Volume 16 Number 20 1988 Nucleic Acids Research

CpG islands of the X chromosome are gene associated

Myriam Alcalay+ and Daniela Toniolo[1]*

International Institute of Genetics and Biophysics, CNR, Via Marconi 10, 80125 Naples and [1]Istituto di Genetica Biochimica ed Evoluzionistica, CNR, Via Abbiategrasso, 207 Pavia, Italy

Received July 26, 1988; Revised and Accepted September 15, 1988 Accession no. X12458

ABSTRACT

Unmethylated CpG rich islands are a feature of vertebrate DNA: they are associated with housekeeping and many tissue specific genes. CpG islands on the active X chromosome of mammals are also unmethylated. However, islands on the inactive X chromosome are heavily methylated. We have identified a CpG island in the 5' region of the G6PD gene, and two islands forty Kb 3' from the G6PD gene, on the human X chromosome. Expression of the G6PD gene is associated with concordant demethylation of all three CpG islands. We have shown that one of the two islands is in the promoter region of a housekeeping gene, GdX. In this paper we show that the second CpG island is also associated with a gene, P3. The P3 gene has no homology to previously described genes. It is a single copy, 4 kb gene, conserved in evolution, and it has the features of a housekeepin two genes is within the CpG island and that sequences in the islands have promoter function.

INTRODUCTION

The dinucleotide CpG is rare in vertebrate DNA occurring at only one-fifth of the frequency expected from base composition. However, about 1% of the genome is rich in CpGs: in this fraction of DNA the CpGs are not methylated and can be cleaved by methyl-CpG sensitive restriction enzymes, hence the definition of these regions as HpaII tiny fragment (HTF) fraction (1). Since CpGs in bulk DNA are often methylated at cytosine, it has been proposed that the rarity of CpG is caused by the high frequency of deamination of 5-methylcytosine to give thymine. On the other hand, HTF regions have a high CpG content because they are not methylated. Cloning and sequencing of a number of these regions have given insight into the properties and the genomic distribution of the HTF fraction. It occurs as islands of clustered CpGs, 1-2 kb long, dispersed in the genome. There are about 30.000 islands per haploid genome in mammals, which could be regularly

spaced along the genome; however, variation from this value is great (2). No extensive sequence homology has been described between different HTF islands, their common feature being the clustered CpGs and the high G+C content. Many CpG islands are associated with genes: all sequenced housekeeping genes have islands at their 5' ends, but also many tissue-specific genes have CpG islands (3, 4). A comparison of the approximate number of islands (about 30,000) with the calculated number of genes (20-50,000) suggests that a high proportion of islands will be gene-associated.

CpG islands are not methylated: an exception are CpG islands on the inactive X chromosome of mammals. The HPRT, PGK and G6PD genes have been studied in detail: the CpG clusters in their 5' regions were methylated on the inactive X but unmethylated on the active X and demethylated when the locus on the inactive X is reexpressed (2, 5, 6, 7). Methylation studies of X-linked genes and several random sequences have shown that no global methylation differences exist between active and inactive X chromosome DNA and the inactive state of X-linked genes has been correlated only with methylation of CpG islands (2, 8, 9). These findings have suggested that CpG islands, regularly spaced along the X chromosome, could be control elements for the maintenance of X chromosome inactivation (7).

Forty kb downstream from the human G6PD gene we have identified two HTF islands whose behaviour with respect to X chromosome inactivation is identical to the one in the 5' region of the G6PD gene (2, 10). We have shown that one of the two islands is at the 5' of a housekeeping gene, which we have called GdX (11). We now show that also the second CpG island is at the 5' end of a gene, P3. The P3 gene has no homology to previously described genes. We also show that initiation of transcription of the two neighbouring GdX and P3 genes is within the islands, and that sequences in the islands have promoter function.

MATERIALS AND METHODS

Isolation and characterization of cDNA clones

cDNA clones were isolated from cDNA libraries from placenta (12) or human teratocarcinoma cells (a gift from Dr. J. Skowronski, Cold Spring Harbor

Laboratory, CSH, USA). The hybridization probe for the initial screening was the genomic probe pGdP3, shown in Fig. 1. In subsequent screenings to obtain full length cDNAs, subclones or fragments of previously isolated cDNAs were used. Small scale plasmid or phage preparations were analyzed by restriction mapping by standard procedures (13). The genomic and cDNA clones from lambda libraries were subcloned in the plasmids pEMBL8 or pUC18 for further analysis.

DNA sequencing

The sequence of the genomic DNA was determined by the procedure of Maxam and Gilbert (14) or by the dideoxy method modified to sequence directly from the plasmid pUC18 (15).

The sequence of the cDNA clones was determined by the dideoxy method. The sequence in Fig. 2 was obtained by sequencing both strands from genomic clones and more than one cDNA clone from different libraries.

Southern blot

DNA (10-20 ug) from tissues or from cultured cells, prepared as described (10), was digested with the restriction enzymes indicated, run on 1% agarose gels, transferred to nylon filters (Zetabind, AMF-CUNO) and hybridized at 65°C in 5X SSPE, 1% SDS, 2X Denhardt's and 100 ug/ml herring sperm DNA for 18-20 hrs (13). Filters were washed at 65°C in 2X SSC and 0.2X SSC (stringent washing). Hybridization of human probes with DNA from species other than mammals were washed in 2X SSC and 1X SSC at 60°C.

