G-quadruplex formation of FXYD1 pre-mRNA indicates the possiblity of regulating expression of its protein product

Hansraj Dhayan, Anwar R Baydoun and Andreas Kukol[*]
School of Life and Medical Sciences, University of Hertfordshire, Hatfield, AL10 9AB, United Kingdom

* to whom correspondence should be addressed: Tel: +44(0) 1707 284543; FAX: +44(0) 1707 284870; Email: a.kukol@herts .ac.uk

Short title: G-quadruplex formation of FXYD1

Abstract

G-quadruplexes are higher-order nucleic acid structures formed of square-planar arrangements of four guanine bases held together by Hoogsteen-type hydrogen bonds. Stacks of guanine tetrads are stabilised by intercalating potassium ions. FXYD1 encodes for phospholemman, a regulatory subunit of the cardiac Na^+/K^+-ATPase. Computational sequence analysis of FXYD1 pre-mRNA predicted the formation of stable intramolecular G-quadruplexes in human and orthologue sequences. Multiple sequence alignment indicated that G-rich sequences are conserved in evolution suggesting a potential role of G-quadruplexes in FXYD1 gene expression. The existence of a non-functional alternative splicing product indicated that the G-quadruplex formation may control alternative splicing. Quadruplex formation of human and bovine oligonucleotides was confirmed *in vitro* by native polyacrylamide gel electrophoresis and intrinsic fluorescence emission spectroscopy. Taking together the evolutionary conservation of G-quadruplex forming sequences with the confirmation of G-quadruplex formation in vitro by two FXYD1 homologues the results point to a potential role of these structures in regulating the expression of FXYD1 and thus regulate indirectly the activity of the cardiac Na^+/K^+ -ATPase.

Keywords: G-quadruplex; pre-mRNA; splicing; sodium-pump; fluorescence; bioinformatics

Introduction

Over the last decade *in-silico* analysis of the human genome has revealed many potential sequences that can fold into G-quadruplexes, higher order secondary structures formed by nucleic acids rich in guanine residues. These are formed under cationic conditions[1, 2], with a large fraction falling into gene promoter regions, UTR regions as well as exonic and intronic partsof pre-mRNAs [3-5]. G-quartets,originally characterised by Gellert, Lipsett & Davies [6], are square planar arrangements of four guanine residues stabilised by hydrogen bonds. Stacks of G-tetrads form G-quadruplexes[7], which are further stabilised by cations, in particular potassium, located between a pair of tetrads.

Some of the proposed biological roles of G-quadruplexes include up-regulation of genes by keeping promoter or upstream regions of genes in a more open conformation, providing easy access for transcription factors to bind [8]. Alternatively, down-regulation of genes has also been proposed to be associated with G-quadruplexes. This can be used for therapeutic intervention, for example the proto-oncogene *c-myc* showed reduced expression following the stabilisation of G-quadruplexes in its promoter [9]. Apart from DNA, G-quadruplex formation of RNA has been reported. G-quadruplex formation in RNA is as important as in DNA and the high occurrence of G-quadruplex in UTR regions lead to hypothesizing on their role as translational regulators [10, 11]. Kumari*et al.* [12] reported that a G-quadruplexwithin the 5'-UTR of the *NRAS* oncogene reduced expression of the latter. G-quadruplexes have been confirmed to regulated gene expression by influencing pre-mRNA splicing [16, 17]. Marcel *et al.* [13] reported that the formation of a G-quadruplex in the pre-mRNA of tumour suppressor protein, P53, leads to alternative splicing. G-quadruplex formation in hTERT intron 6 led to alternative splicing patterns, which caused down regulation of the activity of telomerase in A549 carcinoma cells [14].

Previous studies support the notion that conformational changes within mRNA molecules have the potential of regulating protein formation [15, 16] such as phospholemman (PLM) encoded by the *FXYD1* gene. Proteins of the FXYD family are involved in regulating ion transport across membranes [17, 18]. PLM (FXYD1) that is mainly expressed in cardiac tissue acts as a modulator of the cardiac Na^+/K^+-ATPase (NKA) [19-21]. PLM is fundamentally important in the physiological regulation of cardiomyocytes and as such a potential therapeutic target for the treatment of cardiac hypertrophy and heart failure [22]. The 72-residue single-span transmembrane protein forms alpha helical tetramers *in vitro* [23] and *in vivo* [24, 25].The tetramer has the potential to interact with NKA, which was proposed to lead to a subsequent dissociation of the tetramer [26]. Further *in vivo* studies have shown that there is a delicate balance between monomer and tetramer, which also depends on the phosphorylation of PLM [25]. X-ray crystallography studies of NKA in other tissues and species have shown that monomeric FXYD1 (PLM) homologs, such as FXYD2 in porcine renal tissue [27] and FXYD10 in the shark rectal gland [28] act as a third subunit of NKA.

