nature portfolio

Peer Review File

Combined biosynthesis and site-specific incorporation of phenylalanine derivatives from aryl aldehydes or carboxylic acids

Corresponding Author: Professor Aditya Kunjapur

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
Review Comments
1.Research Contributions
The study proposes an engineered bacterial platform for the biosynthesis and site-specific incorporation of non-standard amino acids (nsAAs). Key contributions include: 1.Utilizing synthetic biology approaches to enable E. coli to directly synthesize ncAAs from aromatic aldehydes or carboxylic acids. 2.Successfully incorporating novel ncAAs such as 4,4-L-biphenylalanine and 4-azido-L-phenylalanine, with validation of high-fidelity protein translation.
2.Significance and Limitations
The study presents methodological innovations and offers new insights into genetic code expansion. However, the following limitations exist: 1. Insufficient substrate scope: The work primarily focuses on aromatic aldehyde substrates, with limited exploration of carboxylic acid compounds as precursors. Carboxylic acids are generally more stable, readily available, and have higher practical value. 2.Inadequate testing of substrate broad-spectrum compatibility: The study does not fully investigate the synthetic pathway's tolerance to aromatic carboxylic acids with different substituents (e.g., p-methyl-, p-fluoro-, p-methoxy-substituted benzoic acids etc) or product yields. Expand substrate diversity and provide detailed conversion rates and product analyses.
3.Data Support for Conclusions
The manuscript provides sufficient experimental data to support partial conclusions, including synthesis and site-specific incorporation of different ncAAs. Data confirm that engineered bacteria can synthesize target ncAAs, with incorporation verified via fluorescent labeling and mass spectrometry. However, data on substrate adaptability, enzymatic efficiency, and potential byproducts are limited.
4.Methodological Reliability and Reproducibility
Experimental methods align with bioengineering standards, employing multi-enzyme cascade reactions successfully implemented in E. coli. Methodological descriptions can be improved, particularly regarding substrate optimization and biosynthesis pathway selection. Clarify substrate selection criteria: Why were the tested aromatic aldehydes/carboxylic acids chosen? Are other aromatic carboxylic acids potentially applicable? Optimize key enzymes in the biosynthesis pathway: How do different substrates affect enzymatic efficiency? Is directed evolution needed to enhance activity? Include reaction kinetic data (e.g., Km, Vmax) to ensure methodological rigor.

The manuscript is of high overall quality with reasonable study design and data-supported conclusions. However, further experimental optimization is required, especially for broad-spectrum carboxylic acid testing (The use of aromatic aldehydes is relatively conventional.). It is expected to have a positive impact on industrial protein manufacturing and synthetic biology applications. I recommend acceptance after major revisions.

Reviewer #2

(Remarks to the Author)
I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career

Researchers who co-review manuscripts.

Reviewer #3

(Remarks to the Author)
The manuscript titled “Combined biosynthesis and site-specific incorporation of phenylalanine derivatives from aryl aldehydes or carboxylic acids in engineered bacteria” by Anderson et al. describe the development of a single engineered bacterial host capable of carrying out both the semi-synthesis of phenylalanine derivatives and their site-specific incorporation into target proteins. This platform pathway broadly accommodates achiral aryl aldehydes and carboxylic acids, generating nsAAs widely used in genetic code expansion. However, the manuscript in its current form is too preliminary for publication in Nature Communications. The use of a tandem enzyme system comprising SrCAR, TTA (s-ObiH), PSDH, and TyrB for nsAA production is not novel. This approach has been previously reported in ChemCatChem in the article titled 'A One-Pot Biocatalytic Cascade to Access L-Phenylalanine Derivatives from Aldehydes or Carboxylic Acids” doi: 10.1101/2024.12.06.627276 Although previous biosynthesis platforms have been limited to a narrow range of substrates, the platform in this study also has its own limitations in recognizing the substrates. Because the authors only demonstrated the complete biosynthesis and protein incorporation of only 1e and 4e. Additionally, the screening and selection of enzymes seems somewhat disorganized, and their introduction complicates the system, potentially hindering future development. The authors should consider testing more diverse substrates to show the platform’s efficacy. Additionally, to fully demonstrate the platform's potential, it would be essential to showcase one or two practical applications beyond the synthetic auxotrophy example. And the authors should consider the demonstration of the yield of GFP incorporated with nsAA in biosynthesis group the direct feeding group. Otherwise, how do we know the efficacy of this platform combing the complicate biosynthesis pathway and GCE. Once these additional experiments and revisions are incorporated, the manuscript will be more likely to meet the high standards of Nature Communications.

I have some other comments to further improve this manuscript:
1. In Figure 1, Figure 2b and so on, what do AKRs and ADHs stand for? The full terms should be provided the first time they appear in the article.
2. In Figure 2, the legend refers to L-TTA, but the figure shows TTA. The naming should be consistent.
3. In Figure 2, although s-TTA refers to L-TTA with a SUMO tag, it should be written as "SUMO-tagged L-TTA (s-TTA)" the first time s-TTA appears in the article.
4. In Figure 2c, why is the conversion in the in vivo group much lower than that of the purified enzyme? Also, what do the new peaks in the HPLC trace represent? You might consider adding p-values to the bar chart to indicate the statistical significance of the differences.
5. In Figure 2d, the Y-axis should represent concentration, as in other figures, rather than peak area.
6. In Figure 3a and the text, you should provide the full names of PSDH and AT the first time they appear in the article.
7. Are the concentrations of nsAAs in the bar chart intracellular or extracellular? In additionl I didn’t see a standard curve for the compounds in this article; how were their concentrations determined?
8. In Figure 3b, the product names are not labeled, but they are listed as 1e to 4e in Figure 3c. I suggest mentioning the names of the corresponding products on the Y-axis, such as "concentration of products (e)."
9. In Figure S2, which antibody was used for the Western blot? I didn’t find any information about this. It doesn’t appear to be an antibody against SUMO, as the band for the group without the SUMO tag is present.
10. Why didn’t you test the aldehyde stability of 4b, only 1b-3b (Figures S4-S6)?
11. Why didn’t you test s-ObiH in different strains in Figure 3c? It showed better conversion than s-RpPSDH in Figure 3b. Additionally, PbTTA demonstrated better conversion for 1b-3b, but s-ObiH seemed to perform better for 2b-4b in Figure 3b when coupled with PSDH and AT. It will be better for giving more explanation for this part.
12. SrCAR shows better results in Figure 4a, but why did you choose MavCAR for the following assays in Figure 4b? The screening and selection of relevant enzymes throughout the article need more logical consistency.
13. In line 183, you mention “1 mM of 4a (due to lower solubility) to RARE.Δ16,” but in the following study (Figure 4b), you used a higher concentration of 4 mM, which showed better results than 1 mM. This suggests that higher concentrations might be more effective at the beginning.
14. In line 218, you mention “frame UAG codon targeted for suppression (Fig. 5a),” but it is actually a TAG codon in Figure 5a. The description should be consistent, and the specific position of the TAG codon should be mentioned in the GFP reporter.
15. The SDS-PAGE gel, the mass spectrum of full GFP, and its yield are missing in the GCE section.
16. Where is the "d" part in the legend of Figure 6? It seems to be missing or not labeled.

