

# Synthesis of Hybrid Cyclopeptides through Enzymatic Macrocyclization

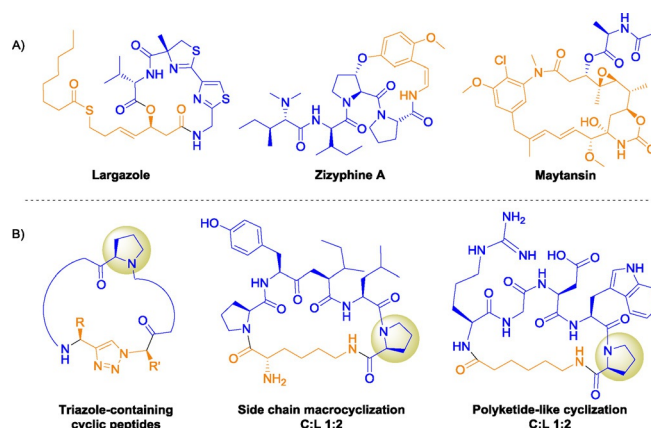
Emilia Oueis,<sup>[a]</sup> Brunello Nardone,<sup>[a]</sup> Marcel Jaspars,<sup>[c]</sup> Nicholas J. Westwood,<sup>\*,[a]</sup> and James H. Naismith<sup>\*,[a, b]</sup>

Natural products comprise a diverse array of molecules, many of which are biologically active. Most natural products are derived from combinations of polyketides, peptides, sugars, and fatty-acid building blocks. Peptidic macrocycles have attracted attention as potential therapeutics possessing cell permeability, stability, and easy-to-control variability. Here, we show that enzymes from the patellamide biosynthetic pathway can be harnessed to make macrocycles that are hybrids of amino acids and a variety of manmade chemical building blocks, including aryl rings, polyethers, and alkyl chains. We have made macrocycles with only three amino acids, one of which can be converted to a thiazoline or a thiazole ring. We report the synthesis of 18 peptide hybrid macrocycles, nine of which have been isolated and fully characterized.

Natural products, in particular secondary metabolites, derived from various natural resources such as bacteria, fungi, marine sources, and plants, have been used as medicines since ancient times. Today they have utility directly as drugs, for example paclitaxel, or as active scaffolds, for example the  $\beta$ -lactam ring.<sup>[1]</sup> The advent of high-throughput combinatorial chemistry has coincided with a decline in attention towards natural products by pharmaceutical companies.<sup>[2]</sup> Recent progress in phenotypic screening and metabolomic technologies has rekindled interest in natural products.<sup>[1a,2–3]</sup> In addition to vast structural and chemical diversity, natural products are often stereochemically complex; this feature has allowed them to evolve both specificity and potency.<sup>[4]</sup> Macrocycles, in particular, whether peptidic or polyketides, are especially appealing, owing to their inherent chemical stability and structural rigidity,<sup>[5]</sup> the later proper-

ty reducing the entropic penalty for binding, thus raising affinity.

Several natural macrocycles possess useful biological and medicinal activities.<sup>[6]</sup> Hybrid peptide macrocycles show increased chemical and structural diversity, for example, largazole (PKS/NRPS hybrid), zizyphines (cyclopeptide alkaloids), and maytansin (macrolide)<sup>[6c,7]</sup> (Figure 1). A drawback of peptidic macrocycles is their polarity, both in the side chain and the backbone, where it is an inescapable consequence of the peptide bond. High polarity hinders transport across lipid bilayers,<sup>[8]</sup> strategies such as methylation, prenylation, and heterocyclization aim to overcome these problems.<sup>[9]</sup> The synthesis of macrocyclic compounds where the backbone is a hybrid of both peptidic and non-peptidic scaffolds is desirable, as it permits diversification and exquisite control of the physicochemical properties.<sup>[10]</sup>



**Figure 1.** Structures of hybrid molecules. A) Natural hybrid macrocyclic compounds.<sup>[6c,7]</sup> B) Hybrid cyclopeptides previously synthesized using PatGmac.<sup>[11]</sup> C:L indicates the ratio of the enzymatic macrocyclization product to the linear cleavage product.

