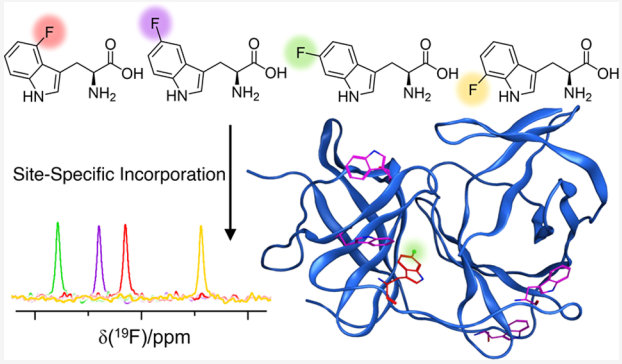
JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
pubs.acs.org/JACS
Article

Genetic Encoding of Fluoro-L-tryptophans for Site-Specific Detection of Conformational Heterogeneity in Proteins by NMR Spectroscopy

Haocheng Qianzhu,[§] Elwy H. Abdelkader,[§] Gottfried Otting,* and Thomas Huber*

Cite This: *J. Am. Chem. Soc.* 2024, 146, 13641−13650
Read Online

ABSTRACT: The substitution of a single hydrogen atom in a protein by fluorine yields a site-specific probe for sensitive detection by ^{19}F nuclear magnetic resonance (NMR) spectroscopy, where the absence of background signal from the protein facilitates the detection of minor conformational species. We developed genetic encoding systems for the site-selective incorporation of 4-fluorotryptophan, 5-fluorotryptophan, 6-fluorotryptophan, and 7-fluorotryptophan in response to an amber stop codon and used them to investigate conformational heterogeneity in a designed amino acid binding protein and in flaviviral NS2B-NS3 proteases. These proteases have been shown to present variable conformations in X-ray crystal structures, including flips of the indole side chains of tryptophan residues. The ^{19}F NMR spectra of different fluorotryptophan isomers installed at the conserved site of Trp83 indicate that the indole ring flip is common in flaviviral NS2B-NS3 proteases in the apo state and suppressed by an active-site inhibitor.



INTRODUCTION

Fluorine atoms in proteins present site-specific probes that are readily detected by ^{19}F NMR spectroscopy. Any structural perturbation caused by the fluorine is minimized when only a single hydrogen atom in the protein is replaced by fluorine. In particular, fluorotryptophans (F-Trp) with a single fluorine atom in the six-membered ring of the side chain indole are very similar in size and chemical properties to their parent canonical amino acid, tryptophan (Trp). The tryptophanyl-tRNA synthetases of prokaryotic and eukaryotic organisms are known to accept F-Trp instead of Trp, and proteins with uniform substitution of Trp for 4-fluorotryptophan (4F-Trp), 5-fluorotryptophan (5F-Trp), 6-fluorotryptophan (6F-Trp), and 7-fluorotryptophan (7F-Trp) have long been produced in *Escherichia coli* and yeast, typically with the help of tryptophan auxotrophs or specifically evolved strains to aid the incorporation of F-Trp instead of Trp.[1−4] To obtain homogeneous protein preparations, however, these methods typically aim to replace every Trp residue by F-Trp, leading to a signal for each F-Trp residue in the ^{19}F NMR spectrum, creating potential signal overlap and the need to assign the signals to specific residues before spectral changes can be interpreted in terms of site-specific responses. Enzymes made with uniform substitution of Trp for F-Trp usually preserve their structure and function,[5−11] but the presence of multiple F-Trp residues often attenuates their enzymatic rates, although enhancements have also been reported.[5]

To reap the full benefits of F-Trp incorporation requires the capability of site-selectively installing single F-Trp residues in proteins in the presence of multiple Trp residues. We recently reported an aminoacyl-tRNA synthetase (RS) mutant that enables genetic encoding of 7F-Trp as the target noncanonical amino acid (ncAA) in response to an amber codon. The enzyme was selected from a library of pyrrolysyl-RS mutants derived from the archaeal organism *ISO-G1* (G1PylRS). It enables the replacement of single Trp residues by 7F-Trp in high yield and purity in the presence of multiple tryptophan residues at other sites in the protein, thus producing the desired outcome of a protein where a single hydrogen atom is replaced by a fluorine atom.[12] The selectivity of the enzyme was unexpected, as the structural difference between F-Trp and Trp is small (the C−F bond being only 0.3 Å longer than the C−H bond and the van der Waals radius of fluorine larger by 0.15 Å). Unsurprisingly, our initial attempts to select an RS enzyme for the incorporation of other F-Trp isomers were unsuccessful.

Received: March 16, 2024
Revised: April 21, 2024
Accepted: April 23, 2024
Published: April 30, 2024

https://doi.org/10.1021/jacs.4c03743
J. Am. Chem. Soc. 2024, 146, 13641−13650

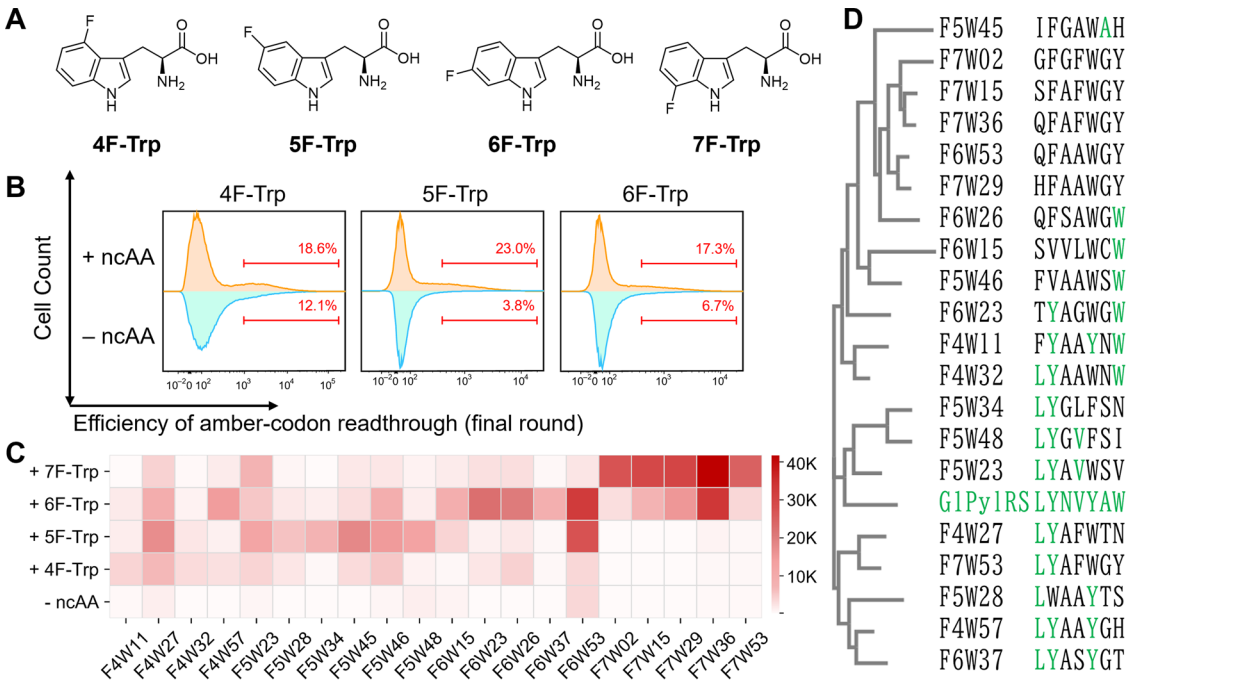


Figure 1. Development of functional and specific *G*1PylRS for F-Trps. (A) Chemical structure of 4F-Trp, 5F-Trp, 6F-Trp, and 7F-Trp. (B) Histogram of the third FACS selection rounds to identify *G*1PylRS enzymes active for 4F-Trp, 5F-Trp, and 6F-Trp, respectively. The horizontal axis plots the level of red fluorescence observed by expression of the mCherry red fluorescent protein (RFP) gene preceded by an amber stop. The vertical axis represents the cell count. The difference in RFP fluorescence intensity of cells grown with and without F-Trp serves as an indicator of the presence of RS enzymes specific to F-Trp in the gene pool. (C) Heatmap of activities from 20 selected F-Trp tRNA synthetases when 4F-Trp, 5F-Trp, 6F-Trp, or 7F-Trp is supplied either at 1 mM concentration or without addition of any ncAA, showing relative activity and cross-specificity. (D) Cladogram of mutation site residues in the selected F-Trp tRNA synthetases. Residues identical to the wild-type *G1*PylRS are highlighted in green.

The simplicity of the one-dimensional (1D) ^{19}F NMR spectrum afforded by the incorporation of a single fluorine atom into a protein presents unprecedented opportunities to detect minor conformational species, which would easily be overlooked in more complex NMR spectra. As a point in case, the incorporation of 7F-Trp to replace Trp69 or Trp83 in the Zika virus NS2B-NS3 protease led to the unexpected observation of minor peaks at different intensity (ca. 25 and 50% of the main peak, respectively) that exchanged with the main peaks, limiting the lifetimes of the main peaks to about 0.1 to 1 s and coalescing with the main peak upon addition of a covalently binding active-site inhibitor. Furthermore, 7F-Trp replacing Trp50 showed no signal in the apo protein, although a signal appeared in the complex with the inhibitor.[12] The different dynamics reported by different Trp residues indicate that they sense local rather than global conformational changes.

