Protein Ligations

DOI: 10.1002/anie.201200907

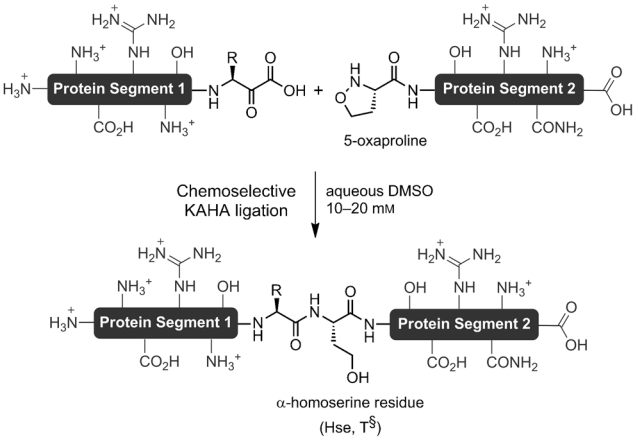
Chemical Protein Synthesis by Chemoselective α-Ketoacid–Hydroxylamine (KAHA) Ligations with 5-Oxaproline**

*Vijaya R. Pattabiraman, Ayodele O. Ogunkoya, and Jeffrey W. Bode**

Total chemical synthesis is an important method for preparing proteins of biological interest and makes possible the incorporation of unnatural amino acids and site-specific modifications.[1] Currently, chemical protein synthesis is served almost exclusively by the remarkable native chemical ligation (NCL),[2] which allows unprotected peptide segments to be chemoselectively coupled under aqueous conditions. The requirements of the NCL, an N-terminal cysteine or cysteine-surrogate and C-terminal thioesters, have encouraged a search for alternative amide-forming ligation reactions.[3] Although several mechanistically unique amide formations[4] have emerged from these studies, none has yet proven to be readily applicable for protein synthesis by the combination of unprotected segments.

As part of our efforts to develop a general peptide-forming ligation we reported that peptide α-ketoacids and N-terminal hydroxylamines undergo chemoselective, reagentless couplings to give backbone amide bonds (KAHA ligation).[5] We have demonstrated that the KAHA ligation is suitable for the synthesis of medium-sized peptides by ligations of unprotected segments, but we have struggled to perform the reactions in aqueous media typically employed for peptide solubilization and handling.[6] Mechanistic studies of the ligation of α-ketoacids and O-unsubstituted hydroxylamines[7] revealed a complicated pathway that is likely disturbed by water. In contrast, the ligations of O-Bz hydroxylamines give clean and rapid amide-forming ligations in water but are not suitable for α-peptide-derived substrates owing to facile elimination.[8] These observations led us to search for O-substituted hydroxylamines that are stable to the conditions for ligation and solid-phase peptide synthesis while at the same time sufficiently reactive in ligations performed in aqueous solvent.

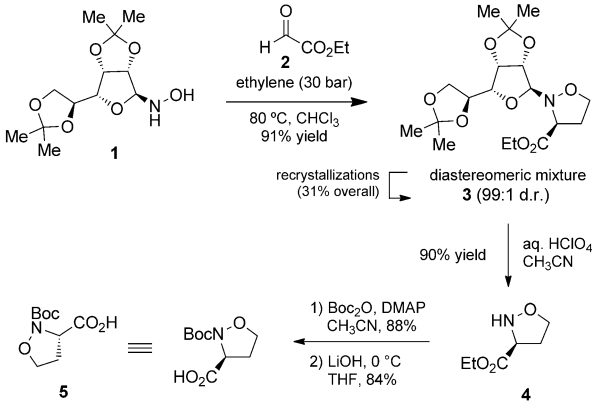
We are now pleased to report that 5-oxaproline is a stable, easily prepared and incorporated *N*-hydroxyamino acid that gives clean ligations in the presence of water, thereby leading to a homoserine residue at the ligation site (Scheme 1). We



Scheme 1. The α-ketoacid–5-oxaproline ligation. R = aminoacid side-chain, DMSO = dimethylsulfoxide.

demonstrate utility of KAHA ligation with 5-oxaproline by the total chemical synthesis of two small proteins from *Mycobacterium*, the prokaryotic-ubiquitin-like protein (Pup, 63 residues) and one of its target proteins, probable cold shock protein A (cspA, 66 residues), by the ligation of unprotected peptide segments prepared by Fmoc solid-phase peptide synthesis (SPPS).

