This is a repository copy of Application of serine integrases for secondary metabolite pathway assembly in Streptomyces.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/162639/

Version: Published Version

Article:
Gao, Hong, Taylor, Gabrielle, Evans, Stephanie K. et al. (2 more authors) (2020)
Application of serine integrases for secondary metabolite pathway assembly in Streptomyces. Synthetic and Systems Biotechnology. pp. 111-119. ISSN 2405-805X

https://doi.org/10.1016/j.synbio.2020.05.006

Reuse
This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can’t change the article in any way or use it commercially. More information and the full terms of the licence here: https://creativecommons.org/licenses/

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.
Application of serine integrases for secondary metabolite pathway assembly in *Streptomyces*

Hong Gao*,1, Gabrielle Taylor, Stephanie K. Evans, Paul C.M. Fogg, Margaret C.M. Smith**

*Department of Biology, University of York, York, YO10 5DD, United Kingdom

**Corresponding author.

1 Corresponding author.

E-mail addresses: h.gao@tees.ac.uk (H. Gao), Maggie.smith@york.ac.uk (M.C.M. Smith).

† Present address: School of Health & Life Science, Teesside University, Middlesbrough, TS1 3BA; National Horizons Centre, Darlington campus, Teesside University, Darlington, DL1 1HG.

https://doi.org/10.1016/j.synbio.2020.05.006

Received 18 March 2020; Received in revised form 27 May 2020; Accepted 28 May 2020

© 2020 The Author(s). Production and hosting by Elsevier B.V. on behalf of KeAi Communications Co., Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/BY-NC-ND/4.0/).

**ARTICLE INFO**

Serine integrases have been shown to be efficient tools for metabolic pathway assembly. To further improve the flexibility and efficiency of pathway engineering via serine integrases, we explored how multiple orthogonally active serine integrases can be applied for use in vitro for the heterologous expression of complex biosynthesis pathways in *Streptomyces* spp., the major producers of useful bioactive natural products. The results show that multiple orthogonal serine integrases efficiently assemble the genes from a complex biosynthesis pathway in a single in vitro recombination reaction, potentially permitting a versatile combinatorial assembly approach. Furthermore, the assembly strategy also permitted the incorporation of a well-characterised promoter upstream of each gene for expression in a heterologous host. The results demonstrate how site-specific recombination based on orthogonal serine integrases can be applied in *Streptomyces* spp.

**ABSTRACT**

Serine integrases have been shown to be efficient tools for metabolic pathway assembly. To further improve the flexibility and efficiency of pathway engineering via serine integrases, we explored how multiple orthogonally active serine integrases can be applied for use in vitro for the heterologous expression of complex biosynthesis pathways in *Streptomyces* spp., the major producers of useful bioactive natural products. The results show that multiple orthogonal serine integrases efficiently assemble the genes from a complex biosynthesis pathway in a single in vitro recombination reaction, potentially permitting a versatile combinatorial assembly approach. Furthermore, the assembly strategy also permitted the incorporation of a well-characterised promoter upstream of each gene for expression in a heterologous host. The results demonstrate how site-specific recombination based on orthogonal serine integrases can be applied in *Streptomyces* spp.

1. Introduction

Phage-encoded serine integrases catalyse site-specific integration of DNA into bacterial host chromosomes in a highly controllable and predictable way [1], making these proteins powerful tools for molecular genetics. Since the first serine integrases were discovered and described in the 1990s, integrases such as those from *Streptomyces* phage φC31 [2] and φBT1, and mycobacteriophage Bxb1, have been used as genome integration vectors in bacteria and in other organisms, including humans, mice and yeasts [3–6]. Moreover, the pool of available integrase proteins is still expanding, driven by the recent upsurge in synthetic biology applications of this protein family [7], such as in genetic memory devices [8].

Serine integrases bind to specific sites in the DNA that are 40–60 bp in length, bring these sites together, cut them, and then rejoin the sites to produce a recombinant product. Integration occurs between the so-called attachment sites derived from the phage (attP) and bacterial chromosome (attB) to yield the recombinant sites (attL and attR). The mechanism has been described in detail in previous publications [9–11]. During recombination the DNA is cut in both strands at specific locations in both attP and attB to yield 2 bp overhangs; as this 2 bp overhang is identical in both attP and attB, complementarity is maintained in the recombinants and the DNA backbone can be rejoined. Studies have shown that this 2 bp sequence can be changed to other sequences and recombination efficiency is unaffected, but only if both attP and attB have the same 2-bp sequence [12,13]. This central 2-bp site specificity has permitted the use of a single integrase to assemble DNA fragments together in a predictable order in a single recombination reaction [11,14]. This is evident in the study of Colloms et al., who developed the serine integrase recombinational assembly (SIRA) method that used φC31 to assemble up to 5 genes into a functional pigment biosynthetic pathway; the efficiency dropped from 87% for the 3-gene assembly to 18% for the 5-gene assembly [11]. This is because integrase attempts and fails to complete recombination between sites with non-identical crossover sites but recombination only proceeds to completion in reactions where the central 2-bp site is identical in both attP and attB.

