Published in final edited form as:

Org Chem Front. 2017 April 1; 4(4): 495–499. doi:10.1039/C6QO00819D.

De novo synthesis of alkyne substituted tryptophans as chemical probes for protein profiling studies

R. N. Nair^{a,b}, J. J. Rosnow^a, T. A. Murphree^a, M. E. Bowden^a, S. R. Lindemann^c, and A. T. Wright^a

^aBiological Sciences Division, Pacific Northwest National Laboratory, Richland, WA.

^bR. N. Nair. Current address – David Geffen School of Medicine, Department of Psychiatry and Behavioral Sciences, University of California, Los Angeles.

^cS. R. Lindemann. Current address – Department of Food Science, Purdue University, West Lafayette, IN.

Abstract

De novo synthesis of alkynated tryptophan moieties as chemical probes for protein profiling studies, via an unexpected synthesis of Michael acceptor 12 is reported. Friedel Craft's alkylation of 12 and alkyne substituted indoles gave alkynated tryptophan moieties, which perform as chemical probe by incorporating into actively translating *Escherichia coli* cells, whereby the alkyne moiety enables multimodal analyses through click chemistry mediated attachment of reporting groups.

Bioisosteric amide bond replacement and mimicking of peptide secondary structure have been the most popular approach for developing peptidomimetic agents.^{1, 2} However, over the years, incorporation of unnatural α -amino acid derivatives has gained interest.^{3, 4} Non-proteinogenic α -amino acids with restricted conformational mobility have great potential to elucidate biologically active conformations of peptides.^{5, 6} There are several examples ranging from simple α -methylated to various α , β -amino acids.⁷ The Orphan and Tirrell research groups have demonstrated that incorporation of an unnatural amino acid possessing a terminal alkyne facilitates click chemistry of a fluorescent reporter to track active protein synthesis; also known as BONCAT (biorthogonal noncanonical amino acid tagging).⁸⁻¹¹ The use of nonproteinogenic α -amino acids has not only led to characterization of, and modification to, activity, stability, bioavailability and binding specificity of proteins and peptides, but has also enabled studies on protein folding, protein function and signal transduction.^{5, 12-14}

A major focus on synthesis of non-proteinogenic amino acids has emphasized modifications to the α -carbon. However, amino acids with aromatic side chains (histidine, phenylalanine, tyrosine and tryptophan) allow modification on the ring structure, thereby maintaining the functionality of the amino acid. Our group has developed chemical probes for important

† Footnotes relating to the title and/or authors should appear here.

Electronic Supplementary Information (ESI) available. CCDC 1516420. For experimental procedures, characterization data (¹H, ¹³C, MS and crystallographic data in CIF) of all compounds see DOI: 10.1039/x0xx00000x

metabolites, including unnatural amino acids, adding terminal alkynes capable of undergoing click-chemistry with an azide modified reporter group. These probes allow multimodal studies such as imaging and mass spectrometry-based proteomic studies which is the primary focus of research in our group.¹⁵⁻¹⁹ We now sought to develop an alkynylated version of tryptophan to evaluate both incorporation of unnatural amino acids into proteins, and characterize amino acid-protein binding.

Tryptophan residues are known to bind cations through π -interactions. This interaction has led to several approaches on the synthesis of Trp regioisomers where the alanine unit is not attached to the indole C-3, but to C-2, C-4 and C-7.²⁰ At the onset of our research we identified several reports on the synthesis of functionalized Trp moieties with a variety of substituents on the indole ring, including NO₂, halides, alkyl chains, ethers, and others.²¹⁻²³ To the best of our knowledge, no report of an alkyne substituted Trp moiety directly attached to C-2 or C4-C7 positions of the indole ring (Fig.1, A) has been reported, although propargylation of hydroxy tryptophan has been recently reported (Fig.1, B).²⁴

The retrosynthetic approach depicted in Scheme 1 shows that Pd catalyzed Sonogashira coupling should be an accessible route to incorporate an alkyne group on the indole ring. A similar approach has been successfully accomplished on iodotyrosine^{25, 26} where the amino acid was protected using 9-BBN, and Bonjoch et. al. reported the synthesis of 9-BBN protected hydroxy Trp and Tyr to study their activity against malignant melanoma.²⁷ We decided to use a similar approach in protecting bromo-tryptophan. 5- & 6- bromo-tryptophan were subjected to 9-BBN protection to give compounds **1** & **2**, and further, the indole N-H was protected using Boc anhydride to give **3** & **4** in near quantitative yields (Scheme 3).

