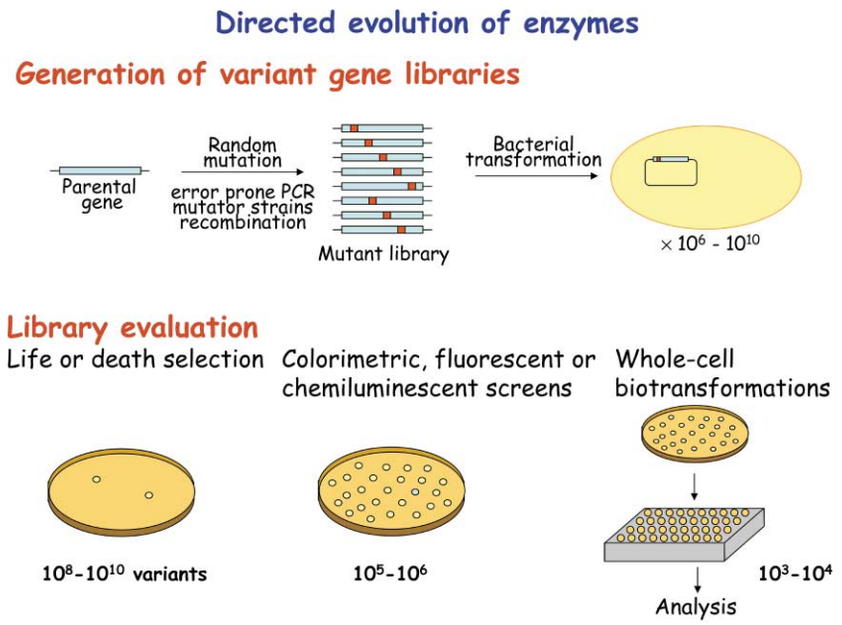

Fig. 1 Strategies for the directed evolution of enzymes involving (a) generation of libraries of variant genes and (b) high-throughput screening of the libraries using different analytical methods.

basis that they make important non-bonding interactions with the substrate.[4]

The past 5–10 years has witnessed the emergence of 'directed evolution' strategies in order to change the properties of enzymes in a targeted manner.[5] Directed evolution of enzymes combines two distinct events, namely (i) random mutagenesis of the parent gene to provide a library of variant enzymes and (ii) methods for screening the individual members of the library for specific characteristics (*e.g.* reactivity towards a specific substrate, resistance to product inhibition, thermal/organic solvent stability) (Fig. 1). The power of this approach is that it requires no prior knowledge of the structure (or in some cases sequence) of the enzyme since the variants are selected simply on their ability to carry out a specific transformation, without regard to prior assumptions as to which particular amino acids might be important. The real challenge in this area is to devise high-throughput screens that are able to handle sufficiently large libraries (*ca.* 10^5–10^9) such that significant improvements over the activity of the wild-type enzyme can be achieved. Such levels of throughput (*i.e.* $>10^5$ samples per day) have been achieved using *in vivo* 'life or death' selection and *in vitro* detection (*e.g.* colorimetric/fluorescence detection, capillary array electrophoresis) but new methods are required to expand the range of enzyme activities that can be detected.[6]

The aim of this article is to provide an insight into the field of 'directed evolution' and to consider how this technology can best be employed to develop new enzyme-catalysed reactions and processes for the future. To assist in this objective, the authors have provided a personal case study from their own laboratory in which many of the current themes surrounding directed evolution have been encountered.

Deracemisation reactions: the development of a platform technology

Inspired by the early work of Hafner *et al.*,[7] and Soda *et al.*,[8] we set out to develop general and practical methods for the cyclic deracemisation of chiral compounds *via* an oxidation–reduction sequence (Fig. 2). In this catalytic cycle the combination of an enantioselective oxidase, together with a non-selective chemical reducing agent, results in the conversion of a racemic mixture to a single enantiomer. Provided that the oxidase enzyme is highly enantioselective, only 7 cycles are required to achieve an ee of >99%.[9] We have successfully applied this method to the deracemisation of a variety of

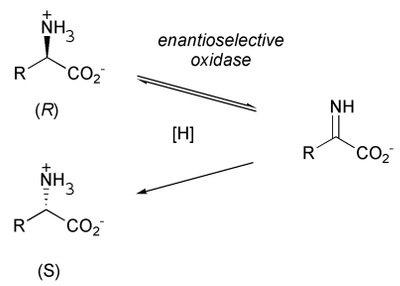

Fig. 2 Deracemisation of α-amino acids using an enantioselective amino acid oxidase in combination with a non-selective chemical reducing agent.

