Molecular Characterization and Tissue Distribution of a Novel Member of the S100 Family of EF-Hand Proteins[†,‡]

Alexey V. Gribenko, James E. Hopper, and George I. Makhatadze*

Department of Biochemistry and Molecular Biology, Penn State University College of Medicine, Hershey, Pennsylvania 17033

Received July 16, 2001; Revised Manuscript Received October 17, 2001

ABSTRACT: We have isolated from a human prostate cDNA library a cDNA encoding a novel member of the S100 family of EF-hand proteins. The encoded 99-amino acid protein, designated S100Z, is capable of interacting with another member of the family, S100P. S100Z cDNA was cloned into a bacterial expression system, and the S100Z protein was purified to homogeneity from bacterial lysates by a combination of hydrophobic column and gel-filtration chromatography. Direct amino acid sequencing of the 20 N-terminal amino acids confirmed that the sequence of the recombinant protein is identical to the sequence deduced from the cDNA. Low-resolution structural data have been obtained using circular dichroism and fluorescence spectroscopies, and equilibrium analytical centrifugation. These results show that S100Z is a dimeric, predominantly α-helical protein. Addition of calcium to a solution of S100Z changes the fluorescence intensity of the protein, indicating that S100Z is capable of binding calcium ions. Analysis of the calcium-binding isotherm indicates the existence of two calcium-binding sites with apparent affinities on the order of 5×10^6 and 10^2 M^{-1}. Binding of calcium results in conformational changes and exposure of hydrophobic surfaces on the protein. Using a PCR-based assay, we have detected differences in the expression level of S100Z mRNA in various tissues. The highest levels were found in spleen and leukocytes. S100Z gene expression appears to be deregulated in some tumor tissues, compared to expression in their normal counterparts.

S100 proteins represent a group of small proteins that contain two calcium-binding motifs (EF-hands) in their primary sequence (*1*, *2*). Unlike another typical calcium-binding protein, calmodulin, S100 proteins are characterized by tissue and cell type-specific expression and exhibit various metal binding properties (*3*). Although the exact biological functions of the S100 proteins remain to be determined, they seem to be associated with a wide variety of human diseases and tumors (*3*).

S100P is one of the least studied members of the S100 family, and it appears to be associated with prostate and breast cancer development (*4*, *5*). Our previous work has implicated S100P as a calcium-dependent conformational switch (*6*). We have also demonstrated that it is capable of binding a model peptide (melittin) in a calcium-dependent manner (*7*). To gain insight into the biological role of S100P in living cells, we set out to identify S100P binding partners. The method that we chose was the yeast two-hybrid screen (*8*).

S100P expression was shown to be differentially regulated in androgen-dependent and androgen-independent prostate cancer cell lines (*4*). On this basis, we chose to screen the human prostate cDNA library for possible S100P-interacting proteins. We identified a cDNA clone encoding a novel protein of 99 amino acids. The amino acid sequence of the protein is highly similar to the sequences of members of the S100 family. We have cloned the cDNA for this protein into a bacterial expression system and overproduced the recombinant protein in *Escherichia coli*. Using biophysical methods, we show that the recombinant protein possesses unique secondary and tertiary structure, is capable of binding two calcium ions per subunit with different affinity, forms noncovalent dimers in solution, and physically interacts with S100P in vitro. PCR-based assays demonstrate differences in the expression levels of the new gene in several human tissues, with the highest levels of expression found in leukocytes and spleen. Expression of the gene also appears to be aberrant in certain human tumors.

MATERIALS AND METHODS

Construction of Plasmids for the Yeast Two-Hybrid Screen

DNA-Binding Domain Plasmid Expressing the GAL4 DNA-Binding Domain (BD)−S100P Fusion Protein. The S100P gene (*9*) was amplified by PCR using forward (AVGR01, Table 1) and reverse (AVGR02, Table 1) primers, generating unique *Eco*RI and *Bam*HI restriction sites for directional cloning into the first multiple cloning site of the pBridge vector (Clontech, Palo Alto, CA). The PCR product was purified using the QIAquick PCR purification kit

[†] This work was supported in part by grants from the National Institutes of Health (G.I.M.), Research Corp. (G.I.M.), the South Plains Foundation (G.I.M.), and the Penn State Cancer Center (G.I.M. and J.E.H.).

[‡] The nucleotide sequence reported in this paper has been submitted to GenBank with accession number AF437876.

* To whom correspondence should be addressed: Department of Biochemistry and Molecular Biology, Penn State University College of Medicine, 500 University Dr., Hershey, PA 17033. Phone: (717) 531-0712. Fax: (717) 531-7072. E-mail: makhatadze@psu.edu.

10.1021/bi0114731 CCC: $20.00
Published on Web 11/29/2001

Table 1: Sequences of the Primers Used in This Study

primer	sequence
AVGR01	5′-CGAATTCATGACGGAACTAGAG
AVGR02	5′-CCCGGATCCTCATTTGAGTCCTGT
AVGR03	5′-GAGACAGCATAGAATAAGTGC
AVGR04	5′-ATAAGAATGCGGCCGCATGACGGAACTAGAG
AVGR05	5′-GAAGATCTTCATTTGAGTCCTGTCTT
AVGR06	5′-ATGGGCCATATGGCTTCTAGC
AVGR07	5′-TTTGGAATCACTACAGGGATG
AVGR08	5′-CGCCGGATCCGCATGCCCACCCAGCTCGAG-ATGGCCATGGAC
AVGR09	5′-GCCAATGCAGGATCCCTGCAGTATTCCAGG-CAGACCCCGTCCTGAGAGCGA
AVGR10	5′-ATGCCCACCCAGCTCGAGATGGCCATGGAC
AVGR11	5′-CTCAAGCAGCTTTCTCCAAAGCCCCCAAGC

(Qiagen), digested with *Eco*RI and *Bam*HI restriction endonucleases, and ligated into pBridge linearized with *Eco*RI and *Bam*HI. The BL21(DE3) *E. coli* strain was transformed with the ligation reaction using an *E. coli* pulser (Bio-Rad) and plated on 2×YT agar plates containing 100 μg/mL ampicillin. Plasmid DNA containing a *Sac*I restriction site unique to the S100P gene was sequenced on an ABI 100 genetic analyzer using the pBridge sequencing primer (AVGR03, Table 1). The resulting construct was designated pBridge-S100P$_1$.[1]

DNA-Binding Domain Plasmid Expressing a Second Copy of the S100P Gene. To allow for the formation of the S100P heterodimer between the *GAL4* DNA-binding domain−S100P fusion protein and the native S100P, we cloned another S100P cDNA copy into the multiple cloning site II of pBridge. The complete S100P ORF was amplified by PCR using forward (AVGR04, Table 1) and reverse (AVGR05, Table 1) primers generating *Not*I and *Bgl*II restriction sites, respectively. The PCR product was purified with the QIAquick PCR purification kit (Qiagen), digested with *Not*I and *Bgl*II restriction endonucleases, and ligated into the pBridge-S100P$_1$ plasmid linearized with *Not*I and *Bgl*II. The ligation reaction was used to transform the BL21(DE3) *E. coli* strain, and transformants were selected for ampicillin resistance on 2×YT agar plates containing 100 μg/mL ampicillin. Plasmids isolated from several ampicillin-resistant colonies were digested by *Sac*I restriction endonuclease. Plasmids that produced two fragments upon digestion with *Sac*I (i.e., potentially carrying two copies of S100P) were sequenced on an ABI 100 genetic analyzer using primer AVGR06 (Table 1). The resulting construct, the bait plasmid, was designated pBridge-S100P$_2$.

Activation Domain (AD) Plasmid Expressing the GAL4 AD−S100P Fusion Protein. The PCR product used to create the pBridge-S100P$_1$ plasmid was ligated into the unique *Eco*RI and *Bam*HI restriction sites of the pGAD424 vector (Clontech). The S100P recombinant was confirmed by *Sac*I digestion and automated DNA sequencing using the pGAD424 sequencing primer (AVGR07), as described above for the pBridge-S100P$_1$ construct. The resulting construct was designated pGAD424-S100P.

