ROYAL SOCIETY
OF CHEMISTRY

COMMUNICATION

Check for updates

Cyclic peptide production using a macrocyclase with enhanced substrate promiscuity and relaxed recognition determinants†

Cristina N. Alexandru-Crivac,[ab] Christian Umeobika,[a] Niina Leikoski,[c] Jouni Jokela, [ID][c] Kirstie A. Rickaby,[a] André M. Grilo,[ab] Peter Sjö, [ID][d] Alleyn T. Plowright,[e] Mohannad Idress,[ab] Eike Siebs,[a] Ada Nneoyi-Egbe,[a] Matti Wahlsten, [ID][c] Kaarina Sivonen, [ID][c] Marcel Jaspars, [ID][a] Laurent Trembleau, [ID]*[a] David P. Fewer [ID]*[c] and Wael E. Houssen [ID]*[abf]

Macrocyclic peptides have promising therapeutic potential but the scaling up of their chemical synthesis is challenging. The cyanobactin macrocyclase PatG$_{mac}$ is an efficient tool for production but is limited to substrates containing 6–11 amino acids and at least one thiazoline or proline. Here we report a new cyanobactin macrocyclase that can cyclize longer peptide substrates and those not containing proline/thiazoline and thus allows exploring a wider chemical diversity.

Macrocyclic peptides show great promise as therapeutics because they have an increased target binding affinity and selectivity, are more stable against proteolytic enzymes and often have higher membrane permeability when compared to their linear counterparts.[1,2] These scaffolds can modulate therapeutic targets that involve extended binding surfaces such as protein–protein interactions[3] known to be challenging for small molecule drugs.[4–6] However, macrocyclic peptides are under-exploited due to challenges in their chemical production on a large scale. Several chemical ring closing strategies have been proposed including metathesis, click chemistry, metal-templated chelation and the use of auxiliaries but all have variable success rates and inherent limitations.[7]

[a] Marine Biodiscovery Centre, Department of Chemistry, University of Aberdeen, Meston Walk, Aberdeen AB24 3UE, UK. E-mail: l.trembleau@abdn.ac.uk, w.houssen@abdn.ac.uk

[b] Institute of Medical Sciences, University of Aberdeen, Aberdeen AB25 2ZD, UK

[c] Microbiology and Biotechnology Division, Department of Food and Environmental Sciences, University of Helsinki, P.O.Box 56, Biocenter 1, Viikinkaari 9, 00014, Finland. E-mail: david.fewer@helsinki.fi

[d] Medicinal Chemistry, Respiratory, Inflammation & Autoimmunity, Innovative Medicines and Early Development Biotech Unit, AstraZeneca, Pepparedsleden 1, Mölndal, 43183, Sweden

[e] Medicinal Chemistry, Cardiovascular and Metabolic Diseases Innovative Medicines and Early Development Biotech Unit, AstraZeneca, Pepparedsleden 1, Mölndal, 43183, Sweden

[f] Pharmacognosy Department, Faculty of Pharmacy, Mansoura University, Mansoura 35516, Egypt

† Electronic supplementary information (ESI) available: Experimental section, Fig. S1–S60 and Tables S1–S3. See DOI: 10.1039/c7cc05913b
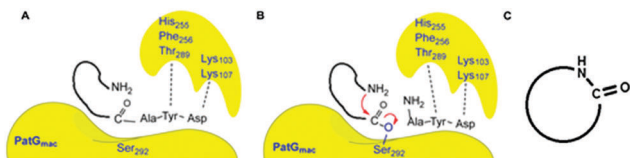
Biochemical approaches include sortase-mediated ligation,[8] split-intein circular ligation of peptides and proteins[9,10] and the use of macrocyclases from the biosynthetic pathways of ribosomally synthesized and post-translationally modified peptides (RiPPs).[11] The former approach incorporates a pentapeptide in the product, while the second cannot be used for substrates containing more than one non-natural amino acid.[12] Several macrocyclases from RiPP classes have been identified including butelase-1,[13] GmPOPB,[14] and PCY1[15] which are involved in the formation of cyclotides, α-amanitin, and segetalin A respectively. The difficulty in the recombinant production or the requirement for long recognition sequences has impeded the wider application of these enzymes. Other RiPP macrocyclases are derived from azole-containing cyclic peptide cyanobactin pathways.

