Analyzing the Genetic Characteristics of a Tryptophan-overproducing Escherichia Coli

Dongqin Ding
Tianjin Institute of Industrial Biotechnology Chinese Academy of Sciences
Danyang Bai
Tianjin Institute of Industrial Biotechnology Chinese Academy of Sciences
Jinlong Li
Tianjin Institute of Industrial Biotechnology Chinese Academy of Sciences
Zhitao Mao
Tianjin Institutes of Industrial Biotechnology Chinese Academy of Sciences
Yaru Zhu
Tianjin Institute of Industrial Biotechnology Chinese Academy of Sciences
Pi Liu
Tianjin Institutes of Industrial Biotechnology Chinese Academy of Sciences
Jianping Lin
Tianjin Institutes of Industrial Biotechnology Chinese Academy of Sciences
Hongwu Ma
Tianjin Institute of Industrial Biotechnology Chinese Academy of Sciences
Dawei Zhang (zhang_dw@tib.cas.cn)
Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences https://orcid.org/0000-0001-5555-6151

Research

Keywords: L-trp, genome sequencing, protein structure, transcriptome analysis, transcriptional regulator

Posted Date: December 9th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-122063/v1

Version of Record: A version of this preprint was published at Bioprocess and Biosystems Engineering on March 22nd, 2021. See the published version at https://doi.org/10.1007/s00449-021-02552-4.

Abstract

Background: L-tryptophan (L-trp) production in *Escherichia coli* has been developed by employing random mutagenesis and selection for a long time, but this approach produces an unclear genetic background.

Results: We generated the L-trp overproducer TPD5 by combining an intracellular L-trp biosensor and fluorescence-activated cell sorting (FACS) in *E. coli*, and succeeded in elucidating the genetic basis for L-trp overproduction. The most significant identified positive mutations affected TnaA (deletion), AroG (S211F), TrpE (A63V), and RpoS (nonsense mutation Q33*). The underlying structure-function relationships of the feedback-resistant AroG (S211F) and TrpE (A63V) mutants were uncovered based on protein structure modeling and molecular dynamics simulations, respectively. According to transcriptomic analysis, the global regulator RpoS not only has a great influence on cell growth and morphology, but also on carbon utilization and the direction of carbon flow. Finally, by balancing the concentrations of the L-trp precursors serine and glutamine based on the above analysis, we further increased the titer of L-trp to 3.18 g/L with a yield of 0.18 g/g.

Conclusions: The analysis of the genetic characteristics of an L-trp overproducing *E. coli* provides valuable information on L-trp synthesis and elucidates the phenotype and complex cellular properties in a high-yielding strain, which opens the possibility to transfer beneficial mutations and reconstruct an overproducer with a clean genetic background.

Background

Ideal strains can be created through targeted and nontargeted strain improvement efforts, which have both been proved to be powerful approaches for the construction of strains with high-performing phenotypes[1-3]. Although hampered by the complex cellular network of metabolic pathways and their multiple regulation, nontargeted strain improvement allows one to access an enormous number of genotypes if a proper selective pressure is applied, and to screen for desired and robust strains even in the absence of detailed knowledge on the genetic basis or physiology of the producer strains[4, 5]. However, while nontargeted evolution has some major advantages over classical genetic engineering, it also has some inherent limitations[6]. Due to the randomness of mutagenesis, strains developed using nontargeted strategies generally have an unclear genetic background, resulting in challenging and slow further improvement, especially in the later stages of strain improvement. In addition, through traditional random mutagenesis often introduces undesirable mutation into the genome, resulting in strains that do not offer the best possible performance[7].Moreover, the resulting unknown genetic background precludes the rapid transfer of positive mutations to other strains or species.

With the development of systems biology tools, in particular next generation sequencing technology, increasing numbers of potentially valuable mutations are being identified[8, 9]. Comparative genomic analysis of superior production strains with their parental ancestors can identify the underlying

mutations, greatly increasing our understanding of the high-yield phenotype. This lays a foundation for avoiding unnecessary harmful mutations and identifying beneficial changes, finally making it possible to transfer these beneficial mutations to other strains and generate a more robust overproducer[10, 11]. However, it is still difficult to identify epistatic interactions between the large number of mutations, especially in strains derived from random mutagenesis. For a comprehensive mutation analysis, it is necessary to assess transcriptional changes to obtain a broader understanding of the relationship between genes affected by the mutations[12]. Through complete genome and transcriptome sequencing, we were able to obtain comprehensive profiles of genomic changes, encompassing point mutations and indels, as well as transcriptomic changes, to guide the subsequent mutation analysis.

L-tryptophan (L-trp) is a nutritionally essential amino acid as well as a key precursor of various chemicals and bioactive compounds of pharmaceutical interest[13, 14]. However, its biosynthesis requires several precursors, and the development of an efficient L-trp production strain by rational design and modification is challenging due to the interaction of several interrelated metabolic pathways and their complex regulation[15, 16]. During the past decades, continued efforts based on random mutagenesis and selection have improved the efficiency of L-trp production in *E. coli*[17-20]. Nevertheless, the random mutations generate an unclear genetic background, which hinders the ability to further explore their broader potential. Few studies to date have systematically analyzed the relationship between the L-trp overproduction phenotype and the genetic characteristics of *E. coli*. The objective of this study was to identify the specific genetic elements responsible for the L-trp overproduction phenotype, and thus gain insights into the mechanisms through which the L-trp overproducing mutants were derived.

Results and Discussion

Screening L-Trp hyperproducing strains using an intracellular L-Trp biosensor

The biosensor is a small-molecule inducible system that produces a fluorescent readout proportional to the amount of product inside of the cell[2]. Here, we used a previously characterized L-trp biosensor (Fig. 1a and Fig. S1)[21]. To test the applicability of this biosensor, we introduced it into *E. coli* DH5α and induced GFPmut2 expression by adding different concentrations of the Ala-Trp dipeptide in vitro. The results showed that there was a good linear relationship between the Ala-Trp concentration and the fluorescence intensity when the dipeptide concentration was less than 2 mM (Fig. S2). In addition, a clear difference in fluorescence intensity was observed in DH5α carrying the L-trp biosensor with and without the addition of Ala-Trp (Fig. S3). These results indicated that the sensor system was capable of reflecting the difference in L-trp concentration by altered fluorescence intensity. Furthermore, to confirm the feasibility of using the L-trp biosensor response to measure intracellular L-trp, we next investigated whether the biosensor could distinguish strains with different L-trp production capacities via different fluorescence signals. The L-trp sensor plasmid was introduced into the strains TPD1, TPD2, TPD3 and TPD4 with increased L-trp productivity (Table S1). As shown in Fig. S4, with the increased titer of L-trp from TPD1 to TPD4, there was an obvious increase in specific GFP fluorescence in the same order. Therefore, the L-trp biosensor showed a good correlation between L-trp concentration and fluorescence

intensity, both when L-trp was added at different concentrations in vitro and in strains with different L-trp production capacities, suggesting that this device could be used as a screening platform for selecting L-trp hyperproducing strains.

To establish a high-throughput screening technique, we introduced the L-trp sensor into *E. coli* TPD4, which was obtained by deleting *trpR* and overexpressing the *trpEDCBA* operon, along with the *aroG* and *ppsA* genes in *E. coli* KW[22]. To obtain an L-trp hyperproducing strain, a random mutagenesis library of the TPD4 strain was constructed using exposure to atmospheric and room-temperature plasma (ARTP) for 16 s (Fig. S5). In the control group, colonies without ARTP treatment were cultured under the same conditions for comparison. After ARTP mutagenesis and FACS screening, two mutant strains were finally selected from a library of 2×10^5 clones. The isolated strains were cultured in fermentation medium and the L-trp concentrations were determined by HPLC. Strains TPD5 and TPD6 showed similar growth characteristics (Fig. 1b), but respectively produced 3.08 and 1.46 times more L-trp than the parental strain TPD4 (Fig. 1c). In addition, FACS spectrograms of TPD4 and TPD5 showing effective separation of the strains confirmed the functionality of the L-trp biosensor (Fig. 1d), which was consistent with their L-trp production. Using this approach, we successfully increased L-trp biosynthesis to a titer of 2.01 g/L and a yield of 0.124 g/g.

Identification of genetic changes in strain TPD5 by whole-genome sequencing and identification of mutations beneficial for L-Trp overproduction

Whole-genome re-sequencing technology enables the complete characterization of mutant genomes and facilitates the identification of relevant mutations for different phenotypes. Compared with the parental strain TPD4, the screened *E. coli* TPD5 exhibited a 3.08-fold increase of L-trp titer, which greatly improved the L-trp productivity. To determine the genetic basis of the strain's L-trp overproduction phenotype, the whole genome of stain TPD5 was sequenced and re-annotated (see Supplementary Table S3). Combined with mutations found in gaps, a total of 38 mutations were identified across the genome of TPD5, encompassing 3 insertions, 5 deletions, and 23 substitutions. A total of 29 mutations were located within ORFs and 9 in non-ORF regions (Supplementary Tables S4 and S5). Among these mutations, only a small number affected loci with annotated functions related to the L-trp biosynthesis pathway. These include *mlc* (deletion) related to the phosphotransferase transport system (PTS), AcnA (S522G) in the TCA cycle, as well as AroG (S211F), TrpE (A63V) and *tnaA* (deletion) in the L-trp biosynthesis pathway. Some mutations were thought to occupy sites of interest that may be correlated with the strain's phenotype. For example, FlhD (V84F), encoding a flagellar transcriptional regulator, and AcrR (D139N), encoding a HTH-type transcriptional regulator, both of which are involved in the transcriptional regulation of several flagellar operons and modulate swimming motility, and therefore might be responsible for the physiological changes in the cell.

In order to identify genome changes responsible for L-trp overproduction, 28 mutations located in ORFs that alter amino acid sequences were selected as genetic manipulation targets for the introduction of individual back mutations into the chromosome of the parent strain TPD5 to evaluate their effects on L-

trp overproduction. Shake-flask fermentations indicated that a number of these mutations are actually neutral or even harmful (Fig. 2). For instance, AcrR (D139N), ArsB (E303K), GlpQ (R101C), DinD (P7S), OppF (S325A), EntE (A407T), GarL (A190S), DicA (E123K), PolA (G763S), Imp (E288K), LacI (deletion), YdeI (deletion), YgiQ (deletion) and YlbE (insertion) were neutral mutations, with no observable phenotypic changes in the corresponding mutants. PuuP (Y110C) behaved as a harmful mutation, and L-trp production increased to 2.4 g/L after this gene was restored, which may be related to its role as an important growth factor in the synthesis of proteins and nucleic acids. It is likely responsible for the low efficiency in transforming putrescine. In addition to the confirmed neutral and harmful mutations, at least 13 specific mutations were identified as beneficial for L-trp overproduction, with varying contributions.

Among the 13 beneficial mutations, four changes led to a significant decrease of L-trp overproduction after back mutation, including AroG (S211F), RpoS (nonsense mutation Q33*), TrpE (A63V) and TnaA (deletion). Strain TPD5-RM2, in which the AroG (S211F) mutation was restored, reintroducing the wild-type 3-deoxy-D-arabino-heptulosonate-7-phospphate (DAHP) synthase, showed a decrease of L-trp production from 2.01 g/L to 0.86 g/L. The L-trp production of strain TPD5-RM10 was even reduced to 0.09 g/L after restoring the wild-type anthranilate synthase TrpE. Strain TPD5-RM27, in which the degradation pathway tryptophanase TnaA was restored, showed a decrease of L-trp production to 0.9 g/L. In addition, restoring the Q33* nonsense mutation of RpoS, which is a global regulator, reduced the L-trp titer almost 3.3-fold to 0.6 g/L. Our results confirmed that only four mutations among the selected ones were significantly beneficial for L-trp production, accounting for a small part of the total mutations.

Insights into the AroG S180F mutation that leads to L-Trp overproduction

AroG is a DAHP synthetase isozyme that catalyzes the first committed step in the biosynthesis of aromatic amino acids and vitamins. It condenses the two important precursors phosphoenolpyruvate (PEP) and erythrose-4-phosphate (E4P) to form DAHP[23]. However, feedback inhibition by L-phenylalanine (L-phe) results in insufficient DAHP synthase activity to efficiently channel PEP and E4P into the L-trp biosynthesis pathway[24, 25]. In order to remove the feedback inhibition of AroG enzyme, Ger et al. isolated the mutant S180F through screening of phe-resistant colonies[26]. We also found that AroG in the screened strain TPD5 was mutated, but not at the residue S180. Instead, the residue S211 was mutated to F.

As can be seen in Fig. 2, AroG (S211F) had a significant impact on the production of L-trp. In order to explore whether S211F relieved the L-phe feedback inhibition of AroG in the same way as S180F, we analyzed the effect of the mutations at residue 118 and residue 211 on the binding of L-phe. Fig. 3a shows that under normal conditions, the inhibitor L-phe interacts directly with residue 180 (S) and residue 6 (Q) of wild-type AroG. When S180 was mutated to F, the phenyl ring of F blocked the binding of the inhibitor L-phe (Fig. 3b), thus releasing L-phe inhibition. Similarly, although there is no direct interaction between S211 and the inhibitor L-phe, replacement of S211 with F also blocks the binding of the inhibitor L-phe via steric hindrance from the phenyl ring (Fig. 3c). Therefore, AroG (S211F) in strain TPD5 relieved the feedback inhibition of L-phe in the same way as the previously reported S180F mutant, increasing the

conversion of PEP and E4P into DAHP, which increased the metabolic flux into the shikimate pathway. The feedback-resistant AroG (S211F) may be a good choice for the modification of other strains that produce aromatic amino acids or vitamins.

Insights into the TrpE A63V mutation that leads to L-Trp overproduction

Anthranilate synthetase (TrpE), which produces anthranilate from chorismate and glutamine in the first step of the terminal branch of the L-trp pathway in *E. coli*, is subject to feedback inhibition by the end-product L-trp[27, 28]. However, considering that TrpE is a key enzyme in the L-trp branch of the pathway, releasing its feedback inhibition is crucial for increasing L-trp production. In order to explore whether TrpE (A63V) is resistant to L-trp feedback-inhibition, we carried out molecular dynamics (MD) simulations of wild-type TrpE and mutant TrpE (A63V). The wild-type TrpE protein structure for MD simulations was constructed using the network server I-TASSER[29], on which Schrödinger[30] was used to construct the initial structure of the mutant TrpE (A63V). The binding sites of the substrate chorismate and the inhibitor L-trp after docking are shown in Fig. 4a.

In the absence of L-trp, the wild-type TrpE protein was in its normal, closed conformation (as shown in Fig. 4b, blue). Accordingly, L-trp bound to a specific position of the TrpE protein, resulting in structural changes causing its active-site pocket to shift from a closed conformation to an open conformation (as shown in Fig. 4b, yellow). In the open conformation, TrpE could not bind the substrate chorismate, thus preventing its normal catalytic function. By contrast, the structure of the mutant TrpE (A63V) did not change and remained in a closed conformation regardless of the presence (Fig. 4c, green) or absence (Fig. 4c, pink) of L-trp. This conformation was the same as that of wild-type TrpE performing the normal catalytic function (Fig. 4c, blue). Thus, the feedback-resistance of TrpE was achieved by replacing the residue A63 with V, explaining the reason for the great change in L-trp production caused by this mutation. Here, the mechanism of TrpE inhibition by L-trp was analyzed, which enables us to clearly understand the mechanism of feedback inhibition, laying a foundation for further improvement of TrpE protein in the future.

Effects of the RpoS nonsense mutation Q33* on L-Trp production

RpoS encodes the alternative sigma factor σ^s, a subunit of RNA polymerase that acts as the master regulator of many stationary-phase genes for adaptation to nutrient deprivation and other stresses in *E. coli*[31]. Genome-wide analyses of RpoS-dependent gene expression showed that up to 10% of genes in *E. coli* are under direct or indirect control of RpoS, among which over 130 genes are positively controlled and a surprisingly large number are negatively regulated by this sigma factor[32]. Although the RpoS regulon is a large, conserved system that is critical for adaptation to a variety of stresses and metabolic conditions, its effect on specific metabolic pathways, such as L-trp biosynthesis pathway, remains incompletely characterized. In order to determine the effects of the nonsense mutation RpoS (Q33*) on the increase of L-trp production, we fermented strains TPD5 and TPD5-RM3 and observed the differences in the fermentation process.

