J Ind Microbiol Biotechnol
DOI 10.1007/s10295-011-1049-x

Recombinatorial biosynthesis of polyketides

Antonio Starcevic · Kerstin Wolf · Janko Diminic · Jurica Zucko · Ida Trninic Ruzic · Paul F. Long · Daslav Hranueli · John Cullum

Received: 13 July 2011 / Accepted: 15 October 2011

Abstract Modular polyketide synthases (PKSs) from *Streptomyces* and related genera of bacteria produce many important pharmaceuticals. A program called *CompGen* was developed to carry out in silico homologous recombination between gene clusters encoding PKSs and determine whether recombinants have cluster architectures compatible with the production of polyketides. The chemical structure of recombinant polyketides was also predicted. In silico recombination was carried out for 47 well-characterised clusters. The predicted recombinants would produce 11,796 different polyketide structures. The molecular weights and average degree of reduction of the chemical structures are dispersed around the parental structures indicating that they are likely to include pharmaceutically interesting compounds. The details of the recombinants and the chemical structures were entered in a database called *r-CSDB*. The virtual compound library is a useful resource for computer-aided drug design and chemoinformatics strategies for finding pharmaceutically relevant chemical entities. A strategy to construct recombinant *Streptomyces* strains to produce these polyketides is described and the critical steps of mobilizing large biosynthetic clusters and producing new linear cloning vectors are illustrated by experimental data.

Keywords Polyketides · Actinobacteria · Gene clusters · Homologous recombination

Electronic supplementary material The online version of this article (doi:10.1007/s10295-011-1049-x) contains supplementary material, which is available to authorized users.

A. Starcevic · J. Diminic · J. Zucko · I. T. Ruzic · D. Hranueli
Faculty of Food Technology and Biotechnology,
University of Zagreb, Pierottijeva 6, 10000 Zagreb, Croatia

K. Wolf · J. Cullum (✉)
Department of Genetics, University of Kaiserslautern,
Postfach 3049, 67653 Kaiserslautern, Germany
e-mail: cullum@rhrk.uni-kl.de

Present Address:
I. T. Ruzic
Novalis d.o.o., Božidara Adžije 17, 10000 Zagreb, Croatia

P. F. Long
Institute of Pharmaceutical Science, King's College London,
Franklin–Wilkins Building, Stamford Street,
London SE1 9NH, UK

Introduction

Polyketides produced by *Streptomyces* and related genera are of great importance in the pharmaceutical industry (e.g. the antibiotic erythromycin and the immunosuppressant rapamycin). Many are produced by modular polyketide synthases (PKSs), which suggested the idea of combinatorial biosynthesis, whereby different sets of modules can be recombined to produce a vast number of novel hybrid polyketides [1, 11]. Although experiments to generate hybrid polyketides have had some success, in most cases product yield is extremely low preventing industrial exploitation [9]. Studies on the natural evolution of modular PKSs suggest that recombination plays a major role [10]. In vitro construction of hybrid PKSs creates unnatural junctions, which may account for the low product yield as there is evidence for strong purifying selection in PKS modules [30], i.e. most variation in amino acid sequence is not tolerated. Homologous recombination occurs between DNA sequences with high sequence similarity [24] and it is likely that there will be fewer problems when such a junction is used to generate novel polyketides.

Published online: 01 November 2011

A homologous recombination model was developed [26] and used to study single crossovers between gene clusters encoding modular PKSs. The model assumed that homologous recombination required a region of sequence identity in the two parent molecules, where the crossover occurs, flanked on both sides by regions with a high degree of sequence similarity. Recombinants between 12 PKS clusters were studied and it was shown that many recombinants had a genetic structure compatible with the synthesis of novel products.

Many molecular genetic techniques have been established in *Streptomyces* species and it is possible to manipulate modular PKS clusters [1]. The size of such clusters (sometimes over 100 kb) presents challenges and makes recombineering approaches attractive. This requires systems to select for replacement of gene segments. It has been shown that, as in many other bacteria, the *rpsL* gene can be used for counter-selection in *Streptomyces* species [8]. Recombination corresponding to the in silico recombination model [26] requires single crossovers and these would be most conveniently generated using linear vectors. Linear plasmids are common in *Streptomyces* and it is possible to generate derivatives to investigate the functions needed for plasmid maintenance [21, 29].

In this paper, we describe a convenient computer program *CompGen*, based on the annotation program *ClustScan* [25], for producing in silico recombinants and predicting the chemical structure of any products. This was used to explore the chemical space generated. We describe the construction of a cassette system based on *rpsL* and show that it is possible to mobilize a complete antibiotic producing cluster. We also describe the construction of a linear vector suitable for use with the cassette system.

Materials and methods

Polyketide gene clusters

Modular PKS clusters that were used for the in silico recombination were derived from *Streptomyces* and related genera (Supplementary Table 1).

