GENOMICS 24, 20–26 (1994)

Human Neutral Amino Acid Transporter ASCT1: Structure of the Gene (SLC1A4) and Localization to Chromosome 2p13–p15

KAY HOFMANN,* MARIA DÜKER,* THOMAS FINK,† PETER LICHTER,† AND WILHELM STOFFEL*,[1]

*Institut für Biochemie, Universität Köln, Joseph-Stelzmann-Strasse 52, D-50931 Köln, Germany; and †Institut für Virusforschung, Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 280, D-69120 Heidelberg, Germany

Received April 7, 1994; revised July 10, 1994

Screening for cDNAs encoding proteins similar to the sodium-coupled glutamate transporter GLAST1 led to the isolation of a cDNA clone coding for a protein that turned out to be identical to the recently described neutral amino acid transporter ASCT1. The new member of the GLAST-related transporter family does not transport glutamate or aspartate but alanine, serine, cysteine, and threonine instead. The expressed sequence tag *EST02446*, a short cDNA sequence found in the course of a large-scale sequencing project of human brain-derived cDNAs, showed significant similarity to the eukaryotic glutamate transporter GLAST1 and was therefore used as probe in the search for further glutamate transporter cDNAs. Fragments of the cDNA were used for the isolation and characterization of human ASCT1 genomic clones. The ORF of 1572 bp encoding 524 amino acid residues is distributed over 8 exons, which span at least 40 kb of human chromosomal DNA. The ASCT1 gene locus was assigned to chromosome 2p13–p15 by chromosomal *in situ* suppression (CISS) studies. The gene structure is not related to any other previously characterized transporter gene. In contrast to the genes of the sodium-coupled nonglutamate neurotransmitter transporters, it shows no obvious correspondence between intron/exon structure and transmembrane organization. The transcription start site in human liver tissue was determined by primer extension analysis to be located 291 bp upstream of the initiating ATG codon. The DNA region immediately upstream of the transcription start lacks any TATA or CAAT boxes but contains several binding sites for the transcription factors *Sp1* and *Egr-1*. The ASCT1 gene (SLC1A4) structure reported here will facilitate the characterization of the genes of the other members of the GLAST-related transporter family and might be useful in the elucidation of amino acid transport-related defects.

INTRODUCTION

Three related sodium-coupled glutamate/aspartate transporters, GLAST1 (Storck *et al.*, 1992), GLT1 (Pines *et al.*, 1992), and EAAC1 (Kanai and Hediger, 1992), have been discovered recently. They share no homology to other previously characterized neurotransmitter transporters and therefore form a new family of ion-coupled transport proteins. The report of a short cDNA fragment derived from a human brain cDNA library (Adams *et al.*, 1992), exhibiting sequence similarity to GLAST1, suggested the existence of further mammalian members of this protein family. We screened a human cDNA library for clones corresponding to the expressed sequence tag *EST02446* and isolated a cDNA encoding a protein that proved to be identical to the recently described neutral amino acid transporter ASCT1 (Arriza *et al.*, 1993; Shafqat *et al.*, 1993). This is the fourth eukaryotic member of the family of GLAST-related transport proteins.

The known members of this protein superfamily vary considerably in the specificity of transported solutes as well as in their ion requirements. The eukaryotic transporters GLAST1, GLT1, and EAAC1 and the bacterial gltT (Tolner *et al.*, 1992a; Wallace *et al.*, 1990) and gltP (Tolner *et al.*, 1992b) transport glutamate and aspartate while the bacterial transporter dctA (Engelke *et al.*, 1989) can also transport other dicarboxylic acids like malate. The human ASCT1 does not work with glutamate or aspartate but transports alanine, serine, cysteine, and threonine instead. All eukaryotic GLAST-related transporters studied so far use a cotransport of Na^+ ions together with a countertransport of K^+ ions and probably of pH-changing anions to drive the transport process against the gradient of the transported compound (Bouvier *et al.*, 1992; Kanner and Schuldiner, 1987; Klöckner *et al.*, 1993). Their prokaryotic counterparts, however, utilize the pH gradient over the bacterial plasma membrane by coupling the solute transport to a cotransport of protons. The heterogeneity of transport properties is accompanied by a substantial sequence variability of the proteins. The identity of amino acid sequences ranges between 25 and 40% over the whole family. However, there are several regions that exhibit a higher degree of similarity. Most notably, the motif AAxFIAQ, x being a hydrophobic residue, is completely conserved in all members known so far. Moreover, sequence analysis suggests a common transmembrane topology for the whole transporter

The nucleotide sequences reported in this paper have been deposited with the GenBank data repository under Accession Nos. U05229 to U05235.

[1] To whom correspondence and reprint requests should be addressed. Telephone: +49 (221) 478-6980. Fax: +49 (221) 478-6979.

0888-7543/94 $6.00

family although the topology of the C-terminal part following the six putative α-helical transmembranal segments is contradictory in the proposed models (Jording and Pühler, 1993; Kanai and Hediger, 1992; Pines *et al.*, 1992; Storck *et al.*, 1992). For the superfamily of Na^+/Cl^- ion-coupled neurotransmitter transporters with its common 12-transmembrane helix topology, several gene structures have been described (Liu *et al.*, 1992), showing a correspondence between exon/intron structure and transmembrane topology. In contrast, no such data are available for the family of GLAST-related transporters. We report the structure, organization, and chromosomal localization of the human gene for ASCT1, a neutral amino acid transporter and member of the family of GLAST-related transporters.

