

Control of metabolite efflux in microbial cell factories: current advances and future prospects

^{1,2,3,4,*} *Douglas B. Kell*

¹School of Chemistry, ²The Manchester Institute of Biotechnology, ³SYNBIOCHEM, The University of Manchester, 131, Princess St, Manchester M1 7DN, UK.

⁴Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Building 220, DK-2800 Kgs Lyngby, Denmark

*corresponding author dbk@manchester.ac.uk

Keywords: transporters – productivity – flux control – SLCs – transporter engineering

CONTENTS

Control of metabolite efflux in microbial cell factories: current advances and future prospects	1
Keywords	2
Abstract	2
Introduction	3
Membrane transporters	3
Systems biology of metabolic networks	5
A large percentage of orphans, even in well-studied organisms	6
Flux control and metabolic control analysis	6
Role of efflux transporters in metabolic flux control.....	7
Amino acids	7
Fatty acids	7
Riboflavin	8
Why would microbes excrete expensively produced biochemicals?	8
Simple reversibility of facilitated diffusion transporters	8
Osmotic regulation	8
Efflux of substances that assist the uptake of other desirable substances.....	8
Overcoming externally generated sources of toxicity	9

Intercellular signalling: pheromones and quorum sensing	9
Symbiosis and syntrophy	9
A note on permeability units and terminology	10
Finding transporters for a given substrate or product.....	10
The importance of QSARs (Quantitative Structure-Activity Relationships) for drug transporters	13
Transporter bioengineering.....	15
Promoter engineering	15
Control factor expression engineering.....	15
Directed evolution: changing the sequence of the target transporter protein(s)	16
Synthetic biology for efflux transporter engineering	16
Concluding and summarising remarks.....	18
Legends to figures	19
Acknowledgments	19
References.....	20

Keywords

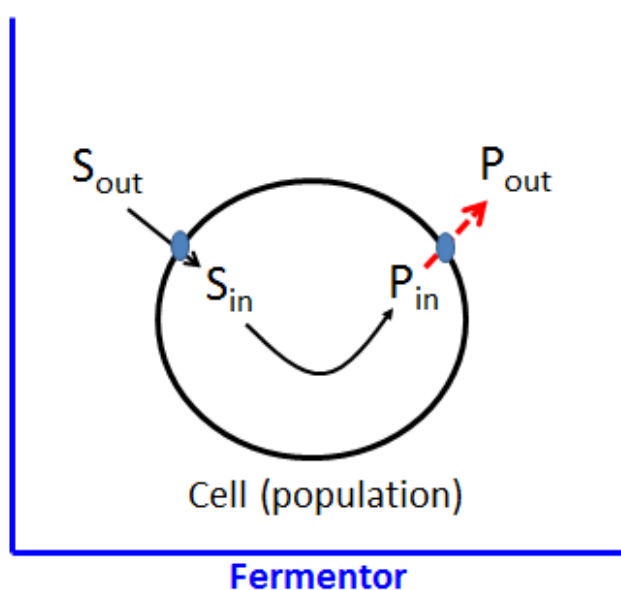
Transporters – productivity – deorphanisation – cheminformatics – efflux – pumps – functional genomics – systems biology

Abstract

The yield and ease of purification of biotechnological products are typically greatly enhanced if they can be secreted into the external medium. Although not universally appreciated, however, small molecule biotechnological products do not ‘float across’ the phospholipid bilayer portion of biological membranes, and thus they need transporters to assist their passage into the extramembrane and extracellular spaces. Some of these transporters may be reversible, equilibrative transporters that might more normally be used for uptake, while others may have an efflux directionality imposed on them via suitable energy coupling mechanisms. Despite the energetic costs of small molecule synthesis, natural evolution does in fact provide a number of mechanisms by which the secretion of such products can actually enhance fitness. Where available these provide useful starting points, and assaying for such activities is crucial. In particular, in a systems biology approach, we first need to identify such/suitable activities before we can seek to increase them. They may then be improved through promoter engineering, via manipulation of control elements, or by directed evolution of the transporter proteins themselves. Modern methods of synthetic biology provide enormous opportunities for all kinds of efflux transporter engineering; they are just beginning to be realised.

Introduction

In a typical fermentation, substrates are provided externally to the cells catalysing the fermentation, and converted to products; initially (at least in most cases) product formation occurs within the cells. However, for a relative density (specific gravity) of 1, cells occupy 1mL.g^{-1} wet weight so for a fermentation that achieves even $100\text{ mg wet cell wt.mL}^{-1}$ only approximately one tenth of the total volume is intracellular (Fig 1). Commonly, intracellular concentrations during the development of processes for industrial bioengineering can soon become toxic [1, 2]. Since thermodynamics dictates that it is standard free energies and concentrations that control the eventual outcome, we need to recognise (Fig 1) that outwith a solid substrate fermentation in which the biomass is the product (as in [3]) the overall titre of product will be much enhanced if its internal concentrations can be decreased by secretion into the larger extracellular space. In addition, it is much easier to purify products if cells are



not present [4]. Hence the desirability for product efflux, that is the focus of this chapter. Although much the same can be said of protein production, we focus here solely on small molecules. An overview of the chapter is given as a mind map [5] in Fig 2.

Figure 1. *The ability to efflux an intracellular product is of great significance in biotechnology, as the extracellular space is normally much greater than the intracellular space in a typical fermentation.*

Membrane transporters

Although this is textbook material, we first rehearse briefly the different types of membrane transporter. A first distinction is whether they are equilibrative (i.e. permitting or catalysing 'facilitated diffusion') or whether their activities are coupled to an external free energy source such as ATP (hydrolysis) or electrochemical gradients; the latter kinds of transporter may then be concentrative in terms of changing the transmembrane ratio of their substrate concentrations (properly, activities) away from 1. Those lacking secondary coupling are referred to as uniporters, symporters cotransport co-substrates, while antiporters act to exchange substrates in opposite directions in a coupled manner. 'Group transfer' reactions involve the direct coupling of a chemical motif to the substrate, as in the PEP-dependent glucose transferases [6, 7] whose external substrate is glucose but whose internal product is glucose-6-phosphate.

We also note here the use and misuse of the term ‘passive’ to describe transport activities; this term has two common but orthogonal meanings. The first is thermodynamic, and means ‘equilibrative’, while the second is mechanistic and is then taken to mean ‘transport through a bilayer’. When the mechanism involves a transporter it is properly known as ‘facilitated diffusion’. Since these two uses of ‘passive’ are often conflated, and consequently cause much unnecessary confusion, we recommend that the term ‘passive’ is simply dropped in the context of transporters [8].



Figure 2. A ‘mind map’ [5] of the layout of this paper

In a similar vein, it is common to refer to ‘influx’ and ‘efflux’ transporters on the basis of the direction of substrate flux observed in their most typical operating conditions. Clearly, however, any reaction is in principle thermodynamically reversible (even if free energy changes are large and negative). We note in particular therefore that while a particular transporter might ‘normally’ be an ‘influx’ transporter if its substrate is provided externally, there is no reason of principle, especially if it is equilibrative, why it would not become an ‘efflux’ transporter if large amounts of the same substrate are made intracellularly in a biotechnological process. Consequently, while the focus is on ‘efflux’ transporters, we shall have plenty to say about the more widely studied ‘influx’ transporters as well.

In general, despite their prevalence (transporters account for one eighth of *E. coli* genes [9], and see later), transporters are comparatively little studied [10] and the substrates of many (and in some organisms most) of them are unidentified [11, 12]. Although our focus here is on microorganisms (and see [13]), the situation is also acute for mammalian transporters, as

these underpin the distributions of xenobiotic pharmaceutical drugs within and between cells [8, 14-20].

One possible reason for the comparative lack of study is the widespread (but erroneous) textbook belief that many substances can actually cross the phospholipid portion of biological membranes at non-negligible rates without the use of protein transporters. This belief is based in part on the fact that such ‘transport’ can indeed be observed in liposomes or across ‘black’ (bilayer) lipid membranes [21, 22] that are simply leaky to such molecules; however, this is not the case in real biomembranes, with a typical protein:phospholipid mass ratio of 3:1. Winter and colleagues [23] provide a particularly nice example, showing that at least 99.5% of the mammalian cell uptake of a drug known as sepantronium bromide occurs only via the SLC35F2 transporter, meaning that maximally 0.5% can be transported by any other means, including through the bilayer.

Although most commentators recognise that charged substances require transporters (including for so-called lipophilic cations [24]), we now know that membrane transporters exist (and are required) even for all kinds of small uncharged molecules that were once widely believed to be capable of crossing biomembranes without transporters; Table 1 provides a set of examples. Notably, the discovery of water transporters (aquaporins) attracted a Nobel Prize (see [25] and [26]).

Molecular class	Selected reference(s)
Alkanes	[27, 28]
Ammonia (NH ₃)	[29]
Carbon dioxide (CO ₂)	[30, 31]
Ethanolamine	[32]
Fatty acids	[33, 34]
Glycerol	[35]
Hydrogen peroxide (H ₂ O ₂)	[36]
Hydroxyurea	[37]
Nitric oxide (NO)	[38]
(Di)oxygen (O ₂)	[39]
Urea	[40, 41]
Water	[35, 42]

Table 1. Some small, neutral molecules for which transporters are known.

Since the role of biomembranes in general is to keep things inside and outside of the compartments that they surround, it is easy to imagine that transporters have an important role in the selective influx and efflux of molecules. Consequently they attract our focus within the context of metabolic network or systems biology. Indeed, having the metabolic network is a *sine qua non* for developing such studies, so we briefly rehearse the systems biology approach to metabolic biotechnology.

Systems biology of metabolic networks

To understand metabolic networks it is first necessary to reconstruct them [43]. This is done in four main steps [44, 45]. The first two are qualitative: they define the network in terms of the actors that are involved (enzymes, metabolites and effectors), and the means by which they are connected (mathematically as a ‘graph’). The result of this is the familiar

biochemical wallchart or the kinds of qualitative representation visible in an online resource such as KEGG [46]. The third step defines the kinetic rate equations governing each step, while the fourth step parametrises them. Because of the extreme stoichiometric constraints [43], i.e. the requirement that in the absence of alchemy all atoms are conserved, a surprisingly large amount can be done with the qualitative network alone, and we shall tend to focus on this in the present chapter. Obviously transporter steps are to be seen simply as ‘extra’ enzymatic steps, as in Figure 1. Reconstructions have become quite advanced by now, with strategies including ‘jamborees’ [47-49], helped by the availability of digital tools (e.g. [50]), including for the reconstruction of non-natural pathways for the biosynthesis of ‘novel’ products [51, 52]. In very few cases are all the kinetic rate equations known with any precision (see e.g. [53] for yeast glycolysis), but when they are one may resort to ODE-type modelling, using tools such as Copasi [54, 55]. Another strategy is to use surrogate kinetic rate equations for each step, lin-log being both common and effective [56]. ‘Convenience’ kinetics [57] are also widely used for this.

A large percentage of orphans, even in well-studied organisms

All of this said, however, in many cases we still lack the ‘parts list’ [58]; in other words, we do not know the substrates for substantial fractions of (‘uptake’ or ‘efflux’) membrane transporters, and even when we know some they may not be the most active or ‘natural’ or ‘real’ substrates. The systematic genome sequencing of microbes finally allowed at least a start to be made on the parts list, and increased it considerably. Thus in *E. coli* sequencing caused the number of proposed genes to be increased from ~1700 to over 4000 [59]. This 2006 survey of *E. coli* [59] listed some 4453 genes, of which 591 were transporters (13.3%)(see also [9, 12]). 337 were ‘known’ and 254 (43% of the total) were ‘predicted’ (i.e. orphan) transporters. In 2009, Hu and colleagues [60] found that about one third of *E. coli* genes were functionally orphans. Ecogene (<http://www.ecogene.org/>) (searched 3rd January 2018) listed 326 transporters, of which 113 were ‘y-genes’ (unassigned)(see also [61]). A Table (<https://ecocyc.org/group?id=biocyc17-4655-3682299327>) at EcoCyc [62] gives 163 “putative” transporters, including 124 y-genes.

In 1998, Paulsen *et al.* [63] summarised this for *S. cerevisiae*, concluding that “among the 258 yeast transporters... a total of 139 {35%} lack genetic or biochemical names and thus lack either a demonstrated transport function or a recognizable physiological function”. The latest version of transportDB [64] <http://www.membranetransport.org/transportDB2/overview.html>, the most comprehensive listing, gives 341 transporters for *S. cerevisiae* S288c, and shows that improvements in identification have been less than substantial. In general, the efflux transporters seem to have evolved to remove natural toxins from the yeast’s environment [65, 66], are highly promiscuous [67, 68], and many are known as pleiotropic drug transporters (PDRs) [69].

The general conclusion of this is that at least 30% of transporters (and in many cases a great deal more) are genuine orphans in terms of actual experimental assessment; deorphanisation is thus the key required first step for improving our understanding of them, even in supposedly well-characterised organisms [10, 70].

Flux control and metabolic control analysis

El-Mansi and colleagues give an excellent description in this volume of the metabolic control analysis (/theory/formalism) of Kacser, Burns, Heinrich and Rapoport [71, 72]. Other reviews

can be found in [73-76] and an online tutorial at <http://dbkgroup.org/metabolic-control-analysis/>. Although not at all new, it still provides an excellent formalism for describing the extent (their flux-control coefficients, FCC) to which individual reactions, including transporter reactions, control metabolic fluxes in biochemical networks. We note that those that do have reasonably high FCCs catalyse reactions that in the steady state are held far from their thermodynamic equilibrium, meaning (by definition) that these steps exert regulatory control. Increasing the rate of such steps will by definition increase fluxes. Both influx and efflux transporters tend to have this property (for an early example, see [77]).

Role of efflux transporters in metabolic flux control

Since we reviewed the role of transporters in biotechnology not too long ago [13], with many historical examples, we start with just one main generic example here, followed by a couple not previously discussed by us in any detail.

Amino acids. The classic example of the role of efflux transporters in improving the yield of an important fermentation product (more than 2M tonnes p.a. [78]) is represented by the glutamate fermentation carried out in particular by *Corynebacterium glutamicum* [78-80]. Following the initial discovery of the microbial/fermentative production of glutamate [81], various empirical findings in the 1960s and 1970s [82, 83] showed that a variety of treatments, involving biotin limitation, or the addition of weak surfactants such as 'acetylated corn oil' or Tween, or the use of certain auxotrophs, would enhance the efflux of glutamate in producer strains. Soon enough, however, it was recognised that this was not due to a general membrane-leakiness, because it was very selective for glutamate (and was even against a glutamate concentration gradient!), but that it was due to a change in membrane tension that activated a mechanosensitive glutamate efflux pump encoded by a gene called NCgl1221 (a homologue of the *E. coli* *yggB* gene, now known as *mscS*, the mechansensitive channel of small conductance) [84-89]. Similar efflux pumps are now known to be involved in the export of product during a variety of other amino acid fermentations [79, 90], such as those for lysine [91-95], isoleucine [96, 97], serine [98], threonine [99-102], cysteine [103], methionine [104], alanine [105] and others [103, 106, 107]. Thus, the general role of product exporters in enhancing fermentation yields (as per Figure 1) is clear [2, 103, 108]. We return to this later.

Fatty acids. In contrast to amino acids, it used to be widely believed that both short- and long-chain fatty acids could simply diffuse through (the bilayer portion) of cellular membranes. Since that is not the case (see above) we devote a short section to longer-chain fatty acid uptake and efflux.

In mammalian systems, fatty acids are taken up by members of the SLC27 family [109], also known as fatty acid transport proteins (FATPs), of which there are 6 types with different expression profiles). All FATP members have a highly conserved, 311-amino acid signature sequence known as the FATP sequence.

Many microbes will use fatty acids as sole source of carbon, and thus microbes also contain fatty acid transporters (e.g. [110]), the main one in Gram-negative organisms being known as FadL [33, 34, 111, 112]. Yeasts also contain fatty acid transporters; these are mainly homologues of mammalian FAT1 [113-115]. It is of particular interest that the Fat1p homologue in the oleaginous yeast *Yarrowia lipolytica* is involved in the export of fatty acids [116-119].

Riboflavin. Riboflavin is a vitamin and nutritional supplement that is nowadays produced biosynthetically [120, 121]. It is of interest to us here because a main producer, *Ashbyi gossypii*, contains a very active (influx and) efflux transporter for it [122]. It is also of interest, and a general point worth nothing for fungi, that disruption of its vacuolar uptake system (by disrupting the vacuolar ATPase) ensures that all of the biosynthesised riboflavin is indeed excreted [123]. The other major industrial producer is *B. subtilis* [120], where an efflux transporter from *Streptomyces davawensis* has been used to improve productivity [124]. Of course, since a vitamin is by definition not synthesised therein, all human cells require [125] and express [126] the relevant uptake transporters, in this case of the SLC52 family [127].

Why would microbes excrete expensively produced biochemicals?

