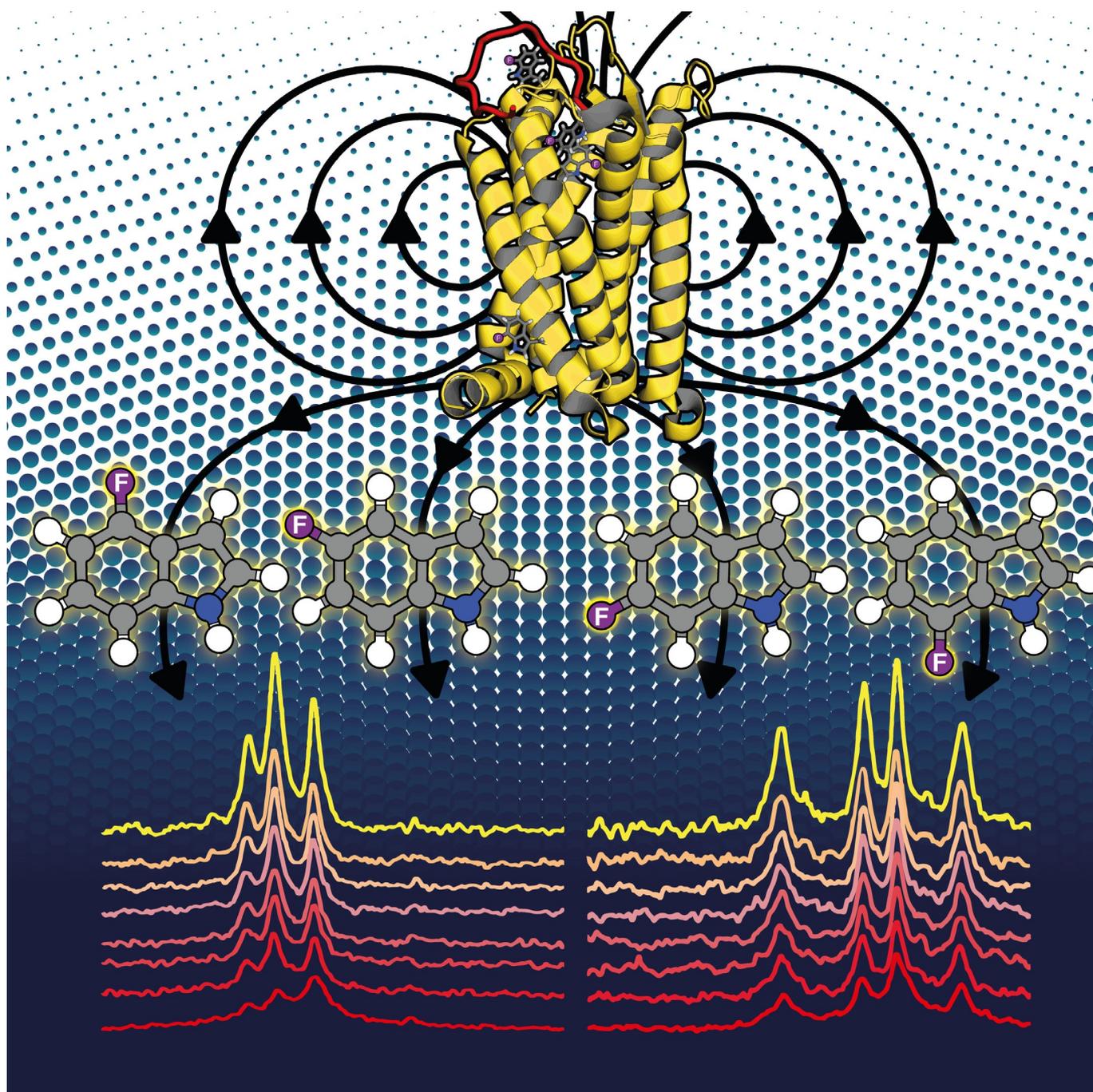


Peptides

Mixed Fluorotryptophan Substitutions at the Same Residue
Expand the Versatility of ^{19}F Protein NMR SpectroscopyCalem Kenward,^[a] Kyungsoo Shin,^[a] and Jan K. Rainey^{*[a, b]}

