Journal of Biomolecular NMR (2020) 74:61–69
https://doi.org/10.1007/s10858-019-00290-0

ARTICLE

2-Fluorotyrosine is a valuable but understudied amino acid for protein-observed ^{19}F NMR

Peter D. Ycas[1] · Nicole Wagner[1] · Noelle M. Olsen[1] · Riqiang Fu[2] · William C. K. Pomerantz[1]

Received: 28 March 2019 / Accepted: 13 November 2019 / Published online: 23 November 2019

Abstract

Incorporation of ^{19}F into proteins allows for the study of their molecular interactions via NMR. The study of ^{19}F labeled aromatic amino acids has largely focused on 4-,5-, or 6-fluorotryptophan, 4-fluorophenylalanine, (4,5, or 6FW) or 3-fluorotyrosine (3FY), whereas 2-fluorotyrosine (2FY) has remained largely understudied. Here we report a comparative analysis with different fluorinated amino acids. We first report the NMR chemical shift responsiveness of five aromatic amino acid mimics to changes in solvent polarity and find that the most responsive, a mimic of 3FY, has a 2.9-fold greater change in chemical shift compared to the other amino acid mimics in aprotic solvents including the 2FY mimic. We also probed the utility of 2FY for ^{19}F NMR by measuring its NMR relaxation properties in solution and the chemical shift anisotropy (CSA) of a polycrystalline sample of the amino acid by magic angle spinning. Using protein-observed fluorine NMR (PrOF NMR), we compared the influence of 2FY and 3FY incorporation on stability and pK_a perturbation when incorporated into the KIX domain of CBP/p300. Lastly, we investigated the ^{19}F NMR response of both 2FY and 3FY-labeled proteins to a protein–protein interaction partner, MLL, and discovered that 2FY can report on allosteric interactions that are not observed with 3FY-labeling in this protein complex. The reduced perturbation to pK_a and similar but reduced CSA of 2FY to 3FY supports 2FY as a suitable alternative amino acid for incorporation into large proteins for ^{19}F NMR analysis.

Keywords ^{19}F NMR · Protein-observed NMR · Fluorotyrosine · Protein–protein interactions · Allostery

Introduction

^{19}F NMR is an enabling technique for studying the dynamic structural properties of biomolecules including protein folding, allostery, and protein–ligand interactions (Danielson and Falke 1996; Arntson and Pomerantz 2016). In addition, both fluorinated small molecules and proteins have been used in screening campaigns for small molecule discovery (Arntson and Pomerantz 2016; Norton et al. 2016; Dalvit and Vulpetti 2018). The utility of ^{19}F NMR is due in part to a strong signal from the ^{19}F nucleus (83% sensitivity relative to ^{1}H), the absence of background signals in complex biological matrices, and a large chemical shift response based on changes in chemical environment (Arntson and Pomerantz 2016; Hull and Sykes 1975). PrOF NMR is a relatively fast method for detecting changes to small ^{19}F-labeled proteins, experiments can typically be completed in one to two minutes on a 11.7 T magnet with a ^{19}F-tuned cryoprobe (Urick et al. 2015; Hawk et al. 2016). PrOF NMR can provide quantitative affinity measurements for weak binding ligands and inform on binding site location and allosteric effects (Mishra et al. 2014; Song et al. 2007). In many cases, due to a similar atomic radius with ^{1}H, direct incorporation of ^{19}F into proteins (H to F substitution) only moderately perturbs protein structure. However, this is dependent on the type and location of the particular amino acid incorporated (Arntson and Pomerantz 2016).

Replacement of a single type of amino acid with a monofluorinated analog via metabolic labeling is an efficient method for labeling aromatic amino acids which are frequently found at the interface of protein–protein

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10858-019-00290-0) contains supplementary material, which is available to authorized users.

✉ William C. K. Pomerantz
wcp@umn.edu

[1] Department of Chemistry, University of Minnesota, 207 Pleasant St. SE, Minneapolis, MN 55455, USA

[2] National High Magnetic Field Lab, 1800 East Paul Dirac Dr., Tallahassee, FL 32310, USA

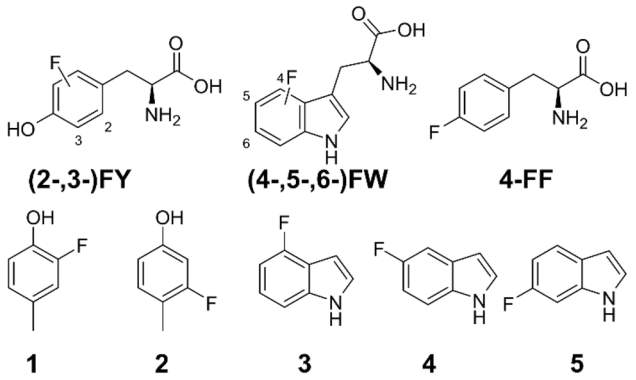

Fig. 1 The most commonly used fluorinated amino acids for PrOF NMR, 3-fluorotyrosine (3FY), 5-fluorotryptophan (5FW), and 4-fluorophenylalanine (4FF) are shown at the top. Below are small molecule mimics of fluorinated aromatic amino acids used in this study (**1**–**5**)

interactions, making this an attractive method for small molecule inhibitor discovery using ^{19}F NMR (Arntson and Pomerantz 2016; Mishra et al. 2014; Bogan and Thorn 1998). 3-Fluorotyrosine (3FY), 5-fluorotryptophan (5FW), and 4-fluorophenylalanine (4FF) are the most heavily used aromatic amino acids for metabolic labeling (Fig. 1). As such, their NMR properties have been well-characterized. Selection of fluorinated amino acids for PrOF NMR can be influenced by the dispersion of the ^{19}F resonances in a fluorine-labeled protein. In the case of 4,5, and 6-fluorotryptophan labeling of lactate dehydrogenase by Rule et al., the dispersion of resonances ranges varied from ~ 2.5 to 6 ppm with the largest chemical shift dispersion for 5FW (Rule et al. 1987). Alternatively, 4-fluorotryptophan-labeling of the apelin receptor led to the largest chemical shift dispersion over the other fluorinated isomers of tryptophan (Kenward et al. 2018). For the first bromodomain of BRD4, 3FY-labeled BRD4 has a chemical shift dispersion of 11 ppm, whereas the 5FW-labeled BRD4 has a dispersion of only 1.4 ppm (Mishra et al. 2014). More direct comparisons of the NMR properties of fluorinated amino acids within proteins will inform on the best choice of ^{19}F-labeled amino acid for PrOF NMR.

