Received: 2 August 2021 | Revised: 8 October 2021 | Accepted: 12 October 2021

DOI: 10.1002/pro.4206

METHODS AND APPLICATIONS

Enhancing protein perdeuteration by experimental evolution of *Escherichia coli* K-12 for rapid growth in deuterium-based media

Vinardas Kelpšas | **Claes von Wachenfeldt**

The Microbiology Group, Department of Biology, Lund University, Lund, Sweden

Correspondence
Claes von Wachenfeldt, The Microbiology Group, Department of Biology, Lund University, Sölvegatan 35, Lund SE-223 62, Sweden.
Email: claes.von_wachenfeldt@biol.lu.se

Funding information
Swedish Research Council, Grant/Award Number: 2019-05578_3; Jörgen Lindström's Foundation; Erik Philip-Sörensens stiftelse; Carl Tryggers Stiftelse för Vetenskaplig Forskning

Abstract

Deuterium is a natural low abundance stable hydrogen isotope that in high concentrations negatively affects growth of cells. Here, we have studied growth of *Escherichia coli* MG1655, a wild-type laboratory strain of *E. coli* K-12, in deuterated glycerol minimal medium. The growth rate and final biomass in deuterated medium is substantially reduced compared to cells grown in ordinary medium. By using a multi-generation adaptive laboratory evolution-based approach, we have isolated strains that show increased fitness in deuterium-based growth media. Whole-genome sequencing identified the genomic changes in the obtained strains and show that there are multiple routes to genetic adaptation to growth in deuterium-based media. By screening a collection of single-gene knockouts of nonessential genes, no specific gene was found to be essential for growth in deuterated minimal medium. Deuteration of proteins is of importance for NMR spectroscopy, neutron protein crystallography, neutron reflectometry, and small angle neutron scattering. The laboratory evolved strains, with substantially improved growth rate, were adapted for recombinant protein production by T7 RNA polymerase overexpression systems and shown to be suitable for efficient production of perdeuterated soluble and membrane proteins for structural biology applications.

KEYWORDS

adaptive experimental evolution, deuteration, isotope labeling, neutron crystallography

1 | INTRODUCTION

The discovery of deuterium in 1931 by Urey et al. stimulated numerous investigations of the biological effects of this stable isotope of hydrogen.[1,2] Deuterium (^{2}H or D) consists of one proton, one neutron, and one electron, in contrast to ordinary hydrogen (^{1}H or protium) which lacks the neutron. Because of the neutron present in the nucleus, deuterium has roughly twice the atomic weight of protium (2.014101 vs. 1.007782).[3] The natural abundance of deuterium in the environment is only about 0.015% of hydrogen atoms in ocean water.[4] Due to the higher mass of ^{2}H compared to ^{1}H, deuterium enriched water is called heavy water (^2H_2O or D_2O). Protium and

deuterium, as well as water and heavy water, have close but clearly distinguishable physical properties. For example, heavy water has higher melting temperature (3.8°C),[5] higher boiling point (101.4°C),[5] and higher viscosity.[6] Deuterium isotope effects are observed at the single enzyme level (kinetic isotope effects) where the rate of cleavage of covalent bonds are affected as well as in the cumulative effect on cellular reactions leading to growth rate reduction or complete inhibition of growth. These isotope effects are due to both solvent (D_2O) and macromolecular isotope effect. Bacteria, yeast, and single cell algae can be grown in fully deuterated media, whereas there is no report on the growth of insect or mammalian cells in fully deuterated medium.[7]

Structural biology of proteins probed by biophysical tools such as nuclear magnetic resonance spectroscopy, small angle scattering, and neutron protein crystallography (NPC) are greatly enhanced by deuterium labeling either of the solvent or of the proteins themselves, or both. NPC is an attractive method for locating hydrogen atoms in protein structures due to that the coherent scattering lengths of protium and deuterium for neutrons are similar in magnitude to those of carbon, nitrogen, and oxygen.[8–10] However, the scattering length of protium is negative while that of deuterium, carbon, nitrogen, and oxygen are positive. In addition, the incoherent scattering length of deuterium is much smaller compared to protium, reducing the isotropic background intensity.[11] Therefore, NPC benefit in signal and reduction in noise from perdeuteration, where all solvent and protein protium atoms (exchangeable- and non-exchangeable) are replaced by deuterium.[11] Protein perdeuteration is typically carried out by growing the *Escherichia coli* or yeast strain producing the recombinant protein in heavy water-based medium supplied with a perdeuterated carbon source such as glycerol. The recombinant protein is then purified by standard methods using protiated reagents and solvents. Before or after crystallization, the H_2O solvent is replaced by D_2O, back-exchanging the labile protons with deuterons.

While *E. coli* can grow in heavy water based minimal media, the growth rate and biomass yield are significantly reduced.[12–14] The precise mechanisms behind this are not known, however the deuterium kinetic isotope effect is one important factor. A recent study showed that growth in deuterated growth medium results in global changes in the *E. coli* proteome however with no apparent stress response induction.[14] It has been reported that growth of *E. coli* in fully deuterated minimal media can be improved if the cells are pre-adapted to gradually increasing deuterium content; however, this adaptation is often lost when cells are grown in ordinary medium.[15] In a recent study, we showed, by using genome-wide random mutagenesis, that *E. coli* strains with increased tolerance to growth in deuterated minimal medium can be isolated.[13]

Experimental evolution has been successfully utilized in several studies to generate *E. coli* strains with improved traits.[16,17] The conceptually simple procedure involves continued propagation of a bacterial population under a selective pressure. Faster growing clones arise from random mutations during replication of the chromosome. When the ratio of the growth rates of the evolved and parental strain as they compete with one another increase, the frequency of clones with improved fitness in the population will rise. Thus, growth rate equates directly to fitness. This method is particular well suited for bacteria such as *E. coli* that can have a large population size and a short generation time. When a strong selective pressure can be used experimental evolution can produce strains with desired traits in a relatively short time (see, e.g., References 16,18,19).

Here we report evolutionary adaptation of *E. coli* MG1655 K-12 cells to growth in deuterated minimal medium with perdeuterated glycerol as the sole carbon and energy source. The genetic changes in adapted strains were identified by whole genome sequencing. We show that selected evolved strains can be used for production of recombinant perdeuterated soluble and membrane proteins. In addition, we show that further modifications such as deletion of the protease encoding genes *lon* and *ompT*, does not improve protein yield, but can instead compromise growth during recombinant protein production. Furthermore, we provide results of screening 3,985 single-gene deletion mutants from the Keio collection, to provide a basis for future engineering of strains for bio-deuteration.

2 | RESULTS AND DISCUSSION

2.1 | Experimental evolution via serial passage of *E. coli* cultures in perdeuterated growth medium

E. coli K-12 derivatives are the most widely used strains by the biotechnology industry. The *E. coli* K-12 strain MG1655 was used in this study because its genome has been completely sequenced and is well annotated.[20] The growth rate in deuterated minimal glycerol medium (DD-M9) reduced approximately three times compared to growth in minimal glycerol medium (H-M9).[13] Inspection of cells by phase-contrast microscopy showed that there are no apparent differences in cell size or other morphological changes in DD-M9 grown cells as compared to H-M9 grown cells (Figure 1, Figure S1). The

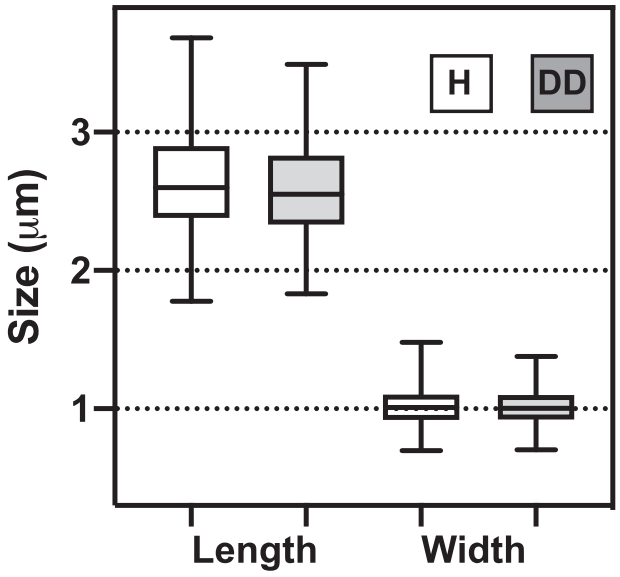

FIGURE 1 Effect of deuterium on cell size of *Escherichia coli* K-12 MG1655. Cell size (length and width) of *E. coli* MG1655 grown in H-M9 and DD-M9. Cells were analyzed at $OD_{600} = 0.5$. Hundred cells were measured using phase contrast microscopy for each sample. The cell size data were plotted using box-and-whisker chart and the bottom and top of the box are the first and third quartiles, the line inside the box is the median. The whiskers go down to the minimum and up to the maximum value

OD_{600} was proportional to the final dry weight of cells (-Table S1). The biomass yield in DD-M9 was reduced by approximately 40% compared to cells grown in H-M9. Thus, deuterium affects both the growth rate and final biomass yield in the used medium.

