

An RT-PCR artifact in the characterization of bacterial operons

Maritza Guacucano

Laboratorio de Bioinformática y Microbiología Molecular
Universidad de Santiago, Santiago, Chile
Tel: 56-2-6786073
Fax: 56-2-7355580
E-mail: mguacucano@hotmail.com

Gloria Levican

Instituto de Ciencias Biomédicas
Universidad de Chile, Santiago, Chile
Tel: 56-2-6786073
Fax: 56-2-7355580
E-mail: glevican@canela.med.uchile.cl

David S. Holmes

Laboratorio de Bioinformática y Microbiología Molecular
Universidad de Santiago, Santiago, Chile
Tel: 56-2-6811542
Fax: 56-2-6812108
E-mail: dsholmes@hotmail.com

Eugenia Jedlicki *

Instituto de Ciencias Biomédicas
Universidad de Chile, Santiago, Chile
Tel: 56-2-678-6258
Fax: 56-2-737-6320
E-mail: ejedlick@machi.med.uchile.cl

Financial support: Fondecyt 1980665.

Keywords: *Acidithiobacillus ferrooxidans*, co-transcription, polycistronic mRNA.

During the course of experiments involving RT-PCR to investigate the co-transcription of the *pet* and *res* operons of the bacterium *Acidithiobacillus ferrooxidans* we observed that the mRNA of certain gene pairs can be amplified by RT-PCR in the absence of added RT-primer. We provide experimental evidence that eliminates, as explanations for the observed results, the possible global contamination of the RNA preparation by genomic DNA or by contaminating single stranded DNA or RNA fragments. We suggest that the results can be explained by self-priming of the RNA perhaps as a result of secondary structure configurations that provide a suitable 3'-terminus to prime the reverse transcriptase. This type of artifact can be recognized by carrying out a control in which exogenously added RT-primers are left out of the reaction. Such a control is essential if RT-PCR is to be used to investigate and correctly identify operons. RT-PCR is also an important tool for the amplification of eukaryotic mRNAs and our observations may also be applicable to these systems.

Several efforts are underway to organize the information that is pouring out of the microbial genome sequencing projects. These projects include, for example, EcoCyc, (<http://ecocyc.pangeasystems.com/ecocyc>), WIT (<http://wit.Integratedgenomics.com/IGwit>) and Kegg (<http://www.genome.ad.jp/kegg/kegg2.html>) where genetic and metabolic data are being linked to provide a comprehensive view of the complete metabolism of an organism. One of the striking facts to emerge from these projects is the preponderance of genes, as many as 70-80%, that are clustered in a manner that suggests they are operons. This view is derived from the observation that, in a given gene cluster, the genes are tightly spaced on the same DNA strand or even overlapping, allowing little room for individual gene promoters and DNA regulatory regions. In some cases the organization of the genes within a cluster may be completely or partially conserved between several different microorganisms suggesting they participate in evolutionary conserved regulatory control circuits, strengthening the possibility that they are operons.

* Corresponding author

The abundance of such gene clusters and the probability that a knowledge of their regulation will contribute overwhelmingly to an understanding of global gene expression and cellular physiology substantially increases the importance of experiments that validate or negate the affirmation that a gene cluster is indeed an operon. An operon can be defined as two or more co-transcribed genes whose expression is regulated by the same control elements. A common first step to experimentally identify an operon is to demonstrate co-transcription of the candidate gene cluster and the most frequent way to do this is via RT-PCR of whole cell RNA. RT-PCR is a two step process in which a specific RNA is first reverse transcribed into a cDNA using a gene specific primer, and is then amplified via PCR using primers designed from genes within the gene cluster under investigation (for example, Gupta, 1999). Successful PCR amplification is taken as a signal that the genes are co-transcribed into a polycistronic mRNA and, consequently, are members of an operon.

In this paper we demonstrate that RT-PCR can generate an artifact that could potentially lead to the misinterpretation of experiments designed to investigate co-transcription of genes that, in turn, could result in the misidentification of operons.