A less-well-characterized amino acid is 2-fluorotyrosine (2FY), which remains underutilized for ^{19}F NMR. A rare example was the use of 2FY as an NMR probe for allosteric interactions of MerR (Song et al. 2007). However, most studies to date have used 3FY as a ^{19}F NMR reporter, including the first NMR analysis of a fluorinated protein, the 86 kDa alkaline phosphatase (Hull and Sykes 1975). Here we carry out a comparative analysis of this amino acid with 3FY and provide information on its nuclear relaxation properties, chemical shift responsiveness, and chemical shift anisotropy to evaluate potential advantages of using 2FY-labeled proteins in PrOF NMR experiments.

To evaluate chemical shift responsiveness in the absence of protein structure, we first measured the responsiveness of the fluorine chemical shift to solvents of differing polarity for five different aromatic amino acid mimics of tyrosine and tryptophan (**1–5**, Fig. 1). We next measured ^{19}F T_2*, and T_1 of (**1**–**5**) in water and an 80% glycerol solution to mimic a 12 kDa protein. Relaxation due to the chemical shift anisotropy (CSA) also leads to broadening of ^{19}F resonances (Gerig 2001), therefore, we measured the ^{19}F CSA of a polycrystalline sample of 2FY to compare with 3FY by solid state ^{19}F NMR. Previous studies by Ulrich have determined the CSA properties of a number of commonly incorporated fluorinated aromatic amino acids and determined reduced anisotropy values ranging from 47.1 ppm for 5FW to − 74.6 ppm for 3FY (Durr et al. 2008). These studies did not include measurements of 2FY, which we believe could be a valuable ^{19}F nucleus for PrOF NMR and is studied here. Finally, to guide researchers in selecting amino acids for PrOF NMR, we determined the pK_a of both amino acids and response to ligand binding when incorporated into a 12 kDa protein using ^{19}F NMR.

Results and discussion

We examined fluoroindoles, and fluorocresols which mimic the side chains of fluorinated tryptophans and tyrosines respectively (Rule et al. 1987). 4,5, and 6-fluoroindole **(3–5)** and 2, and 3-fluorocresol **(1** and **2)** were dissolved in solvents of increasing polarity and their change in chemical shift (Δδ) are presented in Fig. 2. These mimics were selected over the amino acids to enable solubility across

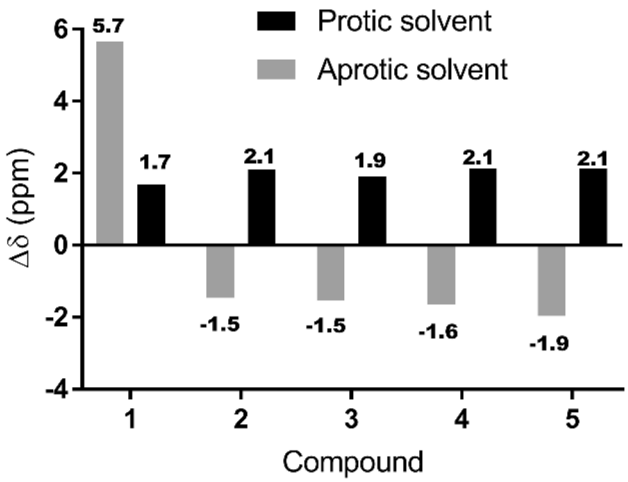

Fig. 2 Δδ of each ^{19}F nucleus in response to changes in solvent polarity, aprotic response is measured as (δacetone–δhexanes), and protic solvents are measured as (δH_2O-δmethanol). 3FY mimic, **1,** exhibits a 3-fold greater chemical shift change relative to the next largest chemical shift perturbation

a range of solvents. Upon conducting these experiments, we observed differences in chemical shift trends between conditions using polar protic solvents and aprotic solvents and thus separate our analysis. The small molecule mimic of 3FY, (2-fluoro-*p*-cresol, **1**), demonstrated the largest change in chemical shift ($\Delta\delta = 5.7$ vs. 1.7–2.1 ppm for the other analogs) in aprotic solvents upon increasing solvent polarity from hexanes to DMSO. Previous computational analysis by Isley et al. suggests this responsiveness is due to the differences in the *s*-trans and *s*-cis conformations of the phenolic O–H leading to calculated differences in dispersions of greater than 10 ppm (Isley et al. 2016). Such large changes can occur from dipolar interactions with the fluorine atom and the hydroxyl group in the *s*-cis conformation. 2FY would be less sensitive to such an effect due to the position of the fluorine on the aromatic ring which is farther from the hydroxyl group. This result is also consistent with analyses from Dalvit et al. showing fluorine chemical shifts of *o* and *p*-fluorophenols (but not *m*) are particularly sensitive reporters of non-covalent interactions based on resonance and field/inductive effects (Dalvit et al. 2014). Large differences in chemical shift were not observed when **1** was tested in protic solvents of increasing polarity from methanol to H_2O resulting in a $\Delta\delta$ of 1.7 ppm. In fact, this was the smallest chemical shift change measured. This may be due to a screening of the intramolecular interaction of the hydroxyl group with the ortho fluorine.