PatGmac, the macrocyclization domain of PatG from the biosynthetic pathway of patellamides,<sup>[12]</sup> recognizes octapeptides containing a thiazoline ring at the C-terminal adopting a *cis* conformation followed by the recognition sequence AYD(GE).<sup>[12b]</sup> PatGmac is as a versatile tool capable of tolerating a range of different precursor peptides;<sup>[11,13]</sup> a proline can replace the thiazoline ring,<sup>[11,13a]</sup> the amino acids of the core peptide are variable almost at will,<sup>[11,13]</sup> PatGmac has been reported to make macrocycles of up to 22 amino acids,<sup>[13b]</sup> and can tolerate non-natural amino acids and triazole amide bond mimics.<sup>[11a,13a]</sup> In an elegant paper by Schmidt and co-workers,<sup>[11b]</sup> the chemistry of the *N*-terminus of the core peptide was exten-

[a] Dr. E. Oueis, Dr. B. Nardone, Prof. N. J. Westwood, Prof. J. H. Naismith  
School of Chemistry, University of St Andrews  
BSRC, North Haugh, St Andrews KY16 9ST (UK)  
E-mail: njw3@st-andrews.ac.uk  
naismith@st-andrews.ac.uk

[b] Prof. J. H. Naismith  
State Key Laboratory of Biotherapy  
Sichuan University (P. R. China)

[c] Prof. M. Jaspars  
Marine Biodiscovery Centre, Department of Chemistry  
University of Aberdeen, Old Aberdeen AB24 3UE (UK)

Supporting Information and the ORCID identification number(s) for the author(s) of this article can be found under <http://dx.doi.org/10.1002/open.201600134>.

© 2016 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

sively evaluated to study the cyclization patterns of lysine and its analogues. Depending on the length of the sequence and the amino acid side chain, both backbone and side-chain cyclization were observed along with a linear product. The synthetic utility of PatGmac derives precisely from its promiscuity towards the core sequence. The cyclic dehydratases, PatD and LynD, which make thiazoline and oxazoline rings from cysteine and serine (threonine), respectively, have also been shown to tolerate the presence of non-natural amino acids,<sup>[13a]</sup> but it is not known how they tolerate non-amino acid groups. Here, we have used enzymes from cyanobactin biosynthesis to generate eighteen hybrid peptidic/non-peptidic macrocycles that explore a wide range of chemical variation. Nine of these macrocycles (**1–6, 7a–c**, Figure 2) were made at sufficient scale to allow full characterization.

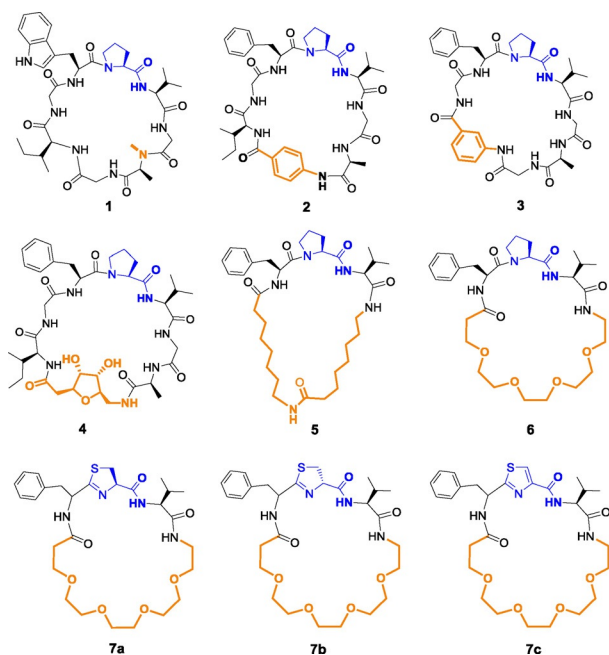


Figure 2. Structures of fully characterized macrocycles **1–6**, **7a–c**.

