Protein Science (1998), 7:211–215. Cambridge University Press. Printed in the USA.

FOR THE RECORD

Cloning, overexpression, purification, and spectroscopic characterization of human S100P

ALEXEY GRIBENKO, MARIA M. LOPEZ, JOHN M. RICHARDSON III, AND GEORGE I. MAKHATADZE

Department of Chemistry and Biochemistry, Texas Tech University, Lubbock, Texas 79409-1061

(RECEIVED September 23, 1997; ACCEPTED October 31, 1997)

Abstract: The calcium-binding protein S100P has been found to be associated with human prostate cancer. We have overexpressed S100P in *Escherichia coli* using a T7 expression system. A rapid two-step procedure for the isolation of overexpressed S100P leads to a preparation of >95% pure protein with a yield of ~150 mg per liter of culture. The structural integrity of recombinant S100P was analyzed using CD and fluorescence spectroscopic techniques. The far-UV CD shows that secondary structure of recombinant S100P consists predominantly of α-helical structure. Both near-UV CD and tyrosine fluorescence spectra show that aromatic residues are involved in the formation of a specific, well packed structure, indicating that the recombinant S100P protein adopts a compact folded conformation. Ca^{2+} has a profound effect on S100P structure. Near-UV CD and fluorescence intensity of both internal (tyrosine) and external (ANS) probes suggest significant structural rearrangements in the tertiary structure of the molecule. The similarity of far-UV CD spectrum of S100P in the presence and in the absence of Ca^{2+} suggests that Ca^{2+} binding has only minor effects on secondary structure.

Keywords: bacterial expression; Ca^{2+}-binding; circular dichroism; fluorescence spectroscopy; human S100P

Calcium plays an important role in regulating cell growth. The Ca^{2+} concentration in cells varies with the cell development stage. Precise balance of the amount of Ca^{2+} in the cells is controlled by a number of calcium-binding proteins, including the S100 subfamily. S100 proteins primarily function to mediate Ca^{2+}-dependent signal transduction pathways involved in cell growth and differentiation, cell cycle regulation, and metabolic control (Shafer & Heizmann, 1996). A characteristic feature of S100 proteins is their cell-type specific expression and their association with both neurological and neoplastic human diseases (Hilt & Kligman, 1991; Shafer & Heizmann, 1996). Differences in expression of S100A protein have been documented for cardiomyopathies (Baudier et al., 1986), whereas differences in expression of S100B have been correlated to Alzheimer's disease (Marshak, 1990; Marshak & Pena, 1992; Marshak et al., 1992), Down's syndrome (Allore et al., 1988), and epilepsy (Griffin et al., 1995). Other members of the S100 subfamily of Ca^{2+}-binding proteins are believed to be associated with metastasis (CAPL, Davies et al., 1993), psoriasis (psoriasin, Hoffmann et al., 1994), cystic fibrosis (MRP14, Renaud et al., 1994), and melanoma (calcyclin, Weterman et al., 1993). S100P protein, the least studied S100 protein, was first isolated from human placenta (Becker et al., 1992; Emoto et al., 1992). Unknown cellular function and the absence of correlation with human disease limited interest to this protein. Recently, however, Averboukh et al. (1996) showed that expression of S100P can be abolished by androgen deprivation in an androgen-responsive prostate cancer cell line, LNCaP-FGC. In contrast, the expression of S100P in androgen-independent cell lines LNCaP-R, DU145, and PC3 differs; it is either constitutively expressed or down regulated (Averboukh et al., 1996). This suggests that differences in the amount of S100P protein in androgen-dependent and androgen-independent prostate cancer cell lines might mediate different pathways of prostate cancer development (Averboukh et al., 1996). These pathways are possibly mediated by the interactions of S100P with other proteins in the prostate cancer cells and these interactions are probably associated with calcium binding. As a first step in elucidating the structural and biochemical role of S100P in prostate tumor development, we have overexpressed the S100P protein in *Escherichia coli* using a T7 expression system and developed a rapid S100P purification method.

