Supporting Information

Genetically Encoded Pentafluorophenylalanine Enables Quantitative Probing of Local Protein Malleability by ^{19}F NMR

Nathan Paul,[1] Adarshi P. Welegedara,[2] Rebecca L. Frkic,[2] Luca Macri,[1] Toby R. C. Thompson,[3] James L. Baber,[4] Edan Habel,[1] Elwy H. Abdelkader,[2] Haocheng Qianzhu,[1] Nicholas F. Chilton,[1,3] Colin J. Jackson,[2] Ad Bax,[4] Thomas Huber,[1] and Gottfried Otting[2*]

[1] Research School of Chemistry, The Australian National University, Canberra, ACT 2601, Australia

[2] Centre of Excellence for Innovations in Peptide and Protein Science, Research School of Chemistry, The Australian National University, Canberra, ACT 2601, Australia

[3] Department of Chemistry, The University of Manchester, Manchester M13 9PL, United Kingdom

[4] Laboratory of Chemical Physics, NIDDK, NIH, Bethesda, MD 20892-0520, United States

Table of contents:

Supplementary tables

a) Materials

The non-canonical amino acids (ncAA) 2,3,4,5,6-pentafluoro-L-phenylalanine (F_5Phe, Cat. No.: A142712) was purchased from Ambeed (USA). 4-Fluoro-L-phenylalanine (4F-Phe, Cat. No.: J53633) was purchased from AK Scientific Inc. (USA). The amino acids were used without further purification. Rapamycin (Cat. No.: A656002) was purchased from Ambeed (USA) and aprotinin was purchased from Sigma-Aldrich (Australia; Cat. No.: 10236624001). The enzymes BsaI-HFv2, BsmBI-v2, and T4 DNA ligase were purchased from New England Biolabs (NEB, USA).

b) Molecular cloning

Coding sequences were synthesized as gBlocks (Integrated DNA Technologies, USA) containing flanking BsaI restriction enzyme sites to enable Golden Gate assembly. Overhangs followed the EcoFlex Modular Cloning (MoClo) toolkit[1] nomenclature with 5′-CATA overhang (terminal A corresponding to the start codon methionine) and 3′-TCGA overlap downstream of the TAA stop codon, thereby enabling direct insertion into the pTU1 position. A pCDF-based acceptor plasmid (pCDF T7 GFP:DO) was designed with matching pTU1-compatible overhangs for compatibility with the cloning toolkit and containing a T7 promoter-lacO and transcriptional terminator. This vector incorporates a constitutively expressed green

fluorescence protein (GFP) dropout cassette, which is replaced by the gBlock insert via Golden Gate assembly. Assemblies were transformed into *E. coli* DH10B-ALT,[2] and correct colonies were identified by loss of GFP fluorescence under blue-light illumination. Plasmids were isolated using the ISOLATE II plasmid mini kit (Meridian Bioscience, USA) and verified by whole-plasmid sequencing (Plasmidsaurus, USA). Alternatively, clonal genes were synthesised from Twist Bioscience (USA) and cloned into a pCDF vector.

To compare the performance different aminoacyl-tRNA synthetases (RS), cloning into the vector pRSF followed the same strategy with modifications as for the G1 tRNA gene, which contains an internal BsaI site. To avoid restriction enzyme interference, the tRNA cassette was synthesized as a gBlock containing the proK promoter and terminator flanked by BsmBI sites and treated as a pTU1-B module. The RS coding sequence was cloned as a pTU1-A module with the same T7 promoter-lacO as the pCDF constructs. Both transcription units were subsequently assembled into a pRSF-TU2-A2 backbone by BsmBI Golden Gate assembly. The T7 promoter-lacO sequence used in these pRSF constructs differs from the version previously deposited with Addgene (USA) by featuring the lac operator (lacO) downstream of the T7 promoter to reduce basal (leaky) expression and promote more consistent cell growth.

c) *In vivo* performance of G1-, *Mb*- and *Mm*-PylRS enzymes with F_5Phe

Electrocompetent *E. coli* B95.ΔAΔ*fabR* cells[3] were co-transformed with the pRSF plasmids for the respective RS enzymes and pCDF plasmid containing an amber-interrupted RFP reporter gene (RFP-TAG; Addgene #229524). The RS enzymes evaluated comprised variants of the pyrrolysyl-tRNA synthetases (PylRS) from the methanogenic archaeon ISO4-G1 (G1 variant 09),[4] *Methanosarcina barkeri* (*Mb*; variants B5 and D6)[5] and *Methanosarcina mazei* (*Mm*; variant AS).[6] In addition, a chimeric Com2-AS variant was generated by grafting the engineered tRNA binding domain (Com2) onto the *Mm*-AS scaffold.[7]

Transformants were recovered for 1 h in LB medium and plated on LB-agar supplemented with kanamycin (25 mg/L) and spectinomycin (25 mg/L). Overnight cultures were inoculated into fresh LB medium (1:100), containing kanamycin (25 mg/L) and spectinomycin (25 mg/L). Cultures were grown to an OD_{600} of 0.6–1, at which point 1 mM F_5Phe was added. Cultures were cooled to 25 °C prior to induction with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) and expression for 16 h at 25 °C. Following expression, 50 µL of cell culture was transferred to a 384-well black wall with clear-bottom plate (Corning, USA) for OD_{600} and fluorescence measurements. RFP fluorescence was measured using excitation at 585 nm and emission at 615 nm on an Infinite 200 Pro M Plex microplate reader (Tecan,

Switzerland). Each PylRS system was tested in three biological replicates. Fluorescence values were normalised to OD_{600}. Data analysis (mean and standard deviation) was performed using GraphPad Prism (version 11).

d) Protein expression and purification

Protein expression

Genes of interest were expressed from pCDF plasmids in *E. coli* B95.ΔAΔ*fabR* cells. To produce protein with F_5Phe or 4F-Phe, the cells were co-transformed with the pCDF plasmid encoding the gene of interest with an amber (TAG) codon and a pRSF plasmid encoding the requisite tRNA-synthetase/tRNA pair. Wild-type genes without amber codon were expressed using pCDF alone.

Cultures were grown overnight at 37 °C in LB medium supplemented with spectinomycin (25 mg/L) and, where appropriate, kanamycin (25 mg/L). Overnight cultures were diluted 1:100 into fresh LB medium containing the same antibiotics and grown to an OD_{600} of 0.6–1. For amber suppression expressions, the relevant non-canonical amino acid (F_5Phe or 4F-Phe) was added at a final concentration of 1 mM prior to cooling the cultures to 25 °C. Cell pellets were harvested by centrifugation (4,000 × *g*, 4 °C, 15 min).

Protein purification

The cell pellets were resuspended in buffer A (50 mM Tris-HCl pH 7.5, 300 mM NaCl, 5% glycerol, 10 mM imidazole) and lysed by sonication (Omni-Ruptor 4000, USA; 12 min, 50% pulse, 60% power) on an ice-water slurry. Lysates were clarified by centrifugation (15,000 × *g*, 4 °C, 60 min) and applied to a 1 mL His GraviTrap column (Cytiva, USA). Columns were washed with 30 column volumes of buffer A, and proteins were eluted with 5 column volumes buffer B (buffer A containing 300 mM imidazole). Eluted proteins were buffer exchanged using PD-10 desalting columns (Cytiva. USA) and concentrated using 3 or 10 kDa molecular weight cut-off (MWCO) Amicon ultrafilters (Millipore, USA). Final buffer conditions were as follows:

RFP-TAG, PpiB, H11 and Arc-NL: 50 mM HEPES pH 7.5, 100 mM NaCl;

GB1: 20 mM MES pH 6.5, 100 mM NaCl;

FKBP-12: 50 mM MES pH 6.5, 150 mM NaCl;

MBP: 20 mM sodium phosphate buffer pH 7.2, 100 mM NaCl;

DENV protease: 50 mM HEPES pH 7.5, 100 mM NaCl.

The constructs of GB1, Arc-NL, and MBP contained TEV protease cleavage sites to remove fusion proteins or the His_6 tag. For TEV cleavage, samples were buffer exchanged into buffer C (50 mM Tris-HCl pH 8.0, 300 mM NaCl, 1 mM DTT) and incubated with His_6-TEV protease (1:10 w/w TEV:target protein) at 4 °C for 16 h. Following cleavage, samples were centrifuged (5,000 × *g*, 4 °C, 20 min) to remove precipitated protein and applied to a 1 mL His GraviTrap column (Cytiva, USA). For constructs in which the His_6 tag was removed (e.g. GB1, MBP, Arc-NL), the cleaved target proteins were collected in the flow-through.

For C-terminally His_6-tagged constructs (e.g. Arc-NL-H_6), the His_6-TEV protease co-elutes with the target protein following cleavage. To address this, samples were subjected to heat treatment (20 min at 55 °C) to selectively precipitate the TEV protease, followed by centrifugation (5,000 × *g*, 4 °C, 20 min). The clarified supernatant was then purified again by Ni^{2+} affinity purification as described above.

Some proteins were further purified by size exclusion chromatography (SEC) using a Superdex 75 increase 10/300 GL column on an ÄKTA pure 25 system (Cytiva, USA). SEC was performed using the following buffers:
NTGG-PpiB variants used for protein stability assays: 50 mM HEPES pH 7.5, 100 mM NaCl;
PpiB with F_5Phe in position 4 for protein crystallization: 50 mM imidazole pH 7.0, 1 mM EDTA, and 1 mM DTT;
Arc-NL variants for protein stability assays, biolayer inferometry, and NMR: 50 mM HEPES pH 7.5, 100 mM NaCl.

The present work used two versions of PpiB, referred to as NTGG-PpiB and PpiB. The NTGG-PpiB construct was used for all experiments except X-ray crystallography. It contains two glycine residues after the N-terminal methionine to facilitate its complete removal by host cell peptidases, thereby improving sample homogeneity for NMR analysis. In contrast, the native construct of PpiB without the additional glycine residues was used for crystallography as in previously reported X-ray structures of PpiB.[8,9]

The present work also used two versions of Arc-NL, referred to as Arc-NL and Arc-NL-H_6, where the untagged construct was used only for NMR analysis of the complex with the H11 F70/F_5Phe mutant (Figure 8c). The C-terminally His_6-tagged Arc-NL construct was used for all other experiments. Initial expression trials indicated that Arc-NL could not be robustly expressed in the absence of an N-terminal fusion partner. We obtained good yields of soluble protein using RFP as the expression tag.

For co-crystallization of Arc-NL with F_5Phe in position 267 in complex with H11, proteins were mixed in 1.2:1 molar ratio (H11:Arc-NL) and subjected to SEC purification in PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, and 1.8 mM KH_2PO_4, pH 7.4) to isolate the complex in equimolar stoichiometry.

e) Protein mass spectrometry

Intact protein mass spectrometry was performed either on an Orbitrap Elite Hybrid Ion Trap-Orbitrap mass spectrometer (Thermo Fisher Scientific) or Orbitrap Fusion Tribrid mass spectrometer (Thermo Fisher Scientific) connected to a Thermo Fisher Scientific UltiMate 3000 HPLC system equipped with ZORBAX 300SB-C3, 3.5 μm, 4.6 x 50 mm HPLC column (Agilent Technologies). Approximately 10-50 pmol of sample was injected using a 500 μL/min linear gradient of solvent A (0.1% (v/v) formic acid in water) and solvent B (0.1% (v/v) formic acid in acetonitrile), ramping solvent B from 5% solvent B at the start to 80% after 12 min. Data were collected using an electrospray ionization (ESI) source in positive ion mode. Protein intact mass was determined by deconvolution using the program Xcalibur 3.0.63 (Thermo Fisher Scientific).

f) NMR spectroscopy

All NMR spectra were recorded at 25 °C unless mentioned otherwise. The NMR samples were prepared in their respective final buffers, supplemented with 10% D_2O and 0.1 mM TFA. The ^{19}F chemical shifts were referenced to internal TFA (-75.25 ppm).[10] The probe heads of the 400 and 600 MHz NMR spectrometers featured a strong background signal of fluorinated polymer that overlapped with the ^{19}F NMR signals of F_5Phe, requiring manual baseline corrections after Fourier transformation. In some of the 1D ^{19}F NMR experiments, the background signal was attenuated by recording the spectrum with the 3-pulse sequence 90^o-δ_1-90^o-δ_2-90^o-acquisition with δ_1 = 2 μs, δ_2 = 200 μs, and the phase cycle of the NOESY experiment. The 600 MHz NMR spectrometer with high-pressure accessory was equipped with a conventional ^{19}F NMR probe, where the background signal was suppressed about three-fold by selective irradiation at -190 ppm during the recovery delay.

The flip rates of aromatic rings, *k*, were determined from the ratio of cross-peak intensity to diagonal peak intensity, *r*, detected in [^{19}F,^{19}F]-NOESY spectra recorded with sufficiently short mixing time, τ_m, to produce smaller cross-peaks than diagonal peaks (see

Figure 4c in the main text). The ratio *r* was determined by peak integration in one-dimensional slices in the δ_2 dimension. The flip rates *k* were calculated using the equation[11]

$$k = \ln(1 + r)/(1 - r)]/(2\ \tau_m) \quad (1)$$

Uncertainty ranges were obtained from the values determined from slices through different diagonal and cross-peaks.

For ring flip rates sufficiently fast to produce broad and incompletely resolved signals of the *ortho* and *meta* fluorines but showing a narrower signal for the *para* fluorine (see, e.g., Figures 6a and b in the main text), the flip rate was estimated from the difference in signal line widths:

$$k = \pi(\Delta_{om} - \Delta_p) \quad (2)$$

where Δ_{om} is the full line width at half maximum of the ^{19}F NMR signals of the *ortho* or *meta* fluorines and Δ_p is the corresponding line width of the signal of the *para* fluorine.

g) Protein thermal stability

Circular dichroism (CD)

The stabilities of the protein folds of GB1 and Arc-NL were assessed by thermal denaturation monitored by CD spectroscopy using a Chirascan spectrophotometer (Applied Photophysics) and a 1 mm quartz cuvette. The CD signal was monitored at 220 nm for GB1 and 208 nm for Arc-NL over a temperature range of 20 to 94 °C. In all cases, the protein concentration was 5 μM in PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, and 1.8 mM KH_2PO_4, pH 7.4) and measurements were recorded for 4 seconds per point in 1 °C increments. Blanks containing PBS measured in the same cuvette were subtracted from the data. The temperature denaturation was reversible for Arc-NL but not for GB1. The data analysis for the protein melting temperature (T_m) was performed using GraphPad Prism (version 11). Data were analysed by non-linear regression using the Boltzmann sigmoidal equation. The inflection point (V_{50}) was defined as T_m and reported to the nearest degree.

Differential scanning fluorometry (DSF)

The stabilities of PpiB variants were assessed by thermal denaturation monitored by DSF using a QuantStudio real-time PCR thermocycler (Thermo Fisher Scientific). DSF screening was performed following the protein thermal shift (PTS) dye kit (Thermo Fisher Scientific) protocol using a final protein concentration of 0.5 mg mL^{-1} in a volume of 20 µL. Samples were heated from 25 to 95 °C at 0.05 °C s^{-1}. The data analysis was performed using the protein thermal shift software v1.4 by Thermo Fisher Scientific. The midpoint of the unfolding transition (T_m) was obtained from the maximum of the first derivative of fluorescence.

h) Affinity measurements by BLI

Binding kinetics of Arc-NL with the nanobody H11 were measured on a Gator Pivot (Gator Bio) at 30 °C using Anti-His (HIS) probes (Gator Bio). The experiments used the Arc-NL construct with C-terminal His_6-tag (either wild-type or with F_5Phe in positions 267 or 271). Probes were pre-equilibrated in K-buffer (Gator Bio) prior to loading the proteins. The proteins were used at a concentration of 5 µg mL^{-1} to a response of 0.6–0.8 nm, followed by 60 seconds waiting for baseline stabilization in K-buffer. The association with H11 was measured by dipping the sensors into wells containing untagged H11 at 80–2.5 nM in a two-fold dilution series for 180 seconds, followed by transfer into buffer-only wells for 180 seconds to measure the dissociation. Reference probes (ligand-loaded in buffer-only and unloaded probes dipped into analyte at the highest concentration) were included to account for baseline drift and background subtraction. Data were reference-subtracted, aligned and globally fitted to a 1:1 binding model to determine association (k_{on}) and dissociation (k_{off}) rate constants, as well as the equilibrium dissociation constant (K_D). The reported kinetic parameters represent the mean of two independent experiments. Global fits yielded R^2 values > 0.99 and χ^2 values < 1.51 indicative of strong agreement between experimental data and the binding model (Table S1).

i) DFT calculations of the energies of ring rotation

Energy barriers were calculated for rotation of the phenyl ring in $PhCH_2CH_3$ and its perfluorophenyl derivative, representing the equivalent forms of phenylalanine. Energy surface scans were performed for the rotation of the C-C-C-C dihedral angle centred on the Ph-C bond over 90° in steps of 5°. All other coordinates were optimised at each geometry. The optimisations and energy evaluations used restricted Kohn-Sham density functional theory[12] in Orca 6.0.1 [13] with the ωB97X functional[14,15] and the D3 dispersion correction[16] with Becke-Johnson damping.[17] The aug-cc-pVDZ basis set[18] was used for all atoms, with the RIJCOSX approximation[19] applied to two-electron integrals.

j) Protein crystallization and structure determination

PpiB containing F_5Phe in position 4 was concentrated to 11 mg/mL using a 10 kDa molecular weight cut-off (MWCO) centrifugal concentrator prior to crystallizing. Rod-shaped crystals were obtained after several weeks using the hanging drop vapor diffusion method against a reservoir solution containing 35% PEG 3350, 0.2 M sodium acetate trihydrate, and 0.1 M Tris-HCl, pH 8.0. The crystallization drop comprised 1.5 µL protein and 1 µL well solution. Crystals were harvested without further cryoprotection and flash cooled for data collection.

