Pergamon

Bioorganic & Medicinal Chemistry Letters, Vol. 4, No. 21, pp. 2551-2554, 1994

0960-894X/94 $7.00+0.00

0960-894X(94)00364-5

Directed Evolution of Novel Biosynthetic Pathways: Construction of a Randomly Mutated Gene Library in an *Escherichia coli* Proline Auxotroph and Selection for Growth on Potential Proline Precursors

John D. Sutherland*, Esther J. Wilson and Martin C. Wright.

The Dyson Perrins Laboratory and The Oxford Centre for Molecular Sciences,
South Parks Road, Oxford OX1 3QY, U.K..

Abstract

By introducing a plasmid library into a highly mutagenising *Escherichia coli* proline auxotroph the evolutionary scenario of gene doubling and mutagenesis has been mimicked allowing for evolved growth on certain potential proline precursors.

Introduction

Current biological catalysts are the result of Darwinian evolution by mutation and selection for novel or improved function and we have been interested in ascertaining whether presumed modes of biosynthetic pathway evolution can be mimicked to produce new biosynthetic pathways to materials of interest currently made by conventional synthesis. In an attempt to demonstrate the feasibility of this approach we have focussed, initially, on the simple natural metabolite, proline **1**. In previous papers we have described the growth of an *Escherichia coli* proline auxotroph, strain X7026, on the unnatural precursor Δ^1-pyrroline-2-carboxylic acid **2**[1,2,3,4]. Interestingly it was found that *E. coli* already possesses the metabolic machinery to convert **2** to proline by direct reduction and not *via* initial isomerisation to Δ^1-pyrroline-5-carboxylic acid **3**, Fig. 1.

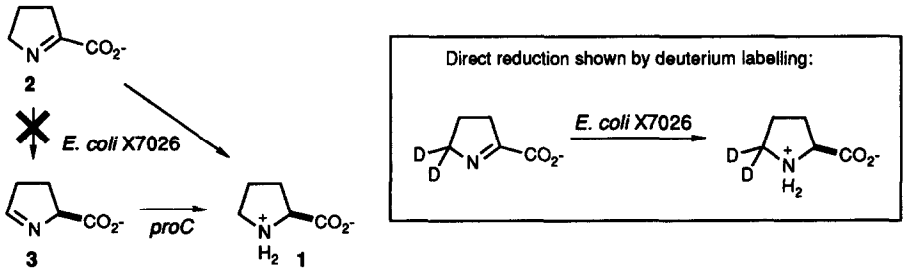

Fig. 1. Conversion of Δ^1-Pyrroline-2-carboxylate 2 to Proline 1 in *E. coli*

In an attempt to identify the enzyme responsible for the *in vivo* reduction we selected for plasmids from a random *E. coli* library which conferred a faster growth phenotype on **2**. The common insert in the plasmids selected was associated with membrane permeability and not with the reduction *per se*. At the same time as conducting further experiments to identify the enzyme responsible for reducing **2** we have sought to extend the approach to other potential proline precursors.

Results and Discussion

A variety of other compounds which can potentially be converted into proline by some sort of biotransformation have been considered, Fig. 2.

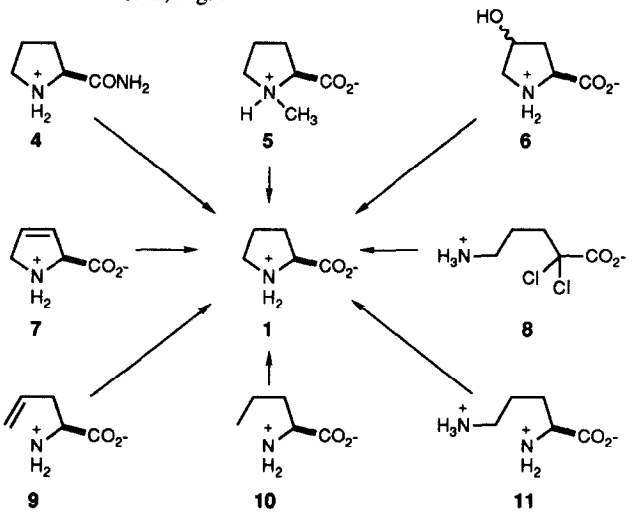

Fig. 2. Potential Proline Precursors

The potential biotransformations required to convert these precursors to proline are briefly indicated in Table 1.

Table 1: Potential Biotransformations Producing Proline

Conversion	Biotransformation
4 to **1**	Amide bond hydrolysis
5 to **1**	Demethylation by nucleophilic displacement or oxidation/hydrolysis
6 to **1**	Elimination and isomerisation to **2** followed by reduction
7 to **1**	Isomerisation to **2** or **3** followed by reduction
8 to **1**	Hydrolysis to **2** followed by reduction or reduction followed by nucleophilic cyclisation
9 to **1**	Protonation and ring closure
10 to **1**	Oxidative C-H bond insertion
11 to **1**	Transamination of α- or δ-amino group to give **2** or **3** respectively followed by reduction.

