

## A General Scheme For Incorporating Nonnatural Functionality Into Peptides

Bernard Cuenoud and Alanna Schepartz\*  
Department of Chemistry, Yale University  
New Haven, Connecticut 06511 USA

**Abstract:** We describe a convergent synthesis of  $\Delta$ GCN4<sub>229-248</sub>-terpyridyl **1**, which contains a terpyridyl ligand attached covalently through a disulfide bond to the proposed DNA binding domain of the yeast transcriptional activator protein GCN4.

One of the many powerful applications of solid-phase peptide synthesis<sup>1</sup> is the construction of peptides containing functional groups not normally found in nature. Historically, the nonnatural functionality is introduced into the growing resin-bound peptide chain in the form of a suitably protected amino acid derivative or acylating agent.<sup>2</sup> Before experimentation can begin, the peptide must be cleaved from the resin, deprotected, purified by HPLC, and characterized by amino acid analysis and mass spectroscopy. If the experiments indicate that modifications to the original design are necessary, one must begin anew synthesizing another peptide. This involves non-trivial expenditures of resources as well as the production of considerable waste since peptides are normally synthesized in amounts which far exceed those required for biological experiments. Moreover, it requires that the nonnatural functionality be robust enough to withstand rigorous side chain deprotection and resin cleavage steps.

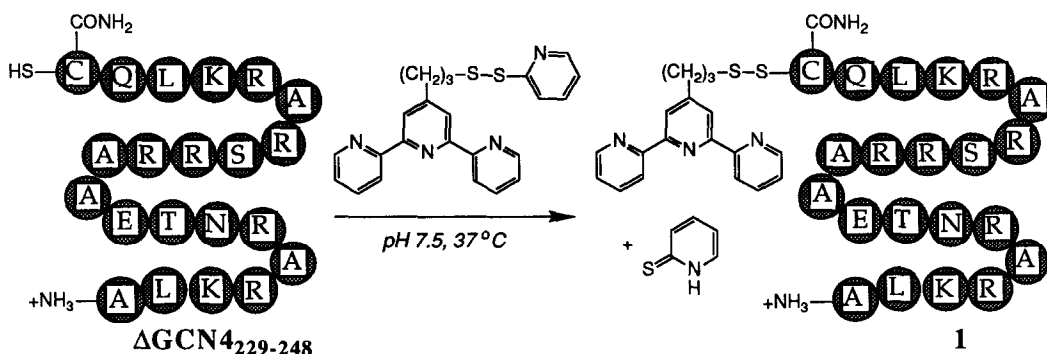


Figure 1

A more pragmatic strategy would involve synthesis of a single peptide bearing a unique and unambiguous locus of reactivity. Once synthesized, purified, and characterized, the peptide could be reacted with one or more reagents capable of covalently linking new functionality to the unique site of reactivity. If subsequent experimentation indicates that modifications to the original design are necessary, one need only synthesize a

different reagent, not an entire peptide.<sup>3</sup> An attractive choice for a unique locus of reactivity on a peptide is the thiol group of a cysteine residue. Conversion of a free thiol into a mixed disulfide is a straightforward procedure resulting in a covalent bond which is exceedingly stable in the absence of excess thiol.<sup>4</sup> Thus, reaction of a peptide thiol with the appropriate activated disulfide in a disulfide interchange reaction would permit the peptide to be modified by a multitude of functional groups.<sup>5</sup> In this letter we describe the use of this strategy to synthesize  $\Delta$ GCN4-terpyridyl conjugate **1**.

GCN4<sup>6</sup> is a transcriptional activator protein responsible for regulating amino acid biosynthesis in yeast.<sup>7</sup> As a member of the "leucine zipper"<sup>8</sup> class of DNA binding proteins, it binds DNA as a dimer.<sup>9</sup> Each monomeric protein contains two domains responsible for DNA recognition: a basic domain consisting of ~25 amino acids which contacts the DNA directly,<sup>10</sup> and a domain containing 30 amino acids responsible for assembling two independently translating protein molecules into a parallel coiled coil, facilitating DNA binding.<sup>11</sup> The peptide *ALKRARNTTEAARRSRARKLQC-NH<sub>2</sub>* ( $\Delta$ GCN4<sub>229-248</sub>) was chosen to constitute a minimal DNA-binding domain of GCN4.<sup>11</sup> It consists of GCN4 residues 229 through 248 with an additional cysteine at the carboxyl terminal to permit formation of a mixed disulfide.  $\Delta$ GCN4<sub>229-248</sub> was synthesized using an Applied Biosystems 430-A Synthesizer and *tert*-butyloxycarbonyl protected amino acid derivatives, purified by HPLC<sup>12</sup> and characterized by amino acid analysis<sup>13</sup> and FAB-mass spectroscopy.<sup>14</sup>

