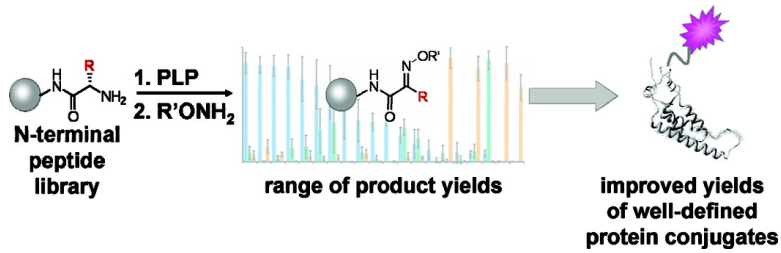
Article

Optimization of a Biomimetic Transamination Reaction

Rebecca A. Scheck, Michel T. Dedeo, Anthony T. Iavarone, and Matthew B. Francis

J. Am. Chem. Soc., **2008**, 130 (35), 11762-11770 • DOI: 10.1021/ja802495w • Publication Date (Web): 07 August 2008



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 1 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
ARTICLES

Published on Web 08/07/2008

Optimization of a Biomimetic Transamination Reaction

Rebecca A. Scheck, Michel T. Dedeo, Anthony T. Iavarone, and Matthew B. Francis*

Department of Chemistry, University of California, Berkeley, California 94720-1460, and Materials Sciences Division, Lawrence Berkeley National Laboratories, Berkeley, California 94720-1460, and QB3/Chemistry Mass Spectrometry Facility, University of California, Berkeley, California 94720-1460

Received April 11, 2008; E-mail: francis@cchem.berkeley.edu

Abstract: For a range of protein substrates, N-terminal transamination offers a convenient way to install a reactive ketone or aldehyde functional group at a single location. We report herein the effects of the identity of N-terminal residues on the product distribution generated upon reaction with pyridoxal 5′-phosphate (PLP). This study was accomplished through the combination of solid-phase peptide synthesis with detailed liquid chromatography−mass spectrometry analysis. Many N-terminal amino acids provided high yields of the desired transaminated products, but some residues (His, Trp, Lys, and Pro) generated adducts with PLP itself. N-terminal Cys and Ser residues were observed to undergo β-elimination in addition to transamination, and the transamination product of N-terminal Gln was resistant to subsequent oxime formation attempts. The information generated through the screening of peptide substrates was successfully applied to a protein target, changing an initially unreactive terminus into one that could be modified in high (70%) yield. Thus, these studies have increased our predictive power for the reaction, both in terms of improving conversion and suppressing reaction byproducts. An initial set of guidelines that may be used to increase the applicability of this reaction to specific proteins of interest is provided.

Introduction

Through the attachment of synthetic functional groups to specific locations on protein targets, bioconjugation has provided a critically important tool for the study of biological function[1–3] and the creation of new materials.[4–7] The utility of such conjugates relies not only on the identity of the chemical modification, but also on the ability to control its placement and the number of times it is displayed on a protein surface. Thus, a number of powerful techniques have been developed to address the site-specificity of bioconjugation reactions. One successful approach has been to create bioconjugates that are genetically encodable, encompassing techniques such as green fluorescent protein fusions,[8,9] tetracysteine-based biarsenical labels,[10,11] non-native amino acid incorporation,[12–15] and several enzyme-based conjugation strategies.[16–20] The expression of proteins displaying unique cysteine residues on their surface is a well-established way to gain access to site-specific conjugates, either through reaction with small molecule electrophiles[1] or C-terminal thioesters.[21,22] As a complement to these techniques, which require genetic manipulation of the protein sequence, it is also possible to create bioconjugates through chemical reactions that target native reactive groups on the protein surface. Historically, many of these methods have suffered from poor site-selectivity due to the frequency with which certain residues, such as lysine, are displayed on the protein surface.[1] Recent work has led to the development of several bioconjugation

(1) Hermanson. G. T. *Bioconjugate Techniques*, 1st ed.; Academic Press: San Diego, 1996.
(2) Francis, M. B. *Chem. Biol.* **2007**, *2*, 593–634.
(3) Evans, M. J.; Cravatt, B. F. *Chem. Rev.* **2006**, *106*, 3279–3301.
(4) Wang, Q.; Lin, T.; Tang, L.; Johnson, J. E.; Finn, M. G. *Angew. Chem., Int. Ed.* **2002**, *41*, 459–462.
(5) Miller, R. A.; Presley, A. D.; Francis, M. B. *J. Am. Chem. Soc.* **2007**, *129*, 3104–3109.
(6) Hooker, J. M.; Kovacs, E. W.; Francis, M. B. *J. Am. Chem. Soc.* **2004**, *126*, 3718–3719.
(7) Broyer, R. M.; Quaker, G. M.; Maynard, H. D. *J. Am. Chem. Soc.* **2008**, *130*, 1041–1047.
(8) Tsien, R. Y. *Annu. Rev. Biochem.* **1998**, *67*, 509–544.
(9) Zhang, J.; Campbell, R. E.; Ting, A. Y.; Tsien, R. Y. *Nat. Rev. Mol. Cell Biol.* **2002**, *3*, 906–918.
(10) Griffin, B. A.; Adams, S. R.; Tsien, R. Y. *Science* **1998**, *281*, 269–272.
(11) Luedtke, N. W.; Dexter, R. J.; Fried, D. B.; Schepartz, A. *Nat. Chem. Biol.* **2007**, *3*, 779–784.
(12) Kirshenbaum, K.; Carrico, I. S.; Tirrell, D. A. *ChemBioChem* **2002**, *3*, 235–237.
(13) Cornish, V. W.; Mendel, D.; Schultz, P. G. *Angew. Chem., Int. Ed. Engl.* **1995**, *34*, 621–633.
(14) Carrico, Z. M.; Romanini, D. R.; Mehl, R. A.; Francis, M. B. *Chem. Commun.* **2008**, *10*, 1205–1207.
(15) Strable, E.; Prasuhn, D. E.; Udit, A. K.; Brown, S.; Link, A. J.; Ngo, J. T.; Lander, G.; Quispe, J.; Potter, C. S.; Carragher, B.; Tirrell, D. A.; Finn, M. G. *Bioconjugate Chem.* **2008**, *19*, 866–875.
(16) Chen, I.; Howarth, M.; Lin, W. Y.; Ting, A. Y. *Nat. Methods* **2005**, *2*, 99–104.
(17) Meier, J. L.; Mercer, A. C.; Rivera, H.; Burkart, M. D. *J. Am. Chem. Soc.* **2006**, *128*, 12174–12184.
(18) Carrico, I. S.; Carlson, B. L.; Bertozzi, C. R. *Nat. Chem. Biol.* **2007**, *3*, 321–322.
(19) Popp, M. W.; Antos, J. M.; Grotenbreg, G. M.; Spooner, E.; Ploegh, H. L. *Nat. Chem. Biol.* **2007**, *3*, 707–708.
(20) Connor, R. E.; Varshavsky, A. K. P.; Tirrell, D. A. *ChemBioChem* **2008**, *9*, 366–369.
(21) Dawson, P. E.; Muir, T. W.; Clarklewis, I.; Kent, S. B. H. *Science* **1994**, *266*, 776–779.
(22) Muir, T. W. *Annu. Rev. Biochem.* **2003**, *72*, 249–289.

