Series
DOI 10.1007/s12038-014-9482-4

What history tells us XXXV. Enhancers: Their existence and characteristics have raised puzzling issues since their discovery

MICHEL MORANGE

Centre Cavaillès, République des Savoirs USR 3608, Ecole Normale Supérieure, 29 rue d'Ulm, 75230 Paris Cedex 05, France

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

1. Introduction

Enhancers are sequences of DNA that can stimulate the rate of transcription initiation at large distances from the promoters, in 5′ or 3′ positions relative to the gene, and in one or another orientation. Enhancers are now part and parcel of the description of transcription in metazoans. Enhancers are formed by the modular association of short DNA sequences corresponding to the binding sites of different transcription factors.

Enhancers participate in the developmental control of gene expression, and the localization and description of the putative five hundred thousand (or more) enhancers present in the human genome are a major issue in genomics. The evolutionary role of enhancer modification is also under scrutiny.

This continuous interest in enhancers for the last thirty years or more should not mask the fact that their existence and characteristics have always been puzzling for molecular biologists, and that the vision of enhancers has constantly evolved.

2. The discovery of enhancers

In December 1981, Walter Schaffner and his group introduced the term 'enhancer' to designate a 72 bp repeat close to the origin of replication of the SV40 virus which stimulates transcription initiation of the rabbit β-globin gene (Banerji *et al.* 1981). They described and generalized the characteristics of this enhancer – action at distance, in either orientation and upstream or downstream of the gene – and proposed some mechanistic explanations for its action: enhancers might change the superhelical density of DNA, bind to the nuclear matrix, or provide entry sites for RNA polymerase II.

This article had been preceded by a series of observations on different organisms (sea urchins, yeast, the SV40 simian virus) revealing that, in eukaryotes, sequences localized far upstream of the transcription initiation site are required for a full transcription rate – often in parallel with the puzzling observation that the promoter sequences (sequences immediately upstream of the transcription initiation site) did not appear themselves to be indispensable (Grosschedl and Birnstiel 1980; Benoist and Chambon 1980; Gruss *et al.* 1981; Struhl 1981).

What was new in the article by Schaffner and colleagues was the demonstration that the SV40 enhancer was active on a heterologous promoter, the introduction of the term, and a precise description of the three characteristics that define an enhancer. Similar observations were made in parallel by the same group on the enhancer of the polyoma virus (de Villiers and Schaffner 1981). Simultaneously, Pierre Chambon's group published similar results obtained with the SV40 enhancer, although they observed a dependence of the enhancing effect on the distance between the promoter and the enhancer. The latter observation was in agreement with the hypothesis that Chambon favoured – that enhancers are the

Keywords. Enhancer; genomics; loop; promoter; RNA polymerase II; transcription factors; transcription in eukaryotes

Published online: 20 October 2014

site of entry of RNA polymerase II, an enzyme which he studied in parallel (Moreau *et al.* 1981).

Many publications over the next two years confirmed these early observations (for a review, see Khoury and Gruss 1983). The existence of enhancers was demonstrated in other viruses, but they could be also upstream or downstream of cellular genes, one of the first and best examples being the enhancers present in the immunoglobulin genes coding for the light and heavy chains of antibodies (Gillies *et al.* 1983; Banerji *et al.* 1983; Queen and Baltimore 1983; Neuberger 1983). Evidence for the existence of enhancers was also obtained in the case of the globin genes, but the diversity of the genes and of their regulation made the observations more complex. The role of these cellular enhancers was confirmed by the creation of transgenic mice (Grosschedl *et al.* 1984). These studies revealed that the enhancing effect was cell-dependent (de Villiers *et al.* 1982; Spandidos and Wilkie 1983). The specificity of viral enhancers for cells and organisms could explain the host specificity of these viruses. The term 'enhancer' was limited to metazoans. It was not adopted for yeast, and the acronym UAS for 'upstream activating sequences' was preferred, since these sequences did not seem to be active in the 3′ position (Struhl 1982).

3. The context of the discovery

The fact that transcriptional regulation was different and more complex in eukaryotes than in prokaryotes was not a surprise, coming only four years after the totally unexpected discovery that eukaryotic RNAs were spliced following transcription!

