Published in final edited form as:

Angew Chem Int Ed Engl. 2018 September 03; 57(36): 11634–11639. doi:10.1002/anie.201805191.

Leveraging the Knorr Pyrazole Synthesis for the Facile Generation of Thioester Surrogates for use in NCL

Dillon T. Flood,

Department of Chemistry, The Scripps Research Institute, 10550 N. Torrey Pines Road, La Jolla, CA 92037 (USA)

Jordi C. J. Hintzen,

Department of Chemistry, The Scripps Research Institute, 10550 N. Torrey Pines Road, La Jolla, CA 92037 (USA)

Michael J. Bird,

Department of Chemistry, The Scripps Research Institute, 10550 N. Torrey Pines Road, La Jolla, CA 92037 (USA)

Philip A. Cistrone,

Department of Chemistry, The Scripps Research Institute, 10550 N. Torrey Pines Road, La Jolla, CA 92037 (USA)

Jason S. Chen, and

Director of the Automated Synthesis Facility, The Scripps Research Institute, 10550 N. Torrey Pines Road, La Jolla, CA 92037 (USA)

Philip E. Dawson

Department of Chemistry, The Scripps Research Institute, 10550 N. Torrey Pines Road, La Jolla, CA 92037 (USA), Dawson@scripps.edu

Abstract

Facile synthesis of C-terminal thioesters is integral to native chemical ligation (NCL) strategies for chemical protein synthesis. Here we introduce a new method of mild peptide activation, which leverages SPPS on an established resin linker and a classical heterocyclic chemistry to convert C-terminal peptide hydrazides to their corresponding thioesters via an acyl pyrazole intermediate. Peptide hydrazides, synthesized on established trityl chloride resins, can be activated in solution with stoichiometric acetyl acetone (acac), readily proceed to the peptide acyl-pyrazole. Acyl pyrazoles are mild acylating agents and are efficiently exchanged with an aryl thiol, that can then be directly utilized in NCL. The mild, chemoselective, and stoichiometric activating conditions allow this method to be utilized through multiple sequential ligations without intermediate purification steps.

Keywords

NCL; Pyrazole; Peptide Hydrazide; Chemoselective; Chemical Protein Synthesis

C-terminal thioester peptides are obligate intermediates in the synthesis and semi-synthesis of proteins by NCL.^[1] Several routes exist for the introduction of the C-terminal thioester into synthetic peptides, the most straightforward being Boc-SPPS peptide synthesis using thioester linkers.^[2,3] While the Boc-SPPS approach is highly efficient, Fmoc-SPPS has risen in popularity as it alleviates the need for dedicated HF cleavage apparatuses.^[4] The alkyl-thioester linkages employed in Boc-SPPS have limited utility in Fmoc-SPPS as they are degraded during the repeated Fmoc deprotection steps.^[5,6] Thus, multiple routes have been developed to introduce a C-terminal thioester into Fmoc-SPPS peptides.^[7] Some involve incorporating “safety-catch” linkers into the thioester linkages,^[8–13] or utilizing the natural *O*→*S* or *N*→*S* acyl shift that is observed in peptides as a post-SPPS thioesterification route.^[7,14,15] Although these routes have shown utility, it is more convenient and efficient to utilize a C-terminal moiety that is robust to Fmoc-SPPS conditions, and can be activated post synthesis for thioesterification. These methods include, but are not limited to on-resin cyclization of 3,4-diaminobenzoic acid (Dbz) and 3-amino-4-(methylamino)benzoic acid (MeDbz) to their corresponding N-acyl urea (**Figure 1A**),^[16,17] or solution phase oxidation of Dbz and C-terminal hydrazides to the corresponding benzotriazole and acyl-azide respectively (**Figure 1B**).^[18–22] While these solution-based methods have been employed in the synthesis of complex proteins, the required oxidants which are typically used in large excess cannot be employed in concert with N-terminal thiazolidines, a common peptide protecting strategy utilized in the synthesis of proteins requiring multiple ligations, aryl amines, or other redox-sensitive residues which can be incorporated via SPPS. Thus, the required activating agents may preclude their use as a general strategy.^[21] A robust method for the generation Fmoc-SPPS peptide thioesters compatible with multiple ligations would find broad application in synthetic protein chemistry.