To test the incorporation of an alkyne group on the indole ring, compound **1** was subjected to Sonogashira coupling using Pd (II) catalyst, Cu(I), base and TMS-acetylene. MS analysis was employed to analyse the reaction, which did not show any alkyne incorporated product. We therefore decided to study the effects of Pd catalyst, base, solvent and temperature on all four substrates to enhance the Sonogashira coupling reaction.

None of the conditions listed in Table 1 gave indication of product formation, and in most cases dehalogenation of starting material was observed. Further, bromides **2** and **4** were also subjected to the Molander trifluoroboronates,²⁸ but to no avail. With none of Pd catalysed Sonogashira coupling chemistry giving positive reaction outcomes, we further investigated if 5-hydroxy tryptophan would serve as a better substrate for alkyne incorporation.

As shown in Scheme 2, 5-hydroxy tryptophan was protected with 9-BBN to give **5**. We envisioned alkylating the hydroxy position of **5** to give **6** via direct *o*-alkylation of the hydroxy indole, using Karageorge and Macor's strategy,^{29, 30} however the desired *o*-alkylation was not achieved with this approach. Zhu, et. al.²⁴ have shown that using standard protecting groups like Boc (for amine) and ester (for acid), propargylation of 5-hydroxy tryptophan is possible. We speculated that the ionic nature of **5**, with 9-BBN protecting group, may have caused a decreased solubility in acetone leading to the failure. As a control, 5-hydroxy indole on treatment with propargylbromide in the presence of Cs₂CO₃ in acetone

gave *o*-alkylated product **14** (see table 3, entry 3).³¹ Therefore, we began to investigate the possibility of a *de novo* tryptophan synthesis from indole and substituted indoles.

It is well known that in the presence of a strong base the C-3 indole position is activated to promote nucleophilic attack via S_N2 reaction. For example, strong bases like *n*-BuLi³² and EtMgBr³³ have been used to bring about nucleophilic substitution at the indole C-3 position. As shown in Scheme 3a, we subjected indole to freshly prepared EtMgBr to successfully synthesize 3-benzyl indole (**7**) and 3-propargyl indole (**8**). With this success, we envisioned that bromide **10** (Scheme 3b), which can be obtained from serine methyl ester, would be an excellent substrate to bring about the synthesis of protected tryptophan **11**.

Based on the retrosynthetic approach shown in Scheme 3b, in our quest to synthesize bromide **10**, the amino group of methyl serine was protected as benzophenone imine **9** and subjected to Appel reaction using CBr₄ and PPh₃. However, instead of brominating the hydroxy group of **9** to give bromide **10**, the reaction yielded formation of intermediate **12** (Scheme 3c), an excellent Michael acceptor. Although a known compound,²² formerly used for the *de novo* syntheses of substituted tryptophan, synthesis of **12** via Appel reaction has not been reported to the best of our knowledge. We also observed the reaction to be highly concentration-dependent such that a concentrated reaction mixture (deep red-brown in colour) did not proceed in the forward direction. However, with significant dilution (observed as a bright yellow colour) of the starting reactants we observed complete conversion to **12**.³⁴ With this observation we further developed the process by performing a step-wise scale up of the synthesis of **12** from ester **9** as shown in Table 2.

With significant quantities of compound **12** in hand, we sought to use this key intermediate in our substituted indole synthesis. Bartolucci²³ and Balsamini²¹ have reported an electrophilic aromatic substitution approach using EtAlCl₂ and AlCl₃ as Lewis acids to synthesize substituted tryptophan moieties using compound **12**. We envisioned that a similar approach using alkyne-substituted indoles would enable successful incorporation of the desired alkyne on the aromatic ring of tryptophan.

