J Mol Evol (1994) 39:93–100

Journal of Molecular Evolution

Genetic Basis for Tissue Isozymes of Glutamine Synthetase in Elasmobranchs

Purnima R. Laud, James W. Campbell

Department of Biochemistry and Cell Biology and Laboratory for Biochemical and Genetic Engineering, Rice University, Houston, TX 77251, USA

Received: 12 October 1993 / Accepted: 8 December 1993

Abstract. Tissue-specific isozymes of glutamine synthetase are present in elasmobranchs. A larger isozyme occurs in tissues in which the enzyme is localized in mitochondria (liver, kidney) whereas a smaller form occurs in tissues in which it is cytosolic (brain, spleen, etc.). The nucleotide sequence of spiny dogfish shark (*Squalus acanthias*) liver glutamine synthetase mRNA, derived from its cDNA, shows there are two in-frame initiation codons (AUG) at the N-terminus which will account for the size differences between the two isozymes. Initiation at the up-stream and down-stream sites would yield peptides of 45,406 and 41,869 mol. wts. representing the precursor of the mitochondrial isozyme and the cytosolic isozyme, respectively. The additional N-terminal 29 amino acids present in the mitochondrial isozyme precursor contains two putative cleavage sites based on the Arg-X-(Phe,Ile,Leu) motif. The predicted two-step processing would remove 14 of the 29 N-terminal amino acids. These 14 amino acids can be predicted to form a very strong amphipathic mitochondrial targeting signal. Their removal would yield a mature peptide of 43,680 mol. wt. The calculated mol. wts. based on the derived amino acid sequence are therefore in good agreement with previous estimates of an approximately 1.5–2-kDa difference between the M_rs of the mitochondrial and cytosolic isozymes. A model for the evolution of the mitochondrial targeting of glutamine synthetase in vertebrates is proposed.

Correspondence to: J.W. Campbell

The nucleotide sequence reported will appear in GenBank under accession number U04617

Key words: Glutamine synthetase evolution — Mitochondrial import — Signal sequence evolution — Cleavage site — Tissue-specific isozymes — Elasmobranchs

Introduction

Based on size differences, glutamine synthetase (EC 6.3.1.2) occurs as tissue-specific isozymes in elasmobranchs and certain other cartilagenous fish (Ritter et al. 1987; Smith et al. 1987). These size differences correspond to the metabolic compartmentation of the isozymes. The larger isozyme is mitochondrial in liver and kidney tissues. In liver, glutamine synthetase is the primary ammonia-fixing/detoxifying enzyme in the mitochondrial matrix. Glutamine formed by its action is utilized by carbamoylphosphate synthetase-III, the first enzyme in the modified urea cycle present in elasmobranch liver which functions to provide urea for osmotic purposes (Casey and Anderson 1982; Anderson 1991; Campbell and Anderson 1991). The larger isozyme functions for glutamine synthesis in kidney mitochondria in a substrate cycle between glutamate and glutamine which may be involved in the regulation of ammoniagenesis (King and Goldstein 1983). In spleen, the smaller cytosolic isozyme is thought to function in pyrimidine biosynthesis by providing glutamine for carbamoylphosphate synthetase-II, the first enzyme in the pyrimidine biosynthetic pathway (Anderson 1989). The smaller isozyme is also present in brain, retina, heart, gill, and rectal gland (Smith et al. 1987) and presumably also functions in a biosynthetic capacity in these tissues.

Most compartmental isozymes are encoded by sep-

arate genes which contain the differential targeting information. Examples of mitochondrial and cytosolic isozymes include aspartate aminotransferase (Juretic et al. 1990), malate dehydrogenase (Joh et al. 1987), and phosphoenolpyruvate carboxykinase (Hod et al. 1982). The cytosolic and mitochondrial isozymes of glutamine synthetase are also coded for by separate genes in the fruit fly *Drosophila melanogaster* (Caizzi et al. 1990). Mammalian glutamine synthetase, which is cytosolic in all tissues examined, is coded for by a single gene (Kuo and Darnell 1989). A single gene also encodes avian glutamine synthetase (Patejunas and Young 1987), which occurs in both mitochondrial and cytosolic compartments in a tissue-specific manner (Campbell and Smith 1992). The compartmental isozymes of glutamine synthetase in the stingray *Dasyatis sabina* also appear to be coded for by a single gene. (See Campbell and Anderson 1991.) A single gene has therefore persisted into the vertebrate line of descent and the similarity in sequence between the *Drosophila* cytosolic isozyme and vertebrate glutamine synthetases suggests they shared a common ancestry (Pesole et al. 1991; Kumada et al. 1993).

Whether encoded by two genes or one, cytosolic and mitochondrial isozymes are usually generated from two translation products, one of which contains the mitochondrial targeting information. When a single gene encodes both compartmental isozymes, the translation products may arise from separate mRNAs generated by differential transcription (e.g., Tropschug et al. 1988) or, in at least one instance, by differential translation of a single mRNA (Suzuki et al. 1992). We show here that the glutamine synthetase gene in the spiny dogfish *Squalus acanthias* encodes two in-frame initiation codons at the N-terminus, suggesting that the cytosolic and mitochondrial isozymes of elasmobranch glutamine synthetases could arise by either of the two mechanisms. We have also identified putative cleavage sites in the N-terminal mitochondrial targeting signal that will account for the M_r differences previously reported for the two isozymes in elasmobranchs.