Abstract: The strategy of applying fluorine NMR to characterize ligand binding to a membrane protein prepared with mixtures of tryptophans substituted with F at different positions on the indole ring was tested. The ^{19}F NMR behavior of 4-, 5-, 6-, and 7-fluorotryptophan were directly compared as a function of both micellar environment and fragment size for two overlapping apelin receptor (AR/APJ) segments; one with a single transmembrane (TM) helix and two tryptophan residues, the other with three TM helices and two additional tryptophan residues. Chemical shifts, peak patterns, and nuclear spin relaxation rates were observed to vary as a function of micellar conditions and F substitution position in the indole ring, with the exposure of a given residue to micelle or solvent being the primary differentiating factor. Titration of the 3-TM AR segment biosynthetically prepared as a mixture of 5- and 7-fluorotryptophan-containing isoforms by two distinct peptide ligands (apelin-36 and apela-32) demonstrated site-specific ^{19}F peak intensity changes for one ligand but not the other. In contrast, both ligands perturbed ^1H - ^{15}N HSQC peak patterns to a similar degree. Characterization of multiple fluorotryptophan types for a given set of tryptophan residues, thus, significantly augments the potential to apply ^{19}F NMR to track otherwise obscure modulation of protein conformation and dynamics without an explicit requirement for mutagenesis or chemical modification.

^{19}F NMR spectroscopy has a long history of application for biomolecular characterization, with three key advantages: (i) high sensitivity and wide chemical shift range; (ii) an almost complete absence in biological systems, leading to a lack of background signals; and, (iii) the potential to employ 1D experiments, allowing for decreased experiment times relative to triple-resonance experiments.^[1] In triple-resonance studies, chemical shift overlap in large proteins may prove intractable, even with sparse labeling.^[2] Conversely, ^{19}F NMR can provide quantitative distance restraints in systems of this nature^[3] and track changes in conformation and dynamics that are either undetectable or confounded by signal overlap in triple-resonance experiments.^[4]

^{19}F NMR has been highly valuable in studies of α -helical membrane proteins. Notably, in the study of G protein-coupled receptors (GPCRs), ligand-induced changes in structure, dynamics, and conformational sampling have been made uniquely accessible through ^{19}F NMR.^[5] Here, we employ two overlap-

ping fragments of a class A GPCR, the apelin receptor (AR or APJ): AR55, comprising the extracellular N-terminal tail and first transmembrane (TM) segment (residues 1–55), and AR TM1–3, spanning the N-terminal tail and first three TM segments (residues 1–137) (Figure 1). The structure, backbone dynamics, and topology of AR55 have been comprehensively characterized in a variety of membrane-mimetic micellar environments,^[6] with TM1–3 studied in a more preliminary manner.^[7]

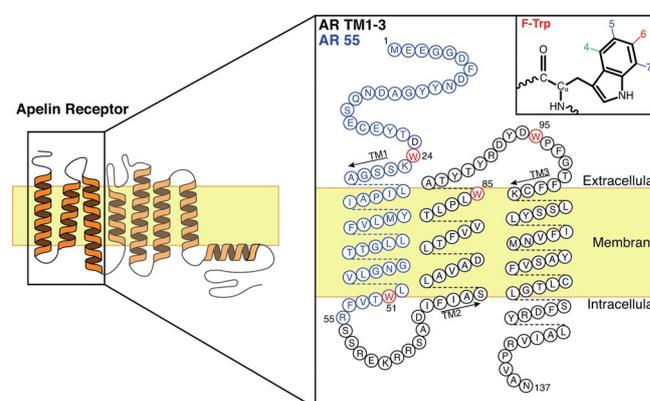


Figure 1. AR fragments employed (AR55, blue, residues 1–55 and TM1–3, residues 1–137) shown in relation to entire apelin receptor (AR) with transmembrane (TM) domains denoted. Tryptophans are red, with fluorotryptophan substitution sites (4–7) illustrated in inset.

Tryptophan is a valuable probe of protein conformation, dynamics, and intermolecular interactions. The typically low relative abundance,^[3,4c] or complete absence,^[4a] of Trp in proteins leads to an inherent sparseness of probe positions coupled with the potential for site-specific probe introduction by mutagenesis. In membrane proteins, further value comes from the long-recognized propensity of Trp to be positioned at the membrane-water interface,^[8] with the corresponding potential to employ Trp to probe changes in protein stability and/or report on intermolecular interactions occurring proximal to this interface.

