ORIGINAL RESEARCH

The homopurine–homopyrimidine DNA sequence flanking the *thyroglobulin* gene does not behave as a transcriptional diode in living cells

Christiane Christophe-Hobertus
Daniel Christophe

Institut de Recherche Interdisciplinaire en Biologie Humaine et Moléculaire (IRIBHM), Institut de Biologie et de Médecine Moléculaires (IBMM), Faculté de Médecine, Université Libre de Bruxelles, Gosselies, Belgium

Abstract: Homopurine–homopyrimidine DNA sequences are found at higher than expected frequency around gene promoters in mammalian genomes. They were reported to act as a transcriptional diode in in vitro assays, as only the homopurine DNA strand was transcribed. We show here that the homopurine–homopyrimidine upstream element of the *thyroglobulin* gene promoter is able to reduce transcription initiated from different promoters, both on transiently transfected plasmids or when integrated into the genome of rat thyroid cells. In living cells, this homopurine–homopyrimidine sequence inhibits transcription of both DNA strands.

Keywords: Pu/Py, promoter, transcription, *thyroglobulin*, thyroid

Introduction

Long homopurine–homopyrimidine (Pu/Py) DNA sequences are found in mammalian genomes at much higher frequency than that estimated by chance alone, raising lots of speculation about their possible function(s).[1] Pu/Py DNA motifs are also found at higher than expected frequency around transcription units,[2] and in many instances in the upstream region.[3] Some of these sequences are localized just next to, or within, gene promoters and have been reported to play a role in the control of promoter strength in a few cases.[4,5] A few have even been reported to display intrinsic promoter function.[6] DNA-binding proteins have also been shown to specifically recognize Pu/Py stretches and to behave as transcriptional regulators in some instances.[7] Pu/Py DNA sequences were also reported to reduce DNA replication[8–10] and to favor DNA recombination.[11,12] In appropriate conditions, repeat-containing Pu/Py DNA sequences are known to adopt a triple-helical structure called H-DNA.[13,14]

Regarding DNA transcription, it has been shown[3,15,16] that a long Pu/Py sequence could selectively block transcription in one direction when assayed in in vitro transcription assays involving a bacteriophage RNA polymerase. Transcription of the DNA strand bearing the homopurine motif (leading to a transcript with a homopyrimidine motif) was observed in this experimental system, but no transcript originating from the opposite homopyrimidine DNA strand was produced. But in *Escherichia coli*, the presence of a homopyrimidine motif in the coding sequence (Py-rich mRNA transcribed from Pu-rich DNA strand) was shown to strongly reduce protein expression.[17] A study conducted in COS-7 cells revealed that transcription of both DNA strands was reduced in living cells, mostly the transcription of the DNA strand bearing the homopyrimidine motif, but that protein expression was affected only in this latter case (when a Pu-rich mRNA was produced from the Py-rich DNA strand).[18] Obviously, the existing

Correspondence: Daniel Christophe
Institut de Recherche Interdisciplinaire en Biologie Humaine et Moléculaire (IRIBHM), Institut de Biologie et de Médecine Moléculaires (IBMM), rue des Professeurs Jeeners et Brachet 12, B-6041 Gosselies, Belgium
Tel +32 2 650 9828
Fax +32 2 650 9820
Email dchristo@ulb.ac.be

differences between assays conducted in in vitro transcription systems or in living cells, and between bacterial and mammalian cells transcription machineries, might explain the observed discordances. However, additional investigations are clearly needed in order to expand our knowledge about the effect of the presence of a Pu/Py DNA sequence on transcription in living cells.

The *thyroglobulin (Tg)* gene, which encodes a key protein involved in thyroid hormone synthesis, is flanked on its proximal 5′ side by a Pu/Py DNA element of variable length and primary sequence depending on the species.[13] No role was found for this motif in the functional study of the Tg gene promoter region, as assayed by transient transfection of reporter DNA constructs into thyroid cells.[19] We show here that this long Pu/Py motif inhibits transcription in rat thyroid cells, both in its natural context or when placed next to a heterologous promoter. Contrary to the previous in vitro observations,[3,15,16] transcription of both DNA strands was reduced.

Materials and methods
DNA constructions

Oligonucleotides used for polymerase chain reaction (PCR) and DNA sequencing were purchased from Eurogentec (Seraing, Belgium). The sequences are shown in Table 1. All PCR-amplified sequences were verified by DNA sequencing in order to exclude the presence of sequence changes introduced during the PCR step.

Tg reporter constructs

The sequences of the 5′ end of the rat Tg gene extending from position −168 to +9 (relative to transcription start as +1) were amplified by PCR from PCCl3 thyroid cell genomic DNA using primers CC313 (forward) and CC118 (reverse), which added XhoI and Hind III restriction sites at 5′ and 3′ ends, respectively, and inserted between the corresponding cloning sites in the bidirectional reporter vector pGL4LUC-RLUC,[20] yielding construct rTgLRL (Figure 1B).The Pu/Py elements of rat and mouse Tg gene 5′ upstream regions were amplified by PCR from PCCl3 cell and mouse tail genomic DNAs, respectively, using the pairs of primers (forward + reverse) CC314 + CC315 and CC425 + CC426, respectively, which added NheI and EcoRI restrictions sites at the 5′ and 3′ ends of the amplicons, respectively, and inserted between the corresponding cloning sites in plasmid rTgLRL, giving constructs Pu/Py-rTgLRL (Figure 1B; Pu/Py is either rat or murine Pu/Py sequence). For the construction of plasmid Py/Pu-rTgLRL