Materials and Methods

Materials. Sharks (*Squalus acanthias*) were collected from the waters off Mt. Desert Island, ME. Liver tissue was "snap frozen" in liquid nitrogen and placed on solid carbon dioxide for shipping. Long-term storage was at −85° C. A shark liver cDNA library, constructed in λZAP II vector, was obtained from Drs. Wilmar L. Salo and Paul M. Anderson, University of Minnesota, Duluth.

Restriction enzymes, T4 DNA ligase, and DNA polymease I (Klenow fragment) were from New England Biolabs or BRL; AmpliTaq DNA polymerase and PCR reagents, from Perkin Elmer-Cetus; PCR and sequencing primers, from Oligo Etc., Inc.; and [^{35}S]- and [^{32}P]dATP, from Amersham. Sequencing reactions were carried out with USB's Sequenase 2.0 using BioRad's Sequi-Gen sequencing cell.

Methods. A shark-specific probe for screening the shark cDNA library was prepared from shark liver glutamine synthetase (GS) mRNA using the polymerase chain reaction (PCR). For this, total RNA was isolated from liver tissue using guanidine isothiocyanate followed by cesium chloride gradient centrifugation (Chirgwin et al. 1979). Poly (A+) mRNA was isolated from the total RNA with oligo-dT chromatography (Aviv and Leder 1972). This mRNA was used as a template for first-strand cDNA synthesis using Pharmacia's kit. An aliquot of the resulting mRNA-cDNA mixture was then used as the template for the PCR. Degenerate oligomeric primers to the highly conserved regions of avian and mammalian GSs were used for the PCR. The "sense" primer started at Trp_{61} of the shark sequence (Trp_{32} of the avian and mammalian enzymes) and encoded the sequence Trp-Ile-Asp-Gly-Thr-Gly. It contained a *Bam*HI site and CG "clamp" at the 5′. The sequence was CGG GGATCC TGG ATX GA(CT) GGX ACX GG where X is inosine and parentheses indicate the bases used. The "antisense" primer started at Asp_{368} (Asp_{339} of the avian and mammalian enzymes) and encoded the sequence Lys-Gly-Tyr-Phe-Glu-Asp. It contained an *Eco*RI site at the 5′ with a GG "clamp." The sequence was GG GAATTC (AG)T (CT)TC (AG)TA XCC (CT)TT. These primers flanked a 890-bp region and included the ATP-binding region (Kim and Rhee 1988).

Conditions used for the PCR were modified from Friedman et al. (1988): denaturation at 94° C for 3 min followed by 30 cycles of denaturation at 94° C for 1 min, annealing at 50° C for 30 s, and elongation at 72° C for 30 s. A final elongation was carried out at 72° C for 10 min. The PCR product was reamplified, purified using the GeneClean kit (Bio101), and digested with *Bam*HI and *Eco*RI. This led to the removal of 177 bp from the N-terminus corresponding to amino acids Trp_{61} to Arg_{120} due to an internal *Bam*HI site in the PCR product. The purified product was then ligated to pBluescript KS vector previously digested with *Bam*HI and *Eco*RI. The resulting plasmid reacted with both full-length chicken (Campbell and Smith 1992) and rat (Mill et al. 1991) GS cDNA clones during Southern blotting and was used to transform competent *E. coli* XL-I Blue cells. Transformants were screened for the presence of the PCR product. A positive transformant shown by sequencing to have the PCR insert was used to produce bulk amounts of the probe. This probe was radiolabeled and used to screen the λZAP II shark cDNA library. Approximately 15,000 plaques were screened.

Six positive clones were initially identified using the shark-specific probe. Five of these tested positive through tertiary screening. Partial sequence data were obtained for all clones and one, designated pSGS6, appeared to encode the complete GS peptide. This clone was fully sequenced by the chain-termination method (Sanger et al. 1977). T3 and T7 primers were used initially and, as sequence information became available, specific synthetic primers were used. Both strands were sequenced.

Software. Routine manipulations of sequence data were with the MacVector program sold by IBI and the U. Wisconsin GCG program. The program Wheel, used for helical wheel projections, was provided by Mr. Martin Jones and Dr. Jere P. Segrest, U. Alabama (Segrest et al. 1990). The program PAUP was used for phylogenetic tree construction (Swofford 1993).

