A Conserved Transcription Motif Suggesting Functional Parallels between \textit{Caenorhabditis elegans} SKN-1 and Cap’n’Collar-related Basic Leucine Zipper Proteins*

Received for publication, March 2, 2000
Published, JBC Papers in Press, April 6, 2000, DOI 10.1074/jbc.M001746200

Amy K. Walker‡, Raymond See§, Ceri Batchelder‡, Thip Kophengnavong‡, J. Timothy Gronninger‡, Yang Shi§, and T. Keith Blackwell‡‡

From the ‡Center for Blood Research and the §Department of Pathology, Harvard Medical School, Boston, Massachusetts 02115

In \textit{Caenorhabditis elegans}, the predicted transcription factor SKN-1 is required for embryonic endodermal and mesodermal specification and for maintaining differentiated intestinal cells post-embryonically. The SKN-1 DNA-binding region is related to the Cap’n’Collar (CNC) family of basic leucine zipper proteins, but uniquely, SKN-1 binds DNA as a monomer. CNC proteins are absent in \textit{C. elegans}, however; and their involvement in the endoderm and mesoderm suggests some functional parallels to SKN-1. Using a cell culture assay, we show that SKN-1 induces transcription and contains three potent activation domains. The functional core of one domain is a short motif, the DIDLID element, which is highly conserved in a subgroup of vertebrate CNC proteins. The DIDLID element is important for SKN-1-driven transcription, suggesting a likely significance in other CNC proteins. SKN-1 binds to and activates transcription through the p300/cAMP-responsive element-binding protein-binding protein (CBP) coactivator, supporting the genetic prediction that SKN-1 recruits the \textit{C. elegans} p300/CBP ortholog, CBP-1. The DIDLID element appears to act independently of p300/CBP, however, suggesting a distinct conserved target. The evolutionarily preservation of the DIDLID transcriptional element supports the model that SKN-1 and some CNC proteins interact with analogous cofactors and may have preserved some similar functions despite having divergent DNA-binding domains.

During development, establishment of cell fates frequently involves conserved regulatory pathways. In the early \textit{Caenorhabditis elegans} embryo, endodermal (intestinal) fates are specified by GATA family transcription factors (END-1, END-3, ELT-2) that are related to those that mediate endoderm development in \textit{Drosophila} (Serpent) and vertebrates (GATA-4/5/6 (1–3). This program is triggered by maternally expressed SKN-1, a predicted transcription factor (4) that also specifies mesodermal lineages (pharynx and some body wall muscle) (5). The presence of consensus SKN-1-binding sites adjacent to the \textit{end-1} gene along with the timing of its expression suggests that SKN-1 may activate \textit{end-1} directly (2). SKN-1 is also required in the endoderm post-embryonically, to prevent differentiated intestinal cells from undergoing severe atrophy (5).

Apparent SKN-1 orthologs have not been identified outside of nematodes, but in its DNA-binding region, SKN-1 is related to a subgroup of basic leucine zipper transcription factors (4, 5), the CNC proteins. A basic DNA-binding region at the SKN-1 COOH terminus is particularly similar to those of CNC proteins, but SKN-1 lacks a zipper dimerization domain (Fig. 1, A and C) and, uniquely, binds DNA as a monomer (4). Its DNA binding requires the adjacent \(\alpha\)-helical “CNC region” (Fig. 1, A and C) (4, 6–8), which is otherwise found only in CNC proteins. An adjacent SKN-1 element that is lacking in CNC proteins, the NH\(_2\)-terminal arm (Fig. 1, A and C), contributes additional binding affinity and specificity (6, 9). These similarities and differences pose the question of whether SKN-1- and CNC-related proteins simply bind DNA through related but divergent mechanisms or might share a closer functional relationship.

