HHS Public Access

Author manuscript

Angew Chem Int Ed Engl. Author manuscript; available in PMC 2018 March 13.

Published in final edited form as:

Angew Chem Int Ed Engl. 2017 March 13; 56(12): 3177–3181. doi:10.1002/anie.201611202.

Pd-mediated Arylation of Lysine in Unprotected Peptides

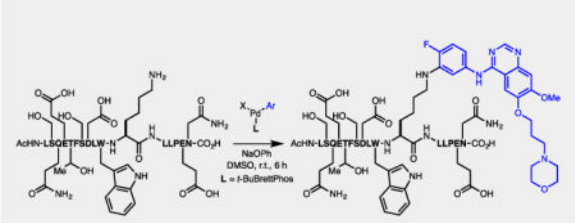
Dr. Hong Geun Lee[a], **Dr. Guillaume Lautrette**[a], **Prof. Bradley L. Pentelute**[*,a], and **Prof. Stephen L. Buchwald**[*,a]

[a]Department of Chemistry, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139 (USA)

Abstract

A mild method for the arylation of lysine in an unprotected peptide is presented. In the presence of a preformed biarylphosphine-supported Pd(II)-aryl complex and weak base, lysine amino groups underwent C–N bond formation at room temperature. The process generally exhibited high selectivity for lysine over other amino acids containing nucleophilic side chains and was applicable to the conjugation of a variety of organic compounds, including complex drug molecules, with an array of peptides. Lastly, this method was also successfully applied to the formation of cyclic peptides via macrocyclization.

Abstract



Keywords

bioconjugation; cross-coupling; phosphane ligand; peptide-macrocyclization; chemoselectivity

The chemical modification of amino acid residues in peptides has recently evolved into a valuable tool for the study of biological macromolecules.[1] The construction of covalent bonds to peptides provides a means of accessing novel macromolecules that contain molecular fragments of interest, such as affinity probes, chromophores, or medicinally active structures. Furthermore, covalent modifications can enhance the therapeutic potential of a bioactive peptide by extending its circulation half-life and augmenting cell permeability.[1a,1b] To broaden the range of synthetically accessible altered biomolecules, additional methods for bioconjugation are needed. Procedures that do not degrade the peptide[2] and provide the amended biomolecule with high levels of site- and regioselectivity are particularly valuable.

[*]Corresponding Authors sbuchwal@mit.edu, blp@mit.edu.

Supporting information for this article is given via a link at the end of the document.((Please delete this text if not appropriate))

We recently reported a new approach for mild and site-selective bioconjugation that leverages the reactivity of organometallic reagents (Figure 1).[3] In the presence of preformed complexes of type LPd(Ar)X (L = biarylphosphine, X = Cl, Br, OTf), cysteine thiol groups formed S–C(sp^2) bonds to provide the respective S-arylated peptide or protein. By employing this operationally simple protocol, the arylation of cysteine-containing biomolecules is completed within minutes and in generally high yields.[4] Moreover, palladium complexes bearing a variety of functionalized aryl groups were compatible with these reaction conditions. Nevertheless, the potential chemical and biological lability of the resultant thioether linkage, as well as the scarcity of cysteine residues in peptides, limits the applicability of this protocol. Consequently, we set out to develop an alternative arylation protocol for the formation of stable bonds to amino acid residues that are more abundant in peptides and proteins.

We considered the arylation of the amino group in lysine as an alternative means of generating a stable bioconjugate. Among the various amino acid residues containing potentially reactive nitrogen atoms, we targeted lysine because of its nucleophilicity and unambiguous site of reactivity. Strategies used for the chemical modification of primary amines in polypeptides include conjugate addition,[5] condensation with an activated ester,[6] addition to a (thio)isocyanate or ketene,[7] and Schiff base formation/derivatization.[8] More recently, the reaction of acyl trifluoroborates with prefunctionalized hydroxylamine esters has also been reported.[9] Despite the utility of these protocols, important limitations still exist, including inadequate chemoselectivity, the requirement for a preactivated nucleophile, or the limited chemical stability of the respective conjugate.