In many cases, possibly the majority, efflux transporters are involved in the removal of potentially cytotoxic drugs or xenobiotics that happened to have been taken up by influx transporters, especially in mammalian cells where they represent a major problem e.g. for cancer chemotherapy (e.g. [128-130]). In microbes, where they have also presumably evolved to remove environmental toxins that had been taken up [131], they can play a leading role in microbial resistance to anti-infectives [132-139]. In each of these cases, it is easy to understand why natural evolution would select for such activities.

By contrast, for substances that are actively biosynthesised by the host, it is rather less obvious why they might evolve an efflux transporter for them instead of simply lowering the rate of synthesis to a level that is adequate to satisfy the requirements of the host. Needless to say, evolution has in fact selected for this active efflux, and it is of interest to seek to understand its basis (if only to replicate it in the selection schemes of the biotechnologist). A number of examples can be given.

Simple reversibility of facilitated diffusion transporters. Although there can be an apparent kinetic irreversibility that follows from the Haldane relation [140], any equilibrative transporter might serve as a selective efflux transporter when the intracellular concentration of its substrates begin to exceed those of the extracellular ones. This appears to be the case, for instance, in at least one of the riboflavin examples given above [122]. It probably becomes true for any substance that cells can both take up and biosynthesise if the uptake transporter is not completely turned off in the latter case. Possibly this counts as neutral evolution [141], but by definition if such activities remain then natural selection was involved. To this end, it provides a rationale for adding the relevant equilibrative “influx” transporters to the cells of interest, even when it is efflux that is desired.

Osmotic regulation. Osmotic stress can modulate the activity of membrane transporters (e.g. [77]), and this is the basis for the mechanosensitive efflux pumps e.g. for glutamate (see above). Thus, as in plants [142], the sudden advent of a raindrop or other source of water [143] under previously dry or drier conditions to which the cells had acclimatised can create a huge osmotic stress for a cell [144]; only cells that can respond with an almost instantaneous secretion of internal osmolytes can easily survive this, and one can imagine (as is the case, see above) that these kinds of channels work precisely because they are ‘mechanically’ sensitive to the osmotic stress generated across the membrane..

Efflux of substances that assist the uptake of other desirable substances. Fermentative production of citric acid by the fungus *Aspergillus niger* is a long-standing and very large-scale process [145]. It too involves active export of the product from the producer strain

using a proton-symporting transporter [146]. Here it would appear that in nature both plants and fungi have evolved citrate efflux to serve as a chelator for metal ions that are necessary for the growth of the host [147-149]. Depending on the pH, an equilibrative transporter may be sufficient [150]. Siderophores are of course another well-known class of compounds excreted by plants and microbes to assist their uptake of iron [151-154]. Dicarboxylate efflux transporters are also known [155], potentially (and experimentally) serving similar purposes [148, 156].

Overcoming externally generated sources of toxicity. Microbes are necessarily exposed to many cytotoxic stresses, including both physical stresses (such as UV radiation) and chemicals. This can account for excretion of substances that compete with uptake transporters for uptake of the toxins (just as anti-metabolites can be used to select high producers, see [157]). It is likely that redox-active coloured compounds such as riboflavin are indeed produced for these purposes; the same is true for astaxanthin production by *Phaffia rhodozyma* [158, 159].

Intercellular signalling: pheromones and quorum sensing. “A pheromone is a chemical excreted by an organism into the environment that acts to elicit a specific response from other organisms of the same species. The importance of pheromones in the life cycle of various species of mammals, insects and fungi is well known. In the past decade, it has become apparent that pheromones influence the behavior and development of prokaryotes.” [160]. Pheromones are thus well recognised as molecules that are secreted by organisms in order to elicit physiological responses from their genetic relatives. It is probable that secreted ‘secondary’ metabolites (so named because their distribution between clades is rather restricted [161]) are made by microbes precisely for these purposes [162-165]. This said, some are clearly intended to send signals (including cytotoxic molecules) to other organisms, and many examples (e.g. [166]) are known in which the presence of organism 1 induces organism 2 to synthesise molecules that are normally cryptic. In enclosed environments, a molecule that is secreted and promotes both its own synthesis and secretion in another cell of the same species can achieve a steady-state concentration of the pheromone that depends on the concentration of cells. Consequently, and for historical reasons, such (pheromonal) behaviour has become known, in prokaryote biology, as ‘quorum sensing’ (see e.g. [167-169]). It provides another general example of secretory processes involving small molecules that have selective advantage for the host, and has applications in biotechnology [170, 171].

Symbiosis and syntrophy. As well as signalling molecules, many natural microbial and other ecosystems involve the exchange of nutrients between organisms, often of different types, necessarily involving secretion or metabolite efflux. Thus, lichens are symbioses between algae and fungi [172, 173], and stable consortia can develop even in novel ecosystems by learning to assist each other [174]. Since this is not a review of microbial ecology, our purpose is only to recognise the role of effluxers in complex ecosystems [175]; in some cases the dependence of at least one organism is absolute, for thermodynamic reasons involving the removal by a second organism of a product secreted by the first; this is referred to as syntrophy [176-178].

In microbes that reproduce by binary fission, selection is usually for growth rate [179], and if circumstances can be arranged to select for it, improved efflux may lead to improved growth rate that can be selected for in a turbidostat (e.g. [180, 181]). In conclusion, as ever in

biology, evolutionary considerations can give useful insights into circumstances in which the secretion of a biosynthesised metabolite is, perhaps surprisingly, beneficial for the host.

A note on permeability units and terminology

Biochemists (and modellers) are used to expressing fluxes in terms of units such as $\text{nmol} \cdot (\text{min} \cdot \text{mg protein})^{-1}$. If the volume of the compartment of interest is known this is sometimes rendered as $\text{nM} \cdot \text{min}^{-1}$ or its dimensional equivalent. However, the convention in transporter studies, especially those in epithelial cells such as Caco-2 cells, is to express such rates in terms of so-called permeability coefficients (P_{app}), with the dimensionality $(\text{cm}) \cdot \text{s}^{-1}$. (Many pharmaceutical drugs have values in Caco-2 cells in the range $1\text{--}10 \text{ cm} \cdot \text{s}^{-1}$ when expressed in these units [182].) Transport efficiency is defined [183, 184] as the initial rate of uptake ($\mu\text{mol} \cdot \text{min}^{-1}$) (or $\mu\text{mol} \cdot (\text{min} \cdot \text{mg})^{-1}$) divided by substrate concentration ($\text{nmol} \cdot \text{cm}^{-3}$). The former thus has the dimensionality of the usual measure of enzyme kinetic power (k_{cat}/K_m) [185] and the SI dimensions of $\text{m}^3 \cdot \text{t}^{-1}$, and when this is divided by an estimate of the membrane area in $(\text{cm})^2$ we obtain the (apparent) permeability P_{app} with the dimensions $\text{length} \cdot \text{time}^{-1}$ (see also [186]).

Finding transporters for a given substrate or product

In line with our ‘four-stage’ systems biology strategy, the first step is to establish qualitatively which transporters might be responsible for the transport of particular substrates (in shorthand: ‘**have substrate, seek transporter(s)**’). A different but related question, albeit with a similar endpoint, (‘**have transporter, seek substrate(s)**’), enquires as to which substrates are used by a given transporter. We shall consider both. One strategy (see e.g. [103]) for the former is purely *in silico*, and uses sequence similarity to suggest transporter substrates. However, this is not particularly reliable as most transporters have undergone divergent evolution [187] such that moderately similar sequences or motifs can have considerably different substrates. (Note that efflux systems do seem to be more strongly conserved than are influx systems [188], and are often highly promiscuous [68, 134].)

A more reliable method for both questions seeks to assess a co-variation of uptake with the loss or gain of a transporter activity. This is particularly convenient if there can be a (growth) selection step, since if the transporter is not essential for cell growth, and sufficient substrate can be added that it is toxic, cells that survive the presence of normally cytotoxic concentrations of the substrate may do so because they lack the relevant uptake transporter (e.g. [23, 157]) or have increased activity of an efflux transporter (e.g. [132, 133, 135, 136, 189–192]) or a factor controlling its expression (see below). This then allows their identification, nowadays typically by sequencing survivors directly.

In the former (uptake) case, we [157] exploited the availability [193] of a knockout collection of bar-coded, haploid cell lines in baker’s yeast, and assessed those exhibiting resistance to substrates at concentration capable of decreasing the wild-type growth rate by 90%. These could easily be observed in 18/26 cases, the others being considered to have multiple uptake transporters such that deleting them individually was without major effect (or any such effects were subsumed by pleiotropy [194, 195]). The parallel analyses of the barcoded mutants were mirrored by a robotic analysis of individual mutant strains.

Superti-Furga and colleagues [23] used a similar but more elegant strategy, exploiting the availability of a (nearly entirely) haploid mammalian cell line KBM7 [196] into whose genome a retrovirus could insert randomly, thereby knocking out a particular gene. They challenged

the cells with a cytotoxic concentration of the candidate anticancer drug sepantronium bromide (YM155), noting that on 122/122 occasions the gene that was knocked out to enable cell survival was an uptake transporter known as SLC35F2. These knockouts were some 500 times more resistant than was the wild type, and the sensitivity of different (diploid) cell lines correlated strongly and negatively ($P=0.0007$) with SLC35F2 expression levels ($R = -0.77$) as judged by transcriptome analysis.

Classical example using untargeted metabolomics

Discovery of the ergothioneine transporter

Dirk Gründemann^{*,†}, Stephanie Harlfinger^{*}, Stefan Golz[‡], Andreas Geerts[§], Andreas Lazar^{*}, Reinhard Berkels^{*,§}, Norma Jung[‡], Andrea Rubbert[‡], and Edgar Schömig^{*,†}

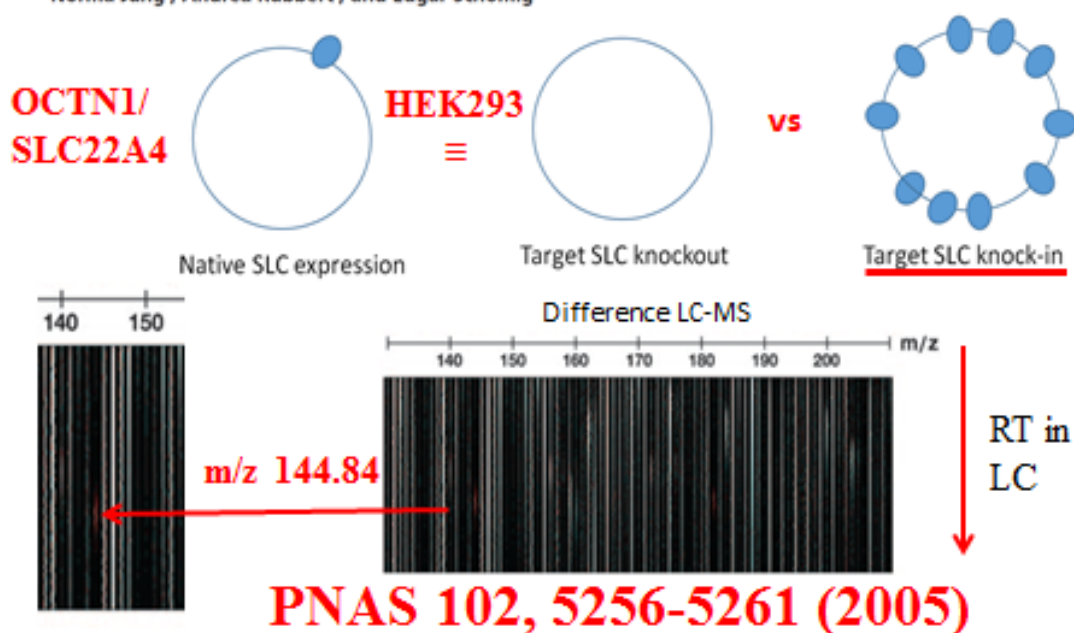


Figure 3. An untargeted metabolomics strategy for determining the substrates of 'orphan' transporters whose 'true' substrates are considered not to be known. Based on [184]

These strategies work well when there is a screening (or better selection) step available for cells that survive a specific treatment, or can at least be differentiated from the rest, e.g. by flow cytometry (see e.g. [159, 197] and below). A more general strategy (Fig 3) is to expose different cell lines, in which the transporter of interest is differentially expressed (including naturally [126]), to a large cocktail of substrates, and see which substrates are differentially taken up. These may then be tested individually. A classic example is that of Gründemann and colleagues [184, 198, 199], who were interested in the function of a transporter previously named OCTN1 (organic cation transporter N1), now referred to as SLC22A4 [200]. It was at the time alleged to be a transporter of the cations carnitine and (the unnatural substrate) tetraethyl ammonium. However, Gründemann and colleagues [184] recognised that the rates of uptake of these substrates were in fact rather miserable (see also [182]). They used what would now (see e.g. [201-205]) be referred to as 'untargeted metabolomics' to determine the differential uptake of substances from pooled serum when it was incubated with HEK293 cells either lacking measurable amounts, or containing cloned-up levels, of SLC22A4. A specific mass of m/z 144.84 was detected as being particularly taken up in the transporter-containing cells, and this was identified as the dipeptide proline betaine (aka

stachydrine), a characteristic constituent of citrus fruits and their juices [206-208]. Gründemann and colleagues [184] then performed some elementary cheminformatics (as in [209]) to look for other molecules that were chemically (structurally) similar to stachydrine, recognising that among them was ergothioneine (also known as 2-mercaptohistidine trimethylbetaine; IUPAC name (2S)-3-(2-Thioxo-2,3-dihydro-1H-imidazol-4-yl)-2-(trimethylammonio)propanoate) (Fig 4). When tested individually, ergothioneine turned out to be by far the best substrate, being taken up ~100x more efficiently than was carnitine, and being accumulated ca 180-fold (its uptake was Na^+ -coupled, presumably with a stoichiometry of at least 2Na^+ per ergothioneine). Although not our subject here, the biology of ergothioneine (an antioxidant) is very interesting (see e.g. [126, 199, 210-212]), with SLC22A4^{-/-} cells being very sensitive to oxidative stress, but the key point is that this analysis shows the likelihood that many transporters were actually selected on the basis of their ability to take up dietary, bioactive compounds that are effectively xenobiotics [213].

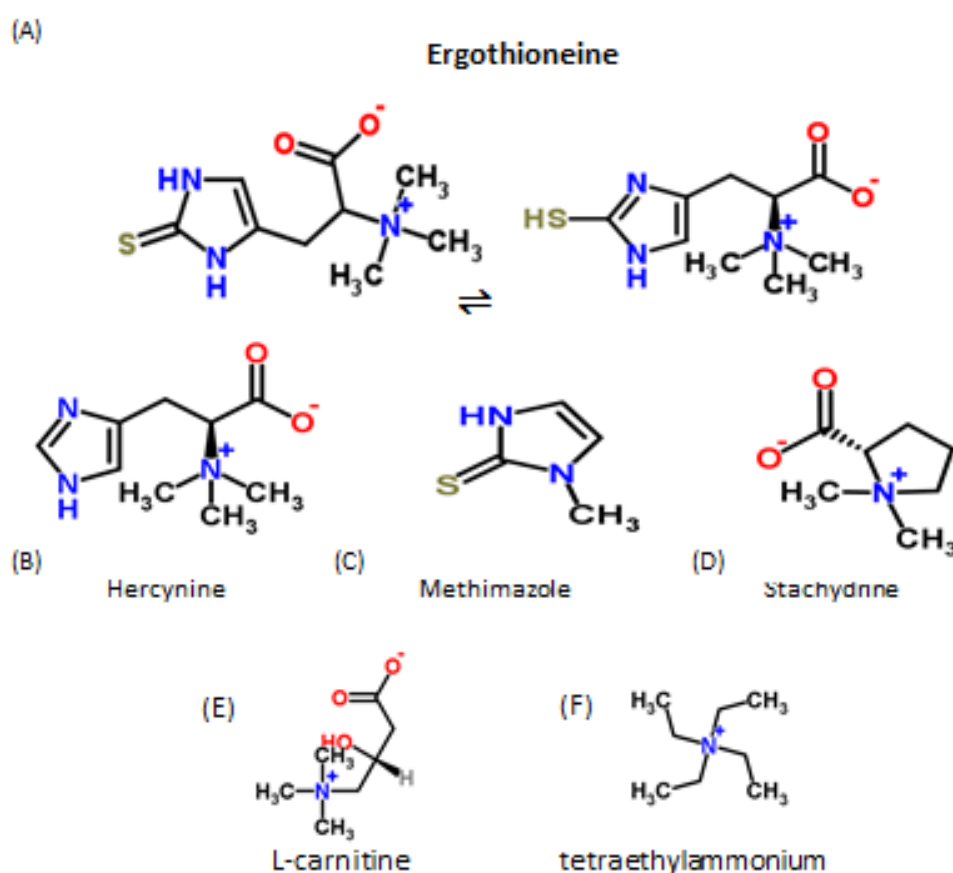


Figure 4. The structures of L-carnitine, ergothioneine, and some related molecules.