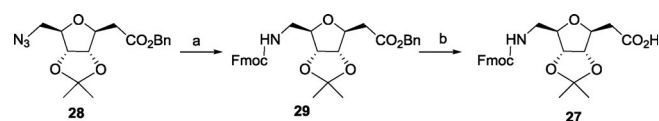
Nineteen analogues (**8–26**) were synthesized based on a common precursor peptide VGAGIGFP, in which one or more amino acids were substituted with non-natural, non-amino acid scaffolds. All but one of the substrates finished with a Pro in the core peptide (one had a Cys) and all possessed a C-terminal AYD, the minimal recognition sequence for PatGmac, or AYD-Doc (8-amino-3,6-dioxaoctanoic acid; Table 1). Peptides **8–10** contain a one-residue substitution involving small non-natural amino acids  $\beta$ -Ala, GABA, and Doc (8-amino-3,6-dioxaoctanoic acid), respectively. Peptide **11** has an *N*-Me amino acid. Peptides **12–18** contain either an isomer of amino benzoic acid (Abz) at different positions or a ribose-derived sugar amino acid (Rib). Finally, peptides **19–26** represent the hybrid peptides, where a polyketide chain 7-aminoheptanoic acid (7Ahp), 8-amino-octanoic acid (8Aoc), or a polyether amino acid (PEG)<sub>4</sub> is introduced at different positions of the peptide. Peptide **26** has a Cys, which requires a heterocyclization reaction

Table 1. Synthetic hybrid precursor peptides and their corresponding PatGmac macrocyclization products.

Peptide	Sequence <sup>[a]</sup>	Product <sup>[b]</sup>	Macrocycle <sup>[c]</sup>
<b>8</b>	VGA- <b><math>\beta</math>-Ala</b> -IGWPAYD-Doc	C	
<b>9</b>	VGA- <b>GABA</b> -IGWPAYD-Doc	C	
<b>10</b>	VGA- <b>Doc</b> -IGWPAYD-Doc	C	
<b>11</b>	VG( <b><i>N</i>-Me</b> )AGIGWPAYD-Doc	C	1
<b>12</b>	VGA-2- <b>Abz</b> -IGFPAYD	C	
<b>13</b>	VGA-3- <b>Abz</b> -IGFPAYD	C	
<b>14</b>	VGA-4- <b>Abz</b> -IGFPAYD	C	2
<b>15</b>	VGAG-2- <b>Abz</b> -GFPAYD	C	
<b>16</b>	VGAG-3- <b>Abz</b> -GFPAYD	C	3
<b>17</b>	VGAG-4- <b>Abz</b> -GFPAYD	C	
<b>18</b>	VGA- <b>Rib</b> -IGFPAYD	C	4
<b>19</b>	<b>8Aoc</b> -AGIGFPAYD	L	
<b>20</b>	V- <b>7Ahp</b> -GAGFPAYD	C	
<b>21</b>	VGAG- <b>7Ahp</b> -PAYD	NR	
<b>22</b>	VGAG- <b>8Aoc</b> -FPAYD	C	
<b>23</b>	V- <b>8Aoc</b> -FPAYD	L	
<b>24</b>	V- <b>8Aoc</b> - <b>8Aoc</b> -FPAYD	C	5
<b>25</b>	V-(PEG) <sub>4</sub> -FPAYD	C	6
<b>26</b>	V-(PEG) <sub>4</sub> -FCAYD	C	7

[a] Highlighted is the minimal recognition sequence where the peptide is cleaved. Non-natural scaffolds in the sequence are in bold. [b] All successful enzymatic reactions afforded exclusively either cyclic or linear products, as detailed. Doc = 8-amino-3,6-dioxaoctanoic acid; GABA =  $\gamma$ -aminobutyric acid; Abz = aminobenzoic acid; Rib = ribose sugar amino acid; 8Aoc = 8-amino-octanoic acid; 7Ahp = 7-aminoheptanoic acid; PEG = polyethylene glycol; C = macrocyclic; L = linear; NR = no reaction. [c] The corresponding isolated macrocycle number.

