

Molecular Cloning of Multiple Isoforms of Synaptojanin 2 and Assignment of the Gene to Mouse Chromosome 17A2-3.1

Li-Fong Seet,^{*,†} Stephen Cho,^{*} Andrew Hessel,^{*} and Daniel J. Dumont^{*,†,‡,1}

^{*}Amgen and [†]Ontario Cancer Institutes, Toronto, Ontario; and [‡]Department of Medical Biophysics, University of Toronto, Canada M5G 2C1

Received March 30, 1998

Synaptojanin 2 is an inositol polyphosphate 5'-phosphatase that appears to be regulated by alternative splicing. By screening mouse cDNA libraries derived from either mouse day 16 embryo or adult liver, we have identified additional synaptojanin 2 cDNAs that represent six new isoforms of the protein. This finding, together with other reports, indicates the presence of eight isoforms of synaptojanin 2. Sequence analysis of our cDNA clones suggests that there are at least two putative initiation sites and at least six different sequences coding for the carboxyl-terminus of the molecule. In addition, we have mapped synaptojanin 2 to mouse chromosome 17 band A2-3.1 by fluorescence *in situ* hybridization. © 1998 Academic Press

The signaling system involving lipid molecules is a key pathway whereby growth factor signals are transduced in the cell (1). In response to a variety of stimuli, the ubiquitous membrane inositol lipid, phosphatidylinositol 4,5-bisphosphate [PtdIns(4,5)P₂] is hydrolyzed, resulting in the production of second messengers inositol (1,4,5)-trisphosphate [Ins(1,4,5)P₃] and diacylglycerol. Ins(1,4,5)P₃ can be further phosphorylated by the specific Ins(1,4,5)P₃ 3-kinase to form inositol 1,3,4,5-tetrakisphosphate (IP₄) which may also serve as a second messenger (2, 3, 4, 5). PtdIns(4,5)P₂ can be additionally phosphorylated by phosphoinositide-3-OH kinase [PI(3)K] to form PtdIns(3,4,5)P₃ which is implicated in a number of cellular processes including regulation of exocytosis, cell adhesion and membrane trafficking (6, 7).

The signals generated by the lipid-derived second

messenger molecules are rapidly terminated following the metabolism of these molecules by inositol polyphosphate 5'-phosphatase enzymes (8, 9, 10). This 5'-phosphatase family of enzymes now includes at least ten members, all characterized by the presence of two conserved amino acid sequences in the catalytic domain. There is increasing evidence that these proteins have important biological functions. For instance, the 43 kDa type I 5'-phosphatase has been shown to be involved in inhibiting cellular transformation (11) while deficiency of the oculocerebrorenal (OCRL) 5'-phosphatase is associated with Lowe's syndrome, a condition affecting the lens, brain and kidneys (12, 13). In addition, Ship, a Src homology 2 (SH2)-containing inositol polyphosphate 5'-phosphatase, has been shown to mediate inhibitory signaling in hematopoietic cells (14, 15). Synaptojanin 1, a 5'-phosphatase predominantly expressed in the brain, is implicated in synaptic vesicle trafficking (16).

Several members of the inositol polyphosphate 5'-phosphatase protein family are characterized by the expression of multiple isoforms. These include Ship, synaptojanin 1 and the recently reported synaptojanin 2. Three isoforms of Ship that migrated as 110, 130 and 145 kDa protein species differing in the amino-terminal SH2 coding sequence have been reported (17). The functional significance of these Ship isoforms has yet to be determined. Similarly, synaptojanin 1 has been detected as two protein species of 145 kDa and 170 kDa (16). Synaptojanin 1 is a protein consisting of three domains: an amino-terminal Sac1 homology domain, a central inositol 5'-phosphatase catalytic domain and a carboxyl-terminal proline-rich domain (16). The Sac1 homology domain is implicated in interaction with actin filaments (18) while the carboxyl-terminal proline-rich domain mediates binding with the Src homology 3 (SH3) domains of a number of molecules, including Grb2, amphiphysin and the SH3p4/8/13 family

¹ Corresponding author: Department of Medical Biophysics, University of Toronto, Canada M5G 2C1. Fax: 416-204-2277. E-mail: ddumont@amgen.com.

of proteins (16, 19, 20, 21). The two synaptotjanin isoforms appear to be generated by the use of alternative stop codons, with the longer 170 kDa isoform containing additional proline-rich sequence motifs (22). The closely related synaptotjanin 2 (23) has also been reported to be alternatively spliced, producing two isoforms that differ in the carboxyl-terminal coding sequence (24). This alternative splicing event appears to be developmentally regulated (24), suggesting that the isoforms may be involved in distinct cellular functions during development.

In this study, we report the identification of six new synaptotjanin 2 isoforms: 2 α , 2 β , 2 γ , 2 δ , 2 ϵ and 2 ζ . These isoforms differ mainly in the carboxyl-terminal coding sequences. We have also localized synaptotjanin 2 by fluorescence *in situ* hybridization (FISH) to the A2-3.1 region of mouse chromosome 17.

MATERIALS AND METHODS

Isolation of 5'-phosphatases expressed in C166 endothelial cells by PCR. Degenerate oligonucleotides were designed based on two conserved coding sequences present within the catalytic domain of 5'-phosphatases: FWFGDLNY (sense 5'-NNNTGGNNNGGNGACC/TNNNAAC/TTA/TCT-3') and RAPAWCD (antisense 5'-NTT/CNTG/ANNNT/CAGNG/CT/AGGTNNNNCC-3') (N represents any of the four nucleotides). Total RNA from C166 cells (25) was prepared with the RNeasy Total RNA (Qiagen) kit. First strand cDNA was obtained using the Moloney murine leukaemia virus reverse transcriptase according to the manufacturer's protocol (Clontech). PCR was performed with the AmpliTaq DNA polymerase (Perkin Elmer). Amplified products were subcloned into the pCR-TRAP vector (GenHunter), sequenced and identified by BLAST searches against the gene database.

