3118–3123 *Nucleic Acids Research, 1997, Vol. 25, No. 15*

Intronic polyadenylation in the human glycinamide ribonucleotide formyltransferase gene

Julie L. C. Kan+ and Richard G. Moran*

Department of Pharmacology and Toxicology and the Massey Cancer Center, Medical College of Virginia, Virginia Commonwealth University, Richmond, VA 23298-0230, USA

Received March 19, 1997; Revised and Accepted June 23, 1997 DDBJ/EMBL/GenBank accession nos AF008653–55

ABSTRACT

The mouse glycinamide ribonucleotide formyltransferase (GART) locus is known to produce two functional proteins, one by recognition and use of an intronic polyadenylation site and the other by downstream splicing. We now report a similar intronic polyadenylation mechanism for the human GART locus. The human GART gene has two potential polyadenylation signals within the identically located intron as that involved in intronic polyadenylation in the mouse gene. Each of the potential polyadenylation signals in the human gene was followed by an extensive polyT rich tract, but only the downstream signal was preceded by a GT tract. Only the downstream signal was utilized. The polyT rich tract which followed the functional polyadenylation site in the human GART gene was virtually identical in sequence to a similarly placed region in the mouse gene. An exact inverted complement to the polyT rich stretch following the active polyadenylation signal was found in the upstream intron of the human gene, suggesting that a hairpin loop may be involved in this intronic polyadenylation.

INTRODUCTION

Multifunctional enzymes have been found in several metabolic pathways in eukaryotes, e.g., purine and pyrimidine synthesis and the interconversion of folates. A comparison of the sequence of the genes encoding these multifunctional proteins across species suggests that they arose by fusion of genomic regions encoding single domain proteins with the loss of individual promoters (1,5,8). In mammals and birds, the GART locus encodes an enzyme of purine synthesis which catalyzes three steps of this pathway [glycinamide ribonucleotide synthetase (GARS), glycinamide ribonucleotide formyltransferase (GART) and aminoimidazole ribonucleotide synthetase (AIRS)] (1,4,8), each of which is expressed as a separate protein in *Escherichia coli* and *Bacillus subtilis* (6,13–15). Paradoxically, the mouse GART gene has also maintained the ability to produce a second mRNA which translates to a monofunctional protein with GARS activity (8). It is not clear why a mechanism to express a separate, monofunctional GARS activity has been retained after the evolutionary selection of a trifunctional protein which also catalyzes the GARS reaction. However, the structure of the mouse GART gene suggests that the smaller protein has been actively retained and that it is not just a vestigial remnant (9).

Our previous studies on the mouse GART locus demonstrated that the mRNAs for monofunctional GART were produced by the utilization of several polyadenylation signals present in intron 11 (9). In the mouse, the transcripts for the trifunctional protein use the same promoter and same initial 11 exons of this gene as does the monofunctional GARS but processing of these transcripts ignores the intron 11 polyadenylation signals and splices in the downstream exons 12–22. Hence, processing of the primary transcripts from the GART gene represents a case in which two functional proteins are produced from a single gene but by alternative intronic polyadenylation and splicing rather than by alternative exon usage. The mechanism whereby intronic polyadenylation occurs for the GART gene remains to be established, although recent studies have directly addressed this issue in the immunoglobulin heavy chain gene (16). In this study, we isolated the human GART genomic locus and searched for *cis* elements common between the human and mouse (9) genes that might be involved in this mechanism. From the patterns of sequence conservation and analysis of frequency of utilization of polyadenylation signals, we noted that an extensive polyT tract located in the intron distal to the last utilized intronic cleavage/polyadenylation signal and a weak 5′ splice donor site were common to both mouse and human GART genes. The GART gene in each species also had a backup motif which could accentuate the effectiveness of these intronic sequence motifs on cleavage/polyadenylation while still allowing downstream splicing.

MATERIALS AND METHODS

Library screening

Human genomic clones for the GART locus were isolated from a λ FixII library constructed from male placental DNA (Stratagene). The library was screened with a cDNA probe, derived by PCR from poly(A)$^+$ selected RNA isolated from CEM human lymphoblastic leukemia cells. This probe was clone HR24

*To whom correspondence should be addressed at: Massey Cancer Center, Medical College of Virginia, MCV Box 980230, Richmond, VA 23298, USA. Tel: +1 804 828 5783; Fax: +1 804 828 5782; Email: rmoran@gems.vcu.edu

+Present address: Howard Hughes Medical Institute, Program in Molecular Medicine and University of Massachusetts Medical Center, 373 Plantation Street, Suite 309, Worcester, MA 01605, USA

(Fig. 3B) derived by 3′ RACE; the sequence has been deposited in GenBank. The library screen was performed using Denhardt's solution hybridization conditions as previously described (12).

