OPINION

Exploiting plug-and-play synthetic biology for drug discovery and production in microorganisms

Marnix H. Medema, Rainer Breitling, Roel Bovenberg and Eriko Takano

Abstract | One of the most promising applications of synthetic biology is the biosynthesis of new drugs from secondary metabolites. Here, we survey a wide range of strategies that control the activity of biosynthetic modules in the cell in space and time, and illustrate how these strategies can be used to design efficient cellular synthetic production systems. Re-engineered versions of secondary metabolite biosynthetic pathways identified from any genomic sequence can then be inserted into these systems in a plug-and-play fashion.

The rapid progress of genome-sequencing projects has revealed thousands of uncharacterized secondary metabolite biosynthetic pathways, many of which are expected to produce new biologically active compounds such as antitumour drugs, cholesterol-lowering agents and antibiotics. These pathways are a potentially rich source for drug discovery; they constitute a giant evolutionary compound library that, in contrast to randomly constructed libraries, has been highly pre-selected for optimal stability, bioavailability and bioactivity. However, these pathways originate from a diverse range of organisms, including slow-growing soil bacteria[1], extremophiles[2], rare plants[3] and species from environmental metagenome samples[4], many of which are uncultured or unculturable. Moreover, the conditions in which the pathways are active and the compounds are produced at levels sufficient for characterization are often unknown[5,6].

Therefore, new concepts are required to exploit this richness. Interestingly, in microbial genomes the genes that code for secondary metabolite biosynthetic pathways are usually highly grouped in gene clusters that often also contain the pathway-specific regulators and transport systems. This allows for their easy identification by *in silico* detection of signature genes or gene domains that are specific for certain classes of pathways[7–9]. However, finding methods for expressing genes to levels sufficient for biochemical characterization is still a large challenge. Current methods include cultivation in different media, heterologous expression and manipulation of the regulators that control the gene clusters[10]. Some researchers have given up on the sequence-first approach and are using proteomics to focus on those gene clusters that are naturally highly expressed[11]. Although these strategies enable the functional characterization of single gene clusters, they can be time-consuming. We suggest that the tools of synthetic biology[12–14] can enable us to tune newly sequenced pathways so that they can be easily plugged into suitably pre-engineered microbial hosts in which the cellular machinery is already optimized for overproduction of compounds synthesized from specific pathway classes. This can then enable us to discover new drugs from these gene clusters in a rapid and high-throughput manner (FIG. 1). Subsequently, several intrinsic levels of cluster modularity can be exploited to further engineer the biosynthetic pathways for optimal production and generation of useful derivatives (BOX 1).

The large-scale engineering that is necessary to make this approach efficient will depend on the ability to exert control over cellular function in both space and time. Here, we outline the design principles of temporal and spatial control (FIG. 2) that we believe will be crucial in order to accomplish this new technology. We discuss both temporal and spatial engineering at different scales, from allosteric control of enzyme activities to fine-tuning of expression patterns and metabolic programmes, and from protein–protein interactions to subcellular organization and microbial communities. Finally, we discuss how these engineering features can be integrated to design versatile host strains that will enable efficient metabolite screening and production.

Exerting temporal control

When manipulating biosynthetic pathways for optimal production capacities, the timing of expression of pathway components is important. During high-level constitutive expression, the continuous metabolic requirements for production are in conflict with the changing demands for cellular survival and growth. Therefore, dynamic regulation of pathway expression in time normally results in more efficient production. For instance, Farmer and Liao showed that dynamically regulating *Escherichia coli* lycopene production by recruiting the nitrogen regulatory system (Ntr) to stimulate production at times of high glycolytic flux resulted in much higher product titres[15]. Synthetic biology can be applied at several scales to achieve optimized exploitation of the metabolic potential of engineered microorganisms: at the rapid scale of allosteric control of enzyme activities and gene expression; at the intermediate scale of temporal fine-tuning of enzyme expression patterns; and at the long-term scale of synchronizing population activities and executing metabolic programmes.

Allosteric control. The cellular machinery can be controlled most rapidly and directly through metabolite concentrations. This enables a rapid response of enzyme activities to changes in metabolite concentrations, but also the rapid adaptation of gene expression levels to a specific metabolic situation[16]. For this purpose, metabolites function as cofactors for regulatory proteins. As part of a synthetic biology strategy, the effector

