Discovery of a New Source of Rifamycin Antibiotics in Marine Sponge Actinobacteria by Phylogenetic Prediction

Tae Kyung Kim,1 Amita K. Hewavitharana,2 P. Nicholas Shaw,2 and John A. Fuerst1*

School of Molecular and Microbial Sciences, University of Queensland, Brisbane, Queensland 4072, Australia,1 and School of Pharmacy, University of Queensland, Brisbane, Queensland 4072, Australia2

Received 7 October 2005/Accepted 31 December 2005

Phylogenetic analysis of the ketosynthase (KS) gene sequences of marine sponge-derived Salinispora strains of actinobacteria indicated that the polyketide synthase (PKS) gene sequence most closely related to that of Salinispora was the rifamycin B synthase of Amycolatopsis mediterranei. This result was not expected from taxonomic species tree phylogenetics using 16S rRNA sequences. From the PKS sequence data generated from our sponge-derived Salinispora strains, we predicted that such strains might synthesize rifamycin-like compounds. Liquid chromatography-tandem mass spectrometry (LC/MS/MS) analysis was applied to one sponge-derived Salinispora strain to test the hypothesis of rifamycin synthesis. The analysis reported here demonstrates that this Salinispora isolate does produce compounds of the rifamycin class, including rifamycin B and rifamycin SV. A rifamycin-specific KS primer set was designed, and that primer set increased the number of rifamycin-positive strains detected by PCR screening relative to the number detectable using a conserved KS-specific set. Thus, the Salinispora group of actinobacteria represents a potential new source of rifamycins outside the genus Amycolatopsis and the first recorded source of rifamycins from marine bacteria.

Rifamycins, a group of antibiotics of the ansamycin family (3, 20), are clinically important antibacterial agents active against gram-positive bacteria. Several semisynthetic rifamycin variants (e.g., rifampin) have been used clinically for the treatment of tuberculosis and other bacterial infections, but resistance threatens their effectiveness (17). Rifamycins are known to be produced by the soil actinobacterial species Amycolatopsis mediterranei, and rifamycin synthesis has been most intensively studied in this species (2, 15, 16). The gene cluster for the synthesis of rifamycin B, a compound forming the basis for all other rifamycins, has been characterized only in A. mediterranei (2, 15). Five polyketide synthase subunits (RifA to RifE) produce rifamycin B using 3-amino-5-hydroxybenzoic acid as a starting point, and then RifF terminates the chain elongation to produce free rifamycin B (20). Rifamycin B is processed further by either natural modification enzymes or artificial semisynthetic protocols to produce biologically active rifamycins (Fig. 1). Widening the horizon of available rifamycin-producing genera of actinobacteria beyond known rifamycin producers could provide considerable assistance to efforts to engineer new rifamycins via genetic recombination by taking advantage of the full diversity of the sequence space of rifamycin-synthesizing actinobacteria (13). We have been able to show that rifamycin-synthesizing actinobacteria are indeed found in a species other than A. mediterranei, and to do this, we have used a new approach applying phylogenetic analysis based on polyketide synthase genes to the prediction of synthesis of polyketide antibiotics in a particular chemical class.

New actinobacteria with biopharmaceutical potential have been increasingly isolated from marine habitats (6, 7, 9, 12). We have previously isolated and described Salinispora strains derived from the marine sponge Pseudoceratina clavata (7), members of a genus previously observed only in marine sediments and formally described recently (10, 12). These strains display antibiotic activity and, significantly, β-ketosynthase (KS) gene fragments, indicating the existence of a polyketide synthase (PKS) gene cluster (7). The Salinispora group, as a relatively newly discovered group of Actinobacteria, is one with great applied potential, and its properties and products are only beginning to be explored. Members of this group have been found in habitats as diverse as marine sediments and marine sponges. Its full range of taxonomic diversity is unknown, but at least two species have been identified.