In this study we investigated the potential of a control mechanism at the level of FXYD1 gene expression. As outlined above G-quadruplex formation of RNA has the ability to control gene expression by interference with splicing of the pre-mRNA. Using a combination of bioinformatics-based sequence analysis, energy calculations and laboratory-based experiments we show that FXYD1 pre-mRNA has the potential to form intramolecular G-quadruplexes and that this potential has been conserved in evolution.

Materials and Methods

Computational analysis

The online G-quadruplex prediction algorithm QGRS mapper (http://bioinformatics.ramapo.edu/QGRS/analyze.php) [29] was used to analyse the G-quadruplex forming potential of *FXYD1* pre-mRNA sequences. Pre-mRNA sequences were retrieved from the Ensembl website. Accession numbers: *H. sapiens (ENST00000351325), M. musculus (ENSMUSG00000036570), C. familiaris(ENSCAFT00000011368), P. troglodytes (ENSPTRT00000020057), B. taurus (ENSBTAG00000017816), R. norvegicus (ENSRNOG00000021079), M. domestica (ENSMODT00000033163), F. catus (ENSFCAG00000008890), O. garnettii (ENSOGAG00000014401), E. caballus (ENSECAG00000014815), A. melanoleuca (ENSAMEG00000000212), P. abelii (ENSPPYG00000009851), O. cuniculus (ENSOCUG00000022123), G. gorilla (ENSGGOT00000026217), S. scrofa (ENSSSCT00000027321), O. aries (ENSOARG00000004709), T. truncates (ENSTTRG00000001446)*. The DNA sequence of a positive control DNA known to form an intramolecular G-quadruplex was obtained from the RCSB Protein Data Bank(PDB-ID: 2KM3). Two negative controls were generated using a sequence randomizer (http://www.cellbiol.com/python.html). The negative control_A has a similar base composition as the positive control, while the negative control_B has the same base composition as the highest scoring human FXYD1 G-quadruplex forming sequence (GQS). The Vienna RNA Package version 2.1.2 [30] was used to predict G-quadruplex and other secondary structure formation. Multiple sequence alignment was performed by multiple alignment of all 17 orthologous pre-mRNA sequences with Clustal-Omega version 1.2.0 [31]. Comparison of the variant FXYD1-009 (retrieved from Ensembl website accession

number: ENST00000589121) and FXYD1-001 transcripts was carried out using the '1000 Genomes' Transcript Comparison option [32]. The effect of mutations in the pre-mRNA sequence were analysed for potential G-quadruplex forming sequences using QGRS mapper.

Oligonucleotides

Purified oligonucleotides were purchased from EurogentecLtd. (Southampton, UK) and used without further modification. Positive control DNA (AGGGCTAGGGCTAGGGCTAGGG) was purified by reversed phase cartridge, negative control_ADNA (CGTGGGGAGATTGGGGA GCGCA) and negative control_B DNA (GGTGTGCGTGTGCGAGCGAGAGAGAGUGG) were both purified by reversed phase cartridge, Human *FXYD1* RNA (GGGAGACUGCGGGUAUUCU GGGGAGAGGG) and bovine *FXYD1* RNA (GGGCGCGGGGGGGUCGGGGAUCGGG) were both purified by reversed phase high performance liquid chromatography. Samples were prepared in G-quadruplex folding buffer and normal buffer to act as controls. G-quadruplex folding buffer was made to have final concentration of K^+ of either 0.05 M or 0.1 M and TrisOAC at 0.02 M, pH 7.5. K^+ solution was made to contain a mixture of KOAc and KCl. Normal buffer contained 0.02 M TrisOAC pH 7.5 only. Solutions were made using RNAase free water that was autoclaved prior to use. The buffer was added to the lyophilised oligonucleotides and shock frozen for future use. Working dilutions were prepared and heated to 95^0C for 10 mins. Following heating, the control samples and the negative DNA controls in the presence of K^+ containing buffer were placed on ice to prevent G-quadruplex formation, while the positive DNA control, human and bovine FXYD1 RNA sequences incubated in G-quadruplex buffer were allowed to anneal for 2.5 hours to allow G-quadruplex formation.