As different integrases only recombine their cognate recombination sites, their activities are expected to be entirely separate and independent of each other, that is, an integrase should not recognise or recombine the substrate sites of the other integrases. This orthogonalility permits the use of different integrases in the same recombination reaction, yielding entirely predictable recombinants depending on the location of recombination sites. Moreover, there should not be any loss
of efficiency of recombination, provided the integrases are sufficiently active in the recombinant buffer used. Some integrases have already been used in pairs (e.g., ϕC31 and Bxb1 integrase [15]) to investigate orthogonality, but information on the expected efficiency of using more than two integrases is not yet available.

Streptomyces is the primary producer of natural products, which have played important roles in healthcare all over the world. Biosynthetic gene clusters (BGCs) of natural products are large and complex (even more than 100 kb) with multiple layers of regulation. Synthetic biology offers a way of exploring BGCs for which the product has not yet been characterised, that is, cloning the genes of the BGC into vectors that integrate into a heterologous Streptomyces host such as S. coelicolor or S. lividans and under the control of well-characterised promoters. These tractable Streptomyces hosts have worked well in the biosynthesis of the complex natural products. Cloning each gene downstream of a promoter is a laborious and time-consuming process, especially when the BGC contains more than 20 genes. We propose a new strategy in which the genes are amplified using primers incorporating a promoter and assembled by in vitro fragment assembly using orthogonal integrases.

Previously our lab has used the erythromycin biosynthesis pathway as a model system for the expression of BGCs in a heterologous host [16]. The three polyketide synthase (PKS) genes eryAII, eryAll and eryAIII in the erythromycin BGC were cloned into three orthologous integrating plasmids, which were based on the int/attP loci from phages TG1, SV1 and φBT1, respectively. Following integration, 6-deoxyerythronolide B (6-DEB), the first intermediate produced by the three PKS enzymes, could be detected in the fermentation broth. The results demonstrated sequential integration of multiple orthogonal integrating vectors is a reliable method to clone large genes required to synthesise many natural products.

Here we set out to demonstrate that multiplexed integration could be an option to improve the efficiency of recombination in the assembly of metabolic pathways. In this paper, results are presented for the use of orthogonal serine integrases for pathway level gene assembly. The results demonstrate improved strategies for the manipulation of BGCs in synthetic biology and for natural product discovery.

### 2. Materials and Methods

#### 2.1. Bacterial strains and culture conditions

*E. coli* strain Top10 (F− mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 lacX74 proA1 araD139 Δ(ara-leu)7697 galE15 galK16 rpsL150 lacIq1288Δ (F− lacY1) [M15] M1152 [15]) was used for plasmid propagation and subcloning. *E. coli* strain ET12567(pUZ8002) was used as the donor host in plasmid conjugations from *E. coli* to Streptomyces as described previously [17]. The *E. coli* strains were grown in Luria-Bertani broth (LB) or on LB agar at 37 °C.

*Streptomyces coelicolor* M1152 [Δact 36d Δpck Δapd (CB12987)] [18] and *Streptomyces lividans* TK24 (str-6 SLP2-6LP3) [19] were used as the recipients in conjugation. The strains were maintained on Soya Mannitol (SM) agar at 30 °C. Conjugations were performed on SM containing 10 mM MgCl₂ and tryptic soy broth (TSB) medium was used for the preparation of genomic DNA [19]. Antibiotic concentrations for *E. coli* were 150 μg/ml hygromycin, 50 μg/ml apramycin, 50 μg/ml kanamycin and 100 μg/ml ampicillin; and for *Streptomyces* were 50 μg/ml hygromycin, 50 μg/ml apramycin, 50 μg/ml kanamycin, 20 μg/ml erythromycin (in DMSO, 120 μg/ml erythromycin for *S. lividans* TK24) and 25 μg/ml nalidixic acid for selection.

#### 2.2. DNA manipulation

*E. coli* transformation and gel electrophoresis were carried out as described previously [20]. Plasmid DNA extraction from *E. coli* was performed using a QIAprep spin miniprep kit (Qiagen, Germany) according to the manufacturer’s protocol. Restriction enzymes were purchased from New England BioLabs (NEB, USA), and were used according to the manufacturer’s instructions. In-Fusion® cloning (Clontech®, USA) and TA cloning (CloneJET PCR Cloning Kit, Thermo Scientific, USA) were used for joining DNA fragments. Genomic DNA preparation from *Streptomyces* was performed following the salting out procedure in the *Streptomyces* manual [19]. Polymerase Chain Reaction (PCR) was carried out using Phusion® High-Fidelity DNA Polymerase (NEB, USA) according to the manufacturer’s instructions. The plasmids
Table 2