MATERIALS AND METHODS

Vectors, chemicals, and enzymes. pBluescript II SK(+) was obtained from Stratagene (La Jolla, CA), and the human hippocampus cDNA library in λgt10 and the human placenta genomic library in λEMBL3-SP6/T7 were obtained from Clontech (Palo Alto, CA). Human DNA samples were prepared from blood of white Caucasian healthy volunteers, and human liver mRNA was prepared from biopsy tissue samples obtained from the University Hospital, Cologne, Germany. Oligonucleotides were prepared using an Applied Biosystems 380A oligonucleotide synthesizer. [γ-^{32}P]ATP, [α-^{32}P]dATP, and [^{35}S]dATP were purchased from Amersham–Buchler (Braunschweig, Germany), *Taq* DNA polymerase was from Gibco–BRL (Bethesda, MD), and restriction endonucleases and other enzymes were from Boehringer Mannheim (Germany) unless stated otherwise. DNA sequencing was performed using the T7 sequencing kit from Pharmacia (Uppsala, Sweden). Digoxigenin-11–dUTP and anti-digoxigenin conjugated to rhodamine were purchased from Boehringer (Mannheim, Germany). Biotin-11–dUTP was purchased from Sigma, and avidin conjugated to FITC was obtained from Vector Laboratories.

cDNA cloning. Oligonucleotide primers C1s, TCAGATTCCTCCTGGGCCTCCTCGCCCCAT, and C1as, TCCAGGATAATGGCAATGGTGAGGACCCC, both derived from the sequence of *EST02446* (Adams *et al.*, 1992), were used for PCR amplification of a 320-bp cDNA fragment from a human hippocampus cDNA library. Screening of the above cDNA library with the randomly labeled PCR fragment (PCR320) yielded the clone λc10 with an insert size of 3 kb. The cDNA insert was released by *Eco*RI digestion, subcloned into pBluescript II SK(+), and mapped by double digestion with the restriction enzymes *Eco*RI, *Bam*HI, *Sac*I, *Xba*I, and *Xho*I. The coding portion of the insert was sequenced using the dideoxy chain termination method (Sanger *et al.*, 1977).

Genomic cloning. Four independent and partially overlapping genomic clones were obtained by screening a human placental genomic DNA library with several ASCT1 cDNA fragments as hybridization probes. Application of the 320-bp PCR fragment (PCR320), previously used for cDNA screening, yielded the genomic clones λg2 and λg12. A further hybridization probe was obtained by PCR using λc10 as template and the oligonucleotides C2s, GCTGCTCACCGTGTCCGGGGTGCTGGCGGGC, and C2as, GCAAACAGGACCAATCCTAAAATGTTCATCC, as primers. The resulting 555-bp fragment (PCR555) contains the 5′ coding region of the cDNA clone λc10. Application of this PCR procedure was necessary since the original λc10 contains the artifact of foreign cDNA fused to the 5′ end of the authentic ASCT1 coding region. Screening of the genomic library with the fragment PCR555 yielded the genomic clones λg6 and λg10. For further characterization, the four independent genomic clones were mapped using the enzymes *Eco*RI, *Bam*HI, *Sac*I, *Xba*I, and *Xho*I. Genomic fragments containing coding sequences were identified by Southern blot analysis with cDNA fragments and cDNA-derived oligonucleotides as hybridization probes. Positive genomic fragments were isolated, if necessary submapped with *Pst*I, *Hin*dIII, *Kpn*I, and *Pvu*II, and subsequently sequenced. Mapping, hybridization analysis, and cloning were carried out according to standard procedures (Sambrook *et al.*, 1989).

Genomic Southern blot. For estimating the size of the gap between genomic clones λg6 and λg12, human leukocyte-derived genomic DNA samples of three unrelated individuals were digested using *Eco*RI, *Bam*HI, and *Sac*I, respectively, and separated on a 0.7% agarose gel. Fragments were transferred to a nylon membrane by capillary blotting and hybridized with either a 400-bp *Kpn*I fragment of λg6 (containing exon 3 and flanking sequences) or a 200-bp *Rsa*I fragment of the cDNA clone λc10 (containing exons 3 and 4).

Genomic PCR amplifications. The Southern blot suggested a short gap length between the genomic clones λg6 and λg12. This result was confirmed by PCR analysis of genomic DNA samples of three unrelated individuals using the primers G6s, AGCAGCTCTGGAAATGTAACCCATGAAAAG, and G12as, TCTAGCATTTACTCATCACTCTTGGTTTCT. The identity of the resulting 352-bp product (PCR350) was confirmed by blunt-end cloning into pBluescript and subsequent DNA sequencing. To study the authenticity of the in-frame stop codon within exon 5, DNA samples of three unrelated individuals were studied by PCR using the primers G41s, GGCATCATGTTCCTTGTTGGAAGCAAGAT, and G42as, GTCGCAAATGGGGCGAGGAGGCCCAGGAG. The resulting 200-bp fragment was blunt-end cloned into pBluescript and analyzed by DNA sequencing.

Primer extension analysis. Determination of the transcription start site was carried out using the antisense oligonucleotide PE1as, CGGGGGTCTCCGGGCCTGGCGCAGGAGCAGGTG, which is located 138 bp upstream of the initiating ATG codon and poly(A)$^+$ RNA prepared from human liver RNA. The 5′-labeled oligonucleotide was hybridized to 6 μg of RNA. The primer extension reaction was performed following the standard procedure (McKnight and Kingsbury, 1982).

