Dear Valeria Souza.

Thank you for dealing with our manuscript so quickly. We have addressed all of your and the reviewer's comments and uploaded both tracked and clean copies of the manuscript. Details of how we have addressed the comments are given below, point by point.

All the best,

Matt Hutchings

Reviewer 1

Minor points that should be addressed by the authors before the manuscript can be accepted.

1. The authors identify the borders of the antimycin cluster by obtaining mutant strains with deletions in neighboring genes, both upstream and downstream the ant cluster. The upstream border was identified by deleting genes STRS4_02194 and STRS4_02195, and the downstream border by deleting genes STRS4_02212 and STRS4_02214-STRS4_02217. None of these deletions has an effect on antimycin production. No mention is made of STRS4_02213. Is there a reason why this gene was left intact in these experiments? Why was it not considered for deletion instead of genes further downstream? In this respect Figure 1 should be improved, as there are only four black arrows to the right of STRS_02212 but these are labeled as corresponding to five genes (02213-02217). One of these black arrows is misaligned.

> We did not delete STRS4_002213 because it was not annotated as a coding sequence in our original RAST (Rapid Annotation using Subsystem Technology) analysis. A later manual annotation by the Genome Analysis Centre identified this CDS and it is included in the published genome annotation. The predicted 135 amino acid product of STRS4_002213 consists of a Roadblock / L7 family domain which is predicted to regulate dynein-type motor proteins involved in bacterial motility. Since streptomycetes are non-motile it is hard to suggest what its function might be but L7 domain-containing proteins are encoded by other streptomcyetes and are not associated with secondary metabolite clusters. STRS4_002213 is not a transcriptional regulator and it is highly unlikely to be involved in antimycin biosynthesis so we do not think it worthwhile to go back and delete this gene.

2. The authors should state in the figure legends what the error bars in figures 3, 5, 6, 8 and S4 represent (I assume it is the standard deviation).

> Standard deviation.

3. It is incorrect to describe operons antAB and antFG as such, since the gene order

should correspond to the direction of transcription. These two operons should be referred to as the antBA and antGF operons throughout the text and in Figure S1 (B)

> Agreed, we have changed this.

4. Figure 8 describes the effect of replacing the terminal Ala-Ala motif at the C-terminus of sigma-AntA on expression of the target operons. The authors should discuss the statistically significant increase in antH and antO expression levels (relative those of the wild type strain, i.e. gray and white bars) when the (antA-AA) construct is used to complement the antA deletion strain, particularly since this is not observed for the antF and antG transcripts. This difference is much higher in the case of the antO transcript levels, yet the figure marks this difference as statistically less significant (two asterisks, $P<0.01$) than that for antH (three asterisks, $P<0.001$). Are the authors sure about these analyses?

> There is no easy explanation for why the AntA-AA construct induces higher expression of antH and antO (the first and last genes of the antHIJKLMNO operon) than we observe in the wild-type. Crucially for this experiment though we only really need to compare the AntA-AA and AntA-DD carrying strains as they are identical apart from those two changes. It may be that making an in-frame deletion of the antB gene (which is between the promoter and the antA gene) increases antA expression in the constructs relative to wild-type but we do not see an increase in antGF expression.

5. Figure S2, the blue shading should be corrected to include the DVL amino acids in the sigma4 domain of the S. ambofaciens AntA orthologue.

> OK, we have changed this.

6. The paragraph from lines 116 to 132 should be rewritten to improve clarity (i.e. it described the generation of two constructs not just “a construct”). This should be explained more clearly.

>OK, we have clarified this paragraph in the Materials and Methods section.

Reviewer 2

I have a few minor criticisms/questions that should be addressed/answered before acceptance for publication (particularly point 2):

1. Summary, line 3: Presumably this sentence should read: “activities against the ANTI-apoptotic machineries inside human cells”, not “the apoptotic machineries”.

>Thanks, we have changed this.

2. Lines 47 and 287-293: The authors claim that AntA belongs to the ECF family of sigma factors, but the only data that substantiate this are presented in the SI, which is a pity.