Oligonucleotides used in this study.

| Oligonucleotide | Sequence (5′-3′) |
|-----------------|-----------------|
| pHG1A-for       | GAAACGGCAGCTGATGATATTTAGGAGGATGTAATGAGACAGGCTGGTCCCG |
| pHG1A-rev       | GTGTTGGGGGCGTACGAGCTCTTCTTGATCTCATCCGTCG |
| pHG1-for        | ACTAGAGGAGCTAGGCTCTTCAAGTTGAGAGGATGATGGAG |
| pHG1-rev        | CAGGTTGAGAGCTCTTCTTGATCTCATCCGTCG |
| pHG3-for        | CCGCGGCGGACCCCTCCAGGATCTGCAGGTAATGAGACAGGCTGGC |
| pHG3-rev        | TCTTGGTATGACCTACACTTGACCAGGGAAGCCGCGAAT |
| pHG9A-for       | GCCCGCGCGGACCCCTCCAGGATCTGCAGGTAATGAGACAGGCTGGC |
| pHG9A-rev       | GCGGACCCCTCCAGGATCTGCAGGTAATGAGACAGGCTGGC |
| pHG2B0-for      | GGGGTTAGAGCTACCAAGCTCATGATGAGGAGGATGATGGAG |
| pHG2B0-rev      | GTATTACCTATCATGATGAGGAGGATGATGGAG |
| pHG2B2-for      | GTATTACCTATCATGATGAGGAGGATGATGGAG |
| pHG2B2-rev      | GTATTACCTATCATGATGAGGAGGATGATGGAG |
| pHG2B3-for      | GTATTACCTATCATGATGAGGAGGATGATGGAG |
| pHG2B3-rev      | GTATTACCTATCATGATGAGGAGGATGATGGAG |
| pHG2B45-for     | GTATTACCTATCATGATGAGGAGGATGATGGAG |
| pHG2B45-rev     | GTATTACCTATCATGATGAGGAGGATGATGGAG |
| pHG3C6-for      | GCCCCGGGGCATCCATGCATTCAACCCCTCCAGGCAAA |
| pHG3C6-rev      | GCCCCGGGGCATCCATGCATTCAACCCCTCCAGGCAAA |
| pHG3C45-for     | AAAATTACAAAGTTTTCAACCCTTGATTTGAATTAGCGGTCAAATAATTTGTAATTCGTTT |
| pHG3C45-rev     | AAAATTACAAAGTTTTCAACCCTTGATTTGAATTAGCGGTCAAATAATTTGTAATTCGTTT |
| pHG3C23-for     | GTGTTTATGACCTAGGAGGATGATGGAG |
| pHG3C23-rev     | GTGTTTATGACCTAGGAGGATGATGGAG |
| pHG3C1-for      | ACCGGAACGCGGTCTACAACTTTGGCCACAGCCAGGGAGGAGGAGGTTGGCGC |
| pHG3C1-rev      | ACCGGAACGCGGTCTACAACTTTGGCCACAGCCAGGGAGGAGGAGGTTGGCGC |
| pHG9A          | GCCCCGGGGCATCCATGCATTCAACCCCTCCAGGCAAA |
| pHG9A-for       | GCCCCGGGGCATCCATGCATTCAACCCCTCCAGGCAAA |
| pHG9A-rev       | GCCCCGGGGCATCCATGCATTCAACCCCTCCAGGCAAA |
| pHG3-for        | CTGTTTATGACCTAGGAGGATGATGGAG |
| pHG3-rev        | CTGTTTATGACCTAGGAGGATGATGGAG |
| pHG2R2-for      | GCCCCGGGGCATCCATGCATTCAACCCCTCCAGGCAAA |
| pHG2R2-rev      | GCCCCGGGGCATCCATGCATTCAACCCCTCCAGGCAAA |

Restriction enzyme recognition sites were underlined. The attB/attP sites were indicated in bold letters.

used in this study are listed in Table 1 and the primers used are listed in Table 2.

2.3. Plasmids

2.3.1. Construction of eryF expression plasmid pHG1

The eryF gene was amplified from Saccharopolyspora erythraea BIOT-0666 genomic DNA using the primer pair pHG1A-for/pHG1A-rev, and inserted into In-Fusion cloning into pBFB20 [16] cut with NheI and PacI to form the plasmid pHG1A. The 3785 bp fragment containing the ϕC31 int/attP and hygromycin resistance gene was amplified from plasmid pH2F27C [16], using the primer pair pHG1-for and pHG1-rev. Plasmid pHG1A was digested with XbaI and NheI, and the 5608 bp fragment was ligated with the 3785 bp PCR fragment from pH2F27C by In-Fusion cloning to give the plasmid pHG1 (Fig. 1).

