Site-specific Protein Modifications through Pyrroline-Carboxy-Lysine Residues

Weijia Ou[#], Tetsuo Uno[#], Hsien-Po Chiu, Jan Grünewald, Susan E. Cellitti, Tiffany Crossgrove, Xueshi Hao, Qian Fan, Lisa Quinn, Paula Patterson, Linda Okach, David H. Jones, Scott A. Lesley, Ansgar Brock, Bernhard H. Geierstanger*

Genomics Institute of the Novartis Research Foundation

10675 John-Jay-Hopkins Drive

San Diego, CA 92121-1125, USA

*To whom correspondence should be addressed. E-mail: bgeierst@gnf.org, Phone: 1-858-812-1633, Fax: 1-858-812-1746

[#]Contributed equally

SUPPORTING INFORMATION

1. Production of Pcl Proteins

1.1 Pcl Incorporation into Proteins in *E. coli*.

Pyl incorporation was accomplished in *E. coli* cells by co-transformation of two plasmids, pAra-pylSTBCD containing the Pyl biosynthetic genes *PylB*, *PylC* and *PylD*, and the pyrrolysyl-tRNA *PylT* and the pyrrolysyl-tRNA synthetase gene *PylS*, as well as an expression plasmid for the protein of interest as described(1). Alternatively, Pcl incorporation was accomplished with pARA-pylSTCD (lacking *PylB*) and the addition of *D*-ornithine to the growth media.

Plasmids. Existing expression plasmids with single TAG sites in the thioesterase domain of human fatty acid synthase (hFAS-TE) in pMH4(2) and mouse tumor necrosis factor alpha (mTNF-α) in a pET vector(3) were used, while new constructs for human FK506-binding protein 1A (FKBP) and mouse epidermal growth factor (mEGF) in pET vectors were cloned by the PIPE method(4). TAG codons for single-site incorporation of Pcl were introduced individually at the positions indicated in Table S1. FAS-TE constructs with altered +1 codons were also constructed by PIPE with the following codon changes designed to introduce a purine at the adjacent base to reduce termination efficiency(5): CTG2223ATT, TCC2307AGC, CGG2351AGG, and CAG2373AAT. Additional incorporation experiments were performed with several proprietary targets that were prepared with similar procedures.

Table S1. Residues targeted individually for Pcl substitution in protein expressed in *E. coli*.

Protein	**Sequence**
hFAS-TE (His_6TEV- 2221-2502)	MGSDKIHHHHHHENLYFQG-S**L**^{2222}LVNPEGPTLMRLNSVQSSERPLFLVHPIEGST**T**^{2255}V FHSLASRLSIPTYGLQCTRAAPLDSIHSLAAYYIDCIRQVQPEGPYRVAG**Y**^{2307}SYGACVAF EMCSQLQAQQSPAPTHNSLFLFDGSPT**Y**^{2343}VLA**Y**^{2347}TQS**Y**^{2351}RAKLTPGSEAEAETEAICF FV**Q**^{2373}Q**F**^{2375}TDMEHNRVLEALLPLKGLEERVAAAVDLIIKS**H**^{2408}QGLDRQELSFAARSFY YKLRAAEQYTPKAKYHGNVMLLRAK**T**^{2450}GGA**Y**^{2454}GEDLGADYNLSQVCDGKVSVHVI EGDHRTLLEGSGLESIISIIHSSLA
mTNF-α (His_6Xa- 1-156)	MRGSHHHHHHGSGIEGR-LRSSSQNSSD**K**^{11}PVAHVVANH**Q**^{21}VEEQLEWLSQRANALLA NGMDLKDNQLVVPADGLYLVYSQVLFKGQGCPDYVLLTHTVSRFAISYQEKVNLLSAV KSPCPKDTPEGAELKPWYEPIYLGGVFQLEKGDQLSAEVNLPKYLDFAESGQVYFGVIAL
FKBP (His_6HRV3C- 1-107)	MGSSHHHHHHLEVLFQGP-GVQVETISPGDGRTFPKRGQTCVVHY**T**^{27}GMLEDGKKFDSS RDRNKPF**K**^{47}FMLGKQEVIRGWEEGVAQMSVGQRAKLTISPDYAYGATGHPG**I**^{90}IPPHAT LVFDVELLKLE
mEGF (1-53- His_6)	M-NSYPGCPSS**Y**^{10}DGYCLNGGVCMHIESLDS**Y**^{29}TCNCVIGYSGDRCQTRDLRWWELR-LEHHHHHH

Protein Expresssion and Purification. hFAS-TE was expressed and purified as previously described(1). mTNF-α, FKBP, and mEGF were expressed in BL21 (DE3) cells (Invitrogen). Cells were grown in TB media (Sigma) at 37 °C with the addition of 5 mM D-ornithine. When cultures reached OD_{600} of 0.5, they were moved to 30 °C and induced with 1 mM IPTG and 0.2% (w/v) arabinose. Cells were harvested 12-16 hours later and pellets were stored at -20 °C until use. All wildtype proteins were grown in LB at 37 °C with induction for 3 hours using protocols otherwise identical to the mutant proteins. FKBP and hFAS-TE pellets were lysed by sonication in 25 mM Tris, 300 mM NaCl, 5 mM imidazole, 5% glycerol, pH 8.0. The lysate was centrifuged at 30,000 x g for 20 minutes at 4°C. The supernatant was loaded to a Ni-NTA (Qiagen) column which was washed with 10 column volumes (CVs) 25 mM Tris, 20 mM imidazole, pH 8.0 and eluted with 3 CV 25 mM Tris, 250 mM imidazole, pH 8.0. For mTNF-α, pellets were resuspended in 20 mM Tris, 50 mM NaCl, 1 mg/ml lysozyme, pH 8.0 and lysed by sonication. After centrifugation, an appropriate amount of 50% Ni-NTA slurry was added to the cleared lysate and gently mixed by shaking at 4 °C for 1 hour. The slurry was then moved to a column format, washed with 20 CV of 25 mM imidazole in PBS (pH 8.0), and the protein eluted with 2.5 CV 250 mM imidazole in PBS (pH 8.0). For mEGF, the cell pellet was lysed in 20 mM Tris pH 8.5 and centrifuged. The pellet was resuspended by sonication in 20 mM Tris pH 8.5, 2% (v/v) Triton-X100. After another round of centrifugation, the pellet was solubilized by sonication in 8 M urea, 20 mM Tris, 10 mM β-mercaptoethanol, pH 8.5. After centrifugation, the supernatant was diluted 2-fold with refolding buffer (100 mM Tris, 4 mM reduced glutathione, 0.4 mM oxidized glutathione, 20% (v/v) ethanol, pH 8.5) and dialyzed against refolding buffer overnight at 4 °C using a 3500 MWCO dialysis cassette (Thermo Scientific). Insoluble protein was removed by centrifugation at 30,000 x g for 20 minutes. To the supernatant were added β-mercaptoethanol to 2 mM and 10 mM imidazole and an appropriate amount of Ni-NTA slurry. The protein was then purified as for mTNF-α. For modifications, the proteins were buffer exchanged by PD-10 desalting column (GE Healthcare) or 3500 MWCO dialysis cassette (Thermo Scientific) (hFAS-TE into 20 mM imidazole, 200 mM NaCl, pH 6.5; FKBP into 20 mM HEPES, 150 mM NaCl, pH 7.0; mTNF-α and mEGF into PBS, pH 7.4) and concentrated by Amicon Ultra filters (Millipore). The purity of the protein preparations were assessed by SDS-PAGE, Pcl incorporation was verified by LC-MS, and proteins were quantified by Bradford assay (Thermo Scientific) or A_{280}. Protein yields based on Bradford measurements for the preparations shown in Fig. 2*A* are listed in Table S2.

Table S2. Protein expression yields. Expression yields for hFAS-TE Pcl mutants are shown in Fig. 2*C*.

Protein	mTNF-a			FKBP			mEGF		
	WT	Gln21Pcl		WT	Ile90Pcl		WT	Tyr10Pcl	
D-Orn	-	-	+	-	-	+	-	-	+
Yield [mg/L]	74	1.4	14	122	4.4	50	9.1	0.2	6.7

1.2. Production of Pcl Proteins in Mammalian Cells.

Pcl incorporation into proteins in mammalian cells was accomplished by transient transfection of HEK293F cells as described (1). In addition to human retinal binding protein 4 (RBP4), Pcl incorporation into mouse erythropoietin (EPO), into two mouse Fc domain constructs and into human IgG4 was accomplished. Pcl was also incorporated into several proprietary targets following similar procedures.

Cloning. Coding regions of human retinol binding protein (hRBP4) and mouse erythropoietin (EPO) were cloned into the pRS expression vector under the control of a CMV promoter with a His tag in the C-terminal. Two mIgG1 Fc constructs, Fc1 and Fc2, and an antibody construct in a human IgG4 frame (hIgG4) were also cloned into the pRS vector for transient expression in mammalian cells. Residues for mutation to a TAG codon in these target genes were selected based on their solvent-exposure in existing structural models and introduced into the target genes by site-directed mutagenesis (Table S3).

Table S3. Residues targeted individually for Pcl substitution in hRBP4 and mEPO expressed in mammalian cells.

hRBP4	mEPO
Y51	R28
F62	A54
W93	R58
Y116	K76
W117	Q89
F122	Q100
Y140	P112
Y191	K140
Y199	T149
	R155
	R190

Cell culture and transfection. HEK293F cells were grown in suspension in 293 Freestyle expression media at 37^{o}C under 5% CO_2. One day before transfection, cells were split to 0.7x10^{6} cells/ml. Plasmid DNA was prepared using Qiagen Maxi plasmid preparation kit. Transfections were carried out by PEI method. PEI was mixed with plasmid DNA in a ratio of 2 to 1 in Opti-MEM and the DNA complex was added to HEK293F cells at 1 ug plasmid DNA/ml cell culture and the cells were cultured for four days before harvest. For Pcl incorporation studies, 293F cells were co-transfected with a target gene plasmid containing TAG codon in a designated position indicated in Table S3 and three plasmids required for Pcl biosynthesis, pCMVpyS, pCMVpyC and pCMVpyD. The ratio of PEI to total amount of plasmid DNA was always 2 to 1. *D*-ornithine was added to culture media at 5 mM 5 hours after transfection.

Protein purification and analysis. Four days after transfection, cell cultures were centrifuged at 2000g for 20 min and media were collected for purification. For purification of His tagged hRBP4 and mEPO proteins, the media were loaded to Ni-NTA columns equilibrated previously with 20mM TrisHCl (pH 7.5), 150 mM NaCl containing 10 mM imidazole. The columns were washed with the same buffer and eluted with an elution buffer (20mM TrisHCl 7.5, 150 mM NaCl, 300 mM imidazole). Fc and hIgG4 constructs were purified by Protein A-affinity chromatography. Eluted proteins were assayed by Bradford method and analyzed on SDS-PAGE. In some cases, media and purified proteins were analyzed by western Blot with antibody to His tag or to the target proteins.

The protein yields for hRBP4 Phe62Pcl, Phe122Pcl and Tyr199Pcl were between 4 and 8 mg/L, about 20% of the yield for wild-type hRBP4. Expression levels of all eleven mutant mEPO proteins are similar and between 10 to 20% of wt protein that expresses at ~40 mg/L (Fig. 2*C*). The yield of the hIgG4 Pcl mutant used in Fig. 5*C*,*D* was 3 mg/L. Pcl incorporation was confirmed by mass spectrometric (MS) analysis.

2. Site-specific Derivatization of Pcl and Pyrrolysine Proteins

2.1. Reaction Conditions and Experimental Details for Fig. 3

Reagents. 2-Amino-benzyaldehyde (2-ABA) and 2-amino-acetophenone (2-AAP) were purchased from Sigma and used without further purification. The synthesis of the 2-AAP-PEG8 (TU3205-044) is described in Section 7.1. The synthesis of the 2-ABA analog TU3627-014 is described in Section 6.1. hRBP4 Phe122Pcl was prepared as described above. The expected mass of the protein is 23167.7 Da considering three disulfide bonds and cyclization of the N-terminal glutamine to pyroglutamate.

Site-specific modification of Pcl in hRBP4 by 2-ABA and 2-AAP (Fig. 3B-E). Ten microliter of hRBP4 Phe122Pcl protein stock solution (0.17 mM) was mixed with 89 μl of 200 mM sodium acetate buffer, pH 5.0 and 1 μl of a 10 mM 2-ABA (Fig. 3*B*) or 2-AAP (Fig. 3*C*) solution. The final concentrations in the reaction mixtures were 17 μM hRBP4 Phe122Pcl protein and 0.1 mM 2-ABA or 0.1 mM 2-AAP. After 12 hours at room temperature, mass spectra of the reaction mixtures predominantely showed the expected mass of conjugated protein (Fig. 3*B*,*C*): For 2-ABA (Fig. 3*B*), the relative intensity of the correctly conjugated peak (23720.2 Da observed, 23270.8 Da = 23167.7 Da + 103.1 Da expected) was ~95% of the total intensity; only ~5% of the protein remained unreacted (23167.2 Da observed, 23167.7 Da expected). The sodiated product peak at 23290.8 Da as well as other small features were not taken into account in the calculations. For 2-AAP (Fig. 3*C*), 95% appeared reacted (23283.6 Da observed, 23284.8 Da = 23167.7 Da + 117.1 Da expected); ~5% was unreacted (23166.8 Da observed, 23167.7 Da expected). The observed mass increases upon conjugation were consistent with the structure of the Pcl protein adducts shown in Fig. 4. Derivatization of the Pcl residues

was confirmed by LC-MS analysis of tryptic digests (see below).

17 μM hRBP4 Phe122Pcl protein was also incubated with 80 mM 2-ABA for 12 hours at 22 ^{o}C and pH 7.5. At this very high reactant to protein ratio (4700-fold) and at pH 7.5, attachment of multiple reactant molecules was observed (Fig. 3*D*): A small amount of unmodified protein (<4% of the total intensity in the mass spectrum) was observed at 23168.8 Da; ~20% of the total intensity corresponded to adducts with one 2-ABA molecule (23270.0 Da obs., 23170.8 = 23167.7 Da + 103.1 Da expected), ~6% adducts with two 2-ABA molecules (23371.6 Da), ~31% adducts with three 2-ABA molecules (23476.0 Da observed, 23477.0 Da = 23167.7 Da + 3 x 103.1 Da expected), ~6% with four at around 23580 Da, ~22% with five (23682.8 Da observed, 23683.2 Da = 23167.7 Da + 5 x 103.1 Da expected), ~5% with six (~23750 Da) and ~8% adducts with seven 2-ABA molecules (23889.6 Da observed, 23889.4 Da = 23167.7 Da + 7 x 103.1 Da expected). Precipitation of an identical sample at pH 5.0 suggested multiple reactions as well. Similarly, hRBP4 modified with O-methyl-phenylalanine (OMePhe)(6) instead of Phe62 (6.5 μM) showed a similar pattern of conjugates (Fig. 3*E*) when reacted with 100 mM 2-ABA (15400-fold molar excess) at pH 7.5 (Fig. 3*D*): Unmodified protein (23120.8 Da observed, 23121.6 Da expected) was observed with ~42% of the total intensity in the mass spectrum; ~35% of the intensity corresponded to adducts with two 2-ABA molecule (23327.2 Da observed, 23327.8 Da = 23121.6 Da + 2 x 103.1 Da expected) and ~14% adducts with four 2-ABA molecules (23532.8 Da observed, 23534.0 Da = 23121.6 Da + 4 x 103.1 Da). Small amounts of adducts with one (~4%) and three (~3%) 2-ABA were observed also.

These data suggest that at very large molar excess of reactant over protein conjugation of 2-ABA, in addition to Pcl, can also occur at other reactive groups on the protein. These additional adducts were predominantly observed as modifications with multiples of two 2-ABA molecules (+ 206 Da, + 412 Da, + 618 Da) suggesting a limited number of modification sites or conjugations at a single additional site with 2-ABA dimers, tetra- and hexamers.

Derivatization of Pcl incorporated into hRBP4 with 2-AAP-PEG8 (Fig. 3F-I). hRBP4 Phe122Pcl was derivatized efficiently at a single site with a polyethylene glycol (PEG) derivative of 2-AAP (TU3205-044) (Fig. 3*A*,*F*,*G*). Wild-type hRBP4 did not react even when a large excess (up to 2300:1 reagent to protein) of the 2-AAP-PEG8was applied (Fig. 3*H*,*I*).

For the reaction performed at pH 7.5 (Fig. 3*F*), 10 μL of hRBP4 Phe122Pcl (0.22 mg/mL, in PBS, pH 7.5) were diluted with 10 μl of 10x PBS. 2 μl of a 100 mM stock solution of 2-AAP-PEG8 (in water) were added to a final concentration of 9.1 mM. The protein concentration was 4.3 μM resulting in a 2100 molar excess of reactant to protein. For reactions performed at pH 5.0 (Fig. 3*G*), 10 μL of hRBP4 Phe122Pcl stock solution were diluted with 90 μl of 200 mM sodium acetate buffer at pH 5.0. Ten microliter of a 100 mM 2-

AAP-PEG8 stock solution were added to a final concentration of 9.1 mM. The protein concentration was 0.86 μM, respectively, resulting in a 10500 molar excess of reactant to protein. The reactions were allowed to proceed for 14 hours at room temperature followed by 72 hours at 4 degree °C before mass spectra were obtained of the reaction mixtures. Fig. 3*F* shows the mass spectra of pH 7.5 reaction mixture at a 2100 to 1 PEG to protein ratio, indicating ~94% completeness for the reaction (~6% unreacted protein, 23167.6 Da observed, 23167.7 Da expected) and the observed mass increase of 556.6 Da was identical to the expected value (reacted protein: 23724.2 Da observed, 23724.3 Da = 23167.7 Da + 556.6 Da expected). Similarly, Fig. 3*G* shows the mass spectra of the pH 5.0 reaction mixture at the 10500 to 1 ratio with ~4% unreacted protein remaining (Unreacted protein: 23167.8 Da observed, 23167.7 Da expected; reacted protein: 23724.2 Da observed, 23724.3 Da = 23167.7 Da + 556.6 Da expected).

To test the reactivity of wild-type hRBP4 protein, test reactions were setup in parallel to those above. For the pH 7.5 reaction (Fig. 3*H*), the final wild-type protein concentration was 20 μM and the 2-AAP-PEG8 concentration was 9.1 mM resulting in a molar ratio of 460 to 1. The pH 5.0 reaction (Fig. 3*I*), was performed at 4 μM protein concentration and 9.1 mM 2-AAP-PEG8 (2300 to 1 reactant to protein). In both reactions only unmodified wild-type hRBP4 protein was observed at the expected mass. These data demonstrated that the coupling reaction of 2-AAP-PEG8 was highly specific for the presence of a Pcl residue in the target protein.

Derivatization of Pyrrolysine (Pyl) and Pcl incorporated into mEGF (Fig. 3J,K). Under low oxygenation conditions and in the presence of the *pylB* gene, pyrrolysine (Pyl) is generated in *E. coli* and Pyl or a mixture of Pyl and Pcl can be incorporated into a protein(1). For example, mEGF Tyr10TAG expressed in *E. coli* BL21(DE3) cells in the presence of 5 mM *D*-ornithine when co-transformed with pAra-pylSTBCD, resulted in such a mixture with both Pcl and Pyl incorporated at position 10 of mEGF as confirmed by ESI-MS (Fig. 3*J*; expected mass of mEGF Tyr10Pcl = 7296 Da; expected mass of mEGF Tyr10Pyl = 7310 Da). For this particular sample, Pcl incorporation was prevalent over Pyl incorporation. To study the reactivity of Pyl with 2-ABA reagents, the mEGF Tyr10Pcl/Pyl mixture (10 μM) was reacted with 1 mM of TU3627-014 (see Section 6.1 for synthesis) in 10 x PBS (pH 7.0) and 1% (v/v) DMSO at 22 °C for 16 hours. Completion of the reaction and formation of the protein conjugates was confirmed by ESI-MS of the reaction mixture (Fig. 3*K*). The ratio between the Pcl and Pyl adduct resembled the ratio of Pcl and PYL in the unreacted protein (Fig. 3*K*), suggesting similar reactivity for Pcl and Pyl.

2.2. Mass Spectrometric Verification of 2-ABA/Pcl Adduct Formation

Specific derivatization of the Pcl residue in hRBP4 Phe122Pcl by 2-ABA was verified by LC-MS analysis after tryptic digestion. Intially, hRBP4 Phe122Pcl (17 μM) was reacted with 10 mM 2-ABA for 16 hrs at 22 ^{o}C in 200 mM sodium acetate buffer, pH 5.0. 23269.2 Da were measured for the derivatized hRBP4. The unmodified hRBP4 had a mass of 23166.8 Da and therefore the observed mass increase of 102.4 Da (expected +103 Da) demonstrated that the hRBP4 has been modified with 2-ABA. At least 96% of the peak intensities in the mass spectrum were due to 2-ABA adduct of Pcl, which indicates that the reaction went to near completion.

The LC-MS analysis of the tryptic digest of the 2-ABA-derivatized hRBP4 Phe122Pcl protein was performed and MS/MS analysis (Fig. S1) identified the expected YWGVASF*LQK peptide, wherein F* had a mass consistent with that of non-reduced 2-ABA-modificed Pcl as shown in Fig. 4***A***. The assigned MS/MS spectrum of YWGVASF*LQK (F* = Pcl-2-ABA adduct) is shown in Fig. S1*A* and tabulated in Table S4. Unmodified precursor was not observed in the EIC (Fig. S1*B*). MS data demonstrated that the observed reaction with 2-ABA occurred site-specifically with the Pcl residue incorporated at the desired TAG site at residue 122.

Table S4. Verification of site-specific modifications of Pcl by 2-ABA.

Monoisotopic precursor ion mass 1376.725Da (calc. 1376.724Da, error 0.8ppm)
Mascot assigned fragment ions in bold.

