seminars in **VIROLOGY**, Vol 4, 1993 : pp 3-13

The SV40 enhancer: transcriptional regulation through a hierarchy of combinatorial interactions

Winship Herr

The simian DNA tumor virus SV40 regulates synthesis of early and late viral transcripts from two divergent and overlapping promoters. The SV40 early promoter exhibits all the hallmarks of eukaryotic promoters, containing elements that function when positioned near the start site of transcription and other positionally flexible elements that constitute a transcriptional enhancer. The architecture of the SV40 enhancer has been elucidated by four complementary approaches: (1) saturation mutagenesis in vitro; (2) viral genetic selection in vivo; (3) synthetic multimerization of its component parts; and (4) detailed protein binding studies. The SV40 enhancer represents a cluster of cell-specific regulatory elements that cooperate with one another in a functional hierarchy. The basic units of enhancer structure, enhansons, are the building blocks and represent protein binding sites. The functional units are proto-enhancers, which consist of one or more enhansons. Proto-enhancers serve as regulatory modules that can cooperate with one another or duplicates of themselves in flexible arrangements to create a potent enhancer.

Key words: SV40 / proto-enhancer / enhanson / enhancer / transcription

VIRUSES ARE excellent tools for examining the mechanisms by which gene expression is regulated within the viral host. Features that make the study of viruses advantageous include smaller genomes than those of the host cells, regulatory programs that may unfold rapidly during a relatively short infectious cycle, and genetic selection for altered viral growth properties, which can result in changes in the regulatory apparatus. Because each virus depends, albeit to differing extents, on the host cell regulatory machinery to control its own gene expression, understanding viral gene regulation leads to an understanding of cellular gene regulation.

In mammalian cells, papovaviruses such as the murine polyomavirus and simian SV40 virus have served as excellent probes to study the structure and arrangements of *cis*-acting elements in mammalian promoters. Indeed, the amazing ability of enhancers to activate transcription when positioned thousands of base pairs away from the transcriptional start site was discovered in studies of the SV40 early promoter.[1,2] When viruses like SV40 infect a cell, they are entirely dependent on the host cell machinery to initiate transcription of the viral genome. Once transcription is initiated, the early viral gene products, such as the large T antigen of SV40, can alter the patterns of gene expression resulting in the case of SV40 in reduced early gene expression and elevated late gene expression.[3] Because the early promoter is entirely dependent on cellular transcription factors to direct transcription, its structure is an accurate reflection of the structure of cellular promoters, with one exception. Whereas in cellular promoters regulatory elements such as enhancers can be located very far from the start site of transcription, in SV40, which contains a small genome, the various regulatory elements are all located within a few hundred base pairs of the transcriptional start site.

This chapter describes the structure and functional properties of the SV40 enhancer. This enhancer is one of the best characterized complex regulatory elements to activate transcription in mammalian cells. Its characterization resulted from four complementary experimental approaches: (1) a detailed and systematic point mutagenesis of the entire SV40 regulatory region;[4] (2) viral genetic selection for improved enhancer function;[5,6] (3) synthetic multimerization of its component parts;[7-10] and (4) detailed protein binding studies.[11-15] Systematic point mutagenesis, together with the protein binding studies, identified sequence motifs required for enhancer function. These motifs subsequently became known as enhansons, because they represent the basic units or building blocks of the enhancer. In contrast, viral genetic selection and multimerization of enhancer elements revealed the existence of subenhancer modules that display little activity on their own but can cooperate with one another or duplicates of themselves

From the Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA

1044-5773/93/010003 + 11$8.00/0

to create an effective enhancer. These modules are referred to as proto-enhancers and may consist of one or two enhansons. A surprising feature of proto-enhancers is that, over a span of about 100 bp, they can cooperate with one another to create an effective enhancer with little constraint on the precise spacing between paired proto-enhancers.

Below I describe how the different experimental approaches led to our current appreciation of the structure of the SV40 enhancer and how the elucidation of this structure contributed to our understanding of transcriptional regulation.

The SV40 early promoter

Figure 1 illustrates features of the SV40 genome and regulatory region. SV40 contains a relatively small double-stranded circular DNA genome of about 5200 bp.[3] The exact size varies depending on the presence or absence of a variably sized tandem duplication within the enhancer region; the prototypical 776 strain of SV40 contains a 72 bp duplication as shown in Figure 1, but other variants either contain no enhancer duplication or duplications ranging in size from 64 to 93 bp (see ref 16 for references). Emanating from the regulatory region are two divergent transcription units, one early (E) and one late (L), each encompassing roughly half of the genome. The respective early and late promoters overlap and share *cis*-acting elements (see ref 17). Because the late stages of SV40 infection are more complicated than the early stages, the late promoter is less well characterized than the early promoter.

Figure 2 shows in more detail the structure of an SV40 early promoter lacking any enhancer duplication together with the nucleotide sequence of the enhancer region. As is typical of promoters in general, the SV40 early promoter contains elements that are involved in positioning the start site of transcription and others that only regulate the levels of transcription. The first category includes the transcriptional initiation sites and the TATA box,