2.3.2. Backbone vectors for DNA assembly

The fragment containing TGI attP-lacZa-SPBc attP was amplified from the plasmid pLT5-lacZa (Table 1) using the primer pair pHG2B0-for and pHG2B0-rev, then cloned into pH1 cut with BglII by In-Fusion to form the plasmid pHG2B0 (Fig. 1). The plasmid pHG2B0 contains the eryF gene, and the lacZa reporter gene flanked with the TGI attP [21] and SPBc attP [22] sequences. The plasmid pHG2B0 was used as the vector backbone for the assembly of the eryB genes.

One of the correctly assembled plasmids was named pHG2R2 (Fig. 1) and this plasmid was used to construct the backbone vector, pHG9A (Fig. 1), for receipt of the assembled eryC genes. To construct pHG9A, the lacZa reporter gene flanked with the TGI attP and SPBc attP sequences was inserted between the hygromycin resistance gene and the ϕC31 attP/int locus in pHG2R2, allowing the attP sites to be again used to integrate the DNA fragment assembly, in this case the eryC genes. Plasmid pHG9A was therefore constructed as follows: First, to insert restriction enzyme sites AsISI and Nsil, the primer pair pHG3C-for and pHG3C-rev was used to amplify a fragment of approximate 1 kb size from pHG2R2, which was then cloned back into pHG2R2 digested by BbcII (two sites, located at 7385 and 8395 bp separately) to give plasmid pHG3F (Fig. S1). Next, the fragment TGI attP-lacZa-SPBc attP was amplified from the plasmid pHG9A using the primer pair pHG9A-for and pHG9A-rev, then cloned into pHG3F cut with AsISI and Nsil by In-Fusion to form the plasmid pHG1 (Fig. 1).

2.4. DNA assembly

2.4.1. Assembly of eryB genes

All the eryB genes were amplified from S. erythraea BIOT-0666 genomic DNA using the primer pairs listed in Table 2 (pHG2B2-for/rev, pHG2B3-for/rev, pHG2B45-for/rev, pHG2B6-for/rev and pHG2B7-for/rev). As eryBIV and eryBV are adjacent genes in the native erythromycin biosynthesis gene cluster S. erythraea BIOT-0666 and apparently translationally coupled, they were amplified together so that their
coupled status was maintained.

Firstly, an actIp promoter fragment was amplified using the primer pair pHG2A-for/pHG2A-rev (Table 2), then cloned into pJET1.2/blunt, to generate pHG2A (Fig. S1). Then, the five gene fragments: eryBII, eryBIII, the fragment containing eryBIV and eryBV, eryBVI, and eryBVI, were cloned individually into pHG2A downstream of the actIp promoter (Fig. 2), to give plasmids pHG2B2, pHG2B3, pHG2B45, pHG2B6 and pHG2B7, respectively. Gene cassettes, each encoding one of the eryB genes or the eryB operon and the upstream actIp promoter were amplified from plasmids pHG2B2 - pHG2B7 using the primer pairs listed in Table 2 (TG1 attB F, Bxb1 attB R, Bxb1 attP F, Int9 attB R, Int9 attP F, Int4 attB R, Int4 attP F, Int7 attB R, Int7 attP F, and SPBc attB R). Each primer contains the sequence of an attP or attB site and, at the 3’ end, a sequence that binds to vector derived DNA. The amplified gene cassettes can thus be tagged with different attP/attB sites. In vitro recombination between all matching attP-attB pairs using purified integrases joins the cassettes together end-to-end in a specific order and inserts the resulting linear assembly into the pHG2B0 vector, replacing the lacZa. This system is extremely versatile as the use of the vector binding sites in the primer sequences allows the researcher to generate different orders of assembled genes and thus optimise expression of the gene assembly. The strategy to amplify the DNA fragments to be used for assembly is shown in Fig. 2A.

The assembly strategy for the assembly of eryB genes is shown in Fig. 2B. In this assembly, six integrases were used: TG1 [23], Bxb1 [24], Int9, Int4, Int7 [8] and SPBc [22], each recognising only its cognate attP and attB sites located in the amplified gene cassettes.

2.4.2. Assembly of eryC genes

The eryC genes were assembled using a similar strategy to that used for the eryB genes. The eryC genes were amplified from S. erythraea BIOT-0666 genomic DNA using the primer pairs listed in Table 1 (pHG3C1-for/rev, pHG3C23-for/rev, pHG3C45-for/rev and pHG3C6-for/rev). Among them, eryCII and eryCIII, and eryCIV and eryCV are located adjacent to each other and are possibly translationally coupled in the native erythromycin biosynthesis gene cluster. Their putative operon organisation was maintained for the assembly and the two gene pairs were amplified generating two separate fragments. Thus, four DNA fragments: eryCII, the fragment containing eryCII and CIII, the fragment containing eryCV and eryCIV, and eryCV and CII were cloned individually into pHG2A, to give plasmids pHG3C1, pHG3C23, pHG3C45 and pHG3C6. In each plasmid, the gene fragment containing the eryC gene(s) was under the control of the actIp promoter. Each cassette used in assembly was amplified from plasmids pHG3C1 – pHG3C6 using the primer pairs listed in Table 2 (TG1 attB F, Bxb1 attB R, Bxb1 attP F, Int9 attB R, Int9 attP F, Int4 attB R, Int4 attP F, Int7 attB R, Int7 attP F and SPBc attB R). As for the assembly of the eryB genes, the primers contain the attP or attB sites and a region at the 3’end that binds to vector sequences derived from pHG2A. In vitro recombination between all matching attP-attB pairs via the integrases joins the cassettes together end-to-end in a pre-determined order and inserts the resulting linear assembly into the pHG9A vector, replacing the lacZa.