As can be seen from the fermentation results, when RpoS was restored, the color of the fermentation broth and cell pellet changed from light brown to pink (Fig. 5a and Fig. S6). In addition, the cell morphology at different fermentation stages was observed by optical microscopy, and it was found that the cells of TPD5-RM3 deteriorated and became abnormal and irregular from 12 h (Fig. S7). Furthermore, cells of TPD5 and TPD5-RM3 fermented for 12 h and 38 h were observed by scanning electron microscopy. As can be seen in Fig. 5b and c, strain TPD5-RM3 exhibited obvious cell surface shrinkage after fermentation for 12 h, and obvious holes appeared on the surface of a small number of cells, which may lead to the leakage of intracellular lysates, leading to the adhesion of cells. This phenomenon was more obvious at 38 h of fermentation, and almost all the cells of strain TPD5-RM3 collapsed, presenting a withered state. Compared with strain TPD5, in which only individual cells showed surface shrinkage, strain TPD5-RM3 showed dramatically reduced vitality. The normal expression of RpoS resulted in an observable phenotypic alteration of the L-trp producing strain, which was specifically visible in the color of the fermentation broth, cell morphology, cell size, etc. This change may be due to the direct or indirect regulation of some cell membrane proteins by RpoS[33]. Moreover, the premature termination of RpoS translation due to the nonsense mutation Q33* impaired cell growth on glucose (Fig. 5d, left) and strongly altered the distribution of intracellular fluxes. After RpoS was recovered, the production of L-trp decreased significantly to only 0.42 g/L, representing a 79.93% reduction (Fig. 5d, middle), while the concentration of acetic acid as a by-product in strain TPD5-RM3 was significantly increased to 2.2 times that of strain TPD5 (Fig. 5e, right). This difference indicates that the carbon flow of cells was affected by RpoS, redirecting more metabolic flux to by-product biosynthesis. Therefore, it can be inferred that RpoS not only has a direct impact on the overall morphology of cells on the macro level, but also significantly influences the carbon flow of strains on the micro level.

Transcriptomic analysis of the role of the global regulator RpoS in the L-Trp biosynthesis pathway

In order to obtained deeper insights into the genetic link between RpoS and the L-trp biosynthesis pathway, comparative transcriptomic analysis was conducted with cells of *E. coli* TPD5 and TPD5-RM3 harvested during the stationary growth phase. Fig. 6a shows the heatmap of differentially expressed genes, from which it can be seen that the genes expressed in strain TPD5-RM3 were significantly different from those expressed in strain TPD5. A total of 469 genes showed significant variation at the genome-wide transcriptional level between TPD5 and TPD5-RM3 (163 up- and 306 downregulated genes; Fig. 6b and Supplementary Table S6). The correlation of gene expressing levels between the two strains was 0.846 (Fig. S8). Gene Ontology (GO) functional annotation and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis were carried out. GO functional annotation classified the differentially expressed genes into three categories: biological process, cellular component, and molecular function, while KEGG pathway analysis classified the differential genes into 20 categories with the most significant difference in biosynthesis of secondary metabolites (Fig. 6c, d and Supplementary Table S7).

To determine the functional classes of the differentially expressed genes in stain TPD5-RM3, we carried out a homology-based annotation specifically for all the genes and GO terms related to the dataset

shown in Table S7. Significant differences were found in the expression of genes in the categories biological process and molecular function (Fig. 6c). In particular, the enrichment of terms related to metabolic process were distinct (GO:0008152, GO:0009987, GO:0071704, GO:0044238, GO:0044237), indicating that RpoS has an extensive effect on gene expression in metabolic pathways, thus directly affecting the production of most primary metabolites. Moreover, the set of differentially expressed genes was mapped to the KEGG pathways in *E. coli* W3110 (Supplementary Table S8), showing that a series of enriched pathways exhibited significant differential expression. The transcripts of genes involved in the biosynthesis of secondary metabolites, amino acid and carbon metabolites were significantly different from those in strain TPD5, and thus may directly or indirectly affect the synthesis of L-trp.

The differentially expressed genes nearly all fell into different functional pathways, including biosynthesis of secondary metabolites (49 genes), carbon metabolism (31 genes), biosynthesis of amino acids (30 genes), purine metabolism (19 genes), glycolysis/gluconeogenesis (11 genes), pentose phosphate pathway (8 genes), as well as glycine, serine and threonine metabolism (8 genes). To further analyze the differential expression information, we first focused on genes in the pathways directly related to L-trp biosynthesis (Fig. 7). In the L-trp biosynthesis pathway, transcription of the trp genes, which are organized in a single operon, was strongly upregulated 1.54- (*trpE*), 2.34- (*trpD*), 3.43- (*trpC*), 2.43- (*trpB*) and 3.36-fold (*trpA*) in *E. coli* TPD5-RM3. In contrast to the higher expression of the trp operon in *E. coli* TPD5-RM3, most genes involved in precursor supply (PEP, E4P, PRPP (5-phospho-α-D-ribose 1-diphosphate), glutamine, L-serine (L-ser)) presented lower expression levels. These included *pgi*, *pfkA/B*, *fbaA*, *gapA*, *pgk*, *gpmA/M*, and *eno* in the glycolysis pathway, *ppsA* and *fbp* in the gluconeogenesis pathway, as well as *zwf*, *gnd*, rpiA/B, *tktB*, *talB*, *deoB* and *phnN* in the pentose phosphate pathway and almost all genes in the TCA cycle. Thus, the insufficiency of crucial precursors, energy or cofactors derived from the central metabolic pathways may have restricted biomass accumulation and L-trp production. Moreover, pathway enrichment analysis also revealed that most genes involved in the glucose absorption system, including *ptsH*, *ptsI*, *crr*, *frsA*, *hfq*, *sgrR*, *mlc*, *crp*, and *relA-uspD*, showed low expression levels in *E. coli* TPD5-RM3, leading to inadequate nutritional supply. Overall, RpoS attenuates glucose uptake at the beginning of the entire metabolic pathway by reducing the expression level of genes related to the glucose uptake system. Furthermore, the transcription levels of most genes in the glycolysis pathway and pentose phosphate pathway were reduced as a whole, leading to severely weakened metabolic flux in the central metabolic pathways under the premise of inadequate nutrient uptake, which made the cells unable to produce sufficient precursors, energy, and cofactors for L-trp biosynthesis. Therefore, the production of L-trp was very low even though the gene expression intensity of the terminal branch of the L-trp pathway was strong.

Metabolic engineering of TPD5-RM4 for improved L-Trp production

The terminal branch of the L-trp pathway starts from the condensation of chorismate and glutamine and ends with the condensation of indole and L-ser to L-trp (Fig. 1a). However, after the analysis of strain TPD5-RM4, it was found that the concentration of L-ser and its substrate phosphoserine (P-ser) was low, which may be caused by feedback inhibition of the enzyme D-3-phosphoglycerate dehydrogenase (3-

PGDH, encoded by *serA*), leading to insufficient supply of the precursors. In addition, considering that no mutations or manipulations have been made in the L-ser biosynthesis pathway in strain TPD5, the H344A/N364A double mutant of the key enzyme SerA was expressed from the chromosome of strain TPD5-RM4, resulting in strain TPD5-RM4S, which is resistant to feedback inhibition by L-ser. After the overexpression of SerA (H344A/N364A), the concentration of P-ser and L-ser increased significantly (Fig. 8a). As expected, the production of the final product L-trp was further increased to 2.94 g/L (Fig. 8b).

Although the production of L-trp was improved by overexpressing SerA (H344A/N364A), we found that the concentration of glutamate in strain TPD5-RM4S greatly increased to 8.86 mM, which was 368-fold higher than that of strain TPD5-RM4 (Fig. 8a). Glutamate can be converted by glutamine synthetase GlnA to produce glutamine, which is a crucial precursor involved in the first step of the terminal branch of L-trp biosynthesis. Therefore, in order to further expand the metabolic flux of the L-trp pathway, we attempted to convert the large amount of glutamate of strain TPD5-RM4S into glutamine, thus further promoting the synthesis of L-trp. Accordingly, we introduced GlnA (Y405F)[34] from *Corynebacterium glutamate* into TPD5-RM4S to construct stain TPD5-RM4G. Compared with strain TPD5-RM4S, the concentration of glutamate in strain TPD5-RM4G was reduced to 0.24 mM, successfully transforming the accumulated glutamate into glutamine. The final TPD5-RM4G strain produced L-trp at a titer of 3.18 g/L in shake flask fermentations. Thus, we accomplished a further 57.83% increase of the L-trp titer compared with strain TPD5 by altering the expression levels of genes in the L-ser pathway and boosting the conversion of glutamate to glutamine.

Conclusion

The genetic analysis of overproducing strains with an unclear background is crucial for understanding and conferring a high-performance phenotype on industrial microorganisms. A number of new mutations with the potential to be transferred to construct new overproducer strains for other biochemicals were identified. The comprehensive analysis of the feedback-resistance mechanism of the mutants AroG (S221F) and TrpE (A63V) not only helps us understand the proteins better at a theoretical level, but also provides the possibility to further improve their performances for broader applications. The feedback-inhibition mechanism of TrpE was analyzed in depth, providing a new strategy for the production of many valuable chemicals whose biosynthetic pathway is strictly controlled by a feedback inhibition mechanism. Notably, the global regulator RpoS was found to not only induce morphological changes and modulate cell growth, but also regulate the L-trp biosynthesis pathway. Therefore, researchers concerned with engineering the metabolic pathway of a target product should also take into account global regulators, especially when the modification enters a bottleneck stage.

Materials and Methods

Bacterial strains, plasmids and culture conditions

The strains and plasmids used in the study are listed in supplementary Table S1, And the primers are listed in supplementary Table S2.

Biological replicates were defined as samples inoculated from distinct colonies. The cultures were grown in lysogeny broth (LB) and M9 minimal medium.

Construction of the intracellular L-Trp biosensor

The plasmid ptrc99a-tnaC encoding the L-trp sensor contained GFPmut2 fused downstream of the *tnaCAB* regulatory element *tnaC*, which was amplified from *E. coli* MG1655 genomic DNA with tnaC-F and tnaC-R primers. TnaC senses the intracellular L-trp concentration and converts the input signal into a fluorescence output signal.

Measurement of fluorescence

A single colony was cultured in 5 mL of LB medium overnight at 37 °C. The resulting seed culture was used to inoculate 5 mL of M9 medium at a ratio of 1:50 LB and incubated at 37 °C and 200 rpm. When the OD_{600} of the culture reached 0.4, 1 mM IPTG was added to induce expression of the biosensor, and various amounts of ala-trp dipeptide were then added to individual culture samples to control the intracellular L-trp concentration. After 5 h of incubation, the fluorescence (483 nm excitation, 525 nm emission for GFPmut2) of 200 µl of the diluted cultures was monitored on a multimode microplate reader.

Molecular docking and molecular dynamic simulations

The crystal structure of AroG from *E. coli* (PDB ID: 5CKS) was downloaded from the Protein Data Bank (PDB)[35], and the 3D complex structure of wild-type TrpE was constructed using I-TASSER[36]. The Glide algorithm[37] was applied to dock L-phe and L-trp into AroG and TrpE, respectively. The 3D models of the AroG (S211F) and TrpE (A63V) mutants were constructed based on the wild-type structure using Schrödinger 2018. Then, the 3D models of wild-type TrpE and TrpE (A63V) were used as the initial complex for MD simulations. First, energy minimization of the constructed model was conducted using AMBER18[30], which was also used to carry out molecular dynamics simulation of the final model using the ff14SB and gaff2 force field with the hydrogen mass repartitioning (HMR) method[38]. After heating and equilibration, a 500 ns MD simulation was executed without any restrictions. The complete simulation methodology used in this work is available in the supporting information. The MD trajectories were collected for further analysis and identification of relevant binding poses.

Fermentation conditions

Strains were cultured overnight at 37 °C in LB medium with appropriate antibiotics and then transferred into a 500 mL shake flask containing 30 mL of fermentation medium ($(NH_4)_2SO_4$ 10 g/L, KH_2PO_4 5 g/L, yeast extract 2 g/L, $FeSO_4{\cdot}7H_2O$ 15 mg/L, $MnSO_4{\cdot}H_2O$ 15 mg/L, $CuSO_4{\cdot}5H_2O$ 4 mg/L, $CoCl_2{\cdot}6H_2O$ 4 mg/L, $ZnSO_4{\cdot}7H_2O$ 4 mg/L, glycine betaine 1 g/L, biotin 30 µg/L, $MgSO_4{\cdot}7H_2O$ 1 g/l, and the pH value

was adjusted to 7.0) at a ratio of 1:20. Tetracycline was added at a final concentration of 15 mg/L to promote plasmid retention.

HPLC quantification of L-Trp

The cell culture was centrifuged at 12,000×g for 5 min, and the supernatant was passed through a 0.22 μm filter to obtain analysis samples. Amino acid analysis was carried out on an 1100 HPLC instrument (Agilent, Santa Clara, USA) equipped with a Zorbax Eclipse-AAA column. The mobile phase consisted of two solvent A (40 mM Na_2HPO_4, pH 7.8) and solvent B (ACN: MeOH: water =45:45:10, v/v/v) at a flow rate of 2 mL/min, in gradient elution mode as follows: 0-1 min, 100% A; 9.8 min: 43% A +57% B; 10 min: 100% B; 12 min: 100% B; 12.5 min: 100% A. The retention time of L-trp was 26 min.

Library preparation for strand-specific transcriptome sequencing

Cells from 5 mL of culture grown in 100 mL of fermentation medium were collected for RNA isolation in the stationary phase. RNA degradation and contamination were monitored on 1% agarose gels. RNA purity was checked using the NanoDrop® spectrophotometer (IMPLEN, CA, USA). The RNA concentration was measured using a Qubit® RNA Assay Kit in a Qubit® 2.0 Fluorometer (Life Technologies, CA, USA). RNA integrity was assessed using the RNA Nano 6000 Assay Kit of the Bioanalyzer 2100 system (Agilent Technologies, CA, USA). A total amount of 3 μg RNA per sample was used as input material. Sequencing libraries were generated using the NEBNext® Ultra™ Directional RNA Library Prep Kit for Illumina® (NEB, USA) following the manufacturer's recommendations, and index codes were added to attribute sequences to each sample.

Declarations

Acknowledgements

Not applicable.

Authors' contributions

D.D. and D.Z. initiated the project. D.D. and D.B. performed the experiments. J.L. performed the computational work. D.D., D.B., J.L., Z.M., JP.L., H.M. and D.Z. analyzed data and wrote the manuscript, which was revised and approved by all authors.

Funding

This work was supported by the National Key R&D Program of China (2020YFA0907800), Tianjin Synthetic Biotechnology Innovation Capacity Improvement Project (TSBICIP-KJGG-004-03), the National Natural Science Foundation of China (NSFC 31800086), National Key R&D Plan Special Project for "Synthetic biology" (2018YFA0901402, 2019YFA0904901).

Availability of data and materials

All data generated or analyzed during this study are included in this published article [and its supplementary information files].

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

[1] Tianjin Institutes of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin, 300308, P.R. China. [2] Key Laboratories of Systems Microbial Biotechnology, Chinese Academy of Sciences, Tianjin, 300308, P. R. China. [3] National Technology Innovation Center of Synthetic Biology, Tianjin, 300308, P. R. China. [4] Biodesign Center, Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin, 300308, P. R. China. [5] Biodesign Center, Key Laboratory of Systems Microbial Biotechnology, Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin, 300308, China. [6] College of Pharmacy, Nankai University, Haihe Education Park, 38 Tongyan Road, Tianjin 300353, P. R. China.

References

1. Yu AQ, Pratomo Juwono NK, Foo JL, Leong SSJ, Chang MW. Metabolic engineering of *Saccharomyces cerevisiae* for the overproduction of short branched-chain fatty acids. Metab Eng. 2016; 34:36-43.
2. Rogers JK, Church GM. Genetically encoded sensors enable real-time observation of metabolite production. Proc Natl Acad Sci USA. 2016; 113:2388-2393.
3. Pontrelli S, Fricke RCB, Teoh ST, Lavina WA, Putri SP, Fitz-Gibbon S, et al. Metabolic repair through emergence of new pathways in *Escherichia coli*. Nat Chem Biol. 2018; 14:1005-1009.
4. Barata D, van Blitterswijk C, Habibovic P. High-throughput screening approaches and combinatorial development of biomaterials using microfluidics. Acta Biomater. 2016; 34:1-20.
5. Xiao Y, Bowen CH, Liu D, Zhang F. Exploiting nongenetic cell-to-cell variation for enhanced biosynthesis. Nat Chem Biol. 2016; 12:339-344.
6. Dragosits M, Mattanovich D. Adaptive laboratory evolution – principles and applications for biotechnology. Microb Cell Fact. 2013; 12:64.