The discovery of enhancers would not have been possible without the development of genetic engineering techniques in previous decades: use of restriction enzymes, construction of chimeric molecules containing an enhancer from one species, a promoter from another, and eventually a reporter gene from a third one to make the assay of the transcriptional activity more sensitive. Also essential was the possibility of modifying these DNA sequences (by *in vitro* mutagenesis), and of reintroducing these constructions in cells by transfection (or microinjection in the case of *Xenopus* oocytes), in organisms by transgenesis, or in the rapidly developed *in vitro* transcription systems.

That the first enhancer was discovered in a virus was no accident. Viruses were considered as the gateway to the study of eukaryotes, as bacteriophages had been thirty years before in the study of bacteria.

The impact of the discovery of enhancers, whose existence was at odds with what was known of the regulation of transcription in prokaryotes, was strengthened by the immediate explanation that it provided for the oncogenic potential of some retroviruses (Hayward *et al.* 1981), exactly at the same time when oncogenes were discovered (Bishop 1981). The transforming power of these viruses was puzzling since they did not harbour oncogenes of cellular origin as most transforming retroviruses were shown to do. It was initially proposed that they could provide efficient promoters, by insertion in the genome of their long terminal repeats, to the cellular genes situated close to the site of integration. The replacement in the previous sentence of 'promoters' by 'enhancers' obviated the requirement for very precise integration of the retrovirus, and also explained why the insertion could occur downstream of the proto-oncogene.

The early explanatory models also relied deeply on the scientific context in which enhancers were discovered. The study of prokaryotic and eukaryotic RNA polymerases was intense, and the possibility of RNA polymerase sliding along DNA to the precise site of transcriptional initiation, instead of reaching it by diffusion in three dimensions, had recently been proposed and demonstrated in the case of *E. coli* (Winter *et al.* 1981). The same was true for the role of the nuclear matrix in transcription, and for the structure of the nucleus, within which there were potentially transcriptional factories. Interest in the structure of chromatin had been renewed by the studies of Harold Weintraub and Mark Groudine (1976) and by the subsequent discovery of nuclease-hypersensitive sites (Varshavsky *et al.* 1978; Scott and Wigmore 1978; Wu *et al.* 1979) correlated with the absence of nucleosomes on DNA: the 72 bp repeat of the SV40 enhancer was a site of hypersensitivity to nucleases, and was shown by electron microscopy to be devoid of nucleosomes (Saragosti *et al.* 1980). The role of superhelicity in DNA – induced in eukaryotes by the binding of nucleosomes – was actively sought, while enzymes modifying it were purified and characterized (Champoux 1978).

4. From a structural regulatory element to an ensemble of binding sites for transcription factors

In the first two years after their discovery, enhancers were seen as structural elements, and the explanation for their functions was sought at the level of chromatin organization and localization in the cell nucleus.

Looking for proteins that bind to enhancers was not immediately considered important. The paucity of appropriate techniques to purify the protein factors that interact with enhancer sequences did not fully explain this lack of interest.

One example is quite illustrative: the polyoma enhancer was shown to be inactive in mouse embryonal carcinoma cells, but mutations were characterized that made the enhancer active, permitting the replication of the virus in these cells (Vasseur *et al.* 1980; Katinka *et al.* 1980). A retrospective interpretation is that mutations created sites for factors present in embryonal carcinoma cells, or relieved the binding of inhibitory factors. But at the time, sequence comparisons and general explanations were sufficient objectives for this work.

This vision progressively changed as rapidly accumulating evidence showed that the enhancing effect differed from one cell type to another: the simplest explanation was that this cell specificity resulted from the presence (or absence) of cell factors that bind to the enhancer sequences and are responsible for the enhancing effect (Laimins *et al.* 1982; Weiher *et al.* 1983). In parallel, the study of genes whose expression was stimulated by external signals such as the addition of steroid hormones or heat shock showed that the activation was associated with the presence of short specific DNA sequences located at variable distances from the site of transcriptional initiation, with which specific transcription factors interacted (Chandler *et al.* 1983).

In the following years, with the rapid development of *in vitro* transcription systems, many of these transcription factors were isolated, their precise binding sites characterized *in vitro* and *ex vivo* by footprinting experiments, and their role in the enhancing effect demonstrated.