Data of this F_5Phe mutant were collected on the MX3 beamline at the Australian Synchrotron, and diffraction images were processed using Xia2[20] in space group P 1. For all structures, data were truncated using Aimless[21] until statistics were of suitable quality in the outer shell. The structures were solved by molecular replacement (Phaser MR,[22] CCP4) using the structure of wild-type PpiB (PDB ID 2NUL)[23] as the search model. The non-canonical F_5Phe residue was built using chemical ID PF5 in Coot.[24] Refinement was carried out in Phenix.refine.[25] Multiple rounds of refinement and model building in Coot were carried out until R-factors converged and the electron density was modelled as best as possible. MolProbity[26] was used for structure validation. Coordinates were uploaded to the Protein Data Bank under accession ID 26AO.

Arc-NL containing F_5Phe in position 267 and in complex with equimolar H11 was concentrated to 5 mg/mL using a 10 kDa molecular weight cut-off (MWCO) centrifugal concentrator prior to crystallizing. Crystals were obtained after a few days using the hanging drop vapor diffusion method against a reservoir solution containing 30% PEG 3350, 0.2 M magnesium chloride hexahydrate and 0.1 M Tris-HCl, pH 8.0. The crystallization drop was composed of 1 µL protein and 1 µL well solution. Crystals were harvested without further cryoprotection and flash cooled for data collection.

Data of this F_5Phe mutant were collected on the MX3 beamline at the Australian Synchrotron, and diffraction images were processed using XDS[27] in space group P 1 2_1 1. The structure was determined as for PpiB, using wild-type Arc-NL and H11 (PDB ID 8QF4)[28] as separate search models. Model building and the insertion of the F_5Phe (chemical ID: PF5) were performed in Coot.[24] Refinement was carried out using Refmac[29] and Phenix.refine until R-factors converged.[25] MolProbity[26] was used for structure validation. Coordinates have been uploaded to the Protein Data Bank under accession ID 26AP.

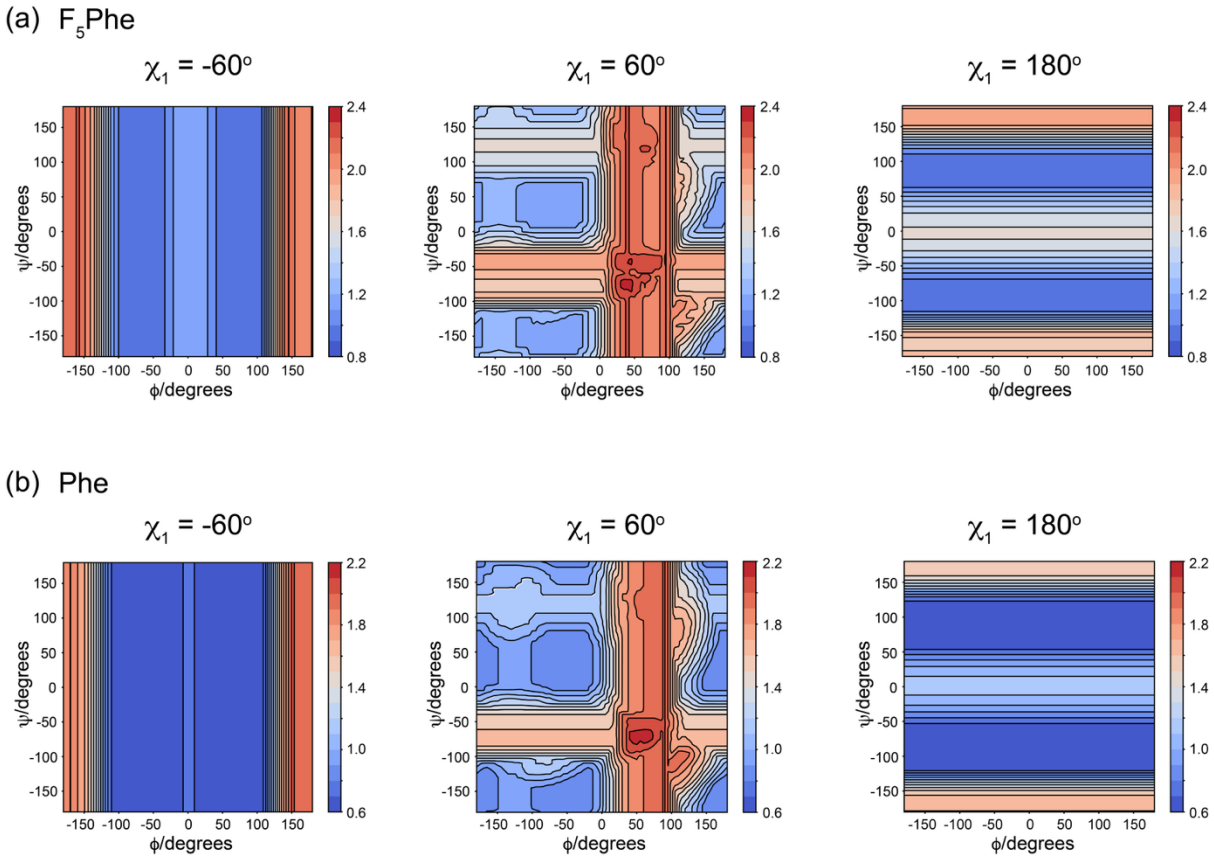

Figure S1. Maximal van der Waals violations (in Å) of a full ring rotation of (a) a F_5Phe and (b) a Phe residue about the C^{β}–C^{γ} axis (χ_2 angle) as a function of backbone conformation (ϕ and ψ angles). Only the van der Waals violations between the ring atoms and the backbone atoms were considered (i.e. ignoring short distances with the $C^{\beta}H_2$ group). For each residue, three separate calculations were performed, starting with the χ_1 angle in one of the three staggered conformations as indicated. For each χ_2 angle passed by the ring rotation, the χ_1 angle was allowed to vary by up to ±30° to minimize the van der Waals violations without deviating greatly from the staggered conformation of the C^{α}–C^{β} bond. The van der Waals radii used were 1.2 Å for hydrogen, 1.7 Å for carbon, 1.55 Å for nitrogen, 1.52 Å for oxygen, and 1.47 Å for fluorine.

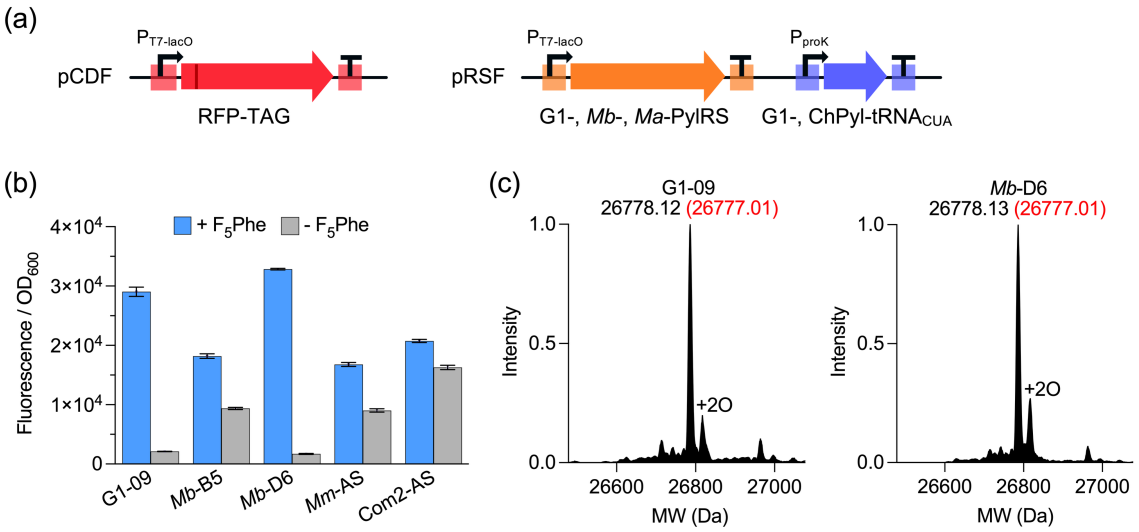

Figure S2. *In vivo* performance of G1-, *Mb*-, and *Mm*-PylRS enzymes for the incorporation of the non-canonical amino acid (ncAA) F_5Phe. (a) Schematic of the pCDF and pRSF dual plasmid system used to assess *in vivo* performance. (b) Level of RFP expression levels reported by fluorescence intensities of the *E. coli* cells with or without F_5Phe provided in the growth medium. Residual fluorescence in the absence of F_5Phe indicates background amber suppression. (c) Intact protein mass spectrometry confirming F_5Phe incorporation fidelity and orthogonality for the two best performing PylRS enzymes, showing the calculated masses in red. For the other enzymes, mass spectrometry indicated a significant level of misincorporation of phenylalanine or glutamine consistent with the level of RFP fluorescence observed in the absence of F_5Phe.

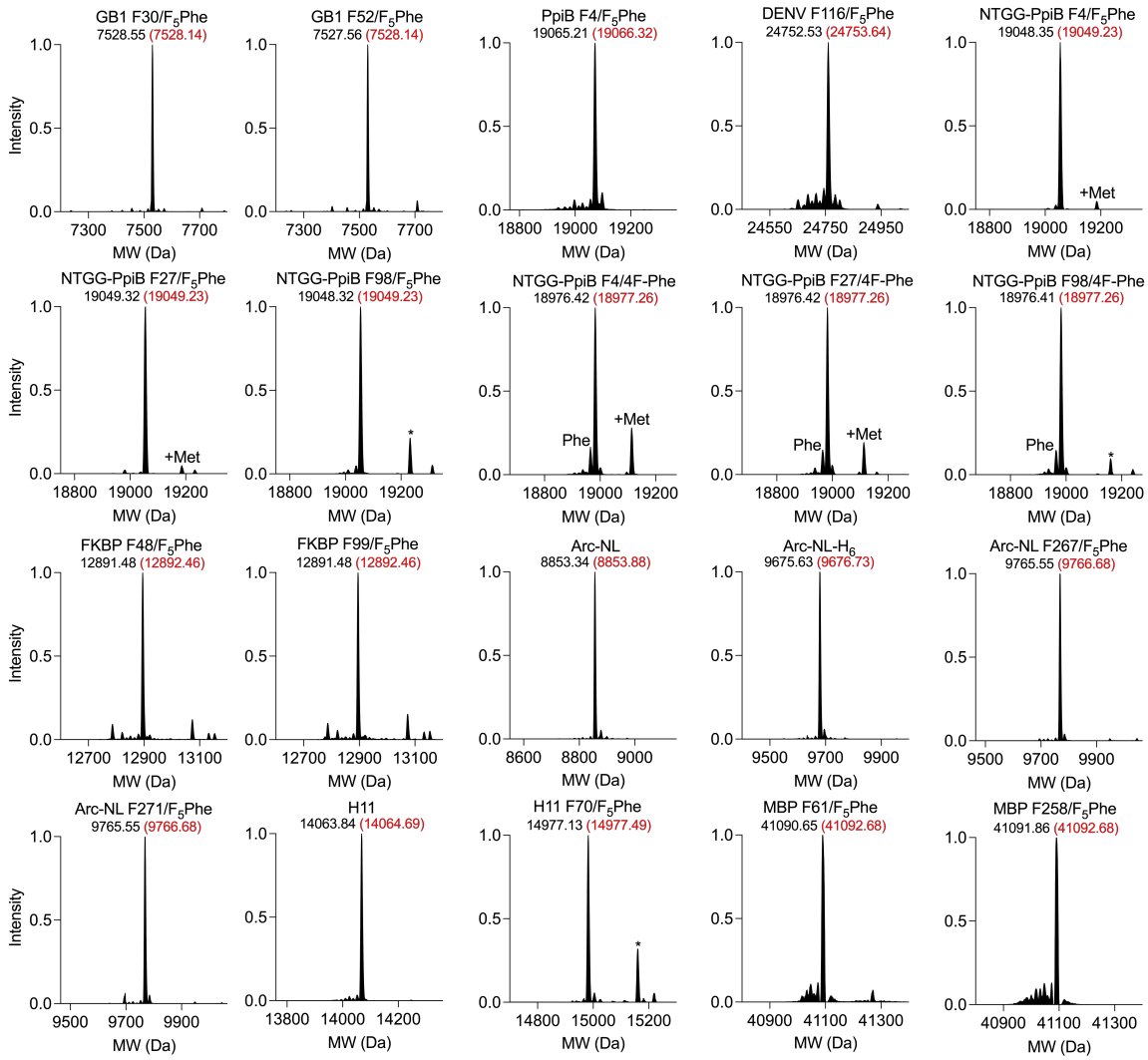

Figure S3. Intact protein mass spectrometry analysis of purified proteins. The observed and expected masses of the full-length proteins are shown in black and red, respectively. Most proteins display close to complete loss of the N-terminal methionine (Met). Peaks corresponding to misincorporation of Phe are labeled. Peaks with an asterisk (*) correspond to a mass increase of +178 Da, indicating α-N-gluconoylation.[30,31]

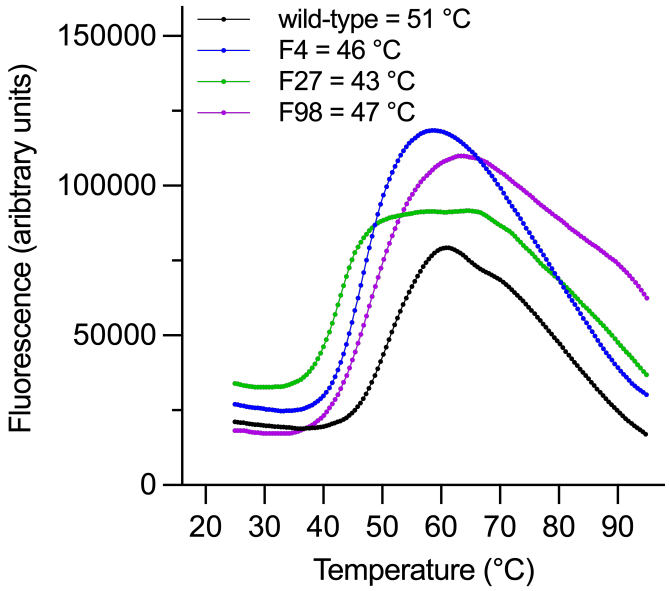

Figure S4. Thermal stability of PpiB and F_5Phe mutants measured by DSF protein thermal shift assays. Representative DSF traces with individual data points are shown with apparent melting temperatures (T_m). The proteins were measured in 50 mM HEPES pH 7.5, 100 mM NaCl. The melting temperatures shown in the insert correspond to the maximum of the first derivative of each curve.

