

Expanded Genetic Code

A Genetically Encoded Boronate-Containing Amino Acid**

Eric Brustad, Mark L. Bushey, Jae Wook Lee, Dan Groff, Wenshe Liu, and Peter G. Schultz*

Organoboronates have attracted considerable interest as synthetic intermediates for Suzuki cross-coupling reactions,^[1,2] copper catalyzed heteroatom alkylation reactions,^[3] asymmetric reductions,^[4] Diels–Alder reactions,^[5] and a variety of other transformations. Boronic acids are also known to form reversible covalent complexes with diols,^[6] amino alcohols,^[7] amino acids,^[8,9] alkoxides,^[10] and hydroxamic acids.^[11] This latter property has been exploited in the synthesis of ligands for the selective recognition of sugars^[12–14] and for the development of potent serine protease inhibitors.^[15–18] In addition, boronates are finding utility as boron neutron capture agents to kill tumor cells.^[19] Despite their novel chemical properties, boronic acids are not known to occur naturally in polypeptides, either as posttranslational modifications or as cofactors. Therefore, the ability to genetically encode this functional group would provide useful tools for biomolecular recognition and for the selective chemical modification of proteins. Herein, we describe a general methodology for the site-specific incorporation of *p*-boronophenylalanine (**1**) into proteins directly in *E. coli* (Figure 1). We also demonstrate the utility of this system for boronate-mediated protein modification as well as the single step scarless purification of proteins.

We have previously described methodology for the addition of nonnatural amino acids to the genetic code of *E. coli*,^[20] yeast,^[21] and mammalian cells.^[22] This method is based on the generation of an orthogonal tRNA/aminoacyl-tRNA synthetase (aaRS) pair that allows the site-selective incorporation of novel amino acids into proteins in response to unique nonsense and frameshift codons. To selectively incorporate *p*-boronophenylalanine into proteins in *E. coli*, we have used a *Methanococcus jannaschii* (*Mj*) derived amber suppressor tyrosyl tRNA (*MjtRNA*^{Tyr}_{CUA})/tyrosyl-tRNA synthetase (*MjTyrRS*) pair that has been previously shown to function efficiently in *E. coli*, but does not cross-react with any of the endogenous tRNAs and aminoacyl-tRNA synthe-

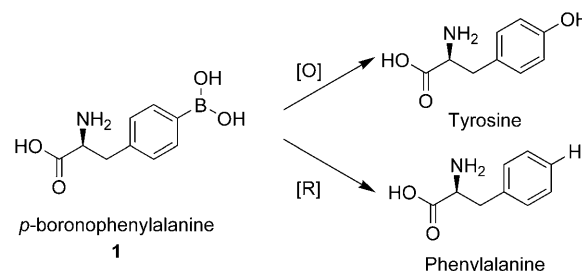


Figure 1. Structure of *p*-boronophenylalanine (**1**). Oxidation [O] of the amino acid leads to tyrosine while reduction [R] produces phenylalanine.

tases.^[23,24] To alter the specificity of the *MjTyrRS* synthetase to selectively recognize **1**, two active site NNK saturation mutagenesis libraries of > 10⁹ diversity were each subjected to iterative rounds of positive and negative selections. The design of these libraries and the selection methodology has been described elsewhere.^[20,25]

After selection (three positive and two negative rounds) numerous clones were obtained that permitted cells harboring a chloramphenicol acetyl transferase selection plasmid to survive on 120 μg mL⁻¹ of chloramphenicol only in the presence of 1 mM **1**. Absence of the amino acid precluded cell growth in this system. As shown in Table S1 in the Supporting Information, of the seven clones sequenced, three unique sequences were obtained displaying considerable consensus. Positions 65 and 162 showed complete convergence to alanine and glutamate, respectively, while position 155 maintained the wildtype glutamine for all clones. Positions 32, 70, and 158 were enriched for Ser/Gly, His/Met, or Ser/Ala, respectively. Analysis of the wildtype *MjTyrRS* crystal structure complexed with tyrosine provides some rationale for the possible roles of these mutations.^[26] Tyr32 and Asp158 make critical hydrogen bonds to the phenolic oxygen of the bound tyrosine. Replacement of these amino acids with a smaller serine residue removes the determinants necessary for binding to tyrosine, while maintaining hydrogen bonding functionality that may interact with the boronate group. Mutation of Leu162 to Glu adds an additional hydrogen bonding residue that could interact with the boronate functionality. The most common sequence (1BG11, designated B(OH)₂PheRS) was used for all subsequent experiments.

