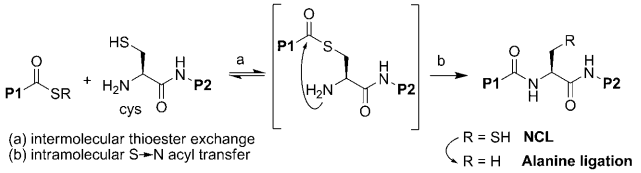
Peptide Coupling

DOI: 10.1002/anie.201005513

Insights into the Finer Issues of Native Chemical Ligation: An Approach to Cascade Ligations**

*Zhongping Tan, Shiying Shang, and Samuel J. Danishefsky**

We have been studying the total synthesis of complex glycopeptide and glycoprotein targets.[1] Pursuant to this goal, we hope to discover broadly useful methods to join large peptide and glycopeptide fragments while minimizing the need for side-chain protection.[2] A field-changing contribution to the problem of polypeptide ligation, termed native chemical ligation (NCL), was provided by Kent and coworkers in 1994 (Scheme 1).[3] NCL involves the merger of a



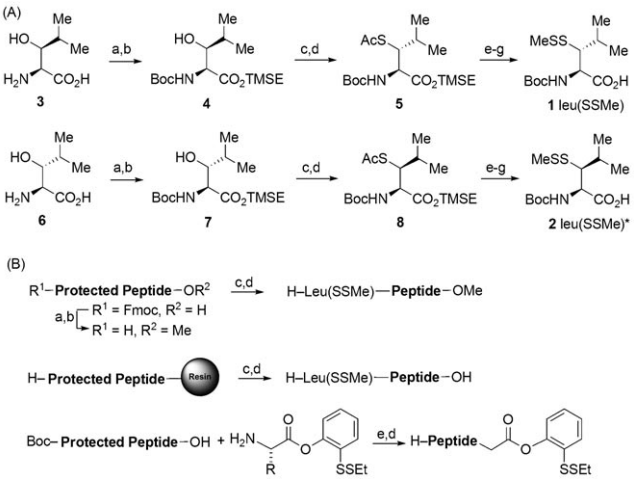
Scheme 1. NCL and alanine ligation.

peptide domain possessing a C-terminal thioester fragment with a second peptide bearing an N-terminal cysteine residue. The key mechanistic features of NCL—transthioesterification and S→N acyl transfer—are outlined in Scheme 1. Clearly if the primary thio group could be desulfurized, the NCL method can be used to accommodate alanine ligation.[4] Of course, for this to work well, other sulfur moieties within the construct must withstand the desulfurization reaction. A major advance in this regard was accomplished in a metal-free fashion, using classical mechanistic insights in free radical mediated desulfurization.[5]

We then set about to apply, more generally, the overall logic of NCL to other proteogenic amino acids. The thought was to synthesize non-proteogenic amino acids, bearing strategically placed thiol groups, to serve as the N-terminal residues in the ligation event. In this way the logic of NCL could, in principle, be broadly extended. In each case, the concluding step would exploit our metal-free desulfurization method. Indeed, this was accomplished for valine, lysine, and threonine ligations.[6]

We next turned to the possibility of extending NCL logic to accomplish leucine ligation.[7] We first prepared the diastereomeric β-thioleucine surrogates **1** [leu(SSMe)] and **2** [which we term leu(SSMe)*] (Scheme 2a). The synthesis of **1** commenced with commercially available (2*S*,3*S*)-3-hydroxyleucine (**3**) and passed through **4** and **5** as shown in Scheme 2. Compound **2** was prepared through an analogous sequence from (2*S*,3*R*)-3-hydroxyleucine (**6**).[4c,6a]

In addition to studying the feasibility and quality of the projected leucine ligations, we anticipated that the availability of two epimeric leucine surrogates (compounds **1** and **2**) as potential probes, could also provide a basis for studying rather subtle, otherwise hidden issues of native chemical ligation. Accordingly, we prepared the peptides described below (Scheme 2b). In all cases, the acyl donor component of the ligation was presented as a masked thiol ester of a type we had previously described.[2a] The required peptides (vide infra) with C-terminal methyl esters or free carboxylic acids and N-terminal leucine surrogates were prepared using HATU-



