J|A|C|S
ARTICLES
Published on Web 12/23/2003

Fluorogenic Peptide Sequences − Transformation of Short Peptides into Fluorophores under Ambient Photooxidative Conditions

Gary L. Juskowiak,[†] Shawn J. Stachel,[‡] Parcharee Tivitmahaisoon,[§] and David L. Van Vranken*,[†]

Contribution from the Department of Chemistry, University of California, Irvine, California 92697-2025

Received August 18, 2003; E-mail: dlvanvra@uci.edu

Abstract: Long-lived proteins are susceptible to nonenzymatic chemical reactions and the evolution of fluorescence; however, little is known about the sequence-dependence of fluorogenesis. We synthesized a library of over half a million octapeptides and exposed it to light and air in pH 7.4 buffer to identify fluorogenic peptides that evolve under mild oxidative conditions. The bead-based peptide library was composed of the general sequence H_2N-Ala-$(Xxx)_6$-Ala-resin, where Xxx was one of nine representative amino acids: Asp, Gly, His, Leu, Lys, Pro, Ser, Trp, and Tyr. Next, we selected five highly fluorescent beads from the library and subjected them to microsequencing, revealing the sequence of the unreacted peptide. All five of the fluorogenic sequences were ionic; lacked Tyr, His, and Leu; and most of the sequences contained only one Trp. We then synthesized the five soluble peptides corresponding to the fluorogenic peptide sequences and exposed them to photooxidative conditions. In general, the soluble peptides reacted slowly, generating nonfluorescent monooxygenated and dioxygenated products. However, one peptide (H_2N-AlaLysProTrpGlyGlyAspAla-$CONH_2$) evolved into a highly fluorescent photoproduct as well as a nonfluorescent monooxygenated photoproduct. The fluorescent photoproduct consisted of a 2-carboxy-quinolin-4-yl moiety fused to the N-terminus of GlyGlyAspAla. The formation of this photoproduct requires cleavage of the peptide backbone and a dramatic reorganization of tryptophan. This work demonstrates that sequencing unreacted peptide on beads can reveal sequences with unique nonenzymatic reactivity. The study also confirms that peptide fluorogenesis is dependent on sequence and not merely on the presence of tryptophan. The potential importance of fluorogenic peptide sequences is two-fold. First, fluorogenic sequences that arise through mutation could prove to be hot spots for human aging. Second, fluorogenic sequences, particularly those compatible with intracellular conditions, may serve as fluorescent tags for proteins or as fluorescent biomaterials.

1. Introduction

The correlation between peptide sequence and protein conformation remains a cornerstone of modern structural biology, yet there have been few attempts to correlate peptide sequence with nonenzymatic chemical reactivity. Interest in sequence-dependent peptide reactivity was spurred by the discovery that certain peptide sequences (e.g., Asn-Gly) are prone to rapid clocklike reactions.[1] Most intracellular proteins are turned over rapidly, allowing little opportunity for nonenzymatic reactions to occur. In contrast, many extracellular proteins (e.g., collagens, elastins, proteoglycans, lens proteins) have relatively long lifetimes and slowly react through a combination of enzymatic and nonenzymatic reactions that result in cross-linking and the evolution of visible fluorescence.[2] Amino acids, like tryptophan and tyrosine, are particularly susceptible to oxidative processes that generate fluorescent chromophores, so it would not be surprising to find that some peptide sequences are particularly prone to fluorogenesis. Such sequences would be hot spots for molecular aging, especially in denatured or partially denatured proteins. Thus, the identification of fluorogenic peptide sequences could provide insight into the complicated nonenzymatic oxidative reactions of human proteins, such as those that lead to cataracts. Ultimately, peptide sequences that efficiently form fluorophores in vivo could be used as unobtrusive alternatives to green fluorescent protein.[3]

[†] University of California.

[‡] Present address: Merck Research Laboratories, Medicinal Chemistry Department, West Point, Pennsylvania 19486.

[§] Present address: Roche Bioscience, 3401 Hillview Avenue, Palo Alto, California 94304.

(1) (a) Robinson, A. B.; McKerrow, J. H.; Cary, P. *Proc. Natl. Acad. Sci. U.S.A.* **1970**, *66*, 753−7. (b) Robinson, N. E.; Robinson, A. B. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 944−9. (c) Clarke, S. *Int. J. Pept. Protein Res.* **1987**, *30*, 808−21.

(2) (a) Nagy, I.; Nagy, K. *Mech. Aging Dev.* **1980**, *14*, 245−51. (b) Bron, A. J.; Vrensen, G. F. J. M.; Koretz, J.; Maraini, G.; Harding, J. J. *Ophthalmologica* **2000**, *214*, 86−104. (c) Zeng, H. S.; Macaulay, C.; Mclean, D. I.; Palcic, B. *Photochem. Photobiol.* **1995**, *61*, 639−45. (d) Bjorksten, J.; Tenhu, H. *Exp. Gerontol.* **1990**, *25*, 91−5.

(3) (a) Bastiaens, P. I. H.; Pepperkok, R. *Trends Biochem. Sci.* **2000**, *25*, 631−7. (b) Girotti, M.; Banting, G. *J. Cell Sci.* **1996**, *109*, 2915−26. (c) Simpson, J. C.; Wellenreuther, R.; Poustka, A.; Pepperkok, R.; Wiemann, S. *EMBO Rep.* **2000**, *1*, 237−92. (d) Thomas, C. L.; Maule, A. J. *J. Gen. Virol.* **2000**, *81*, 1851−5.

10.1021/ja0379971 CCC: $27.50

The discovery of nonenzymatic peptide reactions is generally based on serendipity,[4] but combinatorial peptide libraries can greatly accelerate this process.[5,6] Biological libraries offer a higher diversity than bead libraries, but bead libraries offer two distinct advantages. First, bead libraries are compatible with a wider range of reaction conditions; second, they are not subject to artifactual transformations by enzymes in vivo. Several laboratories have used bead-based combinatorial libraries to identify short peptides that catalyze synthetic reactions.[7] However, there have been no studies of intrinsic nonenzymatic peptide transformations using bead-based combinatorial libraries under biological conditions. We set out to identify fluorogenic peptides, particularly under conditions where light might play a catalytic role in the presence of molecular oxygen. Such photocatalytic conditions complement the photoreactivity studies that have been carried out on a relatively small number of tryptophan-containing peptides,[8] usually using photosensitizers or high-intensity light. The aim of this study was to identify fluorogenic peptide sequences by exposing a bead-based peptide library to ambient photooxidative conditions, determine which peptides exhibit fluorogenesis as soluble peptides, and, where possible, elucidate the structure of the fluorophores that are formed.

