
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

A *de Novo* Designed Peptide Ligase: A Mechanistic Investigation

Alan J. Kennan,[†] V. Haridas, Kay Severin,[‡] David H. Lee, and M. Reza Ghadiri*

Contribution from the Departments of Chemistry and Molecular Biology and the Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, California 92037

Received April 20, 1999

Abstract: A 33-residue *de novo* designed peptide ligase is reported which catalyzes the template-directed condensation of suitably activated short peptides with catalytic efficiencies in excess of 10^5 ($[k_{\text{cat}}/K_{\text{m}}]/k_{\text{uncat}}$). The ligase peptide, derived from natural and designed α -helical coiled-coil proteins, presents a surface for substrate assembly via formation of a hydrophobic core at the peptide interface. Charged residues flanking the core provide additional binding specificity through electrostatic complementarity. Addition of the template to an equimolar fragment solution results in up to 4100-fold increases in initial reaction rates. Dramatic decreases in efficiency upon mutation of charged residues or increase in ionic strength establishes the importance of electrostatic recognition to ligase efficiency. Although most of the increase in reaction efficiency is due to entropic gain from binding of substrates in close proximity, mechanistic studies with altered substrates demonstrate that the system is highly sensitive to precise ordering at the point of ligation. Taken together these results represent the first example of a peptide catalyst with designed substrate binding sites which can significantly accelerate a bimolecular reaction and support the general viability of α -helical protein assemblies in artificial enzyme design.

The exquisite efficiency of nature's catalytic machinery has long been the envy of organic chemists. Although many natural enzymes have proven useful in synthetic applications,¹ expansion of this repertoire to include custom-tailored catalysts remains an area of active research. Study of re-engineered enzymes²⁻⁴ and catalytic antibodies⁵ has produced effective catalysts for even difficult chemical transformations. Yet despite

recent advances in the understanding of protein structure and function,^{6,7} *de novo* construction of synthetic peptide catalysts remains elusive.⁸ Benner and co-workers reported a designed helix bundle which catalyzed oxaloacetate decarboxylation with impressive efficiency.^{8c} In contrast, initially impressive reports of designed cyclic peptides with protease activity^{9,10} were later refuted or retracted.^{11,12} Other noteworthy systems exploit

[†] Present address: Department of Chemistry, Colorado State University, Fort Collins, CO 80523.

[‡] Present address: Department of Chemistry, University of Munich, Germany.

(1) Wong, C.-H.; Whitesides, G. M. *Enzymes in Synthetic Organic Chemistry*; Pergamon: New York, 1984.

(2) Douglas, K. T. *Curr. Opin. Biotechnol.* **1992**, *3*, 370-377.

(3) Abrahmsen, L.; Tom, J.; Burnier, J.; Butcher, K. A.; Kossiakoff, A.; Wells, J. A. *Biochemistry* **1991**, *30*, 4151-4159.

(4) Kaiser, E. T. *Angew. Chem., Int. Ed. Engl.* **1988**, *27*, 913-922.

(5) (a) Kirby, A. J. *Acta Chem. Scand.* **1996**, *50*, 203-210. (b) MacBeath, G.; Hilvert, D. *Chem. Biol.* **1996**, *3*, 433-445. (c) Schultz, P. G.; Lerner, R. A. *Science* **1995**, *269*, 1835-1842.

(6) Dahiyat, B. I.; Mayo, S. L. *Science* **1997**, *278*, 82-87.

(7) Bryson, J. W.; Betz, S. F.; Lu, H. S.; Suich, D. J.; Zhou, H. X.; O'Neil, K. T.; DeGrado, W. F. *Science* **1995**, *270*, 935-941.

(8) For reviews, see: (a) Corey, M. J.; Corey, E. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 11428-11434. (b) Baltzer, L. *Curr. Opin. Struct. Biol.* **1998**, *8*, 466-470. Some successful peptide catalysts: (c) Johnsson, K.; Allemann, R. K.; Widmer, H.; Benner, S. A. *Nature* **1993**, *365*, 530-532. (d) Broo, K. S.; Brive, L.; Ahlberg, P.; Baltzer, L. *J. Am. Chem. Soc.* **1997**, *119*, 11362-11372.

(9) Atassi, M. Z.; Manshouri, T. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 8282-8286.

(10) Hahn, K. W.; Klis, W. A.; Stewart, J. M. *Science* **1990**, *248*, 1544-1547.

prosthetic groups such as hemes, porphyrins, flavins, and pyridoxamines in order to elicit the corresponding activity in peptide-based scaffolds,¹³ although the resulting catalytic properties are often not appreciably distinct from those of the isolated prosthetic group. Most other reported peptide catalysts exhibit only small rate enhancements.

The dramatic rate accelerations exhibited by intramolecular reactions have been the subject of intense study for decades.¹⁴ Results from numerous investigations¹⁵ suggest that rate increases on the order of 10^8 are achievable. As a result, use of intramolecularity to accelerate reactions has been exploited in a wide variety of enzyme mimics.¹⁶ Cyclodextrins¹⁷ and other synthetic hosts^{16,18} capable of performing reactions on bound guests have exhibited rate increases of up to 10^6 in comparison to the uncatalyzed systems. More recently, these principles have been applied to the more challenging catalysis of bimolecular reactions, albeit with significantly smaller accelerations.^{16a} Although binding of two substrates to a template results in catalysis due to an increase in effective concentration, the accompanying intrinsic increase in binding affinity for covalently linked substrates typically results in severe product inhibition.^{16a} Sanders and Walter have described a cyclic porphyrin trimer that catalyzed a Diels–Alder reaction by a factor of 1000, but displayed no turnover.^{19a} Product inhibition was avoided in catalysis of an acyl transfer reaction by a similar template, but in this case the acceleration was only a factor of

11.^{19b} Rebek and co-workers investigated one host which catalyzes amide bond formation between its substrates by a factor of 160^{20a} and another which accelerates the Diels–Alder reaction between quinone and cyclohexadiene by 200-fold.^{20b} Terfort and von Kiedrowski reported a template which employed electrostatic recognition to catalyze imine formation with modest (<10-fold) rate increases.²¹ Kelly et al. prepared a template for S_N2 catalysis which avoided inhibition via product precipitation but exhibited only limited (6-fold) acceleration.²² Various examples of nucleic acid directed oligonucleotide condensations have also resulted in small (≤ 10) rate enhancements.²³ The system described here displays rate accelerations of up to 4100-fold in the templated-directed ligation of two peptides.

