












# In vitro Cas9-assisted editing of modular polyketide synthase genes to produce desired natural product derivatives

Kei Kudo <sup>1,6</sup>, Takuya Hashimoto <sup>1,6</sup>, Junko Hashimoto <sup>2</sup>, Ikuko Kozone <sup>2</sup>, Noritaka Kagaya <sup>1</sup>, Reiko Ueoka <sup>1</sup>, Takehiro Nishimura <sup>3</sup>, Mamoru Komatsu <sup>4</sup>, Hikaru Suenaga <sup>1</sup>, Haruo Ikeda <sup>4</sup>✉ & Kazuo Shin-ya <sup>1,3,5</sup>✉

One major bottleneck in natural product drug development is derivatization, which is pivotal for fine tuning lead compounds. A promising solution is modifying the biosynthetic machineries of middle molecules such as macrolides. Although intense studies have established various methodologies for protein engineering of type I modular polyketide synthase(s) (PKSs), the accurate targeting of desired regions in the PKS gene is still challenging due to the high sequence similarity between its modules. Here, we report an innovative technique that adapts in vitro Cas9 reaction and Gibson assembly to edit a target region of the type I modular PKS gene. Proof-of-concept experiments using rapamycin PKS as a template show that heterologous expression of edited biosynthetic gene clusters produced almost all the desired derivatives. Our results are consistent with the promiscuity of modular PKS and thus, our technique will provide a platform to generate rationally designed natural product derivatives for future drug development.

<sup>1</sup>National Institute of Advanced Industrial Science and Technology (AIST), 2-4-7 Aomi, Koto-ku, Tokyo, Japan. <sup>2</sup>Japan Biological Informatics Consortium (JBIC), 2-4-32 Aomi, Koto-ku, Tokyo, Japan. <sup>3</sup>Technology Research Association for Next Generation Natural Products Chemistry, 2-4-7 Aomi, Koto-ku, Tokyo, Japan. <sup>4</sup>Kitasato Institute for Life Sciences, Kitasato University, 1-15-1 Kitasato, Minami-ku, Sagami-hara, Kanagawa, Japan. <sup>5</sup>Biotechnology Research Center, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo, Japan. <sup>6</sup>These authors contributed equally: Kei Kudo, Takuya Hashimoto. ✉email: [iked@ls.kitasato-u.ac.jp](mailto:iked@ls.kitasato-u.ac.jp); [kshinya@aist.go.jp](mailto:kshinya@aist.go.jp)

Recently, middle-weight molecular compounds have attracted increasing attention as new compounds to fill the gap in molecular weight between macromolecules, such as antibodies, and low-molecular-weight compounds. In particular, as Fabrizio Giordanetto et al. reported<sup>1</sup>, macrocycles can attack large surface of target, such as protein–protein interaction (PPI), due to their size and complexity. Both their incomplete rigidity and moderate flexibility influence the overall structure of macrocycles, which are essential properties for drug development, resulting in suitable membrane permeability<sup>2</sup> and desirable overall pharmacokinetics<sup>3,4</sup>. Even today, natural products are of great importance as drug leads<sup>5</sup>; however, the significant decrease in the rate of the discovery of novel compounds has become a big problem. In addition, difficulties in derivatizing macrocyclic natural compounds by chemical synthesis due to their structural complexity are also a bottleneck in drug development<sup>6</sup>. Therefore, various congeners will be produced from known useful natural compounds by genetic manipulation.

Among natural products, macrolides, such as erythromycin and rapamycin, are considered good models of macrocycles. These compounds are biosynthesized via an assembly line process that is catalysed by modular polyketide synthases (PKSs) (Supplementary Fig. 1). Despite the usefulness of macrolides in drug development, the genetic modification of modular PKS gene clusters is extremely difficult. The multi-modular architecture of PKSs may have evolved from a single module PKS by gene duplication events<sup>7–9</sup>. Accordingly, the sequences of the modules in modular PKS clusters are highly homologous (more than 70% on average). With these contexts, many have tried to modify type I PKS to produce new analogues of targeted polyketides and proposed several genetic strategies, which still have potential limitations on applying them to larger type I PKS compounds<sup>10–13</sup>. Among these reports, Aleksandra Wlodek et al. recently published a recombination-based methodology to derivatize rapamycin<sup>10</sup>. However, the desired compounds could not be obtained since the homologous recombination occurred so randomly across the PKS modules that could not be under controlled. Alternatively, Robert McDaniel et al. manipulated 6-deoxyerythronolide B synthase by utilising a pair of unique restriction sites as an editing point<sup>11</sup>. The availability of suitable restriction sites is key to their methodology; however, the usage of restriction sites limits its broad applicability since it becomes increasingly difficult to find or introduce artificial restriction sites as the number of modules increases.

For the in-will production of new polyketide skeletons by applying former or even future knowledge to PKS engineering, the credible editing of all type I PKS genes is necessary. However, since the well-known methods described above have obstacles due to the high sequence similarities among PKS modules/domains, we developed a methodology that employs the CRISPR-Cas9 reaction in vitro to avoid unnecessary off-target reactions and simplify sequence analyses of edited constructs. To precisely engineer against highly repetitive modular PKS genes, the establishment of a precise DNA digestion and subsequent seamless cloning method that does not depend on recombination or restriction digestion is of great importance. Herein, we report the in vitro module editing technique and its application to produce the desired rapamycin derivatives.

## Results

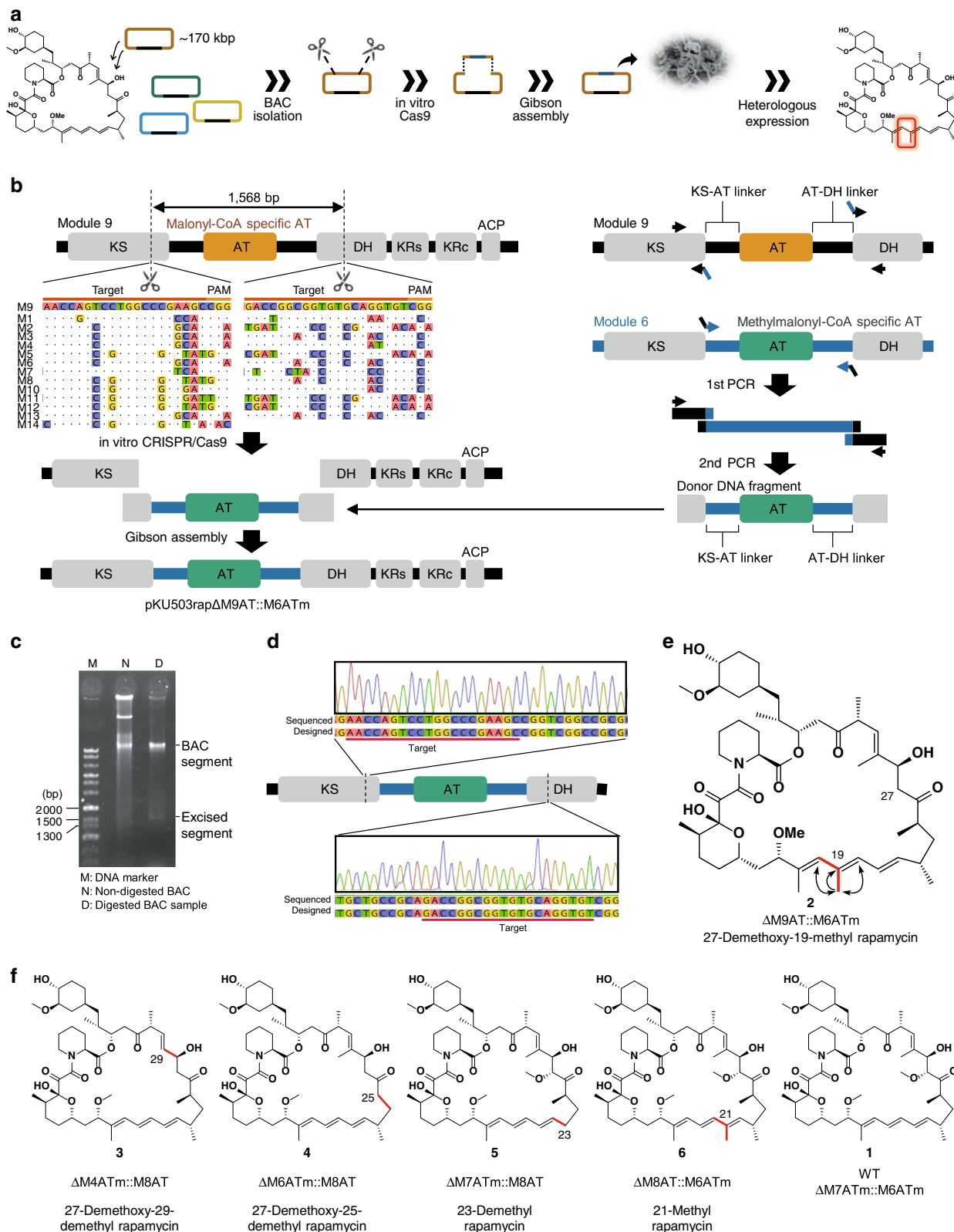
**Concept of in vitro module editing.** Based on the heterologous expression system we established that is applicable to biosynthetic clusters with over 200 kb cloned into the BAC vector<sup>14</sup>, we developed the idea of in vitro module editing, which applies a combination of both CRISPR-Cas9<sup>15</sup> and Gibson assembly<sup>16</sup> to

the BAC clone in vitro (Fig. 1a). Once we construct vectors harbouring artificial gene clusters, they can be expressed in suitable hosts, such as *Streptomyces avermitilis* SUKA strains<sup>17</sup>, to validate the productivity. Because the heterologous expression experiment does not allow unexpected mutations anywhere in the biosynthetic genes, this methodology guarantees accurate editing, which is its most important advantage. As a proof of concept, we selected rapamycin (1) (Supplementary Fig. 1) to establish the in vitro module editing system since some rapamycin derivatives<sup>10</sup> can be index for target derivatives (for a review of chemically synthesised derivatives of rapamycin, see ref. 18). If we can accurately edit the modular PKS of rapamycin, whose modules show higher homology to each other (the average homology between 14 KS domains is 91.6%) than those of other systems such as the erythromycin PKS (86.1% homology between 6 KS domains), this technique will be applicable for most natural products.

We established the heterologous production of rapamycin in *S. avermitilis* SUKA strains, since the production yield of rapamycin was high enough that rapamycin could be isolated, and was better in these strains than in other host strains. The biosynthetic gene cluster for rapamycin from *S. rapamycinicus* NRRL5491<sup>19</sup> was cloned into the BAC vector pKU503 to yield pKU503rap (166 kb, accession number: LC566301). In addition, the expression of the positive transcriptional regulator gene, *rapH*<sup>20</sup>, in the host strain (designated SUKA33::*rapH* or SUKA34::*rapH*) led to the greatest production of rapamycin (8.6 mg l<sup>-1</sup>) under our conditions (Supplementary Fig. 1).