α-amino acids in high yield (70–85%) and ee (>98%) using both D- and L-amino acid oxidases together with a variety of reducing agents ($NaCNBH_3$; $NaBH_4$; amine–boranes: Pd/C–ammonium formate).[10,11] Because the reaction is in essence a stereoinversion process, we have also shown that it can be used to interconvert diastereomeric amino acids bearing more than one stereogenic centre.[12] Encouraged by this success with α-amino acids, we sought to broaden the range of substrates amenable to deracemisation and considered amines in view of their increasing value as chiral intermediates for pharmaceutical drugs and also ligands for asymmetric catalysis.[13] The key issue at the outset was the availability of enantioselective amine oxidases that could be applied to deracemisation reactions. Unfortunately, we could not find any convincing literature precedent for the presence of enantioselective Type II amine oxidases in microbial sources (some Type I enzymes were known to be enantioselective but these amine oxidases generate an enzyme-bound imine and hence were unsuitable for use in deracemisation reactions). Thus we decided to undertake a directed evolution approach to evolve enzymes that had the characteristics required for our deracemisation reactions.

Directed evolution of an amine oxidase

The starting point for any successful directed evolution project is a reliable and robust high-throughput screen. *In vitro* screens, based upon the production of a coloured or fluorescent product, have the advantage that they can be used to screen colonies directly on an agar plate thus enabling 10^5–10^6 clones to be assayed in a short period of time. We developed a colorimetric screen, based upon capture of the hydrogen peroxide generated

