Amplification of GC-rich *ADAMTS-2* and *URG4/URGCP* promoter regions with optimized combination of PCR enhancers

Meltem ALPER[1], Esra TOKAY[2], Feray KÖÇKAR[2],*
[1]Aksaray Vocational School of Technical Sciences, Aksaray University, Aksaray, Turkey
[2]Department of Biology, Faculty of Science and Literature, Balıkesir University, Balıkesir, Turkey

Abstract: PCR is an effective and widely used method for the amplification of target DNA fragments in vitro. The templates that have high guanine-cytosine (GC) content, complex secondary structures, or short tandem repeats are difficult to amplify by the conventional PCR methods. Human genome regulatory regions such as promoters are rich in terms of GC base compositions. To amplify these regions some special modifications or additives are required. Using different PCR strategies such as hot-start/touchdown PCR or the utilization of various additives into the reaction mixture such as organic molecules can improve PCR yield. In the present study we optimized the PCR conditions for the amplification of the human *ADAMTS-2* gene promoter region featuring an extremely high GC content and a secondary structure. We show that a combination of three additives, betaine, dimethyl sulfoxide, and 7-deaza GTP, is essential to obtain a correct PCR amplicon. To demonstrate whether the optimized PCR condition was applicable to other GC-enriched regions, a similar procedure was repeated for the amplification of the human *URG4/URGCP* gene promoter.

Key words: GC-rich PCR, PCR enhancers, promoter amplifications, robust PCR

1. Introduction

Polymerase chain reaction (PCR) is a conventional method for the amplification of target DNA regions. A large variety of parameters can influence the outcome of a PCR amplification. These variables include cycling parameters as well as the concentrations of Mg^{+2}, H^+, dNTP, primers, templates, and added enhancers (Innis and Gelfand, 1990; Hecker and Roux, 1996). Most DNA does not require special modifications or additives for amplification. However, the templates that have tandem repeats, high guanine-cytosine (GC) content, and strong secondary structures are difficult to amplify with basic PCR conditions. Increased hydrogen bonding with neighboring cytosine and guanine generates complex hairpins and loops in the GC-rich regions (McDowell et al., 1998; Musso et al., 2006). In the human genome there are often GC-rich sequences including important regulatory domains such as promoters, enhancers, and control elements (Wilson, 1997; Mamedov et al., 2008). The PCR amplification of the GC-rich DNA is often problematic because of stable secondary structures in the DNA that are resistant to melting. These secondary structures cause DNA polymerases to stall, resulting in incomplete and nonspecific amplification (Sahdev et al., 2007).

To amplify these problematic GC-rich DNA fragments, several approaches have been developed. Organic molecules such as dimethyl sulfoxide (DMSO), glycerol, polyethylene glycol, formamide, betaine, 7-deaza GTP, and dUTP have been included in the reaction mixture and have been shown to improve the amplification of GC-rich DNA sequences (Pamp and Medrano, 1991; Sun et al., 1993; Mutter and Baynton, 1995; Turner and Jenkins, 1995; Baskaran et al., 1996; Henke et al., 1997; Chakrabarti and Schutt, 2001; Kang et al., 2005; Musso et al., 2006). Moreover, different PCR strategies were improved, such as hot-start PCR, step-down PCR, and a hot-start method combined with touch-down to increase PCR yield (Korbie and Mattek, 2008).

ADAMTS-2 (a disintegrin and metalloproteinase with thrombospondin type I, motif II) is a metalloproteinase playing a key role in the processing of fibrillar collagen precursors into mature collagen molecules by cleaving of the amino terminus of the type I, II, III, and V procollagens. This maturation process allows the correct fibril formation (Colige et al., 2005). The functional analysis of the human ADAMTS-2 protein has been widely studied and mutations in the *ADAMTS-2* gene have been illustrated with Ehlers–Danlos syndrome type VIIC in humans (Colige et al.,

---

* Correspondence: feraykockar@hotmail.com

2004). Antiangiogenic and antitumor activities of this gene have also been identified in recent studies (Dubail et al., 2010).

The human upregulator of cell proliferation (*URG4/URGCP*; GenBank NM_017920) encodes 922 amino acids in cytoplasm. Overexpression of *URG4/URGCP* in HepG2 cells promotes hepatocellular growth and survival in tissue culture and in soft agar, and accelerated tumor development in nude mice. *URG4/URGCP* is strongly expressed in hepatocellular carcinoma, gastric cancer, and osteosarcoma. It may be a putative oncogene that contributes importantly to multistep carcinogenesis and cell cycle regulation (Tufan et al., 2002; Dodurga et al., 2012).

To find out the molecular mechanisms involved in the regulation of the *ADAMTS-2* and *URG4/URGCP* genes it is necessary to amplify their promoter regions. Due to the high GC content, multiple CpG island regions, short tandem repeats, and perfect hairpin structures, the *ADAMTS-2* and *URG4/URGCP* promoter regions are extremely difficult to amplify by conventional PCR.

In the present study 658 bp of the putative human *ADAMTS-2* gene promoter were amplified using a PCR-based strategy with conventional Taq DNA polymerase. Different PCR enhancers, Taq DNA polymerase enzymes, and PCR methods were used to obtain the correct amplicon. Because of the failure of these approaches, optimized concentrations of some PCR enhancers such as DMSO, 7-deaza GTP, and betaine were introduced into the reaction mixture as enhancers. In this case a specific PCR product yield was obtained. The PCR fragment was cloned into a pGEMT Easy Vector system and confirmed by automated DNA sequencing.

Afterwards, to demonstrate whether the optimized PCR condition was applicable for other similar GC-enriched regions, the same procedure was repeated for the amplification of the putative human *URG4/URGCP* promoter. We obtained a high yield of specific PCR product including the combinations of the PCR enhancers as mentioned above. The PCR fragment was also cloned into the pGEMT Easy Vector and confirmed by automated DNA sequencing. In both cases, we obtained a specific PCR amplification, demonstrating that these low-cost additives can be used for the amplification of a variety of DNA templates with extremely high G and C deoxyribonucleotide content and complex secondary structures.

2. Materials and methods
2.1. Genomic DNA isolation from cultured cells
Saos-2 cells were kindly provided by Dr Kenneth Wann (Cardiff School of Biosciences, Cardiff, UK). Monolayers of the cells were maintained in Dulbecco's modified Eagle medium (Sigma) containing 10% heat inactivated (56 °C for 1 h) fetal calf serum (Invitrogen) and 2 mM L-glutamine (HyClone). The cultures were maintained at 37 °C in a humid incubator with a 5% (v/v) CO_2 atmosphere. Genomic DNA was prepared from Saos-2 cells using a Blood & Cell Culture DNA Mini Kit (QIAGEN) according to the manufacturer's instructions.

2.2. Bioinformatic analysis
To determine the putative human *ADAMTS-2* and *URG4/URGCP* promoter regions we obtained the sequences approximately 1.5 kb upstream of the translation start site of the *ADAMTS-2* gene (GenBank accession no. NM_014244) and the human *URG4/URGCP* gene (GenBank accession no. NM_017920) using the genome browser feature of the National Center for Biotechnology Information. Potential promoter sequences of these genes were submitted to Genomatix (https://www.genomatix.de/) to predict potential core promoter elements, transcription factors, and putative islands. GC content and CpG nucleotide composition of the *ADAMTS-2* and *URG4/URGCP* promoters were identified by using the bioinformatic tool EMBOSS CpGPlot /CpGReport/ Isochore (http://www.ebi.ac.uk/Tools/emboss/cpgplot/). Secondary structure analysis was also performed to display hairpin and loop structures using the web-based mfold program (http://mfold.rna.albany.edu/?q=mfold/DNA-Folding-Form).

