Evolving a mitigation of the stress response pathway to change the basic chemistry of life

Isabella Tolle[1†], Stefan Oehm[2†], Michael Georg Hoesl[3], Christin Treiber-Kleinke[4], Lauri Peil[5, 6], Abdul-Rahman Adamu Bukari[7], Torsten Semmler[8], Juri Rappsilber[5,9], Aleeza Gerstein[*7,10], Nediljko Budisa[*1,11]

[1] Technische Universität Berlin, Institut für Chemie, Müller-Breslau Str. 10, 10623 Berlin, Germany

[2] Pulcra Chemicals GmbH, Isardamm 79-83, 82538 Geretsried, Germany

[3] Clariant Produkte Deutschland GmbH., Semmelweisstraße 1, 82152 Planegg, Germany

[4] Freie Universität Berlin, Institut für Chemie und Biochemie, Arnimallee 20, 14195 Berlin, Germany

[5] Wellcome Centre for Cell Biology, University of Edinburgh, Edinburgh, UK

[6] Icosagen Cell Factory OÜ, Eerika tee 1, Õssu, 61713 Tartumaa, Estonia (current address)

[7] Department of Microbiology, University of Manitoba, R3T 2N2 Winnipeg, MB, Canada

[8] Robert Koch Institut, Nordufer 20, 13353 Berlin (Germany)

[9] Technische Universität Berlin, Institut für Biotechnologie, Gustav-Meyer-Allee 25, 13355 Berlin (Germany)

[10] Department of Statistics, University of Manitoba, R3T 2N2 Winnipeg, MB, Canada

[11] Department of Chemistry, University of Manitoba, R3T 2N2 Winnipeg, MB, Canada

†These authors have contributed equally to this study

*Author to whom correspondence should be addressed:

E-Mail: aleeza.gerstein@umanitoba.ca; nediljko.budisa@tu-berlin.de; nediljko.budisa@umanitoba.ca

ABSTRACT

Billions of years of evolution have produced only slight variations in the standard genetic code, and the number and identity of proteinogenic amino acids have remained mostly consistent throughout all three domains of life. These observations suggest a certain rigidity of the genetic code and prompt musings as to the origin and evolution of the code. Here we conducted an adaptive laboratory evolution (ALE) to push the limits of the code restriction, by evolving *Escherichia coli* to fully replace tryptophan, thought to be the latest addition to the genetic code, with the analog L-β-(thieno[3,2-*b*]pyrrolyl)alanine ([3,2]Tpa). We identified an overshooting of the stress response system to be the main inhibiting factor for limiting ancestral growth upon exposure to β-(thieno[3,2-*b*]pyrrole ([3,2]Tp), a metabolic precursor of [3,2]Tpa, and Trp limitation. During the ALE, *E. coli* was able to "calm down" its stress response machinery, thereby restoring growth. In particular, the inactivation of RpoS itself, the master regulon of the general stress response, was a key event during the adaptation. Knocking out the *rpoS* gene in the ancestral background independent of other changes conferred growth on [3,2]Tp. Our results add additional evidence that frozen regulatory constraints rather than a rigid protein translation apparatus are Life's gatekeepers of the canonical amino acid repertoire. This information will not only enable us to design enhanced synthetic amino acid incorporation systems but may also shed light on a general biological mechanism trapping organismal configurations in a status quo.

SIGNIFICANCE STATEMENT

The (apparent) rigidity of the genetic code, as well as its universality, have long since ushered explorations into expanding the code with synthetic, new-to-nature building blocks and testing its boundaries. While nowadays even proteome-wide incorporation of synthetic amino acids has been reported on several occasions[1–3], little is known about the underlying mechanisms.

We here report ALE with auxotrophic *E. coli* that yielded successful proteome-wide replacement of Trp by its synthetic analog [3,2]Tpa accompanied with the selection for loss of RpoS[4] function. Such laboratory domestication of bacteria by the acquisition of *rpoS* mitigation mutations is beneficial not only to overcome the stress of nutrient (Trp) starvation but also to evolve the paths to use environmental xenobiotics (e.g. [3,2]Tp) as essential nutrients for growth.

We pose that regulatory constraints rather than a rigid and conserved protein translation apparatus are Life's gatekeepers of the canonical amino acid repertoire (at least where close structural analogs are concerned). Our findings contribute a step towards understanding possible environmental causes of genetic changes and their relationship to evolution.

Our evolved strain affords a platform for homogenous protein labeling with [3,2]Tpa as well as for the production of biomolecules[5], which are challenging to synthesize chemically. Top-down synthetic biology will also benefit greatly from breaking through the boundaries of the frozen bacterial genetic code, as this will enable us to begin creating synthetic cells capable to utilize an expanded range of substrates essential for life.

INTRODUCTION

Evolution is an ongoing and pervasive process for all life on the planet. The mechanisms of evolution have yielded extant living organisms with immense biodiversity. Vast diversity exists along nearly all possible axes imaginable in body form, structure, and function. Given the potential for adaptation to explore different trajectories and organisms to live in virtually all known ecosystems on the planet, it is perhaps most unexpected not to find variation in the building blocks themselves: all life is based on the same 20 proteinogenic amino acids. The extant genetic code, which appeared >3 billion years ago, relates each of the 20 canonical amino acids to specific nucleotide triplets of genes and mRNAs and is an almost invariant ("frozen") feature of life on Earth[6]. However, the genetic code likely rather started with a few small, simple amino acids that can be produced via a proto-metabolism, evolving gradually over time to add more complex amino acids to the repertoire with diverse functionalities. While the discovery of Selenocysteine and Pyrrolysine, the 21st and 22nd proteinogenic amino acids, implies some flexibility to the code, they are incorporated only at very specific positions in certain proteins. This begs the fundamental questions of why evolution stopped at these 20 amino acids[7–9] and whether it is possible to alter the number and identity of these amino acids in a proteome-wide manner? Can a lineage with an altered set of canonical amino acids be evolved from an existing species? If we were able to experimentally alter natural species step by step so that they adopt different genetic codes with synthetic amino acids - then we would pave the way for the creation of artificial diversity and synthetic life.

Tryptophan is thought to be the latest addition to the genetic code with the highest metabolic cost of production[10,11]. Trp is one of only two canonical amino acids (cAAs) encoded by a single codon[12]. It is the least abundant amino acid in eukaryotic proteomes including humans (1.24% and 1.22%, respectively) and viruses (1.19%), and the second least abundant behind cysteine in both archean (1.03%) and bacterial proteomes (1.27%)[13]. By contrast, the most prevalent amino acid in all life forms, leucine, has an abundance between 8.84 % (viruses) and 10.09% (bacteria)[13]. Accordingly, Trp-residues have long been targets to demonstrate the chemical mutability of the genetic code by replacing them with noncanonical synthetic analogs[1,2]. In 2014, Yu et al.[14] sequenced and analyzed a strain from an adaptation experiment of a Trp-auxotrophic *Bacillus subtilis* strain towards the noncanonical amino acid (ncAA) 4-fluoro-tryptophan (4-F-Trp) published more than 30 years earlier[2]. Only a few rounds of selection were necessary to give rise to a 4-F-Trp-tolerating strain with mutations of rpoC, in the sigma factors *sigI* [15] and *sigB* [16] as general stress sigma factors in *B. subtilis*. Most recently, a similar experiment performed with Trp-auxotrophic *Escherichia coli* revealed point mutations in the genes encoding non-essential vegetative sigma factors *rpoA* and *rpoC*, which help to reconfigure transcription under different environmental conditions, including stress[17].