prior to macrocyclization. The peptides were all synthesized through solid-phase peptide synthesis by using the Fmoc strategy and commercially available reagents (see the Supporting Information). The fully protected Fmoc-sugar amino acid **27** was synthesized in a similar approach to the one reported in the literature<sup>[14]</sup> for the synthesis of the Boc-protected analogue starting from *D*-ribose. Minor changes, such as the use of benzyl ester protection in the synthesis, were required to produce the desired compound **27** in good yields (see Scheme 1 and the Supporting Information). The azide of com-



Scheme 1. Synthesis of ribose sugar amino acid derivative **27**. Reagents and conditions: a)  $\text{Ph}_3\text{P}$ , THF/ $\text{H}_2\text{O}$ , rt, O/N (i); FmocOSu, DIEA,  $\text{CH}_2\text{Cl}_2$ , rt, O/N, 56% (ii); b)  $\text{Et}_3\text{SiH}$ , Pd/C, MeOH, rt, 10 min, 60%.

pound **28** was selectively reduced under the Staudinger reaction conditions to afford the corresponding amine, which was immediately Fmoc protected to yield **29**. Benzyl deprotection<sup>[15]</sup> gave **27** in 62% yield. Compound **27** was used for peptide synthesis, and deprotection of the alcohols was accomplished during the final acidic cleavage of the peptide.

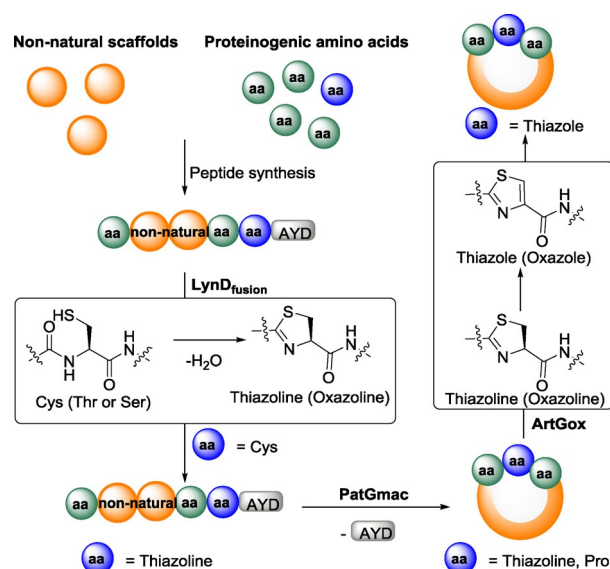
Substrates **12–26** were purified before reaction with the enzyme(s), and substrates **8–11** were used without any further purification after cleavage from the resin.

The macrocyclization reactions were set up on small scale (< 1 mL reactions) with PatGmac in bicine buffer at pH 8.1 and 37 °C. The reactions were followed by MALDI-MS to monitor progress. MS–MS fragmentation data established the structure of the different products with the presence of the PV dipeptide fragment being indicative of macrocycle formation. Table 1 summarizes the results of the reactions; in brief, only peptide **21** gave no reaction, whereas **20** and **23** yielded a linear cleaved product. The reaction rate for these transformations was still slow, with reaction times ranging between 2–4 weeks. Nine representative cyclic hybrid peptides (**1–6**, **7a–c**; Figure 2), products of peptides **11**, **14**, **16**, **18**, **24**, **25**, and **26**, were purified from a large-scale reaction and fully characterized by using NMR spectroscopy. One of the products of peptide **26** (**7a**) was subsequently oxidized.

Substrates with  $\beta$ -Ala and GABA (**8**, **9**) processed essentially as native peptides, perhaps unsurprisingly because these are very similar to amino acids. More interestingly, a longer chain in that position, namely Doc (**10**), also gave only the macrocycle. The inclusion of a *N*-methylated amino acid in peptide **11** proceeded to afford the corresponding macrocycle **1**. The incorporation of a rigid aromatic scaffold, 2-, 3-, or 4-aminobenzoic acid at either position 4 or 5 of the core peptide (**12–17**), proceeded normally. Remarkably, the presence of a sugar amino acid (Rib, **18**), known to cause conformational restraints in peptidic analogues,<sup>[16]</sup> afforded exclusively the macrocyclic product **4**.

