npg

REVIEW ARTICLE

Spontaneous and induced mutations to rifampicin, streptomycin and spectinomycin resistances in actinomycetes: mutagenic mechanisms and applications for strain improvement

Richard H Baltz

Chemical mutagenesis continues to be an important foundational methodology for the generation of highly productive actinomycete strains for the commercial production of antibiotics and other secondary metabolites. In the past, the determination of frequencies of chemically induced resistance to rifampicin (RifR), spectinomycin (SpcR) and streptomycin (StrR) have served as surrogate markers to monitor the efficiencies and robustness of mutagenic protocols. Recent studies indicate that high level RifR, SpcR and StrR phenotypes map to specific regions of the *rpoB*, *rpsE* and *rpsL* genes, respectively, in actinomycetes. Moreover, mutagenesis to RifR can occur spontaneously at many different sites in *rpoB*, and all six types of base-pair substitutions, as well as in-frame deletions and insertions, have been observed. The RifR/*rpoB* system provides a robust method to rank mutagenic protocols, to evaluate mutagen specificity and to study spontaneous mutagenesis mechanisms involved in the maintenance of high G + C content in *Streptomyces* species and other actinomycetes.
The Journal of Antibiotics (2014) **67**, 619–624; doi:10.1038/ja.2014.105; published online 13 August 2014

INTRODUCTION

The passing of Prof Piero Sensi, the discoverer of rifamycin and its commercial derivative rifampicin, has given us a pause to reflect on the impact of these discoveries on human medicine and the progression of scientific knowledge. As a graduate student at the University of Illinois, Urbana-Champaign, I was intrigued by observations on the heat-induced mutations in bacteriophage T4 DNA during heat denaturation of transforming DNA.[1] This led to further studies that uncovered the mechanisms of heat-induced mutagenesis causing GC to AT transitions and GC to CG transversions.[2,3] GC to AT transitions were caused by cytosine deamination, and rates of mutation were site-specific due to differences in Arrhenius activation energies. GC to CG transversions had uniform Arrhenius activation energies very similar to those for depurination (35 kcal mol^{-1}). Subsequent studies by Schaaper *et al.*[4] demonstrated that depurination is mutagenic in *Escherichia coli* induced for the SOS response, and occur at G or A residues causing G to T or A to T transversions. Both GC to AT transitions and GC to CG transversions in T4 were also acid catalyzed, as would be expected from chemical studies on cytosine deamination and depurination.[2,3] The rates of heat-induced mutations observed in T4 and the rates of cytosine deamination and depurination from chemical studies suggest that at physiological temperatures and pHs these two processes may have significant roles in spontaneous mutation.

After joining Eli Lilly and Company in 1974, I initiated studies on the fundamental mechanisms of mutagenesis in *Streptomyces* species to better understand how to optimize the mutagenic process for strain development.[5–7] We demonstrated that the tylosin-producing *Streptomyces fradiae* has both error-free and error-prone DNA repair pathways and an adaptive response to *N*-methyl-*N'*-nitro-*N*-nitroso-quanidine (MNNG) treatment similar to those in *E. coli*. Furthermore, *S. fradiae* has a *recA* gene functionally complementable by *E. coli recA*.[8] These studies were facilitated by the use of antibiotic-resistance markers, including resistance to Rifampicin (RifR), spectinomycin (SpcR) and streptomycin (StrR), to monitor the frequencies of mutation in different strains and with different mutagenic agents.

I have been involved in the development of many molecular genetic tools to manipulate actinomycetes over the years, but I have not lost sight of the importance of random chemical mutagenesis for strain improvement. Although many new genetic approaches have proven successful at improving secondary metabolite production in specific strains over the past two decades, chemical mutagenesis remains a robust foundational method to generate strains that produce very high titers of secondary metabolites during commercial fermentation. It is in this context that I consider the role of mutagenesis to RifR in furthering the understanding of mutagenic mechanisms and applying the information to efficient strain improvement.

CognoGen Biotechnology Consulting, Sarasota, FL, USA
Correspondence: Dr RH Baltz, CognoGen Biotechnology Consulting, 7636 Andora Drive, Sarasota, FL 34238, USA.
E-mail: rbaltz923@gmail.com

Rifampicin is an important antibiotic used in combination therapy to treat infections caused by *Mycobacterium tuberculosis*. It blocks transcription by binding to the β-subunit of RNA polymerase encoded by the *rpoB* gene. Rifampicin cannot be used as a single agent because of the relatively high spontaneous RifR mutation frequencies observed in strains of *M. tuberculosis* not previously exposed to the antibiotic, which range from 10^{-8} to 10^{-7}.[9] The relatively high spontaneous mutation frequencies can be attributed at least in part to the large number of mutable codons and specific base-pair substitutions that lead to amino-acid substitutions (>20) associated with the RifR phenotype in *M. tuberculosis*.[9,10]