The *de novo* design of synthetic enzymes poses several challenges of escalating complexity. The primary barrier to design of peptide catalysts is creation of unique and predictable protein tertiary structures.²⁴ Protein-like scaffolds have been reported which serve as metal ion receptors,²⁵ heme and porphyrin binding agents,^{13c,26} and models for biological electron-transfer systems.²⁷ Despite these advances, the construction of catalysts possessing well-defined substrate binding sites, a further prerequisite for effective design, remains problematic, even using selection protocols from randomized peptide libraries.²⁸ Yet the formidable task of substrate binding pales in comparison with that of active site engineering. Precision-tuned enzyme active sites employ a variety of mechanisms to effect their catalytic efficiencies: approximation (proximity effects), orientation, strain, concerted functional group catalysis, solvent effects, etc.^{14,29–30} It is therefore unrealistic to expect that catalysts with enzyme-like active sites can be engineered from first principles, given the current level of design capabilities. Nonetheless, a rational hierarchical approach founded on simple and testable design concepts can

(11) Corey, D. R.; Phillips, M. A. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 4106–4109.

(12) (a) Wells, J. A.; Fairbrother, W. J.; Otlewski, J.; Laskowski, M., Jr.; Burnier, J. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 4110–4114. (b) Corey, M. J.; Hallakova, E.; Pugh, K.; Stewart, J. M., *Appl. Biochem. Biotechnol.* **47**, 199–210.

(13) (a) Kuang, H.; Distefano, M. D. *J. Am. Chem. Soc.* **1998**, *120*, 1072–1073. (b) Roy, R. S.; Imperiali, B. *Protein Eng.* **1997**, *10*, 691–698. (c) Robertson, D. E.; Farid, R. S.; Moser, C. C.; Urbauer, J. L.; Mulholland, S. E.; Pidikiti, R.; Lear, J. D.; Wand, A. J.; DeGrado, W. F.; Dutton, P. L. *Nature* **1994**, *368*, 425–432. (d) Mihara, H.; Tomizaki, K.; Nishino, N.; Fujimoto, T. *Chem. Lett.* **1993**, 1533–1536. (e) Sasaki, T.; Kaiser, E. T. *J. Am. Chem. Soc.* **1989**, *111*, 380–381. (f) Barbier, B.; Brack, A. *J. Am. Chem. Soc.* **1988**, *110*, 6880–6882. (g) Gutte, B.; Daeumigen, M.; Wittschieber, E. *Nature* **1979**, *281*, 650–655. (h) Chakravarty, P. K.; Mathur, K. B.; Dhar, M. M. *Indian J. Chem.* **1974**, *12*, 464–470.

(14) (a) Bruice, T. C.; Lightstone, F. C. *Acc. Chem. Res.* In press. (b) Houk, K. N.; Tucker, J. A.; Dorigo, A. E. *Acc. Chem. Res.* **1990**, *23*, 107–113. (c) Page, M. I.; Jencks, W. P. *Gazz. Chim. Ital.* **1987**, *117*, 455–460. (d) Menger, F. M. *Acc. Chem. Res.* **1985**, *18*, 128–134.

(15) (a) Jencks, W. P. *Adv. Enzymol.* **1975**, *43*, 219–410. (b) Page, M. I.; Jencks, W. P. *Proc. Natl. Acad. Sci. U.S.A.* **1971**, *68*, 1678–1683. (c) Bruice, T. C.; Pandit, U. K. *J. Am. Chem. Soc.* **1960**, *82*, 5858–5865. (d) Pandit, U. K.; Bruice, T. C. *J. Am. Chem. Soc.* **1960**, *82*, 3386–3390.

(16) (a) Kirby, A. J. *Angew. Chem., Int. Ed. Engl.* **1996**, *35*, 707–724. (b) Murakami, Y.; Kikuchi, J.-I.; Hisaeda, Y.; Hayashida, O. *Chem. Rev.* **1996**, *96*, 721–758.

(17) (a) Breslow, R. *Acc. Chem. Res.* **1995**, *28*, 146–153. (b) Tee, O. S. *Adv. Phys. Org. Chem.* **1994**, *29*, 1–85. (c) D'Souza, V. T.; Bender, M. L. *Acc. Chem. Res.* **1987**, *20*, 146–152. (d) Zhang, B.; Breslow, R. *J. Am. Chem. Soc.* **1997**, *119*, 1676–1681.

(18) (a) Asakawa, M.; Brown, C. L.; Menzer, S.; Raymo, F. M.; Stoddart, J. F.; Williams, D. J. *J. Am. Chem. Soc.* **1997**, *119*, 2614–2627. (b) Fan, E.; Vicent, C.; Hamilton, A. D. *New J. Chem.* **1997**, *21*, 81–85. (c) Kennan, A. J.; Whitlock, H. W. *J. Am. Chem. Soc.* **1996**, *118*, 3027–3028. (d) Muehe, M.-S.; Goebel, M. W. *Angew. Chem., Int. Ed. Engl.* **1996**, *35*, 2126–2129. (e) Mattei, P.; Diederich, F. *Angew. Chem., Int. Ed. Engl.* **1996**, *35*, 1341–1344. (f) Jubian, V.; Veronese, A.; Dixon, R. P.; Hamilton, A. D. *Angew. Chem., Int. Ed. Engl.* **1995**, *34*, 1237–1239. (g) Perrault, D. M.; Chen, X.; Anslyn, E. V. *Tetrahedron* **1995**, *51*, 353–362. (h) Pieters, R. J.; Huc, I.; Rebek, J., Jr. *Tetrahedron* **1995**, *51*, 485–498. (i) Smith, J.; Ariga, K.; Anslyn, E. V. *J. Am. Chem. Soc.* **1993**, *115*, 362–364. (j) McCurdy, A.; Jimenez, L.; Stauffer, D. A.; Dougherty, D. A. *J. Am. Chem. Soc.* **1992**, *114*, 10314–10321. (k) Zimmerman, S. C.; Zeng, Z.; Wu, W.; Reichert, D. E. *J. Am. Chem. Soc.* **1991**, *113*, 183–196. (l) Diederich, F.; Lutter, H. D. *J. Am. Chem. Soc.* **1989**, *111*, 8438–8446.

(19) (a) Walter, C. J.; Sanders, J. K. M. *Angew. Chem., Int. Ed. Engl.* **1995**, *34*, 217–219. (b) Mackay, L. G.; Wylie, R. S.; Sanders, J. K. M. *J. Am. Chem. Soc.* **1994**, *116*, 3141–3142.

(20) (a) Huc, I.; Pieters, R. J.; Rebek, J., Jr. *J. Am. Chem. Soc.* **1994**, *116*, 10296–10297. (b) Kang, J.; Rebek, J., Jr. *Nature* **1997**, *385*, 50–52.

(21) Terfort, A.; Kiedrowski, G. v. *Angew. Chem., Int. Ed. Engl.* **1992**, *31*, 654–656.

(22) Kelly, T. R.; Bridger, G. J.; Zhao, C. *J. Am. Chem. Soc.* **1990**, *112*, 8024–8034.

(23) (a) von Kiedrowski, G. *Angew. Chem., Int. Ed. Engl.* **1986**, *25*, 932–935. (b) Luo, P.; Leitzel, J. C.; Zhan, Z.-Y. J.; Lynn, D. G. *J. Am. Chem. Soc.* **1998**, *120*, 3019–3031. (c) Zhan, Z.-Y. J.; Lynn, D. G. *J. Am. Chem. Soc.* **1997**, *119*, 12420–12421.