2.3. Genomic PCR
The *ADAMTS-2* and *URG4/URGCP* putative promoter regions were amplified using Saos-2 genomic DNA as the template. The primers were designed 760 bp and 545 bp upstream of the translational start sites of the *ADAMTS-2* and *URG4/URGCP* genes based on the genomic sequence mentioned above. The sequences of both forward and reverse primers and the size of the amplified gene products are listed in the Table. The basic reaction mixture consisted of 100 ng/µL of each primer set, 10 mM of each dNTP, 1X PCR buffer with KCl, and 1 U of Taq DNA polymerase. As for cycling conditions, the initial denaturation was performed at 94 °C for 3 min, followed by 35 cycles of denaturation at 94 °C for 30 s, gradient annealing at 60 °C/62 °C/65 °C/68 °C for 30 s, and an extension at 72 °C for 60 s, with a final extension at 72 °C for 10 min. To amplify these GC-rich promoters, some enhancers such as ethylene glycol, formamide, 7-deaza GTP, and DMSO were added to the PCR mix. A hot-start PCR was also performed with a combination of touch-down PCR (TD-PCR). Cycling conditions consisted of two separate phases. After 10 min of initial denaturation at 95 °C (Taq DNA polymerase was added to the reaction mixture after this

Table. Forward and reverse primer sets for PCR amplification of *URG4/URGCP* and *ADAMTS-2* promoters.

Gene	Primer sequence (5' – 3')	Product length (bp)
ADAMTS-2 (−658/+110)	Frw: TTCGAAGATCCATGGCAGCCGGACT Rev: CTC GAG GACTTCAGA GGAAGAGGAACTGG	760
ADAMTS-2 (−805/+110)	Rev: CTCGAGGACAGAATGCATACGTG	915
ADAMTS-2 (−1257/+110)	Rev: CTCGAGCTGCTGTGTAAGTCGTGTTTCCC	1367
URG4/URGCP (−482/+63)	Frw: AAGCTTCGACGCCATGAGCGCAGCGAGGTC Rev: CTCGAGGTGCTTCGCATTGCCTAGAGCT	548
URG4/URGCP (−1500/+63)	Rev :ATCTCCTTCATTTCATTACTCCCTCGAG	1563
URG4/URGCP (−736/+63)	Rev: CTCGAGACAATGGTTTACATTGTCCA	1101
URG4/URGCP (−1038/+63)	Rev: CTCGAGTAATTGGTGGAGCCCAGATT	799
H-β2	Frw: TTTCTGGCCTGGAGGCTATC Rev: CATGTCTCGATCCCACTTAACT	294
GAPDH	Frw: CCCTTCATTGACCTCAACTACATGG Rev: AGTCTTCTGGGTGGCAGTGATGG	455

step or another hot-start DNA polymerase was added to the reaction mixture at the beginning of the reaction), the first cycling phase began with an annealing temperature of Tm 10 °C and then the annealing temperature was decreased by 1 °C per cycle until the Tm of the primers was reached for a total of 10–15 cycles. Phase 2 of the TD-PCR program is a generic amplification stage of 20 or 25 cycles using the final annealing temperature reached in Phase 1. PCR products were size-fractioned on 0.8% agarose gel and visualized in the gel imaging system. The verification of the PCR conditions were controlled using GAPDH primers as positive controls (Table). A negative control was also included in all PCR reactions to control for genomic DNA contamination.

2.4. Cloning and sequence analysis of the PCR fragments
PCR products were purified using a gel extraction kit (Fermentas) and cloned into the pGEMT Easy Vector system with a T:A cloning strategy. Then 3 U/µL T4 DNA ligase, 1X ligation buffer, 50 ng of the pGEM-T Easy plasmid vector, and purified PCR product were added into tubes at a final volume of 20 µL. The ligation reaction was incubated at 4 °C overnight and then transformed into the *E. coli* DH5α strain. Recombinant colonies were screened with blue-white screening and then automated DNA sequencing was performed by RefGen (Ankara).

3. Results

3.1. Putative promoter prediction and amplification of the human *ADAMTS-2* gene
A putative *ADAMTS-2* promoter region at position −1400/+110 relative to the transcription start site (TSS) was chosen for the amplification. Figure 1A indicates 1400 bp upstream of the putative promoter region of the *ADAMTS-2* gene. The GC content and CpG nucleotide composition of the putative promoter region were determined and are presented in Figure 1B using the EMBOSS CpGPlot/CpGReport/Isochore bioinformatic tool. The −658/+110 promoter region was chosen for further cloning studies to eliminate the severe secondary structure present in the 1500-bp promoter region. However, the −658/+110 region also included potential GC islands such as the region spanning 500 bp (−450/+50) from the translation start site, with an observed-to-expected ratio of CpG 1.4 (Figure 1B). The secondary structure prediction of these regions indicated the formation of 11 independent structures owing to its high internal energy in the range of −158.86 to −165.67 kcal/mol, as indicated in Figure 1C.