Table 1 Sequences of the oligonucleotides used as primers in PCR, qPCR, and RT-qPCR experiments

Name	Sequence (5′........3′)
CC118	CGAAGCTTCAGTGTCCCATCTGAGTCCC
CC313	CGCTCGAGCTCGAATTCACCCTACTGATTACTCAAGTATTCTT
CC314	GCGCTAGCAATTTTCAAGAATAGTGTCTACAGCT
CC315	GCGAATTCTGCCAGAGCTTTGTACCCTT
CC322	GCCTCGAGCAGGTGTAAATGTGTGAATCAGA
CC323	GCAAGCTTCGCGCAGCGTTCACAAAG
CC324	GCAAGCTTCAGGTGTAAATGTGTGAATCAGA
CC325	GCCTCGAGCGCGCAGCGTTCACAAAG
CC329	GCGCTAGCAGTCAAGAGCTTGGTGGAACTT
CC330	GCGAATTCCCTGGACAATCAACCAGCAAAG
CC425	GCGCTAGCAATTTTCTACAGTATTACCTACAGATATTAGG
CC426	GCGAATTCTATAAAACTGTGGAAGATAAAATACCTC
CC427	GCGAATTCAATTTTCTACAGTATTACCTACAGATATTAGG
CC428	GCGCTAGCTATAAAACTGTGGAAGATAAAATACCTC
CC430	GCGGATCCTATAAAACTGTGGAAGATAAAATACCTC
CC433	GCGGATCCGCGCAGCGTTCACAAAG
CC434	GCGGATCCAGGTGTAAATGTGTGAATCAGA
CC435	GCCTCGAGATTTTCTACAGTATTACCTACAGATATTAGG
CC439	AGGACGCTCCAGATGAAAT
CC444	(T)$_{29}$ GTTGTTA
CC445	CAAGATCCGCGAGATTCTCATT
CC446	TTTCAAGAATAGTGTCTACAGCTGAATTGCTCTAA
CC447	GGGACAAGAAATGTGAGGCTAGCACTAA
CC448	AGGGGCATGTCGACCTTATGTGTAAA
CC449	TGCTTTAGAGCAATTCAGCTGTAGACACT
CC450	CTGCCAGCTACACAAGCAACGAA
CC451	CCCTCTGTGTCTACACACCAGCATT

Abbreviations: PCR, polymerase chain reaction; qPCR, quantitative PCR; RT-qPCR, quantitative reverse transcription PCR.

(Figure 1B) in which the murine Pu/Py sequence was placed in the reverse orientation relative to the Tg promoter, the same strategy as mentioned earlier was followed except that the pair of primers CC427 + CC428 was used in the PCR which resulted in the addition of EcoRI and NheI restrictions sites at the 5′ and 3′ ends of the amplicon, respectively. The sequence extending from −3469 to −3696 upstream of the rat Tg gene were amplified by PCR from PCCl3 cell genomic DNA, using primers CC329 (forward) and CC330 (reverse), which added NheI and EcoRI restriction sites in 5′ and 3′, respectively, and inserted between the corresponding cloning sites in plasmid rTgLRL, giving construct rDNA-rTgLRL (Figure 1B). For stable transformation of PCCl3 cells, the *Renilla* luciferase (RLUC) reporter gene and rat Tg promoter sequences with or without the Pu/Py region were excised from plasmid Pu/Py-rTgLRL or rTgLRL, respectively, using Asp718 and HindIII, and inserted into (Asp718 + HindIII)-cleaved pGL4.17 vector (Promega Corp, Madison, WI, USA).
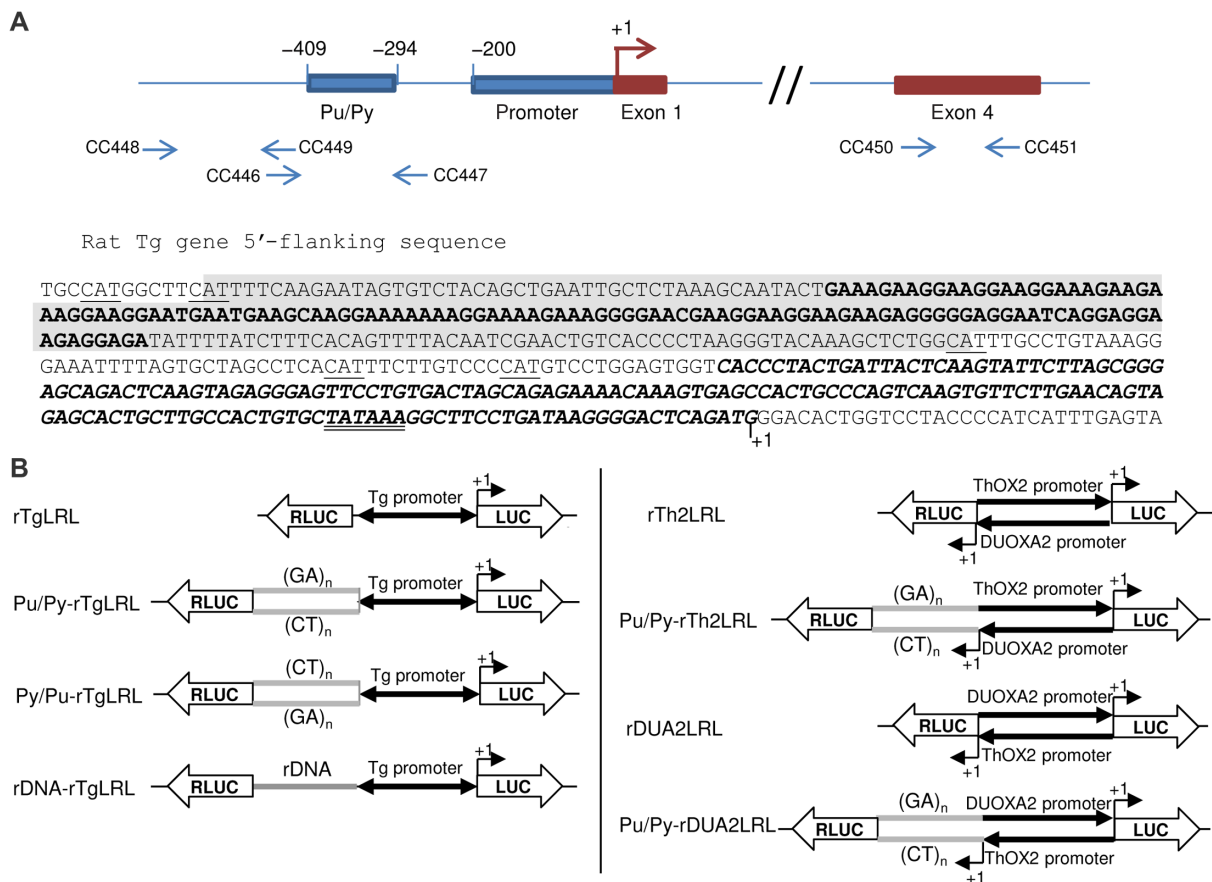
Figure I Genomic DNA sequences and reporter constructs used in the study.

