Access to phosphoproteins and glycoproteins through semi-synthesis, Native Chemical Ligation and $N \rightarrow S$ acyl transfer†

Jinit Masania,[a] Jiejin Li,[b] Stephen J. Smerdon[b] and Derek Macmillan*[a]

Received 30th June 2010, Accepted 9th August 2010
DOI: 10.1039/c0ob00363h

Peptide thioesters are important tools for protein synthesis and semi-synthesis through their use in Native Chemical Ligation (NCL). NCL can be employed to assemble site-specifically modified proteins that can help elucidate the mechanisms of biomolecular processes. In this article we explore the compatibility of phosphopeptide synthesis and glycopeptide synthesis with thioester production through $N \rightarrow S$ acyl transfer.

Introduction

For chemical reactions to have an impact on biology they are often required to exhibit remarkable robustness, selectivity, and reactivity on encountering a specific target.[1] One such reaction, Native Chemical Ligation (NCL), has emerged as being particularly useful for assembling proteins tools that can probe biomolecular processes *in vitro*[2] and *in vivo*.[3] Hundreds of biomolecular studies have utilised NCL though application to proteins containing post-translational modifications (PTM's) such as phosphorylation and glycosylation took longer to emerge.[4] This was likely due, in part, to the plethora of available systems for study that did not require access to the more chemically complex PTM's and the practical barriers to studying the importance of *e.g.* protein glycosylation as well as the difficulty in producing the required chemical components. NCL generally requires the availabilty of a peptide thioester and this is most readily achieved using *tert*-butyloxycarbonyl-based solid phase peptide synthesis which utilises strongly acidic conditions that may not be compatible with chemically fragile protein appendages.[5] The thioester combines with a second peptide component adorned with a suitable coupling partner, usually (but not limited to) an N-terminal cysteine residue, resulting in a native peptide bond in the assembled protein (Fig. 1a).[6] The fact protein thioesters and protein fragments containing N-terminal cysteines can be expressed, albeit in unmodified form, from bacteria in addition to being chemically synthesised renders NCL an attractive method for studying the mechanisms of biomolecular reactions and interactions by chemists and biologists alike.

Recently we observed that peptide thioesters, the key tools for native chemical ligation, can be prepared following Fmoc-based solid phase peptide synthesis of peptides terminating in an Xaa-Cys amino acid sequence, particularly where Xaa = histidine, cysteine or glycine (Fig. 1b).[7] Thioesters are formed upon heating thioester precursors to 50–60 °C under mildly acidic conditions (usually approximately pH 5) in an excess of thiol component such as sodium 2-mercaptoethanesulfonate (MESNa). While we have demonstrated that *N*-glycopeptide analogues appeared stable we were keen to explore the effect of the reaction conditions on peptides containing potentially more chemically fragile motifs such as *O*-linked glycosylation and phosphorylation. If phosphopeptide and glycopeptide thioesters can be prepared then this constitutes a very simple route to these important tools for chemical biology, enabling us to dissect the biological importance of these indispensible classes of post-translational modification since homogeneous samples of stoichiometrically modified glyco- and phosphoproteins are not generally available from natural sources.[8]

Glyco/Phospho peptide synthesis

Glycoproteins are characterised by a covalent linkage between carbohydrates and the side-chains of certain amino acid residues

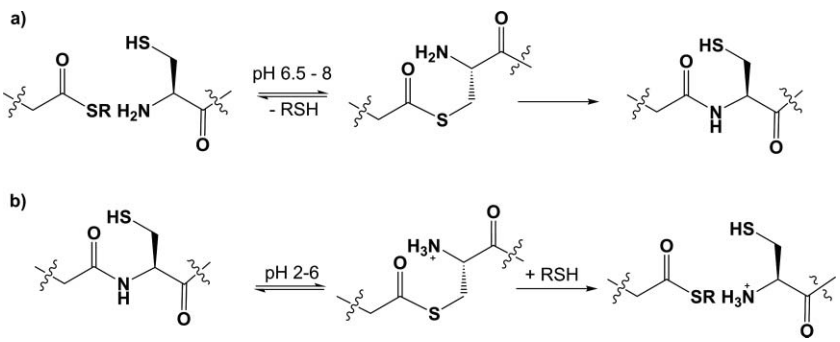
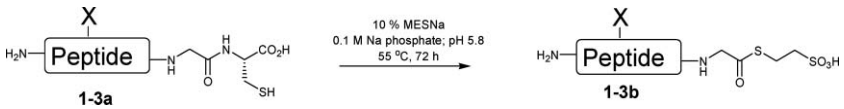
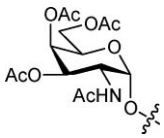
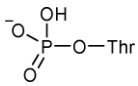
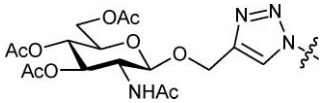


Fig. 1 a) Native chemical ligation and b) thioester formation through $N \rightarrow S$ acyl shift.

[a]*Department of Chemistry, University College London, 20 Gordon Street, London, UK WC1H 0AJ. E-mail: d.macmillan@ucl.ac.uk.; Tel: +44 (0)20 7679 4684*
[b]*Division of Molecular Structure, MRC National Institute for Medical Research, The Ridgeway, London, UK NW7 1AA*
† This paper is part of an *Organic & Biomolecular Chemistry* web theme issue on chemical biology.

