Journal of Fluorine Chemistry 261-262 (2022) 110014

Contents lists available at ScienceDirect

Journal of Fluorine Chemistry

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

Development of a single culture *E. coli* expression system for the enzymatic synthesis of fluorinated tyrosine and its incorporation into proteins

Noelle M. Olson [a,1], Jorden A. Johnson [a,1], Kerstin E. Peterson [a], Stephen C. Heinsch [b], Andrew P. Marshall [a], Michael J. Smanski [b], Erin E. Carlson [a], William C.K. Pomerantz [a,*]

[a] Department of Chemistry, University of Minnesota, 207 Pleasant St. SE, Minneapolis MN 55455, USA
[b] Department of Biochemistry, Molecular Biology, and Biophysics, University of Minnesota, 1479 Gortner Avenue, St. Paul, MN 55108

ARTICLE INFO

Keywords:
Fluorine NMR
Fluorotyrosine
Metabolic labeling
Bromodomain
Tyrosine phenol lyase

ABSTRACT

Current experiments that rely on biosynthetic metabolic protein labeling with ^{19}F often require fluorinated amino acids, which in the case of 2- and 3-fluorotyrosine can be expensive. However, using these amino acids has provided valuable insight into protein dynamics, structure, and function. Here, we develop a new in-cell method for fluorinated tyrosine generation from readily available substituted phenols and subsequent metabolic labeling of proteins in a single bacterial expression culture. This approach uses a dual-gene plasmid encoding for a model protein BRD4(D1) and a tyrosine phenol lyase from *Citrobacter freundii*, which catalyzes the formation of tyrosine from phenol, pyruvate, and ammonium. Our system demonstrated both enzymatic fluorotyrosine production and expression of ^{19}F-labeled proteins as analyzed by ^{19}F NMR and LC-MS methods. Further optimization of our system should provide a cost-effective alternative to a variety of traditional protein-labeling strategies.

1. Introduction

Enrichment of aromatic amino acids at protein binding interfaces has led to the development of several selective labeling strategies to enable the detection and characterization of biomolecular interactions including those with proteins, nucleic acids, carbohydrates, and small molecules [1–7]. Protein-observed fluorine NMR (PrOF NMR) is one approach that has gained traction for characterizing protein-protein and protein-small molecule binding events. This approach has predominately used monofluorinated aromatic amino acids, where the fluorine NMR chemical shift is highly responsive to ligand binding due to the sensitivity, large chemical shift range, and 100% isotopic abundance of the ^{19}F nucleus. Fluorine is almost completely absent from native biological systems, making it an ideal biorthogonal probe with virtually no background interference from endogenous nuclei [8]. Additionally, non-natural amino acids in which hydrogen atoms have been replaced by fluorine are incorporated into proteins without dramatically altering their conformation or function, although this is not always the case [9–13]. The lack of perturbation is consistent with the van der Waals radii of fluorine and hydrogen being 1.47 and 1.20 Å, respectively [14]. While ^{19}F NMR typically requires high-label incorporation, low levels of fluorination have been shown to enhance protein stabilization [15] and improve NMR spectra when significant perturbation results [16].

When selecting amino acid positions to incorporate fluorine into proteins, it is important to consider minimizing the complexity of the ^{19}F NMR spectrum and maintaining the protein's structural and functional integrity. Regarding investigations into protein-protein interactions (PPIs), aromatic amino acids make an excellent choice since they are enriched at binding interfaces [17], increasing the sensitivity of PrOF NMR to the detection of binding events. Being low in abundance, they result in an easily interpretable spectrum with few resonances. Existing methods for incorporating non-natural amino acids into proteins include site-selective incorporation via Amber suppression [18–20], post-translational side chain modification through either chemical conjugation [9,21] or enzyme-catalyzed reactions [22], or biosynthetic, sequence-selective replacement of the natural amino acid via metabolic labeling [16]. For PrOF NMR studies, we and others have metabolically labeled proteins in *E. coli* with 5-fluorotryptophan (5FW), 3-fluorotyrosine (3FY, Fig. 1A), 2-fluorotyrosine (2FY, Fig. 1A), 2-fluorophenylalanine (2FF), 3-fluorophenylalanine (3FF), and 4-fluorophenylalanine (4FF) [20,23,24].

For tyrosine and phenylalanine labeling of proteins, the natural

* Corresponding author.
E-mail address: wcp@umn.edu (W.C.K. Pomerantz).
[1] These authors contributed equally to the manuscript.

https://doi.org/10.1016/j.jfluchem.2022.110014
Received 25 April 2022; Received in revised form 22 June 2022; Accepted 26 June 2022
Available online 28 June 2022

amino acid is replaced in the expression media by the non-natural fluorinated analogue (e.g., 2FY, 3FY, or, 2/3/4FF[25]). When protein expression is induced, the endogenous aminoacyl tRNA synthetase recognizes the fluorinated amino acid and incorporates it into the protein of interest. This method can use the DL39(DE3) *E. coli* strain, an auxotrophic cell line for phenylalanine and tyrosine. While this system has produced a highly fluorinated model protein in our lab, 3FY-BRD4(D1) (>95% ^{19}F incorporation) [26], protein yields tend to be reduced (5−15 mg/L) in comparison to unlabeled BRD4(D1) expressions. Alternatively, the small molecule additive glyphosate can be used in combination with non-auxotrophic *E. coli* expression strains to induce auxotrophy but is required in significant amounts (1 g/L). High concentrations of amino acids (mM) can be used without the need for glyphosate [23]; however, all of the afroementioned methods for tyrosine labeling rely on the use of expensive fluorinated amino acids ($178/g racemic 3FY, Ambeed).

In contrast to tyrosine labeling, tryptophan labeling takes advantage of the endogenous biosynthetic capabilities of a standard *E.coli* expression cell line, BL21*(DE3) [27]. Instead of replacing the tryptophan in the defined media with 5-fluorotryptophan (5FW, $875/g, Sigma), an affordable tryptophan precursor, 5-fluoroindole ($52/g, Sigma) is added, which is converted into enantiomerically pure 5FW via the endogenous tryptophan synthase in the last step of the Shikimate pathway [27]. This method is robust, and in the case of the model protein employed herein, BRD4(D1), we routinely obtain yields of 60−90 mg/L with >95% ^{19}F incorporation [24,26]. This yield is enough for around 150 PrOF NMR experiments from a 1 L expression using a ^{19}F-specific inverse cryoprobe. Although 5FW-labeling is more efficient and cost effective, it is a low-abundant amino acid. In the case of BRD4 (D1), there are only three tryptophan residues, thereby limiting the capability of 5FW-BRD4(D1) to report on binding events. In some proteins, such as α-synuclein, there are no tryptophans, requiring an additional tryptophan to be added [28] or alternative amino acids to be fluorine labeled [29,30]. Tyrosine residues are more abundant than tryptophan. In the case of BRD4(D1) there are seven tyrosines, three of which, Y97, Y98, and Y139, are located in the protein-protein interaction binding site. Therefore, 2FY/3FY-labeled-BRD4(D1) provides additional information on binding site location and protein dynamics (Fig. 1B). As such, our lab sought to develop an efficient biosynthetic in-cell system for both fluorotyrosine production and protein labeling using cheap and readily accessible precursors (e.g., 2-fluorophenol, $2/g, Sigma). As an additional advantage, this approach would generate the enantiomerically pure amino acid, versus the more commonly used racemic version. We use BRD4(D1) here as a model protein given our past success in metabolic labeling it with 3FY.