To evolve strains with increased growth rate in a deuterated growth medium three independent populations of *E. coli* MG1655 from the same seed culture were serially passaged in DD-M9 medium until the culture growth rate increased. Initially, the cultures were passaged every second day as illustrated in Figure 2. At the 17th subculturing, corresponding to approximately 90 generations, we observed that cultures took less time to become visually turbid. From this point onwards, the cultures were diluted daily. The passaging frequency made sure that the cultures were mainly under selective pressure for exponential growth. During this part of the experiment, the cultures were growing at similar growth rates, as indicated by that the OD_{600} value before each subculturing was similar for the replicates. All tested lineages performed better than the parental strain in terms of growth rate in DD-M9 after an average of 418 generations (Figure 3a). In order to examine that this improvement is not only due to phenotypic adaptation, cultures were streaked onto LA-plates to isolate single colonies. Three

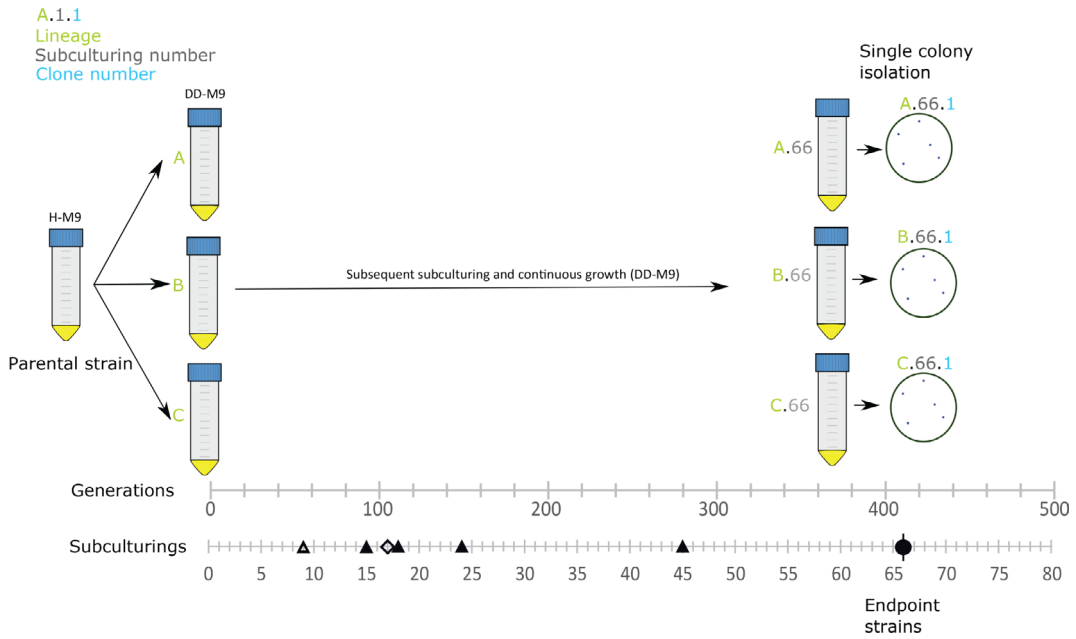

FIGURE 2 Outline of the experimental evolution via serial passage procedure. The *Escherichia coli* parental strain culture was inoculated into three different TPP TubeSpin bioreactor tubes (lineage A, B, and C) containing DD-M9 growth medium and incubated at 37°C, 200 rpm. The indicated generation number is an estimation and does not represent an exact number. Subculturing was done every second day (48 hr) for 17 subculturings (indicated by diamond) and after that once a day. Filled triangles indicates time points at which the *glpK* gene was analyzed and found to contain mutations. The triangle at the 9th subculturing. indicate that no mutation in *glpK* was detected at this sampling point

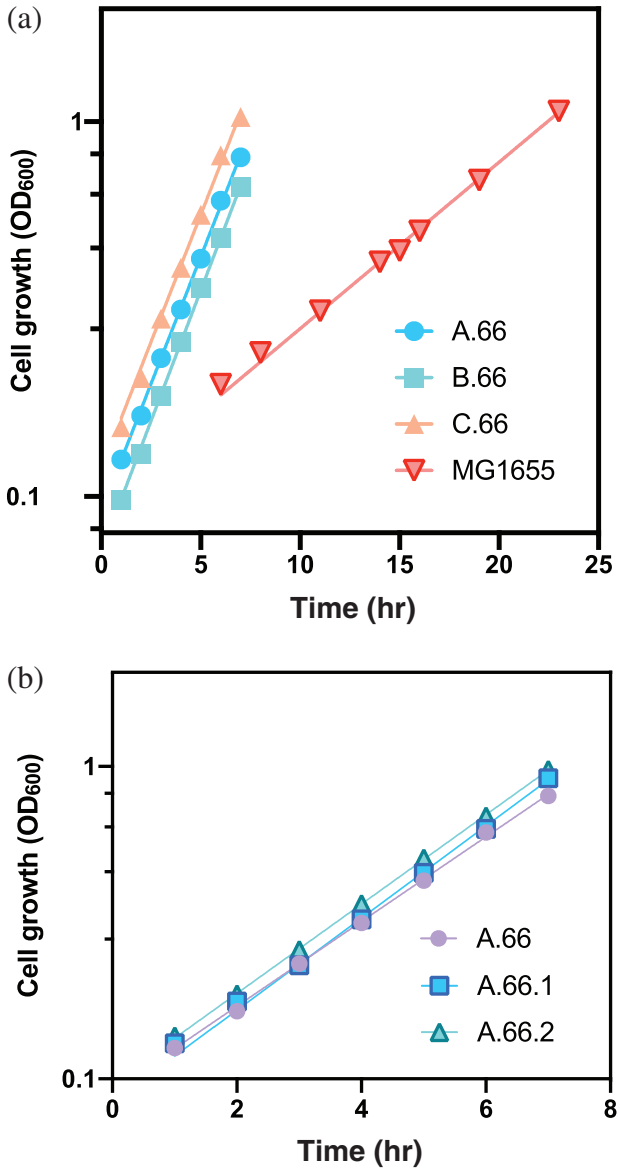

FIGURE 3 Growth of the evolved lineages compared to the parental strain in DD-M9. (a) Exponential growth in DD-M9 of lineages A (circles), B (squares), and C (triangles point up), compared to the parental MG1655 strain (triangles point down). The specific growth rates are: 0.31 ± 0.01 hr^{-1} (A.66), 0.32 ± 0.01 hr^{-1} (B.66) 0.31 ± 0.02 hr^{-1} (C.66), and 0.102 ± 0.004 hr^{-1} (MG1655). (b) Growth curves of the heterogenous culture of lineage A at the 66th (circles) subculturing and of two isolates (A.66.1, squares and A.66.2, triangle) from the lineage A culture. The specific growth rates are: 0.34 ± 0.03 hr^{-1} (A.66.1), 0.33 ± 0.01 hr^{-1} (A.66.2). Cells were grown in DD-M9. For each strain, three biological replicates were analyzed. The mean is plotted and the error bars represents the *SD*. Error bars are not shown when they are shorter than the size of the symbol. The lines represent the fit to exponential growth. The R^2 values for both fits are ≥.99

single colonies from each lineage were isolated, re-streaked and used for growth rate measurements and whole-genome sequencing. The sampled clones grew indistinguishable to the lineage, it was isolated from indicating that the isolated clones were representative of the improved growth phenotype of the culture population (Figure 3b).

2.2 | Genome analysis of evolved strains

As shown above the improved growth rates of the evolved strains appeared to be stably maintained. Next, we set out to identify mutations by whole genome sequencing of the clones isolated from the three parallel evolutionary experiments. Two clones from each of the three lineages (named A, B, and C) were sequenced at the 66th subculturing (after approximately 412, 426, 420 generations, respectively, and referred to as lineage A.66, B.66, and C.66). The A and C lineages had a similar number of mutations, seven and six, respectively (Table 1). However, the B lineage contained a very high number of mutations; 282 and 417 in the two sequenced isolates. Of those mutations six were also present in the analyzed clones from the A and C lineages (Table 1). The two clones of the B lineage appeared to be so-called hypermutator strains. These strains had mutations in genes associated with DNA replication (e.g., *dnaQ*, *dnaA*, and *dnaE*) and repair (*uvrD*) that could explain the high number of observed mutations (Table S2). Thus, it is likely that all strains from the B lineage have increased mutation rate, which is undesired for recombinant protein production.

Improved fitness can be explained by an accumulation of fixed beneficial mutations in the genome. Since we used three different lineages, we assume that modifications in coding regions of the same gene in these lineages will pinpoint mutations which are selected due to specific growth media composition or the presence of deuterium. The observed mutations in the *glpK*, *ilvG*, *rpoS*, and *trkH* genes were common to at least two lineages. Surprisingly, no mutation was common to all analyzed strains. Three lineages showed mutations in genes (*rpoB* and *rpoC)* encoding subunits of the DNA-dependent RNA polymerase. The *rpoB and rpoC* mutations are likely to have global effects on gene expression. The mutations are listed in Table 1 and further discussed below.

2.3 | Mutations in *glpK* encoding glycerol kinase

All but one of the sequenced strains acquired missense mutations in the *glpK* gene encoding glycerol kinase, a key enzyme in glycerol metabolism (Table 1). Four different single base pair substitutions in *glpK* were found.

TABLE 1 Mutations identified in selected populations after adaptation in perdeuterated minimal medium[a]

		A.66.		B.66.		C.66.				
Position[b]	Gene	1	2	1	2	1	2	Coding region change[c]	Protein change[d]	Description
86,887	*ilvI*							1258C>T	Leu420Phe	Valine biosynthesis
442,487	*thiI*							939G>A	No change	tRNA uridine 4-sulfurtransferase
569,566	*ybcK*							665 T>A	Val222Glu	Putative recombinase
1,935,180	*zwf*							1135G>A	Asp379Asn	Glucose-6-P-dehydrogenase
1,935,533	*zwf*							782G>T	Arg261Leu	Glucose-6-P-dehydrogenase
2,416,713	*pta*							1967A>G	Asp656Gly	Phosphate acetyltransferase
2,867,401	*rpoS*							151 del A	Thr51fs	Sigma S
2,877,560	*rpoS*							992A>C	Ser331*	Sigma S
2,912,771	*relA*							881G>A	Gly294Glu	GDP/GTP pyrophosphokinase
3,078,276	*cmtA*							584A>G	Asp195Gly	Putative mannitol permease
3,740,542	*avtA*							838G>A	Ala280Thr	Valine-pyruvate aminotransferase
3,941,830	*rrsC*							-	-	16S ribosomal RNA
3,951,533	*ilvG*							974 T>C	Leu325Ser	Isoleucine biosynthesis
3,951,541	*ilvG*							982 del T	fs	Isoleucine biosynthesis
3,951,542	*ilvG*							983 del G	fs	Isoleucine biosynthesis
3,966,721	*rho*							305G>A	Arg102His	Transcription termination
4,033,383	*trkH*							239 T>A	Leu80Gln	K^+ transporter
4,033,611	*trkH*							467G>A	Gly156Asp	K^+ transporter
4,116,309	*glpK*							914G>C	Gly305Ala	Glycerol kinase
4,116,351	*glpK*							692G>A	Gly231Asp	Glycerol kinase
4,116,703	*glpK*							520A>G	Met174Val	Glycerol kinase
4,117,063	*glpK*							160 T>A	Trp54Arg	Glycerol kinase
4,181,336	*rpoB*							92A>G	Gln31Arg	RNAP subunit β
4,186,099	*rpoC*							749–750 ins TCCG CTGGT	LevValPro(252–254) ins	RNAP subunit β′
4,272,883	*uvrA*							989 T>G	Phe330Cys	Excision nuclease

[a]Two clones from each of three populations were selected and sequenced. The total number of identified mutations in the hypermutator clones B.66.1 and B.66.2 were 281 and 416, respectively. Genes with mutations that are shared between at least two clones are color coded and those that are only present in one clone are shown in grey.
[b]Nucleotide position in the reference sequence Genbank entry: U00096.3.
[c]Deletion (del), insertion (ins).
[d]Non-sense mutation (*). Frame shift (fs).