10.1021/ja802495w CCC: $40.75

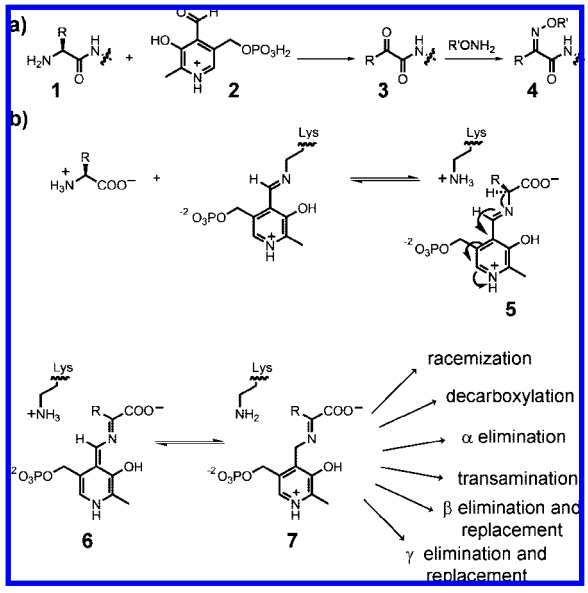


Figure 1. Survey of PLP-mediated reaction pathways. (a) General scheme for PLP-mediated N-terminal transamination of proteins and peptides. (b) Range of reactions mediated by PLP in the biological context.

techniques that offer a high degree of selectivity without the need for genetic engineering.[23–26]

We have recently applied a site-specific biomimetic transamination reaction[27–29] mediated by the biological cofactor pyridoxal 5′-phosphate (PLP, **2**) to the modification of the N-terminus of peptides and proteins.[30,31] The transformation of the N-terminal amino group (**1**) into a ketone or aldehyde (**3**) provides a versatile orthogonal handle for further reaction with alkoxyamines containing functional groups of interest (Figure 1a). Despite the abundant display of lysine side chains on the surface of most proteins, transamination is highly specific to the N-terminus. This is most likely due to the lower pK_a of the α-proton of the N-terminal amino acid, relative to the ε-proton of lysine.

From a mechanistic standpoint, PLP-mediated transamination most likely begins with the formation of a Schiff base at the N-terminus (**5**), followed by tautomerization (**5** to **6**) and subsequent hydrolysis, as is the case for many PLP-mediated reactions in biological settings (Figure 1b).[32,33] A brief summary of known enzymatic reactions that are PLP-mediated is provided in the Supporting Information (Figure S1). In the absence of an enzyme active site, this transformation could be considered to be independent of the identity of the side chain of N-terminal residue. This is certainly an oversimplified view, since our earliest evidence for this reaction resulted from the observation that angiotensin I (N-terminal Asp) underwent decarboxylation along with PLP-mediated transamination.[30] This indicates that the Schiff base intermediate formed between PLP and the N-terminus of a protein or peptide can lead to the formation of alternative products in a fashion that may vary on a case-by-case basis.

The appeal of this reaction is its ability to modify a variety of proteins one time in a single location, but ultimately it is the application at hand that will determine the importance of the overall product yield and the potential for byproduct formation. This technique has already proven to be useful in cases where the sequence is unknown,[31] and when high yields are not crucial. For example, this method has been applied successfully to control the orientation and polydispersity of protein conjugates immobilized on surfaces.[34] However, other types of applications require maximum conversion of a protein target to the oxime product, and thus the ability to predict the reaction outcome is critical.[35] Additionally, applications that require the labeling of complex protein or peptide mixtures would also profit from the ability to predict the behavior of this reaction for various N-terminal amino acids. To facilitate the use of PLP in all of these cases, we present here a more thorough characterization of the effects of the N-terminal residues on the product distribution generated through reaction with PLP. This is achieved using an approach that couples solid-phase peptide synthesis with liquid chromatography−mass spectrometry to assay the reaction products. The results of this study were then extended to the modification of a protein substrate. These studies have increased our predictive power for the reaction, both in terms of improving overall yields and suppressing reaction byproducts. They have also furthered our understanding of the fundamental reaction pathways that are available to N-terminal protein residues.

Results and Discussion

Solid-Phase Approach to the Study of PLP-Mediated Transamination. Initial attempts to study the product distribution of PLP-mediated transamination on small peptide substrates were hampered by the difficulties associated with removing a large excess of PLP. This is less problematic for full-sized proteins because a variety of size exclusion techniques can be used to remove the excess small molecules. We turned to a solid-phase approach in which we could subject the resin-bound peptide substrate to PLP reaction conditions, and then cleave the product mixture after washing away all of the reagents (such as PLP and alkoxyamines used for further derivitization). This was accomplished using PEG-grafted poly(styrene) resin in conjunction with standard Fmoc-based solid phase peptide synthesis techniques. After incubation for 18−20 h with 10 mM PLP, followed by treatment with 250 mM *O*-benzylhydroxylamine, the peptides were cleaved using 5% sodium hydroxide and analyzed by liquid chromatography in-line with mass spectrometry (Figure 2a). As seen in Figure 2b, it is possible to track the progress of the reaction through analysis of the base peak ion chromatogram (BPI), which reports the mass of the highest intensity peak as a function of retention time. After treatment with benzylalkoxyamine, two peaks with identical spectra were observed (Figure 2c). These were assigned as the

(23) Joshi, N. S.; Whitaker, L. R.; Francis, M. B. *J. Am. Chem. Soc.* **2004**, *126*, 15942–15943.
(24) Antos, J. M.; Francis, M. B. *J. Am. Chem. Soc.* **2004**, *126*, 10256–10257.
(25) Tilley, S. D.; Francis, M. B. *J. Am. Chem. Soc.* **2006**, *128*, 1080–1081.
(26) Shaunak, S.; Godwin, A.; Choi, J.-W.; Balan, S.; Pedone, E.; Vijayarangam, D.; Heidelberger, S.; Teo, I.; Zloh, M.; Brocchini, S. *Nat. Chem. Biol.* **2006**, *2*, 312–313.
(27) Dixon, H. B. F. *J. Protein Chem.* **1984**, *3*, 99–108.
(28) Fields, R.; Dixon, H. B. F. *Methods Enzymol.* **1972**, *25*, 409–419.
(29) Wu, P.; Brand, L. In *Flourescence Spectroscopy*; Brand, L., Johnson, M. L., Eds.; Academic Press, 1997; Vol. 278, pp 321−330.
(30) Gilmore, J. M.; Scheck, R. A.; Esser-Kahn, A. P.; Joshi, N. S.; Francis, M. B. *Angew. Chem., Int. Ed.* **2006**, *45*, 5307–5311.
(31) Scheck, R. A.; Francis, M. B. *ACS Chem. Biol.* **2007**, *2*, 247–251.
(32) Eliot, A. C.; Kirsch, J. F. *Annu. Rev. Biochem.* **2004**, *73*, 383–415.
(33) Snell, E. E. *J. Am. Chem. Soc.* **1945**, *67*, 194–197.
(34) Christman, K. L.; Broyer, R. M.; Tolstyka, Z. P.; Maynard, H. D. *J. Mater. Chem.* **2007**, *17*, 2021–2027.
(35) Esser-Kahn, A. P.; Francis, M. B. *Angew. Chem., Int. Ed.* **2008**, *47*, 3751–3754.

