Production of Hygromycin A Analogs in *Streptomyces hygroscopicus* NRRL 2388 through Identification and Manipulation of the Biosynthetic Gene Cluster

Nadaraj Palaniappan,¹ Sloan Ayers,² Shuchi Gupta,¹ El-Sayed Habib,² and Kevin A. Reynolds¹,*
¹Department of Chemistry
Portland State University
P.O. Box 751
Portland, Oregon 97207
²Departments of Medicinal Chemistry
Virginia Commonwealth University
Richmond, Virginia 23219

Summary

Hygromycin A, an antibiotic produced by *Streptomyces hygroscopicus* NRRL 2388, offers a distinct carbon skeleton structure for development of antibacterial agents targeting the bacterial ribosomal peptidyl transferase. A 31.5 kb genomic DNA region covering the hygromycin A biosynthetic gene cluster has been identified, cloned, and sequenced. The hygromycin gene cluster has 29 ORFs which can be assigned to hygromycin A resistance as well as regulation and biosynthesis of the three key moieties of hygromycin A (5-dehydro-α-L-lucurafuranose, (E)-3-(3,4-dihydroxyphenyl)-2-methylacrylic acid, and 2L-2-amino-2-deoxy-4,5-0-methylene-neo-inositol). The predicted Hyg26 protein has sequence homology to short-chain alcohol dehydrogenases and is assigned to the final step in production of the 5-dehydro-α-L-lucurafuranose, catalyzing the reduction of α-L-lucurafuranose. A hyg26 mutant strain was generated and shown to produce no hygromycin A but 5′-dihydroxyhygromycin A, 5′-di-hydromethoxyhygromycin A, and a 5′-dihydroyhygromycin A product lacking the aminocyclitol moiety. To the best of our knowledge, these shunt metabolites of biosynthetic pathway intermediates have not previously been identified. They provide insight into the ordering of the multiple unusual steps which compromise the convergent hygromycin A biosynthetic pathway.

Introduction

Hygromycin A (compound 1) (Figure 1) is an antibiotic first isolated from the fermentation broth of several strains of *Streptomyces hygroscopicus* in 1953 [1]. A second structurally unrelated antibiotic, the aminoglycoside hygromycin B, was later isolated from *S. hygroscopicus* [2]. Early studies demonstrated that hygromycin A had a relatively broad spectrum of activity against gram-positive and -negative bacteria [1, 3]. Almost three decades later, Guerrero and Modolell demonstrated that the mode of action was inhibition of the ribosomal peptidyl transferase activity. Initial studies also demonstrated that hygromycin A (1) blocked the binding of either chloramphenicol or lincomycin to the ribosomes [4] and (2) bound more tightly than chloramphenicol. More recent footprinting experiments have shown that macrodilides only block binding of hygromycin A to the ribosome if they contain a mycarose unit [5]. Crystallographic evidence indicates that in such macrodilides, the C5-disaccharide group extends from the polypeptide exit channel into peptidyl transferase center [6]. Hygromycin A is not a macrolide and thus offers a distinct carbon skeleton and binding mode to other antibiotics that target the bacterial ribosome. As such, it represents a promising starting point for generating new antibiotics to treat infections with drug-resistant pathogens.

Hygromycin A has also been reported to have additional activities and potential applications. It exhibits hemagglutination inactivation activity and high antireponemal activity [7, 8], leading to the possible application of hygromycin A-related compounds for the treatment of swine dysentery, a severe mucohemorrhagic disease thought to be caused by *Serpulina (Treponema) hyodysenteriae* [8, 9]. Hygromycin A has also been reported to possess an immunosuppressant activity in the mixed poor lymphocyte reaction but does not work via suppression of interleukin 2 production [10]. Most recently, methoxyhygromycin A (compound 2) (Figure 1), an analog of hygromycin A produced in the same fermentation broth, has been shown to have herbicidal activity and has led to the suggestion that it could be developed as a biological agent for weed control [11].

Semisynthetic programs based on hygromycin A and its attractive biological properties have been reported [12-14]. This work has led to the synthesis of over 100 analogs and determination of their activity, both in terms of MICs for *Serpulina hyodysenteriae* and their ability to inhibit protein synthesis in an *E. coli* cell extract. The resulting structure activity relationship (SAR) has revealed that the unusual aminocyclitol moiety is an important component for the antibacterial activity, while the 5-dehydro-α-L-lucurafuranose moiety is not essential and can be replaced with hydrophobic allyl group. Reduction in the antibacterial activity is also observed with replacement in methyl group of central (E)-3-(3,4-dihydroxyphenyl)-2-methylacrylic acid moiety with propyl, allyl, or hydrogen [14, 15]. For the most part, these structural analogs were prepared by a semisynthetic method with hygromycin A as a starting point. A total synthesis of hygromycin A and C2′-epi-hygromycin A has also been reported [12, 16]. Several multistep syntheses of the 2L-2-amino-2-deoxy-4,5-O-methylene-neo-inositol moiety have also been reported [17]. The most recent synthesis of this was enantioselective and accomplished in 14 steps with an overall 12% yield [18].

We have sought a complementary approach of deciphering the biosynthetic process with a long-term aim of using this as an economical means for generating hygromycin analogs for further development. In our preliminary work, we determined the biosynthetic origins of the three unusual and structurally distinct moieties of hygromycin A. Mannose was shown to provide the 5-dehydro-α-L-lucurafuranose moiety, 4-hydroxybenzoic acid and propionic acid the central (E)-3-(3,4-dihydroxyphenyl)-2-methylacrylic acid moiety, and glucose and...
methionine the 2L-2-amino-2-deoxy-4,5-O-methylene-neo-inositol moiety [19]. A convergent biosynthetic pathway was proposed (Figure 2) from these observations, and we now report how this formed the basis for a PCR approach that has led to the identification of the hygromycin A biosynthetic gene cluster of *S. hygroscopicus* NRRL 2388. The gene cluster has been cloned, sequenced, and analyzed, and putative assignments for the majority of corresponding gene products in the biosynthesis of the three structural moieties of hygromycin A have been proposed. Furthermore, manipulation of the gene cluster has been shown to give rise to hygromycin A analogs and shunt products that have not previously been identified. These findings provide important insights into the order of the biosynthetic steps and are a promising step toward the long-term objective of this work.

