Bidirectional imprinting of a single gene: GNAS1 encodes maternally, paternally, and biallelically derived proteins

BRUCE E. HAYWARD, VERONICA MORAN, LISA STRAIN, AND DAVID T. BONTHRON*

Human Genetics Unit, University of Edinburgh, Western General Hospital, Edinburgh EH4 2XU, United Kingdom

Edited by Victor A. McKusick, Johns Hopkins Hospital, Baltimore, MD, and approved October 26, 1998 (received for review August 26, 1998)

ABSTRACT The GNAS1 gene encodes the α subunit of the guanine nucleotide-binding protein Gα, which couples signaling through peptide hormone receptors to cAMP generation. GNAS1 mutations underlie the hormone resistance syndrome pseudohypoparathyroidism type Ia (PHP-Ia), so the maternal inheritance displayed by PHP-Ia has raised suspicions that GNAS1 is imprinted. Despite this suggestion, in most tissues Gα is biallelically encoded. In contrast, the large G protein XLαs, also encoded by GNAS1, is paternally derived. Because the inheritance of PHP-Ia predicts the existence of maternally, rather than paternally, expressed transcripts, we have investigated the allelic origin of other mRNAs derived from GNAS1. We find this gene to be remarkable in the complexity of its allele-specific regulation. Two upstream promoters, each associated with a large coding exon, lie only 11 kb apart, yet show opposite patterns of allele-specific methylation and monoallelic transcription. The more 5′ of these exons encodes the neuroendocrine secretory protein NESP55, which is expressed exclusively from the maternal allele. The NESP55 exon 11 kb 5′ to the paternally expressed XLαs exon. The transcripts from these two promoters both splice onto GNAS1 exon 2, yet share no coding sequences. Despite their structural unrelatedness, the encoded proteins, of opposite allelic origin, both have been implicated in regulated secretion in neuroendocrine tissues. Remarkably, maternally (GNAS1), paternally (XLαs), and biallelically (Gαs) derived proteins all are produced by different patterns of promoter use and alternative splicing of GNAS1, a gene showing simultaneous imprinting in both the paternal and maternal directions.

Activation of several hormone receptors of the seven-transmembrane domain family is coupled to production of the second messenger cAMP via a heterotrimeric G protein, Gα, which stimulates adenylate cyclase. The guanine nucleotide-binding α subunit of Gα is encoded by 13 exons of the GNAS1 gene, located at chromosome 20q13.2-q13.3 (1, 2). Null mutations of GNAS1 have been found in patients with the hormone-resistance syndrome pseudohypoparathyroidism type Ia (PHP-Ia), which also has a 50% reduction in erythrocyte Gα activity (3–5). However, similar reductions in Gα activity also are found in individuals who are not hormone resistant, but who show the characteristic skeletal phenotype associated with PHP-Ia, known as Albright hereditary osteodystrophy (6). The clinical picture in these individuals, who may be relatives of patients with typical PHP-Ia, sometimes is referred to as pseudopseudohypoparathyroidism (PPHP).

The variable phenotype associated with GNAS1 mutations, even within one family, suggests the influence of other genetic factors on the presence or absence of endocrine disturbance. One suggestion, based on a review of published PHP/PPHP families with two or more generations, is that genomic imprinting plays a role (7). Full-blown PHP-Ia was observed in 60 of 60 instances of maternal transmission, whereas PPHP alone was seen in all three instances of paternal transmission. The pedigrees presented in a large recent study of PHP/PPHP in France provide strong further support for this inheritance bias (even though its authors did not favor the imprinting hypothesis) (8). Although there has been one suggested exception to the maternal inheritance rule (9), other reports are also consistent with this model (10, 11).