The speed and signal-to-noise of a ^{19}F NMR spectrum is determined by the relaxation properties of the ^{19}F nucleus, leading us to examine the relaxation properties of different amino acids mimics. A short T_1 allows for a shorter experiment time while a long T_2 results in narrower resonances leading to an increase in signal-to-noise. The T_2 is comprised of relaxation due to dipole–dipole interactions and relaxation due to CSA. Fluorine-proton dipole relaxation is governed by the local environment around the ^{19}F nucleus as well as the rotational correlation times (τ_C). T_2*, an estimate of T_2, can be determined by measuring the linewidth at half height ($\nu_{1/2}$) of the resonance according to Eq. 1 and accounts for local inhomogeneities in the magnetic field.

$$T_2^* = \left(\pi \nu_{1/2}\right)^{-1} \tag{1}$$

The NMR relaxation properties of the fluorinated amino acid mimics, **1–5** were investigated by measuring T_1, and T_2* in aqueous solutions as well as in 80% glycerol to mimic their spectral behavior in a 12 kDa protein (Ye et al. 2015). These results are reported in Table 1. The T_1 values between the cresols and indoles dissolved in 80% glycerol ranges from 400 to 800 ms, a corresponding d1 increase to ~ 400 ms in a 500 scan NMR experiment corresponds to an increase in only about 3 min of experiment time between the two extremes. T_2* correlate inversely with the width of the ^{19}F resonance,

Table 1 ^{19}F NMR properties of fluorinated amino acid mimics

Small molecule	δ (ppm)	80% Glycerol		Aqueous	
		T_2*	T_1	T_2*	T_1
1	− 139.0	**15**	**0.5**	106	3.0
2	− 117.0	**64**	**0.4**	106	3.0
3	− 123.8	**53**	**0.5**	80	3.8
4	− 126.6	**32**	**0.8**	160	6.3
5	− 123.2	**40**	**0.7**	80	4.0

T_2* (ms) and T_1 (s) are relaxation in the transverse and longitudinal direction, respectively. Values in bold were measured in an 80% glycerol solution to simulate the environment of a 12 kDa protein, values in black were measured in an aqueous environment. T_1 and T_2^*measurements were made at 11.74 T by inversion recovery and with a ^{1}H decoupled pulse sequence, respectively

these values are largest for the 2FY mimic, **2**. We conclude that 3FY may ultimately be more responsive to changes in chemical environment leading to better resonance dispersion for more buried resonances or manifested in larger responses from binding of hydrophobic ligands in an NMR screen. Alternatively, 2FY may lead to less broadened resonances from longer T_2 relaxation. Such an effect supports including 2FY as a complimentary probe to 3FY for use in PrOF NMR.

We further investigated the effects of resonance linewidth by determining the chemical shift anisotropy of 2FY. Due to the large effect of CSA on T_2 relaxation in ^{19}F at high field strengths it is important to consider these effects in the context of PrOF NMR. The majority of ^{19}F NMR tyrosine studies have used 3FY, and the CSA values of this amino acid have previously been measured (Arntson and Pomerantz 2016; Hull and Sykes 1975). Given the larger T_2* value determined for the 2FY mimic, **2**, in the 80% glycerol solution, versus the 3FY mimic, **1,** we anticipated a lower CSA. We measured CSA values of 2FY by proton-decoupled solid-state NMR with magic angle spinning. Tensor values were determined using the software HBA (Eichele 2015). The isotropic (δ_{iso}) peak, the anisotropy (Δ), and the asymmetry (η) are defined as described in Eqs. 2–4.

$$\delta_{iso} = \frac{\delta_{11} + \delta_{22} + \delta_{33}}{3} \tag{2}$$

$$\delta_{11} - \delta_{iso} = \Delta \text{if} \left|\delta_{11} - \delta_{iso}\right| > \left|\delta_{33} - \delta_{iso}\right|, \quad \text{otherwise} \delta_{33} - \delta_{iso} = \Delta \tag{3}$$

$$\eta = \frac{\delta_{22} - \delta_{11}}{\delta_{33} - \delta_{iso}} \tag{4}$$

The values presently measured for 2FY and 3FY along with previously measured values by Ulrich and co-workers are displayed in Table 2 (Durr et al. 2008). The isotropic chemical shifts were similar to solution state measurements

Table 2 CSA tensor properties measured for polycrystalline (D/L) 2FY and 3FY

Amino acid	δ_{11} (ppm)	δ_{22} (ppm)	δ_{33} (ppm)	δ_{iso} (ppm)	Δ (ppm)	η
2FY	−53.1	−96.8	−179.6	−109.9	−69.8 ± 2.3	0.63 ± 0.05
3FY	−78.1	−116.2	−212.2	−135.5	−76.6 ± 1.3	0.49 ± 0.02
4FF*	−51.9	−118.4	−162.0	−110.8	58.9	0.74
5FW*	−77.5	−148.2	−148.2	−124.6	47.1	0
3FY*	−79.9	−115.7	−209.7	−135.1	−74.6	0.48

Literature reported values for the commonly incorporated 4FF, 5FW, and 3FY are included for reference. Data was collected at 14.1 T spinning at 12.5 kHz to determine the isotropic position

*Values were obtained from Durr et al. (2008)

for both samples of 2FY and 3FY. For the polycrystalline sample of 3FY both the anisotropy and asymmetry values were also similar to literature reports. In the case of 2FY, the anisotropy relative to 3FY was slightly reduced supportive of CSA effects reducing the observed linewidth of this amino acid and provides further evidence that 2FY may be suitable for incorporation into a large protein. The asymmetry values of both 2FY and 3FY measured in the solid state are > 0.4, suggesting restricted side chain motion in the crystalline material (Schanda et al. 2011). However, Durr et al., pointed out that the conditions used for crystallization of the samples can produce significant variability in this value (Durr et al. 2008).