Yeast Two-Hybrid Screen

The yeast strain used in the screen was AH109 (Clontech). As a first step, yeast cells were transformed with pBridge-S100P$_1$, pBridge-S100P$_2$, or the pBridge vector (for use as a negative control) using the one-step procedure (*10*). Cells were plated on synthetic complete (SC) agar plates lacking tryptophan and containing 2% glucose. Colonies growing on the selection plates were picked and grown in liquid SC medium lacking tryptophan with 2% glucose. These liquid cultures were transformed with the human prostate MATCHMAKER cDNA library (Clontech) by the high-efficiency yeast transformation method (*11*). The cells were then plated on SC agar selection plates lacking tryptophan, leucine, and histidine, containing 10 mM 3-aminotriazole. Colonies growing on the selection plates were picked, inoculated into SC liquid medium lacking leucine, and cultured for at least 20 generations to allow the loss of the S100P bait plasmids carrying the *TRP1* nutritional marker. Cells which had lost the TRP1 plasmid were identified by replica plating. Plasmids from such cells were isolated and used to transform *E. coli* strain MC1066 (*12*) to further select for the *LEU2* nutritional marker (MC1066 has a defect in the leucine biosynthesis pathway, which can be complemented by the yeast *LEU2* gene contained on the library plasmids). DNA from the bacterial colonies growing on the minimal plates was isolated and used to transform yeast carrying pBridge, pBridge-S100P$_1$, or pBridge-S100P$_2$ to verify the yeast two-hybrid interactions. Two DNA minipreps (a and b) from the MC1066 colonies were used in tests for the repeated two-hybrid interactions. Transformants were plated on SC plates lacking tryptophan, leucine, and either histidine, adenine, or both. At this step, the plates lacking histidine also contained 25 mM 3-aminotriazole, which increased the stringency (*13*). Colonies growing on the selection plates were assayed for β-galactosidase activity as described previously (*13*). Activation domain plasmids from the colonies exhibiting reporter gene activity were recovered and sequenced on an ABI 100 genetic analyzer using the primer AVGR07 (Table 1).

Cloning of S100Z into the Expression Vector pGia

S100Z cDNA was amplified by PCR using the library plasmid obtained in the yeast two-hybrid screen as a template. The forward primer (AVGR08, Table 1) generated *Bam*HI and *Sph*I sites upstream of the putative starting ATG codon. The initiator ATG codon was determined by aligning the S100Z sequence with the sequences of the other S100 family members. The reverse primer (AVGR09, Table 1) provided *Bam*HI and *Pst*I restriction sites 80 bases downstream of the TAA stop codon. The PCR product was purified using a QIAquick PCR purification kit (Qiagen), digested with *Sph*I and *Bam*HI enzymes, and ligated into pGia linearized with *Sph*I and *Bam*HI. The presence of the S100Z insert was confirmed by *Sac*I digestion and automated DNA sequencing of the entire gene using the T7 primer, as described above for the pBridge-S100P$_1$ construct. The resulting plasmid was designated pGia-S100Z.

[1] Abbreviations: G3PDH, glyceraldehyde-3-phosphate dehydrogenase; MTC panels, multiple-tissue cDNA panels; CD, circular dichroism; AD, GAL4 transcriptional activation domain; BD, GAL4 DNA-binding domain; pBridge-S100P$_1$, DNA-binding domain vector pBridge carrying S100P cDNA expressed as a BD fusion; pBridge-S100P$_2$, DNA-binding domain vector pBridge carrying S100P cDNA expressed as a BD fusion and another copy expressed as itself; pGIA-S100Z, expression vector pGIA carrying S100Z cDNA; pGAD424-S100P, transcriptional activation domain vector carrying S100P cDNA expressed as an AD fusion.

Expression and Purification of the Recombinant S100Z

BL21(DE3) *E. coli* competent cells were transformed with the pGia-S100Z plasmid and grown in 1 L of 2×YT medium at 37 °C to an optical density of 1.1 at 600 nm. S100Z synthesis was induced by the addition of 1 mM isopropyl β-D-thiogalactopyranoside to the culture medium, and the culture was incubated for an additional 6 h. Induced cells were harvested by centrifugation at 7500*g* for 30 min at 4 °C. Cell pellets were resuspended in 40 mL of 50 mM sodium acetate buffer (pH 4.0) and passed through the French pressure cell. Cell debris was removed by centrifugation at 39000*g* for 60 min at 4 °C. The supernatant was adjusted to a final calcium concentration of 10 mM and applied on a FastFlow phenyl-Sepharose (Pharmacia) column (3 cm × 10 cm) equilibrated in 50 mM sodium acetate (pH 4.0) and 10 mM $CaCl_2$. The flow through was passed through the column two more times, and the column was washed with 10 column volumes of equilibration buffer. Bound S100Z was eluted from the column with a 6 M urea solution. S100Z was refolded by a stepwise dialysis against decreasing concentrations of urea and several changes of Milli-Q water and subsequently lyophilized. Lyophilized protein was dissolved in 5% acetic acid and 2% β-mercaptoethanol and applied to a Sephadex G-75 gel-filtration column (2.5 cm × 100 cm). S100Z-containing fractions were pooled together and lyophilized. The S100Z purity was higher than 95% as judged from Coomassie staining of SDS polyacrylamide gels. Purified S100Z was subjected to 20 N-terminal sequencing cycles to confirm the identity of the recombinant protein.

Circular Dichroism Spectroscopy

Circular dichroism experiments were performed on a Jasco J-715 automatic recording spectropolarimeter as described previously (*9*). Far-UV CD spectra were recorded in a 1 mm rectangular quartz cell. The buffer that was used consisted of 5 mM Tris-HCl (pH 7.5), 0.2 mM EDTA, and 3 mM DTT (buffer B), with or without 5 mM $CaCl_2$. The protein concentration was 0.44 mg/mL (38 μM). For recording the far-UV CD spectrum of the unfolded S100Z, the buffer also contained 6.5 M guanidinium hydrochloride and the protein concentration was 0.22 mg/mL (19 μM). Near-UV CD spectra of S100Z were recorded in a 1 cm rectangular quartz cell. A 2.1 mg/mL (182 μM) solution of S100Z in buffer A with or without 5 mM $CaCl_2$ was used in all near-UV CD experiments. Each far-UV and near-UV CD experiment consisted of a minimum of five accumulations, and average values are reported. The molar ellipticity of S100Z was calculated as

$$[\Theta] = \frac{\Theta M_w}{10lC} \quad (1)$$

where $[\Theta]$ is the experimental ellipticity, M_w is the mean molecular mass of amino acid residues in S100Z (taken to be 116.6 Da), C is the concentration of S100Z in the solution in milligrams per milliliter, and l is the optical path length in centimeters.

Fluorescence Spectroscopy

Fluorescence measurements were performed on a FluoroMax spectrofluorimeter using DM3000F software (SPEX Industries, Inc.) as described previously (*9*). A quartz cell with a 1 cm path length was used. The buffer used in all experiments was 25 mM Tris-HCl (pH 7.5), 0.2 mM EDTA, and 3 mM DTT (buffer A) with or without 5 mM $CaCl_2$. The fluorescence emission spectrum of the unfolded S100Z was recorded in the presence of 6.5 M guanidinium hydrochloride. Fluorescence emission spectra of S100Z were recorded using an excitation wavelength of 276 nm, and a protein concentration of 39 μM. Emission spectra of the external fluorescence probe, 8-anilino-1-naphthalenesulfonic acid (ANS), were recorded using a 30 μM ANS solution in buffer A, and an excitation wavelength of 350 nm. The protein concentration in the ANS fluorescence experiments was 9.5 μM. All emission spectra were recorded using a step resolution of 0.5 nm and an integration time of 3 s.

The Ca^{2+} titrations of S100Z were carried out using tyrosine residues of the protein as internal fluorescence probes. Samples (1250 μL) of a 26 μM solution of S100Z in buffer A were titrated with increasing amounts of calcium until the calcium concentration exceeded 50 mM. An excitation wavelength of 276 nm and an emission wavelength of 305 nm were used in all experiments. The calcium concentration at each titration point was estimated using a computer algorithm taking into account constants for the association between calcium and EDTA and is reported as pCa ($-\log[Ca^{2+}]$). The experiment was carried out in triplicate, and average values are reported. The experimental data were fitted to the binding equation, assuming there are two calcium-binding sites per S100Z monomer and fluorescence changes are proportional to the fractional saturation for each site:

$$Q = \frac{K_1[Ca^{2+}]}{1 + K_1[Ca^{2+}]} + \frac{K_2[Ca^{2+}]}{1 + K_2[Ca^{2+}]} \quad (2)$$

where Q is the sum of fractional saturations of sites 1 and 2, K_1 is the association constant for the first binding site, K_2 is the association constant for the second binding site, and $[Ca^{2+}]$ is the free calcium concentration.