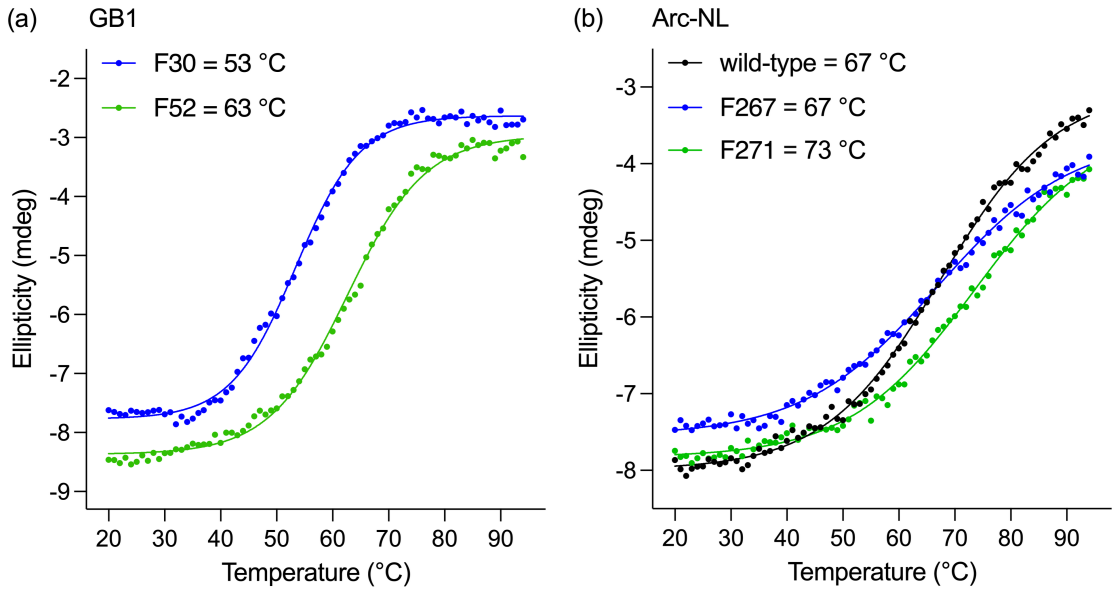

Figure S5. Thermal stability of GB1 and Arc-NL F_5Phe mutants by CD spectroscopy. Individual data points are shown and solid lines represent non-linear fits, from which apparent melting temperatures (T_m) were determined. Conditions: 5 µM protein in PBS (pH 7.4), 1 mm quartz cuvette, Chirascan spectrophotometer (Applied Photophysics). (a) CD signal of GB1 containing F_5Phe in positions F30 (blue) and F52 (green) monitored at 220 nm. CD experiments of the wild-type protein determined a denaturation temperature of about 81 °C.[32] (b) CD signal of Arc-NL wild-type (black) and containing F_5Phe at positions F267 (blue) and F271 (green) monitored at 208 nm.

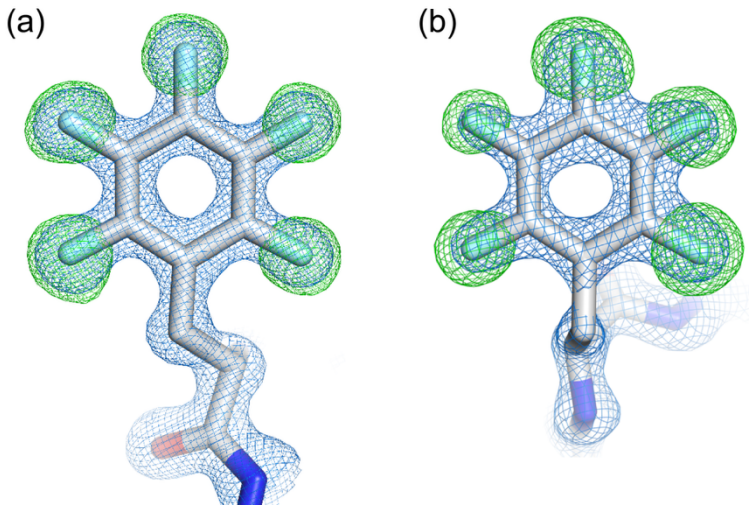

Figure S6. Electron density supporting modeling of the F_5Phe residue in protein crystal structures. The residue is shown in a stick representation, with $2mF_o$-DF_c electron density shown by blue mesh contoured to 2σ, and mF_o-DF_c omit density shown by green mesh. (a) Position 4 of PpiB. Omit density contoured to 4σ. (b) Position 267 of Arc-NL. Omit density contoured to 5σ.

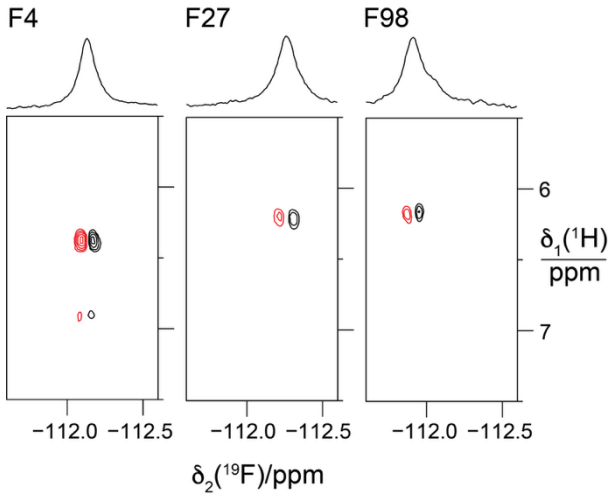

Figure S7. ^{1}H-^{19}F correlation spectra of PpiB produced with 4F-Phe replacing F4, F27, or F98 (corresponding to the sites analyzed in Figure 4 of the main text). The spectra were recorded on a 500 MHz NMR spectrometer using the INEPT pulse sequence 90°(^{1}H) – (Δ + t_1)/2 – 180°(^{1}H,^{19}F) – (Δ - t_1)/2 – 90°(^{1}H,^{19}F) – acquisition (t_2) with $\Delta = t_{1max}$ = 10 ms and t_{2max} = 43.5 ms. The total recording times were 2 h, 14 h, and 1 h for the spectra of 4F-Phe in positions 4, 27, and 98, respectively. The respective protein concentrations were 1.2 mM, 1.2 mM, and 1.6 mM. The ^{19}F resonances observed in 1D ^{19}F NMR spectra (processed with 10 Hz line broadening) are shown at the top. In the case of 4F-Phe in position 4, the relatively narrow ^{19}F NMR signal enabled the observation also of a second cross-peak of weaker intensity as expected for a correlation to a pair of degenerate *ortho* protons, which are coupled with the ^{19}F spin via the smaller $^{4}J_{HF}$ coupling.

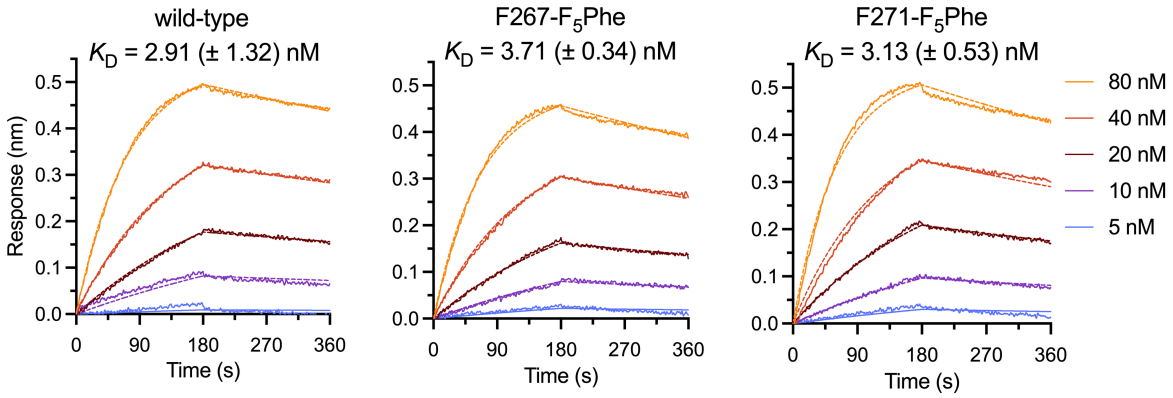

Figure S8. BLI affinity measurements of immobilized Arc-NL-H_6 (wild-type) with F_5Phe in position F267 or F271 binding to untagged H11 nanobody. Measurements (solid line) were fitted to a 1:1 binding model (dashed line). The K_D values were determined by two replicates with the range shown.

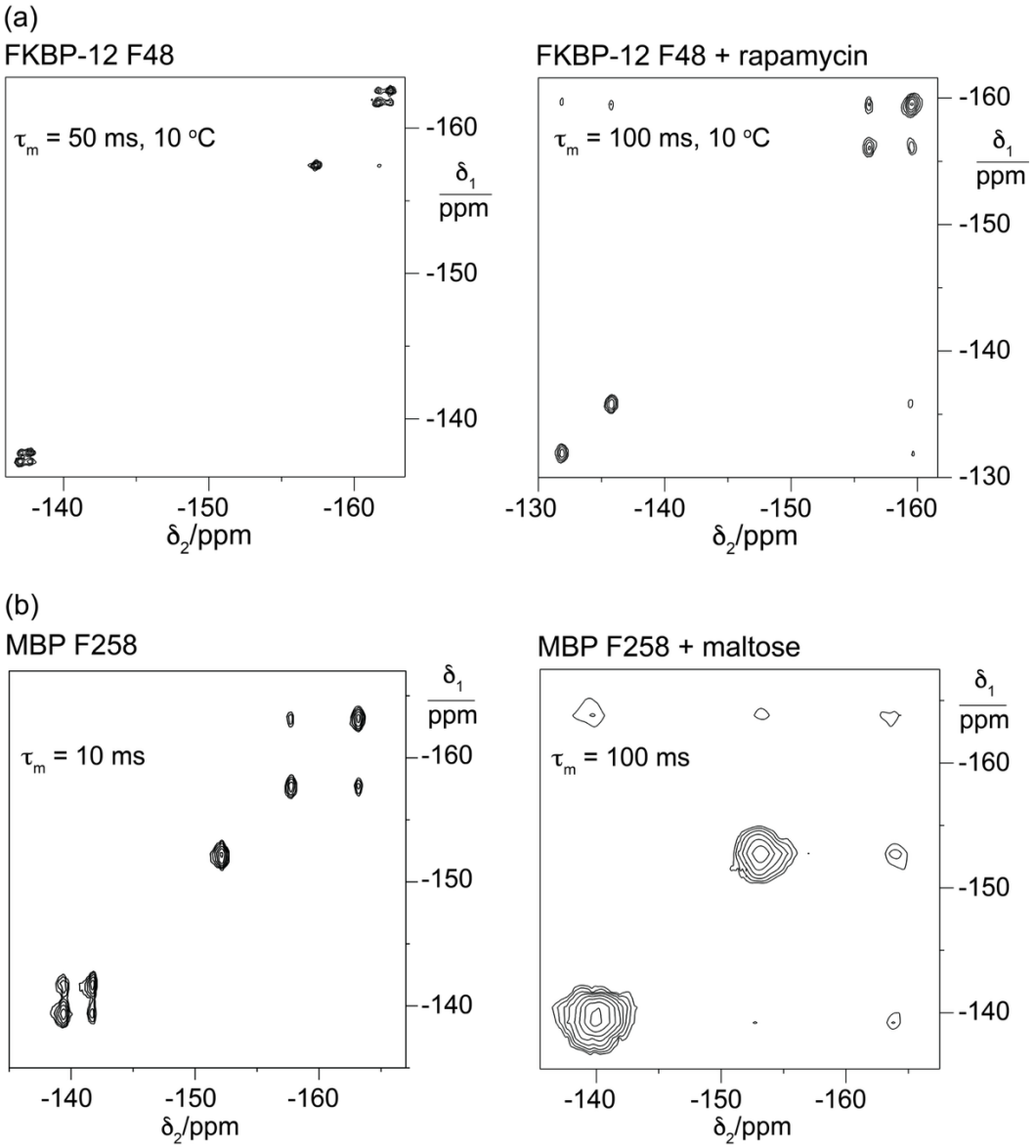

Figure S9. [^{19}F,^{19}F]-NOESY spectra of F_5Phe residues in FKBP-12 and maltose binding protein. Mixing times are indicated in the spectra. See Table S4 for further experimental details. (a) FKBP-12 F48/F_5Phe. The spectra were recorded at 10 °C, as the protein was prone to precipitation. Left panel: free protein. Processed with 70 Hz line broadening in the δ_2 dimension. The intensity ratio of cross-peaks to diagonal peaks indicates aromatic ring flips at a frequency of 11 s^{-1}. Right panel: same as left panel, but in the presence of rapamycin. Processed with 100 Hz line broadening in the δ_2 dimension. The absence of exchange cross-peaks for the *ortho* fluorines indicates ring-flip rates below 1 s^{-1}. (b) Maltose binding protein (MBP) F258/F_5Phe. Left panel: free protein. Spectrum recorded in 21 minutes and processed with 100 Hz line broadening in the δ_2 dimension. The intensity ratios indicate a flip rate of 45 s^{-1}. Right panel: same as left panel, but in the presence of a 2-fold excess of maltose. Spectrum recorded in 42 minutes and processed with 600 Hz line broadening in both dimensions.

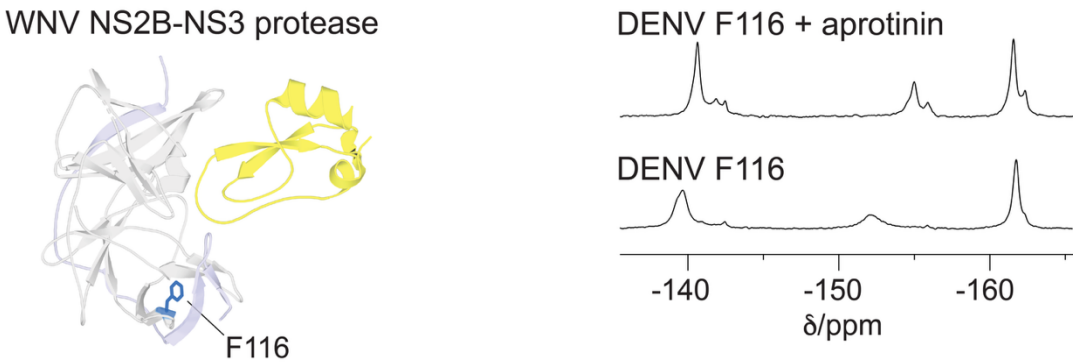

Figure S10. 1D ^{19}F NMR spectra of the dengue virus (serotype 2) NS2B-NS3 protease mutant F116/F_5Phe, free and in complex with aprotinin. The cartoon is of the homologous protease from the West Nile virus (WNV) in complex with aprotinin (PDB ID 2IJO),[33] showing NS2B in violet, NS3 in grey, and aprotinin in yellow. As a crystal structure of the DENV protease is available only in an open form, where the C-terminal part of NS2B is in an inactive conformation,[34] the WNV structure is shown to illustrate the active closed conformation. This structure indicates that the fluorine atoms in the F116/F_5Phe mutant are at least 7 Å from the nearest atom of the inhibitor. ^{19}F NMR of the DENV protein shows significant shifts in response to aprotinin and narrower peaks in the complex than in the free protein.