#	b	b^{++}	b*	b^{*++}	b^{0}	b^{0++}	Seq.	y	y^{++}	y*	y^{*++}	y^{0}	y^{0++}	#
1	164.1	82.5					Y							10
2	350.1	175.6					W	1214.7	607.87	1197.6	599.3	1196.7	598.8	9
3	407.2	204.1					G	1028.6	514.8	1011.6	506.3	1010.6	505.8	8
4	506.2	253.6					V	971.6	486.3	954.5	477.8	953.6	477.3	7
5	577.3	289.1					A	872.5	436.8	855.5	428.2	854.5	427.7	6
6	664.3	332.7			646.3	323.7	S	801.5	401.2	784.4	392.7	783.5	392.2	5
7	990.5	495.7			972.5	486.7	F	714.4	357.7	697.4	349.2			4
8	1103.6	552.3			1085.6	543.3	L	388.2	194.6	371.2	186.1			3
9	1231.6	616.3	1214.6	607.8	1213.6	607.3	Q	275.2	138.1	258.1	129.6			2
10							K	147.1	74.1	130.1	65.5			1

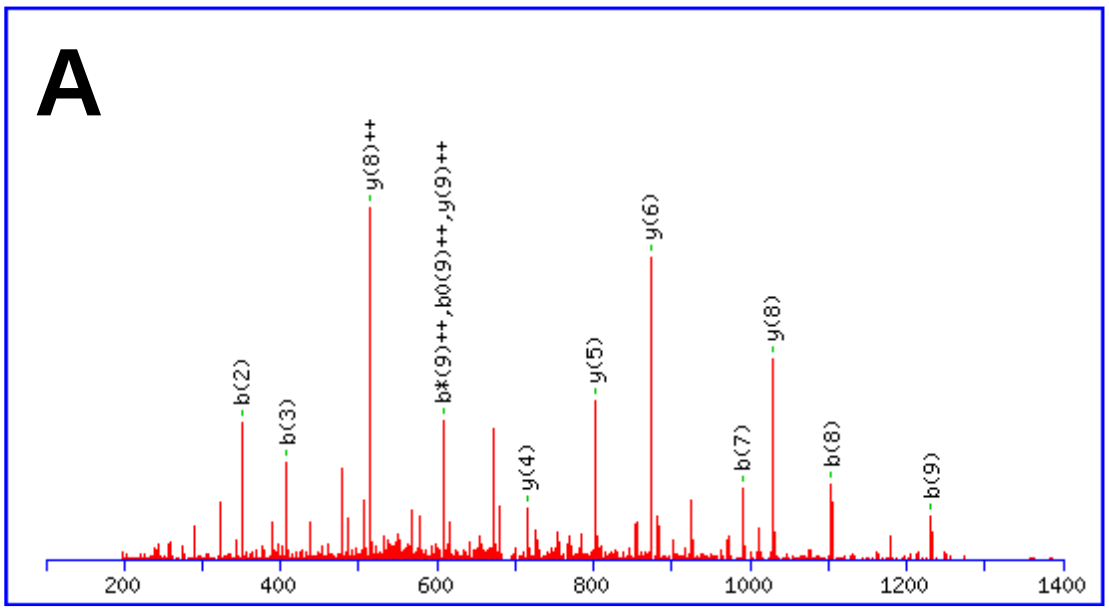
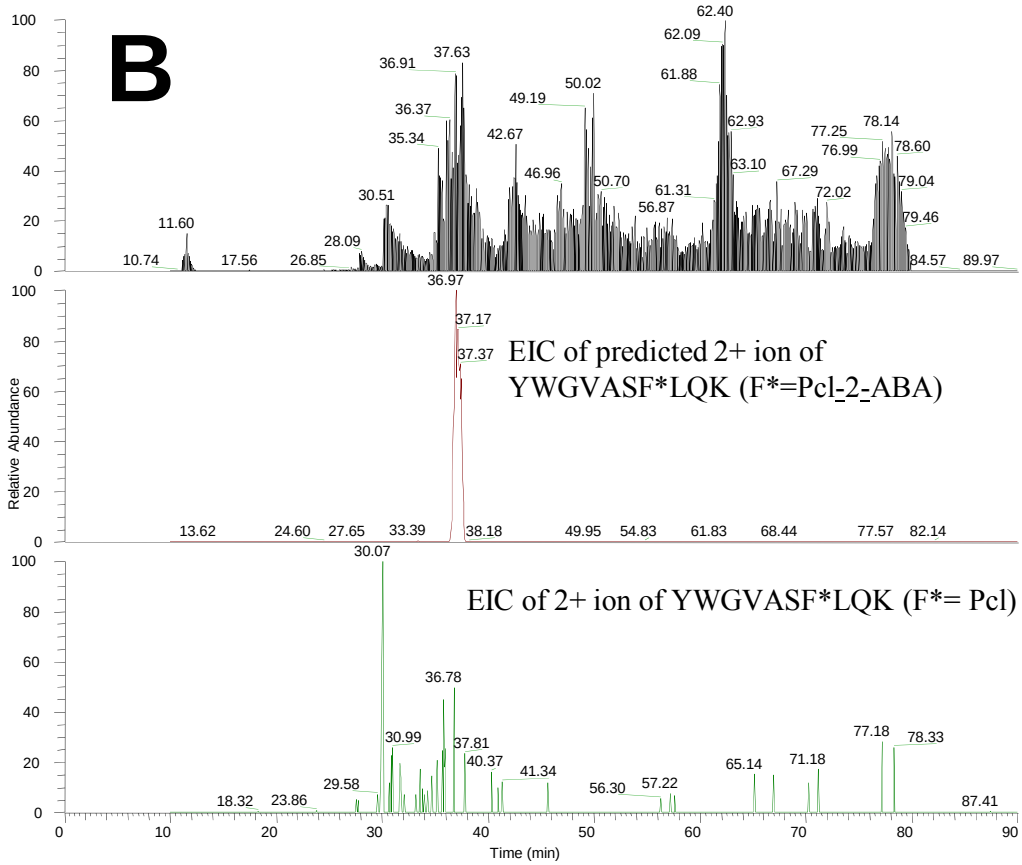


Fig. S1. Mass spectrometric verification of 2-ABA conjugation to Pcl. LC-MS analysis of the tryptic digest of the 2-ABA-derivatized hRBP4 Phe122Pcl protein verified Pcl specific derivatization through MS/MS analysis. (*A***)** Tandem MS of the peptide YWGVAS(Pcl-2-ABA)LQK. (*B*) TIC and EIC of 2+ ions of YWGVASF*LQK (F* = Pcl and Pcl-2-ABA adduct. (*C***)** MS spectrum showing 3+ and 2+ precursors of YWGVASF*LQK at m/z 459.92 (3+) and 689.37 (2+) respectively. (F* = Pcl-2-ABA adduct).


B
EIC of predicted 2+ ion of YWGVASF*LQK (F*=Pcl-2-ABA)
EIC of 2+ ion of YWGVASF*LQK (F*= Pcl)
Relative Abundance
Time (min)

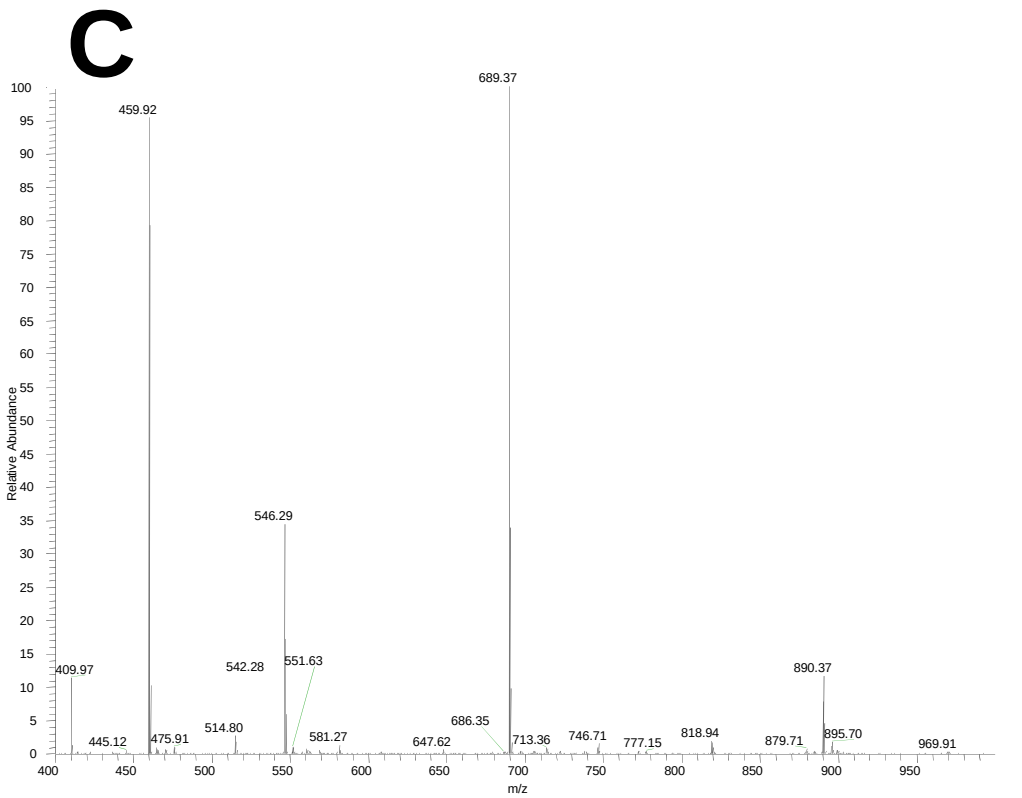

C
Relative Abundance
m/z

Fig. S1

3. Structural Characterization of Pcl Adducts by Nuclear Magnetic Resonance (NMR) Spectroscopy

As a model system for the conjugation of Pcl proteins with 2-amino-benzyaldehyde (2-ABA) and 2-amino-acetophenone (2-AAP) reagents, the reaction of Pcl amino acid and its isomer (S)-2-amino-6-(3,4-dihydro-2H-pyrrole-5-carboxamido)hexanoic acid and the resulting adducts with 2-ABA and 2-AAP were studied by standard 1D and 2D nuclear magnetic resonance (NMR) spectroscopy. In addition, the adduct with 3,4-dihydro-2H-pyrrole-2-carboxylic acid was also studied to further confirm the observations with Pcl.

3.1. Reagents and Synthesis of Pcl Amino Acids

2-AAP, 2-ABA and sodium cyanoborohydride were purchased from Sigma and used without further purification. D_2O and d6-DMSO were purchased from Cambridge Isotope Laboratory. Pcl and its isomer (S)-2-amino-6-(3,4-dihydro-2H-pyrrole-5-carboxamido)hexanoic acid as well as 3,4-dihydro-2H-pyrrole-2-carboxylic acid were synthesized as described(1). The synthetic Pcl amino acid ((2S)-2-amino-6-(3,4-dihydro-2H-pyrrole-2-carboxamido)hexanoic acid) is a diastereomeric mixture but in some cases the stereochemistry of Pcl adducts is drawn based on the stereochemistry that is observed for Pyl in proteins(7).

3.2. NMR Data Collection

For the reaction with 2-ABA and subsequent characterization of the products, NMR data was acquired at 300 K on a Bruker Avance NMR instrument (Bruker Biospin, Billerica, MA) equipped with a $^{1}H/^{13}C/^{19}F/^{31}P$-QNP-cryoprobe operating at a ^{1}H frequency of 400.13MHz. ^{1}H 1D spectra were typically recorded with 16 scans, relaxation delay of 5 s, 16384 complex data points with a sweep width of 12 ppm. ^{1}H-^{1}H double quantum filtered (DQF) COSY spectra were typically recorded with 4 scans, 256 t_1 experiments and ^{1}H-^{1}H ROESY spectra (200 ms mixing time) with 8 scans and 512 t_1 experiments. ^{1}H-^{13}C HMBC spectra were typically recorded with 32 scans and 256 t_1 experiments. The experiment was optimized for the detection of three-bond-correlations using a 50 ms delay (J = 10 Hz) for the evolution of long-range couplings (pulse program: hmbcgpndqf). ^{1}H-^{13}C HMQC spectra were typically recorded with 4 scans, 128 t_1 experiments using a spectral width of 222 ppm in the carbon dimension and 12 ppm or 7.5 ppm in the proton dimension.

For the characterization of the reduced adduct, all spectra were recorded at 300 K on a Bruker Avance 600 MHz instrument equipped with a $^{1}H/^{13}C/^{15}N$-TXI-cryoprobe. ^{1}H spectra were typically recorded with 64 scans, a relaxation delay of 2 s, 16384 complex data points with a sweep width of 14 ppm and excitation sculpting for water suppression. ^{1}H-^{1}H DQF-COSY spectra were typically recorded with 16 scans, 1024 t_1 experiments using a spectral width of 10 ppm. ^{1}H-^{13}C HMQC spectra were typically recorded with 8 scans, 256 t_1 experiments using a spectral width of 160 ppm in the carbon dimension and 10 ppm in the proton dimension.

^{1}H-^{13}C HMBC spectra were typically recorded with 88 scans, 256 t_1 experiments at 300K using a spectral width of 180 ppm in the carbon dimension and 10 ppm in the proton dimension, and a delay of 50 ms.

3.3. NMR studies of the Reaction of Pcl and its Isomer with 2-ABA

The reactions of Pcl and of its isomer (S)-2-amino-6-(3,4-dihydro-2H-pyrrole-5-carboxamido)hexanoic acid with 2-ABA were studied as follows: 1.0 mg of Pcl or (S)-2-amino-6-(3,4-dihydro-2H-pyrrole-5-carboxamido)hexanoic acid were dissolved in 0.5 mL PBS in D_2O. 10 μL of 10 mM 3-(trimethylsilyl)propionic acid (TSP) in D_2O were added as internal standard for referencing and for concentration determination by NMR. 3.7 mg of 2-aminobenzaldehyde was dissolved in 0.5 mL PBS in D_2O and 10 μL of 10 mM TSP. NMR signals of starting materials (Table S5, S6) were assigned using standard NMR methods with ^{1}H 1D, ^{1}H-^{1}H COSY and ^{1}H-^{1}H ROESY experiments.

Table S5. NMR signal assignments of unreacted Pcl and 2-ABA (for numbering, see Fig. S2*A*)

Atom number	Shift (ppm)	H's	Type	J (Hz)
19	1.31	2	m	-
18	1.47	2	quin	7.34
4	1.68	1	m	-
20	1.74	2	m	-
4	2.16	1	m	-
3	2.57	2	m	-
17	3.14	2	t	6.97
21	3.59	1	t	6.24
5	4.51	1	m	-
12, 10	6.76	2	m	-
11	7.34	1	ddd	8.50, 7.03, 1.59
13	7.53	1	ddd	8.19, 159
2	7.76	1	m	-
7	9.68	1	s	-

Table S6. NMR signal assignments of unreacted (S)-2-amino-6-(3,4-dihydro-2H-pyrrole-5-carboxamido)hexanoic acid and 2-ABA (for numbering, see Fig. S2*B*)

Atom number	Shift (ppm)	H's	Type	J (Hz)
19	1.33	2	m	-
18	1.5	2	q	7.02
20	1.77	2	m	-
3	1.88	2	m	-
4	2.69	2	m	-
17	3.21	2	t	6.94
21	3.61	1	m	-
2	3.85	2	m	-
12, 10	6.76	2	m	-
11	7.34	1	ddd	8.44, 7.05, 1.61
13	7.53	1	ddd	8.18, 1.61
7	9.67	1	s	-

To initiate the reaction, 325 μL of Pcl solution was mixed with 175 μL of the 2-ABA solution (approximately a 1:1 molar ratio) and the resulting solution was transferred to an NMR tube The reaction was allowed to proceed at room temperature and NMR spectra were periodically acquired at the indicated times (Fig. S2*A*). Signals for Pcl (dots) and for 2-ABA (stars) quickly disappeared as the reaction proceeded to completion with all Pcl being converted and only leaving a small amount of residual 2-ABA because of the slight excess. In the first time point acquired 0.5 hrs after mixing, two new species were detected at a ratio of

approximately 2:1 (representative resonances are marked by arrows). The minor species completely converted to the major species over the course of several days.

The reaction of the Pcl isomer (S)-2-amino-6-(3,4-dihydro-2H-pyrrole-5-carboxamido)hexanoic acid with 2-ABA was studied under identical conditions. (S)-2-amino-6-(3,4-dihydro-2H-pyrrole-5-carboxamido)hexanoic acid was dissolved in D_2O (as above), mixed with 2-ABA at an approximate molar ratio of 1:1, and transferred into an NMR tube. The reaction was allowed to proceed at room temperature and NMR spectra were acquired at the indicated times (Fig. S2*B*). In contrast to the Pcl sample, (S)-2-amino-6-(3,4-dihydro-2H-pyrrole-5-carboxamido)hexanoic acid did not react readily with 2-ABA. Even after 17 days most of the starting material (stars for 2-ABA; dots for (S)-2-amino-6-(3,4-dihydro-2H-pyrrole-5-carboxamido)hexanoic acid) was still present with only a small amount converted into a new species (arrows). This species could not be further characterized. However, the NMR results of the two reactions clearly indicated that the reactivity of Pcl with 2-ABA was much higher than that of (S)-2-amino-6-(3,4-dihydro-2H-pyrrole-5-carboxamido)hexanoic acid. This further supports the previous assignment of the Pcl structure(1).

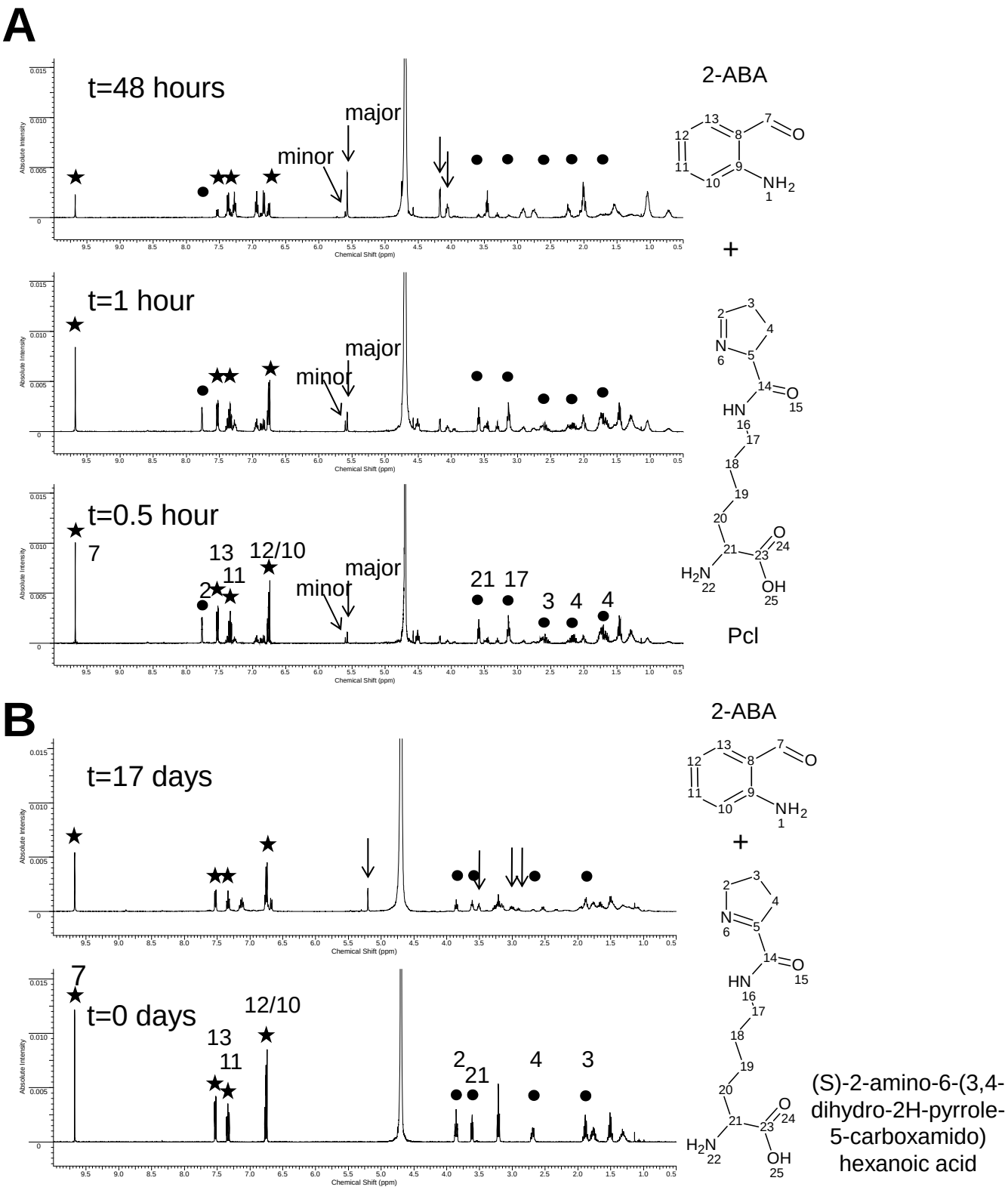


Fig. S2. Reactions of Pcl and its isomer (S)-2-amino-6-(3,4-dihydro-2H-pyrrole-5-carboxamido)hexanoic acid with 2-ABA as monitored by NMR. (*A***)** 1D NMR spectra of the Pcl and 2-ABA reaction collected at the indicated time after mixing. (*B***)** 1D NMR spectra of a mixture of (S)-2-amino-6-(3,4-dihydro-2H-pyrrole-5-carboxamido)hexanoic acid with 2-ABA at various times after mixing. Resonances of unreacted 2-ABA are indicated by stars and those of unreacted Pcl or of its isomer are indicated by dots. Resonances of the Pcl-2-ABA adduct or of uncharacterized products (in b) are highlighted by arrows.

3.4. NMR Characterization of the 2-ABA/Pcl Adduct

The final reaction product of Pcl with 2-ABA was characterized by ^{1}H and ^{13}C 1D, 2D ^{1}H-^{1}H COSY, ^{1}H-^{1}H ROESY, ^{1}H-^{13}C HMBC and ^{1}H-^{13}C HMQC NMR spectroscopy with samples prepared in D_2O and d6-DMSO. The sample in d6-DMSO was purified by HPLC. Signal assignments for proton and carbon resonances and the observed correlations in ^{1}H-^{13}C HMBC spectra in d6-DMSO are summarized in Table S7.

Table S7. ^{1}H-^{13}C HMBC correlations observed for the major form of the 2-ABA/Pcl adduct (in d6-DMSO)

	Correlations between	^{1}H	^{13}C		Correlations between	^{1}H	^{13}C		Correlations between	^{1}H	^{13}C
No.	F2 Atom	F1 Atom	F2 (ppm)	F1 (ppm)	No.	F2 Atom	F1 Atom	F2 (ppm)	F1 (ppm)	No.	F2 Atom	F1 Atom	F2 (ppm)	F1 (ppm)
1	1	2	5.77	72.69	27	1	9	5.77	148.47	53	17	18	2.86	26.47
2	4	2	2	72.73	28	2	9	4.06	148.42	54	19	18	1.03	26.58
3	4	2	1.84	72.83	29	7	9	5.42	148.5	55	20	18	1.48	26.36
4	5	2	3.78	72.59	30	11	9	7.09	148.43	56	20	18	1.34	26.48
5	7	2	5.42	72.72	31	12	9	6.7	148.55	57	17	19	2.86	22.49
6	1	3	5.76	32.36	32	13	9	7.21	148.39	58	17	19	2.37	22.95
7	2	3	4.06	32.2	33	1	10	5.77	115.57	59	18	19	0.72	22.37
8	4	3	1.82	32.22	34	7	10	5.43	115.78	60	18	19	1.07	22.53
9	4	3	2.01	32.42	35	11	10	7.09	115.53	61	20	19	1.49	22.52
10	5	3	3.78	32.41	36	12	10	6.7	115.47	62	20	19	1.34	22.55
11	2	4	4.06	26.55	37	13	10	7.21	115.5	63	21	19	2.93	22.55
12	3	4	1.82	26.41	38	13	11	7.21	129.27	64	18	20	1.02	30.46
13	5	4	3.78	26.55	39	10	12	6.66	118.14	65	18	20	0.73	30.63
14	7	4	5.43	26.38	40	1	13	5.77	128.39	66	19	20	1.02	30.46
15	2	5	4.06	64.16	41	7	13	5.43	128.48	67	20	20	1.5	30.66
16	4	5	1.85	63.96	42	11	13	7.09	128.49	68	21	20	2.93	30.83
17	4	5	2.01	64.04	43	12	13	6.7	128.11	69	19	21	1.02	54.23
18	10	7	6.66	71.92	44	4	14	2	173	70	20	21	1.34	53.98
19	13	7	7.21	71.95	45	4	14	1.84	173.19	71	20	21	1.48	54.02
20	17	7	2.87	71.9	46	5	14	3.78	173.11	72	20	23	1.35	169.78
21	17	7	2.38	71.99	47	17	14	2.87	173.28	73	20	23	1.49	169.78
22	1	8	5.76	120.55	48	17	14	2.37	173.29	74	21	23	2.94	169.73
23	7	8	5.42	120.43	49	18	17	1.04	39.08					
24	10	8	6.67	120.29	50	18	17	0.73	39.4					
25	11	8	7.09	120.35	51	19	17	1.04	39.08					
26	12	8	6.7	120.32	52	17	18	2.37	26.26					

The numbering of the atoms and selected through-bond correlations in the HMBC are shown in Fig. 4*B* and Fig. S3. In contrast to the observations in the Pcl starting material, the two methylene protons on carbon 17 were observed at two different chemical shift values. These protons also showed a heteronuclear through-bond correlation in the HMBC spectra to carbon 7 suggesting the formation of a covalent bond between nitrogen 16 and carbon 7. The proton on carbon 7 resonated at 5.43 ppm (Fig. 4*B*; proton 7 is the signal of the 2-ABA/Pcl adduct highlighted in Fig. S2*A*) and showed correlations to carbon 2, 8, 9 and 13. Similarly the proton on carbon 2 exhibited a HMBC correlation to carbon 9 and carbon 5. These observations along with all the other NMR correlations for the two samples characterized in D_2O and in D6-DMSO were consistent with the

structure in Fig. 4*B* drawn for the major product of the 2-ABA/Pcl adduct.