With all necessary elements in hand, we set forth a test reaction by reacting the indole with **12** in the presence of AlCl₃ (Table 3; entry 1) in anhydrous DCM at 0 °C. Under these conditions, alkylation at the C3 position went smoothly in less than 2h as monitored by TLC. Next, 5-ethynyl indole (Table 3; entry 2) was subjected to alkylation conditions to give 5-ethynyl tryptophan, which was otherwise inaccessible with other reaction routes as described earlier. To extend the scope of this methodology we decided to study other terminal alkyne groups on the C5 position of indole. 5-hydroxy indole was subjected to *o*-alkylation using propargyl bromide and Cs₂CO₃ to give a mixture of *O*-alkylated (**14**) and *O*-,*N*-alkylated (**15**) products, which were isolated by column chromatography. These two compounds also gave desired tryptophan moieties in good yields when reacted with **12** in presence of AlCl₃ (Table 3; entries 3 and 4). Table 3 entries 5 and 6 show that long chain terminal alkynes on indole rings are accessible via this approach. Purity of the starting indole, compound **12** and maintaining anhydrous conditions was very important for the success of this methodology. Performing this reaction with impure starting material or non-anhydrous solvent resulted in a competing side reaction where dimerization of compound **12**

led to formation of side product **13**,³⁵ which was confirmed by X-ray crystallographic studies CCDC 1516420 (Fig 2).

To study the application of alkynylated tryptophans in protein synthesis, **19** was subjected to standard deprotection conditions using HCl and LiOH over a two-step process giving alkyne-substituted tryptophan **24** (Scheme 4). *Escherichia coli* BL21 cells were grown at 37 °C to an OD₆₀₀ of 1.0 on M9 media supplemented with 22 mM glucose and either 200 μM L-tryptophan or 200 μM **24**. To confirm that **24** is incorporated into proteins in actively translating *E. coli*, the cells were lysed and the fluorescent reporter group, azido-rhodamine, was added to proteins incorporating **24** via click chemistry. Proteins were separated by gel electrophoresis, and analysed for fluorescence (Figure 2).

As shown in figure **3a**, proteins isolated from *E. coli* grown in the presence of compound **24** displays strong fluorescence compared to the no alkyne control and no click chemistry control. Figure **3b** shows that after coomassie staining an equal amount of protein was loaded into each lane, thus the fluorescence signal is dependent on incorporation of **24** into proteins, demonstrating that compound **24** is a functional non-proteinogenic amino acid.

In summary, we have synthesized new alkyne substituted tryptophan moieties. Sonogashira couplings on 9-BBN protected 5- & 6-bromotryptophan was not successful, nor was *o*-alkylation of 9-BBN protected 5-hydroxy tryptophan (**5**). Bromination of the hydroxy group of **9** via Appel reaction yielded Michael acceptor **12**, an excellent substrate for the synthesis of alkyne substituted tryptophans via electrophilic aromatic substitution. We have applied this strategy to synthesize new alkyne-substituted unnatural tryptophan moieties and demonstrated their utility within a biological system. We further envision incorporation of diazirine species into the alkyl chain, enabling applications in protein profiling studies that require a covalent bond formation, such as proteomic analysis of proteins that bind tryptophan (e.g., TrpR).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We acknowledge the support of this work by the Laboratory Directed Research and Development Program at Pacific Northwest National Laboratory (PNNL), and the Foundational Scientific Focus Area program that is sponsored by the U.S. Department of Energy (DOE), Office of Biological and Environmental Research (OBER), as part of BER's Genomic Science Program. The authors would also like to thank Dr. Phillip Koech for in-depth discussions on the chemistry.

Notes and References

1. Perdih A, Dolenc MS. Curr. Org. Chem. 2007; 11:801–832.
2. Perdih A, Dolenc MS. Curr. Org. Chem. 2011; 15:3750–3799.
3. Chen PR, Groff D, Guo J, Ou W, Geierstanger B, Schultz PG. Angew. Chem. 2009; 48:4052–4055. [PubMed: 19378306]
4. Groff D, Chen PR, Schultz PG. ChemBioChem. 2010; 11:1066–1068. [PubMed: 20422671]

