Chemical Ligation

International Edition: DOI: 10.1002/anie.201505274
German Edition: DOI: 10.1002/ange.201505274

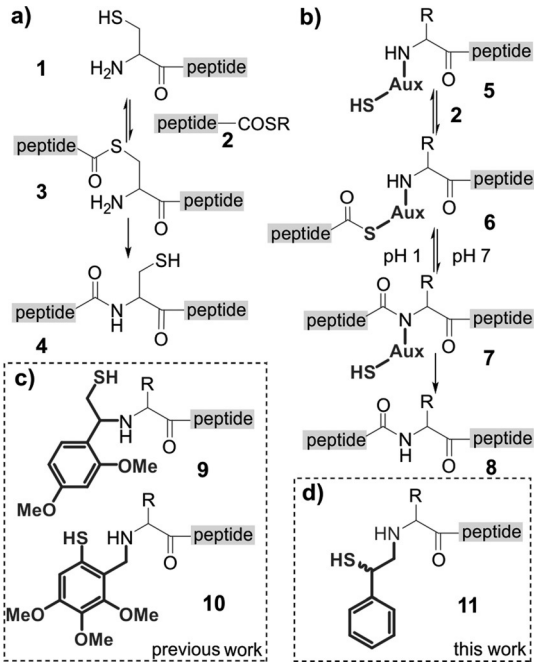
A Type of Auxiliary for Native Chemical Peptide Ligation beyond Cysteine and Glycine Junctions

*Simon F. Loibl, Ziv Harpaz, and Oliver Seitz**

Abstract: *Native chemical ligation enables the chemical synthesis of proteins. Previously, thiol-containing auxiliary groups have been used to extend the reaction scope beyond N-terminal cysteine residues. However, the N-benzyl-type auxiliaries used so far result in rather low reaction rates. Herein, a new N^{α}-auxiliary is presented. Consideration of a radical fragmentation for cleavage led to the design of a new auxiliary group which is selectively removed under mildly basic conditions (pH 8.5) in the presence of TCEP and morpholine. Most importantly and in contrast to previously described auxiliaries, the 2-mercapto-2-phenethyl auxiliary is not limited to Gly-containing sites and ligations succeed at sterically demanding junctions. The auxiliary is introduced in high yield by on-resin reductive amination with commercially available amino acid building blocks. The synthetic utility of the method is demonstrated by the synthesis of two antimicrobial proteins, DCD-1L and opistoporin-2.*

The native chemical ligation (NCL) reaction is a frequently used tool in the total chemical synthesis of proteins.[1] A thiol-exchange reaction between an unprotected peptide thioester **2** and an unprotected cysteine peptide **1** and the subsequent intramolecular S→N acyl shift in the thioester-linked intermediate **3** (Scheme 1 a) to give a native peptide bond are hallmarks of this chemoselective segment coupling. However, many proteins lack cysteine residues.

Two approaches have been developed to extend the scope of NCL beyond ligation at cysteine. In auxiliary-mediated NCL a mercapto group is attached to the N-terminus of the C-terminal segment by means of a cleavable tether (**5** in Scheme 1 b).[2] To obtain the native peptides **8** the auxiliaries are removed after ligation by photolysis[3] or, more commonly, acidolysis.[4] Low yields at junctions other than Gly–Gly as well as a tendency for cleavage of the established amide bond during acidolytic removal of the auxiliary are recurring problems of the commonly used *N*-benzyl-type auxiliaries (e.g., **9**, **10**). Alternatively, the ligation–desulfurization strategy provides access to ligation junctions that do not contain cysteine.[5] This method relies on thiolated amino acids at the N-terminal end of the C-terminal NCL partner. Unfortunately, only 4 out of the 13 ligation sites accessible with the reported side-chain-thiolated pecursors can be established by means of commercially available amino acids.[5c,h,m] The requirement for the preparation of the noncommercial building blocks by laborious multistep procedures (>50 steps for precursors of Arg,[5j] Asp,[5k] Gln,[5i] Glu,[5l] Leu,[5g] Lys,[5e] Phe,[5b] Thr[5f]) is a nuisance. This probably is the reason why the application of the ligation–desulfurization currently is restricted to a few expert laboratories.



Scheme 1. Overview of a) native chemical ligation and b) auxiliary-mediated peptide ligation. c) Examples of currently used auxilaries; d) structure of the 2-mercapto-2-phenethyl auxiliary designed by us.