Scheme 2. Synthesis of peptide substrates. Reagents and conditions: A) a) Boc_2O, Na_2CO_3, THF/H_2O, RT, 91%; b) TMSE-OH, DCC, DMAP, CH_2Cl_2, 0°C→RT, 99%; c) MsCl, Et_3N, CH_2Cl_2, 0°C; d) AcSK (excess), DMF, RT, 40°C→60°C, 82% over two steps; e) NaOH, MeOH, 0°C; f) MMTS, DIEA, CH_2Cl_2, RT, 79% over two steps; g) TBAF, THF, RT, 98%. B) a) MeOH, DCC, DMAP, CH_2Cl_2; b) piperidine, CH_2Cl_2; c) Boc-Leu(SSMe)-OH, HATU, DIEA, DMSO; d) TFA/H_2O/TIS (95:2.5:2.5); e) EDCI, HOOBt, $CHCl_3$/TFE. TMSE = trimethylsilylethyl, DCC = dicyclohexylborane, DMAP = 4-dimethylaminopyridine, AcSK = potassium thioacetate, DMF = dimethylformamide, MMTS = methane methylthiosulfonate, DIEA = ethyldiisopropylamine, TBAF = tetrabutylammonium fluoride, Boc = *tert*-butyloxycarbonyl, HATU = *O*-(7-azabenzotriazol-1-yl)tetramethyluronium hexafluorophosphate, DMSO = dimethylsulfoxide, TFA = trifluoroacetic acid, TIS = triisopropylsilane, EDCI = *N*′-(3-dimethylaminopropyl)-*N*-ethylcarbodiimide, TFE = 2,2,2-trifluoroethanol.

[*] Prof. S. J. Danishefsky
Department of Chemistry, Columbia University
3000 Broadway, New York, NY 10027 (USA)
Fax: (+1) 212-772-8691
E-mail: s-danishefsky@ski.mskcc.org
Dr. Z. Tan, Dr. S. Shang, Prof. S. J. Danishefsky
Laboratory for Bioorganic Chemistry
Sloan-Kettering Institute for Cancer Research
1275 York Avenue, New York, NY 10065 (USA)

[**] Support for this research was provided by the National Institutes of Health (CA28824 to S.J.D.).

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201005513.

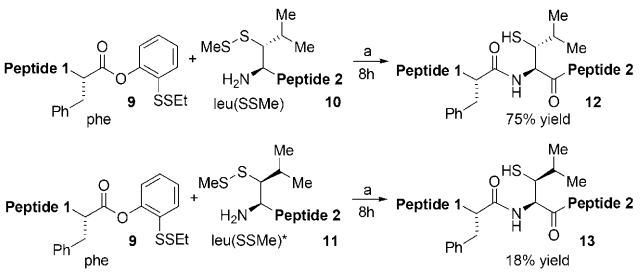
© 2010 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

mediated peptide coupling reactions in the presence of 1 equivalent of synthetic leucine precursor.[1c] The peptides bearing *ortho*-disulfide phenolic esters at the C-termini were synthesized using EDCI mediation under the non-epimerizing conditions developed by Sakakibara and co-workers. (Scheme 2b)[8]

We began by comparing the quality of ligation of the two leucine epimers by joining both **10** and **11** with peptide **9**, presenting a C-terminal phenylalanine residue (Scheme 3).[6a] The impact of thioleucine configuration on the ligation was striking. Thus, under standard conditions, peptide **10** readily underwent coupling with **9** to afford **12**. Within 15 min, the reaction had achieved ca. 50% conversion (by UPLC-MS), and a 75% yield was obtained as shown. In the case of peptide **11**, less than 7% conversion was observed after 15 min. After 30 h, an 18% yield of **13** could be obtained. The mechanistic implications of this type of finding will be discussed below.

