Pergamon

Bioorganic & Medicinal Chemistry Letters, Vol. 6, No. 16, pp. 1961-1966, 1996

0960-894X/96 $15.00 + 0.00

PII: S0960-894X(96)00351-4

RAPID ASSEMBLY OF "SUBTILIGASE" SUBSTRATES TO ELUCIDATE OPTIMAL LIGATION JUNCTURES

J. Kevin Judice,* Angela K. Namenuk, and John P. Burnier

Department of Bioorganic Chemistry, Genentech, Inc.
460 Point San Bruno Blvd., South San Francisco, CA 94080

Abstract: A method allowing for the rapid selection of optimal ligation junctures for the enzyme "subtiligase" is described. To expedite production of the glycolate ester substrate peptides for subtiligase, we have modified the protocols for compatibility with Fmoc chemistry. The utility of this approach is demonstrated in a model system.

Several innovative strategies for incorporating unnatural amino acids into proteins have been described over the past decade.[1] Such proteins have already been employed in the study of enzyme mechanism, protein stability, ion channel function, and other aspects of protein structure and function. The report by Wells and colleagues describing a subtilisin variant which catalyzes peptide segment condensation ("subtiligase")[1c] and its use in assembling mutants of RNase[2] has increased interest in this methodology, and encouraged us to undertake the studies described here.

Subtiligase catalyzes peptide bond formation between a peptide fragment with a free amino terminus and a glycolate acceptor peptide, as shown in Figure 1a. An integral aspect of synthesizing large peptides or proteins with subtiligase is the selection of junction points for linking peptide segments together. This process is essentially a simple retrosynthetic analysis, wherein particular regions of the linear peptide sequence are "disconnected" according to the likelihood that the forward segment condensation will be catalyzed by subtiligase. General guidelines have been developed by probing ligation efficiencies with combinations of P_X-$P_{X'}$ residues;[3] however, we have found that there is often subtlety in subsite interactions, making accurate selection of optimal ligation junctures difficult.

An illustrative case of subsite selection is shown in Figure 1b. This sequence corresponds to the DNA-encoded residues 1-43 from the coagulation protease Factor IX. Analysis of this peptide according to established guidelines[1c,3] allowed general areas to be targeted as probable ligation junctures, but many of these had "violations" of one sort or another and thus predicting the best of them *a priori* was not trivial. As such, we found it most efficient to synthesize short model peptides and evaluate their efficacy in subtiligase-catalyzed ligations prior to undertaking syntheses of longer (20-35 residue) fragments.

a. $...P_4\text{-}P_3\text{-}P_2\text{-}P_1$-glc-F-amide + $H_2N\text{-}P_1'\text{-}P_2'\text{-}P_3'\text{-}P_4'...$ ***subtiligase***

glycolate ester **nucleophile**

$...P_4\text{-}P_3\text{-}P_2\text{-}P_1\text{-}P_1'\text{-}P_2'\text{-}P_3'\text{-}P_4'...$

product

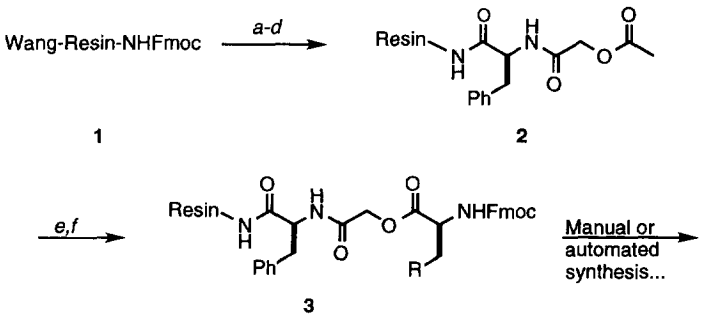
b. YNSGKLEEFVQGNLERECMEEKCSFEEAREVFENTERTTEFWK

Site 1 **Site 2** **Site 3**

Figure 1: ***a.*** *Subtiligase-catalyzed peptide bond formation. Subsites that may confer specificity to subtiligase are indicated. Numbering is analogous to that used in protease literature; residues P4 to P1 arise from the glycolate ester peptide while P_1' to P_4' come from the peptide containing the free α-amino group (nucleophile).* ***b.*** *The sequence of residues 1-43 of the coagulation protease Factor IX. Possible ligation juncture regions are indicated as sites 1-3.*