Analytical Ultracentrifugation Experiments

Analytical ultracentrifugation experiments were performed on a Beckman XLA ultracentrifuge. The absorbance in the cell was measured at 276 nm. Samples were allowed to equilibrate at 25 000 and 30 000 rpm at 4 °C. Two different protein concentrations (0.55 and 1.1 mg/mL S100Z in buffer A with or without 5 mM $CaCl_2$) were used in the experiments. Cells with an optical path length of 1.2 cm were used. Absorbance data were globally fitted to the equation

$$A(r) = A_0 e^{HM(r^2 - r_0^2)} + E \quad (3)$$

where $A(r)$ is the absorbance at any given radius r, A_0 is the absorbance at a reference radius r_0, M is the molecular mass of the species in the cell, and E is the baseline offset. H is the constant equal to $[(1 - \bar{v}\rho)\omega^2]/(2RT)$, where $\bar{v}$ is the partial specific volume of S100Z (0.755 cm^3/g), calculated according to ref *14*, ρ is the density of the solution (assumed to be 1 g/mL), and ω is the rotor angular velocity (*15*).

Electrophoresis

Isoelectric focusing experiments were performed on a PhastSystem gel electrophoresis system (Pharmacia). S100P

Table 2: Results of the Experiments Confirming S100P Expression and Dimerization in Yeast[a]

	pGAD424		pGAD424-S100P	
	growth on His^- medium	β-galactosidase activity	growth on His^- medium	β-galactosidase activity
pBridge	no	no	no	no
pBridge-S100P_1	no	no	yes	yes
pBridge-S100P_2	no	no	yes	yes

[a] To test S100P expression and dimerization in yeast, AH109 cells were cotransformed with the DNA-binding domain plasmids and transcriptional activation domain plasmids in pairwise combinations as indicated in the table. Expression of the reporter genes was assessed by growth of the transformants on the His^- medium and development of the product in β-galactosidase assays.

and S100Z were mixed together at a final concentration of 100 μM, incubated at room temperature for 24 h, and run on PhastGel IEF 3-9 gels (Pharmacia). Gels were fixed with 20% trichloroacetic acid and stained by the fast Coomassie Blue method as recommended by manufacturer. Isoelectric points of the samples were determined using the isoelectric focusing calibration kit (Pharmacia).

Lyophilized samples of S100P and S100Z were dissolved in 25 mM Tris-HCl (pH 7.5), 0.2 mM EDTA buffer containing 6 M urea and 2% β-mercaptoethanol. Samples were centrifuged at 13000*g* to remove any insoluble material. Forty microliters of an S100P solution was mixed with either 200 or 400 μL of S100Z, diluted with 10 mL of 25 mM Tris-HCl (pH 7.5), 0.2 mM EDTA, and 3 mM DTT, and concentrated to a final volume of 130 μL using Vivaspin 20 concentrators (Vivaproducts, Westford, MA). The samples were analyzed by electrophoresis on a 20% native polyacrylamide gel in Tris-glycine buffer (pH 8.3) under nondenaturing conditions as described elsewhere (*16*).

PCR Analysis of S100Z Gene Expression in Human Tissues

Human multiple-tissue cDNA panels I and II and the human tumor MTC panel (Clontech) were used as templates in PCR-based assays of S100Z expression in human tissues. PCRs were carried out as follows. Each reaction mixture contained 5 μL of tissue-specific cDNA as a template (or 5 μL of milli-Q water as a negative control), 0.4 μM 5′ (AVGR10) and 3′ (AVGR11) S100Z gene-specific primers, 0.2 mM dNTPs, 5 mM $MgCl_2$, and 2.5 units of HotStarTaq DNA polymerase (Qiagen) in the final reaction volume of 50 μL. The thermal cycler (MJ Research) was programmed for a total of 38 cycles with an initial activation step of 15 min at 95 °C to activate HotStarTaq DNA polymerase, a 30 s denaturation step at 94 °C, and a 2 min annealing/extension step at 68 °C. During cycles 26, 30, 34, and 38, the annealing/extension time was increased to 5 min, and upon completion of cycles 26, 30, 34, and 38, the temperature was dropped to 15 °C and 5 μL samples were removed for further analysis. PCR products were analyzed by electrophoresis on a 2% agarose gel in TAE buffer. DNA bands were visualized by staining the gels with an ethidium bromide solution (0.5 μg/mL). PCR products obtained from the leukocyte, lung, colon, ovary, and prostate cDNAs were further amplified in separate PCRs under the same conditions described above, except that the primer concentrations were increased to 5 μM. Products of the amplification reactions were sequenced on an ABI 100 genetic analyzer using primer AVGR11 (Table 1) to confirm their identities.

RESULTS AND DISCUSSIONS

Yeast Two-Hybrid Screen. We performed a yeast two-hybrid screen of the human prostate cDNA library using S100P as bait in search for possible S100P binding partners. Before proceeding with the actual screen, we determined whether GAL4−S100P fusion proteins could be expressed in yeast and whether fusing S100P to the GAL4 DNA-binding domain would affect the dimerization properties of S100P. Another preliminary step was a test for the S100P interaction with the GAL4 transcriptional activation domain alone. We cloned a copy of the S100P cDNA into an activation domain vector (AD, pGAD424), which also carries the GAL4 transcriptional activation sequence, creating a pGAD424-S100P plasmid. When this plasmid was used together with any of the pBridge-S100P constructs to cotransform yeast cells, the yeast transformants did grow on the histidine-deficient medium in the presence of 10 mM 3-AT, indicating an interaction of two S100P subunits (one from the BD−S100P fusion and another from the AD−S100P fusion). The results of these experiments are presented in Table 2. As can be seen from these results, cells with either pBridge-S100P_1 or pBridge-S100P_2 are not able to grow on the selection plates when they are cotransformed with the pGAD424 vector. Similarly, when cells bearing pGAD424-S100P are cotransformed with the pBridge vector, growth on plates lacking histidine is not observed. On the contrary, cells cotransformed with either pBridge-S100P_1 or pBridge-S100P_2 and pGAD424-S100P grow on the selection plates lacking histidine. These results indicate that the S100P fusions are expressed in yeast, fusing S100P to the GAL4 sequences does not prevent dimerization of the protein, and the S100P bait alone does not activate the reporter gene sufficiently to allow cells to grow on the selection plates. Therefore, the yeast two-hybrid screen was used to search for the putative S100P-binding partners.

As described in Materials and Methods, we have transformed yeast AH109 carrying pBridge-S100P_2 with the MATCHMAKER human prostate cDNA library. Out of 7×10^7 primary *TRP1*$^+$*LEU2*$^+$ transformants plated on selective media, 19 exhibited growth within 1 week. These candidates were further tested to verify interaction as described in Materials and Methods. *LEU2*-bearing plasmids isolated from 14 of them (labeled 1P−14P) were capable of activating transcription of at least two of the reporter genes when they were transformed back into yeast carrying either pBridge-S100P_1 or pBridge-S100P_2 but not pBridge alone; three candidates (labeled 15P−17P) activated transcription of the reporters only in the presence of pBridge-S100P_2, but not pBridge-S100P_1, and two candidates did not activate transcription in the presence of either pBridge-S100P_1 or

```
CGGCCGCGTCGACGTGTCTTCTCCCCGGGTTTGGTGGCCTGCTTCTGGAGTGGTCAGTT
CTGCTGCCGAC

ATG CCC ACC CAG CTC GAG ATG GCC ATG GAC ACC ATG ATT AGA ATC
 M   P   T   Q   L   E   M   A   M   D   T   M   I   R   I
     P   T   Q   L   E   M   A   M   D   T   M   I   R   I

TTC CAC CGC TAT TCT GGC AAG GCA AGG AAG AGA TTC AAG CTC AGC
 F   H   R   Y   S   G   K   A   R   K   R   F   K   L   S
 F   H   R   Y   ?   G

AAG GGG GAA CTG AAA CTG CTC CTG CAG CGA GAG CTC ACG GAA TTC
 K   G   E   L   K   L   L   L   Q   R   E   L   T   E   F

CTC TCG TGC CAA AAG GAA ACC CAG TTG GTT GAT AAG ATA GTG CAG
 L   S   C   Q   K   E   T   Q   L   V   D   K   I   V   Q

GAC CTG GAT GCC AAT AAG GAC AAC GAA GTG GAT TTT AAT GAA TTC
 D   L   D   A   N   K   D   N   E   V   D   F   N   E   F

GTG GTC ATG GTG GCA GCT CTG ACA GTT GCT TGT AAT GAT TAC TTT
 V   V   M   V   A   A   L   T   V   A   C   N   D   Y   F

GTA GAA CAA TTG AAG AAG AAA GGA AAA TAA
 V   E   Q   L   K   K   K   G   K  stop

AGGAAGTTCGTGAATGCTAATCAAGGGCTAAATAAAGTGGAAAGCAGCATCCTTGAAAA
TCATCGCTCTCAGGACGGGGTCTGCCTGGAATAGCACTGAATGTGTTTAGTAGTGCCTT
TGGCTTGGGGGCTTTGGAGAAAGCTGCTTGAGCCCTTCCCATGCTCATAAATTAGGTAA
GGAGCAGTGAGGAAGTGTCTACAGCAACCTCAGCCAATGTGATGTGCCTCTACAGCTGT
ATGAAGTAGATTCTCTGATTTTATCTTGAAATCCACACTTTCACAAAAAAGAAAAATCT
TTTACCATTTCTCATTNAAAAAGTAATATAAGNTTATTGTAGGGGGAAATAATCAAATG
TAATCTCATTACCCAAAATAATCCTGTACTTCTTATTCNATTTTCCTAGGCCTGTTCTA
TTTTTATGGTCTAAAATGAAATCNCCAAAAANTNTTTGTTTTTTTAGTCNNTTATAAAT
CACTGNTNTTAAAAAAAAAAAAAAAAAAAAAACNNGGANTTTTTTTTNTNAAACCCNG
TNTTTATTGGTGGGCCATTATTTTT
```