Cloning of synaptotjanin 2 α . One clone obtained by the above procedure contained a then novel 5'-phosphatase sequence. A 1.9 kb mouse placenta expressed sequence tag (EST, AMGEN database) with partial sequence identity to this clone was used to design PCR primers for Rapid Amplification of 5'-cDNA Ends (5' RACE) from a mouse brain Marathon-Ready cDNA library (Clontech). An amplified product containing a potential start ATG codon was obtained. The 5' translated and untranslated sequence was further confirmed by performing 5' RACE on a mouse adult heart Marathon-Ready cDNA library (Clontech). One amplified product has identical untranslated sequence to that obtained with the adult brain cDNA library. Two other amplified products contained 5' sequences present in the 2 β putative isoforms.

Cloning of additional synaptotjanin 2 isoforms. A 16 day mouse embryo λ EXlox cDNA library (Novagen) and an adult mouse λ gt11 cDNA library (Clontech) were screened for additional synaptotjanin isoforms. A *Stu*I DNA fragment corresponding to nucleotide 855-2945 of synaptotjanin 2 α was used as a probe to hybridize to DNA from 1×10^6 plaques of each library.

Chromosome *in situ* hybridization. A mouse synaptotjanin 2 genomic fragment was obtained by screening a bacterial artificial chromosome (BAC) mouse embryonic stem cell library (Genome Systems, Inc.) with a cDNA probe corresponding to nucleotide 1-639 of synaptotjanin 2 α . The genomic clone was verified by sequence analysis of a *Sma*I-digested fragment that hybridized to a probe corresponding to nucleotide 1-594 of synaptotjanin 2 α . For fluorescence *in situ* hybridization (FISH), the genomic fragment was labeled by nick translation with digoxigenin dUTP and hybridized to mouse embryo fi-

broblast metaphase chromosomes in 50% formamide, 10% dextran sulfate and $2 \times$ SSC. Specific hybridization was detected by incubation with fluoresceinated anti-digoxigenin antibody and counterstaining with DAPI (4',6-diamidino-2-phenyl-indole dihydrochloride). Confirmation for chromosome 17 was obtained by co-hybridization with a probe specific for the telomeric region of chromosome 17 (Genome Systems, Inc.).

RESULTS

Isolation of Synaptotjanin 2 α

In order to identify 5'-phosphatases expressed in the C166 yolk-sac fps/fes transformed endothelial cell line (25), we have used the polymerase chain reaction (PCR) (26) with degenerate primers corresponding to nucleotide sequences coding for amino acid residues that are highly conserved within the catalytic domain of known inositol and phosphatidylinositol polyphosphate 5'-phosphatases. PCR was performed on first-strand cDNA prepared from C166 mRNA as template. Amplified products, averaging about 250 bp in size, were cloned and sequenced. Using this approach, we identified the expression of the putative mouse homologues of synaptotjanin 1, 51C (GenBank Accession No. L36818) and synaptotjanin 2. Subsequent searching of the AMGEN database revealed a 1.9 kb mouse placenta expressed sequence tag (EST) that has sequence identity to synaptotjanin 2. This EST contains the putative 3' coding sequence of the synaptotjanin 2 gene, up to and including the polyadenylation sequences. To further extend the 5' sequence of the novel gene, we used the Rapid Amplification of 5' cDNA Ends (5' RACE) on mouse brain and heart Marathon-Ready cDNA libraries. The 5' RACE was performed with a series of oligonucleotide primers designed from the EST sequence. From the deduced amino acid sequence of the primer-extended cDNAs, we identified one putative open reading frame initiating at nucleotide 283 that encodes a protein with an approximate molecular mass of 140 kDa (1216 amino acids). The compiled 4.1 kb cDNA sequence and its predicted amino acid sequence are shown in Fig. 1.

The amino acid sequence of synaptotjanin 2 is about 40 to 60% homologous to other members of the 5'-phosphatase family, the greatest being to synaptotjanin 1 (16). Like synaptotjanin 1, it contains an amino-terminal Sac1 homology domain, a central inositol 5'-phosphatase domain and a carboxyl-terminal proline-rich region. Rat synaptotjanin 2 (23) contains over 95% amino acid sequence identity to our clone. There are, however, three major differences: the presence of additional 5' sequence indicating a more upstream initiation codon, the absence of a 46 amino acid sequence after Lys982 and no sequence homology after Lys1151. A recently reported mouse synaptotjanin 2 sequence (24) appears to be the homologue of the above rat sy-

