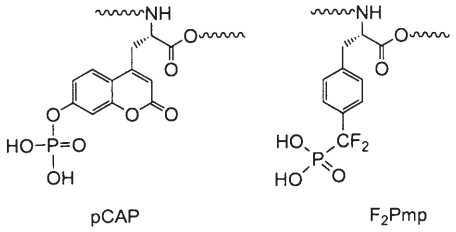
DOI: 10.1002/cbic.200800046

Identifying Selective Protein Tyrosine Phosphatase Substrates and Inhibitors from a Fluorogenic, Combinatorial Peptide Library

Sayantan Mitra[a] and Amy M. Barrios*[a, b]

Protein tyrosine phosphatases (PTPs) are increasingly recognized as enzymes that exhibit exquisite substrate selectivity and play critical, nonredundant roles in cellular signaling.[1–4] In vivo, tyrosine phosphorylation is both dynamic and tightly regulated, controlled by the opposing actions of protein tyrosine kinases (PTKs) and PTPs.[5,6] Although several PTPs have been identified as attractive therapeutic targets in human diseases including autoimmunity, obesity, diabetes, and cancer,[1,7,8] the development of potent, selective PTP inhibitors remains a significant challenge because of the high homology of the PTP catalytic domains and a lack of information about the molecular basis for substrate and inhibitor recognition.[7,9,10]

Strategies for studying PTP selectivity include measuring the small changes in absorbance and fluorescence upon dephosphorylation of phosphotyrosine (pY),[11–13] mass spectrometry techniques (ECLIPSE),[14] radiolabeled phosphosubstrates,[15] and anti-pY antibodies.[16,17] These approaches, although valuable, have limited sensitivity, employ discontinuous assays, and can require specialized equipment or isotopic enrichment. Alternatively, catalytically incompetent enzymes have been used to "trap" substrates from cell lysates[18] or combinatorial libraries of phosphopeptides.[19–21] Libraries of peptides containing difluorophosphonomethylphenylalanine (F_2Pmp; Scheme 1), a



Scheme 1.

nonhydrolyzable analogue of pY, have also been used.[22–24] These efforts generally identify only high affinity substrates and have the limitation that affinity is not always equal to catalytic turnover. Finally, innovative approaches using assays linked to chymotrypsin-mediated cleavage[25–27] or tyrosinase-catalyzed oxidation[28,29] of the dephosphorylated substrate have been used in both discontinuous[25,28,29] and continuous assays[26,27] for PTP substrate selectivity. These assays, even when run continuously, are not direct assays for PTP activity and can require the addition of several different reagents and significant data workup.

In light of the challenges highlighted above, there remains a great need for new tools to facilitate PTP substrate selectivity profiling. An ideal assay would allow direct, continuous measurement of dephosphorylation by catalytically competent PTP domains, provide a simple, highly sensitive readout, and be amenable to high-throughput screening and incorporation into fully diverse combinatorial peptide libraries. We have recently described the use of phosphorylated coumaryl amino acids as pY analogues in peptide-based PTP substrates.[30,31] The increase in fluorescence upon dephosphorylation of the pCAP residue (Scheme 1) can be readily followed in real time and is linear over a large concentration range.[30] Notably, substrates containing the pCAP residue are readily hydrolyzed by a variety of PTPs with kinetic parameters similar or superior to that of analogous pY-containing peptides.[30,31] Herein we describe the development of two pCAP-based, fully diverse, positionally scanned combinatorial peptide substrate libraries and their utility in profiling PTP substrate selectivity.

By scanning one position at a time and presenting an equimolar mixture of all possible amino acids in the other positions of the substrate sequence, a positional scanning approach allows determination of amino acid preferences at each position with no inherent bias due to interactions with neighboring positions.[32,33] The downside to this is that individual peptides based on the resulting library profiles must be synthesized and characterized to validate the profile. As a template for our library, we selected the sequence surrounding Tyr992 of the EGF receptor, DADE-pY-L, a well-established, general PTP substrate.[13] We initially scanned four amino acids N-terminal to the phosphorylated residue using a library of compounds with the sequence XXXX-pCAP-LAA (N-terminal library, Figure 1). Based on preliminary results indicating that Ile was preferred over Ala at the −3 position, we scanned the C-terminal subsite preferences using a library with the sequence DIDE-pCAP-XXXX (C-terminal library, Figure 1). To validate the utility of the combinatorial libraries, we profiled the N- and C-terminal substrate selectivity of TCPTP, a well-characterized member of the PTP family of enzymes.