To confirm the incorporation of **1** into proteins, an amber mutation (TAG) was substituted for Lys7 of a C-terminal 6XHis tagged variant of the Z-domain of staphylococcal protein A (Figure S1 in the Supporting Information). Expression experiments were carried out in DH10B *E. coli* cells harboring plasmids containing the amber Z-domain gene as well as the *MjtRNA*^{Tyr}_{CUA} and the evolved B(OH)₂PheRS. Protein was expressed in 2xYT media in the presence or

[*] E. Brustad, Dr. M. L. Bushey, Dr. J. W. Lee, D. Groff, Prof. W. Liu,^[†] Prof. Dr. P. G. Schultz
Department of Chemistry and The Skaggs Institute for Chemical Biology, The Scripps Research Institute
10550 N. Torrey Pines Rd. La Jolla, CA 92037 (USA)
Fax: (+1) 858-784-9440
E-mail: schultz@scripps.edu

[†] Current address: Department of Chemistry, Texas A&M University, College Station, TX 77842 (USA)

[**] We thank Dr. Vaughn Smider, Chang Liu, Jeremy Mills, and Emily Remba for assistance and helpful discussion. This work is supported by NIH grant R01 GM062159, US Department of Energy, Division of Materials Sciences, under Award No. DE-FG03-00ER46051, and the Skaggs Institute for Chemical Biology.

Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/anie.200803240>.

absence of 1 mM amino acid **1** followed by nickel affinity purification. Analysis by SDS-PAGE and subsequent coomassie staining showed that protein was only produced in the presence of the boronate amino acid (Figure S2a). Protein yields were typically around 15 mg L^{-1} of expressed cell culture. The expected mass of the boronate-containing Z-domain (Z-domain-K7(**1**)) is 7824 ($M+H$ protein mass minus the N-terminal methionine which is cleaved post-translationally); however, electrospray ionization (ESI) mass spectrometry of the protein showed two peaks corresponding to the loss of one or two water molecules (7807 and 7788, respectively, Figure S2b). Analysis of the Z-domain crystal structure shows that position 7 lies in close proximity to Thr2 and Ser3 on a flexible N-terminal loop.^[27] These masses are consistent with the formation of boronic esters with the hydroxy groups of these residues. To further confirm the incorporation of **1**, we substituted Ala82 with **1** in an engineered cysteine-free T4 lysozyme (T4L-A82(**1**)). This location was chosen because it is remote to any free hydroxy groups in the protein structure. ESI of this protein showed the expected mass corresponding to the intact boronate amino acid (calcd. 18721, obsd. 18722; Figure S3a) with no dehydration products observed.

p-Boronophenylalanine has been used in solid phase peptide synthesis as a precursor to tyrosine and phenylalanine through the oxidation or reduction of the boronate functionality, respectively (Figure 2a).^[28] This reactivity allowed the incorporation of the amino acid to be confirmed by chemical methods. Oxidation of Z-domain-K7(**1**) for two hours with excess hydrogen peroxide led to the disappearance of both dehydration peaks in the mass spectrum described above to give a single ESI peak corresponding to the expected mass of tyrosine at position 7 (calcd. 7798 [$M+H$], obsd. 7798; Figure S2c). Similarly, overnight reduction of Z-domain-K7(**1**) with excess silver diammonium nitrate yielded the expected mass of the phenylalanine product (calcd. 7782 [$M+H$], obsd. 7782; Figure S2d). Using hydrogen peroxide to oxidize the boronate amino acid is not selective in the presence of other easily oxidized residues such as methionine. Indeed, attempts to oxidize T4L-A82(**1**) (calcd. 18693 [$M+H$]) gave a mass approximately 80 daltons larger than expected (obsd. 18772, Figure S3b) corresponding to oxidation of each of the five methionine residues to methionine sulfoxide. Selective oxidation of the boronate moiety was achieved by using one equivalent of potassium peroxymonosulfate (oxone)^[29] which afforded the expected tyrosine product with no observed oxidation of the endogenous methionines (Figure S3c).