Cloning large libraries of chromosome fragments [94, 214], and seeing their effects on export or productivity, or assessing genes with increased transcription in response to increased external metabolite concentrations [108, 215] provide for other methods (see e.g. [103]).

Competition assays provide yet another means of assessing which transporters take up which drugs if an uptake assay is available for at least one, known substrate. Traditionally, radio-isotopically labelled drugs would be used, but nowadays fluorescent assays (e.g. [216-220]) or mass spectrometric assays [221] are more common.

Finally, a major trend for the detection of the production or uptake of a substance, especially in microbes, is the development of screening or selection based on various types of biosensors (e.g. [12, 197, 222-236]). These powerful techniques, that give optical readouts amenable to cell sorting techniques [159], are now becoming much more common.

The importance of QSARs (Quantitative Structure-Activity Relationships) for drug transporters

Ultimately, we would like some kind of a mathematical model that could predict the rate of uptake of any small molecule by a particular transporter molecule (i.e. its turnover number k_{cat}), with the total maximum uptake rate per cell or organelle at saturating external concentration (V_{max}) being given by the product of its expression level and the k_{cat} . We note, of course, that transporter kinetics are particularly complex, with strong interactions between internal and external concentrations of the various substrates and inhibitors. However, the usual practice is to perform experiments under initial rate conditions, and to assume (or determine the) Michaelis-Menten kinetics. Given at least one molecule that is a substrate, it is possible (as with the example of Gründemann and colleagues [184] mentioned above) to assess structurally related molecules, on the basis (the 'molecular similarity' principle, e.g. [237]) that structurally similar molecular will tend to have similar effects.

This is not a review of cheminformatics (see e.g. [238, 239]), but, in brief, a similarity comparison is typically done as follows. The structure of the interrogating molecule, provided in the form of a SMILES [240] or InChI [241-243] string, is encoded therefrom as a string of 1s and 0s. Modern software suites such as RDKit [244] (www.rdkit.org/), CDK [245] or KNIME (which includes them both) [246-248] allow one to do this automatically. The same is done for the molecule(s) with which the interrogating molecule is to be compared. Although more complex comparisons are occasionally used (e.g. [249-251]), strings are typically compared on the basis of the number of bits they have in common relative to the total, a true metric (between 0 and 1) known as the Jaccard or Tanimoto similarity. Although this is a continuous function, Tanimoto similarities above 0.8 are commonly found (i) to be resistant to the precise encoding used, and (ii) indeed to have broadly similar effects in most cases (review at [213]). An alternative encoding uses calculated molecular 'descriptors' or parametrised properties [252] such as polarity, number of hydrogen bond donors, and so on (CDK includes 22, for instance [253]). Of course fingerprints can also be combined with descriptors. Note, however, that 'pure' structural similarity analyses do not take any pharmacological activities into account, and such methods are referred to as 'unsupervised' learning methods (Figure 5), an important subset of which includes clustering methods (see e.g. [253-256]).

QSAR analysis counts as a 'supervised' method (Figure 5), in which a set of candidate drugs, that may be chosen 'actively' on the basis of previous knowledge, are assessed for their potency in an assay of interest, e.g. the determination of the k_{cat} of a transporter for which they may be a substrate or the K_i if they are inhibitors. This 'potency' (or whatever it is that we are trying to predict) is known as the output or objective function. The result of a series of such measurements using various drugs is a set of paired values of structure (encoded as a bitstring of 1s and 0s as described above) and the value of the objective function. Any number of the modern methods of machine learning may then be used to

construct a nonlinear mapping of these inputs onto the output; in [248] we compared three, viz a multivariate linear statistical method known as partial least squares [257], a potent general method known as random forests [258], and an evolutionary algorithm known as genetic programming (see e.g. [259-262]). Overtraining is avoided by testing the model on a subset of molecules not used in the training; if successful, the model is said to generalise, and can with reasonable confidence (but see [263-265]) be used to make predictions *in silico* on any unseen molecules (that can, of course, themselves be tested experimentally). Possibly the most extensive uptake transporter studies (e.g. [266, 267]) have been carried out on SLC15 [268] members (previously known as PEPT1 and PEPT2), and responsible e.g. for the intestinal uptake of penicillins [269] and cephalosporins [270, 271].

Unsupervised and Supervised Learning methods

Unsupervised

- No objective function
- Typically used in assessing similarities
- Related to clustering methods
- Not susceptible to overtraining, though many algorithms are not deterministic

Supervised

- Objective (output) function used
- Trained with paired inputs and outputs
- Can (if done properly) generalise and extrapolate
- Can be susceptible to overtraining

Figure 5. *Unsupervised and supervised learning methods in cheminformatics*

As mentioned, vitamin transporters, by definition, are required by essentially all mammalian cells (and are expressed therein; see e.g. data in [126]). Another example here is provided by Ray and colleagues [272, 273], who noted that a variety of molecules, including Janus kinase (JAK) inhibitors such as fedratinib, were transported by and/or were inhibitors of two of the three main thiamine transporters (SLC19 family [274]), thereby inducing Wernicke's encephalopathy (due to thiamine deficiency). A kind of QSAR model (in this case a pharmacophore model) suggested that 2,4-diaminopyrimidine-containing compounds can in fact adopt a conformation matching several key features of thiamine. This led to the discovery that the antibiotic trimethoprim also potently inhibits thiamine uptake [272]. Metformin is also both a substrate and an inhibitor of one of the transporters [275].

In general, provided we know the QSAR of relevant transporters and their expression profiles, we can expect to be able to predict their utility in the cellular uptake or efflux of any substrates of interest whose structure is known and can thus be encoded in a similar way to that used in the construction of the model.

Transporter bioengineering

Armed with knowledge of an (uptake or) efflux transporter we wish to incorporate or improve in our host organism, three main non-empirical strategies for transporter engineering can be identified:

- Varied (usually increased) expression by promoter engineering
- Varied (usually increased) expression by other control factor (e.g. transcription factor) engineering
- Varied (usually increased) activity per molecule via directed evolution or synthetic biology, this also potentially including varying substrate specificities

We consider each in turn.

Promoter engineering

This is nowadays very well known, and does not demand much space. Well-established inducible promoter systems include various forms of tet [276, 277], while in *E. coli* IPTG is still widely used at small scales. Varying the promoter sequence can allow an almost continuous tuning of expression levels [278-282], even in difficult high-G+C-containing actinobacteria [283, 284].

Control factor expression engineering

As well as promoter sequences, a very great many other features can determine the steady-state expression level of a target protein of interest (e.g. [285-288]). These include mRNA stability [289, 290], codon usage [291-294], riboswitches [12, 295], ribosome binding site potency [296, 297] and even the RNA polymerase itself [298] (including photoswitchable variants [299]).

By definition, transcription factors can have major effects on the expression of multiple genes and hence pathway fluxes, a particularly clear example being the role of the myb transcription factor driving phenylpropanoid and anthocyanin biosynthesis in plants [300, 301]. Transcription factor engineering rather lags behind in prokaryotes (see e.g. [197, 302, 303]), but an example of present interest is the use of marA to improve solvent tolerance (geraniol) in *E. coli* [304], as this acts, at least in part, by increasing the expression level of the enormous (770kDa) [305] and otherwise somewhat intractable (but see [306]) tolC/acrAB efflux transporter. Mutations in *marR* [307] and σ^{70} (*rpoD*) [308] can have similar effects. Of course one cannot also fail to mention the variants of CRISPR/Cas9 gene editing technology that allow almost unlimited tinkering (e.g. in yeast [309, 310]).

While microbes even of the same genotype are purposely highly heterogeneous physiologically [311, 312], even in the same media, the biotechnologist might wish to turn on and off pathways in 'all' cells at once; 'quorum sensing' methods are one means to seek to do this [169, 313].

Finally, we would also recognise the potential benefits of engineering the production or uptake (e.g. [314-318]) of compatible solutes [319-322].

Directed evolution: changing the sequence of the target transporter protein(s)

Ever since the recognition of mutation itself, it has been obvious that changes in the primary sequence of a protein are more or less tightly coupled to changes in its properties [323, 324], whether they be k_{cat} [325], thermotolerance [326], substrate specificity [327], solvent tolerance [328], or any other properties.

Examples involving transporter engineering are legion [13]. A very striking one includes the cloning and engineering of an NTP transporter to allow *E. coli* to take up non-natural nucleotide triphosphates, to encode “a form of life that can stably store genetic information using a six-letter, three-base-pair alphabet” [329]. As to efflux transporters, the engineering towards influx of substances normally pumped out, for instance via uncoupled variants of the LmrP ‘efflux’ transporter in lactobacilli [330-333] is notable. Other examples have changed the specificity [334] and promiscuity [335] of transporters.

With regard to efflux transporters more generally, one sixth of the transporters of *E. coli* are efflux transporters [336]), including as many as 37 ‘multidrug resistant’ (MDR) transporters [337], most commonly from the Major Facilitator Superfamily [338, 339]. Any of these are candidates for manipulation to cause the efflux of biotechnological products. Known ones include those for dipeptides [340], antibiotics [132, 133, 135, 136], and solvents [133, 341, 342]. Broadly similar statements are true in *S. cerevisiae*, where a variety of efflux pumps help remove intracellular toxicants of all kinds [343-348]. Given the importance of horizontal gene transfer in natural evolution (e.g. [74, 349-351]), it is not surprising that even interkingdom transfer of efflux genes can be effective, e.g. the use of yeast genes in causing xenobiotic tolerance in plants [352].

Table 2 gives some examples in which efflux has been selected by exposing cells to potentially inhibitory concentrations of the substances to be effluxed.

Class of substance	Selected reference(s)
Reviews	[353]
alkanes	[189, 354-358] (see also [28])
arenes	[359-361]
short-chain alcohols	[190, 306]
short-chain fatty acids	[362, 363]
long-chain fatty acids	[110, 364]
terpenoids	[189, 365, 366]

Table 2. Some substances for which efflux pumps have been selected via tolerance to the substances in question

Consequently, when the product of interest is not close in structure to a known substrate (product) of any native transporter, it may be worth starting with a transporter for which it is; various resources allow one to seek these (e.g. [367-373]). In particular, modern methods of synthetic biology are far more powerful than are the classical methods of directed evolution, and clearly represent the future.

Synthetic biology for efflux transporter engineering

Whichever of the above general strategies are chosen (promoters, other control elements, the target protein itself), improving them always involves changing the host’s DNA sequence. In the past, and partly because of the enormous number of possible sequences [374-376] this was done rather empirically, using methods such as error-prone PCR (ePCR) [377-379]

to introduce mutations. Although showing the utility of the general directed evolution strategy, this had three highly undesirable consequences: (i) there was no control over which mutations were made, (ii) the search could only be local, as high mutation rates necessarily introduced stop codons [379, 380], and (iii) the reliance on selection of local ‘winners’ as starting points for the next generation inevitably meant that search was soon trapped in local minima from which it was impossible to escape (as was evident from many published studies showing a lack of further improvement after 3 or so generations, despite quite poor k_{cat} values) [324].

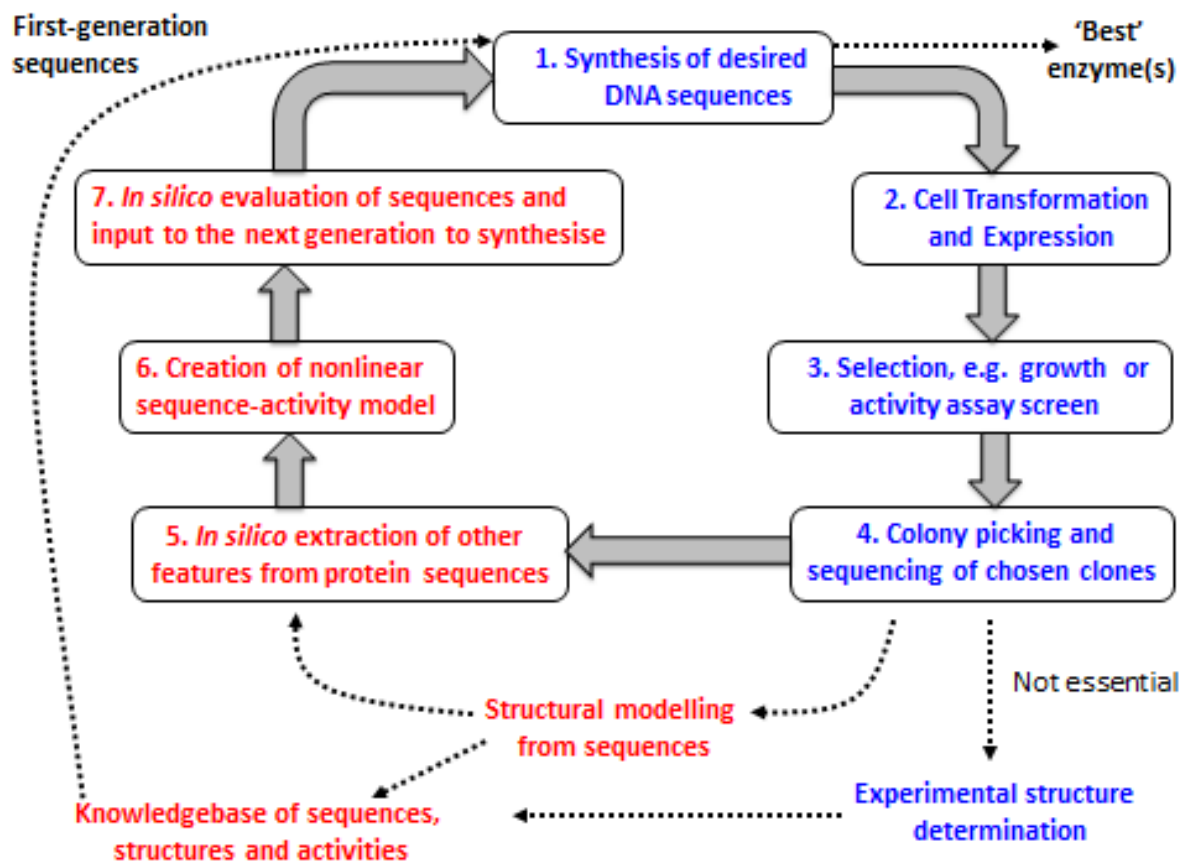


Figure 6. A generic scheme for modern synthetic biology. Parts in red are computational, in blue experimental. Based on a figure in [324] via a CC-BY licence.

The conceptual solution, well known to those studying evolutionary and related algorithms for purposes of optimisation (e.g. [324, 376, 381-385]), is that one has to combine exploitation (local search) with exploration (wider forays), and that consequently it can be helpful to know where one is in the search space (i.e. the genotype [262, 386]).

The experimental solution to this is to make the DNA in a statistically deterministic manner, which means synthetically [387], as part of a synthetic biology (synbio) pipeline (Figure 6) [324, 388]. Thus we have described methods for the controlled generation and assembly of DNA/protein sequences [288, 294, 389-391] designed to navigate these very large search spaces ‘intelligently’ [324]. While specificity is largely (but not at all completely) based on residues at or near the active site, we note in particular that raising k_{cat} requires contributions from residues that may be very distant from the active site (e.g. [324, 325, 392, 393]). The opportunities afforded by ‘deep mutational scanning’ [394-399], the coupled deep

sequencing, and the 'deep' learning [400, 401] of structure(sequence)-activity relationships [324, 402] are enormous.

A particularly nice example of the application of synbio to transporter engineering comes from Sommer and colleagues [12], who deployed and evolved riboswitches that could detect either thiamine or xanthine, thereby enabling a selection for microbes presented with metagenomes that had acquired transporters for those substrates therefrom. Presumably a related strategy based on transcriptional events that were turned off rather than on by a riboswitch [403] could equally be applied to efflux transporters. Related strategies serve to highlight other important genes in the network of interest [404].

Concluding and summarising remarks

In this comparatively short review (albeit we have tried to give many citations), the aim has been to rehearse the value of efflux transporter engineering to biotechnology. The key points are as follows:

- Essentially nothing 'floats freely' through any phospholipid bilayer that may be present in the plasma or other membrane of producer cells
- Consequently, there is typically a kinetic restriction or barrier to effluxing product by producers of high intracellular concentrations thereof, unless a suitable transporter is, or can be arranged, to be present
- The first step in a systems biology strategy is to make a model of the organism of interest, and discover which transporters might have the desired activity, and whether native activities can be increased or if it is necessary to add exogenous genes.
- Even equilibrative transporters (often labelled as 'influx' transporters) can be useful for these purposes if their kinetics are sufficiently great.
- There can be sound evolutionary (natural selection) reasons why a cell might naturally choose to efflux expensively synthesised product; the biotechnologist is wise to make use of these where they exist
- In favourable cases it may be possible to pump out the product of interest using an efflux transporter that is coupled to cellular sources of free energy.
- The methods of synthetic biology offer almost unlimited opportunities for efflux transporter engineering, and thereby for learning the sequence-structure-activity relationships of transporters and their substrates of interest

Legends to figures

Figure 1. The ability to efflux an intracellular product is of great significance in biotechnology, as the extracellular space is normally much greater than the intracellular space in a typical fermentation.