Results and Discussion

Derived Amino Acid Sequence of pSGS6

The nucleotide and derived amino acid sequences of shark liver clone pSGS6 are shown in Fig. 1. This clone contains one open reading frame coded for by 1,212 nu-

```
CTGTCAACAAGATTTTCATTCAGTAGCCCTTTCCAGCACAGCCCCTATCAAGGAACTTTACAGTCTATTTTTGGTGGAGG

ATG CGG ATC TGT CGC AGT TTC CTG TTT CTG GTA AAA AAA TGC GGC AAT ATC ACG CCG ACG
 M  R  I  C  R  S  F  L  F  L  V  K  K  C  G  N  I  T  P  T    20

ATT TGG CGG AAT CAA CAT ACT TAC AAG ATG GCC ACG TCA GCC AGC GCC AAT CTC AGC AAA
 I  W  R  N  Q  H  T  Y  K  M  A  T  S  A  S  A  N  L  S  K    40

ATC GTC AAG AAG AAT TAC ATG GAG CTG CCC CAA GAT GGC AAG GTG CAA GCG ATG TAC ATC
 I  V  K  K  N  Y  M  E  L  P  Q  D  G  K  V  Q  A  M  Y  I    60

TGG ATA GAC GGC ACA GGG GAG GCC GTC CGC TGC AAG ACC AGA ACC TTG GAC AAT GAG CCC
 W  I  D  G  T  G  E  A  V  R  C  K  T  R  T  L  D  N  E  P    80

AAG AGC ATT GCC GAA CTC CCA GAA TGG AAC TTC GAT GGC TCA AGT ACG TAT CAG TCA GAG
 K  S  I  A  E  L  P  E  W  N  F  D  G  S  S  T  Y  Q  S  E   100

GGG TCC AAC AGC GAC ATG TAC CTG GTT CCA TCT GCC ATG TTC CGG GAT CCT TTC CGT AGG
 G  S  N  S  D  M  Y  L  V  P  S  A  M  F  R  D  P  F  R  R   120

GAT CCA AAC AAG CTC GTC CTC TGT GAG GTC CTC AAG TAT AAC AGG AAG CCA GCA GAA TCT
 D  P  N  K  L  V  L  C  E  V  L  K  Y  N  R  K  P  A  E  S   140

AAT CTT AGA CAC TCA TGC CAG AAA ATC ATG TCC ATG ATC GCA AAT GAA TAT CCA TGG TTT
 N  L  R  H  S  C  Q  K  I  M  S  M  I  A  N  E  Y  P  W  F   160

GGA ATG GAA CAA GAG TAC ACT TTG CTG GGA ACG GAC GGT CAT CCC TTT GGA TGG CCT TCC
 G  M  E  Q  E  Y  T  L  L  G  T  D  G  H  P  F  G  W  P  S   180

AAT TGC TTT CCT GGA CCA CAA GGG CCC TAT TAC TGT GGA GTT GGT GCA GAC AAA GCT TAC
 N  C  F  P  G  P  Q  G  P  Y  Y  C  G  V  G  A  D  K  A  Y   200

GGC AGA GAT ATT GTC GAG GCT CAC TAC CGG GCG TGT CTG TAT GCT GGA ATT GAA CTC AGT
 G  R  D  I  V  E  A  H  Y  R  A  C  L  Y  A  G  I  E  L  S   220

GGA ACC AAT GCT GAA GTT ATG GCT GCT CAG TGG GAA TAC CAA GTT GGA CCT TGT GAA GGT
 G  T  N  A  E  V  M  A  A  Q  W  E  Y  Q  V  G  P  C  E  G   240

ATC CAG ATG GGT GAC CAC TTG TGG ATT TCC AGG TTT ATT CTG CAC AGG GTG TGC GAG GAC
 I  Q  M  G  D  H  L  W  I  S  R  F  I  L  H  R  V  C  E  D   260

TTC GGT ATC ATT GCT AGC TTT GAC CCT AAG CCC ATT CCT GGC AAC TGG AAT GGT GCT GGG
 F  G  I  I  A  S  F  D  P  K  P  I  P  G  N  W  N  G  A  G   280

TGC CAC ACT AAC TTT AGC ACC AAA GCC ATG CGG GAT GAT GGA GGG TTG AAG TAC ATT GAA
 C  H  T  N  F  S  T  K  A  M  R  D  D  G  G  L  K  Y  I  E   300

GAC TCA ATT GAA AAA CTG GGC AAG AGG CAT CAG TAC CAC ATT CGT GCC TAT GAT CCT AAA
 D  S  I  E  K  L  G  K  R  H  Q  Y  H  I  R  A  Y  D  P  K   320

GGA GGG TTG GAC AAT GCT AGA GCT TTG ACA GGC CAC CAT GAA ACC TCA AAT ATC AAT GAG
 G  G  L  D  N  A  R  R  L  T  G  H  H  E  T  S  N  I  N  E   340

TTC TCA GCT GGT GTT GCC AAT AGA GGA GCC AGC ATC CGA ATC CCT CGA TCC GTT GGC CAG
 F  S  A  G  V  A  N  R  G  A  S  I  R  I  P  R  S  V  G  Q   360

GAC AAG AAA GGC TAC TTT GAA GAC CGC CGT CCA TCT GCT AAT TGT GAC CCT TAT GCA GTC
 D  K  K  G  Y  F  E  D  R  R  P  S  A  N  C  D  P  Y  A  V   380

ACA GAA GCA TTG GTC CGC ACA TGC CTA TTG GAT GAG TCT GGG GAC AAG CCT ATT GAG TAC
 T  E  A  L  V  R  T  C  L  L  D  E  S  G  D  K  P  I  E  Y   400

AAC AAA AAT TAA GCAAAATAATGCACTAATGGACCTGGCATTTGTAAGCAGTGATAGCTGTTGAAATGTGGGACC
 N  K  N  *

TTTGGGTCTCTACTCTACTCCTATACTGTACAGGTGCTAAAGGGGGAGGGTCAGAAGGGTTTTATTTGTATTTCAGAAC
CTAATTTCTTCTGTTGTTATCTGGAAGGTGAGGAATGAGGCTTGCGATAGGACAACAAAACTGTTCTCTTATTTAGAAC
AGTTAATAACTCTTCAAGTTGACTGGTCAACTTCACATGGACACTAACCTGTAATTTAATGCAGTCAACATTGATATTA
AAGCAGGGCTTTTGTTTTTTTATTAAATACAGCATTAATATTATGGTCAAGTGACATCTCATCTTACTAAAAAGCTTCT
TGGATCTGTACTGATGCACAAATTACTAATACTGTAAACTGCACAATTTTTTCGTAACTTTTAACTAGTTTACATTTTC
AGATCAATTTTTTTGCAGTGCTGGGAGAAATGTTCTGTAAATGGCTGGTCTAGCACCAGGGAAAAAAGCTTGTGACTTC
TTGTTGGGACTGCATACTAGTCCCATACATTAAATTGCTTCAGAACACTGCATGTGTTAACCCACTTAAATGGGAACTA
GGATACGTAAGATCACAAAATTGGCATACTGCATTCCTGTACAATAATTTTAACTTTGTTATAGTCTTGACCCACCTAA
TGTTAGTGCTCCTGTCCAATTGCTATAGGGTCCAAGGGCACTGGATTCTTCCAGCTTTCTTAAATTATGCCTTTGTTCA
TTGTACACTTACTTTAATGACTCTCATGGGACTTGTCCTAAAAGATCAAAAATGCTCAGCTAATCTCTAGTTTTACAAA
CATTTGTAAAATCTGAGCTGTGCCACTATAACTGGCTGTGCATTCTAATTTTTTGGGGGTCGGGGGGGGGGGGGAAG
AAGTGGGTTGATGAGCAATATGCGACTGGTCCTTCCTGCAGTGAGTTACTGTTCACTTGTCACTTGACCTTTTTTAAAC
TATTTTAAGGCATTACAGGGAAAAGTTCAGTTGTAAAAATTTGTTTCCATGTTTTATTATTTCTGTAGTCCCTGCCTAG
GTACACTAACTGGGATTAGCTCCTCCCAGTGCTACATTTTTGTGCAACAGTTTCTGAAGCATGCTGTAAGCATTGATGT
ATTTTACCGTATTTTGATGGTACAGATTTGTTTCGGGGTTCAACTTTGCTGTAGGTTTTTTAATTAAGGAAAAACAAAT
AGAACTTTTTTTAAATTGCTGTCTCACTGTCTTGTATTTTGAATTGTCATTGCTAAAATTGAGTAAAGTAGTTTTTAAT
CAATTGCTTTCGTATGAAACAAATCTGCAGTACACATGCTGGTTGAGTGTTCCAAGCTGTTGGAAGTTTGTGTCTATGT
GCCCACATGATGCCTCGACGCTATCGAGGCAAACCTATACGTAAAAGGGTCATGTG
```