The Knorr pyrazole synthesis, a classic route to substituted heterocycles first reported in 1883,^[23] suggests a robust method for chemoselective access to peptide acyl pyrazoles (**Figure 1C**). These weakly activated species are readily exchanged to their corresponding peptide thioesters under standard NCL conditions (**Figure 1C**). Under acidic conditions, the exquisite chemoselectivity and fast kinetics of the initial hydrazone imine formation^[24–26] and the intense thermodynamic drive towards aromaticity allow for peptide-hydrazide precursors to be converted into their corresponding acyl-pyrazole, using stoichiometric quantities of acetyl acetone (acac) even in the presence of sensitive functional groups (**Figure 1C**).^[27–30] Herein, we present an efficient method of Fmoc-SPPS peptide thioester generation, which utilizes mild and selective conditions at stoichiometric quantities, allowing for multiple ligations to be performed sequentially in one pot. This method has potential to provide a general route to the synthesis of large complex proteins requiring multiple native chemical ligations.

N-acyl heterocycles are known to act as efficient acylating agents^[31,32] but *N*-acylpyrazoles have been largely overlooked due to their modest reactivity, about 20 times less reactive than the corresponding *N*-acylimidazole.^[33,34] Although *N*-acylpyrazoles are widely studied in the realm of organic heterocyclic synthesis, their utility has been focused on ligands, and mildly activated substrates for organometallic transformations.^[35–39] The relatively low

reactivity towards endogenous nucleophiles^[33] and the ease of synthetic access presented an opportunity to adapt Knorr's synthesis for mild and selective *in situ* thioesterification.

To optimize thioester conversion, the effect of pH on pyrazole formation and subsequent thioesterification was studied. Model peptide-hydrazide (**1**) Ac-LYRAG-NH-HN₂ was synthesized by Fmoc-SPPS on trityl chloride resin (TentaGel, Rapp Polymere), according to previously published procedures.^[21] Peptide (**1**) was exposed to acac (2.5 equiv) in a solution of 6 M GdmCl and 2% v/v thiophenol while buffered at varying pH (**Figure 2B**). While acidic conditions markedly sped up the reaction, at higher pHs the desired reaction was not observed. At pH 7 the initial hydrazone was formed and trapped in moderate yields but no cyclization to the pyrazole was observed. The reaction at pH 8.5 yielded only starting material with minute amounts of hydrazone intermediate being observed, consistent with mechanistic studies on the pH dependence of hydrazone formation.^[24,25]

Various additives known to speed up the initial hydrazone ligation step were explored to assess if they would allow for efficient pyrazole formation and thioesterification.^[24,40] These additives included nucleophilic aniline catalysts, electrophilic Lewis acid catalysts, and metal acac complexes as pyrazole activators (**SI**). Although aniline catalysis allowed for the formation of the hydrazone intermediate quickly at higher pH, it also inhibited pyrazole cyclization thus trapping the hydrazone species (**SI**). Of the 11 electrophilic Lewis-acids tested, most significantly slowed thioesterification, with Ga(NO₃)₃ mildly increasing the thioesterification rate (**SI**). A study of the ability of metal acac complexes to induce pyrazole cyclization and thioesterification revealed that such complexes were inferior to the Lewis acids in promoting thioesterification (**SI**). Finally, no thioester formation was observed when hexafluoroacetylacetone or methyl vinyl ketone were used to induce pyrazole formation.

Upon consulting Knorr's initial report and the large body of subsequent literature, our inability to enhance pyrazole formation can be readily rationalized.^[23,27–30] Acid catalysis has a marked effect on both the rate of pyrazole formation (**Figure 3**), and subsequent thioesterification. Acidic aqueous conditions facilitate pyrazole formation as both the imine formation step and pyrazole cyclization step involve protonating a ketone oxygen to activate the corresponding carbonyl carbon for nitrogenous attack (**Figure 3**). Also, it is known that *N*-protonation of *N*-acylpyrazole (pK_a ~2.5)^[41] lowers the resonance energy of the amide bond by ~3 kcal/mol.^[39] Thus one would expect that even partial protonation of the pyrazole nitrogen would increase the rate of thiolysis. It was also observed that although aniline catalyzes hydrazone formation,^[24,25] it completely shut down the subsequent pyrazole cyclization. The lack of cyclization could be due to the relative stability of the aniline-iminium species which, may be kinetically stable towards attack by the amide nitrogen and subsequent cyclization (**Figure 3**). Similarly, most Lewis acids are known to stabilize the enolate which can be unproductive to pyrazole cyclization as the protonated ketone species is more reactive than the enolate form in this mechanism (**Figure 3**).^[42] Through similar reasoning, it is unsurprising that the metal acac complexes were unable to form the pyrazole as the acac ligands are stabilized enolates. Although none of these routes proved fruitful, they provided insight into the transformation of peptide acyl hydrazides to peptide acyl pyrazole under relevant conditions.