At the end of the 1980s, a new picture of enhancers had emerged. These sequences of variable length (from 50 bp to more than 1000 bp), and situated at distances from the transcription initiation site that could be considerable, are formed by the modular association of short (less than 10 bp) sequences specific for different transcription factors. The same sequences are also found in the promoters (Maniatis *et al.* 1987; Müller *et al.* 1988).

This new vision did not do away with the previous interpretations of the enhancing effect: it did not exclude that enhancers might contribute to the structural organization of chromatin, to the binding of RNA polymerases, or to the localization of the transcribed genes in a specific part of the nucleus. But the favoured model was that enhancers and promoters were put in close proximity by the formation of a DNA loop, and co-activated transcription. The activating effect of enhancers was the direct consequence of an increase in the local concentration of transcription factors at the transcriptional initiation site. It was demonstrated in *E. coli* that binding of a multimeric repressor at two different operator sites could generate a DNA loop. This evidence was recurrently used to support the formation of loops between enhancers and promoters in eukaryotes, although the loops described in prokaryotes had characteristics – small, dependence on the pitch of the DNA helix – that were not obvious for the loops formed between enhancers and promoters.

5. Enhancers: Twenty-five years later

Twenty-five years have passed, and the description of the structure and function of enhancers has not dramatically changed during this time (Blackwood and Kadonaga 1998; Pennacchio *et al.* 2013).

The notion of enhancers might have disappeared, or at least lost its importance. If what matters is simply a high concentration of different transcription factors close to the transcription initiation site, the fact that these factors are bound to DNA sequences that are close to (in the sequence) or at huge distances from the promoter might have appeared as secondary.

Such was not the case. The first reason is that many researchers have not abandoned the idea that there is something special about enhancers, that there are additional reasons for their existence. The recent discovery of enhancer RNAs, long non-coding RNAs transcribed from enhancers, supports this interpretation (Orom and Shiekhattar 2013).

Another reason is that enhancers have found a new life in developmental biology (Levine *et al.* 2014). Enhancers were shown to control the tissue in which a gene is activated during development, but also the stage at which the gene is switched on (Garabedian *et al.* 1986). Not only have these observations been fully confirmed, but, after the rise of Evo-Devo, enhancers are now considered as the privileged sites of mutation responsible for morphogenetic transformations (Carroll 2008; Levine 2010).

In parallel, the growing interest in epigenetics, and the need to complement the information provided by the Human Genome Project, have led to the development of new programmes such as ENCODE, for the identification of 'functional' sequences, and in particular regulatory sequences, within 'junk' DNA. Conserved sequences located outside of the coding regions have been discovered, and a large fraction of them correspond to enhancer sequences (Pennacchio *et al.* 2006). The use of the technique of chromosome conformation capture has provided evidence for the occurrence of loops between enhancers and promoters (Sanyal *et al.* 2012). The topology of mammalian developmental enhancers is the subject of active study (De Laat and Duboule 2013). In addition, mutations in enhancers have been associated with the development of diseases (Kleinjan and Lettice 2008; Sakabe *et al.* 2012).

6. Conclusion

The scientific context in which enhancers were initially described explains the different types of mechanisms that were proposed to justify their existence – such as their role in the structural organization of chromatin. Most of all, enhancers were the reflection of the sudden discovery of the huge differences existing between metazoans and other organisms, as exemplified by the discovery of splicing, and their existence retrospectively explained the difficulties encountered in the study of transcription in metazoans in previous years.

Between 1981 and 1983, enhancers were considered as 'transcriptional regulatory signals' and 'activating sequences'. These expressions referred to the experiments that revealed their role. But an enhancer is not *per se* a 'signal' or

an 'activating sequence'. Molecular biology mostly remained one-dimensional.

The subsequent advent of transcription factors changed the vision of enhancers, but generated other difficulties. Since these factors can bind to promoters as well as to enhancers, and since the distance between enhancers and promoters is highly variable, but can be reduced to zero, the physiological significance of enhancers faded, and the use of the term became purely empirical – based on the distance of the regulatory sequences from the transcription initiation site.

