Cell Survival Enabled by Leakage of a Labile Metabolic Intermediate

Encarnación Medina-Carmona [1,2] Luis I. Gutierrez-Rus,[1] Fadia Manssour-Triedo,[1] Matilda S. Newton [†,3,4] Gloria Gamiz-Arco,[1] Antonio J. Mota,[5] Pablo Reiné,[6] Juan Manuel Cuerva,[6] Mariano Ortega-Muñoz [6] Eduardo Andrés-León [7] Jose Luis Ortega-Roldan,[2] Burckhard Seelig [3,4] Beatriz Ibarra-Molero,[1] and Jose M. Sanchez-Ruiz [*,1]

[1]Departamento de Quimica Fisica, Facultad de Ciencias, Unidad de Excelencia de Quimica Aplicada a Biomedicina y Medioambiente (UEQ), Universidad de Granada, Granada, Spain

[2]School of Biosciences, University of Kent, Canterbury, United Kingdom

[3]Department of Biochemistry, Molecular Biology and Biophysics, University of Minnesota, Minneapolis, MN

[4]BioTechnology Institute, University of Minnesota, St Paul, MN

[5]Departamento de Quimica Inorganica, Facultad de Ciencias, Unidad de Excelencia de Quimica Aplicada a Biomedicina y Medioambiente (UEQ), Universidad de Granada, Granada, Spain

[6]Departamento de Quimica Organica, Facultad de Ciencias, Unidad de Excelencia de Quimica Aplicada a Biomedicina y Medioambiente (UEQ), Universidad de Granada, Granada, Spain

[7]Unidad de Bioinformática, Instituto de Parasitología y Biomedicina "Lopez Neyra", CSIC, Armilla, Granada, Spain

[†]Present address: Department of Molecular, Cellular, and Developmental Biology, University of Colorado Boulder, Boulder, CO; and Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, Boulder, CO.

***Corresponding author:** E-mail: sanchezr@ugr.es.

Associate editor: Belinda Chang

Abstract

Many metabolites are generated in one step of a biochemical pathway and consumed in a subsequent step. Such metabolic intermediates are often reactive molecules which, if allowed to freely diffuse in the intracellular milieu, could lead to undesirable side reactions and even become toxic to the cell. Therefore, metabolic intermediates are often protected as protein-bound species and directly transferred between enzyme active sites in multi-functional enzymes, multi-enzyme complexes, and metabolons. Sequestration of reactive metabolic intermediates thus contributes to metabolic efficiency. It is not known, however, whether this evolutionary adaptation can be relaxed in response to challenges to organismal survival. Here, we report evolutionary repair experiments on *Escherichia coli* cells in which an enzyme crucial for the biosynthesis of proline has been deleted. The deletion makes cells unable to grow in a culture medium lacking proline. Remarkably, however, cell growth is efficiently restored by many single mutations (12 at least) in the gene of glutamine synthetase. The mutations cause the leakage to the intracellular milieu of a highly reactive phosphorylated intermediate common to the biosynthetic pathways of glutamine and proline. This intermediate is generally assumed to exist only as a protein-bound species. Nevertheless, its diffusion upon mutation-induced leakage enables a new route to proline biosynthesis. Our results support that leakage of sequestered metabolic intermediates can readily occur and contribute to organismal adaptation in some scenarios. Enhanced availability of reactive molecules may enable the generation of new biochemical pathways and the potential of mutation-induced leakage in metabolic engineering is noted.

Key words: **auxotrophy rescue, prototrophy restoration, metabolic innovation, labile metabolic intermediates, evolutionary repair experiments, laboratory evolution.**

Introduction

Outline of This Work

Many metabolic pathways are of ancient origin and have evolved over long periods of time (Noda-Garcia et al. 2018). Yet, new pathways can also emerge in short timescales in response, for instance, to the presence of anthropogenic chemicals in the environment (Copley 2009). Determining the molecular mechanisms responsible for the emergence and evolution of metabolic pathways is important for our understanding of how life evolves. Furthermore, detailed information about these mechanisms may suggest new tools for synthetic biology and metabolic engineering. Re-enacting the natural evolution of metabolic pathways in the laboratory is certainly challenging. However, fundamental mechanisms of metabolic adaptation can be revealed through experiments with cells in which a crucial enzyme has been deleted. The deletion

blocks a metabolic pathway, makes the cells auxotrophic for its final product, and allows evolutionary experiments on the restoration of prototrophy to be performed in the laboratory. A number of such evolutionary repair experiments (Patrick et al. 2007; Mcloughlin and Copley 2008; Kim et al. 2010; Digianantonio and Hecht 2016; Digianantonio et al. 2017; Kim et al. 2019) have provided support for the proposed role of enzyme promiscuity in the emergence of new biochemical pathways (Schulenburg and Miller 2014; Copley 2015; Noda-Garcia et al. 2018; Peracchi 2018). Note that we are using the term "enzyme promiscuity" in a general sense, that is, including the enzyme capability to catalyze a given chemical reaction with different substrates (substrate scope), as well as the capability to catalyze different reactions (catalytic promiscuity). Certainly, the low-level promiscuous activities displayed by many enzymes often have no physiological role (Khersonsky and Tawfik 2010; Copley 2015), yet, their enhancement upon mutation may allow the organism to meet new challenges by patching novel pathways. For instance, Kim et al. (2019) found that a four-step pathway assembled from promiscuous enzymes could restore growth in a knockout *Escherichia coli* strain in which an enzyme crucial for the synthesis of an essential cofactor, pyridoxal phosphate, had been deleted. Also, Patrick et al. (2007) screened about a hundred single-gene knockouts from the Keio collection and found that about 20% of these auxotrophs could be rescued by overexpression of noncognate genes.

Here, we report evolutionary repair experiments with *E. coli* in which an enzyme crucial for the biosynthesis of the amino acid proline has been deleted. Strikingly, our results depart from similar studies in the literature in that we find reversion to prototrophy through a mechanism that does not involve the recruitment of a low-level promiscuous activity or a bypass by tweaking metabolic flux (Kim et al. 2019) but, rather, the mutation-induced leakage to the intracellular milieu of a labile metabolic intermediate. Specifically, the rescuing mutations cause the leakage to the intracellular milieu of a highly reactive phosphorylated intermediate common to the biosynthetic pathways of glutamine and proline.

The rescue mechanism discovered in this work emerges as particularly striking when considering that labile (i.e., highly reactive) metabolic intermediates are typically protected as protein-bound species and directly transferred between enzyme active sites in multi-functional enzymes, multi-enzyme complexes, and metabolons (Srere 1987; Pareek et al. 2021). Sequestration of labile metabolites helps prevent undesirable side reactions and potentially toxic effects and thus contributes to metabolic efficiency. Mutations that impair metabolite protection could therefore be expected to be deleterious and efficiently purged during the course of evolution. However, our results support that leakage of one sequestered intermediate may readily occur and even enable organismal survival in some scenarios.

In the next section of this introduction, we provide the general metabolic context for this work by expounding the rationale behind the assumption of extensive metabolite sequestration and the supporting experimental evidence. In subsequent sections of this introduction, we supply basic information on the proline and glutamine biosynthetic pathways, highlight the main features of the rescue mechanism uncovered in this work, and anticipate some of its more relevant implications.

Restriction of Free Diffusion (i.e., "Sequestration") of Metabolic Intermediates

Metabolic pathways are generally depicted as linear (or cyclic) sequences of chemical species connected by arrows. Each arrow is typically labeled with the name of the enzyme (or with the name of the gene encoding the enzyme) that catalyzes the chemical transformation defined by the arrow. Many chemical species in a metabolic pathway are therefore the product of the reaction catalyzed by one enzyme as well as the substrate of another enzyme in the pathway. These transient species are often referred to as metabolic intermediates. In a diluted solution, the metabolic intermediate that is produced by one enzyme would diffuse to eventually encounter the next enzyme in the pathway. However, such "diffusive metabolism" is unlikely to occur to a substantial extent within a cell, given that the cytosol is a crowded environment where complex and tightly regulated networks of interconnected pathways take place. It has long been recognized (Srere 1987; Pareek et al. 2021) that, in this scenario, free diffusion of metabolic intermediates may have detrimental effects. For instance, the chemical instability of intermediates may impair the flux through the metabolic pathway and give rise to processes that interfere with other pathways. Furthermore, a substantial number of metabolic intermediates are actually toxic for the cell, and their free concentration must be kept below certain levels (Lee et al. 2020). Examples include, not only reactive oxygen species, but also metabolites with reactive groups that can chemically modify DNA and proteins, as well as metabolites that can aberrantly trigger neurotransmitter receptors. It has been noted (Lee et al. 2020) that cross-referencing metabolites with toxicity databases reveals many metabolites with toxic profiles.

Of course, known organisms have evolved a diversity of molecular mechanisms to prevent or limit to a substantial extent the free diffusion of metabolic intermediates, that is, to "sequester" them. These mechanisms have been extensively reviewed in the literature (see, for instance; Srere 1987; Miles et al. 1999; Huang et al. 2001; Tullman-Ercek 2015; Schmitt and An 2017; Grunwald 2018; Kastritis and Gavin 2018; Pareek et al. 2021; Zhang and Fernie 2021) and only a brief summary will be provided here. In enzymes that catalyze multi-step reactions, intermediates may remain bound in the active site and thus prevented to interact with other molecules present in the intracellular milieu. In some bifunctional enzymes, intermediates produced in one active site are transferred to a second active site through an intramolecular tunnel, thus minimizing their release to the intracellular milieu.

Enzymes that catalyze successive steps in a metabolic pathway often form a multi-enzyme complex in such a way that the intermediate (the product of the first step and the substrate of the second) is directly channeled between the two active sites. Cluster channeling of intermediates occurs in metabolons, which are temporary associations of sequential metabolic enzymes in which multiple copies of each enzyme are present. Already about 35 years ago, Paul Srere summarized evidence of metabolons in 14 metabolic pathways (see table 1 in Srere 1987).

A prominent role of enzyme association in metabolic pathways should not come as a surprise, given that interaction proteomics studies indicate that proteins rarely function in isolation and that most proteins in a cell (possibly about 80% or more) interact with other proteins (Gavin et al. 2006; Berggård et al. 2007; Gupta et al. 2022). Indeed, experimental evidence for multi-enzyme complexes and metabolons has come primarily from analyses of protein–protein interactions (for a recent review, see Schmitt and An 2017). Such studies have supported the existence of associations between enzymes in many key biochemical pathways, including, for instance, glycolysis, the Krebs cycle, oxidative phosphorylation, and pathways of amino acid metabolism (see Kastritis and Gavin 2018, and references therein). Direct evidence for metabolite transfer between active sites in enzyme complexes and metabolons is certainly more challenging to obtain. Yet, a recent review on plant metabolism (Zhang and Fernie 2021) has listed 18 metabolic pathways in which protein–protein interactions supporting channeling have been determined, with direct evidence of channeling being available in 7 of them (see table 1 in Zhang and Fernie 2021), including the Krebs cycle and glycolysis (see fig. 2 in Zhang and Fernie 2021).

Overall, it emerges that restriction of free diffusion (i.e., sequestration) of metabolic intermediates is widespread and helps prevent undesirable side processes and toxic effects, thus ensuring metabolic efficiency (Srere 1987; Miles et al. 1999; Huang et al. 2001; Tullman-Ercek 2015; Schmitt and An 2017; Kastritis and Gavin 2018; Pareek et al. 2021; Zhang and Fernie 2021). Certainly, we may expect different metabolic intermediates to be sequestered to different extents, depending, among other factors, on the molecular mechanism responsible for the restriction of free diffusion. In fact, metabolons are sometimes supposed to be more prone to allow metabolite "leakage," to use a term that has been used in the literature of the field (Pareek et al. 2021), as compared, for instance, with tunneling in multi-functional enzymes. It is also clear that sequestration relies on protein–metabolite interactions and, therefore, that mutations in the proteins involved could impair sequestration and increase the leakage of the metabolite to the intracellular milieu. Yet, those mutations may be expected to be detrimental and to be efficiently purged by natural selection. Contrary to this expectation, however, the results we report here support that the mutation-induced leakage of metabolic intermediates may readily occur and contribute to organismal adaptation in some scenarios.

Metabolic Intermediate Sequestration in the Biosynthetic Pathways of Proline and Glutamine

In most cells, the amino acids proline and glutamine are synthesized from glutamate, which is in turn synthesized from α-ketoglutarate, an intermediate in the Krebs cycle (fig. 1). The first step of both biosynthetic pathways is the activation of glutamate through phosphorylation to yield a highly reactive metabolic intermediate, γ-glutamyl phosphate. In the proline biosynthetic pathway, glutamate phosphorylation is performed by γ-glutamyl kinase; while in the glutamine biosynthetic pathway, it is carried out by a different enzyme, glutamine synthetase.

γ-Glutamyl phosphate is a highly reactive molecule, generally thought to exist only as a protein-bound species (Csonka and Leisinger 2007; Fichman et al. 2015). γ-Glutamyl phosphate free in solution quickly undergoes a spontaneous intramolecular chemical alteration to yield a cyclized molecule, which is useless for amino acid biosynthesis (Csonka and Leisinger 2007; Fichman et al. 2015). Therefore, the enzymes responsible for the first and second steps of the proline biosynthetic pathway, γ-glutamyl kinase, and γ-glutamyl phosphate reductase are believed to form a complex in *E. coli*, thus allowing the channeling of the labile γ-glutamyl phosphate directly from the kinase active site to the reductase active site and preventing undesirable side reactions (Smith et al. 1984; Csonka and Leisinger 2007; Fichman et al. 2015). This notion is supported by the fact that the orthologs of these enzymes in animals and vascular plants form a single bifunctional enzyme, pyrroline-5-carboxylate synthase (Csonka and Leisinger 2007; Fichman et al. 2015).

In the case of the glutamine biosynthetic pathway, glutamine synthetase (encoded by *glnA*) catalyzes the two steps of the synthesis of glutamine from glutamate. Therefore, the γ-glutamyl phosphate formed by phosphorylation of glutamate remains bound to the active site of glutamine synthetase until ammonia binds and reacts with it to yield glutamine, which is released from the active site. It is worth noting that the glutamine thus synthesized is not only used in protein synthesis, but, in addition, its amide group provides the nitrogen required for the biosynthesis of many metabolites (Eisenberg et al. 2000; Stadtman 2004; see also fig. 1 in Stadtman et al. 1968). That is, the second step of the reaction catalyzed by glutamine synthetase provides the main route for the assimilation of ammonia, a major pool of inorganic nitrogen in cells. In fact, glutamine synthetase is a crucial and very ancient enzyme for nitrogen assimilation in all cells (Kumada et al. 1993; De Carvalho Fernandes et al. 2022).

Rescue of Proline Auxotrophy by Mutations in the Gene of Glutamine Synthetase

Deletion of γ-glutamyl kinase, the first enzyme in the proline biosynthetic pathway, prevents the formation of γ-glutamyl phosphate in this biochemical route, thus blocking proline biosynthesis and making *E. coli* cells unable to grow in a medium lacking proline (fig. 2). That is, deletion of γ-glutamyl kinase makes *E. coli* cells auxotrophic for proline.

