The SCAN Domain Mediates Selective Oligomerization*

Received for publication, January 7, 2000, and in revised form, February 29, 2000
Published, JBC Papers in Press, March 21, 2000, DOI 10.1074/jbc.M000119200

Christoph Schumacher‡§, Hubert Wang‡, Christian Honer‡, Wei Ding‡, James Koehn‡, Quentin Lawrence‡, Christopher M. Coulis‡, Lei Lei Wang‡, Dennis Ballinger‡, Benjamin R. Bowen‡, and Susanne Wagner‡§

From the ‡Novartis Institute for Biomedical Research, Summit, New Jersey 07901 and §Myriad Genetics, Salt Lake City, Utah 84108

The SCAN domain is described as a highly conserved, leucine-rich motif of approximately 60 amino acids found at the amino-terminal end of zinc finger transcription factors. Although no specific biological function has been attributed to the SCAN domain, its predicted amphipathic secondary structure led to the suggestion that this domain may mediate protein-protein associations. A yeast two-hybrid screen identified members of two SCAN domain protein families that interact with the SCAN domain of the zinc finger protein ZNF202. The interacting ZNF191 protein represents the family of SCAN domain-containing zinc finger proteins, whereas the novel SDP1 protein establishes a new family of genes that encode an isolated SCAN domain. Isolated SCAN domain proteins may form asymmetric homodimers in solution. Biochemical binding studies confirmed the associations of ZNF191 and SDP1 with ZNF202 and established the SCAN domain as a selective hetero- and homotypic oligomerization domain. SCAN mediated protein associations might therefore represent a new regulatory mechanism of transcriptional activity.

The SCAN or leucine-rich domain, originally identified by its homology with similar elements in several zinc finger transcription factors, consists of approximately 60 amino acids and is rich in leucine and glutamic acid residues (1). Most SCAN domain sequences are linked to Cys2-His2 zinc finger motifs through their carboxyl-terminal end. Although the function of the SCAN domain has not yet been elucidated, the predicted amphipathic structure of the domain led to the suggestion that SCAN box elements have the capacity to interact with other proteins, in particular with components of the transcriptional machinery (1).

The zinc finger protein ZNF2021 is expressed in two common splice variants, here referred to as m1 and m3 (2). Whereas the m1-splice form encodes a full-length protein of 648 amino acids with a SCAN box, a KRAB repression domain, and eight Cys2-His2 zinc finger motifs, the 133 amino acid product of the m3-splice form encompasses only the SCAN domain.2 These splice forms are conserved in the murine ZNF202 homolog, suggesting that the SCAN motif itself is an independent functional domain.3 The existence of other genes that encode SCAN elements as an isolated structural feature is further demonstrated by the recent identification of the murine Leap1/PCG-2 gene (GenBank™ accession AF106473) (3).

In order to study the function of the SCAN domain, we performed an extensive yeast two-hybrid screen for the identification of SCAN binding proteins. The screen, supported by biochemical association studies, suggested that SCAN motifs have the ability to associate selectively with each other. Furthermore, the first human gene encoding an isolated SCAN domain was identified. The formation of SCAN domain-mediated protein complexes may therefore modulate the biological function of transcription factors.

EXPERIMENTAL PROCEDURES

Two-hybrid Screen—A Gal4-DNA binding domain (Gal4-DBD) fusion construct was generated by ligating a cDNA fragment of ZNF202 (amino acids 1–199) encompassing the SCAN domain into the Gal4p DNA-binding domain vector pGBT-C (bait Z1) (4). Additional Gal4-DBD fusion constructs contained the SCAN domain of SDP1, amino acids 1–179, or ZNF191, amino acids 1–204. The Gal4-DBD constructs ("baits") were transformed into the mating type yeast strain J692. Gal4 activation domain (Gal4-AD) libraries ("prey") from human B-cell, liver, kidney, and brain cDNA were obtained from CLONTECH and transformed into the a-mating type yeast strain J693. Individual Gal4-AD fusion constructs of ZNF202 encoded amino acids 1–44 (Z6), 1–171 (Z7), 202–328 (Z8), and 235–278 (Z10). Yeast strains containing the bait plasmid and the activation domain library were mated on filters and plated on minimal media lacking tryptophan, leucine, and histidine but containing 25 μM 3-amino-1,2-triazole. After incubation for approximately 8 days at 30 °C, the colonies that grew on the triplet dropout media were subjected to a fluorescent-labeled dye primer sequencing on ABI 377 sequencers (Applied Biosystems, Inc.). The assembled sequence was confirmed by sequencing a representative EST cDNA (GenBank™ accession number N75095; IMAGE Consortium cDNA clone 284448 obtained from Genome Systems).

