Sex-Determining Region Y-Related Protein SOX13 Is a Diabetes Autoantigen Expressed in Pancreatic Islets

Helen Kasimiotis, Mark A. Myers, Anthony Argentaro, Sabine Mertin, Shahnaz Fida, Tania Ferraro, Jane Olsson, Merrill J. Rowley, and Vincent R. Harley

The SOX (sex-determining region [SRY]-type high mobility group [HMG] box) family of transcription factors play key roles in determining cell fate during organ development. In this study, we have identified a new human SOX gene, SOX13, as encoding the type 1 diabetes autoantigen, islet cell antigen 12 (ICA12). Sequence analysis showed that SOX13 belongs to the class D subgroup of SOX transcription factors, which contain a leucine zipper motif and a region rich in glutamine. SOX13 autoantibodies occurred at a significantly higher frequency among 188 people with type 1 diabetes (18%) than among 88 with type 2 diabetes (6%) or 175 healthy control subjects (4%). Deletion mapping of the antibody epitopes showed that the autoantibodies were primarily directed against an epitope requiring the majority of the protein. SOX13 RNA was detected in most human tissues, with the highest levels in the pancreas, placenta, and kidney. Immunohistochemistry on sections of human pancreas identified SOX13 in the islets of Langerhans, where staining was mostly cytoplasmic. In mouse pancreas, Sox13 was present in the nucleus and cytoplasm of β-cells as well as other islet cell types. Recombinant SOX13 protein bound to the SOX consensus DNA motif AACAAT, and binding was inhibited by homodimer formation. These observations—along with the known molecular interactions of the closely related protein, rainbow trout Sox23—suggest that SOX13 may be activated for nuclear import and DNA binding through heterodimer formation. In conclusion, we have identified ICA12 as the putative transcription factor SOX13 and demonstrated an increased frequency of autoantibody reactivity in sera from type 1 diabetic subjects compared with type 2 diabetic and healthy control subjects. Diabetes 49:555-561, 2000

Type 1 diabetes results from the destruction of pancreatic islet β-cells by an autoimmune response to β-cell constituents, which is initiated by an unknown environmental agent acting on a susceptible genetic background. Identification of GAD (1), a family of tyrosine phosphatase-like proteins variously designated islet cell antigen 512 (ICA512)/IA-2/IA-2α (2,3), and insulin (4) as autoantigens in type 1 diabetes has led to the proposal of several mechanisms for the initiation and progression of islet autoimmunity. These include molecular mimicry between GAD and coxsackievirus B (5) or between ICA512 and rotavirus (6), hyperexpression of GAD in response to metabolic stress (7), and lack of tolerance to (pro)insulin as a consequence of genetic variation of proinsulin expression levels in the thymus (8,9). The multifactorial nature of type 1 diabetes and the association of numerous environmental agents with the disease (10) suggest that there may be multiple pathways leading to β-cell autoimmunity. To attain a better understanding of the pathogenesis of type 1 diabetes, the autoimmune response to islet β-cells needs to be fully characterized. In a search for novel diabetes-associated autoantigens, the screen of an islet cDNA expression library with type 1 diabetes sera that identified ICA512 also yielded a second clone, designated ICA12, which encoded a protein that reacted with type 1 diabetes sera by Western blot (11). Here, sequencing of the ICA12 cDNA revealed 99% sequence similarity to the high mobility group (HMG) domain of mouse Sox13 (12,13) and identified ICA12 as a partial cDNA of the novel human gene SOX13. Sex-determining region Y (SRY)-type HMG box proteins (SOX proteins) are a large family of transcription factors related to the testis-determining factor SRY through their HMG domain (14), an 80 amino acid motif that binds and bends DNA (15,16). Evidence from clinical mutations in SRY, SOX9, and SOX10 and gene knockout experiments of Sox4 in the mouse suggest that SOX proteins play key roles in determining cell fate during organ development (17,18). In this article, we describe the gene sequence, expression in human tissues, and biochemical properties of SOX13 and the reactivity of autoantibodies to SOX13 from subjects with diabetes. Our results demonstrate that SOX13 is a minor type 1 diabetes autoantigen expressed in pancreatic islets and would function as a DNA-binding transcription factor through the formation of heterodimers.
**RESEARCH DESIGN AND METHODS**