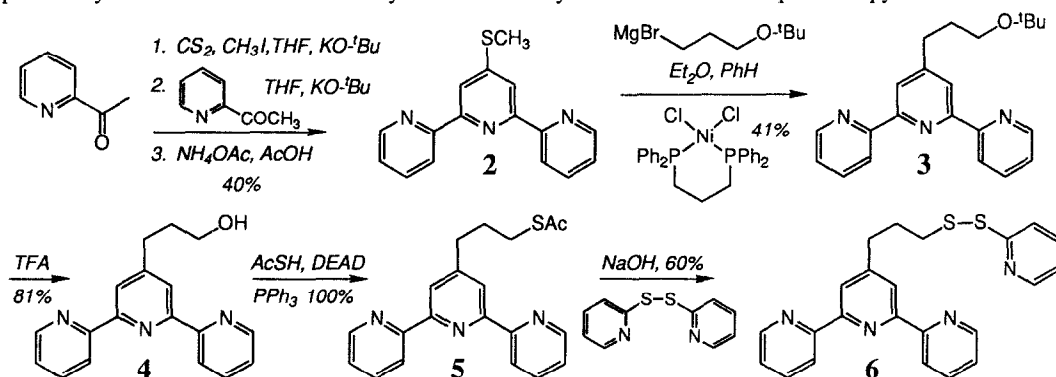


Figure 2

The synthesis of a terpyridyl ligand which can be linked to  $\Delta$ GCN4<sub>229-248</sub> through a disulfide bond is illustrated in Figure 2. Following the general procedure of Potts,<sup>15</sup> the potassium enolate of acetyl pyridine was reacted with one equivalent of carbon disulfide followed by two equivalents of methyl iodide to generate the  $\alpha$ -oxoketenethioacetal. This species was not isolated, but was reacted with an additional equivalent of acetyl pyridine potassium enolate and then cyclized with ammonium acetate in acetic acid to provide 4'-methylthio-2,2':6,2''-terpyridyl **2**.<sup>16</sup> Treatment of this compound with the Grignard reagent derived from the *tert*-butyl ether of 3-bromopropanol<sup>17</sup> in the presence of Ni(II)Cl<sub>2</sub>(PPh<sub>3</sub>)<sub>2</sub> produced none of the desired *tert*-butyl ether **3**. However, the more exchange-inert catalyst Ni(II)Cl<sub>2</sub>dppp (dppp=1,3-bis(diphenylphosphino)propane)<sup>18</sup> provided **3** in a 41% yield.<sup>19</sup> The *tert*-butyl ether was removed by use of trifluoroacetic acid (81%) and the alcohol **4**<sup>20</sup> was converted into the thiolacetate **5**<sup>21</sup> using the procedure of Volante (100%).<sup>22</sup> The thiolacetate was hydrolyzed, and reacted *in situ* with 2-thiopyridonedisulfide to provide the mixed disulfide **6**.<sup>23</sup>

$\Delta$ GCN4229-248 reacted readily with activated disulfide **6** to generate  $\Delta$ GCN4229-248-terpyridyl **1**.  $\Delta$ GCN4229-248 (1 mM) was incubated with terpyridyl **6** (5 mM) in a degassed 40% aqueous acetonitrile solution (10 mM phosphate, pH 7.5) at 37 °C, and the extent of reaction was monitored by reverse phase HPLC.<sup>24</sup> Monomeric  $\Delta$ GCN4229-248 elutes with a retention time of 12.2 minutes under these conditions, while the  $\Delta$ GCN4229-248 disulfide dimer (generated in small amounts) elutes at 15.5 minutes. Over a period of three hours, we observe the appearance of a new peak at 21.4 minutes. This peak was isolated and shown to be the desired product **1** by UV/VIS spectroscopy,<sup>25</sup> amino acid analysis,<sup>26</sup> FAB-MS<sup>27</sup> and by the quantitative conversion into  $\Delta$ GCN4229-248 upon treatment with excess dithiothreitol.<sup>28</sup> Future studies will focus on the ability of **1** to self-assemble into a sequence-specific DNA binding complex.