Reviewer #4

(Remarks to the Author)
Summary:
This study presents an innovative engineered bacterial platform that combines biosynthesis and site-specific incorporation of phenylalanine-derived non-standard amino acids (nsAAs) from aryl aldehyde or carboxylic acid precursors. The authors demonstrate the design of a multi-enzyme cascade in E. coli to bypass the traditional reliance on nsAA supplementation, enabling direct conversion of simple, achiral precursors into functionalized nsAAs. Key achievements include the successful coupling of biosynthesis with orthogonal translation systems (OTS) for high-fidelity incorporation of nsAAs into proteins and the extension of synthetic auxotrophy to aldehyde-dependent biocontainment. The work addresses critical limitations in scaling genetic code expansion for industrial and environmental applications.
Conclusion:

This manuscript presents a powerful platform for cost-effective nsAA production and incorporation by bridging metabolic engineering and genetic code expansion, with transformative potential for synthetic biology and biotechnology. While technical strengths are evident, addressing the major concerns—particularly host robustness and quantitative analysis of escape frequency—would solidify the work's impact. With appropriate revisions, this study is suited for publication in Nature Communications.

Major Concerns:
1. Host strain stability and robustness: Limited data are provided on the long-term stability of engineered strains under fermentation conditions. Experiments assessing growth performance of many constructed strains (e.g., RARE.Δ16 strain expressing a cascade of enzymes for the biosynthesis of phenylalanine derivates or the final biocontainment strain derived from DEP.e5) in bioreactors would strengthen industrial relevance.
2. Quantitative analysis of escape frequency: While synthetic auxotrophy is achieved in Fig.6, quantitative analysis of escape frequency for DEP.e5 transformed with biosynthesis pathway is not addressed. A side-by-side comparison for escape frequency between DEP.e5 (requires 4e) and DEP.e5 containing the biosynthesis pathway (requires 4b) would enhance the manuscript.
3. Incomplete characterization of byproducts: For carboxylic acid pathways, minor azide reduction to 4-amino-phenylalanine (pAF) was observed. A deeper analysis of side reactions and strategies to mitigate unwanted modifications is warranted.
4. Substrate Scope and Pathway Efficiency: While the platform claims broad substrate specificity, data for bulky substrates (e.g., 4b/4a) show low titers (e.g., 5 μM 4e). The feasibility of scaling these yields for industrial applications remains unclear. Additional optimization (e.g., enzyme engineering or OTS evolution) should be discussed.

Minor Comments:
1. Fig. 2C: Labeling of HPLC traces (e.g., "in vivo s-ObiH + 1 mM (b)") is ambiguous. Specify which substrate corresponds to each trace.
2. Fig. 6: Simplify schematic annotations (e.g., "GFP fragment composition") to improve readability. The figure legend for panel d is missing.
3. Clarify the rationale for using MOPS EZ Rich media with 2% glucose in biosynthesis assays (line 113)—does this medium optimize precursor uptake or enzyme activity?

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The revised manuscript has made substantial progress in the field of genetic code expansion and synthetic biology. The authors have developed a reliable and broadly applicable bacterial platform that enables in vivo biosynthesis and site-specific incorporation of ncAAs, reflecting important progress in integrating metabolic engineering with translation regulation. The authors have thoroughly addressed the previous review comments, significantly expanded the substrate scope, and clarified the research mechanism through detail analysis. The experimental data are sufficient and well-supported, and the overall quality and academic value of the work have been improved. I recommend acceptance for publication.

Reviewer #2

(Remarks to the Author)
I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Reviewer #3

(Remarks to the Author)
The authors have provided thorough and satisfactory responses to my previous comments and have appropriately revised the manuscript. I have no further suggestions for improvement and find the work to be noteworthy, significant, and suitable for publication in Nature Communications.

Reviewer #4

(Remarks to the Author)
The authors have satisfactorily addressed each of the comments raised during the review process.

Point-by-point responses to reviewers:

We thank the reviewers for their valuable comments that have strengthened our paper. We found their analysis of our manuscript to be fair, rigorous, and helpful. Our responses to reviewer feedback are colored in blue, and modifications to the manuscript text included below are colored in red.

Reviewer #1 (Remarks to the Author):

Review Comments

1.Research Contributions
The study proposes an engineered bacterial platform for the biosynthesis and site-specific incorporation of non-standard amino acids (nsAAs). Key contributions include: 1.Utilizing synthetic biology approaches to enable E. coli to directly synthesize ncAAs from aromatic aldehydes or carboxylic acids. 2.Successfully incorporating novel ncAAs such as 4,4-L-biphenylalanine and 4-azido-L-phenylalanine, with validation of high-fidelity protein translation.

2.Significance and Limitations
The study presents methodological innovations and offers new insights into genetic code expansion. However, the following limitations exist:
1. Insufficient substrate scope: The work primarily focuses on aromatic aldehyde substrates, with limited exploration of carboxylic acid compounds as precursors. Carboxylic acids are generally more stable, readily available, and have higher practical value.
2.Inadequate testing of substrate broad-spectrum compatibility: The study does not fully investigate the synthetic pathway's tolerance to aromatic carboxylic acids with different substituents (e.g., p-methyl-, p-fluoro-, p-methoxy-substituted benzoic acids etc) or product yields. Expand substrate diversity and provide detailed conversion rates and product analyses.