The studies above test the innate responsiveness and relaxation properties of the amino acids. To further test the proposal that a lower anisotropy of 2FY was sufficient to produce narrower fluorine resonances in proteins and to evaluate the chemical shift responsiveness of 3FY and 2FY, we incorporated both amino acids into the 12 kDa KIX domain of CBP/p300. The KIX domain is involved in protein–protein interactions with several transcription factors, notably the oncogenic proteins: the mixed lineage leukemia (MLL) protein, c-Myb, and phosphorylated kinase-inducible-domain (pKID). KIX has previously been studied by ^{19}F NMR using 3FY and 4FF labeling and thus serves as a useful model protein for anlaysis (Pomerantz et al. 2012).

^{19}F NMR spectra of 2FY- and 3FY-labeled KIX were obtained to make these comparisons (Fig. 3). Fluorine-labeled proteins were produced using standard metabolic labeling methods in the DL39(DE3) auxotrophic cell line (Frieden et al. 2004). The label incorporation of both 3FY and 2FY were consistently high (> 90%) as determined by ESI–MS. Under the same expression conditions, 3FY tended to yield a higher amount of protein relative to 2FY (14 vs. 8 mg/L), respectively. ^{19}F NMR resonances for 3FY-KIX were previously assigned through a combination of site directed mutagenesis of Y to F mutations for Y658, Y650, and Y640. Y631 and Y649 could not be determined with mutagenesis due to poor protein expression so a direct binding analysis of MLL was used to assign Y631. Y649 was assigned using solvent accessibility determination by titrating Gd·EDTA and observing a broadening of solvent exposed residues (Pomerantz et al. 2012). The same set of experiments were performed on 2FY-KIX to assign the observed resonances. These experiments are described further in the supporting information.

In the spectrum of 2FY-KIX only three of the five resonances could be observed. The dispersion of resonances in

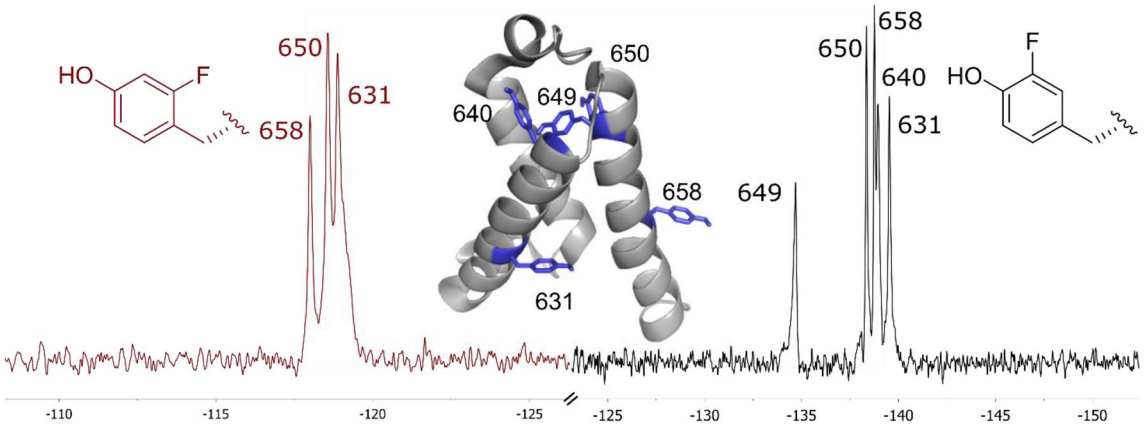

Fig. 3 PrOF NMR spectra of 2FY (maroon) and 3FY KIX (black). We hypothesize that the broader resonances in 3FY-KIX (649 and 640) are not observed in 2FY KIX due to a restricted rotation of the sidechains. PDB ID: 2AGH

Table 3 T_2 and T_1 relaxation properties of 3FY and 2FY KIX

Resonance	3FY, T_2 (ms)	3FY, T_1 (s)	2FY, T_2 (ms)	2FY, T_1 (s)
649	5.4	0.3	–	–
650	15.1	0.6	3.1	0.6
658	14.5	1.1	5.4	1.2
640	9.9	1.1	–	–
631	8.8	0.6	3.9	0.7
Whole Protein	9.9	0.8	3.5	0.8

T_1 measurements were made using inversion recovery, T_2 measurements were made using a CPMG pulse sequence. Spectra were acquired at 14.1 T, 0.3 Hz line broadening was applied

2FY-KIX is 1.4 ppm compared to the 3FY which is 5.6 ppm; most resonances are grouped within 1.6 ppm of each other in 3FY-KIX with the exception for Y649. This result is consistent with our model study above as Y650, Y631, and Y658 are solvent exposed while Y649 is significantly buried in KIX as determined by solvent accessibility studies and structural analysis (Hawk et al. 2016; Pomerantz et al. 2012). T_2 measurements of 2FY and 3FY-KIX indicate that CSA is not the dominant factor in linewidth as even though the anisotropy of 2FY-KIX is expected to be less than 3FY-KIX, the measured T_2 values of 2FY-KIX are shorter and the observed linewidth is broader (Table 3). The increase in line broadening in 2FY may be why Y649 and Y640 are not observed in the 2FY-KIX spectrum. The broadening of these resonances is a sum of dipole–dipole relaxation and chemical shift anisotropy effects which contribute to the T_2 term (Gerig 2001). However, a slow chemical exchange process from restricted rotation can also lead to resonance broadening. Given the results of the T_2* studies on the fluorocresols, **1** and **2**, and the solid state NMR analysis of 3FY and 2FY, this indicates that the fluorine nuclei of Y649 and Y650 have more restricted side chain rotation in 2FY compared to 3FY KIX.

The impact of fluorination may also change the stability of the protein. Previous work has noted small perturbations of fluorination on stability (e.g., a 2.8 °C decrease in thermal stability of 3FY-BRD4-BD1) (Mishra et al. 2014), but this is expected to be protein dependent. In fact, we previously measured an unexpected increase in stability of ~ 9 °C from unlabeled to 3FY-labeled KIX. To test the impact of fluorination via 2FY labeling on the stability of KIX, thermal melt curves of the protein were obtained using circular dichroism spectroscopy, measuring the absorption of circularly polarized light at 222 nm while increasing temperature.