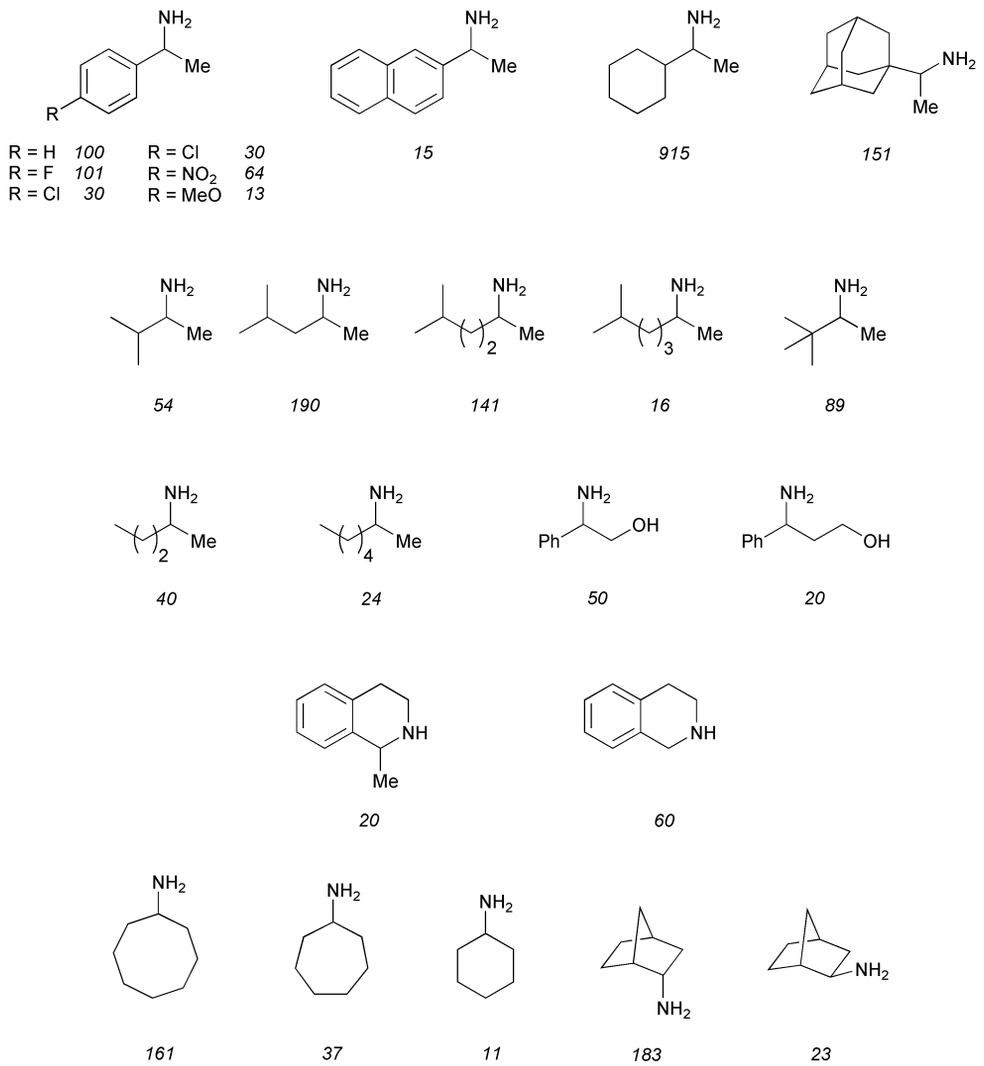

Fig. 3 Range of amines that are substrates for the Asn336Ser variant of the amine oxidase from *Aspergillus niger.*, The numbers in italics below the structures indicate the activity relative to α-methylbenzyl amine.

by the oxidase enzyme, and demonstrated that it could be used to detect colonies expressing amino acid or amine oxidases. The next decision concerned the starting gene to be used for generation of the library of variants. It was known that the monoamine oxidase (MAO-N) from *Aspergillus niger* catalysed the oxidation of simple alkyl amines including benzylamine although no data was reported for any chiral amines.[14] A plasmid harbouring the MAO-N gene was obtained, subcloned into *Escherichia coli*, and used to express the wild-type enzyme which was found to have very low, but detectable activity towards α-methylbenzylamine, the amine chosen for our model studies. Interestingly there was also evidence that the wild-type enzyme was enantioselective although the intrinsic rates were very low. Next the MAO-N gene was randomly mutated, using the *E. coli* XL1-Red mutator strain,[15] and the library of variants (*ca.* 150 000) screened against (*S*)-α-methylbenzylamine as a substrate using the agar-plate colorimetric assay. The frequency of mutation was adjusted such that individual clones possessed *ca.* 1–2 nucleotide mutations per gene. Approximately 30 clones were selected at this stage, on the basis of their activity in the screen, and amongst these one clone in particular was found to have very high (*S*)-selectivity. Small scale growth of this clone and partial purification of the amine oxidase revealed that this variant possessed approximately 47 fold greater activity, and 6 fold greater enantioselectivity towards α-methylbenzylamine compared with the wild-type enzyme, a substantial improvement from a single round of directed evolution. This variant was sufficiently catalytically active to be used in combination with ammonia–borane to deracemise α-methylbenzylamine yielding the (*R*)-enantiomer in 77% yield and 93% ee.[16]

Sequencing of this variant revealed that it possessed a single mutation (Asn336Ser). The variant amine oxidase was subsequently purified to homogeneity, by Ni affinity chromatography, and examined for its reactivity towards a panel of *ca.* 50 different chiral amines. Interestingly, more than 60% of these structurally different amines were found to be substrates, some with greater activity than *a*-methylbenzylamine itself (Fig. 3). For those substrates where the individual enantiomers were available we determined the enantioselectivity and found that in all cases the enzyme was highly (*S*)-selective ($E > 10$). By contrast, the wild-type enzyme was active towards only *ca.* 15% of the substrates, confirming that the Asn336Ser variant possessed a substantially different substrate specificity.[17] Analysis of the relationship between structure and reactivity led us to propose a simple model to summarise the effect of directed evolution of the wild-type amine oxidase (Fig. 4). Overall, this first round of directed evolution had resulted in the identification of a variant with substantially improved catalytic activity and enantioselectivity together with significantly broadened substrate specificity.

