The Hematopoietic Transcription Factor SCL Binds the p44 Subunit of TFIIH

(Received for publication, August 12, 1998, and in revised form, October 28, 1998)

Xian-Feng Zhao‡ and Peter D. Aplan§

From the ‡Departments of Pediatrics and Cancer Genetics, Roswell Park Cancer Institute, Buffalo, New York 14263 and the §Department of Pediatrics, State University of New York, Buffalo, New York 14222

SCL is a basic domain helix-loop-helix (bHLH) oncoprotein that is involved in T-cell acute lymphoblastic leukemia as well as in normal hematopoiesis. Although it is believed that SCL functions as a tissue-specific transcription factor, no molecular mechanism has thus far been identified for this putative function. In this report, we show that SCL interacts with p44, a subunit of the basal transcription factor TFIIH. The minimal region of SCL that interacts with p44 was mapped to a 101-amino acid sequence that includes, but is not limited to, the bHLH region; the SCL-binding site of p44 is located in the carboxyl-terminal half of p44. This interaction was confirmed by glutathione S-transferase fusion protein pull-down assays and a co-immunoprecipitation assay. As analyzed with a yeast two-hybrid system, p44 interacts specifically with SCL, but not with the other class A or B bHLH proteins tested. E2A did not compete with p44 for SCL binding, as demonstrated by an in vitro binding assay. These findings document a previously unsuspected interaction between SCL and a subunit of the basal transcription factor TFIIH, suggesting a potential means by which SCL might modulate transcription.

Since its isolation from the multipotential DU528 stem cell leukemia cell line, which carries a t(1;14)(p32;q11) chromosome translocation (1–3), the SCL (TCL3 or tal-1) gene has been studied in the context of both normal and abnormal hematopoiesis. The observation that the SCL locus is frequently disrupted in T-cell acute lymphoblastic leukemia cells (4–6), resulting in aberrant SCL expression, has led to speculation that the gene product of SCL is an oncprotein. Indeed, recent transgenic mouse models have confirmed that unscheduled SCL expression leads to aggressive T-cell malignancies (7, 8). Moreover, despite its initial identification in leukemic cells, targeted disruption of the SCL locus has demonstrated that SCL expression is absolutely required for normal hematopoietic development (9–12).

SCL belongs to the basic domain helix-loop-helix (bHLH) family of proteins (13). Three principal classes of bHLH proteins have been identified (14). Class A bHLH proteins include the E proteins E2A/ITF1 (immunoglobulin transcription factor 1), E2-2/ITF2, and HEB (an E protein related to E2A/ITF1 and E2-2/ITF2); these proteins are expressed ubiquitously. In contrast, the expression of class B bHLH proteins, such as MyoD and SCL, is restricted to specific organs or tissues. The myogenic (MyoD, myogenin, myf5) bHLH proteins have been studied extensively (15) and serve as a useful paradigm for the actions of tissue-restricted bHLH proteins. MyoD forms a heterodimer with the ubiquitously expressed E2A proteins (16), binds specific DNA sequence at the regulatory regions of genes coding for muscle-specific proteins (such as muscle creatine kinase), and activates transcription of these genes (17). Although unproven, it is thought that SCL activates transcription of genes required for normal hematopoietic development in an analogous fashion (18).

Several forms of the SCL protein have been found in mammalian cells (19); the full-length form of SCL has a molecular mass variously reported to be between 42 and 49 kDa, whereas an amino-terminally truncated form of 22–26 kDa is produced by translation of alternatively spliced transcripts (20) in both normal and leukemic cells. Although a transcription activation domain has been mapped to its amino terminus (21), unscheduled synthesis of an SCL protein that lacks the transactivation domain leads to T-cell leukemia in transgenic mice, indicating that this transcription domain is not required for leukemogenesis (8). Similar to many other bHLH proteins, SCL has been shown to bind several E proteins, including E2-2/ITF2 (22), E12, and E47 (two alternatively spliced forms of E2A/ITF1) (23). When bound to E12, the SCL/E12 complex preferentially binds to a CAGATG nucleotide sequence in vitro (24). Although SCL is suspected to function as a sequence-specific transcription factor, no target genes for SCL have thus far been convincingly identified, and any mechanism by which SCL may control transcription remains unknown.

