De Novo Enzyme Cascades in Whole Cells for the Synthesis of Chiral Cyclic Amines.

Lorna J Hepworth, Scott P. France, Shahed Hussain, Peter Both, Nicholas J Turner, and Sabine L. Flitsch

*ACS Catal.*, **Just Accepted Manuscript** • DOI: 10.1021/acscatal.7b00513 • Publication Date (Web): 14 Mar 2017

Downloaded from http://pubs.acs.org on March 16, 2017

**Just Accepted**

"Just Accepted" manuscripts have been peer-reviewed and accepted for publication. They are posted online prior to technical editing, formatting for publication and author proofing. The American Chemical Society provides "Just Accepted" as a free service to the research community to expedite the dissemination of scientific material as soon as possible after acceptance. "Just Accepted" manuscripts appear in full in PDF format accompanied by an HTML abstract. "Just Accepted" manuscripts have been fully peer reviewed, but should not be considered the official version of record. They are accessible to all readers and citable by the Digital Object Identifier (DOI®). "Just Accepted" is an optional service offered to authors. Therefore, the "Just Accepted" Web site may not include all articles that will be published in the journal. After a manuscript is technically edited and formatted, it will be removed from the "Just Accepted" Web site and published as an ASAP article. Note that technical editing may introduce minor changes to the manuscript text and/or graphics which could affect content, and all legal disclaimers and ethical guidelines that apply to the journal pertain. ACS cannot be held responsible for errors or consequences arising from the use of information contained in these "Just Accepted" manuscripts.
De Novo Enzyme Cascades in Whole Cells for the Synthesis of Chiral Cyclic Amines.

Lorna J. Hepworth\textsuperscript{a}, Scott P. France\textsuperscript{a}, Shahed Hussain\textsuperscript{a,b}, Peter Both\textsuperscript{a}, Nicholas J. Turner\textsuperscript{a}, Sabine L. Flitsch\textsuperscript{a,*}.

\textsuperscript{a}. School of Chemistry, University of Manchester, Manchester Institute of Biotechnology, 131 Princess Street, Manchester M1 7DN, United Kingdom.

\textsuperscript{b}. present address: Dr. Reddy’s Laboratories, Chirotech Technology Centre, 410 Cambridge Science Park, Milton Road, Cambridge, CB4 0PE, United Kingdom.

ABSTRACT

The increasing diversity of reactions mediated by biocatalysts has led to development of multi-step \textit{in vitro} enzyme cascades, taking advantage of generally compatible reaction conditions. The construction of pathways within single whole cell systems is much less explored, yet has many advantages. Herein we report the generation of a successful whole cell \textit{de novo} enzyme cascade for the diastereoselective and/or enantioselective conversion of simple, linear keto acids into valuable cyclic amine products. The pathway starts with carboxylic acid reduction that triggers a transamination, imine formation and subsequent imine reduction. Construction and optimization of the system was achieved by standard genetic manipulation and the cascade required only starting material, amine donor and whole cell catalyst with cofactors provided internally by glucose metabolism. A panel of synthetic keto acids provided access to piperidines in high conversions (up to 93\%) and ee (up to 93\%).
KEYWORDS: biocatalysis, chiral piperidines, enzyme cascades, gene duplication, whole cell biotransformation

The exploitation of synthetic biology methodologies to produce pharmaceutically and industrially relevant synthetic products, using the design principles of biosynthetic retrosynthesis, 1–3 is a highly promising method for chemical synthesis. The ever-expanding biocatalytic toolbox presents an opportunity to design and build cascades that have no biological origin for the production of high value chemicals. A number of enzymatic cascades have been described during the past several years 4–13 although very few have been used as single whole cell biocatalysts, until recently. 14–20 Whereas natural biosynthetic pathways have evolved to be biocompatible, de novo synthetic cascades can potentially raise problems in whole cells, such as generation of bioreactive intermediates, metabolic degradation, transport issues through membranes, heavy cofactor requirements and control of activity of the individual biocatalysts. Although these issues make development of whole cell systems more complex, the overall benefit of having a single (whole cell) biocatalyst that potentially provides biochemical cofactors by glucose metabolism makes cascades in single cells very attractive. Here, we describe the development of a four-enzyme, four-step cascade that was chosen as it allows for the synthesis of high value chiral amines from easily accessible keto acids. The cascade also challenges existing whole cell systems by generating a reactive aldehyde intermediate and by (multi)-stoichiometric requirements for intracellular cofactors ATP and NAD(P)H.