7. Wang G, Shi T, Chen T, Wang X, Wang Y, Liu D, et al. Integrated whole-genome and transcriptome sequence analysis reveals the genetic characteristics of a riboflavin-overproducing *Bacillus subtilis*. Metab Eng. 2018; 48:138-149.
8. Liu L, Li Y, Li S, Hu N, He Y, Pong R, et al. Comparison of next-generation sequencing systems. J Biomed Biotechnol. 2012; 2012:251364.
9. Ansorge WJ. Next-generation DNA sequencing techniques. N Biotechnol. 2009; 25:195-203.
10. McKie SJ, Maxwell A, Neuman KC. Mapping DNA topoisomerase binding and cleavage genome wide using next-generation sequencing techniques. Genes (Basel). 2020; 11.
11. Hancock-Hanser BL, Frey A, Leslie MS, Dutton PH, Archer FI, Morin PA. Targeted multiplex next-generation sequencing: advances in techniques of mitochondrial and nuclear DNA sequencing for population genomics. Mol Ecol Resour. 2013; 13:254-268.
12. Schena M, Shalon D, Davis RW, Brown PO. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science (80-). 1995; 270:467-470.
13. Chen L, Zeng AP. Rational design and metabolic analysis of *Escherichia coli* for effective production of L-tryptophan at high concentration. Appl Microbiol Biotechnol. 2017; 101:559-568.
14. Porter RJ, Mulder RT, Joyce PR, Luty SE. Tryptophan and tyrosine availability and response to antidepressant treatment in major depression. J Affect Disord. 2005; 86:129-134.
15. Du L, Zhang Z, Xu Q, Chen N. Central metabolic pathway modification to improve L-tryptophan production in *Escherichia coli*. Bioengineered. 2019; 10:59-70.
16. Liu Q, Cheng Y, Xie X, Xu Q, Chen N. Modification of tryptophan transport system and its impact on production of L-tryptophan in *Escherichia coli*. Bioresour Technol. 2012; 114:549-554.
17. Wang J, Cheng LK, Wang J, Liu Q, Shen T, Chen N. Genetic engineering of *Escherichia coli* to enhance production of L-tryptophan. Appl Microbiol Biotechnol. 2013; 97:7587-7596.
18. Shen T, Liu Q, Xie X, Xu Q, Chen N. Improved production of tryptophan in genetically engineered *Escherichia coli* with TktA and PpsA overexpression. J Biomed Biotechnol. 2012; 2012:605219.
19. Cheng LK, Wang J, Xu QY, Zhao CG, Shen ZQ, Xie XX, et al. Strategy for pH control and pH feedback-controlled substrate feeding for high-level production of L-tryptophan by *Escherichia coli*. World J Microbiol Biotechnol. 2013; 29:883-890.
20. Dodge TC, Gerstner JM. Optimization of the glucose feed rate profile for the production of tryptophan from recombinant *E coli*. J Chem Technol Biotechnol. 2002; 77:1238-1245.
21. Fang M, Wang T, Zhang C, Bai J, Zheng X, Zhao X, et al. Intermediate-sensor assisted push-pull strategy and its application in heterologous deoxyviolacein production in *Escherichia coli*. Metab Eng. 2016; 33:41-51.
22. Chen Y, Liu Y, Ding D, Cong L, Zhang D. Rational design and analysis of an *Escherichia coli* strain for high-efficiency tryptophan production. J Ind Microbiol Biotechnol. 2018; 45:357-367.
23. Rodriguez A, Martinez JA, Flores N, Escalante A, Gosset G, Bolivar F. Engineering *Escherichia coli* to overproduce aromatic amino acids and derived compounds. Microb Cell Fact. 2014; 13:126.

24. Smith LC, Ravel JM, Lax SR, Shive W. The control of 3-deoxy-D-arabino-heptulosonic acid 7-phosphate synthesis by phenylalanine and tyrosine. J Biol Chem. 1962; 237:3566-3570.
25. Li H, Liang C, Chen W, Jin JM, Tang SY, Tao Y. Monitoring in vivo metabolic flux with a designed whole-cell metabolite biosensor of shikimic acid. Biosens Bioelectron. 2017; 98:457-465.
26. Ger YM, Chen SL, Chiang HJ, Shiuan D. A single Ser-180 mutation desensitizes feedback inhibition of the phenylalanine-sensitive 3-deoxy-D-arabino-heptulosonate 7-phosphate (DAHP) synthetase in *Escherichia coli*. J Biochem. 1994; 116:986-990.
27. Zhao ZJ, Zou C, Zhu YX, Dai J, Chen S, Wu D, et al. Development of L-tryptophan production strains by defined genetic modification in *Escherichia coli*. J Ind Microbiol Biotechnol. 2011; 38:1921-1929.
28. Fang MY, Zhang C, Yang S, Cui JY, Jiang PX, Lou K, et al. High crude violacein production from glucose by *Escherichia coli* engineered with interactive control of tryptophan pathway and violacein biosynthetic pathway. Microb Cell Fact. 2015; 14:8.
29. Zhang Y. I-TASSER server for protein 3D structure prediction. BMC Bioinformatics. 2008; 9:40.
30. D.A. Case IYB-S, S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III, V.W.D. Cruzeiro, T.A. Darden,, R.E. Duke DG, M.K. Gilson, H. Gohlke, A.W. Goetz, D. Greene, R Harris, N. Homeyer, Y. Huang,, S. Izadi AK, T. Kurtzman, T.S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D.J., Mermelstein KMM, Y. Miao, G. Monard, C. Nguyen, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R., Qi DRR, A. Roitberg, C. Sagui, S. Schott-Verdugo, J. Shen, C.L. Simmerling, J. Smith, R. Salomon-, Ferrer JS, R.C. Walker, J. Wang, H. Wei, R.M. Wolf, X. Wu, L. Xiao, D.M. York and P.A. Kollman, et al: 2018.
31. Landini P, Egli T, Wolf J, Lacour S. sigmaS, a major player in the response to environmental stresses in *Escherichia coli*: role, regulation and mechanisms of promoter recognition. Environ Microbiol Rep. 2014; 6:1-13.
32. Patten CL, Kirchhof MG, Schertzberg MR, Morton RA, Schellhorn HE. Microarray analysis of RpoS-mediated gene expression in *Escherichia coli* K-12. Mol Genet Genomics. 2004; 272:580-591.
33. Zhang Y, Shi C, Yu J, Ren J, Sun D. RpoS regulates a novel type of plasmid DNA transfer in *Escherichia coli*. PLoS ONE. 2012; 7:e33514.
34. Liu Q, Zhang J, Wei XX, Ouyang SP, Wu Q, Chen GQ. Microbial production of L -glutamate and L -glutamine by recombinant *Corynebacterium glutamicum* harboring Vitreoscilla hemoglobin gene vgb. Appl Microbiol Biotechnol. 2008; 77:1297-1304.
35. Burley SK, Berman HM, Bhikadiya C, Bi C, Chen L, Di Costanzo L, et al. RCSB Protein Data Bank: biological macromolecular structures enabling research and education in fundamental biology, biomedicine, biotechnology and energy. Nucleic Acids Res. 2019; 47:D464-D474.
36. Roy A, Kucukural A, Zhang Y. I-TASSER: a unified platform for automated protein structure and function prediction. Nat Protoc. 2010; 5:725-738.
37. Schrödinger Release 2018-1: Glide S, LLC, New York, NY: 2018.
38. Hopkins CW, Le Grand S, Walker RC, Roitberg AE. Long-time-step molecular dynamics through hydrogen mass repartitioning. J Chem Theory Comput. 2015; 11:1864-1874.

Figures

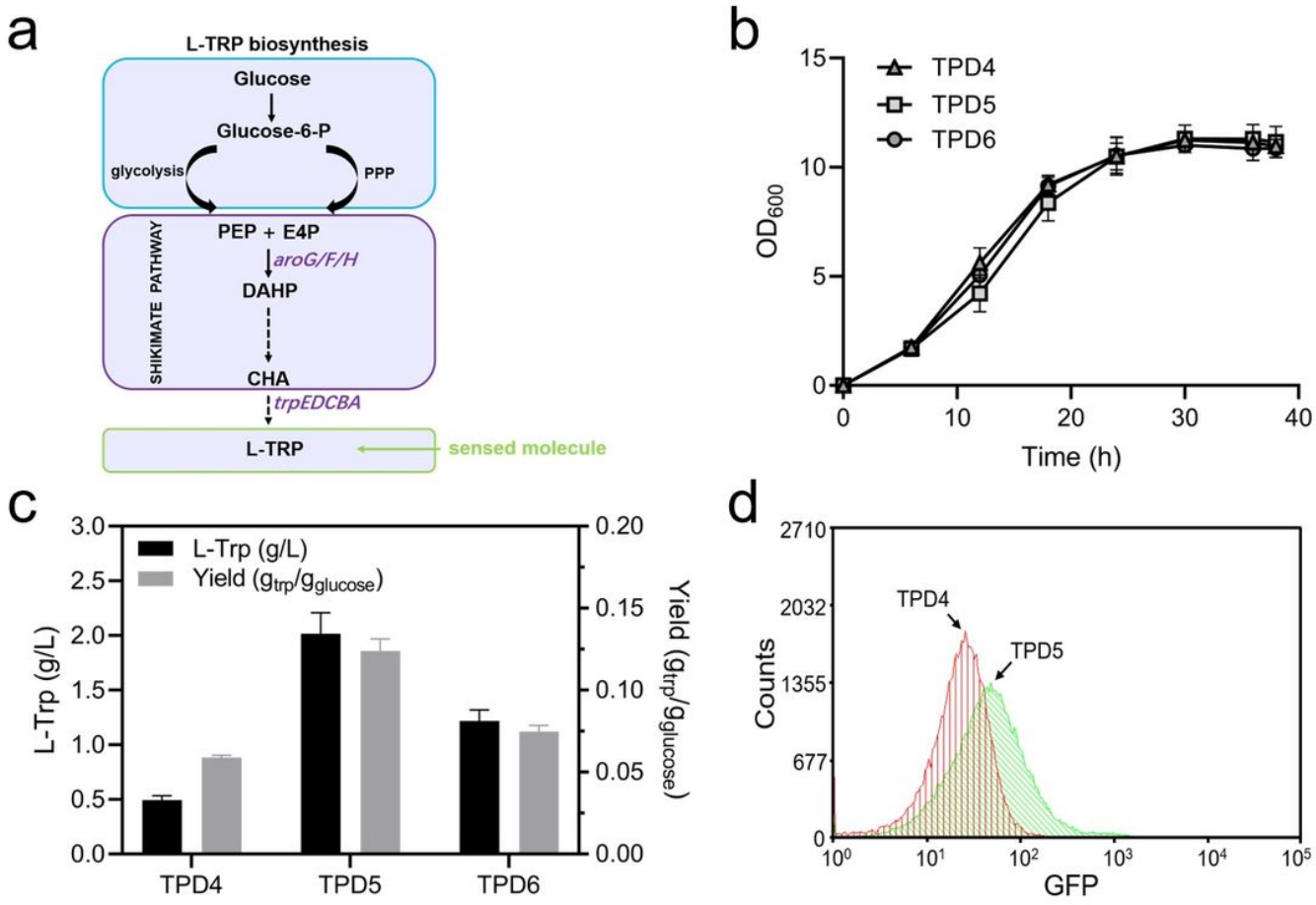

Figure 1

Optimization of L-trp biosynthesis using random mutagenesis and FACS-based selection via an intracellular L-trp biosensor. (a) L-trp is produced from glucose by converting PEP and E4P into DAHP and then onward to L-trp. PEP (phosphoenolpyruvate); E4P (erythrose-4-phosphate); DAHP (3-deoxy-D-arabino-heptulosonate 7-phosphate); CHA (chorismate). (b) Growth of strains TPD4, TPD5 and TPD6 in shake flasks. (c) L-trp production and glucose conversion rate of strains TPD4, TPD5 and TPD6 in shake flasks. (d) FACS discrimination between two L-trp producers (TPD4 and TPD5) using the intracellular L-trp biosensor as a screening tool. Experiments were conducted in triplicate and the data represent the means ± SD.

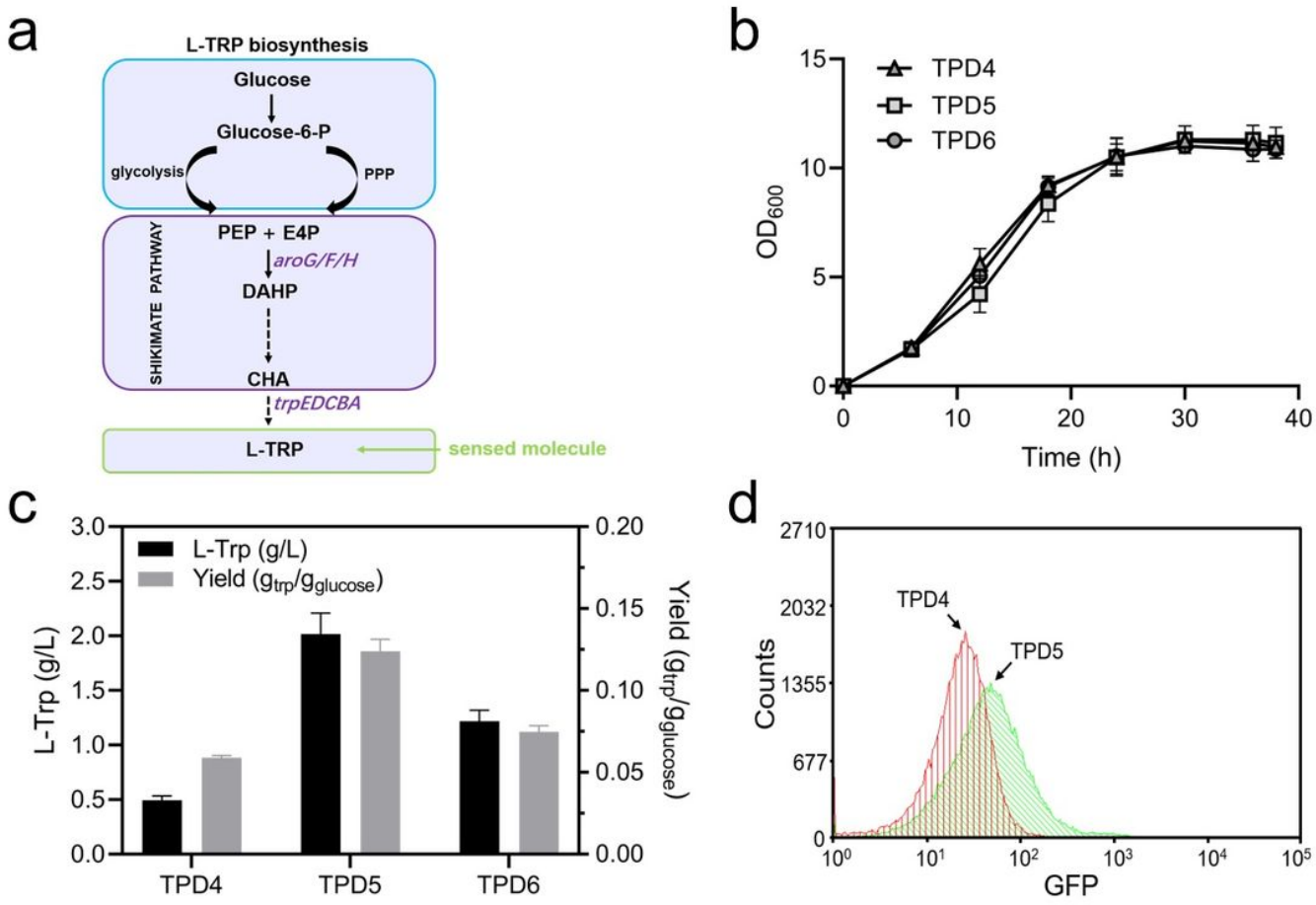

Figure 1

Optimization of L-trp biosynthesis using random mutagenesis and FACS-based selection via an intracellular L-trp biosensor. (a) L-trp is produced from glucose by converting PEP and E4P into DAHP and then onward to L-trp. PEP (phosphoenolpyruvate); E4P (erythrose-4-phosphate); DAHP (3-deoxy-D-arabino-heptulosonate 7-phosphate); CHA (chorismate). (b) Growth of strains TPD4, TPD5 and TPD6 in shake flasks. (c) L-trp production and glucose conversion rate of strains TPD4, TPD5 and TPD6 in shake flasks. (d) FACS discrimination between two L-trp producers (TPD4 and TPD5) using the intracellular L-trp biosensor as a screening tool. Experiments were conducted in triplicate and the data represent the means ± SD.

