Lawrence Berkeley National Laboratory
LBL Publications

Title
Cost-efficient Bayesian emulation of quark-gluon plasma modeling with variable statistical precision

Permalink
https://escholarship.org/uc/item/9hr9q0xd

Journal
Physical Review C, 114(2)

ISSN
2469-9985

Authors
Ehlers, R
Ji, Y
Jacobs, PM
et al.

Publication Date
2026-08-01

DOI
10.1103/jqzf-d1tz

Copyright Information
This work is made available under the terms of a Creative Commons Attribution License, available at https://creativecommons.org/licenses/by/4.0/

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

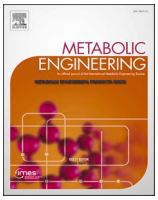
Metabolic Engineering 99 (2027) 102570



Contents lists available at ScienceDirect

Metabolic Engineering

journal homepage: www.elsevier.com/locate/meteng



Orchestrating polyketide synthase-derived aldehyde metabolism for *Streptomyces* biomanufacturing

Qingyun Dan [a,b,c,d,e,*,1], Xina Wang [f,1], Yan Chiu [a,g,1], Soonkyu Hwang [b,c,d], Justin J. Baerwald [a,b,c], Yilin Li [a,b,c,d], Lucas M. Waldburger [b,c,h,i,j], Chang Dou [c,d,k], Yiou Rong [a], Xavier B. Tao [c,f], Chenyi Li [a,b,c,d], Ning Qin [a,b,c,d], Zhen Liu [a], Seokjung Cheong [a,b,c,d], Yan Chen [b,c], Jennifer W. Gin [b,c], Matthias Schmidt [b,c,d,l], Rachel E. Sullivan [g], Tyler W.H. Backman [b,c], Christopher J. Petzold [b,c], Jay D. Keasling [a,b,c,d,h,**]

[a] *Department of Chemical and Biomolecular Engineering, University of California, Berkeley, Berkeley, CA, 94720, USA*
[b] *Joint BioEnergy Institute, Emeryville, CA, 94608, USA*
[c] *Biological Systems and Engineering Division, Lawrence Berkeley National laboratory, Berkeley, CA, 94720, USA*
[d] *QB3, University of California, Berkeley, Berkeley, CA, 94720, USA*
[e] *Department of Chemical and Biomolecular Engineering, Johns Hopkins University, Baltimore, MD, 21218, USA*
[f] *Department of Chemistry, University of California, Berkeley, CA, 94720, USA*
[g] *Department of Molecular and Cell Biology, University of California, Berkeley, Berkeley, CA, 94720, USA*
[h] *Department of Bioengineering, University of California, Berkeley, Berkeley, CA, 94720, USA*
[i] *Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, Berkeley, CA, 94720, USA*
[j] *Department of Electrical Engineering and Computer Science, University of California, Berkeley, Berkeley, CA, 94720, USA*
[k] *Advanced BioProcess Development Unit, Lawrence Berkeley National Laboratory, Emeryville, CA, 94608, USA*
[l] *Institute of Applied Microbiology, Aachen Biology and Biotechnology, RWTH Aachen University, Aachen, 52062, Germany*

A B S T R A C T

Advances in metabolic engineering have enabled the integration of diverse non-native reactions into microbial hosts, providing sustainable alternatives to petrochemical manufacturing. We previously developed a modular polyketide synthase (PKS)-based platform in *Streptomyces albus* for producing 17 diols, amino alcohols, and hydroxy acids. In this platform, terminal thioester reductases (TRs) release aldehyde intermediates that are subsequently converted by post-PKS enzymes to control product chemistry. Here we found that PKS-TR strains accumulated hundreds of genome-wide single-nucleotide variants, including missense mutations on conserved PKS and TR residues leading to impaired enzyme activity. Based on the hypothesis of aldehyde toxicity, we coupled TR catalysis with downstream alcohol dehydrogenases (ADHs) through gene co-integration or PKS-TR-ADH fusion to limit aldehyde accumulation. These strategies reduced genomic variability and increased diol production, reaching a total diol titer of 14.2 g/L in bioreactors, including 10.8 g/L 1,3-butanediol. Extending this approach, engineered strains incorporating PKS-TR co-integration with transaminases or the *padA* aldehyde dehydrogenase achieved 79% higher total amino alcohol titers and 145% higher 2-methyl-3-hydroxyhexanoic acid titer than previous reports, while eliminating diol byproducts from the optimized hydroxy acid pathway. These results establish TR-mediated aldehyde release as a programmable branch point and orchestrated downstream catalysis as a general strategy for improving the stability, productivity, and selectivity of PKS-based *Streptomyces* biomanufacturing.

1. Introduction

As one of the most ancient technologies, microbial fermentation has been used by humans to produce food and beverages for over 9000 years (McGovern et al., 2004). Until the 1930s, many bulk chemicals were also produced from biomass fermentation. Following the development of cracking and reforming processes to produce olefin and aromatic building blocks, the petroleum industry came to dominate the

* Corresponding author. Department of Chemical and Biomolecular Engineering, Johns Hopkins University, Baltimore, MD, 21218, USA.
** Corresponding author. Department of Chemical and Biomolecular Engineering, University of California, Berkeley, Berkeley, CA, 94720, USA.
E-mail addresses: qdan@jhu.edu (Q. Dan), keasling@berkeley.edu (J.D. Keasling).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.ymben.2026.102570
Received 28 August 2026; Received in revised form 12 October 2026; Accepted 17 October 2026
Available online 17 October 2026
1096-7176/© 2026 The Authors. Published by Elsevier Inc. on behalf of International Metabolic Engineering Society. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

production of bulk and specialty chemicals (Alper, 1999). As a result, modern petrochemical manufacturing produces tens of thousands of molecules at the scale of nearly 1 billion tonnes per year, emitting approximately 1.30 Gt of CO_2-equivalent in 2020 (Cullen et al., 2024; Schüth and Schunk, 2026). Thus, to reduce fossil CO_2 emissions and save lives, our world requires a shift from petroleum-dependent industry toward a more sustainable bioeconomy (Liu et al., 2021).

Biomanufacturing holds great promise for replacing selected petrochemical production routes, as demonstrated by the commercial-scale microbial production of a series of diols, including natural metabolite 1,3-propanediol and nonnatural 1,4-butanediol (Biebl et al., 1999; Nakamura and Whited, 2003; Yim et al., 2011; Ewing et al., 2022). Despite this progress, development of biosynthetic platforms for many nonnatural chemicals remains challenging and requires more

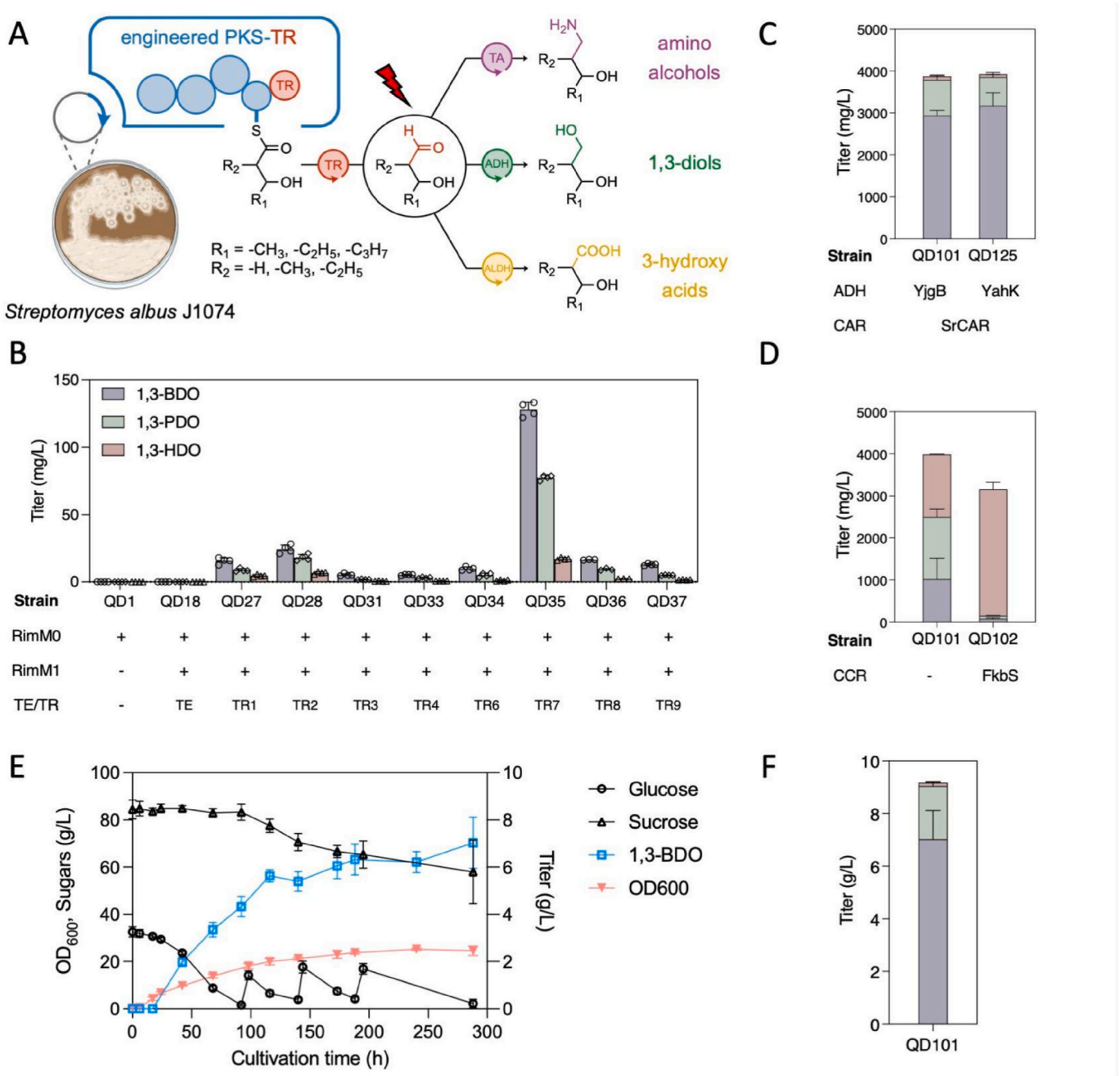


Fig. 1. Construction and optimization of the rimocidin PKS-TR platform for diol production in *Streptomyces albus* J1074. (A) Schematic representation of the RimPKS-TR platform. TR-mediated product release generates 3-hydroxyaldehydes that can be converted by transaminases (TAs), alcohol dehydrogenases (ADHs), or aldehyde dehydrogenases (ALDHs) into amino alcohols, 1,3-diols, or 3-hydroxy acids, respectively. The red lightning symbol represents aldehyde-associated cellular stress. (B) Effects of terminal TR selection on diol production. Strains were cultivated in R5 medium in 24-well plates for 3 days. The TE domain is from the 6-deoxyerythronolide B synthase (DEBS); TR information is listed in (Dan et al., 2025). $n = 3$ for QD36, $n = 4$ for the rest groups. (C) Diol production by strains expressing YjgB or YahK ADH and SrCAR. Strains were cultivated in shake flasks containing R5 + 2% glucose for 10 days. Compared to QD101, QD125 presented a slight increase in 1, 3-BDO production but lower titers of 1,3-PDO and 1,3-HDO. $n = 3$. (D) Effect of the FkbS CCR on diol product profile. 12-day cultivation results in R5 + 2% glucose + 100 mM L-valine were presented here. $n = 2$ for QD101, $n = 4$ for QD102. (E) Profiles of sugar concentrations, cell growth and 1,3-BDO production during fed-batch cultivation of QD101 with glucose feeding. $n = 4$. (F) Final diol composition of QD101 after 12 d of fed-batch cultivation. $n = 4$. Bars represent means, error bars indicate standard deviation.

sophisticated enzyme tools compatible with programmable pathway design. Modular polyketide synthases (PKSs) are assembly-line megasynthases widely present in both prokaryotes and eukaryotes, producing secondary metabolites with substantial structure complexity from simple coenzyme A (CoA) building blocks (Cortes et al., 1990; Donadio et al., 1991). Because of their assembly-line architecture, PKS engineering is suitable for combinatorial biosynthesis and *de novo* production for designer molecules (Menzella et al., 2005). Over the years, we have demonstrated the versatility of combinatorial PKSs as biosynthetic platforms through the bioproduction of nonnatural methyl/ethyl ketones (Yuzawa et al., 2018), hydroxy acids (Schmidt et al., 2025), lactones (Wang et al., 2026), and lactams (Lee et al., 2025). Successful implementation of these diverse PKS pathways in *Streptomyces albus*, *Escherichia coli* and *Pseudomonas putida* further highlights their adaptability across different microbial hosts.