(24) (a) Schneider, J. P.; Lombardi, A.; Degrado, W. F. *Folding Des.* **1998**, *3*, R29–R40. (b) Tuchscherer, G.; Scheibler, L.; Dumy, P.; Mutter, M. *Biopolymers* **1998**, *47*, 63–73. (c) Kohn, W. D.; Hodges, R. S. *Trends Biotechnol.* **1998**, *16*, 379–389.

(25) (a) Dieckmann, G. R.; McRorie, D. K.; Tierney, D. L.; Utschig, L. M.; Singer, C. P.; O'Halloran, T. V.; Penner-Hahn, J. E.; DeGrado, W. F.; Pecoraro, V. L. *J. Am. Chem. Soc.* **1997**, *119*, 6195–6196. (b) Klemba, M.; Regan, L. *Biochemistry* **1995**, *34*, 10094–10100. (c) Ghadiri, M. R.; Case, M. A. *Angew. Chem., Int. Ed. Engl.* **1993**, *32*, 1594–1597. (d) Handel, T. M.; Williams, S. A.; DeGrado, W. F. *Science* **1993**, *261*, 879–885. (e) Ghadiri, M. R.; Soares, C.; Choi, C. *J. Am. Chem. Soc.* **1992**, *114*, 4000–4002. (f) Regan, L.; Clarke, N. D. *Biochemistry* **1990**, *29*, 10878–10883.

(26) (a) Rojas, N. R. L.; Kamtekar, S.; Simons, C. T.; Mclean, J. E.; Vogel, K. M.; Spiro, T. G.; Farid, R. S.; Hecht, M. H. *Protein Sci.* **1997**, *6*, 2512–2524. (b) Choma, C. T.; Lear, J. D.; Nelson, M. J.; Dutton, P. L.; Robertson, D. E.; DeGrado, W. F. *J. Am. Chem. Soc.* **1994**, *116*, 856–85.

(27) (a) Sasaki, T.; Kaiser, E. T. *Biopolymers* **1990**, *29*, 79–88. (b) Mutz, M. W.; Case, M. A.; Wishart, J. F.; Ghadiri, M. R.; McLendon, G. L. *J. Am. Chem. Soc.* **1999**, *121*, 858–859. (c) Winkler, J. R.; Gray, H. B. *Chem. Rev.* **1992**, *92*, 369–379. (d) Isied, S. S.; Ogawa, M. Y.; Wishart, J. F. *Chem. Rev.* **1992**, *92*, 381–394.

(28) Ku, J.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 6552–6556.

(29) (a) Jencks, W. P. *Catalysis in Chemistry and Enzymology*; Dover: New York, 1987. (b) Fersht, A. *Enzyme Structure and Mechanism*; W. H. Freeman: New York, 1985. (c) Walsh, C. *Enzymatic Reaction Mechanisms*; W. H. Freeman: New York, 1979.

(30) (a) Kirby, A. J. *Adv. Phys. Org. Chem.* **1980**, *17*, 183–278. (b) Albery, W. J.; Knowles, J. R. *Angew. Chem., Int. Ed. Engl.* **1977**, *16*, 285–293. (c) Bruice, T. C. *Annu. Rev. Biochem.* **1976**, *45*, 331–373.

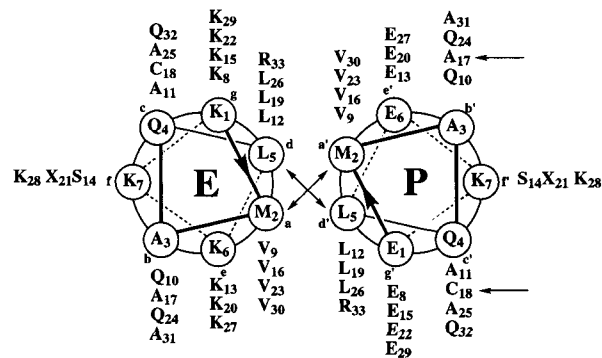
be exploited to evaluate and advance the scope and function of a given design. At present, the best approaches to rational design are those founded on well-characterized folding motifs which display sufficient kinetic and thermodynamic stability to reliably position substrate(s) and reactive functionalities.

In the present work we utilize one such well-known protein folding motif: the α -helical coiled coil (*vide infra*). Many elegant studies from laboratories of DeGrado,³¹ Hodges,³² Kim,³³ and others³⁴ have elucidated the influence of sequence variations on the structure, orientation, and aggregation state of coiled-coils. These advances, augmented by our own experiences in de novo design of artificial peptides and proteins,³⁵ have served as the design platform for the ligase system described here.³⁶

Design Principles

The ligase sequence is based on natural and designed coiled-coil proteins,^{33d–e,37} which have served as the framework for a number of synthetic architectures including our previous self-replicating peptide systems.³⁵ The coiled-coil³⁸ is a folding motif in which two or more α -helical peptides wrap around each other with a left-handed superhelical twist.³⁹ Coiled-coil sequences are characterized by a seven residue repeat, denoted (abcdefg)_n, in which the first and fourth (a and d) positions are occupied by hydrophobic residues whose side chains establish a densely packed hydrophobic core at the interface between helices.^{33f} The combination of hydrophobic and van der Waals interactions which comprise this “knobs-into-holes” packing³⁸ are the primary driving forces for dimerization. Charged residues in the fifth and seventh (e and g) positions of the heptad repeat provide additional specificity in the oligomerization, either through salt-bridge formation or electrostatic repulsion.^{32a,33b,e,40}

In light of the strong binding affinities of coiled-coil peptides, we reasoned that a template helix should prove capable of binding and preorganizing suitably complementary short peptide



E:	Ar- KMAQLKKKQVQALKSKVA -CLKXKQVQALKKKVAQR-CONH ₂
E*:	Ar- KMAQLKKKQVQALKSKVA -SLKXKQVQALKKKVAQR-CONH ₂
S ₁ :	Ar- EMAQLEKEVQALESEVA -COSBn
S ₂ :	Ar- EMAQLEKEVQALESEVA -COSCH ₂ CH ₂ CONH ₂
S ₃ :	Ar- EMAQLEKEVQALESEVA -CONHOH
S ₄ :	Ar- RMKQLEEKVYELLSKVA -COSBn
S ₅ :	Ar- KMAQLKKKQVQALKSKVA -COSBn
S ₆ :	Ar- KMAQLKKKQVQALKSKVP -COSBn
S ₇ :	H ₂ N- CLEXEVQALEKEVAQR -CONH ₂
S ₈ :	H ₂ N- CLEXEVARLKKLVGE -CONH ₂
S ₉ :	H ₂ N- GLEXEVQALEKEVAQR -CONH ₂
S ₁₀ :	H ₂ N-D-Cys LEXEVQALEKEVAQR -CONH ₂
S ₁₁ :	H ₂ N-Hcy LEXEVQALEKEVAQR -CONH ₂
S ₁₂ :	(CHO)HN- CLEXEVQALEKEVAQR -CONH ₂
S ₁₃ :	HSCH ₂ CH ₂ - LEXEVQALEKEVAQR -CONH ₂

