Dissection of a Human Septin: Definition and Characterization of Distinct Domains within Human SEPT4

Wanius Garcia,*,‡ Ana Paula Ulian de Araújo,‡ Mario de Oliveira Neto,‡ Michel R. M. Ballestero,‡ Igor Polikarpov,‡ Manami Tanaka,§ Tomoo Tanaka,|| and Richard Charles Garratt*,‡

Centro de Biotecnologia Molecular e Estrutural (CBME), Instituto de Física de São Carlos (IFSC), Universidade de São Paulo (USP), Av. Trabalhador São Carlense 400, centro, Box 369, São Carlos, SP, 13560-970, Brazil, National Institute of Advanced Industrial Science and Technology, Higashi, Tsukuba Science City, Ibaraki 305-8565, Japan, and Tokai University School of Medicine, Isehara, Kanagawa, 259-1193, Japan

Received July 31, 2006

ABSTRACT: The septins are a conserved family of guanosine-5′-triphosphate (GTP)-binding proteins. In mammals they are involved in a variety of cellular processes, such as cytokinesis, exocytosis, and vesicle trafficking. Specifically, SEPT4 has also been shown to be expressed in both human colorectal cancer and malignant melanoma, as well as being involved in neurodegenerative disorders. However, many of the details of the modes of action of septins in general remain unclear, and little is known of their detailed molecular architecture. Here, we define explicitly and characterize the domains of human SEPT4. Regions corresponding to the N-terminal, GTPase, and C-terminal domains as well as the latter two together were successfully expressed in *Escherichia coli* in soluble form and purified by affinity and size-exclusion chromatographies. The purified domains were analyzed by circular dichroism spectroscopy, fluorescence spectroscopy, dynamic light scattering, and small-angle X-ray scattering, as well as with bioinformatics tools. Of the three major domains that comprise SEPT4, the N-terminal domain contains little regular secondary structure and may be intrinsically unstructured. The central GTPase domain is a mixed α/β structure, probably based on an open β sheet. As defined here, it is catalytically active and forms stable homodimers *in vitro*. The C-terminal domain also forms homodimers and can be divided into two regions, the second of which is α-helical and consistent with a coiled-coil structure. These studies should provide a useful basis for future biophysical studies of SEPT4, including the structural basis for their involvement in diseases such as cancer and neurodegenerative disorders.

Septins were first identified in *Saccharomyces cerevisiae* as proteins required for the completion of the cell cycle (*1*). Septins purified from yeast form filaments of variable length that are 7−9 nm in diameter. However, their mechanism of assembly is still not fully elucidated (*2*). Similar filaments have also been isolated from *Drosophila* and mammalian brain tissue (*3*, *4*). At least 12 mammalian septins have been identified, and a recent paper suggests the utilization of a generic nomenclature (*5*). The sequences of all members can be divided into three domains: a variable N-terminal domain, a central GTPase domain, and a C-terminal region that generally includes sequences characteristic of coiled coils (*6*).

The central GTPase domain is reasonably well-conserved (>35% sequence identity between yeast and mammalian septin homologues, for example), but this is not the case for the N- and C-terminal domains, which vary greatly both within and between species. However, the level of sequence conservation among GTPase domains of septins from the same species is considerably higher, with a minimum sequence similarity of 70% (*7*). The GTPase domain is characterized by the presence of a P-loop motif (G1 or Walkers's A box) close to its N terminus (*8*). Furthermore, additional GTPase motifs (G3 and G4) are also frequently present, both of which are directly associated with substrate binding in small GTPases (*8*, *9*). Both the binding of GTP and its hydrolysis have been experimentally demonstrated for several septins *in vitro* (*10*−*12*).

In contrast, little is known about the highly variable N-terminal domain in terms of both its structure and biological role. This domain varies considerably from one septin to another, particularly in terms of its length. For example, the N-terminal domains of human SEPT3 and SEPT4 are 24 and 119 residues in length, respectively. Although the function of this domain is still largely obscure, its variability from one septin to another is suggestive of a role in determining selectivity.

The C-terminal domains of many lower eukaryotic and mammalian septins present typical sequences characteristic of left-handed coiled coils (*6*, *11*, *13*). These are expected to be involved in protein−protein interactions, which may be important for the formation of homo- or heteroligomers. However, there is evidence that the oligomerization of septins and particularly the formation of filaments are based on interactions that also involve the GTPase domain (*11*, *13*).

* To whom correspondence should be addressed. E-mail: wanius@if.sc.usp.br (W.G.); richard@if.sc.usp.br (R.C.G.).
‡ Universidade de São Paulo (USP).
§ National Institute of Advanced Industrial Science and Technology.
|| Tokai University School of Medicine.

10.1021/bi061549z CCC: $33.50
Published on Web 10/26/2006

The association of some septins with membranes has been attributed to a well-characterized polybasic sequence residing between the N-terminal domain and the GTPase, typically about 9 residues prior to the P loop. This polybasic sequence is known to bind phosphatidylinositol 4,5-bisphosphate [PtdIns(4,5)P_2] and probably confers membrane-binding capability as reported for the mouse SEPT4 (*10*). During spermatogenesis, mouse SEPT4 was essential for correct mitochondrial architecture and in establishing the annulus within the tail region of sperm (*14*). These data may also be indicative of a role for SEPT4 that is dependent upon specific protein−membrane interactions.

The *SEPT4* gene resides at 17q23 and is physiologically expressed mainly in normal cerebral tissue (*15*, *16*). Interestingly, however, under certain pathological conditions, human SEPT4 is also expressed in colorectal and urologic cancer and malignant melanoma (*17*). This unique pathology-related expression profile has never been observed among other mammalian septin family members. A recent paper showed that the impaired expression of human SEPT4 directly affected cell division and caused selective G2 arrest in cultivated human cancer cells (*17*). This is consistent with a role for SEPT4 in cytokinesis, as has been reported for other septin family members (*18*, *19*). Furthermore, SEPT4 together with SEPT1 and SEPT2 are accumulated in tau-based filamentous deposits known as neurofibrillary tangles and glial fibrils in Alzheimer's disease (*20*). The SEPT4 protein is also involved in the formation of cytoplasmic inclusions as well as induction of cell death in α-synuclein-associated neurodegenerative disorders (*21*). Clearly, further studies of the detailed mechanisms of action of SEPT4 are relevant for a complete understanding of a series of important physiological and pathological processes.

Despite these considerable advances in the understanding of the cell biology of different members of the septin family, there has been relatively little progress concerning their detailed structural properties. Indeed, more is probably known about the ultrastructure of the heterpolymeric fibers, which are formed *in vivo* and *in vitro*, than about the atomic structure of any individual septin. Certainly, one of the factors limiting progress in this field is the instability of recombinantly expressed septins and their tendency to aggregate (*11*). One approach to this problem is to accurately define and express separately the individual structural domains (or pairs thereof) in an attempt to generate products that are both more stable and soluble when expressed in *Escherichia coli*.

In the present study, the full-length form of human SEPT4 and its individual N-terminal, GTPase, and C-terminal domains (SEPT4-N, SEPT4-G, and SEPT4-C, respectively) as well as the latter two together in a single construct (SEPT4-GC) were successfully defined. They were all subsequently expressed in *E. coli* in soluble form and purified by a two-step protocol using both affinity and size-exclusion chromatography (SEC). The structural integrity of the purified products was analyzed by native gel electrophoresis, circular dichroism (CD) spectroscopy, fluorescence spectroscopy, dynamic light scattering (DLS), and small-angle X-ray scattering (SAXS) and compared with the results of computational sequence analysis. The catalytic activity of the SEPT4-G product was also demonstrated.

MATERIALS AND METHODS

Materials. The bacterial expression vector pGEX-5X-1, Glutathione Sepharose 4B, and Factor Xa were purchased from Amersham Pharmacia Biotech. The bacterial expression vector pET28a(+) and nickel−nitrilotriacetic acid (Ni−NTA) resin were purchased from NOVAGEN. Restriction endonucleases, isopropyl-β-D-thiogalactopyranoside (IPTG), and T4 DNA ligase were obtained from Invitrogen, Inc. Guanosine-5′-triphosphate (GTP), polyethyleneimine (PEI)-cellulose for thin-layer chromatography (TLC), and protein standards used as sodium dodecyl sulfate−polyacrylamide gel electrophoresis (SDS−PAGE) markers were from Sigma Aldrich. Guanosine-5′-[γ-^{32}P]triphosphate ([γ-^{32}P]GTP, 5000 Ci/mmol) was purchased from Amersham Pharmacia Biotech. Native gels for electrophoresis [8−25% (w/v) gradient polyacrylamide], Superdex-200, and Superdex-75 resin were purchased from Amersham Pharmacia Biotech. All other chemicals were of analytical grade.

Computational Methods for Domain Definitions and Structure Prediction. The N-terminal domain (SEPT4-N) was defined as the common sequence between two splice variants of SEPT4, namely, isoform 1 (NCBI accession number NP 536342) and isoform 4 (NP 004565) (*16*). To define the boundaries of the GTPase domain, the human SEPT4 (SEPT4-G) sequence was initially aligned with a further eight human septin sequences using Clustal-X (*22*) and subsequently compared with an alignment of nine small GTPases of known structure. The P-loop motif, clearly recognizable in all cases, was taken as a fixed reference point for both alignments and used to identify a four-residue hydrophobic stretch characteristic of the N terminus of small GTPases. The corresponding position in SEPT4 was taken as the definition of the beginning of the SEPT4-G domain. This resulted in a gap of 24 residues (including the polybasic region) between the C terminus of the N-terminal domain and the N terminus of the GTPase domain. The C terminus of the SEPT4-G domain was defined with reference to a CAAX motif (where A is an aliphatic residue and X is a serine/leucine/methionine) (*23*). According to this definition, the C-terminal domain contains only 62 residues. However, uncertainties in the exact definition of the C terminus of the SEPT4-G domain suggested that the expression of the C-terminal domain in conjunction with the preceding GTPase domain, in the form of a single molecule, would also be of value.

The methods used for general secondary-structure prediction were PHD (*24*), PSIPRED (*25*), and SSpro (*26*). Paircoil (*27*) was used for coiled-coil prediction, and IUPred (*28*) was used for the identification of intrinsically unstructured regions.