Sequence analysis

Human GART cDNA sequence (1) that corresponds to exons 11–14 of the mouse gene (9) were mapped to restriction fragments of λ clones by Southern blot hybridization to the cDNA probe used for the library screen and with 25mer oligonucleotides corresponding to this region. Hybridizing restriction fragments were subcloned into pBluescript SKII(+). Double-stranded plasmid DNA was isolated by the alkaline lysis method and partially sequenced using Sequenase 2.0 (Amersham).

Determination of the use of polyadenylation signals

Two independent methods were used to determine the use of polyadenylation signals, 3′ RACE and ribonuclease protection assays. 3′ RACE was performed as previously described (7,9) using poly(A)$^+$ selected RNA from CEM cells, reverse transcribed using an oligo dT primer containing sequence of an anchor primer. The gene-specific primers used for 3′ RACE were 5′-ctcaagctctag-gactggaggtgttccatgcaggc-3′ (P1), 5′-cttcatgatagcgtaagtttgg-3′ (P2) and 5′-actgaagatgagaatactggtc-3′ (P3). The PCR products obtained were gel purified and subcloned into the pCRII vector (Invitrogen). Constructs used for ribonuclease protection assays were two adjacent *Sau*3AI fragments of 0.9 and 0.4 kb from the genomic subclone HJK32 ligated into the *Bam*HI site of pBluescript SKII (+) and named HS1 and HS13, respectively (Fig. 3B). The subclone HJ1 was created by ligating the *Rsa*I fragment from clone HS13 into the *Eco*RV site of pBluescript SK (+). Templates for *in vitro* transcription were linearized with either *Fok*I, *Bst*71I, *Nde*I or *Eco*RV (Fig. 3B) and were transcribed using either T3 or T7 RNA polymerase. Ribonuclease protection assays were performed essentially as described (2). Briefly, 30 µg of total RNA from CEM cells was hybridized overnight at 50°C with an *in vitro* transcribed RNA probe. Approximately 5×10^5 c.p.m. of labeled RNA probe was used for each 30 µl reaction. The hybridized RNA was treated with 100 µg/ml of RNAse A at 4°C for 30 min followed by incubation in 250 µg/ml of proteinase K at 42°C for 15 min. The samples were extracted with phenol–chloroform, then precipitated with ethanol. The precipitated RNA was resuspended in loading buffer and fractionated on a 6% polyacrylamide sequencing gel.

RESULTS

Isolation and partial characterization of the human GART locus

A total of 1.5×10^6 λ plaques from a human male placenta λ FIXII genomic library were screened using a radiolabeled probe containing human cDNA sequence corresponding to the end of the GARS domain and the beginning of the AIRS domain. Twenty two genomic clones were isolated. The alignment of four clones which overlapped in the region of interest is depicted in Figure 1A. Restriction mapping and Southern blot analysis using an oligo-nucleotide probe corresponding to the 3′-segment of the GARS domain identified a hybridizing *Bgl*II–*Bam*HI fragment on clone λHGAG1 which was then subcloned into the *Bam*HI–*Sty*I sites of pBluescript SKII (+). A *Bam*HI–*Hin*dIII subfragment (Fig. 1A) which contained human GART exon 14 sequence was also subcloned. The sizes of the exons 11–14 and the junctional sequences in this region of the human GART gene were determined by limited sequence analysis (Table 1). The exon numbering presented here was arbitrarily set to the assignments previously found for the mouse GART locus (9) to permit a facile comparison of these two genes. Exons 11–14 in the human GART gene were identical in size and in the exact site of interruption of the coding region by introns previously reported for the mouse gene (9). In the human GART gene, as in the mouse, exon 11 constituted the end of the coding region for the GARS domain. For the smaller (monofunctional GARS) transcript, the stop codon, taa, was part of the 5′-splice donor site used to generate the trifunctional transcript, and the 3′-untranslated region of the monofunctional GARS message was contiguous with the end of exon 11 (see below). Thus, as was previously found to be the case in the mouse, the monofunctional GARS message was produced from primary transcripts from this human gene by cleavage and polyadenylation within what constitutes intronic sequence for the longer transcript encoding the trifunctional GARS–AIRS–GART protein.