In order to expedite production of these model peptides, we have adapted the synthesis of glycolate esters to allow the use of Fmoc protocols. The previously reported method[1c,2] employs a Boc protecting group strategy. This method is effective, but the necessity of (a) conducting one step in the scheme at elevated temperature and (b) employing HF cleavage to liberate the resulting glycolate peptide led us to develop a more streamlined approach. Our modified approach is shown in Scheme 1. Rink resin[4] (**1**) is charged with Fmoc-Phe (or Fmoc-Lys),[5] deprotected under standard conditions, and the resulting free amine is coupled to acetoxyacetic acid to produce **2**. This resin was stable for at least a year at ambient temperature, allowing for bulk preparation and storage.

Scheme 1



Reagents: *(a) 20% piperidine in DMA. (b) Fmoc-Phe, BOP, NMM, DMA. (c) 20% piperidine/DMA. (d) acetoxyacetic acid, DIPC, CH_2Cl_2. (e) $N_2H_4 \cdot H_2O$, DMA. (f) DIPC, DMAP, Fmoc-AA, CH_2Cl_2.*

Synthesis of the tetramer glycolate series begins with deprotection of the acetoxy group with hydrazine in DMA overnight. After liberation of the glycolate hydroxyl group, the first amino acid is coupled using DIPC in CH_2Cl_2, with catalytic DMAP to form **3**.[6] The peptide is then extended using standard Fmoc chemistry protocols.[7] Yields of the resulting glycolate ester peptides following cleavage and purification were similar to those observed for the analogous peptides prepared by Boc chemistry,[8] indicating minimal loss of peptide from the resin due to ester hydrolysis or aminolysis.

A partial list of peptides used to examine ligation junctures in the Factor IX gla domain is shown in Table 1. The predicted results are compared to the actual data; in some cases, sites which predicted to work do not, and in others the reverse is true, emphasizing the need to explore subsite specificity with model peptides before undertaking long syntheses.

Site	(P4-P1) glycolate ester	(P1'-P4') nucleophile	Ligation Efficiency Prediction[a]	Result[b]
1	suc EFVQ glc F amide	GNLE	yes	+++
	suc VQGN glc F amide	LERE	yes?	-
	suc GNLE glc F amide	RECM	yes?	+
2	suc EEKC glc F amide	SFEE	yes	-
	suc EKCS glc F amide	FEEA	yes?	-
3	suc AREV glc F amide	FENT	no	-
	suc EARE glc F amide	VFEN	no	+++

Table 1: *Ligation reactions listed according to success observed against predicted results. (a) Predictions are based on previously published guidelines; "yes" signifies a potential ligation site with both P1 and P1' residues listed among those preferred by subtiligase and no other residues listed as disfavored; "yes?" signifies a site with P1 and P1' either preferred or of unknown tendency, with at least one other residue classified as disfavored; and "no" signifies a site with at least one violation at P1 or P1'. (b) Yields categorized as follows: "+++" indicates a yield of >50% product based on the nucleophile as limiting reagent; "+", product yield between 10-50%; and "-", no product detected. Ligations were carried out at ambient temperature in 100 mM tricine pH 8.0 with 10 mM $CaCl_2$ containing 4 mM glycolate ester peptide, 4 mM nucleophilic peptide, and 10 μM subtiligase.*

The ligation sites that score well in Table 1 also perform well in the context of 10-12 residue peptides (not shown). This has been true in a variety of cases: following identification of optimal ligation sites with this method, we have had good results in assembling longer peptides and small proteins with subtiligase. It is important to note, however, that the success of these tetrameric model peptides in subtiligase-catalyzed ligations is necessary but not sufficient to guarantee success when full-length fragments are condensed. We have found that ligation junctures successful in the context of tetramer probes occasionally fail in longer peptides (ca. 15-25 residues in length); the usual cause is insolubility of the longer fragments, but formation of secondary structures may also be problematic.[9]