Vasella et al. have reported an asymmetric synthesis of (*S*)-5-oxaproline.[9] Using a modified version of this procedure, we prepared compound **3** by cycloaddition of ethylene and the nitrone formed from L-gulose hydroxylamine **1** and ethyl glyoxalate (**2**) at 30 bar (Scheme 2).[10] The cycloadduct **3** was obtained as a 7:3 mixture of diastereomers, which gave a single stereoisomer after two recrystallizations from hexanes. The relative and absolute configuration of **3** was



Scheme 2. Synthesis of enantiopure (*S*)-*N*-Boc-5-oxaproline **5**. DMAP = 4-dimethylaminopyridine, Boc = *tert*-butoxycarbonyl.

[*] Dr. V. R. Pattabiraman, A. O. Ogunkoya, Prof. Dr. J. W. Bode
Laboratorium für Organische Chemie
Department of Chemistry and Applied Biosciences, ETH Zürich
Wolfgang Pauli Strasse 10, 8093 Zürich (Switzerland)
E-mail: bode@org.chem.ethz.ch
Homepage: http://www.bode.ethz.ch

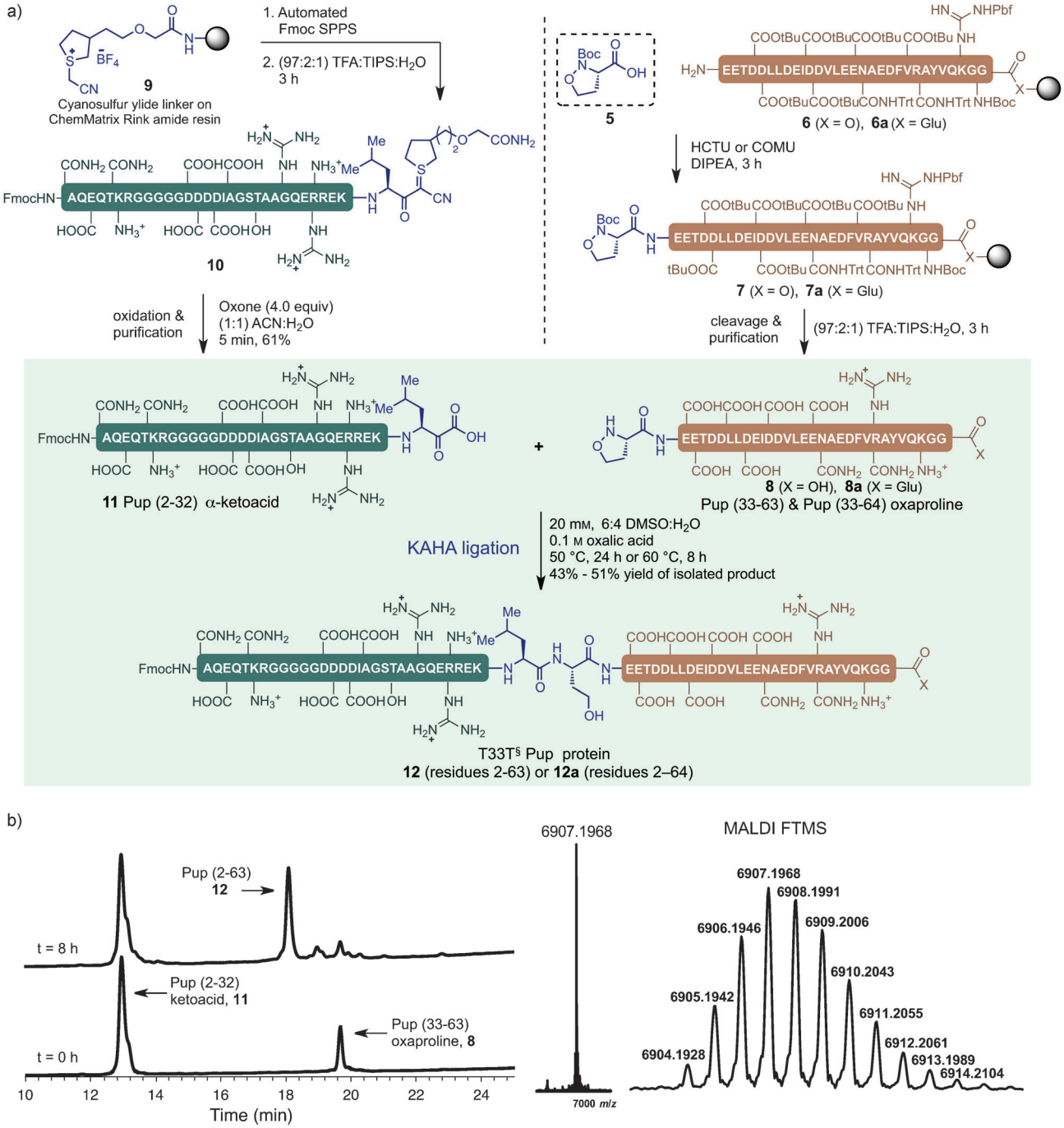
[**] This work was supported by the Swiss National Science Foundation (200021_131957) and an ETH Fellowship to V.R.P. (ETH FEL 06-10-1). We are grateful to Dr. Xiangyang Zhang and Rolf Häfliger of the LOC Mass Spectrometry Service for analyses and Prof. Andrea Vasella for helpful discussions.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201200907.