In this experimental vein, we have recently achieved a complete Trp substitution in the *Escherichia coli* proteome by its surrogate L-β-(thieno[3,2-*b*]pyrrolyl)-alanine ([3,2]Tpa, Figure 1) using adaptive laboratory evolution (ALE)[3]. Selective pressure was exerted on a Trp-auxotrophic strain which, in order to grow, was forced to perform conversion of [3,2]Tp and L-Ser to [3,2]Tpa through the reaction catalyzed by the endogenous enzyme Trp synthase (TrpS). However, when we tried to deduce the underlying adaptation mechanisms using 'omics data, we realized that the genomic configuration of the ancestral strain we used was flawed. Specifically, it led to several additional selection pressures and subsequently

led to off-target mutations that rendered a clear deduction of the genetic basis of the adaptation to [3,2]Tpa impossible (manuscript in prep).
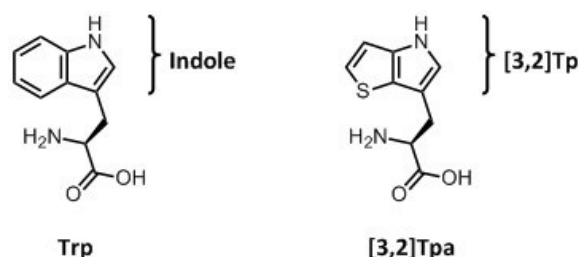

Figure 1 I Chemical structures of Tryptophan and L-β-(thieno[3,2-*b*]pyrrolyl)alanine ([3,2]Tpa). The brackets indicate the structures of the precursors indole and β-thieno[3,2-*b*]pyrrole ([3,2]Tp).

Here, we have used what we learned from the first ALE experiment to design a novel ALE experiment using an *E. coli* strain with an optimized genomic configuration[18]. Using the new experimental framework, we were able to demonstrate that *E. coli* can succeed in overcoming the frozen state of the genetic code to replace tryptophan with the analog [3,2]Tpa, and identify that the primary adaptive mechanism is an alternation of the RpoS-mediated general stress response. The co-occurrence of mutations in stress regulators of Gram-positive bacteria supports the *E. coli* adaptation model towards Trp-analog usage presented here.

The general stress response is activated upon entering stationary phase and provides cross-protection to a wide spectrum of stresses *E. coli* are confronted with in their natural environments (*e.g.*, nutrient limitation, competition of resources, fluctuating temperature, pH, osmolarity)[19–23] The general stress response is mediated by the alternative sigma factor RpoS (σ^S), which interacts with the core RNA polymerase (RNAP)[24]. RpoS governs the expression of a vast set of genes[25] and shares control of various promoters with the vegetative sigma factor RpoD (σ^{70}). The RpoS response is regulated at multiple levels that include transcription, translation, degradation, and activity[26].

What happened during the evolution of bacteria and other microbes and organisms over several billions of years cannot be fully compared to the phenomenon of proteome-wide ncAA insertion experiments over the last 50 years[27]. Contemporary selection pressure of ncAA use and disposal is much more intense; selection is largely for survival in hostile minimal media environments rather than for traits providing fitness in slowly evolving populations in natural settings. In summary, what is occurring in our lifetime is an evolutionary process intensified by anthropogenic influences rather than the slower, random course of natural evolution.

During experimentally designed long-term serial dilutions, i.e. ALE, the exposure of bacterial metabolic prototypes to high concentrations of ncAAs for extended periods creates a severe selection pressure and leads to higher levels of incorporation. Knowledge of the intermediate steps in this important process would be revealing—how many steps are there from parent strain to fully adapted bacteria?

RESULTS

Adaptive Laboratory Evolution (ALE) on [3,2]Tp

We constructed an MG1655 derivative strain of *E. coli*, where the tryptophanase, as well as the majority of the Trp operon (*trpLEDC*), was knocked-out while retaining the tryptophan synthase (*trpBA*) on the chromosome under its natural regulation of *trpR* (MG1655 *ΔtnaA ΔtrpLEDC*). We refer to this new starting strain as TUB00 (Figure S1 A). The ALE was performed in new minimal medium (NMM)[28].

In the first phase of the experiment, the concentration of indole was continually reduced; after only 29 passages (~ 190 generations) the population was capable of growing in medium without the addition of any indole (Figure 2A). To avoid Trp contamination from commercial amino acid preparations, as described in other experiments[1], in the second phase the other 19 canonical amino acids were removed from the medium in so-called metabolic blocks. After another 97 passages (~ 630 generations) and a reduction of the cultivation temperature from 37°C to 30°C, methionine was the only canonical amino acid left in the medium. Within 10 additional serial dilutions and a total of 138 passages, the last canonical amino acid (cAA) was removed from the medium without loss of growth. During the last passages, the cultivation temperature was gradually increased back to 37°C (phase 3, Figure 2A). After a total of 170 passages, the adaptation experiment was terminated and the resulting strain was termed TUB170. We denote the media compositions in the following format throughout: NMM(a/b/c; a = number of cAAs; b = [3,2]Tp concentration in µM; c = indole concentration in µM). The ALE was initiated with NMM(19/25/1) and concluded with NMM(0/25/0) (Figure 2A).

The growth rates of TUB00, TUB85 (isolated after 85 passages in phase 2), and TUB170 were compared in different media compositions (Figure 2B). TUB170 had a significantly higher growth rate than TUB00 and TUB85 in all NMM environments including the starting condition medium (NMM (19/30/1, Figure 2B). Interestingly, the growth rate of TUB170 increased when supplied with 100 µM [3,2]Tp (NMM(0/100/0)) compared to 30 µM [3,2]Tp (NMM(0/30/0)). A higher efficiency of the tRNA^{Trp} aminoacylation to [3,2]Tpa-tRNA^{Trp} or of [3,2]Tp conversion to [3,2]Tpa by the endogenous TrpBA could be a reason for this observation. Relative to TUB00 and TUB85, TUB170 had a significantly decreased growth rate in the standard rich medium LB (Figure 2B), indicative of a growth trade-off. TUB85 had the same growth rate as TUB00 except in the environment it was isolated in (NMM (13/30/0)), when it had an intermediate growth rate (Figure 2B), suggesting that this isolate contains a beneficial mutation that is specific to that environment.
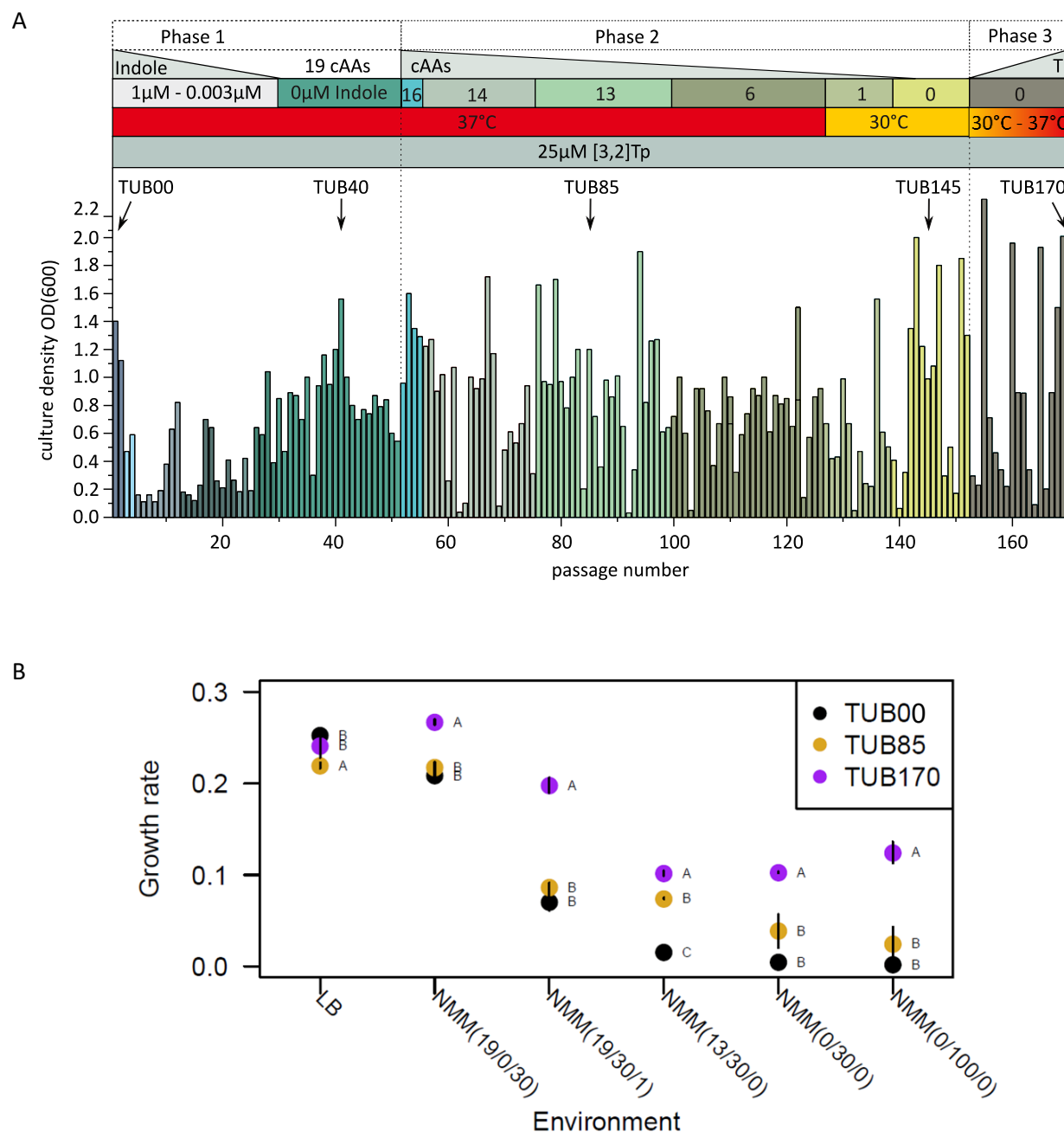
Figure 2 I Adaptive Laboratory Evolution (ALE) on [3,2]Tp. A) Overview of optical densities throughout the adaptation experiment and the corresponding cultivation conditions. Phase 1: the indole concentration was reduced from 1µM to 0µM in several steps. Phase 2: All other 19 cAAs were removed according to metabolic blocks, whereby the temperature needed to be reduced to 30°C for removal of the last amino acid, Met. Phase 3: The temperature was gradually increased back to 37°C. B) Doubling times of TUB00, TUB85, and TUB170. Data were collected at 30 °C. The letters A, B, C indicate Tukey test results (statistical significance, p < 0.05), when different strains have different letters they are statistically different.