Phylogenetic analysis of KS gene sequences from marine-sponge-derived Salinispora strains indicated that the polyketide synthase gene sequence most closely related to that of Salinispora was the rifamycin B synthase of A. mediterranei, a result not expected from taxonomic phylogenetics using 16S rRNA sequences. The closest relatives to the Salinispora strains on the basis of 16S rRNA sequence analysis were members of the genus Micromonospora. From the KS gene sequence data generated from our sponge-derived Salinispora strains, we predicted that such strains might synthesize rifamycin-like compounds on the basis of the relationship of their KS genes to those of Amycolatopsis. Liquid chromatography-tandem mass spectrometry (LC/MS/MS) analysis was applied to one sponge-derived Salinispora strain to test the hypothesis of rifamycin synthesis. The analysis demonstrated that this Salinispora was indeed able to produce compounds of the rifamycin class, including rifamycin B and rifamycin SV. The availability of the KS gene sequences for both Salinispora and A. mediterranei made the design of a rifamycin-specific PCR primer set possible, and PCR using this primer set was applied to screening of new Salinispora strains for KS genes associated with rifamycin synthesis. Chemical analysis confirmed that strains positive for...
PCR with rifamycin-specific primers were also positive for synthesis of rifamycins. Using the rifamycin-specific primer set also increased the number of PKS-synthesizing strains detected by such PCR screening beyond those detectable with a simple KS-specific primer set. Rifamycins, previously known to be produced only by soil actinobacteria, have now been found to also be produced by marine bacteria—Salinispora isolated from the marine sponge Pseudoceratina clavata. Significantly, Salinispora forms a potential new source of rifamycins and polyketide synthesis gene clusters specific to rifamycin synthesis outside the genus Amycolatopsis. This opens the way to insights from comparative genomic analysis of rifamycin synthesis gene clusters and, for the first time, to gene recombination engineering. Such engineering can exploit the sequence and operon diversity hidden in rifamycin gene clusters from diverse bacteria (5).

MATERIALS AND METHODS

Bacterial strains and culture. Salinispora strains were isolated from the Great Barrier Reef sponge Pseudoceratina clavata and cultured for maintenance purposes on starch-yeast-extract-peptone (SYP) medium with artificial seawater (ASW) at 28°C (7). Amycolatopsis mediterranei (Australian Collection of Microorganisms no. 2548) was maintained on starch-yeast-extract-peptone medium with ASW (7). The GenBank accession numbers resulting from that study are AAT68460 to AAT68464. Using the PHI and PSI-BLAST algorithms at NCBI (7), the GenBank accession numbers were aligned with the PHYLIP package (4). One thousand bootstrap resamplings of neighbor-joining (PROTDIST and NEIGHBOR), parsimony (PROTPARS), and maximum likelihood (PROML) analyses were performed using SEQBOOT and CONSENSE within PHYLIP. A type II KS (ActI of Streptomyces coelicolor) was used as the outgroup in that analysis.

Preparation of extracts for LC/MS/MS analysis. Salinispora M403 was cultured on SYP agar medium with ASW, and A. mediterranei was cultured on SYP agar medium without ASW at 28°C for 4 weeks. The mycelial cell mass was harvested from the medium by scraping using a glass slide, and the mass was weighed. Absolute ethanol was added in sufficient volume to submerge all the mycelia. The mycelium and ethanol mixture was incubated for 90 min at room temperature with rotation. The liquid phase was clarified by centrifugation, and then the extraction of the cell pellet with ethanol was repeated. The combined extracts were filtered using a 0.22-µm-pore-size filter and then stored at −20°C.

LC/MS/MS analysis of extracts. Rifamycin SV sodium salt was purchased from Sigma (St. Louis, MO), and rifamycin B was obtained by Lupin Ltd (Tarakpur, Maharashtra, India). All solvents used were high-performance liquid chromatography (HPLC) grade. Ethanol extracts were evaporated using a freeze drier and then reconstituted initially with methanol and then with water, resulting in a final composition of 45% (vol/vol) methanol-water and representing a 10-fold concentration of the original extracts. The samples were filtered through 0.22-µm filters, and 100-µl aliquots were injected into the column using an autosampler. The separation was carried out using a binary solvent gradient consisting of a 0.1% (wt/vol) aqueous solution of ammonium acetate (A) and a 0.1% (wt/vol) acetonitrile solution of ammonium acetate (B) delivered at 200 µl/min by an Agilent Binary HPLC system with a Cogent Aclarity-C18 (2.0- by 50-mm; 5-µm) HPLC column (Microsoft Technology Corporation, Long Branch, NJ). The separated compounds were detected and analyzed using an API 3000 tandem mass spectrometer equipped with a turbo ion spray interface and supported by Analyst 1.4 software (Applied Biosystems, Foster City, CA). The mass spectrometer was operated in Q1-total ion current mode within the range 200 to 800 atomic mass units. The fragmentation pattern of each molecular ion was obtained by Q3-MS2 negative-ion mode within the 200- to 800-atomic-mass-unit range. The fragmentation of ion 755 (rifamycin B) was monitored for the first 3.2 min of the chromatographic run, and that of ion 698 (rifamycin SV) was monitored for the remainder of the run.

