MOLECULAR AND CELLULAR BIOLOGY, Sept. 1994, p. 6056–6067
0270-7306/94/$04.00+0

Vol. 14, No. 9

Reconstitution of Transcriptional Activation Domains by Reiteration of Short Peptide Segments Reveals the Modular Organization of a Glutamine-Rich Activation Domain

MASAFUMI TANAKA* AND WINSHIP HERR

Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724

Received 28 February 1994/Returned for modification 21 April 1994/Accepted 9 June 1994

The POU domain activator Oct-2 contains an N-terminal glutamine-rich transcriptional activation domain. An 18-amino-acid segment (Q^{18}III) from this region reconstituted a fully functional activation domain when tandemly reiterated and fused to either the Oct-2 or GAL4 DNA-binding domain. A minimal transcriptional activation domain likely requires three tandem Q^{18}III segments, because one or two tandem Q^{18}III segments displayed little activity, whereas three to five tandem segments were active and displayed increasing activity with increasing copy number. As with natural Oct-2 activation domains, in our assay a reiterated activation domain required a second homologous or heterologous activation domain to stimulate transcription effectively when fused to the Oct-2 POU domain. These results suggest that there are different levels of synergy within and among activation domains. Analysis of reiterated activation domains containing mutated Q^{18}III segments revealed that leucines and glutamines, but not serines or threonines, are critical for activity in vivo. Curiously, several reiterated activation domains that were inactive in vivo were active in vitro, suggesting that there are significant functional differences in our in vivo and in vitro assays. Reiteration of a second 18-amino-acid segment from the Oct-2 glutamine-rich activation domain (Q^{18}II) was also active, but its activity was DNA-binding domain specific, because it was active when fused to the GAL4 than to the Oct-2 DNA-binding domain. The ability of separate short peptide segments derived from a single transcriptional activation domain to activate transcription after tandem reiteration emphasizes the flexible and modular nature of a transcriptional activation domain.

In eukaryotes, transcriptional activators display considerable modularity, frequently consisting of separable domains for binding to DNA or for interacting with the transcriptional machinery (reviewed in reference 32). These domains can often be exchanged between activators, creating activators with new promoter specificities (1, 32, 47).

Much is known about the structure of DNA-binding domains and how they recognize DNA (reviewed in reference 30). DNA-binding domains can be classified into structurally related families that have been conserved during evolution. In contrast, little is known concerning the structure of the regions responsible for transcriptional activation. Although in many cases there is no evidence that these regions represent structurally defined domains, they are nevertheless referred to as activation domains because of their structural independence from DNA-binding domains. Unlike DNA-binding domains, however, the primary sequence of activation domains is rarely well conserved, although the relative abundance of particular amino acids, such as acidic, glutamine, or proline residues, within such domains permits a loose classification according to amino acid composition (reviewed in reference 24).

A variety of transcriptional activation domains have been characterized in detail by mutagenesis. Typically, in stark contrast to their effects on DNA-binding domains, deletions fail to identify discrete activation domain boundaries; instead, activity is lost gradually as more sequences are deleted, as first illustrated with the yeast activator GCN4 (16, 17). This finding led to the suggestion that activation domains are composed of multiple redundant elements (16). Consistent with this view, amino acid substitutions only rarely have dramatic effects on the activity of transcriptional activation domains (4, 11–13, 34).

These properties are reminiscent of those of transcriptional enhancers, which are structurally complex and consist of a hierarchy of interacting regulatory elements, as exemplified by the enhancer in simian virus 40 (SV40) (reviewed in reference 14). Individual enhancer elements display little, if any, activity, but their activity can be amplified by tandem reiteration (29, 37, 50). Tandem reiteration is a valuable tool by which to elucidate the structure of enhancers, in part because the activity of tandemly reiterated enhancer elements is very sensitive to point mutagenesis, since point mutations are repeated in each copy of the enhancer element (29).

Here, we have tested whether reiteration of activation domain subsegments and subsequent mutagenesis can identify elementary self-cooperating units within an activation domain. As a test, we used subsegments of a glutamine-rich transcriptional activation domain derived from the human B-cell transcriptional activator Oct-2.

MATERIALS AND METHODS

Effector constructs. The empty expression vector pCG and the pCG*oct*-2, pCG*oct*-2ΔN^6 (Δ.POU.C), pCG*oct*-2$\Delta N^6 \Delta C^2$ (Δ.POU.Δ) (45), and pCG-GAL4(1–94) (43) expression vectors have been described previously. Expression constructs containing reiterated 18-, 19-, or 20-amino-acid glutamine-rich segments were constructed by the same strategies as described below for the Oct-2 Q^{18}III construct except that for the Q^{18}III(S/T→A), Q^{18}III(L→A), and Q^{18}III(Q→A) constructs synthetic oligonucleotides rather than PCR products were used to create the reiterated segments.

For Q^{18}III, a DNA fragment encoding the 18-amino-acid

* Corresponding author.

Oct-2 segment was generated by PCR using 5′ and 3′ oligonucleotide primers containing the sequences CTGAAGACA CACCAAATCTATTCCAG and CTGAAGACTGTGGTTGG GAGGTCAGAA, respectively. These 5′ and 3′ primers anneal to *oct*-2 nucleotides 487 to 504 and 530 to 542, respectively (3). Both primers contain *Bbs*I recognition sites (underlined): *Bbs*I cleavage of the amplification products produces ACCA and TGGT 5′ overhangs (boldface). Multimers of the *Bbs*I fragment were prepared by self-ligation, and fragments with the desired number of Q^{18}III segments were purified by polyacrylamide gel electrophoresis.

Fragments containing multiple Q^{18}III segments were then subcloned into a set of intermediate vectors for subsequent transfer N or C terminal to the Oct-2 POU domain or C terminal to the GAL4 DNA-binding domain. To place the tandem Q^{18}III segments N terminal or C terminal of the Oct-2 POU domain, they were subcloned into the N-terminal and C-terminal adaptor vectors pU*Coct*-2Nad and pU*Coct*-2Cad described in the accompanying article (43). To place the tandem Q^{18}III segments N terminal of the Oct-2 POU domain, the Q^{18}III segment-containing pU*Coct*-2Nad *Xba*I-*Sty*I fragment was inserted into the unique *Xba*I site positioned immediately N terminal to the Oct-2 POU domain in pCG*oct*-2ΔN^6. To place the multiple Q^{18}III segments C terminal to the Oct-2 POU domain, the Q^{18}III segment-containing pU*Coct*-2Cad *Sal*I-*Bam*HI fragment was recombined with the appropriate pCG*oct*-2 derivative.

To place multiple Q^{18}III segments C terminal of the GAL4(1–94) DNA-binding domain, the ligated and gel-purified *Bbs*I Q^{18}III segment-containing fragment was subcloned into the intermediate vector pUC-GAL4ad containing matching ACCA and TGGT overhangs, in which the 5′ ACCA sequence is preceded by the sequence TCTAG to create an *Xba*I recognition site and the ACCA at the 3′ end of the reiterated Q^{18}III sequence is followed by the sequence TAGTCTTCGGATCC, containing a TAG termination codon (underlined) and a *Bam*HI recognition site. The resulting *Xba*I-*Bam*HI Q^{18}III segment-containing fragment was then cloned between the *Xba*I and *Bam*HI sites of pCG-GAL4(1–94) (43).

The *Escherichia coli* expression plasmids are derivatives of the T7 polymerase expression vector pET11c (39). *Xba*I-*Bam*HI fragments containing the tandem Q^{18}III segment-coding sequences were prepared from the appropriate pCG construct and cloned between the pET11c *Nde*I and *Bam*HI sites with the use of a short *Nde*I-*Xba*I adaptor (annealed TATGGCTT and CTAGAAGCCA oligonucleotides). The resulting coding sequences in the *E. coli* expression plasmids are identical to those in the pCG plasmids. Plasmids are referred to by the vector name (e.g., pCG- or pET11c-) followed by the name of the effector protein.

Reporter constructs. The β-globin reporter construct p$\beta\Delta^{36}$6×oct$^+$sph$^-$ (p$\beta\Delta^{36}$6×B20dpm2/−52) has been described previously (45) and contains the human β-globin promoter truncated immediately upstream of the TATA box and six copies of the SV40 B element octamer motif inserted 52 bp upstream of the transcriptional start site. The c-*fos* GAL4-binding site reporter pc-*fos*−56/4×G17M contains a fusion of the c-*fos* promoter to the chloramphenicol acetyltransferase (CAT) gene and is described in the accompanying article (43). The internal reference plasmid pα4×(A+C) has been described previously (45). The p$\beta\Delta^{36}$2×oct/hept construct is similar to p$\beta\Delta^{36}$6×oct$^+$sph$^-$ except that, in place of the six SV40-derived octamer sites, it contains two tandem copies of the octamer-heptamer (underlined) containing sequence TGCCTCATGAGTATGCAAATCAGC (20).

Transient-expression assay. HeLa cell transient-expression assays were performed as described in the accompanying article (43). Each transfection mix, with transfection done by calcium phosphate coprecipitation, contained the β-globin or c-*fos* reporter plasmid (0.5 μg), internal reference plasmid pα4×(A+C) (4 ng), pCG effector plasmid (0.75 μg), and pUC119 carrier DNA (up to 10 μg of total DNA). To adjust the expression of different effector proteins to comparable levels, some of the pCG effector plasmids were diluted up to sixfold with the empty pCG vector. Each titration of effector concentration was a threefold serial dilution of the effector expression plasmid with the empty pCG vector. To measure β-globin reporter expression by RNase protection, the α-globin probe α132 and β-globin probe β350 were used together, and to measure c-*fos*-CAT reporter expression, the α-globin probe α98 and c-*fos* probe *fos*128 were used together. The results were quantitated with either a BAS1000 or BAS2000 Fuji bio-image analyzer.

The quality and quantity of the transiently expressed POU domain- or GAL4 DNA-binding domain-containing proteins were monitored by an electrophoretic mobility retardation assay as described before (43). Titration curves performed in parallel were used to establish the linearity of the DNA-binding assays. Complexes generated by truncated forms of the transiently expressed proteins, which were less than 25% of the total octamer site-binding activity, were included in the measurement of transient protein expression.

