Chapter IV

INCORPORATION OF PHOTO-PROLINE INTO PROTEINS IN *ESCHERICHIA COLI*

4.1 Contributions

Mona Shahgholi performed mass spectrometry of proinsulin peptides to determine photo-proline incorporation efficiencies. Jeff Jones, Brett Lomenick, Ting Yu Wang, and Tsui-Fen Chou assisted with determining photo-pro photoconversion by mass spectrometry. Edgar Manriquez-Sandoval and Anneliese Faustino performed PPIase enrichments, fractionation, and mass spectrometry, with input from Stephen D. Fried. Stephanie L. Breunig prepared all plasmids, performed all other experiments, participated in proteomics data analysis, and analyzed all other data. We thank Gracie Wang for assistance with UV irradiation.

4.2 Abstract

Photo-crosslinking non-canonical amino acids contain light-activated side chains that form covalent crosslinks to nearby molecules upon irradiation. Diverse non-canonical residues carrying photo-active moieties have characterized interactions between proteins and other biomolecules. Here, we report the ribosomal incorporation of photo-proline, a diazirine-containing proline analog, into recombinantly-expressed proinsulin in *Escherichia coli*. We found that up to 74% of proline residues could be replaced with photo-proline, and sought to leverage this photo-crosslinking amino acid to detect protein-protein interactions at proline positions in live cells. However, initial efforts did not detect crosslinking between

E. coli peptidyl-prolyl isomerases and photo-proline – containing client proteins. This chapter expands the photo-crosslinking non-canonical amino acids available for incorporation into proteins, and informs future efforts to apply photo-proline in studying protein-protein interactions.

4.3 Introduction

Protein-protein interactions (PPIs) are involved in a plethora of cellular processes and involve a wide range of protein interfaces and binding affinities.[1] One method among many[2] for studying protein-protein interaction partners involves photo-crosslinking non-canonical amino acids (ncAAs).[3] These residues include a light-activatable moiety that, upon irradiation, generates a reactive intermediate capable of forming covalent bonds with nearby molecules. Photo-crosslinking ncAAs have been used, for instance, to identify protein-protein interactions,[4] map interaction sites,[5] and generate cross-linked protein biomaterials.[6]

Photo-crosslinking ncAAs are often incorporated into proteins in a site-specific manner. In these cases, the translational machinery is engineered such that a particular codon (traditionally the amber stop codon) instead encodes the ncAA of interest. *p*-Benzoylphenylalanine (Bpa) was first introduced as a crosslinking ncAA in a chemically synthesized peptide in 1986,[7] then incorporated site-specifically into proteins using an engineered *Methanococcus jannaschii* tyrosyl-tRNA synthetase/tRNA pair in 2002.[8,9] Other photo-crosslinking amino acids incorporated by nonsense suppression include *p*-azidophenylalanine (pAzF),[10] 4'-[3-(trifluoromethyl)-3H-diazirin-3-yl]-l-phenylalanine

(TfmdPhe),[11] 3′-azibutyl-N-carbamoyl-lysine (AbK),[12] and 3-(3-methyl-3H-diazirine-3-yl)-propaminocarbonyl-Nε-L-lysine (DiZPK).[4]

Residue-specific replacement approaches have similarly led to the incorporation of photo-crosslinking ncAAs into proteins. These techniques rely upon promiscuity inherent to the endogenous translational machinery of the host organism: the amino acid analog replaces its canonical parent across the proteome. In particular, the diazirine-containing ncAAs photo-methionine, photo-leucine, and photo-lysine are accepted by the translational machinery of mammalian cells, and have been used to detect protein-protein interaction partners.[13,14] Both photo-leucine and photo-methionine can also label proteins expressed in *E. coli*.[15,16]

A similar photo-crosslinking amino acid derivative is photo-proline (photo-pro, Figure 4.1a), which includes a diazirine functionality at the C^{γ} position of the pyrrolidine side chain. Photo-pro has been introduced into chemically-synthesized peptides[17–20] and small molecule inhibitors,[21] and was first used to demonstrate the interaction between a peptidomimetic antibiotic and its target LptD.[17] However, its incorporation into ribosomally-expressed proteins has not yet been reported.

Here, we demonstrate the incorporation of photo-pro into proteins expressed in *E. coli*. We show that photo-pro incorporation is enabled by high osmotic stress conditions or *proP* overexpression. We further describe our efforts to detect the substrates of peptidyl-prolyl isomerases (PPIases) in *E. coli* via photo-pro crosslinking; however, initial experiments

have not identified PPIase client proteins. Together, this chapter facilitates the production of recombinant proteins that include photo-pro residues, and may assist future efforts to apply this residue in studying protein-protein interactions *in vivo*.

4.4 Results and discussion

4.4.1 Incorporation of photo-proline

In Chapter II of this thesis, we described our efforts to identify an expanded set of proline analogs that could be incorporated into proteins expressed in *E. coli* (Figure 2.S1, Table 2.S1). We found that the non-canonical proline (ncPro) analog photo-pro could replace proline residues in recombinant proinsulin with an incorporation efficiency of 37% (Table 2.S1); *proS* overexpression and 0.3 M NaCl were used promote ncPro incorporation. In these initial incorporation studies, we also observed by mass spectrometry the presence of an ion corresponding to the loss of two protons from the proline-containing peptide (Figure 4.1b-c). A degradation product of photo-proline is 3,4-dehydroproline (dhp, Figure 4.1a, Figure 4.S1), so we concluded that this ion likely corresponds to the dhp-containing peptide. Because of its potential utility in photo-crosslinking studies, we sought to characterize and improve photo-pro incorporation.

The M157Q and C443G mutants of the *E. coli* prolyl-tRNA synthetase (ProRS) increase the incorporation of 4*S*-aminoproline and piperidine-2-carboxylic acid, respectively.[22,23] However, overexpression of either mutant did not enhance photo-pro incorporation (Figure 4.1d, Table 4.1). We then screened photo-pro and NaCl concentrations for improved incorporation into proinsulin. Because proline is an osmoprotectant, high osmotic stress

conditions (such as high salt concentrations) increase expression of proline transporters[24] that enhance non-canonical proline (ncPro) uptake and incorporation.[23] Increasing the concentrations of both photo-pro and NaCl increased incorporation (Figure 4.1e-f, Table 4.1): 48% of peptides contained photo-pro when we increased the concentration of the ncPro to 1.0 mM (0.3 M NaCl); 0.5 M NaCl (0.3 mM photo-pro) increased incorporation further to 61%. No incorporation of photo-pro could be detected without added NaCl, highlighting the importance of amino acid uptake.

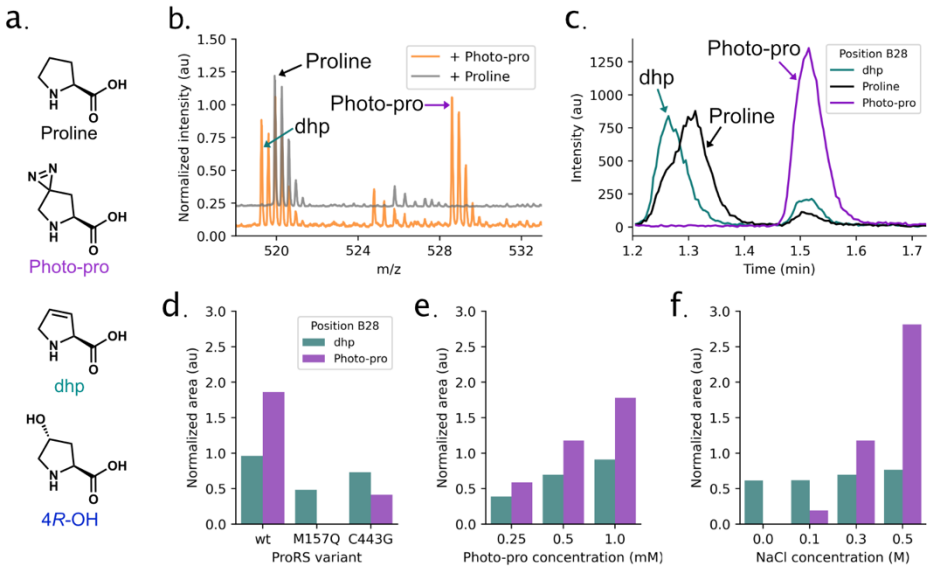


Figure 4.1. Improved incorporation of photo-pro. a. The chemical structures of proline and proline analogs described in this chapter. **b**. Mass spectra of the proinsulin peptide RGFFYT**P**KTRRE (obtained by Glu-C digestion; $[M+3H]^{+3}$ ion) after proinsulin expression in the presence of proline or photo-proline. **c**. Extracted ion chromatograms (EIC; m/z = 520.3, proline; 519.3, dhp; 528.3, photo-pro; each corresponds to the $[M+3H]^{+3}$ ion) of the +photo-pro sample shown in (b). **d-f**. Photo-pro and dhp incorporation vary across ProRS mutants (0.5 mM photo-pro, 0.3 M NaCl; d), photo-pro concentration (wild-type ProRS, 0.3 M NaCl; e), and NaCl concentration (wild-type ProRS, 0.5 mM photo-pro; f). Shown is the area under the curve corresponding to each ncPro peptide from the EIC, normalized to the proline-containing peptide in that sample.