Upon testing different thiol sources, it was observed that mercaptophenyl acetic acid (MPAA) remarkably sped up the thioesterification reaction, nearing completion in one hour (**Figure 4**). This is most likely due to a combination of the increased solubility of MPAA and the lower pKa of the aryl thiol which may be more prone to exchange with the acyl pyrazole at low pH.^[43] The effect of acac concentration on the rate of thioesterification was then probed. The concentration of acac has no effect on the overall reaction progress, at concentrations relevant to NCL, when at least stoichiometric quantities acac are used. Peptide (**1**) at 1 mM, exposed to 10 eq or 1 eq acac was fully converted in about one hour (**Figure 4**). Interestingly, formation of the initial hydrazone was observed to be slower when using 1 eq acac but the subsequent, rate determining, intramolecular cyclization steps proceeded nearly identically and went to completion after one hour (**Figure 4**).

The applicability of this method for the conversion of peptide hydrazides with bulky or sterically hindered C-terminal amino acids to their corresponding thioesters was assayed using proline (**2**) and valine (**3**) as C-terminal residues. Peptides **2** and **3** at 1 mM in 6 M GdmCl at pH 3, with 200 mM MPAA were activated by the addition of 2.5 eq acac in water. Proline containing peptide **2**, was converted to the thioester much more slowly (**Figure 5**). Interestingly, the peptide was converted to the acyl pyrazole at a similar rate to peptide **1** but thiolysis proceeded much more slowly, probably due to steric hindrance. Peptide **3**, with a C-terminal valine was converted to the corresponding thioester at rates similar to the glycine containing peptide **1** (**Figure 5**). Also, peptide **4** containing an N-terminal thiazolidine was converted to the corresponding thioester quickly and without unwanted side products or degradation of the thiazolidine protecting group. Importantly, epimerization upon activation and thioesterification was determined to be below detectable limits (<2%) by HPLC (**SI**). Also, as with other thioester methods we expect C-terminal Asp or Glu may isomerize or hydrolyze during thioesterformation and ligation.^[44]

The observation that peptide hydrazides could be converted into thioesters quickly and with stoichiometric activating agent suggested a potential to use this chemistry to perform multiple sequential ligations, with peptides containing unprotected cysteine and hydrazide moieties without intermediate purification steps (**Figure 6**). Peptide **1** was dissolved in 6 M GdmCl (5 mg/ml) containing MPAA (200 mM) and acac (1 eq) was added to initiate thioesterification. After incubation (2 hrs at 25° C), **5** (1 eq) was dissolved in 6 M GdmCl, 200 mM K₂HPO₄, pH 9, containing TCEP (50 mM), and added to the thioesterification solution. The resulting reaction mixture was brought to pH 7 to initiate ligation. After incubation (6 hrs at 25° C) analysis by HPLC-MS clearly showed formation of peptide **6** (**Figure 6B**). The crude reaction mixture was acidified (pH 3–4), MPAA was added (200 mM) to increase thiol concentration, and acac was added (1 equiv.) to initiate thioesterification of **6**. After incubation (2 hrs at 25° C), the second ligation of **5** (1 equiv.) was initiated as described above. The reaction mixture was analyzed by HPLC-MS after incubation (6 hrs at 25° C) and clearly showed the formation of peptide **7** (**Figure 6**). This demonstrates the utility of this methodology to perform sequential ligations in a single pot without intermediate purification or isolation steps.