A structural characterization of the minor form observed in the reaction mixture (Fig. S2*A*) was also attempted. The analysis was inconclusive because of the low concentration and the fact that it slowly converted to the major form. The results indicated that both minor and major forms have one proton on carbon 7 with very similar chemical shifts (arrows at 5.6 ppm in Fig. S2*A*). Chemical shifts for carbon 7 and through-space ROESY correlations were similar in both forms as well. The two methylene protons on carbon 17 were degenerate, meaning they exhibited the same chemical shift, for the minor form of the 2-ABA/Pcl adduct as well as for unreacted Pcl. These protons however had distinct chemical shifts in the major form of the 2-ABA/Pcl adduct. These protons also showed a HMBC correlation to carbon 7 in both minor and major forms suggesting a covalent bond between carbon 7 and nitrogen 16. The observation of degenerate chemical shifts of the methylene protons on carbon 17 in the minor form may imply that the covalent bond between carbon 7 and nitrogen 16 could be semi-stable, and that the NMR observations for the minor form were the result of chemical exchange between two or more species.

A protonated form of the 2-ABA/Pcl adduct (Fig. S3*B*) similar to the proposed structure of the adduct of Δ^1-pyrroline-5-carboxylate with 2-ABA (8) does not appear to be consistent with the minor form because of the presence of the HMBC correlation of the methylene protons on carbon 17 to carbon 7 and the lack of a significant chemical shift difference for carbon 7 (as expected upon the protonation of nitrogen 6; see Section 3.7 and Table S11) between minor and major form.

3.5. NMR Characterization of the 2-ABA/Pcl Adduct after Reduction

Solutions of synthetic Pcl and 2-ABA in D_2O were mixed as described above and the reaction was allowed to proceed to completion as determined by NMR. Aliquots of a freshly prepared sodium cyanoborohydride ($NaBH_3CN$) solution in D_2O were subsequently added to the NMR tube and generated a new species upon the reduction of the 2-ABA/Pcl adduct. The intensity of the unique 2-ABA/Pcl adduct resonance at 5.6 ppm (Fig. S2*A*, arrows) decreased as more $NaBH_3CN$ aliquots were added. The resonance eventually completely disappeared suggesting complete reduction. The NMR sample of the final reduced Pcl/2-ABA adduct was repeatedly lyophilized and redissolved in dry D_2O in order to remove residual HOD and to concentrate the sample. The final sample was reconstituted in 0.5 mL of dry D_2O. Signal assignments of protons and carbon resonances and heteronuclear through-bond correlations from the ^1H-^{13}C HMBC spectra are summarized in Table S8. The numbering of the atoms is shown in Fig. 4*C* and Fig. S3. The key differences compared to the major form of the non-reduced 2-ABA/Pcl adduct were the lack of a heteronuclear through-bond correlation between the methylene protons on carbon 17 and carbon 7 and the absence of a through-bond correlation between the protons on carbon 2 and carbon 9. Comparing the integrals of the NMR signals

between the reduced and the non-reduced 2-ABA/Pcl adduct (Table S9) demonstrated that the number of protons on carbon 7 increased from one to two upon treatment with $NaBH_3CN$. Chemical shift differences (Table S9) and all other NMR results were consistent with the structure shown in Fig. 4*C* and in Fig. S3*C* for the reduced 2-ABA/Pcl adduct.

Table S8. ^{1}H-^{13}C HMBC correlations observed for the reduced form of the 2-ABA/Pcl adduct (in D_2O)

	Correlations between		^{1}H	^{13}C		Correlations between		^{1}H	^{13}C		Correlations between		^{1}H	^{13}C
No.	F2 Atom	F1 Atom	F2 (ppm)	F1 (ppm)	No.	F2 Atom	F1 Atom	F2 (ppm)	F1 (ppm)	No.	F2 Atom	F1 Atom	F2 (ppm)	F1 (ppm)
1	4	2	1.78	54.88	24	12	8	6.87	124.04	47	17	18	2.97	27.95
2	4	2	2.29	55.04	25	7	9	3.89	145.27	48	19	18	1.31	27.85
3	7	2	3.58	54.79	26	7	9	3.58	145.3	49	20	18	1.83	27.91
4	7	2	3.89	54.86	27	11	9	7.24	145.3	50	17	19	2.97	21.71
5	2	3	3.27	23.34	28	12	9	6.87	145.33	51	18	19	1.35	21.66
6	2	3	2.6	23.37	29	13	9	7.21	145.25	52	18	19	1.42	21.69
7	4	3	1.76	23.3	30	11	10	7.25	116.51	53	20	19	1.83	21.71
8	4	3	2.29	23.39	31	12	10	6.87	116.63	54	20	19	1.88	21.77
9	5	3	3.22	23.26	32	12	11	6.87	128.99	55	21	19	3.72	21.71
10	2	4	3.27	30.38	33	13	11	7.22	129.05	56	18	20	1.35	30.27
11	5	4	3.21	30.37	34	10	12	6.9	119.25	57	18	20	1.42	30.28
12	2	5	3.27	67.08	35	7	13	3.58	130.66	58	19	20	1.3	30.15
13	4	5	2.3	66.97	36	7	13	3.89	130.71	59	19	20	1.35	30.27
14	4	5	1.77	67.02	37	11	13	7.24	130.59	60	21	20	3.72	30.16
15	7	5	3.89	66.98	38	12	13	6.87	130.69	61	19	21	1.31	54.42
16	7	5	3.58	67.05	39	4	14	1.76	177.53	62	20	21	1.83	54.63
17	2	7	2.61	56.85	40	4	14	2.29	177.58	63	20	21	1.89	54.67
18	5	7	3.21	56.79	41	5	14	3.21	177.47	64	20	23	1.85	175.18
19	13	7	7.22	56.83	42	17	14	2.97	177.56	65	20	23	1.9	175.23
20	7	8	3.58	123.98	43	18	17	1.37	38.49	66	21	23	3.73	175.21
21	7	8	3.89	124.01	44	18	17	1.42	38.57					
22	10	8	6.9	124.02	45	19	17	1.37	38.49					
23	11	8	7.25	124.05	46	19	17	1.34	38.59					

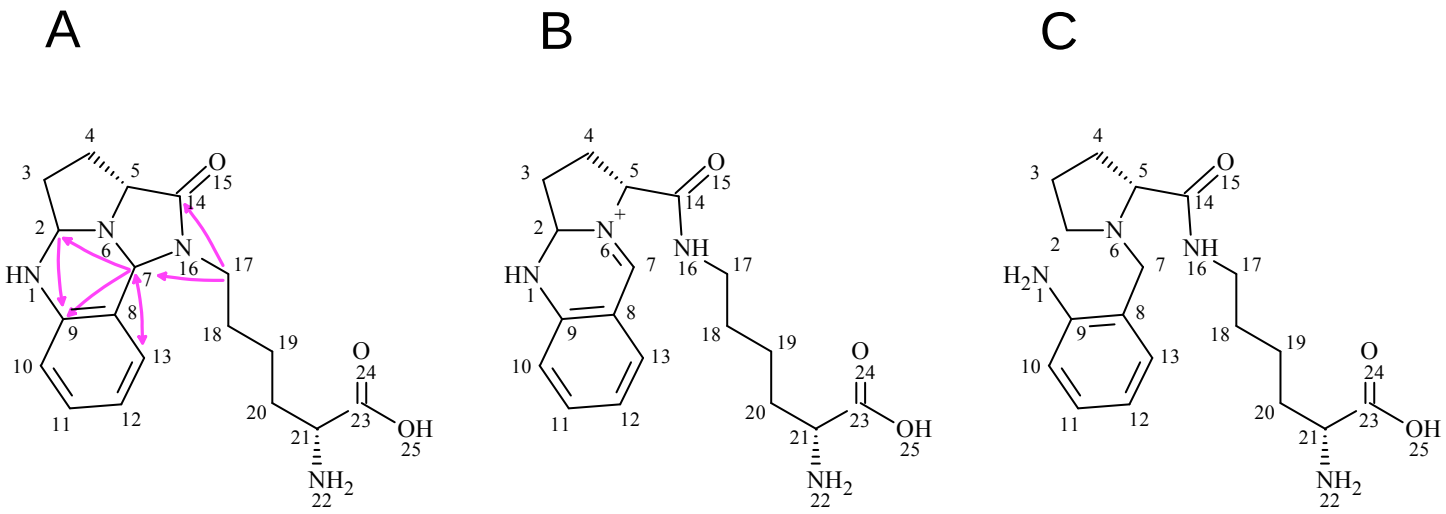


Fig. S3. Structures and numbering schemes for 2-ABA/Pcl adducts as characterized by NMR. (*A*) Major form of the 2-ABA/Pcl adduct with characteristic HMBC through-bond interactions highlighted. (*B*) A protonated form of the 2-ABA/Pcl adduct similar to the proposed structure of the adduct of Δ^1-pyrroline-5-carboxylate with 2-ABA (8). (*C*) 2-ABA/Pcl adduct after reduction.

Table S9. Comparison of reduced and non-reduced 2-ABA/Pcl adduct: Proton signal assignments and integrals from 1D 1H NMR spectra of samples in D_2O. Integrals were calculated relative to proton 17.

	Adduct		Reduced Form	
Atom Number	Shift (ppm)	Integrals	Shift (ppm)	Integrals
2a	4.17	1.05	3.26	1.20
2b	-	-	2.61	1.15
3a	2.01	2.12	1.91	1.02
3b			1.84	1.09
4a	2.24	1.11	2.27	0.92
4b	2.01	1.06	1.77	1.19
5	4.06	0.99	3.21	1.00
7a	5.57	1.00	3.89	1.01
7b	-	-	3.58	1.15
10	6.83	1.09	6.9	0.94
11	7.28	1.24	7.25	1.13
12	6.94	1.23	6.87	0.95
13	7.37	1.28	7.22	1.00
17a	2.91	1.06	2.97	2.11
17b	2.75	1.01		
18a	1.04	1.04	1.41	1.15
18b	0.72	1.04	1.32	1.06
19a	1.04	1.04	1.32	1.06
19b	1.04	1.04	1.32	1.06
20a	1.54	1.05	1.91	1.02
20b	1.54	1.05	1.84	1.09
21	3.46	1.13	3.71	1.17

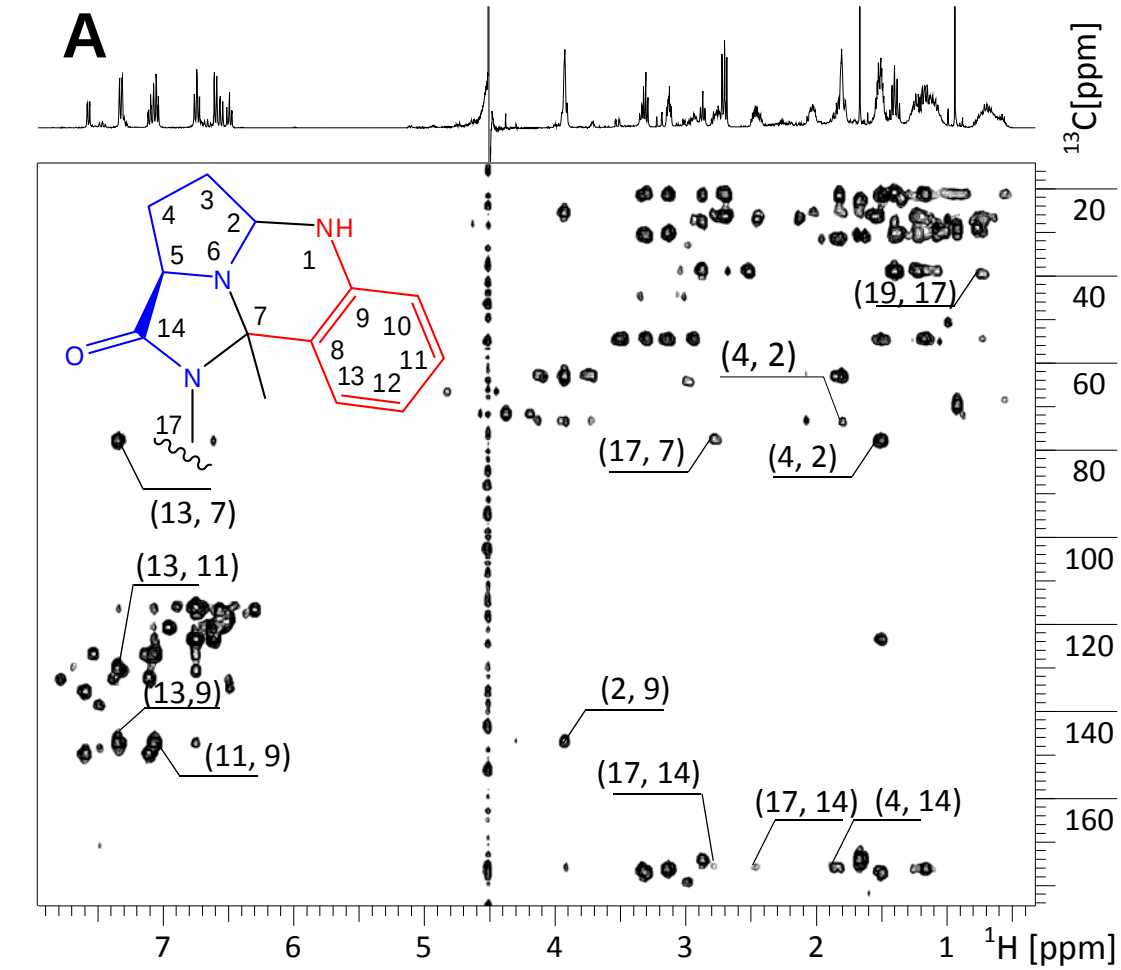
3.6. NMR Characterization of the 2-AAP/Pcl Adduct

Pcl was also reacted with 2-AAP. Briefly, 2-AAP and Pcl were mixed together in D_2O at an approximate 1:1 molar ratio and the reaction was allowed to proceed at room temperature. The reaction mixture contained three forms at approximately 47%, 33% and 20% as determined by NMR. Analysis of the reaction products were complicated by the presence of impurities in the Pcl substrate and consequently the major form was purified by reverse-phase HPLC. ^{1}H-^{13}C HMBC (Fig. S4, Table S10) as well as all other NMR data collected for the major form after purification were consistent with the structure shown in Fig. S4*B*. Similar to the NMR characterization of the 2-ABA/Pcl adduct, the methylene protons on carbon 17 were observed at two different chemical shift values. As in the non-reduced 2-ABA/Pcl adduct, proton H17 (2.77) ppm showed an HMBC correlation to carbon 7 suggesting a covalent bond between nitrogen 16 and carbon 7. The proton on carbon 2 exhibited a HMBC correlation to carbon 9. These observations and all other NMR observations for the sample characterized in D_2O were consistent with the structure in Fig. S4*B* drawn for the major product of the 2-AAP/Pcl adduct. The NMR analysis of the other two forms was inconclusive.

The 2-AAP/Pcl adduct was also compared to the non-reduced and reduced 2-AAP adduct with 3,4-dihydro-2H-pyrrole-2-carboxylic acid (Section 3.7, Fig. S5). In the case of 3,4-dihydro-2H-pyrrole-2-carboxylic acid, the 2-AAP adduct was protonated (Fig. S5*A*) similar to the proposed structure of the adduct of Δ^{1}-pyrroline-5-carboxylate with 2-ABA (8). Because of the protonation of nitrogen 6, the chemical shift for carbon 7 was significantly perturbed (171.6 ppm, Table S11). Chemical shift comparison for carbon 7 with the 2-AAP/Pcl (77.8 ppm) and 2-ABA/Pcl adduct (71.9 ppm) indicated that both adducts were not protonated further supporting their structures as shown in Fig. 4*A*.

Table S10. Through-bond correlations for the 2-AAP/Pcl adduct

	Correlations between		^{1}H	^{13}C		Correlations between		^{1}H	^{13}C		Correlations between		^{1}H	^{13}C
No.	F2 Atom	F1 Atom	F2 (ppm)	F1 (ppm)	No.	F2 Atom	F1 Atom	F2 (ppm)	F1 (ppm)	No.	F2 Atom	F1 Atom	F2 (ppm)	F1 (ppm)
1	3	2	2.07	73.42	15	12	8	6.75	123.3	29	17	18	2.45	27.18
2	3	2	1.8	73.43	16	26	8	1.51	123.44	30	19	18	0.74	26.9
3	4	2	1.8	73.43	17	2	9	3.93	146.83	31	20	18	1.16	27.07
4	5	2	3.92	73.5	18	11	9	7.05	147.2	32	17	19	2.76	21.36
5	2	3	3.93	25.51	19	13	9	7.33	146.99	33	17	19	2.44	21.42
6	5	3	3.93	25.51	20	12	10	6.75	116.33	34	20	19	1.18	21.42
7	2	5	3.93	63.05	21	13	11	7.33	130.15	35	21	19	3.13	21.4
8	3	5	2.08	62.93	22	10	12	6.61	120.87	36	19	20	0.73	29.73
9	3	5	1.82	62.98	23	11	13	7.06	126.72	37	21	20	3.14	30.19
10	4	5	1.82	62.98	24	4	14	1.85	175.89	38	19	21	0.72	54.54
11	13	7	7.33	77.93	25	5	14	3.92	175.65	39	20	21	1.17	54.46
12	17	7	2.77	77.67	26	17	14	2.46	175.69	40	20	23	1.17	176.04
13	26	7	1.51	77.82	27	17	14	2.77	175.85	41	21	23	3.13	176
14	10	8	6.6	123.24	28	19	17	0.73	39.75					



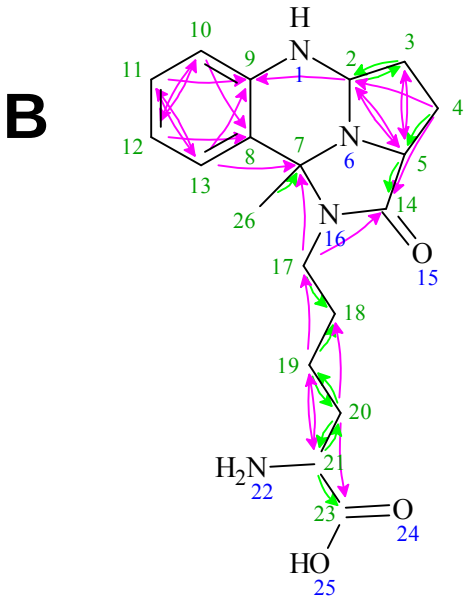


Fig. S4. NMR characterization of 2-AAP/Pcl adduct. (*A*) ^{1}H-^{13}C HMBC spectra with selected through-bond correlations highlighted. See (*B*) for numbering. (*B*) Structure of the 2-AAP/Pcl adduct with characteristic correlations highlighted by arrows – green and magenta are the two and three bond heteronuclear correlations respectively.

3.7. NMR Characterization of the 2-AAP Adduct with 3,4-dihydro-2H-pyrrole-2-carboxylic acid

The product of 3,4-dihydro-2H-pyrrole-2-carboxylic acid reacted with 2-AAP was characterized by ^{1}H and ^{13}C 1D, 2D ^{1}H-^{1}H COSY, ^{1}H-^{1}H ROESY, ^{1}H-^{13}C HMBC and ^{1}H-^{13}C HMQC NMR spectroscopy which were acquired on a 400 MHz Bruker spectrometer with a sample prepared in d6-DMSO. The product was purified from the reaction mixture by reverse-phase HPLC. In contrast to the Pcl adduct reactions, the adduct with 3,4-dihydro-2H-pyrrole-2-carboxylic acid did not form a structure with a fourth ring. There was no indication of a proton at carbon 7 which had a carbon chemical shift of 170.99 ppm and there was an HMBC correlation from proton 5 to carbon 7 (Fig. S5*A*, Table S11). These data indicated the presence of a double bond between carbon 7 and nitrogen 6. Similar to previous observations, the proton on carbon 2 exhibited a HMBC correlation to carbon 9. All the NMR observations for the sample characterized in d6-DMSO were consistent with the structure shown in Fig. S5*A*.

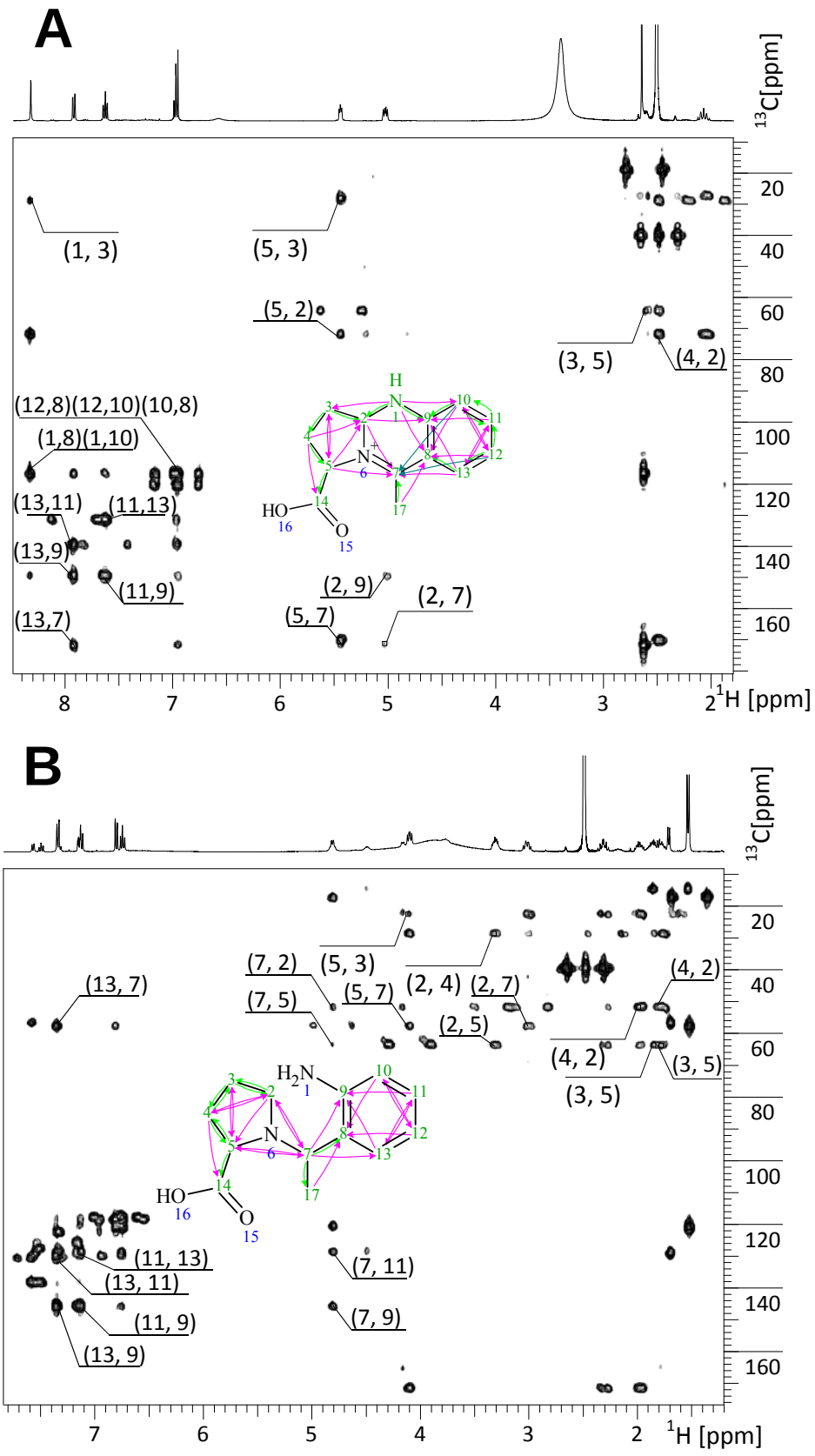
The 2-AAP adduct with 3,4-dihydro-2H-pyrrole-2-carboxylic acid was reduced with $NaBH_3CN$, purified by reverse-phase HPLC, and characterized by NMR spectroscopy as above (Fig. S5*B*, Table S12). Similar to the 2-ABA/Pcl and 2-AAP/Pcl adduct results, the bond between nitrogen 1 and carbon 2 was broken resulting in a large change in chemical shift for carbon 2 (71.8 ppm in the non-reduced compared to 51.7 ppm in the reduced adduct). In parallel, the double bond between carbon 7 and nitrogen 6 was reduced to a single bond. Evidence of a proton on carbon 7 (H7-4.83 ppm) and up-field carbon chemical shift from 170.99 ppm to 62.05 ppm for carbon 7 supported the structure shown in Fig. S5*B*.