Many of current research programmes such as the acrimoniously debated ENCODE (Graur *et al.* 2013) are precisely that – empirical – which in part explains the importance now accorded to the search for enhancers.

The rationale for their existence has progressively (but not fully) shifted from a mechanistic to an evolutionary one. Most enhancers have been conserved during evolution, which means that their presence gives a positive selective value. But the way in which chromatin landscape restricts enhancers to correct targets is yet to be fully understood.

Acknowledgements

I am indebted to David Marsh and the two anonymous reviewers for their critical reading of the manuscript.

References

Banerji J, Rusconi S and Schaffner W 1981 Expression of a β-globin gene is enhanced by remote SV40 DNA sequences. *Cell* **27** 299–308

Banerji J, Olson L and Schaffner W 1983 A lymphocyte-specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. *Cell* **33** 729–740

Benoist C and Chambon P 1980 Deletions covering the putative promoter region of early mRNAs of simian virus 40 do not abolish T-antigen expression. *Proc. Natl. Acad. Sci. USA* **77** 3865–3869

Bishop JM 1981 Enemies within: the genesis of retrovirus oncogenes. *Cell* **23** 5–6

Blackwood EM and Kadonaga JT 1998 Going the distance: a current view of enhancer action. *Science* **281** 60–63

Carroll SB 2008 Evo-Devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* **134** 25–36

Champoux JJ 1978 Proteins that affect DNA conformation. *Annu. Rev. Biochem.* **47** 449–479

Chandler VL, Maier BA and Yamamoto KR 1983 DNA sequences bound specifically by glucocorticoid receptor in vitro render a heterologous promoter hormone responsive in vivo. *Cell* **33** 489–499

de Laat W and Duboule D 2013 Topology of mammalian developmental enhancers and their regulatory landscapes. *Nature* **502** 499–506

de Villiers J and Schaffner W 1981 A small segment of polyoma virus DNA enhances the expression of a cloned β-globin gene over a distance of 1400 base pairs. *Nucleic Acids Res.* **9** 6251–6264

de Villiers J, Olson L, Tyndall C and Schaffner W 1982 Transcriptional 'enhancers' from SV40 and polyoma virus show a cell type preference. *Nucleic Acids Res.* **10** 7965–7976

Garabedian MJ, Shepherd BM and Wensink PC 1986 A tissue-specific transcription enhancer from the *Drosophila* yolk protein 1 gene. *Cell* **45** 859–867

Gillies SD, Morrison SL, Oi VT and Tonegawa S 1983 A tissue-specific transcription enhancer element is located in the major intron of a rearranged immunoglobulin heavy chain gene. *Cell* **33** 717–728

Graur D, Zheng Y, Price N, Azevedo RBR, Zufall RA and Elhaik E 2013 *Genome Biol. Evol.* **5** 578–590

Grosschedl R and Birnstiel ML 1980 Spacer DNA sequences upstream of the T-A-T-A-A-A-T-A sequence are essential for promotion of H2A histone gene transcription *in vivo*. *Proc. Natl. Acad. Sci. USA* **77** 7102–7106

Grosschedl R, Weaver D, Baltimore D and Constantini F 1984 Introduction of a μ immunoglobulin gene into the mouse germ line: specific expression in lymphoid cells and synthesis of functional antibody. *Cell* **38** 647–658

Gruss P, Dhar R and Khoury G 1981 Simian virus 40 tandem repeated sequences as an element of the early promoter. *Proc. Natl. Acad. Sci. USA* **78** 943–947

Hayward WS, Neel BG and Astrin SM 1981 Activation of a cellular *onc* gene by promoter insertion in ALV-induced lymphoid leukosis. *Nature* **290** 475–480

Katinka M, Yaniv M, Vasseur M and Blangy D 1980 Expression of polyoma early functions in mouse embryonal carcinoma cells depends on sequence rearrangements in the beginning of the late region. *Cell* **20** 393–399

Khoury G and Gruss P 1983 Enhancer elements. *Cell* **33** 313–314

Kleinjan DA and Lettice LA 2008 Long-range gene control and genetic diseases. *Adv. Genet.* **61** 339–388