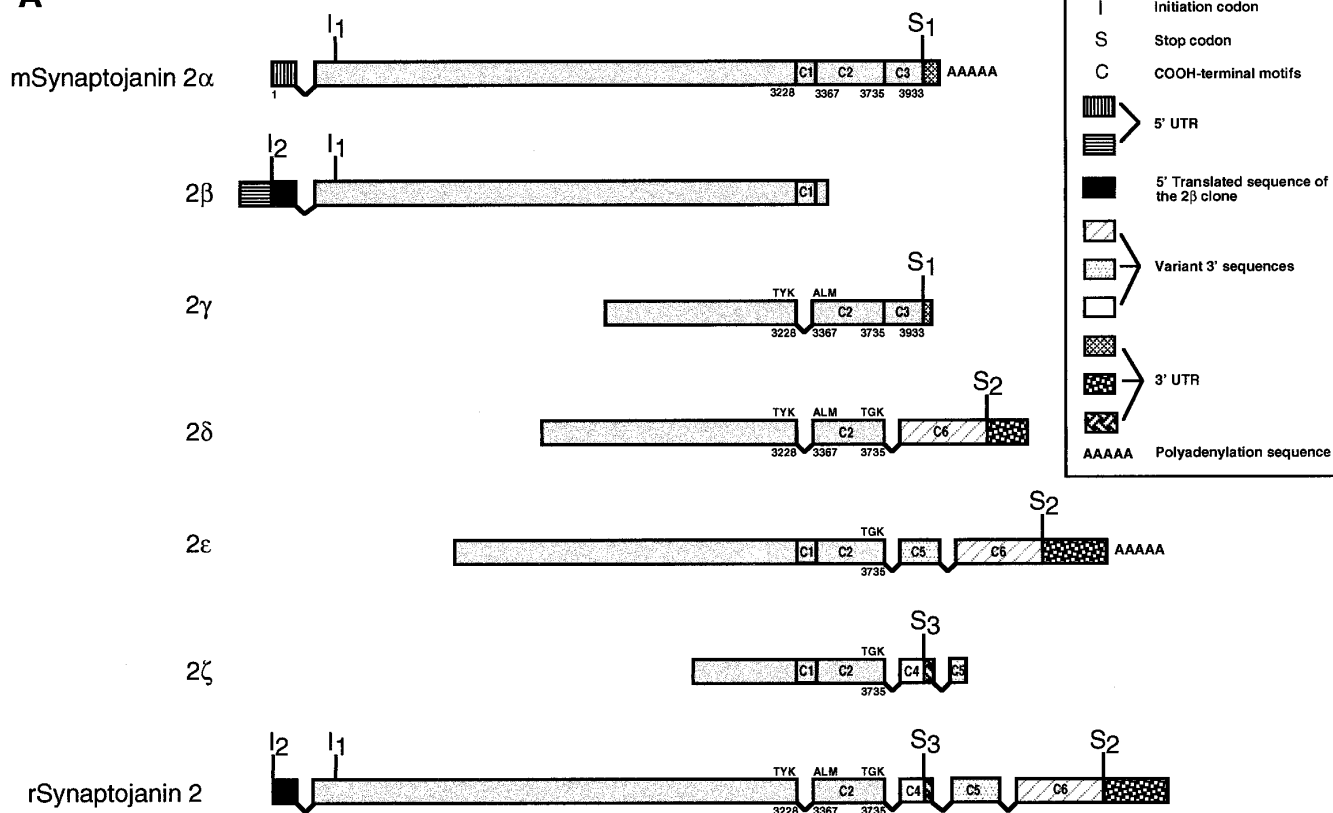
1	AAAACCCATACACATGATGTCTGTGGGCGGAGCTATTGTCTGCTTTCCTTTTTCGCGTG	2101	CAGGTGAAGGAGAGGAATGAAGACTACCGGGAGATCAGCCACAACTCTCCCTCCCTTCG	626
61	GGTGTGAGCCAGATTGGCAAGCAGCTGTGAATGAACAGCAGTGAACGAGCCACTCGCT		Q V K E R N E D Y R E I T H K L S F P S	
121	TGCTGATGTCTCTCTGATGTGTTTCCCGCCAGCTCCGGAAGAGAAAGATCATTA	2161	GGGGAACACATATTTTCACATGATTACGTGTTTGGTGTGGCGATTTCACCTACCGTATT	646
181	GGACTGTATAGCAAGCTGACGGATGCCTATGCTGCTCGGGGAGCTGAGGTACAAATCC		G R N I F S H D Y V F W C G D F N Y R I	
241	GGTGGCTCCCTTGAGCTTCTGCTGTGGTGACAGGCTGCATGTCAGTGGGCGAAT	2221	GATCTTACTTACGAAGAAGTCTCTTATTTGTTAAACGCCAAGACTGGAAGAAGCTTATG	666
	M S V G R I	2281	GAATTTGATCAGTTACAGTTACGAATCAAGTGGAAAAATTTTAAAGACTTTCATGAA	686
301	CCAGATCCAGAGATCTACAAAATCACTGCCATGATGTACCCCTGCAGGAAGAGGCC		E F D Q L Q L Q L Q K S S G K I F K D F H E	
	P D A E I Y K I T A T E L Y P L T L Q E A	2341	GGAGCCGTTAACCTCGGACCCACCTACAAGTATGACCTGGATCAGCTGCCATACACACA	706
361	AAGGAAGAGGACCGCTGCCACCTTAAAGAAATCTGAGCTCAGSGGTGTCTATTTC		G A V N F G P T Y K Y D V G S A A Y D T	
	K E E D R L P L T L K K I L S S G V P Y F	2401	AGTGACAAGTGCCTACCCAGCCTGGACAGACAGGCTGTGTGGTGGAGGAAGAAGCAT	726
421	GCATGGCCCAATGATGGCGCTGCTTCGATCTGACCATCAGGGCTCAGAAACAGGGTGAT		S D K C R T P A W T D R V L W R K K H	
	A W P N D G A C A F D L T I R A Q K Q G D	2461	CCATATGATAAGACAGCTGGTGAACCTCAACCTTCTAGACAGCATCTAGACGGCAGTCCC	746
481	GACGGCTCTGAATGGGGACCTTCTTCTTGGAAACAGCATATGATGCTGCTCGCGG		P Y D K T A G E L N L L D S D L D G D P	
	D G S E W G T S F P F W N Q L L H V P L R	2521	CAAACTCAGACACTGGTCTCAGGACCTTGAATACTACGGCGTGCAGAGCTGCAG	766
541	CAGCACCAGGTGAATGTCTAATCGTGTGCTGAAAGTCATCTGTGGGTGGTGACCATC		Q I R H T W S P G T L K Y V G R A E L Q	
	Q H Q V N C H N N W L L K V I C G V V T I	2581	CGCTGTGATCAGACACTGTGCTGCCATTGTGGAGTGGAGGTTCAAGAGTGGATGTA	786
601	CGCAGCATATATGCTCCCAACAGGCAAGCGCTGTCTCATCTCTCGCATCAGCTGT		A S D H R P V L A I V E V E V Q E V D V	
	R T V Y A S H K Q A K A C L I S R I S C	2641	GGAGCCCGGAGAGGCTTCCAGGAAGTGTCTCTGTCCAAGGCCGCTGGATGCCACC	806
661	GAACGCGCAGGTGCTGCTTCTCACCCTGGTGTGAACGATGATGCCACGTGTCCAAC		G A R E R V F Q E V S S V Q G P L D A T	
	E R A G A G A R P L T R G V N D D G H V S N	2701	GTTGTTGAACCTCCAGCTTCAACCTAGAGAAATAATTTCCAGAGGACCTG	826
721	TTTGTGGAGACAGCAGACGATTTACATGGATGATGAGTATCGCTTGTGCCAGATC		V V V N L Q Q S P T L E E K N E F P E D L	
	F V E T E Q T I Y M D D G V S S P V Q I	2761	CGCAGACAGCTTATCGAGCTTGGGAATATGGGAGCTTATTCTAGTCAGGATCAAC	846
781	CGAGGCTCCGTTCCGCTGTTCTGGGAGCAACCAAGACTTCAGGTGGCTCCCATCATCTG		R T E L M Q T L G N Y G T I I L V R I N	
	R G S V P L P F I D R P V L Q V G S H L	2821	CAAGGCAAGTGTGGTGCAGCTTTCAGACAGCCACTCGCTCTCAGTGTGCTGGATGTG	866
841	AGACTGCACAGAGGCTAGGGGCAACACTCTCTGCTTTTGAAGGCACATGCTGTCTCTG		Q G M L V T F A D S H S A L S V L D V	
	R L E R G L G A N A P A F E R H M V L L	2881	GATGGTATGAAGGTGAAGAGGCGCTGGAAGTTCGACCAAGACCAAGATTGGCTA	886
901	AAGGAGCAATACGGTAAGCAGGTGTGTGAACCTGCCGGTACGACAGGCGGTGAAGAG		D G M K V K G R A V K I R P K T K D W L	
	K E Q Y G K Q V V V N L P G S R G G E	2941	GAAGCGCTGAGAGGAGCTTCCGGAAGAGGACAGCATGAGCCCTGTGTCTCCACC	906
961	GTGCTCAACAGAGCTTCAAGAAGTGTCTGGGCTTCTTGCACGCGGCTGACACACT		E G L R E E L L R K R D S M A P D S P T	
	V L N R A P F K L L N W A S C H A G D T	3001	GCCAACCTCTGTGTTGGAGGAGAACTTGTACTCTCGAGTCTGGACTATGATCCGAA	926
1021	ATGATAAATTTGACTTCCATCAGTTTGCACAAAGTGAAGACTAGAGAAATTTGAGAAC		A N S C L L E E N F D F S S L D Y E S E	
	M I N F D P F H Q P A K G R K L E K L E N	3061	GGGGATGTCTTGAAGAGGATGAAGACTATTAGTGGATGGGTTTGGCCAGCTGTGATG	946