Interestingly, these results also suggest that an appreciable level of dhp-containing peptide is the result of co-translational dhp incorporation, rather than photolysis after photo-pro incorporation. We observed the dhp peptide across all ProRS variants and NaCl concentrations tested, independent of photo-pro incorporation (Figure 4.1d,f). Its prevalence also increased with photo-pro concentration (Figure 4.1e), presumably due to the unavoidable correlation between concentrations of photo-pro and its thermal degradation products under these culturing conditions. We note that dhp is known to incorporate well into recombinant proteins in *E. coli*.[23] The low (~5%) levels of dhp produced due to thermal degradation over the course of a photo-pro labeling experiment (Figure 4.S1) appear to compete with photo-pro for incorporation at proline codons.

Table 4.1. Photo-pro incorporation efficiencies under osmotic stress conditions

Protein	ProRS	Photo-pro conc. (mM)	NaCl conc. (M)	Fraction proline	Fraction photo-pro	Fraction dhp
6xHis-PI	wt	0.5	0.3	0.26	0.49	0.25
6xHis-PI	M157Q	0.5	0.3	0.67	n.d.*	0.33
6xHis-PI	C443G	0.5	0.3	0.47	0.19	0.34
H27R-PI	wt	0.25	0.3	0.51	0.30	0.20
H27R-PI	wt	0.5	0.3	0.35	0.41	0.24
H27R-PI	wt	1	0.3	0.27	0.48	0.25
H27R-PI	wt	0.5	0	0.62	n.d.*	0.38
H27R-PI	wt	0.5	0.1	0.55	0.11	0.34
H27R-PI	wt	0.5	0.5	0.22	0.61	0.17

*n.d., not detected.

4.4.2 ProP overexpression leads to ncPro incorporation

High osmotic stress conditions necessary for photo-pro incorporation might interfere with *E. coli* cell physiology,[25] limiting the use of photo-pro to study protein-protein interactions

in vivo. We therefore pursued an alternative approach for photo-pro incorporation, namely the direct overexpression of the proline transporter *proP*[26] (Figure 4.2a-b). We found that expressing *proP* in proline medium, then inducing proinsulin expression after a medium shift into photo-pro medium, led to 73% photo-pro incorporation into proinsulin even in the absence of osmotic stress conditions. Using the *proP* overexpression approach, we obtained a 91% incorporation efficiency for the proline analog 4*R*-hydroxyproline (4*R*-OH, Figure 4.1a). These results are nearly identical to 4*R*-OH incorporation using osmotic stress conditions,[27] suggesting that *proP* overexpression is a general approach for ncPro uptake and incorporation (Figure 4.2c-d, Table 4.2).

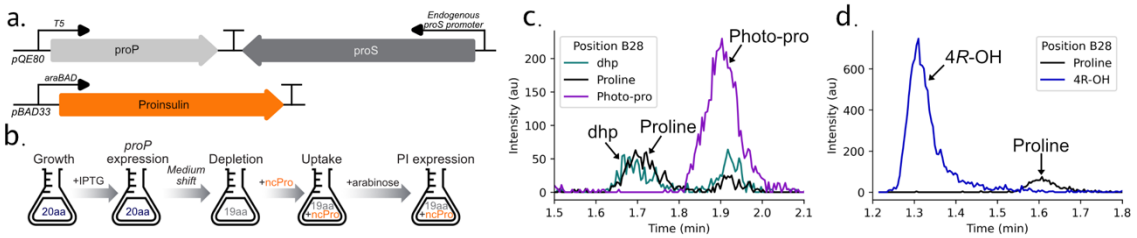


Figure 4.2. *proP* overexpression leads to proline analog incorporation. a-b. The constructs (a) and expression protocol (b) used for ncPro incorporation into recombinantly-expressed proinsulin by *proP* overexpression. **c-d**. Extracted ion chromatograms after photo-pro (c) and 4*R*-OH (d) incorporation. Proinsulins containing ncPro residues were digested with Glu-C and analyzed by LC-ESI-MS. The $[M+3H]^{+3}$ ion corresponding to the peptide RGFFYT**P**KTRRE was monitored (m/z = 520.3, proline; 519.3, dhp; 528.3, photo-pro; 525.3, 4*R*-OH). Dhp and proline ions that co-eluted with the photo-pro ion (~1.9 min) were presumed to be degradation products of photo-proline formed during mass spectrometry.

Table 4.2. Proline analog incorporation efficiencies with *proP* overexpression

ncPro	ncPro conc. (mM)	Fraction proline	Fraction photo-pro	Fraction dhp	Fraction 4*R*-OH
Photo-pro	1.0	0.135	0.743	0.122	–
4*R*-OH	1.0	0.095	–	–	0.905

4.4.3 PPIase substrates were not detected with photo-proline

Peptidyl-prolyl isomerases (PPIases) are chaperones that promote amide *cis-trans* isomerization of proline residues. We wondered if the crosslinking capabilities of photo-pro could be leveraged to identify PPIase substrates in *E. coli*. Here, photo-pro was incorporated proteome-wide using the conditions that led to 74% incorporation into recombinant proinsulin (Figure 4.2), and samples irradiated to promote crosslinking at photo-pro positions. 3xFLAG-tagged PPIases, which was overexpressed (along with *proP*) in proline-containing medium prior to photo-pro labeling, was visualized by immunoblotting. We expected that client proteins could be enriched and identified by immunoprecipitation and mass spectrometry.

We were unable to detect the presence higher molecular weight species indicative of successful covalent crosslinking between a PPIase and its substrate by western blot (Figure 4.3a). Mass spectrometry was similarly unsuccessful: after PPIase immunoprecipitation, we did not observe an increase in the number of client proteins identified in the photo-pro treated sample, compared with the proline control (Figure 4.3b-c). Further, we could not confidently identify PPIase:client crosslinked peptides by mass spectrometry (data not shown).

Control photolysis experiments of solubilized photo-pro–proinsulin indicated 69-92% photoconversion by mass spectrometry after UV irradiation under similar conditions. We attempted to detect crosslinking to free, protected amino acids (N-boc and methyl ester

protected derivatives of cysteine, tyrosine, and glutamine); however, crosslinking at photo-pro positions was not identified (data not shown).

