Pergamon

Tetrahedron Letters, Vol. 37, No. 25, pp. 4313-4316, 1996

0040-4039/96 $15.00 + 0.00

PII: S0040-4039(96)00833-7

BIOTIN AND FLUORESCEIN LABELING OF BIOMOLECULES BY ACTIVE ESTERS OF 1-PHENYLPYRAZOLIN-5-ONES[1,2]

Jonathan N. Kremsky[3a], Malcolm Pluskal[3b], Shelagh Casey, Heather Perry-O'Keefe, Steven A. Kates*, and Nanda D. Sinha[3c*]

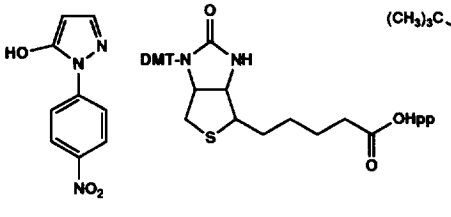
PerSeptive Biosytems, Inc., 500 Old Connecticut Path, Framingham, MA 01701, USA.

Key words: biotin, fluorescein, oligonucleotides, peptide, PNA, probes, labeling

Abstract: The synthesis of 1-phenylpyrazolin-5-one carboxylate esters of biotin and fluorescein and their specific reactivity with amines under mild conditions to label various synthetic biopolymers such as oligonucleotides, peptides and PNA's is described.

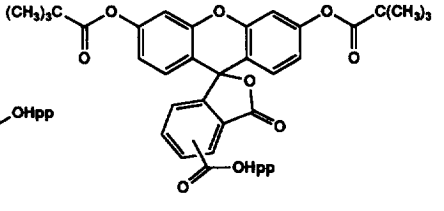
Interest in the use of synthetic biomolecules as probes has increased due to the success of combinatorial chemistry[4] and the proliferation of instrumentation able to detect minute samples of non-radioactive labels.[5] Introduction of a label can be achieved using an activated ester *via* classical solution methods or by post synthetic modification of a resin bound biopolymer. The advantages of the latter method are that the reactions can be driven to completion with excess of reagent which can be removed simply by washes and filtration and is amenable to automation. A disadvantage of active esters in solid-phase methods is that the reagents must be soluble and stable in solution.

The relative reactivities of different substituted enol esters of highly crystalline, 4-substituted 1-phenylpyrazolin-5-ones (Xpp) and their successful use as active esters in solid-phase peptide synthesis has been previously reported.[6] These esters are extremely stable with respect to hydrolysis and their byproducts are of low acidity and soluble in *N,N*-dimethylformamide (DMF) and CH_3CN. Xpp's react specifically with amines and are inert under mild conditions towards the presence of hydroxyl groups in organic solutions, attractive properties for on-line labeling of synthetic biomolecules. These characteristics of Xpp's as well as their solubility in organic solvents make them a potential alternative to *N*-hydroxysuccinimide esters or isothiocyanates which tend to hydrolyze during reaction conditions. To illustrate the utility of this class of heterocyclic compounds, 1-(4'-nitrophenyl)-pyrazolin-5-one (Hpp) **1** active esters of biotin **2** and fluoroscein **3** were chosen to incorporate non-radioactive labels into amine-containing oligonucleotides, peptides and sequences of peptide nucleic acids (PNA).



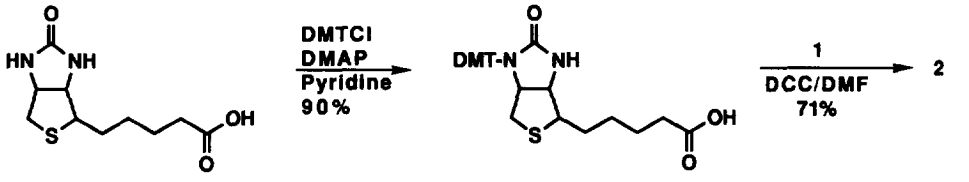
Hpp 1

DMT-Biotin-Hpp 2



Fluor-Hpp 3

Hpp **1** was obtained by the methods described by Hudson.[6c] Using a modification of the procedure of Zhang and Case, treatment of D-(+)-biotin with 4,4'-dimethoxytrityl chloride (DMTCl) and a catalytic amount of 4-dimethylaminopyridine (DMAP) in pyridine afforded 1-*N*-(4,4'-dimethoxytrityl)-D-(+)-biotin[7] in 90% yield (Scheme 1).[8] Condensation of DMT-D-(+)-biotin with *N,N'*-dicyclohexylcarbodiimide (DCC) and **1** in DMF gave DMT-D-(+)-biotin-Hpp **2**[9] in 71% yield. Treatment of 5(6)-carboxyfluorescein with pivaloyl chloride[10] followed by the addition of DCC and Hpp in DMF afforded *bis*-(pivaloylfluorescein)-Hpp active ester **3**[11] in 76% overall yield.



