Organic & Biomolecular Chemistry

PAPER

View Article Online
View Journal | View Issue

Cite this: *Org. Biomol. Chem.*, 2014, **12**, 8542

Insights into the programmed ketoreduction of partially reducing polyketide synthases: stereo- and substrate-specificity of the ketoreductase domain†

Ishin Soehano,‡[a] Lifeng Yang,‡[a] Feiqing Ding,‡[b] Huihua Sun,[a] Zhen Jie Low,[a] Xuewei Liu[b] and Zhao-Xun Liang*[a]

Received 20th August 2014,
Accepted 3rd September 2014

DOI: 10.1039/c4ob01777c

www.rsc.org/obc

One of the hallmarks of iterative polyketide synthases (PKSs) is the programming mechanism which is essential for the generation of structurally diverse polyketide products. In partially reducing iterative PKSs (PR-PKSs), the programming mechanism is mainly dictated by the ketoreductase (KR) domain. The KR domain contributes to the programming of PR-PKSs through selective reduction of polyketide intermediates. How the KR domain achieves the selective ketoreduction remains to be fully understood. In this study, we found that the KR domain of the (*R*)-mellein-synthesizing PR-PKS SACE5532 functions as a B-type KR domain to generate (*R*)-hydroxyl functionalities. Comparative studies of the KR domains of SACE5532 and NcsB suggested that the two KR domains have distinct substrate preferences towards simple *N*-acetylcysteamine thioester (SNAC) substrates. We further found that the substrate preference of $KR_{SACE5532}$ can be switched by swapping several motifs with KR_{NcsB}, and that swapping of the same motifs in the full length SACE5532 resulted in a reprogramming of the PKS. Together, the results advance our understanding of the programming of iterative PR-PKSs by providing new support to the hypothesis that the programmed ketoreduction is accomplished by differential recognition of polyketide intermediates.

Introduction

Iterative PKSs are multifunctional enzymes that assemble structurally diverse products using a single module of catalytic domains. The use of a single module of catalytic domains makes the iterative PKSs mechanistically distinct from the modular PKSs.[1–3] Iterative PKSs are categorized into non-reducing (NR), highly reducing (HR) and partially reducing (PR)-PKSs according to the composition of catalytic domains and oxidation levels of the polyketide intermediates.[2,4] The best known PR-PKSs are the bacterium or fungus-derived 6-MSAS, NcsB and AziB, which synthesize 6-methylsalicylic acid (6-MSA),[5–8] 2-hydroxyl-5-methyl-naphthoic acid (hmNPA)[9,10] and 5-methyl-naphthoic acid (mNPA)[11] respectively. In a recent effort to explore the diversity of PR-PKSs, we discovered a (*R*)-mellein-synthesizing PR-PKS (SACE5532) from the soil bacterium *Saccharopolyspora erythraea*.[12] Despite the fact that SACE5532 shares the same domain composition and a head-to-tail sequence homology with NcsB, AziB and 6-MSAS, SACE5532 synthesizes (*R*)-mellein, an aromatized pentaketide product that is likely to be generated by a distinct programming mechanism. SACE5532 was postulated to selectively reduce the β-keto group of the diketide and tetraketide intermediates to generate hydroxyl groups at C2 and C6 positions (Fig. 1). The ketoreduction pattern is different from those of 6-MSAS and NcsB, with 6-MSAS only reducing the triketide intermediate and NcsB reducing the triketide and pentaketide intermediates.[12]

The structural diversity of PR-PKS products arises from the programming of the C–C bond forming and processing catalytic domains, mainly through controlling carbon-chain length, ketoreduction and cyclization pattern. The carbon-chain length of the products is most likely to be determined by the substrate-binding pocket of the ketosynthase (KS) domain, as observed in other PKSs,[13–15] whereas the cyclization pattern is likely to be governed by the positions of the hydroxyl and keto functionalities without using a product template (PT) domain.[16] While the NADPH-dependent ketoreduction is catalysed by the KR domain, how the ketoreduction pattern is determined remains to be fully understood.[7,17–19] Despite the

[a]*School of Biological Sciences Nanyang Technological University, 60 Nanyang Drive, 637551, Singapore. E-mail: zxliang@ntu.edu.sg; Fax: (+)65 6791 3856*
[b]*School of Mathematics and Physics, Nanyang Technological University, 21 Nanyang Link, 637371, Singapore*
†Electronic supplementary information (ESI) available. See DOI: 10.1039/c4ob01777c
‡These authors contributed equally to this work.

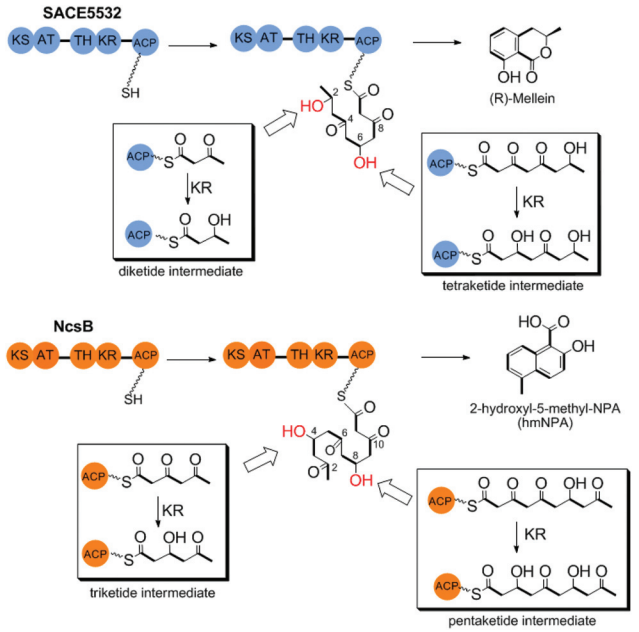

Fig. 1 Postulated selective ketoreduction in SACE5532 and NcsB. SACE5532 reduces the keto group of the diketide and tetraketide intermediates, whereas NcsB reduces the keto group of the triketide and pentaketide intermediates. In addition to the essential ketosynthase (KS), acyltransferase (AT) and acyl carrier protein (ACP) domains, the PR-PKSs also contain a dehydratase-like thiohydrolase (TH) domain[20] and a ketoreductase (KR) domain.