Table 1 Conversion of modified peptides into their corresponding MESNa thioesters



Compound	Peptide	X	Mass calc.	Mass obs.	Isolated yield
1b	GlyCAM-1 (residues 75-82: H-GSSQLEE**T**SGC-OH)		1447.5	1448.4	37%
2b	Chk2 FHA domain (residues 63-72: H-LETVS*p***T**QELGC-OH)		1280.3	1281.2	54%
3b	CD52(N3Aha) (H-GQ**Aha**ETSQTSSPSGC-OH)		1800.8	1801.8	52%

within a protein sequence. In the case of *N*-linked glycoproteins a β-*N*-glycosidic linkage connects *N*-acetyl glucosamine to the sidechain amide group of asparagine in Asn-Xaa-Ser/Thr consensus motifs. *O*-linked glycosylation occurs most commonly as an α-*O*-glycosidic linkage between *N*-acetyl galactosamine and the hydroxyl groups of serine or threonine though *N*-acetyl glucosamine, mannose, xylose and fucose are also found linked to proteins *via* serine and threonine.[9] Several groundbreaking studies have shown that the total chemical synthesis of certain native glycopeptide components of glycoproteins is possible, though not yet routine,[10] and the difficulty associated with glycopeptide synthesis has been a barrier to much biological research using entirely chemical approaches. Since many therapeutic proteins are glycoproteins numerous semi-synthetic strategies for their preparation have been explored,[11] and recent developments in engineered cell culture employing yeast or *campylobacter jejuni* suggests that protein therapeutics with a human-like glycosylation profile can be achieved.[12] An inherent disadvantage of such systems is that, like glycoproteins (and phosphoproteins) of human origin, they are heterogeneous with respect to glycosylation (or phosphorylation) site-occupancy. In the case of phosphoproteins, many are hyperphosphorylated on their serine and threonine sidechains and consequently determining the contribution of specific phosphorylation events to protein activity is a lengthy process.

A powerful feature of NCL is that a synthetic component can be stoichiometrically and site-specifically adorned with phosphorylation or glycosylation, retaining native linkages. Indeed the majority of native, chemically synthesised, *N*/*O*-linked glycoproteins have been prepared using this method.[13] The use of NCL to dictate the identity and the position of glycosylation allows control and versatility that is limited only by the imagination of the synthetic chemist, rendering NCL the most widely used non cell-based approach to glyco- and phosphoprotein synthesis. Since it is widely accepted that approximately fifty amino acid residues represents the limit for efficient automated solid phase peptide synthesis (SPPS) the optimisation of recombinant methods for the production of one or the other of the required components has also been vigorously pursued.[14] Microorganism-derived thioesters are most frequently generated by a commercially available intein-fusion expression system from New England Biolabs.

The peptides employed in this study (Table 1) were all prepared using standard Fmoc-based SPPS on commercially available preloaded Fmoc-Cys(Trt)-NovaSyn-TGT resin (Merck Biosciences) and the desired post-translational modifications were introduced using commercially available building blocks. Fmoc-GalNAc(OAc)$_3$-Thr-OH (Dextra Laboratories/Merck Biosciences) was used to introduce GalNAc at position T82 of glycosylation dependant cell adhesion molecule-1 (GlyCAM-1) residues 75–82 (sequence **1a**: GSSQLEE**T**SGC) as a representative *O*-glycopeptide.[15] Fmoc-Thr(PO(OBn)OH)-OH (Merck Biosciences) introduced phosphothreonine into residues 63–72 of human checkpoint kinase 2 (Chk2), which was selected as a representative phosphopeptide (sequence **2a**: LETVS*p***T**QELGC).[16] Additional unnatural modifications containing *O*-glycosidic linkages could be introduced employing the suitably protected L-azidohomoalanine derivative Fmoc–Aha–OH (Chiralix). Peptides containing natural PTM's (**1–2a**) were assembled, cleaved from the solid support, isolated using standard peptide synthesis protocols (see experimental details) and subjected to thioester formation. Under the reaction conditions (Table 1) we predict that the thioester precursor, each containing a C-terminal Gly-Cys motif, initially undergoes an *N*→*S* acyl shift which is facilitated by the acidic pH. Following this the excess thiol (MESNa) drives the reaction towards thioester formation, displacing the terminal cysteine residue by transthioesterification. Consequently we would expect the *N*→*S* acyl shift to be favoured by low pH and the transthioesterification to be accelerated by higher pH. In practice low pH (pH <2) tends to increase thioester hydrolysis and so, for practical purposes, operating between pH 2 and pH 6 appears to be most beneficial. Consequently, all peptides were incubated at 55 °C in 0.1 M sodium phosphate buffer; pH 5.8, containing 10% w/v MESNa for 72 h. In each case thioester formation was readily observed after 24 h though, in the case of glycopeptides, slow thioester hydrolysis also appeared to be

occurring upon prolonged exposure to the reaction conditions. The hydrolysed product could be isolated in small but significant quantities by reverse phase HPLC and ultimately compromised the isolated yield for the desired product.[19] Product hydrolysis appeared more severe at lower pH (in 10% v/v AcOH; pH ~2) though was still a relatively minor side reaction over the course of 72 h. Importantly, there was no evidence that the phosphate group or glycosidic linkages were being cleaved during the course of the reaction (Fig. 2), A peptide containing azidohomoalanine was elaborated to a further *O*-linked glycopeptide (**3a**) on solid phase employing two equivalents of the propargyl glycoside of *N*-acetylglucosamine **4** (Fig. 3). The propargyl glycoside was introduced using Cu(I) mediated triazole formation[17] to the Aha-containing analogue of the cell surface glycopeptide CD 52 (**3a**, sequence: GQ**Aha**ETSQTSSPSGC).[18] Following synthesis and HPLC purification the thioester precursor **3a** was isolated in an unoptimised, though reasonable, yield of 35% and then converted to thioester **3b** in 52% yield.