Results and Discussion

Cloning and Sequencing of the Hygromycin A Biosynthetic Gene Cluster

Previous studies on hygromycin A have suggested a convergent biosynthesis from 5-dehydro-β-L-fucofuranose, (E)-3-(3,4-dihydroxyphenyl)-2-methylacrylic acid, and 2L-2-amino-2-deoxy-4,5-O-methylene-neo-inositol [19]. The 5-dehydro-β-L-fucofuranose moiety was shown to be derived from mannose, and a pathway proceeding from a nucleoside diphosphate (NDP) activated mannose, through NDP-4-keto-6-deoxymannose and...
NDP-L-fucose, was proposed. Genes encoding proteins with homology to GDP-D-mannose-4,6-dehydratase (MDH) have been identified in the nystatin, candicidin, with homology to GDP-D-mannose-4,6-dehydratase and NDP-L-fucose, was proposed. Genes encoding proteins with sequence homology to putative MDH enzymes. The partial gene product encodes a protein with significant structural similarities to hygromycin A (Figure 1), also contains a gene encoding a putative MDH. A BLAST search for similar gene products in two reported genome sequences from S. coelicolor and S. avermitilis failed to produce sequences encoding primary metabolic proteins with significant homology. These observations suggested a gene encoding a putative MDH as a target for identification of the hygromycin A biosynthetic gene cluster of S. hygroscopicus.

Accordingly, a pair of degenerate primers based on the highly conserved motifs observed by creating an alignment of predicted MDHs was used to amplify a portion of hyg5 from genomic DNA of the hygromycin A producer S. hygroscopicus NRRL 2388. Sequencing confirmed that the PCR product encoded a protein with homology to putative MDH enzymes. The partial hyg5 fragment was then used to screen a cosmid library of S. hygroscopicus NRRL 2388 to identify hygromycin A biosynthetic gene cluster. Cosmid clone 2F1 was identified through screening the first 600 cosmid clones and sequence analysis suggested that only a portion (15 kb) of the predicted biosynthetic gene cluster was present. A PCR fragment of hyg14 was obtained with this cosmid clone as a template and used to further screen the library, resulting in the identification of three more overlapping cosmid clones, 12G10, 17E3, and 15A10, respectively. The entire putative hygromycin A biosynthetic gene cluster was obtained by shotgun sequencing 17E3 cosmid and primer walking 7 kb of cosmid 15A10 cosmid clone. The sequence of the entire gene cluster and the proposed function for individual ORFs is summarized in Table 1, and nucleotide sequence submitted to GenBank with accession number DQ314862.

A contiguous 40 kb DNA region encoding the putative hygromycin biosynthetic gene cluster was thus sequenced adequately on both strands with these three cosmids (2F1, 17E3, and 15A10) and analyzed for putative open reading frames (ORFs) with DS Gene software (Accelrys) and Frame program. A total of 29 ORFs, designated hyg1–29, were aligned with homologous sequences of GenBank by using BLAST programs. The predicted functions of these hyg gene products and their corresponding homologs are listed in Table 1. Genes adjacent to hyg1 encoded a hypothetical protein and a putative glucose dehydrogenase, respectively, while those adjacent to hyg29 encoded a putative pyruvate dehydrogenase and transcriptional regulator. These genes flanking the 29 hyg ORFs were highly homologous to genes identified in the genome sequences of S. coelicolor and S. avermitilis, could not readily be assigned a dedicated function in the biosynthesis of hygromycin A, and likely are involved in other cellular processes.

Gene Products Putatively Required for 5-Dehydro-L-Fucofuranose Biosynthesis

Analysis of the hyg cluster indicated the presence of several putative genes believed to be involved in the proposed biosynthetic pathway to the 5-dehydro-L-fucofuranose moiety from NDP-mannose (Figure 2). The hyg5 gene used to identify the biosynthetic gene cluster encodes a putative MDH that would catalyze the first step in this process, the conversion of an activated NDP-mannose into NDP-4-keto-6-deoxymano- nose. The subsequent step was proposed to be a conversion to NDP-L-fucose (involving epimerization at the C-3 and C-5 positions of the hexose ring and an NADPH-dependent reduction at the C-4 position) and is likely catalyzed by the hyg23 gene product, which has clear homology (>60% identity) to L-fucose synthetases. The proposed next step, conversion from the pyranose to the furanose form, was proposed to follow a similar mechanism to that established for the UDP-galactopyranosyltransferase. None of the hyg genes encoded proteins with sequence homology to this protein and no clear candidate for this step can be identified. The hyg20 gene product encodes proteins with homology to transglucosylases in databases and may catalyze this step. Interestingly, Hyg20 has homology to Ata16 protein, which is encoded by the antibiotic A201A gene cluster. The structure of antibiotic A201A suggests a pyranose-furanose ring might also be required in the biosynthetic pathway. The final step in the predicted pathway to 5-dehydrofucofuranose is an
oxidation and is assigned to the hyg26 gene product that has homology to a family of short-chain dehydrogenases (Table 1). As described below, subsequent genetic experiments have confirmed this assignment. A glycosyltransferase encoded by hyg16 is expected to be responsible for formation of the glycosidic linkage between the 5-dihydro-α-L-fucofuranose and the 4-hydroxy group of the (E)-3-(3,4-dihydroxyphenyl)-2-methylacrylic acid moiety (the juncture in the biosynthetic process in which this link is formed has yet to be determined). A similar glycosidic linkage is required in the biosynthesis of antibiotic A201A and may be carried out by the Ata5 protein, which has 63% identity to the Hyg16 (Table 1).