Like SKN-1, many CNC proteins are involved in the development of endoderm or function of endodermal or mesodermal cells. \textit{Drosophila} CNC is required for specification of pharyngeal segments (10), and vertebrate p45 NF-E2 is involved in hematopoiesis (11). In mice, different knockouts of the \(NF1\) (NF-E2 related factor-1; also LCRF-1/TC11) gene either cause a fetal liver microenvironmental defect (12) or appear to block cell-to-cell induction of the mesoderm, a function usually ascribed to endodermal cells (13). Both \(NF1\) and \(NF2\) directly induce expression of detoxification enzymes (14–17), a pathway that is markedly stimulated in the liver and intestine (14). Supporting the idea that some CNC protein functions might parallel those of SKN-1, neither they nor their dimerization partners, the Maf proteins, have been identified. Genetic evidence suggests that SKN-1 may interact functionally with CBP-1, the \textit{C. elegans} ortholog of the p300/CBP transcription coactivators (20). p300/CBP proteins are metazoan histone acetyltransferases that are involved in developmental and inducible gene expression and that are recruited by numerous activators (21), including the CNC protein p45 NF-E2 (22). \textit{C. elegans} CBP-1 is required for specification of all non-neuronal embryonic developmental lineages (20). Endodermal differentiation can be restored in embryos that lack either CBP-1 or SKN-1, however, by inhibition of histone deacetylases

\[ \text{SKN-1} + \text{CBP-1} \rightarrow \text{active transcription complex} \]

---

*This work was supported by National Institutes of Health Grants GM50900 (to T. K. B.), GM58012 (to Y. S.), and DK09416 (to A. K. W.).

The abbreviations used are: CNC, Cap’n’Collar; CBP, cAMP-responsive element-binding protein-binding protein; CMV, cytomegalovirus; TK, thymidine kinase; CAT, chloramphenicol acetyltransferase; GST, glutathione S-transferase; PCAF, p300/CBP-associated factor.
(20). This suggests that in the endoderm, SKN-1 might recruit the CBP-1 histone acetyltransferase activity directly. It remains possible, nevertheless, that lack of histone deacetylases simply restores expression of downstream genes independently of possible histone acetyltransferase recruitment by SKN-1.

We have addressed these issues by investigating how SKN-1 regulates transcription. *C. elegans* cell lines have not been developed, but given the conservation of the eukaryotic mRNA transcription machinery (23), SKN-1 would be predicted to be active in mammalian cells. Using transfection assays, we show here that SKN-1 is a powerful activator of transcription that contains three transactivation domains. The functional core of one domain consists of a short sequence (the DIDLID element) that is specific to SKN-1-related proteins and to the NRF group of CNC proteins. The p300/CBP proteins appear to be direct cofactors of SKN-1, but not to be critical for the activity of the DIDLID element. The conservation and transcriptional function of the DIDLID element suggest that SKN-1 and the NRF CNC proteins share a common transcriptional target in addition to p300/CBP and, during evolution, may have maintained some parallel functions in endodermal cells.

**EXPERIMENTAL PROCEDURES**

**Plasmid Constructions**—For analysis of full-length SKN-1, a FLAG epitope was added at its NH$_2$ terminus. A SpeI site was first created by polymerase chain reaction, which added a linker of LV prior to the initial methionine, and then the FLAG epitope (DYKDDDDK) was added as an RI/SpeI fragment prior to cloning into the CMV-based vector CS2 (24). To delete the DIDLID motif (ΔDID), an SKN-1-(113–391) fragment was generated by polymerase chain reaction and then substituted for residues 100–391. Activation domain B was disrupted (ΔAD(B)) by deleting a restriction fragment encoding residues 154–205. The double deletion (ΔDID,ΔAD(B)) was made by swapping appropriate restriction fragments. Each deleted region was sequenced. Point mutations were constructed with the QuickChange kit (Stratagene) and confirmed by sequencing. Gal4-SKN-1 fusions were constructed by creating BamHI/XhoI sites at the ends of SKN-1 sequences using the polymerase chain reaction. These were ligated into the COOH terminus of the Gal4 DNA-binding domain within pSG424 (25), creating an intervening linker of ISRA. Junctions were sequenced for accuracy. The polymerase chain reaction was performed using Pfu polymerase (Stratagene). The SKN-1/TK-CAT reporter was constructed by blunting and multimerizing the SK1 oligonucleotide (4), which corresponds to the preferred SKN-1-binding site. A fragment with four SK1 sites in the same orientation was then ligated into a blunted SalI site in the TK-CAT reporter (26).

