# Highly Active and Specific Tyrosine Ammonia-Lyases from Diverse Origins Enable Enhanced Production of Aromatic Compounds in Bacteria and Saccharomyces cerevisiae 

Jendresen, Christian Bille; Stahlhut, Steen Gustav; Li, Mingji; Gaspar, Paula; Siedler, Solvej; Förster, Jochen; Maury, Jerome; Borodina, Irina; Nielsen, Alex Toftgaard

Published in:
Applied and Environmental Microbiology

Link to article, DOI:
10.1128/AEM.00405-15

Publication date:
2015

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):
Jendresen, C. B., Stahlhut, S. G., Li, M., Gaspar, P., Siedler, S., Förster, J., Maury, J., Borodina, I., \& Nielsen, A. T. (2015). Highly Active and Specific Tyrosine Ammonia-Lyases from Diverse Origins Enable Enhanced Production of Aromatic Compounds in Bacteria and Saccharomyces cerevisiae. Applied and Environmental Microbiology, 81(13), 4458-4476. https://doi.org/10.1128/AEM.00405-15

[^0]
# Highly Active and Specific Tyrosine Ammonia-Lyases from Diverse Origins Enable Enhanced Production of Aromatic Compounds in Bacteria and Saccharomyces cerevisiae 

(1) Christian Bille Jendresen, Steen Gustav Stahlhut, Mingji Li, Paula Gaspar, Solvej Siedler, Jochen Förster, Jérôme Maury, Irina Borodina, Alex Toftgaard Nielsen

The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Hørsholm, Denmark


#### Abstract

Phenylalanine and tyrosine ammonia-lyases form cinnamic acid and $p$-coumaric acid, which are precursors of a wide range of aromatic compounds of biotechnological interest. Lack of highly active and specific tyrosine ammonia-lyases has previously been a limitation in metabolic engineering approaches. We therefore identified 22 sequences in silico using synteny information and aiming for sequence divergence. We performed a comparative in vivo study, expressing the genes intracellularly in bacteria and yeast. When produced heterologously, some enzymes resulted in significantly higher production of $p$-coumaric acid in several different industrially important production organisms. Three novel enzymes were found to have activity exclusively for phenylalanine, including an enzyme from the low-GC Gram-positive bacterium Brevibacillus laterosporus, a bacterial-type enzyme from the amoeba Dictyostelium discoideum, and a phenylalanine ammonia-lyase from the moss Physcomitrella patens (producing $230 \mu \mathrm{M}$ cinnamic acid per unit of optical density at $600 \mathrm{~nm}\left[\mathrm{OD}_{600}\right]$ ) in the medium using Escherichia coli as the heterologous host). Novel tyrosine ammonia-lyases having higher reported substrate specificity than previously characterized enzymes were also identified. Enzymes from Herpetosiphon aurantiacus and Flavobacterium johnsoniae resulted in high production of $p$-coumaric acid in Escherichia coli (producing $440 \mu \mathrm{M} p$-coumaric acid $\mathrm{OD}_{600}$ unit $^{-1}$ in the medium) and in Lactococcus lactis. The enzymes were also efficient in Saccharomyces cerevisiae, where $p$-coumaric acid accumulation was improved 5 -fold over that in strains expressing previously characterized tyrosine ammonia-lyases.


Small organic molecules of biotechnological interest include aromatic structures that are derived from $p$-coumaric acid ( pHCA ). pHCA can be formed from phenylalanine either through deamination to cinnamic acid (CA) by phenylalanine ammonialyase (PAL; EC 4.3.1.24) and subsequent hydroxylase activity by trans-cinnamate-4-monooxygenase or through deamination of tyrosine by tyrosine ammonia-lyase (TAL; EC 4.3.1.23) (Fig. 1). Considering that the route from phenylalanine requires activity of a P450 enzyme, which has been shown to be the limiting step in previous engineering strategies ( 1,2 ), deamination of tyrosine for production of pHCA may be preferred in microbial cell factories. Tyrosine ammonia-lyases (TAL) have been employed for the production of plant biochemicals and aromatic compounds, e.g., for pHCA itself ( $3-5$ ), or in the production of stilbenes, such as resveratrol $(6,7)$, flavonoids, such as naringenin $(8,9)$, cinnamoyl anthranilates (10), or plastic precursors, such as $p$-hydroxystyrene $(11,12)$, yet the TAL reaction may be the limiting step ( 10,13 ).

Due to the significant interest in production of phenolic compounds in biotechnological processes, we aimed at increasing the range of characterized phenylalanine ammonia-lyases (PAL) and tyrosine ammonia-lyases (TAL), and in particular at identifying the optimal TAL for use in microbial cell factories. Several enzymes have previously been characterized both in vivo and in vitro, and the data are available from both patent and scientific literature ( $3,14-32$ ). However, the conditions of kinetic analysis are often quite different and may not represent the extent to which the enzymes perform in vivo. Therefore, we decided to evaluate a range of previously identified enzymes, and expanded our analysis to novel enzymes selected from available sequence data. Identification of TAL enzymes based solely on sequence information is not straightforward, since the sequences do not provide unambig-
uous prediction of substrate specificity. This is illustrated by the fact that hypothetical genes whose predicted products have homology to aromatic amino acid ammonia-lyases are often simply annotated as being histidine ammonia-lyases without experimental evidence of their substrate specificity.

Phenylalanine and tyrosine ammonia-lyases are members of a family that either have a lyase activity that forms $\alpha, \beta$-unsaturated acids from amino acids by elimination of ammonia or a mutase activity that forms $\beta$-amino acids. The aromatic amino acid am-monia-lyases (here XAL) in this family are classified by their substrate specificity as being histidine ammonia-lyases (HAL; histidase; EC 4.3.1.3), tyrosine ammonia-lyases (TAL; EC 4.3.1.23), phenylalanine ammonia-lyases (PAL; EC 4.3.1.24), or phenylalanine/tyrosine ammonia-lyases (PAL/TAL; EC 4.3.1.25) (Fig. 1).

[^1]



TAM




urocanic acid

FIG 1 Aromatic amino acid ammonia-lyases and aminomutases. Phenylalanine (Phe) may be converted into $\beta$-phenylalanine ( $\beta$-Phe) by phenylalanine aminomutase (PAM) or deaminated to cinnamic acid (CA). Similarly, tyrosine (Tyr) may be converted to $\beta$-tyrosine by tyrosine aminomutase (TAM) or $p$-coumaric acid (pHCA) by tyrosine ammonia-lyase (TAL). pHCA may also be formed from CA by cinnamate 4-hydroxylase (C4H). Histidine (His) may be deaminated to urocanic acid by histidine ammonia-lyase (HAL).

Enzymes categorized as acting on either of the structurally similar amino acids tyrosine or phenylalanine often have some activity for the other substrate as well $(16,32)$ or for similar compounds, such as 3,4-dihydroxy-L-phenylalanine (L-dopa) (22). Homologous and structurally similar enzymes are tyrosine 2,3 -aminomutases (TAM; EC 5.4.3.6) and phenylalanine aminomutase (PAM; EC 5.4.3.11) (33-35). The enzyme activities are difficult to predict based on primary sequence, as all enzymes contain a prosthetic group, 3,5-dihydro-5-methylidene-4H-imidazol-4-one (MIO), formed by the cyclization of the sequential amino acids alanine, serine, and glycine (22, 36-38).

Here we compared previously described and commonly used en-zymes-RsTAL from Rhodobacter sphaeroides (1,39), RcTAL from Rhodobacter capsulatus (9, 18), BagA from Streptomyces sp. (40), SeSam8 from Saccharothrix espanaensis (21), RmXAL from Rhodotorula mucilaginosa (41-43), and other fungal enzymes-for enzymatic activity in vivo together with 11 uncharacterized enzymes.

## MATERIALS AND METHODS

Bioinformatic methods. The nonredundant protein database in GenBank was searched using BLASTP 2.2.28+ using four known enzymes having

TAL activity: RsTAL from Rhodobacter sphaeroides, RmXAL from Rhodotorula mucilaginosa, S_bagA from Streptomyces sp., and SeSam8 from Saccharothrix espanaensis. To limit the list of BLAST hits to the most likely correctly annotated sequences, duplicates were removed and the resulting hits were limited to those having homology over the central part of the sequence covering the catalytically important residues, ranging from the alanine-serine-glycine triplet forming the MIO prosthetic group (A149 in RsTAL) (see Fig. S1 in the supplemental material) to the tyrosine stabilizing the MIO group (Y303 in RsTAL) and to a minimum length of 300 amino acids. We grouped the hits into clusters using H-CD-HIT in the CD-HIT suite (44), with three hierarchical rounds of clustering: $90 \%$, $60 \%$, and $40 \%$ amino acid identity, yielding 107 clusters. Assignment to phyla was done by reconstructing taxonomy trees from the NCBI taxonomy database. Synteny was analyzed using GCView (45) with input sequences being PSI-BLAST results for RsTAL (GI 126464011) from Rhodobacter sphaeroides, PYP (GI 409511), 4-coumarate-CoA ligase (4CL; GI 121590003) from Halorhodospira halophila, and the urocanate hydratase (hutU) protein of Bacillus subtilis (GI 603769) retrieved from the nonredundant database using standard parameters. Phylogenetic trees were constructed in MEGA6 (46) using the maximum-likelihood method based on a JTT matrix-based model (47) on Clustal Omega (48)-aligned sequences and an initial tree constructed by the neighbor-joining method.

Strains and media. (i) Media. Escherichia coli strains were routinely grown at $37^{\circ} \mathrm{C}$ on plates or in liquid culture in either lysogeny broth (LB) or $2 \times$ YT medium ( 16 g liter ${ }^{-1}$ Bacto tryptone, 10 g liter ${ }^{-1}$ Bacto yeast extract, 5 g liter ${ }^{-1} \mathrm{NaCl}$; adjusted to pH 7.0 with NaOH ) with antibiotics when appropriate. Liquid cultures were incubated with aeration in an orbital shaker ( 250 rpm ). For analysis of production rates, E. coli strains were grown in M9 minimal medium with $0.2 \%$ glucose. Growth was measured by following the optical density at $600 \mathrm{~nm}\left(\mathrm{OD}_{600}\right)$.

Saccharomyces cerevisiae strains were routinely grown in synthetic dropout medium lacking histidine for selection of DNA integration or lacking uracil for selection of episomal plasmids. For production analysis, strains were grown either in the Feed-In-Time (FIT) synthetic fed-batch medium M-Sc.syn-1000 from m2p-labs GmbH (Germany) or in a defined minimal medium (49) based on Delft medium (50) supplemented with 76 mg liter $^{-1}$ uracil and 380 mg liter $^{-1}$ leucine or 25 mg liter ${ }^{-1}$ histidine and 75 mg liter ${ }^{-1}$ leucine at $30^{\circ} \mathrm{C}$ and 250 rpm .

For molecular biology procedures, Lactococcus lactis strains were cultivated as batch cultures (flasks) without aeration in M17 medium (Difco, USA) supplemented with $0.5 \%$ glucose ( $\mathrm{wt} / \mathrm{vol}$ ) at $30^{\circ} \mathrm{C}$. To assess pHCA production, strains were grown as static cultures in chemically defined medium (CDM) (51) containing $1 \%$ glucose ( $\mathrm{wt} / \mathrm{vol}$ ) without pH control (initial pH 6.5 or 7.0 ) and supplemented with 1.7 or 3.7 mM L-tyrosine. Plasmid selection was achieved by addition of $5 \mu \mathrm{~g} \mathrm{ml}^{-1}$ chloramphenicol to the growth medium. Growth was monitored by measuring $\mathrm{OD}_{600}$.
(ii) DNA and molecular biology tools. DNA oligonucleotide primers (listed in Table 1) were purchased from Integrated DNA Technologies. Synthetic genes, codon optimized for expression in E. coli, L. lactis or S. cerevisiae, were ordered from GeneArt (Life Technologies). General molecular techniques were performed essentially as described elsewhere (52). Restriction enzymes, T4 DNA ligase, DNA polymerases, and Gibson Assembly master mix were obtained from New England BioLabs or Thermo Scientific and used according to the supplier's instructions. DNA purification was done using Macherey-Nagel kits, except for plasmid purification from L. lactis. Lactococcal plasmid DNA isolation was carried out using the QIAprep spin miniprep kit (Qiagen, United Kingdom) for small-scale purifications with the following modifications: the bacterial pellet was resuspended in Birnboim A solution (53), containing 5 mg $\mathrm{ml}^{-1}$ lysozyme and $30 \mu \mathrm{~g} \mathrm{ml}^{-1} \mathrm{RNase} \mathrm{I}$, and incubated at $55^{\circ} \mathrm{C}$ for 10 to 15 min before addition of buffer P2.
(iii) Cloning for expression in Escherichia coli. Genes optimized for E. coli (see Table S1 in the supplemental material) were amplified using the oligonucleotides shown in Table 1 and cloned into pCDFDuet-1 (Novagen) as follows. The plasmid was digested with NdeI and BglII and gel
TABLE 1 DNA oligonucleotides used in this study