Because of the high GC content and complex secondary structures, an amplification of 768 bp of the 5' flanking region of the *ADAMTS-2* gene was carried out using different PCR strategies and Taq DNA polymerase enzymes to obtain a proper amplicon. Saos-2 genomic
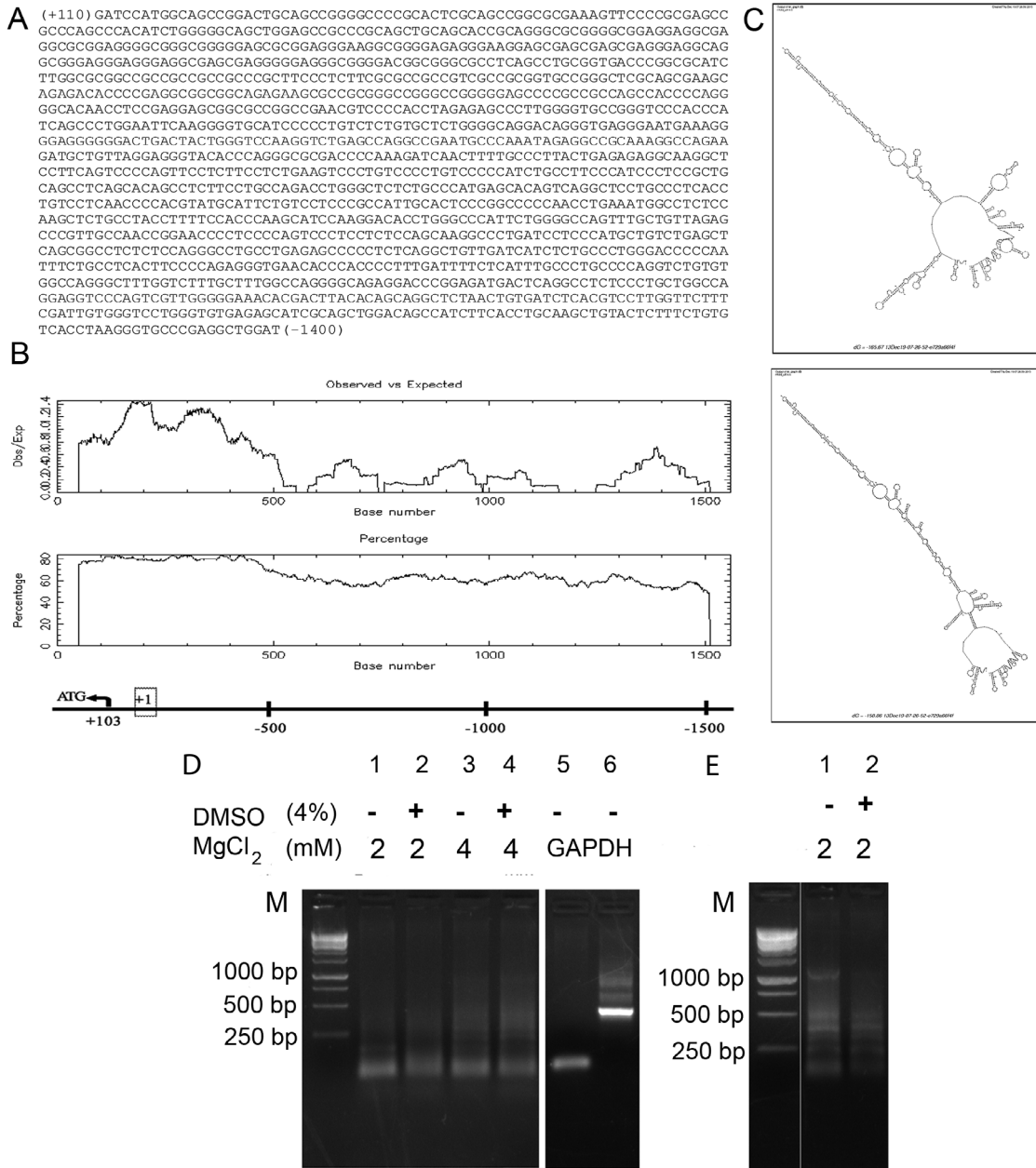
A
```
(+110)GATCCATGGCAGCCGGACTGCAGCCGGGGCCCCGCACTCGCAGCCGGCGCGAAAGTTCCCCGCGAGCC
GCCCAGCCCACATCTGGGGGCAGCTGGAGCCGCCCGCAGCTGCAGCACCGCAGGGGCGCGGGGCGGAGGAGGCGA
GGCGCGGAGGGGCGGGCGGGGGAGCGCGGAGGGGAAGGCGGGGAGAGGGAAGGAGCGAGCGAGCGAGGGAGGCAG
GCGGGAGGGAGGGAGGCGAGCGAGGGGGAGGGCGGGGACGGGGCGCCTCAGCCTGCGGTGACCCGGCGCATC
TTGGCGCGGCCGCCGCCGCCGCCCGCTTCCCTCTTCGCGCCGCCGTCGCCGCGGTGCCGGGCTCGCAGCGAAGC
AGAGACACCCCGAGGCGGCGGCAGAGAAGCGCCGCGGGCCGGGCCGGGGGAGCCCCGCCGCCAGCCACCCCAGG
GGCACAACCTCCGAGGAGCGGCGCCGGCCGAACGTCCCACCTAGAGAGCCCTTGGGGTGCCGGGTCCCACCCA
TCAGCCCTGGAATTCAAGGGGTGCATCCCCCTGTCTCTGTGCTCTGGGGCAGGACAGGGTGAGGGAATGAAAGG
GGAGGGGGGACTGACTACTGGGTCCAAGGTCTGAGCCAGGCCGAATGCCCAAATAGAGGCCGCAAAGGCCAGAA
GATGCTGGTTAGGAGGGTACACCCAGGGCGCGACCCCAAAGATCAACTTTTGCCCTTACTGAGAGAGGCAAGGCT
CCTTCAGTCCCCAGTTCCTCTTCCTCTGAAGTCCCTGTCCCCTGTCCCCCATCTGCCTTCCCATCCCTCCGCTG
CAGCCTCAGCACAGCCTCTTCCTGCCAGACCTGGGCTCTCTGCCCATGAGCACAGTCAGGCTCCTGCCCTCACC
TGTCCTCAACCCCACGTATGCATTCTGTCCTCCCGCCATTGCACTCCCGGCCCCCAACCTGAAATGGCCTCTCC
AAGCTCTGCCTACCTTTTCCACCCAAGCATCCAAGGACACCTGGGCCCATTCTGGGGCCAGTTTGCTGTTAGAG
CCCGTTGCCAACCGGAACCCCTCCCCAGTCCCTCCTCTCCAGCAAGGCCCTGATCCTCCCATGCTGTCTGAGCT
CAGCGGCCTCTCTCCAGGGCCTGCCTGAGAGCCCCCTCTCAGGCTGTTGATCATCTCTGCCCTGGGACCCCCAA
TTTCTGCCTCACTTCCCCAGAGGGTGAACACCCACCCCCTTTGATTTTCTCATTTGCCCTGCCCCAGGTCTGTGT
GGCCAGGGCTTTGGTCTTTGCTTTGGCCAGGGGCAGAGGACCCGGAGTGACTCAGGCCTCTCCCTGCTGGCCA
GGAGGTCCCAGTCGTTGGGGGAAACACGACTTACACAGCAGGCTCTAACTGTGATCTCACGTCCTTGGTTCTTT
CGATTGTGGGTCCTGGGTGTGAGAGCATCGCAGCTGGACAGCCATCTTCACCTGCAAGCTGTACTCTTTCTGTG
TCACCTAAGGGTGCCCGAGGCTGGAT(-1400)
```

