Title

Genetically Enabling Phosphorus Fluoride Exchange Click Chemistry in Proteins.

Journal

Chem, 10(6)

ISSN

2451-9294

Authors

Cao, Li
Yu, Bingchen
Li, Shanshan
et al.

Publication Date

2024-06-13

DOI

10.1016/j.chempr.2024.02.010

Peer reviewed

HHS Public Access

Author manuscript

Chem. Author manuscript; available in PMC 2024 July 22.

Published in final edited form as:

Chem. 2024 June 13; 10(6): 1868–1884. doi:10.1016/j.chempr.2024.02.010.

Genetically Enabling Phosphorus Fluoride Exchange Click Chemistry in Proteins

Li Cao[1,2], **Bingchen Yu**[1,2], **Shanshan Li**[1], **Pan Zhang**[1], **Qingke Li**[1], **Lei Wang**[1,3,*]

[1]Department of Pharmaceutical Chemistry, the Cardiovascular Research Institute, and Hellen Diller Family Comprehensive Cancer Center, University of California, San Francisco, San Francisco, CA 94158, USA

[2]These authors contributed equally

[3]Lead contact

SUMMARY

Phosphorus Fluoride Exchange (PFEx), recently debuted in small molecules, represents the forefront of click chemistry. To explore PFEx's potential in biological settings, we developed amino acids PFY and PFK featuring phosphoramidofluoridates and incorporated them into proteins through genetic code expansion. PFY/PFK selectively reacted with nearby His, Tyr, Lys, or Cys in proteins, both *in vitro* and in living cells, demonstrating that proximity enabled PFEx reactivity without external reagents. The reaction with His showed unique pH-dependent properties and created thermally sensitive linkages. Additionally, Na_2SiO_3 enhanced PFEx reactions with Tyr and Cys. PFEx, by generating defined covalent P-N/O linkages, extends the utility of phosphorus linkages in proteins, aligning with nature's use of phosphate connectors in other biomolecules. More versatile and durable than SuFEx, PFEx in proteins expands the latent bioreactive arsenal for covalent protein engineering and will facilitate the broad application of this potent click chemistry in biological and biomedical fields.

eTOC blurb

Phosphorus Fluoride Exchange (PFEx) is a cutting-edge click reaction recently developed in small molecules. In this work, we expanded PFEx to proteins by genetically encoding two latent bioreactive amino acids, PFY and PFK, in *E. coli* and mammalian cells. PFY/PFK reacts with Tyr,

*Correspondence: Lei.Wang2@ucsf.edu.

ARTHOR CONTRIBUTIONS

L.C. identified PFYRS, incorporated PFY and PFK into proteins followed with characterization, studied protein cross-linking *in vitro* and in cells, investigated various factors affecting PFEx in proteins, designed experiments, analyzed data, and prepared figures. B.Y. synthesized PFY and PFK, wrote the synthesis procedures, and identified PFKRS. S.L. incorporated PFK in mammalian cells and studied mNb6(PFK) cross-linking with the Spike protein. P.Z. analyzed PFK hydrolysis. Q.L. helped with flow cytometric analysis of PFY incorporation. L.W. conceived and supervised the project and analyzed the data. L.C. and L.W. wrote the manuscript.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online.

DELCARATION OF INTERESTS

L.W. is a Scientific Advisor for Enlaza Therapeutics.

His, Lys, and Cys in proteins through proximity-enabled PFEx reaction, creating specific P-N/O linkages in vitro and in vivo. The introduction of PFEx into proteins paves the way for its broad applications in biological and biomedical research.

INTRODUCTION

Click chemistry has revolutionized the synthesis of functional molecules and found wide applications in biomolecular research and engineering.[1,2,3] Phosphorus Fluoride Exchange (PFEx) has recently been developed and heralded as the new frontier of click chemistry.[4] Catalyzed by Lewis base and in the presence of silicon additive, the P(V)-F bonds are exchanged with incoming nucleophiles to produce stable tetrahedral P(V)–O and P(V)–N linked products in PFEx reactions. PFEx stands out for its highly chemoselectivity towards alcohols and amines, and its orthogonality to other click chemistries like Copper-catalyzed Azide-Alkyne Cycloaddition (CuAAC) and Sulfur Fluoride Exchange (SuFEx), enabling successive reactions on a single molecule. Previously, P(V)-F containing compounds have also been synthesized and employed in activity-based protein profiling, especially for serine hydrolases, though they often exhibit cytotoxicity.[5–8] The expansion of click chemistry from small molecules to biomacromolecules has significantly transformed our research, manipulation, and engineering of biomolecules and their functions, as demonstrated by the profound impact of CuAAC and SuFEx on biological studies.[9,10] Having just debuted in small molecules and showing potential for generating intricate structures,[4] PFEx is now poised to make significant strides in the realm of proteins and other biomolecules.

Introducing new covalent bonding capabilities into proteins is significantly advancing protein research, engineering, and applications.[11] This innovation is achieved through incorporation of latent bioreactive unnatural amino acids (Uaas) into proteins via genetic code expansion.[12,13] These Uaas react with nearby target natural amino acid residues through proximity-enabled reactivity, creating covalent linkages within or between proteins and various biomolecules.[14,15] Such covalent modifications have introduced novel properties to proteins,[13,16,17] helped identify elusive biomolecular interactions,[15,18–22] and contributed to the development of covalent protein therapeutic candidates.[23–26] While various functionalities have been explored for latent bioreactive Uaas,[27–36] challenges remain in complex biological settings, especially regarding cytotoxicity and biocompatibility.[37] To date, SuFEx-based latent bioreactive Uaas have shown promising biocompatibility across a range of proteins and applications.[20,23,24,34,38] However, these Uaas are limited in their ability to form stable linkages with Cys, Ser, or Thr in proteins, and the linkages they form with other residues generally lack responsiveness to environmental changes.[39,40] New biocompatible functional groups and chemistries are increasingly sought after to meet the growing demands of targeting a diverse range of biological entities, enhancing detection sensitivity, and improving the specificity and effectiveness of therapeutics. Phosphoramidofluoridates, utilized in small molecule PFEx reactions, exhibit thermal stability and resistance to hydrolysis under biological conditions,[4,6] positioning them as promising candidates for latent bioreactive applications. PFEx is unique in its creation of phosphorus-based linkages, akin to those naturally occurring in biomolecules, offering potential enhanced biocompatibility, a feature not achievable with existing latent bioreactive

functional groups. Additionally, the versatility of PFEx in reacting with various nucleophiles could allow for targeting multiple amino acid residues, broadening its scope for multi-target applications. The emergent nature of PFEx reactions and the resulting linkages present opportunities for innovative covalent manipulation in biochemical research. However, the current methodology for PFEx in small molecules, requiring a strongly basic catalyst and a water-insoluble silicon additive,[4] poses challenges for its direct application to proteins and other biological settings.

In this study, we have developed two latent bioreactive Uaas: phosphoramidofluoridate tyrosine (PFY) and phosphoramidofluoridate lysine (PFK). Demonstrating exceptional biocompatibility, these Uaas have been successfully incorporated into proteins in both *E. coli* and mammalian cells through genetic code expansion. Utilizing proteins integrated with PFY and PFK, we established that PFEx can be activated within proteins solely through proximity effects, obviating the need for external reagents. This intrinsic activation enables the formation of covalent phosphorus linkages between interacting proteins, a process we have demonstrated effectively *in vitro* and, more excitingly, directly within living cells. Impressively, PFEx extends its targeting capability to include amino acid residues Lys, His, Tyr, and Cys, and exhibits unique pH and thermal sensitivities. Additionally, the water-soluble Na_2SiO_3 was found to enhance PFEx reaction with Tyr and Cys residues. These characteristics of PFEx greatly enhance its utility for the versatile covalent manipulation of biomacromolecules, offering new prospects to advance research and engineering applications.

RESULTS

Design, synthesis, and genetic incorporation of PFY into proteins in *E. coli* and mammalian cells

We designed PFY as a tyrosine analog to contain a phosphoramidofluoridate (Figure 1A), because phosphoramidic difluorides decompose at room temperature in several hours while phosphoramidofluoridates are bench stable.[4,41] We hypothesized that the phosphoramidofluoridate side chain of PFY would react with nucleophilic side chains of natural amino acid residues via PFEx when they are brought into close proximity, effect of which would be able to activate the PFEx reactivity of PFY (Figure 1B).

To synthesize PFY, the dimethylphosphoramidic difluoride **2** was first prepared from **1** through fluoride-chloride halogen exchange,[4] and was then directly used for PFEx reaction with a protected tyrosine **3** followed by deprotection to give PFY in overall yield of 57% (Figure 1C, Supplementary information). To evaluate PFY's cytotoxicity, we added different concentrations of PFY to *E. coli* or mammalian cells. The growth of *E. coli* was not affected by 4 mM of PFY in the growth media (Figure S1A), and cell viability assay showed that HEK-293T cells could tolerate 2 mM PFY well (Figure S1B), suggesting no obvious toxicity of PFY to both bacterial and mammalian cells.

To genetically encode PFY into proteins in *E. coli*, we identified a mutant pyrrolysyl-tRNA synthetase (PylRS) specific for PFY to pair with the pyrrolysyl-tRNA ($tRNA^{Pyl}$) for PFY incorporation through genetic code expansion via amber suppression. As PFY is similar

to Uaas NpY,[42] FSY,[34] and mFSY[38] in structure, we tested if synthetases specific for these Uaas could incorporate PFY. Using a GFP-based fluorescence assay we found that synthetases evolved for NpY and mFSY both were able to incorporate PFY, and the mFSY-specific synthetase showed higher incorporation efficiency (Figure S2). We thus used this synthetase for subsequent experiments and named it PFYRS for clarity.

