Available online at www.scholarsresearchlibrary.com

Scholars Research Library

Der Pharmacia Lettre, 2010: 2 (1) 131-140
(http://scholarsresearchlibrary.com/archive.html)

ISSN 0974-248X
USA CODEN: DPLEB4

Future of Combinatorial biocatalysis in a drug discovery

Devendra Kumar Singh

Birla Institute of Technology Mesra, Ranchi, India

Abstract

The published applications of combinatorial biocatalysis have continued to expand at a growing rate. This is exemplified by the variety of enzyme catalysts and whole-cell catalysts used for the creation of libraries through a wide range of biocatalytic reactions, including acylation, glycosylation, halogenation, oxidation and reduction. These biocatalytic methods add the capability to perform unique chemistries or selective reactions with complex or labile reagents when integrated with classical combinatorial synthesis methods. Thus, applications towards the production of libraries *de novo*, the expansion of chemically derived combinatorial libraries, and the generation of novel combinatorial reagents for library synthesis can be achieved. Theoretically, these results illustrate what is already evident from nature: that complex, biologically active, structurally diverse compound libraries can be generated through the application of biocatalysis alone or in combination with classical organic synthesis approaches.

Keywords: combinatorial biocatalysis; drug discovery; nonaqueous enzymology; enzyme

---

Introduction

Drug discovery is a high-stakes game. Each year, pharmaceutical companies spend billions of dollars in search of the elusive new blockbuster drug. The traditional route to new drugs is to screen large sources of chemical entities for promising lead compounds; possible sources include in-house libraries of existing chemicals, fermentation broths, plant extracts and compounds from marine organisms[1]. Once a promising lead compound is identified in a preliminary screen, it is turned over to chemists for modification. At considerable time and expense (the industry rule of thumb is two weeks and US$7500 per derivative[2]), the original lead structure is modified in an effort to improve its pharmacological properties. Desirable characteristics can include increased potency, higher selectivity, reduced toxicity and greater bioavailability. The odds of success are low but the potential payoff is high. Although estimates vary, it is generally accepted that only

one compound in 10 000 will succeed as a new pharmaceutical, with an overall time from discovery to market of 12 years and a total cost of approximately US$350 million[3, 4]. Success does, however, have its rewards: the most successful drugs have markets in the range of US$1 billion a year and produce immeasurable benefits to society.

These incentives provide a powerful driving force to develop new technologies that streamline and expedite the drug-discovery process. One such technology, which has the potential radically to change the way that new drugs are discovered, is combinatorial chemistry. The overall aim of combinatorial chemistry is to make new molecules faster, more cheaply and in numbers large enough for high-throughput screening. There are many variations on this theme, most involving automated synthesis to assemble all possible combinations of a given set of chemical building blocks (e.g. amino acids, nucleotides, sugars and simple organic rings), followed by screening and sorting the combinatorial library to isolate and identify any active product(s)[3, 4, 5, 6, 7]. Generally, the more diverse the products, the more likely it is that the library will contain something useful and novel. The improvement in throughput offered by combinatorial methods is enormous: a typical estimate places the number of compounds synthesized as approximately 3300 per chemist-month at a cost of US$12 per compound7]. The great promise of combinatorial chemistry has spawned many new drug-discovery companies and triggered the development of in-house efforts at nearly all the major pharmaceutical companies.

Combinatorial biocatalysts

Combinatorial biocatalysis is a powerful technology, especially useful for creating focused libraries of complex natural products or synthetic compounds . This technology uses nature's catalysts in an effort to recreate and enhance the exquisite diversity of organic molecules generated and recycled in nature. Nature has long practiced solution-phase divergent synthesis to create the unparalleled complexity of natural products. The broad array of chemistries required for the production and degradation of organic biomolecules all must occur under mild and uniform conditions within the living cell. Some of these reactions (such as hydroxylations of non-activated carbons, mild and selective oxidative reactions, decarboxylation, etc.) are difficult to reproduce using purely chemical means under any conditions. Because most natural products are polyfunctional and chiral, a high degree of catalytic selectivity is important. Yet, all natural product structures must be efficiently synthesized and recycled, so catalytic versatility is important for survival. Thus, by necessity, enzymes have evolved to catalyze a broad array of reactions with high catalytic efficiency, high selectivity, and few byproducts on the full breadth of chemical structures observed in nature.