Facilitating combination of the benefits of ^{19}F NMR with those of Trp residues as probes, *Escherichia coli* readily take up 4-, 5-, 6-, or 7-fluoroindole precursors for fluorotryptophan biosynthesis (4F-, 5F-, 6F-, or 7F-Trp, respectively; positions denoted in Figure 1).^[3,9] In the case of AR55 and AR TM1–3, yields of 2–25 mg of purified protein per liter of ^{15}N and ^{19}F -indole supplemented minimal medium were achieved. For either AR fragment, supplementation with 7-fluoroindole was the least perturbing to protein production while 4-fluoroindole consistently led to suboptimal yields, suggesting possible toxicity. High biosynthetic F-Trp incorporation (> 95%) was apparent in all cases, evidenced by an absence of Trp $\text{He}1\text{-Ne}1$ cross-peaks in ^1H - ^{15}N HSQC experiments resulting from replacement to a given F-Trp containing Ne at 0.4% ^{15}N natural abundance (Figure S1).

Based on ^1H - ^{15}N HSQC peak patterns, incorporation of F-Trp appears minimally perturbing to either AR fragment (Figure S1). Much more variation is apparent for AR55 between do-

[a] C. Kenward, Dr. K. Shin, Prof. J. K. Rainey
Department of Biochemistry & Molecular Biology
Dalhousie University, Halifax, Nova Scotia B3H 4R2 (Canada)
E-mail: jan.rainey@dal.ca

[b] Prof. J. K. Rainey
Department of Chemistry, Dalhousie University
Halifax, Nova Scotia B3H 4R2 (Canada)

Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under <https://doi.org/10.1002/chem.201705638>.

decylphosphocholine (DPC) and sodium dodecyl sulfate (SDS) micellar conditions, despite very similar TM domain structuring, topology, and backbone dynamics in each micelle.^[6b] Correspondingly, AR55 backbone amide ¹H and ¹⁵N chemical shift perturbations localize to residues in proximity to W24 and W51 residues (D23, K25, S26, L50 and T52, from prior assignments,^[6] Figure S1). Hence, biosynthetic F-Trp incorporation appeared minimally perturbing to overall protein structure for each AR fragment.

Each F-Trp type gave distinct ¹⁹F NMR spectroscopic behavior, both in terms of chemical shift range and peak pattern. In all cases, the observed shielding consistently increased from 6F- to 4F- to 5F- to 7F-Trp (Figure 2). This deviates from behavior reported for a single buried Trp in a globular protein, where the relative shielding of 4F- versus 6F-Trp was reversed.^[3] Overlapping 4F-, 5F- and 6F-Trp ¹⁹F chemical shift ranges have also been reported in globular proteins.^[10] Hence, F-Trp ¹⁹F NMR behavior appears quite context-dependent.