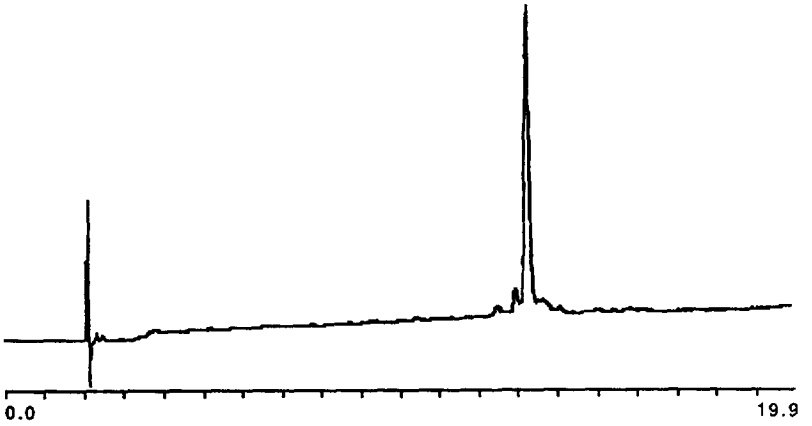


Figure 2: *HPLC trace of the crude material after synthesis of inoc-YNSGKLγγFVQ glc F-amide; "inoc" is isonicotinylcarbamate, and the "γ" symbol stands for γ-carboxyglutamic acid.*

In summary, the use of model peptides with subtiligase is strongly advised, and the modified chemistry reported here for production of glycolate esters should facilitate this. To date, chemistry described in Scheme 1 has been used to generate over 70 short glycolate peptides with a wide range of P1 residues, emphasizing the generality of the method. Although our main goal was the rapid production of these tetrameric species, we also synthesized several longer peptides to further explore the utility of this chemistry. Peptides of ten to twenty residues in length were generated in acceptable yields and with high fidelity (Figure 2). Thus, for the rapid generation of tetrameric species and the occasional production of longer peptides, the method described here is a significant advance in terms of rapidity and avoidance of non-standard solid phase conditions.[10]

Acknowledgments

We thank Dave Jackson and Jim Wells for discussions; and Dan Suich and Bill DeGrado for sharing their results prior to publication.

Experimental

Preparation of Glycolate Ester Resins A 10 g (0.5 meq/g, 5 mmol) portion of Rink resin (Advanced Chemtech) was swollen with 100 mL washes of CH_2Cl_2 and dimethylacetamide (DMA). The resin was treated with 20% piperidine in DMA for 15 min, then washed 5 times with DMA, once with CH_2Cl_2, and resuspended in 10 mL of CH_2Cl_2 . Fmoc-Phe (5.8 g, 15 mmol) was activated in a separate vessel with 6.6 g of benzotriazol-1-yloxy-tris(dimethylamino)phosphonium hexafluorophosphate (BOP; 15 mmol) and 3.3 mL (30 mmol) of N-methylmorpholine (NMM) in 10 mL DMA; after 10 min, this solution was added to the resin and the resulting suspension agitated for 1 h. After washing with DMA and CH_2Cl_2 as before, the resin was treated with 20% piperidine, washed as before, then resuspended in 20 mL CH_2Cl_2. To this was added a solution of 3.5 g (30 mmol) of acetoxy acetic acid in 10 mL of CH_2Cl_2, followed by 30 mL of a 1 M solution of 1,3-diisopropylcarbodiimide (DIPC) in CH_2Cl_2. After 30 min the resin was washed as before and resuspended in 40 mL DMA. To this was added 5 mL (100 mmol) of hydrazine monohydrate; this suspension was allowed to stand for 6-12 h; the resin was then washed as before and resuspended in CH_2Cl_2. To this was added a solution of the amino acid corresponding to the P1 residue (20 mmol, 4 equiv) in CH_2Cl_2, followed by 20 mL of 1 M DIPC in CH_2Cl_2 and 0.1 mol % DMAP for 1 h. Resin substitution was checked at this point.[8] Peptide synthesis was then continued under normal Fmoc chemistry conditions.[7]