To achieve this goal, we use a tyrosine phenol lyase (TPL) from a thermophilic organism *Citrobacter freundii* (*Cf.* TPL) that can produce tyrosine analogues from substituted phenols. TPL is a pyridoxal 5′-phosphate (PLP)-dependent carbon-carbon lyase that catalyzes the forward and reverse reaction of phenol, pyruvate, and ammonia to form L-tyrosine and water in six mechanistic steps [31]. TPL has been used to produce tyrosine both as an isolated enzyme and in cells, and mechanistic and kinetic studies have revealed that the enzyme has a wide substrate scope. In terms of fluorinated tyrosine analogues, it has been used to make all possible mono-, di-, tri-, and tetra- isomers of tyrosine from their respective fluorophenols. To date, this has only been done with the isolated enzymes with reaction times of 3 days - 4 weeks [32, 33]. Given the success of tyrosine production in cells, we chose to evaluate if a TPL expression system could be developed using bacteria for both production of modified tyrosine analogues, and to demonstrate a method for using this enzyme system in cells for metabolic labeling of proteins. We focus our analysis on using our model protein BRD4(D1) here and describe the strengths and limitations of this approach.

2. Results and discussion

2.1. Fluorophenol toxicity and cellular uptake

We found that the addition of large amounts of fluorinated phenol to the growth medium resulted in a bacteriostatic effect on *E.coli* (**Figs. S2, S3**). Therefore, for the success of our expression system, we first needed to verify that cells could take up fluorophenol at a given concentration without affecting their growth or viability. Cultures containing BL21* (DE3) cells were grown to an optical density at 600 nm (OD_{600}) of 0.3, after which increasing amounts of a representative phenol, 2-fluorophenol, were added and the OD_{600} continued to be monitored. Cultures containing up to 3 mM 2-fluorophenol grew at a similar rate as the culture free from 2-fluorophenol. However, at the highest concentration of phenol tested (6 mM), growth of the culture stalled. When this culture was harvested and resuspended in fresh media without the presence of 2-fluorophenol, growth resumed, and the culture reached the exponential growth phase (Fig. S2). Similar behavior was observed with 3-fluorophenol with a large growth reduction at lower concentrations between 3 mM and 6 mM which also recovered upon resuspension in fluorophenol free media (Fig. S3). This growth recovery indicated that fluorophenols had a bacteriostatic effect on the cells at high concentrations and our ideal protein expression system should aim to keep the concentration of phenol below this level. However, for amino acid production in the absence of protein expression, we have found this to be less of a concern, *vide infra*.

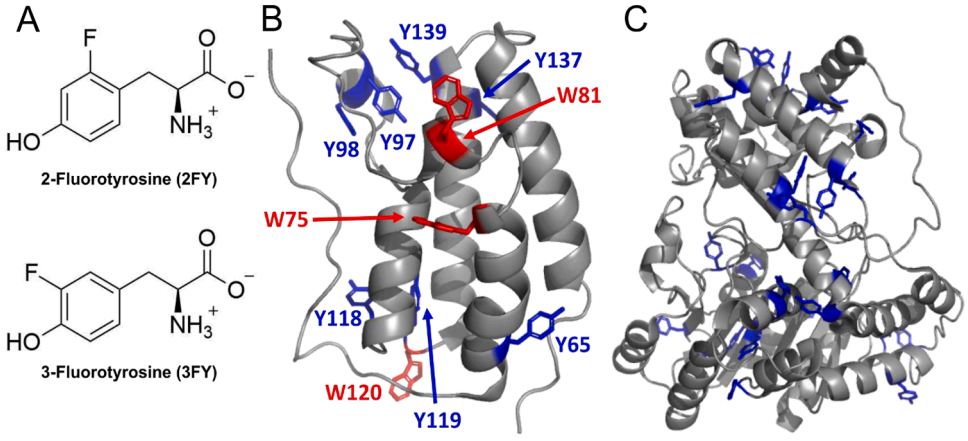

Fig. 1. (A) Structures of fluorinated amino acids 2FY and 3FY used for the expression of ^{19}F-labeled proteins through sequence-selective metabolic incorporation. (B) Ribbon diagram of BRD4(D1) (PDB: 2OSS) with seven native tyrosines and three native tryptophans highlighted in blue and red, respectively. C) Ribbon diagram of *Cf.* TPL (PDB: 2EZ1) with 23 native tyrosine residues highlighted in blue.

2.2. Plasmid design

Unlike tryptophan synthase, the TPL gene is not present in the *E. coli* BL21*(DE3) genome, and therefore needed to be incorporated via an engineered plasmid. We designed and built two plasmids, pSCH9.28 and pSCH9.29, that provide constitutive TPL expression and inducible BRD4 (D1) expression. Expression of BRD4(D1) via a T7 promoter is controlled by the addition of isopropyl β-D-1-thiogalactopyranoside (IPTG), which induces the T7 RNA polymerase genomically encoded in the *E. coli* strains used in this study. The production of TPL is expected to be slightly different from each plasmid, with TPL expression controlled by promoter J23101 in pSCH9.28 and by J23108 in pSCH9.29, with measured strengths in *E. coli* of 1 relative expression units (REU) and 0.20 REU, respectively [34,35]. These plasmids were transformed and cultured in *E. coli* BL21*(DE3). Culture samples were analyzed using sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) to assess the levels of protein production after IPTG induction. We observed the overexpression of TPL (MW 51,441 kDa) and His_6-BRD4 (D1) (MW 17,549 Da) (Fig. 3B). Due to the better performance for fluorine incorporation, we focus our subsequent analyses on using the pSCH9.29 plasmid.

2.3. Assessment of tyrosine analogue production

After verification of TPL production from our engineered plasmids, we investigated whether *E. coli* cells expressing TPL could efficiently convert differentially substituted fluorinated phenols to their respective tyrosine analogues. *E. coli* strains BL21*(DE3) and tyrosine auxotrophic DL39(DE3) cells were transformed with our lead dual-gene plasmid and cultured in a nutrient-rich, Luria-Bertani (LB) media until they reached the exponential growth phase. At this point, the cells were swapped into a defined media containing phenol, PLP, and pyruvate at pH=8 using our optimized conditions and literature protocols [36]. The cultures were incubated at 40°C for 2 h at which point ^{19}F NMR was used to quantify production of each tyrosine analogue. Use of 25 mM phenol resulted in monofluorinated tyrosine analogues (2FY and 3FY) with higher yields (3.9−6.4 mg/25 mL and 4.4−5.9 mg/25 mL) than the difluorinated tyrosine analogues with 2,6 FY yielding 2.2−2.8 mg/25 mL and 3,5 FY only detected in trace amounts by ^{19}F NMR (Fig. 2C, S15, S16, Table S2). Higher yields for difluorinated tyrosine may have been obtained if the reaction was allowed to proceed longer given previous success using purified TPL to produce 3,5-difluorotyrosine in cell-free systems with 85% conversion in 4 days [32]. Additionally, there was a similar yield of fluorotyrosine between the auxotrophic cell line (DL39) and the BL21* cell line, with only 2FY showing a significant difference (p-value = 0.0003). In the NMR experiments, in some cases a second aromatic fluorine resonance was observed close to the synthesized amino acid chemical shift (Figs. S7−S20). It was unclear if this was a rotameric species, a bound state resonance with cellular material, or a chemical modification. With the exception of 2,6 FY for which the rotamers were supported based on a solvent dependence in the NMR integrations (Fig. S8), these resonances were not included in quantitation. Due to the low signal to noise in our NMR experiments and in some cases observation of rotameric species and unconfirmed resonances, quantitative LC-MS using authentic standards of 2FY and 3FY was used as a secondary method to assess 2FY and 3FY production from the cultures. In this case, quantitative MS gave closely agreeing fluorotyrosine yields with our NMR measurements (n=3, p>0.05) (Fig. 2C).