We have previously determined that multiple polyadenylation signal sequences were present in intron 11 of the mouse GART gene, most of which could be used to generate the heterogeneous 3′-termini of the monofunctional GARS mRNAs (9). There was also a striking 24 nt polyT tract in the mouse gene immediately downstream of the most 3′ polyadenylation signal. Detailed sequence analysis of intron 11 in the human GART locus was performed to determine whether any or all of these structural features were conserved between mouse and human genes. In the first 900 nt of intron 11, immediately downstream of the last GARS domain-encoding exon in the human gene, there were two candidate polyadenylation signal sequences, which are referred to below as poly(A) site I (ATTAAA) and II (AATAAA) (Fig. 1B). Interestingly, there was a polyT tract downstream of each of these potential polyadenylation signals, although the GT rich tract associated with polyadenylation sites was present only for site II, located immediately upstream of the polyT tract.

Utilization of intronic polyadenylation signals

3′ RACE was performed with poly(A)$^+$ RNA from CEM cells to determine the use of the two intronic polyadenylation signal sequences. A gene-specific PCR primer derived from exon 11 and two from intron 11 were individually used in conjunction with the 3′ anchor primer for these reactions. The position of these PCR primers relative to the two potential polyadenylation signals in human intron 11 are shown in Figure 2A. PCR products were cloned into pCRII and multiple clones from each reaction were sequenced. The positions of cleavage/polyadenylation detected by 3′ RACE are summarized in Figure 2B. PCR reactions with the anchor primer and either intron primer 2 (P2) or exon 11 primer (P1) would be capable of identifying mRNAs that used either of the two potential polyadenylation signals present in intron 11. PCR products corresponding to mRNAs that used poly(A) site I were not represented in the clones sequenced, but cDNAs corresponding to poly(A) site II usage were found. The frequent, if not exclusive, use of poly(A) site II among transcripts terminating in intron 11 was confirmed in subsequent 3′ RACE experiments using intron primer 3 (P3) and an anchor primer (Fig. 2B). Although 3′-RACE with the P1 primer would also detect the trifunctional transcript, gel bands within this size range were not analyzed.