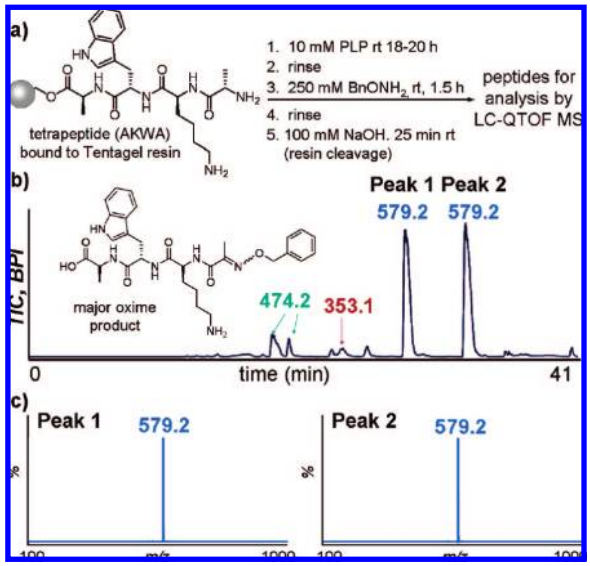


Figure 2. A solid-phase approach to study PLP-mediated transamination of peptides. (a) General modification scheme for using a solid-phase approach to study PLP-mediated transamination. Unless otherwise specified, this method was used for all experiments outlined in this work. (b) Base peak ion (BPI) chromatogram for a sample of AKWA treated with PLP and *O*-benzylalkoxyamine. The main peaks correspond to the oxime product and are labeled in blue. Transaminated products are labeled in green, and the peak labeled in red is due to the oxime formed between PLP and the alkoxyamine. Unmodified AKWA has an expected *m/1* of 475.2 and has a retention time of approximately 12.5 min. For AKWA, no unmodified species are observed after reaction with PLP. (c) Comparison of combined mass spectra for the two oxime products observed in peak one and peak two. The MS/MS spectra also match closely (see Supporting Information).

syn- and *anti*- isomers of the oxime product. In support of this conclusion, fragmentation of this ion (*m/z* 579.2) from both chromatographic peaks yielded MS/MS spectra that were consistent with oxime product (see Supporting Information Figure S2).

N-Terminal Screen for PLP-Mediated Transamination. We used this system to create a tetrapeptide library that displayed each of the 20 natural amino acids at the N-terminus. The backbone peptide ***X-KWA*** was chosen in order to confirm that no reaction occurred with internal Lys residues, and to provide a convenient fluorescent handle (Trp*)*. Quantification of the resulting crude product mixtures was accomplished through integration of the extracted ion chromatograms for each species (Figure 3). Many N-termini gave primarily the expected oxime product (blue). Notably, the residues Ala and Gly, Asp, Glu, and Asn led to very high conversion to the oxime. Beyond these residues, however, it was more difficult to distinguish trends that reflect how the side chain identity influenced the level of conversion. As was observed previously, no modification of internal Lys side chains could be detected.

Unexpected PLP-Mediated Elimination Products. Our initial work with angiotensin I indicated that when a peptide with N-terminal Asp was treated with PLP, decarboxylation of the side chain leads to a pyruvamide product that is identical to the one obtained for N-terminal Ala. We were able to confirm that this is the case through a comparison of base peak ion chromatograms (Figure 4a,b) and MS/MS spectra for AKWA and DKWA. We were surprised to find that identical oxime products were also obtained for CKWA rather than the expected thiazolidine product, as has been previously reported for other aldehydes.[36,37] When treated with PLP, peptides containing N-terminal Cys underwent a β-elimination of the thiol that resulted in the same Ala-derived oxime product (Figure 4c).[38] Close investigation of the product mixture resulting from SKWA indicated that this type of elimination was also possible for N-terminal Ser, albeit with lower yield (Figure 4d). It is interesting to note that, in this case, two peaks with identical spectra were observed for the remaining unmodified peptide, which could be due to the presence of diasteromers resulting from PLP-dependent epimerization of the N-terminal residue. Elimination of the hydroxyl group from N-terminal Thr was not observed, which was likely due to the increased steric bulk of this residue. In each of these cases, the observed elimination products reacted smoothly with alkoxyamines, and were therefore highly compatible with the PLP-mediated transamination methodology.

Products Resistant to Oxime Formation. It was unexpected to find that the oxime formation step—which is widely considered to be a "click" reaction between a ketone or aldehyde and an alkoxyamine[39]—is not quantitative in all cases. Although oxime linkages are generally very stable under physiological conditions,[40,41] it is possible that for certain substrates the reverse reaction (hydrolysis) may be favored over formation of the oxime. Alternatively, the ketone or aldehyde intermediates may interact directly with functionalities on the amino acid side chain, thus becoming unreactive toward attack by the alkoxyamine. As seen in Figure 3, the ketone intermediate (green bars) can be relatively inert to nucleophilic attack by benzylalkoxyamine. This is most clear in the case of N-terminal Gln, which transaminates with high yield (see Supporting Information Figure S3), but is resistant to oxime formation (Figure 5a). However, the product distribution obtained for QKWA can be changed if the side-chain protecting groups remain on the resin-bound peptide during exposure to PLP. In one experiment, these groups were removed by treatment with TFA immediately before incubation with the alkoxyamine, leading to increased levels of oxime formation (Figure 5b). This indicates that the Gln side chain interacts with the ketone intermediate that is formed during reaction with PLP (possibly forming a cyclic hemiaminal). If this interaction does not occur during reaction with PLP, the expected product distribution is observed. This behavior is also observed, albeit to a lesser extent, if the Gln side chain is free during reaction with PLP, and the crude product is exposed to identical acidic conditions before incubation with alkoxyamine. This indicates that the intermediate structure is sensitive to acidic conditions, and thus it might be possible to bias the formation of expected product through changes in the reaction conditions. From a practical perspective, however, we would not recommend the use of N-terminal Gln for applications that require high conversion to the oxime product.

Covalent N-Terminal PLP-Adducts. In our previous work, we have observed that small amounts of byproducts can be formed through covalent attachment of PLP to the N-terminus.[31] In the present work, these PLP adducts were found to be minor

(36) Tam, J. P.; Yu, Q. T.; Miao, Z. W. *Biopolymers* **1999**, *51*, 311–332.
(37) Hang, H. C.; Bertozzi, C. R. *Acc. Chem. Res.* **2001**, *34*, 727–736.
(38) Mihara, H.; Esaki, N. *Appl. Microbiol. Biotechnol.* **2002**, *60*, 12–23.
(39) Kolb, H. C.; Finn, M. G.; Sharpless, K. B. *Angew. Chem., Int. Ed.* **2001**, *40*, 2004–2021.
(40) Calzadilla, M.; Cordova, T.; Malpica, A. *J. Phys. Org. Chem.* **1999**, *12*, 708–712.
(41) Rosenberg, S.; Silver, S. M.; Sayer, J. M.; Jencks, W. P. *J. Am. Chem. Soc.* **1974**, *96*, 7986–7998.