Expanding product chemistry remains a major challenge in PKS retrobiosynthesis. One key limitation is that PKS product type is largely determined by the terminal enzymatic domain, often a thioesterase (TE) catalyzing thioester hydrolysis, macrocyclization or Claisen-like condensation (Horsman et al., 2016). To expand PKS-accessible chemical design space toward nonnatural diols, we included aldehyde-forming terminal thioester reductases (TRs) into combinatorial PKS design, and established a rimocidin (Rim) PKS-TR biosynthetic platform in *Streptomyces albus* J1074 (Dan et al., 2025) (Fig. 1a). Through genome mining and phylogenetic analysis, we assembled a library of PKS-associated terminal TR domains (Mullowney et al., 2018) and fused selected TRs to the first two modules of Rim PKS using serine recombinase-assisted genome engineering (SAGE), thereby programming reductive PKS product release (Merrick et al., 2018). Specifically, we sequentially introduced three biosynthetic components into *S. albus* J1074: the Rim loading module (M0) promiscuously accepting acetyl-, propionyl- and butyryl-starter units was first integrated into phiC31 bacterial attachment (*attB*) site, followed by VWB *attB* site integration of a chimeric malonyl-CoA-specific Rim extension module fused to selected TR (M1-TR) (Supplementary Fig. 1a). This RimM0M1-TR platform produced 1,3-butanediol (1,3-BDO), 1,3-pentanediol (1,3-PDO) and 1, 3-hexanediol (1,3-HDO), likely through endogenous alcohol dehydrogenase- (ADH-) catalyzed reduction of the released 3-hydroxyaldehyde PKS products (Supplementary Fig. 1b). Using a canonical PKS acyltransferase (AT) exchange strategy in RimM1-TR, we further incorporated methyl or ethyl branches at the C2 position to produce α-branched diols. Alternatively, *fkbS*, which encodes a crotonyl-CoA carboxylase/reductase (CCR) that converts crotonyl-CoA to (2*S*)-ethylmalonyl-CoA and butyryl-CoA (Erb et al., 2007; Yu et al., 2019), was introduced to pSAM2 *attB* site to increase 1,3-HDO product ratio, demonstrating product fine-tuning through precursor pool regulation. Lastly, to improve titers and redirect product chemistry, standalone post-PKS ADHs (*yahK* and *yjgB*) or transaminases (TAs) were integrated at phiBT1 *attB* site, enabling the production of 1008 mg/L diols and 554 mg/L amino alcohols in shake flasks.

Our PKS-TR platform opens new technical possibilities for programmable access to precisely controlled product chemistry. An equally important challenge is microbial host engineering to improve the titer, yield and rate (TRY) of target bioproducts. In fatty acid megasynthase biomanufacturing in yeasts, the effects of substrates, key intermediates, cofactors and products on host genetic and metabolomic stability have been heavily investigated. For instance, Gao et al. reported methanol toxicity during free fatty acid production by the methylotropic yeast *Ogataea polymorpha* disrupted phospholipids homeostasis, which could be rescued by restoring phospholipid metabolism (Gao et al., 2022). In contrast, how newly introduced PKS pathways affect host genetics and metabolism remains poorly understood, particularly in non-model hosts in metabolic engineering, such as *S. albus*.

To address these questions, we expanded our previous TR screening efforts in chimeric RimPKS-TR pathways in *S. albus* J1074 and discovered an unexpected inverse relationship between genetic integration efficiency and diol titer. Whole-genome sequencing revealed hundreds of genome-wide mutations across diol-producing strains with consistent patterns, mostly single-nucleotide polymorphisms (SNPs). AlphaFold3-based structure analysis (Abramson et al., 2024), protein and genomic language model-based residue naturalness scoring (Cornman et al., 2024; Lin et al., 2023) and biochemical characterization of PKSs or TRs suggested a linkage between the mutations and impaired enzyme function. Hypothesizing that mutation accumulation resulted from toxicity of TR-mediated aldehyde release, we designed two approaches to orchestrate downstream aldehyde metabolism. Co-integrating *rimPKS-TR* with the *yjgB* ADH or constructing a PKS-TR-YjgB fusion reduced mutation accumulation, and the co-integration strain increased the total diol titer from 9.1 g/L to 14.2 g/L in bioreactors. Furthermore, we extended this strategy by co-integrating TAs or aldehyde dehydrogenases (ALDHs), contributing to 79% increase in amino alcohol titers and 145% increase in hydroxy acid titers with substantially reduced mutation accumulation. Notably, the PadA ALDH-containing strain had a high product selectivity for hydroxy acid production by eliminating diol byproducts. Taken together, these findings present a thorough investigation of PKS-TR engineering effects on *S. albus* genetic stability and productivity, and provide a general strategy for orchestrating TR utilization in PKS retrobiosynthesis.

2. Results

2.1. Identification of a putative TR for improved diol production in engineered Streptomyces albus

We previously performed initial screening of the activity of eight selected TR sets in diol production, revealing substantial variability in diol titers depending on the TR used (Dan et al., 2025). Additionally, the introduction of TR7, a putative TR originating from *Streptomyces* sp. TRM70350 (Supplementary Fig. 2), resulted in undetectable diol production in the QD35 (RimM0M1-TR7) strains. Notably, out of the eight tested TRs, the success rate in integrating the *rimM1-TR7* gene was significantly less compared to any other TR. Further investigation by whole-genome sequencing of a representative strain, QD35m0, revealed the presence of RimM1-TR7 mutations leading to early translation termination (Supplementary Fig. 3), resulting in incapable diol production.

To better understand TR7 peculiarity, we developed a *S. albus* J1074 cultivation platform for rapid 3-day diol production detection in 24-well plates and subsequently performed extensive screening of QD35 transformants (Fig. 1b). This experiment led to the discovery of three representative colonies with a record diol titer, compared to previously best-performing QD27 (RimM0M1-TR1, coelimycin PKS TR from *Streptomyces coelicolor* M145) and QD28 (RimM0M1-TR2, a *Streptomyces* sp. NRRL-B24891 TR) strains (Awodi et al., 2017). We named these TR7-carrying strains QD35m1, QD35m2 and QD35m3 (Fig. 1b; Supplementary Fig. 4). Whole-genome sequencing revealed TR7 mutations in QD35m2 and QD35m3, as well as Rim PKS mutations in all three strains (Supplementary Table 1). Nonetheless we decided to test QD35m1 engineering for improved diol titers, because of its wild-type TR7 and optimal abundance of RimM0 and RimM1-TR7 proteins (Supplementary Fig. 5). We introduced *yjgB* ADH in QD35m1 to improve the conversion of aldehyde intermediates into diols. A carboxylic acid reductase (CAR) gene (Gahloth et al., 2017), *SrCAR*, was co-integrated to facilitate recycling of hydroxy acids to aldehydes, and eventually diol production. The resulting strain QD101 produced 2939.0 mg/L 1, 3-BDO, 849.1 mg/L 1,3-PDO, and 87.4 mg/L 1,3-HDO in R5 + 2% glucose medium in shake flasks after 10-day cultivation (Fig. 1c). In parallel, QD125 was constructed from QD35m1 with an alternative ADH gene, *yahK*. Compared to QD101, QD125 showed similar overall diol productivity (Fig. 1c).

To test whether the RimPKS-TR7-based platform retained PKS product tunability, QD102 was constructed by introducing the *fkbS* CCR

gene into QD101, with expectation to increase butyryl-CoA flux and shift the product pool towards 1,3-HDO. As expected, QD102 exhibited a dramatic shift in diol composition: 1,3-BDO and 1,3-PDO titers were reduced to 78.4 mg/L and 70.2 mg/L, respectively, and 1,3-HDO production increased to 3005.5 mg/L after 12 days of cultivation in R5 + 2% glucose medium supplemented with 100 mM L-valine, a direct butyryl-CoA precursor (Gläser et al., 2021). In comparison, QD101 produced roughly equal ratios of different diols: 1019.7 mg/L 1,3-BDO, 1473.1 mg/L 1,3-PDO, and 1490.4 mg/L 1,3-HDO (Fig. 1d). Altering the CoA substrate concentration changed the diol product distribution without significantly compromising overall product titers, confirming that the optimized RimPKS-TR7 platform retains product tunability.

Next, QD101 and QD125 were selected for scale-up because of the industrial relevance of 1,3-BDO (Liu et al., 2024). Fed-batch cultivation of QD101 in a 1-L bioreactor, using R5 as the starting medium and glucose as supplemental carbon source, resulted in 7.0 g/L 1,3-BDO production over 12-day cultivation to a final OD_{600} of 24.5 (Fig. 1e). Together with 2.0 g/L 1,3-PDO and 129.3 mg/L 1,3-HDO production, the total diol titer reached 9.1 g/L (Fig. 1f). Unexpectedly, we observed slow but reproducible consumption of sucrose as a second carbon source in bioreactors. R5 medium contains 103 g/L sucrose which serves primarily as an osmotic stabilizer for *Streptomyces* growth (Elibol and Mavituna, 1998), and its consumption by *S. albus* J1074 was not observed in our shake-flask cultivation experiments (Supplementary Fig. 6a–b). In comparison to QD101 fed-batch results, similar production trends and titers were observed with QD125 strain when glucose (Supplementary Fig. 6c) or sucrose (Supplementary Fig. 6d) was used as the supplemental carbon source.

2.2. Effect of TR integration on engineered Streptomyces albus genomes

Next we aimed to systematically analyze genome-wide mutations in the engineered *S. albus* J1074 strains. We found that the predominantly used phiC31 site-specific integration system integrates duplicate *rimM0* genes in both the native *attB* and another pseudo-*attB* site at 100% efficiency across all tested *S. albus* strains (Supplementary Fig. 7). In comparison, *rimM1-TR* genes were integrated into the native VWB *attB* site as a single copy (Supplementary Fig. 7).

Notably, while all strains accumulated single nucleotide mutations, insertions and deletions to some extent, whole genome sequencing revealed a consistently marked increase in SNPs following *rimM1-TR* integration compared to controls (Fig. 2a; Supplementary Fig. 8), suggesting a TR-specific effect on genome stability. In contrast, the frequency of genome-wide insertions and deletions remained comparable across strains regardless of TR integration or the choice of TR, indicating that TR-associated mutagenesis primarily results in nucleotide substitutions (Supplementary Fig. 9). Furthermore, we observed substantially higher SNP mutation frequencies for TR7-integrated QD35 strains compared to TR1-carrying QD27 and TR2-carrying QD28 strains (Fig. 2a–b). Consistent with these results, TR-integrated strains accumulate predominantly missense mutations, with QD28 exhibiting 308 missense mutations out of 401 total SNPs and QD35m1 exhibiting 694 missense mutations out of 896 total SNPs (Supplementary Fig. 8). Together with the fact of lower genetic integration efficiency and higher diol production titers of QD35 mutants, these results suggest a positive correlation between SNP frequency and TR activity. Together with no confirmed detection of hydroxy aldehydes, the immediate PKS-TR products, during *S. albus* cultivation, we hypothesized that these aldehydes are highly reactive, may be toxic to *S. albus* cells and impose strong selective pressure on the host, ultimately leading to higher mutation rates.