<table>
<thead>
<tr>
<th>Gene</th>
<th>Amino Acids</th>
<th>Proposed Function</th>
<th>Homologous Protein in Database</th>
<th>Identity/Similarity</th>
<th>Accession Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>hyg1</td>
<td>375</td>
<td>AfsR regulatory protein</td>
<td>afsR (S. coelicolor)</td>
<td>37/52</td>
<td>BAA14186</td>
</tr>
<tr>
<td>hyg2</td>
<td>399</td>
<td>p-hydroxybenzoate hydrolase</td>
<td>4-hydroxybenzoate hydrolase (B. japonicum)</td>
<td>51/66</td>
<td>NP773060</td>
</tr>
<tr>
<td>hyg3</td>
<td>323</td>
<td>Regulatory protein</td>
<td>putative regulatory protein (S. griseus)</td>
<td>42/56</td>
<td>CAAB8515</td>
</tr>
<tr>
<td>hyg4</td>
<td>173</td>
<td>Chorismate lyase</td>
<td>Chorismate pyruvate lyase (Methanopyrus kandleri)</td>
<td>28/44</td>
<td>NP613406</td>
</tr>
<tr>
<td>hyg5</td>
<td>352</td>
<td>Mannose dehydratase</td>
<td>GDP-D-mannose dehydratase (Kineococcus radiotolerans)</td>
<td>67/77</td>
<td>ZP00227152</td>
</tr>
<tr>
<td>hyg6</td>
<td>249</td>
<td>Methyltransferase</td>
<td>Ata11 protein (Saccharothrix mutabilis subsp. capreolus)</td>
<td>31/42</td>
<td>CAD62205</td>
</tr>
<tr>
<td>hyg7</td>
<td>579</td>
<td>Unknown</td>
<td>hypothetical protein (M. avium)</td>
<td>51/67</td>
<td>NP963248</td>
</tr>
<tr>
<td>hyg8</td>
<td>411</td>
<td>Aminotransferase</td>
<td>aminotransferase (Prochlorococcus marinus)</td>
<td>33/52</td>
<td>NP939396</td>
</tr>
<tr>
<td>hyg9</td>
<td>83</td>
<td>ACP</td>
<td>AtaPKS2 protein (Saccharothrix mutabilis subsp. capreolus)</td>
<td>34/55</td>
<td>CAD62191</td>
</tr>
<tr>
<td>hyg10</td>
<td>378</td>
<td>β-ketoacyl synthase 1</td>
<td>AtaPKS3 protein (Saccharothrix mutabilis subsp. capreolus)</td>
<td>54/64</td>
<td>CAD62192</td>
</tr>
<tr>
<td>hyg11</td>
<td>252</td>
<td>Unknown</td>
<td>AtaPKS4 protein (Saccharothrix mutabilis subsp. capreolus)</td>
<td>48/59</td>
<td>CAD62193</td>
</tr>
<tr>
<td>hyg12</td>
<td>451</td>
<td>CoA-ligase</td>
<td>Ata18 protein (Saccharothrix mutabilis subsp. capreolus)</td>
<td>48/61</td>
<td>CAD62194</td>
</tr>
<tr>
<td>hyg13</td>
<td>96</td>
<td>ACP</td>
<td>Ata19 protein (Saccharothrix mutabilis subsp. capreolus)</td>
<td>43/64</td>
<td>CAD62195</td>
</tr>
<tr>
<td>hyg14</td>
<td>347</td>
<td>3-hydroxyacyl ACP dehydratase</td>
<td>Ata22 protein (Saccharothrix mutabilis subsp. capreolus)</td>
<td>43/59</td>
<td>CAD62196</td>
</tr>
<tr>
<td>hyg15</td>
<td>256</td>
<td>3-ketoacyl ACP reductase</td>
<td>Ata4 protein (Saccharothrix mutabilis subsp. capreolus)</td>
<td>54/67</td>
<td>CAD62198</td>
</tr>
<tr>
<td>hyg16</td>
<td>431</td>
<td>Glycosyltransferase</td>
<td>Ata5 protein (Saccharothrix mutabilis subsp. capreolus)</td>
<td>63/74</td>
<td>CAD62199</td>
</tr>
<tr>
<td>hyg17</td>
<td>330</td>
<td>myo-inositol dehydrogenase</td>
<td>hypothetical protein (M. avium)</td>
<td>41/56</td>
<td>NP962413</td>
</tr>
<tr>
<td>hyg18</td>
<td>363</td>
<td>myo-inositol-1-phosphate synthase</td>
<td>Myo-inositol-1-phosphate synthase (Thermobifida fusca)</td>
<td>78/87</td>
<td>ZP00291606</td>
</tr>
<tr>
<td>hyg19</td>
<td>416</td>
<td>Transmembrane protein</td>
<td>Ata9 protein (Saccharothrix mutabilis subsp. capreolus)</td>
<td>50/63</td>
<td>CAD62203</td>
</tr>
<tr>
<td>hyg20</td>
<td>370</td>
<td>Transglucosylase</td>
<td>Ata16 protein (Saccharothrix mutabilis subsp. capreolus)</td>
<td>60/71</td>
<td>CAD62189</td>
</tr>
<tr>
<td>hyg21</td>
<td>213</td>
<td>Phosphotransferase</td>
<td>Ard2 protein (Saccharothrix mutabilis subsp. capreolus)</td>
<td>56/66</td>
<td>CAD62197</td>
</tr>
<tr>
<td>hyg22</td>
<td>391</td>
<td>Acetyltransferase</td>
<td>AtaPKS1 protein (Saccharothrix mutabilis subsp. capreolus)</td>
<td>45/52</td>
<td>CAD27643</td>
</tr>
<tr>
<td>hyg23</td>
<td>325</td>
<td>Fucose synthase</td>
<td>EpiA (M. avium)</td>
<td>62/74</td>
<td>NP960166</td>
</tr>
<tr>
<td>hyg24</td>
<td>287</td>
<td>Unknown</td>
<td>hypothetical protein (M. avium)</td>
<td>43/58</td>
<td>NP962413</td>
</tr>
<tr>
<td>hyg25</td>
<td>213</td>
<td>myo-inositol-1-phosphatase</td>
<td>sstD (S. griseus)</td>
<td>44/59</td>
<td>CAAB7013</td>
</tr>
<tr>
<td>hyg26</td>
<td>271</td>
<td>Short chain dehydrogenase</td>
<td>hypothetical protein (M. avium)</td>
<td>53/70</td>
<td>NP961344</td>
</tr>
<tr>
<td>hyg27</td>
<td>401</td>
<td>DAHPS synthase</td>
<td>DAHP synthase (Actinosynnema pretiosum)</td>
<td>62/71</td>
<td>AAC13561</td>
</tr>
<tr>
<td>hyg28</td>
<td>570</td>
<td>ABC transporter</td>
<td>Putative ABC transporter (S. avermitilis)</td>
<td>78/86</td>
<td>NP822846</td>
</tr>
<tr>
<td>hyg29</td>
<td>378</td>
<td>Methyltransferase</td>
<td>carboxymethyltransferase (Thermobifida fusca)</td>
<td>45/59</td>
<td>ZP00293268</td>
</tr>
</tbody>
</table>