We next examined the utility of this amino acid as an orthogonal handle to selectively modify proteins. Boronic acids are known to form strong reversible interactions with polyhydroxylated compounds in aqueous solution at physiological pH.^[6,30] The interaction of boronates with sorbitol and glucamine (**2**; Figure 2a) for example are of very high affinity and provide a useful method to selectively label proteins under conditions that are relatively benign to biological systems. To examine the ability of polyhydroxylated molecules to bind boronate-containing proteins, we first synthesized a glucamine-functionalized fluorescent reporter. NHS-

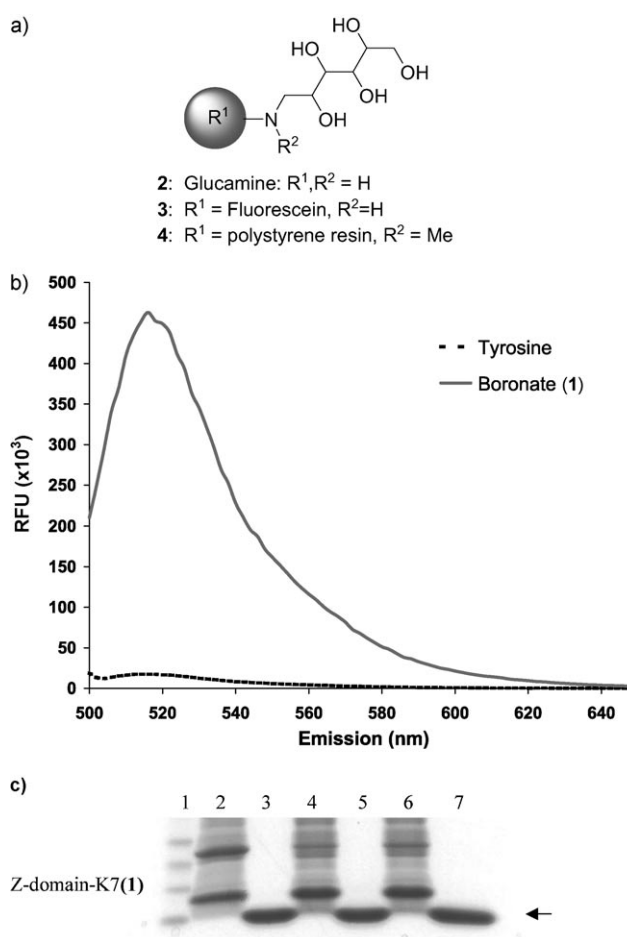


Figure 2. a) Structures of glucamine (**2**), fluorescein–glucamine probe (**3**), and immobilized glucamine affinity purification resin (**4**); XUS43594.00 Dow Chemicals). b) Selective fluorescent labeling of Z-domain-K7(**1**) (solid line) over Z-domain-K7Y (dotted line) with fluorescein–glucamine dye (**3**). c) Affinity purification of 6XHis tagged Z-domain-K7(**1**) using nickel-NTA resin or boronate affinity resin (**4**). Lane 1: protein markers; lane 2: flow through of Ni-NTA purified protein; lane 3: elution of Ni-NTA purified protein; lane 4: flow through of boronate affinity resin; lane 5: 200 mM hydrogen peroxide elution of boronate affinity resin; lane 6: flow through of boronate affinity resin; lane 7: 1 M sorbitol elution of boronate affinity resin. Z-domain is indicated by the arrow.