| Oligonucleotide | Gene | Direction | Sequence | Restriction site ${ }^{a}$ | Use (reference) ${ }^{\text {b }}$ |
| :---: | :---: | :---: | :---: | :---: | :---: |
| CBJP483 | RstaL | Forward | CATCTTAGTATATTAGTTAAGTATAAGAAGGAGATATACATATGCTGGCAA TGAGCCCT |  | Ec |
| CBJP484 | RstaL | Reverse | TGGCCGGCCGATATCCAATTGATTAAACCGGACTCTGTTGC |  | Ec |
| CBJP485 | S_BagA | Forward | CATCTTAGTATATTAGTTAAGTATAAGAAGGAGATATACATATGAAAATTG <br> ATGGTCGTGGTCTG |  | Ec |
| CBJP486 | S_BagA | Reverse | TGGCCGGCCGATATCCAATTGATTACAGATTACCGCCTGCAC |  | Ec |
| CBJP487 | RmXAL | Forward | CATCTTAGTATATTAGTTAAGTATAAGAAGGAGATATACATATGGCACCG AGCGTTGATAGC |  | Ec |
| CBJP488 | RmXAL | Reverse | TGGCCGGCCGATATCCAATTGATTAGGCCATCATTTTAACCAGAACC |  | Ec |
| CBJP535 | SeSam8 | Forward | CATCTTAGTATATTAGTTAAGTATAAGAAGGAGATATACATATGACCCAG GTTGTTGAACG |  | Ec |
| CBJP536 | SeSam8 | Reverse | TGGCCGGCCGATATCCAATTGATTAGCCAAAATCTTTACCATCTGC |  | Ec |
| CBJP537 | BlPAL | Forward | CATCTTAGTATATTAGTTAAGTATAAGAAGGAGATATACATATGAGCCA GGTTGCACTG |  | Ec |
| CBJP538 | BIPAL | Reverse | TGGCCGGCCGATATCCAATTGATTAATCATTCACATTCTGATCATGAATTTC |  | Ec |
| CBJP539 | R_XAL | Forward | CATCTTAGTATATTAGTTAAGTATAAGAAGGAGATATACATATGCGTAGCG <br> AACAGCTGA |  | Ec |
| CBJP540 | R_XAL | Reverse | TGGCCGGCCGATATCCAATTGATTAGGCCAGCAGTTCAATCA |  | Ec |
| CBJP541 | PpPAL | Forward | CATCTTAGTATATTAGTTAAGTATAAGAAGGAGATATACATATGCACGA TGATAACACCAGC |  | Ec |
| CBJP542 | PpPAL | Reverse | TGGCCGGCCGATATCCAATTGATTAACAGCTTGCGCGTGCGCTAAAC |  | Ec |
| CBJP543 | LbTAL | Forward | CATCTTAGTATATTAGTTAAGTATAAGAAGGAGATATACATATGCCT <br> CGTTTTTGTCCGA |  | Ec |
| CBJP544 | LbTAL | Reverse | TGGCCGGCCGATATCCAATTGATTAATCGTTCGGGGTCATAACC |  | Ec |
| CBJP545 | IITAL | Forward | CATCTTAGTATATTAGTTAAGTATAAGAAGGAGATATACATATGACCACCT CCATTATTGC |  | Ec |
| CBJP546 | IITAL | Reverse | TGGCCGGCCGATATCCAATTGATTATGCCGGTTCTTGATACAG |  | Ec |
| CBJP547 | DdPAL | Forward | CATCTTAGTATATTAGTTAAGTATAAGAAGGAGATATACATATGATCGAAA CCAACCACA |  | Ec |
| CBJP548 | DdPAL | Reverse | TGGCCGGCCGATATCCAATTGATTACAGGTTCAGGTTAATAAAGTCCA |  | Ec |
| CBJP549 | SrXAL | Forward | CATCTTAGTATATTAGTTAAGTATAAGAAGGAGATATACATATGAGCAC CCCGAGCGCA |  | Ec |
| CBJP550 | SrXAL | Reverse | TGGCCGGCCGATATCCAATTGATTATGCGGTCGGAGGGGTAAC |  | Ec |
| CBJP551 | L_XAL | Forward | CATCTTAGTATATTAGTTAAGTATAAGAAGGAGATATACATATGACCC <br> TGACCCCGAC |  | Ec |
| CBJP552 | L_XAL | Reverse | TGGCCGGCCGATATCCAATTGATTAGTTAAAGCTGCTAATGGTGCT |  | Ec |
| CBJP553 | HaTAL1 | Forward | CATCTTAGTATATTAGTTAAGTATAAGAAGGAGATATACATATGAGCAC CACCCTGATTC |  | Ec |
| CBJP554 | HaTAL1 | Reverse | TGGCCGGCCGATATCCAATTGATTAGCGAAACAGAATAATACTACGCA |  | Ec |
| CBJP555 | FjTAL | Forward | CATCTTAGTATATTAGTTAAGTATAAGAAGGAGATATACATATGAACACC <br> ATCAACGAATATCTG |  | Ec |
| CBJP556 | FjTAL | Reverse | TGGCCGGCCGATATCCAATTGATTAATTGTTAATCAGGTGGTCTTTT ACTTTCTG |  | Ec |
| CBJP559 | $\mathrm{His}_{6}$-tag | Forward | TATGGCCCACCATCATCACCACCATGAGAACCTCTACTTCCA |  | His |
| CBJP560 | $\mathrm{His}_{6}$-tag | Reverse | GATCTGGAAGTAGAGGTTCTCATGGTGGTGATGATGGTGGGCCA |  | His |
| CBJP561 | RsTAL | Forward | CACCACCATGAGAACCTCTACTTCCAGATGCTGGCAATGAGCCCT |  | His |
| CBJP562 | S_BagA | Forward | CACCACCATGAGAACCTCTACTTCCAGATGAAAATTGATGGTCGTGGTCTG |  | His |
| CBJP563 | RmXAL | Forward | CACCACCATGAGAACCTCTACTTCCAGATGGCACCGAGCGTTGATAGC |  | His |
| CBJP564 | SeSam8 | Forward | CACCACCATGAGAACCTCTACTTCCAGATGACCCAGGTTGTTGAACG |  | His |
| CBJP565 | BlPAL | Forward | CACCACCATGAGAACCTCTACTTCCAGATGAGCCAGGTTGCACTG |  | His |
| CBJP566 | R_XAL | Forward | CACCACCATGAGAACCTCTACTTCCAGATGCGTAGCGAACAGCTGA |  | His |
| CBJP567 | PpPAL | Forward | CACCACCATGAGAACCTCTACTTCCAGATGCACGATGATAACACCAGC |  | His |
| CBJP568 | LbTAL | Forward | CACCACCATGAGAACCTCTACTTCCAGATGCCTCGTTTTTGTCCGA |  | His |
| CBJP569 | IITAL | Forward | CACCACCATGAGAACCTCTACTTCCAGATGACCACCTCCATTATTGC |  | His |
| CBJP570 | DdPAL | Forward | CACCACCATGAGAACCTCTACTTCCAGATGATCGAAACCAACCACA |  | His |
| CBJP571 | SrXAL | Forward | CACCACCATGAGAACCTCTACTTCCAGATGAGCACCCCGAGCGCA |  | His |
| CBJP572 | L_XAL | Forward | CACCACCATGAGAACCTCTACTTCCAGATGACCCTGACCCCGAC |  | His |
| CBJP573 | HaTAL1 | Forward | CACCACCATGAGAACCTCTACTTCCAGATGAGCACCACCCTGATTC |  | His |
| CBJP574 | FjTAL | Forward | CACCACCATGAGAACCTCTACTTCCAGATGAACACCATCAACGAATATCTG |  | His |


| CBJP575 | pCBJ229 | Forward | TAATCAATTGGATATCGGCCGGCCA | Plasmid amplification |
| :---: | :---: | :---: | :---: | :---: |
| CBJP576 | pCBJ229 | Reverse | CATCTGGAAGTAGAGGTTCTCATGGTGGTG | Plasmid amplification |
| CBJP637 | SeSam8 | Forward | ATCTGTCAUAAAACAATGACCCAGGTTGTTGAACG | Sc |
| CBJP638 | SeSam8 | Reverse | CACGCGAUTCAGCCAAAATCTTTACCATCTGC | Sc |
| CBJP645 | HaTAL1 | Forward | ATCTGTCAUAAAACAATGAGCACCACCCTGATTC | Sc |
| CBJP646 | HaTAL1 | Reverse | CACGCGAUTCAGCGAAACAGAATAATACTACGCA | Sc |
| CBJP647 | FjTAL | Forward | ATCTGTCAUAAAACAATGAACACCATCAACGAATATCTG | Sc |
| CBJP648 | FjTAL | Reverse | CACGCGAUTCAATTGTTAATCAGGTGGTCTTTTACTTTCTG | Sc |
| CBJP649 | $\mathrm{HaTAL1}_{\text {Sc }}$ | Forward | ATCTGTCAUAAAACAATGTCCACCACCTTGATTTTGA | Sc |
| CBJP650 | $\mathrm{HaTAL1}_{\text {Sc }}$ | Reverse | CACGCGAUTCATCTGAACAAGATGATGGATCTCAA | Sc |
| CBJP651 | $\mathrm{FjTAL}_{\text {sc }}$ | Forward | ATCTGTCAUAAAACAATGAACACCATCAACGAATACTTG | Sc |
| CBJP652 | $\mathrm{FjTAL}_{\text {sc }}$ | Reverse | CACGCGAUTCAGTTGTTAATCAAGTGATCCTTAACTTTTTGG | Sc |
| CBJP741 | RtXAL, RtXALmut1-EP18Km-6, RtXALmut2-RM120-1 | Forward | CATCTTAGTATATTAGTTAAGTATAAGAAGGAGATATACATATGGCACCGA GCCTGGATAG | Ec |
| CBJP742 | RtXAL, RtXALmut1-EP18Km-6,RtXALmut2-RM120-1 | Reverse | TGGCCGGCCGATATCCAATTGATTAGGCCAGCATTTTCAG | Ec |
| CBJP743 | TcXAL | Forward | CATCTTAGTATATTAGTTAAGTATAAGAAGGAGATATACATATGTTCATCG AAACCAATGTTG | Ec |
| CBJP744 | TcXAL | Reverse | TGGCCGGCCGATATCCAATTGATTAAAACATTTTACCAACTGCAC | Ec |
| CBJP745 | RcTAL-var1 | Forward | CATCTTAGTATATTAGTTAAGTATAAGAAGGAGATATACATATGCTGG <br> ATGC <br> AACCATTGG | Ec |
| CBJP746 | RcTAL-var1 | Reverse | TGGCCGGCCGATATCCAATTGATTATGCCGGAGGATCCGCT | Ec |
| CBJP747 | RcTAL-var2 | Forward | CATCTTAGTATATTAGTTAAGTATAAGAAGGAGATATACATATGACCC <br> TGCA <br> GAGCCAGAC | Ec |
| CBJP748 | RcTAL-var2 upstream part | Reverse | CGATCAATGGTTTCCGGACT | Ec |
| CBJP749 | RcTAL-var2 middle part | Forward | GTGCAAGTCCGGAAACCATTGATCGTATTGTTGCCGTTCTG | Ec |
| CBJP750 | RcTAL-var2 middle part | Reverse | CGATCACGCAGATCACGTGCGGTCAG | Ec |
| CBJP752 | HaTal2 | Forward | CATCTTAGTATATTAGTTAAGTATAAGAAGGAGATATACATATGCGTC <br> ATCAGGTTACC | Ec |
| CBJP753 | HaTal2 | Reverse | TGGCCGGCCGATATCCAATTGATTAGGCAAACAGAATAATACTACG | Ec |
| CBJP754 | RtXAL, RtXALmut1-EP18Km-6,RtXALmut2-RM120-1 | Forward | ATCTGTCAUAAAACAATGGCACCGAGCCTGGATAG | Sc |
| CBJP755 | RtXAL, RtXALmut1-EP18Km-6,RtXALmut2-RM120-1 | Reverse | CACGCGAUTCAGGCCAGCATTTTCAGCA | Sc |
| CBJP758 | RcTAL-var 1 | Forward | ATCTGTCAUAAAACAATGCTGGATGCAACCATTGG | Sc |
| CBJP759 | RcTAL-var1 | Reverse | CACGCGAUTCATGCCGGAGGATCCGCT | Sc |
| CBJP760 | RcTAL-var2 | Forward | ATCTGTCAUAAAACAATGACCCTGCAGAGCCAGAC | Sc |
| CBJP761 | RcTAL-var2 | Reverse | CACGCGAUTCATGCAGGCGGATCTGCG | Sc |
| CBJP762 | HaTAL2 | Forward | ATCTGTCAUAAAACAATGCGTCATCAGGTTACC | Sc |
| CBJP763 | HaTAL2 | Reverse | CACGCGAUTCAGGCAAACAGAATAATACTACGCAG | Sc |
| CBJP812 | PcXAL | Forward | CATCTTAGTATATTAGTTAAGTATAAGAAGGAGATATACATATGCCGAG CCGTATTGATTATTACAC | Ec |
| CBJP813 | PcXAL | Reverse | TGGCCGGCCGATATCCAATTGATTAGGCTTTAATGCTTTTCACCAGCA | Ec |
| CBJP815 | PcXAL | Forward | ATCTGTCAUAAAACAATGCCGAGCCGTATTGATTATTACAC | Sc |
| CBJP816 | PcXAL | Reverse | CACGCGAUTCAGGCTTTAATGCTTTTCACCAGCA | Sc |
| CBJP828 | RtXALmut1-EP18Km-6, RtXALmut2-RM120-1 | Reverse | TTGCACGCAGATCAATGGCCT | Ec |
| CBJP829 | RtXALmut1-EP18Km-6, RtXALmut2-RM120-1 | Forward | CTGCAGGCCATTGATCTGCGTGCAACCGAATTTGAGTTCAAAAAACAGTTTGG | Ec |
| CBJP830 | RtXALmut2-RM120-1 | Reverse | TCCAGCAGTGCTTTCTGCAGGCTAATTGCACCTTCGGTACGGGTATCTG | Ec |
| CBJP831 | RtXALmut2-RM120-1 | Forward | TTAGCCTGCAGAAAGCACTGCTGGAACATCTGCTGTGTGGTGTTCTGCC GAGCAGCT | Ec |
| TAL1_fw | RsTAL | Forward | AGTGCAGGUAAAACAATGAGCCCTCCGAAACCGGCAGTTGAACTGG | Sc |
| TAL1_rv | RsTAL | Reverse | CGTGCGAUTTAAACCGGACTCTGTTG | Sc |
| TAL2_fw | S_BagA | Forward | AGTGCAGGUAAAACAATGAAAATTGATGGTCGTGGTCTGACCAT TAGCCAGACCG | Sc |
| TAL2_rv | S_BagA | Reverse | CGTGCGAUTTACAGATTACCGCCTGC | Sc |
| TAL3_fw | RmXAL | Forward | AGTGCAGGUAAAACAATGGCACCGAGCGTTGATAGC | Sc |


TABLE 1 (Continued)

| Oligonucleotide | Gene | Direction | Sequence | Restriction site ${ }^{a}$ | Use (reference) ${ }^{\text {b }}$ |
| :---: | :---: | :---: | :---: | :---: | :---: |
| TAL3_rv | RmXAL | Reverse | CGTGCGAUTTAGGCCATCATTTTAAC |  | Sc |
| TAL4_fw | SeSam8 | Forward | AGTGCAGGUAAAACAATGACCCAGGTTGTTGAACGTCAGG |  | Sc |
| TAL4_rv | SeSam8 | Reverse | CGTGCGAUTTAGCCAAAATCTTTACC |  | Sc |
| TAL5_fw | BlPAL | Forward | AGTGCAGGUAAAACAATGAGCCAGGTTGCACTGTTTG |  | Sc |
| TAL5_rv | BIPAL | Reverse | CGTGCGAUTTAATCATTCACATTCTG |  | Sc |
| TAL6_fw | R_XAL | Forward | AGTGCAGGUAAAACAATGCGTAGCGAACAGCTGACC |  | Sc |
| TAL6_rv | R_XAL | Reverse | CGTGCGAUTTAGGCCAGCAGTTCAAT |  | Sc |
| TAL7_fw | PpPAL | Forward | AGTGCAGGUAAAACAATGCACGATGATAACACCAGCCCG |  | Sc |
| TAL7_rv | PpPAL | Reverse | CGTGCGAUTTAACAGCTTGCGCGTGC |  | Sc |
| TAL8_fw | LbTAL | Forward | AGTGCAGGUAAAACAATGCCTCGTTTTTGTCCGAGCATGTATCTGC |  | Sc |
| TAL8_rv | LbTAL | Reverse | CGTGCGAUTTAATCGTTCGGGGTCAT |  | Sc |
| TAL9_fw | IITAL | Forward | AGTGCAGGUAAAACAATGACCACCTCCATTATTGCATTTGG |  | Sc |
| TAL9_rv | IITAL | Reverse | CGTGCGAUTTATGCCGGTTCTTGATA |  | Sc |
| TAL10_fw | DdPAL | Forward | AGTGCAGGUAAAACAATGATCGAAACCAACCACAAA |  | Sc |
| TAL10_rv | DdPAL | Reverse | CGTGCGAUTTACAGGTTCAGGTTAAT |  | Sc |
| TAL11_fw | SrXAL | Forward | AGTGCAGGUAAAACAATGAGCACCCCGAGCGCA |  | Sc |
| TAL11_rv | SrXAL | Reverse | CGTGCGAUTTATGCGGTCGGAGGGGT |  | Sc |
| TAL12_fw | L_XAL | Forward | AGTGCAGGUAAAACAATGACCCTGACCCCGACCG |  | Sc |
| TAL12_rv | L_XAL | Reverse | CGTGCGAUTTAGTTAAAGCTGCTAAT |  | Sc |
| TAL13_fw | HaTAL1 | Forward | AGTGCAGGUAAAACAATGAGCACCACCCTGATTCTG |  | Sc |
| TAL13_rv | HaTAL1 | Reverse | CGTGCGAUTTAGCGAAACAGAATAAT |  | Sc |
| TAL14_fw | FjTAL | Forward | AGTGCAGGUAAAACAATGAACACCATCAACGAATATCTGAGC |  | Sc |
| TAL14_rv | FjTAL | Reverse | CGTGCGAUTTAATTGTTAATCAGGTG |  | Sc |
| PPGK1_fw | PPGK1 promoter | Forward | CGTGCGAUGGAAGTACCTTCAAAGA |  | Sc (49) |
| PPGK1_rv | PPGK1 promoter | Reverse | ATGACAGAUTTGTTTTATATTTGTTG |  | Sc (49) |
| PTEF1_fw | PTEF1 | Forward | ACCTGCACUTTGTAATTAAAACTTAG |  | Sc (49) |
| PTEF1_rv | PTEF1 | Reverse | CACGCGAUGCACACACCATAGCTTC |  | Sc (49) |
| pIntFwdU | S. cerevisiae chromosomal DNA | Forward | ACCCAAUTCGCCCTATAGTGAGTCG |  | Sc; integration verification |
| XI-5 down_out | S. cerevisiae chromosomal DNA | Reverse | CCCAAAAGCAATCCAGGAAAAACC |  | Sc; integration verification |
| Pnis_1 - |  |  | GGTGAGTGCCTCCTTATAATTTATTTTG |  | Ll |
| Pnis_2 |  |  | AAGCTTTCTTTGAACCAAAATTAGAAAACC |  | Ll |
| TAL_Rs_fwl | RsTALLl | Forward | CAAAATAAATTATAAGGAGGCACTCACCATGCTTGCTATGTCACC ACCAAAACC |  | Ll |
| TAL_Rs_rv2 | RsTALLl | Reverse | GGTTTTCTAATTTTGGTTCAAAGAAAGCTTTTAAACTGGTGATTGTT GTAATAAATG |  | Ll |
| TAL_Rr_fwl | RmXALLl | Forward | CAAAATAAATTATAAGGAGGCACTCACCATGGCTCCATCAGTTGAT TCAATTGC |  | Ll |
| TAL_Rr_rv2 | RmXALLl | Reverse | GGTTTTCTAATTTTGGTTCAAAGAAAGCTTTTAAGCCATCATTTTAA <br> CTAAAACTGG |  | Ll |
| TAL_4_fw1 | SeSam8 | Forward | CATGTCATGACCCAGGTTGTTGAACG | BspHI | Ll |
| TAL_4_rv1 | SeSam8 | Reverse | GCTCTAGATTAGCCAAAATCTTTACCATC | XbaI | Ll |
| TAL_6_fw1 | R_XAL | Forward | GCGGTCTCCCATGCGTAGCGAACAGCTGAC | BsaI | Ll |
| TAL_6_rv1 | R_XAL | Reverse | GCTCTAGATTAGGCCAGCAGTTCAATCAG | XbaI | Ll |
| TAL_13_fw1 | HaTAL1 | Forward | GCGGTCTCCCATGAGCACCACCCTGATTCTG | BsaI | Ll |
| TAL_13_rv1 | HaTAL1 | Reverse | GCTCTAGATTAGCGAAACAGAATAATACTACG | XbaI | Ll |
| TAL_14_fw1 | FjTAL | Forward | CATGTCATGAACACCATCAACGAATATC | BspHI | Ll |
| TAL_14_rvl | FjTAL | Reverse | GCTCTAGATTAATTGTTAATCAGGTGGTC | XbaI | Ll |