Three of these mutations have been reported before and results in an increased glycerol kinase activity. Some of the previously studied GlpK variants result in increased growth rate in glycerol minimal medium.[21–23] In the A and C lineages the two analyzed clones had different mutations in *glpK*, in contrast the same mutation was present in the analyzed clones of the hypermutator strains (Table 1).

Previous laboratory evolution experiments have revealed that specific point mutations in the *glpK* gene promote growth by improving glycerol utilization in regular glycerol minimal medium.[19] It is thus likely that the evolved strains with GlpK variants with improved glycerol utilization also show improved growth in H-M9. To test this, we conducted a growth experiment in H-M9 with two of the evolved strains (A.66.1 and A.66.2) and compared it to the parental MG1655 strain. The evolved strain A.66.1 showed a 39% increase in growth rate growth ($\mu = 0.486 \pm 0.001\ hr^{-1}$ [$n = 2$]) compared to the wild-type strain ($\mu = 0.349 \pm 0.006\ hr^{-1}$ [$n = 2$]). Strain

A.66.2 showed an improvement of 74% in growth rate ($\mu = 0.607 \pm 0.02\ hr^{-1}$ [$n = 2$]) compared to the wild type (Figure S2). Next, we repeated the growth experiment and switched the carbon source to glucose. In this case, the difference in growth rate between the evolved strain (A.66.1) and the wild type was less than 10% ($\mu = 0.580 \pm 0.009\ hr^{-1}$ [$n = 2$] vs. $\mu = 0.628 \pm 0.013\ hr^{-1}$ [$n = 2$]; Figure S3). This suggests that mutations affecting glycerol utilization are causal to the improved growth also in the perdeuterated DD-M9 medium.

2.4 | Structure–function predictions of glycerol kinase variants

In order to understand how the adaptive GlpK mutants improve fitness during growth on minimal glycerol medium the detected amino acid substitutions were mapped to the X-ray crystal structure of GlpK (PDB entry 1GLF). It was observed that most substitutions are located at or in proximity to the tetramer interface in GlpK (Figure S4). Such substitutions have been found to affect tetramerization of GlpK, which is required for the allosteric inhibition by fructose-1,6-bisphosphate (FBP).[24]

Since most GlpK variations found here were located at the tetramer interface, it is likely that the increased fitness of those mutants is due to abolished FBP inhibition. We performed Sanger sequencing on colonies isolated from intermediate evolutionary steps, specifically targeting the *glpK* gene. In addition to the detected mutations in the experimental evolution end-point isolates, we found four additional missense mutations in *glpK*, all predicted to alleviate the inhibition of GlpK by FBP (Table 2).

2.5 | The *ilvG* pseudogene is restored in evolved strains

The *E. coli* K-12 MG1655 strain has a frameshift mutation that introduces a premature stop codon (UGA) in the *ilvG* gene making it non-functional.[26] We observed mutations in the *ilvG* gene in all lineages. A.66.1, C.66.1, and C.66.2 have single base pair deletions of the T or the G of the premature stop codon (TGA) that restores the *ilvG* reading frame. The hypermutator strains, B.66.1 and B.66.2, have a single base pair substitution eight bases upstream of the premature stop codon. This missense mutation

TABLE 2 Observed polymorphism within *glpK* in the evolved strains

Strain[a]	Coding region change	Protein change	Description
A.18.1	685A>C	N229T	Homotetramer interface
A.66.1	914G>C	G305A	Decreased allosteric regulation by FBP.[b]
A.24.4	1061C>T	A354V	Close to G305
A.15.2 A.18.2 A.18.3 A.24.5 A.45.1 A.45.2 A.66.2	692G>A	G231D	Homotetramer interface. Increased activity and decreased allosteric regulation by FBP.[c]
B.15.3 B.18.1 B.18.2 B.18.3 B.24.2 B.24.1 B.45.1 B.45.2B.66.1 B.66.2	520A>G	M174V	Close to the homotetramer interface
C.18.2	519A>C	K173N	Homotetramer interface
C.45.1	689 T>C	I230T	Homotetramer interface
C.45.2 C.66.2	160 T>C	W54R	Homotetramer interface

[a]Sampling for sequencing was done at the 9th, 15th,18th, 24th, 45th, and 66th subculturing. No mutation in *glpK* was observed at the 9th subculturing.
[b]A G305S substitution in GlpK alleviating the allosteric inhibition by fructose-1,6-bisphosphate (FBP) was described in References 23,25.
[c]This substitution was described in Reference 21.

results in a substitution of Leu325 for Ser but does not restore the reading frame. Biosynthesis of the branched-chain amino acids isoleucine, leucine, and valine are closely related. Three different acetohydroxy acid synthetase isoenzymes are present in *E. coli*.[27] These enzymes participate in the branched-chain amino acid biosynthetic pathway. Isoenzyme I and III are feedback inhibited by valine, whereas isoenzyme II is not. A functional *ilvG* gene encodes the large subunit of acetohydroxy acid synthetase II. *E. coli* K-12 strains cannot form an active isoenzyme II and are therefore sensitive to growth inhibition by valine renders *E. coli* K-12 auxotrophic for isoleucine.[27] The evolved strains were tested for valine sensitivity by assaying growth in minimal medium supplemented with different concentrations of valine (0–1 mg/ml). The parental strain and the evolved strains without mutations in *ilvG* did not grow in the presence of added valine. In contrast, strains carrying the single base pair deletion that restored the *ilvG* reading frame were able to grow in the presence of 1 mg/ml valine (Table 3). The hypermutator strains, B.66.1 and B.66.2, showed intermediate sensitivity to valine (Table 3). The A66 and C66 isolates have the *ilvG* frames restored due to loss of one base of the stop codon. The hypermutator strains (B66.1 and B.66.2) may have the *ilvG* restored by a missense mutation before the stop codon resulting in a second reading initiated from an AUG codon upstream of the stop codon (Figure S5). The putative initiation codon is proceeded by a sequence that may function as a ribosome binding sequence. This reading frame could encode the C-terminal part of IlvG which may partially rescue the non-functional truncated IlvG (Figure S5).

Previous studies found that *ilvG* disruption causes polar effect on downstream genes of the *ilvGMEDA* operon that encodes enzymes required for isoleucine and valine biosynthesis.[28] Restoring *ilvG* could alleviate the polar effect, increasing transcription of downstream genes. This suggests that either in deuterated conditions downstream genes is important in producing branched chain amino acids, or that valine production and degradation is unbalanced, leading to an accumulation of valine that inhibits isoleucine biosynthesis. It has been observed that restoring *ilvG* abolish oscillations in growth in high-density fermenter cultures.[29] Such oscillations might be enhanced in deuterium-based medium increasing the selective pressure for restoring *ilvG*. Moreover, valine biosynthesis and degradation could be differently affected by the deuterium isotope effect, and thus amplifying these oscillations.

TABLE 3 Sensitivity to valine in different strains

	Growth with valine (mg/ml)[a]			
Strain	**0**	**0.1**	**0.5**	**1**
MG1655	+	−	−	−
A.66.1	+	+	+	+
A.66.2	+	−	−	−
B.66.1	+	+	+	−
B.66.2	+	+	−	−
C.66.1	+	+	+	+
C.66.2	+	+	+	+

[a]Growth (+), no growth (−).

2.6 | Mutations in *rpoB* and *rpoC* RNA polymerase β and β′ subunits, respectively

The DNA-dependent RNA polymerase (RNAP) is central to transcription and transcriptional regulation and mutations affecting RNAP can arise due to a range of different selective conditions. We observed, in common with previous long-term adaptive laboratory evolution studies, mutations in the genes encoding the RNAP catalytic subunits β and β′ (*rpoB* and *rpoC*)[30] (Table 1). In this work, we found a missense mutation in *rpoB* which was not reported previously. In *rpoC*, a duplication between position 749 and 750, resulting in duplication of amino acid residues LVP at 252–254 positions in RpoC. Several mutations in *rpoC* have been found during adaptive evolution to glycerol minimal media.[19] However, these mutations, substitution at position 2,249 and deletion at position 3,132...3158,[19] are far away from the one obtained in this study.

Mutations in *rpoB* that leads to resistance to rifampicin also lead to altered expression of several genes.[31] The RpoB substitution (Gln31Arg) is in close proximity to the rifampicin-binding site. Since the mode of action of rifampicin is to directly block the path of the elongating RNA transcript, these mutations are likely to affect the transcription mechanism itself.

2.7 | Mutations in *rpoS* encoding RNA polymerase sigma factor (σ^{38})

The *rpoS* gene encode an alternative sigma factor (RpoS or σ^S or σ^{38}) that mediates the general stress response and is mainly active in the stationary phase.[32] We observed that in two instances *rpoS* was mutated. In strain A.66.1 a deletion of adenine at position 151 in *rpoS* results in a frameshift mutation which in turn leads to premature termination. It is expected that this mutation would lead to a loss of RpoS function. The C.66.1 strain has a mutation in the stop codon which results in a

predicted 41 amino acid residue extension of RpoS. Both variants are expected to negatively affect RpoS function. To experimentally test this, we made use of that transcription of the catalase encoded by the *katE* gene is RpoS dependent.[33] To assay for catalase activity hydrogen peroxide was applied on bacterial patches and the presence or absence of vigorous bubbling (oxygen) was recorded. The wild-type strain was used as the reference, and a strain deleted for *rpoS* was used as a negative control. As expected, the wild type gave a positive catalase reaction and the *rpoS* null mutant showed almost no reaction with hydrogen peroxide. The two strains, A.66.1 and C.66.2 with mutations in *rpoS* were both catalase negative most likely due to a non-functional RpoS (Table 4).

Mutations in *rpoS* were not reported in previous adaptive laboratory evolution studies using glycerol minimal media. However, *rpoS* mutations have been detected during selection on succinate, and during nitrogen and glucose limiting conditions.[34–36] A large number of genes in *E. coli* are positively controlled by RpoS and several genes are negatively controlled.[32] The RpoS variants found here are both non-functional according to the catalase assay (Table 4). Thus, lack of RpoS will reduce expression of hundreds of genes and by so liberate resources for improved growth in deuterated media.