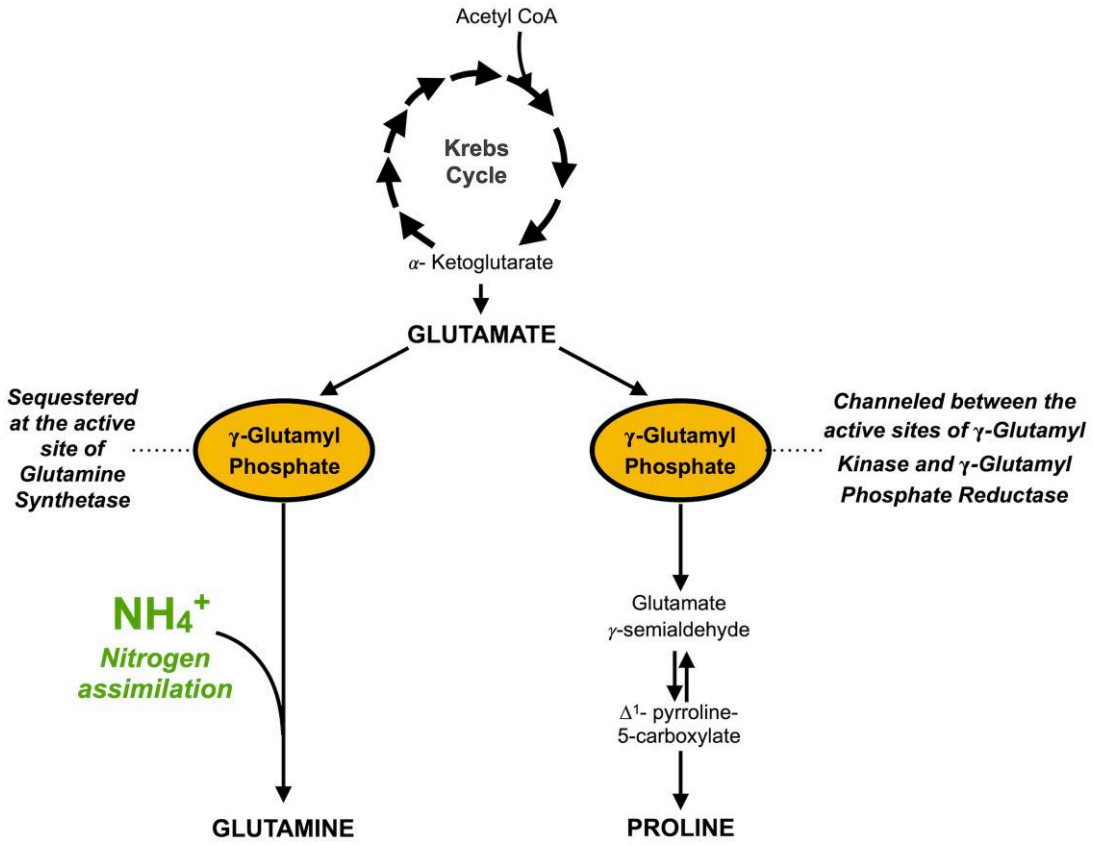

FIG. 1. Annotated flow chart for the glutamine and proline biosynthetic pathways. The first step of both pathways is the activation of glutamate through phosphorylation to yield γ-glutamyl phosphate, a highly reactive metabolic intermediate which is protected as a protein-bound species in the glutamine biosynthetic route and through channeling in the proline biosynthetic route. Note that the second step of glutamine biosynthesis provides a major route for the incorporation of inorganic nitrogen into metabolism. Glutamine synthetase is thus a central enzyme in nitrogen metabolism.

Remarkably, we have found that the capability to grow in the absence of proline in the culture medium (i.e., proline prototrophy) is efficiently restored by many single mutations in the gene of glutamine synthetase.

As already mentioned, efficient biosynthesis of glutamine requires that the labile γ-glutamyl phosphate intermediate remains bound to the active site of glutamine synthetase until ammonia binds and the second step of the synthesis takes place. Furthermore, glutamine synthetase is a very old enzyme, plausibly even older than the last universal common ancestor (Kumada et al. 1993; De Carvalho Fernandes et al. 2022). Therefore, sequestration of γ-glutamyl phosphate in its active site is likely a primordial adaptation. Yet, we show here that the rescuing mutations allow γ-glutamyl phosphate to leak from the active site of glutamine synthetase to the intracellular milieu (fig. 2) while preserving the key role of this enzyme in nitrogen assimilation. Furthermore, we provide evidence that, despite its lability, a substantial fraction of the leaked γ-glutamyl phosphate diffuses to reach γ-glutamyl phosphate reductase, thus restoring proline biosynthesis (fig. 2).

Overall, our results show that this adaptation can be relaxed and that leakage of sequestered intermediates to the intracellular milieu can readily occur and contribute to organismal adaptation in some scenarios. Rescue of a block in an essential route by leakage of a sequestered intermediate from a different route, as demonstrated here, is one such scenario. Another interesting possibility is that leakage can contribute to metabolic innovation, in the sense that it may enable the emergence of new metabolic phenotypes (chapter 2 in Wagner 2011). Models of metabolic pathway emergence and evolution often emphasize the acquisition of new reactions through horizontal gene transfer and promiscuous enzyme functionalities (Pál et al. 2005; Schulenburg and Miller 2014; Copley 2015; Noda-Garcia et al. 2018; Peracchi 2018). However, labile metabolic intermediates may undergo a variety of chemical and enzymatic transformations. Their leakage may thus provide an additional source of new reactions and enable the generation of new biochemical pathways. We suggest that mutation-induced leakage may find applications in metabolic engineering.

Results and Discussion

Proline Auxotrophy Caused by Deletion of the First Enzyme of the Proline Bioshynthetic Pathway is Efficiently Rescued by Many Single Mutations in the Gene of Glutamine Synthetase

In most organisms, proline is synthesized from glutamate through a pathway that involves four steps (Csonka and

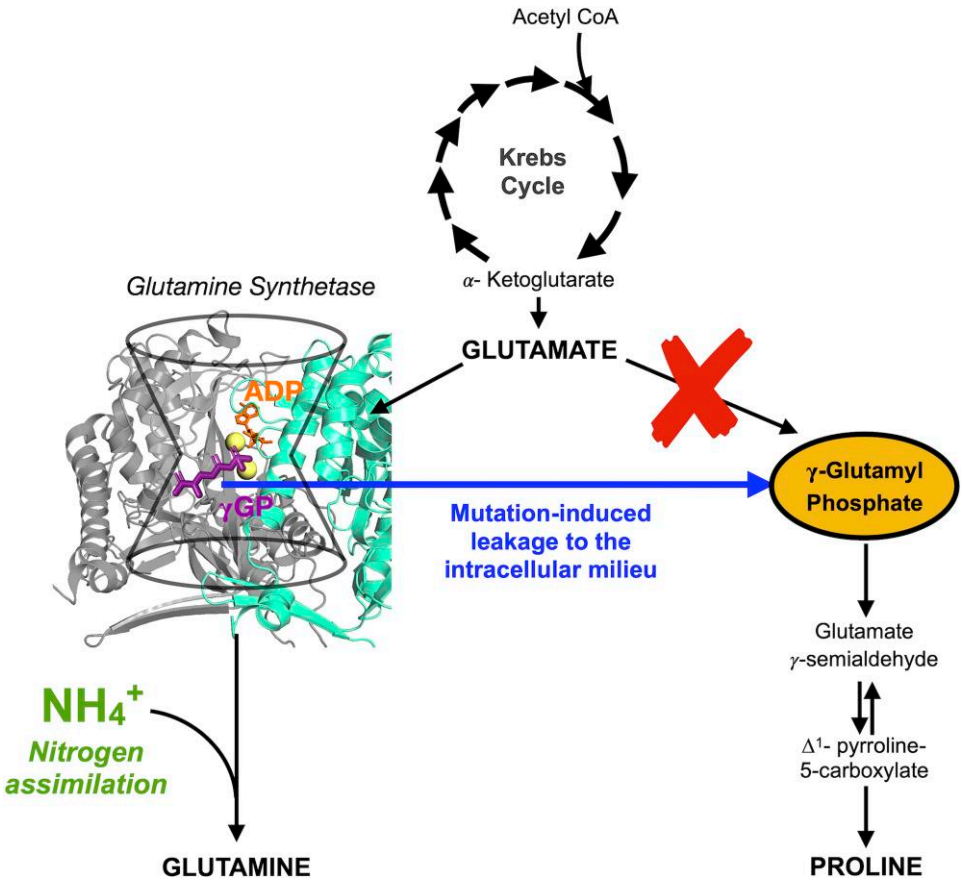

Fig. 2. Mechanism of auxotrophy rescue (i.e., prototrophy restoration) uncovered in this work. Deletion of γ-glutamyl kinase in *Escherichia coli* prevents the activation of glutamate to γ-glutamyl phosphate in the biochemical route to proline (cross), thus blocking proline biosynthesis and making cells unable to grow in the absence of proline in the culture medium. Growth could in principle be enabled by the γ-glutamyl phosphate formed in the glutamine biosynthetic pathway but, for this to happen, some of the γ-glutamyl phosphate tightly bound to the active site of glutamine synthetase (labeled γGP in this figure) must leak to the intracellular milieu and diffuse to reach γ-glutamyl phosphate reductase in the proline biosynthetic pathway. Glutamine synthetase is a very ancient enzyme and it is crucial for nitrogen metabolism. Wild-type *E. coli* glutamine synthetase does not significantly leak γ-glutamyl phosphate to the intracellular milieu, likely reflecting the evolutionary adaptation for efficient glutamine biosynthesis and nitrogen assimilation. However, glutamine synthetase variants with suitable mutations enable γ-glutamyl phosphate leakage and thus prototrophy restoration, while keeping sufficient levels of glutamine biosynthesis and nitrogen assimilation.

Leisinger 2007; Fichman et al. 2015; fig. 3). The first step is the phosphorylation of glutamate to γ-glutamyl phosphate, which is followed by reduction to γ-glutamyl semialdehyde, cyclization to Δ^1-pyrroline-5-carboxylate, and further reduction to proline. While the cyclization step is spontaneous, the first, second, and fourth steps are catalyzed by the enzymes γ-glutamyl kinase, γ-glutamyl phosphate reductase, and Δ^1-pyrroline-5-carboxylate reductase, respectively. In *E. coli*, the genes encoding these enzymes are known as *proB*, *proA*, and *proC*, respectively (see fig. 3).

Here, we have studied a *ΔproB* strain from the Keio collection (Baba et al. 2006). This *E. coli* Keio knockout cannot produce γ-glutamyl kinase, the first enzyme of the proline biosynthetic pathway, and, as a result, cannot grow in the absence of proline in the culture medium. However, upon plating the knockout strain onto minimal medium, we found that a small number of colonies were visually apparent after only a few days. Specifically, we prepared several hundred plates, which implies that we screened for survival on the order of 10 million cells (see Materials and Methods for details). We examined the plates for 4 days. Colonies appeared during days 2, 3, and 4 (supplementary fig. S1, Supplementary Material online) and most of the colonies, up to a total of 30, showed robust growth and gave rise to many new colonies upon replating onto minimal medium (supplementary fig. S2, Supplementary Material online).

Early studies (Berg and Rossi 1974) reported that some proline auxotrophs blocked at the first or second steps in the proline biosynthetic pathway could revert to prototrophy as a result of mutations in the arginine biosynthetic pathway (see supplementary fig. S3, Supplementary Material online). Briefly, impaired activity of the enzyme encoded by the *argD* gene leads to accumulation of the substrate of this enzyme, *N*-acetylglutamate γ-semialdehyde, which, with the assistance of a promiscuous activity of the enzyme encoded by the *argE* gene, is

transformed into γ-glutamyl semialdehyde, thus bypassing the block in proline biosynthesis.

Surprisingly, the sequencing of our rescued *E. coli* cells did not detect any mutations in the enzymes involved in arginine biosynthesis, but showed that most of our rescues had single mutations in the gene for glutamine synthetase, *glnA*. We performed Illumina whole-genome sequencing on the seven rescues that appeared first (i.e., on the second day after plating), and it showed that there were no additional mutations outside *glnA*. (Two rescue strains did not yield any mutations, suggesting that, in these cases, rescue is linked to alterations that cannot be easily detected by random fragment sequencing, such as gene duplications or transposon insertions, for instance.) Furthermore, it appears that a single mutation was sufficient for rescue via glutamate synthetase; 19 (out of 21) of the rescues that appeared on the third and fourth days after plating had just single mutations in *glnA*, as shown by Sanger sequencing. In total, we identified unique 15 mutations in the glutamine synthetase gene of the *E. coli* cells that reversed to proline prototrophy. Some of those mutations occurred independently up to three times. In order to test whether the mutations are sufficient for rescue, we complemented the *ΔproB* knockout cells with a plasmid expressing the mutated variants of glutamine synthetase, as well as the wild-type enzyme as control. We observed efficient growth in the absence of proline for 12 of the identified mutations, while no growth was observed upon complementation with the wild-type enzyme (fig. 4*A* and supplementary fig. S4, Supplementary Material online). Note that we did not include isopropyl β-D-thiogalactopyranoside (IPTG) in these complementation experiments and, therefore, that the observed rescues of proline auxotrophy were linked to the amounts of glutamine synthetase variants produced by the background transcription levels of the plasmid used. RT-qPCR experiments (supplementary fig. S5, Supplementary Material online) indicate that these background transcription levels are actually several-fold above the natural endogenous levels of mRNA for the wild-type glutamine synthetase.

As a further test that the rescue of proline auxotrophy is linked to glutamine synthetase, we performed cell growth experiments in minimal medium supplemented with 2 mM glutamine. In a typical example of negative metabolic feedback, glutamine inhibits the transcription of the *glnA* gene that encodes glutamine synthetase (Stadtman 2004). Therefore, the presence of glutamine in the culture medium will decrease the levels of glutamine synthetase and impair any rescue mechanism linked to this enzyme. Indeed, we ascertained, using several of the original rescues, that the presence of glutamine in the culture medium prevents rescue (fig. 4*B* and supplementary fig. S6, Supplementary Material online). As a control, the addition of glutamine has no effect on the growth of these cells if the culture medium is complemented with proline.

Finally, it is worth noting that we did not use any mutagenic protocol to introduce mutations in the *E. coli* genome. Therefore, the rescuing mutations in the gene of glutamine synthetase arise spontaneously in the cell population or, more likely, they were already present at a very low level in the population. This is not surprising, given that our plating experiments involve the screening for survival of a huge number of cells (probably >10 million, since each plate screens about 10^5 cells and several hundred plates were prepared: see Materials and Methods for details). Therefore, even a few rescuing mutations present at very low levels will be revealed by these experiments.

A Leakage Mechanism to Explain the Rescue of Proline Auxotrophy by Variants of Glutamine Synthetase

Glutamine synthetase is the key enzyme in nitrogen assimilation, it is found in all cells, and the glutamine synthetase gene is one of the oldest existing genes (Kumada et al. 1993; de Carvalho Fernandes et al. 2022). It catalyzes the ATP-dependent reaction of glutamate with ammonia to yield glutamine (Eisenberg et al. 2000; Stadtman 2004), which in turn provides the nitrogen required for the biosynthesis of many metabolites. The reaction occurs via two steps (fig. 5*A*): the first step is the formation of the intermediate γ-glutamyl phosphate, which subsequently reacts with ammonia in a second step to yield glutamine with the concomitant release of phosphate (Eisenberg et al. 2000). Therefore, it seems reasonable to hypothesize that the mutations responsible for the restoration of proline prototrophy in the *ΔproB* knockouts cause the leakage of the intermediate from the active site of the synthetase to the intracellular milieu and that diffusion then allows the intermediate to become the substrate of γ-glutamyl phosphate reductase. Indeed, the rescuing mutations define a clear structural pattern, which seems consistent with the leakage hypothesis. Bacterial glutamine synthetases are dodecamers of identical subunits (fig. 5*B*) with active sites formed by residues from two monomers (Almassy et al. 1986). The rescuing mutations appear in all cases at the interaction surface between two glutamine synthetase monomers and close to the active site (fig. 5*D*) lending credence to the notion that they may enable some molecules of the intermediate to leak out of the active site before being transformed into glutamine.