To evaluate the incorporation specificity of PFY by the $tRNA^{Pyl}$/PFYRS pair in *E. coli*, we expressed the Z_{spa} affibody (Afb) gene containing a TAG codon at site 36 (Afb-36TAG) together with the $tRNA^{Pyl}$/PFYRS genes in *E. coli*. No full-length Afb was detected when PFY was not added in the cell culture; in the presence of 2 mM PFY, full-length Afb36PFY protein was produced in the yield of 5.0 mg/L (Figure 1D). We then purified the Afb(36PFY) protein, digested it with proteases, and analyzed with tandem mass spectrometry (MS) (Figure 1E). A series of b and y ions clearly indicated that PFY was incorporated at the TAG-specified position 36, but the dimethylamino group of PFY was hydrolyzed. This observation is consistent with the known hydrolysis of phosphoramidate group under acidic conditions used in liquid chromatography (LC) during MS analysis, and was also confirmed by subjecting the amino acid to LC-MS analysis conditions (Figure S3).[42] To further verify PFY incorporation, we used the $tRNA^{Pyl}$/PFYRS pair to incorporate PFY into another protein, ubiquitin (Ub), at sites 51 and 54, respectively. The Ub proteins were purified and analyzed using electrospray ionization time-of-flight MS (ESI-TOF-MS). For the Ub(51PFY) intact protein (Figure 1F), a peak observed at 9530 Da corresponds to Ub containing intact PFY at site 51 (expected 9530.7 Da); the other peak observed at 9503 Da corresponds to Ub containing PFY at site 51 but with the dimethylamino group of PFY hydrolyzed (expected 9503.6 Da); no peaks corresponding to Ub containing other amino acids at site 51 were observed. Similar results were also obtained for the Ub(54PFY) intact protein, in which PFY was incorporated at a different site 54 (Figure S4). The data obtained from SDS-PAGE, Western blot, and intact MS collectively indicate that the fidelity of PFY incorporation in *E. coli* exceeded 99.9%. These results demonstrate that the identified $tRNA^{Pyl}$/PFYRS pair incorporated PFY into proteins with high efficiency and specificity in *E. coli*.

We further evaluated PFY incorporation in mammalian cells. The $tRNA^{Pyl}$/PFYRS genes were co-transfected with the EGFP gene containing a TAG stop codon at the permissive site 182 [EGFP(182TAG) gene] into HEK-293T cells. Only when PFY was added to the growth media did cells show green GFP fluorescence, suggesting the expression of full-length GFP by incorporating PFY at the TAG codon (Figure 1G); Western blot analysis of the cell lysate further confirmed full-length GFP expression only in the presence of PFY (Figure 1H). These cells were next analyzed by flow cytometry. The percentage of GFP-positive cells increased with PFY concentration and incubation time (Figure 1I). The total fluorescence intensity in cells also increased with PFY concentrations, showing significant increases of 472-fold at 0.1 mM, 1194-fold at 0.5 mM, 1770-fold at 1.0 mM, and 2526-fold at 2.0 mM (Figure 1J). Based on the results from Western blot analysis and flow cytometry, the fidelity of PFY incorporation in mammalian cells was found to exceed 99.9%. These results indicate that the $tRNA^{Pyl}$/PFYRS pair was able to incorporate PFY in mammalian cells efficiently and specifically.

PFY reacts with His, Tyr, Lys, and Cys in proteins through proximity-enabled PFEx reaction

We assessed whether PFY could react with natural amino acid side chains in proteins through proximity-enabled PFEx reaction. The Afb-Z protein pair was used, which binds in moderate affinity (K_d 6 μM). To place PFY and target residue in proximity upon Afb-Z binding, we incorporated PFY in Afb at site 36 and mutated different amino acids at site 6 of the Z protein on the basis of the Afb-Z complex structure (Figure 2A).[43] Considering that the phosphoramidofluoridate in PFY is a weak electrophile, we decided to test amino acid residues with nucleophilic side chains. Maltose binding protein (MBP) was fused to the N-terminus of the Z protein to better separate the Afb and Z proteins of similar molecular weights. The purified Afb(36PFY) and MBP-Z(6X) proteins were incubated in PBS buffer for 16 hours and then analyzed by SDS-PAGE and Western blot under denatured conditions (Figure 2B, 2C). Among 12 amino acid residues tested, Lys and Cys showed relatively weak crosslinking, while His and Tyr showed strong crosslinking. To validate the crosslink, the cross-linked protein samples were analyzed by tandem MS. Covalently crosslinked peptides of Afb(36PFY) and MBP-Z(6X) were clearly identified, and fragmented ions unambiguously indicate that the incorporated PFY specifically cross-linked with the target His, Tyr, Lys, or Cys placed in proximity (Figure 2D, 2E and Figure S5).

To show PFEx reactivity in a different protein context, we incorporated PFY into a nanobody SR4 that binds the Spike protein of SARS-CoV-2. Based on the crystal structure of the SR4-Spike complex,[44] we incorporated PFY at site 57 of SR4 to target a proximal Tyr505 of the Spike protein. As a control, PFY was separately incorporated at site 54 of SR4, which is further away from Tyr505. Indeed, SR4(57PFY) cross-linked the Spike protein efficiently, while SR4(54PFY) did not, confirming that PFY reactivity was proximity-driven (Figure S6).

Besides *in vitro* cross-linking, we further evaluated PFY reactivity in proteins directly in mammalian cells. We incorporated PFY into *E. coli* glutathione transferase (ecGST), a homodimeric protein, at its dimer interface and expressed the mutant ecGST in HEK-293T cells to test PFY's ability to crosslink ecGST into a covalent dimer in a mammalian cellular environment. Specifically, in light of the structure of ecGST (Figure 2F),[45] we incorporated PFY at site 103 of ecGST to target the proximal His106 in the other monomer. The nearby Lys107 was mutated to Ala so that His106 was the only reachable target. His106 was also mutated to Tyr, Lys, or Cys for testing, and to Ala as the negative control. The mutant ecGST was expressed in HEK-293T cells, and cell lysates were analyzed by Western blot to detect ecGST dimeric cross-linking (Figure 2G). As expected, no dimeric cross-linking was observed for Ala. Strong dimeric cross-linking was observed for His and Tyr, while weak dimeric cross-linking was observed for Lys and Cys, consistent with what was observed in the Afb-Z protein system. Together, these results indicate that PFY was able to react with His, Tyr, Lys, and Cys placed in proximity in proteins both *in vitro* and in cells.

Opposing pH effects: PFY reaction with Tyr vs. His

We next studied the effect of pH on PFY reaction in proteins. Purified Afb(36PFY) was incubated with MBP-Z(6X) at pH 7.4 or 8.8 overnight, followed by SDS-PAGE analysis (Figure 3A). The cross-linking efficiency of the two proteins was determined by

densitometric measurement of the crosslink band and the MBP-Z band. The cross-linking efficiency between PFY and Lys or Cys did not change significantly upon pH change. The cross-linking efficiency between PFY and Tyr increased from 33.7% to 44.5% when pH was shifted to basic (Figure 3B), which may be explained by the enhanced deprotonation of the phenol side chain of Tyr. Unexpectedly, the cross-linking efficiency between PFY and His decreased from 35.3% to 21.3% when pH was increased to 8.8 (Figure 3B). To determine whether the PFY-His linkage becomes unstable at pH 8.8, Afb(36PFY) was initially crosslinked with MBP-Z(6H). Subsequently, aliquots of the crosslinked protein were separately exposed to pH 7.4 or 8.8 for 16 hours. The percentage of crosslinked product remained consistent across these pH levels (Figure S7), indicating that the PFY-His linkage was stable when shifting from pH 7.4 to 8.8. We also monitored the MBP-Z(6His) protein under these pH conditions using circular dichroism (CD) spectroscopy and found that these pH levels had a negligible impact on its secondary structures (Figure S8). These results suggest that the reduced crosslinking efficiency observed at pH 8.8 was likely due to a less efficient reaction under basic conditions, rather than instability of the PFY-His linkage or MBP-Z-(6His). Therefore, PFY reaction with His showed an opposite pH-dependent effect compared to PFY reaction with Tyr in proteins.

We further compared the pH effects on the reactions of PFY and FSY with His in proteins. Uaa FSY reacts with His through proximity-enabled SuFEx click chemistry.[34] Purified MBP-Z(6His) protein was incubated with either Afb(36PFY) or Afb(36FSY) at various pHs overnight, followed by SDS-PAGE analysis (Figure 3C). The cross-linking efficiency between FSY and His increased from pH 6.0 to 7.4 and plateaued at 8.8, while the cross-linking efficiency between PFY and His had the opposite trend, plateauing from pH 6 to 6.5 and then decreasing from pH 6.5 to 8.8 (Figure 3D). These data indicate that the reaction of PFY with His demonstrates a pH effect that is also opposite to that of the FSY reaction with His. Together, these results demonstrate that PFY's reaction with His exhibited atypically increased reactivity at acidic pHs.

Contrasting thermostability: PFY-Tyr and PFY-Cys vs. PFY-His linkages

We discovered unexpectedly that the reaction products of PFY with Tyr and His showed divergent thermostability. We first incubated Afb(36PFY) with MBP-Z(6Tyr) or MBP-Z(6His) at 37 °C for 16 hours to allow cross-linking. Aliquots of samples were treated with or without 95 °C heating for 10 minutes followed with SDS-PAGE (Figure 4A, 4B). MBP-Z(6Tyr) demonstrated strong cross-linking with Afb(36PFY) regardless the heating, showing 27.5% for the sampled heated 95 °C and 26.8% for the sample maintained at 37 °C (Figure 4C), indicating that the resultant P(V)-O linkage was thermostable under the test condition. In contrast, MBP-Z(6His) exhibited strong cross-linking, yielding 37.4% in the absence of heat treatment. However, when exposure to 95 °C, there was a significant reduction in the cross-linked product, dropping to 5.0% (Figure 4D). This outcome indicates that the associated P(V)-N linkage was unstable at 95 °C. Further, we subjected the Afb(36PFY)/MBP-Z(6His) cross-linking product to various temperatures for durations of 5 and 10 minutes. We observed appreciable breakage of the P(V)-N linkage at 50 °C for 5 minutes, with increased breakage noted with longer exposure times and higher temperatures (Figure 4E). We then tested incubation of Afb(36PFY) with MBP-Z(6His) or MBP-Z(6Tyr)

at 4 °C for 16 hours followed with SDS-PAGE analysis without heating the samples. At this low temperature, both His and Tyr were capable of forming crosslinks with PFY, but His showed greater efficiency (21.4%) in comparison to Tyr (5.3%) (Figure 4C, 4D). Therefore, His reacted with PFY efficiently at low temperature 4 °C but its P(V)-N linkage was unstable at temperature higher than 50 °C, while Tyr needed higher temperature 37 °C to react with PFY efficiently and its P(V)-O linkage remained stable at 95 °C. Using another protein pair which afforded robust cross-linking of PFY with Cys, we confirmed that the P(V)-S linkage was also stable at 95 °C (Figure S9).