Northern Blot Analysis—Labeled SDP1 probes were synthesized in vitro in the presence of [α-32P]dCTP (Amersham Pharmacia Biotech) from a cDNA template encompassing the open reading frame using random primers and Klenow enzyme (Promega) and were used to probe multiple human tissue Northern blots (CLONTECH). Hybridization analysis was carried out in Quickhyb solution (Stratagene) at 65 °C and

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EMBL Data Bank with accession number(s) N75095, U68536, and AF204271.

‡ To whom correspondence should be addressed: Novartis Institute for Biomedical Research, MCD Research, 556 Morris Ave., Summit, NJ 07901. Tel.: 908-277-4797; Fax: 908-277-4756; E-mail: christopher.schumacher@pharma.novartis.com.

§ Both authors contributed equally to this work.

1 The abbreviations used are: ZNF, zinc finger protein; Gal, galactosidase; DBD, DNA binding domain; AD, activation domain; SDP, SCAN-domain protein; PAGE, polyacrylamide gel electrophoresis; GST, glutathione S-transferase.

2 S. Wagner, unpublished results.

3 S. Morham, L. Huwyler, and M. Toth, unpublished results.
visualized by autoradiography. The blots were then stripped and reprobed with a \(^{32}\)P random prime-labeled glyceraldehyde-3-phosphate dehydrogenase cDNA probe (CLONTECH) to confirm equal loading.

**Gel Filtration Chromatography, Sedimentation Equilibrium Analysis, and Chemical Cross-linking of the ZNF202-m3 Protein**—A cDNA amplicon encoding the ZNF202-m3 splice form was subcloned into the NcoI site of the pFastBac1 expression vector (Novagen). The expressed scan protein of 133 amino acids was purified over a cation exchange column using a POROS HS20 column (Perkin-Elmer) and analyzed by fast protein liquid chromatography. The analytical size exclusion chromatography was performed on a Superdex-75 HR 10/30 column (Amersham Pharmacia Biotech) equilibrated with a buffer containing 50 mM HEPES, pH 7.5, 100 mM NaCl, and 2 mM dithiothreitol. A protein sample of approximately 1.0 mg in 100 μl of column buffer was loaded on the column and was eluted at a flow rate of 0.5 ml/min at 4 °C. Fractions of 0.4 ml were analyzed by absorbance at 280 nm and by the Bradford method for protein detection (Bio-Rad). The column was calibrated with bovine serum albumin (67 kDa), ovalbumin (45 kDa), chymotrypsinogen (25 kDa), and ribonuclease A (13.7 kDa) as protein standards (Amersham Pharmacia Biotech).

For the postcrosslinking measurement of sedimentation equilibrium, 1.0 mg of purified protein sample was layered on top of a 5–20% linear sucrose gradient established in equilibration buffer. As molecular weight standards, aldolase (154 kDa) and ovalbumin (43 kDa) were analyzed onto a second gradient, whereas bovine serum albumin (67 kDa) was eluted onto a third gradient. Centrifugation was carried out in a Beckman SW40 rotor at 23,000 rpm for 3 h at 4 °C. Fractions of 0.4 ml were collected from the bottom to the top of the gradient, and the presence of protein was analyzed by the Bradford method.

For chemical crosslinking experiments, a purified ZNF202-m3 sample of 1.6 mg/ml (0.1 mg protein) concentration was incubated with 0.4 or 2 mM bis(sulfosuccinimidyl)suberate (Pierce) for 30 min at room temperature. The crosslinking reaction was subsequently quenched with 50 mM Tris-HCl, pH 7.5, and analyzed by denaturing SDS-PAGE followed by Coomassie Blue staining. A recombinant Htra protease (a dodecamer in solution) and bovine serum albumin (a monomer) were used as cross-linking control samples in parallel reactions (11).