The analytical benefit of single-atom substitutions and the question of how far fluorine could have altered the protein behavior encouraged us to seek new aminoacyl-tRNA synthetases for the genetic encoding of other F-Trp isomers. Here we report how we were successful with the identification of RS enzymes that are selective for 4F-Trp, 5F-Trp, and 6F-Trp. In addition, we evaluated the polyspecificity profile of some of these RS enzymes for different fluorinated F-Trp amino acids with the aim to use the same cell line for producing proteins with different F-Trp isomers. We present two examples of the utility of these amino acids. In the first example, we substituted the tryptophan residue in the binding site of the ancestral amino acid binding protein AncCDT-1[13] by the different F-Trp isomers to detect and discriminate between the binding of ornithine, histidine, lysine, and arginine by simple 1D ^{19}F NMR spectra. We also extended the ^{19}F NMR analysis of conformational heterogeneity of the conserved residue Trp83 in the Zika virus (ZIKV) NS2B-NS3 protease to the same Trp residue in the related flaviviral proteases from dengue virus serotype 2 (DENV2), West Nile virus (WNV), Murray valley encephalitis virus (MVEV), and Japanese encephalitis virus (JEV). The combined data suggest that the indole side chain at this site is prone to a ring flip in all of these flaviviral apo proteins.

■ RESULTS

Library Screening for 4F-Trp, 5F-Trp, and 6F-Trp tRNA Synthetases. We hypothesized that our previous failure to identify tRNA synthetases specific for 4F-Trp (4F-TrpRS) arose from (i) suppression of protein expression caused by the toxicity of 4F-Trp at 2.5 mM concentration and (ii) insufficient sampling of the RS library in the first selection round, where only 0.3% of the total population had been collected (estimated to contain 2.7×10^7 variants). We therefore carried out new sets of fluorescence-activated cell sorting (FACS) selection experiments for 4F-TrpRS, 5F-TrpRS, and 6F-TrpRS, reducing the F-Trp amino acid concentration in the cell cultures to 1 mM and collecting five times more cells in the first selection round. Other parameters were kept unchanged, such as the use of *E. coli* DH10B cells co-transformed with the tRNA synthetase plasmid, pBK-*G*1RS, which contains the *G*1PylRS library, and the selection plasmid pBAD-H6RFP, which harbors the gene of mCherry red fluorescent protein (RFP) preceded by an amber stop codon and an N-terminal His$_6$ tag. The tRNA

https://doi.org/10.1021/jacs.4c03743
J. Am. Chem. Soc. 2024, 146, 13641−13650

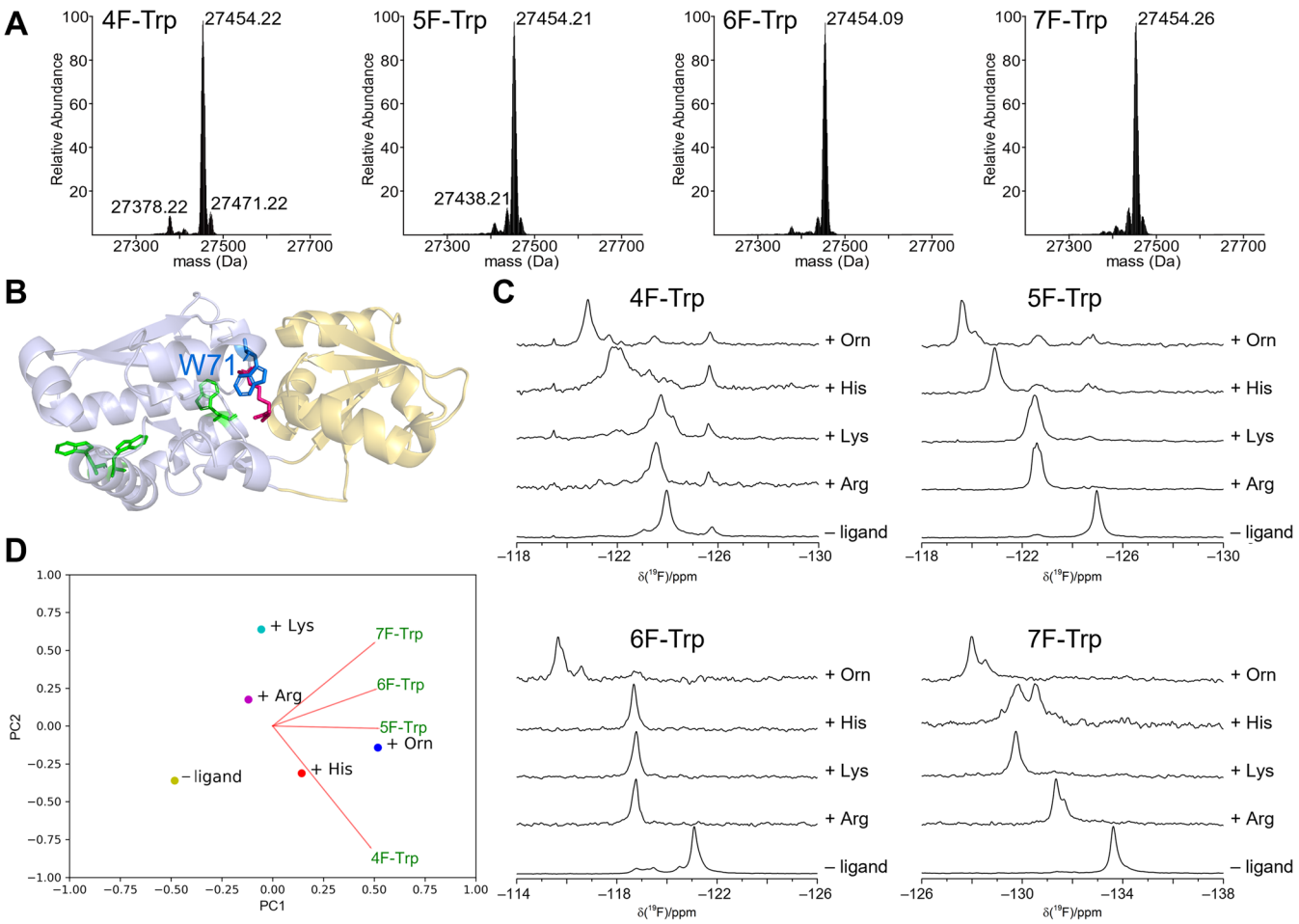


Figure 2. AncCDT-1 with site-specific replacement of Trp71 with four different F-Trp isomers for ligand binding detection and discrimination. (A) Intact protein mass spectrum of AncCDT-1 produced with 4F-Trp, 5F-Trp, 6F-Trp, or 7F-Trp. The expected mass for all samples is 27454.5 Da if they contain a single Trp to F-Trp substitution. Minor species appearing in some of the spectra with −18 or +18 Da mass correspond to low level of misincorporation of tryptophan at the amber position or replacement of Trp by F-Trp at native tryptophan sites. (B) Crystal structure of the closed conformation of AncCDT-1 (PBD: 5T0W)[13] showing the substrate binding site and location of Trp71 (blue), with other tryptophan sites colored in green and the amino acid ligand in red. The two domains are colored gray and yellow. (C) Spectral regions of 1D ^{19}F NMR spectra showing the F-Trp71 signals of AncCDT-1 in complex with ligands recorded in 50 mM HEPES, pH 7.5, and 150 mM NaCl. (D) Principal component analysis of chemical shift changes observed in the spectra shown in (C). Red lines show the principal axes in feature space, indicating high complementarity in the data obtained with the four F-Trp isomers.

synthetase library consists of *G*1PylRS variants containing the following mutations: site saturation of Leu124, Tyr125, Ala221, and Trp237; Asn165 mutated to a set of eight amino acids (Gly, Ala, Val, Ser, Asn, Thr, Ile, Asp), Val167 to a set of seven amino acids (Gly, Ala, Val, Ser, Cys, Leu, Phe), and Tyr204 to aromatic amino acids (Phe, Tyr, Trp).[14] Alternating positive and negative selection successfully enriched cells featuring the desired RS functionality for all three F-Trp variants (Figures 1B and S1). After three selection rounds, multiple colonies of cells producing high RFP fluorescence were isolated and sequenced. Remarkably, most of the candidates isolated showed negligible activity with any of the canonical amino acids compared with their activity toward the F-Trps they had been selected for (Figures 1C and S2). Ending the selection process at early rounds, the remaining library pool showed notable diversity. Among the candidates showing the highest enzyme activity, 15 new *G*1PylRS mutants were identified, including four 4F-TrpRS, six 5F-TrpRS, and five 6F-TrpRS enzymes (Figure 1D and Table S1).