TFIIH is a multicomponent basal transcription factor complex that is also known to function in certain DNA repair pathways (25). Nine subunits have been identified within the TFIIH holoenzyme complex; various enzymatic activities, including DNA repair, helicase (26), and cyclin-dependent kinase (27) activities, have been identified. The p62, p52, p44, and p34 subunits are thought to constitute the “core” of the TFIIH transcription machinery (28). Although the p44 and p34 subunits have no defined enzymatic activity, their zinc finger structures suggest that they may be DNA-binding proteins (29) that might mediate interactions with soluble transcription factors.

Here we report the identification of an unexpected interaction between SCL and the p44 subunit of TFIIH. This observation provides a link between SCL and the basal transcription machinery, suggesting that SCL may exert its suspected transcription regulatory effects through an interaction with TFIIH.
SCL Binds TFIIH

Yeast Two-hybrid Screen—An SCL cDNA derived from clone 67 (20) was subcloned into the pGBT9 two-hybrid vector (CLONTECH) with the SCL coding sequence downstream of the GAL4 DNA-binding domain. This vector, named pGBT8SCL, was cotransformed into Saccharomyces cerevisiae YRG-2 yeast cells (the relevant genotype of the YRG-2 strain is *Mats ura3-52 his3-200 leu2-3,112 gal4-101 trp1-101 ura3-52 ade2-101 lys2-801 trp1-901 his3-200 leu2-3,112 gal4-101 LYS2::UAS2-TATATAGAGATC-3' and LYS2::UAS2-TATATAGAGATC-3' together with cDNA library plasmid DNA using Stratagene reagents and protocols. The cDNA libraries were constructed in either the pGAD10 (CLONTECH) or the pAD-GAL4 (Stratagene) vector, which contained the GAL4 activation domain fused to random library cDNAs. Since the pGBT9 vector encodes TRP1 and the pGAD10 or pAD-GAL4 vector encodes LEU2, transcription/translation reactions were performed using the *in vitro* protein. The cDNA derived from clone 67 (20) was subcloned into the pRC/CMV vector (Stratagene) and used as the DNA template for the production of SCL protein. The E2-5 DNA (30) was a gift from Dr. Adam Goldfarb. The *in vitro* transcription/translation reactions were performed using the TNT® coupled Reticultocyte lysate systems with T7 RNA polymerase (Promega). The ^35^-labeled proteins were analyzed by 10% SDS-PAGE. GST Fusion Protein Interaction Assay—GST-fusion proteins were incubated at room temperature for 30 min with GST-Sepharose (Amersham Pharmacia Biotech) and detected with anti-SCL monoclonal antibody BTL73 (a gift from Dr. Karen Pulford) (33), horseradish peroxidase-conjugated anti-mouse IgG, and Western blot Chemiluminescence Reagent Plus (NEN Life Science Products). SCL and p44 Deletion Mutants—Most of the SCL deletion mutants were derived from pGBT8SCL using exonuclease III to create nested deletion mutants (Erase-A-Base, Promega). Briefly, 5 μg of pGBT8SCL was digested with *PstI* and *BamHI*, and the carboxyl-terminal deletion mutants were generated with exonuclease III and S1 nuclease. For the amino-terminal deletion mutants, a derivative clone of pGBT8SCL, which lacks the EcoRI site located 3' to the SCL cDNA, was first digested with EcoRI, and the 5' overhangs were filled in with Klenow fragment of DNA polymerase I and α-thiol deoxyribonucleotides before digestion with *NotI*. The deletion mutants were sequentially generated with exonuclease III and S1 nuclease. Mutant N100 was produced by utilizing the *ApaI* restriction site within the SCL cDNA, whereas mutants N174, N186, C262, and C275 were PCR-amplified DNA fragments which were recloned into the EcoRI site of pGBT9. The amino-terminal p44 deletion mutant was constructed by subcloning the carboxyl-terminal portion of the p44 cDNA derived from the human thymus cDNA library (see “Results”) into pGBT9, whereas the carboxyl-terminal deletion mutant was generated by *PstI* digestion of pGBT9p44 to delete the 3'-portion of the full-length p44 cDNA, followed by religation of the plasmid. Construction of Additional GAL4-BHLH Expression Vectors—Partial cDNA sequences of HEB (amino acids 277–882; GenBank accession number M80627), E2A/ITF1 (amino acids 349–582; GenBank® accession number X52078), and E21-2TTF2 (amino acids 210–622; GenBank® accession number X52079) isolated from the full-length p44 cDNA were PCR-amplified using primers 5'GTGGAATTCACCATGACTGAGAAG-7.5′ and 5'ATATCCCTCTTAATGAATGA-3′. The cDNA was amplified using the primers described above and subcloned into pGBT9 and pGAD424.