RifR mutations occur at relatively high frequencies in other bacteria, and the mutations map to the same central region of *rpoB* gene as observed in *M. tuberculosis*. Early studies in *E. coli* analyzed 42 RifR mutants, and 35 had single mutations in the *rpoB* gene; the others had multiple mutations.[11] The 35 single mutations mapped to 17 alleles in 14 codons. Two were in-frame deletions, one was an in-frame insertion and the others were base-pair substitutions. More recently, Garibyan *et al.*[12] have identified 69 different base-pair substitutions among *E. coli* RifR mutants that map to 37 base pairs located in 24 codons of the *rpoB* gene. All six types of base-pair substitutions (that is, GC to AT and AT to GC transitions; and GC to CG, GC to TA, AT to CG and AT to TA transversions) were observed. This observation indicated that the *rpoB*/RifR system can provide a simple and reliable method to study spontaneous and induced mutational mechanisms in *E. coli* and possibly other bacteria.

Recently, Ochi and colleagues[13–15,17] have been studying the effects of *rpoB* mutations on secondary metabolite yields in actinomycetes. They observed that certain RifR mutations mapping to *rpoB* are associated with enhanced production of secondary metabolites or activation of cryptic secondary metabolite biosynthetic gene clusters. Ochi and colleagues[17,18] have also isolated many StrR-resistant actinomycetes that map to the *rpsL* gene, and have shown that certain StrR mutations are associated with improved secondary metabolite production. Thus spontaneous mutations to RifR and StrR have important applications for strain improvement and for the discovery of novel secondary metabolites by genome mining.[17–19]

In this report, I summarize what is known about the distribution of mutable sites in the *rpoB*, *rpsL* and *rpsE* genes, and how this information can be used to monitor mutagenic mechanisms and mutagen potency for actinomycete strain improvement. The current information indicates that induction of RifR mutations in the *rpoB* gene offers a robust method to monitor mutagenesis, thus emphasizing an important nonclinical aspect of the discoveries of rifamycin and RifR.

EXPERIMENTAL PROCEDURES
Nucleotide and protein searches
Nucleotide and protein searches were carried out by BLASTn and BLASTp (http://blast.ncbi.nlm.nih.gov/Blast.cgi).[20] Codon and amino-acid usage patterns for *Streptomyces avermitilis* and *E. coli* were determined by using the website: http://wishart.biology.ualberta.ca/BacMap.

RESULTS
Early studies on mutagenesis at Eli Lilly and Company
Studies on mutagenic mechanisms and mutagen potency at Eli Lilly and Company in the late 1970s and early 1980s indicated that mutation induction in the tylosin-producing *S. fradiae* could be monitored by measuring the frequencies of StrR, RifR and SpcR mutants among survivors of chemical or UV light mutagenesis.[5–7] After treatment with the most potent mutagen, MNNG, under conditions yielding $\sim 10\%$ surviving fractions, the relative frequencies for StrR, RifR and SpcR were approximately 10^{-6}, 10^{-5} and 10^{-4}. The RifR and SpcR frequencies were increased by about 10-fold by mutagenizing in the presence of chloramphenicol, which blocked the expression of an adaptive response.[7] As mutation to SpcR offered the largest dynamic range to test mutagen potency, it was chosen to rank mutagenic agents, and to validate the reproducibility and robustness of mutagenic protocols for strain improvement.[21,22]

In the early 1970s, little was known about the targets for such mutations, let alone the numbers and variety of mutable sites. In addition, little was known about the fundamental mechanisms of mutagenesis in actinomycetes. The state of the art on mutagenic mechanisms in the mid-1970s was reviewed by Drake and Baltz,[23] and a classic paper on mutagenic mechanisms focused on reversion of amber and ochre mutations in the *lacI* gene of *E. coli* was published in 1977.[24] The choice of mutagens explored at Lilly was heavily influenced by these publications, but we did not have a system to measure mutagen-induced base-pair substitution specificities. However, we developed a model for optimum mutagenesis by applying Poisson statistics and correlated the frequencies of SpcR and RifR mutations with optimum frequencies of positive yield mutations while minimizing negative mutations.[21] In this way, we could monitor the effectiveness and robustness of the mutagenesis protocols by measuring the frequencies of SpcR and/or RifR as surrogate markers prior to submitting mutagenized cell populations for high-throughput fermentation screening.