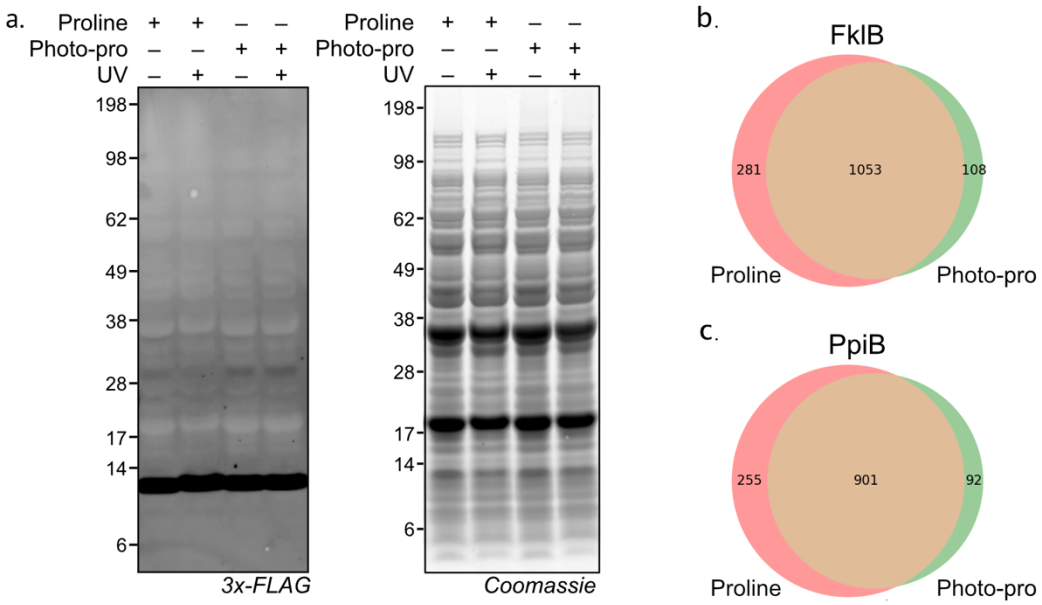


Figure 4.3. PPIase:client interactions were not detected using photo-proline. **a**. Western blot and SDS-PAGE analysis to detect PpiC crosslinked to its substrates. PpiC-3xFLAG and ProP were expressed in proline medium, then shifted to medium that instead contained photo-pro. Cells were irradiated (365 nm, 15 min), and proteins analyzed by SDS-PAGE. PpiC was detected using the 3xFLAG epitope; proline and –UV conditions were used as negative controls. No higher molecular weight species could be detected in +photo-pro/+UV conditions compared to the controls. Similar results were obtained for the eight other PPIases tested (not shown). **b, c**. FklB-3xFLAG (b) and PpiB-3xFLAG (c) were expressed, possible substrates labeled with photo-pro, and cells irradiated as described in (a). Each PPIase, potentially cross-linked to its substrates by photo-pro, was enriched by immunoprecipitation. Samples were analyzed by bottom-up proteomics. We did not observe an increase in the number of hits for the samples labeled with photo-pro, compared to those labeled with proline.

Dissociation constants for PPIase:substrate interactions have been reported in the low micromolar to upper nanomolar range (for example, Ref. 28–30). This interaction is

perhaps too weak for sufficient crosslinking with photo-pro; we do note that successful photo-pro labeling was detected between two components with a binding constant in the tens of nanomolar range.[20] Perhaps carbene stabilization (for instance, by installing fluoro substituents about the pyrrolidine ring) would better promote cross-linking and substrate identification. The diazirine moiety, while small, might interrupt the interaction with the PPIase, or may interfere with client protein folding, solubility, and stability. Finally, optimized photolysis conditions, especially in the context of the cellular environment of *E. coli*, may improve crosslinking efficiency.

4.5 Conclusion

In conclusion, we have detailed the incorporation of the proline analog photo-pro into proteins expressed in *E. coli*. We find that photo-pro incorporation requires high osmotic stress conditions or *proP* overexpression. Despite obtaining good incorporation efficiencies in recombinantly-expressed proinsulin, we have thus far been unable to use photo-pro to identify PPIase substrates *in vivo*. Together, this work enables the production of recombinant proteins containing photo-pro at proline positions, and informs future efforts to use this residue in detecting protein-protein interactions *in vivo*.

4.6 Materials and methods

4.6.1 Chemicals

All chemicals were purchased from MilliporeSigma.

4.6.2 Enzymes

2x Q5 master mix (New England Biolabs) was used for all PCR applications. Gibson assembly was performed with the Repliqa HiFi assembly mix from Quantabio.

4.6.3 Strains and plasmids

The proline-auxotrophic *E. coli* strain CAG18515 was obtained from the Coli Genetic Stock Center (CGSC) at Yale University. Strain DH10B was used for standard cloning operations; electrocompetent CAG18515 were transformed with purified plasmid products.

Plasmids pQE80PI-proS (which encodes 6xHis-PI and proS) and pQE80_H27R-PI_proS (which encodes H27R-PI and *proS*) were described in Chapter II of this thesis and references therein. Installation of the M157Q and C443G mutations into *proS* were also described in Chapter II.

4.6.4 Primers

AmpR-GA_fwd: GCAGTGTTATCACTCATGG
AmpR-GA_rev: GACAGTAAGAGAATTATGCAGTG
SLB7008_vec_rev: AGTTAATTTCTCCTCTTTAATGAATTC
SLB7008_vec_fwd: GGATCCGCATGCGAG

SLB7008_ins_fwd: GAATTCATTAAAGAGGAGAAATTAACTATGCTGAAAAGGA
AAAAAGTAAAAC

SLB7008_ins_rev: CTCGCATGCGGATCCTTATTCATCAATTCGCGGATG

SLB7009_vec_fwd: GGTACCCGGGGATCCTCTAG

SLB7009_vec_rev: GAGCTCGAATTCGCTAGCCC

4.6.5 Nucleotide and amino acid sequences

The sequences of H27R-PI and *proS* are described in Chapter II of this thesis.

6xHis-PI:

ATGAGAGGATCGCATCACCATCACCATCACCGCTTTGTGAACCAGCACCTGT
GCGGTAGCCACCTGGTGGAAGCTCTGTACCTGGTTTGCGGTGAGCGTGGTTTC
TTCTACACGCCaAAGACCCGCCGTGAAGCTGAAGATCTGCAGGTGGGCCAGG
TAGAACTGGGCGGTGGTCCGGGTGCCGGCTCTCTGCAACCGCTGGCACTGGA
AGGTTCCCTGCAAGCGCGTGGTATCGTAGAGCAGTGCTGTACTTCTATCTGCT
CCCTGTACCAGCTGGAGAACTACTGTAATTAA

The sequence of the 6xHis leader peptide is underlined, proinsulin is in **bold**. The A-chain and B-chain in mature insulin are colored **red** and **blue**, respectively.

MRGSHHHHHHR**FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQV**
GQVELGGGPGAGSLQPLALEGSLQARGIVEQCCTSICSLYQLENYCN

proP: The gene encoding *proP* is in UPPERCASE. The nucleotide sequence corresponding to the Gibson Assembly overlap regions used to install *proP* into plasmid pQE80_proP_proS are in lowercase.

gaattcattaaagaggagaaattaactATGCTGAAAAGGAAAAAAGTAAAACCGATTACCCTT
CGTGATGTCACCATTATTGATGACGGTAAACTGCGTAAAGCCATTACCGCAG
CATCACTGGGTAATGCAATGGAATGGTTCGATTTTGGTGTTTATGGTTTTGTT
GCTTACGCATTAGGTAAAGTTTTTTTCCCGGGGGCTGACCCCAGCGTGCAGAT