fluorescein was conjugated to glucamine in DMF quantitatively in one step to afford dye **3** (Figure 2a) which was then incubated with T4 lysozyme containing the boronate amino acid at position 82 (or tyrosine at the same location as a negative control) in 50 mM CHES buffer, pH 8.5. After 1 h at room temperature, excess dye was removed by exhaustive washing in a 10 kDa molecular weight cutoff Amicon centrifugal filter device and fluorescence was imaged. As shown in Figure 2b, only protein containing the boronate amino acid showed fluorescence after labeling. The yield of this protein labeling was determined to be approximately 61% based on the absorbance of the dye at 494 nm ($\epsilon_{494} \approx 68000 \text{ cm}^{-1} \text{ M}^{-1}$) and of the protein at 280 nm ($\epsilon_{280} \approx 24750 \text{ cm}^{-1} \text{ M}^{-1}$). Yields were lower than expected due to slight air oxidation of the boronic acid under these conditions.

We next asked whether this chemistry could be used as a general strategy for the affinity purification of protein from cell lysates. XUS43594.00 resin (**4**) from Dow Chemicals is an *N*-methylglucamine conjugated resin (Figure 2a) designed for the removal of free borate from water supplies. To test whether a boronate amino acid would be sufficient for the affinity purification of a protein, 6XHis tagged Z-domain with either **1** or tyrosine at position 7 was expressed. After overnight expression in 2xYT media, clarified cell lysates in 50 mM CHES buffer, pH 8.5, were incubated with *N*-methylglucamine resin (**4**) for 4 h at room temperature before being loaded onto a disposable polypropylene column and washed exhaustively with a high salt (1M NaCl) wash buffer to remove any proteins that may bind to the resin through nonspecific interactions. The resin was then split and protein was eluted by the addition of excess sorbitol (1M) or by oxidation of the boronate to tyrosine using excess hydrogen peroxide. As shown in Figure 2c (lanes 5 and 7), Z-domain protein was isolated in high purity using either elution condition. Purification yields were comparable to that using Ni-NTA/His Tag purification (Figure 2c, lane 3). Of 1 mg of the Z-domain-K7(**1**) loaded onto 3 mL of the *N*-methylglucamine resin, 0.96 mg of protein was recovered in the sorbitol elution fraction corresponding to >95% protein recovery. Mass spectral analysis of the purified protein showed that protein eluted using high concentrations of competing sorbitol contained the unmodified boronate amino acid whereas hydrogen peroxide elution yielded exclusively protein containing tyrosine at that position (Figures S4a and S4b). No protein was found in either *N*-methylglucamine elution fraction when tyrosine was used in place of the boronate amino acid (Figure S5, lanes 5 and 7). It is significant to note that due to the ability of **1** to be oxidized or reduced to tyrosine or phenylalanine, respectively, this methodology allows the purification of native protein sequences. In one step, a protein can be isolated based on the unique chemistry of the nonnatural amino acid alone without the need for commonly used protein affinity tags such as 6XHis or fusion proteins. Selective oxidation can be subsequently used to generate a tyrosine residue, yielding a native protein sequence with no modification remaining from the purification process.

Finally, we explored the palladium mediated Suzuki coupling between boronate containing proteins and a fluorescent aryl iodide reporter based on the bodipy scaffold (**5**) (Figure 3a). The synthesis of **5** was carried out by using a previously reported method with minor modifications.^[31] Suzuki couplings with proteins have already been reported with a synthetic W-domain peptide containing a *p*-iodophenylalanine nonnatural amino acid and a fluorescent boronic acid compound in the presence of a water soluble palladium catalyst (Na₂PdCl₄).^[32] In addition Kodama et al. have demonstrated the first example of a palladium catalyzed organometallic reaction on a biosynthetically incorporated *p*-iodophenylalanine amino acid.^[33] To achieve Suzuki coupling of **5** with T4L-A82(**1**), a variety of basic buffered solutions and palladium catalysts were screened. In the end, coupling was achieved in moderate (ca. 30%) yield using the Pd⁰ dibenzylidene acetone (Pd-DBA) catalyst in 20 mM EPPS,