Mutagenesis in the *rpsE* gene encoding ribosomal protein S5
In a recent study, spontaneous mutations to SpcR in the daptomycin-producing *Streptomyces roseosporus* were sequenced, and all were mapped to the *rpsE* gene.[25] Each of the nine mutations caused amino-acid changes in a restricted region (positions 43–52) in loop 2 of the S5 ribosomal protein. The targeted peptide region (VAKVVKGGRR) is rich in basic and aliphatic amino acids. Of the nine mutations characterized, seven were amino-acid substitutions caused by transversions at G residues in codons for V43, A44 or K48, including five G to C and two G to T transversions. The other two mutations were in-frame deletions, causing ΔR51 or ΔV46-G49. This study also established that the SpcR phenotype displayed in *rpsE* mutants is recessive to SpcS, so the SpcR/SpcS system could be used for dominance selection, as had been demonstrated with StrR/StrS *rpsL* alleles in *S. roseosporus*.[26]

This small data set of SpcR mutations does not provide much insight into the high frequencies of MNNG and ethyl methanesulfonate (EMS) induced mutations to SpcR in *S. fradiae*,[5–7] as these mutagens induce base-pair substitutions almost exclusively as GC to AT transitions in *E. coli*.[12,24] Inspection of the 10 codons present in the SpcR region of the *rpsE* gene, however, indicated that 14 different G to A or C to T transitions at position 1 or 2 of codons would cause amino-acid substitutions (not shown). Therefore, it seems likely that exploring a larger set of SpcR mutations in actinomycetes will reveal sites for GC to AT transition mutations induced by MNNG or EMS. It is also possible that at least some of the mutagen-induced SpcR mutations in *Streptomyces* occur by in-frame deletions. It is known that MNNG causes deletion mutations at high frequencies in repetitive regions of the polyketide synthase genes encoding tylactone in *S. fradiae*.[27,28]

Mutagenesis in the *rpsL* gene encoding ribosomal protein S12
Mutants displaying a high level StrR phenotype map to the *rpsL* gene in bacteria, including actinomycetes. The *rpsL* gene encodes the

ribosomal protein S12, and the StrR phenotype is recessive to StrS in bacteria, including actinomycetes.[26] Ochi and colleagues[16,17] have been studying the effects of *rpsL* mutations on secondary metabolite production in actinomycetes for some time. They mapped the frequencies and locations of *rpsL* mutations in *S. avermitilis* and *Saccharopolyspora erythraea*, and noted that StrR mutations occurred at 10 nucleotide sites in eight codons, but most involved changes in lysine codons at amino-acid positions 43 and 88,[16] as has also been noted in other actinomycetes.[29,30] The StrR mutations occurred spontaneously by all six types of base-pair substitutions (Table 1). However, when we consider that each of the six types of base-pair substitutions has two configurations in the sense strand (for example, GC to AT transitions can be G to A or C to T), then two of the four transitions (G to A and T to C) and one of eight transversions (T to G) were not observed. A to G transitions distributed over three mutable sites were quite frequent, whereas C to T transitions at two mutable sites were infrequent. As most of the StrR mutations were located in lysine codons (AAG), and occur at all three positions, it seemed that further analysis of the relative frequencies at each site might shed light on the spontaneous mutation process. Table 2 shows that of the 63 mutations observed, 42 (67%) occurred by A to G transitions at codon positions 1 and 2. Thus it appears that AT to GC transitions may occur at higher frequencies than the individual transversion mutations. This may be one mechanism to maintain high G + C content in actinomycetes (see Discussion).

Mutagenesis in the *rpoB* gene encoding the β subunit of RNA polymerase

Garybyan *et al.*[12] generated 500 spontaneous and mutagen-induced RifR mutants in wild-type and mutator strains (*mutS*, *mutT* and *mutYM*) of *E. coli*. They noted that spontaneous mutations in wild-type *E. coli* arose at a frequency of 7.6×10^{-8} by all six base-pair

substitution pathways distributed over 37 mutable sites, whereas EMS induced only GC to AT transition mutations (40 of 40) at eight mutable sites at a frequency of 4.4×10^{-4}. This specificity was consistent with earlier studies on mutagen specificity in the *lacI* gene of *E. coli*.[24] Strains containing mutator mutations showed substantially higher spontaneous mutation frequencies and powerful enrichment for specific types of base-pair substitutions. Strains carrying *mutS*, *mutT* or *mutYM* mutations showed predominantly GC to AT transitions, AT to CG transversions or GC to TA transversions, respectively. This type of analysis may be useful for future strain improvement programs in actinomycetes (see Discussion).