Mutations occurred in a wide range of genes

To characterize genetic changes that arose during the ALE, single colony isolates from different experimental phases were isolated and whole genome sequenced: TUB40 (phase 1), TUB85 (phase 2), TUB145 (phase 2), and TUB170 (phase 3). We identified 42 mutations within these four isolates, which provides a snapshot of the evolutionary dynamics of the ALE (Figure 3A, Table S4). An examination of the shared mutations (defined as any mutation that was present in at least two isolates) and private mutations (those present in only a single isolate) indicates that multiple lineages were likely present for the majority of the 170 transfers of the ALE (Figure 3B). Nine mutations were found in all sequenced isolates, indicating they arose in the first 40 transfers in the same genetic background (lineage) that spawned all sequenced isolates. These nine mutations occurred in genes involved in amino acid biosynthesis and processing (*aroG, astB,* and *leuS*), cell morphogenesis (*fliD, cvrA*), membrane proteins (*ylcJ, yejM*), oxidoreduction (*ydfI*), and in the promoter region of *yobF*, a small protein with unknown function.

At some point before transfer 40, there was a split into at least two lineages. One led to TUB40, which has five private mutations. The second lineage was ancestral to the other three sequenced isolates (TUB85, TUB145, TUB170). The second lineage contains 13 shared single nucleotide variants in kinases (*glcK*, *pyrH*, *phoR*), phosphatases (*gpp*), reductases (*nfeF*), proteases (*lon*, *ftsH*), stress response factors (*rpoS, lrhA, gpp, sanA, sgrR*), and a mutation in a pseudogene (*ybfI*) and two large deletions. The larger is a 23 kb deletion involving 35 genes including the Rac prophage. The second is a 6 kb deletion involving 10 genes, beginning within the putative defective integrase of the Qin prophage (*intQ*) and ending within the putative selenite reductase (*ynfE*). No private mutations were identified in TUB85, indicating that this isolate is a common ancestor to both TUB145 and TUB170.
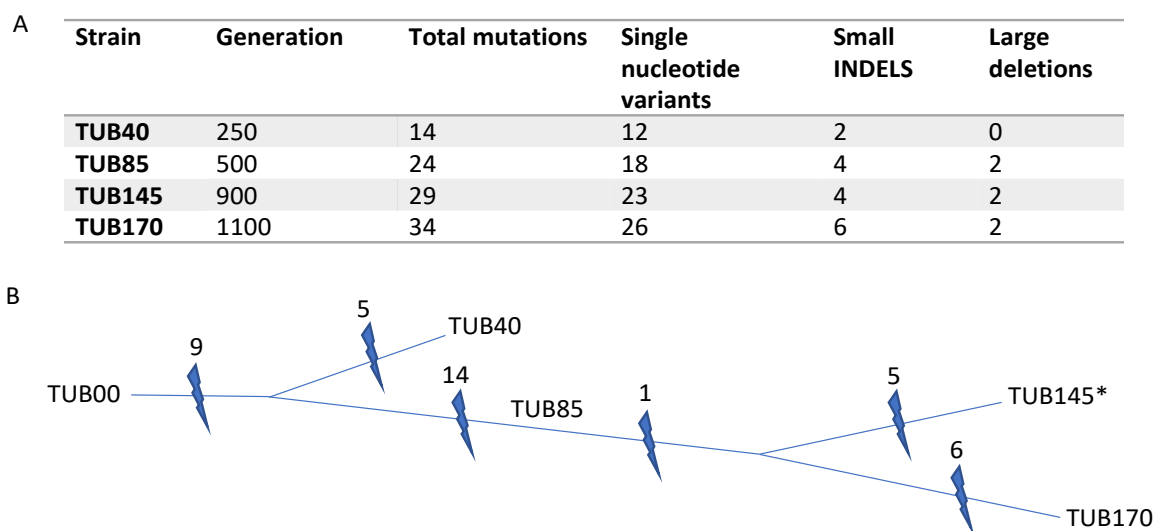
A

Strain	Generation	Total mutations	Single nucleotide variants	Small INDELS	Large deletions
TUB40	250	14	12	2	0
TUB85	500	24	18	4	2
TUB145	900	29	23	4	2
TUB170	1100	34	26	6	2

B

Figure 3 I Genomic analysis of the ALE. A) The number of mutations identified in each sequenced isolate. Generations are estimated as described elsewhere[29]. B) A dendrogram depiction of the evolutionary history of the sequenced strains from the ALE experiment as they accumulate mutations (lightning bolt).

Only a single common mutation, in *ppc* (a phosphoenolpyruvate carboxylase), was found in both TUB145 and TUB170, suggesting that these two lineages split shortly after transfer 85. An additional four (TUB145) and nine (TUB170) private mutations were found in these two isolates. Interestingly, two genes with different mutations were found in these two terminal isolates: *cysN* (sulfate adenylyltransferase) and *gltA* (citrate synthetase), suggestive of strong selection for variants in these genes.

Mutations involved in stress response

The mutation in the stress response gene *rpoS* (present in TUB85, TUB145, TUB170) particularly caught our attention. The mutation occurred at D118 (Figure 4A), which abolishes the salt bridge between D118 in σ^S and R275 in the RpoC-subunit of the core-RNAP (Figure 4B). The aspartate is conserved in most sigma factors (σ^S, σ^D, σ^H) and is an important binding motif. To verify that an altered stress response is beneficial for the adaptation to [3,2]Tp, the *rpoS* gene was deleted in TUB00 via CRISPR/Cas9. We reasoned that reduced σ^S-mediated stress response will improve the growth on the indole-analog compared to the wildtype. Therefore, the *ΔrpoS* strain, TUB170, and TUB00 were grown in media with varying indole concentrations. The optical density of the *ΔrpoS* strain was higher than that of the ancestral strain under all conditions (Figure 4 C), indicating increased tolerance towards the adverse effects of [3,2]Tp supplementation. Notably, even in media completely devoid of a Trp source (i.e., NMM(19/30/0) and NMM(0/30/0)), considerable growth of the *ΔrpoS* strain can be detected.