Although combinatorial biocatalysis is a relatively recent addition to the field of synthetic chemistry, it has probably been a staple of nature for millions of years. Nature synthesizes biomolecules of unparalleled structural complexity by encoding enzymes that catalyse a myriad of reactions. These reactions are often carried out on low molecular weight synthons, the biological equivalent of lead compounds in drug discovery (Fig. 1a). The products of these reactions induce a variety of important functions, which, in turn, play a major role in determining the characteristic traits of the cell. Over evolutionary time, random mutations to DNA result in the expression of modified enzymes, which may transform new precursors or transform existing precursors in different ways, to produce new organic biomolecules. If these new biomolecules provide a survival advantage for the cell, the enzymatic pathway that produced them is

preserved. Through this process, the vast diversity of natural products has been produced, from which many important therapeutics and agrichemicals have been identified.

Steps involved in the combinatorial synthesis of new biomolecules from synthons (precursors) in nature and from lead compounds in vitro

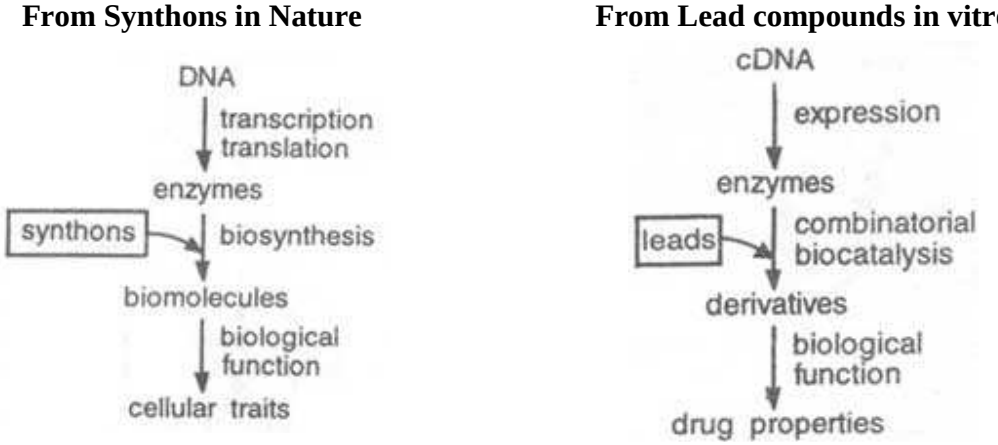

Fig. 1. Pathways for the combinatorial synthesis of new biomolecules from synthons in nature (a) and from lead compounds *in vitro* (b).

However, the evolutionary time scales necessary for the creation of new organic molecules are hardly appropriate for drug-discovery programmes. Combinatorial biocatalysis aims to accelerate this natural model and apply it to the discovery and optimization of pharmaceuticals and agrichemicals. As shown in Fig. 1b, this approach can take advantage of the huge diversity of natural catalysts (enzymes and whole cells), as well as the rapidly growing supply of recombinant and engineered enzymes, for the direct derivatization of interesting pharmaceutical or agrichemical leads. (NB Combinatorial biocatalysis should be distinguished from combinatorial biosynthesis, which refers to the generation of novel molecules derived from natural products by the genetic engineering of biosynthetic pathways in living microorganisms[8].) As in the cell, the mild, uniform, nontoxic, noncorrosive conditions for enzymic catalysis allow simple reactors to compartmentalize library synthesis for individual compounds; the cell membrane is replaced by the walls of a multiwell microplate. The wide range of enzyme activities in a typical living cell can be replaced by one (or more) enzyme(s) for a specific conversion (or combination of conversions) of interest. Derivative libraries generated in solution can then be screened for the biological activity of interest, or for suitability as a clinical drug or field agrichemical.