The high conservation of the regions for RifR mutations among bacteria, the large number of mutable sites, the ease at selecting RifR mutants and the inclusion of all six types of base-pair substitutions among the RifR mutations suggested that the *rpoB*/RifR system might be a general method to study mutagenic processes in other bacteria, including those lacking well-developed genetic systems.[12]

Actinomycetes include numerous strains of academic and commercial interest, and induced mutagenesis remains an important general approach applicable to industrial strain improvement. Ochi and colleagues[13–15,17] have demonstrated that certain RifR mutants of different actinomycetes enhance secondary metabolite production, and in some cases activate silent or cryptic pathways, an approach that has important implications for microbial genome mining for new and novel secondary metabolites for drug development and other applications.[15,18,19] In an attempt to identify all of the mutable sites in *rpoB* that lead to enhanced secondary metabolite production, they sequenced 248 mutant *rpoB* alleles from seven different actinomycetes, including five *Streptomyces* species, *Sa. erythraea* and *Amycolatopsis orientalis*.[14] A side benefit of this work (reported here) was that it provided a data set to test the notion that the *rpoB*/RifR system could be used to explore the processes of mutation in actinomycetes as exemplified in *E. coli*.[12] The work of Tanaka *et al.*[14] demonstrated that spontaneous RifR mutations in actinomycetes can arise by all six base-pair substitution pathways, and in addition seven of the 248 mutations contained different in-frame deletions in a region of *rpoB* that encodes a 10 amino-acid peptide spanning Gln421-Asn430. If we examine all possible transitions (4) and transversions (8) on the sense strand, only T to G transversions were not observed (Table 3). The GC to AT transition mutations occurred at six sites and AT to GC transitions occurred at three sites. Both type of transition mutations were relatively abundant compared to transversion mutations, and the highest frequencies were accounted for C to T transitions and A to G transitions on the sense

Table 1 Spontaneous StrR mutations in *rpsL* genes in actinomycetes[a]

Mutation pathway	Base-pair substitutions[b]	Mutable sites	Mutations observed	Mutations per site	Normalized mutations per site
Transition	GC to AT	2	4	2	0.4
	G to A	0	0	0	
	C to T	2	4	2	
	AT to GC	3	42	14	2.7
	A to G	3	42	14	
	T to C	0	0	0	
Transversion	GC to CG	4	6	1.5	0.3
	G to C	3	5	1.7	
	C to G	1	1	1	
	GC to TA	2	12	6	1.2
	G to T	1	10	10	
	C to A	1	2	2	
	AT to CG	1	3	3	0.6
	A to C	1	3	3	
	T to G	0	0	0	
	AT to TA	2	6	3	0.6
	A to T	2	6	3	
	T to A	0	0	0	
Total		14	73	5.2	1.0

[a]Data calculated from Tanaka *et al.*[15]
[b]Base-pair substitutions (underlined) and individual base substitutions observed in the sense strand.

Table 2 Spontaneous StrR mutations in AAG lysine codons in the *rpsL* genes of *S. avermitilis* and *S. erythraea*

Codon position	Mutation	Type of mutation	Mutants observed
1	AAG to GAG(Glu)	Transition	11
	AAG to CAG(Thr)	Transversion	0
	AAG to TAG(Stop)	Transversion (lethal)	0
2	AAG to AGG(Arg)	Transition	31
	AAG to ACG(Thr)	Transversion	3
	AAG to ATG(Met)	Transversion	6
3	AAG to AAA(Lys)	Transition (silent)	0
	AAG to AAC(Asn)	Transversion	2
	AAG to AAT(Asn)	Transversion	10

Data compiled from Tanaka *et al.*[16]

Table 3 Spontaneous RifR mutations in *rpoB* genes in actinomycetes[a]

Mutation pathway	Base-pair substitutions[b]	Mutable sites	Mutations observed	Mutations per site	Normalized mutations per site
Transition	GC to AT	6	116	19.3	2.5
	G to A	2	18	9	
	C to T	4	98	24.5	
	AT to GC	4	47	11.8	1.6
	A to G	3	46	15.3	
	T to C	1	1	1	
Transversion	GC to CG	4	15	3.8	0.5
	G to C	1	1	1	
	C to G	3	14	4.7	
	GC to TA	7	12	1.7	0.2
	G to T	3	7	2.3	
	C to A	4	5	1.3	
	AT to CG	1	1	1	0.1
	A to C	1	1	1	
	T to G	0	0	0	
	AT to TA	5	15	3	0.4
	A to T	4	13	3.3	
	T to A	1	2	2	
Total		27	206	7.6	1.0

[a]Data calculated from Tanaka *et al.*[15]
[b]Base-pair substitutions (underlined) and individual base substitutions observed in the sense strand.