SNP analysis revealed a predominance of two amino acid substitutions: aspartate to asparagine (D to N) resulting from GAC to AAC codon change, and serine to leucine (S to L) from TCG to TTG codon change (Fig. 2c; Supplementary Tables 1–2). Strand-specific mutation rate analysis revealed that G to A and C to T transitions were the most frequent base substitutions (Supplementary Fig. 10). Because these transitions directly generate the observed D to N and S to L substitutions, their prevalence could initially be attributed to a general mutation bias at G:C base pairs in the GC-rich *S. albus* genome. However, this bias persisted even after accounting for strain-specific base substitution frequencies and native codon usage in *S. albus*, indicating that the observed pattern cannot be fully explained by background mutational processes (Supplementary Table 2). Moreover, the affected GAC and TCG codons are frequently used in *S. albus* J1074, excluding rare-codon usage as a likely explanation (Supplementary Table 3). Furthermore, these specific amino acid substitutions are consistently enriched in all measured TR-containing strains, reinforcing the idea of an unexpectedly conserved mutagenesis pattern linked to TR integration (Supplementary Table 4). These mutations were also detected within the RimPKS-TR genes integrated at SAGE target sites (Supplementary Table 1). The basis for both the amino acid preference and the involvement of the integrated constructs remains unknown.

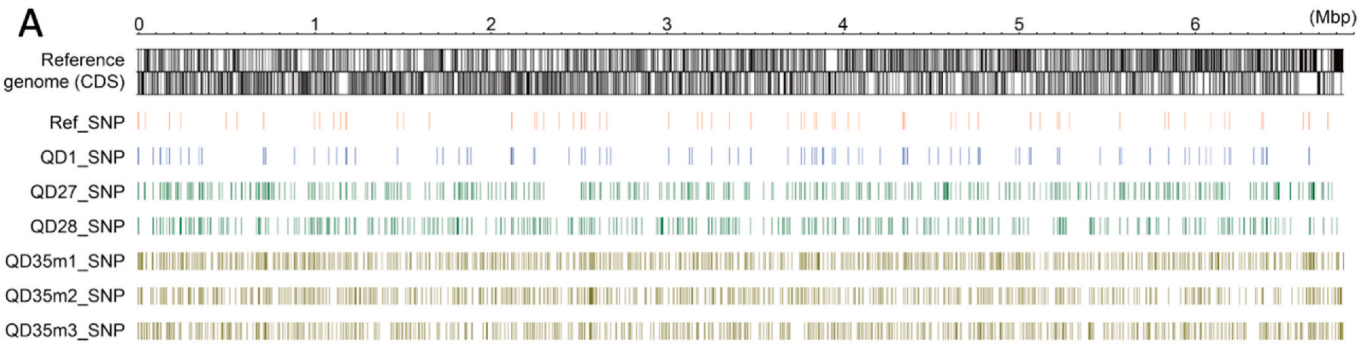
To assess whether these RimPKS-TR substitutions affect evolutionarily constrained positions, we performed conservation analysis on mutated PKS residues in QD35m1 strain (RimM0 D1044N, D1519N, D1531N, and RimM1 D886N, D1140N, D1495N) using two independent language models, ESM2 (Lin et al., 2023) and gLM2 (Cornman et al., 2024). Among five of the six mutated positions in QD35m1, both models assigned high naturalness probability to the wild-type aspartate (Fig. 2d; Supplementary Fig. 11). We expanded the probability analysis to aspartate variations in QD35m2 and QD35m3, and observed consistently high naturalness probabilities (Supplementary Fig. 11). These results indicate strong sequence constraint at these sites across naturally evolved sequences, and suggest chances of mutagenesis occurring at functionally relevant positions. Next, we mapped these mutations onto AlphaFold3-predicted structures of RimM0 and RimM1 respectively (Abramson et al., 2024). In the RimM0 structure, all D1044, D1519 and D1531 residues localize near the hypothetical ketosynthase (KS)-AT–acyl carrier protein (ACP) interfaces (Fig. 2e) (Cogan et al., 2021). In RimM1, D886 also sits in the KS-AT linker domain and D1495 maps to the ACP domain (Fig. 2e). The recurrence of these D to N substitutions across TR-containing strains, together with language model predictions and their structural locations, supports the hypothesis that they may alter PKS function.

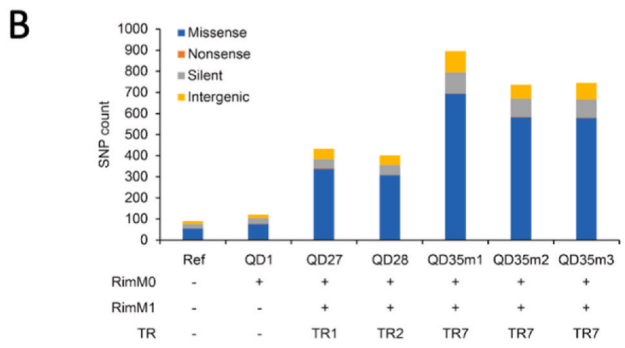
Finally, we evaluated the effects of TR7 S173L mutation present in QD35m2 and TR7 D128N mutation in QD35m3. The two TR7 variants were expressed in *E. coli*, purified and analyzed by native PAGE, revealing quick aggregation of TR7 S173L after size-exclusion chromatography purification (Supplementary Fig. 12), indicating the importance of S173 for proper folding. Furthermore, purified TR7 wild-type enzyme and variants were subjected to NADPH consumption assay. D128N retained catalytic activity comparable to the wild-type enzyme; in contrast, the S173L variant exhibited markedly reduced NADPH consumption activity, demonstrating its functional importance (Fig. 2f). Together with gLM2 and ESM2 predictions, structural analysis (Supplementary Fig. 13) and previous literature reports on homologous enzymes (Chhabra et al., 2012; Reid et al., 2003), these results demonstrate S173 as a catalytic residue in TR7. Because the S173L variant showed impaired enzyme activity, we next aimed to reduce mutation rates for improved diol production.

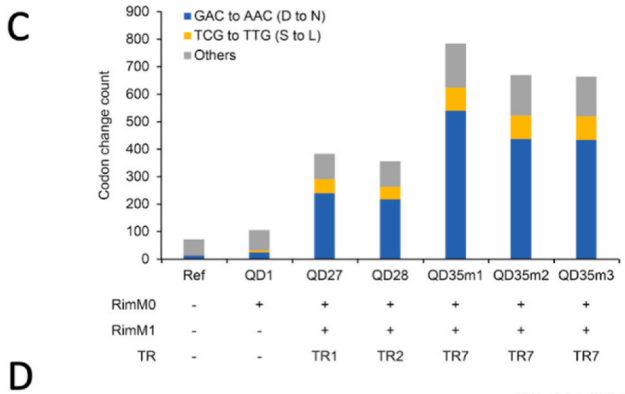
2.3. Detoxification strategies by orchestrating post-TR ADH catalysis

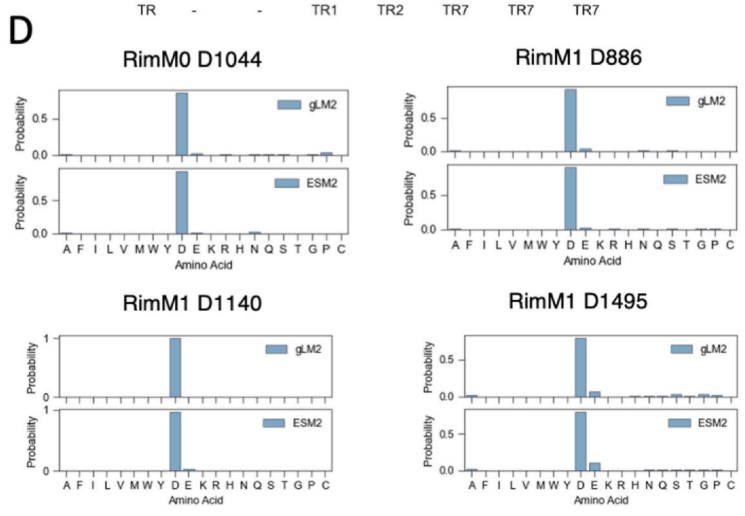
We hypothesized that mitigating aldehyde accumulation would be key to engineered *Streptomyces* genome stability and optimal diol production. To test this hypothesis and improve diol titers, we incorporated post-TR ADH catalysis for reducing potentially toxic hydroxyaldehydes to diols, and investigated the effects of both ADH co-integration and PKS-TR-ADH fusion strategies (Fig. 3a).

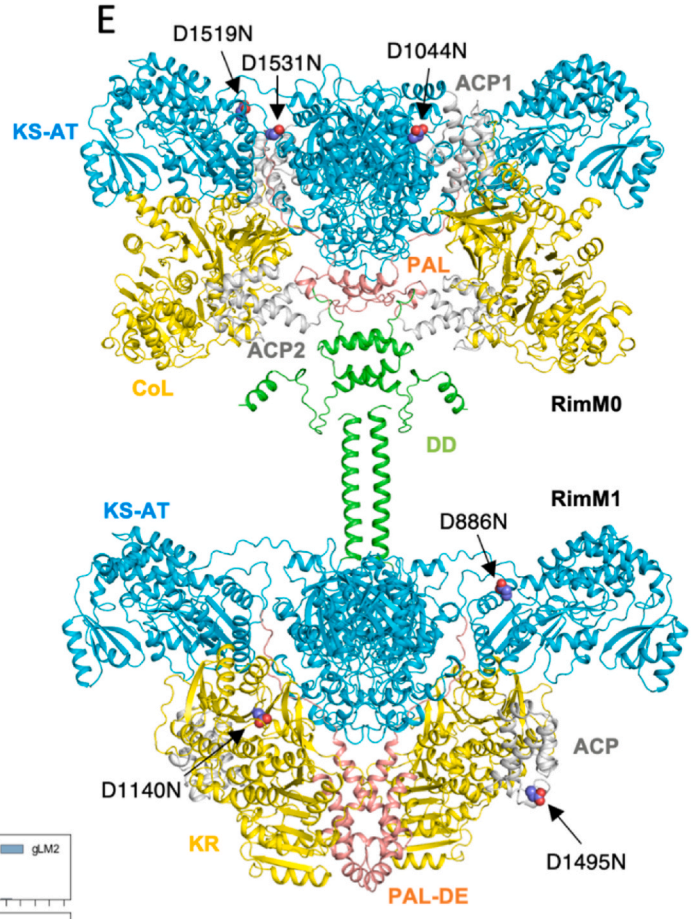
Unsure of the effects of dual phiC31 integration sites on genetic stability of QD35 strains with a *PrpsL(RO)* promoter for RimM0

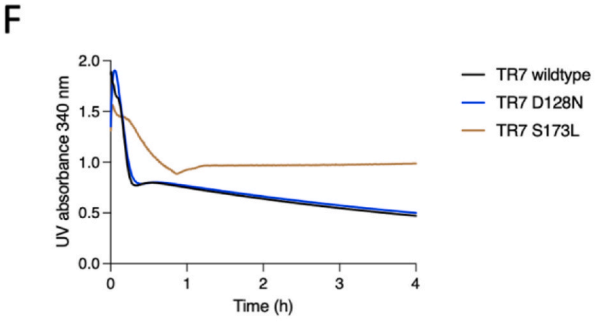












(caption on next page)

Fig. 2. TR7-associated genome-wide mutation accumulation and functional effects of PKS-TR mutations. (A) Genome-wide distribution of single nucleotide polymorphisms (SNPs) across engineered *S. albus* J1074 strains relative to the reference genome. Each colored tick denotes one SNP position. QD35 strains carrying RimM1-TR7 contained substantially more SNPs than the RimM0 controls (QD1) and strains carrying the previously characterized TR1 (QD27) or TR2 (QD28). (B) RimM1-TR7 integration resulted in a significant increase in genome-wide SNPs that are mainly missense mutations. (C) Distribution of codon substitutions among the detected variants. GAC to AAC (D to N) and TCG to TTG (S to L) substitutions were the predominant codon changes in RimM1-TR containing strains. (D) Model-assigned amino acid probabilities at RimM0 D1044 and RimM1 D886, D1140 and D1495 positions, calculated independently using gLM2 and ESM2. Both models strongly favored the native aspartate at each position, indicating that these positions may be functionally important. (E) AlphaFold 3 models of RimM0 and RimM1 showing the locations of observed D to N substitutions. The Asp residues are shown as spheres. Structural domains are colored as indicated. CoL: CoA ligase; KS: ketosynthase; AT: acyltransferase; PAL: post-AT linker; ACP: acyl carrier protein; KR: ketoreductase; DE: dimerization element; DD: docking domain. (F) *In vitro* NADPH consumption assay results for TR7 wild-type (WT), D128N and S173L mutants. D128N exhibited activity similar to WT, whereas S173L showed substantially reduced NADPH consumption.