Additional general stress response proteins were also identified. Two of the mutations that arose before transfer 40 and were present in all sequenced isolates occurred in the promoter regions of the transcription units *yobF-cspC* and *lrhA*. The stress protein CspC stabilizes the *rpoS* mRNA[30], while the DNA-binding transcriptional dual regulator LrhA represses *rpoS* translation[31]. A threonine to proline mutation in *gpp* was found in TUB85, TUB145, and TUB170, which the Protein Variation Effect Analyzer (PROVEAN)[32,33] predicts to be deleterious. GppA hydrolyzes the 5' phosphate from pppGpp to produce ppGpp, thereby fine-tuning the ratio between the two nucleotides[34]. This alarmone accumulates in response to amino acid (or other) starvation as part of the stringent response, which is also connected to the σ^S-mediated general stress response[35]. The synthesis and stability of RpoS is enhanced by (p)ppGpp, thus boosting the expression of genes involved in stress protection[36,37], and accumulation of (p)ppGpp or RpoS results in diminished fitness under nutrient-limited conditions[38]. Furthermore, (p)ppGpp increases the availability of core RNAP[39], while indirectly decreasing cellular RpoD, thereby promoting binding of RpoS to RNAP[40,41].
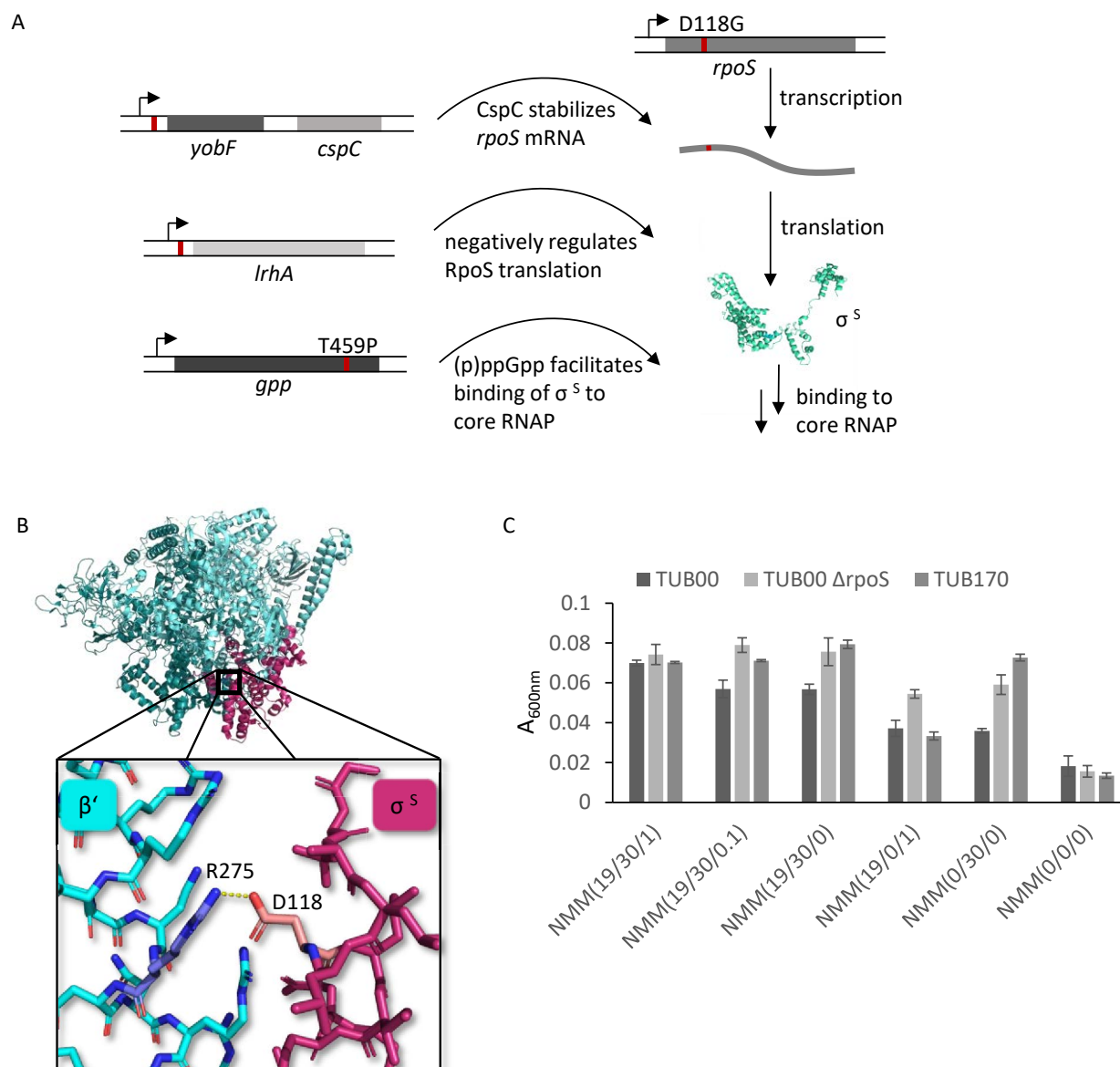
Figure 4 I Mutations involved in the RpoS-mediated stress response. A) Schematic representation of the mutations directly involved in the general stress response and their impact on the stress response. B) The interaction between pseudo-Asp118 (rose) of the sigma factor S and Arg275 (blue) of the RpoC subunit of the core RNAP (PDB: 6OMF). The salt bridge is abolished by the mutation to Gly. C) Optical densities of the *ΔrpoS* mutant compared to the ancestral strain and the adapted strain TUB170 in the presence of varying indole concentrations after 48 h. NMM(19/30/1): starting conditions, NMM(19/30/0): end of phase 1, NMM(19/0/1): without [3,2]Tp, NMM(0/30/0): end conditions, NMM(0/0/0): without amino acids or precursors.

Mutations deactivating proteases accompany the adaptation

Three mutations that arose in protease genes were predicted to be deleterious by PROVEAN (*lon* and *ftsH* were identified in TUB85, TUB145, and TUB170; *clpP* was found only in TUB170). We hypothesize that due to its physicochemical differences to Trp, the incorporation of [3,2]Tpa into proteins potentially yields locally or globally misfolded proteins, as reduction of protein stability upon [3,2]Tpa incorporation has

been previously observed.[42] Such misfolded proteins are usually cleared by proteases. However, in our novel strain, every single Trp position is replaced by [3,2]Tpa, presumably giving rise to a vast number of sub-optimally folded proteins. During the adaptation experiment, it could thus be advantageous for the bacterial system to lower this proteolytic activity in order to enhance the half-life of some affected proteins. The accumulation of mutations in proteases, particularly early in the ALE, implies that impaired protein folding is a concern when replacing Trp with [3,2]Tpa in the proteome and that downregulation of proteolytic activity might be beneficial for adaptation towards ncAA-usage.

Proteomic analysis reveals major upregulation of stress proteins

Proteomic analysis was performed using stable isotope labeling by amino acids in cell culture (SILAC). Since SILAC relies on the labeling of proteins by the isotope-labeled amino acids $^{13}C_6$$^{15}N_4$-arginine and $^{13}C_6$$^{15}N_2$-lysine, the biosynthesis of arginine and lysine in the ancestral strain TUB00 was disabled via phage P1 transduction using strains from the Keio collection as donors[43]. The resulting strain was designated TUB00dKO and used as the reference in all SILAC experiments. By mixing the digested proteins from cells cultivated with and without the isotope-labeled amino acids, followed by quantification via high-resolution mass spectrometry, expression differences of proteins from cells cultured under different conditions can be analyzed.

To characterize the impact of [3,2]Tp addition on the proteome of the ancestral strain (TUB00), cells were cultivated in indole-rich medium without any [3,2]Tp (NMM(19/0/30)) and compared to the reference (TUB00dKO) cultured in medium supplied with [3,2]Tp mirroring the starting conditions, but also supplemented with the two isotope-labeled amino acids (NMM(17/30/1) + $^{13}C_6$$^{15}N_4$-arginine and $^{13}C_6$$^{15}N_2$-lysine; i.e., NMM(19/30/1) with two labeled amino acids).