Table 4 Distribution of GC to AT and AT to GC transition mutations observed in *rpoB* genes in seven actinomycetes

Nucleotide position in sense strands	Transition in codon	Occurrence per seven actinomycetes	Numbers observed	Total mutants	Mutants per seven actinomycetes
1298	C to T	3	3,2,1,0,0,0,0	6	0.9
1309	C to T	7	22,8,5,4,4,3,1	47	6.7
1318	C to T	1	11,0,0,0,0,0,0	11	1.6
1325	C to T	4	23,9,1,1,0,0,0	34	4.9
1279	G to A	1	2,0,0,0,0,0,0	2	0.3
1319	G to A	5	7,4,3,1,1,0,0	16	2.4
1265	T to C	1	1,0,0,0,0,0,0	1	0.14
1276	A to G	1	2,0,0,0,0,0,0	2	0.3
1280	A to G	3	6,2,1,0,0,0,0	9	1.3
1310	A to G	5	29,3,1,1,1,0,0	35	5.0

strand. The higher number of mutable sites (six) for GC to AT transitions may help explain the higher frequencies of MNNG-induced mutations to RifR relative StrR (two mutable sites) observed in early studies at Lilly. If we look more closely at the six mutable sites for GC to AT transitions, it is clear that some sites are more mutable spontaneously than others (Table 4). GC to AT transition mutations at specific sites occurred in one to seven actinomycetes, and the total mutations per site ranged from 2 to 47. This >20-fold range strongly suggests that site-specific mutation rates depend on local DNA sequence context, which was well documented in early studies on mutagenic mechanisms in the *lacI* gene in *E. coli*.[24] Likewise, the site-specific spontaneous mutation rates for AT to GC transitions ranged

over 30-fold, and the A to G sites in the sense strand showed conspicuously higher spontaneous mutation frequencies than the single T to C transition. These wide differences in site-specific mutation frequencies were not observed for the eight site-specific transversions.

DISCUSSION

Over the past two decades, many new genetic and molecular genetic methods have been applied to strain improvement for secondary metabolite production in actinomycetes.[14,15,17,18,22,31,32] Some of the methods, including manipulation of positive and negative regulatory circuits, improving the robustness of transcription and translation, targeted gene or pathway amplifications, and heterologous expression are particularly applicable to early-stage strain improvement, including expression of cryptic secondary metabolite pathways from genome mining for drug discovery.[17–19,32] However, random chemical mutagenesis remains a robust general method that can be coupled with other approaches to generate strains capable of very high-level fermentation production for manufacturing.[21,22] An advantage of random mutagenesis is that it is agnostic to gene function: it requires no prior knowledge of the complex metabolic interactions that influence product yield, and functions effectively without a 'rational' hypothesis. It only requires optimization and balance between the types of base-pair substitutions, deletions and insertions induced.

In the 1970s and early 1980s, comparative mutagenesis studies were carried out at Eli Lilly and Company to identify the most efficient mutagenic agents and to optimize the dosage to maximize beneficial mutations while minimizing deleterious mutations.[5–7,21,22] At that time, Lilly was developing or manufacturing the actinomycete products monensin, narasin, tobramycin, tylosin and vancomycin, as well as the fungal products penicillin V and cephalosporin C, the starting material for multiple semisynthetic cephalosporin antibiotics. Therefore, the identification of highly efficient mutagenesis protocols was crucial for the development of economical production processes for these antibiotics. As surrogate markers, the frequencies of mutation to SpcR, RifR and StrR were initially compared, and SpcR was chosen for routine monitoring of mutagen treatments because of the dynamic range measurable. Mutation frequencies under optimal conditions ranged over three orders of magnitude from $\sim 10^{-7}$ for UV light to $\sim 10^{-4}$ for MNNG.[21,22] With the current information presented here, it appears that monitoring mutagenesis frequencies and pathways to RifR and StrR may provide a comprehensive approach to study mutagenic mechanisms and efficiencies, because all six type of base-pair substitutions can be measured. This may also be true for SpcR, but a much larger set of spontaneous and induced SpcR mutations needs to be analyzed before this judgment can be made. Having the ability to monitor all six base-pair substitution pathways provides a means to determine which mutagenic agents might cause mutations other than the common GC to AT transitions in actinomycetes. For instance, whereas MNNG and EMS cause GC to AT transitions exclusively in *E. coli*, 4-nitroquinoline-1-oxide causes GC to TA transversions ($\sim 10\%$) in addition to GC to AT transitions ($\sim 90\%$), and UV light causes substantial levels of transversions at GC and AT base pairs in addition to GC to AT transitions.[24] The latter two mutagens have been evaluated for mutations to SpcR, but not to RifR.[21,22]