Notes: (**A**) Schematic view of the organization of the 5′ end of the rat Tg gene with the locations of the primer pairs used in the (RT-)qPCR experiments, and sequence of the rat Tg gene 5′-flanking region showing the homopurine–homopyrimidine (Pu/Py) region used in our reporter constructs (light gray shaded) and the Pu/Py sequence itself (bold characters), the CAT motifs (underlined; ATG on complementary strand: see text), and the Tg promoter region used in the reporter constructs (bold italics, double-underlined TATA box and TSS indicated as "+1"). (**B**) Schematic representation of the reporter constructs used in this study. Firefly luciferase (LUC) and *Renilla* luciferase (RLUC) genes are positioned in a head-to-head configuration. $(GA)_n$ and $(CT)_n$: Pu/Py region in forward and reverse orientation, respectively; rDNA: rat genomic DNA fragment (see text).

Abbreviations: Pu/Py, homopurine–homopyrimidine; PCR, polymerase chain reaction; qPCR, quantitative PCR; RT-qPCR, quantitative reverse transcription PCR; TSS, transcription start site.

ThOX2–DUOXA2 reporter constructs

The bidirectional promoter region of ThOX2 and DUOXA2 genes[20] was amplified by PCR from PCCl3 cell genomic DNA using the pairs of primers (forward + reverse) CC324 + CC325 or CC322 + CC323 which added HindIII and XhoI restriction sites at 5′ and 3′ or 3′ and 5′ ends, respectively, of the 190 bp long amplicon. The amplified sequences were inserted between the corresponding cloning sites in the bidirectional reporter vector pGL4LUC-RLUC, yielding construct rTh2LRL or rDUA2LRL, respectively (Figure 1B). The Pu/Py element of the mouse Tg gene 5′ upstream region was amplified by PCR from mouse tail genomic DNA using primers CC435 (forward) and CC430 (reverse), which added XhoI and BamHI restriction at 5′ and 3′ ends, respectively, and the bidirectional promoter region of ThOX2 and DUOXA2 genes was amplified by PCR from PCCl3 cell genomic DNA using the pairs of primers (forward + reverse) CC324 + CC433 or

CC434 + CC323 which added HindIII and BamHI restriction sites at 5′ and 3′ or 3′ and 5′ ends, respectively. The amplified Pu/Py and promoter sequences were inserted together into (XhoI + HindIII)-cleaved bidirectional reporter vector pGL4LUC-RLUC, yielding construct Pu/Py-rTh2LRL or Pu/Py-rDUA2LRL, respectively (Figure 1B).

Cell culture and DNA transfections

PCCl3 cells were cultured as described by Fusco and coworkers.[21] The cells were trypsinized and seeded at 30%–40% confluence the day preceding transfection. Co-transfections were performed using 3.3 µL of FuGENE 6 transfection reagent (Roche Diagnostics GmbH, Mannheim, Germany) and 1.1 µg of DNA per dish (diameter: 3.5 cm). To each individual cell culture dish were added 1 µg of firefly (LUC) + *Renilla* (RLUC) luciferases reporter construct and 100 ng of pTPO-SEAP,[22] containing the secreted alkaline

phosphatase coding sequence under the control of the human thyroperoxidase gene promoter and used as internal control for normalization in regard to transfection efficiency. In order to obtain PCCl3 cell clones containing copies of the reporter construct integrated within their genome, transfection was performed by adding to the dish (diameter: 6 cm) 1 µg of the reporter plasmid (pGL4.17 vector backbone) and FuGENE 6 transfection reagent at a ratio of FuGENE 6 (µL) to DNA (µg) of 3:1. After 24 hours, cells stably transformed with the reporter construct were selected in culture medium supplemented with 400 µg/mL of Geneticin (Thermo Fisher Scientific, Waltham, MA, USA) and maintained afterward in the presence of 200 µg/mL of Geneticin.