Chromosomal mapping. Chromosome preparations were obtained from phytohemagglutinin-stimulated human lymphocytes cultured for 72 h. After denaturation at 70°C for 2 min in 70% deionized formamide/2× SSC, the slides were dehydrated in ice-cold ethanol. Hybridization with labeled DNA fragments at 37°C and posthybridization washes were performed as described previously (Lichter *et al.*, 1988).

Genomic clone λg12 was labeled by nick-translation using digoxigenin-11–dUTP (Kessler *et al.*, 1990) or biotin-11–dUTP (Langer *et al.*, 1981). This probe was used for chromosomal *in situ* suppression (CISS) hybridization to elongated metaphase chromosomes as described previously (Lichter *et al.*, 1988, 1990). Photographs of an electronic superposition of FITC and DAPI images were taken directly from the video screen.

Computational sequence analysis. Database searching and protein comparisons were performed using the software package of the Wisconsin Genetics Computer Group (Devereux *et al.*, 1984). For hydrophobicity analysis, the program PROFILEGRAPH (Hofmann and Stoffel, 1992) was used. Restriction maps were verified and refined by the program COMAP (K. Hofmann, unpublished).

RESULTS AND DISCUSSION

EST02446 Encodes a Neutral Amino Acid Transporter

The recent discovery of the three related Na+/K+-coupled glutamate transporters GLAST1, GLT1, and EAAC1 (Kanai and Hediger, 1992; Pines *et al.*, 1992; Storck *et al.*, 1992) raised the question of the molecular diversity of mammalian glutamate transporters. Since not all pharmacologically described high-affinity glutamate transport systems (Ferkaney and Coyle, 1986; Fletcher and Johnston, 1991; Robinson *et al.*, 1991) have their counterparts in the three cloned transporter species, there is a considerable probability that more

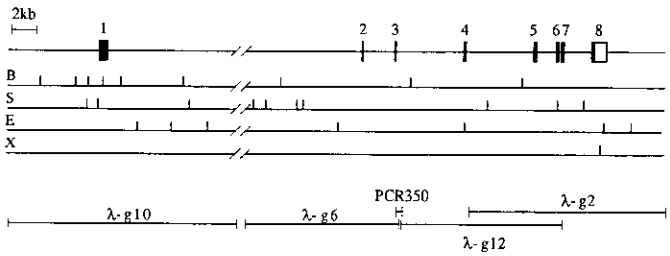

FIG. 1. Organization of SLC1A4. The upper trace shows the arrangement of the eight exons (depicted by boxes and numbers). Filled boxes represent protein coding regions, 5′- and 3′-untranslated regions are represented by open boxes. The result of restriction mapping using four enzymes is shown in the four traces labeled B (*Bam*HI), S (*Sac*I), E (*Eco*RI), and X (*Xho*I), respectively. In the bottom trace, the positions of the four genomic phage clones λg10, λg6, λg12, and λg2 are denoted. The label PCR350 indicates the position of the PCR fragment used to determine the gap between λg6 and λg12.

glutamate transporting proteins remain to be found. The expressed sequence tag *EST02446,* a short cDNA sequence found in the course of a large-scale sequencing project of human brain-derived cDNAs (Adams *et al.,* 1992), showed significant similarity to all three eukaryotic glutamate transporters. For that reason, *EST02446* seemed to be a good candidate for being part of a further glutamate transporter.

Screening of a human hippocampus cDNA library with an *EST02446*-derived DNA probe yielded the cDNA clone λc10, encoding 486 amino acids of a protein with 40–44% sequence identity to the three eukaryotic glutamate transporters. Very recent reports of the cloning of the ASC-type neutral amino acid transporter ASCT1 (Arriza *et al.,* 1993; Shafqat *et al.,* 1993) revealed that the protein encoded by our λc10 clone was identical to ASCT1 transporter.

Cloning of the Human ASCT1 Gene (SLC1A4)

Fragments of the ASCT1-encoding cDNA clone λc10 were used to screen a human placental genomic DNA library. In two rounds of screening, the four partially overlapping phage clones λg10, λg6, λg12, and λg2 were obtained (Fig. 1). These four clones span 52 kb of the human genome and harbor the complete coding region of the ASCT1 gene. The length of the gene is at least 40 kb; it is organized in eight exons interrupted by seven introns (Table 1). The size of the exons varies considerably. While exon 1 contains 527 coding bp and additionally 291 bp 5′-untranslated region, exon 2 consists of only 43 bp, coding for 14 amino acids. Intron sizes range between 276 bp for intron 6 and >19.5 kb for intron 1. Since the size of the gap between λg10 and λg6 could not be determined by Southern blot analysis and no contig-forming phage clone was obtained for that region, the exact size of intron 1 is unknown.

The size of the gap between λg6 and λg12 was estimated by genomic Southern blot analysis. DNA probes made from exons 3 and 4, respectively, hybridized to the same 14-kb *Sac*I fragment and 10-kb *Eco*RI fragment, suggesting a very small distance between the two genomic clones (data not shown). This result was confirmed by PCR amplification using one primer from λg6 and λg12 each. A 350-bp fragment containing regions from both genomic clones was obtained.