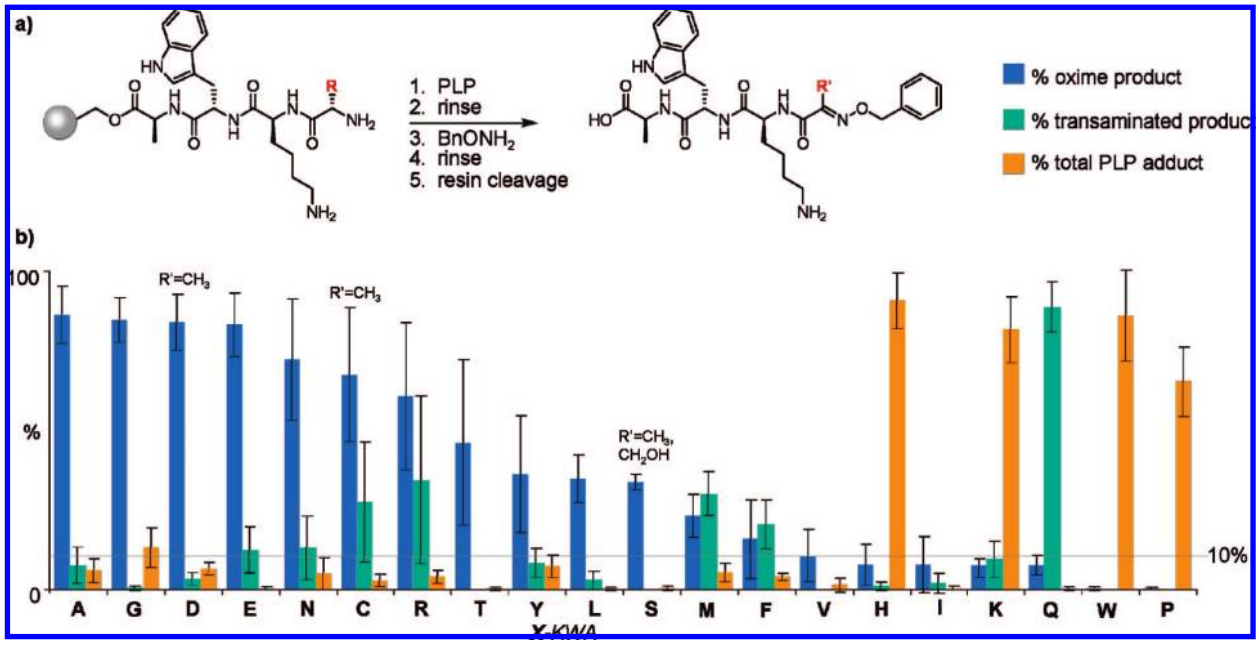


Figure 3. Effects of N-terminal amino acid on PLP-mediated transamination. (a) General method for screening a 20 member tetrapeptide library for PLP reactivity. Reaction conditions are outlined in Figure 2a. (b) Product distribution for tetrapeptides reacted with PLP. Percentages are based on integration of extracted ion chromatograms. The expected oxime product is in blue, transaminated product that did not form an oxime is in green, and PLP adducts are in orange. The sum of all modified and unmodified (not shown) species is 100%. Error bars represent the standard deviation for three replicate experiments.

products in most cases (orange bars, Figure 3), but were highly prevalent for His, Lys, Pro, and Trp residues. A detailed view of the distribution of up to four different PLP adducts is shown for a selection of peptides that were screened (Figure 6). The ratios of different PLP products varied significantly based on the identity of the terminus. In the case of N-terminal Trp and His, complete conversion to a PLP adduct is likely due to a Pictet-Spengler reaction, as previously described for other aldehyde substrates.[42,43]

N-terminal Lys and Pro are two interesting residues that afford a PLP adduct as the major product. In both cases, these covalent adducts can undergo further reaction with alkoxyamines, albeit with varying levels of conversion. For N-terminal Lys, a mass change corresponding to transamination (−1 amu) and cyclic imine formation (−18 amu), was observed after treatment with PLP (see Supporting Information, Figure S4). This intermediate leads to markedly different reactivity as compared to other N-termini, presumably *via* tautomerization to a cyclic enamine. The observed PLP addition products can thus be explained by the nucleophilic addition of this species to the aldehyde group of PLP. Although this final product can still undergo reaction with benzylalkoxyamine, the variable conversion of this reaction prevents this pathway from being used as a practical means for protein modification. Indeed, this is consistent with the observed reactivity for a protein substrate, RNase A, containing an N-terminal Lys residue (see Supporting Information, Figure S5).

As Pro is the only N-terminus that is a secondary amine, it is likely that tautomerization after iminium formation leads to a favorable ring opening hydrolysis step. This yields a pyridoxamine derivative that remains covalently bound to the peptide segment. As was the case for Lys, this Pro derived product can still be modified with benzylalkoxymine. We have found that N-terminal Pro residues follow similar reaction pathways with other aldehydes, such as glyoxylic acid. (see Supporting Information, Figure S6). Further investigation of this unique reactive pathway is being pursued by our laboratory.

As Trp, His, Lys, and Pro tend to form covalent PLP adducts in moderately high yield, we do not consider them to be optimal substrates for PLP-mediated transamination. Conversely, the remaining sixteen N-termini do not form appreciable amounts of these products, and thus remain suitable substrates for this modification technique.

Effect of PLP Concentration on Overall Conversion. We performed a series of experiments to determine whether the product distribution for a given N-terminus can be altered by changing the reaction conditions, such as the PLP concentration. As previously indicated, reaction of N-terminal Lys with PLP leads to a large amount of undesired PLP adduct. As the concentration of PLP was decreased for KKWA, a substantial reduction in the amount of PLP adduct was observed, along with a dramatic increase of the desired oxime product (Figure 7a). The same trend was observed when the PLP concentration was varied for the AKWA substrate (Figure 7b). However, in the case of N-terminal Ala, the amount of undesired PLP-bound products was so minimal that there is only a nominal gain in overall desired reactivity. These results indicate that, at least for peptide substrates, it is possible to shift the product distribution solely by changing the reaction conditions. It should be anticipated that the ideal concentration for optimal conversion to the oxime will depend on the identity of the N-terminal side chain, and thus we recommend that this screen be performed for new targets.

Effect of Internal Sequence on Product Distribution. In the past we have demonstrated that it is possible to modify proteins containing N-terminal Val with good yield.[30] This result contrasts with those generated from our tetrapeptide screen, in

(42) Li, X. F.; Zhang, L. S.; Hall, S. E.; Tam, J. P. *Terahedron. Lett.* **2000**, *41*, 4069–4073.
(43) de la Figuera, N.; Fiol, S.; Fernandez, J. C.; Forns, P.; Fernandez-Forner, D.; Albericio, F. *Synlett* **2006**, 1903–1907.

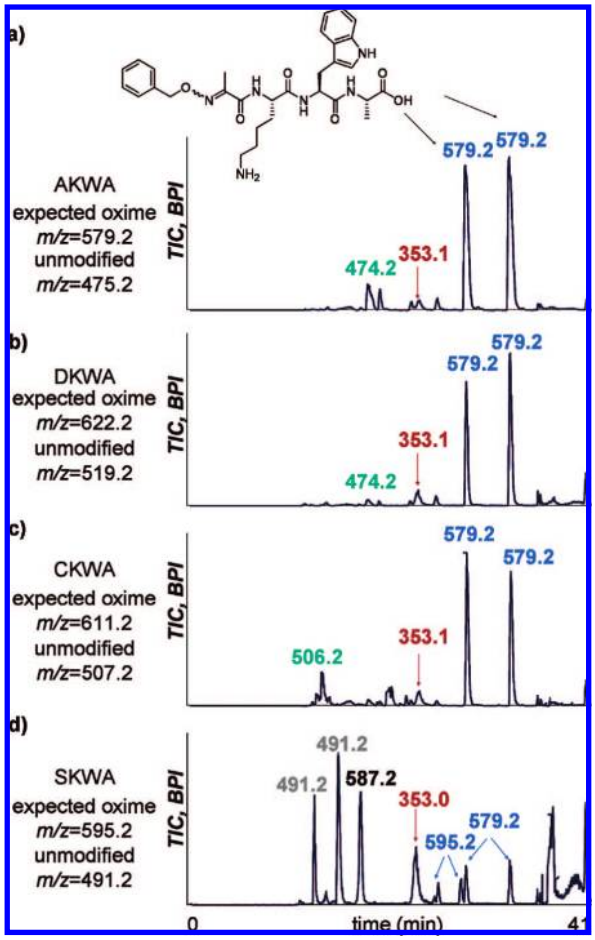


Figure 4. BPI chromatograms for (a) AKWA, (b) DKWA, (c) CKWA, and (d) SKWA after reaction with PLP. In all cases, identical oxime products were observed due to a PLP-dependent β-elimination pathway. Unmodified masses are labeled in gray, transaminated masses are in green, and oxime products are labeled in blue. Unidentified masses are indicated in black. The mass labeled in red (353.1) corresponds to the oxime formed between residual PLP and benzylalkoxyamine.