Plasmid Construction. Recombinant DNA techniques were performed using conventional protocols. cDNA corresponding to the full-length SEPT4 and the predicted domains was amplified by polymerase chain reaction (PCR) using a MJ Research, Inc. thermal cycler PTC-100, using SEPT4 as a DNA template (*16*). The DNA encoding the SEPT4, N-terminal domain (SEPT4-N), GTPase domain (SEPT4-G), C-terminal domain (SEPT4-C), and GTPase domain plus C-terminal domain (SEPT4-GC) were amplified, producing five fragments of around 1500, 360, 820, 1000, and 200 bp, respectively. In the case of SEPT4, the amplification product

Table 1: Plasmids Used in This Study

plasmid designation	plasmid construction	residues
pGST-SEPT4	SEPT4	1−478, 478 amino acids
pET-SEPT4-N	SEPT4-N	1−119, 119 amino acids
pET-SEPT4-G	SEPT4-G	144−416, 273 amino acids
pET-SEPT4-GC	SEPT4-GC	144−478, 335 amino acids
pET-SEPT4-C	SEPT4-C	417−478, 62 amino acids

and the expression vector pGEX-5X-1 were digested with *Bam*HI and *Xho*I and used to transform *E. coli* DH5α after ligation with T4 DNA ligase. The new vector construct was named pGST-SEPT4 and produced the target protein in fusion with GST. In the remaining cases, both the amplified products and the expression vector pET28a(+) were digested with *Nde*I and *Bam*HI and used to transform the *E. coli* DH5α strain after ligation with T4 DNA ligase. Transformants were confirmed by PCR and restriction analysis. The new vector constructs were named pET-SEPT4-N, pET-SEPT4-G, pET-SEPT4-C, and pET-SEPT4-GC, producing the SEPT4-N, SEPT4-G, SEPT4-C, and SEPT4-GC proteins fused to a His tag. All plasmids were sequenced by the dideoxy chain termination method (*29*) using an ABI Prism 377 automated DNA sequencer (Perkin-Elmer) following the protocol of the manufacturer. Table 1 shows a summary of the plasmids used in this study.

Expression and Purification of Recombinant SEPT4, SEPT4-N, SEPT4-G, SEPT4-C, and SEPT4-GC. The expression plasmid pGST-SEPT4 was used to transform *E. coli* AD202-competent cells. A total of 500 μL of an overnight culture of *E. coli* AD202 pGST-SEPT4 was inoculated into 500 mL of fresh LB medium containing ampicilin (100 μg/mL) and kanamycin (50 μg/mL). The culture was grown at 37 °C to mid-log phase ($OD_{600\ nm} = 0.8$). At this point, the culture was induced with 0.2 mM IPTG, and the incubation was continued for an additional 16 h at 20 °C. After incubation, the culture was immediately harvested by centrifugation at 6000*g* for 20 min at 4 °C. The pelleted cells were resuspended in 50 mM Tris-HCl at pH 8.0 buffer containing 25 mM NaCl, 5% glycerol, 0.1 mM GTP, 1 mM PMSF, 10 μM Leupeptin, and 1 μM Pepstatin and disrupted by sonication. The suspension was then centrifuged at 20000*g* for 20 min at 4 °C to separate the cell debris. The supernatant containing recombinant GST-SEPT4 was gently mixed with Glutathione Sepharose 4B Fast Flow (2.0 mL) at 4 °C for 20 min. After incubation, the resin was packed into a column and washed several times with 50 mM Tris-HCl at pH 8.0, 25 mM NaCl, 5% glycerol, and 2 mM $CaCl_2$. The recombinant fusion protein was cleaved with the Factor Xa protease (50 units) while bound to the column at 10 °C. The released SEPT4 was eluted, pooled, and loaded onto a Superdex-200 column. Elution was carried out with 25 mM Tris-HCl, 25 mM NaCl, 0.1 mM ethylenediaminetetraacetic acid (EDTA), and 5% glycerol at pH 8.0, and the collected fractions were analyzed by 15% SDS−PAGE.

The expression plasmids carrying the SEPT4-N, SEPT4-G, SEPT4-C, and SEPT4-GC gene fragments were used to transform *E. coli* strain BL21(DE3)-competent cells. A total of 500 μL of an overnight culture of *E. coli* BL21(DE3) harboring the pET-SEPT4-N or pET-SEPT4-C plasmid was inoculated into 500 mL of fresh LB medium containing kanamycin (50 μg/mL). The culture was agitated at 37 °C until attaining an optical density between 0.5 and 0.6 at $\lambda = 600$ nm. Subsequently, IPTG was added to a final concentration of 0.5 mM. The incubation was prolonged for 4 h at 37 °C (or 10 h at 20 °C in the case of SEPT4-C). After centrifugation, the pelleted cells were resuspended in PBS buffer (pH 7.4) containing 1.0 mM PMSF and disrupted by the addition of 0.1 mg/mL lysozyme for 30 min at 4 °C followed by sonication. The suspension was then centrifuged at 20000*g* for 20 min at 4 °C, and both the pellet and supernatant were analyzed on SDS−PAGE to check the solubility of the recombinant fusion protein. The supernatant containing the recombinant SEPT4-N or SEPT4-C domain was applied to a nickel-affinity column equilibrated with PBS buffer. After the unbound proteins were washed out with PBS buffer (pH 7.4) containing 10 mM imidazole, the recombinant proteins were eluted from the column with the same buffer containing 100 mM imidazole. The recombinant SEPT4-N or SEPT4-C was applied to a Superdex-75 gel-filtration column at 4 °C. Elution was carried out with the buffer containing 25 mM Tris-HCl and 20 mM NaCl at pH 7.8, and the fractions were analyzed by 15% SDS−PAGE.

A total of 500 μL of an overnight culture of *E. coli* BL21-(DE3) harboring the pET-SEPT4-G (or pET-SEPT4-GC) plasmid was inoculated into 500 mL of fresh LB medium containing kanamycin (50 μg/mL). The culture was incubated at 37 °C while shaking until attaining an optical density of 0.5 at $\lambda = 600$ nm. Subsequently, IPTG was added to a final concentration of 0.4 mM. The incubation was prolonged for a further 10 h at 20 °C. After centrifugation, the pelleted cells were resuspended in buffer containing 25 mM Tris-HCl at pH 7.8, 20 mM NaCl, 1.0 mM PMSF, 5% glycerol, 0.1 mM GTP, and 0.05% tween and disrupted by the addition of 0.1 mg/mL lysozyme for 30 min at 4 °C followed by sonication. The suspension was then centrifuged at 20000*g* for 20 min at 4 °C, and the pellet and supernatant were analyzed on SDS−PAGE to check the solubility of the recombinant fusion protein. The supernatant containing the recombinant SEPT4-G (or SEPT4-GC) domain was applied to a nickel-affinity column equilibrated with the same buffer. After the unbound proteins were washed out with the same buffer containing 5 mM imidazole, the recombinant protein was eluted from the column by increasing the imidazole concentration to 100 mM. The recombinant SEPT4-G (or SEPT4-GC) was then applied to a Superdex-200 size-exclusion column. Elution was carried out at 4 °C using a modified buffer containing 25 mM Tris-HCl at pH 7.8, 20 mM NaCl, 5% glycerol, and 0.1 mM EDTA, and the fractions were analyzed by 15% SDS−PAGE.

Protein Concentration. The concentrations of the recombinant proteins were determined by UV absorbance at $\lambda = 280$ nm (spectrophotometer U-2001 Hitachi) using a theoretical extinction coefficient based on the amino acid composition (*30*). The theoretical extinction coefficients employed were $\epsilon_{280} = 45\,570\ M^{-1}\ cm^{-1}$ for SEPT4, $\epsilon_{280} = 21\,030\ M^{-1}\ cm^{-1}$ for SEPT4-N, $\epsilon_{280} = 21\,980\ M^{-1}\ cm^{-1}$ for SEPT4-G, $\epsilon_{280} = 1490\ M^{-1}\ cm^{-1}$ for SEPT4-C, and $\epsilon_{280} = 24\,785\ M^{-1}\ cm^{-1}$ for SEPT4-GC.

Nucleotide Content. The nucleotide content of the purified SEPT4, SEPT4-G, and SEPT4-GC was estimated using the method described by Seckler et al. (*31*). Briefly, cold $HClO_4$ was added to a solution of SEPT4, SEPT4-G, or SEPT4-GC (at a known concentration) to a final concentration of

0.5 M and left on ice for 15 min. The resulting precipitate was removed by centrifugation at 16000*g* at 4 °C for 15 min. The optical density of the supernatant was measured spectrophotometrically, and the nucleotide (GDP and/or GTP) concentration was determined assuming an extinction coefficient of $\epsilon_{254nm} = 12\,400\ M^{-1}\ cm^{-1}$.

CD Spectroscopy. CD spectra were collected using a Jasco J-715 spectropolarimeter. The protein solution was approximately 10 μM in the cases of SEPT4-N and SEPT4-C (in 25 mM Tris-HCl and 20 mM NaCl at pH 7.8) and 9 μM in the cases of SEPT4-G and SEPT4-GC (in 25 mM Tris-HCl at pH 7.8, 20 mM NaCl, 0.1 mM EDTA, and 5% glycerol). All data were collected using a 0.1 cm quartz cuvette at 10 °C, and the spectra were recorded over the wavelength range from 195 to 250 nm and were determined as an average of 32 scans. Spectra were transformed to molar ellipticity (θ) using the mean weight residue and concentration prior to the secondary-structure analysis. To obtain structural information, the CD data were analyzed by three different programs. Spectra were deconvoluted using the SELCON3 (*32*), CONTIN (*33*), and CDS (*34*) programs using different databases.

DLS. Protein solutions (between 0.4 and 1.0 mg/mL) were centrifuged at 16000*g* for 15 min at 4 °C and immediately loaded into a quartz cuvette prior to measurement. In the cases of SEPT4-N and SEPT4-C, measurements were made in 25 mM Tris-HCl and 20 mM NaCl at pH 7.8, while for SEPT4, SEPT4-G, and SEPT4-GC, measurements were made with the addition of 5% glycerol and 0.1 mM EDTA included in the buffer. Light scattering was performed using a DynaPro Molecular Sizing instrument, with DYNAMICS control and analysis software (DYNAMICS, version 5, Proteins Solutions, Inc.). Data collection times of 2 s were used in all cases, for a minimum of 60 acquisitions at 4 °C.

Analysis of SEPT4-G and SEPT4-GC by Native Gel Electrophoresis. SEPT4-G (~30 μM) and SEPT4-GC (~18 μM) were subjected to electrophoresis on an 8−25% (w/v) gradient polycrylamide gel at pH 8.8 using the Phast System (Amersham Bioscience) at 4 °C and stained following standard protocols. Protein standards of known hydrodynamic radii (thyroglobulin, 8.5 nm; ferritin, 6.1 nm; catalase, 5.22 nm; lactate dehydrogenase, 4.4 nm; and bovine serum albumin, 3.55 nm) were electrophoresed under the same conditions. The mobility of the individual bands of the protein standards were plotted as the retardation factor (R_f) versus the Stokes radii (R_h). The linear equation obtained from this calibration was employed to calculate the R_h of SEPT4-G and SEPT4-GC.

Fluorescence Spectroscopy. Fluorescence emission measurements were performed on a steady-state spectrofluorimeter, model K2 ISS, equipped with a refrigerated circulator. The three tryptophans of SEPT4-N and two tryptophans of SEPT4-G and SEPT4-GC were excited at 295 nm at 10 °C. Fluorescence emission spectra were recorded from 305 to 450 nm from samples loaded into a 1.0 cm path-length quartz cuvettte. The SEPT4-N concentration was approximately 5 μM in 25 mM Tris-HCl buffer at pH 7.8 containing 20 mM NaCl. The SEPT4-G and SEPT4-GC concentrations were approximately 6 μM in 25 mM Tris-HCl buffer at pH 7.8 containing 20 mM NaCl, 0.1 mM EDTA, and 5% glycerol.