Nucleotide Sequence of SLC1A4

All coding portions of SLC1A4 together with some of their flanking regions were characterized by DNA sequencing. The 5′ untranslated region, exon 1, intron 1, and exon 2 are presented in Fig. 2. The nucleotide sequence of the eight exons encoding the human ASCT1 protein was identical to the λc10 cDNA sequence (confirmed by the original ASCT1 sequence (Arriza *et al.,* 1993)), apart from one major difference. At position 164 counted from the start of exon 5, thymine has replaced the cytosine present in the cDNA sequence. This mutation would change the codon for Arg^{323} into a stop codon, resulting in a grossly truncated protein. Interestingly, this mutation was found on both independent genomic clones λg2 and λg12. To check the authenticity of this inframe stop codon and to exclude the operation of processes such as RNA editing, we amplified the region under consideration by PCR using genomic DNA of three unrelated individuals as template and primers located within exon 5. All three DNA samples yielded 200-bp PCR fragments with sequences identical to the cDNA, containing the

TABLE 1

Splice Junction Sequences, Exon Sizes, and Estimated Intron Sizes of the Human ASCT1 Gene (SLC1A4)

Exon		Exon/intron junction		Intron size (kb)
No.	Position*	Donor sequence	Acceptor sequence	
1	1–527	GCCAG gtaaca	caacag AAACC	>19.5
2	528–570	GTACG gtaagg	tcacag TATGC	~2.5
3	571–634	AAAAG gtaaag	ctgcag ATCCC	~5.4
4	635–802	ATGTG gtgagt	ccctag GTACG	~5.9
5	803–1037	TCCAG gtgagt	ttctag CTCAG	~2.0
6	1038–1233	ATTCT gtaagt	ccttag AGTGA	276 bp
7	1234–1419	ATTGT gtaagt	caccag GGACC	~2.5
8	1420–1652[a]			

[a] Exon positions in coding sequence.

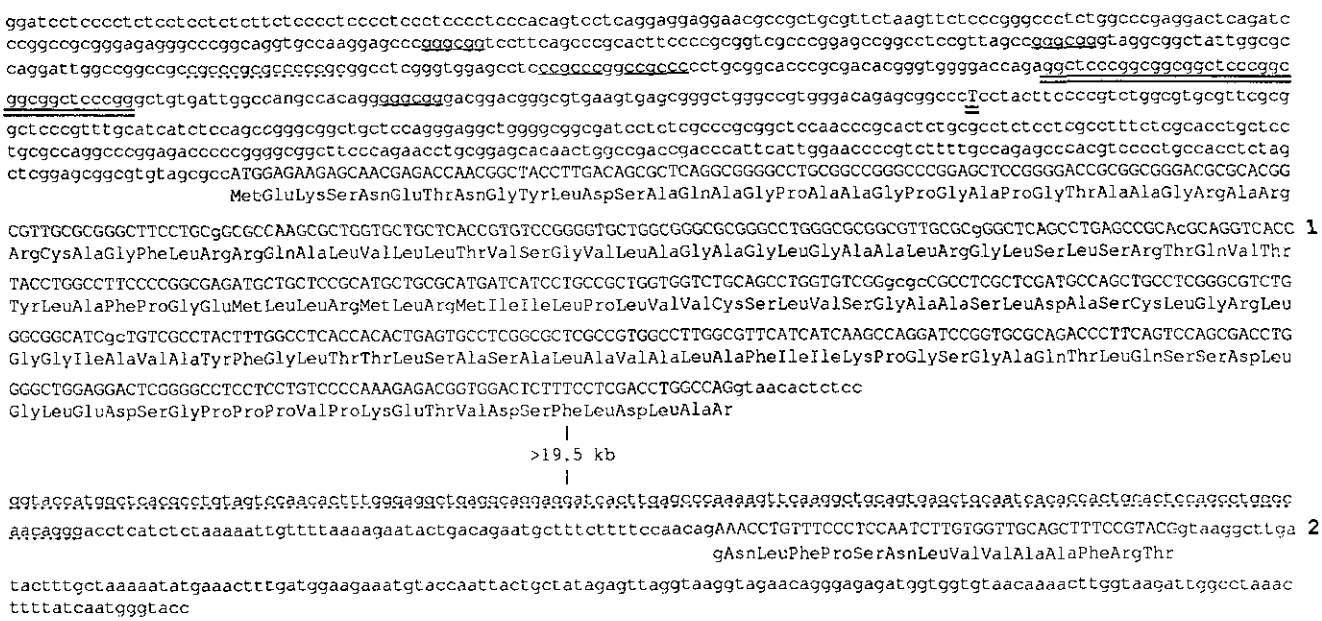

FIG. 2. Nucleotide and deduced amino acid sequence of SLC1A4. The complete nucleotide sequences of the eight exons and large parts of the intron sequences flanking the exons were determined. Those of the 5′ untranslated region, exon 1, intron 1, and exon 2 are presented. The other sequence data are deposited in the GenBank repository, Accession Nos. U05229 to U05235. The distances between the exons were calculated from the restriction map. The major site of transcription initiation, located 291 bp upstream of the start codon ATG, is denoted by a double-underlined capital T. Five consensus *Sp1* recognition sequences at positions −55, −152, −160, −234, and −291 relative to the transcription start are underlined. Two overlapping recognition sequences for the transcription factor *Egr-1* in the region −193 to −179 are denoted by a dotted line. A direct repeat consisting of two complete copies and one truncated copy of the sequence GGCTCCCGGCGGC in the region −113 to −78 is double underlined. Two copies of the dispersed *Alu*-repeat contained in introns 2 and 5 are emphasized by dotted lines. The C→T mutation observed in exon 5 is denoted by the letter T above the sequence.

unaltered codon for Arg^{323}. Therefore, we conclude that in the DNA used for the preparation of the genomic library, at least one allele of SLC1A4 codes for a mutated, probably dysfunctional protein.