The total amount of phenol added to each culture, if converted completely to fluorotyrosine, would largely exceed the amount of racemic 2FY or 3FY typically added in the auxotrophic expression method (60−80 mg/L). From our experience, this optimal amount is roughly based on the number of tyrosine residues in the target protein; 10 mg/L per tyrosine residue, of which His_6-BRD4(D1) has eight due to an added tyrosine for the affinity tag. However, for our protocol, we based our conditions on previously described methods for optimal cell-

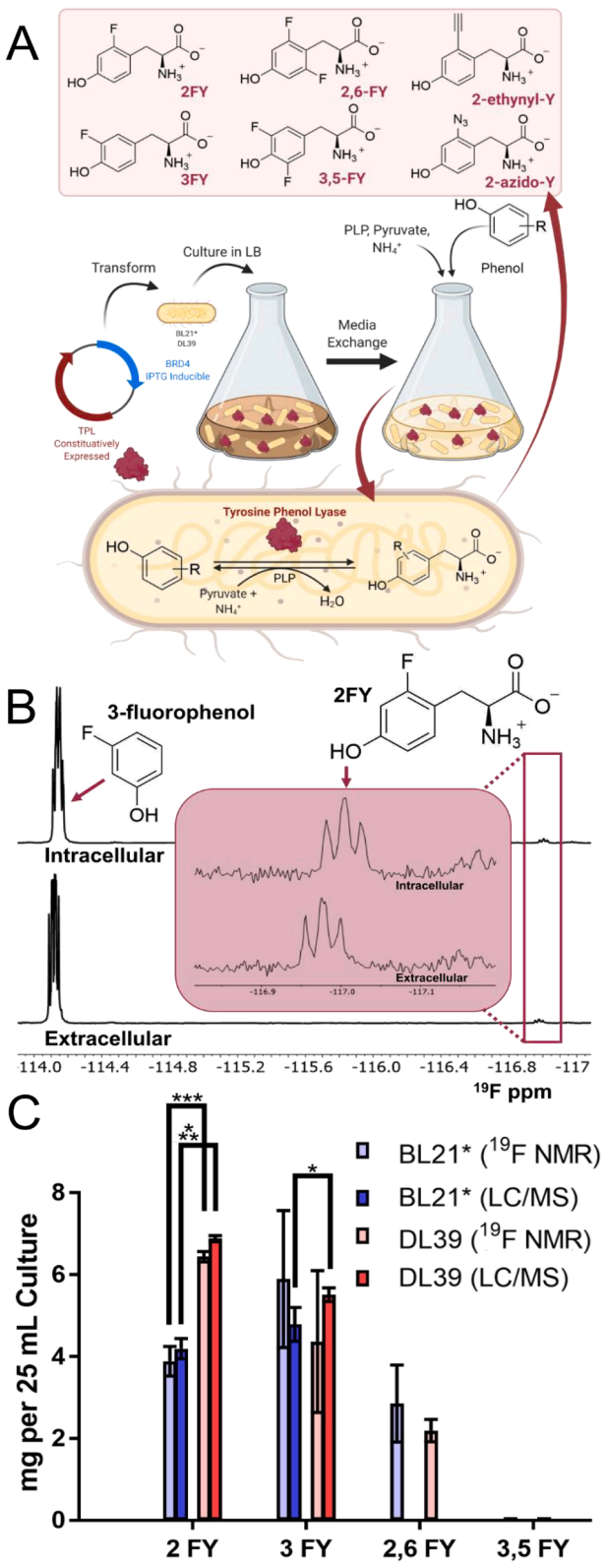

Fig. 2. Quantification of fluorotyrosine production. (A) Workflow for conversion of fluorophenol to fluorotyrosine in *E. coli* expressing TPL. Created with BioRender.com. (B) Representative ^{19}F NMR spectra of intra- and extracellular culture extracts. C) ^{19}F NMR and LC-MS quantification of substituted tyrosine analogues in 25 mL BL21*(DE3) or DL39(DE3) E. coli cultures. (n=3, * = p<0.05, *** = p<0.001)

culture-based production of tyrosine using TPL [36]. Given this large excess of phenol, the reactions converting monofluorinated phenols to fluorotyrosine (2FY and 3FY) had overall low conversion rates of 1.4−2.6% and 2.6−5.6%, respectively. However, it should be noted that this is a conservative estimate of fluorotyrosine production since any phenol that is produced and subsequently translated into proteins was not captured by this quantification. Additionally, we do not yet know if the incorporation of fluorotyrosine into TPL, which has 23 native tyrosine residues (Fig. 1C), affects its ability to perform the conversion, which will be the focus of a future study. As one way to increase yields, we tested the effect of adding tyrosine to the reaction culture, to suppress the reverse reaction of TPL, as the enzyme may preferentially use canonical tyrosine as a substrate in the conversion back to phenol. However, adding L-tyrosine only had a small, though statistically significant, difference on the amount of fluorotyrosine produced (Fig. S6). Finally, using both our ^{19}F NMR and LC-MS quantification, we further showed that the amount of fluorotyrosine produced in the whole cell reaction was scalable to larger 250 mL cultures in the BL21*(DE3) strain (Fig. S5). Extrapolation of 2FY production to a 1 L culture volume gave closely agreeing yields for both 25 mL (156 ±15 mg/L) and 250 mL (156 mg/L) reactions, as well as closely agreeing values across the two quantification methods (Fig. S5). Given that 2FY production should be enantiomerically pure, this amount is in excess to the amount of racemic 2FY needed for our auxotrophic system for fluorine-labeling BRD4(D1).