GTPase Activity Assay. GTP hydrolysis was determined by the appearance of the hydrolysis products of [γ-^{32}P]GTP according to previously described methods (*35*), with slight modifications. The reaction mixture contained 1 μM of SEPT4-G together with 50 nM [γ-^{32}P]GTP (either in the presence or absence of an additional 1 μM unlabeled GTP), 25 mM Tris-HCl at pH 7.8, 20 mM NaCl, 5% glycerol, and 0.5 mM $MgCl_2$ in 25 μL. The reaction was performed at 18 °C for 10 min and was stopped by adding 1 μL of 150 mM EDTA. Subsequently, 1 μL of sample from the reaction mixture was spotted onto a PEI-cellulose TLC plate and was developed by ascending chromatography using 1 M lithium phosphate/0.3 M sodium phosphate (pH 3.8). The TLC plate was visualized by autoradiography. Control experiments were performed in the absence of SEPT4-G or Mg^{2+}.

SAXS Data Collection. For SAXS measurements, the SEPT4-G concentrations used were approximately 3 and 9 mg/mL in 25 mM Tris-HCl buffer at pH 7.8 containing 20 mM NaCl, 0.1 mM EDTA, and 5% glycerol. The samples were centrifuged for 10 min at 13000*g* (4 °C) prior to measurement, which were carried out at the LNLS, Campinas, Brazil, using the synchrotron SAXS beamline (*36*). The wavelength of the incoming monochromatic X-ray beam was $\lambda = 1.488$ Å. A 1D X-ray position-sensitive detector (PSD) was utilized to record the scattering intensity as a function of the modulus of the scattering vector q (here $q = 4\pi \sin \theta/\lambda$, where θ is half the scattering angle). The parasitic scattering from air and beamline windows were subtracted from the total measured intensities. The sample−detector distance (1031.5 mm) was adjusted to record the scattering intensity for q values ranging from 0.01 to 0.25 Å^{-1}. The samples were encapsulated inside a cell with two thin parallel mica windows. Data analysis was performed using the GNOM program package (*37*, *38*). Model simulations were done using DAMMIN (*39*, *40*). At least 10 independent reconstructions starting from random approximations yielded reproducible results.

RESULTS

Domain Definitions of SEPT4. Figure 1A shows schematically the methodology used to define the boundaries of the three SEPT4 domains used in this study. The N-terminal domain was defined from residues 1−119 based on the sequential identity of this region observed for the SEPT4 splice variants known as isoforms 1 and 4. This sequence identity, which terminates abruptly after residue 119, was taken to be indicative of a possible independent domain, here defined as SEPT4-N.

Parts B and C of Figure 1 show regions of the multiple-sequence alignment used to define the N and C termini of the GTPase domain. The crystal structures of small GTPases showed the presence of a consensus stretch of four hydrophobic residues corresponding to the N-terminal β strand, immediately prior to the P loop. A similar pattern was observable in all human septin sequences and was used as a definition of the N terminus of the GTPase domain, SEPT4-G. This definition leaves a gap of 24 residues between the end of SEPT4-N and the beginning of SEPT4-G, which encompasses the polybasic region. The C terminus of SEPT4-G was defined in reference to the CAAX motif shown in Figure 1C. The resulting definition of SEPT4-G encompasses residues 144−416. SEPT4-GC therefore runs continuously from residue 144 to 478.

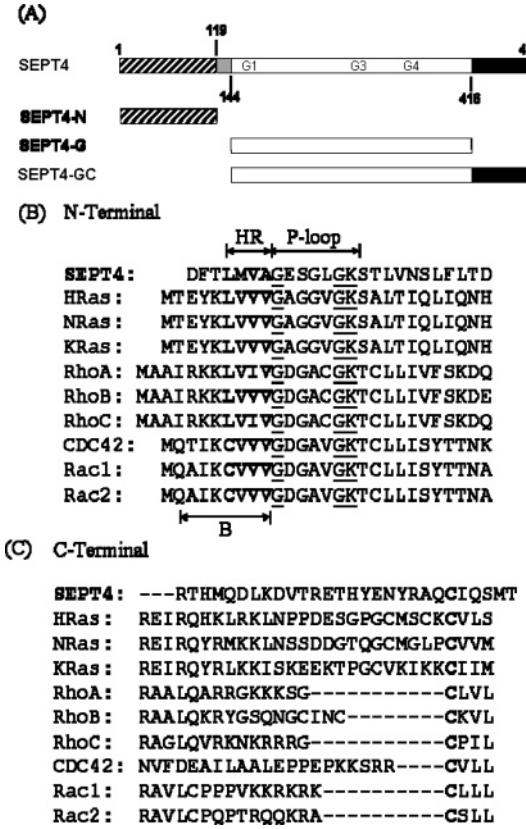

FIGURE 1: Domain structure of human SEPT4. (A) Human SEPT4 is characterized by a large N-terminal domain (residues 1−119, SEPT4-N), a central GTPase domain (residues 144−416, SEPT4-G), and a C-terminal domain including a coiled coil (residues 417−478, SEPT4-C). SEPT4-N is followed by a linker sequence of 24 residues prior to the N terminus of SEPT4-G. (B) Multiple-sequence alignment of the N-terminal region of SEPT4-G with the similar region observed in small GTPases of known three-dimensional structure. The P loop is indicated. HR refers to the hydrophobic region observed within the N-terminal β strand of small GTPases (B). (C) Multiple-sequence alignment of the C-terminal region of SEPT4-G with the corresponding region in small GTPases. C indicates the conserved cystine residue of the CAAX motif seen in small GTPases. The NCBI accession numbers are CAG47067 (H-Ras), P01111 (N-Ras), P01116 (K-Ras), NP001655 (RhoA), CAA68228 (RhoB), P08134 (RhoC), P25763 (CDC42), AAM21111 (Rac1), and P15153 (Rac2).

Table 2: Secondary-Structure Prediction[a]

domains	α helix (%)	β sheet (%)	coil (%)
SEPT4	30−36	11−17	53−54
SEPT4-N	0−19	0−8	81−92
SEPT4-G	36−40	19−25	37−41
SEPT4-C	48−58	0−0	42−52

[a] The computational methods PHD, PSIPRED, and SSpro are compared. The maximum and minimum values obtained for each type of secondary structure are given.

The results of the secondary-structure prediction for the complete SEPT4 sequence and its individual domains are shown in Table 2. The results indicate that overall SEPT4 is composed of roughly equal percentages of regular secondary structure (α helices plus β sheets) and random coil. However, this large percentage of the coil structure is principally due to the N-terminal domain, which is predicted to have at most 19% regular secondary structure. This domain is also predicted to be an intrinsically unstructured region as defined by IUPred (see the Supporting Information). The GTPase domain on its own shows roughly equal proportions of α helix, β sheet, and coil, while the secondary structure of the C-terminal domain is dominated by α helix, as confirmed by the Paircoil prediction (see the Supporting Information).

Subcloning, Expression, and Purification of Recombinant SEPT4, SEPT4-N, SEPT4-G, SEPT4-C, and SEPT4-GC. DNA amplification produced a ∼1500 bp fragment, corresponding to the encoding region of SEPT4. The expression vector containing this fragment was used to transform *E. coli* AD202 cells, which expressed GST-SEPT4 after IPTG induction. An additional band on SDS−PAGE corresponding to a protein of ∼82 kDa was observed in the crude bacterial lysates, consistent with that expected for the recombinant protein (data not shown). The greater part of the GST-SEPT4 was found in the pellet after lysis under all conditions tested. That which remained in the soluble fraction was isolated by affinity chromatography, pooled for cleavage by Factor Xa, and further purified by Superdex-200 SEC (peak 1 in Figure 2A). SEPT4 eluted in the peak corresponding to the void volume of the column. However, upon SDS−PAGE under reducing conditions, this peak migrated as a largely homogeneous band with a molecular weight between 45 and 66 kDa, consistent with that expected for an SEPT4 monomer (Figure 3A). SEPT4 was demonstrated by N-terminal sequencing of the first five amino acids (data not shown). The position of the major peak on the size-exclusion column is therefore indicative of the formation of SEPT4 oligomers or aggregates. After purification, the final yield was approximately 0.5 mg of SEPT4/L of bacterial culture. Purified SEPT4 only contained 0.25 ± 0.10 nucleotide bound under the purification conditions described here.

After induction with IPTG, *E. coli* transformed with the relevant vectors yielded additional bands of approximately 14.5, 33, 9, and 41 kDa corresponding to SEPT4-N, SEPT4-G, SEPT4-C, and SEPT4-GC, respectively, which were most prominent within the crude bacterial lysates (data not shown). These major products of expression were all recovered in the supernatant. After Ni−NTA affinity chromatography, the SEPT4-N eluted in the major peak from the Superdex-75 size-exclusion column (peak 1 in Figure 2B) prior to imidazole (peak 2). The apparent molecular weight of SEPT4-N determined from linear regression of the protein standards was estimated to be approximately 19 kDa. The SEPT4-C after affinity chromatography was also purified by Superdex-75 chromatography (peak 1 in Figure 2E). The apparent molecular weight of SEPT4-C determined from linear regression was estimated to be approximately 16 kDa.

In the case of SEPT4-G, the results of the Superdex-200 SEC, after initial purification on a Ni−NTA affinity column, are shown in Figure 2C. SEPT4-G appears as a single symmetrical dominant peak in the elution profile (peak 1) followed by GTP and excess imidazole (peaks 2 and 3). The apparent molecular weight in this case was estimated to be 76 kDa. The recombinant SEPT4-N and SEPT4-C proved to be stable, while the SEPT4-G domain was more susceptible to aggregation at temperatures above 30 °C. Final yields for the three expression systems were typically 8, 10, and 5 mg/L culture medium for SEPT4-N, SEPT4-C, and SEPT4-G, respectively. Purified SEPT4-G contained 0.9 ± 0.1 nucleotide bound under the purification conditions described here. The left inset of Figure 2C shows the characteristic