1081	CTGTTGAGACCTCAGTTACAGCTACACTGGGAAGACTTCGGCGTGTTCGGAAGGCGAG		G D V L E E D E D Y L V D G F G Q P V V	
	L L R P Q L L H W E D P G V F A K G E	3121	TCAGACAGTGAAGCTCGGTGGAGCAACTCTTCGACACCATGAGCTCTTTCGACACCGCC	966
1141	AATGTAAGTCAACGTTTCAGAAAGGCACTTCGCGGATGAACCTGTCTCGACTGTCTGGAT		S D S E L G D N S D S D T C M A T S S L T P A	
	N V S P R P F Q K G T L R M N C L D C L D	3181	AGCAAGTCTCCCGCTTGGCTAAAGAGGACACATCAACATCAAAAGACGAGCTCAC	986
1201	AGAACAACATCTGCAAGTCTTCACTGCTTGAAGTCTCTCACTGCAAGTCTGAGAGC		S K S P A L A K K K Q H P T Y K D D A H	
	R T N T V Q C F I A L E V L H L Q L E S	3241	CTGGTACCTTAAAGCAGGAGCTGGAAGTGTGCTGGGAATTTTCGCGCCCTTCCGAGC	1006
1261	TTGGGCTAAATTCAGAGCCCATATTGACCGTTTGTGGAGTCTTCAAGAGGCTGTGG		L V T L K Q E L E V A G N F R H R S P S	
	L G L N S K P I I D R P V E S F K A M W	3301	AGGTCCCTGTCGTTCCCAATAGGCTCGGCCCTACCCACACAGAGACGCCCCCTCT	1026
1321	TCTCTGAATGGGACAGCTGAGCAAGTGTTCACAGGAGCAGGCGCTTGGGAAGGAAG		R S L S V P N R P R P P P O R P P P	
	S L N G H S L S K V F T G S R A L E G K	3361	CCAAGTGGCTTAATGGTGAAGAGTCAAGCTCAGACGGCTCCATCTTCTGCGACTCAT	1046
1381	GCCAGAGTGGGAAGCTGAAGGATGGGGCCGATCCATGTCTCGACCATCCAGTCCAAC		E T G L M V K K S A S D A S I S S G T H	
	A K V G K L K D G A R S M S R T I Q S N	3421	GGACAAATTTCCATCTTTCAGACAGCGAAATTTCTCCAGGAGCAGCCCAAGCAACCC	1066
1441	TTCTTCGACGGGTTGAAGCAGGAGCCATCAAGCTACTGCTAGTCCGAGATGTCTACAAT		G Q Y S I L Q T A K L L E G A P O Q P P	
	F F D G V K Q E A I K L L L V G D V Y N	3481	AAGGTAGAAGTGAATAGTAAACCTTACAAGCTCAACAGATCAAAACCAACAGCT	1086
1501	GAAGAGTCTACAGACAAAGGAGGATGCTGCTGGACAAACAGCGCCCTTCTGGCGACCC		K A R T G I S K P Y N V K Q I K T T N A	
	E E S T D K G R M L L D N T A L L A T P	3541	CAGGAGGCAAGGAGCTATCCGCTGTCTTCCGGAAGTACGGAGGGGTCCCGGAATCA	1106
1561	AGGATCTTGAAGGCCATGACAGACGCGCATCGGAATTCAGAAATTTCAAGCGGATCCAG		Q E A E A A I R C L L E A S G G V P E S	
	R I L K A M T E R Q S E F T N F K R I Q	3601	GCCCCAGTGCCATACCCCTGAGAAACCAAGGCTCTTCAAGCCAGAGGCGACCCCTGGG	1126
1621	ATTGCTTGGGACCTTGAATGTGAACGAGGAAAGCAATTCCTGACAAATCTCCTGGGG		A P G A I P L R N Q G S S K P E A T L G	
	I A V G T W N V N G G K Q F R S N L L G	3661	CCCCAGCCCTGCGCCGCGCTGTCTCAAGGCTTCCCATATGAAGAACAACCAATTTG	1146
1681	ACGGCTGAGCTCAGGACTGGCTCTAGATGCTCTCAGTGTGAGGAGCAGTGGACTCC		P P A L P E R R P A P R V E T M M K K P T L	
	T A E L T D W L L D A P Q L S G A V D S	3721	AGGAGGACAGGAAGGTGTACTCGGCTCTCCAGTGTCTGCGTGAAGAGCTGCGCTCT	1166
1741	CAGGATGATGGCAGTCTGCTGATGATTTTGCATCGGGTTGAGGAGATGGTGAACCTG		R R T G K V Y S G I S Q C L R E E L R S	
	Q D D G S P A D V F A I G F E E M V E L	3781	GCTGCTGCACCCACACGAGTGTGCTGACAGGACTCGGAGACCTTAAACACAGATGG	1186
1801	AGTCCGGGAATATTGTCAATGCCAGCACCAACAGGAAGATGTGGGGCGAGCAGCTT		A A C T P H A V S A Q D C G D L N N R W	
	S A G N I V N A S T T N R K M W G E Q L	3841	AGAATGCCAGTCTTCTGCACTATATCACACAAAAATGAAGAACGATCTCTCTCAGT	1206
1861	CAGAAAGCCATCTCCGCTCCCATCGGTACATCTCTTGACCTCCGACAGCTGTGGGG		R M P R F S H Y I H T K K K N V S L S	
	Q K A I S R S H R Y I L L T S A Q L V G	3901	TTTCAAGATCTGCTGCTAAAGTTTCGAGATGAAGTGTAGAGTACGTTGGTACAAAAAG	1217
1921	GTCTGTCTTTACATCTTTGTACGTCCTGACAGCTCCGCTTTCATCAGAGAGCTGGCCATC		F Q D L W L K F R R *	
	V C L Y I F V R P Y H V P F I R D V A I	3961	CATCTTTGTCTTCGGCTTCAAAATGTTCAAGTGGTGTGACGTCCACAGGAATTAGATCCC	
1981	GACACCGTGAAGACCGGATGGGGGAAAGCGGGAATAAGGGTGTCTGGGCGATCCGC	4021	CAATGCTAAATCTGAAGAAAGCAAGGTGTCAATAAACAAGGGAAGCCAAAAA	
	D T V K T G M G G K A G N K G A V G I R	4081	AAAAAAAAAAAAAAAAAAAA	
2041	TTCCAGCTCCACAGCAGGATTTCTGCTGTGTGACAGCTGACGCGTGGCAGTCT			
	F Q L H S T S F C F V C S H L T A G Q S			