The Design-Build-Test-Analyze (DBTA) cycle 21 was adapted to the generation of novel whole cell in vivo biocatalytic cascades as outlined in Figure 1: (1) designing of the pathway via biocatalytic retrosynthetic analysis, (2) selection of homologs from each enzyme class to test (e.g. based on substrate scope), (3) generation of a library of plasmids containing various gene combinations, (4) introduction of a single plasmid into a suitable chassis and expression
of cascade proteins and (5) testing the system by screening of each whole cell system against appropriate substrate(s). Using this workflow, a variety of pathway scenarios can be quickly tested without need for enzyme isolation and purification and the most successful whole cells taken through to process optimization.

Figure 1. Overview of the Design-Build-Test workflow for the generation of de novo enzyme cascades in whole cell systems. The design follows the principles of biocatalytic retrosynthesis with the catalyst then built at the genetic level and directly tested as a whole cell system. Results from the test phase are fed back to the design and the cycle is repeated until the desired whole cell catalyst is generated.

Chiral secondary amines are ubiquitous functionalities found in biologically active compounds but are not readily accessible through natural biosynthetic pathways. Chemical synthesis of cyclic chiral amines is also challenging and often involves the use of precious transition metal catalysts requiring increased downstream processing, the use of high
pressures or undesirable solvent systems, and frequently results in the production of unwanted intermediate racemates. As such, a facile enzymatic route to these scaffolds, in particular chiral piperidine derivatives, is of great interest.

Disconnecting the target piperidine structures through a biocatalytic retrosynthetic approach suggested keto acids as starting materials for a cascade consisting of carboxylic acid reductase (CAR), ω-transaminase (ω-TA) and imine reductase (IRED) enzymes (Figure 1). Keto acids are stable compounds which are easily obtainable through scalable chemical routes but are potentially also accessible through biosynthetic pathways.

The enzyme cascade is initiated by the ATP- and NADPH-dependent reduction of linear keto acids to the corresponding keto aldehydes, catalyzed by CAR. CARs require posttranslational modification with coenzyme A through the actions of a 4'-phosphopantetheinyl transferase (Sfp), which needs to be included in the whole cell system.

The products are substrates for transaminases, with the less hindered aldehyde group selectively undergoing transformation. The resulting aminoketone products then spontaneously cyclize to imines that are substrates for imine reductases.

The design and construction of the cascade started with generating suitable expression vectors that would programme the host cell to express the chosen biocatalysts. Based on our previous work we started with pPB01, a plasmid devoid of regulatory elements such that each gene could be regulated independently through a monocistronic operon design.

BioBrick cloning strategies were used in the combinatorial assembly of the expression constructs pLH01-08 [pPB01/ATA + CAR + IRED + BsSfp] (Figure 2) using different homologs, which were then used to transform E. coli BL21 (DE3) cells. Protein expression from the new plasmid systems harbored within the cells was induced with isopropyl-β-D-1-thiogalactopyranoside (IPTG) and SDS-PAGE analysis confirmed soluble expression of all enzymes (see Supporting Information, SI).
Figure 2. Plasmids pLH01E08 and pLH09E10 were generated using homologs in a combinatorial fashion for the first and second Design-Build-Test cycles, respectively. Introduction of the plasmids into *E. coli* resulted in transformation of 1a-e to 4a-e via the reaction cascade shown. (ATA-117, ω-transaminase from *Arthrobacter sp.*; SitATA, mutated version of ATA-117; NCAR, carboxylic acid reductase from *Nocardia sp.*; MCAR, carboxylic acid reductase from *Mycobacterium marinum*; (S)-IRED, imine reductase from *Streptomyces sp.* GF3546; (R)-IRED, imine reductase from *Streptomyces sp.* GF3587; *BsSfp*, 4’-phosphopantetheinylation transferase from *Bacillus subtilis*.)