2.8 | Mutations in *trkH* encoding a potassium ion transporter

The *trkH* gene encodes a low affinity high rate potassium ion transporter.[37] This membrane protein interacts with the cytoplasmic protein TrkA.[38] The TrkH/TrkA complex requires the proton-motive force and ATP to transport K^+ and H^+ ions.[39] We observed three missense mutations in the *trkH* gene in the evolved strains. Since there is no structural information on *E. coli* TrkH we aligned it with the homologue from *Vibrio parahaemolyticus* and mapped the positions of the observed amino acid changes on the TrkH X-ray structure (PDB entry 3PJZ)[40] (Figure S6). The Leu80Gln substitution is predicted to be located in an α-helix close to the pore helix which is responsible for translocating ions through the membrane. The Gly156Asp mutation is located in a loop region that is in close proximity to the interaction face with TrkA.[41] How these variants may affect K^+ uptake is not apparent. However, as mutations in *trkH* were observed in several of the evolved strains it is likely that they have a phenotypic effect on K^+ influx rates.

TABLE 4 Catalase activity in different strains

Strain	Catalase reaction[a]	*rpoS* genotype[b]
MG1655	+	Wild type
A.66.1	−	Thr51fs
A.66.2	+	Wild type
B.66.1	+	Wild type
B.66.2	+	Wild type
C.66.1	+	Wild type
C.66.2	−	*331Ser
MG1655(DE3) Δ*rpoS*	−	*rpoS*::kan

Note: Each test was carried out on three replicates.
[a](+) Vigorous reaction, (−) faint reaction.
[b]frameshift (fs), Stop codon (*).

2.9 | Perdeuterated protein production in evolved strains

To investigate if the evolved strains are suitable for recombinant protein perdeuteration two strains (A.66.2 and C.66.2) were lysogenized with the λ bacteriophage DE3 that carries the T7 gene 1 encoding the T7 RNA polymerase under control of the IPTG inducible lacUV5 promoter. The resulting strains were named, DAG1 (DE3) (Deuterium Adapted Glycerol) and DAG2(DE3), respectively. Next, recombinant protein production of superfolder green fluorescent protein (sfGFP) and the glycolytic enzyme triose-phosphate isomerase (TIM) from *Leishmania mexicana* in the deuterium-based growth medium (DD-M9) was tested. The evolved strains produced levels of TIM comparable to that of the parental strain (MG1655(DE3); Figure 4). Subcellular fractionation showed that the proportion of TIM recovered in the soluble fraction was close to 100% in the wild type and in the evolved strains (Figure 4). Production of sfGFP was evaluated by SDS-PAGE and by measuring fluorescence intensity of intact cells (Figure 5). Both evolved strains showed a higher yield of sfGFP compared to the parental strain at the analyzed time points (Figure 5b). This shows that the evolved strains are suitable for recombinant perdeuterated protein production. We noticed that the induction time is of critical importance for optimal yield of both TIM and sfGFP. Induction in the early exponential growth phase ($OD_{600} \sim 0.5$) resulted in considerably lower protein yield compared to when induction was done at a later time point ($OD_{600} \sim 1.2$; Figure 6). Next, we investigated the general feasibility of using the evolved strains for recombinant protein perdeuteration. Two membrane proteins, the outer membrane protein OmpF and the *Aquifex aeolicus* leucine transporter (LeuT) which serve as a model for neurotransmitter sodium symporters were produced in strain DAG1(DE3).

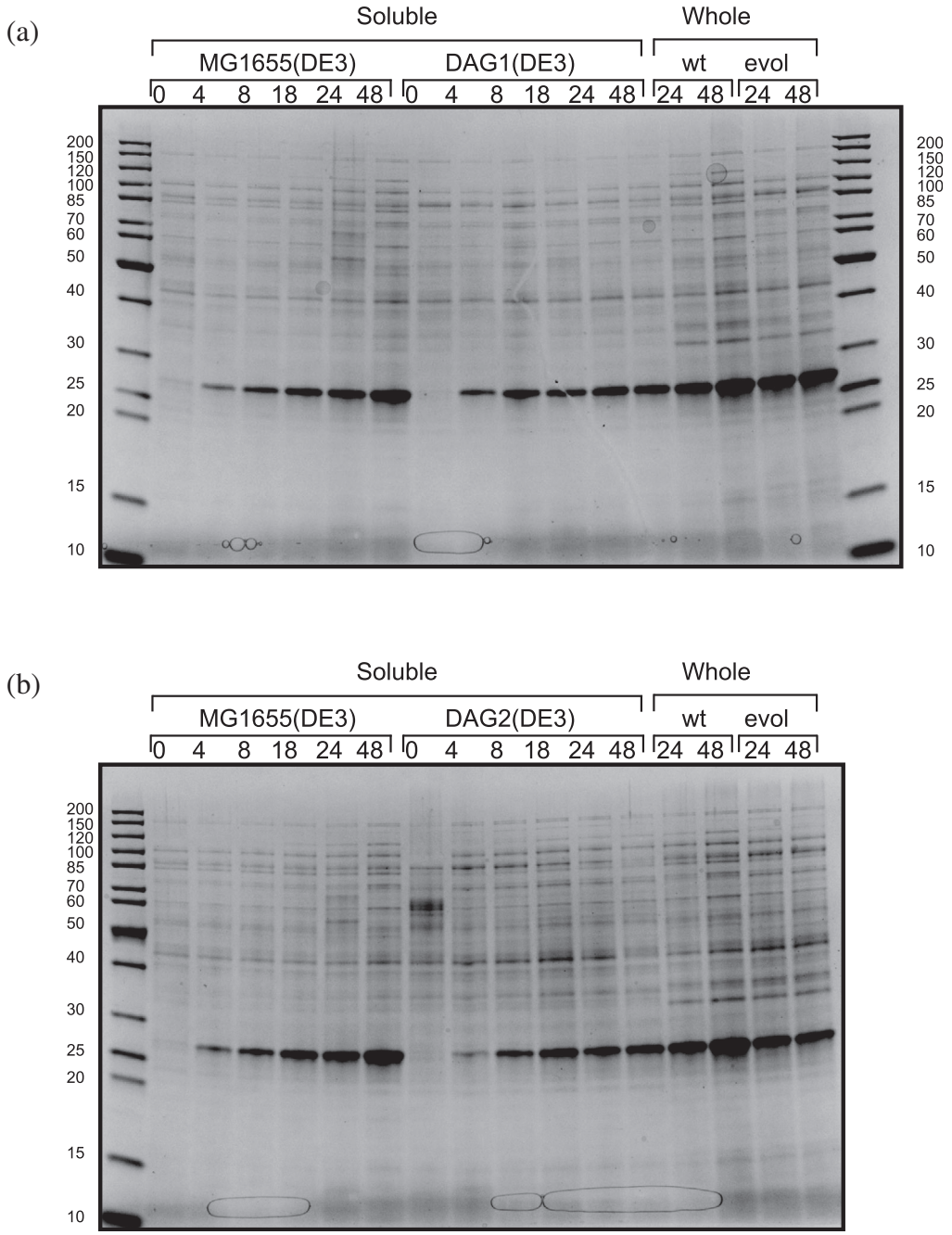

FIGURE 4 Recombinant production of TIM in wild type and two evolved *Escherichia coli* strains. SDS–PAGE analysis of whole cell and soluble fraction of lysed cells at the indicated time points after induction with IPTG. The expression of the gene encoding TIM was induced when OD_{600} was 1.2. Each lane was loaded with a sample corresponding to 10 µl cell culture. The molecular mass standards, in kilodaltons, are indicated on the left. (a) Wild type (wt, MG1655(DE3)) and evolved strain (evol, DAG1(DE3)), (b) (wt, MG1655(DE3)), and evolved strain (evol, DAG2(DE3))

SDS-PAGE analysis showed that the evolved strain produced high levels of OmpF (Figure S7a). LeuT could also be expressed in the adapted strain in deuterated minimal medium at about 0.34 mg per liter of growth medium (-Figure S7b). In addition, the DNA binding tryptophan repressor protein (TrpR) could also successfully be produced in the evolved strain (Figure S7c).

2.10 | Inactivation of the protease encoding genes *lon* and *ompT* does not improve protein yield

The *E. coli* B-strain, BL21(DE3) is the most commonly used host strain for recombinant protein production for structural biology according to data from the Protein