**Acyltransferase domain-targeted module editing.** The minimum PKS module contains a set of three domains, a ketosynthase (KS), an acyltransferase (AT), and an acyl carrier protein (ACP), that catalyse a 2-carbon chain elongation (Supplementary Fig. 1). The sequence of AT domains controls the selection of methyl (alkyl) malonyl-CoA or malonyl-CoA as the extender units, which determines the  $\alpha$ -branched or non-branched substructures. Our first aim was to change the extender unit for a particular module of rapamycin PKS by replacing the AT domain with AT domains from the different modules of the same rapamycin PKS (i.e., the AT domain of module 6 for the methylmalonyl extender unit and the AT domain of module 8 for the malonyl extender unit. See also Supplementary Fig. 2).

We first targeted module 9 to replace a malonyl-CoA unit with a methylmalonyl-CoA unit. The corresponding position is the triene moiety of rapamycin which is reported to play a significant role in the biological activity<sup>21,22</sup>. To precisely cut out the nucleotide sequence encoding the AT domain in module 9 (M9AT) of the BAC clone pKU503rap, the CRISPR-Cas9 system was employed in vitro (Fig. 1b). A pair of sgRNAs was designed to target the region flanking M9AT (Supplementary Tables 1 and 2). The sequence alignment of the PKS modules was used to avoid the homologous region, especially the 3' regions of target sequences, which may cause off-target cleavage in other modules. As a result, a band for the digested BAC segment and another band corresponding to the targeted 1.6 kbp segment were observed by gel electrophoresis (Fig. 1c). The digested BAC segment was then subjected to Gibson assembly with a donor DNA fragment. The donor DNA fragment was prepared by overlap PCR and consisted of an exchanged domain (i.e., an AT domain from another module) flanked by unchanged sequences (Fig. 1b and Supplementary Tables 1 and 2). The M9AT region with the KS-AT linker and AT-DH linker was replaced with the corresponding region of M6AT (Fig. 1b and Supplementary Fig. 3). The desired clone was screened by colony PCR (Supplementary Tables 1 and 2), and the sequence of the edited region was confirmed by Sanger sequencing (Fig. 1d).



**Fig. 1** The scheme for in vitro module editing and AT-exchange. **a** Concept art of in vitro module editing. **b** Practical workflow for the M9AT-exchange experiment. The target sequences of sgRNAs and the equivalent regions of repeated modules within the rapamycin PKS are aligned. The donor DNA fragment was prepared by overlap PCR of three fragments. Dotted lines indicate the excised position. **c** Gel electrophoresis of Cas9-digested pKU503rap. The experiments were performed five times independently with similar results. **d** Sequence confirmation of pKU503rap $\Delta$ M9AT::M6ATm. **e** Structure of the product for pKU503rap $\Delta$ M9AT::M6ATm. Arrows indicate the observed HMBC correlations. **f** Other AT-swapped rapamycin derivatives produced in this study.

The BAC vector pKU503 harbouring the edited artificial rapamycin biosynthetic gene cluster (pKU503rap $\Delta$ M9AT::M6ATm) was first transformed into *S. lividans* TK24 carrying the linear plasmid SAP1, which was then transferred into *S. avermitilis* SUKA34::rapH by exogenous conjugation<sup>23</sup>. As a result of cultivation, a UV absorption peak (280 nm) consistent with a triene was observed, as well as an MS peak corresponding to dehydroxy rapamycin (920.5489 (C<sub>52</sub>H<sub>82</sub>NO<sub>13</sub>Na<sup>+</sup>, [M + Na]<sup>+</sup>: -0.6 mmu)) (Supplementary Fig. 2) in the HR-ESI-MS data. Finally, the structure of the produced compound was confirmed by analysis of its NMR spectra such as HMBC and DQF-COSY (Supplementary Data 1–7). Analyses of the <sup>1</sup>H spin couplings together with the <sup>1</sup>H-<sup>13</sup>C long-range couplings, especially key HMBC correlations from a new methyl proton signal to C-18, C-19 and C-20, revealed that the methyl residue was attached to C-19. The oxymethine carbon, C-27, in rapamycin was replaced by a methylene carbon in this rapamycin derivative. These analyses established the structure as 27-demethoxy-19-methyl rapamycin (**2**) (Fig. 1e). This result indicated that the PKS function was altered as intended. In the same manner, we performed a series of AT exchanges targeting modules 4, 6, 7 and 8 (Supplementary Fig. 2). As a result, we succeeded in constructing each desired clone and observed the MS peaks and the MS/MS fragmentation patterns corresponding to 27-demethoxy-29-demethyl rapamycin (**3**), 27-demethoxy-25-demethyl rapamycin (**4**), 23-demethyl rapamycin (**5**) and 21-methyl rapamycin (**6**) (Fig. 1f and Supplementary Figs 2 and 4), all of which are novel rapamycin analogues. Compound **6** was isolated from scaled-up culture, and its chemical structure was confirmed by NMR analyses (Supplementary Data 8–14). These results indicate that the substrate tolerance of the PKS domains allows them to accept both  $\alpha$ -methylated and  $\alpha$ -demethylated biosynthetic intermediates.

**Production of module deletion and insertion derivatives.** Following the production of AT-exchanged derivatives, we turned to module deletion (Fig. 2a) and insertion (Fig. 2b) to further examine the promiscuity of PKS modules towards upstream ACP-bound substrates. We carried out single module deletion processed targeting modules 6, 8 and 10 (Fig. 2a and Supplementary Figs 3 and 5). Artificial constructs designated pKU503rap $\Delta$ M6, pKU503rap $\Delta$ M8 and pKU503rap $\Delta$ M10 were prepared and introduced into SUKA strains. As a result, MS peaks corresponding to three new 29-membered rapamycin derivatives (**7**, **8** and **9**) were observed (Fig. 2c and Supplementary Fig. 6). The structures of the compounds **7** and **9** were determined by NMR analyses (Supplementary Data 15–28), and the production of compound **8** was much lower than that of **7** and **9** (Supplementary Table 3).

Using our *in vitro* module editing technique, we also confirmed the production of derivatives in the previous research<sup>10</sup> to prove whether we can rationally target specific regions. Unlike their results of accidental recombination products, we accurately prepared pKU503rap $\Delta$ M3, pKU503rap $\Delta$ M2ACP- $\Delta$ M4KR, pKU503rap $\Delta$ M3KS- $\Delta$ M6KS and pKU503rap $\Delta$ M2AT- $\Delta$ M8AT, which corresponded to ISOM-4280, 4144, 4185 and 4193 (producing 29-, 27- and 25-membered rings and acyclic rapamycin analogues), respectively. As a result of the heterologous expression in SUKA33::rapH, the production of three of the four derivatives (**10**, **11** and **12**, but not the 25-membered ring rapamycin analogue) was confirmed by the combination of HR-ESI-MS and NMR analyses (Supplementary Fig. 6 and Supplementary Data 29–49).

For the expansion of the macrocycle, we inserted module 12 between modules 2 and 3 to biosynthesize a 33-membered ring derivative **13**, which is the same compound as that produced by

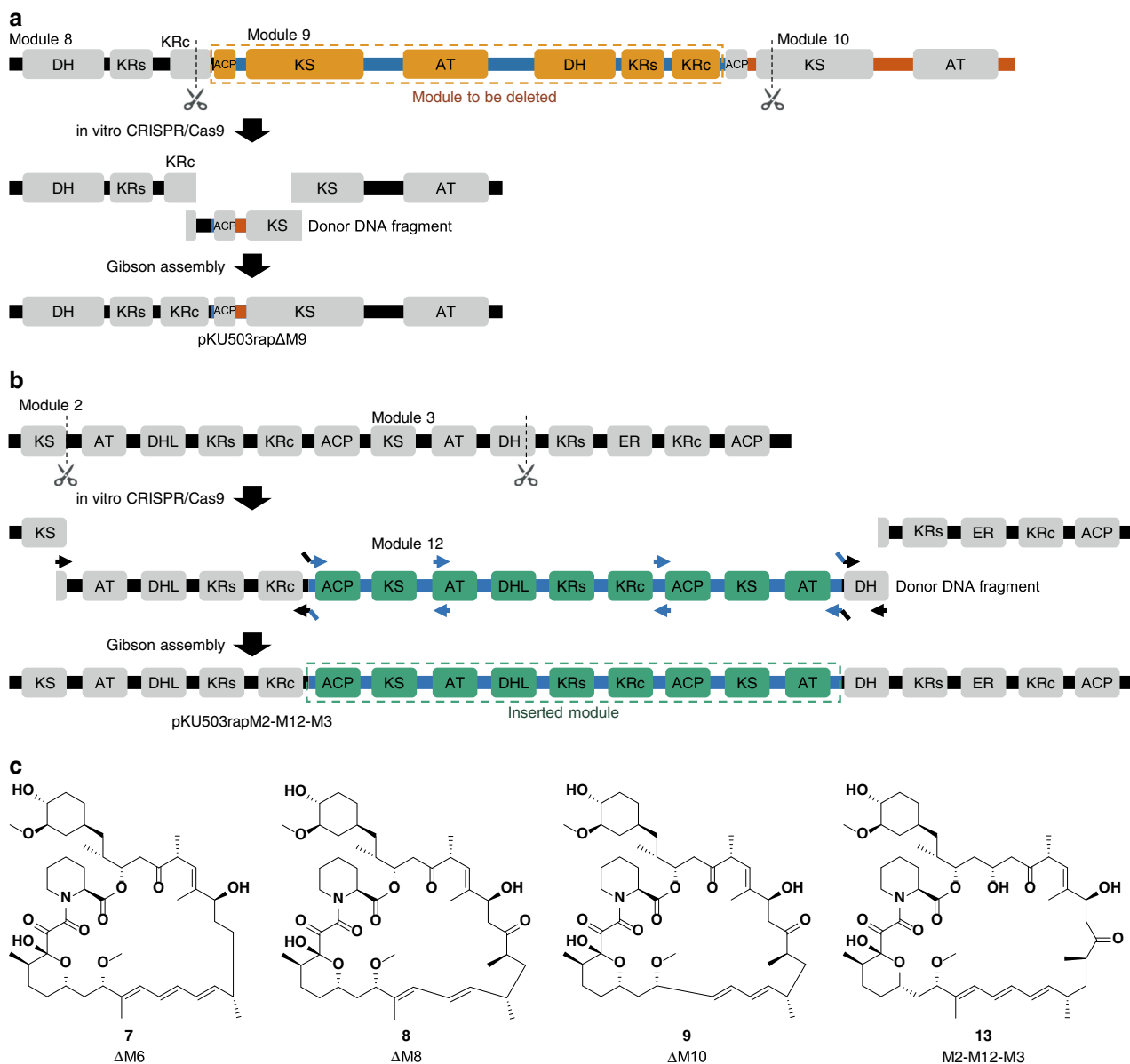
ISOM-4309<sup>10</sup>. Since the length of the donor DNA fragment was much greater than the length of one module unit (~11 kbp), the donor DNA fragment was divided into five segments that could be amplified by PCR, and they were concatenated by Gibson assembly to afford the desired sequence. Using this long donor DNA fragment led to the construction of pKU503rap\_M2-M12-M3 (Fig. 2b). As shown in this example, *in vitro* module editing can accurately generate the intended products even with large gene units. As expected, this module insertion derivative was the same as that produced in Aleksandra Wlodek's study (Supplementary Fig. 6 and Supplementary Data 50–56).