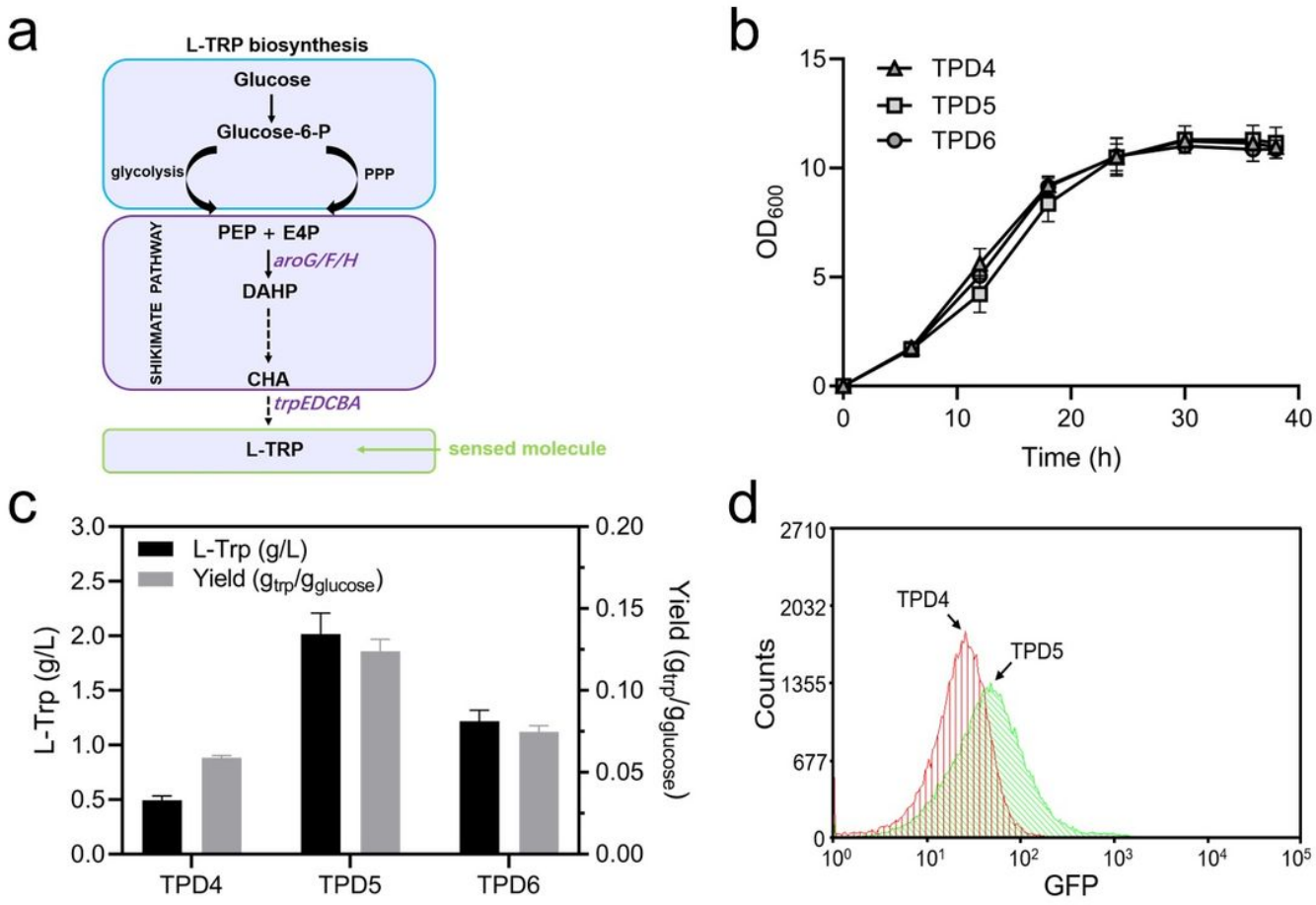

Figure 1

Optimization of L-trp biosynthesis using random mutagenesis and FACS-based selection via an intracellular L-trp biosensor. (a) L-trp is produced from glucose by converting PEP and E4P into DAHP and then onward to L-trp. PEP (phosphoenolpyruvate); E4P (erythrose-4-phosphate); DAHP (3-deoxy-D-arabino-heptulosonate 7-phosphate); CHA (chorismate). (b) Growth of strains TPD4, TPD5 and TPD6 in shake flasks. (c) L-trp production and glucose conversion rate of strains TPD4, TPD5 and TPD6 in shake flasks. (d) FACS discrimination between two L-trp producers (TPD4 and TPD5) using the intracellular L-trp biosensor as a screening tool. Experiments were conducted in triplicate and the data represent the means ± SD.

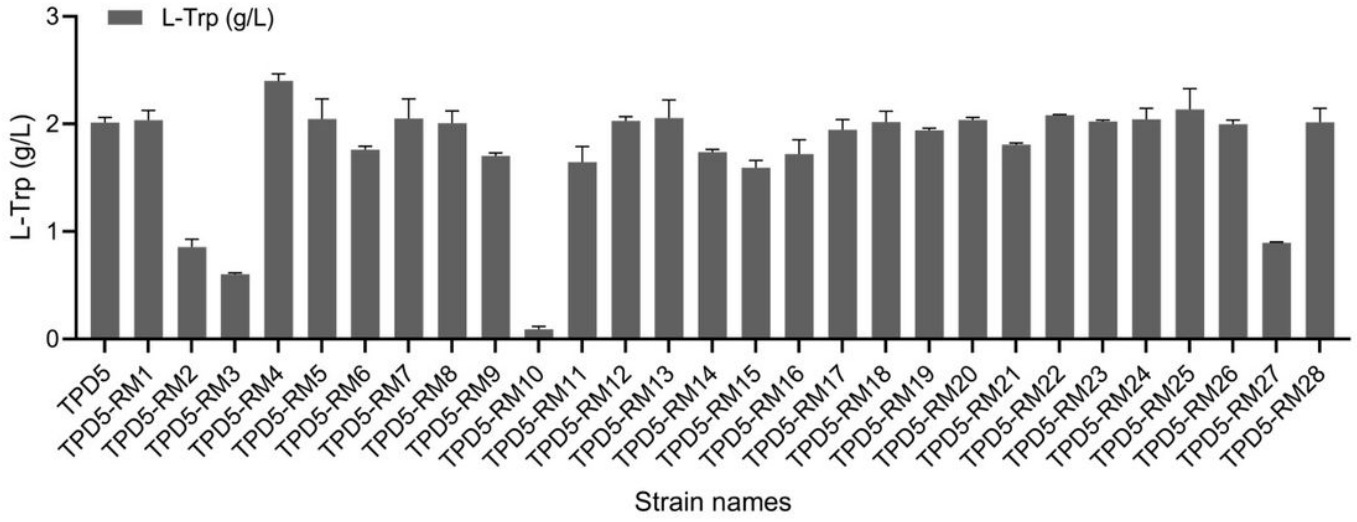

Figure 2

Fermentation results for strain TPD5 and the strains reconstructed by back mutation. Experiments were conducted in triplicate and the data represent the means ± SD.

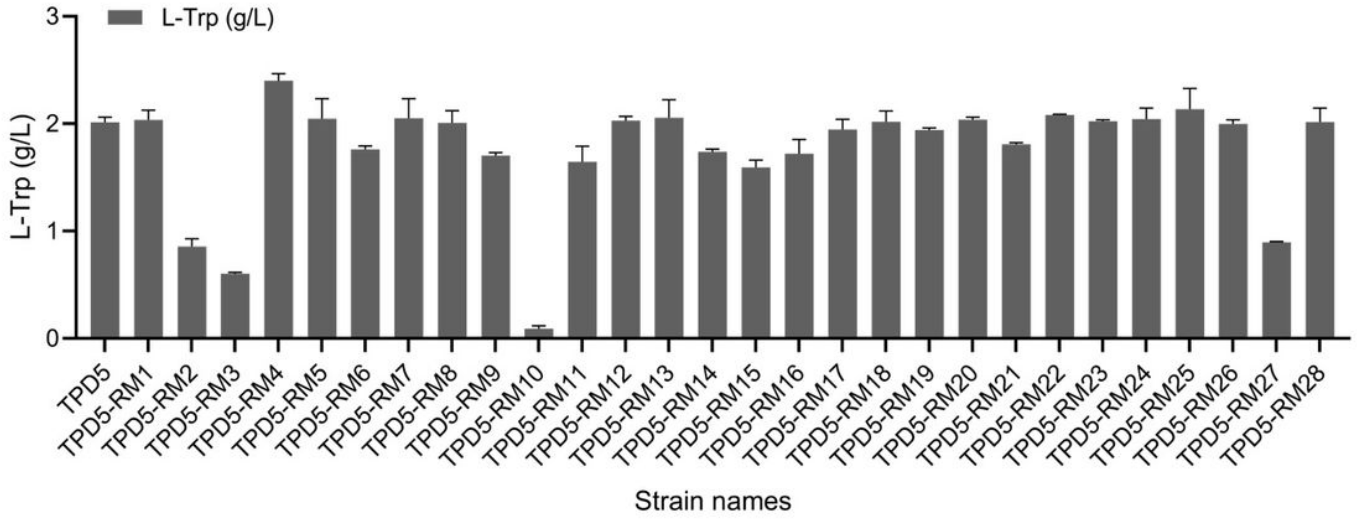

Figure 2

Fermentation results for strain TPD5 and the strains reconstructed by back mutation. Experiments were conducted in triplicate and the data represent the means ± SD.

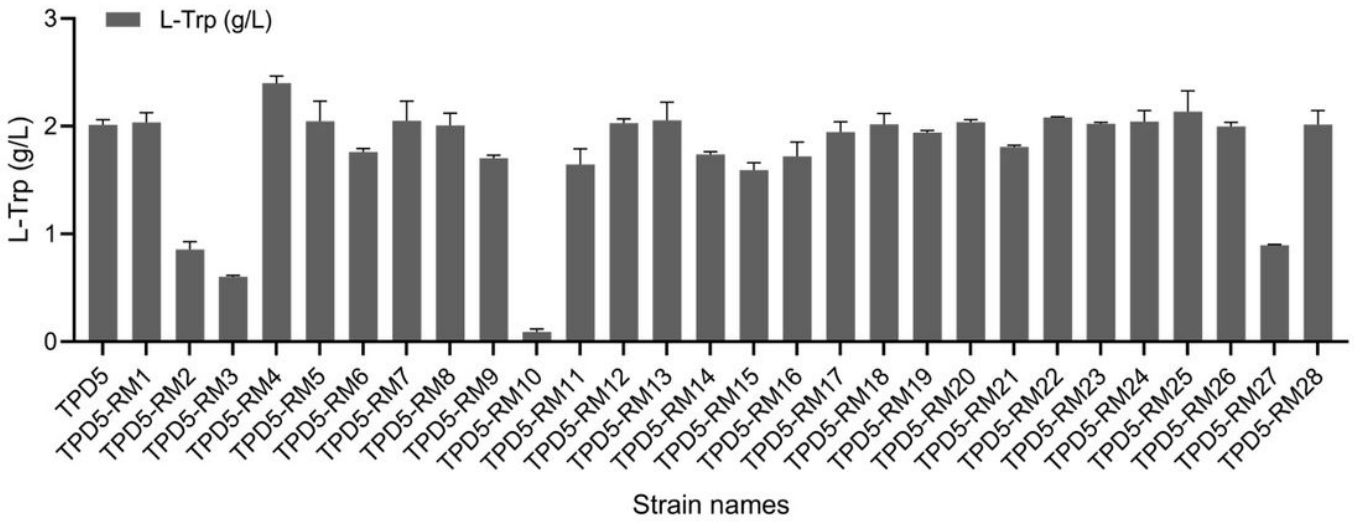

Figure 2

Fermentation results for strain TPD5 and the strains reconstructed by back mutation. Experiments were conducted in triplicate and the data represent the means ± SD.

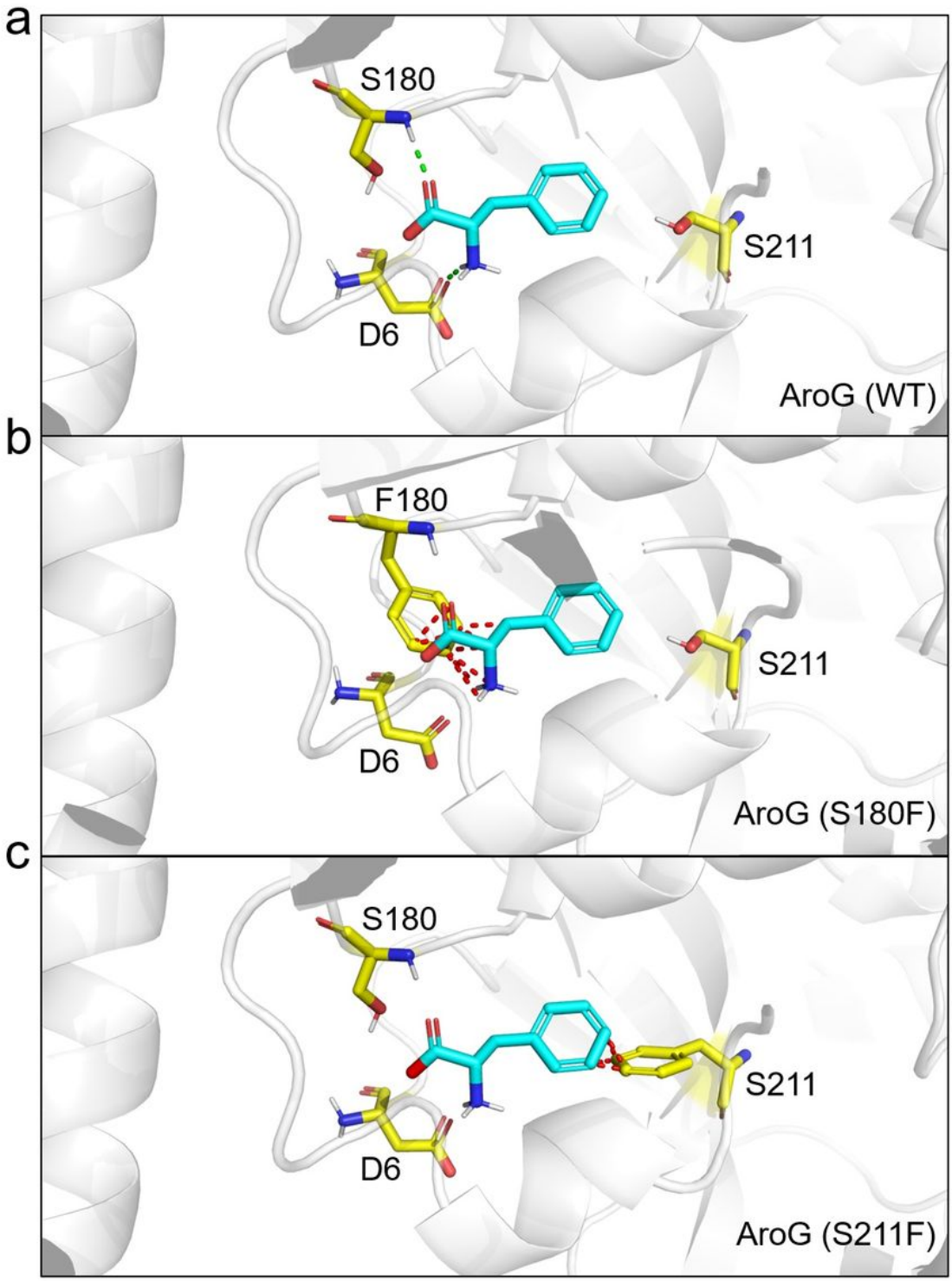

Figure 3

Binding mode of the inhibitor L-phe with AroG. (a) Binding mode of inhibitor L-phe with AroG (wild-type, WT) (PDB: 5CKS). Green dotted lines represent hydrogen-bonding interactions between L-phe and AroG (WT). After S180 (b) or S211 (c) was replaced with F, the binding of L-phe to AroG was blocked. The red dotted lines represent obstruction via steric hindrance.

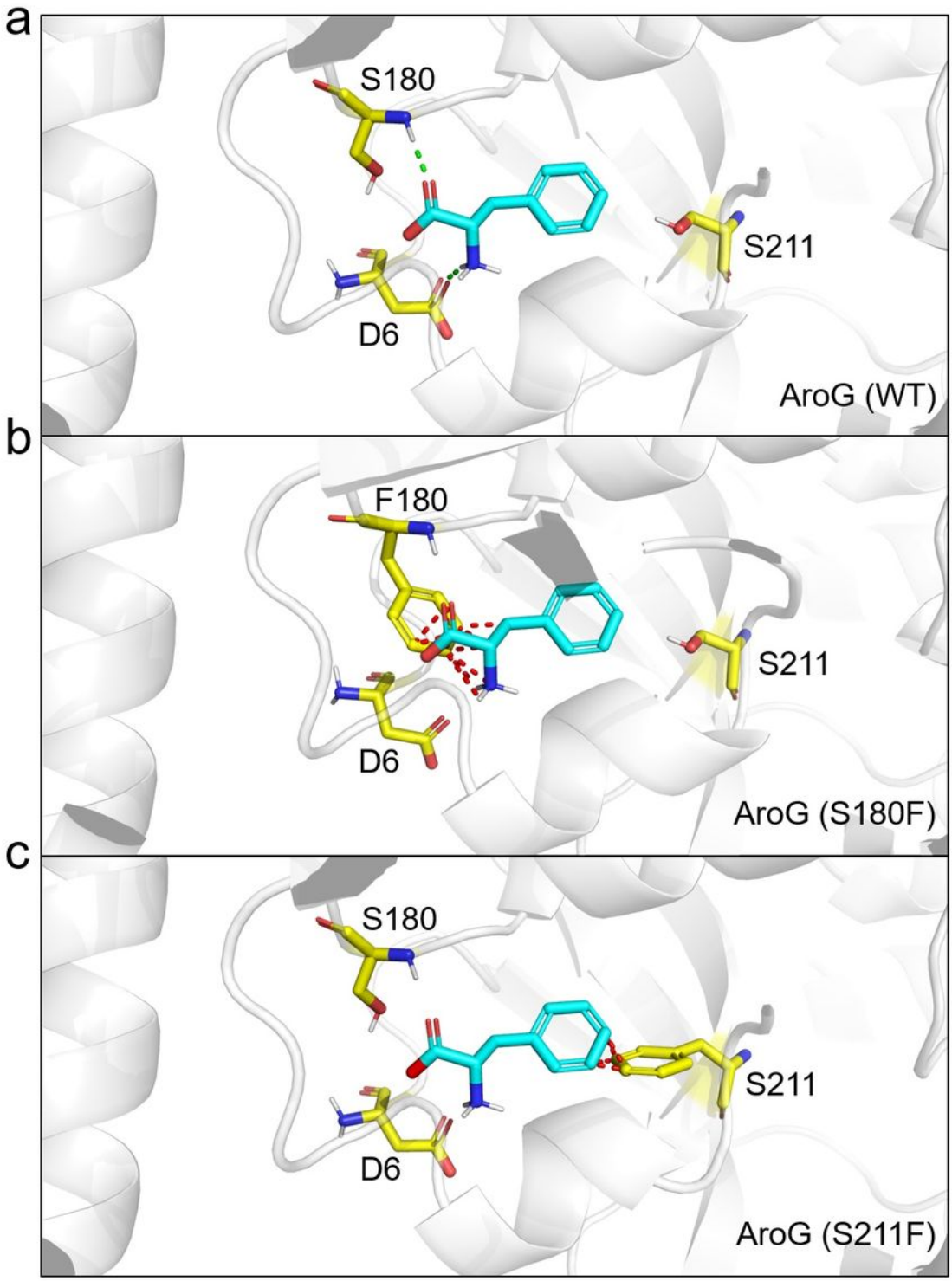

Figure 3

Binding mode of the inhibitor L-phe with AroG. (a) Binding mode of inhibitor L-phe with AroG (wild-type, WT) (PDB: 5CKS). Green dotted lines represent hydrogen-bonding interactions between L-phe and AroG (WT). After S180 (b) or S211 (c) was replaced with F, the binding of L-phe to AroG was blocked. The red dotted lines represent obstruction via steric hindrance.