RESULTS

Yeast Two-hybrid Interaction Assay—Log-phase YRG-2 yeast cells were made competent for transformation by suspension in TE buffer (10 m M Tris–HCl and 1 m M EDTA, pH 7.5) and 100 m M lithium acetate, pH 7.5. One μg of plasmid DNA was added to 120 μl of yeast competent cells before adding 600 μl of TE buffer, 100 m M lithium acetate, and 40% polyethylene glycol 4000. The mixture was incubated at 30 °C for 90 min. The cells were pelleted, resuspended in 50 μl of water, and incubated for 15 min at 42 °C. The cells were plated in duplicate on –Leu/Trap and –Leu/TrapHis SD plates. The plates were incubated for 5 days at 30 °C, and growth was monitored daily. The colonies grown on the –Leu/Trap plate were transferred to a Whatman filter, and the cells were disrupted in liquid nitrogen; β-galactosidase activity was detected with Z buffer (60 m M NaHPO4·H2O, 40 m M Na2HPO4, 10 m M KCl, and 1 m M MgSO4·7H2O) containing 0.27% β-mercaptoethanol and 5-bromo-4-chloro-3-indolyl-β-D-galactoside. Quantitative Assay of β-Galactosidase—Five replicate colonies were picked from –Leu/Trap SD plates containing cotransformants and incubated in –Leu/Trap SD medium, which was incubated at 30 °C with shaking until the culture reached mid-log phase. The cells were pelleted, resuspended in Z buffer, and disrupted with liquid nitrogen. One-hundred-μl samples were mixed with 700 μl of Z buffer containing 0.27% β-mercaptoethanol before 160 μl of 4 m M o-nitrophenyl β-D-galactopyranoside in Z buffer was added. The reaction was incubated at 30 °C until a yellow color developed. Four-hundred μl of 1 m Na2CO3 was added to stop the reaction, and the incubation time was recorded. The absorbance of the supernatant was measured at 420 nm and used to calculate the β-galactosidase units.
GST fusion proteins were again precipitated by centrifugation activity in YRG-2 yeast cells without an additional fusion protein. However, since we found that the GAL4 DNA-hybrid system. This fusion protein was also found to generate fusion proteins with both the full-length and truncated forms of p44 fused to GST for use in an two-hybrid system as described above. To determine if these deletion mutants could still bind p44, expression of the HIS3 and lacZ reporter genes was monitored by growth on histidine-deficient medium and production of the β-galactosidase, respectively (Fig. 4). No obvious change in the SCL/p44 interaction was detected when the N-terminal 100 amino acids were deleted (Fig. 4). However, both the N156 and N174 (which closely approximates the naturally occurring 22–26-kDa isoform of SCL) mutants produced decreased level of the β-galactosidase activity and less robust growth on histidine-deficient medium, suggesting a decreased ability to interact with p44. The SCL/p44 interaction was completely absent in the N186 mutant, which retained the entire bHLH domain, demonstrating that the bHLH domain was not sufficient to mediate the interaction. The region between amino acids 100 and 186 was necessary for full binding activity, whereas the region between amino acids 174 and 186 highlighted residues that were absolutely required for the SCL/p44 interaction.