Enhanced durability of PFY compared to FSY in proteins

The durability of the latent bioreactive Uaa's reactivity in proteins is important for their in vivo applications, particularly in covalent protein therapeutics. To evaluate and compare the durability of PFY and FSY, we incubated Afb(36PFY) or Afb(36FSY) in PBS buffer (pH 7.4) at 37 °C. Aliquots of samples were taken out at different incubation times and cross-linked with MBP-Z(6His) protein to determine the Afb protein's remaining reactivity. As shown in Figure 4F, the reactivity of Afb(36FSY) decreased significantly over time, with its cross-linking efficiency dropping from 60.2% on day 0 to 47.7% on day 1, 27.5% on day 3, and 17.9% on day 5. In contrast, Afb(36PFY) maintained its reactivity well throughout the incubation period.

Varied effects of Na_2SiO_3: PFY reaction with Tyr and Cys vs. His

We further explored potential reagents to catalyze the PFEx reaction in proteins. PFEx reactions among small molecules require a catalyst, for which 1,5,7-triazabicyclo[4.4.0]dec-5-ene (TBD) has been reported to be most efficient, and the presence of the silicon-containing additive hexamethyldisilazane.[4] Using Afb(36PFY) cross-linking with MBP-Z(6X) and MBP-Z(24PFY) cross-linking with Afb(7X), we found that TBD had no catalytic effect for PFY reactions in proteins (Figure S10), possibly because the solvent for proteins has to be aqueous while that for small molecules is mainly the aprotic acetonitrile. The additive hexamethyldisilazane is insoluble in water and thus cannot be used for proteins. We reasoned that the water soluble Na_2SiO_3 might function similarly as SiO_2 is often used to scrub fluoride.

Indeed, when Na_2SiO_3 was added to the reaction of MBP-Z(24PFY) with Afb(7X) (Figure 5A), we found that Na_2SiO_3 significantly enhanced the cross-linking between PFY and Cys (Figure 5B). The cross-linking efficiency rose from 2.3% to 9.6% with 1 mM Na_2SiO_3, reached 34.7% with 2 mM Na_2SiO_3, and remained at 34.5% with 5 mM Na_2SiO_3. However, further increasing Na_2SiO_3 to 10 mM resulted in decreased cross-linking. A similar trend was observed for the cross-linking between PFY and Tyr: efficiency increased from 2.2% to 8.7% with 1 mM Na_2SiO_3 and to 11.7% with 2 mM Na_2SiO_3, while further increasing Na_2SiO_3 to 5 mM reduced the efficiency to 3.8% (Figure 5C). Notably, the addition of Na_2SiO_3 reduced the cross-linking efficiency between PFY and His (Figure 5D): efficiency declined from 29.3% to 18.9% with 1 mM Na_2SiO_3 and further to 9.0% with 2 mM Na_2SiO_3. In this protein context, the effect of Na_2SiO_3 on PFY reaction with Lys was not pronounced, possibly because these two sites were not optimal for their crosslink. To assess how the concentration of Na_2SiO_3 affects protein stability, we conducted CD

spectroscopy on Afb(7X) proteins at various Na_2SiO_3 levels (Figure S11). The structure of Afb proteins consists of three α-helices, whose denaturation is detectable via CD. We found that 1 mM Na_2SiO_3 had a negligible impact on Afb(7H), whereas 2 mM Na_2SiO_3 led to increased unfolding. Concentrations of 5 and 10 mM Na_2SiO_3 caused significant denaturation in both Afb(7H) and the variants Afb(7Y) and Afb(7C). To assess the impact of Na_2SiO_3 on the stability of the PFY-His linkage, MBP-Z(24PFY) was first crosslinked with Afb(7H), and aliquots of this crosslinked protein were treated with varying concentrations of Na_2SiO_3. Analysis revealed that the percentage of crosslinked product remained relatively unchanged regardless of Na_2SiO_3 exposure (Figure S12), indicating that Na_2SiO_3 did not significantly affect the stability of the PFY-His linkage. Consequently, the reduced crosslinking efficiency at higher Na_2SiO_3 concentrations (5–10 mM) appears to be linked to protein unfolding. Notably, at a lower concentration of 1 mM Na_2SiO_3, there was an increase in the crosslinking efficiency of Afb(7C) and Afb(7Y), but a decrease for Afb(7H).

We next added Na_2SiO_3 to the reaction of Afb(36PFY) with MBP-Z(6X) (Figure 5E). In this protein context, PFY reaction with Lys or Cys was suboptimal, and thus the effect of Na_2SiO_3 was similarly not obvious. Contrastingly, for Tyr and His which yielded robust cross-linking with PFY, we found that the cross-linking efficiency between PFY and Tyr also increased from 24.7% to 45.7% when 1 mM Na_2SiO_3 was added (Figure 5F); consistently, the same concentration of Na_2SiO_3 reduced the cross-linking efficiency between PFY and His from 36.9% to 18.8% (Figure 5G). Collectively, these results indicate that Na_2SiO_3 could boost the PFY reaction with Tyr and Cys but diminish the PFY reaction with His.

Design, synthesis, and genetic incorporation of PFK with a flexible long side chain

PFY was designed as a Tyr analog with a rigid side chain and length close to those of canonical amino acids. To provide flexibility and longer reaction radius, we designed PFK by installing the phosphoramidofluoridate onto the Lys backbone, resulting in an extension of the final side chain length (Figure 6A). We synthesized PFK starting from dimethylphosphoramidic dichloride through fluoride-chloride halogen exchange, direct PFEx reaction with a protected 4-hydroxybenzoic acid, and conjugation to the protected Lys followed with deprotection (Supplemental Information).

To genetically encode PFK into proteins, we reasoned that FSKRS, the synthetase specific for a structurally similar Uaa FSK,[20] should be able to incorporate PFK. When the $tRNA^{Pyl}$/FSKRS gene was co-transformed with a nanobody mNb6 gene[46] containing an amber stop codon at site 54 [mNb6(54TAG)] in *E. coli* cells, full-length mNb6 protein was produced when 1 mM of PFK was added to the growth media (Figure 6B), indicating PFK incorporation. To evaluate PFK incorporation in mammalian cells, we co-expressed the $tRNA^{Pyl}$/FSKRS gene and the EGFP(182TAG) gene in HEK-293T cells. Western blot analysis of the cell lysates showed that full-length EGFP was produced only when PFK was added to cell culture (Figure 6C). We further transfected the $tRNA^{Pyl}$/FSKRS gene into the HeLa-GFP(182TAG) reporter cell line that contains the genome-integrated GFP(182TAG) gene.[47] Flow cytometric analysis of these cells showed that the total fluorescence intensity in cells increased 380 fold when 1 mM PFK was added (Figure S13). Based on Western blot and flow cytometric analysis, the fidelity of PFK incorporation in mammalian cells exceeded

99.7%. These results indicate that the $tRNA^{Pyl}$/FSKRS pair was able to incorporate PFK into proteins in both *E. coli* and mammalian cells. We renamed FSKRS to PFKRS for clarity.

PFK expands protein cross-linking unreachable by PFY *in vitro* and in cells

To confirm that PFK could react with target residue unreachable by PFY, we first incorporated PFK into nanobody mNb6 to evaluate the *in vitro* cross-linking of mNb6 with its binding target – the Spike protein of SARS-CoV-2. Based on the structure of mNb6-Spike complex (Figure 6D),[46] we decided to incorporate PFK into mNb6 at sites 50–59 individually to target Tyr351 of the Spike protein's receptor binding domain (RBD). PFK-incorporated mNb6 mutant proteins were purified and incubated with the Spike RBD followed with Western blot analysis. Robust cross-linking of mNb6 with the Spike RBD was detected when PFK was incorporated at site 54 but no other sites (Figure 6E). Tandem MS analysis of the cross-linked proteins confirmed that PFK reacted with the target Tyr351 as expected (Figure 6F). This site-specific cross-linking indicated that PFK-based PFEx reaction was also proximity-driven and non-random. In contrast, when the shorter PFY was incorporated at site 54, no cross-linking of mNb6 with the Spike RBD was detected (Figure 6G).

We next incorporated PFK into ecGST expressed in mammalian cells to verify its reactivity and compare with PFY in cells. At the dimer interface of ecGST, we incorporated PFK at site 103 to target Cys10 of the other monomer (Figure 6H).[45] The Cα-Cα distance between site 103 and site 10 is 4.2 Å longer than that between site 103 and site 106, which was used to test PFY cross-linking earlier. Nearby His106, Lys107 and Tyr157 were mutated to Ala so that Cys10 was the only target residue in proximity for PFK103. After expressing this mutant ecGST(103PFK) in HEK-293T cells, we analyzed the cell lysate via Western blot, revealing substantial dimeric ecGST cross-linking with a dimer to monomer (D/M) ratio of 29.6% (Figure 6I). Substituting Cys10 with His resulted in strong dimeric ecGST cross-linking, evident from a D/M ratio of 52.4%. Conversely, mutating Cys10 to Tyr led to a modest D/M ratio of 4.1% for dimeric cross-linking, and no dimeric cross-linking was observed when Cys10 was replaced with Ala in the negative control. These findings confirmed that PFK was able to react with Cys, His, and Tyr when placed in proximity in mammalian cells. Furthermore, substituting PFK with PFY at site 103 yielded varied results: no ecGST dimeric cross-linking for Cys, considerably weaker dimeric cross-linking for His (D/M ratio 11.6%), and notably stronger cross-linking for Tyr (D/M ratio of 13.6%) (Figure 6I, 6J). This variation in cross-linking efficiency can be attributed to the difference in side chain lengths of PFK and PFY. Within the fixed distance between sites 103 and 10, the longer PFK could reach the shorter side chains of His and Cys, whereas the shorter PFY was better suited to the longer side chain of Tyr. Taken together, these results indicate that PFK, due to its long and flexible side chain, was able to react with protein residues that are unreachable by PFY. Consequently, PFY and PFK act complementarily, targeting residues at varying distances in proteins.