DNA cloning and sequencing. The 3.2 kb ICA12 cDNA (11) was sequenced using an ABI377 DNA sequencer. The 5' end of human SOX13 was determined by 5'RACE using three antisense primers (SP1 5'-G C A G C C A G T T C T T C T G C T; SP2 5'-T C G G T T G A A G T C C A G C T T C T G C T-3'; and SP3 5'-T C G G T T G A A G T C C A G C T T C T G C T-3') and a 5'RACE kit (Boehringer Mannheim, Mannheim, Germany) with human placental RNA.

**Production of recombinant SOX13.** Five fragments of the ICA12 cDNA clone derived by polymerase chain reaction were inserted into the EcoRI/Sall sites of pET23b vector, which contains a T7 epitope tag at the amino terminus (Promega, Madison, WI). They encoded SOX13 (amino acids 66–604), which is the originally described ICA12 clone, and SOX13aLZ-Q (amino acids 327–604), which lacks the leucine zipper domain and the glutamine-rich region. Three further constructs were used for epitope mapping were SOX13(66–210), SOX13(209–334), and SOX13(485–604). For expression in bacteria, Escherichia coli BL21 cells were transformed with various PET23b-SOX13 and PET7-SOXQ and PET7-SRY control DNA. Two-milliliter cultures of SOX13(66–604) were grown for 6 h before isopropanilthylgalactoside induction, and 5 ml cultures of SOX9 and the various SOX13 constructs were grown for 2–3 h before induction. Soluble extracts were prepared (19), and protein expression of SOX13 was confirmed by Western blot analysis using SOX13 antisera and T7-tag antibody (Novagen, Madison, WI). The amount of recombinant SOX13 protein was estimated by filtration against bovine serum albumin on Coomassie blue stained SDS-PAGE gels. For production of [35S]-labeled protein, RNA was transcribed in vitro from various linearized pET23b-SOX13 plasmids with T7 polymerase, then translated in vitro by rabbit reticulocyte lysate (Promega) in the presence of [35S] methionine (Dupont, Boston, MA).

**Detection of autoantibodies to SOX13.** Autoantibodies to SOX13 (anti-SOX13) were detected by immunoprecipitation (20) using [35S]-labeled SOX13(66–604) or [35S]-labeled SOX13(115–604) as antigen. The two preparations gave similar results, and so the data were combined. The sera had previously been collected and tested for anti-GAD and anti-IA-2 for other studies (21–23). The sera were from 188 subjects with type 1 diabetes with a mean ± SD age at diagnosis of 29 ± 16 years and a duration of diabetes of 10 ± 11 years. 81 subjects with type 2 diabetes with a mean ± SD age at diagnosis of 52 ± 14 years and duration of 8 ± 7 years; and 175 healthy control subjects. Diagnoses were according to World Health Organization criteria (24), and use of the sera was approved by the Monash University Standing Committee on Ethics in Research on Humans. To compensate for variations between different batches of labeled proteins, a single reference serum was included in each assay, and the activity of each serum was expressed as a reference units (RU) calculated as a percentage of the activity of the reference after correction for nonspecific binding. A positive result was defined as greater than the mean ± 2 SD of the control subjects. Serum was set to >0 RU were considered to have high levels of antibody.