Based on the success with the production of mono- and difluorinated tyrosine in our cellular experiment, we next attempted to generate a broader set of tyrosine analogues amenable to copper-catalyzed azide/alkyne cycloaddition (CuAAC) functionalization. We performed the

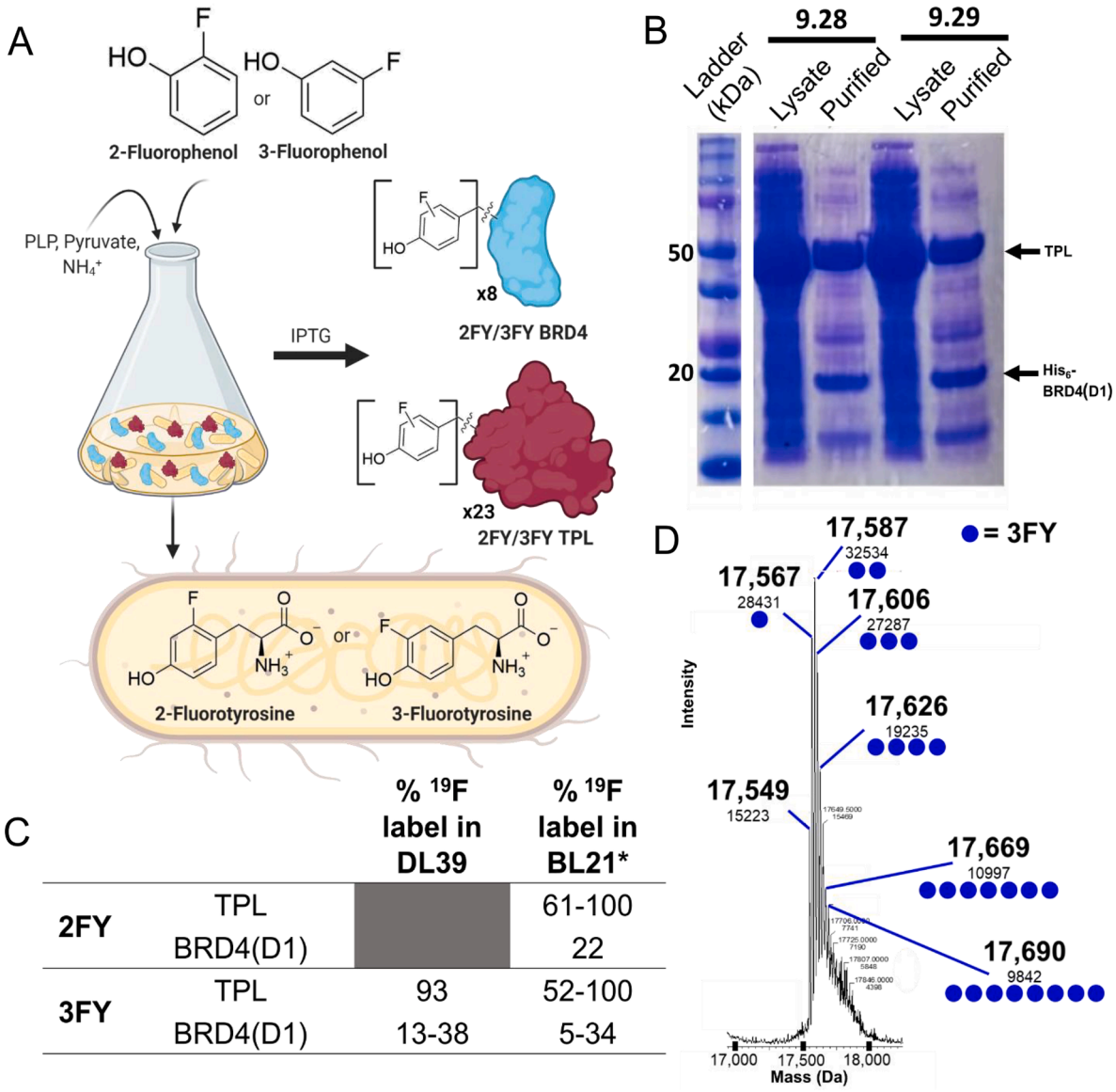

Fig. 3. Determination of protein labeling with fluorotyrosine using a TPL/BRD4(D1) dual-gene plasmid. (A) Single-culture workflow for in-cell fluorotyrosine production and labeled protein expression. Created with BioRender.com. (B) SDS-PAGE analysis of 3FY expressions using plasmid 9.28 and 9.29 after lysis by sonication and isolation of 3FY-His$_6$-BRD4(D1) using a Ni-NTA column. Modified from Johnson, 2020, with permission [37]. (C) Percent ^{19}F label incorporation into BRD4(D1) and TPL determined by intact protein LC-MS of proteins isolated from the workflow in Fig. 3A and both BL21*(DE3) and DL39(DE3) E. coli strains. (D) Deconvoluted mass spectrum of His$_6$-BRD4(D1) isolated from 3FY expression in the BL21*(DE3) cell line. We observed the intact protein mass for unlabeled protein ([M+H]$^+$ 17,549) as well as masses that closely match the +18 Da mass shift (replacement ^{1}H with ^{19}F) in multiple fluorinated species. Using the expected [M+H]$^+$ values of the fluorinated protein states for comparison, we approximated a total 42% ^{19}F incorporation in 3FY-BRD4(D1) (Equation S1, Table S3). Due to the limited resolution of the mass spectrometer, some observed masses fell between the expected +18 Da spacing of n and n+1 fluorinated species. In these cases, it was assumed that the signal intensity was identical for both n x FY His$_6$-BRD4(D1) species for the purpose of determining total ^{19}F incorporation

whole-cell reaction procedure with 2-ethynylphenol, 3-ethynylphenol, and 2-azidophenol. LC-MS analysis detected trace amounts of 2-ethynyltyrosine and 2-azidotyrosine present in the cultures. However, without authentic 2-ethynyltyrosine and 2-azidotyrosine standards, we were unable to quantify their amounts. Further optimization of conditions, or enzyme engineering would be necessary to increase the production of these clickable amino acids, and thus these amino acids were not pursued further.

2.4. Assessment of tyrosine analogue incorporation into proteins

The next step in our strategy was to see if the fluorinated tyrosine analogues were incorporated into proteins expressed by the cells. To accomplish this, we transformed the dual-gene plasmid construct containing an IPTG inducible gene for BRD4(D1) in addition to a constitutively expressed TPL gene (9.29) into both BL21*(DE3) and DL39(DE3) *E. coli* strains. Cells were first cultured in LB media until the exponential growth phase, then exchanged into a defined media containing 0.7 mM of the desired fluorophenol, 0.7 mM pyruvate, and 0.4 mM PLP at a pH of 8. Following the metabolic labeling protocol of Gee et al., the culture was allowed to recover at 37°C for 90 min [24]. At the 45-min and 90-min time points, the cultures were batch-fed an additional 0.7 mM and 0.35 mM fluorophenol and pyruvate, respectively. In total, 1.7 mM fluorophenol was added to the expression cultures. This amount, if completely consumed and converted to fluorotyrosine, would exceed the molar concentration of 2FY or 3FY used in the established protocol for metabolic labeling (80 mg/L or 0.4 mM), which was to account for the overall low conversion rates of fluorophenol by TPL that we observed. After the 90-min recovery period, the temperature was reduced to 20°C for 30 min after which 0.5 mM IPTG was added to induce BRD4(D1) expression. After 18 h, SDS-PAGE of the cultures showed the overexpression of both TPL and BRD4(D1) (Fig. 3B). Cells were isolated, lysed and proteins purified using both Anion Exchange Chromatography (AEC) and Ni affinity chromatography.