Materials and Methods

A. ferrooxidans ATCC19859 was grown in FeII medium at pH 2.4 as described previously (Yates and Holmes, 1987). RNA was isolated from cells in mid-log as described by Hagen and Young (1978). All solutions and materials were pre-treated with 0.1% diethylpyrocarbonate. 50 µg of purified RNA was suspended in 300 µl of 5 mM MgCl₂-0.1M NaAcetate pH 5.2 and treated with RNase-free DNaseI (Gibco) at 210 U/ml for 30 min at 37°C. It was then extracted with phenol-chloroform and precipitated at -18°C by addition of 1/10 vol of 2.5 M NaAcetate pH 5.5 and 2 vols of 95% EtOH.

An aliquot of the RNA was treated as follows: 10 µg of RNA in 10 µl of 10mM Tris-1mM EDTA pH 7 was incubated for 2 hr at 37°C with 1 µg of RNaseA (Amersham). RT-PCR was carried using Superscript II (Gibco) and Taq polymerase (Gibco) according to the manufacturer's recommendations. The reverse transcription reaction was routinely carried out at 42°C according to the manufacturer's recommendations, but reactions were also tested at 56°C with no difference in the results.

The following RT-primers were used (the annealing temperatures are indicated in brackets, followed by the position of the 5'-end nucleotide of the primer relative to GenBank accession AF220499):

A = 5'AGAAAAAACCCGCTAATCAGCTGGAG (76°C, 2202),

B = 5'CACCAAAACACCAGCAGCAA (60°C, 6491).

The following PCR primers were used:

1 = 5'AACCAGGTATGCAGGTAACGGTGC (74°C, 1578),

2 = 5'CAACAACAGCAATGATCCTGCATAGTATAG (76°C, 2167),

3 = 5'ACACCTTGGACTTACGCTTGC (64°C, 1021),

4 = 5'TCAGCGCAATACTGCAGC (56°C, 6021),

5 = 5'GGTTCACCCAATGTCCCTCGGC (68°C, 6429).

Each experiment was accompanied by the following controls, a reaction in which: 1) No RNA or DNA template was added, 2) genomic DNA but not RNA was added, 3) RNA but not reverse transcriptase was added and 4) RNAase treated RNA was added. Reactions 1, 3 and 4 did not yield detectable product (results not shown).

Results and Discussion

Figure 1A shows a map of the pet and the res operons of the bacterium *Acidithiobacillus ferrooxidans*. Between the two operons is an intergenic spacer of 620 bp that includes a characteristic rho-independent-like stop for transcription of the upstream operon and a promoter and +1 start of transcription for the downstream operon (GenBank Acc. no. AF220499). During our initial analyses of these operons several RT-PCR experiments were carried out in an effort to investigate the co-transcription of the respective genes within the operons and to substantiate the proposed lack of co-transcription between the operons.

While carrying out these experiments we recorded the occurrence of the following artifact. Using the RT primer A to convert the RNA of the pet operon to cDNA (Figure 1A), we detected appropriately sized PCR amplification products between petA and petB (using PCR primers 1 and 2) and between nad-1 and petB (using PCR primers 3 and 2) (Figure 1B). These results are to be expected if nad-1, petA and petB are co-transcribed. However, we also detected a PCR amplification product between resC and resB (using PCR primers 4 and 5). This result was not expected because the RNA corresponding to these genes had not been converted into cDNA in this experiment. Similarly, if RNA corresponding to the res operon is converted into cDNA using the RT primer B, a PCR amplification product of the expected size is detected between resC and resB (using PCR primers 4 and 5) as expected, and an unexpected PCR product is also detected between petA and petB (using PCR primers 1 and 2) but not between nad-1 and petB (using PCR primers 3 and 2) (Figure 1B). This appears to indicate that petA and petB are co-transcribed as part of the same operon as resB and resC.