DISCUSSION

We investigated whether PFEx could be introduced in proteins and function in biocompatible conditions to generate new covalent linkages in proteins. Through designing, synthesizing, and genetic encoding latent bioreactive amino acids PFY and PFK that contain phosphoramidofluoridate, we succeeded in introducing the new click PFEx reaction into proteins. We demonstrated that PFY or PFK, once incorporated, could target nearby His, Tyr, Lys, or Cys residues in proteins through proximity-enabled PFEx reactivity, enabling precise covalent cross-linking of interacting proteins both *in vitro* and in cells, without needing external reagents.

PFEx reaction has recently been developed for small molecules only,[4] and our study has expanded this new click reaction into proteins for the first time. In small molecule context, PFEx showed chemoselectivity for alcohols and amines, especially phenols.[4] Consistently, in proteins, we observed robust cross-linking with Tyr (phenol side chain) and His (imidazole side chain) using PFY and PFK. Lys, with its primary amine side chain, also showed reliable but weaker cross-linking. PFEx also tolerated various functional groups in small molecules including amides,[4] a feature we confirmed in the protein context as PFY/PFK did not react with the protein backbone or other tested amino acids.

Our findings reveal intriguing aspects of PEFx click chemistry in a protein context. The recent report on small molecule PFEx reactions describes the use of the polar aprotic organic solvent acetonitrile, typically requiring a strongly basic catalyst TBD and a silicon additive, hexamethyldisilazane, to accelerate the reaction.[4] However, both acetonitrile and TBD can denature proteins,[48] and hexamethyldisilazane, being water-insoluble, is not suitable for aqueous protein solutions. Our results demonstrate that PFEx in proteins is feasible in protic aqueous solvents and even within the complex cellular environments, paving the way for its potential in biological applications. Moreover, the PFEx reaction in proteins proceeded without an external catalyst, suggesting that mere proximity of reactants was sufficient for activation, similar to our previously reported SuFEx reaction in proteins.[34] Interestingly, a water-soluble silicon reagent, Na_2SiO_3, boosted the PFEx reaction between PFY and Cys/Tyr in proteins, but reduced it between PFY and His. This differential effect merits further investigation. Additionally, Cys reliably cross-linked with PFY, particularly with Na_2SiO_3 assistance, indicating that thiol groups, previously unmentioned in small molecule PFEx studies,[4] could be viable targets for PFEx in small molecules as well. Interestingly, among the ten consecutive sites of nanobody mNb6, only site 54 facilitated the cross-linking of PFK with the Spike protein (Figure 6E). Apart from protein binding and side chain interactions, it is possible that specific environmental factors at site 54 in the binding interface – whether involving electrostatic or electric field effects – contributed to the promotion of PFK reactivity. This possibility is akin to the “Sleeping Beauty” effect, a unique trait of folded proteins originally observed by Sharpless *et al.* in their SuFEx studies with small molecules and proteins.[49,50]

PFEx in proteins offers more versatility and functionality compared to the existing SuFEx. It introduces P-N/O linkages, preferred in native biomacromolecules, as opposed to the S-N/O linkages characteristic of SuFEx. While SuFEx targets Lys/His/Tyr, PFEx expands

this range to include Cys, enhancing its versatility as a proximity-enabled chemistry. PFEx also enables thermally reversible P-N linkages, a feature not present in SuFEx. At low temperatures, efficient His cross-linking with PFY was observed, which was reversible at higher temperatures. This thermal sensitivity provides reversible linkage options, potentially opening innovative applications. Unlike SuFEx and other nucleophilic reactions that are more reactive at basic pH, PFEx with His showed higher reactivity at acidic pHs. This unique pH behavior could offer a novel approach for targeting acidic microenvironments, such as the tumor microenvironment, with high selectivity. Lastly, PFY demonstrated enhanced durability in proteins compared to FSY, its SuFEx counterpart, a property beneficial for in vivo applications.

In conclusion, introducing PFEx in proteins is a significant advancement for biological and biomedical research. Site-specific incorporation of PFY or PFK in proteins enabled us to generate defined covalent crosslinks between interacting proteins, both *in vitro* and in living cells. This extends the use of phosphorus linkages in proteins, paralleling nature's preference for phosphate connectors in various biomolecules.[51] The remarkable chemoselectivity and unique properties of PFEx, compared to established click reactions,[4] will significantly expand our capabilities in covalent protein engineering, a field with vast potential in chemical biology and therapeutics.[37,11] Analogous to the transformative effects of CuAAC and SuFEx click chemistry in protein studies,[3,9,10] the introduction of PFEx into proteins is poised to open new pathways in biological research, synthetic biology, and biotherapeutics.

EXPERIMENTAL PROCEDURES

Resource availability

Lead contact—Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Lei Wang (lei.wang2@ucsf.edu).

Materials availability—Unique reagents generated in this study are available from the lead contact with a completed material transfer agreement.

Data and code availability—All data supporting this study are available within the main article and supplemental information. This study did not generate any datasets or original code.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

L.W. acknowledges support from the National Institutes of Health (NIH) (R01GM118384 and R01CA258300).

REFERENCES

1. Kolb HC, Finn MG, and Sharpless KB (2001). Click Chemistry: Diverse Chemical Function from a Few Good Reactions. Angew. Chem. Int. Ed Engl 40, 2004–2021. 10.1002/1521-3773(20010601)40:11<2004::AID-ANIE2004>3.0.CO;2-5. [PubMed: 11433435]

2. Dong J, Krasnova L, Finn MG, and Sharpless KB (2014). Sulfur(VI) fluoride exchange (SuFEx): another good reaction for click chemistry. Angew. Chem. Int. Ed Engl 53, 9430–9448. 10.1002/anie.201309399. [PubMed: 25112519]

3. Devaraj NK, and Finn MG (2021). Introduction: Click Chemistry. Chem. Rev 121, 6697–6698. 10.1021/acs.chemrev.1c00469. [PubMed: 34157843]

4. Sun S, Homer JA, Smedley CJ, Cheng Q-Q, Sharpless KB, and Moses JE (2023). Phosphorus fluoride exchange: Multidimensional catalytic click chemistry from phosphorus connective hubs. Chem 9, 2128–2143. 10.1016/j.chempr.2023.05.013. [PubMed: 38882554]

5. Tuin AW, Mol MAE, Van Den Berg RM, Fidder A, Van Der Marel GA, Overkleeft HS, and Noort D (2009). Activity-Based Protein Profiling Reveals Broad Reactivity of the Nerve Agent Sarin. Chem. Res. Toxicol 22, 683–689. 10.1021/tx8004218. [PubMed: 19226147]

6. Liu Y, Patricelli MP, and Cravatt BF (1999). Activity-based protein profiling: the serine hydrolases. Proc. Natl. Acad. Sci. U. S. A 96, 14694–14699. 10.1073/pnas.96.26.14694. [PubMed: 10611275]

7. Higson AP, Ferguson MAJ, and Nikolaev AV (1999). Synthesis of 6-N-Biotinylaminohexyl Isopropyl Phosphorofluoridate: A Potent Tool for the Inhibition/Isolation of Serine Esterases and Proteases. Synthesis 1999, 407–409. 10.1055/s-1999-3419.

8. Aminoff D, Bochar DA, Fuller AA, Mapp AK, Showalter HDH, and Kirchhoff PD (2009). Research into selective biomarkers of erythrocyte exposure to organophosphorus compounds. Anal. Biochem 392, 155–161. 10.1016/j.ab.2009.05.049. [PubMed: 19497294]

9. Barrow AS, Smedley CJ, Zheng Q, Li S, Dong J, and Moses JE (2019). The growing applications of SuFEx click chemistry. Chem. Soc. Rev 10.1039/c8cs00960k.

10. Wu P (2022). The Nobel Prize in Chemistry 2022: Fulfilling Demanding Applications with Simple Reactions. ACS Chem. Biol 17, 2959–2961. 10.1021/acschembio.2c00788. [PubMed: 36279258]

11. Cao L, and Wang L (2022). New covalent bonding ability for proteins. Protein Sci. 31, 312–322. 10.1002/pro.4228. [PubMed: 34761448]

12. Wang L, Brock A, Herberich B, and Schultz PG (2001). Expanding the Genetic Code of *Escherichia coli*. Science 292, 498–500. 10.1126/science.1060077. [PubMed: 11313494]

13. Xiang Z, Ren H, Hu YS, Coin I, Wei J, Cang H, and Wang L (2013). Adding an unnatural covalent bond to proteins through proximity-enhanced bioreactivity. Nat. Methods 10, 885–888. 10.1038/nmeth.2595. [PubMed: 23913257]

14. Li S, Wang N, Yu B, Sun W, and Wang L (2023). Genetically encoded chemical crosslinking of carbohydrate. Nat. Chem 15, 33–42. 10.1038/s41557-022-01059-z. [PubMed: 36216893]

15. Sun W, Wang N, Liu H, Yu B, Jin L, Ren X, Shen Y, and Wang L (2023). Genetically encoded chemical crosslinking of RNA in vivo. Nat. Chem 15, 21–32. 10.1038/s41557-022-01038-4. [PubMed: 36202986]

16. Li JC, Nastertorabi F, Xuan W, Han GW, Stevens RC, and Schultz PG (2019). A Single Reactive Noncanonical Amino Acid Is Able to Dramatically Stabilize Protein Structure. ACS Chem. Biol 14, 1150–1153. 10.1021/acschembio.9b00002. [PubMed: 31181898]

17. Hoppmann C, Maslennikov I, Choe S, and Wang L (2015). In Situ Formation of an Azo Bridge on Proteins Controllable by Visible Light. J. Am. Chem. Soc 137, 11218–11221. 10.1021/jacs.5b06234. [PubMed: 26301538]