© 2011 Macmillan Publishers Limited. All rights reserved

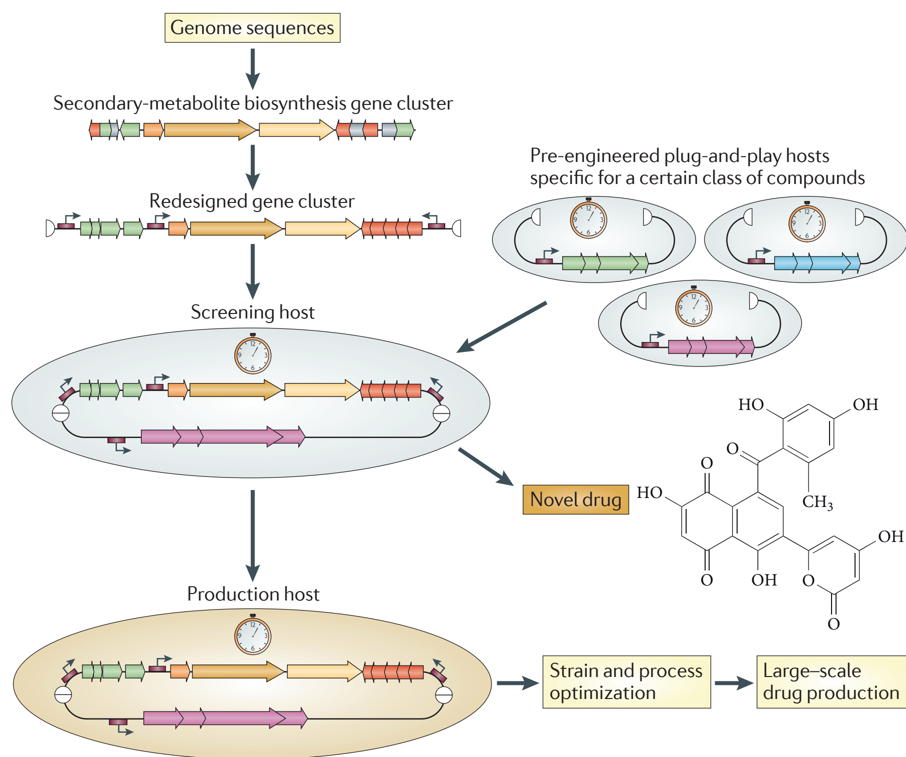


Figure 1 | **Pipeline for plug-and-play expression of unknown biosynthetic pathways.** Overview of the proposed pipeline for plug-and-play execution of secondary metabolite biosynthetic pathways. Gene clusters of interest are selected from the genomic databases and are redesigned for streamlined expression in pre-engineered screening hosts. These are specifically optimized for the broad chemical class of the compound that is encoded by each pathway in terms of the levels and timing of gene expression (signified by the clocks) and the spatial parameters that are crucial for maximizing the metabolic fluxes towards the end compound. For each screening host, a complementary production host is available, to which any compounds with useful bioactivities can be transferred. Further synthetic tuning of these hosts will then lead to efficient production of these new compounds.

specificity of these regulators can be altered so that they can be used as switches in different contexts, beyond their natural regulatory role[17]. For example, the L-arabinose-binding regulator AraC has been engineered to bind D-arabinose instead, through the construction of mutant libraries[18]. Unfortunately, protein engineering has not advanced far enough to allow straightforward creation of regulators that bind less closely related small molecules. However, metabolites can also be bound by riboswitches, which are regulatory RNA elements that reside in the non-coding regions of mRNAs and regulate their translation through changes in their folding pattern that are induced by the binding of a small molecule. Compared with proteins, these elements have the advantage that they can be engineered more easily to bind to a specific small molecule, mainly because RNAs can be synthesized on large scales more straightforwardly. Recently, Dixon *et al.* mutated an adenine-specific riboswitch that was cloned in front of a chloramphenicol resistance gene and screened it with ligands derived from a library of small molecules to select for riboswitch–ligand pairs that could activate the resistance[19]. In this way, they were able to engineer riboswitches that were specific for two rather distantly related molecules, ammeline and azacytosine. Therefore, it should be possible to create switches that can respond to many different sorts of metabolic input.

For example, riboswitches could be engineered to recognize the final product of a precursor biosynthesis operon in a drug discovery strain in order to activate this operon when the cell runs out of precursors. If the intermediates of plugged-in pathways are known, various riboswitches could be engineered to recognize each intermediate and intelligently control the expression of the enzymes governing each step accordingly (FIG. 2). In such a self-regulating system, toxic accumulation of intermediates would be prevented and metabolic fluxes would be automatically adjusted to changing conditions.

Similarly, in gene cluster optimization for industrial drug production, the transcriptional units that encode the transporters could be regulated by a riboswitch that recognizes one of the late intermediates, in order to have the transporters available in time to avoid toxic intracellular build-up of the end product.

Timing of enzyme expression. The generation of biosynthetic intermediates that cannot be processed downstream, or the production of enzymes by the cell before their substrate becomes available, is energetically wasteful. Enzyme production (and the associated gene transcription) is one of the most costly processes in terms of energy and resources in a microbial cell, and its optimization has clear evolutionary benefits[20]. Therefore, fine-tuning of enzyme expression through the engineering of transcriptional units, promoters and ribosome-binding sites (RBSs) is needed in the development of screening strains to achieve optimally timed fluxes towards the metabolic precursors (BOX 1). This phenomenon of 'just-in-time' gene expression was first reported in *E. coli* amino acid biosynthesis, in which enzymes are consecutively produced in distinct phases in an order corresponding to their sequence in the pathway[21]. In this manner, wrongly timed superfluous enzyme production is prevented. Similar just-in-time gene expression has been observed in the biosynthesis of starter units used by exotic polyketide synthases (PKSs) and precursors used by non-proteinogenic amino acid non-ribosomal-peptide synthetases (NRPSs)[22]. Such sequential activity patterns can be generated by differential binding affinities of transcription factors to the promoters of the genes involved[23]. A library of transcription factor binding sites with different strengths would allow the use of this coordinated regulation in artificial biosynthetic operons. Combining the construction of the libraries with *in silico* modelling of regulatory outputs based on experimentally determined regulatory responses could facilitate the predictable design of entire regulatory networks without the need for extensive tuning[24]. The approach could be extended to temporal ordering of whole operons by adding yet another layer of regulation in which the expression of synthetic-pathway-specific regulators acting on these operons is also ordered sequentially.

By using synthetic regulation to optimize the timing of enzyme production, the cell can direct its limited resources towards production of the right enzymes at the right time. This is important because core biosynthetic enzymes are often rate limiting for

© 2011 Macmillan Publishers Limited. All rights reserved

total production, as is the case in cephalosporin biosynthesis[25]. Although in current industrial practice the production of drugs is continuous after the start-up phase, such just-in-time expression could also be implemented to achieve continuous oscillation of pathway expression in synchrony with other cellular programmes[26]. This can sometimes lead to more efficient flux, as occurs naturally in yeast glycolysis, for example[27].

Population synchronization of metabolic programmes. To obtain maximal production of compounds that are encoded by pathways plugged into a strain, it would be desirable to fully exploit the biosynthetic resources of the whole population synchronously. Synchronicity of compound production ensures that the maximal fraction of cells is contributing to biosynthesis and avoids the loss of resources that would be the unavoidable consequence of population heterogeneity. Currently, synchronized gene expression is not always achieved in industrial fermentations. Interestingly, 'synchronized bacterial clocks' in which quorum sensing is applied to synchronize oscillatory gene expression in large populations have been generated recently[28]. Such a synchronization system could also be applied to synchronize the activation of genetic programmes in drug-producing bacteria.