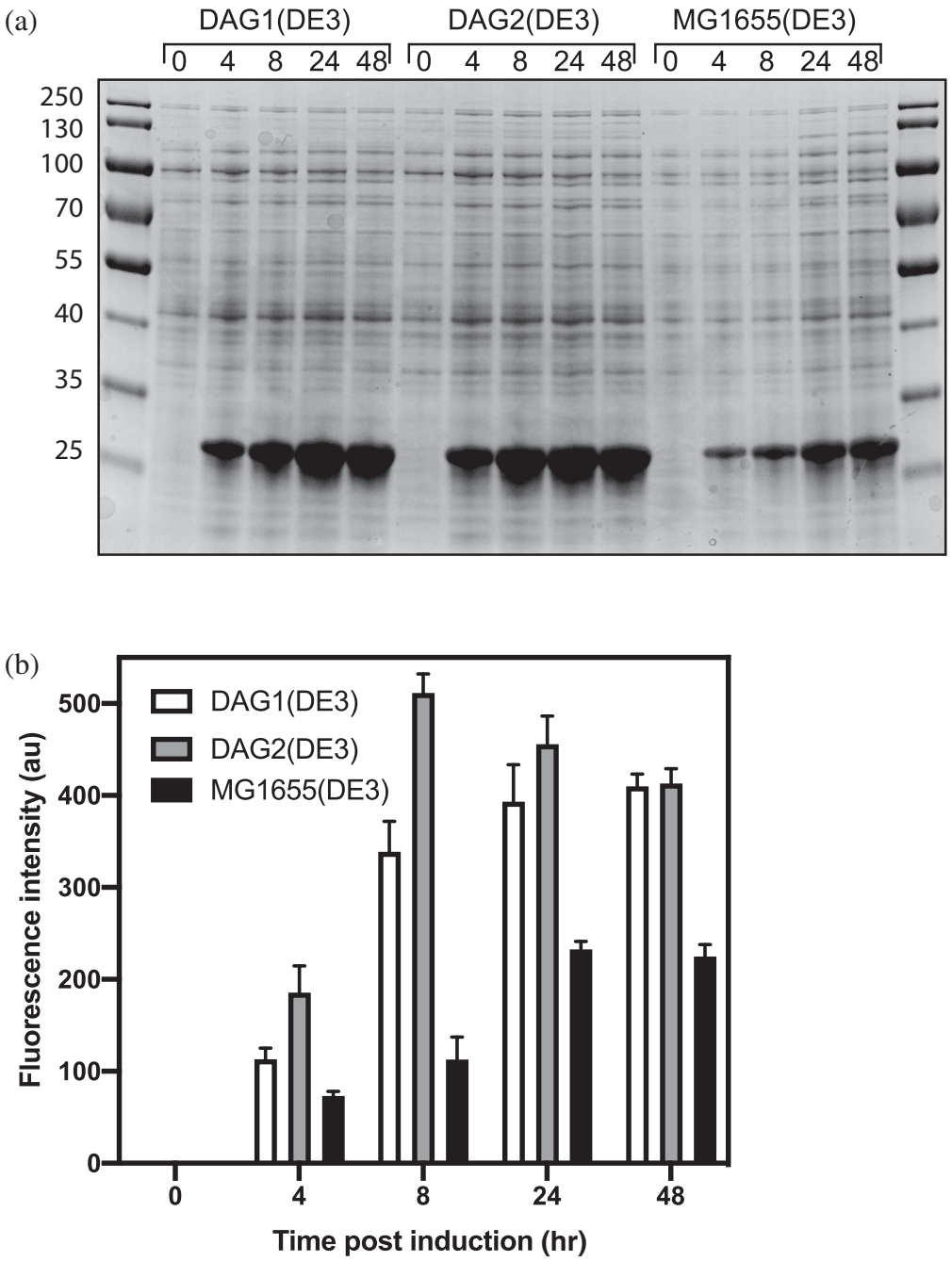

FIGURE 5 Recombinant production of sfGFP in wild type and two evolved *Escherichia coli* strains. (a) SDS–PAGE analysis of the soluble fraction of lysed cells at the indicated time points after induction with IPTG. Each lane was loaded with a sample corresponding to 19-μl cell culture. The positions of molecular mass standards, in kilodaltons, are indicated on the left. (b) Absolute fluorescence intensities (au, arbitrary units) in cells harvested at different time points after inducing expression of the gene encoding sfGFP. Three replicates were analyzed and the mean fluorescence intensity is plotted. The error bars represents the *SD*. A.66.2 (DE3) white bars, C.66.2(DE3) grey bars, and MG1655(DE3) black bars

Data Bank (http://www.rcsb.org/). One reason why *E. coli* BL21(DE3) is considered to be particularly suitable for protein production is that it contains fewer proteases.[42] The cytoplasmic AAA+ protease Lon is not produced due to insertion of a transposon in the *lon* promoter region and the gene encoding the outer membrane OmpT protease is absent in *E. coli* B-strains.[43,44] However, evidence that these proteases are directly linked to reduced recombinant protein yield are limited.[45,46] Moreover, the Lon protease plays a central role in protein quality control in most bacteria as it degrades aberrant proteins and functions as both a protease and a chaperone.[47] It has been found that in the Bl21(DE3)-derived strain C43(DE3) (also known as a Walker strain) which is adapted for membrane protein production expression of the *lon* gene has been restored.[48] To test the effect of *lon* and *ompT* on recombinant protein production in one of the evolved strains these genes were deleted in strain DAG1(DE3). Production of sfGFP was followed by fluorescence measurements. The analysis shows that both *lon* and *ompT* negatively affect the concentration of produced sfGFP (Figure 7). Both *ompT*[49] and *lon*[50] have been associated with the cellular response to heat shock and protein overexpression and could those negatively affect recombinant protein production.

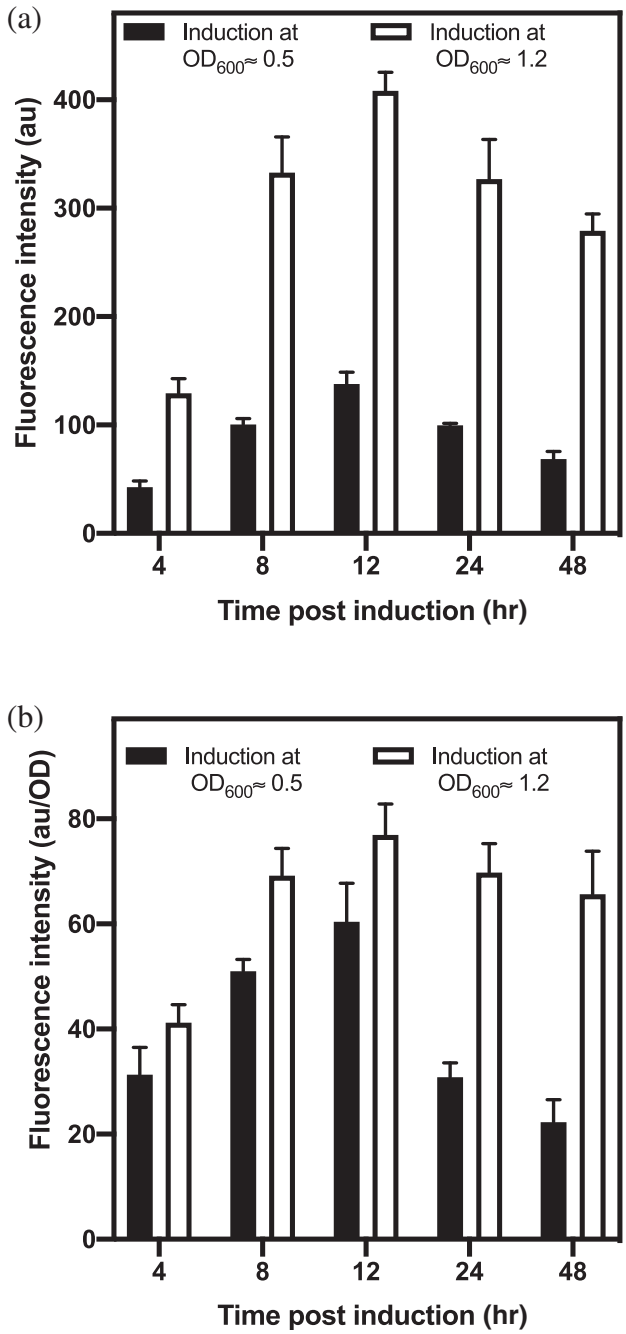

FIGURE 6 Time of induction affect yield of sfGFP. Absolute (a) and relative (b) fluorescence intensities in cells at different time points after inducing expression of the gene encoding sfGFP. Three replicates were analyzed and the mean fluorescence intensity is plotted. The error bars represents the *SD*. Induction was done either at an OD_{600} of 0.5 or at an OD_{600} of 1.2 as indicated. au, arbitrary units

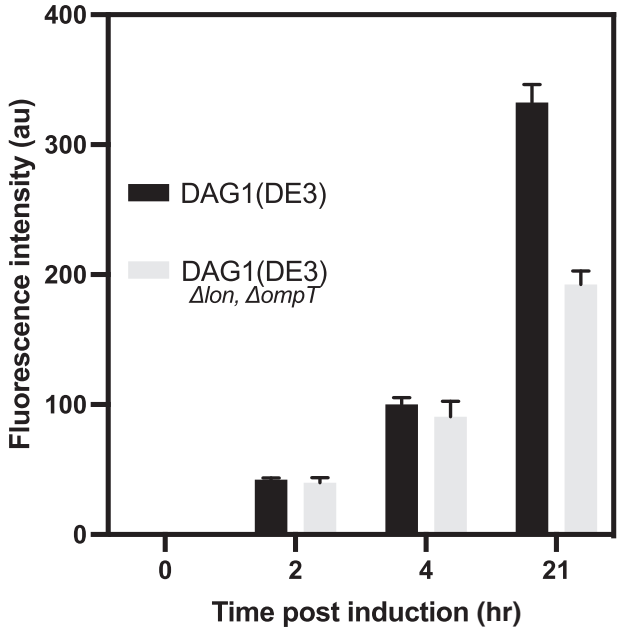

FIGURE 7 Effect of the protease encoding genes *lon* and *ompT* on sfGFP production. Fluorescence intensities in cells at different time points after inducing expression of the gene encoding sfGFP. Three replicates were analyzed and the mean fluorescence intensity is plotted. The error bars represents the *SD*

2.11 | Are there any genes that are essential for growth in deuterated medium?

To find out if there are any genes that are essential for growth on perdeuterated minimal glycerol medium we profiled the Keio single-gene deletion mutant collection. The 3,985 mutants were first grown on LA plates supplemented with kanamycin and then transferred using a 96-pin tool to M9-glycerol agar plates (H-M9 agar). Fifty-nine mutants were not able to grow on minimal medium. Many of these effect functions involved in amino acid or nucleotide biosynthesis. The obtained results largely agree with what has been reported previously.[51] However, some of the mutants found not to grow on H-M9 agar, for example, *ΔilvE* and *ΔmetE* were reported to grow in liquid minimal glycerol medium.[51] However, in agreement with our observations they were found to be essential for growth on minimal medium with glucose.[52] Next, the mutant strains were pinned from H-M9 agar to deuterated minimal glycerol agar plates (D-M9 agar). None of the strains failed to grow on D-M9 agar. A few strains that were found to grow poorly on D-M9 agar as well as strains not particularly affected by the deuterated medium were pinned in parallel onto H-M9 and D-M9 plates. Strains that grew poorly on the ordinary growth medium generally showed very poor growth on the deuterated growth medium. Images of colonies after 136 hr of growth were overlaid to estimate the relative colony size as a proxy of growth. Some deletion mutations (e.g., *lpd*, *pal*, *dksA*, *ptsH*, *truA*, *bioB*, *cysK*, *lipA*, *fepA*, *ldcA*, and *dnaT*) had a particular large effect on the relative colony size (Figure 8, Figure S8).