To assess the activity range of this *in vivo* cascade several commercially available, or readily synthesized, keto acid substrates (1a-e) were screened against the cells harboring pLH01-08. It was envisaged that the use of a single whole cell acting as a biocatalyst would eliminate the need to supplement the reactions with expensive exogenous cofactors (NAD(P)⁺, PLP), cofactor regeneration systems, or an external driving force for the ω-TA reaction (lactate dehydrogenase, LDH). Although glucose is present in all respiring cells, it was expected that supplying additional glucose would be required to relieve any metabolic burden on the
cells and enable the cascade reaction to proceed successfully. An excess of racemic D/L-alanine or isopropylamine (IPA) was also supplied to the whole cell in order to drive the reductive amination to completion. Encouragingly, GC analysis of the reaction mixtures revealed the formation of chiral secondary amine for all keto acids tested. In fact, subsequent analysis indicated that the amine product freely diffused out of the cell into the reaction buffer, which also offers several advantages with regards to biocatalyst recycling and product extraction.\textsuperscript{37,38} Comparison of the different vectors pLH01-08 showed that the best conversions were obtained using plasmid pLH02 (Table 1). For further details regarding pLH01 and pLH03-pLH08, see Supporting Information. The ee and de values of amine products 4 are consistent with known IRED selectivities for the reduction of imine 3.\textsuperscript{5,29} Notably, the product of rac-1e with the reported cascade affords rac-4e with high cis-diastereoselectivity, revealing that CAR and TA can convert both enantiomers of their respective substrates in this case during the course of the reaction. Furthermore, The (R)-IRED has previously been shown to be sensitive to pre-existing chirality in a cyclic imine substrate, possibly due to steric factors between the incoming NADPH cofactor and substrate ring substituents.\textsuperscript{5} The stereogenic centre of imine 3e is capable of overriding the inherent selectivity of the IRED and both enantiomers are converted to afford the cis-diastereomers.
Table 1. Synthesis of Mono- and Disubstituted Piperidines Through the Use of Single Whole Cell Expressing Plasmid pLH02

<table>
<thead>
<tr>
<th>Entry</th>
<th>Substrate</th>
<th>Consumption of 4%/</th>
<th>Conv. to imine 3%/</th>
<th>Conv. to keto alcohol%/</th>
<th>Conv. to 4 4%/</th>
<th>4 ee%/ (abs. config.)</th>
<th>4 de%/ (cis or trans)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1a</td>
<td>&gt;99</td>
<td>-</td>
<td>-</td>
<td>47</td>
<td>30 (S)</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>1b</td>
<td>&gt;99</td>
<td>3</td>
<td>47</td>
<td>50</td>
<td>90 (S)</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>1c</td>
<td>&gt;99</td>
<td>51</td>
<td>42</td>
<td>7</td>
<td>93 (S)</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>1d</td>
<td>&gt;99</td>
<td>n.d.</td>
<td>46</td>
<td>54</td>
<td>30 (S)</td>
<td>&gt;98 (cis)</td>
</tr>
<tr>
<td>5</td>
<td>1e</td>
<td>&gt;99</td>
<td>24</td>
<td>19</td>
<td>57</td>
<td>-</td>
<td>&gt;98 (cis)</td>
</tr>
</tbody>
</table>

Reaction conditions: 3 mM substrate, 40 mg mL⁻¹ wet E. coli BL21 (DE3) cells expressing plasmid pLH02, 250 mM D/L-alanine, 50 mM glucose, 500 mM pH 7 NaP, buffer. Reaction incubated at 30 °C (250 rpm) for 24 h to allow full conversion of keto acid substrate, therefore conversion is based on imine:amine:keto alcohol product ratio (absolute conversion to 4a was calculated against a calibration curve using decane as internal standard, conversion to imine and keto alcohol was therefore not calculated for this example). n.d., not detected.