FIG. 1. Molecular structure of the novel synaptotjanin 2 α isoform. The predicted amino acid sequence is shown below the nucleotide sequence. The SacI homology region is in boldface italics and the conserved sequence motifs that define the 5'-phosphatase family of enzymes are indicated in boldface. The consensus PXXP sequences that may mediate SH3 domain interaction are underlined.

naptojanin 2 isoform and shows similar sequence deviation from our clone. A further Blast search revealed the presence of a human brain cDNA (Accession No. AB002346; 27) which has 86% nucleotide sequence identity to our clone. The report, however, indicates the presence of an open reading frame initiating downstream at Met299 of our sequence and significant sequence deviation after Lys1151. These three reported sequences may represent differentially spliced isoforms of synaptotjanin 2. Hence, we will refer to our compiled sequence as synaptotjanin 2 α .

Cloning of Additional Synaptotjanin 2 Isoforms

Northern blot analysis of synaptotjanin 2 indicated the presence of multiple transcripts (data not shown and 23). In an attempt to isolate other putative synaptotjanin 2 isoforms, we screened a murine liver cDNA library and an embryo day 16 cDNA library. Nine positive clones were obtained from the mouse liver cDNA library. In particular, three partial synaptotjanin 2 cDNA clones, which we have designated as 2 β , 2 γ and 2 δ (3.7, 1.8 and 2.8 kb respectively), were found to have