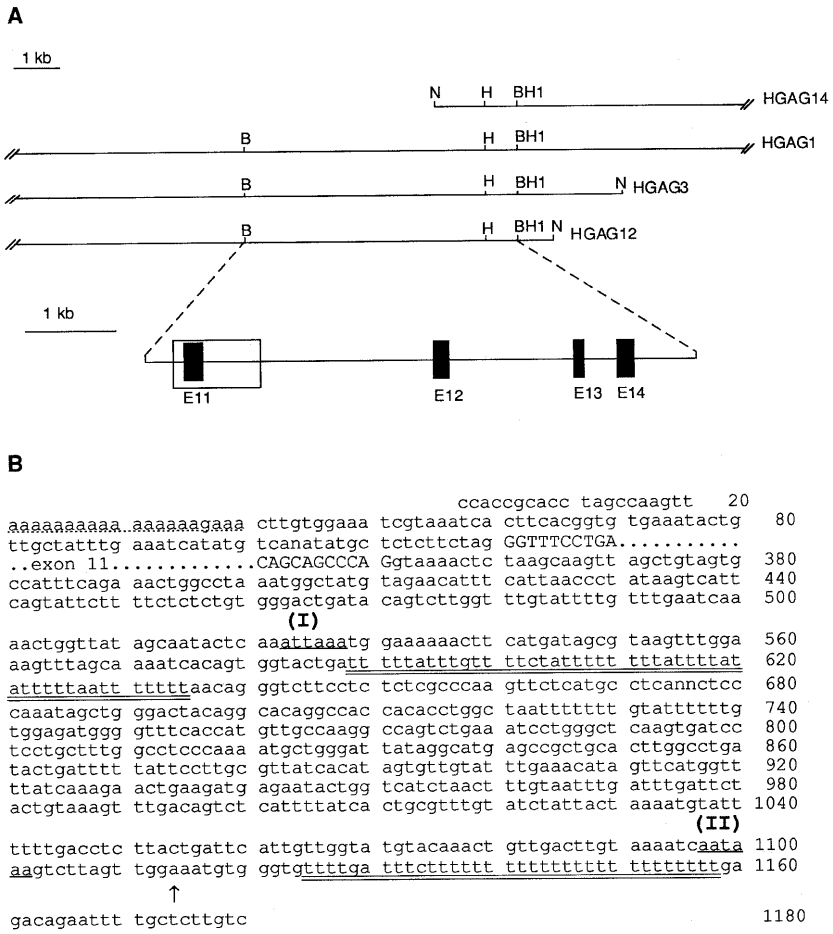


Figure 1. (**A**) Alignment of human genomic clones isolated for the GART locus. The positions of exons 11–14 are shown as filled boxes and the site of polyadenylation is noted by an arrow. The region whose sequence is summarized in (B) is denoted by an open box. Restriction sites mapped in this region were *Bgl*II (B), *Hin*dIII (H) and *Bam*HI (BHI). The *Not*I (N) site shown for some phage is the cloning site for the λ gt11 library. (**B**) Nucleotide sequence of exon 11 (upper case letters) and the surrounding intron sequence (lower case letters). The two polyadenylation signal sequences are underlined and denoted as poly(A) sites I and II. The polyT tracts are double underlined. An inverted complementary sequence to the nt 1405–1438 polyT tract is indicated by the dotted line.

Table 1. Partial intron/exon boundaries of the human GART gene

Exon	Exon size	Domain	5′ Intron	Exon sequence	3′ Intron sequence	Intron size (nt)
11	231	GARS	tctcttctag	GGTTTC.....GCCCAG	gtaaaactct	2200
12	95	AIRS	aattctttag	GAGTTT.....GATCAG	gtcagtgatc	1300
13	110	AIRS	tctccaacag	GCTGTA.....CTAAAG	gtaatcagac	343
14	199	AIRS	tcttacatag	ATTGCC.....TCCTTG	gtatgatcag	ND

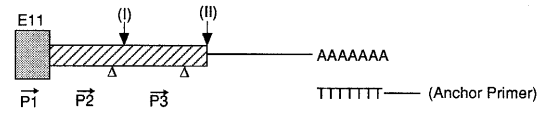
ND, not determined.

Identification and characterization of an unexpected transcript from the human GART locus

Several of the 3′ RACE clones obtained using the exon 11 primer did not represent the use of either of the polyadenylation signals in intron 11. These PCR products initiated from the primer used, contained the remaining exon 11 sequence followed by all of exons 12 and 13, a stretch of 83 nt that did not have homology to exon 14, followed immediately by 39–41 nt of exon 14 and a polyA tail (the sequences described are deposited in GenBank) (Fig. 3A). This set of PCR products represented a novel transcript that was not detected among the mouse GART transcripts (8) and appeared to represent an alternatively spliced RNA product.

Partial sequencing of λ subclones from this genomic region indicated that the unexpected 83 nt constituted the most 3′ end of intron 13, which was a total of 343 nt in length. Hence, this splice variant used an alternative 3′-splice acceptor site than that used to generate the trifunctional GARS–AIRS–GART mRNA; this region of the two transcripts was designated exons 14A and 14, respectively (Fig. 3A). This splice variant originated from use of a cryptic 3′-splice acceptor site in intron 13, coupled to the use of

A



B

Primers Used	Clones Sequenced	Termination at SiteI	Termination at Site II	Termination in Exon 14
P1+Anchor	18	0	0	3
P2+Anchor	24	0	4	N.A.
P3+Anchor	15	N.A.	8	N.A.

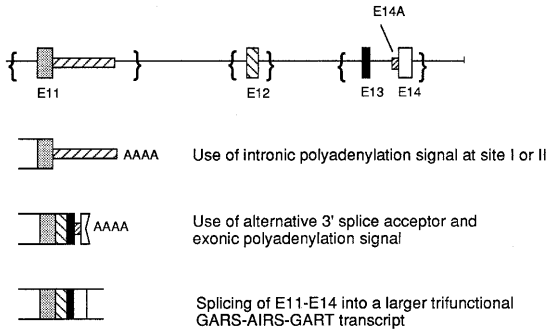
Figure 2. (**A**) Schematic representation of the 3′ RACE PCR. Exon 11 (E11, shaded box) and the adjacent intron (cross-hatched box) of the human GART gene are aligned with the positions of the gene specific primers used for PCR (horizontal arrows). The expected sites of polyadenylation for site I (I) and site II (II) are depicted by vertical arrows. The position of the polyadenylation signals (ATTAAA or AATAAA) are depicted as triangles. (**B**) Summary of clones analyzed. Clones that would not be obtained in each set of primer reactions are indicated as N.A. (not applicable). Note that the number of clones sequenced exceeds the number of clones that are specific to the GART gene. These clones represented non-specific transcripts amplified in the 3′ RACE.

a polyadenylation signal present within coding region in exon 14. In addition, 3′ RACE experiments did not detect a splice product which used the exon 14A 3′ splice acceptor site but not the polyadenylation signal sequence in exon 14.