Rifamycin ketosynthase-specific primer set design and amplification. A rifamycin-specific degenerate primer set (Rifr and Rifr) was designed using an alignment of KS sequences of A. mediterranei, Salinispora strains, and other reference actinobacteria. A typical 50-µl reaction mixture contained 180 ng to 1.5 µg of genomic DNA, 200 ng of Rifr (5'-RTGGAYCBBAGACGCGG-3') forward primer, 200 ng of Rifr (5'-RCGGCTCSCGMGCCACCGA-3') reverse primer, 0.2 mM of each deoxynucleoside triphosphate, 5% (vol/vol) dimethyl sulfoxide, and 2.5 µl of AmpI Tag Gold (Roche, CA). The cycling consisted of 94°C for 5 min (activation of AmpliTaq), followed by 30 cycles of denaturation (94°C for 1 min), annealing (56°C for 1 min), and extension (72°C for 1 min), and then amplification was terminated with a final extension (72°C for 10 min). Bands of the expected size (ca. 510 bp) were excised, and then DNA was extracted using NucleoSpin Extract II kits (BD Biosciences, CA). The nucleotide sequences were determined using ABI BigDye terminator version 3.1 with ABI 3730xl sequencing platforms. The resulting sequences were matched with the GenBank database using BLASTX with default parameters (1).

RESULTS

Phylogenetic analysis. The NCBI protein database was searched via PHI and PSI-BLAST using the sequences retrieved from Salinispora by PCR with KS-specific primer sets as queries. The top matched sequences were type I KS domains of rifamycin B synthase of A. mediterranei (more than 80% identity and 90% similarity), followed by KS domains from the genus Streptomyces. The Salinispora homologues shared amino acid sequence motifs characteristic of KS domains that were derived from examination of multiple alignments of type I KS domains. These motifs included the KS consensus amino acid sequences (M/V)DPQQR and TXCSSSLV (including the active-site cysteine) and E(A/G)HGTGT (including the active-site histidine). A rifamycin synthase-specific pattern (VARER) was detected within the KS sequences of both rifamycin and Salinispora KSs, including the unique amino acid valine at position 162 (Fig. 2). In phylogenetic analysis, Salinispora KS
FIG. 2. Amino acid sequence alignment of KS genes from marine sponge *Salinispora* strains SW15, M412, M403, M413, and SW17 and those from the reference bacteria. The arrows indicate sequences used for the rifamycin-specific primer set (Riff and Rifr). The arrowheads point to the active-site cysteine and histidine. The rifamycin KS-specific motif is marked by the open box.
sequences clustered with the KS domains of the rifamycin B synthesis PKS from *A. mediterranei*, as supported by significant (99/85/89% [neighbor-joining/parsimony/maximum likelihood]) bootstrap values. *Salinispora* KS sequences and the rifamycin KS sequences were members of a larger cluster with other actinobacterial KS sequences, a cluster supported by 87/98/94% bootstrap values (Fig. 3). The genus *Micromonospora* is the group most closely related to *Salinispora* according to the 16S rRNA sequences, but this relationship was not found in our analysis of KS sequences. The TS tree did not show any significant relation between *Salinispora* KSs and *Micromonospora* KS (Myc A1). The KS amino acid sequence tree topology, however, was consistent with 16S rRNA trees at a coarser level, since phylum-level groups of *Actinobacteria, Firmicutes* (low-G+C percentage gram-positive bacteria), and *Cyanobacteria* could be distinguished in both cases. The *Salinispora* KS gene amino acid sequence displayed 90% similarity to the amino acid sequence of the RfB gene of *A. mediterranei* but only 74% amino acid similarity to that of the KS gene of *Micromonospora*.