purified. The genes were inserted by isothermal assembly using Gibson Assembly master mix (New England BioLabs), and transformed into chemically competent DH5 $\alpha$ (laboratory strain) or NEB5 $\alpha$ (New England BioLabs), selecting for resistance to $50 \mu \mathrm{~g} \mathrm{ml}^{-1}$ spectinomycin in LB medium. A few genes were constructed by assembly of multiple fragments: the gene encoding RsTAL-var2 was assembled from two fragments PCR amplified from the gene encoding RcTAL-varl and the synthetic double-stranded DNA fragment CBJBB6 (see Table S1 in the supplemental material). The genes encoding RtXAL-EP18Km-6 and RtPALmut2-RM120-1 were constructed by assembly of two and three, respectively, PCR products amplified from the gene encoding RtXall. Plasmids carrying genes encoding His-tagged versions of the proteins were made as follows. Partly complementary oligonucleotides CBJP559 and CBJP560 were ligated into pCDFDuet-1 digested with NdeI and BglII, forming a new histidine tag. The resulting plasmid pCBJ229 was PCR amplified with CBJP575 and CBJP576, and this linear DNA fragment was combined with individual genes amplified with the oligonucleotides shown in Table 1 as described previously. Clones were verified by sequencing and electroporated into the E. coli expression strain BL21(DE3)/pLysS (Invitrogen/Life Technologies), selecting for pLysS with $34 \mu \mathrm{~g} \mathrm{ml}^{-1}$ chloramphenicol and for expression plasmids with $50 \mu \mathrm{~g} \mathrm{ml}^{-1}$ spectinomycin. A control strain carrying pCDFDuet-1 was also made. The resulting plasmids and strains are shown in Tables 2 and 3, respectively.
(iv) Cloning for expression in Saccharomyces cerevisiae. All genetic constructions were made in E. coli DH5 $\alpha$ grown in LB containing $100 \mu \mathrm{~g}$ $\mathrm{ml}^{-1}$ ampicillin for plasmid maintenance. Fourteen of the genes cloned for expression in E. coli and the TEF1 promoter were amplified using the uracil-containing primers listed in Table 1. The promoter fragment and each of the genes were combined into the integrative EasyClone vector, pCfB391 (49) following the uracil-specific excision reagent (USER) cloning method (54). The resulting plasmids are listed in Table 2. Plasmids were digested by NotI and purified by using NucleoSpin gel and PCR cleanup kits (Macherey-Nagel) following manual instructions. DNA linearized plasmids ( 300 to 700 ng ) were transformed into S. cerevisiae CEN.PK102-5B (MATa ura3-52 his3D1 leu2-3/112 MAL2-8c SUC2) by the lithium acetate transformation protocol (55). Yeast transformants were selected on synthetic complete (SC) dropout medium without histidine, resulting in strains STC0 to STC14. Correct insertions of TAL genes were verified by yeast colony PCR by primers pIntFwdU and XI-5 down_out. Alternatively, for plasmid-borne expression, 12 genes, including those encoding HaTAL1 and FjTAL optimized for S. cerevisiae (see Table S1 in the supplemental material), were amplified using the oligonucleotide also listed in Table 1 and inserted by uracil excision into the vector pCfB132 together with the PPGK1 promoter amplified by primers PPGK1_fw and PPGK1_rv (49). The finished plasmids were transformed into S. cerevisiae CEN.PK102-5B selecting for growth on synthetic dropout medium plates lacking uracil. Control strains with either the integration of pCfB 391 or carrying pCfB 1322 were also made. The resulting plasmids and strains are shown in Tables 2 and 3, respectively.
(v) Cloning for expression in Lactococcus lactis. The synthetic $\operatorname{RsTAL}_{L l}$ and $\mathrm{RmXAL}_{L l}$ genes (see Table S1 in the supplemental material) were cloned into the nisin-inducible expression vector $\mathrm{pNZ8048}$ (56) as follows. RsTAL $L_{L l}$ and $\mathrm{RmXAL}_{L l}$ genes and the vector were PCR amplified using the primers listed in Table 1 and assembled in a single-tube isothermal reaction using the Gibson Assembly master mix (New England BioLabs). Reaction products were ethanol precipitated and suspended in double-distilled water before transformation into $L$. lactis by electroporation as described by Holo and Nes (57). The synthetic genes encoding SeSam8, R_XAL, HaTAL1, and FjTAL were amplified by PCR using the primer pairs listed in Table 1, digested with specific restriction enzymes, and cloned between the NcoI and XbaI restriction sites of pNZ8048. The plasmids were obtained and maintained in NZ9000 (56), and the gene sequences of the different constructs were verified by sequencing. To assess pHCA production, TAL expression vectors were transformed into $\mathrm{NZ9000} \mathrm{\Delta ldh} \mathrm{\Delta ldhB}(86)$. A control strain was also constructed by trans-
formation of NZ9000 $\Delta l d h \Delta l d h B$ with empty expression vector pNZ 8048. Plasmids and strains are listed in Tables 2 and 3.

Analysis of production. (i) Heterologous TAL and PAL expression and analysis of pHCA and CA production in E. coli. Unless otherwise stated, pHCA production was carried out as follows. E. coli BL21(DE3) pLysS strains harboring recombinant plasmids were precultured in 3 ml of $2 \times \mathrm{YT}$ with appropriate antibiotics and incubated at $37^{\circ} \mathrm{C}$ and 250 rpm overnight. The following day, each preculture was transferred into 2 ml of M9 minimal medium with appropriate antibiotics to a final $\mathrm{OD}_{600}$ of 0.05 and cultured at $37^{\circ} \mathrm{C}$ and 300 rpm in 24 -well deep-well plates (Enzyscreen B.V., Netherlands) until the $\mathrm{OD}_{600}$ reached $\sim 0.6$. Then, isopropyl $\beta-\mathrm{D}-1$ thiogalactopyranoside (IPTG) was added at a final concentration of 1 mM , and the cells were grown for 3 h at $30^{\circ} \mathrm{C}$, followed by addition of 2 mM substrate (tyrosine, phenylalanine, or histidine). Finally, the culture was incubated at $30^{\circ} \mathrm{C}$ for 24 h . In the case of the time course experiments with E. coli, samples were collected every hour for 14 h and after 24 h from substrate addition. Collected samples were harvested by centrifugation at $16,200 \times g$ for 10 min , and the supernatant was filtered through $0.2-\mu \mathrm{m}$ filters and analyzed by using high-performance liquid chromatography (HPLC) as described below. Experiments were carried out at least in triplicate.
(ii) Heterologous TAL and PAL expression and analysis of pHCA and CA production in S. cerevisiae. For screening the 14 different chromosomally integrated TAL and PAL genes, yeast cultures were grown in 0.5 ml SC medium lacking histidine at $30^{\circ} \mathrm{C}$ and 250 rpm for 24 h in a 96-well deep-well plate, and then $50 \mu \mathrm{l}$ of this preculture was inoculated into 0.5 ml Delft medium with $2 \%$ glucose or into FIT medium, both of which were supplemented with 76 mg liter $^{-1}$ uracil, 380 mg liter ${ }^{-1}$ leucine, and 10 mM tyrosine or phenylalanine. Yeast cultures were cultivated in 96 -well deep-well plates at $30^{\circ} \mathrm{C}$ and 250 rpm for 72 h . Samples for HPLC were taken at the endpoint of cultivation, and the $\mathrm{OD}_{600}$ of the yeast cultures was measured in a microtiter plate reader after cultures had been diluted 2 to 20 times. A similar procedure was used for the analysis of strains carrying genes on episomal plasmids, except that selection was done in medium without uracil.
(iii) Heterologous TAL and PAL expression and analysis of pHCA and CA production in L. lactis. For heterologous expression of cloned tyrosine ammonia-lyases, L. lactis strains were grown in CDM, and nisin ( $1.5 \mu \mathrm{~g}$ liter ${ }^{-1}$ ) was added at an $\mathrm{OD}_{600}$ of 0.3 to 0.4 . Samples $(1 \mathrm{ml})$ of cultures were collected at different points during growth and centrifuged $\left(16,100 \times \mathrm{g}, 10 \mathrm{~min}, 4^{\circ} \mathrm{C}\right)$, and the supernatants were stored at $-20^{\circ} \mathrm{C}$ until analysis by HPLC.

Analytical methods. The production of pHCA, cinnamic acid, and urocanic acid was measured by HPLC on a Thermo HPLC setup and quantified using standards (Sigma-Aldrich). Samples were analyzed using a gradient method with two solvents: 10 mM ammonium formate ( pH 3.0) (A) and acetonitrile (B) at $1.5 \mathrm{ml} \mathrm{min}^{-1}$, starting at $5 \%$ B. From 0.5 min after injection to 7 min , the fraction of $B$ increased linearly from $5 \%$ to $60 \%$, and between 9.5 and 9.6 min , the fraction of B decreased back to $5 \%$, remaining there until 12 min . pHCA and CA were quantified by measuring absorbance at 333 nm and 277 nm , respectively. Titers were calculated by correlating $\mathrm{OD}_{600}$ to grams (dry weight) of cells (CDW) liter ${ }^{-1}$, using previously measured correlation or literature values $(58,59)$.

Protein purification and kinetic analysis. Strains expressing Histagged versions of the enzymes were grown in LB medium overnight at $37^{\circ} \mathrm{C}$, diluted into fresh LB with 1 mM IPTG, and propagated overnight (approximately 18 h ) at $30^{\circ} \mathrm{C}$. Cells were harvested by centrifugation at $9,800 \times g$ for 8 min and disrupted using an EmulsiFlex-C5 homogenizer (Avestin) into a buffer ( 50 mM Tris- $\mathrm{HCl}, 10 \mathrm{mM}$ imidazole, 500 mM $\mathrm{NaCl}, 10 \%$ glycerol [pH 7.5]). The supernatant was clarified by centrifugation at $10,000 \times g$ for 10 min at $4^{\circ} \mathrm{C}$ and loaded onto nickel-nitrilotriacetic acid $\left(\mathrm{Ni}^{2+}-\mathrm{NTA}\right)$ resin columns on an Äkta Pure system connected to an F9-C fraction collector (GE). Finally, the fractions containing the purified protein was dialyzed overnight against a buffer containing 50 mM Tris- $\mathrm{HCl}(\mathrm{pH} 7.5), 150 \mathrm{mM} \mathrm{NaCl}$, and $10 \%$ glycerol, flash-frozen in liquid