Figure 2. A 'mind map' [5] of the layout of this paper.

Figure 3. An untargeted metabolomics strategy for determining the substrates of 'orphan' transporters whose 'true' substrates are considered not to be known. Based on [184].

Figure 4. The structures of L-carnitine, ergothioneine, and some related molecules.

Figure 5. Unsupervised and supervised learning methods in cheminformatics

Figure 6. A generic scheme for modern synthetic biology. Parts in red are computational, in blue experimental. Based on a figure in [324] via a CC-BY licence.

Acknowledgments

I thank the BBSRC (grants BB/P009042/1 and BB/M017702/1) for financial support, and Irina Borodina, Zachary King and Bernhard Palsson for useful discussions..

References

1. **Fletcher E, Krivoruchko A, Nielsen J.** Industrial systems biology and its impact on synthetic biology of yeast cell factories. *Biotechnol Bioeng* 2016;113(6):1164-1170.
2. **Jezierska S, Van Bogaert INA.** Crossing boundaries: the importance of cellular membranes in industrial biotechnology. *J Ind Microbiol Biotechnol* 2017;44(4-5):721-733.
3. **Peñaloza W, Davey CL, Hedger JN, Kell DB.** Physiological Studies on the Solid-state Quinoa Tempe Fermentation, Using On-line Measurements of Fungal Biomass Production. *J Sci Food Agric* 1992;59:227-235.
4. **Krämer R.** Production of amino acids: physiological and genetic approaches. *Food Biotechnol* 2004;18(2):171-216.
5. **Buzan T.** *How to mind map*. London: Thorsons; 2002.
6. **Erni B.** Group translocation of glucose and other carbohydrates by the bacterial phosphotransferase system. *Int Rev Cytol* 1992;137:127-148.
7. **Gabor E, Gohler AK, Kosfeld A, Staab A, Kremling A et al.** The phosphoenolpyruvate-dependent glucose-phosphotransferase system from *Escherichia coli* K-12 as the center of a network regulating carbohydrate flux in the cell. *Eur J Cell Biol* 2011;90(9):711-720.
8. **Kell DB, Oliver SG.** How drugs get into cells: tested and testable predictions to help discriminate between transporter-mediated uptake and lipoidal bilayer diffusion. *Front Pharmacol* 2014;5:231.
9. **Ren Q, Paulsen IT.** Large-scale comparative genomic analyses of cytoplasmic membrane transport systems in prokaryotes. *J Mol Microbiol Biotechnol* 2007;12(3-4):165-179.
10. **César-Razquin A, Snijder B, Frappier-Brinton T, Isserlin R, Gyimesi G et al.** A call for systematic research on solute carriers. *Cell* 2015;162(3):478-487.
11. **Arai M, Okumura K, Satake M, Shimizu T.** Proteome-wide functional classification and identification of prokaryotic transmembrane proteins by transmembrane topology similarity comparison. *Protein Sci* 2004;13(8):2170-2183.
12. **Genee HJ, Bali AP, Petersen SD, Siedler S, Bonde MT et al.** Functional mining of transporters using synthetic selections. *Nat Chem Biol* 2016;12(12):1015-1022.
13. **Kell DB, Swainston N, Pir P, Oliver SG.** Membrane transporter engineering in industrial biotechnology and whole-cell biocatalysis. *Trends Biotechnol* 2015;33:237-246.
14. **Dobson PD, Kell DB.** Carrier-mediated cellular uptake of pharmaceutical drugs: an exception or the rule? *Nat Rev Drug Disc* 2008;7:205-220.
15. **Dobson P, Lanthaler K, Oliver SG, Kell DB.** Implications of the dominant role of cellular transporters in drug uptake. *Curr Top Med Chem* 2009;9:163-184.
16. **Kell DB, Dobson PD, Bilsland E, Oliver SG.** The promiscuous binding of pharmaceutical drugs and their transporter-mediated uptake into cells: what we (need to) know and how we can do so. *Drug Disc Today* 2013;18(5/6):218-239.
17. **Kell DB.** Finding novel pharmaceuticals in the systems biology era using multiple effective drug targets, phenotypic screening, and knowledge of transporters: where drug discovery went wrong and how to fix it. *FEBS J* 2013;280:5957-5980.
18. **Kell DB.** What would be the observable consequences if phospholipid bilayer diffusion of drugs into cells is negligible? *Trends Pharmacol Sci* 2015;36(1):15-21.
19. **Kell DB.** The transporter-mediated cellular uptake of pharmaceutical drugs is based on their metabolite-likeness and not on their bulk biophysical properties: Towards a systems pharmacology *Perspect Sci* 2015;6:66-83.
20. **Kell DB.** How drugs pass through biological cell membranes – a paradigm shift in our understanding? *Beilstein Magazine* 2016;2(5):http://www.beilstein-institut.de/download/628/609_kell.pdf.
21. **Jain MK.** *The bimolecular lipid membrane*. New York: Van Nostrand Reinhold; 1972.
22. **Tien HT, Ottova-Leitmannova A (editors).** *Planar lipid bilayers (BLMs) and their applications*. New York: Elsevier; 2003.

23. **Winter GE, Radic B, Mayor-Ruiz C, Blomen VA, Trefzer C et al.** The solute carrier SLC35F2 enables YM155-mediated DNA damage toxicity. *Nat Chem Biol* 2014;10:768-773.
24. **Barts PWJA, Hoeberichts JA, Klaassen A, Borst-Pauwels GWFH.** Uptake of the lipophilic cation dibenzylidimethylammonium into *Saccharomyces cerevisiae*. Interaction with the thiamine transport system. *Biochim Biophys Acta* 1980;597(1):125-136.
25. **Agre P.** Aquaporin water channels (Nobel lecture). *Angew Chem Int Ed Engl* 2004;43(33):4278-4290.
26. **Benga G.** The first discovered water channel protein, later called aquaporin 1: molecular characteristics, functions and medical implications. *Mol Aspects Med* 2012;33(5-6):518-534.
27. **Call TP, Akhtar MK, Baganz F, Grant C.** Modulating the import of medium-chain alkanes in *E. coli* through tuned expression of FadL. *J Biol Eng* 2016;10:5.
28. **Grant C, Deszcz D, Wei YC, Martinez-Torres RJ, Morris P et al.** Identification and use of an alkane transporter plug-in for applications in biocatalysis and whole-cell biosensing of alkanes. *Sci Rep* 2014;4:5844.
29. **Wang J, Fulford T, Shao Q, Javelle A, Yang H et al.** Ammonium transport proteins with changes in one of the conserved pore histidines have different performance in ammonia and methylamine conduction. *PLoS One* 2013;8(5):e62745.
30. **Kaldenhoff R, Kai L, Uehlein N.** Aquaporins and membrane diffusion of CO₂ in living organisms. *Biochim Biophys Acta* 2014;1840:1592-1595.
31. **Kai L, Kaldenhoff R.** A refined model of water and CO₂ membrane diffusion: Effects and contribution of sterols and proteins. *Sci Rep* 2014;4:6665.
32. **Penrod JT, Mace CC, Roth JR.** A pH-sensitive function and phenotype: evidence that EutH facilitates diffusion of uncharged ethanolamine in *Salmonella enterica*. *J Bacteriol*, Research Support, U.S. Gov't, P.H.S. 2004;186(20):6885-6890.
33. **Black PN, DiRusso CC.** Transmembrane movement of exogenous long-chain fatty acids: proteins, enzymes, and vectorial esterification. *Microbiol Mol Biol Rev* 2003;67(3):454-472, table of contents.
34. **van den Berg B.** The FadL family: unusual transporters for unusual substrates. *Curr Opin Struct Biol*, Review 2005;15(4):401-407.
35. **Ishibashi K, Kondo S, Hara S, Morishita Y.** The evolutionary aspects of aquaporin family. *Am J Physiol Regul Integr Comp Physiol* 2011;300(3):R566-576.
36. **Bienert GP, Chaumont F.** Aquaporin-facilitated transmembrane diffusion of hydrogen peroxide. *Biochim Biophys Acta* 2014;1840(5):1596-1604.
37. **Walker AL, Franke RM, Sparreboom A, Ware RE.** Transcellular movement of hydroxyurea is mediated by specific solute carrier transporters. *Exp Hematol* 2011;39(4):446-456.
38. **Herrera M, Garvin JL.** Aquaporins as gas channels. *Pflugers Arch* 2011;462(4):623-630.
39. **Wang Y, Tajkhorshid E.** Nitric oxide conduction by the brain aquaporin AQP4. *Proteins* 2010;78(3):661-670.
40. **Shayakul C, Cl  men  on B, Hediger MA.** The urea transporter family (SLC14): Physiological, pathological and structural aspects. *Mol Aspects Med* 2013;34(2-3):313-322.
41. **Strugatsky D, McNulty R, Munson K, Chen CK, Soltis SM et al.** Structure of the proton-gated urea channel from the gastric pathogen *Helicobacter pylori*. *Nature*, Research Support, N.I.H., Extramural

Research Support, Non-U.S. Gov't

Research Support, U.S. Gov't, Non-P.H.S. 2013;493(7431):255-258.

42. **Day RE, Kitchen P, Owen D, Bland C, Marshall L et al.** Human aquaporins: Regulators of transcellular water flow. *Biochim Biophys Acta* 2013;1840:1492-1506.
43. **Palsson B  .** *Systems biology: properties of reconstructed networks*. Cambridge: Cambridge University Press; 2006.

44. **Kell DB, Knowles JD.** The role of modeling in systems biology. In: Szallasi Z, Stelling J, Periwé V (editors). *System modeling in cellular biology: from concepts to nuts and bolts*. Cambridge: MIT Press; 2006. pp. 3-18.
45. **Kell DB.** Metabolomics, modelling and machine learning in systems biology: towards an understanding of the languages of cells. The 2005 Theodor Bücher lecture. *FEBS J* 2006;273:873-894.
46. **Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M et al.** Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Res* 2014;42(1):D199-205.
47. **Herrgård MJ, Swainston N, Dobson P, Dunn WB, Arga KY et al.** A consensus yeast metabolic network obtained from a community approach to systems biology. *Nat Biotechnol* 2008;26(10):1155-1160.
48. **Thiele I, Palsson BØ.** Reconstruction annotation jamborees: a community approach to systems biology. *Mol Syst Biol* 2010;6:361.
49. **Thiele I, Swainston N, Fleming RMT, Hoppe A, Sahoo S et al.** A community-driven global reconstruction of human metabolism. *Nat Biotechnol* 2013;31(5):419-425.
50. **Swainston N, Smallbone K, Mendes P, Kell DB, Paton NW.** The SuBliMinaL Toolbox: automating steps in the reconstruction of metabolic networks. *Integrative Bioinf* 2011;8(2):186.
51. **Carbonell P, Parutto P, Baudier C, Junot C, Faulon JL.** Retropath: Automated Pipeline for Embedded Metabolic Circuits. *Acs Synth Biol* 2014;3(8):565-577.
52. **Carbonell P, Parutto P, Herisson J, Pandit SB, Faulon JL.** XTMS: pathway design in an eXTended metabolic space. *Nucleic Acids Res* 2014;42:W389-W394.
53. **Smallbone K, Messiha HL, Carroll KM, Winder CL, Malys N et al.** A model of yeast glycolysis based on a consistent kinetic characterization of all its enzymes. *FEBS Lett* 2013;587:2832-2841.
54. **Bergmann FT, Hoops S, Klahn B, Kummer U, Mendes P et al.** COPASI and its applications in biotechnology. *J Biotechnol* 2017.
55. **Hoops S, Sahle S, Gauges R, Lee C, Pahle J et al.** COPASI: a COMplex PATHway Simulator. *Bioinformatics* 2006;22(24):3067-3074.
56. **Smallbone K, Simeonidis E, Broomhead DS, Kell DB.** Something from nothing: bridging the gap between constraint-based and kinetic modelling. *FEBS J* 2007;274:5576-5585.
57. **Liebermeister W, Klipp E.** Bringing metabolic networks to life: convenience rate law and thermodynamic constraints. *Theor Biol Med Model* 2006;3:41.
58. **Hanson AD, Pribat A, Waller JC, de Crécy-Lagard V.** 'Unknown' proteins and 'orphan' enzymes: the missing half of the engineering parts list--and how to find it. *Biochem J*, Research Support, N.I.H., Extramural

Research Support, Non-U.S. Gov't

Research Support, U.S. Gov't, Non-P.H.S.

Review 2010;425(1):1-11.

59. **Riley M, Abe T, Arnaud MB, Berlyn MK, Blattner FR et al.** *Escherichia coli* K-12: a cooperatively developed annotation snapshot--2005. *Nucleic Acids Res* 2006;34(1):1-9.
60. **Hu P, Janga SC, Babu M, Diaz-Mejia JJ, Butland G et al.** Global functional atlas of *Escherichia coli* encompassing previously uncharacterized proteins. *PLoS Biol* 2009;7(4):e96.
61. **Ito M, Baba T, Mori H, Mori H.** Functional analysis of 1440 *Escherichia coli* genes using the combination of knock-out library and phenotype microarrays. *Metab Eng* 2005;7(4):318-327.
62. **Keseler IM, Mackie A, Santos-Zavaleta A, Billington R, Bonavides-Martínez C et al.** The EcoCyc database: reflecting new knowledge about *Escherichia coli* K-12. *Nucleic Acids Res* 2017;45(D1):D543-D550.

63. **Paulsen IT, Sliwinski MK, Nelissen B, Goffeau A, Saier MH, Jr.** Unified inventory of established and putative transporters encoded within the complete genome of *Saccharomyces cerevisiae*. *FEBS Lett* 1998;430(1-2):116-125.
64. **Elbourne LDH, Tetu SG, Hassan KA, Paulsen IT.** TransportDB 2.0: a database for exploring membrane transporters in sequenced genomes from all domains of life. *Nucleic Acids Res* 2017;45(D1):D320-D324.
65. **Del Sorbo G, Schoonbeek HJ, De Waard MA.** Fungal transporters involved in efflux of natural toxic compounds and fungicides. *Fungal Genetics and Biology* 2000;30(1):1-15.
66. **Perlin MH, Andrews J, Toh SS.** Essential Letters in the Fungal Alphabet: ABC and MFS Transporters and Their Roles in Survival and Pathogenicity. *Advances in Genetics, Vol 85* 2014;85:201-253.
67. **Prasad R, Banerjee A, Khandelwal NK, Dhamgaye S.** The ABCs of *Candida albicans* Multidrug Transporter Cdr1. *Eukaryot Cell* 2015;14(12):1154-1164.
68. **Wong K, Ma J, Rothnie A, Biggin PC, Kerr ID.** Towards understanding promiscuity in multidrug efflux pumps. *Trends Biochem Sci* 2014;39(1):8-16.
69. **Lamping E, Baret PV, Holmes AR, Monk BC, Goffeau A et al.** Fungal PDR transporters: Phylogeny, topology, motifs and function. *Fungal Genetics and Biology* 2010;47(2):127-142.
70. **Diallinas G.** Dissection of Transporter Function: From Genetics to Structure. *Trends Genet* 2016;32(9):576-590.
71. **Kacser H, Burns JA.** The control of flux. In: Davies DD (editor). *Rate Control of Biological Processes Symposium of the Society for Experimental Biology Vol 27*. Cambridge: Cambridge University Press; 1973. pp. 65-104.
72. **Heinrich R, Rapoport TA.** A linear steady-state treatment of enzymatic chains. General properties, control and effector strength. *Eur J Biochem* 1974;42:89-95.
73. **Kell DB, Westerhoff HV.** Metabolic control theory: its role in microbiology and biotechnology. *FEMS Microbiol Rev* 1986;39:305-320.
74. **Kell DB, van Dam K, Westerhoff HV.** Control analysis of microbial growth and productivity. *Symp Soc Gen Microbiol* 1989;44:61-93.
75. **Heinrich R, Schuster S.** *The regulation of cellular systems*. New York: Chapman & Hall; 1996.
76. **Fell DA.** *Understanding the control of metabolism*. London: Portland Press; 1996.
77. **Walter RP, Morris JG, Kell DB.** The roles of osmotic stress and water activity in the inhibition of the growth, glycolysis and glucose phosphotransferase system of *Clostridium pasteurianum*. *J Gen Microbiol* 1987;133:259-266.
78. **Sano C.** History of glutamate production. *Am J Clin Nutr*, Historical Article Review 2009;90(3):728S-732S.
79. **Eggeling L, Sahm H.** New ubiquitous translocators: amino acid export by *Corynebacterium glutamicum* and *Escherichia coli*. *Arch Microbiol* 2003;180(3):155-160.
80. **Tryfona T, Bustard MT.** Mechanistic understanding of the fermentative L-glutamic acid overproduction by *Corynebacterium glutamicum* through combined metabolic flux profiling and transmembrane transport characteristics. *J Chem Technol Biotechnol* 2004;79(12):1321-1330.
81. **Kinoshita S, Udaka S, Shimamoto M.** Studies on amino acid fermentation, part I. Production of L-glutamic acid by various microorganisms. *J Gen Appl Microbiol* 1957;3:193–205.
82. **Hirasawa T, Kim J, Shirai T, Furusawa C, Shimizu H.** Molecular Mechanisms and Metabolic Engineering of Glutamate Overproduction in *Corynebacterium glutamicum*. *Subcellular Biochem* 2012;64:261-281.
83. **Vertès AA, Inui M, Yukawa H.** The Biotechnological Potential of *Corynebacterium glutamicum*, from Umami to Chemurgy. *Microbiol Monogr* 2013;23:1-49.
84. **Nakamura J, Hirano S, Ito H, Wachi M.** Mutations of the *Corynebacterium glutamicum* NCgl1221 gene, encoding a mechanosensitive channel homolog, induce L-glutamic acid production. *Appl Environ Microbiol* 2007;73(14):4491-4498.