Eighty-seven (87) proteins were up- or downregulated (log_2-value > ± 2.5) in the absence of [3,2]Tp (Figure S2). Eight of the 36 upregulated proteins are related to chemotaxis (e.g., CheW, CheA) and motility (e.g., FliC, FliN, FliA, FliL; Table 1), a common response upon entering stationary phase and also a part of the general stress response[44–46]. The majority of the 51 most downregulated proteins are either directly connected to the general stress response under the control of the alternative sigma factor RpoS or via participation of the DNA-binding transcriptional dual regulator H-NS (Figure S2, cluster plot). H-NS is involved in fine-tuning RpoS expression under stress conditions, both by supporting RpoS-dependent expression[47–50] over expression mediated by the housekeeping sigma factor RpoD[51–53] and by destabilizing RpoS protein[48,54,55]. Downregulated proteins involved in the RpoS-mediated general stress response include proteins from the glutamate-dependent acid resistance system: GadC, GadB, GadA (Table 1). GadAB is a glutamate decarboxylase, which scavenges one proton upon decarboxylation of glutamate to γ-aminobutyrate (GABA)[56,57], which is then secreted by the glutamate/GABA antiporter GadC.[56,58] Another downregulated protein connected to the general stress response is the multidrug efflux transporter MdtEF (Table 1), involved in detoxification and described to be induced by indole.[59]

Table 1 I Overview of up- and downregulated proteins in response to [3,2]Tp-supplementation. Isolates from the beginning of the ALE (TUB00, cultivated in NMM(19/0/30)) are compared to isolates from the middle (TUB85, cultivated in NMM(13/30/0)) and the end of the experiment (TUB170, cultivated in NMM(0/30/0)). The reference(TUB00dKO) was cultivated in a medium supplied with [3,2]Tp mirroring the starting conditions (NMM(19/30/1)). The change in abundance is reported as log2-fold and is visualized in the form of a heat map, where yellow indicates upregulation and blue indicates downregulation, white indicates no change in abundance.

			Legend		
			upregulated		downregulated
			Change of abundance in		
Function	**Gene**	**Protein**	**the absence of [3,2]Tp**	**ALE isolate vs TUB00**	
			TUB00 log2-fold	**TUB85 log2-fold**	**TUB170 log2-fold**
chemotaxis	*cheW*	Chemotaxis protein CheW	-4.5	n.a.	0
	cheA	Chemotaxis protein CheA(L)	-3.7	0	0
motility	*fliC*	Flagellar filament structural protein	-4.8	1	4.8
	fliA	RNA polymerase, sigma 28 (sigma F) factor	-3.4	n.a.	0
	fliL	Flagellar protein FliL	-1.9	n.a.	n.a.
	fliG	Flagellar motor switch protein	-3	n.a.	0
acid resistance	*gadC*	L-glutamate:4-aminobutyrate antiporter	5.3	6.1	6.5
	gadB	Glutamate decarboxylase B	5.6	6.9	8.4
	gadA	Glutamate decarboxylase A	3	5.9	0
multidrug efflux transporter	*mdtE*	Multidrug efflux pump membrane fusion protein	5.5	4	7
	mdtF	Multidrug efflux pump RND permease	3.8	2.9	3.9
Pho regulon	*phoU*	Phosphate signaling protein	-1.1	-4.2	-3
	phoA	Alkaline phosphatase	n.a.	-3.6	-2.6
	phoB	DNA-binding transcriptional dual regulator PhoB	0.7	-2.7	-2.8
	pstS/ phoS	Phosphate ABC transporter periplasmic binding protein	1.2	-3.2	-2.9
	pstB	Phosphate ABC transporter ATP binding subunit	-0.6	-3.1	-3.1

			Legend		
			upregulated		downregulated
			Change of abundance in		
Function	**Gene**	**Protein**	the absence of [3,2]Tp TUB00 log2-fold	**ALE isolate vs TUB00**	
				TUB85 log2-fold	**TUB170 log2-fold**
	ugpB	sn-glycerol 3-phosphate ABC transporter periplasmic binding protein	0.2	-2.5	-3.6
Leu biosynthesis	leuA	2-isopropylmalate synthase	0.2	-4.6	-3.1
	leuD	3-isopropylmalate dehydratase subunit LeuD	-3.2	-4.5	-3.1
	leuB	3-isopropylmalate dehydrogenase	0.1	-4.5	-1.7
	leuC	3-isopropylmalate dehydratase subunit LeuC	-3.5	-4	-2.4
Trp synthesis	trpA	Tryptophan synthase subunit α	-2.3	-2.2	-1.6
	trpB	Tryptophan synthase, β subunit dimer	-1.7	-1.9	-1.6
RpoS repressor	lrhA	DNA-binding transcriptional dual regulator LrhA	-1.4	-2.5	-1.1
rpoS stabilizer	cspC	CspA family stress protein CspC	-1	3.9	3.7
galactose catabolism	galE	UDP-glucose 4-epimerase	5.7	5.3	5
	galK	Galactokinase	5.2	4.9	4.7
	galT	Galactose-1-phosphate uridylyltransferase	4.6	4.8	4.8
protection oxid. and osmotic stress	osmC	Peroxiredoxin OsmC	4.6	3	3.2
pyruvate metabolism	poxB	Pyruvate oxidase	4	6	3.9

To investigate how the expression profile changed over the course of the adaptation experiment we cultivated the adapted strains TUB85 and TUB170 in media analogous to the starting conditions (NMM(19/30/1)), as well as in their adapted media (NMM(13/30/0) and NMM(0/30/0), respectively) to distinguish between effects caused by the removal of the cAAs from the media and those responsible for adaptation to [3,2]Tp. TUB85 and TUB170 samples yielded similar expression profiles under all three culturing conditions, with the only major difference being an expected upregulation of amino acid biosynthesis enzymes in media lacking cAA supplementation (Figures S3 and S4). Thus, while the ancestral

strain reacted to the presence of [3,2]Tp by inducing the RpoS-mediated general stress response, both adapted strains seem to re-regulate this response and restore an unstressed ground state similar to the one observed for the ancestral strain in the medium lacking [3,2]Tp (Table 1). Compared to the ancestral strain cultivated in [3,2]Tp-supplemented media, 105 proteins are up- or downregulated in TUB85 and TUB170 in the presence of [3,2]Tp. Of these, 39 proteins are upregulated (Figure S3) and comprise proteins under the regulation of the Pho-regulon (PhoU, PhoA, PstS, PhoB, PhoS, PstB, UgpB), which are usually responsible for P_i uptake in phosphate limiting environments[60]. As the cultivation media used in this study contain phosphate buffer in non-limiting concentrations, it is not clear why these proteins are upregulated. We hypothesize this could be a response to the mutation in *sgrR* which activates the small RNA gene *sgrS* under glucose-phosphate stress conditions. Induction of the Pho Regulon has been shown to suppress the growth defect of an *Escherichia coli sgrS* mutant[61]. Others have shown a link between the Pho regulon and the synthesis of polyphosphate and guanosine tetraphosphate (ppGpp), a key phosphate-containing molecule involved in the stringent response, which is responsible for the inhibition of RNA synthesis when there is a deficiency in the amino acid availability[62]. The Pho regulon is repressed in *E. coli* mutants that cannot accumulate ppGpp although the actual mechanism relating these two is unclear[63].

Proteins from the leucine biosynthesis pathway are also upregulated in TUB85 and TUB170. This pathway is activated by the transcriptional regulator LeuO[64], which also indirectly inhibits *rpoS* translation by repressing the transcription of the sRNA DsrA.[65] DsrA, in turn, stabilizes *rpoS* mRNA[66] and also activates its translation[67–69]. Perhaps not surprisingly, another prominent protein upregulation in the adapted strains is that of the tryptophan synthase TrpBA, which is responsible for the Trp synthesis from indole and serine[70] (or in this case [3,2]Tpa synthesis from the precursor [3,2]Tp and serine[3,42]). Furthermore, LrhA, for which a mutation in the promoter region was detected, is also upregulated. LrhA is a repressor of *rpoS* translation[31].

The most prevalent feature among the 66 downregulated proteins revolves around the σ^S-mediated general stress response (Figure S4), thus reversing the effect [3,2]Tp elicits at the beginning of the adaptation experiment. One such example is the downregulation of the *rpoS* mRNA-stabilizing protein CspC, which is accompanied by a mutation in its promoter region (Figure 4 A).

DISCUSSION

Here we set out to ask whether we could evolve a lineage with an altered set of canonical amino acids from an existing species, a feat that could pave the way for the creation of artificial diversity and synthetic life. To accomplish this, we performed an ALE with an *E. coli* MG1655 derivative to obtain cells with chemically evolved proteomes in which one amino acid has been completely replaced. We chose tryptophan (Trp) as our canonical target for substitution by the physicochemically similar, non-canonical, analog L-β-(thieno[3,2-*b*]pyrrolyl)alanine ([3,2]Tpa)[3]. Since Trp is rarely an essential catalytic moiety in *E. coli* enzymes, it is well suited for such experiments. The initial medium was NMM supplemented with

indole and amino acids. After 170 transfers, our terminal evolved strain was able to grow in NMM + 25 µM [3,2]Tp without additional supplements.