Schmidt and co-workers have previously shown the tolerance of PatGmac for polyketide-like scaffolds such as alkyl chains as part of the core peptide.<sup>[11b]</sup> Peptide **19**, having a *N*-terminal 8Aoc unit, gave only linear peptide, whereas **20**, which has a *N*-terminal Val preceding the alkyl chain of 7Ahp, yielded exclusively the cyclic product. Comparing substrates **19** and **20** indicated that an amino acid at the *N*-terminus was important with our sequence for efficient macrocyclization (Table 1). Schmidt and co-workers,<sup>[11b]</sup> showed both cyclic and linear products were obtained with an *N*-terminal amino-alkyl chain of a different peptidic sequence (Figure 1). Peptide **21** with the 7Ahp chain immediately before the required Pro in the core peptide failed to react, but when a Phe was introduced in between 8Aoc and Pro (**22**), the macrocycle was obtained. It seems that an additional amino acid is required *N*-adjacent to the terminal Pro or thiazoline of the core peptide for it to be processed by the enzyme (see Table 1 and the Supporting Information).

We observed a length requirement for the macrocyclization reaction. Compound **23**, which has an 8Aoc unit between a *N*-terminal Val and the Phe–Pro at the C-terminus, gave only a linear product, whereas **24** with two molecules of 8Aoc, yielded only macrocycle **5**. Substrate **25** was designed with a pegylated chain (PEG)<sub>4</sub> and yielded the corresponding macrocycle **6**. Full conversion of substrates **25** and **26** was difficult to achieve, even after long incubation times (4 weeks, adding more enzyme), reaching only to about 70% conversion to macrocycle (the remainder was unreacted starting material). In **26**, the Pro was replaced with a Cys, which was converted to thiazoline with the engineered LynD enzyme<sup>[17]</sup> (establishing



Scheme 2. General reaction scheme of the enzymatic transformations.

this enzyme tolerated PEG) and subsequently macrocyclized to **7** (Scheme 2).

Compounds **1** and **2** were present as a single conformer, whereas compounds **3–6** were all present as a mixture of two conformers in approximately 2:1 (**3**, **4**) or 3:2 (**5**, **6**) ratios, as confirmed by exchange spectroscopy (EXSY) experiments (see the Supporting Information). Compound **7**, containing the thiazoline ring, shows two separable peaks in the HPLC spectrum, in an 8:2 (**7a/7b**) ratio. Both compounds have identical HRMS and MS–MS fragmentation data. The NMR spectra of these two compounds are, nonetheless, different, but confirm the same peptidic sequence [cyclic V(PEG)<sub>4</sub>FThz]; however, an EXSY experiment showed no exchange. We postulate that **7a** and **7b** are epimers. As thiazoline with the natural (*R*) *L*-configuration is a substrate for the oxidase enzyme,<sup>[12d]</sup> **7a** and **7b** were incubated in the presence of this enzyme overnight. **7a** afforded compound **7c**, which was isolated, purified, and characterized by NMR as the thiazole, confirming the oxidase enzyme also tolerates non-peptide groups (Scheme 2). **7b** was, however, not enzymatically oxidized. We suggest the thiazoline stereo-center of **7a** is (*R*) *L*, whereas it has inverted to a (*S*) *D*-configuration in **7b**. This has precedent; the thiazoline moiety of piperidine thiopeptides is known to be spontaneously epimerized from the (*R*) *L*- to the (*S*) *D*-configuration.<sup>[18]</sup>

