Modular control of L-tryptophan isotopic substitution via an efficient biosynthetic cascade†

Clayton M. Thompson, Allwin D. McDonald, Hanming Yang, Silvia Cavagnero* and Andrew R. Buller*†

Isotopologs are powerful tools for investigating biological systems. We report a biosynthetic-cascade synthesis of Trp isotopologs starting from indole, glycine, and formaldehyde using the enzymes L-threonine aldolase and an engineered β-subunit of tryptophan synthase. This modular route to Trp isotopologs is simple and inexpensive, enabling facile access to these compounds.

Isotopically substituted amino acids are valuable tools for the mechanistic and structural analysis of biological systems. Among the standard amino acids, L-tryptophan (Trp) is the most structurally complex and serves as the precursor to diverse natural products and clinically used compounds. Selective isotopic substitution of Trp is often essential for determining the metabolic fate of individual atoms and the kinetic properties of enzymes that manipulate Trp. Further, the low prevalence of Trp in proteins makes this amino acid attractive for selective substitution methodologies for protein NMR. In each of these cases, however, access to the requisite Trp isotopolog is a central hurdle.

While there are diverse synthetic methods that can produce amino acids, making isotopically substituted compounds comes with unique challenges. For instance, synthetic routes should be planned starting from the cheapest possible isotopically substituted precursor. The typical concerns over catalyst loading and atom economy are secondary to yield with respect to the expensive, isotopically labelled or substituted reagents. Traditional synthetic routes that can access isotopically substituted Trp include the asymmetric Strecker synthesis (Scheme 1) and reductive amination. These approaches, however, often require isotopically enriched starting materials that are not commercially available or are prohibitively expensive. Additionally, most applications of amino acids like Trp benefit from high enantiopurity and chiral small molecule catalysts that are often difficult to obtain.

Enzymatic Trp synthesis is an attractive alternative with the potential for both high selectivity and low environmental impact. For example, previous studies on enzymatic mechanisms utilized a (2-2H)Trp that was prepared biocatalytically from indole and S-methyl cysteine in D2O using the enzyme tryptophanase (Enzyme Commission (EC) 4.1.99.1, Scheme 1). Although this procedure was successful, it afforded only 40% yield after a 2-day reaction. The synthesis of a [3,3-2H2]Trp presented a more substantial challenge. To make this isotopolog, methionine γ-lyase (EC 4.4.1.11) was used to catalyze the exchange of the three protons on the 2- and 3-positions of Trp.

Scheme 1. Synthetic and proposed routes to L-tryptophan (Trp). KPi = K2HPO4, potassium phosphate buffer.

†Electronic supplementary information (ESI) available. See DOI: 10.1039/d0ob00868k
S-ethyl-l-cysteine with D$_2$O solvent over three days in a 47% yield, followed by reaction with indole and tryptophanase.$^{10}$ Methionine γ-lyase accomplishes this labeling through desaturation at C3 during the course of its reversible reaction. Unfortunately, the resulting Trp was composed of a mixture of mono, di-, and tri-deutero isotopologs. Biocatalytic approaches have also been used to prepare various $^{13}$C-substituted Trp by the action of tryptophan synthase (EC 4.2.1.20) on the correspondingly substituted $^{13}$C-Ser.$^{11}$ This $^{13}$C-Ser was prepared from isotopically substituted glycine (Gly) and formaldehyde using the enzyme Ser hydroxymethyltransferase (SHMT, EC 2.1.2.1). This method is particularly attractive, as Gly isotopologs are relatively cheap. However, long reactions times (>80 h) were required, and contamination with unreacted Gly necessitated tert-butyloxycarbonyl (Boc)-protection followed by column chromatography to isolate the $^{13}$C-Ser, reducing yields of the combined process to <30%. Here, we propose an affordable and efficient method to generate Trp starting from Gly, formaldehyde, and indole in a one-pot, two-step enzymatic cascade reaction that leads to the facile synthesis of enantiopure Trp isotopologs.