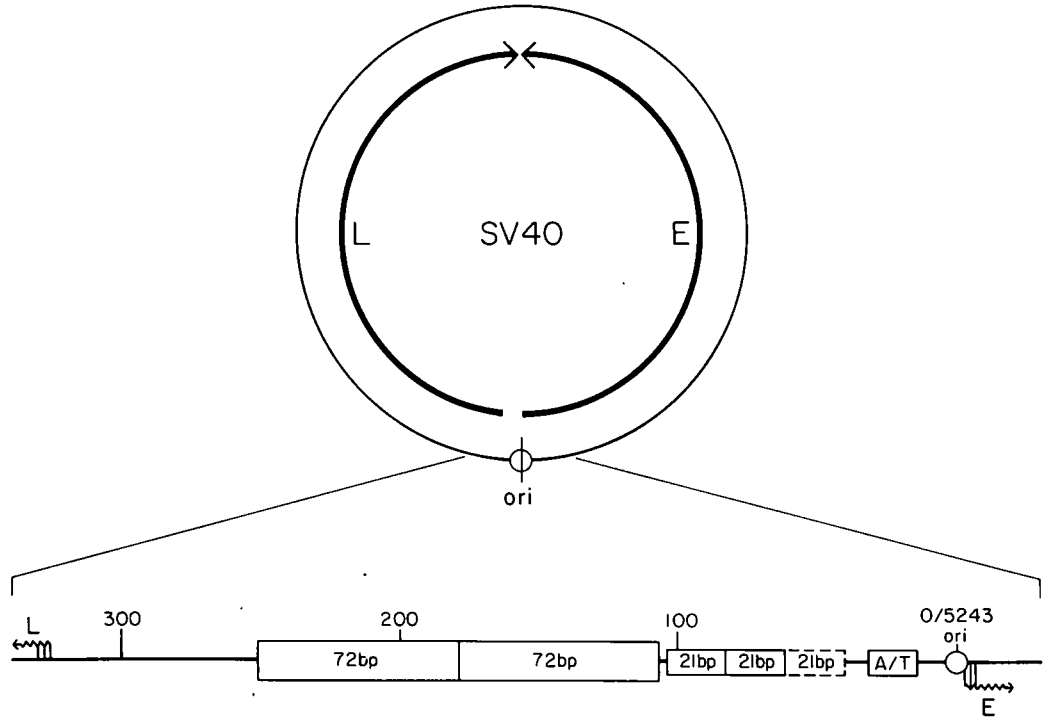

Figure 1. The regulatory region of the SV40 genome. A representation of the circular SV40 genome shows the positions of the early (E) and late (L) transcription units. Below wavy arrows indicate the positions of the major early and late transcription start sites on an enlargement of the regulatory region. Nucleotide coordinates given are for the strain 776 SV40 genome which contains 5243 bp. The center of the origin of replication (ori) is indicated by a circle, whereas the 17 bp A:T rich segment containing at least two TATA elements (A/T), and the 21 bp and 72 bp repeats that form proximal promoter and enhancer regions are enclosed by rectangles.

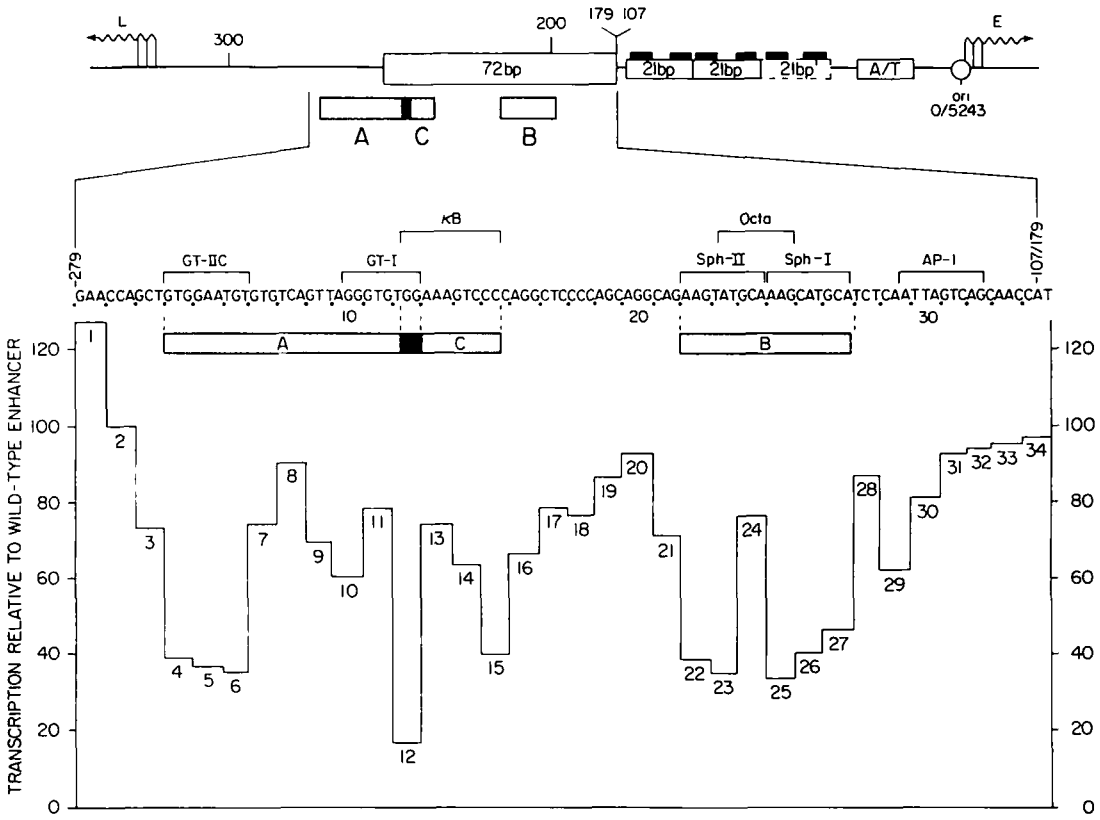

Figure 2. Structure of the SV40 enhancer region. (Top) The SV40 early promoter shown is similar to that shown in Figure 1 except that it contains only one 72 bp element and includes the positions of SP1 binding sites, indicated by black boxes. (Middle) Bracketed above the nucleotide sequence of the SV40 enhancer region are the positions of enhansons; the positions of the three A, B, and C proto-enhancers are indicated by the boxes below the sequence. The A and C proto-enhancers overlap by two nucleotides as indicated by the solid shading. (Bottom) The bar graph shows the relative activity in HeLa cells of the pA series of 34 tandem triple point mutations as determined by Zenke *et al.*[4] The three base pairs mutated in each pA mutant are indicated by the series of dots below the nucleotide sequence of the enhancer region; the positions of pA10, pA20, and pA30 mutations are indicated. Note that the first mutation pA1 results in a 25% increase in activity relative to the wild-type enhancer, whereas the pA2 mutation has no effect.