The requirement for maternal inheritance for development of PHP-Ia suggests maternal allele-specific expression of GNAS1, at least in some tissues. However, we previously showed that in human fetal tissues, GNAS1 is biallelically expressed (12). In the mouse, there are some discrepancies between the results of different studies. Williamson et al. (13) found Gnas (which lies within a candidate imprinted region of distal chromosome 2) to be bialleically expressed in almost all tissues; in mice with uniparental disomy for this region, evidence for preferential expression of the paternal Gnas allele was found only in the renal glomerulus. Neither the tissue specificity nor the preferential paternal activity are as predicted by the physiology and the maternal inheritance of PHP-Ia. However, in a study of Gnas knockout mice (14), only maternal inheritance of an exon 2-disrupted Gnas resulted in hormone resistance, and the selectivity of this resistance for certain hormones (including parathyroid hormone) was suggested to result from biallelic expression of Gnas in most tissues, but from maternal expression in the proximal renal tubule.

In addition to these studies of murine Gnas, it recently has emerged that the confusion over whether or not human GNAS1 is imprinted also results in part from an incomplete picture of the structure and function of this gene. We have shown that ~35 kb upstream of the body of GNAS1 is a region of maternal allele-specific methylation (15). This region includes a paternally active alternative promoter and a large first exon that is spliced onto GNAS1 exon 2 in place of the normal exon 1. The resulting paternal transcripts appear to encode a protein with a large N-terminal extension, in-frame with the G protein sequences encoded by GNAS1 exons 2–13. The homologue of this large G protein in the rat is termed XLαs and is expressed specifically in neuroendocrine cells (16). In these cells, XLαs is tightly associated with the membrane of the trans-Golgi network and cofractionates with secretogranin II. It is hypothesized that this large G protein may play a role in the regulation of formation of secretory vesicles.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database [accession nos. AJ009849 (human genomic NESP55 sequence) and AJ010163 (mouse NESP55 cDNA sequence)].

This paper was submitted directly (Track II) to the Proceedings office. Abbreviations: PHP, pseudohypoparathyroidism; NESP, neuroendocrine secretory protein; PPHP, pseudopseudohypoparathyroidism; RT-PCR, reverse transcription–PCR; RACE, rapid amplification of cDNA ends.

To whom reprint requests should be addressed. e-mail: D.Bonthron@ed.ac.uk.

© 1998 by The National Academy of Sciences 0027-8424/98/0515475-6$2.00/0 PNAS is available online at www.pnas.org.
Although GNAS1 therefore is clearly imprinted, the maternal inheritance of PHP-Ia still suggests that there may be a gene product that is exclusively maternally rather than paternally derived. After a search aimed at identifying additional GNAS1 transcripts that may be imprinted, we now report the presence of an additional promoter and first exon, located 11 kb upstream of the XLαs exon. Remarkably, and in complete contrast to the XLαs exon, this region is paternally methylated, and the promoter is active only on a maternal allele. The new exon includes the entire coding region of the human homologue of the bovine neuroendocrine secretory granule protein NESP55 (17). Like the XLαs exon, the NESP55 exon is spliced onto GNAS1 exons 2–13. GNAS1 is thus a gene of bizarre complexity from the point of view of its imprinting. It encodes distinct paternally (XLαs), maternally (NESP55), and biallelically (Gαs) expressed proteins. The paternally and maternally active promoters are separated by only 11 kb. Furthermore, the two structurally unrelated paternal and maternal protein products, XLαs and NESP55, both appear to be involved specifically in the formation of secretory granules in neuroendocrine cells.

MATERIALS AND METHODS

cDNA Cloning. Reverse transcription–PCR (RT-PCR), 5′ rapid amplification of cDNA ends (RACE), and DNA sequencing (on both strands) were performed as described (15, 18). Mouse NESP55 clones were obtained by RT-PCR on RNA from the pancreatic islet cell line βTC3 (19), using primers dGAGCTCAGCTAATTACACAC (designed after inspection of mouse expressed sequence tag sequences apparently representing the murine homologue of bovine NESP55) and dTAGTCCACTGAACTGTTTCC (Gnas exon 5). The first 280 nt of the murine NESP55 cDNA sequence (GenBank accession no. AJ010163) were determined from the IMAGE cDNA clone 746837, and the remainder by direct sequencing of uncloned RT-PCR products (so as to eliminate PCR-generated errors). The human NESP55 sequence was determined from a BglII subclone of the PAC 309f20 and the remainder by direct sequencing of RT-PCR products from fetal muscle.