Further rounds of 'real time' directed evolution

An attractive aspect of directed evolution is that variants selected from the first round of evolution can form the starting

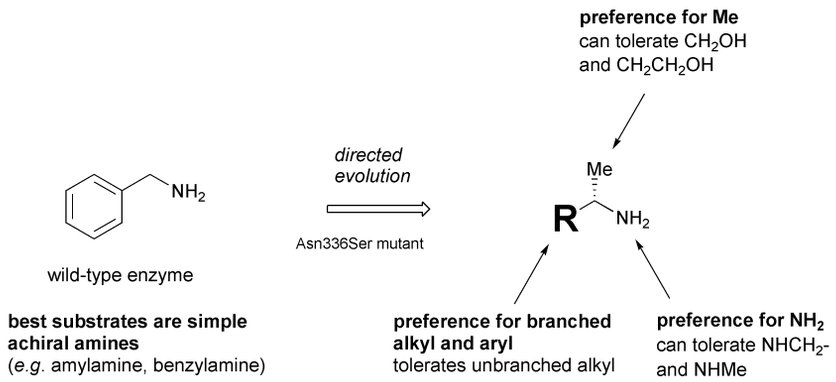

Fig. 4 Schematic illustration of the change in substrate of specificity of the Asn336Ser mutant relative to the wild-type enzyme.

