Complementation Cloning of S2P, a Gene Encoding a Putative Metalloprotease Required for Intramembrane Cleavage of SREBPs


*Department of Molecular Genetics
University of Texas Southwestern Medical Center
Dallas, Texas 75235
†Department of Biochemistry
Dartmouth Medical School
Hanover, New Hampshire 03755

Summary

We report the cloning of a gene, S2P, that encodes a putative metalloprotease required for intramembrane proteolysis of sterol-regulatory element-binding proteins (SREBPs) at Site-2. SREBPs are membrane-bound transcription factors that activate genes regulating cholesterol metabolism. The active NH2-terminal domains of SREBPs are released from membranes by sequential cleavage at two sites: Site-1, within the lumen of the endoplasmic reticulum; and Site-2, within a transmembrane segment. The human S2P gene was cloned by complementation of mutant CHO cells that cannot cleave SREBPs at Site-2 and are cholesterol auxotrophs. S2P defines a new family of polytopic membrane proteins that contain an HEXXXH sequence characteristic of zinc metalloproteases. Mutation of the putative zinc-binding residues abolishes S2P activity. S2P encodes an unusual metalloprotease that cleaves proteins within transmembrane segments.

Introduction

Animal cells in tissue culture synthesize cholesterol actively when grown in cholesterol-free medium, and they repress this synthesis when cholesterol is supplied externally in the form of cholesterol-carrying lipoproteins such as low density lipoproteins (LDL). Uptake of LDL cholesterol is mediated by LDL receptors, which are also synthesized actively in the absence of sterols and suppressed when sterols are abundant (Goldstein and Brown, 1990). Insights into the mechanism of this feed-back regulation have emerged recently from studies of a family of membrane-bound transcription factors designated as sterol-regulatory element-binding proteins (SREBPs) (Brown and Goldstein, 1997).

SREBPs are tripartite proteins of approximately 1150 amino acids. The NH2-terminal segment of ~480 amino acids is a transcription factor of the basic helix-loop-helix-leucine zipper family (bHLH-Zip). This is followed by a membrane attachment domain of ~80 amino acids containing two membrane-spanning helices separated by a short hydrophilic segment of ~31 amino acids. The third segment is a COOH-terminal domain of ~590 amino acids that performs a regulatory role (Sakai et al., 1997). The SREBPs are oriented on membranes of the endoplasmic reticulum (ER) and nuclear envelope in a hairpin configuration with the NH2-terminal and COOH-terminal segments facing the cytosol and the hydrophilic loop projecting into the lumen (Brown and Goldstein, 1997).

In sterol-depleted cells, a two-step proteolytic process releases the NH2-terminal segment of the SREBP so that it can enter the nucleus where it binds to sterol-regulatory elements (SREs) in the promoters of genes encoding the LDL receptor and several enzymes of cholesterol synthesis, including 3-hydroxy-3-methylglutaryl coenzyme A (HMG CoA) synthase, HMG CoA reductase, farnesyl dipiphosphate synthase, and squalene synthase (Ericsson et al., 1996; Brown and Goldstein, 1997). In addition, the NH2-terminal segments of the SREBPs bind to elements in the promoters of acetyl CoA carboxylase and fatty acid synthase, activating synthesis of fatty acids (Tontonoz et al., 1993; Kim and Spiegelman, 1996; Lopez et al., 1996; Magana and Osborne, 1996; Shimano et al., 1996). In transgenic mice, SREBP-1a increases the mRNA for stearoyl CoA desaturase-1, thereby increasing production of unsaturated fatty acids (Shimano et al., 1996).

Sterols repress gene transcription by blocking the proteolytic release of SREBPs from membranes (Wang et al., 1994). Proteolysis begins in the ER where a protease clips the SREBPs at Site-1, which is between the leucine and serine of the sequence RSVLS in the hydrophilic luminal loop (Duncan et al., 1997). The putative sterol sensor is a polytopic membrane protein designated SCAP (SREBP cleavage-activating protein), which stimulates Site-1 cleavage in sterol-depleted cells (Hua et al., 1996). The action of SCAP is abolished when sterols build up in cells, and this prevents the NH2-terminal domain of the SREBPs from entering the nucleus and activating transcription. The Site-1 cleavage reaction breaks the covalent bond between the two transmembrane domains of the SREBPs, but it is not sufficient to release the NH2-terminal segment of the membrane. A second protease, designated the Site-2 protease (S2P), cleaves the protein within the first transmembrane segment, thereby releasing the NH2-terminal segment with a portion of the transmembrane segment attached (Sakai et al., 1996). The activity of S2P is not directly regulated by sterols, but it cannot act unless Site-1 cleavage has taken place (Sakai et al., 1996; Brown and Goldstein, 1997). Neither the Site-1 protease nor the Site-2 protease has been identified to date.

In the current studies, we isolate a gene required for Site-2 proteolysis by complementation cloning in a line of Chinese hamster ovary (CHO) cells that lack this protease. The defective cell line, designated M19 cells, was isolated by Hasan et al. (1994), who mutagenized CHO cells with y-irradiation and then selected cells that could not synthesize cholesterol or take it up from plasma LDL. The selection procedure made use of amphotericin B, an antibiotic that kills mammalian cells by forming complexes with cholesterol in cell membranes (Chang...
and Chang, 1982). When wild-type CHO cells are placed in sterol-free medium, they activate sterol synthesis, and the membrane retains sufficient cholesterol to allow killing by amphotericin B. M19 cells, which cannot synthesize cholesterol, become depleted of cholesterol when placed in cholesterol-free medium. They are therefore resistant to killing by amphotericin B (Hasan et al., 1994).