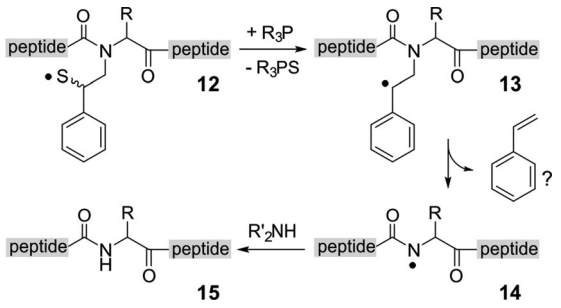
In the pursuit of a more generally applicable ligation method that is suitable for use with commercially available amino acid building blocks, we revisited the auxiliary approach. The commonly used 2-mercapto-1-phenethyl auxiliaries such as **9**[4b,g] contain a branching point at the α-carbon atom (Scheme 1 c). This dramatically reduces the rate of S→N acyl transfer at non-glycine ligation junctions. Alternatively, 2-mercaptobenzyl auxiliaries such as **10**[4f] are used in extended NCL. However, ligation rates are rather low because the reaction must proceed via a six-membered rather than five-membered transition state. Furthermore, a commonly observed problem is the reversibility of the S→N acyl shift when tertiary amide bonds are exposed to acidic conditions required for auxiliary removal.[4i]

Our auxiliary design is based on a 2-mercapto-2-phenethyl scaffold **11** (Scheme 1 d). This auxiliary allows native chemical ligation via a five-membered transition state, and

[*] S. F. Loibl, Z. Harpaz, Prof. Dr. O. Seitz
Institut für Chemie, Humboldt-Universität zu Berlin
Brook-Taylor-Strasse 2, 12489 Berlin (Germany)
E-mail: oliver.seitz@chemie.hu-berlin.de

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201505274.

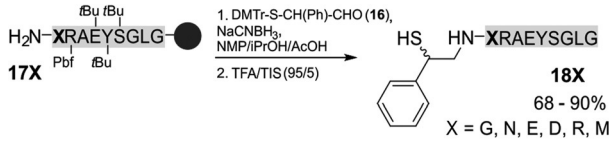
because it does not contain α-branching the reaction should occur rapidly. The removal should be feasible under non-acidic conditions. We envisioned a radical β-fragmentation pathway (Scheme 2). According to our hypothesis the thiyl radical **12** would, in analogy to the radical desulfurization



Scheme 2. Hypothetic mechanism for the removal of a 2-mercapto-2-phenethyl auxiliary.

procedure,[6] form the benzyl radical **13** upon treatment with phosphine. In the absence of a reactive hydrogen donor, β-fragmentation would deliver the amide radical **14** along with styrene which may react further under the cleavage conditions. We expected that the nitrogen-centered amide radical should be readily scavenged by amines such as morpholine which are not reactive enough to quench the benzyl radical **13**.[7]

Since the auxiliary is removed after ligation, its chirality does not affect the final product. Therefore, the 2-mercapto-2-phenethyl auxiliary lends itself to a facile introduction by reductive amination of the racemic alkyl aldehyde with the peptide N-terminus in the last step of solid-phase peptide synthesis (Scheme 3). The auxiliary was attached to the resin-



Scheme 3. Synthesis of auxiliary-substituted peptides by reductive amination during solid-phase synthesis. DMTr: 4,4′-dimethoxytrityl; NMP: *N*-methyl-2-pyrrolidone; TFA: trifluoroacetic acid; TIS: triisopropylsilane.

bound glycine model peptide **17G** by using an excess of racemic aldehyde **16** (see Scheme S1 in the Supporting Information for the synthesis) and $NaCNBH_3$ (Scheme 3). UPLC-MS analysis of the crude material obtained after TFA cleavage showed near-quantitative conversion to the desired auxiliary-substituted peptide **18G** within four hours (Figure S2). A similar procedure furnished the asparagine, glutamate, aspartate, arginine, and methionine peptides **18X**.