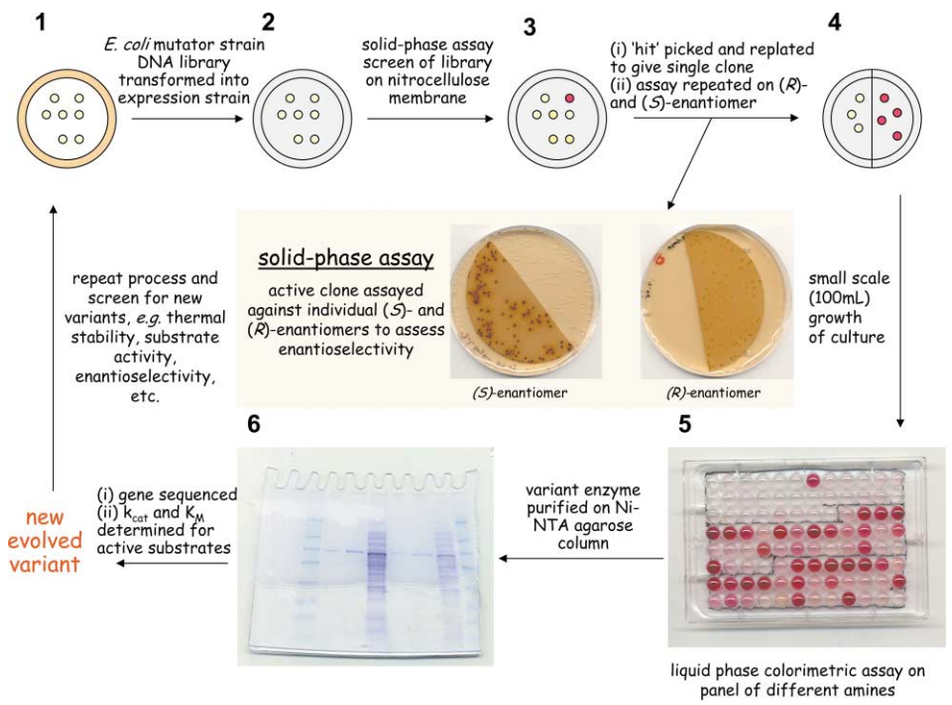

Fig. 5 Strategy for iterative evolution of enzymes in 'real time' by employing hierarchical screening protocols; *Step 1*: Library of variants generated from parent gene using *E. coli*, XL1-Red mutator strain. *Step 2*: DNA library transformed into *E. coli*, expression strain and colonies streaked and grown on nitrocellulose membrane on agar plate. *Step 3*: Membrane lifted, colonies partially lysed (freeze–thaw) and then screened for activity using colorimetric assay containing specific amine substrate. *Step 4*: 'Hits' picked and replated to give single colonies. Active colonies checked for enantioselectivity towards individual (*R*)- and (*S*)-enantiomers of amine substrate. *Step 5*: Variants that are both highly active and enantioselective grown on small scale (100 ml) and screened against range of chiral amines (*ca.*, 32) in 96-well microtitre plate using liquid-phase colorimetric assay. This step allows differentiation between 'expression mutants' and 'specificity mutants'. *Step 6*: Individual variants purified using Ni-affinity column and subjected to full kinetic characterisation (k_{cat} and K_M). Gene sequenced to identify positions of mutations. *Step 7*: Steps 1–6 repeated varying target amine substrate or introducing additional selection criteria (*e.g.*, thermostability, high substrate concentration).

point for further rounds of mutagenesis/selection in the anticipation than additional mutations might lead to further changes in an additive manner. These further cycles can be carried out rapidly (3–4 weeks), in a highly automated fashion, in order to fine tune the enzyme towards substrates of specific interest. Fig.5 illustrates the approach we have recently developed involving initial passage through the mutator strain, to generate new libraries (*ca.* 10000–300000 clones) which are then screened for specific characteristics including substrate specificity, thermal stability, enantioselectivity *etc*. Libraries can also be generated at this stage using error-prone PCR or combinatorial mutagenesis in which several amino acids are simultaneously randomised. Any identified 'hits', which typically arise with a frequency of *ca.* 1 in 10000 using the mutator strain, are rapidly checked against a panel of substrates (*ca.* 32 structurally different compounds) in order to roughly define the substrate range and enantioselectivity. At this stage we are also able to identify 'expression mutants' *i.e.* those variants in which the protein expression levels have been enhanced without any change in specific activity or substrate specificity. Variants of particular interest are subsequently purified to homogeneity, *via* Ni-chelation chromatography, and then fully characterised against specific chiral amines of interest including determination of k_{cat} and K_M values. Further rounds of evolution can then be carried out until an enzyme possessing suitable characteristics is obtained. In this way we have been able to evolve new types of amine and amino acid oxidases, *i.e.* enzymes that possess greater activity towards secondary, rather than primary, amines or variants that are able to oxidise sterically demanding substrates. The goal is to rapidly generate a family of (*S*)- and (*R*)-selective amine and amino acid oxidases that possess complementary substrate specificities allow us to deracemise a wide range of chiral amines and amino acids respectively.

Future directions

Directed evolution represents a powerful strategy for altering the properties of an enzyme in a rapid and targeted manner. However, the challenge now for applying this technology to biocatalysis is to identify the best strategy for evolving enzymes

such that they ultimately possess all of the characteristics required for ultimate use in a large scale bioprocess.[18] The ideal biocatalyst should possess most, if not all of the following properties; high k_{cat}, K_M in the millimolar range, high enantioselectivity ($E > 50$), broad substrate specificity, high tolerance towards organic solvents and stability at elevated temperatures. Each of these characteristics can probably be screened for by adjusting the assay conditions and in principle it may be possible to select for more than one property in a single step.[19] For example in our experience, based upon the directed evolution of amine and amino acid oxidases, variants selected primarily for their activity in the solid-phase assay against a single enantiomer substrate, often turn out to be also highly enantioselective. Since high catalytic activity is a fundamental prerequisite for any biocatalyst, this observation is guiding our thinking in terms of the way forward for the directed evolution of enzymes. Perhaps the most exciting aspect of this area is that directed evolution can now be carried out in 'real time', at a rate that is commensurate with the expectations of those wishing to develop and optimise enzymes for applications in the commercial synthesis of fine chemicals and chiral intermediates.[20] Undoubtedly, as we learn more about the best strategies for evolving an enzyme in a particular direction, this process of optimisation will become faster, more automated, and hence increasingly attractive to synthetic organic chemists.