Scheme 1

To demonstrate the utility of this strategy for efficient labeling, various biopolymers were synthesized followed by treatment with the activated esters. Oligodeoxynucleotides were assembled on a PerSeptive Biosystems 8909 Synthesizer at 0.2 μmol scale on controlled pore glass (CPG) using β-cyanoethyl-*N, N*-diisopropylphosphoramidite[12] with benzoyl (Bz) protection on the exocyclic amines with standard DNA protocols. An M13 forward primer (5'-3') AGG GTT TTC CCA GTC ACG AC was synthesized followed by coupling to *N*-monomethoxytritylaminohexa-6-oxycyanoethyl-*N,N*-diisopropylaminophosphoramidite linker. Cleavage of the 5'-MMT group was achieved with a 3% solution of dichloroacetic acid in CH_2Cl_2 for 3 min to give the oligo derivatized with a 6-amino-*n*-hexyl linker at the 5'-terminus (M13-NH_2). The same sequence without an amino linker was also synthesized (M13-OH). To label the oligonucleotides both M13-NH_2 and M13-OH were exposed to a 50 mM solution of **3** in CH_3CN for 10 min followed by treatment with concentrated aqueous ammonia for 18 h at 55 °C to release the DNA from the solid support and remove the Bz and pivaloyl side-chain protecting groups.

HPLC analysis[13] of the crude products from the M13-NH_2 synthesis indicated the usual pattern of a full length product (35 min) and failure sequences (27-29 min) at 260 nm and two absorptions at 495 nm (35 min) indicating the presence of the deprotected fluorescein chromophore (Figure 1). The two peaks from the desired sequence are due to the 5(6)-isomers of the original fluorescein. Previous experience with a fluoresceinated phosphoramidite in our group indicates that the two isomers in conjunction with polyacrylamide gel electrophoresis (PAGE) as sequencing probes possess identical mobility. HPLC analysis[13] of M13-OH oligonucleotide did not exhibit a 495 nm peak, confirming that the Hpp activated esters are specific for labeling amino groups.

The peptide nucleic acid sequence CCT AAT GCA GGA GTC GCAT was assembled on a modified PerSeptive Biosystems 8909 as described by Christensen[14] on Boc-benzhydrylamine polyethylene glycol-polystyrene (Boc-BHA-PEG-PS) resin. Chain elongation was carried out *via* N^{α}-*t*Boc removal with trifluoroacetic acid (TFA)–*m*-cresol (1:19, 2 x 3 min), neutralization with a pyridine–DMF (1:19) wash, acylation with 0.1 M *O*-(7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HATU) 0.1 M dicyclohexylmethylamine, 0.15 M *sec*-collidine and 0.1 M monomer in *N*-methylpyrrolidinone (NMP) for

15 min., capping with acetic anhydride–*sec*-collidine–DMF (1:1.2:17:8) for 5 min, followed by a piperidine–DMF (1:19) wash for reversal of the unwanted acylation of the nucleobases. The PNA-resin was treated with 0.1 M solution of **2** in DMF and 0.2 M solution of *N,N*-diisopropylethylamine (DIEA) in DMF for 15 min. Release of the PNA from the solid support and removal of side-chain protecting groups was accomplished with 2 x 1 h treatments of TFA–trifluoromethanesulfonic acid (TFMSA)–*m*-cresol–thioanisole [6:2:1:1] to give the labeled biopolymer in 83% purity *via* HPLC analysis[13,15] (Figure 2). To examine the sensitivity of the labeled PNA sequence, biotinylated- CCT AAT GCA GGA GTC GCAT was used to probe a blot of a pBR322 digest[16]. The signal was detectable to 0.5 ng of target with an exposure time of 21 min with no evidence of false negatives.

Finally, an analog sequence of prothrombin (1-9) H-ANKGY(OAl)LEEV-NH_2 was carried out in the C→N on PEG-PS resin *via* a PAL linker[17] on a PerSeptive Biosystems 9050Plus continuous-flow synthesizer. Treatment of the N^{α}-amino group with 4 equiv. of **3** and 6 equiv. of DIEA in DMF for 1 h at 25 °C, final cleavage of the anchoring linkage and the *t*Bu, Boc, and Trt side-chain protecting groups with TFA–anisole–β-mercaptoethanol (95:3:2) for 2 h at 25 °C, and treatment with aqueous ammonia for 2 h at 25 °C afforded the fluorescent peptide which was confirmed by HPLC analysis[13] and MALDI-TOF mass spectrometry[18].