Computer programs

The recombination model for the *CompGen* module was implemented as a PERL program on the server. Maximal common subsequences between two clusters were found using a standard suffix tree-based algorithm [2]. A region symmetrically located around the common subsequence was aligned using the Needleman–Wunsch algorithm [17] from the EMBOSS package [22]. The recombination parameters can be set by the user and the parameters used in this paper were minimum length of sequence identity 27 bp, length of surrounding region 100 bp, nucleotide identity in surrounding region 75%. The *ClustScan* program [25] uses an XML format to describe the location of genes, modules and domains as well as the biosynthetic order and the chemistry produced by each module. The recombination program generates the XML description of the recombinant clusters from those of the parent clusters. The graphical user interface was written in Java as a module of the *ClustScan* program. It has functions to select the two clusters to be recombined and the recombination parameters (Supplementary Fig. 1a). The annotated sequences are shown in a workspace window (Supplementary Fig. 1b). The recombinants generated by the server program can be viewed in a workspace window of the client. A cartoon overview (Fig. 1a) shows the location of the crossover in comparison to modules. A details window (Supplementary Fig. 1c) shows the sequences used for the recombination and the values of the recombination parameters. The sequences can be filtered (Supplementary Fig. 1d) to eliminate those that are unlikely to produce recombinant products. The predicted chemical structures of the products were generated as isomeric SMILES (Simplified Molecular Input Line Entry System) [28] (Supplementary Fig. 1e) using the same procedure as in *ClustScan*, but showing all the recombinants in one window. A Perl program was written to calculate the molecular weight of polyketides from the SMILES description as well as the degree of reduction: unreduced (keto groups), hydroxyls, double bonds and fully reduced were given scores of 0, 3.3, 6.7 and 10, respectively, and the average degree of reduction for all positions in the molecule calculated.

Mobilization of the actinorhodin cluster

Standard methods were used for genetic manipulation of *Streptomyces* [12] and *E. coli* [23]. An *Eco*RI-*Hin*dIII cassette was constructed carrying the ends of the actinorhodin biosynthesis cluster separated by a wild-type *rpsL*$^+$ gene. The left end of the actinorhodin biosynthesis cluster (1,129 bp, coordinates 5511250–5512378 in the *S. coelicolor* chromosome, accession number AL645882) was amplified using PCR primers carrying *Eco*RI and *Bgl*II restriction sites (GGA ATT CCA CGA CAC AGG AGA CGA AGG, GAA GAT CTT GCA CGG ACT TCG AGA CG). Similarly the right-hand end (1,116 bp, coordinates 5539867–5540982) was amplified using primers with *Bgl*II and *Hind*III sites (GAA GAT CTG GCG TCC TGT GCT TCC TG, CCC AAG CTT AGC AAC AGC GTG GCA ACC). The amplified fragments were digested with *Eco*RI-*Bgl*II and *Bgl*II-*Hind*III, respectively, and cloned into *Eco*RI-*Hind*III digested pUC18 [18] vector to give plasmid pKW40. The *rpsL* gene of *S. coelicolor* A3(2) was

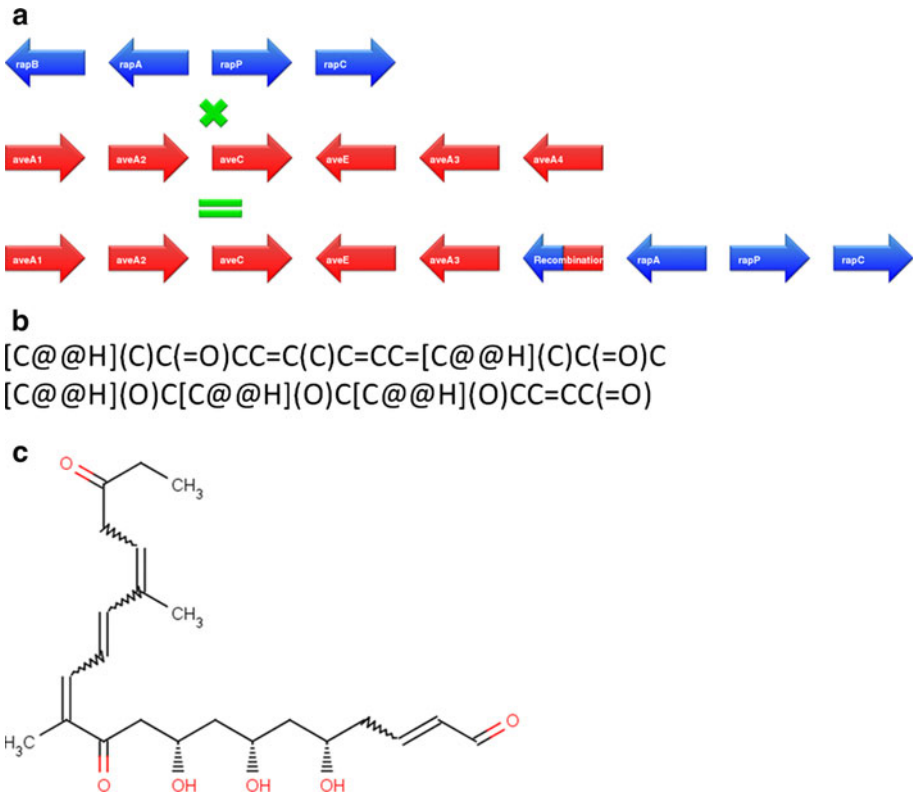

Fig. 1 **a** Screen shot of the cartoon window of *CompGen* program showing generation of a recombinant between the rapamycin cluster (*rapB*, module 9) and the avermectin cluster (*aveA4*, module 12). The recombination parameters were minimum length of sequence identity 27 bp, length of surrounding region 100 bp, nucleotide identity in surrounding region 75%. **b** The predicted chemical structure of the product in a SMILES format. **c** The chemical structure displayed using ChemAxon (http://www.chemaxon.com/products/)

available as a 640-bp *Bam*HI fragment [20] produced by a PCR reaction with primers carrying *Bam*HI sites (CGG GATCCC GGT GTC GCC CAT TTG T, CGG GAT CCG CCC TTA CGA GGC ATT CT). This fragment was cloned into the *Bgl*II site of pKW40 to give pKW040. The mobilizing cassette was excised from pKW040 as an *Eco*RI-*Hin*dIII fragment and cloned into *Streptomyces* vector pIJ903 [14] to give plasmid pKW140. pKW140 was transformed into the Str^R strain *S. lividans* 66 TK64 [7] and transformants selected on thiostrepton-containing medium. Single transformant colonies were transferred to medium containing thiostrepton and streptomycin to produce independent streptomycin-resistant derivatives.