**Transfections and Analysis**—HeLa cells were used in all transfections and maintained in Dulbecco’s modified Eagle’s medium plus 10% fetal calf serum. Cells were transfected transiently by calcium phosphate for 18 h at 3% CO$_2$ and then placed in 5% CO$_2$ and harvested 24–26 h later. Transfections for CAT assays were performed in six-well plates and brought to 10 µg of total DNA with BlueScript. Transfections for Western blotting and electrophoretic mobility shift assays were performed in 100-mm plates using 20 µg of total DNA. CAT assays were performed as described (25). An internal control was omitted because some SKN-1 constructs influenced overall gene expression, but not transfection efficiency (data not shown), presumably through squelching. Samples were normalized for total protein concentration. Six independent values were generated for each data point, of which a representative is shown in each figure. Each fell in the linear range for CAT and represented an average of three experiments unless otherwise indicated, with error bars showing S.D. values. Cotransfection of a β-galactosidase vector revealed that in these assays, E1A expression did not cause detectable apoptosis (data not shown).

**Protein Analysis**—Lysates for Western and electrophoretic mobility shift assay analyses were prepared as described (27), and SKN-1 was expressed by *in vitro* translation using Promega kits. For Western blotting, proteins were separated on a 7.5% gel and transferred to nitrocellulose membranes (Schleicher & Schuell), which were probed with an anti-SKN-1 monoclonal antibody (a gift of J. Priess) and visualized by enhanced chemiluminescence (Amersham Pharmacia Biotech). Electrophoretic mobility shift assays were carried out as described (9) with an extract equivalent of 10% of a 100-cm plate/lane. Binding of SKN-1 to a $^{32}$P-labeled SK1 oligonucleotide (4) was assayed and was competed either by excess unlabeled SK1 DNA or by a mutant site (MSK1) to which SKN-1 does not bind specifically (4). Glutathione S-transferase (GST)-p300 proteins were expressed as described (28) and then coupled to glutathione-agarose beads (Sigma) and incubated for 2 h with $^{35}$S-labeled SK1 in 25 mm Heps, 250 mm NaCl, 25 mm EDTA, 0.1% Nonidet P-40, and protease inhibitor mixture as described (28). Beads were washed in phosphate-buffered saline plus 0.1% Nonidet P-40 five times, and then proteins were eluted in sample buffer and electrophoresed on an SDS-polyacrylamide gel.

**RESULTS**

**A Short Transactivation Motif Specific to SKN-1 and Some CNC-related Basic Leucine Zipper Proteins**—When expressed
by transfection in human (HeLa) cells, SKN-1 strongly activated a reporter containing four SKN-1-binding sites (SKN-1/TK-CAT) (Fig. 2), but not the corresponding control reporter lacking those sites (data not shown). SKN-1/TK-CAT was not activated by the SKN-1 DNA-binding domain alone (SKN-1(449–533)) or by SKN-1 residues 1–391, which encompass its similarity to a C. elegans SKN-1-related gene (SRG-1) (5) outside of the DNA-binding domain (data not shown) (Fig. 2). When fused to the yeast Gal4 DNA-binding domain, however, residues 1–391 (Gal4-SKN-(1–391)) strongly activated a Gal4-based reporter (E1B-CAT) (Fig. 3A) (28). They did not activate a control reporter lacking Gal4 sites (data not shown), suggesting that this SKN-1 region can bind and recruit transcription complexes. Residues 1–448 similarly constituted a powerful activator that was comparable to the viral protein VP16 (Fig. 3A).