References and Notes

1. (a) Schultz and colleagues have developed a method based on suppression of stop codons by chemically aminoacylated suppressor t-RNAs. This method has allowed single-site insertions of many different unnatural amino acids in a variety of proteins. For leading references, see Cornish, V. W.; Mendel, D.; Schultz, P. G. *Ang. Chem. Int. Ed. Eng.* **1995**, *34*, 621. (b) Kent and his co-workers have developed improved peptide synthesis protocols which allow for the direct assembly of small proteins by solid-phase peptide synthesis; the synthesis of HIV protease is illustrative: Milton, R. C.; Milton, S. C.; Kent, S. B. H. *Science* **1992**, *256*, 1445. They have more recently reported several segment condensation approaches; for example, see Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. H. *Science* **1994**, *266*, 776. Also see Fitzgerald, M. C.; Chernushevich, I.; Standing, K. G.; Kent, S. B. H.; Whitman, C. P. *J. Am. Chem. Soc.* **1995**, *117*, 11075, and references therein. (c) Wells and colleagues have developed an enzyme-catalyzed segment condensation approach based on a subtilisin variant ("subtiligase") first described in Abrahmsen, L.; Tom, J.; Burnier, J.; Butcher, K. A.; Kossiakoff, A.; Wells, J. A. *Biochemistry* **1991**, *30*, 4151.

2. Jackson, D. Y.; Burnier, J.; Quan, C.; Stanley, M.; Tom, J.; Wells, J.A. *Science* **1994**, *265*, 243.

3. Chang, T. K.; Jackson, D. Y.; Burnier, J.; Wells, J. A. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 12544.

4. Rink, H. *Tetrahedron Lett.* **1987**, *28*, 3787.

5. Previous reports[1c,2,3] have described the use of glycolate phenylalanyl amide substrates. We have found that the use of glycolate lysyl amide substrates provides a solubility advantage without

compromising efficiency in subtiligase-catalyzed bond formations. Jackson, D. Y.; Judice, J. K., unpublished results.

6. Racemization of the P_1 residue α-carbon is a potential problem under these conditions;[11] however, we do not generally observe any, as judged by the absence of diastereomeric peaks in HPLC/MS analyses. The exception is cysteine, where complete racemization was observed under these conditions (two closely spaced HPLC peaks of equal integration and identical mass spectra), consistent with the known propensity of activated cysteine esters to racemize. For peptides containing cysteine at the P1 site, we instead used Mitsunobu conditions (Richter, L. S.; Gadek, T. R. *Tetrahedron Lett.* **1994**, *35*, 4705) to accomplish the ester bond formation step. No racemization was observed under these conditions (data not shown). Analysis of the crystal structure of subtiligase[1c] suggests that D-residues will not be well tolerated at the P1 site. To probe this experimentally, we made a pair of peptides with identical sequences except for D- and L-residues at the P1 site; only the peptide with the L residue at position P1 was a substrate for subtiligase. This suggests that even if racemization occurs and is not detected, the probability of obtaining a ligated product that contains D-residues is low.

7. Either manual or automated synthesis could be used to extend the glycolate ester peptides after attachment of the first residue. Machine synthesis was carried out according to protocols described in the ABI 431A User's Guide (Applied Biosystems). Protocols for manual synthesis evolved over time, trending toward shorter coupling and deprotection times to allow more rapid synthesis; a good general review can be found in Fields, G. B.; Noble, R. L. *Int. J. Pept. Protein Res.* **1990**, *35*, 161.

8. After coupling of the Fmoc-Phe to the Rink resin, the resin substitution was determined spectrophotometrically by quantifying the fluorenylmethylpiperidine adduct formed upon removal of the Fmoc group; see Meienhofer, J.; Waki, M.; Heimer, E. P.; Lambros, T. J.; Makofske, R. C.; Chang, C.-D. *Int. J. Pept. Protein Res.* **1979**, *13*, 35. Peptides were cleaved from the resin using 95% trifluoroacetic acid/5% triisopropylsilane for 2 h at ambient temperature. Following removal of volatiles from the cleavage mixture, the peptides were triturated with diethyl ether, washed from the resin by rinsing with water, 10% acetic acid in water, and 100% acetic acid, then frozen and lyophilized. Yields of full-length peptides were determined after purification by reversed-phase HPLC and calculated as a percentage based on original resin substitution.

9. Jackson, D. Y.; Scheidig, A.; Judice, J. K., unpublished results. See also ref 10. Members of the Wells lab at Genentech are exploring solutions to this problem.

10. DeGrado and his coworkers have developed a related method for producing glycolate esters using Fmoc chemistry: Suich, D.J.; Ballinger, M.; Wells, J.A.; DeGrado, W.F.; *Tetrahedron Lett.*, in press.

11. Bodanszky, M. *Principles of Peptide Synthesis, 2nd ed.*; Springer-Verlag: New York; 1993, pp 179-185.

(*Received in USA* 29 *May* 1996; *accepted* 22 *July* 1996)