Natural peptide hybrid macrocycles have interesting and useful biological properties,<sup>[6c]</sup> but are, of course, limited to natural building blocks. The routine synthesis of macrocycles composed of peptide and non-peptide building blocks has wide potential in medicinal chemistry, allowing access to novel scaffolds and the fine-tuning of properties. The patellamide biosynthetic pathway has enzymes that catalyze a number of valuable chemical reactions including heterocyclization, oxidation, and macrocyclization. The enzymes in the pathway largely separate recognition from catalysis, and are thus assumed to have a high degree of promiscuity.<sup>[12c]</sup> We have used these enzymes to synthesize diverse macrocycles containing methylat-



ed amide, sugars, non-natural amino acids, aromatic rings, thiazole, thiazoline, polyether moieties, and polyketide-like chains. The overall yield of the enzymatic processes (starting from substrate) varied from 27 to 66% (Table S2). More interestingly, the enzymes can synthesize a product with just three amino acids. Thus, we have established that these enzymes are tools that, when combined with chemical synthesis of substrates, can produce highly modified, complex, and diverse peptide/non-peptide hybrid macrocycles.

## Acknowledgements

We thank Dr. G. Mann and Dr. A. Bent for supplying the enzymes and the useful discussions, and Dr. T. Lebl for the useful NMR discussions. This work was supported by the European Research Council (339367), UK Biotechnology and Biological Sciences Research Council (K015508/1), the Wellcome Trust [TripleTOF 5600 mass spectrometer (094476), the MALDI TOF-TOF Analyzer (079272A/A), 700 NMR], and the EPSRC UK National Mass Spectrometry Facility at Swansea University. J.H.N., M.A.J., and N.J.W. are named on patents that have been filed by St. Andrews and Aberdeen Universities to commercialize the enzymes (PatG) and (LynD) used in the study. No income derives from the patent.