Compared to cysteine bioconjugation, we anticipated that lysine conjugation would present additional challenges as a consequence of the lower nucleophilicity of the amino group and the lower acidity of the palladium-amine complex. To address the latter, we envisioned the use of a weak base to effect the requisite deprotonation. Importantly, this base would need to be relatively mild in order to avoid degradation of the polypeptide. In addition, a judiciously chosen ligand would be required to facilitate the desired C-N reductive elimination in preference to other palladium-mediated bond-forming processes. Herein we report the development of conditions to address these challenges, resulting in a general and selective protocol for the conjugation of aryl groups to peptidic lysine residues.

In initial optimization studies, a model peptide containing a lysine residue was exposed to preformed complexes of type LPd(Ar)Br (Ar = 4-anisyl) supported by a series of biarylphosphine ligands in the presence of sodium phenoxide as the base (Table 1, entries 1–6).[10] Sodium phenoxide was selected due to its widespread availability and relatively low basicity (pK_a[BH] = 10).[11] Notably, stability studies indicated that greater than 95% of the peptide substrate remained intact in the presence of sodium phenoxide (see the Supporting Information). It was discovered that the complex supported by *t*-BuBrettPhos exhibited the most pronounced reactivity at room temperature (Entry 5).[12] Unfortunately, a palladium complex bearing an electron deficient aryl group (4-CO_2Me-Ph) afforded the arylation product in low yield due to competitive phenol and aryl ether formation under these conditions (Entry 7).[13] Improved results were obtained when *t*-BuBrettPhos was replaced with the less bulky BrettPhos, presumably due to the reduced propensity of the latter to

facilitate C–O bond forming reductive elimination (Entry 8). Analysis of the reaction mixture by LC/MS indicated that partial arylation of the supporting ligand occurred, due to a previously observed arylative rearrangement of the palladium complex.[14] Thus, more than one equivalent of the palladium complex was required for full conversion of the peptide to the arylated product.[15]

Next, we investigated the selectivity of this methodology by utilizing peptide substrates containing a lysine and another nucleophilic residue (Table 2). Throughout this analysis, the position of modification was unambiguously determined by tandem MS/MS analysis. Although the presence of a cysteine residue was not tolerated due to competitive base-mediated dehydroalanine formation, the current protocol was shown to be completely selective towards the modification of a lysine residue in the presence of serine, tyrosine, methionine, histidine, or tryptophan residues (Entries 1–5). Amino acid residues containing an amide (Entry 6) or a guanidine (Entry 7)[17] side chain could be used in this protocol, although diarylation was also observed. Similarly, the presence of a primary amine at the N-terminus[18] and an amide at the C-terminus gave rise to the corresponding diarylation product (Entries 8 and 9).[19] However, these side reactions could be completely suppressed by employing the Pd complex as the limiting reagent (Entries 6–9).

To demonstrate the utility of this developed protocol, we investigated the arylation of a complex bioactive peptide using LPd(Ar)X complexes with a variety of aryl groups (Scheme 1). We focused our attention on a tumor-suppressing peptide that targets a p53-MDM2 interaction.[20,21] Using this method, an aryl group derived from the corresponding chloride, bromide, and triflate could be coupled to the respective peptide in comparable yields (**1–3**). Complex functional molecules, such as natural product derivatives (**4**[22] and **5**), conjugation/affinity tags (**6** and **7**), chromophores (**8** and **9**), and complex drug molecules containing a chlorine atom (**10**–**15**), were successfully appended with high efficiency. Importantly, this protocol could be conducted on a larger scale without diminishing yield (**8**), and greater than 95% of the Pd-containing species could be conveniently removed from the sample by HPLC purification (153 ppm). The triflate derived from fluorescein (**9**) underwent undesired coupling with sodium phenoxide in addition to the desired coupling with the peptide. In accordance with the results in Table 1, this side reaction could be suppressed by the use of BrettPhos in place of *t*-BuBrettPhos.