In vitro transcription assay. HeLa cell fractions were prepared by the method of Sumimoto et al. (41). Nuclear extracts were fractionated over phosphocellulose (Whatman P11) to obtain 0.1, 0.3, 0.5, and 0.85 M KCl step eluates. The 0.1 and 0.85 M eluates were further fractionated over DEAE-cellulose (Whatman DE52) to prepare TFIIA and TFIID fractions. RNA polymerase II was purified to apparent homogeneity by the procedure of Reinberg and Roeder (35). Protein concentrations were measured with Bio-Rad dye, using bovine serum albumin as the standard.

The POU domain-containing effector proteins were expressed in *E. coli* at room temperature and purified by DNA affinity chromatography (18). The purity of the proteins, as determined by Coomassie blue staining after polyacrylamide gel electrophoresis, was at least 80% except for the Q^{18}III(Q→A) mutant protein, which was approximately 50% pure. The DNA-binding activity of each preparation was measured by electrophoretic mobility retardation, and the relative protein concentration was subsequently compared by polyacrylamide gel electrophoresis and Coomassie blue staining. For the entire set of effector proteins used for in vitro transcription, the relationship between the DNA-binding activity and the protein concentration of each effector protein preparation was within twofold. In the in vitro transcription reactions, the levels of effector protein were normalized to their level of DNA-binding activity.

In vitro transcription reaction mixtures contained 1.2 μl of TFIIA fraction (1.4 mg of protein per ml), 4.8 μl of 0.3 to 0.5 M P11 fraction (0.7 mg of protein per ml), 1.6 μl of TFIID fraction (0.8 mg of protein per ml), 0.4 μl of purified RNA polymerase II (approximately 0.1 mg/ml with 0.1 mg of bovine serum albumin carrier per ml), 20 ng of β-globin reporter plasmid, 80 ng of pUC119, and up to 50 ng of effector protein. The TFIID fraction contained approximately 10 μg of TATA box-binding protein per ml, as estimated by immunoblot (21). Reactions were performed in 20 μl of 8 mM Tris-HCl (pH 7.9 at 4°C)–4 mM HEPES (*N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid)-KOH (pH 7.9 at room temperature)–80 mM KCl–7 mM $MgCl_2$–0.4 mM EDTA–0.4 mM dithiothreitol–4

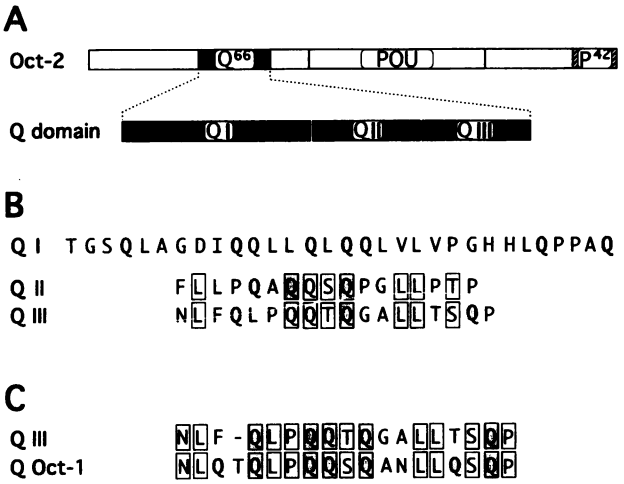

FIG. 1. The 66-amino-acid Oct-2 glutamine-rich transcriptional activation domain contains an internally repeated structure. (A) Diagram of Oct-2, showing the locations of the 66-amino-acid glutamine-rich (Q^{66}) and 42-amino-acid proline-rich (P^{42}) activation domains and the central DNA-binding POU domain. The Q^{66} domain is subdivided into QI, QII, and QIII segments, according to the clustering of glutamine residues. (B) Sequence of the QI, QII, and QIII segments aligned to maximize the similarity of the arrangement of glutamine residues (shaded). Identities and serine-threonine similarities between segments QII and QIII in this alignment are boxed. (C) Sequence comparison between Oct-2 segment QIII and a 19-amino-acid glutamine-rich sequence from Oct-1 is shown, with residues shaded and boxed as described above.

mM 2-mercaptoethanol–2.6% (wt/vol) polyvinyl alcohol–12% (vol/vol) glycerol–0.4 mM ribonucleoside triphosphates at 30°C for 60 min. The resulting β-globin transcripts were analyzed by RNase protection of a radiolabeled probe that is protected over 286 nucleotides by correctly initiated transcripts. To monitor transcript recovery during the experimental manipulations, a constant amount of nonradiolabeled β-globin coding-strand RNA truncated at the human β-globin *Nco*I recognition site (+49 relative to the initiation site) was added to each sample at the end of the reaction.

RESULTS

Oct-2 glutamine-rich transcriptional activation domain contains an internally repeated substructure. Figure 1 shows an illustration of Oct-2 (3, 25, 36), showing the positions of two previously mapped transcriptional activation domains that flank the central DNA-binding POU domain: a 66-amino-acid N-terminal glutamine-rich activation domain (Q domain) and a 42-amino-acid C-terminal proline-rich activation domain (P domain) (43). By sequence analysis, the 66-amino-acid Q domain can be subdivided according to the clustering of glutamine residues into three segments (26): an N-terminal 31-amino-acid segment, which we refer to as QI; a 17-amino-acid QII segment; and a C-terminal 18-amino-acid QIII segment (Fig. 1A and B). In Fig. 1B, the sequences of these three Q-rich segments are aligned to maximize the similarity in the arrangement of glutamine residues (shown shaded). This alignment reveals considerable similarity between the QII and QIII segments; in addition to three identically positioned glutamine residues, there are three identically positioned leucine residues and two identically positioned serine or threonine residues (boxed residues in Fig. 1B).

Human Oct-1, a broadly expressed POU domain protein closely related in structure to Oct-2 (40), also contains an N-terminal region rich in glutamine residues, and this region can functionally replace the Oct-2 Q domain (45). Figure 1C shows a sequence alignment of the Oct-2 QIII segment with a similar 19-amino-acid segment from the Oct-1 N-terminal Q-rich region, referred to as the Q-Oct-1 segment. When a single-amino-acid gap is placed in the Oct-2 QIII segment, these Oct-1- and Oct-2-derived segments are identical at 13 of 19 positions, and one of the nonidentities is threonine for serine, a conservative difference (boxed residues in Fig. 1C).

The sequence alignments shown in Fig. 1 suggest that the Oct-1 and Oct-2 glutamine-rich regions share a sequence motif which in the Oct-2 Q domain is reiterated at least twice. We therefore tested whether artificial reiteration of one of the Oct-2 segments could reconstitute a functional transcriptional activation domain.

Reiteration of the 18-amino-acid Oct-2 QIII segment can create a potent transcriptional activation domain. Because of its similarity to other segments of both the Oct-1 and Oct-2 Q domains, we first chose to reiterate the 18-amino-acid Oct-2 QIII segment, which we refer to as Q^{18}III. We first replaced the N-terminal Oct-2 Q domain with one through five tandem copies of Q^{18}III. In our transcriptional activation assay, wild-type Oct-2 requires both the N-terminal Q domain and C-terminal P domain to activate transcription effectively (43, 45). We therefore asked whether the reiterated Q^{18}III segments could functionally replace the Q domain and cooperate with the C-terminal sequences of Oct-2 to activate transcription.

The transcriptional activity of these proteins was assayed in a transient-expression assay in HeLa cells, in which Oct-1 but not Oct-2 is expressed. The recombinant activators were expressed from the cytomegalovirus promoter-based expression plasmid pCG (45), and their activity was assayed on a reporter plasmid containing six copies of the SV40 enhancer octamer sequence ATGCAAAG placed immediately upstream of the β-globin promoter TATA box. The SV40 enhancer octamer site is a natural binding site for both Oct-1 and Oct-2, but this reporter displays little if any activity in HeLa cells in the absence of Oct-2 because Oct-1 does not activate mRNA transcription as effectively as Oct-2 (45, 47). As an internal reference, a plasmid containing the human α-globin gene, which does not respond to Oct-1 or Oct-2, was included in the assay.

To assay reporter gene and effector protein expression, the transfected cells were divided into two portions. Reporter gene expression was measured in an RNase protection assay that measured the levels of correctly initiated β-globin reporter and α-globin reference transcripts. The quantity and quality of effector protein expression were assayed by electrophoretic mobility retardation.

Figure 2 shows the results of such an assay. The structure of the effector proteins is illustrated in Fig. 2C, together with the results of the electrophoretic mobility retardation (Fig. 2A) and RNase protection (Fig. 2B) assays. As shown in Fig. 2A, the effector proteins containing one to five tandem copies of the 18-amino-acid Q^{18}III segment were expressed at similar levels and with little evident degradation (Fig. 2A, lanes 4 to 8). In the RNase protection assay, Oct-2 activated β-globin expression 20-fold (compare lanes 1 and 2, Fig. 2B); the response of the β-globin promoter to changes in Oct-2 expression was linear at concentrations up to fivefold higher than those shown in Fig. 2 (data not shown). Therefore, we compared the levels of RNA when normalized to the less than threefold variation in protein expression observed in Fig. 2A. The results of such normalization are shown in Fig. 2D.