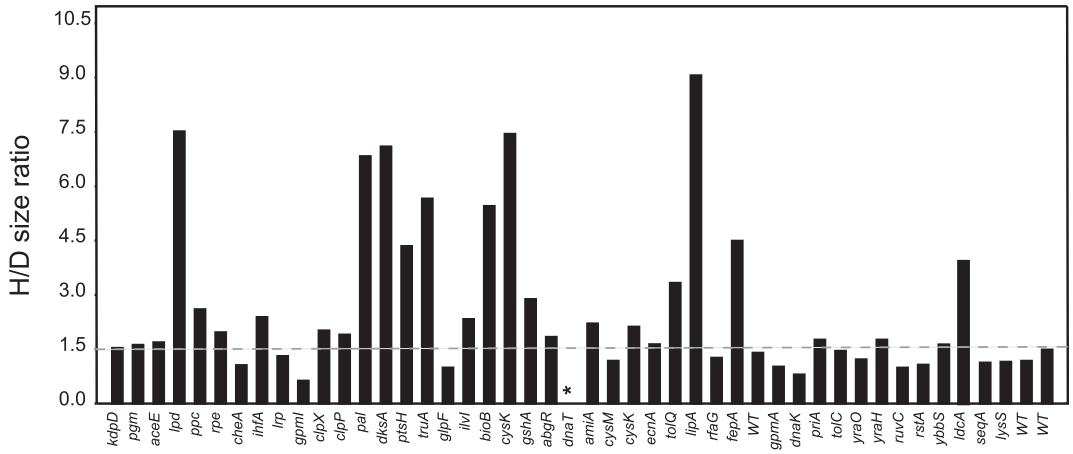

FIGURE 8 Relative growth of selected mutants from the single-gene knockout collection. The indicated strains were spotted onto H-M9 and D-M9 agar plates. Each bar represents the relative colony size for growth on H-M9 and D-M9 agar plates for the indicated strains. WT is the wild-type reference *Escherichia coli* K-12 BW25113 strain. The line indicates H/D size ration of the wild type. For the strain deleted for *dnaT* growth on D-M9 agar plates was very poor and the star (*) indicates that the H/D size ratio was very large but that an exact value could not be plotted. Images of the colonies from which the data was obtained is shown in Figure S8

The profiling presented here show that growth on deuterated medium in some but not all cases lead to a reduction of growth as compared to a non-deuterated medium. The data also indicate that there is no single gene which is essential for growth on deuterated medium. However, we cannot fully rule this out as the Keio collection does not comprise all genes and some mutants may rapidly accumulate suppressor mutations that restore growth on deuterated media. Some mutants, for example those deleted for *bioB* (biotin synthase) or *dnaT* (primosomal protein), showed very poor growth on both media. This shows that also strains which have limited growth capabilities on defined ^{1}H media can grow in deuterated growth medium. If the poorer growth on deuterated medium is due to that the kinetic isotope effect has the same impact on all metabolic pathways, we would aspect to see a constant ratio of growth in H versus D medium for all tested strains. This was not observed as shown in Figure 8 suggesting that some cellular reactions are more affected than others by the presence of the heavy isotope of hydrogen.

When particular proteins, as, for example, enzymes, are recombinantly produced in *E. coli*, it is sometimes required to eliminate related endogenous enzyme activities to avoid contaminating activities. The findings presented here can be used to predict growth properties of single knock out mutant for use as host for production of specific deuterated proteins.

3 | CONCLUSIONS

We show here that *E. coli* K-12 strains with increased fitness in deuterium-based growth media can be isolated by adaptive experimental evolution. Analysis of the obtained strains reveal that there are multiple routes to genetic adaptation to growth in deuterium-based media and that the improved fitness also translates to nondeuterated minimal growth medium. Thus, mutations are not primarily selected due to specific effects of deuteration but rather to the general growth restricting effects imposed by the deuterium isotope effect. These findings are further corroborated by that no specific gene was found to be essential for growth in deuterated minimal medium. Nevertheless, we show that evolved strains can be of practical use for recombinant production of perdeuterated proteins for structural biology applications. Glycerol-d_8 is commonly used as a carbon source for deuterium labeling and the evolved strains show enhanced growth rates in particular in minimal media with glycerol as carbon source. One of the evolved strains was adapted for the widely used T7 RNA polymerase overexpression systems and used in production of perdeuterated recombinant soluble and membrane-proteins.

4 | MATERIALS AND METHODS

4.1 | Bacterial strains, oligonucleotides, plasmids, culture conditions, and transformation

The bacterial strains, plasmids and oligonucleotides used in this study are listed in Tables S3, S4 and S5, respectively. Media were supplemented when needed with the appropriate antibiotic at the following concentrations: 100 μg/ml ampicillin, 50 μg/ml chloramphenicol, 50 μg/ml kanamycin, 50 μg/ml spectinomycin. Lysogeny broth

(LB) (10 g/L NaCl, 5 g/L Difco yeast extract, 10 g/L Difco tryptone) made up with H_2O, pH 7.6, was used as a rich medium. The defined M9 minimal medium (here referred to as H-M9) consisted of (42.7 mM $Na_2HPO_4 \cdot 2H_2O$, 22 mM KH_2PO_4, 8.6 mM NaCl, 107 mM NH_4Cl, 1 mM $MgSO_4 \cdot 7\ H_2O$, 0.1 mM $CaCl_2$, 2 mg/L thiamine HCl, 0.018 mM $FeCl_3 \cdot 6H_2O$ with glycerol 0.8% [w/v] as sole carbon and energy source). To induce recombinant gene expression 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to the growth media. When deuterated media were prepared all solutions were made in heavy water (99.8% D-atom, Sigma-Aldrich) and then filtered with 0.22 μm sterile filter (VWR Vacuum filtration unit). In the deuterated media, either regular glycerol (medium referred to as D-M9) or glycerol d-8 (98% D-atom, Sigma-Aldrich; medium referred to as DD-M9) was used. All salts used in D-M9 and DD-M9 were the same as in M9, but instead of $Na_2HPO_4 \cdot 2H_2O$ anhydrous Na_2HPO_4 was used. For perdeuterated protein production, a 1 M stock solution of IPTG prepared in heavy water was used. Plates and liquid cultures were incubated at 30°C or 37°C. Strains were revived from glycerol freezer stocks by streaking onto lysogeny broth agar (LA) and incubation at 37°C. Plasmid DNA was transformed into electrocompetent cells which were made using the glycerol/mannitol density step centrifugation method.[53]

4.2 | Growth experiments

Overnight bacterial cultures were set-up by inoculating a few colonies from an LA plate of the appropriate strain into 25 ml H-M9, and incubated at 37°C, 200 rpm. The overnight cultures were diluted approximately 30 times in fresh medium to an optical density at 600 nm (OD_{600}) of 0.1 and incubated at 37°C with shaking (200 rpm). When growth was performed in deuterated media, the following steps were carried out to minimize carry-over of ordinary hydrogen. The overnight cultures were harvested by centrifugation (8 min at 8000× *g* at 20°C), the supernatant was carefully removed, and the pellet suspended in 10 ml D-M9. The cultures were then incubated at 37°C, until it reached an OD_{600} between 1 and 3. The cultures were diluted to an OD_{600} of 0.1 into 25 ml DD-M9 and incubated at 37°C, 200 rpm. Cultures of 25 ml total volume were cultivated in 250 ml baffled Erlenmeyer flasks (Bellco) and 10 ml cultures were grown in 100 ml Erlenmeyer flasks (Schott DURAN). Flasks were placed in a 37°C incubator. Growth rate experiments were performed by measuring the optical density at 600 nm (OD_{600}) of duplicate cultures over several time points at cell densities between 0.05 and 2. The specific growth rate (μ) was calculated as the slope of the linear best-fit line through a plot of ln (OD_{600}) versus time (hr). The generation time (or doubling), t_d, is equal to ln2/μ Growth curves are shown as semilogarithmic plots of time versus OD.

4.3 | Phase contrast microscopy, cell counting, and dry mass estimation

All cultures were set-up as described under "growth experiments." For microscopy and cell counting 0.5 ml of culture was withdrawn and placed on ice until analyzed (<30 min). At the same time point, OD_{600} was recorded. Phase contrast images of *E. coli* cells (OD_{600} between 0.3 and 0.5) were taken using an (ZEISS Axio Imager2) equipped with ORCA-FLASH4.0 V2 Digital CMOS Camera (Hamamatsu). Prior to imaging, cells were loaded onto a microscope slide covered with a thin pad of 1% (wt/vol) agarose in PBS (83.2 mM NaCl, 2.2 mM KH_2PO_4, 20.9 mM K_2HPO_4 [pH 7.5]). Hundred cells per growth medium (H-M9 and DD-M9) were analyzed for cell length and width. A *t*-test was used to determine whether the difference in cell size were statistically significant. To establish the relation between OD_{600} values and cell number each culture was measured at various OD_{600} values and then cell numbers were counted using a Bürker chamber (Hirschmann) with a chamber depth of 0.01 mm. Cells in 10 squares with area size of 0.01 mm^2 were counted, to obtain a concentration of cells/ml of culture. The culture was diluted if there were more than 30 cells per square. Cell numbers were normalized against OD_{600}. For the determination of the relation between dry cell weight and OD_{600} value of stationary phase cultures 3 × 1 ml of cultures were collected into pre-weighed glass falcon tubes and spun down at 4500 rpm for 8 min. The supernatant was carefully removed and the tubes were incubated for 48 hr at 85°C. The tubes were weighted using a scale with readability of 0.0001 g and dry cell weight was estimated and normalized against OD_{600} values.

4.4 | Evolutionary adaptation

The parental strain *E. coli* MG1655 (K-12) was streaked on an LA plate and incubated overnight at 37°C. The following day a few colonies were used to inoculate 25 ml H-M9 in a 250 ml baffled Erlenmeyer flask that was incubated at 37°C and 200 rpm until the culture reached an OD_{600} of 3. The culture was centrifuged for 8 min at 8000× *g*, the supernatant was removed and the cell pellet suspended in 5 ml DD-M9. The suspended cells were used to inoculate 15 ml DD-M9 to an OD_{600} of 0.1. The culture

was divided into three equal parts (5 ml) in three 50 ml TPP® TubeSpin bioreactor tubes which were placed at 37°C inclined at a 45° angle and rotated at 200 rpm. The cultures are referred to as three separate lineages (lineage A, B, and C). The cultures were grown until they reached an OD_{600} of approximately 5, then the exact OD_{600} was recorded and used to estimate the number of generations passed from the start of the experiment. Next, each lineage was used to inoculate 5 ml fresh DD-M9 medium. Cultures up to the 11th subculturing were started at an OD_{600} of 0.1 (1:50 dilution), after this point, an OD_{600} of 0.05 (1:100 dilution) was used. Subculturing was done every second day (48 hr) for 17 subculturings and after that, once a day. Periodically, for every third subculturing, 660 µl of the culture was withdrawn and mixed with 340 µl of 85% glycerol and stored at −80°C. Samples were inspected for contamination by streaking for single colonies on LA plates. At different time points, single colonies were isolated from the cultures by repeatedly (three times) streaking and incubating on LA plates. Single colonies were used to streak a bacterial lawn on LA plates which was later used to prepare a glycerol stock for long-term freezer storage at −80°C.