The thermal melt curves of unlabeled, 2FY, and 3FY-KIX are shown in Fig. 4 obtained under the same conditions. The midpoint of each curve is the thermal melt temperature (T_m), in the unlabeled protein the T_m was measured at 63 ± 2 °C, for 3FY this point is 71 ± 1 °C. These values track well with

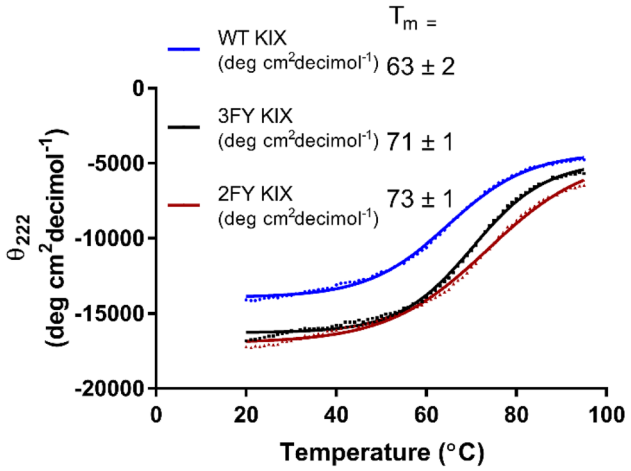

Fig. 4 Thermal melt curves of unlabeled, 3FY- and 2FY-KIX. 2FY and 3FY are 8°–10° more stable compared to unlabeled KIX. These values represent an average of three independent experiments

published values of 63–65 °C and 72.9–74.3 °C for unlabeled and 3FY-labeled protein respectively (Pomerantz et al. 2012; Gee et al. 2018). The thermal stability of 2FY-labeled KIX is 73 ± 1 °C. This difference in stabilization is within error of 3FY-KIX. Additionally, the slope of unfolding was lower for both 2FY and 3FY-KIX relative to the unlabeled protein but were within error relative to one another. These results support an increase in stability. To further probe this effect, ^{1}H-^{15}N HMQC data of the unlabeled, 3FY, ^{15}N-, and 2FY-^{15}N labeled protein indicate no large changes in the general resonance dispersion of the protein (Figure S16), supporting a more subtle structural effect for enhancing stability. However, we do note a reduction in the intensity of several resonances supporting a change in the dynamics of both 3FY- and 2FY-KIX relative to the unlabeled protein. The general increase in stability from fluorination may be due to an equal contribution from each tyrosine or a large contribution from primarily one. At present the contribution from fluorinated amino acid stabilization cannot be directly understood from these studies.

Another consideration for amino acid selection beyond magnetic resonance parameters are physicochemical perturbations incurred by fluorination. This is particularly relevant for tyrosines, whose side chain pK_a can be significantly altered upon fluorination. We used ^{19}F NMR to measure the impact of fluorination on the pK_a of each observable tyrosine resonance from 3FY and 2FY incorporation into KIX. The electron withdrawing character of fluorine is known to reduce the pK_a of tyrosines; however, we were interested in if this was a uniform drop across all tyrosine residues, or if there were more dramatic decreases in particular sites in the protein. The pK_a values for each fluorinated tyrosine phenol were determined by measuring the change in chemical shift as a function of pH and fitting this relationship with a

sigmoidal function (e.g. Y631, Fig. 5), these pK_a values are listed in Table 4.

These pK_a values are in close agreement with pK_a values previously measured for the free amino acids (2FY = 9.2, 3FY = 8.4) with 2FY being closer to the $pK_a$of Y (pK_a = 10.1) though there is a distinct drop in pK_a of Y631 relative to the free amino acid in both 2FY and 3FY of between 0.3 and 0.6 pK_a units respectively (Seyedsayamdost et al. 2006; Faleev et al. 2003). Such a pK_a perturbation can be significant, as the population of the conjugate base of tyrosine may be significantly increased at physiological pH, particularly for 3FY. Such a potential perturbation can be mitigated with 2FY.

As a final test for using our 2FY-labeled protein, functional perturbations of the proteins resulting from fluorination was investigated by titration with a native peptide ligand from MLL (Fig. 6). This is a small peptide fragment from the transcriptional activation domain with a reported K_d of 2.8 µM (Goto et al. 2002). Both 2FY and the previously reported 3FY-KIX PrOF NMR spectra exhibited resonance effects in the slow exchange regime upon MLL binding which is expected for a ligand of this affinity (Pomerantz et al. 2012). Binding is detected at the Y631 resonance, as this residue is in the MLL site moving 890 Hz upfield in 2FY and downfield 612 Hz in 3FY KIX (Figure S15). Of further interest, alteration of the allosterically controlled binding site resonance for Y650 is only observed in 2FY (Y650 moves 108 Hz downfield).