We agreed that investigation of additional substrates would improve the manuscript. At the time of initial manuscript submission, our combined system for biosynthesis and incorporation exhibited some imbalance and burden. We have now improved relative heterologous gene expression and expanded our substrate scope to include 14 new chemistries, reaching a total of 18 chemistries investigated in the paper (see Fig. 4). We define a "chemistry" here as a unique R group – this means the actual number of substrates we test is greater than 18 because for many R groups we report both the aldehyde and the carboxylic acid as successful substrates.

To further highlight the value of our combined nsAA semi-synthesis and incorporation system while also addressing opportunities to show broader substrate scope, we have demonstrated the synthesis of 6 nsAAs that are not commercially available using precursors that are commercially available. Finally, we scaled up the biosynthesis and incorporation of one of these 6 nsAAs from 300 µL in 96-well plate format to 50 mL culture volume in 250 mL flask format for investigation of the purified reporter protein that contains the incorporated nsAA. We obtained a good protein yield of 51 mg/L using our combined system, demonstrating its practical value (Fig. 4g).

These new results strengthen the manuscript and should address most of the elements of the reviewer's comment. Rather than directly measure and report conversion rates, especially now that there are so many unique substrates, to help audiences understand what biosynthetic steps appear to be limiting in the cellular context we now report observed intermediates and byproducts for each of our 4 initially tested chemistries. We have added Results and associated text that discusses why we believe the TTA reaction was the bottleneck for chemistries 1-3 (see lines 176-193

in our revised manuscript – also copied below for convenience - and Supplemental Figures 9-13).

"To gain insight into what step might be limiting nsAA titers, we analyzed expected intermediates and potential byproducts. Although our lack of standards for the β-OH nsAA and keto acid intermediates limited precise quantification of their concentration, we observed that the remaining peak area of the β-OH nsAA intermediate upon expression of the entire biosynthetic pathway was less than 5% that seen upon expressing only a TTA (**Fig. S9**). Additionally, based on comparable UV-absorbance of a few available β-OH nsAAs and their corresponding L-Phe derivatives, we estimate that for each of the four chemistries less than 100 μM β-OH nsAA accumulated as detected in the culture media (**Table S6**). Collectively, these results strongly suggest that the PSDH is not the limiting step. To better understand whether the limiting step was upstream or downstream of this step, we next quantified the remaining aldehyde for each of the four chemistries, as well as the alcohol and carboxylic acid byproducts due to oxidation or reduction in the cellular environment (**Fig. S10-S13**). The sum of the remaining reactant aldehyde and its associated redox byproducts in the RARE.Δ16 host amounts to 29.6%, 30.3%, 55.6%, and 11.7% of the supplemented 1 mM **1b**, **2b**, **3b**, and **4b**, respectively. For chemistries **1-3**, this indicates that the TTA-catalyzed reaction is the limiting step when the full biosynthetic pathway is expressed in live cells. For chemistry **4**, due to an incomplete mass balance and high hydrophobicity of pathway compounds, it is less clear if any individual enzymatic reaction, low solubility, or membrane association could be limiting."

3.Data Support for Conclusions
The manuscript provides sufficient experimental data to support partial conclusions, including synthesis and site-specific incorporation of different ncAAs. Data confirm that engineered bacteria can synthesize target ncAAs, with incorporation verified via fluorescent labeling and mass spectrometry. However, data on substrate adaptability, enzymatic efficiency, and potential byproducts are limited.

As described in the previous comment, we have investigated additional substrates for compatibility with the cascade. We thank the reviewer for mentioning the opportunity to highlight potential byproducts. This data is now included in Supplemental Figures 9-13.

4.Methodological Reliability and Reproducibility
Experimental methods align with bioengineering standards, employing multi-enzyme cascade reactions successfully implemented in E. coli. Methodological descriptions can be improved, particularly regarding substrate optimization and biosynthesis pathway selection. Clarify substrate selection criteria: Why were the tested aromatic aldehydes/carboxylic acids chosen?

The reviewer has good rationale to question the initial substrate selection criteria when only 4 chemistries were tested, though we attempted to explain that the R groups of the 4 were quite distinct along axes of sterics and electrostatics. However, now that our revised manuscript reports many more chemistries, we still include brief

justification of the initial 4 model chemistries as corresponding to nsAAs with established uses while exhibiting varying electrostatics and size for prototyping, but our revised manuscript now places less emphasis on these 4 chemistries and any restricted substrate selection criteria.

Are other aromatic carboxylic acids potentially applicable?

As mentioned in a previous comment, we have expanded the screen to show that other carboxylic acids (and aldehydes) are applicable.

Optimize key enzymes in the biosynthesis pathway: How do different substrates affect enzymatic efficiency? Is directed evolution needed to enhance activity? Include reaction kinetic data (e.g., Km, Vmax) to ensure methodological rigor.

Regarding optimization, our originally submitted manuscript reported the screening of 6 different TTA variants and 6 different CAR variants, identifying overall polyspecific variants for broad screening purposes. Experiments conducted during manuscript revision indicate that it was more important to optimize heterologous gene expression. We now show that the variants chosen here extend to exhibit activity on a broad range of chemistries in the cellular context.

The questions about directed evolution and reaction kinetic data are good questions that, if we may paraphrase, seem to be about identifying and reporting the bottleneck step(s), which can vary depending on the chemistry. We have chosen to approach this by investigating additional substrates and by including data on intermediate/byproduct formation for the initial 4 chemistries used for prototyping (particularly the peak areas of the intermediate beta-hydroxylated phenylalanines and concentrations of aldehyde redox products). We think this data better achieves understanding of the rate limiting steps in the current cellular context, and it complements our prior *in vitro* analyses of reactions featuring TTA homologs on their own or in the context of a phenylserine dehydratase in previously published work. Among a series of differences, key differences in the cellular context where biochemical parameters may not be as helpful for determination of how to improve the system are the abundance of co-substrates, co-factors, and heterologous enzymes. Collectively, these new analyses and discussion should more directly help accomplish the reviewer's interest in providing more detailed characterization and knowledge of what step to address next for future improvement of the system (as well as the improvement in heterologous expression already made during revision).

The manuscript is of high overall quality with reasonable study design and data-supported conclusions. However, further experimental optimization is required, especially for broad-spectrum carboxylic acid testing (The use of aromatic aldehydes is relatively conventional.). It is expected to have a positive impact on industrial protein manufacturing and synthetic biology applications. I recommend acceptance after major revisions.