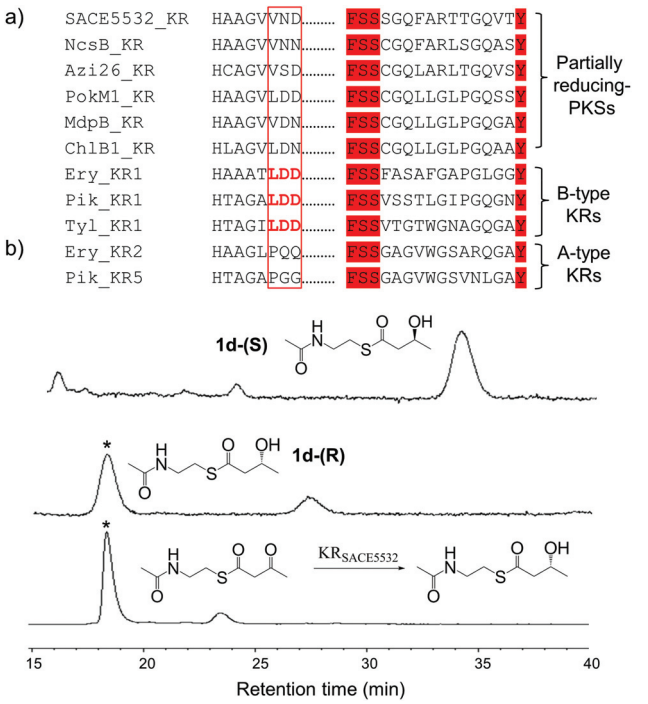

Fig. 2 Stereo-specific ketoreduction of the $KR_{SACE5532}$ domain. (a) Partial sequence comparison of the KR domains of PR-PKSs and representative type-A and B KR domains from modular PKSs. The motifs containing the signature "LDD" motifs for type-B KR domains are shown in the frame (Azi26_KR of AziB_KR: KR domain of AziB; PokM1_KR: KR domain of PokM1; MdpB_KR: KR domain of MdpB; ChlB1_KR: KR domain of ChlB1; Ery_KR1: KR domain of the first module of erythromycin PKS; Pik_KR1: KR domain of the first module of pikromycin PKS; Tyl_KR1: KR domain of the first module of tylosin PKS; Ery_KR2: KR domain of the second module of erythromycin PKS; Pik_KR5: KR domain of the fifth module of pikromycin PKS). (b) Chiral HPLC analysis of the product from the reduction of acetoacetyl-SNAC by $KR_{SACE5532}$.

fact that PR-PKSs are some of the first PKSs to be discovered and characterized,[7,17–19] the function of individual catalytic domains of PR-PKSs is difficult to dissect and much remains to be understood. Based on our recent observation that the ketoreductase (KR) domain of SACE5532, but not the KR domain of NcsB, is able to reduce diketide substrates,[12] we proposed that the ketoreduction patterns of SACE5532 and NcsB are solely governed by the KR domains through differential substrate recognition. In this study, the stereo- and substrate-specificity of the $KR_{SACE5532}$ domain were further probed to yield fresh evidence supporting the hypothesis that differential substrate recognition is central for achieving the programmed ketoreduction.

Results and discussion

A large body of knowledge about the stereo-specificity of the KR domains of fatty acid synthases, type II PKS and type I modular PKSs has been accumulated.[21–38] Enzymatically active KR domains are classified into A-type and B-type according to the stereochemistry of the β-hydroxyl group generated by ketoreduction.[26,31–35,39–41] Signature motifs of the A- or B-type KR domains have been identified. However, because the signature motifs found in the KR domains of modular PKSs are usually absent in the KR domains of most PR-PKSs (Fig. 2a), it is not immediately evident from sequence analysis whether the KR domains of PR-PKSs produce (*R*)- or (*S*)-configured hydroxyl groups. To establish the stereo-preference of the $KR_{SACE5532}$ domain, we cloned the stand-alone $KR_{SACE5532}$ domain (1170–1638 aa of SACE5532) according to newly assigned domain boundaries[12] and synthesized the two *N*-acetylcysteamine thioester (SNAC) enantiomers **1d**-(*R*) and **1d**-(*S*). As revealed by chiral HPLC analysis, incubation of the diketide substrate acetoacetyl-SNAC (**1c**) with $KR_{SACE552}$ and NADPH produced **1d**-(*R*) (Fig. 2b), suggesting that $KR_{SACE5532}$ is a B-type KR domain. Generation of the (*R*)-configured diketide product is fully consistent with the (*R*)-configured chiral centre in (*R*)-mellein. Moreover, a survey of the sequences of more than 100 putative PR-PKSs revealed that although the signature motif "LDD" of B-type KR domains is not strictly conserved in most PR-PKSs, a similar motif (V/L/Y)(N/S/D/E)(D/N/S) can still be found at the same position (Fig. 2a). Considering that the KR domains of PR-PKSs also lack a highly conserved tryptophan residue found in A-type KR domains[26,32,39] we propose that 6-MSAS-like PR-PKSs contain B-type KR domains that generate (*R*)-configured hydroxyl groups.

Having established the stereochemistry of the KR domains, we next tested whether $KR_{SACE5532}$ selectively reduces the

diketide (**1a**) and tetraketide (**3a**) intermediates and KR_{NcsB} prefers the triketide (**2b**) and pentaketide (**4b**) intermediates (Fig. 3a). To probe the substrate preference of the KR domains, we started by synthesizing a series of tri-, tetra- and pentaketide-SNAC substrates. Our effort to prepare the structurally faithful acyl-SNAC mimics of **2b**, **3a** and**4b** was not successful because the products were highly prone to spontaneous cyclization. To circumvent the problem, we synthesized the simple tri-, tetra- and pentaketide-SNAC substrates **2c**, **3c** and **4c** instead (Fig. 3b). The SNAC-substrates **2c**, **3c** and **4c** share the same carbon chain length with the postulated PKS intermediates **2b**, **3a** and **4b**, but lack the hydroxyl or keto groups beyond the β-position. In addition to $KR_{SACE5532}$, KR_{NcsB} (1162–1641 aa of NcsB) was cloned and shown to be enzymatically active using the non-specific substrate *trans*-1-decalone.[12,22] By *in vitro* enzymatic assays, we found that $KR_{SACE5532}$ and KR_{NcsB} exhibited clear preferences towards three of the four SNAC substrates. $KR_{SACE5532}$ reduced the diketide substrate **1c** with a k_{cat}/K_M of 2.1 M^{-1} s^{-1}, in contrast to the negligible activity for KR_{NcsB}. In contrast, KR_{NcsB} exhibited higher activity than $KR_{SACE5532}$ towards the triketide **2c** and the pentaketide **4c** (Fig. 3b). Note that the facile reduction of **4c** by KR_{NcsB} (2.8×10^2 M^{-1} s^{-1}) favours a mechanism whereby cyclization of the polyketide intermediate occurs only after the ketoreduction. The only substrate that did not show the anticipated preference is the tetraketide **3c**, with KR_{NcsB} showing moderately higher activity. Enzymatic assays of the KR activity with full-length SACE5532 and NcsB showed similar results, with the two KR domains exhibiting strong preference for **1c**, **2c** and **4c** (data not shown). Despite the drawback associated with the use of the simple SNAC substrates, the ability to differentiate the SNAC substrates indicates that the carbon-chain length of the polyketide intermediate is likely an important structural factor in substrate recognition. For **3c**, we speculate that the lack of differentiation is due to the absence of γ- and δ-hydroxyl/keto groups or the tethering phosphopantetheinyl. Although it can be argued that the ACP domain may play a role in substrate recognition and that the observed substrate preferences may not reflect the reduction of ACP-tethered substrates, the ACP domain does not affect substrate specificity because the $ACP_{SACE5532}$ and ACP_{NcsB} domains are interchangeable, which we will discuss later on. Hence, despite the limitation of the simple SNAC substrates used in the study, the observations suggest that the KR domains have distinct substrate preferences.