It was pointed out some time ago that the range of potential amino-acid substitutions at the most common actinomycete codons (those containing G or C in position 3) inducible by mutagens that induce primarily or exclusively GC to AT transitions (for example,

MNNG and EMS) is limited, and could be augmented by a factor of two by a mutagenic protocol that induces AT to CG transversions.[22] In studies by the Miller group,[12] it was shown that *E. coli mutT* mutants showed elevated spontaneous mutation rates by making AT to CG transversions almost exclusively. A survey of sequenced *Streptomyces* genomes indicated that *mutT* homologs are present and highly conserved within the genus. However, their protein products show very poor homology to *E. coli* MutT, and there are no reports on the phenotype of *mutT* mutants in *Streptomyces*. MutT-family homologs are encoded by some other actinomycetes, but those encoded in *Amycolatopsis* and *Saccharopolyspora* species are distantly related (37 and 34% amino-acid identities, respectively) to the MutT homologs encoded by *Streptomyces* species, indicating that they are not orthologs, but paralogs. Gene disruption of *mutT* homologs in actinomycetes offers a possible opportunity to broaden the approaches to robust mutagenesis in actinomycetes, but genetic studies will be needed to determine whether any gene disruptions of *mutT* homologs in actinomycetes cause mutator phenotypes of elevated spontaneous AT to CG transversions. The *rpoB*/RifR system could be an effective means to characterize the *mutT* mutants.

Regarding the SpcR mutations, only a small number have been sequenced from a single *Streptomyces* species, and all contained transversions at GC sites or in-frame deletions in the *rpsE* gene.[25] As in-frame RifR deletions also occurred in the *rpoB* gene, both markers are suitable for analyzing deletion mutagenesis. It would be useful to expand the numbers of SpcR mutations in several actinomycetes to determine the full spectrum of available base-pair substitution pathways. Data on high frequency mutagenesis to SpcR induced by MNNG and EMS, and inspection of potential mutable sites in the SpcR locus in *S. roseosporus* leading to amino-acid substitutions suggest the existence of several sites for GC to AT transition mutations. Furthermore, it is possible that some SpcR mutants will show enhanced secondary metabolite production, as observed with some mutations in *rpsL* and other genes involved in ribosome function.[16,17,33,34]

More extensive studies on spontaneous mutation to RifR, SpcR and StrR may shed light on how actinomycetes maintain high G + C content in DNA. *Streptomyces* species have G + C content of ~70:50:90% in codon positions 1:2:3, averaging ~70% overall. Furthermore, G is enriched over C in position 1 of the sense strand (~42% vs 27%), and C is enriched over G in positions 2 (~28% vs 22%) and 3 (~55% vs 36%).[35] This type of enrichment for high G + C content results from forces acting primarily at the nucleotide level, rather than at the codon or amino-acid level, and results in preferential use of certain amino acids and reduced use of others when compared to proteins from microorganisms containing lower G + C content.[36] For example, when comparing frequencies of basic amino acids in proteins from *Streptomyces* species, Arg has elevated usage and Lys reduced usage as compared to microorganisms with lower G + C content. Of the six Arg codons, CGC and CGG are the most widely used (~40 and 36%). Of the other four codons, the three with two GC base pairs (CGT, AGG and CGA) are used less frequently (~9, 7 and 6%, respectively), whereas the AGA codon is used only 2% of the time.[36] Of the two Lys codons, AAG is used 91% and AAA 9% of the time. The average G + C content of Arg codons is 72%, whereas the average for Lys codons is 17%. The enrichment for the use of Arg cannot be accounted for simply converting Lys codons to Arg codons, because AGA and AGG Arg codons, which could be generated by AT to GC transitions from AAA and AAG Lys codons, are used only 9% of the time. If we compare the use of aliphatic amino acids from the complete *S. avermitilis* genome with those from

the *E. coli* genome, Ala, which has codons averaging 83% G + C, is used 13% of the time in *S. avermitilis* proteins compared to 9.5% in *E. coli*, whereas Ile, which has codons averaging 11% G + C, is used 3% of the time in *S. avermitilis* and 6% in *E. coli*. This is consistent with the interpretation that the evolutionary driving force for high G + C DNA is acting primarily at the nucleotide level, ultimately driving the preferential use of codons with high G + C content.[36]