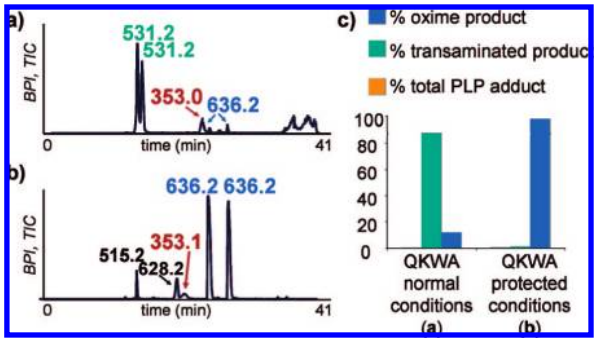


Figure 5. BPI chromatogram for QKWA (a) under normal reaction conditions and (b) under conditions where the peptide side chains are protected during reaction with PLP, and later removed with TFA before incubation with $BnONH_2$. (c) Quantification of products formed in (a) and (b) (oxime peaks are in blue, transaminated peaks are in green, PLP adduct peaks are in orange, and unmodified peaks are in gray). The change in product distribution indicates the presence of a stable intermediate that impedes oxime formation. The unidentified species with the base peak ion 515.2 is an impurity commonly observed in modified and unmodified peptide samples of QKWA. The mass labeled in red (353.1) corresponds to the oxime formed between residual PLP and $BnONH_2$.

which Val is one of the poorest residues (10% conversion to oxime) for this transformation (see Figure 3). This suggests that

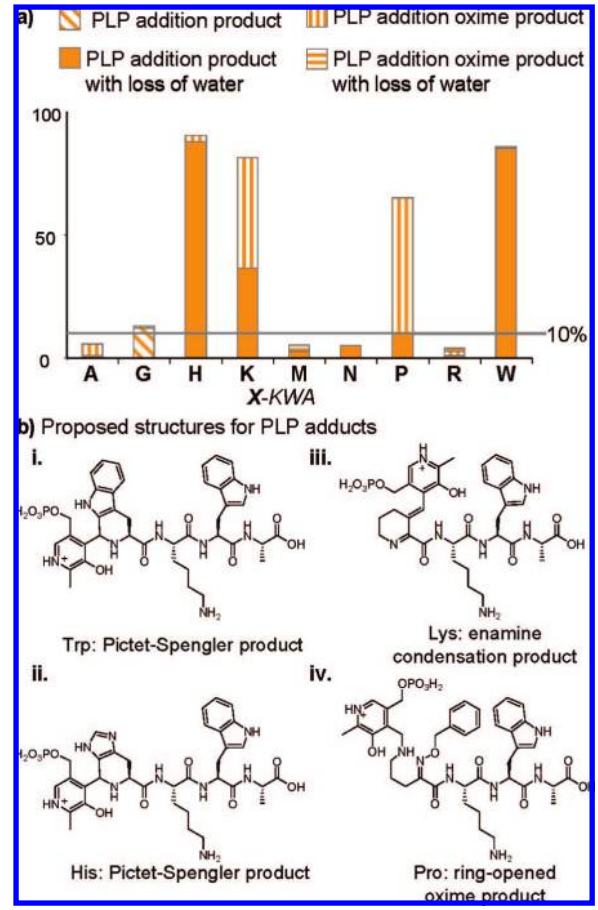


Figure 6. Types of PLP adducts. (a) The distribution of PLP adducts is shown for several tetrapeptides corresponding to data shown in Figure 3. Percentages are based on integration of extracted ion chromatograms. Error bars are omitted here, but can be found in Figure 3. The major PLP-bound product varies per residue. In most cases, the total amount of PLP adduct is not significant (ca. 10%; see Figure 3 for all termini.) All PLP adducts for KKWA include an additional loss of 18 mass units, likely due to cyclic imine formation (see Supporting Information, Figure S4). (b) Proposed structures for the primary PLP adducts of peptides containing N-terminal (i) Trp, (ii) His, (iii) Lys, and (iv) Pro. The PLP products for Lys and Pro can undergo further reaction with alkoxyamines.

the internal sequence can influence the reactivity of the terminal residues. To begin to probe this type of behavior, we made a number of significant changes to the two internal positions of an LXXA tetrapeptide. In this case, Leu was chosen as the terminal residue because it performs with middling success (approximately 30% conversion to oxime) in the XKWA series, and does not lead to a significant amount of side products. We found that the internal sequence does have a significant effect on the product distribution resulting from reaction with PLP (Figure 8a.) The negatively charged (−ED−) sequence, as well the aromatic and aliphatic sequences (−FW− and −AI−) gave poor or no conversion to the oxime product. The nucleophilic sequence (−CQ−) was the only peptide in this screen to yield a significant amount of transaminated product that did not form an oxime. This could be due to the fact that an internal cysteine is well positioned to make a 6-membered ring through reaction with one of the PLP Schiff base tautomers, which would be less reactive toward oxime formation. In the case of a positively charged internal sequence (−HK−), the conversion to oxime product was moderate. From these results, it is tempting to conclude that His is an optimal choice for a second residue. However, further variations at the second position, (while

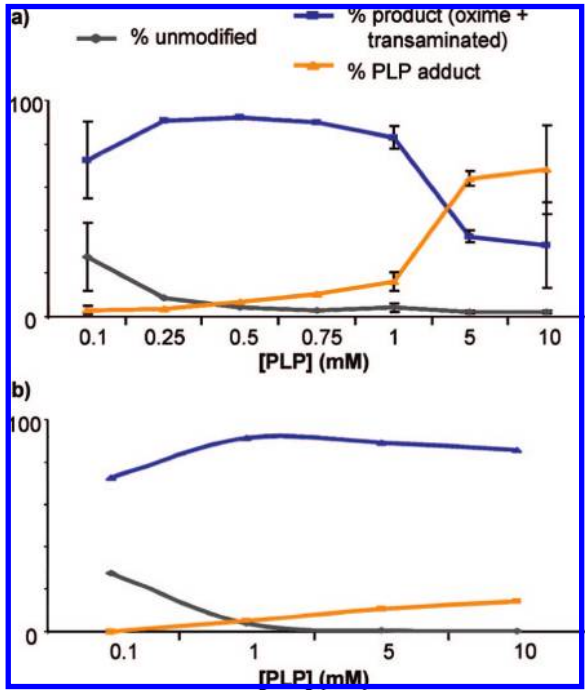


Figure 7. Effect of PLP concentration on product distribution for (a) KKWA and (b) AKWA. In both cases, the amount of PLP adduct (orange) decreases with lower PLP concentration. For KKWA, this leads to a dramatic gain in desired product formation (dark blue). In the case of AKWA, conversion to desired product is high at all PLP concentrations. Product distributions were calculated based on integrations from extracted ion chromatograms.