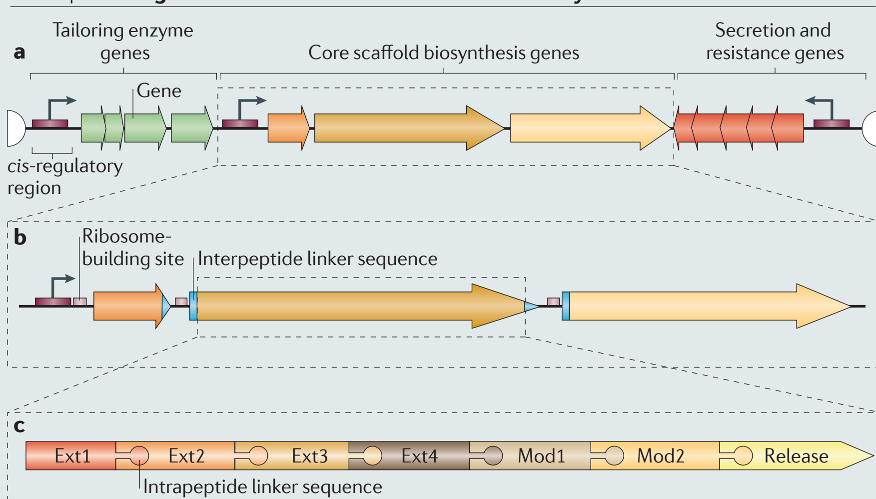
A particular benefit of this synchronization during drug production would be the ability to use distinct metabolic programmes sequentially; cells could be designed to focus on biomass accumulation in the start-up phase and then switch to a distinct drug-production phase. Such metabolic programmes also occur in natural producers of secondary metabolites, such as *Streptomyces coelicolor*, which has been shown to undergo a major metabolic shift at the transition from exponential to stationary phase[29,30]. However, in a synthetically designed metabolic programme, one would not want to wait for depletion of metabolic resources to occur before drug production is started, but would want instead to switch to this programme before the end of the exponential growth phase, so that full-scale drug production can be coupled to well-adjusted slow growth that is just high enough to maintain biomass levels.

Exerting spatial control

The optimal engineering of spatial aspects of microbial metabolism is crucial to the effective design of cellular machinery: the local concentrations of enzymes and metabolites determines the rate at which the biosynthetic reactions in a pathway can proceed[31]. The more tightly metabolites are channelled between reaction centres, the more efficiently enzyme resources are used. This is most important when the intermediates and/or end products are hydrophobic and thus limited in their diffusion. However, even the spatial arrangement of genes on a bacterial chromosome seems to be crucial, because mRNA diffusion from the site of transcription in *E. coli* was also recently shown to be very limited[32]. Similarly to temporal control, spatial control can be exerted at multiple scales: at the miniature scale of protein complexes; at the intermediate scale of subcellular organization; and at the large scale of entire microbial communities (FIG. 2).

Scaffolding of enzyme complexes. At the scale of proteins, synthetic protein scaffolds — which assemble proteins into macromolecular complexes — have been successfully used to increase local enzyme concentrations and, thus, to increase the metabolic flux through a biosynthetic pipeline in an *E. coli* host[33]. In this example, scaffolds were used to link different numbers of interaction domains from metazoan signalling proteins to biosynthetic

Box 1 | **Building blocks at different levels of modularity**



Secondary metabolite biosynthetic pathways display modularity at different levels of organization. Modules at every level can be implemented as building blocks that can be recombined in new ways in the synthetic engineering of secondary metabolism, following the way in which these clusters were naturally engineered during evolution[73].

The first level of organization of biosynthetic gene clusters is their subdivision into multiple operons (see the figure, panel **a**), each of which is often responsible for the biosynthesis of a distinct part of the end product. In the course of evolution, nature has often recombined such operons in new ways to generate new compounds. For example, the hybrid antibiotic simocyclinone has probably evolved from the fusion of aminocoumarin-producing operons and anthracycline-producing operons[73]. This strategy will be mimicked in synthetic engineering.

Going one level down, operons are also modular, consisting of many different genes (see the figure, panel **b**). The composition of these operons can be altered in various ways. Thus, biosynthetic steps can be removed or new functionalities can be added. For instance, tailoring enzymes that alter the backbone of a compound can be introduced heterologously[74].

At the final level, the core scaffolds of many secondary metabolites are synthesized by polyketide synthases (PKSs) or non-ribosomal-peptide synthetases (NRPSs). These contain several modules that catalyse the incorporation of a specific precursor into the growing polyketide or peptide chain. Genomic analyses show that the PKSs present in the genome of *Streptomyces avermitilis*, for example, have largely evolved through recombination of these modules and the domains that they contain[75]. This process has been mimicked in synthetic approaches by rearranging modules from, for example, erythromycin, rapamycin and pikromycin polyketide synthesis[37].

Several tailoring enzymes can also be incorporated into the synthases themselves as internal protein domains (see the figure, panel **c**; Ext represents a chain extension domain and Mod represents a modification or tailoring domain). Well-known examples are methyltransferase and oxidation domains. However, examples of transporter, formyltransferase, sulphotransferase, nitroreductase, lanthionine synthetase and cytochrome p450 oxidoreductase domains fused to these large polypeptides can also be found in genome sequences, opening up possibilities for exploitation by synthetic engineering.