TABLE 2 Plasmids

| Name | Parent vector | Description ${ }^{a}$ | Source or reference |
| :---: | :---: | :---: | :---: |
| pCDFDuet-1 |  |  | Novagen |
| pCBJ215 | pCDFDuet-1 | pCDFDuet-1 MCS2::RsTAL, Sp ${ }^{\text {r }}$ | This work |
| pCBJ216 | pCDFDuet-1 | pCDFDuet-1 MCS2::S_BagA, Sp ${ }^{\text {r }}$ | This work |
| pCBJ217 | pCDFDuet-1 | pCDFDuet-1 MCS2::RmXAL, Sp ${ }^{\text {r }}$ | This work |
| pCBJ218 | pCDFDuet-1 | pCDFDuet-1 MCS2::SeSam8, $\mathrm{Sp}^{\text {r }}$ | This work |
| pCBJ219 | pCDFDuet-1 | pCDFDuet-1 MCS2::BlPAL, $\mathrm{Sp}^{\text {r }}$ | This work |
| pCBJ220 | pCDFDuet-1 | pCDFDuet-1 MCS2::R_XAL, Sp ${ }^{\text {r }}$ | This work |
| pCBJ221 | pCDFDuet-1 | pCDFDuet-1 MCS2::PpPAL, $\mathrm{Sp}^{\text {r }}$ | This work |
| pCBJ222 | pCDFDuet-1 | pCDFDuet-1 MCS2::LbTAL, $\mathrm{Sp}^{\text {r }}$ | This work |
| pCBJ223 | pCDFDuet-1 | pCDFDuet-1 MCS2::IlTAL, $S^{\text {r }}$ | This work |
| pCBJ224 | pCDFDuet-1 | pCDFDuet-1 MCS2::DdPAL, $\mathrm{Sp}^{\text {r }}$ | This work |
| pCBJ225 | pCDFDuet-1 | pCDFDuet-1 MCS2::SrXAL, Sp ${ }^{\text {r }}$ | This work |
| pCBJ226 | pCDFDuet-1 | pCDFDuet-1 MCS2::L_XAL, Sp ${ }^{\text {r }}$ | This work |
| pCBJ227 | pCDFDuet-1 | pCDFDuet-1 MCS2::HaTAL1, Spr | This work |
| pCBJ228 | pCDFDuet-1 | pCDFDuet-1 MCS2::FjTAL, Sp ${ }^{\text {r }}$ | This work |
| pCBJ229 | pCDFDuet-1 | pCDFDuet-1-His6; cloning vector for expression of $\mathrm{His}_{6}$-tagged proteins, $\mathrm{Sp}^{\text {r }}$ | This work |
| pCBJ230 | pCBJ229 | pCBJ229::RsTAL, Sp ${ }^{\text {r }}$ | This work |
| pCBJ233 | pCBJ229 | pCBJ229::SeSam8, Sp ${ }^{\text {r }}$ | This work |
| pCBJ242 | pCBJ229 | pCBJ229::HaTAL1, $\mathrm{Sp}^{\text {r }}$ | This work |
| pCBJ243 | pCBJ229 | pCBJ229::FjTAL, Sp ${ }^{\text {r }}$ | This work |
| pCfB132 |  | Episomal replication vector with USER cassette derived from pESC-URA (Agilent); Amp ${ }^{\mathrm{r}}$, URA | 49 |
| pCBJ278 | pCfB132 | pCfB132::PPGK1->SeSam8, Amp ${ }^{\text {r }}$, URA | This work |
| pCBJ279 | pCfB132 | pCfB132::PPGK1->HaTAL1, Amp ${ }^{\text {r }}$, URA | This work |
| pCBJ280 | pCfB132 | pCfB132::PPGK1->FjTAL, Amp ${ }^{\text {r }}$, URA | This work |
| pCBJ281 | pCfB132 | pCfB132::PPGK1->HaTAL1Sc, Amp ${ }^{\text {r }}$, URA | This work |
| pCBJ282 | pCfB132 | pCfB132::PPGK1->FjTALSc, Amp ${ }^{\text {r }}$, URA | This work |
| pCBJ295 | pCDFDuet-1 | pCDFDuet-1 MCS2::RtXAL, Sp ${ }^{\text {r }}$ | This work |
| pCBJ296 | pCDFDuet-1 | pCDFDuet-1 MCS2::TcXAL, Sp ${ }^{\text {r }}$ | This work |
| pCBJ297 | pCDFDuet-1 | pCDFDuet-1 MCS2::RcTAL-var1, $\mathrm{Sp}^{\text {r }}$ | This work |
| pCBJ298 | pCDFDuet-1 | pCDFDuet-1 MCS2::RcTAL-var2, $\mathrm{Sp}^{\text {r }}$ | This work |
| pCBJ299 | pCDFDuet-1 | pCDFDuet-1 MCS2::HaTAL2, Sp ${ }^{\text {r }}$ | This work |
| pCBJ300 | pCDFDuet-1 | pCDFDuet-1 MCS2::PcXAL, Sp ${ }^{\text {r }}$ | This work |
| pCBJ301 | pCDFDuet-1 | pCDFDuet-1 MCS2::RtXAL-EP18Km-6, Sp ${ }^{\text {r }}$ | This work |
| pCBJ302 | pCDFDuet-1 | pCDFDuet-1 MCS2::RtXAL-RM120-1, Sp ${ }^{\text {r }}$ | This work |
| pCBJ303 | pCfB132 | pCfB132::PPGK1->RtXAL, Amp ${ }^{\text {r }}$, URA | This work |
| pCBJ304 | pCfB132 | pCfB132::PPGK1->RcTAL-var1, Amp ${ }^{\text {r }}$, URA | This work |
| pCBJ305 | pCfB132 | pCfB132::PPGK1->RcTAL-var2, Amp ${ }^{\text {r }}$, URA | This work |
| pCBJ306 | pCfB132 | pCfB132::PPGK1-> HaTAL2, Amp ${ }^{\text {r }}$, URA | This work |
| pCBJ307 | pCfB132 | pCfB132::PPGK1->PcXAL, Amp ${ }^{\text {r }}$, URA | This work |
| pCBJ308 | pCfB132 | pCfB132::PPGK1->RtXAL-EP18Km-6, Amp ${ }^{\text {r }}$, URA | This work |
| pCBJ309 | pCfB132 | pCfB132::PPGK1->RtXAL-RM120-1, Amp ${ }^{\text {r }}$, URA | This work |
| pCfB391 |  | pXI-5-HIS5, yeast integrative plasmid containing USER cassette; TADH1 and TCYC1 from S. cerevisiae, Amp ${ }^{\text {r }}$ | 49 |
| pCfB860 | pCfB391 (pXI-5-HIS5) | RsTAL, Amp ${ }^{\text {r }}$, HIS | This work |
| pCfB861 | pCfB391 (pXI-5-HIS5) | S_BagA, Amp ${ }^{\text {r }}$, HIS | This work |
| pCfB862 | pCfB391 (pXI-5-HIS5) | RmXAL, Amp ${ }^{\text {r }}$, HIS | This work |
| pCfB863 | pCfB391 (pXI-5-HIS5) | SeSam8, Amp ${ }^{\text {r }}$, HIS | This work |
| pCfB864 | pCfB391 (pXI-5-HIS5) | BlPAL, Amp ${ }^{\text {r }}$, HIS | This work |
| pCfB865 | pCfB391 (pXI-5-HIS5) | R_XAL, Amp ${ }^{\text {r }}$, HIS | This work |
| pCfB866 | pCfB391 (pXI-5-HIS5) | PpPAL, Amp ${ }^{\text {r }}$, HIS | This work |
| pCfB867 | pCfB391 (pXI-5-HIS5) | LbTAL, Amp ${ }^{\text {r }}$, HIS | This work |
| pCfB868 | pCfB391 (pXI-5-HIS5) | IlTAL, Amp ${ }^{\text {r }}$, HIS | This work |
| pCfB869 | pCfB391 (pXI-5-HIS5) | DdPAL, Amp ${ }^{\text {r }}$, HIS | This work |
| pCfB870 | pCfB391 (pXI-5-HIS5) | SrXAL, Amp ${ }^{\text {r }}$, HIS | This work |
| pCfB871 | pCfB391 (pXI-5-HIS5) | L_XAL, Amp ${ }^{\text {r }}$, HIS | This work |
| pCfB872 | pCfB391 (pXI-5-HIS5) | HaTAL1, Amp ${ }^{\text {r }}$, HIS | This work |
| pCfB873 | pCfB391 (pXI-5-HIS5) | FjTAL, Amp ${ }^{\text {r }}$, HIS | This work |
| pNZ8048 |  | $\mathrm{Cm}^{\mathrm{r}}$; inducible expression vector carrying $\mathrm{P}_{\text {nis }}$ | 56 |
| pNZ_RsTAL ${ }_{L l}$ | pNZ8048 | $\mathrm{Cm}^{\mathrm{r}}$; derivative of pNZ8048 carrying the Rhodobacter sphaeroides TALencoding gene codon optimized for expression in L. lactis | Gaspar et al., unpublished |

TABLE 2 (Continued)

| Name | Parent vector | Description ${ }^{a}$ | Source or reference |
| :---: | :---: | :---: | :---: |
| pNZ_RmXAL ${ }_{L l}$ | pNZ8048 | $\mathrm{Cm}^{\mathrm{r}}$; derivative of pNZ8048 carrying the Rhodotorula mucilaginosa PAL/ TAL-encoding gene codon optimized for expression in L. lactis | This work |
| pNZ_SeSam8 | pNZ8048 | $\mathrm{Cm}^{\mathrm{r}}$; derivative of pNZ8048 carrying the SeSam8-encoding gene codon optimized for expression in E. coli | This work |
| pNZ_R_XAL | pNZ8048 | $\mathrm{Cm}^{\mathrm{r}}$; derivative of pNZ8048 carrying the R_XAL-encoding gene codon optimized for expression in E. coli | This work |
| pNZ_HaTAL1 | pNZ8048 | $\mathrm{Cm}^{\mathrm{r}}$; derivative of pNZ8048 carrying the HaTAL1-encoding gene codon optimized for expression in E. coli | This work |
| pNZ_FjTAL | pNZ8048 | $\mathrm{Cm}^{\mathrm{r}}$; derivative of pNZ8048 carrying the FjTAL-encoding gene codon optimized for expression in E. coli | This work |

${ }^{a} \mathrm{Cm}^{\mathrm{r}}$, chloramphenicol resistance; $\mathrm{Sp}^{\mathrm{r}}$, spectinomycin resistance; $\mathrm{Amp}^{\mathrm{r}}$, ampicillin resistance.
nitrogen, and stored at $-80^{\circ} \mathrm{C}$. The purification and protein sizes were assessed by SDS-PAGE (see Fig. S2 in the supplemental material).

Enzymatic assays were performed in $200-\mu$ l volumes in wells in a UV transparent 96-well plate, by following the increase in absorbance at 315 nm (pHCA) or 295 nm (CA). The pH optima were determined in 50 mM potassium phosphate buffer for pH 6.0 to 8.0 and in 50 mM potassium borate buffer for pH 8.5 to 10.5 . The reaction mixtures contained $2 \mu \mathrm{~g}$ of purified protein, and the reactions were initiated with 1 mM tyrosine or 6 mM phenylalanine after equilibration to $30^{\circ} \mathrm{C}$. The enzymatic activity was expressed in units per gram, where 1 unit is defined as $1 \mu \mathrm{~mol}$ substrate converted per minute. We did not observe any production in the absence of enzymes under any conditions.

The kinetic constants $K_{m}$ and $V_{\max }$ were determined from assays containing from $1.56 \mu \mathrm{M}$ to $200 \mu \mathrm{M}$ tyrosine or from $193 \mu \mathrm{M}$ to 25 mM phenylalanine.

Nucleotide sequence accession numbers. The codon-optimized genes (see Table S1 in the supplemental material) have been deposited in GenBank under accession numbers KR095285 to KR095310.

## RESULTS AND DISCUSSION

Using synteny and phylogeny to identify enzymes. We examined the diversity of known and hypothetical XAL enzymes, aiming at identifying enzymes from the protein sequence databases that differed from the previously characterized enzymes. For this purpose, we created a protein sequence subset from the nonredundant protein sequence database at GenBank by identifying homologs to four nonsimilar known genes with TAL or PAL activity: those encoding RsTAL from R. sphaeroides (1, 60), S_BagA from Streptomyces (40), RmXAL from Rhodotorula mucilaginosa (61, 62), and SeSam8 from Saccharothrix espanaensis $(21,63)$.

In order to assess the true diversity of XAL enzymes, and to further counter the challenge of sequence bias, we condensed the data set for 4,729 unique sequences into 107 representative clusters, each containing 1 to 1,201 sequences with more than $40 \%$ sequence identity to the first cluster representative sequence.

Sequencing bias is illustrated by the fact that $99 \%$ of sequences found in the largest cluster, making up one fourth of all sequences, could be identified as belonging to proteobacteria. To a large extent, each cluster represented sequences found in a single phylum, but a single phylum could be represented in several clusters. For example, a single strain may have two homologs in its genome, one being a HAL and the other a PAL. Overall, the majority of sequences were of bacterial origin, with proteobacteria contributing to $47 \%$ of all sequences that could be assigned to phyla, reflecting the overrepresentation of proteobacteria in the database.

Figure 2 shows a phylogenetic tree containing representatives of each cluster together with previously characterized proteins
(see Table S2 in the supplemental material) (all functionally characterized TAMs, PAMs, and TALs listed in BRENDA [http://www .brenda-enzymes.info] [64], 10 PALs, and three HALs) and enzymes chosen for further study here.

Since functionally related genes may show synteny (colocalization of genetic loci) and cluster together (high linkage) in microbial genomes, we examined the combinatorial presence of characteristic proteins that could hint at the activity encoded by the candidate gene. Some bacteria use pHCA as a chromophore in the light-sensing photoactive yellow protein (PYP) $(18,65)$, and homologs to PYP were found in close proximity to a TAL homolog in Idiomarina loihiensis, Halorhodospira halophila, and Curvibacter. R. sphaeroides, Rhodobacter capsulatus, Salinibacter ruber (66), Rhodopseudomonas palustris, Sorangium cellulosum, Leptospira biflexa, Leptothrix cholodnii, Gemmatimonas aurantiaca, and Haliangium ochraceum carry PYP gene homologs elsewhere in their chromosomes, while other bacteria (Rhodospirillum centenum, Methylobacterium extorquens, Methylobacterium sp. strain 4-46, Methylobacterium radiotolerans, Methylobacterium populi, and Methylobacterium chloromethanicum) have a PYP homolog as part of a larger protein.

The enzyme 4-coumaroyl-CoA ligase (4CL) forms the activated thioester from pHCA and CoA, required for example for formation of pHCA-bound PYP or for formation of flavonoids or stilbenes. A homolog to a 4CL may act on either pHCA or CA, but not on urocanic acid, and is thus linked to PAL and TAL activity but not HAL activity. HAL activity is the first step in the degradation of histidine to urocanic acid, which is then processed by the enzyme urocanase (urocanate hydratase), whose gene is commonly annotated as $h u t U$ in bacteria.

Thus, we used linkage (within a distance of five genes) of either a 4CL gene or $h u t U$ as a criterion for predicting enzymatic function, and the clusters containing XAL homologs closely linked to 4CL or hutU are shown in Fig. 2. No cluster contained two XALs with different linkages. The largest clusters carried sequences homologous to HALs and linkage to hutU, showing that the HAL functionality is more abundant in the sequence database (Table 4). However, close inspection showed that, the catalytically important MIO-forming amino acid triad and a conserved tyrosine (Y300 in RsTAL) (22) were absent from many sequences, while the immediate surrounding amino acid residues were conserved. Some of the coding sequences are genetically linked to histidine degradative genes, and it is unclear if their products are catalytically active as histidases independent of a MIO-based mechanism, e.g., by using an alternative prosthetic group, such as formation of

TABLE 3 Strains

| Strain | Description | Reference or source |
| :---: | :---: | :---: |
| E. coli |  |  |
| CBJ772 | DH5 $/$ /pCBJ215 | This work |
| CBJ773 | DH5 $/$ /PCBJ216 | This work |
| CBJ774 | DH5 / /pCBJ217 | This work |
| CBJ775 | NEB5 $/$ /pCBJ218 | This work |
| CBJ776 | NEB5 $/$ /pCBJ219 | This work |
| CBJ777 | NEB5 $/$ /pCBJ220 | This work |
| CBJ778 | NEB5 $/$ /pCBJ221 | This work |
| CBJ779 | NEB5 $/$ /pCBJ222 | This work |
| CBJ780 | NEB5 $/$ /pCBJ223 | This work |
| CBJ781 | NEB5 $/$ /pCBJ224 | This work |
| CBJ782 | NEB5 $\alpha / \mathrm{pCBJ} 225$ | This work |
| CBJ783 | NEB5 $/$ /pCBJ226 | This work |
| CBJ784 | NEB5 $/$ /pCBJ227 | This work |
| CBJ785 | NEB5 $/$ /pCBJ228 | This work |
| CBJ786 | BL21(DE3)/pLys p pCDFDuet-1 | This work |
| CBJ787 | BL21(DE3)/pLysS pCBJ215 | This work |
| CBJ788 | BL21(DE3)/pLysS pCBJ216 | This work |
| CBJ789 | BL21(DE3)/pLysS pCBJ217 | This work |
| CBJ790 | BL21(DE3)/pLysS pCBJ218 | This work |
| CBJ791 | BL21(DE3)/pLysS pCBJ219 | This work |
| CBJ792 | BL21(DE3)/pLysS pCBJ220 | This work |
| CBJ793 | BL21(DE3)/pLysS pCBJ221 | This work |
| CBJ794 | BL21(DE3)/pLysS pCBJ222 | This work |
| CBJ795 | BL21(DE3)/pLysS pCBJ223 | This work |
| CBJ796 | BL21(DE3)/pLysS pCBJ224 | This work |
| CBJ797 | BL21(DE3)/pLysS pCBJ225 | This work |
| CBJ798 | BL21(DE3)/pLysS pCBJ226 | This work |
| CBJ799 | BL21(DE3)/pLysS pCBJ227 | This work |
| CBJ800 | BL21(DE3)/pLysS pCBJ228 | This work |
| CBJ801 | NEB5 $/$ /pCBJ229 | This work |
| CBJ802 | NEB5 $/$ /pCBJ230 | This work |
| CBJ805 | NEB5 $/$ /pCBJ233 | This work |
| CBJ814 | NEB5 $/$ /pCBJ242 | This work |
| CBJ815 | NEB5 $/$ /pCBJ243 | This work |
| CBJ827 | BL21(DE3)/pLysS pCBJ229 | This work |
| CBJ828 | BL21(DE3)/pLysS pCBJ230 | This work |
| CBJ831 | BL21(DE3)/pLysS pCBJ233 | This work |
| CBJ840 | BL21(DE3)/pLysS pCBJ242 | This work |
| CBJ841 | BL21(DE3)/pLysS pCBJ243 | This work |
| CBJ889 | NEB5 $/$ /pCBJ278 | This work |
| CBJ890 | NEB5 $/$ /pCBJ279 | This work |
| CBJ891 | NEB5 $/$ /pCBJ280 | This work |
| CBJ892 | NEB5 $/$ /pCBJ281 | This work |
| CBJ893 | NEB5 $/$ /pCBJ282 | This work |
| CBJ938 | NEB5 $/$ /pCBJ295 | This work |
| CBJ939 | NEB5 $/$ /pCBJ296 | This work |
| CBJ940 | NEB5 / /pCBJ297 | This work |
| CBJ941 | NEB5 / /pCBJ298 | This work |
| CBJ942 | NEB5 /pCBJ299 | This work |
| CBJ943 | NEB5 $/$ /pCBJ300 | This work |
| CBJ944 | NEB5 $/$ /pCBJ301 | This work |
| CBJ945 | NEB5 $/$ /pCBJ302 | This work |
| CBJ946 | BL21(DE3)/pLysS pCBJ295 | This work |
| CBJ947 | BL21(DE3)/pLysS pCBJ296 | This work |
| CBJ948 | BL21(DE3)/pLysS pCBJ297 | This work |
| CBJ949 | BL21(DE3)/pLysS pCBJ298 | This work |
| CBJ950 | BL21(DE3)/pLysS pCBJ299 | This work |
| CBJ951 | BL21(DE3)/pLysS pCBJ300 | This work |
| CBJ952 | BL21(DE3)/pLysS pCBJ301 | This work |
| CBJ953 | BL21(DE3)/pLysS pCBJ302 | This work |
| CBJ962 | NEB5 $/$ /pCBJ303 | This work |