FIGURE 1: Sequence of candidate 9P. DNA sequence of the cDNA insert of clone 9P. The sequence of the open reading frame starting with the first methionine after GAL4-AD is given in triplets. The polyadenine tail of mRNA is double-underlined. The translation of the open reading frame is given in bold and italicized. Results of the direct N-terminal amino acid sequencing of the recombinant protein produced in *E. coli* are given below the translation of the open reading frame.

pBridge-S100P$_2$. None of the candidates exhibited expression of the reporter genes when cotransformed with the pBridge vector. Of these 17 candidates that were capable of activating the reporter gene activity, candidate 9P exhibited the most rapid growth on the selective media (within 1 day) and the most rapid development of the blue color in the β-galactosidase assay (within 2 h). Therefore, the analysis of this candidate was pursued.

Characterization of Candidate 9P. The DNA sequence of the cDNA insert of candidate 9P is shown in Figure 1. A BLAST search of the human genome with the DNA sequence of the candidate revealed that fragments of the sequence are located on chromosome 5 (reference number NT_006713.3), indicating that this is not a cloning artifact and homologous sequences are indeed present in the human genome. The cDNA insert of the candidate contains an open reading frame encoding a polypeptide of 99 amino acids (starting with the first in-frame ATG codon after the GAL4 activation domain sequence, Figure 1) fused to the GAL4 transcriptional activation domain sequence. A BLAST search with the polypeptide sequence did not reveal any matches in the databases. The predicted polypeptide sequence of the candidate is homologous to the members of the S100 family of calcium-binding proteins (Figure 2), being most similar to S100A1 and S100P (57% of the sequence identical and 76% similar and 52% of the sequence identical and 71% similar, respectively). Residues in the putative C-terminal EF-hand ("classical" EF-hand) that are known to be involved in calcium binding (*17*) are conserved (D63, N65, D67, and E74). Conservation of amino acid residues is also evident in the putative N-terminal EF-hand (atypical EF-hand, Figure 2). There are two cysteines in the novel protein sequence (C48 and C86), providing the potential to form covalently linked dimers and possibly higher-order oligomers. The predicted isoelectric point of the protein is 8.6, which is unusual for other members of the S100 family, since most of them are acidic. This high pI value can be attributed to the four lysine residues present at the C-terminus of the polypeptide. Without the four lysines, the predicted pI would be 5.4. These lysines are not due to the erroneous translation of the poly-A tail of the mRNA, because the poly-A tail of the mRNA is clearly defined in the cDNA sequence (Figure 1). Secondary structure estimates using the joint prediction option of the PELE secondary structure prediction software (*18*) suggest that 67% of the amino acid residues of the polypeptide are in an α-helical conformation. Sequence

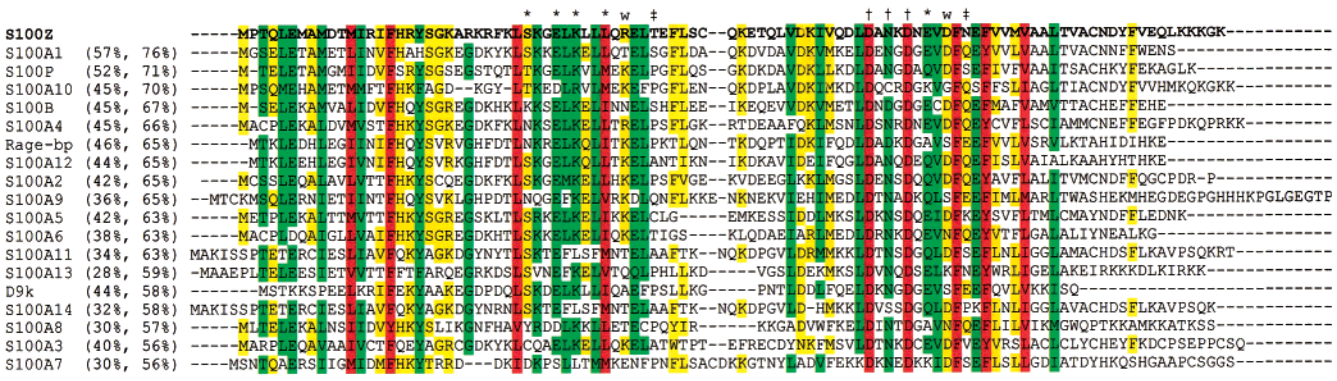

FIGURE 2: Sequence alignment of candidate 9p and 16 other members of the S100 family. Sequences of the S100 proteins were aligned using CLUSTALW software (*52*, *53*). Residues predicted to be involved in calcium coordination (according to ref *27*) are marked as follows: ∗, backbone carbonyl; w, water molecule; †, monodentate side chain carboxylate or amide; and ‡, bidentate side chain carboxylate. Numbers in parentheses indicate the percent sequence identity and the percent sequence similarity between S100Z and the corresponding member of the S100 family. Residues conserved throughout the entire S100 family are highlighted in red; residues conserved in at least 75% of the family members are highlighted in green, and residues conserved in at least 60% of the family members are highlighted in yellow.

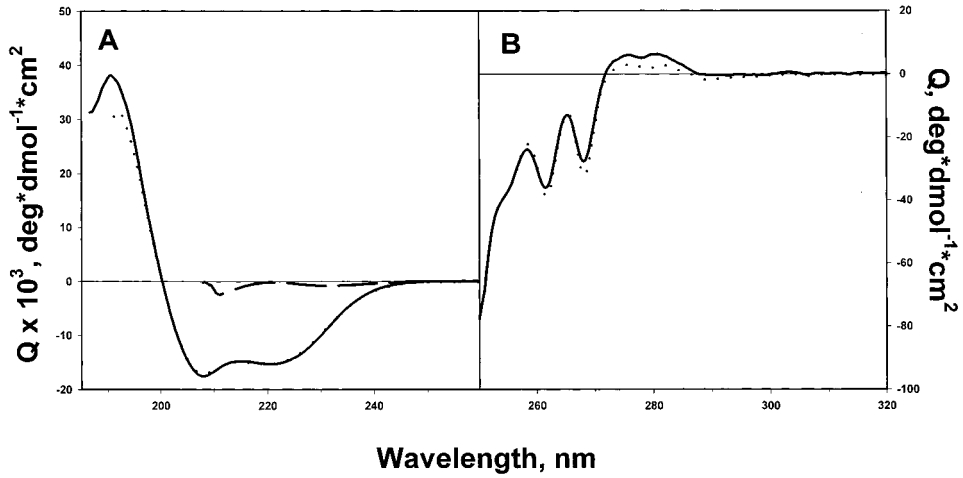

FIGURE 3: Circular dichroism spectra of recombinant S100Z. (A) Far-UV CD spectrum of S100Z. (B) Near-UV CD spectrum of S100Z. Dotted lines show the spectra of the apo form of the protein, and solid lines show the spectra of S100Z in the presence of 5 mM $CaCl_2$. The dashed line corresponds to the far-UV CD spectrum of S100Z in the presence of 6.5 M guanidinium chloride.

homology, the length of the polypeptide, and the presence of the EF-hand calcium-binding motifs in the primary sequence strongly suggest that the novel protein is a member of the S100 family. There is an accepted nomenclature for the S100 proteins located on chromosome 1 (*19*). Other S100 genes are located on chromosome 4 (S100P) or chromosome 21 (S100B). However, no DNA sequences corresponding to the novel protein are located on chromosome 1, chromosome 4, or chromosome 21. This argues against the new gene being another *S100A*, *S100P*, or *S100B* gene (*19*). We therefore name the novel S100 gene *S100Z*.