A variety of parameters concerning both protein expression and reaction conditions were investigated, with the following achieving the highest conversion to amine product 4 when expressing plasmid pLH02 in E. coli BL21 (DE3) cells: 16 h protein expression at 20 °C after induction with IPTG (0.8 mM) at OD₆₀₀ 0.6 in LB growth media, followed by whole cell biotransformations (24 h, 30 °C) using 3 mM substrate in 500 mM sodium phosphate buffer containing glucose (50 mM) and either D/L-alanine (250 mM) or IPA (100 mM).
In order to gain further insight into the factors affecting the yield of amine 4 production, a number of optimization experiments were performed using 1a as substrate (for more details see SI). One side-reaction observed was metabolism of the keto aldehydes 2 to the corresponding keto alcohols, which occurred for all substrates using this cascade system (Table 1). The mismatch of the kinetics of the first two steps of the cascade appeared to lead to accumulation of the reactive aldehydes 2, an issue that may have been resolved in natural pathways through evolution.\textsuperscript{38,39} The design cycle of the whole cell system provided us with a number of options to increase \textit{\textit{\omega}}TA activity, including change of promoter or finding more active transaminases, both of which would require extensive screening.

An alternative way in which Nature amplifies the dosage of a given protein in response to environmental stimuli is through gene duplication.\textsuperscript{40,41} Since gene duplication could boost the slow step, \textit{\textit{\omega}}-TA conversion of keto aldehyde, this approach was investigated further. A duplicate ATA-117 gene was introduced into the original construct to give an alternate 5-gene plasmid (pLH09) (Figure 2) in the second DBTA cycle. Gratifyingly, biotransformations of 1a using cells harboring pLH09 showed a 14\% increase in amine production compared with pLH02 (Table 2). This gene duplication strategy was also used to circumvent accumulation of the imine intermediate; a duplicate (\textit{R})-IREDA gene was cloned into the original construct to give a new plasmid pLH10 (Figure 2). Increasing the expression level of (\textit{R})-IREDA resulted in even greater improvements with respect to conversion to amine 4 product (up to 93\% conversion when keto acid 1e was used) (Table 2). Interestingly, the addition of a duplicate (\textit{S})-IREDA gene into constructs pLH01 and pLH03 also drastically improved the conversion to amine 4a seen for these cascades, increasing the conversion from < 5 \% to conversions comparable to those seen for construct pLH02 (See SI). It is thought that an improvement in the IRED step moderates accumulation of the imine intermediate 3, whilst also driving the \textit{\textit{\omega}}-TA reaction towards amine 4 production. These results show that
using gene duplication, varying expression levels for bottleneck proteins could be rapidly assessed, which is an attractive alternative to testing promoter/terminator pairs of varying strengths.\(^{42}\)