A variety of reactive functional groups were tolerated and demonstrated the robustness of this method. These functional groups include arylamines (**10**, **11**, and **15**), alkylamines (**11**, **12**, and **13**) an amidine, (**12**), a ketone (**13**), a carbamate (**14**), a carboxylic acid (**15**), and a diverse array of heterocycles (**10**, **11**, **12**, **14**, and **15**). For a number of cases (entries **1**, **2**, **3**, **5**, and **13**), we isolated the respective LPd(Ar)X complexes as bench stable reagents for peptide arylation.[3] In other cases, the palladium complexes were generated *in situ* from the aryl (pseudo)halide and cyclooctadiene-ligated palladium(0) complex **A** and used without isolation.[23] We anticipate that this *in situ* protocol would find utility in cases where the isolation of the metal complex is non-trivial or rapid diversification of a peptide target is desired. In all cases, the desired monoarylation product was observed as the major product, although in some cases, diarylation products (**1**, **2**, **3**, **6**, and **10**) or regioisomeric products (**5** and **9**) were detected as minor by-products.[24] Notably, Pd triflate complexes had a higher

propensity to undergo these undesired pathways, presumably due to the cationic nature of the metal center.

Peptide stapling has proven to be an invaluable tool in increasing the stability, cell-permeability, and enhancing the α-helicity of peptides.[25] While numerous chemical approaches have been reported, only a limited number of strategies can utilize native amino acid residues as a handle for macrocyclization of both synthetic and recombinantly expressed peptides.[26] Furthermore, these approaches often result in the formation of amides or thioethers that can be proteolytically or oxidatively unstable, respectively. Recently, we have reported a lysine stapling strategy with highly electron deficient arenes that proceeds via S_NAr to form chemically and biologically stable constructs that addresses the aforementioned concerns.[27]

In the presence of a stapling reagent derived from 1,2-bis(4-bromophenoxy)ethane, p53 peptide with an additional lysine residue underwent facile macrocyclization (Scheme 2). The protocol provided access to both [i, i+4] and [i, i+7] stapled products with comparable efficiencies. In addition to the unreacted starting material, side products derived from arylation with a mono-organometallic species account for the rest of the mass balance.[28] It is expected that the organometallic reagent-based approach will allow for the straightforward modulation of the length and identity of the linkers, a feature that is not feasible with a S_NAr reaction-based strategy.

In conclusion, we have discovered a general lysine arylation method based on the use of preformed or *in situ* generated LPd(Ar)X complexes. The reaction allows for the formation of N–aryl conjugates, which are more stable than the corresponding S–aryl conjugates.[27] Furthermore, this protocol operates under mild conditions and is selective over most other nucleophilic amino acid residues. The success of this method stems from the use of the biarylphosphine ligands BrettPhos and *t*-BuBrettPhos in conjunction with the mildly basic sodium phenoxide. We have used this strategy to functionalize complex peptide substrates with a variety of biologically important small molecules and peptide macrocyclization.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors would like to thank Merck & Co. and the National Institutes of Health (R01GM110535 B.L.P.) for financial support of this project. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. We also thank Dr. Shane W. Krska and Dr. Craig A. Parish (Merck & Co.) for helpful suggestions and encouragement. Insightful discussion as well as technical assistance from Dr. Fayçal Touti, Mr. Chi Zhang, and Mr. Peng Dai (MIT) are gratefully acknowledged. We also thank Dr. Yiming Wang (MIT), Dr. Michael Pirnot (MIT), and Dr. Nicholas White (MIT) for the assistance with the preparation of the manuscript.