Northern blot

HeLa cell total RNA in Fig. 3, lanes 2 and 12, was prepared in our laboratory by the guanidine-HCl method (13). Total RNA from HL60 cells (16) was isolated by cell lysis in 4M guanidine-thiocyanate and sedimentation through 5.7M CsCl (12). RNA from HeLa (Fig. 3, lane 6) and choriocarcinoma cells (JEG) (17) were a gift by Dr. J. Chou (NIH, Bethesda, Md, USA). RNA from human teratocarcinoma cells NTera2/D1 (18), PA-1, Ca-Ma and COS-1 (19) were a gift by Dr. J. Skowronski. Human fibroblast RNA was a gift from Dr. A. Simeone (IIGB, Naples, Italy), and normal human thyroid RNA was a gift from Dr. A. Fusco (II Medical School, Naples, Italy). All the RNAs were total

cellular RNAs except the RNA from NTera2/D1 teratocarcinoma cells, which was polyA+. 5ug samples were electrophoresed through a 1.5% agarose gel containing formaldehyde (13). The RNA was transferred to nylon filters and hybridized with the probes indicated using the same hybridization conditions described for Southern blots.

S1 mapping

To map the transcription initiation of the P3 gene a 450 bp fragment was obtained by AvaI digestion of a genomic subclone (pEB9R40). The fragment was 5' end-labelled with T4 polynucleotide kinase and gamma-^{32}P ATP (13). After digestion with NaeI, fractionation in 5% polyacrylamide gel and electroelution, the 260 bp 5' end-labelled fragment was hybridized in Formamide-Pipes buffer (40mM Pipes pH 6.4, 1mM EDTA, 400mM NaCl, 80% Formamide) with 50 ug of total HeLa RNA at 55°C, overnight. S1 nuclease was added at a concentration of 400 or 1000 U/ml in S1 buffer (0.28M NaCl, 0.05M Na Acetate pH 4.6, 4.5M ZnSO4, 20ug/ml single strand DNA) and incubated at 30°C or 37°C for 2 hours.

To map transcription initiation of the GdX gene a 650 bp fragment was obtained by HindIII/PvuII digestion of a genomic subclone (pMLGd15), ^{32}P-labelled as above. After digestion with TaqI, fractionation in 5% polyacrylamide gels and electroelution, the labelled fragment was hybridized with 50 ug of total HeLa RNA at 52°C, overnight. S1 nuclease was added at the same concentrations as above and incubated for 2 hours at 28°C or 35°C.

Use of higher temperature of S1 digestion would decrease the total amount of protected fragments. Protected fragments were fractionated on 8% polyacrylamide/urea sequencing gel (13) and exposed to autoradiographic film.

Primer extension

PolyA+ RNA from HeLa cells was selected by chromatography on oligo(dT) cellulose (13). Two oligonucleotides were synthetized. A 25 bp oligonucleotide ATCGCGGAGCAGGGTCGGGGCCAGA (oligo-P3) complementary to nucleotide 74 to 98 of the P3 cDNA and 20 bp oligonucleotide AGCTGCTTCAGGCTGGACAC (oligo-GdX) complementary to nucleotide 100 to 120 of the GdX cDNA. The oligonucleotide were 5' end labelled with T4 polynucleotide

kinase and gamma-^{32}P-ATP and purified by fractionation in a 20% polyacrylamide/urea gel, and elution of the radioactive band at 37°C, overnight. One pmole of each labelled oligonucleotide was hybridized to 20 ug of HeLa polyA+ RNA in Pipes-NaCl buffer (10mM Pipes, pH 6.9, 400mM NaCl, 1mM EDTA). Hybridization was for 18 hours at 60°C for oligo-P3 and at 59°C or 55°C for oligo-GdX. The extension reaction was performed with 20U of AMV-reverse transcriptase in 30 ul of 50mM Tris-HCl pH 8.3, 5mM $MgCl_2$, 50mM KCl, 1mM DTT, 1mM dNTP at 42°C for 2 hours. The products of the reaction were fractionated on 8% polyacrylamide/urea sequencing gel and exposed to autoradiographic film for 3-6 days.

CAT assay

Different portions of the P3 and GdX CpG islands (shown in Fig. 7A and B) were cloned in the plasmid pEMBL8CAT (gift of Drs. G. Ciliberto and R. Cortese, EMBL, Heidelberg). One pmole of each CAT-plasmid was transfected in HeLa cells by the $CaPO_4$ procedure. 48 hrs later extracts were prepared and analyzed (20).

RESULTS

Isolation of cDNAs and mapping of exons

Using a single copy genomic probe (pGdP3) from the region between the two CpG islands identified downstream from the G6PD gene we screened cDNA libraries from placenta and teratocarcinoma cells. From both libraries we obtained full length cDNA clones.