© 2011 Macmillan Publishers Limited. All rights reserved

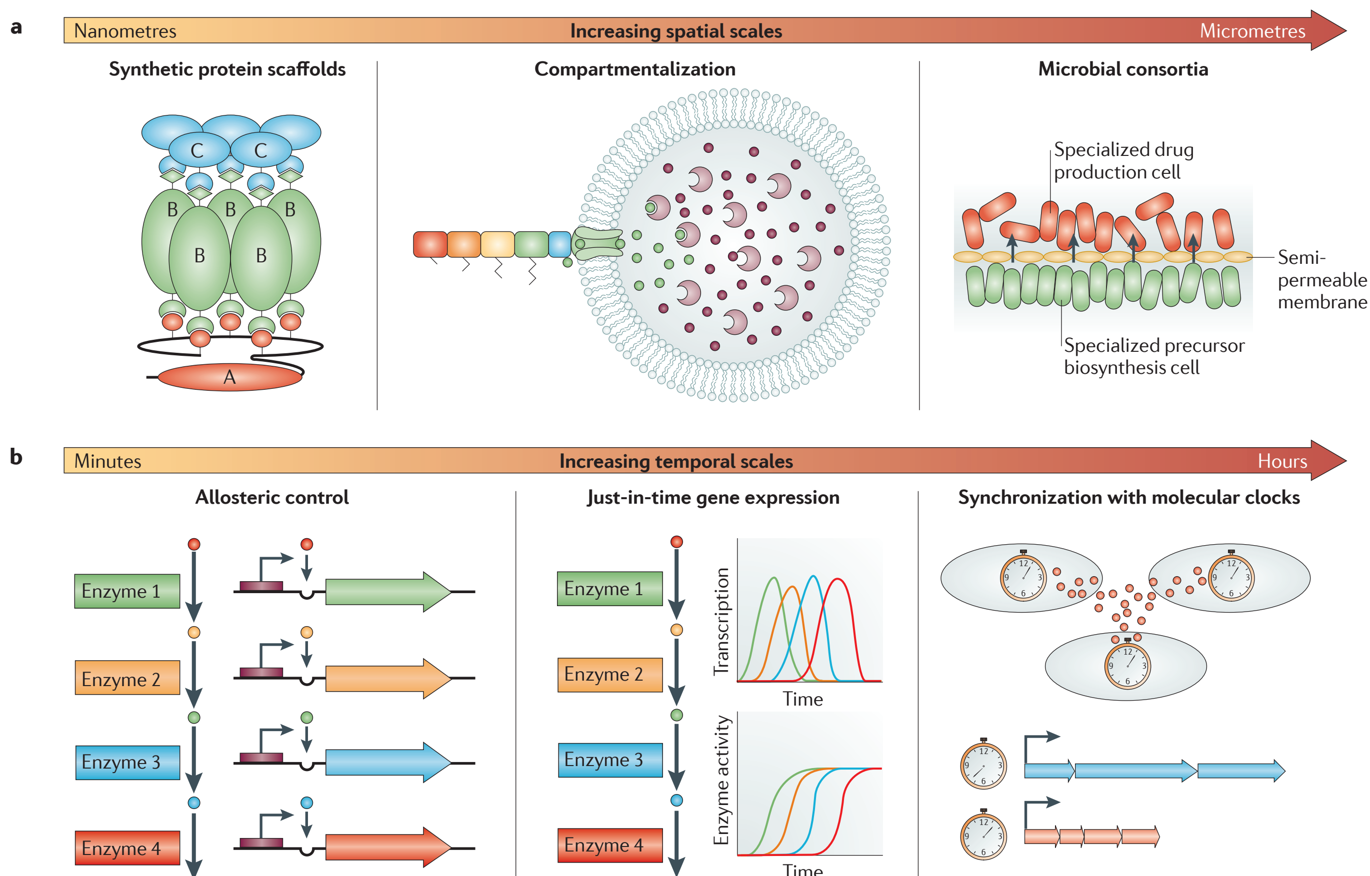


Figure 2 | **Controlling space and time on different scales in an optimized plug-and-play system.** Cellular systems can be controlled at different scales in space and time. **a** | In space, protein scaffolding can be implemented to physically co-localize enzymes (A, B and C) from a certain biosynthetic pathway, and the number of linkers connecting the enzymes can be used to tune stoichiometry. Compartmentalization of pathways can also improve local enzyme and substrate concentrations and thus increase pathway fluxes, but it can also make outward transport more efficient and prevent self-toxicity. Microbial consortia can be of great value because they allow the metabolic systems of each cell type to be specifically tuned to a part of the production pathway and therefore avoid the need for trade-offs in a single cell type. **b** | In time, transcription can be rapidly controlled on the basis of metabolite concentrations through allosteric control, which allows for intelligent tuning of the expression of each enzyme based on the metabolic state of the pathway. One level higher, just-in-time gene expression can ensure that genes are only transcribed when the proteins that they encode are needed, so that protein translation is optimally coupled to the implemented biosynthetic pathway. Finally, synchronization of cell populations can help to synchronize metabolic programmes and to directly use as large a part of the population as possible for compound production.

enzymes catalysing subsequent steps of the *Saccharomyces cerevisiae* mevalonate pathway[33]. The protein stoichiometry in the scaffolds could be tuned to the catalytic efficiencies of the subsequent enzymes by varying the number of interaction domains that linked them, in order to optimize the flux through the pathway and avoid build-up of intermediates. In the optimal configuration, production titres increased almost 80-fold, even though enzyme expression was lower than in the native configuration[33]. Recently, the scaffolding strategy was also successfully used to obtain unprecedented high titres of glucaric acid in *E. coli*[34]. Such protein scaffolds — complemented with protein fusions, if beneficial — could well be used to divert and increase metabolic fluxes towards the precursors for certain chemical classes in the development of screening strains. This may prove very useful given that precursor biosynthesis is a serious bottleneck in the optimization of product titres. For example, when biosynthesis of the core structure of the important antibiotic erythromycin was optimized using codon-optimized synthetic PKS genes, these genes were so efficient that they had to be downregulated again to avoid metabolic imbalance of precursor levels, which were too low to constitute an adequate supply to the PKSs[35].

In the secondary metabolite biosynthetic pathways, NRPSs and type I modular PKSs also function as scaffolds with docking domains that facilitate module–module interactions. Scaffolding has proved to be crucial for efficient pathway flux in such systems. In attempts to construct synthetic PKSs, only those with functional interaction domain pairs showed substantial activity[36,37]. These assembly lines can also be customized to physically link tailoring enzymes to the domains of the core scaffold, thereby increasing yields of the final product. In such a procedure, the natural linker domains could be replaced with linker domains from other protein systems. This strategy could relieve the need for highly product-specific tailoring enzymes, because the co-localization of the enzymes would probably reduce the rates of unwanted side reactions.

Cellular compartmentalization. An aspect that should not be underestimated is the subcellular organization of enzyme complexes. For instance, the hybrid NRPS–PKS complexes that produce bacillaene in *Bacillus subtilis* assemble into an organelle-like megacomplex[38]. The observation that this megacomplex is membrane associated

© 2011 Macmillan Publishers Limited. All rights reserved

indicates that this configuration may be important not only for increasing local enzyme concentrations, but also for efficient transport of the product out of the cell[38]. In fungi, an important part of penicillin biosynthesis is known to take place in the peroxisome[39]. Also, the biosynthesis and export of aflatoxin in *Aspergillus parasiticus* have recently been shown to be highly coordinated through accumulation in specific vesicles and subsequent transport to the vacuole[40]. Because the last two biosynthetic steps leading to the final product are performed in the vesicles, the compartmentalization is likely to prevent self-toxicity, as is the case for many plant secondary metabolites[41]. Another important advantage of such dedicated subcellular organizations seems to be the ability to maintain rapid fluxes through high local concentrations of enzymes and intermediates[39]. Therefore, compartmentalization should be considered seriously for the construction of efficient screening and production strains.