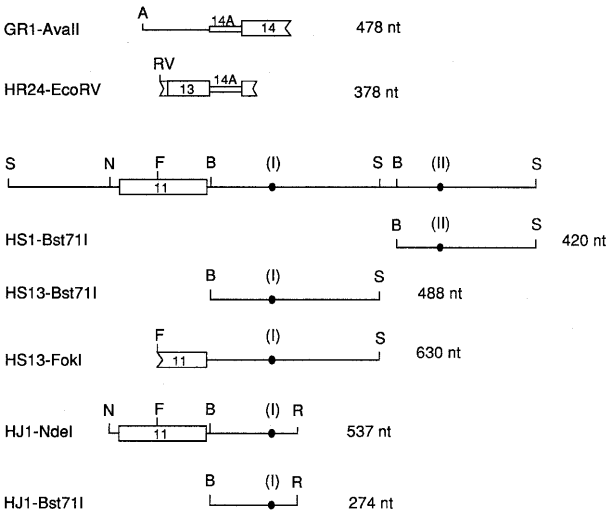
Quantitation of transcripts from the human GART gene

The presence of three possible transcripts from the human GART locus [a monofunctional GARS using intron 11 polyadenylation signal(s), the splice variant from polyadenylation in exon 14, and a trifunctional transcript] prompted an investigation of the relative levels of these transcripts in human cells. Ribonuclease protection assays using several antisense RNA probes (shown in Fig. 3B) allowed estimation of the relative levels of these transcripts in total RNA from human CEM leukemic cells. Two probes, HR24 and GR1 (Fig. 3B), were used to determine the relative message level of the splice variant that used exon 14A and that for the trifunctional GARS–AIRS–GART transcript.