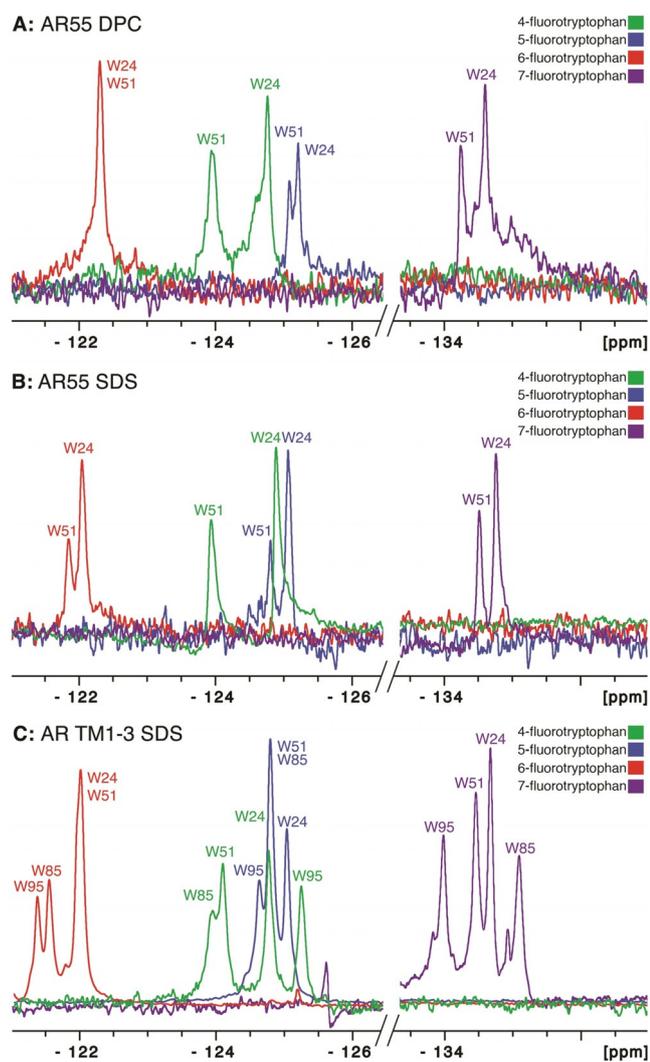


Figure 2. 1D ¹⁹F NMR spectra for given biosynthetically F-Trp-labeled AR fragment isoform in DPC or SDS micelles (as indicated).

In AR55, although ¹H-¹⁵N HSQC peak patterns differ distinctly between micelle types (Figure S1), ¹⁹F NMR spectra exhibited only minor perturbation between micelles (Figure 2). ¹⁹F NMR spectroscopy, therefore, provides a more readily interpretable probe of localized perturbations at Trp sites versus the ¹H-¹⁵N HSQC, which is convoluted by both local (e.g., individual resonance perturbation) and global (e.g., varying convolution by overlap with resonances of other residues) effects even in this relatively simple single TM-helix membrane protein fragment.

Using site-directed mutagenesis, ¹⁹F chemical shift assignment was straightforward through comparison of 1D ¹⁹F NMR spectra of W24F AR55^[11] and W95Y AR TM1-3 mutants to wild-type (Figure S2, with assignments annotated in Figure 2). To unambiguously assign chemical shifts, 1D ¹⁹F spectra were collected for isoforms of each mutant substituted with one F-Trp type. For future studies, non-overlapping F-Trp isoform mixtures (vide infra) would streamline assignment. Chemical shift assignment makes it clear that different F-Trp types have distinct peak overlap patterns and ordering of peak positions. Following from context-dependent F-Trp behavior in globular proteins,^[10] membrane-mimetic conditions (e.g., AR55 in DPC versus SDS micelles) and protein construct (e.g., 6F-Trp in AR55 versus TM1-3) both appear to modulate F-Trp ¹⁹F NMR behavior differently and unpredictably for a given F-Trp type.

In contrast to chemical shift behavior, linewidths for a given Trp site relative to the other(s) in a given protein were consistent both between F-Trp types and between conditions (Figure 2). This followed the trend, from broadest to narrowest, of W85 ≈ W95 > W51 > W24. To test for potential inherent spin relaxation differences and sensitivity to both environment and motion between the four F-Trp types, T₁ and T₂ spin relaxation time constants were measured (Figure 3). Each F-Trp type exhibits distinct relaxation behavior, with 4F- and 7F-Trp having similar and shorter T₁ and T₂ values than 5F- and 6F-Trp when compared for a given sample, with individual residues following the trend expected^[12] on the basis of linewidth.

Based upon Mn²⁺-induced PRE in AR55,^[6b] the indole Hε1-Nε1 of W51 is more protected from solvent than that of W24 in both DPC and SDS micelles (Figure S3). This is consistent with the topology of full-length AR predicted on the basis of the recent crystal structure.^[13] W85 is crystallographically predicted to be fully membrane-embedded and W95 to be membrane-proximal in extracellular loop 1 (ECL1; Figure 1). A greater degree of solvent exposure and a corresponding decrease in conformational restriction versus a micelle-embedded positioning is consistent^[14] with the sharper lines and longer T₂ values observed for ¹⁹F nuclei in W24 versus W51 in all instances (Figure 3). The ¹⁹F T₁ behavior also reflects this trend, with slightly elevated values for W24 relative to W51. In TM1-3, W85 generally follows the same trend as W51 while W95 follows that of W24. In short, regardless of F-Trp type, T₁ and T₂ are reflective of the topology of the residue in a micellar environment.