Mutations in the Amino Acid Binding Site and Cross-Specificity Analysis of F-Trp tRNA Synthetases. To rationalize the remarkable specificity of the selected F-TrpRS enzymes for the single fluorine atom, we compared mutations in the active sites of 20 functional candidates across all F-TrpRS enzymes (including five identified from 7F-Trp selection) and visualized them in a sequence cladogram (Figure 1D). Surprisingly, the F-TrpRS sequences do not cluster according to the substrates they were selected for. Furthermore, the data suggest cross-specificity between more closely related F-Trp isomers. For example, F6W53 (mutation set QFAAWGY in positions 124, 125, 165, 167, 204, 221, 237) has high activity with 6F-Trp and differs by only a single residue from F7W29 (HFAAWGY) and F7W36 (QFAFWGY), suggesting that these specifically selected F-TrpRSs are potentially active with other F-Trps. We therefore assessed the substrate cross-specificity of 20 functional tRNA synthetases with 4F-Trp, 5F-Trp, 6F-Trp, and 7F-Trp (Figures 1C and S2). Besides the consistent incorporation of the substrates they were selected for, all F-TrpRS enzymes indeed display some degree of cross-specificity with other F-Trp

https://doi.org/10.1021/jacs.4c03743
J. Am. Chem. Soc. 2024, 146, 13641−13650

analogues, particularly for F-Trp isomers containing fluorine in a neighboring position. For example, F7W29 and F7W36 also accept 6F-Trp as substrate. F6W53 is more reactive with 5F-Trp than 7F-Trp. Interestingly, just two mutations, A167S and H237T, change F4W57 (LYAAYGH) to F6W37 (LYASYGT) but lead to the complete loss of activity with 4F-Trp. To produce a protein overexpression system with genetic encoding of the F-Trp isomers, we cloned the genes of F4W27 and F7W36 into a high-copy number pRSF plasmid for high yield protein production.[15] The resulting pRSF-G1F4W27 plasmid was subsequently used for the site-specific incorporation of 4F-Trp and 5F-Trp, and pRSF-G1F7W36 was used to install 6F-Trp and 7F-Trp.

Heterogeneous Ligand Binding Detected by ^{19}F NMR. The ancestral protein of the cyclohexadienyl dehydratase (AncCDT-1) is a monomeric two-domain protein, which undergoes a large structural change from an open to a closed conformation upon amino acid binding between the domains.[13] AncCDT-1 contains four tryptophan residues, with Trp71 near the amino acid binding site. To produce AncCDT-1 with Trp71 replaced by one of the four F-Trp isomers, we co-transformed the expression plasmid pCDF-AncCDT-1-W71TAG and the tRNA synthetase plasmid pRSF-G1F4W27 or pRSF-G1F7W36 into the *E. coli* strain B-95.ΔAΔ*fabR*.[16] Following our previously reported expression protocol,[12] we supplied the growth medium with fluoroindoles instead of the F-Trp amino acids, relying on the natural TrpB activity to produce the respective F-Trp amino acids. 4-, 5-, 6-, or 7-fluoroindole was supplied at a final concentration of 1 mM when the cell cultures had reached the optical density at 600 nm (OD_{600}) of 0.5. After 30 min incubation with fluoroindoles, overexpression of AncCDT-1 was induced at room temperature. All F-Trp mutants of AncCDT-1 were obtained in high yield (7.2 mg purified protein per liter of cell culture with 4F-Trp; 23 mg/L with 5F-Trp; 7.4 mg/L with 6F-Trp; 9 mg/L with 7F-Trp), and the proteins were homogeneous as indicated by intact protein mass spectrometry (Figure 2A). Notably, substitution of native tryptophan sites with F-Trp was negligible in all four samples. For the 4F-Trp mutant, a minor impurity arose from amber stop codon suppression attributed to glutamyl-tRNA, which recognizes a codon similar to the amber stop codon, and low level of misincorporation of Gln instead of F-Trp, resulting in a mass difference of −76 Da.

The proteins were analyzed using a 600 MHz NMR spectrometer equipped with a ^{19}F cryoprobe. 1D ^{19}F NMR spectra were recorded for each sample in its ligand-free state and after the addition of either ornithine, histidine, lysine, or arginine (Figure 2C). The NMR spectra of the apo protein show one major signal together with some minor peaks, which were insensitive to the presence of ligands and may indicate some level of misincorporation of F-Trp residues at canonical Trp sites. The misincorporation can be suppressed by supplying additional tryptophan in the culture media.[12] Importantly, the major peaks responded to the addition of the different amino acids with large changes in chemical shifts but, unexpectedly, some of the amino acids added split the main signal, indicating that they caused conformational heterogeneity in the binding site of this engineered non-natural protein.

The spectral changes observed in the NMR spectra of the different F-Trp isomers present fingerprints of the different amino acid ligands that can be used for their identification. This is quantified by a principal component analysis (PCA) of all chemical shift changes observed. The principal axes in feature space highlight the degree to which the different F-Trp probes provide complementary information (Figure 2D).

Unraveling Tryptophan Conformation Dynamics in Flaviviral NS2B-NS3 Proteases. In the past two decades, structural biology of the flaviviral NS2B-NS3 proteases has yielded much insight into their function and druggability, with numerous crystal structures resolved for proteases from different related flaviviruses, including the ZIKV, WNV, MVEV, JEV, yellow fever virus, and the four serotypes of DENV (Table S2). In summary, all proteases share similar folds, with the largest differences observed in the C-terminal segment of NS2B (NS2Bc) forming either a closed or an open conformation. In the closed conformation, NS2Bc forms a β-hairpin and associates with the NS3 protease (NS3p) to form part of the substrate binding site. The association is loose.[18−21] In the open conformation, NS2Bc is remote from the active site, or its structure is sufficiently disordered that no electron density can be discerned in the crystal structures. The closed conformation is required for full enzymatic activity of the proteases. The population of the closed conformation is promoted by ligand molecules binding to the active site. Initial crystallization experiments succeeded with constructs containing an artificial G_4SG_4 peptide linker between NS2Bc and NS3p, which ensured the correct 1:1 stoichiometry between NS2B and NS3p. Cumulative NMR and crystallography data suggest, however, that the G_4SG_4 peptide linker promotes the population of the open state.[17−19,21−25]

Close inspection of the crystal structures further reveals structural differences involving the side chains of the tryptophan residues, which display alternative conformations that are independent of the large conformational change of NS2Bc between open and closed conformations (Figures 3 and S3). These structural variations are remarkable as six tryptophan sites (Trp61* in NS2Bc and Trp5, Trp50, Trp69, Trp83, and Trp89 in NS3p) are strictly conserved in the aforementioned flaviviral proteases (Figure S4), and high sequence conservation is a strong indicator for structural and

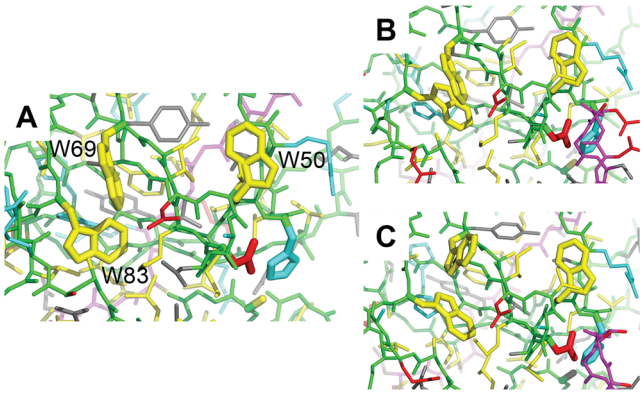


Figure 3. Crystal structures of different DENV4 NS2B-NS3 constructs.[17] The backbone of NS2B and NS3p is shown in magenta and green, respectively. The tryptophan side chains and the residues of the catalytic triad are shown in bold stick representations. (A) 5YVJ, showing the canonical orientation of the tryptophan side chains. (B) 5YVU, showing the side chain of Trp83 in a flipped orientation. The structure is of the protein shown in (a) but bound to aprotinin. (C) 5YW1, showing the side chains of Trp69 and Trp83 in flipped orientations. The structure is of a construct without covalent linkage between NS2B and NS3, in complex with aprotinin.

https://doi.org/10.1021/jacs.4c03743
J. Am. Chem. Soc. 2024, 146, 13641−13650

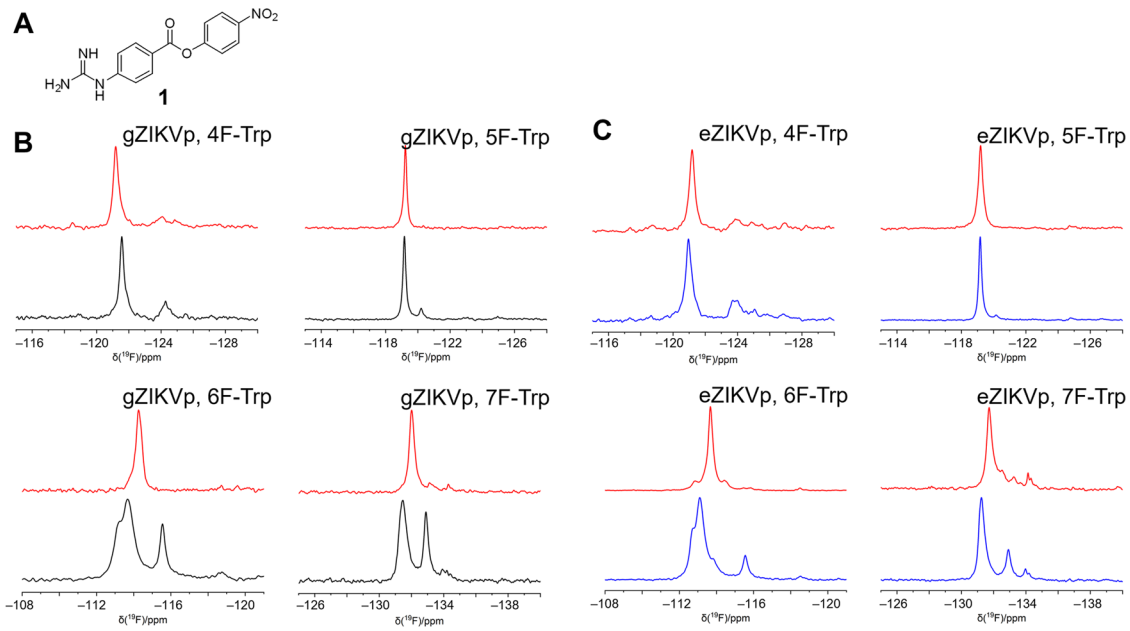


Figure 4. Conformational heterogeneity detected by F-Trp incorporation in linked and unlinked ZIKV NS2B-NS3 protease constructs. (A) Chemical structure of the broadband protease inhibitor 4-nitrophenyl-4-guanidinobenzoate (**1**). (B) 1D ^{19}F NMR spectra showing the F-Trp signals with residue Trp83 replaced in gZIKVp, where NS2Bc and NS3p are linked by a G_4SG_4 linker, before and after addition of the inhibitor **1** (black and red spectra, respectively). The ^{19}F NMR spectra were recorded at 25 °C on a 600 MHz NMR spectrometer. The buffer used contained 20 mM MES, pH 6.5, and 150 mM NaCl. (C) Same as (B), but for unlinked eZIKVp samples, where the linker contains the protease recognition site for self-cleavage.