B

C

D

E

Figure 1. (A) Putative human *ADAMTS-2* promoter region spanning –1400/+110 bp; (B) GC contents and CpG nucleotide composition graphs of the *ADAMTS-2* promoter. First, 500 bp of the 5' UTR and promoter region showed extremely rich G+C content (80%). The *ADAMTS-2* promoter contains a CpG island region spanning 500 bp (–450/+50) from the translation start site, with an observed-to-expected ratio of CpG of 1.4. (C) Predicted graphs of the *ADAMTS-2* secondary structure with the highest and lowest internal energy. (D) Agarose gel electrophoresis of the PCR amplification of the human *ADAMTS-2* promoter (768-bp region). Effects of $MgCl_2$ concentration and DMSO on PCR amplification with Taq DNA polymerase: 1–4: different $MgCl_2$ concentrations (2–4 mM) with or without 4% DMSO, 5: nontemplate control, 6: GAPDH as positive control. (E) PCR amplification with Dream Taq Green DNA Polymerase: 1: 2 mM $MgCl_2$ with 4% DMSO, 2: 2 mM $MgCl_2$ without DMSO. (F) PCR amplification with QIAGEN LongRange PCR Enzyme Mix (907-bp region): 1: nontemplate control, 2: GAPDH as positive control, 3: 2.5 mM $MgCl_2$ with 20% Q Solution. (G) PCR amplification with Platinum Pfx DNA Polymerase (1360-bp region): 1: 1 mM $MgSO_4$ without DMSO, 2: 1 mM $MgSO_4$ with 2% DMSO, 3: 2 mM $MgSO_4$ with 4% DMSO. (H) PCR amplification with Expand High Fidelity Taq DNA Polymerase (1360-bp region): 1: 2 mM $MgCl_2$ with 10% DMSO, 2: 2 mM $MgCl_2$ with 4% DMSO, 3: 2 mM $MgCl_2$ with 2% DMSO, 4: 2 mM $MgCl_2$ with 10% DMSO, 5: 2 mM $MgCl_2$ with 4% DMSO and Q Solution, 6: 2 mM $MgCl_2$ with 2% Q Solution. (I) Hot-start combined with touch-down PCR: 1–4: Different $MgCl_2$ concentrations (2 mM–4 mM) with or without 10% DMSO, 5: GAPDH as positive control. (J) PCR amplification using a combination of 2 mM $MgCl_2$, 50 µM/L 7-deaza GTP, 10% DMSO, and 20% betaine (0.26 M) in the reaction mixture with conventional Taq DNA Polymerase, M: 1-kb DNA ladder (Thermo SM0311).
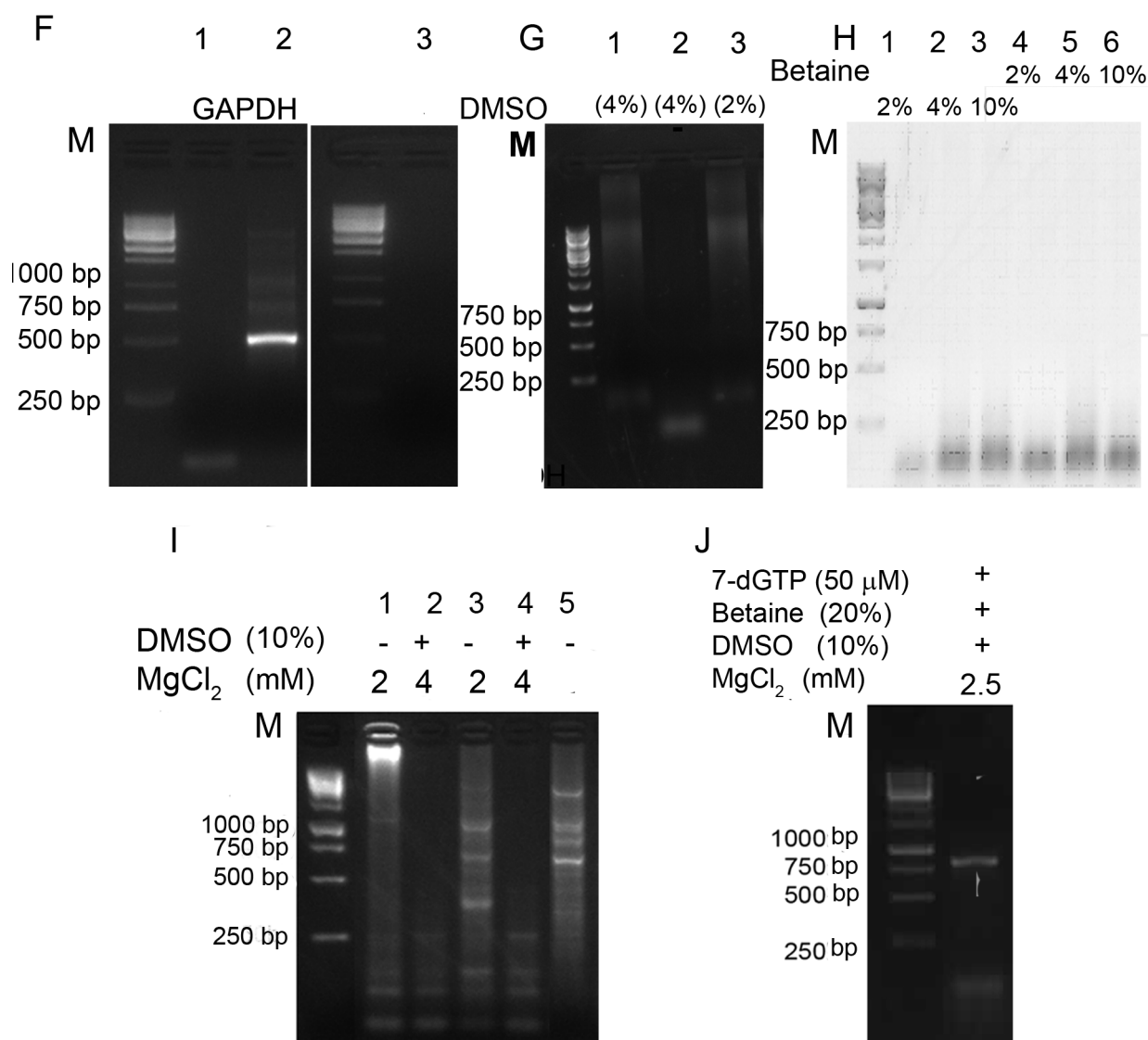
Figure 1. (Continued).

DNA was used as an initial template for the amplification of 768 bp of the putative human *ADAMTS-2* promoter with a GC content of 80%. PCR was carried out as described in Section 2. First, cycling conditions and MgCl$_2$ concentrations were optimized. Figure 1D shows the PCR amplification results of the human *ADAMTS-2* promoter for varying MgCl$_2$ concentrations within the range of 2–4 mM and 4% DMSO. As can be seen in Figure 1D, neither the optimization of MgCl$_2$ concentration nor the addition of 4% DMSO to the reaction mixture could provide a successful amplification with the conventional Taq DNA polymerase enzyme, but the positive control worked. Thus, different GC-rich targeting DNA polymerase enzyme systems were used in the reactions.

Dream Taq Green DNA polymerase (Thermo Scientific) was used in combination with 10X Dream Taq Green Buffer containing KCl, MgCl$_2$, and (NH$_4$)$_2$SO$_4$ at a ratio optimized for a robust performance in PCR. As shown in Figure 1E, the amplification resulted in products derived from regions other than the target DNA location, leading to multiple bands on the gel electrophoresis. Among the high-fidelity DNA polymerase enzyme and buffer systems, the QIAGEN LongRange provides a highly specific long-range PCR amplification, even from GC-rich and complex genomic DNA sequences. It was attempted to amplify 907 bp of the *ADAMTS-2* promoter region using 2 U of a high-fidelity enzyme mix (QIAGEN), 2.5 mM MgCl$_2$, and 1X Q solution as an enhancer. As shown in Figure 1F,

the *ADAMTS-2* promoter region could not be amplified even when the positive control worked. Another high-fidelity DNA polymerase enzyme, Platinum Pfx (Life Technologies), provides a hot-start method for problematic and/or GC-rich sequences and has 26 times higher fidelity than conventional Taq DNA polymerase. PCR was carried out to amplify 1360 bp of the *ADAMTS-2* promoter region using 1 U of Platinum Pfx DNA Polymerase, 1–2 mM MgSO$_4$, DMSO (2%–4%), and a hot-start PCR cycling condition performed as described before Section 2. As indicated in Figure 1G, the correct amplicon could not be obtained. We tried to amplify 1360 bp of the *ADAMTS-2* promoter region according to the Expand High Fidelity Enzyme System (Roche) instructions. Briefly, we used 2.6 U of DNA polymerase and 2 mM MgCl$_2$, DMSO (2%, 4%, 10%) or PCR$_x$ Enhancer solutions (2%, 4%, 10%) (Life Technologies). As indicated in Figure 1H, this high-fidelity enzyme and enhancer system could not solve the amplification problem.