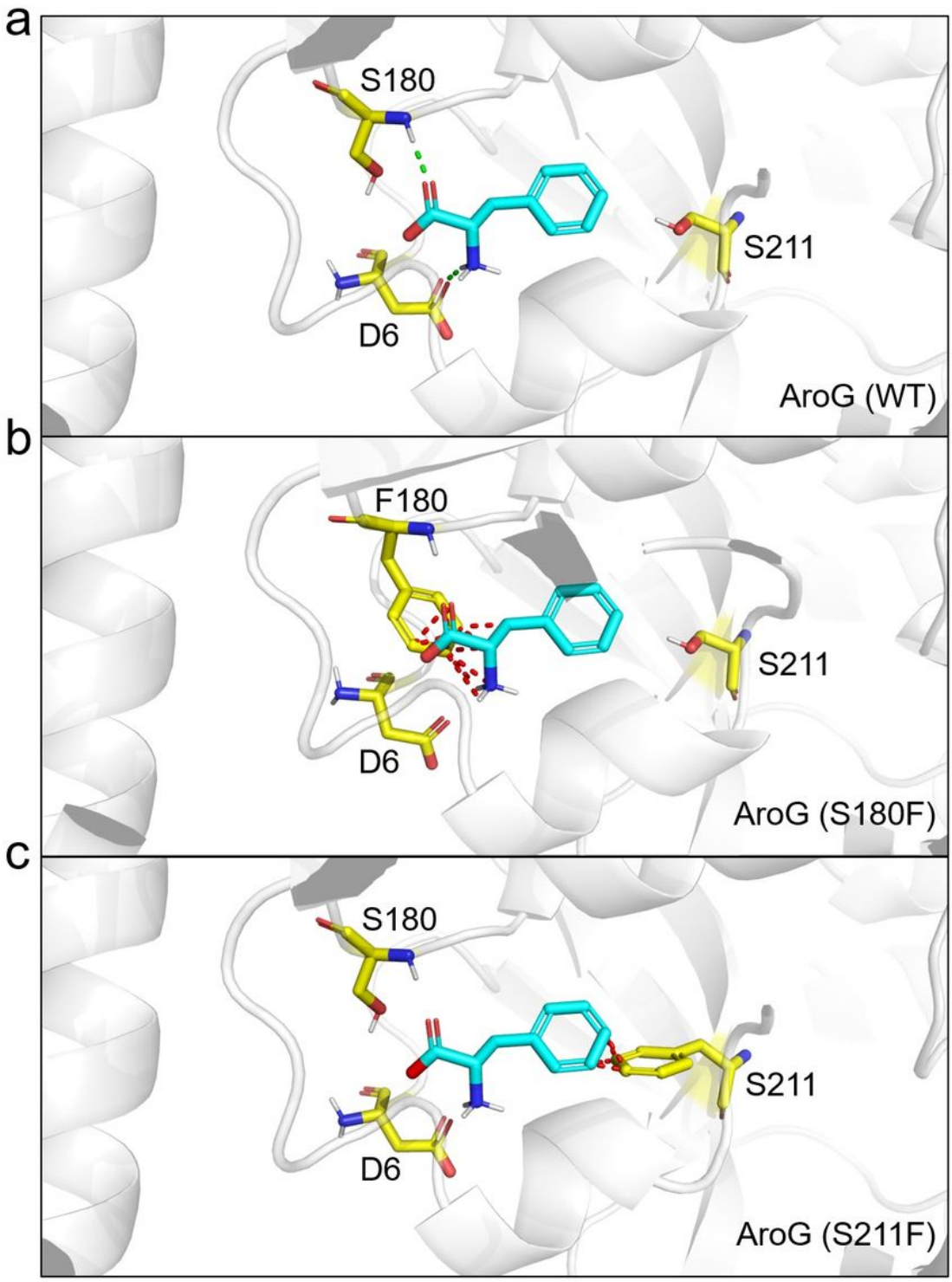

Figure 3

Binding mode of the inhibitor L-phe with AroG. (a) Binding mode of inhibitor L-phe with AroG (wild-type, WT) (PDB: 5CKS). Green dotted lines represent hydrogen-bonding interactions between L-phe and AroG (WT). After S180 (b) or S211 (c) was replaced with F, the binding of L-phe to AroG was blocked. The red dotted lines represent obstruction via steric hindrance.

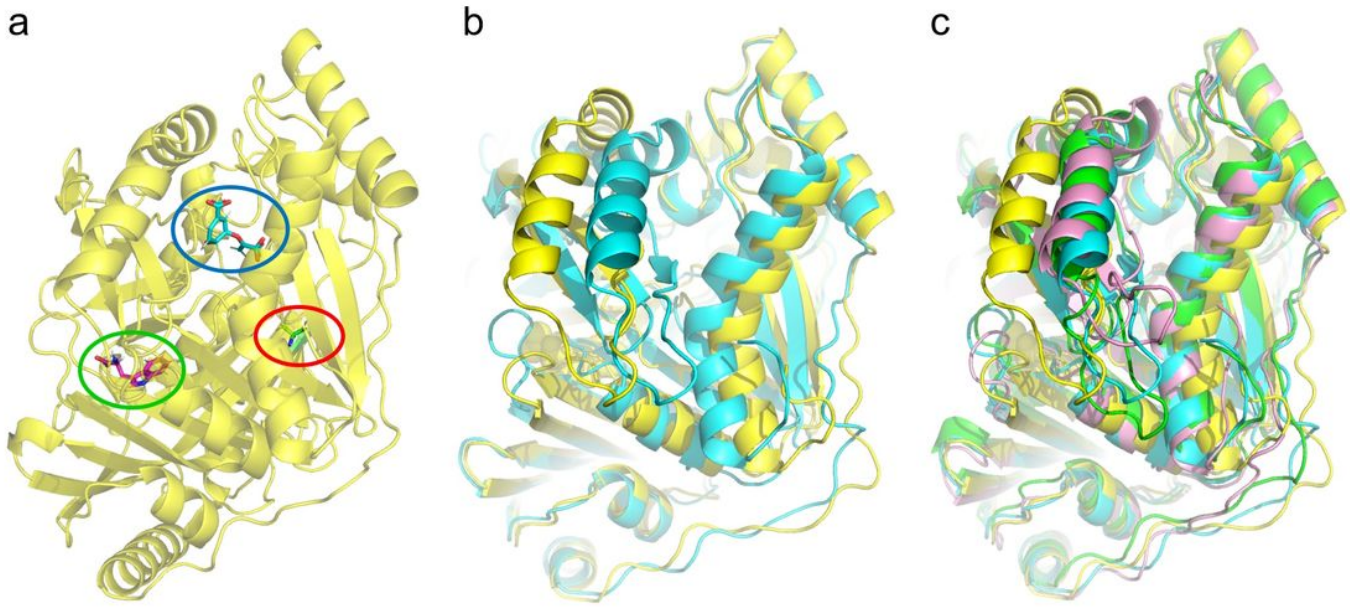

Figure 4

Modeled 3D structure of wild-type TrpE and mutant TrpE (A63V) according to MD simulations. (a) Model structure of the wild-type TrpE. The blue circle indicates the binding site of the substrate chorismate, the green circle indicates the binding site of the inhibitor L-trp, the red circle indicates the location of A63. (b) Wild-type TrpE structure in the absence (blue) and presence (yellow) of the inhibitor L-trp. (c) Structure of the mutant TrpE (A63V) in the absence (pink) and presence (green) of the inhibitor L-trp. Structure of wild-type TrpE in the absence (blue) and presence (yellow) of L-trp for comparison.

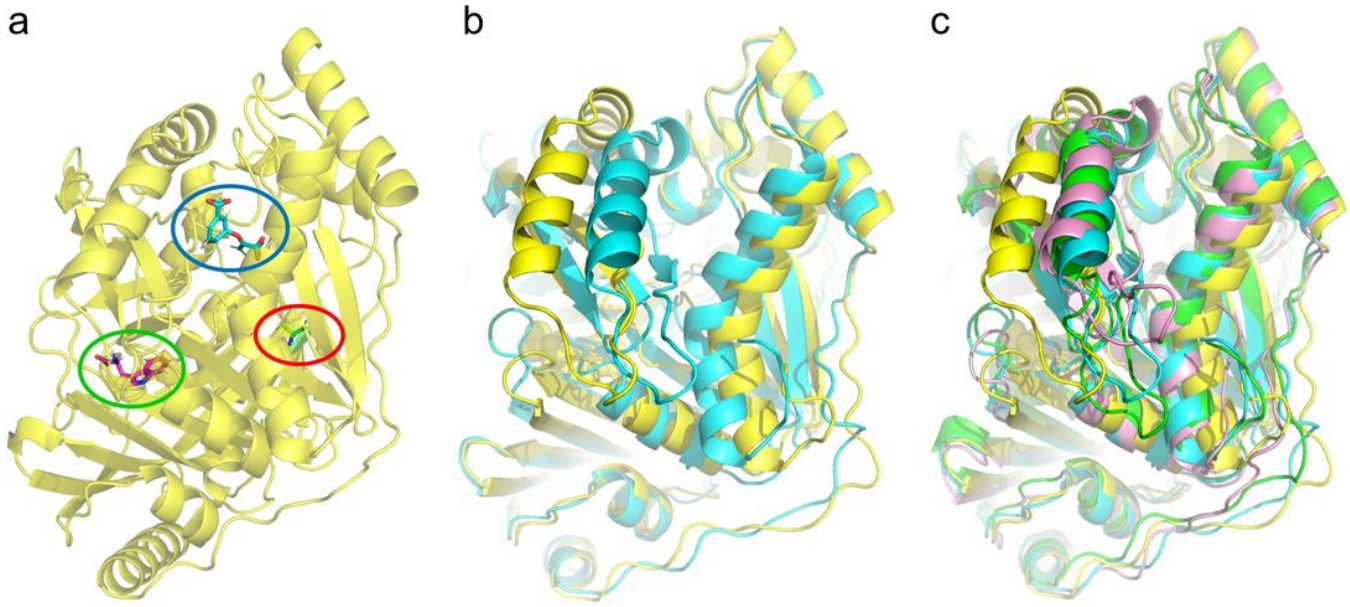

Figure 4

Modeled 3D structure of wild-type TrpE and mutant TrpE (A63V) according to MD simulations. (a) Model structure of the wild-type TrpE. The blue circle indicates the binding site of the substrate chorismate, the green circle indicates the binding site of the inhibitor L-trp, the red circle indicates the location of A63. (b) Wild-type TrpE structure in the absence (blue) and presence (yellow) of the inhibitor L-trp. (c) Structure of the mutant TrpE (A63V) in the absence (pink) and presence (green) of the inhibitor L-trp. Structure of wild-type TrpE in the absence (blue) and presence (yellow) of L-trp for comparison.

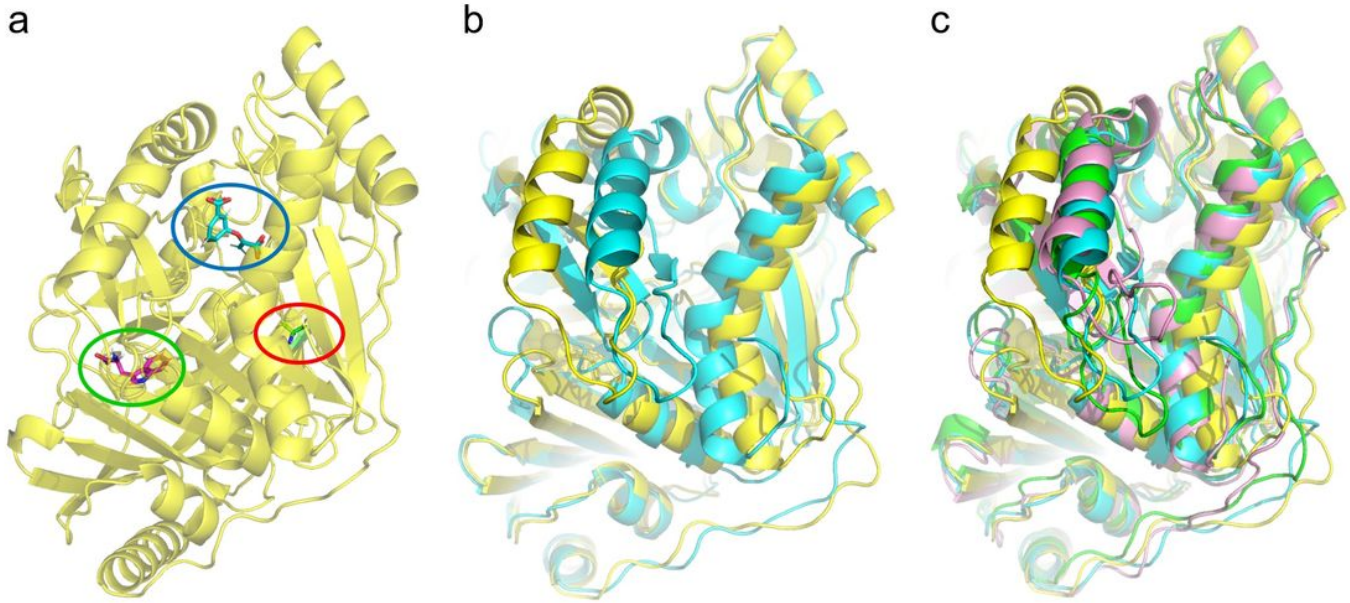

Figure 4

Modeled 3D structure of wild-type TrpE and mutant TrpE (A63V) according to MD simulations. (a) Model structure of the wild-type TrpE. The blue circle indicates the binding site of the substrate chorismate, the green circle indicates the binding site of the inhibitor L-trp, the red circle indicates the location of A63. (b) Wild-type TrpE structure in the absence (blue) and presence (yellow) of the inhibitor L-trp. (c) Structure of the mutant TrpE (A63V) in the absence (pink) and presence (green) of the inhibitor L-trp. Structure of wild-type TrpE in the absence (blue) and presence (yellow) of L-trp for comparison.