The assembly strategy is shown in Fig. 3A. In the eryC assembly, five integrases were used: TG1 [23], Bxb1 [24], Int9, Int7 [8] and SPBc [22], each recognising only its cognate attP and attB sites located in the amplified gene cassettes.

2.4.3. Assembly reactions

The enzymes were prepared as previously described [21,25]. Recombination reactions contained one of four different buffers: Buffer 1 (pH3.1 RxE buffer [26]), buffer 2 (10 mM Tris-HCl [pH 7.5], 5 mM spermidine·3HCl, 2 mM dithiothreitol, 0.1 mM EDTA, 0.1 mg/ml...
bovine serum albumin and 4.5% glycerol, adapted from Refs. [11]), buffer 3 (Bxb1 RxE buffer [26]) and buffer 4 (TGI RxE buffer [26]), in 20 µl final volume. The final concentrations of integrases were: 200 nM for TGI and Bxb1, 100 nM for Int9, Int4 and Int7, and 50 nM for SPBr.

Recombination substrates were used at 50 ng each per reaction. Reactions were incubated at 30 °C overnight and then heated (10 min, 75 °C) to denature integrases. The reaction mixtures were transformed into E. coli Top10 and blue-white screening was used to select colonies containing assembled plasmids.

2.5. Production and analysis of erythromycin intermediates

Strains were cultured in 25 ml of R2YE medium [19] as seed culture medium for 3 days at 30 °C, then 1.25 ml of inoculum was transferred to 25 ml of R2YE medium fed with 0.2 ml of propan-1-ol at 30 °C for fermentation. The extraction process was carried out as described previously [16].

Dried extracts were re-suspended in methanol (200 µl), and an aliquot (5 µl) was injected for analysis by high-pressure liquid chromatography (Acuity I class; Waters, Elstree, UK) on a reverse-phase BEH C18 column (100 by 2.1 mm, 1.7 µm, Waters; including a Waters separation guard; 4.5% glycerol, adapted from Refs. [27, 28]) and B (0.1% [vol/vol] formic acid in methanol) were mixed to give 40% B initially and held at 40% B upon injection for 0.2 min, and then a linear gradient was applied from 40 to 100% B over 3.8 min. The solvent was then held at 100% B for 0.34 min prior to reversion back to 40% B and equilibration for 1.32 min before the next injection. Total method duration was 5.7 min. Samples contained 5 µM internal standard (IS) Roxithromycin (Rox; Sigma-Aldrich, Gillingham, UK). Calibration standards containing a fixed amount of the internal standard, Rox (5 µM), and increasing amounts of Erythromycin A (EryA; Sigma Aldrich, Gillingham, UK), namely 0.2–100 µM, were prepared and used to quantify Erythronolide B (EB), 6-deB and 3-O-alpha-L-Mycarosylerythronolide B (MEB). Both EB and 6-deB were provided by Isomerase Therapeutics, purified from fermentations with blocked mutants of S. erythraea. Standards for EB and 6-deB were used for qualitative but not quantitative purposes. Analytes were identified by their mass spectra and MS/MS fragmentation profiles using electrospray ionization mass spectrometry (Fusion Orbitrap, Thermo Fisher Sci, Loughborough, UK; positive-ion mode; capillary voltage 3500 V; mass range, 150–1700 m/z; Orbitrap resolution: 240000; HCD fragmentation).

Values for analytes were calculated using Xcalibur software (v. 4.0.27.10, Thermo Fisher Sci, Loughborough, UK) by plotting normalised areas of extracted ion chromatograms (EICs) corresponding to the sodium ion adducts of the molecular ion of 6-deB ([M + Na] + = 409.256 m/z), EB ([M + Na] + = 409.256 m/z), MEB ([M + Na] + = 596.330); and the protonated ions of EryA ([M + H] + = 734.468) and Rox ([M + H] + = 837.531).