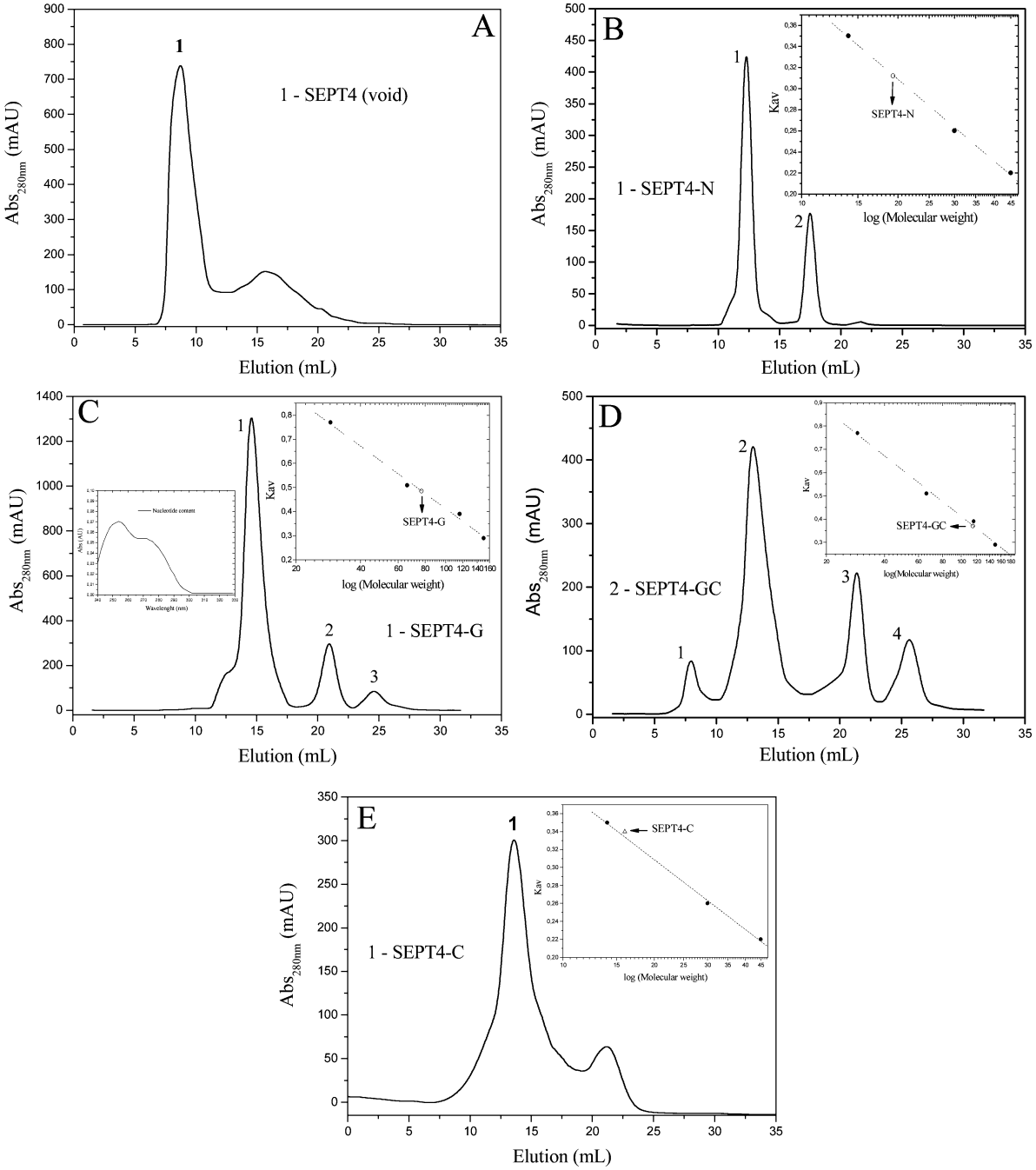

FIGURE 2: SEC of recombinant domains. (A) SEC of purified SEPT4 on Superdex-200. (B) SEC of purified SEPT4-N on Superdex-75. The elution volumes of standard proteins were used to calculate the K_{av} values (K_{av} = [elution volume − column void volume]/[total column volume − column void volume]). The standard proteins of known molecular weight were ovoalbumin (45 kDa), carbonic anhydrase (29 kDa), and lactoalbumin (14.2 kDa). (Right inset) K_{av} versus log(molecular weight). (C) SEC of purified SEPT4-G on Superdex-200. The standard proteins of known molecular weight were alcohol dehydrogenase (150 kDa), β-galactosidase (116 kDa), bovine serum albumin (66 kDa), and carbonic anhydrase (29 kDa). (Right inset) K_{av} versus log(molecular weight). (D) SEC of purified SEPT4-GC on Superdex-200. (Right inset) K_{av} versus log(molecular weight). (E) SEC of purified SEPT4-C on Superdex-75. (Right inset) K_{av} versus log(molecular weight).

spectrum for the free nucleotide recovered from the purified SEPT4-G domain.

SEPT4-GC eluted in the major peak (peak 2 in Figure 2D) in between peak 1 (which corresponded to a small fraction of nucleotide-free aggregated SEPT4-GC) and peaks 3 and 4 corresponding to GTP and imidazole, respectively. The apparent molecular weight of SEPT4-GC determined from linear regression of the protein standards was estimated to be 115 kDa, and typical yields were of the order of 3 mg/L culture. Purified SEPT4-GC contained 0.8 ± 0.1 nucleotide bound under the purification conditions described here. Similar to SETP4-G, SEPT4-GC was also susceptible to aggregation at temperatures above 30 °C. Parts B and C of Figure 3 show the 15% SDS−PAGE under reducing conditions for the final purified SEPT4-N, SEPT4-G, SEPT4-GC, and SEPT4-C.

Native Gel Electrophoresis. To examine the oligomeric state of SEPT4-G and SEPT4-GC in solution, both were subject to electrophoresis under nondenaturing conditions. In the case of SEPT4-G, a major band that migrated similarly

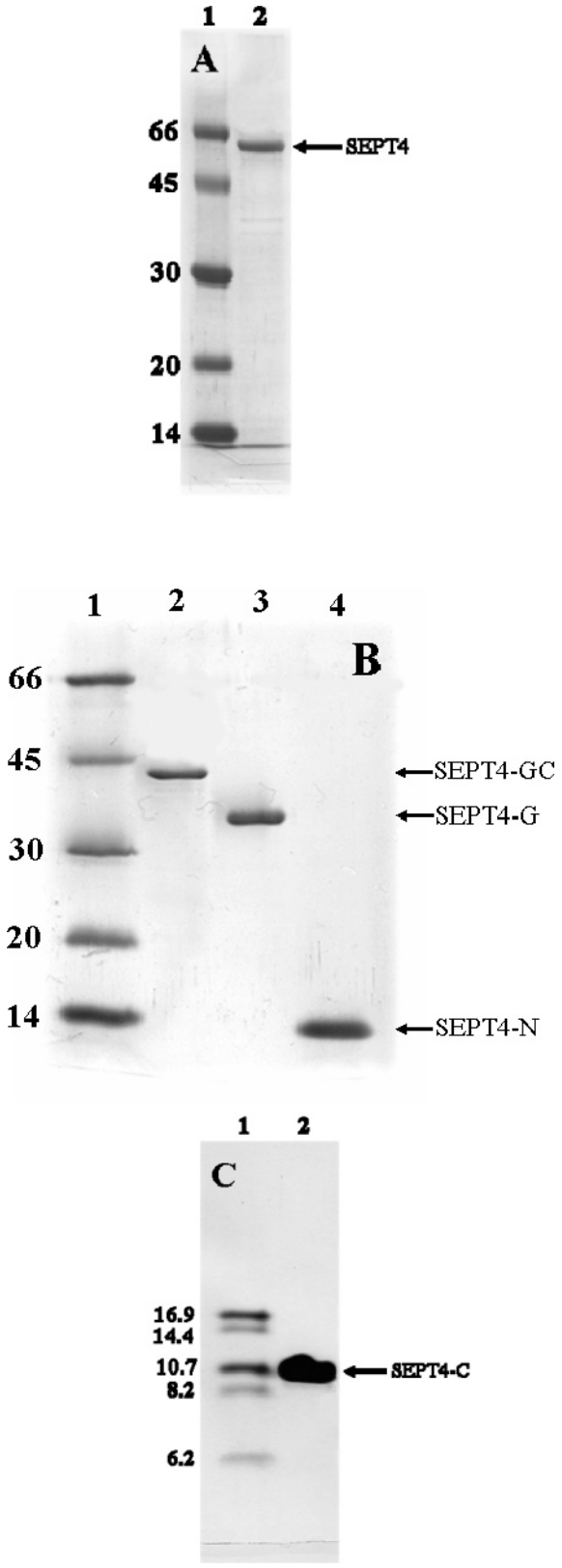

FIGURE 3: Gel electrophoresis under denaturing conditions (SDS−PAGE). (A) Coomassie-stained 15% SDS−PAGE showing SEPT4 after purification. Lane 1, molecular-weight standards; and lane 2, SEPT4. (B) Coomassie-stained 15% SDS−PAGE showing SEPT4-GC, SEPT4-G, and SEPT4-N after purification by affinity chromatography and SEC. Lane 1, molecular-weight standards; lane 2, SEPT4-GC; lane 3, SEPT4-G; and lane 4, SEPT4-N. (C) Coomassie-stained 15% SDS−PAGE showing SEPT4-C after purification. Lane 1, molecular-weight standards; and lane 2, SEPT4-C.

to bovine serum albumin (67 kDa) was clearly detectable (Figure 4A). The hydrodynamic radius of SEPT4-G calculated from the R_f value was 3.8 ± 0.2 nm, corresponding to a M_r of 74 ± 11 kDa for a spherical particle. This is most consistent with a dimer for SEPT4-G in solution under native conditions, whose expected molecular weight is 66 kDa. In the case of SEPT4-GC, a major band was observed, which migrated in a position below lactate dehydrogenase (R_h = 4.4 nm, 140 kDa). The hydrodynamic radius of SEPT4-GC calculated from the R_f value was 4.2 ± 0.3 nm, corresponding to an estimated M_r of 94 ± 18 kDa for a spherical particle. This is also most consistent with a dimer for SEPT4-GC in solution under native conditions, whose molecular weight is 82 kDa.

CD Spectroscopy. CD spectroscopy was used to probe the secondary structure of the expressed SEPT4-N, SEPT4-G, SEPT4-C, and SEPT4-GC products to assess their folding in terms of secondary structure and to draw comparisons (Figure 5A). The CD spectrum of SEPT4-N is characterized by a minimum at 201 ± 1 nm, typically found in structures with a low content of regular secondary structure. This is markedly different to that observed for the recombinant SEPT4-G, which shows two minima at 210 ± 1 and 221 ± 1 nm, a maximum near 196 ± 1 nm, and a negative to positive crossover at 201 ± 1 nm. The negative ellipticities exhibited around 210 and 221 nm are indicative of the presence of α helices. The CD spectrum of SEPT4-GC is characterized by a positive band at 195 ± 1 nm and two negative bands at 210 ± 1 and 222 ± 1 nm, together with a negative to positive crossover at 202 ± 1 nm. These bands are also characteristic of α-helical structures. In the case of SEPT4-C, the CD spectrum is characterized by two minima at 221 ± 1 and 206 ± 1 nm, once again indicative of α helix.

The percentages of α helix, β strand, and coils (turns and irregular structures) were obtained by deconvolution of the experimental spectra using three independent methods and expressed as a mean and standard deviation. In the case of SEPT4-G, the deconvolution yielded 30 ± 4% α helix, 24 ± 2% β strand, and 46 ± 5% turns/irregular structures, indicating that SEPT4-G is composed of a mixture of α and β secondary structures. The corresponding results obtained for SEPT4-GC were 35 ± 6% α helix, 20 ± 5% β strand, and 45 ± 4% turns/irregular structures, indicative of a similar structure. This is to be expected given that the greater part of the SEPT4-GC structure is the GTPase domain itself. In the case of SEPT4-C, the deconvolution yielded 58 ± 5% α helix and 42 ± 4% irregular structures (see Table 3).

