Human ZFM1 Protein Is a Transcriptional Repressor That Interacts with the Transcription Activation Domain of Stage-specific Activator Protein

Di Zhang and Geoffrey Childs

From the Department of Molecular Genetics, Albert Einstein College of Medicine, Bronx, New York 10461

Stage-specific activator protein (SSAP) is the transcription factor responsible for the activation of the sea urchin late H1 gene at the mid-blastula stage of embryogenesis. SSAP contains an extremely potent transcription activation domain that functions 4–5-fold better than VP16 in a variety of mammalian cell lines. We used the two-hybrid screening technique to identify human cDNAs from an HL60 cell-derived cDNA library that encode proteins that interact with the transcription activation domain of SSAP. One of these cDNAs encodes ZFM1, a protein previously identified at the locus linked to multiple endocrine neoplasia type 1 (MEN1) and as presplicing factor SF1. Functional assays establish the ZFM1 protein as a transcriptional repressor. ZFM1 protein represses Gal4-GQC-mediated transcription, and this activity requires both a repression domain found in the N-terminal 137 amino acids of the protein, as well as a GQC interaction region. The physiological significance of repression mediated by ZFM1 comes from the ability of its specific repression domain to function when fused to Gal4 and tethered to promoters containing Gal4 binding sites. The activity is unique in that activated but not basal transcription levels are affected.

Different families of histone genes in sea urchins are expressed with distinct temporal patterns during early embryogenesis, making this system ideal for studying mechanisms of temporal gene expression. In recent years, much progress has been made in studying the expression of late H1 histone subtype genes (1–4). The late H1-β gene is transcribed at low levels until it is transcriptionally activated at the mid-blastula stage, and reaches its peak level of expression in 24-h late-blastula stage embryos. The correct temporal activation of late H1 gene expression is determined by an enhancer located 220–280 base pairs upstream in its promoter region (1). Stage-specific activator protein (SSAP)1 is a 43-kDa polypeptide that can specifically bind to this enhancer (2). Early in development, SSAP is present as a monomer; however, it undergoes a posttranslational modification at about 12 h after fertilization, and dimerization coincides with the activation of late H1 gene (2). Synthetic SSAP mRNA injected into zygotes transactivates reporter genes containing SSAP binding sites (3). Significantly, this transactivation also occurs in a temporal-specific manner.

The DNA binding activity of SSAP maps to its N-terminal 180 amino acids (3). This domain contains two RNA recognition motifs, which recognize both double-stranded and single-stranded DNA in a sequence-specific manner (3). In addition to this novel DNA binding domain, SSAP contains an extremely potent transcription activation domain consisting of amino acids 181–404 (4). This activation domain consists of a central glycine/glutamine-rich sequence and a C-terminal region rich in serine/threonine and basic amino acids (referred to as the GQC domain). In a variety of mammalian cell lines, Gal4-GQC fusion protein can transactivate the expression of Gal4-responsive reporter genes as much as 4–5-fold better than Gal4-VP16 (4). This activity requires the presence of both the central G/Q-rich domain as well as the C-terminal domain. The GQ region alone cannot activate transcription, whereas the C-terminal region has minimal activity (4). The activation domain of a transcription factor may promote transcription at several different steps during the formation of a preinitiation complex by recruiting basal transcription factors or during promoter clearance and elongation (5–7). To do this, the activators must interact with multiple targets in the transcriptional apparatus. These targets include not only basal factors in the transcription machinery, but also various co-activators or adaptors, which facilitate interactions between activator and basal transcription factors or RNA polymerase II (8, 9). In vitro binding assays have shown that the GQC domain can physically interact with several basal transcription factors of RNA polymerase II, including TBP, TFIIB, TFIIF74, and dTAFII110 (4). On the other hand, the GQC domain has self-squelching activity and the ability to squelch VP16- and E1A-driven reporter genes. This suggests that the activation domains of SSAP, VP16, and E1a share some common targets necessary for maximal transcription. It is believed that these targets may not be basal transcription factors, instead, they may be specific adaptors or coactivator proteins (4). Although the ability of the activation domain of SSAP to drive such high levels of transcription may be explained by stronger association with common targets, it is also possible that it may interact with unique protein targets not shared by VP16 or E1a. It has been well documented that different classes of activation domains may interact with distinct coactivators or adaptors to stimulate transcription (8). For example, the glutamine-rich activation domain of SP1 binds specifically to TAFII110 (10, 11), whereas the acidic activation domain of VP16 cannot stimulate transcription without interaction with TAFII40 (12). Similarly, a group of transcription factors require p300/CBP coactivator family members for max-
imal activity, whereas other transcription factors act independently of this regulatory network (13–15).

In this paper, a two-hybrid screening technique is employed to identify human protein(s) that can interact with the activation domain of SSAP. The rationale for this screen comes from the extremely potent activation domain used as a "bait" in this study. We believed that the GQC activation domain would interact with interesting transcriptional mediators. This paper describes the properties of one SSAP-interacting protein, ZFM1. ZFM1 was previously cloned as a nuclear protein at a locus linked to multiple endocrine neoplasia type 1 (16). It was also identified as a presplicing factor SF1 (17). Here, we show that ZFM1 functions as a transcription repressor.

**EXPERIMENTAL PROCEDURES**

**Plasmids**—To generate pBTM116-GQC as bait for two-hybrid screening, an EcoRI fragment derived from pSG424-GQC (4), which contains the entire GQC domain (amino acids 181–404), was ligated into pBTM116 vector cut with EcoRI. The resulting construct was sequenced to verify that the GQC domain was cloned in the correct orientation and in frame with the LexA DNA binding domain.

To create pGAD424-ZFM1-E for yeast two-hybrid assays, two primers (5'-GGG GTC GAC CAG ACC ACA TGG CGA CCG GAG CGA AC-3') and (5'-GCC TTC GAC TCA CTT GTC ACG GTC CTT GTA GTC CCA ATG GGC GCG GAA AGT-3') were used in a PCR reaction to amplify ZFM1-E from pBluescript SK phagemid DNA. This PCR product was cut by SalI and ligated into the SalI site of pGAD424 to fuse ZFM1-E in frame with the Gal4 activation domain.

To generate pCR3.1-ZFM1-A expression construct for functional assays, two PCR fragments that correspond to N-terminal (between nucleotides 383–1607) and C-terminal (between nucleotides 836 and 2251) regions of ZFM1-A, respectively, were created. The N-terminal piece was amplified using template ZFM1-E in pBluescript SK and two primers (5'-GGG GTC GAC CCG CCA CCA TGA GCG CAG CGA AC-3') and (5'-CAT GCC CAC CTG TCA GGC C-3'), and the C-terminal piece was amplified using template CE9 and two primers (5'-TTC GTC GGG GGC GGA TCC AAG CTT ACA TCA-3') and (5'-GAC GCG TCG ACT CAC TTG TCA-3'). The two non-overlapping portions were joined together and ligated to pcR3.1 vector (Invitrogen) in a three-piece ligation reaction. In the resulting plasmid, pCR3.1-ZFM1-A, the expression of ZFM1-321–478 with a FLAG-tag on its C terminus is under the control of cytomegalovirus promoter, and there is a FLAG-tag in frame with the LexA DNA binding domain. The two non-overlapping portions were joined together and ligated to pcR3.1 vector (Invitrogen) in a three-piece ligation reaction. In the resulting plasmid, pCR3.1-ZFM1-A, the expression of ZFM1-321–478 with a FLAG-tag on its C terminus is under the control of cytomegalovirus promoter, and there is a FLAG-tag in frame with the LexA DNA binding domain. The two non-overlapping portions were joined together and ligated to pcR3.1 vector (Invitrogen) in a three-piece ligation reaction. In the resulting plasmid, pCR3.1-ZFM1-A, the expression of ZFM1-321–478 with a FLAG-tag on its C terminus is under the control of cytomegalovirus promoter, and there is a FLAG-tag in frame with the LexA DNA binding domain. The two non-overlapping portions were joined together and ligated to pcR3.1 vector (Invitrogen) in a three-piece ligation reaction. In the resulting plasmid, pCR3.1-ZFM1-A, the expression of ZFM1-321–478 with a FLAG-tag on its C terminus is under the control of cytomegalovirus promoter, and there is a FLAG-tag in frame with the LexA DNA binding domain. The two non-overlapping portions were joined together and ligated to pcR3.1 vector (Invitrogen) in a three-piece ligation reaction. In the resulting plasmid, pCR3.1-ZFM1-A, the expression of ZFM1-321–478 with a FLAG-tag on its C terminus is under the control of cytomegalovirus promoter, and there is a FLAG-tag in frame with the LexA DNA binding domain. The two non-overlapping portions were joined together and ligated to pcR3.1 vector (Invitrogen) in a three-piece ligation reaction. In the resulting plasmid, pCR3.1-ZFM1-A, the expression of ZFM1-321–478 with a FLAG-tag on its C terminus is under the control of cytomegalovirus promoter, and there is a FLAG-tag in frame with the LexA DNA binding domain. The two non-overlapping portions were joined together and ligated to pcR3.1 vector (Invitrogen) in a three-piece ligation reaction. In the resulting plasmid, pCR3.1-ZFM1-A, the expression of ZFM1-321–478 with a FLAG-tag on its C terminus is under the control of cytomegalovirus promoter, and there is a FLAG-tag in frame with the LexA DNA binding domain. The two non-overlapping portions were joined together and ligated to pcR3.1 vector (Invitrogen) in a three-piece ligation reaction. In the resulting plasmid, pCR3.1-ZFM1-A, the expression of ZFM1-321–478 with a FLAG-tag on its C terminus is under the control of cytomegalovirus promoter, and there is a FLAG-tag in frame with the LexA DNA binding domain. The two non-overlapping portions were joined together and ligated to pcR3.1 vector (Invitrogen) in a three-piece ligation reaction. In the resulting plasmid, pCR3.1-ZFM1-A, the expression of ZFM1-321–478 with a FLAG-tag on its C terminus is under the control of cytomegalovirus promoter, and there is a FLAG-tag in frame with the LexA DNA binding domain. The two non-overlapping portions were joined together and ligated to pcR3.1 vector (Invitrogen) in a three-piece ligation reaction. In the resulting plasmid, pCR3.1-ZFM1-A, the expression of ZFM1-321–478 with a FLAG-tag on its C terminus is under the control of cytomegalovirus promoter, and there is a FLAG-tag in frame with the LexA DNA binding domain. The two non-overlapping portions were joined together and ligated to pcR3.1 vector (Invitrogen) in a three-piece ligation reaction. In the resulting plasmid, pCR3.1-ZFM1-A, the expression of ZFM1-321–478 with a FLAG-tag on its C terminus is under the control of cytomegalovirus promoter, and there is a FLAG-tag in frame with the LexA DNA binding domain. The two non-overlapping portions were joined together and ligated to pcR3.1 vector (Invitrogen) in a three-piece ligation reaction. In the resulting plasmid, pCR3.1-ZFM1-A, the expression of ZFM1-321–478 with a FLAG-tag on its C terminus is under the control of cytomegalovirus promoter, and there is a FLAG-tag in frame with the LexA DNA binding domain.
most yeast activation domains are acidic (25). The GQC activation domain has a region rich in glycine and glutamine (GQ) and a C-terminal region (C) rich in serine, threonine, and basic amino acids (4). We also made two constructs that individually express LexA fusions to either the GQ or the C regions. As expected from the result in mammalian cells, neither LexA-GQ nor LexA-C can produce β-galactosidase activity by driving LacZ gene expression in yeast. For this reason, we decided to use the intact activation domain as bait and to inhibit growth of the parent strain used for our screen by addition of the competitive inhibitor of the histidine synthetase 3-AT to the medium. The background growth of LexA-GQC transformants can be eliminated by adding as little as 5 mM 3-AT in the medium. Therefore, we modified the original two-hybrid selection technique by screening library on plates containing 50 or 100 mM of 3-AT to suppress growth of cells not containing true interacting proteins.

A human cDNA library derived from an undifferentiated leukemia cell line (HL60) was screened using this LexA-GQC bait. From 2.85 x 10^7 yeast transformants, 38 colonies grew on His^- plates with 50 mM 3-AT. These colonies quickly turn dark blue using the β-galactosidase filter assay. Twenty of the 38 turned out to be false positives in genetic tests. The remaining 18 are true positives because their ability to activate LacZ and HIS3 genes depends solely on the presence of LexA-GQC. We also tested the interaction between these 18 fusion proteins and GAL4-GQC fusion protein in a different yeast strain, HF7c, which possesses a upstream activating sequence with Gal4 binding sites upstream of the LacZ gene. All 18 positives still interact with the GQC bait but not with plasmids encoding either Gal4 DNA binding domain alone or irrelevant Gal4 fusion proteins. Each fusion protein was also tested for interaction with LexA-GQ and LexA-C alone. There is detectable but weak interaction between the positives with the C domain, but none of them interacts with the GQ domain alone (data not shown). The plasmids from most of these positives were successfully isolated and sequenced. Nine of the 18 positives have identical inserts that encode a fragment corresponding to part of the ZFM1-A protein (GenBank™ accession no. D26120) (16), between amino acids 321 and 484 (ZFM1-(321–484)), which is the subject of this paper.

ZFM1 was first described by Toda et al. (16) as a nuclear protein at the locus tightly linked to a putative tumor suppressor gene for multiple endocrine neoplasia type 1. ZFM1 was subsequently isolated by Kramer and colleagues (17) as a splicing factor SF1 potentially functioning in pre-spliceosome assembly. Both groups reported that there are multiple splicing variants of ZFM1 in different tissues or cell types. At the same time, we tried to obtain cDNAs of ZFM1 for functional studies by screening a HeLa cell cDNA library using ZFM1-(321–484) as a probe. A number of overlapping cDNA clones were isolated. Interestingly, we isolated a new splicing variant not reported previously. This ZFM1 variant, designated as ZFM1-E, lacks 434 nucleotides between nucleotide position 1785 and 2218 of ZFM1-A and contains an open reading frame that differs from other reported ZFM1 isoforms by deleting part of the proline-rich domain but retaining the same C-terminal tail as ZFM1-B (Fig. 1). The exon-intron junction sites at positions of both 1785 and 2218 show a good match to the consensus splice site sequence. We constructed a clone encoding the full-length ZFM1-A variant by joining an N-terminal fragment amplified from ZFM1-E and a C-terminal fragment amplified

![Image](http://www.jbc.org/)

**Fig. 1.** Schematic representation of the domain structure of ZFM1-A and a newly identified splicing variant of ZFM1. Numbers above the cross-hatched boxes, which represent ZFM1 coding frames, refer to the amino acid sequence of ZFM1-A according to Toda et al. (16). The functional motifs found in ZFM1 and sequence homologies with other proteins are aligned and shown at top by closed boxes. The putative nuclear localization signal (NLS) is indicated by an asterisk. The region on ZFM1 that interacts with the GQC domain (amino acids 321–484) is indicated by a solid bar below the box representing ZFM1-A. ZFM1-E is a new variant that splices out nucleotides 1785–2218 of the ZFM1-A variant and resumes translation for 20 C-terminal amino acids of ZFM1-B (16), SLPAAAAMARMVRTRFRAHW*. This splice variant is missing about half of the region of homology with the WT1 and EGR2 proteins.
from CE9 (Fig. 1). Like ZFM1-(321–484), both ZFM1-E and ZFM1-A can interact with the GQC domain in a yeast two-hybrid assay (Table 1).

**ZFM1 Binds to the GQC Domain in Vitro**—We used GST pull-down assays to confirm the direct association of ZFM1 and GQC. ZFM1-(321–478), which covers most of the region of ZFM1-A that binds to GQC in yeast was fused in frame with glutathione S-transferase to express the GST-ZFM1-(321–478) chimeric protein in bacteria. Radiolabeled in vitro translated GQC was incubated with glutathione-Sepharose beads coupled with either GST-ZFM1-(321–478) or GST alone. After extensive washing, proteins retained on the beads are eluted and separated by SDS-PAGE. The GQC domain specifically associated with GST-ZFM1-(321–478) but not with GST alone (Fig. 2A). In parallel, we did the reverse experiment. In vitro translated radiolabeled ZFM1-E was incubated with GST-GQC protein. As in the previous experiment, ZFM1-E can bind to GST-GQC but not to GST alone (Fig. 2B).

**Overexpression of ZFM1-A in Mammalian Cells Represses the Transactivation of Reporter Constructs Driven by Gal4-GQC**—The GQC domain can function as a potent transcriptional activation domain in a variety of mammalian cell lines (4). To assess the consequences of an interaction between GQC and ZFM1, we asked whether overexpression of ZFM1 can modulate the transcriptional activity of the GQC domain. When pCR3.1-ZFM1-A is transfected into HepG2 cells, we detected a 70–80-kDa protein band that corresponds to the size predicted as ZFM1-A in Western blots using M2 antibody. Moreover, the level of expression of tagged ZFM1-A correlates well with the amount of DNA transfected (Fig. 3C). The open reading frame of ZFM1 contains a potential nuclear localization signal near its N terminus (16). We confirmed the nuclear localization of ZFM1 protein by immunostaining and Western blot analysis of the nuclear and cytoplasmic fractions of HepG2 cell extract (data not shown). When increasing amounts of pCR3.1-ZFM1-A are cotransfected into HepG2 cells along with pSG424-GQC and G5E1BCAT (a CAT reporter with five Gal4 DNA binding sites upstream in its promoter), we observed that the reporter expression is repressed by expression of ZFM1-A in a dose-dependent manner (Fig. 3A). At the highest levels of ZFM1-A, transcription driven by Gal4-GQC is repressed by 4–5-fold compared with cells containing endogenous levels of ZFM1 variants. In Fig. 3, we used a level of Gal4-GQC that is within the linear range of a titration curve of Gal4-GQC activator. (4). Higher levels of Gal4-GQC exhibit self-squelching, whereas lower levels give linear decreases in activation levels (4). We observe very similar 4–5-fold repression when different

![Fig. 2. ZFM1 binds to the GQC domain in vitro. A, in vitro translated [35S]methionine-labeled GQC protein was incubated with either GST-ZFM1-(321–478) fusion protein or GST alone immobilized on Sepharose beads, and proteins bound to washed beads were fractionated by SDS-PAGE and visualized by autoradiography. One-fiftieth of input of labeled GQC domain is shown in the left lane. B, in parallel, in vitro GST pull-down experiment was also performed using in vitro translated labeled ZFM1-E and GST-GQC.](http://www.jbc.org/)

![Fig. 3. Overexpression of ZFM1-A in mammalian cells represses the transactivation of G5E1BCAT reporter genes driven by Gal4-GQC. A, CAT assays showing the effect of overexpression of ZFM1-A on transactivation of reporter gene driven by the Gal4-GQC transactivator. HepG2 cells were cotransfected with activator pSG424-GQC (100 μg), reporter G5E1BCAT (2 μg), control plasmid pGK-β-gal (2 μg), and increasing amounts of pCR3.1-ZFM1-A (0–15 μg). pCL-neo was used as carrier to keep equal amount of total DNA in each transfection mixture. The activation of reporter is expressed as relative CAT activity after transfection efficiency is normalized. The CAT activity of the transfection when zero ZFM1-A is cotransfected was arbitrarily set at 100. Each transfection was repeated at least three times, and the mean and standard error are shown in the figure. B, CAT assays showing the effect of overexpression of ZFM1-A on transactivation of G5E1BCAT reporter only. Cotransfection experiments were performed using DNA mixture similar to that in A, except without the activator pSG424-GQC. The CAT activity of the transfection when no ZFM1-A is cotransfected was arbitrarily set at 100. C, Western blot showing the expression of ZFM1-A in transfected cells. Equivalent amount of normalized cell extracts from transfection samples shown above in which 0–15 μg of pCR3.1-ZFM1-A were cotransfected were loaded on 10% SDS-PAGE. The expression of ZFM1-A was detected using M2 antibody, which recognizes the FLAG-tag on the transfected ZFM1-A.](http://www.jbc.org/)
amount of Gal4-GQC (from 10 ng to 5 \(\mu\)g) are cotransfected with increasing amounts of pCR3.1-ZFM1-A (data not shown). The reporter gene used in this experiment, G5E1BCAT, has a minimal TATA-containing promoter. Without Gal4-GQC-activated transcription, we asked if the minimal activity of the reporter alone could be repressed by overexpression of ZFM1-A. The observation that the expression of the reporter alone cannot be repressed similarly indicates that repression of transcription from this reporter requires the presence of the Gal4-GQC protein (Fig. 3B).

The GQC activation domain represents a class of transcriptional activation domain rich in glycine, glutamine, serine, and threonine (4). EWS, which is frequently involved in translocations in Ewing’s Sarcoma tumors, also contains a potent transcriptional activation domain with similar amino acid composition (26). We observed that ZFM1-A can also interact directly with this domain of EWS. Similarly, overexpression of ZFM1-A can repress transcription of reporter genes driven by Gal4-EWS in transient transfection experiments. We speculate that ZFM1 will act on a variety of cellular activators containing this type of activation domain.

We also tested the ability of ZFM1-A to repress CAT reporter genes controlled by either the thymidine kinase promoter or the SV40 early promoter (TKCAT and SV40CAT, respectively). These promoters utilize distinct activator proteins, and the resulting reporters have different degrees of basal CAT activity. The Gal4-GQC protein can activate transcription from G5TKCAT an additional 12-fold using 10 ng of pSG424-GQC expression vector in transient cotransfection experiments (Fig. 4); however, G5SV40CAT is not responsive to overexpression of Gal4-GQC (data not shown). We therefore tested for the ability of ZFM1-A to repress Gal4-GQC-activated transcription from G5TKCAT. We observed that the expression from the G5TKCAT reporter activated by Gal4-GQC is repressed in a dose-dependent manner similar to that for the G5E1BCAT reporter (Fig. 4). At the highest levels of ZFM1-A transfected,
activation potential of Gal4-(1–147). However, when G5TKCAT is used as reporter, CAT expression is dramatically repressed. At the highest levels of Gal4-ZFM1-A tested, there is a greater than 10-fold reduction in CAT activity. These results clearly demonstrate that ZFM1 can function as a transcriptional repressor when tethered to the TK promoter by the Gal4 DNA binding domain. The marginal repression observed using TK without the Gal4 upstream activating sequence implies that overexpression of ZFM1 in this case has little if any effect via inhibition of splicing of CAT transcripts. We also tested the repression activity of Gal4-ZFM1-A fusion protein in other cell lines. We observed nearly identical titration curves in HeLa, U2OS, and NIH3T3 cells, indicating that the repression of activated transcription is not cell line-specific (data not shown).

As a first step toward understanding the mechanism of active repression mediated by ZFM1, we questioned whether the minimal TATA box promoter on G5E1BCAT is repressed by GAL4-ZFM1-A. Overexpression of Gal4-ZFM1-A fails to repress the low level basal transcription driven by the E1B TATA box (Fig. 5B). This suggests that ZFM1 functions on some step(s) related to activated but not basal transcription.

The Repression Domain on ZFM1 Maps to Its N-terminal Domain—To characterize the regions on ZFM1 required for its repression function, a number of constructs were made in which different regions of ZFM1-A were fused in frame with the Gal4 DNA binding domain. We also tested the splicing...
variant, Gal4-ZFM1-E, for repressor activity. This ZFM1 isoform represses TK promoter activity about 2-fold better than Gal4-ZFM1-A. We then utilized a number of constructs to further localize the amino acid sequences with repressor activity. Gal4-ZFM1-(1–79) and Gal4-ZFM1-(80–528), which span the ABC and D regions of Toda et al. (16), respectively, contain repression activity that is nearly as potent as full-length protein. Additional overlapping deletion constructs within this region confirmed that the repression domain maps to the first 137 amino acids of ZFM1-A just N-terminal to the hnRNP K homology domain. Although there are many splicing variants of ZFM1, the open reading frames of all isotypes but one initiate from the same methionine and contain the entire repression domain (16, 17). Toda et al. (16) reported a novel cDNA encoding what they refer to as the HDEF protein. This subtype contains a novel noncoding H exon spliced to the D exon at nucleotide 619. This protein is predicted to initiate translation at an internal methionine (Met-116) of the ZFM1-A subtype at nucleotide 728. Although we do not have cDNAs encoding this subtype, we would predict that it may not function as a transcriptional repressor. Perhaps regulated expression of this subtype acts as a dominant negative inhibitor of other ZFM1 subtypes or it has a role in splicing. The first 137 amino acids of ZFM1-A, which constitute the repression domain, are rich in charged amino acids (34.3% charged residues) (Fig. 6B). This is a characteristic of several well characterized transcriptional repressors, such as proteins containing Kruppel-associated boxes and E4BP4 (27, 28).

Surprisingly, in addition to a repression domain, ZFM1 has a cryptic transcription activation domain within the region homologous to the Wilms’ tumor suppressor gene. Gal4-ZFM1-(393–503) can activate G5TKCAT at least 10-fold better than Gal4-(1–147) (Fig. 6A). In light of the repressor activity of the full-length protein, the function of this region of the protein if any is not clear. Perhaps different ZFM1 subtypes or specific regulated post-translational modifications can allow this cryptic transcription activation domain to function in vivo.

The Repression Effect of ZFM1 on GQC-mediated Transactivation Depends on the Presence of the GQC Interaction Region on ZFM1—Our results suggest that two separate regions on ZFM1-A are involved in its ability to repress GQC-mediated transactivation: an N-terminal repression domain and a GQC interaction region in the central portion of the protein just beyond the hnRNP K homology domain and zinc knuckle motifs.
C levels of expression of transfected ZFM1-(1–320). Western blots using standard error are shown in the figure. Transfection was repeated at least three times, and the mean and the CAT activity of the reporter alone, which was set at 100. Each (100 ng). The value of expression is shown as relative CAT activity to

interaction with GQC but retains the entire repression domain.

interaction with GQC but retains the entire repression domain

on Gal4-GQC-driven transcription from G5E1BCAT reporter. We did not observe any repression (Fig. 7 A).

FIG. 7. Repression of Gal4-GQC-driven transcription requires the interaction domain of ZFM1-A as well as the repression domain. A, CAT assays showing the effect of overexpression of ZFM1-(1–320) on Gal4-GQC-driven transcription from G5E1BCAT reporter. The indicated amounts of pCR3.1-ZFM1-(1–320) were cotransfected into HepG2 cells with G5E1BCAT reporter (2 μg) and pSG424-GQC (100 ng). The value of expression is shown as relative CAT activity to the CAT activity of the reporter alone, which was set at 100. Each transfection was repeated at least three times, and the mean and standard error are shown in the figure. B, Western blot to show the levels of expression of transfected ZFM1-(1–320). Western blots using M2 antibody were done as described for Fig. 3C.

To assess whether the GQC interaction region is critical to mediate repression, we also tested a construct expressing a variant of ZFM1-A, which lacks the region responsible for interaction with GQC but retains the entire repression domain (ZFM1-(1–320)). As expected, ZFM1-(1–320) failed to interact with the GQC domain in the yeast two-hybrid assay (Table I). When increasing amounts of pCR3.1-ZFM1-(1–320) were cotransfected into HepG2 cells along with pSG424-GQC and G5E1BCAT, we did not observe any repression (Fig. 7A). We do not believe that this is due to poor expression of transfected ZFM1-(1–320) because we detected an overexpressed 40-kDa protein band corresponding to transfected ZFM1-(1–320) on Western blots probed with M2 antibody (Fig. 7B). Moreover, ZFM1-(1–320) fused with the Gal4 DNA binding domain (pSG424-ZFM1-(1–320)) could still repress G5TKCAT expression (Fig. 6A) and ZFM1-(1–320) contained the nuclear localization signal. These results strongly suggest that the repression effect of ZFM1-A on GQC-mediated transactivation depends on the presence of an intact GQC interaction region on ZFM1-A. The association of ZFM1-A and the GQC domain recruits ZFM1-A to the proximity of the promoter where the N-terminal repression domain may function.

**DISCUSSION**

The GQC region (amino acids 181–404) of the transcription factor SSAP of sea urchins functions as an extremely potent activation domain in a variety of mammalian cell lines. In most lines tested, Gal4-GQC activates reporter genes 4–5-fold better than Gal4-VP16 (4). It belongs to a class of transcription activation domain rich in glycine, glutamine, serine, threonine, and basic amino acids. This activation domain most resembles those found in the proteins EWS and TLS, which are involved in translocations in many Ewing's sarcoma and liposarcoma tumors (26, 29). To better understand the mechanisms by which the GQC domain activates transcription to such high levels, we attempted to identify targets that it interacts with in the cell nucleus. We utilized the yeast two-hybrid approach and isolated several human proteins that can specifically interact with this domain. We believe that among these interacting proteins are interesting transcriptional mediators that may positively or negatively regulate the activity of the GQC domain. ZFM1, which was repeatedly picked during our screening of an HL60 cell-derived cDNA library is one such example. ZFM1 protein can interact with the GQC domain in *in vitro* GST pull-down experiment, and it contains several structural motifs that are characteristic of proteins involved in transcription. It contains a nuclear localization signal in its N terminus. It also contains in its N-terminal half an hnRNP K homology domain and a zinc knuckle motif, which are involved in binding to nucleic acids. Although ZFM1 has been shown to bind to RNA nonspecifically (17), it does not exclude the possibility that ZFM1 may also bind to DNA. The hnRNP K protein is an example of a protein with just such a dual function in RNA metabolism and in transcription. At least one splice variant of hnRNP K binds to the single-stranded CT element of the c-Myc promoter and functions both *in vivo* and *in vitro* as a transcriptional activator (30). In its C-terminal half, ZFM1 contains a hydrophobic proline-rich domain, which is characteristic of the activation domain for some transcription factors (31). Within this proline-rich domain, ZFM1 shows significant homology to part of Wilm’s tumor suppressor gene product (WT1) and early growth response 2 proteins (EGR2) (16). Both WT1 and EGR2 are presumed to be transcription factors (32, 33). Similar to transcription activators, some classes of transcription regulators can have a modular structure with a transferable repression domain. The repression domain of ZFM1 belongs to a group of repression domains characterized by charged amino acids. A human leucine zipper family protein E4BP4 and proteins containing Krüppel-associated boxes are included in this group (27, 28).

Also contained within the proline-rich domain of ZFM1 is a PPLP motif recognized by the WW domain of the formbind protein FBP11 and overlapping it a site recognized by the Abl Src homology domain (34). The two motifs are found in all splice variants of ZFM1 including ZFM1-E, which suggests that ZFM1 is regulated by FBP11 and Abl. ZFM1 was selected during expression screening for FBP11-binding proteins and subsequently shown to bind Abl with somewhat lower affinity (34). Among the other ligands binding to the WW domain of FBP11 are the transcriptional regulator ATRBF1, methyl-CpG DNA-binding protein (WPB10), a trithorax-related protein (WPB7), the HLH-like protein NDP1 (WPB8), and SRPK1 (a serine/threonine kinase involved in phosphorylation or SR pre-mRNA splicing factors). All of these are nuclear proteins of varied function and suggest that FBP11 could modulate either the transcriptional repressor or pre-mRNA splicing activity of ZFM1.

It appears that ZFM1 is not a global repressor. Overexpression of ZFM1 has little effect on transcription driven by the TK promoter, which is regulated by CCAAT-binding proteins and SPI (35). It also does not exert any effect on the SV40 early promoter (36) and, by inference, the transcriptional
SSAP Interacts with Transcriptional Repressor ZFM1

activators binding to them. However, when tethered to the TK promoter as a Gal4 fusion protein, ZFM1 can inhibit transcription. ZFM1 does not apparently function through direct interactions with general transcription factors, and, in this regard, it is a unique class of repressor protein. Several well-known active repressors have been shown to function through interaction with basal transcription factors. For example, the repression activity of unliganded thyroid hormone receptor correlates to its binding to TBPF and TFIIB (37, 38). Dr1 functions as repressor by interacting with TBPF and thus excluding TFIIB from entering the preinitiation complex (39, 40). The evidence that ZFM1 functions differently is that it cannot repress a minimal promoter which contains only a TATA box, initiator element, and upstream Gal4 upstream activating sequence (Fig. 5B). If the repression activity of ZFM1 does have promoter selectivity, it is likely therefore that ZFM1 exerts its repression activity not by contacting TBPF, TFIIB, or TFIIF74, the known general transcription factor targets of in vitro interactions with GQC (4), but perhaps by contacting and inhibiting either a TAF1 protein component of TFIID or an as yet unknown coactivator of transcription that coordinates the activity of a family of activator proteins in which GQC is a member. Perhaps some of the other genes isolated in our two-hybrid screen encode this molecule.

Kramer et al. (17) have recently isolated a mammalian splicing factor SF1, which is identical to ZFM1. SF1 is one of several proteins that is required for the formation of a presplicing complex in the early stage of the process in the splicing of nuclear pre-mRNA. Recombinant SF1 protein expressed from baculovirus in insect cells can stimulate pre-spooling complex formation in a dose-dependent manner in vitro experiments. However, we do not think that the repression activity of ZFM1 we observed is due to inhibition of splicing of reporter CAT gene pre-mRNA for the following reasons. First, overexpression of ZFM1-A only represses transcription of reporters driven by Gal4-GQC but does not significantly affect identical CAT genes driven by TK or SV40 promoters. If the repression effect is due to splicing, we should observe the same relative decrease in expression regardless of the promoter driving CAT expression. Second, the potential of ZFM1 to function as a transcriptional repressor depends on its binding to the promoter region. Gal4-ZFM1-A will inhibit CAT expression dramatically when it is cotransfected with a CAT reporter with five Gal4 binding sites upstream in its promoter (G5TKCAT). In contrast, Gal4-ZFM1-A only marginally reduces CAT expression when co-transfected with a CAT reporter with no Gal4-binding site (G0TKCAT). If ZFM1 only functions in splicing, we should observe the same degree of repression regardless of how many Gal4 binding sites are present in the promoter region of the CAT reporter. Therefore, we strongly believe that ZFM1 functions in transcription as a repressor, and we speculate that ZFM1 may have a dual role, functioning in the splicing process as well.

More and more evidence has supported the idea that mRNA splicing and processing occurs cotranscriptionally (41–44). ZFM1 maybe an important factor which links transcription and RNA splicing. Proteins with dual roles in both transcription control and splicing have already been identified. In addition to the previously mentioned hnRNP K protein is WT1, a tumor suppressor gene product involved in Wilms’ tumor (45). WT1 has two isoforms, which differ by the presence of a 3-amino acid KTS insertion between zinc fingers 3 and 4. The KTS form has been shown to bind to DNA and function as a transcription repressor in transient transfection assays. In contrast, the +KTS form has been shown to associate with splicing factors by confocal microscopy and immunoprecipitation experiments. Similarly, it is an intriguing possibility that the different isoforms of ZFM1 may participate in either transcription or splicing. The subcellular localization of individual ZFM1 isoforms under various metabolic conditions should be carefully investigated and compared with other transcription factors and components of the splicing machinery. We detect at least eight different protein species in HepG2 cells using our polyclonal antibody against the N terminus of ZFM1-A (data not shown) and at least seven PCR products are generated from HeLa RNA amplified using primers spanning exons 12 and 14B (17). Most of the splice variants of ZFM1 have distinct C-terminal amino acid sequences. We currently have no functional role(s) for the C-terminal domain of ZFM1. Perhaps it is important for recruiting additional proteins to either transcription or splicing complexes. Our future experiments will be aimed at understanding the mechanism of ZFM1 induced repression of activated transcription and the cellular targets it uses to lower transcription from certain classes of activators.

Acknowledgments—We thank Dr. R. Sterngrinzl for yeast vectors and L. A. Childs for Dr. G. V. Kalpana for the HL60 DNA library, Richard Torres for help with two-hybrid screening, Dr. C. P. Yang for mammalian cell lines and help with tissue culture, Dr. A. J. Berk for CAT reporters, and Dr. Tatsushi Toda for CE9 DNA. We are grateful to members of the Childs laboratory for helpful discussion and comments on the manuscript.

REFERENCES

1. Lai, Z. C., DeAngelo, D. J., DiLiberto, M., and Childs, G. (1989) Mol. Cell. Biol. 9, 2315–2323
2. DeAngelo, D. J., DeFalco, J., and Childs, G. (1993) Mol. Cell. Biol. 13, 3746–3758
3. DeAngelo, D. J., DeFalco, J., Rybacki, L., and Childs, G. (1995) Mol. Cell. Biol. 15, 1254–1264
4. DeFalco, J., and Childs, G. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 5802–5807
5. Choy, B., and Green, M. R. (1993) Nature 366, 531–536
6. Narayan, S., Widen, S. G., Beaud, W. A., and Wilson, S. H. (1994) J. Biol. Chem. 269, 12755–12763
7. Yankulov, R., Blau, J., Purton, T., Roberts, S., and Bentley, D. L. (1994) Cell 77, 749–759
8. Tjian, R., and Maniatis, T. (1994) Cell 77, 5–8
9. Triendenberg, S. J. (1995) Curr. Opin. Genet. Dev. 5, 190–196
10. Hoey, T., Weinzierl, R. O., Gill, G. J., Chen, J. L., Dynlacht, B. D., and Tjian, R. (1993) Cell 72, 247–260
11. Gill, G., Fischat, E., Tseng, Z. H., and Tjian, R. (1994) Proc. Natl. Acad. Sci. U. S. A. 91, 192–196
12. Goodrich, J. A., Hoey, T., Thut, C. J., Admon, A., and Tjian, R. (1993) Cell 73, 519–530
13. Arany, Z., Newsome, D., Oldread, E., Livingston, D. M., and Eckner, R. (1995) Nature 374, 81–84
14. Janne, R., and Hunter, T. (1996) Curr. Biol. 6, 951–954
15. Ogryzko, V. V., Schiltz, R. L., Rassanova, V., Howard, B. H., and Nakatani, Y. (1996) Cell 87, 853–958
16. Toda, T., Iida, A., Miwa, T., Nakamura, Y., and Imai, T. (1994) Hum. Mol. Genet. 3, 465–470
17. Arning, S., Gruter, P., Blone, B., and Kramer, G. (1996) RNA 2, 794–810
18. Luban, J., Bossolt, K. L., Franke, E. K., Kalpana, G. V., and Goff, S. P. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 1615–1619
19. Vejtek, A. B., Hollenberg, S. M., and Cooper, J. A. (1993) Cell 74, 205–214
20. Schiestl, R. H., and Gietz, R. D. (1989) Curr. Biol. 16, 339–346
21. Hill, J., Donald, K. A., Griffiths, D. E., and Donald, K. A. (1991) Proc. Natl. Acad. Sci. U. S. A. 88, 10907–10911
22. Pontelli, A. S., Pardee, T. S., and Struhl, K. (1995) Mol. Cell. Biol. 15, 983–988
23. Struhl, K. (1995) Annu. Rev. Genet. 29, 651–674
24. Crozz, T., Aman, P., Mandahl, N., and Ron, D. (1993) Nature 363, 640–644
25. Cowell, I. G., and Hurst, H. C. (1994) Nature Cell Biol. 22, 59–65
26. Margolín, J. F., Friedman, J. R., Meyer, W. K., Vissing, H., Thiesen, H. J., and Rauscher, F. J. (1994) Proc. Natl. Acad. Sci. U. S. A. 91, 4509–4513
27. Delattre, O., Zucman, J., Plougastel, B., Desmaze, C., Melot, T., Peter, M., Kvar, H., Joubert, I., de Jong, P., Rouleau, G., Aurias, A., and Thomas, G. (1992) Nature 359, 162–165
28. Broder, E., and Levens, D. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 5830–5835
29. Mermod, N., O’Neill, R. A., Kelly, T. J., and Tjian, R. (1989) Cell 58, 741–753
30. Hewitt, S. M., Frazer, G. C., Wu, Y.-J., Rauscher, F. J., III, and Saunders, G. F. (1996) J. Biol. Chem. 271, 5885–5892
31. Joseph, L. J., Le Beau, M. M., Jamieson, G. A., Jr., Acharya, S., Shows, T. B., Rowlcy, J. D., and Suhkatme, V. P. (1998) Proc. Natl. Acad. Sci. U. S. A. 85,
34. Bedford, M. T., Chan, D. C., and Leder, P. (1997) EMBO J. 16, 2376–2383
35. Jones, K. A., Yamamoto, K. R., and Tjian, R. (1985) Cell 42, 539–572
36. Dynan, W. S., and Tjian, R. (1983) Cell 35, 79–87
37. Fondell, J. D., Brunel, F., Hisatake, K., and Roeder, R. G. (1996) Mol. Cell. Biol. 16, 281–287
38. Banahmad, A., Ha, I., Reinberg, D., Tsai, S., Tsai, M. J., and O'Malley, B. W. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 8832–8836
39. Yeung, K. C., Inostroza, J. A., Mermelstein, F. H., Kannabiran, C., and Reinberg, D. (1994) Genes Dev. 8, 2097–2109
40. Inostroza, J. A., Mermelstein, F. H., Ha, I., Lane, W. S., and Reinberg, D. (1992) Cell 70, 477–489
41. Wuarin, J., and Schibler, U. (1994) Mol. Cell. Biol. 14, 7219–7225
42. Bauren, G., and Wieslander, L. (1994) Cell 76, 183–192
43. McCracken, S., Feng, N., Yankulov, R., Ballantyne, S., Pan, G., Greenblatt, J., Patterson, S. D., Wickens, M., and Bentley, D. L. (1997) Nature 385, 357–361
44. Yuryev, A., Patturajan, M., Litingtung, Y., Joshi, R. V., Gentile, C., Gebara, M., and Corden, J. L. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 6975–6980
45. Larsson, S. H., Charleu, J. P., Miyagawa, K., Engelkamp, D., Rassoulzadegan, M., Ross, A., Cuzin, F., van Heyningen, V., and Hastie, N. D. (1995) Cell 81, 391–401
Human ZFM1 Protein Is a Transcriptional Repressor That Interacts with the Transcription Activation Domain of Stage-specific Activator Protein
Di Zhang and Geoffrey Childs

J. Biol. Chem. 1998, 273:6868-6877.
doi: 10.1074/jbc.273.12.6868

Access the most updated version of this article at http://www.jbc.org/content/273/12/6868

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC’s e-mail alerts

This article cites 45 references, 20 of which can be accessed free at http://www.jbc.org/content/273/12/6868.full.html#ref-list-1