2. Results and Discussion

2.1. Library Design and Synthesis. A octapeptide library corresponding to H_2N-Ala-$(Xxx)_6$-Ala-NH-resin was prepared from nine representative Fmoc-protected amino acids, excluding those with sulfur: Asp, Gly, His, Leu, Lys, Pro, Ser, Trp, and Tyr. The hexapeptide sequence was flanked by alanines to minimize chemistry specific to the N- and C-termini. The peptide library was prepared on Argogel-NH_2 resin (120−230 μm, 0.41 mmol/g) using the split/couple/pool method of Hruby and Lam.[9] The last step in the synthesis involved acidic side-chain deprotection using trifluoroacetic acid with a mixture of cation scavengers. The library was prescreened for fluorescent beads that might have resulted from the harsh deprotection conditions or from side reactions concomitant to peptide synthesis.

The split/couple/pool method of library construction generates a Poisson distribution of beads. As a convenient scale, we chose to work on 1−2 g of library. Each gram of TentaGel (mean diameter 164 μm) has about 5×10^5 beads.[10] On the basis of its similarity to TentaGel, we assumed that a 1.4 g library of Argogel-NH_2 would consist of about 7×10^5 beads, only slightly larger than the number of possible peptides. In the design of our library, we were forced to choose between two opposing considerations. A library composed of smaller (e.g., 90 μm) more numerous beads would increase the chance that all sequences would be represented, whereas a library composed of larger (120−230 μm) less numerous beads would facilitate bead handling and sequencing. We chose to work with larger beads that would be more amenable to microsequencing.

2.2. Aging of the Library. We used a tungsten−halogen lamp as a light source. In contrast to Hanovia lamps typically used in photochemistry, tungsten−halogen lamps are common light sources in everyday life. The halogen catalyzes the redeposition of tungsten on the filament, but does not dramatically alter the spectral output. The lamp has a maximum output around 650 nm and less than one-tenth of the maximum intensity at 400 nm.[11] The use of Pyrex further reduces the UV exposure of the library. Pyrex has only 50% transmittance to light at 300 nm, but over 90% transmittance at 340 nm and above. The main chromophores, Trp and Tyr, absorb UV light between 280 and 290 nm with the absorbance of Trp greater than Tyr. However, at the levels of exposure used in this experiment, the starting peptides should absorb very little of the light. Thus, in this experiment, light is expected to play no more than a catalytic role.

The peptide library was suspended in pH 7.4 phosphate buffer (which solvates the Argogel resin) in a Pyrex vessel with a fritted bottom, illuminated with a tungsten−halogen lamp, and gently agitated with air. Heat from the lamp kept the solution between 40 and 50 °C. After 2 days, the beads were examined under a long wavelength UV lamp with a broad emission centered at 350 nm. Beads with varying levels of fluorescence were present, and the five most visually fluorescent beads (Figure 1) were removed with a Pasteur pipet and submitted for gas-phase Edman microsequencing. All but one of the beads gave satisfactory sequencing data (Table 1). The third amino acid on bead **5** could not be identified.

2.3. The Fluorogenic Hexapeptide Sequences: Trends and Motifs. We initially expected to find hydrophobic amino acid sequences rich in tryptophan. Solvent exposure often facilitates fluorescence quenching, so it is reasonable to expect that hydrophobic residues might shield the chromophore. Also, tryptophan is known to generate fluorescent kynurenine derivatives at wavelengths above 300 nm.[8b,12] The results of sequencing were contrary to our initial expectations. The sequences were not hydrophobic, they were abundant in aspartic acid and serine, and every sequence contained at least one lysine. The sequences were also not abundant in tryptophan as four out of the five sequences had only one tryptophan. Histidine, tyrosine, and leucine were absent from all of the sequences. The protonated side-chain of histidine is the most efficient quencher of tryptophan fluorescence, but at pH 7.4 histidine and tyrosine would be expected to quench with similar efficiency.[13] A

(4) (a) Flatmark, T. *Acta Chem. Scand.* **1964**, *18*, 1656−61. (b) Flatmark, T.; Sletten, K. *J. Biol. Chem.* **1968**, *243*, 1623−9.
(5) (a) Lam, K. S.; Liu, R. W.; Miyamoto, S.; Lehman, A. L.; Tuscano, J. M. *Acc. Chem. Res.* **2003**, *36*, 370−7. (b) Lam, K. S.; Lebl, M.; Krchňák, V. *Chem. Rev.* **1997**, *99*, 4411−48. (c) Wu, J. Z.; Ma, Q. Y. N.; Lam, K. S. *Biochemistry* **1994**, *33*, 14825−33. (d) Lou, Q.; Leftwich, M. E.; Lam, K. S. *Bioorg. Med. Chem.* **1996**, *4*, 677−82.
(6) Nefzi, A.; Ostresh, J. M.; Giulianotti, M.; Houghten, R. A. *J. Comb. Chem.* **1999**, *1*, 195−8.
(7) (a) Guerin, D. J.; Miller, S. J. *J. Am. Chem. Soc.* **2002**, *124*, 2134−6. (b) Copeland, G. T.; Miller, S. J. *J. Am. Chem. Soc.* **2001**, *123*, 6496−502. (c) Sigman, M. S.; Vachal, P.; Jacobsen, E. N. *Angew. Chem., Int. Ed.* **2000**, *39*, 1279−81. (d) Porter, J. R.; Wirschun, W. G.; Kuntz, K. W.; Snapper, M. L.; Hoveyda, A. H. *J. Am. Chem. Soc.* **2000**, *122*, 2657−8.
(8) (a) Benassi, C. A.; Scoffone, E.; Galiazzo, G.; Iori, G. *Photochem. Photobiol.* **1967**, *6*, 857−66. (b) Pirie, A.; Dilley, K. J. *Photochem. Photobiol.* **1974**, *19*, 115−8. (c) Bent, D. V.; Hayon, E. *J. Am. Chem. Soc.* **1975**, *97*, 2612−9. (d) Holt, L. A.; Milligan, B.; Rivett, D. E.; Stewart, F. H. C. *Biochim. Biophys. Acta* **1977**, *499*, 131−8. (e) Dillon, J. *Photochem. Photobiol.* **1980**, *32*, 37−9. (f) Tassin, J. D.; Borkman, R. F. *Photochem. Photobiol.* **1980**, *32*, 577−85. (g) Simat, T. J.; Steinhart, H. *J. Agric. Food Chem.* **1998**, *46*, 490−8. (h) McKim, S.; Hinton, J. F.; Deghenghi, R. *Biospectroscopy* **1997**, *3*, 317−23.
(9) Lam, K. S.; Salmon, S. E.; Hersh, E. M.; Hruby, V. J.; Kazmierski, W. M.; Knapp, R. J. *Nature* **1991**, *354*, 82−4.
(10) Rapp, W. *Rapp Polymere: US Pricelist 2001*; Rapp Polymere GmbH: Tübingen, 2001; p 4.
(11) Braudaway, G. W.; Wong, H.-S. P. "Color Calibration for the TDI Pro Scanner"; http://www.research.ibm.com/image_apps/colorsci.html, 12/12/01.
(12) Pirie, A. *Biochem. J.* **1971**, *125*, 203−8.
(13) Chen, Y.; Barkley, M. D. *Biochemistry* **1998**, *37*, 9976−82.