S100P cDNA, kindly provided by Dr. Lydia Averboukh, was amplified by PCR. Forward (S100N; 5′-GGGGCAATTCCATATG ACGGACTAGA) and reverse (S100E; 5′-CCGGAATTCGTCAT TTGAGTCC) primers were used to generate unique restriction sites for directed cloning of the coding region of S100P into the T7 expression vector pVEX (Presper et al., 1989). S100N forward primer generated an *Nde* I site and S100E reverse primer generated an *Eco*R I site immediately following the UGA stop codon (Fig. 1). PCR was performed using 1 μg of each primer, 5 ng of template (cDNA), Taq buffer (10 mM KCl, 20 mM Tris-HCl,

Reprint requests to: George I. Makhatadze, Department of Chemistry and Biochemistry, Texas Tech University, Lubbock, Texas 79409-1061; e-mail: gia@ttu.edu.

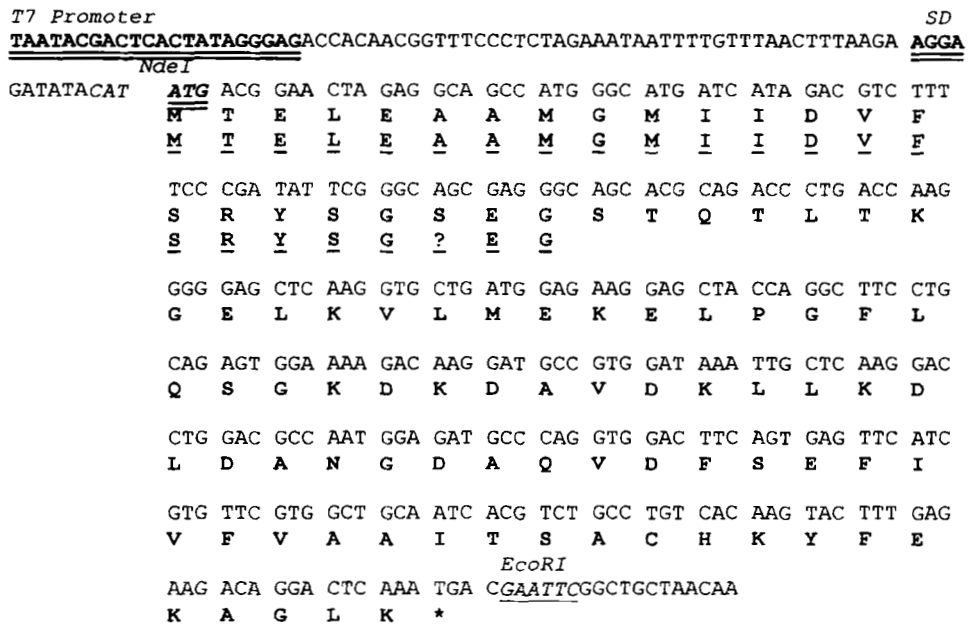

Fig. 1. First row, sequence of pVEX-S100P construct for the expression of the S100P protein in *E. coli*. T7 promoter, Shine–Dalgarno sequence (SD) and the ATG start codon are double-underlined. Second row, amino acid sequence of S100P deduced from the DNA sequence. Third row, N-terminal sequence of the recombinant S100P obtained from the direct protein sequencing. DNA sequence was performed on an ABI Prism 310 genetic analyzer using AmplyTAQ FS dye terminator chemistry. Protein sequencing was performed using Portan 20/20 protein sequencer with an on-line Beckman Gold HPLC system utilizing standard Edman degradation.

pH 8.8, 2 mM $MgSO_4$, 10 mM $(NH_4)_2SO_4$, 0.1% Triton X-100), 1 U of Taq polymerase, 200 μM dNTP in a 100 μL reaction volume. PCR was done on a Perkin-Elmer thermal cycler for 35 cycles using 90 s at 94 °C for denaturation, 120 s at 45 °C for annealing, and 120 s at 72 °C for extension. The reaction mixture was extracted by phenol/chloroform/isoamyl alcohol and purified by Qiaquick PCR purification kit (Qiagen). The PCR fragment was treated with *Nde* I and *Eco*R I, gel purified by Qiaquick Gel Extraction (Qiagen), and ligated into pVEX that had been linearized by *Nde* I and *Eco*R I. BL21 (DE3) strain of *E. coli* was transformed with the ligation reaction by means of electroporation (*E. coli* Pulser, Bio-Rad). Plasmid DNA isolated from several ampicillin-resistant colonies was digested with *Sap* I to identify recombinant plasmids that contained the S100P gene. These clones were sequenced on an ABI Prism 310 genetic analyzer using a T7 primer. Obtained sequence was identical to the S100P structural gene sequence (Fig. 1).