Construction of linear vector pKW201

pST200 [27] carries the 298-bp left terminal *Bam*HI fragment of linear plasmid SLP2 cloned into *Bam*HI-*Hin*cII cut vector pUC18 [18]. This was used to construct the linear vector pCL204, which contains the replication origin of linear plasmid pSLA2 as a 6-kb *Mlu*I fragment derived from plasmid pQC48 [21] and a thiostrepton resistance gene (*tsr*) for selection in *Streptomyces* strains between two copies of the SLP2 end fragment. The SLP2 ends are separated by the *E. coli* vector pUC18 so that the vector can be obtained as a circular plasmid from *E. coli* and can replicate as a linear plasmid in *Streptomyces* strains losing the pUC18 sequences. The unique *Hin*dIII site of pCL204, which is present in the pUC18 sequences, was removed by digestion with *Hin*dIII, treating with Klenow enzyme to remove the 5′ overhangs and relegation. A linker molecule (sequences of the two strands, TGT CTG CAA TTG GTC GTA AGC TTG TCG TG and ACA GAC GTT AAC CAG CAT TCG AAC AGC AC) containing sites for *Mun*I and *Hin*dIII was ligated into the unique *Pme*I site to yield vector pKW201 (Fig. 6).

Measurement of actinorhodin concentration in medium

Ten-millilitre cultures were grown for 3 days in YEME medium [12] at 30°C with shaking. After removal of mycelium by centrifugation, the pH of the culture medium was lowered to 2.5 by addition of HCl. The medium was extracted with half the volume of 1:1 chloroform/methanol and the actinorhodin concentration estimated by measuring OD_{542} and using a molar extinction coefficient of actinorhodin of 18,600 [3].

Results

Modelling recombination of PKS clusters

The *CompGen* program was written to implement the recombination model and determine whether recombinants will produce a polyketide product. Recombination sites

were found by using a suffix tree approach [2] to identify regions of identity and the Needleman–Wunsch algorithm [17] to assess degree of similarity in the surrounding regions. Figure 1a shows an example of such a recombination in a cartoon form: the left-hand part of the recombinant cluster is derived from the avermectin cluster and the right-hand part from the rapamycin cluster.

The most difficult problem was to predict whether recombinants would synthesize a polyketide product. The *ClustScan* program package was developed [25] to annotate PKS clusters and uses a novel data structure to link the DNA sequence, the polypeptide sequences and the chemical structure of the product. *CompGen* was integrated as a module of *ClustScan* (generating the new version of *ClustScan* called *ClustScan*-Professional; http://bioserv.pbf.hr/cms/) and utilizes this data structure to allow inheritance of the biosynthetic information from the parental clusters to the recombinant clusters. This allows *CompGen* to predict whether recombinants will produce a hybrid polyketide and also to predict the chemical structure of the product. In this paper, we use the most stringent criteria for product formation: the recombinant must have a loading domain, a contiguous set of modules in the biosynthetic pathway and a terminal module.

A graphical user interface allows the user to model recombination between pairs of clusters. For individual recombinants, an overview of the recombination is provided in a cartoon form (Fig. 1a) and the details can be viewed in a details window. Predicted products can be displayed and exported as isomeric SMILES [28] (Fig. 1b), which is a standard input format for chemoinformatics programs. For instance, the chemical structure can be displayed (Fig. 1c).

CompGen was used to model recombination between all possible pairs of 47 modular PKS clusters derived from *Streptomyces* and related genera. Of the 1,081 possible cluster pairs, 777 were predicted to yield one or more recombinants. A total of 20,187 recombinants were predicted to synthesize a polyketide product with 11,796 different chemical compounds being produced. The predicted linear polyketide backbones of the products were exported from the program in a SMILES format and were assessed for chemical diversity; two parameters were used. The molecular weights of the predicted recombinant compounds were calculated and the average was similar to those of the parents (473 vs. 481). The average degree of reduction was also calculated: each β-carbon atom in the polyketide backbone can have one of four degrees of reduction, which were given a score in the range of 0–10: keto group, hydroxyl group, double bond or fully reduced. When the two parameters are plotted against each other (Fig. 2) it can be seen that there is considerable chemical diversity and the recombinant molecules are clustered around the parent molecules, which are compounds of known pharmaceutical interest. The *r-CSDB* database (http://bioserv.pbf.hr/cms/) gives details of the parental clusters and the chemical structures of the recombinant molecules.

Mobilization of clusters

A general strategy for the construction of mobilization cassettes was developed (Fig. 3). PCR primers were constructed to amplify regions of about 1 kb in length at each end of an antibiotic biosynthesis cluster. The primers contained suitable restriction sites so that the two PCR fragments could be cloned into a standard *E. coli* vector (pUC18). A *Streptomyces rpsL* gene [8] was then cloned between the two end fragments completing the mobilization cassette that can be sub-cloned into different vectors as an *Eco*RI-*Hind*III fragment.