In the analysis of frequencies of different base-pair substitutions arising spontaneously in actinomycetes in *rpoB* and *rpsL*, two patterns are emerging. Whereas all six base-pair substitution pathways can be measured, the two transition pathways account for higher frequencies of mutations than the transversion pathways. The GC to AT pathway gave specific frequencies at different sites that ranged over 20-fold. This is consistent with a cytosine deamination mechanism that has been shown previously in bacteriophage T4 to have site-specific mutation rates dependent on context-specific Arrhenius activation energies. This pathway would be counter to a mutational force that maintains high G + C content. On the other hand, the AT to GC transition pathway accounted for an even higher number of RifR and StrR mutants. This might be accounted for in part by a DNA repair mechanism that preferentially inserts C residues across from apurinic sites (lacking A or G) following depurination, unlike preferential insertion of A residues across from apurinic sites as observed in *E. coli*,[4] or it might be due to a DNA polymerase that has a relatively high transition error rate at A residues during DNA replication. Elevated spontaneous AT to GC mutagenesis could reverse the effects of the GC to AT pathway, particularly at third positions of actinomycete codons where GC to AT transitions are silent, but where G + C content is maintained at ~90%. The numbers of mutations analyzed so far are relatively low, so it would be useful to examine many additional spontaneous RifR, StrR and SpcR mutants in different actinomycetes to see if these preliminary trends hold, and to gain additional insights on the mutagenic processes in actinomycetes to help explain the driving forces to maintain high G + C content, and to enhance the prospects of improving mutagenic protocols for industrial strain improvement.

CONFLICT OF INTEREST
The author declares no conflict of interest.

1 Baltz, R. H. Biological properties of an improved transformation assay for native and denatured T4 DNA. *Virology* **70,** 52–64 (1976).
2 Baltz, R. H., Bingham, P. M. & Drake, J. W. Heat mutagenesis in bacteriophage T4: the transition pathway. *Proc. Natl Acad. Sci. USA* **73,** 1269–1273 (1976).
3 Bingham, P. M., Baltz, R. H., Ripley, L. S. & Drake, J. W. Heat mutagenesis in bacteriophage T4: the transversion pathway. *Proc. Natl Acad. Sci. USA* **73,** 4159–4163 (1976).
4 Schaaper, R. M., Kunkel, T. A. & Loeb, L. A. Infidelity of DNA synthesis associated with bypass of apurinic sites. *Proc. Natl Acad. Sci. USA* **80,** 487–491 (1983).
5 Stonesifer, J. & Baltz, R. H. Mutagenic DNA repair in *Streptomyces*. *Proc. Natl Acad. Sci. USA* **82,** 1180–1183 (1985).
6 Baltz, R. H. & Stonesifer, J. Mutagenic and error-free DNA repair in *Streptomyces*. *Mol. Gen. Genet.* **200,** 351–355 (1985).
7 Baltz, R. H. & Stonesifer, J. Adaptive response and enhancement of *N*-methyl-*N*′-nitro-*N*-nitrosoguanidine mutagenesis by chroramphenicol in *Streptomyces fradiae*. *J. Bacteriol.* **164,** 944–946 (1985).