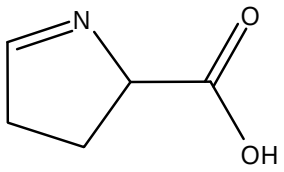
Table S11. ^{1}H-^{13}C HMBC correlations observed for the non-reduced 2-AAP adduct of 3,4-dihydro-2H-pyrrole-2-carboxylic acid (in d6-DMSO)

	Correlations between	^{1}H	^{13}C			Correlations between	^{1}H	^{13}C			Correlations between	^{1}H	^{13}C	
No.	F2 Atom	F1 Atom	F2 (ppm)	F1 (ppm)	No.	F2 Atom	F1 Atom	F2 (ppm)	F1 (ppm)	No.	F2 Atom	F1 Atom	F2 (ppm)	F1 (ppm)
1	1	2	8.32	71.69	14	10	7	6.96	171.37	27	13	9	7.92	149.22
2	3	2	2.61	71.52	15	12	7	6.96	171.37	28	1	10	8.32	116.39
3	3	2	2.06	71.83	16	13	7	7.92	171.51	29	11	10	7.63	116.14
4	4	2	2.5	71.74	17	17	7	2.64	171.57	30	12	10	6.98	116.21
5	5	2	5.44	71.71	18	1	8	8.32	116.39	31	10	11	6.97	139.06
6	1	3	8.32	28.81	19	10	8	6.98	116.21	32	12	11	6.97	139.06
7	5	3	5.44	28.08	20	12	8	6.98	116.21	33	13	11	7.92	139.25
8	3	4	2.63	27.16	21	13	8	7.92	116.56	34	10	12	6.96	120.16
9	3	4	2.06	27.32	22	17	8	2.64	116.22	35	11	13	7.63	131.24
10	3	5	2.61	64.24	23	1	9	8.32	148.95	36	12	13	6.97	130.31
11	4	5	2.51	64.31	24	2	9	5.01	149.55	37	4	14	2.5	169.86
12	2	7	5.03	171.2	25	10	9	6.96	149.19	38	5	14	5.44	170.17
13	5	7	5.44	170.17	26	11	9	7.62	149.08					

Table S12. ^{1}H-^{13}C HMBC correlations observed for the reduced 2-AAP adduct of 3,4-dihydro-2H-pyrrole-2-carboxylic acid (in d6-DMSO)

Correlations between ^{1}H ^{13}C					Correlations between ^{1}H ^{13}C					Correlations between ^{1}H ^{13}C				
No.	F2 Atom	F1 Atom	F2 (ppm)	F1 (ppm)	No.	F2 Atom	F1 Atom	F2 (ppm)	F1 (ppm)	No.	F2 Atom	F1 Atom	F2 (ppm)	F1 (ppm)
1	3	2	1.82	51.7	15	3	5	1.85	63.62	29	7	9	4.82	145.47
2	3	2	1.77	51.78	16	3	5	1.78	63.74	30	11	9	7.14	145.49
3	4	2	1.99	51.72	17	4	5	2.31	63.58	31	13	9	7.34	145.53
4	4	2	2.29	51.91	18	4	5	2	63.85	32	12	10	6.74	117.47
5	7	2	4.82	51.69	19	7	5	4.82	63.63	33	13	11	7.35	129.68
6	2	3	3.01	22.69	20	2	7	3.03	57.77	34	10	12	6.81	118.32
7	4	3	2.32	22.46	21	5	7	4.11	57.7	35	7	13	4.83	128.43
8	4	3	1.98	22.53	22	10	7	6.8	57.83	36	11	13	7.13	128.43
9	5	3	4.11	22.35	23	13	7	7.34	57.72	37	4	14	1.99	171.34
10	2	4	3.32	28.62	24	17	7	1.54	57.89	38	4	14	2.31	171.42
11	3	4	1.84	28.67	25	7	8	4.82	120.15	39	5	14	4.11	171.33
12	3	4	1.76	28.85	26	10	8	6.8	120.43	40	7	17	4.82	17.44
13	5	4	4.1	28.74	27	12	8	6.75	120.05					
14	2	5	3.31	63.71	28	17	8	1.53	120.5					





3,4-dihydro-2H-pyrrole-2-carboxylic acid

Fig. S5. NMR characterization of 2-AAP adduct with 3,4-dihydro-2H-pyrrole-2-carboxylic acid in its non-reduced (*A*) and reduced form (*B*). ^{1}H-^{13}C HMBC spectra with selected through-bond correlations indicated. Green and magenta arrows correspond to two and three bond correlations respectively. See inserts for numbering and structures of the respective adducts with characteristic correlations highlighted by arrows.

4. Proposed Reaction Mechanism of 2-ABA/Pcl and 2-APP/Pcl Adduct Formation

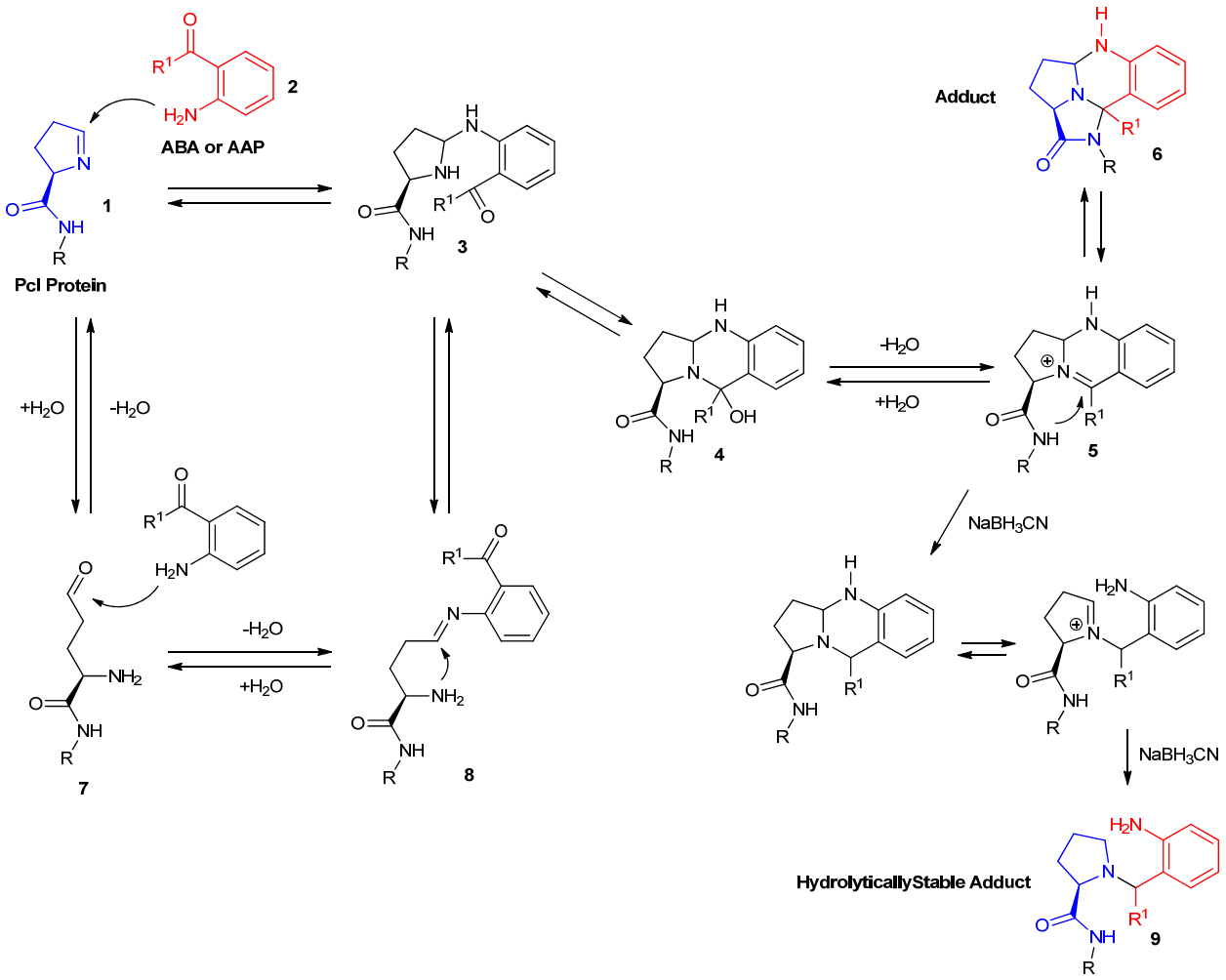
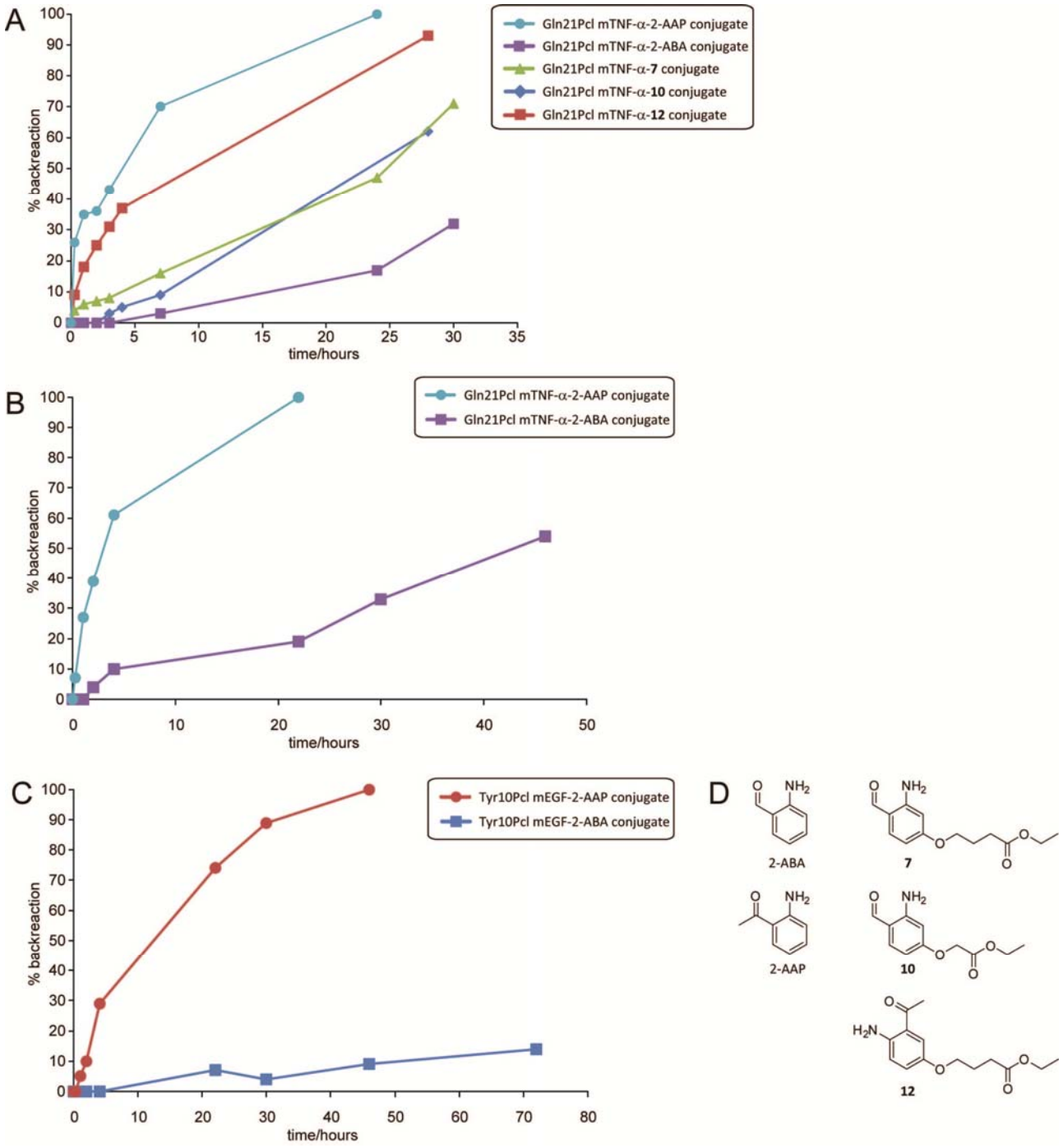


Fig. S6. Proposed reaction mechanism of 2-ABA/Pcl and 2-APP/Pcl adduct formation. Adduct formation likely begins with addition of the amino group of 2-ABA or 2-AAP to the imine moiety of Pcl, followed by formation of the second ring (**4**) via intramolecular cyclization, formation of iminium intermediate **5** via dehydration, and finally intramolecular Mannich reaction to close the third ring, resulting in adduct **6**. Alternatively, the pyrroline ring of Pcl may open to form linear amino aldehyde **7**, to which 2-ABA or 2-AAP adds to form linear imine **8**, followed by intramolecular cyclization to intermediate **3**. Net one water molecule is lost in the adduct formation by either routes. All the steps involved are reversible and adduct **6** can be hydrolyzed back to Pcl and 2-ABA, or 2-AAP. Adduct **6** can be reduced to hydrolytically stable adduct **9** by two hydrides from $NaBH_3CN$ via iminium intermediate **5**, which likely exists in minute amount in equilibrium with adduct **6**. We observe that 2-ABA or 2-AAP conjugation reactions with Pcl proteins turn yellow and the yellow color disappear upon reduction with $NaBH_3CN$, suggesting the presence of yellow species in the reaction system, which is speculated to be iminium **5**. The intensity of the yellow color qualitatively correlates to yields of conjugation reactions. The yellow color appears within 5 minutes under pH 5, consistent with the proposed reaction mechanism where the step(s) from **1** to **3** as well as the step from **4** to **5** are all proton catalyzed. R^1=H, 2-ABA; R^1=CH_3, 2-AAP.

5. Stability of 2-ABA/Pcl and 2-AAP/Pcl Adducts



E

Reaction conditions	Observed mass (calculated mass) [Da]	
	Uncoupled Leu2222Pcl hFAS-TE	2-AAP-Leu2222Pcl hFAS-TE conjugate
a) $NaBH_3(CN)$-reduced	Not detected (33253)	33373.20 (33370)
b) $NaBD_3(CN)$-reduced	Not detected (33255)	33374.80 (33372)
c) Non-reduced	33254.40 (33251)	Not detected (33368)

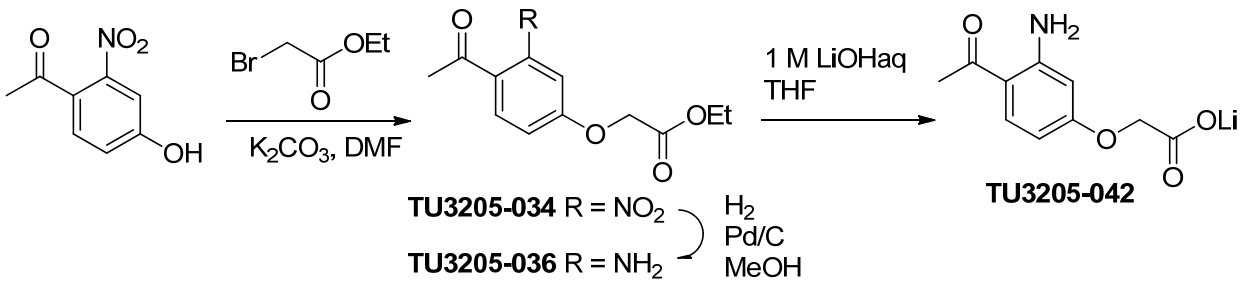
Fig. S7

Fig. S7. Stability of Pcl protein conjugates with 2-AAP and 2-ABA and their derivatives. Stability measurements of non-reduced Pcl protein adducts (*A-D*): (*A*) Mass spectrometric determination of the stability of small molecule Pcl adducts with Gln21Pcl mTNF-α. After confirming quantitative adduct formation, excess reactant was removed by size exclusion chromatography using NAP-5 column (GE Healthcare). The stability of the Gln21Pcl mTNF-α protein conjugates was monitored by LC-MS over time at 37°C and pH 7.4. Data points correspond to the normalized peak intensity of the charge-state deconvoluted ESI/MS spectrum. Side-by-side comparison of the stability of 2-ABA and 2-AAP conjugates of Gln21Pcl mTNF-α (*B*) and of Tyr10Pcl mEGF (*C*). (*D*) Structures of the compounds used for derivatization of Gln21Pcl mTNF-α and Tyr10Pcl mEGF. Stability of Pcl protein conjugates after reduction (*E*): LC-MS analysis indicated that Pcl conjugate formation with 2-AAP can be made irreversible by $NaBH_3CN$ reduction. Leu2222Pcl hFAS-TE was quantitatively derivatized with 2-AAP, excess reagent was removed and LC-MS data was acquired after incubating the samples for 4 days at 22 °C. Unmodified protein was not detectable in the samples reduced with $NaBH_3CN$ (a) or with $NaBD_3CN$ (b) indicating the formation of a hydrolytically stable form. In the non-reduced control sample (c) only unmodified protein was detectable indicating complete reversion of the conjugation reaction.

Sample preparation for experiment shown in **Fig. S7*E***: Conjugation was carried out by adding 1 mM of 2-AAP to 10 μM of Leu2222Pcl hFAS-TE in PBS at pH 7.4. After 24 h at 22 °C, the Pcl protein conjugate was buffer-exchanged into 10 mM sodium phosphate buffer, pH 7.5 using NAP-5 column (GE Healthcare) to remove excess reagent 2-AAP. One aliquot of the protein conjugate was set aside as control sample (c). The remaining Pcl protein conjugate was concentrated to about 100 μM using an Amicon Ultra-4 Centrifugal Filter Unit with 10 kDa cutoff (Millipore Corporation, Bedford, MA). Freshly prepared solutions of 200 mM $NaBH_3CN$ or $NaBD_3CN$ (dissolved in 10 mM phosphate buffer, pH 7.5) were added to a final concentration of 20 mM. All three samples were then incubated at 22 °C for 4 days followed by LC-MS analysis.

6. Synthesis of Pcl Reactive Precursors for the Preparation of Conjugates

6.1. Synthesis of lithium 2-(4-acetyl-3-aminophenoxy)acetate (TU3205-042)



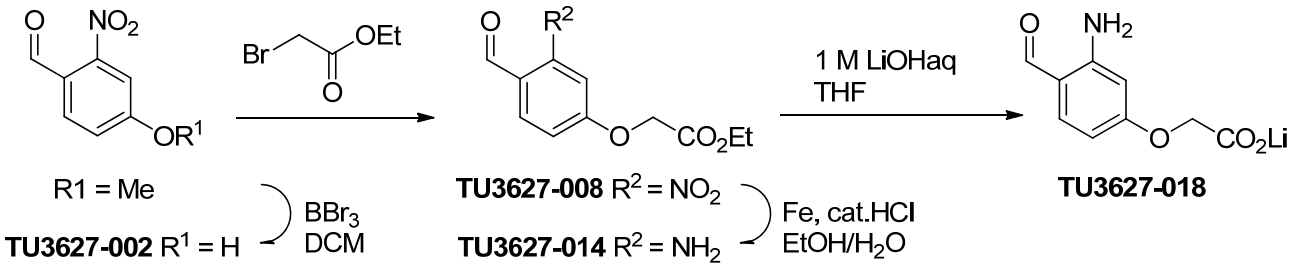
1-(4-Hydroxy-2-nitrophenyl)ethanone (Carbocore, 181 mg, 1.00 mmol), ethyl 2-bromoacetate (183 mg, 1.10 mmol), K_2CO_3 (138 mg. 1.00 mmol) and DMF (5 mL) were combined in a 20 mL glass vial and stirred at 60 ^oC for 2 h, at which point LC-MS analysis showed a clean complete reaction. The reaction mixture was diluted with water and extracted with EtOAc, washed with sat. NaCl(aq). The reaction was repeated using 1-(4-hydroxy-2-nitrophenyl)ethanone (0.802 g, 4.43 mmol), ethyl 2-bromoacetate (0.813 g, 4.87 mmol), K_2CO_3 (0.612 g. 4.43 mmol) and DMF (25 mL), and worked up in the same way. The combined EtOAc extracts were dried over anhydrous $MgSO_4$, filtered and concentrated under reduced pressure, affording ethyl 2-(4-acetyl-3-nitrophenoxy)acetate (TU3205-034) as dark yellow oil. MS (ESI^+) (m/z): calcd. for $C_{12}H_{14}NO_6^+$ (MH^+) 268.08, found 268.10 . ^{1}H-NMR (400 MHz, $CDCl_3$, δ): 1.317 (t, *J*=7.2 Hz, 3H), 2.512 (3H, s), 4.294 (q, *J*=7.2 Hz, 2H), 4.722 (s, 2H), 7.192 (dd, *J*=8.4, 2.4 Hz, 1H), 7.455 (d, *J*=8.4 Hz, 1H), 7.462 (d, *J*=2.8 Hz, 1H). ^{13}C-NMR (100 MHz, $CDCl_3$, δ): 14.133, 29.695, 61.922, 65.499, 110.105, 119.735, 129.461, 130.154, 147.793, 159.356, 167.474, 198.352.

TU3205-034 (111 mg) in MeOH (5 mL) was hydrogenated using 10% palladium on charcoal (11 mg) under atomospheric pressure of H_2 at ambient temperature. After 30 min LC-MS analysis revealed a clean complete reaction. The reaction was repeated using TU3205-034 (1.45 g), 10% palladium on charcoal (140 mg), and MeOH (80 mL). The two reaction mixtures were combined, and the spent catalyst was removed by filtration through a celite pad. The filtrate was concentrated under reduced pressure, affording ethyl 2-(4-acetyl-3-aminophenoxy)acetate (TU3205-036) as dark yellow oil. MS (ESI^+) (m/z): calcd. for $C_{12}H_{16}NO_4^+$ (MH^+)238.11, found 238.10. ^{1}H-NMR (400 MHz, $CDCl_3$, δ): 1.295 (t, *J*=7.0 Hz, 3H), 2.507 (s, 3H,), 4.269 (q, *J*=7.2 Hz, 2H), 4.603 (s, 2H), 6.047 (d, *J*=2.8 Hz, 1H), 6.236 (dd, *J*=8.8, 2.8 Hz, 1H), 6.396 (br.s, 2H), 7.647 (d, *J*=9.2 Hz, 1H). ^{13}C-NMR (100 MHz, $CDCl_3$, δ): 14.135, 27.658, 61.532, 64.939, 100.237, 104.030, 113.586, 134.286, 152.468, 162.294, 168.310, 199.055.