Apparently, the AntA only shows similarity to ECF sigma factors in conserved regions 2 and 4. Could the authors state the level of amino acid identity/similarity in these regions to experimentally verified ECF sigma factors, or show an alignment to substantiate this point? Currently, it is not clear to me why the authors believe that AntA belongs to the ECF family of sigma factors. A phylogenetic tree is shown in Figure S3, but the text is too small to be able to see where the Ant family of putative ECFs sits with respect to known sigma factors. And how "significant" is this clustering? I believe that this statement needs further justification. However, in my view, the paper would still be worthy of publication even if a convincing bioinformatic case could not be made that AntA belongs to the ECF family.

> AntA only contains Sigma factor regions 2 and 4 which is characteristic of the ECF family and is not true of any other sigma factors. However, an alignment with representatives of all ECF sub-families shows that AntA forms a distant clade and therefore represents a new branch of this ECF family. The alternative would be to classify AntA as a distinct family of sigma factors separate from ECF and we do not think that our analysis justifies this conclusion.

3. Line 228: The use of a RT-PCR primer corresponding to sequences "300 bp upstream of the putative start codon" does not eliminate the possibility of a transcriptional start site located in the region of interest ("line 228 – "to exclude possible intragenic promoters"). RT-PCR will only indicate transcriptional read-through. The text needs to be modified accordingly.

> OK, we have changed this.

4. Line 248: At this stage of the work, this ought to read "Since antA is the only PUTATIVE regulatory gene in the ant gene cluster…..". i.e. "Putative" should be inserted.

> OK, we have changed this.

5. Line 264: "artificially" does not seem to be a very appropriate term. "precociously" or "prematurely" better?

> We have removed the word "artificially".

Reviewer 3

Experimental design

Lines 102-103: please be more specific about what primer combinations were used to verify gene knockouts. Not all PCRs are equally diagnostic…

> OK, we have added this information.

Validity of the findings

The conclusion that no regulators outside of sigma-AntA are needed for expression of the antFG and antH-O operons seems reasonable – however, one cannot exclude the

possibility that sigma-AntA has other target genes outside of the antimycin cluster, and the products of these may contribute to cluster activation. There are two relatively straight-forward experiments that would help to address this possibility. The first is searching the S. albus genome for other sigma-AntA-like promoters (given the very strong consensus sequence generated by the authors). There isn't any obvious resistance gene encoded in the cluster (and given that antimycins inhibit bacterial cytochrome c reductase, presumably some sort of resistance mechanism is required? – a pump maybe?) – could this be under the control of sigma-AntA elsewhere in the chromosome?

> Scanning the complete published *S. albus* genome with the AntA -10 and -35 sequences (using GLEME2 – part of MEME) returns only two significant hits, the antFG and antHIJKLMNO promoters suggesting there are no other AntA targets. We have described this analysis in the text.

The second experiment would involve taking advantage of cosmid 213 which appears to contain the entire antimycin biosynthetic cluster, and moving this into a heterologous host like S. coelicolor M1152. Successful expression of antimycin would support the idea that all necessary genes are contained on this cosmid (although it would not exclude the possibility that global transcription regulators may contribute to antimycin expression). If this was attempted but failed to result in antimycin production, this might suggest that other regulatory (or metabolic) elements are required – but would be useful information to include regardless. To this reviewer, this is an important experiment to include in a revised manuscript.

> Agreed. To test this we introduced cosmid 213 into *S. lividans*, wild-type *S. coelicolor* and the optimised *S. coelicolor* expression host strains M1146, M1152 and M1154 but could not detect antimycin production suggesting at least one additional regulator is required to activate expression of the *antBA* and *antCDE* gene clusters. We have described this in the text, at the end of the section titled "Antimycin production is dependent on the orphan ECF sigma factor σ^{AntA}."

Comments for the author

In the submitted manuscript, Seipke and colleagues describe the first in-depth genetic characterisation of the antimycin biosynthetic cluster. They effectively mapped the ends of the biosynthetic cluster (although previous bioinformatics analyses had provided strong support for the experimentally determined boundaries), and determined that cluster activity required the AntA sigma factor.