The N-terminal library profile for TCPTP (Figure 2) shows slight preferences for Phe and Leu at the −4 position and Ala at −1, a marked preference for hydrophobics at −3, and little selectivity at the −2 position. Interestingly, acidic amino acids do not appear to be preferred at any of the N-terminal sub-

[a] Dr. S. Mitra, Prof. Dr. A. M. Barrios
Department of Chemistry, University of Southern California
Los Angeles, CA 90089 (USA)

[b] Prof. Dr. A. M. Barrios
Present Address:
Department of Medicinal Chemistry, University of Utah
Salt Lake City, UT 84112 (USA)
Fax: (+1) 801-581-7087
E-mail: Amy.Barrios@utah.edu

Supporting information for this article is available on the WWW under http://www.chembiochem.org or from the author.

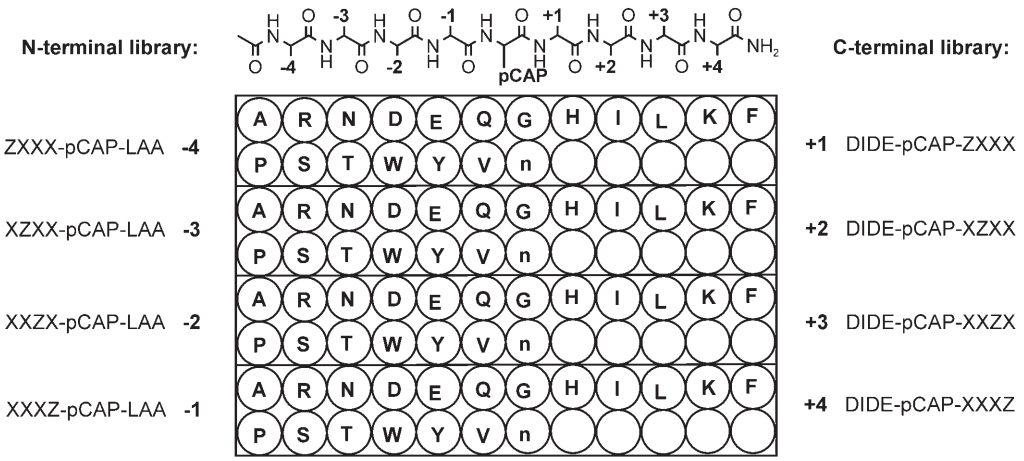


Figure 1. The N-terminal and C-terminal libraries in 96-well plate format. Z denotes the positionally addressed amino acid. X denotes an equimolar mixture of all amino acids (the 20 naturally occurring amino acid excluding Cys and Met, but including the Met isostere, norleucine, n). Residues N-terminal to the pCAP moiety are denoted as −1, −2, −3, etc., whereas residues C-terminal to pCAP are +1, +2, +3, etc.

