Series

What history tells us
XIV.
Regulation of gene expression by non-coding RNAs: the early steps

Michel Morange

Centre Cavaillès and IHPST, Ecole normale supérieure, 29 rue d'Ulm, 75230 Paris Cedex 05, France

(Fax, 33-144-323941; Email, morange@biologie.ens.fr)

1. Introduction

RNA interference has become one of the tools favoured by biologists to ascertain rapidly the function of a gene in cells and organisms. In addition, microRNAs have a natural role in the control of differentiation and development. A "missing continent of regulation" has emerged quite recently. It is characterized by the sophistication of the mechanisms involved, as well as by its numerous connections with regulatory mechanisms operating at other levels: consider for instance the role of small RNAs in the establishment and maintenance of the chromatin state (Grewal and Elgin 2007). The award of the 2006 Nobel Prize to Andrew Fire and Craig Mello, who demonstrated the role of double-stranded RNAs in genetic interference and their potential utilization to inactivate gene expression (Fire *et al* 1998), has been at the centre of a controversy (Bots *et al* 2006): the first observations on post-transcriptional gene silencing – co-suppression – were made in plants at the beginning of the 1990s (Napoli *et al* 1990; van der Krol *et al* 1990), but did not attract the attention of biologists before they were reproduced in animals, in *Caenorhabditis elegans*, emphasizing the separation between animal and plant cell and molecular biology. More generally, the discovery appeared tardy. How was it possible that such a continent of regulation had escaped attention? Technical difficulties, such as the non-adaptation of traditional methods of molecular biology to small RNAs, have been pointed out.

The truth is that regulation by RNAs was considered very early, as early as the structures and relations between DNA, RNA and proteins were described. But the models proposed were shown to be either wrong or too vague; and regulation by RNAs was considered more as an experimental tool than as a natural mechanism. In any case, the sophistication of extant mechanisms was never anticipated, and the simplest hypothesis – the inhibition of initiation of translation by an antisense RNA – was favoured even if data were regularly at odds with this simple model. We will consider three successive episodes in the prehistory of the notion of regulatory RNA: the (wrong) identification of the lactose and λ repressors as RNAs in 1959; the role of RNAs in the Britten-Davidson model of gene regulation proposed in 1969; and the use of antisense RNAs as tools to inactivate genes initiated by Harold Weintraub in 1984 with not only the successes, but also the difficulties it rapidly encountered.

2. Repressor is a RNA

One always looks retrospectively at historical papers and designs and sees in them what one expects to find: the roots of our present knowledge. Consider the 1961 paper published in the *Journal of Molecular Biology* by François Jacob and Jacques Monod on the "genetic regulatory mechanisms in the synthesis of proteins" and Figure 6 in which the operon model is presented (Jacob and Monod 1961). Two models are in fact proposed, one in which regulation takes place at the transcriptional level, and a second one in which the repressor inhibits translation. In both models, the repressor is represented as an RNA, and not as a protein.

There were experimental reasons to support the hypothesis that the repressor was an RNA. They derived from work by different researchers on lactose (Pardee and Prestidge 1959) and the λ system (Bertani 1957; Jacob and Campbell 1959) with protein synthesis inhibitors, analogues of amino acids, and carbon starvation. But such an identification raised a

Keywords. antisense RNA; repetitive sequences; repressor; siRNA

difficulty acknowledged by Jacob and Monod (1961): it was impossible to imagine how the repressor could interact with the inducer (lactose or a derivative of lactose).

This (incorrect) identification of the repressor with an RNA is either omitted from later descriptions (Pardee 1985; with the exception of Müller-Hill 1996) or considered as a transient error without any significance. The repressor became a protein when Jacob and his collaborators showed that bacterial suppressor strains, known to act at the protein synthesis level, could relieve some of the mutations occurring in the repressor gene: it was the proof that the repressor gene was translated into a protein (Jacob *et al* 1962). In addition the development of the allosteric theory by Jean-Pierre Changeux, Jeffries Wyman and Monod rapidly transformed the repressor into the prototype of an allosteric protein (Monod et al 1963).