Table 1. Gene Annotation of hyg Gene Cluster and Comparison with Homologous Proteins in the Public Database

Gene Products Putatively Required for (E)-3-(3,4-Dihydroxyphenyl)-2-Methylacrylic Acid Biosynthesis

Based on the results from previous biosynthetic studies with 13C labeled precursors, we had proposed that this central (E)-3-(3,4-dihydroxyphenyl)-2-methylacrylic acid moiety was derived from a condensation of methylmalonyl-ACP with 3,4-dihydroxybenzoyl CoA or 4-hydroxybenzoyl CoA [19]. The 3-keto group of the resulting 3-(3,4 dihydroxyphenyl)-3-oxopropanoyl CoA (ACP) would then be processed by reduction and dehydration steps in a manner similar to that observed for polyketide biosynthetic processes (Figure 2). The gene products of hyg9–15, hyg27, hyg4, and hyg22 can be assigned
Hygromycin A Biosynthetic Gene Cluster

Putative roles in this process. As antibiotic A201A has the same structural moiety as hygromycin A, with the exception of the 3-hydroxy substituent, similar enzymes are anticipated to be required for its biosynthesis. At the time this project was initiated, a partial sequence of the antibiotic A201A gene sequence had been reported. Analysis showed that ard1 and ard2 were resistance determinants, while ataP3, ataP5, ataP4, and ataP7 were likely involved in formation of the N,N-di-methyl-3-amino-3′-deoxyadenosine moiety, which is not present in hygromycin A (the full sequence of the ataPKS1 gene was also reported). Additional sequence information from this cluster has now been deposited and is consistent with our prediction; clear homologs of hyg9–15 and hyg22 are observed (Table 1).

The pathway to this moiety is proposed to start with 4-hydroxybenzoic acid, derived from chorismate via the enzyme chorismate lyase. A putative chorismate lyase is encoded by hyg4 (Table 1). The chorismate required for this enzyme reaction presumably is generated by the shikimate pathway. The committed step of this pathway, which generates aromatic amino acids and other primary metabolites in plants and bacteria [29], is catalyzed by 3-deoxy-D-arabino-heptulosonate 7-phosphate (DAHP) synthase. DAHP synthase is subject to feedback inhibition by aromatic acids by an allosteric mechanism [30]. Genes encoding homologs of DAHP synthase are often found in antibiotic biosynthetic gene clusters where a shikimate metabolite is used as a biosynthetic precursor. Presumably, these homologs are not inhibited by aromatic amino acids and help ensure an adequate supply of shikimate/chorismate during secondary metabolism [31]. Consistent with these observations, a gene (hyg27) encoding a DAHP synthase homolog is noted in the hygromycin biosynthetic gene cluster. The hyg2 gene product encodes a protein with homology to 4-hydroxybenzoate hydroxylase and is assigned a role in the formation of 3,4-dihydroxybenzoic acid. It is not clear if the hydroxylation occurs directly on 4-hydroxybenzoic acid or at a later juncture in the biosynthesis of the (E)-3-(3,4-dihydroxyphenyl)-2-methylacrylic acid moiety. This hydroxylation is not required in antibiotic A201A biosynthesis, and our analysis revealed that there is no Hyg2 homolog encoded by the corresponding biosynthetic gene cluster.