4.5 | DNA sequencing

Illumina sequencing technology was used for whole genome shot-gun sequencing and Sanger sequencing was used to confirm mutations in specific genes of interest. Genomic RNA-free DNA was purified using the DNeasy Blood and Tissue Kit (Qiagen) according to the manufacturer's instructions. Quality control and library preparation were performed by GATC (Konstanz, Germany) and sequenced on a genome sequencer Illumina HiSeq with 2× 150 bp paired end read output. Reads, approximately 10 million reads per sample, were mapped using CLC genomics workbench (Qiagen) to the reference genome sequence of *E. coli* MG1655 (Genbank entry code: U00096.3) obtained from the NCBI genome repository and corrected for the identified variants in the laboratory stocked MG1655 strain.[13] More than 99% of the reads could be mapped to the reference sequence. Only variants detected at a frequency above 95% were considered. In order to confirm and identify particular mutations, appropriate stretches of DNA were amplified by PCR and then subjected to Sanger sequencing (Eurofins). Primers for sequencing and amplification are listed in Table S5.

4.6 | Other genetic modifications

Deletion of *lon* and *ompT* was done using the λ-Red recombinase-mediated gene deletion method.[54] A PCR product harboring 50 bp end sequences homologous to *ompT* and a kanamycin resistance marker was amplified with plasmid pKD4 as template and primers ompT_up and ompT_down. For deletion of *lon* PCR was done using primers lon_up and lon_down with plasmid pKD3 as template generating a fragment with a chloramphenicol resistance marker. The *lon* and the *ompT* PCR products were transformed into *E. coli* DA1 and *E. coli* MG1655, respectively. The recipient strains both harbored the λ-Red recombinase expression plasmid pKD46, and subsequently transformants were selected on media plates containing kanamycin or chloramphenicol. Deletion of *lon* and *ompT* was verified by amplifying the appropriate chromosomal region using primer pair lon_for/lon_rev and ompT_for/ompT_rev, respectively. Phage P1 transduction was done according to[55] and used to transduce the *lon* and *ompT* mutations to strain DAG1 (DE3). The kanamycin resistance gene was eliminated by transforming the appropriate strain with plasmid pCP20 that carries the Flp recombinase gene.[56] To make the double *lon*, *ompT* mutant Δ*lon*::cap was transduced into strain DAG1(DE3) Δ*ompT*::frt. In order to test recombinant protein production capabilities, the gene for T7 RNA polymerase was introduced into strains of interest by lysogenizing with the lambda phage DE3 using the λDE3 lysogenization kit as previously described.[13] Phage P1 transduction was used to transduce the *rpoS*::kan mutation from the Keio collection of deletion mutants to the MG1655 background according to Reference 55.

4.7 | Valine resistance assay

Overnight cultures grown at 37°C in H-M9 was used to inoculate H-M9 medium supplemented with different concentrations of L-valine: 1000 µg/ml, 200 µg/ml, 40 µg/ml, 8 µg/ml. All tubes were inoculated to an OD_{600} of 0.01 and incubated at 37°C, 200 rpm. Cultures were visually inspected for turbidity after incubation for 24 and 48 hr. As a reference, a culture grown in H-M9 without added valine was used.

4.8 | Catalase test

Strains were streaked on LA plates and incubated overnight at 37°C. Before the catalase assay was carried out the plates were incubated at 4°C for 48 hr to induce *rpoS* expression. 10 µl of a 6% (vol/vol) hydrogen peroxide solution was applied onto bacterial patches. Vigorous foaming was an indication of high catalase activity and linked to the presence of the RNA polymerase sigma factor RpoS, while faint foaming was considered to be due

to lack of RpoS activity that is required for catalase gene expression.

4.9 | Protein production

E. coli MG1655(DE3), *E. coli* DAG1(DE3), and *E. coli* DAG2(DE3) were transformed with plasmids pNIC28_Lm_TIM[57] (encoding triosephosphate isomerase of *L. mexicana*) and pETM14_sfGFP (encoding super folder GFP).[13] Cultures were inoculated and set-up as outlined under "growth experiments" in DD-M9. At an OD_{600} of 0.5 or at an OD_{600} of 1.2 IPTG to a final concentration of 0.5 mM was added. At different time points 1 ml of culture was harvested by centrifugation at 8000× *g* for 8 min at 4°C. The supernatant was carefully removed, and the cell pellet was suspended in 0.25 ml cold 50 mM Tris–HCl, 100 mM NaCl, 10 mM EDTA buffer (pH 8.0) and stored at −20°C until used. Samples were thawed on ice and sonicated using a Vibra Cell disruptor. Samples were then centrifuged at 20,000× *g* for 45 min at 4°C. Supernatant of 40 µl was mixed with 10 µl 5X SDS-PAGE sample buffer (60 mM Tris–HCl [pH 6.8], 25% glycerol, 2% SDS, 0.1% bromophenol blue, 0.7 M 2-mercaptoethanol) and incubated at 95°C for 10 min. Each sample of 6 µl of was loaded per lane on an SDS-PAGE gel (Any kD™ Criterion™ TGX™). Tris–glycine–SDS (25 mM Tris, 192 mM glycine, 0.1% SDS, pH 8.3) was used as running buffer. Bands were visualized by using Coomassie Brilliant Blue (SimplyBlue™ SafeStain, Bio-Rad) according to the manufacturer's protocol. Cell pellets for sfGFP quantitation were suspended in 0.5 ml of cold 50 mM Tris–HCl, 100 mM NaCl, 10 mM EDTA buffer (pH 8.0). Of note, 300 µl was transferred to another tube and sonicated as described above. This served as the whole cell lysate. Whole cell lysate of 100 µl was centrifuged 20,000× *g* for 45 min at 4°C. The supernatant was used as soluble fraction. Soluble fraction of 40 µl was mixed with 10 µl 5X SDS-PAGE sample buffer, heated at 98°C for 10 min. Samples were spun down for a few seconds and placed on ice. Sample of 6 µl was loaded per lane on a Criterion™ TGX™ Any KD™ (BIO-RAD) precast gels. Measurement of sfGFP fluorescence on whole cells was done by resuspending cells in 0.5 ml 50 mM Tris, 100 mM NaCl, 10 mM EDTA buffer (pH 8.0). Samples were then diluted 100-fold in 100 mM NaCl, 100 mM sodium-phosphate buffer (pH 7.5), and OD_{600} as well as sfGFP fluorescence (excitation 485 nm and emission 510 nm) were measured. Fluorescence was measured using a RF-5301 Spectrofluorophotometer (Shimadzu) and fluorescence intensity was recorded after subtraction of background fluorescence for the buffer used.

4.10 | Keio collection screen

The Keio collection which consists of 3,985 of nonessential gene disruptions in *E. coli* K-12 BW25113 background[58] was used to screen for genes essential for growth on deuterated solid minimal medium. It was obtained from the National Bioresource Project (NIG, Japan). Screening was carried out on solid media in Nunc® Omnitray (Sigma) filled with 50 ml ordinary M9 (H-M9) or deuterated M9 (D-M9) minimal medium supplemented with 1.5% agar. A 96 solid pin multi-blot replicator was used to transfer colonies grown on LA plates onto H-M9 agar plates. After incubation, the colonies were transferred from H-M9 to D-M9 plates. All plates were incubated at 37°C. ImageJ[59] was used to calculate relative spot sizes. The spot size area is an arbitrary number which is normalized to the length of the left corner of the Omnitray plate.

ACKNOWLEDGMENTS

The authors acknowledge the financial support from Carl Tryggers Stiftelse, Erik Philip-Sörensens stiftelse, Jörgen Lindström's Foundation and the Swedish Research Council Grant 2019-05578_3.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

AUTHOR CONTRIBUTIONS

Vinardas Kelpšas: Formal analysis (equal); investigation (lead); methodology (equal); validation (equal); visualization (equal); writing – original draft (supporting); writing – review and editing (supporting). **Claes von Wachenfeldt:** Conceptualization (equal); data curation (equal); formal analysis (equal); funding acquisition (lead); investigation (equal); methodology (equal); project administration (lead); resources (equal); supervision (lead); validation (equal); visualization (equal); writing – original draft (lead); writing – review and editing (lead).

ORCID

Claes von Wachenfeldt https://orcid.org/0000-0001-8950-3218

REFERENCES

1. Urey HC. The separation and properties of the isotopes of hydrogen. Science. 1933;78:566–571.
2. Lewis GN. The biology of heavy water. Science. 1934;79:151–153.
3. Cohen ER, Cvitaš T, Frey JG, et al. Quantities, Units and Symbols in Physical Chemistry. 3rd ed. Cambridge, UK: The Royal Society of Chemistry; 2007. https://doi.org/10.1039/9781847557889