We first analyzed fractions of cellular lysates purified by AEC that contained TPL. Intact protein LC-MS was used to determine the degree to which TPL was labeled by ^{19}F. Due to the low signal-to-noise ratio and resolution of the spectra, we were not able to distinguish individual fluorinated TPL species and thus estimated the ^{19}F incorporation based on the dominant protein mass. Replicates ranged in labeling efficiency in the 2FY and 3FY systems (61−100% and 52−100%, respectively), demonstrating that our one-pot expression workflow was successful at both synthesizing fluorinated amino acids and incorporating them into proteins (Fig. 3C). However, this is a more incomplete and variable label incorporation than is accomplished using our fluorotyrosine-containing media, which typically achieves > 90%. We attribute this to the high abundance of tyrosine residues in the TPL sequence (23 total) and its expression in high levels in our cultures as evidenced by SDS-PAGE (Fig. 3B).

We next analyzed the BRD4(D1) purified by Ni affinity chromatography by intact protein LC-MS to determine if ^{19}F could be incorporated into a target protein when its expression is induced in the same culture as TPL. This analysis revealed the presence of multiple fluorinated BRD4 (D1) species with ^{19}F incorporation estimated to be 5−42% for 3FY (Fig. 3C and 22% for 2FY. This is a drastically lower label incorporation than our standard protocol and also noticeably lower than the ^{19}F incorporation into TPL in the same culture. We hypothesize that the fluorotyrosine that is produced in the culture by TPL is depleted by the expression of TPL based on its high number of tyrosine residues and constitutive expression, leaving an inadequate amount of fluorotyrosine available for protein incorporation once the expression of BRD4(D1) is induced. To examine this, we attempted to boost ^{19}F incorporation into BRD4(D1) by supplementing the 3FY that was produced enzymatically with additional racemic 3FY (Figs. S37−39). We were unable to increase BRD4(D1) labeling efficiency under these conditions, although TPL reached near complete labeling. In addition, using a higher fluorophenol concentration to match the levels used in our fluorotyrosine isolation experiments led to reduced overall cell mass and protein production (Fig. S4). We also hypothesize that the pH of the expression media, which was inspired by the previously determined optimal value of 8 for the whole-cell tyrosine synthesis via TPL [36], is too high in comparison to the standard expression conditions of BRD4(D1) at a pH of 7.4. Future investigations will optimize the phenol concentrations and expression conditions needed to produce more highly fluorinated BRD4(D1) in higher yields.

Despite lower and inconsistent fluorine-labeling of BRD4(D1), PrOF NMR spectra of the isolated 3FY-BRD4(D1) and 2FY-BRD4(D1) provided further evidence of ^{19}F protein labeling (Fig. 4, S26). Comparison of the isolated protein's spectra to that reported for 3FY-BRD4 led to a similar spectrum (Fig. 4). The protein spectra were also responsive to ligand binding. In the case of 3FY-labeled BRD4(D1), a PrOF NMR ligand titration using a small molecule inhibitor of BRD4(D1) (1) showed resonances corresponding to Y97 and Y98, both near the binding site, respond to increased ligand concentration, indicating that the labeled BRD4(D1) is functional (Fig. 4). However, the low signal to noise in the spectra results from low ^{19}F incorporation and heterogeneity. These results support a need for further optimization for future PrOF NMR applications.

3. Conclusion

Here we have shown the feasibility of a single-culture system for the production and metabolic labeling of proteins with fluorinated tyrosine from inexpensive fluorophenols (~$2/g, Sigma) within cells. In this study, three different fluorinated amino acids were produced in addition to two clickable amino acid at low levels. For subsequent protein labeling, the lower level of fluorine incorporation may be sufficient for some NMR applications where fractional labeling is needed for maintaining native protein integrity [16] or for increasing the stability of engineered proteins. Future studies will seek to optimize this system for higher fluorotyrosine incorporation to provide a cost-efficient alternative labeling strategy for applications that necessitate high levels of ^{19}F incorporation. Our strategy has the potential to be modified for other types of protein-labeling strategies which would use other substituted

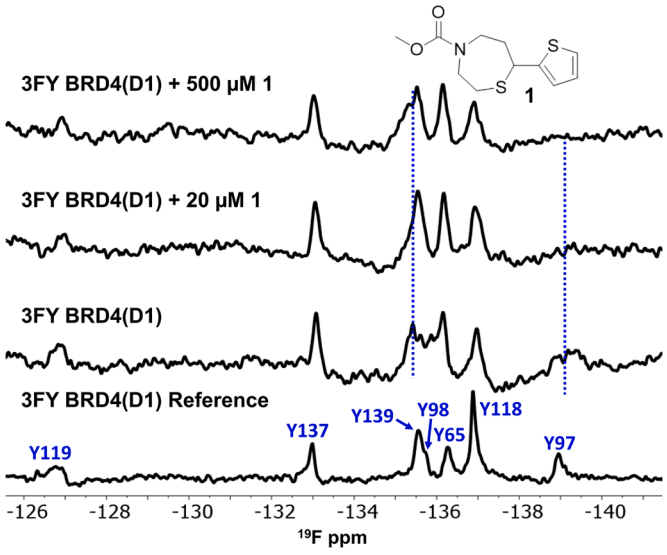

Fig. 4. PrOF NMR evidence of BRD4(D1) fluorination. PrOF NMR titration experiment of 3FY-BRD4(D1) isolated from our TPL expression system with thiazepane 1. The reported 3FY-BRD4(D1) spectrum is shown on bottom for comparison. Titration of 1 showed a slight response of resonances corresponding to Y97 and Y98, both of which are located in the protein-protein interaction binding site of BRD4(D1).

phenols to produce alternative unnatural and hard-to-access tyrosine analogues, such as our attempts with "click"-able tyrosines or ^{13}C-labeled phenols for C-F TROSY NMR methods [38]. Further aims of our approach also focus on optimizing the expression levels of TPL or introducing an inducible system to our plasmid constructs to turn off TPL expression when needed.

4. Experimental section

4.1. Plasmid construction and strain engineering

The tyrosine-phenol lyase coding sequence was amplified from pTZTPL [39]. 6xHis-TEV-BRD4 was amplified from pNIC24-BsaI-BRD4. In the case of BRD4, the coding sequence was amplified in two fragments, generating a synonymous C344T mutation to disable a SapI site to enable domestication of the CDS into our Golden Gate cloning pipeline. In both cases the amplified coding sequences were cloned into pMJS_CDS [40] via an AarI TypeIIS assembly reaction (Thermo). The BRD4 expression cassette includes a strong T7 promoter, and strong RBS. The expression cassette includes either a high or medium strength constitutive promoter and a medium strength RBS. All genetic parts, including their reported or calculated strengths and their DNA sequence, are reported in Supplementary Table S1. All promoters, RBS, and terminator parts were synthesized (IDT, Coralville IA). CDS were cloned into expression cassettes via a SapI TypeIIS assembly reaction. Expression cassettes containing CDS were combined into a medium-copy pMB1 based multi-cistron expression vector [40] via an AarI TypeIIS assembly reaction. All cloning and assembly steps were carried out in NEB Stable *E. coli*. *E. coli* BL21 (DE3) was transformed with the multi-cistron expression plasmid for fluorinated BRD4 incorporation experiments.