assigned by X-ray crystallographic analysis. Removal of the chiral auxiliary gave **4**, which was Boc-protected and saponified to provide (*S*)-*N*-Boc-5-oxaproline **5**, ready for use in SPPS.

As a first test of the oxaproline ligation for chemical protein synthesis, we targeted the preparation of Pup, which is a modifier protein from *Mycobacterium tuberculosis* that fills the role of ubiquitin in prokaryotes.[11] For Pup protein, we chose Leu (residue 32) and Thr (residue 33) as the ligation site. This disconnection requires two modified peptide segments (ca. 30 residues each), which we anticipated to be readily synthesized by Fmoc SPPS. In the ligated protein, Thr (residue 33) would be substituted with its structural isomer homoserine (Hse, T[§]).

The preparation of the oxaproline segment of Pup (33–63) **8** began with automated SPPS to give peptide **6** on-resin with side-chain protecting groups and an N-terminal free amine group (Scheme 3 a, right). At this stage, oxaproline monomer



Scheme 3. a) Synthesis of Pup (33–63) **8** and Pup (33–64) **8a** oxaproline segments & Pup (2–32)-α-ketoacid segment **11** and KAHA ligation to give synthetic Pup proteins **12** and **12a**. b) HPLC traces of the Pup ligation and high-resolution mass spectrum of the synthetic Pup protein (2–63). TFA = trifuluoroacetic acid, TIPS = triisopropylsilane, Fmoc = 9-fluorenylmethoxycarbonyl, Trt = triphenylmethyl, Pbf = 2,2,4,6,7-pentamethyldihydro-1-benzofurane-5-sulfonyl, HCTU and COMU are uronium-type coupling reagents.

5 was coupled manually by using HCTU as the coupling reagent for three hours to give **7**. Side-chain deprotection and cleavage of the peptide from the resin and purification proceeded smoothly to yield Pup (33–63) **8**. No special handling procedures were required for the coupling, TFA cleavage, or purification of oxaproline-containing peptide **8**. These results demonstrate that 5-oxaproline behaves like canonical amino acids during SPPS.