Advantages of combinatorial biocatalysis

Biocatalysis offers a wide variety of synthetic possibilities within the expanding purview of combinatorial chemistry. The types of reactions catalysed by enzymes and microorganisms are summarized in Table 1 [9], and a representative scheme of possible transformations of a given lead is shown in Fig. 2. When applied to the modification and optimization of existing lead structures, biocatalytic reactions have some distinct advantages over synthetic chemical reactions (A). By necessity, enzymes have evolved to catalyse reactions with high efficiency, high yield and few byproducts. Enzymatic reactions are also highly selective, which is particularly

important if the aim is to preserve some structural features of a lead molecule while modifying others. The high regioselectivity of enzymatic reactions also affords the opportunity for specific combinatorial modification of lead molecules with multiple copies of the same functional group. Furthermore, high selectivity avoids the need to protect and deprotect other reactive functional groups on the molecule, a multistep process that increases the complexity and reduces the overall yield of a transformation. Many of these advantages of combinatorial biocatalysis allow extensive reaction schemes to be readily automated using simple, inexpensive equipment.

Table 1. Biocatalytic reactions available for combinatorial synthesis

Reaction type	Specific reactions
Introduction of functional groups	Carbon–carbon-bond formation
	Hydroxylation
	Halogenation
	Halohydrin formation
	Cycloadditions
	Addition of amines
Modification of existing functionalities	Oxidation of alcohols to aldehydes and ketones
	Reduction of aldehydes and ketones to alcohols
	Oxidation of sulfides to sulfoxides
	Oxidation of amino groups to nitro groups
	Oxidation of thiols to thioaldehydes
	Hydrolysis of nitriles to amides and carboxylic acids
	Replacement of amino groups with hydroxyl groups
	Lactonization
	Isomerization
	Epimerization
	Dealkylation
	Methyl transfer
Addition onto functional groups	Esterification
	Carbonate formation
	Carbamate formation
	Glycosylation
	Amidation
	Phosphorylation

The large size and vast structural diversity of libraries produced using combinatorial biocatalysis stem from both the wide selection of available biocatalytic transformations (Table 1) and the large number of building blocks that can be added onto the lead molecule in the process of library generation. For example, over 60 acyl donors of different types are available for enzymatic acylations, including vinyl and trifluoroethyl esters, vinyl carbonates, and bifunctional acyl donors. A wide variety of structural elements, including aromatic, carbocyclic, heterocyclic and aliphatic fragments of different polarity, can be introduced into the lead compound using these diverse building blocks.