GGTTGCTGCACTTGCCACTTTCTCCGTTCCCTTTCTGATTCGACCGCTTGGCGG
ACTCTTCTTTGGTATGTTGGGCGATAAATATGGTCGCCAGAAGATCCTCGCTA
TCACTATTGTGATTATGTCGATCAGTACGTTCTGTATTGGCTTAATACCGTCCT
ACGACACGATTGGTATTTGGGCACCGATTCTGCTGTTGATCTGTAAGATGGCA
CAAGGTTTCTCGGTCGGCGGTGAATATACCGGGGCGTCGATATTTGTTGCGG
AATACTCCCCTGACCGTAAACGTGGCTTTATGGGCAGCTGGCTGGACTTCGGT
TCTATTGCCGGGTTTGTGCTGGGTGCGGGCGTGGTGGTGTTAATTTCGACCAT
TGTCGGCGAAGCGAACTTCCTCGATTGGGGCTGGCGTATTCCGTTCTTTATCG
CTCTGCCGTTAGGGATTATCGGGCTTTACCTGCGCCATGCGCTGGAAGAGACT
CCGGCGTTCCAGCAGCATGTCGATAAACTGGAACAGGGCGACCGTGAAGGTT
TGCAGGATGGCCCGAAAGTCTCGTTTAAAGAGATTGCCACTAAATACTGGCG
CAGCCTGTTGACATGTATTGGTCTGGTAATTGCCACCAACGTGACTTACTACA
TGTTGCTGACCTATATGCCGAGTTATTTGTCGCATAACCTGCATTACTCCGAA
GACCACGGGGTGCTGATTATTATCGCCATTATGATCGGTATGCTGTTTGTCCA
GCCGGTGATGGGCTTGCTGAGTGACCGTTTTGGCCGTCGTCCGTTTGTGCTAC
TTGGTAGTGTTGCCCTGTTTGTGTTGGCGATCCCGGCGTTTATTCTGATTAACA
GTAACGTCATCGGCCTGATTTTTGCCGGGTTACTGATGCTGGCGGTGATCCTT
AACTGCTTTACGGGCGTTATGGCTTCTACCTTGCCAGCGATGTTCCCGACGCA
TATCCGTTACAGCGCGCTGGCGGCGGCATTTAATATTTCGGTGCTGGTTGCCG
GTCTGACGCCAACGCTGGCGGCCTGGCTGGTCGAAAGCTCGCAGAATCTGAT
GATGCCTGCCTATTACCTGATGGTAGTGGCGGTGGTTGGTTTAATCACCGGCG
TAACCATGAAAGAGACGGCAAATCGTCCGTTGAAAGGTGCGACACCGGCGG
CGTCAGATATACAGGAAGCGAAGGAAATTCTCGTCGAGCATTACGATAATAT
CGAGCAGAAAATCGATGATATTGACCACGAGATTGCCGATTTGCAGGCGAAA
CGTACCCGCCTGGTGCAGCAACATCCGCGAATTGATGAATAAggatccgcatgcgag

MLKRKKVKPITLRDVTIIDDGKLRKAITAASLGNAMEWFDFGVYGFVAYALGKV
FFPGADPSVQMVAALATFSVPFLIRPLGGLFFGMLGDKYGRQKILAITIVIMSISTF
CIGLIPSYDTIGIWAPILLLICKMAQGFSVGGEYTGASIFVAEYSPDRKRGFMGSW
LDFGSIAGFVLGAGVVVLISTIVGEANFLDWGWRIPFFIALPLGIIGLYLRHALEET
PAFQQHVDKLEQGDREGLQDGPKVSFKEIATKYWRSLLTCIGLVIATNVTYYML

LTYMPSYLSHNLHYSEDHGVLIIIAIMIGMLFVQPVMGLLSDRFGRRPFVLLGSVA
LFVLAIPAFILINSNVIGLIFAGLLMLAVILNCFTGVMASTLPAMFPTHIRYSALAA
AFNISVLVAGLTPTLAAWLVESSQNLMMPAYYLMVVAVVGLITGVTMKETANR
PLKGATPAASDIQEAKEILVEHYDNIEQKIDDIDHEIADLQAKRTRLVQQHPRIDE

4.6.6 Cloning

pQE80_proP_proS: We installed the *proP* gene in place of proinsulin in plasmid pQE80_H27R-PI_proS using a three-part Gibson Assembly approach. The vector was amplified as two fragments which split the ampicillin resistance gene using the following primer sets: AmpR-GA_fwd / SLB7008_vec_rev & SLB7008_vec_fwd / AmpR-GA_rev. *proP* was amplified from the *E. coli* genome using colony PCR with primers SLB7008_ins_fwd & SLB7008_ins_rev.

pBAD33_PPIase-3xFLAG: Genes encoding nine *E. coli* proteins associated with PPIase activity (*fklB*, *fkpA*, *fkpB*, *ppiA*, *ppiB*, *ppiC*, *slyD*, *surA*, and *tig*), with a C-terminal GS-linker and 3x-FLAG tag, were ordered as g-Block gene fragments from Integrated DNA technologies (IDT). They were installed into vector pBAD33 (amplified with primers SLB7009_vec_fwd & SLB7009_vec_rev) by Gibson Assembly. The amino acid sequences of each tagged PPIase are given in Table 4.S1.

4.6.7 Incorporation of photo-pro into proinsulin under osmotic stress conditions

A single colony of strain CAG18515 harboring plasmid pQE80_H27R-PI_proS or pQE80PI-proS was used to inoculate a culture of Luria Bertani (LB) medium supplemented with ampicillin. The culture was grown overnight at 37°C until stationary phase was

reached, then diluted 1:100 into 1x M9 medium, supplemented with all twenty amino acids. The medium composition of M9 is as follows: 8.5 mM NaCl, 18.7 mM NH_4Cl, 22 mM KH_2PO_4, 47.8 mM Na_2HPO_4, 0.1 mM $CaCl_2$, 1 mM $MgSO_4$, 3 mg L^{-1} $FeSO_4$, 1 μg L^{-1} trace metals [Cu^{2+}, Mn^{2+}, Zn^{2+}, MoO_4^{2-}], 35 mg L^{-1} thiamine HCl, 10 mg L^{-1} biotin, 20 mM D-glucose, 100 mg L^{-1} ampicillin, 50 mg L^{-1} of each L-amino acid.

The culture was grown at 37°C until it reached OD ~0.8, after which it was subjected to a medium shift: cells were pelleted via centrifugation (5 kg, 5 min, 4°C) and washed twice with ice-cold 0.9% NaCl. Washed cells were resuspended in 1.25x M9 –Pro, a 1.25x concentrated form of M9 that omits proline. The culture was split into 4 mL aliquots, and incubated for 30 min at 37°C to deplete residual proline. A 1 mL solution containing 1.25-5.0 mM ncPro and 0-2.5 M NaCl was added (0.25-1.0 mM ncPro and 0-0.5 M NaCl working concentrations). Cultures labeled with photo-pro were covered with aluminum foil to limit exposure to light. After 30 min of incubation at 37°C to allow for ncPro uptake, proinsulin expression was induced by the addition of 1 mM IPTG. Cultures were incubated for 2.5 h at 37°C, after which cells were harvested via centrifugation and stored at -80°C until further processing.

4.6.8 Incorporation of photo-pro into proinsulin with proP overexpression

A single colony of strain CAG18515 harboring plasmid pQE80_proP_proS and pBAD-PI was used to inoculate a culture of LB medium supplemented with ampicillin and chloramphenicol. The culture was grown overnight at 37°C until stationary phase was reached, then diluted 1:100 into 1x M9 medium, supplemented with all twenty amino acids.

The culture was grown at 37°C until it reached OD ~0.2, then 0.5 mM IPTG (to induce *proP* expression) was added. After 1.5 h incubation at 37°C, cells were pelleted via centrifugation (5 kg, 5 min, 4°C), washed twice with ice-cold 0.9% NaCl, then resuspended in 1x M9 –Pro (M9 medium as described above that lacks proline). The culture was incubated for 30 min at 37°C to deplete residual proline, after which 1 mM photo-pro, 4*R*-OH, or proline was added. Samples labeled with photo-pro were covered with aluminum foil to limit exposure to light. After 30 min incubation at 37°C to enable proline uptake, 0.1% arabinose was added to induce proinsulin expression. Cultures were incubated for 3 h at 37°C, then centrifuged and pellets stored at -80°C until further processing.

4.6.9 Incorporation of photo-pro with proP overexpression for PPIase crosslinking

A single colony of strain CAG18515 harboring plasmid pQE80_proP_proS and pBAD-PPIase was used to inoculate a culture of LB medium supplemented with ampicillin and chloramphenicol. Here, PPIase in pBAD-PPIase represents one of nine PPIases (Table 4.S1) with a C-terminal 3xFLAG tag. The culture was grown overnight at 37°C until stationary phase was reached, then diluted 1:100 into 1x M9 medium, supplemented with all twenty amino acids.

The culture was grown at 37°C until it reached OD ~0.2 (or ~0.4 for proteomics studies), then 0.5 mM IPTG (to induce *proP* expression) and 0.1% arabinose (to induce expression of the PPIase) were added. After 2 h incubation at 37°C, cells were pelleted via centrifugation (5 kg, 5 min, RT), washed twice with 1x M9 –Pro, then resuspended in 1x M9 –Pro. The culture was incubated for 30 min at 37°C to deplete residual proline, after

which 1 mM photo-pro or proline was added. Samples labeled with photo-pro were covered with aluminum foil to limit exposure to light. Cultures were incubated for 1.5 h at 37°C, then irradiated with UV light (see section 4.6.12 below).