The mobilization strategy was tested using the actinorhodin biosynthesis cluster of *S. lividans* 66; the PCR primer sequences were deduced from the sequence of the closely related strain *S. coelicolor* A3(2). After construction of the cassette, it was sub-cloned in the *E. coli–Streptomyces* shuttle vector pIJ903 [14] to produce pKW140 (Fig. 4), which was transformed into TK64, a streptomycin-resistant strain of *S. lividans* 66 [7], which carries a resistant allele of *rpsL* in the chromosome. As the wild-type allele of *rpsL* in the mobilization cassette is dominant, this results in a streptomycin-sensitive strain [8]. Streptomycin-resistant derivatives were selected and about 25% of them overproduced actinorhodin, which could be seen as a deep blue pigment around the colonies (Fig. 5). These were candidates for mobilization events as an increased copy number results in blue colonies [15]. This was confirmed using pulsed-field gel electrophoresis (PFGE) after digestion of total DNA with the enzyme *Ase*I (Fig. 6a). Ten of 12 blue colonies gave two visible plasmid bands, whereas pIJ903 and pKW140 only have a single visible band. This is expected if mobilization has occurred, because the actinorhodin cluster contains an *Ase*I site. The presence of actinorhodin cluster sequences on these two DNA fragments was confirmed by Southern blotting (Fig. 6b). As the plasmids are shuttle vectors, they could be transformed into *E. coli* and their structure was confirmed by restriction analysis (data not shown). The strains with blue colonies, which carried a copy of the actinorhodin cluster on the plasmid, also produced much more pigment in liquid culture. The amount of actinorhodin in the medium for 24 clones was estimated spectrophotometrically. The parent *S. lividans* strain without plasmid produced 0.3 mg/l. The 24 clones produced amounts between 0.4 and 48.7 mg/l (median value 3.8 mg/l).

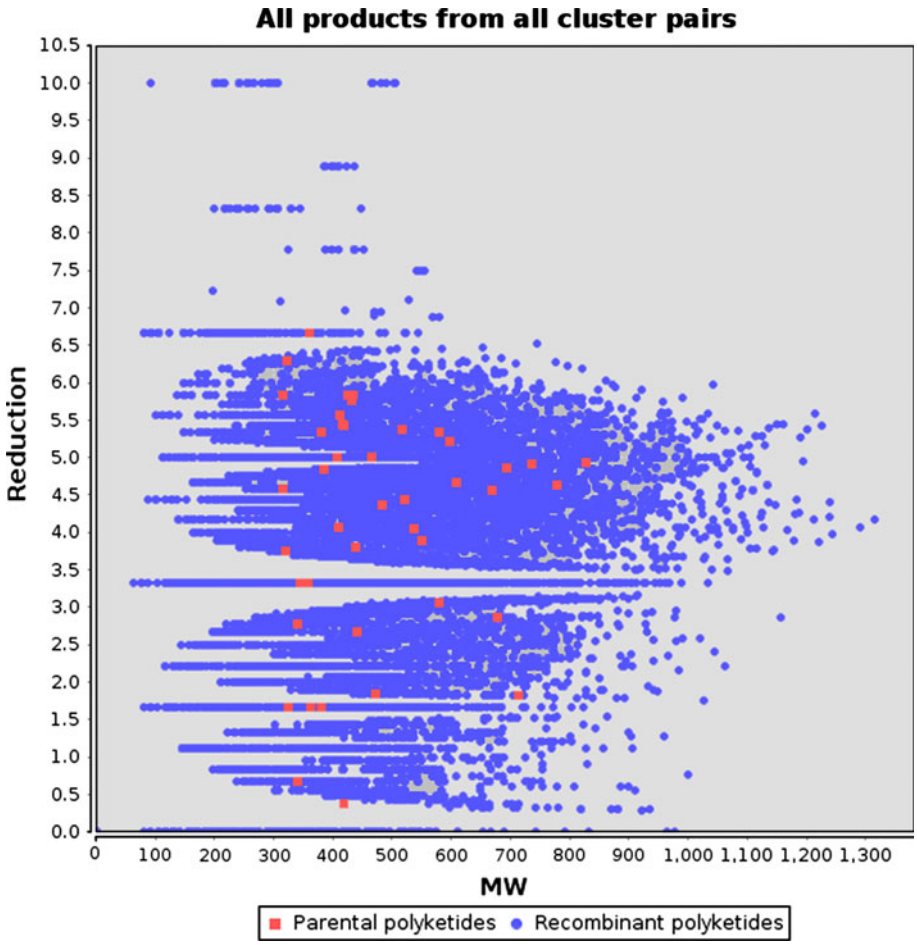

Fig. 2 Properties of the predicted recombinant polyketides (*blue*) and the parents (*red*). The linear polyketide backbones of 11,796 chemically different recombinant molecules and the 47 parent molecules were used to compute the molecular weight and the degree of reduction on a scale of 0–10. The same recombination parameters were used as in Fig. 1