A dramatic decrease in the SCL/p44 interaction was noted between the C-terminal mutants C262 and C275, demonstrating a requirement for this region. Interestingly, the principal difference between mutants C275 and C262 is a polyclongeny tract consisting of 10 consecutive glycine residues, the loss of which essentially eliminated SCL binding activity, suggesting that this polyclongeny tract is necessary for the SCL/p44 interaction. Therefore, the minimal region of SCL that could bind p44 was between amino acids 174 and 275, a region including, but not limited to, the bHLH region. Of note, the interaction between p44 and full-length SCL produced twice as much β-galactosidase as did the SCL/TP44 interaction, indicating that the interaction between SCL and p44 is of similar magnitude to that seen with a well known SCL-binding protein; the SCL/ITF2 interaction showed the strongest binding for SCL among the E proteins (data not shown).

To determine which portion of the p44 protein binds SCL, p44 mutants were used in a yeast two-hybrid assay (Fig. 5).
Although p44 was originally isolated as an N-terminally truncated form, full-length p44 also bound SCL, and we wondered whether the amino terminus of p44 might also contain a region that could bind SCL. A PstI site was utilized to delete the C-terminal region of p44, and the N-terminus-only mutant (Dp44) was tested in the yeast two-hybrid system. The lack of interaction with this mutant demonstrates that the SCL-binding region of p44 is located at the carboxyl terminus (Fig. 5).

Specificity of the SCL/p44 Interaction among bHLH Proteins—Since the p44-binding site of SCL encompasses its bHLH region, it is easily conceivable that the bHLH and flanking regions of other bHLH proteins could also bind p44. When several bHLH proteins were tested in the yeast two-hybrid system, only SCL was shown to bind p44 (Table I). In addition, we were able to use the yeast two-hybrid system to demonstrate previously undescribed interactions between SCL and NHLH1 (neural HLH protein 1, previously named NSCL) (34) as well as an SCL homodimeric interaction (Table I).

The SCL/p44 Interaction Is Not Disrupted by E2A—Since SCL and E2A form stable heterodimers (23, 40) and that interaction is mediated through the bHLH domain, we wondered whether these E proteins might interfere with SCL/p44 binding. Equal amounts of in vitro translated SCL and E2A proteins were prebound by a 15-min incubation at 37 °C. The SCL-E2A complex was then incubated with the GST-p44 fusion protein and precipitated with glutathione-Sepharose beads as described above. The coprecipitated material was washed extensively in the presence of 0.1% Nonidet P-40. As shown in Fig. 6, the presence of in vitro translated E2A did not prevent the interaction of SCL and p44.

**DISCUSSION**

Using a yeast two-hybrid system, we have isolated several SCL-binding proteins, both expected and unexpected, from three different human cDNA libraries. As expected, the most prevalent SCL-binding partners recovered from these screens were E proteins, including E2A/ITF1, E2-2/ITF2, and HEB (14). We also used the yeast two-hybrid system to demonstrate previously unreported interactions between SCL and NHLH1 as well as a homodimeric interaction of SCL with itself. A human form of the DRG protein was also isolated from a human thymus cDNA library; this was not surprising given that mouse DRG has been isolated from a mouse erythroleu-
kemia cell line (MEL) (35) using the 22–26-kDa form of SCL as the bait protein in a yeast two-hybrid screen.

However, we were surprised by the isolation of the p44 subunit of TFIIH from the human thymus cDNA library. Although interactions between bHLH proteins (e.g., c-Myc and MyoD) and basal transcription machinery have been documented (41, 42), those studies involved the interaction with TBP, the first basal transcription factor that is recruited to an activated promoter (43). The interaction between a bHLH protein and TFIIH, which is recruited to the transcription initiation complex relatively late and functions in promoter clearance and elongation of transcription (25), has not been previously reported.