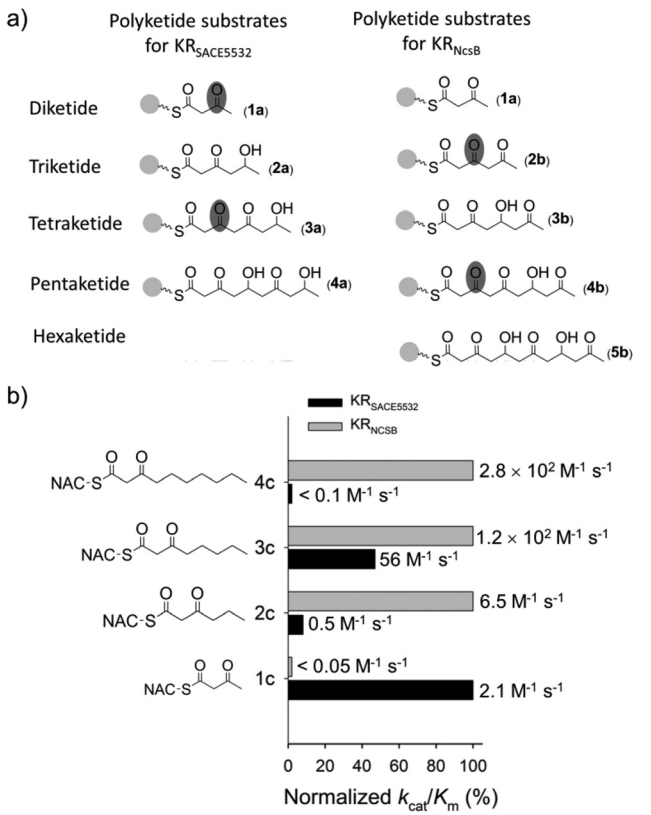

Fig. 3 Differential recognition of polyketide substrates by $KR_{SACE5532}$ and KR_{NcsB} domains. (a) Putative polyketide intermediates for SACE5532 and NcsB. Note that only the intermediates that bear a β-keto group are shown. The shaded keto groups are hypothesized to be reduced by the KR domain of SACE5532 or NcsB. (b) Comparison of the efficiency of the two KR domains in reducing the four acyl-SNAC substrates. The assay conditions are described in ESI.†

The ketoreduction catalysed by KR domains is believed to be assisted by a triad of Tyr, Ser and His residues.[40,42] How the ACP-tethered substrates are bound and recognized by the KR domains is less clear. To identify the residue and regions important for substrate recognition, we constructed a homology model for $KR_{SACE5532}$ using a type-B KR structure (PDB ID: 2FR0[40]) as the template. The structural model features a cleft for accommodating NADPH and likely acyl substrates (Fig. 4a). Judging from their locations, we identified all potential residues that are likely to be involved in direct binding of the acyl substrates. The identified 51 residues are contained in four fragments or motifs. Motif I consists of a short loop preceding the essential catalytic residue Y1536. The region occupied by motif I has been shown to be important for controlling the stereochemistry of the KR domain of the PKS in hypothemycin biosynthesis.[43] Motif II consists of a potentially flexible C-terminal motif. Motifs I and II form the walls of a narrow cleft in which the nicotinamide moiety of NADPH and the β-keto group of the acyl substrate are likely to be bound. Motif III encompasses a short helix and a loop situated on the top of motif I, whereas motif IV is located next to motif II and is probably involved in the binding of ACP and the phosphopantetheinyl moiety. We asked the question of whether the substrate preference of $KR_{SAC3352}$ can be altered by replacing some of the residues from the four motifs with the corresponding residues from KR_{NcsB}. The corresponding residues in KR_{NcsB} were identified by sequence alignment and secondary structure analysis (Fig. S1†). Substrates **1c** and **4c** were used as probes to report the switching of substrate preference. A switch of substrate preference would be indicated by a loss of activity towards **1c** and a concomitant gain of activity towards **4c**. The $KR_{SACE5532}$ variants harbouring different number of residue replacements were constructed using site-directed

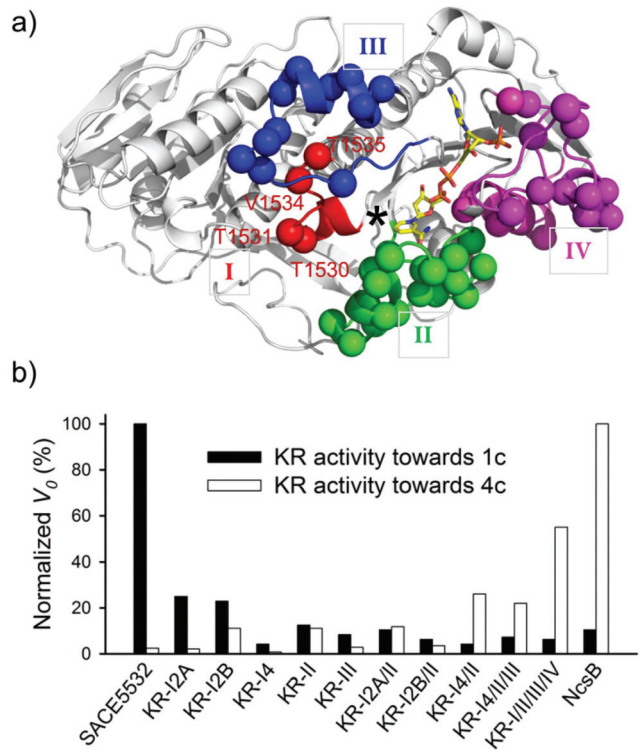

Fig. 4 Identification of the regions in $KR_{SACE5532}$ that govern substrate specificity. (a) Coloured are the four motifs that are likely to govern substrate specificity. The α-carbon of the residues from the motifs I–IV that are different between $KR_{SACE5532}$ and KR_{NcsB} are represented by the balls. NADPH is shown as sticks. The active site is indicated by the asterisk. (b) Enzymatic activities of the chimeric KR towards substrates **1c** and **4c** in relation to $KR_{SACE5532}$ and KR_{NcsB}. The chimeras were derived from $KR_{SACE5532}$. The residues that are replaced in the mutants are listed in Table S5.† The assay conditions are described in ESI.†

mutagenesis and overlapping PCR methods as described in the Experimental section.[46]