The pyrazole activation method was utilized to chemically synthesize a SpA protein domain, a model protein engineering target. Fragment **8** was synthesized on hydrazine-functionalized

chlorotriyl resin (TentaGel, Rapp Polymere) and fragment **10** was synthesized on Rink amide Chem-Matirx resin, both by previously reported protocols (**Figure 7**).^[7,21] Fragment **8** was dissolved to 5 mM in 6 M GdmCl with 200 mM MPAA, 2.5 eq acac (from a 150 mM stock in water) were added to the mixture and the reaction mixture was stirred for 4 hours to form thioester fragment **9** (**Figure 7A**). Cysteine-containing fragment **10** (1.2 eq) was dissolved to 6 mM in 6 M GdmCl with 200 mM K₂HPO₄ at pH 8.5, and added to the mixture of **9**. This ligation mixture was brought to pH 7 with the addition of 1 M NaOH and allowed to sit overnight (18 hrs). Analysis by LCMS (for mass spectra see Supporting Information) reveal complete conversion from hydrazide **8** to thioester **9**, and complete conversion from thioester **9** to the target protein **11** (**Figure 7B**).

Next, we chose to utilize the pyrazole activation method to chemically synthesize SDF-1a^[45,46] analog **15** which contains several sensitive selenomethionine residues. Hydrazide fragment **12** was synthesized on hydrazine-functionalized chlorotriyl resin (TentaGel, Rapp Polymere) and cysteine containing fragment **14** was synthesized on Wang Chem-matirx resin, both by previously reported protocols (**Figure 7**).^[7,21] Fragment **12** was dissolved to 20 mg/ml in 6 M GdmCl with 200 mM MPAA, 2.5 eq acac (from a 100 mM stock in water) were added to the mixture and the reaction mixture was stirred for 4 hours to form thioester peptide **13** (**Figure 8B**). Cysteine-containing fragment **14** (0.8 eq) was dissolved to 20 mg/ml in 6 M GdmCl with 200 mM K₂HPO₄ at pH 8.5 and 50 mM TCEP, and added to the mixture of **13**. This ligation mixture was brought to pH 7 with the addition of 1 M NaOH and allowed to sit overnight (18 hrs). A crude deconvoluted mass spectra of the reaction mixture reveals that while small amounts of fragment **13** still remain, protein **15** was generated as the major product (**Figure 8B**). After purification by mass directed preparatory HPLC, **18** was recovered (18 mg, 75%) demonstrating that this approach is robust and selective enough for use on large proteins containing sensitive selenomethionine (SeM) amino acids.

The selective manipulation of the peptide C-terminus remains a significant synthetic challenge with broad utility in peptide and protein chemistry. Maintaining applicability with standard Fmoc-SPPS and the wide range of unnatural side chains and post-translational modifications still present a critical challenge for the straightforward assembly of multiple peptide segments by NCL. A general Fmoc-SPPS compatible synthesis of peptide thioesters does not exist as most of the previously described methods are incompatible with at least one common moiety incorporated in chemical protein synthesis. The chemistry presented above quickly and selectively transforms peptide hydrazides into a peptide-thioesters without the use of superstoichiometric or redox active activating agents. We envision that the method presented herein will provide a robust and general approach for the generation of C-terminal peptide thioesters for use in chemical protein synthesis.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This work was funded by the National Institutes of Health (NIH R01 AI113867). Author J. C. J. Hintzen acknowledges support from the Stichting Fundatie van de Vrijvrouwe van Renswoude (AV 20170242). The authors would like to thank Elizabeth Billings at the Scripps Mass Spectrometry Core for their help in protein mass analysis. The authors would also like to thank Wolfgang Rapp, of Rapp Polymere, for his insight into SPPS resin selection.