18. Shu X, Liao Q-Q, Li S-T, Liu L, Zhang X, Zhou L, Zhang L, Coin I, Wang L, Wu H, et al. (2023). Detecting Active Deconjugating Enzymes with Genetically Encoded Activity-Based Ubiquitin and Ubiquitin-like Protein Probes. Anal. Chem 95, 846–853. 10.1021/acs.analchem.2c03270. [PubMed: 36595388]

19. Yang B, Tang S, Ma C, Li S-T, Shao G-C, Dang B, DeGrado WF, Dong M-Q, Wang PG, Ding S, et al. (2017). Spontaneous and specific chemical cross-linking in live cells to capture and identify protein interactions. Nat. Commun 8, 2240. 10.1038/s41467-017-02409-z. [PubMed: 29269770]

20. Liu J, Cao L, Klauser PC, Cheng R, Berdan VY, Sun W, Wang N, Ghelichkhani F, Yu B, Rozovsky S, et al. (2021). A Genetically Encoded Fluorosulfonyloxybenzoyl- L - lysine for Expansive Covalent Bonding of Proteins via SuFEx Chemistry. J. Am. Chem. Soc 143, 10341–10351. 10.1021/jacs.1c04259. [PubMed: 34213894]

21. Tang H, Dai Z, Qin X, Cai W, Hu L, Huang Y, Cao W, Yang F, Wang C, and Liu T (2018). Proteomic Identification of Protein Tyrosine Phosphatase and Substrate Interactions in Living

Mammalian Cells by Genetic Encoding of Irreversible Enzyme Inhibitors. J. Am. Chem. Soc 140, 13253–13259. 10.1021/jacs.8b06922. [PubMed: 30247891]

22. Liu C, Wu T, Shu X, Li S, Wang DR, Wang N, Zhou R, Yang H, Jiang H, Hendriks IA, et al. (2021). Identification of Protein Direct Interactome with Genetic Code Expansion and Search Engine OpenUaa. Adv. Biol 5, 2000308. 10.1002/adbi.202000308.
23. Klauser PC, Chopra S, Cao L, Bobba KN, Yu B, Seo Y, Chan E, Flavell RR, Evans MJ, and Wang L (2023). Covalent Proteins as Targeted Radionuclide Therapies Enhance Antitumor Effects. ACS Cent. Sci 9, 1241–1251. 10.1021/acscentsci.3c00288. [PubMed: 37396859]
24. Li Q, Chen Q, Klauser PC, Li M, Zheng F, Wang N, Li X, Zhang Q, Fu X, Wang Q, et al. (2020). Developing Covalent Protein Drugs via Proximity-Enabled Reactive Therapeutics. Cell 182, 85–97.e16. 10.1016/j.cell.2020.05.028. [PubMed: 32579975]
25. Yu B, Li S, Tabata T, Wang N, Cao L, Kumar GR, Sun W, Liu J, Ott M, and Wang L (2022). Accelerating PERx reaction enables covalent nanobodies for potent neutralization of SARS-CoV-2 and variants. Chem 8, 2766–2783. 10.1016/j.chempr.2022.07.012. [PubMed: 35874165]
26. Cao YJ, Yu C, Wu K-L, Wang X, Liu D, Tian Z, Zhao L, Qi X, Loredo A, Chung A, et al. (2021). Synthesis of precision antibody conjugates using proximity-induced chemistry. Theranostics 11, 9107–9117. 10.7150/thno.62444. [PubMed: 34522229]
27. Liu J, Li S, Aslam NA, Zheng F, Yang B, Cheng R, Wang N, Rozovsky S, Wang PG, Wang Q, et al. (2019). Genetically Encoding Photocaged Quinone Methide to Multitarget Protein Residues Covalently in Vivo. J. Am. Chem. Soc 141, 9458–9462. 10.1021/jacs.9b01738. [PubMed: 31184146]
28. Yu B, Cao L, Li S, Klauser P, and Wang L (2023). Proximity-Enabled Sulfur Fluoride Exchange Reaction in Protein Context. Chem. Sci, 10.1039.D3SC01921G. 10.1039/D3SC01921G.
29. Chen X-H, Xiang Z, Hu YS, Lacey VK, Cang H, and Wang L (2014). Genetically Encoding an Electrophilic Amino Acid for Protein Stapling and Covalent Binding to Native Receptors. ACS Chem. Biol 9, 1956–1961. 10.1021/cb500453a. [PubMed: 25010185]
30. Hoppmann C, and Wang L (2016). Proximity-enabled bioreactivity to generate covalent peptide inhibitors of p53–Mdm4. Chem. Commun 52, 5140–5143. 10.1039/C6CC01226D.
31. Furman JL, Kang M, Choi S, Cao Y, Wold ED, Sun SB, Smider VV, Schultz PG, and Kim CH (2014). A genetically encoded aza-Michael acceptor for covalent cross-linking of protein-receptor complexes. J. Am. Chem. Soc 136, 8411–8417. 10.1021/ja502851h. [PubMed: 24846839]
32. Xuan W, Li J, Luo X, and Schultz PG (2016). Genetic Incorporation of a Reactive Isothiocyanate Group into Proteins. Angew. Chem 128, 10219–10222. 10.1002/ange.201604891.
33. Xuan W, Shao S, and Schultz PG (2017). Protein Crosslinking by Genetically Encoded Noncanonical Amino Acids with Reactive Aryl Carbamate Side Chains. Angew. Chem. Int. Ed Engl 56, 5096–5100. 10.1002/anie.201611841. [PubMed: 28371162]
34. Wang N, Yang B, Fu C, Zhu H, Zheng F, Kobayashi T, Liu J, Li S, Ma C, Wang PG, et al. (2018). Genetically Encoding Fluorosulfate-L-tyrosine to React with Lysine, Histidine, and Tyrosine via SuFEx in Proteins in vivo. J. Am. Chem. Soc 140, 4995–4999. 10.1021/jacs.8b01087. [PubMed: 29601199]
35. Xu Y, Rahim A, and Lin Q (2022). Spontaneous Orthogonal Protein Crosslinking via a Genetically Encoded 2-Carboxy-4-Aryl-1,2,3-Triazole. Angew. Chem. Int. Ed 61, e202202657. 10.1002/anie.202202657.
36. Xiang Z, Lacey VK, Ren H, Xu J, Burban DJ, Jennings PA, and Wang L (2014). Proximity-Enabled Protein Crosslinking through Genetically Encoding Haloalkane Unnatural Amino Acids. Angew. Chem. Int. Ed 53, 2190–2193. 10.1002/anie.201308794.
37. Wang N, and Wang L (2022). Genetically encoding latent bioreactive amino acids and the development of covalent protein drugs. Curr. Opin. Chem. Biol 66, 102106. 10.1016/j.cbpa.2021.102106. [PubMed: 34968810]
38. Klauser PC, Berdan VY, Cao L, and Wang L (2022). Encoding latent SuFEx reactive meta-fluorosulfate tyrosine to expand covalent bonding of proteins. Chem. Commun 58, 6861–6864. 10.1039/D2CC01902G.
39. Yang B, Wang N, Schnier PD, Zheng F, Zhu H, Polizzi NF, Ittuveetil A, Saikam V, DeGrado WF, Wang Q, et al. (2019). Genetically Introducing Biochemically Reactive Amino Acids

Dehydroalanine and Dehydrobutyrine in Proteins. J. Am. Chem. Soc 141, 7698–7703. 10.1021/jacs.9b02611. [PubMed: 31038942]

40. Mukherjee H, Debreczeni J, Breed J, Tentarelli S, Aquila B, Dowling JE, Whitty A, and Grimster NP (2017). A study of the reactivity of S(VI)-F containing warheads with nucleophilic amino-acid side chains under physiological conditions. Org. Biomol. Chem 15, 9685–9695. 10.1039/c7ob02028g. [PubMed: 29119993]
41. Smaliy RV, Chaikovskaya AA, and Pinchuk AM (2006). Reactions of isocyanatophosphoryl difluoride with π-abundant nitrogen heterocycles and carbonyl compounds. Russ. Chem. Bull 55, 585–587. 10.1007/s11172-006-0297-9.
42. Hoppmann C, Wong A, Yang B, Li S, Hunter T, Shokat KM, and Wang L (2017). Site-specific incorporation of phosphotyrosine using an expanded genetic code. Nat. Chem. Biol 13, 842–844. 10.1038/nchembio.2406. [PubMed: 28604697]
43. Högbom M, Eklund M, Nygren P-Å, and Nordlund P (2003). Structural basis for recognition by an *in vitro* evolved affibody. Proc. Natl. Acad. Sci 100, 3191–3196. 10.1073/pnas.0436100100. [PubMed: 12604795]
44. Li T, Cai H, Yao H, Zhou B, Zhang N, van Vlissingen MF, Kuiken T, Han W, GeurtsvanKessel CH, Gong Y, et al. (2021). A synthetic nanobody targeting RBD protects hamsters from SARS-CoV-2 infection. Nat. Commun 12, 4635. 10.1038/s41467-021-24905-z. [PubMed: 34330908]
45. Nishida M, Harada S, Noguchi S, Satow Y, Inoue H, and Takahashi K (1998). Three-dimensional structure of Escherichia coli glutathione S-transferase complexed with glutathione sulfonate: catalytic roles of Cys10 and His106. J. Mol. Biol 281, 135–147. 10.1006/jmbi.1998.1927. [PubMed: 9680481]
46. Schoof M, Faust B, Saunders RA, Sangwan S, Rezelj V, Hoppe N, Boone M, Billesbølle CB, Puchades C, Azumaya CM, et al. (2020). An ultrapotent synthetic nanobody neutralizes SARS-CoV-2 by stabilizing inactive Spike. Science 370, 1473–1479. 10.1126/science.abe3255. [PubMed: 33154106]
47. Wang W, Takimoto JK, Louie GV, Baiga TJ, Noel JP, Lee K-F, Slesinger PA, and Wang L (2007). Genetically encoding unnatural amino acids for cellular and neuronal studies. Nat. Neurosci 10, 1063–1072. 10.1038/nn1932. [PubMed: 17603477]
48. Gekko K, Ohmae E, Kameyama K, and Takagi T (1998). Acetonitrile-protein interactions: amino acid solubility and preferential solvation. Biochim. Biophys. Acta BBA - Protein Struct. Mol. Enzymol 1387, 195–205. 10.1016/S0167-4838(98)00121-6.
49. Cheng Y, Li G, Smedley CJ, Giel M-C, Kitamura S, Woehl JL, Bianco G, Forli S, Homer JA, Cappiello JR, et al. (2022). Diversity oriented clicking delivers β-substituted alkenyl sulfonyl fluorides as covalent human neutrophil elastase inhibitors. Proc. Natl. Acad. Sci 119, e2208540119. 10.1073/pnas.2208540119. [PubMed: 36070343]
50. Zheng Q, Woehl JL, Kitamura S, Santos-Martins D, Smedley CJ, Li G, Forli S, Moses JE, Wolan DW, and Sharpless KB (2019). SuFEx-enabled, agnostic discovery of covalent inhibitors of human neutrophil elastase. Proc. Natl. Acad. Sci 116, 18808–18814. 10.1073/pnas.1909972116. [PubMed: 31484779]
51. Westheimer FH (1987). Why Nature Chose Phosphates. Science 235, 1173–1178. 10.1126/science.2434996. [PubMed: 2434996]