When *E. coli* X7026 was plated on minimal media containing compounds **4-11** in place of proline it was found that only compound **4** supported growth and then only to a small extent. It was reasoned that the machinery to effect the desired biotransformations was not in place and that growth would only be possible if mutation and/or gene doubling took place. Gene doubling can be simulated by introduction of a plasmid library of *E. coli* X7026 into the same strain. Accordingly a large library ($2x10^4$ colonies) was constructed in *E. coli* X7026 by subcloning 5-10kb *Sau* 3A fragments of *E. coli* X7026 chromosomal DNA into the *Bam* HI site of the high copy-number plasmid, pJS62[1]. This library was then used to inoculate minimal liquid media containing compounds **4-11** and trace amounts of proline to allow for some initial growth. After growth to saturation plasmid DNA was prepared and used to retransform *E. coli* X7026 which was then plated on minimal solid medium containing the appropriate compound. By this procedure faster growing colonies were isolated only for compound **4** the implication being that if growth on the other precursors is to be possible then mutation of existing, doubled genes will be necessary. The further characterisation of these colonies which grow faster on **4** will be discussed in the following paper.

Several methods of producing large numbers of mutations were investigated. Of these the use of mutator strains[5] proved to be simplest and most effective. To produce a mutator strain which was also a non-reverting proline auxotroph it was necessary to introduce a suitable mutator allele into an *E. coli* proline auxotroph. A P1 lysate of strain SJ1860 (*proAB*$^+$, *argE*$^+$, *zaf*-13::Tn10, *mutD5*) was prepared and used to transduce strain AB1157[6] (another proline auxotroph carrying a smaller chromosomal deletion) to tetracycline resistance. Eighteen colonies were produced by transduction, these were tested for mutator activity by reversion of the Arg$^-$ phenotype. Fourteen of the colonies showed mutator activity, giving reversion frequencies of $1x10^{-5}$ to $5x10^{-5}$ after overnight growth in rich medium[7]. One of the mutator$^+$, Pro$^-$ colonies was labelled MW1000 and was used in further experiments[8]. To engineer a situation mimicking gene doubling and mutation we have simply introduced the X7026 chromosomal library into the mutator strain MW1000. After growth in rich medium vast numbers of doubled and mutated gene permutations are possible. The *mutD5* allele encodes a mutant ε-subunit of DNA polymerase III resulting in production of a polymerase deficient in proof-reading and therefore error-prone (*ca.* 10^3 > wild-type)[9]. Plasmid replication however depends on both DNA polymerases III and I[10], DNA polymerase I only being required for part of the plasmid sequence. To verify that *mutD5* mutator strains can mutate harboured plasmids we carried out an initial study using the blue/white selection for *lacZ* mutants afforded by plasmid pJS62 in the presence of IPTG and X-gal. Both inactivating mutants (blue to white obtained at $3{\cdot}3x10^{-2}$ efficiency) and reversions (white to blue, obtained for three different reversions at efficiencies of $7{\cdot}1x\ 10^{-5}$, $7{\cdot}2x10^{-5}$ and $2{\cdot}0x10^{-6}$) were sequenced. The high efficiency of these mutations suggests that the *mutD5* allele is indeed capable of producing a broad spectrum of plasmid mutations in the polylinker coding region and therefore that the mutated library in MW1000 should contain the perturbations we require. We estimate that a diversity of $>10^{10}$ has been produced.

In order to simulate an evolutionary scenario, *E. coli* strain MW1000 harbouring the library was grown under mutating conditions in minimal medium[11] containing compounds **4-11**, cycles of growth were carried out with progressively decreasing amounts of proline also in the medium, Fig. 3.