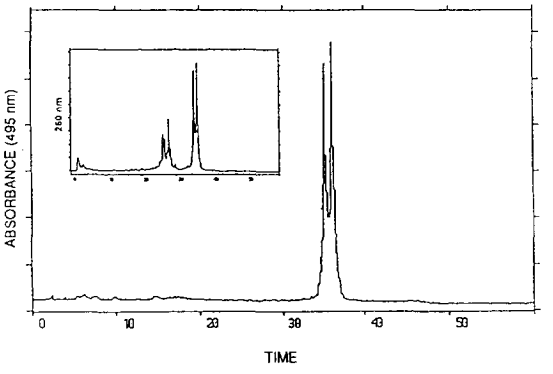


Figure 1. HPLC chromatograms of crude fluoresceinated oligodeoxynucleotide M13-NH_2 (AGG GTT TTC CCA GTC ACG AC) directly after cleavage monitored at 495 nm and 260 nm (inset).

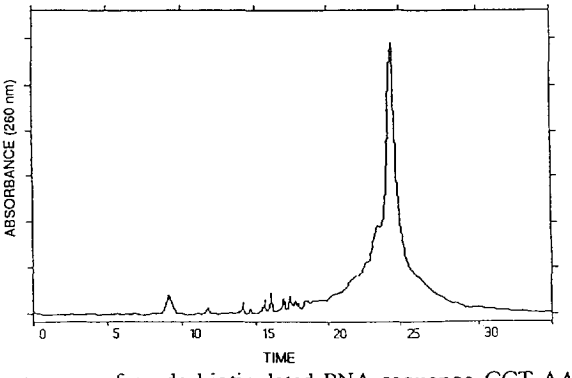


Figure 2. HPLC chromatogram of crude biotinylated PNA sequence CCT AAT GCA GGA GTC GCAT directly after cleavage.

To conclude, carboxylic acid esters of 1-(4-nitrophenyl)pyrazolin-5-one of biotin and fluorescein have been shown to label synthetic biomolecules efficiently. The reaction is specific for amine groups even in the presence of hydroxyl groups. Finally, the stability and handling of these reagents allow for derivatization of the solid supports to occur *via* automation.