Characterization of the 5′ End of SLC1A4

The site of transcription initiation was determined by primer extension analysis using human liver poly(A)$^+$ RNA. One major transcription start was determined 291 bp upstream of the ATG start codon (Fig. 3). The nucleotide sequence immediately upstream of the transcription start site is extremely GC-rich (80% of the proximal 250 bp) and lacks TATA and CAAT boxes in proper spacing to the transcription start. There are, however, several consensus sequences for *Sp1* binding at positions −55, −152, −160, −234, and −291 as well as consensus sequences for other transcription factors. At positions −193 and −187, there are two overlapping recognition sites for members of the *Egr*-family of transcription factors comprising *Egr-1* (*krox24*), *Egr-2* (*krox20*), *Egr-3,* and *Egr-4* (*NGFI-C*). These early growth response genes involved in cell differentiation have been reported to be involved in gene expression in brain and muscle (Li *et al.,* 1993). The observed relative lack of well-defined *cis*-elements, common to genes with low tissue specificity, agrees with the expression of ASCT1 mainly in brain, muscle, and pancreas, but it is also detectable in heart, liver, and kidney (Arriza *et al.,* 1993; Shafqat *et al.,* 1993).

As an interesting feature a direct repeat was found in the proximal promoter region of the SLC1A4 at position −78 to −113. It consists of two copies of the sequence GGCTCCCGGCGGC, followed by a third truncated copy. Single copies of this GC-rich sequence, especially of the core sequence GGCGGCGGCTCC, occur in upstream regions of several genes, notably in most early

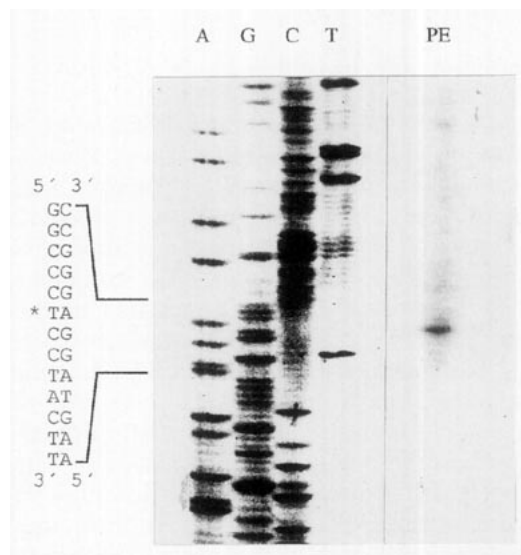

FIG. 3. Primer extension mapping of the transcription start site. The transcription start site was analyzed by primer extension using the oligonucleotide PE1as and 6 μg human liver poly(A)$^+$ RNA (lane PE). The exact position was determined by aligning the corresponding sequence ladder (lanes **A, G, C,** and **T**). The main site of transcription initiation is marked by an asterisk.

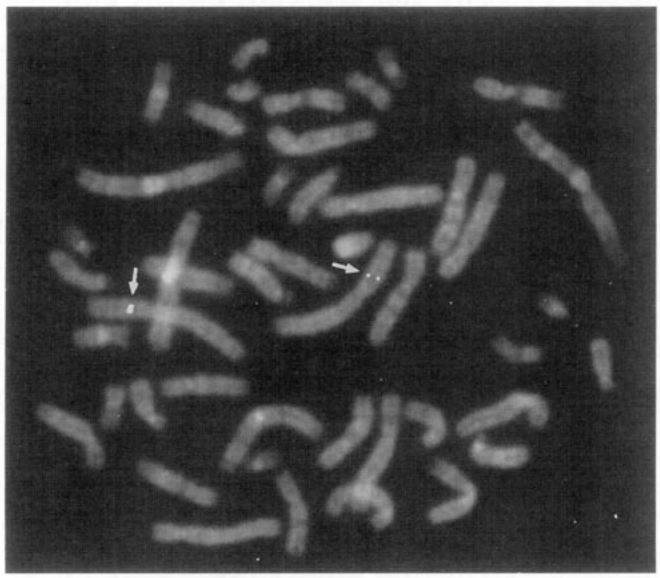

FIG. 4. Localization of ASCT1 to human chromosome 2p13–p15. The digoxigenin-labeled insert of the genomic clone λg12 was hybridized to human metaphase chromosomes and detected by anti-digoxigenin antibodies conjugated to rhodamine. Chromosomes were counterstained with DAPI. The hybridization signal was assigned to chromosome 2p13–p15 (arrowheads).

growth response genes (*Egr* family, *junD*) where they are present in reverse orientation and are highly conserved between species. A similar repeat structure is found only in the 5′ untranslated region of the transcription factor *Egr-1* (Suggs *et al.,* 1990).

The significance of this observation remains to be investigated. The upstream region of SLC1A4 contains an unusually large open reading frame of 960 bp in reverse orientation. The corresponding polypeptide sequence, which has a high content of Gly and Arg, shows no detectable similarity to any characterized protein.

Chromosomal Localization of the SLC1A4

To localize the gene for human ASCT1, we applied the technique of chromosomal *in situ* suppression hybridization. The genomic clone λg12, spanning exons 4 to 7, was hybridized to elongated metaphase chromosomes. Detection of the biotinylated or digoxigenin-labeled probes with avidin or antibodies coupled to fluorescent dyes resulted in specific signals in the short arm of chromosome 2 (Fig. 4). For quantitative evaluation 17 metaphase spreads were analyzed. In all cases specific hybridization signals were observed on both homologs of chromosome 2 (see Fig. 4). Comparison of the hybridization signal to chromosomal bands obtained by DAPI staining allowed us to narrow the region of the human ASCT1 locus to chromosome 2p13–p15. This region, which also contains the gene for the early growth response gene *Egr-4* (*NGFI-C*) (Crosby *et al.,* 1992; Holst *et al.,* 1993), has been shown to contain a constitutive fragile site (*FRA2E*) (Yunis and Soreng, 1984). Knowledge of the gene organization and the chromosomal localization of human ASCT1 will be useful in the elucidation of hereditary defects connected to altered cellular uptake of neutral amino acids.