Luciferase assays

LUC and RLUC activities were measured between 24 and 48 hours after transfection using the Dual-Luciferase® Reporter Assay System as recommended by the manufacturer (Promega Corp, Madison, WI, USA) and a TD-20/20 luminometer (Turner Designs, Sunnyvale, CA, USA). Secreted embryonic alkaline phosphatase (SEAP) activity present in the culture medium was measured at the same time using the Phospha-Light™ System (Tropix Inc. [Thermo Fisher], Bedford, MA, USA). All transfections were performed in duplicate in each individual experiment. LUC and RLUC activities were expressed in arbitrary units as displayed on the luminometer after a 10-second integration time and normalized against SEAP production in the individual transfected dishes. The RLUC/LUC ratio was arbitrarily set to the value of one for the construct containing only the promoter sequence. Statistical analysis of the data was performed using the two-tailed Mann–Whitney test available on the VassarStats web site for statistical computation (http://vassarstats.net/index.html).

qPCR and RT-qPCR experiments

Total RNA was obtained from cultured PCCl3 cells (dish diameter: 10 cm) using the RNeasy Mini Kit from Invitrogen (Life Technologies Europe BV, Gent, Belgium). Residual DNA was removed from the total RNA preparations using the Turbo DNA-free kit (Thermo Fisher Scientific). RNA was reverse transcribed into cDNA by random priming using the Superscript IIRT system from Invitrogen. RNA from PCCl3 cells transiently transfected with a reporter construct was obtained identically but reverse transcribed using an oligo-dT primer (T_{15}). Genomic DNA was obtained from PCCl3 cells using a standard protocol.[23] For the analysis of natural transcripts qPCR (quantitative PCR) or RT-qPCR (quantitative reverse transcription PCR) mixes contained 10 ng of genomic DNA or 6.25 ng of cDNA, respectively, and 200 nM of each primer (Table 2) in the iQ™ SYBR Green Supermix (Bio-Rad Laboratories, Hercules, CA, USA). Reactions (final volume: 20 µL) were performed in a C1000™ Thermal Cycler equipped with a CFX96™ Real-Time System (Bio-Rad Laboratories) using an initial activation step (95°C, 3 minutes) and 40 repetitions of a two-step cycle, first step at 95°C for 10 seconds and second step at 60°C for 1 minute, and were followed by a melting curve analysis of the produced amplicons. For the quantification of LUC and RLUC transcripts, 25 ng of cDNA was used with identical primer concentrations, basal reaction mix, and initial activation step, and then 40 repetitions of a two-step cycle, first step at 95°C for 10 seconds and second step at 56°C for 1 minute, were performed and followed by a melting curve analysis of the produced amplicons. The primer pairs (forward + reverse) used were as follows: LUC: CC445 + CC444 (amplicon size: 255 bp); RLUC: CC439 + CC 444 (amplicon size: 268 bp). Reactions were done in duplicate in each experiment and the experiments were repeated at least once.

Results and discussion

The 5′ homopurine–homopyrimidine DNA sequence of the *thyroglobulin* gene reduces reporter gene expression

The Tg gene promoter is flanked on its 5′ side by a long Pu/Py DNA sequence in several mammalian species (Figure 1A), including man.[13] This classical TATA box-containing promoter directs the unidirectional production of Tg mRNA in thyroid cells. However, by inserting the minimal Tg gene promoter region in a bidirectional reporter plasmid harboring both the LUC and the RLUC genes in a head-to-head configuration

Table 2 Absence of stable antisense transcripts originating from the Tg promoter in rat thyroid cells

1 Target sequence	2 Cq value of qPCR on genomic DNA	3 Cq value of RT-qPCR on total RNA	4 Cq value of qPCR on total RNA
Upstream of Pu/Py	26.72±0.2	37.5±1.5	38.5±0.3
Pu/Py region	26.93±0.18	38.5±0.5	Not detected
Exon 4 of Tg gene	25.92±0.05	19.06±0.02	Not detected

Notes: Cq values (mean of duplicates ± half-range) obtained in representative qPCR (columns 2 and 4) and RT-qPCR (column 3) experiments involving genomic DNA (column 2) or total RNA (columns 3 and 4) are shown. The pairs of primers used were as follows: upstream of Pu/Py: CC448 + CC449; Pu/Py region: CC446 + CC447; exon 4 of Tg gene: CC450 + CC451.
Abbreviations: Pu/Py, homopurine–homopyrimidine; PCR, polymerase chain reaction; qPCR, quantitative PCR; RT-qPCR, quantitative reverse transcription PCR.