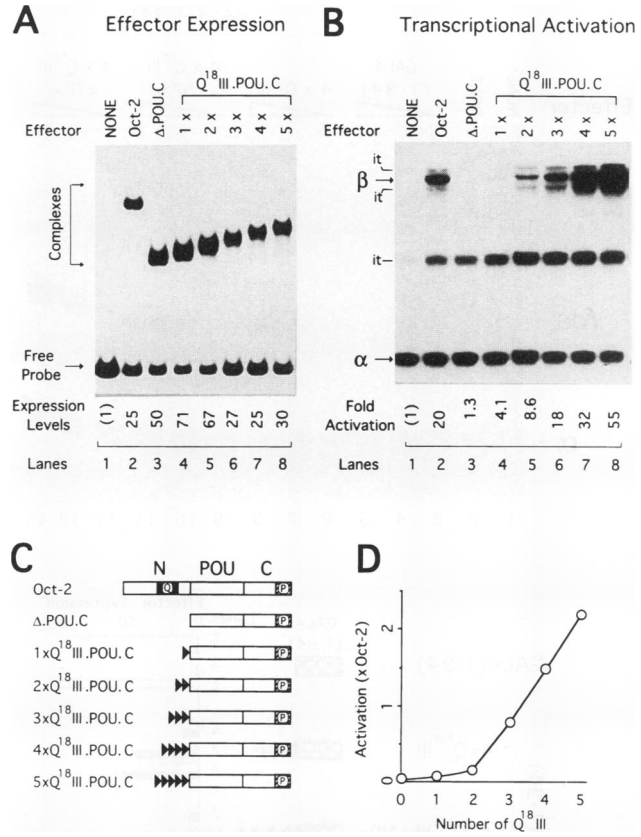

FIG. 2. Reiterated $Q^{18}III$ segments can activate transcription effectively in conjunction with the Oct-2 C-terminal P^{42} domain. The β-globin reporter $p\beta\Delta^{36}6\times oct^{+}sph^{-}$ and the α-globin internal reference $p\alpha4\times(A+C)$ were transfected into HeLa cells along with a pCG expression plasmid for each effector indicated. (A) Effector expression as assayed by electrophoretic mobility retardation with extracts prepared from a portion of the transfected cells. The identity of the expressed effector is indicated above each lane; lane 1, no effector was expressed. The fold increase in DNA-binding activity relative to the endogenous Oct-1 activity in lane 1 is listed underneath each lane. (B) Transcriptional activity as assayed by RNase protection. The positions of fragments generated by the correctly initiated α-globin (α) and β-globin (β) transcripts and incorrect transcripts (it) are indicated. The identity of the effector is indicated above each lane; lane 1, no effector. The fold increase in β-globin transcripts relative to the level in lane 1 is indicated below each lane. (C) Structures of effectors with reiterated $Q^{18}III$ segments. Each solid triangle indicates one copy of the $Q^{18}III$ segment. (D) Graph showing the levels of transcriptional activation by effectors with one to five tandem $Q^{18}III$ segments, after normalization for effector protein expression. The levels are shown relative to the level of wild-type Oct-2 activation.

Consistent with previous results (45), removal of the N-terminal Q domain from Oct-2 resulted in nearly complete loss of activation of the correct transcript (β), although for unknown reasons, an incorrect transcript (it) was still activated (Fig. 2B, lane 3). Activity was restored, however, by the progressive addition of $Q^{18}III$ segments (Fig. 2B, lanes 4 to 8). Indeed, three tandem $Q^{18}III$ segments restored wild-type Oct-2 levels of activation (Fig. 2B, compare lanes 2 and 6). The normalized transcriptional activation values shown in Fig. 2D show that the response to increasing numbers of $Q^{18}III$ segments is bimodal. Addition of just one or two $Q^{18}III$ segments results in only a small, albeit reproducible, increase in reporter gene activation. The low level of activation by two tandem $Q^{18}III$ segments in $2\times Q^{18}III$ can be clearly seen in Fig. 2B (lane 5), because this protein was very well expressed (see Fig. 2A, lane 5). After two copies, however, the rate of increased activation with respect to increased $Q^{18}III$ copy number is linear and much greater than with one to two $Q^{18}III$ segments (see Fig. 2D). These results show that, in the context of the Oct-2 POU and C-terminal activation domains, reiteration of the Oct-2 $Q^{18}III$ segment, which itself displays little activity, can create an effective activation domain once there are three tandem segments.

Sets of reiterated $Q^{18}III$ segments synergize to activate transcription autonomously when flanking the Oct-2 POU domain. The ability of sets of reiterated $Q^{18}III$ segments to activate transcription in the absence of a heterologous activation domain was tested by placing four copies of the $Q^{18}III$ segment ($4\times Q^{18}III$) both N and C terminal of the DNA-binding Oct-2 POU domain. As shown in Fig. 3, neither the Oct-2 POU domain (lanes 11 to 13) nor the $4\times Q^{18}III$ activation domain placed either N terminal (lanes 8 to 10) or C terminal (lanes 5 to 7) of the Oct-2 POU domain stimulated transcription from the β-globin promoter effectively. But when individual $4\times Q^{18}III$ activation domains flank the Oct-2 POU domain, as in $4\times Q^{18}III.POU.4\times Q^{18}III$ (Fig. 3, lanes 2 to 4), reporter gene expression is activated as much as 200-fold (lane 4). Thus, the $4\times Q^{18}III$ activation domain can synergize either with itself or with the Oct-2 C-terminal activation domain (Fig. 2), displaying synergistic interactions similar to those exhibited by the wild-type Oct-2 Q and P activation domains (43, 45).

Reiterated $Q^{18}III$ segments are active when linked to a heterologous DNA-binding domain or with altered spacing between segments. The ability of reiterated $Q^{18}III$ segments to function independently of the POU domain was tested by fusing the $4\times Q^{18}III$ activation domain to the unrelated DNA-binding domain from the yeast activator GAL4. Unlike the POU domain, which binds the octamer motif as a monomer (19), the GAL4 DNA-binding domain binds DNA as a dimer (23). To assay the activity of the GAL4-$4\times Q^{18}III$ chimera, we used a c-*fos* promoter reporter with four GAL4-binding sites. The results of such an assay are shown in Fig. 4. The GAL4-$4\times Q^{18}III$ chimera but not the GAL4 DNA-binding domain alone (residues 1 to 94) activated the c-*fos* promoter effectively (Fig. 4, compare lanes 6 to 8 with lanes 3 to 5). In contrast to the apparent requirement for two $4\times Q^{18}III$ activation domains to activate transcription effectively when linked to the Oct-2 POU domain, only a single $4\times Q^{18}III$ activation domain is required to activate transcription when linked to the GAL4 DNA-binding domain (compare Fig. 3 and 4). This difference may result because GAL4 binds DNA as a dimer and thus naturally presents two activation domains per DNA-binding site; these domains may synergize with each other, as do the two separate activation domains linked to the monomer-binding Oct-2 POU domain.

Figure 4 also shows that neither the precise spacing nor junction sequence between $Q^{18}III$ segments is critical, because insertion of two amino acids between each $Q^{18}III$ segment, either asparagine and proline, as in $Q^{18}III<NP>$ (lanes 9 to 11), or threonine and proline, as in $Q^{18}III<TP>$ (lanes 12 to 14), had only a small positive effect on transcriptional activation. This spacing result, together with the weak but significant activity of a single copy of $Q^{18}III$ (see Fig. 2), suggests that the $Q^{18}III$ sequence contains a functional transcriptional activation element and thus perhaps a unit for contact with a receptor molecule in the transcriptional machinery.

$Q^{18}III$-related and mutant segments. An advantage of the reiteration approach described here is that the positive or negative effects of sequence changes are magnified by segment

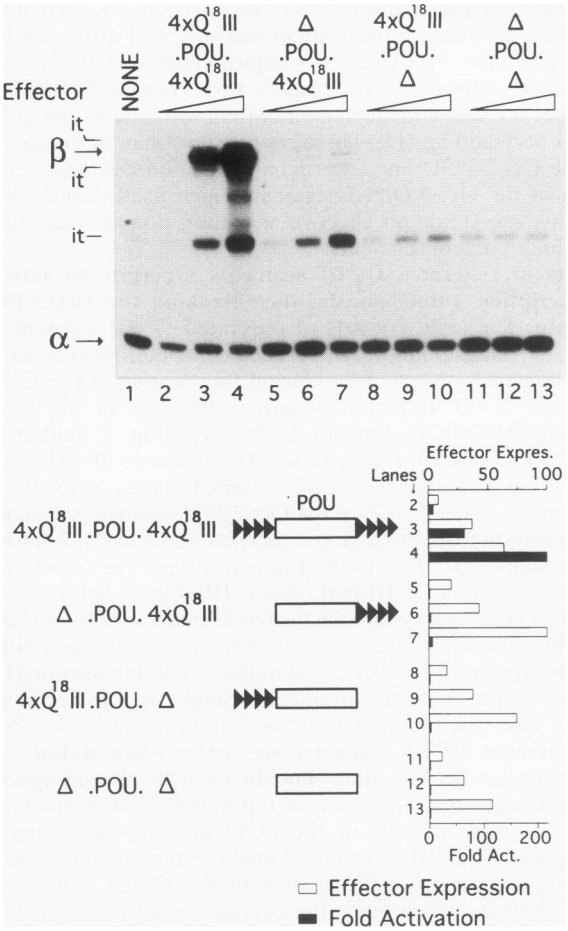

FIG. 3. Four tandemly reiterated $Q^{18}III$ segments synergize with a duplicate of themselves to activate transcription effectively. The β-globin reporter $p\beta\Delta^{36}6\times oct^{+}sph^{-}$ and the α-globin internal reference pα4×(A+C) were transfected into HeLa cells along with a pCG expression plasmid for the Oct-2 POU domain alone (Δ.POU.Δ, lanes 11 to 13) or for four reiterated $Q^{18}III$ segments N (4×Q^{18}III.POU.Δ, lanes 8 to 10) or C (Δ.POU.4×Q^{18}III, lanes 5 to 7) terminal of the POU domain and both N and C terminal of the POU domain (4×Q^{18}III.POU.4×Q^{18}III, lanes 2 to 4). Each set of three lanes represents a threefold titration of transfected expression vector concentration. Lane 1, empty pCG vector. The results of RNase protection assays are shown. The positions of fragments generated by the correctly initiated α-globin (α) and β-globin (β) transcripts and incorrect transcripts (it) are indicated. Underneath are shown the structures of the effectors and the quantitation of effector expression determined by electrophoretic mobility retardation (open bars) relative to that of the sample in lane 7, which is set at 100% and is 35-fold higher than the endogenous level of activity from Oct-1 in lane 1, and transcriptional activation (solid bars) relative to that of the sample without effector (lane 1).