Outside of their DNA-binding domains, SKN-1, NRF1, and NRF2 are highly related within a short NH₂-terminal motif, which we refer to as “DIDLID” after part of its sequence (Fig. 1, A and B). This motif is also present in SRG-1 and an SKN-1 homolog in the nematode Caenorhabditis briggsae (Fig. 1B), but among CNC proteins, it appears to be restricted to NRF1 and NRF2 orthologs (data not shown). The DIDLID element appears to be present only in SKN-1- and CNC-related proteins, and SKN-1 and NRF1 are more closely related within it (92%) (Fig. 1, B and C) than in their DNA-binding basic regions (86%), suggesting that it has a specific and conserved function. The NRF2 DIDLID element is located adjacent to a domain that is lacking in SKN-1 and that can retain NRF2 in the cytoplasm, but the DIDLID element is not involved in this interaction (29). The DIDLID element includes alternating charged and hydrophobic residues (Fig. 1B), but is not predicted to form an amphipathic α-helix. It is reminiscent, however, of short helical protein-protein interaction modules that are involved in transcription, such as the LXXLL motif in nuclear receptor co-activators (30, 31) and the LDIFS motif in E2A proteins (32), suggesting that the DIDLID element might also have a transcriptional function.

Supporting this idea, activation by the SKN-1 NH₂ terminus was not substantially diminished by COOH-terminal deletions to position 112 (Gal4-SKN-(1–112)), but was abrogated by further removal of the DIDLID element (Gal4-SKN-(1–98)) (Fig. 3A). Within residues 1–112, only the 14-residue DIDLID element activated significantly on its own, to ~10% of the level induced by SKN-1(1–391) or VP16 (Gal4-SKN-(99–112)) (Fig. 3A), suggesting that it forms the functional core of the SKN-1 NH₂-terminal activation domain (domain A) (Fig. 1A). Deletion of the DIDLID motif from full-length SKN-1 (ΔDID) decreased transcription of SKN-1/TK-CAT by 60% (Fig. 4A), but did not impair SKN-1 expression or DNA binding (Fig. 4, B and C). Transactivation by SKN-1 was comparably decreased by mutation of its third glutamic acid to alanine (D105A), but not arginine (D105R) (Fig. 4A), suggesting that the hydrophilicity of this residue might be more important than its charge.
Conserved Transactivation Motifs in C. elegans SKN-1

**Figure 4. Analysis of activation domains in full-length SKN-1.** A, activity of SKN-1 mutants. Vectors expressing SKN-1 or the indicated mutants (2.0 μg) were assayed for activation of the SKN-1/TJTCAT reporter (1.0 μg) (see Fig. 2). Sequence domains are indicated as described in the legend to Fig. 1. B, DNA binding by SKN-1 mutants. Extracts from cells transfected with the indicated SKN-1 derivatives were tested by electrophoretic mobility shift assay along with in vitro translated (IVT) SKN-1 for binding to a labeled SKN-1-binding site (4). Only bound complexes are shown, with a nonspecific species indicated by an arrow. C, expression of SKN-1 derivatives, assayed by Western blotting of the six protein samples analyzed in B. By this assay, the DIDLID point mutants analyzed in A were also expressed comparably to SKN-1 (not shown).

**DISCUSSION**

The presence of a DNA-binding domain in SKN-1 (4) and its localization to nuclei (36) suggested previously that SKN-1 is likely to regulate transcription. We have shown here that SKN-1 is a potent activator of transcription when it binds its cognate site (Fig. 2). SKN-1 interacts with two regions of human p300 in vitro (Fig. 5A), and p300/CBP proteins contribute to its activity (Fig. 5, B and C). Given the extensive similarity between p300/CBP proteins and their C. elegans ortholog CBP-1, these data support the model that SKN-1 may recruit CBP-1 directly to promoters as a cofactor in vivo (20). Overexpression of p300 potentiated SKN-1 activation of domains B and C (Fig. 5D), suggesting that it can be recruited directly or indirectly by them. Domain B was not inhibited by E1A mutants in which binding to either p300/CBP (p300mut) or PCAF (E55) in particular was impaired (Fig. 5E), implicating each of these histone acetyltransferases in its activity. Domain C was partially inhibited by these E1A mutants (Fig. 5E), however, suggesting either that it may require both histone acetyltransferases simultaneously or might act on an independent E1A target.