Native polyacrylamide gel electrophoresis (PAGE)

Separate gels were prepared for samples incubated in G-quadruplex (0.05 M and 0.1 M K^+) and non-G-quadruplex folding buffer. The gels were prepared using 40% (w/v) acrylamide (acrylamide 38% (w/v), bisacrylamide 2% (w/v)), 10% ammonium persulfate (APS), sterile distilled water (s.d.w), 10 x TRIS-borate-EDTA (TBE), and (N,N,N',N'-Tetramethylethylenediamine) TEMED. Gels used for G-quadruplex samples were prepared with 10 X TBE + KCl + KOAc made to contain the same concentration of K^+ as in the folding buffer. Samples used for electrophoresis were made at a final oligonucleotide concentration of 3 µM at a volume of 50 µl. 8 µl of tracking dye containing 1 x TBE + 20% (w/v) sucrose + 10% (w/v) Ficoll + 10mM EDTA and 0.25% (w/v) bromophenol blue was added to each tube and the contents of each tube were thoroughly mixed. 12 µl of sample were loaded into each well. The buffers were pre-chilled at 4^oC and electrophoresis was performed at 140 V for 3-4 hours. The gels were stained using SYBR Green I RNA stain S9430 and SYBR Green I nucleic S32717 and were exposed at 254 nm for 15.5s for digital snapshots. The R_f value was calculated as the ratio of the distance migrated by each samples to the distance migrated by the tracking dye. The R_f values were calculated using the software Gene Tool Syngene (Synoptics Ltd).

Fluorescence spectroscopy

Samples were prepared at a final oligonucleotide concentration of 1.5 µM for RNA species and 5.0 µM for DNA species, which was above the minimum detectable concentrations for either RNA or DNA. Buffers for non G-quadruplex samples were 0.02 M TrisOAC only and buffers for G-quadruplex samples were 0.02 M TrisOAC + 0.1 M K^+. The fluorescence emission spectrum was measured over the wavelength range 300-500 nm (step size 0.2 nm) using a Perkin Elmer LS 55 fluorimeter with a fluor cuvette Type C quartz glass having a 10 mm light path (Whatman International Ltd Maidstone, England). Measurements were repeated three times to ensure reproducibility. Samples were excited at 260 nm and both excitation and emission slit widths were set at 5 nm. The scan rate was 150 nm/min. Emission spectra of buffers were also measured. The original spectra were smoothed with a moving average of 30 data points. UV/VIS spectra of each sample in a wavelength range of 200-400 nm were recorded using a dual beam UV/VIS CARY 100 (Varian Inc.) spectrophotometer.

Results and Discussion

In-sillico analysis of *FXYD1* pre-mRNA and orthologuesrevealed potential G-quadruplex forming sequences.

The sequences of FXYD1 pre-mRNA and orthologues were obtained from the Ensembl database and analysed with QGRS mapper[29]. QGRS-mapper scans the sequence for characteristic G-rich motifs and provides a score that indicates the potential for G-quadruplex formation.

The results in table 1 show that all seventeen orthologues have sequences that can potentially fold into a G-quadruplex, in the following referred to as predicted G-Quadruplex forming Sequences (GQS). The G-scores are in the same region as the positive control, which has a G-score of 42. The positive control is a DNA oligonucleotide definitively known to form G-quadruplexes with the 3-dimensional structure obtained by NMR spectroscopy [33]. The negative controls are randomised DNA oligonucleotides of the same base composition as the positive control (A) and human FXYD1 GQS (B) respectively. The negative controls did not show G-quadruplex forming potential with a G-score of zero, as would be expected due to the absence of four stretches of consecutive guanine bases. With the exception of *S. Scrofa, R. Norvegicus* and *G. gorilla*, all orthologue sequences have comparable G-scores to the positive control. Most of the QGS listed in table 1 occur in intronic parts of the pre-mRNA. For instance the *H. sapiens*GQS is intronic between exons 6 and 7, a position shared with four other orthologues.