### RESULTS

**A Yeast Two-hybrid Screen Identifies Interacting SCN Domain Proteins**—A yeast two-hybrid approach was used to search for proteins that interact with the SCN domain in vivo. The amino-terminal end of ZNF202 (amino acids 1–199) was fused in frame to the DNA binding domain of Gal4 and was used as bait (Z1) to screen independently brain, kidney, liver and heart cDNA libraries (Fig. 1A). The screen resulted in the isolation of 48 interacting cDNA clones, 42 of which encoded a SCAN homology domain (Fig. 1B). Five cDNA clones corresponded to ZNF191, 28 to SDP1 (SCN domain protein 1), 4 to SDP2, and 5 to SDP3. Whereas ZNF191 encodes a SCAN-domain-containing zinc finger protein, the novel SDP1 cDNA of 0.9 kilobases was found to encode an}
isolated SCAN box without adjacent zinc finger motifs. The presence or absence of zinc finger motifs in SDP2 and SDP3 is unknown since the respective cDNAs are incomplete.

An additional yeast two-hybrid screen was performed with a SDP1 bait and a kidney library that resulted in the isolation of five SCAN domain encoding novel cDNAs (SDP1–4), or to unrelated cDNAs (others) are indicated. C, Gal4 DNA-binding domain-linked fragments of ZNF202, SDP1, and ZNF191 were tested for their abilities to interact with activation domain-linked fragments of ZNF202. Occurrence (+) or absence (−) of growth is indicated.

In order to further define the ZNF202 binding domain, additional ZNF202 Gal4 activation domain fusion constructs were directly probed with SCAN domain-derived cDNAs (SDP1–4), or to unrelated cDNAs (others) are indicated. C, Gal4 DNA-binding domain-linked fragments of ZNF202, SDP1, and ZNF191 were tested for their abilities to interact with activation domain-linked fragments of ZNF202. Occurrence (+) or absence (−) of growth is indicated.

In order to determine the stoichiometry of the ZNF202 SCAN domain complex in solution, a bacterially expressed ZNF202 SCAN protein complex in solution, a bacterially expressed SCAN domain-containing proteins interact in a cellular milieu. Various expression constructs of ZNF202 and SDP1 with epitope tags were employed in co-immunoprecipitation/immunoblot experiments in order to demonstrate intracellular associations. Fig. 4A shows an anti-ZNF202[KRAB] immunoblot of lysates from cells transfected with pGal4-ZNF202-m1 alone or with pGal4-ZNF202-m1 plus pXpress-SDP1. Lysates from these cells contain an appropriately sized (~110 kDa) immunoreactive band (3rd and 4th lanes). The anti-Xpress immunoprecipitation of the pGal4-ZNF202-m1/pXpress-SDP1 cotransfection shows a robust band (2nd lane), indicative of intracellular association between ZNF202 and SDP1, whereas a control antibody yielded no band (1st lane). Fig. 4B represents the reciprocal of Fig. 4A. Instead of coprecipitating ZNF202 with antibodies to tagged forms of SDP1, Xpress-SDP1 was precipitated with anti-ZNF202[KRAB] antibody (2nd lane). Thus, in lysates from appropriately transfected cells, ZNF202 precipitates SDP1 and SDP1 precipitates ZNF202 precisely as was seen in affinity purification studies and in far Western experiments.

The ZNF202-m3 Protein Forms an Asymmetric Homodimer—In order to determine the stoichiometry of the ZNF202 SCAN protein complex in solution, a bacterially ex-
pressed ZNF202-m3 protein was purified over a cation exchange column and subjected to size exclusion chromatography. A fast protein liquid chromatography experiment using a prepacked Superdex-75 HR column eluted the SCAN protein in a 52-kDa molecular size fraction as calculated by plotting the logarithm of the molecular mass of the protein standards against the elution volume (Fig. 5A). In contrast, an SDS-PAGE analysis of this protein separated the denatured and thus monomeric form at 15-kDa molecular size as detected by Coomassie Blue gel staining (Fig. 5A, inset). In addition, the postcentrifugation sedimentation equilibrium of the ZNF202-m3 protein through a 5–20% sucrose gradient was analyzed. As shown in Fig. 5B, the sedimentation profile of the m3 protein showed a well defined peak at approximately 48 kDa in the presence of 100 mM KCl. Thus, the migratory properties of the ZNF202-m3 protein subjected to size exclusion chromatography...