Hybridization of the P32-labelled cDNA clones to restriction digests of DNA from previously isolated lambda phages (11) showed that the P3 gene spans about 4 kb of genomic DNA, less than 1 kb upstream from the neighbouring GdX gene. Sequencing the cDNA and the genomic DNA (Fig. 2) has allowed the precise location of exons, of the only intron and of the exon-intron boundaries. The sequence of 4739 nucleotides was obtained using both chemical and enzymatic sequencing methods from the different cDNA clones, shown in Fig. 1 and subclones of the corresponding genomic DNA.

From this analysis we determined that the P3 gene is divided into 2

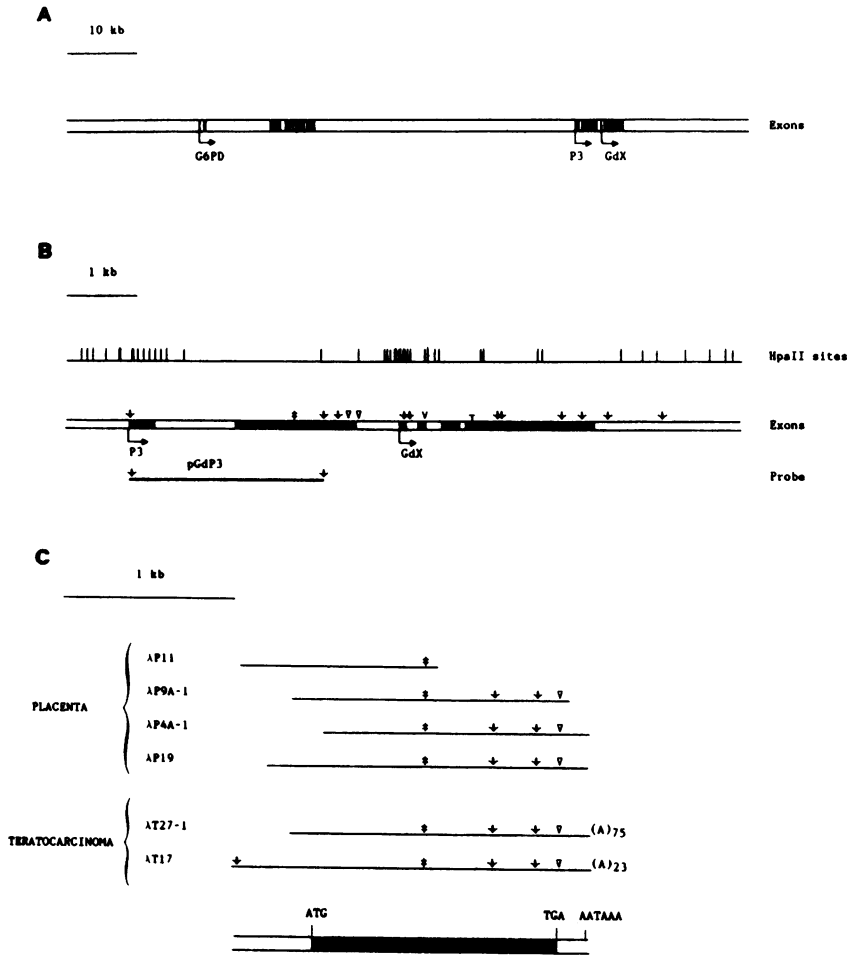

Fig. 1. A: Relative location of the G6PD, P3 and GdX genes. B: Location of the P3 and GdX genes relative to the two CpG islands. Vertical bars are HpaII sites. Blacked in regions are exons. Arrows below the map indicate the direction of transcription. Symbols above the map are restriction enzyme sites: : Pst I; : XhoI; : Hind III; : EcoRI; : KpnI. C: Map of some of the P3 cDNAs isolated. Below is a schematic representation of the P3 mRNA: in black is the coding region. Restriction enzyme symbols are as in B.

exons. The first exon is small (350 bp), the second is 1806 bp long. One polyadenylation site (AATAAA) can be recognized at the 3' end followed by a polyA tail in some of the cDNAs, 16-17 residues downstream from the polyadenylation site (Fig. 1). The sequences at the 5' and 3' intron