which serves as a binding site for the TATA-box binding protein (TBP) component of the general transcription factor complex TFIID (reviewed in ref 18). In SV40, both the early and late promoters contain multiple start sites. In the SV40 early promoter, a cluster of multiple start sites arise in part because there are at least two TATA boxes within a stretch of 17 A : T base pairs (the A/T box), which form part of the overlapping SV40 origin of DNA replication. The preferential use of the TATA boxes is affected by the spacing between them and upstream regulatory elements.[19]

The second class of promoter elements, the regulatory elements, can also be divided into two groups depending on the positional flexibility they display with respect to the transcriptional start site. In the SV40 early promoter, two perfect (21 bp) and one imperfect (21 bp') 21 bp repeats serve as proximal promoter elements and contain binding sites (black boxes) for the zinc-finger transcription factor Sp1.[20,21] Except under unusual circumstances,[22] these proximal promoter elements, unlike enhancers, do not activate transcription when positioned at a distance from the transcriptional start site. Immediately upstream of the 21 bp repeats lies the enhancer. Although in the prototypic SV40 strain the enhancer contains a tandem 72 bp duplication, studies of the SV40 enhancer have focused on the

structure of a non-duplicated enhancer because its activity is only a fewfold less than that of the duplicated form, and it is easier to make site-directed point mutations in a sequence lacking a tandem duplication.

The SV40 enhancer contains multiple sequence motifs involved in enhancer function

To analyze the structure and function of the SV40 enhancer, Chambon and colleagues scanned the entire SV40 enhancer region with a series of 34 tandem 3 bp mutations, and assayed the ability of each mutant to activate transcription.[4] Figure 2 shows the relative activity of each triple point mutant in a transient expression assay in the human cervical carcinoma cell line HeLa. Surprisingly, although the SV40 enhancer augments the activity of the SV40 early promoter by 100- to 1000-fold, any one set of three point mutations only reduced transcription by at most sixfold. This result suggested that SV40 enhancer function is not the result of or dependent on any one regulatory element but instead results from the combined activities of multiple elements. A search for reiterated sequences within the SV40 enhancer or sequence similarities between the SV40 enhancer and other enhancers revealed a series of sequence motifs about 10 bp in length.[4] Those sequence motifs that correlate with positions that are sensitive to point mutagenesis are indicated above the sequence of the SV40 enhancer in Figure 2. Together, the effects of the point mutations and the identification of sequence motifs led to a picture of the SV40 enhancer in which many individual sequence motifs work together to constitute a fully active enhancer, but the relationship between motifs was unknown.

The SV40 enhancer contains multiple elements that can functionally substitute for one another

In parallel with the point mutagenesis study of Zenke *et al*,[4] Y. Gluzman, J. Clarke, and I showed, using viral genetic selections, that the SV40 enhancer is composed of at least three separate elements or modules. Each of these can cooperate with the others or duplicates of itself to enhance transcription. These experiments originated from an analysis of the significance of two 8 bp stretches of alternating purines and pyrimidines (Pu/Py), one just upstream of the 72 bp element (ATGTGTGT) and the other within the 72 bp element (GCATGCAT) which were hypothesized to form Z DNA.[23] Although Z DNA formation is probably not important for SV40 enhancer function, a combination of double point mutations within each of the two Pu/Py segments debilitated SV40 enhancer function in a transient assay and debilitated SV40 viral growth.[5] Passage of mutant viral stocks, however, resulted in viral isolates with improved growth properties. Characterization of 18 independent isolates revealed consistent duplication of the mutated enhancer region; these duplications were responsible for the improved growth and resulted in improved enhancer function.[5]

Figure 3 shows the region that was duplicated in each of the 18 growth revertants which varied from 45 bp (in revertant duplication rd45) to 135 bp (in rd135) in length. In each case, other than a tandem duplication, the nucleotide sequence of the enhancer, including the deleterious point mutations, was unaltered. Examination of the new sequence created by the junction between each tandem duplication did not reveal any consistent sequence that could suggest that the junction sequences were responsible for restoring activity; in particular they did not recreate sequences resembling the Pu/Py sequences that were originally mutated. The mutated sequences themselves were not consistently duplicated, suggesting that amplification of weakened elements was also not responsible for the restored activity. Instead, the most striking feature of the revertant duplications was that a central 15 bp sequence (the stippled region in Figure 3) was contained in all 18 duplications. This 15 bp sequence spans a 10 bp element shared with other viral and cellular enhancers that has subsequently been termed the κB motif.[24,25] The consistent pattern of duplication of the central 15 bp sequence suggested that duplication of this 'core' region could compensate for the loss of function caused by the Pu/Py point mutations.[5]

In a reciprocal experiment, in which the κB motif was mutated and the Pu/Py sequences were left intact, SV40 growth was debilitated and subsequent revertants contained tandem sequence duplications that consistently encompassed either one or both of the regions containing the wild-type Pu/Py segments but not necessarily the mutated κB region.[6] This and other viral selections[26] showed that duplication of any one of three distinct regions within the SV40 enhancer, called A, B, and C, can compensate for loss of function in the other two regions. Thus, the