functional importance. In the case of phenylalanine and tyrosine residues, which also feature high sequence conservation, 180° ring flips on the microsecond to millisecond time scale are commonly observed, but these flips result in no net structural change.[26] In contrast, a flip of the indole ring of a buried tryptophan residue presents a significant structural change of the protein structure and is probably unfavorable in energy. It is therefore remarkable that the body of crystal structures accumulated for the flaviviral proteases indicates that tryptophan ring flips can occur in these proteins for indole side chains with limited or, as in the case of Trp83, no solvent exposure (Table S2).

The crystal structures indicate that the tryptophan ring flips are quite independent of each other. For example, the structure 5YVU shows a conformational change only for the side chain of Trp83, whereas the side chains of Trp69 and Trp83 are both changed in the crystal structure 5YW1 (Figure 3). It is remarkable that these unusual tryptophan side chain conformations can become the predominant species in the crystals of NS2B-NS3 of dengue serotype 4. Other flaviviral proteases reveal alternative tryptophan conformations only for the side chain of Trp50, which is near the active-site residue His51 and more solvent-exposed than the other tryptophan side chains. The flip of the Trp50 side chain appears to be facile, as its noncanonical conformation has occasionally been captured in crystals of almost all flaviviral proteases (Table S2).

Probing Conformational Heterogeneity with F-Trp. Using ^{19}F NMR, we previously detected additional minor signals for 7F-Trp installed at the sites of Trp69 and Trp83 of the ZIKV NS2B-NS3 protease (ZIKVp), which indicate alternative conformations in slow exchange with the main conformation.[12] Different amounts of minor species (not exceeding 50%) were detected for the side chains of Trp69 and Trp83. As expected for an increased flip rate with lifetimes in the microsecond-to-millisecond range, the ^{19}F NMR signal of 7F-Trp in position 50 of ZIKVp was broadened beyond detection, but it became observable after adding the active-site inhibitor 4-nitrophenyl-4-guanidinobenzoate (Figure 4A), which is known to rigidify the protease.[12]

Having gained the capability of installing different F-Trp isomers, we probed the conformational heterogeneity of different flaviviral proteases by installing different F-Trp isomers at the site of Trp83. In contrast to the AncCDT-1 protein, expression of the proteases delivered significantly lower yields with fluoroindoles than fluorotryptophans, and therefore most samples were prepared with F-Trp. Using the construct of ZIKVp with the G_4SG_4 peptide linker connecting NS2Bc with NS3p (gZIKVp, in the literature also referred to as gZiPro) and supplying the F-Trp isomers in 1 mM concentration, in addition to 4 mM tryptophan supplemented to suppress the substitution of native tryptophan residues with F-Trp, we obtained expression yields of at least 10% of that of the wild-type protein with minimal amber suppression by glutamyl-tRNA, negligible misincorporation of Trp, and little background due to additional F-Trp residues according to intact protein mass spectrometry (Figure S5). 1D ^{19}F NMR spectra in the presence and absence of inhibitor **1** (Figure 4A) showed variable degrees of heterogeneity in the apo protein, demonstrating sensitivity toward the position of the fluorine atom. The heterogeneities were observed both for the linked gZIKVp constructs (Figure 4B) and the unlinked eZIKVp construct (Figure 4C), where the NS2B-NS3 protease was expressed with a linker containing the protease recognition site for autodigestion. ^{19}F−^{19}F NOESY spectra confirmed that exchange occurs between major and minor signals with exchange lifetimes between 0.1 and 0.5 s for the major species (Figure S7). The relative populations of heterogeneities varied between gZIKVp and eZIKVp. Figure 4 shows that the

https://doi.org/10.1021/jacs.4c03743
J. Am. Chem. Soc. 2024, 146, 13641−13650

populations of the minor species are also sensitive to the position of the fluorine atom in the indole ring. In all cases, the heterogeneities were largely suppressed by the addition of the inhibitor **1** (Figure 4).

Trp83 is relatively far from the substrate binding site, making it a good choice for probing the structural heterogeneity of the flaviviral proteases free from direct interactions with active-site ligands. For comparison, we also produced F-Trp variants in position 83 of the NS2B-NS3 proteases from WNV, DENV2, JEV, and MVNV with the G_4SG_4 linker (referred to as gWNVp, gDENV2p, gJEVp, and gMVEVp, respectively). The selectivity of incorporation was confirmed by mass spectrometry (Figure S6). Like with ZIKVp, the 1D ^{19}F NMR spectra all showed substantial heterogeneity, which in the cases of gWNVp and gDENV2p was readily suppressed by the addition of inhibitor **1** (Figure 5A,B). In the cases of gJEVp and gMVEVp, sample

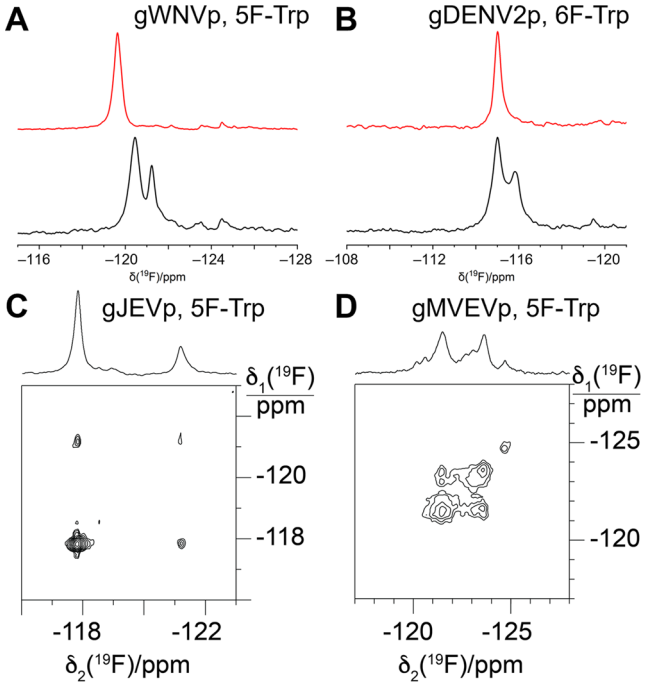


Figure 5. Inhibitor **1** reduces the conformational heterogeneity in different flaviviral proteases. ^{19}F NMR spectra were recorded without (black) and with (red) inhibitor **1**. (A) gWNVp with Trp83 replaced by 5F-Trp. To improve the solubility, the measurements were performed in 50 mM Tris-HCl, pH 7.5, and 100 mM NaCl. (B) gDENV2p with Trp83 replaced by 6F-Trp. Spectra recorded in 20 mM MES, pH 6.5, and 150 mM NaCl. Conformation exchange also appears in the NS2B-NS3 proteases of Japanese and Murray Valley encephalitis viruses with the G_4SG_4-linked construct (C) gJEVp and (D) gMVEVp. ^{19}F−^{19}F NOESY experiments were recorded with 200 ms mixing time in 50 mM Tris-HCl, pH 7.5, 100 mM NaCl. Trp83 was replaced by 5F-Trp. 1D ^{19}F NMR spectra are shown at the top.

precipitation prevented experiments of the complexes with inhibitor **1**, but conformational exchange in the apo proteins was confirmed by exchange cross-peaks in 2D ^{19}F−^{19}F NOESY spectra (Figure 5C−D). Clearly, conformational heterogeneity at the site of Trp83 is common in flaviviral proteases.

Comparing Conformational Dynamics between Constructs. The wild-type flaviviral NS2B-NS3 proteases contain no covalent linkage between NS2Bc and NS3p, and it is known that the commonly used G_4SG_4 linker facilitates crystallization but promotes the open conformation of the proteases. To break the linkage, constructs can be made with a protease recognition site in the linker for enzymatic self-cleavage. In the case of the ZIKV protease construct with enzymatic self-cleavage site (eZIKVp), it has been shown for a closely related construct that, following self-cleavage, an N-terminal tetrapeptide segment of NS3p can bind to the active site and prevent the binding of the inhibitor aprotinin.[27,28] While the heterogeneities in eZIKVp appear smaller than in gZIKVp, the spectral changes observed in Figure 4C indicate that inhibitor **1** can still bind to eZIKVp.