It is possible to use existing compartments of bacteria or fungi and their known protein-targeting systems for this purpose, as pioneered by Bayer *et al.*[42], who were able to increase production of a synthetic methyl halide biosynthetic pathway when the crucial enzyme was targeted to the vacuole. However, this would be limited to species with well-characterized targeting systems and could also lead to problems caused by competition between the engineered proteins and native organellar proteins. Therefore, a more ambitious aim would be to custom design bacterial organelles that could be introduced into desirable hosts[43]. The recent discovery and functional dissection of a gene cluster governing the biogenesis of magnetosomes in magnetospirilli[44] offers a possible template for the future biogenesis of such organelles in model organisms such as *E. coli*. An orthogonal protein translocation system could then be introduced into these compartments to set them apart entirely for their engineered purpose; in such an effort, it would be crucial to avoid crosstalk with the native secretion systems, which could have a serious negative impact on cellular fitness. Intriguingly, we recently discovered genes that encode NRPSs fused to major facilitator transporters[45], opening up the possibility that megasynthase assembly lines can be directly coupled to the membranes of synthetic organelles to allow immediate shuttling of the polyketide or peptide scaffold into the organelle, where it can then be further modified by a range of tailoring enzymes.

An interesting alternative to membranous organelles is also available: the operon that is responsible for generating proteinaceous microcompartments in *Salmonella enterica* has recently been characterized, and specific amino-terminal sequences were shown to direct proteins to these compartments[46]. These microcompartments were successfully expressed heterologously in *E. coli* cells, and GFP could be targeted to them by fusion to an amino-terminal targeting sequence[47]. The advantage of these compartments is that they strongly co-localize enzymes while allowing limited and selective diffusion of small molecules in and out of the compartment, as has been shown for the related carboxysome microcompartments of *Halothiobacillus neapolitanus*[48]. Interestingly, diffusion selectivity can be altered by mutating the shell proteins[48].

Designing microbial consortia. Once new drugs enter the production phase, there is no reason why the whole production process should be carried out by a single strain. Instead, because the biological route towards the production of a drug consists of multiple distinct steps, a further route of exploration is to see whether it could be beneficial to let these distinct steps be performed by dedicated cell types. This would allow the metabolism of each cell type to be specifically tuned to each step instead of forcing a compromise between them in one cell[49]. Microbial communities in which different cell types have distinct roles can be engineered[50]. For example, specialized cells could be constructed to supply biochemical precursors for other specialized cells that convert these precursors into the final product. These cells could then be attached to a surface in a multilayered biofilm[51], or they could be attached to two sides of a semipermeable membrane which allows diffusion of the intermediates but not of the final product between the two populations (FIG. 2). Recent experiments show that such spatial sequestering of syntrophic co-cultures allows stable growth of both populations because it prevents one cell type being outcompeted by the other[52]. However, two or more types of cell can be grown together stably in a fermentor if the cells are engineered to control gene expression of a suicide protein in one of the strains, based on detection of relative cell density of the other strain by a quorum sensing system, such as LuxI–LuxR or LasI–LasR[53,54]. Alternatively, an obligatory cooperative system can be implemented in which each cell type supplies an essential metabolite to the other. For instance, a consortium has been constructed consisting of two *S. cerevisiae* strains that lack either lysine or adenine biosynthesis and supply one another with the essential resource to allow growth[55]. Such approaches may be the most practical strategies by which to pioneer the large-scale industrial production and harvesting of compounds in the short-term, because spatial sequestering would require technical innovations beyond standard fermentor growth.

Suitable host strains

All of these control strategies can now be used to design microbial systems specifically tuned for drug discovery, and these can subsequently be upgraded to specialized production strains for the large-scale production of each drug. Choosing optimal host strains as a basis for these strategies is of vital importance.

Substantial efforts have focused on transplanting secondary metabolite biosynthetic pathways to fast-growing industrial microorganisms such as *E. coli* or *S. cerevisiae*[56–59]. However, production titres of, for instance, polyketides remained low owing to metabolic constraints inherent in the host, because required precursors were present at low intracellular levels[59]. The first efforts

Glossary

Metabolic flux
The flow of metabolites through a metabolic system.

Modularity
The extent to which a system can be subdivided into 'modules' that can be recombined in new ways.

Non-ribosomal-peptide synthetase
A large multidomain megasynthase enzyme that functions as an assembly line to produce peptides, usually containing non-proteinogenic amino acids.

Orthogonal
Functioning entirely uncoupled from the native system, without crosstalk.

Quorum sensing
Intercellular communication between members of a microbial colony or community, carried out through small signalling molecules.

Secondary metabolite
A metabolite that is characteristic of particular strains or species and that has a specific function outside the primary metabolism that governs core cellular functions.

Tailoring enzyme
An enzyme that modifies the core scaffold of a compound by, for example, oxidoreduction, methylation or glycosylation.

Type I modular PKS
A large multidomain megasynthase enzyme that functions as an assembly line to produce a polyketide.

© 2011 Macmillan Publishers Limited. All rights reserved

to redirect metabolic flux towards these precursors show promising results, with titres of the polyketide phloroglucinol being increased approximately fourfold[60].

Unfortunately, transplanting biosynthetic pathways to a new biological context usually yields unpredictable outcomes, because the metabolic fluxes are not tuned to provide the necessary precursors[61]. Bringing parts of primary metabolism under synthetic regulatory control offers a possible solution. This has recently been applied in the engineering of biofuel cells, in which the fatty acid degradation pathway was knocked out and genes governing the biosynthetic steps towards fatty esters and fatty alcohols were overexpressed at the same time[62]. Alternatively, one could attempt to redirect the metabolic fluxes by specific gene knockouts guided by computational modelling of metabolism[63].

In the short term, current industrial production strains such as actinomycetes that are optimized using classical strain engineering[64] may still be useful starting points for engineering, given their proven track record. But for a future principled approach to the problem, using genome-minimized production hosts seems to be the most exciting approach, given that they waste only a minimal amount of cellular resources on reactions other than their designed purpose. In addition to the recent successes in reducing existing genomes with a top-down approach[65,66], the first cell controlled by an entirely synthetic genome has been constructed recently[67–69]. This may soon empower researchers to construct synthetic systems for this type of application from the ground up.