(see construct rTgLRL in Figure 1B), we observed an efficient expression of both reporter genes, presumably as a consequence of the production of promoter upstream transcripts (PROMPTs) with appropriate stabilization signals by the antisense transcriptional activity of the promoter.[24] When inserting the Pu/Py sequence in its natural upstream position relative to the Tg promoter in the bidirectional reporter construct (see construct Pu/Py-rTgLRL in Figure 1B), we took care not to include potential translation initiation codons as they would interfere with RLUC expression by initiating translation upstream from the RLUC ORF (those codons appearing as CAT on the top strand in Figure 1A are underlined). As a control, we also constructed another plasmid in which a fragment of intergenic DNA of similar size (and also devoid of potential translation initiation codons), corresponding to the sequences located between −3,469 and −3,696 relative to the Tg gene TSS, was inserted at the same position relative to the Tg promoter (see construct rDNA-rTgLRL in Figure 1B). This construct would allow us to confirm that any observed effect with the Pu/Py-containing construct was actually related to its Pu/Py content and not only to the size of the insert. The choice of using bidirectional reporter constructs was dictated by the ability of the studied Pu/Py element to adopt a triplex structure (H-DNA) when present on supercoiled plasmid DNA[13], which could possibly affect the transfection efficiency of the plasmids bearing this element as compared to the control plasmids devoid of the Pu/Py DNA sequence. As previous work had demonstrated that the presence of the Pu/Py motif did not influence Tg promoter activity,[19] the expression of the LUC reporter gene could serve as a plasmid-specific internal control for transfection efficiency of the different reporter constructs used. Indeed, as shown in Figure 2A, substantial differences were observed in the level of LUC expression from the different reporter constructs even after normalization for overall transfection efficiency by using a co-transfected reporter construct encoding SEAP. The effect of the presence of the Pu/Py element on the expression of the RLUC reporter gene was thus evaluated by calculating the ratio of RLUC expression relative to LUC expression for each of the individual reporter plasmids. An arbitrary ratio of 1 was set for the parental reporter construct containing only the promoter region. As shown in Figure 2B, after transfection into PCCl3 thyroid cells, the construct containing the Pu/Py element in its natural upstream location relative to the Tg promoter (rPu/Py-rTgLRL) displayed a substantially decreased RLUC/LUC expression ratio as compared to both parent (Tg promoter alone, rTgLRL) and control DNA insert (rDNA-rTgLRL) constructs.

The murine Pu/Py sequence homologous to that of the rat Tg gene promoter displays only limited primary sequence conservation with its counterpart in the rat (72% of base identity; random Pu/Py sequences scoring 50% base identity by chance only). Replacing the rat Pu/Py element by its murine counterpart resulted in a very similar reduction in antisense gene expression (Figure 2B; construct mPu/Py-rTgLRL), which supported the view that the Pu/Py asymmetry rather than the primary DNA sequence itself was responsible for the observed effect on gene expression.

To investigate whether the effect on gene expression was dependent on the orientation of the Pu/Py sequence relative to the promoter, we also placed the same Pu/Py element in the opposite orientation relative to the Tg promoter in construct Py/Pu-rTgLRL (Figure 1B). As shown in Figure 2B (construct mPy/Pu-rTgLRL), RLUC expression was also decreased in this case, but to a lesser extent than when the Pu/Py motif was positioned in the other orientation. However, it must be noted that when comparing the levels of RLUC expression from Pu/Py and mPy/Pu-rTgLRL constructs normalized against the co-expressed SEAP reporter (Figure 2A), the difference is lower than when comparing the ratio of RLUC to LUC expression from the same constructs (Figure 2B) due to the decreased LUC expression from mPy/Pu-rTgLRL construct as compared to mPu/Py-rTgLRL (Figure 2A). Obviously, the effect of the Pu/Py element on gene expression was not strictly orientation-dependent, contrary to what was observed previously in vitro.[3,15,16] In addition, in our experiments the Pu/Py element exerted its main effect on the opposite transcriptional direction as compared to what had been observed in these in vitro experiments. According to these previous reports, the production of RLUC reporter transcripts containing a homopyrimidine RNA motif from the constructs Pu/Py-rTgLRL, which leads to the lower productions of RLUC in our experiments, should not be inhibited. In COS-7 cells transcription of both DNA strands was also reduced, in agreement with our observations on reporter gene expression in rat thyroid cells, but protein expression was not affected accordingly.[18] It is conceivable that cell-specific factors may be involved in these distinct outcomes.