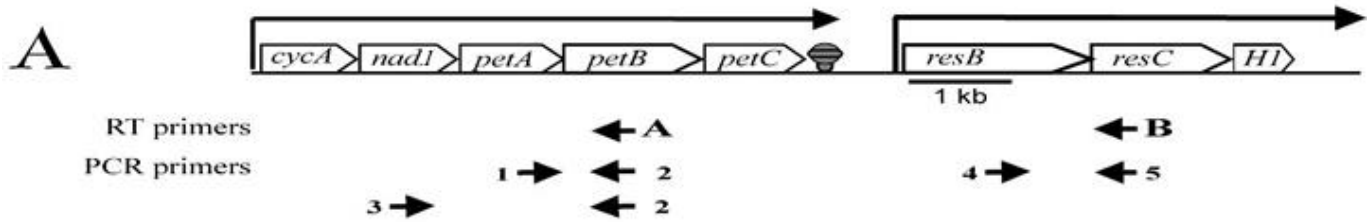


Figure 1 A. Map of the *pet* and *res* operons of *A. ferrooxidans* indicating the genes present and their direction of transcription (GenBank Acc. AF22049). The locations and orientations of RT- and PCR primers are indicated.

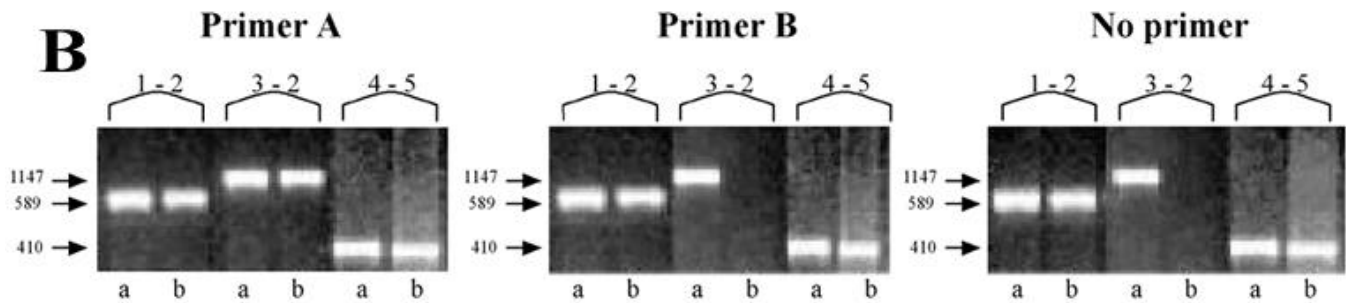


Figure 1 B. Agarose gel electrophoresis result of RT-PCR experiments using the primers indicated **a** = using genomic DNA as the template. **b** = using RNA as the template. Expected sizes of the products are indicated to the left, scale in bps.

We repeated the above experiments in the absence of any added RT primer. As can be seen from inspection of Figure 1B the gene pairs *petA*-*petB* and *resB*-*resC* are amplified both with and without added RT-primer but the gene pair *nad1*-*petB* is not amplified. We repeated this experiment with and without added RT-primer for an additional eight gene pairs and found consistently that without added primer short gene regions of <1000 bps are consistently amplified without added RT-primer but those longer than that are not so amplified (data not shown).

That amplification can occur in the absence of exogenous RT-primers immediately eliminates two possible explanations for the results. First, the results cannot be explained by the possibility that genes of the *pet* and *res* operons are repeated elsewhere in the genome but in a configuration that would generate polycistronic mRNA capable of explaining the amplifications. In addition, we do not favor this possibility because no such configuration has been found by inspection of the >90% complete genome sequence of *A. ferrooxidans* (Selkov et al. 2000).

Second, the possibility that cross-hybridization of the RT-primers to the mRNA of both operons can be eliminated as an explanation. This possibility is also ruled out because the RT-primers were designed to be specific for their target and because they do not generate inappropriate products when

only genomic DNA template is used as a control.