On the other hand, the leakage hypothesis expounded above would seemingly face several serious problems. γ-Glutamyl phosphate is a high-energy phosphorylated compound that is extremely unstable and prone to undergo side reactions. In particular, it is known to readily cyclize when exposed to water to yield 5-oxopyrrolidine-2-carboxylate (pyroglutamic acid; Csonka and Leisinger 2007; Fichman et al. 2015). In fact, it is generally assumed that γ-glutamyl phosphate can only exist as a protected enzyme-bound intermediate (Csonka and Leisinger 2007; Fichman et al. 2015). In the case of the multistep reaction mechanism of glutamine synthetase, γ-glutamyl phosphate is an active-site-bound transient reaction intermediate. Since glutamine synthetase is a very ancient enzyme (Kumada et al. 1993; De Carvalho Fernandes et al. 2022), sequestration of γ-glutamyl phosphate as a bound intermediate in its active site likely reflects a primordial adaptation which would seem a priori difficult to

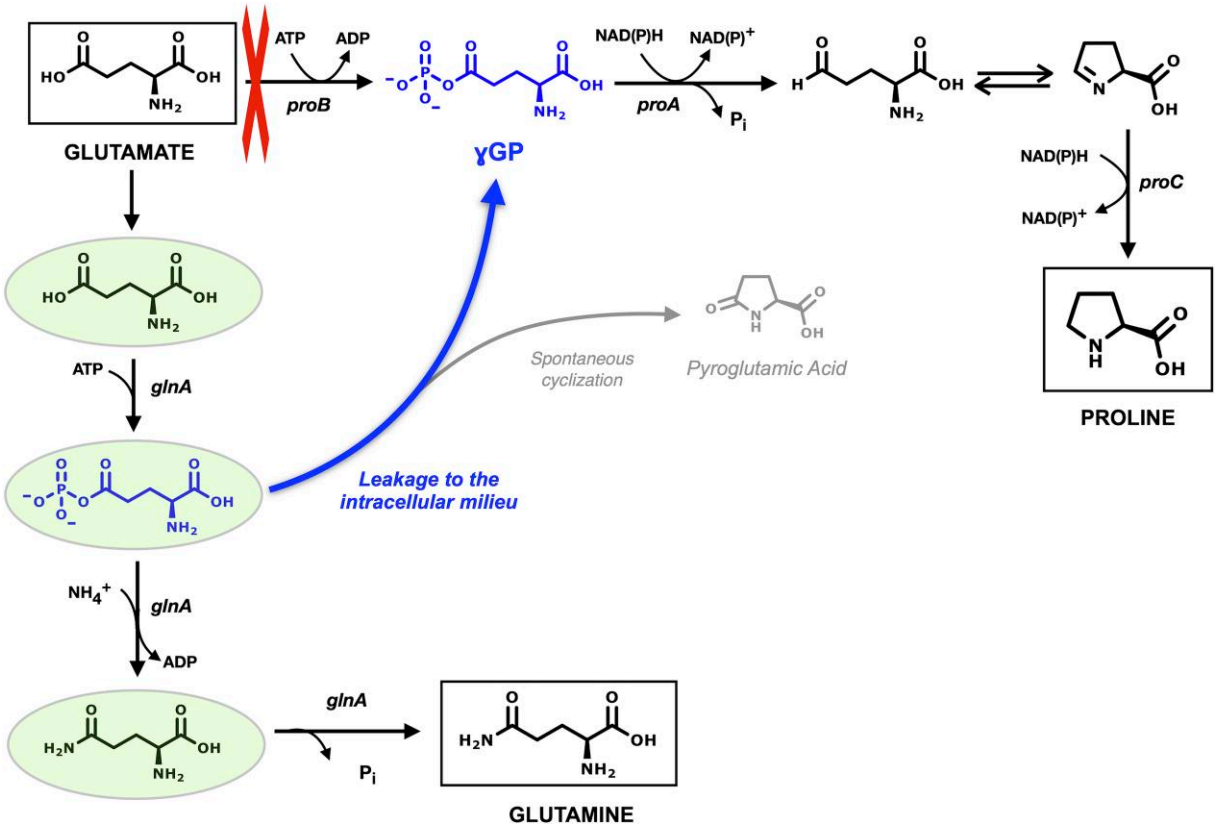

Fig. 3. Biosynthetic pathways of proline and glutamine showing the mechanism of rescue of proline auxotrophy. In most organisms, synthesis of proline from glutamate involves four steps. Three of these steps require catalysis by enzymes, which, in *Escherichia coli*, are encoded by the *proB*, *proA* and *proC* genes. See text for further details. The cross denotes the deletion of the *proB* gene, encoding γ-glutamyl kinase. This block is rescued by mutations in the gene encoding glutamine synthetase that cause the leakage of γ-glutamyl phosphate (γGP), a high-energy phosphorylated compound. The leaked γGP can either undergo spontaneous cyclization to yield 5-oxopyrrolidine-2-carboxylic acid (pyroglutamic acid) or diffuse to become substrate of γ-glutamyl phosphate reductase, the enzyme encoded by the *proA* gene. See text for details. Note that glutamine biosynthesis from glutamate occurs at the active site of glutamine synthetase. To show this fact, glutamine synthetase is represented by a oval enclosing the molecules bound to its active site. Leakage of γ-glutamyl phosphate from the active site of the synthetase and its transfer to the active site of γ-glutamyl phosphate reductase (the product of the *proA* gene) is highlighted in blue.

impair. Furthermore, mutational changes in the gene of glutamine synthetase could be expected to be severely constrained by the need to preserve the essential role of this enzyme in nitrogen assimilation.

However, the experiments and calculations we describe below support that, despite its extreme lability, γ-glutamyl phosphate does indeed leak from the active site of glutamine synthetase variants to the intracellular milieu and diffuses to reach γ-glutamyl phosphate reductase, thus bypassing the block in protein biosynthesis generated by the deletion of γ-glutamyl kinase.

In Vitro and In Vivo Detection of the Mutation-Induced Leakage of γ-Glutamyl Phosphate from Glutamine Synthetase

Catalytic features of glutamine synthetase are subject to complex modulation related, among other factors, to adenylation-deadenylation and the existence of taut and relaxed conformations (Stadtman 2004). Here, however, we are only concerned with the possibility of leakage of γ-glutamyl phosphate from the multi-step reaction mechanism of the enzyme, and this can be easily demonstrated through the detection of 5-oxopyrrolidine-2-carboxylate (pyroglutamic acid), the product resulting from γ-glutamyl phosphate cyclization when released from the synthetase active site to water (fig. 6A). We have thus prepared the wild-type enzyme and its single-mutant variants following the same procedure, and we have studied in vitro the catalytic reaction under the same solvent conditions. Both NMR and mass spectrometry detected that all of the mutations responsible for the restoration of proline prototrophy cause an increase in pyroglutamic acid concentration consistent with leakage of the γ-glutamyl phosphate intermediate (fig. 6B).

As was to be expected, there is a substantial trade-off between the leakage of the intermediate and the normal activity of glutamine synthetase (compare the two plots in fig. 6B). That is, as the leakage of the intermediate increases, the amount of glutamine produced by the enzyme decreases. Still, the production of glutamine is not fully impaired, and the mutations that rescue proline biosynthesis should not completely block glutamine biosynthesis.

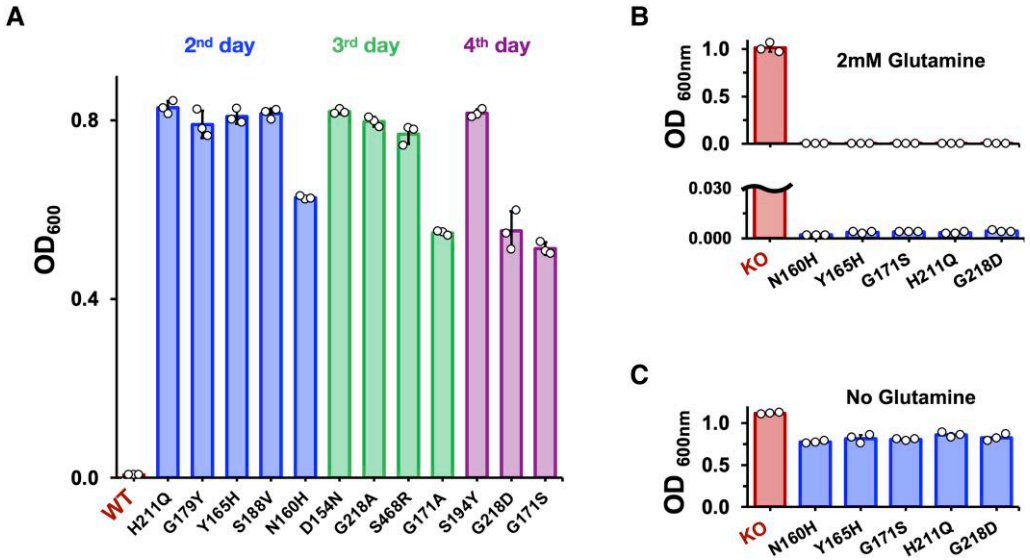

FIG. 4. Linking rescue of proline auxotrophy to variants of glutamine synthetase. Cells were allowed to grow in minimal medium for 14 h at 37° C before determining turbidity (optical density at 600 nm) to assess growth. In all cases, average values and standard deviation from the three independent determinations are given. The original values for the three biological replicates are also shown with open circles. (*A*) *ΔproB* knockout cells complemented with plasmid expressing the mutated variants of glutamine synthetase, as well as the wild-type enzyme as control, in minimal medium. Auxotrophy is rescued by variants of glutamine synthetase (labeled according to the day of appearance of the original rescues) but not by the wild-type enzyme (labeled WT). (*B*) Same as in (*A*), but using some of the original rescues of proline auxotrophy and adding 2 mM glutamine to the minimal medium (see also supplementary fig. S6, Supplementary Material online). KO refers to a control experiment with the parent *ΔproB* knockout cells using the same protocol, except that the culture medium was supplemented with proline to enable growth. Due to negative feedback inhibition of glutamine synthetase biosynthesis, rescue of auxotrophy is inhibited by the addition of glutamine to the culture medium. This is more clearly seen in blow-up plot shown below. (*C*) Same as in (*B*), but without adding glutamine to the minimal medium. This is simply a control for the experiments shown in (*B*). Growth occurs in the absence of both proline and glutamine.

Mass-spectrometry detection of pyroglutamic acid can also be used to explore the leakage of γ-glutamyl phosphate in vivo. Several processes can generate pyroglutamic acid in living cells (Kumar and Bachhawat 2012). In order to probe specifically its formation from glutamate activation, we let several of the original rescues grow in minimal medium to an optical density of 0.7, added isotopically labeled glutamate to a 1-mM concentration and, after 10 min, used mass spectrometry to detect the corresponding isotopically labeled pyroglutamic acid. As a control experiment, we used the same protocol with the parent *ΔproB* knockout cells, supplementing the culture medium with 50 μM proline to enable growth. We found significantly higher levels of pyroglutamic acid in the rescued cells, when compared with the control (fig. 6C). Glutamate can yield pyroglutamic acid through a nonenzymatic process (Kumar and Bachhawat 2012) but this is expected to be slow and occur at the same rate in the rescues and the controls. Formation of pyroglutamic acid from the γ-glutamyl phosphate produced by γ-glutamyl kinase is not possible in these cells, since the encoding gene, *proB*, has been deleted. We conclude, therefore, that the enhanced levels of pyroglutamic acid in the cells that rescue proline auxotrophy (fig. 6C) reflect the leakage of γ-glutamyl phosphate from the glutamine synthetase variants.

It is worth noting at this point that Blank et al. (2014) reported the rescue of auxotropy in a *ΔproB E. coli* strain by mutations in the *glnL* gene, the product of which regulates the activation of transcription of the gene encoding glutamine synthetase, *glnA*. The specific molecular mechanisms behind this rescue were not discussed by Blank et al. (2014). One intriguing possibility is, however, that level of intermediate leakage for the wild-type glutamine synthetase, while very low, it is greater than zero and that an increase in the expression levels of this enzyme may enable proline auxotrophy rescue. To explore this possibility, we performed growth experiments in minimal medium with *ΔproB* knockout cells complemented with a plasmid harboring the wild-type glutamine synthetase gene under overexpression induced by 0.4 mM IPTG. We did not observe growth in any of the three independent replicates. Clearly, in our experiments, levels of leakage sufficient to lead to prototrophy restoration are not produced by a highly expressed wild-type glutamine synthetase.

Transfer of the Leaked γ-Glutamyl Phosphate from Glutamine Synthetase Variants to γ-Glutamyl Phosphate Reductase

For the leakage to be effective in restoring proline prototrophy, a significant fraction of the γ-glutamyl phosphate leaked from the glutamine synthetase active site must be able to reach the γ-glutamyl phosphate reductase before undergoing side chemical alterations, cyclization in particular, that will render the molecule useless as a substrate of the reductase. Transfer of γ-glutamyl phosphate