Fig. 2. SDP1 mRNA expression profile and SCAN domain alignment. A, radiolabeled cDNA probe for SDP1 was hybridized to tissue mRNA blots and detected by autoradiography. Molecular size markers are indicated on the left in kilobase units. The blot was subsequently hybridized to a glycerol-3-phosphate dehydrogenase probe to confirm equal RNA loading. Skeletal muscle, peripheral blood leukocyte; Small Int, small intestine. B, alignment of the SCAN domain amino acid sequences of SDP1, ZNF191, and ZNF202. Identical residues are boxed. Residues that are conserved in the published consensus (Cons.) sequence are indicated in bold (3).
phy and sedimentation analysis suggest that the SCAN domain of ZNF202 forms either a trimeric complex or a modestly asymmetric dimer in a native environment. Chemical cross-linking experiments, however, did not reveal a trimeric complex but rather a dimeric species (Fig. 5C). Thus, the combination of gel filtration, sedimentation equilibrium, and cross-linking experiments of the ZNF202-m3 protein support the presence of an asymmetric dimeric protein.

**Selective SCAN Domain Association Abilities of ZNF202, SDP1, and ZNF191**—In a preliminary attempt to evaluate the selectivity of the SCAN domain-mediated protein associations, we transcribed and translated in vitro cDNA templates for ZNF202-m1, ZNF191, SDP1, and three unrelated SCAN domain-containing zinc finger proteins indicated as mLDS-1, Zfp110, and mFPM315. The cell lysates were incubated with GST-fused SCAN box sequences derived from ZNF202, SDP1, or ZNF191, and the recaptured fusion proteins were analyzed for binding in vitro produced radiolabeled protein (Fig. 6). All three GST fusion proteins associated strongly with radiolabeled SDP1. Conversely, the GST-SDP1 construct readily precipitated ZNF202, SDP1 itself, ZNF191, mLDS-1, and Zfp110. Therefore, the SCAN domain encoded by SDP1 revealed the highest association affinity that was also observed in the far Western experiment (Fig. 3C). Furthermore, SDP1 showed the ability to self-associate, which confirmed the interaction detected in the yeast two-hybrid screen. The fusion proteins of ZNF202 and ZNF191 revealed similar affinities for the radiolabeled proteins yet differed in their affinity for mLDS-1 and Zfp110. Whereas the ZNF202 fusion protein interacted weakly with mLDS-1, the ZNF191 fusion protein precipitated Zfp110. In comparison, the GST-SDP1 construct bound both mLDS-1 and Zfp110. None of the fusion proteins interacted with mFPM315. Therefore, the individual affinity of the fusion proteins to the unrelated SCAN proteins suggests the ability of SCAN domains to associate selectively with each other.

**DISCUSSION**

The SCAN domain is a sequence motif of unknown function that is common to zinc finger proteins. We provide evidence that the SCAN domain mediates selective protein oligomerization. A yeast two-hybrid screen identified homotypic and heterotypic interactions between SCAN domain-containing proteins. Biochemical binding experiments confirmed the interactions in vitro as well as in vivo and revealed selectivity in the SCAN domain-mediated protein associations as individual SCAN domains exhibited distinct differences in their ability to bind to other SCAN domain proteins. The structural...
components within a SCAN domain protein that contribute to binding selectivity and affinity remain to be elucidated. Sequence alignments of SCAN domains show a high degree of amino acid conservation that suggests that relatively few amino acids contribute to binding specificity (3). A hydrodynamic analysis of the ZNF202-m3 protein by gel filtration chromatography and postcentrifugation measurement of sedimentation equilibrium suggests the formation of a trimeric or asymmetric dimeric protein complex in solution. Dynamic light scattering investigations of the ZNF202-m3 protein in solution could not distinguish between these two oligomerization states.\(^4\) In vitro chemical cross-linking experiments, however, indicated a dimeric protein in identical buffer conditions. Therefore, we presume the ZNF202-m3 protein to form an asymmetric dimer in solution.

The state of oligomerization in SCAN domain-mediated homo- or heterotypic protein associations may share similarities with the BTB/POZ domain-containing proteins (12, 13). The GAGA protein has been shown to form BTB/POZ domain-mediated dimers, tetramers, and oligomers in vitro, whereas the BTB/POZ domain from the promyelocytic leukemia zinc finger oncprotein crystallized in a tightly intertwined dimer (14, 15). Furthermore, the BTB/POZ domain formed dimer-dimer interactions in the crystals that suggested a mode of higher order protein oligomerization (16).

