Burton, S., Cowan, D.A., Woodley, J.M. (2002). The search for the ideal biocatalyst. NATURE BIOTECHNOLOGY. 20:35-46

The search for the ideal biocatalyst

Stephanie G Burton, Don A Cowan, and John M Woodley

Abstract

While the use of enzymes as biocatalysts to assist in the industrial manufacture of fine chemicals and pharmaceuticals has enormous potential, application is frequently limited by evolution-led catalyst traits. The advent of designer biocatalysts, produced by informed selection and mutation through recombinant DNA technology, enables production of process-compatible enzymes. However, to fully realize the potential of designer enzymes in industrial applications, it will be necessary to tailor catalyst properties so that they are optimal not only for a given reaction but also in the context of the industrial process in which the enzyme is applied.

The past two decades have led to major advances in our understanding of the subtleties of protein structure-function interrelationships. Mechanisms of protein stability in aqueous and nonaqueous environments[1,2], the links between conformational mobility, structural integrity and activity[3,4], and the complexities of substrate specificity have all succumbed to the onslaught of advanced molecular methods, including crystallography[5-7], site-specific mutagenesis, gene shuffling, and protein evolution[8-11]. Scientists are now in a position to visualize, if not design, catalytic systems that approach the functional "ideal".

The "ideal" catalyst is typically considered by the biochemist in terms of turnover number (k_{cat}) or, for a given process, in terms of maximum specificity constant (A_{cat}/KM). However, from a bioprocess viewpoint, each bioprocess is constrained by a set of conditions dictated by the specific properties of the substrates, products, and the bioconversion reaction. Thus, while for any given bioprocess it is clearly possible to specify a set of catalyst properties that would constitute the ideal for that process, only broad generalizations for ideality may otherwise be identified.

In this review, we discuss molecular properties of enzymes from a bioprocess viewpoint. A paradigm for design based on the ideal process determining the desired features of the catalyst is presented. "Ideal" characteristics (generic and process-specific) and methodologies for seeking or engineering these are discussed. We then review the extent to which significant changes in protein functional properties have been achieved by application of these methods, together with issues that require more research and/or development.

Ideal processes

Bioprocess engineers develop processes not only against fixed capital and operating expenditure, but also rapid, and often, very aggressive timelines. This is particularly the case for pharmaceutical products where patent expiry on the product demands a robust and scalable process with short development times[12]. Conventionally, once a synthetic route is fixed and the biocatalytic step defined, bioprocess design and operation are centered on the properties of the

biocata-lyst. There are two reasons for this. First, biocatalysts are typically characterized by tight operational specifications. Second, the biocat-alyst can often represent a significant proportion of the process operating costs. Increasingly, with more expensive substrates and improved biocatalyst production this contributing proportion is being reduced, and with expansion of the range of suitable biocata-lysts, it may become desirable as well as possible to center bioprocess design and operation on the reaction properties, rather than the bio-catalyst (see Fig. 1). Thus, the conventional paradigm can be reversed and the biocatalyst can be designed to fit the process. This alternative approach leads us to ask: what are the ideal properties we would want of a "designer" biocatalyst, and how would we identify these ideal properties?

The flexibility to design the biocatalyst to fit the process still carries a number of catalyst-independent constraints, set by the thermodynamics of the reaction and the properties of reactants and products (Table 1). In particular, the water solubility characteristics of the reactants and products determine the medium for operation[13], and their stability as a function of pH and temperature determine the range in which the process must operate. In addition, high product concentrations are important in optimizing product recovery.

Under the ideal process paradigm (Fig. 1), preselection of optimal process conditions can supersede the constraints of any particular biocatalyst. A range of technologies contributes to the preselection process (Fig. 2). Having selected the "ideal" process conditions, a growing range of molecular technologies are available for identification or engineering of the "ideal" biocatalyst.

Definitions

$\mathbf{T_m}$ Protein melting (denaturation) temperature, typically determined by physical methods such as fluorescence, circular dichroism (CD) or calorimetry.
$\mathbf{T_{opt}}$ Apparent temperature optimum for enzyme activity, for a defined set of reaction conditions.
$\mathbf{k_{cat}}$ Turnover number = maximum number of substrate molecules converted to product per unit time per active site; a function of the rate constants for the conversion of enzyme-substrate complex to enzyme + product.
$\boldsymbol{K_M}$ Michaelis constant, representative of the dissociation of the enzyme-substrate complex, taken as an indication of substrate binding to the active site.
$\boldsymbol{k_c/K_M}$ Specificity constant; indicates the rate of association of enzyme and substrate and also indicates specificity for competing substrates.

Catalyst production

The key to the process architecture defined in a particular case is whether the catalyst is to be used as a whole cell or an isolated enzyme. For the majority of enzyme-catalyzed reactions, use of an isolated (immobilized) enzyme is the ideal provided that isolation is simple. However, for membrane-bound and/or cofactor-dependent enzymes and/or those that are multiprotein

complexes, isolation may not be possible, and here the alternative whole-cell process will prove the preferred option. It is beyond the scope of this review to consider in depth the presentation of the enzyme within a reaction system and the very important aspects of free versus cell-encapsulated biocatalysts, but the reader is directed to a number of relevant publications[14-16].

Whether a whole-cell or an isolated-enzyme biocatalyst is required, expression in a rapidly growing, robust host is the first requirement. Constitutive, rather than inducible, expression is preferred. High levels of expression will minimize the fermentation volume, but may lead to inclusion body formation. Developments in host-vector systems, coupled with highly evolved fermentation technologies, have resulted in bacterial and fungal fermentations capable of yielding recombinant enzymes at levels of upward of 10 g/L, in stark comparison with laboratory-based expression of recombinant enzymes, where yields of 10-100 mg/L are frequently reported. In all cases, the cost of media imposes a need to minimize fermentation volumes by high levels of expression. Protease activities, which can result in substantial loss of biocatalyst yield, also need to be minimized.