The Pup (2–32)-α-ketoacid peptide **11** was prepared by using our previously established procedure for the Fmoc SPPS of peptide α-ketoacids.[12] The cyanosulfurylide linker was loaded on ChemMatrix Rink amide resin to give **9** (Scheme 3a, left). Chain extension of Pup residues (2–32) on an automated peptide synthesizer was uneventful and pure cyanosulfurylide peptide **10** was obtained after treating the resin with TFA and HPLC purification. Oxidation of Pup (2–32) cyanosulfurylide **10** with Oxone gave peptide α-ketoacid **11**, which was isolated by preparative HPLC in 61% yield. The chromophoric Fmoc group on the N terminus was left intact to aid in ligation reaction monitoring.

With both peptide segments in hand, we undertook a brief survey of reaction conditions for the α-ketoacid–oxaproline ligation to form Pup protein **12** (Table 1). Ligations were performed with 1.0 equiv oxaproline **8** and 1.5 equiv α-ketoacid segment **11** in DMSO:water with oxalic acid (0.1M). While not necessary, we have observed that oxalic acid has a beneficial effect for both peptide solubilization and KAHA ligation. In each experiment an aliquot was taken at four hours and analyzed by HPLC.

Table 1: Survey of conditions for KAHA ligations to synthesize Pup (2–63) from α-ketoacid **11** and oxaproline peptide **8**.

Entry	Solvent[a]	Ligated Pup protein **12** at 4 h [%][b]		
		a) 60 °C	**b**) 50 °C	**c**) 40 °C
1	DMSO	9%	–[c]	–
2	9:1 DMSO:H_2O	56%	8%	–
3	8:2 DMSO:H_2O	70%	–	–
4	6:4 DMSO:H_2O	76%	27%	20%

[a] All ligation reactions were performed with 1.5 equiv Pup (2–32)-α-ketoacid **11** and 1.0 equiv Pup (33–63) oxaproline **8** at 20 mM with 0.1 M oxalic acid. [b] % conversion determined by analytical HPLC. [c] Condition not examined.

Water had a dramatic influence on the rate of the ligations; increasing amounts of water led to an increased amount of ligated Pup (2–63) protein **12** (Table 1). The ligations were faster at higher temperatures. We also observed that increasing the amount of α-ketoacid to 2.0 equivalents reduced the reaction times. Importantly, unprotected asparagine, glutamine, lysine, tryptophan, tyrosine, serine, and threonine side-chain functional groups did not interfere with the ligation. In all the reactions studied the ligated product was obtained, thus suggesting that the KAHA ligation with 5-oxaproline can be performed under a variety of conditions as needed for the protein of interest.

We chose the 6:4 DMSO:water mixtures containing oxalic acid (0.1M) for ligations to give Pup on a preparative scale. Pup (2–63) **12** was prepared by a clean and rapid ligation at 60 °C for eight hours (Scheme 3b). Purification of the ligation mixture by HPLC provided Pup (2–63) **12** in 51% yield. High-resolution mass spectrometry (FTMS) measurement and MS/MS analysis confirmed the exact mass and sequence identity of the synthetic Pup protein (2–63). Neither epimerization nor asparagine hydrolysis were detected despite the slightly elevated reaction temperature. Lower temperatures were also viable for preparative ligation, and we prepared Pup protein (2–64) **12a**, which contains the C-terminal glutamic acid that is necessary for linking of Pup protein to any target protein for proteasomal degradation,[13] by ligations at 50 °C for 24 h. The ligated Pup protein (2–64) **12a** was isolated in 43% yield after HPLC purification (see the Supporting Information).