In the following step, the 3,4-dihydroxybenzoic acid (or 4-hydroxybenzoic acid if the hydroxylation occurs at a later step in the pathway) is presumably activated by conversion to a thioester. The hyg12 gene product is assigned to this step based on its homology to a family of CoA-ligases (Table 1). This enzyme may catalyze formation of either a coenzyme A or an acyl carrier protein (ACP) thioester; the predicted sequences of the hyg9 and hyg13 gene products are ACP homologs. One of these ACPs is presumably required for making methylmalonyl ACP for condensation with the activated 3,4-di-hydroxybenzoyl thioester. The other ACP is also likely involved in one of the subsequent steps. The sequence analysis revealed that hyg22 encodes a putative acyl transferase (AT) with some low sequence similarity to AT domains in type I modular PKSs and may catalyze the methylmalonyl CoA-methylmalonyl ACP interconversion. The presumed AT atAPKS1 encoded by the antibiotic A201A biosynthetic gene cluster has sequence homology to both Hyg22 and to AT domains in modular type I polyketide synthase (PKS) and likely has a similar role. Sequence analysis of both Hyg21 and the atAPKS1 revealed conserved sequence motifs characteristic of methylmalonyl-CoA-specific rather than malonyl CoA-specific ATs [32–34]. In contrast, hyg10 (and the homolog from the antibiotic A201A cluster) encodes a β-ketoacyl synthase (KS) with low homology to discrete KS proteins in type II PKS, but not KS domains in modular type I PKSs. It is presumed that this KS protein, whose sequence is quite distinct from other known KS proteins, is required for catalyzing the decarboxylative condensation reaction with methylmalonyl ACP. A multiple sequence alignment of the predicted Hyg10 and AtaPKS3 with the type II PKS KS proteins revealed the presence of the two highly conserved active site histidines [35], which are presumably required for catalyzing the methylmalonyl ACP decarboxylation. However, in both Hyg10 and AtaPKS3, a serine residue is observed in place of the highly conserved nucleophile cysteine residue (required for formation of the acyl thioester intermediate in the catalytic cycle of these enzymes). Although serine acts as the active site nucleophile in ATs for generating enzyme bound acyl ester intermediates [35, 36], we are unaware of such a role for this residue catalyzing C-C bond formation in a β-ketoacyl synthase. A naturally occurring KS domain containing serine at the active site in place of cysteine has been reported in the loading module of the pimaricin PKS cluster of S. natalensis [37] but most likely catalyzes a decarboxylation (C-C bond cleavage) in the same way as KS-O domains [38]. The hyg11 gene product exhibited 48% identity to atAPKS4 of antibiotic A201A gene cluster and very low amino acid identity to several putative type II KS proteins in the database. The highly conserved catalytic triad of KS proteins and domains was not observed in either Hyg11 or A201A, and no clear role can be assigned based on the sequence analysis.

The 3-(3,4-dihydrophenyl)-3-hydroxy-2-methyl-propiionyl ACP proposed to be generated by Hyg10 may be converted to (E)-3-(3,4-dihydroxyphenyl)-2-methylacryl ACP by the action of a 3-hydroxylacyl ACP dehydratase (Hyg14) and 3-ketoacyl ACP reductase (Hyg15) (Figure 2). Similar roles can be envisioned for the ata2 and ata4 gene products (with 59% and 67% homology to Hyg14 and Hyg15, respectively), from antibiotic A201A gene cluster.

Overall, these analyses reveal the presence of a set of homologous genes in the both hygromycin A and A201A gene clusters, which can be assigned to the synthesis of the (E)-3-(3,4-dihydroxyphenyl)-2-methylacrylic acid (hygromycin) and (E)-3-(4-hydroxyphenyl)-2-methylacrylic acid (antibiotic A201A) moiety. Many aspects of this proposed process appear unusual, including the use of a discrete AT to load methylmalonyl CoA and a KS protein with an active site serine residue. Nonetheless, the exact role of these proteins and the specific ordering of the biosynthetic steps remains to be determined.

Gene Products Putatively Required for 2L-2-Amino-2-Deoxy-4,5-O-Methylene-Neo-Inositol Biosynthesis

Previously, we had shown that the aminocycloitol portion of hygromycin A was derived from glucose and that the
incorporation pattern for D-[1,2-\textsuperscript{13}C\textsubscript{2}]glucose was consistent with a proposed pathway leading from glucose 6-phosphate through a \textit{myo}-inositol intermediate (Figure 2). The first two steps of this pathway would require a \textit{myo}-inositol-1-phosphate synthase (hyg18 gene product) and \textit{myo}-inositol phosphatase (hyg25 gene product). We have previously noted that these two steps are also required in the biosynthetic pathway that generates the \textit{scylo}-inosamine-derived moiety of streptomycin moiety. StsD a putative \textit{myo}-inositol phosphatase involved in the streptomycin biosynthetic pathway has 44% identity to Hyg25 (Table 1). The pathways diverge after \textit{myo}-inositol, with a proposed oxidation of C5 in the case of hygromycin A. A putative inositol dehydrogenase (encoded by \textit{hyg17}) is likely responsible for catalyzing this step. A subsequent transamination catalyzed by the hyg8 gene product is then proposed to provide 2L-2-amino-2-deoxy-\textit{neo}-inositol [19]. The Hyg8 sequence shows homology to class III pyridoxal-phosphate-dependent aminotransferases and low homology to L-glutamine:\textit{scylo}-inosose aminotransferase StsC, which catalyzes a similar reaction in streptomycin biosynthesis [39].

One of the intriguing features of the aminocyclitol portion of hygromycin A is the presence of a C-4 and C-5 methylene bridge. At which stage in the aminocyclitol biosynthetic process the methylene bridge is formed remains to be determined. In a previous biosynthetic study, we have shown labeling of this methylene carbon of hygromycin by \textit{L}-[\textit{methyl}-\textsuperscript{13}C\textsubscript{1}]methionine, consistent with a pathway utilizing \textit{S}-adenosylmethionine and a methyl transferase (hyg6 or hyg29 gene product). The structure of methoxyhygromycin (Figure 1) suggests that the methyl group may be transferred onto the C5 hydroxy group and that the penultimate intermediate in this pathway would be 2L-2-amino-2-deoxy-5-O-methyl-\textit{neo}-inositol (Figure 2). Formation of the final 2L-2-amino-2-deoxy-4,5-O-methylene-\textit{neo}-inositol would require an oxidation and represents a highly unusual conversion in sugar biochemistry. Examination of the hygromycin biosynthetic gene cluster does not reveal a clear candidate protein for catalysis of this step. If formation of the methylene bridge occurs before formation of the amide bond of hygromycin A, then methoxyhygromycin (2), which is reportedly less active than hygromycin [7] and observed in all fermentations of the \textit{S. hygroscopicus} strain, may represent a shunt metabolite. Alternatively, methoxyhygromycin may be the penultimate intermediate with methylene bridge formation representing the last step in the biosynthetic pathway.