Whereas wild-type CHO cells will grow in cholesterol-free medium, M19 cells cannot. In addition to exogenous cholesterol, the M19 cells require external supplies of unsaturated fatty acids for optimal growth. These cells were unable to increase the amount of mRNA for HMG CoA-synthase when deprived of sterols, nor were they able to increase the number of LDL receptors. Cell-fusion experiments showed that the defect in the M19 cells was recessive (Hasan et al., 1994).

The nature of the defect in M19 cells became clear recently when it was found that they lack the ability to cleave SREBPs at Site-2 (Sakai et al., 1996). When deprived of sterols, these cells cleave SREBPs at Site-1, but the NH2-terminal domains remain attached to the membrane in an intermediate form. Since the NH2 terminus cannot enter the nucleus, the cells are unable to increase transcription of the genes for enzymes in the cholesterol biosynthetic pathway or the LDL receptor. They also have a reduced transcription of the gene for stearoyl CoA desaturase-1, which may explain the requirement for exogenous unsaturated fatty acids (Hasan et al., 1994; J. Pai, M. S. B., and J. L. G., unpublished data).

Hasan et al. (1994) took the first steps toward cloning the gene for the Site-2 protease (S2P) by transfecting human genomic DNA into M19 cells and isolating three primary transfectant clones that acquired the ability to grow in cholesterol-free media. They isolated genomic DNA from one of the clones, designated HfT1M19(c), and transfected it into fresh M19 cells and again selected for growth in sterol-free medium. This yielded seven secondary transfectant clones, designated HfT2M19(a-g), which regained the ability to induce LDL receptors and HMG CoA synthase when switched to lipid-depleted medium.

Using the technique of inter-Alu PCR, we here identify the human gene that complements the defect in HfT2M19(f) cells. Cloning and sequencing of the cDNA encoded by the S2P gene reveals that it encodes an unusual type of polytopic membrane protein that contains a consensus metalloprotease metal binding site of the HEXXXH type (Hooper, 1994; Rawlings and Barrett, 1995). We find that the S2P gene is deleted in M19 cells. When transfected into the mutant cells, S2P restores cleavage of SREBPs and relieves the growth requirement for exogenous cholesterol. Mutagenesis of the putative metal-binding site destroys the activity of S2P. The data suggest that S2P encodes a novel form of metalloprotease that is adapted to cleave proteins within transmembrane segments.

Results

Table 1 lists the steps that were used to isolate the human gene encoding S2P by complementation cloning in M19 cells. Through the use of primers complementary to repetitive human Alu sequences, we were able to identify three separate human DNA sequences that were present in the HfT2M19(f) cells (Table 1). One of these sequences was present in the third generation cell line (designated HfT3M19(a)), and this was used to isolate a PAC clone of human DNA that was subsequently shown to complement the defect in the M19 cells. Sequencing of this 100 kb PAC clone revealed one sequence encoding a previously undescribed homolog of a zinc-finger DNA binding protein and two additional sequences corresponding to two expressed sequence tags (ESTs) in the human database. The full-length cDNA corresponding to one of these two ESTs was able to restore sterol-independent growth in M19 cells. The plasmid that contained this cDNA was named pS2P.

Plasmid pS2P contained a cDNA insert with an open reading frame corresponding to S19 amino acids (Figure 1). The human cDNA did not contain an in-frame terminator codon upstream of the putative initiator methionine, but we subsequently isolated a hamster cDNA encoding S2P, and this cDNA contained an in-frame terminator codon 72 nucleotides upstream of the initiator methionine. We therefore believe that the sequence shown in Figure 1 corresponds to the full length of the protein.

The most important feature of S2P is the sequence HEIGH that begins at residue 171 of the human sequence (Figure 1). This corresponds to the consensus sequence (HEXXH), which has been found in 15 of 30

<table>
<thead>
<tr>
<th>Table 1. Steps in Isolation of Human Gene Encoding Site-2 Protease (S2P) by Complementation in M19 Cells</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Mutagenize CHO-K1 cells by γ irradiation, select for amphotericin B resistance, and isolate cholesterol auxotrophs, designated M19 cells (Hasan et al., 1994).</td>
</tr>
<tr>
<td>2. Transfect total human fibroblast genomic DNA into M19 cells and select clones of transfected M19 cells that grow in lipid-deficient medium (Hasan et al., 1994).</td>
</tr>
<tr>
<td>3. Transfect genomic DNA from one of the three primary transfectant clones, clone HfT1M19(c), into M19 cells and select for growth in lipid-deficient medium (Hasan et al., 1994).</td>
</tr>
<tr>
<td>4. Transfect genomic DNA from one of seven secondary transfectant clones, clone HfT2M19(f), into 2 x 10⁶ M19 cells and select for growth in sterol-deficient medium. Retain tertiary transfectant clone, designated HfT3M19(a).</td>
</tr>
<tr>
<td>5. Perform inter-Alu PCR on genomic DNA from HfT2M19(f) cells and isolate three PCR products, designated Probes 1-3.</td>
</tr>
<tr>
<td>6. Screen genomic DNA from clone HfT3M19(a) with Probes 1-3 by PCR. HfT3M19(a) genome contains sequence corresponding to Probe 1, but not Probes 2 and 3.</td>
</tr>
<tr>
<td>7. Screen human genomic PAC library with Probe 1 and isolate two positive clones (127A16 and 167B14) out of 144,000.</td>
</tr>
<tr>
<td>8. Transfect DNA from PAC clone 127A16 into M19 cells and select for growth in sterol-deficient medium.</td>
</tr>
<tr>
<td>9. Screen genomic DNA from seven surviving clones with Probe 1 (six positive clones, 1 negative clone).</td>
</tr>
<tr>
<td>10. Sequence PAC 127A16 and compare with database sequences to identify ESTs.</td>
</tr>
<tr>
<td>11. Use one identified EST to screen for full-length S2P cDNA from human HeLa cell library.</td>
</tr>
<tr>
<td>12. Sequence S2P cDNA.</td>
</tr>
</tbody>
</table>
Intramembrane Cleavage of SREBPs

Figure 2. Hydropathy Plots of Human S2P (A) and Human Endothelin Converting Enzyme-1 (B)

The residue-specific hydrophathy index was calculated over a window of 20 residues by the method of Kyte and Doolittle (1982), using the Genetics Computer Group Sequence Analysis Software Package, Version 8.1 (Devereux et al., 1984). Arrows denote potential N-linked glycosylation sites. The asterisk (*) denotes the sequence corresponding to the consensus pentapeptide metal binding site for zinc metalloproteases (Rawlings and Barrett, 1995).