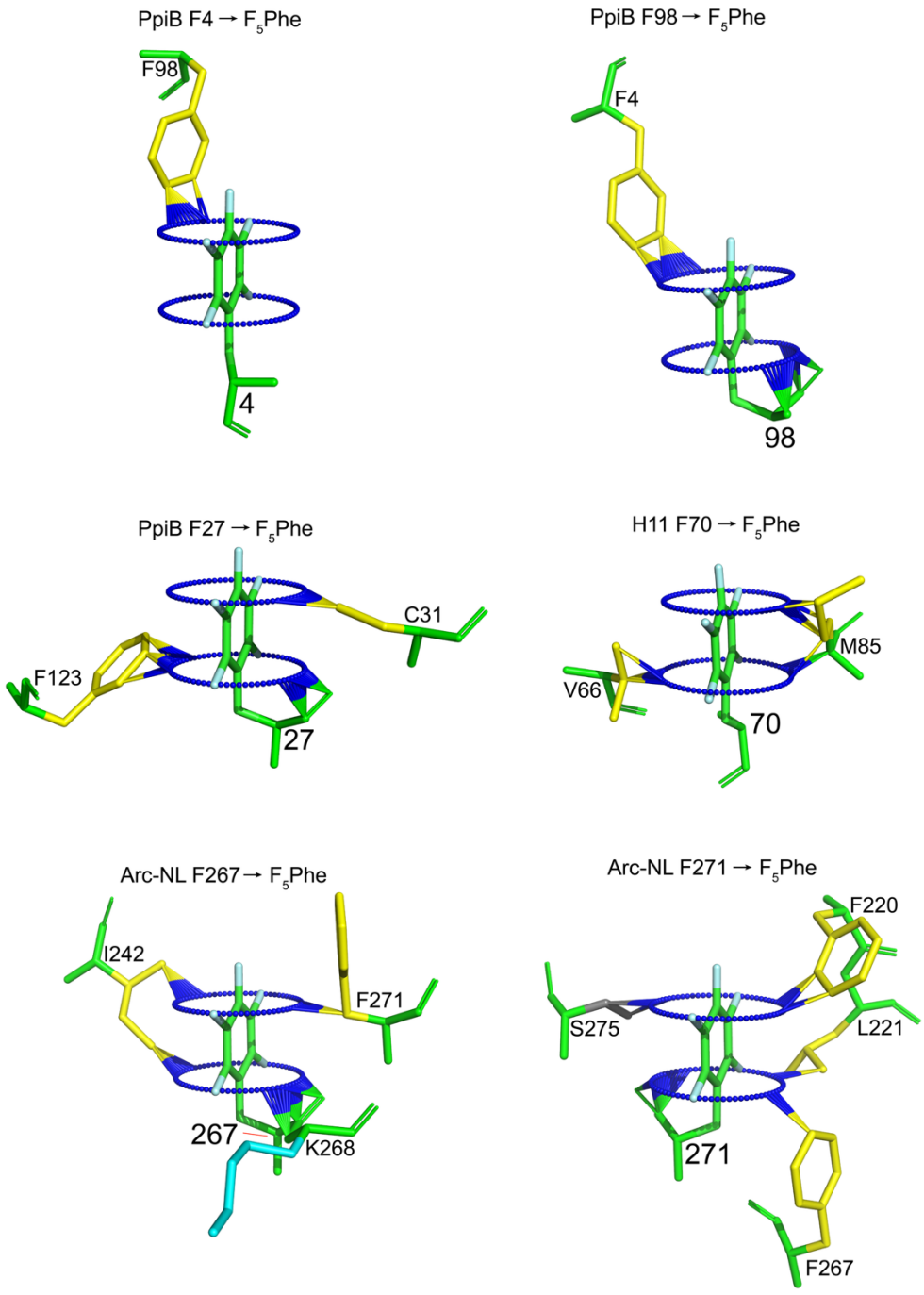

Figure S11 – continued

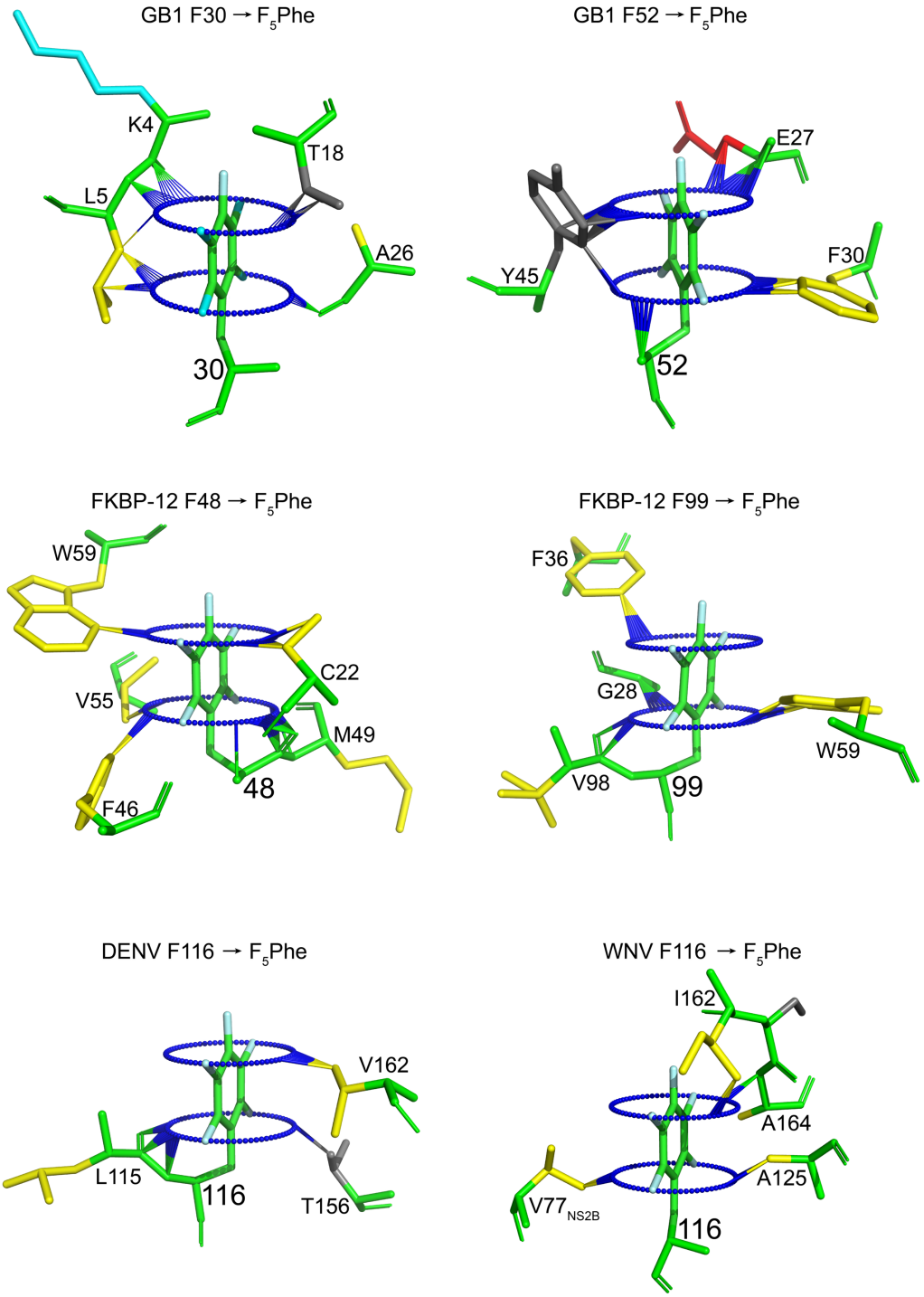

Figure S11 – continued

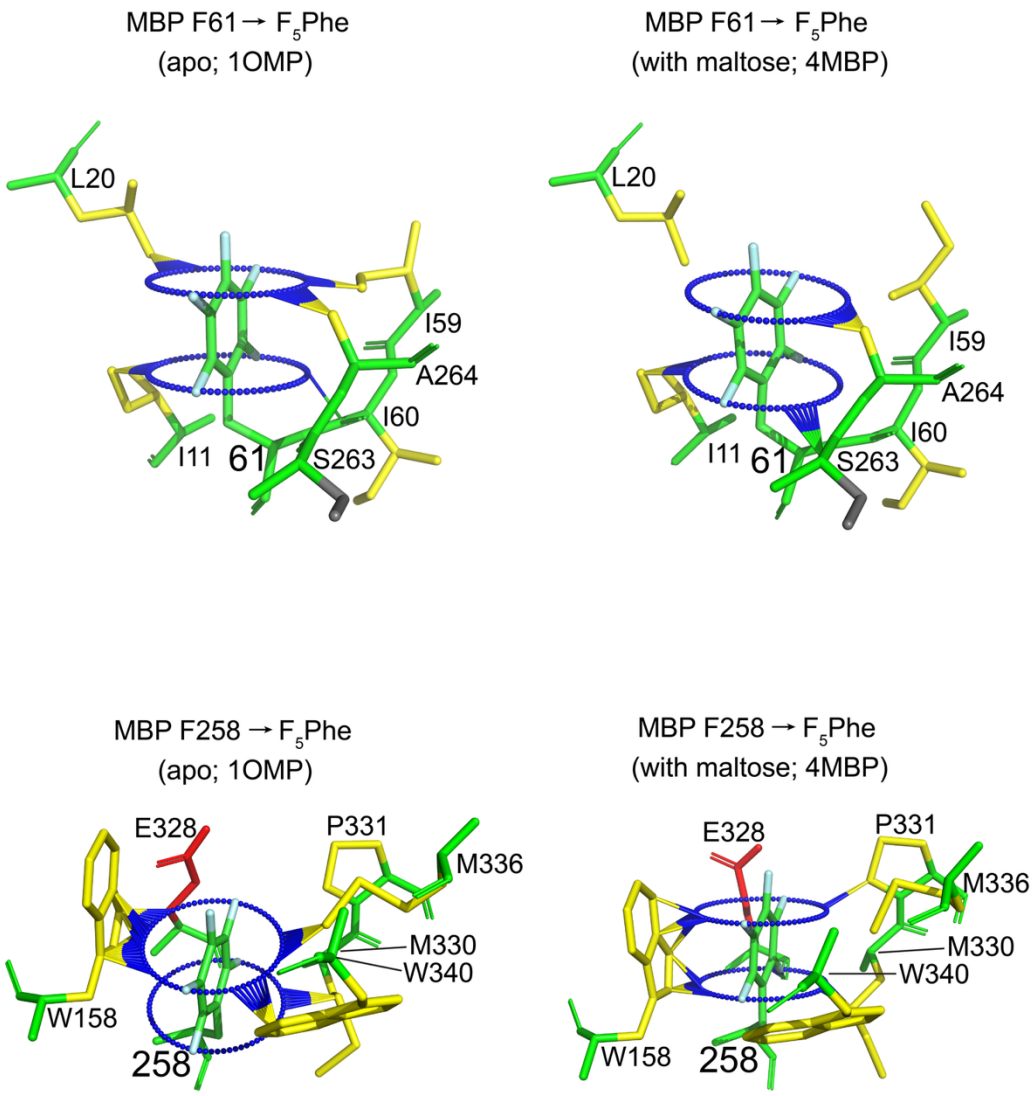

Figure S11. Modeling the ring rotation of the proteins in this work. The Phe→F_5Phe mutation was modeled on the crystal structures by extending the C–H bonds of the phenyl group and renaming the hydrogen to fluorine. The C_6F_5 group was rotated in steps of 5°. Rings of blue spheres highlight the resulting fluorine positions, which display non-physical bonds in PyMOL[35] for very short inter-atomic distances, highlighting the main energy barriers for the ring rotation. The clashing atoms are listed in Table 1 of the main text. The display shows the F_5Phe and nearby residues in a stick representation. Color code: backbone atoms and F_5Phe residue: green (fluorine in cyan); hydrophobic side chains (Ala, Cys, Ile, Leu, Met, Phe, Pro, Trp, Val): yellow; hydrophilic side chains (Asn, Gln, Ser, Thr, Tyr): gray; negatively charged side chains (Asp, Glu): red; positively charged side chains (Arg, His, Lys): cyan. The crystal structures used were 2NUL for PpiB,[23] 8QF4 for the Arc-NL/H11 complex,[28] 2QMT for GB1,[36] 1FKB for FKBP-12,[37] 2FOM for the DENV protease,[34] 2IJO for the WNV protease,[33] and 1OMP and 4MBP for MBP.[38,39]

Table S1. Nucleotide and corresponding amino acid sequences of the proteins used in the current study.

Gene	Used in	DNA sequence	Amino acid sequence[a]
RFP-TAG	Figure S2	ATGGCTTCTATGACCGGTCATCACCATCACCATCA CTAGGCCAGTAGTGAAGACGTTATCAAGGAGTTTA TGCGTTTCAAAGTACGTATGGAGGGTAGTGTTAAC GGACACGAATTTGAGATCGAGGGAGAGGGGGAAGG TCGTCCTTACGAGGGAACTCAAACGGCCAAATTAA AGGTGACCAAAGGTGGGCCCTTGCCATTCGCGTGG GACATCTTGTCACCCCAGTTCCAGTACGGGTCGAA GGCATACGTAAAACACCCAGCGGACATTCCTGACT ATCTTAAGTTATCTTTCCCGGAAGGTTTTAAATGG GAACGCGTGATGAACTTTGAGGATGGGGGGGTTGT TACGGTGACACAAGACTCCTCATTGCAAGATGGAG AGTTTATCTATAAAGTCAAACTTCGCGGCACCAAT TTTCCATCTGACGGTCCTGTAATGCAGAAAAAAAC AATGGGCTGGGAAGCCTCCACAGAACGTATGTACC CCGAAGATGGAGCTTTAAAGGGCGAAATTAAAATG CGCTTAAAACTTAAAGACGGCGGCCATTACGACGC CGAAGTGAAAACGACGTATATGGCTAAGAAACCCG TCCAGCTTCCGGGAGCCTATAAAACTGACATCAAA CTGGATATTACATCACACAACGAAGATTATACTAT TGTCGAACAGTACGAACGCGCCGAAGGCCGCCATT CAACGGGAGCATAA	MASMTGHHHHHHXASSEDVIKEFMRF KVRMEGSVNGHEFEIEGEGEGRPYEG TQTAKLKVTKGGPLPFAWDILSPQFQ YGSKAYVKHPADIPDYLKLSFPEGFK WERVMNFEDGGVVTVTQDSSLQDGEF IYKVKLRGTNFPSDGPVMQKKTMGWE ASTERMYPEDGALKGEIKMRLKLKDG GHYDAEVKTTYMAKKPVQLPGAYKTD IKLDITSHNEDYTIVEQYERAEGRHS TGA
G1-4FF02	Figure S7	ATGGTGGTGAAATTTACCGATAGCCAGATTCAGCA TCTGATGGAATATGGTGATAATGATTGGAGCGAAC CGAATTTGAAGATGCAGCAGCACGTGATAAAGAAT TTAGCAGCCAGTTTAGCAAACTGAAAAGCGCCAAT GATAAAGGCCTGAAAGATGTTATTGCAAATCCGCG TAATGATCTGACCGATCTGGAAAACAAAATTCGCG AAAAACTGGCAGCCCGTGGTTTTATTGAAGTTCAT ACCCCGATTTTTGTGAGCAAAAGCGCACTGGCAAA AATGACCATTACCGAAGATCATCCGCTGTTCAAAC AGGTGTTTTGGATTGATGATAAACGTGCACTGCGT CCGATGCATGCAATGAATGGTATGAAAGTTATGCG TGAACTGCGCGATCATACCAAAGGTCCGGTTAAAA TCTTTGAAATTGGTAGCTGCTTTCGCAAAGAAAGC AAAAGCAGTACCCATCTGGAAGAATTTACCATGCT GAACCTGTTCGAAATGGGTCCTGATGGTGATCCGA TGGAACATCTGAAAATGTATATTGGCGATATCATG GATGCCGTTGGTGTTGAATATACCACCAGTCGTGA AGAATCAGATGTTTTTGTTGAAACCCTGGACGTGG AAATTAATGGCACCGAAGTTGCAAGCGGTGCTGTT GGTCCGCATAAACTGGATCCGGCACATGATGTGCA TGAACCGTATGCAGGTATTGGTTTTGGTCTGGAAC GTCTGCTGATGCTGAAAAATGGTAAAAGCAATGCA CGCAAAACCGGCAAAAGTATTACCTATCTGAATGG CTACAAACTGGATTAA	MVVKFTDSQIQHLMEYGDNDWSEAE FEDAAARDKEFSSQFSKLKSANDKG LKDVIANPRNDLTDLENKIREKLAA RGFIEVHTPIFVSKSALAKMTITED HPLFKQVFWIDDKRALRPMHAMNGM KVMRELRDHTKGPVKIFEIGSCFRK ESKSSTHLEEFTMLNLFEMGPDGDP MEHLKMYIGDIMDAVGVEYTTSREE SDVFVETLDVEINGTEVASGAVGPH KLDPAHDVHEPYAGIGFGLERLLML KNGKSNARKTGKSITYLNGYKLD
G1-F5F09[b]	Figure S2	ATGGTGGTGAAATTTACCGATAGCCAGATTCAGCA TCTGATGGAATATGGTGATAATGATTGGAGCGAAG CCGAATTTGAAGATGCAGCAGCACGTGATAAAGAA TTTAGCAGCCAGTTTAGCAAACTGAAAAGCGCCAA TGATAAAGGCCTGAAAGATGTTATTGCAAATCCGC GTAATGATCTGACCGATCTGGAAAACAAAATTCGC GAAAAACTGGCAGCCCGTGGTTTTATTGAAGTTCA TACCCCGATTTTTGTGAGCAAAAGCGCACTGGCAA AAATGACCATTACCGAAGATCATCCGCTGTTCAAA CAGGTGTTTTGGATTGATGATAAACGTGCACTGCG	MVVKFTDSQIQHLMEYGDNDWSEAEF EDAAARDKEFSSQFSKLKSANDKGLK DVIANPRNDLTDLENKIREKLAARGF IEVHTPIFVSKSALAKMTITEDHPLF KQVFWIDDKRALRPMHAMNVCKVMRE LRDHTKGPVKIFEIGSCFRKESKSST HLEEFTMLTLGEMGPDGDPMEHLKMY IGDIMDAVGVEYTTSREESDVYVETL DVEINGTEVASGGVGPHKLDPAHDVH