**Keywords:** biotransformation • cyanobactin • hybrid macrocycles • macrocyclization • peptides

- [1] a) D. A. Dias, S. Urban, U. Roessner, *Metabolites* **2012**, 2, 303; b) H. F. Ji, X. J. Li, H. Y. Zhang, *EMBO rep.* **2009**, 10, 194.
- [2] A. L. Harvey, R. Edrada-Ebel, R. J. Quinn, *Nat. Rev. Drug Discovery* **2015**, 14, 111.
- [3] a) D. G. I. Kingston, *J. Nat. Prod.* **2011**, 74, 496; b) D. J. Newman, G. M. Cragg, *J. Nat. Prod.* **2012**, 75, 311.
- [4] J. Hong, *Curr. Opin. Chem. Biol.* **2011**, 15, 350.
- [5] a) E. M. Driggers, S. P. Hale, J. Lee, N. K. Terrett, *Nat. Rev. Drug Discovery* **2008**, 7, 608; b) S. H. Joo, *Biomol. Ther. (Seoul)* **2012**, 20, 19; c) D. J. Newman, G. M. Cragg in *Macrocycles in Drug Discovery* (Eds.: J. R. Murphy, C. J. Harris), The Royal Society of Chemistry, London, **2015**, pp. 1.
- [6] a) A. T. Bockus, C. M. McEwen, R. S. Lokey, *Curr. Top. Med. Chem.* **2013**, 13, 821; b) X. Yu, D. Sun, *Molecules* **2013**, 18, 6230; c) L. A. Wessjohann, E. Ruijter, D. Garcia-Rivera, W. Brandt, *Mol. Diversity* **2005**, 9, 171; d) A. K. Yudin, *Chem. Sci.* **2015**, 6, 30.
- [7] a) L. Du, C. Sánchez, B. Shen, *Metab. Eng.* **2001**, 3, 78; b) J. K. Nunnery, E. Mevers, W. H. Gerwick, *Curr. Opin. Biotechnol.* **2010**, 21, 787.
- [8] a) Y.-U. Kwon, T. Kodadek, *Chem. Biol.* **2007**, 14, 671; b) H. Lodish, A. Berk, S. L. Zipursky, P. Matsudaira, D. Baltimore, J. Darnell, *Molecular Cell Biology*, 4th ed., W. H. Freeman, New York, **2000**.
- [9] a) T. R. White, C. M. Renzelman, A. C. Rand, T. Rezai, C. M. McEwen, V. M. Gelev, R. A. Turner, R. G. Lington, S. S. F. Leung, A. S. Kalgutkar, J. N. Bauman, Y. Zhang, S. Liras, D. A. Price, A. M. Mathiowetz, M. P. Jacobson, R. S. Lokey, *Nat. Chem. Biol.* **2011**, 7, 810; b) J. Chatterjee, B. Laufer, H. Kessler, *Nat. Protocols* **2012**, 7, 432.
- [10] a) T. Sonoda, T. Nogami, J. Oishi, M. Murata, T. Niidome, Y. Katayama, *Bioconjugate Chem.* **2005**, 16, 1542; b) T. Uhlig, T. Kyprianou, F. G. Martinielli, C. A. Oppici, D. Heiligers, D. Hills, X. R. Calvo, P. Verhaert, *EuPA Open Proteomics* **2014**, 4, 58.
- [11] a) E. Oueis, M. Jaspars, N. J. Westwood, J. H. Naismith, *Angew. Chem. Int. Ed.* **2016**, 55, 5842; *Angew. Chem.* **2016**, 128, 5936; b) J. A. McIntosh, C. R. Robertson, V. Agarwal, S. K. Nair, G. W. Bulaj, E. W. Schmidt, *J. Am. Chem. Soc.* **2010**, 132, 15499.
- [12] a) M. Jaspars, *Chem. Commun.* **2014**, 50, 10174; b) J. Koehnke, A. Bent, W. E. Houssen, D. Zollman, F. Morawitz, S. Shirran, J. Vendome, A. F. Nneoyiegbe, L. Trembleau, C. H. Botting, M. C. M. Smith, M. Jaspars, J. H. Naismith, *Nat. Struct. Mol. Biol.* **2012**, 19, 767; c) J. Koehnke, A. F. Bent, W. E. Houssen, G. Mann, M. Jaspars, J. H. Naismith, *Curr. Opin. Struct. Biol.* **2014**, 29, 112; d) E. W. Schmidt, J. T. Nelson, D. A. Rasko, S. Sudek, J. A. Eisen, M. G. Haygood, J. Ravel, *Proc. Natl. Acad. Sci. USA* **2005**, 102, 7315.
- [13] a) E. Oueis, C. Adamson, G. Mann, H. Ludewig, P. Redpath, M. Migaud, N. J. Westwood, J. H. Naismith, *ChemBioChem* **2015**, 16, 2646; b) D. Sardar, Z. Lin, E. W. Schmidt, *Chem. Biol.* **2015**, 22, 907; c) W. E. Houssen, A. F. Bent, A. R. McEwan, N. Pieiller, J. Tabudravu, J. Koehnke, G. Mann, R. I. Adaba, L. Thomas, U. W. Hawas, H. Liu, U. Schwarz-Linek, M. C. M. Smith, J. H. Naismith, M. Jaspars, *Angew. Chem. Int. Ed.* **2014**, 53, 14171; *Angew. Chem.* **2014**, 126, 14395.
- [14] R. M. Van Well, H. S. Overkleeft, M. Overhand, E. Vang Carstenen, G. A. van der Marel, J. H. van Boom, *Tetrahedron Lett.* **2000**, 41, 9331.
- [15] C. Bughin, G. Masson, J. Zhu, *J. Org. Chem.* **2007**, 72, 1826.
- [16] M. D. P. Risseuw, M. Overhand, G. W. J. Fleet, M. I. Simone, *Tetrahedron: Asymmetry* **2007**, 18, 2001.
- [17] J. Koehnke, G. Mann, A. F. Bent, H. Ludewig, S. Shirran, C. Botting, T. Lebl, W. E. Houssen, M. Jaspars, J. H. Naismith, *Nat. Chem. Biol.* **2015**, 11, 558.
- [18] S. Schoof, H.-D. Arndt, *Chem. Commun.* **2009**, 7113.

Received: October 27, 2016

Published online on December 13, 2016