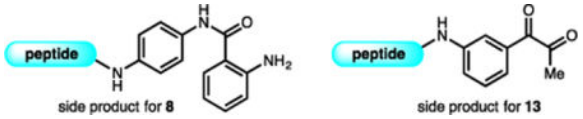
References

1. for review:a) Hu Q, Berti F, Adamo R. Chem Soc Rev. 2016; 45:1691–1719. [PubMed: 26796469] b) Boutureira O, Bernardes GJL. Chem Rev. 2015; 115:2174–2195. [PubMed: 25700113] c) Lau YH, de Andrade P, Wu Y, Spring DR. Chem Soc Rev. 2015; 44:91–102. [PubMed: 25199043]

2. Baslé E, Joubert N, Pucheault M. Chem Biol. 2010; 17:213–227. [PubMed: 20338513]
3. Vinogradova EV, Zhang C, Spokoyny AM, Pentelute BL, Buchwald SL. Nature. 2015; 526:687–691. [PubMed: 26511579]
4. For another application of this approach for a challenging cross-couplingLee HG, Milner PJ, Placzek MS, Buchwald SL, Hooker JM. J Am Chem Soc. 2015; 137:648–651. [PubMed: 25565277]
5. a) Yuan Y, Zhang J, Cao Q, An L, Liang G. Anal Chem. 2015; 87:6180–6185. [PubMed: 25986852] b) Schmitthenner, HF., Beach, S., Weidman, C., Barrett, T. PCT Int Appl. WO 2015017815. 2015.
6. a) Chen X, Muthoosamy K, Pfisterer A, Neumann B, Weil T. Bioconjugate Chem. 2012; 23:500–508.b) Robinson MA, Charlton ST, Garnier P, Wang XT, Davis SS, Perkins AC, Frier M, Duncan R, Savage TJ, Wyatt DA, Watson SA, Davis NG. Proc Natl Acad Sci U S A. 2004; 101:14527–14532. [PubMed: 15448212] c) Diethelm S, Schafroth MA, Carreira EM. Org Lett. 2014; 16:3908. [PubMed: 25019948] D) Corbani M, Trueba M, Stoev S, Murat B, Mion J, Boulay V, Guillon G, Manning M. J Med Chem. 2011; 54:2864–2877. [PubMed: 21428295]
7. a) Micewicz ED, Ratikan JA, Waring AJ, Whitelegge JP, McBride WH, Ruchala P. Bioorg Med Chem Lett. 2015; 25:4419–4427. [PubMed: 26384289] b) Thomas B, Fiore M, Daskhan GC, Spinelli N, Renaudet O. Chem Commun. 2015; 51:5436–5439.c) Chan AOY, Ho CM, Chong HC, Leung YC, Huang JS, Wong MK, Che CM. J Am Chem Soc. 2012; 134:2589–2598. [PubMed: 22288779]
8. a) Tanaka K, Fukase K, Katsumura S. Synlett. 2011:2115–2139.b) McFarland JM, Francis MB. J Am Chem Soc. 2005; 127:13490–13491. [PubMed: 16190700] c) Cal PMSD, Vicente JB, Pires E, Coelho AV, Veiros LF, Cordeiro C, Gois PMP. J Am Chem Soc. 2012; 134:10299–10305. d). [PubMed: 22642715]
9. Noda H, Erős G, Bode JW. J Am Chem Soc. 2014; 136:5611–5614. [PubMed: 24684235]
10. For selected examples using metal phenoxide as a base for C–N cross coupling:a) Alcazar-Roman LM, Hartwig JF. J Am Chem Soc. 2001; 123:12905–12906. [PubMed: 11749551] b) Shekhar S, Hartwig JF. Organometallics. 2007; 26:340–351.c) Schulte JP, Tweedie SR. Synlett. 2007:2331–2336.d) Hoi KH, Organ MG. Chem Eur J. 2012; 18:804–807. [PubMed: 22180129] e) Brusoe AT, Hartwig JF. J Am Chem Soc. 2015; 137:8460–8468. [PubMed: 26065341]
11. Other bases such as metal alkoxides or phosphazenes also promoted the reaction.
12. Fors BP, Dooleweerdt K, Zeng Q, Buchwald SL. Tetrahedron. 2009; 65:6576–6583. [PubMed: 20740063]
13. The undesired cross-coupled products were confirmed by GC/MS analyses.
14. a) Maimone TJ, Milner PJ, Kinzel T, Zhang Y, Takase MK, Buchwald SL. J Am Chem Soc. 2011; 133:18106–18109. [PubMed: 21999801] b) Milner PJ, Maimone TJ, Su M, Chen J, Müller P, Buchwald SL. J Am Chem Soc. 2012; 134:19922–19934. [PubMed: 23153301]
15. To avoid *in situ* modification of the ligand, premodified ligands, such as HGPhos or AlPhos, were also tested. Unfortunately, these alternatives did not promote the reaction at all, presumably due to the highly hindered nature of these ligands.
16. To confirm the validity of this method, the product was isolated in one case (Scheme 1, 8) and a correlation curve between the TIC integration area and the concentration of the sample was generated. The integration area of the sample was fitted to show that the yield acquired from the indicated method matches within less than 3% error. In cases wherein the baseline separation is inefficient, extracted ion count (EIC) analysis was used. For selected cases both TIC and EIC analyses were conducted to demonstrate that the results are comparable (<3% error). See supporting information for details.
17. While a number of Cu-catalyzed reactions are reported for the arylation of a guanidine moiety from aryl halides, the respective Pd-catalyzed process has not been reported except for the arylation of 2-aminopyrimidine. See ref 10c.
18. a) King SM, Buchwald SL. Org Lett. 2016; 18:4128–4131. [PubMed: 27498618] b) Hopkins BA, Smith GF. c) Sciammetta N. Org Lett. 2016; 18:4072–4075. [PubMed: 27508926]
19. In cases where the double arylation was observed, we attempted to detect the monoarylation product at the undesired position by extracting the target mass from the total ion current chromatogram. We only observed the undesired regioisomeric monoarylated product for entry 9