The identification of the repressor with an RNA was not the simple result of ambiguous experiments. There were two more fundamental reasons to adopt the hypothesis that the repressor was an RNA. The first was the clear separation it introduced between structural and regulatory genes, an essential distinction for Monod and Jacob. Not only did these two categories of genes have different functions, but experiments suggested that their products were of a different nature. The second reason was that it was easy to imagine how a nucleic acid could interact with another nucleic acid. As early as 1956, Alexander Rich had shown that it was possible to form double helices with two complementary RNA chains (Rich and Davies 1956). It was the same reason that pressed Francis Crick to propose the hypothesis of an adaptor necessary for the translation of a nucleic acid into a protein (Judson 1996); a hypothesis confirmed with the discovery of soluble RNAs, now called transfer RNAs. It explains why the hypothesis that the repressor contained an RNA fraction was repeatedly proposed in the years which followed the initial description (see, for instance, Miller and Sobell 1966). There was a simple chemical model to explain how nucleic acids could interact; there was no equivalent to explain an interaction between a nucleic acid and a protein.

3. The Britten-Davidson model of gene regulation

The same reasons explain why Roy Britten and Eric Davidson gave RNAs a major role in gene regulation in their 1969 model. The scientific context was different. The objective was to elaborate models of gene regulation able to explain the processes of differentiation and development that many molecular biologists, trained in the study of bacteria and bacteriophages, were now directly addressing. The models had to explain the global changes in gene activity occurring during these processes, the distinction emphasized by embryologists between determination and differentiation, as well as the possibility for these regulatory mechanisms to evolve to generate the complex morphological and functional changes occurring in evolution. Most embryologists considered the operon model as simplistic. The mechanisms which were looked for had to exist only in organisms where differentiation and development were taking place.

Early molecular observations on higher organisms had also revealed the existence of phenomena specific to these organisms: the abundance of repetitive sequences; their high and differential expression during development; the fact that most of the corresponding RNAs were not exported to the cytoplasm, and their action had therefore to be limited to the nucleus (for a review, *see* Weinberg 1973). It was also considered that RNA might be a component of the chromosomes (Huang and Bonner 1965).

In their model, Britten and Davidson emphasized the role of activator RNAs, the products of integrator genes, positively controlling, by interacting with receptor genes, the expression of batteries of producer genes (Britten and Davidson 1969, 1971). This model was praised by the embryologist Conrad Waddington (Waddington 1969). Ten years later, the new experimental results were considered by Davidson and Britten as supporting their early model, even if they proposed new functions for the integrator RNAs (Davidson and Britten 1979). Only some years later, Davidson admitted that there was still a lack of experimental proof in favour of his model, and the role of repetitive sequences (Davidson and Posakony 1982). Retrospectively, the connection established in this model between regulatory RNAs and repetitive sequences was a wonderful anticipation, since it is widely admitted today that the control by siRNAs originated as a weapon against the repetitive sequences brought by viruses and transposons (Plasterk 2002).

The model of Britten and Davidson was not mechanistic. Nothing was said about the way these integrator RNAs regulated gene expression. But this absence of a mechanism was not really noticed, as if the possibility for RNA to interact with DNA and to control its activity was considered obvious.

4. Harold Weintraub and the use of antisense RNAs to control gene expression

In 1984, Jonathan Izant and Harold Weintraub, and almost simultaneously John Rubenstein and Jean-François Nicolas, showed that it was possible to inhibit the expression of a gene using antisense RNA (Izant and Weintraub 1984; Rubenstein *et al* 1984). The latter could be produced by transfection in cells of antisense constructs or, as demonstrated very rapidly in amphibians and *Drosophila*, by the direct injection of antisense RNAs produced *in vitro* (Melton 1985; Rosenberg *et al* 1985). The same year, Takashi Mizuno and Masayori Inouye demonstrated that in prokaryotes the expression of the major outer membrane protein OmpF is controlled by the production of an antisense RNA (Mizuno *et al* 1984).

The experiment of Weintraub led to the generalized use of antisense RNA to control gene activity. It was the puzzling observation that sense RNAs might have the same effect as antisense RNAs that finally led – after a long delay – to the discovery of double-stranded interfering RNAs. Whereas the production of antisense RNAs was well adapted to research, the use of antisense oligonucleotides with modified bases became, some years later, the favoured approach for therapeutic uses (Stein and Cheng 1993; Wagner 1994).

The use of antisense RNA was introduced by Weintraub as a way to ascertain the function of genes at a time when the knockout strategy had not attained its present efficiency. Both strategies are used today, siRNAs being preferred for the rapidity of their use in cell systems, and in some organisms such as *C. elegans* in which they can be easily introduced.

When Harold Weintraub performed his experiments, there were many data showing that it is possible to form a double helix between complementary RNA sequences, and to inhibit gene expression by antisense nucleic acids.