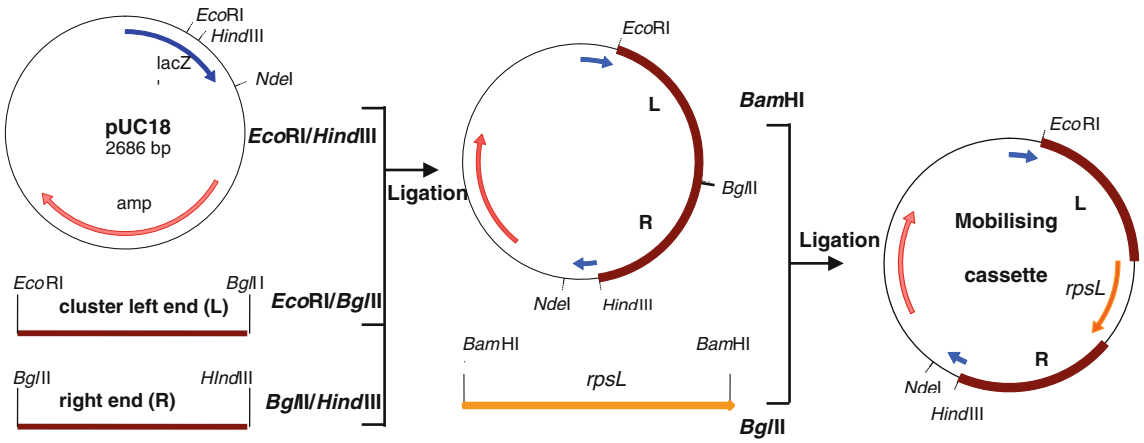

Fig. 3 Construction of mobilizing cassettes

Construction of a linear plasmid vector

The linear plasmid vector pKW201 (Fig. 7), which was designed to carry mobilization cassettes, was constructed as a circular plasmid in *E. coli* and becomes linear after transformation into a suitable *Streptomyces* host strain. The bidirectional replication origin was derived from linear plasmid pSLA2 [21]. The linear ends were derived from plasmid SLP2 [29] and a thiostrepton resistance gene (*tsr*) was used as a selective marker in *Streptomyces*. There are unique cloning sites for *Hind*III and *Mun*I in a linker region, which can be used for cloning *Hind*III-*Eco*RI mobilizing cassettes. DNA from pKW201 was transformed into *S. lividans* 66 and undigested total DNA from

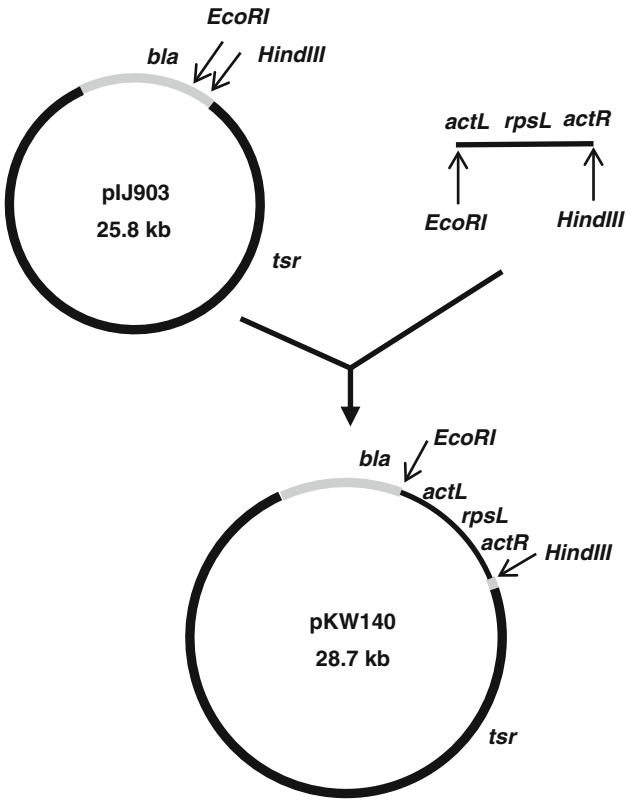

Fig. 4 Construction of pKW140. A mobilizing cassette (plasmid pKW040) for the actinorhodin cluster was constructed using the strategy shown in Fig. 3. The cassette was excised from pKW040 as an *Eco*RI-*Hind*III fragment and cloned into the *E. coli–Streptomyces* shuttle vector pIJ903. The parts of the plasmids shown in *black* are *Streptomyces*-derived sequences and the parts shown in *grey* are derived from *E. coli* vector pBR327, thus allowing replication and selection of pKW140 in both *Streptomyces* species and *E. coli*

Fig. 5 *S. lividans* 66 strains transformed with plasmid pKW140: the *right-hand side* shows a clone that was selected for Str^R, the *left-hand side* is a non-selected transformant. PCR was used to amplify sequences at each end of the actinorhodin biosynthesis cluster from *S. coelicolor* A3(2) and these sequences were assembled in an *Eco*RI-*Hind*III cassette separated by a wild-type $rpsL^+$ gene. The cassette was cloned into a low copy number *E. coli–Streptomyces* shuttle vector (pIJ903) and the resulting plasmid (pKW140) was introduced into a streptomycin-resistant (*rpsL*) mutant of *S. lividans* 66 [which is closely related to *S. coelicolor* A3(2)], strain TK64 selecting transformants with thiostrepton. When Str^R derivatives were selected, about 25% showed intense *blue colonies*, because of overproduction of actinorhodin

transformants was used for PFGE (Fig. 8a). The location of the plasmid band was confirmed by Southern blotting with a *tsr* probe (Fig. 8b). These experiments showed a single plasmid band (tracks 2–5) at a position consistent with the presence of linear molecules, whereas circular plasmid DNA migrated much more slowly (track 7).