References

1 For recent reviews see: J. P. Stambuli and J. F. Hartwig, *Curr. Opin. Chem. Biol.*, 2003, **7**, 7–420; J. F. Traverse and M. L. Snapper, *Drug Discovery Today*, 2002, **7**, 1002.

2 D. B. Berkowitz, M. Bose and S. Choi, *Angew. Chem., Int. Ed.*, 2002, **41**, 1603; M. Müller, T. W. Mathers and A. P. Davis, *Angew. Chem., Int. Ed.*, 2001, **40**, 3813.

3 The number of permutations (N) of an enzyme, containing X amino acids, in which a single amino acid is replaced by any one of the remaining 19 amino acids is given by the equation: $N = 19X!/[(X - 1)!]$.

4 G. P. Horsman, A. M. F. Liu, E. Henke, U. T. Bornscheuer and R. J. Kazlauskas, *Chem. Eur. J.*, 2003, **9**, 1933.

5 For recent reviews see: *Directed Molecular Evolution of Proteins*, eds. S. Brakmann and K. Johnsson, Wiley VCH, Weinheim, 2002; E. T. Farinas, T. Bulter and F. H. Arnold, *Curr. Opin. Biotech.*, 2001, **12**, 545; J. D. Sutherland, *Curr. Opin. Chem. Biol.*, 2000, **4**, 263.

6 M. T. Reetz, *Angew Chem, Int Ed*, 2002, **41**, 1335.

7 E. W. Hafner and D. Wellner, *Proc. Nat. Acad. Sci.*, 1971, **68**, 987.

8 J. W. Huh, K. Yokoigawa, N. Esaki and K. Soda, *J. Ferment., Bioeng.*, 1992, **74**, 189; J. W. Huh, K. Yokoigawa, N. Esaki and K. Soda, *Biosci. Biotech. Biochem.*, 1992, **56**, 2081.

9 K. Faber, *Chem. Eur. J.*, 2001, **7**, 5005.

10 T. Beard and N. J. Turner, *Chem. Commun*, 2002, 246.

11 F.-R. Alexandre, D. P. Pantaleone, P. P. Taylor, I. G. Fotheringham, D. J. Ager and N. J. Turner, *Tetrahedron Lett.*, 2002, **43**, 707.

12 A. Enright, F.-R. Alexandre, G. Roff, I. G. Fotheringham, M. J. Dawson and N. J. Turner, *Chem. Commun.*, 2003, 2636.

13 M. Berger, B. Albrecht, A. Berces, P. Ettmayer, W. Neruda and M. Woisetschläger, *J. Med. Chem.*, 2001, **44**, 3031; S. France, D. J. Guerin, S. J. Miller and T. Lectka, *Chem. Rev.*, 2003, **103**, 2985.

14 S. O. Sablin, V. Yankovskaya, S. Bernard, C. N. Cronin and T. P. Singer, *Eur. J. Biochem.*, 1998, **253**, 270.

15 A. Greener, M. Callahan and B. Jerpseth, *Mol. Biotechnol.*, 1997, **7**, 189.

16 M. Alexeeva, A. Enright, M. J. Dawson, M. Mahmoudian and N. J. Turner, *Angew. Chem., Int. Ed.*, 2002, **41**, 3317.

17 R. Carr, M. Alexeeva, A. Enright, T. S. C. Eve, M. J. Dawson and N. J. Turner, *Angew. Chem., Int. Ed.*, 2003, **42**, 4807.

18 N. J. Turner, *TIBTECH*, 2003, **21**, 474.

19 D. R. Yazbeck, J. Tao, C. A. Martinez, B. J. Kline and S. Hu, *Adv. Synth. Catal.*, 2003, **345**, 524.

20 G. Huisman and D. Gray, *Curr. Opin. Biotech.*, 2002, **13**, 352; R. N. Patel, *Curr. Opin. Biotech.*, 2001, **12**, 587.