where less than 1% of the C-terminus monoarylation product was detected. See supporting information for details.

20. Kussie PH, Gorina S, Marechal V, Elenbaas B, Moreau J, Levine AJ, Pavletich NP. Science. 1996; 274:948–953. [PubMed: 8875929]
21. To secure complete buffering, 20 equiv. of the base was used. As seen previously, no significant degradation of the peptide substrate was observed under these conditions. See supporting information.
22. Palitoylation is known to promote the cell penetration of pepides. See ref 9 and references therein.
23. Lee HG, Milner PJ, Colvin MT, Andreas L, Buchwald SL. Inorg Chim Acta. 2014; 422:188–192.
24. With the exception of these examples, the only peptidic species observed were unreacted starting material and the desired products. In addition, side reactions were observed on the organic molecules during the conjugation with 8 (3%) and 13 (6%) to a small extent. See supporting information for details.



25. For a recent review see reference 1c.
26. a) Shepherd NE, Hoang HN, Abbenante G, Fairlie DP. J Am Chem Soc. 2005; 127:2974–2983. [PubMed: 15740134] b) Phelan JC, Skelton NJ, Braisted AC, McDowell RS. J Am Chem Soc. 1997; 119:455–460.c) Fujimoto K, Kajino M, Inouye M. Chem Eur J. 2008; 14:857–863. [PubMed: 17969217] d) Jo H, Meinhardt N, Wu Y, Kulkarni S, Hu X, Low KE, Davies PL, DeGrado WF, Greenbaum DC. J Am Chem Soc. 2012; 134:17704–17713. [PubMed: 22998171] e) Spokoyny AM, Zou Y, Ling JJ, Yu H, Lin YS, Pentelute BL. J Am Chem Soc. 2013; 135:5946–5949. [PubMed: 23560559] f) Wang Y, Chou DH. Angew Chem. 2015; 127:11081–11084.Angew Chem Int Ed. 2015; 54:10931–10934. Also, see Ref 3.
27. Lautrette G, Touti F, Lee HG, Dai P, Pentelute BL. J Am Chem Soc. 2016; 138:8340–8343. [PubMed: 27332147]
28. See supporting information for details.