**Acknowledgement:** We gratefully acknowledge support from the National Foundation for Cancer Research, Merck & Co., Inc., and the David and Lucile Packard Foundation (fellowship to A.S.). NMR spectra were obtained through the auspices of the Northeast Regional NSF/NMR Facility at Yale University, which was supported by Chemistry Division Grant CHE 7916210. B.C. gratefully acknowledges the Bristol-Myers Squibb Company for a Predoctoral Fellowship.

#### REFERENCES AND NOTES

- Merrifield, R.B. *Adv. Enzymol.* **1969**, *32*, 221-296. Kent, S.B.H. *Ann. Rev. Biochem.* **1988**, *57*, 957-989.
- For some recent examples, see: Sasaki, T.; Kaiser, E.T. *J. Am. Chem. Soc.* **1989**, *111*, 380-381. Ruan, F.; Chen, Y.; Hopkins, P.B. *J. Am. Chem. Soc.* **1990**, *112*, 9403-9404. Sluka, J.P.; Griffin, J.H.; Mack, D.P.; Dervan, P.B. *J. Am. Chem. Soc.* **1990**, *112*, 6369-6374. Lieberman, M.; Sasaki, T. *J. Am. Chem. Soc.* **1991**, *113*, 1470-1471.
- Several related strategies for the synthesis of modified oligonucleotides have been described: Draper, D.E.; *Nucl. Acids Res.* **1984**, *12*, 989-1002. Zuckerman, R.; Corey, D.; Schultz, P. *Nucl. Acids Res.* **1987**, *15*, 5305-5321. Gibson, K.J.; Benkovic, S.J. *Nucl. Acids Res.* **1987**, *15*, 6455-6467. Fidanza, J.A.; McLaughlin, L.W. *J. Am. Chem. Soc.* **1989**, *111*, 9117-9119. Sproat, B.S.; Lamond, A.I.; Beijer, B.; Neuner, P.; Ryder, U. *Nucl. Acids Res.* **1989**, *17*, 3373-3386. MacMillan, A.M.; Verdine, G.L. *J. Org. Chem.* **1990**, *55*, 5931-5933. Povsic, T.J.; Dervan, P.B. *J. Am. Chem. Soc.* **1990**, *112*, 9428-9430.
- Creighton, T.E. *Proteins*; W.H. Freeman and Co.: New York. **1984**; pp. 52-53.
- A related strategy has been used for linking oligonucleotides to proteins: Zuckerman, R.N.; Schultz, P.G. *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 1766-1770. Corey, D.R.; Schultz, P.G. *Science* **1987**, *238*, 1401-1403.
- Penn, M.D.; Galgoci, B.; Greer, H. *Proc. Natl. Acad. Sci. USA* **1983**, *80*, 2704-2708. Hinnebusch, A.G.; Fink, G.R. *Proc. Natl. Acad. Sci. USA* **1983**, *80*, 5374-5378.
- Hope, I.A.; Struhl, K. *Cell* **1986**, *46*, 885-894. Hope, I.A.; Struhl, K. *Cell* **1985**, *43*, 177-188. Hill, D.E.; Hope, I.A.; Macke, J.P.; Struhl, K. *Science* **1986**, *234*, 451-457.
- Vinson, C.R.; Sigler, P.B.; McKnight, S.L. *Science* **1989**, *246*, 911-916. Landschultz, W.H.; Johnson, P.F.; McKnight, S.L. *Science* **1988**, *240*, 1759-1764. Oas, T.G. *et al. Biochemistry* **1990**, *29*, 2891-2894. For reviews, see: McKnight, S.L. *Scientific American* **1991**, April, 54-64. Johnson, P.F.; McKnight, S.L. *Ann. Rev. Biochem.* **1989**, *58*, 799-839.
- Hope, I.A.; Struhl, K. *EMBO J.* **1987**, *6*, 2781-1784.
- Sellers, J.W.; Struhl, K. *Nature* **1989**, *341*, 74-76. Agre, P.; Johnson, P.F.; McKnight, S.L. *Science* **1989**, *246*, 922-926. Gartenberg, M.R. *et al. Proc. Natl. Acad. Sci. USA* **1990**, *87*, 6034-6038. Shuman, J.D.; Vinson, C.R.; McKnight, S.L. *Science* **1990**, *249*, 771-774. Oakley, M.G.; Dervan, P.B. *Science* **1990**, *248*, 847-850. Weiss, M.A. *et al. Nature* **1990**, *347*, 575-578. Saudek, V. *et al. Biochemistry* **1991**, *30*, 1310-1317.
- O'Shea, E.K.; Rutkowski, R.; Kim, P.S. *Science* **1989**, *243*, 538-542. O'Neil, K.T.; Hoess, R.H.; DeGrado, W.F. *Science* **1990**, *249*, 774-781. Talanian, R.V.; McKnight, C.J.; Kim, P.S. *Science* **1990**, *249*, 769-771.