TABLE 3 (Continued)

| Strain | Description | Reference or source |
| :---: | :---: | :---: |
| CBJ963 | NEB5 $/$ /pCBJ304 | This work |
| CBJ964 | NEB5 $/$ /pCBJ305 | This work |
| CBJ965 | NEB5 $/$ /pCBJ306 | This work |
| CBJ966 | NEB5 $/$ /pCBJ307 | This work |
| CBJ967 | NEB5 $/$ /pCBJ308 | This work |
| CBJ968 | NEB5 $\alpha /$ PCBJ309 | This work |
| S. cerevisiae |  |  |
| CEN.PK102-5B | MATa ura3-52 his3D1 leu2-3/112 MAL2-8c SUC2 | Verena Siewers (Chalmers University) |
| STC0 | Integration of pCfB 391 into CEN.PK102-5B | This work |
| STC1 | Integration of pCfB860 into CEN.PK102-5B | This work |
| STC2 | Integration of pCfB 861 into CEN.PK102-5B | This work |
| STC3 | Integration of pCfB 862 into CEN.PK102-5B | This work |
| STC4 | Integration of pCfB 863 into CEN.PK102-5B | This work |
| STC5 | Integration of pCfB 864 into CEN.PK102-5B | This work |
| STC6 | Integration of pCfB 865 into CEN.PK102-5B | This work |
| STC7 | Integration of pCfB866 into CEN.PK102-5B | This work |
| STC8 | Integration of pCfB867 into CEN.PK102-5B | This work |
| STC9 | Integration of pCfB 868 into CEN.PK102-5B | This work |
| STC10 | Integration of pCfB869 into CEN.PK102-5B | This work |
| STC11 | Integration of pCfB 870 into CEN.PK102-5B | This work |
| STC12 | Integration of pCfB871 into CEN.PK102-5B | This work |
| STC13 | Integration of pCfB 872 into CEN.PK102-5B | This work |
| STC14 | Integration of pCfB 873 into CEN.PK102-5B | This work |
| CBJ969 | CEN.PK102-5B/pCBJ278 | This work |
| CBJ970 | CEN.PK102-5B/pCBJ279 | This work |
| CBJ971 | CEN.PK102-5B/pCBJ280 | This work |
| CBJ972 | CEN.PK102-5B/pCBJ281 | This work |
| CBJ973 | CEN.PK102-5B/pCBJ282 | This work |
| CBJ981 | CEN.PK102-5B/pCfB132 | This work |
| CBJ982 | CEN.PK102-5B/pCBJ303 | This work |
| CBJ983 | CEN.PK102-5B/pCBJ304 | This work |
| CBJ984 | CEN.PK102-5B/pCBJ305 | This work |
| CBJ985 | CEN.PK102-5B/pCBJ306 | This work |
| CBJ986 | CEN.PK102-5B/pCBJ307 | This work |
| CBJ987 | CEN.PK102-5B/pCBJ308 | This work |
| CBJ988 | CEN.PK102-5B/pCBJ309 | This work |
| L. lactis |  |  |
| NZ9000 | MG1363 pepN::nisRK | 56 |
| NZ9000 ${ }^{\text {d }}$ ldh LldhB | NZ9000 derivative containing the double deletion of $l d h$ and $l d h B$ | 86 |
| NZ9000 ${ }^{\text {d }}$ ldh $4 l d h B ~ p N Z ~$ | NZ9000 ${ }^{\text {S }}$ ldh ${ }^{\text {dldhB/pNZ8048 }}$ | This work |
| NZ9000 ${ }^{\text {l }}$ ldh $\Delta l d h B \_$RsTAL ${ }_{L l}$ | NZ9000 $\Delta l d h \Delta l d h B / \mathrm{pNZ}$-RsTAL ${ }_{L l}$ | This work |
| NZ9000 ${ }^{\text {d }}$ ldh ${ }^{\text {d }}$ ldhB_RmXAL ${ }_{L l}$ | NZ9000 $\Delta l d h \Delta l d h B / \mathrm{pNZ}_{-} \mathrm{RmXAL}_{L l}$ | This work |
| NZ9000 ${ }^{\text {d }}$ ldh ${ }^{\text {d }}$ ldhB_SeSam8 | NZ9000 ${ }^{\text {d }}$ ldh $\Delta l d h B / \mathrm{pNZ}$ _SeSam8 | This work |
| NZ9000 ${ }^{\text {l }}$ ldh $\Delta l d h B \_R \_X A L$ | NZ9000 $\Delta l d h \Delta l d h B / p N Z \_R \_X A L$ | This work |
| NZ9000 ${ }^{\text {d }}$ ldh $\Delta l d h B$ _HaTAL1 | NZ9000 $\Delta l d h \Delta l d h B / p N Z \_H a T A L 1 ~$ | This work |
| NZ9000 ${ }^{\text {a }}$ ldh $\Delta l d h B$ _FjTAL | NZ9000 $\Delta l d h \Delta l d h B / \mathrm{pNZ}$ _FjTAL | This work |

dehydroalanine, as first suggested for aromatic amino acid ammo-nia-lyases ( $15,67,68$ ), or if they catalyze an alternative reaction.

From Fig. 2 it is evident that there are large groups of potential enzymes which do not show a high degree of homology to enzymes that have been functionally characterized, while many previously characterized enzymes cluster together. It is also clear that there are large groups of enzymes for which no representative has hitherto been characterized. The known HALs as well as the hutUlinked genes from both eukaryotes and prokaryotes form a large group, which is separated from the enzymes that have TAM, PAM, TAL, or PAL activity. This was also found to be true in the cases where a PAL and a HAL are from the same organism.

Aiming at functionally characterizing a wide spectrum of PAL and TAL enzymes, we chose to examine 22 proteins (Table 5), of which half were uncharacterized, representing different clusters. We included two near-identical TALs from R. capsulatus (18), a TAL from R. sphaeroides $(1,22,26,28,60,69)$ which is highly similar with regard to primary sequence ( $51 \%$ amino acid identity) and enzyme kinetics, and a third representative from this cluster (cluster 53), R_XAL, from Rheinheimera sp. strain A13L. Two homologs, from the extreme halophilic bacterium Salinibacter ruber and from the spirochete Leptospira biflexa, grouped with the previously characterized SeSam8 from Saccharothrix espanaensis (21). Because the sequences contain active-site residues similar to those


FIG 2 Phylogenetic relationship between enzymes of aromatic amino acid ammonia-lyase homologs. Included are representative from clusters of sequences that are at least $40 \%$ identical (after iteration of similar clustering based on $90 \%$ and $60 \%$ identities) and enzymes with known PAL, TAL, HAL, PAM, and TAM functions. The number of sequences in each cluster is indicated by the number after the vertical line. Characterized enzymes are shown using the following color code: red, PAL; blue, TAL; purple, PAL/TAL; yellow, HAL; light red, PAM; light blue, TAM; white, no activity. Enzymes subject to analysis in this study are marked by circles, while enzymes with activity shown elsewhere in the literature are marked by square symbols. Some groups of enzymes did not carry the sequences required for the formation of the 3,5-dihydro-5-methylidene-4H-imidazol-4-one (MIO) prosthetic group.
of TALs (23), and since the chromosomes furthermore encode PYP homologs, they were predicted to be TALs and included in this study (SrXAL, LbTAL, and SeSam8). The Idiomarina loihiensis genome encodes the homolog IlTAL, as well as both a PYP and
a 4CL. BagA from Streptomyces sp. showed the greatest homology to BlPAL from the Gram-positive organism Brevibacillus laterosporus.

We selected PpPAL from the biotechnologically relevant moss

TABLE 4 The 15 largest sequence clusters and phylogeny of selected TAL genes

| Order of cluster size | Size | Top phylum (a) (\%) ${ }^{\text {a }}$ | Linkage within cluster ${ }^{b}$ | Enzyme(s) investigated |
| :---: | :---: | :---: | :---: | :---: |
| 1 | 1,201 | Proteobacteria (99) | hutU |  |
| 2 | 763 | Firmicutes (59), Bacteroidetes (24) | hutU |  |
| 3 | 504 | Streptophyta (100) |  | PpPAL |
| 4 | 410 | Proteobacteria (99) | 4CL | IITAL |
| 5 | 360 | Actinobacteria (98) | hutU |  |
| 6 | 257 | Proteobacteria (100) | hutU |  |
| 7 | 170 | Chordata (75), Proteobacteria (3) | hutU |  |
| 8 | 96 | Bacteroidetes (100) |  | FjTAL |
| 9 | 81 | Bacteroidetes (83), Proteobacteria (16) | hutU |  |
| 10 | 74 | Actinobacteria (100) | hutU |  |
| 11 | 47 | Euryarchaeota (98) | hutU |  |
| 12 | 42 | Actinobacteria (68), Proteobacteria (32) | 4CL | RstaL, RcTAL, R_XAL |
| 13 | 41 | Basidiomycota (100) |  | PcXAL |
| 14 | 41 | Proteobacteria (93) | hutU |  |
| 15 | 40 | Firmicutes (60), Proteobacteria (20) | hutU |  |
| 26-27 | 16 | Actinobacteria (75), Firmicutes (19) |  | S_BagA, BIPAL |
| 28-29 | 15 | Cyanobacteria (67), Proteobacteria (13) |  | HaTAL, L_XAL |
| 39-42 | 7 | Basidiomycota (100) |  | RmXAL, RtXAL, TcXAL |
| 39-42 | 7 | Proteobacteria (100) |  | DdPAL |
| 39-42 | 7 | Actinobacteria (100) |  | SeSam8 |
| 44-50 | 6 | Spirochaetes (100) |  | LbTAL |
| 56-66 | 3 | Bacteroidetes (100) |  | SrXAL |

${ }^{a}$ Percentages of sequences that could be assigned to a phylum are shown. Most sequences within a cluster could be assigned to a primary phylum, and in case of less than $90 \%$ assignment, the second most represented phylum is shown as well. The largest clusters are formed by bacterial sequences that are generally linked to hutU sequences.
${ }^{b}$ Three of the selected enzymes used in this study originate from three large bacterial clusters with no linkage to $h u t U$, and there was linkage to 4CL for IITAL and RsTAL within the cluster, this was not the case for FjTAL.

Physcomitrella patens subsp. patens, and the fungal enzymes of RmXAL, TcXAL (25), PcXAL (70), and RtXAL, including two mutants of the latter for which improved enzymatic characteristics have been reported $(17,71)$. A cyanobacterial sequence from

Leptolyngbya was chosen together with two other sequences from the genome of the green nonsulfur bacterium (Chloroflexi) Herpetosiphon aurantiacus. Interestingly, the organism has a third homolog ( $31 \%$ to $32 \%$ identity over 474 to 503 amino acids) in its

TABLE 5 Enzymes analyzed in this study

| ID | Organism | Protein GI | Length ${ }^{\text {a }}$ |
| :---: | :---: | :---: | :---: |
| RmXAL ${ }^{\text {b }}$ | Rhodotorula mucilaginosa (Rhodotorula rubra) | 129592 | 713 |
| TcXAL ${ }^{\text {b }}$ | Trichosporon cutaneum | 77375521 | 552 |
| PcXAL ${ }^{\text {b }}$ | Phanerochaete chrysosporium | 259279291 | 506 |
| RtXAL ${ }^{\text {b }}$ | Rhodosporidium toruloides | 129593 | 567 |
| RtXALmut1-EP18Km-6 ${ }^{\text {b }}$ | Rhodosporidium toruloides | 21503942 | 567 |
| RtXALmut2-RM120-1 ${ }^{\text {b }}$ | Rhodosporidium toruloides | 29719264 | 552 |
| RcTAL-var1 ${ }^{\text {b }}$ | Rhodobacter capsulatus | 155708849 | 506 |
| RcTAL-var2 ${ }^{\text {b }}$ | Rhodobacter capsulatus | 534410416 | 567 |
| RsTAL ${ }^{\text {b }}$ | Rhodobacter sphaeroides | 126464011 | 523 |
| S_BagA ${ }^{\text {b }}$ | Streptomyces | 359308109 | 529 |
| SeSam8 ${ }^{\text {b }}$ | Saccharothrix espanaensis | 433607630 | 510 |
| PpPAL ${ }^{\text {c }}$ | Physcomitrella patens subsp. patens | 168011366 | 714 |
| DdPAL ${ }^{\text {c }}$ | Dictyostelium discoideum | 66822311 | 529 |
| $\mathrm{BlPAL}^{c}$ | Brevibacillus laterosporus | 339009660 | 550 |
| R_XAL ${ }^{\text {c }}$ | Rheinheimera | 336316228 | 516 |
| ITTAL ${ }^{\text {c }}$ | Idiomarina loihiensis | 56459245 | 515 |
| $\mathrm{LbTAL}^{\text {c }}$ | Leptospira biflexa serovar patoc | 183220142 | 514 |
| L_XAL ${ }^{\text {c }}$ | Leptolyngbya | 427415639 | 567 |
| SrXAL ${ }^{\text {c }}$ | Salinibacter ruber | 83816043 | 517 |
| FjTAL ${ }^{\text {c }}$ | Flavobacterium johnsoniae | 146298870 | 506 |
| HaTAL1 ${ }^{\text {c }}$ | Herpetosiphon aurantiacus | 159898407 | 552 |
| HaTAL2 ${ }^{\text {c }}$ | Herpetosiphon aurantiacus | 159898927 | 552 |

[^2]genome, which clusters together with the HALs (cluster 43). We furthermore chose to include DdPAL from Dictyostelium discoideum as a representative of a rare protein from an amoeba.

Finally, we included a bacterial sequence that showed little similarity to other known enzymes, FjTAL from Flavobacterium johnsoniae. This clusters with enzymes from Bacteroides but is otherwise grouped most closely to previously described PALs, TAMs, and PAMs mainly from actinobacteria and proteobacteria (Fig. 2). IITAL and FjTAL represent the two largest clusters with only uncharacterized coding sequences in the analysis (Table 4), and only the former cluster contains homologs linked to 4CL homologs.

Prediction of enzyme specificity from sequence or synteny. Aromatic amino acid ammonia-lyases and mutases share a range of residues that have been shown through structural studies, mutational analysis, and simulation to be important for substrate binding and catalysis (see Fig. S1 in the supplemental material) ( $36,72,73$ ). Catalysis requires the formation of MIO and tyrosine residues Y60 and Y300 (RsTAL numbering) (26), and the MIO interacts with several other amino acid side chains and residues, including L153, G204, F350, and Q436 (22), which are all conserved in the selected enzymes.

Substrate binding takes place to accommodate the carboxylic acid group through R303 and N435 and the aliphatic part through Y60 and G67 to the ring structure through L90, L153, M405, N 432 , and Q436. The far end of the substrate binding pocket may be hydrophobic in the case of phenylalanine or hydrophilic in the case of tyrosine or histidine as a result of the presence of either aliphatic residues or charged residues (H89 and L90 in RsTAL). Mutational studies have confirmed that this last region is important for specificity and that this may be changed through mutation (22, 23). We included the RtXAL-RM120-1 mutant, where a mutation in this site resulted in increased specificity for tyrosine. The residues in this region are, however, not sufficient for predicting the substrate specificity. FjTAL, for example, has the highest sequence similarity to HALs and TAMs, while S_BagA, IITAL, HaTAL1, and HaTAL2 do not conform to any of the previously defined sequences. Mutation of V409 has also been found to change the substrate specificity (60); however, this amino acid is not conserved within enzymes of the same specificity. A reaction mechanism has also been hypothesized to be determined by N432, which was an asparagine for TALs and TAMs and a glutamine for PALs and PAMs (28). However, there is no such consistency between substrate specificity and this residue for the newly identified enzymes, since S_BagA, IITAL, HaTAL1, and HaTAL2 contain a glutamine in this position. The N435-Q436 dyad does, however, appear to allow a sequence-based distinction between HALs (QE or TE sequence) and all other enzymes (NQ).