The near future

In the coming years, technological advances will probably continue at the rapid pace we have witnessed recently. Not only will DNA sequencing advance more and more quickly, but we will also attain an unprecedented understanding of cellular systems at the levels of RNA, proteins and metabolites. From existing cellular systems, we can identify thousands of biologically active compounds, all refined by millions of years of evolution but mostly uncharacterized in terms of their chemical structure and medicinal activity. At the same time, the costs of DNA synthesis are already dropping rapidly[70], and highly efficient techniques for DNA assembly have become available[71]. This inaugurates a new age in which the complete *de novo* synthesis of re-engineered gene clusters becomes economically feasible. If we grasp this opportunity now and start to use all available dimensions of cellular complexity to design bacterial cells that are specialized for drug discovery and production, this will allow unprecedented access to the riches that nature provides. Achieving this ambitious aim will require the integration of knowledge from many disciplines, from bioorganic chemistry and cell biology to bioinformatics and microbial ecology[72]. But when the concepts described in this article are applied in a concerted fashion on a large scale, synthetic drug production may become reality sooner than many people could have imagined just a few years ago.

Marnix H. Medema is at the Department of Microbial Physiology and the Groningen Bioinformatics Centre, University of Groningen, Nijenborgh 7, 9747AG Groningen, The Netherlands.

Rainer Breitling is at the Groningen Bioinformatics Centre, University of Groningen, Nijenborgh 7, 9747AG Groningen, The Netherlands, and at the Institute of Molecular, Cell and Systems Biology, College of Medical, Veterinary and Life Sciences, Joseph Black Building, University of Glasgow, Glasgow G12 8QQ, UK.

Roel Bovenberg is at the Centre for Synthetic Biology, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Nijenborgh 7, 9747AG Groningen, The Netherlands, and at the DSM Biotechnology Center, Alexander Fleminglaan 1, 2613AX Delft, The Netherlands.

Eriko Takano is at the Department of Microbial Physiology, University of Groningen, Nijenborgh 7, 9747AG Groningen, The Netherlands.

Correspondence to E.T.
e-mail: e.takano@rug.nl

doi:10.1038/nrmicro2478
Published online 29 December 2010

1. Lautru, S., Deeth, R. J., Bailey, L. M. & Challis, G. L. Discovery of a new peptide natural product by *Streptomyces coelicolor* genome mining. *Nature Chem. Biol.* **1**, 265–269 (2005).
2. Lentzen, G. & Schwarz, T. Extremolytes: natural compounds from extremophiles for versatile applications. *Appl. Microbiol. Biotechnol.* **72**, 623–634 (2006).
3. Ro, D. K. *et al.* Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **440**, 940–943 (2006).
4. Brady, S. F., Simmons, L., Kim, J. H. & Schmidt, E. W. Metagenomic approaches to natural products from free-living and symbiotic organisms. *Nat. Prod. Rep.* **26**, 1488–1503 (2009).
5. Scherlach, K. & Hertweck, C. Triggering cryptic natural product biosynthesis in microorganisms. *Org. Biomol. Chem.* **7**, 1753–1760 (2009).
6. Gottelt, M., Kol, S., Gomez-Escribano, J. P., Bibb, M. & Takano, E. Deletion of a regulatory gene within the *cpk* gene cluster reveals novel antibacterial activity in *Streptomyces coelicolor* A3(2). *Microbiology* **156**, 2343–2353 (2010).
7. Starcevic, A. *et al.* ClustScan: an integrated program package for the semi-automatic annotation of modular biosynthetic gene clusters and *in silico* prediction of novel chemical structures. *Nucleic Acids Res.* **36**, 6882–6892 (2008).
8. Weber, T. *et al.* CLUSEAN: a computer-based framework for the automated analysis of bacterial secondary metabolite biosynthetic gene clusters. *J. Biotechnol.* **140**, 13–17 (2009).
9. Li, M. H., Ung, P. M., Zajkowski, J., Garneau-Tsodikova, S. & Sherman, D. H. Automated genome mining for natural products. *BMC Bioinformatics* **10**, 185 (2009).
10. Zerikly, M. & Challis, G. L. Strategies for the discovery of new natural products by genome mining. *Chembiochem.* **10**, 625–633 (2009).
11. Bumpus, S. B., Evans, B. S., Thomas, P. M., Ntai, I. & Kelleher, N. L. A proteomics approach to discovering natural products and their biosynthetic pathways. *Nature Biotech.* **27**, 951–956 (2009).
12. Purnick, P. E. & Weiss, R. The second wave of synthetic biology: from modules to systems. *Nature Rev. Mol. Cell Biol.* **10**, 410–422 (2009).
13. Khalil, A. S. & Collins, J. J. Synthetic biology: applications come of age. *Nature Rev. Genet.* **11**, 367–379 (2010).
14. Lu, T. K., Khalil, A. S. & Collins, J. J. Next-generation synthetic gene networks. *Nature Biotech.* **27**, 1139–1150 (2009).
15. Farmer, W. R. & Liao, J. C. Improving lycopene production in *Escherichia coli* by engineering metabolic control. *Nature Biotech.* **18**, 533–537 (2000).
16. Holtz, W. J. & Keasling, J. D. Engineering static and dynamic control of synthetic pathways. *Cell* **140**, 19–23 (2010).
17. Galvao, T. C. & de Lorenzo, V. Transcriptional regulators à la carte: engineering new effector specificities in bacterial regulatory proteins. *Curr. Opin. Biotechnol.* **17**, 34–42 (2006).
18. Tang, S. Y., Fazelinia, H. & Cirino, P. C. AraC regulatory protein mutants with altered effector specificity. *J. Am. Chem. Soc.* **130**, 5267–5271 (2008).
19. Dixon, N. *et al.* Reengineering orthogonally selective riboswitches. *Proc. Natl Acad. Sci. USA* **107**, 2830–2835 (2010).
20. Wagner, A. Energy costs constrain the evolution of gene expression. *J. Exp. Zool. B Mol. Dev. Evol.* **308**, 322–324 (2007).
21. Zaslaver, A. *et al.* Just-in-time transcription program in metabolic pathways. *Nature Genet.* **36**, 486–491 (2004).
22. Sattely, E. S., Fischbach, M. A. & Walsh, C. T. Total biosynthesis: *in vitro* reconstitution of polyketide and nonribosomal peptide pathways. *Nat. Prod. Rep.* **25**, 757–793 (2008).
23. Chechik, G. *et al.* Activity motifs reveal principles of timing in transcriptional control of the yeast metabolic network. *Nature Biotech.* **26**, 1251–1259 (2008).
24. Ellis, T., Wang, X. & Collins, J. J. Diversity-based, model-guided construction of synthetic gene networks with predicted functions. *Nature Biotech.* **27**, 465–471 (2009).
25. Malmberg, L. H. & Hu, W. S. Identification of rate-limiting steps in cephalosporin C biosynthesis in *Cephalosporium acremonium*: a theoretical analysis. *Appl. Microbiol. Biotechnol.* **38**, 122–128 (1992).
26. Fung, E. *et al.* A synthetic gene-metabolic oscillator. *Nature* **435**, 118–122 (2005).
27. Dano, S., Madsen, M. F. & Sorensen, P. G. Quantitative characterization of cell synchronization in yeast. *Proc. Natl Acad. Sci. USA* **104**, 12732–12736 (2007).
28. Danino, T., Mondragon-Palomino, O., Tsimring, L. & Hasty, J. A synchronized quorum of genetic clocks. *Nature* **463**, 326–330 (2010).
29. Nieselt, K. *et al.* The dynamic architecture of the metabolic switch in *Streptomyces coelicolor*. *BMC Genomics* **11**, 10 (2010).
30. Alam, M. T. *et al.* Metabolic modeling and analysis of the metabolic switch in *Streptomyces coelicolor*. *BMC Genomics* **11**, 202 (2010).
31. Conrado, R. J., Varner, J. D. & DeLisa, M. P. Engineering the spatial organization of metabolic enzymes: mimicking nature's synergy. *Curr. Opin. Biotechnol.* **19**, 492–499 (2008).
32. Llopis, P. M. *et al.* Spatial organization of the flow of genetic information in bacteria. *Nature* **466**, 77–81 (2010).
33. Dueber, J. E. *et al.* Synthetic protein scaffolds provide modular control over metabolic flux. *Nature Biotech.* **27**, 753–759 (2009).
34. Moon, T. S., Dueber, J. E., Shiue, E. & Prather, K. L. Use of modular, synthetic scaffolds for improved production of glucaric acid in engineered *E. coli*. *Metab. Eng.* **12**, 298–305 (2010).
35. Menzella, H. G. *et al.* Redesign, synthesis and functional expression of the 6-deoxyerythronolide B polyketide synthase gene cluster. *J. Ind. Microbiol. Biotechnol.* **33**, 22–28 (2006).
36. Menzella, H. G. *et al.* Combinatorial polyketide biosynthesis by *de novo* design and rearrangement of modular polyketide synthase genes. *Nature Biotech.* **23**, 1171–1176 (2005).