We first focused on the residues from motif I because of their proximity to the active site and reported roles in substrate recognition.[43] Only four residues are different for motif I between $KR_{SACE5532}$ and KR_{NcsB}. The two double mutants KR_{I2A} (T1530L, T1531S) and KR_{I2B} (V1534A, T1535S) exhibited significantly lower enzymatic activity than $KR_{SACE5532}$, whereas the replacement of the four residues resulted in 96% loss of activity towards **1c** (Fig. 4b). Although the replacement of individual residues in motifs II and III did not have significant effects, their enzymatic activity was substantially decreased when the entire motif II or III was swapped. It is not surprising that low enzymatic activity towards **1c** was also observed for the variants that contain two, three or all four swapped motifs. Differential scanning fluorometry (DSF) measurements showed that the chimeras have similar unfolding temperatures, indicating that the loss of enzymatic activity was not due to a change of protein stability (data not shown). Importantly, when the enzymatic activity of the variants towards **4c** was measured, we noticed that the replacement of the residues V1534 and T1535 from motif I led to a gain of enzymatic activity towards **4c**, with KR_{I2B} showing a notable 16% activity relative to KR_{NcsB}. Comparable activity (~17%) was observed when the entire motif II was swapped (KR_{II}). The chimera ($KR_{I4/II}$) with motifs I and II swapped exhibited further increased enzymatic activity (24%) for **4c**, whereas the replacement of the residues in motif III alone or in combination with other changes did not seem to improve the activity. The last chimera ($KR_{I/II/III/IV}$) with all four motifs swapped (a total of 51 replacements) is a reasonably efficient ketoreductase domain that exhibited 56% activity towards **4c** (Fig. 4b). A second-order rate constant of 1.6×10^2 M^{-1} S^{-1} was obtained for $KR_{I/II/III/IV}$, compared to the rate constant of 2.8×10^2 M^{-1} S^{-1} for KR_{NcsB}. In conjunction with the observations that none of the single, double or triple mutants exhibited significant activity towards **4c**, the results indicate that the substrate specificity is collectively determined by a large number of residues in the vicinity of the active site. Some residues, such as the ones from motif IV, may play indirect yet important roles by positioning substrate and cofactor-binding residues. These observations seem to suggest that evolution of new substrate specificity for the KR domains by protein engineering may require the replacement of more than a few residues. On an encouraging note, the remarkable tolerance of $KR_{SACE5532}$ to residue replacement indicates that the substrate-binding pocket is highly plastic and evolvable. Such evolvability is likely attributed to the facts that the four motifs are composed of mainly surface-exposed segments and that the overall stability of the protein scaffold is mainly governed by other parts of the protein. Such decoupling of stability and function has been observed in other enzymes with high adaptability and is consistent with the perpetual need of nature to evolve KR domains with new substrate specificity.[47,48]

Having shown that substrate specificity can be altered by motif-swapping in the $KR_{SACE5532}$ domain, we finally tested how the alteration of substrate specificity affects the ketoreduction pattern and product formation in the full-length PKS. We prepared the full PKS construct ($SACE5532(KR_{I/II/III/IV})$) with the KR domain replaced by the $KR_{I/II/III/IV}$ domain. Ketoreduction assays showed that the integrated $KR_{I/II/III/IV}$ domain is enzymatically active with anticipated preference for substrates **2c** and **4c** (data not shown). Instead of producing (*R*)-mellein or hmNPA, $SACE5532(KR_{I/II/III/IV})$ produced triacetic acid lactone (TAL) and 6-MSA as the major products (Fig. 5a). In contrast to SACE5532, the production of TAL and 6-MSA indicates that the diketide intermediate is no longer reduced by the $KR_{I/II/III/IV}$ domain. The production of 6-MSA suggests that a portion of the triketide intermediate is now reduced, confirming that the ketoreduction pattern has indeed been altered (Fig. 5b). The results also suggest that further optimization is needed for the $KR_{I/II/III/IV}$ domain to suppress the production of TAL that is derived from the unreduced triketide. However, we found that further improvement by swapping the residues situated beyond the four motifs is challenging because the swapping often led to a loss of protein stability. It will be crucial in the future to have the high-resolution crystal structure of $KR_{SACE5532}$ or KR_{NcsB} to guide the optimization.

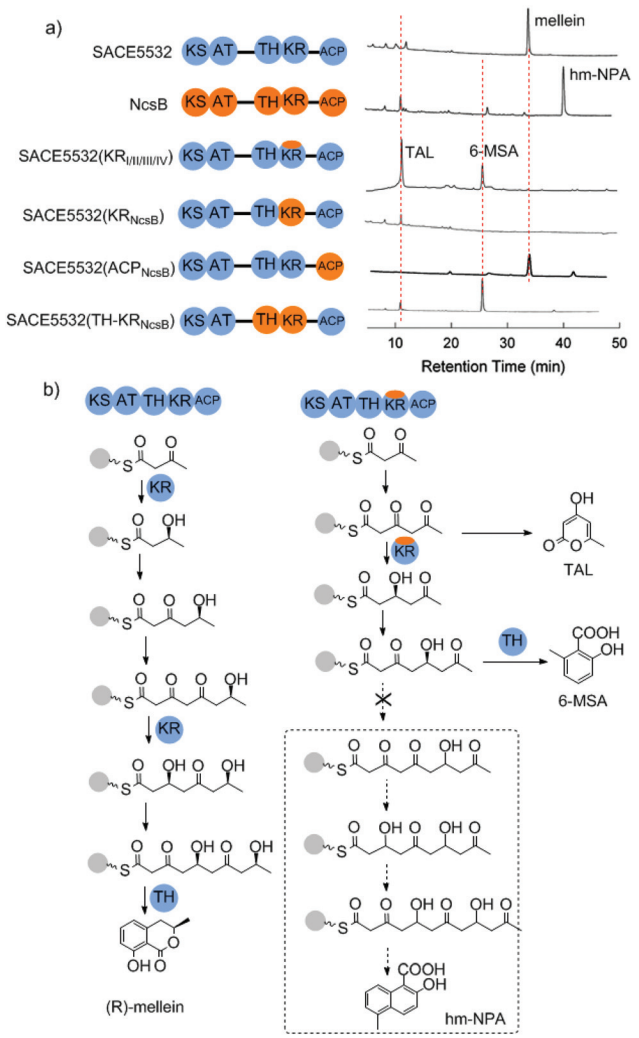

Fig. 5 Domain swapping between SACE5532 and NcsB resulted in the alteration of the ketoreduction pattern. (a) Product analysis by HPLC. Reaction conditions are described in the Experimental section and described elsewhere.[44,45] (b) Postulated mechanism for the formation of observed products.