Analysis of Methylation in the NESP55 Region. A 4.5-kb BglII genomic subclone in pUC18 was KpnI + FspI-digested to yield a 2.6-kb probe fragment containing the entire NESP55 exon. Southern blotting was performed as described (15, 20).

Analysis of Allele-Specific GNAS1 Transcription. Exon 5 FokI polymorphism analysis of fetal RNA samples by RT-PCR was performed essentially as described (15). The downstream primer (exon 6, dCCGTGATGCTCATAGAAATTC) was 5′ fluorescent-labeled. The upstream primer was located in exon 1 (dCCATGGGCTGCCTCGGGAACA), the XLαs exon (dGGATGCTCCTGCTGGAACCG), or the NESP55 exon (dGGATGCTAATGGAGACGCCGT). PCR product (30 μl) was digested with 15 units of EcoNI (New England Biolabs) in a total volume of 200 μl for 2 hr at 37°C. To 100 μl of this digest was added 0.5 unit of FokI. After an additional hour at 37°C, heat inactivation at 65°C was performed, followed by ethanol precipitation and resuspension in 15 μl of Tris-EDTA, 2 μl of which was added to 10 μl of formamide buffer and 4 μl was loaded on a denaturing polyacrylamide gel run on an ALF sequencer (Pharmacia).

RESULTS

Cloning of Human and Murine NESP55-Encoding cDNAs. We used a combination of database searching and 5′ RACE to search for GNAS1 mRNA species derived from promoters other than those so far identified. First, all identifiable GNAS1 expressed sequence tag sequences were examined. Only one (derived from IMAGE cDNA clone 301896) had a 5′ end not previously accounted for.

Next, 5′ RACE was performed on human fetal muscle RNA, using a primer located in GNAS1 exon 5. Products were size-selected to be larger than expected for exon 1-containing mRNAs. A library of cloned RACE products then was screened for the presence or absence of the XLαs exon. One XLαs-negative RACE clone was > 1.3 kb in size and was characterized further. After sequencing, its 5′ end proved to include the entire unknown 5′ sequence of IMAGE clone 301896. However, unlike clone 301896, the RACE clone also contained an additional 0.9 kb of sequence inserted between the unknown 5′ sequence and exon 2. This additional region encodes the human homologue of the bovine NESP55 (17). Bovine NESP55 is encoded by mRNAs in which exons 2–13 of GNAS1 appear to make up the 3′ untranslated region. Just as for bovine NESP55, the stop codon of the human NESP55 ORF is 5′ to the splice junction with exon 2, so that none of the Gα coding region is included in the NESP55 ORF. Whether the exon 2–13 portion of the human (or bovine) NESP55 mRNA also can be translated, by use of internal ribosome entry and an alternative initiator codon within exon 2 (21), is at present unclear.

For comparative purposes, the mouse NESP55 homologue also was sequenced. An IMAGE cDNA clone (746837) apparently encoding the murine NESP55 homologue was identified by BLAST searching. The complete sequence of mouse NESP55 then was determined from this cDNA clone and from RT-PCR products derived from the βTC3 pancreatic islet cell-derived line (19).