The enzymes with dual TAL and PAL activity were generally of eukaryotic origin, although the bacterial R_XAL was capable of both reactions. We speculate that bacterial enzymes are more specific, as they take part in secondary metabolite formation, where the product is generally linked to a single downstream reaction.

FjTAL, whose specificity could not be inferred from phylogeny or primary sequence, may be involved in the formation of flexirubin pigments in F. johnsoniae, giving rise to the bacterium's yellow color, as it is found close to genes characteristic of flexirubin biosynthesis $(74,75)$. The gene encoding BlPAL, which shows homology to the TAL BagA, is present in a genomic region encoding polyketide synthases. HaTAL1 and HaTAL2 are very similar, while the third homolog encoded in the genome is more distinct and is
associated with histidine degradation. However, none of three TAL homologs present in the Herpetosiphon aurantiacus genome belongs to the proposed biosynthetic loci (76). Since sequence motifs as well as synteny were not sufficient to predict substrate specificity for all proposed enzymes, we performed an in vivo screening.

Screening for $p$-coumaric acid production in Escherichia coli. The coding sequences of all selected proteins were codon optimized and expressed in an E. coli B strain that was grown in minimal medium optionally supplemented with a 2 mM concentration of either tyrosine, phenylalanine, or histidine, and the resulting supernatants were analyzed for pHCA and CA formation (Fig. 3A). We were able to measure product formation for all enzymes except L_XAL and SrXAL. The enzymes RsTAL, RcTALvar1, RcTAL-var2, R_TAL, IlTAL, S_BagA, SeSam8, LbTAL, HaTAL1, HaTAL2, and FjTAL all had TAL activities more than 10 -fold higher than their PAL activities, except for R_XAL. The two variant enzymes from R. capsulatus shared a TAL-over-PAL preference, but one variant (RcTAL-var1) resulted in a 4 -foldhigher specific production than the other despite only small sequence differences. RcTAL-var1 has a ten-amino-acid N -terminal extension, a single residue change, and a single residue gap near the C-terminal end compared to RcTAL-var2. The exogenous addition of the substrate tyrosine or phenylalanine to the growth medium caused an increased production of the products pHCA and CA, respectively (see Table S3 in the supplemental material). We also tested histidine as a potential substrate, but no urocanic acid was formed (data not shown), confirming that none of the selected enzymes had HAL activity when the genes were heterologously expressed. The highest production of pHCA was observed from the strains expressing SeSam8, FjTAL, and HaTAL1, with concentrations of $0.54,0.44$, and 0.13 mM pHCA OD 600 unit $^{-1}$ ( $1.6,1.3$, and $0.38 \mathrm{mmol} \mathrm{pHCA}[\mathrm{g} \mathrm{CDW}]^{-1}$ ) while producing 18 , 0.5 , and $1.1 \mu \mathrm{M} \mathrm{CA} \mathrm{OD} 600$ unit $^{-1}(55,1.5$, and $3.2 \mu \mathrm{~mol} \mathrm{CA} \mathrm{[g}$ $\mathrm{CDW}]^{-1}$ ), respectively. In order to determine the maximum rate of catalysis under the chosen reaction conditions, we followed the in vivo production of pHCA from E. coli after induction in late exponential growth phase with exogenous tyrosine added. The results (Fig. 4) show patterns similar to those obtained with the endpoint data (Fig. 3) and also show that SeSam8 and FjTAL are equally effective, while HaTAL1 produces pHCA at a rate around one third that of the other enzymes. SeSam8 and FjTal show pHCA productivities $\left(43 \pm 12 \mu \mathrm{M} \mathrm{OD}_{600}\right.$ unit $^{-1} \mathrm{~h}^{-1}$ and $36 \pm 4.2$ $\mu \mathrm{M} \mathrm{OD}_{600}$ unit $^{-1} \mathrm{~h}^{-1}$ ) that were not significantly different, while HaTAL1 gave a significantly lower productivity ( $6.2 \pm 0.53 \mu \mathrm{M}$ $\mathrm{OD}_{600}$ unit $^{-1} \mathrm{~h}^{-1}$ ).

The yeast enzymes RmXAL, TcXAL, PcXAL, and RtXAL had almost equal TAL and PAL activities, given the substrate availability and product secretion in vivo. RtXAL had almost equal preference for the TAL and PAL reactions, in accordance with published results, and while one of the mutants, RtXAL-EP18Km-6, showed little difference in production from the wild type, the other mutant, RtXAL-RM120-1, had a clearly altered preference for the TAL reaction, as expected, but also an $86 \%$ diminished overall activity. Lastly, BlPAL, PpPAL, and DdPAL all exclusively had PAL activity. Based on whether there was a $>10$-fold preference for a reaction, we designated the enzymes as TAL, PAL, or XAL (less than 10 -fold preference for either) (see Fig. S3 in the supplemental material).

Phylogenetically, HAL enzymes form a cluster separate from other enzymes (Fig. 2). Ritter and Schulz proposed that the eukary-


FIG 3 Production of p-coumaric acid (pHCA) and cinnamic acid (CA) in E. coli, L. lactis, and S. cerevisiae expressing aromatic amino acid ammonia-lyases. (A) Twenty-two open reading frames, codon optimized for E. coli, were expressed in E. coli. The pHCA and CA titers were measured from culture supernatants of cells grown in M9 minimal medium with glucose and supplemented with 2 mM tyrosine or phenylalanine, respectively. Some enzymes had both phenylalanine ammonia-lyase activity (PAL) and tyrosine ammonia-lyase activity (TAL), while others showed a high degree of substrate specificity. A gray background indicates enzymes previously studied in literature and also examined here. (B) Selected genes were expressed in L. lactis grown in medium containing both phenylalanine and tyrosine ( 1.7 mM ). The pHCA production was increased when the cells were grown in medium with MOPS (morpholinepropanesulfonic acid) and adjusted to pH 7. (C) Genes were also expressed chromosomally in S. cerevisiae from the TEF1 promoter. Titers are from growth in synthetic fed-batch medium with the addition of either tyrosine or phenylalanine. The pHCA production was increased when the gene encoding FjTAL was codon optimized for S. cerevisiae and expressed from the PGK1 promoter on an episomal plasmid (FjTALSc-2 $\mu$ ). "Control" indicates the titers reached from a strain containing the native plasmid with no gene inserted or from S. cerevisiae without a gene integrated.
otic PALs evolved from a bacterial HAL early in time, while the bacterial PAL/TALs emerged independently later, also from bacterial HALs (38). The cyanobacterial PALs have been suggested to be related to an intermediate in this evolution (24). The previously described homologs from the cyanobacteria Anabaena variabilis
(AvPAL) and Nostoc punctiforme (NpPAL) have not had any measurable TAL activity (24), yet we identified the TALs, HaTAL1 and HaTAL2, that show greater sequence similarity to AvPAL and NpPAL than to any previously characterized TALs, suggesting that there has been convergent evolution of TAL and PAL activity.


FIG 4 Production of pHCA in E. coli with tyrosine ammonia-lyases SeSam8, HaTAL1, and FjTAL. Genes encoding SeSam8, HaTAL1, or FjTAL were induced in cells after 3 h of growth in M9 minimal medium with glucose. After another 4 h , in the late exponential phase, tyrosine was added to 2 mM (time zero). The cell density reached stationary phase $\left(\mathrm{OD}_{600}\right.$ of around 1.4$)$ within 2 h thereafter.

The included plant enzyme, PpPAL from moss, clusters together with other plant enzymes that generally have a preference for the PAL reaction over the TAL reaction, and this enzyme shows a perfect specificity for phenylalanine. The PAL from $D$. discoideum (DdPAL) is shorter ( 529 amino acids) than the previously described PAL enzymes of eukaryotes (over 700 amino ac-
ids), as it lacks around 50 residues in the N -terminal end and around 180 residues in the C-terminal region compared to others (see Table S2 in the supplemental material). Its closest known homolog ( $49 \%$ identity) is the bacterial PAL enzyme StlA (19). The amoeba has been suggested to have acquired other prokaryotic genes by horizontal gene transfer (77), and this also is a likely origin of DdPAL.

Since substrate specificity is important for product purity, previous work focused on changing the specificity of yeast enzymes for the TAL reaction. RtXAL-RM120-1 did appear to have a higher TAL/PAL ratio than its parent ( 0.9 versus 3.0 ), but this ratio was still lower than what was observed for the novel bacterial enzymes, FjTAL and HaTAL1, where only trace amounts of CA could be detected.

Kinetic characterization of selected TAL enzymes. We further wanted to examine if the in vivo production observed in E. coli reflected the enzyme kinetic parameters, as the apparent in vivo activity also reflects substrate availability and product secretion. SeSam8, FjTAL, and HaTAL1, which showed the highest in vivo TAL activity while still being specific, were examined together with the previously characterized enzyme RsTAL. The four enzymes were purified or partially purified by identical means using N -terminal histidine tagging $(19,26)$ (see Fig. S2 in the supplemental material).

After the tyrosine concentration that resulted in maximal activity with a given enzyme concentration had been determined, the pH optima of RsTAL and SeSam8 and the novel enzymes HaTAL1 and FjTAL were determined (Fig. 5). The maximal enzymatic activities were found at pH 9.0 or above, where they were found to level off for all enzymes except HaTAL1, which showed a


FIG 5 pH optima for RsTAL, SeSam8, HaTAL1, and FjTAL. The pH optima for the TAL and PAL reactions were measured from pH 6.0 to 10.5 in $0.5-\mathrm{pH}-\mathrm{unit}$ intervals and expressed as nanomoles of product produced per minute per milligram of protein.

TABLE 6 Kinetic analysis of four selected TAL homologs using tyrosine or phenylalanine as the substrate

| Enzyme (organism) | Substrate | $K_{m}(\mu \mathrm{M})$ | $k_{\text {cat }}\left(\mathrm{s}^{-1}\right)$ | $k_{\text {cat }} / K_{m}\left(\mathrm{mM}^{-1} \mathrm{~s}^{-1}\right)$ | TAL/PAL | Reference or source ${ }^{a}$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| RsTAL (Rhodobacter sphaeroides) | Tyr | 8.4 | 0.026 | 2.38 | 88 | This work |
|  | Phe | 2100 | 0.075 | 0.0272 |  |  |
| SeSam8 (Saccharothrix espanaensis) | Tyr | 4.7 | 0.015 | 3.05 | 1,200 | This work |
|  | Phe | 6300 | 0.016 | 0.00249 |  |  |
| HaTAL1 (Herpetosiphon auranticus) | Tyr | 24 | 0.076 | 2.74 | 500 | This work |
|  | Phe | 33000 | 0.21 | 0.00552 |  |  |
| FjTAL (Flavobacterium johnsoniae) | Tyr | 6.7 | 0.023 | 2.99 | 2,400 | This work |
|  | Phe | 6600 | 0.0094 | 0.00123 |  |  |
| RsTAL ${ }^{b}$ (Rhodobacter sphaeroides) | Tyr | 74.2 | 4.32 | 1.15 | 51 | 22 |
|  | Phe | 11400 | 13.1 | 58.2 |  |  |
| RsTAL ${ }^{\text {b }}$ (Rhodobacter sphaeroides) | Tyr | 31.4 | 3.4 | 108 | 268 | 23 |
|  | Phe | 5204 | 2.1 | 0.404 |  |  |
| RsTAL ${ }^{\text {b }}$ (Rhodobacter sphaeroides) | Tyr | 60 | 0.02 | 0.333 | 19 | 60 |
|  | Phe | 560 | 0.01 | 0.018 |  |  |
| RsTAL ${ }^{\text {b }}$ (Rhodobacter sphaeroides) | Tyr | 301 | 0.08 | 0.265 | 90 | 26 |
|  | Phe | 6170 | 0.02 | 0.00297 |  |  |
| RsTAL ${ }^{\text {b }}$ (Rhodobacter sphaeroides) | Tyr | 100 | 0.9 | 10 | 44 | 28 |
|  | Phe | 2700 | 0.6 | 0.226 |  |  |
| SeSam8 (Saccharothrix espanaensis) | Tyr | 15.5 | 0.015 | 0.968 | 728 | 21 |
|  | Phe | 2860 | 0.0038 | 0.00133 |  |  |

${ }^{a}$ Data from literature were added for comparison for the previously characterized enzymes RsTAL and SeSam8.
${ }^{b}$ RsTAL from the literature (GI 46193106) differs in 8 positions from the RsTAL in this study.
decrease in activity at high pH . Similar pH optima were observed for the PAL and TAL reactions. The TAL activity was highest for HaTAL1 and lowest for SeSam8, reaching 95 and $18 \mathrm{nmol} \mathrm{mg}{ }^{-1}$ $\mathrm{min}^{-1}$, respectively. The PAL activity was highest for RsTAL and lowest for FjTAL, reaching 65 and $2 \mathrm{nmol} \mathrm{mg}^{-1} \mathrm{~min}^{-1}$, respectively, in the presence of 6 mM phenylalanine.

The kinetic constants $K_{m}$ and $V_{\max }$ were determined for the enzymes with either tyrosine or phenylalanine as a substrate at their optimal pH , as summarized in Table 6. SeSam8, HaTAL1, and FjTAL had very little activity for phenylalanine, with $k_{\text {cat }} / K_{m}$ values of $2.49,5.52$, and $1.23 \mathrm{M}^{-1} \mathrm{~s}^{-1}$, respectively. HaTAL1 had high turnover numbers for both tyrosine $\left(0.076 \mathrm{~s}^{-1}\right)$ and phenylalanine $\left(0.21 \mathrm{~s}^{-1}\right)$ but very low affinity for phenylalanine (33 $\mathrm{mM})$. FjTAL was the enzyme with the lowest activity for phenylalanine, resulting in a 2,400 -fold-higher specificity constant, $k_{\text {cat }} /$ $K_{m}$, between the two substrates. In comparison, SeSam8 was found to have a 1,200 -fold difference in specificity between the substrates.

The kinetic constants are in line with previous analysis of SeSam8 (21), but the importance of standardized assays is evident when data for RsTAL in the literature, where dramatically different kinetic constants for the same enzyme have been presented, are compared (Table 6). These differences may have been due to different expression and assay conditions.

The high pH optimum of the reactions may be a reflection of the reaction mechanism and is generally identical for all enzymes and similar to what has been observed previously for RsTAL,

SeSam8, and other homologous enzymes (18, 20, 21, 23, 27, 7881). Alkaline pH may also be used for biocatalysis, where tyrosine produced by fermentation at physiological pH can be converted to pHCA in high titers by subsequent reaction at alkaline pH in the presence of TAL-expressing E. coli cells or cell paste (82).

Production of $p$-coumaric acid in Lactococcus lactis. The performance of the novel enzymes were tested in a Gram-positive bacterium. Genes encoding HaTAL1, FjTAL, and four other enzymes with TAL activity in E. coli were expressed individually in $L$. lactis from plasmids using a nisin-inducible promoter, and the production was measured in culture supernatants (Fig. 3B). Even though the genes encoding RsTAL and RmXAL had been specifically codon optimized for $L$. lactis, $\mathrm{RsTAL}_{L l}$ and $\mathrm{RmXAL}_{L l}, \mathrm{FjTAL}$ showed by far the highest specific production of pHCA $(15 \mu \mathrm{M}$ $\mathrm{OD}_{600}$ unit $^{-1}, 43 \mu \mathrm{~mol}[\mathrm{~g} \mathrm{CDW}]^{-1}$ ), a 5 -fold increase in specific production over $\mathrm{RmXAL}_{L l}$, the second-best enzyme. The values were lower than those achieved in E. coli, and the production of pHCA could be slightly increased when the concentration of tyrosine in the medium was increased from 1.7 mM to 3.7 mM and the pH of the medium was increased from 6.5 to $7.0 . \mathrm{RmXAL}_{L l}$ was the only enzyme resulting in production of CA (see Table S3 in the supplemental material).