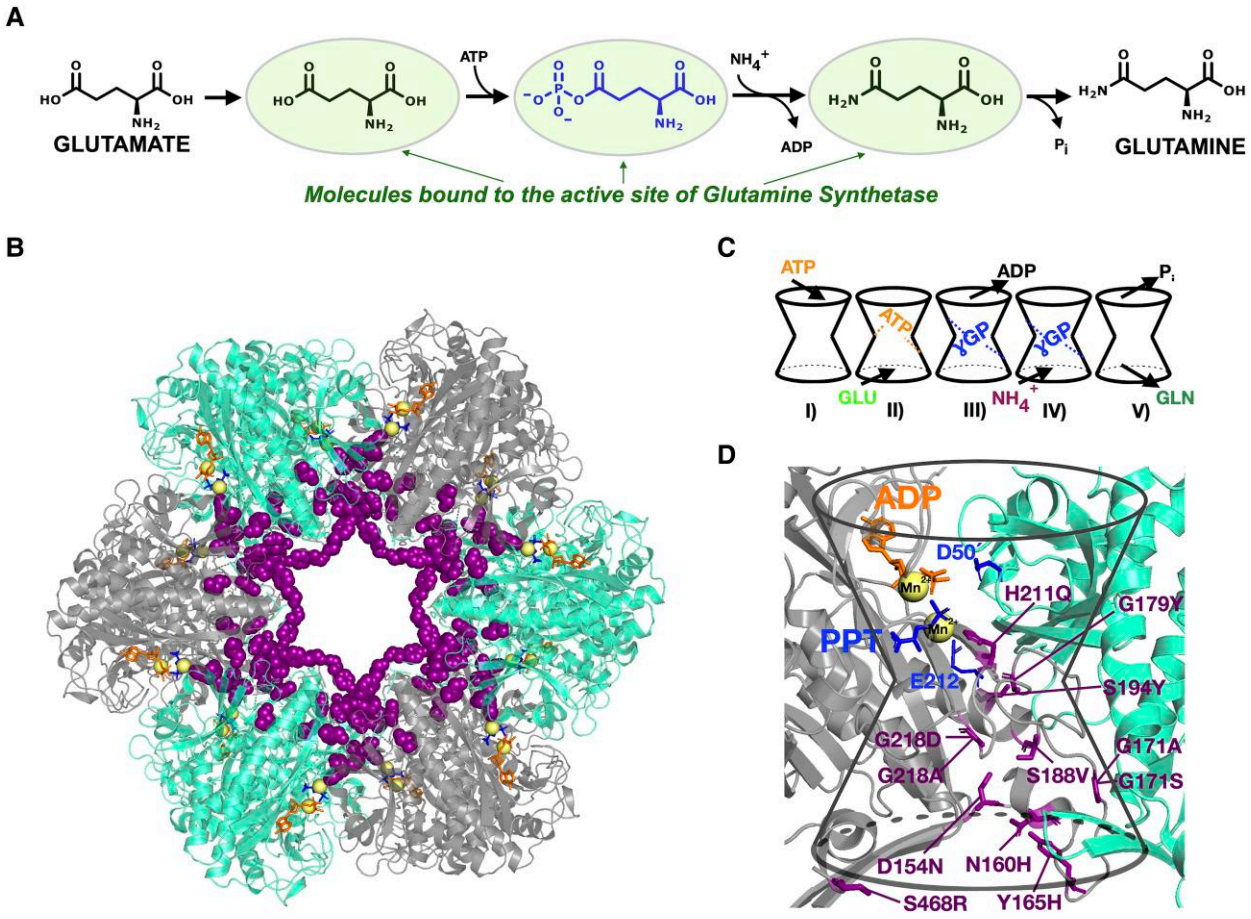

FIG. 5. Location of the mutations that rescue proline biosynthesis in the structure of glutamine synthetase. (*A*) Glutamine biosynthetic pathway. Note that glutamine biosynthesis from glutamate occurs at the active site of glutamine synthetase (the product of the *glnA* gene). To show this fact, glutamine synthetase is represented by a oval enclosing the molecules bound to its active site. The molecule of γ-glutamyl phosphate is highlighted in blue. (*B*) Bacterial glutamine synthetases form dodecamers of identical subunits with two face-to-face rings (in the top-view shown, only one ring is clearly visible) with the active sites located at the interfaces between the subunits at each hexameric ring (Eisenberg et al. 2000). The mutations that rescue *ΔproB* (original residues shown with purple spheres) accumulate at the interfaces between subunits. For clarity, neighboring subunits are shown in different color. (*C*) Cartoon representation of the mechanism of glutamine biosynthesis from glutamate highlighting the several chemical species involved. Note that the active site of glutamine synthetase is often described in the literature as a "bifunnel." (*D*) The interface between two subunits showing the active site region. Bound manganese cations (yellow spheres), ADP (orange), and the inhibitor phosphinothricin (PPT, blue), a substrate analog, are shown. The carboxylic acid residues D50' and E212 are highlighted in blue to indicate the approximate location of the ammonia-binding site. Purple labels with original residues also in purple identify the 12 positions for which rescuing mutations were found. The structure shown in (*B*) and (*D*) corresponds to the glutamine synthetase from *Salmonella typhimurium* (PDB ID: 1FPY), which has 97.9% sequence identity with the *Escherichia coli* enzyme. The residues at the positions labeled are identical in both enzymes.

between the two enzymes can be explored in vitro through a modification of the standard activity assay for γ-glutamyl phosphate reductase (Smith et al. 1984). Since γ-glutamyl phosphate is a highly unstable molecule, it is supplied in the standard assay by the previous enzyme of the proline biosynthetic pathway, γ-glutamyl kinase (the product of the *proB* gene: see fig. 3). We have modified this coupled assay by replacing the kinase with one of our glutamine synthetase variants (fig. 7*A*). In this way, activity is only observed if a sufficient amount of the γ-glutamyl phosphate leaked from the synthetase reaches the reductase before undergoing cyclization. Indeed, reductase activity is observed in these coupled assays (fig. 7*B* and supplementary fig. S7, Supplementary Material online) for all 12 variants of glutamine synthetase that we found to efficiently restore proline prototrophy. In contrast, no reductase activity is detected when using wild-type glutamine synthetase. These coupled assays demonstrated that substantial levels of γ-glutamyl phosphate can diffuse from the synthetase variants to the reductase in vitro. The same scenario must apply in vivo for rescue of proline auxotrophy to occur. This is consistent with the fact that the reductase activity values determined in the coupled assay with the glutamine synthetase variants match qualitatively the general trend observed for the efficiency of the different mutations to restore proline prototrophy (supplementary figs. S7C and *D*, Supplementary Material online).

To summarize, survival of *ΔproB* knockouts in a medium lacking proline is enabled by mutations that cause

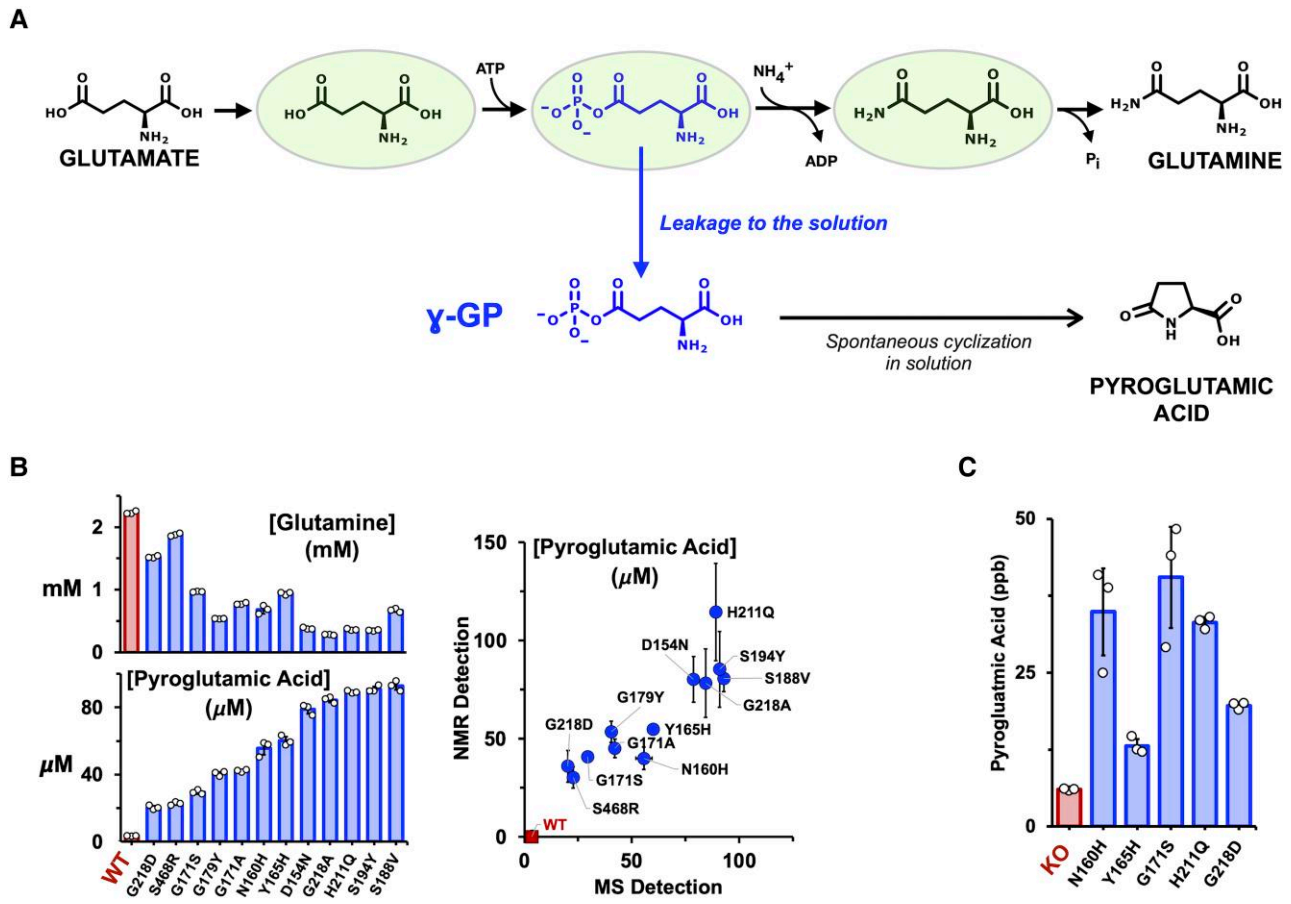

FIG. 6. Mutation-induced leakage of γ-glutamyl phosphate (γ-GP) from glutamine synthetase. (*A*) Leakage of γ-GP from the enzyme to the bulk solution can be demonstrated through the detection of the product of its spontaneous cyclization, pyroglutamic acid. (*B*) To solutions of 4 mM glutamate and 25 mM NH_4Cl at pH 7, variants of glutamine synthetase were added to a final concentration of 0.5 μM. The solutions were analyzed by mass spectroscopy to detect pyroglutamic acid and the glutamine after 12 h of reaction at 25 °C (plots on the left). The experiment was repeated with 3 mM glutamate and detection of pyroglutamic acid by NMR. The plot at the right shows an excellent congruence of the data obtained from NMR with the data obtained from mass spectrometry. (*C*) Determination of leakage in vivo. Original rescues of proline auxotrophy were grown in minimal medium, supplemented with isotopically labeled glutamate and labeled pyroglutamate was determined by mass spectrometry. KO refers to a control experiment with the *ΔproB* knockout cells in a culture medium supplemented with proline to enable growth. The concentrations of pyroglutamic acid shown correspond to the solution obtained after cells separated by centrifugation were completely resuspended in 1 ml of water/acetonitrile 40:60 (v/v). In (*B*) and (*C*), values for three biological replicates are shown, together with the corresponding average values and standard deviations.

the leakage to the intracellular milieu of a sequestered phosphorylated intermediate common to the biosynthetic pathways of proline and glutamine, thus generating a new route to synthesize proline. This mechanism of prototrophy restoration is depicted in figure 2 and, with more detail, in figure 3.

In-cell Diffusion of γ-Glutamyl Phosphate Competes Favorably With its Chemical Alteration

Rescue of proline auxotrophy by the mechanism shown in figures 2 and 3 requires that the timescale associated with the cyclization of γ-glutamyl phosphate is not substantially smaller than the timescale associated with its diffusion across a *E. coli* cell. In other words, transfer of γ-glutamyl phosphate from the synthetase to the reductase to efficiently occur in vivo, cyclization must not be much faster than diffusion over the typical length of an *E. coli* cell, because otherwise most of the leaked γ-glutamyl phosphate molecules would be rendered useless through cyclization before reaching the reductase. The simple calculations we provide below support that this condition is met and that diffusion may actually be several-fold faster than cyclization.

The timescale associated with the diffusion of a molecule over a distance L can be calculated by (see chapter 3 in Phillips et al. 2012):

$$t_{\text{diffusion}} \approx \frac{L^2}{D}$$

where D is the diffusion coefficient and L, for the problem of interest here, would be the approximate linear size of an

E. coli cell (about 1 μm). The diffusion coefficient can be easily estimated using the Stokes–Einstein equation:

$$D = \frac{k_B T}{6\pi\eta R}$$

where k_B is the Boltzmann constant, η is the viscosity, and R is the radius of the molecule (represented by a sphere). There are four single carbon–carbon bonds in γ-glutamyl phosphate, and these bonds span about half the length of the molecule. Using $R = 6$ Å (i.e., about four times the length of a single carbon–carbon bond), a diffusion coefficient of $D = 4.1 \times 10^{-10}$ m^2/s is estimated for γ-glutamyl phosphate, which leads to a diffusion timescale of 2.4 ms. This calculation of D, however, uses the viscosity of pure water at 25 °C (0.89×10^{-3} Pa·s), while diffusion may be slowed down in a crowded intracellular environment. Except for large macromolecules or molecules involved in binding with slow-diffusing components, crowding within cells decreases the diffusion coefficient by only about 4-fold (see chapter 14 in Phillips et al. 2012 and Verkman 2002) a correction that would increase the calculated diffusion time to about 10 ms. Overall, the time required for a γ-glutamyl phosphate molecule to diffuse across a distance comparable to the size of an *E. coli* cell is expected to be on the order of a few milliseconds.

The timescale associated with a chemical reaction with first-order kinetics can be described by the decay time defined as $1/k$, where k is the first-order rate constant, which in turn can be estimated from the activation free-energy barrier ($\Delta G^{\ddagger}$) using the Eyring equation from transition state theory:

$$k = \frac{k_B T}{h} \exp\left(-\frac{\Delta G^{\ddagger}}{RT}\right)$$

where h is the Planck constant. The cyclization reaction of γ-glutamyl phosphate involves the intramolecular nucleophilic attack of the amine nitrogen on the carbonyl carbon. We have used density functional theory as implemented in the GAUSSIAN09 suite (see Materials and Methods and supplementary figs. S8 and S9, Supplementary Material online for details) to calculate the activation free energy barrier for this intramolecular attack process with both, two negative charges and one negative charge on the phosphate group. The computations result in $\Delta G^{\ddagger}$ values of 14.2 and 11.8 kcal/mol, respectively, which translate into rate constants of $k = 2.4 \times 10^2$/s and $k = 1.4 \times 10^4$/s. These k values, however, would describe the rate of the reaction when the amine group is fully deprotonated, while at neutral pH, most of the amine group will be in the protonated form, which is not competent for the reaction. Assuming that the pK value of the amine group in γ-glutamyl phosphate is similar to that in glutamic acid (9.7), the fraction of deprotonated amine at pH 7 is estimated to be about 2×10^{-3}. Correction for this fraction leads to estimated rate constants of about 0.5 and 28/s which yield decay times for the cyclization process of about 2 s and 36 ms. Therefore, cyclization of γ-glutamyl phosphate is not predicted to be substantially faster than its diffusion across an *E. coli* cell, which was calculated above to occur on a timescale of a few milliseconds. Actually, cyclization is predicted to be at least somewhat slower. Certainly, the correction for protonation of the amine we have applied (i.e., multiplying by the fraction of deprotonated amine) implicitly assumes that the nucleophilic attack is rate determining while protonation-deprotonation of the amine is fast and remains at equilibrium. It is conceivable that amine deprotonation could become rate-determining at neutral pH, but as with any change in rate determining step, this would only decrease the overall rate of the reaction, thus increasing the gap between the predicted timescale for cyclization and the time associated with diffusion across an *E. coli* cell.

Overall, these order-of-magnitude calculations confirm that the efficient in vivo transfer of γ-glutamyl phosphate from the synthetase to the reductase is reasonable from the physical and chemical viewpoints. Yet, they also support that cyclization is not orders of magnitude slower than diffusion, which explains why leakage of γ-glutamyl phosphate from glutamine synthetase can be detected in vivo through the experimental determination of the cyclization product, pyroglutamic acid (fig. 6*C*).