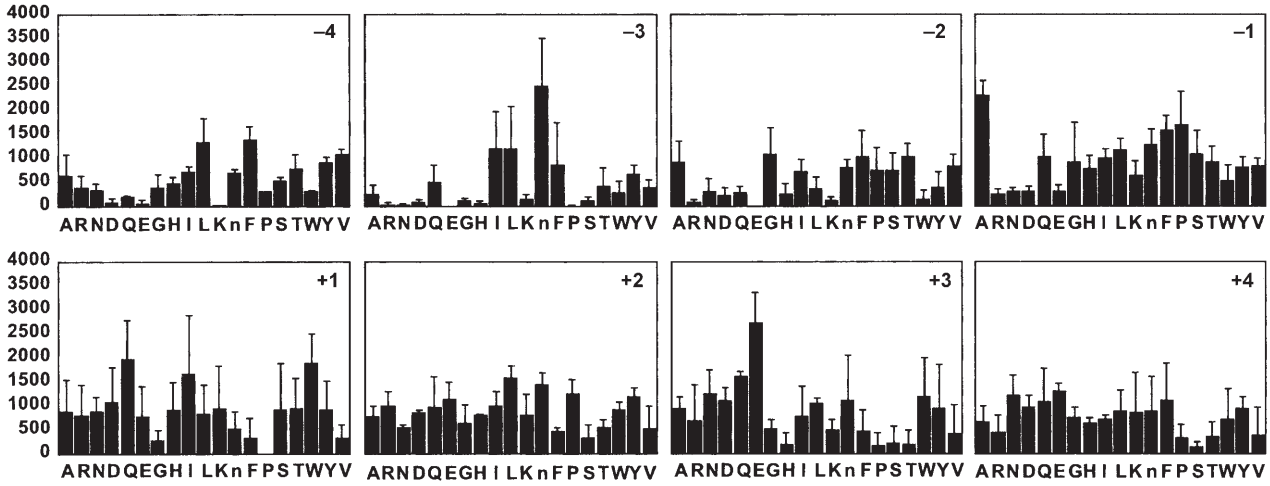


Figure 2. The N- and C-terminal library profiles of TCPTP substrate selectivity. The *y*-axis shows the increase in relative fluorescence intensity produced upon substrate hydrolysis. The amino acids at each position are spatially addressed along the *x*-axis. The numbering scheme represents the relative position with respect to the pCAP residue, as defined in Figure 1.

sites, contrary to the results of others.[15] More recently, a dual preference for substrates containing either acidic or hydrophobic amino acids N-terminal to the phosphorylated residue has been reported, which is in agreement with the results from our library.[28] On the C-terminal side, TCPTP displays a marked bias against Pro at the +1 position, a preference for Glu at the +3 position, and little selectivity at the +2 and +4 positions.

In order to validate the utility of the TCPTP profile, we synthesized peptide substrates based on the most and least preferred amino acids at each position. The library parent sequence is a reasonably good substrate of TCPTP (Table 1). When the "best" amino acids at each position are combined, the resulting peptide, FnGA-pCAP-QLEE, is turned over nearly twice as fast as the parent library sequence. On the other hand, when some of the "worst" amino acids at each position are combined, no hydrolysis of the resulting peptide, DPHR-pCAP-VWKR, was observed. A peptide based on mediocre amino acids at each position, VFDQ-pCAP-HESP, was turned over tenfold less efficiently than the "best" substrate. These results validate the library profiles for TCPTP, demonstrating that combinations of the best, mediocre, or worst amino acids at each position do indeed result in good, average, or poor sub-

Table 1. Kinetic parameters for TCPTP-catalyzed hydrolysis of several peptide substrates.

Substrate	k_{cat} [s^{-1}]	K_m [mM]	k_{cat}/K_m [$M^{-1}s^{-1}\times10^4$]
Ac-DIDE-pCAP-LAA-NH_2	92 ± 5	0.32 ± 0.02	29 ± 3
Ac-FnGA-pCAP-QLEE-NH_2	160 ± 30	0.27 ± 0.05	60 ± 15
Ac-VFDQ-pCAP-HESP-NH_2	11.5 ± 0.5	0.28 ± 0.03	4.1 ± 0.5
Ac-DPHR-pCAP-VWKR-NH_2	NR	NR	NR

NR = no reaction.

strates. It is not surprising that the sequences obtained from the profiles bear little resemblance to known biological substrates as, in a biological context, substrate hydrolysis is governed by a host of factors in addition to the primary sequence surrounding the pY residue, including subcellular localization of both enzyme and substrate, the local secondary and tertiary protein structure surrounding the pY residue, and interactions between domains other than the catalytic domain with the substrate. Nevertheless, the pCAP substrate library profiles provide extremely useful information for both in vitro assays and PTP inhibitor design.