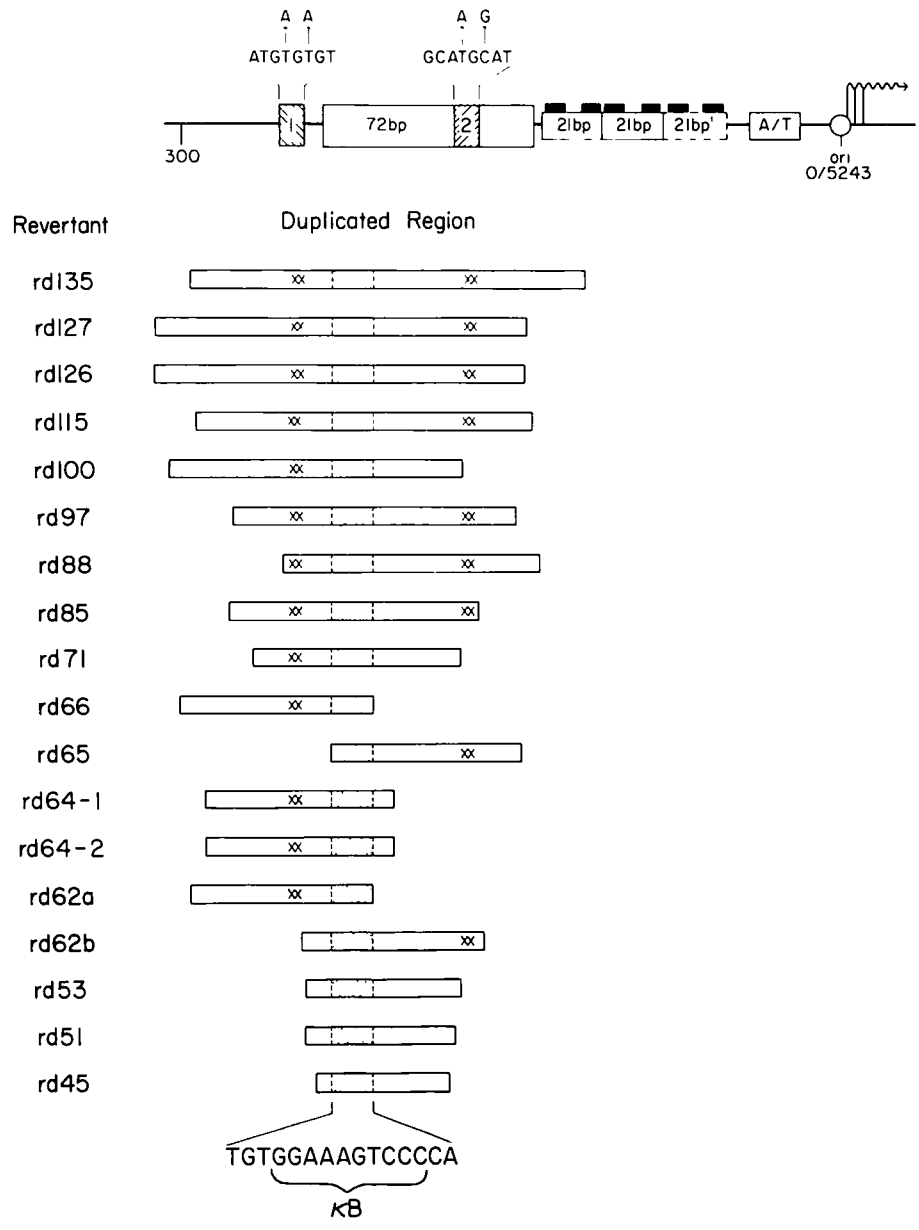

Figure 3. Tandem sequence duplications present in 18 viral revertants of an SV40 variant mutated in the enhancer region Pu/Py segments.[5] The two sets of Pu/Py double point mutations are shown above the diagram of the SV40 early promoter. Each revertant duplication (rd) is differentiated by the size of the duplicated sequence in bp. Two different 62 bp duplications (rd62a and rd62b) and two independent but identical 64 bp duplications (rd64-1 and rd64-2) were obtained. The sequence that is duplicated in each revertant enhancer is indicated by the position and size of the open bars. Within the open bars, the XXs indicate the presence of the original point mutations, and the stippled portion indicates the 15 bp region that is shared by all of the duplications. This 15 bp sequence shown at the bottom of the diagram contains the 10 bp κB proto-enhancer (bracketed).

SV40 enhancer contains distinct modules that do not function effectively on their own but can either cooperate with one another or duplicates of themselves to enhance transcription effectively. The heterogeneous size of the sequence duplications in the viral revertants indicated that these modules can interact with one another without strict spacing requirements between modules.

The A, B, and C regions contain independent enhancer elements

Although the pattern of revertant duplication suggested that the SV40 enhancer contains multiple elements, they did not precisely map the boundaries of the elements nor did they address whether these elements are fully independent of other sequences within the SV40 enhancer. To address these points we and Schaffner and colleagues used a strategy first described by Veldman *et al*[27] in which an enhancer is created by multimerization of synthetic oligonucleotides.[7,8] Figure 4 shows the reporter construct containing the human β-globin transcription unit that we used to assay enhancer activity, together with the sequences of the A, B, and C regions that were commonly duplicated in the different revertant duplication patterns. Multimerization of 17 bp sequences for the B and C regions or a 21 bp sequence for the A region resulted in potent enhancers when assayed in the SV40 host cell line CV-1.[7] In the one case where it was tested, the B element, two or more copies were required for effective activity,[7] consistent with the duplicated structure of the viral revertants. These results showed that A, B, and C regions do represent independent enhancer elements that can function autonomously when present in multiple copies. These elements have subsequently been referred to as proto-enhancers,[10] because they do not possess enhancer activity on their own but can readily form an enhancer when present in two or more copies.

The synthetic enhancers were invaluable for fine structure mapping of the proto-enhancers. First, it was possible to study the structure of one proto-enhancer in isolation of the others. And, second, whereas the original Pu/Py and κB mutations within the A, B, and C proto-enhancers had only three- to fivefold effects in the context of the entire SV40 enhancer, they reduced activity to undetectable levels when incorporated into each copy of a multimerized enhancer.[7] Figure 2 shows the boundaries of the A, B, and C proto-enhancers as deduced from such