4.6.10 Inclusion body isolation

Proinsulin is expressed in the inclusion body fraction in *E. coli*. Cell pellets were thawed and lysed with B-PER Complete (Thermo Fisher Scientific) for 1 h at room temperature with shaking, then centrifuged (14 kg, 10 min) and the supernatant discarded. The pellet (containing insoluble proinsulin) was washed once with Triton wash buffer (2 M urea, 20 mM Tris, 1% Triton X-100, pH 8.0), and twice with ddH_2O. The pellet was resuspended in solubilization buffer (8 M urea, 300 mM NaCl, 50 mM NaH_2PO_4, pH 8.0), and proinsulin was allowed to dissolve for 4.5 h at room temperature with shaking. Samples were centrifuged, and the supernatant removed for analysis (described in section 4.6.11 below).

4.6.11 LC-ESI-MS

To assess levels of photo-pro incorporation into proinsulins, samples were digested with Glu-C, which results in a peptide fragment containing ProB28 (RGFFYT**P**KTRRE). A 20 µL aliquot of proinsulin was subjected to cysteine reduction (5 mM DTT, 55°C for 20 min) and alkylation (15 mM iodoacetimide, RT for 15 min in the dark), prior to 10-fold dilution into 100 mM NH_4HCO_3, pH 8.0 (100 µL final volume). Digestion was started with addition of 0.6 µL Glu-C (0.5 µg μL^{-1} in ddH_20) at 37°C for 2.5 h. The digestion reaction was

quenched by adding 10 µL of 5% TFA. Peptides were analyzed by LC-ESI-MS at the Multi User Mass Spectrometry Laboratory at Caltech.

The $[M+3H]^{+3}$ ion corresponding to the peptide RGFFYT**P**KTRRE was monitored (m/z = 520.3, proline; 519.3, dhp; 528.3, photo-pro; 525.3, 4*R*-OH). Analog incorporation was calculated from the EIC by comparing the area under the curve (AUC) of the ncPro form of the peptide with the AUC of the canonical proline peptide.

Ions corresponding to dhp (m/z = 528.3) and proline (519.3) that co-eluted with the photo-pro ion were presumed to be degradation products formed during LC-MS and discarded from incorporation efficiency analysis. Further, the dhp peptide could not be resolved from the proline peptide. This led to an over-estimation of proline-peptide levels due to the presence of dhp-peptide isotopes at the same m/z as the proline-peptide. The incorporation efficiencies for photo-pro reported here thus represent under-estimated values.

4.6.12 UV irradiation

Cultures were aliquoted in 6-well tissue culture plates (5 mL culture per well) and irradiated for 15 min with 365 nm light (Black-Ray®XX-20BLB UV Bench Lamp, 20 Watt, P/N 95-0045-05, 10 cm distance) at 4 °C. Samples were then transferred to 15-50 mL tubes and centrifuged. Cell pellets were stored at -80°C until further processing and analysis.

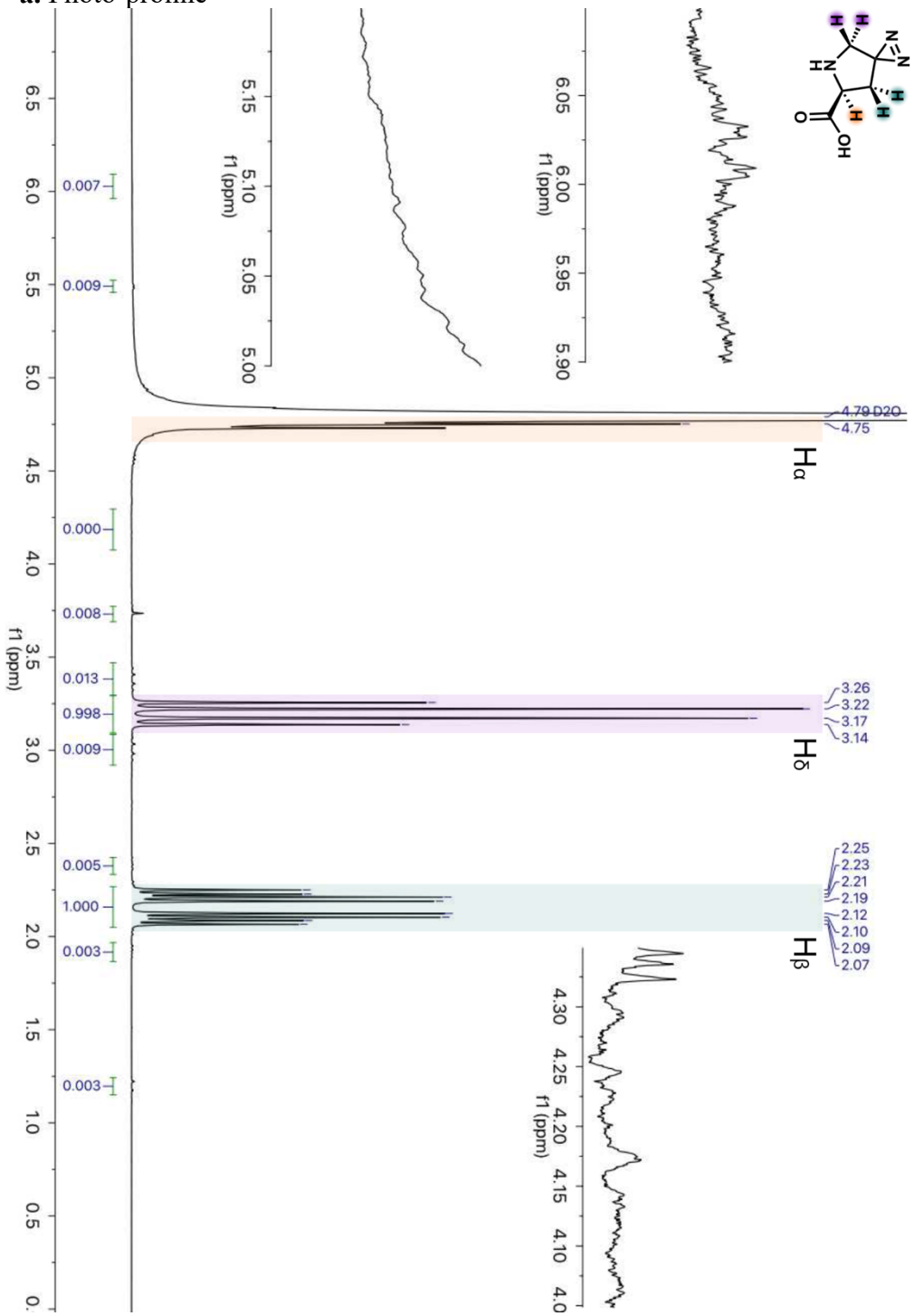
4.6.13 Western blot

Proteins separated by SDS-PAGE (NuPAGE 4-12% Bis-Tris, Thermo Fisher Scientific) were transferred onto a nitrocellulose membrane by an iBlot 2 gel transfer device (Thermo Fisher Scientific) using the manufacturer's protocols. Blots were blocked with 1 g dry milk in 20 mL PBST for 1 h at room temperature, then 4 µL of a primary mouse M2 anti-FLAG antibody was added (1:5,000 dilution, 13 h, 4°C). The blot was washed four times with 20 mL PBST prior to the addition of the secondary antibody (goat anti-mouse–AF647 conjugate; 1:10,000 dilution, 2.5 h, RT). Blots were washed as before, then imaged on a Typhoon Trio (GE Healthcare) with a 633 nm laser and 670 nm bandpass filter.