PatG$_{mac}$ from the patellamide pathway has so far been the only structurally characterized cyanobactin macrocyclase.[16,17] The enzyme is tolerant to changes in the substrate sequence[18] and has been successfully used to produce a library of macrocyclic peptides in vitro.[19] PatG$_{mac}$ is a subtilisin-like protease that recognizes a three-residue signal at the C-terminus of the substrate, cleaves it off and forms an acyl complex with the residue at the P1 position. The enzyme accepts the substrate in a bent conformation in which the N-terminal amino group is in close proximity to the acyl complex (Fig. 1).[20] This conformation is facilitated by the conformational properties of the Pro or thiazoline residues at the P1 position. This may explain why, for cyanobactins, a Pro or a heterocyclized Cys residue is always found at the P1 position.[21]

However, PatG$_{mac}$ can cyclize only substrates containing 6–11 amino acids and it requires that the substrate have Pro or thiazoline before the recognition signal which is incorporated in the cyclic product.[16] Here we report a new cyanobactin macrocyclase that can cyclize longer substrates and accepts (Cys)- and (Ser)-pseudoprolines at position P1. Oxazolidine-based pseudo-proline rings can adopt an s-cis amide conformation analogous to Pro/thiazoline residues.

Genome mining studies identified a cryptic cyanobactin gene cluster in the draft genome of *Oscillatoria* sp. PCC 6506.[22,23]

Fig. 1 A cartoon showing the substrate binding with PatG$_{mac}$ and the steps of enzyme-mediated cyclization. (A) Binding of the enzyme with the recognition signal. (B) Cleavage of the recognition signal and formation of an acyl complex with the residue at position P1. (C) Formation of an amide bond and cyclization of the peptide.

The 12.9 kb cryptic *osc* gene cluster encodes two precursor peptides and six biosynthetic enzymes (Fig. 2 and Table S1, ESI†). Analysis of the strain's methanolic extracts by LC-ITMS/QTOF revealed that OscE2 encodes a 17-mer prenylated cyclic peptide in which Cys residues are heterocyclized and oxidised to thiazoles (Fig. S1, ESI†). The monoisotopic mass of the protonated putative cyanobactin was 1896.7287 m/z (calc. for $C_{87}H_{110}N_{21}O_{22}S_3$ m/z 1896.72909, Δ −0.2 ppm). The presence of an isoprenoid unit was verified by the neutral loss of 68 Da in the MS and MS^2 spectra. The position of prenylation (on Trp between Thr and Gly) was determined based on the presence of the product ions 9–12 in the QTOF MS^2 spectrum (Table S2, ESI†). The second precursor peptide OscE1 encodes a 25-mer cyclic peptide containing three thiazoles and no prenylated residue. A mass corresponding to the double protonated molecule m/z 1275.9993 (calc. for $C_{113}H_{149}N_{29}O_{34}S_3$ m/z 1275.99864, Δ 0.5 ppm) was detected and the fragmentation is shown in Table S2 (ESI†).

Oscillatoria sp. PCC 6506 cells were labelled with ^{15}N and ^{34}S in order to further confirm the structures of the putative cyanobactins (Fig. S2 and S3, ESI†). ^{15}N-labelled protonated cyanobactins A and B show an increase of 21 and 29 mass units, respectively, equating to the same number of nitrogen atoms in these molecules (Fig. S2, ESI†). Similarly ^{34}S-labelled protonated cyanobactins A and B show an increase of 6 mass units equating to 3 sulphur atoms in both molecules (Fig. S3, ESI†). Altogether all the results are in full agreement with the structures predicted from the gene cluster. We named these cyanobactins oscillacyclamides A and B respectively (Fig. 2).