A hot-start PCR method combined with TD-PCR was also performed with Taq DNA polymerase to amplify the target DNA fragment as described in Section 2. As indicated in Figure 1I, with varying MgCl$_2$ concentrations (2–4 mM) and with or without 10% DMSO, nonspecific PCR products were obtained in the reactions.

Finally, to overcome the problems associated with the amplification of the GC-rich *ADAMTS-2* promoter region, 50 µM 7-deaza GTP, 5% DMSO, and 20% betaine (0.26 M) were included in the reaction tubes at given concentrations. As presented in Figure 1J, by combining these additives, the region at position –658/+110 bp could be amplified. The resulting fragment was cloned into a pGEM-T Easy Vector system (Promega). The amplified fragment was confirmed by automated DNA sequence analysis.

3.2. PCR amplification of the putative *URG4/URGCP* promoter region

The putative *URG4/URGCP* promoter region was determined with in silico analysis, showing an extremely high GC content, with 80% in the first 300 bp as shown in Figure 2A. This region also contains two probable CpG island regions spanning 400 bp from the translation start site, with an observed-to-expected ratio ranging from 0.6 to 1 as shown in Figure 2B. In addition, a potential CpG island is present at around 900 bp of the *URG4/URGCP* promoter with an observed-to-expected ratio ranging from 1 to 2. Secondary structure prediction analysis indicated four independent secondary structures owing to its highest and lowest internal energy in the range of –78.02 to –79.6 kcal/mol as shown in Figure 2C. First we tried to amplify the region at position –1516/+63 of the putative human *URG4/URGCP* promoter by the conventional PCR method, as described in Section 2, with

the specific primer pairs shown in the Table. Figure 2D shows the PCR amplification results for 1500 bp of the human *URG4/URGCP* promoter for varying MgCl$_2$ and DMSO concentrations within the range of 1–2 mM and 2%–10%, respectively. The target DNA fragment could not be amplified with the addition of varying concentrations of DMSO in the PCR amplifications as an enhancer. Furthermore, nonspecific amplifications were observed in agarose gel electrophoresis. In order to facilitate the amplification, the target region was shortened to 548 bp and we also tried to amplify the –485/+63 region of the *URG4/URGCP* promoter by PCR using specific primers (Table) and introducing different PCR enhancers into the buffer composition. Formamide was used in reactions with a 2.5% concentration to amplify 1500 bp and 548 bp of the promoter regions. Higher concentrations of formamide were not used in the reactions because of inhibitory effects on Taq DNA polymerase and presumably other DNA polymerases (Varadaraj and Skinner, 1994). As shown in Figure 2E, the target DNA fragment could not be amplified with this composition. Another PCR additive used to enhance the amplification of GC-rich sequences is ethylene glycol; it was also introduced into the reaction mixture at different concentrations (2% 6%, 8%, and 10%) to improve PCR yields. As shown in Figure 2F, ethylene glycol could not provide a successful amplification. We also tried to amplify the *URG4/URGCP* promoter region with the Phusion High Fidelity DNA Polymerase Enzyme (Life Technologies). Briefly, to amplify 736-bp and 1038-bp promoter regions, 1X Phusion Master Mix with 1.5% DMSO and 50 µM 7-deaza GTP was used in the reactions. Unlike the standard PCR cycling conditions, denaturation steps were performed at 98 °C. As can be seen in Figure 2G, the target DNA region could not be amplified.

To demonstrate whether the optimized PCR conditions for the putative human *ADAMTS-2* promoter are applicable to other GC-rich *URG4/URGCP* promoter regions, a similar procedure was repeated. Primer sets were used specifically in the –485/+63 region of the *URG4/URGCP* promoter. Similarly, the 548-bp region of the putative human *URG4/URGCP* promoter was successfully amplified introducing the same concentrations of the PCR additives into the reaction mixture as shown in Figure 2H. The resulting fragment was purified and cloned into a pGEM-T Easy Vector system and confirmed by automated DNA sequence analysis.

4. Discussion

The expression of eukaryotic protein-coding genes can be regulated at several levels. Regulation at the level of transcription initiation is the most critical one. Promoter regions, proximal regulator elements, enhancers, silencers, insulators, or locus control regions are *cis*-acting

transcriptional regulatory DNA elements (Köçkar et al., 2001). In humans, approximately 60% of all promoters fall near a CpG island (Venter et al., 2001; Maston et al., 2006). The presence of a CpG island is the most reliable indicator for predicting the presence of a gene (Ioshikhes and Zhang, 2000; Maston et al., 2006). CpG island promoters may often lack TATA boxes and display heterogeneous TSSs. According to the in silico analysis, the *ADAMTS-2* gene has a TATA-less and probably CpG island promoter having 80% G+C content and a CpG island located in the