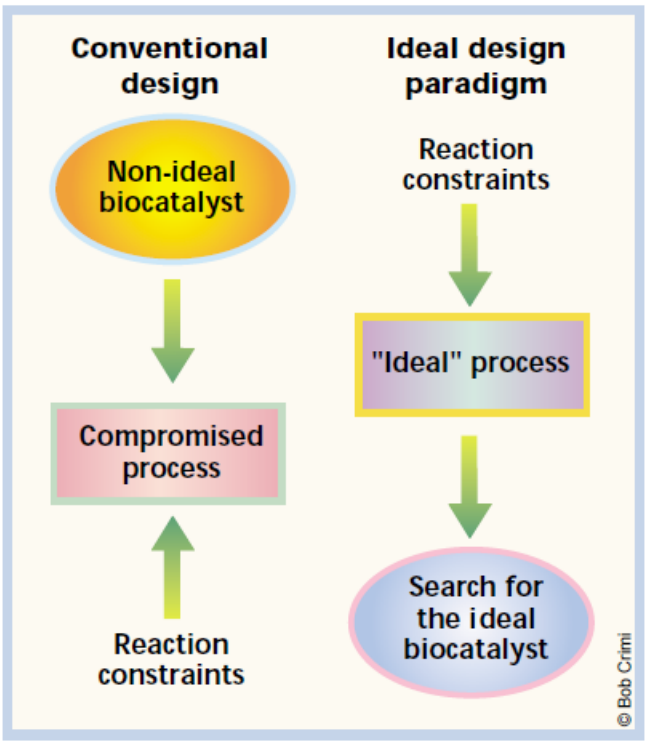

Figure 1. The paradigm shift? Traditionally, the characteristics of the enzyme have constrained the operational limits of the process. New approaches to enzyme discovery offer the prospect of designing the biocatalyst to suit the optimal parameters of the process.

Reactors

Ideal reactors are those with high space/time yield to reduce the capital costs. A packed bed of immobilized enzyme has clear advantages in that the voidage is low: 34% compared with up to

over 80-90% for a stirred-tank reactor[17]. The hydrodynamics of a packed-bed reactor, characterized by plug flow, have a number of implications, particularly in the generation of concentration gradients of substrate and product. As a consequence, both low enzyme K_M and substrate inhibition are critical, so as to capitalize on the concentration gradient generated from column input to output.

Similarly, pH adjustment is not possible during a single pass through a packed bed, so it is essential that the pH/activity profile for the enzyme is sufficiently broad to accommodate any pH changes. Critically, multiphasic mixtures cannot be handled in a packed bed, although a fluidized bed is a reasonable compromise (S. Beveridge *etal.,* unpublished data).

Table 1. Constraints influencing the operating conditions of bioconversion processes

Type of constraint	Description
Reaction-specific constraints dependent on the nature of reactants and products	Water solubility of reactants and products (which is dependent on pH and temperature)
	Stability of reactants and products
	Chemical nature of reactants and products
	Production or consumption of gases
	Production or consumption of acid or alkali
	Reaction equilibrium (which is in turn dependent on pH and temperature)
Biocatalyst-specific constraints dependent on the evolved or modified properties of the biocatalyst	The pH range over which optimal activity and stability of biocatalyst can be maintained
	Temperature range for biocatalyst optimal activity and stability
	Concentration range of reactants/products that can be tolerated by biocatalyst without inhibition or saturation effects

Product recovery

The ideal process involves the minimum number of downstream steps. The catalyst must be easily recovered (immobilization by retention inside the packed bed is optimal). A bottleneck for many current processes is the large volume of water from which the product must be removed, necessitating a concentration step before isolation[18].

In the ideal process, the catalyst would withstand concentrations of the product high enough to allow a single isolation step without prior product concentration.

Key biocatalyst properties

Turnover. Considering that the primary objective of any industrial biotransformation process is a high degree of substrate conversion to product in the minimum possible time, the turnover of an enzyme represents a key factor in the concept of the ideal biocatalyst. As many enzymes have evolved for very specific and selective metabolic roles, the intrinsic activity or turnover of native enzymes depends fundamentally on the nature and mechanism of the *in vivo* reaction catalyzed and the position of *in vivo* reaction equilibria with respect to metabolic flux. Thus, a particular enzyme may catalyze a desired bioconversion reaction, but with a k_{cat} too low for practical application, in which case improvement of the turnover using the novel technologies now

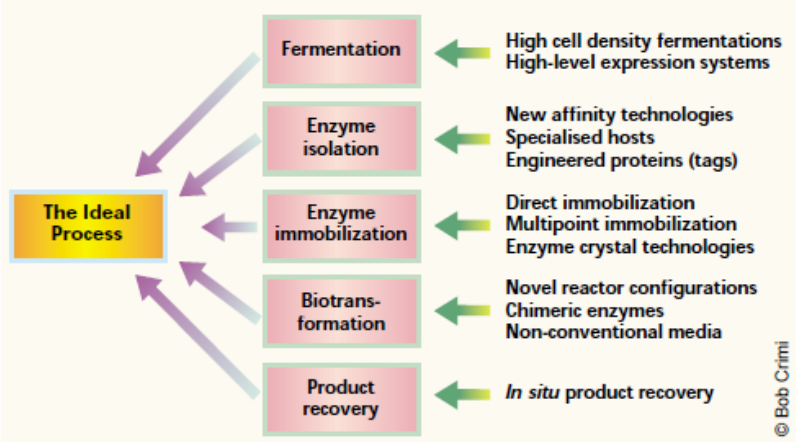

Figure 2. Contributions to the design of the ideal biotransformation process. Each unit operation can be enhanced by the development of key technologies. For example, high cell density fermentations lead to reduced fermentation volumes and simpler dewatering processes.

available (e.g., directed evolution, gene shuffling, and combinations of rational and combinatorial methods) may be required[19]. Although turnover characteristics of enzymes typically used in bioconversions vary widely, high turnovers are characteristic of most current commercial biocatalytic processes.

In an operational biocatalyst, the turnover may be influenced not only by the genetically determined functional capacity of the protein, but also by the microenvironment of the enzyme, and specifically by the form in which the biocatalyst is formulated and applied. Thus process conditions (e.g., the type and use of immobilized enzymes) have considerable influence over the maximum achievable turnover exhibited by the biocatalyst. Such chemical modifications frequently result in lowering of k_{cat} values, attributed to a range of effects, including variations in protein flexibility and allosteric properties (conformational), decreased accessibility to the enzyme active site[20] (steric), and changes in the enzyme microenvironment by the selective attraction or exclusion of water, ions, solutes, and so on[21].

The operating temperature of a biocatalytic process is frequently dictated by factors relating to the reaction rather than the catalyst, such as substrate/product solubilities, reagent stabilities, or the need to reduce undesirable side reactions. Arrhenius kinetics suggest that the highest feasible reaction temperature should be selected, but such selection is always constrained by biocatalyst stability.

Fortunately, evolution has provided a relatively wide thermal range of native biocatalysts, and enzymes derived from psychrotolerant (cold-active) to hyperthermophilic microorganisms may retain high levels of activity and/or stability across a temperature range from below 10°C to

around 100°C, as demonstrated for a set of functionally homologous enzymes in Figure 3. Where the reaction temperature is dictated by the process components, a priori selection of an enzyme from the appropriate thermal group (Fig. 3) may minimize subsequent requirements for enhancing either turnover or molecular stability.