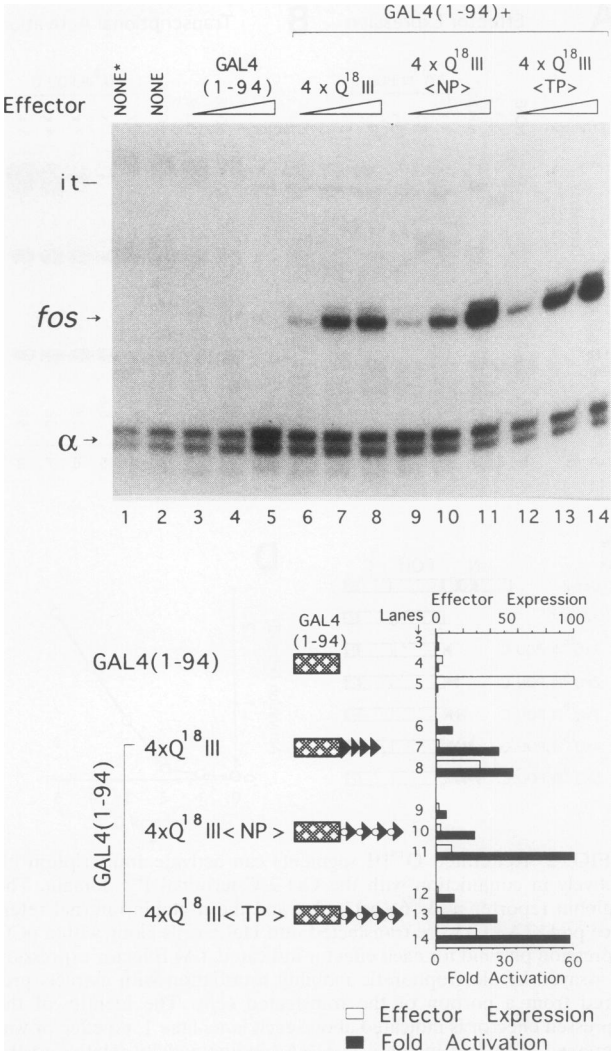

FIG. 4. Reiterated $Q^{18}III$ segments are active when linked to a heterologous DNA-binding domain or with altered spacing between segments. The results of a transient HeLa cell expression assay of GAL4 fusion proteins containing reiterated $Q^{18}III$ segments are shown along with the structures of the effector proteins and quantitation of the results. HeLa cells were transfected with the c-*fos*/−56 (lane 1) or c-*fos*/−56/4×G17M (lanes 2 to 14) reporter, α-globin internal reference, and the empty pCG effector expression plasmid (lanes 1 and 2) or pCG effector expression plasmids encoding the GAL4 DNA-binding domain alone [GAL4(1–94), lanes 3 to 5] or fused to four tandem copies of the $Q^{18}III$ segment without spacer (4×Q^{18}III, lanes 6 to 8) or with two amino acid spacers (4×Q^{18}III<NP>, lanes 9 to 11, or 4×Q^{18}III<TP>, lanes 12 to 14). The positions of RNase-protected fragments corresponding to correctly initiated α-globin (α) and c-*fos* (*fos*) transcripts and incorrect *fos*-vector transcripts (it) are indicated. At the bottom are shown the structures of the effectors and the quantitation of effector protein expression determined by electrophoretic mobility retardation (open bars) relative to that of the sample in lane 5, which is set at 100%, and transcriptional activation (solid bars) relative to the level without effector (lane 2).

reiteration. We therefore tested the structural requirements for activation by reiterated copies of Q-rich segments, using the two different kinds of sequence changes illustrated in Fig. 5. In the first approach, the activities of the $Q^{18}III$-related Oct-1 and Oct-2 sequences described in Fig. 1 were assayed. We tested the activity of the upstream 17-amino-acid Oct-2 QII segment; to retain the same reiterated segment size as in the $Q^{18}III$ constructs, we added an N-terminal asparagine (N) residue to the QII sequence (see Fig. 5A). This segment is called $Q^{18}II$. We also tested the activity of the 19-amino-acid Oct-1 segment that resembles the Oct-2 QIII sequence (see Fig. 1C) and called it Q^{19}Oct-1, as shown in Fig. 5A.

In the second strategy, amino acid substitutions were made in the wild-type $Q^{18}III$ sequence, as shown in Fig. 5B. We first constructed two relatively subtle $Q^{18}III$ mutants, called

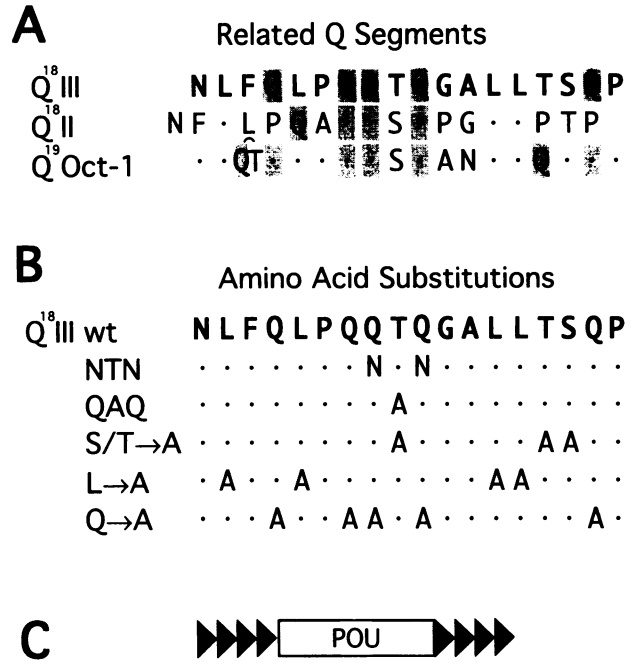

FIG. 5. Amino acid sequence of Q^{18}III-related and mutant segments. (A) Related segments. The Q^{18}III amino acid sequence is shown at the top, with the Q^{18}II and Q^{19}Oct-1 sequences aligned below it. Dots indicate residues that are identical to the Q^{18}III sequence; glutamine residues are shaded, and in the Q^{19}III alignment, two residues (QT) take the place of one residue to maintain the best alignment with the Q^{18}III sequence. (B) Q^{18}III segment amino acid substitutions. The mutant segments are listed below the wild-type Q^{18}III segment sequence, with amino acid substitutions indicated. (C) Diagram of the effector used to study the activity of the Q^{18}III-related and mutant segments. The solid triangles denote Q segments, and the shaded box is the Oct-2 POU domain.

Q^{18}III(NTN) and Q^{18}III(QAQ). In these two substitution mutants, either two glutamine residues were substituted with asparagine residues [Q^{18}III(NTN)] or a single threonine residue was substituted with an alanine residue [Q^{18}III(QAQ)] in the central QTQ sequence. This sequence was targeted for mutagenesis because, in the sequence alignment shown in Fig. 5A, a QSQ variant of this QTQ sequence is present in the same position in the QII and Q-Oct-1 sequences, suggesting a role in activation. However, as shown below, these substitutions had little effect on activation of transcription. We therefore made more extensive changes by substituting particular sets of amino acids (e.g., all glutamines) for alanine residues. In this manner, the three serine and threonine residues (S/T→A), four leucine residues (L→A), and five glutamine residues (Q→A) were separately mutated to alanine, as shown in Fig. 5B.

The activities of the two variant and five mutant Q^{18}III segments were measured by placing four reiterated copies of each segment both N terminal and C terminal of the Oct-2 POU domain, as illustrated in Fig. 5C. Their transcriptional activities were measured in vivo in the transient-expression assay and in vitro in a reconstituted transcription assay on two different β-globin promoter constructs. Below, we show the results obtained in the in vivo assay with the β-globin construct used in the experiments for Fig. 2 and 3, which contains six copies of the relatively low-affinity SV40 octamer site. This construct can be used in the repression assay described below, which is critical for demonstrating that the transcriptionally inactive proteins are not inactive simply because they fail to reach the promoter properly in vivo. For the in vitro assay, we show the results obtained with the β-globin promoter construct containing two copies of the higher-affinity immunoglobulin heavy-chain promoter octamer site. With this construct, a saturating level of activator protein is more easily attained, and therefore the maximal activation potential of the different reiterated activation domains is more easily compared. Both reporters, however, display qualitatively identical responses to the various Q^{18}III-related POU domain activators when compared both in vivo and in vitro (data not shown).

Some Q^{18}III mutations affect transcriptional activation in vivo. Figure 6 shows a comparison of the activity of the Q^{18}III segment and its related and mutant derivatives in vivo and in vitro. In the in vivo assay (Fig. 6A), each effector was assayed at three different concentrations. In every case, the full-length effector protein was faithfully expressed and, except for the Q^{18}III(Q→A) glutamine substitution mutant, which was expressed approximately fivefold less well than its wild-type counterpart, the Q^{18}III, Q^{18}III-related, and mutant Q^{18}III proteins were expressed at similar levels. In Fig. 6C, the fold activation of reporter gene expression is plotted against the level of effector protein expression as measured by electrophoretic mobility retardation.

Wild-type Oct-2 activated transcription 15- to 20-fold in this assay, whereas the Oct-2 POU domain alone displayed little if any activity (Fig. 6A, lanes 2 to 7; Fig. 6C). The Q^{18}III protein afforded up to 100-fold activation of reporter gene expression (Fig. 6A, lanes 8 to 10). When compared at similar protein expression levels, the Q^{18}III protein activated expression 10-fold better in vivo than did the wild-type Oct-2 protein (see Fig. 6C). The two Q^{18}III-related segments, Q^{18}II and Q^{19}Oct-1, however, exhibited little or no activity (Fig. 6A, lanes 11 to 16). This result shows that not all reiterated segments derived from glutamine-rich activation domains are able to activate transcription from the β-globin promoter effectively when fused to the Oct-2 POU domain.

Of the amino acid substitution mutants, the two less drastic mutations, Q^{18}III(NTN) and Q^{18}III(QAQ), had little effect on transcriptional activation (Fig. 6A, lanes 17 to 22). Replacement of the three serine and threonine residues by alanine residues in Q^{18}III(S/T→A) also had little, if any, effect on Q^{18}III activity at the two lower levels of effector protein expression (Fig. 6A, lanes 23 and 24); at the highest level of effector protein expression, however, this mutant protein was reproducibly twofold less active than the wild-type protein (Fig. 6A, lane 25; see also Fig. 6C), suggesting a concentration-dependent defect in transcriptional activation.