4. Gröning M, van Duren M, Andreescu L. Metrological characteristics of the conventional measurement scales for hydrogen and oxygen stable isotope amount ratios: the δ-scales. Combining and Reporting Analytical Results. Royal Society of Chemistry; 2006; p. 62–72. https://doi.org/10.1039/9781847557582-00062
5. O'Neil MJ. The Merck index: an encyclopedia of chemicals, drugs, and biologicals. Royal Society of Chemistry; 2013.
6. Assael MJ, Monogenidou SA, Huber ML, Perkins RA, Sengers JV. New international formulation for the viscosity of heavy water. J Phys Chem Ref Data. 2021;50(3):033102-1–033102-20. http://dx.doi.org/10.1063/5.0048711
7. Katz JJ, Crespi HL. Deuterated organisms: Cultivation and uses. Science. 1966;151:1187–1194.
8. O'Dell WB, Bodenheimer AM, Meilleur F. Neutron protein crystallography: A complementary tool for locating hydrogens in proteins. Arch Biochem Biophys. 2016;602:48–60.
9. Kelpšas V, Caldararu O, Blakeley MP, et al. Neutron structures of Leishmania mexicana triosephosphate isomerase in complex with reaction-intermediate mimics shed light on the proton-shuttling steps. IUCrJ. 2021;8(4):633–643. http://dx.doi.org/10.1107/s2052252521004619
10. Casadei CM, Gumiero A, Metcalfe CL, et al. Heme enzymes. Neutron cryo-crystallography captures the protonation state of ferryl heme in a peroxidase. Science. 2014;345:193–197.
11. Fisher SJ, Blakeley MP, Howard EI, et al. Perdeuteration: Improved visualization of solvent structure in neutron macromolecular crystallography. Acta Cryst D. 2014;70:3266–3272.
12. Meilleur F, Dauvergne MT, Schlichting I, Myles DAA. Production and X-ray crystallographic analysis of fully deuterated cytochrome P450cam. Acta Cryst D. 2005;61:539–544.
13. Kelpšas V, von Wachenfeldt C. Strain improvement of *Escherichia coli* K-12 for recombinant production of deuterated proteins. Sci Rep. 2019;9:17694.
14. Opitz C, Ahrné E, Goldie KN, Schmidt A, Grzesiek S. Deuterium induces a distinctive *Escherichia coli* proteome that correlates with the reduction in growth rate. J Biol Chem. 2019;294:2279–2292.
15. Meilleur F, Weiss KL, Myles DA. Deuterium labeling for neutron structure-function-dynamics analysis. Methods Mol Biol. 2009;544:281–292.
16. Sievert C, Nieves LM, Panyon LA, et al. Experimental evolution reveals an effective avenue to release catabolite repression via mutations in XylR. Proc Natl Acad Sci U S A. 2017;114:7349–7354.
17. LaCroix RA, Sandberg TE, O'Brien EJ, et al. Use of adaptive laboratory evolution to discover key mutations enabling rapid growth of *Escherichia coli* K-12 MG1655 on glucose minimal medium. Appl Environ Microbiol. 2015;81:17–30.
18. Mundhada H, Seoane JM, Schneider K, et al. Increased production of L-serine in *Escherichia coli* through adaptive laboratory evolution. Metab Eng. 2017;39:141–150.
19. Herring CD, Raghunathan A, Honisch C, et al. Comparative genome sequencing of *Escherichia coli* allows observation of bacterial evolution on a laboratory timescale. Nat Genet. 2006;38:1406–1412.
20. Blattner FR, Plunkett G 3rd, Bloch CA, et al. The complete genome sequence of *Escherichia coli* K-12. Science. 1997;277:1453–1462.
21. Anderson MJ, DeLaBarre B, Raghunathan A, Palsson BO, Brunger AT, Quake SR. Crystal structure of a hyperactive *Escherichia coli* glycerol kinase mutant Gly230→ Asp obtained using microfluidic crystallization devices. Biochemistry. 2007;46:5722–5731.
22. Applebee MK, Joyce AR, Conrad TM, Pettigrew DW, Palsson BØ. Functional and metabolic effects of adaptive glycerol kinase (GLPK) mutants in *Escherichia coli*. J Biol Chem. 2011;286:23150–23159.
23. Yao R, Liu Q, Hu H, Wood TK, Zhang X. Metabolic engineering of *Escherichia coli* to enhance acetol production from glycerol. Appl Microbiol Biotechnol. 2015;99:7945–7952.
24. Liu WZ, Faber R, Feese M, Remington SJ, Pettigrew DW. *Escherichia coli* glycerol kinase: Role of a tetramer interface in regulation by fructose 1, 6-bisphosphate and phosphotransferase system regulatory protein IIIglc. Biochemistry. 1994;33:10120–10126.
25. Pettigrew DW, Liu WZ, Holmes C, Meadow ND, Roseman S. A single amino acid change in *Escherichia coli* glycerol kinase abolishes glucose control of glycerol utilization in vivo. J Bacteriol. 1996;178:2846–2852.
26. Favre R, Wiater A, Puppo S, Iaccarino M, Noelle R, Freundlich M. Expression of a valine-resistant acetolactate synthase activity mediated by the *ilvO* and *ilvG* genes of *Escherichia coli* K-12. Mol Gen Genet. 1976;143:243–252.
27. Lawther RP, Calhoun DH, Adams CW, Hauser CA, Gray J, Hatfield GW. Molecular basis of valine resistance in *Escherichia coli* K-12. Proc Natl Acad Sci U S A. 1981;78:922–925.
28. Parekh BS, Hatfield GW. Growth rate-related regulation of the *ilvGMEDA* operon of *Escherichia coli* K-12 is a consequence of the polar frameshift mutation in the *ilvG* gene of this strain. J Bacteriol. 1997;179:2086–2088.
29. Andersen DC, Swartz J, Ryll T, Lin N, Snedecor B. Metabolic oscillations in an *E. coli* fermentation. Biotechnol Bioeng. 2001;75:212–218.
30. Conrad TM, Frazier M, Joyce AR, et al. RNA polymerase mutants found through adaptive evolution reprogram *Escherichia coli* for optimal growth in minimal media. Proc Natl Acad Sci U S A. 2010;107:20500–20505.
31. Jin DJ, Gross CA. Mapping and sequencing of mutations in the *Escherichia coli rpoB* gene that lead to rifampicin resistance. J Mol Biol. 1988;202:45–58.
32. Battesti A, Majdalani N, Gottesman S. The RpoS-mediated general stress response in *Escherichia coli*. Annu Rev Microbiol. 2011;65:189–213.
33. Tanaka K, Handel K, Loewen PC, Takahashi H. Identification and analysis of the rpoS-dependent promoter of *katE*, encoding catalase HPII in *Escherichia coli*. Biochim Biophys Acta Gen Subj. 1997;1352:161–166.
34. Chen G, Patten CL, Schellhorn HE. Positive selection for loss of RpoS function in *Escherichia coli*. Mutat Res. 2004;554:193–203.
35. Charusanti P, Conrad TM, Knight EM, et al. Genetic basis of growth adaptation of *Escherichia coli* after deletion of pgi, a major metabolic gene. PLoS Genet. 2010;6:e1001186.
36. Notley-McRobb L, King T, Ferenci T. *rpoS* mutations and loss of general stress resistance in *Escherichia coli* populations as a consequence of conflict between competing stress responses. J Bacteriol. 2002;184:806–811.
37. Schlösser A, Meldorf M, Stumpe S, Bakker EP, Epstein W. TrkH and its homolog, TrkG, determine the specificity and kinetics of cation transport by the Trk system of *Escherichia coli*. J Bacteriol. 1995;177:1908–1910.

38. Bossemeyer D, Borchard A, Dosch DC, et al. K+-transport protein TrkA of *Escherichia coli* is a peripheral membrane protein that requires other *trk* gene products for attachment to the cytoplasmic membrane. J Biol Chem. 1989;264:16403–16410.
39. Rhoads D, Epstein W. Energy coupling to net K+ transport in *Escherichia coli* K-12. J Biol Chem. 1977;252:1394–1401.
40. Cao Y, Jin X, Huang H, et al. Crystal structure of a potassium ion transporter, TrkH. Nature. 2011;471:336–340.
41. Cao Y, Pan Y, Huang H, et al. Gating of the TrkH ion channel by its associated RCK protein TrkA. Nature. 2013;496:317–322.
42. Rosano GL, Ceccarelli EA. Recombinant protein expression in *Escherichia coli*: Advances and challenges. Front Microbiol. 2014;5:172.
43. SaiSree L, Reddy M, Gowrishankar J. IS186 insertion at a hot spot in the lon promoter as a basis for lon protease deficiency of *Escherichia coli* B: Identification of a consensus target sequence for IS186 transposition. J Bacteriol. 2001;183:6943–6946.
44. Donch J, Chung YS, Greenberg J. Locus for radiation resistance in *Escherichia coli* strain B/r. Genetics. 1969;61:363–370.
45. Grodberg J, Dunn JJ. ompT encodes the *Escherichia coli* outer membrane protease that cleaves T7 RNA polymerase during purification. J Bacteriol. 1988;170:1245–1253.
46. Waegeman H, de Lausnay S, Beauprez J, Maertens J, de Mey M, Soetaert W. Increasing recombinant protein production in *Escherichia coli* K12 through metabolic engineering. New Biotech. 2013;30:255–261.
47. Wohlever ML, Baker TA, Sauer RT. Roles of the N domain of the AAA+ Lon protease in substrate recognition, allosteric regulation and chaperone activity. Mol Microbiol. 2014;91:66–78.
48. Kwon S-K, Kim SK, Lee DH, Kim JF. Comparative genomics and experimental evolution of *Escherichia coli* BL21(DE3) strains reveal the landscape of toxicity escape from membrane protein overproduction. Sci Rep. 2015;5:16076.
49. Gill R, DeLisa MP, Shiloach M, Holoman TR, Bentley WE. OmpT expression and activity increase in response to recombinant chloramphenicol acetyltransferase overexpression and heat shock in *E. coli*. J Mol Microbiol Biotechnol. 2000;2:283–289.
50. Jürgen B, Lin HY, Riemschneider S, et al. Monitoring of genes that respond to overproduction of an insoluble recombinant protein in *Escherichia coli* glucose-limited fed-batch fermentations. Biotech Bioeng. 2000;70:217–224.
51. Joyce AR, Reed JL, White A, et al. Experimental and computational assessment of conditionally essential genes in *Escherichia coli*. J Bacteriol. 2006;188:8259–8271.
52. Patrick WM, Quandt EM, Swartzlander DB, Matsumura I. Multicopy suppression underpins metabolic evolvability. Mol Biol Evol. 2007;24:2716–2722.
53. Warren DJ. Preparation of highly efficient electrocompetent *Escherichia coli* using glycerol/mannitol density step centrifugation. Anal Biochem. 2011;413:206–207.
54. Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc Natl Acad Sci U S A. 2000;97:6640–6645.
55. Thomason LC, Costantino N, Court DL. *E. coli* genome manipulation by P1 transduction. Curr Protoc Mol Biol. 2007;1.17.1–1.17.8. https://doi.org/10.1002/0471142727.mb0117s7
56. Cherepanov PP, Wackernagel W. Gene disruption in *Escherichia coli*: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. Gene. 1995;158:9–14.
57. Kelpšas V, Lafumat B, Blakeley MP, Coquelle N, Oksanen E, von Wachenfeldt C. Perdeuteration, large crystal growth and neutron data collection of *Leishmania mexicana* triose-phosphate isomerase E65Q variant. Acta Crystallogr F. 2019;75:260–269.
58. Baba T, Ara T, Hasegawa M, et al. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: The Keio collection. Mol Syst Biol. 2006;2:2006.0008.
59. Schneider CA, Rasband WS, Eliceiri KW. NIH image to ImageJ: 25 years of image analysis. Nat Methods. 2012;9:671–675.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Kelpšas V, von Wachenfeldt C. Enhancing protein perdeuteration by experimental evolution of *Escherichia coli* K-12 for rapid growth in deuterium-based media. Protein Science. 2021;30:2457–73. https://doi.org/10.1002/pro.4206