−450/+50 region. Furthermore, 1250 bp of the putative human *ADAMTS-2* promoter region contains perfect hairpin structures and the GC percentages of these hairpins are on average 75%. This region also includes the repeated segment GAGG, constituting 34% of the 150-bp region starting at +98. It is well known that templates that have very rich AT or GC base compositions or long short tandem repeats are difficult to amplify by PCR. Among these, templates with GC content greater than 65% frequently give very weak signals when amplified under
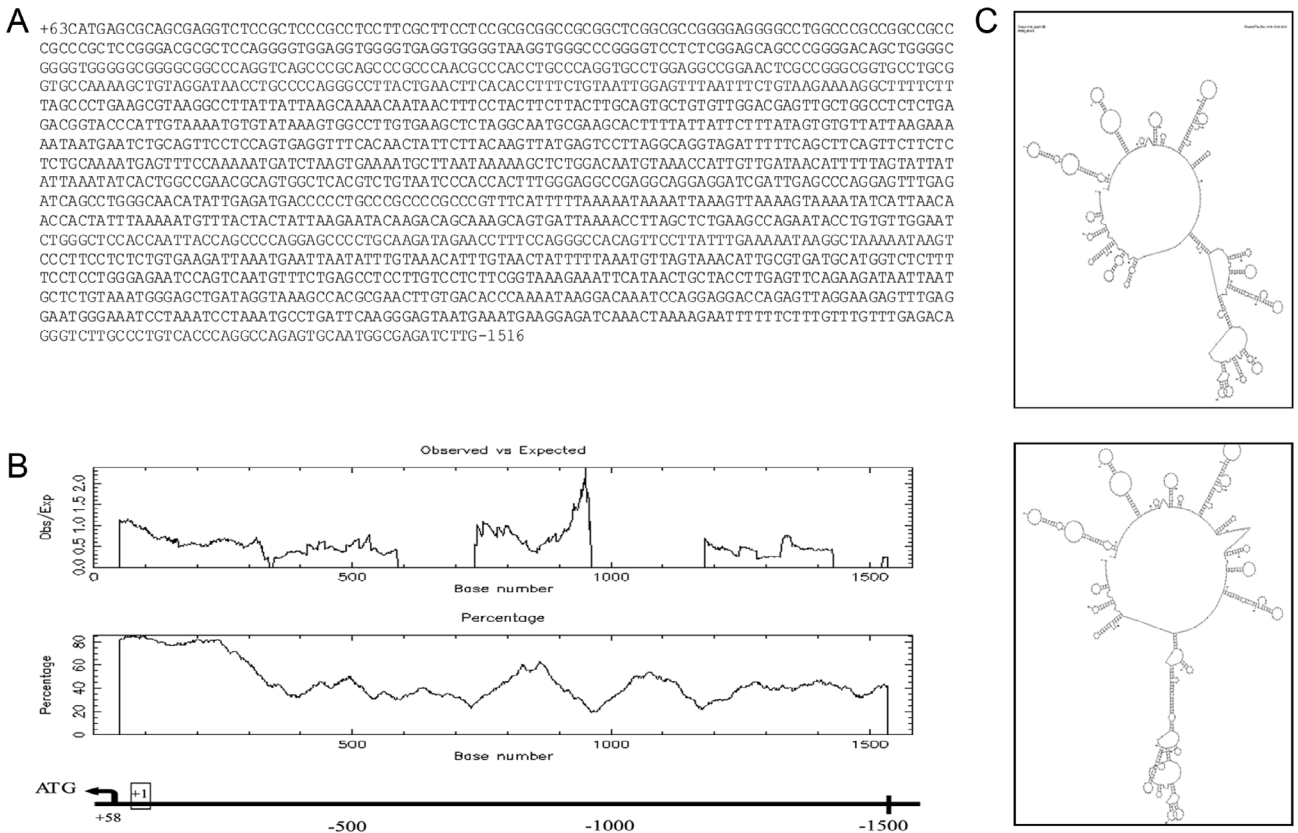
A

```
+63 CATGAGCGCAGCGAGGTCTCCGCTCCCGCCTCCTTCGCTTCCTCCGCGCGGCCGCGGCTCGGCGCCGGGGAGGGGCCTGGCCCGCCGGCCGCC
CGCCCGCTCCGGGACGCGCTCCAGGGGTGGAGGTGGGGTGAGGTGGGGTAAGGTGGGCCCGGGGTCCTCTCGGACCAGCCCGGGGACAGCTGGGGC
GGGGTGGGGGCGGGGCGGCCCAGGTCAGCCCGCAGCCCGCCCAACGCCCACCTGCCCAGGTGCCTGGAGGCCGGAACTCGCCGGGCGGTGCCTGCG
GTGCCAAAAGCTGTAGGATAACCTGCCCCAGGGCCTTACTGAACTTCACACCTTTCTGTAATTGGAGTTTAATTTCTGTAAGAAAAGGCTTTTCTT
TAGCCCTGAAGCGTAAGGCCTTATTATTAAGCAAAACAATAACTTTCCTACTTCTTACTTGCAGTGCTGTGTTGGACGAGTTGCTGGCCTCTCTGA
GACGGTACCCATTGTAAAATGTGTATAAAGTGGCCTTGTGAAGCTCTAGGCAATGCGAAGCACTTTTATTATTCTTTATAGTGTGTTATTAAGAAA
AATAATGAATCTGCAGTTCCTCCAGTGAGGGTTTCACAACTATTCTTACAAGTTATGAGTCCTTAGGCAGGTAGATTTTTCAGCTTCAGTTCTTCTC
TCTGCAAAATGAGTTTCCAAAAATGATCTAAGTGAAAATGCTTAATAAAAAGCTCTGGACAATGTAAACCATTGTTGATAACATTTTTAGTATTAT
ATTAAATATCACTGGCCGAACGCAGTGGCTCACGTCTGTAATCCCACCACTTTGGGAGGCCGAGGCAGGAGGATCGATTGAGCCCAGGAGTTTGAG
ATCAGCCTGGGCAACATATTGAGATGACCCCCTGCCCGCCCCGCCCGTTTCATTTTTAAAAATAAAATTAAAGTTAAAAGTAAAATATCATTAACA
ACCACTATTTAAAAATGTTTACTACTATTAAGAATACAAGACACGAAGACAGTGATTAAAACCTTAGCTCTGAAGCCAGAATACCTGTGTTGGAAT
CTGGGCTCCACCAATTACCAGCCCCAGGAGCCCCTGCAAGATAGAACCTTTCCAGGGCCCACAGTTCCTTATTTGAAAAATAAGGCTAAAAATAAGT
CCCTTCCTCTCTGTGAAGATTAAATGAATTAATATTTGTAAACATTTGTAACTATTTTTAAATGTTAGTAAACATTGCGTGATGCATGGTCTCTTT
TCCTCCTGGGAGAAATCCAGTCAATGTTTCTGAGCCTCCTTGTCCTCTTCGGTAAAGAAATTCATAACTGCTACCTTGAGTTCAGAAGATAATTAAT
GCTCTGTAAATGGGAGCTGATAGGTAAAGCCACGCGAACTTGTGACACCCAAAATAAGGACAAATCCAGGAGGACCAGAGTTAGGAAGAGTTTGAG
GAATGGGAAATCCTAAATCCTAAATGCCTGATTCAAGGGAGTAATGAAATGAAGGAGATCAAACTAAAAGAATTTTTTCTTTGTTTGTTTGAGACA
GGGTCTTGCCCTGTCACCCAGGCCAGAGTGCAATGGCCGAGATCTTG -1516
```