BL21 (DE3) carrying pVEX-S100P plasmid were grown to an optical density of ~0.8 at 600 nm at 37 °C in 2 L of 2×YT medium containing 100 μg/mL ampicillin. S100P production was induced by the addition of isopropyl β-D-thiogalactopyranoside to a final concentration of 1 mM, and the culture was incubated for an additional 7 h. Induced cells were centrifuged for 30 min at 7,500 × *g* at 4 °C. Cell pellets were resuspended in 40 mL of 20 mM Tris-HCl, pH 7.5, buffer containing 1 mM DTT and 1 mM EDTA, and passed twice through a French pressure cell. Cell debris was removed by centrifugation at 39,000 × *g* at 4 °C for 60 min. S100P was located in the soluble fraction. The supernatant fraction was diluted 1:1 with the equilibration buffer (6 mM Tris-HCl, pH 7.5) and applied to a Fast Flow Q-Sepharose column (2 × 10 cm). The column was washed with 5–10 volumes of equilibration buffer and bound protein was eluted in a linear gradient prepared from equilibration buffer and equilibration buffer containing 0.6 M NaCl. S100P-containing fractions eluted at ~150 mM NaCl. S100P-containing fractions were pooled, dialyzed against water, and lyophilized. Lyophilized protein was dissolved in 50 mM Tris-HCl, pH 7.5, 100 mM KCl. β-Mercaptoethanol was added to final concentration of 2% followed by incubation at room temperature for 30 min. Reduced S100P was applied to a Sephadex G-75 column (2.5 × 100 cm) equilibrated with 50 mM Tris-HCl, pH 7.5, 100 mM KCl. The S100P-containing fractions were pooled. S100P had an apparent homogeneity of better than 95%, as judged from Coomassie staining of SDS polyacrylamide gels. This two column purification scheme allows for the rapid isolation of ~150 mg electrophoretically pure protein from 1 L of culture.

Purified recombinant S100P was subject to 23 N-terminal sequencing cycles. The analysis shows that the obtained N-terminal sequence is identical to that determined from the DNA sequence (Fig. 1).

Figure 2A shows the far-UV CD spectrum of recombinant S100P protein in the absence of divalent cations. The spectrum exhibits ellipticity minima at 222 and 208 nm and an ellipticity maximum at 195 nm. Far-UV CD profiles with these characteristics are indicative of predominant α-helical protein secondary structure (Yang et al., 1986; Kuwajima, 1995; Woody, 1995). Analysis of the averaged far-UV CD spectrum using the algorithm of Provencher and Glockner (1981) gave an estimate of (65 ± 7)% α-helix, (4 ± 8)% β-sheet, (10 ± 4)% β-turn, and (25 ± 6)% other. These estimates are close to the fraction of helical structure in two homologous proteins of the S100 family for which three-dimensional structure is known. S100B has 50% sequence identity with S100P and 66% of residues involved in the formation of α-helix (Drohat et al., 1996; Kilby et al., 1996). Rat calcyclin has 35% sequence identity with S100P and has 58% of residues in α-helical conformation