Fig. 1 shows the complete human NESP55 sequence, as determined from both cDNA and genomic clones. The initiation codon indicated is the first ATG occurring after an in-frame upstream stop codon at nucleotide 685. It is also the only ATG within the long NESP55 ORF. The predicted NESP55 translation product has 245 aa (compared to 241 aa for bovine NESP55 and 253 aa for murine NESP55) and a molecular mass of 28,031 Da. Like its bovine homologue, human NESP55 is rich in glutamic acid (15%), arginine (11%), and proline (11%). The C-terminal two-thirds of the protein (amino acids 75–245) contains a number of repetitive motifs. One particularly repetitive region lies at the middle of the predicted protein (amino acid residues 102–143). This sequence is a dipeptide repeat in which every second residue is glutamic acid (E), of the form E→T/P). Most of the differences between human and bovine NESP55 sequences occur within the repetitive regions, whereas the first 50 aa are extremely highly conserved (Fig. 2). Analysis of the codon preference within the human NESP55 sequence suggests that selective pressure is imposed at the DNA and not just at the protein level. For example, of 37 Glu codons 30 are GAG and seven GAA (expected, 21 GAG, 16 GAA); of 14 Thr codons 12 are ACC, two ACT, 0 ACA (expected, five ACC, three ACT, two ACG, four ACA); of 12 Lys codons 10 are AAG and two AAA (expected, seven AAG, five AAA); of 24 Ser codons 20 are TCC or AGC (expected, nine TCC, six AGC); of 17 Leu codons nine are CTC (expected, three). We previously have described another tandemly repeated sequence within the coding region of the neighboring XLαs exon (15).

Genomic Organization of NESP55. PCR assays indicated that the entire NESP55-specific part of the cDNA might be encoded by a single exon, contained within the same PAC (P1 artificial chromosome) clone, 309f20, that contains the XLαs exon and exons 1–13 (15). After hybridization of NESP55- and XLαs-specific oligonucleotide probes to pulse-field gel electrophoresis-separated MluI and NotI digests, it appeared that the NESP55 and XLαs exons were separated by a small (11 kb) MluI fragment. Long-range PCR on the PAC clone as template DNA then was used to confirm that the NESP55 exon lies approximately 11 kb upstream of the XLαs exon (Fig. 3). A
4.5-kb BglII fragment containing the NESP55 exon was subcloned and used to confirm the complete sequence of the RACE clone.

Allele-Specific Methylation in the Region of the NESP55 Exon. A restriction map of the 4.5-kb BglII fragment containing the NESP55 exon is shown in Fig. 3. To assess methylation in this region, the 2.6-kb BglII–FspI NESP55 fragment was used as probe in Southern blotting. As previously, we used parthenogenetic lymphoblastoid cell DNA derived from the chimeric patient F.D. (20) as a tool to identify allele-specific methylation patterns (15). To analyze NESP55, parthenogenetic F.D. or normal female control lymphoblastoid cell line (LCL) DNA was double-digested with BglII and one of the methylation-sensitive enzymes XhoI, NgoMI, MluI, FspI, or SacII. Fig. 4 shows that for all of these enzymes, in normal LCL DNA the 4.5-kb BglII fragment is partially resistant to cleavage, indicating methylation of these sites on at least one allele (paternal). In the parthenogenetic LCL DNA, all of the enzymes, except MluI, cleave the BglII fragment to completion or nearly to completion. This finding is seen most clearly for the FspI and SacII sites, which lie close together a few hundred bp downstream of the NESP55 exon (Fig. 4, last four lanes of blot). Thus, these sites are unmethylated on the maternal allele, and methylated on the paternal one. This finding is in striking contrast to the opposite pattern at the XLαs exon only 11 kb away, where there is methylation only of the maternal allele (15) (see also Fig. 3 legend).

NESP55 Is Expressed Only from the Maternal Allele. We previously showed that while transcription from GNAS1 exon 1 is biallelic, XLαs-derived transcripts are exclusively paternal in origin (15). Given the opposite methylation patterns of the NESP55 and XLαs exons, we investigated whether they also show opposite patterns of allele-specific expression. For this purpose, RT-PCR was performed by using fetal tissues known to be informative for a FokI polymorphism in exon 5. The design of these experiments is illustrated in Fig. 5a. To make the GNAS1 RT-PCR specific for NESP55, XLαs, or exon 1, it is necessary to amplify across exon 3, which is alternatively spliced. The resulting doublet PCR product yields a complicated pattern after FokI digestion (Fig. 5b). To allow easier...
Restriction sites whose methylation status has been examined are indicated above by − or + to indicate unmethylated or methylated status, respectively, on maternal and paternal alleles. (Upper) The large map is complete only for NotI (N) and Ascl (A) sites. (Lower) On the expanded view, BglII (B), MluI (M), NgoMIV (Ng), PvuII (P), SacII (S), FspI (F), and XhoI (X) sites are indicated. The hatched bar indicates the probe used in Southern blotting in the present study. The methylation indicated for the XLαs exon and exon 1 is as described (15).