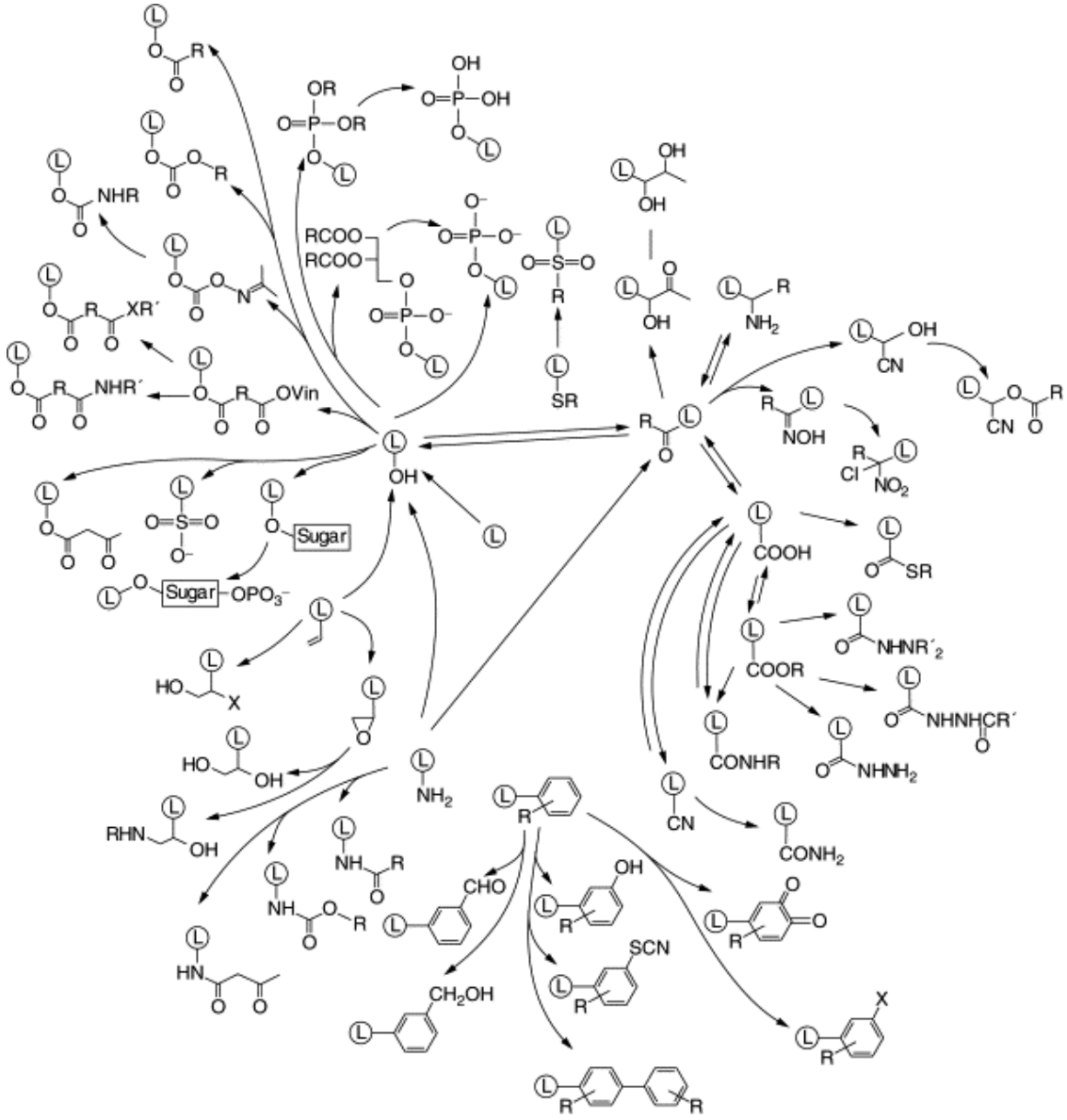

Fig. 2. Possible biocatalytic transformations of lead compounds (the lead scaffold is represented by L). The pathways shown are only a subset of all possible reactions and are intended to illustrate just some of the derivatives produced by biocatalysis (note: X=Cl, Br or I; Vin=vinyl).

Biocatalysis in nonaqueous media

It should be noted that the breadth of *in vitro* enzymatic reactions is expanded significantly by using nonaqueous media. Enzymes are catalytically active in many organic solvents[10, 11] and can catalyse reactions that would ordinarily involve complex biocatalytic pathways *in vivo* and/or energetic intermediates to overcome thermodynamic barriers in water. Moreover, the high solubility of most natural products in organic solvents, in contrast to water, enables higher substrate concentrations to be used and, hence, higher reaction productivities to be obtained.

Although enzymes retain their activity in organic solvents, their catalytic efficiencies are often orders of magnitude lower than in water[10, 11]. Nevertheless, recent work has demonstrated that suitable manipulation of an enzyme's microenvironment can lead to dramatic improvements in enzyme activity in organic solvents. One striking example of this is the influence of nonbuffer salts on subtilisin-catalysed transesterification of *N*-acetyl-L -phenylalanine ethyl ester with *n*-propanol in hexane. The k_{cat}/K_m of subtilisin catalysis is increased over 3700-fold upon lyophilization of the enzyme in the presence of KCl to provide a biocatalyst preparation containing 98% w/w salt[12]. This technique has been extended to nonserine proteases, including thermolysin, with a similar degree of activation[13]. For example, salt-activated thermolysin has been used for the synthesis of 2′-ester derivatives of paclitaxel; a 20-fold activation, compared with native thermolysin, was achieved for paclitaxel acylation in tertiary-amyl alcohol[9, 14]. Salt-activated thermolysin was also effective for derivatization of adenosine and 3,6-dihydroxytropane (Y. L. Khmelnitsky *et al*., unpublished). Salt-activated subtilisin has been used for the selective acylation of bergenin, adenosine, 3-hydroxytropane and bicyclo(2.2.2)oct-5-ene-2,3-dimethanol. Hence, salt activation may represent a general technique for activating enzymes capable of acylating complex molecules in organic media.