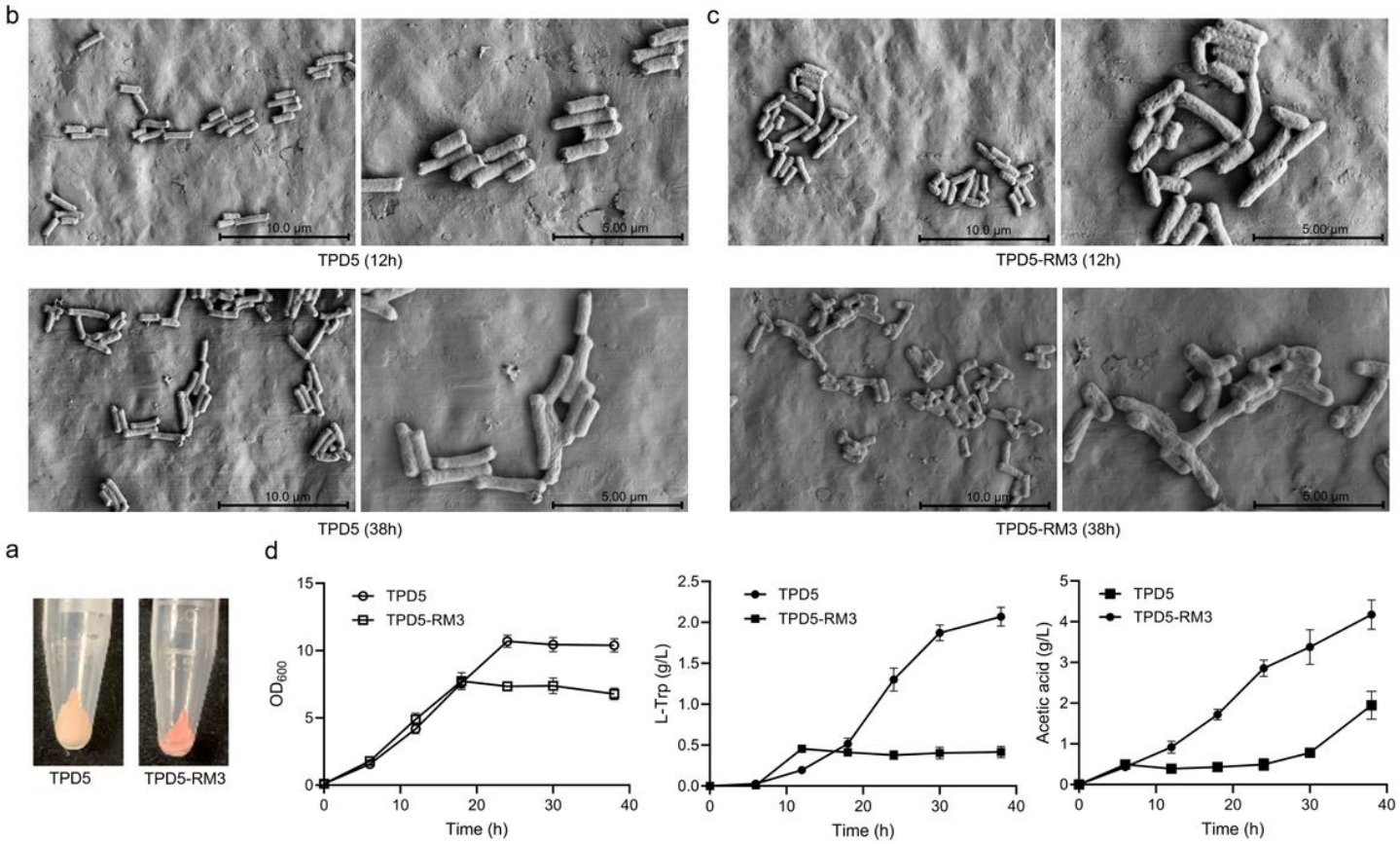

Figure 5

Effects of the global regulator RpoS in strains TPD5 and TPD5-RM3. (a) Cell pellet of strains TPD5 (left) and TPD5-RM3 (right) after fermentation for 38 h. (b) Scanning electron micrographs of strain TPD5 fermented for 12 and 38 h. (c) Scanning electron micrographs of strain TPD5-RM3 fermented for 12 and 38 h. (d) Fermentation results of strains TPD5 and TPD5-RM3. Growth curves (left), L-trp production (middle), and accumulation of acetic acid (right) of strains TPD5 and TPD5-RM3. Experiments were conducted in triplicate and the data represent the means ± SD.

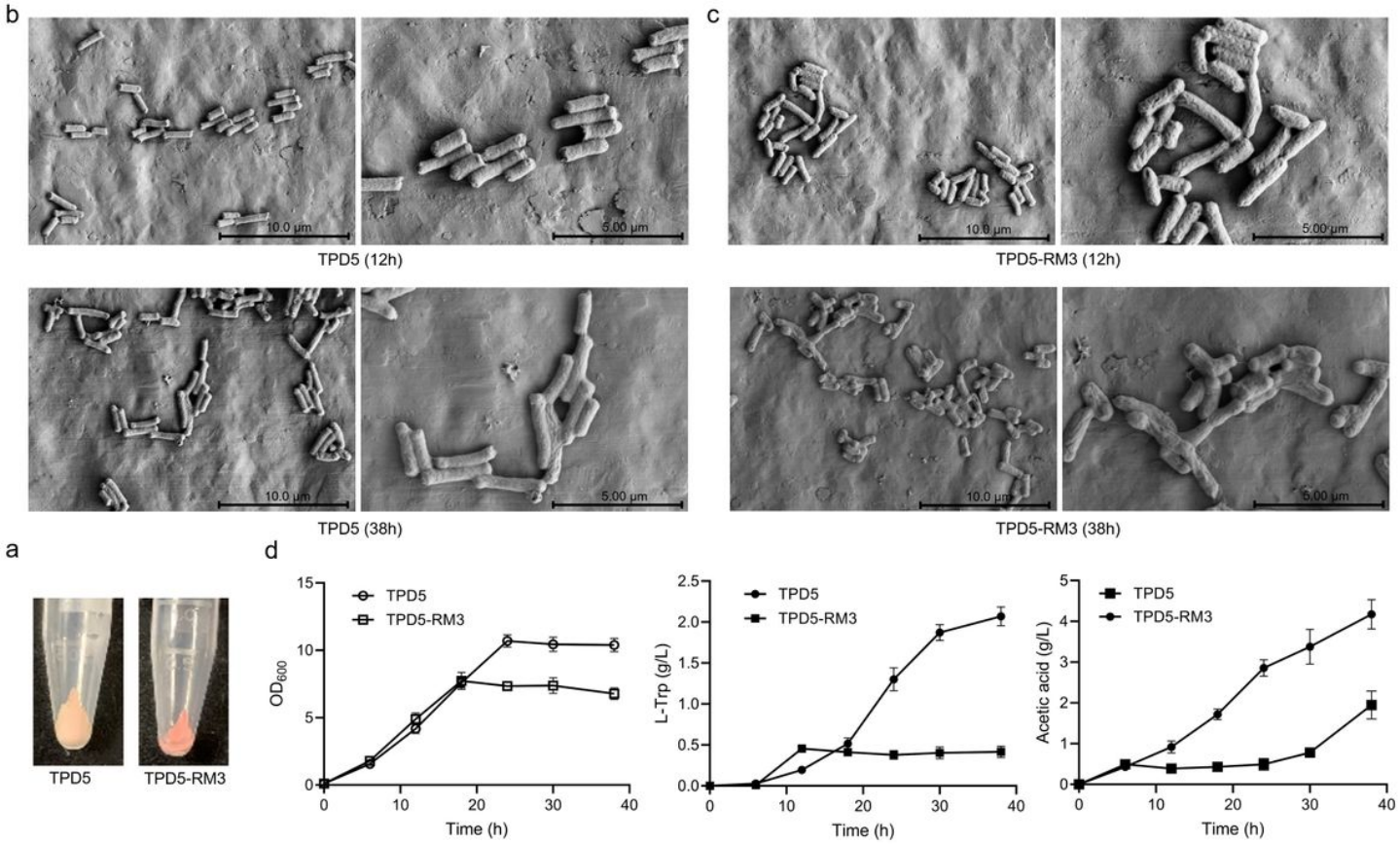

Figure 5

Effects of the global regulator RpoS in strains TPD5 and TPD5-RM3. (a) Cell pellet of strains TPD5 (left) and TPD5-RM3 (right) after fermentation for 38 h. (b) Scanning electron micrographs of strain TPD5 fermented for 12 and 38 h. (c) Scanning electron micrographs of strain TPD5-RM3 fermented for 12 and 38 h. (d) Fermentation results of strains TPD5 and TPD5-RM3. Growth curves (left), L-trp production (middle), and accumulation of acetic acid (right) of strains TPD5 and TPD5-RM3. Experiments were conducted in triplicate and the data represent the means ± SD.

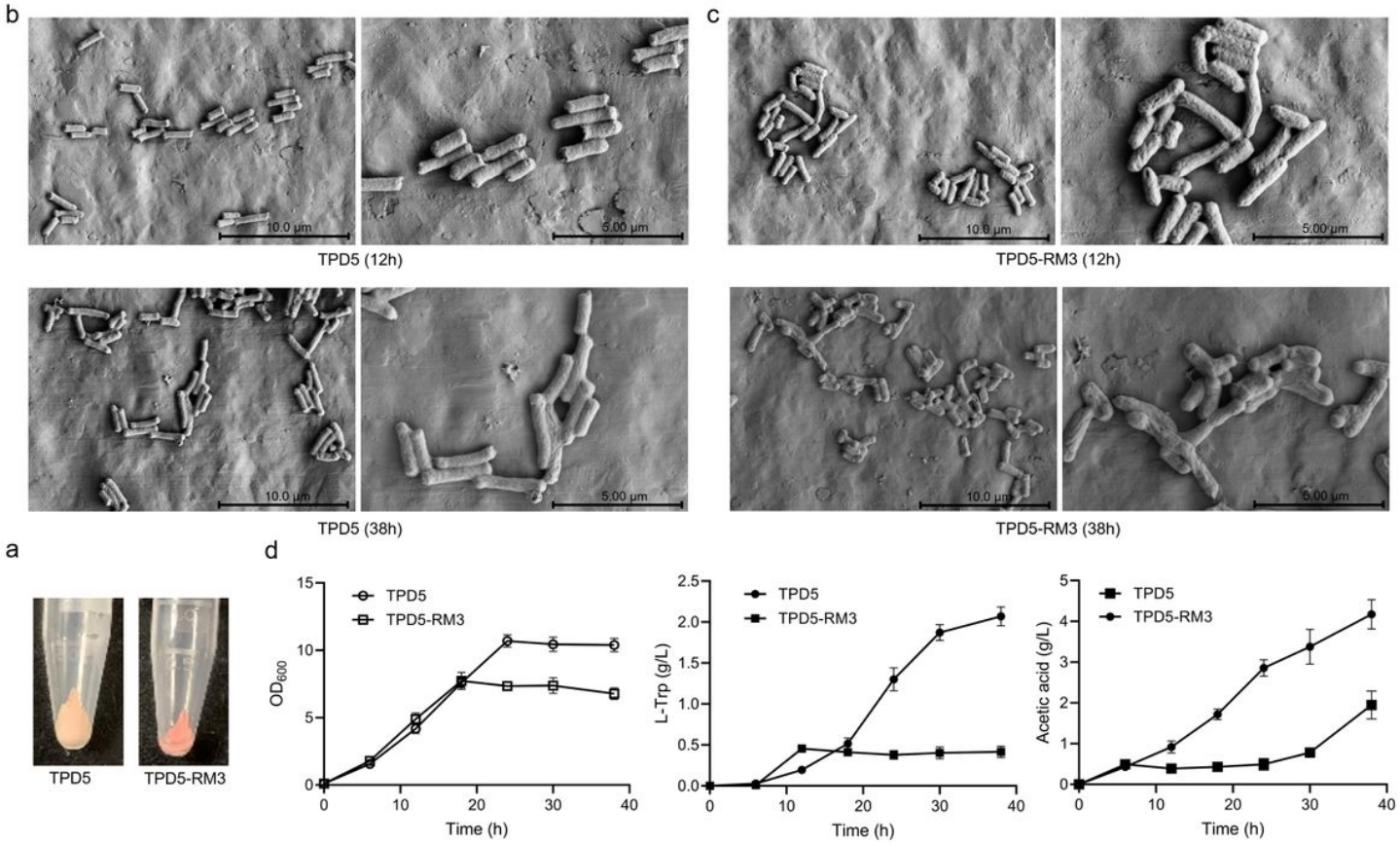

Figure 5

Effects of the global regulator RpoS in strains TPD5 and TPD5-RM3. (a) Cell pellet of strains TPD5 (left) and TPD5-RM3 (right) after fermentation for 38 h. (b) Scanning electron micrographs of strain TPD5 fermented for 12 and 38 h. (c) Scanning electron micrographs of strain TPD5-RM3 fermented for 12 and 38 h. (d) Fermentation results of strains TPD5 and TPD5-RM3. Growth curves (left), L-trp production (middle), and accumulation of acetic acid (right) of strains TPD5 and TPD5-RM3. Experiments were conducted in triplicate and the data represent the means ± SD.

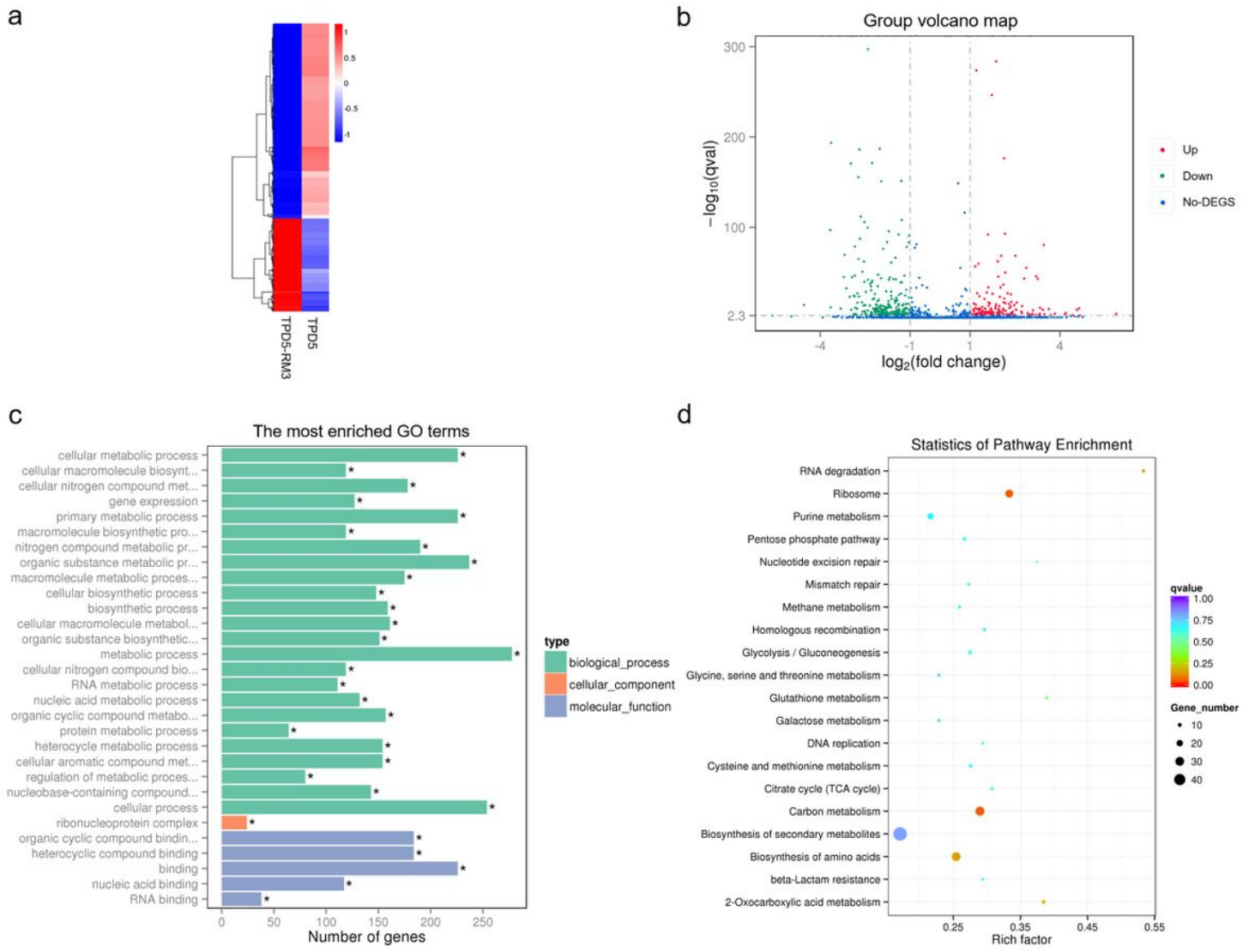

Figure 6

Transcriptomic analysis of differences between strains TPD5-RM3 and TPD5. (a) Heat maps showing the differentially expressed genes in strains TPD5-RM3 and TPD5. (b) Volcano plot showing the differentially expressed genes between TPD5-RM3 and TPD5. Red dots indicate upregulated genes and green dots indicate downregulated genes. (c) Gene Ontology (GO) functional analysis of differentially expressed genes. According to gene function, the differentially expressed genes can be divided into three categories: biological processes, cellular components and molecular function. (d) Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis of the differentially expressed genes. The size of dots represents the number of differentially expressed genes in each pathway, and the color of dots corresponds to different q-value ranges.