Meanwhile, the absence of the hexaketide product hmNPA for SACE5532(KR$_{I/II/III/IV}$) indicates that further extension of the carbon chain is hindered, most likely because the tetraketide intermediate is not recognized by the KS$_{SACE5532}$ domain.[15] The release of 6-MSA by the TH$_{SACE5532}$ domain also seems to suggest that the TH domain is able to tolerate some degree of structural change in the product. It should be noted that we did not observe the production of 6-MSA by SACE5532(KR$_{NcsB}$), a chimera generated by replacing the KR domain with KR$_{NcsB}$. One plausible reason for the complete lack of 6-MSA is the alteration of the quaternary PKS structure and dynamics caused by the disruption of TH-KR domain interface. We also found that the replacement of the ACP domain by ACP$_{NcsB}$ did not affect the ketoreduction pattern and production of mellein, indicating that the ACP domain is not involved in differential substrate recognition. The result is consistent with recent studies which showed that the ACP domains of iterative PKSs are interchangeable. Given the modest decrease in mellein production, the ACP domain is likely to play a simple role enhancing the efficiency of intermediate shuttling through electrostatic-guided docking.[49–54] And lastly, we found that the SACE5532(TH–KR$_{NcsB}$) construct generated by swapping the TH-KR didomain produced 6-MSA and a negligible amount of TAL. The didomain-swapping result not only confirms that the KR domain controls the ketoreduction pattern but also indicates that the TH-KR domain interface is important for the function of PR-PKSs.

In summary, we have established that the KR domain of the mellein-synthesizing SACE5532, and likely other PR-PKSs, functions as a B-type KR domain to generate (*R*)-hydroxyl functionalities. We showed that the KR domains of SACE5532 and NcsB have distinct substrate preferences towards simple SNAC substrates. We found that a KR$_{SACE5532}$ variant that contains motifs grafted from KR$_{NcsB}$ exhibited altered substrate specificity and can cause a reprogramming of the full length PKS. The motif-swapping results also demonstrate the plasticity of the KR active site and the potential of reprogramming the PKSs by motif-grafting. In contrast to the domain-swapping method[55–59] that often disrupts domain–domain interface and causes a loss of function, the motif-swapping or grafting method may hold some advantages for PKS engineering. The observations lend further support to the hypothesis that the KR domain of PR-PKSs achieves programmed ketoreduction by differential recognition of polyketide intermediates. Further work, such as the determination of the crystal structures of enzyme–substrate complexes, will be required to fully establish the molecular basis of the substrate preferences exhibited by the KR domains.

Experimental section

Molecular cloning and DNA manipulation

The cloning of KR$_{SACE5532}$ and KR$_{NcsB}$ domains was described previously.[12] The single, double and quadruple mutants were generated by PCR using KAPAHiFi™ DNA polymerase with the primers listed in Table S2.† The templates used for the double and quadruple mutants were pET-KR$_{SACE5532}$ and the synthesized double mutant pET-KR$_{SACE5532}$ KR$_{I2A}$ respectively. Similarly, mutations were introduced into other expression plasmids such as pET-KR$_{II}$ to yield KR$_{I2A/II}$. The PCR conditions used for site-directed mutagenesis are: denaturing at 95 °C for 5 minutes, 30 cycles at 98 °C for 50 seconds, gradient annealing temperatures (60–85 °C) for 50 seconds, 72 °C for 5 minutes, and 10 minutes at 72 °C. To construct the KR$_{SACE5532}$ variants that contain the motifs from KR$_{NcsB}$ (*e.g.* KR$_{III}$ and KR$_{I4/II/III}$), an overlapping PCR reaction was performed. The three-step PCR reaction eventually generated DNA fragments which were subsequently used for ligation into the expression vector pET28(b)+. For the PCR reactions, specific primers with 15 bases of overlapping sequences were used (Table S3†). Briefly, the first round of PCR used the forward primer A containing the NdeI restriction site and the reverse

primer B containing the replacement motif sequences. The second round of PCR generated the DNA fragment consisting of the replacement motif with the remaining C-terminal sequences. The forward primer that contains the replacement motif sequences and the reverse primer that contains an XhoI restriction site were utilized in the second round of PCR. The PCR conditions are: denaturing at 95 °C for 5 minutes, 30 cycles at 98 °C for 50 seconds, gradient annealing temperatures (50–75 °C) for 50 seconds, and 72 °C for 1 minute, and 7 minutes at 72 °C. The products of the first and second PCR were gel-purified separately. The third round of PCR employed 70 ng each of PCR products as templates with the forward primer A and the reverse primer D (see Table S3†). The 1.5 kb PCR product was gel-purified and digested with NdeI and XhoI restriction enzymes. The digested DNA fragment was gel-purified and ligated into pET28b(+). The resulted plasmid was then transformed into the Top10 competent cells and later into Rosetta (DE3) cells for protein expression. The genes that encode the $KR_{SACE5532}$ variants KR_{II} and $KR_{I/II/III/IV}$ were obtained from GenScript Corporation (NJ, USA) with the DNA sequence optimized for protein expression in *E. coli*. The genes were provided as pUC57-based plasmids with NdeI and XhoI restriction sites added to the N- and C-termini and subcloned into pET28b(+).

To construct the full length SACE5532($KR_{I/II/III/IV}$), Gibson Assembly™ Cloning kit (New England BioLabs) was used. Using this technique, various DNA fragments and the vector backbone were assembled using overlapping primers that include 15–25 nucleotide overlapping sequences and 15–25 nucleotide gene-specific sequences (Table S4†). The PCR-amplified DNA fragments and the linearized vector pET28(b)+ were incubated with the Gibson assembly master mix at 50 °C for 45 minutes. This Gibson-assembled mixture was subsequently transformed into NEB 5α competent cells according to the manufacturer's instructions.

To construct the ACP domain swapping variant SACE5532 (ACP_{NcsB}), the plasmids pUC57-NcsB and PUC57-SACE5532 were digested with ClaI and XhoI. The 498 bp ACP_{NcsB} fragment was ligated into the linearized PUC57-SACE5532 plasmid to yield PUC57-SACE5532(ACP_{NcsB}). The 5.2 kb PKS gene was digested and ligated into the expression vector pET28b(+) to generate pET-SACE5532(ACP_{NcsB}). Similarly, to replace the TH-KR di-domain of SACE5532 (900–1638 aa) with that of NcsB (886–1641 aa), the plasmid pET-SACE5532(TH-KR_{NcsB}) was generated. To replace the KR domain of SACE5532 (1170–1638 aa) with that of NcsB (1162–1641 aa), the $KR_{SACE5532}$ DNA region was replaced by the corresponding DNA sequence from NcsB gene to yield the plasmid pET-SACE5532(KR_{NcsB}).