keeping Lys as the third residue), indicated that a number of residues at the second position all gave high yields. These included residues that had a positive charge, negative charge, and residues that were nonpolar (Figure 8) Notably, the comparison of the product distribution for LFWA and LFKA indicated that the third residue can also have a significant effect on product distribution. Similar results were obtained for the same type of peptides that contained Ala and Gly at the N-terminus. In the case of Met, a larger change in product distribution was observed as the second residue was varied. These results together suggest that the effect of the internal sequence is complex, and appears to result from the synergistic relationship between the internal residues and the N-terminus itself. Certainly these results also indicate that it is possible to generate "winning" sequences that will lead to excellent conversion to oxime product with few side products generated. Current work in our laboratory is beginning to address the subtleties of these effects through the use of split-pool synthesis to generate a much larger peptide library that can cover a broad sequence space.

Applying Peptide Results to Protein Substrates. This solid-phase approach for screening aspects of PLP- mediated transamination proved to be a convenient way to learn about the major pathways of this reaction that would not have been possible to access through work with protein substrates exclusively. Years of experience in our laboratory, however, have shown that peptide substrates often behave differently from proteins. In order to validate the peptide results with a protein substrate, we turned to the tobacco mosaic virus coat protein (TMVP). Our laboratory has utilized TMVP for a number of materials applications; all of these require site-specific conjugation to functional groups.[5,44,45] Unfortunately, no modification of bacterially expressed TMVP occurs under PLP-mediated transamination conditions (Figure 9a,b). In fact, in our hands, the N-terminus of recombinant TMVP is also inert to oxidation of its N-terminal Ser residue with periodate.[46] Based on our peptide data, we set out to engineer a TMV mutant that would be highly reactive under PLP conditions.

The first change we made to the terminus of TMV was to mutate Ser to Ala. Based on our screening results, this alteration should lead to a more reactive terminus (Figure 3). However, we found that S1A was still unreactive to the PEG-alkoxyamine after treatment with PLP. The next step was to extend the N-terminus so that it would be more solvent-exposed, and hopefully lead to higher conversion. Mutants were prepared with

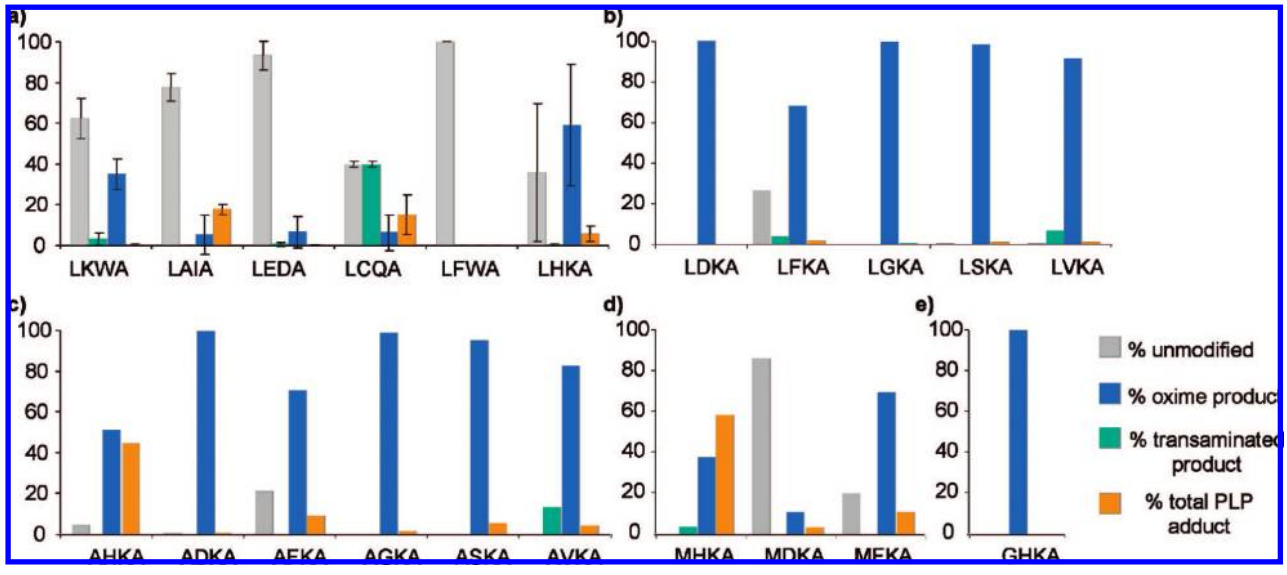


Figure 8. Effects of internal sequence on product distribution. (a) Series in which significant changes were made to the LXXA internal residues. A related set of experiments was performed for a series in which only a single residue was varied for (b) LXKA, (c) AXKA, and (d) MXKA. The results for (e) GHKA are also shown. Product distributions were calculated based on integrations from extracted ion chromatograms. Error bars represent the standard deviations resulting from either 3-fold (LKWA), 4-fold or 6-fold (LHKA) replicate experiments. Experiments shown without error bars were performed once.

Table 1. Guide for Using the PLP-Mediated Transamination Reaction

N-terminus	Expected Product
A, G, D, E, N	Highest conversion to desired oxime
C, R, T, Y, L, S, M, F, or V	Varying levels of conversion to desired oxime; β-elimination can occur with C and S
Q	Transaminated product produced, but turnover to oxime is minimal
W or H	Pictet-Spengler product in high yield
P	PLP adduct formed with some subsequent oxime formation
K	Many possible products. Expect cyclic imine and resulting PLP adducts. Some species may be further modified with alkoxyamine

Outcome	Troubleshooting
High conversion to desired product	None!
Low conversion to oxime product	Try different concentrations of PLP (higher and lower)
High conversion to undesired PLP adducts	Try using lower PLP concentrations
No reactivity, nominal reactivity, or entirely undesired reactivity	Consider site-directed mutagenesis to improve reactivity

N-terminal Ala, Cys. Gly, and Ser, with and without an additional one amino acid spacer before the new terminus. We found that the conversion for all of these termini generally matched the peptide results, in that Ala and Gly give the highest conversion, and Cys and Ser led to more modest yields. It was also observed that a linker separating the new terminal residue from the native terminus gave improved conversion, likely due to improved accessibility. Optimal conversion to the oxime product was obtained for these mutants using higher concentrations of PLP (30 mM). However, even with these higher concentrations, no formation of PLP adduct was observed by ESI-MS (see Supporting Information, Figure S8). These results suggest that secondary and tertiary structure of proteins can lead to an N-terminal environment that is simply not compatible with modification by PLP. This is a key issue with protein substrates, because it is difficult to predict how these environmental conditions will affect reactivity. We believe this must be addressed through empirical results generated through careful mutations made to the terminus of a protein-of-interest, as we have done here with TMV.