The conservation of the DIDLID element across evolution (Fig. 1, B and C) supports the model that SKN-1 and the CNC proteins evolved from a common precursor (4). This conservation is particularly striking because the DIDLID motif is separate from the DNA-binding domain and was maintained despite divergences in how these proteins bind DNA (Fig. 1C).
Point mutants in the DIDLID element dramatically decreased SKN-1-driven transcription in mammalian cells (Fig. 4A), supporting the idea that it mediates a highly specific interaction that is common to these proteins. In apparent contradiction to our findings, an NRF2 fragment that contained the DIDLID element (8). The CNC protein most analogous to SKN-1 may instead form a pocket, suggesting a common protein-protein interaction (8). The CNC protein most analogous to SKN-1 may instead form a pocket, suggesting a common protein-protein interaction (8). The CNC protein most analogous to SKN-1 may instead form a pocket, suggesting a common protein-protein interaction (8). The CNC protein most analogous to SKN-1 may instead form a pocket, suggesting a common protein-protein interaction (8). The CNC protein most analogous to SKN-1 may instead form a pocket, suggesting a common protein-protein interaction (8). The CNC protein most analogous to SKN-1 may instead form a pocket, suggesting a common protein-protein interaction (8). The CNC protein most analogous to SKN-1 may instead form a pocket, suggesting a common protein-protein interaction (8). The CNC protein most analogous to SKN-1 may instead form a pocket, suggesting a common protein-protein interaction (8).

Acknowledgments—We thank Martin Victor and Blackwell laboratory members for critically reading the manuscript.

REFERENCES
1. Zhu, J., Fukushige, T., McGhee, J. D., and Rothman, J. H. (1998) Genes Dev. 12, 3869–3884
2. Zhu, J., Hill, R. J., Heid, P. J., Fukuyama, M., Sugimoto, A., Priess, J. R., and Rothman, J. H. (1997) Genes Dev. 11, 2683–2686
3. Fukushige, T., Hawkins, M. G., and McGhee, J. D. (1998) Dev. Biol. 198, 286–292
4. Blackwell, T. K., Bowerman, B., Priess, J., and Weintraub, H. (1994) Science 266, 621–628
5. Bowerman, B., Eaton, B. A., and Priess, J. R. (1995) Cell 80, 1061–1075
6. Carroll, A. S., Gilbert, D. E., Liu, X., Pelczer, I., Pal, S., and Walker, S. (1998) Proc. Natl. Acad. Sci. U. S. A. 95, 8455–8460
7. Rupert, P. B., Daughdrill, G. W., Bowerman, B., and Matthews, B. W. (1998) Nat. Struct. Biol. 5, 484–491
8. Kohlenov, T., Carroll, A. S., and Blackwell, T. K. (1999) Mol. Cell. Biol. 19, 3039–3050
9. Mohler, J., Mahaffey, J. W., Deutsch, E., and Vani, K. (1995) Development 121, 237–247
10. Shivalasani, R. A., Rosenblatt, M. F., Zucker-Franklin, D., Jackson, C. W., Hunt, P., Saris, C. J., and Orkin, S. H. (1995) Cell 81, 695–704
11. Chan, J. Y., Kwong, M., Lu, R., Chang, J., Wang, B., Yen, T. S., and Kan, Y. W. (1998) EMBO J. 17, 1779–1787
12. Farmer, S. C., Sun, C. W., Winnier, G. E., Wagner, G., Ellenberger, T. E., and Blackwell, T. K. (1997) Genes Dev. 11, 2227–2238
13. Lo, M.-C., Ha, S., Pelczer, I., Pal, S., and Walker, S. (1998) Proc. Natl. Acad. Sci. U. S. A. 95, 8455–8460
14. Itoh, K., Chiba, T., Takahashi, S., Ishii, T., Igarashi, K., Katoh, Y., Ohyama, T., Hayashi, N., Satoh, K., Hayama, I., Yamamoto, M., and Nabeshima, Y. (1997) Biochem. Biophys. Res. Commun. 236, 313–322
15. Venugopal, R., and Jiang, W. K. (1998) Oncogene 17, 3145–3156
16. Kwon, M., Kan, Y. W., and Chan, J. Y. (1999) J. Biol. Chem. 274, 37491–37498
17. Alam, J., Stewart, D., Touchard, C., Boinapally, S., Chai, A. M., and Cook, J. L. (1999) J. Biol. Chem. 274, 26071–26078
18. Blank, V., and Andrews, N. C. (1997) Trends Biochem. Sci. 22, 437–441
19. Mutohshii, H., Shavit, J. A., Igarashi, K., Yamamoto, M., and Engel, J. D. (1997) Nucleic Acids Res. 25, 2953–2959
20. Shi, Y., and Mello, C. (1998) Genes Dev. 12, 943–955
Conserved Transactivation Motifs in C. elegans SKN-1