Recombinant protein expression and purification

For the stand-alone KR domains, the plasmids pET-$KR_{SACE5532}$ and pET-KR_{NcsB} were transformed into *E. coli* Rosetta (DE3) competent cells. The cells were plated on LB medium supplemented with 50 μg ml^{-1} kanamycin. A single colony was used to inoculate the LB medium supplemented with 50 μg ml^{-1} kanamycin. The culture was kept in a shaking incubator overnight at 37 °C at 180 rpm. An 8 ml aliquot was transferred to 800 ml of LB medium supplemented with 50 μg ml^{-1} kanamycin. The culture was kept at 37 °C at 200 rpm till the OD_{600} reached ~0.6 (~3 hours) before it was cooled down to 16 °C and induced with 0.2 mM IPTG. After incubation at 16 °C for an additional ~20 hours at 160 rpm, the cells were harvested. The cell pellet was re-suspended in the lysis buffer (50 mM NaH_2PO_4 (pH 8.0), 300 mM NaCl, 20 mM imidazole, and 10% (v/v) glycerol) and lysed by sonication for 10 minutes. After centrifugation at 20 000 rpm for 30 minutes at 4 °C, the supernatant was filtered using a 0.45 μm membrane and loaded onto a Ni-NTA column. The column was washed using the lysis buffer and wash buffer supplemented with 10, 50 and 200 mM of imidazole before the protein was eluted out with the elution buffer containing 500 mM imidazole. Proteins were desalted and exchanged into the Tris buffer (50 mM Tris-HCl (pH 8.0), 150 mM NaCl, and 10% (v/v) glycerol) using the PD10 desalting column. The protein was concentrated, flash frozen in liquid nitrogen, and stored in −80 °C freezer.

The expression and purification of SACE5532 were described previously[12] and the expression and purification of NcsB and chimeric PKS proteins were carried out using a similar procedure. By co-transformation, the PKSs were co-expressed with the 4′-phosphopantetheinyl transferase Sfp to facilitate the phosphopantetheinylation of the ACP domain.

Synthesis of SNAC substrates

The synthesis of acetoacetyl-SNAC (**1c**) was described previously.[12] The synthetic method for **1d**(*S*) and **1d**(*R*) was adopted from a previously reported procedure and the structures of the products were established by MS and NMR spectroscopy.[60] For the synthesis of **2c**, **3c** and **4c**, a 0.5 M solution of Meldrum's acid (1 eq.) in CH_2Cl_2 was prepared in a round bottomed flask at room temperature. Pyridine (2 eq.) and DMAP (20 mol%) were added and the reaction was stirred for 10 minutes at room temperature. The reaction was cooled to 0 °C, 1 eq. of the desired acid chloride was added slowly, and the reaction was warmed to room temperature and stirred overnight. The reaction was washed with 0.1 M HCl and then with water. The organic layers were dried over Na_2SO_4, filtered, and the solvent was removed. After workup of the reaction, the product was purified by flash column chromatography (7 : 1 hexanes–ethyl acetate) to give the synthetic intermediates **2ca**, **3ca** and **4ca** (ESI Scheme 1†). This material was then refluxed in toluene (0.5 M) with 1 eq. of *N*-acetylcysteamine for three hours. The reaction was cooled and the solvent was removed *in vacuo*. The product was purified by flash column chromatography using 1 : 1 CH_2Cl_2 : ethyl acetate to give **2c**, **3c** and **4c**.

3-Oxohexanoyl-SNAC (2c). ^{1}H NMR ($CDCl_3$, 400 MHz): δ 6.07 (br, 1H), 3.67 (s, 2H), 3.41–3.46 (m, 2H), 3.05–3.09 (m, 2H), 2.50 (t, J = 7.2 Hz, 2H), 1.95 (s, 3H), 1.56–1.65 (m, 2H), 0.89–0.95 (m, 3H); ^{13}C NMR ($CDCl_3$, 100 MHz): δ 202.2, 192.4, 170.5, 99.3, 57.2, 39.2, 29.2, 23.1, 16.9, 13.5; HRMS (ESI) m/z $[M + Na]^+$ calcd for $C_{10}H_{17}NO_3SNa$ 254.0827, found 254.0831.

3-Oxooctanoyl-SNAC (3c). ^{1}H NMR ($CDCl_3$, 400 MHz): δ 6.07 (br, 1H), 3.68 (s, 2H), 3.42–3.46 (m, 2H), 3.05–3.09 (m, 2H), 2.51 (t, J = 7.2 Hz, 2H), 1.96 (s, 3H), 1.54–1.61 (m, 2H), 1.25–1.31 (m, 4H), 0.87–0.89 (m, 3H); ^{13}C NMR ($CDCl_3$, 100 MHz): δ 202.4, 192.4, 170.5, 99.1, 57.2, 39.2, 31.1, 29.2, 23.2, 23.1, 22.4, 13.9; HRMS (ESI) m/z $[M + Na]^+$ calcd for $C_{12}H_{21}NO_3SNa$ 282.1140, found 282.1143.

3-Oxodecanoyl-SNAC (4c). ^{1}H NMR ($CDCl_3$, 400 MHz): δ 6.07 (br, 1H), 3.65 (s, 2H), 3.41–3.46 (m, 2H), 3.05–3.09 (m, 2H), 2.51 (t, J = 7.2 Hz, 2H), 1.96 (s, 3H), 1.56–1.65 (m, 2H), 1.20–1.30 (m, 8H), 0.89–0.95 (m, 3H); ^{13}C NMR ($CDCl_3$, 100 MHz): δ 202.4, 192.4, 170.5, 99.1, 57.2, 39.2, 31.6, 31.5, 29.2, 29.0, 28.9, 23.4, 22.6, 14.0; HRMS (ESI) m/z $[M + Na]^+$ calcd for $C_{14}H_{25}NO_3SNa$ 310.1453, found 310.1451.

Enzymatic assays

The assays of the ketoreductase activity were conducted using a Shimadzu UV-Vis 1650 spectrophotometer following the procedures described previously.[44,45] The assays were performed in 100 mM HEPES buffer (pH 8.5) in a total volume of 200 μl containing the following components: 1 mg ml^{-1} ketoreductase, 1 mM **1c** and **2c** or 0.1 mM **3c** and **4c** substrate. The reaction mixtures were incubated at 25 °C and the reaction was initiated by adding NADPH solution (0.25 mM). The progress of the reaction was followed by monitoring at 341 nm wavelength continuously. Initial velocity (V_0) values were obtained from the slopes of the progress curve and the extinction coefficient of NADPH. The values of V_0 at five different substrate concentrations were used for the measurement of the second-order rate constant k_{cat}/K_M.