References

- [1]. Dawson PE, Muir TW, Clark-Lewis I, Kent SB, Science 1994, 266, 776–779. [PubMed: 7973629]
- [2]. Schnölzer M, Alewood P, Jones A, Alewood D, Kent SBH, Int. J. Pept. Res. Ther 2007, 13, 31–44.
- [3]. Hackeng TM, Griffin JH, Dawson PE, Proc. Natl. Acad. Sci 1999, 96, 10068–10073. [PubMed: 10468563]
- [4]. Muttenthaler M, Albericio F, Dawson PE, Nat. Protoc 2015, 10, 1067–1083. [PubMed: 26086408]
- [5]. Camarero Julio A. and Mitchell Alexander R., Protein Pept. Lett 2005, 12, 723–728. [PubMed: 16305540]
- [6]. Li X, Kawakami T, Aimoto S, Tetrahedron Lett 1998, 39, 8669–8672.
- [7]. Behrendt R, White P, Offer J, J. Pept. Sci 2016, 22, 4–27. [PubMed: 26785684]
- [8]. Ingenito R, Bianchi E, Fattori D, Pessi A, J. Am. Chem. Soc 1999, 121, 11369–11374.
- [9]. Shin Y, Winans KA, Backes BJ, Kent SBH, Ellman JA, Bertozzi CR, J. Am. Chem. Soc 1999, 121, 11684–11689.
- [10]. Backes BJ, Ellman JA, J. Org. Chem 1999, 64, 2322–2330.
- [11]. Backes BJ, Virgilio AA, Ellman JA, J. Am. Chem. Soc 1996, 118, 3055–3056.
- [12]. Kenner GW, McDermott JR, Sheppard RC, J. Chem. Soc. Chem. Commun 1971, 0, 636–637.
- [13]. Rowland MM, Best MD, Chem. Biol 2010, 17, 1166–1168. [PubMed: 21095564]
- [14]. Botti P, Villain M, Manganiello S, Gaertner H, Org. Lett 2004, 6, 4861–4864. [PubMed: 15606085]
- [15]. Nagaïke F, Onuma Y, Kanazawa C, Hojo H, Ueki A, Nakahara Y, Nakahara Y, Org. Lett 2006, 8, 4465–4468. [PubMed: 16986926]
- [16]. Blanco-Canosa JB, Dawson PE, Angew. Chem 2008, 120, 6957–6961.
- [17]. Blanco-Canosa JB, Nardone B, Albericio F, Dawson PE, J. Am. Chem. Soc 2015, 137, 7197–7209. [PubMed: 25978693]
- [18]. Wang P, Shaw KT, Whigham B, Ramage R, Tetrahedron Lett. 1998, 39, 8719–8720.
- [19]. Wang P, Layfield R, Landon M, Mayer RJ, Ramage R, Tetrahedron Lett. 1998, 39, 8711–8714.
- [20]. Liao Y, Kong Y, Hu N, Jin Z, Wang P, Chem. Lett 2010, 39, 196–197.
- [21]. Zheng J-S, Tang S, Qi Y-K, Wang Z-P, Liu L, Nat. Protoc 2013, 8, 2483–2495. [PubMed: 24232250]
- [22]. Fang G-M, Li Y-M, Shen F, Huang Y-C, Li J-B, Lin Y, Cui H-K, Liu L, Angew. Chem. Int. Ed 2011, 50, 7645–7649.
- [23]. Knorr L, Berichte Dtsch. Chem. Ges 1883, 16, 2597–2599.
- [24]. Dirksen A, Dawson PE, Bioconjug. Chem 2008, 19, 2543–2548. [PubMed: 19053314]
- [25]. Kölmel DK, Kool ET, Chem. Rev 2017, 117, 10358–10376. [PubMed: 28640998]
- [26]. Kalia J, Raines RT, Angew. Chem. Int. Ed Engl 2008, 47, 7523–7526. [PubMed: 18712739]
- [27]. Habraken CL, Moore JA, J. Org. Chem 1965, 30, 1892–1896.
- [28]. Kirchhoff WH, J. Am. Chem. Soc 1967, 89, 1312–1316.
- [29]. Svenstrup N, Simonsen KB, Thorup N, Brodersen J, Dehaen W, Becher J, J. Org. Chem 1999, 64, 2814–2820. [PubMed: 11674350]
- [30]. Heller ST, Natarajan SR, Org. Lett 2006, 8, 2675–2678. [PubMed: 16774229]
- [31]. Katritzky AR, Suzuki K, Singh SK, 2004, 24.
- [32]. ROMANI S, MORODER L, BOVERMANN G, WÜNSCH E, Synthesis 1985, 1985, 738–742.
- [33]. Kashima C, HETEROCYCLES 2003, 60, 437.