**Editing of reductive loop domains.** Depending on the module, there may be additional processing enzymes, namely, a ketoreductase (KR), a dehydratase (DH), and an enoyl reductase (ER). This set of contiguous domains is called a reductive loop<sup>24</sup>, and it regulates the oxidation level around the  $\beta$ -keto groups of nascent chains (Supplementary Fig. 1). The manipulation of reductive loops leads to dramatic changes in the structure of the product and often impacts the biological activities of the macrolide compounds. Therefore, we focused on editing the triene moiety. We targeted module 7 to make a tetraene moiety, as follows. We constructed two clones; one was the loss-of-function mutant of the NADPH binding site in the ER domain in which GGVGMA was replaced with SPVGMA<sup>25</sup>, and the other was the mutant with the entire ER domain deleted. The KR domain consists of a structural subunit (KRs) and a catalytic subunit (KRc). The ER domain resides in the linker between KRs and KRc, as shown in Fig. 3a<sup>26</sup>. To construct the M7ER-deleted mutant, the M7ER region was replaced with the corresponding linker in M8KR.

The resultant constructs, pKU503rap\_M7ER<sup>0</sup> and pKU503rap $\Delta$ M7ER, were transformed into SUKA54, and the culture extract was monitored by UPLC-TOF-MS. A peak with a UV absorption characteristic of a tetraene moiety was observed, and its MS signal corresponded to the desired new rapamycin derivative (**14**) (Fig. 3c and Supplementary Fig. 7). The structure of **14** was established by NMR spectroscopy (Supplementary Data 57–63).

We next tried to remove the conjugated triene functionality and produce a 1,5-diene moiety by targeting the olefin at the centre of the triene substructure (Fig. 3c). The DH-KR didomain of module 9 was swapped into the DH-ER-KR tridomain of module 7 (Fig. 3b). Based on previous studies, we exchanged the two reductive loops between the portion immediately after the conserved RYW sequence at the AT-DH linker and the (A/V)(Q/R)W sequence at the KR-ACP linker (Fig. 3b)<sup>24</sup>. The resulting construct, pKU503rap $\Delta$ M9DH-KR::M7DH-ER-KR, was expressed in SUKA34::rapH, and the extract of the culture was monitored by UPLC-TOF-MS. The peak showing a UV absorption characteristic of a triene moiety disappeared, while the HR-ESI-MS results were consistent with the reduced rapamycin derivative (**15**) (Supplementary Fig. 7). As expected, NMR analysis revealed the chemical structure of **15** to be 19,20-dihydro-rapamycin (Supplementary Data 64–70), which to our knowledge is a newly isolated rapamycin derivative.

In addition, we replaced the inactive DH-ER-KR tridomain of module 3 with the active one of module 7 to produce 32-desketorapamycin (**16**) (Fig. 3c) which could not be obtained by the homologous recombination-based system<sup>10</sup>. The construct, pKU503rap $\Delta$ M3DH-ER-KR::M7DH-ER-KR, was prepared, as well as pKU503rap $\Delta$ M9DH-KR::M7DH-ER-KR, and was expressed in SUKA34::rapH. Analysis of the culture extract using UPLC-TOF-MS indicated the product possessed distinctive rapamycin UV absorption and had a consistent *m/z* value of **16** (Supplementary Fig. 7). The chemical structure of compound **16**



**Fig. 2** Module deletion and insertion. Schematic representations of module deletion (**a**) and insertion (**b**). **c** Rapamycin derivatives produced in this study.

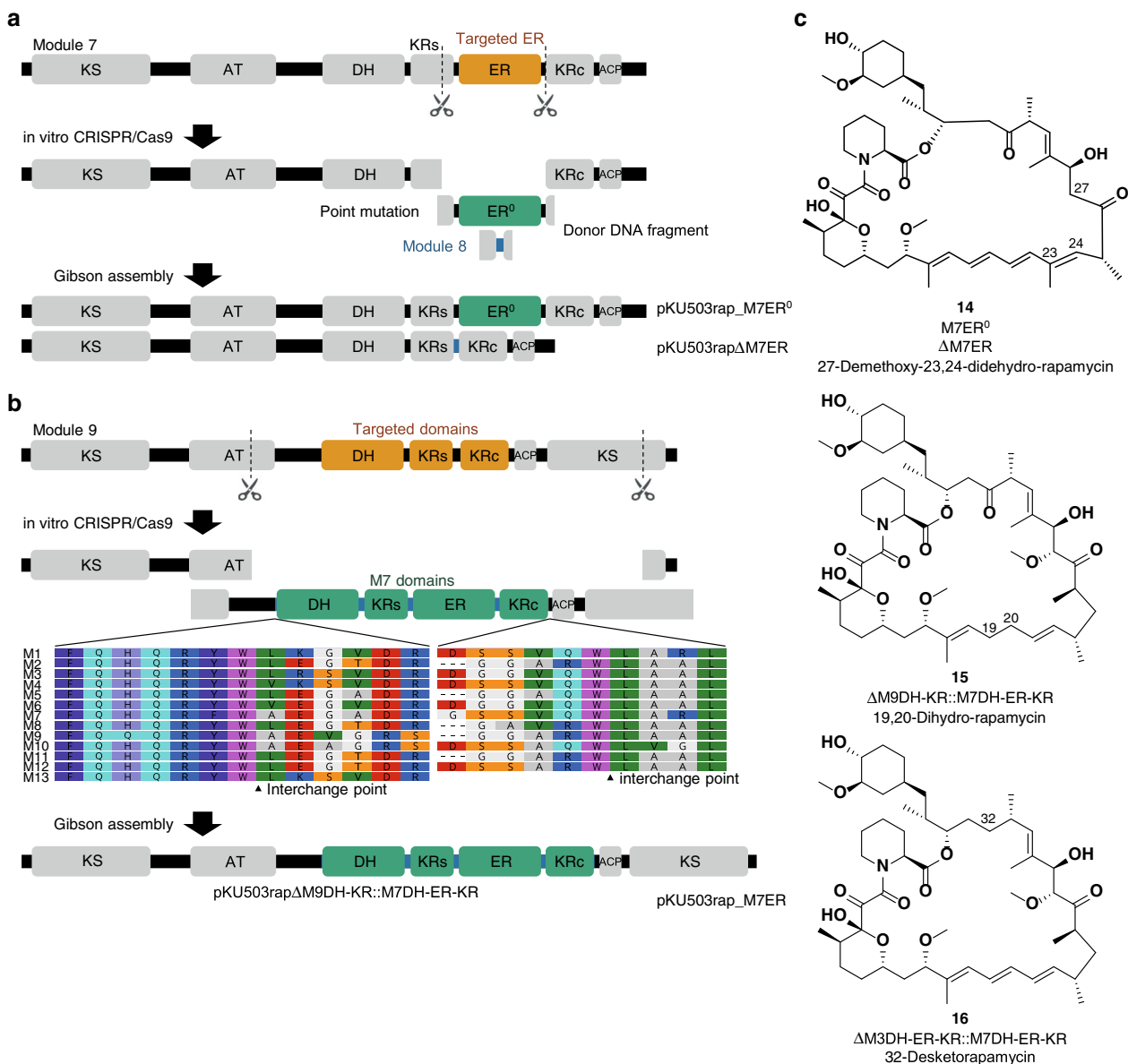
was further confirmed by NMR analysis (Supplementary Data 71–77). Compound **16** was chemically synthesised from rapamycin via reduction of the ketone at the C-32<sup>27</sup>, and have been tried clinical application of coronary stent therapy<sup>28</sup>.

**Inversion of stereochemistry and double modification.** Based on the successful exchange of the reductive loop, we tried to change the absolute configuration of a hydroxy functional group. The configuration of the  $\beta$ -hydroxy groups is controlled by the KR domains which are classified into A-type and B-type<sup>29</sup>; an A-type KR domain results in the L configuration, while a B-type KR domain affords the D configuration at the  $\beta$ -position of the ACP-bound nascent precursor. Since all the KR domains in rapamycin PKS are B-type domains, we adopted A-type KR domains from other macrolides to invert the hydroxy groups of rapamycin. We targeted the KR domain in module 11 (M11KR) for replacement by two A-type KR domains, the bafilomycin KR domain from module 2 (bfmM2KR) and the leucomycin KR domain from module 7 (lcmM7KR) (Fig. 4a). The interchange points were the

same as those in the cases of pKU503rapΔM9DH-KR::M7DH-ER-KR (Supplementary Fig. 3).

Based on the results of the heterologous expression in SUKA34::rapH, these artificial biosynthetic gene clusters were revealed to produce a rapamycin derivative that showed a triene-derived UV absorption band and an HR-MS and MS/MS signal consistent with demethyl rapamycin (Supplementary Figs 7 and 8). The NMR spectra indicated that the planar structure of this derivative is 16-O-demethyl rapamycin (Supplementary Data 78–84). The stereochemistry of **17** was determined as (16R)-16-O-demethyl rapamycin since **17** was identical to the chemically prepared (16R)-epimer (Supplementary Fig. 9, for more detail, please see Determination of the configuration of **17** section in Supplementary Information)<sup>30</sup>.