OscG$_{mac}$ was expressed in *E. coli* and purified (Fig. S4 and S5, ESI†). A series of 19 linear peptide substrates (Table 1) was used to test the enzyme substrate specificity. Some of these substrates were made as part of our ongoing efforts to chemo-enzymatically produce the anti-HIV depsipeptide homophymine A (Fig. S53, ESI†) and to test the substrate specificity of different cyanobactin macrocyclases. The peptide substrates were designed to include the PatG$_{mac}$ recognition signal (Ala-Tyr-Asp) or the PirG$_{mac}$ (from the piricyclamide pathway[24]) signal (Phe-Ala-Gly-Asp) at the C-terminus. The recognition signal of PatG$_{mac}$ resembles those present in OscE1 and OscE2 for OscG$_{mac}$ (Ser-Phe/Tyr-Asp) and follows the order of a small amino acid, an aromatic residue and an acidic residue. The crystal structure of PatG$_{mac}$[16] showed that the substrate P1′ alanine Cα and side chain make only a few hydrophobic interactions with the enzyme and sits in a cavity that appears to be large enough for bulkier residues. The natural recognition signal of the homologous macrocyclase, PirG$_{mac}$, contains a bulkier residue (Phe) in P1′. In our hands, PatG$_{mac}$ did not process any of the substrates containing the recognition signal of its homologue PirG$_{mac}$. In comparison, OscG$_{mac}$ shows higher substrate promiscuity and could process substrates containing both of these recognition signals (Table 1). Cyclic products of the enzymatic reactions contain 6–30 amino acids and each of the 20 canonical amino acids was incorporated at least once into the products.

The enzyme tolerability for substrates containing D-amino acids in positions P2, P6 and P7 was investigated (substrates 4–9, 11 and 12; Table 1). A maximum of three D-amino acids per substrate was incorporated that are neutral, small hydrophobic, basic or acidic. Remarkably, all substrates except those containing D-amino acids in the P2 position were processed by OscG$_{mac}$ (Table 1). This may be due to the fact that the P2 position is close to position P1 which is conformationally critical to macrocyclisation. Having a D-stereocentre in position P2, thus, may prevent the substrate adopting the bent conformation.

In effect, the prerequisite by PatG$_{mac}$ for having Pro or thiazoline in position P1 limits the diversity of the cyclic peptides that can be generated. In order to overcome this limitation, sequences containing (Cys)- and (Ser)-pseudoprolines at P1 were tested (sequences 2, 17 and 18; Table 1). None of these sequences were processed by PatG$_{mac}$. However, OscG$_{mac}$ was able to macrocyclise the des(dimethyl) pseudoproline-containing peptides (sequences 2 and 18; Table 1) as observed using LC-MS analyses (Fig. S8 and S9 and S49 and S50, ESI†). Peptide 17, which contains a dimethyl-pseudoproline known to be the preferred substrate for facile reversion to serine,[25–27] was not processed most probably for steric reasons.

We tried to estimate the catalytic rate of OscG$_{mac}$ but obtaining reliable kinetics is complicated owing to the fact that OscG$_{mac}$ is present at >20 mol%[28] and that the catalytic rate of the enzyme is substrate dependent. We are working on understanding the factors affecting the rate of cyclisation using a mixed experimental and theoretical approach.[20]

We aligned the sequence of OscG$_{mac}$ with those of 7 other cyanobactin macrocyclases listed in Table S3 (ESI†) (Fig. S54, ESI†) and created a Phyre2 model[29] for OscG$_{mac}$ using the crystal structure of PatG$_{mac}$ (PDB 4AKS) as a template (Fig. S55, ESI†).
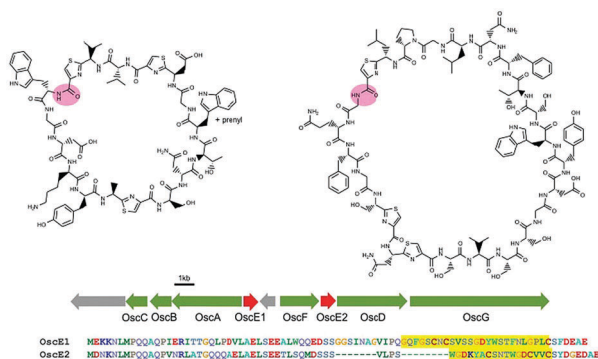

Fig. 2 The *osc* gene clusters and sequences of OscE1 and OscE2 precursor peptides. Structures of oscillacyclamide A (left) and oscillacyclamide B (right) are shown. Core sequences are highlighted in yellow.