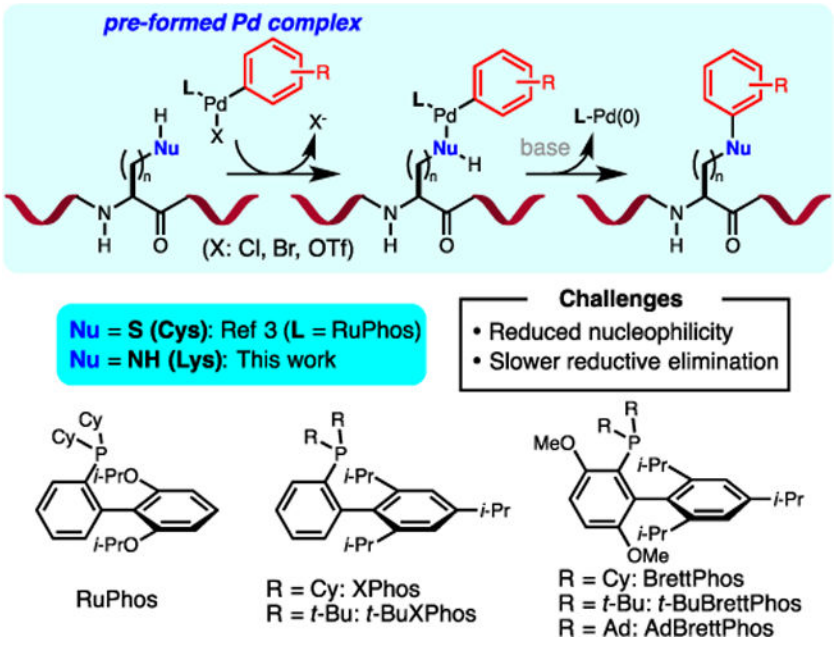
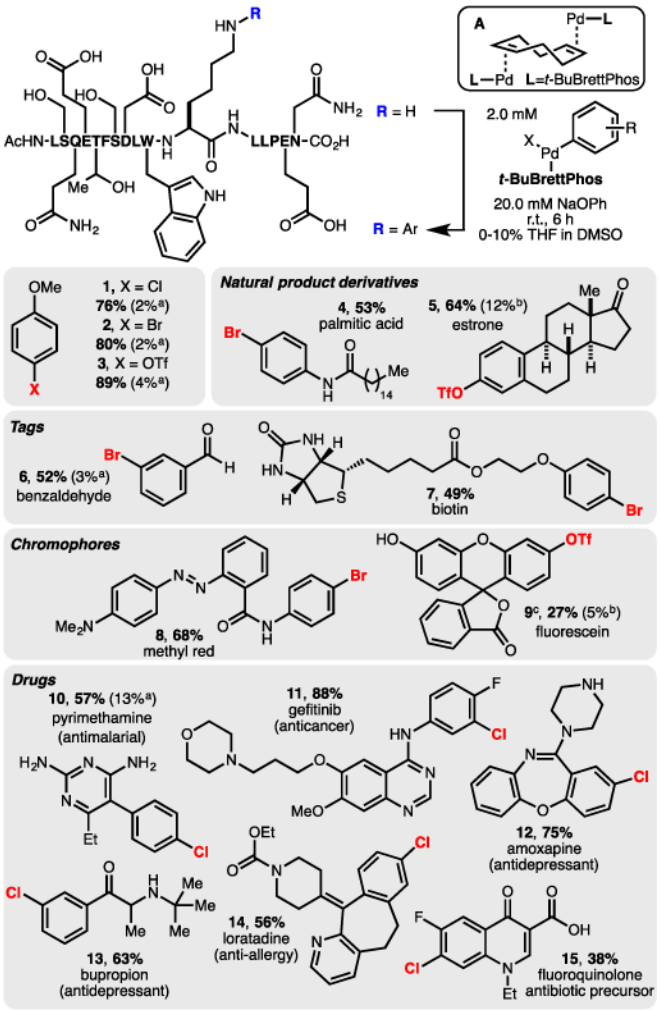