Resistance Genes
The presence of one or more resistance genes within actinomycete biosynthetic gene clusters is well documented [40]. Analysis of the \textit{hyg} gene cluster revealed several candidates genes whose predicted products, a methyltransferase (Hyg8 or Hyg29), a transmembrane protein (Hyg19), a phosphotransferase (Hyg21), and an ABC transporter (Hyg28), are likely to be involved in providing resistance to hygromycin A.

It has been shown that antibacterial activity of hygromycin A involves inhibition of the ribosomal peptidyl transferase, with a binding site on the large ribosomal subunit closely related to that observed for chloramphenicol and lincomycin [4]. We speculate that methylation of the nucleotide at the binding site of the \textit{S. hygroscopicus} ribosome by a methyltransferase would provide resistance to hygromycin A. Ribosomal methylation is a well-established mechanism of resistance to this class of antibiotics, and genes encoding the appropriate methyl transferases are often located within the antibiotic biosynthetic gene cluster [41-43].

The \textit{hyg21} gene product is predicted to be a phosphotransferase and has high amino acid sequence similarity with \textit{ard2} gene product from the A201A gene cluster (66% homology). It has been shown that \textit{Ard2} protein catalyzes an ATP-dependent phosphorylation of the C2 hydroxyl group in the furanose moiety of antibiotic A201A, thereby inactivating the aminonucleoside antibiotic [44]. Hyg21 may provide resistance to hygromycin A by a similar mechanism. It has previously been established that hygromycin A is a substrate of Acr A/B efflux pump in \textit{E. coli} and that this is the major cause for its ineffectiveness against enteric gram-negative bacteria such as \textit{E. coli} and \textit{Salmonella} where this efflux pump is widespread [14]. Furthermore, Hyg19 and Hyg28 have amino acid sequence homology to Ata9 (65% similarity) and \textit{Ard1} (86% similarity), and the ability of \textit{Ard1}, a member of the ABC transporter superfamily, to provide resistance to antibiotic A201A has previously been established [45]. Therefore, we propose that the Hyg19 (a transmembrane protein) and Hyg28 (an ABC transporter) may provide an additional layer of resistance to \textit{S. hygroscopicus} by catalyzing hygromycin A efflux.

Regulatory Genes
Comparison of \textit{hyg} genes with the databases revealed two genes that likely regulate hygromycin A biosynthesis. The \textit{hyg1} gene encodes a protein with 52% sequence similarity to the AfsR transcriptional regulatory protein of \textit{S. coelicolor} A3(2). A clear helix-turn-helix motif can be identified at the N-terminal region of Hyg1, which is proposed to be a transcriptional activator for the pathway. A similar role is envisioned for Hyg3, which has amino acid sequence homology to the StrR, a pathway-specific activator that regulates streptomycin biosynthesis in \textit{S. griseus} [46].

Production of 5′-Dihydrohygromycin Analogs
A genetic experiment was used to confirm that the \textit{hyg} biosynthetic gene cluster is responsible for hygromycin A biosynthesis. Allelic replacement of the \textit{hyg26} gene in \textit{S. hygroscopicus} by \textit{aac(3)IV} resistance marker (confering apramycin resistance) and \textit{oriT} led to the SCH30 mutant. As described above, Hyg26 has homology to a family of short-chain dehydrogenases and is proposed to catalyze the final oxidative step in the predicted pathway to 5-dehydrofucarofuranose moiety of hygromycin A.

As shown in Figure 4B, HPLC analyses of \textit{S. hygroscopicus} NRRL 2388 reveal the clear presence of hygromycin A (1) and methoxyhygromycin A (2) (the 5′ epi-meric forms of these two compounds appear as smaller shoulder peaks at a slightly earlier retention time). Production levels under the fermentation conditions used were typically 350 mg/l methoxyhygromycin A and 880 mg/l hygromycin A. In contrast, the SCH30 mutant generates no detectable levels of hygromycin A or...
Hygromycin A Biosynthetic Gene Cluster

methoxyhygromycin A but generates three new peaks (3, 4, and 5 in Figure 4C). The three new fermentation products were purified and characterized and shown to be 5'-dihydromethoxyhygromycin A (3), 5'-dihydromethoxyhygromycin A (4), and (E)-3-(3-hydroxy-4-O-β-fucofuranosylphenyl)-2-methylacrylic acid (5).

5'-dihydromethoxyhygromycin A (3) was produced in the SCH30 mutant with fermentation titers of approximately 180 mg/l. MS analyses revealed an m/z of 514 (M + H)+ for 3 (Figure 5B), 2 Da greater than the 512 (M + H)+ observed for hygromycin A and consistent with the proposed structure. Careful LC-MS analyses of the hygromycin A products from the wild-type strain revealed that 3 may also be produced albeit at very low levels (less than 1% of hygromycin A). Two peaks (each with an m/z of 516 [M + H]+ by LC-MS analyses), one of which coelutes with 4, were obtained by a nonstereoselective reduction of methoxyhygromycin A by using NaBH₄. Proton NMR analyses of the purified 5'-dihydromethoxyhygromycin A (4) revealed a methoxy singlet at 3.52 ppm, absent in 3 and hygromycin but present in methoxyhygromycin [7, 11]. The proton resonances for the α-L-fucofuranose moiety of 4 were the same as those observed for 3, including the presence of an H-5° proton coupled to both the H-6° methyl protons and H-3° proton. Analysis of the 13C NMR of dihydromethoxyhygromycin A showed the disappearance of the ketone peak (C-5') at 210.1 ppm and appearance of a new peak at 69.31 ppm, which corresponds to the new alcohol carbon peak (C-5°). Small changes in the resonances of the other furanose carbons were also observed. The HMOC (heteronuclear multiple quantum coherence) spectrum analysis was also consistent with structure of 4, with the new carbon signal at 69.31 ppm (C-5°) displaying a cross peak at 3.79 ppm (H-5°) in 1H dimension. Chemical reduction of methoxyhygromycin A with sodium borohydride provided a diastereomeric mixture of dihydromethoxyhygromycin A (4). MS and 1H NMR analysis were consistent with this being a diastereomeric mixture of 5'-dihydromethoxyhygromycin A (4). This mixture was cojected in the HPLC with a sample of naturally produced dihydromethoxyhygromycin A from the mutant and one of the peaks of the diastereomeric mixture coeluted with that of the natural one. The 13C NMR spectrum of the chemically obtained dihydromethoxyhygromycin A was exactly similar to that obtained from the mutant, confirming its structure.