Table 1 Sequences of the synthetic peptide substrates used for testing the functionality and promiscuity of OscG$_{mac}$ and the resultant cyclic products

Peptide	Sequence	Product	Cyclization with OscG$_{mac}$	Figure (ESI)
1	H-Asn-Glu-Phe-Met-Gln-Thr-Gly-Ser-Tyr-Ser-Gly-Pro-Ala-Tyr-Asp-Gly-OH	Cyclo-[Asn-Glu-Phe-Met-Gln-Thr-Gly-Ser-Tyr-Ser-Gly-Pro-]	Yes	S6 and S7
2	H-Val-Gly-Ala-Gly-Ile-Gly-Phe-Ser(ψPro)-Ala-Tyr-Asp-Gly-OH	Cyclo-[Val-Gly-Ala-Gly-Ile-Gly-Phe-Ser(ψPro)-]	Yes	S8 and S9
3	Fmoc-Ala(Amino)-Ser-Lys-Leu-Gln-Ile-Asp-Pro-Ala-Tyr-Asp-Gly-OH	Cyclo-[Dpr(Fmoc)-Ser-Lys-Leu-Gln-Ile-Asp-Pro-]	Yes	S10–S12
4	Fmoc-Ala(Amino)-Ser-Lys-Leu-Gln-Ile-D-Asp-Pro-Ala-Tyr-Asp-Gly-OH	Cyclo-[Dpr(Fmoc)-Ser-Lys-Leu-Gln-Ile-D-Asp-Pro-]	No	S13 and S14
5	Fmoc-Ala(Amino)-Ser-D-Lys-Leu-Gln-Ile-Asp-Pro-Ala-Tyr-Asp-Gly-OH	Cyclo-[Dpr(Fmoc)-Ser-D-Lys-Leu-Gln-Ile-Asp-Pro-]	Yes	S15–S18
6	Fmoc-Ala(Amino)-D-Ser-Lys-Leu-Gln-Ile-Asp-Pro-Ala-Tyr-Asp-Gly-OH	Cyclo-[Dpr(Fmoc)-D-Ser-Lys-Leu-Gln-Ile-Asp-Pro-]	Yes	S19–S22
7	Fmoc-Ala(Amino)-Ser-D-Lys-Leu-Gln-Ile-D-Asp-Pro-Ala-Tyr-Asp-Gly-OH	Cyclo-[Dpr(Fmoc)-Ser-D-Lys-Leu-Gln-Ile-D-Asp-Pro-]	No	S23 and S24
8	Fmoc-Ala(Amino)-D-Ser-Lys-Leu-Gln-Ile-D-Asp-Pro-Ala-Tyr-Asp-Gly-OH	Cyclo-[Dpr(Fmoc)-D-Ser-Lys-Leu-Gln-Ile-D-Asp-Pro-]	No	S25 and S26
9	Fmoc-Ala(Amino)-D-Ser-D-Lys-Leu-Gln-Ile-Asp-Pro-Ala-Tyr-Asp-Gly-OH	Cyclo-[Dpr(Fmoc)-D-Ser-D-Lys-Leu-Gln-Ile-Asp-Pro-]	Yes	S27–S30
10	Fmoc-Ala(Amino)-Ser-Lys-Leu-Gln-Ile-Ala-Pro-Ala-Tyr-Asp-Gly-OH	Cyclo-[Dpr(Fmoc)-Ser-Lys-Leu-Gln-Ile-Asp-Pro-]	Yes	S31–S34