The ^{19}F NMR spectra of the corresponding unlinked constructs of WNVp and MVEVp (eWNVp, and eMVEVp, respectively) with F-Trp isomers in position 83 revealed heterogeneities closely similar to the linked constructs (Figure S8). For all unlinked and F-Trp-labeled constructs, intact protein mass spectrometry detected the individual NS2B and NS3p fragments without any full-length protease, confirming the enzymatic activities of these constructs (Figures S5 and S6). This suggests that the linked and unlinked constructs of these proteases are similarly suitable for the enzymatic assays underpinning drug discovery.

pH Dependence of Structural Heterogeneity Sensed by Trp83. The closed conformation of the NS2B-NS3 proteases is known to be promoted by electrostatic attraction of negatively charged residues in the tip of the β-hairpin of NS2Bc to positive charges in substrates and inhibitors binding at the active site.[19,23] The positive charge contributed by inhibitor **1** following reaction with the active-site serine residue presents a point in case. It may thus be expected that the closed conformation is favored at lower pH, where the active-site histidine side chain carries a positive charge.[19] In our constructs of different proteases made with different F-Trp isomers in position 83, lower pH changed the relative populations of minor and major species relatively little (Figure S9), suggesting that the pH change does not rigidify the proteases to the same extent as the binding of an inhibitor.

■ DISCUSSION

Benefits of Single-Site-Selective Installation of F-Trp Residues. F-Trp residues deliver readily accessible site-specific information from simple 1D ^{19}F NMR spectra. This has been exploited with great success in a multitude of applications such as measuring diffusion coefficients of proteins in live cells,[29,30] protein folding/unfolding equilibria,[31−36] and dynamics arising from ligand binding,[37] at protein interfaces,[38] or as reflected in interconversion rates between different conformations.[39] Furthermore, the ^{19}F chemical shifts of F-Trp residues are highly sensitive to their chemical environment, which is advantageous for ligand binding studies.[40−45] ^{19}F NMR signals can still be detected by solution NMR of systems up to 500 kDa.[46] Taking advantage of a ^{19}F−^{13}C spin pair in fluorotyrosine, ^{19}F−^{13}C TROSY experiments produce remarkably narrow ^{13}C signals in high-molecular-weight systems,[47] which has also been demonstrated for ^{13}C-labeled F-Trp.[48] In an important recent development, solid-state NMR has been demonstrated to afford accurate distance measurements between the ^{19}F spins of F-Trp residues in proteins[49] in the distance range up to 20 Å.[50] In solution, measurements of distances up to 24 Å to a paramagnetic nitroxide label have been reported.[51] Recently developed EPR methods have measured accurate distances in the range of 5−20 Å with exquisite sensitivity that enables in-cell analyses of proteins at

https://doi.org/10.1021/jacs.4c03743
J. Am. Chem. Soc. 2024, 146, 13641−13650

low concentrations.[52−54] In protein engineering, F-Trp residues have been used to modify protein properties.[55] An exciting recent finding is that 7F-Trp can form photoinduced fluorescent protein cross-links with phenylalanine and tyrosine.[56]

All of these applications will benefit fundamentally from the capability to install single F-Trp residues site-specifically, as the single-atom substitution of a hydrogen for a fluorine atom minimizes the perturbation of protein structure, whereas global substitution of Trp by F-Trp can compromise protein stability even when the structural changes are small.[39,57−59] Single-site substitutions also obliterate the need for additional experiments to attribute the effects to a specific individual F-Trp residue. As a major benefit, the single-site incorporation enables the detection of weak ^{19}F NMR signals from minor conformational species, and the populations of the different species can be determined by simple integration of the ^{19}F NMR resonances. Importantly, the capability of installing different F-Trp isomers makes it possible to control for any structural or functional impact that may be caused by the fluorine atom.

The present work shows that 4F-Trp, 5F-Trp, and 6F-Trp can be site-specifically installed in proteins by genetic encoding using *ISO-G*1PylRS enzymes related to the previously published PylRS enzyme for 7F-Trp,[12] despite their close similarity to natural tryptophan. Pyrrolysyl tRNA/tRNA synthetase pairs have been shown to enable ncAA incorporation not only in bacteria but also in yeast[60] and baculovirus-infected insect[61] and mammalian cells.[62,63] Therefore, we expect that the tools developed here will also prove valuable in other protein expression systems. On a practical note, producing proteins with single F-Trp residues is more straightforward than the uniform substitution of tryptophan, as there is no need for auxotrophic bacterial strains or the inhibition of tryptophan biosynthesis by inhibitors such as 3β-indoleacrilic acid[64] or glyphosate.[33,65,66] Furthermore, shutting down endogenous Trp synthesis, either by overloading the medium with F-Trp or using inhibitors like glyphosate can enhance protein degradation and potentially modify the F-Trp amino acid.[67] As bacterial cells efficiently convert fluorinated indoles to F-Trp, cost can be saved by provision of the corresponding fluorinated indoles, which is feasible both in vivo[51,68−72] and in cell-free synthesis.[12] In our hands, the protein yields were reduced by at most 10-fold relative to those obtained for the wild-type proteins. Higher yields could potentially be achieved by controlling cellular factors that regulate F-Trp concentration in *E. coli* cells. For example, the transcriptional repressor TrpR downregulates the expression of tryptophan synthase, and Trp permease (Mtr) and multidrug efflux pumps (MdtF, MdtO, and MdtK) modulate fluoroindol/F-Trp homeostasis.[73]

Tryptophan Ring Flips in Flaviviral Proteases. Our analysis of flaviviral NS2B-NS3 proteases produced with a site-specifically installed F-Trp residue to replace Trp83 revealed multiple conformational species as a common feature of these proteases. The chemical exchange between major and minor species and the disappearance of minor species after addition of an inhibitor proves that the minor signals are not due to misincorporation at different Trp sites.

We attribute the minor species to alternative conformations of the indole ring systems of Trp side chains, which limit the lifetime of the major species to about 0.1−0.5 s, with different Trp sites displaying different rates of the conformational exchange.[12] The slow rate of interconversion puts the exchange rate on an entirely different time scale from the exchange between open and closed conformations of NS2Bc. For example, an unlinked construct of DENV2p has been shown to populate the open state to only 4% and the lifetime of the closed conformation was estimated to be about 8 ms, which is too short to resolve separate NMR signals.[25] Also in the presence of a peptide linker between NS2Bc and NS3p, the exchange dynamics between open and closed conformations are too fast to resolve separate signals for the two species.[23]

Rapid dynamics of NS2Bc have also been established for ZIKVp, where solution NMR showed that the $Gly_4SerGly_4$ linked construct (gZIKVp) features NS2Bc in a mostly open conformation,[74] whereas it appears to be in a closed conformation in the unlinked construct (bZiPro), with dynamic exchange processes rendering the NMR signals of residues 67−90 too broad to allow resonance assignment.[28] In a different unlinked construct, where the protein was expressed with an enzymatically cleaved linker (eZiPro), ^{15}N relaxation data show NS2Bc in a highly dynamic state with no evidence of exchange broadening.[28]

Several crystal structures of flaviviral NS2B-NS3 proteases demonstrate that these proteins can crystallize with tryptophans displaying noncanonical indole ring conformations as the major species (Figure 3 and Table S2). The observation of multiple side chain conformations is unexpected, as the tryptophan residues are conserved between the different proteases, and sequence conservation of buried Trp residues is a common signature of structural importance, as they act as anchor points for stable protein folds owing to their large size and the rigidity of the indole ring structure. In other proteins, F-Trp residues show no evidence of multiple conformations as observed, e.g., in the stable proteins GB1 and NT* domain.[75] The present work shows that the conformational variability of the flaviviral proteases is also greatly restricted in the presence of inhibitor **1**, whereas it is significant in the apo proteins. The conformational heterogeneity of the apo proteins may play a yet unknown functional role or merely be a signature of viral genome evolution, for which protease activity is vital, whereas long-term stability and very high activity is of lesser biological importance.

F-Trp as a Fluorescence Probe. F-Trp residues are attractive not only as NMR probes but also as fluorescent probes. The fluorescence properties of some of the F-Trp isomers compare favorably to those of tryptophan. For example, the absorption of 5F-Trp is red-shifted, which enables selective excitation of its fluorescence at wavelengths between 305 and 310 nm,[76] and its fluorescence decay kinetics in proteins are usually monoexponential,[77] which renders 5F-Trp attractive as an energy donor in FRET experiments.[77,78] In addition, the time dependence of 5F-Trp fluorescence reflects the dynamics of the dielectric relaxation in the fluorophore environment, rendering it a marker of the heterogeneity of the environment,[79] including the detection of local protein−water dynamics.[80] In contrast, 4F-Trp features an exceptionally low fluorescence quantum yield and can thus be used as a "knock-out" fluorescence analog of Trp (Figures S10 and S11).[81−83] The possibility to install single F-Trp residues opens the door to their use as site-specific fluorescent probes, for example, for measuring the solvent accessibility of the side chain by Stern−Volmer fluorescence quenching experiments with potassium iodide.[84]

https://doi.org/10.1021/jacs.4c03743
J. Am. Chem. Soc. 2024, 146, 13641−13650

F-Trp in Protein Engineering. Finally, installing F-Trp residues may prove useful for biotechnological applications. In some examples, the global substitution of Trp for F-Trp has been shown to increase the stability of specific proteins.[85] Targeted substitution of individual Trp residues will make this approach much more powerful.

CONCLUSIONS

We anticipate that the genetic encoding of different fluorotryptophan isomers will greatly advance NMR, fluorescence, and biotechnological applications, especially as the production of proteins can be achieved with inexpensive fluoroindoles. To encourage the uptake of this technology, the requisite plasmids have been deposited at Addgene (Watertown, MA) (#207620 and #207621).