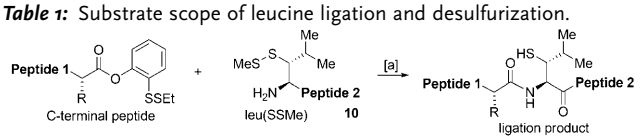
With the preferred leucine amino acid surrogate established (i.e. **1**), we next probed the versatility of the leucine ligation protocol. Not surprisingly, the rate and efficiency of ligation was found to be qualitatively dependent on the level of steric hindrance at the C-terminus residue. Thus, under standard conditions, peptide **10** underwent rapid ligation with the C-terminal glycine peptide, **14**, to furnish adduct **15** in 95% conversion within 2 h (Table 1, entry 1). Although ligation with the C-terminal alanine peptide, **16**, was somewhat less facile, a reasonable conversion of ligation product was obtained (85%, Table 1, entry 2). As expected, peptides **18** and **20**, presenting C-terminal valine and proline residues, respectively, were significantly less reactive as acyl donors and gave accordingly lower yields (Table 1, entries 4 and 5). The ability to efficiently convert the β-thioleucine surrogates to leucines in the context of their primary ligation products, upon exposure to our standard metal-free desulfurization conditions, is shown in entries 3 and 4 in Table 1.[5]

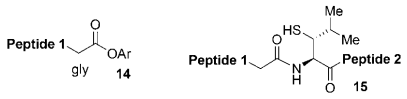
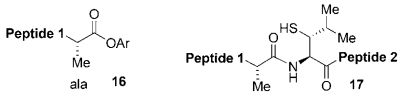
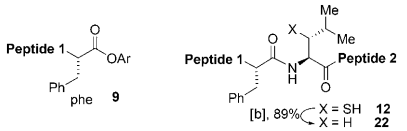
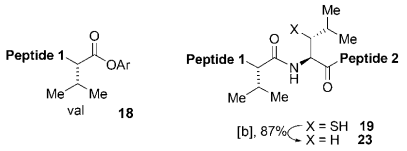
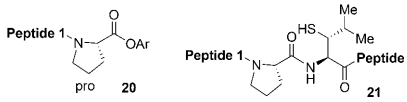
The substantial difference in the performance of peptides terminating in surrogate **1** and surrogate **2** as acyl acceptor encouraged us to probe more intensively into some of the mechanistic issues associated with native chemical ligation.[9] The widely held notion is that the transthioesterification step is rate determining, and that the acyl transfer of the C-terminal coupling component to nitrogen is rapid (see Scheme 1). This accounts for the inability to clearly identify any intermediate thioesters en route to ligation.[9]



Scheme 3. Leucine ligation with two leu(SSMe) diastereomers. Reagents and conditions: a) 6 M Gn·HCl, 100 mM NaH_2PO_4, 50 mM TCEP, pH 7.5. **Peptide 1**: GKHLNSAERVE–; **Peptide 2**: –RKKLQDVHNFVALG-OMe. Gn·HCl = guanidine hydrochloride, TCEP = tris(2-carboxyethyl)phosphine.

Table 1: Substrate scope of leucine ligation and desulfurization.

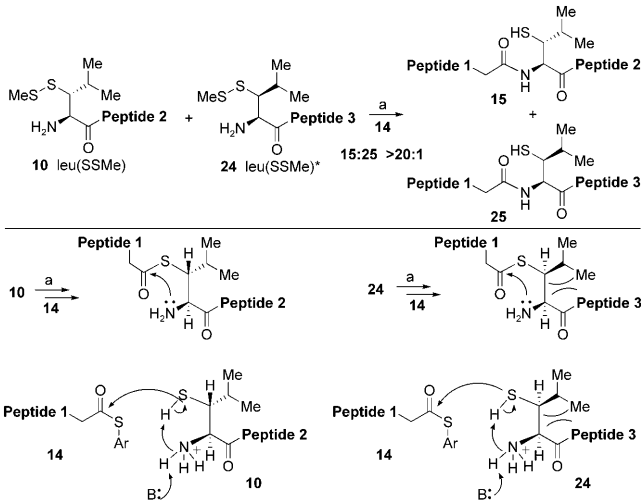


Entry	C-Terminal peptide	Ligation product	Conv. [%]	*t* [h]
1			95	2
2			85	2.5
3			83	2.5
4			50	8
5			21	9

[a] 6 M Gn·HCl, 100 mM NaH_2PO_4, 50 mM TCEP, pH 7.5; ratio C-terminal peptide/**10** = 1.5:1. [b] TCEP, VA-044, *t*BuSH, 0.8 h. **Peptide 1**: GKHLNSAERVE–; **Peptide 2**: –RKKLQDVHNFVALG-OMe.