Although the phosphopeptide thioester **2b** was isolated in only 58% yield, the starting material was also recovered in 38% yield (thioester yield = 85% based on recovered starting material) and so the mass balance for the reaction was good considering the small quantities of materials employed, slight thioester hydrolysis, and losses associated with handling, analysis and HPLC purification, which are common in peptide synthesis.

Phosphoprotein semi-synthesis

While we had shown that glycopeptide and phosphopeptide thioesters could be prepared using $N \rightarrow S$ acyl shift, it was important to demonstrate that the fidelity of peptides and indeed proteins containing such PTM's was retained following exposure to the conditions required for thioester formation. To this end we prepared a stoichiometrically and site-specifically phosphorylated variant of the fork-head associated (FHA) domain of human Checkpoint Kinase 2 (Chk2). Chk2 is serine/threonine kinase that transduces signals from DNA breaks (Fig. 4).[16] The protein is comprised of three key domains.[20] The first in an N-terminal serine/threonine cluster domain (SCD), in which initial phosphorylation of Thr-68 is intimately linked to kinase activation since substitution of Thr-68 with alanine affords Chk2 with a severely impaired response to DNA damage. Thr-68 is known to

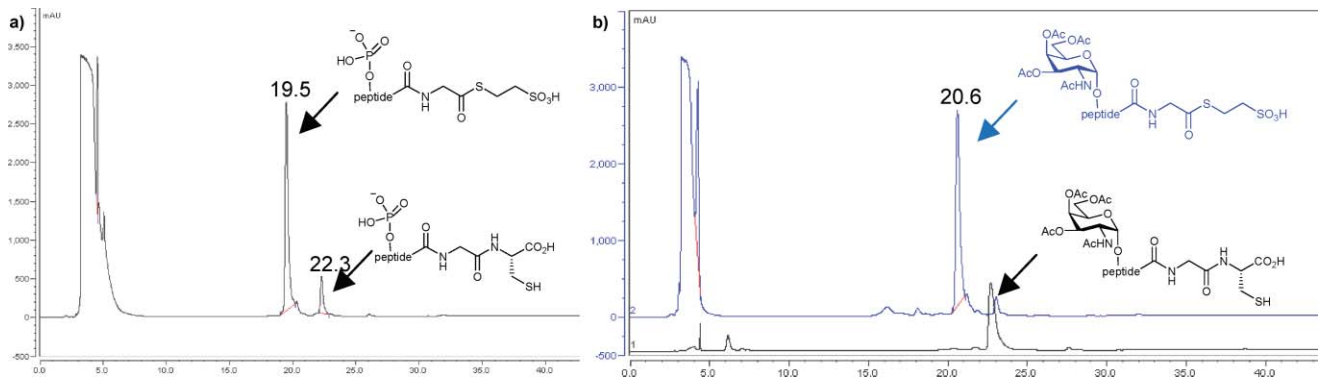


Fig. 2 Semi-preparative HPLC purification of thioesters from reaction mixtures. a) Purification of phosphopeptide thioester (t_R = 19.5 min) from unreacted starting material (t_R = 22.3 min). b) purification of GlyCAM-1 derived glycopeptide (upper trace, t_R = 20.6 min) from the reaction mixture (the starting thioester precursor (lower trace) is shown for comparison). In each case the thioester is the major product.