A possible explanation, that cannot be ruled out solely by the observation that PCR products can be generated in the absence of RT-primer, is that the RNA preparation is contaminated with double stranded DNA making it possible for the PCR amplification to proceed from a DNA template. Our preparations are treated extensively with DNase and, in control experiments, we do not observe the amplification of gene pairs that are known not to be in operons, eliminating the possibility of global DNA contamination. This view is substantiated by a second control in which the RNA preparation is treated with RNase and, in this case, there is no PCR amplification of any gene pairs tested. In addition, in the absence of added reverse transcriptase there is no detectable product (data not shown), suggesting that PCR primer pairs and Taq polymerase alone are not sufficient to generate product which would not be the case if contaminating double stranded DNA were present.

Another possible explanation is that the RNA is capable of self-priming. This might occur if the RNA preparation is contaminated by single stranded RNA or DNA fragments that can base pair with the mRNA of the operons, or if fold-back regions of the mRNA are present that can provide an appropriate 3' terminus for reverse transcriptase to act on. This could typically be a 3'-OH or, as has been shown for

retron reverse transcriptase, a 2'-OH (for example, Shimamoto et al. 1999).

We consider it unlikely that self-priming results from contamination of short random segments of DNA or RNA that can serve as primers. We have observed that other genes, used as negative controls in our experiments, are not self-primed, demonstrating that self-priming is not a general phenomenon as would be expected if the RNA was being randomly primed. We favor the possibility that self-priming is occurring as a result of the secondary structure of the mRNAs which provide suitable termini to prime the reverse transcriptase.

Whatever the explanation for the observed results, it remains evident that RT-PCR can generate results that are not only difficult to interpret but could also result in the misdiagnosis of operons. We suggest that investigators, using RT-PCR to investigate co-transcription of genes, should include a standard control in which exogenous RT-primers are left out of the reaction.

We have restricted our discussion so far to the relevance of the RT-PCR artifact in prokaryotic systems, however RT-PCR is also used extensively for the amplification of eukaryotic mRNAs (for example, Kako et al. 2000) and is a frequently used technique to detect and/or quantitate the presence of mRNAs of medical importance (for example, Cham et al. 2000). Since self-priming of eukaryotic mRNAs may also occur, the same control recommended for prokaryotic systems should be carried out during RT-PCR experiments of eukaryotic mRNAs.

Acknowledgements

Work supported by Fondecyt 1980665.

References

Cham, F., Heyndrickx, L., Janssens, W., Vereecken, K., De Houwer, K., Coppens, S., Van Der Auwera, G., Whittle, H. and Van Der Groen, G. (2000). Development of a one-tube multiplex reverse transcriptase-polymerase chain reaction assay for the simultaneous amplification of HIV type 1 group M gag and env heteroduplex mobility assay fragments. *AIDS Research Human Retroviruses* 16:1503-1505.

Gupta, A. (1999). RT-PCR Characterization of long multi-gene operons and multiple transcript gene clusters in bacteria. *BioTechniques* 27:966-972.

Hagen, F. and Young, E. (1978). Effect of RNase III on efficiency of translation of bacteriophage T7 lysozyme mRNA. *Journal of Virology* 26:793-804.

Kako, K., Tsumori, K., Ohmasa, Y., Takahashi, Y. and Munekata, E. (2000). The expression of Cox17p in rodent tissues and cells. *European Journal of Biochemistry*

267:6699-6707.

Selkov, E., Overbeek, R., Kogan, Y., Chu, L., Vonstein, V., Holmes, D., Silver, S. and Fonstein, M. (2000). Functional analysis of gapped microbial genomes: principles with examples from *Thiobacillus ferrooxidans*. *Proceedings of the National Academy of Science USA* 97:3509-3514.

Shimamoto, T., Kawanishi, H., Tsuchiya, T., Inouye, S. and Inouye, M. (1998). In vitro synthesis of multicopy single-stranded DNA, using separate primer and template RNAs, by *Escherichia coli* reverse transcriptase. *Journal of Bacteriology* 180:2999-3002.

Yates, J. and Holmes, D.S. (1987). Two families of repeated DNA sequences in *Thiobacillus ferrooxidans*. *Journal of Bacteriology* 169:1861-1870.