We began by considering the synthesis of Trp from Ser and indole. Previous efforts used wild-type tryptophan synthase from *Salmonella typhimurium*, which is a multi-enzyme complex that is tedious to prepare for synthetic applications.$^{12}$ The β-subunit of the complex (TrpB) is responsible for the pyridoxal phosphate (PLP)-dependent condensation of Ser and indole. Recently, the TrpB protein from the hyperthermophilic archæon *Pyrococcus furiosus* (PfTrpB) was engineered for activity outside of its native complex.$^{13,14}$ Detailed studies have subsequently shown that a stand-alone variant containing eight mutations, PFTrpB$_{B9}$, has structural and kinetic properties that are nearly identical to those of the parent complex.$^{15}$ We overexpressed PFTrpB$_{B9}$ in *E. coli* and purified the enzyme using Ni-affinity chromatography.$^{15}$ This procedure routinely yielded >500 mg protein per L culture. This high expression level is particularly important for practical synthetic applications. To achieve the quantitative deuteration at the 2-position of Trp, we deployed PFTrpB$_{B9}$ for the condensation of Ser and indole in deuterated buffer (Scheme 1, Route 1). This biocatalytic reaction produced [2-2H]Trp in 89% isolated yield (18.2 mg, Fig. 1, Table S1†). Given the high yield and excellent properties (expression, speed, thermal stability, etc.) of the catalyst, this synthesis of [2-2H]Trp represents a significant improvement over previous routes. However, synthesis of more sophisticated Trp isotopologs, including the desirable 2-$^{13}$C substitution, is hampered by the significant cost of the corresponding Ser isotopolog.

Previous efforts to synthesize isotopologs of Ser from Gly utilized the PLP-dependent SHMT enzyme, which uses the expensive tetrahydrofolate cofactor to deliver formaldehyde. This route was further stymied by the need for [Boc]-derivatization and isolation of the Ser isotopolog.$^{11}$ We hypothesized that Ser could be produced from formaldehyde and Gly using an l-threonine aldolase (LTA, EC 4.1.2.5). We reasoned that even though the synthetic direction of the LTA reaction is unfa-
indole. The catalyst loadings to achieve this conversion were fairly high for enzymatic reactions, 20 µM of each enzyme, corresponding to 0.4 mol% catalyst. In practical terms, however, the high catalyst loadings are compensated for by the high expression of each enzyme.

This new biocatalytic cascade is highly modular with respect to isotope incorporation (Scheme 1, Route 2). The (2-13C)Trp substitution can be incorporated from commercially available (2-13C)Gly, and (2-2H) substitution (deuteration at the α-position) is achieved via exchange with a deuterated solvent, as shown above. Substituting at the 3-position can be controlled via the source of formaldehyde, which is commercially available as a variety of isotopologs. To demonstrate this the new cascade methodology, we employed the two-enzyme cascade to synthesize the complex (2-13C, 2, 3, 3-2H3)Trp isotopolog starting from the appropriate isotopically enriched precursors. This particular Trp isotopolog was chosen because of its potential as a resonance-assignment aid for protein NMR spectroscopy. In addition, this compound showcases the modular nature of the PfTrpB289-LTA cascade. In order to characterize the extent of 13C incorporation at position 2 of Trp and D incorporation at positions 2 and 3, we carried out UPLC-MS as well as 1H and 13C NMR analysis (Fig. 1 and S5-S8f). These data show a complete absence of the 2-H in (2-2H)Trp. Combined, the above factors result in a more intense, sharper 2-13C resonance.

In conclusion, we have developed a simple and modular one-pot method for the synthesis of enantiopure Trp isotopologs. The simpler (2-2H)Trp can be accessed through a single enzymatic reaction. Independent control of substitution at the 2- and 3-positions can be achieved through a dual enzyme cascade, as demonstrated by the synthesis of (2-13C, 2, 3, 3-2H3)Trp. We note that each of the enzymes, PfTrpB289 and TmLTA, can be overexpressed to exceptionally high protein density in E. coli (0.5–1.0 g protein per L culture) facilitating preparative-scale reactions. We envision that this modular approach can be implemented with any source of selectively or uniformly substituted glycine, formaldehyde, or indole, which are all commercially available, to yield diverse isotopically-substituted Trps. The resulting products promise to serve as valuable tools in mechanistic biochemical studies as well as structural-biology investigations.

**Conflicts of interest**

A.R.B is an inventor on a patent on the use of PfTrpB289 for the synthesis of tryptophans.

**Acknowledgements**

We are grateful to Jon Ellis for helpful scientific discussions and to Aadhishe Kasat and Meghan Campbell for the synthesis of FDVA. S. C. thanks the National Institute of Health (grant R01GM125995) and the UW-Madison Graduate School.
for funding. A. R. B. gratefully acknowledges support from the National Institute of Health New Innovator Award (DP2GM137417). The Bruker AVANCE 400 NMR spectrometer was supported by NSF grant CHE-1048642, and the Bruker AVANCE 600 NMR spectrometer was supported by NIH grant S10 OD012245.

References