We also investigated whether the same effect could be observed when transcription was driven by other promoters. To do this, the Pu/Py element was inserted on either side of the bidirectional ThOX2–DUOXA2 promoter region[20] (see constructs Pu/Py-rTh2LRL and Pu/Py-rDUA2LRL in Figure 1B) and the expression of the reporter genes driven by these constructs were compared to those obtained with the parent plasmids (constructs rTh2LRL and rDUA2LRL,
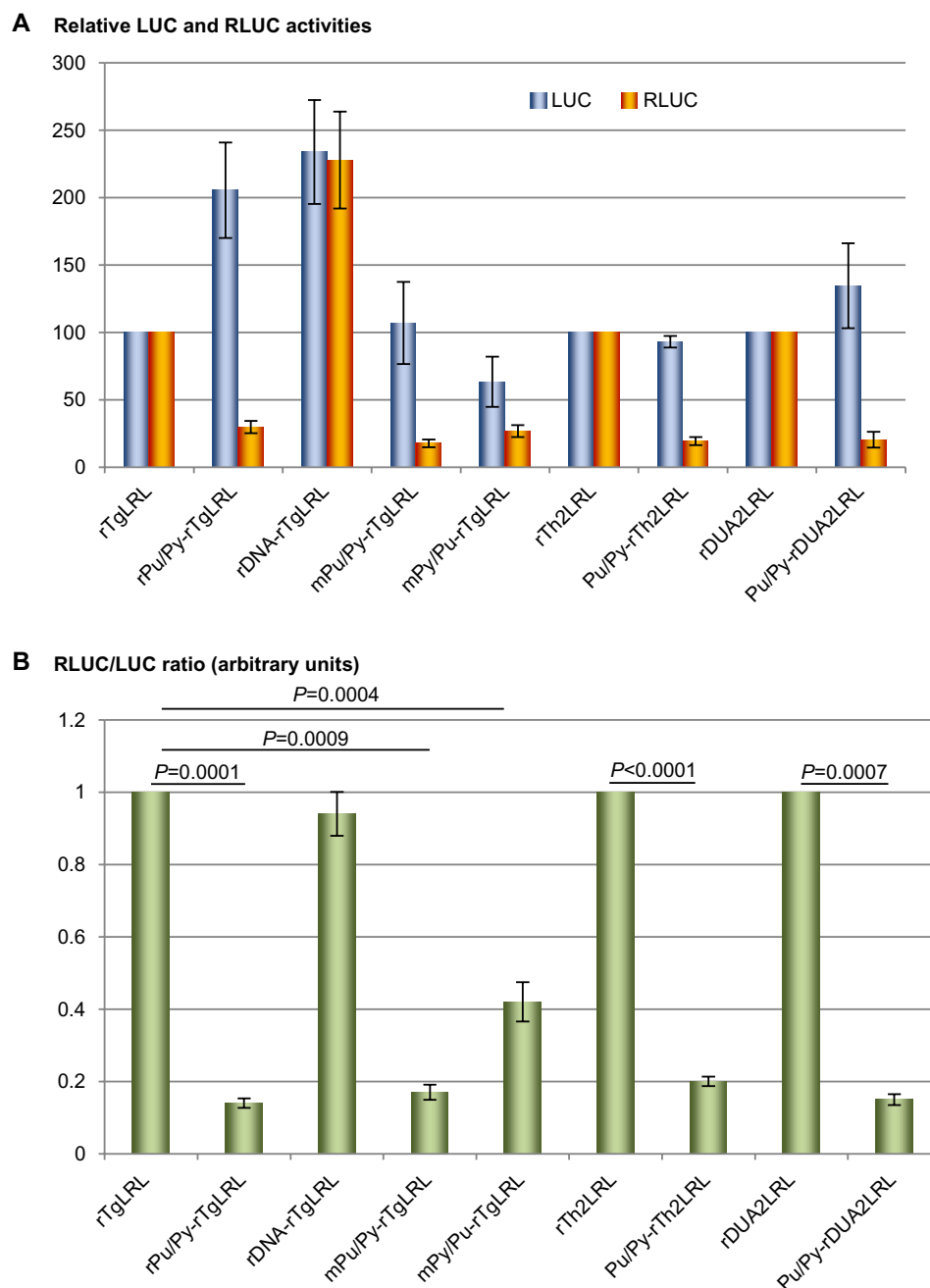
A   **Relative LUC and RLUC activities**

B   **RLUC/LUC ratio (arbitrary units)**

Figure 2 The presence of the Pu/Py element reduces reporter gene expression.
Notes: (**A**) The reporter constructs (see Figure 1B for detailed structures) were transiently transfected in rat thyroid PCCl3 cells and the expressions of LUC and RLUC were measured individually in the same transfected cells. The LUC and RLUC values were arbitrarily set to 100% for the construct containing only the promoter region. (**B**) The same results were expressed as ratios of RLUC/LUC production for each of the constructs. The ratio was arbitrarily set to 1 for the construct containing only the promoter region. The results shown are mean ± SD from at least three independent experiments. Statistical analysis was performed using two-tailed Mann–Whitney test.
Abbreviations: Pu/Py, homopurine–homopyrimidine; LUC, firefly luciferase; RLUC, *Renilla* luciferase; SD, standard deviation.

respectively, in Figure 1B). As shown in Figure 2B, the presence of the Pu/Py sequence decreased the relative level of RLUC expression driven by either the ThOX2 promoter or the DUOXA2 promoter. Thus, the Pu/Py element of the Tg promoter inhibited gene expression from both of these promoters as well as antisense gene expression from the Tg promoter itself.

The homopurine–homopyrimidine DNA motif negatively controls mRNA accumulation

Before it could be concluded that the presence of the Pu/Py element reduced transcriptional activity in our transfections, we had to exclude the possibility that the presence of the Pu/Py sequence in the 5′ UTR (untranslated region) of the mRNA

prevented efficient translation of the RLUC ORF. A decreased RLUC production could indeed result either from a lower amount of RLUC mRNA or from a reduced translational efficiency of this mRNA. We thus conducted RT-qPCR experiments on mRNA preparations from transfected cells with the aim of quantifying both LUC and RLUC transcripts. The relative amounts of LUC and RLUC mRNAs were measured in RNA preparations from cells transfected with the Tg promoter reporter constructs containing or not the Pu/Py element. We devised a specific strategy relying on the use of an anchored oligo-dT primer in the RT-qPCR experiments in order to prevent the amplification of non-polyadenylated LUC or RLUC coding sequence. The absence of amplification when reverse transcription of the RNAs was omitted confirmed the specificity of our measure. The results of the RT-qPCR experiments were expressed as $2^{(Cq\,RLuc-CqLuc)}$ and this value was arbitrarily set to 1 in the case of the Pu/Py-less reporter construct. As depicted in Figure 3, the value of $2^{\Delta Cq}$ was higher for the Pu/Py-containing construct, indicating that the RLUC/LUC mRNA ratio was significantly decreased in RNA preparations from cells transfected with the Pu/Py-containing reporter construct as compared to the Pu/Py-less reporter transfected cells (the higher $2^{\Delta Cq}$ value reflected the increased Cq value obtained for RLUC compared with LUC cDNAs, which corresponded to a decreased amount of RLUC cDNA compared with LUC cDNA). Thus, the accumulation of transcripts originating from the Tg promoter was clearly reduced on the strand bearing the Pu/Py sequence as compared to the opposite direction.