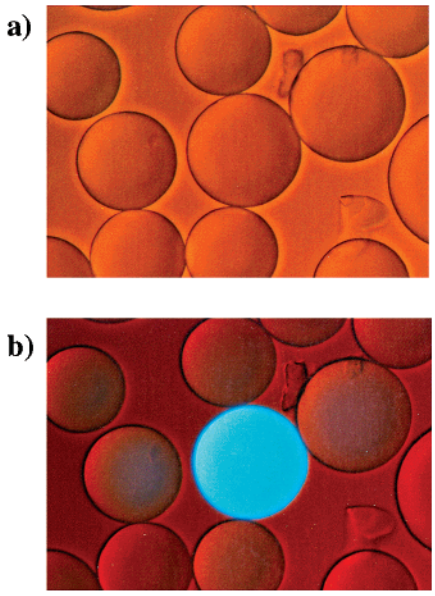

Figure 1. An example of bead selection (peptide **4**) under (a) normal light and (b) normal and long wave UV light. Beads are 120−230 μm in diameter.

Table 1. Fluorogenic Peptide Sequences[a]

1	H_2N-Ala-**Lys-Lys-Trp-Asp-Lys-Ser**-Ala-CONHR
2	H_2N-Ala-**Lys-Trp-Pro-Ser-Lys-Pro**-Ala-CONHR
3	H_2N-Ala-**Asp-Asp-Pro-Ser-Trp-Lys**-Ala-CONHR
4	H_2N-Ala-**Lys-Pro-Trp-Gly-Gly-Asp**-Ala-CONHR
5	H_2N-Ala-**Pro-Xxx-Lys-Ser-Gly-Trp**-Ala-CONHR

[a] For the bead library, R = Argogel; for the solution-phase peptides, R = H.

complete sequence was not obtained for peptide **5**, as position 3 was not analyzable. This could be attributed to an amino acid with a high level of reactivity. We chose to focus on the variant **5a** with tryptophan, the most reactive amino acid in the library, at the third amino acid position.

Molecular fluorogenesis is a prominent aspect of human aging. While rigid conformations would influence reactivity, the presence of sequences corresponding to peptides **1**−**5** in human proteins could indicate hot spots for human aging. A BLAST[blastp] search was performed on human protein sequences in the NCBI nr database against fluorogenic sequences **1**−**5** (expectation value = 2×10^8).[14] The only exact match was on peptide sequence **3** (AspAspProSerTrpLys) as it was found in both brain and lung forms of the enzyme tubulin-tyrosine ligase, but the protein is not known to be associated with a photooxidative disease state.

2.4. Solution-Phase Photooxidation of Peptides 1−5. We then set out to determine which of the five peptide sequences exhibited intrinsic fluorogenic behavior outside of the unique reaction milieu of the bead library. We anticipated that the reaction conditions of the library would contain reactive oxygen species and fragmentation products generated by various bead-bound sequences. Peptides with C-terminal carboxamides corresponding to sequences **1**−**5** were synthesized on Rink MBHA amide resin (Table 1, R = H).

Each of the peptides **1**−**5** was dissolved at 1 mM in pH 7.4 phosphate buffered saline in a Pyrex flask, irradiated with a tungsten−halogen lamp, aerated, and maintained at 40 °C. Under these conditions, the light served to warm the reactions and offer the potential for photocatalysis. The local concentration of peptide on the Argogel beads was significantly higher at approximately 400 mM, but the preparative photooxidation at 1 mM was more representative of biological conditions. The reactions were periodically monitored with a handheld UV lamp with a broad emission centered at 350 nm and by analytical HPLC with UV (254 nm) and fluorescence detection. The chromatograms for the photoreactions of peptides **1**−**5** after 6−9 days are shown in Figure 2.

In contrast to their behavior on the beads, peptides **1**, **2**, and **3** (Figure 2a−c) did not exhibit significant evolution of fluorescence in the solution phase. However, peptide **3** generated products significantly faster than peptides **1** and **2**. Thus, the unique fluorogenic transformation of these three peptides may require the involvement of additional species that were present in the library milieu, but absent from the solution-phase reactions. Alternatively, fluorogenesis could require impurities that were separated from the free peptides, but not from the bead-bound peptides. The minor products detectable by LCMS exhibited masses 16 and 32 higher than the starting peptide corresponding to monooxygenation and dioxygenation products, probably of the tryptophan residue.

In contrast to peptides **1**−**3**, free peptides **4** and **5a** generated highly fluorescent photoproducts, which, in both cases, represented minor components of the reaction mixture. Peptide **4** gave a relatively clean distribution of products in comparison to peptide **5a**; therefore, the fluorogenic transformation of peptide **4** was studied further. The batch of peptide **5a** used in the photooxidation had two major impurities with retention times of 13.6 and 23.0 min, neither of which exhibited significant fluorescence (Figure 2f). The photoreaction of peptide **5a** generated a large number of products, one of which was fluorescent. We attribute the major peak at 13.6 min to the original impurity rather than the new fluorescent product, although it is possible that the original impurity and the fluorescent product coelute.