For the determination of the stereo-selectivity of the $KR_{SACE5532}$ domain, a total of 800 μl reaction mixture of 2 mM substrate **1c**, 0.3 mM NADPH and 3.5 mg $KR_{SACE5532}$ in 100 mM HEPES buffer (pH 8.5) was incubated at 26 °C for 80 minutes. The reaction mixture was then supplemented with 30 μl of NADPH (10 mM) and incubated at 26 °C for 80 minutes. The reaction was quenched with 80 μl of 1.2 mM HCl and the solution was vortexed to remove the enzyme. The reaction mixture was extracted twice with equal volumes of ethyl acetate. The mixture was spun at 14 800 rpm for 15 min and the organic layer was separated and subjected to chiral HPLC analysis. The two (*R*) and (*S*)-configured standards and the reaction extract were dissolved in 100 μl isopropanol. The samples were analysed using normal phase HPLC equipped with a Chiralpak IA® column (150 × 4.60 mm) (Chiral Technologies). The HPLC conditions were: linear gradient; hexane to isopropanol ratio (v/v) of 90 to 10; 30 °C; 1 ml min^{-1} flow rate (254 nm).

The enzymatic assays for the wild type PKSs were performed following the procedure described previously.[44,61] To test the enzymatic activity of the chimeric PKSs, a reaction mixture containing 3 μl of $MgCl_2$ (1 M), 1 μl of acetyl-CoA (100 mM), 100 μl of co-purified PKS/Sfp (15 mg ml^{-1}), 3 μl of malonyl-CoA (100 mM), 5 μl of NADPH (10 mM) and 88 μl of reaction buffer (50 mM Tris (pH 8.0), 150 mM NaCl and 10% (v/v) glycerol) was incubated at room temperature. A UV-Vis spectrophotometer was used to follow the depletion of NADPH at 340 nm wavelength typically for two hours. At the end of the reaction, the reaction was stopped by quenching with 20 μl of 1.2 M HCl and the solution was vortexed to precipitate the enzyme. The reaction mixture was extracted twice with an equal volume of ethyl acetate. The mixture was spun at 14 800 rpm for 15 min and the organic layer was separated and subjected to HPLC analysis. The conditions for HPLC analysis are similar to the ones described previously.[44,45,61]

Construction of a structural model for $KR_{SACE5532}$

The structural model of $KR_{SACE5532}$ was constructed using the first KR domain of erythromycin synthase DEBS (PDB ID: 2FR0) as the template[40,62] and the software Rosetta Suite.[63,64] Energy minimization was subsequently performed using the Rosetta protein design.

Acknowledgements

This work was supported by a Tier II grant from the Ministry of Education (MOE) of Singapore to Z.X. Liang.

Notes and references

1 A. M. Hill, *Nat. Prod. Rep.*, 2006, **23**, 256–320.
2 B. Shen, *Curr. Opin. Chem. Biol.*, 2003, **7**, 285–295.
3 C. Khosla, D. Herschlag, D. E. Cane and C. T. Walsh, *Biochemistry*, 2014, **18**, 2875–2883.
4 R. J. Cox, *Org. Biomol. Chem.*, 2007, **5**, 2010–2026.
5 I. Schmitt, S. Kautz and H. T. Lumbsch, *Mycol. Res.*, 2008, **112**, 289–296.
6 L. Shao, X.-D. Qu, X.-Y. Jia, Q.-F. Zhao, Z.-H. Tian, M. Wang, G.-L. Tang and W. Liu, *Biochem. Biophys. Res. Commun.*, 2006, **345**, 133–139.
7 J. Beck, S. Ripka, A. Siegner, E. Schiltz and E. Schweizer, *Eur. J. Biochem.*, 1990, **192**, 487–498.
8 P. Lu, A. Zhang, L. M. Dennis, A. M. Dahl-Roshak, Y. Q. Xia, B. Arison, Z. An and J. S. Tkacz, *Mol. Gen. Genomics*, 2005, **273**, 207–216.
9 W. Liu, K. Nonaka, L. P. Nie, J. Zhang, S. D. Christenson, J. Bae, S. G. Van Lanen, E. Zazopoulos, C. M. Farnet, C. F. Yang and B. Shen, *Chem. Biol.*, 2005, **12**, 293–302.
10 B. Sthapit, T. J. Oh, R. Lamichhane, K. Liou, H. C. Lee, C. G. Kim and J. K. Sohng, *FEBS Lett.*, 2004, **566**, 201–206.
11 Q. F. Zhao, Q. L. He, W. Ding, M. C. Tang, Q. J. Kang, Y. Yu, W. Deng, Q. Zhang, J. Fang, G. L. Tang and W. Liu, *Chem. Biol.*, 2008, **15**, 693–705.
12 H. Sun, C. L. Ho, F. Ding, I. Soehano, X.-W. Liu and Z.-X. Liang, *J. Am. Chem. Soc.*, 2012, **134**, 11924–11927.
13 I. Abe, S. Oguro, Y. Utsumi, Y. Sano and H. Noguchi, *J. Am. Chem. Soc.*, 2005, **127**, 12709–12716.
14 Y. Tang, S.-C. Tsai and C. Khosla, *J. Am. Chem. Soc.*, 2003, **125**, 12708–12709.
15 K. Watanabe, C. C. C. Wang, C. N. Boddy, D. E. Cane and C. Khosla, *J. Biol. Chem.*, 2003, **278**, 42020–42026.