We next examined the reactivity of the 2-mercapto-2-phenethyl auxiliary in NCL reactions. We compared the rates of eight different ligation reactions (Figure 1a). The conjugation proceeded much faster than reported for previous

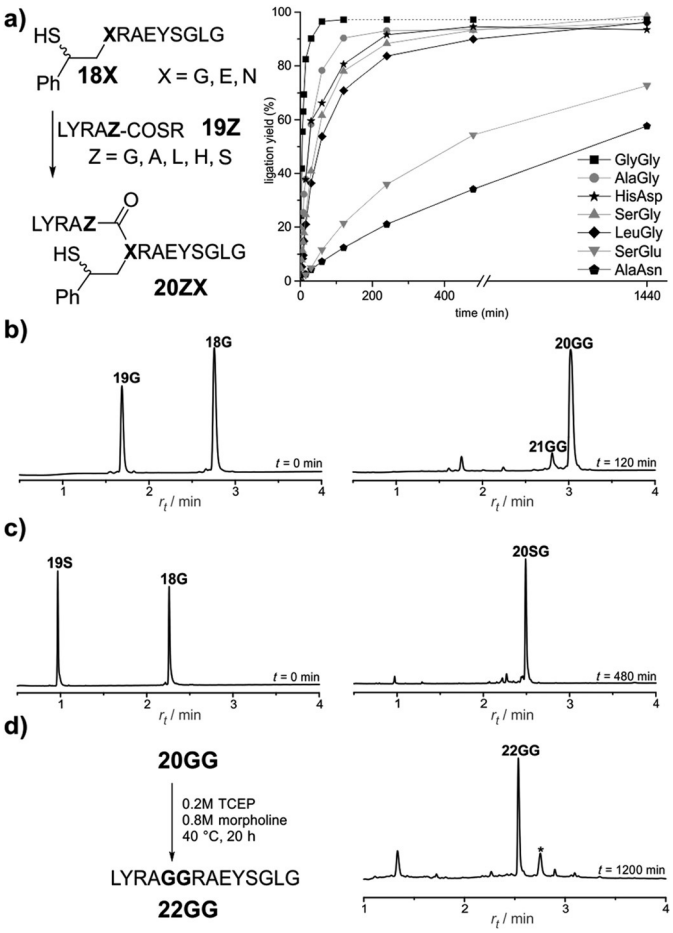


Figure 1. a) Time course of peptide ligation between peptides **18X** and thioesters **19Z ; c**onditions: 5 mM peptides, 20 mM TCEP, 100 mM phosphate, 3 vol% PhSH,[8] 25 °C, pH 7.5. UPLC analysis (UPLC: ultra-performance liquid chromatography; $\lambda = 210$ nm) of b) Gly–Gly ligation and c) Ser–Gly ligation, and d) of auxiliary cleavage from purified ligation product **20GG** after 20 h.

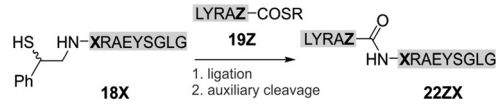
auxiliaries. As expected, formation of the Gly–Gly junction was particularly swift and reached completion within 60 min. As previously observed for other native peptide couplings,[9] the reaction afforded small amounts (about 6%) of the N- and S-acylated ligation product **21GG** (Figure 1b). Double acylation was not observed in the sterically more challenging Ser–Gly ligation (Figure 1c). The reactions at Ala–Gly and Leu–Gly junctions required 2–4 h to provide ≥80% product (see also Figure S18). Most noteworthy, the 2-mercapto-2-phenethyl auxiliary enabled ligations at junctions that did not contain glycine—reactions that are difficult if not impossible with the previous auxiliaries. Formation of the His–Asp peptide bond was particularly rapid (90% yield after 4 h) but also the Ala–Asn and Ser–Glu ligations proceeded in useful ≥50% yield at low concentration (5 mM) within 8–24 h (see also Figure S19) Higher yields were obtained when the reactions were performed with excess thioester (vide infra).

The next was to remove the auxiliary from the purified ligation product. Based on our previous experience[10] we anticipated that the benzyl radical **13** should be readily formed at high concentrations of triscarboxyethylphosphine

(TCEP) probably upon exposure to light or oxygen. We screened various amines which were envisioned to act as amide radical scavengers (Figure S22). We found that treatment of an auxiliary-modified ligation product such as **20GG** with aqueous solutions of TCEP (0.2 M) and excess morpholine (0.8 M) at pH 8.5 induced a smooth reaction which led to quantitative cleavage of the auxiliary within 20 h (Figure 1 d). Of note, the presence of TEMPO, a classical radical scavenger, resulted in complete suppression of the cleavage reaction (Figure S24).