Several functional domains have been mapped to portions of the SCL protein; the amino-terminal region is proline-rich and has been shown to contain a transcription activation domain (21), whereas the bHLH domain is involved in both DNA binding and dimerization with other bHLH proteins (44). The p44-binding region of SCL was mapped to a region that includes the bHLH domain as well as amino acid residues immediately N- and C-terminal to the bHLH domain. Although the bHLH domain was included in the region essential for binding, it is possible that the bHLH domain may only serve as an arm to link the N- and C-terminal flanking regions together to form a binding domain for p44. Support for this possibility is found in the observation that only SCL, among several bHLH proteins tested, was able to bind p44 in the yeast two-hybrid system, indicating that the bHLH structure alone is not sufficient for this interaction. For example, LYL1 (a bHLH protein named by virtue of its involvement in a chromosomal translocation associated with lymphoblastic leukemia) has a bHLH region that is quite similar (49/55 identical amino acids and four conservative substitutions) to SCL (1). Despite this, an interaction between LYL1 and p44 was not detected with this system, suggesting that the interaction between SCL and p44 requires amino acids proximal and distal to the bHLH region. Our results also show that the C-terminal polyglycine tract is important for the interaction; this is not surprising because glycine-rich domains have been shown to mediate other protein/protein interactions (45). However, the SCL C-terminal glycine tract itself is not sufficient for the SCL/p44 interaction (see Fig. 4). The SCL-binding region of p44 has not yet been clearly defined, the above two motifs are potential targets for future studies.

It is not clear how the interaction between SCL and p44

![Fig. 4. Mapping of regions of SCL that bind p44. A series of SCL deletion mutants fused to GBD were tested in the yeast two-hybrid system. SCL amino acid (aa) residues are indicated; the solid black bars represent the bHLH domain. The interaction between SCL mutants and p44 was quantitated by observing the colony growth on –Leu/Trp/His SD plates and β-galactosidase assay. Interaction between SCL and ITF2 was used as a positive control. +++, colony growth to full size 2 days after transformation; +++, colony growth to full size 3 days after transformation; ++, colony growth to full size 4 days after transformation; +, colony growth to full size 5 days after transformation; –, no colony growth 5 days after transformation.](image)

![Fig. 5. Carboxyl terminus of p44 binds SCL. p44 and its truncated mutants were fused to GAD and tested in the yeast two-hybrid system with GBD-SCL. Δp44 represents the carboxyl-terminally truncated mutant, whereas Tp44 is the amino-terminally truncated mutant. The interaction was demonstrated by the colony growth on the –Leu/Trp/His SD plate; growth on the –Leu/Trp SD plate was used to monitor the transformation efficiency. GBD/SCL and GAD/ITF2 interactions were used as positive controls.](image)

![Fig. 6. E2A does not interfere with SCL/p44 interaction. In vitro translated E2A (*E2-5), SCL (*SCL), or SCL and E2A (*SCL + *E2-5) were incubated with GST or GST-p44, and interactions were detected by coprecipitation as described in the legend to Fig. 2B. The positions of SCL are indicated with two arrows.](image)
might function in a biological context. However, some insight may be gained through the study of previously reported interactions between nonessential transcription factors and elements of the basal transcription complex. Several proteins in addition to those of the basal transcription complex have been shown to interact with TFIIH subunits. These proteins include p53, which has been shown to interact in vitro with the XPB subunit of TFIIH (46); this interaction is thought to modulate either the nucleotide excision repair activity of TFIIH (47) or p53-mediated apoptosis (48). Additionally, the transcription activation domain of VP16 has been shown to bind the p62 subunit of TFIIH (49), resulting in activation of transcription from an adenovirus type 2 major late promoter. With regard to interactions between bHLH proteins and basal transcription factors, the bHLH transcription factor c-Myc has been shown to interact with p44, thus modulating the activity of the basal transcription complex. Several proteins in this category, such as SCL, have been found to interact with TFIIH subunits. One hypothesis regarding the functional relevance of an SCL/p44 interaction is that SCL binds the basal transcription factor TBP (50). Therefore, one hypothesis regarding the functional relevance of an SCL/p44 interaction is that SCL binds the basal transcription factor TBP (50).