12. Column: Vydac C18, 300 Å, 22 x 250 mm (The Separations Group). Mobile phase A: 99.5% H<sub>2</sub>O, 0.5% CF<sub>3</sub>COOH. Mobile phase B: CH<sub>3</sub>CN. Gradient: 0 to 100% B in 400 min. Retention time 85 min.
13. Expected: Asx<sub>1</sub>Thr<sub>1</sub>Ser<sub>1</sub>Glx<sub>2</sub>Ala<sub>5</sub>Leu<sub>2</sub>Lys<sub>2</sub>Arg<sub>6</sub>Cys<sub>1</sub>, found: Asx<sub>1,0</sub>Thr<sub>0,9</sub>Ser<sub>0,7</sub>Glx<sub>2,0</sub>Ala<sub>5,0</sub>Leu<sub>2,0</sub>Lys<sub>2,0</sub>Arg<sub>6,6</sub>Cys<sub>1,1</sub>
14. Calculated for [M+H] C<sub>99</sub>H<sub>185</sub>N<sub>44</sub>O<sub>27</sub>S: 2454.42; found, 2454.3 (*meta*-nitrobenzyl alcohol + trifluoroacetic acid matrix).
15. Potts, K.T.; Usifer, D.A.; Guadalupe, A.; Abruna, H.D. *J. Am. Chem. Soc.* **1987**, *109*, 3961-3967.
16. *1H-NMR* (250 MHz, CDCl<sub>3</sub>) δ 8.67 (d<sub>app</sub>, J=4.8 Hz, 2H), 8.59 (d<sub>app</sub>, J=8.0 Hz, 2H), 8.29 (s, 2H), 7.83 (dt, J=1.8, 7.8 Hz, 2H), 7.32 (ddd, J=1.2, 4.8, 7.5 Hz, 2H), 2.64 (s, 3H). *FT-IR* (neat, cm<sup>-1</sup>) 1578, 1559, 1541, 1465, 1391, 785. *MP*: 122-122.5 °C (lit<sup>15</sup> mp 118 °C)
17. Prepared by treatment of 3-bromopropanol with isobutylene and Amberlyst H-15. Alexakis, A.; Gardette, M.; Colin, S. *Tetrahedron Lett.* **1990**, *29*, 2951-2954.
18. Wenkert, *et al. J. Org. Chem.* **1985**, *50*, 1125-1126.
19. *1H-NMR* (250 MHz, CDCl<sub>3</sub>) δ 8.69 (d<sub>app</sub>, J=4.8 Hz, 2H), 8.62 (d<sub>app</sub>, J=8.0 Hz, 2H), 8.33 (s, 2H), 7.84 (dt, J=1.8, 7.8 Hz, 2H), 7.31 (ddd, J=1.2, 4.8, 7.5 Hz, 2H), 3.40 (t, J=6.3 Hz, 2H), 2.88 (t<sub>app</sub>, J=7.8 Hz, 2H), 2.02 (m, 2H), 1.18 (s, 9H); *FT-IR* (neat, cm<sup>-1</sup>) 2970, 2932, 2866, 1582, 1567, 1558, 1468, 1465, 1361, 1198, 1088, 1073, 792, 741; *HR-MS* (FAB, glycerol) M+H calculated 348.2075, observed 348.2081.
20. *1H-NMR* (250 MHz, CD<sub>3</sub>OD) δ 8.66 (d<sub>app</sub>, J=4.7 Hz, 2H), 8.60 (d<sub>app</sub>, J=8.0 Hz, 2H), 8.25 (s, 2H), 7.98 (dt, J=1.8, 7.8 Hz, 2H), 7.45 (ddd, J=1.2, 4.9, 7.6 Hz, 2H), 3.65 (t, J=6.4 Hz, 2H), 2.92 (t<sub>app</sub>, J=7.8 Hz, 2H), 2.00 (m, 2H); *FT-IR* (neat, cm<sup>-1</sup>) 3274 (b), 2918, 1584, 1564, 1472, 1404, 1073, 790; *HR-MS* (FAB, glycerol) M+H calculated 292.1449, observed 292.1461.