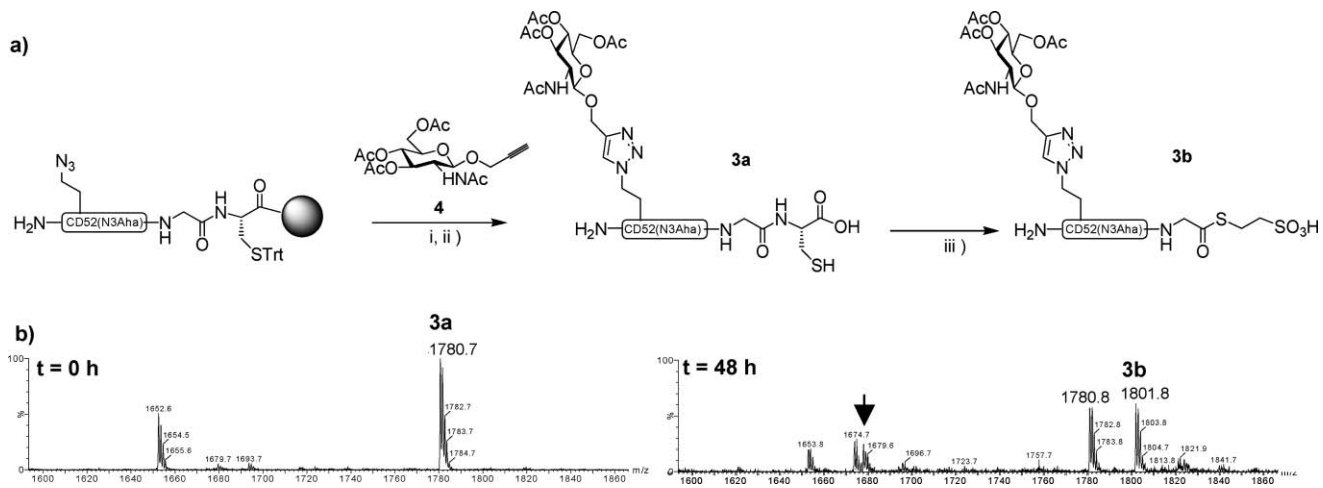


Fig. 3 a) Peptide modification of Aha containing peptides by solid-phase copper(I) mediated azide-alkyne ligation (click chemistry) and subsequent conversion to thioester. *Reagents and conditions*: i) sodium ascorbate, $CuSO_4$, 9 : 1 : 1 $CHCl_3$–EtOH/50 mM Na phosphate buffer; pH 7, 37 °C, 24 h. ii) TFA/EDT/H_2O (95 : 2.5 : 2.5), 4 h. iii) 10% w/v MESNa, 0.1 M sodium phosphate buffer, 55 °C, 72 h. b) MS analysis of the crude reaction mixture during thioester formation in a β-*O*-GlcNAc containing CD52 analogue. Thioester hydrolysis is indicated by the arrow.

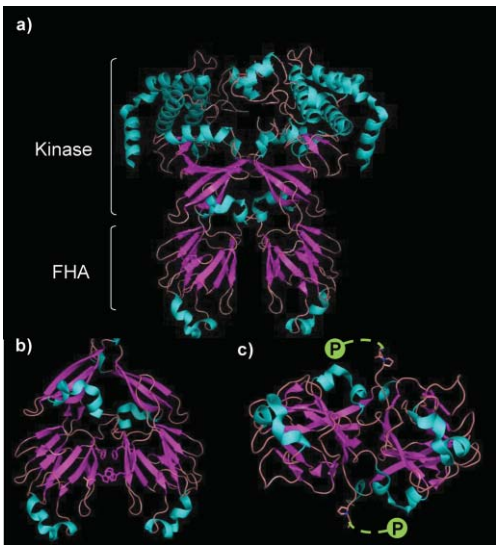


Fig. 4 a) Crystal structure of the human Chk2 dimer (residues 92-501, PBD ID 316U).[20] b) The Chk2 FHA domain showing hydrophobic contacts at the dimer interface. c) Alternative view of the FHA domain showing the relative positions of *p*Thr68 residues in an unresolved region of the crystal structure.

be phosphorylated by ataxia telangiectasia-mutated kinase (ATM) following DNA damage and this crucial phosphorylation event is believed to initiate Chk2 homo-dimerisation. Downstream of the SCD is a fork-head associated domain (FHA) which generally mediates protein-protein recognition events with cellular targets by binding to phosphothreonine motifs, and is believed to regulate Chk2 activation.[20,21] Thr-68 phosphorylation converts weak Chk2 self-association into a strong homodimeric complex that stimulates autophosphorylation and kinase activation. Once activated, the C-terminal kinase domain phosphorylates substrates which ultimately initiate apotosis. Since the disordered SCD has yet to be observed by crystallography, the mechanism of Chk2 activation and regulation by Thr-68 phosporylation has remained controversial. Previously we studied phosphate dependent FHA domain dimerisaton and showed that stoichiometric Thr-68 phosphorylation allowed the FHA domain to form a tight dimeric complex that could be separated upon treatment with lambda phosphatase.[21] In that study the thioester employed (sequence H_2N-LETVS*p*TQELY-SBn) was prepared using a sulfonamide safety catch linker. Notably the yield for synthesis of this short peptide was variable and often very low (approx. 1 mg recovered from a 0.05 mmol scale synthesis, yield ~ 2%). Encouragingly the thioester precursor employed in the present study H_2N-LETVS*p*TQELGC-CO_2H (isolated in 23%) yield was smoothly converted to thioester **2b** (Fig. 3a) in 57% yield. In order to facilitate thioester formation a Y72G substitution was required in the peptide sequence since we anticipated that formation of a tyrosyl thioester would be less efficient.

The Chk2 FHA domain (Chk2 residues 73–219) was expressed as a glutathione S-transferase (GST) fusion protein and isolated from *E. coli*, then treated with factor Xa to reveal an N-terminal cysteine (calculated mass = 17191.1 Da, observed mass = 17188.3 Da) as previously described.[22] The resulting protein fragment Ser73Cys-FHA was ligated to the phosphothreonine containing synthetic peptide under non-denaturing NCL conditions (see experimental section for further details) employing a 10-fold excess of the synthetic peptide. A new species corresponding to the ligated product (*p*T68-FHA) was observed by mass spectrometry (calculated mass = 18329.2 Da, measured mass 18327.8 Da) and further characterisation of the ligation product was undertaken using size-exclusion chromatography in combination with multi-angle laser light scattering (SEC-MALLS). As noted previously SEC-MALLS revealed that the stoichiometrically phosphorylated FHA domain formed a strong homodimeric complex (Fig. 5) with an average molecular weight of 3.108×10^4 Da for ligated FHA. Treatment of pT68-FHA with lambda phosphatase resulted in the formation of the monomeric species with an observed molecular mass of 1.788×10^4 Da confirming the specificity of the interaction.