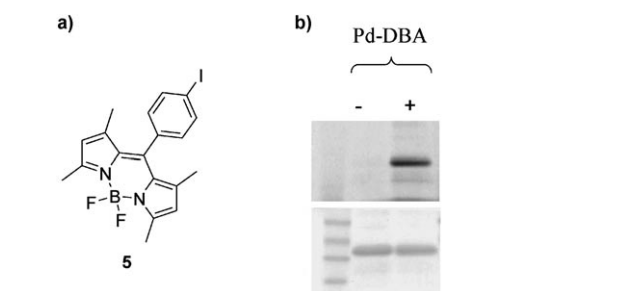


Figure 3. a) Structure of iodinated bodipy reporter molecule. b) Suzuki coupling of 50 μM T4L-A82(**1**) to 1 mM reporter molecule **5** in 20 mM EPPS, pH 8.5, 70°C, for 12 h in the absence (–) or presence (+) of 1 mM Pd-DBA. Top: bodipy fluorescence; bottom: coomassie staining of the same gel.

pH 8.5, at 70°C (Figure 3b); no coupling between the boronate protein and bodipy **5** was observed in the absence of palladium catalyst. The high temperature needed to attain significant cross-coupling may preclude this as a general strategy for the site selective labeling of proteins as many proteins are not stable at these conditions. The development of more efficient water soluble palladium catalysts may improve this reactivity significantly.^[34,35]

In summary, we have demonstrated the successful addition of the boronate functionality to the genetic code of *E. coli* in high yield and efficiency in response to the amber stop codon (TAG). The biosynthetic incorporation of this amino acid into proteins allows for selective chemistry on the protein surface including oxidation, reduction, Suzuki coupling reactions, as well as the formation of covalent boronic esters with polyhydroxylated compounds. We have also shown that this amino acid can be used in concert with a polyhydroxylated solid support to purify native protein sequences in a one step scarless affinity purification procedure. Furthermore, the ability of boronic acids to bind diols and reactive serine residues suggests that this technology could lead to the development of boronate-containing antibodies that specifically recognize and covalently bind various glycoproteins or proteases. It should also be possible to form intramolecular serine–boronate crosslinks in proteins to enhance stability. And finally, the unique chemistry of this functionality may allow for the *in vivo* labeling of boronate containing proteins with polyhydroxylated reporter molecules.

Experimental Section

Materials and Methods can be found in the Supporting Information.

Received: July 4, 2008

Published online: September 24, 2008

Keywords: affinity purification · boronic acids · nonnatural amino acids · protein modification

[1] N. Miyaura, A. Suzuki, *Chem. Rev.* **1995**, *95*, 2457.

[2] A. Suzuki, *J. Organomet. Chem.* **1999**, *576*, 147.

[3] D. M. T. Chan, K. L. Monaco, R. H. Li, D. Bonne, C. G. Clark, P. Y. S. Lam, *Tetrahedron Lett.* **2003**, *44*, 3863.