85. **Hashimoto K, Nakamura K, Kuroda T, Yabe I, Nakamatsu T et al.** The protein encoded by NCgl1221 in *Corynebacterium glutamicum* functions as a mechanosensitive channel. *Biosci Biotechnol Biochem* 2010;74(12):2546-2549.
86. **Hashimoto K, Murata J, Konishi T, Yabe I, Nakamatsu T et al.** Glutamate is excreted across the cytoplasmic membrane through the NCgl1221 channel of *Corynebacterium glutamicum* by passive diffusion. *Biosci Biotechnol Biochem* 2012;76(7):1422-1424.
87. **Nakayama Y, Yoshimura K, Iida H.** A gain-of-function mutation in gating of *Corynebacterium glutamicum* NCgl1221 causes constitutive glutamate secretion. *Appl Environ Microbiol* 2012;78(15):5432-5434.
88. **Yamashita C, Hashimoto K, Kumagai K, Maeda T, Takada A et al.** L-Glutamate secretion by the N-terminal domain of the *Corynebacterium glutamicum* NCgl1221 mechanosensitive channel. *Biosci Biotechnol Biochem* 2013;77(5):1008-1013.
89. **Yao W, Deng X, Liu M, Zheng P, Sun Z et al.** Expression and localization of the *Corynebacterium glutamicum* NCgl1221 protein encoding an L-glutamic acid exporter. *Microbiol Res* 2009;164(6):680-687.
90. **Mitsuhashi S.** Current topics in the biotechnological production of essential amino acids, functional amino acids, and dipeptides. *Curr Opin Biotechnol* 2014;26:38-44.
91. **Bröer S, Krämer R.** Lysine excretion by *Corynebacterium glutamicum*. 1. Identification of a specific secretion carrier system. *Eur J Biochem*, Research Support, Non-U.S. Gov't 1991;202(1):131-135.
92. **Bröer S, Krämer R.** Lysine excretion by *Corynebacterium glutamicum*. 2. Energetics and mechanism of the transport system. *Eur J Biochem*, Research Support, Non-U.S. Gov't 1991;202(1):137-143.
93. **Kelle R, Laufer B, Brunzema C, Weuster-Botz D, Krämer R et al.** Reaction engineering analysis of L-lysine transport by *Corynebacterium glutamicum*. *Biotechnol Bioeng* 1996;51(1):40-50.
94. **Vrljić M, Sahm H, Eggeling L.** A new type of transporter with a new type of cellular function: L-lysine export from *Corynebacterium glutamicum*. *Mol Microbiol* 1996;22(5):815-826.
95. **Bellmann A, Vrljić M, Pátek M, Sahm H, Krämer R et al.** Expression control and specificity of the basic amino acid exporter LysE of *Corynebacterium glutamicum*. *Microbiology*, Research Support, Non-U.S. Gov't 2001;147(Pt 7):1765-1774.
96. **Hermann T, Krämer R.** Mechanism and Regulation of Isoleucine Excretion in *Corynebacterium glutamicum*. *Appl Environ Microbiol* 1996;62(9):3238-3244.
97. **Xie X, Xu L, Shi J, Xu Q, Chen N.** Effect of transport proteins on L-isoleucine production with the L-isoleucine-producing strain *Corynebacterium glutamicum* YILW. *J Ind Microbiol Biotechnol* 2012;39(10):1549-1556.
98. **Mundhada H, Seoane JM, Schneider K, Koza A, Christensen HB et al.** Increased production of L-serine in *Escherichia coli* through Adaptive Laboratory Evolution. *Metab Eng* 2017;39:141-150.
99. **Lee KH, Park JH, Kim TY, Kim HU, Lee SY.** Systems metabolic engineering of *Escherichia coli* for L-threonine production. *Mol Syst Biol* 2007;3:149.
100. **Diesveld R, Tietze N, Fürst O, Reth A, Bathe B et al.** Activity of exporters of *Escherichia coli* in *Corynebacterium glutamicum*, and their use to increase L-threonine production. *J Mol Microbiol Biotechnol* 2009;16(3-4):198-207.
101. **Dong X, Quinn PJ, Wang X.** Metabolic engineering of *Escherichia coli* and *Corynebacterium glutamicum* for the production of L-threonine. *Biotechnol Adv* 2011;29(1):11-23.
102. **Dong X, Quinn PJ, Wang X.** Microbial metabolic engineering for L-threonine production. *Subcell Biochem* 2012;64:283-302.
103. **Eggeling L.** Exporters for production of amino acids and other small molecules. *Adv Biochem Eng Biotechnol* 2017;159:199-225.
104. **Trötschel C, Deutenberg D, Bathe B, Burkovski A, Krämer R.** Characterization of methionine export in *Corynebacterium glutamicum*. *J Bacteriol* 2005;187(11):3786-3794.

105. **Hori H, Yoneyama H, Tobe R, Ando T, Isogai E et al.** Inducible L-alanine exporter encoded by the novel gene *ygaW* (*alaE*) in *Escherichia coli*. *Appl Environ Microbiol* 2011;77(12):4027-4034.
106. **Marin K, Krämer R.** Amino Acid Transport Systems in Biotechnologically Relevant Bacteria. *Microbiol Monogr* 2007;5:289-325.
107. **Van Dyk TK.** Bacterial efflux transport in biotechnology. *Adv Appl Microbiol* 2008;63:231-247.
108. **Kind S, Kreye S, Wittmann C.** Metabolic engineering of cellular transport for overproduction of the platform chemical 1,5-diaminopentane in *Corynebacterium glutamicum*. *Metab Eng* 2011;13(5):617-627.
109. **Anderson CM, Stahl A.** SLC27 fatty acid transport proteins. *Mol Aspects Med* 2013;34(2-3):516-528.
110. **Villalba MS, Alvarez HM.** Identification of a novel ATP-binding cassette transporter involved in long-chain fatty acid import and its role in triacylglycerol accumulation in *Rhodococcus jostii* RHA1. *Microbiology* 2014;160(Pt 7):1523-1532.
111. **Dirusso CC, Black PN.** Bacterial long chain fatty acid transport: gateway to a fatty acid-responsive signaling system. *J Biol Chem* 2004;279(48):49563-49566.
112. **van den Berg B, Black PN, Clemons WM, Jr., Rapoport TA.** Crystal structure of the long-chain fatty acid transporter FadL. *Science* 2004;304(5676):1506-1509.
113. **Færgeman NJ, DiRusso CC, Elberger A, Knudsen J, Black PN.** Disruption of the *Saccharomyces cerevisiae* homologue to the murine fatty acid transport protein impairs uptake and growth on long-chain fatty acids. *J Biol Chem* 1997;272(13):8531-8538.
114. **Zou Z, DiRusso CC, Ctrnacta V, Black PN.** Fatty acid transport in *Saccharomyces cerevisiae*. Directed mutagenesis of FAT1 distinguishes the biochemical activities associated with Fat1p. *J Biol Chem* 2002;277(34):31062-31071.
115. **Lorenz S, Guenther M, Grumaz C, Rupp S, Zibek S et al.** Genome Sequence of the Basidiomycetous Fungus *Pseudozyma aphidis* DSM70725, an Efficient Producer of Biosurfactant Mannosylerythritol Lipids. *Genome announcements* 2014;2(1).
116. **Dulermo R, Gamboa-Melendez H, Dulermo T, Thevenieau F, Nicaud JM.** The fatty acid transport protein Fat1p is involved in the export of fatty acids from lipid bodies in *Yarrowia lipolytica*. *FEMS Yeast Res* 2014;14(6):883-896.
117. **Dulermo R, Gamboa-Melendez H, Ledesma-Amaro R, Thevenieau F, Nicaud JM.** Unraveling fatty acid transport and activation mechanisms in *Yarrowia lipolytica*. *Biochim Biophys Acta* 2015;1851(9):1202-1217.
118. **Ledesma-Amaro R, Kerkhoven EJ, Revuelta JL, Nielsen J.** Genome scale metabolic modeling of the riboflavin overproducer *Ashbya gossypii*. *Biotechnol Bioeng* 2014;111(6):1191-1199.
119. **Ledesma-Amaro R, Dulermo R, Niehus X, Nicaud JM.** Combining metabolic engineering and process optimization to improve production and secretion of fatty acids. *Metab Eng* 2016;38:38-46.
120. **Schwechheimer SK, Park EY, Revuelta JL, Becker J, Wittmann C.** Biotechnology of riboflavin. *Appl Microbiol Biotechnol* 2016;100(5):2107-2119.
121. **Revuelta JL, Ledesma-Amaro R, Lozano-Martínez P, Díaz-Fernández D, Buey RM et al.** Bioproduction of riboflavin: a bright yellow history. *J Ind Microbiol Biotechnol* 2017;44(4-5):659-665.
122. **Förster C, Revuelta JL, Krämer R.** Carrier-mediated transport of riboflavin in *Ashbya gossypii*. *Appl Microbiol Biotechnol* 2001;55(1):85-89.
123. **Förster C, Santos MA, Ruffert S, Krämer R, Revuelta JL.** Physiological consequence of disruption of the VMA1 gene in the riboflavin overproducer *Ashbya gossypii*. *J Biol Chem* 1999;274(14):9442-9448.
124. **Hemberger S, Pedrolli DB, Stolz J, Vogl C, Lehmann M et al.** RibM from *Streptomyces davawensis* is a riboflavin/roseoflavin transporter and may be useful for the optimization of riboflavin production strains. *BMC Biotechnol* 2011;11:119.
125. **Udhayabanu T, Subramanian VS, Teafatiller T, Gowda VK, Raghavan VS et al.** SLC52A2 [p.P141T] and SLC52A3 [p.N21S] causing Brown-Vialetto-Van Laere Syndrome in

an Indian patient: First genetically proven case with mutations in two riboflavin transporters. *Clin Chim Acta* 2016;462:210-214.

126. **O'Hagan S, Wright Muelas M, Day PJ, Lundberg E, Kell DB.** GeneGini: assessment via the Gini coefficient of reference "housekeeping" genes and diverse human transporter expression profiles *Cell Syst* 2018;in the press.

127. **Yonezawa A, Inui K.** Novel riboflavin transporter family RFVT/SLC52: identification, nomenclature, functional characterization and genetic diseases of RFVT/SLC52. *Mol Aspects Med* 2013;34(2-3):693-701.

128. **Montanari F, Ecker GF.** Prediction of Drug-ABC Transporter Interaction - Recent Advances and Future Challenges. *Adv Drug Deliv Rev* 2015;86:17-26.

129. **Zinzi L, Contino M, Cantore M, Capparelli E, Leopoldo M et al.** ABC transporters in CSCs membranes as a novel target for treating tumor relapse. *Front Pharmacol* 2014;5:163.

130. **Bugde P, Biswas R, Merien F, Lu J, Liu DX et al.** The therapeutic potential of targeting ABC transporters to combat multi-drug resistance. *Expert Opin Ther Targets* 2017;21(5):511-530.

131. **Deparis Q, Claes A, Foulquié-Moreno MR, Thevelein JM.** Engineering tolerance to industrially relevant stress factors in yeast cell factories. *FEMS Yeast Res* 2017;17(4).

132. **Blair JM, Richmond GE, Piddock LJV.** Multidrug efflux pumps in Gram-negative bacteria and their role in antibiotic resistance. *Future Microbiol* 2014;9(10):1165-1177.

133. **Fernandes P, Ferreira BS, Cabral JM.** Solvent tolerance in bacteria: role of efflux pumps and cross-resistance with antibiotics. *Int J Antimicrob Agents* 2003;22(3):211-216.

134. **Lewinson O, Adler J, Sigal N, Bibi E.** Promiscuity in multidrug recognition and transport: the bacterial MFS Mdr transporters. *Mol Microbiol*, Research Support, Non-U.S. Gov't

Review 2006;61(2):277-284.

135. **Li XZ, Plésiat P, Nikaido H.** The challenge of efflux-mediated antibiotic resistance in Gram-negative bacteria. *Clin Microbiol Rev* 2015;28(2):337-418.

136. **Maira-Litrán T, Allison DG, Gilbert P.** An evaluation of the potential of the multiple antibiotic resistance operon (*mar*) and the multidrug efflux pump *acrAB* to moderate resistance towards ciprofloxacin in *Escherichia coli* biofilms. *J Antimicrob Chemother* 2000;45(6):789-795.

137. **Pence MA, McElvania TeKippe E, Burnham CA.** Diagnostic assays for identification of microorganisms and antimicrobial resistance determinants directly from positive blood culture broth. *Clinics in laboratory medicine* 2013;33(3):651-684.

138. **Rahman T, Yarnall B, Doyle DA.** Efflux drug transporters at the forefront of antimicrobial resistance. *Eur Biophys J* 2017;46(7):647-653.

139. **Saidijam M, Benedetti G, Ren Q, Xu Z, Hoyle CJ et al.** Microbial drug efflux proteins of the major facilitator superfamily. *Curr Drug Targets* 2006;7(7):793-811.

140. Fundamentals of enzyme kinetics, 2nd ed., (1995).

141. **Kimura M.** *The neutral theory of molecular evolution*. Cambridge: Cambridge University Press; 1983.

142. **Kell A, Glaser RW.** On the mechanical and dynamic properties of plant cell membranes: their role in growth, direct gene transfer and protoplast fusion. *J Theor Biol* 1993;160(1):41-62.

143. **Ouyang Y, Li X.** Recent research progress on soil microbial responses to drying–rewetting cycles. *Acta Ecol Sin* 2013;33:1-6.

144. **Morbach S, Krämer R.** Body shaping under water stress: osmosensing and osmoregulation of solute transport in bacteria. *ChemBioChem* 2002;3(5):384-397.

145. **Soccol CR, Vandenberghe LPS, Rodrigues C, Pandey A.** New perspectives for citric acid production and application. *Food Technol Biotechnol* 2006;44(2):141-149.

146. **García J, Torres N.** Mathematical modelling and assessment of the pH homeostasis mechanisms in *Aspergillus niger* while in citric acid producing conditions. *J Theoret Biol* 2011;282(1):23-35.