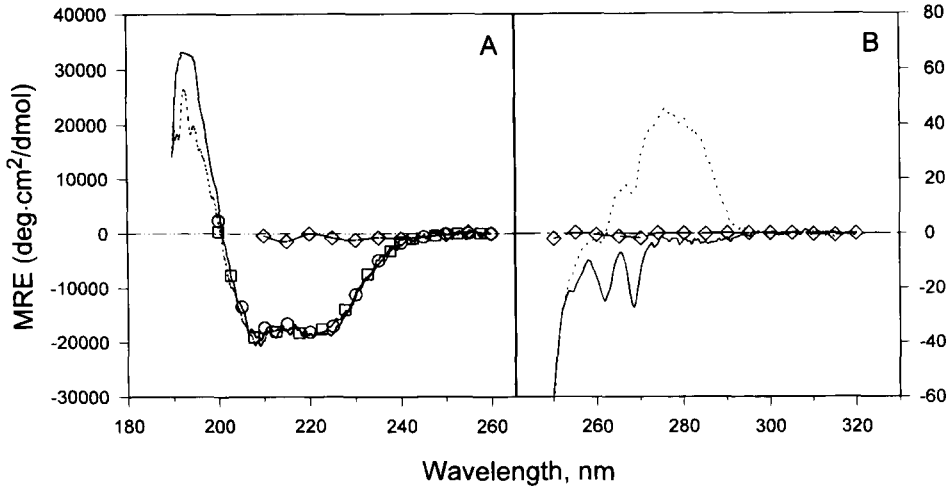

Fig. 2. CD spectra of the recombinant S100P protein. **A:** Far-UV CD spectra in 50 mM Tris/HCl, 0.2 mM EDTA, 1 mM DTT, pH 7.5, in the absence (apo-S100P; solid line, 0.94 mg/mL; circles, 0.17 mg/mL) of Ca^{2+} and in the presence of 5 mM $CaCl_2$ (holo-S100P; dotted line, 0.94 mg/mL; squares, 0.17 mg/mL), and in the presence of 6 M GdmCl (dashed line and diamonds). Each presented spectrum is an average of six independent measurements (calculated SD, 7%). **B:** Near-UV CD spectra in 50 mM Tris/HCl, 0.2 mM EDTA, 1 mM DTT, pH 7.5, in the absence (apo-S100P; solid line) of Ca^{2+} and in the presence of 5 mM $CaCl_2$ (holo-S100P; dotted line), and in the presence of 6 M GdmCl (dashed line and diamonds). Each presented spectrum is an average of two independent measurements at S100P concentration of 4.5 and 3.7 mg/mL. Spectrum obtained at 10 times lower concentration (not shown) is virtually superimposable. CD measurements of the far- and near-UV spectra were performed on JASCO J-20 spectropolarimeter equipped with photoelastic modulator and lock-in amplifier. Signal from the lock-in amplifier was passed through microvolt ammeter (Kethley/Metrabyte) and collected on a PC computer via a 12-bit A/D board. Electrical signal was calibrated in ellipticity units using known values for the ellipticity of aqueous solution of d-(+)-10-camphorsulphonic acid (Yang et al., 1986). Quartz cells with 1-mm or 0.1-mm (far-UV) and 10-mm (near-UV) light pathlength were used. Constant temperature of the protein solution (25 °C) was maintained using a circulating water bath. Protein concentration was determined spectrophotometrically using an extinction coefficient of $E^{0.1\%,280nm}_{1cm} = 0.292$, calculated according to Gill and von Hippel (1989) as described by Pace et al. (1995). Light scattering was corrected according to Winder and Gent (1971).

(Potts et al., 1995). Addition of 6 M guanidinium hydrochloride (GdmCl) leads to a dramatic decrease in the magnitude of ellipticity (Fig. 2A), indicating complete unfolding of S100P (Privalov et al., 1989; Kuwajima, 1995; Woody, 1995).

The near-UV CD spectrum of recombinant S100P in the absence of divalent cations is shown in Figure 2B. The spectrum has characteristic ellipticity minima at 268 and 262 nm, negative peaks that likely correspond to the ^1L_b electronic transitions in phenylalanine (Phe) residues (Strickland, 1974). The existence of such peaks can be interpreted to indicate that at least some of phenylalanine residues of S100P are buried in specific and tightly packed environments. It was proposed that minima at 268 and 263 nm can serve as markers for the native fold (Pain, 1996). This appears to be the case for the recombinant S100P as evidenced from the effect of 6 M GdmCl on the near-UV CD spectrum (Fig. 2B).

Similar effects of 6 M GdmCl were observed by fluorescence spectroscopy. Figure 3A shows emission spectra of S100P in the absence and presence of 6 M GdmCl. In the absence of denaturant, the emission spectra has a maximum at 303 nm, typical of tyrosine fluorescence (Lakowitcz, 1983). Addition of 6 M GdmCl leads to an increase in the emission spectrum intensity and confirms the results of CD spectroscopy that the recombinant S100P possesses both secondary and tertiary structure, i.e., is uniquely folded.