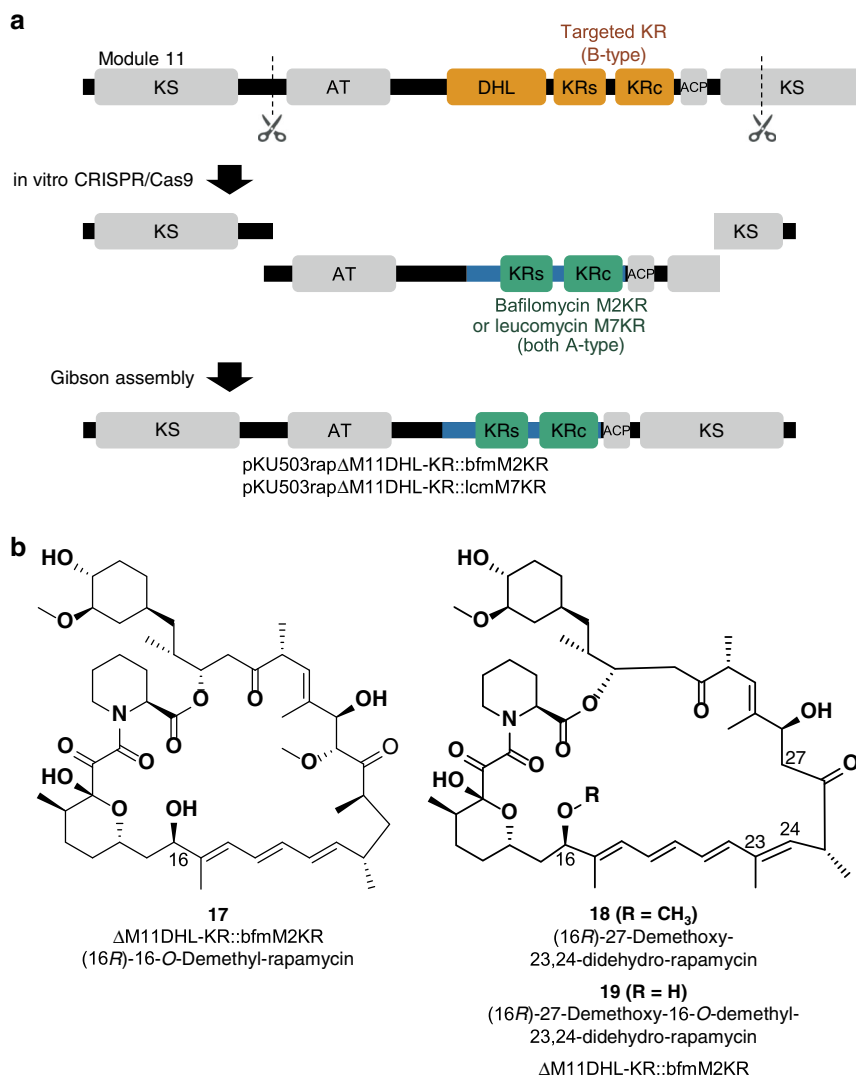
To optimise candidate compounds for drug development, modifying several substructures in the skeleton is often necessary. Therefore, we attempted to simultaneously edit multiple modules. The combined modification construct involving both the ER domain dysfunctional mutation of module 7 (pKU503rap\_M7ER<sup>0</sup>,



**Fig. 3 Schematic representation of the swapping of the reductive loop domains. a** Constructions of the ER-inactivation and ER-deletion mutants. **b** Construction of the reductive loop-swapped mutant. **c** Structures of the rapamycin derivatives produced in this section.

see the former section) and the stereochemical inversion mutation of M11KR was prepared via replacement with bfmM2KR. As a result of the heterologous expression of this double-modified construct in *SUKA34::rapH*, while its production was relatively low (Supplementary Table 3), two rapamycin derivatives showing tetraene-derived UV absorption bands were detected. Based on the MS/MS analyses (Supplementary Fig. 8), these derivatives were deduced to be 27-demethoxy-23,24-didehydro-rapamycin (**18**) and 27-demethoxy-16-*O*-demethyl-23,24-didehydro-rapamycin (**19**). This result suggested that type-I PKS is highly flexible enough to accept multiple editing. Compound **18** was deduced to have a methoxy group at the C-16 position, while the corresponding methylated derivative of compound **17** was not observed. The responsible methyltransferase, RapM, was reported to accept various alkyl group donors other than *S*-adenosyl methionine<sup>31</sup>. While the relaxed specificity of RapM against polyketide substrates may describe this contradiction, further biochemical and structural analyses are required.

**Biological activities.** Rapamycin shows its immunosuppressive activity via enhancing the interaction between FKBP12 and mTOR resulting in the formation of three components complex (FKBP12-rapamycin-mTOR)<sup>21</sup>. We employed fluorescent based PPI detecting technology Fluoppi (Fluorescent based technology detecting Protein-Protein Interactions) analysis to evaluate the binding activities of rapamycin derivatives simultaneously to FKBP12 and mTOR. Fluoppi harbours a tetrameric fluorescent protein hAG tag and an oligomeric assembly helper Ash tag to create detectable fluorescent foci in cells when there are interactions between two proteins fused to each tag<sup>32</sup>. As the results, **15**, in which the triene substructure was replaced by a 1,5-diene moiety, showed potent binding enhancing activity with comparable level to that of rapamycin (**1**) (Supplementary Fig. 10 and Table 1). This result was interesting since the triene moiety in rapamycin is considered to play a significant role for the binding enhancing activities<sup>21</sup>. This, to our knowledge, is the first report of the potency of **15**, although its detailed mode of action



**Fig. 4** Inversion of the stereochemistry of one of the hydroxyl groups. **a** Schematic representation of the swapping of the reductive loop of module 11 with the corresponding region of module 2 from bafilomycin PKS. **b** Structures of **17**, **18** and **19**.

**Table 1** Biological activities of rapamycin and its derivatives to induce the formation of FKBP-mTOR complex.

	EC <sub>50</sub> (nM)
<b>1</b>	9.0
<b>2</b>	45
<b>7</b>	320
<b>9</b>	34
<b>10</b>	21
<b>11</b>	n/a
<b>13</b>	73
<b>14</b>	2130
<b>15</b>	6.3
<b>16</b>	38
<b>17</b>	58

EC<sub>50</sub> values are calculated from three independent biological replicates. n/a, not active. Source data are provided as a Source Data file.

mechanisms are remains unclear. The other derivatives, **2**, **9**, **10**, **13**, **16** and **17**, showed relatively weaker activities for the induction of PPI between FKBP12 and mTOR than that of **1** (Table 1). Notably, **7** and **14** were much less potent than **1**, and **11** had

completely lost the ability to induce the formation of FKBP12-mTOR complex. It seemed that the geometrical changes of the whole macrocyclic ring structures reflected the effect of these derivatives, however, further analysis is required to design a derivative with desired activity.

In addition, we also performed the cytotoxic studies of these compounds against T lymphoma Jurkat cells, human ovarian adenocarcinoma SKOV-3 cells and malignant pleural mesothelioma MESO-1 cells, since rapamycin was reported to exhibit cytostatic effects against cancer cell lines<sup>33</sup>. As shown in Supplementary Fig. 11, cells treated with rapamycin (**1**) and most of its derivatives gradually reduced their viabilities and kept plateau levels at middle range indicating the cytostatic effects, while etoposide, utilised as a positive control, showed typical cytotoxic activities. In the dose-response curve of **1** and **16**, plateau levels for Jurkat, SKOV-3 and MESO-1 cells were around 25%, 60% and 75%, respectively. This is consistent with the previous literature<sup>34</sup> that **16** has comparable activity to **1**. Compounds **17**, **7**, **13**, **9**, **15** and **14**, in ascending order of the minimal inhibitory concentration, slightly suppressed the cell viabilities. Compounds **2**, **10** and **11** showed cytotoxic activities against the tested cell lines only at unusually high concentrations (over 10 μM) implying that biological activities of these

compounds were weakened. In addition, the PPI-inductive activities and the cytostatic effects seemed to have no significant relationship each other (Table 1 and Supplementary Fig. 10). For example, contrary to the potent PPI inductive activity, the cytostatic effects of **15** was ~100-fold lower than that of **1** and was almost the same magnitude of cytostaticity as that of **14** of which PPI-inductive activity is quite low. Thus, the derivatization by engineering of rapamycin PKS has potential to lead other type of active compound, proving the usefulness of module editing techniques.

## Discussion

While the in vivo CRISPR/Cas9 system is commonly employed in genome editing in eukaryotic cells (especially in mammalian cells), it is not applicable for the engineering of bacterial modular PKSs due to the occurrence of unpredictable recombination events. To date, only one application of the in vitro Cas9-assisted gene editing has been reported to manipulate the biosynthetic gene<sup>35</sup>. Yunkun Liu et al. developed a strategy for gene deletion and marker gene insertion on a cosmid vector of which size is up to 40 kb<sup>35</sup>. For another example, the combination of in vitro Cas9-digestion of the genomic DNA and Gibson assembly is capable of capturing up to 100 kb of targeted biosynthetic gene cluster<sup>36</sup>. In this study, we demonstrated that our in vitro module editing could introduce target DNA fragments from endogenous (rapamycin) and/or exogenous (bafilomycin and leucomycin) PKS genes into desired sites precisely even when the BAC exceeds 160 kb. Since modular PKSs are among the most difficult targets, this methodology should be applicable to other large biosynthetic genes, such as most other PKSs and non-ribosomal peptide synthetases. In the conventional method, the derivatives of the target compounds were produced by the wild-type producer. Many strategies for genome editing of *Streptomyces* using CRISPR/Cas9 technology in vivo have been developed<sup>37,38</sup>; however, it is still challenging to introduce a large DNA fragment into non-model host strains and to avoid off-targets especially in the PKS genes. In contrast, our approach is based on the high probability of heterologous expression in a versatile host. One of the most important characteristics of our system, which uses heterologous expression, is that we can produce a variety of derivatives of the target compounds with the same platform.

Engineering modular PKS enzymes has been a long-standing goal. The manipulation of PKS modules may lead to the loss of the intimate domain/module interactions required for precise nascent chain transfer, elongation reactions and  $\beta$ -keto processing<sup>39</sup> (Supplementary Fig. 1). Moreover, in the traditional understanding of the PKS reaction cascade, KSs are considered strict gatekeepers that recognise specific substrates for the chain elongation reactions<sup>39</sup>. However, from our experimental results on rapamycin along with the linear polyketides aureothin<sup>12</sup> and monensin<sup>13</sup>, substrate promiscuity is broad enough to accept various unnatural intermediates. Although there are still few successful examples of designed PKS, these cases are undoubtedly setting the stage for utilising the substrate tolerance of modular PKSs. These prospects encouraged us to establish a theoretical combinatorial biosynthesis for the production of desired polyketides based on this design, which includes computer simulations to find ways to increase the biological activities. By employing in vitro module editing, we can assess any designed genetic manipulation of a natural product without significant limitation. Further investigations on the module boundary and non-native interactions between exogenous modules have been conducted to conceptualise the combinatorial biosynthesis.