		TCCGATGCATGCAATGAATGTTTGTAAAGTTATGC GTGAACTGCGCGATCATACCAAAGGTCCGGTTAAA ATCTTTGAAATTGGTAGCTGCTTTCGCAAAGAAAG CAAAAGCAGTACCCATCTGGAAGAATTTACCATGC TGACCCTGGGAGAAATGGGTCCTGATGGTGATCCG ATGGAACATCTGAAAATGTATATTGGCGATATCAT GGATGCCGTTGGTGTTGAATATACCACCAGTCGTG AAGAATCAGATGTTTATGTTGAAACCCTGGACGTG GAAATTAATGGCACCGAAGTTGCAAGCGGTGGGGT TGGTCCGCATAAACTGGATCCGGCACATGATGTGC ATGAACCGTGGGCAGGTATTGGTTTTGGTCTGGAA CGTCTGCTGATGCTGAAAAATGGTAAAAGCAATGC ACGCAAAACCGGCAAAAGTATTACCTATCTGAATG GCTACAAACTGGATTAA	EPWAGIGFGLERLLMLKNGKSNARKT GKSITYLNGYKLD
Mb-B5	Figure S2	ATGGATAAAAAGCCGCTGGATGTTCTGATTAGCGC AACCGGTCTGTGGATGAGCCGCACCGGTACACTGC ATAAAATCAAACATCATGAAGTGAGCCGCAGCAAA ATCTATATTGAAATGGCATGTGGTGATCATCTGGT GGTGAATAATAGCCGTAGCTGTCGTACCGCACGTG CATTTCGTCATCACAAATATCGTAAAACCTGTAAA CGTTGCCGTGTTAGCGACGAAGATATTAACAATTT TCTGACCCGTAGCACCGAAAGCAAAAATTCAGTTA AAGTTCGTGTTGTGAGCGCTCCGAAAGTTAAAAAG GCAATGCCGAAAAGCGTTAGTCGTGCACCGAAACC GCTGGAAAATAGCGTTAGCGCAAAAGCAAGCACCA ATACCAGCCGTAGCGTTCCGAGTCCGGCAAAAAGC ACCCCGAATAGCAGCGTTCCGGCAAGCGCACCGGC ACCGAGCCTGACACGTAGCCAGCTGGATCGTGTTG AAGCACTGCTGAGTCCGGAAGATAAAATCAGCCTG AATATGGCAAAACCGTTTCGTGAACTGGAACCGGA ACTGGTTACCCGTCGTAAAAATGATTTTCAGCGTC TGTATACCAACGATCGCGAAGATTATCTGGGTAAA CTGGAACGTGATATTACCAAATTTTTCGTGGATCG CGGTTTCCTGGAAATCAAAAGCCCGATTCTGATTC CGGCAGAATATGTTGAACGTATGGGCATTAATAAC GATACCGAACTGAGCAAGCAGATTTTTCGCGTTGA TAAAAATCTGTGTCTGCGTCCGATGCTGGCACCGA CACTGTATAACTATCTGCGCAAACTGGATCGTATT CTGCCTGGTCCGATTAAAATCTTTGAAGTTGGTCC GTGCTATCGCAAAGAAAGTGATGGTAAAGAACACC TGGAAGAGTTTACCATGGTTGGTTTTGCACAGATG GGTAGCGGTTGTACCCGTGAAAATCTGGAAGCACT GATTAAAGAGTTTCTGGACTATCTGGAAATCGACT TTGAAATTGTTGGTGATAGCTGCATGGTTTATGGT GATACCCTGGATATTATGCATGGTGATCTGGAACT GAGTAGCGCAGTTGTTGGTCCGGTTAGCCTGGATC GCGAATGGGGTATTGATAAACCGTGGATTGGTGCA GGTTTTGGTCTGGAACGTCTGCTGAAAGTTATGCA CGGCTTTAAAAACATTAAACGTGCAAGCCGTTCCG AGAGCTATTACAATGGTATTAGCACCAACCTGTAA	MDKKPLDVLISATGLWMSRTGTLHKI KHHEVSRSKIYIEMACGDHLVVNNSR SCRTARAFRHHKYRKTCKRCRVSDED INNFLTRSTESKNSVKVRVVSAPKVK KAMPKSVSRAPKPLENSVSAKASTNT SRSVPSPAKSTPNSSVPASAPAPSLT RSQLDRVEALLSPEDKISLNMAKPFR ELEPELVTRRKNDFQRLYTNDREDYL GKLERDITKFFVDRGFLEIKSPILIP AEYVERMGINNDTELSKQIFRVDKNL CLRPMLAPTLYNYLRKLDRILPGPIK IFEVGPCYRKESDGKEHLEEFTMVGF AQMGSGCTRENLEALIKEFLDYLEID FEIVGDSCMVYGDTLDIMHGDLELSS AVVGPVSLDREWGIDKPWIGAGFGLE RLLKVMHGFKNIKRASRSESYYNGIS TNL
Mb-D6	Figure S2	ATGGATAAAAAGCCGCTGGATGTTCTGATTAGCGC AACCGGTCTGTGGATGAGCCGCACCGGTACACTGC ATAAAATCAAACATCATGAAGTGAGCCGCAGCAAA ATCTATATTGAAATGGCATGTGGTGATCATCTGGT GGTGAATAATAGCCGTAGCTGTCGTACCGCACGTG CATTTCGTCATCACAAATATCGTAAAACCTGTAAA CGTTGCCGTGTTAGCGACGAAGATATTAACAATTT TCTGACCCGTAGCACCGAAAGCAAAAATTCAGTTA AAGTTCGTGTTGTGAGCGCTCCGAAAGTTAAAAAG GCAATGCCGAAAAGCGTTAGTCGTGCACCGAAACC GCTGGAAAATAGCGTTAGCGCAAAAGCAAGCACCA	MDKKPLDVLISATGLWMSRTGTLHKI KHHEVSRSKIYIEMACGDHLVVNNSR SCRTARAFRHHKYRKTCKRCRVSDED INNFLTRSTESKNSVKVRVVSAPKVK KAMPKSVSRAPKPLENSVSAKASTNT SRSVPSPAKSTPNSSVPASAPAPSLT RSQLDRVEALLSPEDKISLNMAKPFR ELEPELVTRRKNDFQRLYTNDREDYL GKLERDITKFFVDRGFLEIKSPILIP AEYVERMGINNDTELSKQIFRVDKNL CLRPMLAPTLYNYLRKLDRILPGPIK

		ATACCAGCCGTAGCGTTCCGAGTCCGGCAAAAAGC ACCCCGAATAGCAGCGTTCCGGCAAGCGCACCGGC ACCGAGCCTGACACGTAGCCAGCTGGATCGTGTTG AAGCACTGCTGAGTCCGGAAGATAAAATCAGCCTG AATATGGCAAAACCGTTTCGTGAACTGGAACCGGA ACTGGTTACCCGTCGTAAAAATGATTTTCAGCGTC TGTATACCAACGATCGCGAAGATTATCTGGGTAAA CTGGAACGTGATATTACCAAATTTTTCGTGGATCG CGGTTTCCTGGAAATCAAAAGCCCGATTCTGATTC CGGCAGAATATGTTGAACGTATGGGCATTAATAAC GATACCGAACTGAGCAAGCAGATTTTTCGCGTTGA TAAAAATCTGTGTCTGCGTCCGATGCTGGCACCGA CACTGTATAACTATCTGCGCAAACTGGATCGTATT CTGCCTGGTCCGATTAAAATCTTTGAAGTTGGTCC GTGCTATCGCAAAGAAAGTGATGGTAAAGAACACC TGGAAGAGTTTACCATGGTTGCATTTATGCAGATG GGTAGCGGTTGTACCCGTGAAAATCTGGAAGCACT GATTAAAGAGTTTCTGGACTATCTGGAAATCGACT TTGAAATTGTTGGTGATAGCTGCATGGTTTATGGT GATACCCTGGATATTATGCATGGTGATCTGGAACT GAGTAGTGCCGGTGTTGGTCCGGTTAGCCTGGATC GCGAATGGGGTATTGATAAACCGACCATTGGTGCA GGTTTTGGTCTGGAACGTCTGCTGAAAGTTATGCA CGGCTTTAAAAACATTAAACGTGCAAGCCGTTCCG AGAGCTATTACAATGGTATTAGCACCAACCTGTAA	IFEVGPCYRKESDGKEHLEEFTMVAF MQMGSGCTRENLEALIKEFLDYLEID FEIVGDSCMVYGDTLDIMHGDLELSS AGVGPVSLDREWGIDKPTIGAGFGLE RLLKVMHGFKNIKRASRSESYYNGIS TNL
Mm-AS	Figure S2	ATGGATAAAAAGCCGCTGAATACCCTGATTAGCGC AACCGGTCTGTGGATGAGCCGTACCGGCACCATTC ATAAAATCAAACATCATGAAGTGAGCCGCAGCAAA ATCTATATTGAAATGGCATGTGGTGATCATCTGGT GGTGAATAATAGCCGTAGCAGTCGTACCGCACGTG CACTGCGTCATCACAAATATCGTAAAACCTGTAAA CGTTGCCGTGTTAGTGATGAGGATCTGAACAAATT TCTGACCAAAGCCAATGAAGATCAGACCAGCGTTA AAGTTAAAGTTGTTAGCGCACCGACACGTACCAAA AAGGCAATGCCGAAAAGCGTTGCCCGTGCACCGAA ACCGCTGGAAAATACCGAAGCAGCACAGGCACAGC CGAGCGGTAGCAAATTTTCACCGGCAATTCCGGTT AGCACCCAAGAAAGCGTTAGCGTTCCGGCAAGCGT GAGCACCAGCATTAGCAGCATTTCAACCGGTGCAA CCGCAAGCGCACTGGTTAAAGGTAATACCAATCCG ATTACCAGCATGAGCGCACCGGTTCAGGCAAGTGC ACCGGCACTGACCAAAAGTCAGACCGATCGTCTGG AAGTTCTGCTGAATCCGAAAGATGAAATTAGCCTG AATAGCGGTAAACCGTTTCGTGAACTGGAAAGCGA ACTGCTGAGCCGTCGTAAAAAGGATCTGCAGCAGA TTTATGCCGAAGAACGTGAAAACTATCTGGGTAAA CTGGAACGTGAGATTACCCGTTTCTTTGTTGATCG TGGTTTCCTGGAAATCAAAAGCCCGATTCTGATTC CGCTGGAATATATTGAACGTATGGGCATTGATAAC GATACCGAACTGAGCAAACAAATCTTTCGTGTGGA CAAAAACTTTTGTCTGCGTCCGATGCTGGCACCGA ATCTGTATAACTACCTGCGCAAACTGGATCGTGCC CTGCCGGATCCGATTAAAATCTTTGAAATTGGTCC GTGCTACCGCAAAGAAAGTGATGGTAAAGAACACC TGGAAGAATTTACCATGCTGGCATTTAGCCAGATG GGTAGCGGTTGTACCCGTGAAAATCTGGAAAGCAT TATTACCGATTTTCTGAACCATCTGGGCATCGATT TCAAAATTGTTGGTGATAGCTGCATGGTGTATGGT GATACCCTGGATGTTATGCATGGTGATCTGGAACT GAGTAGCGCAGTTGTTGGTCCGATTCCTCTGGATC GCGAATGGGGTATTGATAAACCGTGGATTGGTGCA GGTTTTGGTCTGGAACGCCTGCTGAAAGTTAAACA	MDKKPLNTLISATGLWMSRTGTIHKI KHHEVSRSKIYIEMACGDHLVVNNSR SSRTARALRHHKYRKTCKRCRVSDED LNKFLTKANEDQTSVKVKVVSAPTRT KKAMPKSVARAPKPLENTEAAQAQPS GSKFSPAIPVSTQESVSVPASVSTSI SSISTGATASALVKGNTNPITSMSAP VQASAPALTKSQTDRLEVLLNPKDEI SLNSGKPFRELESELLSRRKKDLQQI YAEERENYLGKLEREITRFFVDRGFL EIKSPILIPLEYIERMGIDNDTELSK QIFRVDKNFCLRPMLAPNLYNYLRKL DRALPDPIKIFEIGPCYRKESDGKEH LEEFTMLAFSQMGSGCTRENLESIIT DFLNHLGIDFKIVGDSCMVYGDTLDV MHGDLELSSAVVGPIPLDREWGIDKP WIGAGFGLERLLKVKHDFKNIKRAAR SESYYNGISTNL

		CGACTTCAAAAACATTAAACGTGCAGCACGTAGCG AGAGCTATTACAATGGTATTAGTACCAACCTGTAA	
Com2-Mm-AS	Figure S2	ATGAATAAAAAGCCGCTGTATACCCTGATTTCGGC GACCGGGCTGTGGATGAGCCGTACGGGTACGATTC ATAAAATTAAACACCACGAAATTTCGCGTAGCAAA ATTTACATTGAAATGGCATGTGGCGATCATCTGGT TGTCAACAATAGTCGTAGTAGCCGTCCGGCACGTG CACTGAAATACCTGAAATACCGCAATACATGTAAA CGTTGTCGTTGGAGCGATGAGGATCTGAATAAATT TCTGACCAAAGCAAATGAAGACCAAACCAGCGTTA AAGTGAAAGTTGTAAGCGCACCGACCCGCACCAAA AAGGCCATGCCGAAAAGCGTTGCTCGTGCACCGAA ACCGCTGGAAAACAGCGAAGCAGCACAGGCACAGC CGAGCGGCAGCAAATTTAGCCCGGCAATTCCAGTA TCAACCCAAGAAAGCGTTAGCGTCCCAGCAAGTGT AAGCACCAGCATCAGCAGCATTAGCACCGGTGCAA CCGCAAGTGCACTGGTGAAAGGAAACACAAATCCT ATTACCAGCATGTCCGCACCAGTACAGGCAAGCGC ACCAGCACTGACCAAACGTCAAACCGATCGTCTGG AAGTTCTGCTGAACCCGAAAGATGAAATCTCACTG AATAGCGGCAAACCGTTCCGTGAACTGGAAAGCGA ACTGCTGAGCCGTCGTAAAAAGGATCTGCAACAAA TTTATGCCGAAGAGCGTGAAAACTACCTGGGGAAA CTGGAACGTGAAATCACCCGTTTTTTCGTTGATCG TGGTTTTCTGGAAATTAAATCACCTATTCTGATCC CGCTGGAATATATTGAACGTATGGGAATTGATAAC GATACTGAGCTGAGCAAACAGATTTTTCGCGTTGA TAAAAATTTCTGTCTGCGTCCGATGCTGGCACCGA ATCTGTATAACTACCTGCGTAAACTGGATCGTGCG CTGCCGGATCCGATTAAAATTTTTGAGATTGGTCC GTGTTACCGTAAAGAAAGCGACGGTAAAGAACATC TGGAAGAATTTACCATGCTGGCGTTTTCTCAAATG GGTAGCGGTTGTACACGTGAAAATCTGGAAAGCAT TATTACCGATTTCCTGAATCATCTGGGAATTGATT TTAAAATCGTCGGTGATAGCTGCATGGTTTATGGT GATACTCTGGATGTGATGCATGGTGACCTGGAACT GAGCAGCGCAGTTGTTGGGCCGATTCCGCTGGACC GTGAATGGGGTATTGATAAACCGTGGATTGGTGCG GGTTTTGGTCTGGAACGCCTGCTGAAAGTGAAACA TGATTTTAAAAACATCAAACGCGCAGCACGTAGCG AGAGCTATTATAATGGAATTAGTACGAATCTGTAA	MNKKPLYTLISATGLWMSRTGTIHKI KHHEISRSKIYIEMACGDHLVVNNSR SSRPARALKYLKYRNTCKRCRWSDED LNKFLTKANEDQTSVKVKVVSAPTRT KKAMPKSVARAPKPLENSEAAQAQPS GSKFSPAIPVSTQESVSVPASVSTSI SSISTGATASALVKGNTNPITSMSAP VQASAPALTKRQTDRLEVLLNPKDEI SLNSGKPFRELESELLSRRKKDLQQI YAEERENYLGKLEREITRFFVDRGFL EIKSPILIPLEYIERMGIDNDTELSK QIFRVDKNFCLRPMLAPNLYNYLRKL DRALPDPIKIFEIGPCYRKESDGKEH LEEFTMLAFSQMGSGCTRENLESIIT DFLNHLGIDFKIVGDSCMVYGDTLDV MHGDLELSSAVVGPIPLDREWGIDKP WIGAGFGLERLLKVKHDFKNIKRAAR SESYYNGISTNL
G1Pyl-tRNA$_{CUA}$	Figure S2	GGAGGGCGCTCCGGCGAGCAAACGGGTCTCTAAAA CCTGTAAGCGGGGTTCGACCCCCCGGCCTTTCGCC A	
ChPyl-tRNA$_{CUA}$	Figure S2	GGAAACGTGATCATGTAGATCGAACGGACTCTAAA TCCGTTCAGTGGGGTTAGATTCCCCACGTTTCCGC CA	
GB1 F30TAG	Figure 7 Figure S5	ATGGCTTCTATGACCGGTATGACCTACAAACTGAT CCTGAACGGTAAAACCCTGAAAGGTGAAACCACCA CCGAAGCGGTTGACGCGGCGACCGCGGAAAAAGTT TAGAAACAGTACGCGAACGACAACGGTGTTGACGG TGAATGGACCTACGACGACGCGACCAAAACCTTCA CCGTTACCGAAGAAAACCTGTATTTTCAGGGCCAT CATCATCACCATCACTAA	MASMTGMTYKLILNGKTLKGETTTEA VDAATAEKVXKQYANDNGVDGEWTYD DATKTFTVTEENLYFQGHHHHHH