Analysis of the amino acid sequence of S100 family of proteins indicates that there are two putative EF-hand motifs (Hilt & Kligman, 1991). Indeed, Ca^{2+}-binding properties have been demonstrated for a number of S100 proteins (Schafer & Heizmann, 1996), and their functional importance has been proposed. Ca^{2+}-binding is the only known function of S100P. To monitor the ability of recombinant S100P to bind Ca^{2+} and the effect of Ca^{2+} on S100P structure, we recorded CD and fluorescence spectra of S100P in the presence of saturating concentrations of Ca^{2+} (5 mM). Figure 2 shows the effect of Ca^{2+} on the far-UV CD spectrum of recombinant S100P. Spectra in the presence (holo-S100P) and absence (apo-S100P) of Ca^{2+} are indistinguishable within the experimental error. In contrast, both the near-UV CD spectrum and the intrinsic tyrosine fluorescence intensity spectrum were significantly affected by addition of Ca^{2+} (Figs. 2, 3).

The addition of Ca^{2+} leads to an appearance of major peak at 277 nm and of minor peak at 283 nm in the near-UV CD spectrum. This is consistent with the ^1L_b electronic transitions in tyrosine residues (e.g., Strickland, 1974). Characteristic ellipticity minima at 263 and 268 nm observed in the absence of Ca^{2+} were less intense in the presence of Ca^{2+}. We interpret these dramatic changes in the near-UV CD spectra of S100P as changes in the tertiary structure of S100P upon binding of Ca^{2+}. This involves the burial of tyrosine residues as evident from the appearance of the 277 and 283 nm peaks, and is further supported by the changes in Tyr-fluorescence emission (Fig. 3) and UV absorbance (data not shown) spectra.

To gain more low-level structural detail on the effect of Ca^{2+} on the structure of the recombinant S100P, we followed the changes in the fluorescence intensity of the external probe, 8-anilino-1-naphthalenesulfonic acid (ANS). Fluorescence intensity of ANS changes upon binding to nonpolar protein surfaces, a property that has been used widely to study protein folding and unfolding reac-

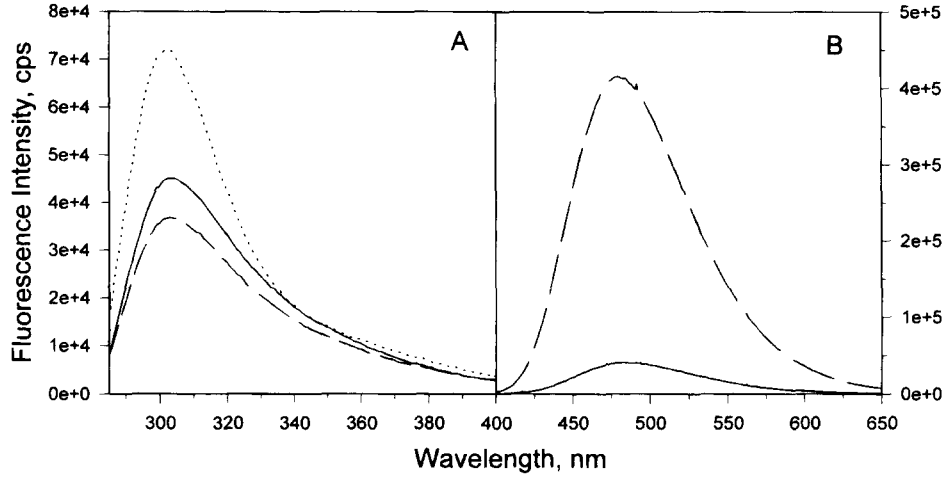

Fig. 3. Fluorescence emission spectra of S100P solution at 25°C in 50 mM Tris/HCl, 0.2 mM EDTA, 1 mM DTT, pH 7.5, in the absence (apo-S100P, solid line) and in the presence of 5 mM $CaCl_2$ (holo-S100P, dashed line), and in the presence of 6 M GdmCl (dotted line). **A:** Tyrosine emission spectra after excitation at 282 nm. **B:** ANS (40 μM) emission spectra after excitation at 355 nm. Emission spectra were recorded on a FluoroMax spectrofluorometer with DM3000F software in 1-cm cells. Protein concentration in all cases was 0.14 mg/mL.

tions (e.g., Goto et al., 1979; Teschke et al., 1993; Jones et al., 1994). Figure 3B presents fluorescence emission spectra of ANS in a solution containing S100P with or without Ca^{2+}. The large increase in ANS fluorescence intensity upon the addition of Ca^{2+} indicates structural rearrangements in the S100P molecule in agreement with the changes in the near-UV CD. Structural changes in the S100P molecule upon addition of Ca^{2+} ions are in accord with the observed changes in the structures of apo and Ca^{2+} forms of calmodulin. It has been shown by Finn et al. (1995) that addition of Ca^{2+} ions to calmodulin causes rearrangements of the secondary structure elements leading to exposure of hydrophobic surface, without, however, significant perturbations of the amount of secondary structure.