4.2. Whole cell transformation of fluorinated phenols to tyrosine derivatives

BL21*(DE3) or DL39(DE3) *E. coli* were transformed with the 9.29 plasmid and plated on LB agar plates containing kanamycin and incubated at 37°C overnight. Primary cultures were inoculated using single colony (5 mL LB media containing kanamycin) and grown at 37°C, 220 rpm overnight. To inoculate secondary cultures, 2 mL of primary culture was added to a 25 mL secondary LB culture containing kanamycin and grown at 37°C, 220 rpm while monitoring the OD_{600}. When the cell density reached an OD_{600} of 0.6−0.8, the cultures were harvested by centrifugation (6000 x *g*, 10 min, 4°C) and the cell pellets resuspended in 25 mL of defined media[24] with a pH of 8 containing 25 mM NH_4Cl, 25 mM Pyruvate, 25 mM Pyridoxal 5-Phosphate, and 25 mM of substituted phenol. The whole-cell reaction was allowed to proceed at 40°C, 225 rpm for 2 h after which the cells were pelleted by centrifugation (8000 x *g*, 8 min, 4°C). The supernatant was removed and diluted 5x with LC-MS grade MeOH. The remaining cell pellet was resuspended in 2 mL LC-MS grade MeOH and was incubated at 4°C overnight with rotation. The pellet extract was centrifuged to remove the cell debris and the organic layer was removed for further analysis. Both supernatant (extracellular) and pellet (intracellular) samples were then assessed via quantitative ^{19}F NMR and LC-MS.

4.3. Expression of BRD4(D1) using Dual TPL/BRD4(D1) plasmid system

BL21*(DE3) or DL39(DE3) *E. coli* strains were transformed with the 9.29 plasmid and plated on LB agar plates containing kanamycin and incubated at 37°C overnight. Primary cultures were inoculated using single colony (5 mL LB media containing kanamycin) and grown at 37°C, 220 rpm overnight. To inoculate secondary cultures, 5 mL of primary culture was used for every 250 mL LB media containing kanamycin. These cultures grown at 37°C, 220 rpm while monitoring the OD_{600}. When the cell density reached an OD_{600} of 0.6−0.8, the cultures were harvested by centrifugation (6000 x *g*, 10 min, 4°C) and the cell pellets were resuspended in the appropriate volume of defined media [24] with a pH of 8 containing 0.7 mM NH_4Cl, 0.7 mM Pyruvate, 0.4 mM Pyridoxal 5-Phosphate, and 0.7 mM of substituted phenol. The cultures were incubated at 37°C for 45 min, after which an additional 0.7 mM phenol and 0.7 mM pyruvate were added. After incubating 45 more minutes at 37°C, 220 rpm, 0.35 mM phenol and 0.35 mM pyruvate were added, and the temperature was reduced to 20°C. After 30 additional minutes at 220 rpm, 0.5 mM IPTG was added to the cultures to induce BRD4(D1) expression. After 18 h, the cultures were harvested by centrifugation (8000 x *g*, 8 min, 4°C) and the pellets were stored at -80°C until purification.

4.4. ^{19}F NMR quantification of fluorotyrosine production

A 436.5 μL sample of either supernatant or pellet sample was diluted with 110 μL D_2O and 4.4 μL 125 mM 5-fluoroindole in DMSO to serve as an internal standard (1 mM final concentration). All ^{19}F NMR spectra were obtained on a Bruker Avance III HD 500 (HD-500) (NS=64, D1=10, O1P=-120, SW=100, AQ=0.5).

4.5. LC-MS quantification of tyrosine derivative production

Standard curves were made from either dissolving 2- or 3- fluorinated tyrosine into 100% LC-MS grade water at a concentration of 1 mg/mL. From this stock, concentrations from 0.01 to 0.1 mg/mL were made with a final ratio of 50:50 water:acetonitrile and 2 μL injections were used. Analysis was performed on an Agilent 1290 UPLC with an Agilent Zorbax Eclipse Plus HILIC column 1.8 μm (2.1 mm x 50 mm) coupled to an Agilent 6540 UHD QTOF. A 5 min gradient was performed as follows: 0−1 min isocratic 100% B, 1−4 min 100-50% B, 4-5 min isocratic at 50% B. Flowrate was held at 0.4 mL/min. A = 95:5% LC-MS grade water:acetonitrile (0.1% ammonium acetate, adjusted to pH 4 with formic acid), B = 95:5% LC-MS grade acetonitrile:water, (0.1% ammonium acetate, adjusted to pH 4 with formic acid). Mass spectra were collected in profile mode over the range *m/z* 100−1700 with a scan rate of 5 spectra/s. MS parameters in positive electrospray ionization mode were as follows: vcap, 3500 V; nozzle voltage, 0 V; fragmentor voltage, 100 V; drying gas temp, 325°C; drying gas flow, 10 L/min; sheath gas temp, 400°C; sheath gas flow, 12 L/min; nebulizer pressure; 20 psig. MzMine2 was used to process the data and calculate areas under the curve.

4.6. Anion exchange chromatography

Culture pellets (100 mL) were resuspended in 10 mL Lysis buffer (50 mM Phosphate, 300 mM NaCl, pH=7.4) and 0.5 mg/mL PMSF and lysed by sonication (45% amplitude 30 s on, 60 s off for 12 cycles). Cell debris was removed by centrifugation at 10,000 x *g* for 10 min at 4°C. After gravity filtration, the supernatant was loaded onto a Zeba™ Spin Desalting Column, 40K MWCO, and desalted into Ion Exchange equilibration buffer (20 mM Bis-Tris, pH=6.9). The desalted cell lysate was then loaded onto an AKTAgo 2721705 FPLC and underwent anion exchange chromatography using a Cytiva HiTrap Q FF 1 mL column with a two-step gradient: 0−40% B Elution buffer (1 M NaCl, 20 mM Bis-Tris, pH=6.9) over 15 CV followed by 40−100% B over 10 CV. After verification of protein purity via SDS-PAGE, fractions were collected and underwent concentration and desalting. Protein fractions were concentrated in a 500 μL Amicon Ultra filter, MWCO 50 kDa, at 14,000 × *g* for 4 min and buffer exchanged into Milli-Q water by 2-3 wash/spin cycles. Samples were then submitted for LC-MS analysis on the LTQ instrument.

4.7. LC-MS analysis of ^{19}F incorporation into BRD4(D1) and TPL

For UPLC-MS analysis of proteins, a Waters Acquity UPLC coupled to a Waters Synapt G2 HDMS quadrupole orthogonal acceleration time-of-flight mass spectrometer was used (Waters Corp., Milford, MA USA). A

Waters Acquity UPLC Protein BEH C4 2.1 mm x 100 mm column (1.7 μm diameter particles) at 50°C was used for the following 30 min linear gradient separation at a flow rate of 0.400 mL/min using A: water containing 0.1% formic acid and B: methanol containing 0.1% formic acid: 3% B, 0 min to 5 min; 3% B to 97% B, 5 min to 15 min; 97% B, 15 min to 16 min; 97% B to 55% B, 16 min to 17 min; 55% B 17 min to 24 min.; 55% B to 3% B, 24 min to 25 min. Mass spectra were collected in profile mode over the range m/z 300−2500 every 0.1 s during the chromatographic separation. MS parameters in positive electrospray ionization mode were as follows: capillary, 0.6 kV; sampling cone, 30.0 V; extraction cone, 5.0 V; desolvation gas flow, 800 L/h; source temperature, 100°C; desolvation temperature, 350°C; cone gas flow, 20 L/h; trap CE, off. Lockspray (on-the-fly mass calibration) configuration consisted of infusion of a 5 mg/mL solution of leucine-enkephalin and acquisition of one mass spectrum (0.2s scan, m/z 50−1200) every 10 s. Three lockspray m/z measurements of protonated (positive ionization mode) leucine-enkephalin were averaged and used to apply a mass correction to measured m/z values during the course of the analysis.