Figure 1. Amino Acid Sequences of Human and Hamster S2P and Comparison with Related Sequences from Other Organisms

Amino acid numbers are indicated for those S2P sequences whose complete opening reading frames are known. Residues identical to the human sequence are indicated by closed boxes. The overbar denotes the consensus pentapeptide metal binding site (HEXXH) for zinc metalloproteases (Rawlings and Barrett, 1995). Cysteine residues in the cysteine-rich region are denoted by dots below the sequence. GenBank accession numbers for human and hamster S2P are AF019612 and AF019611, respectively. The sequences for C. elegans (GenBank number C11831), D. melanogaster (GenBank number AA391707), S. solfataricus P2 (Sensen et al., 1996; GenBank number Y08257, ORF C04034), and S. mansoni (GenBank number N33645) were obtained from the indicated source.

A hydropathy plot of human S2P revealed extensive hydrophobicity, especially over the NH₂-terminus, which includes the HEIGH sequence (Figure 2A). This region contains up to four or five putative transmembrane sequences, including a strongly hydrophobic sequence at the extreme NH₂-terminus. After the putative transmembrane region located at amino acids 210–230, the protein becomes relatively hydrophilic until the COOH-terminal region, which is again hydrophobic. The hydropathy profile of S2P is unlike that of any previously reported metalloprotease, and it raises the possibility identity over 119 residues); a flatworm, Schistosoma mansoni (27% identity over 117 residues); and an archaean, Sulfolobus solfataricus P2 (25% identity over 366 residues). All of these proteins share the HEXXH consensus with the exception of the partial sequence from S. mansoni, which does not extend into this region.

Another conspicuous feature of human S2P is a stretch of 23 contiguous serine residues beginning at amino acid 114 (Figure 1). This polyserine stretch is shorter in the hamster sequence (17 serines). It is much shorter in the Drosophila sequence, and it is precisely deleted from the Sulfolobus sequence. The available sequence from C. elegans does not extend to this region. Human and hamster S2P also contain a cysteine-rich sequence (residues 285–377 of the human sequence). Of these 92 residues, 10 are cysteines. This sequence presence in the other species cannot be evaluated because of insufficient data. We note that cysteine-rich sequences are present in other families of metalloproteases, although none of them show direct sequence resemblance to this region of S2P (Hooper, 1994; Rawlings and Barrett, 1995).

A hydropathy plot of Human S2P revealed extensive hydrophobicity, especially over the NH₂-terminus, which includes the HEIGH sequence (Figure 2A). This region contains up to four or five putative transmembrane sequences, including a strongly hydrophobic sequence at the extreme NH₂-terminus. After the putative transmembrane region located at amino acids 210–230, the protein becomes relatively hydrophilic until the COOH-terminal region, which is again hydrophobic. The hydrophathy profile of S2P is unlike that of any previously reported metalloprotease, and it raises the possibility
that S2P and its relatives constitute a new family of these enzymes (see Discussion).

For comparative purposes, Figure 2B shows a hydropathy plot for endothelin converting enzyme-1 (ECE-1), a membrane-bound metallopeptase that is found in the Golgi apparatus (Xu et al., 1994). In contrast to the S2P, ECE-1 contains only a single hydrophobic segment, and the rest of the protein is quite hydrophilic. Similar results are obtained in the hydropathy plots for other membrane-bound human metallopeptases, including alanine aminopeptidase N, nephrilysin, and the matrix metallopeptase-14 precursor (Rawlings and Barrett, 1995; Will and Hinzmann, 1995).

Hydropathy plots of the S2P-related proteins from the other organisms shown in Figure 1 tended to show hydrophobicities that resemble those of the human and hamster proteins (data not shown). A detailed comparison of these hydrophobicities awaits the availability of full-length sequences.

To confirm that the M19 cells harbor a mutation in the S2P gene, we performed genomic Southern blots (Figure 3). DNA from wild-type CHO-7 cells showed evidence for a single S2P gene that was detected with probes from the 5' region and the 3' region of the hamster S2P gene. In contrast, DNA from the M19 cells failed to hybridize with probes from either end of the gene. The M19 cell DNA did hybridize with probes for an unlinked gene, ACAT (right panel of Figure 3), demonstrating the integrity of the M19 genomic DNA preparation. Reverse transcriptase-PCR analysis showed that wild-type hamster cells express an S2P mRNA whose cDNA is 4.4 kb in length, which is consistent with the size of the human cDNA clone. As expected, the same PCR primers failed to amplify any S2P mRNA in extracts of M19 cells (data not shown). We conclude from these experiments that the M19 cells have a deletion that encompasses the entire S2P gene.

Genomic mapping by fluorescence in situ hybridization (FISH) revealed that the human S2P gene is located on the X chromosome (data not shown). Consistent with this observation was the finding that PAC clone 127A16, which encodes S2P, contained a sequence-tagged site (STS) that was reported to be X-linked (WI-6111, GenBank accession number G03877).

We obtained a commercially prepared filter containing poly(A)^+ RNA from a variety of human tissues. When this filter was probed with the human 32P-labeled S2P cDNA, it revealed an mRNA of 4.4 kb that was present in every organ studied (Figure 4).