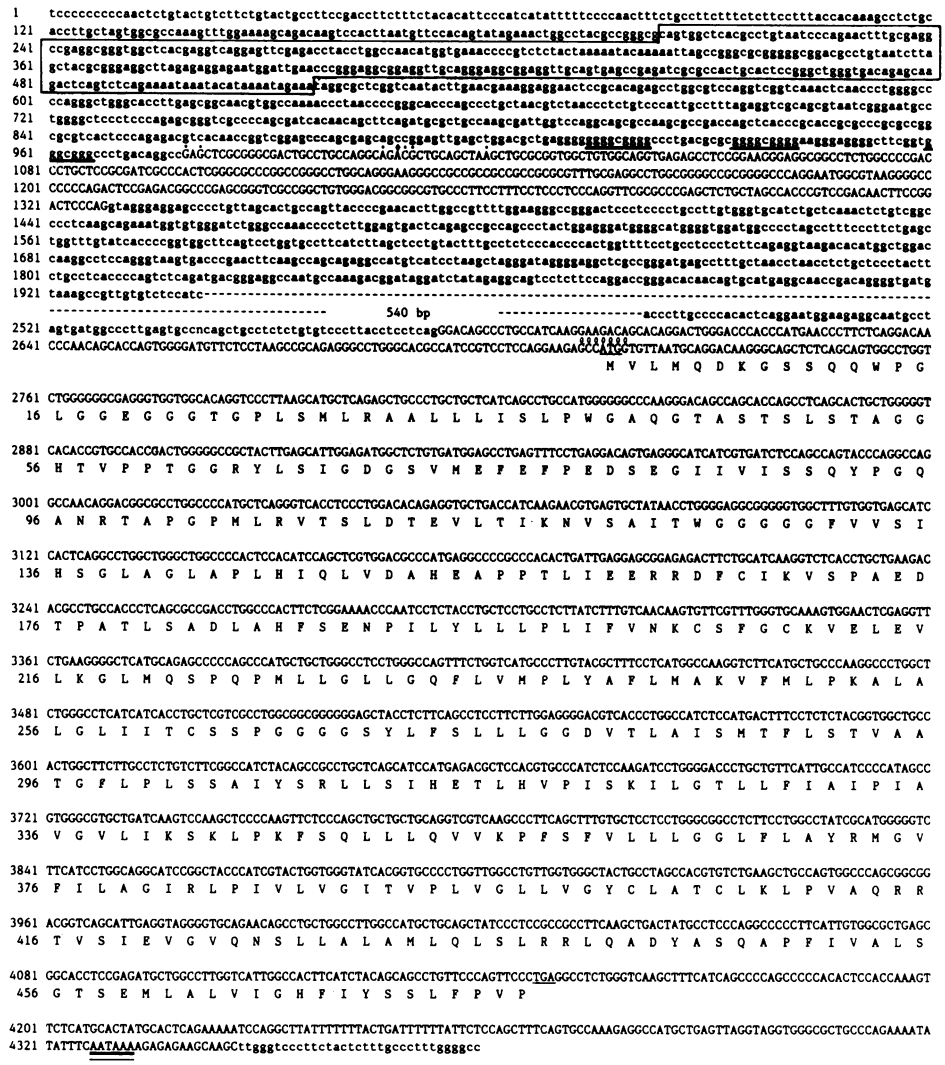

Fig. 2: Nucleotide sequence of the P3 gene. The cDNA sequence is in uppercase letters; introns and flanking regions are in lowercase letters; ATG and TAG codons are underlined, the GCCATGG, at the initiation of translation (22), is indicated by white dots over the sequence. The aminoacids encoded by the cDNA are shown by one-letter symbols below the nucleotide sequence. Black dots over the sequence indicate minor sites of initiation of transcription; the black diamond indicate the major site of initiation of transcription. Binding sites for transcription factor Sp1 are underlined. The polyadenylation signal is indicated by a double underlining. The large boxed sequence in the 5' flanking region is an Alu sequence.

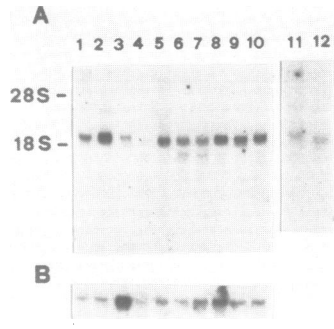

Fig. 3: Northern blot. A: Hybridization of the pP9A-1 cDNA to RNA from various sources. Lane 1: human fibroblasts; lanes 2, 6 and 12: Hela cells; lane 3: NTera 2/D1 (undifferentiated human teratocarcinoma cells); lane 4: HL60 (undifferentiated human myeloid cells); lane 5: human T lymphocytes; lane 7: Ca-Ma (human mammary carcinoma cells); lane 9: PA-1 (human neuroblastoma cells); lane 8: COS-1 (monkey fibroblasts transformed with replication-defective SV40); lane 10: JEG (human choriocarcinoma cells); lane 11: normal human thyroid. B) Hybridization of a chicken ß-actin cDNA probe (27) to the same filter shown in A.

boundaries are in agreement with the consensus sequence for exon-intron junctions of eukaryotic genes (21).

Southern and Northern hybridization

Hybridization of P32-labelled cDNA or genomic clones to Southern blots showed that the P3 gene is unique in the human genome: only the bands expected from restriction analysis of the genomic clones appear (not shown).

In Northern blots the P3 gene probes hybridize to one RNA species of 2.1 kb expressed at similar levels in all cell types analyzed (Fig. 3A), with the exception of the RNAs from teratocarcinoma and HL60 cells, which give a fainter hybridization signal compared to the rest of the RNAs analyzed. The amount of P3 RNA in teratocarcinoma cells does not change following differentiation (not shown).

The P3 gene is conserved

The pGdP3 probe as well as a cDNA corresponding to the coding region of the P3 gene (see below) was hybridized to Southern blots of DNA from several animal species: monkey, mouse, rat, bovine, horse, chicken, Xenopus laevis, a reptile (T. sex lineatus), D. melanogaster and the yeasts S.