A**B**

C1 DDHLVTLKQLELVAGNFRHRSRSLSV**PNRPRPPHPORPPPT**
 C2 ALMVKKSSADASISSGTHGQYSLQTAALL**PGAPQOPPKARTGISKPYNVKQIKTTNAQEA**AAIRCLLEASGGVPESAPGAIPLRNQGSSKPEATLGP**FALPRRPA**RVPTMKKPTLRRTGK
 C3 VYSGISQCLREELRSAACTPHVAQAQDCGLNNRWMPFESHYIHTKKWKNVLSFPQDLWLKFRF
 C4 IVPCNSQASQPCLLLRHFVTVAAQRLTPIDASGSSV
 C5 **PML**PEENFEPQVHFTMASQEMLETPPPITAT**P**IPVPKPRTLQPGKGVGGRPSSGKPEPEDEAPSVT
 C6 GTVESPPPEAQEAPSLA**PKV**PRRKRKSAPAAFLQLVQLQNSQVLQGLTCSSSSPPSLKPDTHPLCLQVALGTSSARSPETHGPRVTEPEAASFHGNYPDPFWSLHHKPLLNNTWLSKSEPLDVGSRNPERTHTEPAQVNASLAERGLPPDHGKDLSHWVTASNKDKRTTLGV

FIG. 2. Synaptojanin 2 isoforms. (A) Schematic of synaptojanin 2 cDNAs. Six possible variants of synaptojanin 2 were identified in this study. The predicted carboxyl-terminal region of mouse (m)Synaptojanin 2 that vary amongst the isoforms are represented by sequence blocks designated from C1 to C6 (see text). The sequence of the reported rat (r)Synaptojanin 2 isoform (23) is also represented here for comparison. Numbers shown below the bars indicate the nucleotide positions of the synaptojanin 2α cDNA sequence as well as positions where putative splicing is predicted to have occurred. The last three amino acid residues at the putative splice site before C1 and the first and last three amino acids of C2 are shown above the bars. I₁ and I₂ indicate the alternative initiation sites. S₁, S₂ and S₃ indicate the stop codons that are determined by the different carboxyl-terminal sequences. Synaptojanin 2α and the 2γ have identical 3' UTR. Similarly, 2δ and 2ε share identical 3' UTR. (B) Predicted coding sequences of the six carboxyl-terminal sequence regions. C1, C2 and C3 correspond to nucleotides 3229–3366, 3367–3735 and 3736–3930 of the synaptojanin 2α cDNA sequence, respectively. PXXP sequence motifs are in boldface and underlined. The accession numbers for 2α, 2β, 2γ, 2δ, 2ε, 2ζ are AF041862, AF041857, AF041860, AF041858, AF041859, and AF041861 respectively.

sequences that are unique in comparison to synaptojanin 2α (Fig. 2A). Clone 2β is distinct from 2α in that it contains a putative initiation site 252 bp upstream from that predicted in 2α (Fig. 2A, I₂). This initiation site and the associated 5' coding sequence corresponds to that reported for rat synaptojanin 2 (23). Clone 2γ, however, differs from 2α by the absence of a 138 bp (nucleotide 3228–3366; 46 amino acids) 3' sequence, which is hereby designated as the C1 carboxyl-terminal

sequence. Clone 2δ is similarly deleted of the C1 sequence and contains a unique 528 bp (176 amino acids) 3' sequence (designated as C6) that substituted for the last 198 bp (66 amino acids) of 2α (designated as C3).

Screening of the day 16 mouse embryo cDNA library produced three positive clones. Two of the partial synaptojanin cDNA clones, which we have designated as 2ε and 2ζ (3.8 and 1.4 kb respectively), were found to contain additional unique sequence motifs. Clone 2ε

has a unique 204 bp (68 amino acids) 3' sequence (designated as C5) that is flanked by the C2 sequence at its 5' end and the C6 sequence at its 3' end. It is noted that the previously reported human brain EST (27) carries a 3' sequence that has 70% nucleotide sequence identity to this C5 to C6 sequence, suggesting that this is the human homologue of the 2 ϵ isoform. In addition, the recently reported partial mouse synaptojanin 2 sequence (24) contains 100% amino acid sequence identity to C5 and C6, confirming that this sequence does indeed represent an isoform of synaptojanin 2. The clone 2 ζ had yet another unique 123 bp (41 amino acids) sequence, designated as the C4 sequence, substituting for the C3 sequence of 2 α . The predicted amino acid sequence of C4 is 92% identical to the homologous region of rat synaptojanin 2 (23) and 100% identical to a similar region in the recently reported mouse sequence (24), indicating that both the latter clones are isoforms similar to 2 ζ . The 3' untranslated region (3' UTR) of the 2 ζ cDNA is similar to that reported for rat synaptojanin 2. Close examination of the 3' UTR of rat synaptojanin 2 reveals the presence of both C5 and C6 sequences following C4. The C5 and C6 sequences are in a different reading frame from C4 and are separated from the C4 stop codon by 26 nucleotides (Fig. 2A).

Our results therefore indicate the presence of several potential alternative RNA splice sites in the synaptojanin 2 gene. There may be at least one splice site at the 5' end, resulting in multiple initiation sites for the gene (Fig. 2A, I₁ and I₂). Alternatively, the different 5' untranslated regions (5' UTRs) may be the result of different promoter usage. The 3' end appears to be the most complex as the carboxyl-terminal region of the product is potentially encoded by at least 6 exons, which when alternatively spliced, can produce multiple isoforms. There are three potential splice sites in the synaptojanin 2 α sequence, at nucleotides 3228, 3367 and 3735 (Fig. 2A). It is noted that all of the clones reported here and the isoforms reported elsewhere (23, 24) have one 3' sequence in common, namely C2, suggesting that this sequence may be important to the function of the protein. In addition, all of the unique sequences identified in the synaptojanin 2 cDNA clones, except for C3 and C6, contain proline-rich regions (Fig. 2B), suggesting that alternative splicing at the carboxyl-terminus may lead to differential binding to SH3 domain-containing proteins (28). Furthermore, the presence of both C5 and C6 sequences after the C4 stop codon in the 2 ζ clone and the reported rat synaptojanin 2 cDNA (23) suggests a possible 3' exon structure for synaptojanin 2 in which the C1 exon(s) precedes that of C2 which is followed by the exon(s) of C4, C5 and finally C6.