16 J. M. Crawford, T. P. Korman, J. W. Labonte, A. L. Vagstad, E. A. Hill, O. Kamari-Bidkorpeh, S.-C. Tsai and C. A. Townsend, *Nature*, 2009, **461**, 1139–1143.
17 J. B. Spencer and P. M. Jordan, *Biochem. J.*, 1992, **288**, 839–846.
18 C. J. Child, J. B. Spencer, P. Bhogal and P. M. ShoolinginJordan, *Biochemistry*, 1996, **35**, 12267–12274.
19 Q. Zhang, B. Pang, W. Ding and W. Liu, *ACS Catal.*, 2013, **3**, 1439–1447.
20 T. Moriguchi, Y. Kezuka, T. Nonaka, Y. Ebizuka and I. Fujii, *J. Biol. Chem.*, 2010, **285**, 15637–15643.
21 H. Fu, S. Ebertkhosla, D. A. Hopwood and C. Khosla, *J. Am. Chem. Soc.*, 1994, **116**, 4166–4170.
22 L. H. Østergaard, L. Kellenberger, J. Cortés, M. P. Roddis, M. Deacon, J. Staunton and P. F. Leadlay, *Biochemistry*, 2002, **41**, 2719–2726.
23 A. T. Hadfield, C. Limpkin, W. Teartasin, T. J. Simpson, J. Crosby and M. P. Crump, *Structure*, 2004, **12**, 1865–1875.
24 S. Bali, H. M. O'Hare and K. J. Weissman, *ChemBioChem*, 2006, **7**, 478–484.
25 R. Castonguay, W. G. He, A. Y. Chen, C. Khosla and D. E. Cane, *J. Am. Chem. Soc.*, 2007, **129**, 13758–13769.
26 A. T. Keatinge-Clay, *Chem. Biol.*, 2007, **14**, 898–908.
27 T. Maier, M. Leibundgut and N. Ban, *Science*, 2008, **321**, 1315–1322.
28 C. R. Valenzano, R. J. Lawson, A. Y. Chen, C. Khosla and D. E. Cane, *J. Am. Chem. Soc.*, 2009, **131**, 18501–18511.
29 D. H. Kwan, M. Tosin, N. Schlager, F. Schulz and P. F. Leadlay, *Org. Biomol. Chem.*, 2011, **9**, 2053–2056.
30 J. T. Zheng and A. T. Keatinge-Clay, *J. Mol. Biol.*, 2011, **410**, 105–117.
31 J. T. Zheng and A. T. Keatinge-Clay, *MedChemComm*, 2013, **4**, 34–40.
32 J. T. Zheng, S. K. Piasecki and A. T. Keatinge-Clay, *ACS Chem. Biol.*, 2013, **8**, 1964–1971.
33 J. T. Zheng, C. A. Taylor, S. K. Piasecki and A. T. Keatinge-Clay, *Structure*, 2010, **18**, 913–922.
34 P. Javidpour, A. Das, C. Khosla and S. C. Tsai, *Biochemistry*, 2011, **50**, 7426–7439.
35 P. Javidpour, T. P. Korman, G. Shakya and S. C. Tsai, *Biochemistry*, 2011, **50**, 4638–4649.
36 J. T. Zheng, D. C. Gay, B. Demeler, M. A. White and A. T. Keatinge-Clay, *Nat. Chem. Biol.*, 2012, **8**, 615–621.
37 M. Häckh, M. Müller and S. Lüdeke, *Chem. – Eur. J.*, 2013, **19**, 8922–8928.
38 A. Garg, C. Khosla and D. E. Cane, *J. Am. Chem. Soc.*, 2013, **135**, 16324–16327.
39 A. T. Keatinge-Clay, *Nat. Prod. Rep.*, 2012, **29**, 1050–1073.
40 A. T. Keatinge-Clay and R. M. Stroud, *Structure*, 2006, **14**, 737–748.
41 S. K. Piasecki, J. Zheng, A. J. Axelrod, M. E. Detelich and A. T. Keatinge-Clay, *Proteins: Struct., Funct., Bioinf.*, 2014, **82**, 2067–2077.
42 S. Smith and S. C. Tsai, *Nat. Prod. Rep.*, 2007, **24**, 1041–1072.
43 H. Zhou, Z. Z. Gao, K. J. Qiao, J. J. Wang, J. C. Vederas and Y. Tang, *Nat. Chem. Biol.*, 2012, **8**, 331–333.
44 R. Kong, L. P. Goh, C. W. Liew, Q. S. Ho, E. Murugan, B. Li, K. Tang and Z.-X. Liang, *J. Am. Chem. Soc.*, 2008, **130**, 8142–8143.
45 M. Kotada, R. Kong, I. Qureshi, Q. Ho, H. Sun, C. W. Liew, L. P. Goh, P. Cheung, Y. Mu, J. Lescar and Z.-X. Liang, *J. Biol. Chem.*, 2008, **284**, 15739–15749.
46 R. Higuchi, B. Krummel and R. Saiki, *Nucleic Acids Res.*, 1988, **16**, 7351–7367.
47 E. Dellus-Gur, A. Toth-Petroczy, M. Elias and D. S. Tawfik, *J. Mol. Biol.*, 2013, **425**, 2609–2621.
48 V. Anantharaman, L. Aravind and E. V. Koonin, *Curr. Opin. Chem. Biol.*, 2003, **7**, 12–20.
49 J. Bruegger, B. Haushalter, A. Vagstad, G. Shakya, N. Mih, C. A. Townsend, M. D. Burkart and S.-C. Tsai, *Chem. Biol.*, 2013, **20**, 1135–1146.
50 J. Crosby and M. P. Crump, *Nat. Prod. Rep.*, 2012, **29**, 1111–1137.
51 J. Lim, R. Kong, E. Murugan, C. L. Ho, Z.-X. Liang and D. Yang, *PLoS One*, 2011, **6**, e20549.
52 M. Leibundgut, S. Jenni, C. Frick and N. Ban, *Science*, 2007, **316**, 288–290.
53 Z. X. Liang, M. Jiang, Q. Ning and B. M. Hoffman, *J. Biol. Inorg. Chem.*, 2002, **7**, 580–588.
54 Z. X. Liang, I. V. Kurnikov, J. M. Nocek, A. G. Mauk, D. N. Beratan and B. M. Hoffman, *J. Am. Chem. Soc.*, 2004, **126**, 2785–2798.
55 T. Liu, Y.-M. Chiang, A. D. Somoza, B. R. Oakley and C. C. C. Wang, *J. Am. Chem. Soc.*, 2011, **133**, 13314–13316.
56 H.-H. Yeh, S.-L. Chang, Y.-M. Chiang, K. S. Bruno, B. R. Oakley, T.-K. Wu and C. C. C. Wang, *Org. Lett.*, 2013, **15**, 756–759.
57 X. Zhu, F. Yu, X.-C. Li and L. Du, *J. Am. Chem. Soc.*, 2006, **129**, 36–37.
58 K. M. Fisch, W. Bakeer, A. A. Yakasai, Z. Song, J. Pedrick, Z. Wasil, A. M. Bailey, C. M. Lazarus, T. J. Simpson and R. J. Cox, *J. Am. Chem. Soc.*, 2011, **133**, 16635–16641.
59 A. L. Vagstad, A. G. Newman, P. A. Storm, K. Belecki, J. M. Crawford and C. A. Townsend, *Angew. Chem., Int. Ed.*, 2013, **52**, 1718–1721.
60 T.-Y. Lin, L. S. Borketey, G. Prasad, S. A. Waters and N. A. Schnarr, *ACS Synth. Biol.*, 2013, **2**, 635–642.
61 H. Sun, R. Kong, D. Zhu, W. Lu, Q. Ji, C. W. Liew, J. Lescar, G. Zhong and Z.-X. Liang, *Chem. Commun.*, 2009, 7399–7401.
62 T. Schwede, J. Kopp, N. Guex and M. C. Peitsch, *Nucleic Acids Res.*, 2003, **31**, 3381–3385.
63 R. Das and D. Baker, *Annu. Rev. Biochem.*, 2008, **77**, 363–382.
64 A. Leaver-Fay, M. Tyka, S. M. Lewis, O. F. Lange, J. Thompson, R. Jacak, K. Kaufman, P. D. Renfrew, C. A. Smith, W. Sheffler, I. W. Davis, S. Cooper, A. Treuille, D. J. Mandell, F. Richter, Y. E. A. Ban, S. J. Fleishman, J. E. Corn, D. E. Kim, S. Lyskov, M. Berrondo, S. Mentzer, Z. Popovic, J. J. Havranek, J. Karanicolas, R. Das, J. Meiler, T. Kortemme, J. J. Gray, B. Kuhlman, D. Baker and P. Bradley, *Methods Enzymol., 487: Comp. Meth., Pt C*, 2010, 545–574.