1. Antimycin expression: transcript levels for the different gene clusters were assessed after 18 (vegetative) or 42 (aerial/spores) h. However in S. coelicolor, secondary metabolism typically (or at least historically is thought to) initiates during the onset of aerial hyphae formation. Investigating a timepoint between 18 and 42h (when aerial hyphae formation is initiating) may show expression that is higher still – or it may not, which in itself would be interesting (as it would shake up this dogma a bit).

> We do not think this would add significantly to the manuscript, especially in terms of the amount of work involved.

2. It would be helpful to include information about where the transcription start points are

for the antG and antH clusters relative to their translation start sites, and how large the intergenic region is separating these two operons.

> The distance between the TSS and ATG or GTG translational start codons are shown in Figure 7A. Regarding the intergenic regions, there are 297 base pairs separating the predicted translational start codons of antG and antH and there is 39 base pairs separating the promoter regions we show in Figure 7A. We have added this information to the Figure legend.

3. Line 295: replace “potential” with “obvious”. What is it about STRS4_02195 that makes it a reasonable anti-sigma factor candidate? There exist orphaned anti-anti-sigma factors (e.g. BldG), so it may still be possible that sigma-AntA has a cognate anti-sigma factor, but that it is found elsewhere in the chromosome. Furthermore, on the ‘StrepDB’ website, there is another (small) gene annotated in the region between antB and antC (02198). Is this gene transcribed along with antA and antB? (indicating the exact transcription start site for this operon would be useful as well) Could this possibly encode an anti-sigma factor?

> As described in the manuscript, STRS4_02195 is the only gene product in the ant cluster with no assigned function and, like many anti-sigma factors, it is a transmembrane protein that shows no significant homology to other anti-sigma factors. Although orphaned anti-anti sigma factors exist we do not know of any orphan anti-sigma factors so it is very unlikely that AntA has an orphan anti-sigma. Unfortunately we have been unable to map the TSS for *antBA* presumably because of mRNA secondary structure but there is no CDS between the *antBA* genes, this must be a mis-annotation and we will ask the curator of StrepDB to amend this error.

4. Figure 3: there is an enormous difference in expression levels between antF and antG. Any speculation as to what is going on there?

> No, but if we had to speculate we would say there must be secondary structure which reduces read through to *antF* (the second gene in the operon).

Editorial comments:

Line 25: what hydrophobic groove is being discussed? Why is antimycin binding here significant?

> The hydrophobic groove of Bcl2 - binding here inhibits Bcl2 activity, we have amended the text to make it clearer.

Lines 48-49 and 53-54: repetitive

> OK, we have changed this.

Line 74: please change “ClpX” to “ClpXP”, and provide an accompanying reference

> OK, we have changed this.

Throughout: micro, delta, gamma and degree symbols failed to convert properly in the PDF

> This is a problem with the Peer J submission which can hopefully be corrected by their editorial team.

Line 82: italicise “E. coli”

> OK, we have changed this.

Line 109: instead of “T4 DNA polymerase” did the authors mean “T4 DNA ligase”?

> Yes, we have changed this.

Line 96: please remove the rogue bracket

> OK, we have changed this.

Line 270: mention that transcriptional start sites were mapped using 5’-RLM RACE

> OK, we have changed this.

Line 276: does “sequence identity” refer to amino acid sequence identity?

>In this case it is nucleotide sequence identity, we have made this clear.

Table 1: the spacing of the plasmid information makes it challenging to figure out what information corresponds to any given plasmid

> We have added borders to the Table cells.

For all RT-PCR results, what do the error bars represent? Standard deviation? Standard error?

> Standard deviation.

Figure 1: change to “cosmid 456 comprised additional DNA”

> OK, we have changed this.

Figure 7: the “zero nucleotide identity” shared between antAB and the AntA-targeted promoters is modestly overstated, as there are two nucleotides that are shared between these. Perhaps replace “zero” with “little”. The first two sentences in section (B) are also completely repetitive.

> OK, we have changed this.