MLL has been shown to be a positive allosteric regulator at the pKID binding site (improves binding by ~ 2 fold) (Goto et al. 2002). Because Y650 is involved in the interaction of pKID with KIX (Brüschweiler et al. 2013), it was expected to be perturbed by allosteric binding at the MLL site; however, with 3FY-labeled KIX, no resonance perturbation was observed in a titration with MLL. The lack of resonance perturbation at this site could be due to a lack of involvement of the aromatic amino acid in the allosteric network. Alternatively, 3FY-labeling of KIX may have led to stabilization of the protein or the reduction in the cooperativity of folding, such that the dynamic change was no longer observable. Fortunately, 2FY does not seem to suffer from the loss of allosteric reporting and as a result, via PrOF NMR, the Y650 bound state resonance begins to appear in the presence of MLL, indicating that the Y650 resonance is now able to report on allosteric effects from MLL binding and may more closely resemble the native protein than 3FY-KIX. Future binding studies are required to further probe the allosteric network

Table 4 pK_a values of phenolic proton measured by ^{19}F NMR of 2FY- or 3FY-labeled KIX and computational determination of unlabeled KIX

Tyrosine resonance	3FY pK_a	2FY pK_a	Y pK_a
Free amino acid	8.4 Seyedsayamdost et al. (2006)	9.2 Faleev et al. (2003)	10.0 Edsall et al. (1958)
Y649	7.1	–	12.8
Y650	7.2	9.3	10.2
Y658	5.7[a]	9.2	9.9
Y640	6.9	–	10.6
Y631	7.9	8.9	10.5

Unlabeled KIX pK_a values were determined using the protein preparation module of Maestro

[a]This pK_a is only an estimate based on a small an incomplete resonance response to the pH ranges studies

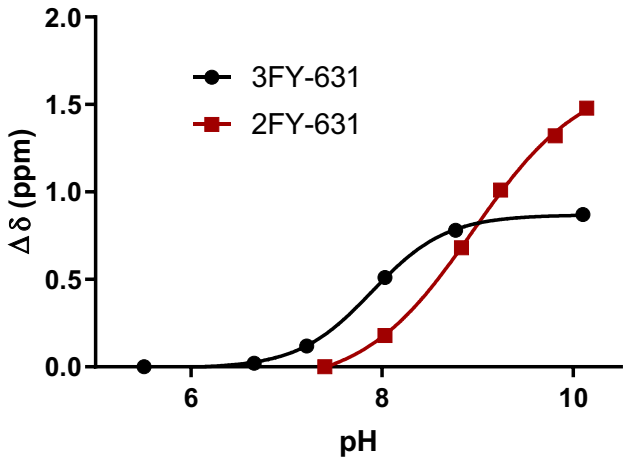

Fig. 5 Δδ of Y631 in 3FY- and 2FY-KIX vs pH. A sigmoidal curve was fit to determine the pK_a values of each phenolic proton (7.9 and 8.9, respectively)

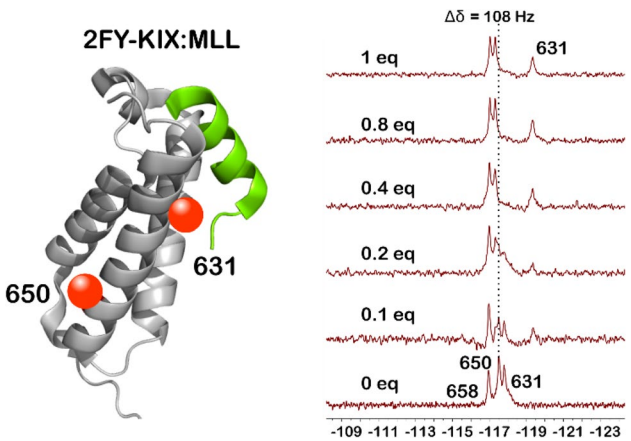

Fig. 6 Titration of the MLL peptide with 2FY-KIX. The resonance in the MLL binding site (Y631) broadens and reappears 890 Hz upfield in 2FY-KIX.Y650 which is a residue in the allosterically controlled binding site, moves downfield 108 Hz in 2FY, this allostery is not observed with 3FY-KIX (Pomerantz et al. 2012). Red spheres indicate resonances perturbed upon binding of KIX, blue spheres are not perturbed when bound. PDB ID: 2AGH

using 2FY-KIX. However, we anticipate that 2FY-labeled KIX will be a useful protein construct to study allosteric binding interactions of both native and synthetic ligands.

Conclusions

We have investigated the ^{19}F NMR properties of 4,5, and 6 fluoroindoles and 2, and 3 fluorocresols and compared their T_1, T_2*, spectral width, and responsiveness to change in their chemical shift. The change in chemical shift response was largest for the 3FY mimic, with a 2.9-fold larger dynamic range than all other fluorine labeled small molecules. Alternatively, the measured T_2* value of the 2FY mimic was larger than the other fluorinated aromatic amino acid mimics, and a slight reduction of CSA values of 2FY compared to 3FY indicate that 2FY resonances will be less affected by line-broadening when incorporated into high molecular weight proteins and measured at high field strengths, if the side chain rotation is not restricted. The pK_a values of each tyrosine in a 12 kDa KIX domain of the CBP/p300 protein were measured using ^{19}F NMR and were consistent with previously measured values for the free amino acids. 2FY consistently had a higher pK_a than 3FY at different sites within the protein serving as a better analog of native tyrosine residues. The increased thermal stability of both 2FY and 3FY relative to unlabeled KIX indicates a general fluorine stabilization effect of the protein, though this is not observed in all other proteins. Furthermore, titration studies with a native ligand (MLL) displayed the expected direct binding interactions at the orthosteric site, however, only 2FY KIX reported on the allosteric change at Y650. Together, these data suggest that 3FY-labeling may provide a more responsive signal in some cases; however, 2FY labeling may provide a more wild type protein behavior relative to 3FY with respect to the pK_a of the phenolic proton. At high field strengths in large proteins 2FY resonances may be narrower due to a reduced CSA if side chain rotation is not restricted. Recent advances in ^{19}F-^{13}C TROSY experiments have further reduced the detrimental effects of CSA using 3FY-labeled proteins. In this 2D-experiment, good resolution and signal can be obtained on large proteins (180-kDa), though this requires ^{13}C labeling of the amino acid side chain (Boeszoermenyi et al. 2019). Our studies here indicate that 2FY should also be a suitable amino acid for this technique and should be considered for ^{19}F labeling.