Independent of F-Trp type, the two Trp residues in AR55 generally exhibited less chemical shift dispersion in DPC than in SDS micelles (Figure 2). This is contrary to the fact that the DPC headgroup has a greater dipole moment than that of SDS

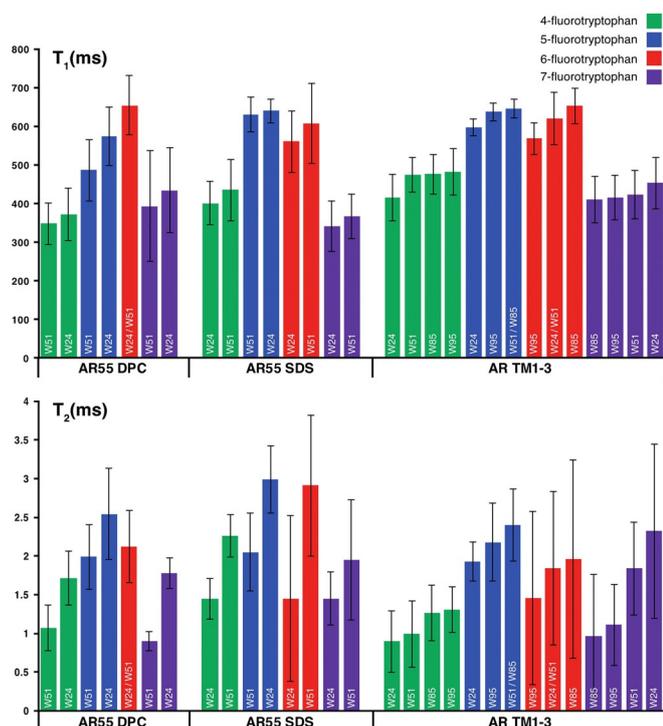


Figure 3. ^{19}F T_1 and T_2 spin relaxation constants at 11.7 T for given biosynthetically F-Trp-labeled AR fragment isoform in indicated micelle.

(≈ 15 D versus ≈ 4.7 D, respectively^[11]). In contrast to this micelle-dependent behavior, the Trp residues shared between AR55 and AR TM1-3 in SDS had highly similar ^{19}F NMR chemical shifts and spin relaxation time constants for a given F-Trp type. Given the increased protein size of TM1-3 versus AR55 (≈ 17.7 kDa versus ≈ 7.3 kDa, respectively, at natural abundance), this may imply similar overall protein-micelle complex sizes, an inference testable in future by, e.g., pulsed field gradient diffusion NMR^[6b] or other hydrodynamics measurements. The apparent lack of distinct environments for W24 and W51 in TM1-3 versus AR55 also indicates that the divide-and-conquer strategy^[15] may be applicable more generally to streamline membrane protein ^{19}F NMR chemical shift assignment.

The distinct chemical shift ranges and variation in residue-specific chemical shifts observed as a function of F-Trp type, amino acid location within both protein constructs, and micellar environment imply that incorporation of multiple F-Trp types in a given sample provide an improved ability to track changes to the sample arising from, e.g., intermolecular interactions or conformational change. Such mixtures may be readily achieved by mixing of different isoforms or through providing multiple fluoroindole precursors during biosynthesis. The value of applying F-Trp mixtures also follows from dramatically different ^{19}F chemical shift dispersion for 5F- versus 6F-Trp for the two native Trp residues of transthyretin, exhibiting distinct sensitivity to environment despite unperturbed folding,^[10b] as well as a greater 4F-Trp chemical shift dispersion versus 5F- or 6F-Trp in lysozyme.^[10a] Multi-site perturbation should, therefore, be straightforwardly and efficiently tracked by 1D ^{19}F NMR experiments following multiple F-Trp types instead of

relying upon a single F-Trp type which may or may not be an effective probe in all conditions. Based on indole precursor costs, incorporation efficiencies, and chemical shift values and ranges observed, we tested this strategy with AR TM1-3 isoforms containing mixtures of 5F- and 7F-Trp.