To approach this question more precisely, we designed a competition experiment that would allow for rough comparison of the acyl acceptor qualities of the two epimeric N-termini under identical conditions. In the event, peptides **10** and **24**, bearing N-terminal leu(SSMe) and leu(SSMe)* residues were treated with the C-terminal glycine peptide **14** under standard ligation conditions. The results are shown in Scheme 4. Since the yield of the ligation of the N-terminal leu(SSMe)* epimer is poor (Scheme 3, vide supra), it is not possible to extract hard numbers for the relative rates of acyl acceptor reactivity of epimers leu(SSMe) and leu(SSMe)*. By examining the relative amount of ligation products in the early stages of the experiment, we estimate the ratio to be at least 20:1 in favor of fast reacting epimer leu(SSMe). In the limiting case, assuming the inherent acyl acceptor properties of the thiol groups in leu(SSMe) and leu(SSMe)* were nearly the same, and each intermediate suffered S→N acyl transfer with comparable efficiency, the ratio of effective ligation would be ca. 1:1. The actual comparative yields reported in Scheme 3 strongly suggest that the acyl transfer step is substantially slower in the case of leu(SSMe)* relative to leu(SSMe). This finding can be rationalized, since in the case of leu(SSMe), S→N acyl transfer requires a *trans* relationship of the substituents on the 5-membered ring, while in the case of the epimeric leu(SSMe)* system, the corresponding isopropyl and peptidic residues are *cis*. The attenuated rate of acyl transfer in the case of leu(SSMe)* presumably renders its thioester intermediate more vulnerable to competitive adventitious hydrolysis.

 © 2010 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim



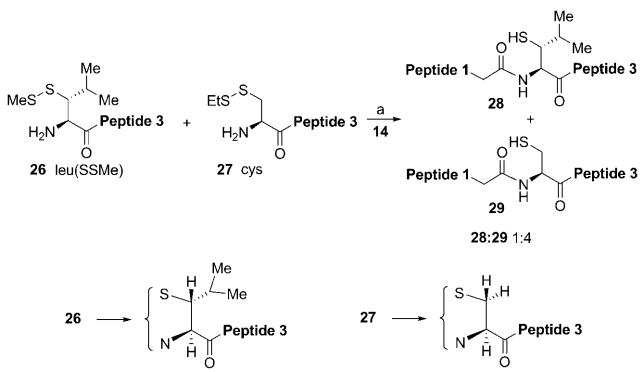
Scheme 4. The competition reaction between leu(SH) and leu(SH)* to form peptide-coupling products. a) 6 M Gn·HCl, 100 mM NaH_2PO_4, 50 mM TCEP, pH 7.5. **Peptide 1**: GKHLNSAERVE−; **Peptide 2**: RKKLQDVHNFVALG-OMe; **Peptide 3**: −RKKLQDVHNFVALG-OH.

Furthermore, the rates of the transthioesterification steps may also be quite different for the two epimeric surrogates. For intramolecular transfer to occur, the proton on the thiol acyl acceptor group must be transferred to a putative "base" or to the bulk solvent under weakly basic conditions (pH ≈7.5). A possibility in this regard is that a proton of the neighboring NH_3^+ group is donated to medium, thus allowing the nitrogen to remove the proton from the sulfur. Once again, in the case of leu(SSMe), such a transfer would require a *trans* relationship of the large substituents, while in the case of leu(SSMe)*, they are *cis* (see Scheme 4).

We then posed the question as to how the productive epimer, leu(SSMe), might compare with cysteine in a competitive experiment. Required substrates were prepared as described and the experiment crafted as shown in Scheme 5.