Chromosomal Mapping of Mouse Synaptojanin 2

In order to determine the synaptojanin 2 genomic locus, we obtained a BAC synaptojanin 2 genomic clone

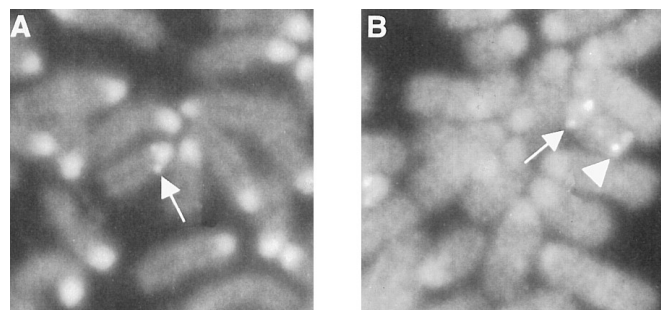


FIG. 3. Chromosomal localization of the synaptojanin 2 gene by fluorescence *in situ* hybridization. The chromosomal location of the mouse synaptojanin 2 gene was determined by using a mouse synaptojanin 2 genomic fragment as a digoxigenin-labeled probe on mouse metaphase chromosome spreads. (A) Synaptojanin 2 is located on mouse chromosome 17. The hybridization signal produced by the synaptojanin 2 genomic probe is indicated by the arrow. The DAPI-stained chromosomes revealed that synaptojanin 2 maps adjacent to the heterochromatic region on chromosome 17 on band A2-3.1. (B) Synaptojanin 2 co-localizes with a chromosome 17 specific probe. To verify the chromosomal localization of synaptojanin 2, a probe specific for the telomeric region of chromosome 17 was used in a co-hybridization experiment. The signal generated by this control probe is indicated by an arrowhead.

by screening a mouse embryonic stem cell library. Partial sequencing of this BAC clone confirmed that it contains the synaptojanin 2 gene. This genomic DNA was used to map the genomic locus of synaptojanin 2 by fluorescence *in situ* hybridization (FISH). The signal specific for synaptojanin 2 was detected on chromosome 17 in 73 of 80 mouse embryo fibroblast metaphase chromosome spreads analyzed (Fig. 3A). The identity of the chromosome was confirmed by co-hybridization of the synaptojanin 2 probe with a chromosome 17-specific probe (Fig. 3B). DAPI staining revealed that synaptojanin 2 maps immediately adjacent to the heterochromatic region on chromosome 17, corresponding to band A2-3.1, in a region close to the mouse T-locus.

DISCUSSION

Synaptojanin 2 is an inositol 5'-phosphatase consisting of an amino-terminal Sac1 homology domain, a catalytic domain and a carboxyl-terminal region that is alternatively spliced. We report here the identification of several synaptojanin 2 cDNA clones that differ in both the amino- and carboxyl-terminal coding sequences. Our data suggest the existence of two alternative initiation sites for the gene. Initiation from these alternative start sites results in isoforms that vary in the expression of 84 amino acids in the amino-terminal of the Sac1 homology domain. A third putative initiation site has been suggested for a human cDNA sequence isolated from brain (27). The human sequence differs from the synaptojanin 2 α sequence by the inser-

tion of a 57 nucleotide sequence at nucleotide 1155 of synaptojanin 2 α . This insertion sequence contains a stop codon, resulting in an open reading frame that begins 298 amino acids downstream of the predicted initiation site in synaptojanin 2 α . In this case, the putative translated product lacks most of the Sac1 homology domain.

The Sac1 homology domain was originally identified in yeast in the Sac1p protein and it is thought to modulate both the secretory pathway and actin cytoskeletal activity (29, 30). Recent data indicate that Sac1p is associated with the endoplasmic reticulum (31) and has a role in mediating ATP transport through the endoplasmic reticulum (32). To date, synaptojanin 1 and 2 are the only mammalian proteins known to contain this Sac1 homology domain. A possible role for this domain has been suggested by the demonstration that the Sac1 homology region of synaptojanin 1 interacts with actin *in vitro* (18). Hence, truncation or deletion of the Sac1 homology domain may modify the function of synaptojanin 2.

The carboxyl-terminal sequence of synaptojanin 2 is encoded by at least six exons. Alternative splicing of these exons in specific combinations can result in the expression of multiple synaptojanin 2 isoforms. We report in this study the identification of six new isoforms that differ at the carboxyl-terminus. If these isoforms are also alternatively spliced at the amino-terminus, resulting in initiation at either of the two potential start sites mentioned above, then at least twelve possible isoforms can be generated. This complexity is reflected by the multiple transcripts that hybridized to synaptojanin 2 α probes in the Northern analyses (data not shown and 23).

The carboxyl-terminal of synaptojanin 2 is also characterized by the expression of the PXXP sequence motif that mediates interaction with SH3 domains (28). We have evidence that the proline-rich carboxyl-terminal region of synaptojanin 2 α have the ability to bind to the SH3 domains of a variety of molecules *in vitro* (Seet, L. F., and Dumont, D. J., unpublished data). The interaction with SH3 domains is likely to be critical to the function and/or regulation of the protein. For instance, synaptojanin 1 is thought to facilitate synaptic vesicle recycling via its interaction with the SH3 domain of amphiphysin, a presynaptic protein implicated in endocytosis (16). It has also been suggested that synaptojanin 1 function is regulated by similar SH3-mediated interactions with members of the SH3p4/8/13 family of protein molecules (19, 20, 21).

As the carboxyl-terminal isoforms of synaptojanin 2 differ in the expression of the PXXP sequence motif, the individual isoforms may interact with distinct SH3 domain-containing molecules. This may in turn influence the cellular location, catalytic activity and possibly signal transduction of the different isoforms. This

phenomenon has been indicated for synaptojanin 1 where the additional SH3-binding domains of the 170 kDa isoform is thought to be responsible for the increased membrane affinity of the protein relative to the 145 kDa isoform (22).