In the initial phase of the experiment (before transfer 40, generation 250) the concentration of indole was reduced from 1 µM to zero. *E. coli* likely adapted to both the presence of the non-canonical amino acid precursor [3,2]Tp as well as concomitant Trp starvation. We observed three mutations in genes with a direct link to amino acid metabolism (*astB, aroG,* and *leuS*). The first mutation introduced a stop codon leading to a truncation of the N-succinylarginine dihydrolase (AstB). *astB* is a member of the ammonia-producing arginine succinyltransferase (*astCADBE*) operon which is necessary for arginine degradation during nitrogen-limited growth and also contributes to the degradation of other amino acids. As the growth medium used in this study includes a non-limiting ammonia source (Ammonium sulfate), it is essential to reduce the ammonia producing capacity of this operon while maintaining its ability to perform its other functions. Disruption of *astB* has been shown to eliminate succinylarginine dihydrolase activity and prevent arginine utilization without impairing ornithine catabolism[71]. This reduced dependency on arginine as a nitrogen source enables arginine to serve other relevant functions such as the biosynthesis of the most common polyamines, putrescine and spermidine, which are required for optimal growth through their involvement in several physiological processes[72].

In the second phase of the experiment, the number of canonical amino acids was decreased from 19 to 0. Between transfer 40 (generation 250) and transfer 85 (generation 500), the evolving population acquired many mutations in genes related to chemotaxis and flagella synthesis. It is typical for *E. coli* to cease synthesis of chemotaxis and flagella upon entering stationary phase or being exposed to stresses[44–46] while increasing the production of stress-related proteins and drug exporters. In this state, *E. coli* redirects its metabolism from producing primarily growth-related (macro)molecules to using its limited resources for maintenance and survival[23].

In this second phase, the evolved population also acquired a mutation in *rpoS* that lead to RpoS inactivation contributing to improved adaptability. In particular, mutation of the RpoS Asp118 residue abolished RpoS' capacity for repression, diminishing the RpoS-mediated general stress response. *rpoS* mutations observed in a laboratory context should be evaluated critically, as storage and shipment of strains in stab cultures or glycerolized cultures on filter disks are known to select for *rpoS* mutants[73,74]. Nevertheless, the *rpoS* mutation and concomitant downregulation of RpoS protein detected in this study are unlikely to be such storage-related artifacts. Samples taken from the ALE experiment were immediately frozen at -80°C and isolates used for later characterization and analysis were directly derived from these frozen fossils. Indeed, deletion of *rpoS* in the ancestral strain is sufficient for the growth of the un-adapted strain on [3,2]Tp, and *rpoS* mutations have been identified in previous ALE experiments in nutrient-limited conditions, such as in glucose-limited chemostats.[75] A similar experiment was conducted recently with a genome-reduced MG1655 derivative showing growth defects in minimal medium. The deletion of *rpoS* restored the growth rate of the unevolved strain to 80% of the growth rate of the evolved strain[76]. High amounts of RpoS protein typical in stationary phase cells or cells under stress cause competition between RpoS and the vegetative sigma factor RpoD for a limited number of RNA polymerase core subunits. Mutations in *rpoS* therefore likely alleviate the competition in favor of nutrient scavenging through increased expression of RpoD-dependent genes.

In a natural habitat prone to fluctuations in environmental conditions such as temperature, osmolarity, or pH, loss-of-function of RpoS is likely to be disadvantageous. Loss of function of RpoS can also be detrimental under certain stress conditions[77]. For example, *rpoS* mutants evolved under osmotic stress insert IS*10* in the promoter of the *otsBA* operon, thereby partially restoring the wildtype response to osmotic stress by altering the normally RpoS-dependent expression to a RpoS-independent one[78]. In other instances, the appearance of attenuated *rpoS* mutations in long-term stationary phase cultures[79], as well as in glucose-limited chemostats under acidic pH[75] suggests less drastic mutations may be beneficial. These findings, together with the fact that *rpoS* deletions seem to be rare in natural *E. coli* populations[80], indicate that the nature of the trade-off between reduced stress resistance and increased nutrient scavenging is part of a very careful equilibrium.

In the isolates TUB85 and TUB170, we observed the upregulation of proteins from the leucine biosynthesis pathway (*leuLABCD* operon). Transcription of this operon is not only activated by LeuO but also guanosine tetraphosphate (ppGpp) which activates the binding of RNA polymerase to the *leuLABCD* promoter (*leuLp*) under conditions of leucine starvation. Interestingly, the leucine supplement was suspended for these two strains. ppGpp is produced by guanosine-5'-triphosphate,3'-diphosphate phosphatase (GppA) and its gene acquired a T459P mutation in our ALE experiment. Thus, the upregulation of proteins from the leucine biosynthesis pathway could be the hyperactivation of the *leuLABCD* operon transcription by both LeuO and ppGpp. Additionally, aside from activating leucine biosynthesis, LeuO indirectly inhibits *rpoS* translation, thereby deactivating the general stress response.

E. coli mitigated its stress response in favor of cell growth when adapted to a progressive depletion of an essential canonical substrate. The proteomic data revealed that the levels of stress-related proteins subside to levels corresponding to the ancestral strain in relaxing media during the time-course of the adaptation experiment. This discovery is supported by genomic data identifying mutations related to four stress proteins, including two up- or downregulated proteins and the master regulator RpoS itself. In order to neglect the negative effects of [3,2]Tpa incorporation, the adapted strain lowers its protein quality management by reducing its proteolytic activity as evinced by mutations in three proteases.

Our study shows that adaptation to proteome-wide ncAA-insertions is highly pleiotropic. Pleiotropic interactions can result from changes in the metabolic pathways that are networked in the cells; e.g. a change in the concentration of an enzyme or protein could lead to the reconfiguration of intracellular networking. These interactions are an integral part of the adaptation mechanisms associated with proteome-wide insertions and provide the means to attenuate stress responses, modulate folding pathways or other functions. Translational activity is only one of the many biological functions of canonical amino acids; ncAA integration will influence not only proteomes but also cell physiology, signaling, and metabolic and structural processes.

Our findings contribute a step towards understanding possible environmental causes of genetic changes and their relationship to evolution. The knowledge of how cellular metabolism can be regulated in response to alterations of the environment uncovered in this study might help guide engineering efforts and the creation of synthetic cells.

ACKNOWLEDGMENTS

DECLARATION OF INTERESTS

The authors declare no competing interests.

AUTHOR CONTRIBUTIONS

I. T. wrote the final version of the manuscript and provided samples for additional and control experiments. M. H., N. B., and S. O. designed the experimental setup, and S. O. performed the serial dilution experiments and associated analytical/microbiological procedures. A. G., M. H., and S. O. contributed to the writing of the manuscript at various stages. A. G., T. S., and A-R. A. B. performed genetic analyses and constructed pedigrees. L. P. and J. R. performed proteomic analyses, and together with S. O. interpreted the data. C. T.-K. performed chromosomal gene deletions and the corresponding growth assays. N. B. conceived the project, directed the experimental work, and wrote the first version of the manuscript.

DATA AND MATERIALS AVAILABILITY

Materials and methods, as well as additional data, are available in the separate SI file.

REFERENCES

1.   Bacher, J. M. & Ellington, A. D. Selection and characterization of escherichia coli variants capable of growth on an otherwise toxic tryptophan analogue. *J. Bacteriol.* **183**, 5414–5425 (2001).

2.   Tze-fei Wong, J. Membership mutation of the genetic code: Loss of fitness by tryptophan (amino acid analogue/4-fluorotryptophan). *Proc. Nati. Acad. Sci. USA* **80**, 6303–6306 (1983).

3.   Hoesl, M. G. *et al.* Chemical Evolution of a Bacterial Proteome. *Angew. Chemie - Int. Ed.* **54**, 10030–10034 (2015).

4.   Schellhorn, H. E. Function, Evolution, and Composition of the RpoS Regulon in Escherichia coli. *Front. Microbiol.* **11**, 560099 (2020).