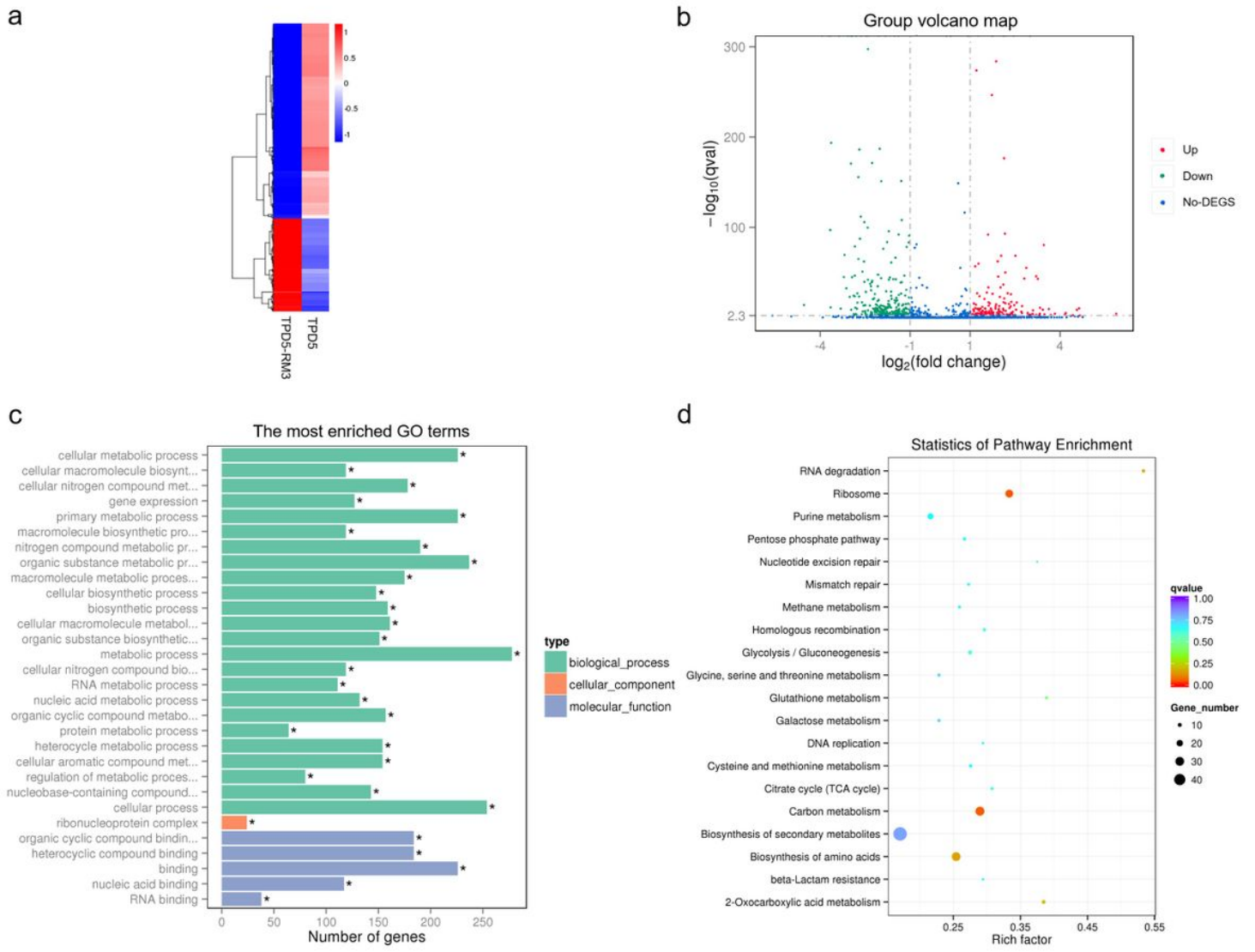

Figure 6

Transcriptomic analysis of differences between strains TPD5-RM3 and TPD5. (a) Heat maps showing the differentially expressed genes in strains TPD5-RM3 and TPD5. (b) Volcano plot showing the differentially expressed genes between TPD5-RM3 and TPD5. Red dots indicate upregulated genes and green dots indicate downregulated genes. (c) Gene Ontology (GO) functional analysis of differentially expressed genes. According to gene function, the differentially expressed genes can be divided into three categories: biological processes, cellular components and molecular function. (d) Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis of the differentially expressed genes. The size of dots represents the number of differentially expressed genes in each pathway, and the color of dots corresponds to different q-value ranges.

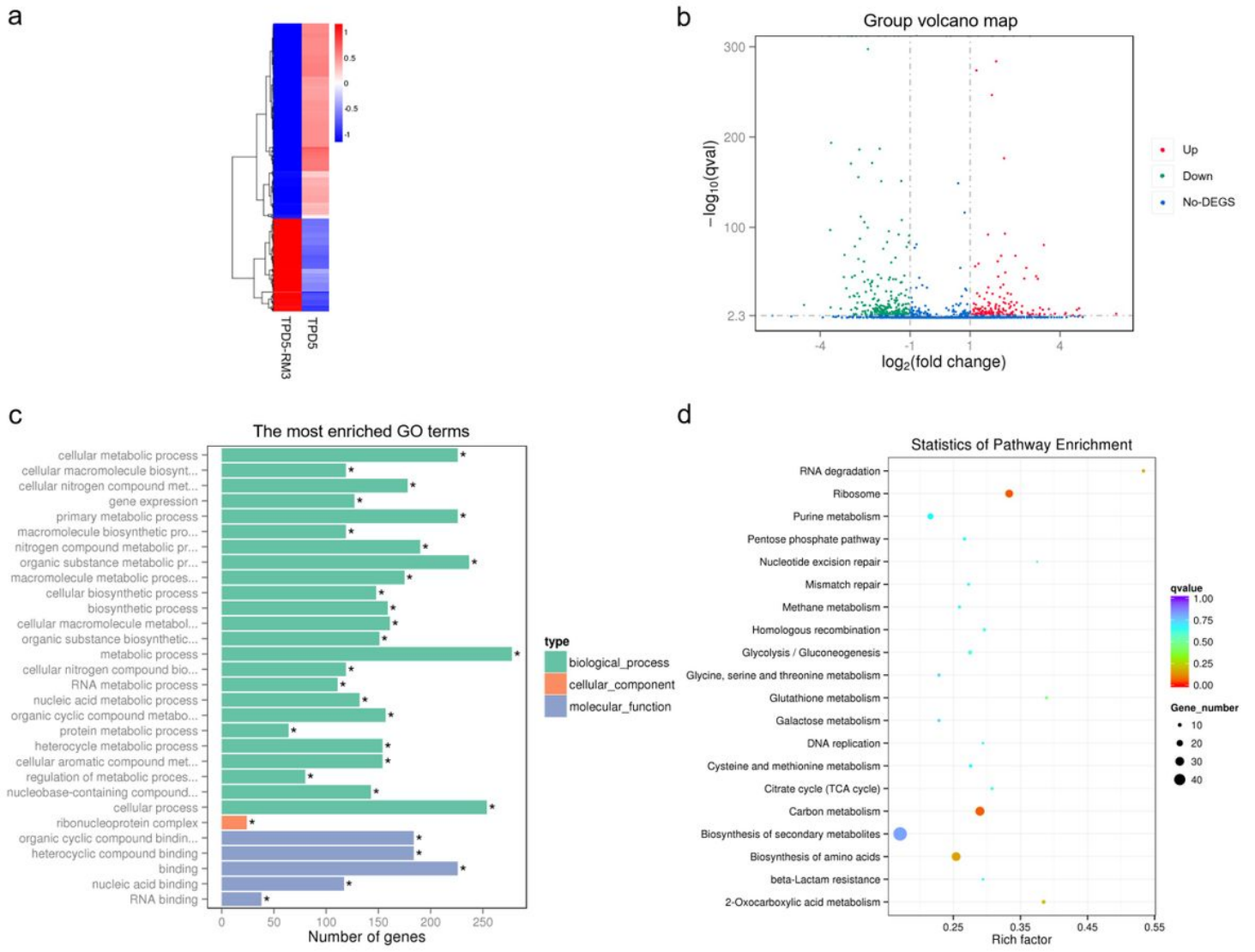

Figure 6

Transcriptomic analysis of differences between strains TPD5-RM3 and TPD5. (a) Heat maps showing the differentially expressed genes in strains TPD5-RM3 and TPD5. (b) Volcano plot showing the differentially expressed genes between TPD5-RM3 and TPD5. Red dots indicate upregulated genes and green dots indicate downregulated genes. (c) Gene Ontology (GO) functional analysis of differentially expressed genes. According to gene function, the differentially expressed genes can be divided into three categories: biological processes, cellular components and molecular function. (d) Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis of the differentially expressed genes. The size of dots represents the number of differentially expressed genes in each pathway, and the color of dots corresponds to different q-value ranges.

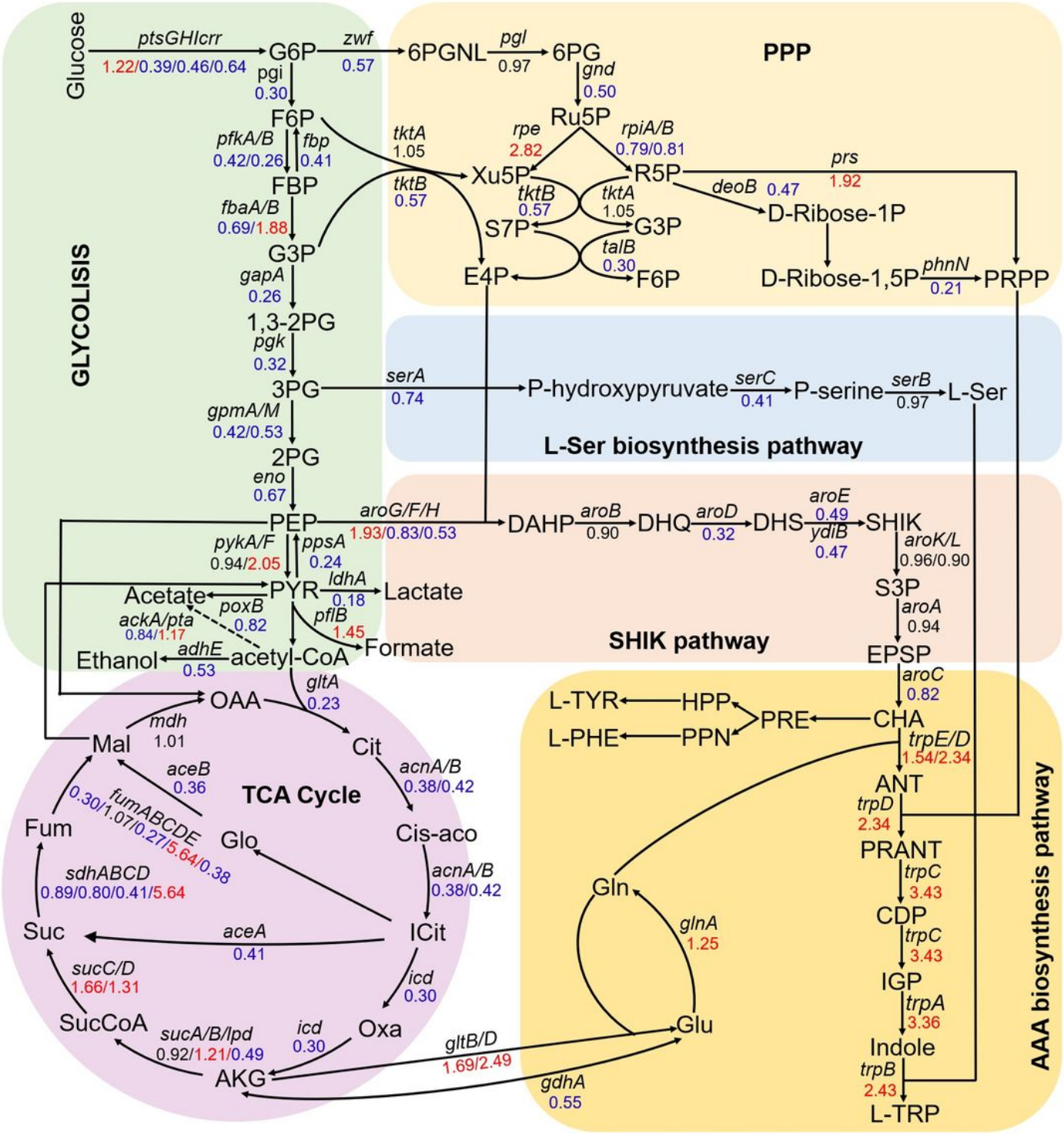

Figure 7

Schematic overview of the expression profiles of genes involved in L-trp biosynthesis. The numbers above or underneath the arrows are the ratios of the comparative expression levels in TPD5-RM3 and the control TPD5. Red indicates upregulation, blue indicates downregulation, and black indicates no notable change. G6P, Glucose-6-phosphate; F6P, Fructose-6-phosphate; FBP, Fructose-1,6-bisphosphate; G3P, Glyceraldehyde-3-phosphate; 3PG, Glycerate-3-phosphate; 2PG, Glycerate-2-phosphate; PEP,

Phosphoenolpyruvate; PYR, Pyruvate; Cit, Citric acid; Cis-aco, Cis-aconitase; ICit, Isocitrate; Oxa, Oxalosuccinate; AKG, α-ketoglutarate; SucCoA, Succinyl-CoA; Suc, Succinate; Fum, Fumarate; Mal, Malate; OAA, Oxaloacetate; 6PGNL, 6-phosphate-glucose lactone; 6PG, 6-phosphate-gluconate; Ru5P, Ribulose-5-phosphate; Xu5P, Xylulose-5-phosphate; R5P, Ribose-5-phosphate; S7P, Sedoheptulose-7-phosphate; E4P, Erythrose-4-phosphate; PRPP, 5-phospho-α-D-ribose 1-diphosphate; L-Ser, L-serine; DAHP, 3-Deoxy-D-arabinoheptulosonate 7-phosphate; DHQ, 3-Dehydroquinate; DHS, 3-Dehydroshikimate; SHIK, Shikimate; S3P, Shikimate 3-phosphate; EPSP, 5-Enolpyruvylshikimate 3-phosphate; CHA, Chorismate; PRE, Prephenate; HPP, 4-Hydroxyphenylpyruvate; L-TYR, L-tyrosine; PPN, Phenylpyruvate; L-PHE, L-phenylalanine; ANT, Anthranilate; PRANT, N-(5-phosphoribosyl)-anthranilate; CDP, 1-(o-carboxyphenylamino)-1’-deoxyribulose 5’-phosphate; IGP, (1S,2R)-1-C-(indol-3-yl)glycerol 3-phosphate; L-TRP, L-tryptophan.

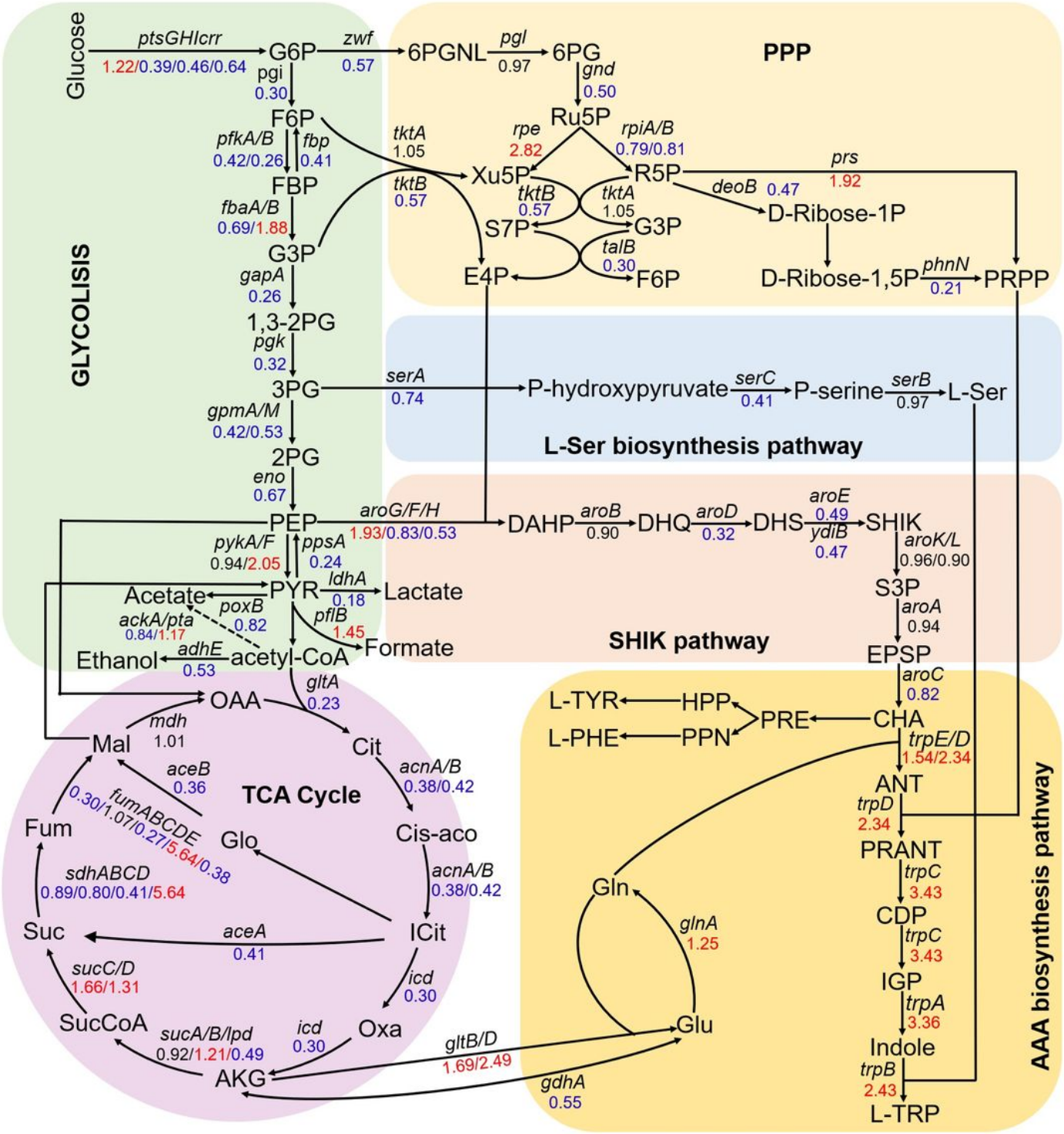

Figure 7

Schematic overview of the expression profiles of genes involved in L-trp biosynthesis. The numbers above or underneath the arrows are the ratios of the comparative expression levels in TPD5-RM3 and the control TPD5. Red indicates upregulation, blue indicates downregulation, and black indicates no notable change. G6P, Glucose-6-phosphate; F6P, Fructose-6-phosphate; FBP, Fructose-1,6-bisphosphate; G3P, Glyceraldehyde-3-phosphate; 3PG, Glycerate-3-phosphate; 2PG, Glycerate-2-phosphate; PEP,

Phosphoenolpyruvate; PYR, Pyruvate; Cit, Citric acid; Cis-aco, Cis-aconitase; ICit, Isocitrate; Oxa, Oxalosuccinate; AKG, α-ketoglutarate; SucCoA, Succinyl-CoA; Suc, Succinate; Fum, Fumarate; Mal, Malate; OAA, Oxaloacetate; 6PGNL, 6-phosphate-glucose lactone; 6PG, 6-phosphate-gluconate; Ru5P, Ribulose-5-phosphate; Xu5P, Xylulose-5-phosphate; R5P, Ribose-5-phosphate; S7P, Sedoheptulose-7-phosphate; E4P, Erythrose-4-phosphate; PRPP, 5-phospho-α-D-ribose 1-diphosphate; L-Ser, L-serine; DAHP, 3-Deoxy-D-arabinoheptulosonate 7-phosphate; DHQ, 3-Dehydroquinate; DHS, 3-Dehydroshikimate; SHIK, Shikimate; S3P, Shikimate 3-phosphate; EPSP, 5-Enolpyruvylshikimate 3-phosphate; CHA, Chorismate; PRE, Prephenate; HPP, 4-Hydroxyphenylpyruvate; L-TYR, L-tyrosine; PPN, Phenylpyruvate; L-PHE, L-phenylalanine; ANT, Anthranilate; PRANT, N-(5-phosphoribosyl)-anthranilate; CDP, 1-(o-carboxyphenylamino)-1’-deoxyribulose 5’-phosphate; IGP, (1S,2R)-1-C-(indol-3-yl)glycerol 3-phosphate; L-TRP, L-tryptophan.