Discussion

The *CompGen* program allowed the easy generation of theoretical recombinants with control of the recombination parameters. The identification of "natural" recombination sites is not trivial; on average, there are about 19 productive recombinants per cluster pair. The number of recombinants is proportional to the square of the number of clusters so that a threefold increase in the number of clusters will result in nearly ten times as many recombinants. Even with the 47 clusters used here, there are already almost 12,000 unique compounds. An important feature of the program is that it generates the chemical structures of the linear backbone of products and allows the export of the compounds in a standard SMILES format [28]. We wrote a program to calculate simple chemical parameters of the compounds: molecular weight and degree of reduction. Figure 2 shows that the recombinant compounds possess a high degree of chemical variability and are clustered around the parent compounds, which are of known pharmaceutical interest. The SMILES format also allows the use of more sophisticated chemoinformatics programs to look for promising compounds using a range of strategies. The recombinatorial approach will allow the mining of the complete range of potential natural chemical diversity. It is likely that the compounds currently produced by microorganisms have been selected to fulfil specific functions such as inhibiting competitors and other pharmacological effects are more or less incidental e.g. the immune suppressor rapamycin was first isolated because of its weak antifungal activity [5]. Thus, many chemical structures with useful pharmacological properties would not be selected in nature.

In silico studies can predict new polyketide products, but this is only useful if it is possible to generate strains producing them. *Streptomyces* species are attractive for this purpose as they have a well-developed molecular genetics and there is extensive experience of industrial fermentation. A further critical advantage is that *Streptomyces* chromosomes [4] and many plasmids [13] are linear so that

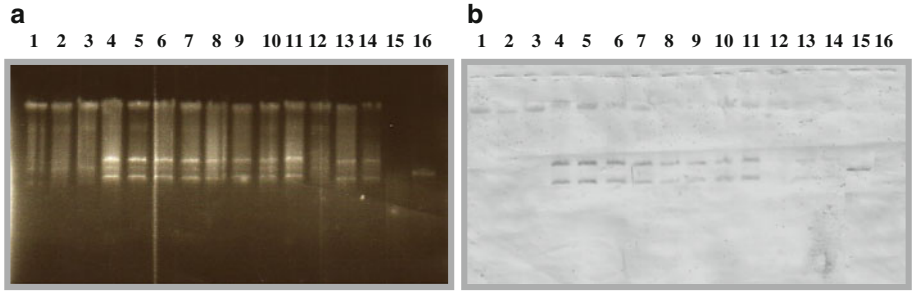

Fig. 6 *Ase*I digests of plasmids carrying the actinorhodin cluster. **a** PFGE gel of DNA digested with *Ase*I: *track 1*, *S. coelicolor* M146; *track 2*, *S. lividans* TK64; *tracks 3–14*, independent blue streptomycin-resistant clones of *S. lividans* TK64 (pKW140); *track 15*, plasmid pKW140; *track 16*, plasmid pIJ903. PFGE program: 6 V/cm for 14 h, pulse times ramped 2–6 s. **b** Southern blot of the gel from (**a**) hybridized with the 2.8-kb *Eco*RI-*Hin*dIII fragment from plasmid pKW040 carrying both ends of the actinorhodin cluster and the *rpsL* gene

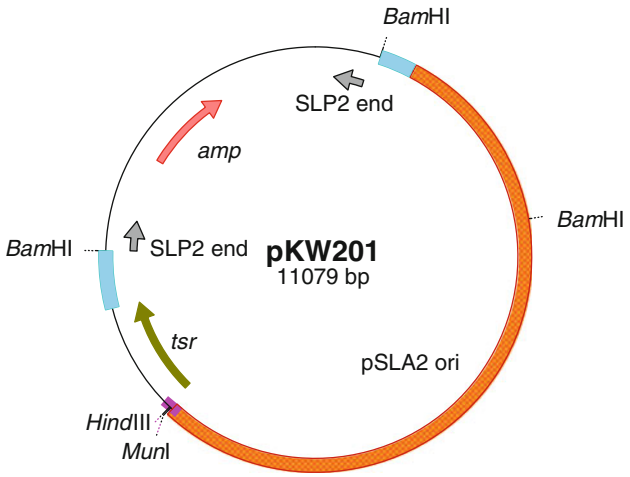

Fig. 7 Map of plasmid pKW201

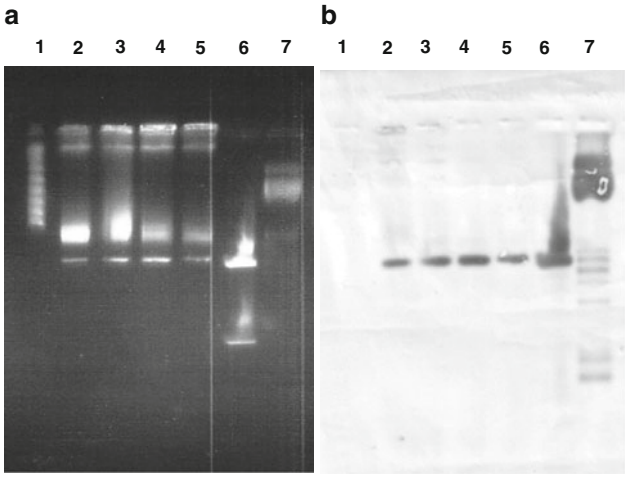

Fig. 8 PFGE of pKW201 in *S. lividans* 66. **a** *track 1*, *λ-Hin*dIII; *tracks 2–3*, two transformants of pKW201 into *S. lividans* TK64; *tracks 4–5*, two transformants of pKW201 into *S. lividans* 1326; *track 6*, pCL204 digested with *Hin*dIII and *Xho*I; *track 7*, pCL204 undigested (circular DNA). PFGE program: 5 V/cm, 16 h, ramped 20–40 s. **b** Southern blot of the gel from (**a**) hybridized with the *tsr* gene