Fig. 1. Nucleotide sequence of shark liver glutamine synthetase cDNA clone pSGS6 and its corresponding amino acid sequence. *Numbers at right* refer to amino acids.

cleotides plus 80 of the 5′- and 1,462 of the 3′-flanking nucleotides to total 2,754 nucleotides. This is within the size range of 2.8–3.2 kb for most vertebrate GS mRNAs (Smith and Campbell 1988). However, spiny dogfish shark GS mRNA was previously shown by Northern blotting to be an exception at around 4.3 kb (Smith et al. 1987). Based on the latter value, pSGS6 would thus appear to have only 60–65% of the total GS mRNA nucleotides even though it contains the complete coding sequence.

The open reading frame of pSGS6 encodes 403 amino acids and one termination codon. This is 30 amino acids longer than other vertebrate GSs. Twenty-nine of these additional amino acids are at the N-terminus and presumably contain the mitochondrial targeting signal. (See below.) One additional amino acid (Asn_{401}) is at the C-terminus. There are two in-frame initiation codons at the N-terminus which are presumably responsible for the differential targeting information in the mRNAs of the shark GS mitochondrial and cytosolic isozymes. The upstream codon has a single Kozak consensus nucleotide at the initiation site (A at −3) whereas the downstream start codon has two (A at −3 and C at −4) (Kozak 1984). Initiation at the former would give rise to mitochondrial preGS (preGS is the designation used for GS mRNA translation product) and, at the latter, to cytosolic GS.

Evolution of Animal GSs

The GS gene is considered to be one of the oldest functioning genes (Kumada et al. 1993) and the primary structure of GS encoded by it has been highly conserved throughout eukaryote evolution. This is especially true for the vertebrate enzymes. For example, avian GS shows 87–90% sequence identity with the mammalian enzymes even though their respective lines of descent diverged some 300 myr ago (Benton 1990). This marked conservation of vertebrate GSs is further emphasized by the shark enzyme. There is an 82% sequence identity between it and the mammalian enzyme and 83% between it and the chicken enzyme. The sequence similarity (includes conservative replacements) between the shark and the mammalian and avian enzymes is around 91%. The identity between the *Drosophila* cytosolic isozyme which may have shared a common ancestor with the vertebrate peptide is 60% (similarity of 74%) despite a divergence of over 500 myr of their ancestral lines (Pesole et al. 1991; Benton 1990). A phylogenetic tree for the known animal GSs is shown in Fig. 2. As can be seen in this figure, the two vertebrate mitochondrial GSs (shark and chicken) and the cytosolic mammalian ancestral GS each arose independently from an ancestral GS sharing a common ancestor with the fly cytosolic isozyme.

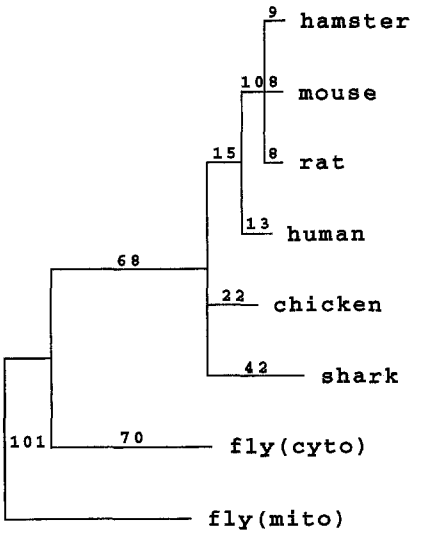

Fig. 2. Phylogenetic tree for animal glutamine synthetases based on 358 amino acid residues common to all peptides. "Bootstrap" analysis of 200 replicates was performed using the PAUP program (Swofford 1993). The *Drosophila* mitochondrial peptide was used as the outgroup. *Numbers* indicate branch lengths. Amino acid sequences used are referenced in Campbell and Smith (1992).