Fluorescence Spectroscopy. Fluorescence spectroscopy was used to probe the tertiary structure of SEPT4-N, SEPT4-G, and SEPT4-GC. Figure 5B shows the emission spectrum of SEPT4-N after excitation at 295 nm. It is characterized by a maximum at 356 ± 1 nm, which is typical of solvent-exposed tryptophan residues. Given that free tryptophan typically presents a similar emission maximum, this strongly suggests that all three tryptophan residues of SEPT4-N are exposed to the solvent. Figure 5B also shows the emission spectrum of SEPT4-G and SEPT4-GC, which are characterized by a maximum at 334 ± 1 and 332 ± 1 nm, respectively. This emission maximum is typical of tryptophan residues buried inside the protein core, indicating that both tryptophans present in the SEPT4-G and SEPT4-GC se-

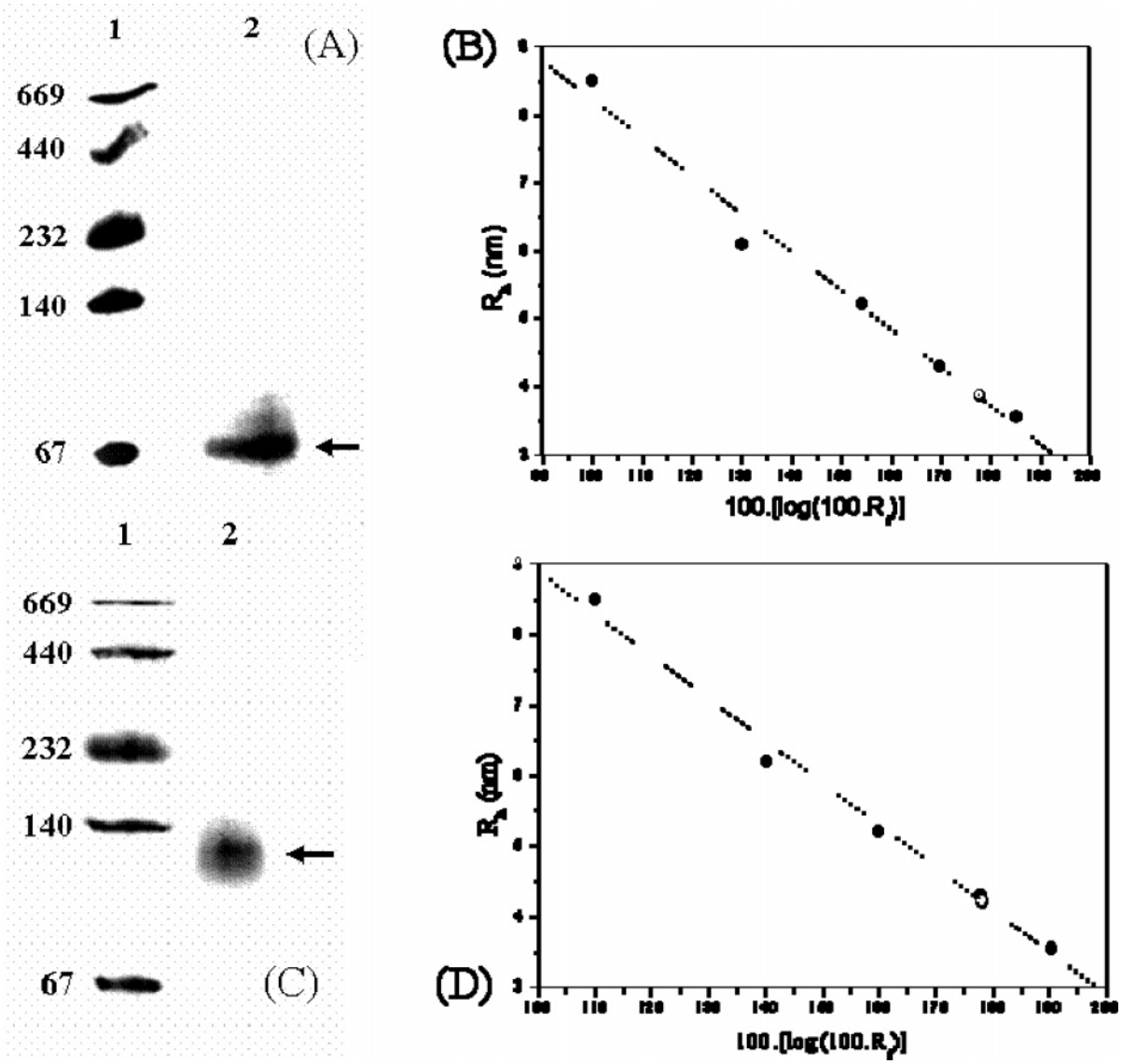

FIGURE 4: Native gel electrophoresis. (A) Native gel showing recombinant SEPT4-G (lane 2). Molecular-weight standards are shown in lane 1. The standard proteins of known molecular weight were thyroglobulin (669 kDa), ferritin (440 kDa), catalase (232 kDa), lactate dehydrogenase (140 kDa), and albumin (67 kDa). The corresponding linear regression is shown in B. (C) Native gel showing recombinant SEPT4-GC and (D) corresponding linear regression.

quence have become hidden from the solvent upon protein folding.

DLS. When the purified SEPT4 was analyzed by DLS, the observed profile was characteristic of a monodisperse solution, whose major peak, corresponding to a hydrodynamic radius of approximately 7.8 ± 0.5 nm, almost certainly corresponds to an oligomeric form of the protein or to aggregates (see Table 4). This is consistent with the SEC profile described above. A second, minor peak (corresponding to particles with a hydrodynamic radius of approximately 30 nm) probably represents high-molecular-weight aggregates. When purified SEPT4-N was analyzed by DLS, the observed profile was characteristic of a largely monodisperse solution, whose principal peak, corresponding to >97% of the mass, had an estimated hydrodynamic radius of 2.3 ± 0.2 nm (Table 4). Similar monodisperse profiles were also observed for SEPT4-G and SEPT4-GC. In these cases, the major peak representing >95 and >92% of the total mass corresponded to hydrodynamic radii of 4.0 ± 0.3 and 4.8 ± 0.3 nm, respectively. In the case of SEPT4-C, the observed profile was also characteristic of a largely monodisperse solution, whose peak corresponded to a particle with an estimated hydrodynamic radius of 2.0 ± 0.1 nm (Table 4).

SAXS Measurements. To obtain information about the tertiary structure of SEPT4-G and its molecular shape, we submitted the domain to SAXS analysis. The X-ray scattering curve of SEPT4-G obtained is shown in Figure 6A. The radius of gyration (R_g) evaluated with the Guinier approximation was $R_g = 31.6 \pm 0.2$ Å. The Guinier plot was linear in the qR_g range up to 1.71. The distance distribution function, $p(r)$, evaluated by the indirect Fourier transform with GNOM (*37*, *38*), estimated the radius of gyration and the maximum dimension of the molecule to be 33.6 ± 0.2 Å and 110 Å, respectively (inset of Figure 6A). The shape of the domain was determined from X-ray scattering data with DAMMIN (*39*, *40*), using a real-space algorithm. The results of 10 independent simulated annealing computations were compared with each other, and the best model is given in Figure 6B. The molecular dimensions of the domain by

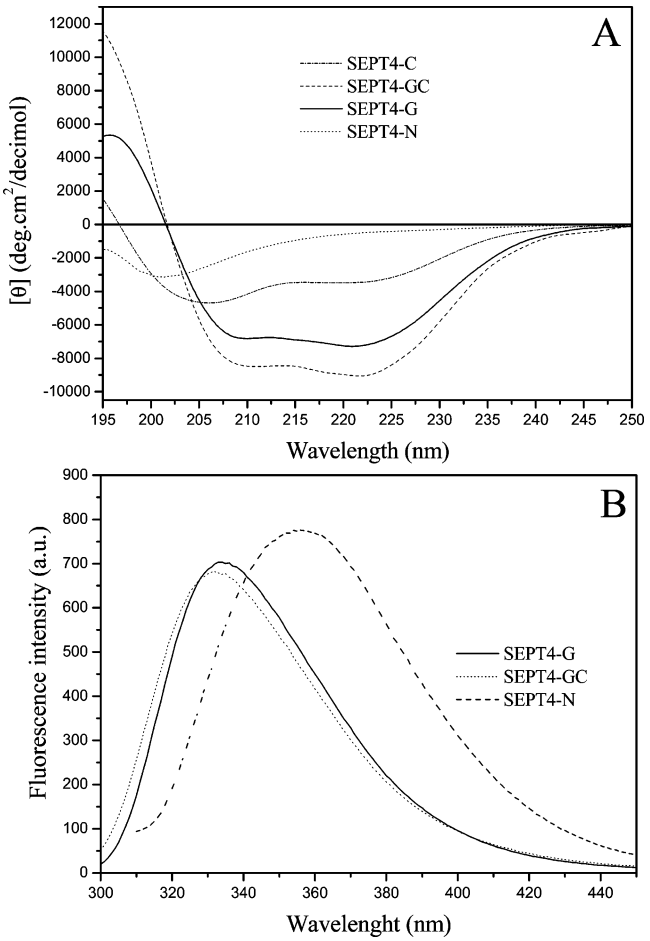

FIGURE 5: CD spectrum and fluorescence spectroscopy of recombinant domains. (A) Far-UV CD spectra of SEPT4-N (···), SEPT4-G (—), SEPT4-GC (- - -), and SEPT4-C (·-·). All measurements were made in 25 mM Tris-HCl at pH 7.8 and 20 mM NaCl at 10 °C. In the case of SEPT4-G and SEPT4-GC, 5% glycerol and 0.1 mM EDTA were included in the buffer. (B) Data were collected at 10 °C using an excitation wavelength of 295 nm and emission spectra collected from 305 and 450 nm, for SEPT4-N (···), SEPT4-G (—), and SEPT4-GC (- - -). All data were collected in 25 mM Tris-HCl at pH 7.8 and 20 mM NaCl. In the case of SEPT4-G and SEPT4-GC, 0.1 mM EDTA and 5% glycerol were added to the buffer.

Table 3: CD Deconvolution

domains	α helix (%)	β sheet (%)	coil (%)
SEPT4-G	30 ± 4	24 ± 2	46 ± 5
SEPT4-GC	35 ± 6	20 ± 5	45 ± 4
SEPT4-C	58 ± 5	0	42 ± 4

this approach were approximately 110 Å, in the longest dimension, by 58 Å, perpendicular to it.

GTPase Activity Assay. The result of the incubation of SEPT4-G in the presence of [γ-^{32}P]GTP was analyzed by TLC. Complete hydrolysis of the [γ-^{32}P]GTP occurred within 10 min only in the presence of magnesium. In the absence of SEPT4-G or magnesium ions, [γ-^{32}P]GTP hydrolysis was undetectable (see the Supporting Information).