With a high yield bacterial expression system and rapid protein purification scheme in hand, the production of isotopically labeled protein to facilitate structure determination by NMR spectroscopy becomes practical. Large changes in the signal of near-UV CD and fluorescence emission spectra, on the other hand, provide important tools for the analysis of Ca^{2+}-binding affinity of human S100P. Studies involving site-directed mutagenesis to determine the structure–function relationships of this protein are under way.

Acknowledgments: We thank Dr. Lydia Averboukh for the S100P cDNA, Drs. Norman Meadow and Nemat Keyhani for pVEX vector, and Dr. James Harman for critical review and fruitful discussions. This work was supported in part by the Petroleum Research Fund, TTU Institute for Biotechnology minigrant, and the South Plains Foundation (to G.I.M.). M.M.L. is supported by a postdoctoral fellowship from the Ministry of Education of Spain. J.M.R. is Howard Hughes Medical Institution undergraduate fellow.

References

Allore R, O'Hanlon D, Price R, Neilson K, Willard HF, Cox DR, Marks A, Dunn RJ. 1988. Gene encoding the beta subunit of S100 protein is on chromosome 21: Implications for Down syndrome. *Science 239*:1311–1313.

Averboukh L, Liang P, Kantoff PW, Pardee AB. 1996. Regulation of S100P expression by androgen. *Prostate 29*:350–354.

Baudier J, Glasser N, Gerard D. 1986. Ions binding to S100 proteins. I. Calcium- and zinc-binding properties of bovine brain S100aa S100a (ab) and S100b (bb) protein: Zn^{2+} regulates Ca^{2+} binding on S100b protein. *J Biol Chem 261*:8192–8203.

Becker T, Gerke V, Kube E, Weber K. 1992. S100P, a novel Ca(2+)-binding protein from human placenta. *Eur J Biochem 207*:541–547.

Davies BR, Davies MP, Gibbs FE, Barraclough R, Rudland PS. 1993. Induction of the metastatic phenotype by transfection of a benign rat mammary epithelial cell line with the gene for p9Ka a rat calcium-binding protein but not with the oncogene EJ-ras-1. *Oncogene 8*:999–1008.

Drohat AC, Amburgey JC, Abildgaard F, Starich MR, Baldisseri D, Weber DJ. 1996. Solution structure of rat apo-S100B (beta beta) as determined by NMR spectroscopy. *Biochemistry 35*:11577–11588.

Emoto Y, Kobayashi R, Akatsuka H, Hidaka H. 1992. Purification and characterization of a new member of the S-100 protein family from human placenta. *Biochem Biophys Res Commun 182*:1246–1253.

Finn BE, Evenas J, Drakenberg T, Waltho JP, Thulin E, Forsen S. 1995. Calcium-induced structural changes and domain autonomy in calmodulin. *Nature Struct Biol 2*:777–783.

Gill SC, von Hippel PH. 1989. Calculation of protein extinction coefficients from amino acid sequence data. *Anal Biochem 182*:319–326.

Goto Y, Azuma T, Hamaguchi K. 1979. Refolding of the immunoglobulin light chain. *J Biochem (Tokyo) 85*:1427–1438.

Griffin WS, Yeralan O, Sheng JG, Boop FA, Mrak RE, Rovnaghi CR, Burnett BA, Feoktistova A, Van Eldik LJ. 1995. Overexpression of the neurotrophic cytokine S100 beta in human temporal lobe epilepsy. *J Neurochem 65*:228–33.

Hilt DC, Kligman D. 1991. The S-100 protein family: A biochemical and functional overview. In: Heizmann CW, ed. *Novel calcium binding proteins*. Berlin/Heidelberg: Springer-Verlag. pp 65–104.