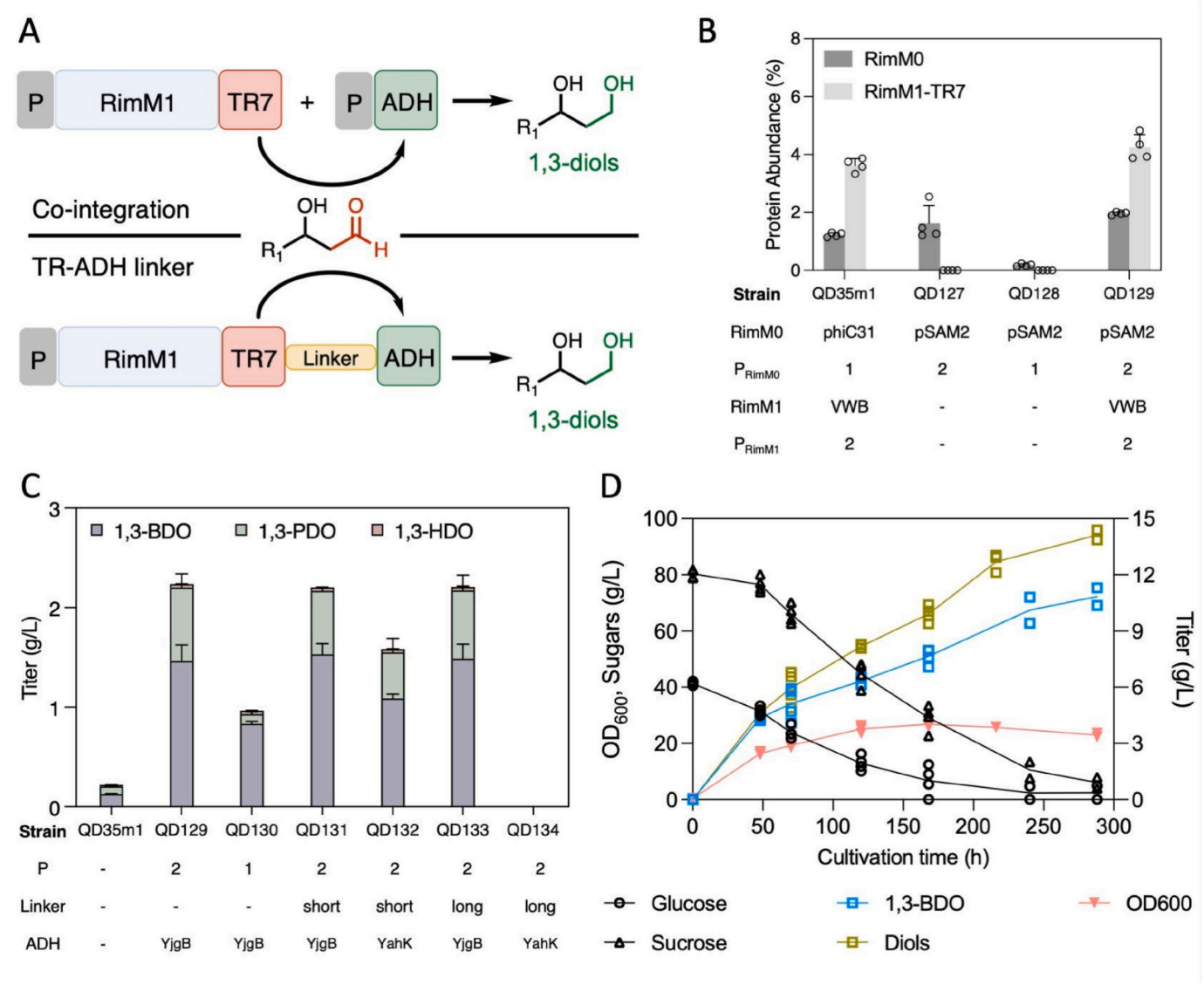


Fig. 3. Coordinated ADH expression improves diol production in engineered *S. albus*. (A) Schematic of two strategies designed to accelerate the conversion of RimPKS-TR-release 3-hydroxyaldehydes: co-integration of a standalone ADH, and a TR-linker-ADH fusion. P: promoter. (B) Relative abundances of RimM0 and RimM1-TR7 proteins. The genomic integration sites and promoter choices (1: *PrpsL(RO)*; 2: *Pgapdh(EL)*) are indicated below the graph. $n = 4$. (C) Diol production by strains expressing standalone or TR-fused YjgB or YahK ADH. Strains were cultivated for 3 d in R5 + 1% glucose in 24-well plates. $n = 4$ for QD35m1, $n = 2$ for QD131, $n = 3$ for the rest groups. (D) Profiles of sugar concentrations, cell growth and diol production during batch cultivation of QD129, achieving diol titers up to 14.2 g/L $n = 4$. Bars represent means, error bars indicate standard deviation.

expression, we switched to the pSAM2 integration system with a validated single integration site for *rimM0* and generated two *S. albus* J1074 strains (Supplementary Fig. 14). QD127 uses a constitutive *Pgapdh(EL)* promoter, whereas QD128 has a *PrpsL(RO)* promoter for RimM0 expression. Next we performed VWB-based co-integration of *rimM1-TR7* and *yjgB* ADH genes in QD127 and QD128, generating QD129 and QD130 strains, respectively. Consistent with our previous findings on promoter strength, RimM0 protein abundance in QD127 was higher than it in QD128 (Fig. 3b). We confirmed that our QD129 strain had comparable RimM0 and RimM1 protein abundances to the reference

QD35m1 strain, making these two excellent samples for diol production comparison (Fig. 3b). To evaluate whether physical linkage of TR and ADH domains could facilitate aldehyde intermediate tunneling between two active sites, we designed TR7-YjgB ADH fusion proteins using two different linker architectures, yielding strains QD131 with a shorter and artificial "GGGGSGGGGS" linker, and QD133 with a longer "LGAAG-PATVVQASAAALD" linker. The longer linker is the natural RimM1-M2 linker. In parallel, analogous TR7-YahK ADH fusion proteins were designed to yield strains QD132 with the shorter linker and QD134 with the longer linker to assess enzyme-specific effects on the fusion strategy.

Subsequent 3-day cultivation experiments revealed that all tested strains except for QD134 produced more diols than QD35m1 in a 24-well cultivation plate, with QD129 presenting the highest titers (Fig. 3c). With 1462.5 mg/L 1,3-BDO, 739.2 mg/L 1,3-PDO, and 33.8 mg/L 1,3-HDO production, this total diol titer marks a more than 10-fold increase compared to QD35m1. QD130 produced less diols than QD129, likely due to lower RimM0 protein abundance. As to TR-ADH fusion, the TR7-YjgB fusion strains, QD131 and QD133, achieved similar titers to QD129, confirming the effectiveness of fusion design. However, the strain harboring the longer linker between TR7 and YahK (QD134) did not produce any diol (Fig. 3c), suggesting that both ADH and fusion linker choices can critically influence the outcome.

Finally, we chose strain QD129 with the best performance for 1-L bioreactor cultivation experiments. After 12-day cultivation, we achieved a total diol titer of 14.2 g/L, with 10.8 g/L 1,3-BDO production and a maximum OD_{600} of 23.0, presenting a PKS-TR engineering record (Fig. 3d). Moreover, genome analysis revealed scarce SNPs in QD129 (Supplementary Fig. 15), demonstrating optimal strain stability. Taken together, these results demonstrate that post-TR ADH co-integration design alleviates cellular stress, improves pathway stability, and