Recently, drug screenings have shifted from target-based screenings to phenotypic screenings since the candidates

obtained by target-based screening have low successful rate to become clinical drugs than those expected. In addition, more first-in-class drugs have been discovered from phenotypic screenings than from target-based screenings<sup>40</sup>. As recently developed advanced techniques, superior phenotypic screening models that have not been used previously, such as patient-derived iPS cells and patient-derived tissue organoids, can be employed. In these phenotypic models, natural products often afford better results. However, the identification of the exact target molecule is essential in modern drug development. Due to the difficulties associated with target identification, it has become a bottleneck in natural product-based drug development. Our technique simplifies the introduction of new functionalities such as hydroxy groups at desired positions for synthesising labelled derivatives that can act as bait. These scaffolds can also be used to synthesise the payload of antibody-drug conjugates. Very recently, 27-*O*-desmethyl rapamycin derivative was reported to show specific activity against mTORC1, indicating the derivatization of potent natural product such as rapamycin could induce alternative activity<sup>41</sup>. Therefore, this technique is likely to become indispensable in natural product-based drug development. We are currently succeeded in applying this technique to not only other PKS compounds, pikromycin and leucomycin, but also a non-ribosomal peptide compound, YM-254890, indicating that this technique can be widely available for the derivatization of natural compounds. The paradigm that the substrate specificity of PKS is flexible now become common in derivatizing the skeleton of natural products.

## Methods

**Bacterial strains and culture conditions.** The growth of *Escherichia coli* strains DH5 $\alpha$ , NEB10 $\beta$  (New England Biolabs Japan Inc., Tokyo, Japan) and GM2929 *hdsS::Tn10*<sup>42</sup> was achieved using LB containing 10 g of tryptone, 5 g of yeast extract and 5 g of NaCl in 1 l of deionized water adjusted to pH 7.5, and the solid medium (LA) was prepared by adding 15 g of agar to 1 l of LB medium. In the production of rapamycin and its derivatives, *Streptomyces avermitilis* SUKA carrying the biosynthetic gene cluster for each derivative was inoculated into a 50-ml test tube containing 15 ml of GSY medium containing 5 g of glucose, 15 g of soya flour, and 5 g of yeast extract in 1 l of deionized water, and the cells were cultured with reciprocal shaking at 320 rpm at 27 °C for 2 days. A 375  $\mu$ l aliquot of the vegetative culture was used to inoculate a 125-ml flask containing 15 ml of production medium containing 40 g of  $\beta$ -cyclodextrin, 20 g of Pharmamedia, 5 g of glycerol, 21.2 g of MES, 5 mg of ZnSO<sub>4</sub>·7H<sub>2</sub>O, 5 mg of CuSO<sub>4</sub>·5H<sub>2</sub>O, and 5 mg of MnCl<sub>2</sub>·4H<sub>2</sub>O in 1 l of deionized water adjusted to pH 6.0. The fermentation culture was carried out with a rotary shaker at 24 °C for 5 days at 180 rpm. Large-scale preparations of the products were carried out in a baffled 500-ml flask containing 100 ml of the production medium.

**Genome sequencing and construction of the BAC library.** Genome sequencing of the *S. hygroscopicus* NRRL5491 strain was performed using a PacBio RS II (Pacific Biosciences, Menlo Park, CA), and the sequence data were assembled using HGAP2 (Pacific Biosciences). A BAC library of *S. hygroscopicus* NRRL5491 was constructed following a previously reported protocol. Each of the obtained BAC clones was stored in a 384-well plate containing Plusgrow II (Nacalai Tesque, Kyoto, Japan) (100  $\mu$ g ml<sup>-1</sup> ampicillin and 20% glycerol) at -80 °C. A clone carrying the targeted gene cluster was screened by PCR amplification using two pairs of primers (rapF1: AACAGCCGAAGAAATGGCTGTGC/rapR1: GGCCCTCTCGAACTCCGTACCTC and rap F2GGTGGTTTCGTCATGCCTGTCTG/rapR2 GCTCTCCTTGAGCATCAGCCACTG) to amplify the upstream and downstream regions of the gene cluster, respectively. A clone, pKU503rap, was selected, and the inserted sequence was confirmed by end-sequencing.

**Introduction of edited BACs into *S. avermitilis* SUKA.** All clones were introduced into *S. lividans* TK24 carrying the SAPI1 vector (containing the synthetic sequence of *attB*<sub>PK38-1</sub>, *attB*<sub>R4</sub>, *attB*<sub>BT1</sub>, *attB*<sub>PC31</sub> and *attB*<sub>TG1</sub> of *S. avermitilis* MA-4680). All clones were integrated into the SAPI1 vector because the bacteriophage attachment sites (*attB*) were located in the SAPI1 vector but not in the chromosome.

The transfer of the SAPI1 vector containing the edited gene cluster for rapamycin was conducted as described in ref. 23. The objective transconjugants were confirmed by the antibiotic-resistance phenotype and the size of the linear plasmid by contour-clamped homogeneous electric field (CHEF)<sup>43</sup> electrophoresis using SAPI1::pKU503rap as the control linear plasmid.



**Cloning of *rapH* and expression in SUKA33, 34.** pKU565*tsr*-2cistron<sup>23</sup> was digested with *Hind*III and *Nhe*I, and then a small fragment was cloned into the same site of pKU592AT-*aac*(3)*IV* to give the expression vector pKU592AT-*aac*(3)*IV*-2cistron.

The *rapH* gene was amplified from pKU503rap as a template with PCR using a pair of primers (rapH\_Xba\_1F: GGCTCTAGATATGCCTGCCGTGGAGTGCTA TGAAGTGGACGCGG and rapH\_Hind\_1R: AAGAAGCTTCTAGACGAGTT CGGCGGACAGGTGCTGGCGGCC). The amplified DNA fragment was digested with *Hind*III/*Xba*I and cloned into the same site of pKU592AT-*aac*(3)*IV*-2cistron to afford pKU592AT-2cis-*rapH*. The resultant plasmid was introduced into *S. avermitilis* SUKA33 and 34 by intergeneric conjugation and integration. The integration of the plasmid was confirmed based on its apramycin resistance and by PCR analysis. The resistance marker (*aac*(3)*IV*) flanking the *loxP* sequences was removed by the expression of a gene encoding Cre recombinase. The *cre*-expression plasmid, pKU471::*pheS*<sup>A229G/T278A</sup>, was introduced into the exconjugants by conjugation. The exconjugants were grown on yeast extract-malt extract medium<sup>44</sup> containing 2% D-xylose to induce the expression of *cre*. After sporulation, the 4-chlorophenylalanine-resistant and thio streptom-sensitive and apramycin-sensitive progenies were selected, and the desired clones were defined as SUKA33::*rapH* and SUKA34::*rapH*, respectively.

**Cloning of *rapA*, *rapB* and *rapC*.** To prepare the PCR template, pKU503rap was digested with *Fsp*AI (ThermoFisher Scientific, Waltham, MA) to yield fragments containing *rapA* (27.0 kbp), *rapB* (28.2 kbp) and *rapC* (23.7 kbp), and each fragment was purified by CHEF electrophoresis. The largest fragment of *Nru*I-digested pKU518 was ligated with the purified *rapA*, *rapB* and *rapC* fragments to give pKU518rapA, pKU518rapB and pKU518rapC, respectively.

**Preparation of donor DNA fragments.** The donor DNA fragments were prepared by overlap PCR using a certain number of short fragments as template DNA. Each fragment was amplified by PCR with pKU518rapA, pKU518rapB or pKU518rapC as the appropriate template DNA using the primer pairs listed in Supplementary Tables 1 and 2. The resultant fragment DNAs were spliced by overlap PCR using the pairs of primers corresponding to the far ends.

**In vitro module editing of the rap gene cluster on BAC.** All sgRNAs were prepared using the EnGen sgRNA Synthesis Kit, *S. pyogenes* (New England Biolabs Japan Inc., Tokyo, Japan), according to the manufacturer's instructions. The target sequences are listed in Supplementary Table 2. For the in vitro Cas9 reaction, pKU503rap, a pair of sgRNAs (see Supplementary Table 1) and Cas9 (New England Biolabs Japan Inc., Tokyo, Japan) were mixed according to the manufacturer's instructions, and the reaction mixture was incubated at 37 °C for 120 min. The excision of the desired DNA fragment was confirmed by agarose gel electrophoresis (Fig. 1c). The purified BAC segment and the donor DNA fragment were subjected to Gibson assembly. After 50 min of incubation at 50 °C, the reaction mixture was treated with phenol/chloroform/isoamyl alcohol and then subjected to 2-propanol precipitation. The ligated DNA was transformed into *E. coli* NEB10 by electroporation. The transformants were selected on LA containing 100 µg ml<sup>-1</sup> tetracycline. The desired plasmids were further selected by colony PCR using a pair of primers listed in Supplementary Table 1 and confirmed by restriction digestion.

**UPLC/UV/MS analysis of the metabolites from the transformant.** Five millilitres of culture broth were extracted with 5 ml of *n*-butanol (*n*-BuOH), 1.5 ml of the *n*-BuOH extract was dried in vacuo, and the residue was dissolved in 400 µl of dimethyl sulfoxide. Analytical UPLC and HR-ESI-MS (positive mode) were performed using a Waters ACQUITY UPLC System (Waters, Taunton, MA) in conjunction with a BEH ODS column (2.1 i.d. × 100 mm, Waters), a Waters ACQUITY UPLC photodiode array eλ detector (Waters), and a XevoG2 ToF system (Waters). Mobile phase A was water +0.1% formic acid, and mobile phase B was acetonitrile +0.1% formic acid. The elution program was 5–100% B over 5 min and 90% B for 1 min at a flow rate of 0.8 ml min<sup>-1</sup>. Data acquisitions and analyses were performed by using MassLynX V4.1 software (Waters). The MS/MS analysis was performed with the Orbitrap Fusion Lumos Tribrid system (Thermo Fisher Scientific). Data acquisitions and analyses were performed by using Xcalibur V4.0.27.42 software (Thermo).

**Treatment of rapamycin with trifluoroacetic acid and water.** To a solution of rapamycin (1.0 mg) in CH<sub>2</sub>Cl<sub>2</sub> (0.5 ml) at -40 °C was added trifluoroacetic acid (20 µl) followed by a drop of water. After 10 min and the reaction mixture was partitioned between ethyl acetate and saturated aqueous sodium bicarbonate. Layers were separated and organic layer successively washed with brine and dried over anhydrous sodium sulphate. The resulting crude material was subjected to UPLC-MS and MS/MS analyses.