Transcription factors are composed of functional domains that contribute to DNA binding, ligand binding, and transcriptional activation or repression (17, 18). One important regulatory mechanism affecting both DNA binding affinity as well as transcriptional activation is oligomerization (19, 20). Whereas zinc finger transcription factors of the Cys\(_2\)-His\(_2\) class are reported to bind DNA in monomeric form, SCAN domain-mediated association with other zinc finger proteins may modulate their biological function (21–23). The juxtaposition of zinc finger motifs through protein oligomerization may lead to altered DNA binding activities. For example, the self-oligomerization of the amino-terminal BTB/POZ domain of the GAGA zinc finger protein resulted in increased DNA binding affinity and transcriptional activity (14, 24). Recently, BTB/POZ domain-mediated oligomerization of transcription factors has been suggested to serve as combinatorial codes for gene expression (25).

The identification of the SDP1 protein establishes a new family of proteins characterized by an isolated SCAN domain. The amino acid sequence conservation between SDP1 and PGC-2 is high, and consequently SDP1 may represent the human ortholog of the murine PGC-2 protein (3). A third protein that contains only a SCAN domain is encoded by the m3 splice-form of ZNF202.

Single domain proteins that function as intermolecular regulators of transcription have been described previously. The PGC-2 protein was shown to interact with the nuclear receptor peroxisome proliferator-activated receptor \(\gamma\). Upon association, PGC-2 increases the transcriptional activity of peroxisome proliferator-activated receptor \(\gamma\) in the absence of any intrinsic transcriptional activity of PGC-2 (3). The biological effect of peroxisome proliferator-activated receptor \(\gamma\) may be modulated through competitive binding of PGC-2 and other transcriptional cofactors. Conversely, the DNA binding activities of some basic helix-loop-helix motif-containing transcription factors were shown to be inhibited by Id proteins. The Id family of helix-loop-helix proteins, which lacks a basic DNA-binding domain, functions as a negative regulator of basic helix-loop-helix proteins through the formation of inactive heterodimers with intact basic helix-loop-helix transcription factors (26, 27). Similarly, isolated SCAN domain proteins such as SDP1 or ZNF202-m3 may modulate the formation of functional SCAN domain zinc finger transcription factors.

The elucidation of the biological consequences arising from associations between SCAN domain proteins may offer new insights into the regulation of zinc finger transcription factors. Transient transfection experiments with appropriate reporter gene constructs may address the impact of SCAN domain-mediated oligomerization on gene expression. Particularly, the association of two different splice forms, such as the m1 and m3 forms of ZNF202, may represent a novel biological mechanism to regulate transcriptional activity.

Acknowledgments—We thank Drs. Kenton Zavitz and Mathew Toth for critical reading of the manuscript.

\(^{4}\) K. Clark and T. Stams, unpublished results.
Selective SCAN Domain Oligomerization

Azorin, F. (1999) J. Biol. Chem. 274, 16461–16469
15. Ahmad, K. F., Engel, C. K., and Privé, G. G. (1998) Proc. Natl. Acad. Sci. U. S. A. 95, 12123–12128
16. Li, X., Peng, H., Schultz, D. C., Lopez-Guisa, J. M., Rauscher, F. J., III, and Marmorstein, R. (1999) Cancer Res. 59, 5275–5282
17. Lefstin, J. A., and Yamamoto, K. R. (1998) Nature 392, 885–888
18. Tan, S., and Richmond, T. J. (1998) Curr. Opin. Struct. Biol. 8, 41–48
19. Chen, L. (1999) Curr. Opin. Struct. Biol. 9, 48–55
20. Wolberger, C. (1998) Curr. Opin. Genet. & Dev. 8, 552–559
21. Pavletich, N. P., and Pabo, C. O. (1991) Science 252, 809–817
22. Fairall, L., Schwabe, J. W., Chapman, L., Finch, J. T., and Rhodes, D. (1994) Nature 366, 483–487
23. Pavletich, N. P., and Pabo, C. O. (1993) Science 261, 1701–1707
24. Katsani, K. R., Hajibagheri, M. A., and Verrijzer, C. P. (1999) EMBO J. 18, 698–708
25. Kobayashi, A., Yamagiwa, H., Hoshino, H., Muto, A., Sato, K., Morita, M., Hayashi, N., Yamamoto, M., and Igarashi, K. (2000) Mol. Cell. Biol. 20, 1733–1746
26. Sun, X. H., Copeland, N. G., Jenkins, N. A., and Baltimore, D. (1991) Mol. Cell. Biol. 11, 5603–5611
27. Benezra, R., Davis, R. L., Lockshon, D., Turner, D. L., and Weintraub, H. (1990) Cell 61, 49–59
28. Williams, A. J., Blacklow, S. C., and Collins, T. (1999) Mol. Cell. Biol. 19, 8526–8535