**Mapping of the antibody epitopes of SOX13.** Recombinant proteins were purified by T7-tag antibody affinity chromatography (Novagen, Madison, WI), separated by 10% SDS-PAGE, and transferred to nitrocellulose filters. Filters were blocked in 5% skim milk powder and then incubated with doubling dilutions of disease sera at 1:1,000 dilution. Bound antibody was detected using horseradish peroxidase (HRP)-conjugated anti-human immunoglobins (Silenus, Hawthorn, Australia) and color development. Bacterial lysates containing the HMG domain of SOX13(66–210), SOX13(209–334), and SOX13(485–604) were subjected to Western blot analysis using SOX13 antisera and the membrane was blocked and the proteins detected in the same manner.

**Sequence comparisons reveal a high degree of amino acid similarity to the mouse SOX13.** The DNA sequence of the ICA12 cDNA and 5′RACE products indicate a mRNA transcript length of 3,577 bp (Fig. 1A). The SOX13 sequence has been submitted to Genbank under accession code AF116571. A single open reading frame of 604 amino acids contains a leucine heptad repeat (leucine zipper domain) adjacent to a region rich in glutamine (20 of 39 amino acids), an HMG domain, and an acidic COOH-terminal (16/40 amino acids) (Fig. 1B). Sequence comparisons reveal a high degree of amino acid sequence similarity to the mouse SOX13 deduced from the mouse SOX13 described by Kido et al. (12) throughout the entire 595 amino acid open reading frame (Fig. 1C), indicating that the ICA12 cDNA is the human homologue of SOX13 from mouse. However, our identification of a 604 amino acid open reading frame for ICA12, called SOX13 hereafter, is in conflict with a recent report by Roose et al. (31), which describes an open reading frame for human SOX13 of 890 amino acids. This longer open reading frame can be explained by a missing G nucleotide at base pair 1950 (our numbering) in the sequence reported by Roose et al. (accession number AF083105), which results in a change in the open reading frame. We have confirmed the presence of this G at base pair 1950 in the ICA12 cDNA and genomic DNA from >30 individuals (A.A., personal communication). A sec-
FIG. 1. Sequence analysis of human SOX13 gene and protein. A: Diagrammatic representation of the 3,577 nucleotide SOX13 mRNA transcript comprising an 1812 nucleotide open reading frame (boxed) flanked by 5' (162 nt) and 3' (1603 nt) untranslated regions (UTR), shown as lines. The shaded regions represent LZ, leucine zipper domain; Q, the glutamine-rich region; HMG, the HMG DNA binding motif; acidic, the acid-rich tail. *Polyadenylation signal at 3564. The ICA12 clone is truncated at the 5' end and lacks the 5' UTR and the first 65 amino acids of the open reading frame. B: The deduced amino acid sequence of SOX13 protein. Leucine residues of the leucine zipper are boxed; the glutamine-rich region is underlined with a solid line; the HMG domain is shown in bold; and the acid-rich tail is underlined with a dashed line. C: Alignment of the amino acids of human SOX13 with related SOX proteins. The BLAST search for proteins related to human SOX13 identified mouse Sox13, Xenopus Sox12, and rainbow trout Sox23 (12,33,32). D: Amino acid comparison of the HMG domain of human SOX13 with other SOX proteins. Gaps represent the same amino acid at this position as for SOX13. Underlined amino acids are identical through all SOX proteins listed. Upper-case (SOX) is human; lower-case (Sox) is mouse. The accession numbers for each of the following sequences are: SOX5, X65657; Sox6, X65659; Sox17, D4947; SOX20, AB006687; Sox18, L35032; Sox1, X94126; Sox3, X94125; Sox2, U31967; Sox12, X53772; SOX9, X74056; SOX10, AF006501; and SOX21, AF107044.

ond discrepancy is the insertion of three nucleotides in the human SOX13 sequence reported by Roose et al. (accession number AF083105) at three separate positions between base pairs 1052 and 1171. SOX13 shares high amino acid similarity to Sox23 from rainbow trout (32) and to Sox12 from Xenopus (33) (Fig. 1C), which are more similar to SOX13 than to any other SOX protein. Hence, mouse Sox13, rainbow trout Sox23, and Xenopus Sox12 may be orthologues of human SOX13. The leucine zipper/glutamine-rich region and the HMG domain of these proteins are highly similar, suggesting a functional role in vertebrates. Among mammals, the SOX13 HMG domain sequence shows 90% identity to SOX5 and Sox6 (Fig. 1D) and therefore belongs to the class D subgroup of SOX proteins, which also contain leucine zipper and glutamine-rich regions.