The AR has two known cognate peptide ligands, apelin and apela, each of which is found in multiple bioactive isoforms of varying length.^[16] The physiological effects of AR activation differ as a function both of ligand identity and of isoform size for a given ligand.^[16] Similarly-sized but physicochemically^[16] and sequentially (Figure S4) distinct apelin-36 and apela-32 ligands (at natural abundance) were independently titrated into TM1-3 (uniformly ^{15}N -enriched, biosynthetically 5F- and 7F-Trp labeled isoforms) samples in SDS micelles. From previous mutagenesis studies, TM1-3 contains residues in the N-terminal tail and ECL1 essential for apelin binding.^[17] Apelin analogue-AR interactions in the N-terminal tail and ECL1 were also recently crystallographically observed.^[13] To date, apela-AR interactions remain uncharacterized.

As would be anticipated, the requisite incorporation of only one F-Trp type at each Trp site in a given TM1-3 protein provides 1D ^{19}F NMR spectra in the absence of ligand that are simply F-Trp concentration-dependent linear combinations of the individual 5F- and 7F-Trp spectra (Figure 4). Titration by each ligand similarly resulted in minor perturbation to subsets of TM1-3 ^1H - ^{15}N HSQC cross-peaks (Figure S4). In contrast, distinct ^{19}F NMR behavior was observed in each titration (Fig-

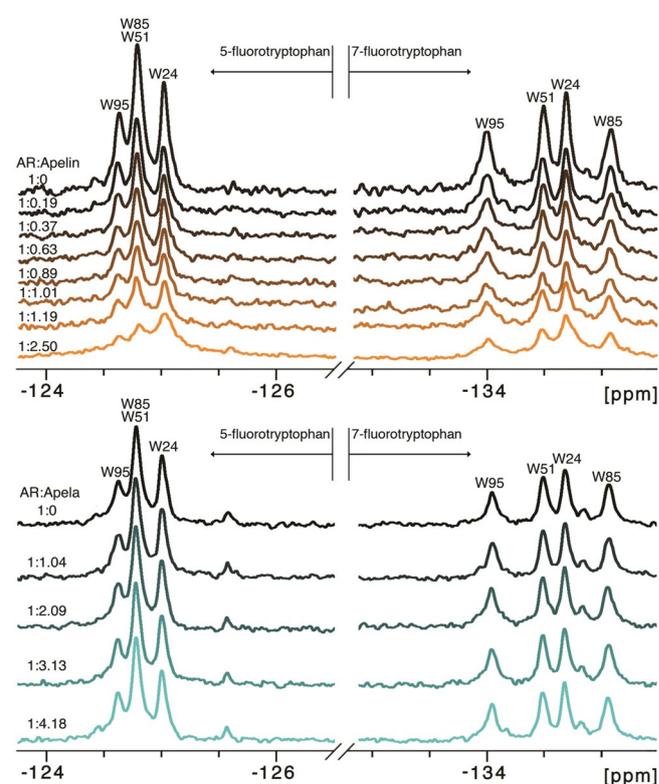


Figure 4. Titration of 5F- and 7F-Trp-labeled AR TM1-3 isoform mixtures in SDS micelles with apelin-36 or apela-32 (stoichiometry denoted) followed by ^{19}F NMR.

ures 4 and S5). Addition of apela-32 resulted in no change to 1D ^{19}F NMR spectra, even at 4.2:1 molar excess, while apelin-36 addition to 2.5:1 molar excess modulated peak intensity in a residue-dependent manner without, unexpectedly, significant ^{19}F chemical shift perturbation for either 5F- or 7F-Trp. Of the four Trp residues, W95 exhibited the most significant signal intensity decrease for both 5F- and 7F-Trp ^{19}F peaks with W51 being overlapped with W85 in the 5F-Trp region but clearly attenuated in the 7F-Trp region of the ^{19}F NMR spectrum (Figures 4 and S5).