Bigger Picture

Click chemistry has transformed the synthesis of small molecules and the manipulation of large biomolecules, significantly impacting both chemistry and biology. At the forefront of this innovation is Phosphorus Fluoride Exchange (PFEx), recently developed for in vitro use with small molecules. We have now expanded PFEx's potential into biological applications by introducing two novel amino acids embedded with phosphoramidofluoridates into proteins through genetic code expansion. These amino acids enable the formation of precise covalent linkages in proteins via the PFEx reaction, exhibiting unique properties and proving effective both in vitro and in live cells. This advancement not only facilitates covalent protein engineering using nature's preferred phosphate connectors but also lays the groundwork for PFEx's wider application in biological and biomedical research, which can deepen our understanding of biological processes and open up new therapeutic possibilities.

Highlights

- Bioreactive Uaas PFY and PFK are genetically encoded in E. coli and mammalian cells
- PFY and PFK react with His, Tyr, Lys, and Cys via proximity-enabled PFEx reaction
- PFEx reaction with His results in unique pH and thermal properties
- Na_2SiO_3 enhances PFEx reactions with Tyr and Cys

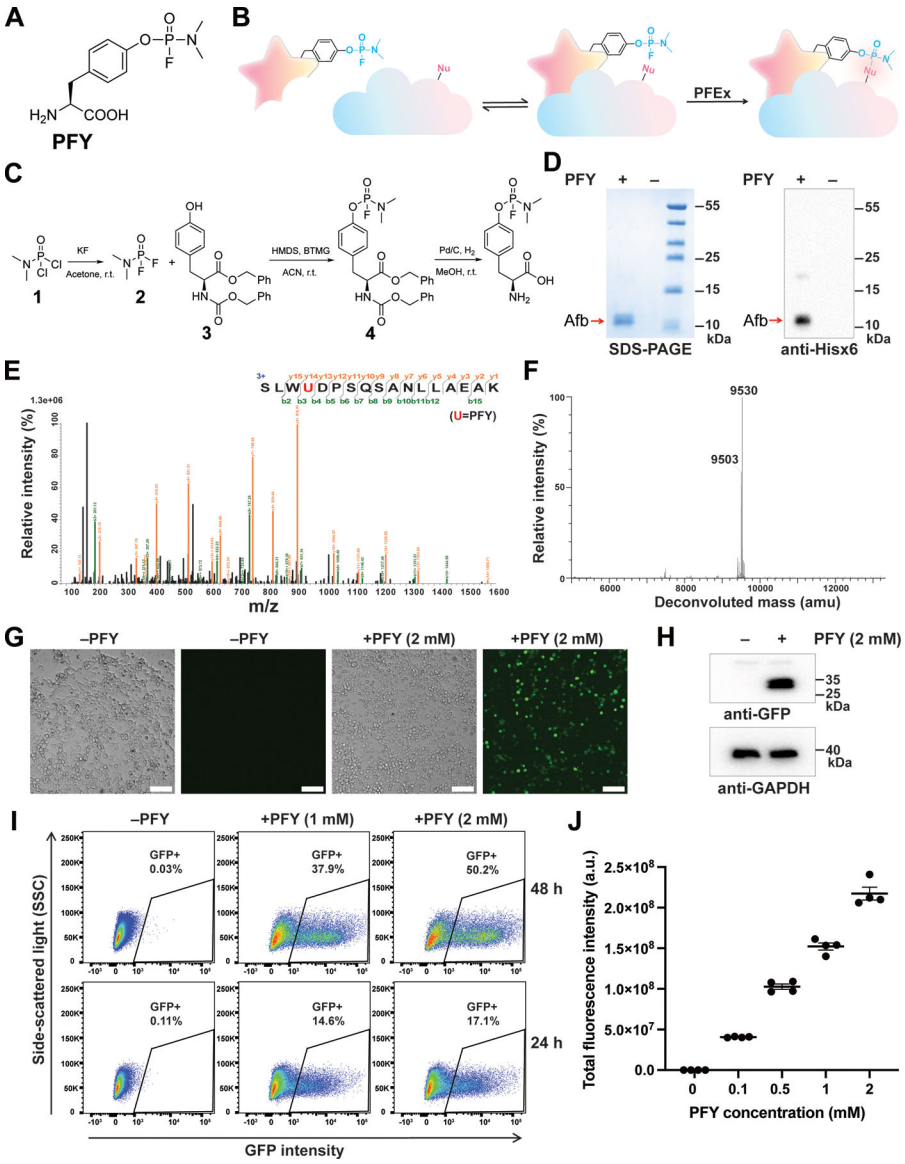

Figure 1. Design, synthesis, and genetic incorporation of PFY into proteins in *E. coli* and mammalian cells.

(A) Structure of PFY.

(B) PFY to react with nucleophilic residues in close proximity via proximity-enabled PFEx reactivity.

(C) Chemical synthesis of PFY.

(D) SDS-PAGE and Western blot analysis of PFY incorporation into Afb(36TAG) by the $tRNA^{Pyl}$/PFYRS in *E. coli*.

(E) Tandem MS spectrum of the Afb(36PFY) tryptic peptide containing residue 36 verifying PFY incorporation at site 36.

(F) ESI-TOF-MS spectrum of intact Ub(51PFY) protein verifying PFY incorporation. The peak at 9530 Da corresponds to the incorporation of intact PFY at site 51; the peak at 9503 Da corresponds to PFY incorporation at site 51 but with its dimethylamino group hydrolyzed.

(G) Bright field and fluorescence microscopic images of HEK-293T cells. The cells were transfected with the $tRNA^{Pyl}$/PFYRS and the EGFP(182TAG) gene, and grown in the absence and presence of PFY. Scale bar, 100 μm.

(H) Western blot analysis of HEK-293T cells in (G).

(I) Flow cytometric analysis of HEK-293T cells, which were transfected with the $tRNA^{Pyl}$/PFYRS and the EGFP(182TAG) gene and grown in the presence of various concentrations of PFY for 24 h and 48 h.

(J) Flow cytometric quantification of the total fluorescence intensity of the HEK-239T cells in (I) incubated with PFY for 48 h.

See also Figures S1–S4.

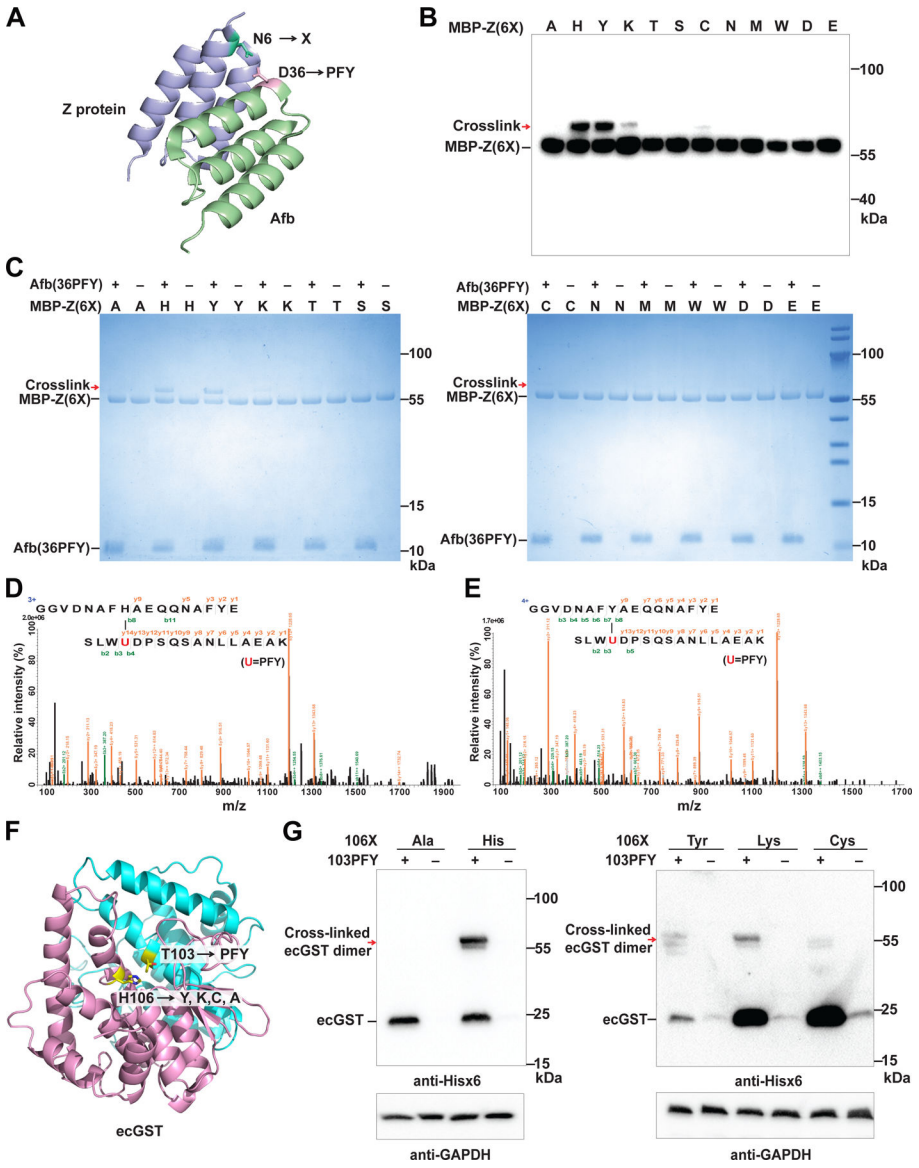

Figure 2. PFY reacts with proximal His, Tyr, Lys, and Cys in proteins through PFEx.