Expression and Purification of Recombinant S100Z. S100Z cDNA was cloned into a bacterial expression system, and the protein product was purified to homogeneity as described in Materials and Methods. The yield of the soluble recombinant S100Z was ~20−25 mg/L of bacterial culture. The recombinant protein was subjected to the 20 N-terminal sequencing cycles. As is illustrated in Figure 1, 20 N-terminal amino acid residues of the protein are virtually identical to the ones predicted from the cDNA sequence (the only difference being the absence of the N-terminal methionine). To verify that the recombinant protein possesses secondary and tertiary structures and to determine if S100Z is able to bind calcium, we have used several biophysical methods: circular dichroism and fluorescence spectroscopies and equilibrium analytical centrifugation.

Circular Dichroism. Circular dichroism spectroscopy is a method widely used to obtain low-resolution structural data. Far-UV CD spectroscopy is used to estimate secondary structure of the proteins; ellipticity minima at 208 and 222 nm indicate that a protein is largely in an α-helical conformation, while changes in the near-UV region of the CD spectrum reflect changes in the environment of the aromatic residues (*20*). The circular dichroism spectra of S100Z are presented in Figure 3. The far-UV CD spectrum of the apoprotein displays two characteristic minima at 208 and 222 nm, indicating that S100Z is predominantly α-helical. Addition of 6.5 M guanidinium hydrochloride results in the loss of the far-UV CD signal, indicating that under these conditions S100Z is completely unfolded (*20*). The near-UV CD spectrum of the apoprotein exhibits ellipticity minima at 262 and 268 nm, indicating that phenylalanine residues of S100Z are in the unique packed environment (*20*, *21*). Taken together, these results indicate that recombinant S100Z possesses both secondary and tertiary structure reflective of a folded state.

Addition of millimolar calcium concentrations to the solution of S100Z does not produce significant changes in the far-UV CD spectrum, suggesting that calcium has no effect on the secondary structure of the protein. This is also the case with other S100 family members. Calcium does not induce formation of the α-helical structure but rather causes reorientation of the helices with respect to each other (*22*−*28*). The near-UV CD spectrum displays a small decrease in ellipticity between 272 and 287 nm upon addition of 5 mM $CaCl_2$. This result indicates that there is no dramatic change in the environment of the tyrosine residues in the presence of calcium, but does not necessarily rule out the possibility that S100Z is capable of binding the ion. For example, calcium-responsive conformational changes could occur without significantly affecting tyrosine residues of S100Z and would therefore not be detected by CD spectroscopy. In contrast to CD, fluorescence spectroscopy is a more sensitive method of investigating possible calcium binding to S100Z.

Fluorescence Spectroscopy. Fluorescence spectroscopy is a useful tool for monitoring conformational changes in proteins upon their interaction with ligands (*29*). Fluorescent properties of the aromatic residues in the protein (primarily tyrosines and tryptophans) often change upon structural perturbations, resulting in either a change in fluorescence intensity, a shift in the maximum emission wavelength, or both. Furthermore, use of an external fluorescence probe, such as 8-anilino-1-naphthalenesulfonic acid (ANS), for example, can provide additional information about structural changes that are taking place in a protein upon ligand binding. Binding of ANS to hydrophobic surfaces is thought to produce increases in ANS fluorescence intensity (*29*). If any hydrophobic surfaces become exposed upon protein−ligand interaction, addition of the ligand (in our case calcium) to the protein solution will produce an increase in ANS fluorescence intensity. Using ANS as an external fluorescence probe is particularly suitable for studies on S100 proteins, which are thought to expose hydrophobic surfaces upon interaction with calcium (*30*−*32*).

To study calcium binding by S100Z, we have used both internal (tyrosines) and external (ANS) fluorescence probes. S100Z has two tyrosine residues in its amino acid sequence (Y19 and Y89) and no tryptophans, and upon excitation at 276 nm produces a characteristic fluorescence emission spectrum (Figure 4). Addition of 6.5 M Gdm-HCl produces a dramatic increase in the emission spectrum intensity, indicating that structural changes are taking place in the presence of the denaturant; i.e., S100Z becomes unfolded. These results further support the conclusion from the CD experiments (Figure 3) that recombinant S100Z is folded. Addition of 5 mM calcium to the solution of the native protein results in an increase in fluorescence intensity, as compared to that of the apo form of the protein. Because calcium itself does not have fluorescence, the increase is most likely due to structural rearrangements in the S100Z molecule that are taking place upon interaction with calcium. Therefore, we conclude that S100Z is capable of binding the ion. Assuming that the observed changes in fluorescence intensity in response to calcium are proportional to the degree of calcium binding, we fit the data to the binding equation (eq 2). Results of the nonlinear regression analysis produce apparent Ca^{2+} binding constants on the order of $\sim 5 \times 10^6$ M^{-1} for the high-affinity site and 10^2 M^{-1} for the low-affinity site (Figure 4, inset). Furthermore, the increase in the ANS fluorescence intensity upon addition of Ca^{2+} to the S100Z solution (results not shown) indicates that calcium binding to S100Z exposes hydrophobic surfaces.

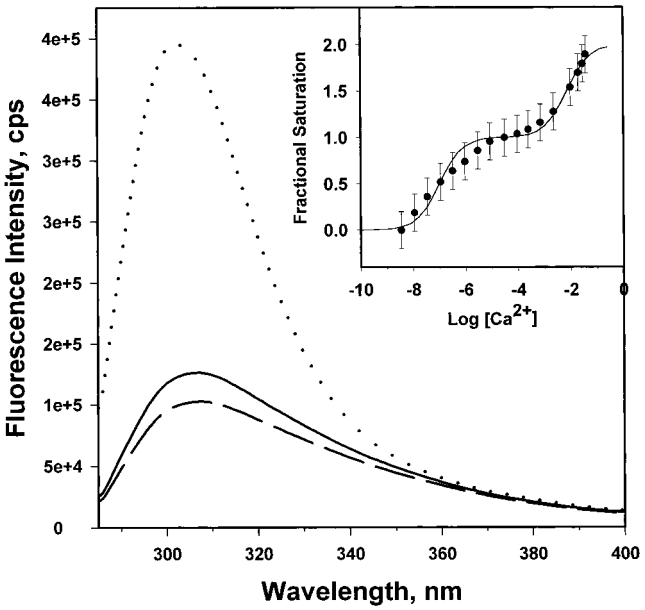

FIGURE 4: S100Z−calcium interactions monitored by tyrosine fluorescence. Emission spectra of the internal fluorescence probes (Y19 and Y89): (− − −) emission of the apoprotein, (—) emission of S100Z in the presence of 5 mM $CaCl_2$, and (•••) emission in the presence of 6.5 M guanidinium hydrochloride. The inset shows changes in fractional saturation (assumed to be proportional to the fluorescence changes) plotted as a function of calcium concentration: (●) average of three independent experiments and (—) fit of the experimental data to the binding equation (eq 2) with apparent association constants of 5×10^6 M^{-1} for the high-affinity site and 10^2 M^{-1} for the low-affinity site.