The lack of ^{19}F chemical shift perturbation implies that neither ligand changes the local electronic environment significantly. This is not trivially interpretable since, at present, the relative contributions from the various effects underlying ^{19}F chemical shifts remain controversial.^[1] Given favorable SDS micelle binding by both apela-32^[18] and apelin-36,^[19] one potential reason for HSQC spectral perturbation would be nonspecific binding of a given ligand to the micelle, indirectly modulating AR structure and/or dynamics. This may explain the behavior seen with apela-32, but is hard to rationalize with respect to the specific modulation of F-Trp behavior by apelin-36. In the case of apelin, modulation of local dynamics within the ECL1 segment containing W95 (Figure 1) seems the most likely source of the observed decrease in W95 ^{19}F peak intensity without local perturbation at the other proximal Trp positions. It is, thus, possible that apela and apelin bind to the AR with distinct mechanisms, corresponding to signaling differences.^[16] Further characterization of the interaction of each ligand with full-length AR is required to definitively test for distinct binding mechanisms.

In summary, mixtures of membrane protein isoforms that contain differently-substituted F-Trp residues at each Trp position were readily biosynthesized. Although the focus here has been on ^{19}F NMR spectroscopy, it should also be noted that the modified photophysics of F-Trp versus Trp residues^[10b,20] imply broader applicability of these approaches for fluorescence spectroscopy and resonance energy transfer characterization. Both ^{19}F chemical shift and spin dynamics were consistent between AR constructs as a function of position. Furthermore, each F-Trp type was relatively unaffected by micellar detergent headgroup (e.g., anionic SDS versus zwitterionic DPC). 7F-Trp was the most effectively introduced F-Trp type and exhibited the most distinct chemical shift range, falling at least 9 ppm upfield of that of the next most upfield (5F-Trp) type. Disparate titration behavior by two peptidic ligands for this GPCR was also apparent through ^{19}F NMR spectroscopy despite similar, minor perturbation to ^1H - ^{15}N HSQC spectra in each titration. Thus, incorporation and simultaneous tracking of the behavior of multiple types of F-Trp provides a highly versatile means to characterize complex or challenging to study protein systems, even in the absence of full heteronuclear NMR-based chemical shift assignment.

Acknowledgements

Thanks to Dr. Mike Lumsden for NMR spectrometer support; Bruce Stewart for technical assistant; Dr. David Langelaan for PRE analysis; and, Carley Bekkers and Dr. Muzaddid Sarker for early experiments. This work was supported by grants (to J.K.R.) from the Canadian Institutes for Health Research (CIHR, MOP-111138) and Nova Scotia Health Research Foundation (MED-SSG-2015-10041) and infrastructure funding (to J.K.R.) from the Natural Sciences and Engineering Research Council of Canada (NSERC) and Dalhousie Medical Research Foundation. K.S. was supported by an NSERC Alexander Graham Bell Canadian Graduate Scholarship and a Killam Predoctoral Scholarship; J.K.R. by a CIHR New Investigator Award.

Conflict of interest

The authors declare no conflict of interest.

Keywords: biosynthesis · membrane proteins · NMR spectroscopy · peptides · protein–protein interactions