This is consistent with the phylogenetic analysis of KS gene amino acid sequences.

**LC/MS/MS data analysis.** The chromatograms monitored using total ion current mode showed that both the *Salinispora* M403 and *A. mediterranei* extracts contained peaks that eluted with retention times similar to those of the standard compounds rifamycin B and SV. These peaks were absent from the blank extracts. The electrospray negative-ion spectrum of each peak resembled those of the standard rifamycins B and SV (data not shown), confirming the identities of the peaks in *Salinispora* M403 and *A. mediterranei* extracts. To further explore the identities of the peaks present in extracts, the sample and standard chromatograms were monitored using MS2 (fragmentation using the third quadrupole). The fragmentation of the molecular ion 755 (rifamycin B) was monitored for the first 3.2 min of the run, and that of the molecular ion 696.8 (rifamycin SV) was monitored for the remainder of the chromatographic run. The chromatograms and the MS2 spectra of the rifamycin B and SV standards and those of the samples (Fig. 4 and 5) displayed marked similarity, further confirming the identities of rifamycins B and SV in the samples. The presence of an *m/z* 273 naphthofuran fragment, which is considered the structurally most useful signature fragment of rifamycins (19), provides conclusive evidence for the identification of rifamycins in *Salinispora* M403 extracts (Fig. 5). The structure of this fragment is shown in Fig. 1. The fragmentation of rifamycin B (*m/z* 755) occurs initially by the loss of an acetate group to produce rifamycin SV (*m/z* 696.8) and then subsequently by fragmentation of the ansa group to produce the *m/z* 273 fragment, which is observed for both rifamycin B and SV. The observation that the HPLC mass chromatograms and the mass spectra (Fig. 5) obtained for the extracts of *A. mediterranei*, which is known to produce rifamycins B and SV (14, 16), resemble those of *Salinispora* M403 extracts (Fig. 5) further confirms the presence of these compounds in *Salinispora*.

**Rifamycin-specific primers.** Comparison of rifamycin KS gene sequences between *A. mediterranei* and *Salinispora* strains made the design of rifamycin-specific primers possible (Fig. 2). The rifamycin-specific primer set discriminatingly amplified rifamycin KS genes from *A. mediterranei* and *Salinispora* strains, but not from reference strains. The reference strains included strains producing PKS, but not rifamycin, and non-*Salinispora* sponge isolates in which KS genes had been detected using a primer set designed to be specific only for ketosynthase in general. The rifamycin-specific primer set increased the number of PKS-synthesizing strains detected by PCR screening to 9 out of 10 *Salinispora* isolates, while only five strains were detected within this set of 10 isolates when the general KS primer set was used (7).

**DISCUSSION**

We have demonstrated by chemical analysis that a marine-sponge-derived actinobacterium, *Salinispora* M403, can produce both rifamycin B and rifamycin SV. This is consistent with our evidence that five strains of these *Salinispora* isolates possess genes for the ketosynthase enzyme, which forms the key part of the polyketide synthase enzyme complex. This is also consistent with our finding that these genes are most closely related to those of *A. mediterranei*, which is a known producer of rifamycins B and SV. While they are most closely related to rifamycin KS genes of *A. mediterranei*, the KS genes of *Salinispora* strains are not identical, suggesting potential genetic diversity among rifamycin synthesis genes from different genera. Members of the *Salinispora* group of marine actinobacteria thus constitute a potential new source of antibiotics and other biopharmaceutically active compounds in the rifamycin class and also a possible source of new rifamycin synthesis genes. The only known source of rifamycins supported by published evidence for rifamycin production is *A. mediterranei* (14, 16), and this species has been subjected to considerable genetic analysis of its polyketide synthase gene cluster (2, 15). Other sources have been reported in documentation accompanying culture collection databases (e.g., within cultures held in the American Type Culture Collection, *Streptomyces albovinaceus* ATCC 12951, *Micromonospora lacastris* ATCC 21974, ATCC 21975, and *Micromonospora chalcea* ATCC 21994, all without any supporting chemical data accompanying annotations). Except for *Micromonospora lacastris* ATCC 21975 (8), 16S rRNA sequence data do not appear to have been released for these culture collection strains reported to produce rifamycins, so that their taxonomic status remains uncertain within the context of modern molecular phylogenetics of actinobacteria. *Salinispora* is a taxonomically well-defined genus level group, at least one member of which produces rifamycins as confirmed by chemical analysis and supported by gene sequence phylogenetics. It thus provides a second model, in addition to the existing *A. mediterranei* model, for the study of rifamycin production and its genomics and genetics, which may prove useful for comparative purposes (13). The expansion of rifamycin polyketide synthase gene sequences by the inclusion of *Salinispora*, as well as *Amycolatopsis*, has also made a rifamycin-specific primer set possible. We found that application of these high-specificity primers could increase the detection rate of PCR screening for certain KS genes. However, even these highly specific primers are not free from PCR bias, because when we repeated the chemical analysis of the remaining *Salinispora* strains, we detected production of rifamycins from all 10 *Salinispora* strains.