Because intermolecular transthioesterification is supposed to be rate determining, one might anticipate that the primary thiol of the cysteine residue would react more rapidly than the secondary thiol of the leu(SSMe) surrogate.[9] Indeed, as shown in Scheme 5, peptide **29**, arising from ligation of the cysteine-bearing peptide, **27**, was found to be the predominant product (**29**/**28** ≈ 4:1). Thus, the observed margin of selectivity was less than that expected on the basis of "A-value" comparisons between the H and isopropyl substituents. Apropos of the arguments outlined above, explanations for this apparent discrepancy may be proposed.

Putative ring character is clearly involved in the intramolecular S→N acyl transfer step in NCL. As argued above, it could also be involved in the required de-protonation of the thiol group en route to transthioacylation. In either case such ring formation, in the case of the fast reacting pre-leucine compound, is favored by an argument similar to the classical Thorpe–Ingold effect, since the acyclic array is more substituted.[10] The corresponding pertinent acylic ensemble in the case of the N-terminal cysteine is substantially less substituted. Therefore it might benefit far less from cyclization.



Scheme 5. The competition reaction between leu(SH) and cys to form peptide-coupling products. a) 6 M Gn·HCl, 100 mM NaH_2PO_4, 50 mM TCEP, pH 7.5. **Peptide 1**: GKHLNSAERVE−; **Peptide 3**: −RKKLQDVHNFVALG-OH.

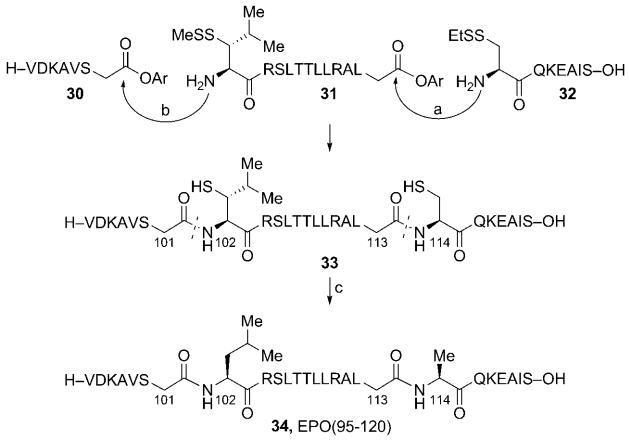
Finally, we were able to demonstrate the complexity-building capacity of this chemistry. The thought was to develop a kinetically controlled re-iterative coupling strategy, which would exploit the reactivity differential between the N-terminal thioleucine surrogate and cysteine acyl acceptors.[11] In keeping with a major concern of our laboratory, we envisioned, as a goal, the rapid assembly of the human erythropoietin (hEPO) peptide, **34**, from three individual fragments through sequential cysteine and thioleucine ligations.

The experiment was crafted as shown in Scheme 6. We first undertook to connect **32** and **31**. The resulting gross product was capped by a second ligation to **30**. In the event, a 61% yield of the desired sequential double ligation product **33** was isolated. This product indeed corresponds to the sequence **32** + **31** + **30**. We were unable to find the single ligation product corresponding to the self-coupling of **31**, though it could well have been missed (perhaps due to further oligomerization). Interestingly, we also did not detect a double ligation product arising from a sequence **31** + **31** + **30**. One product, implicated in ca. 5% yield,[12] presumably corresponds to initial cyclization (i.e. thiolactonization) of **31**. We also identified a monoligation product arising from the coupling of **32** with **30**. In principle, it is the thiolactonization side reaction, which is responsible for "leftover" **32** and **30**.

It is well to note that every self-coupling of **31** + **31** serves to deplete two equivalents of the "competing substrates". Hence, in principle, even a 4:1 acyl acceptor reactivity ratio of **32**/**31** would be leveraged to provide an ca. 8:1 factor "against" the hypothetical (unobserved) product of a single ligation of **31** + **30**.[13] Obviously, any further level of oligomerization of **31** would provide greater leveraging of the selectivity for forming **33** rather than the unobserved products (**31** + **31** + **30** or **31** + **30**).