Conclusion

Native Chemical Ligation is a powerful reaction for the synthesis of chemical biology tools such as specifically modified proteins and recently we observed a straightforward method for the synthesis of the thioester component involving $N \rightarrow S$ acyl transfer. While there are still obstacles to overcome associated with this method, an advantage is that no specialised resins, reagents or linkers are needed in order to bring about the synthesis of the required thioesters which may encourage further exploration of optimal reaction conditions and stimulate its wider use in NCL-type processes. We were concerned that the acidic conditions employed may limit application of thioester formation *via* $N \rightarrow S$ acyl shift to short peptides lacking acid sensitive post-translational modifications so in this study we investigated an *O*-linked glycopeptide (**1a**) and a phosphopeptide (**2a**). Encouragingly both precursors were transformed to the corresponding MESNa thioesters under typical reaction conditions providing access to multi-milligram quantities of thioester products. Although thioester hydrolysis was noted as a minor side reaction we did not observe cleavage of phosphates or glycosides. Phosphopeptide thioester **2b** was then joined to the FHA domain of Chk2 by Native Chemical Ligation and formation of a phosphate dependent dimer was observed. These results suggest that the fidelity of the phosphopeptide was retained following thioester formation and that the preferred Y72G substitution was tolerated.

We have presented the first application of thioester formation through $N \rightarrow S$ acyl shift of Gly-Cys terminated peptides, in combination with NCL, to the assembly of peptides and proteins bearing native post-translational modification and are hopeful that, through further study, this route will become a practical route to valuable thioester tools for chemical biology.

Experimental Details

General peptide synthesis procedure

NovaSyn-TGT resin, pre-loaded with Fmoc-Cys(Trt)-OH loading = 0.20 mmol g^{-1}, (250 mg, 0.05 mmol) was transferred to an automated peptide synthesiser (ABI 433A) reaction vessel for peptide chain elongation employing ten equivalents of each amino acid (Fastmoc protocol), HBTU/HOBt as coupling reagents and *N*,*N*-diisopropylethylamine (DIPEA) as base. When modified amino acids were to be introduced into the sequence five equivalents of

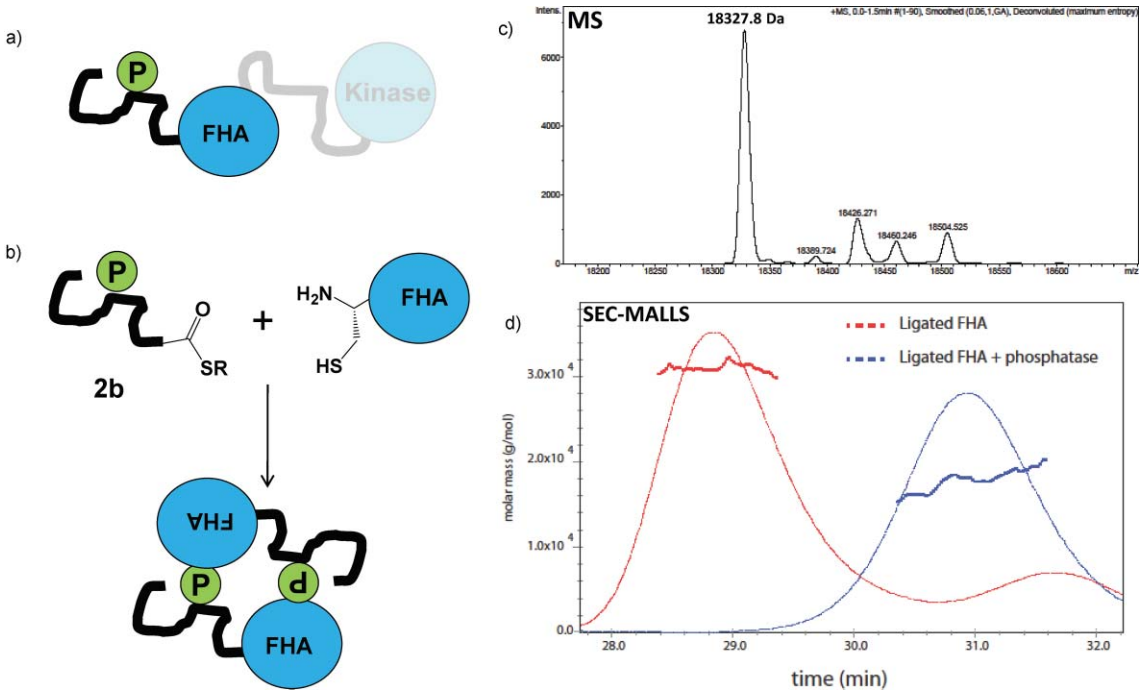


Fig. 5 a) cartoon representation of Chk2 from the N-terminal SCD (containing *p*Thr-68) to the C-terminal kinase domain. b) NCL reaction between phosphopeptide **2b** and bacterially expressed FHA domain to form the phosphate dependent homodimer. c) Mass spectrum of the ligated phosphoprotein (calculated mass = 18329.2 Da, measured mass 18327.8 Da). d) SEC-MALLS analysis of the ligation product. The faster eluting homodimer has an approximate mass of 31 KDa (red trace). Upon treatment with lambda phosphatase the monomeric FHA dominates (approximate mass = 18 KDa).

each modified amino acid was coupled (2 h) manually (coupling verified through the Kaiser ninhydrin test) then synthesis was continued in automated fashion. Following synthesis the resin was transferred to a glass vial and treated with trifluoroacetic acid: ethanedithiol: water (95: 2.5 : 2.5 = 4.0 mL) for 5 h, after which time the resin was filtered off and the filtrate was poured into cold diethyl ether (40.0 mL). The precipitate was collected by centrifugation at 3000 rpm, 4 °C for 15 min. The ether layer was then decanted and the precipitate was washed with cold diethyl ether (40.0 mL). After centrifugation the white precipitate was then dissolved in water and purified by semi-preparative reverse phase (RP)-HPLC (gradient: 5→50% acetonitrile/45 min). Fractions containing thioester precursors (**1–3a**) were identified by LC-MS and pooled. The purified products were obtained as fluffy white solids after lyophilisation in yields ranging from 20–45%.