11	Fmoc-Ala(Amino)-Ser-Lys-Leu-Gln-Ile-D-Ala-Pro-Ala-Tyr-Asp-Gly-OH	Cyclo-[Dpr(Fmoc)-Ser-Lys-Leu-Gln-Ile-D-Ala-Pro-]	No	S35 and S36
12	Fmoc-Ala(Amino)-D-Ser-D-Lys-Leu-Gln-Ile-D-Asp-Pro-Ala-Tyr-Asp-Gly-OH	Cyclo-[Dpr(Fmoc)-D-Ser-D-Lys-Leu-Gln-Ile-D-Asp-Pro-]	No	S37 and S38
13	H-His-Gln-Trp-Gly-Trp-Leu-Val-Gly-Gly-Thr-Asp-Pro-Phe-Ala-Gly-Asp-OH	Cyclo-[His-Gln-Trp-Gly-Trp-Leu-Val-Gly-Gly-Thr-Asp-Pro-]	Yes	S39 and S40
14	H-Asp-Trp-Gly-Thr-Phe-Cys-Val-Gln-Glu-Asp-Gly-Glu-Gly-Asn-Cys-Lys-Glu-Trp-Tyr-Glu-Leu-Pro-Phe-Ala-Gly-Asp-Asp-Ala-Glu-OH	Cyclo-[Asp-Trp-Gly-Thr-Phe-Cys-Val-Gln-Glu-Asp-Gly-Glu-Gly-Asn-Cys-Lys-Glu-Trp-Tyr-Glu-Leu-Pro-] (disulfide bond)	Yes	S41 and S42
15	H-Val-Gly-Ala-Gly-Ile-Gly-Phe-Pro-Ala-Tyr-Asp-OH	Cyclo-[Val-Gly-Ala-Gly-Ile-Gly-Phe-Pro-]	Yes	S43 and S44
16	H-Val-Ala-Gly-Ile-Gly-Phe-Pro-Ala-Tyr-Asp-OH	Cyclo-[Val-Ala-Gly-Ile-Gly-Phe-Pro-]	Yes	S45 and S46
17	H-Val-Gly-Ala-Gly-Ile-Gly-Phe-Ser(ψ^{Me,Me}Pro)-Ala-Tyr-Asp-Gly-OH	Cyclo-[Val-Gly-Ala-Gly-Ile-Gly-Phe-Ser(ψ^{Me,Me}Pro)-]	No	S47 and S48
18	H-Val-Ala-Gly-Ile-Gly-Phe-Cys(ψPro)-Ala-Tyr-Asp-OH	Cyclo-[Val-Ala-Gly-Ile-Gly-Phe-Cys(ψPro)-]	Yes	S49 and S50
19	H-Gly-Val-Cys-Gly-Glu-Thr-Cys-Val-Gly-Gly-Thr-Cys-Asn-Thr-Pro-Gly-Cys-Thr-Cys-Ser-Trp-Pro-Val-Cys-Thr-Arg-Asn-Gly-Leu-Pro-Ala-Tyr-Asp-Gly-Gly-Leu-Glu-His-His-His-His-His-His-OH (two disulfide bonds)	Cyclo-[Gly-Val-Cys-Gly-Glu-Thr-Cys-Val-Gly-Gly-Thr-Cys-Asn-Thr-Pro-Gly-Cys-Thr-Cys-Ser-Trp-Pro-Val-Cys-Thr-Arg-Asn-Gly-Leu-Pro-] (two disulfide bonds)	Yes	S51 and S52