DISCUSSION

Domain Definitions for Human SEPT4. It has been recognized for some time that effectively all known septins are composed of three major structural domains, although

Table 4: DLS[a]

domains	R_h (nm)	percent mass (%)	M_r (kDa)
SEPT4	7.8 ± 0.5	<85	422 ± 106
SEPT4-N	2.3 ± 0.2	>97	22 ± 5
SEPT-G	4.0 ± 0.3	>95	83 ± 17
SEPT-GC	4.8 ± 0.3	>92	130 ± 19
SEPT4-C	2.0 ± 0.1	100	17 ± 2

[a] DLS of recombinant SEPT4-N (~1 mg/mL) and SEPT4-C (~1 mg/mL) in 25 mM Tris-HCl and 20 mM NaCl buffer at pH 7.8; SEPT4-G (~1 mg/mL) in 25 mM Tris-HCl, 20 mM NaCl, 5% glycerol, and 0.1 mM EDTA buffer at pH 7.8; and SEPT4-GC (~0.5 mg/mL) and SEPT4 (~0.8 mg/mL) in the same buffer used for SEPT4-G.

the size and sequence of both the N- and C-terminal domains varies considerably from septin to septin (*6*). However, at present, there is no known three-dimensional structure for any septin from any species or isolated domain thereof, although crystals of a mutant human SEPT1 in the form of a fusion protein with GST have been recently reported (*41*). Furthermore, there is a surprising dearth of physicochemical characterization of septins in general, particularly in terms of spectroscopic analyses of highly purified material. One of the reasons for this may be a certain difficulty in precisely defining the boundaries to the three domains, which is necessary for building constructs suitable for their heterologous expression. This is an approach that should, in principle, simplify structural characterization given that most full-length septins are susceptible to aggregation (*11*). Here, we have successfully expressed, purified, and structurally characterized recombinant proteins corresponding to all three of the domains of human SEPT4. Furthermore, two additional constructs, which correspond to the entire protein and a fusion of the GTPase domain with the C-terminal domain, have also been successfully produced. This provides a basis for further structural and functional studies, allowing for a better understanding of the physiological role of the individual domains as well as their participation in pathological processes such as cancer and neurodegenerative disorders (*17*, *20*, *21*).

The final purified SEPT4 migrated with the expected molecular mass (55 kDa) under denaturing conditions on SDS−PAGE, but both DLS experiments and SEC indicated that the product was either a large oligomer or in the form of aggregates. Upon gel filtration, the protein eluted as a single peak with a molecular weight significantly greater than 150 kDa, the largest molecular-weight standard applied to the column. On the other hand, the major peak seen in the DLS profile was centered at a hydrodynamic radius of approximately 8 nm, which, on the basis of the assumption of a spherical particle, corresponds to a molecular mass of around 450 kDa, consistent with a large particle. However, because there is little information available at present concerning the molecular shape of septin oligomers, this estimate should be treated with some caution. Furthermore, SEPT4 did not enter the gel under native electrophoresis, also testifying to its high molecular weight.

Besides this difficulty in accurately determining the size and nature of the recombinant SEPT4 molecule, this product also presented a series of other characteristics, which were unfavorable for its further study. These included a very low yield for the expression system described here (typically only 0.5 mg/L), a low nucleotide content (0.25 ± 0.1), and

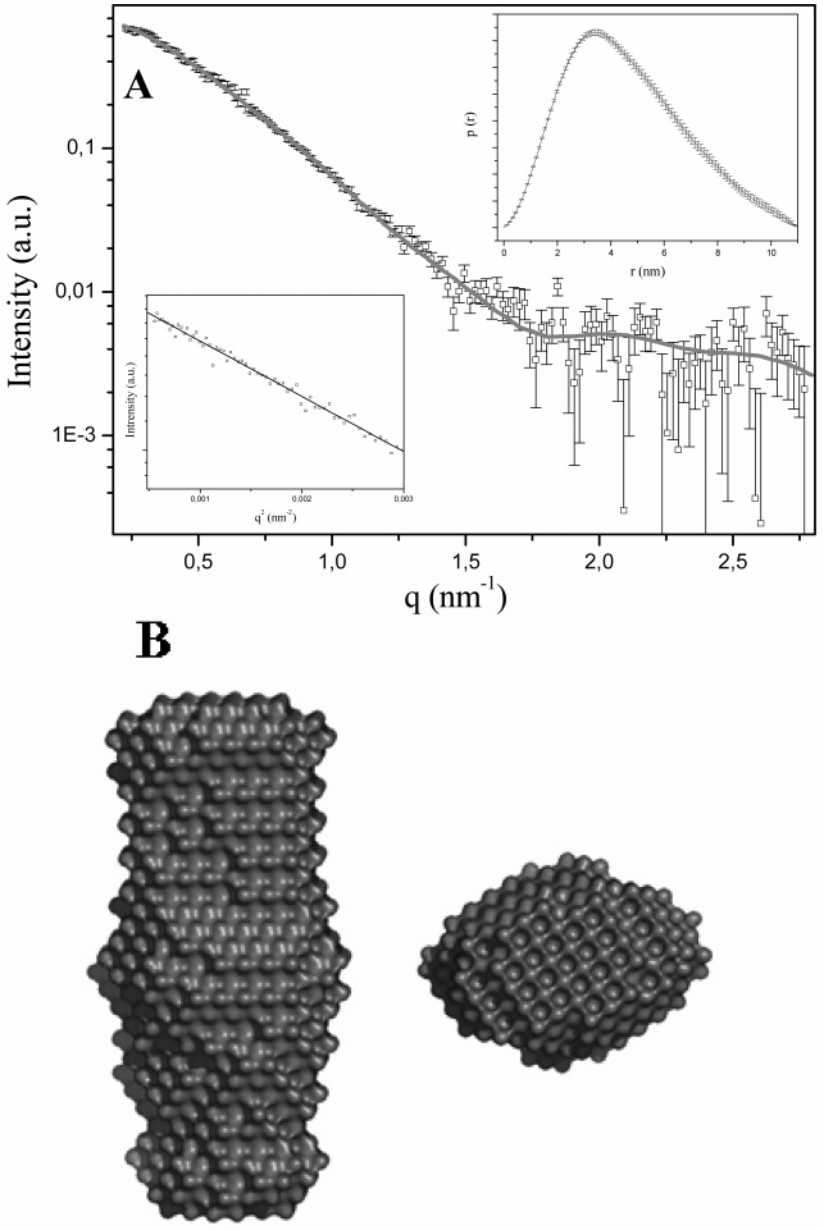

FIGURE 6: SAXS measurements. (A) SAXS curve of the SEPT4-G in 25 mM Tris-HCl, 20 mM NaCl, and 0.1 mM EDTA buffer at pH 7.8 and 5% glycerol superimposed with the theoretical scattering curve from the restored low-resolution model. The radius of gyration obtained from the Guinier plot is $R_g = 31.6 \pm 0.2$ Å, which is just slightly smaller than $R_g = 33.6 \pm 0.2$ Å obtained from the integral analyses of the scattering curve using the method implemented in GNOM. (Inset) Distance distribution, $p(r)$, of SEPT4-G computed by indirect Fourier transform with GNOM (*37*, *38*). (B) Shape of SEPT4-G in solution was restored from synchrotron X-ray scattering data using an *ab initio* simulated annealing algorithm implemented in DAMMIN (*39*). Only the best model is shown.

instability that leaded to a tendency to aggregate and then finally precipitate over time. Similar results have been described for other single septins *in vitro* (*11*). For these reasons, emphasis was placed on attempts to dissect the molecule into its principal component domains.

As yet, no specific function has been attributed to the N-terminal domain of SEPT4, although it is one of the largest such domains among human septins. In the present study, the definition of SEPT4-N was based on the common sequence observed for this region in two different splice variants (isoforms 1 and 4) and corresponded to the first 119 amino acid residues of the whole protein. According to the definitions used here, the N-terminal domain is followed by a linker sequence of 24 residues (120−143) prior to the N terminus of the central GTPase domain (SEPT4-G). The latter was defined with respect to small GTPases, which consistently present a hydrophobic stretch of four residues, corresponding to part of the N-terminal β strand, immediately adjacent to the P loop. The known structures for small GTPases suggest that the presence of the β strand is essential for structural integrity, and the strand typically starts a few residues prior to the hydrophobic stretch. For this reason, the N terminus of the GTPase domain was defined as being residue Asp144, three residues prior to the beginning of the hydrophobic stretch (LMVA), which starts at Leu147. The C terminus of the GTPase domain was less easy to define unambiguously but was guided by a CAAX sequence common to many small GTPases at their C termini. By this criterion, SEPT4-G terminates at residue Thr416. The final definition of the domain structure is shown in Figure 1A.

Both SEPT4-N and SEPT4-G were expressed as soluble products. Previous attempts to express the GTPase domain

using a construct that generated a product starting at residue Tyr124 and terminating at Arg408 were unsuccessful in that the product was always insoluble under the conditions used (data not shown). This former construct includes the greater part of the linker sequence between the N-terminal domain and the GTPase domain used in the present study. The linker includes the functionally important polybasic motif (HRKS-VKK, residues 135−141) associated with phospholipid binding (*10*). It seems to be important to eliminate this region to generate a soluble SEPT4-G product, possibly because of a tendency for membrane association via the polybasic region in the absence of the N-terminal domain. Therefore, in structural terms, it is probably convenient to consider the region containing the polybasic sequence as independent from the GTPase domain, although functionally, these two regions have been shown to be coupled (*10*).

SEPT4-N Contains Little Regular Secondary Structure and May Be Intrinsically Unstructured. The final purified SEPT4-N migrated with the expected molecular mass under denaturing conditions on SDS−PAGE (∼14.5 kDa, including the His tag), although this provides no information concerning its native oligomeric state. Two approaches were taken to determine the molecular mass of the soluble product under native conditions after expression and purification. Upon gel filtration, SEPT4-N eluted as a single peak of apparent molecular weight ∼ 19 kDa (Figure 2B). This is suggestive of a monomer, despite being slightly above the expected value. In contrast, the peak seen in the DLS profile is centered at a hydrodynamic radius of 2.3 ± 0.2 nm (see Table 4), which, on the basis of the assumption of a spherical particle, corresponds to a molecular mass of 22 ± 5 kDa. This is roughly midway between the theoretical mass of a dimer (29 kDa) and that of a monomer. These ambiguous results may be explicable if the shape of the molecular envelope was significantly nonglobular. To investigate this further, SEPT4-N was subject to sequence analysis, CD, and fluorescence spectroscopy.

The CD spectrum of SEPT4-N (Figure 5A) indicates a very low content of regular secondary structure, and this is consistent with the theoretical predictions given in Table 2. Predicted values varied between 81 and 92% of random coil, and certainly, one of the reasons for this is the unusually high proline content of the N-terminal domain (17 prolines in 119 residues, corresponding to 14.3% of the sequence).