5.   Kuthning, A. *et al.* Towards Biocontained Cell Factories: An Evolutionarily Adapted Escherichia coli Strain Produces a New-to-nature Bioactive Lantibiotic Containing Thienopyrrole-Alanine. *Sci. Rep.* **6**, 1–7 (2016).

6.   Crick, F. H. C. The origin of the genetic code. *J. Mol. Biol.* **38**, 367–379 (1968).

7.   Hartman, H. & Smith, T. F. The evolution of the ribosome and the genetic code. *Life* **4**, 227–249 (2014).

8.   Koonin, E. V. & Novozhilov, A. S. Origin and Evolution of the Universal Genetic Code. *Annu. Rev. Genet.* **51**, 45–62 (2017).

9.   Kun, Á. & Radványi, Á. The evolution of the genetic code: Impasses and challenges. *BioSystems* **164**, 217–225 (2018).

10.   Trifonov, E. N. The origin of the genetic code and of the earliest oligopeptides. *Res. Microbiol.* **160**, 481–486 (2009).

11.   Fournier, G. P. & Alm, E. J. Ancestral Reconstruction of a Pre-LUCA Aminoacyl-tRNA Synthetase Ancestor Supports the Late Addition of Trp to the Genetic Code. *J. Mol. Evol.* **80**, 171–185 (2015).

12.   Lajoie, M. J. *et al.* Expand Biological Functions. *Science* **342**, 357–360 (2013).

13.   Kozlowski, L. P. Proteome-pI: proteome isoelectric point database. *Nucleic Acids Res.* **45**, D1112–D1116 (2017).

14.   Yu, A. C. S. *et al.* Mutations enabling displacement of tryptophan by 4-fluorotryptophan as a canonical amino acid of the genetic code. *Genome Biol. Evol.* **6**, 629–641 (2014).

15.   Zuber, U., Drzewiecki, K. & Hecker, M. Putative sigma factor sigI (ykoZ) of Bacillus subtilis is induced by heat shock. *J. Bacteriol.* **183**, 1472–1475 (2001).

16.   Hecker, M., Pané-Farré, J. & Uwe, V. SigB-Dependent General Stress Response in Bacillus subtilis and Related Gram-Positive Bacteria . *Annu. Rev. Microbiol.* **61**, 215–236 (2007).

17.   Agostini, F. *et al.* Multiomics Analysis Provides Insight into the Laboratory Evolution of Escherichia coli toward the Metabolic Usage of Fluorinated Indoles. *ACS Cent. Sci.* **7**, 81–92 (2021).

18.   Blattner, F. R. *et al.* The complete genome sequence of Escherichia coli K-12. *Science* **277**, 1453–1462 (1997).

19.   Jenkins, D. E., Schultz, J. E. & Matin, A. Starvation-induced cross protection against heat or H2O2

challenge in Escherichia coli. *J. Bacteriol.* **170**, 3910–3914 (1988).

20. Jenkins, D. E., Chaisson, S. A. & Matin, A. Starvation-induced cross protection against heat or H2O2 challenge in Escherichia coli. *J. Bacteriol.* **170**, 3910–3914 (1990).

21. Lange, R. & Hengge-Aronis, R. Identification of a central regulator of stationary-phase gene expression in Escherichia coli. *Mol. Microbiol.* **5**, 49–59 (1991).

22. McCann, M. P., Kidwell, J. P. & Matin, A. The Putative sigma factor KatF has a central role in development of starvation-mediated general resistance in Escherichia coli. *J. Bacteriol.* **173**, 4188–4194 (1991).

23. Hengge, R. Stationary-Phase Gene Regulation in Escherichia coli §. *EcoSal Plus* **4**, (2014).

24. Gottesman, S. Trouble is coming: Signaling pathways that regulate general stress responses in bacteria. *J. Biol. Chem.* jbc.REV119.005593 (2019) doi:10.1074/jbc.rev119.005593.

25. Weber, H. *et al.* Genome-Wide Analysis of the General Stress Response Network in. *Society* **187**, 1591–1603 (2005).

26. Battesti, A., Majdalani, N. & Gottesman, S. The RpoS-Mediated General Stress Response in Escherichia coli . *Annu. Rev. Microbiol.* **65**, 189–213 (2010).

27. Cowie, D. B. & Cohen, G. N. Biosynthesis by Escherichia coli of active altered proteins containing selenium instead of sulfur. *BBA - Biochim. Biophys. Acta* **26**, 252–261 (1957).

28. Budisa, N. *et al.* High-level Biosynthetic Substitution of Methionine in Proteins by its Analogs 2-Aminohexanoic Acid, Selenomethionine, Telluromethionine and Ethionine in Escherichia coli - Budisa - 2005 - European Journal of Biochemistry - Wiley Online Library. *Eur. J. Biochem.* **230**, 788–796 (1995).

29. Lenski, R. E., Rose, M. R., Simpson, S. C. & Tadler, S. C. Long-Term Experimental Evolution in Escherichia coli . I . Adaptation and Divergence During 2 , 000 Generations Author (s): Richard E . Lenski , Michael R . Rose , Suzanne C . Simpson and Scott C . Tadler Published by : The University of Chicago Press f. *Am. Nat.* **138**, 1315–1341 (1991).

30. Cohen-Or, I., Shenhar, Y., Biran, D. & Ron, E. Z. CspC regulates rpoS transcript levels and complements hfq deletions. *Res. Microbiol.* **161**, 694–700 (2010).

31. Peterson, C. N., Ruiz, N. & Silhavy, T. J. RpoS proteolysis is regulated by a mechanism that does not require the SprE (RssB) response regulator phosphorylation site. *J. Bacteriol.* **186**, 7403–7410 (2004).

32. Choi, Y. A fast computation of pairwise sequence alignment scores between a protein and a set of single-locus variants of another protein. 414–417 (2012) doi:10.1145/2382936.2382989.

33. Choi, Y., Sims, G. E., Murphy, S., Miller, J. R. & Chan, A. P. Predicting the Functional Effect of Amino Acid Substitutions and Indels. *PLoS One* **7**, (2012).

34.   Mechold, U., Potrykus, K., Murphy, H., Murakami, K. S. & Cashel, M. Differential regulation by ppGpp versus pppGpp in Escherichia coli. *Nucleic Acids Res.* **41**, 6175–6189 (2013).

35.   Grace, E. D. *et al.* DksA and ppGpp Regulate the σS Stress Response by Activating Promoters for the Small RNA DsrA and the Anti-Adapter Protein IraP. *J. Bacteriol.* **200**, 1–12 (2017).

36.   Gentry, D. R., Hernandez, V. J., Nguyen, L. H., Jensen, D. B. & Cashel, M. Synthesis of the stationary-phase sigma factor sigma s is positively regulated by ppGpp. *J. Bacteriol.* **175**, 7982–7989 (1993).

37.   Spira, B. & Ospino, K. Diversity in E. coli (p)ppGpp Levels and Its Consequences. *Front. Microbiol.* **11**, 1759 (2020).

38.   King, T., Ishihama, A., Kori, A. & Ferenci, T. A regulatory trade-off as a source of strain variation in the species Escherichia coli. *J. Bacteriol.* **186**, 5614–5620 (2004).

39.   Barker, M. M., Gaal, T., Josaitis, C. A. & Gourse, R. L. Mechanism of regulation of transcription initiation by ppGpp. I. Effects of ppGpp on transcription initiation in vivo and in vitro. *J. Mol. Biol.* **305**, 673–688 (2001).

40.   Durfee, T., Hansen, A.-M., Zhi, H., Blattner, F. R. & Jin, D. J. Transcription profiling of the stringent response in Escherichia coli. *J. Bacteriol.* **190**, 1084–1096 (2008).

41.   Traxler, M. F. *et al.* The global, ppGpp-mediated stringent response to amino acid starvation in Escherichia coli. *Mol. Microbiol.* **68**, 1128–1148 (2008).

42.   Budisa, N. *et al.* Proteins with β-(thienopyrrolyl)alanines as alternative chromophores and pharmaceutically active amino acids. *Protein Sci.* **10**, 1281–1292 (2001).

43.   Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc. Natl. Acad. Sci.* **97**, 6640–6645 (2000).

44.   Dong, T., Yu, R. & Schellhorn, H. Antagonistic regulation of motility and transcriptome expression by RpoN and RpoS in Escherichia coli. *Mol. Microbiol.* **79**, 375–386 (2011).