One important application of the substrate specificity profiles described above is the design of potent, selective PTP inhibitors. By replacing the pCAP residue of the optimal substrate with a nonhydrolyzable pY mimic, we obtained a TCPTP inhibitor with an IC_{50} value near 400 nM (Figure 3). The peptide, FnGA-F_2Pmp-QLEE, did not inhibit other members of the PTP family of enzymes, CD45 and YopH, even at a concentration of 250 μM. However, PTP1B, a PTP with a catalytic domain that is approximately 74% identical to TCPTP[34] was also inhibited with an IC_{50} value of 300 nM. This is not surprising given the high homology between the two catalytic domains and the difficulty in achieving selectivity between these two enzymes. However, because even small changes in tyrosine phosphorylation can have significant impact on cellular signaling pathways, we believe that our approach to identifying PTP substrates and inhibitors has great potential. Future work with these libraries will focus on obtaining selectivity among homologous enzymes.

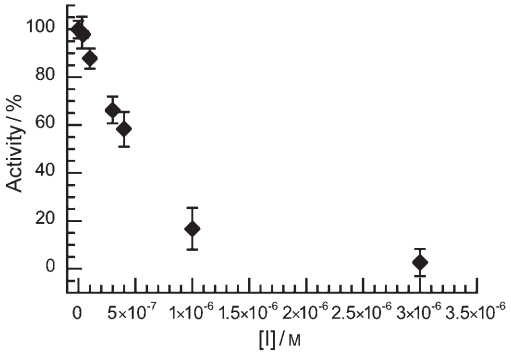


Figure 3. Inhibition of TCPTP by Ac-FnGA-F_2Pmp-QLEE-NH_2.

In summary, we have developed a new approach to rapidly and efficiently profile the substrate selectivity of PTPs using a highly sensitive fluorogenic mimic of pY incorporated into positionally scanned peptide libraries. This approach overcomes several of the challenges associated with profiling PTP substrate selectivity. The fluorescent readout is highly sensitive, direct and continuous and requires no specialized equipment. Furthermore, the positionally addressed library yields data that is easy to deconvolute, requiring little workup after data collection. The selectivity profiles thus obtained are useful in designing both optimal peptide substrates and selective inhibitors for a given enzyme. Given the difficulty of identifying potent, selective PTP substrates and inhibitors, this facile, substrate-library based approach has the potential to be a powerful tool in the study of biological tyrosine phosphorylation.

Acknowledgement

The authors thank Dr. Nunzio Bottini for many helpful discussions and Ms. Caitlin Hubbard for generous technical assistance. This work was supported in part by grants from the USC Zumberge Research and Innovation Fund and the American Cancer Society (IRG-58-007-48).

Keywords: combinatorial libraries · enzymes · inhibitors · phosphopeptides · substrate selectivity