Equilibrium Analytical Ultracentrifugation. Most S100 proteins exist as dimers in solution, a notable exception being calbindin D9k, which exists as a monomer (*33*−*35*). To analyze the oligomerization state of S100Z, we have used equilibrium analytical ultracentrifugation. In analytical ultracentrifugation experiments, the distribution of the light-absorbing species in the cell along the rotor radius depends on the molecular mass of the species, the stoichiometry of the association (if any association is taking place), and the equilibrium constant of the association (*36*, *37*). Equilibrium analytical ultracentrifugation experiments with S100Z were performed in the presence or absence of calcium to test for possible effects of calcium on the oligomerization properties of S100Z. As illustrated in Figure 5, centrifugation profiles in the presence of 0.2 mM EDTA or 5 mM Ca^{2+} overlap, indicating that calcium has no effect on the apparent molecular mass of S100Z. A global fit of six experiments at two different rotor speeds to the model describing sedimentation of a single ideal species (eq 2) produces an apparent molecular mass of the species in the cell of 22.1 ± 0.1 kDa. From the expected molecular mass of the S100Z monomer with the N-terminal methionine removed (11 430 Da) and the apparent molecular mass obtained in analytical centrifugation experiments, we conclude that under our experimental conditions S100Z exists as a dimer in solution.

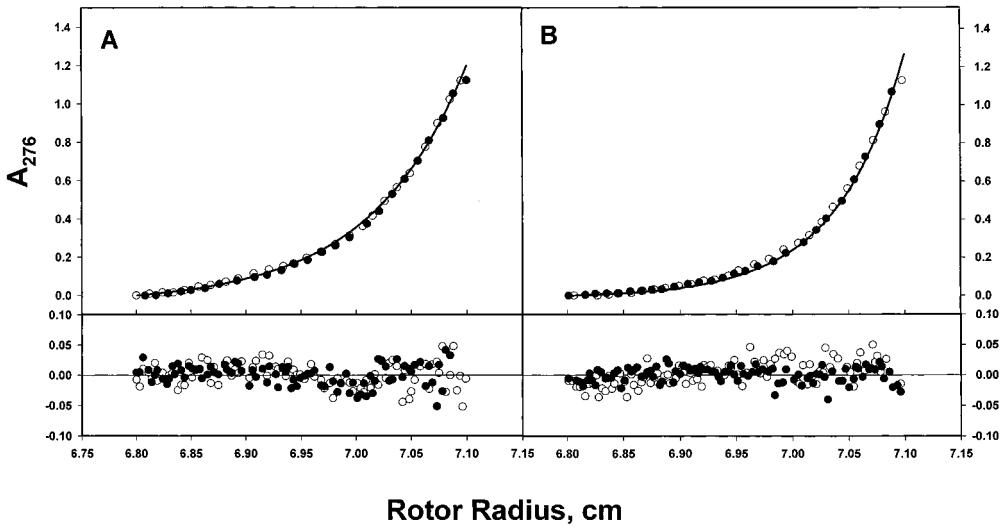

FIGURE 5: Analytical ultracentrifugation profiles of S100Z. (A) Samples were equilibrated at (A) 25 000 and (B) 30 000 rpm: (○) experimental data obtained in the presence of 0.2 mM EDTA, (···) fit of the experimental data obtained in the presence of 0.2 mM EDTA to eq 2, (●) experimental data obtained in the presence of 5 mM $CaCl_2$, and (—) fit of the experimental data obtained in the presence of 5 mM $CaCl_2$ to eq 2. Residuals of the fits shown are also included using the same symbols.

Physical Interaction of S100Z with S100P. Yeast two-hybrid experiments clearly established that S100Z and S100P can interact in the context of the yeast two-hybrid system. To demonstrate physical interaction between S100Z and S100P in vitro, we took advantage of the fact that the two proteins have very different isoelectric points (predicted pI for S100P of 4.8, and a predicted pI for S100Z of 8.6). When S100P or S100Z is incubated alone and allowed to focus, S100P bands are observed just above the bands corresponding to the soybean trypsin inhibitor (pI = 4.55). S100Z bands focus at the other end of the gel, close to the lentil lectin basic band (pI = 8.65). If a complex is formed between these two proteins in a 1:1 ratio, it will have a predicted isoelectric point of 5.6. Indeed, when the two proteins were incubated together and allowed to focus, in addition to the two bands corresponding to S100P and S100Z, a third band migrating between bovine carbonic anhydrase B (pI = 5.85) and β-lactoglobulin (pI = 5.20) was observed (data not shown). This indicates that S100P and S100Z are able to form a complex under in vitro conditions. To further confirm this result, the S100P/S100Z mixture was subjected to PAGE on a 20% gel under nondenaturing conditions (pH 8.3), as shown in Figure 6. Under these conditions, the S100P, being an acidic protein, migrates very fast (lane 1). In contrast, S100Z (lane 2) will not enter the gel. However, the S100P/S100Z mixture exhibits two distinct bands (lanes 3 and 4), one corresponding to S100P and the other to the S100P−S100Z complex (Figure 6).

Expression of the S100Z Gene in Human Tissues. One of the properties of S100 proteins is their tissue-specific expression (*3*); therefore, we analyzed the tissue distribution of S100Z mRNA expression using a variation of the RT-PCR analysis, multiple-tissue cDNA (MTC) panels (*38*−*46*). The use of RT-PCR for the analysis of gene expression has a major advantage over the more traditional RNA blotting techniques. RT-PCR can be 1000−10000-fold more sensitive than Northern blotting, allowing the detection of very rare transcripts (*47*, *48*). MTC panels provide an additional advantage by allowing more accurate assessment of tissue-specific gene expression due to the cDNA normalization and

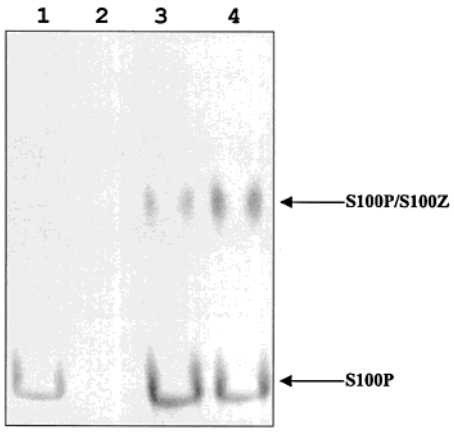

FIGURE 6: S100P−S100Z heterocomplex formation as shown by PAGE under nondenaturing conditions: lane 1, S100P alone; lane 2, S100Z alone (does not enter the gel); lane 3, S100P and S100Z samples mixed at a 1:5 ratio; and lane 4, S100P and S100Z samples mixed at a 1:10 ratio. See Materials and Methods for details.

have been successfully used in a number of studies (*38*−*46*). The amount of cDNA in each panel has been normalized to the expression levels of several housekeeping genes to account for possible differences in the levels of transcriptional activity between different tissues. Therefore, differences in the amount of product obtained in PCRs under identical conditions are likely to represent true differences in corresponding mRNA levels between different tissues represented on the MTC panel. The relative abundance of the mRNA of interest can be estimated from the number of cycles after which the PCR product is observed and/or band intensity, provided that the reaction has not reached saturation, i.e., primers and dNTPs have not been depleted. It was shown (*38*−*46*, *49*) that 22−25 cycles is sufficient to detect high-abundance mRNA and 25−30 cycles is sufficient to detect high- to medium-abundance mRNAs; after 30−35 cycles, medium- to low-abundance mRNA will produce a distinctive band, and after 38 cycles, even very rare transcripts can be detected.