- [4] S. W. Huang, B. Wang, Z. X. Shan, D. J. Zhao, *Synth. Commun.* **2000**, *30*, 2423.
- [5] K. Ishihara, H. Yamamoto, *Eur. J. Org. Chem.* **1999**, 527.
- [6] J. P. Lorand, J. O. Edwards, *J. Org. Chem.* **1959**, *24*, 769.
- [7] G. Springsteen, C. E. Ballard, S. Gao, W. Wang, B. H. Wang, *Bioorg. Chem.* **2001**, *29*, 259.
- [8] L. K. Mohler, A. W. Czarnik, *J. Am. Chem. Soc.* **1994**, *116*, 2233.
- [9] L. K. Mohler, A. W. Czarnik, *J. Am. Chem. Soc.* **1993**, *115*, 7037.
- [10] A. N. Cammidge, K. V. L. Crépy, *Tetrahedron* **2004**, *60*, 4377.
- [11] L. Lamandé, D. Boyer, A. Munoz, *J. Organomet. Chem.* **1980**, 329, 1.
- [12] T. D. James, K. Sandanayake, S. Shinkai, *Angew. Chem.* **1996**, *108*, 2038; *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 1910.
- [13] T. D. James, K. Sandanayake, S. Shinkai, *Nature* **1995**, *374*, 345.
- [14] W. Wang, X. M. Gao, B. H. Wang, *Curr. Org. Chem.* **2002**, *6*, 1285.
- [15] J. Adams, M. Behnke, S. W. Chen, A. A. Cruickshank, L. R. Dick, L. Grenier, J. M. Klunder, Y. T. Ma, L. Plamondon, R. L. Stein, *Bioorg. Med. Chem. Lett.* **1998**, *8*, 333.
- [16] G. S. Weston, J. Blazquez, F. Baquero, B. K. Shoichet, *J. Med. Chem.* **1998**, *41*, 4577.
- [17] W. Q. Yang, X. M. Gao, B. H. Wang, *Med. Res. Rev.* **2003**, *23*, 346.
- [18] D. A. Matthews, R. A. Alden, J. J. Birktoft, S. T. Freer, J. Kraut, *J. Biol. Chem.* **1975**, *250*, 7120.
- [19] Y. Kinashi, S. Masunaga, K. Ono, *Int. J. Radiat. Oncol. Biol. Phys.* **2002**, *54*, 562.
- [20] L. Wang, A. Brock, B. Herberich, P. G. Schultz, *Science* **2001**, *292*, 498.
- [21] J. W. Chin, T. A. Cropp, J. C. Anderson, M. Mukherji, Z. W. Zhang, P. G. Schultz, *Science* **2003**, *301*, 964.
- [22] W. S. Liu, A. Brock, S. Chen, S. B. Chen, P. G. Schultz, *Nat. Methods* **2007**, *4*, 239.
- [23] L. Wang, T. J. Magliery, D. R. Liu, P. G. Schultz, *J. Am. Chem. Soc.* **2000**, *122*, 5010.
- [24] L. Wang, P. G. Schultz, *Chem. Commun.* **2002**, 1.
- [25] J. M. Xie, W. S. Liu, P. G. Schultz, *Angew. Chem.* **2007**, *119*, 9399; *Angew. Chem. Int. Ed.* **2007**, *46*, 9239.
- [26] T. Kobayashi, K. Sakamoto, T. Takimura, R. Sekine, K. Vincent, K. Kamata, S. Nishimura, S. Yokoyama, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 1366.
- [27] D. Y. Zheng, J. M. Aramini, G. T. Montelione, *Protein Sci.* **2004**, *13*, 549.
- [28] D. S. Kemp, D. C. Roberts, *Tetrahedron Lett.* **1975**, *16*, 4629.
- [29] K. S. Webb, D. Levy, *Tetrahedron Lett.* **1995**, *36*, 5117.
- [30] G. Springsteen, B. H. Wang, *Tetrahedron* **2002**, *58*, 5291.
- [31] C. Tahtaoui, C. Thomas, F. Rohmer, P. Klotz, G. Duportail, Y. Mely, D. Bonnet, M. Hibert, *J. Org. Chem.* **2007**, *72*, 269.
- [32] A. Ojida, H. Tsutsumi, N. Kasagi, I. Hamachi, *Tetrahedron Lett.* **2005**, *46*, 3301.
- [33] K. Kodama, S. Fukuzawa, H. Nakayama, T. Kigawa, K. Sakamoto, T. Yabuki, N. Matsuda, M. Shirouzu, K. Takio, K. Tachibana, S. Yokoyama, *ChemBioChem* **2006**, *7*, 134.
- [34] D. Badone, M. Baroni, R. Cardamone, A. Ielmini, U. Guzzi, *J. Org. Chem.* **1997**, *62*, 7170.
- [35] K. H. Shaughnessy, R. S. Booth, *Abstr. Pap. Am. Chem. Soc.* **2004**, 227, U1534.