Figure 1. Helical wheel representation of the catalyst–product dimer **E·P**, along with peptide sequences employed in these experiments. The complementary electrostatic interface formed by lysine and glutamic acid residues at the e and g positions (highlighted in bold) ensures specificity in the fragment assembly process (see text). To facilitate HPLC analysis 4-acetamidobenzoic acid (ABA) was coupled to the N-termini of **S**_{1–7} and **E** (denoted Ar) and to lysine₂₂ side chains of **S**_{7–13} and **E** (denoted X). Arrows indicate the position of the new Ala17–Cys18 amide bond in the product. Hcy = homocysteine.

fragments. Given appropriate chemical activation of the fragments, it seemed likely that their condensation should be catalyzed by the template, due to the increase in effective concentration in the ternary complex. This strategy, coupled with the ligation chemistry of Kent et al.⁴¹ (*vide infra*), had proven successful in our earlier studies³⁵ and more recently in other peptide replication systems.⁴² Furthermore, the incorporation of complementary charges at the e and g positions was expected to enhance substrate binding while disrupting the inhibitory homomeric assemblies.^{35b,c}

The designed catalyst–product dimer **E·P** is represented as a helical wheel in Figure 1. The specific sequences chosen combine the valine–leucine interface of the self-replicating systems^{35b} with the designed heterodimeric coiled-coil of Kim et al.^{33e} The e and g positions of the catalyst (**E**) are occupied by lysine, while those of the product (**P**) contain glutamic acids. These residues form an electrostatic recognition interface, composed of 9 pairs of matched charges, in which residue i in an e or g position contacts i′+5 or i′−5, respectively, on the opposite strand. The solvent-exposed residues were chosen to promote solubility and helix formation. The peptide bond between residues 17 and 18 of **P** was selected as the site for amide bond formation, since its location on the solvent-exposed face of the helix should permit ligation without disruption of the critical interface. In choosing among a variety of fragment

(40) (a) Monera, O. D.; Kay, C. M.; Hodges, R. S. *Biochemistry* **1994**, *33*, 3862–3871. (b) Graddis, T. J.; Myszkka, D. G.; Chaiken, I. M. *Biochemistry* **1993**, *32*, 12664–12671.

(41) Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. H. *Science* **1994**, *266*, 776–779.

(31) (a) Hill, R. B.; DeGrado, W. F. *J. Am. Chem. Soc.* **1998**, *120*, 1138–1145. (b) Boice, J. A.; Dieckmann, G. R.; DeGrado, W. F.; Fairman, R. *Biochemistry* **1996**, *35*, 14480–14485. (c) Lovejoy, B.; Choe, S.; Cascio, D.; McRorie, D. K.; DeGrado, W. F.; Eisenberg, D. *Science* **1993**, *259*, 1288–1293.

(32) (a) Zhou, N. E.; Kay, C. M.; Hodges, R. S. *J. Mol. Biol.* **1994**, *237*, 500–512. (b) Monera, O. D.; Zhou, N. E.; Kay, C. M.; Hodges, R. S. *J. Biol. Chem.* **1993**, *268*, 19218–19227. (c) Zhou, N. E.; Kay, C. M.; Hodges, R. S. *Biochemistry* **1992**, *31*, 5739–5746.

(33) (a) Eckert, D. M.; Malashkevich, V. N.; Kim, P. S. *J. Mol. Biol.* **1998**, *284*, 859–865. (b) Lumb, K. J.; Kim, P. S. *Science* **1995**, *268*, 436–439. (c) Harbury, P. B.; Kim, P. S.; Alber, T. *Nature* **1994**, *371*, 80–83. (d) Harbury, P. B.; Zhang, T.; Kim, P. S.; Alber, T. *Science* **1993**, *262*, 1401–1407. (e) O’Shea, E. K.; Lumb, K. J.; Kim, P. S. *Curr. Biol.* **1993**, *3*, 658–667. (f) Hu, J. C.; O’Shea, E. K.; Kim, P. S.; Sauer, R. T. *Science* **1990**, *250*, 1400–1403.

(34) (a) Cohen, C.; Parry, D. A. D. *Proteins: Struct., Funct., Genet.* **1990**, *7*, 1–15. (b) Wendt, H.; Berger, C.; Baici, A.; Thomas, R. M.; Bosshard, H. R. *Biochemistry* **1995**, *34*, 4097–4107. (c) Myszkka, D. G.; Chaiken, I. M. *Biochemistry* **1994**, *33*, 2363–2372. (d) Engel, M.; Williams, R. W.; Erickson, B. W. *Biochemistry* **1991**, *30*, 3161–3169. (e) Hecht, M. H.; Richardson, J. S.; Richardson, D. C.; Ogden, R. C. *Science* **1990**, *249*, 884–891.

(35) (a) Lee, D. H.; Severin, K.; Ghadiri, M. R. *Curr. Opin. Chem. Biol.* **1997**, *1*, 491–496. (b) Severin, K.; Lee, D. H.; Martinez, J. A.; Ghadiri, M. R. *Chem. Eur. J.* **1997**, *3*, 1017–1024. (c) Lee, D. H.; Granja, J. R.; Martinez, J. A.; Severin, K.; Ghadiri, M. R. *Nature* **1996**, *382*, 525–528. (d) Severin, K.; Lee, D. H.; Martinez, J. A.; Vieth, M.; Ghadiri, M. R. *Angew. Chem., Int. Ed. Engl.* **1998**, *37*, 126–128. (e) Lee, D. H.; Severin, K.; Yokobayashi, Y.; Ghadiri, M. R. *Nature* **1997**, *390*, 591–594.

(36) Severin, K.; Lee, D. H.; Kennan, A. J.; Ghadiri, M. R. *Nature* **1997**, *389*, 706–709.

(37) O’Shea, E. K.; Klemm, J. D.; Kim, P. S.; Alber, T. *Science* **1991**, *254*, 539–544.

(38) Crick, F. H. *Acta Crystallogr.* **1953**, *6*, 689–697.

(39) Harbury, P. B.; Plecs, J. J.; Tidor, B.; Alber, T.; Kim, P. S. *Science* **1998**, *282*, 1462–1467.