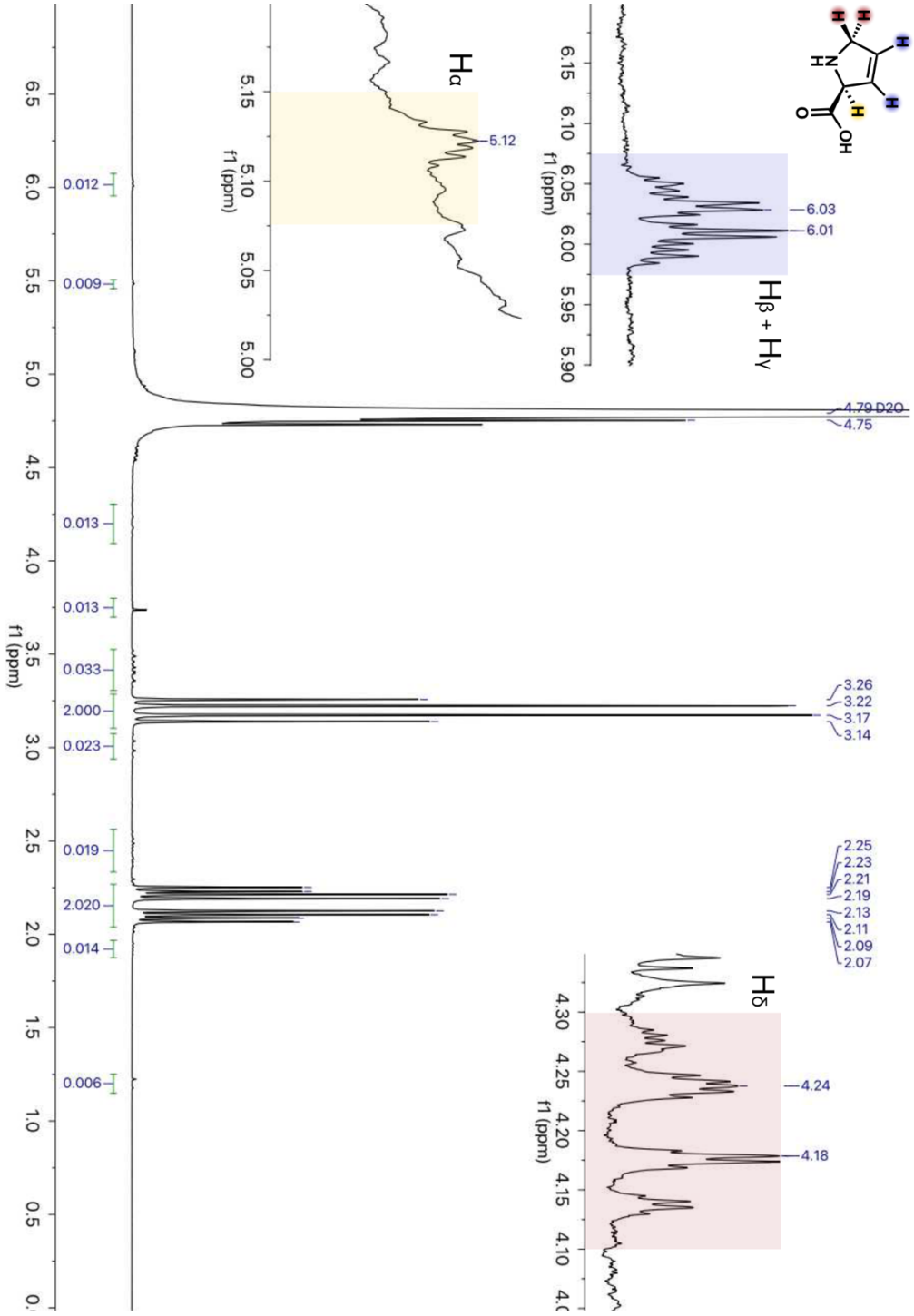
4.7 Supplementary figures and tables

Figure 4.S1. Thermal degradation products of photo-pro include dhp. ^{1}H NMR spectra of photo-pro **a**. immediately after dissolution in D_2O; **b**. after heating to 37°C for 3 h, dark (mimicking culturing conditions); **c**. heated photo-pro + 10% dhp.