Mitigation of the Deleterious Effects Associated With the Mutations in the Glutamine Synthetase Gene that Rescue Proline Auxotrophy

We have shown that proline auxotrophy caused by the deletion of glutamate kinase is rescued by a diversity of mutations in the gene of glutamine synthetase that allow the leakage to the intracellular milieu of the labile metabolic intermediate γ-glutamyl phosphate. These mutations enable cell survival in the absence of proline in the culture medium, but they are also expected to have some deleterious effects. First of all, since the leaked γ-glutamyl phosphate does not react with ammonia to yield glutamine, the concentration of free ammonia in the cell could increase and potentially reach toxic levels. More immediately, the rescuing mutations impair the production of glutamine (fig. 6*B*), a molecule that is crucial not only because it is required for protein synthesis but also because it is a major entry point of inorganic nitrogen in metabolism (fig. 1). Despite this, rescue of proline auxotrophy is very efficient, and in particular, cell growth does not require supplementation of the culture medium with glutamine. This suggests that the impairment of glutamine production linked to the rescue of proline auxotrophy is corrected immediately by the regulation mechanisms of the glutamine biosynthesis pathway. In particular, some of these regulation mechanisms work at the level of transcription to increase the expression levels of glutamine synthetase in response to decreased levels of glutamine (Stadtman 2004). To test this hypothesis, we have used RT-qPCR to quantify the messenger RNA arising from the transcription of the glutamine synthetase gene in the original *ΔproB*

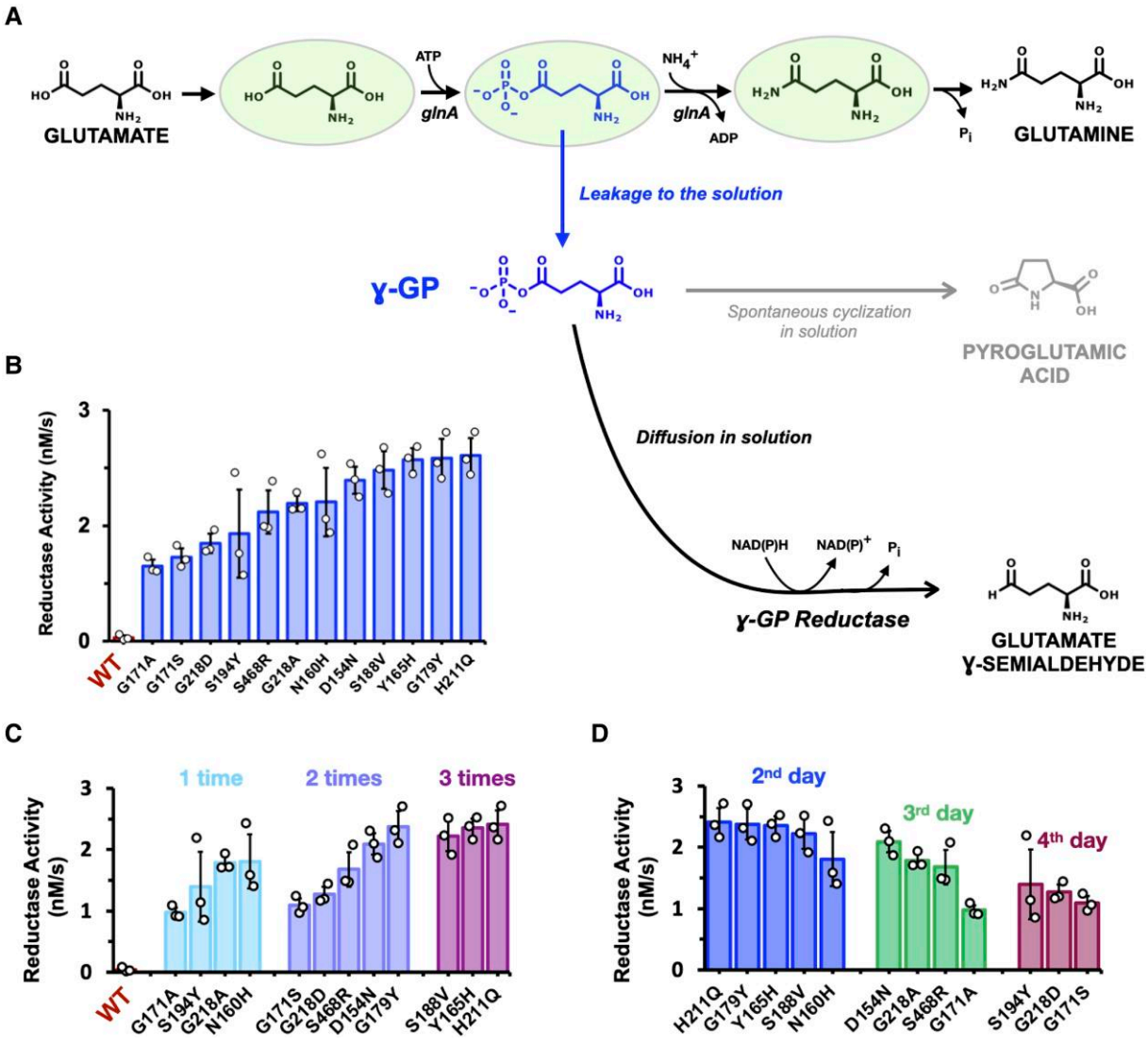

FIG. 7. Diffusion of γ-glutamyl phosphate (γ-GP) from glutamine synthetase to γ-GP reductase as detected via a coupled assay with γ-GP reductase. (*A*) The leaked γ-GP from the glutamine synthase can diffuse and be processed by the γ-GP reductase, despite diffusion competing with the spontaneous cyclization of γ-GP. This coupled in vitro assay simulates in solution the molecular process responsible for the rescue of a block in proline biosynthesis by γ-GP leakage from glutamine biosynthesis (fig. 2). (*B*) The assay of γ-GP reductase coupled with glutamine synthetase was performed as described in Materials and Methods using wild-type synthetase and the single-mutant variants responsible for the rescue of proline biosynthesis (fig. 3). All variants of glutamine synthetase lead to activity values much higher than the wild-type level, which was undetectable. The values shown are the average of at least three biological replicates (supplementary fig. S7, Supplementary Material online) and the original replicate values are also shown with open circles. Error bars are standard deviations. (*C*) and (*D*) show γ-glutamyl phosphate reductase activity in the coupled assay with glutamine synthetase variants (reported in *B*) versus efficiency of restoration of proline prototrophy, as measured by the number of times the mutation appears after plating of the knockout strain onto minimal medium (*C*) and the day of the first appearance of the mutation (*D*).

knockouts, as well as in several of the proline auxotrophy rescues that display mutations in the gene of glutamine synthetase. In agreement with our hypothesis, we found the levels of messenger RNA to be consistently several-fold higher in the rescues as compared with the original knockouts (supplementary fig. S10, Supplementary Material online).

Finally, one interesting question at this point is whether it would be possible to combine rescuing mutations with completely knockout glutamine production. That is, whether it would be possible to transform glutamine synthetase into just a glutamate kinase without compromising cell survival. The answer is: probably not. Since glutamine is essential, it would have to be included in the culture medium to allow cell growth, and the concomitant negative metabolic feedback would decrease the levels of glutamine synthetase variants and impair the restoration of proline prototrophy.

Molecular and Evolutionary Analyses of Mutation-induced Leakage of γ-glutamyl Phosphate from Glutamine Synthetase

We have shown that mutations in the gene of glutamine synthetase induce the leakage of the active-site-bound

γ-glutamyl phosphate intermediate to the intracellular milieu and, furthermore, that a substantial amount of the intermediate can reach γ-glutamyl phosphate reductase, thus rescuing proline auxotrophy caused by the deletion of the enzyme encoded by the *proB* gene, γ-glutamyl kinase. Yet, it remains to be understood why the leakage and the subsequent rescue of proline auxotrophy can be enabled by (at least) 12 single mutations. This large number of rescued single mutations may at first seem surprising, since a completely new functionality is expected to require the emergence of complex molecular machinery that is unlikely to be generated by a single mutation. The crucial point to note here, however, is that intermediate leakage does not involve the emergence of a new molecular functionality. Rather, as it is expounded below, it is the result of the impairment of a previous evolutionary adaptation and can therefore be achieved in many different ways.

Glutamine synthetase is a key enzyme for nitrogen assimilation, and it has therefore evolved for efficient glutamine synthesis. Substantial leakage of the labile γ-glutamyl phosphate intermediate would impair glutamine biosynthesis. Consequently, modern glutamine synthetases are very likely to display evolutionary adaptations that minimize such leakage. In particular, the specific amino acid residues present at the active site, their spatial arrangement, and the active-site conformational dynamics are expected to have been naturally selected to ensure strong and specific interaction with the γ-glutamyl phosphate intermediate, efficient ammonia binding, and the subsequent fast reaction of the active-site-bound intermediate with ammonia to yield glutamine. However, it is known that most modifications in any optimized complex system consisting of interacting parts and processes will lead to a less fit system (Kirschner and Gerhart 2005). In the case of glutamine synthetase, it appears reasonable that many single mutations at the active site impair the molecular mechanisms that ensure intermediate sequestration, provided, of course, that such mutations are disruptive in the structural and energetic senses and likely cause, therefore, substantial perturbations. Indeed, as we discuss below in some detail, the mutations in glutamine synthetase that we have found to rescue proline auxotrophy appear to fit this scenario.

Disruptive mutations are generally destabilizing and, given that protein stability is marginal, they tend to be rejected during evolution (i.e., purged by purifying natural selection). As a result, correlations between the mutational effects on stability and the frequency of the amino acid occurrences in sequence alignments are often observed (see Godoy-Ruiz et al. 2006, Risso et al. 2015 and references therein) with the strongly disruptive and destabilizing mutations occurring rarely. We therefore explored the frequency of occurrence in modern glutamine synthetases of the new residues introduced by the mutations that rescue proline auxotrophy. To this end, we performed a BLAST search in the UniProt database using the sequence of *E. coli* glutamine synthetase as a query. We recovered 998 sequences annotated as glutamine synthetases belonging to different species of γ-proteobacteria, the class of bacteria to which *E. coli* belongs. These 998 sequences were aligned using MEGA (Tamura et al. 2021). As shown by the results summarized in table 1, the new amino acid residues generated by the mutations we found to rescue proline auxotrophy rarely appear in the sequences of modern glutamine synthetases from γ-proteobacteria. Actually, in most of the mutated positions, the new residues never appear in the sequence alignment we analyzed.

On the other hand, the *E. coli*, wild-type residues at the considered positions are generally well conserved in γ-proteobacteria, being the consensus residues (i.e., the most frequent residues) in the alignment used, with the only exception of position 154 (table 1). It is to be noted that, for positions of low sequence diversity, the consensus residue is expected to match the ancestral residue (Risso et al. 2014). As a metric of sequence diversity, we provide

Table 1. Evolutionary and Structural Analysis of Mutations that Rescue Proline Auxotrophy.

Mutations in glutamine synthetase that rescue proline auxotrophy in *ΔproB* knockout *E. coli* cells	Statistical analyses of an alignment of 998 sequences of glutamine synthetases from γ−proteobacteria				Buried surface area of the amino acid side chain in the 3D structure of glutamine synthetase (Å²)	
	Number of amino acid occurrences at the position hosting the mutation			Shannon entropy of the position		
	E. coli WT residue	Consensus residue	New residue		*E. coli* WT residue	New residue
D154N	D: 273	S: 607	N: 0	0.94	107.2	112.7
N160H	N: 688	N: 688	H: 0	1.09	122.9	149.6
Y165H	Y: 576	Y: 576	H: 2	1.31	146.3	115.4
G171A	G: 927	G: 927	A: 71	0.26	–	57.2
G171S	G: 927	G: 927	S: 0	0.26	–	66.9
G179Y	G: 998	G: 998	Y: 0	0	–	190.6
S188V	S: 846	S: 846	V: 1	0.67	57.1	91.0
S194Y	S: 503	S: 503	Y: 0	1.21	44.0	163.4
H211Q	H: 998	H: 998	Q: 0	0	95.2	98.7
G218A	G: 918	G: 918	A: 0	0.21	–	54.4
G218D	G: 918	G: 918	D: 0	0.21	–	104.2
S468R	S: 862	S: 862	R: 0	0.50	8.9	14.7

in table 1 values of Shannon entropy calculated from the sequence alignment. In general, Shannon entropy values range between 0, when there is full conservation of a given amino acid at the considered position, and 3, when the 20 amino acids appear with equal frequency. The values of Shannon entropy given in table 1 indicate, therefore, low to very low sequence diversity at the positions hosting the rescuing mutations. Furthermore, previous studies (Risso et al. 2014) support that discrepancies between ancestral residues and consensus residues tend to occur at positions with Shannon entropy values substantially above one. Furthermore, as we have noted above, the wild-type residues in *E. coli* glutamine synthetase at the position hosting the rescuing mutations match the corresponding consensus residues, except for position 154. Hence, it appears safe to assume that the wild-type residues in *E. coli* glutamine synthetase at the positions hosting the rescuing mutations agree in most cases with the residues at the corresponding positions in the glutamine synthetase from the common ancestor of γ-proteobacteria.

The interpretation expounded in the preceding paragraph, together with the sequence statistics information collected in table 1, then supports that the mutations that rescue proline auxotrophy in *E. coli ΔproB* knockout cells have rarely occurred during the evolution of γ-proteobacteria, as was to be expected if the mutations are disruptive in the structural sense. Accessible surface area calculations (see Materials and Methods for details) are consistent with this view, as, in many cases, the buried surface area for the side chain of the new amino acid is large, that is, more than a hundred of $Å^2$. The important point to note here is that the molecular environment of a given site in a modern protein may be expected to have been optimized during evolution to establish stabilizing interactions with the ancestral amino acid residue at that site (Risso et al. 2015). Consequently, burial of a different amino acid at the site likely brings about the disruption of stabilizing interactions and the energetically costly structural reorganization of the local environment to accommodate the new residue.

Certainly, for a few of the rescuing mutations, the buried surface areas do not seem to conform to the pattern described above. In particular, for the S468R mutation, both the wild-type residue and the new residue at position 468 are fairly exposed to the solvent and show little surface burial. In this case, however, the disruptive character of the mutation is likely related to the existence of clusters of charged residues in the neighborhood of position 468, which will be severely disrupted by the introduction of the positive charge on the arginine side chain (see supplementary fig. S11, Supplementary Material online). Also, comparatively low buried surface areas are obtained for alanine and serine residues at loop positions 171 and 218, which originally bear a glycine (see data for mutations G171A, G171S, and G218A in table 1). Yet, glycine does not have a side chain and, unlike all other amino acids, does not impose restrictions on backbone dihedral angles, thus providing a flexible link that allows a diversity of loop conformations (Scott et al. 2007; Gamiz-Arco et al. 2019). Plausibly, therefore, substituting another residue for glycine may impose an alternative local conformation to the polypeptide chain, thus bringing about a substantial (and potentially disruptive) conformational change. In addition, the replacement that eliminates glycine may impair the active-site dynamics required for efficient glutamine production.

Overall, the molecular and evolutionary analyses reported in this section support that the mutations that rescue proline auxotrophy disrupt structure and possibly dynamics in the active site of glutamine synthetase, thus impairing previous evolutionary adaptations for efficient sequestration of the γ-glutamyl phosphate intermediate.

On the Possibility that Leakage of γ-Glutamyl Phosphate from Glutamine Synthetase may have been Occasionally Fixed During Natural Evolution

One intriguing possibility is that mutations in glutamine synthetase leading to intermediate leakage may have been fixed sporadically during natural evolution. One plausible scenario for this would involve genetic drift. That is, a population bottleneck may prevent a slightly deleterious, leakage-enabling mutation from being purged, and the mutation will be permanently fixed by subsequent alterations in the *proB* gene that impair γ-glutamyl kinase activity. Another possibility is that alterations that impair to some extent γ-glutamyl kinase activity are preserved through genetic drift and then fixed upon mutations that cause intermediate leakage in glutamine synthetase.