A



B



C

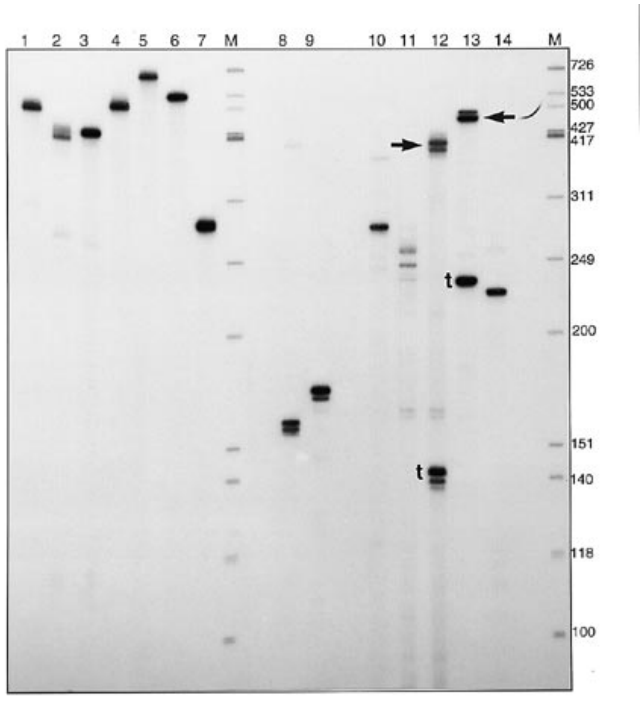


Figure 3. (**A**) Schematic representation of the possible transcripts from the human GART gene. The top drawing depicts the region of the human GART genomic locus. The bracketed areas of this drawing indicate the genomic regions that have been sequenced and deposited in GenBank. The boxes represent exons which are separated by introns (lines). The horizontal box at the end of exon 11 represents the 3′ untranslated region of the monofunctional GARS. The horizontal box before exon 14 represents the alternative 3′ splice acceptor site detected by PCR. (**B**) Schematic representation of the transcripts used for ribonuclease protection probes. The length of each probe is denoted in nucleotides next to each diagram. (**C**) Ribonuclease protection analysis for transcripts from CEM cells. The probes used were as follows: lanes 1 and 8, GR1–*Ava*II; lanes 2 and 9, HR24–*Eco*RV; lanes 3 and 10, HS1–*Bst*71I; lanes 4 and 11, HS13–*Bst*71I; lanes 5 and 12, HS13–*Fok*I, lanes 6 and 13, HJ1–*Nde*I; lanes 7 and 14, HJ1–*Bst*71I. CEM total RNA (30 µg) was used for protection in lanes 8–14, and lanes 1–7 were loaded with 1×10^3 c.p.m. of probe. The bands corresponding to the trifunctional (exon 11/12 spliced) transcripts are noted by a 't', and the monofunctional transcripts by arrows in lanes 12 and 13. Marker lanes (M) are ^{32}P-end-labeled *Hin*fI-digested fragments of ϕX174 DNA. At the exposure shown or at much longer exposures, distinct bands were not detected in the control reactions containing 30 µg of yeast tRNA (data not shown).

Surprisingly, only protected fragments (162 and 175 nt, lanes 8 and 9 of Fig. 3C) corresponding to the trifunctional transcript were detected, and the fragments expected for the splice variant (245 and 258 nt) were too infrequent to be seen in these experiments (Fig. 3C). The results of these ribonuclease protection experiments agreed with Northern blots in which a signal corresponding to the message containing sequence unique to intron 14A was not detectable even after long exposures of the film (data not shown).

Ribonuclease protection assays using an antisense probe produced by linearization of subclone HS1 with *Bst*71I (Fig. 3B), allowed the assessment of the use of polyA site II. A major protection product of ~283 nt was detected (Fig. 3C, lane 10) which corresponded to a fragment size expected for cleavage and polyadenylation at the site predicted by the 3′-RACE product which would use polyA site II. The minor band at 378 nt coincided with the size expected for unspliced transcript.

A subclone, HS13, representing a *Sau*3AI fragment of the human GART locus that contains sequence of exon 11 and the first 436 nt of intron 11 was linearized by either *Fok*I or *Bst*71I and antisense probes were transcribed from these substrates. Ribonuclease protection assays using the antisense probe HS13-*Fok*I would allow the detection of transcripts that represented the use of poly(A) sites I and II as well as the trifunctional GARS–AIRS–GART message. On the other hand, the antisense probe HS13-*Bst*71I would not hybridize to mRNA containing exon 11 but would allow the detection of transcripts using either poly(A) sites I or II. As cleavage and polyadenylation occurs within 10–30 nt from the polyadenylation signal (17), the expected size of products for use of polyA site I were 330–360 and 178–209 nt and for use of polyA site II were 587 and 435 nt for the *Fok*I and *Bst*71I probes, respectively. Neither of these sets of products was obtained (Fig. 3C). Rather, protected fragments of 137, 397 and 410 nt were obtained using the *Fok*I probe and fragments of 247 and 260 nt were detected using the *Bst*71I probe (Fig. 3C, lanes 11 and 12). The 137 nt band was generated by protection of only the exon 11 sequence of the trifunctional transcript. With both of these probes, the size of the protected fragments would place the site of cleavage 79 and 101 nt downstream from polyadenylation signal sequence I, substantially longer than the maximal expected distance of 30 nt for use of poly(A) site I. However, these fragments also do not represent the protection fragments expected from use of the poly(A) site downstream. Instead, the cleavage sites that generated the products obtained in these reactions mapped exactly in the first of two polyT tracts in intron 11 (Fig. 3B). Hence, it appeared that all of the products obtained in these reactions corresponded to transcripts which used the downstream polyA site II that had been shortened by spurious RNAse A digestion due to breathing of the polyT stretch (2). The use of poly(A) signal I appeared to be excluded by these experiments, although the interpretation of the data was clouded by the presence of the upstream polyT tract in the protection probe.

A construct that had the polyT stretch removed, HJ1, was used for additional ribonuclease protection experiments to confirm these conclusions. Linearization of HJ1 with *Nde*I or *Bst*71I was used to make the antisense probes HJ1–*Nde*I and HJ1–*Bst*71I (shown in Fig. 3B). Protection of the *Nde*I probe by CEM mRNA produced fragments of 231 and 468 nt (Fig. 3C, lane 13) whereas protection of the *Bst*71I probe resulted in a single band of 228 nt (lane 14). The 231 nt band represented, again, the protection of only exon 11 sequence of the larger trifunctional GARS–AIRS–GART transcript. The 468 nt band spans the length of exon 11 and all of the intronic sequence present in this subclone, indicating the protection of a monofunctional transcript terminating downstream of site II. The 228 nt band derived from HJ1–*Bst*71I, likewise indicated the exclusive use of the intronic downstream site II. Thus, poly(A) site I was not used and appeared to represent a non-functional polyadenylation signal.