In vitro, antisense oligodeoxynucleotides had been shown to be capable of inhibiting RNA translation (Zamecnik and Stephenson 1978 ; Stephenson and Zamecnik 1978). The possibility of inhibiting translation by the formation of a hybrid with DNA was the basis of the method of "DNA.mRNA hybrid-arrested cell-free translation" extensively used by molecular biologists to establish the correspondence between RNAs and the coding parts (exons) of the genes (Paterson *et al* 1977). In the same way, the interaction between mRNA and the 3' terminal region of the 16S ribosomal RNA in prokaryotes had been demonstrated by competition with an oligonucleotide complementary to the ribosomal RNA (Taniguchi and Weissmann 1978).

In prokaryotes, transposition (Simons and Klechner 1983), as well as the replication of colicin and antibiotic resistance plasmids, was shown to be controlled by small antisense RNAs (Tomizawa *et al* 1981; Rosen *et al* 1981). The role of snRNP, and in particular U1 RNA, in splicing had demonstrated that, despite their interactions with proteins, RNAs produced in the nucleus of eukaryotic cells were able to form antiparallel double helices with small RNAs (Lerner *et al* 1980 ; Rogers and Wall 1980).

So there were excellent arguments in favour of a simple model in which gene expression might be controlled by the binding of small antisense RNAs and oligonucleotides at the initiation site of translation. In his article, Harold Weintraub hypothesized that antisense RNAs probably had a general regulatory role within organisms.

But this model of regulation never became dominant. The first reason was that the use of antisense RNAs or oligonucleotides was from the beginning conceived of as a tool, adapted to different purposes, not as a naturally occurring mechanism. The purpose was so precise that sometimes it obscured the possibility that the phenomenon that was observed had a wider significance, and other possible utilizations: the successes rapidly met by this new strategy paradoxically masked its biological significance.

An additional reason is that the existence of a regulation by antisense RNAs was only very slowly established. In 1984, when Weintraub designed his experiments, there were a few rare cases demonstrated in bacteria. Some years later, although the applicability of the experimental approach of Weintraub had been confirmed in different experimental systems such as *Dictyostelium discoideum* (Knecht and Loomis 1987), clear evidence for such a regulation in eukaryotes was still lacking (Green *et al* 1986), and the number of observations in bacteria was increasing very slowly (Liao *et al* 1987; for a general review: Takayama and Inouye 1990). The evidence for regulation of translation by antisense RNAs was only provided in *C. elegans* nine years after the experiment of Weintraub (Lee *et al* 1993). The obvious possibility of such a mechanism of regulation from the physicochemical point of view, despite the absence of clear evidence in its favour, explains that the major reviews written on it were focused on the optimization of the process, and its physicochemical characteristics more than on its biological significance (Eguchi *et al* 1991).

But the third reason to consider regulation by antisense RNA an oddity was, from its early days, the accumulation of puzzling observations, which prevented the emergence of a simple model and misdirected researchers. For instance, antisense RNAs could apparently interfere with gene expression at different levels: transcription, export from the nucleus, translation and degradation of the mRNAs. Antisense effects were observed at the initiation site of translation, but also at the 5' and 3' ends of the mRNA. This suggested that the action of antisense RNAs and oligonucleotides might be indirect instead of direct, through, for instance, the proteins known to interact with these parts of the mRNA to control its stability and capacity to be translated. Additional observations often made the picture even more obscure. What was the significance of the unwinding and nucleotide-modifying activity specifically acting on double-stranded RNAs (Rebagliati and Melton 1987; Bass and Weintraub 1987, 1988 ; Kimelman and Kirschner 1989)? Some other observations, such as the existence of a protein kinase activated by double-stranded RNA, induced by interferon and involved in the general inhibition of translation, gave double-stranded RNA a well-defined, circumvented function, and blinded researchers to its additional roles (Farrell *et al* 1977). Even the discovery of the catalytic role of RNAs acted as a decoy, diverting research in a new, totally different direction.

The idea that RNAs might have important regulatory functions was not absent from the minds of biologists. The opposite was true – the hypothesis of antisense RNAs

was proposed as early as 1962 (Rich 1962) –, but clear, unambiguous facts were hard to come by. There were obstacles preventing the discovery of the regulatory role of non-coding RNAs. But, in contrast to a model now favoured by historians, the obstacles were not to be found in the domination of one antagonistic model and the requirement for a scientific revolution to displace it. Rather, the obstacles were in a well-adopted practice, and in complex facts. Practical applications of the antisense principle were developed before the evidence was obtained that organisms had used similar principles of regulation. But no one would have imagined that such a simple principle of regulation would require such sophisticated machinery.