In summary, the chemistry described above started with the central concept of native chemical ligation. To extend this core idea, we have devised an effective route to generating the required C-terminal thioester acyl donor through an intramolecular O→S transfer in the same step where the cysteine (or surrogate) thiol is exposed. Sterochemically defined leucine ligation surrogates were synthesized and used

 © 2010 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim



Scheme 6. Synthesis of EPO(95-120) peptide (**34**). Reagents and conditions: a) 6 M Gn·HCl, 100 mM NaH_2PO_4, 50 mM TCEP, pH 7.5, 0.5 h; b) 6 M Gn·HCl, 100 mM NaH_2PO_4, 50 mM TCEP, pH 7.5, 0.5 h; MESNa, H_2O/MeCN (1:1), 1 min, 61 % over two steps. c) TCEP, VA-044, *t*BuSH, 1 h, 82 %. MESNa = sodium-2-mercaptoethane sulfonate, VA-044 = 2,2′-azobis[2-(2-imidazolin-2-yl)propane] dihydrochloride.

to probe important mechanistic issues involved in native chemical ligation. Selectivities in acyl acceptor efficacies in NCL were explored from several considerations. First, evidence has been brought to bear that the rate of the S→N intramolecular acyl transfer (Scheme 1, step b) can be undermined by imposing a steric hindrance restraint at the stage where a formation of mechanistically required ring is required. Similar issues may also impact on the first step of the process, that is, transthioesterification (Scheme 1, step a). We postulate that, perhaps, even in this opening step, quasi ring formation may be a central element in directing a deprotonation of the acyl acceptor thiol group (see Scheme 4, structure **24**). Finally, we further suggest that increased substitution in the acyclic system can perhaps be a factor in the transthioesterification step by favoring ring formation in the transfer of the thiol proton by intramolecular means (see Scheme 5).[10]

A combination of these considerations served to establish a selectivity margin between N-terminal cysteine, itself, and the N-terminal thiol containing precursor of leucine. Combining the principles, we realized a rather promising sequential ligation, wherein peptide A (cf. **31**) couples with peptide B (cf. **32**) to give peptide C (not purified), which in turn couples with peptide D (cf. **30**) to give rise to peptide F (cf. **33**). The implications of these findings for the synthesis of critical, biologically relevant peptides and glycopeptides are being pursued.[14]

Received: September 2, 2010
Published online: November 4, 2010

Keywords: desulfurization · kinetic control · leucine · native chemical ligation · peptides

[1] a) B. Wu, Z. Tan, G. Chen, J. Chen, Z. Hua, Q. Wan, K. Ranganathan, S. J. Danishefsky, *Tetrahedron Lett.* **2006**, *47*, 8009; b) J. Chen, G. Chen, B. Wu, Q. Wan, Z. Tan, Z. Hua, S. J. Danishefsky, *Tetrahedron Lett.* **2006**, *47*, 8013; c) Z. Tan, S. Shang, T. Halkina, Y. Yuan, S. J. Danishefsky, *J. Am. Chem. Soc.* **2009**, *131*, 5424; d) Y. Yuan, J. Chen, Q. Wan, Z. P. Tan, G. Chen, C. Kan, S. J. Danishefsky, *J. Am. Chem. Soc.* **2009**, *131*, 5432; e) C. Kan, J. D. Trzupek, B. Wu, G. Chen, Z. P. Tan, Y. Yuan, S. J. Danishefsky, *J. Am. Chem. Soc.* **2009**, *131*, 5438; f) P. Nagorny, B. Fasching, X. Li, G. Chen, B. Aussedat, S. J. Danishefsky, *J. Am. Chem. Soc.* **2009**, *131*, 5792.

[2] a) J. D. Warren, J. S. Miller, S. J. Keding, S. J. Danishefsky, *J. Am. Chem. Soc.* **2004**, *126*, 6576; b) B. Wu, J. Chen, J. D. Warren, G. Chen, Z. Hua, S. J. Danishefsky, *Angew. Chem.* **2006**, *118*, 4222; *Angew. Chem. Int. Ed.* **2006**, *45*, 4116; c) B. Wu, J. D. Warren, J. Chen, G. Chen, Z. Hua, S. J. Danishefsky, *Tetrahedron Lett.* **2006**, *47*, 5219; d) G. Chen, Q. Wan, Z. Tan, C. Kan, Z. Hua, K. Ranganathan, S. J. Danishefsky, *Angew. Chem.* **2007**, *119*, 7527; *Angew. Chem. Int. Ed.* **2007**, *46*, 7383; e) Q. Wan, J. Chen, Y. Yuan, S. J. Danishefsky, *J. Am. Chem. Soc.* **2008**, *130*, 15814; f) X. Li, S. J. Danishefsky, *J. Am. Chem. Soc.* **2008**, *130*, 5446.