GB1 F52TAG [c]	Figure 7 Figure S5	ATGGCTTCTATGACCGGTATGACCTACAAACTGAT CCTGAACGGTAAAACCCTGAAAGGTGAAACCACCA CCGAAGCGGTTGACGCGGCGACCGCGGAAAAAGTT TTCAAACAGTACGCGAACGACAACGGTGTTGACGG TGAATGGACCTACGACGACGCGACCAAAACCTAGA CCGTTACCGAAGAAAACCTGTATTTTCAGGGCCAT CATCATCACCATCACTAA	MASMTGMTYKLILNGKTLKGETTTEA VDAATAEKVFKQYANDNGVDGEWTYD DATKTXTVTEENLYFQGHHHHHH
PpiB F4TAG	Figure 3 Figure S6	ATGGTTACTTAGCACACCAATCACGGCGATATTGT CATCAAAACTTTTGACGATAAAGCACCTGAAACAG TTAAAAACTTCCTGGACTACTGCCGCGAAGGTTTT TACAACAACACCATTTTCCACCGTGTTATCAACGG CTTTATGATTCAGGGCGGCGGTTTTGAACCGGGCA TGAAACAAAAAGCCACCAAAGAACCGATCAAAAAC GAAGCCAACAACGGCCTGAAAAATACCCGTGGTAC GCTGGCAATGGCACGTACTCAGGCTCCGCACTCTG CAACTGCACAGTTCTTCATCAACGTGGTTGATAAC GACTTCCTGAACTTCTCTGGCGAAAGCCTGCAAGG TTGGGGCTACTGCGTGTTTGCTGAAGTGGTTGACG GCATGGACGTGGTAGACAAAATCAAAGGTGTTGCA ACCGGTCGTAGCGGTATGCATCAGGACGTGCCAAA AGAAGACGTTATCATTGAAAGCGTGACCGTTAGCG AGCACCACCATCATCACCACTAA	MVTXHTNHGDIVIKTFDDKAPETVKN FLDYCREGFYNNTIFHRVINGFMIQG GGFEPGMKQKATKEPIKNEANNGLKN TRGTLAMARTQAPHSATAQFFINVVD NDFLNFSGESLQGWGYCVFAEVVDGM DVVDKIKGVATGRSGMHQDVPKEDVI IESVTVSEHHHHHH
NTGG-PpiB	Figure S4	ATGGGCGGTGTTACTTTCCACACCAATCACGGCGA TATTGTCATCAAAACTTTTGACGATAAAGCACCTG AAACAGTTAAAAACTTCCTGGACTACTGCCGCGAA GGTTTTTACAACAACACCATTTTCCACCGTGTTAT CAACGGCTTTATGATTCAGGGCGGCGGTTTTGAAC CGGGCATGAAACAAAAAGCCACCAAAGAACCGATC AAAAACGAAGCCAACAACGGCCTGAAAAATACCCG TGGTACGCTGGCAATGGCACGTACTCAGGCTCCGC ACTCTGCAACTGCACAGTTCTTCATCAACGTGGTT GATAACGACTTCCTGAACTTCTCTGGCGAAAGCCT GCAAGGTTGGGGCTACTGCGTGTTTGCTGAAGTGG TTGACGGCATGGACGTGGTAGACAAAATCAAAGGT GTTGCAACCGGTCGTAGCGGTATGCATCAGGACGT GCCAAAAGAAGACGTTATCATTGAAAGCGTGACCG TTAGCGAGCACCACCATCATCACCACTAA	MVTFHTNHGDIVIKTFDDKAPETVKN FLDYCREGFYNNTIFHRVINGFMIQG GGFEPGMKQKATKEPIKNEANNGLKN TRGTLAMARTQAPHSATAQFFINVVD NDFLNFSGESLQGWGYCVFAEVVDGM DVVDKIKGVATGRSGMHQDVPKEDVI IESVTVSEHHHHHH
NTGG-PpiB F4TAG	Figure 4 Figure S4 Figure S7	ATGGGCGGTGTTACTTAGCACACCAATCACGGCGA TATTGTCATCAAAACTTTTGACGATAAAGCACCTG AAACAGTTAAAAACTTCCTGGACTACTGCCGCGAA GGTTTTTACAACAACACCATTTTCCACCGTGTTAT CAACGGCTTTATGATTCAGGGCGGCGGTTTTGAAC CGGGCATGAAACAAAAAGCCACCAAAGAACCGATC AAAAACGAAGCCAACAACGGCCTGAAAAATACCCG TGGTACGCTGGCAATGGCACGTACTCAGGCTCCGC ACTCTGCAACTGCACAGTTCTTCATCAACGTGGTT GATAACGACTTCCTGAACTTCTCTGGCGAAAGCCT GCAAGGTTGGGGCTACTGCGTGTTTGCTGAAGTGG TTGACGGCATGGACGTGGTAGACAAAATCAAAGGT GTTGCAACCGGTCGTAGCGGTATGCATCAGGACGT GCCAAAAGAAGACGTTATCATTGAAAGCGTGACCG TTAGCGAGCACCACCATCATCACCACTAA	MGGVTXHTNHGDIVIKTFDDKAPETV KNFLDYCREGFYNNTIFHRVINGFMI QGGGFEPGMKQKATKEPIKNEANNGL KNTRGTLAMARTQAPHSATAQFFINV VDNDFLNFSGESLQGWGYCVFAEVVD GMDVVDKIKGVATGRSGMHQDVPKED VIIESVTVSEHHHHHH

NTGG-PpiB F27TAG	Figure 4 Figure S4	ATGGGCGGTGTTACTTTCCACACCAATCACGGCGA TATTGTCATCAAAACTTTTGACGATAAAGCACCTG AAACAGTTAAAAACTAGCTGGACTACTGCCGCGAA GGTTTTTACAACAACACCATTTTCCACCGTGTTAT CAACGGCTTTATGATTCAGGGCGGCGGTTTTGAAC CGGGCATGAAACAAAAAGCCACCAAAGAACCGATC AAAAACGAAGCCAACAACGGCCTGAAAAATACCCG TGGTACGCTGGCAATGGCACGTACTCAGGCTCCGC ACTCTGCAACTGCACAGTTCTTCATCAACGTGGTT GATAACGACTTCCTGAACTTCTCTGGCGAAAGCCT GCAAGGTTGGGGCTACTGCGTGTTTGCTGAAGTGG TTGACGGCATGGACGTGGTAGACAAAATCAAAGGT GTTGCAACCGGTCGTAGCGGTATGCATCAGGACGT GCCAAAAGAAGACGTTATCATTGAAAGCGTGACCG TTAGCGAGCACCACCATCATCACCACTAA	MGGVTFHTNHGDIVIKTFDDKAPETV KNXLDYCREGFYNNTIFHRVINGFMI QGGGFEPGMKQKATKEPIKNEANNGL KNTRGTLAMARTQAPHSATAQFFINV VDNDFLNFSGESLQGWGYCVFAEVVD GMDVVDKIKGVATGRSGMHQDVPKED VIIESVTVSEHHHHHH
NTGG-PpiB F98TAG	Figure 4 Figure S4	ATGGGCGGTGTTACTTTCCACACCAATCACGGCGA TATTGTCATCAAAACTTTTGACGATAAAGCACCTG AAACAGTTAAAAACTTCCTGGACTACTGCCGCGAA GGTTTTTACAACAACACCATTTTCCACCGTGTTAT CAACGGCTTTATGATTCAGGGCGGCGGTTTTGAAC CGGGCATGAAACAAAAAGCCACCAAAGAACCGATC AAAAACGAAGCCAACAACGGCCTGAAAAATACCCG TGGTACGCTGGCAATGGCACGTACTCAGGCTCCGC ACTCTGCAACTGCACAGTAGTTCATCAACGTGGTT GATAACGACTTCCTGAACTTCTCTGGCGAAAGCCT GCAAGGTTGGGGCTACTGCGTGTTTGCTGAAGTGG TTGACGGCATGGACGTGGTAGACAAAATCAAAGGT GTTGCAACCGGTCGTAGCGGTATGCATCAGGACGT GCCAAAAGAAGACGTTATCATTGAAAGCGTGACCG TTAGCGAGCACCACCATCATCACCACTAA	MGGVTFHTNHGDIVIKTFDDKAPETV KNFLDYCREGFYNNTIFHRVINGFMI QGGGFEPGMKQKATKEPIKNEANNGL KNTRGTLAMARTQAPHSATAQXFINV VDNDFLNFSGESLQGWGYCVFAEVVD GMDVVDKIKGVATGRSGMHQDVPKED VIIESVTVSEHHHHHH
FKBP-12 F48TAG	Figure 7 Figure S9	ATGGGTGTTCAGGTTGAAACCATTAGTCCTGGTGA TGGTCGTACCTTTCCGAAACGTGGTCAGACCTGTG TTGTTCATTACACCGGTATGCTGGAAGATGGCAAA AAGTTTGATAGCAGCCGTGATCGTAATAAGCCGTT TAAATAGGTTCTGGGTAAACAAGAAGTTATTCGCG GTTGGGAAGAGGGTGTTGCACAGATGAGCGTTGGT CAGCGTGCAAAACTGACCATTTCACCGGATTATGC ATATGGTGCAACCGGTCATCCGGGTATTATTCCGC CTAATGCAACCCTGATTTTTGATGTTGAACTGCTG AAACTGGAAGGTGGTAGCCATCACCATCATCATCA TTAA	MGVQVETISPGDGRTFPKRGQTCVVH YTGMLEDGKKFDSSRDRNKPFKXVLG KQEVIRGWEEGVAQMSVGQRAKLTIS PDYAYGATGHPGIIPPNATLIFDVEL LKLEGGSHHHHHH
FKBP-12 F99TAG	Figure 7	ATGGGTGTTCAGGTTGAAACCATTAGTCCTGGTGA TGGTCGTACCTTTCCGAAACGTGGTCAGACCTGTG TTGTTCATTACACCGGTATGCTGGAAGATGGCAAA AAGTTTGATAGCAGCCGTGATCGTAATAAGCCGTT TAAATTCGTTCTGGGTAAACAAGAAGTTATTCGCG GTTGGGAAGAGGGTGTTGCACAGATGAGCGTTGGT CAGCGTGCAAAACTGACCATTTCACCGGATTATGC ATATGGTGCAACCGGTCATCCGGGTATTATTCCGC CTAATGCAACCCTGATTTAGGATGTTGAACTGCTG AAACTGGAAGGTGGTAGCCATCACCATCATCATCA TTAA	MGVQVETISPGDGRTFPKRGQTCVVH YTGMLEDGKKFDSSRDRNKPFKFVLG KQEVIRGWEEGVAQMSVGQRAKLTIS PDYAYGATGHPGIIPPNATLIXDVEL LKLEGGSHHHHHH
RFP-Arc-NL-H_6	Figure S5 Figure S8 Table S5	ATGGCCAGTAGTGAAGACGTTATCAAGGAGTTTAT GCGTTTCAAAGTACGTATGGAGGGTAGTGTTAACG GACACGAATTTGAGATCGAGGGAGAGGGGGAAGGT CGTCCTTACGAGGGAACTCAAACGGCCAAATTAAA GGTGACCAAAGGTGGGCCCTTGCCATTCGCGTGGG ACATCTTGTCACCCCAGTTCCAGTACGGGTCGAAG GCATACGTAAAACACCCAGCGGACATTCCTGACTA	MASSEDVIKEFMRFKVRMEGSVNGHE FEIEGEGEGRPYEGTQTAKLKVTKGG PLPFAWDILSPQFQYGSKAYVKHPAD IPDYLKLSFPEGFKWERVMNFEDGGV VTVTQDSSLQDGEFIYKVKLRGTNFP SDGPVMQKKTMGWEASTERMYPEDGA LKGEIKMRLKLKDGGHYDAEVKTTYM

		TCTTAAGTTATCTTTCCCGGAAGGTTTTAAATGGG AACGCGTGATGAACTTTGAGGATGGGGGGGTTGTT ACGGTGACACAAGACTCCTCATTGCAAGATGGAGA GTTTATCTATAAAGTCAAACTTCGCGGCACCAATT TTCCATCTGACGGTCCTGTAATGCAGAAAAAAACA ATGGGCTGGGAAGCCTCCACAGAACGTATGTACCC CGAAGATGGAGCTTTAAAGGGCGAAATTAAAATGC GCTTAAAACTTAAAGACGGCGGCCATTACGACGCC GAAGTGAAAACGACGTATATGGCTAAGAAACCCGT CCAGCTTCCGGGAGCCTATAAAACTGACATCAAAC TGGATATTACATCACACAACGAAGATTATACTATT GTCGAACAGTACGAACGCGCCGAAGGCCGCCATTC AACGGGAGCAGGTGGATCAGAAAACCTGTACTTTC AAGGTGCGATGGGGCCGGGTGTGGACACGCAAATT TTTGAGGATCCCCGCGAATTTTTAAGTCACTTAGA GGAGTATCTTCGCCAAGTTGGCGGTTCAGAAGAGT ACTGGCTGTCTCAGATCCAGAACCACATGAACGGT CCTGCCAAAAAGTGGTGGGAGTTTAAGCAAGGATC CGTCAAAAACTGGGTCGAATTTAAGAAGGAGTTTT TACAGTACAGCGAGGGACACCACCACCACCACCAC TAA	AKKPVQLPGAYKTDIKLDITSHNEDY TIVEQYERAEGRHSTGAGGSENLYFQ GAMGPGVDTQIFEDPREFLSHLEEYL RQVGGSEEYWLSQIQNHMNGPAKKWW EFKQGSVKNWVEFKKEFLQYSEGHHH HHH
RFP-Arc-NL-H$_6$ F267TAG [c,d]	Figure 3 Figure S5 Table S3 Table S5	ATGGCCAGTAGTGAAGACGTTATCAAGGAGTTTAT GCGTTTCAAAGTACGTATGGAGGGTAGTGTTAACG GACACGAATTTGAGATCGAGGGAGAGGGGGAAGGT CGTCCTTACGAGGGAACTCAAACGGCCAAATTAAA GGTGACCAAAGGTGGGCCCTTGCCATTCGCGTGGG ACATCTTGTCACCCCAGTTCCAGTACGGGTCGAAG GCATACGTAAAACACCCAGCGGACATTCCTGACTA TCTTAAGTTATCTTTCCCGGAAGGTTTTAAATGGG AACGCGTGATGAACTTTGAGGATGGGGGGGTTGTT ACGGTGACACAAGACTCCTCATTGCAAGATGGAGA GTTTATCTATAAAGTCAAACTTCGCGGCACCAATT TTCCATCTGACGGTCCTGTAATGCAGAAAAAAACA ATGGGCTGGGAAGCCTCCACAGAACGTATGTACCC CGAAGATGGAGCTTTAAAGGGCGAAATTAAAATGC GCTTAAAACTTAAAGACGGCGGCCATTACGACGCC GAAGTGAAAACGACGTATATGGCTAAGAAACCCGT CCAGCTTCCGGGAGCCTATAAAACTGACATCAAAC TGGATATTACATCACACAACGAAGATTATACTATT GTCGAACAGTACGAACGCGCCGAAGGCCGCCATTC AACGGGAGCAGGTGGATCAGAAAACCTGTACTTTC AAGGTGCGATGGGGCCGGGTGTGGACACGCAAATT TTTGAGGATCCCCGCGAATTTTTAAGTCACTTAGA GGAGTATCTTCGCCAAGTTGGCGGTTCAGAAGAGT ACTGGCTGTCTCAGATCCAGAACCACATGAACGGT CCTGCCAAAAAGTGGTGGGAGTTTAAGCAAGGATC CGTCAAAAACTGGGTCGAATAGAAGAAGGAGTTTT TACAGTACAGCGAGGGACACCACCACCACCACCAC TAA	MASSEDVIKEFMRFKVRMEGSVNGHE FEIEGEGEGRPYEGTQTAKLKVTKGG PLPFAWDILSPQFQYGSKAYVKHPAD IPDYLKLSFPEGFKWERVMNFEDGGV VTVTQDSSLQDGEFIYKVKLRGTNFP SDGPVMQKKTMGWEASTERMYPEDGA LKGEIKMRLKLKDGGHYDAEVKTTYM AKKPVQLPGAYKTDIKLDITSHNEDY TIVEQYERAEGRHSTGAGGSENLYFQ GAMGPGVDTQIFEDPREFLSHLEEYL RQVGGSEEYWLSQIQNHMNGPAKKWW EFKQGSVKNWVEXKKEFLQYSEGHHH HHH
RFP-Arc-NL-H$_6$ F271TAG [c]	Figure S5 Figure S8 Table S5	ATGGCCAGTAGTGAAGACGTTATCAAGGAGTTTAT GCGTTTCAAAGTACGTATGGAGGGTAGTGTTAACG GACACGAATTTGAGATCGAGGGAGAGGGGGAAGGT CGTCCTTACGAGGGAACTCAAACGGCCAAATTAAA GGTGACCAAAGGTGGGCCCTTGCCATTCGCGTGGG ACATCTTGTCACCCCAGTTCCAGTACGGGTCGAAG GCATACGTAAAACACCCAGCGGACATTCCTGACTA TCTTAAGTTATCTTTCCCGGAAGGTTTTAAATGGG AACGCGTGATGAACTTTGAGGATGGGGGGGTTGTT ACGGTGACACAAGACTCCTCATTGCAAGATGGAGA GTTTATCTATAAAGTCAAACTTCGCGGCACCAATT TTCCATCTGACGGTCCTGTAATGCAGAAAAAAACA	MASSEDVIKEFMRFKVRMEGSVNGHE FEIEGEGEGRPYEGTQTAKLKVTKGG PLPFAWDILSPQFQYGSKAYVKHPAD IPDYLKLSFPEGFKWERVMNFEDGGV VTVTQDSSLQDGEFIYKVKLRGTNFP SDGPVMQKKTMGWEASTERMYPEDGA LKGEIKMRLKLKDGGHYDAEVKTTYM AKKPVQLPGAYKTDIKLDITSHNEDY TIVEQYERAEGRHSTGAGGSENLYFQ GAMGPGVDTQIFEDPREFLSHLEEYL RQVGGSEEYWLSQIQNHMNGPAKKWW