Cleavage Site Identification in Mitochondrial preGS

Shark liver mitochondrial preGS is larger than the mature form which is in turn larger than the cytosolic isozyme (Smith et al. 1987). Cleavage sites for the peptidase system that removes the mitochondrial targeting signal during import must therefore reside within the additional 29 N-terminal amino acids. An examination of several mitochondrial targeting-signal cleavage sites has revealed certain conserved motifs (Hendrix et al. 1989; von Heijne et al. 1989; Gavel and von Heijne 1990). One of these, a Arg-X-(Phe,Ile,Leu) motif in which Arg is at the −10 position relative to the final cleavage site is especially common in prepeptides that undergo two-step cleavage (Hendrix et al. 1989). The first cleavage between the −9 and −8 sites results in an octapeptidyl sequence essential for recognition by the second processing peptidase which then cleaves between the −1 and +1 site to yield the mature peptide (Isaya et al. 1992). As can be seen in Fig. 3, an Arg-X-(Phe, Ile,Leu) is present at the N-terminus of shark liver preGS. With this motif, the first cleavage would occur between −9 Ser and −8 Phe. A second cleavage would then be predicted between −1 Cys and +1 Gly. Preliminary attempts to obtain direct sequence data at the N-terminus indicated the mature mitochondrial peptide is blocked. While blocking of N-terminal Gly residues is rare, it does occur (Driessen et al. 1985). The shark motif differs from the consensus motif in that a hydrophobic amino acid (Leu) is present at the −5 position. Ser is most commonly found at this position (Isaya et al. 1992) although Gly and Thr may also be

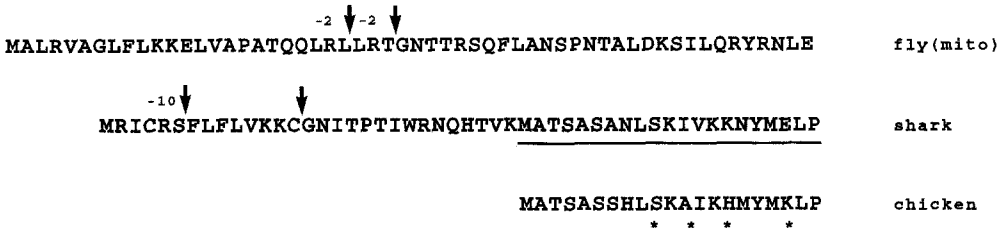

Fig. 3. N-terminal amino acid sequences of animal mitochondrial glutamine synthetases. *Arrows* indicate potential cleavage sites for the shark and fly peptides which contain a presequence. The amino acids thought to be responsible for the targeting function of the mature chicken peptide are indicated by *. The sequence corresponding to the N-termini of avian and mammalian glutamine synthetases is *underlined* for the shark enzyme.

present (Hendrix et al. 1989). Arg is common at the −2 position of many cleavage sites. Lys occurs at this position as well as the −3 position of the second putative cleavage site in shark liver preGS, suggesting the resulting positively charged field substitutes for Arg in peptidase recognition.

The open reading frame of pSGS6 encodes a peptide of calculated mol. wt. 45,406, which is taken to represent the mitochondrial precursor peptide. Two-step processing of this precursor based on the Arg-X-(Phe,Ile, Leu) motif would yield a peptide of calculated mol. wt. 43,680 with an estimated pI of 6.29. As indicated above, initiation of transcription/translation at the downstream start site encoded by pSGS6 would yield a peptide of calculated mol. wt. 41,869 and estimated pI of 5.94, which is presumably the cytosolic isozyme. We had previously estimated differences of 3 kDa between the M_rs of mitochondrial preGS and the mature isozyme and of 2 kDa between the mature mitochondrial and cytosolic isozymes using SDS-PAGE (Smith et al. 1987). The more alkaline pI of the shark mitochondrial isozyme is in keeping with what has been observed for most other mitochondrial and cytosolic isozymes (Hartmann et al. 1991).

Drosophila preGS is the only other animal preGS which has a putative N-terminal mitochondrial targeting signal (Caizzi et al. 1990). Again, as in the shark, the mature mitochondrial isozyme is larger than the cytosolic isozyme so the cleavage site(s) must reside within the additional N-terminal amino acids encoded by the *Drosophila* mitochondrial GS gene. An apparent difference of 2 kDa between the M_rs of *Drosophila* mitochondrial preGS and the mature isozyme has been reported (Caizzi et al. 1990). There is only one sequence resembling the Arg-X-(Phe,Ile,Leu) in *Drosophila* preGS. This is the $Arg_{23}Leu_{24}Leu_{25}$ sequence. However, two-step processing based on this motif would yield a peptide similar in size to that of the cytosolic isozyme. One of the most common motifs for one-step processing of mitochondrial targeting signals is Arg at −2 (Hendrix et al. 1989). Cleavage at either of the −2 Arg sites indicated in Fig. 3 for *Drosophila* preGS would yield mature peptides with calculated mol. wts. equivalent to 2.7–3 kDa less than that of the precursor and larger than that of the cytosolic isozyme. This is within reasonable agreement with the 2-kDa difference between the two isozymes estimated by SDS-PAGE by Caizzi et al. (1990).

preGS Mitochondrial Targeting Signals and Their Evolution

Mitochondrial targeting signals typically have high basic and hydroxy amino acid contents, are generally lacking acidic amino acids and proline (Lemire et al. 1989), and are capable of forming amphipathic helices (von Heijne 1986; Roise et al. 1988). This amphipathic helix may initially insert into mitochondrial membranes and interact with acidic phospholipids (Swanson and Roise 1992; Maduke and Roise 1993). The cardiolipins are especially known to have a high affinity for proteins (Schlame et al. 1991). Following this initial interaction, there is subsequent association with the proteinaceous import machinery (Kiebler et al. 1990) for translocation into the matrix.