Transmembrane Domain Structure of ASCT1

When the first three mammalian glutamate/aspartate transporters were cloned, it was not surprising that they did not belong to the superfamily of Na^+/Cl^--coupled transporters. Several attempts to find new members of this protein family by homology screening had led to the characterization of more than 15 different transport proteins, none of them able to transport glutamate (Amara and Kuhar, 1993; Liu *et al.,* 1992).

All Na^+/Cl^--coupled transporters possess a clearly defined arrangement of 12 hydrophobic domains probably spanning the membrane. For the GLAST-related family of transporters, a clear decision on the transmembrane topology is not possible from sequence considerations alone. In ASCT1, as well as in GLAST1, GLT1, EAAC1, and the prokaryotic H^+/glutamate symporters, there are six hydrophobic regions in the N-terminal half of the molecule, all of them large enough to span a membrane bilayer in α-helical manner. In ASCT1 and the other mammalian family members, there is a seventh extended hydrophobic region that is absent in the prokaryotic transporters, located shortly after the proposed transmembrane helix VI. The most intriguing feature with regard to the prediction of transmembrane topology is a relatively long region of intermediate hydrophobicity within the C-terminal half of all superfamily members. A hydrophobicity analysis of this region with a window width of 9 reveals the presence of six short hydrophobic stretches in almost equal spacing. This feature, which is strictly conserved in both mammalian and prokaryotic GLAST-related transporters, had prompted us to speculate about alternatives to a classical α-helical transmembrane organization (Storck *et al.,* 1992).

In a recent report on the topology of the prokaryotic glutamate transporter dctA from *Rhizobium meliloti,* it was demonstrated that a series of dctA/lacZ and dctA/phoA fusion proteins with their fusion sites located at different points between the six short hydrophobic stretches exhibited an alternating outside–inside pattern of the reporter enzyme activity (Jording and Pühler, 1993). Under the assumption of a common transmembrane topology for the whole protein family, this finding suggests that the short hydrophobic stretches are indeed membrane spanning. Whether they form the core of relatively short transmembrane α-helices or span the membrane in some different way remains subject to speculation. Figure 5 gives a model of ASCT1 transmembrane organization that is in accordance with the existing experimental data. The seventh transmembrane region, which is detectable in only the mammalian members of the family, is drawn as a helix, while five of the six shorter hydrophobic stretches are tentatively drawn as a β-sheet structure.