**Determination of the configuration of 17.** To confirm the configuration of 16-hydroxyl, we synthesised a diastereomeric mixture of 16-*O*-demethyl rapamycin from **1** by the treatment with trifluoroacetic acid and water (see above). Also, using module editing technique, we obtained the deletion mutant of *rapM*, which

encodes the 16-*O*-specific methyltransferase<sup>31,45</sup>, to prepare (16*S*)-16-*O*-demethyl rapamycin (**20**) (Supplementary Data 85–91; indicating the same planar structure to **17**). UPLC analysis revealed that **17** was identical to the former peak of the diastereomeric mixture of 16-*O*-demethyl rapamycin (Supplementary Fig. 9) of which (16*R*)-epimer shows earlier retention time than the (16*S*)-epimer<sup>50</sup>. As expected, **20** was identical to the latter peak, the (16*S*)-epimer (Supplementary Fig. 9). Thus, we concluded that **17** is (16*R*)-16-*O*-demethyl rapamycin.

#### Isolation of compounds

**Isolation of 2 from SUKA34::rapH/pKU503rapΔM9AT::M6ATm.** Four litres of fermentation broth of SUKA34::rapH/pKU503rapΔM9AT::M6ATm was centrifuged to obtain a mycelial cake, which was extracted twice with 500 ml of acetone. The acetone was removed in vacuo, and the residual aqueous layer was extracted twice with ethyl acetate (EtOAc). The resultant EtOAc layer was concentrated in vacuo to afford 1.9 g of crude extract. The crude extract was subjected to medium-pressure liquid chromatography (MPLC) on silica gel (SNAP Ultra 25 g, Biotage, Uppsala, Sweden) eluted with a gradient system of *n*-hexane–EtOAc (0–25% EtOAc) followed by a stepwise solvent system of chloroform (CHCl<sub>3</sub>)–methanol (MeOH) (0, 1, 3, 5, 10, 50 and 90% MeOH). The 3% MeOH fraction (242 mg) was collected and subjected to silica gel MPLC (SNAP Ultra 25 g) with isocratic elution with 3% MeOH in CHCl<sub>3</sub>. The fractions were monitored by UPLC analysis, and fractions containing **2** were collected (48.1 mg). The sample was further purified by preparative reversed-phase HPLC using a CAPCELL PAK MG-II C18 column (5.0 µm, 20 i.d. × 150 mm; Shiseido, Tokyo, Japan). **2** was detected using a 2996 photodiode array detector (Waters) and a 3100 mass detector (Waters) following elution with 80% aqueous acetonitrile, and 4.4 mg of **2** was obtained.

**Isolation of 6 from SUKA34::rapH/pKU503rapΔM8AT::M6ATm.** Six litres of fermentation broth of SUKA34::rapH/pKU503rapΔM8AT::M6ATm was centrifuged to obtain the mycelial cake, which was extracted twice with 500 ml of acetone. The acetone was removed in vacuo, and the residual aqueous layer was extracted twice with EtOAc. The resultant EtOAc layer was concentrated in vacuo to afford 3.3 g of crude extract. The crude extract was subjected to MPLC on silica gel (SNAP Ultra 25 g) eluted with a gradient system of *n*-hexane–EtOAc (0–25% EtOAc) followed by a stepwise gradient of CHCl<sub>3</sub>–MeOH (0, 1, 3, 5, 10, 50 and 90% MeOH). The 3% MeOH fraction (441 mg) was collected and subjected to silica gel MPLC (SNAP Ultra 25 g) with isocratic elution with 3% MeOH in CHCl<sub>3</sub>. The fractions were monitored by UPLC analysis, and fractions containing **6** were collected (209 mg). The sample was further purified by preparative reversed-phase HPLC using a CAPCELL PAK MG-II C18 column (5.0 µm, 20 i.d. × 150 mm; Shiseido) with 70% aqueous acetonitrile, yielding 1.6 mg of **6**.

**Isolation of 7 from SUKA33::rapH/pKU503rapΔM6.** Three litres of fermentation broth of SUKA33::rapH/pKU503rapΔM6 were centrifuged to obtain a mycelial cake, which was extracted twice with 500 ml of acetone. The acetone was removed in vacuo, and the residual aqueous layer was extracted twice with EtOAc. The resultant EtOAc layer was concentrated in vacuo to afford 1.77 g of crude extract. The crude extract was subjected to silica gel MPLC (SNAP Ultra 25 g, Biotage) eluted with a gradient system of *n*-hexane–EtOAc (0–25% EtOAc) followed by a stepwise gradient of CHCl<sub>3</sub>–MeOH (0, 1, 5, 10, 50 and 90% MeOH). The 5% MeOH fraction (269 mg) was collected and subjected to silica gel MPLC (SNAP Ultra 25 g) with isocratic elution with 5% MeOH in CHCl<sub>3</sub>. The fractions were monitored by UPLC analysis, and fractions containing **7** were collected (184 mg). The sample was further purified by preparative reversed-phase HPLC using a CAPCELL PAK MG-II C18 column (5.0 µm, 20 i.d. × 150 mm; Shiseido) with 75% aqueous acetonitrile as the eluent, and 4.5 mg of **7** was obtained.

**Isolation of 9 from SUKA33::rapH/pKU503rapΔM10.** Three litres of fermentation broth of SUKA33::rapH/pKU503rapΔM10 was centrifuged to obtain a mycelial cake, which was extracted twice with 500 ml of acetone. The acetone was removed in vacuo, and the residual aqueous layer was extracted twice with EtOAc. The resultant EtOAc layer was concentrated in vacuo to afford 1.67 g of crude extract. The crude extract was subjected to silica gel MPLC (SNAP Ultra 25 g, Biotage) eluted with a gradient system of *n*-hexane–EtOAc (0–25% EtOAc) followed by a stepwise gradient of CHCl<sub>3</sub>–MeOH (0, 1, 5, 10, 50, 90% MeOH). The 5% MeOH fraction (345 mg) was collected and subjected to silica gel MPLC (SNAP Ultra 25 g) with isocratic elution with 5% MeOH in CHCl<sub>3</sub>. The fractions were monitored by UPLC analysis, and the fractions containing **9** were collected (174 mg). The sample was further purified by Sephadex LH-20 column chromatography (1:1 CHCl<sub>3</sub>/CH<sub>3</sub>OH) to obtain a fraction containing **9** (72.6 mg), which was then subjected to preparative reversed-phase HPLC using a CAPCELL PAK MG-II C<sub>18</sub> column (5.0 µm, 20 i.d. × 150 mm; Shiseido) with 65% aqueous acetonitrile as the eluent, and 12 mg of **9** was obtained.

**Isolation of 10 from SUKA33::rapH/pKU503rapΔM3.** Six litres of fermentation broth of SUKA33::rapH/pKU503rapΔM3, was centrifuged to obtain a mycelial cake, which was extracted twice with 500 ml of acetone. The acetone was removed in vacuo, and the residual aqueous layer was extracted twice with EtOAc. The

resultant EtOAc layer was concentrated in vacuo to afford 2.67 g of crude extract. The crude extract was subjected to silica gel MPLC (SNAP Ultra 25 g, Biotage) eluted with a gradient of *n*-hexane–EtOAc (0–25% EtOAc) followed by a stepwise gradient of CHCl<sub>3</sub>–MeOH (0, 1, 3, 5, 10, 20% MeOH). The 3% MeOH fraction (382 mg) was collected and subjected to silica gel MPLC (SNAP Ultra 25 g) with isocratic elution with 3% MeOH in CHCl<sub>3</sub>. The fractions were monitored by UPLC analysis, and the fractions containing **10** were collected (193 mg). The sample was subjected to Sephadex LH-20 column chromatography (1:1 CHCl<sub>3</sub>/CH<sub>3</sub>OH) and afforded a fraction containing **10** (45.9 mg). This sample was further purified by preparative reversed-phase HPLC using a CAPCELL PAK MG-II C<sub>18</sub> column (5.0 μm, 20 i.d. × 150 mm; Shiseido), and elution with 80% aqueous acetonitrile supplemented with 0.1% formic acid yielded 4.7 mg of **10**.

**Isolation of 11 from SUKA33::rapH/pKU503rapΔM2ACP-ΔM4KR.** Four litres of fermentation broth of SUKA33::rapH/pKU503rapΔM3, was centrifuged to obtain a mycelial cake, which was extracted twice with 500 ml of acetone. The acetone was removed in vacuo, and the residual aqueous layer was extracted twice with EtOAc. The resultant EtOAc layer was concentrated in vacuo to afford 2.79 g of crude extract. The crude extract was subjected to silica gel MPLC (SNAP Ultra 25 g, Biotage) eluted with a gradient of *n*-hexane–EtOAc (0–25% EtOAc) followed by a stepwise gradient of CHCl<sub>3</sub>–MeOH (0, 1, 3, 5, 10, 20% MeOH). The 3% MeOH fraction (341.2 mg) was collected and subjected to silica gel MPLC (SNAP Ultra 25 g) with isocratic elution with 3% MeOH in CHCl<sub>3</sub>. The fractions were monitored by UPLC analysis, and the fractions containing **11** were collected (70.1 mg). The sample was subjected to Sephadex LH-20 column chromatography (1:1 CHCl<sub>3</sub>/CH<sub>3</sub>OH) and afforded a fraction containing **11** (14.8 mg). This sample was further purified by preparative reversed-phase HPLC using a CAPCELL PAK MG-II C<sub>18</sub> column (5.0 μm, 20 i.d. × 150 mm; Shiseido), and elution with 80% aqueous acetonitrile supplemented with 0.1% formic acid yielded 4.7 mg of **11**.

**Isolation of 12 from SUKA33::rapH/pKU503rapΔM2AT-ΔM8AT.** Four litres of fermentation broth of SUKA34::rapH/ΔM2AT-ΔM8AT was centrifuged to obtain the mycelial cake, which was extracted twice with 500 ml of acetone. The acetone was removed in vacuo, and the residual aqueous layer was extracted twice with EtOAc. The resultant EtOAc layer was concentrated in vacuo to afford 1.29 g of crude extract. The crude extract was subjected to MPLC on silica gel (SNAP Ultra 25 g) eluted with a gradient system of *n*-hexane–EtOAc (0–25% EtOAc) followed by a stepwise gradient of CHCl<sub>3</sub>–MeOH (0, 1, 3, 5, 10% MeOH). The 5% MeOH fraction (41.1 mg) was collected and further purified by Sephadex LH-20 column chromatography (1:1 CHCl<sub>3</sub>/CH<sub>3</sub>OH) to obtain a fraction containing **12** (28.8 mg). The sample was further purified by preparative reversed-phase HPLC using a CAPCELL PAK MG-II C<sub>18</sub> column (5.0 μm, 20 i.d. × 150 mm; Shiseido), and elution with 60% aqueous acetonitrile supplemented with 0.1% formic acid yielded 8.3 mg of **12**.