To confirm that synthetic Pup protein (2–63) **12** adopts the expected tertiary structure, a CD spectrum was recorded in 40% trifluoroethanol in water. Synthetic Pup **12** showed the presence of α-helical regions, which are proposed for the natural Pup protein (Figure 1).[14] This finding implies that the incorporation of homoserine in place of threonine does not influence the ability of the Pup protein to form the expected secondary structures.

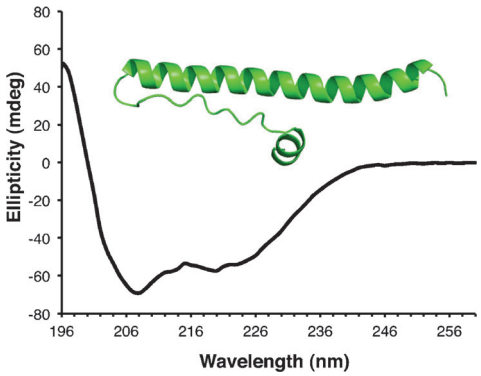
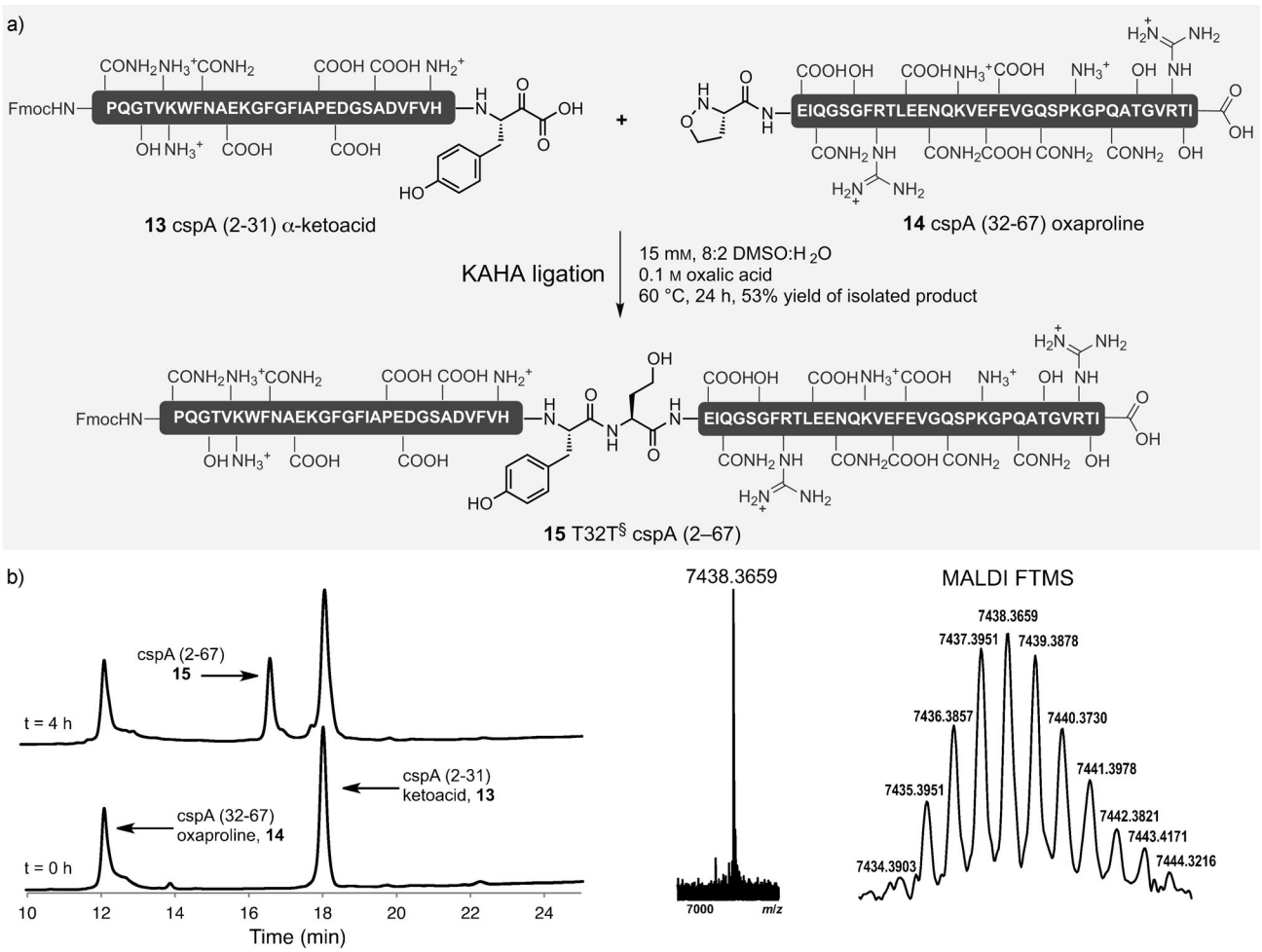