The results of fluorescence spectroscopy measurements (Figure 5B) indicated that all three tryptophans of SEPT4-N are highly solvent-exposed, as characterized by the average emission maximum at 356 ± 1 nm, which is close to the maximum possible red shift and corresponds to that expected for the free amino acid (*42*). Although, on their own, the fluorescence data cannot be used to infer much about the three-dimensional structure of the domain, when these are taken together with the results of the CD spectroscopy and structure prediction, they strongly suggest that the isolated N-terminal domain has no well-defined compact 3D structure and is therefore probably highly nonglobular. Consistent with this conclusion, the H 1D nuclear magnetic resonance spectrum measured at 600 MHz (data not shown) also indicated a low content of regular secondary structure, as inferred from the low dispersion of the H_N and H_α main-chain chemical shifts.

These data may be used to rationalize the mass measurements made by gel filtration and DLS. The most likely explanation appears to be that SEPT4-N is a largely unstructured domain, whose shape is elongated. This would lead to anomolous migration upon gel filtration, leading to an overestimation of the molecular mass (19 kDa compared with 14.5 kDa), and would invalidate the use of a spherical model for the particle in the calculation of the mass from the hydrodynamic radius. Therefore, most likely, SEPT4-N is a monomer in solution and therefore does not contribute to homo-oligomerization of SEPT4 (see below).

The function of SEPT4-N is still unclear; however, the presence of many proline-rich regions are suggestive of a role in the protein−protein interaction, as seen for example in SH3-binding domains (*43*). Recently, it has become recognized that intrinsically unstructured proteins or regions therein (IUPs) are more common than originally believed (*28*, *44*). These regions have a lower density of internal contacts than compact globular structures, and their inherent disorder may confer the ability to interact differentially with alternative molecular partners (*44*). Predictions made with the IUPred server (*28*), designed to locate such regions in protein sequences, identify residues 1−128 as an IUP. When these observations are taken together, they suggest that the extended N-terminal domain, which shares no homology with any other protein in the current database, may therefore be important for conferring unique biochemical signaling properties on human SEPT4.

Oligomerization of SEPT4-G, SEPT4-GC, and SEPT4-C. Upon gel filtration, SEPT4-G eluted as a single peak with an apparent molecular weight of ∼76 kDa (see Figure 2C). This is broadly consistent with a dimer, whose theoretical expected mass is 66 kDa (including the His tag). This is supported by the results of native gel electrophoresis, which showed a principal band that comigrated with bovine serum albumin and was estimated to correspond to a molecular mass of 74 ± 11 kDa (see Figure 4). However, when the peak from the gel-filtration column was analyzed by DLS, a hydrodynamic radius of approximately 4.0 ± 0.3 nm was observed (see Table 4), which, on the basis of the assumption of a spherical particle, would correspond to a molecular mass of 83 ± 17 kDa, which would be more consistent with a trimer than a dimer. However, in this case, an overestimate for the molecular mass is expected in cases where the molecular envelope deviates significantly from a sphere, which would be likely if the SEPT4-G were a dimer. SAXS measurements were therefore used to probe the molecular shape of SEPT4-G directly in solution.

The radii of gyration calculated from both the Guinier regions of the SAXS scattering curve (31.6 Å) and that obtained with GNOM (33.6 Å) are consistent with a dimeric particle. Furthermore, the reconstructed envelope of SEPT4-G (Figure 6B) restored at 20 Å resolution from synchrotron X-ray scattering data shows an elongated particle of approximately 110 by 58 Å. This envelope was calculated assuming a dimeric particle with 2-fold symmetry, which was able to reproduce well both the experimental curve and $p(r)$. The SAXS data are therefore consistent with a dimeric structure in the form of a prolate ellipsoid. This extended molecular form shown in Figure 6B therefore explains why both DLS and SEC tend to overestimate slightly the molecular mass. Although the resolution of the molecular

envelope is extremely limited, it represents the first visualization of a septin domain structure reported to date.

Overall, therefore, the weight of the present evidence seems to favor a homodimeric particle for the free GTPase domain in solution. Furthermore, no evidence was observed from either SEC or right-angle light scattering for the formation of homopolymers of the individual SEPT4-G domain, despite the fact that in its purified state it is saturated with GTP. This is consistent with the observation that the vast majority of single septins do not polymerize *in vitro* (*45*).

In the case of SEPT4-GC, gel filtration, native gel electrophoresis, and DLS analyses yielded estimates of ~115, 94 ± 18, and 130 ± 19 kDa (corresponding to a hydrodynamic radius of 4.8 nm), respectively, for the molecular mass. It is difficult to draw any definitive conclusion concerning the oligomeric state in this case because the estimated values are greatly dispersed around the expected values for a dimer (~82 kDa) and a trimer (~123 kDa) and may be greatly overestimated because of the likely nonspherical nature of the molecule. In this case, difficulty in obtaining highly concentrated samples of SEPT4-GC prohibited the use of SAXS for a fuller description of the oligomers.

In the case of SEPT4-C, gel filtration and DLS analyses yielded estimates of 16 and 17 ± 2 kDa for the molecular mass. These results are approximately twice the calculated molecular weight of monomeric SEPT4-C (~9 kDa). This indicates that SEPT4-C is homodimeric in solution, consistent with the presence of the coiled coil. Similar results of SEC were obtained for the C-terminal fragment (coiled coil, bZIP) of the Epstein−Barr virus for example (*46*). The fact that both the SEPT4-G and SEPT4-C domains appear to be dimers aids in clarifying the nature of the SEPT4-GC product described above, because it strongly suggests that the most likely oligomeric state for the latter is also dimeric. This follows from the following reasoning: if the pairs of molecules that dimerize via the GTPase domain were different from those that dimerize via the C-terminal domain, this would inevitably lead to an oligomer that would have to be at least a tetramer and this is inconsistent with the results described above for SEPT4-GC.

Secondary and Tertiary Structure of SEPT4-G and SEPT4-GC. The fluorescence emission spectra of SEPT4-G and SEPT4-GC are effectively identical with an emission maximum between 332 and 334 nm, indicative of buried tryptophans. This in itself is suggestive of a product with a defined tertiary structure. Visual inspection of the CD spectrum of SEPT4-G indicates the presence of α-helical secondary structure as evidenced by the negative ellipticities around 210 ± 1 and 221 ± 1 nm (Figure 5A). The spectra are strongly characteristic of an α/β protein, with spectral bands attributed to electron transitions in the amide groups of the protein backbone. Deconvolution of the spectrum using three different algorithms yielded the following average values for the secondary structure: 30 ± 4% α helix, 24 ± 2% β sheet, and 46 ± 5% coil (turn plus random). By comparison with small GTPases, the SEPT4-G domain would be expected to fold as a mixed α/β structure based on an open β sheet. Such domains typically contain on the order of ~40% α helix and ~30% β sheet and would thus be expected to contribute both signals to the CD spectrum. These values are not dissimilar to those given above from spectral deconvolution, suggesting that the central GTPase domains of septins are structurally similar to small GTPases in terms of their overall fold and are probably based on an open β-sheet topology.

As can be seen from Figure 5A, the CD spectrum for the SEPT4-GC shows a greater positive ellipticity at 195 ± 1 nm when compared with any of the isolated domains. This is accompanied by an increase in the characteristic negative ellipticities around 210 ± 1 and 222 ± 1 nm. These differences are suggestive of a slightly greater α-helical content for SEPT4-GC compared with the GTPase domain, and this difference can only be attributable to the C-terminal domain. This is reflected by a small increase in α-helical content from 30 to 35% on deconvolution. This was readily confirmed by analysis of the CD spectrum of the isolated C-terminal domain (Figure 5A), which, upon deconvolution, resulted in 58% α helix and 42% coil, entirely consistent with the results of secondary-structure prediction (Table 2).

SEPT4-C (residues 417−478) includes a second proline-rich sequence followed by a short stretch of uninterrupted leucine/isoleucine repeats at every seventh amino acid residue. This is expected to form part of a left-handed coiled coil, as suggested for most other septins and consistent with the results of the Paircoil prediction (see the Supporting Information), which indicates a tendency toward coiled-coil formation as from position Pro447. This explains why the CD spectrum of this domain is not purely α-helical, as can be seen both directly from the results of the deconvolution (Table 3) itself as well as from the shift in the lower wavelength minimum from 208 to 206 nm, indicative of the presence of coil. Presumably, this arises from the N-terminal praline-rich region of the domain.

Interestingly, however, the residues originally identified as characteristic of a leucine-zipper-like coiled coil (Ile453, Leu460, and Leu467) (*16*) do not correspond to the canonical d positions but rather to the a positions of the heptad repeat (*47*). This may be indicative of a capacity to participate in coiled coils containing more than two helices (*48*). This opens up the possibility for the C-terminal domain playing a role in intermolecular interactions either between different septins or with other target molecules. This is consistent with the Paircoil prediction, which yielded only a relatively weak coiled coil probability for this region (50%), and also with the IUPred profile, which indicates that the C-terminal domain may be somewhat unstructured.

In summary, there is a clear sequence signal within the C-terminal domain corresponding to a coiled coil and also experimental evidence from the CD spectrum that this region is at least partly α-helical. However, it is still unclear what role this structure may play in terms of homo- and hetero-oligomeric associations among septins, especially given the fact that the SEPT4-G domain alone is capable of forming dimers in the absence of the C-terminal domain.

GTPase Activity Assay. Several lines of evidence indicate that the final product, SEPT4-G, is correctly folded. These include CD spectroscopy, fluorescence spectroscopy, and GTP binding. Purified SEPT4-G contained 0.9 ± 0.1 nucleotide bound under the purification conditions described here. To test if the GTPase domain as defined here was also catalytically active, samples were incubated with [γ-^{32}P]GTP in the presence and absence of 0.5 mM Mg^{2+}. The production of radioactive free thiophosphate was followed by TLC and

visualized by autoradiography. GTP was completely hydrolyzed in 10 min (see the Supporting Information), demonstrating that SEPT4-G readily binds and hydrolyzes GTP, suggesting that the definition of the GTPase domain described here corresponds to a biologically functional unit. In the absence of SEPT4-G and/or Mg^{2+} (see the Supporting Information), we could not detect GTP hydrolysis, demonstrating catalysis to be Mg^{2+}-dependent as expected (*10*, *12*).

Conclusion. In the current work, a first attempt to structurally characterize the domains of SEPT4 has been undertaken. In summary, SEPT4 can be described as being composed of four different regions: (i) an N-terminal domain (SEPT4-N), which possesses a low content of regular secondary structure, is probably structurally disordered and may be classified as an IUP (intrinsically unstructured region) possible involved in heterologous interactions; (ii) a short basic region that binds PtdIns(4,5)P_2, believed to interact with membranes; (iii) a central GTPase domain (SEPT4-G) composed of a mixed α/β structure, which is catalytically active and forms homodimers *in vitro*; and (iv) a C-terminal domain divided into two parts, one of which is proline-rich and a second that is α-helical and expected to form a coiled coil. These studies should provide a useful basis for future biophysical studies of SEPT4, including the structural basis for their involvement in diseases such as cancer and neurodegenerative disorders.

ACKNOWLEDGMENT

We thank Dr. Claudia Munte for NMR analyses and Suzana Sculaccio and Derminda I. de Moraes for technical assistance. This work was supported by FAPESP via a grant to the Centro de Biotecnologia Molecular Estrutural. The present work was also partially supported by a grant from Tokai University and AIST.