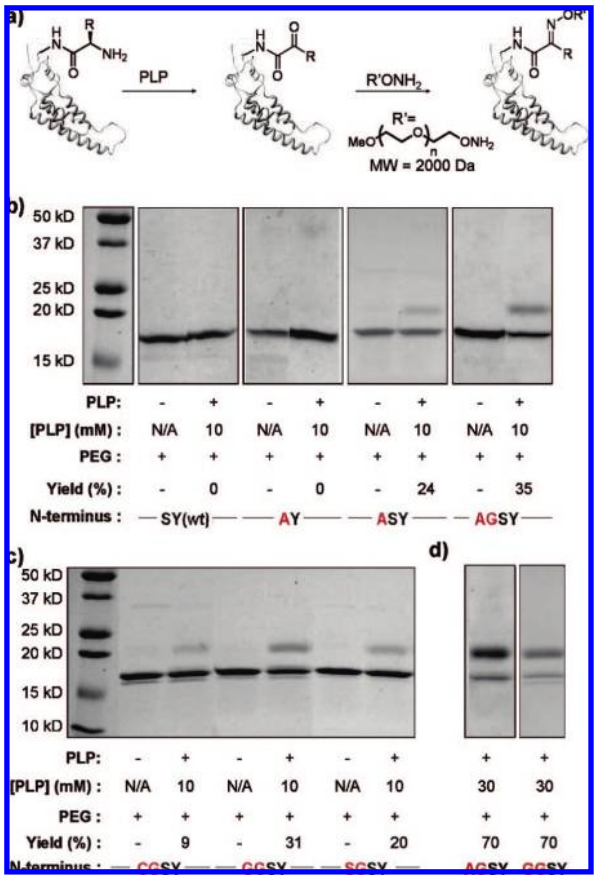


Figure 9. Modification of the TMV capsid protein with PLP. (a) Reaction scheme for PLP modification and oxime formation of TMV. (b) SDS-PAGE analysis of TMV following reaction with PLP and treatment with PEG-alkoxyamine. No reactivity was observed with the wt N-terminus (SY) or an S1A mutant (AY) after 44 h of reaction with PLP. Modification was observed after extension of the N-terminus with A or AG, followed by 24 h treatment with PLP. (c) Reactivity was also observed with N-terminally extended mutants containing N-terminal Cys, Gly, and Ser residues. (d) Higher conversion was attained when the PLP concentration was increased to 30 mM. One letter codes labeled in red indicate sites of mutagenesis.

Conclusion

The results of this study can improve one's ability to predict the success of the reaction and to re-engineer N-terminal sequences to increase reactivity. All observed side products were due to reactions that stem from the N-terminal PLP-Schiff base intermediate; no modification was observed for internal residues. Thus, the value of this reaction for site-specific protein modification remains intact. However, the screening results clearly indicate that the choice of N-terminal residue must be made with care, as is the case for most other selective protein bioconjugation reactions. To help other laboratories who are interested in using this technique, Table 1 summarizes the predicted products for various N-terminal residues and offers troubleshooting suggestions. We demonstrated the utility of these guidelines through the successful N-terminal modification of the TMV coat protein. Our results suggest that optimal product conversion may be attained using slightly different reaction conditions, which may vary based on the identity of the substrate at hand. Current efforts in our laboratory seek to elucidate the complex behavior of the internal sequence identity, as well as to develop PLP derivatives that may improve desired product yields and minimize undesired side reactions.

Materials and Methods

General Procedures. Unless otherwise noted, all chemicals and solvents were of analytical grade and were used as received from commercial sources. Water (dd-H_2O) used in biological procedures or as reaction solvents was deionized using a Barnstead NANOpure purification system (ThermoFisher, Waltham, MA). Pyridoxal 5′-phosphate monohydrate was obtained from Aldrich (St. Louis, MO). All Fmoc-protected amino acids were obtained from Novabiochem (EMD, Germany). Tentagel S OH resin was obtained from Advanced ChemTech (Louisville, KY). The centrifugations required

in spin concentration steps were conducted using an Allegra 64R Tabletop Centrifuge (Beckman Coulter, Inc., Fullerton, CA).

Solid-Phase Peptide Synthesis. Peptides were constructed using standard conditions for Fmoc-based chemistry. The side chain protecting groups used were as follows: Asn(Trt), Asp(tBu), Arg(Pbf), Cys(Trt), Gln(Trt), Glu(tBu), His(Trt), Lys(Boc), Ser(t-Bu), Thr(tBu), Trp(Boc), and Tyr(tBu). The C-terminal amino acid (10 equiv) was preactivated with 5 equivalents of diisopropylcarbodiimide (DIC) and then coupled to the Tentagel S OH resin with 0.1 equivalents of *N,N*-dimethylaminopyridine (DMAP) as an additive. Synthesis was accomplished using 5 equiv of amino acid in *N,N*-dimethylformamide (DMF) with *O*-(benzotriazol-1-yl)-*N,N,N′,N′*-tetramethyluronium hexafluorophosphate (HBTU) (5 equiv) as the coupling reagent with 1-hydroxybenzotriazole (HOBT) (5 equiv) and *N,N*-diisopropylethylamine (DIPEA) (10 equiv) as additives. Peptides were cleaved from the resin with a 25 min incubation with 100 mM sodium hydroxide.

General Procedure for Modification of Resin-Bound Substrates. Fmoc and side-chain protecting groups were removed from approximately 20 mg portions of resin-bound tetrapeptide. After washing, the beads were equilibrated in 1 mL of 50 mM phosphate buffer (pH 6.5) 3 times, for at least 3 min each. After removal of excess buffer, the peptides were treated with 1 mL of 10 mM PLP in 50 mM phosphate buffer, pH 6.5 with 10% DMF for 18−20 h at rt. The beads were washed extensively first with deionized water and then with DMF. The peptides were then incubated with 1 mL of 250 mM *O*-benzylhydroxylamine hydrochloride ($BnONH_2$) for 1.5 h at rt. Excess alkoxyamine was washed away with deionized water followed by DMF. The peptides were cleaved from the resin through incubation with 300 μL of 100 mM sodium hydroxide for 25 min. The resulting peptide solution was added to 700 μL of 50 mM of phosphate buffer at pH 6.5. This solution was diluted 200-fold into dd-H_2O for mass spectrometry analysis.

Liquid Chromatography. LC−MS analyses were performed using a Waters nanoACQUITY UPLC (Milford, MA) liquid chromatograph equipped with C18 trapping (180 μm × 20 mm) and analytical (100 μm × 100 mm) columns and a 10 μL sample loop. Solvent A was 0.1% v/v formic acid/99.9% v/v water and solvent B was 0.1% v/v formic acid/99.9% v/v acetonitrile. Trapping was performed for 2 min with 100% A at a flow rate of 15 μL/min. The elution program consisted of a linear gradient from 15% to 40% B over 25 min, a linear gradient to 95% B over 0.33 min, isocratic conditions at 95% B for 4.67 min, a linear gradient to 1% B over 0.33 min, and then isocratic conditions at 1% B for 7.67 min, at a flow rate of 500 nL/min. The analytical column was maintained at 35 °C and the sample compartment was maintained at 8 °C. The use of the instrument's autosampler for sample injection provided for efficient, unattended analysis in a high-throughput environment.