a. Photo-proline



b. Photo-proline; 37°C, 3 h



c. Photo-proline; 37°C, 3 h + 10% dhp

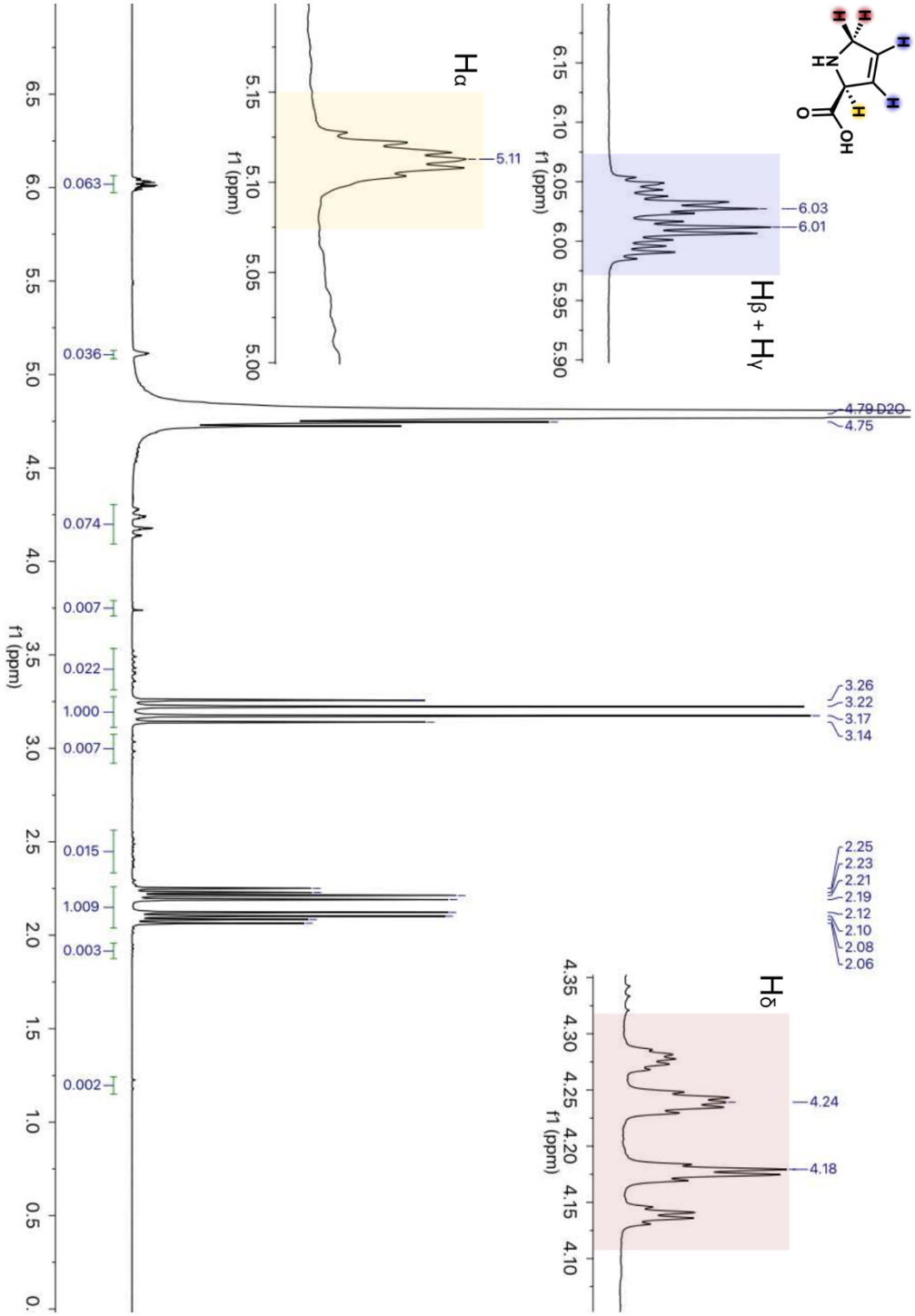


Table 4.S1. 3xFLAG-tagged PPIases

Name	Amino acid sequence*
fklB	MTTPTFDTIEAQASYGIGLQVGQQLSESGLEGLLPEALVAGIADALEG KHPAVPVDVVHRALREIHERADAVRRQRFQAMAAEGVKYLEENAK KEGVNSTESGLQFRVINQGEGAIPARTDRVRVHYTGKLIDGTVFDSSV ARGEPAEFPVNGVIPGWIEALTLMPVGSKWELTIPQELAYGERGAGAS IPPFSTLVFEVELLEILggsggsdykdhdgdykdhdidykddddk
fkpA	MKSLFKVTLLATTMAVALHAPITFAAEAAKPATAADSKAAFKNDDQ KSAYALGASLGRYMENSLKEQEKLGIKLDKDQLIAGVQDAFADKSKL SDQEIEQTLQAFEARVKSSAQAKMEKDAADNEAKGKEYREKFAKEK GVKTSSTGLVYQVVEAGKGEAPKDSDTVVVNYKGTLIDGKEFDNSY TRGEPLSFRLDGVIPGWTEGLKNIKKGGKIKLVIPPELAYGKAGVPGIP PNSTLVFDVELLDVKPAPKADAKPEADAKAADSAKKggsggsdykdhdgd ykdhdidykddddk
fkpB	MSESVQSNSAVLVHFTLKLDDGTTAESTRNNGKPALFRLGDASLSEG LEQHLLGLKVGDKTTFSLEPDAAFGVPSPDLIQYFSRREFMDAGEPEI GAIMLFTAMDGSEMPGVIREINGDSITVDFNHPLAGQTVHFDIEVLEID PALEAggsggsdykdhdgdykdhdidykddddk
ppiA	MFKSTLAAMAAVFALSALSPAAMAAKGDPHVLLTTSAGNIELELDK QKAPVSVQNFVDYVNSGFYNNTTFHRVIPGFMIQGGGFTEQMQQKKP NPPIKNEADNGLRNTRGTIAMARTADKDSATSQFFINVADNAFLDHG QRDFGYAVFGKVVKGMDVADKISQVPTHDVGPYQNVPSKPVVILSA KVLPggsggsdykdhdgdykdhdidykddddk
ppiB	MVTFHTNHGDIVIKTFDDKAPETVKNFLDYCREGFYNNTIFHRVINGF MIQGGGFEPGMKQKATKEPIKNEANNGLKNTRGTLAMARTQAPHSA TAQFFINVVDNDFLNFSGESLQGWGYCVFAEVVDGMDVVDKIKGVA TGRSGMHQDVPKEDVIIESVTVSEggsggsdykdhdgdykdhdidykddddk
ppiC	MAKTAAALHILVKEEKLALDLLEQIKNGADFGKLAKKHSICPSGKRG GDLGEFRQGQMVPAFDKVVFSCPVLEPTGPLHTQFGYHIIKVLYRNgg sggsdykdhdgdykdhdidykddddk
slyD	MKVAKDLVVSLAYQVRTEDGVLVDESPVSAPLDYLHGHGSLISGLET ALEGHEVGDKFDVAVGANDAYGQYDENLVQRVPKDVFMGVDELQV GMRFLAETDQGPVPVEITAVEDDHVVVDGNHMLAGQNLKFNVEVV AIREATEEELAHGHVHGAHDHHHDHDHDGCCGGHGHDHGHEHGGE GCCGGKGNGGCGCHggsggsdykdhdgdykdhdidykddddk

surA	MKNWKTLLLGIAMIANTSFAAPQVVDKVAAVVNNGVVLESDVDGL MQSVKLNAAQARQQLPDDATLRHQIMERLIMDQIILQMGQKMGVKI SDEQLDQAIANIAKQNNMTLDQMRSRLAYDGLNYNTYRNQIRKEMI ISEVRNNEVRRRITILPQEVESLAQQVGNQNDASTELNLSHILIPLPEN PTSDQVNEAESQARAIVDQARNGADFGKLAIAHSADQQALNGGQM GWGRIQELPGIFAQALSTAKKGDIVGPIRSGVGFHILKVNDLRGESKN ISVTEVHARHILLKPSPIMTDEQARVKLEQIAADIKSGKTTFAAAAKE FSQDPGSANQGGDLGWATPDIFDPAFRDALTRLNKGQMSAPVHSSF GWHLIELLDTRNVDKTDAAQKDRAYRMLMNRKFSEEAASWMQEQ RASAYVKILSNggsggsdykdhdgdykdhdidykddddk
tig	MQVSVETTQGLGRRVTITIAADSIETAVKSELVNVAKKVRIDGFRKG KVPMNIVAQRYGASVRQDVLGDLMSRNFIDAIIKEKINPAGAPTYVP GEYKLGEDFTYSVEFEVYPEVELQGLEAIEVEKPIVEVTDADVDGML DTLRKQQATWKEKDGAVEAEDRVTIDFTGSVDGEEFEGGKASDFVL AMGQGRMIPGFEDGIKGHKAGEEFTIDVTFPEEYHAENLKGKAAKF AINLKKVEERELPELTAEFIKRFGVEDGSVEGLRAEVRKNMERELKS AIRNRVKSQAIEGLVKANDIDVPAALIDSEIDVLRRQAAQRFGGNEK QALELPRELFEEQAKRRVVVGLLLGEVIRTNELKADEERVKGLIEEM ASAYEDPKEVIEFYSKNKELMDNMRNVALEEQAVEAVLAKAKVTE KETTFNELMNQQAggsggsdykdhdgdykdhdidykddddk

* Amino acid sequences in UPPERCASE correspond to the indicated PPIase; those in lowercase represent the C-terminal linker and 3xFLAG tag.

4.8 References

(1) Veselovsky, A. V.; Ivanov, Y. D.; Ivanov, A. S.; Archakov, A. I.; Lewi, P.; Janssen, P. Protein–Protein Interactions: Mechanisms and Modification by Drugs. *J Mol Recognit* **2002**, *15* (6), 405–422. https://doi.org/10.1002/JMR.597.

(2) Zhou, M.; Li, Q.; Wang, R. Current Experimental Methods for Characterizing Protein-Protein Interactions. *ChemMedChem* **2016**, *11* (8), 738–756. https://doi.org/10.1002/cmdc.201500495.

(3) Yang, Y.; Song, H.; Chen, P. R. Genetically Encoded Photocrosslinkers for Identifying and Mapping Protein-Protein Interactions in Living Cells. *IUBMB Life* **2016**, *68* (11), 879–886. https://doi.org/10.1002/iub.1560.

(4) Zhang, M.; Lin, S.; Song, X.; Liu, J.; Fu, Y.; Ge, X.; Fu, X.; Chang, Z.; Chen, P. R. A Genetically Incorporated Crosslinker Reveals Chaperone Cooperation in Acid Resistance. *Nat Chem Biol 2011 7:10* **2011**, *7* (10), 671–677. https://doi.org/10.1038/nchembio.644.

(5) Mori, H.; Ito, K. Different Modes of SecY-SecA Interactions Revealed by Site-Directed in Vivo Photo-Cross-Linking. *Proc Natl Acad Sci* **2006**, *103* (44), 16159–16164. https://doi.org/10.1073/pnas.0606390103.

(6) Carrico, I. S.; Maskarinec, S. A.; Heilshorn, S. C.; Mock, M. L.; Liu, J. C.; Nowatzki, P. J.; Franck, C.; Ravichandran, G.; Tirrell, D. A. Lithographic Patterning of Photoreactive Cell-Adhesive Proteins. *J Am Chem Soc* **2007**, *129*, 4874–4875. https://doi.org/10.1021/ja070200b.

(7) Kauers, J. C.; Erickson-Viitanen, S.; Wolfe, H. R.; Degrados, W. F. P-Benzoyl-L-Phenylalanine, A New Photoreactive Amino Acid. *J Biol Chem* **1986**, *261* (23), 10695–10700. https://doi.org/10.1016/S0021-9258(18)67441-1.

(8) Chin, J. W.; Martin, A. B.; King, D. S.; Wang, L.; Schultz, P. G. Addition of a Photocrosslinking Amino Acid to the Genetic Code of *Escherichia coli*. *Proc Natl Acad Sci* **2002**, *99* (17), 11020–11024. https://doi.org/10.1073/pnas.172226299.

(9) Chin, J. W.; Schultz, P. G. In Vivo Photocrosslinking with Unnatural Amino Acid Mutagenesis. *ChemBioChem* **2002**, *3* (11), 1135–1137. https://doi.org/10.1002/1439-7633.