147. **Franz A, Burgstaller W, Schinner F.** Leaching with *Penicillium simplicissimum*: Influence of Metals and Buffers on Proton Extrusion and Citric Acid Production. *Appl Environ Microbiol* 1991;57(3):769-774.
148. **Jones DL.** Organic acids in the rhizosphere - a critical review. *Plant and Soil* 1998;205(1):25-44.
149. **Durrett TP, Gassmann W, Rogers EE.** The FRD3-mediated efflux of citrate into the root vasculature is necessary for efficient iron translocation. *Plant Physiol* 2007;144(1):197-205.
150. **Burgstaller W.** Thermodynamic boundary conditions suggest that a passive transport step suffices for citrate excretion in *Aspergillus* and *Penicillium*. *Microbiology* 2006;152(Pt 3):887-893.
151. **Kell DB.** Iron behaving badly: inappropriate iron chelation as a major contributor to the aetiology of vascular and other progressive inflammatory and degenerative diseases. *BMC Med Genom* 2009;2:2.
152. **De Serrano LO.** Biotechnology of siderophores in high-impact scientific fields. *Biomolecular concepts* 2017;8(3-4):169-178.
153. **Saha M, Sarkar S, Sarkar B, Sharma BK, Bhattacharjee S et al.** Microbial siderophores and their potential applications: a review. *Environmental science and pollution research international* 2016;23(5):3984-3999.
154. **Wilson BR, Bogdan AR, Miyazawa M, Hashimoto K, Tsuji Y.** Siderophores in Iron Metabolism: From Mechanism to Therapy Potential. *Trends Mol Med* 2016;22(12):1077-1090.
155. **Chen J, Zhu X, Tan Z, Xu H, Tang J et al.** Activating C₄-dicarboxylate transporters DcuB and DcuC for improving succinate production. *Appl Microbiol Biotechnol* 2014;98(5):2197-2205.
156. **Li XZ, Zhang WM, Wu MK, Xin FX, Dong WL et al.** Performance and mechanism analysis of succinate production under different transporters in *Escherichia coli*. *Biotechnol Bioproc Eng* 2017;22(5):529-538.
157. **Lanthaler K, Bilsland E, Dobson P, Moss HJ, Pir P et al.** Genome-wide assessment of the carriers involved in the cellular uptake of drugs: a model system in yeast. *BMC Biol* 2011;9:70.
158. **Schroeder WA, Johnson EA.** Carotenoids protect *Phaffia rhodozyma* against singlet oxygen damage. *J Ind Microbiol* 1995;14(6):502-507.
159. **Davey HM, Kell DB.** Flow cytometry and cell sorting of heterogeneous microbial populations: the importance of single-cell analysis. *Microbiol Rev* 1996;60:641-696.
160. **Stephens K.** Pheromones among the prokaryotes. *CRC Crit Rev Microbiol* 1986;13:309-334.
161. **Bu'lock JD.** Intermediary metabolism and antibiotic synthesis. *Adv Microbial Physiol* 1961;3:293-333.
162. **Kell DB, Kaprelyants AS, Grafen A.** On pheromones, social behaviour and the functions of secondary metabolism in bacteria. *Trends Ecol Evolution* 1995;10:126-129.
163. **Goh EB, Yim G, Tsui W, McClure J, Surette MG et al.** Transcriptional modulation of bacterial gene expression by subinhibitory concentrations of antibiotics. *Proc Natl Acad Sci U S A* 2002;99(26):17025-17030.
164. **Raaijmakers JM, Mazzola M.** Diversity and natural functions of antibiotics produced by beneficial and plant pathogenic bacteria. *Annu Rev Phytopathol* 2012;50:403-424.
165. **Yim G, Wang HH, Davies J.** The truth about antibiotics. *Int J Med Microbiol* 2006;296(2-3):163-170.
166. **Peiris D, Dunn WB, Brown M, Kell DB, Roy I et al.** Metabolite profiles of interacting mycelial fronts differ for pairings of the wood decay basidiomycete fungus, *Stereum hirsutum* with its competitors *Coprinus micaceus* and *Coprinus disseminatus*. *Metabolomics* 2008;4(1):52-62.
167. **Fuqua WC, Winans SC, Greenberg EP.** Quorum sensing in bacteria - the *luxR-luxI* family of cell density-responsive transcriptional regulators. *J Bacteriol* 1994;176:269-275.

168. **Dandekar AA, Chugani S, Greenberg EP.** Bacterial quorum sensing and metabolic incentives to cooperate. *Science*, Research Support, N.I.H., Extramural

Research Support, Non-U.S. Gov't 2012;338(6104):264-266.

169. **Whiteley M, Diggle SP, Greenberg EP.** Progress in and promise of bacterial quorum sensing research. *Nature* 2017;551(7680):313-320.

170. **Choudhary S, Schmidt-Dannert C.** Applications of quorum sensing in biotechnology. *Appl Microbiol Biotechnol* 2010;86(5):1267-1279.

171. **Mangwani N, Dash HR, Chauhan A, Das S.** Bacterial quorum sensing: functional features and potential applications in biotechnology. *J Mol Microbiol Biotechnol* 2012;22(4):215-227.

172. **Crittenden PD, Porter N.** Lichen-forming fungi - potential sources of novel metabolites. *Trends Biotechnol* 1991;9:409-414.

173. **Lutzoni F, Pagel M, Reeb V.** Major fungal lineages are derived from lichen symbiotic ancestors. *Nature* 2001;411(6840):937-940.

174. **Senior E, Bull AT, Slater JH.** Enzyme evolution in a microbial community growing on the herbicide Dalapon. *Nature* 1976;263(5577):476-479.

175. **Xiao Y, Angulo MT, Friedman J, Waldor MK, Weiss ST et al.** Mapping the ecological networks of microbial communities. *Nat Commun* 2017;8(1):2042.

176. **McInerney MJ, Sieber JR, Gunsalus RP.** Syntrophy in anaerobic global carbon cycles. *Curr Opin Biotechnol* 2009;20(6):623-632.

177. **Sieber JR, McInerney MJ, Gunsalus RP.** Genomic insights into syntrophy: the paradigm for anaerobic metabolic cooperation. *Annu Rev Microbiol* 2012;66:429-452.

178. **Rotaru AE, Woodard TL, Nevin KP, Lovley DR.** Link between capacity for current production and syntrophic growth in *Geobacter* species. *Front Microbiol* 2015;6:744.

179. **Westerhoff HV, Hellingwerf KJ, van Dam K.** Thermodynamic efficiency of microbial growth is low but optimal for maximal growth rate. *Proc Natl Acad Sci USA* 1983;80(1):305-309.

180. **Markx GH, Davey CL, Kell DB.** The permittistat: a novel type of turbidostat. *J Gen Microbiol* 1991;137:735-743.

181. **Davey HM, Davey CL, Woodward AM, Edmonds AN, Lee AW et al.** Oscillatory, stochastic and chaotic growth rate fluctuations in permittistatically-controlled yeast cultures. *Biosystems* 1996;39:43-61.

182. **O'Hagan S, Kell DB.** The apparent permeabilities of Caco-2 cells to marketed drugs: magnitude, and independence from both biophysical properties and endogenite similarities *PeerJ* 2015;3:e1405.

183. **Gründemann D, Liebich G, Kiefer N, Köster S, Schömig E.** Selective substrates for non-neuronal monoamine transporters. *Mol Pharmacol* 1999;56(1):1-10.

184. **Gründemann D, Harlfinger S, Golz S, Geerts A, Lazar A et al.** Discovery of the ergothioneine transporter. *Proc Natl Acad Sci* 2005;102(14):5256-5261.

185. **Fersht A.** *Structure and mechanism in protein science : a guide to enzyme catalysis and protein folding.* San Francisco: W.H. Freeman; 1999.

186. **Mendes P, Oliver SG, Kell DB.** Fitting transporter activities to cellular drug concentrations and fluxes: why the bumblebee can fly. *Trends Pharmacol Sci* 2015;36:710-723.

187. **Höglund PJ, Nordström KJV, Schiöth HB, Fredriksson R.** The solute carrier families have a remarkably long evolutionary history with the majority of the human families present before divergence of Bilaterian species. *Mol Biol Evol*, Research Support, Non-U.S. Gov't 2011;28(4):1531-1541.

188. **Hassan KA, Fagerlund A, Elbourne LDH, Vörös A, Kroeger JK et al.** The putative drug efflux systems of the *Bacillus cereus* group. *PLoS One* 2017;12(5):e0176188.

189. **Foo JL, Leong SSJ.** Directed evolution of an *E. coli* inner membrane transporter for improved efflux of biofuel molecules. *Biotechnol Biofuels* 2013;6:81.

190. **Foo JL, Jensen HM, Dahl RH, George K, Keasling JD et al.** Improving microbial biogasoline production in *Escherichia coli* using tolerance engineering. *MBio* 2014;5(6):01932-01914.
191. **Ramos JL, Sol Cuenca M, Molina-Santiago C, Segura A, Duque E et al.** Mechanisms of solvent resistance mediated by interplay of cellular factors in *Pseudomonas putida*. *FEMS Microbiol Rev* 2015;39(4):555-566.
192. **Mohamed ET, Wang S, Lennen RM, Herrgard MJ, Simmons BA et al.** Generation of a platform strain for ionic liquid tolerance using adaptive laboratory evolution. *Microb Cell Fact* 2017;16(1):204.
193. **Giaever G, Chu AM, Ni L, Connelly C, Riles L et al.** Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* 2002;418(6896):387-391.
194. **Featherstone DE, Broadie K.** Wrestling with pleiotropy: genomic and topological analysis of the yeast gene expression network. *Bioessays* 2002;24(3):267-274.
195. **Eraly SA.** Striking differences between knockout and wild-type mice in global gene expression variability. *PLoS One* 2014;9(5):e97734.
196. **Bürkstümmer T, Banning C, Hainzl P, Schobesberger R, Kerzendorfer C et al.** A reversible gene trap collection empowers haploid genetics in human cells. *Nat Methods* 2013;10(10):965-971.
197. **Skjoedt ML, Snoek T, Kildegaard KR, Arsovska D, Eichenberger M et al.** Engineering prokaryotic transcriptional activators as metabolite biosensors in yeast. *Nat Chem Biol* 2016;12(11):951-958.
198. **Gründemann D.** The ergothioneine transporter controls and indicates ergothioneine activity--a review. *Prev Med* 2012;54 Suppl:S71-74.
199. **Kerley RN, McCarthy C, Kell DB, Kenny LC.** The potential therapeutic effects of ergothioneine in pre-eclampsia. *Free Radic Biol Med* 2018;117:145-157.
200. **Koepsell H.** The SLC22 family with transporters of organic cations, anions and zwitterions. *Mol Aspects Med* 2013;34(2-3):413-435.
201. **Dunn WB, Broadhurst D, Begley P, Zelena E, Francis-McIntyre S et al.** Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. *Nat Protoc* 2011;6(7):1060-1083.
202. **Cho K, Mahieu NG, Johnson SL, Patti GJ.** After the feature presentation: technologies bridging untargeted metabolomics and biology. *Curr Opin Biotechnol* 2014;28:143-148.
203. **Dunn WB, Erban A, Weber RJM, Creek DJ, Brown M et al.** Mass Appeal: metabolite identification in mass spectrometry-focused untargeted metabolomics. *Metabolomics* 2013;9:S44-S66.
204. **Martin JC, Maillot M, Mazerolles G, Verdu A, Lyan B et al.** Can we trust untargeted metabolomics? Results of the metabo-ring initiative, a large-scale, multi-instrument inter-laboratory study. *Metabolomics* 2015;11(4):807-821.
205. **Di Guida R, Engel J, Allwood JW, Weber RJ, Jones MR et al.** Non-targeted UHPLC-MS metabolomic data processing methods: a comparative investigation of normalisation, missing value imputation, transformation and scaling. *Metabolomics* 2016;12:93.
206. **Lloyd AJ, Beckmann M, Favé G, Mathers JC, Draper J.** Proline betaine and its biotransformation products in fasting urine samples are potential biomarkers of habitual citrus fruit consumption. *Br J Nutr* 2011;106(6):812-824.
207. **Heinzmann SS, Brown IJ, Chan Q, Bictash M, Dumas ME et al.** Metabolic profiling strategy for discovery of nutritional biomarkers: proline betaine as a marker of citrus consumption. *Am J Clin Nutr* 2010;92(2):436-443.
208. **Lang R, Lang T, Bader M, Beusch A, Schlagbauer V et al.** High-Throughput Quantitation of Proline Betaine in Foods and Suitability as a Valid Biomarker for Citrus Consumption. *J Agric Food Chem* 2017;65(8):1613-1619.
209. **O'Hagan S, Swainston N, Handl J, Kell DB.** A 'rule of 0.5' for the metabolite-likeness of approved pharmaceutical drugs. *Metabolomics* 2015;11(2):323-339.

210. **Cheah IK, Halliwell B.** Ergothioneine; antioxidant potential, physiological function and role in disease. *Biochim Biophys Acta* 2012;1822(5):784-793.
211. **Cheah IK, Tang RMY, Yew TSZ, Lim KHC, Halliwell B.** Administration of Pure Ergothioneine to Healthy Human Subjects: Uptake, Metabolism, and Effects on Biomarkers of Oxidative Damage and Inflammation. *Antioxid Redox Signal* 2017;26(5):193-206.
212. **Halliwell B, Cheah IK, Drum CL.** Ergothioneine, an adaptive antioxidant for the protection of injured tissues? A hypothesis. *Biochem Biophys Res Commun* 2016;470(2):245-250.
213. **O'Hagan S, Kell DB.** Consensus rank orderings of molecular fingerprints illustrate the 'most genuine' similarities between marketed drugs and small endogenous human metabolites, but highlight exogenous natural products as the most important 'natural' drug transporter substrates. *ADMET & DMPK* 2017;5(2):85-125.
214. **Daßler T, Maier T, Winterhalter C, Böck A.** Identification of a major facilitator protein from *Escherichia coli* involved in efflux of metabolites of the cysteine pathway. *Mol Microbiol* 2000;36(5):1101-1112.
215. **Fukui K, Koseki C, Yamamoto Y, Nakamura J, Sasahara A et al.** Identification of succinate exporter in *Corynebacterium glutamicum* and its physiological roles under anaerobic conditions. *J Biotechnol* 2011;154(1):25-34.
216. **Holle SK, Ciarimboli G, Edemir B, Neugebauer U, Pavenstädt H et al.** Properties and regulation of organic cation transport in freshly isolated mouse proximal tubules analyzed with a fluorescence reader-based method. *Pflugers Arch* 2011;462(2):359-369.
217. **Kido Y, Matsson P, Giacomini KM.** Profiling of a prescription drug library for potential renal drug-drug interactions mediated by the organic cation transporter 2. *J Med Chem* 2011;54(13):4548-4558.
218. **Chang HC, Yang SF, Huang CC, Lin TS, Liang PH et al.** Development of a novel non-radioactive cell-based method for the screening of SGLT1 and SGLT2 inhibitors using 1-NBDG. *Mol Biosyst* 2013;9(8):2010-2020.
219. **Tegos GP, Evangelisti AM, Strouse JJ, Ursu O, Bologna C et al.** A high throughput flow cytometric assay platform targeting transporter inhibition. *Drug Disc Today Technol* 2014;12:e95-e103.
220. **Ugwu MC, Pelis R, Esimone CO, Agu RU.** Fluorescent organic cations for human OCT2 transporters screening: uptake in CHO cells stably expressing hOCT2. *ADMET & DMPK* 2017;5(2):135-145.
221. **Liang Y, Li S, Chen L.** The physiological role of drug transporters. *Protein Cell* 2015;6(5):334-350.
222. **Binder S, Schendzielorz G, Stäbler N, Krumbach K, Hoffmann K et al.** A high-throughput approach to identify genomic variants of bacterial metabolite producers at the single-cell level. *Genome Biol* 2012;13(5):R40.
223. **D'Ambrosio V, Jensen MK.** Lighting up yeast cell factories by transcription factor-based biosensors. *FEMS Yeast Res* 2017;17(7).
224. **Liu D, Evans T, Zhang F.** Applications and advances of metabolite biosensors for metabolic engineering. *Metab Eng* 2015;31:35-43.
225. **Leavitt JM, Wagner JM, Tu CC, Tong A, Liu Y et al.** Biosensor-Enabled Directed Evolution to Improve Muconic Acid Production in *Saccharomyces cerevisiae*. *Biotechnol J* 2017;12(10).
226. **Libis V, Delépine B, Faulon JL.** Sensing new chemicals with bacterial transcription factors. *1369-5274* 2016;33:105-112.
227. **Mahr R, Gätgens C, Gätgens J, Polen T, Kalinowski J et al.** Biosensor-driven adaptive laboratory evolution of L-valine production in *Corynebacterium glutamicum*. *Metab Eng* 2015;32:184-194.
228. **Mannan AA, Liu D, Zhang F, Oyarzún DA.** Fundamental Design Principles for Transcription-Factor-Based Metabolite Biosensors. *Acs Synth Biol* 2017;6(10):1851-1859.
229. **Mustafi N, Grunberger A, Kohlheyer D, Bott M, Frunzke J.** The development and application of a single-cell biosensor for the detection of L-methionine and branched-chain amino acids. *Metab Eng* 2012;14(4):449-457.