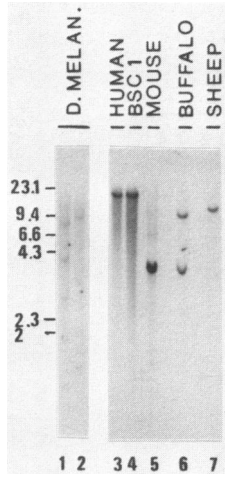

Fig. 4: Hybridization of the pP9A-1 cDNA to DNAs from the fruit fly D. melanogaster, (lanes 1 and 2), human leukocytes (lane 3), monkey BSC1 cells (lane 4), mouse liver (lane 5), buffalo (lane 6) and sheep (lane 7) leukocytes. DNAs (10 ugr) were digested with EcoRI (lane 1, 3-7) or BglI (lane 2). Markers (kb) were HindIII fragments of phage lambda DNA.

cerevisiae and H. capsulatum. The hybridization to some of the species tested is shown in Fig. 4: the probes hybridize to single or a few bands in all the DNAs tested at rather stringent hybridization washing conditions (1xSSC at 60°C), suggesting that the P3 sequence has been considerably conserved during evolution.

Sequence analysis of the cDNA

The nucleotide sequence of the P3 cDNA was compared with the Gene Bank (tape release 48): no homologies were found with published nucleotide sequences. A search for ORFs (Open Reading Frames) in the P3 cDNA showed a long ORF starting from the ATG in position 493 and ending with the stop TGA codon in position 1923 of the cDNA (Fig. 2). The sequence GCCATGG surrounding the ATG, is in good agreement with the eukaryotic consensus sequence for translation initiation sites (22). The ORF is entirely located in the second exon. Searching the protein NBRF Bank (Release 11) for homologies with the P3 ORF did not show homology with any sequenced protein.

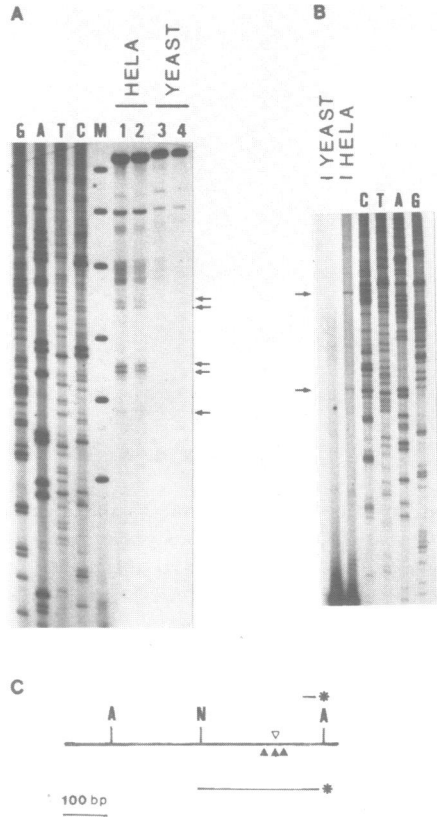

Fig. 5: Analysis of the 5' end of the P3 mRNA. A: S1 mapping with the NaeI(N)-AvaI(A) fragment, labelled at the AvaI site, and Hela cell total RNA (lanes 1 and 2) or control yeast tRNA (lanes 3 and 4). GATC is a sequence ladder of a P3 genomic fragment (pEB9R40). M is HpaII digested pUC18: fragments are 67, 94, 110, 147, 190 and 242 bp. Lanes 1 and 3: S1 digestion was at 30°; lanes 2 and 4: S1 digestion at 37°. B: Primer extension analysis with the oligo nucleotide oligo-P3 (see Materials and Methods) hybridized to yeast tRNA (lane 1) or Hela cell poly A+ RNA (lane 2). GATC is the same sequence ladder shown in A. C: Probes used for the analysis above or below a genomic restriction map. The white arrow above the line indicate the position of the primer extended product. The black arrows below the line indicate the position of the S1 protected fragments.

Mapping the 5' end of the P3 gene

The first exon of the P3 gene is very G+C rich and entirely contained within the HTF island. To confirm that transcription initiation sites of the P3 gene are within the HTF islands we mapped transcription initiation by S1

mapping and primer extension experiments.

A genomic fragment of 260 bp from a NaeI to an AvaI site, 5' end labelled at the AvaI site, was subjected to S1 nuclease digestion after hybridization with HeLa cell total RNA. Two protected fragments were more prominent: 93 and 91, nucleotides long (Fig. 5A). Several minor protected fragments of 85, 95, 94, 124 and 126 nucleotides were also observed. Variation of temperature and enzyme concentration during S1 nuclease digestion only affected the overall intensity of the bands, not the intensity relative to one another. However, in order to obtain specific protection the temperature of S1 digestion had to be raised to 30 degrees. This might be due to the high G+C content of the fragments. In the upper part of the gel many fragments in all lanes are due to unspecific protection and they might cover additional specific sites.

To confirm the results of S1 mapping experiments an oligonucleotide complementary to the cDNA from nucleotide 74 to 99 was hybridized to polyA+ RNA from HeLa cells and extended with AMV reverse transcriptase. An extended product was observed having the same length as one of the two major S1 protected fragments (Fig. 5B). Primer extended products corresponding to the minor S1 resistant fragments were not detected, probably due to the lower sensitivity of the primer extension assay. A 56 bp longer extension product was also synthesized, that did not correspond to a S1 protected band. It may be an indication of an additional site of transcription initiation further upstream.