We also investigated if PROMPTs originating from the Tg promoter and encompassing the Pu/Py sequence
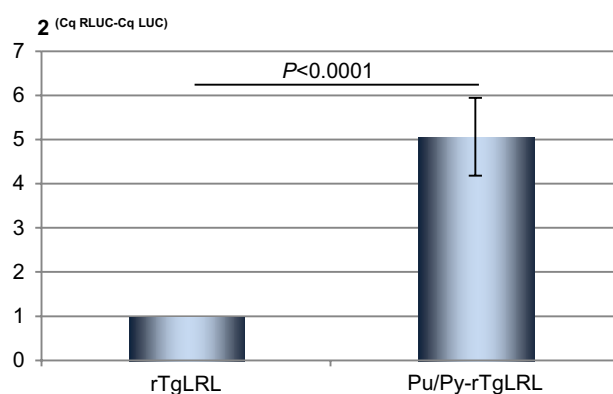
accumulated to a detectable level in rat thyroid cells. RT-qPCR experiments were performed on total RNA. We used two primer pairs directing the amplification of sequences located upstream from the promoter and a pair of primers enabling the amplification of sequences in exon 4 of the Tg gene as a positive control (Figure 1A). When assessed in qPCR experiments using genomic DNA, all three primers pairs directed the amplification of their target sequences with comparable efficiencies (Table 2, column 2). By contrast, only the sequences of exon 4 were significantly amplified in the RT-qPCR experiments (Table 2, column 3), indicating the absence of significant amounts of transcripts encompassing the sequences located upstream from the promoter in the RNA preparations. Omitting the reverse transcription step resulted in the absence of significant amplification in all three cases as expected (Table 2, column 4). Thus, no significant level of PROMPTs encompassing the upstream Pu/Py motif of the Tg gene could be detected in rat thyroid cells, which is in agreement with the adverse effect of this DNA sequence on transcription observed in our previous experiments. Whether limiting the extension of PROMPTs originating from the Tg promoter constitutes the main physiological function of the evolutionary conserved Pu/Py element lying in the 5′ vicinity of the Tg promoter remains, however, to be investigated specifically.

The transcription-blocking effect also occurs in a genomic context

When present on supercoiled plasmid DNA, the Pu/Py sequence we used could adopt a triple-helical structure called H-DNA.[13,14] We thus investigated whether the effect we observed using transiently transfected plasmid DNAs could be reproduced when the DNA constructs were integrated into the cellular genomic DNA. Cell clones stably transformed with either the Pu/Py-containing or the Pu/Py-less Tg promoter reporter construct were established and the relative amounts of LUC and RLUC activity present in pools of these clones were quantified. As shown in Figure 4, the expected difference in the mean ratio of RLUC/LUC expression was observed between pools of cell clones transformed with the Pu/Py-containing or the Pu/Py-less reporter construct. Thus, the presence of the Pu/Py region reduced gene expression from the Tg promoter in a genomic context as well as when assayed using transfected plasmids.

Conclusion

Contrary to previous observations made in vitro using a bacteriophage RNA polymerase,[3,15,16] Pu/Py DNA sequences do
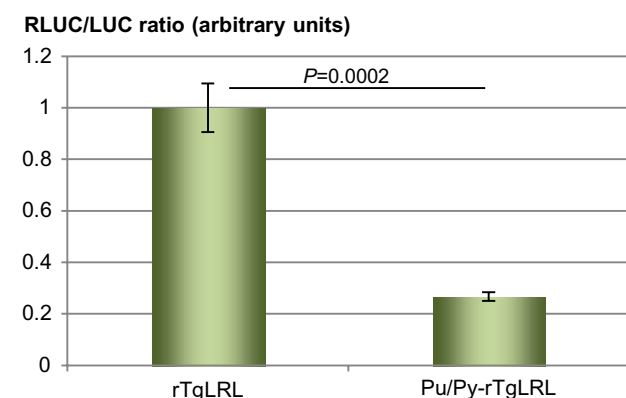

$2^{(Cq\,RLUC-Cq\,LUC)}$

Figure 3 LUC and RLUC mRNAs levels in cells transfected with Pu/Py-containing or Pu/Py-less reporter construct as determined by RT-qPCR.
Notes: The results shown are the means ± SD from six individual measures expressed in the form of $2^{\Delta Cq}$ relatively to the value obtained for the Pu/Py-less reporter construct which was arbitrarily set to 1. Statistical analysis was performed using two-tailed Mann–Whitney test.
Abbreviations: Pu/Py, homopurine–homopyrimidine; LUC, firefly luciferase; RLUC, *Renilla* luciferase; RT-qPCR, quantitative reverse transcription polymerase chain reaction; SD, standard deviation.

RLUC/LUC ratio (arbitrary units)

Figure 4 Assay of the transcription-blocking effect of the Pu/Py sequence in a genomic context.
Notes: Ratios of RLUC/LUC expression in pools of clones of rat thyroid PCCl3 cells transformed with Pu/Py-containing or Pu/Py-less Tg promoter reporter construct are shown. The ratio was arbitrarily set to 1 for the construct containing only the promoter region. Four individual measures of RLUC/LUC ratio in independent pools of transformed cells were made for each construct. Statistical analysis was performed using two-tailed Mann–Whitney test.
Abbreviations: LUC, firefly luciferase; RLUC, *Renilla* luciferase; Pu/Py, homopurine–homopyrimidine.

not appear to behave as transcriptional diodes in mammalian cells. In COS-7 cells[18] and in rat thyroid cells (this study) at least, their presence reduced transcriptional activity on both DNA strands, with some minor differences in the level of transcription of one DNA strand relative to the other depending on the cell type.