Prediction of potential GQS relies on Kikin and colleagues' folding rule: $G_xN_{y1}G_xN_{y2}G_xN_{y3}G_x$ [29]. G represents guanine residues taking part in G-tetrad formation and x is the number of successive guanines. According to the folding rule, x is always constant and should be at least two, as a minimum of two quartets is required to stack on top of each other to form a G-quadruplex. The lower scoring GQS will have two G-tetrads in its G-quadruplex entity, similar to *R. Norvegicus* and *G. gorilla's* GQS. Stability is enhanced by more G-quartets or the number of guanine residues participating in the stacks. N is representative of the other bases involved in the loops of the G-quadruplex, N can be any base including guanine. y1, y2

& y3 are the frequency of the different residues that participate in the three different loops, and can vary. Loop size can affect stability of G-quadruplex ensembles but not as much as the frequency of guanine bases. GQSs having at least three guanine tetrads and loops of equal length connecting them will be highly stable and have high G-scores such as the positive control. The GQS predicted for *S. scrofa* has a G-score of 83, which makes its G-quadruplex twice as stable as that of the +VE control. *S. scrofa* has five G-quartets that form a G-quadruplex of equal loop length of one base each, which make the G-quadruplex highly stable, hence yielding a high G-score. Regardless of the equally sized loops connecting the two G-tetrads inall six of *R. norvegicus*' and *G. gorilla's*GQS, theirscoresarehalf than that of the positive control,due to two G-tetrads only present in the GQS.

RNA secondary structure and minimum free energy calculations

Using a combination of RNA secondary structure calculations and G-quadruplex minimum free energy (MFE) calculations as implemented in the Vienna RNA Package 2.1 it was suggested that the majority of G-quadruplex forming sequences (GQS) may form canonical secondary structures instead [34]. Therefore, we calculated the MFE of the predicted most stable structure as well as the MFE of suboptimal structures for the human and bovine GQS shown in table 1 using the Vienna RNA package.

The results in table 2 show that the positive control and *Bos taurus* QGS are clearly most stable in the G-quadruplex form. For the *Homo Sapiens* QGS the situation is more complex, the most stable form with MFE = -4.3 kcal/mol is a secondary structure based on canonical base pairing, closely followed by another secondary structure, while the G-quadruplex structure has an MFE of -3.51 kcal/mol. These MFE values can be used to estimate the relative amountof secondary structure in equilibrium with G-quadruplex structure assuming that the MFE values correspond approximately to the free enthalpy of folding ΔG. Defining $K_1 = [G_4]/[U]$ as the equilibrium constant for the formation of the G-quadruplex structure from the unfolded state, and $K_2 = [2°]/[U]$ as the equilibrium constant for the formation of secondary structure, then $K_2/K_1 = [2°]/[G_4]$ can be calculated from the MFE values as $K_2/K_1 = \exp(-(MFE_{2°} - MFE_{G4})/RT) = 3.8$ (with R = $1.987 \cdot 10^{-3}$ kcal K^{-1} mol^{-1}, T = 298 K) meaning that the concentration of the secondary structure is approximately four-fold higher than the G-quadruplexstructure.Note that inaccuracies in MFE values are partially compensated for by considering the difference in the calculation above. This result shows that in case of *Homo Sapiens* QGS there is a significant proportion of G-quadruplex structure present, which may increase upon increasing potassium concentration. Note that the energy model of the Vienna RNA package for G-quadruplex structures did not take the potassium ion concentration into account [34].

G-quadruplex forming sequences (QGS) are evolutionary conserved among orthologous FXYD1 pre-mRNA.

The human pre-mRNA was aligned withthe pre-mRNA sequences of seventeen orthologues. The results shown in figure 1 reveal G-rich sequences that align with the human GQS (shown in the first line of table 1), except in *Monodelphisdomestica*. In particular the human GQS aligns with GQS from Canis lupus and Pan troglodytes, while the other organisms have GQS in other parts of the sequence as shown in table 1. Furthermore, the consensus sequence of

the multiple alignment has G-quadruplex forming potential, as it shows four stretches of at least three consecutive Gs. The existence of evolutionary conserved GQS based on a pairwise alignment of two sequences has been proposed as a method of validation and emphasis of their functional significance [35]. The presence of G-rich sequences in orthologues points to an evolutionary conservation of that feature, which supports the hypothesis that G-quadruplex formation, is a control mechanism of FXYD1 pre-mRNA processing.