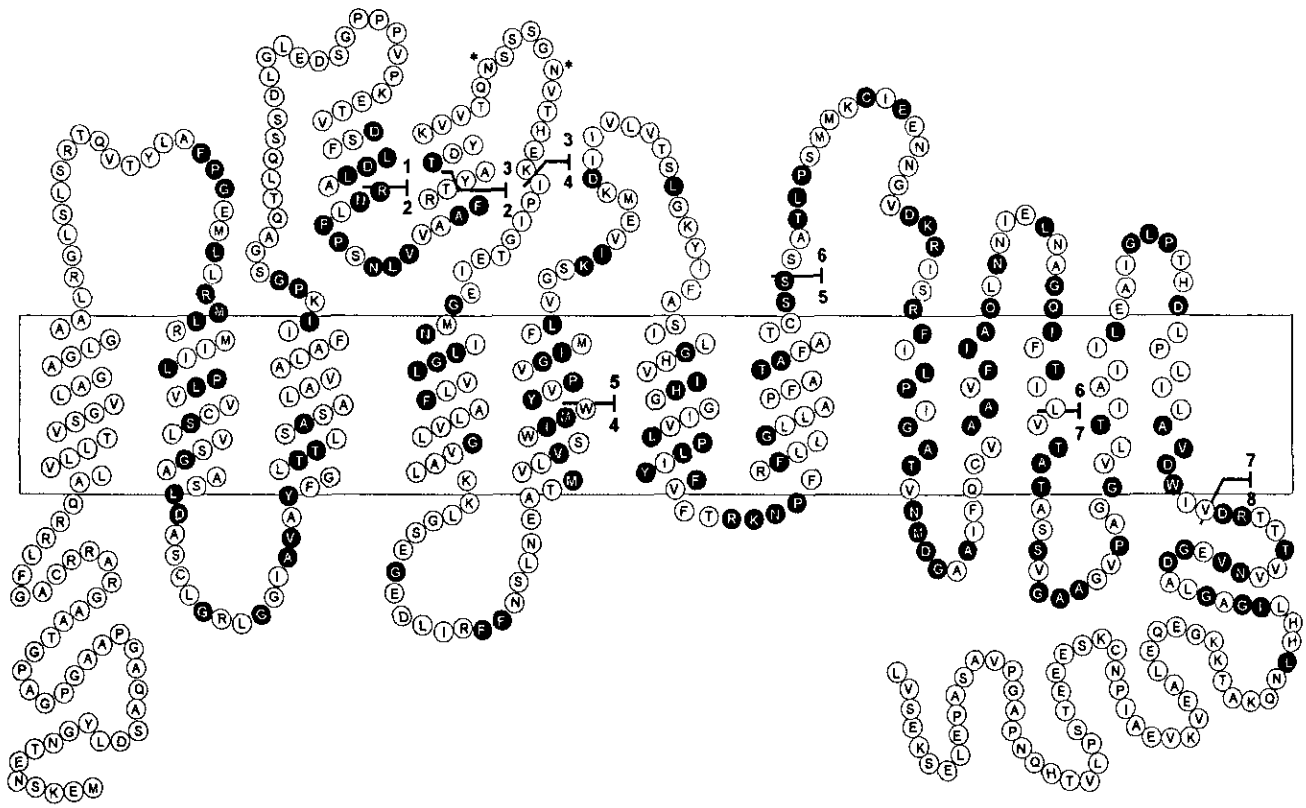

FIG. 5. A model for ASCT1 transmembrane topology. The proposed model for ASCT1 transmembrane topology is defined by seven transmembrane helical segments and five shorter hydrophobic segments, either forming the core of further bilayer-spanning helices or spanning the membrane in a different conformation. Both N- and C-termini are depicted on the cytoplasmic side of the plasma membrane represented by the shaded horizontal bar. The positions of the intron insertions are indicated by black lines with associated exon numbers. Residues conserved throughout all eukaryotic members of the family of GLAST-related transporters are given in white letters on dark background. The two putative glycosylation sites are labeled by asterisks.

Position of the Intron Sites in Relation to Transmembrane Organization

In many transmembrane proteins there is a striking correlation between their transmembrane arrangement and the position of the intron insertions into the sequence. Usually single transmembrane segments are encoded by individual exons while the intron positions are located within the connecting loops, as suggested by the exon-shuffling hypothesis (Gilbert *et al.,* 1986) under the assumption that transmembrane helices are autonomous folding domains (Popot and de Vitry, 1990). The gene family of the Na^+/Cl^--coupled neurotransmitter transporters, which also includes members transporting nonneurotransmitters, is an additional example of the common relation between transmembrane organization and intron site position. The mouse gene for the GABA-transporter, which is relatively compact (12 kb), contains 12 introns, none of which dissects a putative transmembrane domain (Liu *et al.,* 1992). With only one exception, each of the exons contains only a single transmembrane domain.

However, the situation in SLC1A4 is quite different (Fig. 5). Introns 1, 2, and 3 all dissect the nucleotide sequence of the extracellular loop between transmembrane helices III and IV; intron 4 is inserted in the middle of the sequence coding for transmembrane helix V. Introns 5 and 7 are at "conventional" positions within the hydrophilic loops following helix VII and the fifth hydrophobic stretch, while intron 6 dissects the third short hydrophobic region. No correlation between the sites of the insertion of introns interrupting the coding sequence of the transporter and the hydrophobicity profile of its polypeptide chain is apparent in the proposed model of its membrane topology (Fig. 5). Whether the exon/intron structure or the hydrophobicity profile is more reliable in the theoretical prediction of the transmembrane organization of the GLAST-related transporter superfamily can be answered only by topochemical studies.

ACKNOWLEDGMENT

This work has been supported by the Deutsche Forschungsgemeinschaft (Sto32/34-1).

REFERENCES

Adams, M. D., Dubnick, M., Kerlavage, A. R., Moreno, R., Kelley, J. M., Utterback, T. R., Nagle, J. W., Fields, C., and Venter, J. C. (1992). Sequence identification of 2375 human brain genes. *Nature* **355:** 632–634.

Amara, S. G., and Kuhar, M. J. (1993). Neurotransmitter transporters: Recent progress. *Annu. Rev. Neurosci.* **16:** 73–93.

Arriza, J. L., Kavanaugh, M. P., Fairman, W. A., Wu, Y. N., Murdoch, G. H., North, R. A., and Amara, S. G. (1993). Cloning and expression of a human neutral amino acid transporter with structural similarity to the glutamate transporter gene family. *J. Biol. Chem.* **268:** 15329–15332.

Bouvier, M., Szatkowski, M., Amato, A., and Attwell, D. (1992). The glial cell glutamate uptake carrier countertransports pH-changing anions. *Nature* **360:** 471–473.

Crosby, S. D., Veile, R. A., Donis-Keller, H., Baraban, J. M., Bhat, R. V., Simburger, K. S., and Milbrandt, J. (1992). Neural-specific expression, genomic structure, and chromosomal localization of the gene encoding the zinc-finger transcription factor NGFI-C. *Proc. Natl. Acad. Sci. USA* **89:** 4739–4743.

Devereux, J., Haeberli, P., and Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* **12:** 387–394.

Eisenberg, D. (1984). Three-dimensional structure of membrane and surface proteins. *Annu. Rev. Biochem.* **53:** 595–623.

Engelke, T., Jording, D., Kapp, D., and Pühler, A. (1989). Identification and sequence analysis of the *Rhizobium meliloti* dctA gene encoding the C4-dicarboxylate carrier. *J. Bacteriol.* **171:** 5551–5560.

Ferkaney, J., and Coyle, J. T. (1986). Heterogeneity of sodium-dependent excitatory amino acid uptake mechanisms in rat brain. *J. Neurosci. Res.* **16:** 491–503.

Fletcher, E. J., and Johnston, G. A. (1991). Regional heterogeneity of L-glutamate and L-aspartate high-affinity uptake systems in the rat CNS. *J. Neurochem.* **57:** 911–914.

Gilbert, W., Marchionni, M., and McKnight, G. (1986). On the antiquity of introns. *Cell* **46:** 151–154.

Hofmann, K., and Stoffel, W. (1992). PROFILEGRAPH—An interactive graphical tool for protein sequence analysis. *Comput. Applic. Biosci.* **8:** 331–337.