We are grateful for the thoughtful suggestions and positive summary from this reviewer.

Reviewer #2 (Remarks to the Author):

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Thank you for your initiative to contribute to peer-review and gain important skills as a trainee.

Reviewer #3 (Remarks to the Author):

The manuscript titled “Combined biosynthesis and site-specific incorporation of phenylalanine derivatives from aryl aldehydes or carboxylic acids in engineered bacteria” by Anderson et al. describe the development of a single engineered bacterial host capable of carrying out both the semi-synthesis of phenylalanine derivatives and their site-specific incorporation into target proteins. This platform pathway broadly accommodates achiral aryl aldehydes and carboxylic acids, generating nsAAs widely used in genetic code expansion. However, the manuscript in its current form is too preliminary for publication in Nature Communications. The use of a tandem enzyme system comprising SrCAR, TTA (s-ObiH), PSDH, and TyrB for nsAA production is not novel. This approach has been previously reported in ChemCatChem in the article titled 'A One-Pot Biocatalytic Cascade to Access L-Phenylalanine Derivatives from Aldehydes or Carboxylic Acids” doi: 10.1101/2024.12.06.627276

In our revised manuscript, we now more clearly articulate in the Introduction (lines 87-92) how the nature of the system and its performance, including what steps are the bottlenecks, change substantially when we investigate the semi-synthesis pathway in a cellular context and when we couple genetic code expansion. Thus, our earlier publication about the tandem enzyme system in a cell-free environment should not interfere with the novelty of the results in a cellular system.

“The performance of a semi-synthesis pathway in a live cell environment can be met with numerous additional challenges relative to a cell-free environment due to differences in relative enzyme abundance, availability of co-factors and co-substrates, transport and toxicity of pathway metabolites, and competing reactions. Given also that our prior cell-free enzyme cascade did not include the step of GCE, the system that we sought to create in this manuscript is rather distinct.”

Although previous biosynthesis platforms have been limited to a narrow range of substrates, the platform in this study also has its own limitations in recognizing the substrates. Because the authors only demonstrated the complete biosynthesis and protein incorporation of only 1e and 4e. Additionally, the screening and selection of enzymes seems somewhat disorganized, and their introduction complicates the system, potentially hindering future development. The authors should consider testing more diverse substrates to show the platform’s efficacy.

We agreed that investigation of additional substrates would improve the manuscript. At the time of initial manuscript submission, our combined system for biosynthesis and incorporation exhibited some imbalance and burden. We have now improved relative heterologous gene expression and expanded our substrate scope to include 14 new chemistries, reaching a total of 18 chemistries investigated in the paper (see Fig. 4). We define a “chemistry” here as a unique R group – this means the actual number of substrates we test is greater than 18 because for many R groups we report both the aldehyde and the carboxylic acid as successful substrates.

Additionally, to fully demonstrate the platform's potential, it would be essential to showcase one or two practical applications beyond the synthetic auxotrophy example. And the authors should consider the demonstration of the yield of GFP incorporated with nsAA in biosynthesis group the direct feeding group. Otherwise, how do we know

the efficacy of this platform combing the complicate biosynthesis pathway and GCE. Once these additional experiments and revisions are incorporated, the manuscript will be more likely to meet the high standards of Nature Communications.

We thank the review for their thoughtful suggestions and have attempt to implement their proposed improvements. We have expanded testing to 14 diverse substrates, all of which we confirmed nsAA biosynthesis via MS. Additionally, we were able to utilize the platform to begin to screen for and identify compatible orthogonal translation machinery for incorporation of biosynthesized nsAAs. See our Figure 4 and associated Results text (lines 315-382). Given the length of the new additions, we have not included the edited text below and instead kindly ask the reviewer to look at the revised manuscript.

Beyond the synthetic auxotrophy example, we further show a practical application of the system within this section: Access to nsAAs that are not commercially available or that are cost-prohibitive if obtained through custom chemical synthesis. We foresee this system aiding in making genetic code expansion technologies economically feasible for industrial-scale biomanufacturing.

To help address the reviewer's comment about judging the efficacy of the combined system versus traditional supplementation of nsAAs, we now report two kinds of data: (1) Yields of purified GFP proteins where combined nsAA biosynthesis and incorporation was used for 3 different chemistries – including one of the non-commercially available chemistries (Figs. 3h, 3j, and 4g). Our yields are reasonably good when considering prior work in the GCE field – they range from 22-51 mg/L after we optimized heterologous expression. (2) Dose-dependent comparison of how exogenous supplementation of nsAAs compares with biosynthesis for 2 chemistries (Figs. 3g and 3i).

I have some other comments to further improve this manuscript:
1. In Figure 1, Figure 2b and so on, what do AKRs and ADHs stand for? The full terms should be provided the first time they appear in the article.

We now include the definitions for AKRs (aldo-keto reductases) and ADHs (alcohol dehydrogenases) both in the caption for Figure 1 as well as for the first mention of terms in the main text (line 123).

2. In Figure 2, the legend refers to L-TTA, but the figure shows TTA. The naming should be consistent.

At the first mention of L-threonine transaldolases, we now denote the equality of L-TTA and TTA (line 82) and changed all subsequent mention to TTA for consistency with the nomenclature of a SUMO-tagged TTA (s-TTA).

3. In Figure 2, although s-TTA refers to L-TTA with a SUMO tag, it should be written as "SUMO-tagged L-TTA (s-TTA)" the first time s-TTA appears in the article.

We have transitioned to the primary use of TTA (rather than L-TTA) for clarity, as mentioned above, and therefore include the introduction of the SUMO-tagged variants as "SUMO-tagged variant of the TTA (s-TTA)" (line 115).

4. In Figure 2c, why is the conversion in the in vivo group much lower than that of the purified enzyme? Also, what do the new peaks in the HPLC trace represent? You might consider adding p-values to the bar chart to indicate the statistical significance of the differences.

With the purified enzyme condition, we supplemented a higher concentration of precursor and observed a greater peak area of the β-OH nsAA product as a result. Because standards for the β-OH nsAA intermediate products are not commercially available, we showed the HPLC traces as one form of confirmation (in addition to MS). The measurement of product peak areas also allows comparison of TTA variants on the basis of β-OH nsAA peak area. Ultimately, we have moved these traces to the supplemental to be more indicative of a result supporting a protocol, rather than achieving the overarching goal of total biosynthesis of the nsAA and subsequent incorporation (see Fig. S1).