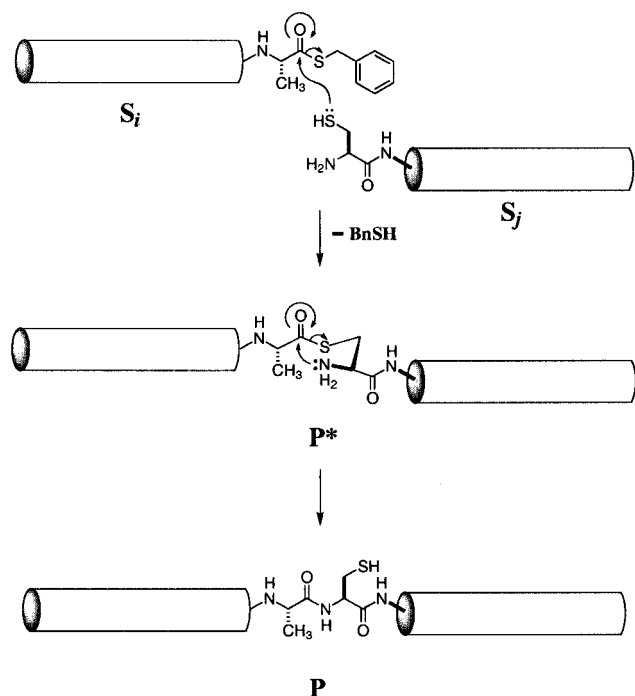


Figure 2. Fragment coupling by Kent ligation.⁴¹ The reaction proceeds via an initial trans-thioesterification between C-terminal thiolester of electrophile S_i and the N-terminal cysteine sulfhydryl group of nucleophile S_j . The resulting intermediate (P^*) rapidly rearranges via intramolecular $S \rightarrow N$ acyl transfer to give the final native peptide P .

coupling methods,^{41,43} we were guided by our previous success with the native chemical ligation strategy developed by Kent et al.⁴¹ (Figure 2). Thus, the C-terminal carboxylic acid of the electrophilic peptide S_1 was activated as the thiolbenzyl ester. Fragment condensation proceeds through an initial transthioesterification in which the thiolbenzyl group is displaced by the sulfhydryl group of the N-terminal cysteine on the nucleophilic fragment S_7 . The resulting intermediate (P^*) then undergoes an intramolecular $S \rightarrow N$ acyl shift to produce the desired amide bond in the product P .

Results

Ligation of S_1 and S_7 . The catalytic power of template E was determined by measuring initial rates of product formation in the reaction of equimolar solutions of S_1 and S_7 ($[S_1] = [S_7] = 200 \mu\text{M}$, pH 7.50) containing various amounts of E (11–430 μM). Rate enhancements up to 4100-fold (compared with the uncatalyzed coupling of S_1 and S_7) are observed, a level of efficiency which required experimentation at 5 °C to permit convenient kinetic monitoring. The reaction rates depend dramatically on the amount of E added (Figure 3a), reflecting catalyst participation in peptide fragment coupling. The kinetic efficiency of the ligation process was established by standard kinetic analysis in which the concentration of each substrate was varied independently, generating a series of Lineweaver–Burk plots (Figure 4). Second-order plots of slopes and intercepts⁴⁴ gave Michaelis constants of $78 \pm 29 \mu\text{M}$ for S_1

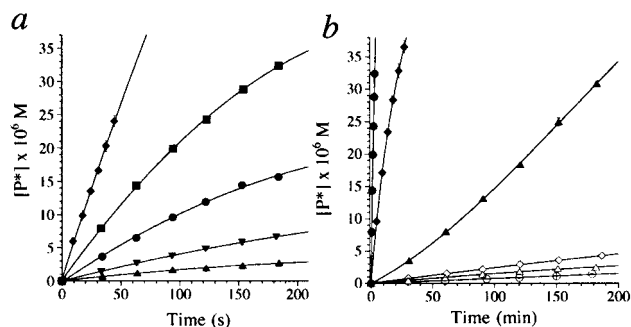


Figure 3. (a) Product formation (P^*) as a function of time for reaction mixtures initially containing 200 μM S_1 , 200 μM S_7 , and 11 μM (▲), 24 μM (▼), 53 μM (●), 104 μM (■), and 430 μM (◆) catalyst E , respectively. (b) Product formation ($P^* + P$) as a function of time for reactions containing equimolar mixtures of fragments: S_1 and S_7 (△), S_4 and S_7 (Δ), and S_1 and S_8 (◇) ($[S_x] = [S_y] = 200 \mu\text{M}$). In the presence of 50 mol % of catalyst, rate enhancements of 9.2-fold (S_4 and S_7 , ▲), 82-fold (S_1 and S_8 , ◇), and 1800-fold (S_1 and S_7 , ●) over the uncatalyzed reaction are observed. The curve which joins the points is intended merely as a visual guide.

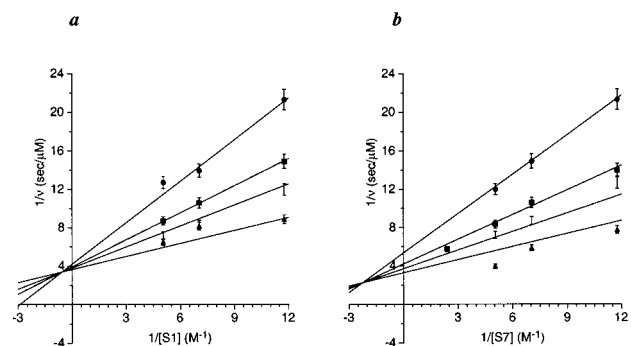


Figure 4. Lineweaver–Burk plots for the ligation of S_1 and S_7 by E . (a) Reactions with the concentration of S_7 held constant at (●) 85 μM , (■) 140 μM , (◆) 200 μM , and (▲) 415 μM while S_1 was varied from 85 to 305 μM . (b) Reactions in which the concentration of S_1 was held constant at (●) 85 μM , (■) 140 μM , (◆) 200 μM , and (▲) 305 μM while S_7 was varied from 85 to 415 μM . All reactions contained 20 μM E .

and $29 \pm 17 \mu\text{M}$ for S_7 , with a k_{cat} of $0.02 \pm 0.004 \text{ s}^{-1}$. Comparison⁴⁵ with the uncatalyzed process ($k_{\text{uncat}} = 0.003 \text{ M}^{-1} \text{ s}^{-1}$) reveals a catalytic efficiency ($[k_{\text{cat}}/K_m]/k_{\text{uncat}}$) of 7×10^5 , a value in good agreement with that obtained earlier from fitting reaction data to the kinetic model (Figure 5).³⁶ It should be noted that this calculation does not include any product release steps, and thus the value of k_{cat} is not precisely analogous to one observed for a real enzyme operating under conditions of multiple substrate turnover. Nonetheless, it serves as a convenient measure of the ligating abilities of template E .

Surprisingly, the $P^* \rightarrow P$ rearrangement is relatively slow ($t_{1/2} > 1 \text{ min}$), resulting in a buildup of the intermediate thiolester which is not observed in reactions of either the uncatalyzed system,^{41,46} similar models,^{29a,47} or most self-replicating peptides.^{35,42} Inhibition of the rearrangement is presumably due to the high stability and reduced conformational freedom of the thiolester in the $E \cdot P^*$ complex. The isolated intermediate P^* is stable under acidic conditions but spontaneously rearranges to

(45) The K_m of the weaker binding substrate is used for the efficiency calculation.

(46) Muir, T. W.; Dawson, P. E.; Fitzgerald, M. C.; Kent, S. B. H. *Chem. Biol.* **1996**, *3*, 817–825.