To show directly that S2P is capable of restoring Site-2 cleavage in M19 cells, we performed a cotransfection experiment (Figure 5) in which M19 cells were transiently transfected with a cDNA encoding SREBP-2 that was tagged at its NH2 terminus with an epitope derived from the herpes simplex virus (HSV) envelope glycoprotein D (Sakai et al., 1996). Membranes and nuclear extracts were isolated and blotted with an antibody against the HSV epitope tag. In the absence of cotransfected S2P, the M19 cells cleaved SREBP-2 to generate the membrane-bound intermediate form (lane 3, upper panel). There was no mature form of SREBP in the nuclear extract (lane 3, lower panel). When steroids were added to the cells, the intermediate disappeared, indicating that it had been produced by the sterol-regulated Site-1 cleavage enzyme (lane 4, upper panel).

To measure the activity of S2P, we cotransfected M19 cells with the HSV-tagged SREBP-2 plasmid plus a plasmid encoding Myc-tagged S2P under control of the CMV promoter. Under these conditions, the intermediate form of SREBP-2 disappeared (upper panel of Figure 5, lane 5), and the mature nuclear form of SREBP-2 appeared (lower panel of Figure 5, lane 5). We also transfected a cDNA encoding the R519A mutant of SREBP-2. This mutation destroys the recognition site for the Site-1 protease (Sakai et al., 1996; Duncan et al., 1997). The R519A mutant SREBP-2 did not give rise to an intermediate form in the absence of pS2P (upper panel, lane 7), nor did it produce the nuclear form in the presence of pS2P (lower panel, lane 7). These data indicate that the cloned...
Intramembrane Cleavage of SREBPs

pS2P will activate cleavage of SREBP-2 only after it is cleaved by the Site-1 protease.

A distinguishing characteristic of the Site-2 protease is its requirement for the D478RSR pattern that is unique to the transmembrane region of SREBPs (Sakai et al., 1996). When we transfected the D478RSR→AS mutant of SREBP-2 into M19 cells, it generated the expected intermediate form, indicating that Site-1 cleavage had occurred (upper panel of Figure 5, lane 11). However, pS2P no longer caused the intermediate to disappear (upper panel, lane 13), nor did it cause the nuclear form to appear (lower panel, lane 13). Thus, the action of the protein encoded by pS2P, like the endogenous Site-2 protease, requires the DRSR sequence.

To further study the activity of pS2P, we used a luciferase reporter assay that monitors the amounts of SREBP-2 in cell nuclei (Figure 6). M19 cells were transfected with a plasmid encoding luciferase driven by a sterol-independent promoter. In the absence of cotransfected pS2P, the SRE-luciferase construct produced low levels of luciferase activity in M19 cells either in the absence or presence of sterols (empty vector). When wild-type pS2P was cotransfected, luciferase activity was high in the absence of sterols (closed bars in Figure 6) and was reduced to basal level in the presence of sterols (stippled bars). These data are consistent with the previous data showing that wild-type pS2P led to cleavage of the intermediate forms of endogenous SREBPs and that the mature nuclear forms activated transcription of the SRE-1 containing promoter.

We also used the luciferase assay to analyze the activity of mutant forms of S2P (Figure 6). For this purpose, we produced a series of plasmids encoding mutant forms of S2P in which residues in the consensus metal-binding site were replaced. Replacement of either of the histidine residues of the HEIGH sequence or alanine abolished the ability of S2P to stimulate transcription of the SRE-luciferase reporter. These data strongly suggest that the activity of S2P depends on the consensus metalloprotease metal-binding site.

In many HEXXH metalloproteases, the glutamic acid of the consensus sequence is believed to activate a water molecule that is used to cleave the peptide bond (Rawlings and Barrett, 1995). Replacement of this glutamic acid by site-directed mutagenesis with other amino acids, such as alanine or glutamine, abolishes protease activity. In three metalloproteases in which the glutamic acid was changed to aspartic acid, activity was
Figure 7. Growth of M19 Cells Transfected with cDNA Encoding Wild-Type or Mutant S2P.

On day 0, M19 cells were set up at $3.5 \times 10^5$ cells/60 mm dish in medium B (containing cholesterol, mevalonate, and oleate) as described in Experimental Procedures. On day 1, the cells were transfected with 5 µg/dish of the CMV-driven pcDNA3 vector containing the indicated human S2P cDNA using the MBS Transfection Kit (Stratagene). The cells were washed once with PBS and refed with medium B 3 hr after transfection. On day 2 and every 1-2 days thereafter, the cells were refed with fresh medium B containing 750 µg/ml G418. On day 14, the cells were divided into two groups and received either medium A (cholesterol-free) containing 500 µg/ml G418 or medium B (containing cholesterol, mevalonate, and oleate) containing 500 µg/ml G418. Fresh medium A or medium B containing G418 was added every 1-3 days. On day 28, the cells were washed with PBS, fixed in 95% ethanol, and stained with crystal violet.

Discussion

In the current studies, we use a cholesterol auxotrophic mutant line of CHO cells as a recipient for complementation cloning of a gene encoding a protein that is required for intramembranous proteolysis of SREBPs at Site-2. The mutant cell line, designated M19, was derived by Hasan et al. (1994), who showed that it has a global regulatory defect, i.e., it cannot increase transcription of genes encoding several sterol biosynthetic enzymes discussed below.