single crossovers generate viable recombinant molecules [19]. The successful engineering of PKS clusters needs effective methods to mobilize them onto various replicons. We used a method in which the ends of a cluster are cloned into the target replicon and homologous recombination with a complete copy of the cluster transfers it onto the replicon. As such recombination events are rare, it is necessary to have a selection for transfer events. We chose the *rpsL* gene as a marker gene, because it had previously been used to select recombination events in a *Streptomyces* species [8]. Successful transfer of the actinorhodin gene cluster [15] was demonstrated in 10 out of 12 actinorhodin-overproducing blue colonies (Figs. 5, 6). There was also higher production in shake flasks (median value of blue 24 clones 3.8 mg/l compared to host strain 0.3 mg/l). However, the production of the blue clones fluctuated a lot, which probably reflected a need to optimize the fermentation conditions e.g. pH control [3]. This provides a general strategy for obtaining overproduction of sequenced clusters and for transferring them to other more convenient hosts. Two of the blue colonies did not show any visible plasmid bands (Fig. 6, tracks 3 and 12) and these clones were not investigated further. However, the fact that they are resistant to thiostrepton suggests that part of pKW140 (Fig. 4) including the *tsr* gene has probably integrated into the chromosome. It is conceivable that the resulting DNA rearrangements resulted in duplication of part of the *act* cluster resulting in overproduction; close examination of the digestion of one of the clones (Fig. 6a, track 12) shows some weak bands that could correspond to such a low copy number DNA amplification. We also describe the construction of the linear vector pKW201, which has restriction sites for the cloning of transfer cassettes.

Recombination between two clusters could be carried out in *Streptomyces* if both are carried on linear DNA molecules. It would be possible to select homoeologous recombination between clusters by incorporation of suitable selective markers flanking the clusters. The different

recombination events could be distinguished rapidly using a set of PCR primers for different regions of the two clusters. If a particular recombinant were wanted, the novel junction could be synthesized using standard PCR-based methods and used to force the desired recombination event. The advantage of such approaches is that, when an additional cluster is cloned, it can be used for recombination with all the existing cloned clusters allowing the number of compounds produced to grow rapidly. Continuing progress in synthetic biology [6, 16] should improve the methods of synthesizing long stretches of DNA and reduce the cost. PKS clusters are of significant size (50–150 kb) so they would need comparable technology to that used for chromosome synthesis of *Mycoplasma* species (1.08 Mb). Once nucleotide sequences of several clusters had been synthesized, they could be used to synthesize all of the recombinants between those clusters. A synthetic biology approach needs suitable vectors (e.g. the linear vector pKW201 described in this paper) and would benefit from the ability to transfer clusters between different plasmid vectors and/or the chromosomes of different species using the *rpsL*-based constructs.

An important question, which can only be answered experimentally, is how high the yield of novel polyketides will be. The fact that novel polyketide synthases, which evolve in nature, must have high enough yields to allow selection to act on the products, suggests that this approach is promising. As discussed above, this approach will allow access to polyketide structures that do not have selective advantages to *Streptomyces* hosts and most of the novel structures will probably have no antibacterial activity. However, any novel antibiotic is likely to kill the host as there will be no resistance mechanism. A strategy to overcome this problem would be to control expression of the new recombinant products. In many cases, this might be achieved by putting the transcriptional activator genes of clusters under control of a promoter that can be switched on, when production is desired.

A major issue for the pharmaceutical industry is maintaining a continuous supply of promising new leads for drug development. We propose that recombinatorial biosynthesis offers a new and exciting strategy whereby large and chemically diverse libraries of polyketides can first be screened in silico and then generated in the laboratory for further new lead development. Given that many polyketides are used clinically as antimicrobials, this new strategy comes at an important time, with ever increasing numbers of pathogens becoming resistant to our current antibiotic armamentarium.

Acknowledgments We thank David Hopwood for plasmid pIJ903 and Stanley N. Cohen for plasmid pQC48. This work was supported by a grant (058-0000000-3475 to D.H.) from the Ministry of Science, Education and Sports, Republic of Croatia and by a cooperation grant of the German Academic Exchange Service (DAAD) and the Ministry of Science, Education and Sports, Republic of Croatia (to D.H. and J.C.).