SUPPORTING INFORMATION AVAILABLE

IUPred (A) and coiled-coil (B) predictions for human SEPT4 and DLS of recombinant domains. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Hartwell, L. H. (1971) Genetic control of the cell division cycle in yeast. IV. Genes controlling bud emergence and cytokinesis, *Exp. Cell Res. 69*, 265−276.
2. Frazier, J. A., Wong, M. L., Longtine, M. S., Pringle, J. R., Mann, M., Mitchison, T. J., and Field, C. (1998) Polymerization of purified yeast septins: Evidence that organized filament arrays may not be required for septin function, *J. Cell Biol. 143*, 737−749.
3. Field, C. M., al-Awar, O., Rosenblatt, J., Wong, M. L., Alberts, D., and Mitchison, T. J. (1996) A purified *Drosophila* septin complex forms filaments and exhibits GTPase activity, *J. Cell Biol. 133*, 605−616.
4. Hsu, S. C., Hazuka, C. D., Roth, R., Foletti, D. L., Heuser, J., and Scheller, R. H. (1998) Subunit composition, protein interactions, and structures of the mammalian brain sec6/8 complex and septin filaments, *Neuron 20*, 1111−1122.
5. Macara, I. G., Baldarelli, R., Field, C. M., Glotzer, M., Hayashi, Y., Hsu, S. C., Kennedy, M. B., Kinoshita, M., Longtine, M., Low, C., Maltais, L. J., McKenzie, L., Mitchison, T. J., Nishikawa, T., Noda, M., Petty, E. M., Peifer, M., Pringle, J. R., Robinson, P. J., Roth, D., Russell, S. E., Stuhlmann, H., Tanaka, M., Tanaka, T., Trimble, W. S., Ware, J., Zeleznik-Le, N. J., and Zieger, B. (2002) Mammalian septins nomenclature, *Mol. Biol. Cell 13*, 4111−4113.
6. Kartmann, B., and Roth, D. (2001) Novel roles for mammalian septins: From vesicle trafficking to oncogenesis, *J. Cell Sci. 114*, 839−844.
7. Martínez, C., Sanjuan, M. A., Dent, J. A., Karlsson, L., and Ware, J. (2004) Human septin−septin interactions as a prerequisite for targeting septin complexes in the cytosol, *Biochem. J. 382*, 783−791.
8. Field, C. M., and Kellogg, D. (1999) Septins: Structural polymers or signaling GTPases? *Trends Cell Biol. 9*, 387−394.
9. Bourne, H. R., Sanders, D. A., and McCormick, F. (1991) The GTPase superfamily: Conserved structure and molecular mechanism, *Nature 349*, 117−127.
10. Zhang, J., Kong, C., Xie, H., McPherson, P. S., Grinstein, S., and Trimble, W. S. (1999) Phosphatidylinositol polyphosphate binding to the mammalian septin H5 is modulated by GTP, *Curr. Biol. 9*, 1458−1467.
11. Sheffield, P. J., Oliver, C. J., Kremer, B. E., Sheng, S., Shao, Z., and Macara, I. G. (2003) Borg/septin interactions and the assembly of mammalian septin heterodimers, trimers, and filaments, *J. Biol. Chem. 278*, 3483−3488.
12. Hillebrand, S., Garcia, W., Cantú, M. D., Araújo, A. P. U., Tanaka, M., Tanaka, T., Garratt, R. C., and Carrilho, E. (2005) In vitro monitoring of GTPase activity and enzyme kinetics studies using capillary electrophoresis, *Anal. Bioanal. Chem. 383*, 92−97.
13. Casamayor, A., and Snyder, M. (2003) Molecular dissection of a yeast septin: Distinct domains are required for septin interaction, localization, and function, *Mol. Cell. Biol. 23*, 2762−2777.
14. Kissel, H., Georgescu, M., Larisch, S., Manova, K., Hunnicutt, G., and Steller, H. (2005) The *Sept4* septin locus is required for sperm terminal differentiation in mice, *Dev. Cell 8*, 353−364.
15. Xie, H., Surka, M., Howard, J., and Trimble, W. S. (1999) Characterization of the mammalian septin H5: Distinct patterns of cytoskeletal and membrane association from other septin proteins, *Cell Motil. Cytoskeleton 43*, 52−62.
16. Tanaka, M., Tanaka, T., Kijima, H., Itoh, J., Matsuda, T., Hori, S., and Yamamoto, M. (2001) Characterization of tissue- and cell-type-specific expression of a novel human septin family gene, Bradeion, *Biochem. Biophys. Res. Commun. 286*, 547−553.
17. Tanaka, T., Kijima, H., Itoh, J., Matsuda, T., and Tanaka, T. (2002) Impaired expression of a human septin family gene Bradeion inhibits the growth and tumorigenesis of colorectal cancer *in vitro* and *in vivo*, *Cancer Gene Ther. 9*, 483−488.
18. Neufeld, T. P., and Ruben, G. M. (1994) The *Drosophila* peanut gene is required for cytokinesis and encodes a protein similar to yeast putative bud neck filament proteins, *Cell 77*, 371−379.
19. Nguyen, T. Q., Sawa, H., Okano, H., and White, J. G. (2000) The *C. elegans* septin genes, unc-59 and unc-61, are required for normal postembryonic cytokineses and morphogenesis but have no essential function in embryogenesis, *J. Cell Sci. 113*, 3825−3837.
20. Kinoshita, A., Kinoshita, M., Akiyama, H., Tomimoto, H., Akiguchi, I., Kumar, S., Noda, M., and Kimura, J. (1998) Identification of septins in neurofibrillary tangles in Alzheimer's disease, *Am. J. Pathol. 153*, 1551−1560.
21. Ihara, M., Tomimoto, H., Kitayama, H., Morioka, Y., Akigunchi, I., Shibasaki, H., Noda, M., and Kinoshita, M. (2003) Association of the cytoskeletal GTP-binding protein Sept4/H5 with cytoplasmatic inclusions found in Parkinson's disease and others synucleinopathies, *J. Biol. Chem. 278*, 24095−24012.
22. Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., and Higgins, D. G. (1997) The Clustal-X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools, *Nucleic Acids Res. 25*, 4876−4882.
23. Hancock, J. F., Paterson, H., and Marshall, C. J. (1990) A polybasic domain or palmitoylation is required in addition to the CAAX motif to localize p21Ras to the plasma membrane, *Cell 63*, 133−139.
24. Rost, B. (1996) PHD—Predicting one-dimensional protein structure by profile-based neural networks, *Methods Enzymol. 266*, 525−539.
25. Jones, D. T. (1999) PSIPRED—Protein secondary structure prediction based on position-specific scoring matrices, *J. Mol. Biol. 292*, 195−202.
26. Baldi, P., Brunak, P., Frasconi, P., Soda, G., and Pollastri, G. (1999) SSPRO—Exploiting the past and the future in the protein secondary structure prediction, *Bioinformatics 15*, 937−946.
27. Berger, B., Wilson, D. B., Wolf, E., Tonchev, T., Milla, M., and Kim, P. S. (1995) Predicting coiled coils by use SF pairwise residue correlations, *Proc. Natl. Acad. Sci. U.S.A. 92*, 8259−8263.

28. Dosztanyi, Z., Csizmok, V., Tompa, P., and Simon, I. (2005) IUPred: Web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content, *Bioinformatics 21*, 3433−3434.
29. Sanger, F., Nicklen, S., and Coulson, R. (1977) DNA sequencing with chain-terminating inhibitors, *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463−5467.
30. Gill, S. C., and von Hippel, P. H. (1989) Calculation of proteins extinction coefficients from amino acid sequence data, *Anal. Biochem. 182*, 319−326.
31. Seckler, R., Wu, G. M., and Timasheff, S. N. (1990) Interactions of tubulin with guanylyl-(β-γ-methylene)diphosphate, *J. Biol. Chem. 265*, 7655−7661.
32. Sreerama, N., and Woody, R. W. (1993) A self-consistent method for the analysis of protein secondary structure from circular dichroism, *Anal. Biochem. 209*, 32−44.
33. Provencher, S. W., and Glockner, J. (1981) Estimation of globular protein secondary structure from circular dichroism, *Biochemistry 20*, 33−37.
34. Johnson, W. C. (1999) Analyzing protein circular dichroism spectra for accurate secondary structures, *Proteins: Struct., Funct., Genet. 35*, 307−312.
35. Tamura, J., Warrener, P., and Collett, M. S. (1993) RNA-stimulated NTPase activity associated with the p80 protein of the pestivirus bovine viral diarrhea virus, *Virology 193*, 1−10.
36. Kellermann, G., Vicentin, F., Tamura, E., Rocha, M., Tolentino, H., Barbosa, A., Craievich, A., and Torriani, I. (1997) The small-angle X-ray scattering beamline of the Brazilian Synchrotron Light Laboratory, *J. Appl. Crystallogr. 30*, 880−883.
37. Svergun, D. I., Semenyuk, A. V., and Feigin, L. A. (1988) Small-angle-scattering-data treatment by the regularization method, *Acta Crystallogr. 44*, 244−250.
38. Svergun, D. I. (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria, *J. Appl. Crystallogr. 25*, 495−503.
39. Svergun, D. I. (1999) Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing, *Biophys. J. 76*, 2879−2886.
40. Petoukhov, M. V., and Svergun, D. I. (2003) New methods for domain structure determination of proteins from solution scattering data, *J. Appl. Crystallogr. 36*, 540−544.
41. Hu, H., Yu, W., Li, S., Ding, X., Yu, L., and Bi, R. C. (2006) Crystallization and preliminary crystallographic studies of human septin 1 with site-directed mutations, *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun. 62*, 128−132.
42. Eftink, M. R. (1994) The use of fluorescence methods to monitor unfolding transitions in proteins, *Biophys. J. 66*, 482−501.
43. Kay, B. K., Williamson, M. P., and Sudol, M. (2000) The importance of being proline—The interaction of proline-rich motifs in signaling proteins with their cognate domains, *FASEB J. 14*, 231−241.
44. Tompa, P., Szasz, C., and Buday, L. (2005) Structural disorder throws new light on moonlighting, *Trends Biochem. Sci. 30*, 484−489.
45. Versele, M., and Thorner, J. (2005) Some assembly required: Yeast septins provide the instruction manual, *Trends Cell Biol. 30*, 271.
46. Morand, P., Budayova-Spano, M., Perrissin, M., Müller, C., and Petosa, C. (2006) Expression, purification, crystallization and preliminary X-ray analysis of a C-terminal fragment of the Epstein−Barr virus ZEBRA protein, *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun. 62*, 210−214.
47. Harbury, P. B., Zhang, T., Kim, P. S., and Alber, T. (1993) A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants, *Science 262*, 1401−1407.
48. Harbury, P. B., Kim, P. S., and Alber, T. (1994) Crystal structure of an isoleucine-zipper trimer, *Nature 371*, 80−83.

BI061549Z