Another technique used to activate enzymes in organic solvents involves solubilizing the enzyme in the presence of small concentrations of surfactants[15, 16]. Although enzymes are insoluble in nearly all organic solvents (but are active to some extent), the presence of the surfactant results in the formation of hydrophobic ion pairs that enable the enzyme to dissolve in hydrophobic solvents. Compared with suspended native enzyme, activation of nearly four orders of magnitude has been achieved for subtilisin catalysis in octane[17]. The solubility of ion-paired enzymes in organic solvents should prove especially valuable for combinatorial biocatalysis using substrates attached to a solid phase.

The many special attributes of enzymatic reactions and biotransformations enable the application of combinatorial biocatalysis to small organic molecules as well as more complex natural products not readily amenable to modification by conventional combinatorial techniques[9]. Several lead molecules used for the generation of libraries employing combinatorial biocatalysis for drug discovery and optimization are shown in Fig. 3. These include nucleosides (e.g. adenosine), flavonoids (e.g. bergenin), alkaloids (e.g. 3,6-dihydroxytropane), polyketides (e.g. erythromycin) and terpenoids and taxanes (e.g. paclitaxel). These diverse structures are among the examples that illustrate how combinatorial biocatalysis has been used to produce novel libraries.

Automated synthesis of bergenin libraries

The derivatization of the flavonoid bergenin demonstrates the use of a broad range of biocatalytic reactions for an automated, iterative synthesis of a 600-member library. Initial screening of candidate biocatalysts, in 96-well plates using a multiprobe liquid-handling robot, identified 16 purified enzymes and 25 microorganisms capable of accepting bergenin as a substrate (P. C. Michels *et al*., unpublished). Using these enzymes, modifications to bergenin and its derivatives could be made to generate a diverse library in two synthetic rounds (Fig. 4). The reactions applied to bergenin can be generally classified as those that introduce new functional groups to this lead molecule (e.g. hydroxylation and halogenation), those that modify existing

functionalities (e.g. oxidations and reductions) and those that add new groups on to existing functionalities (e.g. acylation[18], glycosylation and phosphorylation).