Experimental methods

^{19}F NMR relaxation properties determinations of cresols and indoles

Indoles or cresols were dissolved to a final concentration of 1–25 mM in H_2O. A capillary containing 25 µL 5 mM TFA in D_2O was inserted as an internal reference and locking solvent. Spectra were acquired using a proton-decoupled ^{19}F sequence. A total of 16 scans was taken at an O1P of −75 with a SW of 10 ppm, and a 0.5 s AQ for the TFA reference. The same parameters were used to acquire spectra of the fluorinated cresols and indoles, varying the O1P based on literature values. The indole/cresol spectra were referenced based on the internal trifluoroacetate reference of (−76.55 ppm). T_1 values were determined by inversion recovery on a Bruker Avance III operating at 471 MHz. The 90° pulse was determined, and the variable delay times were set at (0.01, 0.05, 0.1, 0.25, 0.5, 1, 2, 4, 8, 15 s) with a 3 s relaxation delay for the cresols in glycerol and a 15 s relaxation delay for all other small molecules. Acquisition times were 0.5–2.8 s, 896–3676 points were acquired using a sweep width of 1411–3063 Hz with a total of 16 transients. Zero filling was applied to 4096–16384 points. Automatic phase correction using MestreNova was applied along with a Whittaker smoother baseline correction zero filling was applied to 4096 points and linebroading of 0.3 Hz was applied.

T_2* values were determined with ^{1}H decoupled ^{19}F pulse sequence on a Bruker Avance III HD operating at 471 MHz. An acquisition time of 0.5–1 s and a sweep width of 9399 Hz was used with a spectral size of Zero filling was applied to 16384–32768 points, automatic phase correction using the region analysis function in MestreNova, a Whittaker smoother baseline correction was applied, linebroadening of 0.3 Hz was applied. T_{2*} values were determined from the resonance width at half height.

SS-NMR value determination of 2FY and 3FY tyrosine

^{19}F MAS NMR measurements were carried out on a Bruker Avance 600 MHz NMR spectrometer at the National High Magnetic Field Lab (NHMFL) operating at the Larmor frequencies of 564.68 and 600.13 MHz for ^{19}F and ^{1}H, respectively, using a Bruker 4.0 mm H/F/X triple-resonance MAS probe. The ^{19}F NMR signals were polarized by a 90° pulse length of 3 µs and then acquired under high power ^{1}H decoupling using the SPINAL decoupling sequence with a ^{1}H RF field of 83.3 kHz (Fung et al. 2000). 128 and 256 scans were used to accumulate the ^{19}F signals of 2FY and 3FY, respectively, with a recycle delay of 5 s. The sample spinning rate was controlled by a Bruker pneumatic MAS unit with a variation of ± 3 Hz. For each sample, the spectra under two sample spinning rates (12.5 and 13.0 kHz) were used to identify the isotropic positions (δ_{iso}). The ^{19}F chemical shift for flufenamic acid (FFA) was set to −61.5 ppm for reference. The chemical shift anisotropy Δ and the asymmetry (η) were obtained by fitting the sideband manifolds (Herzfeld and Berger 1980).

Protein expression procedure

Fluorinated CBP/p300 KIX was expressed by transforming the DNA containing plasmid into DL39 cells. Single colonies were picked and cultured in 20 mL of LB overnight at 25° shaking at 220 RPM. Fluorotyrosine incorporation was achieved according to published protocols (Gee et al. 2016). Briefly, these primary cultures were transferred to 1L of LB and shaken at 37° for 3–4 h, until an OD_{600} of 0.6–0.8 was achieved. The culture was centrifuged and the media was changed for defined media containing 50 mg of the D/L fluorinated tyrosine. The culture was equilibrated at 37° C for 90 min before cooling to 20° C for 30 min. Protein expression was induced by addition of IPTG at a final concentration of 1 mM for 16–20 h. Purified proteins were isolated using an NTA-column as previously described (Gee et al. 2016).

Protein T_m determination

WT, 3FY-, and 2FY-KIX was concentrated to 30 µM in 10 mM $Na_2HPO_4^{-3}$, 100 mM NaCl PBS. The melting temperature was determined by measuring the θ_{222} while increasing the temperature 1°/min from 20 to 95° C on a J-815 Jasco CD Spectrometer with a 0.1 cm path length. The resulting curve was fit with a sigmoidal function in GraphPad and the midpoint was used to determine the T_m. Reported T_ms are an average of three replicates.

^{19}F NMR of 2-,3-,fluorotyrosine labeled KIX

20–50 µM fluorine labeled protein in 50 mM HEPES, 100 mM NaCl at pH 7.4 was diluted to 5% D_2O, 0.0004% trifluoroacetate as an internal reference. Spectra were collected on a Bruker Avance NEO 600-MHz operating at 564.46 MHz with a 5 mm TCI cryoprobe using a 0.6 s relaxation delay, 0.05 s acquisition time, and 2000-3000 scans with a 40 ppm sweep width. Exponential line broadening of 15–20 Hz was applied. pK_a values of each tyrosine were determined by buffer exchanging protein into various pHs, monitoring the Δδ of each ^{19}F resonance and plotting the resulting curve to a sigmoidal function for each residue. T_1 and T_2 measurements were collected by inversion recovery and a CPMG pulse sequence, respectively. For these experiments a relaxation delay of 4 s for T_2 or 5 s for T_1 was used with a 0.1 s acquisition time. 256 transients were acquired for T_1 measurements and 512 transients were acquired for T_2. 568 points were acquired with zero filling up to 2048 with 10 Hz line-broadening.

Acknowledgements This work was supported by the National Science Foundation (CHE-00038929, CHE-1904071). High resolution solid-state ^{19}F NMR experiments were carried out at the National High Magnetic Field Lab (NHMFL) supported by the NSF Cooperative agreement No. DMR-1644779 and the State of Florida. NW acknowledges a NSF REU summer fellowship (Grant CHE-1359181).