Production of $p$-coumaric acid in Saccharomyces cerevisiae. Yeasts, and in particular S. cerevisiae, are widely used as a cell factory for production of various chemicals (83). However, all known TAL enzymes of fungal origin also have at least a partial PAL activity. Well-described examples include enzymes used in
this study: RmXAL, TcXAL, PcXAL, and RtXAL. Twenty-one genes were cloned and were chromosomally integrated into $S$. cerevisiae and/or expressed from an episomal plasmid based on a $2 \mu$ origin of replication. The resulting strains were analyzed for production of CA and pHCA in minimal defined medium and in FIT fed-batch-simulation medium with or without addition of tyrosine or phenylalanine (see Table S3 in the supplemental material).

The novel enzymes HaTAL1 and FjTAL resulted in the highest production of pHCA, 90 and $89 \mu \mathrm{M} \mathrm{pHCA} \mathrm{OD}{ }_{600}$ unit $^{-1}(133$ and $130 \mu \mathrm{~mol}$ pHCA [g CDW] ${ }^{-1}$ ) (Fig. 3C; also, see Table S3 in the supplemental material) while remaining highly specific for the TAL reaction over the PAL reaction, as no CA could be detected. The specific production was almost 4 -fold higher than what was obtained using the previously reported TAL, RsTAL ( $23 \mu \mathrm{M}$ $\mathrm{pHCA} \mathrm{OD}_{600}$ unit $^{-1} ; 33 \mu \mathrm{~mol} \mathrm{pHCA}[\mathrm{g} \mathrm{CDW}]^{-1}$ ).

As for E. coli, expression of genes encoding PpPAL, DdPAL, and BlPAL gave CA as the sole product, while PpPAL gave the highest specific production of CA $\left(22 \mu \mathrm{M} \mathrm{OD}_{600}\right.$ unit $^{-1} ; 33 \mu \mathrm{~mol}$ CA $\left.[\mathrm{g} \mathrm{CDW}]^{-1}\right)$. RmXAL led to production of both pHCA and CA, favoring pHCA, in contrast to the result for E. coli and L. lactis, where the CA production was equal or higher. RsTAL, S_BagA, SeSam8, R_XAL, IlTAL, LbTAL, FjTAL, and HaTAL1 all gave pHCA as the sole product, unlike in E. coli, where trace amounts of CA could be measured from all enzymes. The titers in defined minimal medium could be increased by the addition of phenylalanine or tyrosine, while the extent was less significant in the FIT medium. An analysis of the dynamics of pHCA production in FIT medium (see Fig. S4 in the supplemental material) by five strains expressing either of RsTAL, S_BagA, RmXAL, HaTAL1, or FjTAL revealed that additionally supplemented tyrosine was consumed fast by S. cerevisiae, without a significant simultaneous production of pHCA . Rather, the majority of the pHCA production was observed in the subsequent part of the growth experiment. Consumption of tyrosine without comparative production of the downstream pathway has previously been seen for resveratrol production in yeast (84). In S. cerevisiae, the previously best-performing TAL enzymes have been reported to be those from Rhodosporidium (RtXAL) $(3,85)$ and Rhodobacter (RcTAL and RsTAL) (7, 9). In comparison, the novel FjTAL and HaTAL1 enzymes gave a specific pHCA production that was three- to fivefold higher while not resulting in any measurable production of CA as a side product (see Table S3 in the supplemental material). The production was further increased by codon optimizing the genes encoding HaTAL1 and FjTAL for S. cerevisiae and expressing these from an episomal plasmid (see Table S3 in the supplemental material). The most significant improvements were observed for FjTAL (Fig. 3C), reaching $200 \mu \mathrm{M} \mathrm{OD}_{600}$ unit $^{-1}\left(300 \mu \mathrm{~mol}[\mathrm{~g} \mathrm{CDW}]^{-1}\right)$ in FIT medium supplemented with tyrosine.

Conclusion. In conclusion, 22 TAL and PAL genes were expressed and screened for product formation in E. coli. Selected enzymes were purified and characterized kinetically, and their efficacy was further demonstrated in the industrially important Gram-positive bacterium L. lactis as well as in the yeast S. cerevisiae. To our knowledge, this constitutes the largest comparative study of TAL homologs for biotechnological purposes reported so far, and it sheds light on the broader phylogenetic distribution of TAL and PAL activities, as the novel enzymes presented here cover a portion of the sequence space not previously covered by characterized enzymes. The study resulted in the identification of enzymes with improved kinetic characteristics over the currently
employed enzymes. The novel TALs outperformed the previously characterized TALs when employed in vivo in a Gram-negative bacterium, a Gram-positive bacterium, and a yeast, while being very specific, as confirmed by in vitro kinetic experiments. We envision a role of the presented enzymes in biochemical production of phenolic compounds, and we are currently employing them in a series of cell factory designs.

## ACKNOWLEDGMENTS

This work was financed by the Novo Nordisk Foundation.
We thank Helle Munck Petersen and Stefan Kol of the Novo Nordisk Foundation Center for Biosustainability at the Technical University of Denmark for help with protein purification. Ana Rute Neves and Jens Nielsen are thanked for helpful discussions. We also appreciate the helpful comments from the reviewers.