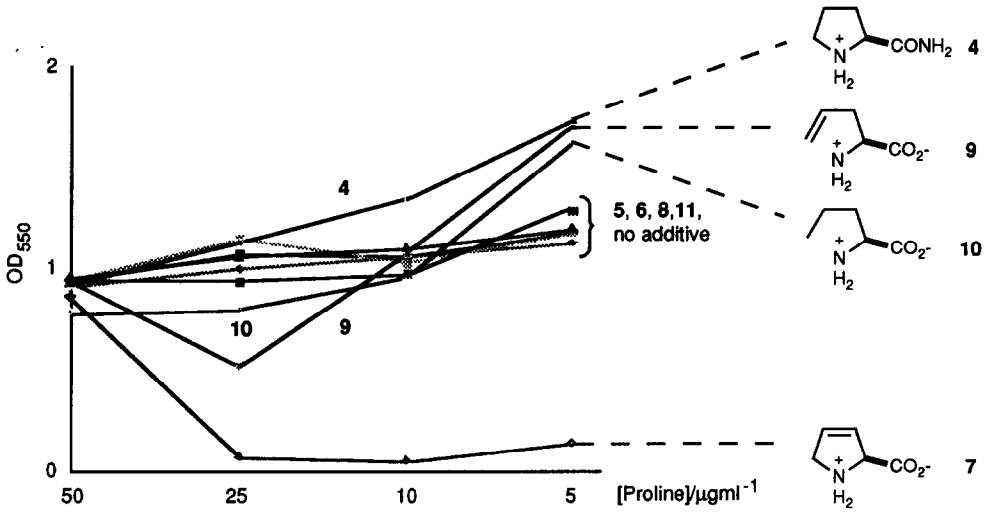

Fig. 3. Evolved Growth on Proline Precursors

As can be seen from Fig. 3, extremely interesting results were obtained from this experiment. After several days, growth on prolinamide **4** showed an increase relative to a control in which no potential proline precursor was added. Coupled with the results from the non-mutant library experiments this suggests that there is a gene product which can hydrolyse **4** to proline and that over expression of the encoding gene or mutagenesis of the gene or its regulatory region increases activity. Identification of this gene is described in the following paper. Dehydroproline **7** is known to be toxic in the absence of endogenous proline[12] and the results of feeding experiments with this compound are in accord with this fact. Perhaps most interestingly, allylglycine **9** and norvaline **10** initially proved detrimental to growth but under conditions of proline starvation and after mutation, supported increased growth relative to the zero control. The biotransformations required to convert these compounds to proline, Table 1, are unusual and further investigation of these mutants should be extremely exciting.

Acknowledgements

We thank the SERC for support and Professor J. E. Baldwin for helpful discussions.

References & Footnotes

1 Sutherland, J. D.; Wilson, E. J.; Wright, M. C. *Bioorg. Med. Chem. Lett.* **1993**, *3*, 1185.

2 Lewis, M. L.; Martin, S. L.; Rowe, C. J.; Sutherland, J. D.; Wilson, E. J.; Wright, M. C. *Bioorg. Med. Chem. Lett.* **1993**, *3*, 1189.

3 Lewis, M. L.; Rowe, C. J.; Sewald, N.; Sutherland, J. D.; Wilson, E. J.; Wright, M. C. *Bioorg. Med. Chem. Lett.* **1993**, *3*, 1193.

4 Lewis, M. L.; Martin, S. L.; Rowe, C. J.; Sutherland, J. D.; Wilson, E. J.; Wright, M. C. *Bioorg. Med. Chem. Lett.* **1993**, *3*, 1197.

5 Degnen, G. E.; Cox, E. C. *J. Bacteriol.* **1974**, *117*, 477.

6 AB1157: F$^-$, *thr-1, ara-14, leuB6,* Δ(*gpt-proA*)62, *lacY1, tsx-33, supE44, galK2,* λ$^-$, rac-, *hisG4* (Oc), *rfbD1, mgl-51, rpsL31, kdgK51, xyl-5, mtl-1, argE3* (Oc), *thi-1*, Howard-Flanders, P.; Simson, E.; Theriot, L. *Genetics* **1964**, *49*, 237.

7 Six of the colonies however also grew on minimal medium lacking proline, this was caused by the proximity of *proAB* to *zaf*-13 on the chromosome resulting in partial cotransduction.

8 In *ΔproAB* strains partial reversion to a Pro$^+$ phenotype is possible by cross-feeding from the arginine biosynthetic pathway, this phenomenon becomes significant only in *argD, argR* mutants in the presence of arginine in the growth medium (Eckhardt, T.; Leisinger, T. *Mol. Gen. Genet.* **1975**, *138*, 225). Three mutations are thus required to obtain a Pro$^+$ strain from MW1000: *argD, argR* and *argE*$^+$ but this was never observed in cultures of the strain even under mutation conditions.

9 Fowler, R. G.; Schaaper, R. M.; Glickman, B. W. *J. Bacteriol.* **1986**, *167*, 130.

10 Gutterson, N. I.; Koshland, J. E. Jnr. *Proc. Natl. Acad. Sci.* **1983**, *80*, 4894.

11 Mutation efficiencies using the *mutD5* allele are highest in rich medium or in minimal medium containing thymidine[9].

12 Leisinger, T. in *Escherichia Coli and Salmonella Typhimurium Cellular and Molecular Biology*, Ed. in Chief Neidhart, F. C.; American Society for Microbiology: Washington, D.C., 1987; Vol 1, pp. 345-351.

(*Received in Belgium* 28 *July* 1994; *accepted* 21 *September* 1994)