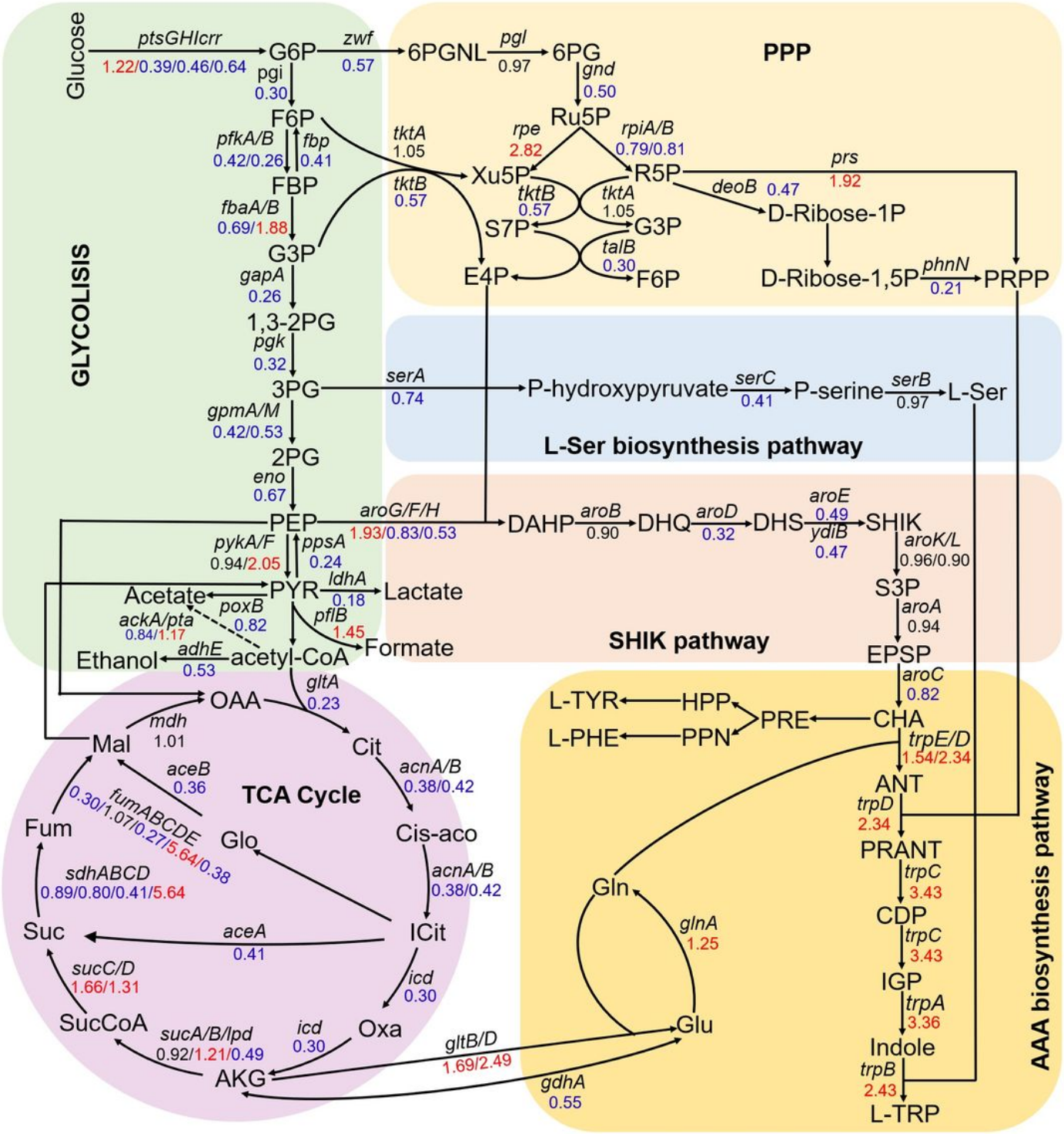

Figure 7

Schematic overview of the expression profiles of genes involved in L-trp biosynthesis. The numbers above or underneath the arrows are the ratios of the comparative expression levels in TPD5-RM3 and the control TPD5. Red indicates upregulation, blue indicates downregulation, and black indicates no notable change. G6P, Glucose-6-phosphate; F6P, Fructose-6-phosphate; FBP, Fructose-1,6-bisphosphate; G3P, Glyceraldehyde-3-phosphate; 3PG, Glycerate-3-phosphate; 2PG, Glycerate-2-phosphate; PEP,

Phosphoenolpyruvate; PYR, Pyruvate; Cit, Citric acid; Cis-aco, Cis-aconitase; ICit, Isocitrate; Oxa, Oxalosuccinate; AKG, α-ketoglutarate; SucCoA, Succinyl-CoA; Suc, Succinate; Fum, Fumarate; Mal, Malate; OAA, Oxaloacetate; 6PGNL, 6-phosphate-glucose lactone; 6PG, 6-phosphate-gluconate; Ru5P, Ribulose-5-phosphate; Xu5P, Xylulose-5-phosphate; R5P, Ribose-5-phosphate; S7P, Sedoheptulose-7-phosphate; E4P, Erythrose-4-phosphate; PRPP, 5-phospho-α-D-ribose 1-diphosphate; L-Ser, L-serine; DAHP, 3-Deoxy-D-arabinoheptulosonate 7-phosphate; DHQ, 3-Dehydroquinate; DHS, 3-Dehydroshikimate; SHIK, Shikimate; S3P, Shikimate 3-phosphate; EPSP, 5-Enolpyruvylshikimate 3-phosphate; CHA, Chorismate; PRE, Prephenate; HPP, 4-Hydroxyphenylpyruvate; L-TYR, L-tyrosine; PPN, Phenylpyruvate; L-PHE, L-phenylalanine; ANT, Anthranilate; PRANT, N-(5-phosphoribosyl)-anthranilate; CDP, 1-(o-carboxyphenylamino)-1’-deoxyribulose 5’-phosphate; IGP, (1S,2R)-1-C-(indol-3-yl)glycerol 3-phosphate; L-TRP, L-tryptophan.

a

Amino acid analyzer test results

Metabolites (mM)	TPD5-RM4	TPD5-RM4S	TPD5-RM4G
P-Ser	0.069	1.676	0.5935
PEA	0.008	0.0035	0.0105
Urea		0.0955	0.5905
Asp	0.08	0.0135	0.0745
Thr	0.013	0.014	0.0405
Ser	0.0175	0.1675	0.214
Glu	0.024	8.8665	0.2405
Gly	0.0205		0.0825
α-ABA		0.058	
Ala	1.403		
Phe		0.254	0.723
NH_3			5.937
GABA	0.1125	0.023	0.0355
MEA	0.107	0.3185	
orn	0.002		0.0845
Hypro pro	0.167	0.048	1.692

b

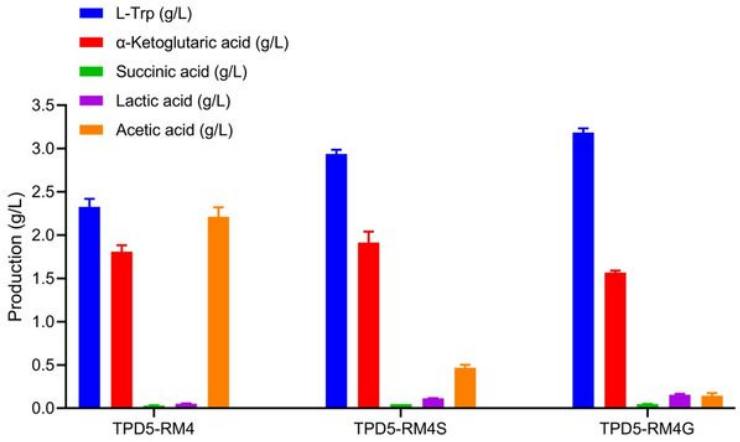

Figure 8

Fermentation results for strains TPD5-RM4, TPD5-RM4S and TPD5-RM4G. (a) The concentration of metabolites analyzed using the amino acid analyzer of the strains TPD5-RM4, TPD5-RM4S and TPD5-RM4G in shake flasks. P-Ser, Phosphoserine; PEA, Phosphoethanolamine; Asp, L-Aspartic acid; Thr, L-Threonine; Ser, L-Serine; Glu, L-Glutamic acid; α-ABA, DL-2-Aminobutyric acid; Ala, L-Alanine; Phe, L-Phenylalanine; GABA, 4-Aminobutyric acid; MEA, 2-Aminoethanol; orn, L-Ornithine; Hypro pro, L-Hydroxyproline. (b) Concentrations of L-trp production, α-ketoglutaric acid, succinic acid, lactic acid and acetic acid of the strains TPD5-RM4, TPD5-RM4S and TPD5-RM4G in shake flasks. The experiments were conducted in triplicate and the data represent the means ± SD.

a

Amino acid analyzer test results

Metabolites (mM)	TPD5-RM4	TPD5-RM4S	TPD5-RM4G
P-Ser	0.069	1.676	0.5935
PEA	0.008	0.0035	0.0105
Urea		0.0955	0.5905
Asp	0.08	0.0135	0.0745
Thr	0.013	0.014	0.0405
Ser	0.0175	0.1675	0.214
Glu	0.024	8.8665	0.2405
Gly	0.0205		0.0825
α-ABA		0.058	
Ala	1.403		
Phe		0.254	0.723
NH_3			5.937
GABA	0.1125	0.023	0.0355
MEA	0.107	0.3185	
orn	0.002		0.0845
Hypro pro	0.167	0.048	1.692

b

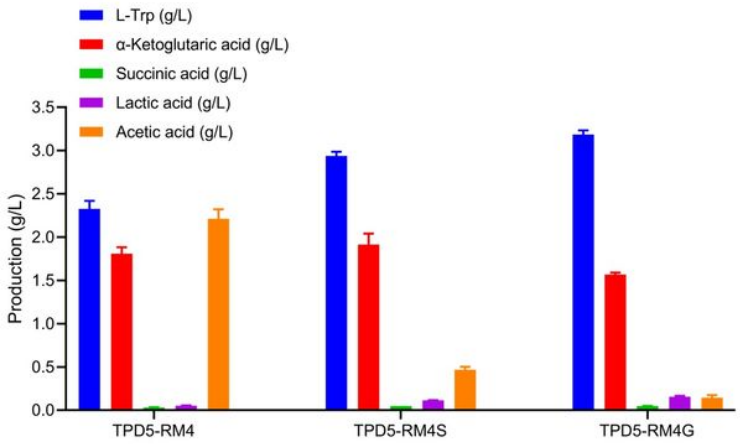

Figure 8

Fermentation results for strains TPD5-RM4, TPD5-RM4S and TPD5-RM4G. (a) The concentration of metabolites analyzed using the amino acid analyzer of the strains TPD5-RM4, TPD5-RM4S and TPD5-RM4G in shake flasks. P-Ser, Phosphoserine; PEA, Phosphoethanolamine; Asp, L-Aspartic acid; Thr, L-Threonine; Ser, L-Serine; Glu, L-Glutamic acid; α-ABA, DL-2-Aminobutyric acid; Ala, L-Alanine; Phe, L-Phenylalanine; GABA, 4-Aminobutyric acid; MEA, 2-Aminoethanol; orn, L-Ornithine; Hypro pro, L-Hydroxyproline. (b) Concentrations of L-trp production, α-ketoglutaric acid, succinic acid, lactic acid and acetic acid of the strains TPD5-RM4, TPD5-RM4S and TPD5-RM4G in shake flasks. The experiments were conducted in triplicate and the data represent the means ± SD.

a

Amino acid analyzer test results

Metabolites (mM)	TPD5-RM4	TPD5-RM4S	TPD5-RM4G
P-Ser	0.069	1.676	0.5935
PEA	0.008	0.0035	0.0105
Urea		0.0955	0.5905
Asp	0.08	0.0135	0.0745
Thr	0.013	0.014	0.0405
Ser	0.0175	0.1675	0.214
Glu	0.024	8.8665	0.2405
Gly	0.0205		0.0825
α-ABA		0.058	
Ala	1.403		
Phe		0.254	0.723
NH_3			5.937
GABA	0.1125	0.023	0.0355
MEA	0.107	0.3185	
orn	0.002		0.0845
Hypro pro	0.167	0.048	1.692

b

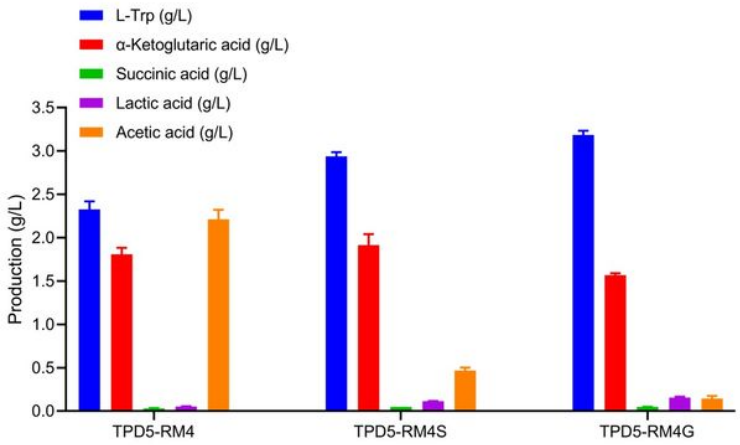

Figure 8

Fermentation results for strains TPD5-RM4, TPD5-RM4S and TPD5-RM4G. (a) The concentration of metabolites analyzed using the amino acid analyzer of the strains TPD5-RM4, TPD5-RM4S and TPD5-

RM4G in shake flasks. P-Ser, Phosphoserine; PEA, Phosphoethanolamine; Asp, L-Aspartic acid; Thr, L-Threonine; Ser, L-Serine; Glu, L-Glutamic acid; α-ABA, DL-2-Aminobutyric acid; Ala, L-Alanine; Phe, L-Phenylalanine; GABA, 4-Aminobutyric acid; MEA, 2-Aminoethanol; orn, L-Ornithine; Hypro pro, L-Hydroxyproline. (b) Concentrations of L-trp production, α-ketoglutaric acid, succinic acid, lactic acid and acetic acid of the strains TPD5-RM4, TPD5-RM4S and TPD5-RM4G in shake flasks. The experiments were conducted in triplicate and the data represent the means ± SD.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile1.docx
- Additionalfile1.docx
- Additionalfile1.docx
- Additionalfile2.xlsx
- Additionalfile2.xlsx
- Additionalfile2.xlsx
- Additionalfile3.xlsx
- Additionalfile3.xlsx
- Additionalfile3.xlsx
- Additionalfile4.xlsx
- Additionalfile4.xlsx
- Additionalfile4.xlsx