In order to explore the occurrence of these scenarios, we have first searched for the occurrence of our rescuing mutations in glutamine synthetase proteins in different *E. coli* strains. To this end, we have used the sequence of the *E. coli glnA* gene as a query for a blast search at the NCBI site of the GenBank + EMBL + DDBJ + PDB + RefSeq databases. We identified 3,659 sequences belonging to *E. coli* strains that produced statistically significant alignments with the query sequence, although many of them displayed low sequence identity with the query. We therefore focused on proteins with only a few (1–4) mutational changes with respect to the query. Although mutations at the active-site region occasionally appeared, none of them agreed with those found in our laboratory evolution experiments. Subsequently, we examined the whole genomes of all *E. coli* strains harboring the 3,659 sequences identified in the original search and found that the *proB* gene was present in all of them. Finally, we searched the set of whole genomes for organisms displaying simultaneous mutational changes in the *glnA* and *proB* genes. We found only one example, the *E. coli* strain RM-103-MS, which has a single mutation in the *glnA* gene and a single mutation in the *proB* gene. However, the latter mutation is synonymous (L16L) and cannot impair the activity of the encoded enzyme, while the former mutation involves a very conservative amino acid substitution (V207I) that can hardly affect the catalytic profile of glutamine

synthetase. Overall, we find that the current sequence databases do not offer any example of *E. coli* strains with the pattern found in our laboratory evolution experiments (i.e., a total loss of the activity associated with the *proB* gene, which is rescued by mutations in the *glnA* gene). However, we cannot rule out that examples of this pattern exist in natural organisms and that they will become known in the future, given the exponential growth of sequence databases.

Subsequently, we considered the sequences of glutamine synthetases from γ-proteobacteria in which we found mutations (with respect to the consensus/ancestral amino acid) that match the mutations that rescue proline auxotrophy in our evolutionary repair experiments. As shown in table 1, this includes 2 instances of the Y165H mutation, 1 instance of the S188V mutation, and 71 instances of the G171A mutation (note that mutations here are defined with respect to the consensus amino acid, which is likely to be the residue in the glutamine synthetase of the common ancestor of γ-proteobacteria, as discussed in the previous section). For each of these instances, we searched the NCBI genome database for the full genome of the specific organism and strain with the glutamine synthetase hosting the mutation. We identified 8 such full genomes, in which furthermore a gene was annotated as *proB*, that is as the gene encoding γ-glutamyl kinase. Interestingly, in one case with alanine at position 171 in the glutamine synthetase (i.e., with the mutation G171A with respect to the consensus/ancestral amino acid), we found significant alterations in the *proB* gene. The case was found in strain DSM 16069/KCTC 12182/SW-125 of *Kangiella koreensis*. The alterations in the *proB* gene consisted of duplication and the presence of stop codons in both gene copies. These stop codons remove about 100 amino acid residues, but do not delete the kinase domain. One possibility is that removal of the ∼100 amino acid segments impairs, but does not fully eliminate, the γ-glutamyl kinase activity and that the low activity is compensated by gene duplication and intermediate leakage from glutamine synthetase. It is also possible, of course, that the genes annotated as *proB* (seemingly on the basis of sequence identity) perform a different kinase function. In any case, these interpretations will remain intriguing speculations until experiments with the glutamine synthetase and the presumed γ-glutamyl kinase of the above-mentioned strain of *K. koreensis* are carried out. Finally, it is interesting to note that the genus *Kangiella* has been reported to display substantial genome reduction (Wang et al. 2018). Therefore, another intriguing speculation is that the doubling of glutamine synthetase as a γ-glutamyl kinase is more likely to have occurred in microorganisms with severe genome reduction.

Conclusion

The deletion of a crucial enzyme in a metabolic pathway makes the cells auxotrophic for its final product and allows evolutionary experiments on the restoration of prototrophy to be performed in the laboratory. A number of such evolutionary repair experiments have been performed in order to probe the fundamental mechanisms of metabolic adaptation (Patrick et al. 2007; Mcloughlin and Copley 2008; Kim et al. 2010; Digianantonio and Hecht 2016; Digianantonio et al. 2017; Kim et al. 2019). These studies have provided support for the proposed role of enzyme promiscuity in the emergence of new biochemical pathways (Schulenburg and Miller 2014; Copley 2015; Noda-Garcia et al. 2018; Peracchi 2018).

Here, we report evolutionary repair experiments on *E. coli* cells in which an enzyme crucial for the biosynthesis of proline has been deleted. Surprisingly, we find reversion to prototrophy through a mechanism that does not involve the recruitment of a low-level promiscuous activity but, rather, the leakage of a labile metabolic intermediate. As summarized in figures 2 and 3, a substantial amount of the γ-glutamyl phosphate leaked from the catalytic site of glutamine synthetase diffuses to reach γ-glutamyl phosphate reductase, thus generating an alternative pathway from glutamate to proline. Rescue of a blocked metabolic pathway through the leakage of a labile intermediate from another pathway is, to the best of our knowledge, unprecedented in the literature. Therefore, we sought to provide convincing and extensive evidence supporting the leakage mechanism. That rescue of proline auxotrophy is linked to glutamine synthetase was established through complementation experiments of *ΔproB* knockout cells with a plasmid expressing variants of glutamine synthetase (fig. 4*A*) and by the fact that depressing the biosynthesis of glutamine synthetase through negative feedback inhibition does preclude restoration of proline prototrophy (fig. 4*B*). That the rescuing variants of glutamine synthetase leak γ-glutamyl phosphate both in vitro and in vivo was demonstrated using NMR and mass spectrometry determinations (fig. 6). Finally, that the leaked γ-glutamyl phosphate can reach the second enzyme in the proline biosynthetic pathway was supported by coupled activity assays (fig. 7) and by comparison between the calculated timescales for chemical alteration of γ-glutamyl phosphate and its diffusion across an *E. coli* cell.

A particularly remarkable aspect of the experimental results reported here is the high proficiency of intermediate leakage at rescuing a block in an essential biochemical pathway. Not only reversion to proline prototrophy is realized by a large number (12 at least) of different single mutations, but also the rescued cells in a culture medium lacking proline show robust growth immediately (supplementary figs. S2 and S4, Supplementary Material online). These results are even more surprising when considering that it is glutamine synthetase that produces γ-glutamyl phosphate through leakage from its active site. Glutamine synthetase is the key enzyme for nitrogen assimilation (see fig. 1 in Stadtman et al. 1968; Eisenberg et al. 2000; Stadtman 2004); it is found in all cells, and the glutamine synthetase gene is one of the oldest existing genes (Kumada et al. 1993; De Carvalho Fernandes et al. 2022). Since glutamine synthetase is a very ancient

enzyme, sequestration of γ-glutamyl phosphate as a bound intermediate in its multi-step reaction mechanism likely reflects a primordial adaptation, which would seem a priori difficult to reverse. Furthermore, mutational changes in the gene of glutamine synthetase could be expected to be severely constrained by the need to preserve the essential role of this enzyme in nitrogen assimilation.

Yet, the surprising rescuing proficiency summarized in the preceding paragraph can be convincingly rationalized. At a fundamental level, leakage does not involve the generation of new molecular functionality, but the impairment of a previous evolutionary adaptation, the sequestering of an intermediate, and damaging a highly optimized feature can be achieved in many different ways. In general, for instance, mutations that weaken the interactions of the bound intermediate with the protein moiety or mutations that decrease the rate of the enzymatic transformation of the intermediate can enhance the likelihood that the intermediate is released to the intracellular milieu. Certainly, mutations that bring about leakage of γ-glutamyl phosphate occur at the active site (fig. 5) and, consequently, impair the normal activity of glutamine synthetase (fig. 6*B*) potentially compromising nitrogen assimilation. However, this would be immediately corrected to some substantial extent by the regulation mechanisms of the glutamine biosynthesis pathway, as we have experimentally shown on the basis of RT-qPCR determinations of the transcription levels (supplementary fig. S10, Supplementary Material online). Furthermore, since some of these regulation mechanisms work at the level of transcription (Stadtman 2004) they may increase the expression levels of glutamine synthetase and consequently the flow through the new pathway generated by leakage.

Sequestration of intermediate chemical species in metabolism is not restricted, of course, to the specific case of the multi-step reaction mechanism of glutamine synthetase but occurs extensively and in a variety of molecular scenarios, as we have expounded in some detail in the introduction section (see subsection "Restriction of free diffusion [i.e., 'sequestration'] of metabolic intermediates"). Extensive sequestration is required for metabolic efficiency, and widespread leakage of metabolic intermediates to the intracellular milieu will likely result in cell death. On the other hand, our results support that leakage of one given intermediate may readily occur without fatal consequences for the organism and that it may even contribute to organismal adaptation in some scenarios. In his classic review on complexes of sequential metabolic enzymes (Srere 1987), Paul Srere considered 520 molecules in the intermediary metabolism of carbohydrate, lipid, amino acid, and nitrogen base and found that 80% of them have only one fate in the cell because they are nutrients, end product, or intermediates produced by only one enzyme and transformed by only one enzyme. The remaining 20% of metabolites have several uses in the cell. Therefore, rescue of a block in an essential route by leakage of a sequestered intermediate from a different route, as demonstrated here with the proline and glutamine pathways, could plausibly occur in many other instances.

Another interesting possibility is that the leakage of metabolic intermediates contributes to the generation of new biochemical pathways. The acquisition of new chemical reactions is essential for metabolic innovation. Models of metabolic pathway emergence and evolution often emphasize the acquisition of new reactions through horizontal gene transfer and promiscuous enzyme functionalities (Pál et al. 2005; Schulenburg and Miller 2014; Copley 2015; Noda-Garcia et al. 2018; Peracchi 2018). However, leakage of metabolic intermediates may provide an additional source of new reactions. Metabolic intermediates are often intrinsically reactive species that not only may become alternative substrates for promiscuous enzymes (Noda-Garcia et al. 2018) but that may also undergo various nonenzymatic chemical transformations (Lerma-Ortiz et al. 2016). As a result, leakage of one intermediate upon mutation could facilitate its transformation, enzymatic or nonenzymatic, into different chemical species, thus providing an additional source of new chemical reactions. Certainly, some of the molecules generated in these new reactions will be toxic and will need to be neutralized by the metabolite damage-control system of the cell (de Crécy-Lagard et al. 2018). Yet, some others could trigger the generation of new biochemical pathways that allow the organism to meet specific challenges.

Additional work will be required to assess the extent to which the leakage mechanism has contributed to the diversity of the products of secondary metabolism or, more generally, to metabolic pathway evolution by generating new metabolites. On the other hand, it appears that the leakage mechanism can be immediately used as a tool in synthetic biology for metabolic engineering. The design of pathways for the production of valuable products (Lee et al. 2018; Cravens et al. 2019) will benefit from the use of highly reactive molecules as precursors. Certainly, highly reactive molecules are most likely to be sequestered in vivo, but this work demonstrated that their availability can be engineered, as leakage-inducing mutations can easily be found. The overall process would likely involve the following steps: (1) a protein crucial for the sequestration of a given labile intermediate is identified; (2) the protein is used as background for the generation of a suitable variant library, likely focusing the mutational variability to the regions that are expected to be involved in interactions with the intermediate; (3) the library is screened in vitro (i.e., standard directed evolution) for leakage of the intermediate with preservation of sufficient levels of the normal enzyme activity, along the general lines of the experiments we show in supplementary figure S6*B*, Supplementary Material online; (4) standard genome editing is used to generate a modified organism that bears suitable leakage-inducing mutations determined in the previous step; and (v) the organism is further modified to express an enzyme that processes the leaked intermediate and the additional enzymes required to complete the engineering of the new pathway.

Materials and Methods

Selection and Identification of Spontaneous Proline Auxotrophy Rescues

Escherichia coli BW25113 ΔproB760::kan (JW0232-1) was obtained from the Keio collection (Baba et al. 2006). Aliquots of frozen stocks of the *ΔproB* knockout strain were streaked onto 240 plates containing M9–glucose medium (agar [2% w/v], M9 salts [1x], glucose [0.4%, w/v], $MgSO_4$ [2 mM], $CaCl_2$ [100 μM]) supplemented with Kanamycin (Kan; 30 μg/mL). Plates were incubated at 37 °C for up to 4 days, and colony formation was monitored. The rescuing activity of each individual clone was verified by restreaking onto new M9-glucose-Kan plates at least three times. Genomic DNA was extracted from isolated colonies for each of the rescue clones with the bacterial genomic DNA isolation kit from Norgen (CAT#17900) and quantified using a Biodrop μLite+.

In order to estimate the number of cells screened for survival in our experiments, we first estimated the number of viable cells obtained upon inserting the streaking loop in the stock of *ΔproB* knockout cells. To this end, a loop that had been previously inserted in the stock was stirred into 500 μl of LB, 10 μl of the solution was immediately plated in LB agar (so that all viable cells could grow despite proline auxotrophy), the number of colony forming units was determined using ImageJ (Schneider et al. 2012) after overnight incubation, and the number of viable cells obtained by loop insertion was calculated by correcting with the appropriate dilution factor. Three independent replicas of this procedure were performed (i.e., starting with three different loops), which led to an average value of $(1.1 \pm 0.1) \times 10^5$ viable cells per loop insertion. In total, we prepared 240 plates, which means 240 loop insertions in the stock of *ΔproB* knockout cells and, therefore, that $240 \times (1.1 \pm 0.1) \times 10^5 = (2.6 \pm 0.5) \times 10^7$ cells were screened. Of course, this calculation is approximate, and it is only expected to provide the order of magnitude. Yet, it is safe to assume that about 10 million cells were screened for survival in our experiment.

Genomic DNA Sequencing

Of the total number of clones, DNA from nine plate samples was sequenced: two control samples (*E. coli* BW25113 ΔproB760:kan (JW0232-1)) and seven samples from rescue clones obtained as explained in the previous section.

These were sequenced on a Novaseq 6000, using an S2 cartridge, generating on average 8,392,966 × 150 nt paired reads. These sequencing results are available at the following link: https://www.ncbi.nlm.nih.gov/bioproject/839218.

For the remaining gDNA samples from rescue clones, the *glnA* gene was amplified by PCR using the PfuUltra high-fidelity DNA polymerase AD from Agilent (CAT#600385) and the following specific primers: glnA.for (5′-AGA CGC TGT AGT ACA GCT CAA ACT C) and glnA.rev (5′-ATG TCC GCT GAA CAC GTA CTG ACG ATG). Amplified products were sequenced by Sanger method using specific primers to *glnA* gene in order to identify the rescue mutations. The nucleotide sequences of all variants can be found in supplementary figure S12, Supplementary Material online.

Bioinformatics Analyses

Paired files obtained after sequencing were evaluated using FASTQC software to identify possible errors, such as low-quality sequences, accumulation of adapter sequences, or unidentified nucleotides (Andrews 2010).

After this quality control filtering, we observed that 99.1% of the reads had an average quality value higher than Q30, and there was no accumulation of adapters, neither repeated sequences nor a high percentage of Ns. After selecting, by cutadapt (Martin 2012), the sequences with an average quality value of at least Q30, Snippy was used to detect variants in the sequenced DNA (Seemann 2015). Briefly, Burrow Wheeler Alignment (Li and Durbin 2009) was used to align the reads against the reference genome (GCA_000750555 from GeneBank). Samtools (Li et al. 2009) was used to sort, mark, and remove duplicate sequences. Once aligned and sorted, depth was calculated to estimate the number of reads that cover each nucleotide, on average, we obtained 412X. Subsequently, variant calling was performed using the default parameters by Freebayes (Garrison and Marth 2012). Of note is that in all samples in which the whole genome was sequenced, including the two controls, a mutation in the *cpxA* gene (342C > A nucleotide change coding for the Asn114Lys amino acid change) appeared when compared with the published reference sequence. Since, however, this mutation also appeared in the controls (the nonrescued *ΔproB* knockouts), it is clear that it is not related to reversion to prototrophy.