Figure 1. CD spectrum of the synthetic Pup protein **12** in 40% trifluoroethanol:water.

Having successfully synthesized Pup, we sought to apply the α-ketoacid–oxaproline ligation to a completely different sequence. We chose probable cold shock protein A (cspA), a 67-residue protein identified as a target protein for pupylation in *Mycobacterium*.[15] By using an analogous approach to the peptide-segment synthesis, cspA (2–31)-α-ketoacid **13** and cspA (32–67)-oxaproline **14** were prepared by automated SPPS and purified by HPLC (Scheme 4). For the ligation reaction of cspA segments **13** and **14**, we selected 8:2 DMSO:H_2O for optimal solubility. The ligation was performed at 15 mM in the presence of oxalic acid (0.1M) at 60 °C. On our first attempt, without any optimization, ligated T32T[§] cspA (2–67) **15** was isolated in 53% yield. Detailed MS/MS and FTMS analysis confirmed the sequence identity and exact mass of the synthetic cspA protein (2–67) T32T[§] with an N-terminal Fmoc group. The CD spectrum of synthetic cspA protein compared well with a literature report.[16]

The N-terminal oxaproline is a significant advance over our prior use of *N*-hydroxyamino acids for the KAHA ligation. The *N*-hydroxyamino acids could not be incorpo-



Scheme 4. a) Synthesis of T32T§ cspA (2–67) **15** by KAHA ligation of cspA (32–67) oxaproline segment **14** & cspA (2–31)-α-ketoacid segment **13**. b) HPLC traces of cspA ligation and high-resolution mass spectrum of T32T§ cspA (2–67) **15**.

rated into peptides by using standard N-protecting groups and were often unstable to purification or the ligation itself.[6b] Furthermore, peptide ligations with O-unsubstituted hydroxylamines did not tolerate aqueous conditions. The 5-oxaproline leads to homoserine residue, a non-natural amino acid, at the ligation site. In most cases, we anticipate that the homoserine residue will serve as a surrogate for threonine, serine,[17] methionine, aspartic acid,[18] and asparagine residues with minimal effect on protein structure and function. The remarkable chemoselectivity and reactivity of the oxaproline ligation likely arises from the relief of the ring strain of the 5-membered oxaproline. The design of other cyclic hydroxylamines for the KAHA ligation should allow for the preparation of synthetic proteins with both natural and unnatural residues at the ligation site.

In conclusion we have developed a variant of the KAHA ligation using 5-oxaproline that is suitable for protein chemical synthesis. The (*S*)-*N*-Boc-5-oxaproline monomer is readily prepared in enantiomerically pure form in multigram quantities by a short synthetic sequence. The 5-oxaproline residue was easily incorporated into peptides and was stable to standard deprotection and handling conditions. Backbone amide-forming ligations with an α-ketoacid peptide were clean and efficient, even in the presence of unprotected peptide side chains. The ligation was demonstrated by the synthesis of two proteins, Pup (63 residues) and cspA (66 residues). The KAHA ligation with 5-oxaproline promises to be the first general method for native amide-bond forming protein ligations, other than NCL, for the chemical synthesis of proteins, glycoproteins, therapeutic peptides, and other structures of biological interest by the chemoselective couplings of unprotected segments.