		ATGGGCTGGGAAGCCTCCACAGAACGTATGTACCC CGAAGATGGAGCTTTAAAGGGCGAAATTAAAATGC GCTTAAAACTTAAAGACGGCGGCCATTACGACGCC GAAGTGAAAACGACGTATATGGCTAAGAAACCCGT CCAGCTTCCGGGAGCCTATAAAACTGACATCAAAC TGGATATTACATCACACAACGAAGATTATACTATT GTCGAACAGTACGAACGCGCCGAAGGCCGCCATTC AACGGGAGCAGGTGGATCAGAAAACCTGTACTTTC AAGGTGCGATGGGGCCGGGTGTGGACACGCAAATT TTTGAGGATCCCCGCGAATTTTTAAGTCACTTAGA GGAGTATCTTCGCCAAGTTGGCGGTTCAGAAGAGT ACTGGCTGTCTCAGATCCAGAACCACATGAACGGT CCTGCCAAAAAGTGGTGGGAGTTTAAGCAAGGATC CGTCAAAAACTGGGTCGAATTTAAGAAGGAGTAGT TACAGTACAGCGAGGGACACCACCACCACCACCAC TAA	EFKQGSVKNWVEFKKEXLQYSEGHHH HHH
RFP-Arc-NL	Figure 6	ATGGCTTCTATGACCGGTCACCACCATCACCATCA CGGTGGTAGCGCCAGTAGTGAAGACGTTATCAAGG AGTTTATGCGTTTCAAAGTACGTATGGAGGGTAGT GTTAACGGACACGAATTTGAGATCGAGGGAGAGGG GGAAGGTCGTCCTTACGAGGGAACTCAAACGGCCA AATTAAAGGTGACCAAAGGTGGGCCCTTGCCATTC GCGTGGGACATCTTGTCACCCCAGTTCCAGTACGG GTCGAAGGCATACGTAAAACACCCAGCGGACATTC CTGACTATCTTAAGTTATCTTTCCCGGAAGGTTTT AAATGGGAACGCGTGATGAACTTTGAGGATGGGGG GGTTGTTACGGTGACACAAGACTCCTCATTGCAAG ATGGAGAGTTTATCTATAAAGTCAAACTTCGCGGC ACCAATTTTCCATCTGACGGTCCTGTAATGCAGAA AAAAACAATGGGCTGGGAAGCCTCCACAGAACGTA TGTACCCCGAAGATGGAGCTTTAAAGGGCGAAATT AAAATGCGCTTAAAACTTAAAGACGGCGGCCATTA CGACGCCGAAGTGAAAACGACGTATATGGCTAAGA AACCCGTCCAGCTTCCGGGAGCCTATAAAACTGAC ATCAAACTGGATATTACATCACACAACGAAGATTA TACTATTGTCGAACAGTACGAACGCGCCGAAGGCC GCCATTCAACGGGAGCAGGTGGATCAGAAAACCTG TACTTTCAAGGTGCGATGGGGCCGGGTGTGGACAC GCAAATTTTTGAGGATCCCCGCGAATTTTTAAGTC ACTTAGAGGAGTATCTTCGCCAAGTTGGCGGTTCA GAAGAGTACTGGCTGTCTCAGATCCAGAACCACAT GAACGGTCCTGCCAAAAAGTGGTGGGAGTTTAAGC AAGGATCCGTCAAAAACTGGGTCGAATTTAAGAAG GAGTTTTTACAGTACAGCGAGGGATAA	MASMTGHHHHHHGGSASSEDVIKEFM RFKVRMEGSVNGHEFEIEGEGEGRPY EGTQTAKLKVTKGGPLPFAWDILSPQ FQYGSKAYVKHPADIPDYLKLSFPEG FKWERVMNFEDGGVVTVTQDSSLQDG EFIYKVKLRGTNFPSDGPVMQKKTMG WEASTERMYPEDGALKGEIKMRLKLK DGGHYDAEVKTTYMAKKPVQLPGAYK TDIKLDITSHNEDYTIVEQYERAEGR HSTGAGGSENLYFQGAMGPGVDTQIF EDPREFLSHLEEYLRQVGGSEEYWLS QIQNHMNGPAKKWWEFKQGSVKNWVE FKKEFLQYSEG
H11-H_6 F70TAG	Figure 6	ATGGGCTCTGAAGTTCAGCTGCTGGAATCTGGTGG TGGTCTGGTTCAGGCGGGTGACTCTCTGCGTCTGT CTTGCGCGGCGTCTGGTCGTACCTTCTCTGCGTAC GCGATGGGTTGGTTCCGTCAGGCGCCGGGTAAAGA ACGTGAATTCGTTGCGGCGATCTCTTGGTCTGGTA ACTCTACCTACTACGCGGACTCTGTTAAAGGTCGT TAGACCATCTCTCGTGACAACGCGAAAAACACCGT TTACCTGCAGATGAACTCTCTGAAACCGGAGGACA CCGCGATCTACTACTGCGCGGCGCGTAAACCGATG TACCGTGTTGACATCTCTAAAGGTCAGAACTACGA CTACTGGGGTCAGGGTACCCAGGTTACCGTTTCTT CTCACCACCACCACCACCACTAA	MGSEVQLLESGGGLVQAGDSLRLSCA ASGRTFSAYAMGWFRQAPGKEREFVA AISWSGNSTYYADSVKGRXTISRDNA KNTVYLQMNSLKPEDTAIYYCAARKP MYRVDISKGQNYDYWGQGTQVTVSSH HHHHH

H11 [c,d]	Figure S6 Figure S8 Table S3 Table S5	ATGCATCATCATCATCACCACGGTGGTAGCGAAAA CCTGTACTTTCAAGGCTCTGAAGTTCAGCTGCTGG AATCTGGTGGTGGTCTGGTTCAGGCGGGTGACTCT CTGCGTCTGTCTTGCGCGGCGTCTGGTCGTACCTT CTCTGCGTACGCGATGGGTTGGTTCCGTCAGGCGC CGGGTAAAGAACGTGAATTCGTTGCGGCGATCTCT TGGTCTGGTAACTCTACCTACTACGCGGACTCTGT TAAAGGTCGTTTCACCATCTCTCGTGACAACGCGA AAAACACCGTTTACCTGCAGATGAACTCTCTGAAA CCGGAGGACACCGCGATCTACTACTGCGCGGCGCG TAAACCGATGTACCGTGTTGACATCTCTAAAGGTC AGAACTACGACTACTGGGGTCAGGGTACCCAGGTT ACCGTTTCTTCTTAA	MHHHHHHGGSENLYFQGSEVQLLESG GGLVQAGDSLRLSCAASGRTFSAYAM GWFRQAPGKEREFVAAISWSGNSTYY ADSVKGRFTISRDNAKNTVYLQMNSL KPEDTAIYYCAARKPMYRVDISKGQN YDYWGQGTQVTVSS
MBP F61TAG	Figure 7 Figure S9	ATGAAACATCACCATCACCATCACCCCATGAGCGA TTACGACATCCCCACTACTGAGAATCTTTATTTTC AGGGCCATATGAAAATCGAAGAAGGTAAACTGGTA ATCTGGATTAACGGCGATAAAGGCTATAACGGTCT CGCTGAAGTCGGTAAGAAATTCGAGAAAGATACCG GAATTAAAGTCACCGTTGAGCATCCGGATAAACTG GAAGAGAAATTCCCACAGGTTGCGGCAACTGGCGA TGGCCCTGACATTATCTAGTGGGCACACGACCGCT TTGGTGGCTACGCTCAATCTGGCCTGTTGGCTGAA ATCACCCCGGACAAAGCGTTCCAGGACAAGCTGTA TCCGTTTACCTGGGATGCCGTACGTTACAACGGCA AGCTGATTGCTTACCCGATCGCTGTTGAAGCGTTA TCGCTGATTTATAACAAAGATCTGCTGCCGAACCC GCCAAAAAACCTGGGAAGAGATCCCGGCGCTGGATA AAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATG TTCAACCTGCAAGAACCGTACTTCACCTGGCCGCT GATTGCTGCTGACGGGGGTTATGCGTTCAAGTATG AAAACGGCAAGTACGACATTAAAGACGTGGGCGTG GATAACGCTGGCGCGAAAGCGGGTCTGACCTTCCT GGTTGACCTGATTAAAAACAAACACATGAATGCAG ACACCGATTACTCCATCGCAGAAGCTGCCTTTAAT AAAGGCGAAACAGCGATGACCATCAACGGCCCGTG GGCATGGTCCAACATCGACACCAGCAAAGTGAATT ATGGTGTAACGGTACTGCCGACCTTCAAGGGTCAA CCATCCAAACCGTTCGTTGGCGTGCTGAGCGCAGG TATTAACGCCGCCAGTCCGAACAAAGAGCTGGCGA AAGAGTTCCTCGAAAACTATCTGCTGACTGATGAA GGTCTGGAAGCGGTTAATAAAGACAAACCGCTGGG TGCCGTAGCGCTGAAGTCTTACGAGGAAGAGTTGG CGAAAGATCCACGTATTGCCGCCACCATGGAAAAC GCCCAGAAAGGTGAAATCATGCCGAACATCCCGCA GATGTCCGCTTTCTGGTATGCCGTGCGTACTGCGG TGATCAACGCCGCCAGCGGTCGTCAGGCTGTCGAT GAAGCCCTGAAAGATGCACAGACTCGTATCACCAA GTAA	MKHHHHHHPMSDYDIPTTENLYF QGHMKIEEGKLVIWINGDKGYNG LAEVGKKFEKDTGIKVTVEHPDK LEEKFPQVAATGDGPDIIXWAHD RFGGYAQSGLLAEITPDKAFQDK LYPFTWDAVRYNGKLIAYPIAVE ALSLIYNKDLLPNPPKTWEEIPA LDKELKAKGKSALMFNLQEPYFT WPLIAADGGYAFKYENGKYDIKD VGVDNAGAKAGLTFLVDLIKNKH MNADTDYSIAEAAFNKGETAMTI NGPWAWSNIDTSKVNYGVTVLPT FKGQPSKPFVGVLSAGINAASPN KELAKEFLENYLLTDEGLEAVNK DKPLGAVALKSYEEELAKDPRIA ATMENAQKGEIMPNIPQMSAFWY AVRTAVINAASGRQAVDEALKDA QTRITK
MBP F258TAG	Figure 7 Figure S9	ATGAAACATCACCATCACCATCACCCCATGAGCGA TTACGACATCCCCACTACTGAGAATCTTTATTTTC AGGGCCATATGAAAATCGAAGAAGGTAAACTGGTA ATCTGGATTAACGGCGATAAAGGCTATAACGGTCT CGCTGAAGTCGGTAAGAAATTCGAGAAAGATACCG GAATTAAAGTCACCGTTGAGCATCCGGATAAACTG GAAGAGAAATTCCCACAGGTTGCGGCAACTGGCGA TGGCCCTGACATTATCTTCTGGGCACACGACCGCT TTGGTGGCTACGCTCAATCTGGCCTGTTGGCTGAA ATCACCCCGGACAAAGCGTTCCAGGACAAGCTGTA TCCGTTTACCTGGGATGCCGTACGTTACAACGGCA AGCTGATTGCTTACCCGATCGCTGTTGAAGCGTTA TCGCTGATTTATAACAAAGATCTGCTGCCGAACCC	MKHHHHHHPMSDYDIPTTENLYF QGHMKIEEGKLVIWINGDKGYNG LAEVGKKFEKDTGIKVTVEHPDK LEEKFPQVAATGDGPDIIFWAHD RFGGYAQSGLLAEITPDKAFQDK LYPFTWDAVRYNGKLIAYPIAVE ALSLIYNKDLLPNPPKTWEEIPA LDKELKAKGKSALMFNLQEPYFT WPLIAADGGYAFKYENGKYDIKD VGVDNAGAKAGLTFLVDLIKNKH MNADTDYSIAEAAFNKGETAMTI NGPWAWSNIDTSKVNYGVTVLPT FKGQPSKPXVGVLSAGINAASPN

		GCCAAAAACCTGGGAAGAGATCCCGGCGCTGGATA AAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATG TTCAACCTGCAAGAACCGTACTTCACCTGGCCGCT GATTGCTGCTGACGGGGGTTATGCGTTCAAGTATG AAAACGGCAAGTACGACATTAAAGACGTGGGCGTG GATAACGCTGGCGCGAAAGCGGGTCTGACCTTCCT GGTTGACCTGATTAAAAACAAACACATGAATGCAG ACACCGATTACTCCATCGCAGAAGCTGCCTTTAAT AAAGGCGAAACAGCGATGACCATCAACGGCCCGTG GGCATGGTCCAACATCGACACCAGCAAAGTGAATT ATGGTGTAACGGTACTGCCGACCTTCAAGGGTCAA CCATCCAAACCGTAGGTTGGCGTGCTGAGCGCAGG TATTAACGCCGCCAGTCCGAACAAAGAGCTGGCGA AAGAGTTCCTCGAAAACTATCTGCTGACTGATGAA GGTCTGGAAGCGGTTAATAAAGACAAACCGCTGGG TGCCGTAGCGCTGAAGTCTTACGAGGAAGAGTTGG CGAAAGATCCACGTATTGCCGCCACCATGGAAAAC GCCCAGAAAGGTGAAATCATGCCGAACATCCCGCA GATGTCCGCTTTCTGGTATGCCGTGCGTACTGCGG TGATCAACGCCGCCAGCGGTCGTCAGGCTGTCGAT GAAGCCCTGAAAGATGCACAGACTCGTATCACCAA GTAA	KELAKEFLENYLLTDEGLEAVNK DKPLGAVALKSYEEELAKDPRIA ATMENAQKGEIMPNIPQMSAFWY AVRTAVINAASGRQAVDEALKDA QTRITK
DENV F116TAG	Figure S10	ATGGCAGATCTGGAACTGGAACGTGCAGCAGATGT TCGTTGGGAAGAACAGGCAGAAATTAGCGGTAGCA GCCCGATTCTGAGCATTACCATTAGCGAAGATGGT AGCATGAGCATCAAAAACGAAGAAGAAGAACAGAC CCTGGGTGGAGGTGGCTCAGGAGGGGGTGGGGCCG GTGTTCTGTGGGATGTTCCGAGTCCGCCTCCGGTT GGTAAAGCAGAACTGGAAGATGGTGCCTATCGTAT TAAACAGAAAGGTATTCTGGGTTACAGCCAGATTG GTGCGGGTGTTTATAAAGAAGGCACCTTTCATACC ATGTGGCATGTTACCCGTGGTGCAGTTCTGATGCA TAAAGGTAAACGTATTGAACCGAGCTGGGCAGATG TTAAAAAGGATCTGATTAGCTATGGTGGTGGTTGG AAACTGGAAGGTGAATGGAAAGAAGGTGAAGAAGT TCAGGTTCTGGCCCTGGAACCGGGTAAAAATCCGC GTGCAGTTCAGACCAAACCGGGTCTGTAGAAAACC AATACCGGCACCATTGGTGCAGTGAGCCTGGATTT TAGTCCGGGTACAAGCGGTAGCCCGATTGTTGATA AAAAGGGTAAAGTTGTTGGCCTGTATGGTAATGGT GTTGTGACCCGTAGCGGTGCATATGTTAGCGCAAT TGCAAATACCGAAAAACATCACCACCATCATCACT AA	MADLELERAADVRWEEQAEISGSSP ILSITISEDGSMSIKNEEEEQTLGG GGSGGGGAGVLWDVPSPPPVGKAEL EDGAYRIKQKGILGYSQIGAGVYKE GTFHTMWHVTRGAVLMHKGKRIEPSW ADVKKDLISYGGGWKLEGEWKEGE EVQVLALEPGKNPRAVQTKPGLXKTN TGTIGAVSLDFSPGTSGSPIVDKK GKVVGLYGNGVVTRSGAYVSAIANTE KHHHHHH

[a] X indicates the position of the residue site-specifically replaced with F_5Phe (in the case of PpiB, also 4F-Phe).