Uniform results were obtained, in a total of 19 tissue samples from six informative fetuses (three each of genotypes +/+ and −/−; the range of tissues was heart, kidney, muscle, spinal cord, brain, eye, intestine, lung, stomach, and gonad). In each case, NESP55 transcripts were exclusively derived from the maternal allele. Examples from two fetuses of opposite genotypes are shown in Fig. 5c, which also shows the contrast with XLαs expression in the same tissue sample. In fetus DM0909, in which the NESP55 products are cleaved by FokI, the XLαs products are FokI resistant. The opposite is true in fetus LF0101. In both fetuses, exon 1 products are partially cleaved, indicating their biallelic origin, as previously shown (15). For these exon 1 products, we also measured the peak areas (data not shown) and found no significant difference in the proportion of cleaved vs. uncleaved product between +/+ and −/− fetuses, suggesting that within the accuracy of our analysis, there is equal transcription of exon 1 from both maternal and paternal alleles.

**DISCUSSION**

Our results indicate that GNAS1 is extraordinarily complex in its pattern of allele-specific expression. The human homologue of the bovine protein NESP55 is encoded by an upstream exon of this gene, which is expressed exclusively from the maternal allele. Remarkably though, only 11 kb separate this maternally active promoter from the exclusively paternally active promoter that directs the synthesis of XLαs (15, 16).

NESP55 is a neuroendocrine secretory granule protein originally identified by screening cDNA libraries from bovine chromaffin cells with antibody against secretogranin II (17). On sucrose gradient centrifugation, NESP55 copurifies with secretogranin II. It is therefore noteworthy that rat XLαs (which, although a product of the same gene, shares no coding sequence with NESP55) also copurifies with secretogranin II on sucrose gradients (16). In addition, immunoreactive NESP55 displays a similar tissue distribution (particularly in adrenal medulla, pituitary, and brain) to that of XLαs mRNA (16, 17). These similarities suggest that GNAS1 may encode more than one protein with a function in the regulated secretory pathway in neuroendocrine tissues. We find, however, that both NESP55 and XLαs mRNAs can be detected by RT-PCR in a wide range of human fetal tissues, suggesting that low-level expression of these transcripts may be more widespread than suggested by previous studies.

The close proximity of the NESP55 and XLαs exons further suggests the possibility of coordinate regulation of these two proteins' expression, albeit in a reciprocal, mutually exclusive allele-specific manner. Given the genomic architecture of this region of GNAS1, one specific possibility requiring scrutiny will be the involvement of antisense transcripts in regulating NESP55 and XLαs expression. In the murine Igf2r gene, a paternally expressed antisense transcript originating from a promoter within intron 2 appears to suppress transcription in the sense direction from the Igf2r structural gene promoter in cis with it (22). A similar mechanism might be important in GNAS1 regulation. If, for example, transcription from the paternally active XLαs promoter were bidirectional, the antisense transcript might maintain the repressed state of the upstream NESP55 promoter on the paternal allele. We have preliminary evidence for the existence of antisense transcripts in this region of GNAS1, but considerable further work will be...
FIG. 5. RT-PCR analysis of GNAS1 expression in human fetal RNA. Genotypes are indicated as follows: $^{+m/-p}$ indicates FokI cleavage site on maternal but not paternal allele. (a) Experimental design. Coding regions are shaded dark gray, and the alternatively spliced exon 3 is light gray. Amplification is performed across the polymorphic FokI site, by using an upstream primer specific for NESP55, XLαs, or exon 1, and a common downstream PCR primer that is fluorescein labeled. In each case, a doublet PCR product, representing transcripts with and without exon 3, is generated (see b). Digestion with FokI alone therefore produces multiple bands (b). In contrast, EcoNI digestion reduces the doublet to a single 155-nt, fluorescein-labeled cleavage product, and the FokI restriction fragment length polymorphism then is scored by a second digest. (b) RT-PCR products analyzed on an ethidium bromide-stained agarose gel. Analysis of RNA from muscle (DM0909) and kidney (LF0101) of two informative fetuses of opposite genotypes is shown. Each pair of lanes is uncut (Left) and FokI-digested (Right). In fetus DM0909, all of the NESP55-specific but none of the XLαs-specific products are cut. Exon 1-specific products, in contrast, are partially cleaved. For fetus LF0101, an opposite picture is seen. The results indicate that NESP55 transcripts are maternally and XLαs transcripts paternally derived, whereas the Gαs-encoding exon 1-containing transcripts are biallelic in origin. (c) Fluorescent analysis on an ALF sequencer (Pharmacia). Tissues from the same two informative fetuses, DM0909 and LF0101, are shown. Each exon-specific product (NESP55, XLαs, or exon 1) is shown as a triplet of traces; uncut, EcoNI-digested, and EcoNI+FokI-digested, from top to bottom. In each case, the doublet of peaks resulting from exon 3 alternative splicing resolves to a single 155-nt peak on EcoNI digestion (position indicated by arrows). The shift to a 74-nt peak on FokI digestion then indicates the proportion of transcripts derived from the FokI(+) allele.
required to determine whether such a mechanism is in fact important.