(10) Chin, J. W.; Santoro, S. W.; Martin, A. B.; King, D. S.; Wang, L.; Schultz, P. G. Addition of P-Azido-L-Phenylalanine to the Genetic Code of *Escherichia coli*. *J Am Chem Soc* **2002**, *124* (31), 9026–9027. https://doi.org/10.1021/ja027007w.

(11) Tippmann, E. M.; Liu, W.; Summerer, D.; Mack, A. V.; Schultz, P. G. A Genetically Encoded Diazirine Photocrosslinker in *Escherichia coli*. *ChemBioChem* **2007**, *8* (18), 2210–2214. https://doi.org/10.1002/cbic.200700460.

(12) Ai, H.; Shen, W.; Sagi, A.; Chen, P. R.; Schultz, P. G. Probing Protein-Protein Interactions with a Genetically Encoded Photo-Crosslinking Amino Acid. *ChemBioChem* **2011**, *12* (12), 1854–1857. https://doi.org/10.1002/cbic.201100194.

(13) Suchanek, M.; Radzikowska, A.; Thiele, C. Photo-Leucine and Photo-Methionine Allow Identification of Protein-Protein Interactions in Living Cells. *Nat Methods* **2005**, *2* (4), 261–267. https://doi.org/10.1038/nmeth752.

(14) Yang, T.; Li, X. M.; Bao, X.; Fung, Y. M. E.; Li, X. D. Photo-Lysine Captures Proteins That Bind Lysine Post-Translational Modifications. *Nat Chem Biol* **2016**, *12* (2), 70–72. https://doi.org/10.1038/nchembio.1990.

(15) Black, D.; Tran, Q.-K.; Keightley, A.; Chinawalkar, A.; McMullin, C.; Persechini, A. Evaluating Calmodulin–Protein Interactions by Rapid Photoactivated Cross-Linking in Live Cells Metabolically Labeled with Photo-Methionine. *J Proteome Res* **2019**, *18* (10), 3780–3791. https://doi.org/10.1021/ acs.jproteome.9B00510.

(16) MacKinnon, A. L.; Garrison, J. L.; Hegde, R. S.; Taunton, J. Photo-Leucine Incorporation Reveals the Target of a Cyclodepsipeptide Inhibitor of Cotranslational Translocation. *J Am Chem Soc* **2007**, *129* (47), 14560–14561. https://doi.org/10.1021/ja076250y.

(17) Srinivas, N.; Jetter, P.; Ueberbacher, B. J.; Werneburg, M.; Zerbe, K.; Steinmann, J.; Van Der Meijden, B.; Bernardini, F.; Lederer, A.; Dias, R. L. A.; Misson, P. E.; Henze, H.; Zumbrunn, J.; Gombert, F. O.; Obrecht, D.; Hunziker, P.; Schauer, S.; Ziegler, U.; Käch, A.; Eberl, L.; Riedel, K.; Demarco, S. J.; Robinson, J. A. Peptidomimetic Antibiotics Target Outer-Membrane Biogenesis in Pseudomonas Aeruginosa. *Science (1979)* **2010**, *327* (5968), 1010–1013. https://doi.org/10.1126/science.1182749.

(18) Urfer, M.; Bogdanovic, J.; Monte, F. Lo; Moehle, K.; Zerbe, K.; Omasits, U.; Ahrens, C. H.; Pessi, G.; Eberl, L.; Robinson, J. A. A Peptidomimetic Antibiotic Targets Outer Membrane Proteins and Disrupts Selectively the Outer Membrane in *Escherichia coli*. *J Biol Chem* **2016**, *291* (4), 1921–1932. https://doi.org/10.1074/jbc.M115.691725.

(19) Vetterli, S. U.; Zerbe, K.; Müller, M.; Urfer, M.; Mondal, M.; Wang, S. Y.; Moehle, K.; Zerbe, O.; Vitale, A.; Pessi, G.; Eberl, L.; Wollscheid, B.; Robinson, J. A. Thanatin Targets the Intermembrane Protein Complex Required for Lipopolysaccharide Transport in *Escherichia coli*. *Sci Adv* **2018**, *4* (11), eaau2634. https://doi.org/10.1126/sciadv.aau2634.

(20) Andolina, G.; Bencze, L. C.; Zerbe, K.; Müller, M.; Steinmann, J.; Kocherla, H.; Mondal, M.; Sobek, J.; Moehle, K.; Malojčić, G.; Wollscheid, B.; Robinson, J. A. A Peptidomimetic Antibiotic Interacts with the Periplasmic Domain of LptD from Pseudomonas Aeruginosa. *ACS Chem Biol* **2018**, *13* (3), 666–675. https://doi.org/10.1021/ acschembio.7B00822.

(21) Zhou, Z.; Hu, T.; Zhou, X.; Wildum, S.; Garcia-Alcalde, F.; Xu, Z.; Wu, D.; Mao, Y.; Tian, X.; Zhou, Y.; Shen, F.; Zhang, Z.; Tang, G.; Najera, I.; Yang, G.; Shen, H. C.; Young, J. A. T.; Qin, N. Heteroaryldihydropyrimidine (HAP) and Sulfamoylbenzamide (SBA) Inhibit Hepatitis B Virus Replication by Different Molecular Mechanisms. *Sci Rep* **2017**, *7* (1), 1–12. https://doi.org/10.1038/srep42374.

(22) Fang, K. Y. Modulating Biophysical Properties of Insulin with Non-Canonical Mutagenesis at Position B28, California Institute of Technology, 2017. https://doi.org/10.7907/Z9V40S6J.

(23) Kim, W.; George, A.; Evans, M.; Conticello, V. P. Cotranslational Incorporation of a Structurally Diverse Series of Proline Analogues in an *Escherichia coli* Expression System. *ChemBioChem* **2004**, *5* (7), 928–936. https://doi.org/10.1002/cbic.200400052.

(24) Grothe, S.; Krogsrud, R. L.; McClellan, D. J.; Milner, J. L.; Wood, J. M. Proline Transport and Osmotic Stress Response in *Escherichia coli* K-12. *J Bacteriol* **1986**, *166* (1), 253–259. https://doi.org/10.1128/jb.166.1.253-259.1986.

(25) Weber, A.; Kögl, S. A.; Jung, K. Time-Dependent Proteome Alterations under Osmotic Stress during Aerobic and Anaerobic Growth in *Escherichia coli*. *J Bacteriol* **2006**, *188* (20), 7165–7175. https://doi.org/10.1128/jb.00508-06.

(26) Sasaki, H.; Sato, D.; Oshima, A. Importance of the High-Expression of Proline Transporter PutP to the Adaptation of *Escherichia coli* to High Salinity. *Biocontrol Sci* **2017**, *22* (2), 121124. https://doi.org/10.4265/bio.22.121.

(27) Lieblich, S. A.; Fang, K. Y.; Cahn, J. K. B.; Rawson, J.; LeBon, J.; Teresa Ku, H.; Tirrell, D. A. 4S-Hydroxylation of Insulin at ProB28 Accelerates Hexamer Dissociation and Delays Fibrillation. *J Am Chem Soc* **2017**, *139*, 8384–8387. https://doi.org/10.1021/jacs.7b00794.

(28) Kovermann, M.; Zierold, R.; Haupt, C.; Löw, C.; Balbach, J. NMR Relaxation Unravels Interdomain Crosstalk of the Two Domain Prolyl Isomerase and Chaperone SlyD. *Biochim Biophysica Acta Proteins Proteom* **2011**, *1814* (7), 873–881. https://doi.org/10.1016/ j.bbapap.2011.03.016.

(29) Saio, T.; Guan, X.; Rossi, P.; Economou, A.; Kalodimos, C. G. Structural Basis for Protein Antiaggregation Activity of the Trigger Factor Chaperone. *Science (1979)* **2014**, *344* (6184). https://doi.org/10.1126/science.1250494.

(30) Quistgaard, E. M.; Weininger, U.; Ural-Blimke, Y.; Modig, K.; Nordlund, P.; Akke, M.; Löw, C. Molecular Insights into Substrate Recognition and Catalytic Mechanism of the Chaperone and FKBP Peptidyl-Prolyl Isomerase SlyD. *BMC Biol* **2016**, *14* (1), 1–25. https://doi.org/10.1186/s12915-016-0300-3.