In contrast to these substitution mutants, the two mutants carrying substitutions of either the four leucine [(Q^{18}III (L→A)] or five glutamine [Q^{18}III(Q→A)] residues displayed little if any activity in the in vivo assay (Fig. 6A, lanes 26 to 31). The inability of the leucine-to-alanine mutant to activate transcription is consistent with previously described results showing that glutamines per se are not sufficient for a glutamine-rich activation domain to activate transcription (11, 13). The relative inactivity of the glutamine-to-alanine mutant, however, shows that at some level, the glutamine residues are important for activation of transcription, an effect that has not been observed with one of the glutamine-rich activation domains in Spl (13). We note, however, that this glutamine-to-alanine mutant, in contrast to the leucine-to-alanine mutant, consistently displayed a low level of activity, suggesting that although the glutamines are very important for the activity of the multimerized Q^{18}III segment, they are not absolutely essential for low levels of activity.

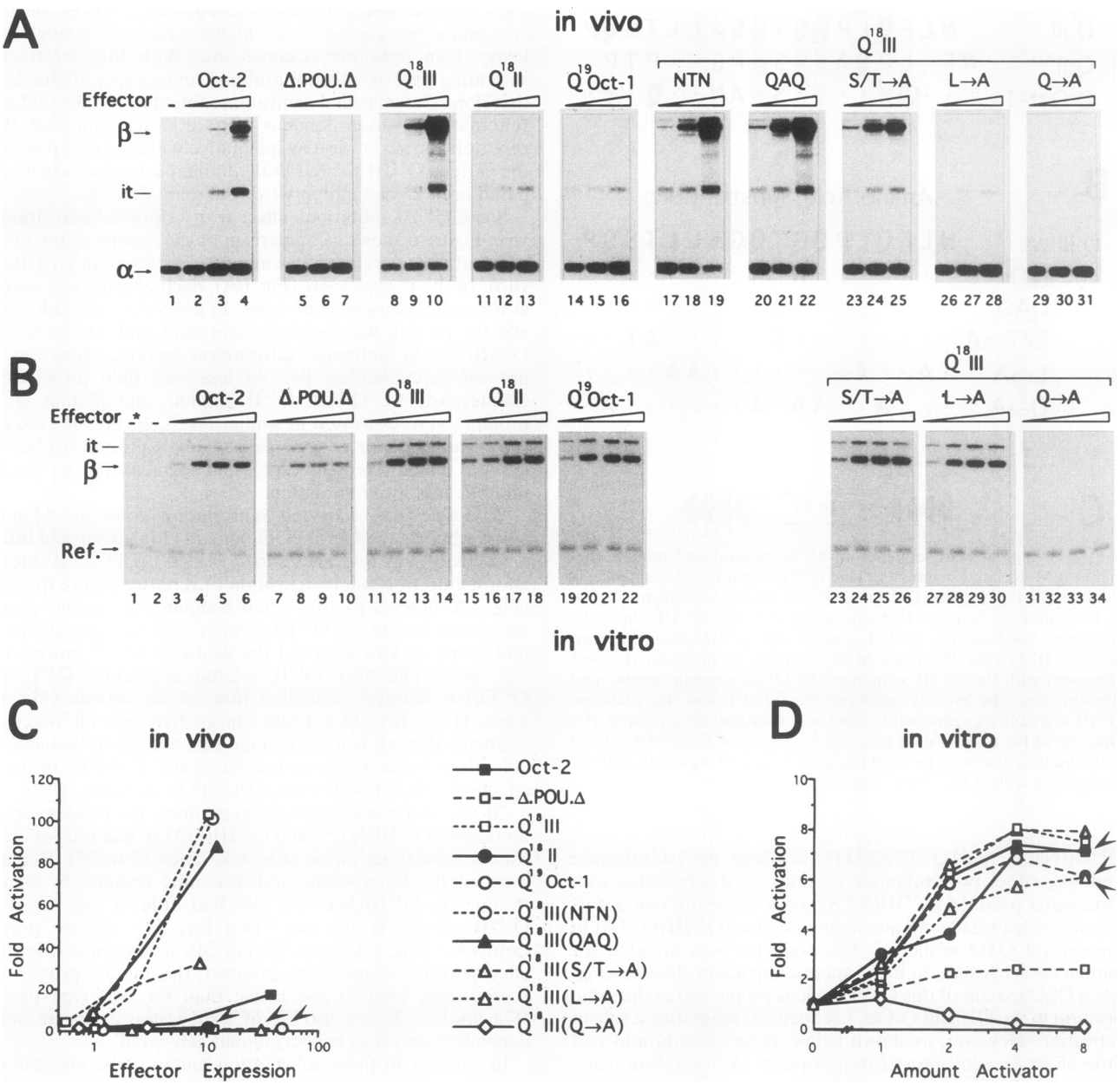

FIG. 6. Related and mutant $Q^{18}III$ segments display different activities in vivo and in vitro. In vivo (A) and in vitro (B) assay of transcriptional activity by POU domain proteins containing N- and C-terminal $4\times Q^{18}III$ activation domain derivatives, and quantitation of the in vivo (C) and in vitro (D) results. (A) RNase protection after transient expression in vivo. HeLa cells were transfected with α-globin internal control, β-globin reporter, and pCG effector expression plasmids, and RNA and protein expression was analyzed as described in the Fig. 2 legend and Materials and Methods. The identity of each effector is listed above each set of three lanes, which represent a threefold titration of effector plasmid concentration. Lane 1, empty pCG expression vector alone. The positions of the RNase-protected fragments for correctly initiated α-globin (α) and β-globin (β) and incorrect β-globin (it) transcripts are indicated. (B) RNase protection after in vitro transcription. The POU domain proteins indicated above each set of four lanes, which represent a twofold serial titration of POU protein concentration, were added to in vitro transcription reaction mixes containing HeLa cell fractions and RNA polymerase II as described in Materials and Methods; for each effector protein, the highest level of protein was approximately 50 ng. The β-globin plasmid $p\beta\Delta^{36}2\times oct/hept$ was used as a template except for lane 1, in which the β-globin plasmid $p\beta\Delta^{36}enh^-$, lacking the 2×oct/hept sites, was used. The positions of RNase-protected fragments corresponding to correctly initiated β-globin experimental transcripts (β), readthrough transcripts (it), and the internal reference (ref) are indicated. (C) Transcriptional activation in vivo is plotted relative to levels of effector protein expression determined by electrophoretic mobility retardation. For clarity, the symbols for the intermediate concentration of effector have been omitted. (D) In vitro transcriptional activation is plotted relative to the amount of added POU domain proteins. The sets of three arrowheads in panels C and D highlight the different in vivo and in vitro activities of the $Q^{18}II$, $Q^{19}Oct\text{-}1$, and $Q^{18}III(L\rightarrow A)$ segments.

Inactive proteins do not fail to reach the target promoter. The electrophoretic mobility retardation analysis of the transiently expressed wild-type and mutant activator proteins showed that they were expressed at comparable levels and without excessive degradation. This assay, however, does not address whether the defective activator proteins are inactive because they are specifically defective in transcriptional activation or because they are unable to reach the promoter in vivo. For example, transcriptionally active proteins might appear inactive if they are not localized in the nucleus properly. To address this ambiguity, we assayed the ability of the inactive Q^{18}III-related proteins to repress the activity of the wild-type Q^{18}III activator when coexpressed in the transient-expression assay at the higher level assayed in Fig. 6A; repression of Q^{18}III activation by a defective activator would suggest that the inactive protein can reach the promoter and occlude the wild-type Q^{18}III activator from the promoter. Repression of Q^{18}III activation would therefore indicate that the defective activators fail to activate transcription because of a specific defect in transcriptional activation and not as a result of improper cellular localization.

Figure 7A shows the results of such a repression assay, with the quantitation shown in Figure 7B. As a reference, we assayed repression by different levels of the transcriptionally inactive Oct-2 POU domain (Δ.POU.Δ) alone, which we have shown previously will repress transcription in a related assay (45). The 47-fold stimulation of the reporter by the wild-type Q^{18}III protein (compare lanes 1 and 2, Fig. 7A) was repressed up to fourfold by increasing levels of expression of the Oct-2 POU domain alone (compare lanes 3 to 5 with lane 2). The four inactive Q^{18}III-related (Q^{18}II and Q^{19}Oct-1) and mutant [Q^{18}III(L→A) and Q^{18}III(Q→A)] segments also repressed transcription (Fig. 7A, compare lanes 6 to 9 with lane 2 and see Fig. 7B); the Q^{18}II segment, however, was a significantly less effective repressor than the Oct-2 POU domain alone. This result may indicate that the Q^{18}II protein is not localized as efficiently to the reporter as the other proteins, but this does not explain the nearly complete lack of activity of this protein in the in vivo assay (Fig. 6A, lanes 11 to 13).

In separate assays (data not shown), the defective Oct-2 POU domain-containing activators did not repress the ability of the wild-type GAL4-Q^{18}III activator (see Fig. 4) to stimulate transcription from GAL4-binding sites, indicating that the repression of Q^{18}III activity shown in Fig. 7 does not result from competition between activation domains but rather from competition for the promoter-binding sites. We therefore conclude that the four defective Q^{18}III-related and mutant POU domain activators are specifically defective for transcriptional activation in vivo and are not defective simply because they cannot reach the promoter appropriately in vivo.

Q^{18} segments that are inactive in vivo can activate transcription in vitro. To dissect the function of the reiterated Q segment activation domains in more detail, we compared the transcriptional activities of a set of Q^{18}III-related and mutant proteins in vitro. In addition to the wild-type Q^{18}III POU domain protein, we chose proteins that were either active [Q^{18}III(S/T→A)] or inactive [Q^{18}II, Q^{19}Oct-1, Q^{18}III(L→A), and Q^{18}III(Q→A)] in the in vivo assay (see Fig. 6A). The wild-type, related, and mutant Q^{18}III POU domain activators were expressed in *E. coli* and purified by DNA affinity chromatography. In addition to a purified activator, the reconstituted in vitro transcription assay mixes contained partially purified TFIIA and TFIID, purified RNA polymerase II, and a crude 0.3 to 0.5 M KCl eluate of a nuclear extract passed over phosphocellulose (see Materials and Methods).