Protein Mutagenesis and Expression

The coding sequence for *E.coli* glutamine synthetase (GS) was inserted into the Kpn1/Xho1-digested pET-45b(+) expression vector (GenScript Biotech, The Netherlands), containing a hexahistidine tag at its N-terminus and an ampicillin resistance gene on the plasmid. Single-mutant variants of glutamine synthetase were prepared by site-directed mutagenesis using the QuikChange Lightning kit (Agilent Technologies, USA).

The gene coding for γ-glutamyl phosphate reductase was ligated into the pHTP1 expression vector (GenScript Biotech) producing a recombinant protein that contains an N-terminal hexahistidine and Kanamycin resistance.

All the recombinant proteins were transformed and expressed in *E.coli* BL21 (DE3) (Invitrogen, USA).

Growth of Keio *ΔproB* Knockouts Complemented with Glutamine Synthetase Variants

Complementation assays were performed to ensure that the glutamine synthetase variants rescued the Keio *ΔproB* strain. For this purpose, the knockout strain was transformed by electroporation with each of the Kpn1/Xho1-digested pET-45b(+) expression vectors (GenScript

Biotech), containing the single glutamine synthetase variants, and selected by the addition of 40 μg/ml kanamycin and 100 μg/ml ampicillin to the minimal medium. All cultures were grown in minimal medium until saturation (about 36 h at 37 °C and shaking). They were then diluted to $OD_{600} = 0.1$ and the optical density (OD_{600}) of all cultures was measured after 14 h (fig. 4 and supplementary fig. S4, Supplementary Material online).

Protein Purification

Cells harboring the plasmids containing the γ-glutamyl phosphate reductase gene or glutamine synthetase genes were grown at 37 °C to an OD_{600} of ∼0.6 in Luria–Bertani medium supplemented with 40 μg/ml kanamycin or 100 μg/ml ampicillin, respectively. Overexpression of all the recombinant proteins were induced with 0.4 mM IPTG and cell growth was continued at 25 °C overnight. The cells were harvested by centrifugation at 6,198 × g for 15 min at 4 °C and the pellets were immediately stored at −80 °C after their resuspension in cold lysis buffer (20 mM $NaPO_4$ pH 7.4/0.5 M NaCl/20 mM Imidazole) supplemented with protease inhibitors. Cells were thawed and disrupted by sonication on ice. The insoluble fractions were removed by centrifugation (27,000 × g, 30 min at 4 °C). Hexahistidine-tagged proteins in the soluble fraction were affinity purified using a nickel-charged His-trap immobilized metal affinity chromatography (IMAC) column (GE Healthcare, UK). Washing was carried out with lysis buffer, and recombinant proteins were eluted with the same buffer, which contained 500 mM imidazole. The eluted glutamine synthetase proteins (wild-type or single-mutant variants) were then dialyzed against buffer 50 mM Tris–HCl, pH 7.0/2 mM $MgCl_2$, while γ-glutamyl phosphate reductase was dialyzed against buffer 50 mM Tris–HCl, pH 7.0/1 mM DTT. Purity was confirmed by SDS-PAGE (see supplementary fig. S13, Supplementary Material online) and concentrations were determined spectrophotometrically using molar extinction coefficients at 280 nm of 36,580 M^{-1} cm^{-1} for glutamine synthetase variants and 17,795 M^{-1} cm^{-1} for γ-glutamyl phosphate reductase. All proteins were stored at −80 °C and at concentrations above 100 μM in the dialysis buffer. No deterioration of the activities was observed upon storing the enzymes at −80 °C, at least over a period of a few months. Activity measurements were thus performed after storage at −80° C, using the coupled assay and metabolite detection by both NMR and mass spectrometry, and the results were similar to those obtained with the proteins prior to freezing. In addition, their possible degradation after freezing was evaluated using SDS-PAGE polyacrylamide gels without observing significant degradation for any of the variants studied.

Experimental In Vitro Determination of Leakage of γ-Glutamyl Phosphate From Glutamine Synthetase Variants

Since γ-glutamyl phosphate is a highly labile species, leakage was determined through the detection of the cyclization product, pyroglutamic acid. Samples were prepared in 50 mM Tris–HCl, 15 mM ATP, 4 or 3 mM L-glutamate, 25 mM NH_4Cl, 50 mM $MgCl_2$ at pH 7.0. The reaction of glutamine synthetase was started with the addition of 300 μg of the enzyme (either the wild type or a single-mutant variant) to a final volume of 1 ml. The solutions were kept at 25 °C for 12 h and their chemical composition was assessed by mass spectrometry or nuclear magnetic resonance, as described below.

Mass Spectrometry

Samples for mass spectrometry were prepared by dilution to 1:2,000 in water/acetonitrile 40/60 (v/v) and 5.0 μL aliquots were injected in an ultraperformance liquid chromatography (UPLC) system consisting of a Quaternary Solvent Manager Acquity (Waters Corporation) equipped with a column Cortecs UPLC Hillic (1.6 μM, 2.1 × 50 mm) and coupled with a triple quadrupole mass spectrometer XEVO-TQS (Waters Corporation). The mobile phase of the UPLC system consisted of H_2O + 0.1% formic acid v/v (solvent A) and acetonitrile + 0.1% formic acid v/v (solvent B). Linear gradient elution (flow rate of 0.2 ml/min) was programmed as follows: 0–4 min (20% A–80%B), 4–5 min (40% A–60%B), 5–5.1 min (20% A–80%B). Electrospray ionization mass spectra (ESI-MS) were acquired in the positive (ESI +). LC-MS/MS acquisition parameters for the target molecules were as follows: L-glutamic acid precursor ions 148.0, product ions 84.0 and 101.9; L-glutamine precursor ion 147.0, product ions 84.0 and 130.0; L-pyroglutamic acid precursor ion 130.0, product ions 56.0 and 84.0. For each glutamine synthetase variant, three independent determinations were performed. The raw data from all mass experiments performed can be found in the additional supplementary document "RawDataMassExperiments."

Nuclear Magnetic Resonance Spectroscopy

^1H NMR spectra were acquired on a Bruker Avance III 600 Hz spectrometer at a proton frequency of 600 MHz equipped with a TCIP cryoprobe at 298 K. Samples were supplemented with 5% D_2O and 0.01 mM DSS (4,4-dimethyl-4-silapentane-1-sulfonic acid). The standard zgesgp pulse sequence from the Bruker library was used with excitation sculpting to allow suppression of the water (Hwang and Shaka 1995). Data were collected with 16,384 points, a spectral width of 16.0242 ppm with 256 scans, 16 dummy scans, and an acquisition time of 1.7 s. Data were processed using MestreNova software. All spectra were automatically phased, baseline corrected using a polynomial function, and calibrated to the center of the DSS peak. The concentrations of glutamate, glutamine, and pyroglutamate were measured by integrating the ^1H peaks corresponding to gamma protons at ^1H frequencies of 2.33, 2.43, and 2.41 ppm, respectively. For each glutamine synthetase variant, three independent determinations were performed. The raw data from all NMR experiments

performed can be found in the additional supplementary document "RawDataNMR."

Assay of γ-Glutamyl Phosphate Reductase Activity Coupled With Glutamine Synthetase

The coupled activity assay we used to assess the diffusion of γ-glutamyl phosphate from glutamine synthetase to γ-glutamyl phosphate reductase is identical to that described by Smith et al. (1984), except for the source of γ-glutamyl phosphate. Specifically, we used variants of glutamine synthetase as a source of γ-glutamyl phosphate, instead of γ-glutamyl kinase. The measurement of the reductase activity was based on the spectrophotometric determination of the oxidation of NADPH. The reaction solution was 50 mM Tris–HCl pH 7.0, 15 mM ATP, 8 mM L-glutamate, 25 mM NH_4Cl, 50 mM $MgCl_2$, 0.15 mM NADPH, and 7 μM γ-glutamyl phosphate reductase. The solution was equilibrated at 25 °C for 5 min, the reaction was initiated by the addition of a glutamine synthetase variant to a concentration of 7 μM. The enzyme activity was determined using the initial reaction rates determined from changes in A_{340} nm and the molar extinction coefficient of NADPH at that wavelength (6221 M^{-1} cm^{-1}). Three independent coupled activity assays of γ-glutamyl phosphate reductase were performed with each glutamine synthetase variant as source of γ-glutamyl phosphate.

Computational Calculation of the Barrier for γ-Glutamyl Phosphate Cyclization

Calculations were performed by means of density functional theory (DFT) using the GAUSSIAN09 suite (Frisch et al. 2010). Structures were optimized using the widely used B3LYP functional (Becke 1988, 1993; Lee et al. 1988) with the Pople's-based 6–31G* basis set (Francl et al. 1982). All minima were fully optimized by the gradient technique and assessed by frequency calculations. Enthalpy and Gibbs free energy values were obtained by taking into account zero-point energies, thermal motion, and entropy contribution at standard conditions (temperature of 298.15 K, pressure of 1 atm). A Polarizable Continuum Model (PCM) (Miertus et al. 1981; Tomasi and Persico 1994) has been taken into account to reproduce the solvent in which reactions and/or physical measurements have been carried out (water).

Mass Spectrometry Determination of Pyroglutamic Acid *In Vivo*

Original rescues were grown at 37 °C in kanamycin-supplemented minimal medium until they reached an optical density of 0.7. As a control, the knockout strain Keio *ΔproB* was grown in minimal medium supplemented with kanamycin and 50 μM proline. ^{13}C- and ^{15}N-labeled glutamic acid (CAS number: 202468-31-3) from Cortecnet was added to all samples to a final concentration of 1 mM. Afterwards, all samples were kept at 25 °C shaked for 10 min. Cells were collected by centrifugation at 6,198xg at 4 °C for 15 min and the supernatant was discarded. Cells were completely resuspended in 1 ml of water/acetonitrile 40:60 (v/v) and centrifuged at 18,407xg for 15 min. 5 μl of the supernatant of each sample was injected in an UPLC system consisting of a Quaternary Solvent Manager Acquity (Waters Corporation) equipped with a column Cortecs UPLC Hillic (1.6 μM, 2.1 × 50 mm) and coupled with a triple quadrupole mass spectrometer XEVO-TQS (Waters Corporation). The mobile phase of the UPLC system consisted of H_2O+ 0.1% formic acid v/v (Solvent A) and acetonitrile + 0.1% formic acid v/v (Solvent B). Gradient elution (flow rate 0.2 ml/min) was programmed as follows: 0–4 min (20% A–80%B), 4–5 min (40% A–60%B), 5–5.1 min (20% A–80%B). ESI-MS were acquired in the positive (ESI+). LC-MS/MS acquisition parameters for the target molecules were as follows: L-pyroglutamic acid precursor ion 130.0, product ion 84.0, [$^{13}C_5{}^{15}N$] L-pyroglutamic acid precursor ion 136.0, product ion 89.0. In order to quantify the concentration of each of the target molecules, a standard curve was prepared with the pure compounds without isotopic labeling (Zhang et al. 2019). For each original rescue and control experiment, three biological replicates were performed. The raw data from all mass experiments performed can be found in the additional supplementary document "RawDataMassExperiments."

Inhibition of Proline Prototrophy Restoration by the Presence of Glutamine in the Culture Medium

Original rescues were grown at 37 °C in kanamycin-supplemented minimal medium until saturation (about 36 h). As a control, the Keio *ΔproB* knockout strain was grown in minimal medium supplemented with kanamycin and 50 μM proline. All cultures were then diluted to $OD_{600} = 0.001$ and grown at 37 °C under shaking. Experiments were carried out in minimal medium (supplemented with kanamycin) in the absence and presence of 2 mM glutamine. In the case of Keio *ΔproB* knockout cells, the medium was additionally supplemented with proline to enable growth. Growth was monitored by measuring the optical density at 600 nm. For each original rescue and control experiment, three biological replicates were performed.

Quantification of Transcription Levels for the *glnA* Gene

Transcription levels for the *glnA* gene were quantified using the quantitative polymerase chain reaction (qPCR). Experiments were carried out with *ΔproB* knockout cells harboring variants of the *glnA* gene with mutations that were previously found in rescue proline auxotrophy, as well as with cells harboring the wild-type *glnA* gene, the latter serving as control samples. For each experiment, a starter culture in minimal medium was prepared and diluted 1:200 in minimal medium. In all cases, the minimal medium was supplemented with kanamycin, and in the case of the control samples, proline at 0.7 μg/ml was also added to allow growth. After the 1:200 dilution, samples

were maintained at 37 °C on shaking until OD values at 600 nm of ~0.9 were reached. Cultures were then centrifuged and the supernatant discarded. mRNA from each sample was extracted using Trizol method and quality validated by $A_{260/A280}$ in NanoDrop 2000c. Reverse transcription was performed with the TaqMan Advanced mRNA cDNA Synthesis kit (Applied Biosystem, Foster City, CA, USA). qPCR was performed on a 96-well plate using the SYBR Green probe (Life Technologies, Carlsbad, CA, USA), with the QuantStudio 6 Flex Real-Time PCR System (Applied Biosystems). qPCR reactions were performed according to the following protocol: 95 °C for 10 min for enzyme activation; followed by 40 cycles of 15 s at 95 °C and 1 min at 60 °C for denaturing and annealing/extension. The primers used were designed using Primer-Blast (NIH) software to generate a PCR product of size between 60 and 150 nt and have a primer melting temperature within the range of 63–68 °C. All samples were run in triplicate, with an NTC (nontemplate control) in each plate. mRNA levels were quantified using the comparative threshold cycle (Ct) method relative to the *ffh* gene (signal recognition particle protein) level as an endogenous control. Relative levels for each sample were calculated as the difference between *glnA* and *ffh* values ($\Delta C_t = C_{t\ glnA} - C_{t\ ffh}$). Three biological replicates were performed for each of the samples and controls in addition to the corresponding instrumental triplicate for each of the measurements.

Surface Area Calculations

Accessible surface areas (ASA) for amino acids in the 3D structure of glutamine synthetase were calculated as described in (Ibarra-Molero et al. 1999) using a radius of 1.4 Å for the spherical probe representing a water molecule. Buried surface areas, as given in table 1, were calculated as the difference between the ASA value for the amino acid side chain when fully exposed to the aqueous solvent and the ASA value of the side chain in the 3D structure. The former ASA value was taken to be that of the side chain in a tripeptide with the same conformation as in the folded structure.

In the case of the new residues generated by the rescuing mutations, surface calculations require a predicted 3D structure for the single-mutation variant. This prediction was carried out using the UCSF Chimera program (Pettersen et al. 2004). The glutamine synthetase wild-type structure with PDB ID: 1FPY was used as template, and the Dunbrack backbone-dependent rotamer library was applied to the structural modeling step for each variant. Subsequently, energy minimization was performed with the minimize structure module implemented in version 1.16 of the UCSF Chimera package (https://www.cgl.ucsf.edu/chimera).

Supplementary Material

Supplementary data are available at *Molecular Biology and Evolution* online.