[3] a) P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. Kent, *Science* **1994**, *266*, 776; b) J. P. Tam, Y. A. Lu, C. F. Liu, J. Shao, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 12485.

[4] a) L. Z. Yan, P. E. Dawson, *J. Am. Chem. Soc.* **2001**, *123*, 526; b) B. L. Pentelute, S. B. Kent, *Org. Lett.* **2007**, *9*, 687; c) D. Crich, A. Banerjee, *J. Am. Chem. Soc.* **2007**, *129*, 10064.

[5] a) Q. Wan, S. J. Danishefsky, *Angew. Chem.* **2007**, *119*, 9408; *Angew. Chem. Int. Ed.* **2007**, *46*, 9248; for the seleno analog, see: b) N. Metanis, E. Keinan, P. E. Dawson, *Angew. Chem.* **2010**, *122*, 7203; *Angew. Chem. Int. Ed.* **2010**, *49*, 7049.

[6] a) J. Chen, Q. Wan, Y. Yuan, J. Zhu, S. J. Danishefsky, *Angew. Chem.* **2008**, *120*, 8649; *Angew. Chem. Int. Ed.* **2008**, *47*, 8521; b) C. Haase, H. Rohde, O. Seitz, *Angew. Chem.* **2008**, *120*, 6912; *Angew. Chem. Int. Ed.* **2008**, *47*, 6807; c) J. Chen, P. Wang, J. L. Zhu, Q. Wan, S. J. Danishefsky, *Tetrahedron* **2010**, *66*, 2277; d) R. Yang, K. K. Pasunooti, F. Li, X. W. Liu, C. F. Liu, *J. Am. Chem. Soc.* **2009**, *131*, 13592.

[7] During the preparation of this manuscript, Brik and co-workers reported a native chemical ligation at leucine. However, these authors did not investigate the relative rates of the two diastereomeric leucine surrogates as the key selectivity versus reported herein. Z. Harpaz, P. Siman, K. S. Kumar, A. Brik, *ChemBioChem* **2010**, *11*, 1232.

[8] S. Sakakibara, *Biopolymers* **1995**, *37*, 17.

[9] a) S. B. Kent, *Chem. Soc. Rev.* **2009**, *38*, 338; b) E. C. Johnson, S. B. Kent, *J. Am. Chem. Soc.* **2006**, *128*, 6640; c) T. W. Muir, *Annu. Rev. Biochem.* **2003**, *72*, 249; d) T. M. Hackeng, J. H. Griffin, P. E. Dawson, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 10068; e) L. E. Canne, S. J. Bark, S. B. H. Kent, *J. Am. Chem. Soc.* **1996**, *118*, 5891.

[10] R. M. Beesley, C. K. Ingold, J. F. Thorpe, *J. Chem. Soc. Trans.* **1915**, *107*, 1080. Of course, in its pure form, the classical Thorpe–Ingold effect refers to the inclusion of a fully substituted (quaternary) carbon in the cyclizing chain. However, as a general matter, heavier substitution (in the absence of steric hindrance in the resulting ring) tends to favor cyclization for reasons of entropy.

[11] D. Bang, B. L. Pentelute, S. B. Kent, *Angew. Chem.* **2006**, *118*, 4089; *Angew. Chem. Int. Ed.* **2006**, *45*, 3985.

[12] This product was not isolated, per se, but inferred as shown in the Supporting Information.

[13] Of course, the 4:1 ratio established for the benchmark case in Scheme 5 may not be applicable to the competition of **32** and **31**.

[14] a) R. M. Wilson, S. J. Danishefsky, *Pure Appl. Chem.* **2007**, *79*, 2189; b) D. P. Gamblin, E. M. Scanlan, B. G. Davis, *Chem. Rev.* **2009**, *109*, 131.

 © 2010 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim