GLI Activates Transcription through a Herpes Simplex Viral Protein 16-Like Activation Domain*

(Received for publication, October 3, 1997, and in revised form, December 1, 1997)

Joon Won Yoon‡, Cheng Zheng Liu‡, Jian Tao Yang‡, Rachel Swart‡, Philip Iannaccone‡, and David Walterhouse‡‡‡

From the ‡Department of Pediatrics and the §Division of Hematology/Oncology, Northwestern University Medical School, Developmental Systems Biology, Children’s Memorial Institute for Education and Research, Children’s Memorial Hospital, Chicago, Illinois 60614

Three proteins have been identified in mammals, GLI, GLI2, and GLI3, which share a highly conserved zinc finger domain with Drosophila Cubitus interruptus and are believed to function as transcription factors in the vertebrate Sonic hedgehog-Patched signaling pathway. To understand the role GLI plays in the Sonic hedgehog-Patched pathway and mechanisms of GLI-induced transcriptional regulation, we have characterized its transcriptional regulatory properties and contributions of specific domains to transcriptional regulation. We have demonstrated that GLI activates expression of reporter constructs in HeLa cells in a concentration-dependent manner through the GLI consensus binding motif and that a GAL4 binding domain-GLI fusion protein activates reporter expression through the GAL4 DNA binding site. GLI-induced transcriptional activation requires the carboxyl-terminal amino acids 1020–1091, which includes an 18-amino acid region highly similar to the α-helical herpes simplex viral protein 16 activation domain, including the consensus recognition element for the human TFIIID TATA box-binding protein-associated factor TAFII31 and conservation of all three amino acid residues believed to contact directly chemically complementary residues in TAFII31. The presence of this region in the GLI activation domain provides a mechanism for GLI-induced transcriptional regulation.

Cubitus interruptus (Ci) functions as a transcription activator in the Drosophila hedgehog-patched signaling pathway to mediate hedgehog signaling specifying segmentation of Drosophila embryos and patterning of imaginal-disk outgrowth (1, 2). Drosophila cAMP response element-binding protein (CREB)-binding protein functions as a coactivator of Ci (2). Three homologous proteins have been identified in mammals, including GLI, GLI2, and GLI3, which share a highly conserved C3-H3 zinc finger domain with Ci and are believed to function as transcription factors in the vertebrate Sonic hedgehog-Patched signaling pathway, each potentially carrying out a specific function during development (3–6). Specific as well as redundant functions have been demonstrated for GLI2 and GLI3 during mammalian skeletal development, whereas GLI but not GLI3 has been shown to activate the HNF-3β enhancer in tissue culture (7, 8). Sonic hedgehog-Patched signaling specifies polarity of limb outgrowth, polarity of the developing neural tube, and somite patterning and mutations in genes in this pathway may relate to understanding causes of human basal cell carcinoma, holoprosencephaly, Pallister-Hall syndrome, Greig syndrome, malignant gliomas, and sarcomas (9–16). Because the GLI family zinc finger domain has been shown to mediate DNA binding to the 9-bp consensus motif GACCACCCA, it is believed that GLI, GLI2, and GLI3 bind identical or similar DNA sequences, and understanding the specificity of function and their specific roles in the Sonic hedgehog-Patched signaling pathway will require an understanding of their transcriptional regulatory properties and the mechanisms responsible for their transcriptional regulation (17–19).

To begin to understand the function of GLI in the Sonic hedgehog-Patched signaling pathway and its mechanism of action, we have studied its transcriptional regulatory properties in tissue culture and identified contributions of specific domains to transcriptional regulation. We find that GLI activates reporter gene transcription through the SV40 promoter and the E1b promoter in a concentration-dependent manner. A transcription activation domain is identified at the carboxyl terminus of the protein which includes an 18-amino acid acidic α-helix highly similar to the herpes simplex viral protein 16 (VP16) transcription activation domain, which targets TAFII31 and Drosophila TAFII31 (20–22). A similar domain is present in the other GLI family proteins. The presence of a VP16-like activation domain in GLI provides a mechanism for transcriptional regulation.

**EXPERIMENTAL PROCEDURES**

GLI Expression Vectors and Reporter Constructs for Cotransfection Experiments—pCMV-GLI was prepared by inserting the full-length human GLI cDNA into the HindIII/XbaI site of the pcDNA3 plasmid (Invitrogen). pCMV-GLI–TAD was prepared by deleting a 1,409-bp AccI fragment from the 3′-end of GLI. pCMV-GLI (−)AT was prepared by removing a 661-bp HindIII/XbaI fragment from the 5′-end of GLI. pM-GLI was prepared by inserting a HindIII/XbaI fragment of the human GLI cDNA into the Smal/XbaI site of the pM plasmid (CLONTECH). pM-GLI151–1106 was prepared by inserting an SmaI/XbaI fragment of human GLI cDNA into the SmaI/XbaI site of the pM plasmid (CLONTECH). pM-GLI1–210 was constructed by inserting polymerase chain reaction-amplified human GLI cDNA corresponding to amino acids 1–210 into the EcoRI/XbaI site of the pM plasmid. pM-GLI76–686 was constructed by removing the AccI fragment from the pM-GLI vector.

To begin to understand the function of GLI in the Sonic hedgehog-Patched signaling pathway and its mechanism of action, we have studied its transcriptional regulatory properties in tissue culture and identified contributions of specific domains to transcriptional regulation. We find that GLI activates reporter gene transcription through the SV40 promoter and the E1b promoter in a concentration-dependent manner. A transcription activation domain is identified at the carboxyl terminus of the protein which includes an 18-amino acid acidic α-helix highly similar to the herpes simplex viral protein 16 (VP16) transcription activation domain, which targets TAFII31 and Drosophila TAFII31 (20–22). A similar domain is present in the other GLI family proteins. The presence of a VP16-like activation domain in GLI provides a mechanism for transcriptional regulation.

© 1998 by The American Society for Biochemistry and Molecular Biology, Inc.
GLI activates transcription through the SV40 promoter after binding to the GACCACCCA consensus binding motif. Panel A, RNase protection assay of transfected cells. The protected GLI mRNA is indicated by the arrow. Lane 1, GLI probe only; lane 2, D259 MG cell RNA (positive control); lane 3, yeast RNA (negative control); lane 4, HeLa cell RNA; lane 5, RNA from HeLa cells transfected with pcMV-GLI; lane 6, RNA from HeLa cells transfected with pcMV-GLI−/−AT; lane 7, RNA from HeLa cells transfected with pcMV-GLI−/−AT and pcG5CAT reporter construct only. The control reporter construct lacking the SV40 enhancer, containing the pCAT-G5 control plasmid (Promega) (17). pCAT-GLI44−/−E was prepared by removing a HindII fragment, which contains the SV40 enhancer, from pCAT-GLI44. The control reporter construct lacking the SV40 enhancer, pcAT control−/−E, was prepared similarly from pcAT control (Promega). The pG5CAT reporter was obtained from CLONTECH.

Preparation of Nested Deletions and DNA Sequencing—DNA subfragments of pM-GLI15–1106, representing serial deletions from the carboxyl terminus, were prepared using exonuclease III/S1 nuclease digestion (23). The DNA sequence of these DNA clones was determined using the dyeoxy chain termination procedure (24).

RNA Protection Assays—Total RNA was isolated using TRI reagent (Molecular Research Center, Inc.). 10 μg of total RNA was incubated with 1 × 106 cpm of the antisense probe. After overnight hybridization at 45 °C, unhybridized RNA was digested with an RNase A/T1 mixture. Protected RNA fragments were analyzed on a 5% denaturing polyacrylamide gel and were visualized by autoradiography.

In Vitro Translation—In vitro translation of pcMV-GLI was completed using nuclease-treated reticulocyte lysate (Ambion Inc.). In vitro translation of pCMV-GLI−/−TAD and pCMV-GLI−/−AT were completed using wheat germ extract (Promega). The newly synthesized proteins were separated on 10% SDS-polyacrylamide.

Electrophoretic Mobility Shift Assays—Nuclear extracts were prepared from 2–4 × 106 cells, and the protein concentration of each extract was determined (25, 26). 5 μg of each nuclear extract was mixed with 5 μl of 2 × binding buffer (20% glycerol (v/v), 20 mM MgCl2, 200 μM EDTA, 1 mM dithiothreitol, 100 mM KCl, 40 mM HEPES, pH 7.9, 100 μM/ml poly(dI-dC), and 20 μM ZnSO4, H2O, and 0 or 1 μl (1.75 pmol) of unlabeled competitor oligonucleotide (27). The mixture was incubated on ice for 5–10 min. 1 μl (8.75 fmol) of the double-stranded 32P-labeled probe (sense strand: 5′-GATCTAAGAGCTCCCGAAGACCACCCACAATTGAGTTGTATGT-3′) was added, and the mixture was incubated on ice for an additional 30 min. The samples were analyzed by electrophoresis, and retarded bands were visualized by autoradiography.

Cell Culture and Transfections—A total of 4 μg of DNA was transfected in each cotransfection experiment using the GLI DNA binding domain-containing reporters, including 1 μg of reporter plasmid, 1 μg of pSV-β-galactosidase (Promega), or pSV-luciferase (Promega) transfection efficiency control plasmid, 0–1 μg of effector plasmid, and pBluescript carrier DNA (Stratagene) in an amount to make up the difference. For the mammalian one-hybrid experiments using the pGLI-expressing effector constructs and the pG5CAT reporter construct, a total of 4 μg of DNA was transfected in each experiment, including 0.4 μg of reporter plasmid, 0.4 μg of pSV-β-galactosidase (Promega) transfection efficiency control plasmid, 0–3.125 μg of pGLI-expressing effector plasmid, and pBluescript carrier DNA (Stratagene) in an amount to make up the difference. In each case the DNA was mixed with 9 μl of LipofectAMINE reagent in 0.3 ml of Opti-MEM (Life Technologies, Inc.), incubated for 45 min at room temperature, and then added to individual 60-mm plates containing 2 ml of Opti-MEM and 4 × 106 cells. Cells were incubated for 5 h in a CO2 incubator before the Opti-MEM was replaced with HeLa culture media. Cells were incubated for 48 h, and lysates were prepared for CAT assays and β-galactosidase assays (Promega).

CAT Assays—CAT assays were performed as described by the manufacturer (Promega) with minor modifications. CAT activity was quantitated by scintillation counter and was normalized by measuring β-galactosidase activity spectrophotometrically (Promega) or by measuring luciferase activity (Promega) with a luminometer.

RESULTS

GLI Functions as a Transcription Activator—HeLa cells were cotransfected with a GLI-expressing effector construct (pcMV-GLI) and a CAT reporter (pCAT-GLI44−/−E), under the control of the SV40 promoter, containing two tandem copies of a previously identified GLI DNA binding motif, GATCTAAGAGCTCCCGAAGACCACCCACAATTGAGTTGTATGT (17). This binding motif, which includes a 9-bp consensus sequence (underlined) identified in all three of the previously identified GLI DNA binding motifs, was selected because it may contain critical binding elements absent from the consensus 9-bp sequence. GLI mRNA was demonstrated in the transfected cells and in the positive control GLI-amplified D259 MG glioblastoma multiforme cell line using RNase protection analysis (Fig. 1A, lanes 1–5). No evidence for GLI mRNA expression was found in HeLa cells or in HeLa cells transfected with the control plasmid pcDNA3 (data not shown). Because GLI3 has been shown to bind to the identical site as GLI and may compete with GLI for binding sites, reverse transcription-polymerase chain reaction was used to look for GLI3 expression in HeLa cells. No GLI3 expression was found in HeLa cells (data not shown). A predicted 150-kDa GLI protein was produced off the pcMV-GLI effector construct and the pCAT-GLI44−/−E reporter construct (black bars); the pCMV-GLI effector construct and the pCAT control−/−E, which lacks the GLI DNA binding sequence (white bars); and the pcDNA3 effector construct, which does not express GLI, and the pCAT-GLI44−/−E reporter (gray bars). The means ± S.D. were calculated from three independent experiments. Stars indicate statistical significance (p < 0.05) using Student’s paired t test comparing CAT activity at each amount of effector DNA with CAT activity using no effector DNA.
GLI activates transcription through the E1b promoter
following binding to the GAL4 consensus binding motif. CAT assays (top) were obtained by transfecting HeLa cells with pM DNA (control) and the pG5CAT reporter or pM-GLI and the pG5CAT reporter. The amount (ng) of pM DNA or pM-GLI used is indicated. Bar graphs (bottom) with means ± S.D. are based on scintillation counting of two independent CAT assay experiments. Stars indicate statistical significance (p ≤ 0.05).

GLI78–686 showed no increase in CAT activity (Fig. 4A). Together these data indicate that a transcription activation domain lies in the carboxy-terminal region of GLI between amino acids 686 and 1106.

To define the transcription activation domain more precisely we then made nested deletions from the carboxyl terminus of GLI using exonuclease III/S1 nuclease digestion (Fig. 4B). Nested deletion products were sequenced, and their ability to drive CAT activity was measured. The smallest segment identified which was required for transcriptional activation included amino acids 1020–1091.

The GLI Transcription Activation Domain Is Similar to the VP16 Activation Domain—The region from 1020 to 1091 includes a negatively charged α-helical region demonstrating 50% similarity with the VP16 α-helical transcription activation domain including conservation of the FXXΦΦ motif (F = phenylalanine; X = any residue; Φ = any hydrophobic residue) general recognition element of acidic activation domains for TAFp31 and conservation of the three residues (Asp472, Phe479, and Leu483) in VP16, and Asp509, Phe522, and Leu526 in GLI that are believed to make direct contacts with TAFp31 (22) (Fig. 5). The region...
surrounding the GLI α-helix is proline-rich. Human GLI3 and Drosophila Ci demonstrate conservation of the FXXFF consensus element and two of the three residues that are believed to make direct contact with TAFII31, whereas mouse gli2 demonstrates conservation of the FXXFF consensus element and one of the three residues believed to make direct contact with TAFII31.

**DISCUSSION**

We have demonstrated that the oncodevelopmental gene GLI functions as a transcriptional activator of the heterologous promoters SV40 and E1b in HeLa cells through the GLI consensus binding motif and the GAL4 consensus binding motif respectively, in a concentration-dependent manner. As increasing amounts of GLI were transfected into HeLa cells greater CAT activity was seen. No evidence for transcription repression was found. In-frame deletion of amino acids 686–1106 from the carboxyl terminus of the GLI protein resulted in reduced transcriptional activation by GLI with retention of DNA binding. This same region drove CAT activity using a GAL4 fusion protein approach. The smallest segment identified by our studies, which was necessary for GLI-induced transcriptional activation, included amino acids 1020–1091. Indeed, Ruppert et al. (28) had suggested previously that the carboxyl-terminal acidic α-helical region of GLI including amino acids 1037–1055 may represent a transcription activation domain based on the fact...
that 6 of 19 amino acids in this region are aspartate or glutamate. Acidic activation domains of VP16, p53, and p65 have been shown to target TAFII31, and each of these acidic activation domains includes an FXXF motif felt to represent the TAFII31 consensus recognition element (22). The FXXF motif is conserved in the GLI transcription activation domain. The VP16 activation domain undergoes an induced transition from a random coil to an $\alpha$-helix upon binding to its target protein TAFII31, and three residues believed to make direct contacts with TAFII31 have been identified (22). The three residues (Asp1040, Phe1048, and Leu1052) are conserved in the GLI transcription activation domain. These structural motifs identified in a region that we demonstrate to function as the GLI transcription activation domain suggest that the mechanism for

**FIG. 4.** Transcriptional activation by specific regions of GLI. Panel A, CAT assay results (top) obtained by transfecting HeLa cells with pM-GLI1–210, pM-GLI78–686, and pM-GLI515–1106 and the pG5CAT reporter. The amount (ng) of pM-GLI1–210, pM-GLI78–686, and pM-GLI515–1106 used is indicated. Bar graphs (bottom) with means ± S.D. are based on scintillation counting of two independent CAT assay experiments. Stars indicate statistical significance ($p \leq 0.05$). Panel B, CAT assay results obtained by transfecting HeLa cells with 625 ng of pM, pM-GLI515–1106 (1544–3540), pM-GLI515–1106 (1544–3318), pM-GLI515–1091, pM-GLI515–1020, and the pG5CAT reporter. Numbers in parentheses indicate nucleotides. Fragments smaller than pM-GLI515–1020 demonstrating low CAT activity were not sequenced.

**FIG. 5.** The GLI transcription activation domain contains a VP16-like activating region. Panel A, comparison of the amino acid sequence of the human GLI transcription activation domain with mouse gli, human GLI3, mouse gli2, Drosophila Ci, and the VP16 TAFII31 binding domain/transcription activating region (shown in yellow). Human GLI demonstrates 50% similarity to the VP16 domain and includes conservation of the three residues (*) that contact TAFII31 directly (shown in green). Two of these residues are conserved in human GLI3 and Drosophila Ci, and one of these residues is conserved in mouse gli2. The FXXF motif is conserved in all of the proteins (GLI amino acids 1045–1052). In addition, GLI residues Asp1037 and Gln1045 are conserved in all of the mammalian Gli family proteins shown. Panel B, left, helical wheel presentations of the $\alpha$-helical TAF binding domains of human GLI and mouse gli (blue), compared with human GLI3 (green), and mouse gli2 (red) demonstrate identical placement of negatively charged residues. Panel B, right, Drosophila Ci helical wheel presentation of the $\alpha$-helix demonstrates a shift in one charged group.
GLI-induced transcriptional activation involves TAF$_p$31 binding. Substitution of alanine for two of the residues believed to make direct contacts with TAF$_p$31 in the VP16 activation domain (Phe$^{379}$ and Leu$^{483}$) has been shown to reduce TAF binding affinity and transcriptional activation (22). Substitutions in this region in the different Gli family members may affect their TAF binding and transcriptional regulatory properties. Conservation of Asp$^{1037}$ and Gln$^{1035}$ in each of the mammalian Gli family members in this region may be of significance for TAF binding.