Acknowledgments

Disclosure

DC is a research director of the Belgian Fonds National de la Recherche Scientifique (FNRS). The authors report no other conflicts of interest in this work.

References

1. Albrecht-Buehler G. Properties and distribution of pure GA-sequences of mammalian genomes. *PLoS ONE.* 2008;3(11):e3818.
2. Lu G, Ferl RJ. Homopurine/homopyrimidine sequences as potential regulatory elements in eukaryotic cells. *Int J Biochem.* 1993;25:1529–1537.
3. Grabczyk E, Fishman MC. A long purine-pyrimidine homopolymer acts as a transcriptional diode. *J Biol Chem.* 1995;270:1791–1797.
4. Glaser RL, Thomas GH, Siegfried E, Elgin SCR, Lis JT. Optimal heat-induced expression of the Drosophila hsp26 gene requires a promoter sequence containing $(CT)_n$-$(GA)_n$ repeats. *J Mol Biol.* 1990;211: 751–761.
5. Raghu G, Tevosian S, Anant S, Subramanian KN, George DL, Mirkin SM. Transcriptional activity of the homopurine-homopyrimidine repeat of the c-Ki-ras promoter is independent of its H-forming potential. *Nucl Acids Res.* 1994;22:3271–3279.
6. Darvish H, Nabi MO, Firouzabadi SG, et al. Exceptional human core promoter nucleotide compositions. *Gene.* 2011;475:79–86.
7. Choi HS, Hwang CK, Song KY, Law P-Y, Wei L-N, Loh HH. Poly(C)-binding proteins as transcriptional regulators of gene expression. *Biochem Biophys Res Commun.* 2009;380:431–436.
8. Rao BS. Regulation of DNA replication by homopurine/homopyrimidine sequences. *Mol Cell Biochem.* 1996;156:163–168.
9. Krasilnikov AS, Panyutin IG, Samadashwily GM, Cox R, Lazurkin YS, Mirkin SM. Mechanisms of triplex-caused polymerization arrest. *Nucl Acids Res.* 1997;25:1339–1346.
10. Krasilnikova MM, Smirnova EV, Krasilnikov AS, Mirkin SM. A new trick for an old dog: TraY binding to a homopurine-homopyrimidine run attenuates DNA replication. *J Mol Biol.* 2001;313:271–282.
11. Kohwi Y, Panchenko Y. Transcription-dependent recombination induced by triple-helix formation. *Gene Dev.* 1993;7:1766–1778.
12. Bacolla A, Collins JR, Gold B, et al. Long homopurine-homopyrimidine sequences are characteristic of genes expressed in brain and the pseudo-autosomal region. *Nucl Acids Res.* 2006;34:2663–2675.
13. Christophe D, Cabrer B, Bacolla A, Targovnik H, Pohl V, Vassart G. An unusually long poly(purine)-poly(pyrimidine) sequence is located upstream from the human thyroglobulin gene. *Nucl Acids Res.* 1985;13:5127–5144.
14. Htun H, Dahlberg JE. Topology and formation of triple-stranded H-DNA. *Science.* 1989;243:1571–1576.
15. Belotserkovskii BP, Liu R, Tornaletti S, Krasilnikova MM, Mirkin SM, Hanawalt PC. Mechanisms and implications of transcription blockage by guanine-rich DNA sequences. *Proc Natl Acad Sci U S A.* 2010;107: 12816–12821.
16. Belotserkovskii BP, Neil AJ, Saleh SS, Shin JH, Mirkin SM, Hanawalt PC. Transcription blockage by homopurine DNA sequences: role of sequence composition and single-strand breaks. *Nucl Acids Res.* 2013;41:1817–1828.
17. Sarkar PS, Brahmachari SK. Intramolecular triplex potential sequence within a gene down regulates its expression in vivo. *Nucl Acids Res.* 1992;20:5713–5718.
18. Ohshima K, Montermini L, Wells RD, Pandolfo M. Inhibitory effect of expanded GAA-TTC triplet repeats from intron 1 of the Friedreich ataxia gene on transcription and replication in vivo. *J Biol Chem.* 1998;273:14588–14595.
19. Christophe D, Gérard C, Juvénal G, et al. Identification of a cAMP-responsive region in thyroglobulin gene promoter. *Mol Cell Endocrinol.* 1989;64:5–18.
20. Christophe-Hobertus C, Christophe D. Delimitation and functional characterization of the bidirectional THOX-DUOXA promoter regions in thyrocytes. *Mol Cell Endocrinol.* 2010;317:161–167.
21. Fusco A, Berlingieri MT, Di Fiore PP, Portella G, Grieco M, Vecchio G. One- and two-step transformations of rat thyroid epithelial cells by retroviral oncogenes. *Mol Cell Biol.* 1987;7:3365–3370.
22. Vilain C, Rydlewski C, Duprez L, et al. Autosomal dominant transmission of congenital thyroid hypoplasia due to loss-of-function mutation of PAX8. *J Clin Endocrinol Metab.* 2001;86:234–238.
23. Sambrook J, Russell DW. *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 2001.
24. Ntini E, Järvelin AI, Bornholdt J, et al. Polyadenylation site-induced decay of upstream transcripts enforces promoter directionality. *Nat Struct Mol Biol.* 2013;20:923–928.