5. In Figure 2d, the Y-axis should represent concentration, as in other figures, rather than peak area.

Like for many papers in natural product biosynthesis, the intermediate beta-hydroxylated chemistries are not commercially available, and thus we cannot quantify them. For this reason, we measure and report product peak areas which is not uncommon in the natural products field. We are using that information only to show relative performance among TTAs for a given substrate, rather than trying to compare substrates among the TTAs. In our revised manuscript, we also mention that we can estimate the beta-hydroxylated Phe concentration on the basis of the corresponding Phe derivative given that the molar extinction coefficients did not change for the limited cases in which we had both the beta-hydroxylated Phe and non-hydroxylated Phe standards. But we do not want to place too much emphasis on these estimates and have kept them in the SI (Table S5).

6. In Figure 3a and the text, you should provide the full names of PSDH and AT the first time they appear in the article.

The full names of PSDH and AT are now detailed in the introduction of the manuscript at the first time they appear.

7. Are the concentrations of nsAAs in the bar chart intracellular or extracellular? In additionl I didn't see a standard curve for the compounds in this article; how were their concentrations determined?

The concentrations of the nsAAs are extracellular concentrations. The concentrations were determined by a standard curve to correlate the peak areas obtained by HPLC to the concentration. This information was already described in the original manuscript Methods – see lines 519-522 and lines 537-541. Based on common practices in the field, and since we already include 48 supplementary figures, we have not included the standard curve that we used for each of the targeted nsAAs multiplied by their ~4-5 precursor metabolites (this would be nearly 100 additional SI figures after expansion of the substrate scope).

8. In Figure 3b, the product names are not labeled, but they are listed as 1e to 4e in Figure 3c. I suggest mentioning the names of the corresponding products on the Y-axis, such as "concentration of products (e)."

We see the reviewer's perspective and have updated the Y-axes on Figs. 2d and 2g to specify "nsAA product concentration (e) (μM)". We hope that this offers clarification for all readers.

9. In Figure S2, which antibody was used for the Western blot? I didn't find any information about this. It doesn't appear to be an antibody against SUMO, as the band for the group without the SUMO tag is present.

This was described in the methods section of our originally submitted manuscript (see Lines 576-578, text copied below for convenience):
"western blot with an HRP-conjugated 6*His, His-Tag Mouse McAB primary antibody"

10. Why didn't you test the aldehyde stability of 4b, only 1b-3b (Figures S4-S6)?

We thank the reviewer for this comment. We have added the aldehyde stability assay for 4b to the supplemental information (Fig. S8).

11. Why didn't you test s-ObiH in different strains in Figure 3c? It showed better conversion than s-RpPSDH in Figure 3b. Additionally, PbTTA demonstrated better conversion for 1b-3b, but s-ObiH seemed to perform better for 2b-4b in Figure 3b when coupled with PSDH and AT. It will be better for giving more explanation for this part.

We thank the reviewer for the question, though there is one minor misunderstanding: The variation in the previous Fig. 3b (updated Fig. 2d) was between co-expressed s-ObiH + s-RpPSDH or s-PbTTA + s-RpPSDH, so the PSDH remained constant between the two conditions. While we did see more optimal pairings between the one TTA and a particular chemistry compared to another, the overall hypothesis of the experiment in Fig. 3c (updated Fig. 2e) was asking if an aldehyde-stabilizing strain was beneficial, which was ultimately accomplished by these results without need for testing both TTAs nor the best TTA for each chemistry. Our earlier data also showed that the difference in performance between the two TTAs was not that large – we come back to this topic in our response to the next comment from the reviewer, which seems related.

12. SrCAR shows better results in Figure 4a, but why did you choose MavCAR for the following assays in Figure 4b? The screening and selection of relevant enzymes throughout the article need more logical consistency.

We ultimately observed high and largely comparable titers for SrCAR, MavCAR, MmCAR, and MabCAR. Our original manuscript describes why we utilized MavCAR in the subsequent experiments – see lines 208-210.
"We proceeded with the *Mav*CAR construct for subsequent experiments as we had known the protein to express well under various conditions."

This and the previous comments give the impression that the reviewer believes we should be using the best TTA and CAR pair for each substrate, and that not doing so

therefore is illogical. If that is an accurate interpretation of the reviewer's comments, then we state the following: We agree that it was valuable to test the substrate specificity of more than one TTA and CAR, but where we disagree is about what to do next, especially when multiple TTAs and CARs exhibit desired polyspecificity and comparable activity. As stated in our introduction, our primary goal is to design one common pathway that exhibits polyspecificity. Ideally, this one pathway would be one common CAR/TTA/PSDH system as we have designed, not a system that requires substitution of different CARs and TTAs to use the pair that performs slightly better for each chemistry. That would require generation of a combinatorial collection of strains rather than one universal strain. Thus, we believe the most logical conclusions to draw from our CAR and TTA experiments is to simply advance forward with one, which need not be "best" one for each tested substrate. We would try to revise our introduction accordingly, but we think it already states our goal and criteria without getting into details about avoiding the need for interchanging enzyme homologs (see Line 76):
"(2) Employ promiscuous enzymes to access diverse nsAA targets from diverse substrates;"

13. In line 183, you mention "1 mM of 4a (due to lower solubility) to RARE.Δ16," but in the following study (Figure 4b), you used a higher concentration of 4 mM, which showed better results than 1 mM. This suggests that higher concentrations might be more effective at the beginning.

We agree with the reviewer that higher concentrations may be necessary, even with the observation of insolubility when substrate is added to the system. We have added the following detail (Line 230) towards this:

"with **4e** detectable in all replicates only when 4 mM **4a** was supplied, even though the addition of 4 mM **4a** results in insolubility at the start of the reaction."

As we continued our investigation, especially with longer-term experiments measuring endpoints at 20 h rather than the 2 h time point used for the data in Figure 4b (updated to Fig. S14), we found that increasing the concentration, even with initial solubility issues, still ultimately drove product conversion. We have included this detail in the main text (Line 227).