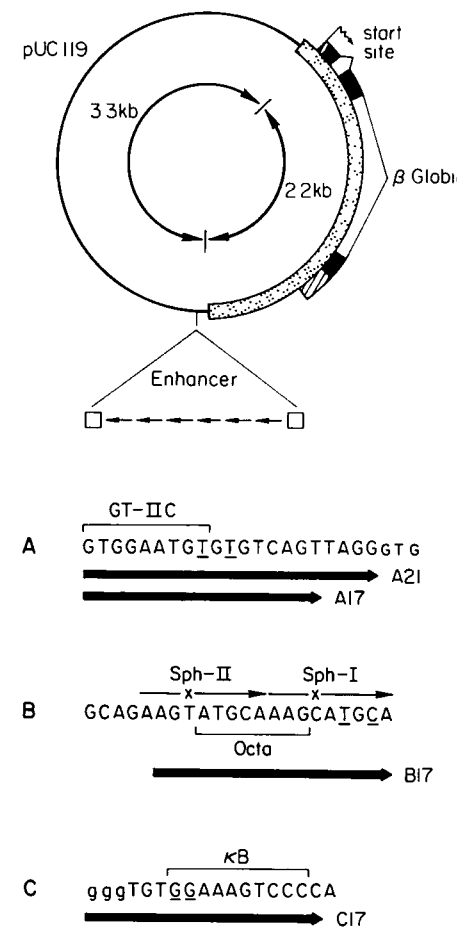

Figure 4. Structure of synthetic multimerized enhancers and sequence of commonly duplicated regions in SV40 enhancer revertants. (Top) The structure of the reporter construct used to assay enhancer function. Multimerized enhancers (usually six tandem copies) were inserted downstream of the 2.3 kb human β-globin transcription unit cloned into pUC119 such that the enhancer lies 2.2 kb downstream and 3.3 kb upstream of the β-globin transcriptional start site. (Bottom) The commonly duplicated A, B, and C sequences (in uppercase letters) in the revertant enhancers. For example, the 15 bp upper case sequence of the C element corresponds to the 15 bp commonly duplicated sequence shown in Figure 3. The adjacent lower case ggg sequence represents flanking SV40 enhancer sequence that was included in the synthetic C17 oligonucleotide. The precise 5′ and 3′ boundaries of the original 21 bp A element sequence was ambiguous due to a 3 bp GTG terminal redundancy which is indicated by the smaller uppercase GTG sequence. The position of enhansons identified in Figure 2 are defined, and the sequences contained in the synthetic oligonucleotides A21, A17, B17, and C17 are indicated by the solid arrows. The position of the single base pair difference between the Sph-I and Sph-II enhansons is marked by an X within each broken arrow.

studies.[7,9,10,28,29] The proto-enhancers lie within the boundaries of the commonly duplicated regions in the revertants, except for the A proto-enhancer, which, for reasons that are not entirely clear, extends several base pairs 3′ of the commonly duplicated A region (compare Figures 2 and 4). The positions of the three proto-enhancers correspond to regions that were sensitive to point mutagenesis in the triple point mutation scan by Zenke *et al.*[4] Indeed the strong effect of the pA12 mutation can be explained by the two base pair overlap of the A and C proto-enhancers (refs 10, 29, 30; M. Tanaka, W. Herr, unpublished results).

The A, B, and C proto-enhancers possess different cell-specific activities

Although, as expected from the viral revertant studies, all three of the synthetic A, B, and C proto-enhancers were active in CV-1 cells, when assayed in other cell types, they each displayed a unique pattern of cell-specific activity.[7,8] This result suggested that each proto-enhancer responds to different sets of cell-specific transcription factors. Consistent with these results, the pattern of inactivation by the series of triple point mutations shown in Figure 2 also varies with cell type.[11,30] These studies further showed that the SV40 enhancer contains proto-enhancers that are silent in CV-1 cells but are active in other cell types. This second category of proto-enhancers includes the octamer motif (labeled Octa in Figure 2), which is active in B lymphocytes.[10,11,28] Thus, the SV40 enhancer represents a cluster of regulatory elements, each responding to different cell-specific factors and only some displaying activity in CV-1 cells.

The arrangement of cell-specific proto-enhancers can explain narrow and broad patterns of transcriptional regulation by enhancers

The properties of proto-enhancers, i.e. being inactive on their own and displaying cell-specific activities, afforded a mechanism by which the combinatorial arrangement of proto-enhancers could lead to both narrow and broad patterns of gene expression[7] as illustrated in Figure 5. This figure shows four hypothetical arrangements of two proto-enhancers, A and B, that are recognized by the cell-specifically expressed enhancer factors α and β, respectively. The

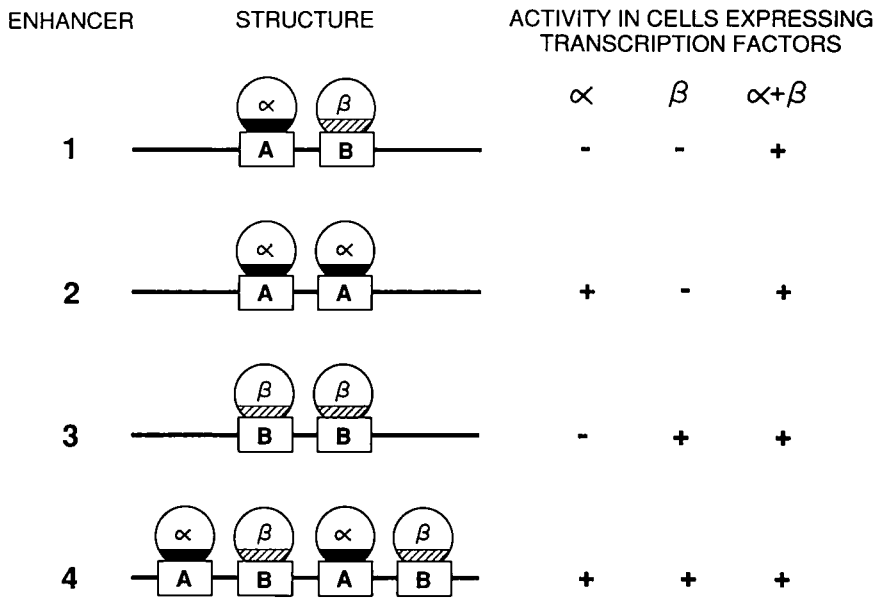

Figure 5. Model predicting the cell type-specific activities of modular enhancers composed of different proto-enhancer arrangements. The two proto-enhancers A and B are recognized by the cell-specific factors α and β, respectively. The activities of the four different arrangements of the A and B proto-enhancers are shown in cells expressing either α alone, β alone, or both factors together.