(A) Structure of the Afb-Z complex (PDB code: 1LP1) showing D36 in Afb for PFY incorporation and N6 in Z protein for mutation to different residues.

(B) Western blot analysis of Afb(36PFY) protein incubation with Z(6X) protein. X represents mutated residues.

(C) SDS-PAGE analysis of Afb(36PFY) protein incubation with Z(6X) protein.

(D) Tandem MS spectrum of Afb(36PFY) cross-linking with Z(6His).

(E) Tandem MS spectrum of Afb(36PFY) cross-linking with Z(6Tyr).

(F) Structure of *ec*GST (PDB code: 1A0F) showing residue T103 in one monomer for PFY incorporation and the proximal residue H106 in the other monomer for mutations.

(G) Western blot analysis of *ec*GST dimeric cross-linking in HEK-293T cells. *ec*GST(103PFY/106X) was expressed in HEK-293T cells and the cell lysate was probed with anti-Hisx6 antibody to detect the Hisx6 tag appended at the C-terminus of *ec*GST.

See also Figures S5–S6.

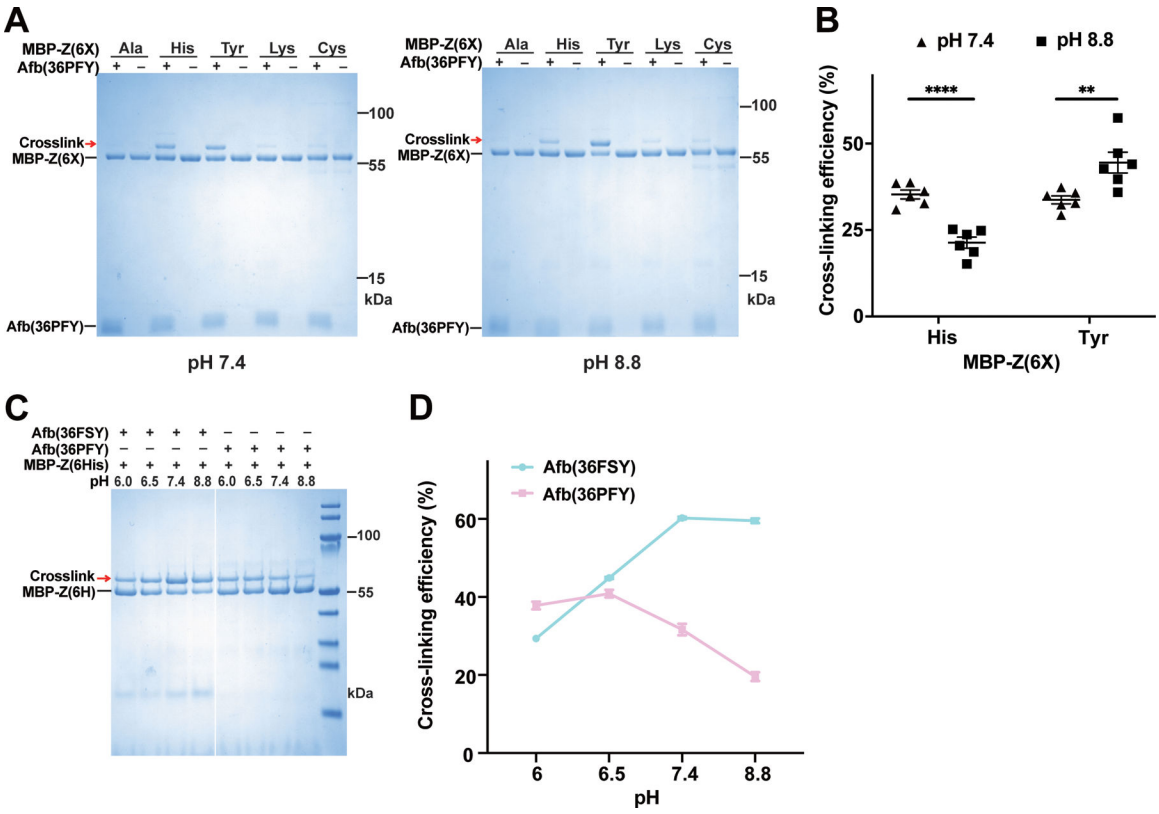

Figure 3. Divergent pH effects on PFY-His cross-linking compared to PFY-Tyr and FSY-His.

(A) SDS-PAGE analysis of Afb(36PFY) cross-linking with MBP-Z(6X) at pH 7.4 and pH 8.8.

(B) Cross-linking efficiency between Afb(36PFY) and MBP-Z(6His) or MBP-Z(6Tyr) at pH 7.4 and pH 8.8. The line and error bar represent mean ± SEM; n = 6 independent experiments. **p < 0.01; ****p < 0.001; unpaired *t*-test.

(C) SDS-PAGE analysis Afb(36PFY) or Afb(36FSY) cross-linking with MBP-Z(6His) at various pH.

(D) Changes in cross-linking efficiency with varying pH for reaction of Afb(36PFY) or Afb(36FSY) with MBP-Z(6His). The line and error bar represent mean ± SEM; n = 5 independent experiments.

See also Figures S7–S8.

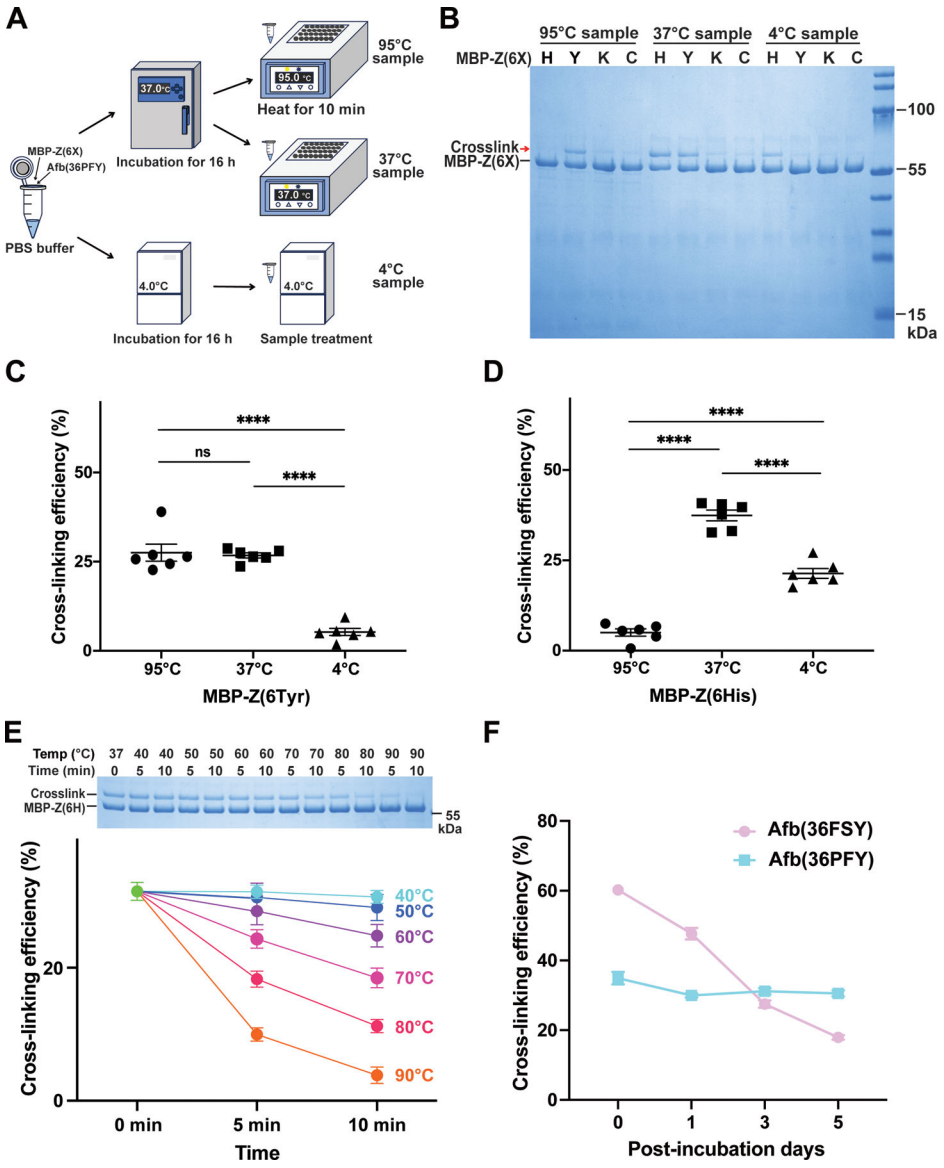

Figure 4. Temperature effects on PFY-generated linkages and PFY durability in proteins.

(A) Scheme showing how Afb(36PFY) was incubated with MBP-Z(6X) and then treated at different temperature.

(B) SDS-PAGE analysis of Afb(36PFY) cross-linking with MBP-Z(6X) under different incubation and treatment temperatures.

(C) Cross-linking efficiency of Afb(36PFY) with MBP-Z(6Tyr) at different temperatures measured from (B) using densitometry.

(D) Cross-linking efficiency of Afb(36PFY) with MBP-Z(6His) at different temperatures measured from (B) using densitometry. For C-D, the line and error bar represent mean ± SEM; n = 6 independent experiments. ns, not significant; ****$p < 0.0001$; unpaired *t*-test.