The final product isolated from the SCH30 mutant was (E)-3-(3-hydroxy-4-O-β-fucofuranosylphenyl)-2-methylacrylic acid (5), which was produced at levels of approximately 125 mg/l. This peak was not observed in fermentations of the wild-type strain. LC-MS revealed an m/z of 363 (M + Na)+ for 5 (Figure 5B), consistent with the proposed structure. Proton and 13C NMR analyses revealed all of the resonances associated with the α-L-fucofuranose and (E)-3-(3,4-dihydroyphenyl)-2-methylacrylic acid moieties of both 3 and 4, but not those of the aminocyclitol moiety, supporting the structural assignment for this compound. Further support was provided by generating a diastereomeric mixture of 5, by
nonstereoselective reduction of hygromycin A followed by base hydrolysis. Two products were formed in this process (5 and 5a in Figure 4A), one of which coeluted with the purified material obtained from the SCH30 strain.

Acid hydrolysis of a mixture of 3 and 3a, provided the hygromycin aglycone lacking the furanose moiety (6), (E)-3-(3,4-dihydroxyphenyl)-2-methyl-N-(2L-2-amino-2-deoxy-4,5-O-methylene-neo-inositol)acrylamide (Figures 1 and 4A). LC-MS analyses failed to show any detectable level of this compound from fermentations of either the wild-type strain or the SCH30 mutant.

Implication for Hygromycin A Biosynthesis
The analysis of hygromycin products made by strain SCH30 provides compelling evidence that the last step in production of the 5-dehydro-L-fucofuranose moiety is an oxidation of L-fucofuranose catalyzed by the hyg26 gene product (Figure 6). It is not possible to determine from these analyses of the SCH30 mutant if this oxidation step occurs before or after formation of the glycosidic bond with the (E)-3-(3,4-dihydroxyphenyl)-2-methylacrylic acid moiety of hygromycin A. Nonetheless, studies in which either 3 or 4 was fed to a hyg23 mutant (this mutant is presumably blocked in formation of the L-fucose and does not produce any detectable hygromycin products) (N.P., S.A., and K.A.R., unpublished data) does not lead to hygromycin production. This suggests that both 3 and 4 are shunt products, and not pathway intermediates. While not conclusive, this observation suggests that the Hyg26 catalyzed oxidation may occur at an earlier stage in the biosynthetic process (Figure 6).

The production of significant levels of 5 in the SCH30 mutant suggests that the glycosidic linkage in hygromycin A can be formed in the absence of the amide linkage. In fermentations of the wild-type strain, the SCH 30, and the hyg23 mutant (blocked in formation of NDP-activated L-fucose), a hygromycin A analog lacking the dehydrofucofuranose moiety (6) was not observed. These observations suggest that the glycosidic bond of hygromycin may both precede and be a prerequisite for formation of the amide linkage (Figure 6). It is not possible to determine from these analyses if there is a specific
step in the conversion of 4-hydroxybenzoic acid to the (E)-3-(3,4-dihydroxyphenyl)-2-methylacrylyl thioester where this glycosidic bond is formed.

The combined yields of 3 and 4 in the SCH30 (320 mg/l) are approximately 25% of levels of 1 and 2 produced by the wild-type strain, suggesting that the enzymes catalyzing the amide and/or glycosidic bond formation have some degree of substrate flexibility. Nonetheless, the significant yields of 5 indicate that amide bond formation is perhaps slower by using a pathway intermediate with a fucofuranose moiety, possibly leading to formation of this shunt metabolite through a hydrolytic process.

Significance
This work describes a new and unusual biosynthetic gene cluster that makes hygromycin A and methoxyhygromycin A, which are structurally unusual aminocyclitol antibiotics produced by Streptomyces hygroscopicus NRRL 2388. This is the first reported gene cluster for a member of this class of important compounds. Among the notable features of the hygromycin A biosynthesis is the convergent nature of the process and the ability to generate a highly unusual aminocyclitol moiety with a methylene bridge that is essential for biological activity. We have established a genetic system for manipulating the hygromycin biosynthetic gene cluster in S. hygroscopicus and have shown that this can be used to provide significant yields of new hygromycin A analogs in a stereoselective and cost-effective manner. Hygromycin A analogs may serve as useful starting points for producing molecules with potential clinical or agricultural applications.

Experimental Procedures

Bacterial Strains, Media, and Culture Conditions
S. hygroscopicus NRRL 2388 and the SCH30 mutant were maintained and grown on ISP2 medium (0.4% yeast extract, 1.0% malt extract, 0.4% dextrose, 2.0% agar at pH 7.2), while the mannitol soy flour (MS) media [47] was used for intergeneric conjugation. All E. coli strains in this study were grown following standard protocols [48].

PCR Amplification of a Partial GDP-D-Mannose 4,6-Dehydratase Gene
Amplification of a partial hyg5 gene encoding a putative MDH from S. hygroscopicus NRRL 2388 was achieved with a set of degenerate primers: Myco-F1, AARCGHGCRCTGATCACYGGA; and Myco-R1, CGSGGBGATTCGTGRTTGAA. These primers were designed based on highly conserved motifs identified by creating a multiple alignment of deposited putative MDH genes of pimJ (CAC20923), nysDIII (AAF71765), amphDIII (AAK73500), and ata12 (CAD27644). Using a GC-rich PCR kit (Roche, Indianapolis, IN) with these primers and genomic DNA of S. hygroscopicus NRRL 2388, we amplified a PCR product of the expected size (~500–600 bp). This fragment was cloned and sequenced, and a BLAST search revealed the predicted product had greater than 78% amino acid sequence similarity to other putative MDHs. This partial hyg5 gene was used as probe to screen a cosmid clone genomic library of S. hygroscopicus NRRL 2388 and led to the identification of the hygromycin A biosynthetic gene cluster.