first enhancer is only active in cells expressing both α and β, because the enhancer contains only a single copy of each of the A and B proto-enhancers, and therefore both proto-enhancers must be active for this enhancer to stimulate transcription. Enhancers 2 and 3, however, contain duplicate copies of one proto-enhancer and are therefore active in cells expressing the α and β factors together and cells expressing only the corresponding α or β factor. The last enhancer, Enhancer 4, contains two copies of both the A and B proto-enhancers and therefore will display activity in all cell types shown. Thus simple tandem duplication of an enhancer containing multiple proto-enhancers has the potential to create a broadly active enhancer (Enhancer 4) from an enhancer (Enhancer 1) with restricted activity.

The SV40 enhancer contains distinct levels of organization

The results from scanning the SV40 enhancer with 3 bp mutations showed that the series of sequence motifs identified in Figure 2 is important for enhancer function.[4] The SV40 growth revertant selections, however, together with the synthetic multimerized enhancers, indicated that these sequence motifs are not sufficient for proto-enhancer function. For example, multimerization of the 17 bp subregion of the A proto-enhancer indicated in Figure 4, which encompasses the entire GT-IIC motif (also referred to as the core A motif in ref 9), did not result in an active enhancer in any cell type tested.[7] An unusual pattern of sequential sequence duplication arising in viral growth revertants of an SV40 enhancer mutated in all three A, B, and C proto-enhancers suggested, however, that a proto-enhancer can be created by tandem duplication of a sequence motif. These patterns of sequence duplication involved a larger duplication of a prior smaller duplication. For example, a 9 bp tandem duplication of the sequence TGTGGAATG, nearly exactly corresponding to the 9 bp GT-IIC motif GTGGAATGT, was subsequently duplicated by larger variably sized duplications in each of three independent revertants of the triple ABC mutant.[6] This pattern of double duplication suggested that the first duplication created a new proto-enhancer, and the second duplication created two copies of this new proto-enhancer.

Analysis of the activity of synthetic multimerized proto-enhancers revealed that this is indeed the case. Within the wild-type SV40 enhancer, the GT-IIC motif cooperates with the GT-I motif to create the A proto-enhancer (see Figure 2),[10] but duplication of the GT-IIC motif alone can also create a functional proto-enhancer.[9,10] Thus, the A proto-enhancer is bipartite, containing heterologous subunits, one of which can cooperate with itself when duplicated. The B proto-enhancer is also bipartite, consisting of the tandemly duplicated, albeit imperfect, Sph motifs (see Figure 4). Not all proto-enhancers are as clearly bipartite, however, as in the case of the octamer and κB proto-enhancers.[10,28,29]

The different structures of the proto-enhancers led to a picture of the SV40 enhancer in which a variety of short sequence motifs, serving as building blocks for the enhancer, come together to cooperate with one another at different levels of a functional hierarchy. These building blocks were termed enhansons, to represent the basic units of enhancer structure.[9] The different ways in which the building blocks come together to form proto-enhancers, led to their classification as A, B, and C type enhansons.[10] Class A enhansons display self-cooperativity as well as the ability to cooperate with heterologous enhansons in the formation of proto-enhancers; this class includes the GT-IIC enhanson. Class B enhansons do not display self-cooperativity; this class includes the GT-I enhanson, which cooperates with the GT-IIC enhanson to create the A proto-enhancer. In contrast, Class C enhansons, which include the κB, octamer, and AP-1[31] enhansons display proto-enhancer activity independently of a second enhanson. For example, the κB enhanson represents the complete SV40 C proto-enhancer (see Figure 2).

Enhansons represent protein binding sites

Consistent with their indivisible, basic unit structure, enhansons represent protein binding sites. Indeed the mapping of protein binding sites that corresponded to the cell-specific activities of enhansons aided greatly in the mapping of these *cis*-regulatory elements.[11-15] Figure 6 depicts the SV40 enhancer region together with proteins that bind to the different SV40 enhansons: TEF-1, TEF-2, NF-κB, Oct-1 and Oct-2, and AP-1. TEF-2,[14] the GT-I enhanson-binding protein, is the only one of these enhancer proteins for which cDNAs have not been obtained. Although TEF-2 and the GT-IIC enhanson-binding protein TEF-1[14,32] cooperate with one another to enhance transcription through the A proto-enhancer, they do not display cooperative DNA

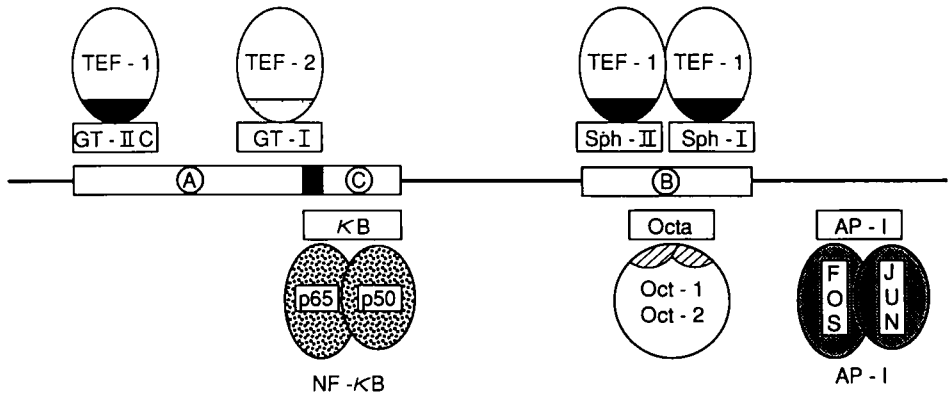

Figure 6. SV40 enhanson-binding proteins. The position of enhansons and their respective enhanson-binding proteins are shown above and below a diagram of the A, B, and C proto-enhancers. Enhansons that cooperate with a second enhanson to create a proto-enhancer (e.g. GT-IIC and GT-I) are shown above the position of the proto-enhancers, whereas enhansons that display proto-enhancer activity on their own (e.g. κB) are shown below.

binding.[14] This result suggests that TEF-1 and TEF-2 cooperate with each other subsequent to binding to the A proto-enhancer, for example, in interactions with other components of the transcriptional apparatus. In contrast, TEF-1 molecules cooperate with one another in binding to appropriately spaced binding sites as in the case of the duplicated GT-IIC enhanson.[14] TEF-1 also possesses a striking ability to recognize dissimilar nucleotide sequences, which results in its ability to bind to both the GT-IIC and Sph enhansons.[14] Thus, the GT-IIC and Sph enhansons, although different at the level of nucleotide sequence, are functionally analogous.