Received: February 2, 2012
Published online: March 14, 2012

Keywords: amides · ligation reactions · peptides · protein synthesis · proteins

[1] a) S. B. H. Kent, *Chem. Soc. Rev.* **2009**, *38*, 338–351; b) B. L. Nilsson, M. B. Soellner, R. T. Raines in *Annual Review of Biophysics and Biomolecular Structure*, *Vol. 34*, Annual Reviews, Palo Alto, **2005**, pp. 91–118.

[2] a) P. E. Dawson, T. W. Muir, I. Clarklewis, S. B. H. Kent, *Science* **1994**, *266*, 776–779; b) P. E. Dawson, *Isr. J. Chem.* **2011**, *51*, 862–867; c) S. N. Bavikar, L. Spasser, M. Haj-Yahya, S. V. Karthi-

keyan, T. Moyal, K. S. Ajish Kumar, A. Brik, *Angew. Chem.* **2012**, *124*, 782 – 787; *Angew. Chem. Int. Ed.* **2012**, *51*, 758 – 763; d) K. Hirano, D. Macmillan, K. Tezuka, T. Tsuji, Y. Kajihara, *Angew. Chem.* **2009**, *121*, 9721 – 9724; *Angew. Chem. Int. Ed.* **2009**, *48*, 9557 – 9560.

[3] a) B. L. Nilsson, L. L. Kiessling, R. T. Raines, *Org. Lett.* **2000**, *2*, 1939 – 1941; b) E. Saxon, J. I. Armstrong, C. R. Bertozzi, *Org. Lett.* **2000**, *2*, 2141 – 2143; c) R. Kleineweischede, C. P. R. Hackenberger, *Angew. Chem.* **2008**, *120*, 6073 – 6077; *Angew. Chem. Int. Ed.* **2008**, *47*, 5984 – 5988; d) J. Offer, *Pept. Sci.* **2010**, *94*, 530 – 541; e) A. Brik, C. H. Wong, *Chem. Eur. J.* **2007**, *13*, 5670 – 5675; f) M.-Y. Lutsky, N. Nepomniaschiy, A. Brik, *Chem. Commun.* **2008**, 1229 – 1231; g) R. Okamoto, S. Souma, Y. Kajihara, *J. Org. Chem.* **2009**, *74*, 2494 – 2501; h) F. B. Dyer, C.-M. Park, R. Joseph, P. Garner, *J. Am. Chem. Soc.* **2011**, *133*, 20033 – 20035; i) G.-M. Fang, Y.-M. Li, F. Shen, Y.-C. Huang, J.-B. Li, Y. Lin, H.-K. Cui, L. Liu, *Angew. Chem.* **2011**, *123*, 7787 – 7791; *Angew. Chem. Int. Ed.* **2011**, *50*, 7645 – 7649; j) P. Wang, S. J. Danishefsky, *J. Am. Chem. Soc.* **2010**, *132*, 17045 – 17051; k) C. P. R. Hackenberger, D. Schwarzer, *Angew. Chem.* **2008**, *120*, 10182 – 10228; *Angew. Chem. Int. Ed.* **2008**, *47*, 10030 – 10074.