© 2011 Macmillan Publishers Limited. All rights reserved

37. Menzella, H. G., Carney, J. R. & Santi, D. V. Rational design and assembly of synthetic trimodular polyketide synthases. *Chem. Biol.* **14**, 143–151 (2007).
38. Straight, P. D., Fischbach, M. A., Walsh, C. T., Rudner, D. Z. & Kolter, R. A singular enzymatic megacomplex from *Bacillus subtilis*. *Proc. Natl Acad. Sci. USA* **104**, 305–310 (2007).
39. Evers, M. E., Trip, H., van den Berg, M. A., Bovenberg, R. A. & Driessen, A. J. Compartmentalization and transport in β-lactam antibiotics biosynthesis. *Adv. Biochem. Eng. Biotechnol.* **88**, 111–135 (2004).
40. Chanda, A. *et al.* A key role for vesicles in fungal secondary metabolism. *Proc. Natl Acad. Sci. USA* **106**, 19533–19538 (2009).
41. Sirikantaramas, S., Yamazaki, M. & Saito, K. Mechanisms of resistance to self-produced toxic secondary metabolites in plants. *Phytochem. Rev.* **7**, 467–477 (2007).
42. Bayer, T. S. *et al.* Synthesis of methyl halides from biomass using engineered microbes. *J. Am. Chem. Soc.* **131**, 6508–6515 (2009).
43. Roodbeen, R. & van Hest, J. C. Synthetic cells and organelles: compartmentalization strategies. *Bioessays* **31**, 1299–1308 (2009).
44. Murat, D., Quinlan, A., Vali, H. & Komeili, A. Comprehensive genetic dissection of the magnetosome gene island reveals the step-wise assembly of a prokaryotic organelle. *Proc. Natl Acad. Sci. USA* **107**, 5593–5598 (2010).
45. Medema, M. H. *et al.* The sequence of a 1.8-Mb bacterial linear plasmid reveals a rich evolutionary reservoir of secondary metabolic pathways. *Genome Biol. Evol.* **2**, 212–224 (2010).
46. Fan, C. *et al.* Short N-terminal sequences package proteins into bacterial microcompartments. *Proc. Natl Acad. Sci. USA* **107**, 7509–7514 (2010).
47. Parsons, J. B. *et al.* Synthesis of empty bacterial microcompartments, directed organelle protein incorporation, and evidence of filament-associated organelle movement. *Mol. Cell* **38**, 305–315 (2010).
48. Cai, F. *et al.* The pentameric vertex proteins are necessary for the icosahedral carboxysome shell to function as a CO_2 leakage barrier. *PLoS ONE* **4**, e7521 (2009).
49. Brenner, K., You, L. & Arnold, F. H. Engineering microbial consortia: a new frontier in synthetic biology. *Trends Biotechnol.* **26**, 483–489 (2008).
50. Weber, W., Daoud-El Baba, M. & Fussenegger, M. Synthetic ecosystems based on airborne inter- and intrakingdom communication. *Proc. Natl Acad. Sci. USA* **104**, 10435–10440 (2007).
51. Stubblefield, B. A. *et al.* Constructing multispecies biofilms with defined compositions by sequential deposition of bacteria. *Appl. Microbiol. Biotechnol.* **86**, 1941–1946 (2010).
52. Kim, H. J., Boedicker, J. Q., Choi, J. W. & Ismagilov, R. F. Defined spatial structure stabilizes a synthetic multispecies bacterial community. *Proc. Natl Acad. Sci. USA* **105**, 18188–18193 (2008).
53. You, L., Cox, R. S. 3rd, Weiss, R. & Arnold, F. H. Programmed population control by cell–cell communication and regulated killing. *Nature* **428**, 868–871 (2004).
54. Balagadde, F. K. *et al.* A synthetic *Escherichia coli* predator–prey ecosystem. *Mol. Syst. Biol.* **4**, 187 (2008).
55. Shou, W., Ram, S. & Vilar, J. M. Synthetic cooperation in engineered yeast populations. *Proc. Natl Acad. Sci. USA* **104**, 1877–1882 (2007).
56. Kealey, J. T., Liu, L., Santi, D. V., Betlach, M. C. & Barr, P. J. Production of a polyketide natural product in nonpolyketide-producing prokaryotic and eukaryotic hosts. *Proc. Natl Acad. Sci. USA* **95**, 505–509 (1998).
57. Watanabe, K. *et al.* Total biosynthesis of antitumor nonribosomal peptides in *Escherichia coli*. *Nature Chem. Biol.* **2**, 423–428 (2006).
58. Pfeifer, B. A., Admiraal, S. J., Gramajo, H., Cane, D. E. & Khosla, C. Biosynthesis of complex polyketides in a metabolically engineered strain of *E. coli*. *Science* **291**, 1790–1792 (2001).
59. Boghigian, B. A. & Pfeifer, B. A. Current status, strategies, and potential for the metabolic engineering of heterologous polyketides in *Escherichia coli*. *Biotechnol. Lett.* **30**, 1323–1330 (2008).
60. Zha, W., Rubin-Pitel, S. B., Shao, Z. & Zhao, H. Improving cellular malonyl-CoA level in *Escherichia coli* via metabolic engineering. *Metab. Eng.* **11**, 192–198 (2009).
61. Keasling, J. D. Synthetic biology for synthetic chemistry. *ACS Chem. Biol.* **3**, 64–76 (2008).
62. Steen, E. J. *et al. Microbial* production of fatty-acid-derived fuels and chemicals from plant biomass. *Nature* **463**, 559–562 (2010).
63. Rocha, I. *et al.* OptFlux: an open-source software platform for *in silico* metabolic engineering. *BMC Syst. Biol.* **4**, 45 (2010).