Phosphopeptide thioester synthesis (2b)

A phosphopeptide corresponding to Chk2 FHA domain residues 63–72 (**2a**) was prepared as described above and isolated, after reverse-phase HPLC as a white fluffy solid (t_R = 22.3 min, 14.5 mg, 23%); calculated mass = 1259.3 Da, Observed mass = 1260.2. **2a** (11 mg, 8.7 μmol) was dissolved in a solution of 0.1 M sodium phosphate buffer; pH 5.8 (3.7 mL) and 300 mg (10% w/v) sodium 2-mercaptoethanesulfonate was added. *tris*(2-carboxyethyl)phosphine (approx 5 mg) was then added and the resulting mixture was agitated in an eppendorf thermomixer at 55 °C for 72 h. The reaction mixture was then purified by semi-preparative (RP)-HPLC (gradient: 5→50% acetonitrile/45 min). The two major species were collected and lyophilised to afford the purified thioester (t_R = 19.5 min, 6 mg, 54%) and the unreacted starting material (4 mg, 36%). LC-MS (ESI-MS) calculated average mass for **2b** = 1280.2 [M], found 1281.2 Da $[MH]^+$.

O-Glycopeptide thioester synthesis

Glycopeptide thioester formation was conducted exactly as described above**: GlyCAM-1 thioester precursor 1a** (t_R = 22.7 min, 28%), LC-MS (ESI-MS) calculated average mass = 1426.4 Da [M], found 1427.4 $[MH]^+$. **GlyCAM-1 thioester** (t_R = 20.6 min, 37%), LC-MS (ESI-MS) calculated average mass = 1447.5 [M], found 1448.4 $[MH]^+$.

2-Acetamido-2-deoxy-3,4,6-Tri-*O*-acetyl-β-D-propargyl glucopyranoside (4)

2-*N*-phthalimido-2-deoxy-1,3,4,6-tetra-*O*-acetyl-β-D-glucosamine (0.25 g, 0.53 mmol), propargyl alcohol (32 μl, 0.53 mmol) and activated 3 Å molecular sieves were stirred in anhydrous DCM (5.0 ml). The mixture was cooled to 0 °C under nitrogen, and borontrifluoride diethyl etherate (64 μl, 0.53 mmol) was added. The ice bath removed and the reaction mixture was stirred for 16 h at room temperature. The reaction mixture was monitored by LC-MS to completion and then diluted with DCM (10.0 ml), washed with water (10.0 ml), $NaHCO_3$ (10.0 ml), NaCl (10.0 ml) and water (10.0 ml). The organic phase was dried with $MgSO_2$, filtered and the solvent removed under reduced pressure to afford the crude propargyl glycoside as a yellow oil. The oil

was dissolved in *n*-BuOH (10.0 ml) and ethylenediamine (116 µl, 1.57 mmol) and the reaction stirred at 90 °C under reflux for 16 h. The resulting orange solution was reduced to an oil and dissolved in 2 : 1 pyrindine/acetic anhydride (30.0 ml) and stirred for 5 h. The reaction mixture was diluted with ethyl acetate (40.0 ml), washed with water (30.0 ml), $NaHCO_3$ (30.0 ml), NaCl (30.0 ml) and water (30.0 ml). The organic phase was dried with $MgSO_2$, filtered and the solvent removed under reduced pressure to afford an orange oil. The oil was purified by flash chromatography over silica (100% EtOAc, R_f 0.4) to afford the product a white solid (80 mg, 40% yield over 3 steps).^{1}H NMR (300 MHz, $CDCl_3$:δ (ppm) = 5.62 (1H, d, J = 8.91 Hz, HN), 5.30 (1H, dd, H3), 5.11 (2H, dd, H4), 4.86 (1H, d, J = 8.40 Hz, H1), 4.37 (2H, s, CH_2), 4.29–4.24 (1H, dd, H6a), 4.15–4.11 (1H, dd, H6b), 3.99 (1H, dd, H2), 3.75–3.69 (1H, m, H5) 2.27 (1H, s, alkyl-H), 2.08, 2.02, 2.02, 1.95 (12H, S, $CH_3 \times 4$). ^{13}C NMR (75 MHz, $CDCI_3$): δ (ppm) = 170.9, 170.7 170.3, 169.4 (4 × RCOR), 98.3, 78.5, 75.4, 72.4, 72.0, 68.5, 61.9 (7C, C1–C6, –CH_2*C*CH), 55.9 (1C, –*C*H_2CCH), 54.3 (1C, CH), 23.4, 20.7, 20.7, 20.6 (4C, CH_3). MS (ESI): m/z 408.30 (M+Na), $C_{17}H_{23}NO_9Na^+$ requires 408.36.

Phosphopeptide ligation

This ligation was performed as previously described. Briefly, the peptide thioester (1 mM) was incubated with 100 µM purified Factor Xa-digested FHA proteins overnight at room temperature in 200 mM phosphate buffer, pH 8.0, 150mM NaCl, 10 mM TCEP, 122 mM MESNa to form the ligated pThr-68 FHA product (pT68-FHA).