(47) (a) Bruce, T. R.; Fedor, L. R. *J. Am. Chem. Soc.* **1964**, *86*, 738–739. (b) Martin, R. B.; Hedrick, R. I. *J. Am. Chem. Soc.* **1962**, *84*, 106–110.

(42) (a) Yao, S.; Ghosh, I.; Zutshi, R.; Chmielewski, J. *Nature* **1998**, *396*, 447–450. (b) Yao, S.; Ghosh, I.; Zutshi, R.; Chmielewski, J. *Angew. Chem., Int. Ed.* **1998**, *37*, 478–481. (c) Yao, S.; Ghosh, I.; Zutshi, R.; Chmielewski, J. *J. Am. Chem. Soc.* **1997**, *119*, 10559–10560.

(43) (a) Tam, J. P.; Lu, Y.-A.; Liu, C.-F.; Shao, J. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 12485–12489. (b) Kemp, D. S.; Carey, R. I. *J. Org. Chem.* **1993**, *58*, 2216–2222.

(44) Cleland, W. W. *Steady-State Kinetics*; 3rd ed.; Boyer, P. D., Ed.; Academic Press: New York, 1970; Vol. 2, pp 1–65.

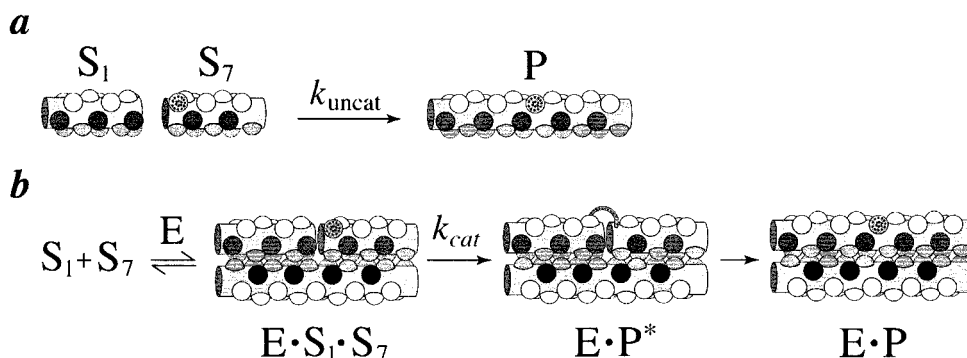


Figure 5. Reaction model. (a) The uncatalyzed background coupling of S_1 and S_7 to produce P proceeds with rate constant k_{uncat} (presumably via intermediate P^* which is not observed directly). (b) The catalytic process. Electrophilic (S_1) and nucleophilic (S_7) peptide fragments are bound by the electrostatically complementary catalyst E , forming the ternary complex $E \cdot S_1 \cdot S_7$ which facilitates ligation (k_{cat}) due to higher reactant effective molarities. Rearrangement of the resulting thioester complex $E \cdot P^*$ gives the heterodimer $E \cdot P$, which dissociates to regenerate free catalyst. Peptide backbones are shown as cylinders and side chains as spheres.

Table 1. Sequence Selectivity of Ligation

fragments	matched charges	k_{rel}^a
$S_1 + S_7$	9	1800 (4100) ^b
$S_1 + S_8$	7	82
$S_4 + S_7$	5	9.2
$S_5 + S_7$	4	1

^a Relative rates of P^* formation using 50 mol % of catalyst E .

^b Maximal rate enhancement in the presence of 215% catalyst.

P immediately upon exposure to MOPS buffer (pH 7.50). Reaction of P^* with a 0.5 M solution of hydroxylamine in MOPS buffer (pH 7.50) produces the hydrolyzed electrophile, the corresponding hydroxamic acid, and the nucleophile S_7 , further supporting the identity of the putative intermediate.

Importance of Electrostatic Complementarity. Charge matching of lysine and glutamic acid side chains in the $E \cdot P^*$ complex was designed to disfavor formation of unproductive homomeric assemblies through electrostatic repulsion, resulting in a high concentration of the productive monomeric species. Consequently a decrease in ligase efficiency (relative to background coupling) would be expected for systems in which one or more mismatches were introduced into this electrostatic recognition interface. The following experiments demonstrate that indeed virtually all of the catalytic advantage of the ligase can be erased by disruption of these interactions (Table 1, Figure 3b).

Ligation of an alternative peptide nucleophile (S_8) in place of S_7 , a substitution which removes two complementary electrostatic contacts, resulted in a 20-fold decrease in the initial rate of product formation. Use of an electrophilic fragment (S_4) with two more mismatches resulted in another order of magnitude decrease in initial velocity. When an electrophile was prepared in which *all* of the e and g residues were mutated to lysine, the fragment coupling rates were the same in the presence or absence of the catalyst E , indicating a total lack of catalysis. To verify that the critical interactions between these side chains were electrostatic in nature, the ligation of the native sequences ($[S_1] = [S_7] = 200 \mu\text{M}$, $[E] = 50 \mu\text{M}$) was carried out in buffer containing increasing amounts of NaCl. As the salt concentration was raised from 0 to 1 M, ligation rates dropped by a factor of 60 (Figure 6).

Mechanistic Studies

The observed remarkable efficiency of peptide catalyst E suggested that the reaction proceed via a highly preorganized catalyst–substrate complex. To gain additional insight into the

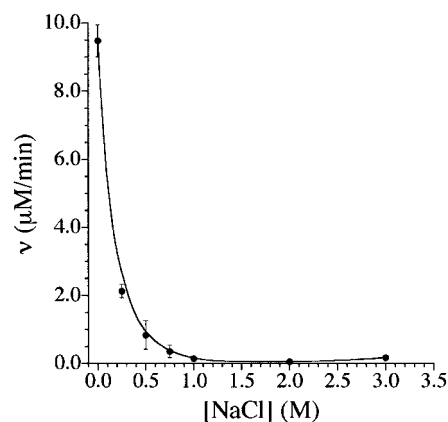


Figure 6. Initial rates of P^* formation as a function of NaCl concentration. Experiments were performed in MOPS buffer (pH 7.50) containing various concentrations of NaCl. All reactions were equimolar in S_1 and S_7 ($[S_1] = [S_7] = 200$) and contained $25 \mu\text{M}$ catalyst E . The curve which joins the points is intended merely as a visual guide.

source of catalysis, we sought to further characterize the functional requirements of the system by evaluating selected substrate variants. As expected, the use of less reactive electrophiles dramatically impacted reaction rates. Both mercapto-propionamide thioester substrate S_2 and C-terminal prolyl-thiolbenzyl ester S_6 proved 100-fold less reactive than the native sequence, while as expected the hydroxamic acid S_3 did not undergo the coupling reaction. Conversely, the reactivity of template E^* , in which the cysteine has been replaced by serine, was indistinguishable from that of E . This is consistent with the positioning of the cysteine on the solvent exposed face of the template, away from the fragment binding interface. The reactivity of the electrophilic fragments in template-directed ligation process was further scrutinized by studying five S_1 analogues that differed only in the identity of the thioester moiety (Figure 7). As expected, the reaction rate profiles displayed an inverse relationship to the $\text{p}K_{\text{a}}$ of the leaving group (Figure 7).