### Table 2. Effect of Gene Duplication on Piperidine Production

<table>
<thead>
<tr>
<th>Substrate</th>
<th>Plasmid</th>
<th>Conv. to imine (%)</th>
<th>Conv. to keto alcohol (%)</th>
<th>Conv. to 4(^a)%[^A]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a</td>
<td>pLH09</td>
<td>-</td>
<td>-</td>
<td>61</td>
</tr>
<tr>
<td>1a</td>
<td>pLH10</td>
<td>-</td>
<td>-</td>
<td>67</td>
</tr>
<tr>
<td>1b</td>
<td>pLH09</td>
<td>n.d.</td>
<td>23</td>
<td>77</td>
</tr>
<tr>
<td>1b</td>
<td>pLH10</td>
<td>n.d.</td>
<td>13</td>
<td>87</td>
</tr>
<tr>
<td>1c</td>
<td>pLH09</td>
<td>67</td>
<td>24</td>
<td>9</td>
</tr>
<tr>
<td>1c</td>
<td>pLH10</td>
<td>21</td>
<td>12</td>
<td>67</td>
</tr>
<tr>
<td>1d</td>
<td>pLH09</td>
<td>n.d.</td>
<td>25</td>
<td>75</td>
</tr>
<tr>
<td>1d</td>
<td>pLH10</td>
<td>11</td>
<td>5</td>
<td>84</td>
</tr>
<tr>
<td>1e</td>
<td>pLH09</td>
<td>21</td>
<td>8</td>
<td>71</td>
</tr>
<tr>
<td>1e</td>
<td>pLH10</td>
<td>n.d.</td>
<td>7</td>
<td>93</td>
</tr>
</tbody>
</table>

**Reaction conditions:** 3 mM substrate, 40 mg mL\(^{-1}\) wet *E. coli* BL21 (DE3) cells expressing plasmid pLH09 or pLH10, 250 mM d/L-alanine, 50 mM glucose, 500 mM pH 7 NaPi buffer. Reactions incubated at 30 °C (250 rpm) for 24 h to allow full conversion of keto acid substrate, therefore conversion is based on imine:amine:keto alcohol product ratio (absolute conversion to 4\(^a\) was calculated against a calibration curve using decane as internal standard, conversion to imine and keto alcohol was therefore not calculated for this example). n.d., not detected.

A hybrid cascade system comprised of two separate whole cell catalysts containing MCAR/\(Bs\)Sfp and (R)-IRED, respectively, and a commercial crude cell lysate containing the transaminase ATA-113 was previously identified as a route to chiral piperidines in high yields.\(^5\) Due to the *in vitro* application of the transaminase it was necessary to provide this system with external cofactor regeneration (GDH, glucose) and a driving force for the
reductive amination reaction (LDH, D/L-alanine), as well as the addition of expensive cofactors (NAD$^+$, PLP) (Figure 3a). Implementing the cascade in a single whole cell abolished the need for this exogenous supplementation but necessitated the use of a different transaminase, as the nucleotide sequence for the ATA-113 gene was not available (Figure 3b).

**Figure 3.** Comparison of the requirements of **a.** multi-component one pot hybrid cascade and **b.** single whole cell *in vivo* cascade. Items highlighted in blue represent reaction components added to facilitate the cascade reaction.

ATA-117 was selected as a suitable alternative to ATA-113 due to previous studies where both transaminases were shown to act on related diketone substrates.$^{43}$ In an effort to assess the performance of ATA-117 against ATA-113 in the cascade reaction, keto acid 1a was tested with hybrid cascade conditions using either ATA-117 or ATA-113, with ATA-113 resulting in significantly higher conversions to 4a than when ATA-117 was used (see SI). This difference in activity between the two transaminase homologs suggests that the lower conversions seen for the *in vivo* cascade reported here compared with the hybrid cascade...
described previously can be attributed to the use of ATA-117 rather than an inherent inferiority with using a single whole cell system containing several biocatalysts.

The use of a single whole cell system also allowed for the rapid screening of further, more complex, cascade substrates due to the relative ease of implementing the \textit{in vivo} system compared with the hybrid cascade reaction. \textit{E. coli} BL21 (DE3) cells harboring pLH10 were used for the production of bicyclic amine 4f and thiomorpholine 4g from keto acids 1f and 1g, and were subsequently compared with hybrid cascade reactions using either ATA-117 or ATA-113 for the transaminase step (Table 3). Reactions using ATA-113 again resulted in higher conversions to product amine than their ATA-117 counterparts, whilst systems utilizing ATA-117 (both the \textit{in vivo} system and the hybrid system) displayed comparable conversions.