4.8. UHPLC-LTQ XL ion trap MS^n

Each protein solution (18 μL) was analyzed by LC-MS on a Thermo Scientific™ LTQ XL™ linear ion trap mass spectrometer equipped with a Dionex Ultimate 3000 HPLC system. Proteins were separated on a Zorbax SB C18 300A column (5 μM, 150 × 0.5 mm, Agilent, Santa Clara, CA) using a gradient of 5− 95% B over 25 min (15 μL/min, A = 0.1% aqueous formic acid, B = 0.1% formic acid in acetonitrile). The mass spectrometer was equipped with a HESI source and operated in full scan, positive ion mode with a capillary temperature of 275°C, the sheath gas set to 8, the spray voltage set at 5V and the tube lens set to 100. All raw m/z data was deconvoluted using ProMass for Xcalibur software according to its large protein parameters.

Funding

We acknowledge support for this research from the NIH (R35 GM140837-02). SCH was supported by the Doctoral Dissertation Fellowship of the UMN Graduate School. DM was supported by the UMN Chemistry Department. JAJ and NMO were supported by a Biotechnology Training Grant: NIH T32GM008347. NMO was also supported by the Wayland E. Noland Excellence Fellowship of the UMN Chemistry Department.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

LTQ LC-ion trap mass spectrometric analysis was performed in the Analytical Biochemistry Shared Resource of the Masonic Cancer Center with support from NIH grant P30-CA077598.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jfluchem.2022.110014.

References

[1] C.R. Buchholz, W.C.K. Pomerantz, 19 F NMR viewed through two different lenses: ligand-observed and protein-observed 19 F NMR applications for fragment-based drug discovery, RSC Chem. Biol. 2 (2021) 1312–1330, https://doi.org/10.1039/d1cb00085c.
[2] K.E. Arntson, W.C.K. Pomerantz, Protein-observed fluorine NMR: a bioorthogonal approach for small molecule discovery, J. Med. Chem. 59 (2016) 5158–5171, https://doi.org/10.1021/acs.jmedchem.5b01447.
[3] A. Divakaran, S.E. Kirberger, W.C.K. Pomerantz, SAR by (Protein-Observed) 19 F NMR, Acc. Chem. Res. 52 (2019) 3407–3418, https://doi.org/10.1021/acs.accounts.9b00377.
[4] S. Wagner, B. Sudhamalla, P. Mannes, S. Sappa, S. Kavoosi, D. Dey, S. Wang, K. Islam, Engineering bromodomains with a photoactive amino acid by engaging 'Privileged' tRNA synthetases, Chem. Commun. 56 (2020) 3641–3644, https://doi.org/10.1039/C9CC09891G.
[5] N.S. Joshi, L.R. Whitaker, M.B. Francis, A three-component Mannich-type reaction for selective tyrosine bioconjugation, J. Am. Chem. Soc. 126 (2004) 15942–15943, https://doi.org/10.1021/ja0439017.
[6] P. Agarwal, J. Van Der Weijden, E.M. Sletten, D. Rabuka, C.R. Bertozzi, A Pictet-Spengler ligation for protein chemical modification, Proc. Natl. Acad. Sci. U. S. A. 110 (2013) 46–51, https://doi.org/10.1073/pnas.1213186110.
[7] G. Fittolani, E. Shanina, M. Guberman, P.H. Seeberger, C. Rademacher, M. Delbianco, Automated glycan assembly of 19F-labeled glycan probes enables high-throughput nmr studies of protein-glycan interactions, Angew. Chem. Int. Ed. 60 (2021) 13302–13309, https://doi.org/10.1002/anie.202102690.
[8] J.T. Gerig, NMR Fluorine, SelfPublished. 21 (2001) 1–35. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1181057&tool=pmcentrez&rendertype=abstract.
[9] J.L. Kitevski-LeBlanc, R.S. Prosser, Current applications of19F NMR to studies of protein structure and dynamics, Prog. Nucl. Magn. Reson. Spectrosc. 62 (2012) 1–33, https://doi.org/10.1016/j.pnmrs.2011.06.003.
[10] G. Xiao, J.F. Parsons, K. Tesh, R.N. Armstrong, G.L. Gilliland, Conformational changes in the crystal structure of rat glutathione transferase M1-1 with global substitution of 3-fluorotyrosine for tyrosine, J. Mol. Biol. 281 (1998) 323–339, https://doi.org/10.1006/jmbi.1998.1935.
[11] D. Alexeev, P.N. Barlow, S.M. Bury, J.D. Charrier, A. Cooper, D. Hadfield, C. Jamieson, S.M. Kelly, R. Layfield, R.J. Mayer, H. McSparron, N.C. Price, R. Ramage, L. Sawyer, B.A. Starkmann, D. Uhrin, J. Wilken, D.W. Young, Synthesis, structural and biological studies of ubiquitin mutants containing (2S, 4S)-5-fluoroleucine residues strategically placed in the hydrophobic core, ChemBioChem 4 (2003) 894–896, https://doi.org/10.1002/cbic.200300699.
[12] B.D. Giaimo, F. Ferrante, A. Herchenröther, S.B. Hake, The histone variant H2A. Z in gene regulation, Epigenet. Chromatin (2019) 1–22, https://doi.org/10.1186/s13072-019-0274-9.
[13] O. Schueler-Furman, Y. Altuvia, H. Margalit, Differential effects of isomeric incorporation of fluorophenylalanines into PvuII endonuclease, Proteins Struct. Funct. Genet. 45 (2001) 55–61, https://doi.org/10.1002/prot.1123.
[14] D. O'Hagan, H.S. Rzepa, Some influences of fluorine in bioorganic chemistry, Chem. Commun. (1997) 645–652, https://doi.org/10.1039/a604140j.
[15] H.Y. Tae, A.J. Link, D.A. Tirrell, Evolution of a fluorinated green fluorescent protein, Proc. Natl. Acad. Sci. U. S. A. 104 (2007) 13887–13890, https://doi.org/10.1073/pnas.0701904104.
[16] J.L. Kitevski-Leblanc, F. Evanics, R. Scott Prosser, Optimizing 19F NMR protein spectroscopy by fractional biosynthetic labeling, J. Biomol. NMR. 48 (2010) 113–121, https://doi.org/10.1007/s10858-010-9443-7.
[17] A.A. Bogan, K.S. Thorn, Anatomy of hot spots in protein interfaces, J. Mol. Biol. 280 (1998) 1–9, https://doi.org/10.1006/jmbi.1998.1843.
[18] L. Wang, A. Brock, B. Herberich, P.G. Schultz, Expanding the genetic code of Escherichia coli, Science 292 (80) (2001) 498–500, https://doi.org/10.1126/science.1060077.
[19] K.V. Loscha, A.J. Herlt, R. Qi, T. Huber, K. Ozawa, G. Otting, Multiple-site labeling of proteins with unnatural amino acids, Angew. Chem. Int. Ed. 51 (2012) 2243–2246, https://doi.org/10.1002/anie.201108275.
[20] N.G. Sharaf, A.M. Gronenborn, Chapter Four-19F-Modified Proteins and 19F-Containing Ligands as Tools in Solution NMR Studies of Protein Interactions. Methods Enzymol, 1st ed., Elsevier Inc., 2015, pp. 67–95, https://doi.org/10.1016/bs.mie.2015.05.014.
[21] T.H. Kim, K.Y. Chung, A. Manglik, A.L. Hansen, R.O. Dror, T.J. Mildorf, D.E. Shaw, B.K. Kobilka, R.S. Prosser, The role of ligands on the equilibria between functional states of a G protein-coupled receptor, J. Am. Chem. Soc. 135 (2013) 9465–9474, https://doi.org/10.1021/ja404305k.
[22] Y. Hattori, D. Heidenreich, Y. Ono, T. Sugiki, K. ichi Yokoyama, E. ichiro Suzuki, T. Fujiwara, C. Kojima, Protein 19F-labeling using transglutaminase for the NMR study of intermolecular interactions, J. Biomol. NMR. 68 (2017) 271–279, https://doi.org/10.1007/s10858-017-0125-6.
[23] N. Carulla, C. Woodward, G. Barany, N. Carulla, G. Barany, C. Woodward, The preparation of 19 F-labeled proteins for NMR studies, Methods Enzymol 380 (2004) 400–415.
[24] C. Gee, K. Arntson, A.K. Urick, N.K. Mishra, L.M.L. Hawk, A. Wisniewski, W.C. K. Pomerantz, Protein observed 19F-NMR for fragment screening, affinity quantification and druggability assessment, Nat. Protoc. 11 (2016) 1414–1427.
[25] E. Lian, Chenyang. Le, Hongbiao. Montez, Bernard. Patterson, Jessica. Harrell, Shannon. Laws, David. Matsumura, Ichiro. Pearson, John. Oldfield, Fluorine- 19 nuclear magnetic resonance spectroscopic study, Biochemistry 33 (1994) 5238–5245.