3. Results

3.1. Production of the erythromycin intermediate EB

The biosynthesis pathway of erythromycin A is shown in Fig. 4. This study was designed to assemble eryB and eryC genes using orthogonal integrases. A strain producing EB was needed to be the starting strain, and then the production of MEB or erythromycin D could be tested after the assembly of eryB or eryC genes, to confirm the assembled pathways are functional. So the eryF expression plasmid pHG1 was introduced by conjugation into the 6-deB producers constructed previously; S. coelicolor M1152:pBF20:pBF22:pBF24 (M1) and S. lividans TK24:pBF20:pBF22:pBF24 (T1) [16]. Genomic DNA from the recipient strains was extracted, and PCR was used to confirm the integration.

Fig. 2. The assembly of DNA fragments. (A) The construction of DNA fragments to be assembled. The eryB or eryC genes were cloned into pJET1.2/blunt, under the control of the promoter act1p. Next, primer pairs (F: forward primer; R: reverse primer) contained attB or attP sites (indicated using triangle) at the 3’ end were used to amplify the DNA fragment used in the assembly reaction. (B) The assembly strategy of pHG2R2. The length of each fragment is: pHG2B0 (10.2 kb), eryBII (1.4 kb), eryBIII (1.7 kb), eryBIV-BV (2.7 kb), eryBVI (1.9 kb) and eryBVI (1.0 kb); (C) The assembly reactions were carried out in four buffers, and the assembled plasmids were checked by digesting with BglII. Stars mark the clones have been assembled correctly. B: pHG2B0, M: NEB Fast DNA Ladder (https://international.neb.com/products/n3238-fast-dna-ladder#Product%20Information). 0.8% agarose gel was used to check the digested bands.
Furthermore, in both M2 and T2, only a tiny amounts of 6-dEB (~1 mg/L) were achieved similar levels of EB production (~10 mg/L) and biomasses 6-dEB produced in M1 was lower than that in T1 [S. coelicolor TK24:pBF20:pBF22:pBF24:pHG1 (T2), the expected ion (m/z = 425.3) indicative of EB could be found (peak 2, Fig. 2C). Six bands after digestion, as shown in the lanes marked with stars in Fig. 2C, suggested that all eryB genes had been incorporated. The assembly reactions showed different efficiencies in the four buffers (Fig. 2C), and using buffer 4, most (75%) of the white colonies contained the correctly assembled plasmids.

Efforts were made to assemble the eryC genes in different orders (Table S1). When the order of fragments was changed from the standard B2–B3–B45–B6–B7, the assembly efficiency dropped. For example, when the eryB genes were assembled in the order B45–B6–B3–B2–B7 (the plasmid was named pHG2R2), all the white colonies from the same reaction in the previously identified optimal buffer 4 did not contain any correctly assembled plasmids, and correctly assembled plasmids were only achieved when buffer 2 was used. The results suggested that the multiplex integration could work, but as a complex reaction including six integrases, further investigation will be required to learn the reaction kinetics, which will be extremely helpful in optimising the reactions.

Next, the assembled plasmids were transferred into M1 and T1 individually. After 7 days fermentation in R2YE medium fed with propan-2-ol, MEB production ability was tested. None of the S. lividans strains produced any MEB in their fermentation broth, while amongst S. coelicolor strains, S. coelicolor M1152:pBF20:pBF22:pBF24:pHG2R2 (M4) produced a very small amount of MEB (~1.4 mg/L, Fig. S2). So, the plasmid pHG2R2 was modified into pHG9A for the assembly of eryC genes.

### 3.2.3. Assembly of eryC genes

The six eryC genes were assembled as described in Materials and Methods in buffer 4. Then eight white colonies were picked randomly; plasmids were extracted and digested with XbaI to check if all the eryC genes had been incorporated. The assembly reactions showed different efficiencies in the four buffers (Fig. 2C), and using buffer 4, most (75%) of the white colonies contained the correctly assembled plasmids.
However, considering the size of the final plasmid increases from 19.2 kb (pHG2R2) after eryBs assembly to 28.1 kb (pHG22A) after eryCs assembly (Fig. 1), the correct plasmids' ratio is still satisfactory. However, when the eryC genes were assembled in different orders (Table S1), no correct clone was obtained. Three plasmids that were incorrectly assembled were picked randomly and sent for sequencing. The sequencing results showed that the SPBc attP site (GTAGTAAGTATCTTAATATAGCTTTATCTGTTTTTTAAGATACTTACTACTTT) frequently recombined with a sequence upstream of the act1p promoters (TGCTCGTGTAGCACCGGTCCGTGAACGCGGTGGAGCCCCTCTTAAG) through the nucleotides TTA. Since there are multiple act1p promoters in the backbone plasmid and the fragments being assembled, there were many opportunities for recombination between the SPBc attP and the off-target site. The net result was that the incorrectly assembled plasmids had unwanted deletions. However, why the recombination between the SPBc attP and the off-target site was much higher in some reactions and not others remains unclear; further exploration is required.