TU3205-036 (1.02 g, 4.30 mmol) was dissolved in THF (17 mL) and treated with 1M LiOH(aq) (4.3 mL) at ambient temperature for 1 h. LC-MS analysis showed a clean complete reaction. The reaction mixture

was concentrated under reduced pressure, affording lithium 2-(4-acetyl-3-aminophenoxy)acetate (TU3205-042) as yellowish solid. MS (ESI^+) (m/z): calcd. for $C_{10}H_{12}NO_4^+$ (H-form, MH^+) 210.08, found 210.10.

6.2. Synthesis of lithium 2-(3-amino-4-formylphenoxy)acetate (TU3627-018)



To 4-methoxy-2-nitrobenzaldehyde (CarboCore, 5.5 g, 30.4 mmol) and 200 mL DCM in a 500 mL round-bottom flask was added borontribromide (24 g, 96 mmol) dropwise with cooling in an ice bath. The reaction was stirred at the same temperature for 30 min and then at ambient temperature for 3 h. The reaction mixture was carefully poured into ice water and let stand at ambient temperature for 3 days. The aqueous mixture was extracted with EtOAc, washed with sat. NaCl(aq), dried over Na_2SO_4, filtered and concentrated under reduced pressure. The crude material was purified by a SiO_2 gel flash chromatography using a linear gradient of 20 to 60% EtOAc in hexanes (Rf. 0.32, 50% EtOAc in hexanes), affording 4-hydroxy-2-nitrobenzaldehyde (TU3627-002) as orange crystals. MS (ESI+) (m/z): calcd. for $C_7H_6NO_4^+$ (MH^+) 168.02, found 168.10. ^{1}H-NMR (400 MHz, DMSO-d6, δ): 7.211 (dd, *J*=2.4, 8.4 Hz, 1H), 7.362 (d, *J*=2.4 Hz, 1H), 7.866 (d, *J*=8.8 Hz, 1H), 9.998 (s, 1H,), 11.466 (s, 1H). ^{13}C-NMR (100 MHz, DMSO-d6, δ): 110.726, 119.683, 120.770, 132.616, 151.225, 162.650, 187.885.

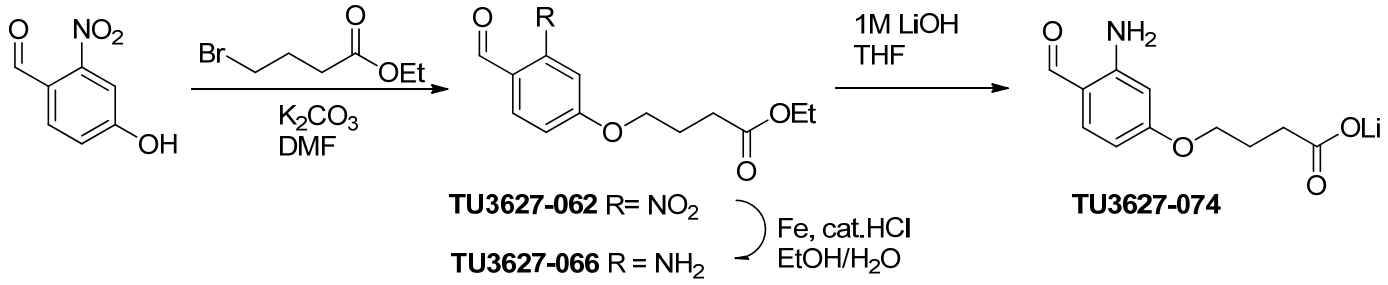
TU3627-002 (1.87 g, 11.2 mmol), ethyl bromoacetate (Aldrich, 1.5 mL, 13 mmol), K_2CO_3 (2.32 g, 16.8 mmol) and DMF (20 mL) were combined and stirred at ambient temperature for 18 h. The reaction mixture was partitioned between EtOAc and H_2O. The organic layer was separated, washed with sat NaCl(aq), dried over $MgSO_4$, filtered and concentrated under reduced pressure. The crude material was purified by a SiO_2 gel flash chromatography using a linear gradient of 5 to 35% EtOAc in hexanes (Rf. 0.32, 35% EtOAc in hexanes), affording ethyl 2-(4-formyl-3-nitrophenoxy)acetate (TU3627-008) as dark yellow oil. MS (ESI^+) (m/z): calcd. for $C_{11}H_{12}NO_6^+$ (MH^+)254.1, found 254.1. ^{1}H-NMR (400 MHz, $CDCl_3$, δ): 1.304 (t, *J*=7.0 Hz, 3H), 4.285 (q, *J*=7.2 Hz, 2H), 4.767 (s, 2H), 7.233 (dd, *J*=2.0, 8.4 Hz, 1H), 7.522 (d, *J*=2.4 Hz, 1H), 7.967 (d, *J*=8.4 Hz, 1H), 10.286 (s, 1H). ^{13}C-NMR (100 MHz, $CDCl_3$, δ): 14.082, 61.971, 65.482, 69.302, 110.427, 119.520, 124.321, 131.525, 151.286, 161.66, 167.172, 186.841.

TU3627-008 was reduced by the method described by Merlic(9). Specifically, TU3627-008 (1.717 g, 6.78 mmol), iron powder (3.79 g, 67.8 mmol), EtOH (45 mL), H_2O (11 mL) and c.HCl (180 μL) were

combined and heated at reflux for 2 h. The reaction mixture was filtered, and the filtrate was concentrated under reduced pressure. The residue was purified by SiO_2 flash chromatography using a linear gradient of 5 to 60% solvent B in solvent A (solvent A: 5% NEt_3 in hexanes; solvent B: 5% NEt_3 in EtOAc) (Rf. 0.34, 35% EtOAc in hexanes), affording ethyl 2-(3-amino-4-formylphenoxy)acetate (TU3627-014) as light yellow crystals. MS (ESI^+) (m/z): calcd. for $C_{11}H_{14}NO_4^+$ (MH^+) 224.10, found 224.10. ^{1}H-NMR (400 MHz, DMSO-d6, δ): 1.216 (t, *J*=7.2 Hz, 3H), 4.173 (q, *J*=7.2 Hz, 2H), 4.777 (s, 2H), 6.168 (d, *J*=2.4 Hz, 1H), 6.248 (dd, *J*=2.4, 8.8 Hz, 1H), 7.183 (br.s, 2H), 7.441 (d, *J*=8.4 Hz, 1H), 9.646 (s, 1H). ^{13}C-NMR (100 MHz, DMSO-d6, δ): 13.976, 60.711, 64.354, 98.521, 103.988, 113.256, 137.699, 152.665, 162.849, 168.176, 191.670.

TU3627-014 (0.608 g, 2.72 mmol) in 5 mL THF was treated with 1M LiOH(aq) (2.72 mL) at ambient temperature. After 20 min the LC-MS analysis revealed a complete reaction. The reaction mixture was concentrated under reduced pressure, affording lithium 2-(3-amino-4-formylphenoxy)acetate (TU3627-018) as yellow solid. MS (ESI^+) (m/z): calcd. for $C_9H_{10}NO_4^+$ (H-form, MH^+) 196.06, found 196.1. MS (ESI^+) (m/z). ^{1}H-NMR (400 MHz, DMSO-d6, δ): 4.132 (s, 2H), 6.110 (d, *J*=2.0 Hz, 1H), 6.145 (dd, *J*=2.0, 8.8 Hz, 1H), 7.155 (br.s, 2H), 7.33 (d, *J*=8.8 Hz, 1H), 9.576 (s, 1H). ^{13}C-NMR (100 MHz, DMSO-d6, δ): 67.675, 98.354, 105.031, 112.386, 137.042, 152.914, 164.589, 169.349, 191.056.

6.3. Synthesis of lithium 4-(3-amino-4-formylphenoxy)butanoate (TU3627-074).



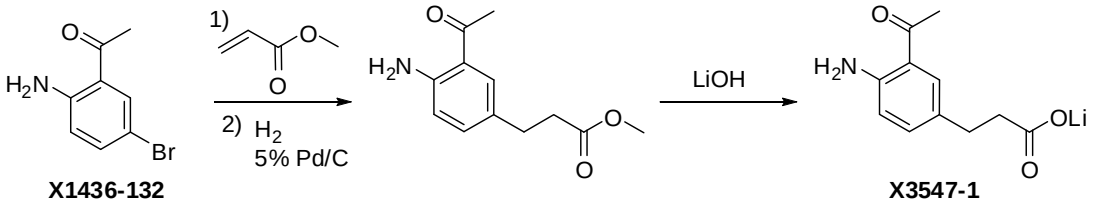
A mixture of 2-nitro-4-hydroxybenzaldehyde (2.85 g, 17.1 mmol), ethyl 4-bromobutanoate (3.66 g, 18.8 mmol), K_2CO_3 (2.84 g, 20.5mmol) and DMF (20 mL) were stirred at ambient temperature for 50 h. The reaction mixture was partitioned between H_2O and EtOAc. The organic layer was separated and washed with sat. $NaHCO_3$(aq). The combined aqueous layers were extracted with EtOAc. The combined organic layers were washed with dilute citric acid(aq), sat NaCl(aq), dried over $MgSO_4$, filtered and concentrated under reduced pressure. The residue was purified by SiO_2 flash chromatography using a linear gradient of 20 to 40%EtOAc in hexanes (Rf: 0.35, 35% EtOAc in hexanes), affording ethyl 4-(4-formyl-3-nitrophenoxy)butanoate (TU3627-062) as yellow oil. MS (ESI^+) (m/z): calcd. for $C_{13}H_{16}NO_6^+$ (MH^+) 282.10, found 282.10. ^{1}H-NMR (400 MHz, $CDCl_3$, δ): 1.261 (t, *J*=7.2 Hz, 3H), 2.170 (quint, *J*=6.8 Hz, 2H), 2.528 (t, *J*=7.2 Hz, 2H), 4.153 (d, *J*=7.2 Hz, 2H), 4.167 (t, *J*=6.4 Hz, 2H), 7.217 (dd, *J*=2.4, 8.4 Hz, 1H), 7.498 (d, *J*=2.4

Hz, 1H), 7.964 (d, *J*=8.8 Hz, 1H), 10.277 (s, 1H). ^{13}C-NMR (100 MHz, $CDCl_3$, δ): 14.190, 24.112, 27.712, 30.291, 32.443, 32.771, 60.648, 68.075, 110.024, 119.397, 123.405, 131.455, 151.572, 162.916, 172.710, 186.943.

TU3627-062 (3.89 g, 13.8 mmol) was reduced as described by Merlic(9). Ethyl 4-(3-amino-4-formylphenoxy)butanoate (TU3627-066) was obtained as a yellow solid after silica gel flash chromatography using a linear gradient of 20 to 60% solvent B in solvent A (solvent A: 5% NEt_3 in hexanes; solvent B: 5% NEt_3 in EtOAc). Rf. 0.51, 50% EtOAc in hexanes. MS (ESI$^+$) (m/z): calcd. $C_{13}H_{18}NO_4^+$(MH^+) 252.12, found 252.20. ^{1}H-NMR (400 MHz, DMSO-d6, δ): 1.177 (t, *J*=7.2 Hz, 3H), 1.962 (quint, *J*=6.8 Hz, 2H), 2.441 (t, *J*=7.2 Hz, 2H), 3.979 (t, *J*=6.4 Hz, 2H), 4.064 (q, *J*=7.2 Hz, 2H), 6.206 (s, 1H), 6.219 (d, *J*=8.8 Hz, 1H), 7.411 (d, *J*=8.8 Hz, 1H), 9.623 (s, 1H). ^{13}C-NMR (100 MHz, DMSO-d6, δ): 14.041, 24.028, 30.002, 59.837, 66.459, 98.075, 104.394, 112.822, 137.558, 152.871, 163.809, 172.404, 191.535.

TU3627-066 (1.00 g), 3.98 mmol in 6 mL THF was treated with 1 M LiOH(aq)(3.98 mL) at ambient temperature for 4 h. Most solvent was removed under reduced pressure, and the resulting cloudy mixture was diluted with doubly deionized water, frozen and lyophilized, affording lithium 4-(3-amino-4-formylphenoxy)butanoate (TU3627-074) as dull yellow solid. MS (ESI$^+$) (m/z): calcd. for $C_{11}H_{14}NO_4^+$ (H-form, MH^+) 224.09, found 224.20 (MH^+). ^{1}H-NMR (400 MHz, DMSO-d6, δ): 1.852 (quint, *J*=6.8 Hz, 2H), 2.040 (t, *J*=6.8 Hz, 2H), 3.949 (t, *J*=6.8 Hz, 2H), 6.191 (dd, *J*=2.4, 8.8 Hz, 1H), 6.239 (d, *J*=2.4 Hz, 1H), 7.208 (br.s, 2H), 7.367 (d, *J*=8.8 Hz, 1H), 9.597 (s, 1H). ^{13}C-NMR (100 MHz, DMSO-d6, δ): 25.624, 33.706, 67.799, 97.997, 104.677, 112.589, 137.427, 153.003, 164.195, 176.575, 191.394.

6.4. Synthesis of lithium 3-(3-acetyl-4-aminophenyl)propanoate (X3547-1).



X1436-132 **X3547-1**

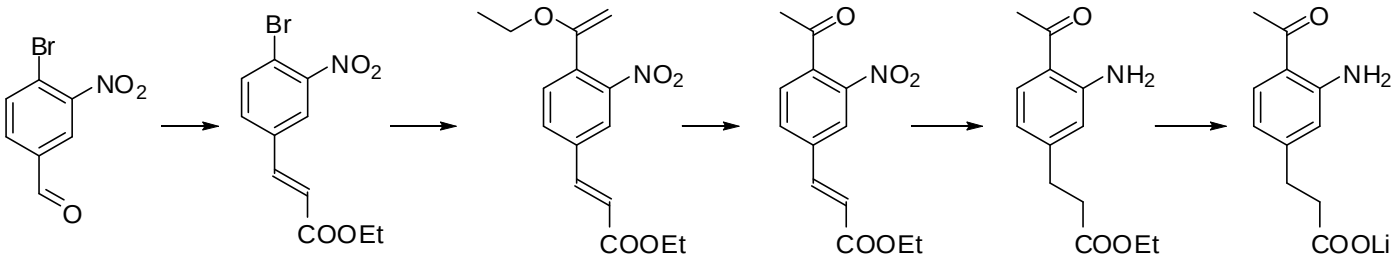
To a mixture of 1-(2-amino-5-bromophenyl)ethanone (642 mg, 3.00 mmol), $Pd(OAc)_2$ (33.7 mg, 0.150 mmol), and P(o-tolyl)$_3$ (137 mg, 0.45 mmol) in DMF(10 mL) in a pressure tube was added methyl acrylate (351 μL, 3.90 mmol) and NEt_3 (1.4 mL). The mixture was flushed with N_2 for 3 min, sealed and heated at 110 °C for 4 h. The reaction mixture was cooled to ambient temperature and partitioned between EtOAc and water. The aqueous layer was extracted once with EtOAc, and the combined organic layers were washed with brine, dried over Na_2SO_4, filtered, and concentrated under reduced pressure. The residue was purified by a SiO_2 flash chromatography (EtOAc/hexanes), affording (*E*)-methyl 3-(3-acetyl-4-aminophenyl)acrylate. MS (ESI$^+$) (m/z): calcd. for $C_{12}H_{14}NO_3^+$ (MH^+)220.10, found 220.24. ^{1}H NMR (400 MHz, MeOD, δ): 7.96 (d, *J*=2.0 Hz, 1H),

7.63 (d, *J*=16.0 Hz, 1H), 6.67 (dd, *J*=8.8, 2.0 Hz, 1H), 6.77 (d, *J*=8.8 Hz, 1H), 6.29 (d, *J*=16.0 Hz, 1H), 3.76 (s, 3H), 2.59 (s, 3H).

(*E*)-Methyl 3-(3-acetyl-4-aminophenyl)acrylate (219 mg, 1.00 mmol) was hydrogenated under atmospheric pressure of hydrogen over 5% Pd on activated carbon (21.9 mg) in MeOH at ambient temperature, affording the product after filtration and concentration (quantitative). The product was used for the next step without further purification. MS (ESI^+) (m/z): calcd. for $C_{12}H_{16}NO_3^+$ (MH^+) 222.11, found 222.25. ^1H NMR (400 MHz, MeOD, δ): 7.59 (d, *J*=2.0 Hz, 1H), 7.13 (dd, *J*=8.4, 2.0 Hz, 1H), 6.67 (d, *J*=8.4 Hz, 1H), 3.64 (s, 3H), 2.81 (t, *J*=7.6 Hz, 2H), 2.60 (t, *J*=7.6 Hz, 2H), 2.54 (s, 3H).

Methyl 3-(3-acetyl-4-aminophenyl)propanoate (221 mg, 0.999 mmol) and 4 M LiOH(aq) (0.275 mL) were added to 3mL THF/H_2O (v/v =3/1). The solvent was removed under reduced pressure after the reaction completed, affording lithium 3-(3-acetyl-4-aminophenyl)propanoate (X3547-1). MS (ESI^+) (m/z): calcd. for $C_{11}H_{14}NO_3^+$ (H-form, MH^+) 208.10, found 208.10. ^1H NMR (400 MHz, MeOD, δ): 7.62 (d, *J*=2.0 Hz, 1H), 7.17 (dd, *J*=8.4, 2.0 Hz, 1H), 6.67 (d, *J*=8.4 Hz, 1H), 2.80 (dd, *J*=7.6, 9.0 Hz, 2H), 2.55 (s, 3H), 2.41 (dd, *J*=7.6, 9.0 Hz, 2H).

6.5. Synthesis of lithium 3-(4-acetyl-3-aminophenyl)propanoate (X3547-8).



X3471-146 **X3471-154A** **X3471-154B** **X3547-2** **X3547-8**

A solution of ethyl triphenylphosphoranylidine acetate (1.742 g, 5.00 mmol) in CH_3CN (15 mL) was added with stirring to 4-bromo-3-nitrobenzaldehyde (1.150 g, 5.00 mmol) in CH_3CN (10 mL). The reaction mixture was heated at reflux overnight. After the mixture was cooled, the solvent was removed under reduced pressure, affording a crude solid. The pure product X3471-146 was obtained as a white solid after silica gel flash column chromatography (hexanes/EtOAc, 9:1) (632.0 mg, 42%).

X3471-146 (632 mg, 2.11 mmol) and $PdCl_2(PPh_3)_2$ (148 mg, 0.21 mmol) were dissolved in DMF (5.0 mL) under N_2 in a Schlenk tube. Tributyl(1-ethoxyvinyl)stannane (711 μL, 2.11mmol) was added with stirring, the mixture was heated at 100 °C overnight. After the mixture was cooled, the solvent was removed under reduced pressure. The residue was diluted with DCM, and washed with H_2O and brine. After removal of DCM, the residue was purified by silica gel flash column chromatography (15% -25% EtOAc in hexanes) to give X3471-154A. X3471-154A was treated with 20 mL 1N HCl(aq) at ambient temperature for 4 hr. Removal of the solvent afforded 286 mg of (*E*)-ethyl 3-(4-acetyl-3-nitrophenyl)acrylate (X3471-154B) (51%, 2

steps). MS (ESI^+) (m/z): calcd. for $C_{13}H_{14}NO_5^+$ (MH^+) 264.09, found 264.07. ^1H NMR (400 MHz, MeOD, δ): 1.34 (t, *J*=7.2 Hz, 3H), 8.31 (d, *J*=1.6 Hz, 1H), 2.57 (s, 3H), 4.27 (q, *J*=7.2 Hz, 2H), 6.74 (d, *J*=16.0 Hz, 1H), 7.67 (d, *J*=8.0 Hz, 1H), 7.76 (d, *J*=16.0 Hz, 1H), 8.04 (dd, *J*'=1.6 Hz, *J*'=8.0 Hz, 1H).

X3471-154B (263 mg, 1.00 mmol) was reduced to ethyl 3-(4-acetyl-3-aminophenyl)propanoate (X3547-2) over 5% Pd on activated carbon (26 mg) in 5.0 mL MeOH under hydrogen at 1 atm to give X3547-2 (quantitative). MS (ESI^+) (m/z): calcd. for $C_{13}H_{18}NO_3^+$(MH^+) 236.13, found 236.30 (MH^+).

The whole product from the previous step (X3547-2) was taken up in THF (3.0 mL) and H_2O (1.0 mL), and treated with 4 M LiOH(aq) (0.248 mL) at ambient temperature. After the reaction was complete indicated by LC-MS, the reaction mixture was lyophilized, affording lithium 3-(4-acetyl-3-aminophenyl)propanoate (X3547-8,). MS (ESI^+) (m/z): calcd. for $C_{11}H_{14}NO_3^+$ (H-form, MH^+) 208.10, found 208.10.

7. Applications. Site-specific Derivatization of Pcl Proteins with Various Entities

To illustrate the versatility of the Pcl/Pyl conjugation chemistry, we conjugated a series of 2-ABA and 2-AAP reagents to Pcl proteins. The coupling reactions were typically carried out in 10x phosphate buffered saline (PBS) at pH 7.4 and 22 °C. The conjugation reaction was started by the addition of 10 μM Pcl-containing protein and 100 μM 2-ABA/2-AAP reagent unless specified differently. Formation of the protein conjugate was verified by electrospray ionization-mass spectrometry (ESI-MS) or matrix-assisted laser desorption/ionization (MALDI). Coupling of 2-ABA-oligonucleotide conjugates was analyzed by gel shift assay using a NuPAGE 4-12% Bis-Tris Gel (Invitrogen, Carlsbad, CA). After coupling, the protein conjugate was dialyzed into 10 mM sodium phosphate buffer (pH 7.5) using 3.5 kDa MW cutoff dialysis cassettes (Thermo Scientific) and concentrated to 100 μM using an Amicon Ultra-4 Centrifugal Filter Unit with 3.5 kDa or 10 kDa MW cutoff (Millipore Corporation, Bedford, MA). A freshly prepared solution of 200 mM $NaBH_3CN$ (dissolved in 10 mM phosphate buffer, pH 7.5) was then added to a final concentration of typically 20 mM. After allowing the reduction reaction to proceed for 1 – 4 hours at 22 °C, the reaction was diluted with six volumes of 10 mM sodium phosphate buffer (pH 7.5). Using a NAP-5 column or PD10 column (GE Healthcare, Piscataway, NJ), the reduced protein conjugate was finally freed from excess $NaBH_3CN$ and buffer exchanged into the desired buffer.

7.1. PEGylation

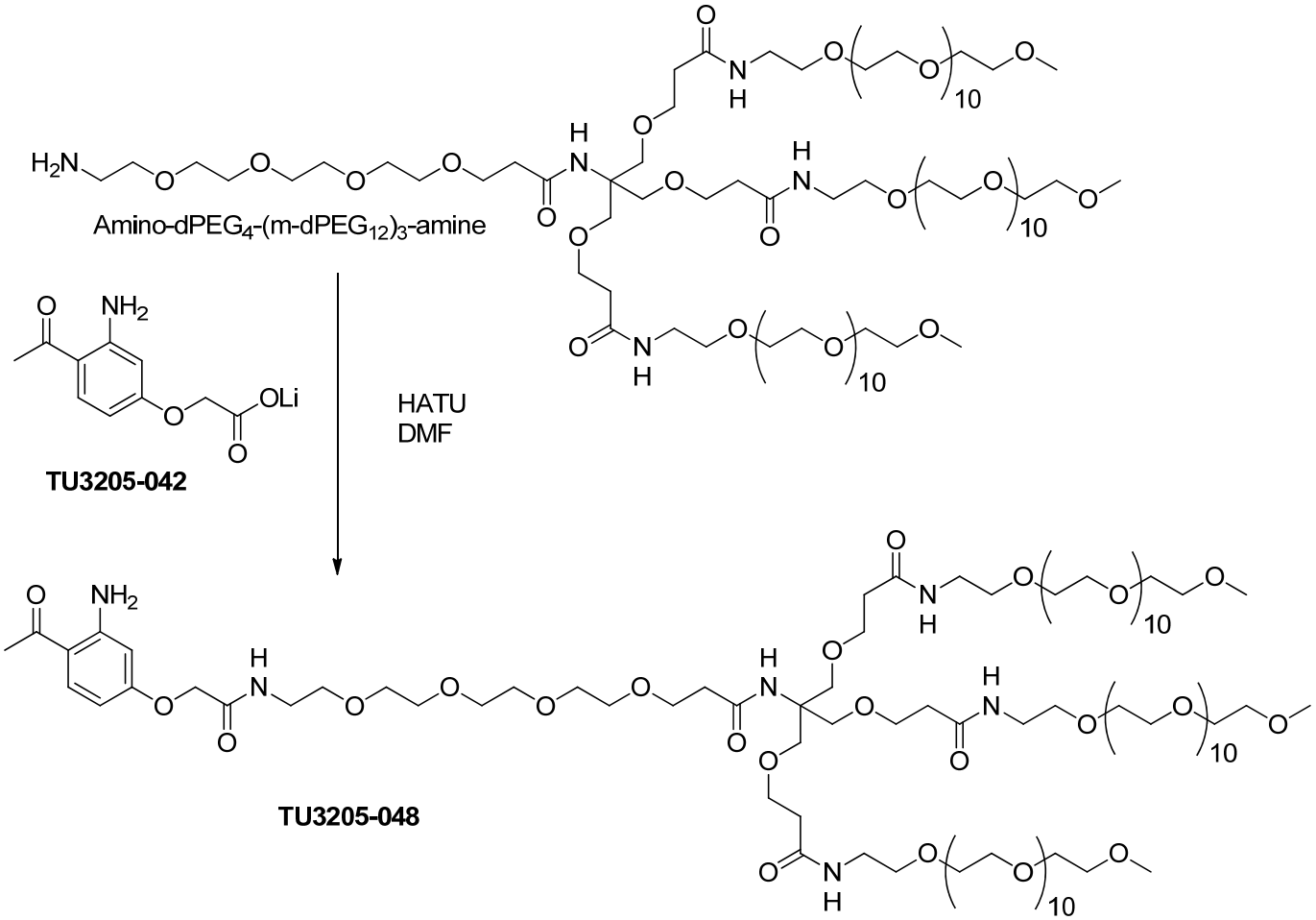
7.1.1. PEGylation reactions

To demonstrate that Pcl protein can be site-specifically and efficiently conjugated with poly(ethylene glycol) polymers (PEGylation), a series of linear and branched PEGs activated with 2-AAP or 2-ABA were

synthesized as described below and tested for conjugation (Fig. 5). Pcl protein PEGylation occured efficiently at pH 7.4 and pH 5.0, at 22 ^{o}C as well as at 4 ^{o}C: For the reactions shown in Fig. 5*A*,*B*, 0.2 mM hFAS-TE Thr2450Pcl was reacted with 1 mM of a 30 kDa-linear 2-ABA (TU3627-084) or 2-AAP-PEG (TU633-120) for 60 hours at 4 ^{o}C at pH 7.4. Similarly, for the reaction shown in Fig. 5*C*,*D*, 10 μM of purified hIgG4 Pcl protein was incubated with 100 μM of 40 kDa-2-ABA-PEG (TU3627-086) in 200 mM Na acetate buffer (pH 5.0) at 22 ^{o}C for 24 hours. $NaBH_3CN$ was added to the mixture at 20 mM at the end of the reactions and the incubation continued at 22 ^{o}C for four more hours. The reaction mixtures were analyzed on SDS-PAGE followed by protein staining.

7.1.2. Synthesis of PEG reagents

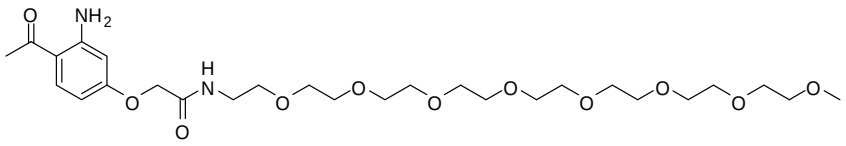
7.1.2.1. Synthesis of 2-AAP-PEGs (TU3205-048, TU3205-044, TU3205-052, TU633-084, TU633-120, TU3627-084)



TU3205-048 (2.4 kDa 2-AAP-PEG)

Lithium 2-(4-acetyl-3-aminophenoxy)acetate (TU3205-042) (12.2 mg) was charged in a 10 mL round bottom flask, and DMF (1mL) and HATU (21.5 mg) were added. The resulting slurry was stirred at the ambient temperature for 35 min. The reaction turned into yellow solution during this period. To the reaction

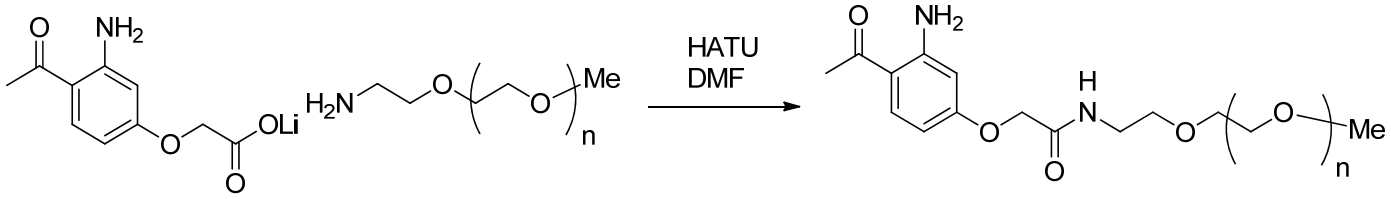
was then added Amino-dPEG$_4$-(m-dPEG$_{12}$)$_3$ (Quanta Biodesign, MW 2209, 100 mg) in 2 mL DMF. The reaction mixture was stirred for 18 h and concentrated under reduced pressure. The residue was purified by a SiO_2 flash chromatography (MeOH in DCM), affording TU3205-048 (2.4 kDa 2-AAP-mPEG) as yellow viscous oil. MS (ESI$^+$) (m/z): calcd. for $C_{109}H_{206}N_6O_{50}$ 800.8 [$(MH_3)^{3+}$/3], 600.8 [$(MH_4)^{4+}$/4], found 800.5, 600.7.



TU3205-044

TU3205-044 (0.5 kDa 2-AAP-PEG)

Lithium 2-(4-acetyl-3-aminophenoxy)acetate (TU3205-042) (55.9 mg, 0.260 mmol) was charged in a 10 mL round bottom flask, and DMF (3 mL) and HATU (98.9 mg, 0.260 mmol) were added. The resulting slurry was stirred at ambient temperature for 40 min. The reaction turned into yellow solution during this period. To the reaction was then added m-dPEG$_8$-amine (Quanta Biodesign, MW 383.5, 100 mg, 0.260 mmol) in 2 mL DMF. After 5 min, LC-MS analysis showed a complete reaction. The reaction mixture was stirred for an additional 40 min and concentrated under reduced pressure. The residue was purified by a SiO_2 flash chromatography (3% MeOH in DCM), affording 129 mg of TU3205-044 (0.5 kDa 2-AAP-mPEG) as yellow viscous oil (86%). MS (ESI$^+$) (m/z): calcd. For $C_{27}H_{47}N_2O_{11}^+$(MH^+) 575.32, found 575.30. ^{1}H-NMR (400 MHz, $CDCl_3$, δ): 2.210 (s, 3H), 3.370 (s, 3H), 3.544 (m. 4H), 3.650 (m, 28H), 4.499 (s, 2H), 6.172 (d, *J*=2.4 Hz, 1H), 6.243 (dd, *J*=8.8, 2.8 Hz, 1H), 6.540 (br.s, 2H). 7.034 (br.s, 1H), 7.649 (d, *J*=8.8 Hz, 1H).

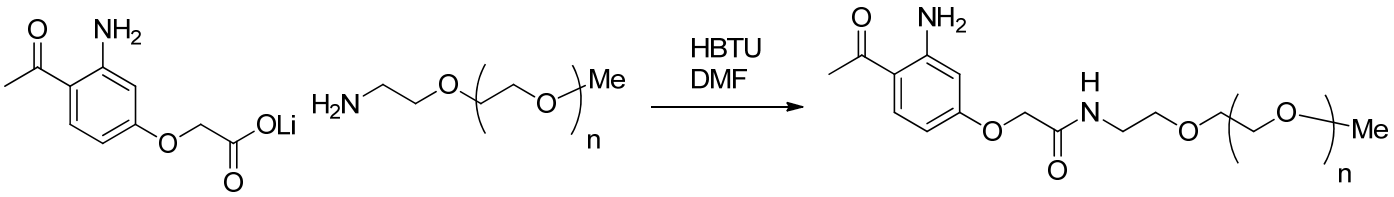


TU3205-042 mPEG-NH$_2$ 23kDa TU3205-052

TU3205-52 (23 kDa 2-AAP-PEG)

Lithium 2-(4-acetyl-3-aminophenoxy)acetate (TU3205-042) (21.5 mg, 100 μmol) and HATU (38.0 mg, 100 μmol) were charged in a 10 mL round bottom flask, and DMF (0.5 mL) was added. The resulting slurry was stirred at the ambient temperature for 50 min. The resulting yellow solution was added to mPEG-NH$_2$ (Laysan Bio, average MW 23 kDa, average n = 520, 0.50 g, 22 μmol) in 5 mL DMF in a 20 mL glass vial. The reaction was shaken at ambient temperature for 2.5 h. The reaction mixture was then diluted with 10 mL water,

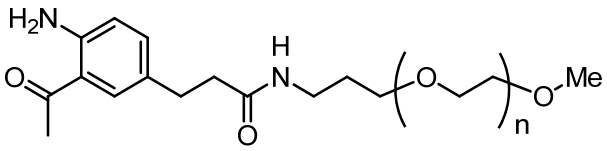
and an aliquot of 2.5 mL each of the solution was applied to PD-10 columns (GE Healthtech) and the desired product was eluted with H_2O according to the supplier's instruction. The pooled aqueous solutions were frozen, and lyophilized, affording 0.352g of TU3205-052 (23 kDa 2-AAP-mPEG) as white solid.



mPEG-NH_2 29kDa **TU633-084**

TU633-084 (29 kDa 2-AAP-PEG)

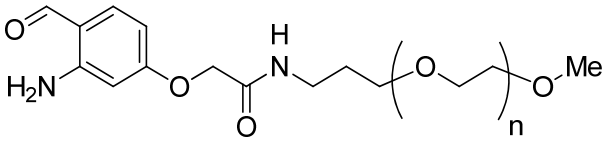
Lithium 2-(4-acetyl-3-aminophenoxy)acetate (187 mg, 0.869 mmol) and HBTU (330 mg, 0.871 mmol) were put in a 20mL glass vial and 10mL DMF was added. The resulting slurry was stirred at ambient temperature. Within 20 min, the reaction turned into yellow solution and after 80 min a 9.5 mL aliquot of the reaction mixture was added to mPEG-NH_2 (Laysan Bio, average MW 28700) dissolved in 40mL DMF in a 100 mL round bottom flask. The reaction was shaken at ambient temperature for 19 h. The reaction mixture was then applied to 24 pieces of PD-10 columns (GE Helthtech) and the desired product was eluted with H_2O according to the supplier's instruction. The pooled eluents were frozen and lyophilized, affording white solid. The solid was dissolved in H_2O, and dialyzed exhaustively against H_2O using a dialysis membrane of MWCO 3500. The dialyzed solution was frozen and lyophilized, affording TU633-084 as white fluffy solid.



TU633-120 (30 kDa 2-AAP-PEG)

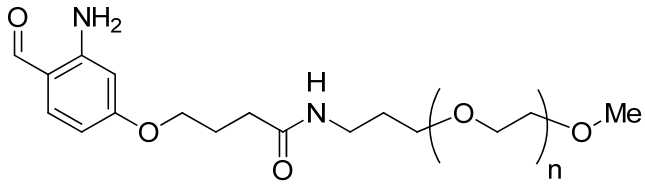
Lithium 3-(3-acetyl-4-aminophenyl)propanoate (139 mg, 0.652 mmol) and HBTU (247 mg, 0.652 mmol) were put in a 20 mL glass vial, and 13mL DMF was added. The resulting slurry was stirred at ambient temperature for 30 min, resulting in a solution of the activated ester. In a 45 mL glass vial was put mPEGamine (NOF Corp., SUNBRIGHT MEPA-30T, average MW 30298, 3.0 g), and 20 mL DMF was added, followed by gentle heating to dissolve mPEGamine in DMF. To the resulting mPEGamine solution was added a 2.2 mL aliquot of the activated ester solution, and the vial was shaken at ambient temperature for 18 h and then at 37 ^oC for 24 h. The reaction mixtures was transferred to dialysis membrane tubing (Fisher Scientific, MWCO 3500) and dialyzed exhaustively against doubly deionized water over 2 days. The dialyzed solution was frozen and lyophilized, affording TU633-120 as white cotton-like solid (Note; the characterization of high MW PEG reagents was obtained only after they were conjugated to the Pcl containing proteins).

7.1.2.2. Synthesis of linear 2-ABA-PEGs (TU3627-024, TU3627-084, TU3627-086)



TU3627-024 (30 kDa 2-ABA-PEG)

TU3627-024 (30 kDa 2-ABA-mPEG) was prepared by the same way as TU633-120 except for using lithium 2-(3-amino-4-formylphenoxy)acetate instead of lithium 3-(3-acetyl-4-aminophenyl)propanoate.



TU3627-084 (30 kDa 2-ABA-PEG)

TU3627-084 (30 kDa 2-ABA-mPEG) was prepared by the same way as TU633-120 except for using lithium 4-(3-amino-4-formylphenoxy)butanoate instead of lithium 3-(3-acetyl-4-aminophenyl)propanoate.

TU3627-086 (40 kDa 2-ABA-PEG)

TU3627-086 (40 kDa 2-ABA-mPEG) was prepared by the same way as TU3627-084 except for using SUNBRIGHT MEPA-40T (NOF Corp., average MW 40036) instead of SUNBRIGHTMEPA-30T. The ^{1}H NMR analysis of 2-ABA-PEGs revealed no unmodified starting PEG in the product. Terminal activity: >95% by ^{1}H-NMR.

7.2. Disaccharide Conjugate

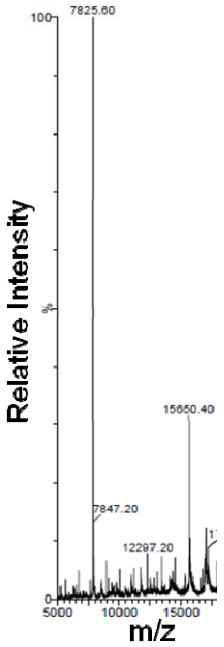
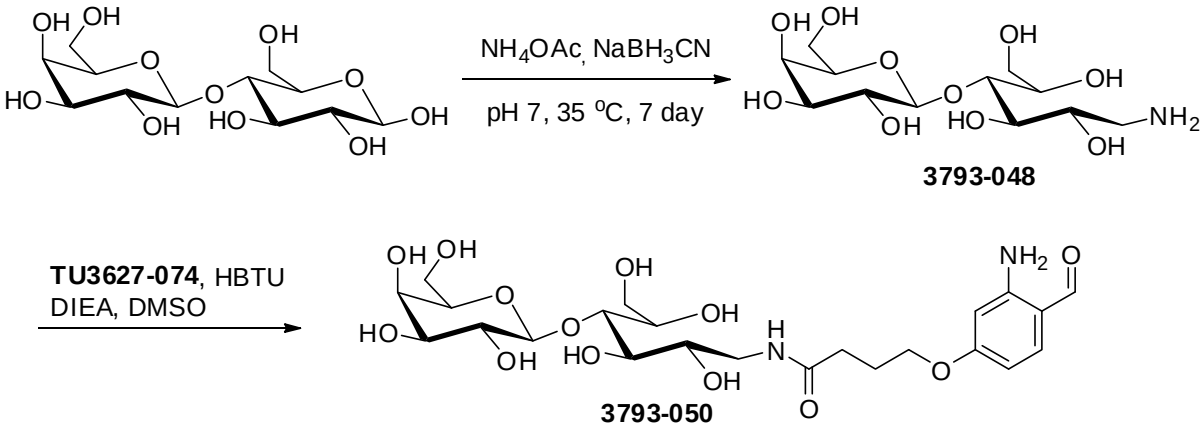


Fig. S8. Disaccharide conjugate. To demonstrate that saccharides can be coupled site-specifically to Pcl proteins, mEGF-Tyr10Pcl (see above) was conjugated with a 2-ABA-disaccharide (3793-050; see below, MW 549 Da). The coupling reaction was carried out by the addition of 1 mM 2-ABA-disaccharide to 10 μM mEGF-Tyr10Pcl mutant protein in PBS and 1% (v/v) DMSO at pH 7.0. The reaction was allowed to proceed at room temperature for 16 hours and analyzed by ESI-MS. The mass spectrum shows the major peak near the expected mass for the conjugated protein (7825.6 Da observed, 7827 Da expected). The expected mass for the non-coupled protein is 7296 Da.

Synthesis of 2-ABA-disaccharide: Gal-Glu-1-amide of 3-amino-4-formylphenoxybutyrate (3793-050).



$NaBH_3CN$ (94.3 mg) was added to an H_2O solution (10 mL) of NH_4OAc (771 mg) and lactose (180mg) at pH 7 and the reaction was stirred at 35°C for 7 days. The reaction mixture was lyophilized, and the residue was dissolved in H_2O, followed by a passage through Dowex 1X8-400 anion exchange resin (OH- form, 45 g) to remove excess $[BH_3CN]^-$, its byproducts and acetate. The eluent was lyophilized, affording the crude product, which was purified by cation exchange chromatography using Dowex 50WX8-400 resin (H^+ form, 30 g), affording the desired product (3793-048) as light yellow powder. ESI-MS calculated for $C_{12}H_{25}NO_{10}$ $[MH]^+$: 344.1; observed: 344.2.

Lithium 3-amino-4-formylphenoxybutyrate (TU3627-074, 6.0 mg) was treated with HBTU (9.9 mg) and DIEA (7.7 μL) in DMSO (200 μL) for 1 h. A DMSO (250 μL) solution of Gal-Glu-1-amine (3793-048, 7.7 mg) was then added to the reaction mixture at ambient temperature, followed by agitation for 1 day. The reaction mixture was lyophilized and purified by preparative MS-triggered HPLC with NH_4OAc elution, affording the desired product (3793-050) as a yellow powder. ESI-MS calculated for $C_{23}H_{36}N_2O_{13}$ $[MH]^+$: 549.2; observed: 549.3.

7.3. Phospholipid Conjugate

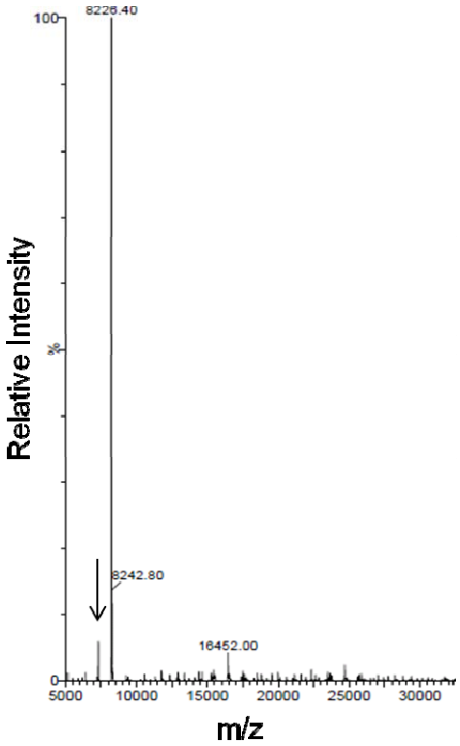
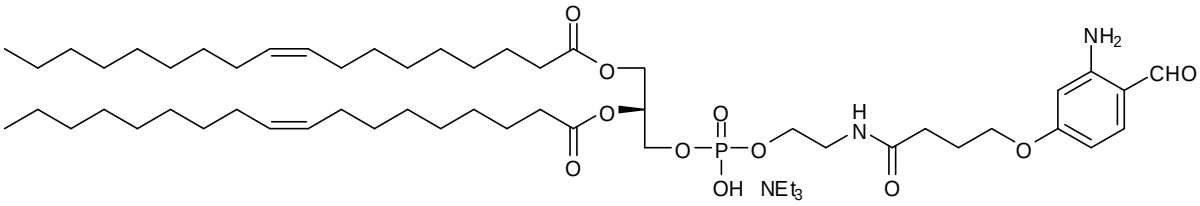


Fig. S9. Phospholipid conjugate. To demonstrate that phospholipids can be coupled site-specifically to Pcl proteins, mEGF-Tyr10Pcl (see above) was conjugated with 2-ABA-DOPE (TU3627-092; see below, MW 949 Da). Coupling of 2-ABA-DOPE to mEGF-Tyr10Pcl was carried out in 20 mM HEPES (pH 7.0) and 1% (v/v) DMSO at 25°C for 16 hours. The conjugation reaction was initiated by the addition of 10 μM mEGF-Tyr10Pcl and 100 μM 2-ABA-DOPE. Formation of the protein conjugate was verified by electrospray ionization-mass spectrometry (ESI-MS) analysis of the conjugation of DOPE to mEGF-Y10Pcl (expected mass of uncoupled protein = 7296 Da; expected mass of the conjugated protein = 8227 Da, observed 8226.4 Da). Unreacted protein is indicated by an arrow.

Synthesis of 2-ABA-DOPE (TU3627-092).



Lithium 4-(3-amino-4-formylphenoxy)butanoate (34 mg) and HBTU (57 mg) were put in a 20 mL glass vial, and 2 mL DMF was added. The reaction was stirred at ambient temperature for 30 min for activation. In a separate 20mL vial was put DOPE (76 mg, 1, 2,-dioleoyl-sn-glycero-3-phosphoethanolamine, NOF Corp.), followed by DIEA (35 μL) and 3 mL DCM. The yellow solution of the activated ester in the first vial was transferred to the second vial, and the reaction was stirred at ambient temperature. After 24 h, the reaction mixture was applied to a 12 g pre-packed SiO_2 column equilibrated with solvent A (solvent A: 5% NEt_3 in DMC, solvent B: 5% NEt_3 in MeOH), and the column was eluted with a linear gradient of 0 to 15 %B in A over 15 min, affording partially purified product as light yellow very viscous oil. This product was purified again by flash chromatography using a 12g SiO_2 column (solvent A: 5% NEt_3 in DMC, solvent B: 5% NEt_3 in MeOH), and eluted with a linear gradient of 2 to 10 % B in A over 15 min. The fraction containing the pure product was concentrated under reduced pressure, affording the title compound as triethylammonium salt. Rf: 0.43 (SiO_2, 10% MeOH in DCM). MS (ESI^+) (m/z): calcd. 949.63, found 949.60 (MH^+). ^{1}H-NMR (400 MHz, $CDCl_3$, δ): 0.875 (t, *J*= 6.8 Hz, 6H), 1.28 (broad multiple peaks, 40H), 1.318 (t, *J*=7.4 Hz, 9H), 1.582 (m, 4H), 2.00 (m, 8H), 2.124 (quint, *J*=6.8 Hz, 2H), 2.28 (m, 4H), 2.373 (t, *J*=7.0 Hz, 2H), 3.042 (q, *J*=7.2 Hz, 6H), 3.476 (2H, m), 4.00 (6H, m), 4.154 (1H, dd, *J*=6.8, 12.0 Hz), 4.373 (1H, dd, *J*=3.2, 12.0 Hz), 5.233 (1H, m), 5.337 (m, 4H), 6.217 (d, *J*=2.0 Hz, 1H), 6.255 (dd, *J*=2.0, 8.4 Hz, 1H), 6.56 (very broad peak, 2H), 7.313 (d, *J*=8.8 Hz, 1H), 7.432 (br.t, *J*=8.4 Hz, 1H), 9.674 (s, 1H), 11.97 (br.s, 1H).

7.4. Biotin Conjugate

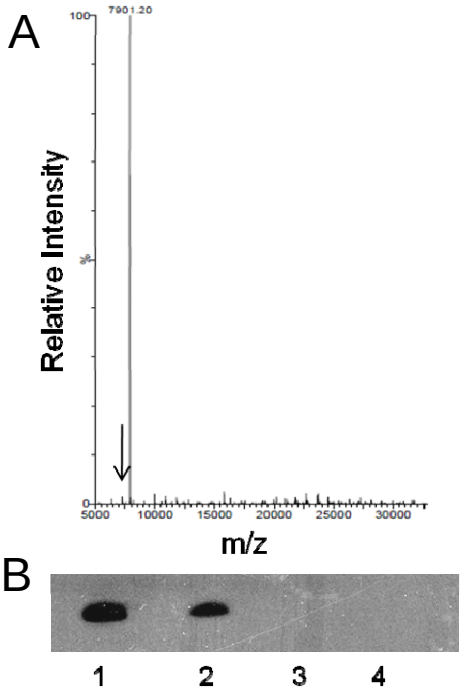
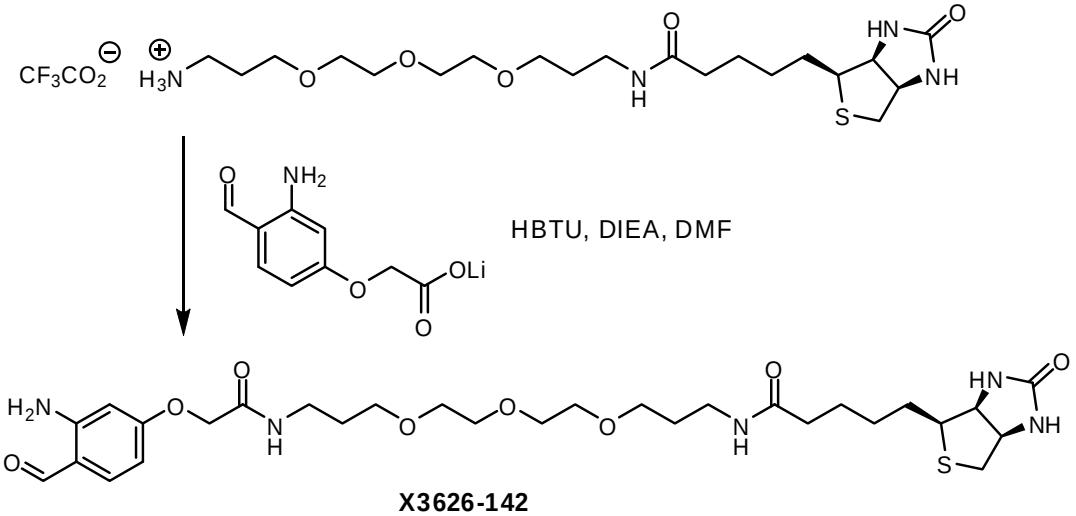


Fig. S10. Biotin conjugate. To demonstrate that biotin can be coupled site-specifically to Pcl proteins, mEGF-Tyr10Pcl (see above) was conjugated with a 2-ABA-biotin reagent (X3642-142, see below, MW 624 Da) as follows: 500 µM 2-ABA-biotin was added to 10 µM mEGF-Tyr10Pcl in PBS (pH 7.0) and 0.5% (v/v) DMSO and reacted at 25°C for 16 hours. Complete formation of the biotin conjugate was verified by ESI-MS (*A*) (expected mass of uncoupled protein = 7296 Da; expected mass of coupled protein = 7902 Da, observed 7901.2 Da). The approximate position of the signal for unmodified protein is indicated by an arrow.

After quantitative coupling, a freshly prepared solution of 200 mM $NaBH_3CN$ (dissolved in PBS, pH 7.0) was added to a final concentration of 20 mM. After allowing the reduction reaction to proceed for 3 hours at 25°C, the reaction was diluted with six volumes of PBS (pH 7.0). Excessive $NaBH_3CN$ was removed by dialysis against PBS (pH 7.0) at 4 °C using a Slide-A-Lyzer dialysis cassette (3,500 Da molecular weight cutoff, Pierce). The reduced biotin conjugate was concentrated using an Amicon Ultra-4 Centrifugal Filter Unit with 3.5 kDa cutoff (Millipore Corporation). After electrophoresis through a NuPAGE 4-12% Bis-Tris gel (Invitrogen), the biotinylated protein was transferred on a polyvinylidene difluoride (PVDF) membrane using an iBlot Gel Transfer System (Invitrogen). The biotin conjugate was then detected with a horseradish peroxidase (HRP) conjugated goat anti-biotin antibody (1:100 dilution, Cell Signaling Technologies) and visualized using a Hyperfilm ECL (GE Healthcare) (*B*). Uncoupled mEGF-Tyr10Pcl and fluorescein-conjugated mEGF-Tyr10Pcl served as negative controls. Lane 1, 20 pmol mEGF-Tyr10Pcl-2-ABA-biotin conjugate; lane 2, 8 pmol mEGF-Tyr10Pcl-2-ABA-biotin conjugate; lane 3, 2 nmol mEGF-Y10Pcl; lane 4, 20 pmol mEGF-Tyr10Pcl-2-ABA-fluorescein conjugate.

Synthesis of 2-ABA-biotin: N-(1-(3-amino-4-formylphenoxy)-2-oxo-7,10,13-trioxa-3-azahexadecan-16-yl)-5-((3aS,4S,6aR)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanamide (X3626-142).



X3626-142

To lithium 2-(3-amino-4-formylphenoxy)acetate (TU3627-018, 45.3 mg, 1.2 equiv), HBTU (77.8 mg, 1.15 equiv) and DIEA(31.1 μL, 1.0 equiv) was added DMF (1.0 mL) and the mixture was stirred for 30 min at room temperature. The resulting solution was added to Biotin-dPEG$^{TM}_3$-NH_3^+TFA (QuantaBiodesign, 100 mg, 1.0 equiv), DIEA (46.6 μL, 1.5 equiv) in DMF(1.0 mL) and stirred at room temperature. The desired product was isolated by preparative HPLC. The title compound was obtained as free base (21.1 mg, 19%) after removal of TFA using anion exchange resin. MS (ESI^+): calcd. for $C_{29}H_{46}N_5O_8S^+$ (MH^+) 624.30, found 624.30.

7.5. Fluorescein Conjugate

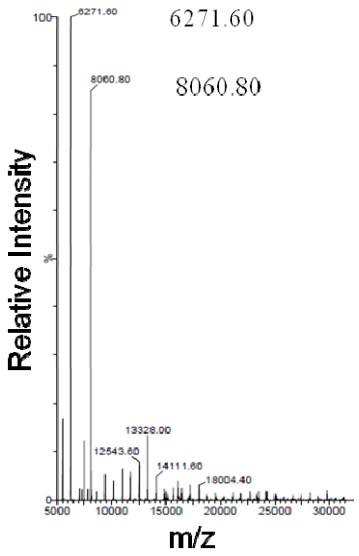
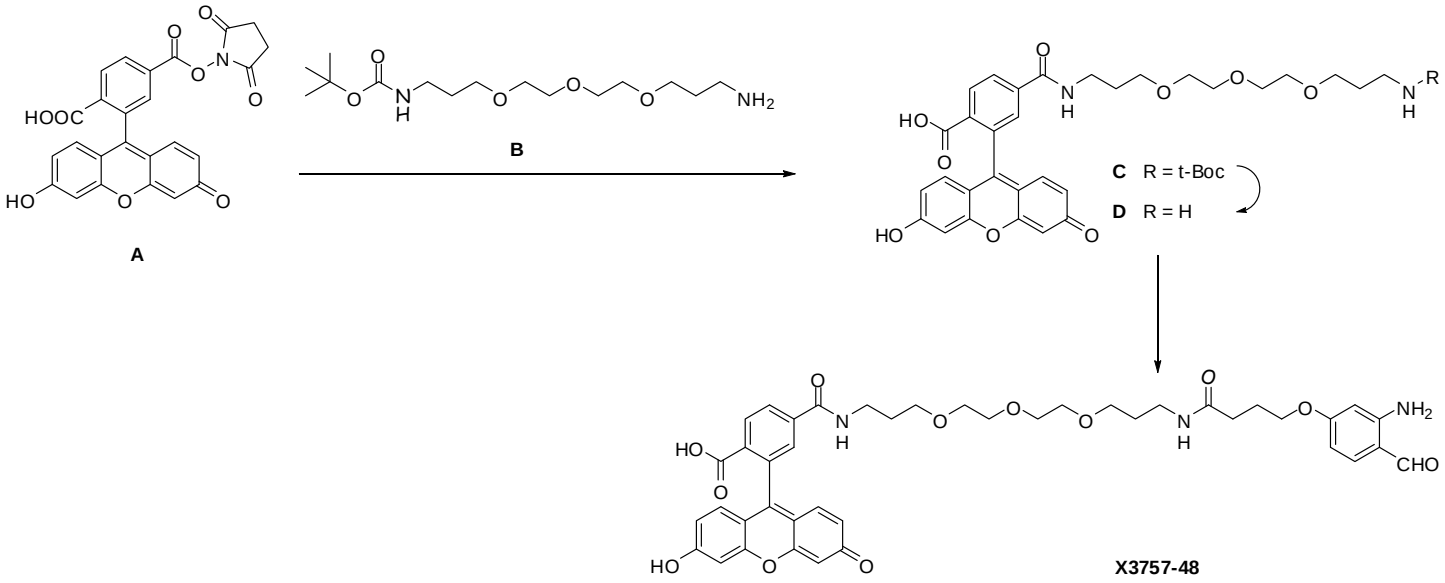


Fig. S11. Fluorescein conjugate. To demonstrate that a fluorescent molecule can be coupled site-specifically to Pcl proteins, mEGF-Tyr10Pcl (see above) was conjugated with a 2-ABA-fluorescein reagent (X3757-48, see below, MW 784 Da). Coupling of 2-ABA-fluorescein was carried out as follows: 1 mM of 2-ABA-fluorescein was added to 10 µM mEGF-Tyr10Pcl in PBS (pH 7.0) and 1% (v/v) DMSO and reacted at 25°C for 16 hours. Formation of the fluorescein conjugate was verified by ESI-MS (expected mass of uncoupled protein = 7296 Da; expected mass of coupled protein = 8062 Da, observed 8060.8 Da).

The fluorescein conjugate was reduced with 20 mM $NaBH_3CN$ for 3 hours at 22 °C. After diluting the reduction reaction with six volumes of PBS (pH 7.0), residual $NaBH_3CN$ was removed by dialysis against PBS (pH 7.0) at 4°C using a Slide-A-Lyzer dialysis cassette (3,500 Da molecular weight cutoff). The conjugate was then concentrated to 1 µM using an Amicon Ultra-4 Centrifugal Filter Unit with 3.5 kDa cutoff. Absorbance spectra in the range of 350 – 700 nm of 1 µM mEGF-Tyr10Pcl-2-ABA-fluorescein and of 10 µM 2-ABA-fluorescein were obtained using a SpectraMax Plus (Molecular Devices). Both samples had an absorbance maximum at 500 nm. Fluorescence spectra were recorded on a SpectraMax GEMINI fluorometer (Molecular Devices). Emission spectra for 1 µM mEGF-Tyr10Pcl-2-ABA-fluorescein and of 10 µM 2-ABA-fluorescein were obtained by maintaining the excitation wavelength at 490 nm, while scanning the emission wavelength from 510 nm to 750 nm using a step size of 2 nm. Both samples had an emission maximum at 522 nm. Excitation maxima at 496 nm were measured for both samples by recording excitation spectra (300 to 510 nm) with a constant emission wavelength of 522 nm.

Synthesis of 2-ABA-Fluorescein (X3757-48).

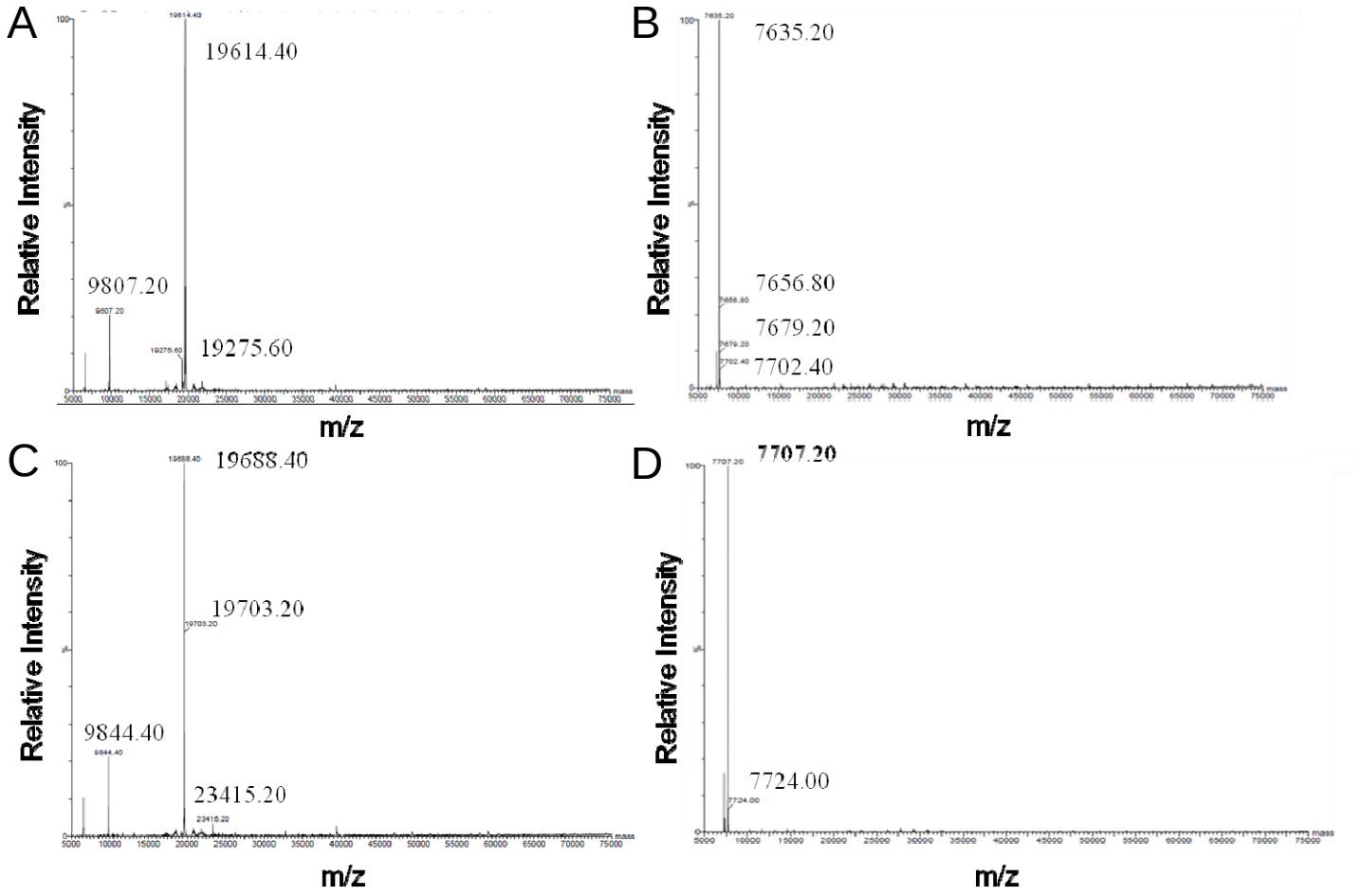


DMF (2.0 mL) was added to NHS-Fluorescein **A** (a mixture of 5- and 6- isomers, 23.6 mg), amine **B** (17.6 mg) and DIEA (8.7 µL). The mixture was stirred at ambient temperature until **A** was consumed, monitored by

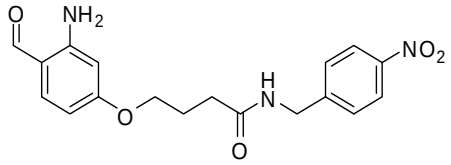
reverse-phase HPLC. The product **C** was isolated by preparative MS-triggered HPLC. MS (ESI^+) (m/z); calcd. for $C_{36}H_{43}N_2O_{11}^+$ (MH^+) 679.29, found 679.30. The product (**C**, 7.4 mg) was dissolved in 3 M HCl (1.0 mL), and the volatiles were removed by evaporation under reduced pressure after stirring for 5 min at ambient temperature. This operation was repeated, resulting in removal of the Boc group, affording amine **D**. MS (ESI^+) (m/z); calcd. for $C_{31}H_{35}N_2O_9^+$ (MH^+) 579.23, found 579.30. To amine **D** was added a premixed solution of lithium 4-(3-amino-4-formylphenoxy)butanoate (2.3 mg), HBTU (3.8 mg), DIEA (7.0 µL) and DMF (2 mL) at ambient temperature. The title product was isolated by preparative MS-triggered HPLC. MS (ESI^+): calcd. for $C_{42}H_{46}N_3O_{12}^+$ (MH^+) 784.31, found 784.30.

7.6. Nitro-phenyl Hapten Conjugates

Fig. S12. Mono- and di-nitro-phenyl hapten conjugates. To demonstrate that nitro-phenyl immune modulators can be coupled site-specifically to Pcl proteins, mTNF-Gln21Pcl and mEGF-Tyr10Pcl (see above) were conjugated with 2-ABA-mono-nitro-phenyl hapten (3793-001, see below, MW 357 Da) and di-nitro-phenyl hapten reagents (TU3627-088, see below, MW 431 Da). 10 µM Pcl protein was reacted with 1 mM 2-ABA-nitro-phenyl hapten in PBS (pH 7.0) and 1% (v/v) DMSO at 22 °C for 16 hours. (*A*) ESI mass spectrum of 3793-001 conjugated to mTNF-Gln21Pcl (expected mass of uncoupled protein = 19275 Da; expected mass of coupled protein = 19614 Da, observed 19614.4 Da). (*B*) ESI mass spectrum of 3793-001 conjugated to mEGF-Tyr10Pcl (expected mass of uncoupled protein = 7296 Da; expected mass of coupled protein = 7635 Da, observed 7635.2 Da). (*C*) ESI mass spectrum of TU3627-088 conjugated to mTNF-Gln21Pcl (expected mass of coupled protein = 19688 Da, observed 19688.4 Da). (*D*) ESI mass spectrum of TU3627-088 conjugated to mEGF-Tyr10Pcl (expected mass of coupled protein = 7709 Da, observed 7707.2 Da).

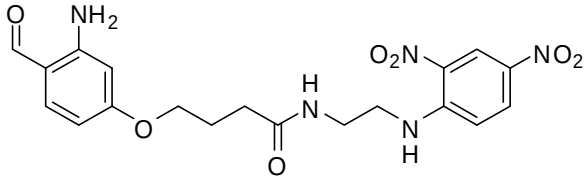


Synthesis of 2-ABA-nitrophenyl hapten: 4-(3-Amino-4-formylphenoxy)-*N*-(4-nitrobenzyl)butanamide (3793-001).



HATU (114.1 mg) was added to a 1 mL DMF solution of lithium 4-(3-amino-4-formylphenoxy)-butanoate (68.7 mg), and the reaction was shaken at ambient temperature for 1 hr. The resulting solution was added to a 1 mL DMF solution of 4-nitrobenzylamine HCl salt (56.6 mg) and triethylamine (84 μL), with another 1 mL DMF to help transfer. The reaction was stirred at ambient temperature for 2 hr. Upon completion indicated by LCMS analysis, the reaction mixture was partitioned between 4 mL of 10% $NaCl_{(aq)}$ and 8 mL of EtOAc. The phases were separated and the aqueous layer was extracted with 8 mL EtOAc. The combined organic layers were dried over Na_2SO_4, filtered and concentrated under reduced pressure, affording crude orange oil. The crude oil was purified by silica gel flash column chromatography with linear gradient elution of 0-10% MeOH in DCM to give the desired product as light yellow solid. MS (ESI^+) (m/z): calcd. for $C_{18}H_{20}N_3O_5^+$ (MH^+) 358.1, found 358.2. ^{1}H NMR (400 MHz, $CDCl_3$, δ): 2.17 (quintet, *J*=6.5 Hz, 2H), 2.47 (t, *J*=7.0 Hz, 2H), 4.02 (t, *J*=5.8 Hz, 2H), 4.54 (d, *J*=6.0 Hz, 2H), 6.03 (d with br, *J*=2.0 Hz, 2H), 6.20 (br, 2H), 6.26 (dd, *J*=2.2 Hz, 8.8 Hz, 1H), 7.35 (d, *J*=8.4 Hz, 1H), 7.40 (d, *J*=8.8 Hz, 2H), 7.40 (d, *J*=8.8 Hz, 2H), 8.14 (d, *J*=8.8 Hz, 2H), 9.70 (s, 1H). ^{13}C NMR (100 MHz, $CDCl_3$, δ): 24.8, 32.5, 42.8, 66.7, 98.9, 105.3 113.9, 123.9, 128.2, 137.8, 145.8, 147.2, 152.2, 164.4, 172.2, 192.0.

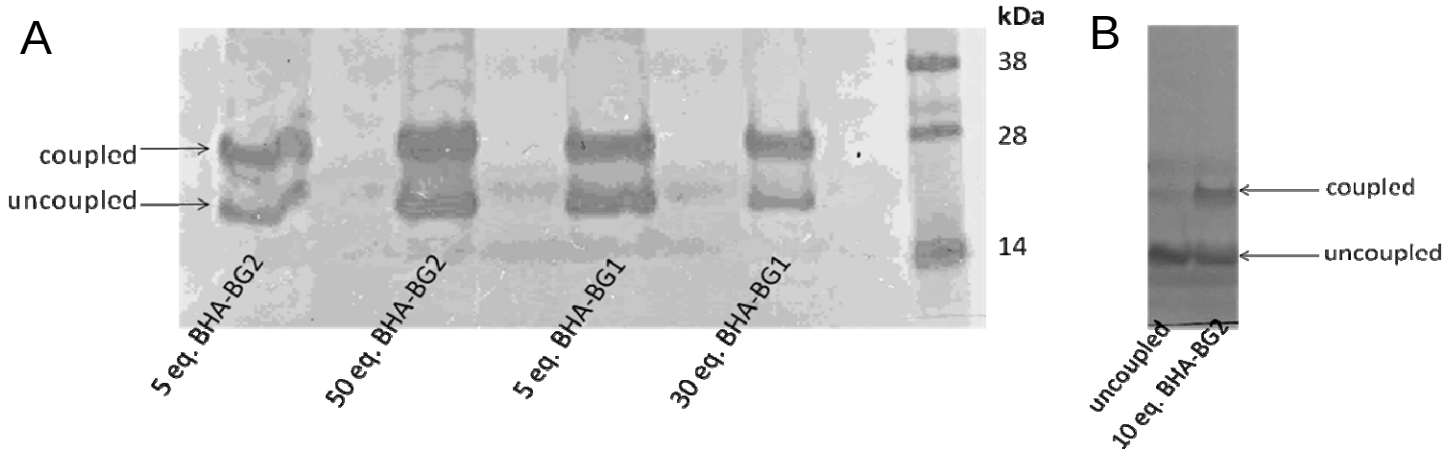
Synthesis of 2-ABA-dinitrophenyl hapten: 4-(3-amino-4-formylphenoxy)-*N*-(2-(2, 4-dinitrophenylamino)ethyl)butanamide (TU3627-088).



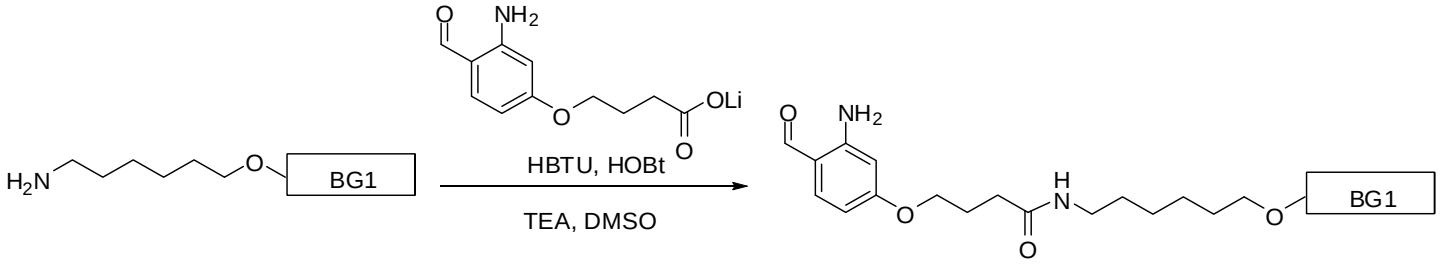
Lithium 4-(3-amino-4-formylphenoxy)butanoate (50 mg), HBTU (80 mg) and DMF (2 mL) were combined in a 20 mL glass vial and stirred at ambient temperature. After 30min, N^1-(2, 4-dinitrophenyl)ethane-1, 2-diamine (Oakwood, 45 mg) was added in one portion and the reaction was stirred at ambient temperature overnight. The reaction mixture was diluted with EtOAc, washed successively with dilute aq. citric acid, water, sat. aq. $NaHCO_3$, water, and sat. aq. NaCl, dried over Na_2SO_4, filtered and concentrated under reduced pressure, affording the title compound as yellow solid. Rf: 0.18 (SiO_2, EtOAc). MS (ESI^+) (m/z): calcd. for $C_{19}H_{22}N_5O_7^+$ (MH^+) 432.15, found 432.20. ^{1}H-NMR (400 MHz, DMSO-d6, δ): 1.914 (quint, *J*=7.0 Hz, 2H), 2.224 (t, *J*=7.0 Hz, 2H), 3.349 (q, *J*=6.0 Hz, 2H), 3.538 (q, *J*=6.0 Hz, 2H), 3.928 (t, *J*=6.4 Hz, 2H), 6.165 (s, 1H), 6.176 (dd, *J*=2.4 Hz, 10.0 Hz, 1H), 7.133 (br.s, 1H), 7.268 (d, *J*=10.0 Hz, 1H), 7.371 (d, *J*=8.4Hz, 1H), 8.167 (t, *J*=5.8 Hz, 1H), 8.250 (dd, *J*=2.8, 8.8 Hz, 1H), 8.835 (d, *J*=2.8 Hz, 1H), 8.913 (t, *J*=5.6Hz, 1H), 9.599 (s, 1H). ^{13}C-NMR (100 MHz, DMSO-d6, δ): 24.489, 31.436, 37.331, 42.745, 66.782, 98.040, 104.339, 112.751, 115.130, 123.472, 123.564, 129.725, 134.731, 137.484, 148.305, 152.824, 163.873, 172.251, 191.400.

7.7. CpG Oligonucleotide Conjugates

Fig. S13. CpG Oligonucleotide conjugates. To demonstrate that oligonucleotides and CpG phosphothioate immune modulators(10) can be coupled site-specifically to Pcl proteins, mTNF-Gln21Pcl and mEGF-Tyr10Pcl (see above) were conjugated with CpG oligonucleotides BHA-BG1 (3647-057; see below) or BHA-BG2 (3597-167; see below). 10 μM Pcl protein was reacted with 100 μM 2-ABA-CpG-oligonucleotides in 10x PBS at pH 7.5 and 22 ^{o}C followed by reduction with 20 mM $NaBH_3CN$ for 4 hours. Excess $NaBH_3CN$ was removed by using a PD10 column. Coupling of the conjugates was analyzed by gel shift assay using a NuPAGE 4-12% Bis-Tris Gel (Invitrogen, Carlsbad, CA). (***A***) Coupling of BHA-BG1 (7.4 kDa) and BHA-BG2 (7.4 kDa) to mTNF-Gln21Pcl (19.3 kDa) was confirmed by gel shift assay. (***B***) Coupling of BHA-BG2 (7.4 kDa) to mEGF-Tyr10Pcl (7.2 kDa) was also confirmed by gel shift assay.



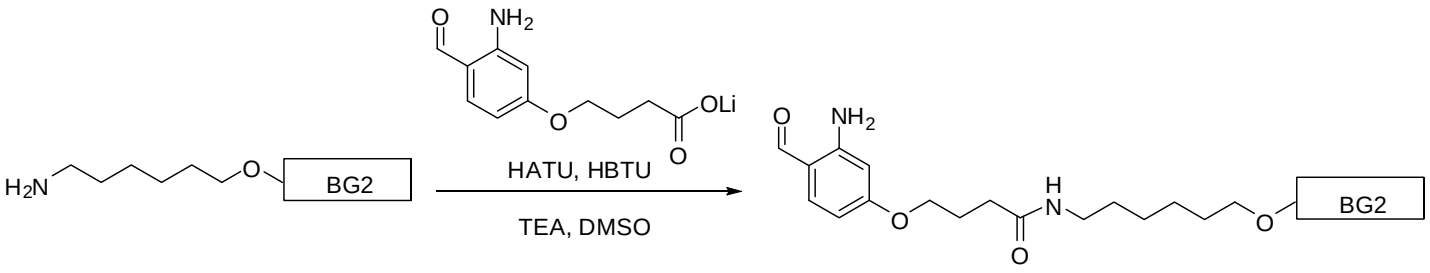
Synthesis of BHA-BG1: 6-(4-(3-Amino-4-formylphenoxy)butanamido)hexyl-5'-*T*C*C*A*T*G*A*C*G*T*T*C*C*T*G*A*C*G*T*T-3' (*: phosphothioate) (3647-057).



BG1 = 5'-*T*C*C*A*T*G*A*C*G*T*T*C*C*T*G*A*C*G*T*T-3' (*: phosphothioate)

HBTU (6.1 mg), HOBt (2.2 mg), and triethylamine (7 μL) were added to a 1 mL DMSO solution of lithium 4-(3-amino-4-formylphenoxy)-butanoate (4.6 mg), and the reaction was shaken at ambient temperature for 1 h. A 276 μL aliquot of the resulting solution was added to a 1.2 mL DMSO solution of amino-modified BG1 phosphothioate oligomer (10) (12.1 mg, 1.8 μmol, purchased from Integrated DNA Technologies, Inc.) and triethylamine (15 μL), and the reaction was shaken at ambient temperature for 2 days. The reaction mixture was diluted with water and dialyzed against water (4 L x10) using Slide-A-Lyzer (Thermo Scientific, MWCO of 3500). The dialyzed solution was lyophilized, affording the desired product as white solid. MS (ESI-Q-TOF) (m/z): calcd. for $C_{211}H_{273}N_{69}O_{108}P_{20}S_{20}$ 6759.7, found 6759.1.

Synthesis of BHA-BG2: 6-(4-(3-amino-4-formylphenoxy)butanamido)hexyl-5'-*T*C*G*T*C*G*T*T*T*T*C*G*G*C*G*C*G*C*G*C*C*G-3' (*:phosphothioate) (3597-167).



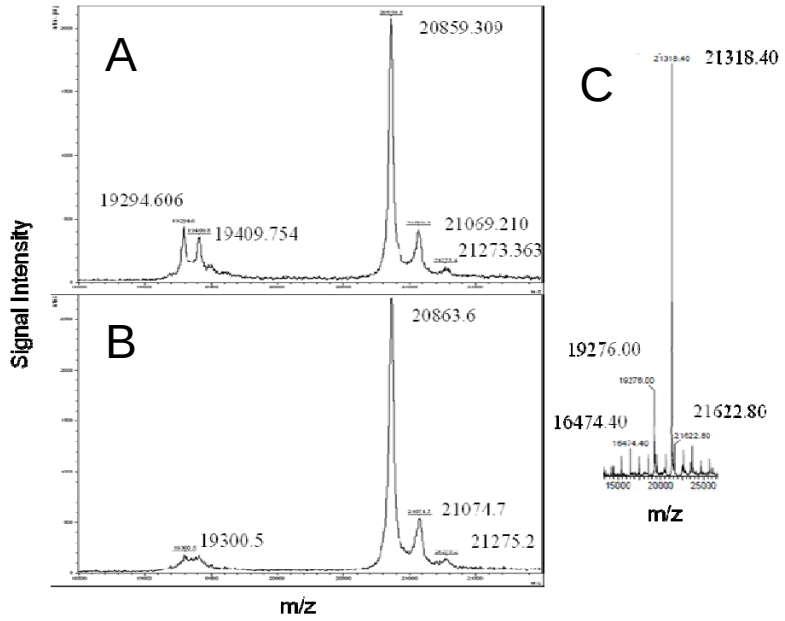
3597-167

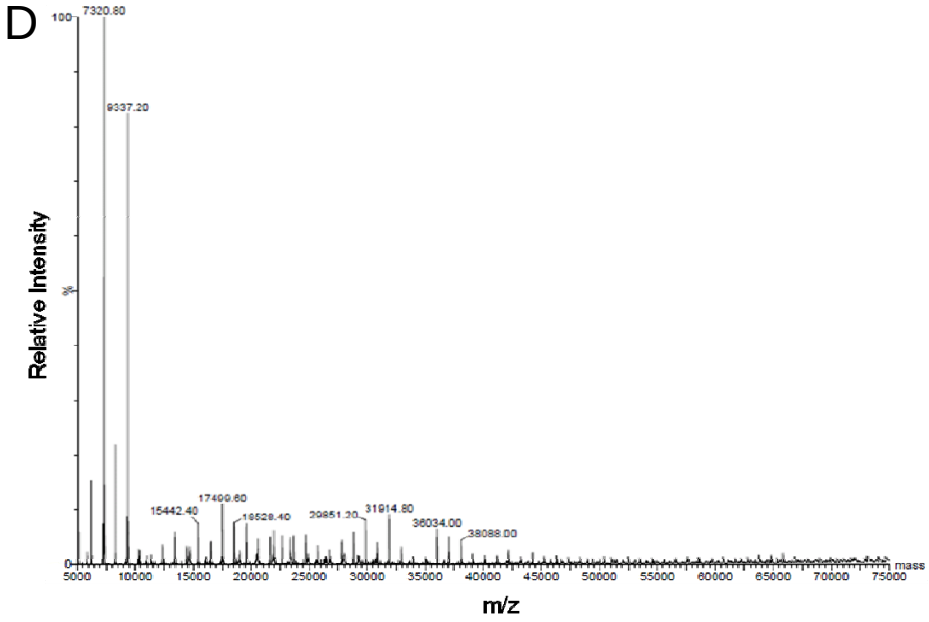
BG2 = 5'-*T*C*G*T*C*G*T*T*T*T*C*G*G*C*G*C*G*C*G*C*C*G-3' (*:phosphothioate)

HATU (7.4 mg) was added to a 1 mL DMSO solution of lithium 4-(3-amino-4-formylphenoxy)-butanoate (5.5 mg), and the reaction was shaken at ambient temperature for 1 h. A 60 μL aliquot of the resulting solution was added to a 0.6 mL DMSO solution of amino-modified BG2 phosphothioate oligomer (6.9 mg, 1.0 μmol) and triethylamine (7.5 μL), and the reaction was shaken at ambient temperature for 20 h. Then another 60 μL aliquot of the activated ester solution that was freshly prepared in the same way except for using HBTU instead of HATU was added to the reaction mixture, and the reaction was shaken for an additional 2 days. The reaction mixture was separated into 3 portions. Each portion was applied to a NAP-25 column equilibrated with H_2O and eluted with H_2O. Every 1 mL fraction was monitored by LC-MS and the fractions containing the desired product were combined and lyophilized, affording the desired product as white solid. MS (ESI-Q-TOF) (m/z) calcd. for $C_{229}H_{296}N_{78}O_{120}P_{22}S_{22}$ 7442.8, found 7442.7.

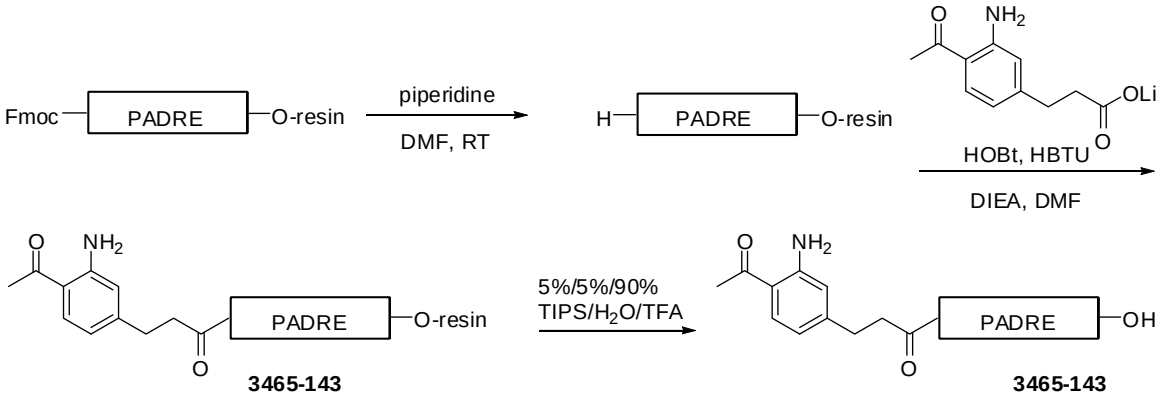
7.8. PADRE Peptide Conjugates

Fig. S14. PADRE Peptide conjugates. To demonstrate that peptides can be coupled site-specifically to Pcl proteins, mTNF-Gln21Pcl and mEGF-Tyr10Pcl (see above) were conjugated with 2-AAP-PADRE and 2-ABA-exPADRE peptides(11). 10 μM Pcl protein was reacted with 100 μM 2-AAP-PADRE or 2-ABA-exPADRE peptide in 10x PBS at pH 7.5 and 22 ^{o}C followed by reduction with 20 mM $NaBH_3CN$ for 4 hours. Excess $NaBH_3CN$ was removed by using a PD10 column. (*A*) MALDI-TOF mass spectrometric analysis of mTNF-Gln21Pcl conjugated with 2-AAP-PADRE (3465-143; see below; MW 1586 Da) at pH 5.0. (*B*) MALDI-TOF mass spectrometric analysis of mTNF-Gln21Pcl conjugated with 2-AAP-PADRE at pH 7.5. The expected mass of uncoupled protein is 19275 Da and the expected mass of coupled protein is 20843 Da (observed 20859.2 Da in (*A*) and 20863.6 Da in (*B*). (*C*) ESI mass spectrometric analysis of mTNF-Gln21Pcl conjugated with 2-ABA-exPADRE (3647-104; see below; MW 2060 Da). The expected mass of uncoupled protein is 19275 Da (observed 19276.0 Da) and the expected mass of coupled protein is 21317 Da (observed 21318.4 Da). (*D*) ESI mass spectrum showing the coupling of 2-ABA-exPADRE to mEGF-Tyr10Pcl (expected mass of uncoupled protein 7296 Da; expected mass of coupled protein 9338 Da, observed 9337.2 Da).





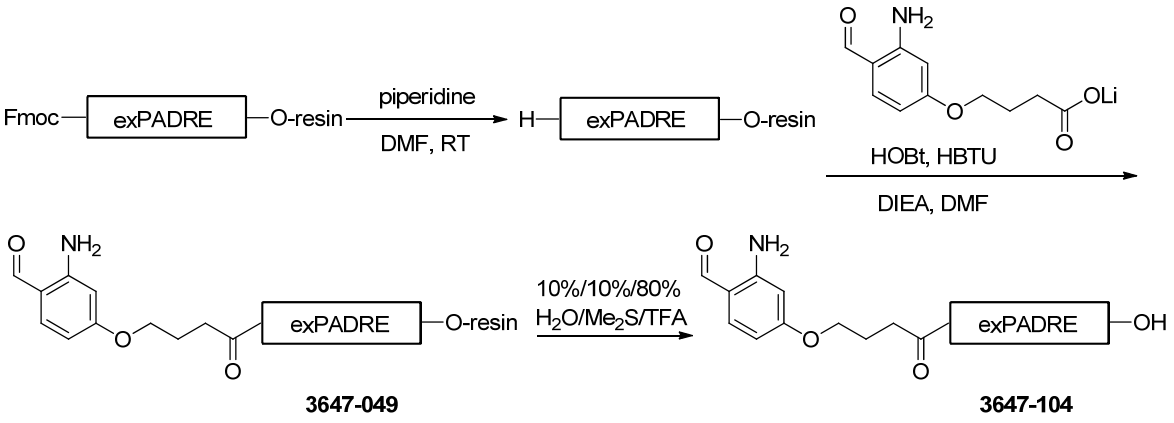
Synthesis of 2-AAP-PADRE: 3-(4-acetyl-3-aminophenyl)propanoyl-Gly(D-Ala)LysChaValAlaAlaTrpThr-LeuLysAla(D-Ala)Gly-OH (3465-143).



PADRE = Gly(D-Ala)LysChaValAlaAlaTrpThrLeuLysAla(D-Ala)Gly

The Fmoc group of Fmoc-PADRE CLEAR resin (105.3 mg, 0.02 mmol, purchased from Peptide International Inc.) was removed by 20% piperidine/DMF (8 mL x3). The resin was washed by DMF (1.5 mL x 5). Then the resin was treated with a 0.5 mL DMF solution of lithium 3-(4-acetyl-3-aminophenyl)propanoate (5.1 mg), HBTU (9.1 mg), and HOBt (3.3 mg) for 7 h at ambient temperature. The resin was washed with DMF and the peptide was cleaved from the resin by TIPS/H_2O/TFA (5/55/90 v/v/v%, 3 mL) for 2 h. The cleavage slurry was filtered through glass wool and most TFA was removed from the filtrate by evaporation. The residue was washed by hexanes (3 mL x 3), and dissolved in 50% $MeCN_{(aq)}$. The crude product was obtained after lyophilization, and purified by preparative HPLC, affording the desired product as light yellow powder. MS (ESI^+) (m/z): calcd. for $C_{77}H_{120}N_{18}O_{18}$ 793.5 [$(M+2H)^{2+}/2$], 529.3 [$(M+3H)^{3+}/3$], found 793.6, 529.4.

Synthesis of 2-ABA-exPADRE: 4-(3-amino-4-formylphenoxy)butanoyl-AlaGlySerArgSerGly(D-Ala)Lys-ChaValAlaAlaTrpThrLeuLysAla(D-Ala)Gly-OH (3647-104).



exPADRE = AlaGlySerArgSerGly(D-Ala)LysChaValAlaAlaTrpThrLeuLysAla(D-Ala)Gly

The Fmoc group of Fmoc-exPADRE CLEAR resin (549.3 mg, 0.1 mmol, purchased from Peptide International Inc.) was removed by 20% piperidine/DMF (8 mL x 3). The resin was washed by DMF (1.5 mL x 5). Then the resin was treated with a 3 mL DMF solution of lithium 4-(3-amino-4-formylphenoxy)-butanoate (34.4 mg), HBTU (45.5 mg), HOBt (16.2 mg), and DIEA (52 μL) for 6 h at ambient temperature. The resin was washed with DMF and the peptide was cleaved from the resin using H_2O/Me_2S/TFA (10/10/80 v/v/v%, 10 mL) for 2 h at ambient temperature. The cleavage slurry was filtered through glass wool, and most of the TFA was removed from the filtrate by evaporation. The residue was neutralized with 1 N $NaOH_{(aq)}$ and diluted with

MeCN, followed by dialysis using SpectraPro™ 7 dialysis membrane (MWCO of 1000) in 50% $MeCN_{(aq)}$ (4 L x 10). The solution remained in the dialysis membrane was lyophilized, affording the desired product as white solid. MS (ESI^+) (m/z): calcd. for $C_{94}H_{150}N_{26}O_{26}$ 1030.6 $[(M+2H)^{2+}/2]$, 687.4 $[(M+3H)^{3+}/3]$, found 1030.7, 687.6.

References:

1. Cellitti SE*, et al.* (2010) D-ornithine Coopts Pyrrolysine Biosynthesis to Make and Insert Pyrroline-Carboxy-Lysine. *Nat Chem Biol*:epublished.
2. Cellitti SE*, et al.* (2008) In vivo incorporation of unnatural amino acids to probe structure, dynamics, and ligand binding in a large protein by nuclear magnetic resonance spectroscopy. *J. Am. Chem. Soc.* 130(29):9268-9281.
3. Grunewald J*, et al.* (2009) Mechanistic studies of the immunochemical termination of self-tolerance with unnatural amino acids. *Proc. Natl. Acad. Sci. USA* 106(11):4337-4342.
4. Klock HE, Koesema EJ, Knuth MW, & Lesley SA (2008) Combining the polymerase incomplete primer extension method for cloning and mutagenesis with microscreening to accelerate structural genomics efforts. *Proteins.* 71(2):982-994.
5. Brown CM, Stockwell PA, Trotman CN, & Tate WP (1990) Sequence analysis suggests that tetra-nucleotides signal the termination of protein synthesis in eukaryotes. *Nucleic Acids Res.* 18(21):6339-6345.
6. Liu W, Brock A, Chen S, Chen S, & Schultz PG (2007) Genetic incorporation of unnatural amino acids into proteins in mammalian cells. *Nat. Methods* 4(3):239-244.
7. Hao B*, et al.* (2004) Reactivity and chemical synthesis of L-pyrrolysine - the 22(nd) genetically encoded amino acid. *Chem. Biol.* 11(9):1317-1324.
8. Vogel HJ & Davis BD (1952) Glutamic gamma-semialdehyde and delta1-pyrroline-5-carboxylic acid, intermediates in the biosynthesis of proline. *J. Am. Chem. Soc.* 74:109-112.
9. Merlic CA, Motamed S, & Quinn B (1995) Structure Determination and Synthesis of Fluoro Nissl Green - an Rna-Binding Fluorochrome. *J. Org. Chem.* 60(11):3365-3369.
10. Vollmer J*, et al.* (2004) Characterization of three CpG oligodeoxynucleotide classes with distinct immunostimulatory activities. *Eur. J. Immunol.* 34(1):251-262.
11. Alexander J*, et al.* (1994) Development of high potency universal DR-restricted helper epitopes by modification of high affinity DR-blocking peptides. *Immunity* 1:751-761.