[b] G1-F5F09 was used for all F_5Phe experiments except those of Figure S2.

[c] The underlined peptide indicates the section removed by TEV protease cleavage.

[d] Amino acid sequence used for protein crystallography.

Table S2. Yields of purified proteins in mg per liter of cell culture and PylRS variant used.

Protein	PylRS variant used	Yield (mg L^{-1} of cell culture)
PpiB F4/F_5Phe	G1-F5F09	3
NTGG-PpiB		67
NTGG-PpiB F4/F_5Phe	G1-F5F09	17[a]
NTGG-PpiB F27/F_5Phe	G1-F5F09	17[a]
NTGG-PpiB F98/F_5Phe	G1-F5F09	9[a]
NTGG-PpiB F4/4F-Phe	G1-4FF02	11
NTGG-PpiB F27/4F-Phe	G1-4FF02	12
NTGG-PpiB F98/4F-Phe	G1-4FF02	15
GB1 F30/F_5Phe	G1-F5F09	15
GB1 F52/F_5Phe	G1-F5F09	13
Arc-NL-H_6 F267/F_5Phe	G1-F5F09	4
Arc-NL-H_6 F271/F_5Phe	G1-F5F09	4.5
Arc-NL-H_6		55
Arc-NL		57
H11		>80
H11-H_6 F70/F_5Phe	G1-F5F09	10
FKBP-12 F48/F_5Phe	G1-F5F09	7
FKBP-12 F99/F_5Phe	G1-F5F09	6
DENV F116/F_5Phe	G1-F5F09	2.5
MBP F61/F_5Phe	G1-F5F09	2
MBP F258/F_5Phe	G1-F5F09	2

[a] Average yield of three independent expression experiments.

Table S3. Crystallographic table of statistics for PpiB containing F_5Phe in position 4 and the Arc-NL complex with H11 containing F_5Phe in position 267 of Arc-NL.

	PpiB F4/F_5Phe	Arc-NL F267/F_5Phe with H11
PDB ID	26AO	26AP
Data collection		
Space group	P 1	P 1 2_1 1
Cell dimensions		
a, b, c (Å)	34.91 39.20 61.62	40.84 48.37 43.71
α, β, γ (°)	78.04 79.29 67.01	90.00 96.57 90.00
Resolution (Å)*	35.68 – 1.15 (1.17 – 1.15)	28.08 – 1.50 (1.53 – 1.50)
R_{merge}	0.077 (1.040)	0.084 (0.742)
R_{pim}	0.053 (0.689)	0.035 (0.306)
$I/\sigma I$	7.7 (1.0)	12.0 (2.5)
$CC_{1/2}$	0.962 (0.470)	0.999 (0.745)
Completeness (%)	94.6 (89.3)	99.8 (99.0)
Redundancy	3.5 (3.3)	6.8 (6.8)
Refinement		
Resolution (Å)	31.92 – 1.15 (1.16 – 1.15)	28.08 – 1.50 (1.56 – 1.50)
No. reflections	98100 (9370)	26954 (2645)
R_{work}/R_{free}	0.1973/0.2260 (0.3038/0.3300)	0.1601/0.1937 (0.2070/0.2528)
No. atoms	2925	2011
Protein	2705	1737
Ligand/ion	0	35
Water	220	239
B-factors (overall)	22.04	19.65
Protein	21.88	18.13
Ligand/ion	-	30.05
Water	24.04	29.24
R.m.s. deviations		
Bond lengths (Å)	0.009	0.006
Bond angles (°)	1.136	0.870

*Values in parentheses are for the highest resolution shell.

Table S4. Parameters of 2D [^{19}F,^{19}F]-NOESY spectra

Protein and site of F_5Phe residue	Concentration (mM)	Figure	Spectrometer (MHz)[a]	t_{1max} (ms)	t_{2max} (ms)
PpiB F4	1.6	4c	400	1.9	109
PpiB F27	2.1	4c	400	1.9	109
PpiB F98	1.0	4c	400	1.3	68
Arc-NL F267 + H11	0.35 (H11 in 1.4-fold excess)	6d	600[a]	1.1	91
Arc-NL F267 at high pressure	1.5	9a	600	2.7	10
Arc-NL F267 + H11 at high pressure	1.5 (H11 in 1.2-fold excess)	9b	600	2.0	10
GB1 F30	0.4	5b	400	1.4	91
GB1 F52	0.8	5c	400	2.8	91
FKBP-12 F48	0.35	S8a	400	6.3	143
FKBP-12 F48 + rapamycin	0.35	S8a	400	1.4	143
MBP F258	0.33	S8b	500[a]	0.9	110
MBP F258 + maltose	0.33	S8b	500[a]	0.4	110

[a] ^{1}H NMR frequency

[b] Equipped with cryoprobe

Table S5. BLI kinetic and affinity data from two independent experiments showing the mean and range of dissociation constants, on-rates and off-rates of wild-type Arc-NL and F_5Phe mutants binding to H11.

				Exp. 1		Exp. 2	
Protein	K_D (nM)[a]	k_{on} (x 10^5 $M^{-1}s^{-1}$) [a]	k_{off} (x 10^{-5} s^{-1}) [a]	χ^2	R^2	χ^2	R^2
wild-type	3.49 (± 0.40)	1.83 (± 0.11)	63.6 (± 3.1)	0.58	1.00	0.83	1.00
F267/F_5Phe	3.71 (± 0.34)	2.03 (± 0.06)	75.4 (± 9.0)	0.88	1.00	0.77	1.00
F271/F_5Phe	3.13 (± 0.53)	1.99 (± 0.20)	63 (± 17)	1.51	1.00	0.59	1.00

[a] Average and range of two independent experiments.

References

(1) Moore, S. J.; Lai, H.-E.; Kelwick, R. J. R.; Chee, S. M.; Bell, D. J.; Polizzi, K. M.; Freemont, P. S. EcoFlex: a multifunctional MoClo kit for *E. coli* synthetic biology. *ACS Synth. Biol.* **2016**, *5*, 1059–1069.

(2) Shis, D. L.; Bennett, M. R. Library of synthetic transcriptional AND gates built with split T7 RNA polymerase mutants. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 5028–5033.

(3) 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*, 9699.

(4) Qianzhu, H.; Tan, Y. J.; Abdelkader, E. H.; Huber, T.; Otting, G. Genetic encoding of fluorinated analogues of phenylalanine for ^{19}F NMR spectroscopy: detection of conformational heterogeneity in flaviviral NS2B-NS3 proteases. *ACS Sens*. **2025**, *10*, 3152–3161.

(5) Galles, G. D.; Infield, D. T.; Clark, C. J.; Hemshorn, M. L.; Manikandan, S.; Fazan, F.; Rasouli, A.; Tajkhorshid, E.; Galpin, J. D.; Cooley, R. B.; Mehl, R. A.; Ahern, C. A. Tuning phenylalanine fluorination to assess aromatic contributions to protein function and stability in cells. *Nat. Commun*. **2023**, *14*, 59.

(6) Lee, Y.-J.; Schmidt, M. J.; Tharp, J. M.; Weber, A.; Koenig, A. L.; Zheng, H.; Gao, J.; Waters, M. L.; Summerer, D.; Liu, W. R. Genetically encoded fluorophenylalanines enable insights into the recognition of lysine trimethylation by an epigenetic reader. *Chem. Commun*. **2016**, *52*, 12606–12609.

(7) Zhang, Q.; Jiang, L.; Niu, Y.; Li, Y.; Chen, W.; Cheng, J.; Ding, H.; Chen, B.; Liu, K.; Cao, J.; Wang, J.; Ye, S.; Yang, L.; Wu, J.; Xu, G.; Lin, J.; Yu, H. Machine learning-guided evolution of pyrrolysyl-tRNA synthetase for improved incorporation efficiency of diverse noncanonical amino acids. *Nat. Commun.* **2025**, *16*, 6648–6648.

(8) Frkic, R. L.; Tan, Y. J.; Abdelkader, E. H.; Maleckis, A.; Tarcoveanu, E.; Nitsche, C.; Otting, G.; Jackson, C. J. Conformational preferences of the non-canonical amino acids (2*S*,4*S*)-5-fluoroleucine, (2*S*,4*R*)-5-fluoroleucine, and 5,5'-difluoroleucine in a protein. *Biochemistry* **2024**, *63*, 1388–1394.

(9) Frkic, R. L.; Tan, Y. J.; Maleckis, A.; Chilton, N. F.; Otting, G.; Jackson, C. J. 1.3 Å crystal structure of *E. coli* peptidyl–prolyl isomerase B with uniform substitution of valine by (2*S*,3*S*)-4-fluorovaline reveals structure conservation and multiple staggered rotamers of CH_2F groups. *Biochemistry* **2024**, *63*, 2602–2608.

(10) Rosenau, C. P.; Jelier, B. J.; Gossert, A. D.; Togni, A. Exposing the origins of irreproducibility in fluorine NMR spectroscopy. *Angew. Chem. Int. Ed*. **2018**, *57*, 9528–9533.

(11) Ernst, R. R.; Bodenhausen, G.; Wokaun, A. Principles of nuclear magnetic resonance in one and two dimensions. Oxford University Press, **1987**.

(12) Kohn, W.; Sham, L. J. Self-consistent equations including exchange and correlation effects. *Phys. Rev*. **1965**, *140*, A1133–A1138.

(13) Neese, F.; Wennmohs, F.; Becker, U.; Riplinger, C. The ORCA quantum chemistry program package. *J. Chem. Phys*. **2020**, *152*, 224108.

(14) Chai, J.-D.; Head-Gordon, M. Systematic optimization of long-range corrected hybrid density functionals. *J. Chem. Phys*. **2008**, *128*, 084106.

(15) Becke, A. D. Density-functional thermochemistry. V. Systematic optimization of exchange-correlation functionals. *J. Chem. Phys*. **1997**, *107*, 8554–8560.

(16) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density dispersion correction (DFT-D) for the 94 elements H–Pu. *J. Chem. Phys*. **2010**, *132*, 154104.

(17) Johnson, E. R.; Becke, A. D. A post-Hartree-Fock model of intermolecular interactions: Inclusion of higher-order corrections. *J. Chem. Phys*. **2006**, *124*, 174104.

(18) Dunning Jr., T. H. Gaussian basis sets foruse in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys*. **1989**, *90*, 1007–1023.

(19) Helmich-Paris, B.; de Souza, B.; Neese, F.; Izsák, R. An improved chain of spheres for exchange algorithm. *J. Chem. Phys*. **2021**, *155*, 104109.

(20) Winter, G. Xia2: an expert system for macromolecular crystallography data reduction. *J. Appl. Cryst*. **2010**, *43*, 186–190.

(21) Evans, P. R.; Murshudov, G. N. How good are my data and what is the resolution? *Acta Cryst. D* **2013,** *69*, 1204–1214.

(22) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. Phaser crystallographic software. *J. Appl. Crystallogr.* **2007,** *40*, 658–674.

(23) Edwards, K. J.; Ollis, D. L.; Dixon, N. E. Crystal structure of cytoplasmic *Escherichia coli* peptidyl-prolyl isomerase: evidence for decreased mobility of loops upon complexation. *J. Mol. Biol*. **1997**, *271*, 258–265.

(24) Emsley, P.; Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Cryst. D* **2004,** *60*, 2126–2132.

(25) Zwart, P. H.; Afonine, P. V.; Grosse-Kunstleve, R. W.; Hung, L. W.; Ioerger, T. R.; McCoy, A. J.; McKee, E.; Moriarty, N. W.; Read, R. J.; Sacchettini, J. C.; Sauter, N. K.; Storoni, L. C.; Terwilliger, T. C.; Adams, P. D. Automated structure solution with the PHENIX suite. *Methods Mol. Biol.* **2008,** *426*, 419–435.

(26) Williams, C. J.; Headd, J. J.; Moriarty, N. W.; Prisant, M. G.; Videau, L. L.; Deis, L. N.; Verma, V.; Keedy, D. A.; Hintze, B. J.; Chen, V. B.; Jain, S.; Lewis, S. M.; Arendall, W. B. 3rd; Snoeyink, J.; Adams, P.D.; Lovell, S. C.; Richardson, J. S.; Richardson, D. C. MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci*. **2018**, *27*, 293–315.

(27) Kabsch, W. *XDS*. *Acta Crystallogr. D. Biol. Crystallogr.* **2010**, *66*, 125–132.

(28) Godoy Muñoz, J. M.; Neset, L.; Markússon, S.; Weber, S.; Krokengen, O. C.; Sutinen, A.; Christakou, E.; Lopez, A. J.; Bramham, C. R.; Kursula, P. Structural characterization of two nanobodies targeting the ligand-binding pocket of human Arc. *PLOS ONE* **2024**, *19*, e0300453.

(29) Murshudov, G. N.; Skubák, P.; Lebedev, A. A.; Pannu, N. S.; Steiner, R. A.; Nicholls, R. A.; Winn, M. D.; Long, F.; Vagin, A. A. *REFMAC*5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D. Biol. Crystallogr.* **2011**, *67*, 355–367.

(30) Geoghegan, K. F.; Dixon, H. B. F.; Rosner, P. J.; Hoth, L. R.; Lanzetti, A. J.; Borzilleri, K. A.; Marr, E. S.; Pezzullo, L. H.; Martin, L. B.; LeMotte, P. K.; McColl, A. S.; Kamath, A. V.; Stroh, J. G. Spontaneous α-*N*-6-phosphogluconoylation of a "His tag" in *Escherichia coli*: the cause of extra mass of 258 or 178 Da in fusion proteins. *Anal. Biochem.* **1999**, 267, 169–184.

(31) Yan, Z.; Caldwell, G. W.; McDonell, P. A.; Jones, W. J.; August, A.; Masucci, J. A. Mass spectrometry determination of a novel modification of the N-terminus of histidine-tagged proteins expressed in bacteria. *Biochem. Biophys. Res. Commun*. **1999**, *259*, 271–282.

(32) Tan, Y. J.; Abdelkader, E. H.; Herath, I. D.; Maleckis, A.; Otting, G. Inter-residue through-space scalar $^{19}F–^{19}F$ couplings between CH_2F groups in a protein. *Magn. Reson*. **2025**, *6*, 131–142.

(33) Aleshin, A. E.; Shiryaev, S. A.; Strongin, A. Y.; Liddington, R. C. Structural evidence for regulation and specificity of flaviviral proteases and evolution of the *Flaviviridae* fold. *Protein Sci*. **2007**, *16*, 795–806.

(34) Erbel, P.; Schiering, N.; D'Arcy, A.; Renatus, M.; Kroemer, M.; Lim, S. P.; Yin, Z.; Keller, T. H.; Vasudevan, S. G.; Hommel, U. Structural basis for the activation of flaviviral NS3 proteases from dengue and West Nile virus. *Nat. Struct. Mol. Biol*. **2006**, *13*, 372–373.

(35) Lilkova, E. The PyMOL Molecular Graphics System, Version 2.0; Schrödinger, LLC, **2015**.

(36) Frericks Schmidt, H. L.; Sperling, L. J.; Gui Gao, Y.; Wylie, B. J.; Boettcher, J. M.; Wilson, S. R.; Rienstra, C. M. Crystal polymorphism of protein GB1 examined by solid-state NMR spectroscopy and X-ray diffraction. *J. Phys. Chem. B* **2007**, 111, 14362–14369.

(37) Van Duyne, G. D.; Standaert, R. F.; Schreiber, S. L.; Clardy, J. Atomic structure of the rapamycin human immunophilin FKBP-12 complex. *J. Am. Chem. Soc*. **1991**, *113*, 7433–7434.

(38) Sharff, A. J.; Rodseth, L. E.; Spurlino, J. C.; Quiocho, F. A. Crystallographic evidence of a large ligand-induced hinge-twist motion between the two domains of the maltodextrin binding protein involved in active transport and chemotaxis. *Biochemistry* **1992**, *31*, 10657–10663.

(39) Quiocho, F. A.; Spurlino, J. C.; Rodseth, L. E. Extensive features of tight oligosaccharide binding revealed in high-resolution structures of the maltodextrin transport/chemosensory receptor. *Structure* **1997**, *5*, 997–1015.