[4] a) V. R. Pattabiraman, J. W. Bode, *Nature* **2011**, *480*, 471 – 479; b) B. Shen, D. M. Makley, J. N. Johnston, *Nature* **2010**, *465*, 1027 – 1032; c) H. Hojo, Y. Matsumoto, Y. Nakahara, E. Ito, Y. Suzuki, M. Suzuki, A. Suzuki, Y. Nakahara, *J. Am. Chem. Soc.* **2005**, *127*, 13720 – 13725; Chemoselective ligations to give non-native amide linkages: d) M. Baca, T. W. Muir, M. Schnoelzer, S. B. H. Kent, *J. Am. Chem. Soc.* **1995**, *117*, 1881 – 1887; e) I. E. Valverde, F. Lecaille, G. Lalmanach, V. Aucagne, A. F. Delmas, *Angew. Chem.* **2012**, *124*, 742 – 746; *Angew. Chem. Int. Ed.* **2012**, *51*, 718 – 722; f) J. P. Tam, Q. Yu, Z. Miao, *Pept. Sci.* **1999**, *51*, 311 – 332; g) G. A. Lemieux, C. R. Bertozzi, *Trends Biotechnol.* **1998**, *16*, 506 – 513.

[5] a) J. W. Bode, R. M. Fox, K. D. Baucom, *Angew. Chem.* **2006**, *118*, 1270 – 1274; *Angew. Chem. Int. Ed.* **2006**, *45*, 1248 – 1252; b) N. Carrillo, E. A. Davalos, J. A. Russak, J. W. Bode, *J. Am. Chem. Soc.* **2006**, *128*, 1452 – 1453.

[6] a) J. Wu, J. Ruiz-Rodriguez, J. M. Comstock, J. Z. Dong, J. W. Bode, *Chem. Sci.* **2011**, *2*, 1976 – 1979; b) S. I. Medina, J. Wu, J. W. Bode, *Org. Biomol. Chem.* **2010**, *8*, 3405 – 3417.

[7] I. Pusterla, J. W. Bode, *Angew. Chem.* **2012**, *124*, 528 – 531; *Angew. Chem. Int. Ed.* **2012**, *51*, 513 – 516.

[8] D. A. Knowles, C. H. Mathews, N. C. O. Tomkinson, *Synlett* **2008**, *18*, 2769 – 2772.

[9] A. Vasella, R. Voeffray, *J. Chem. Soc. Chem. Commun.* **1981**, 97 – 98.

[10] S. Yu, H. Ishida, M. E. Juarez-Garcia, J. W. Bode, *Chem. Sci.* **2010**, *1*, 637 – 641.

[11] a) M. J. Pearce, J. Mintseris, J. Ferreyra, S. P. Gygi, K. H. Darwin, *Science* **2008**, *322*, 1104 – 1107; b) F. Striebel, F. Imkamp, M. Sutter, M. Steiner, A. Mamedov, E. Weber-Ban, *Nat. Struct. Mol. Biol.* **2009**, *16*, 647 – 689.

[12] a) L. Ju, J. W. Bode, *Org. Biomol. Chem.* **2009**, *7*, 2259 – 2264; b) L. Ju, A. R. Lippert, J. W. Bode, *J. Am. Chem. Soc.* **2008**, *130*, 4253 – 4255.

[13] M. Sutter, F. F. Damberger, F. Imkamp, F. H. T. Allain, E. Weber-Ban, *J. Am. Chem. Soc.* **2010**, *132*, 5610 – 5612.

[14] S. H. Liao, Q. Shang, X. C. Zhang, J. H. Zhang, C. Xu, X. M. Tu, *Biochem. J.* **2009**, *422*, 207 – 215.

[15] J. Watrous, K. Burns, W.-T. Liu, A. Patel, V. Hook, V. Bafna, C. E. Barry, S. Bark, P. C. Dorrestein, *Mol. BioSyst.* **2010**, *6*, 376 – 385.

[16] G. D'auria, C. Esposito, L. Falcigno, L. Calvanese, E. Iaccarino, A. Ruggiero, C. Pedone, E. Pedone, R. Berisio, *Biochem. Biophys. Res. Commun.* **2010**, *402*, 693 – 698.

[17] M. Prorok, D. K. Sukumaran, D. S. Lawrence, *J. Biol. Chem.* **1989**, *264*, 17727 – 17733.

[18] Y. Eshdat, A. Dunn, N. Sharon, *Proc. Natl. Acad. Sci. USA* **1974**, *71*, 1658 – 1662.