14. In line 218, you mention "frame UAG codon targeted for suppression (Fig. 5a)," but it is actually a TAG codon in Figure 5a. The description should be consistent, and the specific position of the TAG codon should be mentioned in the GFP reporter.

We thank the reviewer for pointing out what on face value may seem like an inconsistency. However, we are using the terminology correctly based on the specific context. Codons can be on DNA or RNA, although some more rigorous geneticists and molecular biologists may prefer to describe codons as only appearing on RNA – certainly in the context of what is being targeted for suppression since the TAG codon is not being directly suppressed. However, in a diagram of a heterologous expression cassette or plasmid, the genes are listed with their DNA sequence – thus use of TAG is more appropriate here. In any case, we agree with the reviewer that it is important to mention the position. We have updated our Methods section, which described that the reporter consists of a ubiquitin fusion, where the nsAA is incorporated between the

ubiquitin and the GFP, and included a citation. We also now include the full reporter sequence in the SI.

15. The SDS-PAGE gel, the mass spectrum of full GFP, and its yield are missing in the GCE section.

Thank you to the reviewer for pointing out this important detail. With our updated system described in the revised manuscript, we have removed the his-tags from the biosynthesis genes, and therefore successfully purified the GFP products (Fig. S21), confirmed incorporation of the nsAAs by MS and quantified protein product yields via Bradford (Fig. 3h and 3j). Please see the updated text below to complement the updated figures:

“Next, we scaled up the system to 50 mL culture volumes in baffled shake flasks and subsequently purified the Ub-GFP reporter (**Fig. S21**). We achieved yields of 22 mg/L reporter protein containing biosynthesized **2e** (**Fig. 3h**) and 33 mg/L protein containing biosynthesized **3e**, with confirmation by intact protein MS (**Fig. 3j**).”

16. Where is the "d" part in the legend of Figure 6? It seems to be missing or not labeled.

Thank you for pointing out this oversight. We have updated all the figures in the manuscript and ensured proper labeling in the figure captions and manuscript.

Reviewer #4 (Remarks to the Author):

Summary:
This study presents an innovative engineered bacterial platform that combines biosynthesis and site-specific incorporation of phenylalanine-derived non-standard amino acids (nsAAs) from aryl aldehyde or carboxylic acid precursors. The authors demonstrate the design of a multi-enzyme cascade in E. coli to bypass the traditional reliance on nsAA supplementation, enabling direct conversion of simple, achiral precursors into functionalized nsAAs. Key achievements include the successful coupling of biosynthesis with orthogonal translation systems (OTS) for high-fidelity incorporation of nsAAs into proteins and the extension of synthetic auxotrophy to aldehyde-dependent biocontainment. The work addresses critical limitations in scaling genetic code expansion for industrial and environmental applications.

Conclusion:
This manuscript presents a powerful platform for cost-effective nsAA production and incorporation by bridging metabolic engineering and genetic code expansion, with transformative potential for synthetic biology and biotechnology. While technical strengths are evident, addressing the major concerns—particularly host robustness and quantitative analysis of escape frequency—would solidify the work's impact. With appropriate revisions, this study is suited for publication in Nature Communications.

We thank the reviewer for this positive summary and their suggested opportunity to revise as described.

Major Concerns:
1. Host strain stability and robustness: Limited data are provided on the long-term stability of engineered strains under fermentation conditions. Experiments assessing growth performance of many constructed strains (e.g., RARE.Δ16 strain expressing a cascade of enzymes for the biosynthesis of phenylalanine derivates or the final biocontainment strain derived from DEP.e5) in bioreactors would strengthen industrial relevance.

We thank the reviewer for pointing out the importance of showing the strain performance of these engineered strains. While we lack access to a robust bioreactor and labored unsuccessfully to gain such access, we have instead added data to the Supporting Information to show the relative growth of the engineered strains transformed with the various plasmids used in the study. In particular, we show the growth overtime of these strains when induced or not induced at 200 µL/well scale in a 96-well plate, monitored via a plate reader (Fig. S44). We have added commentary on the observations to the main text as follows:

"As this platform could have widespread benefits for industrial biomanufacturing, we sought to evaluate the relative burden of each of the genetic constructs on the metabolically active cells. We therefore monitored the growth of 7 strains, both with and without induction of the genetic constructs at mid-exponential growth phase (Fig. S44). Compared to the RARE.Δ16 strain transformed with a plasmid containing the OTS and reporter, we see minor burden in a two-plasmid system for the additional plasmid containing the TTA and PSDH genes. This results in a 14% decrease in the biomass after 17 h for a system with inducible biosynthetic pathway gene expression

and a 19% decrease for that with constitutive expression. As expected, the burden is further increased with the addition of a third plasmid for expression of the CAR and Sfp, with 39% relative decrease in biomass compared to the one plasmid system. However, ultimately each of the strains were able to grow and accumulate biomass over the 17 h that the optical density was monitored at 200 µL/well scale in a plate reader. Minimizing the number of plasmids by genome integration of one or more of the pathway genes may help to reduce additional burden associated with plasmid maintenance."

Additionally, we show the accumulation of biomass at 50 mL scale in 250 mL baffled shake flasks, wherein a strain provided no substrate or 1 mM of an aldehyde precursor show similar growth curves (Fig. 4d). Monitoring of the fluorescent output over time, because of incorporation of the biosynthesized nsAA into the reporter protein (GFP), shows there is only an increase in the FL for the strain provided a substrate (Fig. 4e).

2. Quantitative analysis of escape frequency: While synthetic auxotrophy is achieved in Fig.6, quantitative analysis of escape frequency for DEP.e5 transformed with biosynthesis pathway is not addressed. A side-by-side comparison for escape frequency between DEP.e5 (requires 4e) and DEP.e5 containing the biosynthesis pathway (requires 4b) would enhance the manuscript.

We agree with the reviewer suggestion that our analysis of escape frequencies should be more quantitative, and indeed this is something that we originally planned to include. The reviewer references a side-by-side comparison between DEP.e5 and DEP.e5 transformed with the biosynthetic pathway. In starting down this path, we realized that there are two relevant questions that could be posed here, and we thought it best to answer them both, in a manner that avoids aldehyde toxicity as a confounding variable.