As can be seen in Fig. 3, the putative shark mitochondrial targeting signal, consisting of the initial 14 N-terminal amino acids, is somewhat atypical in having but a single hydroxy amino acid, Ser_6. However, no acidic amino acid or proline is present in this sequence. The *Drosophila* targeting signal is quite anomalous in containing both a proline, Pro_{17}, and an acidic amino acid, Glu_{13}. Helical wheel projections of the shark's first 15 amino acids are shown in Fig. 4 and, as can be seen, these amino acids form very strong amphipathic helices with a hydrophobic moment of around 0.54/residue. The first seven N-terminal amino acids, which contain the first putative cleavage site, form an especially strong amphipathic helix of hydrophobic moment 0.80/residue (Segrest et al. 1990).

The other known animal mitochondrial isozymes of GS, those of chicken and alligator liver tissues, are synthesized without N-terminal transient targeting sequences (Smith and Campbell 1983, 1987). By com-

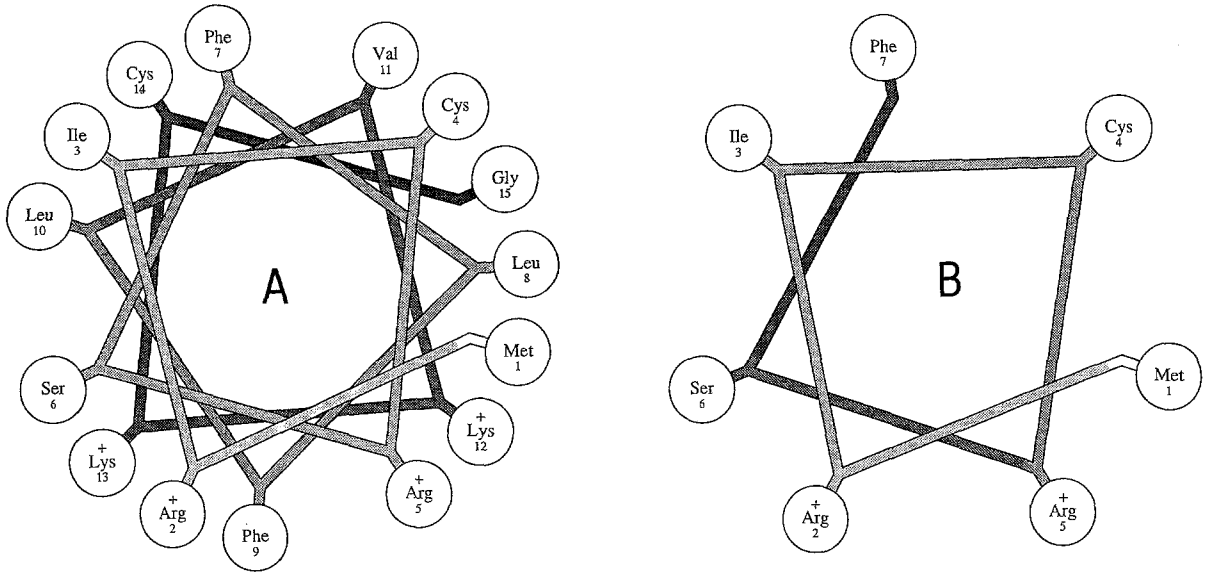

Fig. 4. Helical wheel projections of the N-terminal amino acids of shark preglutamine synthetase. A, amino acids 1–15 and B, 1–7. Projections are with the Wheel program (Segrest et al. 1990). The hydrophobic moment/residue for A is 0.544 and for B, 0.779

paring the N-terminus of chicken mitochondrial GS with the N-termini of the four known cytosolic mammalian enzymes, four amino acid replacements have been identified as possibly having been critical in the conversion of this region to a mitochondrial targeting signal during vertebrate evolution: these were Ser_{10}, Ala_{12}, His_{15}, and Lys_{19} (Fig. 3; Campbell and Smith 1992). It has now been shown experimentally that the N-terminus of at least one other "leaderless" peptide can indeed function as a mitochondrial targeting signal (Arakawa et al. 1990). GS is cytosolic in liver tissue of amphibians, including those which have high levels of GS activity and utilize the enzyme for hepatic ammonia detoxication during periods of water restriction (Campbell et al. 1984). This suggests a mitochondrial-to-cytosolic conversion of hepatic GS during evolution of early tetrapod vertebrates. The highly conserved sequence, Met-Ala-Thr-Ser-Ala-Ser-Ser, at the N-terminus of all higher vertebrate GSs, indicates this conversion involved loss of coding information for all 29 N-terminal amino acids present in shark mitochondrial preGS. The shark cytosolic isozyme would then become ancestral to the GS isozymes of higher vertebrates. This evolutionary sequence would indicate the amino acid replacements identified at the N-terminus of chicken liver GS were therefore necessary to "retarget" the liver enzyme to mitochondria during evolution of the archosaurs.