\textbf{Table 3. Synthesis of Amines 4f and 4g Through Use of \textit{in vivo} or Hybrid Cascade Systems}

<table>
<thead>
<tr>
<th>Cascade System</th>
<th>Conv. to 4f%</th>
<th>Conv. to 4g%</th>
</tr>
</thead>
<tbody>
<tr>
<td>\textit{in vivo} (ATA-117 + MCAR + (R)-IRED)</td>
<td>14</td>
<td>8</td>
</tr>
<tr>
<td>Hybrid (ATA-117 + MCAR + (R)-IRED)</td>
<td>13</td>
<td>8</td>
</tr>
<tr>
<td>Hybrid (ATA-113 + MCAR + (R)-IRED)</td>
<td>30</td>
<td>30</td>
</tr>
</tbody>
</table>
Reaction conditions: \(^{[a]}\) 5 mM substrate, 40 mg mL\(^{-1}\) wet *E. coli* BL21 (DE3) cells expressing plasmid pLH10, 250 mM \(\text{D/L-alanine}\), 50 mM glucose, 500 mM pH 7 NaP buffer. \(^{[b]}\) 5 mM substrate, 75 mg mL\(^{-1}\) *E. coli* BL21 (DE3) cells expressing MCAR, 50 mg mL\(^{-1}\) *E. coli* BL21 (DE3) cells expressions (R)-IRED, 2.5 mg mL\(^{-1}\) ATAE-113 or ATAE-117, 1 mg mL\(^{-1}\) GDH, 0.5 mg mL\(^{-1}\) LDH, 250 mM \(\text{D/L-alanine}\), 100 mM glucose, 1.5 mM NAD\(^+\), 1 mM PLP, 500 mM pH 7 NaP buffer. Reactions incubated at 30 °C (250 rpm) for 24 h to allow full conversion of keto acid substrate, absolute conversion to 4f and 4g was calculated against a calibration curve using decane as internal standard.

Finally, optimized reaction conditions using plasmid pLH10 were used in the preparative-scale synthesis of (S)-4a and (±)-cis-4e, affording 70 mg of 4a (58 % isolated yield, 30 % ee) and 50 mg of 4e (59 % isolated yield, >98 % de) (Scheme 1), thereby demonstrating the utility and future potential of this *in vivo* system for the production of chiral amine building blocks at scale (see SI).

Scheme 1. Isolated yields from preparative-scale biotransformations using *E. coli* cells harboring pLH10

To summarize, ten novel plasmids pLH01-10 were designed and built to express CAR, Sfp, \(\omega\)-TA and IRED in a single cell for the conversion of achiral keto acid substrates into their
valuable chiral amine counterparts, at the expense of D/L-alanine or IPA, and glucose. A panel of keto acids 1a-g were screened against these new constructs, with the expected amines 4a-g produced in all cases. Expression levels of ATA-117 and (R)-IRED proteins were increased through the introduction of gene duplication, leading to increased yields of amine products.

This work demonstrates that it is possible to design and build entirely \textit{de novo} biosynthetic cascades into \textit{E. coli} host cells, with the cells able to meet all cofactor requirements and act as microbial factories for the production of pharmaceutically and industrially relevant chemical scaffolds.

**ASSOCIATED CONTENT**

**Supporting Information**

Nucleotide sequences for plasmid pPB01 and all enzymes used in the study, cloning strategy, general biotransformation protocols, analytical methods, process optimization details and characterization data. This material is available free of charge via the Internet at http://pubs.acs.org.

**AUTHOR INFORMATION**

**Corresponding Author**

Sabine.Flitsch@manchester.ac.uk

**Funding Sources**

The research was supported by Innovative Medicines Initiative Joint Undertaking under grant agreement no 115360, (FP7/2007-2013; and EFPIA companies’ in kind contribution); BBSRC (sLoLa BB/L502005/1 to LJH), (CASE/Pfizer to SPF); EPSRC (CASE/AstraZeneca to SH). NJT and SLF are grateful to the Royal Society for Wolfson Research Merit Awards.
REFERENCES