SOX13 immunological reactivity in type 1 diabetes sera. Autoantibodies to SOX13 were detected in sera by radioimmunoprecipitation (RIP). Reactivity to SOX13 was significantly increased in type 1 diabetic patients compared with control subjects, being present in 34 of 188 (18%) patients with type 1 diabetes and in 8 of 175 (4.6%) nondiabetic control subjects (P = 0.0001, G^2-test with Yates’ correction; Fig. 2A). The frequency of 6.2% (5 of 81) in type 2 diabetes was significantly lower than in type 1 diabetes (P = 0.018), but not significantly different from control subjects (P = 0.81). We conclude that serological anti-SOX13 reactivity is associated with autoimmune type 1 diabetes. Overall, 78% of the diabetic sera were positive for anti-GAD and 30% were positive for anti-IA-2. Eleven of the 33 anti-SOX13 positive sera were negative for both anti-GAD and anti-IA-2, indicating that anti-SOX13 does commonly occur in the absence of the other autoantibody reactivities associated with type 1 diabetes. The remainder 22 anti-SOX13 positive sera were positive for anti-GAD, and 6 of these were also positive for anti-IA-2.

To define the immunoreactive region(s) of SOX13, sera were further tested by RIP using truncated versions of SOX13 (Fig. 2B). Eight of fifty sera from subjects with type 1 diabetes reacted with SOX13(66–604), of which the three with the highest level of antibody reactivity by RIP also reacted with SOX13(327–604). This suggests that autoantibodies are directed to at least two epitopes, one that requires amino acids 66–604 of SOX13, and a second confined within amino acids 327–604 that lacks the LZ-Q domains but includes the HMG domain. Five sera with high levels of antibody reactivity to SOX13 were tested by Western blot for reactivity against the HMG domains of recombinant SRY and SOX9, and all were negative (results not shown).

To directly demonstrate autoantibody reactivity to SOX13, seven sera from subjects with type 1 diabetes positive by RIP to SOX13(66–604) were further tested by Western blotting, using affinity-purified SOX13(66–604) as antigen. Figure 2C shows that four sera strongly positive by RIP detected the 66-kDa band, whereas three sera with lower levels of antibody by RIP did not. One serum that reacted with SOX13(327–604) and SOX13(66–604) by RIP did not bind SOX13(66–604) on the Western blot; the seven control sera also showed no reactivity.

Expression of SOX13. A single 4.2-kb SOX13 transcript was detected in all human tissues examined by Northern blot analysis (Fig. 3A). Quantitation of this band versus actin indicated that expression was highest in the pancreas, placenta, and kidney. Immunohistochemical analysis of human pancreases with a SOX13 antisera showed staining in the cytoplasm of a population of cells within the islets of Langerhans as well as isolated cells in the exocrine pancreas.
The immunostaining was blocked by the peptide used to raise the antisera (data not shown). Double labeling with antibodies to SOX13 and insulin was performed on sections of mouse pancreas. Sox13 staining was apparent in the islets of Langerhans in the cytoplasm of some islet cells, as was also seen in the human islets. In the mouse islets, however, there were also islet cells that showed distinct nuclear staining; many of these cells also expressed insulin (Fig. 3C). This indicates that Sox13 is expressed in islet β-cells as well as in other islet cell types.