5. Gante J. *Angew. Chem.* 1994; 106:1780–1802. See also *Angew. Chem., Int. Ed. Engl.*, 1994, 1733(1717), 1699–1720.
6. Wang, L., Schultz, PG. *Chem. Commun. Cambridge, U. K.*: 2002. p. 1–11. DOI: 10.1039/b108185n
7. Maruoka, K., Ooi, T. *Chem. Rev. Vol. 103.* Washington, DC, U. S.: 2003. p. 3013–3028.
8. Dieterich, J. J. DCH, Gouzer G, Shadrin IY, Ngo JT, Triller A, Tirrell DA, Schuman EM. *Nat Neurosci.* 2010; 13:897. [PubMed: 20543841]
9. Dieterich, A. DCL, Graumann, J. J, Tirrell DA, Shumann EM. *Proc Natl Acad Sci.* 2006; 103:9482–9487. [PubMed: 16769897]
10. Dieterich, J. J. DCL, Link AJ, Graumann J, Tirrell DA, Shumann EM. *Nat Protoc.* 2007; 2:532. [PubMed: 17406607]
11. Hatzenpichler, S. RS, Tavormina PL, Babin BM, Tirrell DA, Orphan VJ. *Environmental Microbiology.* 2014; 16:2568–2590. [PubMed: 24571640]
12. Beene DL, Dougherty DA, Lester HA. *Curr. Opin. Neurobiol.* 2003; 13:264–270. [PubMed: 12850209]
13. Hendrickson TL, de Crecy-Lagard V, Schimmel P. *Annu. Rev. Biochem.* 2004; 73:147–176. [PubMed: 15189139]
14. Lu Y. *Curr. Opin. Chem. Biol.* 2005; 9:118–126. [PubMed: 15811795]
15. Anderson LN, Koeck PK, Plymale AE, Landorf EV, Konopka A, Collart FR, Lipton MS, Romine MF, Wright AT. *ACS Chem. Biol.* 2016; 11:345–354. [PubMed: 26669591]
16. Wiedner SD, Anderson LN, Sadler NC, Chrisler WB, Kodali VK, Smith RD, Wright AT. *Angew. Chem., Int. Ed.* 2014; 53:2919–2922.
17. Sadler NC, Melnicki MR, Serres MH, Merkley ED, Chrisler WB, Hill EA, Romine MF, Kim S, Zink EM, Datta S, Smith RD, Beliaev AS, Konopka A, Wright AT. *ACS Chem. Biol.* 2014; 9:291–300. [PubMed: 24168666]
18. Sadler NC, Angel TE, Lewis MP, Pederson LM, Chauvigne-Hines LM, Wiedner SD, Zink EM, Smith RD, Wright AT. *PLoS One.* 2012; 7:e47996. [PubMed: 23110155]
19. Chauvigne-Hines LM, Anderson LN, Weaver HM, Brown JN, Koeck PK, Nicora CD, Hofstad BA, Smith RD, Wilkins MJ, Callister SJ, Wright AT. *J. Am. Chem. Soc.* 2012; 134:20521–20532. [PubMed: 23176123]
20. Carlier PR, Lam PCH, Wong DM. *J. Org. Chem.* 2002; 67:6256–6259. [PubMed: 12182675]
21. Balsamini C, Diamantini G, Duranti A, Spadoni G, Tontini A. *Synthesis.* 1995:370–372. DOI: 10.1055/s-1995-3915.
22. Balsamini C, Duranti E, Mariani L, Salvatori A, Spadoni G. *Synthesis.* 1990:779–781. DOI: 10.1055/s-1990-27012.
23. Bartolucci S, Bartoccini F, Righi M, Piersanti G. *Org. Lett.* 2012; 14:600–603. [PubMed: 22208950]
24. Zhu B, Ge J, Yao SQ. *Bioorg. Med. Chem.* 2015; 23:2917–2927. [PubMed: 25801160]
25. Nair RN, Lee PJ, Rheingold AL, Grotjahn DB. *Chem. - Eur. J.* 2010; 16:7992–7995. S7992/7991-S7992/7981. [PubMed: 20544759]
26. Walker WHIV, Rokita SE. *J. Org. Chem.* 2003; 68:1563–1566. [PubMed: 12585904]
27. Bonjoch J, Drew MGB, Gonzalez A, Greco F, Jawaid S, Osborn HMI, Williams NAO, Yaqoob P. *J. Med. Chem.* 2008; 51:6604–6608. [PubMed: 18826202]
28. Molander GA, Katona BW, Machrouhi F. *J. Org. Chem.* 2002; 67:8416–8423. [PubMed: 12444619]
29. Karageorge GN, Macor JE. *Tetrahedron Lett.* 2011; 52:1011–1013.
30. Karageorge GN, Macor JE. *Tetrahedron Lett.* 2011; 52:5117–5119.
31. To eliminate the possibility of competing indole NH alkylation, 5 was subjected to Boc protection in DCM and cat. DMAP. Surprisingly, using DCM as solvent resulted in complete solubility of 5, however, Boc group was found to be incorporated on both OH and NH.
32. Sharma R, Chouhan M, Sood D, Nair VA. *Appl. Organomet. Chem.* 2011; 25:305–309.
33. DeGraw JI, Kennedy JG, Skinner WA. *J. Heterocycl. Chem.* 1966; 3:67–69.
34. Note, See supporting information for details on observed color change.