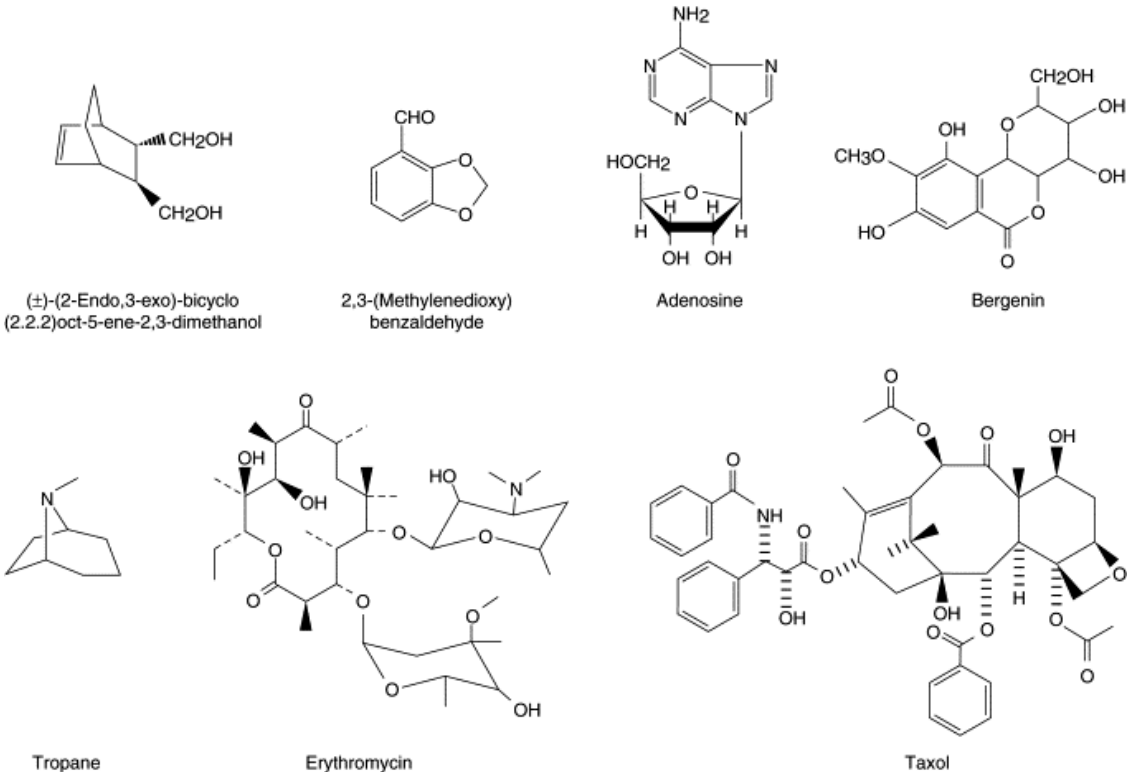

Fig. 3. examples of lead molecules that have been used for the generation of libraries by combinatorial biocatalysis.

Successful generation of each derivative was confirmed by high-throughput mass spectrometry and, for selected derivatives, by ^{13}C nuclear-magnetic-resonance spectroscopy. The resulting library was screened for biological activity against several targets, including xanthine oxidase (a potential target for the treatment of gout and multiple sclerosis) and urokinase (an anticancer target). Inhibitors whose potency was two orders of magnitude better than that of bergenin were detected for these targets in initial screens, and the resynthesis and confirmation of the most potent derivatives is proceeding. The biocatalytic reactions were automated using inexpensive commercial equipment designed for high-throughput screening; however, several modifications were necessary to effect the automation of iterative reaction schemes. Specifically, reactions were performed in commercially available 96-well filter plates that were sealed using a clamping mechanism and a pierceable rubber septum. These parallel reactors accommodated a variety of reaction media (including volatile organic solvents) appropriate for the different synthetic reactions. Robotic grippers, needle manifolds and vacuum boxes on the robotic liquid handler enabled transport, sampling and recovery of reaction products. These modifications allowed the integration of automated synthesis, work-up and product archiving on a single robotic workstation.