Conclusions

The amino acid sequence of shark (*Squalus acanthias*) liver mitochondrial GS, as derived from its mRNA cDNA (clone pSGS6), shows more than 90% sequence similarity (conservative replacements included) with that of mammalian and avian GSs. This again emphasizes the very high degree of structural conservation of vertebrate GSs during their evolution. Phylogenetic tree construction for the known animal GSs indicates the ancestral mammalian, avian, and elasmobranch peptides each arose independently from a peptide sharing a common ancestor with the *Drosophila melanogaster* cytosolic isozyme.

The open reading frame of pSGS6 encodes 403 amino acids plus one termination codon. This is 30 more amino acids than GSs from other vertebrates. The calculated mol. wt. of the encoded peptide is 45,406. There are two in-frame initiation codons at the N-terminus of shark liver GS mRNA. Since shark GS appears to be coded for by a single gene, generation of the larger mitochondrial and smaller cytosolic isozyme of GS is presumably due to initiation at one or the other of these sites. Initiation at the upstream site would result in a peptide with an additional 29 amino acids at the N-terminus. This would then represent the mitochondrial precursor peptide. Two-step processing of this precursor during import, based on an identified Arg-X-(Phe, Ile,Leu) motif present in the additional 29 amino acids, would result in removal of 14 N-terminal amino acids to yield a peptide of 43,680 calculated mol. wt. This would correspond to the mature mitochondrial isozyme. Initiation at the down-stream start cite would then generate the cytosolic isozyme of 41,869 mol. wt. The fact that these calculated mol. wts. for the mitochondrial preGS, the mature mitochondrial isozyme, and the cytosolic isozyme agree with previous estimates of size differences of the respective peptides based on their M_rs supports such a processing mechanism.

An evolutionary model in which the additional N-ter-

minal amino acids are deleted during the early evolution of tetrapod vertebrates followed by the later development of an internal mitochondrial targeting signal at the N-terminus in the archosaur line of descent is proposed.

Acknowledgments. Supported by grants DCB 87-18042 and MCB 92-07880 from the NSF. We thank Drs. Paul Anderson and Wilmar Salo, respectively, Univ. Minnesota, Duluth, for a shark liver GS preparation and the shark liver cDNA library; Dr. John Mill, NIH, for the rat GS cDNA probe; and Mr. Martin Jones and Dr. Jere Segrest, Univ. Alabama, for the Wheel program. We also thank Mr. Anand Kolatkar, Rice Univ., for help with the execution of the latter program.

References

Anderson PM (1989) Glutamine-dependent carbamoyl-phosphate synthetase and other enzyme activities related to the pyrimidine pathway in spleen of *Squalus acanthias* (spiny dogfish). Biochem J 261:523–529

Anderson PM (1991) Glutamine-dependent urea synthesis in elasmobranch fishes. Biochem Cell Biol 69:317–319

Aviv H, Leder P (1972) Purification of biologically active globin messenger RNA by chromatography on oligothymidylic-cellulose. Proc Natl Acad Sci USA 69:1408–1412

Arakawa H, Amaya Y, Mori M (1990) The NH_2-terminal 14–16 amino acids of mitochondrial and bacterial thiolases can direct mature ornithine carbamoyltransferase into mitochondria. J Biochem 107:160–164

Benton MJ (1990) Phylogeny of the major tetrapod groups: morphological data and divergence dates. J Mol Evol 30:409–424

Caizzi R, Bozzetti MP, Caggese C (1990) Homologous nuclear genes encode cytoplasmic and mitochondrial glutamine synthetase in *Drosophila melanogaster.* J Mol Biol 212:17–26

Campbell JW, Anderson PM (1991) Evolution of mitochondrial enzyme systems in fish: the mitochondrial synthesis of glutamine and citrulline. In: Hochachka PW, Mommsen TP (eds): Biochemistry and molecular biology of fishes, vol I. Elsevier, Amsterdam, pp 43–75

Campbell JW, Smith DD Jr (1992) Metabolic compartmentation of vertebrate glutamine synthetase: putative mitochondrial targeting signal in avian liver glutamine synthetase. Mol Biol Evol 9: 787–805

Campbell JW, Vorhaben JE, Smith DD Jr (1984) Hepatic ammonia metabolism in a uricotelic tree frog *Phyllomedusa sauvagei.* Am J Physiol 246:R805–R810

Casey CA, Anderson PM (1982) Subcellular localization of glutamine synthetase and urea cycle enzymes in liver of spiny dogfish (*Squalus acanthias*). J Biol Chem 257:8449–8453

Chirgwin JM, Przbyla AE, MacDonald RJ, Rutter WJ (1979) Isolation of biologically active ribonucleic acids from sources enriched in ribonucleases. Biochemistry 18:5294–5299

Driessen HPC, de Jong WW, Tesser GI, Bloemendal H (1985) The mechanism of N-terminal acetylation of proteins. CRC Crit Rev Biochem 18:281–325

Friedman KD, Rosen NL, Newman PJ, Montgomery RR (1988) Enzymatic amplification of cDNA inserts from lambda gtll libraries. Nucleic Acids Res 16:8718

Gavel Y, von Heijne G (1990) Cleavage-site motifs in mitochondrial targeting peptides. Protein Eng 4:33–37

Hartmann C, Christen P, Jaussi R (1991) Mitochondrial protein charge. Nature 352:762–763

Hendrix JP, Hodges PE, Rosenberg LE (1989) Survey of amino-terminal proteolytic cleavage sites in mitochondrial precursor proteins: leader peptides cleaved by two matrix proteases share a three-amino acid motif. Proc Natl Acad Sci USA 86:4056–4060