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jacs.4c03743.

Experimental procedures describing (a) materials, (b) selection of functional G1PylRS enzymes recognizing 4F-Trp, 5F-Trp, and 6F-Trp, (c) substrate cross-specificity analysis of selected G1PylRS mutants, (d) in vivo protein expression and purification, (e) intact protein mass spectrometry, and (f) NMR spectroscopy; FACS experiments for selection of active and specific G1PylRS enzymes; fluorescence spectra of wild-type (Trp) and F-Trp AncCDT-1 mutants; mutations found in identified G1PylRS variants that recognize F-Trps; and DNA and corresponding amino acid sequences of the proteins (PDF)

AUTHOR INFORMATION

Corresponding Authors

Gottfried Otting − *ARC Centre of Excellence for Innovations in Peptide & Protein Science, Research School of Chemistry, Australian National University, Canberra, ACT 2601, Australia;* orcid.org/0000-0002-0563-0146; Email: gottfried.otting@anu.edu.au

Thomas Huber − *Research School of Chemistry, Australian National University, Canberra, ACT 2601, Australia;* orcid.org/0000-0002-3680-8699; Email: t.huber@anu.edu.au

Authors

Haocheng Qianzhu − *Research School of Chemistry, Australian National University, Canberra, ACT 2601, Australia;* orcid.org/0000-0002-7546-4411

Elwy H. Abdelkader − *ARC Centre of Excellence for Innovations in Peptide & Protein Science, Research School of Chemistry, Australian National University, Canberra, ACT 2601, Australia;* orcid.org/0000-0002-5388-3949

Complete contact information is available at: https://pubs.acs.org/10.1021/jacs.4c03743

Author Contributions

[§]H.Q. and E.H.A. contributed equally to this work.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors thank Dr. Harpreet Vohra and Michael Devoy at the John Curtin School of Medical Research, Australian National University, for technical support on FACS experiments. Financial support by the Australian Research Council, including the Centre of Excellence for Innovations in Peptide and Protein Science (CE200100012) and a Discovery Project (DP230100079) is gratefully acknowledged.

REFERENCES

(1) Browne, D. T.; Kenyon, G. L.; Hegeman, G. D. Incorporation of monofluorotryptophans into protein during the growth of *Escherichia coli*. *Biochem. Biophys. Res. Commun.* **1970**, *39*, 13−19.

(2) Brindle, K.; Williams, S.-P.; Boulton, M. ^{19}F NMR detection of a fluorine-labelled enzyme in vivo. *FEBS Lett.* **1989**, *255*, 121−124.

(3) Bacher, J. M.; Ellington, A. D. Selection and characterization of *Escherichia coli* variants capable of growth on an otherwise toxic tryptophan analogue. *J. Bacteriol.* **2001**, *183*, 5414−5425.

(4) Munier, R. L.; Drapier, A. M.; Thommegay, C. Total substitution of 5- and 6-fluorinated analogues of tryptophane with this amino acid in proteins of *Escherichia coli* − effect of this incorporation on biosynthesis of enzymes. *C. R. Hebd. Seances Acad. Sci., Ser. D* **1967**, 1429−1432.

(5) Pratt, E. A.; Ho, C. Incorporation of fluorotryptophans into proteins of *Escherichia coli*. *Biochemistry* **1975**, *14*, 3035−3040.

(6) Parsons, J. F.; Xiao, G.; Gilliland, G. L.; Armstrong, R. N. Enzymes harboring unnatural amino acids: mechanistic and structural analysis of the enhanced catalytic activity of a glutathione transferase containing 5-fluorotryptophan. *Biochemistry* **1998**, *37*, 6286−6294.

(7) Armstrong, R. N.; Parsons, J. F.; Xiao, G.; Gilliland, G. L. Probing the mechanism and structure of glutathione transferases with unnatural amino acids. *Clin. Chem. Enzymol. Commun.* **1999**, *8*, 203−212.

(8) Mohammadi, F.; Prentice, G. A.; Merrill, A. R. Protein−protein interaction using tryptophan analogues: novel spectroscopic probes for toxin−elongation factor-2 interactions. *Biochemistry* **2001**, *40*, 10273−10283.

(9) Merkel, L.; Schauer, M.; Antranikian, G.; Budisa, N. Parallel incorporation of different fluorinated amino acids: on the way to "Teflon" proteins. *ChemBioChem* **2010**, *11*, 1505−1507.

(10) Chadegani, F.; Lovell, S.; Mullangi, V.; Miyagi, M.; Battaile, K. P.; Bann, J. G. ^{19}F nuclear magnetic resonance and crystallographic studies of 5-fluorotryptophan-labeled anthrax protective antigen and effects of the receptor on stability. *Biochemistry* **2014**, *53*, 690−701.

(11) Liu, Y.; Miao, K.; Dunham, N. P.; Liu, H.; Fares, M.; Boal, A. K.; Li, X.; Zhang, X. The cation-π interaction enables a halo-tag fluorogenic probe for fast no-wash live cell imaging and gel-free protein quantification. *Biochemistry* **2017**, *56*, 1585−1595.

(12) Qianzhu, H.; Abdelkader, E. H.; Herath, I. D.; Otting, G.; Huber, T. Site-specific incorporation of 7-fluoro-L-tryptophan into proteins by genetic encoding to monitor ligand binding by ^{19}F NMR spectroscopy. *ACS Sens.* **2022**, *7*, 44−49.

(13) Clifton, B. E.; Kaczmarski, J. A.; Carr, P. D.; Gerth, M. L.; Tokuriki, N.; Jackson, C. J. Evolution of cyclohexadienyl dehydratase from an ancestral solute-binding protein. *Nat. Chem. Biol.* **2018**, *14*, 542−547.

(14) Abdelkader, E. H.; Qianzhu, H.; George, J.; Frkic, R. L.; Jackson, C. J.; Nitsche, C.; Otting, G.; Huber, T. Genetic encoding of cyanopyridylalanine for in-cell protein macrocyclization by the nitrile−aminothiol click reaction. *Angew. Chem., Int. Ed.* **2022**, *61*, No. e202114154.

(15) Abdelkader, E. H.; Qianzhu, H.; Tan, Y. J.; Adams, L. A.; Huber, T.; Otting, G. Genetic encoding of N^6-(((trimethylsilyl)-methoxy)carbonyl)-L-lysine for NMR studies of protein−protein and protein−ligand interactions. *J. Am. Chem. Soc.* **2021**, *143*, 1133−1143.

(16) Mukai, T.; Hoshi, H.; Ohtake, K.; Takahashi, M.; Yamaguchi, A.; Hayashi, A.; Yokoyama, S.; Sakamoto, K. Highly reproductive