Dpr = 2,3-diaminopropionic acid.

Our alignment shows that the key residues in PatG$_{mac}$ are highly conserved in all of these enzymes. These residues according to the PatG$_{mac}$ sequence are as follows: (1) the Asp57, His127, and Ser292 that form the catalytic triad; (2) the Lys103, Lys107 and Phe256 that make interactions with the Ala-Tyr-Asp recognition signal; (3) the Asn226 that stabilizes the tetrahedral intermediate and (4) the Cys194, Cys233, Cys332 and Cys343 that form two critical disulfide bridges. PatG$_{mac}$'s Cys194–Cys233 disulfide bond bridges two loops resulting in the protrusion of three bulky residues; Met169, Phe193 and Arg195 close to the active site.[16] These residues prevent the binding of substrates that adopt an extended conformation. Interestingly, the corresponding residues in OscG$_{mac}$ (Ile183, Ala207, Cys209) are smaller in size and this may allow extra space for longer substrates to bind (Fig. S54 and S55, ESI†).

In order to test this hypothesis, we generated PatG$_{mac}$ mutants in which the native bulky residues were replaced with those that are smaller from OscG$_{mac}$. Mutants include the single M169I, F193A and R195C; the double M169I F193A, M169I R195C and F193A R195C; and the triple M169I F193A R195C. All mutants were expressed in *E. coli* and purified (Fig. S56, ESI†) and their activities were compared with that of PatG$_{mac}$. Strikingly, the mutant M169I processed substrate 1 (Table 1) which could not be processed with PatG$_{mac}$ or any of its other mutants (Fig. S57, ESI†). In addition, we observed that reactions of the mutants M169I F193A and M169I F193A R195C with substrate 15 (Table 1) resulted in cyclisation of most of the linear precursor peptide, compared to PatG$_{mac}$, which could cyclize only part of it over the same time frame (Fig. S58–S60, ESI†). These results showed that these mutations resulted in slightly altered activity. However, none of these mutants were able to process the 29 mer (substrate 14; Table 1) most probably because of the different recognition signal in this substrate.

The above results show that the cyanobactin macrocyclase OscG$_{mac}$ can process substrates containing Cys- and Ser-pseudoprolines instead of Pro and thiazoline at position P1.

OscG$_{mac}$ has a wider substrate tolerance, including D-amino acids, can process longer peptide substrates and has less stringent recognition determinants when compared to PatG$_{mac}$. We hypothesized that the residues Ile183, Ala207, and Cys209 are responsible for the altered tolerance to certain substrates and showed that mutations of these residues in PatG$_{mac}$ resulted in altered activity and in a specific instance a mutant processed a substrate that could not be processed by PatG$_{mac}$. OscG$_{mac}$, thus, is a very useful biotechnological tool to make modified cyclic peptides.

Conflicts of interest

There are no conflicts of interest to declare.