**Isolation of 13 from SUKA33::rapH/pKU503rap\_M2-M12-M3.** Twelve point five litres of fermentation broth of SUKA33::rapH/pKU503rap\_M2-M12-M3, was centrifuged to obtain a mycelial cake, which was extracted twice with 500 ml of acetone. The acetone was removed in vacuo, and the residual aqueous layer was extracted twice with EtOAc. The resultant EtOAc layer was concentrated in vacuo to afford 7.82 g of crude extract. The crude extract was subjected to silica gel MPLC (SNAP Ultra 25 g, Biotage) eluted with a gradient of *n*-hexane–EtOAc (0–25% EtOAc) followed by a stepwise gradient of CHCl<sub>3</sub>–MeOH (0, 1, 3, 5, 10% MeOH). The 5% MeOH fraction (2.35 g) was collected and subjected to silica gel MPLC (SNAP Ultra 25 g) with stepwise gradient of CHCl<sub>3</sub>–MeOH (0, 2, 3, 5% MeOH). The 3% MeOH fraction (638 mg) was subjected to Sephadex LH-20 column chromatography (1:1 CHCl<sub>3</sub>/CH<sub>3</sub>OH) and afforded a fraction containing **13** (66.5 mg). This sample was further purified by preparative reversed-phase HPLC using a CAPCELL PAK MG-II C<sub>18</sub> column (5.0 μm, 20 i.d. × 150 mm; Shiseido), and elution with 80% aqueous acetonitrile supplemented with 0.1% formic acid yielded 5.0 mg of **13**.

**Isolation of 14 from SUKA54/pKU503rap\_M7ER<sup>0</sup>.** pKU592AT-2cis-*rapH* was introduced into SUKA54/pKU503rap\_M7ER<sup>0</sup> by intergeneric conjugation and integration to afford SUKA54::rapH/pKU503rap\_M7ER<sup>0</sup>. Six litres of fermentation broth of SUKA54::rapH/pKU503rap\_M7ER<sup>0</sup> was centrifuged to obtain a mycelial cake, which was extracted twice with 500 ml of acetone. The acetone was removed in vacuo, and the residual aqueous layer was extracted twice with EtOAc. The resultant EtOAc layer was concentrated in vacuo to afford 2.92 g of crude extract. The crude extract was subjected to silica gel MPLC (SNAP Ultra 25 g, Biotage) eluted with a gradient system of *n*-hexane–EtOAc (0–25% EtOAc) followed by a stepwise gradient of CHCl<sub>3</sub>–MeOH (0, 1, 5, 10, 20, 50 and 90% MeOH). The 5% MeOH fraction (561 mg) was collected and subjected to ODS MPLC (Purif-Pack ODS-100) by using a H<sub>2</sub>O–MeOH stepwise gradient (20, 40, 60, 80, and 100% MeOH). Fractions containing **14** were collected (227 mg). Then, 104.7 mg of this sample was further purified by preparative reversed-phase HPLC using a CAPCELL PAK MG-II C<sub>18</sub> column (5.0 μm, 20 i.d. × 150 mm; Shiseido), and elution with 75% aqueous acetonitrile yielded 7 mg of **14**.

**Isolation of 15 from SUKA34::rapH/pKU503rapΔM9DH-KR::M7DH-ER-KR.** Two point eight litres of fermentation broth of SUKA34::rapH/pKU503rapΔM9DH-

KR::M7DH-ER-KR, was centrifuged to obtain a mycelial cake, which was extracted twice with 500 ml of acetone. The acetone was removed in vacuo, and the residual aqueous layer was extracted twice with EtOAc. The resultant EtOAc layer was concentrated in vacuo to afford 1.44 g of crude extract. The crude extract was subjected to silica gel MPLC (Purif-pack SI-60, 27 g; Shoko Scientific) eluted with a gradient of *n*-hexane–EtOAc (0–25% EtOAc) followed by a stepwise gradient of CHCl<sub>3</sub>–MeOH (0, 1, 3, 5, 10% MeOH). The 3% MeOH fraction (194 mg) was collected and subjected to silica gel MPLC (SNAP Ultra 25 g) with isocratic elution with 3% MeOH in CHCl<sub>3</sub>. The fractions were monitored by UPLC analysis, and the fractions containing **15** were collected (110 mg). The sample was subjected to Sephadex LH-20 column chromatography (1:1 CHCl<sub>3</sub>/CH<sub>3</sub>OH) and afforded a fraction containing **15** (18.3 mg). This sample was further purified by preparative reversed-phase HPLC using a CAPCELL PAK MG-II C<sub>18</sub> column (5.0 μm, 20 i.d. × 150 mm; Shiseido), and elution with 85% aqueous acetonitrile yielded 3.2 mg of **15**.

**Isolation of 16 from SUKA34::rapH/pKU503rapΔM3DH-ER-KR::M7DH-ER-KR.** Seven litres of fermentation broth of SUKA34::rapH/pKU503rapΔM3DH-ER-KR::M7DH-ER-KR, was centrifuged to obtain a mycelial cake, which was extracted twice with 500 ml of acetone. The acetone was removed in vacuo, and the residual aqueous layer was extracted twice with EtOAc. The resultant EtOAc layer was concentrated in vacuo to afford 3.83 g of crude extract. The crude extract was subjected to silica gel MPLC (SNAP Ultra 25 g, Biotage) eluted with a gradient system of *n*-hexane–EtOAc (0–25% EtOAc) followed by a stepwise gradient of CHCl<sub>3</sub>–MeOH (0, 1, 3, 5, 10% MeOH). The 3% MeOH fraction (144 mg) was collected and subjected to Sephadex LH-20 column chromatography (1:1 CHCl<sub>3</sub>/CH<sub>3</sub>OH) and afforded a fraction containing **16** (55.7 mg). This sample was further purified by preparative reversed-phase HPLC using a CAPCELL PAK MG-II C<sub>18</sub> column (5.0 μm, 20 i.d. × 150 mm; Shiseido), and elution with 80% aqueous acetonitrile supplemented with 0.1% formic acid yielded 10.6 mg of **16**.

**Isolation of 17 from SUKA34::rapH/pKU503rapΔM11DHL-KR::bfmM2KR.** Six litres of fermentation broth of SUKA34::rapH/pKU503rapΔM11DHL-KR::bfmM2KR was centrifuged to obtain a mycelial cake, which was extracted twice with 500 ml of acetone. The acetone was removed in vacuo, and the residual aqueous layer was extracted twice with EtOAc. The resultant EtOAc layer was concentrated in vacuo to afford 2.75 g of crude extract. The crude extract was subjected to silica gel MPLC (SNAP Ultra 25 g, Biotage) eluted with a gradient of *n*-hexane–EtOAc (0–25% EtOAc) followed by a stepwise gradient of CHCl<sub>3</sub>–MeOH (0, 1, 3, 5, 10, 50 and 90% MeOH). The 5% MeOH fraction (228 mg) was collected and then subjected to Sephadex LH-20 column chromatography (1:1 CHCl<sub>3</sub>/CH<sub>3</sub>OH) and afforded a fraction containing **17** (33.2 mg). This sample was further purified by preparative reversed-phase HPLC using a CAPCELL PAK MG-II C<sub>18</sub> column (5.0 μm, 20 i.d. × 150 mm; Shiseido), and elution with 60% aqueous acetonitrile yielded 2.2 mg of **17**.

**Isolation of (16S)-16-O-demethyl rapamycin (20) from SUKA34::rapH/pKU503rapΔrapM.** One litre of fermentation broth of SUKA34::rapH/pKU503rapΔrapM was centrifuged to obtain a mycelial cake, which was extracted twice with 500 ml of acetone. The acetone was removed in vacuo, and the residual aqueous layer was extracted with EtOAc twice. The resultant EtOAc layer was concentrated in vacuo to afford 0.74 g of crude solid. The crude extract was subjected to silica gel MPLC (SNAP Ultra 25 g, Biotage) eluted with a gradient of *n*-hexane–EtOAc (0–25% EtOAc) followed by a stepwise gradient of CHCl<sub>3</sub>–MeOH (0, 1, 3, 5, 10, 50 and 90% MeOH). The 5% MeOH fraction (59.8 mg) was collected and then subjected to Sephadex LH-20 column chromatography (1:1 CHCl<sub>3</sub>/CH<sub>3</sub>OH) and afforded a fraction containing **20** (6.4 mg). This sample was further purified by preparative reversed-phase HPLC using a CAPCELL PAK MG-II C<sub>18</sub> column (5.0 μm, 20 i.d. × 150 mm; Shiseido), and elution with 78% aqueous MeOH yielded 2.9 mg of **20**.

**Binding activities of rapamycin derivatives to FKBP and mTOR.** HeLa cells were maintained in DMEM (Fujifilm Wako, Osaka, Japan) supplemented with 10% FBS (Thermo Fisher Scientific, MA, USA) and 1% of Penicillin-Streptomycin (Thermo Fisher Scientific) in a humidified incubator with 5% CO<sub>2</sub> at 37 °C. Harvested cells were plated at 1 × 10<sup>6</sup> cells/well in 6-well plate and incubated for 6 h. Then each 1.25 μg of pAG-mTOR and pAsh-FKBP12 (Medical and Biological Laboratories Co., Ltd., Aichi, Japan) was co-transfected with Lipofectamine 2000 reagent (Thermo Fisher Scientific) overnight. Next day, transfected cells were harvested and seeded in 384-well optical-bottom microplate (PerkinElmer, MA, USA) at a density of 3000 cells/well. After the 24 h incubation at 37 °C, cells were treated with different concentrations of compounds for 1 h. Cells were then fixed with 4% paraformaldehyde and stained with 1 μg/ml of Hoechst33342 (Thermo Fisher Scientific). Cell images of 9 visual fields/well were acquired with ×20 objective lens in Opera Phenix High-Content Screening System (PerkinElmer) and analysed by Harmony4.9 software (PerkinElmer). Because protein-protein interactions are quantified by measuring the aggregation of fluorescent elements in foci, fluorescent intensity in foci was normalised by total fluorescent intensity in each cell. EC<sub>50</sub> values were calculated using TIBCO Spotfire 10.3 software (PerkinElmer).