We have used human MTC panels I and II to study expression of the S100Z gene in normal tissues. Human MTC

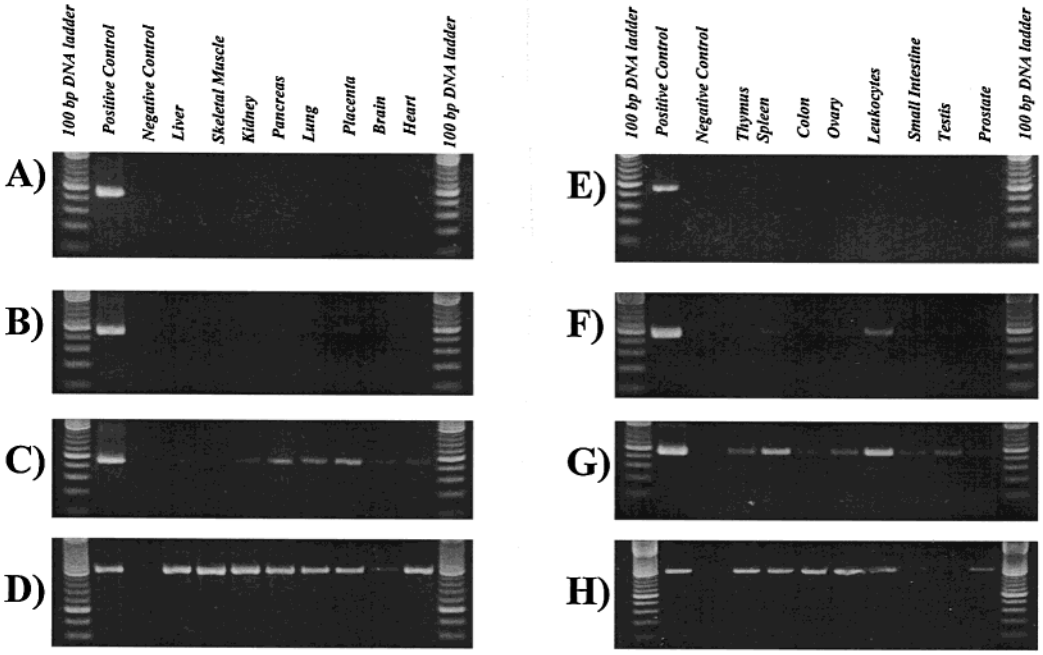

FIGURE 7: PCR-based assays of S100Z gene expression in normal human tissues. Electrophoresis of the PCR products obtained in S100Z cDNA amplification from the MTC panels on 2% agarose gels in TAE buffer. The forward primer for the PCR was designed to anneal to the first 10 codons of the S100Z coding sequence. The reverse primer was designed to anneal in the 3′ untranslated region (120 bases downstream from the translation termination codon) to avoid amplification of homologous cDNA sequences corresponding to the other S100 family members. Therefore, the expected size of the PCR product was 450 bases. As a positive control, primers for amplification of a 983 bp fragment of the glyceraldehyde-3-phosphate dehydrogenase (G3PDH) gene (one of the housekeeping genes used for cDNA normalization in MTC panels) were used. (A−C) PCR products obtained from reactions using cDNAs from human MTC panel I as templates (liver, skeletal muscle, kidney, pancreas, lung, placenta, brain, and heart) using S100Z-specific primers: samples analyzed by gel electrophoresis after 30 cycles (A), 34 cycles (B), 38 cycles (C), and expression of the control gene, G3PDH, in corresponding tissues from human MTC panel I (D) (samples analyzed after 30 PCR cycles). (E−G) PCR products obtained from reactions using cDNAs from human MTC panel II as templates (thymus, spleen, colon, ovary, leukocytes, small intestine, testis, and prostate) using S100Z-specific primers: 30 PCR cycles (E), 34 cycles (F), 38 cycles (G), and expression of the control gene, G3PDH, in corresponding tissues from human MTC panel II (H) (samples analyzed after 30 cycles). Candidate 9P was used as a template in the positive control reactions. The brightest bands on the 100 bp DNA ladder lanes correspond to 500 (bottom) and 1000 bp (top).

panel I includes cDNA from brain, heart, kidney, liver, lung, pancreas, placenta, and skeletal muscle. MTC panel II includes cDNA from colon, ovary, peripheral blood leukocytes, prostate, small intestine, spleen, testis, and thymus. Because abnormal levels of S100 proteins are often observed in human tumors (*3*), and some of them were shown to be differentially expressed in different tumors (*50*), we have also analyzed S100Z expression in several tumor tissues using a human tumor MTC panel. The tumor MTC panel includes cDNA from breast tumor (carcinoma GI-101), lung tumors (carcinoma LX-1 and GI-117), colon tumors (adenocarcinoma CX-1 and GI-112), prostate adenocarcinoma (PC-3), ovarian carcinoma (GI-102), and pancreatic adenocarcinoma (GI-103) propagated as xenografts in nude mice.

The results of PCR amplifications of cDNA from normal tissues are presented in Figure 7. Levels of the S100Z transcript are clearly different between different tissues, indicating that S100Z expression is tissue-specific. The highest levels of S100Z expression are observed in peripheral blood leukocytes where the PCR product can be observed after 30 PCR cycles. cDNA from pancreas, lung, placenta, and spleen produces distinct bands after 34 cycles, indicating medium to low levels of expression of S100Z. The level of expression of S100Z in kidney, brain, heart, thymus, colon, ovary, small intestine, testis, prostate, liver, and skeletal muscle is very low; bands can be observed after only 38 PCR cycles. Nevertheless, the observed PCR bands are produced due to the presence of S100Z mRNA in these tissues and not due to overcycling; no bands are observed in any of the negative controls after 38 PCR cycles. Direct DNA sequencing of the PCR products from leukocyte, lung, colon, ovary, and prostate cDNAs (results not shown) indicates that the amplified product corresponds to the S100Z cDNA and does not represent a known member of the S100 family.

Analysis of S100Z expression in tumor tissues shows that S100Z expression appears to be aberrant at least in some tumors (Figure 8). For example, when cDNA from normal lung is used as a template, the PCR product is clearly visible after 34 cycles, while in both lung carcinomas that we analyzed, the S100Z band can be barely detected after 38 cycles. In cDNA from colon adenocarcinoma CX-1, S100Z expression could not be detected at all, while in normal colon, S100Z cDNA is present. Colon adenocarcinoma GI-112 appears to have somewhat higher levels of S100Z expression, as compared to its normal counterpart. Finally, expression of S100Z could not be detected at all in breast carcinoma GI-101, although we cannot conclude whether there is any difference in S100Z expression between normal and tumor breast tissue as normal breast cDNA was not included in either MTC panel I or II. We did not observe any significant variation in the glyceraldehyde-3-phosphate dehydrogenase (G3PDH) expression profiles between different tissues, with the exception of testis where the level of expression of

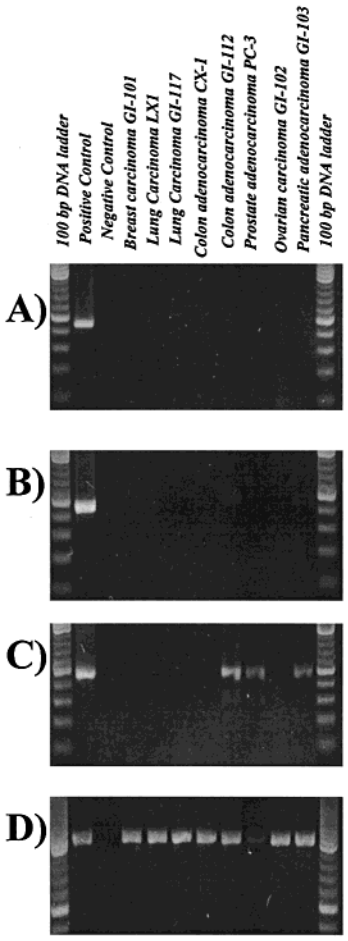

FIGURE 8: PCR-based assays of S100Z expression in human tumor models. cDNAs from the human tumor cDNA panel were used as templates in PCRs: samples analyzed after 30 PCR cycles (A), 34 cycles (B), 38 cycles (C), and expression of the control gene, G3PDH, in the tumor models (D) (samples analyzed after 30 PCR cycles).

G3PDH is known to be low (*51*), indicating that assay conditions were identical for all experiments. Therefore, we conclude that the observed differences in the amount of S100Z PCR product are due to different S100Z expression levels between different tissues.

CONCLUSIONS

Using a yeast two-hybrid screen, we have identified a novel member of the S100 family of calcium-binding proteins, human protein S100Z which is capable of interacting with S100P. To date, S100Z is the only protein known to interact with S100P, although we cannot exclude the possibility that other S100 family members are capable of interacting with either S100P or S100Z. S100Z is a protein with high α-helix content, capable of binding two calcium ions per monomer with different affinities (apparent binding constants of 5×10^{-6} and 10^2 M^{-1}). Upon calcium binding, S100Z exposes additional hydrophobic surfaces, as can be judged from the increase in the ANS fluorescence. S100Z is a dimeric protein, and the dimerization properties of S100Z are not affected by calcium binding. The highest levels of *S100Z* gene expression are found in spleen and leukocytes, and it appears that expression of the gene is aberrant in several human tumor models.

ACKNOWLEDGMENT

We thank Dr. Alok Sil for numerous discussions.