Size-Exclusion Chromatography (SEC)-Multiangle Laser Scattering measurements

Molecular weights and molecular weight distributions were determined by SEC conducted with a JASCO chromatography system composed of dual PU1580 pumps, an UV1575 uv/vis detector, a Dawn HELOS Laser photometer in series with an Optilab rEX interferometric refractometer (Wyatt Technology, Santa Barbara, CA), using a superdex S200 10/300 GL column (Pharmacia) eluted at a flow rate of 0.5 ml min^{-1} in 20 mM Tris (pH 8.0), 150mM NaCl, and 0.5mM TCEP. Molecular weights were analyzed with first-order Zimm using Astra (version 5.3.4.15), at a dn/dc value of 0.186 ml g^{-1}.

References

1 J. A. Prescher and C. R. Bertozzi, *Nat. Chem. Biol.*, 2005, **1**, 13; G. A. Lemieux and C. R. Bertozzi, *Trends Biotechnol.*, 1998, **16**, 506; K. L. Kiick, E. Saxon, D. A. Tirrell and C. R. Bertozzi, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 19; X. Wu and P. G. Schultz, *J. Am. Chem. Soc.*, 2009, **131**, 12497; C. P. R. Hackenberger and D. Schwarzer, *Angew. Chem., Int. Ed.*, 2008, **47**, 10030; A. Dirksen and P. E. Dawson, *Curr. Opin. Chem. Biol.*, 2008, **12**, 760.

2 P. E. Dawson, T. W. Muir, I. Clark-Lewis and S. B. Kent, *Science*, 1994, **266**, 776; G. G. Kochendoerfer and S. B. H. Kent, *Curr. Opin. Chem. Biol.*, 1999, **3**, 665; P. E. Dawson and S. B. H. Kent, *Annu. Rev. Biochem.*, 2000, **69**, 923; S. B. H. Kent, *Chem. Soc. Rev.*, 2009, **38**, 338–351.

3 H. D. Mootz, *ChemBioChem*, 2009, **10**, 2579; J. A. Camarero, D. Fushman, D. Cowburn and T. W. Muir, *Bioorg. Med. Chem.*, 2001, **9**, 2479; I. Giriat and T. W. Muir, *J. Am. Chem. Soc.*, 2003, **125**, 7180; S. W. Lockless and T. W. Muir, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 10999.

4 C. Chatterjee, Robert K. McGinty, J.-P. Pellois and T. W. Muir, *Angew. Chem., Int. Ed.*, 2007, **46**, 2814; C. Chatterjee and T. W. Muir, *J. Biol. Chem.*, 2010, **285**, 11045; R. R. Flavell, M. Huse, M. Goger, M. Trester-Zedlitz, J. Kuriyan and T. W. Muir, *Org. Lett.*, 2002, **4**, 165; M. Huse, M. N. Holford, J. Kuriyan and T. W. Muir, *J. Am. Chem. Soc.*, 2000, **122**, 8337; L. A. Marcaurelle, L. S. Mizoue, J. Wilken, L. Oldham, S. B. H. Kent, T. M. Handel and C. R. Bertozzi, *Chem.–Eur. J.*, 2001, **7**, 1129; K. A. Winans, D. S. King, V. R. Rao and C. R. Bertozzi, *Biochemistry*, 1999, **38**, 11700; Y. Shin, K. A. Winans, B. J. Backes, S. B. H. Kent, J. A. Ellman and C. R. Bertozzi, *J. Am. Chem. Soc.*, 1999, **121**, 11684.

5 D. Macmillan, *Angew. Chem., Int. Ed.*, 2006, **45**, 7668–7672.

6 J. Offer, C. N. C. Boddy and P. E. Dawson, *J. Am. Chem. Soc.*, 2002, **124**, 4642; F. Burlina, D. D. Dixson, R. P. Doyle, G. Chassaing, C. N. Boddy, P. Dawson and J. Offer, *Chem. Commun.*, 2008, 2785; P. Botti, M. R. Carrasco and S. B. H. Kent, *Tetrahedron Lett.*, 2001, **42**, 1831; D. Macmillan and D. W. Anderson, *Org. Lett.*, 2004, **6**, 4659; W. Qian and S. J. Danishefsky, *Angew. Chem., Int. Ed.*, 2007, **46**, 9248; J. Chen, Q. Wan, Y. Yuan, J. Zhu and S. J. Danishefsky, *Angew. Chem., Int. Ed.*, 2008, **47**, 8521; J. Chen, P. Wang, J. Zhu, Q. Wan and S. J. Danishefsky, *Tetrahedron*, 2010, **66**, 2277; C. Kan and S. J. Danishefsky, *Tetrahedron*, 2009, **65**, 9047; D. Crich and A. Banerjee, *J. Am. Chem. Soc.*, 2007, **129**, 10064; C. Haase, H. Rohde and O. Seitz, *Angew. Chem., Int. Ed.*, 2008, **47**, 6807; R. Okamoto and Y. Kajihara, *Angew. Chem., Int. Ed.*, 2008, **47**, 5402.

7 J. Kang and D. Macmillan, *Org. Biomol. Chem.*, 2010, **8**, 1993–2002; J. Kang, J. P. Richardson and D. Macmillan, *Chem. Commun.*, 2009, 407; J. Kang, N. L. Reynolds, C. Tyrrell, J. R. Dorin and D. Macmillan, *Org. Biomol. Chem.*, 2009, **7**, 4918; J. P. Richardson, C.-H. Chan, J. Blanc, M. Saadi and D. Macmillan, *Org. Biomol. Chem.*, 2010, **8**, 1351.