64. Adrio, J. L. & Demain, A. L. Genetic improvement of processes yielding microbial products. *FEMS Microbiol. Rev.* **30**, 187–214 (2006).
65. Posfai, G. *et al.* Emergent properties of reduced-genome *Escherichia coli*. *Science* **312**, 1044–1046 (2006).
66. Komatsu, M., Uchiyama, T., Omura, S., Cane, D. E. & Ikeda, H. Genome-minimized *Streptomyces* host for the heterologous expression of secondary metabolism. *Proc. Natl Acad. Sci. USA* **107**, 2646–2651 (2010).
67. Gibson, D. G. *et al.* Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. *Science* **319**, 1215–1220 (2008).
68. Lartigue, C. *et al.* Creating bacterial strains from genomes that have been cloned and engineered in yeast. *Science* **325**, 1693–1696 (2009).
69. Gibson, D. G. *et al.* Creation of a bacterial cell controlled by a chemically synthesized genome. *Science* **329**, 52–56 (2010).
70. Carlson, R. The changing economics of DNA synthesis. *Nature Biotech.* **27**, 1091–1094 (2009).
71. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods* **6**, 343–345 (2009).
72. Walsh, C. T. & Fischbach, M. A. Natural products version 2.0: connecting genes to molecules. *J. Am. Chem. Soc.* **132**, 2469–2493 (2010).
73. Fischbach, M. A., Walsh, C. T. & Clardy, J. The evolution of gene collectives: How natural selection drives chemical innovation. *Proc. Natl Acad. Sci. USA* **105**, 4601–4608 (2008).
74. Olano, C., Mendez, C. & Salas, J. A. Post-PKS tailoring steps in natural product-producing actinomycetes from the perspective of combinatorial biosynthesis. *Nat. Prod. Rep.* **27**, 571–616 (2010).
75. Jenke-Kodama, H., Borner, T. & Dittmann, E. Natural biocombinatorics in the polyketide synthase genes of the actinobacterium *Streptomyces avermitilis*. *PLoS Comput. Biol.* **2**, e132 (2006).

Acknowledgements
We thank D. Hopwood and C. Voigt for constructive comments and suggestions. This work was supported by the Dutch Technology Foundation (STW), which is the applied-science division of The Netherlands Organisation for Scientific Research (NWO) and the Technology Programme of the Ministry of Economic Affairs (grant STW 10463). R.B. is supported by an NWO–Vidi fellowship, and E.T. is supported by a Rosalind Franklin Fellowship, University of Groningen, the Netherlands.

Competing interests statement
The authors declare competing financial interests, see Web Version for details.

FURTHER INFORMATION
Eriko Takano's homepage: http://www.rug.nl/staff/e.takano/index
Rainer Breitling's homepage: http://tinyurl.com/rbreitling
ALL LINKS ARE ACTIVE IN THE ONLINE PDF

SCIENCE AND SOCIETY

Health biotechnology innovation on a global stage

Halla Thorsteinsdóttir, Monali Ray, Andrew Kapoor and Abdallah S. Daar

Abstract | With increasing globalization, infectious diseases are spreading faster than ever before, creating an urgent need for international collaboration. The rise of emerging economies has changed the traditional collaborative landscape and provided opportunities for more diverse models of collaboration involving developing countries, including North–South, South–South and North–South–South partnerships. Here, we discuss how developing countries can partner with other nations to address their shared health problems and to promote innovation. We look specifically at what drives collaborations and at the challenges that exist for them, and we propose actions that can strengthen these partnerships.

With increasing globalization, pathogens that cause diseases can spread swiftly throughout the world, creating an urgent need for collaboration among nations. HIV/AIDS, influenza (including avian and swine influenzas) and severe acute respiratory syndrome (SARS) have all spread rapidly across national borders. In addition, chronic illnesses such as heart disease, diabetes and cancer, some of which are caused by microorganisms[1], have also become diseases of poverty and are on the rise in low-income populations in developing countries[2,3]. We can no longer view the health problems of developing countries as fundamentally different from those found in high-income nations, and so addressing these problems requires a global approach. To address shared health problems requires investment in research and innovation, as well as active contributions by all affected countries. Modern transportation and communications systems make it ever easier for researchers and entrepreneurs to collaborate wherever new opportunities arise. Specific expertise, exciting research material

© 2011 Macmillan Publishers Limited. All rights reserved