- [34]. Aggarwal JS, Ray JN, J. Chem. Soc. Resumed 1930, 0, 492–493.
- [35]. Wang C, Harms K, Meggers E, Angew. Chem. Int. Ed 2016, 55, 13495–13498.
- [36]. Zhang H-J, Shi C-Y, Zhong F, Yin L, J. Am. Chem. Soc 2017, 139, 2196–2199. [PubMed: 28139912]
- [37]. Kashima C, Miwa Y, Shibata S, Nakazono H, J. Heterocycl. Chem 2003, 40, 681–688.
- [38]. Takahashi Y, Ikeda H, Kanase Y, Makino K, Tabata H, Oshitari T, Inagaki S, Otani Y, Natsugari H, Takahashi H, et al., J. Org. Chem 2017, 82, 11370–11382. [PubMed: 28968504]
- [39]. Meng G, Szostak R, Szostak M, Org. Lett 2017, 19, 3596–3599. [PubMed: 28640630]
- [40]. Kobayashi S, Eur. J. Org. Chem 1999, 1999, 15–27.
- [41]. Haynes WM, CRC Handbook of Chemistry and Physics, 91st Edition, Taylor & Francis Group, 2010.
- [42]. Tocher DA, Appl. Organomet. Chem 2000, 14, 172–173.
- [43]. Johnson ECB, Kent SBH, J. Am. Chem. Soc 2006, 128, 6640–6646. [PubMed: 16704265]
- [44]. Villain M, Gaertner H, Botti P, Eur. J. Org. Chem **n.d.**, 2003, 3267–3272.**n.d.**
- [45]. Campbell JJ, Hedrick J, Zlotnik A, Siani MA, Thompson DA, Butcher EC, Science 1998, 279, 381–384. [PubMed: 9430588]
- [46]. Dealwis C, Fernandez EJ, Thompson DA, Simon RJ, Siani MA, Lolis E, Proc. Natl. Acad. Sci. U. S. A 1998, 95, 6941–6946. [PubMed: 9618518]

A. Dawson, 2015 **B.** Ramage, Wang: 1998, 2007 and Liu (Unprotected): 2013 **C.** This Work **Figure 1:**

Some previously reported routes to C-terminal thioesters and active esters *via* **A**^[17] MeDBZ cyclization and **B**^[18-22] acyl hydrazide oxidation. **C** This work.

Figure 2:

A Peptide thioester formation scheme. **B** Conversion of **1** (1 mM) to corresponding thioester upon addition of 2.5 equiv acac in 6 M GdmCl at varying pHs

Figure 3:
Proposed mechanism of acid catalyzed acyl-pyrazole formation in aqueous solution

Figure 4:

A Conversion of **1** to corresponding thioester with varying thiol sources, 200 mM MPAA, 2% v/v thiophenol or 2% v/v ethyl-3-mercaptopropionate (1 mM peptide **1**, 2.5 mM acac in unbuffered 6 M GdmCl at pH 3). **B** Conversion of **1** to corresponding thioester with varying acac concentrations (1mM peptide **1**, 200 mM MPAA in 6M GdmCl at pH 3)

Figure 5:

A Model peptide variants. **B** Conversion of **1**, **2**, **3** and **4** to corresponding thioester over time (1mM peptide, 2.5 mM acac 200mM MPAA in 6M GdmCl at pH 3)

Figure 6:

A Sequential ligations of **1** and **5** to create peptides **6** and **7**. **B** HPLC-MS chromatograms of sequential ligations outlined above. Thioester formed with 1.0 eq. acac and 200 mM MPAA in 6 M GdmCl at pH 3. Subsequent ligation performed with 1 eq. Cys fragment at pH 7.

Figure 7:

A reaction scheme for the chemical synthesis of SpA domain. Thioester formed with 2.5 eq acac in 6 M GdmCl containing 200 mM MPAA. Ligation performed with 1.2 eq cys fragment **10** in 6 M GdmCl at pH 7. **B** HPLC-MS chromatograms of sequential ligations outlined above. For deconvoluted mass spectra see Supporting Information.

Figure 8:

A reaction scheme for the chemical synthesis of SDF-1a^[45,46] domain. Thioester formed with 2.5 eq acac in 6 M GdmCl containing 200 mM MPAA. Ligation performed with 0.8 eq cys fragment **14** in 6 M GdmCl at pH 7. **B** Crude deconvoluted mass spectra of the reaction mixtures outlined above.