In four of the 15 families of HEXXH-containing metalloproteases, a glutamic acid residue located about 25 residues to the COOH-terminal side of the HEXXH sequence functions as an additional metal coordinating residue (Rawlings and Barrett, 1995). S2P has an aspartic acid at position 203 of the human sequence (see Figure 1). M19 cells do not revert to the wild-type phenotype (Hasan, 1993). Taking advantage of this property, Hasan et al. (1994) transfected human genomic DNA into M19 cells and derived two generations of transfectants whose restoration to wild-type could be explained by the acquisition of a human gene that complements the defect. The gene was then identified in the current studies by inter-Alu PCR, PAC cloning, and large-scale genomic sequencing.

The gene that we designated S2P encodes a novel protein with two outstanding characteristics: (1) it contains the HEXXH consensus sequence that constitutes the metal-binding site in about half of the 30 known families of metalloproteases (Rawlings and Barrett, 1995); and (2) it differs from known metalloproteases in its widespread hydrophobicity, which suggests the existence of several membrane-spanning segments.

Several lines of evidence are consistent with the notion that the product of the S2P gene is indeed the Site-2 protease: (1) the gene is deleted in M19 cells that lack Site-2 protease activity; (2) when expressed in M19 cells by transfection, the cDNA restored cleavage of cotransfected SREBP-2 at Site-2 in a fashion that depended on cholesterol.
upon prior cleavage at Site-1 and upon an intact DRSR recognition sequence; (3) when expressed in M19 cells by transfection, the S2P cDNA restored sterol-regulated transcription activity to genes whose promoters contain an SRE-1; (4) the transfected cDNA restored sterol-independent growth to M19 cells; and (5) all activities of the cDNA were abolished when either of the histidines of the consensus metal-binding site were replaced with phenylalanines. Nevertheless, it remains possible that the protein encoded by S2P is not a protease, but rather a protein that is necessary to activate a protease.

A large number of metalloproteases are known, and these have been divided into 30 families (Hooper, 1994; Rawlings and Barrett, 1995). Approximately half of these families are distinguished by the HEXXH consensus metal-binding site. This site has been defined more extensively as abxHEbbHbc in which a is most commonly valine or threonine; b is an uncharged residue; c is hydrophobic; and x can be any amino acid except proline (Rawlings and Barrett, 1995). The corresponding sequence in human and hamster S2P is GVVHEIGHGI, which matches the extended consensus precisely except that the a residue is glycine instead of valine or threonine. In S2P, two valines occur in the adjacent positions, corresponding to the bx positions in the extended consensus. A list of HEXXXH metalloproteases compiled by Rawlings and Barrett (1995) includes several examples in which a is alanine, but none where it is glycine. Nonetheless, given the enormous variability in the sequence of metalloproteases, it is not difficult to imagine a glycine in this position, especially since S2P is postulated to be a novel metalloprotease that cuts its substrate in an unusual environment, namely, within a transmembrane segment.

The observation that alanine and asparagine cannot replace the glutamic acid of the HEXXXH consensus is consistent with the hypothesis that the COOH group is needed to activate a water molecule in S2P as it does in known metalloproteases. The only result at variance with the interpretation that S2P is metalloprotease comes from the observation that this glutamic acid can be replaced by aspartic acid with preservation of at least some catalytic activity as defined by the ability of the cDNA to correct the metabolic defect in M19 cells. Although aspartic acid theoretically could perform the same water-activating function as glutamic acid, prior studies of three other metalloproteases have shown that this substitution destroys most if not all of the catalytic activity, apparently because the aspartic acid cannot function within the geometrical constraints of the active site (Rawlings and Barrett, 1995).

Each of the HEXXXH families of metalloprotease has one or more distinct signature sequences in addition to the HEXXXH. Often, these sequences include another residue that comes into proximity with the HEXXXH sequence where it contributes an additional coordinating site for zinc. In terms of the linear sequence, these coordinating residues can be located at widely varying distances from the HEXXXH. The coordinating residues also vary. Histidine, tyrosine, and glutamic acid residues are among the most common. In some metalloproteases, an additional coordinating residue cannot be found (Rawlings and Barrett, 1995). The S2P sequence does not include any of the known signature sequences of the metalloprotease families. Again, this is not surprising since S2P is a member of a novel hydrophobic family of proteins.

Taken together, the above considerations are consistent with the notion that S2P is a metalloprotease. A definitive confirmation of this hypothesis will require the purification of the protein and the demonstration that it can cleave its substrate, the intermediate form of an SREBP. This will be a difficult task inasmuch as the purification of polytopic membrane-bound enzymes is exceedingly difficult, even when the proteins are overexpressed through use of recombinant DNA techniques. The problem is accentuated in this case because the substrate, the intermediate form of SREBP, is also an intrinsic membrane protein, and it is likely that it must be presented to the enzyme in its membrane-bound state. The studies would be facilitated by the synthesis of a short peptide that would be cleaved by S2P.

Based on the sequence comparisons in Figure 1, S2P defines a conserved family of hydrophobic membrane proteins that includes examples from archaea, insects, flatworms, and roundworms. All of these sequences were deposited in databases as a result of random DNA sequencing, and none of the proteins have been shown to be a protease. Nevertheless, their striking resemblance to S2P defines them as members of a common family, all of which are likely to be metalloproteases. We note that such evolutionary conservation is not unusual among metalloproteases. Several of the known families include relatives from bacteria and even archebacteria (Rawlings and Barrett, 1995). This extensive conservation of S2P is actually an additional argument that it is a metalloprotease.

If the S2P relatives in other species are proteases, the question arises as to the nature of their substrates. The M19 cells do not have any obvious defect other than auxotrophy for cholesterol and unsaturated fatty acids, suggesting that hamster S2P is not essential for any function other than processing of SREBPs. The same may be true in Drosophila, which is known to produce an SREBP whose mode of processing and target genes are unknown (Theopold et al., 1996). Whether C. elegans has SREBPs is presently unknown.