- [1] J. L. Kitevski-LeBlanc, R. S. Prosser, *Prog. Nucl. Magn. Reson. Spectrosc.* **2012**, *62*, 1.
- [2] a) J. H. Prestegard, D. A. Agard, K. W. Moremen, L. A. Lavery, L. C. Morris, K. Pederson, *J. Magn. Reson.* **2014**, *241*, 32; b) L. E. Kay, *J. Mol. Biol.* **2016**, *428*, 323.
- [3] E. Matei, A. M. Gronenborn, *Angew. Chem. Int. Ed.* **2016**, *55*, 150; *Angew. Chem.* **2016**, *128*, 158.
- [4] a) M. Sarker, K. E. Orrell, L. Xu, M. L. Tremblay, J. J. Bak, X. Q. Liu, J. K. Rainey, *Biochemistry* **2016**, *55*, 3048; b) T. H. Kim, P. Mehrabi, Z. Ren, A. Slijk, C. Ing, A. Bezginov, L. Ye, R. Pomes, R. S. Prosser, E. F. Pai, *Science* **2017**, *355*, eaag2355; c) C. Kyne, P. B. Crowley, *Biochemistry* **2017**, *56*, 5026.
- [5] a) J. Klein-Seetharaman, E. V. Getmanova, M. C. Loewen, P. J. Reeves, H. G. Khorana, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 13744; b) J. J. Liu, R. Horst, V. Katritch, R. C. Stevens, K. Wuthrich, *Science* **2012**, *335*, 1106; c) [1] R. Nygaard, Y. Zou, R. O. Dror, T. J. Mildorf, D. H. Arlow, A. Manglik, A. C. Pan, C. W. Liu, J. J. Fung, M. P. Bokoch, F. S. Thian, T. S. Kobilka, D. E. Shaw, L. Mueller, R. S. Prosser, B. K. Kobilka, *Cell* **2013**, *152*, 532; d) A. Manglik, T. H. Kim, M. Masureel, C. Altenbach, Z. Yang, D. Hilger, M. T. Lerch, T. S. Kobilka, F. S. Thian, W. L. Hubbell, R. S. Prosser, B. K. Kobilka, *Cell* **2015**, *161*, 1101; e) L. Ye, N. Van Eps, M. Zimmer, O. P. Ernst, R. S. Prosser, *Nature* **2016**, *533*, 265.
- [6] a) D. N. Langelaan, T. Reddy, A. W. Banks, G. Delleire, D. J. Dupre, J. K. Rainey, *Biochim. Biophys. Acta Biomembr.* **2013**, *1828*, 1471; b) D. N. Langelaan, A. Pandey, M. Sarker, J. K. Rainey, *J. Phys. Chem. Lett.* **2017**, *8*, 2381.
- [7] A. Pandey, M. Sarker, X. Q. Liu, J. K. Rainey, *Biochem. Cell Biol.* **2014**, *92*, 269.
- [8] J. A. Killian, G. von Heijne, *Trends Biochem. Sci.* **2000**, *25*, 429.
- [9] P. B. Crowley, C. Kyne, W. B. Monteith, *Chem. Commun.* **2012**, *48*, 10681.
- [10] a) C. Lian, H. Le, B. Montez, J. Patterson, S. Harrell, D. Laws, I. Matsu-mura, J. Pearson, E. Oldfield, *Biochemistry* **1994**, *33*, 5238; b) X. Sun, H. J. Dyson, P. E. Wright, *Biochemistry* **2017**, *56*, 5570.
- [11] R. E. Patterson, N. Weatherbee-Martin, J. K. Rainey, *J. Phys. Chem. B* **2017**, *121*, 4768.
- [12] A. Abragam, *The Principles of Nuclear Magnetism*, Clarendon Press, Oxford, UK, **1961**.
- [13] Y. Ma, Y. Yue, Y. Ma, Q. Zhang, Q. Zhou, Y. Song, Y. Shen, X. Li, X. Ma, C. Li, M. A. Hanson, G. W. Han, E. A. Sickmier, G. Swaminath, S. Zhao, R. C. Stevens, L. A. Hu, W. Zhong, M. Zhang, F. Xu, *Structure* **2017**, *25*, 858.

- [14] a) W. E. Hull, B. D. Sykes, *Biochemistry* **1974**, *13*, 3431; b) W. E. Hull, B. D. Sykes, *J. Mol. Biol.* **1975**, *98*, 121.
- [15] N. Bordag, S. Keller, *Chem. Phys. Lipids* **2010**, *163*, 1.
- [16] K. Shin, C. Kenward, J. K. Rainey, *Compr. Physiol.* **2018**, *8*, 407.
- [17] a) N. Zhou, X. Zhang, X. Fan, E. Argyris, J. Fang, E. Acheampong, G. C. DuBois, R. J. Pomerantz, *Virology* **2003**, *317*, 84; b) R. Gerbier, V. Leroux, P. Couvineau, R. Alvear-Perez, B. Maigret, C. Llorens-Cortes, X. Iturrioz, *FASEB J.* **2015**, *29*, 314.
- [18] S. K. Huang, K. Shin, M. Sarker, J. K. Rainey, *Biochim. Biophys. Acta Biomembr.* **2017**, *1859*, 767.
- [19] K. Shin, M. Sarker, S. K. Huang, J. K. Rainey, *Sci. Rep.* **2017**, *7*, 15433.
- [20] S. S. Sarkar, J. B. Udgaonkar, G. Krishnamoorthy, *J. Phys. Chem. B* **2011**, *115*, 7479.

Manuscript received: November 28, 2017

Accepted manuscript online: January 17, 2018

Version of record online: February 13, 2018