[1] T. Hunter, *Harvey Lect.* **1998**, *94*, 81.
[2] T. Mustelin, R. T. Abraham, C. E. Rudd, A. Alonso, J. J. Merlo, *Front. Biosci.* **2002**, *7*, d918.
[3] T. Mustelin, T. Vang, N. Bottini, *Nat. Rev. Immunol.* **2005**, *5*, 43.
[4] A. W. Stoker, *J. Endocrinol.* **2005**, *185*, 19.
[5] A. Alonso, J. Sasin, N. Bottini, I. Friedberg, I. Friedberg, A. Osterman, A. Godzik, T. Hunter, J. Dixon, T. Mustelin, *Cell* **2004**, *117*, 699.
[6] P. A. Cole, A. D. Courtney, K. Shen, Z. Zhang, Y. Qiao, W. Lu, D. M. Williams, *Acc. Chem. Res.* **2003**, *36*, 444.
[7] L. Tautz, M. Pellecchia, T. Mustelin, *Expert Opin. Ther. Targets* **2006**, *10*, 157.
[8] Z.-Y. Zhang, *Curr. Opin. Chem. Biol.* **2001**, *5*, 416.
[9] Z.-Y. Zhang, *Acc. Chem. Res.* **2003**, *36*, 385.
[10] L. Tautz, T. Mustelin, *Methods* **2007**, *42*, 250.
[11] Z. Zhao, N. F. Zander, D. A. Malencik, S. R. Anderson, E. H. Fischer, *Anal. Biochem.* **1992**, *202*, 361.
[12] S. W. Vetter, Y.-F. Keng, D. S. Lawrence, Z.-Y. Zhang, *J. Biol. Chem.* **2000**, *275*, 2265.
[13] Z.-Y. Zhang, A. M. Thieme-Sefler, D. Maclean, D. J. McNamara, E. M. Dobrusin, T. K. Sawyer, J. E. Dixon, *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 4446.
[14] P. Wang, H. Fu, D. F. Snavley, M. A. Freitas, D. Pei, *Biochemistry* **2002**, *41*, 6202.
[15] M. Ruzzene, A. Donella-Deana, O. Marin, J. W. Perich, P. Ruzza, G. Borin, A. Calderan, L. A. Pinna, *Eur. J. Biochem.* **1993**, *211*, 289.
[16] L. Sun, I. Ghosh, T. Barshevsky, S. Kochinyan, M.-Q. Zu, *Methods* **2007**, *42*, 220.
[17] M. Köhn, M. Gutierrez-Rodriguez, P. Jonkheijm, S. Wetzel, R. Wacker, H. Schroeder, H. Prinz, C. M. Niemeyer, R. Breinbauer, S. E. Szedlacsek, H. Waldmann, *Angew. Chem.* **2007**, *119*, 7844; *Angew. Chem. Int. Ed.* **2007**, *46*, 7700.
[18] C. Blanchetot, M. Chagnon, N. Dube, M. Halle, M. L. Tremblay, *Methods* **2005**, *35*, 44.
[19] X. Espanel, R. Hooft van Huijsduijnen, *Methods* **2005**, *35*, 64.
[20] X. Espanel, M. Huguenin-Reggiani, R. Hooft van Huijsduijnen, *Protein Sci.* **2002**, *11*, 2326.
[21] S. Walchli, X. Espanel, A. Harrenga, M. Rossi, G. Cesareni, R. Hooft van Huijsduijnen, *J. Biol. Chem.* **2004**, *279*, 311.
[22] E. Asante-Appiah, K. Ball, K. Bateman, K. I. Skorey, R. Friesen, C. Desponts, P. Payette, C. Bayly, R. Zamboni, G. Scapin, C. Ramachandran, B. P. Kennedy, *J. Biol. Chem.* **2001**, *276*, 26036.
[23] G. Huyer, J. Kelly, J. Moffat, R. Zamboni, Z. Jia, M. J. Gresser, C. Ramachandran, *Anal. Biochem.* **1998**, *258*, 19.
[24] M. C. Pellegrini, H. Liang, S. Mandiyan, K. Wang, A. Yuryev, I. Vlattas, T. Sytwu, Y.-C. Li, L. P. Wennogle, *Biochemistry* **1998**, *37*, 15598.
[25] Y. W. Cheung, C. Abell, S. Balasubramanian, *J. Am. Chem. Soc.* **1997**, *119*, 9568.
[26] M. Nishikata, Y. Yoshimura, Y. Deyama, K. Suzuki, *Biochimie* **2006**, *88*, 879.
[27] M. Nishikata, K. Suzuki, Y. Yoshimura, Y. Deyama, A. Matsumoto, *Biochem. J.* **1999**, *343*, 385.
[28] M. Garaud, D. Pei, *J. Am. Chem. Soc.* **2007**, *129*, 5366.

[29] A.-S. Wavreille, M. Garaud, Y. Zhang, D. Pei, *Methods* **2007**, *42*, 207.
[30] S. Mitra, A. M. Barrios, *Bioorg. Med. Chem. Lett.* **2005**, *15*, 5142.
[31] S. Mitra, A. M. Barrios, *Anal. Biochem.* **2007**, *370*, 249.
[32] R. A. Houghten, C. Pinilla, S. E. Blondelle, J. R. Appel, C. T. Dooley, J. H. Cuervo, *Nature* **1991**, *354*, 84.
[33] C. Pinilla, J. Appel, S. Blondelle, C. Dooley, B. Dorner, J. Eichler, J. Ostresh, R. A. Houghten, *Biopolymers* **1995**, *37*, 221.
[34] L. F. Iversen, K. B. Moller, A. K. Pedersen, G. H. Peters, A. S. Petersen, H. S. Andersen, S. Branner, S. B. Mortensen, N. P. H. Moller, *J. Biol. Chem.* **2002**, *277*, 19982.

Received: January 24, 2008
Published online on April 15, 2008