References

Arntson KE, Pomerantz WCK (2016) Protein-observed fluorine NMR: a bioorthogonal approach for small molecule discovery. J Med Chem 59:5158–5171

Boeszoermenyi A et al (2019) Aromatic 19F-13C TROSY: a background-free approach to probe biomolecular structure, function, and dynamics. Nat Methods 16:333–340

Bogan AA, Thorn KS (1998) Anatomy of hot spots in protein interfaces. J Mol Biol 280:1–9

Brüschweiler S, Konrat R, Tollinger M (2013) Allosteric communication in the KIX domain proceeds through dynamic repacking of the hydrophobic core. ACS Chem Biol 8:1600–1610

Dalvit C, Vulpetti A (2018) Ligand-based fluorine NMR screening: principles and applications in drug discovery projects. J Med Chem 62:2218–2244

Dalvit C, Invernizzi C, Vulpetti A (2014) Fluorine as a hydrogen-bond acceptor: experimental evidence and computational calculations. Chemistry 20:11058–11068

Danielson MA, Falke JJ (1996) Use of 19F NMR to probe protein structure and conformational changes. Annu Rev Biophys Biomol Struct 25:163–195

Durr UH, Grage SL, Witter R, Ulrich AS (2008) Solid state 19F NMR parameters of fluorine-labeled amino acids. Part I: aromatic substituents. J Magn Reson 191:7–15

Edsall JT, Martin RB, Hollingworth BR (1958) Ionization of individual groups in dibasic acids, with application to the amino and hydroxyl groups of tyrosine. Proc Natl Acad Sci 44:505

Eichele K (2015) HBA 1.7.5. Universität Tübingen, Tübingen

Faleev NG, Axenova OV, Demidkina TV, Phillips RS (2003) The role of acidic dissociation of substrate's phenol group in the mechanism of tyrosine phenol-lyase. Biochim Biophys Acta (BBA) 1647:260–265

Frieden C, Hoeltzli SD, Bann JG (2004) The preparation of 19F-labeled proteins for NMR studies. In: Michael L, Holt JJM, Gary KA (eds) Methods in enzymology, vol 380. Academic Press, San Diego, pp 400–415

Fung BM, Khitrin AK, Ermolaev K (2000) An improved broadband decoupling sequence for liquid crystals and solids. J Magn Reson 142:97–101

Gee CT et al (2016) Protein-observed 19F-NMR for fragment screening, affinity quantification and druggability assessment. Nat Protocols 11:1414–1427

Gee CT, Arntson KE, Koleski EJ, Staebell RL, Pomerantz WCK (2018) Dual labeling of the CBP/p300 KIX domain for 19F NMR leads to identification of a new small-molecule binding site. ChemBioChem 19:963–969

Gerig (2001) Fluorine NMR. www.biophysics.org/img/jtg2001-2.pdf

Goto NK, Zor T, Martinez-Yamout M, Dyson HJ, Wright PE (2002) Cooperativity in transcription factor binding to the coactivator CREB-binding protein (CBP). The mixed lineage leukemia protein (MLL) activation domain binds to an allosteric site on the KIX domain. J Biol Chem 277:43168–43174

Hawk LML, Gee CT, Urick AK, Hu H, Pomerantz WCK (2016) Paramagnetic relaxation enhancement for protein-observed 19F NMR as an enabling approach for efficient fragment screening. RSC Adv 6:95715–95721

Herzfeld J, Berger AE (1980) Sideband intensities in NMR spectra of samples spinning at the magic angle. J Chem Phys 73:6021–6030

Hull WE, Sykes BD (1975) Fluorotyrosine alkaline phosphatase: internal mobility of individual tyrosines and the role of chemical shift anisotropy as a 19F nuclear spin relaxation mechanism in proteins. J Mol Biol 98:121–153

Isley WC, Urick AK, Pomerantz WCK, Cramer CJ (2016) Prediction of 19F NMR chemical shifts in labeled proteins: computational protocol and case study. Mol Pharm 13:2376–2386

Kenward C, Shin K, Rainey JK (2018) Mixed fluorotryptophan substitutions at the same residue expand the versatility of 19F protein NMR spectroscopy. Chemistry 24:3391–3396

Mishra NK, Urick AK, Ember SW, Schonbrunn E, Pomerantz WC (2014) Fluorinated aromatic amino acids are sensitive 19F NMR probes for bromodomain-ligand interactions. ACS Chem Biol 9:2755–2760

Norton RS, Leung EWW, Chandrashekaran IR, MacRaild CA (2016) Applications of 19F-NMR in fragment-based drug discovery. Molecules 21:860

Pomerantz WC et al (2012) Profiling the dynamic interfaces of fluorinated transcription complexes for ligand discovery and characterization. ACS Chem Biol 7:1345–1350

Rule GS, Pratt EA, Simplaceanu V, Ho C (1987) Nuclear magnetic resonance and molecular genetic studies of the membrane-bound D-lactate dehydrogenase of *Escherichia coli*. Biochemistry 26:549–556

Schanda P, Huber M, Boisbouvier J, Meier BH, Ernst M (2011) Solid-state NMR measurements of asymmetric dipolar couplings provide insight into protein side-chain motion. Angew Chem Int Ed 50:11005–11009

Seyedsayamdost MR, Reece SY, Nocera DG, Stubbe J (2006) Mono-, di-, tri-, and tetra-substituted fluorotyrosines: new probes for enzymes that use tyrosyl radicals in catalysis. J Am Chem Soc 128:1569–1579

Song L, Teng Q, Phillips RS, Brewer JM, Summers AO (2007) 19F-NMR reveals metal and operator-induced allostery in MerR. J Mol Biol 371:79–92

Urick AK et al (2015) Dual screening of BPTF and Brd4 using protein-observed fluorine NMR uncovers new bromodomain probe molecules. ACS Chem Biol 10:2246–2256

Ye L, Larda ST, Frank Li YF, Manglik A, Prosser RS (2015) A comparison of chemical shift sensitivity of trifluoromethyl tags: optimizing resolution in 19F NMR studies of proteins. J Biomol NMR 62:97–103