SOX13 protein dimerization and DNA binding. Dimerization potential of SOX13 was investigated using bacterial recombinant T7 epitope-tagged SOX13 (amino acids 66–604) and a deletion derivative of SOX13 lacking the leucine zipper domain and glutamine-rich region (SOX13ΔLZ-Q). A Western blot probed with T7 antibody revealed a 66-kDa band for SOX13(66–604) and a 43-kDa band for SOX13ΔLZ-Q (Fig. 4A, lanes 1 and 2). To test if SOX13 can form homodimers, renatured Western blots were probed with 35S-labeled SOX13(66–604) or with 35S-SOX13ΔLZ-Q, produced by in vitro transcription/translation. The 35S-SOX13(66–604) probe bound to SOX13(66–604) but not to SOX13ΔLZ-Q on the Western blot (Fig. 4A, lanes 3 and 4), whereas the 35S-SOX13ΔLZ-Q probe failed to detect either protein (Fig. 4A, lanes 6 and 7). The pattern of proteolytic products of SOX13(66–604) detected by the antibody to the NH2-terminal, T7 epitope was identical to that detected by 35S-SOX13(66–604) (Fig. 4A, lanes 1 and 3), which is consistent with the presence of an NH2-terminal dimerization domain. These data indicate that SOX13 can form homodimers via the leucine zipper/glutamine-rich region.

Recombinant SOX13 was tested for binding to radiolabeled double-stranded DNA bearing the high-affinity SOX binding motif AACAAT (29,34,35). SOX13(66–604) protein bound the AACAAT sequence but not the CCGCGG sequence (Fig. 4B, lanes 5 and 2), which is consistent with the DNA binding specificity of other SOX proteins. However, the DNA binding activity of SOX13 was very low, as micromolar concentrations of protein were required for detectable complex formation. Given that DNA binding activity of rainbow trout Sox23 is completely inhibited by the presence of its leucine zipper region (31), we compared the DNA binding activity of SOX13(66–604) with SOX13ΔLZ-Q. Titration experiments of the two proteins indicate that ~70 times more (2.4 µmol/l) of
SOX13(66–604) protein is required to bind an equivalent amount of DNA as 17 nmol/l SOX13LZ-Q (Fig. 4B, lanes 5 and 6). A SOX13 antibody raised to a peptide located NH₂-terminal to the leucine zipper domain further retards the migration of the SOX13/DNA complex (Fig. 4B, lane 10), indicating that SOX13(66–604) binds DNA, albeit weakly, with an intact leucine zipper domain and consequently as a homodimer. We conclude that SOX13 forms a homodimer through its leucine zipper/glutamine-rich region and that dimerization causes a 70-fold reduction in DNA binding activity by its HMG domain.

**DISCUSSION**

We have demonstrated that the SOX13 gene encodes the type 1 diabetes autoantigen ICA12 and has the biochemical properties of a transcription factor. Antibodies to SOX13 are present in sera from 18% of patients of type 1 diabetes, and the association of these autoantibodies with autoimmune diabetes was confirmed by the low frequency in subjects with type 2 diabetes and healthy control subjects. The dominant antibody epitope of SOX13 requires the majority of the protein, which may be a consequence of dimerization. This suggests that the intact protein, rather than degradation products, is autoantigenic. Similarly, autoantibodies to GAD recognize a conformational epitope and a dimeric configuration of the antigen (36,37). Whether autoimmunity to SOX13 is a component of disease pathogenesis or reflects an autoantibody reaction to sequestered SOX13 that is exposed to the immune system due to β-cell destruction remains to be determined.

The prevalence of autoantibodies to SOX13 is low in comparison to anti-GAD, anti-ICA512, and insulin autoantibodies, which occur in 60–80% of patients with type 1 diabetes (38,39) and have proven to be predictive for the future development of the disease in healthy individuals (22,38,40–42). However, it is worth noting that antibodies to SOX13 can occur independently of these other autoantibody reactivities (43) and so may prove a useful addition to a panel of autoantibody reactivities for the identification of individuals at high risk of developing the disease.