Acknowledgements

I am indebted to David Marsh for critical reading of the manuscript, and to Frédérique Théry for providing important references.

References

Bass B L and Weintraub H 1987 A developmentally regulated activity that unwinds RNA duplexes; *Cell* **48** 607–613

Bass B L and Weintraub H 1988 An unwinding activity that covalently modifies its double-stranded RNA substrate; *Cell* **55** 1089–1098

Bertani L E 1957 The effect of the inhibition of protein synthesis on the establishment of lysogeny; *Virology* **4** 53–71

Bots M, Maughan S and Nieuwland J 2006 RNAi Nobel ignores vital groundwork on plants; *Nature (London)* **443** 906

Britten R J and Davidson E H 1969 Gene regulation for higher cells: a theory; *Science* **165** 349–357

Britten R J and Davidson E H 1971 Repetitive and non-repetituve DNA sequences and a speculation on the origins of evolutionary novelty; *Q. Rev. Biol.* **46** 111–133

Davidson E H and Britten R J 1979 Regulation of gene expression: Possible role of repetitive sequences; *Science* **204** 1052–1059

Davidson E H and Posakony J W 1982 Repetitive sequence transcripts in development; *Nature (London)* **297** 633–635

Eguchi Y, Itoh T and Tomizawa J-I 1991 Antisense RNA; *Annu. Rev. Biochem.* **60** 631–652

Farrell P J, Balkow K, Hunt T and Jackson R J 1977 Phosphorylation of initiation factor eIF-2 and the control of reticulocyte protein synthesis; *Cell* **11** 187–200

Fire A, Xu S, Montgomery M K, Kostas S A, Driver S E and Mello C C 1998 Potent and specific genetic interference by double-strandedRNA in *Caenorhabditis elegans*; *Nature (London)* **391** 806–811

Green P J, Pines O and Inouye M 1986 The role of antisense RNA in gene regulation; *Annu. Rev. Biochem.* **55** 569–597

Grewal S I S and Elgin S C R 2007 Transcription and RNA interference in the formation of heterochromatin ; *Nature (London)* **447** 399–406

Huang R C and Bonner J 1965 Histone-bound RNA, a component of native nucleohistone; *Proc. Natl. Acad. Sci. USA* **54** 960–967

Izant J G and Weintraub H 1984 Inhibition of thymidine kinase gene expression by anti-sense RNA: A molecular approach to genetic analysis; *Cell* **36** 1007–1015

Jacob F and Campbell A 1959 Sur le système de répression assurant l'immunité chez les bactéries lysogènes ; *C. R. Acad. Sci. Paris (Série III)* **248** 3219–3221

Jacob F and Monod J 1961 Genetic regulatory mechanisms in the synthesis of proteins; *J. Mol. Biol.* **3** 318–356

Jacob F, Sussman R and Monod J 1962 Sur la nature du répresseur assurant l'immunité des bactéries lysogènes; *C. R. Acad. Sci. Paris (Série III)* **254** 4214–4216

Judson H F 1996 *The eighth day of creation: Makers of the revolution in biology* (New York: Cold Spring Harbor Laboratory Press) pp 291–292

Kimelman D and Kirschner M W 1989 An antisense mRNA directs the covalent modification of the transcript encoding fibroblast growth factor in Xenopus oocytes; *Cell* **59** 687–696

Knecht D A and Loomis W F 1987 Antisense RNA inactivation of myosin heavy chain gene expression in *Dictyostelium discoideum*; *Science* **236** 1081–1086

Lee R C, Feinbaum R L and Ambros V 1993 The C. elegans heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*; *Cell* **75** 843–854

Lerner M R, Boyle J A, Mount S M, Wolin S L and Steitz J A 1980 Are snRNPs involved in splicing?; *Nature (London)* **283** 220–224

Liao S-M, Wu T-h, Chiang C H, Susskind M M and McClure W R 1987 Control of gene expression in bacteriophage P22 by a small antisense RNA I. Characterization in vitro of the P_{sar} promoter and the sar RNA transcript; *Genes Dev.* **1** 197–203

Melton D A 1985 Injected anti-sense RNAs specifically block messenger RNA translation *in vivo*; *Proc. Natl. Acad. Sci. USA* **82** 144–148

Miller J H and Sobell H M 1966 A molecular model for gene expression; *Proc. Natl. Acad. Sci. USA* **55** 1201–1205

Mizuno T, Chou M-Y and Inouye M 1984 A unique mechanism regulating gene expression: Translational inhibition by a complementary RNA transcript (micRNA); *Proc. Natl. Acad. Sci. USA* **81** 1966–1970