Acknowledgments

This work was supported by Human Frontier Science Program Grant RGP0041/2017 (J.M.S.R. and B.S.), Spanish Ministry of Science and Innovation FEDER Funds Grants RTI2018-097142-B-100 and PID2021-124534OB-100 (J.M.S.R.), National Aeronautics and Space Administration (NASA) Grant 80NSSC18K1277 (B.S.) Grant RYC2021-031155-I (E.M.C.) from Spanish Ministry of Science and Innovation and NextGenerationEU/PRTR and Grant E-BIO-464-UGR-20 (E.M.C.) from FEDER Funds and Consejería de Economía, Conocimiento, Empresas y Universidad de la Junta de Andalucia. E.A.L. was a recipient of a postdoctoral fellowship from the regional Andalusian Government (2020_DOC_00541). We thank the "Centro de Supercomputacion" (ALHAMBRA-CSIRC) of the University of Granada for providing computational resources and Dr Valeria A. Risso for useful discussions, comments on the manuscript and help with the evolutionary and structural analyses of the rescuing mutations.

Author Contributions

E.M.C. carried out the selection and identification of proline auxotrophy rescues, the protein mutagenesis and preparation of protein variants, and the coupled activity assays, under the supervision of B.I.M., who also provided essential input regarding the experimental design of the study; F.M.T. and M.S.N. set up the experimental protocol for the determination of proline prototrophy restoration under the supervision of B.S.; L.I.G.R. carried out preliminary experiments to explore and eventually rule out alternative explanations for proline prototrophy restoration; E.M.C. and G.G.A. carried out the NMR experiments under the supervision of J.L.O.R.; E.M.C. and M.O.M. carried out the mass spectrometry experiments; E.A.L. performed the bioinformatics analyses for genomic DNA sequencing; A.J.M. and P.R. performed the computational calculations on the cyclization barrier under the supervision of J.M.C.; J.M.S.R. conceived and designed the research, and wrote the first draft of the manuscript; all authors discussed the manuscript, suggested modifications and improvements, and contributed to the final version.

Data Availability

Raw sequencing data are available in the Sequence Read Archive (SRA) under the PRJNA839218 BioProject accession number: http://www.ncbi.nlm.nih.gov/bioproject/839218. The raw data from all mass experiments performed can be found in the additional supplementary document "RawDataMassExperiments." The raw data from all NMR experiments performed can be found in the additional supplementary document "RawDataNMR."

Conflict of interest statement: The authors declare no competing interests.

References

Almassy R, Janson C, Hamlin R, Xuong N, Eisenberg D. 1986. Novel subunit—subunit interactions in the structure of glutamine synthetase. *Nature* **323**:304–309.

Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. Available from: http://www.bioinformatics.babraham.ac.uk/projects/fastqc

Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H. 2006. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol*. **2**:1–11.

Becke AD. 1988. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys Rev A (Coll Park)*. **38**:3098–3100.

Becke AD. 1993. A new mixing of hartree–fock and local density-functional theories. *J Chem Phys*. **98**:1372–1377.

Berg CM, Rossi JJ. 1974. Proline excretion and indirect suppression in Escherichia coli and Salmonella typhimurium. *J Bacteriol*. **118**:928–934.

Berggård T, Linse S, James P. 2007. Methods for the detection and analysis of protein-protein interactions. *Proteomics* **7**:2833–2842.

Blank D, Wolf L, Ackermann M, Silander OK. 2014. The predictability of molecular evolution during functional innovation. *Proc Natl Acad Sci U S A*. **111**:3044–3049.

Copley SD. 2009. Evolution of efficient pathways for degradation of anthropogenic chemicals. *Nat Chem Biol*. **5**:559–566.

Copley SD. 2015. An evolutionary biochemist's perspective on promiscuity. *Trends Biochem Sci*. **40**:72–78.

Cravens A, Payne J, Smolke CD. 2019. Synthetic biology strategies for microbial biosynthesis of plant natural products. *Nat Commun*. **10**:1–12.

Csonka LN, Leisinger T. 2007. Biosynthesis of proline. *EcoSal Plus*. **2**:1–20.

de Carvalho Fernandes G, Turchetto-Zolet AC, Pereira Passaglia LM. 2022. Glutamine synthetase evolutionary history revisited: tracing back beyond the last universal common ancestor. *Evolution (N Y)*. **76**:605–622.

de Crécy-Lagard V, Haas D, Hanson AD. 2018. Newly-discovered enzymes that function in metabolite damage-control. *Curr Opin Chem Biol*. **47**:101–108.

Digianantonio KM, Hecht MH. 2016. A protein constructed de novo enables cell growth by altering gene regulation. *Proc Natl Acad Sci U S A*. **113**:2400–2405.

Digianantonio KM, Korolev M, Hecht MH. 2017. A non-natural protein rescues cells deleted for a key enzyme in central metabolism. *ACS Synth Biol*. **6**:694–700.

Eisenberg D, Gill HS, Pfluegl GMU, Rotstein SH. 2000. Structure–function relationships of glutamine synthetases. *Biochim Biophys Acta*. **1477**:122–145.

Fichman Y, Gerdes SY, Kovács H, Szabados L, Zilberstein A, Csonka LN. 2015. Evolution of proline biosynthesis: enzymology, bioinformatics, genetics, and transcriptional regulation. *Biol Rev Camb Philos Soc*. **90**:1065–1099.

Francl MM, Pietro WJ, Hehre WJ, Binkley JS, Gordon MS, DeFrees DJ, Pople JA. 1982. Self-consistent molecular orbital methods. XXIII. A polarization-type basis set for second-row elements. *J Chem Phys*. **77**:3654–3665.

Frisch GW, Trucks HB, Schlegel GE, Scuseria MA, Robb JR, Cheeseman G, Scalmani V, Barone B, Mennucci GA, Petersson H, *et al.* 2010. Gaussian 09, revision B.01. Wallingford (CT): Gaussian Inc.

Gamiz-Arco G, Risso VA, Candel AM, Inglés-Prieto A, Romero-Romero ML, Gaucher EA, Gavira JA, Ibarra-Molero B, Sanchez-Ruiz JM. 2019. Non-conservation of folding rates in the thioredoxin family reveals degradation of ancestral unassisted-folding. *Biochem J*. **476**:3631–3647.

Garrison E, Marth G. 2012. Haplotype-based variant detection from short-read sequencing. arXiv:1207.3907.

Gavin AC, Aloy P, Grandi P, Krause R, Boesche M, Marzioch M, Rau C, Jensen LJ, Bastuck S, Dümpelfeld B, *et al.* 2006. Proteome survey reveals modularity of the yeast cell machinery. *Nature* **440**:631–636.

Godoy-Ruiz R, Ariza F, Rodriguez-Larrea D, Perez-Jimenez R, Ibarra-Molero B, Sanchez-Ruiz JM. 2006. Natural selection for kinetic stability is a likely origin of correlations between mutational effects on protein energetics and frequencies of amino acid occurrences in sequence alignments. *J Mol Biol*. **362**:966–978.

Grunwald P. 2018. *Biocatalysis biochemical fundamentals and applications*. 2nd ed. Singapore: World Scientific.

Gupta OP, Deshmukh R, Kumar A, Singh SK, Sharma P, Ram S, Singh GP. 2022. From gene to biomolecular networks: a review of evidences for understanding complex biological function in plants. *Curr Opin Biotechnol*. **74**:66–74.

Huang X, Holden HM, Raushel FM. 2001. Channeling of substrates and intermediantes in enzyme-catalyzed reactions. *Annu Rev Biochem*. **70**:149–180.

Hwang T, Shaka A. 1995. Water suppression that works. Excitation sculpting using arbitrary wave-forms and pulsed-field gradients. *J Magn Reson A*. **112**:275–279.

Ibarra-Molero B, Loladze VV, Makhatadze GI, Sanchez-Ruiz JM. 1999. Thermal versus guanidine-induced unfolding of ubiquitin. An analysis in terms of the contributions from charge-charge interactions to protein stability. *Biochemistry* **38**:8138–8149.

Kastritis PL, Gavin AC. 2018. Enzymatic complexes across scales. *Essays Biochem*. **62**:501–514.

Khersonsky O, Tawfik DS. 2010. Enzyme promiscuity: a mechanistic and evolutionary perspective. *Annu Rev Biochem*. **79**:471–505.

Kim J, Flood JJ, Kristofich MR, Gidfar C, Morgenthaler AB, Fuhrer T, Sauer U, Snyder D, Cooper VS, Ebmeier CC, *et al.* 2019. Hidden resources in the *Escherichia coli* genome restore PLP synthesis and robust growth after deletion of the essential gene *pdxB*. *Proc Natl Acad Sci U S A*. **116**:24164–24173.

Kim J, Kershner JP, Novikov Y, Shoemaker RK, Copley SD. 2010. Three serendipitous pathways in *E. coli* can bypass a block in pyridoxal-5′-phosphate synthesis. *Mol Syst Biol*. **6**:1–13.

Kirschner M, Gerhart J. 2005. The plausibility of life: resolving Darwin's Dilemma. New Haven (CT): Yale University Press.

Kumada Y, Bensont DR, Hillemannt D, Hostedt TJ, Rochefortt DA, Thompson CJ, Wohllebent W, Tatenovii Y. 1993. Evolution of the glutamine synthetase gene, one of the oldest existing and functioning genes. *Proc Natl Acad Sci U S A*. **90**:3009–3013.

Kumar A, Bachhawat AK. 2012. Pyroglutamic acid: throwing light on a lightly studied metabolite. *Curr Sci*. **102**:288–297.

Lee N, Spears ME, Carlisle AE, Kim D. 2020. Endogenous toxic metabolites and implications in cancer therapy. *Oncogene* **39**:5709–5720.

Lee HM, Vo PNL, Na D. 2018. Advancement of metabolic engineering assisted by synthetic biology. *Catalysts* **8**:1–19.

Lee C, Yang W, Parr RG. 1988. Development of the Colic-Salvetti correlation-energy formula into a functional of the electron density. *Phys Rev B*. **37**:785–789.

Lerma-Ortiz C, Jeffryes JG, Cooper AJL, Niehaus TD, Thamm AMK, Frelin O, Aunins T, Fiehn O, de Crécy-Lagard V, Henry CS, *et al.* 2016. 'Nothing of chemistry disappears in biology': the top 30 damage-prone endogenous metabolites. *Biochem Soc Trans*. **44**:961–971.

Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**:1754–1760.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**:2078–2079.

Martin M. 2012. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J* [Internet]. **17**:10–12. Available from: https://doi.org/10.14806/ej.17.1.200

Mcloughlin SY, Copley SD. 2008. A compromise required by gene sharing enables survival: implications for evolution of new enzyme activities. *Proc Natl Acad Sci U S A*. **105**:13497–13502.

Miertus S, Scrocco E, Tomasi J. 1981. Electrostatic interaction of a solute with a continuum. A direct utilization of AB initio molecular

potentials for the prevision of solvent effects. *Chem Phys*. **55**: 117–129.
Miles EW, Rhee S, Davies DR. 1999. The molecular basis of substrate channeling. *J Biol Chem* [Internet]. **274**:12193–12196. Available from: http://www.jbc.org
Noda-Garcia L, Liebermeister W, Tawfik DS. 2018. Metabolite-enzyme coevolution: from single enzymes to metabolic pathways and networks. *Annu Rev Biochem*. **87**:187–216.
Pál C, Papp B, Lercher MJ. 2005. Adaptive evolution of bacterial metabolic networks by horizontal gene transfer. *Nat Genet*. **37**: 1372–1375.
Pareek V, Sha Z, He J, Wingreen NS, Benkovic SJ. 2021. Metabolic channeling: predictions, deductions, and evidence. *Mol Cell*. **81**: 3775–3785.
Patrick WM, Quandt EM, Swartzlander DB, Matsumura I. 2007. Multicopy suppression underpins metabolic evolvability. *Mol Biol Evol*. **24**:2716–2722.
Peracchi A. 2018. The limits of enzyme specificity and the evolution of metabolism. *Trends Biochem Sci*. **43**:984–996.
Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem*. **25**: 1605–1612.
Phillips R, Kondev J, Theriot J, Garcia H, editors. 2012. Physical biology of the cell. 2nd ed. New York: Garland Sciences. Chapter 3 p. 87–136, chapter 14 p. 543–572.
Risso VA, Gavira JA, Gaucher EA, Sanchez-Ruiz JM. 2014. Phenotypic comparisons of consensus variants versus laboratory resurrections of Precambrian proteins. *Proteins* **82**:887–896.
Risso VA, Manssour-Triedo F, Delgado-Delgado A, Arco R, Barroso-delJesus A, Ingles-Prieto A, Godoy-Ruiz R, Gavira JA, Gaucher EA, Ibarra-Molero B, *et al.* 2015. Mutational studies on resurrected ancestral proteins reveal conservation of site-specific amino acid preferences throughout evolutionary history. *Mol Biol Evol*. **32**:440–455.
Schmitt DL, An S. 2017. Spatial organization of metabolic enzyme complexes in cells. *Biochemistry* **56**:3184–3196.
Schneider CA, Rasband WS, Eliceiri KW. 2012. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods*. **9**:671–675.
Schulenburg C, Miller BG. 2014. Enzyme recruitment and its role in metabolic expansion. *Biochemistry* **53**:836–845.
Scott KA, Alonso DOV, Sato S, Fersht AR, Daggett V. 2007. Conformational entropy of alanine versus glycine in protein denatured states. *Proc Natl Acad Sci U S A*. **104**:2661–2666.
Seemann T. 2015. snippy: Rapid haploid variant calling and core genome alignment. Available from: https://github.com/tseemann/snippy
Smith CJ, Deutch AH, Rushlow KE. 1984. Purification and characteristics of a gamma-glutamyl kinase involved in Escherichia coli proline biosynthesis. *J Bacteriol*. **157**:545–551.
Srere PA. 1987. Complexes of sequential metabolic enzymes. *Annu Rev Biochem*. **56**:89–124.
Stadtman ER. 2004. Regulation of glutamine synthetase activity. *EcoSal Plus*. **1**:1–12.
Stadtman ER, Shapiro BM, Kingdon HS, Woolfolk CA, Hubbard JS. 1968. Cellular regulation of glutamine synthetase activity in Escherichia coli. *Adv Enzyme Regul*. **6**:257–289.
Tamura K, Stecher G, Kumar S. 2021. MEGA11: molecular evolutionary genetics analysis version 11. *Mol Biol Evol*. **38**:3022–3027.
Tomasi J, Persico M. 1994. Molecular interactions in solution: an overview of methods based on continuous distributions of the solvent. *Chem Rev*. **94**:2027–2094.
Tullman-Ercek D. 2015. 'Channeling' Hans Krebs. *Nat Chem Biol*. **11**: 180–181.
Verkman A. 2002. Solute and macromolecule diffusion in cellular aqueous compartments. *Trends Biochem Sci*. **27**:27–33.
Wagner A. 2011. Metabolic innovation. In: The origins of evolutionary innovations: a theory of transformative change in living systems. Oxford (UK): Oxford University Press. p. 18–32.
Wang J, Lu Y, Nawaz MZ, Xu J. 2018. Comparative genomics reveals evidence of genome reduction and high extracellular protein degradation potential in Kangiella. *Front Microbiol*. **9**:1–13.
Zhang P, Chan W, Ang IL, Wei R, Lam MMT, Lei KMK, Poon TCW. 2019. Revisiting fragmentation reactions of protonated α-amino acids by high-resolution electrospray ionization tandem mass spectrometry with collision-induced dissociation. *Sci Rep*. **9**:1–10.
Zhang Y, Fernie AR. 2021. Metabolons, enzyme–enzyme assemblies that mediate substrate channeling, and their roles in plant metabolism. *Plant Commun*. **2**:1–16.