Thus, the *Salinispora* group of actinobacteria represents the first recorded source of rifamycins from marine bacteria. To discover this source, we used an approach of applying phylo-
FIG. 3. Maximum likelihood tree generated using amino acid sequences of KS from marine sponge *Salinispora* strains (boldface) and reference organisms for type I KS. Significant bootstrap values (>50%) are indicated at each node in the following order from top to bottom: neighbor-joining, parsimony, and maximum likelihood. The scale bar represents 0.1 amino acid substitution per site. *Streptomyces coelicolor* ActI (a type II KS) was used as the outgroup.
FIG. 4. Chromatograms and fragmentation patterns of standard compounds rifamycin B and rifamycin SV monitored using MS2 negative-ion mode. (A) Chromatogram of rifamycin B. (B) Fragmentation pattern of the rifamycin B peak. (C) Chromatogram of rifamycin SV and rifamycin S (a breakdown product of rifamycin SV), which is apparent at 6.93 min. (D) Fragmentation pattern of the rifamycin SV peak.
FIG. 5. Chromatograms and fragmentation patterns of *A. mediterranei* and *Salinispora* M403 extracts monitored using MS2 negative-ion mode. (A and D) Chromatograms of *A. mediterranei* (A) and *Salinispora* M403 (D) extracts. (B and E) Fragmentation patterns of *A. mediterranei* (B) and *Salinispora* M403 (E) rifamycin B peaks. (C and F) Fragmentation patterns of *A. mediterranei* (C) and *Salinispora* M403 (F) rifamycin SV peaks.
genetic analysis based on polyketide synthase genes to the prediction of synthesis of polyketide antibiotics in a particular chemical class. Such molecular prediction could conceivably preclude the need for time-consuming nontargeted chemical assay. Significantly for biotechnology, the discovery of rifamycin synthesis by Salinispora strains implies that they are a potential new source, outside the genus Amycolatopsis, of rifamycins and of polyketide synthesis gene clusters specific to rifamycin synthesis. Comparative genomic analysis of rifamycin synthesis gene clusters using strains from the two distinct genus level taxa known to harbor strains producing rifamycins will now be possible. The resulting bank of variation within rifamycin genes may become available for gene recombinatorial engineering and design of synthetic genes on the basis of conserved rifamycin-associated motifs (11). With the availability of new PKS gene clusters and regulatory genes associated with rifamycin synthesis from Salinispora, an advantage could be derived from the study of the full diversity of the sequence space of rifamycin-synthetic genes and associated regulatory genes. If the efficiency of rifamycin production varies among strains, as occurs with synthesis of erythromycin (13), then the availability from diverse bacteria of gene clusters associated with the synthesis and regulation of rifamycin could conceivably be useful for designing increased yields from rifamycin-synthesizing strains. The highly conserved but nonidentical genes for the synthesis of antibiotics in the same rifamycin class could also contribute to research programs aiming to produce novel recombinatorial antibiotics (13).

ACKNOWLEDGMENTS

Lupin Ltd. kindly provided rifamycin B.

T.K.K. was supported by a University of Queensland Graduate School Scholarship. Research in the laboratory of J.A.F. was funded by the Australian Research Council and the University of Queensland School of Molecular and Microbial Sciences and Faculty of Biological and Chemical Sciences.

REFERENCES