Analysis of FXYD1 transcripts suggest alternative splicing

FXYD1 has 10 splice variants as listed in the Ensembl database. Among those only FXYD1-009, based on manual annotation by the HAVANA project [36], expresses a different longer peptide (115 aa), which lacks the membrane targeting signal. The comparison between H. sapiens' FXYD1 genomic sequence and its splice variants FXYD1-009 and FXYD1-001 suggests an alternative 3' splice site, which leads to a partially spliced intron four of the FXYD1-pre mRNA that produces FXYD1-009. The protein product of the splice variant 0009 lacks the signal peptide in its polypeptide chain, which means translocation of FXYD1 to the cell membrane is impaired, thus limiting the association between FXYD1 and the sodium pump.

 A comparison of the of the *FXYD1-01* and *FXYD1*-009 was performed using the '1000 Genomes Transcript comparison' option. *H. sapiens'* highest scoring GQS, GGGAGACUGCGGGUAUUCUGGGGGAGAGGG, was compared in the two pre-mRNA transcripts for mutations. In the second loop of the GQS an A ->G substition was found and a G -> A substitution in the fourth quartet, the latter of which affects G-quadruplex stablity. Computational analysis of the GQS with the A->G substition predicts a low stability G-quadruplex based on the QGRS mapper (G-score 14) and no G-quadruplex formation was predicted based on minimum free energy calculations.

Whilst the exact mechanism of G-quadruplex involvement in pre-mRNA splicing is not understood, it was reported that G-quadruplexes can regulate splicing of RNA [16, 17]. The analysis suggests a potential influence of the G-quadruplex from intron six of FXYD1 in the full splicing of intron four.

Native gel-electrophoresis indicates that human and bovine FXYD1-derived synthetic oligonucleotides form G-quadruplexes *in vitro*.

Native polyacrylamide gel electrophoresis (PAGE) has been used to detect the presence of intramolecular G-quadruplexes due to the faster migration of the more compact G-quadruplex structure compared to the extended oligonucleotide [37, 38].

Figure 2 shows examples of native PAGE of human and bovine GQS together with controls. The oligonucleotide sequences used are given in table 1. Comparing the gel without K^+ ions (Figure 2A) and the gel with 0.1 M K^+ it is apparent that human and bovine QGS (lanes 4 and 5) migrate faster and at the same time show a smear of intensity along the lane. The faster migration indicates G-quadruplex formation in the presence of K^+ of human and bovine FXYD1-derived GQS. Intramolecular G-quadruplexes are compact in shape and confer high

mobility rates in gels in comparison to linear species and intermolecular G-quadruplexes [37, 38]. The smear of intensity can be interpreted as the formation of secondary structures as indicated by the minimum free energy calculations above. The negative control_A sample (lanes 2 in Figure 2) showed in addition to the expected band a band at high molecular mass in the presence of potassium. This can be attributed to the formation of *inter*molecular G-quadruplexes, which is possible in negative control_A since the oligonucleotide has two stretches of four consecutive guanine bases. This could lead to intermolecular G-quadruplexes in either a dimeric or tetrameric arrangement.These observations are also supported by fluorescence emission spectroscopy comparing negative control_A with negative control_B as shown in the next section. The absolute R_f values depend on the molecular mass well as the conformation of the oligonucleotides. Negative control_A (DNA) has the same molecular mass as the positive control and shows the same R_f value as the positive control (DNA) in the absence of potassium. The absolute R_f values of negative control_B with *Homo sapiens* GQS cannot be directly compared as they are chemically different, i.e. DNA vs RNA. In order to obtain samples of the highest purity and stability, DNA was chosen for all control samples.

Native PAGE of short oligonucleotides is experimentally challenging and may lead to variable results. Therefore five different experiments were performed for each K^+ concentration of 0 M, 0.05 M and 0.1 M, and the relative migration distance R_f was calculated (Figure 3). The average of five experiments revealed a clear trend for an increase of the relative migration distance R_f with increasing potassium concentration for human and bovine GQS. The same trend was observed with the positive control, a sequence known to form G-quadruplexes, but not with the negative controls. The difference between R_f values for the positive control and human/bovine QGS sequences is statistically significant at $p < 0.05$ for both 0.05 M and 0.1 M K^+ buffer as shown in table 3.

Fluorescence spectroscopy confirms the formation of G-quadruplex in *FXYD1* pre-mRNA.