35. Being an excellent Michael acceptor, compound 12 can very easily undergo dimerization in presence of strong Lewis Acids. We were able to successfully isolate this product and identify the dimer via X-ray crystallography. See supporting information for more details.

Figure 1.
Representative alkynylated Tryptophan moieties.

Scheme 1.
Synthesis of 9-BBN protected bromo-tryptophans.

Scheme 2.
Attempting OH-propargylation of compound 5.

Scheme 3.

a. Activating C-3 position of indole; **b.** retrosynthetic approach for *de novo* tryptophan synthesis; **c.** synthesis of intermediate **12**.

Figure 2.
X-ray crystal structure of side product **13** (ellipsoid view).

Scheme 4.
Synthesis of alkyne substituted tryptophan **24**.

Figure 3. Poly-acrylamide gel analysis of proteins isolated from *E. coli* cells grown in the presence of L-tryptophan (lane 1 & 3) or compound **24** (lane 2 & 4). a) Fluorescence detection of incorporation of compound **24** via cycloaddition of azido-rhodamine (Lanes 1 & 2 no click chemistry controls, Lanes 3 & 4 click chemistry). b) Coomassie stain of proteins shows that all gel lanes contained equivalent protein loading. L = ECL Rainbow ladder.

Table 1

Failed attempts on Sonogashira coupling of 9-BBN protected 5/6-bromotryptophan.

Compd	Alkyne	Catalyst	Base	Solvent	T (°C)
1	TMS—≡		K ₂ CO ₃		25
2		Pd(PPh ₃) ₄	Cs ₂ CO ₃	THF	80
3		PdCl ₂ (PPh ₃) ₂	TEA	Dioxane	100
4		PdCl ₂ dppf	DiPEA	DMF	120
2	TMS—≡—BF ₃ K	PdCl ₂ (PPh ₃) ₂	Cs ₂ CO ₃	DMF:H ₂ O	80
4				5:1	

Reaction conditions: Standard Sonogashira coupling conditions were followed. Alkyne, Pd catalyst, base and solvent mixed in a reaction vessel, degassed by bubbling N₂ and subjected to heating. Reactions were analysed using LTQ-MS.

Table 2Scale up process: step-wise synthesis of **12**.

Entry	9 (g)	CBr ₄ (g)	PPh ₃ (g)	12 (g)	Yield (%)
1	0.10	0.145	0.185	0.061	65.0
2	0.25	0.365	0.463	0.15	64.0
3	0.50	0.731	0.925	0.32	68.0
4	2.50	3.660	4.630	1.57	67.0
5	10.0	14.631	18.526	5.02	71.0

Table 3

Synthesis of alkyne substituted tryptophan moieties.

Entry			Yield %
1		R = H (11)	63
2		R = $\text{C}\equiv\text{CH}$ (18)	51
3		R = $\text{CH}_2\text{CH}_2\text{OCH}_2\text{C}\equiv\text{CH}$ (19)	69
4		R = $\text{CH}_2\text{CH}_2\text{OCH}_2\text{C}\equiv\text{CH}$ and $\text{CH}_2\text{CH}_2\text{NHC}\equiv\text{CH}$ (20)	73
5		R = $\text{CH}_2\text{CH}_2\text{CH}_2\text{CH}_2\text{OCH}_2\text{C}\equiv\text{CH}$ (21)	71
6		R = $\text{CH}_2\text{CH}_2\text{CH}_2\text{CH}_2\text{C}(=\text{O})\text{OCH}_2\text{C}\equiv\text{CH}$ (22)	64

Reaction conditions: **12** (1 eq), corresponding indole (1.5 eq) and AlCl_3 (2 eq) were added to a reaction flask containing anhydrous DCM at 0 °C. Reaction was monitored by TLC and MS analysis.

^cIsolated yields reported after flash chromatography purification.

^aSee supporting information for synthesis of indole precursors.

^bCorresponding tryptophan moieties.