REFERENCES

1. Chazin, W. J. (1995) *Nat. Struct. Biol. 2*, 707−710.
2. Kretsinger, R. H. (1976) *Annu. Rev. Biochem. 45*, 239−266.
3. Donato, R. (2001) *Int. J. Biochem. Cell Biol. 33*, 637−668.
4. Averboukh, L., Liang, P., Kantoff, P. W., and Pardee, A. B. (1996) *Prostate 29*, 350−355.
5. Guerreiro Da Silva, I. D., Hu, Y. F., Russo, I. H., Ao, X., Salicioni, A. M., Yang, X., and Russo, J. (2000) *Int. J. Oncol. 16*, 231−240.
6. Gribenko, A. V., and Makhatadze, G. I. (1998) *J. Mol. Biol. 283*, 679−694.
7. Gribenko, A. G., Guzman-Casado, M., Lopez, M. M., and Makhatadze, G. I. (2001) *Biophys. J.* (submitted for publication).
8. Fields, S., and Song, O. (1989) *Nature 340*, 245−246.
9. Gribenko, A., Lopez, M. M., Richardson, J. M., III, and Makhatadze, G. I. (1998) *Protein Sci. 7*, 211−215.
10. Chen, D. C., Yang, B. C., and Kuo, T. T. (1992) *Curr. Genet. 21*, 83−84.
11. Gietz, D., St Jean, A., Woods, R. A., and Schiestl, R. H. (1992) *Nucleic Acids Res. 20*, 1425.
12. Casadaban, M. J., Martinez-Arias, A., Shapira, S. K., and Chou, J. (1983) *Methods Enzymol. 100*, 293−308.
13. Vojtek, A. B., and Hollenberg, S. M. (1995) *Methods Enzymol. 255*, 331−342.
14. Makhatadze, G. I., Medvedkin, V. N., and Privalov, P. L. (1990) *Biopolymers 30*, 1001−1010.
15. Johnson, M. L., Correia, J. J., Yphantis, D. A., and Halvorson, H. R. (1981) *Biophys. J. 36*, 575−588.
16. Speicher, D. W. (1998) in *Current Protocols in Protein Chemistry* (Coligan, J. E., Dunn, B. M., Ploegh, H. L., Speicher, D. W., and Wingfield, P. T., Eds.) pp 10.13.11−10.13.19, John Wiley & Sons, New York.
17. Marsden, B. J., Shaw, G. S., and Sykes, B. D. (1990) *Biochem. Cell Biol. 68*, 587−601.
18. King, R. D., and Sternberg, M. J. (1990) *J. Mol. Biol. 216*, 441−457.
19. Schafer, B. W., Wicki, R., Engelkamp, D., Mattei, M. G., and Heizmann, C. W. (1995) *Genomics 25*, 638−643.
20. Woody, R. W. (1995) *Methods Enzymol. 246*, 34−71.
21. Pain, R. (1996) in *Current Protocols in Protein Science*, John Wiley & Sons, New York.
22. Smith, S. P., and Shaw, G. S. (1997) *J. Biomol. NMR 10*, 77−88.
23. Kilby, P. M., Van Eldik, L. J., and Roberts, G. C. (1996) *Structure 4*, 1041−1052.
24. Drohat, A. C., Baldisseri, D. M., Rustandi, R. R., and Weber, D. J. (1998) *Biochemistry 37*, 2729−2740.
25. Drohat, A. C., Amburgey, J. C., Abildgaard, F., Starich, M. R., Baldisseri, D., and Weber, D. J. (1996) *Biochemistry 35*, 11577−11588.
26. Sastry, M., Ketchem, R. R., Crescenzi, O., Weber, C., Lubienski, M. J., Hidaka, H., and Chazin, W. J. (1998) *Structure 6*, 223−231.
27. Potts, B. C., Smith, J., Akke, M., Macke, T. J., Okazaki, K., Hidaka, H., Case, D. A., and Chazin, W. J. (1995) *Nat. Struct. Biol. 2*, 790−796.
28. Matsumura, H., Shiba, T., Inoue, T., Harada, S., and Kai, Y. (1998) *Structure 6*, 233−241.
29. Lakowitcz (1983) *Principles of Fluorescent Spectroscopy*, Plenum Press, New York.
30. Allen, B. G., Durussel, I., Walsh, M. P., and Cox, J. A. (1996) *Biochem. Cell Biol. 74*, 687−694.
31. Fohr, U. G., Heizmann, C. W., Engelkamp, D., Schafer, B. W., and Cox, J. A. (1995) *J. Biol. Chem. 270*, 21056−21061.
32. Baudier, J., and Gerard, D. (1986) *J. Biol. Chem. 261*, 8204−8212.
33. Skelton, N. J., Kordel, J., Forsen, S., and Chazin, W. J. (1990) *J. Mol. Biol. 213*, 593−598.

34. Skelton, N. J., Forsen, S., and Chazin, W. J. (1990) *Biochemistry 29*, 5752−5761.
35. Skelton, N. J., Kordel, J., and Chazin, W. J. (1995) *J. Mol. Biol. 249*, 441−462.
36. Minton, A. P. (1990) *Anal. Biochem. 190*, 1−6.
37. Laue, T. M., and Stafford, W. F., III (1999) *Annu. Rev. Biophys. Biomol. Struct. 28*, 75−100.
38. Hemminki, A., Markie, D., Tomlinson, I., Avizienyte, E., Roth, S., Loukola, A., Bignell, G., Warren, W., Aminoff, M., Hoglund, P., Jarvinen, H., Kristo, P., Pelin, K., Ridanpaa, M., Salovaara, R., Toro, T., Bodmer, W., Olschwang, S., Olsen, A. S., Stratton, M. R., de la Chapelle, A., and Aaltonen, L. A. (1998) *Nature 391*, 184−187.
39. Pennica, D., Swanson, T. A., Welsh, J. W., Roy, M. A., Lawrence, D. A., Lee, J., Brush, J., Taneyhill, L. A., Deuel, B., Lew, M., Watanabe, C., Cohen, R. L., Melhem, M. F., Finley, G. G., Quirke, P., Goddard, A. D., Hillan, K. J., Gurney, A. L., Botstein, D., and Levine, A. J. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 14717−14722.
40. Blechschmidt, K., Schweiger, M., Wertz, K., Poulson, R., Christensen, H. M., Rosenthal, A., Lehrach, H., and Yaspo, M. L. (1999) *Genome Res. 9*, 158−166.
41. Huang, X., and Honkanen, R. E. (1998) *J. Biol. Chem. 273*, 1462−1468.
42. Plager, D. A., Loegering, D. A., Weiler, D. A., Checkel, J. L., Wagner, J. M., Clarke, N. J., Naylor, S., Page, S. M., Thomas, L. L., Akerblom, I., Cocks, B., Stuart, S., and Gleich, G. J. (1999) *J. Biol. Chem. 274*, 14464−14473.
43. Kitagawa, H., Shimakawa, H., and Sugahara, K. (1999) *J. Biol. Chem. 274*, 13933−13937.
44. Johnson, J. D., Mehus, J. G., Tews, K., Milavetz, B. I., and Lambeth, D. O. (1998) *J. Biol. Chem. 273*, 27580−27586.
45. Hallas, C., Pekarsky, Y., Itoyama, T., Varnum, J., Bichi, R., Rothstein, J. L., and Croce, C. M. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*, 14418−14423.
46. Pangalos, M. N., Neefs, J. M., Somers, M., Verhasselt, P., Bekkers, M., van der Helm, L., Fraiponts, E., Ashton, D., and Gordon, R. D. (1999) *J. Biol. Chem. 274*, 8470−8483.
47. Byrne, B. C., Li, J. J., Sninsky, J., and Poiesz, B. J. (1988) *Nucleic Acids Res. 16*, 4165.
48. Mocharla, H., Mocharla, R., and Hodes, M. E. (1990) *Gene 93*, 271−275.
49. Clontech (2000) *User Manual PT3158-1*, Clontech, Palo Alto, CA.
50. Kan-Mitchell, J., Liggett, P. E., Taylor, C. R., Rao, N., Granada, E. S., Danenberg, K. D., White, W. L., Van Eldik, L. J., Horikoshi, T., and Danenberg, P. V. (1993) *Invest. Ophthalmol. Vis. Sci. 34*, 3366−3375.
51. Spanakis, E. (1993) *Nucleic Acids Res. 21*, 3809−3819.
52. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res. 22*, 4673−4680.
53. Higgins, D. G., Bleasby, A. J., and Fuchs, R. (1992) *Comput. Appl. Biosci. 8*, 189−191.

BI0114731