Interestingly, the complete genomic sequence of the yeast Saccharomyces cerevisiae fails to reveal a member of the S2P family. This organism does, however, have an unrelated hydrophobic metalloprotease with multiple membrane-spanning regions. Encoded by the STE24 gene and designated Ste24p, this protease catalyzes the first step in the NH₂-terminal processing of a-factor, a unique type of secreted peptide (Fujimura-Kamada et al., 1997). NH₂-terminal processing of a-factor occurs after COOH-terminal processing, which includes farnesylation, proteolysis, and carboxymethylation. The precise cellular localization of Ste24p is unknown, but it contains a COOH-terminal KXX endoplasmic reticulum retrieval sequence, and therefore, it is suspected to be bound to membranes of the ER. Sequences related to the Ste24p protease were found in humans and in several bacterial species, thereby defining a metalloprotease family (Fujimura-Kamada et al., 1997). The members of the S2P family bear no direct sequence resemblance to

Intramembrane Cleavage of SREBPs

53
Ste24p, nor do they have any of the conserved motifs that define the Ste24p family except for the HEXXXH consensus metal-binding site. Thus, the Ste24p family and the S2P family are separate families of membranous metalloproteases, each of which is evolutionarily conserved.

The current evidence that cleavage of Site-2 of SREBP is within a transmembrane sequence is indirect. It is based on careful estimates of the size of the NH2-terminal fragment that was generated by the Site-2 protease in intact cells transfected with a cDNA encoding a version of SREBP-2 in which the NH2-terminal segment of SREBP-2 (Sakai et al., 1996). To date, Site-2 has not been identified by direct chemical analysis.

The only protein that has been shown definitively to undergo cleavage within a transmembrane segment is the amyloid precursor protein (APP) (Selkoe, 1996). Intramembranous cleavage of APP is carried out by proteases called γ-secretases that cleave at either Sites to generate peptides of 40 or 42 amino acids. In the brain, either of these peptides can form aggregates that lead to the toxic amyloid deposits of Alzheimer’s disease, but the 42 aa peptide is by far the more potent. Recent evidence indicates that the protease that gives rise to the 40 aa amyloid peptide is inhibited by calpain inhibitors (Yamazaki et al., 1997), which would not be expected to inhibit a metalloprotease such as S2P. The nature of the protease that creates the 42 aa peptide is not yet established, and it remains possible that this protease is S2P or one of its relatives.

Experimental Procedures

General Methods

Standard molecular biology techniques were used (Sambrook et al., 1989). Genomic DNA was prepared using the DNA Extraction Kit (Stratagene). The average fragment size of DNA prepared by this method was ~100 kb as determined by pulsed-field gel electrophoresis. Pools of PAC clones containing human genomic DNA and individual PAC clones 127A16 and 167B14 were obtained from Genome Systems, Inc. (St. Louis, MO). DNA from PAC clones was isolated by standard CsCl density gradient centrifugation (Sambrook et al., 1989). Plasmids for transfection were prepared using purification kits (Promega or Qiagen). Lipoprotein-deficient serum (Metherall et al., 1989) and delipidated serum (Hasan et al., 1994). The first round of selection for growth in delipidated serum was repeated. The resultant secondary transfectant cell lines were designated HfT2M19(a–g) cells (Hasan et al., 1994).

Cells were maintained in monolayer culture at 37°C in 9% CO2. Stock cultures of CHO-7 cells, a subline of CHO-K1 cells selected for growth in lipoprotein-deficient serum (Metherall et al., 1989), and HfT2M19(f) cells (see Table 1) were maintained in medium A (1:1 mixture of Ham’s F12 medium and Dulbecco’s modified Eagle’s medium containing 100 U/ml of penicillin, 100 µg/ml streptomycin sulfate, and 5% (v/v) lipoprotein-deficient fetal calf serum). Stock cultures of M19 cells were maintained in medium B (1:1 mixture of Ham’s F12 medium and Dulbecco’s modified Eagle’s medium containing 100 U/ml of penicillin, 100 µg/ml streptomycin sulfate, and 5% (v/v) fetal calf serum, 5 µg/ml cholesterol, 1 mM mevalonate, and 20 µM oleate).

Transfections with HfT2M19 Genomic DNA or PAC Clones

Recipient cells were set up on day 0 at 1 x 10^4 cells per 100 mm dishes in their respective growth medium. On day 1, cells were transfected with either 12.5 µg of genomic DNA/dish or 6.25 µg of PAC DNA/dish using the MBS Transfection Kit (Stratagene). Cells were incubated for 3 hr at 35°C, 3% CO2 in a 1:1 mixture of Ham’s F12 medium and Dulbecco’s modified Eagle’s medium containing 100 U/ml of penicillin, 100 µg/ml streptomycin sulfate, and 6% modified bovine serum (Stratagene). Following incubation, cells were washed once with phosphate-buffered saline, switched to medium B, and returned to 37°C, 9% CO2. On day 2, cells were washed and refed with selection medium C (medium A supplemented with 2 µM compactin and 20 µM oleate). Cells were refed with selection medium C every 2–3 days. On days 21–28, surviving colonies were cloned by limiting dilution. Individual colonies were tested for sterol-mediated regulation by growth in medium A alone or supplemented with 0.5 µg/ml 25-hydroxycholesterol.

PCR Probes

Inter-Alu PCR was carried out as described by Nelson et al. (1989). A primer based on consensus human Alu sequences was used to amplify unique human sequences lying between Alu elements in the genomic DNA prepared from transfected hamster HfT2M19(f) cells (Table 1). The primer used, 5’-GCTGATTCTACGTACCTAGCTGGGCCAC-3’, consisted of the A33 primer described by Chumakov et al. (1992) flanked by an EcoRI site to facilitate cloning of the PCR products. PCR products were cloned and sequenced, and three unique inter-Alu sequences were identified. These were used to design three pairs of specific primers for use as PCR probes: probe pair 1: 5’-TGACACCCCTGGAGGTGGTACG-3’ and 5’-ATCCCCAACCTACCTAGCCCA-3’; probe pair 2, 5’-TTCATGTTGTGTTATGATTCTGC-3’ and 5’-GAACATTCTAACGCTGACCAG-3’; and probe pair 3, 5’-GCCAAAGCTGTGTTCACTAGCCCA-3’ and 5’-GTTACAAACCGTTGGAAAATG-3’.