Mapping the 5' end of the GdX gene

By S1 protection and primer extension we mapped the initiation of transcription of the GdX gene which also appears to be within a CpG island (12).

The results of this analysis are show in Fig. 6. Several sites of initiation of transcription can be mapped at 1, 12 and 24 nucleotides from the first nucleotide of the cDNA extending more 5' (12) (Fig. 6). The analysis of the GdX gene, as in the case of P3, was hampered by the high G+C content of this region. For this reason we cannot exclude the existence of minor sites of initiation of transcription.

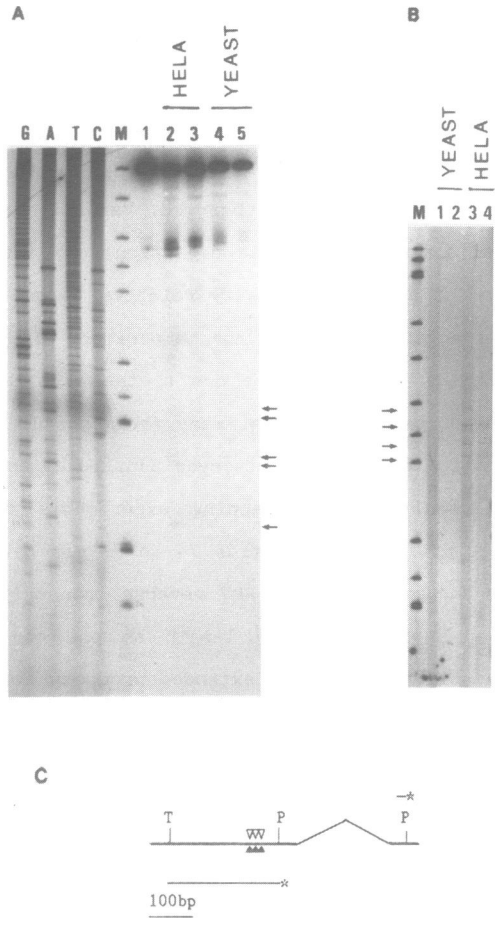

Fig. 6: Analysis of the 5' end of the GdX mRNA. A: S1 mapping with the TaqI(T)-PvuII(P) fragment, labelled at the PvuII site and Hela cell total RNA (lanes 2 and 3) or yeast tRNA (lanes 4 and 5). Lane 1: labelled fragment only. GATC is a sequence ladder of a P3 genomic fragment, M is HpaII digested pEMBL8, fragments are 26, 34, 60, 67, 78, 110, 125, 147, 190 and 242 bp. Lanes 2 and 4: S1 digestion was at 28°C; lanes 3 and 5: S1 digestion was at 35°. B: Primer extension analysis with the oligonucleotides oligo-GdX (see Materials and Methods) hybridized to yeast tRNA (lanes 1 and 2) or Hela cell polyA+ RNA (lanes 3 and 4). Lanes 1 and 3: hybridization at 55°C; lanes 2 and 4: hybridization at 59°C. M is HpaII digested pEMBL8. C: Probes used for the analysis. Above the genomic restriction map is the position of oligo-GdX. The white arrows indicate the position of the primer extended products. Below the genomic restriction map is the probe used for S1 mapping: the black arrows indicate the protected fragments.