Cloning, Sequencing, and Annotation of the Hygromycin A Biosynthetic Gene Cluster
The total genomic DNA of S. hygroscopicus NRRL 2388 was prepared following standard protocols [47]. A genomic library was constructed by using Supercos-1 cosmid vector as recommended in the manufacturer’s protocol (Stratagene, La Jolla, CA). Approximately 3000 cosmids were probed with the DNA of a digoxigenin-labeled partial hyg5. Preparation of the digoxigenin probes and the subsequent hybridization and detection were performed as recommended in the manufacturer’s protocol (Roche). The overlapping cosmids identified by the hyg5 probe were sequenced to completion with the TOPO shotgun subcloning kit (Invitrogen, Carlsbad, CA). Automated DNA sequencing was performed on an ABI Prism 3700 DNA sequencer at DNA core facility of Medical College of Virginia, Virginia Commonwealth University. The DNA sequences were assembled with SeqMan II (DNASTAR, Inc., Madison, WI). The assembled DNA and deduced protein sequences were analyzed with DS Gene software (Accelrys, San Diego, CA) and Frame program [25] and compared with sequences in the public databases with the BLAST suite of programs [26].

Allelic Replacement of hyg26 within the Hygromycin A Biosynthetic Gene Cluster
The hyg26 gene of the hygromycin A biosynthetic gene cluster was replaced with aac(3)IV by the PCR-targeted Streptomyces gene
replacement method [49]. The aac(3)IV resistance marker (conferring apramycin resistance) and oriT were amplified from the pU773 disruption cassette with the primers [49]. The primers were HYG26 (5'-CCGCTTCTGATGCTGGAAGGAGGTGTCGATGATTCCGGGGA-TCCGGCAGC-3' and HYG26-R 5'-CTGCGGCGAACGCTGGACCGTATCAGTGAATTG ACCGTCGCT-3' (sequence homologous to pU773 disruption cassette is shown in bold, italicized text). The resulting PCR product from this set of primers was used to replace hyg26 first in the cosmid clone 17E3 and then in S. hygroscopicus NRRL 2388 to generate the SCH30 mutant (Hyg26: aac(3)IV by the intergenic conjugation method [49]. Allelic replacement of the hyg26 gene in SCH30 mutant was confirmed by PCR amplification and sequencing.

Production and Analysis of Hygromycin A
S. hygroscopicus NRRL 2388 and the SCH30 mutant were cultivated as reported previously [19]. Mycelia were removed by centrifugation, and the resulting supernatant was analyzed directly by HPLC.

HPLC and LC-MS Analyses
The culture filtrates of wild-type and mutant strain were subjected to HPLC and LC-MS analyses with a 5 µm Discovery HS C18 reverse phase column (4.6 x 250 mm, Supelco, Bellaforen, PA) and a methanol-water (containing 0.05% formic acid) gradient from 10:90 to 90:10 over 40 min at 1.0 ml/min flow rate. Hygromycin A and related compounds were detected at 272 nm and by MS analyses with a 5.0 mm discovery HS C18 column. The resulting supernatant was analyzed directly by HPLC. A portion (2.5 ml) of the diastereomeric mixture of products (1:1 mixture of 5'-epimers) was purified by the semipreparative HPLC conditions described above. MS and 1H NMR analysis were consistent with this being a diastereomeric mixture of 5'-dihydroygromycin A (3). Another portion (2.5 ml) of this diastereomeric mixture of 3 was treated with 1.5 ml aqueous NaOH solution, to give a final NaOH concentration of 6 M, under reflux conditions for 6 hr. HPLC analysis revealed that the two peaks for the 5' diastereomers of 3 were lost. Two new peaks were observed, and these were purified by semipreparative HPLC, and MS and 1H NMR analyses were consistent with a diastereomeric mixture of (E)-3-(3-hydroxy-4-O-α-L-fucosanoyl-phenyl)-2-methylacrylic acid (5). The final portion (5 ml) of the remaining diastereomeric mixture of 3 was cooled to 0°C, treated with 0.3 ml concentrated HCl, and stirred overnight at 50°C. The hygromycin A aglycone (6, (E)-3-(3,4-dihydroxyphenyl)-2-methyl-N-(2L-2-amino-2-deoxy-4,6-O-α-L-fucosanoylphenyl)-2-methylacrylic acid) (Figure 1) was precipitated by cooling the reaction mixture to 0°C and subsequently filtering and rinsing with cold acetone. LC-MS analysis of the product was a single peak with the predicted molecular mass (m/z 368 [M + H]+).

Acknowledgments
We are grateful to Dr. Kim Stutzman Engwall (Pfizer) for providing a cosmid library of S. hygroscopicus and to Dr. Steven Brickner (Pfizer) for providing hygromycin samples and degradation conditions. Partial financial support for this work was provided by Pfizer Incorporated. We are grateful to Dr. David Peyton and Ms. Vidya Dhote (Portland State University) for assistance with the 13C NMR analyses and antibacterial assays, respectively.

Received: March 13, 2006
Revised: May 8, 2006
Accepted: May 15, 2006
Published: July 28, 2006

References
Hygromycin A Biosynthetic Gene Cluster


three genes and analysis of transcriptional activity. Nucleic Acids Res. 15, 8041–8056.

Accession Numbers

The sequence reported here has been deposited in the GenBank database under the accession number DQ314862.