To avoid intermediary HPLC purification we developed a more straightforward "one-pot" protocol including peptide ligation followed by subsequent auxiliary removal (Table 1).

Table 1: Yields obtained by one-pot ligation auxiliary cleavage.[a]



19Z	18X	Ligation yield[b] / time	Cleavage[b]	22ZX[c]
19G[d]	**18G**	>99 / 1 h	74	**22GG** (58)
19L[d]	**18G**	86 / 14 h	76	**22LG** (44)
19G[d]	**18N**	94 / 20 h	82	**22GN** (49)
19H[d]	**18D**	>99 / 24 h	79	**22HD** (45)
19L[e]	**18R**	77 / 24 h	88	**22LR** (35)
19F[e]	**18M**	91 / 24 h	89	**22FM** (37)

[a] Conditions: Peptides (2–5 mM) dissolved in buffer (20 mM TCEP, 100 mM $NaHPO_4$, 3 vol% PhSH,[8] pH 7.5) for ligation, then aqueous TCEP (0.1–1.0 M) and morpholine (1–4 M) at pH 8.5. For details see the Supporting Information. [b] Analytical yield in % determined by UPLC-MS analysis. [c] Yield in % of isolated material obtained after HPLC purification over two steps given in brackets. [d] R = $(CH_2)_3CONHCH_2CONH_2$. [e] R = Ph.

A TCEP-containing aqueous solution of morpholine (pH 8.5) was added to the reaction mixture once ligation was complete. To establish the Gly–Gly bond in **22GG** the thioester **19G** and the auxiliary-substituted peptide **18G** were allowed to react at 2 mM concentration in phosphate buffer at pH 7.5. After 2 h the reaction mixture was diluted with a solution of TCEP (200 mM) and morpholine (800 mM) and the mixture was incubated for 36 h at 40 °C. After HPLC purification the native peptide **22GG** was obtained in 58 % overall yield. This yield is high given the 75 % rate of recovery determined for the applied HPLC/lyophilization procedure. For the synthesis of the Leu–Gly and Gly–Asn linkages, the concentration of peptides was increased to 5 mM and guanidine hydrochloride (GnHCl) was added to the ligation buffer. The ligation was complete after 14–20 h. Subsequently, the mixture was diluted with the TCEP/morpholine solution. The auxiliary cleavage proceeded also in the GnHCl-containing buffer and the native peptides were isolated in 44 % and 49 % yield after two steps.

We next turned to one-pot reaction sequences at ligation junctions not containing glycine: His–Asp, Leu–Arg, and Phe–Met. To drive the ligation, further peptide thioester (0.5–1 equiv) was added after 8 h (Leu–Arg, Phe–Met) or 24 h (His–Asp). The ligation products showed a tendency to undergo the undesired N→S acyl shift upon exposure to 0.1 % TFA (Figure S21b). This side reaction points to the advantage provided by the slightly basic (rather than acidic) conditions for removal of the 2-mercapto-2-phenethyl auxiliary. Efficient cleavage of the auxiliary was induced without indications of N→S acyl shift reactions (Figure S21c).

We applied the 2-mercapto-2-phenethyl auxiliary in the synthesis of dermcidin DCD-1L (**26**, Figure 2). This antimicrobial peptide is expressed in human sweat glands and serves as a precursor of several shorter bioactive peptides.[11] The

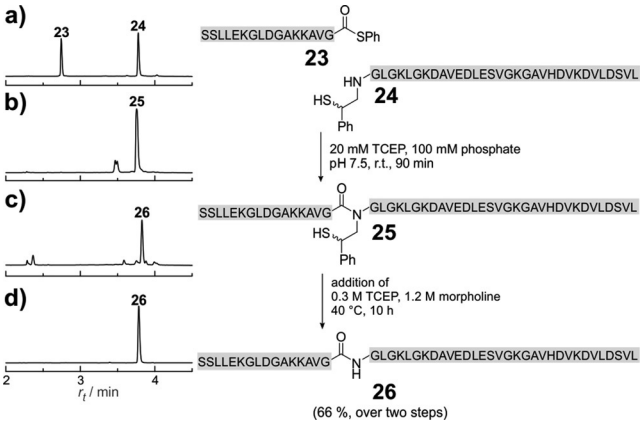


Figure 2. One-pot synthesis of native DCD-1L (**26**) by auxiliary-mediated peptide ligation and subsequent auxiliary removal. UPLC analysis (λ = 210 nm) a) before and b) after ligation, c) of the crude reaction mixture after auxiliary removal and d) of purified **26**.