DNA Sequencing

Genomic DNA and cDNA clones were sequenced by the dyeodeoxy chain termination method (Sanger et al., 1980) using vector and insert-specific primers. Sequencing reactions were performed on Applied Biosystems Model 373A and 377 DNA sequencers.

DNA from PAC clone 127A16 was sheared by sonication, blunted with the Klenow fragment of DNA polymerase (New England Biolabs), and size-fractionated by agarose gel electrophoresis. Fragments of ~1 kb in length were excised from the gel and recovered by Gene Clean (Bio101, Inc.). The fragments were then cloned into the pHoTA vector (5 Prime–3 Prime, Inc.), and random clones were sequenced using vector-specific primers. Approximately 85% of the 100 kb genomic DNA insert was sequenced. Contiguous sequences were assembled using both DNAstar software (Lasergene) and PHRED/PHRAP (Philip Green, University of Washington, Seattle, WA) and compared to the NCBI database using BLAST software (Altschul et al., 1990). Clones corresponding to the ESTs identified by this search strategy were obtained from Research Genetics, Inc., Huntsville, AL and sequenced in their entirety.
cDNA Cloning of Human and Hamster S2P

A human HeLa cDNA library prepared in the λEXlOx vector (Yokoyama et al., 1993) was screened with the 2.4 kb insert from the mouse cDNA clone 34244 (Research Genetics, Inc.), which was released by digestion with NotI/HinDIII. Duplicate filters were hybridized in RapidHyb buffer (Amersham) with 2 × 10^6 cpm/ml of 32P-labeled probe. Of 2 × 10^6 plaques screened, ~120 positive clones were identified, and 12 of these were selected for further analysis. Of these 12 clones, 6 were found to contain full-length inserts based on an insert length of 4.4 kb. These six clones were sequenced using a variety of vector-specific and insert-specific primers, and the resulting data were compiled and analyzed using the DNAStar software package (Lasergene). The resulting cDNA is designated pS2P.

Comparison of the human S2P sequence with mouse ESTs from the NCBI DBEST Database identified highly conserved nucleotide sequences within the 3'-untranslated region of human S2P. These sequences were used to design PCR primers to amplify cDNA reverse-transcribed from total RNA prepared from hamster CHO-7 cells. The resulting PCR product was used to screen a Chinese hamster λEXlOx cDNA library prepared from CHO-7 cells (Hua et al., 1996a). Approximately 2 × 10^6 plaques were screened, and 75 positive clones were identified. Positive clones were recovered, sequenced, and analyzed as described above, and used to screen two additional Chinese hamster λgt22A libraries prepared from SRD-1 cells (Yang et al., 1994) and SRD-3 cells (Yang et al., 1995). The most 5' and 3' cDNA sequences were used to obtain a full-length cDNA sequence by long PCR (Barnes, 1994) of cDNA reverse-transcribed from total RNA prepared from CHO-K1 cells.

Construction of Expression Plasmids

The expression vector pCMV-HSV-S2P encodes amino acids 2-519 of human S2P preceded by an initiator methionine, two tandem copies of the herpes simplex virus (HSV) epitope (QELAPEDPED), and two novel amino acids (ID) encoded by a sequence for the BspDI restriction site. pCMV-HSV-S2P was constructed as follows. First, pCMV-HSV-BP2 (Hua et al., 1996a) was digested with BspDI and XbaI, and the 6.0 kb XbaI-BspDI fragment containing the HSV epitopes was isolated. Second, the human S2P cDNA was cloned into the CMV-driven pDNA3 vector (Invitrogen) at the NotI site. Third, the sequence corresponding to amino acids 2-519 of human S2P in the above vector was amplified by PCR with an NH2-terminal primer flanked by a BspDI site and a COOH-terminal primer flanked by a nonsense codon and an XbaI site. The resulting 1.5 kb BspDI-XbaI fragment was ligated to the above 6.0 kb XbaI-BspDI fragment to generate pCMV-HSV-S2P.

pCMV-Myc-S2P encodes amino acids 2-519 of human S2P preceded by an initiator methionine, two tandem copies of the c-Myc epitope (EOKLISEEDLN), and two novel amino acids (ID) encoded by a sequence for the BspDI restriction site. pCMV-Myc-S2P was constructed as follows. First, a pair of complementary oligonucleotides, (top strand) 5' CTAGCCAATCCAAAAGGTCATCTTCTCAG GAAGACTTACTCTTCTCAGTAT CTTGGG-3', which contained 5' Nhel and 3' BspDI cohesive ends, were annealed as described previously (Hua et al., 1996b). The oligonucleotide encoded the initiator methionine and two tandem copies of the c-Myc epitope. pCMV-HSV-BP2 (SS5-1141) (Sakai et al., 1997) was digested with Nhel and BspDI, and the 90 bp fragment encoding two tandem copies of the HSV epitope was replaced with the two tandem copies of the c-Myc epitope to generate pCMV-Myc-BP2 (SS5-1141). Second, pCMV-Myc-BP2 (SS5-1141) was digested with BspDI and XbaI, and the 6.0 kb XbaI-BspDI fragment containing the c-Myc epitope was isolated. Third, pCMV-HSV-S2P was digested with XbaI and BspDI, and the resulting 1.5 kb BspDI-XbaI fragment was ligated to the above 6.0 kb XbaI-BspDI fragment to generate pCMV-Myc-S2P.