The remaining three enhanson-binding factors shown in Figure 6 activate transcription through proto-enhancers composed of single enhansons. Curiously, however, in each case these enhanson-binding proteins display bipartite characteristics, suggesting that the underlying proto-enhancer structure may also actually be bipartite and consist of two enhansons. For example, the κB enhanson-binding factor NF-κB[25] is a heterodimer composed of p50 and p65 subunits that are related to the proto-oncogene product c-Rel (reviewed in ref 33), and the AP-1 enhanson-binding protein AP-1 is generally a heterodimer composed of one member of the Jun family and one member of the Fos family of leucine zipper proteins (reviewed in ref 34). Furthermore, the lymphoid enhancer factor Oct-2, which like Oct-1 binds the SV40 octamer enhanson as a monomer, contains a bipartite DNA binding domain consisting of the two subsegments of the POU domain, the POU-specific region and the POU homeodomain, and a bipartite activation domain (reviewed in ref 35). Thus, it may be that the different halves of these binding sites should be considered as separate enhansons. For example, variants of the κB motif, in which only one half of the motif is changed, can differentially bind p50 homodimers or p50/p65 heterodimers.[33]

Implications of a hierarchical binary regulatory structure

The structure of the SV40 enhancer emphasizes the use of multiple levels of binary organization to regulate transcription: enhansons creating proto-enhancers, and proto-enhancers creating enhancers. These levels probably evolved because they afford increased flexibility in establishing distinct programs of transcriptional regulation with a limited set of transcription factors. As was shown in Figure 5, different combinations of two proto-enhancers and their corresponding regulatory proteins, can yield four different patterns of cell-specific activity. If the proto-enhancers can now be subdivided into separate enhansons, then the number of different combinations of enhansons and consequently different patterns of activity becomes even larger.

In addition to providing a greater degree of flexibility in transcriptional control, there may have been an evolutionary selection for complex enhancer structures, because such structures may serve to restrict the spontaneous appearance of enhancers by mutation. Unlike proximal promoter elements,

enhancers can activate transcription over very large distances from the transcriptional start site. Nonetheless, transcription factors generally recognize degenerate sets of sequences only about 8-10 bp long. Thus, a structurally simple enhancer, perhaps composed of a single 8-10 bp binding site, might arise frequently due to mutation, all too readily modifying patterns of gene expression from sites distant from the transcriptional start site. But instead, as illustrated by the SV40 enhancer, enhancers require interactions between multiple enhancer-bound transcription factors to stimulate transcription. In this manner, indiscriminate appearance of enhancers is suppressed and yet flexibility in their creation from a limited set of building blocks is enhanced.

Another outcome of the relatively short range interactions between proto-enhancers, compared to the much longer range interactions between an enhancer and the transcriptional start site, is that multiple complex enhancers, each with a distinct cell-specific activity, can regulate transcription of a single gene. In this view, the enhancers represent regulatory islands, buffered from each other's activity, but all capable of communicating with the same proximal promoter complex.

Prospects

The major future challenge in understanding transcriptional regulation by the SV40 enhancer is to provide a molecular description of how the hierarchy of interactions among the different regulatory elements is achieved. For example, one of the most important properties of the SV40 proto-enhancers is the positional flexibility they display without detriment to their ability to cooperate with one another. This flexibility permits the association of many different proto-enhancers within a complex enhancer in a way such that they can all interact with one another; such associations would be difficult, if not impossible, if there existed strict spacing requirements between proto-enhancers. How is such flexibility achieved? One attractive model suggests that proteins bound to the proto-enhancers do not interact directly with one another but instead interact indirectly through binding to the basal transcriptional machinery at the site of transcriptional initiation. As we begin to understand in greater detail the structure and function of transcription factor activation domains, as well as the components of the basal transcriptional apparatus, the mechanisms by which the flexible proto-enhancer interactions are achieved will come into focus.

Acknowledgements

I am indebted to my past and present colleagues who contributed to the studies described. I thank M. Cleary, N. Hernandez, W. Phares, and M. Tanaka for critical comments on the text and J. Reader and J. Duffy for help in preparing the manuscript and figures, respectively.