C

B

Observed vs Expected

Percentage

ATG +1
+58   −500   −1000   −1500

Figure 2. (A) Putative human *URG4/URGCP* promoter region spanning −1516/+63 bp. (B) Bioinformatic analysis revealed a promoter region around −1500 bp including TSS at the +58 bp position for the *URG4/URGCP* gene, and 300 bp of the promoter region showed extremely high G+C content (80%). The *URG4/URGCP* promoter contains CpG island regions of minimum length of 100 bp spanning 150 bp (−100/+50) from the translation start site, with an observed-to-expected ratio of CpG of 0.6. (C) Two predicted secondary structure graphs of the *URG4/URGCP* promoter site. The internal energy required for formation of these constructs is indicated as kcal/mol and the lowest and the highest forms of the secondary structure are shown. (D) Agarose gel electrophoresis of the PCR amplification of the human *URG4/URGCP* promoter (−1516/+63) region: 1: Positive control with GAPDH primers, 2–3: 2% DMSO with 1 mM and 2 mM MgCl$_2$ concentrations, 4–5: 10% DMSO with 1 mM and 2 mM MgCl$_2$ concentrations, 6: nontemplate control. (E) Agarose gel electrophoresis of 1500 bp and 548 bp of human *URG4/URGCP* promoter amplifications with formamide: 1: Positive control with h-β2 primers, 2: 2.5% formamide with 1 mM MgCl$_2$ (1500-bp region), 3: 2.5% formamide with 1 mM MgCl$_2$ concentrations (548-bp region). (F) Agarose gel electrophoresis of the PCR amplification of the human *URG4/URGCP* promoter (548-bp region): 1: Positive control with GAPDH primers, 2–3: 6% ethylene glycol with 1–2 mM MgCl$_2$, 4–5: 8% ethylene glycol with 1–2 mM MgCl$_2$, 6–7: 10% ethylene glycol with 1–2 mM MgCl$_2$, 8–9: 2% ethylene glycol with 1–2 mM MgCl$_2$ concentration, 10: PCR without ethylene glycol with 2 mM MgCl$_2$. (G) Agarose gel electrophoresis of the PCR amplification of the human *URG4/URGCP* promoter (736- and 1038-bp regions) with Phusion High Fidelity PCR Master Mix: 1–2: 1X master mix with 1.5% DMSO and 50 µM 7-deaza GTP (736-bp region and 1038-bp, respectively). (H) PCR amplification with conventional Taq DNA Polymerase using 50 µM 7-deaza GTP, 10% DMSO, and 20% betaine (0.26 M) in the reaction mixture, M: 1-kb DNA ladder (Thermo SM0311).
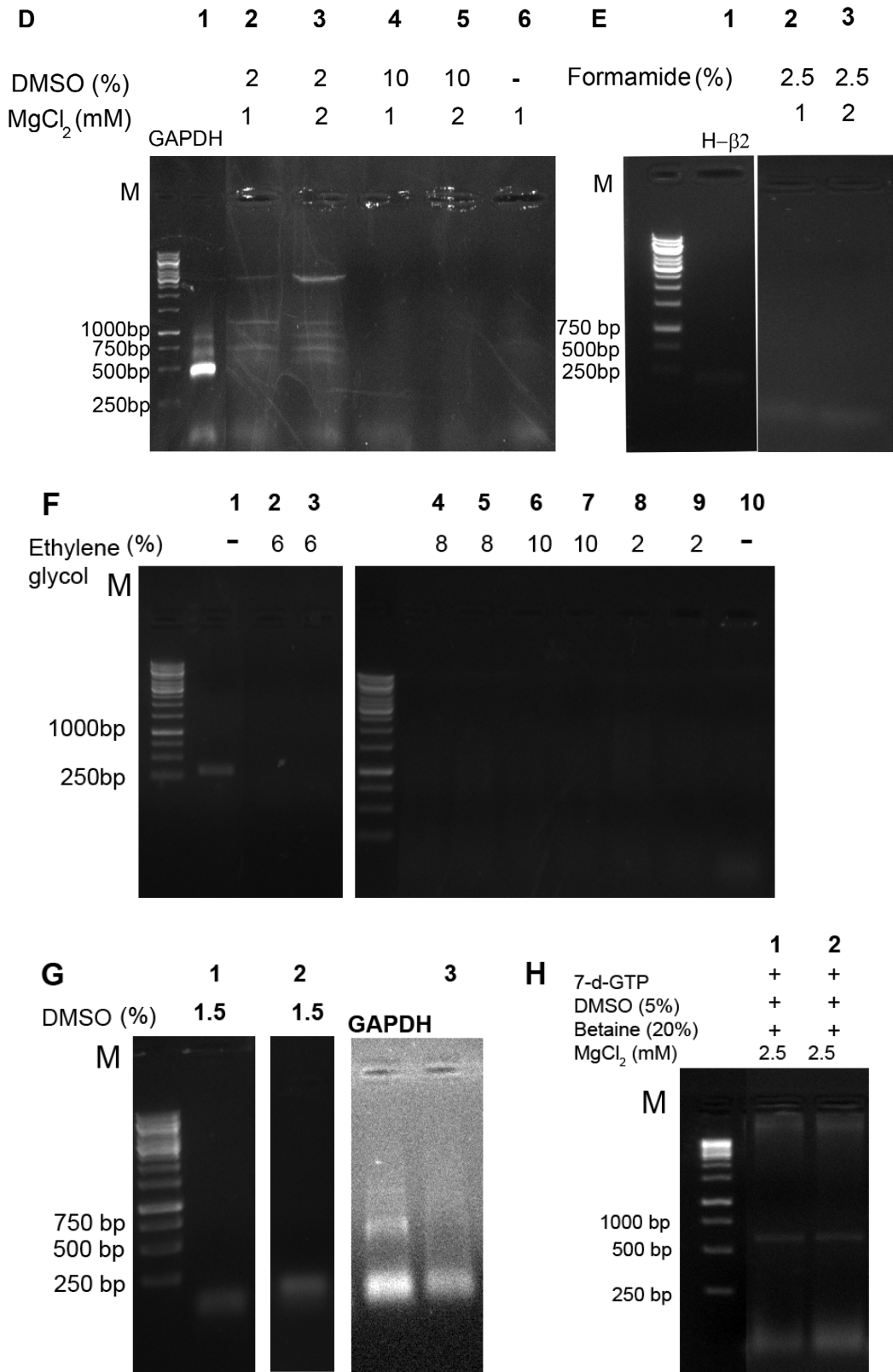
D

	1	2	3	4	5	6	**E**		1	2	3
DMSO (%)		2	2	10	10	-	Formamide (%)			2.5	2.5
MgCl$_2$ (mM)		1	2	1	2	1				1	2

GAPDH

H–β2

F

	1	2	3		4	5	6	7	8	9	10
Ethylene (%) glycol	-	6	6		8	8	10	10	2	2	-

G

	1	2	3
DMSO (%)	1.5	1.5	
			GAPDH

H

	1	2
7-d-GTP	+	+
DMSO (5%)	+	+
Betaine (20%)	+	+
MgCl$_2$ (mM)	2.5	2.5

Figure 2. (Continued).

203

standard conditions. Furthermore, the amplification may result in products derived from regions other than the target DNA location (Sahdev et al., 2007). Because of the high GC content and complex hairpin structures, it was hard to amplify the *ADAMTS-2* and *URG4/URGCP* promoter regions with standard PCR conditions. Despite the optimization of the MgCl$_2$ concentration, buffer pH, and cycling conditions, the target regions could not be amplified. Different additives were introduced into the reaction buffers. In addition to KCl- and MgCl$_2$-containing buffers, buffers with (NH$_4$)$_2$SO$_4$ and MgSO$_4$ were also used in PCR. Ammonium ions can make an amplification reaction more tolerant of nonoptimal conditions and MgSO$_4$ helps to destabilize secondary structures in DNA (Seifi et al., 2012). Previous studies have indicated that formamide and ethylene glycol can be used as PCR enhancers. These compounds were used in varying concentrations in the reactions. Formamide interferes with the formation of hydrogen bonds between the two DNA strands (Varadaraj and Skinner, 1994). Ethylene glycol is able to decrease the melting temperatures of DNA and modulates the annealing process of GC-rich templates (Zhang et al., 2009). These additives and enhancers could not provide a successful amplification.

We tried to amplify these problematic promoter fragments with GC-rich targeted high-fidelity DNA polymerases as well as the conventional Taq DNA polymerase. A hot-start PCR approach was used in combination with a TD-PCR; however, that did not help in getting the correct amplification. In previous studies betaine and DMSO or betaine, DMSO, and 7-deaza GTP combinations were found to enhance amplification by resolving the complex secondary structure formation of the GC-rich regions (Henke et al., 1997; Kang et. al.,

2005; Musso et al., 2006; Ralser et al., 2006). In the attempt to amplify 760 bp of the putative human *ADAMTS-2* promoter region, a 5% DMSO, 20% betaine (0.26 M), and 50 µM 7-deaza GTP combination was included in the reaction mixture. Specific and exclusive amplification of the target sequence was obtained with the addition of these organic molecules, as expected.