protein has a Gly–Gly junction which seems ideal for auxiliary-mediated native chemical ligation. For the synthesis of the N-terminal DCD-1L segment **23** (1–16) the protected peptide acid was prepared on a chlorotrityl resin and converted to the reactive arylthioester **23** by reaction of the fully protected peptide acid with PyBOP and thiophenol in solution (Figure S32).[12] The synthesis of the C-terminal DCD-1L segment **24** (17–48) proved more difficult. The use of pseudoproline dipeptides[13] at positions 31 and 46 significantly increased the efficiency of Fmoc-SPPS. After incorporation of the auxiliary onto the N-terminal Gly residue by on-resin reductive amination, the HPLC-purified peptide **24** was obtained in 13 % overall yield. The ligation of both fragments proceeded smoothly in less than two hours (Figure 2 b). Subsequent addition of an aqueous TCEP/morpholine solution (pH 8.5) at 40 °C triggered the desired auxiliary removal (Figure 2 c). After 10 h, HPLC purification afforded native DCD-1L (**26**) in 66 % yield over two steps and in high purity (Figure 2 d).

In the chemical synthesis of opistoporin-2 (OP2)[14] we targeted a more challenging ligation site (Figure 3). This antimicrobial, α-helical pore-forming peptide lacks a suitably positioned glycine residue. We choose the Ser^{17}–Glu^{18} junction. The reactive arylthioester **27** of the N-terminal OP2 fragment (1–17) was prepared as described for **23**. A small amount of racemized peptide thioester (3 %) was removed during HPLC purification (Figure S35). A pseudoproline (Asn^{25}–Thr^{26}) was used in the SPPS of the C-terminal OP2

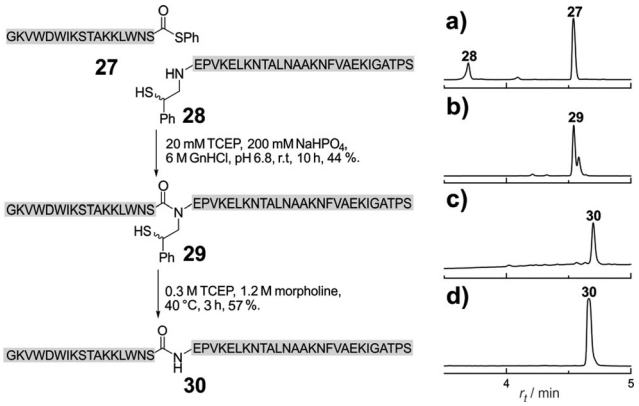


Figure 3. Synthesis of OP2. UPLC analysis ($\lambda = 210$ nm) a) before ligation, b) of purified ligation product (**29**), c) of the crude mixture after auxiliary removal and d) of purified **30**.

segment (18–44). The auxiliary was, again, introduced by reductive amination (Figure S34). The ligation was carried out in 6 M GnHCl at pH 6.8 and ambient temperature. After only 10 h the ligation product **29** was isolated by HPLC purification (44 %, Figure 3b; it appears as a double peak due to separation of the auxiliary's diastereomeric forms). Subsequent treatment with an aqueous solution of TCEP (0.3 M) and morpholine (1.2 M) at pH 8.5 proceeded smoothly within 3 h at 40 °C (Figure 3c) and, after HPLC purification, furnished native **30** in 57 % yield. We performed the peptide ligation and subsequent auxiliary removal in both small and preparative-scale formats as well as in a one-pot format, and obtained similar overall yields (Figures S37, S39, S40).

The previously reported N-auxiliaries were designed for acidolytic or photolytic removal reactions. However, the α-branch at the secondary amine (see **9**, Scheme 1c) and/or the aryl substituents (see **10**, Scheme 1c) introduce bulk that limits the scope and yields of auxiliary-mediated native chemical ligation. Considering a radical cleavage pathway enabled us to draft a new type of auxiliary that does not contain α-branching at the secondary amine. The 2-mercapto-2-phenethyl scaffold can be readily introduced in high yield by means of reductive amination with the resin-bound peptide. This facilitates the application on different ligation sites with just a single compound.