Figure 6B shows the in vitro activity of this set of reiterated

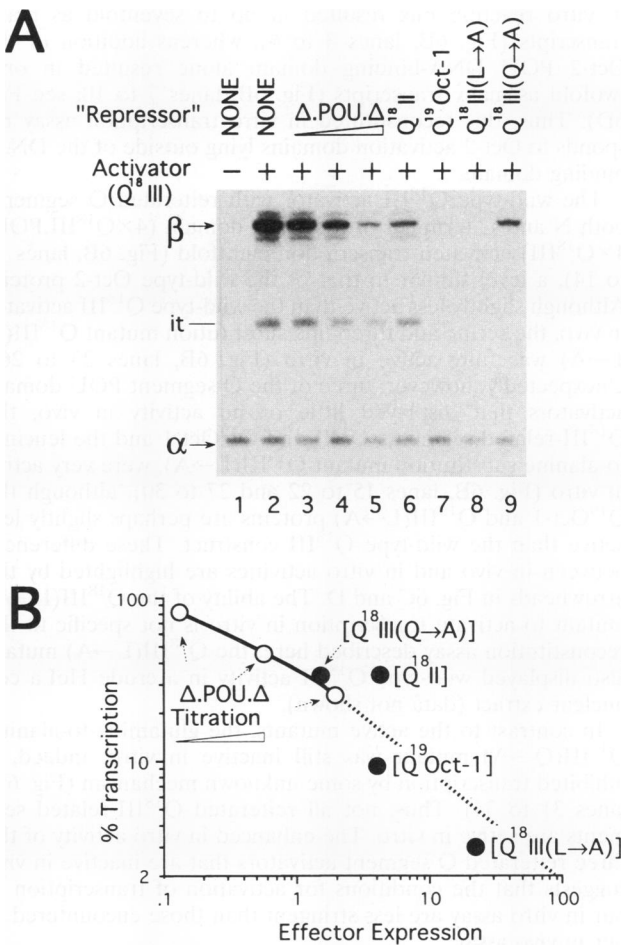

FIG. 7. Transcriptionally inactive reiterated Q segments fused to the Oct-2 POU domain can repress transcription by the active 4×Q^{18}III.POU.4×Q^{18}III activator in vivo in a competition assay. (A) RNase protection. HeLa cells were transfected with the β-globin reporter pβΔ^{36}6×oct^{+}sph^{-} and the α-globin internal reference pα4×(A+C) along with the pCG 4×Q^{18}III.POU.4×Q^{18}III activator and pCG competitor expression plasmids. Lane 1, empty pCG effector plasmid included; lane 2, pCG-4×Q^{18}III.POU.4×Q^{18}III alone; lanes 3 to 5, pCG-4×Q^{18}III.POU.4×Q^{18}III with a serial titration of threefold dilutions of the pCG-Δ.POU.Δ plasmid; lanes 6 to 9, pCG-4×Q^{18}III.POU.4×Q^{18}III and the pCG effector plasmid indicated above each lane. The positions of the RNase-protected fragments for correctly initiated α-globin (α) and β-globin (β) and an incorrect β-globin (it) transcript are indicated. (B) Quantitation of effector repression. Levels of β-globin reporter RNA (percent transcription) are plotted on a log scale relative to the level of β-globin RNA without coexpressed repressor protein (lane 2 in A) and according to the level of effector DNA-binding activity.

Q segment POU domain activators in sets of four twofold titrations of effector DNA-binding activity; Fig. 6D shows a quantitation of the results. In the absence of added activator, addition of the two immunoglobulin octamer sites to the β-globin promoter fails to activate transcription (Fig. 6B, compare lanes 1 and 2). This lack of transcriptional activity is probably because there is very little endogenous Oct-1 activity present in this reconstitution assay; the majority of the Oct-1 elutes from the phosphocellulose column in the 0.1 to 0.3 M KCl fraction, which was not used in the transcription assay (data not shown). Addition of purified wild-type Oct-2 to the

in vitro reaction mix resulted in up to sevenfold as many transcripts (Fig. 6B, lanes 3 to 6), whereas addition of the Oct-2 POU DNA-binding domain alone resulted in only twofold as many transcripts (Fig. 6B, lanes 7 to 10; see Fig. 6D). Thus, this reconstituted in vitro transcription assay responds to Oct-2 activation domains lying outside of the DNA-binding domain.

The wild-type Q^{18}III activator with reiterated Q segments both N and C terminal of the POU domain ($4\times Q^{18}$III.POU.$4\times Q^{18}$III) activated transcription eightfold (Fig. 6B, lanes 11 to 14), a level similar to that of the wild-type Oct-2 protein. Although slightly less active than the wild-type Q^{18}III activator in vivo, the serine and threonine substitution mutant Q^{18}III(S/T→A) was fully active in vitro (Fig. 6B, lanes 23 to 26). Unexpectedly, however, three of the Q segment POU domain activators that displayed little or no activity in vivo, the Q^{18}III-related segments Q^{18}II and Q^{19}Oct-1 and the leucine-to-alanine substitution mutant Q^{18}III(L→A), were very active in vitro (Fig. 6B, lanes 15 to 22 and 27 to 30), although the Q^{19}Oct-1 and Q^{18}III(L→A) proteins are perhaps slightly less active than the wild-type Q^{18}III construct. These differences between in vivo and in vitro activities are highlighted by the arrowheads in Fig. 6C and D. The ability of the Q^{18}III(L→A) mutant to activate transcription in vitro is not specific to the reconstitution assay described here; the Q^{18}III(L→A) mutant also displayed wild-type Q^{18}III activity in a crude HeLa cell nuclear extract (data not shown).

In contrast to the active mutants, the glutamine-to-alanine Q^{18}III(Q→A) mutant was still inactive in vitro; indeed, it inhibited transcription by some unknown mechanism (Fig. 6B, lanes 31 to 34). Thus, not all reiterated Q^{18}III-related segments are active in vitro. The enhanced in vitro activity of the three reiterated Q segment activators that are inactive in vivo suggests that the conditions for activation of transcription in our in vitro assay are less stringent than those encountered in our in vivo assay.

Q^{18}II but not Q^{18}III(L→A) segment is active in vivo when fused to the GAL4 DNA-binding domain. To assay the in vivo activity of reiterated copies of the Q^{18}III-related Q^{18}II and mutant Q^{18}III(L→A) segments when fused to a different DNA-binding domain, we fused four tandem copies to the GAL4 DNA-binding domain. Figure 8 shows that, consistent with the inactivity of the Q^{18}III(L→A) POU domain activator in vivo, this activation domain remains inactive when fused to the GAL4 DNA-binding domain (compare lanes 6 to 8 with lanes 12 to 14). Thus, the ability of this activation domain to activate transcription in vitro appears to be specific to the in vitro assay and not a peculiarity of the activity of the Q^{18}III(L→A) POU domain construct in vivo. Surprisingly, however, the reiterated Q^{18}II activation domain, which displays little in vivo activity when fused to the Oct-2 POU domain (see Fig. 6A), is very active in vivo when fused to the GAL4 DNA-binding domain (Fig. 8, compare lanes 9 to 11 and 6 to 8). This difference is not due to the different promoters (β-globin and c-*fos*) used in these assays because the Q^{18}II POU domain activator is also inactive on a c-*fos* promoter containing octamer sites (data not shown). These results show that functional transcriptional activation domains can be generated by reiteration of different short peptide sequences, but the activities can be DNA-binding domain specific.

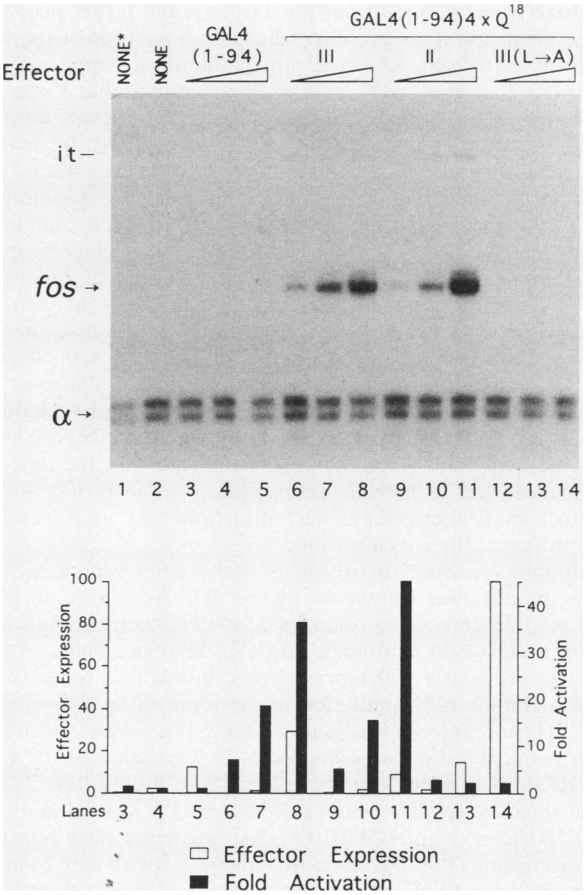

FIG. 8. $4\times Q^{18}$II activates transcription when fused to the GAL4 DNA-binding domain. HeLa cells were transfected with the pc-*fos*/−56/4×G17M reporter, α-globin internal reference, and pCG effector expression plasmids encoding the GAL4 DNA-binding domain alone [GAL4(1–94), lanes 3 to 5] or fused to the $4\times Q^{18}$III (lanes 6 to 8), $4\times Q^{18}$II (lanes 9 to 11), or $4\times Q^{18}$III(L→A) (lanes 12 to 14) activation domain. The positions of RNase-protected fragments corresponding to correctly initiated α-globin (α) and c-*fos* (*fos*) transcripts and incorrect *fos*-vector transcripts (it) are indicated. At the bottom is shown the quantitation of effector protein expression (open bars) determined by electrophoretic mobility retardation and transcriptional activation (solid bars) relative to the level without effector.