45.   Barembruch, C. & Hengge, R. Cellular levels and activity of the flagellar sigma factor FliA of Escherichia coli are controlled by FlgM-modulated proteolysis. *Mol. Microbiol.* **65**, 76–89 (2007).

46.   Pesavento, C. *et al.* Supplementary Material Inverse regulatory coordination of motility and adhesion in Escherichia coli. **2**, 2434–2446 (2007).

47.   Arnqvist, A., Olsen, A. & Normark, S. sigma-S -dependent growth-phase induction of the csgBA promoter in Escherichia coii can be achieved in vivo by sigma-70 in the absence of the nucieoid-associated protein H-NS. *Mol. Microbiol.* **13**, 1021–1032 (1994).

48.   Barth, M., Marschall, C., Muffler, A., Fischer, D. & Hengge-Aronis, R. Role for the histone-like protein H-NS in growth phase-dependent and osmotic regulation of σ(S) and many σ(S)-dependent genes in Escherichia coli. *J. Bacteriol.* **177**, 3455–3464 (1995).

49.   Colland, F. sigmafactor selectivity of Escherichia coli RNA polymerase: role for CRP, IHF and Lrp transcription factors. *EMBO J.* **19**, 3028–3037 (2000).

50.   Grainger, D. C., Goldberg, M. D., Lee, D. J. & Busby, S. J. W. Selective repression by Fis and H-NS at the Escherichia coli dps promoter. *Mol. Microbiol.* **68**, 1366–1377 (2008).

51.   Weber, H., Pesavento, C., Possling, A., Tischendorf, G. & Hengge, R. Cyclic-di-GMP-mediated signalling within the σS network of Escherichia coli. *Mol. Microbiol.* **62**, 1014–1034 (2006).

52.   Shin, M. *et al.* DNA looping-mediated repression by histone-like protein H-NS: specific requirement of Esigma70 as a cofactor for looping. *Genes Dev.* **19**, 2388–2398 (2005).

53.   Waterman, S. R. & Small, P. L. C. Transcriptional expression of *Escherichia coli* glutamate-dependent acid resistance genes *gadA* and *gadBC* in an *hns rpoS* mutant. *J. Bacteriol.* **185**, 4644–4647 (2003).

54.   Yamashino, T., Ueguchi, C. & Mizuno, T. Quantitative control of the stationary phase-specific sigma factor, sigma S, in Escherichia coli: involvement of the nucleoid protein H-NS. *EMBO J.* **14**, 594–602 (1995).

55.   Battesti, A., Tsegaye, Y. M., Packer, D. G., Majdalani, N. & Gottesman, S. H-NS Regulation of IraD and IraM antiadaptors for control of RpoS degradation. *J. Bacteriol.* **194**, 2470–2478 (2012).

56.   Foster, J. W. Escherichia coli acid resistance: Tales of an amateur acidophile. *Nat. Rev. Microbiol.* **2**, 898–907 (2004).

57.   Tramonti, A., De Canio, M. & De Biase, D. GadX/GadW-dependent regulation of the Escherichia coli acid fitness island: Transcriptional control at the gadY-gadW divergent promoters and identification of four novel 42 bp GadX/GadW-specific binding sites. *Mol. Microbiol.* **70**, 965–982 (2008).

58.   Capitani, G. *et al.* Crystal structure and functional analysis of Escherichia coli glutamate decarboxylase. *EMBO J.* **22**, 4027–4037 (2003).

59.   Hirakawa, H., Inazumi, Y., Masaki, T., Hirata, T. & Yamaguchi, A. Indole induces the expression of multidrug exporter genes in Escherichia coli. *Mol. Microbiol.* **55**, 1113–1126 (2005).

60.   Santos-Beneit, F. The Pho regulon: A huge regulatory network in bacteria. *Front. Microbiol.* **6**, 1–13 (2015).

61.   Richards, G. R. & Vanderpool, C. K. Induction of the Pho regulon suppresses the growth defect of an Escherichia coli sgrS mutant, connecting phosphate metabolism to the glucose-phosphate stress response. *J. Bacteriol.* **194**, 2520–2530 (2012).

62.   Chatterji, D. & Ojha, A. K. Revisiting the stringent response, ppGpp and starvation signaling. *Curr. Opin. Microbiol.* **4**, 160–165 (2001).

63.   Loewen, P. C., Hu, B., Strutinsky, J. & Sparling, R. Regulation in the rpoS regulon of Escherichia coli. *Can. J. Microbiol.* **44**, 707–717 (1998).

64. Chen, C. C., Fang, M., Majumder, A. & Wu, H. Y. A 72-Base Pair AT-rich DNA Sequence Element Functions as a Bacterial Gene Silencer. *J. Biol. Chem.* **276**, 9478–9485 (2001).

65. Klauck, E., Böhringer, J. & Hengge-Aronis, R. The LysR-like regulator LeuO in Escherichia coli is involved in the translational regulation of rpoS by affecting the expression of the small regulatory DsrA-RNA . *Mol. Microbiol.* **25**, 559–569 (1997).

66. McCullen, C. A., Benhammou, J. N., Majdalani, N. & Gottesman, S. Mechanism of positive regulation by DsrA and RprA small noncoding RNAs: Pairing increases translation and protects rpoS mRNA from degradation. *J. Bacteriol.* **192**, 5559–5571 (2010).

67. Sledjeski, D. D., Gupta, A. & Gottesman, S. The small RNA, DsrA, is essential for the low temperature expression of RpoS during exponential growth in Escherichia coli. *EMBO J.* **15**, 3993–4000 (1996).

68. Majdalani, N., Cunning, C., Sledjeski, D., Elliott, T. & Gottesman, S. DsrA RNA regulates translation of RpoS message by an anti-antisense mechanism, independent of its action as an antisilencer of transcription. *Proc. Natl. Acad. Sci.* **95**, 12462–12467 (1998).

69. Lease, R. A., Cusick, M. E. & Belfort, M. Riboregulation in Escherichia coli: DsrA RNA acts by RNA:RNA interactions at multiple loci. *Proc. Natl. Acad. Sci.* **95**, 12456–12461 (1998).

70. Lane, A. N. & Kirschner, K. Mechanism of the Physiological Reaction Catalyzed by Tryptophan Synthase from Escherichia coli. *Biochemistry* **30**, 479–484 (1991).

71. Schneider, B. L., Kiupakis, A. K. & Reitzer, L. J. Arginine catabolism and the arginine succinyltransferase pathway in Escherichia coli. *J. Bacteriol.* **180**, 4278–4286 (1998).

72. Charlier, D. & Glansdorff, N. Biosynthesis of Arginine and Polyamines. *EcoSal Plus* **1**, (2004).

73. Sutton, A., Buencamino, R. & Eisenstark, A. rpoS mutants in archival cultures of Salmonella enterica serovar typhimurium. *J. Bacteriol.* **182**, 4375–4379 (2000).

74. Spira, B., De Almeida Toledo, R., Maharjan, R. P. & Ferenci, T. The uncertain consequences of transferring bacterial strains between laboratories - RpoS instability as an example. *BMC Microbiol.* **11**, (2011).

75. Notley-mcrobb, L., King, T. & Ferenci, T. Mutations and Loss of General Stress Resistance in. *Microbiology* **184**, 806–811 (2002).

76. Choe, D. *et al.* Adaptive laboratory evolution of a genome-reduced Escherichia coli. *Nat. Commun.* **10**, (2019).

77. Ferenci, T. The spread of a beneficial mutation in experimental bacterial populations: The influence of the environment and genotype on the fixation of rpoS mutations. *Heredity (Edinb).* **100**, 446–452 (2008).

78. Stoebel, D. M., Hokamp, K., Last, M. S. & Dorman, C. J. Compensatory evolution of gene regulation in response to stress by Escherichia coli lacking RpoS. *PLoS Genet.* **5**, 1–9 (2009).

79.   Zambrano, M. M., Siegele, D. A., Almirón, M., Tormo, A. & Kolter, R. Microbial competition: Escherichia coli mutants that take over stationary phase cultures. *Science (80-.).* **259**, 1757–1760 (1993).

80.   Snyder, E., Gordon, D. M. & Stoebel, D. M. Escherichia coli lacking RpoS are rare in natural populations of non-pathogens. *G3 Genes, Genomes, Genet.* **2**, 1341–1344 (2012).