Hence, these ribonuclease protection experiments gave unambiguous confirmation that site II is the only intronic polyadenylation signal used in intron 11 of the human GART gene. It should also be noted from Figure 3C (lanes 12 and 13) that the monofunctional GARS transcripts are about as frequently represented in CEM RNA as is the trifunctional spiced product.

DISCUSSION

Chicken, mouse and human tissues express two sets of transcripts related to GART, a 1.7–1.9 kb message encoding a monofunctional GARS, and an ~3.4 kb transcript encoding three domains that catalyze the second (GARS), third (GART) and fifth (AIRS) reactions of *de novo* purine synthesis (1,8). We have previously reported that the protein encoded by the smaller set of messages in the mouse is a functional enzyme, in spite of the fact that this GARS reaction is also carried out by the trifunctional enzyme (8). We also showed that both classes of messages come from the same genomic locus (9). The monofunctional GARS transcript was generated by use of the first 11 exons of the mouse GART gene followed by cleavage and polyadenylation within intron 11, whereas for the larger transcript, the intronic polyadenylation signals present in this intron are ignored and exon 11 is spliced to exons downstream. In this study, we demonstrate that intronic polyadenylation is also responsible for generation of the monofunctional GARS transcripts in human cells and we examine the structural features of the human gene to search for common features with the mouse GART locus that might explain this phenomenon.

The most striking commonality between the mouse and human GART genes surrounding the intronic polyadenylation signals is a nearly identical polyT tract:

mouse: GTGTTTTGATTTT$(T)_{24}$T
human: GTGTTTTGATTTC$(T)_{24}$G

This sequence motif is immediately downstream of the only active polyadenylation signal in the human GART intron 11 and is also present close downstream of the most distal intron 11 polyadenylation signal used in the mouse gene. This polyT motif is immediately downstream of the GT tract (apparently required for efficient polyadenylation; 11) and the polyadenylation signal (AATAAA). This degree of conservation of a polyT region within intronic sequence would seem unlikely to be fortuitous, and may reflect an important component in the mechanism for generation of the smaller transcript in human and mouse GART genes.

Intron 10 of the human GART locus was partially sequenced during localization of the intron/exon borders of this region of the gene. These studies revealed the presence of a 20 nt sequence which was an exact inverted complement to a segment of the intron 11 polyT rich tract. This suggested that the polyT tract might be influencing the choice between splicing of exon 11 downstream and that of intronic cleavage/polyadenylation by the formation of a hairpin structure in the unprocessed primary

transcript. This finding prompted us to determine the sequence of the intron immediately upstream of exon 11 in the mouse gene. Somewhat to our surprise, a complement to the mouse polyT tract was not found, at least within this exon, although it might be present elsewhere in the gene. However, for the human GART gene, there is only a single active polyadenylation signal whereas several intronic polyadenylation signals are present and are utilized in the mouse gene. Hence, it seems entirely likely that the formation of a hairpin loop accents the effects of the presence of the single active intronic polyadenylation site in the human gene, whereas the multiple active signals are sufficient to bring about intronic transcript termination in the mouse gene.

The 5′ splice donor site immediately upstream of the site of intronic polyadenylation in both mouse and human GART genes differs from the consensus sequence by the presence of an A nucleotide in the +5 position, a deviation from the G nucleotide normally at this position. The mouse and human immunoglobulin mu heavy chain (IgM) locus, and other immunoglobulin genes that exhibit intronic polyadenylation, all share this deviation from consensus at the 5′ splice donor site immediately upstream of the intronic polyadenylation signal (10). Previous studies on the immunoglobulins have demonstrated that mutation of this A nucleotide to the consensus G nucleotide markedly altered the pattern of splicing versus polyadenylation (10), directly implicating this 'weak' splice site in the phenomenon and suggesting that it is needed to ensure downstream splicing in at least a fraction of the processed transcripts. This similarity suggests that the control of splicing and polyadenylation in the GART and immunoglobulin genes may share a common molecular mechanism.