230. **Qian S, Cirino PC.** Using metabolite-responsive gene regulators to improve microbial biosynthesis. *Current Opinion in Chemical Engineering* 2016;14:93-102.
231. **Rogers JK, Taylor ND, Church GM.** Biosensor-based engineering of biosynthetic pathways. *Curr Opin Biotechnol* 2016;42:84-91.
232. **Schallmeyer M, Frunzke J, Eggeling L, Marienhagen J.** Looking for the pick of the bunch: high-throughput screening of producing microorganisms with biosensors. *Curr Opin Biotechnol* 2014;26:148-154.
233. **Siedler S, Stahlhut SG, Malla S, Maury J, Neves AR.** Novel biosensors based on flavonoid-responsive transcriptional regulators introduced into *Escherichia coli*. *Metab Eng, Research Support, Non-U.S. Gov't* 2014;21:2-8.
234. **Siedler S, Khatri NK, Zsohár A, Kjærboelling I, Vogt M et al.** Development of a Bacterial Biosensor for Rapid Screening of Yeast p-Coumaric Acid Production. *Acs Synth Biol* 2017;6(10):1860-1869.
235. **Younger AKD, Dalvie NC, Rottinghaus AG, Leonard JN.** Engineering Modular Biosensors to Confer Metabolite-Responsive Regulation of Transcription. *Acs Synth Biol* 2017;6(2):311-325.
236. **Zhang J, Jensen MK, Keasling JD.** Development of biosensors and their application in metabolic engineering. *Curr Opin Chem Biol* 2015;28:1-8.
237. **Bender A, Glen RC.** Molecular similarity: a key technique in molecular informatics. *Org Biomol Chem* 2004;2(22):3204-3218.
238. **Gasteiger J (editor)** *Handbook of Chemoinformatics: From Data to Knowledge*. Weinheim: Wiley/VCH; 2003.
239. **Gasteiger J.** *Basic Chemoinformatics: A Textbook*. Weinheim: Wiley/VCH; 2003.
240. **Weininger D.** SMILES, a chemical language and information system .1. Introduction to methodology and encoding rules. *J Chem Inf Comput Sci* 1988;28(1):31-36.
241. **Bachrach SM.** InChI: a user's perspective. *J Cheminform*, Editorial 2012;4(1):34.
242. **Heller S, McNaught A, Stein S, Tchekhovskoi D, Pletnev I.** InChI - the worldwide chemical structure identifier standard. *J Cheminform* 2013;5(1):7.
243. **Williams AJ.** InChI: connecting and navigating chemistry. *J Cheminform* 2012;4(1):33.
244. **Riniker S, Landrum GA.** Open-source platform to benchmark fingerprints for ligand-based virtual screening. *J Cheminform* 2013;5(1):26.
245. **Beisken S, Meini T, Wiswedel B, de Figueiredo LF, Berthold M et al.** KNIME-CDK: Workflow-driven cheminformatics. *BMC bioinformatics* 2013;14:257.
246. **Berthold MR, Cebron N, Dill F, Gabriel TR, Kötter T et al.** KNIME: the Konstanz Information Miner. In: Preisach C, Burkhardt H, Schmidt-Thieme L, Decker R (editors). *Data Analysis, Machine Learning and Applications*. Berlin: Springer; 2008. pp. 319-326.
247. **Mazanetz MP, Marmon RJ, Reisser CBT, Morao I.** Drug discovery applications for KNIME: an open source data mining platform. *Curr Top Med Chem, Review* 2012;12(18):1965-1979.
248. **O'Hagan S, Kell DB.** The KNIME workflow environment and its applications in Genetic Programming and machine learning. *Genetic Progr Evol Mach* 2015;16:387-391.
249. **Arif SM, Holliday JD, Willett P.** Comparison of chemical similarity measures using different numbers of query structures. *J Inf Sci* 2013;39(1):7-14.
250. **Willett P.** The calculation of molecular structural similarity: principles and practice. *Mol Inform* 2014;33(6-7):403-413.
251. **O'Hagan S, Kell DB.** MetMaxStruct: a Tversky-similarity-based strategy for analysing the (sub)structural similarities of drugs and endogenous metabolites. *Front Pharmacol* 2016;7:266.
252. **Todeschini R, Consonni V.** *Molecular descriptors for cheminformatics*. Weinheim: WILEY-VCH Verlag GmbH; 2009.
253. **O'Hagan S, Kell DB.** Analysing and navigating natural products space for generating small, diverse, but representative chemical libraries. *Biotechnol J* 2018;13(1):doi 10.1002/biot.201700503.

254. **Handl J, Knowles J, Kell DB.** Computational cluster validation in post-genomic data analysis. *Bioinformatics* 2005;21:3201-3212.
255. **Xu R, Wunsch D, 2nd.** Survey of clustering algorithms. *IEEE Trans Neural Netw, Comparative Study*

Research Support, Non-U.S. Gov't

Research Support, U.S. Gov't, Non-P.H.S.

Review 2005;16(3):645-678.

256. **Hastie T, Tibshirani R, Friedman J.** *The elements of statistical learning: data mining, inference and prediction, 2nd edition.* Berlin: Springer-Verlag; 2009.
257. **Wold S, Sjöström M, Eriksson L.** PLS-regression: a basic tool of chemometrics. *Chemometr Intell Lab* 2001;58(2):109-130.
258. **Breiman L.** Random forests. *Machine Learning* 2001;45(1):5-32.
259. **Koza JR.** *Genetic programming: on the programming of computers by means of natural selection.* Cambridge, Mass: MIT Press; 1992.
260. **Langdon WB.** *Genetic programming and data structures: genetic programming + data structures = automatic programming!* Boston: Kluwer; 1998.
261. **Langdon WB, Poli R.** *Foundations of genetic programming.* Berlin: Springer-Verlag; 2002.
262. **Kell DB.** Genotype:phenotype mapping: genes as computer programs. *Trends Genet* 2002;18(11):555-559.
263. **Golbraikh A, Tropsha A.** Beware of q^2 ! *J Mol Graph Model* 2002;20(4):269-276.
264. **Golbraikh A, Muratov E, Fourches D, Tropsha A.** Data Set Modelability by QSAR. *J Chem Inf Model* 2014;54(1):1-4.
265. **Fujita T, Winkler DA.** Understanding the Roles of the "Two QSARs". *J Chem Inf Model* 2016;56(2):269-274.
266. **Bailey PD, Boyd CA, Collier ID, George JP, Kellett GL et al.** Affinity prediction for substrates of the peptide transporter PepT1. *Chem Commun (Camb)* 2006(3):323-325.
267. **Foley DW, Rajamanickam J, Bailey PD, Meredith D.** Bioavailability through PepT1: the role of computer modelling in intelligent drug design. *Curr Comput Aided Drug Des* 2010;6(1):68-78.
268. **Smith DE, Cléménçon B, Hediger MA.** Proton-coupled oligopeptide transporter family SLC15: Physiological, pharmacological and pathological implications. *Mol Aspects Med* 2013;34(2-3):323-336.
269. **Bretschneider B, Brandsch M, Neubert R.** Intestinal transport of beta-lactam antibiotics: analysis of the affinity at the H⁺/peptide symporter (PEPT1), the uptake into Caco-2 cell monolayers and the transepithelial flux. *Pharm Res* 1999;16(1):55-61.
270. **Mitsuoka K, Kato Y, Kubo Y, Tsuji A.** Functional expression of stereoselective metabolism of cephalixin by exogenous transfection of oligopeptide transporter PEPT1. *Drug Metab Dispos* 2007;35(3):356-362.
271. **Hironaka T, Itokawa S, Ogawara K, Higaki K, Kimura T.** Quantitative evaluation of PEPT1 contribution to oral absorption of cephalixin in rats. *Pharm Res* 2009;26(1):40-50.
272. **Giacomini MM, Hao J, Liang X, Chandrasekhar J, Twelves J et al.** Interaction of 2,4-Diaminopyrimidine-Containing Drugs Including Fedratinib and Trimethoprim with Thiamine Transporters. *Drug Metab Dispos* 2017;45(1):76-85.
273. **Lepist EI, Ray AS.** Beyond drug-drug interactions: effects of transporter inhibition on endobiotics, nutrients and toxins. *Expert Opin Drug Metab Toxicol* 2017;13(10):1075-1087.
274. **Zhao R, Goldman ID.** Folate and thiamine transporters mediated by facilitative carriers (SLC19A1-3 and SLC46A1) and folate receptors. *Mol Aspects Med* 2013;34(2-3):373-385.
275. **Liang X, Chien HC, Yee SW, Giacomini MM, Chen EC et al.** Metformin Is a Substrate and Inhibitor of the Human Thiamine Transporter, THTR-2 (SLC19A3). *Mol Pharm* 2015;12(12):4301-4310.

276. **Loew R, Heinz N, Hampf M, Bujard H, Gossen M.** Improved Tet-responsive promoters with minimized background expression. *BMC Biotechnol* 2010;10:81.
277. **Urlinger S, Baron U, Thellmann M, Hasan MT, Bujard H et al.** Exploring the sequence space for tetracycline-dependent transcriptional activators: Novel mutations yield expanded range and sensitivity. *Proc Natl Acad Sci U S A* 2000;97(14):7963-7968.
278. **Jensen PR, Hammer K.** Artificial promoters for metabolic optimization. *Biotechnol Bioeng* 1998;58(2-3):191-195.
279. **Hammer K, Mijakovic I, Jensen PR.** Synthetic promoter libraries--tuning of gene expression. *Trends Biotechnol* 2006;24(2):53-55.
280. **Braatsch S, Helmark S, Kranz H, Koebmann B, Jensen PR.** *Escherichia coli* strains with promoter libraries constructed by Red/ET recombination pave the way for transcriptional fine-tuning. *Biotechniques* 2008;45(3):335-337.
281. **Wagner S, Klepsch MM, Schlegel S, Appel A, Draheim R et al.** Tuning *Escherichia coli* for membrane protein overexpression. *Proc Natl Acad Sci U S A* 2008;105(38):14371-14376.
282. **Rohlfhill J, Sandoval NR, Papoutsakis ET.** Sort-Seq approach to engineering a formaldehyde-inducible promoter for dynamically regulated *Escherichia coli* growth on methanol. *Acs Synth Biol* 2017.
283. **Rytter JV, Helmark S, Chen J, Lezyk MJ, Solem C et al.** Synthetic promoter libraries for *Corynebacterium glutamicum*. *Appl Microbiol Biotechnol* 2014;98(6):2617-2623.
284. **Shen J, Chen J, Jensen PR, Solem C.** A novel genetic tool for metabolic optimization of *Corynebacterium glutamicum*: efficient and repetitive chromosomal integration of synthetic promoter-driven expression libraries. *Appl Microbiol Biotechnol* 2017;101(11):4737-4746.
285. **Hockney RC.** Recent developments in heterologous protein production in *Escherichia coli*. *Trends Biotechnol* 1994;12(11):456-463.
286. **Brawand D, Soumillon M, Necsulea A, Julien P, Csárdi G et al.** The evolution of gene expression levels in mammalian organs. *Nature* 2011;478(7369):343-348.
287. **GTEC Consortium, Battle A, Brown CD, Engelhardt BE, Montgomery SB.** Genetic effects on gene expression across human tissues. *Nature* 2017;550(7675):204-213.
288. **Swainston N, Dunstan M, Jervis AJ, Robinson CJ, Carbonell P et al.** PartsGenie: an integrated tool for optimising and sharing synthetic biology parts. 2018:submitted.
289. **Kudla G, Murray AW, Tollervey D, Plotkin JB.** Coding-sequence determinants of gene expression in *Escherichia coli*. *Science* 2009;324(5924):255-258.
290. **Jones DL, Brewster RC, Phillips R.** Promoter architecture dictates cell-to-cell variability in gene expression. *Science* 2014;346(6216):1533-1536.
291. **Angov E.** Codon usage: nature's roadmap to expression and folding of proteins. *Biotechnol J* 2011;6(6):650-659.
292. **Angov E, Legler PM, Mease RM.** Adjustment of codon usage frequencies by codon harmonization improves protein expression and folding. *Methods Mol Biol* 2011;705:1-13.
293. **Plotkin JB, Kudla G.** Synonymous but not the same: the causes and consequences of codon bias. *Nat Rev Genet* 2011;12(1):32-42.
294. **Swainston N, Currin A, Day PJ, Kell DB.** GeneGenie: optimised oligomer design for directed evolution. *Nucleic Acids Res* 2014;12:W395-W400.
295. **Winkler WC, Nahvi A, Roth A, Collins JA, Breaker RR.** Control of gene expression by a natural metabolite-responsive ribozyme. *Nature* 2004;428(6980):281-286.
296. **Salis HM, Mirsky EA, Voigt CA.** Automated design of synthetic ribosome binding sites to control protein expression. *Nat Biotechnol* 2009;27(10):946-950.
297. **Salis HM.** The ribosome binding site calculator. *Methods Enzymol* 2011;498:19-42.
298. **Alper H, Stephanopoulos G.** Global transcription machinery engineering: a new approach for improving cellular phenotype. *Metab Eng* 2007;9(3):258-267.
299. **Han T, Chen Q, Liu H.** Engineered Photoactivatable Genetic Switches Based on the Bacteriophage T7 RNA Polymerase. *Acs Synth Biol* 2017;6(2):357-366.

300. **Butelli E, Titta L, Giorgio M, Mock HP, Matros A et al.** Enrichment of tomato fruit with health-promoting anthocyanins by expression of select transcription factors. *Nat Biotechnol* 2008;26(11):1301-1308.
301. **Zhang Y, Butelli E, Alseekh S, Tohge T, Rallapalli G et al.** Multi-level engineering facilitates the production of phenylpropanoid compounds in tomato. *Nat Commun* 2015;6:8635.
302. **Lange C, Mustafi N, Frunzke J, Kennerknecht N, Wessel M et al.** Lrp of *Corynebacterium glutamicum* controls expression of the *bmFE* operon encoding the export system for L-methionine and branched-chain amino acids. *J Biotechnol* 2012;158(4):231-241.
303. **Fang X, Sastry A, Mih N, Kim D, Tan J et al.** Global transcriptional regulatory network for *Escherichia coli* robustly connects gene expression to transcription factor activities. *Proc Natl Acad Sci U S A* 2017;114(38):10286-10291.
304. **Shah AA, Wang C, Chung YR, Kim JY, Choi ES et al.** Enhancement of geraniol resistance of *Escherichia coli* by MarA overexpression. *J Biosci Bioeng* 2013;115(3):253-258.
305. **Du D, Wang Z, James NR, Voss JE, Klimont E et al.** Structure of the AcrAB-TolC multidrug efflux pump. *Nature* 2014;509(7501):512-515.
306. **Fisher MA, Boyarskiy S, Yamada MR, Kong N, Bauer S et al.** Enhancing tolerance to short-chain alcohols by engineering the *Escherichia coli* AcrB efflux pump to secrete the non-native substrate n-butanol. *Acs Synth Biol* 2014;3(1):30-40.
307. **Pourahmad Jaktaji R, Ebadi R, Karimi M.** Study of Organic Solvent Tolerance and Increased Antibiotic Resistance Properties in *E. coli* *gyrA* Mutants. *Iran J Pharm Res* 2012;11(2):595-600.
308. **Zhang F, Qian X, Si H, Xu G, Han R et al.** Significantly improved solvent tolerance of *Escherichia coli* by global transcription machinery engineering. *Microb Cell Fact* 2015;14:175.
309. **Ronda C, Maury J, Jakociunas T, Jacobsen SA, Germann SM et al.** CrEdit: CRISPR mediated multi-loci gene integration in *Saccharomyces cerevisiae*. *Microb Cell Fact* 2015;14:97.
310. **Stovicek V, Holkenbrink C, Borodina I.** CRISPR/Cas system for yeast genome engineering: advances and applications. *FEMS Yeast Res* 2017.
311. **Kell DB, Ryder HM, Kaprelyants AS, Westerhoff HV.** Quantifying heterogeneity: Flow cytometry of bacterial cultures. *Antonie van Leeuwenhoek* 1991;60:145-158.
312. **Kell DB, Potgieter M, Pretorius E.** Individuality, phenotypic differentiation, dormancy and 'persistence' in culturable bacterial systems: commonalities shared by environmental, laboratory, and clinical microbiology. *F1000Res* 2015;4:179.
313. **Williams TC, Aversch NJH, Winter G, Plan MR, Vickers CE et al.** Quorum-sensing linked RNA interference for dynamic metabolic pathway control in *Saccharomyces cerevisiae*. *Metab Eng* 2015;29:124-134.
314. **Csonka L.** Physiological and genetic responses of bacteria to osmotic stress. *Microbiol Rev* 1989;53(1):121-147.
315. **Farwick M, Siewe RM, Krämer R.** Glycine betaine uptake after hyperosmotic shift in *Corynebacterium glutamicum*. *J Bacteriol* 1995;177(16):4690-4695.
316. **Kempf B, Bremer E.** Uptake and synthesis of compatible solutes as microbial stress responses to high-osmolality environments. *Arch Microbiol* 1998;170(5):319-330.
317. **Weinand M, Krämer R, Murbach S.** Characterization of compatible solute transporter multiplicity in *Corynebacterium glutamicum*. *Appl Microbiol Biotechnol*, Research Support, Non-U.S. Gov't 2007;76(3):701-708.
318. **Ochrombel I, Becker M, Krämer R, Marin K.** Osmotic stress response in *C. glutamicum*: impact of channel- and transporter-mediated potassium accumulation. *Arch Microbiol* 2011;193(11):787-796.
319. **Brown AD.** Compatible solutes and extreme water stress in eukaryotic micro-organisms. *Adv Microb Physiol* 1978;17:181-242.