2.5. Analysis of Products from Photooxidation of Peptide 4. The photooxidation of peptide **4** generates two products of interest: a nonfluorescent photoproduct **A** (RT 15.4 min) and a fluorescent photoproduct **B** (RT 12.0 min) which were analyzed by LCMS prior to purification. Photoproduct **A** ($[M + H]^+$ *m/z* 816) was an oxidation product with a mass 16 higher than the original peptide. The MS/MS (cid) spectrum was rich with peaks consistent with y″ ions and b ions readily allowing the Trp residue to be assigned as the site of oxidation (Figure 3).[15] Two plausible products correspond to addition of an oxygen atom to the tryptophan side chain, oxindolylalanine (Oia) and 3-hydroxypyrrolidinoindoline (Hpi). Much is known about the oxidation and photooxidation of free tryptophan, but not peptides that contain tryptophan.[16] Simat and co-workers have shown that, when Ala-Trp-Ala was oxidized by hydrogen peroxide, Ala-Oia-Ala was the major identifiable product along with *N*-formylkynurenine (NFK).[8g,17] Similarly, Holt and co-workers showed that when AlaGlyTrpVal was irradiated in the presence of oxygen under acidic conditions, the tryptophan

(14) Altschul, S. F.; Madden, T. L.; Schaffer, A. A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D. J. *Nucleic Acids Res.* **1997**, *25*, 3389−402.

(15) (a) Schey, K. L.; Finley, E. L. *Acc. Chem. Res.* **2000**, *33*, 299−306. (b) Finley, E. L.; Dillon, J.; Crouch, R. K.; Schey, K. L. *Protein Sci.* **1998**, *7*, 2391−7.

(16) Creed, D. *Photochem. Photobiol.* **1984**, *39*, 537−62.

(17) Kell, G.; Steinhart, H. *J. Food Sci.* **1990**, *55*, 1120−3.

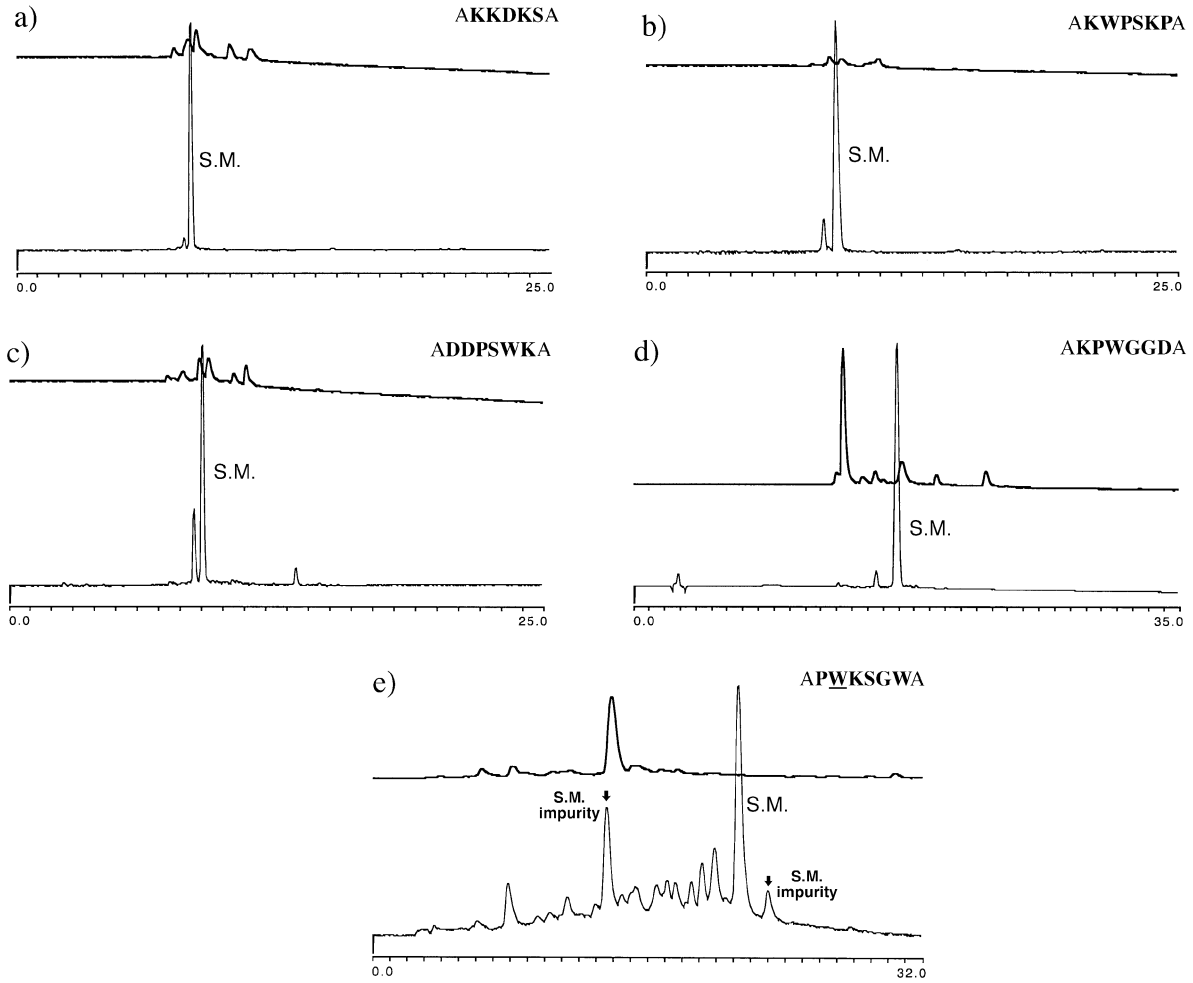

Figure 2. HPLC chromatograms of solution-phase photooxidations of peptides at pH 7.4, 40 °C after about 1 week. The upper chromatogram (bold) is the fluorescence trace (ex 350 nm, em 400 nm), and the lower chromatogram is the UV trace (254 nm). (a) Peptide **1**, 7 days, (b) peptide **2**, 7 days, (c) peptide **3**, 7 days, (d) peptide **4**, 6 days, (e) peptide **5a**, 9 days; the arrows indicate nonfluorescent impurities in the peptide starting material.