References

1. Baltz RH (2006) Molecular engineering approaches to peptide, polyketide and other antibiotics. Nat Biotechnol 24:1533–1540
2. Bieganski P, Riedl J, Carlis JV, Retzel EF (1994) Generalized suffix trees for biological sequence data. Proceedings of the twenty-seventh Hawaii international conference on system sciences, vol V (biotechnology computing), pp 35–44. doi:10.1109/HICSS.1994.323593
3. Bystrykh LV, Fernández-Moreno MA, Herrema JK, Malpartida F, Hopwood DA, Dijkhuizen L (1996) Production of actinorhodin-related "blue pigments" by *Streptomyces coelicolor* A3(2). J Bacteriol 178:2238–2244
4. Chen CW, Huang CH, Lee HH, Tsai HH, Kirby R (2002) Once the circle has been broken: dynamics and evolution of *Streptomyces* chromosomes. Trends Genet 18:522–529
5. Cruz MC, Cavallo LM, Görlach JM, Cox G, Perfect JR, Cardenas ME, Heitman J (1999) Rapamycin antifungal action is mediated via conserved complexes with FKBP12 and TOR kinase homologs in *Cryptococcus neoformans*. Mol Cell Biol 19:4101–4112
6. Gibson DG, Glass JI, Lartigue C, Noskov VN, Chuang RY, Algire MA, Benders GA, Montague MG, Ma L, Moodie MM, Merryman C, Vashee S, Krishnakumar R, Assad-Garcia N, Andrews-Pfannkoch C, Denisova EA, Young L, Qi ZQ, Segall-Shapiro TH, Calvey CH, Parmar PP, Hutchison CA 3rd, Smith HO, Venter JC (2010) Creation of a bacterial cell controlled by a chemically synthesized genome. Science 329:52–56
7. Hopwood DA, Kieser T, Wright HM, Bibb MJ (1983) Plasmids, recombination, and chromosomal mapping in *Streptomyces lividans* 66. J Gen Microbiol 129:2257–2269
8. Hosted TJ, Baltz RH (1997) Use of *rpsL* for dominance selection and gene replacement in *Streptomyces roseosporus*. J Bacteriol 179:180–186
9. Hranueli D, Cullum J, Basrak B, Goldstein P, Long PF (2005) Plasticity of the *Streptomyces* genome: evolution and engineering of new antibiotics. Curr Med Chem 12:1697–1704
10. Jenke-Kodama H, Dittmann E (2009) Bioinformatic perspectives on NRPS/PKS megasynthases: advances and challenges. Nat Prod Rep 26:874–883
11. Khosla C, Kapur S, Cane DE (2009) Revisiting the modularity of modular polyketide synthases. Curr Opin Chem Biol 13:135–143
12. Kieser T, Bibb MJ, Buttner MJ, Hopwood DA (2000) Practical *Streptomyces* genetics. The John Innes Foundation, Norwich
13. Kinashi H, Shimaji-Murayama M, Hanafusa T (1992) Integration of SCP1, a giant linear plasmid, into the *Streptomyces coelicolor* chromosome. Gene 115:35–41
14. Lydiate DJ, Malpartida F, Hopwood DA (1985) The *Streptomyces* plasmid SCP2*: its functional analysis and development into useful cloning vectors. Gene 35:223–235
15. Malpartida F, Hopwood DA (1986) Physical and genetic characterisation of the gene cluster for the antibiotic actinorhodin in *Streptomyces coelicolor* A3(2). Mol Gen Genet 205:66–73
16. Medema MH, Breitling R, Bovenberg R, Takano E (2011) Exploiting plug-and-play synthetic biology for drug discovery and production in microorganisms. Nat Rev Microbiol 9:131–137
17. Needleman SB, Wunsch CD (1970) A general method applicable to the search for similarities in the amino acid sequence of two proteins. J Mol Biol 48:443–453

18. Norrander J, Kempe T, Messing J (1983) Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. Gene 26:101–106
19. Pandza S, Biuković G, Paravic A, Dadbin A, Cullum J, Hranueli D (1998) Recombination between the linear plasmid pPZG101 and the linear chromosome of *Streptomyces rimosus* can lead to exchange of ends. Mol Microbiol 28:1165–1176
20. Paravic A (2001) PhD dissertation, University of Kaiserslautern, Kaiserslautern
21. Qin Z, Shen M, Cohen SN (2003) Identification and characterization of a pSLA2 plasmid locus required for linear DNA replication and circular plasmid stable inheritance in *Streptomyces lividans*. J Bacteriol 185:6575–6582
22. Rice P, Longden I, Bleasby A (2000) EMBOSS: the European molecular biology open software suite. Trends Genet 2000(16): 276–277
23. Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd edn. Cold Spring Harbor Laboratory, Cold Spring Harbor
24. Shen P, Huang HV (1986) Homologous recombination in *Escherichia coli*: dependence on substrate length and homology. Genetics 112:441–457
25. Starcevic A, Zucko J, Simunkovic J, Long PF, Cullum J, Hranueli D (2008) *ClustScan*: an integrated program package for the semi-automatic annotation of modular biosynthetic gene clusters and in silico prediction of novel chemical structures. Nucleic Acids Res 36:6882–6892
26. Starcevic A, Diminic J, Zucko J, Elbekali M, Schlosser T, Lisfi M, Vukelic A, Long PF, Hranueli D, Cullum J (2011) A novel docking domain interface model that can predict recombination between homoeologous modular biosynthetic gene clusters. J Ind Microbiol Biotechnol. doi:10.1007/s10295-010-0909-0
27. Stoll A, Horvat LI, Lopes-Shikida SA, Padilla G, Cullum J (2000) Isolation and cloning of *Streptomyces* terminal fragments. Antonie Van Leeuwenhoek 78:223–226
28. Weininger D (1988) SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. J Chem Inf Comput Sci 28:31–36
29. Xu M, Zhu Y, Zhang R, Shen M, Jiang W, Zhao G, Qin Z (2006) Characterization of the genetic components of *Streptomyces lividans* linear plasmid SLP2 for replication in circular and linear modes. J Bacteriol 188:6851–6857
30. Zucko J, Cullum J, Hranueli D, Long PF (2011) Evolutionary dynamics of modular polyketide synthases, with implications for protein design and engineering. J Antibiot 64:89–92