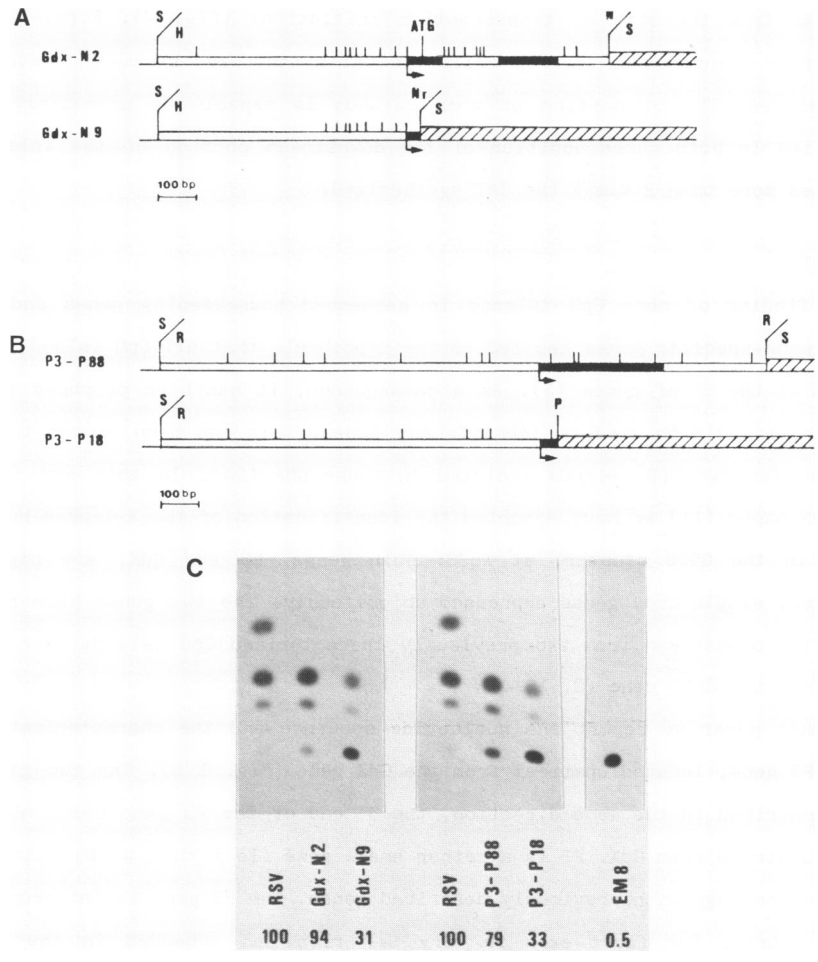

Fig. 7: CAT assay. A and B are restriction maps of the CpG islands at the 5' region of the GdX (A) and P3 (B) genes fused to a CAT gene (etched area). S: SmaI; H: HindIII; Nr: NruI: N: NotI; R: RsaI; P: PstI. Bars are HpaII sites. Blacked in areas are exons. Arrows indicate initiation and direction of transcription. C: CAT activity of the control RSV-CAT (RSV; 28) and pEMBL8CAT plasmids, and of the constructs shown in A and B. Numbers are percent chloramphenicol acetylation, compared to RSV-CAT.

CpG islands have promoter function

To determine if CpG islands can function as promoters, we constructed recombinant plasmids carrying 5' to a CAT gene, without promoter and enhancer, the entire P3 or GdX islands or the 5' portion of the island

(upstream from the site of transcription initiation) (Fig. 7). Forty-eight hours after transfection in HeLa cells the CAT synthesized by each construct was measured. The 5' portion of both islands is capable of promoting CAT synthesis. In both cases addition of the downstream portion of the islands increased more than 3 times the CAT synthesized.

DISCUSSION

The finding of many CpG islands in sequenced housekeeping genes and in some tissue-specific genes has led to the hypothesis that all CpG islands are located at the 5' of genes (3). As a consequence, it has been proposed that CpG islands could be used to identify new genes. This has been proven to be the case for two CpG islands isolated in mouse DNA (24). In this and in a previous paper (11) we have reported the identification of two X-linked human genes, in the G6PD cluster, at Xq28. Both genes, P3 and GdX, are highly conserved, single copy genes expressed ubiquitously. The two genes have been identified downstream from two previously characterized CpG islands, located 40 kb from the G6PD gene (2, 10).

In this paper we report the nucleotide sequence and the characterization of the P3 gene, located upstream from the GdX gene (Fig. 1 B). The two genes are transcribed in the same direction, the 3' end of the P3 gene being about 700 bp upstream from GdX. P3 is a rather small gene (less than 4 kb) with no sequence homology with previously described genes. The P3 gene is interrupted by one intron. The first exon is very G+C rich and embedded in the CpG island. In the second exon we have identified a long ORF of 1430 bp, starting with the ATG in position 493 of the cDNA sequence (underlined in Fig. 2). The first stop codon in frame with the ATG leaves a short 3' non-coding region. The ORF codes for a protein of 477 aminoacids, whose sequence has shown no homology with previously identified proteins. The presence of a long ORF, the ubiquitous expression and the conservation of the P3 sequence in distant animal species suggest that the P3 gene codes for a protein with housekeeping functions. It is not unlikely that the P3 gene corresponds to one of the many genes identified on the human X chromosome, at Xq28, by the study of sex-linked diseases.

The whole DNA region sequenced has a 62% G+C content, that increases in the CpG island. From nucleotide 700 to nucleotide 1,500 the G+C content is 71%. In the same region the CpG content is 10%, close to the GpC (13,1%). The CpGs in the whole sequence, or in the cDNA, are 4,3% (GpCs are 9.2-10.4%). The 5' end of the P3 and GdX trascripts map close to the most G+C rich region of the islands; the first exon and part of the intron of the P3 gene, and the first and second exons of the GdX gene are within the islands. Since this appears to be a common feature of HTF islands (2, 4, 23), the whole island might be important for the expression of the neighbouring genes. Our data indicate that this is the case. The 5' portions of the islands, upstream to the major site of initiation of transcription can promote transcription of the CAT gene after transfection in Hela cells, in the absence of TATA or CAT boxes. However the increase in the CAT synthesized in the constructs carrying the whole islands is in agreement with the hypothesis that sequences of the islands other than the 5' flanking are important for transcription.