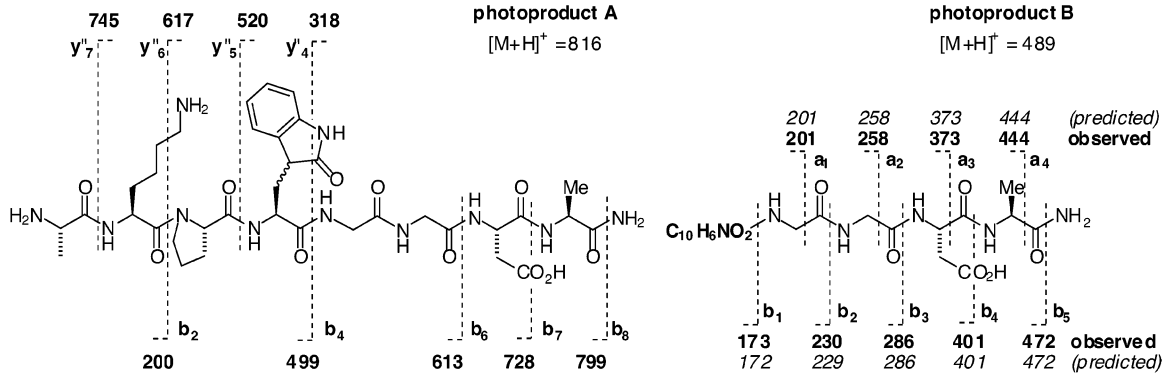

Figure 3. ESI-MS-MS data confirm photoproducts **A** and **B**.

residue was converted to Oia, NFK, and Asp, in nearly equal amounts (4−5% based on amino acid analysis).[8d] Because Oia is the predominant monooxygenation product of tryptophan under both photooxidative conditions and hydrogen peroxide conditions, it is the most plausible structure for photoproduct **A**. It is sobering to consider that in both the Simat and Holt studies, the majority of the oxidation products could not be identified.

Unlike photoproduct **A** which was slightly higher in mass than the starting peptide, photoproduct **B** ($[M + H]^+$ *m*/z 489) was significantly lower in mass. The exact mass of photoproduct **B** was consistent with the molecular formula $C_{21}H_{25}N_6O_8^+$. The MS/MS (cid) spectrum of photoproduct **B** (Figure 3) was rich in peaks, consistent with a ions (iminium ions) and b ions (acylium ions). This abundance of a ions suggests that the N-terminus is readily protonated. The absence of y″ ions is relatively unusual, because most mass spectrometry sequencing is carried out on tryptic fragments with positively charged C-terminal Arg residues or C-terminal Lys residues. Additional peaks at *m*/*z* 329 and *m*/*z* 355 are consistent with a_3-CO_2 and a_3-H_2O, respectively. The fragmentation pattern clearly indicates that photoproduct **B** retains GlyGlyAspAlaCONH$_2$ at the

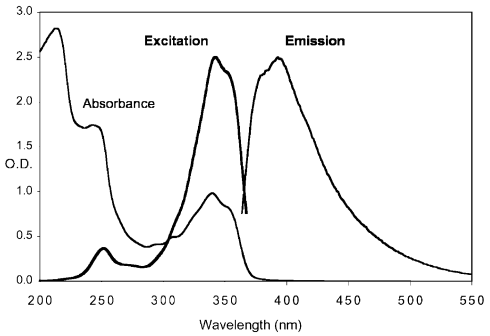

Figure 4. Overlay of absorbance, fluorescence excitation, and fluorescence emission spectra for photoproduct **B**, pH 7.4 PBS.

C-terminus. Thus, the fluorescence of photoproduct **B** must be attributable to the part of the structure ($C_{10}H_6NO_2$) that could not be identified from the MS data. The most direct interpretation is that the 10-carbon fluorophore is derived from the tryptophan residue.

2.6. Isolation of Photoproduct B and Spectroscopic Studies. The fluorogenic transformation of peptide **4** was carried out on a preparative scale. The analytical chromatograms at 220 nm revealed that the fluorescence was due to a minor reaction product with exceptional fluorescent properties rather than an abundant product with modest fluorescent properties. Starch/iodine paper indicated that peroxides were gradually accumulating over the course of the reaction.

Electronic spectra of photoproduct **B** at pH 7.4 revealed an absorption maximum at 339 nm ($\epsilon_{339} = 8800$), an excitation maximum at 340 nm, and an emission maximum at 393 nm (Figure 4). The fluorophore of photoproduct **B** is well suited to stand out in our visual selection process. The excitation maximum at 340 nm overlaps well with the broad emission of our 350 nm UV lamp and with an emission maximum at 393 nm. Most of the light emitted from fluorescence is visible to the human eye.

Unfortunately, our first two attempts to generate fluorescent photoproduct **B** at 50 °C yielded a product that exhibited two closely related sets of signals by ^{1}H NMR yet was homogeneous by HPLC using various mobile phases. While the small quantity of material was suitable for mass spectrometry analysis, it was insufficient to assign the structure of the fluorophore. Fortunately, when the reaction was carried out on a larger scale and at a lower temperature (40 °C), the final product (1.2% yield) exhibited one set of signals by ^{1}H NMR (matching one of the previous sets of signals). The second set of signals could not be regenerated through addition of water, TFA, or triethylamine. The missing set of signals appeared to correspond to an isomeric species because the differences in chemical shift were small. Surprisingly, these differences were centered around the Asp residue rather than the chromophore: Ala8NH (+0.23 ppm), the Asp7NH (−0.13 ppm), Asp7$CH_2\beta$ (−0.08 ppm), AlaHα (+0.06 ppm), Gly6NH (−0.05 ppm), and the Ala$CH_3\beta$ (+0.03 ppm).