Notes and references

1 F. Giordanetto and J. Kihlberg, *J. Med. Chem.*, 2014, **57**, 278–295.
2 E. M. Driggers, S. P. Hale, J. Lee and N. K. Terrett, *Nat. Rev. Drug Discovery*, 2008, **7**, 608–624.
3 M. Gao, K. Cheng and H. Yin, *Biopolymers*, 2015, **104**, 310–316.
4 C. Heinis, *Nat. Chem. Biol.*, 2014, **10**, 696–698.
5 E. A. Villar, D. Beglov, S. Chennamadhavuni, J. A. Porco, D. Kozakov, S. Vajda and A. Whitty, *Nat. Chem. Biol.*, 2014, **10**, 723–731.
6 T. A. F. Cardote and A. Ciulli, *ChemMedChem*, 2016, **11**, 787–794.
7 C. J. White and A. K. Yudin, *Nat. Chem.*, 2011, **3**, 509–524.
8 T. Katoh, Y. Goto, S. Reza and H. Suga, *Chem. Commun.*, 2011, **47**, 9946–9958.
9 A. Tavassoli and S. J. Benkovic, *Nat. Protoc.*, 2007, **2**, 1126–1133.
10 J. E. Townend and A. Tavassoli, *ACS Chem. Biol.*, 2016, **11**, 1624–1630.
11 P. G. Arnison, M. J. Bibb, G. Bierbaum, A. A. Bowers, T. S. Bugni, G. Bulaj, J. A. Camarero, D. J. Campopiano, G. L. Challis, J. Clardy, P. D. Cotter, D. J. Craik, M. Dawson, E. Dittmann, S. Donadio, P. C. Dorrestein, K. Entian, M. A. Fischbach, J. S. Garavelli, U. Göransson, C. W. Gruber, D. H. Haft, T. K. Hemscheidt, C. Hertweck, C. Hill, A. R. Horswill, M. Jaspars, W. L. Kelly, J. P. Klinman, O. P. Kuipers, A. J. Link, W. Liu, M. A. Marahiel, D. A. Mitchell, G. N. Moll, B. S. Moore, R. Müller, S. K. Nair, I. F. Nes, G. E. Norris, B. M. Olivera, H. Onaka, M. L. Patchett, J. Piel, M. J. T. Reaney, S. Rebuffat, R. P. Ross, H. Sahl, E. W. Schmidt, M. E. Selsted, K. Severinov, B. Shen, K. Sivonen, L. Smith, T. Stein, R. D. Süssmuth, J. R. Tagg, G. Tang, A. W. Truman, J. C. Vederas, C. T. Walsh, J. D. Walton, S. C. Wenzel, J. M. Willey and W. A. van der Donk, *Nat. Prod. Rep.*, 2013, **30**, 108–160.
12 K. R. Lennard and A. Tavassoli, *Chem. – Eur. J.*, 2014, **20**, 10608–10614.
13 G. K. T. Nguyen, S. Wang, Y. Qui, X. Hemu, Y. Lian and J. P. Tam, *Nat. Chem. Biol.*, 2014, **10**, 732–738.
14 H. Luo, S. Hong, R. M. Sgambelluri, E. Angelos, X. Li and J. D. Walton, *Chem. Biol.*, 2014, **21**, 1610–1617.
15 C. J. S. Barber, P. T. Pujara, D. W. Reed, S. Chiwocha, H. Zhang and P. S. Covello, *J. Biol. Chem.*, 2013, **288**, 12500–12510.
16 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 and J. H. Naismith, *Nat. Struct. Mol. Biol.*, 2012, **19**, 767–772.
17 N. F. Brás, P. Ferreira, A. R. Calixto, M. Jaspars, W. Houssen, J. H. Naismith, P. A. Fernandes and M. J. Ramos, *Chem. – Eur. J.*, 2016, **22**, 13089–13097.
18 M. S. Donia, B. J. Hathaway, S. Sudek, M. G. Haygood, M. Rosovitz, J. Ravel and E. W. Schmidt, *Nat. Chem. Biol.*, 2006, **2**, 729–735.
19 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 and M. Jaspars, *Angew. Chem., Int. Ed.*, 2014, **53**, 14171–14174.
20 J. Booth, C. Alexandru-Crivac, K. Rickaby, A. Nneoyiegbe, U. Umeobika, A. McEwan, L. Trembleau, M. Jaspars, W. Houssen and D. Shalashilin, *J. Phys. Chem. Lett.*, 2017, **8**, 2310–2315.
21 K. Sivonen, N. Leikoski, D. P. Fewer and J. Jokela, *Appl. Microbiol. Biotechnol.*, 2010, **86**, 1213–1225.
22 M. S. Donia and E. W. Schmidt, *Chem. Biol.*, 2011, **18**, 508–519.
23 N. Leikoski, L. Liu, J. Jokela, M. Wahlsten, M. Gugger, A. Calteau, P. Permi, C. A. Kerfeld, K. Sivonen and D. P. Fewer, *Chem. Biol.*, 2013, **20**, 1033–1043.
24 N. Leikoski, D. P. Fewer, J. Jokela, M. Wahlsten, P. Alakoski and K. Sivonen, *PLoS One*, 2012, **7**, e43002.
25 M. Falorni, S. Conti, G. Giacomelli, S. Cossu and F. Soccolini, *Tetrahedron: Asymmetry*, 1995, **6**, 287–294.
26 A. B. Weinstein, D. P. Schuman, Z. X. Tan and S. S. Stahl, *Angew. Chem., Int. Ed.*, 2013, **52**, 11867–11870.
27 D. Lannotta, N. Castellucci, M. Monari and C. Tomasini, *Tetrahedron Lett.*, 2010, **51**, 4558–4559.
28 C. Sungman, *J. Biol. Chem.*, 1970, **245**, 4814–4818.
29 L. A. Kelley, S. Mezulis, C. M. Yates, M. N. Wass and M. J. E. Sternberg, *Nat. Protoc.*, 2015, **10**, 845–858.