Hoffmann HJ, Olsen E, Etzerodt M, Madsen P, Thogersen HC, Kruse T, Celis JE. 1994. Psoriasin binds calcium and is upregulated by calcium to levels that resemble those observed in normal skin. *J Invest Dermatol 103*:370–375.

Jones BE, Jennings PA, Pierre RA, Matthews CR. 1994. Development of non-polar surfaces in the folding of *Escherichia coli* dihydrofolate reductase detected by 1-anilinonaphthalene-8-sulfonate binding. *Biochemistry 33*: 15250–15258.

Kilby PM, Van Eldik LJ, Roberts GC. 1996. The solution structure of the bovine S100B protein dimer in the calcium-free state. *Structure 4*:1041–1052.

Kuwajima K. 1995. Circular dichroism. *Methods Mol Biol 40*:115–135.

Lakowitcz JR. 1983. *Principles of fluorescence spectroscopy*. New York: Plenum Press.

Marshak DR. 1990. S100 beta as a neurotrophic factor. *Prog Brain Res 86*:169–181.

Marshak DR, Pena LA. 1992. Potential role of S100 beta in Alzheimer's disease: An hypothesis involving mitotic protein kinases. *Prog Clin Biol Res 379*:289–307.

Marshak DR, Pesce SA, Stanley LC, Griffin WS. 1992. Increased S100 beta

neurotrophic activity in Alzheimer's disease temporal lobe. *Neurobiol Aging 13*:1–7.

Pace CN, Vajdos F, Fee L, Grimsley G, Gray T. 1995. How to measure and predict the molar absorption coefficient of a protein. *Protein Sci 4*:2411–2423.

Pain R. 1996. Determining the CD spectrum of a protein. In: Coligan JE, Dunn B, Ploegh HL, Speicher DW, Wingfield PT, eds. *Current protocols in protein science*. John Wiley & Sons. pp 761–723.

Potts BCM, Smith J, Akke M, Macke TJ, Okazaki K, Hidaka H, Case DA, Chazin WJ. 1995. The structure of calcyclin reveals a novel homodimeric fold for S100 Ca^{2+}-binding proteins. *Nature Struct Biol 2*:790–796.

Presper KA, Wong CY, Liu L, Meadow ND, Roseman S. 1989. Site-directed mutagenesis of the phosphocarrier protein IIIGlc a major signal-transducing protein in *Escherichia coli*. *Proc Natl Acad Sci USA 86*:4052–4055.

Privalov PL, Tiktopulo EI, Venyaminov Syu, Griko YuV, Makhatadze GI, Khechinashvili NN. 1989. Heat capacity and conformation of proteins in the denatured state. *J Mol Biol 205*:737–750.

Provencher SW, Glockner J. 1981. Estimation of globular protein secondary structure from circular dichroism. *Biochemistry 20*:33–37.

Renaud W, Merten M, Figarella C. 1994. Increased coexpression of CFTR and S100 calcium binding proteins MRP8 and MRP14 mRNAs in cystic fibrosis human tracheal gland cells. *Biochem Biophys Res Commun 201*:1518–1525.

Schafer BW, Heizmann CW. 1996. The S100 family of EF-hand calcium-binding proteins: Functions and pathology. *Trends Biochem Sci 21*:134–140.

Strickland EH. 1974. Aromatic contributions to circular dichroism spectra of proteins. *CRC Crit Rev Biochem 2*:113–175.

Teschke CM, King J, Prevelige PE Jr. 1993. Inhibition of viral capsid assembly by 1 1'-bi(4-anilinonaphthalene-5-sulfonic acid). *Biochemistry 32*:10658–10665.

Weterman MA, van Muijen GN, Bloemers HP, Ruiter DJ. 1993. Expression of calcyclin in human melanocytic lesions. *Cancer Res 53*:6061–6066.

Winder AF, Gent WLC. 1971. Correction of light-scattering errors in spectrophotometric protein determinations. *Biopolymers 10*:1243–1251.

Woody RW. 1995. Circular dichroism. *Methods Enzymol 246*:34–71.

Yang JT, Wu CS, Martinez HM. 1986. Calculation of protein conformation from circular dichroism. *Methods Enzymol 130*:208–226.