The catalytic process also proved to be very sensitive to the nature of the nucleophilic fragment (Table 2). The N-terminal glycine peptide S_9 did not react under the standard conditions, confirming the critical importance of the sulfhydryl functionality. The diastereoselectivity of the reaction was investigated via *in situ* competition between designed nucleophile S_7 and its epimer S_{10} (bearing an N-terminal D-cysteine) in the presence of equimolar S_1 ($[S_1] = [S_7] = [S_{10}] = 200 \mu\text{M}$). The pronounced selectivity (10:1) in favor of the native species reflects the

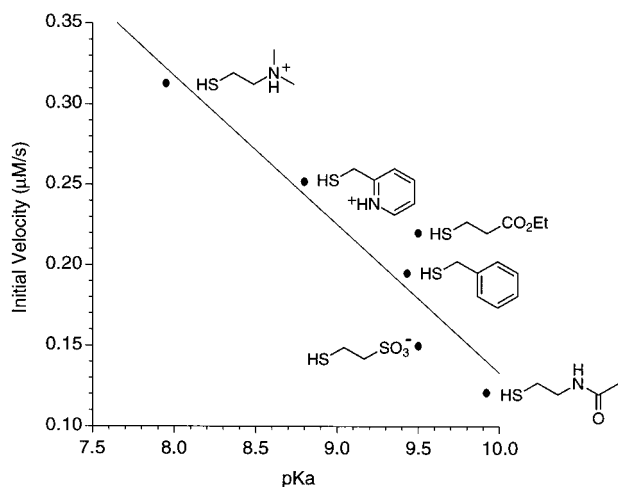


Figure 7. Initial rates of P^* formation vs pK_a of the mercaptan leaving group. Reactions were performed under standard ligation conditions (Experimental Section) except 5 mM TCEP was used as the reducing agent instead of an added mercaptan to minimize contributions to the rate profile from in situ thiol exchange reactions. All reactions were equimolar in S_1 analogs and S_7 ($[S_1] = [S_7] = 100 \mu\text{M}$) and contained $50 \mu\text{M}$ catalyst E .

Table 2. Cysteine Mutant Nucleophiles

Nucleophile	N-terminal residue	k_{rel}^a
S_7		120
S_{12}		60
S_{13}		40
S_{10}		12
S_{11}		1
S_9		0 ^b

^a Relative rates of P^* production at 5°C . ^b No product was observed after 24 h at 25°C .

stringency of conformational constraints at the point of ligation. As a further probe of functional group positioning, the ligation was performed using the homocysteine nucleophile S_{11} , in which an extra methylene unit has been inserted into the side chain. Even this seemingly modest alteration results in a 120-fold decrease in the coupling rate. When the experiment was repeated under denaturing conditions (MOPS buffer containing 6 M Gdn·HCl), the ligation rate of S_7 was only twice that of S_{11} , confirming that the two thiols have intrinsically similar reactivity. The presence of 6 M guanidinium also caused a 450-fold drop in the absolute reaction rate of the designed system (coupling of S_1 and S_7), which further emphasizes that efficient ligation demands properly folded peptides which permit as-

sociation of catalyst and substrate. The formation of helical assemblies such as $E \cdot P^*$ during the ligation process is in keeping with previous circular dichroism (CD) experiments which revealed an increase in helicity of both fragments and catalyst upon mixing.³⁶ The potential catalytic role of the N-terminal ammonium group was evaluated using nucleophiles in which the nitrogen was either formylated (S_{12}) or deleted (S_{13}). The small (2–3-fold) decrease in coupling rates with these fragments further support the notion that precise positioning of the sulfhydryl group is by far the most critical aspect of ligation site organization.

As mentioned above, the catalysis of condensation reactions is usually hampered by product inhibition. The present system case is no exception, and significant reductions in ligation rates are observed upon catalyst saturation. Nonetheless, the viability of catalyst turnover in principle was demonstrated by the reaction of electrophile S_2 with S_7 ($[S_2] = [S_7] = 200 \mu\text{M}$, pH 5.0). Experiments were conducted at room temperature and over a longer time course due to the decreased electrophilicity of the alkylthioester. In the presence of $18 \mu\text{M}$ catalyst, $27 \mu\text{M}$ of P was produced after 20 h, while only $1 \mu\text{M}$ was generated by the background reaction during the same time period.

Discussion

The principle mode of action employed by the ligase system described here is apparently a dramatic increase in effective fragment concentration which results from assembly onto the catalyst recognition surface. The rigidity of the resulting complex is reflected in the slow rearrangement of the intermediate thioester P^* —a reaction which is virtually instantaneous in the absence of the template peptide. Even more impressive are the dramatic rate decreases engendered by subtle modifications in the N-terminal cysteine residue of the nucleophile.

Incorporation of matched charges at the e and g positions of the $E \cdot S_1 \cdot S_7$ complex affords a secondary level of recognition. The dampening effect of NaCl supports the idea that the selectivity stems specifically from the charges on the side chains. The precipitous drop (over 3 orders of magnitude decrease in initial reaction rates) in relative rate upon successive deletion of these interactions points toward a cooperative destabilization in substrate binding, a point further underscored by the eradication of catalysis when one of the coupling fragments contains exclusively mismatched charges on these residues.

Originally we considered the possibility that the N-terminal ammonium group might be playing an active role in catalysis, perhaps as a general acid or through an electrostatic stabilization of a negatively charged transition state. To block these possible interactions with a minimum of structural perturbation, N-formyl nucleophile S_{12} was investigated. The minor impact of this modification upon the initial coupling rate suggests that the free N-terminus plays little or no chemical role in accelerating the transthioesterification reaction (Table 2). The possibility remained however that it might influence the torsional preferences of the side chain in a manner which would have the indirect effect of correctly positioning the key sulfhydryl group. The similar reactivity of nucleophile S_{13} , in which the amine is replaced with hydrogen, argues that such considerations play only a minor role at best (Table 2).

Since the catalytic efficiency of the present system is considerably better than those of our previously reported self-replicating peptides,^{35b} we had hoped that it might prove capable of catalyzing couplings involving amine nucleophiles. The obvious benefit to this would be the ability to ligate peptides with N-terminal residues other than cysteine. Unfortunately all

efforts at ligation with nucleophile **S**₉, in which the cysteine of **S**₇ is replaced by glycine, were unsuccessful. No ligated product was observed, even after extended reaction times or elevated (50 °C) temperature. Thus, further design modification will be required to produce a more general ligase system. Efforts targeted at both extension of suitable substrates and circumvention of product inhibition are currently underway.