**Cytotoxicity assay.** The cytotoxic activities of isolated rapamycin derivatives against T lymphoma Jurkat cells, human ovarian adenocarcinoma SKOV-3 cells

and malignant pleural mesothelioma MESO-1 cells were examined. Jurkat cells were cultured in RPMI1640 medium supplemented with 10% foetal bovine serum, penicillin (50 U/ml), streptomycin (50 µg/ml), and Glutamax. SKOV-3 cells were cultured in DMEM medium supplemented with 10% fetal bovine serum, penicillin (50 U/ml), and streptomycin (50 µg/ml). MESO-1 cells were cultured in RPMI1640 medium supplemented with 10% fetal bovine serum, penicillin (50 U/ml), and streptomycin (50 µg/ml). All cell lines were seeded in a 384-well plate at a density of 1000 cells/well in 20 µl of media and incubated at 37 °C in a humidified incubator with 5% CO<sub>2</sub>. After 4 h, 2-fold dilution samples resolved in DMSO were added to the cell culture at the concentration of 0.5% (0.1 µl) and incubated for 72 h. Cell viabilities were measured using a CellTiter-Glo luminescent cell viability assay and EnVision multilabel plate reader.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

Sequence data that support the findings of this study have been deposited in DDBJ with the accession code LC566301. Source data are provided with this paper. The other data that support the findings of this study are included in this paper. Raw data for spectroscopic analyses are available from the corresponding author upon reasonable request.

Received: 7 November 2019; Accepted: 10 July 2020;

Published online: 11 August 2020

## References

- Giordanetto, F. & Kihlberg, J. Macrocyclic drugs and clinical candidates: what can medicinal chemists learn from their properties? *J. Med. Chem.* **57**, 278–295 (2014).
- Bogdan, A. R., Davies, N. L. & James, K. Comparison of diffusion coefficients for matched pairs of macrocyclic and linear molecules over a drug-like molecular weight range. *Org. Biomol. Chem.* **9**, 7727–7733 (2011).
- Mallinson, J. & Collins, I. Macrocycles in new drug discovery. *Future Med. Chem.* **4**, 1409–1438 (2012).
- Driggers, E. M., Hale, S. P., Lee, J. & Terrett, N. K. The exploration of macrocycles for drug discovery—an underexploited structural class. *Nat. Rev. Drug Discov.* **7**, 608–624 (2008).
- Hann, M. M. & Keseru, G. M. Finding the sweet spot: the role of nature and nurture in medicinal chemistry. *Nat. Rev. Drug Discov.* **11**, 355–365 (2012).
- Bauer, A. & Bronstrup, M. Industrial natural product chemistry for drug discovery and development. *Nat. Prod. Rep.* **31**, 35–60 (2014).
- Medema, M. H., Cimermancic, P., Sali, A., Takano, E. & Fischbach, M. A. A systematic computational analysis of biosynthetic gene cluster evolution: lessons for engineering biosynthesis. *PLoS Comput. Biol.* **10**, e1004016 (2014).
- Zhang, L. et al. Characterization of giant modular PKSs provides insight into genetic mechanism for structural diversification of aminopolyl polyketides. *Angew. Chem. Int. Ed. Engl.* **56**, 1740–1745 (2017).
- Jenke-Kodama, H. & Dittmann, E. Evolution of metabolic diversity: insights from microbial polyketide synthases. *Phytochemistry* **70**, 1858–1866 (2009).
- Wlodek, A. et al. Diversity oriented biosynthesis via accelerated evolution of modular gene clusters. *Nat. Commun.* **8**, 1206 (2017).
- McDaniel, R. et al. Multiple genetic modifications of the erythromycin polyketide synthase to produce a library of novel “unnatural” natural products. *Proc. Natl Acad. Sci. USA* **96**, 1846–1851 (1999).
- Sugimoto, Y., Ding, L., Ishida, K. & Hertweck, C. Rational design of modular polyketide synthases: morphing the aureothin pathway into a luteoreticulin assembly line. *Angew. Chem. Int. Ed. Engl.* **53**, 1560–1564 (2014).
- Kushnir, S. et al. Minimally invasive mutagenesis gives rise to a biosynthetic polyketide library. *Angew. Chem. Int. Ed. Engl.* **51**, 10664–10669 (2012).
- Hashimoto, T. et al. Biosynthesis of quinolidomycin, the largest known macrolide of terrestrial origin: identification and heterologous expression of a biosynthetic gene cluster over 200 kb. *Org. Lett.* **20**, 7996–7999 (2018).
- Jinek, M. et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012).
- Gibson, D. G. et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).
- Komatsu, M. et al. Engineered *Streptomyces avermitilis* host for heterologous expression of biosynthetic gene cluster for secondary metabolites. *ACS Synth. Biol.* **2**, 384–396 (2013).
- Guduru, S. K. R. & Arya, P. Synthesis and biological evaluation of rapamycin-derived, next generation small molecules. *Medchemcomm* **9**, 27–43 (2018).
- Schwecke, T. et al. The biosynthetic gene cluster for the polyketide immunosuppressant rapamycin. *Proc. Natl Acad. Sci. USA* **92**, 7839–7843 (1995).
- Kuscer, E. et al. Roles of rapH and rapG in positive regulation of rapamycin biosynthesis in *Streptomyces hygroscopicus*. *J. Bacteriol.* **189**, 4756–4763 (2007).
- Choi, J., Chen, J., Schreiber, S. L. & Clardy, J. Structure of the FKBP12-rapamycin complex interacting with the binding domain of human FRAP. *Science* **273**, 239–242 (1996).
- Caufield, C. E., Musser, J. H. & Rinker, J. M. Hydrogenated rapamycin derivatives. United States patent US5023262 (1990).
- Kim, J. H., Komatsu, M., Shin-Ya, K., Omura, S. & Ikeda, H. Distribution and functional analysis of the phosphopantetheinyl transferase superfamily in Actinomycetales microorganisms. *Proc. Natl Acad. Sci. USA* **115**, 6828–6833 (2018).
- Kellenberger, L. et al. A polylinker approach to reductive loop swaps in modular polyketide synthases. *ChemBioChem* **9**, 2740–2749 (2008).
- Donadio, S., McAlpine, J. B., Sheldon, P. J., Jackson, M. & Katz, L. An erythromycin analog produced by reprogramming of polyketide synthesis. *Proc. Natl Acad. Sci. USA* **90**, 7119–7123 (1993).
- Zheng, J., Gay, D. C., Demeler, B., White, M. A. & Keatinge-Clay, A. T. Divergence of multimodular polyketide synthases revealed by a didomain structure. *Nat. Chem. Biol.* **8**, 615–621 (2012).
- Strom, T. et al. Structural identification of SAR-943 metabolites generated by human liver microsomes in vitro using mass spectrometry in combination with analysis of fragmentation patterns. *J. Mass Spectrom.* **46**, 615–624 (2011).
- Artzi, N. et al. Sustained efficacy and arterial drug retention by a fast drug eluting cross-linked fatty acid coronary stent coating. *Ann. Biomed. Eng.* **44**, 276–286 (2016).
- Keatinge-Clay, A. T. Stereocontrol within polyketide assembly lines. *Nat. Prod. Rep.* **33**, 141–149 (2016).
- Luengo, J. I., Konialian-Beck, A., Rozamus, L. W. & Holt, D. A. Manipulation of the rapamycin effector domain. selective nucleophilic substitution of the C7 methoxy group. *J. Org. Chem.* **59**, 6512–6513 (1994).
- Law, B. J. C., Struck, A. W., Bennett, M. R., Wilkinson, B. & Micklefield, J. Site-specific bioalkylation of rapamycin by the RapM 16-O-methyltransferase. *Chem. Sci.* **6**, 2885–2892 (2015).
- Watanabe, T. et al. Genetic visualization of protein interactions harnessing liquid phase transitions. *Sci. Rep.* **7**, 46380 (2017).
- Li, J., Kim, S. G. & Blenis, J. Rapamycin: one drug, many effects. *Cell Metab.* **19**, 373–379 (2014).
- Cottens, S. & Sedrani, R. Rapamycin derivatives. WO 96/41807 (1996).
- Liu, Y. et al. In vitro CRISPR/Cas9 system for efficient targeted DNA editing. *mBio* **6**, e01714–e01715 (2015).
- Jiang, W. et al. Cas9-Assisted Targeting of Chromosome segments CATCH enables one-step targeted cloning of large gene clusters. *Nat. Commun.* **6**, 8101 (2015).
- Alberti, F. & Corre, C. Editing streptomycete genomes in the CRISPR/Cas9 age. *Nat. Prod. Rep.* **36**, 1237–1248 (2019).
- Zhao, Y., Li, G., Chen, Y. & Lu, Y. Challenges and advances in genome editing technologies in *Streptomyces*. *Biomolecules* **10**, 734 (2020).
- Kapur, S., Chen, A. Y., Cane, D. E. & Khosla, K. Molecular recognition between ketosynthase and acyl carrier protein domains of the 6-deoxyerythronolide B synthase. *Proc. Natl Acad. Sci. USA* **107**, 22066–22071 (2010).
- Swinney, D. C. & Anthony, J. How were new medicines discovered? *Nat. Rev. Drug Discov.* **10**, 507–519 (2011).
- Schreiber, K. H. et al. A novel rapamycin analog is highly selective for mTORC1 in vivo. *Nat. Commun.* **10**, 3194 (2019).
- Komatsu, M., Uchiyama, T., Omura, S., Cane, D. E. & Ikeda, H. Genome-minimized *Streptomyces* host for the heterologous expression of secondary metabolism. *Proc. Natl Acad. Sci. USA* **107**, 2646–2651 (2010).
- Chu, G., Vollrath, D. & Davis, R. W. Separation of large DNA molecules by contour-clamped homogeneous electric fields. *Science* **234**, 1582–1585 (1986).
- Ikeda, H., Kotaki, H. & Omura, S. Genetic studies of avermectin biosynthesis in *Streptomyces avermitilis*. *J. Bacteriol.* **169**, 5615–5621 (1987).
- Gregory, M. A. et al. Rapamycin biosynthesis: elucidation of gene product function. *Org. Biomol. Chem.* **4**, 3565–3568 (2006).

## Acknowledgements

This work was supported by Japan Agency for Medical Research and Development (AMED) under Grant Number JP19ae0101045 for K.S.

## Author contributions

Conceptualisation, K.K., T.H. and K.S.; data curation, K.K. and T.H.; funding acquisition, K.S.; investigation, K.K., T.H., J.H., I.K., N.K., R.U. and T.N.; methodology, K.K., T.H., M. K., H.I. and K.S.; project administration, K.S.; resources, H.I. and K.S.; supervision, H.I. and K.S.; validation, K.K., T.H., H.S. and K.S.; visualisation, K.K. and T.H.; Writing—original draft, K.K., T.H. and K.S.; Writing—review and editing, K.K., T.H., H.S. and K.S.

**Competing interests**

The authors declare no competing interests.

**Additional information**

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41467-020-17769-2>.

**Correspondence** and requests for materials should be addressed to H.I. or K.S.-y.

**Peer review information** *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020