In order to test whether the optimized concentrations of these three additives could be useful to amplify another GC-rich region, the same procedure was used to amplify the *URG4/URGCP* promoter region containing 80% GC base composition and perfect hairpin structures. When approximately 1500 bp upstream of the *URGCP* gene was analyzed bioinformatically, a perfect hairpin structure between the sequences at positions 203 bp and 903 bp was observed with a GC percentage of 100%. A repeated segment GGGG was also present along the 150-bp region starting at 61 bp. Because of the high GC content and complex hairpin structures, the 1500 bp upstream of the *URGCP* gene region could not be amplified by conventional PCR or by the addition of the different additives into the reactions. Only the region at +63/–485, about 548 bp long, could it be amplified using the same additives at the same concentrations. In both cases, we could obtain specific amplification, demonstrating that the optimized concentrations of these additives can be used for the amplification of GC-rich DNA fragments.

References

Baskaran N, Kandpal RP, Bhargava AK, Glynn MW, Bale A, Weissman SM (1996). Uniform amplification of a mixture of deoxyribonucleic acids with varying GC content. Genome Res 6: 633–638.

Chakrabarti R, Schutt CE (2001). The enhancement of PCR amplification by low molecular-weight sulfones. Gene 274: 293–298.

Colige A, Nuytinck L, Hausser I, van Essen AJ, Thiry M, Herens C, Adès LC, Malfait F, Paepe AD, Franck P et al. (2004). Novel types of mutation responsible for the dermatosparactic type of Ehlers-Danlos syndrome (type VIIC) and common polymorphisms in the *ADAMTS2* gene. J Invest Dermatol 123: 656–663

Colige A, Ruggiero F, Vandenberghe I, Dubail J, Kesteloot F, Van Beeumen J, Beschin A, Brys L, Lapière CM, Nusgens B (2005). Domains and maturation processes that regulate the activity of ADAMTS-2, a metalloproteinase cleaving the aminopropeptide of fibrillar procollagens types I-III and V. J Biol Chem 2005; 280: 34397–34408.

Dodurga Y, Avcı CB, Susluer SY, Satıroğlu Tufan NL, Gündüz C (2012). The expression of URGCP gene in prostate cancer cell lines: correlation with rapamycin. Mol Biol Rep 39: 10173–10177.

Dubail J, Kesteloot F, Deroanne C, Motte P, Lambert V, Rakic JM, Lapière C, Nusgens B, Colige A (2010). ADAMTS-2 functions as anti-angiogenic and anti-tumoral molecule independently of its catalytic activity. Cell Mol Life Sci 67: 4213–4232.

Hecker KH, Roux HR (1996). High and low annealing temperatures increase both specificity and yield in touchdown and stepdown PCR. Biotechniques 20: 478–485.

Henke W, Herdel K, Jung K, Schnorr D, Loening SA (1997). Betaine improves the PCR amplification of GC-rich DNA sequences. Nucleic Acids Res 25: 3957–3958.

Innis MA, Gelfand DH (1990). PCR Protocols: A Guide to Methods and Applications. San Diego, CA, USA: Academic Press.

Ioshikhes IP, Zhang MQ (2000). Large-scale human promoter mapping using CpG islands. Nat Genet 26: 61–63.

Kang J, Lee MS, Gorenstein DG (2005). The enhancement of PCR amplification of a random sequence DNA library by DMSO and betaine: application to in vitro combinatorial selection of aptamers. J Biochem Biophys Methods 64: 147–151.

Köçkar FT, Foka P, Hughes TR, Kousteni S, Ramji DP (2001). Analysis of the *Xenopus laevis* CCAAT-enhancer binding protein alpha gene promoter demonstrates species-specific differences in the mechanisms for both auto-activation and regulation by Sp1. Nucleic Acids Res 29: 362–372.

Korbie DJ, Mattick JS (2008). Touchdown PCR for increased specificity and sensitivity in PCR amplification. Nat Protoc 3: 1452–1456.

Mamedov TG, Pienaar E, Whitney SE, TerMaat JR, Carvill G, Goliath R, Subramanian A, Viljoen HJ (2008). A fundamental study of the PCR amplification of GC-rich DNA templates. Comput Biol Chem 32: 452–457.

Maston GA, Evans SK, Green MR (2006). Transcriptional regulatory elements in the human genome. Annu Rev Genomics Hum Genet 7: 29–59.

McDowell DG, Burns NA, Parkes HC (1998). Localised sequence regions possessing high melting temperatures prevent the amplification of a DNA mimic in competitive PCR. Nucleic Acids Res 26: 3340–3347.

Musso M, Bocciardi R, Parodi S, Ravazzolo R, Ceccherini I (2006). Betaine, dimethyl sulfoxide, and 7-deaza-dGTP, a powerful mixture for amplification of GC-rich DNA sequences. J Mol Diagn 8: 544–550.

Mutter GL, Boynton KA (1995). PCR bias in amplification of androgen receptor alleles, a trinucleotide repeat marker used in clonality studies. Nucleic Acids Res 23: 1411–1418.

Pomp D, Medrano JF (1991). Organic solvents as facilitators of polymerase chain reaction. Biotechniques 10: 58–59.

Ralser M, Querfurth R, Warnatz HJ, Lehrach H, Yaspo ML, Krobitsch S (2006). An efficient and economic enhancer mix for PCR. Biochem Biophys Res Commun 347: 747–751.

Sahdev S, Saini S, Tiwari P, Saxena S, Singh SK (2007). Amplification of GC-rich genes by following a combination strategy of primer design, enhancers, and modified PCR cycle conditions. Mol Cell Probes 21: 303–307.

Seifi T, Ghaedi K, Salamian A, Tanhaei S, Safari F, Hojati Z, Tavassoli M, Baharvand H, Nasr Esfahani MH (2012). Amplification of GC-rich putative mouse pep promoter using betaine and DMSO in ammonium sulfate polymerase chain reaction buffer. Avicenna J Med Biotech 4: 206–209.

Sun YG, Hegamyer G, Colburn NH (1993). PCR-direct sequencing of a GC-rich region by inclusion of 10% DMSO: application to mouse c-jun. Biotechniques 15: 372–374.

Tufan NL, Lian Z, Liu J, Pan J, Arbuthnot P, Kew M, Clayton MM, Zhu M, Feitelson MA (2002). Hepatitis Bx antigen stimulates expression of a novel cellular gene, URG4, that promotes hepatocellular growth and survival. Neoplasia 4: 355–368.

Turner SL, Jenkins FJ (1995). Use of deoxyinosine in PCR to improve amplification of GC-rich DNA. Biotechniques 19: 48–52.

Varadaraj K, Skinner DM (1994). Denaturants or cosolvents improve the specificity of PCR amplification of a G + C-rich DNA using genetically engineered DNA polymerases. Gene 11: 1–5.

Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA et al. (2001). The sequence of the human genome. Science 291: 1304–1351.

Wilson IG (1997). Inhibition and facilitation of nucleic acid amplification. Appl Environ Microbiol 63: 3741–3751.

Zhang Z, Yang X, Meng L, Liu F, Shen C, Yang W (2009). Enhanced amplification of GC-rich DNA with two organic reagents. BioTechniques 47: 775–779.