8 R. J. Payne and C.-H. Wong, *Chem. Commun.*, 2010, **46**, 21; D. P. Gamblin, S. I. v. Kasteren, J. M. Chalker and B. G. Davis, *FEBS J.*, 2008, **275**, 1949; S. I. van Kasteren, H. B. Kramer, H. H. Jensen, S. J. Campbell, J. Kirkpatrick, N. J. Oldham, D. C. Anthony and B. G. Davis, *Nature*, 2007, **446**, 1105.

9 R. G. Spiro, *Glycobiology*, 2002, **12**, 43R–56R.

10 B. Wu, J. Chen, J. D. Warren, G. Chen, Z. Hua and S. J. Danishefsky, *Angew. Chem., Int. Ed.*, 2006, **45**, 4116; G. Chen, Q. Wan, Z. Tan, C. Kan, Z. Hua, K. Ranganathan and S. J. Danishefsky, *Angew. Chem., Int. Ed.*, 2007, **46**, 7383; P. Nagorny, B. Fasching, X. Li, G. Chen, B. Aussedat and S. J. Danishefsky, *J. Am. Chem. Soc.*, 2009, **131**, 5792; Z. Tan, S. Shang, T. Halkina, Y. Yuan and S. J. Danishefsky, *J. Am. Chem. Soc.*, 2009, **131**, 5424–5431; P. H. Seeberger, *Nat. Chem. Biol.*, 2009, **5**, 368; G. a. J. L. Bernardes, B. Castagner and P. H. Seeberger, *ACS Chem. Biol.*, 2009, **4**, 703; T. J. Boltje, T. Buskas and G.-J. Boons, *Nat. Chem.*, 2009, **1**, 611.

11 D. P. Gamblin, E. M. Scanlan and B. G. Davis, *Chem. Rev.*, 2009, **109**, 131.

12 S. R. Hamilton and T. U. Gerngross, *Curr. Opin. Biotechnol.*, 2007, **18**, 387; F. Schwarz, W. Huang, C. Li, B. L. Schulz, C. Lizak, A. Palumbo, S. Numao, D. Neri, M. Aebi and L.-X. Wang, *Nat. Chem. Biol.*, 2010, **6**, 264.

13 D. Macmillan and C. R. Bertozzi, *Angew. Chem., Int. Ed.*, 2004, **43**, 1355; C. P. R. Hackenberger, C. T. Friel, S. E. Radford and B. Imperiali, *J. Am. Chem. Soc.*, 2005, **127**, 12882; H. Hojo, Y. Matsumoto, Y. Nakahara, E. Ito, Y. Suzuki, M. Suzuki, A. Suzuki and Y. Nakahara, *J. Am. Chem. Soc.*, 2005, **127**, 13720; N. Yamamoto, Y. Tanabe, R. Okamoto, P. E. Dawson and Y. Kajihara, *J. Am. Chem. Soc.*, 2008, **130**, 501; C. Piontek, P. Ring, O. Harjes, C. Heinlein, S. Mezzato, N. Lombana, C. Pöhner, M. Püttner, D. V. Silva, A. Martin, Franz X. Schmid and C. Unverzagt, *Angew. Chem., Int. Ed.*, 2009, **48**, 1936; C. Piontek, D. V. Silva, C. Heinlein, C. Pöhner, S. Mezzato, P. Ring, A. Martin, Franz X. Schmid and C. Unverzagt, *Angew. Chem., Int. Ed.*, 2009, **48**, 1941.

14 T. W. Muir, D. Sondhi and P. A. Cole, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 6705; J.-P. Pellois and T. W. Muir, *Curr. Opin. Chem. Biol.*, 2006, **10**, 487; D. Macmillan and L. Arham, *J. Am. Chem. Soc.*, 2004, **126**, 9530; T. J. Tolbert and C.-H. Wong, *Angew. Chem., Int. Ed.*, 2002, **41**, 2171; G. K. Busch, E. W. Tate, P. R. J. Gaffney, E. Rosivatz, R. Woscholski and R. J. Leatherbarrow, *Chem. Commun.*, 2008, 3369; B. L. Pentelute, A. P. Barker, B. E. Janowiak, S. B. H. Kent and R. J. Collier, *ACS Chem. Biol.*, 2010, **5**, 359.

15 D. Macmillan and C. R. Bertozzi, *Angew. Chem., Int. Ed.*, 2004, **43**, 1355.

16 S. Matsuoka, M. Huang and S. J. Elledge, *Science*, 1998, **282**, 1893.
17 V. V. Rostovtsev, L. G. Green, V. V. Fokin and K. B. Sharpless, *Angew. Chem., Int. Ed.*, 2002, **41**, 2596; M. Meldal and C. W. Tornøe, *Chem. Rev.*, 2008, **108**, 2952.
18 M. R. Pratt and C. R. Bertozzi, *J. Am. Chem. Soc.*, 2003, **125**, 6149; A. B. Diekman, E. J. Norton, K. L. Klotz, V. A. Westbrook and J. C. Herr, *Immunol. Rev.*, 1999, **171**, 203.
19 In our hands glycine thioesters appear particularly prone to hydrolysis and as much and 5% hydrolysed product was recovered. In general, very little hydrolysis is observed over a 48 h period.
20 Z. Cai, N. H. Chehab and N. P. Pavletich, *Mol. Cell*, 2009, **35**, 818–829.
21 J. Li, I. A. Taylor, J. Lloyd, J. A. Clapperton, S. Howell, D. MacMillan and S. J. Smerdon, *J. Biol. Chem.*, 2008, **283**, 36019–36030.