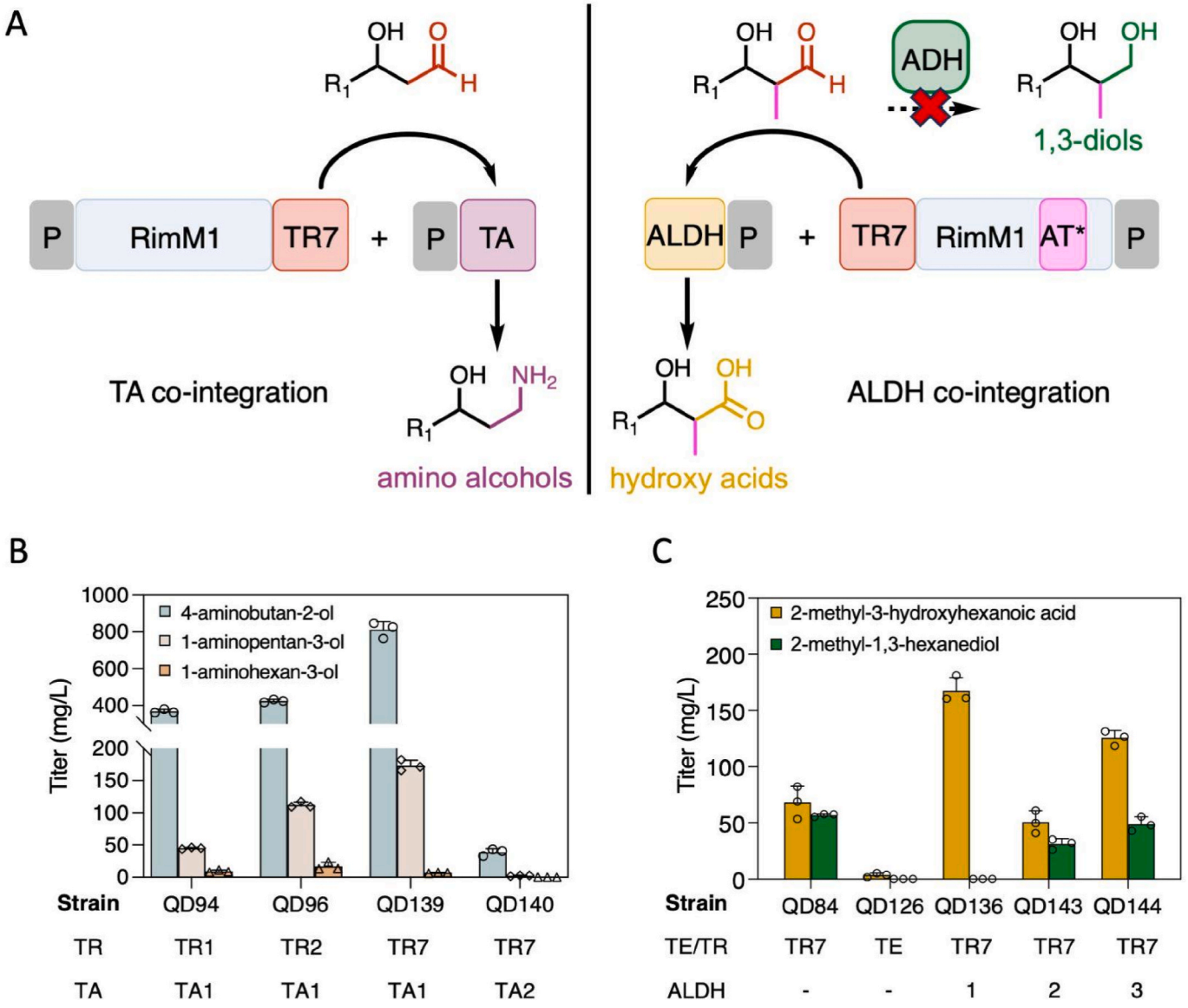


Fig. 4. Expansion of the co-integration strategy for amino alcohol and hydroxy acid production. (A) Schematic of product diversification through post-TR enzyme co-integration and AT exchange. Malonyl-CoA extension generates unbranched 3-hydroxyaldehydes that are converted by TAs into amino alcohols. Methylmalonyl-CoA extension via AT exchange generates a 3-hydroxy-2-methylaldehyde that can be reduced to a competing diol by endogenous ADHs or oxidized to a hydroxy acid by an introduced ALDH. (B) Amino alcohol production by strains expressing TR1/2/7 and TA1/2. TA1 is from *Streptomyces coelicolor* M145 and TA2 is from *Streptomyces* sp. NRRL-B24891. QD139 bearing TR7 and TA1 produced the highest total amino alcohol titer after 7-day cultivation in R5 + 2% glucose. (C) Production of 2-methyl-3-hydroxyhexanoic acid and the competing 2-methyl-1,3-hexanediol by strains carrying different ALDHs after 7-day cultivation in R5 + 2% glucose + 15 mM L-valine. ALDH1: PadA; ALDH2: ChnE; ALDH3: KDHba. QD126 serves as a control with DEBS TE in place of TR7. PadA co-integration eliminated detectable formation of the competing diol. Bars represent means, error bars indicate standard deviation. $n = 3$.

enables high-level, scalable diol production in engineered *Streptomyces albus*.

2.4. Expanding co-integration strategy to amino alcohol and hydroxy acid biosynthesis

With demonstrated use of TR and downstream ADH co-integration, we next investigated whether this strategy could be leveraged to improve other PKS-TR product titers, such as amino alcohols via post-TR transamination and hydroxy acids produced from aldehyde oxidation. To test this, we co-integrated different downstream genes, including TR-cognate TAs and standalone ALDHs, into the QD127 base strain to redirect aldehyde intermediates into distinct product classes (Fig. 4a).

First, we introduced TR and TA co-integration pathways to generate amino alcohol products. We previously reported that downstream expression of a TA (TA1) from coelimycin biosynthetic gene cluster (BGC) in QD94 (RimM0M1-TR1 + TA1) or QD96 (RimM0M1-TR2 + TA1) strain enabled production of amino alcohols with varying chain lengths, including 4-aminobutan-2-ol, 1-aminopentan-3-ol, and 1-aminohexan-3-ol (Dan et al., 2025). In each strain, *rimM0* and *rimM1-TR* genes were integrated in the phiC31 and the VWB integration sites respectively; *TA1* gene was integrated in the phiBT1 site. By applying RimM1-TR7 and TA co-integration into the VWB site, we generated TA1-containing QD139 and homologous TA2-containing QD140 strains from QD127. After 7-day cultivation in R5 + 2% glucose medium in shake flasks, QD139 reached the highest titers of 812.8 mg/L 4-aminobutan-2-ol, 173.2 mg/L 1-aminopentan-3-ol, and 7.0 mg/L 1-aminohexan-3-ol. Compared with QD96 where *TA1* integration occurred after *rimM1-TR2* integration, it marks at least 79% increase in total amino alcohol titers (Fig. 4b).

We then explored the oxidation of aldehyde intermediates to access hydroxy acids. Our primary target was 2-methyl-3-hydroxyhexanoic acid, because α-methyl-branched β-hydroxy acids have great potential as a monomer for structurally diversified polyhydroxyalkanoates (PHAs) (Dong et al., 2019). Previously, we applied (2*S*)-methylmalonyl CoA-specific AT exchange strategy and created QD84 strain (RimM0M1 (M7 AT)-TR7) via phiC31 integration of *rimM0*, VWB integration of *rimM1(M7 AT)-TR7*, and pSAM2 integration of *fkbS*. The QD84 strain produced both 68.2 mg/L 2-methyl-3-hydroxyhexanoic acid and 56.9 mg/L 2-methyl-1,3-hexanediol. Here, we first compared TR7 to the 6-deoxyerythronolide B synthase (DEBS) TE (Tsai et al., 2001), since PKS-TE can directly hydrolyze the PKS intermediate to carboxylic acids without aldehyde en route and is the canonical method for PKS-based hydroxy acid biosynthesis (Schmidt et al., 2025). Unexpectedly, RimM0M1(M7 AT)-TE (QD126 strain) resulted in production of only 3.6 mg/L 2-methyl-3-hydroxyhexanoic acid. Next, we tested expression of ALDH1 (PadA), ALDH2 (ChnE), and ALDH3 (KDHba) aldehyde dehydrogenases with TR7 in strains QD136, QD143, and QD144 respectively, to improve conversion of the aldehyde intermediate into 2-methyl-3-hydroxyhexanoic acid. Among the tested ALDHs, ALDH1 (PadA) resulted in the highest titers of 167.5 mg/L, presenting a 145% titer increase to QD84 and indicating the importance of ALDH choice (Fig. 4c). Additionally, we attempted to avoid diol byproducts that may compromise hydroxy acid titers and complicate product separation. Remarkably, no 2-methyl-1,3-hexanediol was detected in QD136 culture (Fig. 4c), demonstrating highly efficient post-TR reaction control.

We also tested whether screening TRs and ATs could improve the production of 2-ethyl-3-hydroxyhexanoic acid with a previously elusive biosynthesis route other than RimPKS-TR with AT exchange (Dan et al., 2025). Despite the effort, none of the tested (2*S*)-ethylmalonyl CoA-loading ATs or TRs including TR7 presented a significant increase in titers than our previous report (Supplementary Fig. 16a–b), suggesting that other PKS domains, such as KS and ketoreductase (KR), may be the main actors in ethyl branch gatekeeping.

Finally we examined the genomes of representative amino alcohol-producing QD139 and hydroxy acid-producing QD136 strains (Supplementary Fig. 17). Consistent with our findings in the QD129 strain orchestrated for diol production (Supplementary Fig. 15), QD139 and QD136 strains present substantially fewer total SNP mutations compared to previously reported strains. Moreover, for every examined strain with co-integration strategy, observed D to N amino acid substitution bias is absent in QD129 and significantly reduced in QD139 and QD136, regardless of product type (Supplementary Fig. 15c; Supplementary Fig. 17). The slightly higher D to N substitution rates in QD139 and QD136 compared to QD129 can be attributed to different catalytic efficiency of YjgB, TA1 and PadA enzymes used in aldehyde metabolism, in agreement with reported product titers. This observation further backs our hypothesis that the observed SNP pattern is a direct consequence of TR-mediated aldehyde release.

Together, these results establish TR-mediated aldehyde release as a programmable branch point for PKS product diversification. Co-integrating TRs with alcohol dehydrogenases, transaminases and aldehyde dehydrogenases efficiently redirected TR-derived aldehyde intermediates toward diols, amino alcohols and hydroxy acids respectively, while mitigating cellular stress, improving pathway stability and productivity, and limiting competing reactions. This coordinated integration of PKS downstream biocatalysis provides a general strategy for improving pathway performance and expanding PKS-derived chemical space, demonstrating the potential of orchestrated PKS retrobiosynthesis.

3. Discussion

Over the past few decades, advances in genome mining and directed evolution have vastly expanded chemical reaction repertoires available for use in living systems (Chen and Arnold, 2020; Nguyen et al., 2024). The full integration of these novel reactions into cellular metabolism has now become a rapidly emerging field, and it raises a fundamental question: how do microbes respond to the new reactions we engineer them to perform? We began to explore this question using our diol biosynthetic pipelines (Dan et al., 2025), where designed PKS-TRs produce large quantities of aldehyde intermediates that may be toxic to host cells. We hypothesized that this stress resulted in the accumulation of hundreds of genome-wide mutations, mostly SNPs. Through detailed investigations, we adopted a pathway co-integration strategy to mitigate the cellular burden; as a result, we significantly lowered mutation rates and achieved 14.2 g/L diol titers in bioreactors.

Aldehydes are common intermediates in engineered metabolic pathways, including pathways for fatty alcohol and diol biosynthesis. However, to the best of our knowledge, no previous report presents cellular stress and genome instability of the magnitude observed in this study. We reasoned that at least two factors may play a key role for this difference: enzyme architecture for aldehyde production and the detoxification capacity of the host cells. Fatty alcohol biosynthesis in *E. coli* commonly employs fatty acyl-CoA reductase (FAR), a dual functional enzyme converting fatty acids to fatty alcohols through a transient fatty aldehyde intermediate (Steen et al., 2010). As a result, the four-electron, two-step catalytic pattern limits the release of free aldehyde. Alternatively, pathways employing aldehyde-forming acyl-CoA reductases (ACRs) or carboxylic acid reductases (CARs) generally depend on heterologous or endogenous aldehyde reductases (ARs) or ADHs to complete alcohol production (Wang et al., 2020; Youngquist et al., 2013). *E. coli* is well known for possessing several AR or ADH genes with broad substrate scopes including *ybbO* and *yqhD* (Atsumi et al., 2010; Fatma et al., 2016). Deletion of the *ybbO* gene led to more than 90% long-chain fatty alcohol titer loss in engineered *E. coli* strains (Fatma et al., 2016). Although presence of endogenous ADH activity in *S. albus* J1074 was confirmed by diol formation upon PKS-TR expression, this detoxification capacity may be insufficient to alleviate aldehyde accumulation, cellular stress and mutation accumulation observed in our production strains.

Another important observation is that overexpression of TR1/2/7/9

in *E. coli* showed no obvious growth defect, ruling out possible impact of protein-driven toxicity. This result further suggests that catalytic product exposure is the principal source of toxicity. The predominant aldehyde products in our engineered *S. albus* strains include 3-hydroxybutanal, 3-hydroxypentanal, and 3-hydroxyhexanal. Notably, although the cellular toxicity of 3-hydroxy aldehydes has not been characterized, their dehydrated derivatives (α,β-unsaturated aldehydes including crotonaldehyde, trans-2-pentenal and trans-2-hexenal) were all shown to be mutagenic (Eder et al., 1992). These aldehydes form exocyclic deoxyguanosine adducts when tested *in vitro* with nucleosides and mononucleotides (Eder and Hoffman, 1993). Given the observation that the dominant SNP type observed in this study is G:C to A:T transition, this brings up an interesting hypothesis of deoxyguanosine adduct-based mutation mechanism.

Finally, we performed a process-based techno-economic analysis (TEA) to identify the principal cost drivers and performance targets for diol production. The current bulk price of 1,3-BDO ranges widely from $1.0/kg to $4.0/kg, depending on location, manufacturer, purity, and purchase quantity (Peters et al., 2003). At a conservative production capacity of 8500 MT/yr of 1,3-BDO, the estimated total capital investment and net annual operating costs were $137 million and $33 million, respectively (Supplementary Tables 9–10). Fermentation equipment was identified as the major capital investment contributor, whereas operating costs were primarily driven by raw materials and utilities (Supplementary Fig. 18a–b). The resulting minimum selling price (MSP) of $5.9/kg exceeded the reported commercial price range, so we sought to identify key operational parameters for further performance improvement through sensitivity analysis.

Consistent with other biomanufacturing processes, sensitivity analysis identified product yield as the most influential parameter affecting MSP, followed by feedstock costs and fermentation time (Supplementary Fig. 18c) (Baral et al., 2019; Ebrahimian and Mohammadi, 2024). Based on the sensitivity analysis, we also projected potential future cost reduction (Supplementary Fig. 18d). Optimization of selected parameters incrementally contributed to production cost reduction and ultimately resulted in an optimistic MSP of $1.2/kg. This projected price falls within the current market price range, highlighting the importance of further development through multifaceted optimization spanning strain engineering, fermentation, and process design.