DISCUSSION

We have shown that potent transcriptional activation domains can be reconstituted by tandem reiteration of short 18-amino-acid-long peptide segments derived from a glutamine-rich activation domain in the POU domain activator Oct-2. Such reconstitution of a transcriptional activation domain is not limited to glutamine-rich sequences. In an extension of this approach, Seipel et al. (38) showed that duplication of a 12-amino-acid segment derived from the very potent and acidic VP16 activation domain can also generate an activation domain. Indeed, increasing numbers of 8-amino-acid segments from two separate regions of the VP16 activation domain can yield increasingly more active transcriptional activation domains (42). Thus, reiteration of different segments derived from each of two different types of transcriptional activation domains can result in the reconstitution of new transcriptional activation domains. These results suggest that these two different types of activation domain have similar repetitive and flexible structures.

Enhancers and activation domains display structural and functional similarity. The repetitive and flexible nature of activation domains is similar to the repetitive and flexible

nature of transcriptional enhancers. This similarity may not be fortuitous, because enhancers and activation domains are part of the same transcriptional regulatory pathway. In protein-encoding genes, enhancers are *cis*-regulatory signals that determine the assembly of a specific combination of DNA-binding transcription factors, which, through their activation domains, transmit regulatory signals that are ultimately destined to control transcription by RNA polymerase II. The interaction surface of the transcription factors that transmits the signal—i.e., the assemblage of activation domains—is therefore specified by the enhancer to which the transcription factors are bound.

Enhancers display a hierarchical structure (6, 10, 28, 44; reviewed in reference 14). "Proto-enhancers" form a functional unit of enhancer structure; they can cooperate with duplicates of themselves or with other proto-enhancers to activate transcription without strict spacing requirements. Proto-enhancers are composed of one or more enhansons—the basic units of enhancer structure that represent binding sites for the transcription factors. Our studies suggest that transcriptional activation domain synergy is also hierarchical. At one level, as in the reiterated Q^{18}III activation domains, short segments that display little activity on their own can synergize with adjacent copies of themselves (Fig. 2). At a second level, as with the 4×Q^{18}III activation domain, activation domains can synergize with one another from N- and C-terminal sides of the activator (Fig. 3) (43). Although a precise relationship between transcriptional activation domain synergy and enhancer element synergy remains to be elucidated, it is intriguing that in both cases, an increase in the number of an elementary unit can lead to a synergistic increase in transcriptional activity.

At the first level of the activation domain synergy, i.e., generation of a transcriptional activation domain, increasing numbers of reiterated Q^{18}III segments display a bimodal pattern of activation: one and two Q^{18}III segments display little activity, whereas three to five tandem segments display increasing activity (Fig. 2). We observe the same bimodal activation pattern with one to three Q^{18}III segments attached to the GAL4 DNA-binding domain (46), indicating that the relationship between segment number and activity is not strictly DNA-binding domain dependent. This pattern of response to Q^{18}III segment number suggests that three Q^{18}III segments form a minimal activation domain structure and subsequent lengthening of the activation domain results in even stronger activation domains.

The increase in activity for activation domains containing three to five Q^{18}III segments may be due to an increase in affinity of the activation domain for its target or receptor in the transcriptional apparatus. We envision two non-mutually exclusive explanations for an increase in affinity: (i) the affinity increases because there is an increase in the number of surfaces that bind simultaneously to the receptor, such as in the case of DNA-binding proteins containing multiple tandem TFIIIA-like zinc fingers that bind to the DNA simultaneously (31); and (ii) the affinity increases because of an increase in the probability of binding the receptor because of a larger number of interaction surfaces, resulting in a higher concentration of activation domains.

Whichever the mechanism, the synergy among tandemly reiterated Q^{18}III segments is remarkably similar to the synergy observed with tandemly reiterated enhancer elements (29). Because reiterated enhancers can amplify the activity of a single regulatory element, they have been excellent reagents with which to study the structure and function of enhancers and to characterize the activity of transcription factors (14). Reiterated activation domains should also be excellent reagents with which to study the structure and function of activation domains and the targets or receptors for the activation domains in the transcriptional machinery. Indeed, the wild-type and mutant reiterated Q segments described here have already been useful in characterizing a potential Q domain-associated factor (the QA factor [43]) and the promoter selectivity (5) and species specificity (15) of activation domains.

Transcriptional activation domain activity in vitro does not necessarily correspond to activity in vivo. One strategy to identify the receptors of transcriptional activation domains is to study their activity in vitro, where the receptor can potentially be fractionated and purified. One of our unexpected results is that reiterated activation domains fused to the Oct-2 POU domain that fail to activate the β-globin promoter in vivo can activate the same promoter in vitro, suggesting that the requirements for promoter activation in vitro are less stringent than they are in vivo. Curiously, however, the reiterated activation domains that are not active in vivo but activate transcription in our in vitro assay [4×Q^{18}II, 4×Q^{19}Oct-1, and 4×Q^{18}III(L→A)] do activate transcription in vivo, but only under different conditions: the 4×Q^{18}II activation domain is very active in vivo when attached to the GAL4 DNA-binding domain (Fig. 8), and the 4×Q^{19}Oct-1 and 4×Q^{18}III(L→A) activation domains can activate RNA polymerase II small nuclear RNA promoters (5). Thus, perhaps our in vitro assay does not faithfully reproduce the DNA-binding domain specificity and the promoter selectivity that we observe in vivo.

Differences in activation domain activity in vitro and in vivo also exist with other activation domains. Gerber et al. (10a) have shown that homopolymeric stretches of glutamines (or prolines) can activate transcription in vitro but are inactive in vivo unless an SV40 enhancer is provided to boost activity, in which case the best activators display three- to fourfold activation of transcription. Because in our studies only the construct lacking glutamines [4×Q^{18}III(Q→A)] is inactive in vitro, perhaps activation domain function in vitro requires only a sufficient concentration of glutamines. Our results show, however, that more than glutamines are required for effective enhancer-independent activation in vivo of a β-globin or c-*fos* promoter.

Activation domains can display DNA-binding domain specificity. An unexpected result described here is that activation domains, even ones as apparently closely related as the 4×Q^{18}II and 4×Q^{18}III activation domains, can display very different activities depending on the DNA-binding domain and hence the *cis*-regulatory site to which they are attached. Both activation domains are active when fused to the GAL4 DNA-binding domain (Fig. 8), but only the 4×Q^{18}III activation domain (in two copies) was active when attached to the Oct-2 POU domain (Fig. 6). In the accompanying article (43), we describe the opposite DNA-binding domain preference: both the Oct-2 Q and P domains are active when fused to the POU domain, but only the P domain activated transcription effectively when fused to the GAL4 DNA-binding domain. DNA-binding domains also influence the number of activation domains that must be fused to the DNA-binding domain for activity. For example, in our assays, one 4×Q^{18}III activation domain is active when linked to the GAL4 DNA-binding domain but not when linked to the Oct-2 POU domain, perhaps because on our reporters GAL4 binds as a dimer and Oct-2 binds as a monomer. These results emphasize that, although transcription factors display considerable flexibility and modularity (32), there can be dramatic combinatorial effects between activation domains and DNA-binding domains.

Mechanism of transcriptional activation domain function. The structural flexibility observed here with the reiterated transcriptional activation domains is consistent with the results of structural studies of transcriptional activation domains. Analysis of both the VP16 and GAL4 C-terminal transcriptional activation domains reveals little structure at neutral pH (7, 27, 49), yet structure can be induced by placing these activation domains in different environments, such as at an acidic pH (7, 49). These properties are consistent with a flexible activation domain structure that adopts a structure upon binding to the receptor in the transcriptional machinery. The lack of strong sequence conservation among transcriptional activation domains suggests that receptors in the transcriptional machinery accommodate a diverse set of sequences. Such molecular interactions between a diverse set of sequences and a single receptor exist between peptide antigens and major histocompatibility (MHC) class I and class II molecules, and between unfolded proteins and molecular chaperones such as HSP70. The interaction of antigenic peptides with MHC class I and class II molecules is understood in great detail: a large variety of different peptide antigen sequences fit in a rigid groove or cleft on the surface of the MHC molecule (2, 9, 22, 51). Perhaps receptors for transcriptional activation domains in the transcriptional machinery possess a structure similar to that of the antigen-binding cleft of MHC class I or class II molecules to permit binding to a diverse set of transcriptional activation domains. Such molecular interaction would enforce minimal structural requirements on the transcriptional activation domain.

Relaxed activation domain structure provides considerable evolutionary advantages. A central issue in transcriptional regulation of protein-encoding genes is that a large variety of activators, principally sequence-specific DNA-binding proteins, ultimately activate transcription by a single enzyme, RNA polymerase II. This process requires that the activity of a diverse set of DNA sequence-specific transcription factors be focused into interaction with an undoubtedly smaller number of receptor molecules in the transcriptional machinery. If these interactions are flexible, there are fewer constraints on the evolution of different transcriptional activators. The specificity of transcriptional activation would be largely, albeit not entirely, specified by the DNA-binding domain, and the activation domain would funnel the response to a smaller set of molecules, such as coactivators (33) and TBP-associated factors (TAFs) (8, 48).

ACKNOWLEDGMENTS

We thank C. Hinkley for helpful discussions; D. McVey for preparing the purified RNA polymerase II; M. Carey and S. Lobo for advice on extract fractionation; G. Das and C. Hinkley for sharing unpublished results; and M. Cleary, N. Hernandez, A. Stenlund, and W. Tansey for critical readings of the manuscript.

These studies were supported by U.S. Public Health Service grant CA-13106 from the National Cancer Institute.