[26] N.K. Mishra, A.K. Urick, S.W.J. Ember, E. Schonbrunn, W.C. Pomerantz, Fluorinated aromatic amino acids are sensitive 19f NMR probes for bromodomain-ligand interactions, ACS Chem. Biol. 9 (2014) 2755–2760, https://doi.org/10.1021/cb5007344.
[27] P.B. Crowley, C. Kyne, W.B. Monteith, Simple and inexpensive incorporation of 19F-Tryptophan for protein NMR spectroscopy, Chem. Commun. 48 (2012) 10681–10683, https://doi.org/10.1039/c2cc35347d.
[28] G.R. Winkler, S.B. Harkins, J.C. Lee, H.B. Gray, α-Synuclein structures probed by 5-fluorotryptophan fluorescence and19F NMR spectroscopy, J. Phys. Chem. B 110 (2006) 7058–7061, https://doi.org/10.1021/jp060043n.
[29] C. Li, G. Wang, Y. Wang, R. Creager-allen, E.A. Lutz, H. Scronce, K.M. Slade, R.A. S. Ruf, R.A. Mehl, G.J. Pielak, C. Hill, C. Hill, N. Carolina, Protein 19F NMR in Escherichia coli, J. Am. Chem. Soc. 132 (2010) 321–327.
[30] C. Li, E.A. Lutz, K.M. Slade, R.A.S. Ruf, G. Wang, G.J. Pielak, 19F NMR studies of -R synuclein conformation and fibrillation, Biochem. J. 48 (2009) 8578–8584, https://doi.org/10.1021/bi900872p.
[31] D. Milić, T.V. Demidkina, N.G. Faleev, R.S. Phillips, D. Matković-Čalogović, A. A. Antson, Crystallographic snapshots of tyrosine phenol-lyase show that substrate strain plays a role in C-C bond cleavage, J. Am. Chem. Soc. 133 (2011) 16468–16476, https://doi.org/10.1021/ja203361g.
[32] K.R. Ravichandran, L. Liang, J. Stubbe, C. Tommos, Formal reduction potential of 3,5-difluorotyrosine in a structured protein: Insight into multistep radical transfer, Biochemistry 52 (2013) 8907–8915, https://doi.org/10.1021/bi401494f.
[33] M.R. Seyedsayamdost, S.Y. Reece, D.G. Nocera, J.A. Stubbe, Mono-, di-, tri-, and tetra-substituted fluorotyrosines: New probes for enzymes that use tyrosyl radicals in catalysis, J. Am. Chem. Soc. 128 (2006) 1569–1579, https://doi.org/10.1021/ja055926r.
[34] T.I. Yang, W.L. Li, T.H. Chang, C.Y. Lu, J.M. Chen, P.I. Lee, L.M. Huang, L. Y. Chang, Adenovirus replication and host innate response in primary human airway epithelial cells, J. Microbiol. Immunol. Infect. 52 (2019) 207–214, https://doi.org/10.1016/j.jmii.2018.08.010.
[35] K. Temme, D. Zhao, C.A. Voigt, Refactoring the nitrogen fixation gene cluster from Klebsiella oxytoca, Proc. Natl. Acad. Sci. U. S. A. 109 (2012) 7085–7090, https://doi.org/10.1073/pnas.1120788109.
[36] S. Xu, Y. Zhang, Y. Li, X. Xia, J. Zhou, G. Shi, Production of L-tyrosine using tyrosine phenol-lyase by whole cell biotransformation approach, Enzyme Microb. Technol. 131 (2019), 109430, https://doi.org/10.1016/j.enzmictec.2019.109430.
[37] J.A. Johnson, Fragment screening and biophysical method development for BET and non-BET bromodomain inhibitor discovery, 2020.
[38] A. Boeszoermenyi, S. Chhabra, A. Dubey, D.L. Radeva, N.T. Burdzhiev, C. D. Chanev, O.I. Petrov, V.M. Gelev, M. Zhang, C. Anklin, H. Kovacs, G. Wagner, I. Kuprov, K. Takeuchi, H. Arthanari, Aromatic 19 F- 13 C TROSY: a background-free approach to probe biomolecular structure, function, and dynamics, Nat. Methods 16 (2019) 333–340, https://doi.org/10.1038/s41592-019-0334-x.
[39] H. Chen, P. Gollnick, R.S. Phillips, Site-directed mutagenesis of His343→Ala in citrobacter freundii tyrosine phenol-lyase: effects on the kinetic mechanism and rate-determining step, Eur. J. Biochem. 229 (1995) 540–549, https://doi.org/10.1111/j.1432-1033.1995.tb20496.x.
[40] S.Y. Hsu, D. Perusse, T. Hougard, M.J. Smanski, Semisynthesis of the neuroprotective metabolite, serofendic acid, ACS Synth. Biol. 8 (2019) 2397–2403, https://doi.org/10.1021/acssynbio.9b00261.