Perhaps most importantly and in contrast to previously described auxiliaries, the 2-mercapto-2-phenethyl auxiliary is not limited to Gly-containing ligation sites. This opens up new ligation opportunities. To minimize intermediary purification steps, we established a reliable one-pot protocol, which involves peptide ligation and the subsequent auxiliary removal, to yield the desired native peptides. Furthermore, this method enables (selective) auxiliary cleavage under mildly basic conditions (pH 8.5), making it interesting for accessing proteins that bear acid-sensitive modifications. We showed the chemical total synthesis of the two antimicrobial peptides DCD-1 L and opistoporin-2. While ligations at junctions containing β-branched amino acids may still be difficult (Figure S43) we expect that scope and efficiency of the 2-mercapto-2-phenethyl auxiliary method will prove useful for a wider variety of ligation sites than is currently accessible. Since the method relies on radical desulfurization Cys residues need to be protected (see Figures S44, S45). Compared to the currently most frequently applied ligation–desulfurization approach, it certainly is an advantage here that the preparation of noncommercial amino acid building blocks is not required. Based on this and the broader applicability we anticipate that the new auxiliary paradigm will be helpful for future protein syntheses.

Acknowledgements

This work was financially supported by the Deutsche Forschungsgemeinschaft (Se819-15/1, SPP 1623). We acknowledge Luxembourg Bio Technologies Ltd. for providing OxymaPure.

Keywords: desulfurization · N^{α}-auxiliaries · peptide ligation · protein synthesis · radical reactions

How to cite: *Angew. Chem. Int. Ed.* **2015**, *54*, 15055–15059
Angew. Chem. **2015**, *127*, 15269–15273

[1] P. Dawson, T. Muir, I. Clark-Lewis, S. Kent, *Science* **1994**, *266*, 776–779.
[2] L. E. Canne, S. J. Bark, S. B. H. Kent, *J. Am. Chem. Soc.* **1996**, *118*, 5891–5896.
[3] a) T. Kawakami, S. Aimoto, *Tetrahedron Lett.* **2003**, *44*, 6059–6061; b) C. Marinzi, J. Offer, R. Longhi, P. E. Dawson, *Bioorg. Med. Chem.* **2004**, *12*, 2749–2757; c) C. Chatterjee, R. K. McGinty, J.-P. Pellois, T. W. Muir, *Angew. Chem. Int. Ed.* **2007**, *46*, 2814–2818; *Angew. Chem.* **2007**, *119*, 2872–2876; d) C. Nadler, A. Nadler, C. Hansen, U. Diederichsen, *Eur. J. Org. Chem.* **2015**, 3095–3102; e) C. Bello, S. Wang, K. W. Moremen, C. F. W. Becker, *Angew. Chem. Int. Ed.* **2015**, *54*, 7711–7715; *Angew. Chem.* **2015**, *127*, 7823–7828.
[4] a) J. Offer, P. E. Dawson, *Org. Lett.* **2000**, *2*, 23–26; b) P. Botti, M. R. Carrasco, S. B. H. Kent, *Tetrahedron Lett.* **2001**, *42*, 1831–1833; c) T. Kawakami, K. Akaji, S. Aimoto, *Org. Lett.* **2001**, *3*, 1403–1405; d) D. W. Low, M. G. Hill, M. R. Carrasco, S. B. Kent, P. Botti, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 6554–6559; e) C. Marinzi, S. J. Bark, J. Offer, P. E. Dawson, *Bioorg. Med. Chem.* **2001**, *9*, 2323–2328; f) J. Offer, C. N. C. Boddy, P. E. Dawson, *J. Am. Chem. Soc.* **2002**, *124*, 4642–4646; g) D. Macmillan, D. W. Anderson, *Org. Lett.* **2004**, *6*, 4659–4662; h) S. Tchertchian, O. Hartley, P. Botti, *J. Org. Chem.* **2004**, *69*, 9208–9214; i) B. Wu, J. Chen, J. D. Warren, G. Chen, Z. Hua, S. J. Danishefsky, *Angew. Chem. Int. Ed.* **2006**, *45*, 4116–4125; *Angew. Chem.* **2006**, *118*, 4222–4231.
[5] a) L. Z. Yan, P. E. Dawson, *J. Am. Chem. Soc.* **2001**, *123*, 526–533; b) D. Crich, A. Banerjee, *J. Am. Chem. Soc.* **2007**, *129*, 10064–10065; c) C. Haase, H. Rohde, O. Seitz, *Angew. Chem. Int. Ed.* **2008**, *47*, 6807–6810; *Angew. Chem.* **2008**, *120*, 6912–6915; d) K. S. Ajish Kumar, M. Haj-Yahya, D. Olschewski, H. A. Lashuel, A. Brik, *Angew. Chem. Int. Ed.* **2009**, *48*, 8090–8094; *Angew. Chem.* **2009**, *121*, 8234–8238; e) R. Yang, K. K. Pasunooti, F. Li, X.-W. Liu, C.-F. Liu, *J. Am. Chem. Soc.* **2009**, *131*, 13592–13593; f) J. Chen, P. Wang, J. Zhu, Q. Wan, S. J. Danishefsky, *Tetrahedron* **2010**, *66*, 2277–2283; g) Z. Harpaz, P. Siman, K. S. A. Kumar, A. Brik, *ChemBioChem* **2010**, *11*, 1232–1235; h) S. Shang, Z. Tan, S. Dong, S. J. Danishefsky, *J. Am. Chem. Soc.* **2011**, *133*, 10784–10786; i) P. Siman, S. V. Karthikeyan, A. Brik, *Org. Lett.* **2012**, *14*, 1520–1523; j) L. R. Malins, K. M. Cergol, R. J. Payne, *ChemBioChem* **2013**, *14*, 559–

