

Electronic Supplementary Information

Orthogonal ligation: a three piece assembly of a PNA-peptide-PNA conjugate.

Fabienne Burlina, David D. Dixon, Robert P. Doyle, Gérard Chassaing, Christopher N. Boddy, Philip Dawson and John Offer

⁵ Table of content

1- Protocol for the synthesis of the Tmb auxiliary	p S1
2- Protocols for solid-phase synthesis of peptides and PNAs	p S3
3- Protocol for PNA-peptide cysteine-mediated ligation	p S4
4- Protocol for PNA-peptide Tmb-mediated ligation	p S5
¹⁰ 5- Peptide, PNA, conjugate characterization	p S6
Abbreviations and references	p S6

1- Protocol for the synthesis of the Tmb auxiliary

¹⁵ **General Synthetic Techniques.** Reagents were purchased from Aldrich, Acros, and Fischer chemical companies at the highest commercial quality and used without further purification. Deuterated solvents were obtained from Cambridge Isotopes. Reactions were monitored by thin-layer chromatography (TLC) and carried out on 0.25 mm E. Merck silica gel plates (60F-254) using UV light as a visualizing agent and/or ninhydrin stain. Flash column chromatography was performed with E. Merck silica gel
²⁰ (60, particle size 0.040-0.063 mm). ¹H NMR and ¹³C NMR spectra were recorded on a Bruker DPX-300 spectrometer calibrated using residual undeuterated solvent as an internal reference. Mass spectra were recorded on a Bruker 12 Tesla APEX-Qe FTICR-MS with an Apollo II ion source.

3,4,5-trimethoxy-4-methoxybenzyl-thioether (2)

²⁵ 3,4,5-trimethoxythiophenol **1** (45 mg, 0.23 mmol) was dissolved in a mixture of 2 mL tetrahydrofuran and 2 ml methanol. PPh₃ (70 mg, 0.27 mmol) was added and allowed to stir 10 min, followed by slow addition of 60% NaH in mineral oil (26 mg, 0.64 mmol) and let stir 10 min. *p*-methoxybenzylchloride (34 μL, 0.25 mmol) was quickly added and let stir 30 min under argon at room temperature, at which time the solution was evaporated under vacuum. Flash column (10% EtOAc in hexanes) produced 67
³⁰ mg of white solid (95% yield). R_f = 0.27 (silica gel, 20% EtOAc in hexanes); ¹H NMR (300 MHz, CDCl₃) δ 7.20 (d, *J* = 8.7 Hz, 2H, ArH) 6.84 (d, *J* = 8 Hz, 2H, ArH), 6.53 (s, 2H, ArH), 4.04 (s, 2H,

SCH₂PMB), 3.83 (s, 3H, ArOCH₃), 3.78 (s, 9H, ArOCH₃); ¹³C NMR (75 MHz, CDCl₃) δ 159.0, 153.3, 137.5, 130.8, 130.2, 129.9, 114.0, 108.6, 61.1, 56.3, 55.5, 39.9; HRMS: +ESI; *m/z* Calc. for C₁₇H₂₀O₄SNa⁺ 343.0975, found 343.0956.

2,3,4-trimethoxy-6-(4-methoxy-benzylsulfanyl)-benzaldehyde (3)

Thioether **2** (0.476 g, 1.5 mmol) was dissolved in 8 mL *N,N*-dimethylformamide under argon at 0 °C, and POCl₃ (345 μL, 3.7 mmol) was added over 5 min and allowed to stir an additional 20 min. The flask was then removed from the ice bath and immediately put into a 90 °C oil bath and stirred for 2 h. The septum was replaced with a water condenser and saturated NaHCO₃ was added and the temperature increased to reflux for 2 h. The solution was then extracted with ethyl acetate, washed with brine, and dried with Na₂SO₄. Flash column chromatography (5-30% EtOAc in hexanes) gave 0.250 g of an off white solid (48% yield) and 0.185 g (39%) recovered starting material; R_f = 0.25 (silica gel, 30% EtOAc in hexanes); ¹H NMR (300 MHz, CDCl₃) δ 10.35 (s, 1H, ArCOH), 7.37 (d, *J* = 8.7 Hz, 2H, ArH) 6.88 (d, *J* = 8 Hz, 2H, ArH), 6.59 (s, 1H, ArH), 4.11 (s, 2H, SCH₂PMB), 3.99 (s, 3H, ArOCH₃), 3.82 (s, 3H, ArOCH₃), 3.80 (s, 3H, ArOCH₃); ¹³C NMR (75 MHz, CDCl₃) δ 189.4, 158.6, 158.5, 139.9, 138.8, 136.6, 129.2, 129.1, 127.8, 120.6, 105.2, 62.8, 61.4, 56.4, 37.6; HRMS: +ESI; *m/z* Calc. for C₁₈H₂₀O₅SNa⁺ 371.0924, found 371.0913.

***S*-(*p*-methoxybenzyl)-*N*-Boc-2,3,4-trimethoxy-6-mercaptobenzyl glycine (4)**

Aldehyde **3** (196 mg, 0.56 mmol), glycine (35 mg, 0.45 mmol), and NaCNBH₃ (96 mg, 1.5 mmol) were mixed in 10 mL dry methanol and stirred at room temperature for 4 h under argon and monitored for the disappearance of the limiting reagent, glycine, via TLC (40% MeOH in EtOAc). The solution was evaporated under vacuum and redissolved in 5 mL CHCl₃ under argon at room temperature. Triethylamine (95 μL, 0.68 mmol) and di-*t*-butyldicarbonate (195 μL, 0.9 mmol) were added and the reaction stirred for 2 h. Upon complete consumption of the intermediate free amine, the solution was loaded directly onto a silica column and eluted via gradient elution using 0.1% acetic acid and a 0→5% methanol gradient in chloroform to give a white solid (206 mg, 90% yield over two steps). R_f = 0.46 (silica gel, 4% MeOH, 0.1% AcOH in CHCl₃); ¹H NMR (300 MHz, CDCl₃) δ 7.14 (d, *J* = 8.7 Hz, 2H, ArH) 6.80 (d, *J* = 8 Hz, 2H, ArH), 6.61 (s, 1H, ArH), 4.51 (s, 2H, ArCH₂N), 3.96 (s, 2H, SCH₂PMB), 3.86 (s, 3H, ArOCH₃), 3.84 (s, 3H, ArOCH₃), 3.80 (s, 3H, ArOCH₃), 3.75 (s, 3H, ArOCH₃), 1.48 (s, 9H, ^{*t*}Bu); ¹³C NMR (75 MHz, CDCl₃) δ 159.2, 153.7, 153.5, 132.0, 130.4, 129.6,

124.1, 114.3, 111.8, 80.9, 77.6, 61.4, 61.2, 56.3, 55.7, 47.0, 43.5, 43.1, 40.1, 30.1, 28.7; HRMS: +ESI; m/z Calc. for $C_{25}H_{33}NO_8SNa^+$ 530.1819, found 530.1819.

2- Protocols for solid-phase synthesis of peptides and PNAs

5 Compound sequences are mentioned Table S1.

C-terminal amide peptides

Peptides were assembled manually on MBHA-PS resin applying the protocol of *in situ* neutralization/HBTU activation for Boc chemistry.¹ Couplings were carried out with a 10-fold excess
10 of activated amino acid for a minimum of 20 min and monitored by the Kaiser Test. HBTU activation was also used for the coupling of the Boc protected Tmb-Gly residue (1.5 eq. of Tmb-Gly/ peptide-resin) performed as previously reported.² Cys or Tmb-Gly were coupled on the side-chain amino group of lysine residues flanking the peptide (Table S1). The side-chain of these lysine residues were protected by Fmoc, other lysine residues of the peptide were protected by 2-ClZ. After peptide
15 assembly was complete the peptide-resin was treated with HF containing 5% (v/v) *p*-cresol for 1h at 0°C. After evaporation of HF the peptide was precipitated in anhydrous Et₂O, dissolved in HPLC buffer and lyophilized. Peptides were purified by reverse-phase HPLC on a C8 or C18 column using a linear gradient of acetonitrile in water/ 0.1% TFA.

20 C-terminal β-mercaptopropionic acid leucine (MPAL) thioester PNAs (T₄ and CT₄)

TAMPAL resin was prepared according to the previously published procedure.³ Briefly, *S*-trityl-β-mercaptopropionic acid was activated with HBTU (0.9 eq.) in presence of DIEA (3 eq.) and coupled to Leu-PAM resin (0.66 eq.) for 4 h. Trityl was removed by treatment with TFA/H₂O/triethylsilane (95:2.5:2.5, v/v/v), 3 x 1 min. PNA were then assembled on the TAMPAL resin applying the protocol
25 of *in situ* neutralization/HBTU activation for Boc chemistry as reported previously.⁴ Cleavage of the PNA thioester from the resin was performed using the standard HF procedure. After evaporation of HF the PNA was precipitated in anhydrous Et₂O, dissolved in HPLC buffer and lyophilized. The crude PNA-MPAL was analyzed by HPLC on a RP-C8 column heated at 55 °C, using a linear gradient of acetonitrile in water/ 0.1% TFA. The synthesized PNAs were found to be sufficiently pure (> 95%
30 purity) to be used in the ligation reaction without further purification. PNA-MPAL were isolated with 35% yield.

3- Protocol for PNA-peptide cysteine-mediated ligation (NCL): assembly of PenCys-T₄ and TmbPenCys-T₄ (compound 6).

The peptide (0.3 μ mol of PenCys or 2.5 μ mol of TmbPenCys) was mixed with 1 to 1.5 equivalent of T₄ PNA thioester in degassed sodium phosphate buffer (200 mM, pH 8.4) containing 6 M guanidine hydrochloride, 2 mM EDTA, 2% thiophenol and 2% benzylmercaptan (peptide and PNA final concentration: 1 to 2 mM). The pH was checked and carefully adjusted to pH 7 if necessary. The ligation reaction was followed by HPLC (RP-C4 column). The reaction was completed overnight (Fig. S1). The products were purified by HPLC on a RP-C4 column using a linear gradient of acetonitrile in water/ 0.1% TFA and characterized by mass spectrometry (Table S1). Isolated yields: 58% for PenCys-T₄, 55% for TmbPenCys-T₄ (**6**).

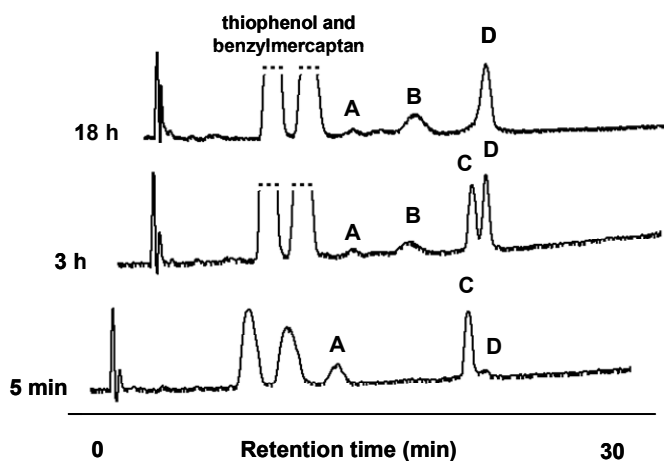
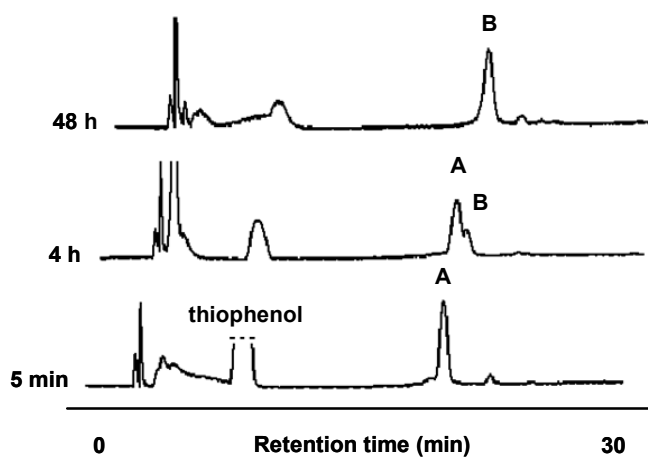


Figure S1. Analytical HPLC of the time course for the ligation reaction between PenCys and PNA T₄. (A=T₄-MPAL, B=T₄-COCH₂SPh, C=PenCys, D=PenCys-T₄).

4- Protocol for PNA-peptide Tmb-mediated ligation (EL): assembly of T₄-TmbPen and CT₄-TmbPenCys-T₄ (compound 7*)

5

Sodium phosphate buffer (200 mM, pH 8.4) containing 6 M guanidine hydrochloride, 2 mM EDTA was carefully degassed under vacuum before addition of TCEP (5 mg/mL) and thiophenol (2%). Peptide TmbPen (0.3 μ mol) and the PNA T₄ (0.45 μ mol) were then successively added (fragment final concentration: 5 to 10 mM). The pH was adjusted to 7.5 by addition of phosphate buffer pH 10. The ligation reaction was monitored by HPLC. Analytical and preparative HPLC were performed on a RP-C4 column using a linear gradient of acetonitrile in water/ 0.1% TFA. Reaction was complete after 48 h and gave the expected T₄-TmbPen conjugate (Fig. S2, isolated yield: 45%). The same procedure was used for the ligation between TmbPenCys-T₄ and the PNA CT₄ to synthesize the CT₄-TmbPenCys-T₄ conjugate (7*). In this case analytical and preparative HPLC were performed using a RP-C18 column (isolated yield: 41%). The Tmb auxiliary was removed after ligation by HF treatment for 30 min at 0°C as described previously.² After HF evaporation, the conjugate CT₄-PenCys-T₄ (compound 7) was precipitated in cold ether, dissolved in HPLC buffer, freeze-dried and characterized by mass spectrometry (Table S1). The Tmb removal reaction was quantitative.



20

Figure S2. Analytical HPLC of the time course for the ligation reaction between TmbPen and T₄-MPAL. (A=TmbPen, B=T₄-TmbPen).

25

5- Peptide, PNA, conjugate characterization

PNAs, peptides, and conjugates were characterized by MALDI-TOF mass spectrometry (positive ion reflector mode using α -cyano-4-hydroxycinnamic acid as matrix) or electrospray ionization (ESI) mass spectrometry.

Table S1. Compound sequence and characterization

Name	Sequence	m/z (M+H) ⁺ measured	m/z (M+H) ⁺ calculated
PNA T ₄	(aeg.T) ₄ -MPAL	1284.5 ^a	1284.5
PNA CT ₄	aeg.C(aeg.T) ₄ -MPAL	1535.7 ^a	1534.6
PenCys	HKRQIKIWFQNRRMKWKKNH ₂ HC	2476.5 ^a	2476.4
TmbPen	HKRQIKIWFQNRRMKWKKNH ₂ TmbG	2644.0±0.5 ^b	2644.3
TmbPenCys	HKRQIKIWFQNRRMKWKKNH ₂ TmbG HC	2747.0±1.0 ^b	2747.4
PenCys-T ₄	HKRQIKIWFQNRRMKWKKNH ₂ (aeg.T) ₄ C	3540.2 ^a	3540.8
T ₄ -TmbPen	HKRQIKIWFQNRRMKWKKNH ₂ (aeg.T) ₄ TmbG	3710.0±1.2 ^b	3709.3
TmbPenCys-T ₄	HKRQIKIWFQNRRMKWKKNH ₂ TmbG (aeg.T) ₄ C	3812.0±0.5 ^b	3812.5
CT ₄ -TmbPenCys-T ₄	HKRQIKIWFQNRRMKWKKNH ₂ aeg.C(aeg.T) ₄ TmbG (aeg.T) ₄ C	5129.0±1.1 ^b	5128.7
CT ₄ -PenCys-T ₄	HKRQIKIWFQNRRMKWKKNH ₂ aeg.C(aeg.T) ₄ G (aeg.T) ₄ C	4914.0±0.6 ^b	4913.3

^a Value of the first isotope of the (M+H)⁺ signal obtained from the MALDI-TOF mass spectra. ^b Value obtained from the reconstituted ESI mass spectra.

Abbreviations: aeg, *N*-(2-aminoethyl)-glycine PNA unit functionalised by thymine (aeg.T) or cytosine (aeg.C); DIEA, diisopropylethylamine; dsDNA, double stranded DNA; EDTA, ethylenediamine tetraacetic acid; EL, extended ligation; HBTU, *N*-[(1*H*-benzotriazol-1-yl)(dimethylamino)methylene]-*N*-methylmethanaminium hexafluorophosphate *N*-oxide; MALDI-TOF MS, matrix-assisted laser desorption/ionization time-of-flight mass spectrometry; MBHA-PS, 4-methylbenzhydrylamine polystyrene; MPAL, β -mercaptopropionic acid leucine; NCL, native chemical ligation; TCEP, tris(2-carboxyethyl)phosphine; TFA, trifluoroacetic acid; Tmb=4,5,6-trimethoxy-2-mercaptobenzyl.

References

- 1 M. Schnölzer, P. Alewood, A. Jones, D. Alewood, S. B. Kent, *Int. J. Pept. Protein. Res.*, 1992, **40**, 180.
- 2 J. Offer, C. N. Boddy, P. E. Dawson, *J. Am. Chem. Soc.*, 2002, **124**, 4642.
- 3 T. M. Hackeng, J. H. Griffin, P. E. Dawson, *Proc. Natl. Acad. Sci. U S A*, 1999, **96**, 10068.
- 4 S. Bregant, F. Burlina, J. Vaissermann, G. Chassaing, *Eur. J. Org. Chem.*, 2001, 3285.