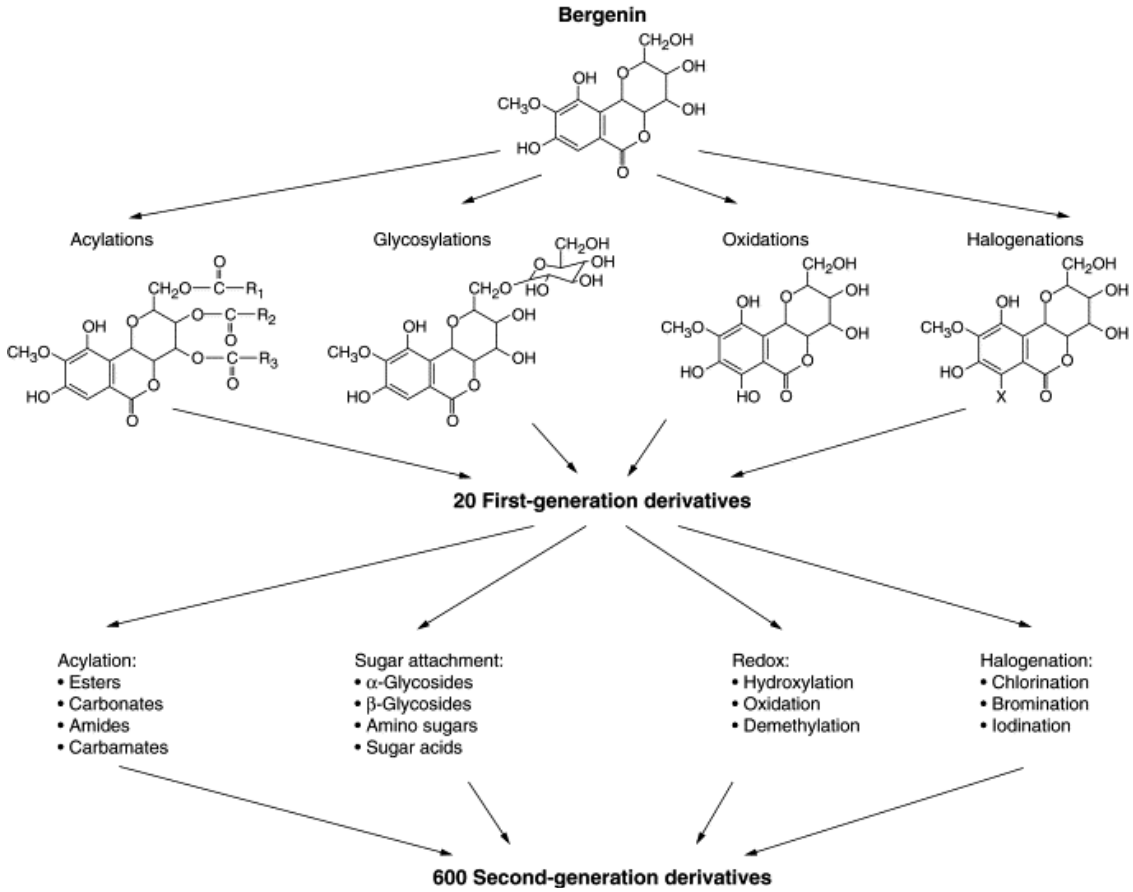

Fig. 4. The process for automated synthesis of bergenin derivatives by combinatorial biocatalysis

Combinatorial acylation of dihydroxytropane

The advantage of enzyme regioselectivity for the generation of libraries with polyfunctional lead molecules has been demonstrated from experiments with the scaffold 3,6-dihydroxytropane, a derivative of tropane (Fig. 3). Automated screening of 75 lipases, proteases and esterases in several organic solvents identified 14 enzymes capable of acylating 3-hydroxytropane. A similar procedure was used to identify 24 enzymes capable of acylating a 6-hydroxy derivative of the tropane scaffold with high efficiency. Thus, enzymes could be grouped according to those specific for the 3 position, those reacting only with the 6 position, and those capable of reacting with both positions. Using catalysts specific for each hydroxyl position, 3,6-dihydroxytropane was selectively derivatized using two heterocyclic vinyl esters and two vinyl carbonates to produce 8 mono- and 16 diacylated derivatives as individual compounds in 24 wells of a deep 96-well plate (P. C. Michels *et al*., unpublished).

Amide libraries from polybenzyl esters

The utility of lipases (particularly one from *Pseudomonas cepacia*) for generating a library of amides was recently demonstrated by Adamczyk *et al*.[19] The aromatic diester dibenzyl-1,2-phenylenedioxydiacetate was used as the core scaffold, and mono-(*N-tert*-butoxycarbonyl) diamines were employed as reaction partners for the diester. Reacting the diester with a mixture of five mono-(*N-tert*-butoxycarbonyl) diamines produced a library of 26 different products,

including 15 bisamides, five monoamide monoesters and small amounts of five monoacid products. This example illustrates that *P. cepacia* lipase has broad specificity for the transformation of benzyl esters to amides, thus enabling the combinatorial synthesis of a bisamide library.

Conclusion and perspectives

The evolution of life has occurred through enzyme-catalysed combinatorial organic synthesis, coupled with natural selection of those biomolecules that possess optimal activities. Mimicking this process, combinatorial biocatalysis harnesses the natural diversity of enzymatic reactions for the iterative synthesis of organic libraries. Combinatorial biocatalysis is therefore a powerful complement to other combinatorial methods for the synthesis of new organic molecules.