Blot Hybridization of Genomic DNA

Genomic DNA (20 μg) from CHO-7 and M19 cells was digested with the indicated restriction enzymes, subjected to electrophoresis in an 0.8% agarose gel, and transferred to Hybond N+ filters (Amersham) by capillary blotting. After transfer, the filters were cross-linked by UV irradiation at 120 mJ in a Stratalinker (Stratagene). Filters were then prehybridized with RapidHyb buffer (Amersham) at 65°C for 30 min and blotted with 32P-labeled probes for 2 hr at 65°C. Probes 1 and 2 (see below) were generated by PCR reactions using hamster genomic DNA as a template and (32P)MCTP. Probe 1 (~2.9 kb), corresponding to exons 4 and 5, and intron 4 of the hamster S2P gene, was amplified with the 5' primer, 5'-GTCAATACG CTTAGCTTCTTCTCCTGGC-3' and the 3' primer, 5'-AAATATTCC TTAATGCTGGACGTGTGTT-3'; Probe 2 (1.1 kb), corresponding to part of exon 11 and part of the 3'-untranslated region of hamster S2P mRNA, was amplified by PCR using the cDNA sequences 5'-GTCCTGGAAATAAAGGCTGTG-3' and the 3' primer, 5'-CAAGTAAAATTTGTGA ATATTT-3'; and Probe 3 (3.0 kb), corresponding to the entire cDNA of hamster ACAT, was digested from the vector pRC-CMV758 using SalI and NotI (Cao et al., 1996) and labeled with 32P by random priming.

cDNA Transfection and Immunoblot Analysis

On day 0, M19 cells were set up at a density of 7 × 10^5 cells/60 mm dish in medium B. On day 1, cells were transfected with 4 μg DNA/dish using Lipofectamin Reagent (Life Technologies) according to the manufacturer's instructions with modifications as follows. We used 4 μg of DNA and 12 μl of Lipofectamin per dish. For duplicate dishes, Lipofectamin/DNA complexes were formed in 0.4 ml of serum-free medium for 15 min at room temperature, and 0.2 ml were added to 1.8 ml of serum-free medium per dish. After incubation for 3 h at 37°C in a 8%-9% CO2 incubator, the existing 2.9 μg of medium in each dish was supplemented with 2 ml of medium A adjusted to give the following final concentrations: 5% newborn lipoprotein-deficient serum, 50 μg/ml of 25-hydroxycholesterol plus 10 μg/ml of cholesterol. After incubation for 20 hr at 37°C, the cells received N-acetyl-leucinyl-leucinyl-norleucine to achieve a final concentration of 25 μg/ml (Wang et al., 1994). After incubation for 3 hr, the cells were harvested and fractionated into a nuclear extract and 10 μg membrane pellet as described (Sakai et al., 1996).

Immunoblot analysis was carried out as previously described (Hua et al., 1996a) using either the Enhanced Chemiluminescense Western Blotting Detection Kit (Amersham) or the SuperSignal Substrate (Pierce). Proteins were blotted with IgG-HSV.Tag, a monoclonal antibody directed against the glycoprotein D epitope of Herpes Simplex Virus (Novagen).

Luciferase Reporter Assay

On day 0, replicate wells of 1 × 10^5 M19 cells/22 mm well were plated in medium A supplemented with 20 μM oleate. On day 1, duplicate wells of cells were cotransfected by the calcium phosphate method with an M5S Kit (Stratagene) with the following plasmids: as previously described (Hua et al., 1996a): 1.0 μg of plasmid containing the indicated CDNA, 0.8 μg of a SRE-1-driven luciferase reporter construct (pSRE-Luciferase), and 0.05 μg of pCMVβ (a plasmid encoding β-galactosidase; Stratagene) in a final volume of 0.1 ml. Transcription of the luciferase reporter cDNA was driven by a promoter consisting of three tandem copies of repeat 2 + 3 of the LDL receptor promoter plus the adenosine E1b TATA box (Hua et al., 1996a). After incubation for 3 hr at 37°C in a 3% CO2 incubator, the cells were washed with PBS and fed with 1 ml of medium A supplemented with 50 μM compactin and 50 μM mevalonate in the presence or absence of steroids (1 μg/ml of 25-hydroxycholesterol plus 10 μg/ml of cholesterol added in 0.2% ethanol). After incubation at 37°C for 16 hr in a 8%-9% CO2 incubator, the cells were lysed with 0.2 ml of 1× Reporter Lysis Buffer (Promega), and aliquots were used for measurement of luciferase (40 μl) and β-galactosidase (40 μl) activities. Luciferase and β-galactosidase assays were carried out with substrates that generate chemiluminescent products, using kits from Promega and Clontech, respectively. Photons were detected in an Optima II Luminoimeter (MGM Instruments). The amount of luciferase activity in transfectedants was normalized to the amount of β-galactosidase activity to correct for transfection efficiency in each experiment.

Acknowledgments

We thank our colleague Roger Schultz for FISH mapping; Terry Franklin for excellent technical assistance with sequence compilation and comparison; Lisa Beatty and Seema Nair for invaluable assistance; Lisa Beatty and Seema Nair for invaluable
help with cell culture; Mark Daris for technical help with immunoblotting; and Jeff Cormier and Michelle Laremore for oligonucleotide synthesis and automated sequencing. This research was supported by grants from the National Institutes of Health (HL-20948) and the Perot Family Foundation. D. N. is supported by Medical Scientists Training Grant GM08014.

Received August 19, 1997; revised September 17, 1997.

References


Intramembrane Cleavage of SREBPs


GenBank Accession Numbers

The accession numbers for the human and hamster S2P cDNA sequences reported in this paper are AF019612 and AF019611, respectively.