References and Notes

1. Portions of this work were presented at the 24th American Chemical Society Northeast Regional Meeting, Burlington, VT, June 19-22, 1994.
2. Amino acids and peptides are abbreviated and designated following rules of the IUPAC-IUB Commission of Biochemical Nomenclature in *J. Biol. Chem.* **1972**, *247*, 977. Additionally, the following abbreviations are used: Al, allyl; BHA, benzhydrylamine; Boc, *tert.*-butyloxycarbonyl; *t*Bu, *tert.*-butyl; Bz, benzoyl; DCC, *N,N'*-dicyclohexylcarbodiimide; DIEA, *N,N*-diisopropylethylamine; DMAP, 4-dimethylamino pyridine; DMF, *N,N*-dimethylformamide; DMT, 4,4'-dimethoxytrityl; DMTCl, 4,4'-dimethoxytrityl chloride; DNA, deoxyribonucleic acid; Fmoc, 9-fluorenylmethyloxycarbonyl; HATU, *O*-(7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate; HOAc, acetic acid; HOBt, 1-hydroxybenzotriazole; Hpp, 1-(4'-nitrophenyl)-pyrazol-5-one; HPLC, high performance liquid chromatography; MALDI-TOF, matrix assisted laser desorption ion time-of flight mass spectrometry,; NMM, *N*-methylmorpholine; NMP, *N*-methylpyrrolidinone; PAGE, polyacrylamide gel electrophoresis; PAL, 5-(4-Fmoc-aminomethyl-3,5-dimethoxyphenoxy)valeric acid; PEG, polyethylene glycol; PNA, peptide nucleic acid; PS, polystyrene, TFA, trifluoroacetic acid; TFMSA, trifluoromethanesulfonic acid; Xpp, 4-substituted 1-phenylpyrazolin-5-one. Amino acid symbols denote the L-configuration unless indicated otherwise.
3. Present address a.) Prime Organics, Inc., 61 Piedmont St., Arlington, MA 02174 b.) Unisyn Technologies, Inc., 25 South St., Hopkinton, MA 01748 c.) Boston Biosystem, Inc., 222 Patridge Lane, Concord, MA 01742
4. For a recent review, see: (a) Gallop, M.A., Barrett, R.W., Dower, R.J., Fodor, S.P.A., Gordon, E.M. *J. Med. Chem.* **1994**, *37*, 1233. (b) Gordon, E.M., Barret, R.W., Dower, R.J., Fodor, S.P.A., Gallop, M.A. *J. Med. Chem.* **1994**, *37*, 1385.
5. Keller, G.H. and Manak, M.M. DNA Probes, Ch.4 &5, Macmillan Publishers, Ltd., U.K. 1989.
6. (a) Hudson, D., U.S. Patent **1993**, 5,233,044. (b) Hudson, D. *In* "Peptides: Chemistry, Structure and Biology, Proceedings of the 11th American Peptide Symposium (J.E. Rivier. and G.Marshall, eds.), Escom, Leiden, 1990, 914. (c) Hudson, D. Peptide Res. **1990**, *3*, 51.
7. ^{1}H NMR (200 Mhz, $CDCl_3$) δ 7.3-6.8 (m, Ar), 5.9(d, CONH), 4.3 (2m, 2H, bridgeheads), 3.75 (s, 6H, $(OCH_3)_2$), 3.0 (m, 1H, HCHS), 2.4 (d, 1H, HCHS), 1.8 (t, 2H, CH_2CO_2), 1.5-1.2 (-CH_2-); FTIR (KBr) 1790 cm^{-1} (CO_2H).
8. Zhang, G. and Case, R. U.S. Patent No. 5,128,476 (1992).
9. ^{1}H NMR (200 Mhz, $CDCl_3$) δ 8.35 (d, 2h, Ar), 7.85 (d, 2H, Ar), 7.67 (d, 1H, H-3), 7.3-6.8 (m, Ar), 6.4 (d, 1H, H-4), 5.3 (d, CONH), 4.3 (2m, 2H, bridgeheads), 3.8 (s, 6H, $(OCH_3)_2$), 3.05 (m, 1H, SCH), 2.5, (t, 2H, CH_2CO_2), 2.4 (m, 1H, HCHS), 2.25 (dd, 1H, HCHS), 1.6-1.2 (m, -$(CH_2)_3$)-; FTIR (KBr): 1695 cm^{-1} (CO_2Ar); mp 125-129 °C.
10. Theisen, P., McCollum, C., Upadhya, K., Jacobson, K., Vu, H. and Andrus, A. *Tetrahedron Lett.* **1992**, *33*, 5033.
11. The ^{1}H NMR spectrum is a series of overlapping aromatic multiplets and the assignments are tentative: ^{1}H NMR ($CDCl_3$) δ 8.8 (meta, s, 1H, H5), 8.3 (H3, 1H, HPP), 8.2 (m, meta + para), 7.8 (d, H4, 1H, HPP), 7.7 (m, meta + para), 7.3 (d, 2H, meta), 7.1 (m, meta + para), 6.8 (m, meta + para), 6.6 (m, para), 1.36 (s, $C(CH_3)_3$); mp 140-144 °C.
12. Sinha, N.D., Biernat, J., McManus, J., Koester, H. *Nucleic Acids Res.* **1984**, *12*, 4539.
13. HPLC analysis was performed with a Waters 600E pump controller with a 490E variable wavelength detector and 860 system for data analysis. Conditions for different products are oligonucleotides: GenPackFax column (Waters) at 65 °C, 0.5 mL/min, Eluent A = 10% CH_3CN in 20 mM $(NH_4)_3PO_4$ (pH 8), Eluent B = eluent A with 1 M NaCl, linear gradient 4:1 to 0:1 in 60 min; peptides and PNA: Waters DeltaPak C_{18} (3.9x150 mm) at 40 °C, 1 mL/min, Eluent A= 0.1% TFA in H_2O, Eluent B = 0.1% TFA in CH_3CN, 1:0 to 1:9 in 30 min.
14. Christensen, L., Fitzpatrick, R., Gildea, B., Warren, W. and Coull, J. *In* "Innovation and Perspectives in Solid Phase Synthesis and Complementary Technologies - Biological and Biomedical Applications" (Epton, R., Ed) Mayflower Worldwide Limited, Birmingham, 1994, 149.
15. As further evidence for the correct structure, matrix assisted laser desorption ion time-of flight mass spectrometry (MALDI-TOF, Voyager, PerSeptive Biosystems) analysis showed MH^+ biotinylated PNA 5956.3 (calcd. 5954.7).
16. Perry-O'Keefe, H. and Kissinger, C.M. *In* "Current Protocols in Molecular Biology" (Ausubel, F.M., et *al.* Eds.) John Wiley, New York, 1992, Suppl. 20, 3.19-1.
17. Albericio, F., Kneib-Cordonier, N., Biancalana, S., Gera, L., Maseda, I., Hudson, D. and Barany, G. *J. Org. Chem.* **1990**, *55*, 3730.
18. The fluoresceinated peptide elutes at 11.7 and 12.7 min whereas the parent peptide elutes at 10 min. MALDI-TOF analysis showed MH^+ 1421 (calcd. 1419).

(*Received in USA* 26 *March* 1996; *revised* 22 *April* 1996; *accepted* 25 *April* 1996)