## REFERENCES

1. Watts KT, Lee PC, Schmidt-Dannert C. 2004. Exploring recombinant flavonoid biosynthesis in metabolically engineered Escherichia coli. Chembiochem 5:500-507. http://dx.doi.org/10.1002/cbic. 200300783.
2. Yan Y, Kohli A, Koffas MA. 2005. Biosynthesis of natural flavanones in Saccharomyces cerevisiae. Appl Environ Microbiol 71:5610-5613. http: //dx.doi.org/10.1128/AEM.71.9.5610-5613.2005.
3. Vannelli T, Wei Qi W, Sweigard J, Gatenby AA, Sariaslani FS. 2007. Production of $p$-hydroxycinnamic acid from glucose in Saccharomyces cerevisiae and Escherichia coli by expression of heterologous genes from plants and fungi. Metab Eng 9:142-151. http://dx.doi.org/10.1016/j .ymben.2006.11.001.
4. Xue Y, Zhang Y, Cheng D, Daddy S, He Q. 2014. Genetically engineering Synechocystis sp. Pasteur Culture Collection 6803 for the sustainable production of the plant secondary metabolite $p$-coumaric acid. Proc Natl Acad Sci U S A 111:9449-9454. http://dx.doi.org/10.1073/pnas . 1323725111.
5. Nijkamp K, Westerhof RG, Ballerstedt H, de Bont JA, Wery J. 2007. Optimization of the solvent-tolerant Pseudomonas putida S12 as host for the production of $p$-coumarate from glucose. Appl Microbiol Biotechnol 74:617-624. http://dx.doi.org/10.1007/s00253-006-0703-0.
6. Wu J, Liu P, Fan Y, Bao H, Du G, Zhou J, Chen J. 2013. Multivariate modular metabolic engineering of Escherichia coli to produce resveratrol from L-tyrosine. J Biotechnol 167:404-411. http://dx.doi.org/10.1016/j .jbiotec.2013.07.030.
7. Wang Y, Halls C, Zhang J, Matsuno M, Zhang Y, Yu O. 2011. Stepwise increase of resveratrol biosynthesis in yeast Saccharomyces cerevisiae by metabolic engineering. Metab Eng 13:455-463. http://dx.doi.org/10.1016 /j.ymben.2011.04.005.
8. Santos CN, Koffas M, Stephanopoulos G. 2011. Optimization of a heterologous pathway for the production of flavonoids from glucose. Metab Eng 13:392-400. http://dx.doi.org/10.1016/j.ymben.2011.02.002.
9. Koopman F, Beekwilder J, Crimi B, van Houwelingen A, Hall RD, Bosch D, van Maris AJ, Pronk JT, Daran JM. 2012. De novo production of the flavonoid naringenin in engineered Saccharomyces cerevisiae. Microb Cell Fact 11:155. http://dx.doi.org/10.1186/1475-2859-11-155.
10. Eudes A, Juminaga D, Baidoo EE, Collins FW, Keasling JD, Loqué D. 2013. Production of hydroxycinnamoyl anthranilates from glucose in Escherichia coli. Microb Cell Fact 12:62. http://dx.doi.org/10.1186/1475 -2859-12-62.
11. Qi WW, Vannelli T, Breinig S, Ben-Bassat A, Gatenby AA, Haynie SL, Sariaslani FS. 2007. Functional expression of prokaryotic and eukaryotic genes in Escherichia coli for conversion of glucose to p-hydroxystyrene. Metab Eng 9:268-276. http://dx.doi.org/10.1016/j.ymben.2007.01.002.
12. Verhoef S, Wierckx N, Westerhof RG, de Winde JH, Ruijssenaars HJ. 2009. Bioproduction of $p$-hydroxystyrene from glucose by the solventtolerant bacterium Pseudomonas putida S 12 in a two-phase water-decanol fermentation. Appl Environ Microbiol 75:931-936. http://dx.doi.org/10 .1128/AEM.02186-08.
13. Lin Y, Yan Y. 2012. Biosynthesis of caffeic acid in Escherichia coli using its endogenous hydroxylase complex. Microb Cell Fact 11:42. http://dx.doi .org/10.1186/1475-2859-11-42.
14. Appert C, Logemann E, Hahlbrock K, Schmid J, Amrhein N. 1994. Structural and catalytic properties of the four phenylalanine ammonia-lyase isoen-
zymes from parsley (Petroselinum crispum Nym.). Eur J Biochem 225:491499. http://dx.doi.org/10.1111/j.1432-1033.1994.00491.x.
15. Schuster B, Rétey J. 1995. The mechanism of action of phenylalanine ammonia-lyase: the role of prosthetic dehydroalanine. Proc Natl Acad Sci U S A 92:8433-8437. http://dx.doi.org/10.1073/pnas.92.18.8433.
16. Rösler J, Krekel F, Amrhein N, Schmid J. 1997. Maize phenylalanine ammonia-lyase has tyrosine ammonia-lyase activity. Plant Physiol 113: 175-179. http://dx.doi.org/10.1104/pp.113.1.175.
17. Gatenby AA, Sariaslani FS, Tang X, Qi WW, Vannelli T. April 2002. Bioproduction of para-hydroxycinnamic acid. US patent US6368837 Bl.
18. Kyndt JA, Meyer TE, Cusanovich MA, Van Beeumen JJ. 2002. Characterization of a bacterial tyrosine ammonia lyase, a biosynthetic enzyme for the photoactive yellow protein. FEBS Lett 512:240-244. http://dx.doi.org /10.1016/S0014-5793(02)02272-X.
19. Williams JS, Thomas M, Clarke DJ. 2005. The gene stlA encodes a phenylalanine ammonia-lyase that is involved in the production of a stilbene antibiotic in Photorhabdus luminescens TT01. Microbiology 151: 2543-2550. http://dx.doi.org/10.1099/mic.0.28136-0.
20. Xiang LK, Moore BS. 2005. Biochemical characterization of a prokaryotic phenylalanine ammonia lyase. J Bacteriol 187:4286-4289. http://dx.doi .org/10.1128/JB.187.12.4286-4289.2005.
21. Berner M, Krug D, Bihlmaier C, Vente A, Müller R, Bechthold A. 2006. Genes and enzymes involved in caffeic acid biosynthesis in the actinomycete Saccharothrix espanaensis. J Bacteriol 188:2666-2673. http://dx.doi .org/10.1128/JB.188.7.2666-2673.2006.
22. Louie GV, Bowman ME, Moffitt MC, Baiga TJ, Moore BS, Noel JP. 2006. Structural determinants and modulation of substrate specificity in phenylalanine-tyrosine ammonia-lyases. Chem Biol 13:1327-1338. http: //dx.doi.org/10.1016/j.chembiol.2006.11.011.
23. Watts KT, Mijts BN, Lee PC, Manning AJ, Schmidt-Dannert C. 2006. Discovery of a substrate selectivity switch in tyrosine ammonia-lyase, a member of the aromatic amino acid lyase family. Chem Biol 13:13171326. http://dx.doi.org/10.1016/j.chembiol.2006.10.008.
24. Moffitt MC, Louie GV, Bowman ME, Pence J, Noel JP, Moore BS. 2007. Discovery of two cyanobacterial phenylalanine ammonia lyases: kinetic and structural characterization. Biochemistry 46:1004-1012. http://dx .doi.org/10.1021/bi061774g.
25. Vannelli T, Xue Z, Breinig S, Qi WW, Sariaslani FS. 2007. Functional expression in Escherichia coli of the tyrosine-inducible tyrosine ammonialyase enzyme from yeast Trichosporon cutaneum for production of $p$ hydroxycinnamic acid. Enzyme Microb Technol 41:413-422. http://dx .doi.org/10.1016/j.enzmictec.2007.03.013.
26. Schroeder AC, Kumaran S, Hicks LM, Cahoon RE, Halls C, Yu O, Jez JM. 2008. Contributions of conserved serine and tyrosine residues to catalysis, ligand binding, and cofactor processing in the active site of tyrosine ammonia lyase. Phytochemistry 69:1496-1506. http://dx.doi.org /10.1016/j.phytochem.2008.02.007.
27. Wang L, Gamez A, Archer H, Abola EE, Sarkissian CN, Fitzpatrick P, Wendt D, Zhang Y, Vellard M, Bliesath J, Bell SM, Lemontt JF, Scriver CR, Stevens RC. 2008. Structural and biochemical characterization of the therapeutic Anabaena variabilis phenylalanine ammonia lyase. J Mol Biol 380:623-635. http://dx.doi.org/10.1016/j.jmb.2008.05.025.
28. Bartsch S, Bornscheuer UT. 2009. A single residue influences the reaction mechanism of ammonia lyases and mutases. Angew Chem Int Ed Engl 48:3362-3365. http://dx.doi.org/10.1002/anie.200900337.
29. Krug D, Müller R. 2009. Discovery of additional members of the tyrosine aminomutase enzyme family and the mutational analysis of CmdF. Chembiochem 10:741-750. http://dx.doi.org/10.1002/cbic. 200800748.
30. Bartsch S, Bornscheuer UT. 2010. Mutational analysis of phenylalanine ammonia lyase to improve reactions rates for various substrates. Protein Eng Des Sel 23:929-933. http://dx.doi.org/10.1093/protein/gzq089.
31. Hsieh LS, Yeh CS, Pan HC, Cheng CY, Yang CC, Lee PD. 2010. Cloning and expression of a phenylalanine ammonia-lyase gene (BoPAL2) from Bambusa oldhamii in Escherichia coli and Pichia pastoris. Protein Expr Purif 71:224-230. http://dx.doi.org/10.1016/j.pep.2010.01.009.
32. Zhu L, Cui W, Fang Y, Liu Y, Gao X, Zhou Z. 2013. Cloning, expression and characterization of phenylalanine ammonia-lyase from Rhodotorula glutinis. Biotechnol Lett 35:751-756. http://dx.doi.org/10.1007/s10529 -013-1140-7.
33. Christenson SD, Liu W, Toney MD, Shen B. 2003. A novel 4-methyl-ideneimidazole-5-one-containing tyrosine aminomutase in enediyne antitumor antibiotic C-1027 biosynthesis. J Am Chem Soc 125:6062-6063. http://dx.doi.org/10.1021/ja034609m.
34. Jin M, Fischbach MA, Clardy J. 2006. A biosynthetic gene cluster for the acetyl-CoA carboxylase inhibitor andrimid. J Am Chem Soc 128:1066010661. http://dx.doi.org/10.1021/ja063194c.
35. Rachid S, Krug D, Weissman KJ, Müller R. 2007. Biosynthesis of (R)-$\beta$-tyrosine and its incorporation into the highly cytotoxic chondramides produced by Chondromyces crocatus. J Biol Chem 282:21810-21817. http: //dx.doi.org/10.1074/jbc.M703439200.
36. Schwede TF, Rétey J, Schulz GE. 1999. Crystal structure of histidine ammo-nia-lyase revealing a novel polypeptide modification as the catalytic electrophile. Biochemistry 38:5355-5361. http://dx.doi.org/10.1021/bi982929q.
37. Calabrese JC, Jordan DB, Boodhoo A, Sariaslani S, Vannelli T. 2004. Crystal structure of phenylalanine ammonia lyase: multiple helix dipoles implicated in catalysis. Biochemistry 43:11403-11416. http://dx.doi.org /10.1021/bi049053+.
38. Ritter H, Schulz GE. 2004. Structural basis for the entrance into the phenylpropanoid metabolism catalyzed by phenylalanine ammonia-lyase. Plant Cell 16:3426-3436. http://dx.doi.org/10.1105/tpc.104.025288.
39. Huang L, Xue Z. June 2006. DNA and amino acid sequence of a tyrosine ammonia lyase enzyme from the bacterium Rhodobacter sphaeroides. US patent US7067302 B2.
40. Zhu Y, Liao S, Ye J, Zhang H. 2012. Cloning and characterization of a novel tyrosine ammonia lyase-encoding gene involved in bagremycins biosynthesis in Streptomyces sp. Biotechnol Lett 34:269-274. http://dx.doi .org/10.1007/s10529-011-0755-9.
41. Vaslet C, Strausberg R, Sykes A, Levy A, Filpula D. 1988. cDNA and genomic cloning of yeast phenylalanine ammonia-lyase genes reveal genomic intron deletions. Nucleic Acids Res 16:11382-11382. http://dx .doi.org/10.1093/nar/16.23.11382.
42. Filpula D, Vaslet CA, Levy A, Sykes A, Strausberg RL. 1988. Nucleotide sequence of gene for phenylalanine ammonia-lyase from Rhodotorula rubra. Nucleic Acids Res 16:11381. http://dx.doi.org/10.1093/nar/16.23 .11381.
43. Hwang EI, Kaneko M, Ohnishi Y, Horinouchi S. 2003. Production of plant-specific flavanones by Escherichia coli containing an artificial gene cluster. Appl Environ Microbiol 69:2699-2706. http://dx.doi.org/10.1128 /AEM.69.5.2699-2706.2003.
44. Huang Y, Niu B, Gao Y, Fu L, Li W. 2010. CD-HIT suite: a web server for clustering and comparing biological sequences. Bioinformatics 26: 680-682. http://dx.doi.org/10.1093/bioinformatics/btq003.
45. Grin I, Linke D. 2011. GCView: the genomic context viewer for protein homology searches. Nucleic Acids Res 39:W353-W356. http://dx.doi.org /10.1093/nar/gkr364.
46. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol 30: 2725-2729. http://dx.doi.org/10.1093/molbev/mst197.
47. Jones DT, Taylor WR, Thornton JM. 1992. The rapid generation of mutation data matrices from protein sequences. Comput Appl Biosci 8:275-282. http://dx.doi.org/10.1093/bioinformatics/8.3.275.
48. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins DG. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol 7:539. http://dx.doi.org/10 .1038/msb.2011.75.
49. Jensen NB, Strucko T, Kildegaard KR, David F, Maury J, Mortensen UH, Förster J, Nielsen J, Borodina I. 2014. EasyClone: method for iterative chromosomal integration of multiple genes in Saccharomyces cerevisiae. FEMS Yeast Res 14:238-248. http://dx.doi.org/10.1111/1567 -1364.12118.
50. Verduyn C, Postma E, Scheffers WA, Van Dijken JP. 1992. Effect of benzoic acid on metabolic fluxes in yeasts: a continuous-culture study on the regulation of respiration and alcoholic fermentation. Yeast 8:501-517. http://dx.doi.org/10.1002/yea.320080703.
51. Poolman B, Konings WN. 1988. Relation of growth of Streptococcus lactis and Streptococcus cremoris to amino acid transport. J Bacteriol 170:700707.
52. Sambrook J, Fritsch EF, Maniatis T. 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
53. Birnboim HC. 1983. A rapid alkaline extraction method for the isolation of plasmid DNA. Methods Enzymol 100:243-255. http://dx.doi.org/10 .1016/0076-6879(83)00059-2.
54. Geu-Flores F, Nour-Eldin HH, Nielsen MT, Halkier BA. 2007. USER fusion: a rapid and efficient method for simultaneous fusion and cloning
of multiple PCR products. Nucleic Acids Res 35:e55. http://dx.doi.org/10 .1093/nar/gkm106.
55. Gietz RD, Schiestl RH. 2007. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. Nat Protoc 2:31-34. http://dx.doi .org/10.1038/nprot.2007.13.
56. Kuipers OP, de Ruyter PGGA, Kleerebezem M, de Vos WM. 1998. Quorum sensing-controlled gene expression in lactic acid bacteria. J Biotechnol 64:15-21. http://dx.doi.org/10.1016/S0168-1656(98)00100-X.
57. Holo H, Nes IF. 1995. Transformation of Lactococcus by electroporation. Methods Mol Biol 47:195-199. http://dx.doi.org/10.1385/0-89603-310 -4:195.
58. Sauer U, Lasko DR, Fiaux J, Hochuli M, Glaser R, Szyperski T, Wuthrich K, Bailey JE. 1999. Metabolic flux ratio analysis of genetic and environmental modulations of Escherichia coli central carbon metabolism. J Bacteriol 181:6679-6688.
59. Gaspar P, Neves AR, Ramos A, Gasson MJ, Shearman CA, Santos H. 2004. Engineering Lactococcus lactis for production of mannitol: high yields from food-grade strains deficient in lactate dehydrogenase and the mannitol transport system. Appl Environ Microbiol 70:1466-1474. http: //dx.doi.org/10.1128/AEM.70.3.1466-1474.2004.
60. Xue Z, McCluskey M, Cantera K, Sariaslani FS, Huang L. 2007. Identification, characterization and functional expression of a tyrosine ammo-nia-lyase and its mutants from the photosynthetic bacterium Rhodobacter sphaeroides. J Ind Microbiol Biotechnol 34:599-604. http://dx.doi.org/10 .1007/s10295-007-0229-1.
61. Hodgins DS. 1971. Yeast phenylalanine ammonia-lyase. Purification, properties, and the identification of catalytically essential dehydroalanine. J Biol Chem 246:2977-2985.
62. Anson JG, Gilbert HJ, Oram JD, Minton NP. 1987. Complete nucleotide sequence of the Rhodosporidium toruloides gene coding for phenylalanine ammonia-lyase. Gene 58:189-199. http://dx.doi.org/10.1016/0378-1119 (87)90375-1.
63. Strobel T, Al-Dilaimi A, Blom J, Gessner A, Kalinowski J, Luzhetska M, Pühler A, Szczepanowski R, Bechthold A, Rückert C. 2012. Complete genome sequence of Saccharothrix espanaensis DSM $44229^{\mathrm{T}}$ and comparison to the other completely sequenced Pseudonocardiaceae. BMC Genomics 13:465. http://dx.doi.org/10.1186/1471-2164-13-465.
64. Schomburg I, Chang A, Placzek S, Söhngen C, Rother M, Lang M, Munaretto C, Ulas S, Stelzer M, Grote A, Scheer M, Schomburg D. 2013. BRENDA in 2013: integrated reactions, kinetic data, enzyme function data, improved disease classification: new options and contents in BRENDA. Nucleic Acids Res 41:D764-D772. http://dx.doi.org/10.1093 /nar/gks1049.
65. Kyndt JA, Vanrobaeys F, Fitch JC, Devreese BV, Meyer TE, Cusanovich MA, Van Beeumen JJ. 2003. Heterologous production of Halorhodospira halophila holo-photoactive yellow protein through tandem expression of the postulated biosynthetic genes. Biochemistry 42:965-970. http://dx.doi .org/10.1021/bi027037b.
66. Memmi S, Kyndt J, Meyer T, Devreese B, Cusanovich M, Van Beeumen J. 2008. Photoactive yellow protein from the halophilic bacterium Salinibacter ruber. Biochemistry 47:2014-2024. http://dx.doi.org/10.1021/bi701486n.
67. Röther D, Poppe L, Viergutz S, Langer B, Rétey J. 2001. Characterization of the active site of histidine ammonia-lyase from Pseudomonas putida. Eur J Biochem 268:6011-6019. http://dx.doi.org/10.1046/j. 0014 -2956.2001.02298.x.
68. Röther D, Poppe L, Morlock G, Viergutz S, Rétey J. 2002. An active site homology model of phenylalanine ammonia-lyase from Petroselinum crispum. Eur J Biochem 269:3065-3075. http://dx.doi.org/10.1046/j. 1432 -1033.2002.02984.x.
69. Kawai Y, Noda S, Ogino C, Takeshima Y, Okai N, Tanaka T, Kondo A. 2013. p-Hydroxycinnamic acid production directly from cellulose using endoglucanase- and tyrosine ammonia lyase-expressing Streptomyces lividans. Microb Cell Fact 12:45. http://dx.doi.org/10.1186/1475-2859-12-45.
70. Xue Z, McCluskey M, Cantera K, Ben-Bassat A, Sariaslani RS, Huang L. 2007. Improved production of $p$-hydroxycinnamic acid from tyrosine using a novel thermostable phenylalanine/tyrosine ammonia lyase enzyme. Enzyme Microb Technol 42:58-64. http://dx.doi.org/10.1016/j .enzmictec.2007.07.025.
71. Tang X. February 2003. Polynucleotide encoding a mutant Rhodotorula glutinis tyrosine ammonia lyase polypeptide. US patent US6521748 B2.
72. Langer M, Reck G, Reed J, Rétey J. 1994. Identification of serine-143 as the most likely precursor of dehydroalanine in the active site of histidine ammo-nia-lyase. A study of the overexpressed enzyme by site-directed mutagenesis. Biochemistry 33:6462-6467. http://dx.doi.org/10.1021/bi00187a011.
73. Pilbák S, Farkas O, Poppe L. 2012. Mechanism of the tyrosine ammonia lyase reaction-tandem nucleophilic and electrophilic enhancement by a proton transfer. Chemistry 18:7793-7802. http://dx.doi.org/10.1002/chem . 201103662.
74. Reichenbach H, Kohl W, Böttger-Vetter A, Achenbach H. 1980. Flex-irubin-type pigments in Flavobacterium. Arch Microbiol 126:291-293. http://dx.doi.org/10.1007/BF00409934.
75. McBride MJ, Xie G, Martens EC, Lapidus A, Henrissat B, Rhodes RG, Goltsman E, Wang W, Xu J, Hunnicutt DW, Staroscik AM, Hoover TR, Cheng YQ, Stein JL. 2009. Novel features of the polysaccharide-digesting gliding bacterium Flavobacterium johnsoniae as revealed by genome sequence analysis. Appl Environ Microbiol 75:6864-6875. http://dx.doi.org /10.1128/AEM.01495-09.
76. Kiss H, Nett M, Domin N, Martin K, Maresca JA, Copeland A, Lapidus A, Lucas S, Berry KW, Glavina Del Rio T, Dalin E, Tice H, Pitluck S, Richardson P, Bruce D, Goodwin L, Han C, Detter JC, Schmutz J, Brettin T, Land M, Hauser L, Kyrpides NC, Ivanova N, Göker M, Woyke T, Klenk HP, Bryant DA. 2011. Complete genome sequence of the filamentous gliding predatory bacterium Herpetosiphon aurantiacus type strain $\left(114-95^{\mathrm{T}}\right)$. Stand Genomic Sci 5:356-370. http://dx.doi.org/10 .4056/sigs. 2194987.
77. Eichinger L, Pachebat JA, Glöckner G, Rajandream MA, Sucgang R, Berriman M, Song J, Olsen R, Szafranski K, Xu Q, Tunggal B, Kummerfeld S, Madera M, Konfortov BA, Rivero F, Bankier AT, Lehmann R, Hamlin N, Davies R, Gaudet P, Fey P, Pilcher K, Chen G, Saunders D, Sodergren E, Davis P, Kerhornou A, Nie X, Hall N, Anjard C, Hemphill L, Bason N, Farbrother P, Desany B, Just E, Morio T, Rost R, Churcher C, Cooper J, Haydock S, van Driessche N, Cronin A, Goodhead I, Muzny D, Mourier T, Pain A, Lu M, Harper D, Lindsay R, Hauser H, et al. 2005. The genome of the social amoeba Dictyostelium discoideum. Nature 435:43-57. http://dx.doi.org/10.1038/nature03481.
78. Kovács K, Bánóczi G, Varga A, Szabó I, Holczinger A, Hornyánszky G, Zagyva I, Paizs C, Vértessy BG, Poppe L. 2014. Expression and properties of the highly alkalophilic phenylalanine ammonia-lyase of thermophilic Rubrobacter xylanophilus. PLoS One 9:e85943. http://dx.doi.org/10 .1371/journal.pone. 0085943.
79. Christenson SD, Wu W, Spies MA, Shen B, Toney MD. 2003. Kinetic analysis of the 4-methylideneimidazole-5-one-containing tyrosine aminomutase in enediyne antitumor antibiotic C-1027 biosynthesis. Biochemistry 42:12708-12718. http://dx.doi.org/10.1021/bi035223r.
80. Hu GS, Jia JM, Hur YJ, Chung YS, Lee JH, Yun DJ, Chung WS, Yi GH, Kim TH, Kim DH. 2011. Molecular characterization of phenylalanine ammonia lyase gene from Cistanche deserticola. Mol Biol Rep 38:37413750. http://dx.doi.org/10.1007/s11033-010-0489-0.
81. Gao ZM, Wang XC, Peng ZH, Zheng B, Liu Q. 2012. Characterization and primary functional analysis of phenylalanine ammonia-lyase gene from Phyllostachys edulis. Plant Cell Rep 31:1345-1356. http://dx.doi.org /10.1007/s00299-012-1253-9.
82. Ben-Bassat A, Sariaslani FS, Huang L, Patnaik R, Lowe DJ. August 2011. Methods for the preparation of para-hydroxycinnamic acid and cinnamic acid at alkaline pH. US patent US8003356 B2.
83. Borodina I, Nielsen J. 2014. Advances in metabolic engineering of yeast Saccharomyces cerevisiae for production of chemicals. Biotechnol J 9:609620. http://dx.doi.org/10.1002/biot. 201300445.
84. Shin SY, Jung SM, Kim MD, Han NS, Seo JH. 2012. Production of resveratrol from tyrosine in metabolically engineered Saccharomyces cerevisiae. Enzyme Microb Technol 51:211-216. http://dx.doi.org/10 .1016/j.enzmictec.2012.06.005.
85. Jiang H, Wood KV, Morgan JA. 2005. Metabolic engineering of the phenylpropanoid pathway in Saccharomyces cerevisiae. Appl Environ Microbiol 71:2962-2969. http://dx.doi.org/10.1128/AEM.71.6.2962-2969.2005.
86. Gaspar P, Vieira A, Kuipers OP, Förster J, Neves AR. 2013. Engineering Lactococcus lactis for the production of plant polyphenols. In Abstr BioMicroWorld 2013: V Int Conf Environ Ind Appl Microbiol, Madrid, Spain, 2 to 4 October 2013.

[^0]:    General rights
    Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

    - Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
    - You may not further distribute the material or use it for any profit-making activity or commercial gain
    - You may freely distribute the URL identifying the publication in the public portal

[^1]:    Received 7 February 2015 Accepted 20 April 2015
    Accepted manuscript posted online 24 April 2015
    Citation Jendresen CB, Stahlhut SG, Li M, Gaspar P, Siedler S, Förster J, Maury J, Borodina I, Nielsen AT. 2015. Highly active and specific tyrosine ammonia-lyases from diverse origins enable enhanced production of aromatic compounds in bacteria and Saccharomyces cerevisiae. Appl Environ Microbiol 81:4458-4476. doi:10.1128/AEM.00405-15.
    Editor: M. J. Pettinari
    Address correspondence to Christian Bille Jendresen, cbj@biosustain.dtu.dk, or Alex Toftgaard Nielsen, atn@biosustain.dtu.dk.
    C.B.J. and S.G.S. contributed equally to this work.

    Supplemental material for this article may be found at http://dx.doi.org/10.1128 /AEM.00405-15
    Copyright © 2015, American Society for Microbiology. All Rights Reserved. doi:10.1128/AEM.00405-15

[^2]:    ${ }^{a}$ Protein length in amino acids.
    ${ }^{b}$ Previously characterized enzyme.
    ${ }^{c}$ Novel enzyme included in this study.