The expression of SOX13 in the β-cells of the islets of Langerhans is consistent with an autoimmune response to SOX13 in type 1 diabetes, which is characterized by autoantibody responses to a variety of islet-cell components (1–4). SOX13 is also widely expressed in other tissues, particularly the heart and placenta, which may reflect expression in the walls of both large and small arteries as observed in the mouse (13). This expression profile in human tissues differs from that of other type 1 diabetes autoantigens, such as GAD, ICA512/IA-2, carboxypeptidase H, and the GM2-1 ganglioside, which are expressed solely in pancreatic islets and other neuroendocrine tissues (44–47).

SOX13 has the required characteristics to act as a transcription factor. Sequence comparisons among vertebrates point to functional roles for the leucine zipper/glutamine-rich region and HMG domain of SOX13. SOX13 contains an unusual leucine zipper motif of the form LX5QLX5QLX5QL, which may mediate specific protein–protein interactions during nuclear import and transcriptional regulation through heterodimerization. In rainbow trout Sox23, the leucine zipper interacts with nucleoporin p62, which in turn interacts with the nuclear pore complex (32). We have demonstrated that SOX13 can form homodimers through the leucine zipper/glutamine-rich region and that homodimerization significantly reduces DNA binding. Sox6 and rtSox23 homodimerize through their leucine zippers, which eliminates
their DNA binding activity. Sox6 homodimerization abolishes its transactivation activity in CHO cells (48). In contrast, we detect residual DNA binding activity by SOX13 homodimers. Heterodimerization of SOX5-SOX6 greatly increases their DNA binding and the subsequent activation by SOX9 at the Col2a1 enhancer (49). By analogy with SOX5 and SOX6, heterodimerization with unknown leucine zipper proteins may modulate SOX13 activity in vivo.

Immunohistochemistry indicates that SOX13 protein can be localized to the cytoplasm as well as the nucleus. In general, SOX proteins tend to localize to the nucleus, the site of DNA binding and the subsequent activation by SOX9 at the Col2a1 enhancer (49). By analogy with SOX5 and SOX6, heterodimerization with unknown leucine zipper proteins may modulate SOX13 activity in vivo.

Immunohistochemistry indicates that SOX13 protein can be localized to the cytoplasm as well as the nucleus. In general, SOX proteins tend to localize to the nucleus, the site of DNA binding, although cytoplasmic staining of SOX9 has been reported (50). Because SOX13 binds DNA poorly as a dimer, dimerization may hinder access of the HMG domain to DNA. Furthermore, the nuclear localization signal shared by SOX protein HMG domains (51) may also be obscured in the SOX13 homodimer. Exposure of the nuclear localization signal to importins may be facilitated through heterodimerization with factors such as p62 nucleoporin, which is known to interact with rtSox23 (32).

The identification of SOX13 as an islet-cell autoantigen further characterizes the autoimmune response in type 1 diabetes and may advance the development of comprehensive diagnostic tests for identifying subjects at risk of developing the disease. Furthermore, the biochemical properties of SOX13 suggest a functional role for this protein as a transcription factor.

ACKNOWLEDGMENTS

This work was funded and supported by an Australian National Health and Medical Research Council Institute Block grant (Key Registration Number 983001) (V.R.H., J.O., S.M., T.F.) and Bayer Corporation, USA (S.F., M.A.M., M.J.R.). H.K. is the recipient of a Monash University departmental scholarship. We thank Daniel Rabin for providing the ICA12 CDNA; Monica Brown for expert technical assistance; Wah Chin Boon and Evan Simpson for the gift of human placenta mRNA; and Paul Zimmet for his collaborative interest. We thank Ian Mackay, Anthony Albiston, and Sharon McDowall for critical reading of the manuscript.

REFERENCES

H. KASIMIOTIS AND ASSOCIATES


39. Lefebvre V, Li P, de Combrugghe B: A new long form of Sox5 (L-Sox5), Sox6 and Sox9 cross-express in chondrogenesis and cooperatively activate the type II collagen gene. EMBO J 17:3713–3723, 1998


D I A B E T E S, V O L . 4 9 , A P R I L 2 0 0 0