4. Conclusions and discussion

In this study, we tested and optimised protocols for multiplexing
integrase to assemble antibiotic pathway genes into integrating plasmids. The erythromycin pathway from *S. erythraea* was chosen as an example in this work. The erythromycin pathway PKS genes had already been cloned into three orthogonal integrating vectors and the production of their product 6-DEB had been demonstrated [16]. The next big challenge was to assemble the smaller genes required for glycosylation of the aglycone and to place each one downstream of a promoter that would express in a heterologous host. Using multiple integrases in an in vitro recombination reaction five eryB DNA fragments (9.0 kb total and assembled plasmid is nearly 20 kb) were initially assembled into an integrating plasmid compatible with the three already expressing the PKS genes. The four fragments encoding the remaining *eryC* genes were also inserted to generate a plasmid of around 30 kbp.

The design of the assembly pathway employed here allowed a combinatorial approach for the order of the DNA fragments being assembled. This is because the primers tagged with the different integrase *attP/B* sites bound universally to the same site in the vector backbones into which the *ery* genes had been cloned. However, when we tried to assemble the *eryB* genes in different orders, the ratio of successfully assembled vectors dropped dramatically. More research on enzyme kinetics is needed. For example, in this study, the activities of the integrases in different buffers were only tested qualitatively. In the future, the activities could be assessed quantitatively, thus to help us understand the enzyme kinetics in the multiplexed assembly reaction.

The application of serine integrases in pathway assembly has been explored in several previous publications [11,14,25], and a couple of review papers have discussed methods used for pathway assembly, including the ones based on serine integrases [10,27,28]. All the previous publications using serine integrases to assemble DNA fragments were based on using a single integrase (Table 3). There are other reports using a combination of a serine integrase and other enzymes to assemble complex pathways, usually with tyrosine integrases, Cre [29,30] or FLP recombinases [31]. As far as we are aware, the work presented here is the first to exploit a multiplex integration system based on using more than one integrase at a time. For five fragments assembled into a 10 kb backbone vector, 75% of colonies were the correctly assembled plasmids, and for four fragments assembled into the 20 kb backbone vector, 50% of colonies contained the desired construct. Although the reliability needs to be further improved, and more investigations are required to avoid the off-target recombination, the high efficiency of this assembly strategy still suggests that there is potential for its use in pathway assembly and synthetic biology.

To the best of our knowledge, this study is the first time that multiplex recombination based on serine integrases has been used to assemble multiple genes into pathways. These results open the way for further investigation of the application and potential of this enzyme family in synthetic biology and natural products biosynthesis.

**Margaret C.M. Smith:** Conceptualization, Writing - review & editing, Supervision, Project administration, Funding acquisition.

**Acknowledgements**

This work was supported by the Biotechnology and Biological Sciences Research Council project grant BB/K003356/1. HPLC-MS analysis was performed on equipment within the Centre of Excellence in Mass Spectrometry (University of York), funded by Science City York (Yorkshire Forward, EP/K039660/1, EP/M028127/1). The strain *Saccharopolyspora erythraea* BIOT-0666 was a gift from Dr Matthew Gregory, Isomerase Therapeutics Ltd, and we also thank Professor Marshall Stalk for the proof reading of the manuscript.

**Appendix A. Supplementary data**

Supplementary data to this article can be found online at https://doi.org/10.1016/j.synbio.2020.05.006.

**References**

[1] Rutherford K, Van Duyne GD. The ins and outs of serine integrase site-specific recombination. Curr Opin Struct Biol 2014;24:125–31.
[2] Thorpe HM, Smith MC. In vitro site-specific integration of bacteriophage DNA catalyzed by a recombinein the resolvase/invertase family. Proc Natl Acad Sci USA 1998:95(10):5505–10.
[3] Saha S, et al. Activation and characterization of a cryptic gene cluster reveals a cyclization cascade for polytetramethyl macro lactams. Chem Sci 2017;8(21):1607–12.
[4] Sosio M, et al. Artificial chromosomes for antibiotic-producing actinomycetes. Nat Biotechnol 2000;18(3):343–5.
[5] Hong Y, Hondalus MK. Site-specific integration of Streptomyces *pc31* integrase-based vectors in the chromosome of *Rhodococcus equi*. FEMS Microbiol Let 2008;287(1):63–8.
[6] Xu Z, et al. Accuracy and efficiency define Bxb1 integrase as the best of fifteen candidate serine recombines for the integration of DNA into the human genome. BMC Biotechnol 2013;13(1):87.
[7] Stark WM. Making serine integrases work for us. Curr Opin Microbiol 2017;38:130–6.
[8] Yang L, et al. Permanent genetic memory with > 1-byte capacity. Nat Methods 2014;11(12):1261–6.
[9] Smith MC. Phage-encoded serine integrases and other large serine recombinases. Mobile DNA iii. American Society of Microbiology; 2015. p. 253–72.
[10] Fogg PC, et al. New applications for phage integrases. J Mol Biol 2014;426(15):2703–16.
[11] Collins SD, et al. Rapid metabolic pathway assembly and modification using serine integrase site-specific recombination. Nucleic Acids Res 2013;42(4):e23.
[12] Smith MC, Till R, Smith MC. Switching the polarity of a bacteriophage integration system. Mol Microbiol 2004;5:16(1):1719–28.
[13] Ghosh P, Kim AI, Hanfull GF. The orientation of mycobacteriophage Bxb1 integration is solely dependent on the central dinucleotide of attP and attB. Mol Cell 2003;12(5):1101–11.
[14] Zhang L, Zhao G, Ding JXt. Tandem assembly of the epothilone biosynthetic gene cluster by in vitro site-specific recombination. Sci Rep 2011;1:141.
[15] Zhu F, et al. Dice, an efficient system for iterative genomic editing in human pluripotent stem cells. Nucleic Acids Res 2013;41(5):e34.
[16] Payed B, et al. Multiplexed integrating plasmids for engineering of the erythromycin gene cluster for expression in Streptomyces spp. and combinatorial biosynthesis. Appl Environ Microbiol 2015;81(24):4802–13.
[17] MacNeil DJ, Characterisation of a unique methyl-specific restriction system in *Streptomyces avermitilis*. J Bacteriol 1998;170(12):5607–12.
[18] Gomez-Escribano JP, Bibb MJ. Engineering Streptomyces coelicolor for heterologous expression of secondary metabolite gene clusters. Microbiol Biotechnol 2011;4(2):207–15.