Mass Spectrometry. The LC was connected online to a quadrupole time-of-flight (Q-Tof) mass spectrometer equipped with a Z-spray electrospray ionization (ESI) source (Q-Tof Premier, Waters). The ion source parameters were as follows: ESI capillary voltage 2.4 kV, nebulizing gas (nitrogen) flow rate 800 L/hr, sample cone voltage 30 V, extraction cone voltage 3 V, ion guide voltage 1 V, source block temperature 80 °C, and nebulizing gas temperature 200 °C. No cone gas was used. The Tof analyzer was operated in "V" mode. Under these conditions, a mass resolving power[47] of 1.0×10^4 was routinely achieved, which is sufficient to resolve the isotopic distributions of the modified and unmodified peptide ions under investigation. Thus, an ion's mass and charge could be determined independently, i.e., the ion charge was determined from the reciprocal of the spacing between adjacent isotope peaks in the *m/z* spectrum. External mass calibration was performed immediately prior to measuring samples, using solutions of sodium formate. Survey scans were acquired over the range *m/z* 100−2000 using a 1.0 s scan integration and a 0.05 s interscan delay. In the data-dependent mode, up to 5 precursor ions exceeding an intensity threshold of 100 counts/second (cps) were selected from each survey scan for tandem mass spectrometry (MS/MS) analysis. Deisotoping and charge state recognition were used to select singly, doubly, and triply charged precursor ions for MS/MS. Collision energies for collisionally activated dissociation (CAD) were automatically selected based on the mass and charge state of a given precursor ion. MS/MS spectra were acquired over the range *m/z* 50−2500 using a 1.0 s scan integration and a 0.05 s interscan delay. Ions were fragmented to achieve a minimum total ion current (TIC) of 15,000 cps in the cumulative MS/MS spectrum for a maximum of 3 s. To minimize the occurrence of redundant MS/MS measurements, real time exclusion was used to preclude reselection of previously analyzed precursor ions over an exclusion width of ± 1 *m/z* unit for a period of 20 s. Data were processed using MassLynx software (version 4.1, Waters).

Construction of TMVP Expression Plasmids. A TMV cDNA clone, pTMVP-S123C, was obtained from laboratory stocks in a PET-20b(+) plasmid. Standard recombinant techniques were used to construct an expression plasmid with pTYB1 vector DNA (NEB, Ipswich, MA) An internal SapI restriction site was removed by introducing a silent mutation (TCT→TCC) using QuikChange mutagenesis (Stratagene/Agilent, Santa Clara, CA). The primers were obtained from IDT (forward 5′- AGCTCTTTCG AGAGCTC-CTC TGGTTTGGTT TGGAC-3′, reverse 5′-GTCCAAACCA AACCAGAGGA GCTCTCGAAA GAGCT-3′). The TMVP-S123C gene was amplified by PCR, using primers obtained from IDT (forward 5′-AGACCATTCA TGTTGTTGCT CAGGTCGC-3′, and reverse 5′-GGTGGTTGCT CTTCCGCAAG TTGCAG-GAC-3′). The amplification product was digested sequentially with SapI and NdeI restriction enzmes (NEB) before ligation into pTYB1 with T4 DNA ligase (NEB). The C-terminal residue of TMV was changed from Thr to Gly using QuikChange mutagenesis to facilitate cleavage by the intein. N-terminal mutants were also made using QuikChange mutagenesis.

Expression and Purification of Recombinant TMV Coat Protein (TMVP). Tuner DE3pLysS competent cells (Novagen) were transformed with TMV, and colonies were selected for inoculation of Terrific Broth cultures. When cultures reached midlog phase as determined by O.D.600, expression was induced by addition of 0.3 mM IPTG (Invitrogen). Cultures were grown for 20 h at 20 °C, harvested by centrifugation, and stored at −80 °C. The following purification steps were performed at 4 °C. Induced cells were thawed and resuspended in 40 mL of Column Buffer (20 mM Tris, pH 8, 500 mM NaCl) and lysed by sonication (Branson Ultrasonics, Danbury, CT.) The cell lysate was cleared by centrifugation at 7000 rpm for 30 min. The cleared lysate was loaded onto a gravity flow column packed with 5 mL of chitin resin equilibrated with 40 mL of Column Buffer. The column was spun on a rotisserie for 1 h to bind the fusion protein to the chitin resin and then washed with an additional 40 mL Column Buffer. This was followed with 15 mL of Cleavage Buffer (100 mM Tris pH 8, 50 mM DTT). The column was capped with 1 mL of cleavage buffer above the resin and left at rt overight. Fractions were analyzed by SDS-PAGE, and fractions containing TMVP were combined and concentrated.

General Procedure for Modification of TMVP. A 1.5 mL Eppendorf tube was charged with a solution of TMV (delivered as 100 μL of a 2 mg ml^{-1} solution in 25 mM potassium phosphate, pH 6.5; 1 equiv) and a solution of PLP (delivered as 100 μL of a 20 mM solution in 25 mM phosphate buffer, pH adjusted to 6.5 with 1 M NaOH; 175 equiv.) The mixture was briefly agitated to

(44) Schlick, T. L.; Ding, Z. B.; Kovacs, E. W.; Francis, M. B. *J. Am. Chem. Soc.* **2005**, *127*, 3718–3723.
(45) Holder, P. G.; Francis, M. B. *Ang. Chem. Int. Ed.* **2007**, *46*, 4370–4373.
(46) Geoghegan, K. F.; Stroh, J. G. *Bioconjugate Chem.* **1992**, *3*, 138–146.
(47) Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S. *Mass Spectrom. Rev.* **1998**, *17*, 1–35.

ensure proper mixing and was incubated without further agitation at 30 °C for 24 h. The PLP was removed from the reaction mixture *via* size exclusion chromatography (Nap-5 desalting column (GE Healthcare) eluting into 25 mM phosphate buffer, pH 6.5.) A 50 uL aliquot of the resulting purified mixture was treated with 50 μL of 0.1 M 2000 MW PEG-ONH_2[44] in a 1.5 mL Eppendorf tube. The mixture was briefly agitated to ensure proper mixing and was incubated without further agitation at rt for 18−20 h. Samples were combined with loading buffer, and analyzed by SDS-PAGE

Gel Analyses. Sodium dodecyl sulfate-poly(acrylamide) gel electrophoresis (SDS-PAGE) was accomplished on a Mini-Protean apparatus from Bio-Rad (Hercules, CA) with 15% gradient polyacrylamide gels (BioRad, CA), following the protocol of Laemmli.[48] All electrophoresis protein samples were mixed with SDS loading buffer in the presence of dithiothreitol (DTT) and heated to 100 °C for 10 min to ensure reduction of disulfide bonds and complete denaturation unless otherwise noted. Commercially available molecular mass markers (Bio-Rad) were applied to at least one lane of each gel for calculation of the apparent molecular masses. Gel imaging was performed on an EpiChem3 Darkroom system (UVP, USA). Protein reaction conversion was estimated from standard optical density measurements of the observed gel bands with Image J software (version 1.34s).

(48) Laemmli, U. K. *Nature* **1970**, *227*, 680–685.

Acknowledgment. This work was generously supported by Chemical Sciences, Geosciences and Biosciences Division, Office of Basic Energy Sciences of the U.S. Department of Energy under Contract No. DE-AC03-765F00098. We also thank the Department of Chemistry at the University of California, Berkeley. The LC−MS instrumentation used in these studies was acquired with support from the National Institutes of Health (grant number 1S10RR022393-01.) We also thank Dante Romanini and members of the Bertozzi and Ellman and Muir laboratories for helpful conversations about peptide synthesis.

Supporting Information Available: Structures for PLP-mediated biological transformations; QTOF-MS/MS analysis for AKWA oxime peaks and QKWA (modified and unmodified.); QTOF-MS spectra for cyclic imine formation of N-terminal Lys; ESI−MS spectra for RNase A under PLP reaction conditions; investigation of the unexpected product distribution obtained for PKWA. This material is available free of charge via the Internet at http://pubs.acs.org.

JA802495W