Finally, our FISH mapping data indicate that synaptojanin 2 is located on mouse chromosome 17 within the t-complex region. This region is known to contain many genes involved in spermatogenesis. As synaptojanin 2 is expressed in the testes, it is possible that synaptojanin 2 activity may affect sperm function.

ACKNOWLEDGMENTS

We thank Vadim Kourkitchi and Sarah Jacobs for technical assistance. We also thank Dave Grosshans and Mike Thomas for sequencing the DNAs. Oligonucleotide synthesis was performed at Amgen Boulder. This work was funded by Amgen and the Medical Research Council (MRC) of Canada (D.J.D.).

REFERENCES

1. Toker, A., and Cantley, L. C. (1997) *Nature* **387**, 673–676.
2. Berridge, M. J. (1993) *Nature* **361**, 315–325.
3. Llinas, R., Sugimori, M., Lang, E. J., Morita, M., Fukuda, M., Niibobe, M., and Mikoshiba, K. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 12990–12993.
4. O'Rourke, F., Matthews, E., and Feinstein, M. B. (1996) *Biochem. J.* **315**, 1027–1034.
5. Cullen, P. J., Hsuan, J. J., Truong, O., Letcher, A. J., Jackson, T. R., Dawson, A. P., and Irvine, R. F. (1995) *Nature* **376**, 527–530.
6. Schiavo, G., Gu, Q.-M., Prestwich, G. D., Sollner, T. H., and Rothman, J. E. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 13327–13332.
7. Kharlund, J. K., Guilherme, A., Holik, J. J., Virbasius, J. V., Chawla, A., and Czech, M. P. (1997) *Science* **275**, 1927–1930.
8. Downes, C. P., Mussat, M. C., and Michell, R. H. (1982) *Biochem. J.* **203**, 169–177.
9. Connolly, T. M., Bross, T. E., and Majerus, P. W. (1985) *J. Biol. Chem.* **260**, 7868–7874.
10. Mitchell, C. A., Brown, S., Campbell, J. K., Munday, A. D., and Speed, C. J. (1996) *Biochem. Soc. Trans.* **24**, 994–1000.
11. Speed, C. J., Little, P. J., Hayman, J. A., and Mitchell, C. A. (1996) *EMBO J.* **15**, 4852–4861.
12. Attree, O., Olivos, I. M., Okabe, I., Bailey, L. C., Nelson, D. L., Lewis, R. A., McInnes, R. R., and Nussbaum, R. L. (1992) *Nature* **358**, 239–242.
13. Zhang, X., Jefferson, A. B., Auethavekiat, V., and Majerus, P. W. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 4853–4856.
14. Ono, M., Bolland, S., Tempst, P., and Ravetch, J. V. (1996) *Nature* **383**, 263–266.
15. Ono, M., Okada, H., Bolland, S., Yanagi, S., Kurosaki, T., and Ravetch, J. V. (1997) *Cell* **90**, 293–301.
16. McPherson, P. S., Garcia, E. P., Slepnev, V. I., David, C., Zhang, X., Grahs, D., Sossin, W. S., Bauerfeind, R., Nemoto, Y., and de Camilli, P. (1996) *Nature* **379**, 353–357.
17. Kavanaugh, W. M., Pot, D. A., Chin, S. M., Deuter-Reinhard, M., Jefferson, A. B., Norris, F. A., Masiaz, F. R., Cousens, L. S., Majerus, P. W., and Williams, L. T. (1996) *Curr. Biol.* **6**, 438–445.

18. Sakisaka, T., Itoh, T., Miura, K., and Takenawa, T. (1997) *Mol. Cell. Biol.* **17**, 3841–3849.
19. de Heuvel, E., Bell, A. W., Ramjaun, A. R., Wong, K., Sossin, W. S., and McPherson, P. S. (1997) *J. Biol. Chem.* **272**, 8710–8716.
20. Ringstad, N., Nemoto, Y., and de Camilli, P. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 8569–8574.
21. Micheva, K. D., Kay, B. K., and McPherson, P. S. (1997) *J. Biol. Chem.* **272**, 27239–27245.
22. Ramjaun, A. R., and McPherson, P. S. (1996) *J. Biol. Chem.* **271**, 24856–24861.
23. Nemoto, Y., Arribas, M., Haffner, C., and DeCamilli, P. (1997) *J. Biol. Chem.* **272**, 30817–30821.
24. Khvotchev, M., and Sudhof, T. C. (1998) *J. Biol. Chem.* **273**, 2306–2311.
25. Wang, S.-J., Greer, P., and Auerbach, R. (1996) *In Vitro Cell. Dev. Biol.-Anim.* **32**, 292–299.
26. Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A., and Arnheim, N. (1985) *Science* **230**, 1350–1354.
27. Nagase, T., Seki, N., Ishikawa, K.-I., Ohira, M., Kawarabayasi, Y., Ohara, O., Tanaka, A., Kotani, H., Miyajima, N., and Nomura, N. (1996) *DNA Res.* **3**, 321–329.
28. Pawson, T. (1995) *Nature* **373**, 573–580.
29. Novick, P., Field, C., and Schekman, R. (1980) *Cell* **21**, 205–215.
30. Cleves, A. E., Novick, P. J., and Bankaitis, V. A. (1989) *J. Cell Biol.* **109**, 2939–2950.
31. Whitters, E. A., Cleves, A. E., McGee, T. P., Skinner, H. B., and Bankaitis, V. A. (1993) *J. Cell Biol.* **122**, 79–94.
32. Mayinger, P., Bankaitis, V. A., and Meyer, D. I. (1995) *J. Cell Biol.* **131**, 1377–1386.