The ^{1}H NMR spectrum of photoproduct **B** was assigned using a combination of COSY, NOESY, HMQC, and HMBC (optimized for10 Hz ^{13}C−^{1}H coupling). The COSY (Figure 5a) and NOESY spectra readily allowed assignment of all of the protons of the tetrapeptide fragment GlyGlyAspAla. The splitting pattern of the nonamide protons above 6 ppm and COSY spectrum in DMSO-d_6 were consistent with an ortho-substituted aromatic. One additional nonexchangeable singlet at 7.13 ppm was also part of the chromophore. The HMQC and HMBC (Figure 5b) spectra allowed assignment of all but two of the ^{13}C resonances. This made it possible to establish the carbon connectivity of much of the chromophore, ultimately allowing us to assign the structure as a 4-aminoquinoline-2-carboxylate (Figure 6). The connectivity between the aromatic chromophore and peptide chain was established by strong NOEs between Gly5 and the aromatic protons. The nonamide exchangeable proton at 6.6 ppm is consistent with a pyridinium salt, and based on crystal structures of 4-aminoquinoline-2-carboxylic acid monohydrate[18] we anticipate that the chromophore is present at the zwitterionic form at physiological pH.

4-Aminoquinoline-2-carboxylic acid was not previously known to arise from tryptophan under any conditions and was not previously reported to be fluorescent. Of course, quinolines,

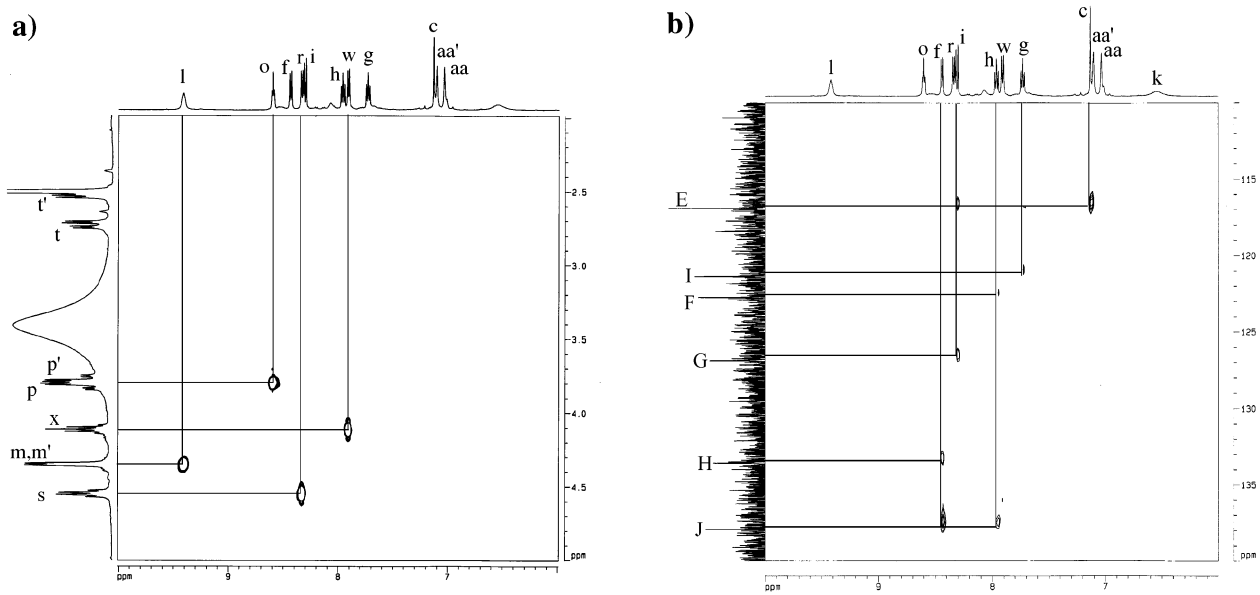

Figure 5. Key regions of the COSY (a) and HMBC (b) spectra for photoproduct **B**.

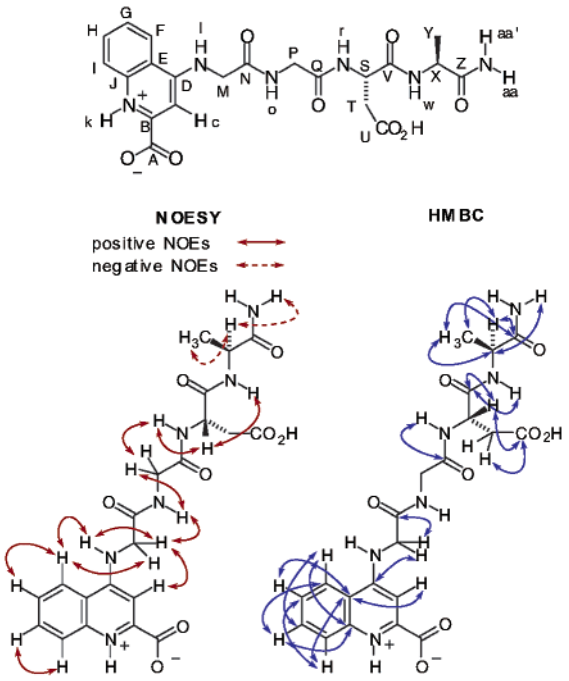

Figure 6. Assigned resonances (upper case = ^{13}C; lower case = ^{1}H) and ^{1}H and ^{13}C correlations from 2D NMR experiments.

in general, are well known to be fluorescent. However, high levels of fluorescence require that the quinoline remain protonated to a significant extent, especially in the most acidic (pK_a −2) of the two excited states.[19] Under such conditions, quinine, a common fluorescence standard, exhibits a fluorescence quantum yield of 0.55.[20] It is satisfying to note that the substituents on the quinoline ring of photoproduct **B** are ideally situated to favor protonation of ring nitrogen. Compounds structurally related to 4-aminoquinoline-2-carboxylic acids, like the endogenous tryptophan metabolite kynurenic acid, antagonize NMDA receptors.[21] Surprisingly, the *N*-glycyl derivative of 4-aminoquinoline-2-carboxylic acid is a potent and selective antagonist of NMDA receptors.[22] It is fortunate that the sequence that generates 4-aminoquinoline-2-carboxylic acid, LysProTrpGlyGlyAsp, is not present in the known human proteome.

When the oxidation of peptide **4** was carried out in the dark, neither of the photoproducts were observed; thus light is essential for the formation of the fluorescent photoproduct **B**. Because peroxides were present during the photooxidation, a control reaction was run to compare the effect of hydrogen peroxide in the presence and absence of light. The fluorescent photoproduct **B** was present in the illuminated reaction, but not in the dark reaction after 5 days of exposure to hydrogen peroxide at pH 7.4. As expected, the nonfluorescent photoproduct **A** was present in both the dark and the illuminated peroxide reactions, but appeared to form faster in the presence of light.