Escherichia coli cells with no specific assignment to the UAG codon. *Sci. Rep.* **2015**, *5*, No. 9699.
(17) Phoo, W. W.; El Sahili, A.; Zhang, Z.; Chen, M. W.; Liew, C. W.; Lescar, J.; Vasudevan, S. G.; Luo, D. Crystal structures of full length DENV4 NS2B-NS3 reveal the dynamic interaction between NS2B and NS3. *Antiviral Res.* **2020**, *182*, No. 104900.
(18) Su, X.-C.; Ozawa, K.; Qi, R.; Vasudevan, S. G.; Lim, S. P.; Otting, G. NMR analysis of the dynamic exchange of the NS2B cofactor between open and closed conformations of the West Nile virus NS2B-NS3 protease. *PLoS Neglected Trop. Dis.* **2009**, *3*, No. e561.
(19) Zhu, L.; Yang, J.; Li, H.; Sun, H.; Liu, J.; Wang, J. Conformational change study of dengue virus NS2B-NS3 protease using ^{19}F NMR spectroscopy. *Biochem. Biophys. Res. Commun.* **2015**, *461*, 677−680.
(20) Brecher, M.; Li, Z.; Liu, B.; Zhang, J.; Koetzner, C. A.; Alifarag, A.; Jones, S. A.; Lin, Q.; Kramer, L. D.; Li, H. A conformational switch high-throughput screening assay and allosteric inhibition of the flavivirus NS2B-NS3 protease. *PLoS Pathog.* **2017**, *13*, No. e1006411.
(21) Agback, T.; Lesovoy, D.; Han, X.; Lomzov, A.; Sun, R.; Sandalova, T.; Orekhov, V. Y.; Achour, A.; Agback, P. Combined NMR and molecular dynamics conformational filter identifies unambiguously dynamic ensembles of Dengue protease NS2B/NS3pro. *Commun. Biol.* **2023**, *6*, No. 1193.
(22) Kim, Y. M.; Gayen, S.; Kang, C.; Joy, J.; Huang, Q.; Chen, A. S.; Wee, J. L. K.; Ang, M. J. Y.; Lim, H. A.; Hung, A. W.; Li, R.; Noble, C. G.; Lee, L. T.; Yip, A.; Wang, Q.-Y.; Chia, C. S. B.; Hill, J.; Shi, P.-Y.; Keller, T. H. NMR analysis of a novel enzymatically active nnlinked Dengue NS2B-NS3 protease complex. *J. Biol. Chem.* **2013**, *288*, 12891−12900.
(23) de la Cruz, L.; Chen, W.-N.; Graham, B.; Otting, G. Binding mode of the activity-modulating C-terminal segment of NS2B to NS3 in the dengue virus NS2B-NS3 protease. *FEBS J.* **2014**, *281*, 1517−1533.
(24) Roy, A.; Lim, L.; Srivastava, S.; Lu, Y.; Song, J. Solution conformations of Zika NS2B-NS3pro and its inhibition by natural products from edible plants. *PLoS One* **2017**, *12*, No. e0180632.
(25) Lee, W. H. K.; Liu, W.; Fan, J.-S.; Yang, D. Dengue virus protease activity modulated by dynamics of protease cofactor. *Biophys. J.* **2021**, *120*, 2444−2453.
(26) Akke, M.; Weininger, U. NMR studies of aromatic ring flips to probe conformational fluctuations in proteins. *J. Phys. Chem. B* **2023**, *127*, 591−599.
(27) Zhang, Z.; Li, Y.; Loh, Y. R.; Phoo, W. W.; Hung, A. W.; Kang, C.; Luo, D. Crystal structure of unlinked NS2B-NS3 protease from Zika virus. *Science* **2016**, *354*, 1597−1600.
(28) Phoo, W. W.; Li, Y.; Zhang, Z.; Lee, M. Y.; Loh, Y. R.; Tan, Y. B.; Ng, E. Y.; Lescar, J.; Kang, C.; Luo, D. Structure of the NS2B-NS3 protease from Zika virus after self-cleavage. *Nat. Commun.* **2016**, *7*, No. 13410.
(29) Williams, S. P.; Haggie, P. M.; Brindle, K. M. ^{19}F NMR measurements of the rotational mobility of proteins in vivo. *Biophys. J.* **1997**, *72*, 490−498.
(30) Ye, Y.; Wu, Q.; Zheng, W.; Jiang, B.; Pielak, G. J.; Liu, M.; Li, C. Quantification of size effect on protein rotational mobility in cells by ^{19}F NMR spectroscopy. *Anal. Bioanal. Chem.* **2018**, *410*, 869−874.
(31) Bann, J. G.; Pinkner, J.; Hultgren, S. J.; Frieden, C. Real-time and equilibrium ^{19}F-NMR studies reveal the role of domain−domain interactions in the folding of the chaperone PapD. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 709−714.
(32) Schuler, B.; Kremer, W.; Kalbitzer, H. R.; Jaenicke, R. Role of entropy in protein thermostability: folding kinetics of a hyperthermophilic cold shock protein at high temperatures using ^{19}F NMR. *Biochemistry* **2002**, *41*, 11670−11680.
(33) Shu, Q.; Frieden, C. Urea-dependent unfolding of murine adenosine deaminase: sequential destabilization as measured by ^{19}F NMR. *Biochemistry* **2004**, *43*, 1432−1439.
(34) Evanics, F.; Bezsonova, I.; Marsh, J.; Kitevski, J. L.; Forman-Kay, J. D.; Prosser, R. S. Tryptophan solvent exposure in folded and unfolded states of an SH3 domain by ^{19}F and ^{1}H NMR. *Biochemistry* **2006**, *45*, 14120−14128.
(35) Li, H.; Frieden, C. Comparison of C40/82A and P27A C40/82A barstar mutants using ^{19}F NMR. *Biochemistry* **2007**, *46*, 4337−4347.
(36) Sun, X.; Dyson, H. J.; Wright, P. E. Fluorotryptophan incorporation modulates the structure and stability of transthyretin in a site-specific manner. *Biochemistry* **2017**, *56*, 5570−5581.
(37) Luck, L. A.; Vance, J. E.; O'Connell, T. M.; London, R. E. ^{19}F NMR relaxation studies on 5-fluorotryptophan- and tetradeutero-5-fluorotryptophan-labeled *E. coli* glucose/galactose receptor. *J. Biomol. NMR* **1996**, *7*, 261−272.
(38) Aramini, J. M.; Hamilton, K.; Ma, L.-C.; Swapna, G. V. T.; Leonard, P. G.; Ladbury, J. E.; Krug, R. M.; Montelione, G. T. ^{19}F NMR reveals multiple conformations at the dimer interface of the nonstructural protein 1 effector domain from influenza A virus. *Structure* **2014**, *22*, 515−525.
(39) Ruben, E. A.; Gandhi, P. S.; Chen, Z.; Koester, S. K.; DeKoster, G. T.; Frieden, C.; Di Cera, E. ^{19}F NMR reveals the conformational properties of free thrombin and its zymogen precursor prethrombin-2. *J. Biol. Chem.* **2020**, *295*, 8227−8235.
(40) Gerig, J. T. Fluorine NMR of proteins. *Prog. Nucl. Magn. Reson. Spectrosc.* **1994**, *26*, 293−370.
(41) Danielson, M. A.; Falke, J. J. Use of ^{19}F NMR to probe protein structure and conformational changes. *Annu. Rev. Biophys.* **1996**, *25*, 163−195.
(42) Leung, E. W. W.; Yagi, H.; Harjani, J. R.; Mulcair, M. D.; Scanlon, M. J.; Baell, J. B.; Norton, R. S. ^{19}F NMR as a probe of ligand interactions with the iNOS binding site of SPRY domain-containing SOCS box protein 2. *Chem. Biol. Drug Des.* **2014**, *84*, 616−625.
(43) Ge, X.; MacRaild, C. A.; Devine, S. M.; Debono, C. O.; Wang, G.; Scammells, P. J.; Scanlon, M. J.; Anders, R. F.; Foley, M.; Norton, R. S. Ligand-induced conformational change of *Plasmodium falciparum* AMA1 detected using ^{19}F NMR. *J. Med. Chem.* **2014**, *57*, 6419−6427.
(44) Ayoub, A. M.; Hawk, L. M. L.; Herzig, R. J.; Jiang, J.; Wisniewski, A. J.; Gee, C. T.; Zhao, P.; Zhu, J.-Y.; Berndt, N.; Offei-Addo, N. K.; Scott, T. G.; Qi, J.; Bradner, J. E.; Ward, T. R.; Schönbrunn, E.; Georg, G. I.; Pomerantz, W. C. K. BET bromodomain inhibitors with one-step synthesis discovered from virtual screen. *J. Med. Chem.* **2017**, *60*, 4805−4817.
(45) Shanina, E.; Siebs, E.; Zhang, H.; Varón Silva, D.; Joachim, I.; Titz, A.; Rademacher, C. Protein-observed ^{19}F NMR of LecA from *Pseudomonas aeruginosa*. *Glycobiology* **2021**, *31*, 159−165.
(46) Cosottini, L.; Zineddu, S.; Massai, L.; Ghini, V.; Turano, P. ^{19}F: A small probe for a giant protein. *J. Inorg. Biochem.* **2023**, *244*, No. 112236.
(47) Boeszoermenyi, A.; Chhabra, S.; Dubey, A.; Radeva, D. L.; Burdzhiev, N. T.; Chanev, C. D.; Petrov, O. I.; Gelev, V. M.; Zhang, M.; Anklin, C.; Kovacs, H.; Wagner, G.; Kuprov, I.; Takeuchi, K.; Arthanari, H. Aromatic ^{19}F-^{13}C TROSY: A background-free approach to probe biomolecular structure, function, and dynamics. *Nat. Methods* **2019**, *16*, 333−340.
(48) Maleckis, A.; Herath, I. D.; Otting, G. Synthesis of ^{13}C/^{19}F/^{2}H labeled indoles for use as tryptophan precursors for protein NMR spectroscopy. *Org. Biomol. Chem.* **2021**, *19*, 5133−5147.
(49) Roos, M.; Wang, T.; Shcherbakov, A. A.; Hong, M. Fast magic-angle-spinning ^{19}F spin exchange NMR for determining nanometer ^{19}F−^{19}F distances in proteins and pharmaceutical compounds. *J. Phys. Chem. B* **2018**, *122*, 2900−2911.
(50) Fritz, M.; Kraus, J.; Quinn, C. M.; Yap, G. P. A.; Struppe, J.; Sergeyev, I. V.; Gronenborn, A. M.; Polenova, T. Measurement of accurate interfluorine distances in crystalline organic solids: a high-frequency magic angle spinning NMR approach. *J. Phys. Chem. B* **2019**, *123*, 10680−10690.
(51) Matei, E.; Gronenborn, A. M. ^{19}F paramagnetic relaxation enhancement: a valuable tool for distance measurements in proteins. *Angew. Chem., Int. Ed.* **2016**, *55*, 150−154.