(E) Thermal sensitivity of the P(V)-N linkage resulting from PFY reaction with His. Top: SDS-PAGE analysis of the Afb(36PFY)/MBP-Z(6His) cross-linking product after it was subjected to various temperatures for durations of 5 and 10 minutes. Bottom: The cross-

linking efficiency of Afb(36PFY) with MBP-Z(6His) was evaluated after their cross-linked product was exposed to various temperatures for 5 and 10 minutes. The line and error bar represent mean ± SEM; n = 3 independent experiments.

(F) PFY exhibited enhanced durability compared to FSY in proteins. The figure shows the remaining reactivity of Afb(36PFY) and Afb(36FSY) after incubation at 37 °C for the indicated number of days. The cross-linking efficiency of Afb(36PFY) and Afb(36FSY) with MBP-Z(6His) was measured to quantify their remaining reactivity. The line and error bars represent mean ± SEM; n = 8 independent experiments.

See also Figure S9.

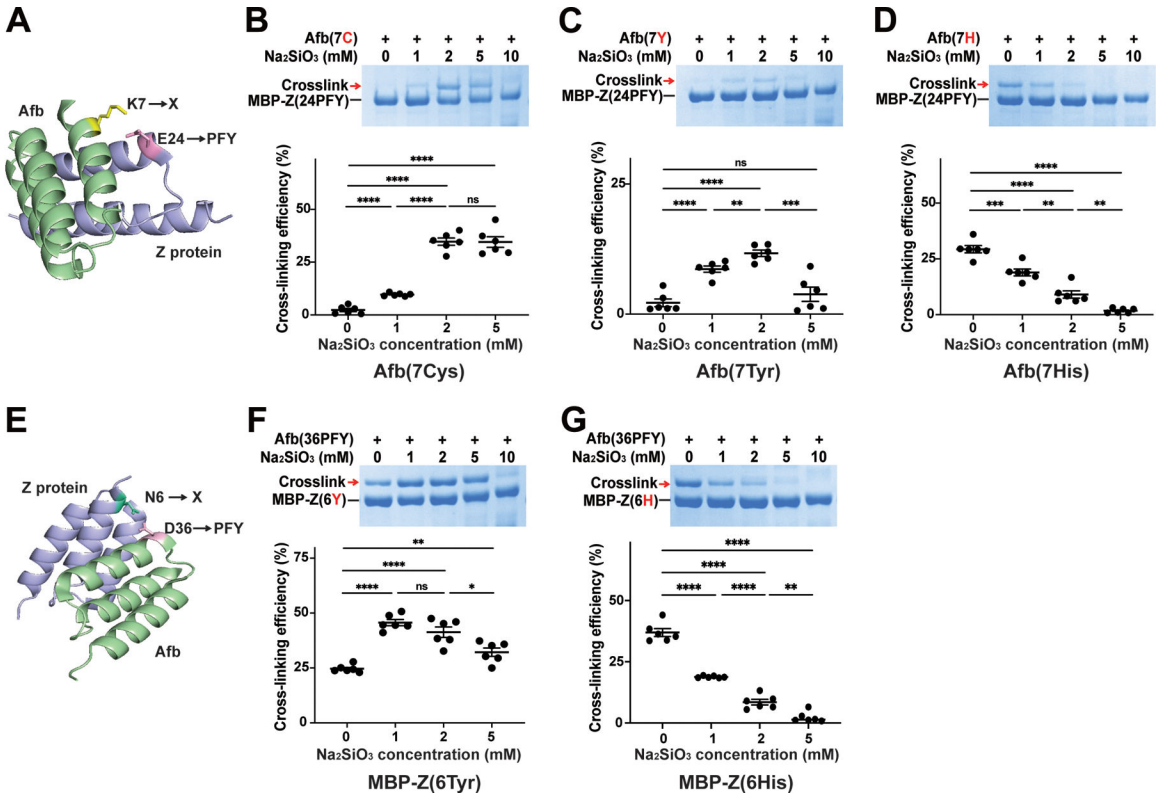

Figure 5. Na_2SiO_3 increases PFY reaction with Tyr and Cys but decreases PFY reaction with His.

(A) Structure of the Afb-Z complex (PDB code: 1LP1) showing E24 in Z protein for PFY incorporation and K7 in Afb for mutation to Cys, Tyr, or His.

(B-D) SDS-PAGE analysis (top panel) and quantification of MBP-Z(24PFY) cross-linking with Afb(7X) (bottom panel) in the presence of different concentration of Na_2SiO_3. (B) Afb(7Cys); (C) Afb(7Tyr); (D) Afb(7His).

(E) Structure of the Afb-Z complex (PDB code: 1LP1) showing D36 in Afb for PFY incorporation and N6 in Z protein for mutation to Tyr, or His.

(F-G) SDS-PAGE analysis (top panel) and quantification of Afb(36PFY) cross-linking with MBP-Z(6X) (bottom panel) in the presence of different concentration of Na_2SiO_3. (F) MPB-Z(6Tyr); (G) MPB-Z(6His).

The line and error bar represent mean ± SEM; n = 6 independent experiments. ns, not significant; *p < 0.05; **p < 0.01; ***p < 0.001; ****p < 0.0001; unpaired *t*-test.

See also Figures S10–S12.

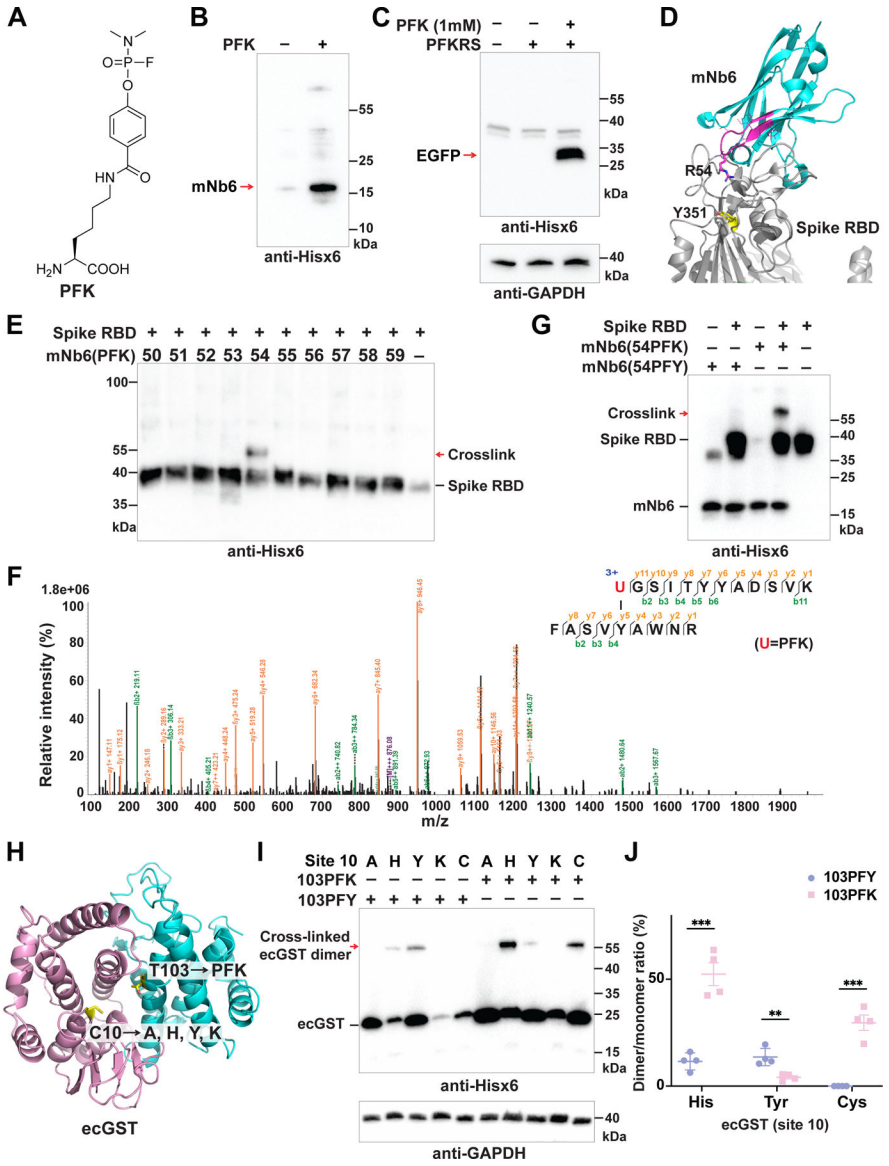

Figure 6. Genetic incorporation of PFK expands protein cross-linking unreachable by PFY *in vitro* and in cells.

(A) Structure of PFK.

(B) Western blot analysis of PFK incorporation in mNb6(54TAG) in *E. coli* cells.

(C) Western blot analysis of PFK incorporation in EGFP(182TAG) in HEK-293T cells.

(D) Structure of nanobody mNb6 binding with the Spike protein of SARS-CoV-2 (PDB code 7KKL). Sites 50–59 for PFK incorporation are colored in purple. R54 in mNb6 and the target residue Y351 in the Spike are shown in stick.

(E) Western blot analysis of cross-linking between the Spike protein's receptor binding domain (RBD) and mNb6 mutants with PFK incorporated at the indicated sites.

(F) Tandem MS spectrum of mNb6(54PFK) cross-linking with the Spike protein's RBD, confirming that PFK reacted with Y351.

(G) Western blot comparison between mNb6(54PFK) and mNb6(54PFY) for cross-linking with the Spike protein's RBD.

(H) Structure of *ec*GST (PDB code: 1A0F) showing residue T103 in one monomer for PFK/PFY incorporation and the proximal residue C10 in the other monomer for mutations.
(I) Western blot analysis of HEK-293T cell lysate of cells expressing ecGST mutants with PFK/PFY incorporated at site 103 and different mutations at site 10.
(J) Quantification of the cross-linked dimer to monomer ratio of ecGST in (I), showing the difference between PFK and PFY in cross-linking His, Tyr, and Cys. The line and error bar represent mean ± SEM; n = 4 independent experiments. **p < 0.01; ***p < 0.001; unpaired *t*-test.
See also Figure S13.