A first question is whether the escape rate of DEP.e5 might have changed due to its transformation with the biosynthetic pathway. For this experiment, we performed a traditional assay of escape frequency by culturing in liquid media that contains BipA in both cases, and by then washing and plating onto permissive (BipA+) or non-permissive (BipA-) media. By keeping the required media component consistent and avoiding aldehyde toxicity, we were able to culture a larger number of cells and thereby use comparable assay detection limits. The experiment showed no detectable escape, added Fig. 4e and to the manuscript as follows:

"To gain insight into whether the transformation of the DEP.e5 strain with a polyspecific biosynthetic pathway could influence escape rates at higher cell densities (and thus lower escape assay detection limits), we simply provided 4e rather than 4b to either the untransformed DEP.e5 strain or to the transformed DEP.e5 strain. We further investigated two different versions of transformed DEP.e5 strains, one transformed with the original plasmid that features inducible expression of the pathway genes (denoted pSA-2) or another that features constitutive expression (denoted pSA-4), which we found further improves the system in terms of growth on biphenylaldehyde. At an assay detection limit of $6.58 \times 10{-}10$ escapees per c.f.u. for DEP.e5, $1.79 \times 10{-}9$ escapees per c.f.u. for DEP.e5 transformed with pSA-2, and $1.45 \times 10{-}9$ escapees per c.f.u. for DEP.e5 transformed with pSA-4 (Fig. 4e), we

observed no c.f.u.s on non-permissive plates throughout 14 days of monitoring (Fig. S47-S48)."

A second question is more simply what the escape rate of the transformed DEP.e5 strain is when grown and selected under permissive media conditions that are based on Biphenylaldehyde. In this case, as documented in the section related to host robustness, the reactive nature of the aldehyde substrate and the requirement that it be converted to an nsAA meant that the strain exhibits a slower doubling time and lower final OD. Given this, we were able to conduct a quantitative assay of escape frequency with an assay detection limit of 3.99×10^{-8} escapees/CFU. At this detection limit, we were pleased to observe no escapees. This experimental data was added to Fig. 4d and to the manuscript as follows:

"To more rigorously quantify the frequency of escape from biocontainment, we repeated the assay with growth of the transformed DEP.e5 strain in liquid media that was supplied with 500 µM 4b for 3 days, a longer culture time under permissive conditions than standard escape frequency assays owing to the slower growth of this strain. After the 3-day incubation period, the culture was washed and plated on permissive or non-permissive plates. On the permissive plates, we observed an average of 5.33 × 107 c.f.u.s/mL (Fig. S45). We observed no c.f.u.s throughout 6 days of monitoring on the non-permissive plates at an assay detection limit of 3.99 × 10−8 escapees per c.f.u. (Fig. 4d) calculated on the basis of the number of cells plated (Fig. S46-S47)."

Thus, we can report that the system as currently designed exhibits no escape from biocontainment, and thus tight reliance on either the aldehyde or the nsAA for growth, but that fitness under these conditions is admittedly an area for improvement in a subsequent study. Such a study could readily leverage adaptive laboratory evolution towards this end.

3. Incomplete characterization of byproducts: For carboxylic acid pathways, minor azide reduction to 4-amino-phenylalanine (pAF) was observed. A deeper analysis of side reactions and strategies to mitigate unwanted modifications is warranted.

We thank the reviewer for pointing out the importance of monitoring potential unwanted byproducts and identifying solutions to combat these undesired modifications. In our assay for the biosynthesis and incorporation of p-Azido-phenylalanine from p-azidobenzoic acid, we observed very minor azide reduction to pAF (only 0.3%). This is a possible side reaction that seems to be a minimally documented, context-dependent phenomenon specific to the azide functionality in certain hosts. Published solutions include chemical restoration of the azide group (Arranz -Gilbert, et al 2022), which was necessary in their work due their observation of significant reduction of externally supplemented pAzF to pAF (at >10% pAF content in the reporter protein). It is also important to note that Ty and pAF differ by only 1 amu, and it is possible that the observed mass is due to minor levels of misincorporation rather than azide reduction.
We included the following in the manuscript to better describe what others have seen and the possibility that it may be misincorporation of Tyr:

"After purification, trypsin digestion, and LC/MS/MS analysis of our reporter protein (Fig. S19), we observed high-fidelity incorporation of biosynthesized 1e (>99%), with either possible azide reduction to 4-amino-phenylalanine (pAF)[44] or misincorporation of tyrosine observed at minor levels (0.3%; pAF and Tyr differ by only 1 amu)."

4. Substrate Scope and Pathway Efficiency: While the platform claims broad substrate specificity, data for bulky substrates (e.g., 4b/4a) show low titers (e.g., 5 μM 4e). The feasibility of scaling these yields for industrial applications remains unclear. Additional optimization (e.g., enzyme engineering or OTS evolution) should be discussed.

We agreed with the reviewer, and see a clear avenue for engineering the platform, particularly for improvements on the 4a/b chemistries and similar bulky chemistries. In addition to the existing brief discussion of the biphenylaldehyde-dependent strain as a selection for evolution that concluded the Discussion section of the manuscript, we have now added many new Results about what step in the pathway may be the bottleneck and associated optimization (Lines 176-193), as well as optimization of heterologous expression also in the Results section (Lines 280-313). For length reasons, we have opted not to expand even further on optimization strategies in the Discussion, and we hope these additions address the reviewer's concerns.

Minor Comments:

1. Fig. 2C: Labeling of HPLC traces (e.g., "in vivo s-ObiH + 1 mM (b)") is ambiguous. Specify which substrate corresponds to each trace.

We thank the reviewer for the additional comments so that we may provide further clarity to all readers. In revised manuscript, Fig. 2c has been moved to Fig. S1c. Here, we have provided further clarification in the figure caption for the labels, specifying what each trace represents for tracking the conversion of 1b to 1c, 2b to 2c, 3b to 3c, and 4b to 4c.

2. Fig. 6: Simplify schematic annotations (e.g., "GFP fragment composition") to improve readability. The figure legend for panel d is missing.

Thank you for providing an example where we can simplify the annotations. We have updated "GFP fragment composition" to "identity at incorporation site" Fig. 3d and 3e. While we may not have shortened the description, we believe this updated definition improves the readability of the figure by improving the clarity of what the data represents.

3. Clarify the rationale for using MOPS EZ Rich media with 2% glucose in biosynthesis assays (line 113)—does this medium optimize precursor uptake or enzyme activity?