(52) Seal, M.; Zhu, W.; Dalaloyan, A.; Feintuch, A.; Bogdanov, A.; Frydman, V.; Su, X.-C.; Gronenborn, A. M.; Goldfarb, D. Gd^{III}-^{19}F distance measurements for proteins in cells by electron-nuclear double resonance. *Angew. Chem., Int. Ed.* **2023**, *62*, No. e202218780.
(53) Judd, M.; Abdelkader, E. H.; Qi, M.; Harmer, J. R.; Huber, T.; Godt, A.; Savitsky, A.; Otting, G.; Cox, N. Short-range ENDOR distance measurements between Gd(III) and trifluoromethyl labels in proteins. *Phys. Chem. Chem. Phys.* **2022**, *24*, 25214−25226.
(54) Bogdanov, A.; Frydman, V.; Seal, M.; Rapatskiy, L.; Schnegg, A.; Zhu, W.; Iron, M.; Gronenborn, A. M.; Goldfarb, D. Extending the range of distances accessible by ^{19}F electron−nuclear double resonance in proteins using high-spin Gd(III) labels. *J. Am. Chem. Soc.* **2024**, *146*, 6157−6167.
(55) Staudt, H.; Hoesl, M. G.; Dreuw, A.; Serdjukow, S.; Oesterhelt, D.; Budisa, N.; Wachtveitl, J.; Grininger, M. Directed manipulation of a flavoprotein photocycle. *Angew. Chem., Int. Ed.* **2013**, *52*, 8463−8466.
(56) Lu, M.; Toptygin, D.; Xiang, Y.; Shi, Y.; Schwieters, C. D.; Lipinski, E. C.; Ahn, J.; Byeon, I.-J. L.; Gronenborn, A. M. The magic of linking rings: discovery of a unique photoinduced fluorescent protein crosslink. *J. Am. Chem. Soc.* **2022**, *144*, 10809−10816.
(57) Minks, C.; Huber, R.; Moroder, L.; Budisa, N. Atomic mutations at the single tryptophan residue of human recombinant annexin V: effects on structure, stability, and activity. *Biochemistry* **1999**, *38*, 10649−10659.
(58) Zhang, Q.-s.; Shen, L.; Wang, E.-d.; Wang, Y.-l. Biosynthesis and Characterization of 4-Fluorotryptophan-Labeled *Escherichia coli* Arginyl-tRNA Synthetase. *J. Protein Chem.* **1999**, *18*, 187−192.
(59) Welte, H.; Zhou, T.; Mihajlenko, X.; Mayans, O.; Kovermann, M. What does fluorine do to a protein? Thermodynamic, and highly-resolved structural insights into fluorine-labelled variants of the cold shock protein. *Sci. Rep.* **2020**, *10*, No. 2640.
(60) Hancock, S. M.; Uprety, R.; Deiters, A.; Chin, J. W. Expanding the genetic code of yeast for incorporation of diverse unnatural amino acids via a pyrrolysyl-tRNA synthetase/tRNA pair. *J. Am. Chem. Soc.* **2010**, *132*, 14819−14824.
(61) Cronin, C. N. Optimization of genetic code expansion in the baculovirus expression vector system (BEVS). *Protein Expression Purif.* **2023**, *210*, No. 106314.
(62) Meineke, B.; Heimgärtner, J.; Lafranchi, L.; Elsässer, S. J. *Methanomethylophilus alvus Mx1201* provides basis for mutual orthogonal pyrrolysyl tRNA/aminoacyl-tRNA synthetase pairs in mammalian cells. *ACS Chem. Biol.* **2018**, *13*, 3087−3096.
(63) Beránek, V.; Willis, J. C. W.; Chin, J. W. An evolved *Methanomethylophilus alvus* pyrrolysyl-tRNA synthetase/tRNA pair is highly active and orthogonal in mammalian cells. *Biochemistry* **2019**, *58*, 387−390.
(64) Leone, M.; Rodriguez-Mias, R. A.; Pellecchia, M. Selective incorporation of ^{19}F-labeled Trp side chains for NMR-spectroscopy-based ligand-protein interaction studies. *ChemBioChem* **2003**, *4*, 649−650.
(65) Kim, H.-W.; Perez, J. A.; Ferguson, S. J.; Campbell, I. D. The specific incorporation of labelled aromatic amino acids into proteins through growth of bacteria in the presence of glyphosate: Application to fluorotryptophan labelling to the H^+-ATPase of *Escherichia coli* and NMR studies. *FEBS Lett.* **1990**, *272*, 34−36.
(66) Pfefferkorn, C. M.; Lee, J. C. 5-Fluoro-D,L-tryptophan as a Dual NMR and Fluorescent Probe of α-Synuclein. In *Intrinsically Disordered Protein Analysis*; Uversky, V. N.; Dunker, A. K., Eds.; Humana Press: Totowa, NJ, 2012; pp 197−209.
(67) Wang, X.; Mercier, P.; Letourneau, P.-J.; Sykes, B. D. Effects of Phe-to-Trp mutation and fluorotryptophan incorporation on the solution structure of cardiac troponin C, and analysis of its suitability as a potential probe for *in situ* NMR studies. *Protein Sci.* **2005**, *14*, 2447−2460.
(68) Kenward, C.; Shin, K.; Rainey, J. K. Mixed fluorotryptophan substitutions at the same residue expand the versatility of 19F protein NMR spectroscopy. *Chem. - Eur. J.* **2018**, *24*, 3391−3396.
(69) Khan, F.; Kuprov, I.; Craggs, T. D.; Hore, P. J.; Jackson, S. E. ^{19}F NMR studies of the native and denatured states of green fluorescent protein. *J. Am. Chem. Soc.* **2006**, *128*, 10729−10737.
(70) Crowley, P. B.; Kyne, C.; Monteith, W. B. Simple and inexpensive incorporation of ^{19}F-tryptophan for protein NMR spectroscopy. *Chem. Commun.* **2012**, *48*, 10681−10683.
(71) Curtis-Marof, R.; Doko, D.; Rowe, M. L.; Richards, K. L.; Williamson, R. A.; Howard, M. J. ^{19}F NMR spectroscopy monitors ligand binding to recombinantly fluorine-labelled b′x from human protein disulphide isomerase (hPDI). *Org. Biomol. Chem.* **2014**, *12*, 3808−3812.
(72) Lu, M.; Ishima, R.; Polenova, T.; Gronenborn, A. M. ^{19}F NMR relaxation studies of fluorosubstituted tryptophans. *J. Biomol. NMR* **2019**, *73*, 401−409.
(73) Agostini, F.; Sinn, L.; Petras, D.; Schipp, C. J.; Kubyshkin, V.; Berger, A. A.; Dorrestein, P. C.; Rappsilber, J.; Budisa, N.; Koksch, B. Multiomics analysis provides insight into the laboratory evolution of *Escherichia coli* toward the metabolic usage of fluorinated indoles. *ACS Cent. Sci.* **2021**, *7*, 81−92.
(74) Mahawaththa, M. C.; Pearce, B. J. G.; Szabo, M.; Graham, B.; Klein, C. D.; Nitsche, C.; Otting, G. Solution conformations of a linked construct of the Zika virus NS2B-NS3 protease. *Antiviral Res.* **2017**, *142*, 141−147.
(75) Maxwell, M.; Tan, Y. J.; Lee, R.; Huber, T.; Otting, G. Electrostatic contribution to ^{19}F chemical shifts in fluorotryptophans in proteins. *Biochemistry* **2023**, *62*, 3255−3264.
(76) Visser, N. V.; Westphal, A. H.; Nabuurs, S. M.; van Hoek, A.; van Mierlo, C. P. M.; Visser, A. J. W. G.; Broos, J.; van Amerongen, H. 5-Fluorotryptophan as dual probe for ground-state heterogeneity and excited-state dynamics in apoflavodoxin. *FEBS Lett.* **2009**, *583*, 2785−2788.
(77) Liu, T.; Callis, P. R.; Hesp, B. H.; de Groot, M.; Buma, W. J.; Broos, J. Ionization potentials of fluoroindoles and the origin of nonexponential tryptophan fluorescence decay in proteins. *J. Am. Chem. Soc.* **2005**, *127*, 4104−4113.
(78) Sarkar, S. S.; Udgaonkar, J. B.; Krishnamoorthy, G. Reduced fluorescence lifetime heterogeneity of 5-fluorotryptophan in comparison to tryptophan in proteins: implication for resonance energy transfer experiments. *J. Phys. Chem. B* **2011**, *115*, 7479−7486.
(79) Toptygin, D.; Gronenborn, A. M.; Brand, L. Nanosecond relaxation dynamics of protein GB1 identified by the time-dependent red shift in the fluorescence of tryptophan and 5-fluorotryptophan. *J. Phys. Chem. B* **2006**, *110*, 26292−26302.
(80) Xu, J.; Chen, B.; Callis, P.; Muiño, P. L.; Rozeboom, H.; Broos, J.; Toptygin, D.; Brand, L.; Knutson, J. R. Picosecond fluorescence dynamics of tryptophan and 5-fluorotryptophan in monellin: slow water−protein relaxation unmasked. *J. Phys. Chem. B* **2015**, *119*, 4230−4239.
(81) Bronskill, P. M.; Wong, J. T. Suppression of fluorescence of tryptophan residues in proteins by replacement with 4-fluorotryptophan. *Biochem. J.* **1988**, *249*, 305−308.
(82) Hott, J. L.; Borkman, R. F. The non-fluorescence of 4-fluorotryptophan. *Biochem. J.* **1989**, *264*, 297−299.
(83) Ross, J. B. A.; Szabo, A. G.; Hogue, C. W. V. Enhancement of Protein Spectra with Tryptophan Analogs: Fluorescence Spectroscopy of Protein−Protein and Protein−Nucleic Acid Interactions. In *Methods in Enzymology*; Academic Press, 1997; Vol. *278*, pp 151−190.
(84) Vos, E. P. P.; Bokhove, M.; Hesp, B. H.; Broos, J. Structure of the cytoplasmic loop between putative helices II and III of the mannitol permease of *Escherichia coli*: a tryptophan and 5-fluorotryptophan spectroscopy study. *Biochemistry* **2009**, *48*, 5284−5290.
(85) Joel, S.; Turner, K. B.; Daunert, S. Glucose recognition proteins for glucose sensing at physiological concentrations and temperatures. *ACS Chem. Biol.* **2014**, *9*, 1595−1602.

https://doi.org/10.1021/jacs.4c03743
J. Am. Chem. Soc. 2024, 146, 13641−13650