4. Materials and methods

4.1. Materials, plasmids and strains

All the plasmids and strains constructed in this study are listed in Supplementary Table 11, except for *E. coli* BL21(DE3), DH5α and DH10β cells which were purchased from New England Biolabs. All gene sequences used in this study are listed in Supplementary Table 12. For plasmid construction, SAGE integration, and *S. albus* J1074 sporulation experiments, we followed previously published protocols (Dan et al., 2025). Plasmid sequences were validated by nanopore sequencing.

4.2. S. albus cultivation in 24-well plates and shake flasks

30 μl of wild-type or engineered *S. albus* J1074 spores were inoculated in 3 ml Tryptic Soy Broth (TSB) with nalidixic acid (25 μg ml^{-1}) in a 10 ml 24-well plate, and incubated at 30 °C, 200 rpm for 2 d to generate the seed culture. For 24-well plate-based cultivation experiments, 100 μl of the seed culture was inoculated in 3 ml R5 medium with 1% *w/v* glucose and nalidixic acid (25 μg ml^{-1}), and incubated at 30 °C, 200 rpm for another 3 d. For shake flask-based cultivation, 1 ml of the seed culture was added in 30 ml R5 with nalidixic acid (25 μg ml^{-1}) and varying concentrations of glucose, and incubated at 30 °C, 200 rpm for a selected amount of time. For all cultivation experiments, 2-5 sterile glass beads (4.5 mm) were added to reduce *Streptomyce* clumping effect.

4.3. S. albus batch and fed-batch cultivation in bioreactors

For seed culture preparation in bioreactor runs, 1 ml of engineered *S. albus* J1074 spores were inoculated in 50 ml TSB with nalidixic acid (25 μg ml^{-1}) in a 250 ml shake flask, and cultivated at 30 °C, 200 rpm for 2 d. 4 Eppendorf DASGIP bioreactors with 1.8 l max working volume were used at a starting volume of 1 l R5 medium with nalidixic acid (25 μg ml^{-1}) and 33 ml seed culture. Dissolved oxygen cascades at a 30% setpoint utilizing air flow rate between 102 and 120 sl/h. The DASGIP MX4x4 module controlled airflow of sterile dry air. Agitation stir rate changes were controlled from 300 to 600 rpm with Rushton impellers spaced at 1.5 diameter. Sampling was taken by discarding the dead-space of the dip-tube sample port. The setpoint for pH was 7 with a deadband range of 0.2; both acid (H_2SO_4, 15% *w/v*) and base (NH_4OH, 15% *w/v*) were controlled using the DASGIP MP8 modules. The cultivation temperature was 30 °C.

For QD101 fed-batch cultivation (Fig. 1E), glucose bolus additions were added with DASGIP MP8 modules in 20 ml increments over an hour of 50% *w/v* glucose on the 4th, 6th and 8th days. For QD125 fed-batch cultivation (Supplementary Fig. 6c–d), glucose or sucrose feeding was controlled utilizing MP8 modules and reactive bolus feeds, and the stir rate was 300-600 rpm.

4.4. HPLC and LC-MS analysis

1,3-PDO, 1,3-HDO, amino alcohol and hydroxy acid product quantification was performed by LC–MS using an Agilent LC/MSD iQ single quadrupole mass spectrometer as previously described (Dan et al., 2025). All authentic standards were commercially available.

Glucose, sucrose and 1,3-BDO product were analyzed and quantified using authentic standards from Sigma on an Agilent 1200 series Infinity HPLC system equipped with a refractive index detector (RID) and an Aminex HPX-87H Column (Bio-Rad). The mobile phase is 4 mM H_2SO_4 (in water) delivered at a flow rate of 0.40 ml min^{-1}. Each chromatographic run lasted 45 min. Both the column compartment and the RID were maintained at 45 °C throughout the analysis.

4.5. Proteomic analysis

Protein was extracted from cell pellets and tryptic peptides were prepared by following established proteomic sample preparation protocol (Chen et al., 2023). Briefly, cell pellets were resuspended in Qiagen P2 Lysis Buffer (Qiagen, Germany) to promote cell lysis. Proteins were precipitated with addition of 1 mM NaCl and 4 x vol acetone, followed by two additional washes with 80% acetone in water. The recovered protein pellet was homogenized by pipetting mixing with 100 mM ammonium bicarbonate in 20% methanol. Protein concentration was determined by the DC protein assay (BioRad, USA). Protein reduction was accomplished using 5 mM tris 2-(carboxyethyl)phosphine (TCEP) for 30 min at room temperature, and alkylation was performed with 10 mM iodoacetamide (IAM; final concentration) for 30 min at room temperature in the dark. Overnight digestion with trypsin was accomplished with a 1:50 trypsin:total protein ratio. The resulting peptide samples were analyzed on an Agilent 1290 UHPLC system coupled to a Thermo Scientific Orbitrap Exploris 480 mass spectrometer for discovery proteomics (Chen et al., 2022). Briefly, peptide samples were loaded onto an Ascentis® ES-C18 Column (Sigma–Aldrich, USA) and were eluted from the column by using a 10 min gradient from 98% solvent A (0.1% formic acid in H_2O) and 2% solvent B (0.1% formic acid in ACN) to 65% solvent A and 35% solvent B. Eluting peptides were introduced to the mass spectrometer operating in positive-ion mode and were measured in data-independent acquisition (DIA) mode with a duty cycle of 3 survey scans from m/z 380 to m/z 985 and 45 Tandem mass spectrometry (MS2) scans with precursor isolation width of 13.5 m/z to cover the mass range. DIA raw data files were analyzed by an integrated software suite DIA-NN (Demichev et al., 2020). The database used in the

DIA-NN search (library-free mode) is *Streptomyces albus* J1074 latest Uniprot proteome FASTA sequences plus the protein sequences of the heterologous proteins and common proteomic contaminants. DIA-NN determines mass tolerances automatically based on first pass analysis of the samples with automated determination of optimal mass accuracies. The retention time extraction window was determined individually for all MS runs analyzed via the automated optimization procedure implemented in DIA-NN. Protein inference was enabled, and the quantification strategy was set to Robust LC = High Accuracy. Output main DIA-NN reports were filtered with a global false discovery rate set at 0.01 (FDR $\leq$ 0.01) on both the precursor level and protein group level. The Top3 method, which is the average MS signal response of the three most intense tryptic peptides of each identified protein, was used to plot the quantity of the targeted proteins in the samples (Ahrné et al., 2013; Silva et al., 2006).

4.6. TR7 mutagenesis, enzyme purification and in vitro reactions

pHIS_*TR7* plasmid harboring D128N or S173L mutation was generated from pHIS_*TR7* wild-type plasmid using Agilent QuikChange XL Site-Directed Mutagenesis Kit. Plasmid sequences were validated by nanopore sequencing. Recombinant TR7 mutant protein expression, purification and enzymatic assays were carried out in the same way as previously described for the wild-type TR7 protein (Dan et al., 2025).

4.7. TR7-carrying genome assembly and annotation

The genome assembly ASM1921982v1 of *Streptomyces* sp. TRM70350 was used for further assembly and annotation of the putative TR7-containing BGC. Because no raw reads for this assembly and no closely related complete genomes were available, a total of 153 contigs were examined and joined based on sequence similarity. All contigs were compared pairwise to identify similar regions by using the Nucmer command (Marçais et al., 2018):

nucmer –maxmatch –minmatch 15 –mincluster 40 –nosimplify -p contigs_all

show-coords -rcl contigs_all.delta > contigs_all.coords)

A total of 3908 pairwise alignments were generated, and the aligned contigs were joined to obtain scaffolds using an in-house python script (scaffold_from_coords.py). Briefly, contigs were processed in descending order of contig length. An unused contig was selected as the seed of a new scaffold, and downstream extensions were constructed iteratively. At each extension step, among all unused contigs exhibiting an overlap with the current contig, the algorithm selected the candidate with the largest overlap length. The next contig was merged by trimming the overlapping region and appending or prepending only the non-overlapping sequence. Specifically, when the Nucmer alignment on the current contig was detected near the end of the current contig (i.e., when the alignment endpoint exceeded half of the current contig length), the algorithm appended the non-overlapping suffix of the next contig; otherwise, it prepended the non-overlapping prefix of the next contig. Incorporated contigs were marked as used, preventing them from being included in multiple scaffolds. The resulting concatenated sequences were written to an output FASTA file as individual scaffold entries. In total, 94 scaffolds were obtained, and they were re-annotated using Prokka v1.14.6 (Seemann, 2014). A scaffold containing the TR7 gene, was then used as an input for antiSMASH8.0 (Blin et al., 2025) to predict the TR7-containing BGC.

4.8. Engineered S. albus genome comparison for mutational analysis

1 ml *S. albus* seed culture was subjected to a Maxwell RSC Instrument for genome extraction. Extracted genomes were submitted to Plasmidsaurus for genome sequencing and assembly. The sequenced genome of engineered *S. albus* strains were compared to the reference *S. albus* J1074 genome (NCBI accession: NC_020990.1) to identify genome-wide mutations including SNP, indels, and large integrations. The sequenced genome were aligned to the reference genome by progressiveMauve (Darling et al., 2010) with the following parameters: default seed weight, determine LCBs (min LCB weight: default), assume collinear genomes, full alignment, iterative refinement, sum-of-pairs LCB scoring (scoring matrix: HOXD default). The information of SNP and indels were exported and processed using in-house perl scripts (SNP_annotator.pl, indel_annotator.pl, and mutation_process.pl).

4.9. Codon mutation rate calculation

Codon usage tables for *S. albus* were generated using an in-house Perl script (codon_usage.pl). In addition, the relative expected frequency (RelExpected) and relative observed frequency (RelObserved) for each codon change type in the QD27 and QD28 strains were calculated using two in-house Perl scripts (SNP_codon_information.pl, codon_frequency.pl). RelExpected refers to the expected frequency of a specific codon change based on the codon usage in *S. albus* and the base-change frequency observed in the strain. RelObserved refers to the frequency of a specific codon change among all observed codon changes.

Codon frequency among total codon = (the number of specific codon in *S. albus* genome)/(a total codon number in *S. albus* genome)

Base change frequency = (the number of specific base changes (e.g. A to G) in the strain)/(a total number of base (e.g. A) in *S. albus* genome)

Expected frequency = *Codon frequency among total codon* × *Base change frequency*.

RelExpected = (Expected frequency of the codon change)/(Expected frequency sum of total codon change)

RelObserved = (The number of the specific codon change (Observed))/(A total number of codon change (Observed))

4.10. Conservation and coevolution analysis

All data and code used in this analysis is publicly available on GitHub (https://github.com/Keasling-Lab/pks_evolutionary_analysis). Conservation and coevolution were assessed for TR7, RimM0 and RimM1 as described in Zhang et al. (2024); Zhang et al. (2024) using ESM2 (esm2_t33_650M_UR50D) and gLM2 (gLM2_650M). Per-position conservation was computed as each model's masked token distribution where each residue was masked in turn, the protein sequence was passed through the model, and the output logits over the 20 AA at the masked position were retained and converted to conditional probabilities by a softmax over those 20 states. Pairwise coevolution was computed from the categorical Jacobian where each position was substituted to each of the 20 AA in turn, all 20 evaluated explicitly rather than by the single mask approximation, and the differences in the output logits at all positions relative to the unperturbed sequence were recorded. The two models were compared at ten residues, namely, TR7 D128 and S173; RimM0 D830, D1044, D1519 and D1531; and RimM1 D417, D886, D1140 and D1495. At each position the masked token AA distributions from the two models were plotted side by side, and a submatrix of the coevolution map spanning $\pm$10 residues around the site was displayed.