The primary rationale for using a MOPS EZ Rich media for the biosynthesis assays is that it is a defined media and therefore reduces experimental variability across separate experiments over time. The second reason is that we can then transition to the use of MOPS EZ Rich media lacking the aromatic amino acids (Y, W, F) which is valuable to reduce the standard amino acids that can lead to misincorporation products at low concentrations while still supporting high specific growth rate.

The following has been added to the Methods section:

“Biosynthesis assays were carried out in MOPS EZ Rich defined media (Teknova M2105) with 2% glucose unless otherwise specified, as it is a defined media and may therefore reduce experimental variability across separate trials. Assays involving nsAA incorporation were carried out in MOPS EZ Rich defined media lacking the aromatic amino acids (Y, W, F) with 2% glucose or 2% glycerol, unless otherwise specified. This media is used to reduce the standard amino acids that can lead to misincorporation products a low concentrations while still supporting growth.”

Summary of Latest Revisions

During processing of our article for typographical and formatting changes needed for acceptance, it came to our attention that we had not included the requisite evidence for confirmation of structurally novel compounds that we had synthesized during revision. It is the editorial policy of *Nature Communications* that all structurally novel compounds are confirmed by either H1-NMR, C13-NMR, or high-resolution mass spectrometry (HRMS).

To meet these guidelines, we have now performed additional experiments in which the 6 structurally novel non-standard amino acids (nsAAs) are biosynthesized by our semi-synthesis pathway in bacterial lysate format, followed by HRMS using an Agilent Revident LC-QTOF system. All 6 target products resulted in observed masses that were within the standard acceptable error limit of 5 ppm between expected and observed mass.

Below, we copy the associated updates to our Results and Methods sections. The SI file contains new figures (Figs. S44-S49) that correspond to the Electrospray Ionization scans for each product. It should be noted that LC/MS of lower resolution had already been performed for these chemistries in the more relevant *in vivo* biosynthesis context, and those chromatograms remain included in the SI.

Revised Results Section:

Lines 339-344:

"In addition to measuring GFP fluorescence, we confirmed nsAA biosynthesis by mass spectrometry for each of the 14 nsAAs shown in **Fig. 4a** (**Fig. S22-43**). We further used high-resolution mass spectrometry (HRMS) to verify the 6 structurally novel nsAA products for which commercial standards are not available, all of which had less than 5 ppm error from expected masses (**Fig. S44-S49**). For chemistries that had commercially available standards (**6, 7, 8, 11, 12, 13, 16**), we quantified nsAA production (**Table S6**)."

Revised Methods Section:

Lines 750-772:

"*Biocatalytic production of novel nsAAs for HRMS confirmation*

To further confirm the synthesis of the six novel chemistries in our panel (**5e**, **10e**, **14e**, **15e**, **17e**, and **18e**) by HRMS, we prepared semi-synthesis reactions in cell-free extracts as previously described[31]. To prepare cell-free extracts, 5 mL starter cultures of RARE.Δ16 pZE-s-ObiH and RARE.Δ16 pZE-s-*Rp*PSDH started from glycerol stocks were grown in LBL media with 30 µg/mL Kan at 37 °C with shaking overnight until the cultures

reached saturation. Then, 250 mL of 2xYT media with 30 µg/mL Kan in a 1 L baffled flask was inoculated from the overnight culture at a 1:100 inoculation ratio and incubated at 37 °C with shaking at 250 RPM. Cultures were induced at OD_{600} = 0.5-0.8 with 0.2 µM aTc and then incubated at 20 °C with shaking at 250 RPM for 24 h. Cells were harvested by centrifugation (7000 × *g* at 4 °C for 6 min). Cells were resuspended in a cold lysis buffer (50 mM HEPES pH 7.5, 0.4 mM PLP) at 5 mL lysis buffer per g cell pellet and sonicated using a QSonica Q125 sonicator with cycles of 5 s at 90% amplitude and 10 s off for 8 min x 2 on ice. Crude lysates were transferred to microcentrifuge tubes (1mL aliquots), centrifuged at 18213 × *g* for 30 min at 4 °C, and subsequently filtered using a 0.22 µM syringe filter to collect the clarified lysates. Next, the total protein concentration of the s-ObiH and s-RpPSDH clarified lysates were measured by Bradford assay using bovine serum albumin as a reference. The reaction was performed with 5 mg/mL s-ObiH lysate and 3 mg/mL s-RpPSDH lysate in reaction buffer containing 50 mM HEPES pH 7.5, 400 µM PLP, 100 mM L-Thr, 50 mM L-Glu, and initialized with addition of 5 mM of aldehyde (**5b**, **10b**, **14b**, **15b**, **17b**, or **18b**) (100 mM stocks prepared in DMSO). Reactions were run at 150 µL scale in triplicate in 96-well plates for 24 h at 30 ºC at 1000 RPM. Samples were quenched with methanol (4x reaction volume) and centrifuged (4000 × *g* at 4 ºC for 12 min) to clear insoluble protein precipitate. Samples were then analyzed by HRMS."

Lines 856-869:

"*HRMS analysis*

High resolution MS analyses were performed using an Agilent 1200 Series HPLC system coupled to an Agilent 6575 Revident Q-TOF mass spectrometer equipped with a dual Agilent Jet Stream electrospray ionization (AJS-ESI) source operated in positive ion mode. Chromatographic separation was achieved on an Agilent RHHD Eclipse C18 column (2.1 x 50 mm, 1.8_µm) maintained at 30 °C with an initial mobile phase of solvent A/B = 95/5 (solvent A, water, 0.1% formic acid; solvent B, acetonitrile, 0.1% formic acid) and was maintained for 5 minutes. A gradient elution was performed (A/B) with: gradient from 95/5 to 50/50 for 5-12 min, 50/50 to 100/0 for 12/13 min and 100/0 to 95/5 for 13-14 min and isocratic flow at 95/5 for 14-15 min. The flow rate was 0.3 ml/min and injection volume was 2 µL. The mass spectrometer was operated in full-scan MS acquisition mode over an m/z range of 30-1000 at a scan rate of 2 spectra/s. Source parameters were set as follows: 2.5 kV capillary voltage, 335 °C drying gas temperature. Continuous mass calibration was performed using reference ions at m/z 121.050873 and 922.009798. Data acquisition and processing were carried out using Agilent MassHunter software."