Holst, C., Skerka, C., Lichter, P., Bialonski, A., and Zipfel, P. F. (1993). Genomic organization, chromosomal localization and promoter function of the human zinc-finger gene pAT133. *Hum. Mol. Genet.* **2:** 367–372.

Jording, D., and Pühler, A. (1993). The membrane topology of the *Rhizobium meliloti* C4-dicarboxylate permease (DctA) as derived from protein fusions with *Escherichia coli* K12 alkaline phosphatase (PhoA) and β-galactosidase (LacZ). *Mol. Gen. Genet.* **241:** 106–114.

Kanai, Y., and Hediger, M. A. (1992). Primary structure and functional characterization of a high-affinity glutamate transporter. *Nature* **360:** 467–471.

Kanner, B. I., and Schuldiner, S. (1987). Mechanism of transport and storage of neurotransmitters. *CRC Crit. Rev. Biochem.* **22:** 1–38.

Kessler, C., Höltke, H., Seibl, R., Burg, J., and Mühlegger, K. (1990). Non-isotopic labeling and detection of nucleic acids. I. A novel DNA labeling and detection system based on digoxigenin–antidigoxigenin ELISA principle (digoxigenin system). *Biol. Chem. Hoppe-Seyler* **371:** 917–927.

Klöckner, U., Storck, T., Conradt, M., and Stoffel, W. (1993). Electrogenic L-glutamate uptake in *Xenopus laevis* oocytes expressing a cloned rat brain L-glutamate/L-aspartate transporter (GLAST1). *J. Biol. Chem.* **268:** 14594–14596.

Langer, P., Waldrop, A., and Ward, D. (1981). Enzymatic synthesis of biotin-labeled polynucleotides: Novel nucleic acid affinity probes. *Proc. Natl. Acad. Sci. USA* **78:** 6633–6637.

Li, Y., Camp, S., Rachinsky, T. L., Bongiorno, C., and Taylor, P. (1993). Promoter elements and transcriptional control of the mouse acetylcholine esterase gene. *J. Biol. Chem.* **268:** 3563–3572.

Lichter, P., Cremer, T., Borden, J., Manuelidis, L., and Ward, D. (1988). Delineation of individual human chromosomes in metaphase cells by in situ suppression hybridization using recombinant DNA libraries. *Hum. Genet.* **80:** 224–234.

Lichter, P., Tang, C., Call, K., Hermanson, G., Evans, G., Housman, D., and Ward, D. (1990). High resolution mapping of human chromosome 11 by in situ hybridization with cosmid clones. *Science* **247:** 64–69.

Liu, Q. R., Mandiyan, S., Nelson, H., and Nelson, N. (1992). A family of genes encoding neurotransmitter transporters. *Proc. Natl. Acad. Sci. USA* **89:** 6639–6643.

McKnight, S. L., and Kingsbury, R. (1982). Transcriptional control signals of a eukaryotic protein coding gene. *Science* **217:** 316–324.

Pines, G., Danbolt, N. C., Bjoras, M., Zhang, Y., Bendahan, A., Eide, L., Koepsell, H., Storm-Mathisen, J., Seeberg, E., and Kanner, B. I. (1992). Cloning and expression of a rat brain L-glutamate transporter. *Nature* **360:** 464–467.

Popot, J. L., and de Vitry, C. (1990). On the microassembly of integral membrane proteins. *Annu. Rev. Biophys. Biophys. Chem.* **19:** 369–403.

Robinson, M. B., Hunter-Ensor, M., and Sinor, J. (1991). Pharmacologically distinct sodium-dependent L-[^{3}H]glutamate transport processes in rat brain. *Brain Res.* **544:** 196–202.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). "Molecular Cloning: A Laboratory Manual," Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sanger, F., Nicklen, S., and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74:** 5463–5467.

Shafqat, S., Tamarappoo, B. K., Kilberg, M. S., Puranam, R. S., McNamara, J. O., Guadano-Ferraz, A., and Fremeau, R. T. (1993). Cloning and expression of a novel Na^+-dependent neutral amino acid transporter structurally related to mammalian Na^+/glutamate cotransporters. *J. Biol. Chem.* **268:** 15351–15355.

Storck, T., Schulte, S., Hofmann, K., and Stoffel, W. (1992). Structure, expression and functional analysis of a Na^+-dependent glutamate/aspartate transporter from rat brain. *Proc. Natl. Acad. Sci. USA* **89:** 10955–10959.

Suggs, S. V., Katzowitz, J. L., Tsai-Morris, C., and Sukhatme, V. P. (1990). cDNA sequence of the human early growth response gene Egr-1. *Nucleic Acids Res.* **18:** 4283.

Tolner, B., Poolman, B., and Konings, W. N. (1992a). Characterization and functional expression in *Escherichia coli* of the sodium/proton/glutamate symport proteins of *Bacillus stearothermophilus* and *Bacillus caldotenax*. *Mol. Microbiol.* **6:** 2845–2856.

Tolner, B., Poolman, B., Wallace, B., and Konings, W. N. (1992b). Revised nucleotide sequence of the gltP gene, which encodes the proton–glutamate–aspartate transport protein of *Escherichia coli* E-12. *J. Bacteriol.* **174:** 2391–2393.

Wallace, B., Yang, Y.-J., Hong, J., and Lum, D. (1990). Cloning and sequencing of a gene encoding a glutamate and aspartate carrier of *Escherichia coli* K-12. *J. Bacteriol.* **172:** 3214–3220.

Yunis, J. J., and Soreng, A. L. (1984). Constitutive fragile sites and cancer. *Science* **226:** 1199–1204.