During the course of our studies on the human GART gene, a splice variant was found that used a polyadenylation signal in exon 14 (Fig. 3A). (The AATAAA found in exon 14 of the human sequence was not preserved in the mouse sequence so that analogous mouse transcripts were not possible.) Transcripts which used the 3′-splice acceptor site of exon 14 and the AATAAA within exon 14 were not found, nor were transcripts using exon 14A without recognition of the exonic polyadenylation signal sequence, although either would have been detected by our experiments. Nevertheless, although this class of transcripts was frequently isolated by 3′ RACE, the message proved to be of low abundance as evidenced by its absence in both Northern blots and ribonuclease protection assays. We concluded that it represented a rare transcript which was overemphasized by PCR.

There is a short but growing list of genes recognized to produce different forms of the same protein by a choice of intronic polyadenylation versus downstream splicing. The two transcripts generated from the GART locus represent an extreme case with two functional proteins being produced, one of which has only GARS activity, the other capable of three of the steps of purine synthesis including the GARS reaction. To our knowledge, all of the other cases analyzed to date involve a switch from one isotype of a protein to another in response to the needs of a tissue-specific or developmental program of gene expression. Although it is not clear whether there is a developmental role for the GART gene, others have implicated this locus in Down's syndrome (3). From the striking sequence similarities and differences in the mouse and human GART genes near the region involved in intronic polyadenylation, this locus seems to be an interesting and informative model system to dissect the mechanisms involved in the choice of intronic polyadenylation versus downstream splicing. Whether and the extent to which the mechanism of this choice in the GART gene is common with other genes which utilize intronic cleavage/polyadenylation will prove useful in understanding the steps involved at a biochemical level.

ACKNOWLEDGEMENTS

We thank Mr Scott Smith for his excellent technical assistance. The study was supported in part by grant CA-27605 from the National Institutes of Health, DHHS and by training grant T32-CA-09564.

REFERENCES

1 Aimi,J., Qiu,H., Williams,J., Zalkin,H. and Dixon,J.E. (1990) *Nucleic Acids Res.* **18**, 6665–6672.
2 Ausubel,F.M., Brent,R., Kingston,R.E., Moore,D.D., Seidman,J.G., Smith,J.A. and Struhl,K. (1994) (eds), *Current Protocols in Molecular Biology.* Massachusetts General Hospital, Harvard Medical School, Boston, MA.
3 Chang,F.H., Barnes,T.S., Schild,D., Gnirke,A., Bleskan,J. and Patterson,D. (1991) *Somat. Cell Mol. Genet.* **17**, 411–420.
4 Daubner,S.C., Schrimsher,J.L., Schendel,F.J., Young,M., Henikoff,S., Patterson,D., Stubbe,J. and Benkovic,S.J. (1985) *Biochemistry* **24**, 7059–7062.
5 Davidson,J.N., Chen,K.C., Jamison,R.S., Musmanno,L.A. and Kern,C.B. (1993) *Bioessays* **15**, 157–164.
6 Ebbole,D.J. and Zalkin,H. (1987) *J. Biol. Chem.* **262**, 8274–8287.
7 Frohman,M.A., Dusch,M.K. and Martin,G.R. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 8998–9002.
8 Kan,J.L.C. (1993) The mouse gene for three steps of *de novo* purine synthesis. Ph.D. dissertation, Department of Biochemistry and Molecular Biology, University of Southern California, Los Angeles.
9 Kan,J.L.C., Jannatipour,M., Taylor,S.M. and Moran,R.G. (1993) *Gene* **137**, 195–202.
10 Kan,J.L.C. and Moran,R.G. (1995) *J. Biol. Chem.* **270**, 1823–1832.
11 Peterson,M.L. and Perry,R.P. (1989) *Mol. Cell. Biol.* **9**, 726–738.
12 Proudfoot,N. (1991) *Cell* **64**, 671–674.
13 Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual, 2nd ed.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
14 Shen,Y., Rudolph,J., Stern,M. and Stubbe,J., Flannigan,K.A. and Smith,J.M. (1990) *Biochemistry* **29**, 218–227.
15 Smith,J.M. and Daum,H.A.,III (1986) *J. Biol. Chem.* **262**, 10632–10636.
16 Smith,J.M. and Daum,H.A.,III (1987) *J. Biol. Chem.* **263**, 10565–10569.
17 Takagaki,Y., Seipelt,R.L., Peterson,M.L. and Manley,J.L. (1996) *Cell* **87**, 941–952.
18 Wickens,M. (1990) *Trends Biochem. Sci.* **15**, 277–281.