320. **Hohmann S, Krantz M, Nordlander B.** Yeast osmoregulation. *Methods in enzymology*, Research Support, Non-U.S. Gov't Review 2007;428:29-45.
321. **Fahnert B.** Folding-promoting agents in recombinant protein production. *Methods Mol Biol* 2004;267:53-74.
322. **Fahnert B.** Using folding promoting agents in recombinant protein production: a review. *Methods Mol Biol* 2012;824:3-36.
323. **Voigt CA, Kauffman S, Wang ZG.** Rational evolutionary design: The theory of *in vitro* protein evolution. *Adv Prot Chem* 2001;55:79-160.
324. **Currin A, Swainston N, Day PJ, Kell DB.** Synthetic biology for the directed evolution of protein biocatalysts: navigating sequence space intelligently. *Chem Soc Rev* 2015;44(5):1172-1239.
325. **Jiménez-Osés G, Osuna S, Gao X, Sawaya MR, Gilson L et al.** The role of distant mutations and allosteric regulation on LovD active site dynamics. *Nat Chem Biol* 2014;10:431-436.
326. **Rocklin GJ, Chidyausiku TM, Goresnik I, Ford A, Houliston S et al.** Global analysis of protein folding using massively parallel design, synthesis, and testing. *Science* 2017;357(6347):168-175.
327. **Herter S, Medina F, Wagschal S, Benhaïm C, Leipold F et al.** Mapping the substrate scope of monoamine oxidase (MAO-N) as a synthetic tool for the enantioselective synthesis of chiral amines. *Bioorg Med Chem* 2017.
328. **Savile CK, Janey JM, Mundorff EC, Moore JC, Tam S et al.** Biocatalytic asymmetric synthesis of chiral amines from ketones applied to sitagliptin manufacture. *Science* 2010;329(5989):305-309.
329. **Zhang Y, Lamb BM, Feldman AW, Zhou AX, Laverne T et al.** A semisynthetic organism engineered for the stable expansion of the genetic alphabet. *Proc Natl Acad Sci U S A* 2017;114(6):1317-1322.
330. **Mazurkiewicz P, Poelarends GJ, Driessen AJM, Konings WN.** Facilitated drug influx by an energy-uncoupled secondary multidrug transporter. *J Biol Chem* 2004;279(1):103-108.
331. **Mazurkiewicz P, Driessen AJM, Konings WN.** Energetics of wild-type and mutant multidrug resistance secondary transporter LmrP of *Lactococcus lactis*. *Biochim Biophys Acta* 2004;1658(3):252-261.
332. **Mazurkiewicz P, Driessen AJM, Konings WN.** What do proton motive force driven multidrug resistance transporters have in common? *Curr Issues Mol Biol*, Review 2005;7(1):7-21.
333. **Schaedler TA, van Veen HW.** A flexible cation binding site in the multidrug major facilitator superfamily transporter LmrP is associated with variable proton coupling. *FASEB J*, Research Support, Non-U.S. Gov't 2010;24(10):3653-3661.
334. **Madej MG, Dang S, Yan N, Kaback HR.** Evolutionary mix-and-match with MFS transporters. *Proc Natl Acad Sci* 2013;110(15):5870-5874.
335. **Khersonsky O, Tawfik DS.** Enzyme promiscuity: a mechanistic and evolutionary perspective. *Annu Rev Biochem* 2010;79:471-505.
336. **Daley DO, Rapp M, Granseth E, Melen K, Drew D et al.** Global topology analysis of the *Escherichia coli* inner membrane proteome. *Science*, Research Support, Non-U.S. Gov't 2005;308(5726):1321-1323.
337. **Nishino K, Yamaguchi A.** Analysis of a complete library of putative drug transporter genes in *Escherichia coli*. *J Bacteriol* 2001;183(20):5803-5812.
338. **Holdsworth SR, Law CJ.** Functional and biochemical characterisation of the *Escherichia coli* major facilitator superfamily multidrug transporter MdtM. *Biochimie* 2012;94(6):1334-1346.
339. **Reddy VS, Shlykov MA, Castillo R, Sun EI, Saier MH, Jr.** The major facilitator superfamily (MFS) revisited. *FEBS J* 2012;279(11):2022-2035.

340. **Hayashi M, Tabata K, Yagasaki M, Yonetani Y.** Effect of multidrug-efflux transporter genes on dipeptide resistance and overproduction in *Escherichia coli*. *FEMS Microbiol Lett* 2010;304(1):12-19.
341. **Segura A, Molina L, Fillet S, Krell T, Bernal P et al.** Solvent tolerance in Gram-negative bacteria. *Curr Opin Biotechnol* 2012;23(3):415-421.
342. **Udaondo Z, Duque E, Fernandez M, Molina L, de la Torre J et al.** Analysis of solvent tolerance in *Pseudomonas putida* DOT-T1E based on its genome sequence and a collection of mutants. *FEBS Lett* 2012;586(18):2932-2938.
343. **Goffeau A, Park J, Paulsen IT, Jonniaux JL, Dinh T et al.** Multidrug-resistant transport proteins in yeast: complete inventory and phylogenetic characterization of yeast open reading frames with the major facilitator superfamily. *Yeast* 1997;13(1):43-54.
344. **Balakrishnan R, Park J, Karra K, Hitz BC, Binkley G et al.** YeastMine--an integrated data warehouse for *Saccharomyces cerevisiae* data as a multipurpose tool-kit. *Database (Oxford)* 2012;2012:bar062.
345. **Cherry JM, Hong EL, Amundsen C, Balakrishnan R, Binkley G et al.** *Saccharomyces* Genome Database: the genomics resource of budding yeast. *Nucleic Acids Res* 2012;40(Database issue):D700-705.
346. **Rogers B, Decottignies A, Kolaczowski M, Carvajal E, Balzi E et al.** The pleiotropic drug ABC transporters from *Saccharomyces cerevisiae*. *J Mol Microbiol Biotechnol* 2001;3(2):207-214.
347. **Dos Santos SC, Teixeira MC, Dias PJ, Sá-Correia I.** MFS transporters required for multidrug/multixenobiotic (MD/MX) resistance in the model yeast: understanding their physiological function through post-genomic approaches. *Front Physiol*, Review 2014;5:180.
348. **dos Santos SC, Sa-Correia I.** Yeast toxicogenomics: lessons from a eukaryotic cell model and cell factory. *Curr Opin Biotechnol* 2015;33:183-191.
349. **Elena SF, Whittam TS, Winkworth CL, Riley MA, Lenski RE.** Genomic divergence of *Escherichia coli* strains: evidence for horizontal transfer and variation in mutation rates. *International microbiology : the official journal of the Spanish Society for Microbiology* 2005;8(4):271-278.
350. **Gluck-Thaler E, Slot JC.** Dimensions of Horizontal Gene Transfer in Eukaryotic Microbial Pathogens. *PLoS Pathog* 2015;11(10):e1005156.
351. **Koonin EV, Makarova KS, Aravind L.** Horizontal gene transfer in prokaryotes: quantification and classification. *Annu Rev Microbiol* 2001;55:709-742.
352. **Remy E, Nino-Gonzalez M, Godinho CP, Cabrito TR, Teixeira MC et al.** Heterologous expression of the yeast Tpo1p or Pdr5p membrane transporters in *Arabidopsis* confers plant xenobiotic tolerance. *Sci Rep* 2017;7(1):4529.
353. **Dunlop MJ, Dossani ZY, Szmidt HL, Chu HC, Lee TS et al.** Engineering microbial biofuel tolerance and export using efflux pumps. *Mol Syst Biol* 2011;7:487.
354. **Doshi R, Nguyen T, Chang G.** Transporter-mediated biofuel secretion. *Proc Natl Acad Sci*, Research Support, U.S. Gov't, Non-P.H.S. 2013;110(19):7642-7647.
355. **Ling H, Chen B, Kang A, Lee JM, Chang MW.** Transcriptome response to alkane biofuels in *Saccharomyces cerevisiae*: identification of efflux pumps involved in alkane tolerance. *Biotechnol Biofuels* 2013;6(1):95.
356. **Chen B, Ling H, Chang MW.** Transporter engineering for improved tolerance against alkane biofuels in *Saccharomyces cerevisiae*. *Biotechnol Biofuels* 2013;6(1):21.
357. **Nishida N, Ozato N, Matsui K, Kuroda K, Ueda M.** ABC transporters and cell wall proteins involved in organic solvent tolerance in *Saccharomyces cerevisiae*. *J Biotechnol* 2013;165(2):145-152.
358. **Lee JO, Cho KS, Kim OB.** Overproduction of AcrR increases organic solvent tolerance mediated by modulation of SoxS regulon in *Escherichia coli*. *Appl Microbiol Biotechnol* 2014;98(20):8763-8773.
359. **Sun X, Zahir Z, Lynch KH, Dennis JJ.** An antirepressor, SrpR, is involved in transcriptional regulation of the SrpABC solvent tolerance efflux pump of *Pseudomonas putida* S12. *J Bacteriol* 2011;193(11):2717-2725.

360. **Fillet S, Daniels C, Pini C, Krell T, Duque E et al.** Transcriptional control of the main aromatic hydrocarbon efflux pump in *Pseudomonas*. *Environmental microbiology reports* 2012;4(2):158-167.
361. **Mingardon F, Clement C, Hirano K, Nhan M, Luning EG et al.** Improving olefin tolerance and production in *E. coli* using native and evolved AcrB. *Biotechnol Bioeng* 2014.
362. **Moschen I, Bröer A, Galić S, Lang F, Bröer S.** Significance of short chain fatty acid transport by members of the monocarboxylate transporter family (MCT). *Neurochem Res* 2012;37(11):2562-2568.
363. **Sá-Pessoa J, Paiva S, Ribas D, Silva IJ, Viegas SC et al.** SATP (YaaH), a succinate-acetate transporter protein in *Escherichia coli*. *Biochem J*, Research Support, Non-U.S. Gov't 2013;454(3):585-595.
364. **Lin MH, Khnykin D.** Fatty acid transporters in skin development, function and disease. *Biochim Biophys Acta*, Research Support, N.I.H., Extramural Review 2014;1841(3):362-368.
365. **Mukhopadhyay A.** Tolerance engineering in bacteria for the production of advanced biofuels and chemicals. *Trends Microbiol* 2015;23(8):498-508.
366. **Turner WJ, Dunlop MJ.** Trade-Offs in Improving Biofuel Tolerance Using Combinations of Efflux Pumps. *Acs Synth Biol* 2015;4(10):1056-1063.
367. **Rahman SA, Cuesta SM, Furnham N, Holliday GL, Thornton JM.** EC-BLAST: a tool to automatically search and compare enzyme reactions. *Nat Meth*, Research Support, Non-U.S. Gov't 2014;11(2):171-174.
368. **Lam SD, Dawson NL, Das S, Sillitoe I, Ashford P et al.** Gene3D: expanding the utility of domain assignments. *Nucleic Acids Res* 2016;44(D1):D404-409.
369. **Dawson NL, Lewis TE, Das S, Lees JG, Lee D et al.** CATH: an expanded resource to predict protein function through structure and sequence. *Nucleic Acids Res* 2017;45(D1):D289-D295.
370. **Carbonell P, Wong J, Swainston N, Takano E, Turner N et al.** Selenzyme: Enzyme selection tool for pathway design. *Bioinformatics* 2018;in press.
371. **Koch M, Duigou T, Carbonell P, Faulon JL.** Molecular structures enumeration and virtual screening in the chemical space with RetroPath2.0. *J Cheminform* 2017;9(1):64.
372. **Swainston N, Batista-Navarro R, Carbonell P, Dobson PD, Vinaixa M et al.** biochem4j: integrated and extensible biochemical knowledge through graph databases. *PLoSOne* 2017;12(7):e0179130.
373. **Delépine B, Duigou T, Carbonell P, Faulon JL.** RetroPath2.0: A retrosynthesis workflow for metabolic engineers. *Metab Eng* 2018;45:158-170.
374. **Moore JC, Jin HM, Kuchner O, Arnold FH.** Strategies for the *in vitro* evolution of protein function: Enzyme evolution by random recombination of improved sequences. *J Mol Biol* 1997;272(3):336-347.
375. **Kell DB.** Scientific discovery as a combinatorial optimisation problem: how best to navigate the landscape of possible experiments? *Bioessays* 2012;34(3):236-244.
376. **Kell DB, Lurie-Luke E.** The virtue of innovation: innovation through the lenses of biological evolution. *J R Soc Interface* 2015;12(2):20141183.
377. **Copp JN, Hanson-Manful P, Ackerley DF, Patrick WM.** Error-prone PCR and effective generation of gene variant libraries for directed evolution. *Methods Mol Biol* 2014;1179:3-22.
378. **McCullum EO, Williams BA, Zhang J, Chaput JC.** Random mutagenesis by error-prone PCR. *Methods Mol Biol* 2010;634:103-109.
379. **Pritchard L, Corne DW, Kell DB, Rowland JJ, Winson MK.** A general model of error-prone PCR. *J Theoret Biol* 2004;234(4):497-509.
380. **Oates MJ, Corne DW, Kell DB.** The bimodal feature at large population sizes and high selection pressure: implications for directed evolution. In: Tan KC, Lim MH, Yao X, Wang L (editors). *Recent advances in simulated evolution and learning*. Singapore: World Scientific; 2003. pp. 215-240.

381. **Sacks J, Welch W, Mitchell T, Wynn H.** Design and analysis of computer experiments (with discussion). *Statist Sci* 1989;4:409-435.
382. **Bäck T, Fogel DB, Michalewicz Z (editors).** *Handbook of evolutionary computation*. Oxford: IOP Publishing/Oxford University Press; 1997.
383. **Corne D, Dorigo M, Glover F (editors).** *New ideas in optimization*. London: McGraw Hill; 1999.
384. **Goldberg DE.** *The design of innovation: lessons from and for competent genetic algorithms*. Boston: Kluwer; 2002.
385. **Knowles J.** Closed-Loop Evolutionary Multiobjective Optimization. *IEEE Computational Intelligence Magazine* 2009;4(3):77-91.
386. **O'Hagan S, Knowles J, Kell DB.** Exploiting genomic knowledge in optimising molecular breeding programmes: algorithms from evolutionary computing. *PLoS ONE* 2012;7(11):e48862.
387. **Heinemann M, Panke S.** Synthetic biology--putting engineering into biology. *Bioinformatics* 2006;22(22):2790-2799.
388. **Carbonell P, Currin A, Dunstan M, Fellows D, Jervis A et al.** SYNBIOCHEM - a SynBio foundry for the biosynthesis and sustainable production of fine and speciality chemicals. *Biochem Soc Trans* 2016;44(3):675-677.
389. **Currin A, Swainston N, Day PJ, Kell DB.** SpeedyGenes: a novel approach for the efficient production of error-corrected, synthetic gene libraries. *Protein Eng Design Sel* 2014;27:273-280.
390. **Currin A, Swainston N, Day PJ, Kell DB.** SpeedyGenes: exploiting an improved gene synthesis method for the efficient production of error-corrected, synthetic protein libraries for directed evolution. *Meth Mol Biol* 2017;1472:63-78.
391. **Swainston N, Currin A, Green L, Breitling R, Day PJ et al.** CodonGenie: optimised ambiguous codon design tools. *Peer J Comp Sci* 2017;3:e120.
392. **Romero-Rivera A, Garcia-Borràs M, Osuna S.** Computational tools for the evaluation of laboratory-engineered biocatalysts. *Chem Comm* 2017;53(2):284-297.
393. **Welch GR, Somogyi B, Damjanovich S.** The role of protein fluctuations in enzyme action: a review. *Prog Biophys Mol Biol* 1982;39(2):109-146.
394. **Araya CL, Fowler DM.** Deep mutational scanning: assessing protein function on a massive scale. *Trends Biotechnol* 2011;29(9):435-442.
395. **Shin H, Cho BK.** Rational Protein Engineering Guided by Deep Mutational Scanning. *Int J Mol Sci* 2015;16(9):23094-23110.
396. **Rubin AF, Gelman H, Lucas N, Bajjalieh SM, Papenfuss AT et al.** A statistical framework for analyzing deep mutational scanning data. *Genome Biol* 2017;18(1):150.
397. **Starita LM, Fields S.** Deep Mutational Scanning: Library Construction, Functional Selection, and High-Throughput Sequencing. *Cold Spring Harb Protoc* 2015;2015(8):777-780.
398. **Kitzman JO, Starita LM, Lo RS, Fields S, Shendure J.** Massively parallel single-amino-acid mutagenesis. *Nat Methods* 2015;12(3):203-206, 204 p following 206.
399. **Matuszewski S, Hildebrandt ME, Ghenu AH, Jensen JD, Bank C.** A Statistical Guide to the Design of Deep Mutational Scanning Experiments. *Genetics* 2016;204(1):77-87.
400. **LeCun Y, Bengio Y, Hinton G.** Deep learning. *Nature* 2015;521(7553):436-444.
401. **Schmidhuber J.** Deep learning in neural networks: an overview. *Neural networks : the official journal of the International Neural Network Society* 2015;61:85-117.
402. **Knight CG, Platt M, Rowe W, Wedge DC, Khan F et al.** Array-based evolution of DNA aptamers allows modelling of an explicit sequence-fitness landscape. *Nucleic Acids Res* 2009;37(1):e6.
403. **Rugbjerg P, Genée HJ, Jensen K, Sarup-Lytzen K, Sommer MOA.** Molecular Buffers Permit Sensitivity Tuning and Inversion of Riboswitch Signals. *Acs Synth Biol* 2016;5(7):632-638.
404. **Cardinale S, Tueros FG, Sommer MOA.** Genetic-Metabolic Coupling for Targeted Metabolic Engineering. *Cell Rep* 2017;20(5):1029-1037.