Hod Y, Utter MF, Hanson RW (1982) The mitochondrial and cytosolic forms of avian phosphoenolpyruvate carboxykinase (GTP) are encoded by different messenger RNAs. J Biol Chem 257:13787–13794

Isaya G, Kalousek F, Rosenberg LE (1992) Amino-terminal octapeptides function as recognition signals for the mitochondrial intermediate peptidase. J Biol Chem 267:7904–7910

Joh T, Takeshima H, Tsuzuki T, Setoyama C, Shimada K, Tanase S, Kuramitsu S, Kagamiyama H, Morino Y (1987) Cloning and sequence analysis of cDNA encoding mammalian cytosolic malate dehydrogenase. J Biol Chem 262:15127–15131

Juretic N, Mattes U, Ziak M, Christen P, Jaussi R (1990) Structure of the genes of two homologous intracellularly heterotopic isoenzymes. Cytosolic and mitochondrial aspartate aminotransferase of chicken. Eur J Biochem 192:119–126

Kiebler M, Pfaller R, Söllner T, Griffths G, Horstmann H, Pfanner N, Neupert W (1990) Identification of a mitochondrial receptor complex required for recognition and membrane insertion of precursor proteins. Nature 348:610–616

Kim KH, Rhee SG (1988) Sequence of peptides from *Saccharomyces cerevisiae* glutamine synthetase. J Biol Chem 263:833–838

King PA, Goldstein L (1983) Renal ammoniagenesis and acid excretion in the dogfish, *Squalus acanthias.* Am J Physiol 245:R581–R589

Kozak M (1984) Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. Nucleic Acids Res 12:857–872

Kuo CF, Darnell JE Jr (1989) Mouse glutamine synthetase is encoded by a single gene that can be expressed in a localized fashion. J Mol Biol 208:45–56

Kumada Y, Benson DR, Hillemann D, Hosted TJ, Rochefort DA, Thompson CJ, Wohlleben W, Tateno Y (1993) Evolution of the glutamine synthetase gene, one of the oldest existing and functioning genes. Proc Natl Acad Sci USA 90:3009–3013

Lemire BD, Fankhauser C, Baker A, Schatz G (1989) The mitochondrial targeting function of randomly generated peptide sequences correlates with predicted helical amphiphilicity. J Biol Chem 264:20206–20215

Maduke M, Roise D (1993) Import of a mitochondrial presequence into protein-free phospholipid vesicles. Science 260:364–367

Mill JF, Mearow KM, Purohit HJ, Haleem-Smith H, King R, Freese E (1991) Cloning and functional characterization of the rat glutamine synthetase gene. Mol Brain Res 9:197–207

Patejunas G, Young AP (1987) Tissue-specific regulation of avian glutamine synthetase expression during development and in response to glucocorticoid hormones. Mol Cell Biol 7:1070–1077

Pesole G, Bozzetti MP, Lanave C, Preparata G, Saccone C (1991) Glutamine synthetase evolution: a good molecular clock. Proc Natl Acad Sci USA 88:522–826

Ritter NM, Smith DD Jr, Campbell JW (1987) Glutamine synthetase in liver and brain tissues of the holocephalan, *Hydrolagus colliei.* J Exp Zool 243:181–188

Roise D, Theiler F, Horvath SJ, Tomich JM, Richards JH, Allison DS, Schatz G (1988) Amphiphillicity is essential for mitochondrial presequence function. EMBO J 7:649–653

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

Schlame M, Beyer K, Hayer-Hartl M, Klingenberg M (1991) Molecular species of cardiolipin in relation to other mitochondrial phospholipids. Eur J Biochem 199:459–466

Segrest JP, DeLoof H, Dohlman JG, Brouillette CG, Ananthara GM (1990) Amphipathic helix motif: classes and properties. Proteins 8:103–117

Smith DD Jr, Campbell JW (1983) Subcellular localization of chicken brain glutamine synthetase and comparison with chicken liver mitochondrial glutamine synthetase. J Biol Chem 258:12265–12268

Smith DD Jr, Campbell JW (1987) Glutamine synthetase in liver of the American alligator, *Alligator mississippiensis*. Comp Biochem Physiol 86B:755–762

Smith DD Jr, Campbell JW (1988) Distribution of glutamine synthetase and carbamoyl-phosphate synthetase in vertebrate liver. Proc Natl Acad Sci USA 85:160–164

Smith DD Jr, Ritter NM, Campbell JW (1987) Glutamine synthetase isozymes in elasmobranch brain and liver tissues. J Biol Chem 262:198–202

Suzuki T, Yoshida T, Tuboi S (1992) Evidence that rat liver mitochondrial and cytosolic fumarases are synthesized from one species of mRNA by alternative translational initiation at two in-phase AUG codons. Eur J Biochem 207:767–772

Swofford DL (1993) Phylogenetic analysis using parsimony, version 3.1. Distributed by the Illinois Natural History Survey, Champaign, IL

Swanson ST, Roise D (1992) Binding of a mitochondrial presequence to natural and artificial membranes: role of surface potential. Biochemistry 31:5746–5751

Tropschug M, Nicholson DW, Hartl F.-U, Köhler H, Pfanner N, Wachter E, Neupert W (1988) Cyclosporin A-binding protein (cyclophilin) of *Neurospora crassa*. One gene codes for both the cytosolic and mitochondrial forms. J Biol Chem 263:14433–14440

von Heijne G (1986) Mitochondrial targeting signals may form amphiphilic helices. EMBO J 5:1335–1342

von Heijne G, Steppuhn J, Herrmann RG (1989) Domain structure of mitochondrial and chloroplastal targeting peptides. Eur J Biochem 180:535–545