References

1. Banerji J, Rusconi S, Schaffner W (1981) Expression of a β-globin gene is enhanced by remote SV40 DNA sequences. Cell 27:299-308
2. Moreau P, Hen R, Wasylyk B, Everett R, Gaub MP, Chambon P (1981) The SV40 72 base pair repeat has a striking effect on gene expression both in SV40 and other chimeric recombinants. Nucl Acids Res 9:6047-6068
3. Tooze J (1980) DNA Tumor Viruses: Molecular Biology of Tumor Viruses. Cold Spring Harbor Laboratory, New York
4. Zenke M, Grundström T, Matthes H, Wintzerith M, Schatz C, Wildeman A, Chambon P (1986) Multiple sequence motifs are involved in SV40 enhancer function. EMBO J 5:387-397
5. Herr W, Gluzman Y (1985) Duplications of a mutated simian virus 40 enhancer restores its activity. Nature 313:711-714
6. Herr W, Clarke J (1986) The SV40 enhancer is composed of multiple functional elements that can compensate for one another. Cell 45:461-470
7. Ondek B, Shepard A, Herr W (1987) Discrete elements within the SV40 enhancer region display different cell-specific enhancer activities. EMBO J 6:1017-1025
8. Schirm S, Jiricny J, Schaffner W (1987) The SV40 enhancer can be dissected into multiple segments, each with a different cell type specificity. Genes Dev 1:65-74
9. Ondek B, Gloss L, Herr W (1988) The SV40 enhancer contains two distinct levels of organization. Nature 333:40-45
10. Fromental C, Kanno M, Nomiyama H, Chambon P (1988) Cooperativity and hierarchical levels of functional organization in the SV40 enhancer. Cell 54:943-953
11. Davidson I, Fromental C, Augereau P, Wildeman A, Zenke M, Chambon P (1986) Cell-type specific protein binding to the enhancer of simian virus 40 in nuclear extracts. Nature 323:544-548
12. Xiao JH, Davidson I, Ferrandon D, Rosales R, Vigneron M, Macchi M, Ruffenach F, Chambon P (1987) One cell-specific and three ubiquitous nuclear proteins bind *in vitro* to overlapping motifs in the domain B1 of the SV40 enhancer. EMBO J 6:3005-3013
13. Xiao JH, Davidson I, Macchi M, Rosales R, Vigneron M, Staub A, Chambon P (1987) *In vitro* binding of several cell-specific and ubiquitous nuclear proteins to the GT-I motif of the SV40 enhancer. Genes Dev 1:794-807
14. Davidson I, Xiao JH, Rosales R, Staub A, Chambon P (1988) The HeLa cell protein TEF-1 binds specifically and cooperatively to two SV40 enhancer motifs of unrelated sequence. Cell 54:931-942
15. Macchi M, Bornert J-M, Davidson I, Kanno M, Rosales R, Vigneron M, Xiao J-H, Fromental C, Chambon P (1989) The SV40 TC-11 (κB) enhanson binds ubiquitous and cell type specifically inducible nuclear proteins from lymphoid and non-lymphoid cell lines. EMBO J 8:4215-4227
16. Shepard A, Clarke J, Herr W (1988) Simian virus 40 revertant enhancers exhibit restricted host ranges for enhancer function. J Virol 62:3364-3370
17. Gidoni D, Kadonaga JT, Barrera-Saldaña H, Takahashi K, Chambon P, Tjian R (1985) Bidirectional SV40 transcription mediated by tandem Sp1 binding interactions. Science 230:511-517

18. Greenblatt J (1991) Roles of TFIID in transcriptional initiation by RNA polymerase II. Cell 66: 1067-1070
19. Takahashi K, Vigneron M, Matthes H, Wildeman A, Zenke M, Chambon P (1986) Requirement of stereospecific alignments for initiation from the simian virus 40 early promoter. Nature 319:121-126
20. Dynan WS, Tjian R (1983) The promoter-specific transcription factor Sp1 binds to upstream sequences in the SV40 early promoter Cell 35:79-87
21 Kadonaga JT, Carner KR, Masiarz FR, Tjian R (1987) Isolation of cDNA encoding transcription factor Sp1 and functional analysis of the DNA binding domain. Cell 51:1079-1090
22. Courey AJ, Holtzman DA, Jackson SP, Tjian R (1989) Synergistic activation by the glutamine-rich domains of human transcription factor Sp1. Cell 59:827-836
23. Nordheim A, Rich A (1983) Negatively supercoiled simian virus 40 DNA contains Z-DNA segments within transcriptional enhancer sequences. Nature 303: 674-679
24. Boshart M, Weber F, Jahn G, Dorsch-Häsler K, Fleckenstein B, Schaffner W (1985) A very strong enhancer is located upstream of an immediate early gene of human cytomegalovirus. Cell 41:521-530
25. Sen R, Baltimore D (1986) Multiple nuclear factors interact with the immunoglobulin enhancer sequences. Cell 46:705-716
26. Clarke J, Herr W (1987) Activation of mutated simian virus 40 enhancers by amplification of wild-type enhancer elements. J Virol 61:3536-3542
27. Veldman GM, Lupton S, Kamen R (1985) Polyomavirus enhancer contains multiple redundant sequence elements that activate both DNA replication and gene expression. Mol Cell Biol 5:649-658
28. Tanaka M, Grossniklaus U, Herr W, Hernandez N (1988) Activation of the U2 snRNA promoter by the octamer motif defines a new class of RNA polymerase II enhancer elements. Genes Dev 2:1764-1778
29. Kanno M, Fromental C, Staub A, Ruffenach F, Davidson I, Chambon P (1989) The SV40 TC-II ($\varkappa$B) and the related H-2K^{B} enhansons exhibit different cell type specific and inducible proto-enhancer activities, but the SV40 core sequence and the AP-2 binding site have no enhanson properties. EMBO J 8:4205-4214
30. Nomiyama H, Fromental C, Xiao JH, Chambon P (1987) Cell-specific activity of the constituent elements of the simian virus 40 enhancer. Proc Natl Acad Sci USA 84:7881-7885
31. Chiu R, Imagawa M, Imbra RJ, Bockoven JR, Karin M (1987) Multiple *cis*- and *trans*- acting elements mediate the transcriptional response to phorbol esters. Nature 329: 648-651
32. Xiao JH, Davidson I, Matthes H, Garnier J-M, Chambon P (1991) Cloning, expression, and transcriptional properties of the human enhancer factor TEF-1. Cell 65:551-568
33. Baeuerle PA (1991) The inducible transcription activator NF-$\varkappa$B: regulation by distinct protein subunits. Biochim Biophys Acta 1072:63-80
34. Curran T, Franza Jr BR (1988) Fos and Jun: the AP-1 connection. Cell 55:395-397
35. Herr W (1992) Oct-1 and Oct-2: differential transcriptional regulation by proteins that bind to the same DNA sequence, in Transcriptional Regulation (McKnight S, Yamamoto K, eds), Cold Spring Harbor Press, New York, in press