HTF islands on the X chromosome are gene associated and their general organization and sequence do not appear different from CpG islands on autosomes. The correlation between methylation of CpG islands on the inactive X chromosome and gene silencing suggests that clustering of CpGs does not merely reflect a passive exclusion of methylases from promoters of housekeeping genes. Instead, conservation of CpG-rich regions might be part of a regulatory mechanism involving constitutive switching on and off of chromosomes or large chromosomal regions. In this respect, X chromosome inactivation could be related to other phenomena involving inactivation of large genomic regions, like genomic imprinting (24, 25, 26). Lack of methylation of these regions could be a requisite for the interaction of housekeeping promoters with cellular factors and for their constitutive expression. Methylation of CpG islands could abolish this interaction and, as a consequence, cause constitutive gene silencing. The role of CpG clusters in this kind of gene regulation could be exclusively the stabilization and maintenance of a chromatin state determined by a primary event different for their methylation. However, we cannot exclude, as it has been proposed for X

inactivation (7), that the methylation state of HTF islands could directly affect the chromatin structure of neighbouring regions.

ACKNOWLEDGEMENTS

We thank Dr. L. Lanfrancone for help with the CAT assay, M. Terracciano, C. Salzano and C. Sole for technical assistance and C. Bagarotti for typing the manuscript. This work was supported by grants from Progetto Finalizzato CNR Ingegneria Genetica e Basi Molecolari delle malattie Ereditarie, and from the Associazione Italiana per la ricerca sul Cancro. M.A. was the recipient of an ENEA fellowship.

+Present address: Laboratorio di Biologia Molecolare, Istituto di Clinica Medica I, Policlinico Monteluce, 06100 Perugia, Italy

*To whom correspondence should be addressed

REFERENCES

1) Bird, A.P. (1986) Nature 321, 209-213.
2) Toniolo, D., Martini, G., Migeon, B.R. and Dono, R. (1988) EMBO J. 7, 401-406.
3) Bird, A.P. (1987) TIG 3, 342-347.
4) Gardiner-Garden, M. and Frommer, M. (1987) J. Mol. Biol. 196, 261-282.
5) Wolf, S.F., Dintzis, S., Toniolo, D., Persico, G., Lunnen, K., Axelman, J. and Migeon, B.R. (1984) Nucleic Acids Res. 12, 9333-9348.
6) Yen, P.H., Patel, P., Chinault, A.C., Mohandas, T. and Shapiro, L.J. (1984) Proc. Natl. Acad. Sci USA, 81, 1759-1763.
7) Riggs, A.D., Singer-Sam, J. and Keith, D.H. (1985) In Razin, A. and Cantoni, G.L. (eds.), Biology and Biochemistry of DNA methylation. Alan R. Liss, New York.
8) Wolf, S.F. and Migeon, B.R. (1985) Nature 314, 467-469.
9) Lindsay, S., Monk, M., Holliday, R., Huschtscha, L., Davies, K.E., Riggs, A.D. and Flavell, R.A. (1985) Ann. Hum. Genet. 49, 115-127.
10) Toniolo, D., D'Urso, M., Martini, G., Persico, M., Tufano, V., Battistuzzi, G. and Luzzato, L. (1984) EMBO J. 3, 1987-1995.
11) Toniolo, D., Persico, M.G. and Alcalay, M. (1985) Proc. Natl. Acad. Sci. USA, 85, 851-855.
12) Persico, M.G., Viglietto, G., Martini, G., Toniolo, D., Paonessa, G., Moscatelli, C., Dono, R., Vulliamy, T., Luzzato L. and D'Urso M. (1986) Nucleic Acids Res. 14, 2511-2522.
13) Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular cloning. A Laboratory Manual. Cold Spring Harbor Laboratoy Press, Cold Spring Harbor, New York.
14) Maxam, A.M. and Gilbert, W. (1980) Methods Enzymol. 65, 499-560.
15) Hattory, M., Hidaka, S. and Sasaky, Y. (1985) Nucleic Acids Res. 13, 7813-7827.

16) Collins, S.J., Gallo, R.C. and Gallagher, R.E. (1977) Nature (London) 270, 347-349.
17) Kohler, P.O. and Bridon, W.E. (1971) J. Clin. Endocrinol. Metab. 32, 683-690.
18) Andrews, P.W. (1984) Dev. Biol. 103, 285-293.
19) Skowronski, J. and Singer, M.F. (1985) Proc. Natl. Acad. Sci. USA, 82, 6050-6054.
20) Rosenthal, N. (1987) Methods Enzymol. 152, 704-720.
21) Mount, S.M. (1982) Nucleic Acids Res. 10, 459-472.
22) Kozak, M. (1986) Cell 44, 283-292.
23) Lavia, P., Macleod, D. and Bird, A. (1987) EMBO J. 6, 2773-2779.
24) Reick, W., Collick, A., Norris, M.L., Barton, S.C. and Surani, M.A. (1987) Nature, 328, 248-251.
25) Sapienza, C., Peterson, A.C., Rossant, J. and Balling, R. (1987) Nature, 328, 251-254.
26) Swain, J.L., Stewart, T.A. and Leder, P. (1987) Cell 50, 719-727.
27) Cleveland, D.W., Lopata, M.A., Mac Donald, R.J., Cowan, N.J., Rutter, W.J. and Kirschener, M.W. (1980) Cell 20, 95-105.
28) Gorman, C.M., Merlino, G.T., Willingham, M.C., Pastan, I. and Howard, B.H. (1982) Proc. Nat. Acad. Sci. USA 79, 6777-6781.