The ideal biocatalyst for bioconversions involving organic and hydrophobic substrates would exhibit high turnover in organic or aqueous-organic media. Many enzymes remain active and stable in organic media, but relatively few exhibit enhanced turnover. Although research using native enzymes has provided insight into the effects of solvents and ionic conditions on enzyme activity[22,23] and control over reaction pathways[24], biocatalysts with significantly improved turnover in organic solvent media have recently been developed (Table 2).

Table 2. Improved kinetic parameters of enzymes modified for biocatalytic activity in organic media

Enzyme and reaction conditions		Kinetic parameters reported				Comments
		V_{max}	k_{cat}	K_M	k_{cat}/K_M	
Subtilisin E and mutants generated by sequential random mutagenesis; ester hydrolysis activity in presence of DMF[50,a]	Native enzyme Aqueous medium 40% DMF	NA[b] NA	21 3.3	0.56 20.9	38 0.16	Variant shows improved k_{cat} in DMF as compared with aqueous medium, and k_{cat}/K_M approaching that of native enzyme in aqueous medium
	Variant PC3 Aqueous medium 40% DMF	NA NA	27 62	0.1 2.96	274 21	
Subtilisin BPN' solubilized by ion pairing; transesterification of APEE[a] and propanol[2]	Aqueous medium Solubilized enzyme in octane medium Suspended enzyme in octane	NA NA NA	NA NA NA	NA NA NA	3,500 370 0.6	Solubilization by ion pairing results in catalytic efficiency that is 60× that of the suspended enzyme and only 10× lower than in aqueous medium
Rhizopus arrhizus lipase-catalyzed esterification of dodecanol and decanoic acid[51]	Water-saturated diisopropyl ether Dry diisopropyl ether	260 240	NA NA	430 33	NA NA	Decreased K_M and similar V_{max} were reported in medium with increased water activity in a hydrophobic solvent system
Mushroom polyphenol oxidase hydroxylation of *p*-cresol[52]	Aqueous solution Chloroform medium	350 812	NA NA	1.98 24.8	NA NA	Improved turnover suggested to be due to increased solubility of products in organic solvent
Subtilisin Carlsberg transesterification of CBZ–alanyl ONP[a] in water-saturated solvents[53]	Toluene Isopropylether Hexane	0.083[c] 0.96[c] 4.36	NA NA NA	NA NA NA	NA NA NA	Significantly increased catalytic efficiency in hexane was attributed to substrate solvation effects
Subtilisin lyophilized in presence of increasing proportions of KCl[54] and applied in anhydrous hexane	Enzyme lyophilized in absence of KCl Enzyme lyophilized in presence of 98% KCl	NA NA	0.027 10.04	260 26.7	0.104 390	Salt-induced activation of enzymes is attributed to protection of the enzyme from solvent deactivation, retention of the active site structure, and retention of water in the enzyme structure
α-Chymotrypsin lyophilized in presence of increasing proportions of KCl[54]	Enzyme lyophilized in absence of KCl Enzyme lyophilized in presence of 98% KCl	NA NA	4×10^{-4} 220×10^{-4}	33 33	0.013 0.67	
Subtilisin transesterification of APEE in presence and absence of salts, in hexane medium preparation	Salt-free enzyme medium[55] Enzyme prepared by lyophilization with 98% KCl	NA NA	0.0069 2.16	8.18 2.24	0.84 966	

[a]Abbreviations: APEE, acetyl pheynl alanyl ethyl ester; DMF, dimethylformamide; ONP, o-nitrophenol.
[b]NA, Not available.
[c]V_{max}/K_M (ml/g/min). Unless otherwise stated, units are as follows: V_{max} (μmol/min/g); k_{cat} (s^{-1}); K_M (mM); k_{cat}/K_M ($M^{-1}s^{-1}$).

Selectivity. Substrate selectivity and affinity are evolved properties of an enzyme, related specifically to its metabolic role. For instance, the substrate range for highly specialized mammalian enzymes may be far narrower than that of extracellular fungal enzymes. In certain cases, broad substrate specificity is essential, as in the use of prote-olytic and lipolytic enzymes in laundry detergents, whereas in the production of most pharmaceuticals and fine chemicals, strict substrate selectivity is required to ensure fidelity in the reaction pathway, and viable conversion yields.

Control and optimization of stereoselectivity in synthetic bio-transformations and chiral resolutions have been seminal focus areas in biocatalysis in recent years, largely as a result of the inefficiency of conventional methods for stereochemical control. Many highly stereoselective biocatalysts are already in large-scale industrial use (see Table 3), and novel moleone reaction type to another, for example, in the modification of indole-3-glycerophosphate synthase, by a combination of rational design and directed evolution, to produce novel phosphoribosyl anthranilate isomerase activity[25]. Substrate selectivity can readily be altered, using molecular or reaction engineering methods, to the extent where compounds of very different steric structure and chemical nature from the natural substrates can be efficiently converted by the biocatalyst . For instance, the D-selectivity of a native hydan-toinase was reversed by mutagenesis and selective screening procedures, giving an L-specific enzyme with no equivalent in nature, able to catalyze the production of L-methionine[26]. Gene-shuffling techniques have been used to randomly recombine genes from two organisms possessing biphenyl dioxygenase activity, yielding a recombinant enzyme capable of oxidative degradation of polychlori-nated biphenyls (PCBs) and simple aromatic monomers[27].

Molecular stability. Biocatalyst stability is a major concern in virtually all bioprocesses, because it may affect process economics at a number of levels. Poor biocatalyst stability will result in longer process operations (resulting from decreasing catalytic efficiency), increased frequency of catalyst replacement and reduced product yields. The causes of reduced biocatalyst stability are intimately associated with the process conditions, and may include extremes of temperature, ionic strength, or pH that are outside the operating "stability window" of the catalyst, or the presence of denaturants, such as substrates/products, organic solvents, surfaces, or interfaces. Some of the options currently available to the bioprocess industry to address the problem of biocatalyst stability are summarized in Table 1.

Table 3. Stereoselective large-scale biocatalytic processes[56,57]

Process	Substrate	Product	Enzyme	Scale (tonnes per annum)	Company
Resolution	Racemic alcohols	Enantiomeric alcohols	Lipase	10^3	BASF
Resolution	Racemic amines	R- and S-amides	Lipase	10^2	BASF
Kinetic resolution	Racemic amino acid amides	L-Amino acids	Amidase	10^2	DSM
Amination	Fumaric acid	L-Aspartic acid	Aspartate ammonia lyase	10^2	DSM

The effect of any deleterious condition on the lifespan of a biocat-alyst is dependent first on the intrinsic molecular stability of the protein, and second on the mode of presentation of the enzyme in the reactor system (e.g., whether the enzyme is cell-encapsulated, immobilized, or crosslinked). Intrinsic molecular stability is a function of both amino acid sequence and tertiary (and quaternary) structure and is dictated, at least in part, by the origins of the protein. Studies of extremophilic organisms and their enzymes have provided extensive insights into both the outer limits of molecular stability and molecular mechanisms responsible for structural stabilization. Proteins of extremely high thermal stability (e.g., *T*m values of 80°-130°C) are typically isolated from hyperthermophilic organisms (growing at up to 115°C), whereas the proteins of psychrophilic organisms (growing to well below 0°C) are generally relatively unstable (e.g., T_m values of 20°-50°C). While the selection of a hyperther-mophile-derived biocatalyst provides a ready means of ensuring high molecular stability, such enzymes have evolved to function optimally at near the organism's growth temperature and exhibit reduced catalytic rates if employed at lower temperatures (see Fig. 3).

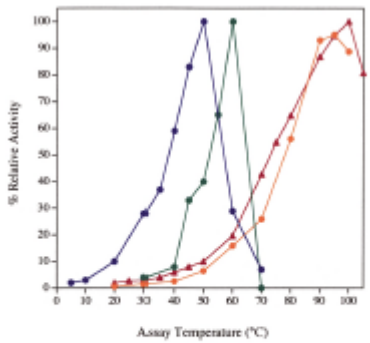

Figure 3. Temperature/activity profiles for caseinolytic activities of extracellular proteases from different thermal sources. Psychrophilic (*Xanthomonas maltophila*, 20°C (●); mesophilic (*Halomonas* spp., 37°C (●) extremely thermophilic (*Thermus aquaticus*, 72°C (●) and hyperthermophilic (*Desulfurococcus mobilis*, 88°C (▲). Data redrawn from refs 75–78, respectively.

Over two decades of intensive study into the mechanisms of protein stability have led both to the identification of a range of molecular stabilization mechanisms (e.g., the presence of salt-bridging networks: see reviews[28-30]) and to the engineering of increased structural stability by random and site-specific mutagenesis, enzyme evolution, and gene shuffling. While small increments in enzyme stability are typically achieved by site-specific mutagenesis[31], some remark-able increases (T_m values enhanced by 10°-15°C) have been achieved by directed enzyme evolution[32,33]. Such changes are not always to the detriment of catalytic activity. For example, the engineering of a 340-fold increase in stability at 100°C in a *Bacillus stearothermophilus* protease by eight site-specific mutations[34] was achieved without compromising catalytic activity, contradicting the dogma that molecular stability and activity are reciprocally related (through global conformational flexibility). Furthermore, there is good evidence for a general correlation between molecular stability to temperature and to other denaturing conditions[35,36], suggesting common initiators and/or pathways of protein denaturation, whether induced by high

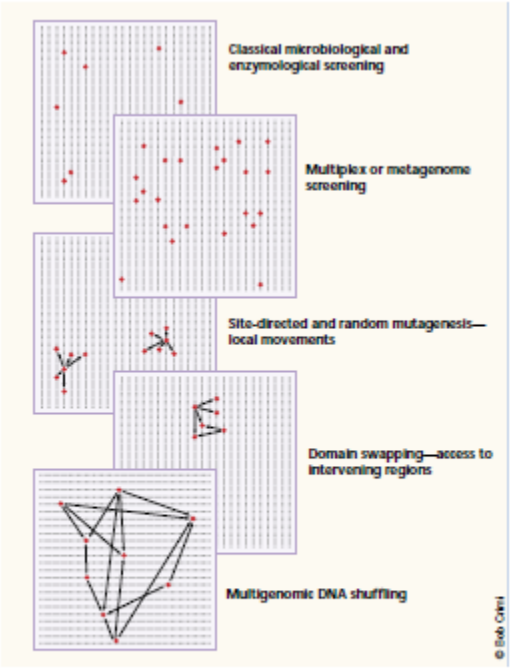

Figure 4. Moving around in sequence space. While classical and modern genomic screening can sample the range of evolved sequences, mutagenic and shuffling technologies can sample a much wider range of sequences outside natural evolutionary boundaries.

temperature, extremes of pH, or the presence of organic solvents or detergents. The practical consequence of this observation is that the selection of a more thermostable enzyme variant (e.g., derived from a thermophilic source) or the engineering of enhanced thermostabil-ity may also result in increased molecular resistance to a range of deleterious conditions.

Virtually all the industrially significant changes in catalyst stability (e.g., in the evolution of xylose isomerases in the glucose conversion/high-fructose syrup industry) have come from strain selection, chemical modification (immobilization, chemical crosslinking), classical mutagenesis, or random mutagenesis. Of these, chemical modification is still the most widely used technology in industry for increasing biocatalyst stability. This preoccupation with traditional technology may reflect the simplicity and reliability of the procedures, and their cost-effectiveness in comparison with more modern and sophisticated genetic technologies. Recent developments in immobilization and related technologies, including cell-surface expression of enzymes[37,38], multipoint covalent immobilization[39], and the application of crosslinked enzyme crystals[40], are likely to ensure that industrial biocatalysis continues to employ such methods. The last two of these methods take advantage of the fact that a restriction of conformational flexibility, as in covalent crosslinking or embedding in a crystal lattice, significantly increases the activation energy of denaturation and prevents enzyme aggregation.

Table 4. Examples of technologies available for enhancing biocatalyst characteristics

Target improvement	Technology available	Example	Activity enhancement
Increased turnover	Solubilization of enzyme	Solubilization of subtilisin (using isooctane + AOT*) for application in organic solvents	Specificity constant k_{cat}/K_M of solubilized subtilisin in octane = $10^3\times$ that of suspended enzyme, and 0.1× that in aqueous medium; stability in octane $10^3\times$ > in aqueous medium[2]
		Biocatalyst plastics using chymotrypsin and subtilisin	Incorporation of chymotrypsin and subtilisin in synthetic polymers gave $10^4\times$ increased reaction rates with high stability, in polar solvents, facilitating efficient peptide synthesis[58]
		Solubilization of chymotrypsin and subtilisin for use in organic solvents using propanol rinsing treatment	Activity of chymotrypsin and subtilisin Carlsberg 10^3 greater than freeze-dried powder and comparable with CLEC preparation[59]
	Molecular imprinting	Imprinting of subtilisin (Carlsberg and mutants) with sucrose, thymidine, and other nucleosides	Acylation of thymidine by subtilisin in THF* with 50× rate enhancement compared with non-imprinted enzyme. Substrate specificity altered and reactivity improved (50–180× enhancement) with alternative nucleosides as substrates, by imprinting with similar nucleophiles[60]
		Induction of enzyme activity by lyophilization of papain and lactoglobulin in presence of transition-state analog	Papain and lactoglobulin imprinted with a transition-state analog showed activity in the catalysis of β-elimination reactions 3× higher than nonimprinted protein[61]
Altered enantioselectivity*	Directed evolution	Improved enantioselectivity in lipases	*Pseudomonas aeruginosa* lipase activity with improved S-enantioselectivity; 2% ee increased to 81% ee after four generations of mutants[62]
		Reversed hydantoinase enantioselectivity	Conversion of D- to L-selectivity with fivefold increase in total activity, in *Arthrobacter* sp. hydantoinase used for production of L-methionine. No natural L-specific hydantoinase is known[25]
		Improved enantioselectivity in esterase-catalyzed stereoselective resolution	Esterase from *Pseudomonas fluorescens* in *E. coli* mutants selected to hydrolyze bulky 3-hydroxyesters, giving 25% ee (ref.63)
Altered functionality	Rational design and directed evolution	New isomerase activity evolved	Conversion of indole-3-glycerol phosphate synthase into phosphoribosylanthranilate isomerase with 2× lower k_{cat} and 15× lower K_M than the native enzyme of *E. coli*[25,64]
	Rational design	Desaturase altered to hydroxylase	Change of activity from double-bond insertion to hydroxylation in oleate desaturase, by site-directed mutagenesis at seven sites, selected by comparative analysis of amino acid sequences of oleate desaturases and unrelated hydroxylases[65]
	Gene shuffling and screening	Modified β-galactosidase activity	DNA shuffling gave an evolved β-galactosidase with 10-fold increased fucosidase and 40-fold less galactosidase activity, due to six amino acid changes[66]
Conversion of non-natural substrates	Directed evolution specificity	Recombination to alter substrate	Genes for biphenyl oxygenase activity from *Pseudomonas pseudoalcaligenes* and *Burkholderia cepacia*, with different substrate specificities, recombined to give dioxygenase activity for efficient PCB conversion[27]
Increased enzyme stability	Directed evolution	Thermostable esterase developed by directed evolution	Improved thermostability (>14°C increase in T_m after six cycles of mutation) without loss of activity, showing that the two properties are not mutually exclusive[33]

\

Table 4. Examples of technologies available for enhancing biocatalyst characteristics (continued)

Target improvement	Technology available	Example	Activity enhancement
		Increased thermostability of subtilisin E	Subtilisin was altered by eight amino acid substitutions, to give functional Thermitase using PCR mutagenesis and staggered extension process recombination[67]
	Genetic and rational approach	Increased thermostability in a protease	Protease activity with half-life 170 min (cf. 1 min in native enzyme) at 100°C, in the presence of denaturants, and wild-type level of activity, in eightfold mutant of thermolysin-like protease from *Bacillus stearothermophilus*[34]
Increased stability and activity	Cell-surface expression	Functional display of levansucrase activity on cell surface of *E. coli*	Levansucrase activity of *Zymomonas mobilis* co-expressed and anchored with ice-nucleation activity in conversion of protein from *Pseudomonas syringae* gave high sucrose to glucose and resistance to protease[38]
Chemical modification to increase stability	Covalent immobilization—multipoint attachment	Hyperstabilization of a thermophilic esterase	Thermophilic esterase from *Bacillus stearothermophilus* modified to give 3×10^8 increased thermostability, compared with native enzyme, by multipoint covalent immobilization through amino groups linked to glyoxyl agarose[68]
	Chemical crosslinking	Use of dextran dialdehyde for crosslinking	Stabilization of *E. coli* penicillin G acylase against thermal inactivation by crosslinking with dextran dialdehyde polymers. Modified enzyme showed ninefold increase in half-life of inactivation at 55°C and no decrease in V_{max} (ref. 69)
	CLEC formation	Crosslinking and chemical modification of subtilisin	Subtilisin crystallized, crosslinked, and chemically converted to give peroxidase activity with higher stability (10×) without loss of enantioselectivity[70]
	Lipid coating of enzymes	Application of glycoside hydrolases in organic medium	Lipid coatings on mannosidase, glucosidase, and glucosaminidase gave hydrophobic nature to enzyme surfaces and increased (80% cf. 10%) transglycosylation activity in isopropyl ether [71]
Chemical modification to increase activity	Modification with monofunctional reagents	Modification of catalase with Brij 35 surfactant	Catalase activity increased 200× in trichloroethylene and 15× in aqueous medium, compared with native enzyme[72]
Modified reaction milieu	Directed evolution	Increased esterase activity in aqueous/organic medium	*p*-Nitrobenzyl esterase activity of subtilisin in 30% DMF increased 50× using random mutagenesis[73]

[a]Abbreviations: AOT, Aerosol OT; CLEC, crosslinked enzyme crystal; ee, enantiomeric excess

Catalyst diversity

Whether a process is designed around the catalyst properties, or the catalyst is selected to suit process needs, the available range of relevant catalysts is a key issue. The range of enzymes potentially available to the bioprocess industry is firstly dictated by evolution, that is, by genetic diversity. Although little consensus exists on true genetic diversity (genetic represented as species diversity), estimates for plant, animal, and microbial diversity (see Table 5) are in the order of 10^7-10^8 species. With an average of 10^4-10^5 open reading frames per genome, many of which encode enzymes, the total enzyme genetic diversity may be as great as 10^{13} distinct functional sequences. This figure is, of course, substantially reduced by the varying, often high, homology between enzymes from related and even unrelated organisms.

This diversity of closely, distantly, or unrelated sequences represents the core biological resource, conceptually considered in terms of "sequence space"[41]. Native sequence space, representing the

diversity of functional (and nonfunctional) protein sequences available from nature, still represents a relatively small proportion of total sequence space—theoretically estimated as the total number of permutations of 20 amino acids (which is 20^{100} for a polypeptide of 100 amino acids). Only a relatively small proportion of native sequence space is currently available to the scientific and bioindustrial community, for the simple reason that a very high proportion of extant species have never been investigated, or, even if investigated, are not readily accessible. The acknowledged reality that a very high proportion of microbial species are currently "unculturable"[42] (Table 5) further restricts access to microbial (bacterial, archaeal, fungal, and unicellular algal) genomes and gene products.

The rapid development of genetic technologies has provided increasingly inventive methods for indirectly accessing new areas of sequence space (Fig. 4). Chemical, site-directed, and random muta-genesis technologies access regions proximal to the native sequences, whereas methods, such as gene and domain swapping, DNA shuffling, and other forms of combinatorial genomics access the regions between the parent sequences.

Modern bioprospecting methods such as multiplex or metagenome cloning[43-46] and similar methods used by companies such as Diversa Corporation (La Jolla, CA) and TerraGen Discovery (Vancouver, BC, Canada) directly access environmental genomes (whether culturable or unculturable). These methods access sequence space in a virtually random manner and have been successfully employed to isolate novel *nah*R (ref. 45) and amylase, lipase, and hemolytic enzyme genes[46].

Perspective

If we accept the new paradigm that the design of an optimized process may proceed without intimate consideration of the limitations of the biocatalyst, and that technologies are available for the subsequent selection or design of the "ideal" catalyst, the following questions are highly relevant: First, where are the limiting factors? And second, what more is required to overcome the limiting factors?

It is evident from the recent literature that *in vitro* enzyme evolution technologies are being very rapidly assimilated. These offer a relatively fast and inexpensive approach to the incremental modification of selected enzyme properties, with a reduced risk that desirable enzyme characteristics will be lost in the process.

A major limitation is the monoselective nature of almost all protein-engineering and enzyme-screening technologies. Single-function screening is essentially incompatible with the requirements of the "ideal process" paradigm, in which a defined set of target properties will require polyfunctional screening and/or engineering strategies. First attempts at multifunction screening have been reported[47].

However, despite the recent development of sophisticated molecular engineering and screening technologies, our ability to move around sequence space in search of the ideal catalyst for a given

Table 5. Estimates of species diversity in nature[43,74]

Group of organisms	Estimated species	Accessible (known) species (as % of total)
Animals (mammals, birds, fishes)	3.5×10^4	>90
Arthropods/Invertebrates	10^6–10^7	10
Nematodes	5×10^5	3
Higher plants	2.7×10^5	>90
Algae[a]	10^4–10^5	[70]
Bryophytes	2.5×10^4	70
Fungi[a]	1.5×10^6	[5]
Bacteria[a]	10^4–10^5	[1–10%]
Archaea[a]	10^3–10^6	[0.1–1%]
Viruses[a]	10^3–10^6	[4]

[a]These values are thought to be underestimated, possibly by one to two orders of magnitude. Estimates of accessible species will be proportionally lower.

process is still limited. Rational design methods face the greatest limitation in scope. For example, despite the intense focus over the past decade on the molecular determinants of protein stability, site-specific protein engineering cannot yet be reliably applied to protein stabilization. None of the "design rules" for thermostability that have evolved from comparative structural studies of more and less stable protein homologs[44,48] can be simply employed to indicate site-specific modifications for reliably enhancing molecular stability. The use of design algorithms for predicting multiple additive muta-tions[49] may provide a future guide to the engineering of hyperther-mostable enzymes. Such algorithms are not currently available for other functional characteristics (solvent stability, alterations in specificity, pH behavior, or turnover, for example) for which even less generic molecular information is available. Direct (i.e., metage-nomic) screening and DNA-shuffling technologies offer the widest access to sequence space. However, both are limited by the fact that any single environmental DNA sample is a poor representation of the enormous genomic diversity present in the multitude of global biotopes.

So are we ready to implement the ideal process paradigm? We conclude that many of the necessary molecular and screening technologies are in place, and there is early evidence of an evolution from monofunctional to multifunctional screening. In our opinion, the full implementation of the latter is the key to the successful acquisition of new "designer" biocatalysts for truly optimized "ideal" bio-processes.

References

1. Halling, P.J. Biocatalysis in low-water media: understanding effects of reaction conditions. *Curr. Opin. Chem. Biol.* 4, 74-80 (2000).
2. Wangikar, P.P, Michels, Clark, D.S. & Dordick, J.S. Structure and function of subtilisin BPN solubilized in organic solvents. *J. Am. Chem. Soc.* 119, 70-76 (1997).
3. Kazlauskas, R.J. Molecular modelling and biocatalysis: explanations, predictions, limitations and opportunities. *Curr. Opin. Chem. Biol.* 4, 81-88 (2000).
4. Selleck, G.A. & Chaudhuri, J. B. Biocatalysis in organic media using enzymes from extremophiles. *Enzyme Microb. Technol.* 25, 471-482 (1999).
5. Schmitke, J.L., Stern, L.J. & Klibanov, A.M. The crystal structure of subtilisin Carlsberg in anhydrous dioxane and its comparison with those in water and ace-tonitrile. *Proc. Natl. Acad. Sci. USA* 94, 4250-4255 (1997).
6. Aguilar, C.F *et al.* Crystal structure of the p-galactosidase from the hyperther-mophilic archaeon *Sulfolobus solfaricus:* resilience as a key factor in thermosta-bility. *J. Mol. Biol.* 271, 789-802 (1997).
7. Villeret, V. *et al.* The crystal structure of *Pyrococcus furiosus* ornithine carbamoyl-transferase reveals a key role for oligomerisation in enzyme stability at extremely high temperatures. *Proc. Natl. Acad. Sci. USA* 95, 2801-2806 (1998).
8. Martin A.B. & Schulze, P.G. Opportunities at the interface of chemistry and biology. *Trends Biochem. Sci.*. 24, M24-M28 (1999).
9. Arnold, F.H. & Volkov, A.A. Directed evolution of biocatalysts. *Curr. Opin. Chem. Biol.* 3, 54-59 (1999).
10. Arnold, FH. Design by directed evolution. *Acc. Chem Res.* 31, 125-131 (1998).
11. Affholter, J. & Arnold, F.H. Engineering a revolution. *ChemTech.* 26, 34-39 (1999).
12. Tapolczay, D.J., Kobylecki, R.J., Payne, L.J. & Hall, B. Extracting order from chaos.*Chem. Ind.* 19, 772-775 (1998).
13. Woodley, J.M. & Lilly, M.D. Process engineering of two-liquid phase biocatalysis.*Prog. Biotechnol.* 8, 147-154 (1992).
14. Wandrey, C., Liese, A. & Kihumbu, D. Industrial biocatalysis: past, present and future. *Org. Process Res. Dev.* 4, 286-290 (2000).
15. Flitsch, S.L., Aitken, S.J., Cho, C.S.Y., Grogan, G. & Staines, A. Biohydroxylation reactions catalysed by enzymes and whole-cell systems. *Bioorg. Chem.* 27, 8190 (1999).
16. Sette, L., Lanzarini, G. & Pifferi, P.G. Whole cell biocataysis for an oil desulphur-ization process. *Fuel Process. Technol.* 52, 145-153 (1997).
17. Woodley, J.M. & Lilly, M.D. Biotransformation reactor selection and operation. In *Applied biocatalysis.* (eds Cabral, J.M.S., Best, D., Boross L. & Tramper J.) 371-393 (Harwood Academic, London; 1994).
18. Chauhan, R.P. & Woodley, J.M. Increasing the productivity of bioconversion processes. *ChemTech* 27, 26-30 (1997)
19. Shanklin, J. Exploring the possibilities presented by protein engineering. *Curr. Opin. Plant Biol.* 3, 243-248 (2000)

20. Tischer, W & Kasche, V. Immobilised enzymes: crystals or carriers? *Trends Biotechnol.* 17, 326-335 (1999).
21. Adlercreutz, P. On the importance of the support material for enzymatic synthesis in organic media—support effects at controlled water activity. *Eur. J. Biochem.*199, 609-614 (1991).
22. Kim, J. & Dordick, J.S. Unusual salt and solvent tolerance of a protease from an extreme halophile. *Biotechnol. Bioeng.* 55, 471-479 (1997).
23. Toba, S. & Mez, K.M. The concept of solvent compatibility and its impact on protein stability and activity enhancement in non-aqueous solvents. *J. Am. Chem. Soc.* 118, 9939-9948 (1997)
24. Jaeger, K.-E. & Reetz, M.T. Directed evolution of enantioselective enzymes for organic chemistry. *Curr. Opin. Chem. Biol.* 4, 68-73 (2000).
25. Altamirano, MM, Blackburn, J.M., Aguayo, C. & Fersht, AR. Directed evolution of a new catalytic activity using alpha/beta barrel scaffold. *Nature* 403, 617-622 (2000).
26. May, O., Nguyen, P.T., & Arnold, F.H. Inverting enantioselectivity and increasing total activity of a key enzyme in a multi-enzyme synthesis creates a viable process for production of L-methionine. *Nat. Biotechnol.* 18, 317-320 (2000).
27. Kumamaru, T., Suenaga, H., Watanabe, T. & Furukawa, K. Enhanced degradation of polychlorinated biphenyls by directed evolution of biphenyl dioxygenase. *Nat Biotechnol.* 16, 663-666 {1998).
28. Hough, D. & Danson, M.J. Extremozymes. *Curr. Opin. Chem. Biol.* 3, 39-36 (1999).
29. Daniel, RM & Cowan, D.A. Biomolecular stability and life at high temperatures. *Cell. Mot. Life Sci.* 57, 250-264 (2000)
30. Jaenicke, R. & Bohm, G. The stability of proteins in extreme environments. *Curr. Opin. Struct. Biol.* 8, 738-748 (1998).
31. Lee, B. & Vasmatzis, G. Stabilization of protein structures. *Curr. Opin. Biotechnol.*. 8,423-428(1999).
32. Kuchner, O. & Arnold. F.H. Directed evolution of enzyme catalysts. *Trends Biotechnol.* 15, 523-530 (1997).
33. Giver, L, Gershenson, A, Freskard, RO. & Arnold, F.H. Directed evolution of a thermostable esterase. *Proc. Natl. Acad. Sci. USA* 95,12809-12813 (1998).
34. Van den Burg, B., Vriend, G., Venema, O.R. & Eijsink, V.G. Engineering an enzyme to resist boiling. *Proc. Natl. Acad. Sci. USA* 95,2056-2060 (1998).
35. Daniel, R.M., Cowan, D.A., Curran, M. & Morgan, H.W. A correlation between protein thermostability and susceptibility to proteolysis. *Biochem. J.* 207, 641-644 (1982).
36. Owusu, R.K. & Cowan, D.A. A correlation between microbial protein thermostability and resistance to denaturation in aqueous-organic solvent two-phase systems. *Enz. Microb. Technol.* 11, 568-574 (1989).
37. Richins, R.D., Kaneva, I., Mulchandari, A., & Chen, W. Biodegradation of organophosphorus pesticides by surface-expressed organophosphorus hydrolase. *Nat Biotechnol.* 15, 984-987 (1997).

38. Jung, H.C., Lebeault J.M. & Pan, J.G. Surface display of *Zymomonas mobilis* lev-ansucrase by using ice-nucleation protein of *Pseudomonas syringae*. *Nat. Biotechnol.* 16, 576-580 (1998).
39. Blanco, R.M. &, Guisan, J.M. Stabilization of enzymes by multipoint covalent attachment to agarose aldehyde gels—borohydride reduction of trypsin agarose derivatives. *Enz. Microb. Technol.* 11, 360-366 (1989).
40. Hanng, D. & Schreier, P Cross-linked enzyme crystals. *Curr. Opin. Chem. Biol.* 3, 35-38 (1999).
41. Stemmer, W.PC. Searching sequence space—using recombination to search more efficiently and thoroughly instead of making bigger combinatorial libraries. *Bio/Technol.* 13. 549-553 (1995).
42. Amann, R.I., Ludwig, W. & Schleiffer K.H. Phylogenetic identification and *in-situ* detection of individual microbial cells without cultivation. *Microbiol. Rev.* 59, 143-169(1995).
43. Cottrell, M.T., Moore, J.A. & Kirchman, D.L. Chitinases from uncultured marine
44. Cowan, D.A. Microbial genomes-the untapped resource. *Trends Biotechnol.* 18, 14-16 (2000).
45. Henne, A., Schmitz, R.A., Bomeke, M., Gottschalk, G. & Daniel, R. Screening of environmental DNA libraries for the presence of genes conferring lipolytic activity on *Escherichia coli*. *Appl. Environ. Microbiol.* 66, 3113-3116 (2000).
46. Rondon, M.R. *et al.* Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. *Appl. Environ. Microbiol.* 66, 2541-2547 (2000).
47. Petrounia, I.P. & Arnold, F.H. Designed evolution of enzymatic properties. *Curr. Opin. Biotechnol.* 11, 325-330 (2000).
48. Jaenicke, R. & Bohm, G. The stability of proteins in extreme environments. *Curr. Opin. Struct. Biol.* 8, 738-748 (1998).
49. Malakauskas, S.M. & Mayo, S.L. Design, structure and stability of a hyperther-mophilic protein variant. *Nat. Struct. Biol.* 5, 470-475 (1998).
50. Chen K. & Arnold, F. Tuning the activity of an enzyme for unusual environments: sequential random mutagenesis of subtilisin E for catalysis in dimethylformamide. *Proc. Natl. Acad. Sci. USA* 90, 5618-5622 (1993).
51. Wfehtje, E. & Adlercreutz, P. Water activity and substrate concentration effects on lipase activity. *Biotechnol. Bioeng.* 55, 798-806 (1997).
52. Burton, S.G., Duncan, J.R., Kaye P.T. & Rose P.D. Activity of mushroom polyphenol oxidase in organic medium. *Biotechnol. Bioeng.* 42, 938-946 (1994).
53. Reimann, A., Robb, D.A. & Halling, P.J. Solvation of CBZ-amino acid nitrophenyl esters in organic media and the kinetics of their transesterification by subtilisin. *Biotechnol. Bioeng.* 43, 1081-1086 (1994).
54. Khemelnitsky, Y.L., Welch, S.H., Clark, D.S & Dordick, D.S. Salts dramatically enhance activity of enzymes suspended in organic solvents. *J. Am. Chem. Soc.*116, 2647-2648 (1994).