It is of interest that the VP16-like domain lies within a proline-rich region at the carboxyl end of the GLI protein. 20% of the amino acid residues in this region (amino acids 752–1081) are prolines. A number of transcription factors contain proline-rich domains that appear to function as transcription activation domains, and the proline-rich region could contribute to transcriptional regulation by GLI (29–31). We were able to eliminate the majority of the proline-rich domain, however, and maintain the transcriptional activation properties of GLI, suggesting that the VP16-like acidic a-helical domain rather than the proline-rich domain represents the critical transcriptional activation domain of GLI.

Transcriptional activation was retained following in-frame deletion of amino acids 2–220 from the amino terminus of the GLI protein. At lower effector concentrations pCMV-GLI$^{-}$IAT appeared to even enhance transcriptional activation relative to pCMV-GLI although overall achieved a comparable level of transcription activation compared with full-length pCMV-GLI. Interestingly, amino acids 68–135 have been identified previously as a region of similarity between Gli family proteins showing 66% amino acid identity between GLI (amino acids 68–135) and GLI3 (amino acids 279–348) (28). Included within this region of GLI (amino acids 67–82) is a conserved domain that has been shown to be functionally significant in Caenorhabditis elegans TRA-1 and has been called the GF or “gain of function” region. Mutations in this region abolish negative regulation of TRA-1 activity (32). Based on the TRA-1 data, conservation of the region in GLI and GLI3, and our current functional experiments, it is possible that the “GF region” at the amino-terminal end of GLI and its immediate vicinity represents an inhibitory domain in GLI as well. Also unique to GLI$^{-}$IAT was the fact that it consistently demonstrated a single shifted complex on gel mobility shift assay compared with wild type GLI, which consistently demonstrated two shifted bands. Two shifted protein complexes may result from partial proteolysis, differing states of protein phosphorylation or glycosylation, or possibly from protein interactions. Removal of the AT region may affect any of these processes.

It has been suggested that mammalian Gli family members function in the Sonic hedgehog-Patched pathway similarly to Ci in the Drosophila hedgehog-patched signaling pathway (15, 33). Decapentaplegic (dpp) represents one of the downstream targets of Ci in this pathway (33). It is believed that low concentrations of Ci repress dpp expression and high concentrations of Ci induce expression of dpp with Ci acting as a repressor or activator of transcription depending on its concentration (33). Recently it has been demonstrated that the 155-kDa Ci protein may be cleaved, resulting in a 75-kDa form that includes the amino terminus through the zinc finger domain but lacks the entire carboxyl end, which includes the transcription activation domain (34). This 75-kDa form of Ci then functions as a transcriptional repressor. Using our in vitro assay conditions we can provide no evidence that GLI2 as a transcription repressor over a wide range of concentrations. It is possible that the three mammalian Gli genes, GLI, GLI2, and GLI3 function cooperatively to carry out comparable functions to ci in the mammalian Sonic hedgehog-Patched signaling pathway. Indeed, it has been shown recently that GLI activates the HNF-3b enhancer whereas GLI3 represses the HNF-3b enhancer (8). Differences in TAF binding or binding affinity could potentially contribute to these different properties. Alternatively it is possible that the three mammalian genes have evolved to add increased layers of complexity to transcriptional regulation in the mammalian Sonic hedgehog-Patched signaling pathway.

Transcription factors are generally considered to be modular, made up of several functional domains. The GLI protein represents an interesting example with each functional module conserved through different evolutionary lines. The GLI protein contains a conserved zinc finger domain found in Drosophila ci and C. elegans tra-1, a conserved 3’-untranslated region translation control element conserved in C. elegans tra-2, and now a conserved transcription activation domain found in herpes simplex virus VP16 (35).

Acknowledgments—We thank K. Kinzler for supplying the full-length human GLI cDNA (pLTR-GLI) and the polyclonal antibody GLI1.
GLI Activates Transcription through a Herpes Simplex Viral Protein 16-Like Activation Domain
Joon Won Yoon, Cheng Zheng Liu, Jian Tao Yang, Rachel Swart, Philip Iannaccone and David Walterhouse

J. Biol. Chem. 1998, 273:3496-3501.
doi: 10.1074/jbc.273.6.3496

Access the most updated version of this article at http://www.jbc.org/content/273/6/3496

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 34 references, 17 of which can be accessed free at http://www.jbc.org/content/273/6/3496.full.html#ref-list-1