21. *1H-NMR* (250 MHz, CDCl<sub>3</sub>) δ 8.68 (d<sub>app</sub>, J=4.7 Hz, 2H), 8.60 (d<sub>app</sub>, J=8.0 Hz, 2H), 8.27 (s, 2H), 7.83 (dt, J=1.8, 7.8 Hz, 2H), 7.31 (ddd, J=1.1, 4.8, 7.5, 2H), 2.94 (t, J=7.2, 2H), 2.85 (t<sub>app</sub>, J=7.9 Hz, 2H), 2.32 (s, 3H), 2.0 (m, 2H); *FT-IR* (neat, cm<sup>-1</sup>) 2936, 1690, 1585, 1564, 1466, 1405, 1131, 789; *HR-MS* (FAB, glycerol) M+H calculated 350.1326, observed 350.1333.
22. Volante, R.P. *Tetrahedron Lett.* **1981**, *33*, 3119-3122. Mitsunobu, O. *Synthesis* **1981**, 1.
23. To a solution of thiolacetate **5** (40.3 mg, 115 μmol, 1 eq) and Aldrithiol<sup>TM</sup>-2 (30.4 mg, 138 μmol, 1.2 eq) in degassed methanol was added dropwise over 1 hr 138 μL of a 1 M aqueous LiOH (138 μmol, 1.2 eq) solution. After 3 h, the reaction mixture was concentrated and purified by chromatography on basic alumina (20% ethyl acetate in hexane) to yield 29 mg (60%) of disulfide **6**. *1H-NMR* (250 MHz, CDCl<sub>3</sub>) δ 8.68 (d<sub>app</sub>, J=3.7 Hz, 2H), 8.59 (d<sub>app</sub>, J=8.0 Hz, 2H), 8.43 (d<sub>app</sub>, J=6.4 Hz, 1H), 8.27 (s, 2H), 7.84 (dt, J=1.8, 7.8 Hz, 2H), 7.69 (d<sub>app</sub>, J=8 Hz, 1H), 7.60 (m, 1H), 7.31 (ddd, J=1.1, 4.8, 7.4 Hz, 2H), 7.02 (t<sub>app</sub>, J=6.6 Hz, 1H), 2.87 (m, 4H), 2.17 (q, J=7.4 Hz, 2H); *FT-IR* (neat, cm<sup>-1</sup>) 3049, 2922, 1582, 1569, 1464, 1442, 1414, 1400, 1111, 787, 668; *HR-MS* (FAB, glycerol) M+H calculated 417.1207, observed 417.1236.
24. Small molecular weight compounds were removed by size exclusion chromatography (Sephadex G-10 spin column) prior to HPLC purification. HPLC Column: Delta-Pak C18, 300 Å, 3.9 x 150 mm (Waters Associates). Flow rate: 1 mL min<sup>-1</sup>. Mobile phase A: 2% CH<sub>3</sub>CN, 0.06% CF<sub>3</sub>COOH, 97.94% H<sub>2</sub>O. B: 80% CH<sub>3</sub>CN + 20% A. Gradient: 90% A to 62% A in 30 min. Detection at 210 nm.
25. λ<sub>max</sub> 288 nm (10 mM phosphate pH 7.4).
26. Expected: Asx<sub>1</sub>Thr<sub>1</sub>Ser<sub>1</sub>Glx<sub>2</sub>Ala<sub>5</sub>Leu<sub>2</sub>Lys<sub>2</sub>Arg<sub>6</sub>, found: Asx<sub>0,9</sub>Thr<sub>0,9</sub>Ser<sub>0,9</sub>Glx<sub>0,9</sub>Ala<sub>4,3</sub>Leu<sub>1,9</sub>Lys<sub>1,9</sub>Arg<sub>6,5</sub> (Cys not determined).
27. Calculated for [M+H] C<sub>117</sub>H<sub>200</sub>N<sub>47</sub>O<sub>27</sub>S<sub>2</sub>: 2759.52; found, 2759.7 (*meta*-nitrobenzyl alcohol + trifluoroacetic acid matrix).
28. Treatment of the isolated product with 15 mM DTT (625 mM Tris-Cl, pH 6.8, 1% SDS, 15% glycerol) for 5 min at 25 °C caused it to co-migrate with authentic ΔGCN<sub>4</sub><sub>229-248</sub> when subject to denaturing gel electrophoresis.

(Received in USA 1 April 1991)