563; k) R. E. Thompson, B. Chan, L. Radom, K. A. Jolliffe, R. J. Payne, *Angew. Chem. Int. Ed.* **2013**, *52*, 9723 – 9727; *Angew. Chem.* **2013**, *125*, 9905 – 9909; l) K. M. Cergol, R. E. Thompson, L. R. Malins, P. Turner, R. J. Payne, *Org. Lett.* **2014**, *16*, 290 – 293; m) L. R. Malins, K. M. Cergol, R. J. Payne, *Chem. Sci.* **2014**, *5*, 260 – 266; n) L. R. Malins, R. J. Payne, *Aust. J. Chem.* **2015**, *68*, 521 – 537.

[6] Q. Wan, S. J. Danishefsky, *Angew. Chem. Int. Ed.* **2007**, *46*, 9248 – 9252; *Angew. Chem.* **2007**, *119*, 9408 – 9412.

[7] J. Hioe, D. Sakic, V. Vrcek, H. Zipse, *Org. Biomol. Chem.* **2015**, *13*, 157 – 169.

[8] E. C. B. Johnson, S. B. H. Kent, *J. Am. Chem. Soc.* **2006**, *128*, 6640 – 6646.

[9] V. Y. Torbeev, S. B. H. Kent, *Angew. Chem. Int. Ed.* **2007**, *46*, 1667 – 1670; *Angew. Chem.* **2007**, *119*, 1697 – 1700.

[10] H. Rohde, J. Schmalisch, Z. Harpaz, F. Diezmann, O. Seitz, *ChemBioChem* **2011**, *12*, 1396 – 1400.

[11] B. Schittek, R. Hipfel, B. Sauer, J. Bauer, H. Kalbacher, S. Stevanovic, M. Schirle, K. Schroeder, N. Blin, F. Meier, G. Rassner, C. Garbe, *Nat. Immunol.* **2001**, *2*, 1133 – 1137.

[12] a) S. Futaki, K. Sogawa, J. Maruyama, T. Asahara, M. Niwa, H. Hojo, *Tetrahedron Lett.* **1997**, *38*, 6237 – 6240; b) R. von Eggelkraut-Gottanka, A. Klose, A. G. Beck-Sickinger, M. Beyermann, *Tetrahedron Lett.* **2003**, *44*, 3551 – 3554.

[13] a) T. Haack, M. Mutter, *Tetrahedron Lett.* **1992**, *33*, 1589 – 1592; b) T. Wöhr, M. Mutter, *Tetrahedron Lett.* **1995**, *36*, 3847 – 3848.

[14] L. Moerman, S. Bosteels, W. Noppe, J. Willems, E. Clynen, L. Schoofs, K. Thevissen, J. Tytgat, J. Van Eldere, J. van der Walt, F. Verdonck, *Eur. J. Biochem.* **2002**, *269*, 4799 – 4810.

Received: June 9, 2015
Revised: September 17, 2015
Published online: November 6, 2015