55. Ru, M.T., Dordick, J.S., Reimer, J.A. & Clark, D.S. Optimising the salt-induced activation of enzymes in organic solvents:effects of lyophilisation time and water content. *Biotechnol. Bioeng.* 63, 233-241 (1999).
56. Schmid, A., Dordick, J.S., Haven, B., Wubbolts, M. & Witholt, B. Industrial biocatal-ysis today and tomorrow. *Nature* 409, 258-268 (2001).
57. Liese, A., Seelbach, K. & Wandrey, C. *Industrial biotransformations.* (Wiley-VCH, Weinheim, Germany; 2000).
58. Wfeng, P., Sergeeva, M.V., Lim, L. & Dordick J.S. Biocatalyst plastics as active and stable materials for biotransformations. *Nat. Biotechnol.* 15, 789-793 (1997).
59. Partridge, J., Halling, P.J. & Moore, B.D. Practical route to high activity enzyme preparations for synthesis in organic media. *Chem. Commun.* 7, 841-842 (1998).
60. Rich, J.O. & Dordick, J.S. Controlling subtilisin activity and selectivity in organic media by imprinting with nucleophilic substrates. *J. Am. Chem. Soc.* 119, 3245-3252 (1997).
61. Slade C.J. & Vulfson, E.N. Induction of catalytic activity in proteins by lyophilization in the presence of a transition state analogue. *Biotechnol. Bioeng.* 57, 211-215 (1998).
62. Reetz, M.T., Zonta, A., Schimossek, K., Liebeton, K. & Jaeger, K.-E. Creation of enantioselective biocatalysts for organic chemistry by *in vitro* evolution. *Angew. Chem., Int. Edn. Engl.* 36, 2830-2833 (1997).
63. Bornscheuer, U.T, Altenbuchner, J. & Meyer, H.H. Directed evolution of an esterase for the stereoselective resolution of a key intermediate in synthesis of epithilones. *Biotechnol. Bioeng.* 58, 554-559 (1998).
64. Sutherland, J.D. Evolutionary optimization of enzymes. *Curr. Opin. Chem. Biol.* 4,263-269 (2000).
65. Broun, P., Shanklin, J., Whittle, E. & Somerville, C. Catalytic plasticity of fatty acid modification enzymes underlying chemical diversity of plant lipids. *Science* 282, 1315-1317 (1998).
66. Zhang, J.H., Dawes, G., & Stemmer, WP. Directed evolution of a fucosidase from a galactosidase by DNS shuffling and screening. *Proc. Natl. Acad. Sci. USA* 94, 4504-4509 (1997).
67. Zhao, H. & Arnold, F Directed evolution converts subtilisin E into a functional equiv-alent of thermitase. *Protein Eng.* 12, 47-52 (1999).
68. Fernandez-Lafuente, R., Cowan, D.A. & Wood, A.N.P. Hyperstabilisation of a ther-mophilic esterase by multipoint covalent attachment. *Enz. Microb. Technol.* 17, 366-372 (1995).
69. Kazan, D., Ertan, H. & Eraslan, A. Stabilisation of *Escherichia coli* penicillin G acy-lase against thermal inactivation by cross-linking with dextran dialdehyde polymers. *Appl. Microbiol. Biotechnol.* 48, 191-197 (1997).
70. Haring, D. & Schreier, P. Novel biocatalysts by chemical modification of known enzymes: cross-linked microcrystals of the semi-synthetic peroxidase seleno-subtilisin. *Agnew. Chem. Int. Edn. Engl.* 37, 2471-2473 (1998).

71. Mori, T. & Okahata, Y. A variety of lipid-coated glycoside hydrolases as effective glycosyl transfer catalysts in homogenous organic solvents. *Tetrahedron Lett.* 38, 1971-1974 (1997).
72. Jene, O., Pearson, J.C. & Lowe, C.R. Surfactant modified enzymes: solubility and activity of surfactant-modified catalase in organic solvents. *Enz. Microb. Technol.* 20, 69-74 (2000).
73. Moore, J.C. & Arnold, F.H. Directed evolution of a *para*-nitrobenzyl esterase for aqueous-organic solvents. *Nat. Biotechnol.* 14, 458-462 (1996).
74. Bull, A.T., Goodfellow, M. & Slater, J.H. Biodiversity as a source of innovation in biotechnology. *Ann. Rev. Microbiol.* 46, 219-252 (1992)
75. Margesin, R. & Schinner, F. Characterization of a metalloprotease from psy-chrophilic *Xanthomonas maltophilia. FEMS Microbiol. Lett.* 79, 257-261 (1991).
76. Gimenez, M.I., Studdert, C.A., Sanchez, J.J. & De Castro, R.E. Extracellular protease of *Natrialba magadii:* purification and biochemical characterisation. *Extremophiles* 4, 181-188 (2000).
77. Cowan, D.A. & Daniel, R.M. Purification and some properties of an extracellular protease (Caldolysin) from an extreme thermophile. *Biochim. Biophys. Acta.* 705,293-305 (1982).
78. Cowan, D.A., Smolenski, K.A., Daniel, R.M. & Morgan, H.W. An extremely thermostable extracellular protease from a strain of the archaebacterium *Desulfurococcus* growing at 88°C. *Biochem. J.* 247, 121-133 (1987).