Monod J, Changeux J-P and Jacob F 1963 Allosteric proteins and cellular control systems; *J. Mol. Biol.* **6** 306–329

Müller-Hill B 1996 *The* lac *operon: A short history of a genetic paradigm* (Berlin: Walter de Gruyter), 89–90

Napoli C, Lemieux C and Jorgensen R 1990 Introduction of a chimeric chalcone synthase gene into petunia results in reversible co-suppression of homologous genes *in trans*; *Plant Cell* **2** 279–289

Pardee A B 1985 Molecular basis of gene expression: Origins from the Pajama experiment; *BioEsaays* **2** 86–89

Pardee A B and Prestidge L S 1959 On the nature of the repressor of β-galactosidase synthesis in *Escherichia coli*; *Biochim. Biophys. Acta* **36** 545–547

Paterson B M, Roberts B E and Kuff E L 1977 Structural gene identification and mapping by DNA.mRNA hybrid-arrested cell-free translation; *Proc. Natl. Acad. Sci. USA* **74** 4370–4374

Plasterk R H A 2002 RNA silencing: The genome's immune system; *Science* **296** 1263–1265

Rebagliati M R and Melton D A 1987 Antisense RNA injections in fertilized frog eggs reveal an RNA duplex unwinding activity; *Cell* **48** 599–605

Rich A and Davies D R 1956 A new two stranded helical structure: polyadenylic acid and polyuridylic acid; *J. Am. Chem. Soc.* **78** 3548–3549

Rich A 1962 On the problems of evolution and biochemical information transfer; in *Horizons in biochemistry* (eds) M Kasha and B Pullman (New York: Academic Press) pp 103–126

Rogers J and Wall R 1980 A mechanism for RNA splicing; *Proc. Natl. Acad. Sci. USA* **77** 1877–1879

Rosen J, Ryder T, Ohtsubo H and Ohtsubo E 1981 Role of RNA transcripts in replication incompatibility and copy number control in antibiotic resistance plasmid derivatives; *Nature (London)* **290** 794–797

Rosenberg U B, Preiss A, Seifert E, Jäckle H and Knipple D C 1985 Production of phenocopies by *Krüppel* antisense RNA injection into *Drosophila* embryos; *Nature (London)* **313** 703–706

Rubenstein J L R, Nicolas J-F and Jacob F 1984 L'ARN non sens (nsARN) : un outil pour inactiver spécifiquement l'expression d'un gene donné *in vivo*; *C. R. Acad. Sci. Paris Série III* **299** 271–274

Simons R W and Kleckner N 1983 Translational control of IS10 transposition; *Cell* **34** 683–691

Stein C A and Cheng Y-C 1993 Antisense oligonucleotides as therapeutic agents – Is the bullet really magical?; *Science* **261** 1004–1012

Stephenson M L and Zamecnik P C 1978 Inhibition of Rous sarcoma viral RNA translation by a specific oligodeoxyribonu cleotide; *Proc. Natl. Acad. Sci. USA* **75** 285–288

Takayama K M and Inouye M 1990 Antisense RNA; *Crit. Rev. Biochem. Mol. Biol.* **25** 155–184

Taniguchi T and Weissmann C 1978 Inhibition of Qβ RNA 70S ribosome initiation complex formation by an oligonucleotide complementary to the 3' terminal region of *E. coli* 16S ribosomal RNA; *Nature (London)* **275** 770–772

Tomizawa J-I, Itoh T, Selzer G and Som T 1981 Inhibition of ColE1 RNA primer formation by a plasmid-specified small RNA; *Proc. Natl. Acad. Sci. USA* **78** 1421–1425

van der Krol A, Mur L, Beld M, Mol J N M and Stuitje A R 1990 Flavonoid genes in petunia: Addition of a limited number of gene copies may lead to a suppression of gene expression; *Plant Cell* **2** 291–299

Waddington C H 1969 Gene regulation in higher cells; *Science* **166** 639–640

Wagner R W 1994 Gene inhibition using antisense oligodeoxynucleotides; Nature **372** 333–335

Weinberg R A 1973 Nuclear RNA metabolism; *Annu. Rev. Biochem.* **42** 329–354

Zamecnik P C and Stephenson M L 1978 Inhibition of Rous sarcoma virus replication and cell transformation by a specific oligodeoxynucleotide; *Proc. Natl. Acad. Sci. USA* **75** 280–284

*e*Publication: 14 August 2008