**Table 3**

| Integrase used | Product synthesised | Pathway assembled | Percentage of correctly assembled | Reference |
|---------------|-------------------|--------------------|-----------------------------------|-----------|
| φBT1          | Epothiline        | 56                 | 9                                 | Not reported | 14 |
| φC31          | Lycopene, β-carotene, and zeaxanthin | Up to 7.8 | 8 | 48% for 7.8 kb assembly | 11 |
| φBb1          | Lycopene         | 3.6                | 3                                | > 96%       | 25 |
| Multiple integrases | EB            | 8.7                | 5                                | 75%         | This study |
|               | MEB            | 8.4                | 4                                | 50%         | |

**CRediT authorship contribution statement**

**Hong Gao:** Methodology, Formal analysis, Investigation, Writing - original draft, Visualization. **Gabrielle Taylor:** Resources, Writing - review & editing. **Stephanie K. Evans:** Resources, Writing - review & editing. **Paul C.M. Fogg:** Resources, Writing - review & editing.
[19] Kieser T, et al. Practical Streptomyces genetics. Norwich: The John Innes Foundation; 2000.
[20] Sambrook J, Russell DW. Molecular cloning: a laboratory manual. New York: Cold Spring Harbor Laboratory Press; 2001.
[21] Morita K, et al. In vitro characterization of the site-specific recombination system based on actinophage TG1 integrase. Mol Genet Genom 2000;282(6):607–16.
[22] Lazarevic V, et al. Nucleotide sequence of the Bacillus subtilis temperate bacteriophage SPβc2. Microbiology 1999;145(5):1055–67.
[23] Morita K, et al. The site-specific recombination system of actinophage TG1. FEMS Microbiol Lett 2009;297(2):234–40.
[24] Mediavilla J, et al. Genome organization and characterization of mycobacteriophage Bxb1. Mol Microbiol 2000;38(5):955–70.
[25] Wang X, et al. Bxb1 integrase serves as a highly efficient DNA recombinase in rapid metabolite pathway assembly. Acta Biochim Biophys Sin 2017;49(1):44–50.
[26] Fogg PC, et al. Genome integration and excision by a new Streptomyces bacteriophage, φJoe. Appl Environ Microb 2017;83(5). e02767-16.
[27] Merrick CA, Zhao J, Rosser SJ. Serine integrases: advancing synthetic biology. ACS Synth Biol 2018;7:299–310.
[28] Zhang JJ, Tang X, Moore BS. Genetic platforms for heterologous expression of microbial natural products. Nat Prod Rep 2019;36:1313–32.
[29] Xu Z, Lee NCO, Dalhuis-Calas F, Malla S, Smith MCM, Brown WRA. Site-specific recombination in Schizosaccharomyces pombe and systematic assembly of a 400kb transgene array in mammalian cells using the integrase of Streptomyces phage q8T1. Nucleic Acids Res 2008;36(1):e9.
[30] Lee NCO, Kim J, Petrov NS, Lee H, Masumoto H, Earnshaw WC, Larionov V, Kouprina N. Method to assemble genomic DNA fragments or genes on human artificial chromosome with regulated kinetochore using a multi-integrase system. ACS Synth Biol 2018;7:63–74.
[31] Neil K, Allard N, Jordan D, Rodrigue S. Assembly of large mobilizable genetic cargo by double recombinase operated insertion of DNA (DROID). Plasmid 2019;104:102419.