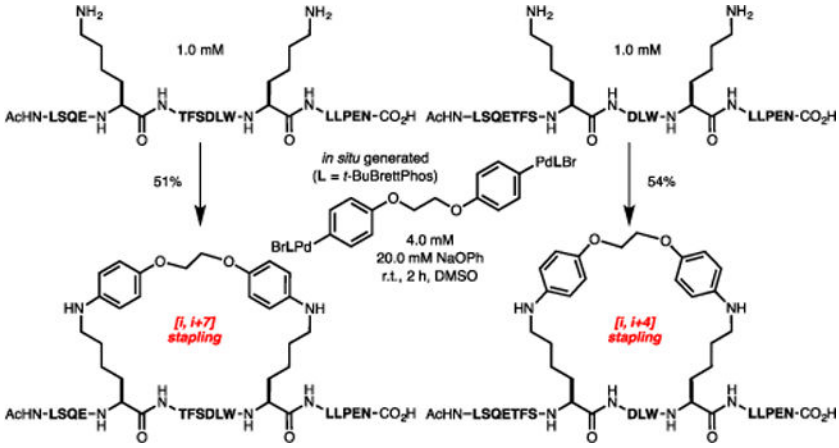
Figure 1.

Bioconjugation strategy using organometallic reagents and the structure of biarylphosphine ligands.



Scheme 1.

Lysine arylation shows broad substrate scope including introduction of tags, chromophores, and drugs. [a] Diarylation product; [b] regioisomeric product; [c] [BrettPhos·Pd(COD)][23] was used in place of **A**.



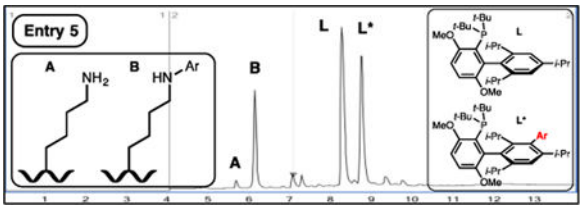
Scheme 2.

Polymetalated Pd complex enables efficient peptide macrocyclization.

Table 1

Conditions for efficient arylation of lysine in unprotected peptides.

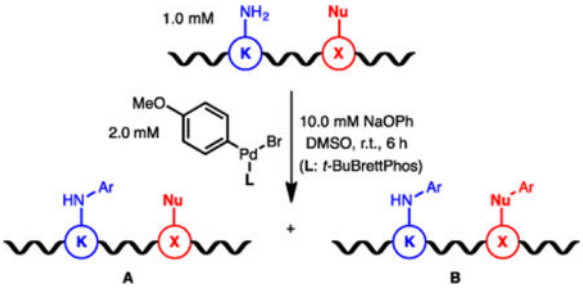
Entry	L	R	Yield (%)[a]
1	RuPhos	4-OMe	0
2	XPhos	4-OMe	0
3	*t*-BuXPhos	4-OMe	93
4	BrettPhos	4-OMe	1
5	*t*-BuBrettPhos	4-OMe	94
6	AdBrettPhos	4-OMe	79
7	*t*-BuBrettPhos	4-CO_2Me	18
8	BrettPhos	4-CO_2Me	71



[a]Yields were calculated by integration of total ion count (TIC) chromatogram.[16]

Table 2

Systematic investigation of lysine arylation in the presence of other nucleophilic side chains.



Entry	Peptide	A / B (%)
1	**AcNH-FLG KGVG SAF-CO_2H**	90/0
2	**AcNH-FLG KGVG *Y* AF-CO_2H**	78 / 0
3	**AcNH-FLG KGVG MAF-CO_2H**	92 / 0
4	**AcNH-FLG KGVG HAF-CO_2H**	62 / 0
5	**AcNH-FLG KGVG WAF-CO_2H**	80 / 0
6	**AcNH-FLG KGVG NAF-CO_2H**	88 / 5 (20/0)[a]
7	**AcNH-FLG KGVG RAF-CO_2H**	37 / 39 / 13[b] (20 / 0)[a]
8	**H_2N-FLAG KGAFG-CO_2H**	86 / 8 (24 / 0)[a]
9	**AcNH-FLAG KGAFG-$CONH_2$**	64 / 29 (18 / 0)[a]

[a] Reaction yield with 0.20 mM of Pd complexes

[b] Triple arylation