The most surprising structural feature of photoproduct **B** is that the Gly5 amino group is only one carbon removed from the phenyl ring, quite different from the tryptophan connectivity

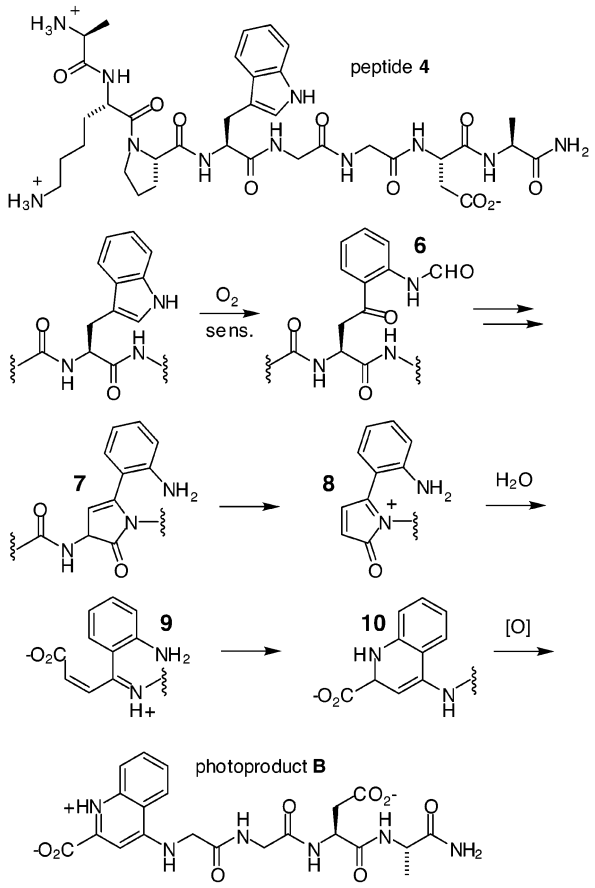

Figure 7. Plausible mechanistic intermediates in the formation of the fluorophore.

from which the chromophore is undoubtedly derived. It is plausible that this connectivity results from condensation of the Gly5 amino group with a kynurenine intermediate **6** as shown in Figure 7. *N*-Formylkynurenines and kynurenines are well-known oxidation products of tryptophan.[23] If the Gly5 amino group condenses with the kynurenine keto group to form enamine **7**, then the N-terminal peptide fragment could eliminate to form an *N*-acyliminium ion **8**. Hydrolysis of the *N*-acyliminium ion would generate an α,β-unsaturated iminium ion **9**, poised to form a dihydroquinoline. Cyclization, followed by oxidation of the electron-rich dihydroquinoline **10**, would generate the ultimate fluorophore.

3. Conclusions

A bead-based library of over half a million hexapeptide sequences was exposed to air and indoor light at pH 7.4 between 40 and 50 °C until beads with strong visible fluorescence emerged. The sequences of the unreacted peptide on each bead were determined. In general, the beads that became fluorescent the fastest had peptide sequences that were ionic; lacked tyrosine, histidine, and leucine; and had one tryptophan residue and one lysine residue. Two of the five peptide sequences exhibited rapid, solution-phase fluorogenesis at 40 °C. Only peptide **4** generated a relatively clean pair of photoproducts, one of which was highly fluorescent. The fluorescent photoproduct **B** from peptide **4** was shown to be a zwitterionic 4-aminoquinoline-2-carboxylic acid derivative generated through

(18) Burd, C. J.; Dobson, A. J.; Gerkin, R. E. *Acta Crystallogr., Sect. C* **1997**, *53*, 602−5.
(19) Schulman, S. G.; Capomacchia, A. C. *J. Am. Chem. Soc.* **1973**, *95*, 2763−6.
(20) Melhuish, W. H. *J. Opt. Soc. Am.* **1964**, *54*, 183−6.
(21) Stone, T. W. *Trends Pharmacol. Sci.* **2000**, *21*, 149−54.
(22) (a) Harrison, B. L.; Baron, B. M.; Cousino, D. M.; McDonald, I. A. *J. Med. Chem.* **1990**, *33*, 3130−2. (b) Harrison, B. L.; Baron, B. M. Ca Patent 2016908, 2001.
(23) Asquith, R. S.; Rivett, D. E. *Biochim. Biophys. Acta* **1971**, *252*, 111−6.

a remarkable peptide cleavage and an unexpected change in chromophore connectivity. 4-Aminoquinoline-2-carboxylic acid is structurally related to kynurenic acid, an enzymatic product in the kynurenine pathway, but it is structurally distinct from fluorescent products that have previously been associated with photooxidation of tryptophan or tryptophan-containing peptides: for example, kynurenine, *N*-formylkynurenine, anthranilic acid, *N*-formylanthranilic acid, and hydroxylated derivatives of these fluorophores.[24] A significant gap remains between the general knowledge that peptides form photoproducts and the structural elucidation of these photoproducts. Previous peptide oxidation studies of tryptophan have been unable to account for the majority of the oxidized tryptophan.[8d,g] This study of fluorogenesis, in particular, the fluorogenesis of peptide **4**, has not closed that gap, but rather the elucidation of the structure of photoproduct **B** has increased our appreciation for the oxidative transformations of tryptophan. This focused study was biased by our choice of time, temperature, light source, and excitation wavelength. More sequences could have been identified by varying any one of these criteria. Ultimately, the solution-phase oxidation studies of the fluorogenic sequences are more laborious than sequencing, yet they are also the most rewarding. Fluorogenic peptide sequence **4**, identified under photooxidative conditions, would not be a useful intracellular fluorescent tag. However, the strategy to sequence unreacted peptides on beads appears to be very promising and could apply to a wide range of chemical modifications of peptides including senescent aging, glycation, and γ-irradiation.