4.11. Techno-economic analysis

The techno-economic analysis (TEA) was performed based on the experimental results using SuperPro Designer (v13.2) and Microsoft® Excel. The flow diagrams of the proposed diol production processes are shown in Fig. S19 and Fig. S20. The material and energy balances of the process were identified in the software and used to determine the number of pieces of equipment. The plant was assumed to operate for 7920 h per year. The currency used in this study (US dollar) was adjusted to the year 2025. Equipment cost was derived from the SuperPro software database and adjusted to 2025 using the Chemical Engineering Plant Cost Index (CEPCI) (Jenkins, 2026). The direct and indirect plant costs were factored from the total equipment costs based

on the literature (Peters et al., 2003). The working capital is set to be 5% of the fixed capital investment. The unit prices of raw materials and co-product credits were collected from literature, vendor quotations, and surveys. Utilities, including electricity, natural gas, steam, and cooling water, were based on the SuperPro database. Depreciation was calculated based on a linear 10-year depreciation schedule to fully recover the depreciation for the 30-year plant. The detailed factors and assumptions are summarized in Supplementary Tables 5–8. In this study, the plant lifetime was estimated to be 30 years, with two years of construction and start-up. The plant was assumed to be 40% equity financed. Income taxes were calculated as 21% of the taxable income. The minimum selling prices of 1,3-BDO under different scenarios were estimated using discounted cash flow rate of return analysis. Moreover, a sensitivity analysis was conducted by fluctuating the key process parameter by 30%.

CRediT authorship contribution statement

Qingyun Dan: Writing – review & editing, Writing – original draft, Project administration, Methodology, Investigation, Conceptualization. **Xina Wang:** Writing – original draft, Methodology, Investigation. **Yan Chiu:** Methodology, Investigation. **Soonkyu Hwang:** Writing – original draft, Methodology, Investigation. **Justin J. Baerwald:** Methodology, Investigation. **Yilin Li:** Writing – original draft, Methodology, Investigation. **Lucas M. Waldburger:** Methodology, Investigation. **Chang Dou:** Methodology, Investigation. **Yiou Rong:** Methodology, Investigation. **Xavier B. Tao:** Methodology, Investigation. **Chenyi Li:** Methodology, Investigation. **Ning Qin:** Methodology, Investigation. **Zhen Liu:** Methodology, Investigation. **Seokjung Cheong:** Methodology, Investigation. **Yan Chen:** Methodology, Investigation. **Jennifer W. Gin:** Methodology, Investigation. **Matthias Schmidt:** Methodology, Investigation. **Rachel E. Sullivan:** Methodology, Investigation. **Tyler W.H. Backman:** Writing – review & editing, Resources, Project administration, Methodology, Investigation. **Christopher J. Petzold:** Writing – review & editing, Resources, Project administration, Methodology, Investigation. **Jay D. Keasling:** Writing – review & editing, Resources, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interests

J.D.K. has financial interests in Ansa Biotechnologies, Apertor Pharma, Berkeley Yeast, BioMia, Cyklos Materials, Demetrix, Lygos, Napigen, QSX Bio, ResVita Bio, and Zero Acre Farms. The other authors declare no competing interests. Correspondence and requests for materials should be addressed to J.D.K (keasling@berkeley.edu). and Q.D. (qdan@jhu.edu). We have filed a patent related to the subject matter of the contribution.

Acknowledgements

This material is based upon work at the Joint BioEnergy Institute (JBEI) supported by the U.S. Department of Energy, Office of Science, Biological and Environmental Research Program under contract DE-AC02-05CH11231 with Lawrence Berkeley National Laboratory. Q.D. and X.W. were also supported by the Bakar fellowship granted to J.D.K.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ymben.2026.102570.

Data availability

All study data are included in the article and/or Supplementary data. All code used in this study is publicly available on GitHub (https://github.com/Keasling-Lab/).

References

Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A.J., Bambrick, J., Bodenstein, S.W., Evans, D.A., Hung, C.-C., O'Neill, M., Reiman, D., Tunyasuvunakool, K., Wu, Z., Žemgulytė, A., Arvaniti, E., Beattie, C., Bertolli, O., Bridgland, A., Cherepanov, A., Congreve, M., Cowen-Rivers, A.I., Cowie, A., Figurnov, M., Fuchs, F.B., Gladman, H., Jain, R., Khan, Y.A., Low, C.M.R., Perlin, K., Potapenko, A., Savy, P., Singh, S., Stecula, A., Thillaisundaram, A., Tong, C., Yakneen, S., Zhong, E.D., Zielinski, M., Žídek, A., Bapst, V., Kohli, P., Jaderberg, M., Hassabis, D., Jumper, J.M., 2024. Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature 630, 493–500. https://doi.org/10.1038/s41586-024-07487-w.

Ahrné, E., Molzahn, L., Glatter, T., Schmidt, A., 2013. Critical assessment of proteome-wide label-free absolute abundance estimation strategies. Proteomics 13, 2567–2578. https://doi.org/10.1002/pmic.201300135.

Alper, J., 1999. Engineering metabolism for commercial gains. Science 283, 1625–1626. https://doi.org/10.1126/science.283.5408.1625.

Atsumi, S., Wu, T.-Y., Eckl, E.-M., Hawkins, S.D., Buelter, T., Liao, J.C., 2010. Engineering the isobutanol biosynthetic pathway in Escherichia coli by comparison of three aldehyde reductase/alcohol dehydrogenase genes. Appl. Microbiol. Biotechnol. 85, 651–657. https://doi.org/10.1007/s00253-009-2085-6.

Awodi, U.R., Ronan, J.L., Masschelein, J., de los Santos, E.L.C., Challis, G.L., 2017. Thioester reduction and aldehyde transamination are universal steps in actinobacterial polyketide alkaloid biosynthesis. Chem. Sci. 8, 411–415. https://doi.org/10.1039/c6sc02803a.

Baral, N.R., Kavvada, O., Mendez-Perez, D., Mukhopadhyay, A., Lee, T.S., Simmons, B.A., Scown, C.D., 2019. Techno-economic analysis and life-cycle greenhouse gas mitigation cost of five routes to bio-jet fuel blendstocks. Energy Environ. Sci. 12, 807–824. https://doi.org/10.1039/c8ee03266a.

Biebl, H., Menzel, K., Zeng, A.-P., Deckwer, W.-D., 1999. Microbial production of 1,3-propanediol. Appl. Microbiol. Biotechnol. 52, 289–297. https://doi.org/10.1007/s002530051523.

Blin, K., Shaw, S., Vader, L., Szenei, J., Reitz, Z.L., Augustijn, H.E., Cediel-Becerra, J.D.D., de Crécy-Lagard, V., Koetsier, R.A., Williams, S.E., Cruz-Morales, P., Wongwas, S., Segurado Luchsinger, A.E., Biermann, F., Korenskaia, A., Zdouc, M.M., Meijer, D., Terlouw, B.R., van der Hooft, J.J.J., Ziemert, N., Helfrich, E.J.N., Masschelein, J., Corre, C., Chevrette, M.G., van Wezel, G.P., Medema, M.H., Weber, T., 2025. antiSMASH 8.0: extended gene cluster detection capabilities and analyses of chemistry, enzymology, and regulation. Nucleic Acids Res. 53, W32–W38. https://doi.org/10.1093/nar/gkaf334.

Chen, K., Arnold, F.H., 2020. Engineering new catalytic activities in enzymes. Nat. Catal. 3, 203–213. https://doi.org/10.1038/s41929-019-0385-5.

Chen, Y., Gin, J., Petzold, C., 2022. Discovery proteomic (DIA) LC-MS/MS data acquisition and analysis V.2. *Protocols.io*. https://www.protocols.io/view/discovery-proteomic-dia-lc-ms-ms-data-acquisition-e6nvwk1z7vmk/v2/ accessed 8.9.26.

Chen, Y., Gin, J., Petzold, C., 2023. Alkaline-SDS cell lysis of microbes with acetone protein precipitation for proteomic sample preparation in 96-well plate format. PLoS One 18, e0288102. https://doi.org/10.1371/journal.pone.0288102.

Chhabra, A., Haque, A.S., Pal, R.K., Goyal, A., Rai, R., Joshi, S., Panjikar, S., Pasha, S., Sankaranarayanan, R., Gokhale, R.S., 2012. Nonprocessive [2 + 2]e- off-loading reductase domains from mycobacterial nonribosomal peptide synthetases. Proc. Natl. Acad. Sci. 109, 5681–5686. https://doi.org/10.1073/pnas.1118680109.

Cogan, D.P., Zhang, K., Li, X., Li, S., Pintilie, G.D., Roh, S.-H., Craik, C.S., Chiu, W., Khosla, C., 2021. Mapping the catalytic conformations of an assembly-line polyketide synthase module. Science 374, 729–734. https://doi.org/10.1126/science.abi8358.

Cornman, A., West-Roberts, J., Camargo, A.P., Roux, S., Beracochea, M., Mirdita, M., Ovchinnikov, S., Hwang, Y., 2024. The OMG dataset: an Open MetaGenomic corpus for mixed-modality genomic language modeling. https://doi.org/10.1101/2024.08.14.607850.

Cortes, J., Haydock, S.F., Roberts, G.A., Bevitt, D.J., Leadlay, P.F., 1990. An unusually large multifunctional polypeptide in the erythromycin-producing polyketide synthase of Saccharopolyspora erythraea. Nature 348, 176–178. https://doi.org/10.1038/348176a0.

Cullen, L., Meng, F., Lupton, R., Cullen, J.M., 2024. Reducing uncertainties in greenhouse gas emissions from chemical production. Nat. Chem. Eng. 1, 311–322. https://doi.org/10.1038/s44286-024-00047-z.

Dan, Q., Chiu, Y., Lee, N., Pereira, J.H., Rad, B., Zhao, X., Deng, K., Rong, Y., Zhan, C., Chen, Y., Cheong, S., Li, C., Gin, J.W., Rodrigues, A., Northen, T.R., Backman, T.W. H., Baidoo, E.E.K., Petzold, C.J., Adams, P.D., Keasling, J.D., 2025. A polyketide-based biosynthetic platform for diols, amino alcohols and hydroxy acids. Nat. Catal. 8, 147–161. https://doi.org/10.1038/s41929-025-01299-5.

Darling, A.E., Mau, B., Perna, N.T., 2010. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS One 5, e11147. https://doi.org/10.1371/journal.pone.0011147.

Demichev, V., Messner, C.B., Vernardis, S.I., Lilley, K.S., Ralser, M., 2020. DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. Nat. Methods 17, 41–44. https://doi.org/10.1038/s41592-019-0638-x.

Donadio, S., Staver, M.J., McAlpine, J.B., Swanson, S.J., Katz, L., 1991. Modular organization of genes required for complex polyketide biosynthesis. Science 252, 675–679. https://doi.org/10.1126/science.2024119.