21. Shikama, N. X., Lyon, J., and La Thangue, N. B. (1997) Trends Cell Biol. 7, 230–236
22. Cheng, X., Reginato, M. J., Andrews, N. C., and Lazar, M. A. (1997) Mol. Cell. Biol. 17, 1407–1416
23. Hampsey, M. (1998) Microbiol. Mol. Biol. Rev. 62, 465–503
24. Turner, D., and Weintraub, H. (1994) Genes Dev. 8, 1434–1447
25. Batchelder, C., Dunn, M. A., Choy, B., Suh, Y., Cassie, C., Shim, E. Y., Shin, T. H., Mello, C., Seydoux, G., and Blackwell, T. K. (1999) Genes Dev. 13, 202–212
26. Weintraub, H., Davis, R., Lockshon, D., and Lassar, A. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 5623–5627
27. Andrews, N. C., and Faller, D. V. (1991) Nucleic Acids Res. 19, 2499
28. Lee, J. S., Galvin, K. M., See, R. H., Eckner, R., Livingston, D., Moran, E., and Shi, Y. (1995) Genes Dev. 9, 1188–1198
29. Itoh, K., Wakabayashi, N., Kato, Y., Ishii, T., Igarashi, K., Engel, J. D., and Yamamoto, M. (1999) Genes Dev. 13, 76–86
30. Heery, D. M., Kalkhoven, E., Hoare, S., and Parker, M. G. (1997) Nature 397, 753–736
31. Torchia, J., Rose, D. W., Inostroza, J., Kamei, Y., Westin, S., Glass, C. K., and Rosenfeld, M. G. (1997) Nature 387, 677–684
32. Massari, M. E., Grant, P. A., Pray-Grant, M. G., Berger, S. L., Workman, J. L., and Murre, C. (1999) Mol. Cell 4, 63–73
33. Reid, J. L., Bannister, A. J., Zegerman, P., Martinez-Balbas, M. A., and Kouzarides, T. (1998) EMBO J. 17, 4469–4477
34. Chakravarti, D., Ogryzko, V., Kan, H. Y., Nash, A., Chen, H., Nakatani, Y., and Evans, R. M. (1998) Cell 96, 393–403
35. Hamamori, Y., Sartorelli, V., Ogryzko, V., Puri, P. L., Wu, H. Y., Wang, J. Y., Nakatani, Y., and Reed, L. (1999) Cell 96, 405–413
36. Bowerman, B., Draper, B. W., Mello, C., and Priess, J. (1990) Cell 64, 443–452
37. Trouche, D., and Kouzarides, T. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 1439–1442
38. Wong, H. K., and Ziff, E. B. (1994) J. Virol. 68, 4910–4920