Experimental Section

Photooxidative Aging of the Bead Library. One-half of the library (1.39 g) was prescreened with a long wave UV lamp (350 nm) for fluorescent beads that might have resulted from the synthesis and deprotection chemistry. Five beads were found and discarded. This portion of the resin was transferred to a large Pyrex solid-phase reaction vessel and suspended in 150 mL of pH 7.4 phosphate buffered saline (0.137 M NaCl, 0.0270 M KCl, and 0.100 M sodium phosphate monobasic with pH adjusted to 7.4 with phosphoric acid). The reaction mixture was illuminated with a 500 W tungsten−halogen lamp with a standard J-type halogen bulb while the bead mixture was agitated with a stream of air bubbles. Heat from the lamp led to a reaction temperature of 40−50 °C. After 48 h of illumination, the beads were filtered. The beads were swelled with CH_2Cl_2 and examined under a long wave UV lamp (350 nm), and five highly fluorescent beads were identified and removed using a Pasteur pipet. The fluorescent beads were then subjected to automated gas-phase Edman sequencing.

General Procedure for Solution-Phase Photooxidation of Peptides 1−5. A Pyrex round-bottom flask was charged with peptide at 1−5 mM in sterile pH 7.4 phosphate buffered saline. The flask was fitted with a condenser and placed 20 cm from a standard 500 W J-type halogen bulb. Aluminum foil was used to reflect light back into the reaction vessel, and the majority of the heat from the halogen lamp was dispersed by a stream of air thereby maintaining a reaction temperature of 40 °C. The reaction was mixed and aerated using a gas dispersion tube. The reaction was followed over 7−20 days using analytical HPLC and by visual inspection with a handheld long wave UV lamp.

(24) Creed, D. *Photochem. Photobiol.* **1984**, *39*, 537−62.

Preparative Photooxidation of Peptide 4 (H_2N-Ala-Lys-Pro-Trp-Gly-Gly-Asp-Ala-$CONH_2$). A 500 mL Pyrex round-bottom flask was charged with 257 mg of peptide **4** at 1 mM in 250 mL of sterile pH 7.4 phosphate buffered saline. The flask was fitted with a condenser and placed 20 cm from a standard 500 W J-type halogen bulb. Aluminum foil was used to reflect light back into the reaction vessel, and excess heat from the halogen lamp was dispersed by a box fan thereby maintaining a reaction temperature of 40 (±2) °C. The reaction was mixed and aerated using a gas dispersion tube. The reaction was followed using analytical HPLC (0−50% CH_3CN, over 30 min; 254 and 220 nm) and stopped when the product peaks were taller than the starting material when monitored at 254 nm (20 days). The orange solution was concentrated in vacuo at 40 °C and lyophilized to yield 2.30 g of crude material, primarily buffer salts. The crude solid was taken up in 5 mL of water, and insoluble solids were removed by filtration through nitrocellulose. The solids were further rinsed with 2 × 5 mL water to yield 8.4 mg of material insoluble in both water and DMSO. The soluble filtrate was subjected to preparative HPLC (0−27% MeOH/0.1% aqueous TFA over 55 min; 220 nm) to afford major photoproduct **A** (49.5 mg, 79% purity, 220 nm), fluorescent photoproduct **B** (6.9 mg, 81% purity, 220 nm), and the starting peptide **4** (76.1 mg, 30% yield). Fluorescent photoproduct **B** was resubjected to HPLC purification (5−30% MeCN/0.1% aqueous TFA over 60 min) and lyophilized to give 4.6 mg of photoproduct **B** ≥ 98% homogeneous by HPLC (254 nm). The sample was dried at 70 °C in vacuo for 6 h to give 2.2 mg of product which retained 15 molar equivalents of water by ^{1}H NMR (1.2% yield).

Photoproduct A: ESI-MS-MS 816(80), 799(15), 728(8), 617(100), 599(40), 529(30), 511(15), 502(10), 483(12), 414(15), 357(8), 300-(38), 282(15), 272(10), 200(5); HRMS (ESI) calcd. for $C_{36}H_{54}N_{11}O_{11}^+$ 816.4004 $[M + H]^+$, found 816.3968.

Photoproduct B: UV(MeOH) λ_{max} (log ϵ) 218(4.0), 247(3.9), 339-(3.7), 353(3.6); ^{1}H NMR (500 MHz, DMSO-d_6) δ 12.46 (br s, 1H), 9.42 (br s, 1H), 8.60 (t, J = 5.5 Hz, 1H), 8.44 (d, J = 8.5 Hz, 1H), 8.34 (d, J = 7.7 Hz, 1H), 8.31 (d, J = 8.6 Hz, 1H), 7.97 (t, J = 7.8 Hz, 1H), 7.91 (d, J = 7.6 Hz, 1H), 7.74 (t, J = 7.7 Hz, 1H), 7.13 (s, 1H), 7.11 (d, J = 10.9 Hz, 1H), 7.03 (d, J = 11.8 Hz, 1H), 6.56 (br s, 1H), 4.54 (dd, J = 6.7, 7.2 Hz, 1H), 4.38 (dd, J = 11.4, 17.7 Hz, 1H), 4.36 (dd, J = 11.4, 17.4 Hz, 1H), 4.13 (q, J = 7.4 Hz, 1H), 3.82 (dd, J = 5.5, 16.5 Hz, 1H), 3.77 (dd, J = 5.7, 16.5 Hz, 1H), 2.73 (dd, J = 5.9, 16.6 Hz, 1H), 2.54−2.49 (m, obscured by DMSO-d_5), 1.15 (d, J = 7.2 Hz, 3H); ^{13}C NMR (125 MHz, DMSO-d_6) δ 174.0, 171.9, 169.9, 168.7, 167.6, 161.0, 157.9, 156.9, 137.9, 133.6, 126.9, 122.7, 121.4, 116.9, 98.7, 49.5, 48.3, 45.5, 42.2, 36.0, 17.8. Fluorescence (pH 7) ex (λ_{max}), 340 nm; em (λ_{max}), 393 nm; ESI-MS-MS 489(100), 472(50), 444(45), 401(100), 373(1), 355(5), 329(30), 286(18), 258(65), 230-(10), 201(20), 173(10); HRMS (ESI) m/z calcd for $C_{21}H_{25}N_6O_8^+$ 489.1734 $[M + H]^+$, found 489.1727.

Acknowledgment. This work was supported by NIH GM54523 and the Alfred P. Sloan Foundation, and G.L.J. was supported by the UC Biotechnology Research and Education Program.

Supporting Information Available: General procedures for peptide synthesis, preparative and analytical HPLC; analytical chromatograms for peptides **1−5** and preparative photooxidation of peptide **4**; spectroscopic data for photoproduct **B** (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA0379971