Dong, H., Liffland, S., Hillmyer, M.A., Chang, M.C.Y., 2019. Engineering in vivo production of α-Branched polyesters. J. Am. Chem. Soc. 141, 16877–16883. https://doi.org/10.1021/jacs.9b08585.
Ebrahimian, F., Mohammadi, A., 2024. Bioprocess design and technoeconomic analysis of 2,3-Butanediol production in wood-based biorefineries. Waste Biomass Valoriz. 15, 6635–6648. https://doi.org/10.1007/s12649-024-02611-3.
Eder, E., Deininger, C., Neudecker, T., Deininger, D., 1992. Environmental and molecular. Mutagenesis 19, 338–345. https://doi.org/10.1002/em.2850190413.
Eder, E., Hoffman, C., 1993. Identification and characterization of deoxyguanosine adducts of mutagenic .beta.-alkyl-substituted acrolein congeners. Chem. Res. Toxicol. 6, 486–494. https://doi.org/10.1021/tx00034a015.
Elibol, M., Mavituna, F., 1998. Effect of sucrose on actinorhodin production by *Streptomyces coelicolor* A3(2). Process Biochem 33, 307–311. https://doi.org/10.1016/S0032-9592(97)00088-5.
Erb, T.J., Berg, I.A., Brecht, V., Müller, M., Fuchs, G., Alber, B.E., 2007. Synthesis of C5-dicarboxylic acids from C2-units involving crotonyl-CoA carboxylase/reductase: the ethylmalonyl-CoA pathway. Proc. Natl. Acad. Sci. 104, 10631–10636. https://doi.org/10.1073/pnas.0702791104.
Ewing, T.A., Nouse, N., van Lint, M., van Haveren, J., Hugenholtz, J., van Es, D.S., 2022. Fermentation for the production of biobased chemicals in a circular economy: a perspective for the period 2022–2050. Green Chem. 24, 6373–6405. https://doi.org/10.1039/d1gc04758b.
Fatma, Z., Jawed, K., Mattam, A.J., Yazdani, S.S., 2016. Identification of long chain specific aldehyde reductase and its use in enhanced fatty alcohol production in *E. coli*. Metab. Eng. 37, 35–45. https://doi.org/10.1016/j.ymben.2016.04.003.
Gahloth, D., Dunstan, M.S., Quaglia, D., Klumbys, E., Lockhart-Cairns, M.P., Hill, A.M., Derrington, S.R., Scrutton, N.S., Turner, N.J., Leys, D., 2017. Structures of carboxylic acid reductase reveal domain dynamics underlying catalysis. Nat. Chem. Biol. 13, 975–981. https://doi.org/10.1038/nchembio.2434.
Gao, J., Li, Y., Yu, W., Zhou, Y.J., 2022. Rescuing yeast from cell death enables overproduction of fatty acids from sole methanol. Nat. Metab. 4, 932–943. https://doi.org/10.1038/s42255-022-00601-0.
Gläser, L., Kuhl, M., Stegmüller, J., Rückert, C., Myronovskyi, M., Kalinowski, J., Luzhetskyy, A., Wittmann, C., 2021. Superior production of heavy pamamycin derivatives using a bkdR deletion mutant of Streptomyces albus J1074/R2. Microb. Cell Fact. 20, 111. https://doi.org/10.1186/s12934-021-01602-6.
Horsman, M.E., Hari, T.P.A., Boddy, C.N., 2016. Polyketide synthase and non-ribosomal peptide synthetase thioesterase selectivity: logic gate or a victim of fate? Nat. Prod. Rep. 33, 183–202. https://doi.org/10.1039/c4np00148f.
Jenkins, S.. 2026 CEPCI updates: May (prelim.) and April (final). https://www.chemengonline.com/2026-cepci-updates-may-prelim-and-april-final/, 8.9.26.
Lee, N., Schmidt, M., Li, C., Filbin, C.J., Klass, S., Yaguchi, A., Fear, A.L., Backman, T.W. H., Abel, B.A., Kim, Y.-M., Kim, W.E., Johnson, C.W., Chen, Y., Gin, J.W., Petzold, C. J., Baidoo, E.E.K., Gao, Y., Burnet, M.C., Babnigg, G., Laible, P.D., Singer, C.A., Michener, W.E., Salvachúa, D., Garcia Martin, H., Haushalter, R.W., Keasling, J.D., 2025. Retrobiosynthesis of unnatural lactams via reprogrammed polyketide synthase. Nat. Catal. 8, 389–402. https://doi.org/10.1038/s41929-025-01325-6.
Lin, Z., Akin, H., Rao, R., Hie, B., Zhu, Z., Lu, W., Smetanin, N., Verkuil, R., Kabeli, O., Shmueli, Y., dos Santos Costa, A., Fazel-Zarandi, M., Sercu, T., Candido, S., Rives, A., 2023. Evolutionary-scale prediction of atomic-level protein structure with a language model. Science 379, 1123–1130. https://doi.org/10.1126/science.ade2574.
Liu, Y., Cruz-Morales, P., Zargar, A., Belcher, M.S., Pang, B., Englund, E., Dan, Q., Yin, K., Keasling, J.D., 2021. Biofuels for a sustainable future. Cell 184, 1636–1647. https://doi.org/10.1016/j.cell.2021.01.052.
Liu, Y., Zhang, C., Zeng, A.-P., 2024. Advances in biosynthesis and downstream processing of diols. Biotechnol. Adv. 77, 108455. https://doi.org/10.1016/j.biotechadv.2024.108455.
Marçais, G., Delcher, A.L., Phillippy, A.M., Coston, R., Salzberg, S.L., Zimin, A., 2018. MUMmer4: a fast and versatile genome alignment system. PLoS Comput. Biol. 14, e1005944. https://doi.org/10.1371/journal.pcbi.1005944.
McGovern, P.E., Zhang, J., Tang, J., Zhang, Z., Hall, G.R., Moreau, R.A., Nuñez, A., Butrym, E.D., Richards, M.P., Wang, Chen-shan, Cheng, G., Zhao, Z., Wang, Changsui, 2004. Fermented beverages of pre- and proto-historic China. Proc. Natl. Acad. Sci. 101, 17593–17598. https://doi.org/10.1073/pnas.0407921102.
Menzella, H.G., Reid, R., Carney, J.R., Chandran, S.S., Reisinger, S.J., Patel, K.G., Hopwood, D.A., Santi, D.V., 2005. Combinatorial polyketide biosynthesis by de novo design and rearrangement of modular polyketide synthase genes. Nat. Biotechnol. 23, 1171–1176. https://doi.org/10.1038/nbt1128.
Merrick, C.A., Zhao, J., Rosser, S.J., 2018. Serine integrases: advancing synthetic biology. ACS Synth. Biol. 7, 299–310. https://doi.org/10.1021/acssynbio.7b00308.
Mullowney, M.W., McClure, R.A., Robey, M.T., Kelleher, N.L., Thomson, R.J., 2018. Natural products from thioester reductase containing biosynthetic pathways. Nat. Prod. Rep. 35, 847–878. https://doi.org/10.1039/c8np00013a.
Nakamura, C.E., Whited, G.M., 2003. Metabolic engineering for the microbial production of 1,3-propanediol. Curr. Opin. Biotechnol. 14, 454–459. https://doi.org/10.1016/j.copbio.2003.08.005.
Nguyen, D.T., Mitchell, D.A., van der Donk, W.A., 2024. Genome mining for new enzyme chemistry. ACS Catal. 14, 4536–4553. https://doi.org/10.1021/acscatal.3c06322.
Peters, M.S., Timmerhaus, K.D., West, R.E., 2003. Plant Design and Economics for Chemical Engineers. McGraw-Hill.
Reid, R., Piagentini, M., Rodriguez, E., Ashley, G., Viswanathan, N., Carney, J., Santi, D. V., Hutchinson, C.R., McDaniel, R., 2003. A model of structure and catalysis for ketoreductase domains in modular polyketide synthases. Biochemistry 42, 72–79. https://doi.org/10.1021/bi0268706.
Schmidt, M., Vilchez, A.A., Lee, N., Keiser, L.S., Pearson, A.N., Thompson, M.G., Zhu, Y., Haushalter, R.W., Deutschbauer, A.M., Yuzawa, S., Blank, L.M., Keasling, J.D., 2025. Engineering *Pseudomonas putida* for production of 3-hydroxyacids using hybrid type I polyketide synthases. Metab. Eng. Commun. 20, e00261. https://doi.org/10.1016/j.mec.2025.e00261.
Schüth, F., Schunk, S.A., 2026. Transitioning of the chemical industry toward a net-zero carbon dioxide emission path. Angew. Chem. Int. Ed. 65, e22234. https://doi.org/10.1002/anie.202522234.
Seemann, T., 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30, 2068–2069. https://doi.org/10.1093/bioinformatics/btu153.
Silva, J.C., Gorenstein, M.V., Li, G.-Z., Vissers, J.P.C., Geromanos, S.J., 2006. Absolute quantification of proteins by LCMSE. Mol. Cell. Proteomics 5, 144–156. https://doi.org/10.1074/mcp.M500230-MCP200.
Steen, E.J., Kang, Y., Bokinsky, G., Hu, Z., Schirmer, A., McClure, A., del Cardayre, S.B., Keasling, J.D., 2010. Microbial production of fatty-acid-derived fuels and chemicals from plant biomass. Nature 463, 559–562. https://doi.org/10.1038/nature08721.
Tsai, S.-C., Miercke, L.J.W., Krucinski, J., Gokhale, R., Chen, J.C.-H., Foster, P.G., Cane, D.E., Khosla, C., Stroud, R.M., 2001. Crystal structure of the macrocycle-forming thioesterase domain of the erythromycin polyketide synthase: versatility from a unique substrate channel. Proc. Natl. Acad. Sci. 98, 14808–14813. https://doi.org/10.1073/pnas.011399198.
Wang, J., Li, C., Zou, Y., Yan, Y., 2020. Bacterial synthesis of C3-C5 diols via extending amino acid catabolism. Proc. Natl. Acad. Sci. 117, 19159–19167. https://doi.org/10.1073/pnas.2003032117.
Wang, Z., Cheong, S., Wang, H., Demarteau, J., Epstein, A.R., Bose, B., Hu, W., Schmidt, M., Keiser, L., Zhao, M., Ge, B., Nava, A.A., Kakumanu, R., Baidoo, E.E.K., Chen, Y., Petzold, C.J., Guo, Y., Baral, N.R., Vora, N., Nordahl, S.L., Liu, Y., Ghanta, R., Cruz-Morales, P., Yin, K., Haushalter, R.W., Persson, K.A., Scown, C.D., Helms, B.A., Keasling, J.D., 2026. Engineered polyketide synthases enable a microbial chassis for recyclable plastics with tunable properties. Nat. Biotechnol. 1–10. https://doi.org/10.1038/s41587-026-03229-7.
Yim, H., Haselbeck, R., Niu, W., Pujol-Baxley, C., Burgard, A., Boldt, J., Khandurina, J., Trawick, J.D., Osterhout, R.E., Stephen, R., Estadilla, J., Teisan, S., Schreyer, H.B., Andrae, S., Yang, T.H., Lee, S.Y., Burk, M.J., Van Dien, S., 2011. Metabolic engineering of Escherichia coli for direct production of 1,4-butanediol. Nat. Chem. Biol. 7, 445–452. https://doi.org/10.1038/nchembio.580.
Youngquist, J.T., Schumacher, M.H., Rose, J.P., Raines, T.C., Politz, M.C., Copeland, M. F., Pfleger, B.F., 2013. Production of medium chain length fatty alcohols from glucose in *Escherichia coli*. Metab. Eng. 20, 177–186. https://doi.org/10.1016/j.ymben.2013.10.006.
Yu, Z., Lv, H., Wu, Y., Wei, T., Yang, S., Ju, D., Chen, S., 2019. Enhancement of FK520 production in Streptomyces hygroscopicus by combining traditional mutagenesis with metabolic engineering. Appl. Microbiol. Biotechnol. 103, 9593–9606. https://doi.org/10.1007/s00253-019-10192-8.
Yuzawa, S., Mirsiaghi, M., Jocic, R., Fujii, T., Masson, F., Benites, V.T., Baidoo, E.E.K., Sundstrom, E., Tanjore, D., Pray, T.R., George, A., Davis, R.W., Gladden, J.M., Simmons, B.A., Katz, L., Keasling, J.D., 2018. Short-chain ketone production by engineered polyketide synthases in Streptomyces albus. Nat. Commun. 9, 4569. https://doi.org/10.1038/s41467-018-07040-0.
Zhang, Z., Wayment-Steele, H.K., Brixi, G., Wang, H., Kern, D., Ovchinnikov, S., 2024. Protein language models learn evolutionary statistics of interacting sequence motifs. Proc. Natl. Acad. Sci. 121, e2406285121. https://doi.org/10.1073/pnas.2406285121.