REFERENCES

1. **Brent, R., and M. Ptashne.** 1985. A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. Cell **43**:729–736.
2. **Brown, J. H., T. S. Jardetsky, J. C. Gorga, L. J. Stern, R. G. Urban, J. L. Strominger, and D. C. Wiley.** 1993. Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. Nature (London) **364**:33–39.
3. **Clerc, R. G., L. M. Corcoran, J. H. LeBowitz, D. Baltimore, and P. A. Sharp.** 1988. The B-cell-specific Oct-2 protein contains POU box- and homeo box-type domains. Genes Dev. **2**:1570–1581.
4. **Cress, W. D., and S. J. Triezenberg.** 1991. Critical structural elements of the VP16 transcriptional activation domain. Science **251**:87–90.
5. **Das, G., and W. Herr.** Unpublished results.
6. **Davidson, I., J. H. Xiao, R. Rosales, A. Staub, and P. Chambon.** 1988. The HeLa cell protein TEF-1 binds specifically and cooperatively to two SV40 enhancer motifs of unrelated sequence. Cell **54**:931–942.
7. **Donaldson, L., and J. P. Capone.** 1992. Purification and characterization of the carboxyl-terminal transactivation domain of Vmw65 from herpes simplex virus type 1. J. Biol. Chem. **267**:1411–1414.
8. **Dynlacht, B. D., T. Hoey, and R. Tjian.** 1991. Isolation of coactivators associated with the TATA-binding protein that mediate transcriptional activation. Cell **66**:563–576.
9. **Fremont, D. H., M. Matsumura, E. A. Stura, P. A. Peterson, and I. A. Wilson.** 1992. Crystal structures of two viral peptides in complex with murine MHC class I H-2K^b. Science **257**:919–927.
10. **Fromental, C., M. Kanno, H. Nomiyama, and P. Chambon.** 1988. Cooperativity and hierarchical levels of functional organization in the SV40 enhancer. Cell **54**:943–953.

10a. **Gerber, H.-P., K. Seipel, O. Georgiev, M. Hofferer, M. Hug, S. Rusconi, and W. Schaffner.** 1994. Transcriptional activation modulated by homopolymeric glutamine and proline stretches. Science **263**:808–811.

11. **Gerster, T., C.-G. Balmaceda, and R. G. Roeder.** 1990. The cell type-specific octamer transcription factor OTF-2 has two domains required for the activation of transcription. EMBO J. **9**:1635–1643.
12. **Gill, G., E. Pascal, Z. H. Tseng, and R. Tjian.** 1994. A glutamine-rich hydrophobic patch in transcription factor Sp1 contacts the dTAF$_{II}$110 component of the *Drosophila* TFIID complex and mediates transcriptional activation. Proc. Natl. Acad. Sci. USA **91**:192–196.
13. **Gill, G., and M. Ptashne.** 1987. Mutants of GAL4 protein altered in an activation function. Cell **51**:121–126.
14. **Herr, W.** 1993. The SV40 enhancer: transcriptional regulation through a hierarchy of combinatorial interactions. Semin. Virol. **4**:3–13.
15. **Hinkley, C., and W. Herr.** Unpublished results.
16. **Hope, I. A., S. Mahadevan, and K. Struhl.** 1988. Structural and functional characterization of the short acidic transcriptional activation region of yeast GCN4 protein. Nature (London) **333**:635–640.
17. **Hope, I. A., and K. Struhl.** 1986. Functional dissection of a eukaryotic transcriptional activator protein, GCN4 of yeast. Cell **46**:885–894.
18. **Kadonaga, J. T., and R. Tjian.** 1986. Affinity purification of sequence-specific DNA binding proteins. Proc. Natl. Acad. Sci. USA **83**:5889–5893.
19. **Klemm, J. D., M. A. Rould, R. Aurora, W. Herr, and C. O. Pabo.** 1994. Crystal structure of the Oct-1 POU domain bound to an octamer site: DNA recognition with tethered DNA-binding modules. Cell **77**:21–32.
20. **LeBowitz, J. H., R. G. Clerc, M. Brenowitz, and P. A. Sharp.** 1989. The Oct-2 protein binds cooperatively to adjacent octamer sites. Genes Dev. **3**:1625–1638.
21. **Lobo, S. M., M. Tanaka, M. L. Sullivan, and N. Hernandez.** 1992. A TBP complex essential for transcription from TATA-less but not TATA-containing RNA polymerase III promoters is part of the TFIIIB fraction. Cell **71**:1029–1040.
22. **Madden, D. R., J. C. Gorga, J. L. Strominger, and D. C. Wiley.** 1992. The three-dimensional structure of HLA-B27 at 2.1 Å resolution suggests a general mechanism for tight peptide binding to MHC. Cell **70**:1035–1048.
23. **Marmorstein, R., M. Carey, M. Ptashne, and S. C. Harrison.** 1992. DNA recognition by GAL4: structure of a protein-DNA complex. Nature (London) **356**:408–414.
24. **Mitchell, P. J., and R. Tjian.** 1989. Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. Science **245**:371–378.
25. **Müller, M. M., S. Ruppert, W. Schaffner, and P. Matthias.** 1988. A cloned octamer transcription factor stimulates transcription from lymphoid-specific promoters in non-B cells. Nature (London) **336**:544–551.

26. **Müller-Immerglück, M. M., W. Schaffner, and P. Matthias.** 1990. Transcription factor Oct-2A contains functionally redundant activating domains and works selectively from a promoter but not from a remote enhancer position in non-lymphoid (HeLa) cells. EMBO J. **9**:1625–1634.
27. **O'Hare, P., and G. Williams.** 1992. Structural studies of the acidic transactivation domain of the Vmw65 protein of herpes simplex virus using ^{1}H NMR. Biochemistry **31**:4150–4156.
28. **Ondek, B., L. Gloss, and W. Herr.** 1988. The SV40 enhancer contains two distinct levels of organization. Nature (London) **333**:40–45.
29. **Ondek, B., A. Shepard, and W. Herr.** 1987. Discrete elements within the SV40 enhancer region display different cell-specific enhancer activities. EMBO J. **6**:1017–1025.
30. **Pabo, C. O., and R. T. Sauer.** 1992. Transcription factors: structural families and principles of DNA recognition. Annu. Rev. Biochem. **61**:1053–1095.
31. **Pavletich, N. P., and C. O. Pabo.** 1991. Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 Å. Science **252**:809–817.
32. **Ptashne, M.** 1988. How eukaryotic transcriptional activators work. Nature (London) **335**:683–689.
33. **Pugh, B. F., and R. Tjian.** 1990. Mechanism of transcriptional activation by Sp1: evidence for coactivators. Cell **61**:1187–1197.
34. **Regier, J. L., F. Shen, and S. J. Triezenberg.** 1993. Pattern of aromatic and hydrophobic amino acids critical for one of two subdomains of the VP16 transcriptional activator. Proc. Natl. Acad. Sci. USA **90**:883–887.
35. **Reinberg, D., and R. G. Roeder.** 1987. Factors involved in specific transcription by mammalian RNA polymerase II: purification and functional analysis of initiation factors IIB and IIE. J. Biol. Chem. **262**:3310–3321.
36. **Scheidereit, C., J. A. Cromlish, T. Gerster, K. Kawakami, C.-G. Balmaceda, R. A. Currie, and R. G. Roeder.** 1988. A human lymphoid-specific transcription factor that activates immunoglobulin genes is a homeobox protein. Nature (London) **336**:551–557.
37. **Schirm, S., J. Jiricny, and W. Schaffner.** 1987. The SV40 enhancer can be dissected into multiple segments, each with a different cell type specificity. Genes Dev. **1**:65–74.
38. **Seipel, K., O. Georgiev, and W. Schaffner.** 1992. Different activation domains stimulate transcription from remote ('enhancer') and promixal ('promoter') positions. EMBO J. **11**:4961–4968.
39. **Studier, F. W., A. H. Rosenberg, J. J. Dunn, and J. W. Dubendorff.** 1990. Use of T7 RNA polymerase to direct expression of cloned genes. Methods Enzymol. **185**:60–89.
40. **Sturm, R. A., G. Das, and W. Herr.** 1988. The ubiquitous octamer-binding protein Oct-1 contains a POU domain with a homeo box subdomain. Genes Dev. **2**:1582–1599.
41. **Sumimoto, H., Y. Ohkuma, T. Yamamoto, M. Horikoshi, and R. G. Roeder.** 1990. Factors involved in specific transcription by mammalian RNA polymerase II: identification of general transcription factor TFIIG. Proc. Natl. Acad. Sci. USA **87**:9158–9162.
42. **Tanaka, M.** Unpublished results.
43. **Tanaka, M., W. M. Clouston, and W. Herr.** 1994. The Oct-2 glutamine-rich and proline-rich activation domains can synergize with each other or duplicates of themselves to activate transcription. Mol. Cell. Biol. **14**:6046–6055.
44. **Tanaka, M., U. Grossniklaus, W. Herr, and N. Hernandez.** 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.
45. **Tanaka, M., and W. Herr.** 1990. Differential transcriptional activation by Oct-1 and Oct-2: interdependent activation domains induce Oct-2 phosphorylation. Cell **60**:375–386.
46. **Tanaka, M., C. Hinkley, and W. Herr.** Unpublished results.
47. **Tanaka, M., J.-S. Lai, and W. Herr.** 1992. Promoter-selective activation domains in Oct-1 and Oct-2 direct differential activation of an snRNA and mRNA promoter. Cell **68**:755–767.
48. **Tanese, N., B. F. Pugh, and R. Tjian.** 1991. Coactivators for a proline-rich activator purified from the multisubunit human TFIID complex. Genes Dev. **5**:2212–2224.
49. **Van Hoy, M., K. K. Leuther, T. Kodadek, and S. A. Johnston.** 1993. The acidic activation domains of the GCN4 and GAL4 proteins are not α helical but form β sheets. Cell **72**:587–594.
50. **Veldman, G. M., S. Lupton, and R. Kamen.** 1985. Polyomavirus enhancer contains multiple redundant sequence elements that activate both DNA replication and gene expression. Mol. Cell. Biol. **5**:649–658.
51. **Young, A. C. M., W. Zhang, J. C. Sacchettini, and S. G. Nathenson.** 1994. The three-dimensional structure of H-2D^b at 2.4 Å resolution: implications for antigen-determinant selection. Cell **76**:39–50.