Laimins LA, Khoury G, Gorman C, Howard B and Gruss P 1982 Host-specific activation of transcription by tandem repeats from simian virus 40 and Moloney murine sarcoma virus. *Proc. Natl. Acad. Sci. USA* **79** 6453–6457

Levine M 2010 Transcriptional enhancers in animal development and evolution. *Curr. Biol.* **20** R754–R763

Levine M, Cattoglio C and Tjian R 2014 Looping back to leap forward: transcription enters a new era. *Cell* **157** 13–25

Maniatis T, Goodbourn S and Fischer JA 1987 Regulation of inducible and tissue-specific gene expression. *Science* **236** 1237–1245

Moreau P, Hen R, Wasylyk B, Everett R, Gaub MP and Chambon P 1981 The SV40 72 base repair repeat has a striking effect on gene expression both in SV40 and other chimeric recombinants. *Nucleic Acids Res.* **9** 6047–6068

Müller MM, Gerster T and Schaffner W 1988 Enhancer sequences and the regulation of gene transcription. *Eur. J. Biochem.* **176** 485–495

Neuberger MS 1983 Expression and regulation of immunoglobulin heavy chain gene transfected into lymphoid cells. *EMBO J.* **2** 1373–1378

Orom UA and Shiekhattar R 2013 Long noncoding RNAs usher in a new era in the biology of enhancers. *Cell* **154** 1190–1193

Pennacchio LA, Ahituv N, Moses AM, Prabhakar S, Nobrega MA, *et al*. 2006 *In vivo* enhancer analysis of human conserved noncoding sequences. *Nature* **444** 499–502

Pennacchio LA, Bickmore W, Dean A, Nobrega MA and Bejerano G 2013 Enhancers: five essential questions. *Nat. Rev. Genet.* **14** 288–295

Queen C and Baltimore D 1983 Immunoglobulin gene transcription is activated by downstream sequence elements. *Cell* **33** 741–748

Sakabe NJ, Savic D and Nobrega MA 2012 Transcriptional enhancers in development and disease. *Genome Biol.* **13** 238

Sanyal A, Lajoie BR and Dekker J 2012 The long-range interaction landscape of gene promoters. *Nature* **489** 109–113

Saragosti S, Moyne G and Yaniv M 1980 Absence of nucleosomes in a fraction of SV40 chromatin between the origin of replication and the region coding for the late leader RNA. *Cell* **20** 65–73

Scott WA and Wigmore DJ 1978 Sites in simian virus 40 chromatin which are preferentially cleaved by endonucleases. *Cell* **15** 1511–1518

Spandidos DA and Wilkie NM 1983 Host-specificities of papillomavirus, Moloney murine sarcoma virus and simian virus 40 enhancer sequences. *EMBO J.* **2** 1193–1199

Struhl K 1981 Deletion mapping a eukaryotic promoter. *Proc. Natl. Acad. Sci. USA* **78** 4461–4465

Struhl K 1982 The yeast *his3* promoter contains at least two distinct elements. *Proc. Natl. Acad. Sci. USA* **79** 7385–7389

Varshavsky AJ, Sundin OH and Bohn MJ 1978 SV40 viral minichromosome: preferential exposure of the origin of replication as probed by restriction endonucleases. *Nucleic Acids Res.* **5** 3469–3477

Vasseur M, Kress C, Montreau N and Blangy D 1980 Isolation and characterization of polyoma virus mutants able to develop in embryonal carcinoma cells. *Proc. Natl. Acad. Sci. USA* **77** 1068–1072

Weiher H, König M and Gruss P 1983 Multiple point mutations affecting the simian virus 40 enhancer. *Science* **219** 626–631

Weintraub H and Groudine M 1976 Chromosomal subunits in active genes have an altered conformation. *Science* **193** 848–856

Winter RB, Berg OG and von Hippel PH 1981 Diffusion-driven mechanisms of protein translocation on nucleic acids. 3. The *Escherichia coli* lac repressor-operator interaction: kinetic measurements and conclusions. *Biochemistry* **20** 6961–6977

Wu C, Bingham PM, Livak KJ, Holmgren R and Elgin SCR 1979 The chromatin structure of specific genes: I. Evidence for higher order domains of defined DNA sequence. *Cell* **16** 797–806