The potential of combinatorial chemistry to revolutionize drug discovery is no longer in question: what remains to be seen is the extent of its impact. Combinatorial chemistry in its various manifestations enables the generation of new molecules with unprecedented speed. As a result, the supply of lead compounds has never been greater. However, the evolution from promising lead to useful drug requires much optimization, ranging from the reorganization of key structural features on the molecule to the reconfiguration of asymmetric centres, and biocatalysis is particularly well suited to these tasks. Moreover, as the structural complexity of the lead compound increases, the advantages of biocatalysis become more important. In particular, natural products, which represent an enormously rich and diverse source of lead compounds for drug discovery, often have very complex structures[20]. The discovery of new natural resources and the expanding capabilities of combinatorial synthesis will ensure that the pipeline of new lead compounds continues to expand. Naturally, combinatorial biocatalysis and its role in drug discovery are expected to grow along with it.

Acknowledgements

I thank J. Krstenansky for many helpful discussions and suggestions.

References

[1] L. P. Schacter *et al.*. *Semin. Oncol.* 19 (**1992**), pp. 613–621.
[2] R. Tsao and K. D. Noonan. *In Vivo* 11 (**1997**), pp. 28–32.
[3] 3. J. Alper. *Science* 264 (**1994**), pp. 1399–1401.
[4] J. C. Hogan, Jr. *Nat. Biotechnol.* 15 (**1997**), pp. 328–330.
[5] N. K. Terrett, M. Gardner, D. W. Gordon, R. U. Kobylecki and J. Steele. *Tetrahedron* 51 (**1995**), pp. 8135–8173.
[6] D. J. Ecker and S. T. Crooke. *Biotechnology* 13 (**1995**), pp. 351–360.
[7] S. Borman. *Chem. Eng. News* 75 (**1997**), pp. 43–62.
[8] C. R. Hutchinson. *Biotechnology* 12 (**1994**), pp. 375–380.
[9] Khmelnitsky, Y. L., Michels, P. C., Dordick, J. S. and Clark, D. S. (**1996**) in *Molecular Diversity and Combinatorial Chemistry: Libraries and Drug Discovery* (Chaiken, I. M. and Janda, K. D., eds), pp. 144–157, American Chemical Society, Washington, DC, USA

[10] Dordick, J. S. (**1990**) in *Applied Biocatalysis* (Vol. 1) (Clark, D. S. and Blanch, H. W., eds), pp. 1–51, Marcel Dekker
[11] A. M. Klibanov. *Trends Biotechnol.* 15 (**1997**), pp. 97–101.
[12] Y. L. Khmelnitsky, S. H. Welch, D. S. Clark and J. S. Dordick. *J. Am. Chem. Soc.* 116 (**1994**), pp. 2647–2648.
[13] Bedell, B. A., Mozhaev, V. V., Clark, D. S. and Dordick, J. S. *Biotechnol. Bioeng*. (in press)
[14] Y. L. Khmelnitsky, C. Budde, J. M. Arnold, A. Usyatinsky, D. S. Clark and J. S. Dordick. *J. Am. Chem. Soc.* 119 (**1997**), pp. 11554–11555.
[15] V. M. Paradkar and J. S. Dordick. *J. Am. Chem. Soc.* 116 (**1994**), pp. 5009–5010.
[16] P. P. Wangikar, P. C. Michels, D. S. Clark and J. S. Dordick. *J. Am. Chem. Soc.* 119 (**1997**), pp. 70–76.
[17] M. V. Sergeeva, V. M. Paradkar and J. S. Dordick. *Enzyme Microb. Technol.* 20 (**1997**), pp. 623–628.
[18] Mozhaev, V. V. *et al*. *Tetrahedron* (in press)
[19] M. Adamczyk, J. C. Gebler and J. Grote. *Bioorg. Med. Chem. Lett.* 7 (**1997**), pp. 1027–1030.
[20] A. D. Kinghorn and E-Y. Seo. *Chem Tech* 26 (**1996**), pp. 46–54.