8 Matsushima, P. & Baltz, R. H. *recA* gene of *Escherichia coli* complements defects in DNA repair and mutagenesis in *Streptomyces fradiae* JS6 (*mcr-6*). *J. Bacteriol* **169**, 4834–4836 (1987).
9 Telenti, A. *et al.* Detection of rifampicin-resistant mutations in *Mycobacterium tuberculosis. Lancet* **341**, 647–650 (1993).
10 Hauck, Y., Fabre, M., Vergnaud, G., Soler, C. & Pourcel, C. Comparison of two commercial assays for the characterization of *rpoB* mutations in *Mycobacterium tuberculosis* and description of new mutations conferring weak resistance to rifampicin. *J. Antimicrob. Chemother.* **64**, 259–262 (2009).
11 Jin, D. J. & Gross, C. A. Mapping and sequencing of mutations in the *Escherichia coli rpoB* gene that lead to rifampicin resistance. *J. Mol. Biol.* **202,** 45–58 (1988).
12 Garibyan, L. *et al.* Use of the *rpoB* gene to determine the specificity of base substitution mutations on the *Escherichia coli* chromosome. *DNA Repair (Amst).* **2,** 593–608 (2003).
13 Hu, H., Zhang, Q. & Ochi, K. Activation of antibiotic biosynthesis by specified mutations in the *rpoB* gene (encoding the RNA polymerase β subunit) of *Streptomyces lividans. J. Bacteriol.* **184**, 3984–3991 (2002).
14 Tanaka, Y. *et al.* Activation and products of the cryptic secondary metabolite biosynthetic gene clusters by rifampin resistance (*rpoB*) mutations in actinomycetes. *J. Bacteriol.* **195**, 2959–2970 (2013).
15 Ochi, K., Tanaka, Y. & Tojo, S. Activating the expression of bacterial cryptic genes by *rpoB* mutations in RNA polymerase or by rare earth elements. *J. Ind. Microbiol. Biotechnol.* **41**, 403–414 (2014).
16 Tanaka, Y. *et al.* Antibiotic overproduction by *rpsL* and *rsmG* mutants of various actinomycetes. *Appl. Environ. Microbiol.* **75**, 4919–4922 (2009).
17 Ochi, K. & Hosaka, T. New strategies for drug discovery: activation of silent or weakly expressed microbial gene clusters. *Appl. Microbiol. Biotechnol.* **97**, 87–98 (2013).
18 Baltz, R. H. Strain improvement in actinomycetes in the postgenomic era. *J. Ind. Microbiol. Biotechnol.* **38**, 657–666 (2011).
19 Bachmann, B. O., Van Lanen, S. G. & Baltz, R. H. Microbial genome mining for accelerated natural products discovery: is a renaissance in the making? *J. Ind. Microbiol. Biotechnol.* **41**, 175–184 (2014).
20 Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
21 Baltz, R. H. in *Manual of Industrial Microbiology and Biotechnology* (eds Demain, A. L. & Solomon, N. A.) 184–190 (ASM, Washington, DC, USA, 1986).
22 Baltz, R. H. New genetic methods to improve secondary metabolite production in *Streptomyces. J. Ind. Microbiol. Biotechnol.* **20**, 360–363 (1998).
23 Drake, J. W. & Baltz, R. H. The biochemistry of mutagenesis. *Annu. Rev. Biochem.* **45**, 11–37 (1976).
24 Coulondre, C. & Miller, J. H. Genetic studies of the *lac* repressor IV. Mutagenic specificity in the *lacI* gene of *Escherichia coli. J. Mol. Biol.* **117**, 577–606 (1977).
25 He, X., Miao, V. & Baltz, R. H. Spectinomycin resistance in *rpsE* mutants is recessive in *Streptomyces roseosporus. J. Antibiot.* **58**, 284–288 (2005).
26 Hosted, T. J. & Baltz, R. H. Use of *rpsL* for dominance selection and gene replacement in *Streptomyces roseosporus. J. Bacteriol.* **179**, 180–186 (1997).
27 Baltz, R. H. & Seno, E. T. Genetics of *Streptomyces fradiae* and tylosin biosynthesis. *Ann. Rev. Microbiol.* **42**, 547–574 (1988).
28 Beckmann, R. J., Cox, K. & Seno, E. T. in *Genetics and Molecular Biology of Industrial Microorganisms* (eds Hershberger, C. L., Queener, S. W. & Hegemann, G.) 176–186 (ASM, Washington, DC, USA, 1989).
29 Kenney, T. J. & Churchwood, G. Cloning and sequence analysis of the *rpsL* and *rpsG* genes of *Mycobacterium smegmatis* and characterization of mutations causing resistance to streptomycin. *J. Bacteriol.* **176**, 6153–6156 (1994).
30 Alexander, D. C. *et al.* Development of a genetic system for lipopeptide combinatorial biosynthesis in *Streptomyces fradiae* and heterologous expression of the A54145 biosynthetic gene cluster. *Appl. Environ. Microbiol.* **76**, 6877–6887 (2010).
31 Baltz, R. H. Genetic methods and strategies for secondary metabolite yield improvement in actinomycetes. *Antonie Van Leeuwenhoek* **79**, 251–259 (2001).
32 Zhu, H., Sandiford, S. K. & van Wezel, G. P. Triggers and cues that activate antibiotic production by actinomycetes. *J. Ind. Microbiol. Biotechnol.* **41**, 371–386 (2014).
33 Wang, G., Hosaka, T. & Ochi, K. Dramatic activation of antibiotic production in *Streptomyces coelicolor* by cumulative drug resistance mutations. *Appl. Environ. Microbiol.* **74**, 2834–2840 (2008).
34 Imai, Y., Fujiwara, T., Ochi, K. & Hosaka, T. Development of the ability to produce secondary metabolites in *Streptomyces* through the acquisition of erythromycin resistance. *J. Antibiot.* **65**, 323–326 (2012).
35 Seno, E. T & Baltz, R. H. in *Regulation of Secondary Metabolism in Actinomycetes* (ed Shipiro, S.) 1–48 (CRC, Boca Raton, USA, 1989).
36 Knight, R. D., Freeland, S. J. & Landweber, L. F. A simple model based on mutation and selection explains trends in codon and amino-acid usage and GC composition within and across genomes. *Genome Biol.* **2**, RESEARCH0010 (2001).