The arrangement of the oppositely imprinted NESP55 and XLαs exons also raises an interesting problem of splicing specificity. Maternally derived nascent transcripts from the NESP55 promoter must contain all of the exons NESP55, XLαs, and 1–13. If any splicing of XLαs to exon 2 occurred in these transcripts, it would show up in RT-PCR experiments as maternally expressed XLαs (as they are for Gsα function). Yet, no severe neuroendocrine phenotype is apparent in these patients. Specifically targeted knockouts of NESP55 and XLαs, will need to be addressed through gene targeting experiments in the mouse.

Some insights into NESP55 function might be obtained from existing mouse models. A neo cassette inserted into exon 2 of \( \text{Gnas} \) results in a severe phenotype on transmission through the parent of either sex (14). On maternal transmission, oedema and major neurological abnormalities occur with 80% lethality. If the exon 2 insertion destabilizes NESP55 mRNA, then some of these additional features (not seen in PHP-Ia) might be attributable to NESP55 deficiency.

Similar questions may be raised about the function of the paternally encoded XLαs. One issue requiring particular scrutiny is that PHP patients with paternally derived mutations in exons 2–13 should lack XLαs function altogether, if such mutations are null for XLαs (as they are for Gsα function). Yet, no severe neuroendocrine phenotype is apparent in these patients. Specifically targeted knockouts of NESP55 and XLαs exons in the mouse might resolve such questions, but these experiments will require careful design if secondary effects on the imprinting of the neighboring, oppositely expressed exon are to be avoided.

Finally, additional questions recently have been raised about the imprinting of the 20q13.3 chromosomal region by linkage studies of PHP-Ib. In this condition, hitherto believed to be of distinct etiology from PHP-Ia, there is isolated parathyroid hormone resistance, but other endocrine functions, as well as Gsα levels, are normal, and Albright osteodystrophy is absent. Four PHP-Ib families show linkage to 20q13.3, with evidence suggesting that as for PHP-Ia, there is a requirement for maternal transmission for manifesting the condition (25). Only a single recombinant excluded \( \text{Gnas} \) from the PHP-Ib region, and because this was with an intragenic marker, it remains quite possible that a structural or regulatory mutation affecting an alternative product of the complex \( \text{Gnas} \) locus could account for PHP-Ib. Further extensive genomic analysis of the region around \( \text{Gnas} \) is needed to address this and other questions.

This work was supported by Wellcome Trust Project Grant 053701 to D.T.B. and by the Ludovici Bequest to the University of Edinburgh. V.M. is supported by a scholarship from CONACYT (Mexican National Council for Science and Technology).