Conclusion

The experiments described above document the design and evaluation of a synthetic peptide capable of ligating two shorter fragments with surprising efficiency and selectivity. The template presents a complementary surface upon which the substrates can assemble, and the corresponding increase in effective concentration results in initial rate increases of up to 4100-fold and catalytic efficiencies in excess of 10⁵. Incorporation of complementary charges in the e and g positions of the heptad repeat is shown to be crucial for optimal catalysis. Kinetic analysis reveals that the ligase efficiency exceeds those of recently described⁴⁸ antibodies (10⁵ versus 10²–10⁴), although it suffers from greater product inhibition and lags behind engineered subtilisin mutants.³ Mechanistic studies with altered fragments support the idea that the principal mode of acceleration is catalysis by approximation, in which reactive fragments are bound in close proximity. As in natural enzymes, catalysis is highly sensitive to the precise positioning of reactive functional groups, with dramatic drops in rate observed for changes as small as addition of a methylene group or inversion of single stereogenic center. Taken together, these results represent the first example of a peptide catalyst with designed substrate binding sites which can significantly accelerate a bimolecular reaction and support the viability of α -helical protein assemblies in artificial enzyme design.

Experimental Section

General. Acetonitrile (HPLC grade), dichloromethane (optima grade), dicyclohexylamine (DCHA), diethyl ether (anhydrous), dimethylformamide (sequencing grade), diisopropylethylamine (peptide synthesis grade), guanidinium hydrochloride (Gdn·HCl), 3-(*N*-morpholino)propanesulfonic acid (MOPS), and silica gel 60 (230–400 mesh) were purchased from Fisher and used without further purification. 4-Acetamidobenzoic acid (ABA, Aldrich), anisole (anhydrous, Aldrich), benzyl mercaptan (BnSH, Aldrich), tris(2-carboxyethyl)phosphine hydrochloride (TCEP, Fluka), deuteriochloroform (CDCl₃, 99.8%, Isotech), hydrogen fluoride (anhydrous, Matheson), 3-mercaptopropionic acid (Fluka), trifluoroacetic acid (TFA, New Jersey Halocarbon), and 2-(1-*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU, Richeleau Biotechnologies) were used without further purification. Commercially available *N*- α -*t*-Boc amino acids for solid-phase peptide synthesis and *p*-methylbenzhydrylamine (MBHA, substitution = 0.87–1.05 mequiv/g) resin were used as obtained from Novabiochem or Bachem. *N*- α -*t*-Boc-L-Ala-SCH₂CH₂CO₂H was pre-

pared as described previously.^{35b} 3-(9-Fluorenylmethyl)mercaptopropionic acid was obtained via the procedure of Woolley and co-workers⁴⁹ Electro spray mass spectrometry was performed on a Sciex API 3 or Sciex API 100 mass analyzer. ¹H NMR spectra were collected on a Bruker AM-250 spectrometer.

Peptide Synthesis. Manual Boc solid-phase peptide synthesis was carried out using MBHA resin on a 0.25–1.0 mmol scale according to the in situ neutralization protocol of Kent.⁵⁰ For electrophilic peptides *N*- α -*t*-Boc-L-Ala-SCH₂CH₂CO₂H was coupled to the resin using the same procedures. The spectroscopic label 4-acetamidobenzoic acid (ABA) was coupled to either the N-terminus of electrophiles or to Lys₂₂ (side chain nitrogen) of nucleophiles in the same fashion. Cleavage from the resin was accomplished using standard HF procedures (10 mL 9:1 HF/anisole per gram of peptidyl resin, 1 h 0 °C), and crude peptides were purified using C₁₈ reverse-phase HPLC with a binary gradient of 99% H₂O/CH₃CN/0.1% TFA (A), 90% CH₃CN/H₂O/0.07% TFA (B). Purity was confirmed by analytical HPLC and mass spectrometry (electrospray or MALDI).

Thiolbenzyl esters of substrates **S**₁–**S**₆ were prepared by transthiolesterification of corresponding thiopropionamide esters obtained directly from peptide synthesis in 6 M Gdn·HCl, 200 mM MOPS, pH 7.50 according to previously reported procedure.^{35b} Thiopropionate nucleophile **S**₁₃ was prepared by coupling 3-(9-fluorenylmethyl)mercaptopropionic acid in place of the N-terminal cysteine of **S**₇. The protecting group was removed by treatment with 20% piperidine/DMF for 20 min, followed by cleavage and purification as above. Catalyst **E** was synthesized by Kent ligation⁴¹ of the appropriate fragments. Both of these procedures have been previously reported.^{35b}

HPLC Analysis. Reverse-phase analytical HPLC was performed using a Zorbax C-8 300 SB column connected to a Hitachi D-7000 diode array HPLC system. Binary gradients of solvents A and B were employed at a flow rate of 1.5 mL/min, with monitoring at 270 nm. Peptide concentrations were determined by comparison to the internal ABA standard.

Ligation Experiments. Reactions were carried out in 0.6 mL Eppendorf tubes. A standard solution of 485.5 μ M ABA was prepared, and stock solutions of electrophiles, nucleophiles, and catalysts were prepared by dissolving the appropriate peptide in 1:1 MOPS:ABA solution, ABA solution, and deionized H₂O, respectively. In a typical experiment, the nucleophile and template were incubated with benzyl mercaptan in degassed MOPS buffer (pH 7.50) at room temperature for 10–15 min to reduce disulfides. The reaction mixture and the electrophile stock were then equilibrated to 5 °C for 30 min, whereupon reaction was initiated by addition of the electrophile. Representative final concentrations of all species were [**S**₁] = [**S**₇] = 200 μ M, [**E**] = 30–100 μ M, [MOPS] = 100 mM, [BnSH] \approx 1.5 mM, [ABA] = 100 μ M. Final reaction volumes were 250–350 μ L, from which 40–50 μ L aliquots were removed and quenched immediately with 5% TFA (50 μ L). Samples were frozen at –78 °C prior to HPLC analysis. Reactions under denaturing conditions were carried out in the same fashion using MOPS buffer which contained 6 M Gdn·HCl.

Acknowledgment. We would like to thank our colleague Alan Saghatelian for his generous assistance in data analysis and preparation of this manuscript and the National Institutes of Health (GM-57690) for financial support.

JA991266C

(49) Foong, L. Y.; You, S.; Jaikaran, D. C. J.; Zhang, Z.; Zunic, V.; Woolley, G. A. *Biochemistry* **1997**, *36*, 1343–1348.

(50) Schnoelzer, M.; Alewood, P.; Jones, A.; Alewood, D.; Kent, S. B. H. *Int. J. Pept. Protein Res.* **1992**, *40*, 180–193.

(48) (a) Hirschmann, R.; Smith, A. B., III; Taylor, C. M.; Benkovic, P. A.; Taylor, S. D.; Yager, K. M.; Sprengeler, P. A.; Benkovic, S. J. *Science* **1994**, *265*, 234–237. (b) Smithrud, D. B.; Benkovic, P. A.; Benkovic, S. J.; Taylor, C. M.; Yager, K. M.; Witherington, J.; Philips, B. W.; Sprengeler, P. A.; Smith, A. B., III; Hirschmann, R. *J. Am. Chem. Soc.* **1997**, *119*, 278–282. (c) Jacobsen, J. R.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 5888–5892.