Fluorescence emission profiles of samples were determined in the presence and absence of K^+ and are shown in Figure 4. Comparison of positive control with human and bovine GQS shows an increase in fluorescence emission intensity in 0.1 M K^+ buffer compared to a buffer without potassium ions, while the negative control_B does not show any increase. G-quadruplex entities have been reported to have increased fluorescence emission intensity in contrast to non-G-quadruplex complexes due to the stacking of G-tetrads [46, 47]. Therefore, the data shown in Figure 4 are a strong indication of G-quadruplex formation of RNA oligonucleotides derived from FXYD1 pre-mRNA. Comparing the human and bovine GQS, it is notable that the human RNA shows higher fluorescence emission even in the absence of potassium ions. A possible explanation is formation of secondary structures, which would alter the fluorescence emission spectrum. This possibility was also indicated by energy calculations and gel electrophoresis results that have been discussed above. However, there is a further increase of the fluorescence intensity in K^+-containing buffer, which cannot be explained by further secondary structure formation.

Interestingly, the negative control_A sample showed an increase in fluorescence intensity similar to the positive control.The increase of fluorescence intensity in the negative control_A sample is due to intermolecular G-quadruplexes that was already indicated by gel electrophoresis experiments. At a final oligonucleotide concentration of 5 µM, as reported

by Moon and Jarsfer [48], intermolecular G-quadruplexes are highly favoured and migrate on gels even slower than linear non-G-quadruplex samples. Therefore, the fluorescence experiments are capable of detecting intra- and inter-molecular G-quadruplexes.

Conclusions

In conclusion, using a computational scan of the FXYD1 pre-mRNA potential G-quadruplex forming sequences (GQS) were identified in *Homo sapiens*, *Bos taurus* and other orthologues. Through energy calculations it was established that the G-quadruplex was either the most stable structure or existent in a significant proportion next to secondary structures. The stability of these G-quadruplex structures is likely higher *in vivo* considering the intracellular K^+ concentration of 120-150 mM. Using native PAGE and fluorescence emission spectroscopy the theoretical calculations were confirmed and the existence of G-quadruplex structures established. Multiple sequence alignment of orthologue GQS confirmed that the G-quadruplex forming potential is conserved in evolution, indicating a possibility that it may occur *in vivo* as a mechanism to control phospholemman expression levels and ultimately the activity of the cardiac sodium-potassium ATPase. Since FXYD1/phospholemman has been proposed as a potential therapeutic target for the treatment of cardiac hypertrophy and heart failure[22] modulating the G-quadruplex formation of FXYD1 pre-mRNA with small molecules or siRNA could have a future in the therapy of heart disease.

Further work should address the consequences of G-quadruplex structure on the splicing of pre-mRNA, which may be a reduced splicing activity or alternative splicing that will cause expression of the variant FXYD1-009. Additionally, the existence of G-quadruplex formation and the consequences of it should be confirmed *in vivo*.

Acknowledgements

<u>References</u>

[1]   M. Franceschin, Eur. J. Org. Chem. (2009) 2225-2238.

[2]   J.L. Huppert, Febs J 277 (2010) 3452-3458.

[3]   J.-D. Beaudoin, J.-P. Perreault, Nucleic Acids Res 38 (2010) 7022-7036.

[4]   J.E. Johnson, K. Cao, P. Ryvkin, L.-S. Wang, F.B. Johnson, Nucleic Acids Res 38 (2010) 1114-1122.

[5]   M.I. Onyshchenko, T.I. Gaynutdinov, E.A. Englund, D.H. Appella, R.D. Neumann, I.G. Panyutin, Nucleic Acids Res 37 (2009) 7570-7580.

[6]   M. Gellert, M.N. Lipsett, D.R. Davies, Proceedings of the National Academy of Sciences of the United States of America 48 (1962) 2013-2018.

[7]   D. Sen, W. Gilbert, Nature 334 (1988) 364-366.

[8]   J.L. Huppert, Philosophical Transactions of the Royal Society a-Mathematical Physical and Engineering Sciences 365 (2007) 2969-2984.

[9]   T.-M. Ou, Y.-J. Lu, C. Zhang, Z.-S. Huang, X.-D. Wang, J.-H. Tan, Y. Chen, D.-L. Ma, K.-Y. Wong, J.C.-O. Tang, A.S.-C. Chan, L.-Q. Gu, Journal of Medicinal Chemistry 50 (2007) 1465-1474.

[10]   A. Bugaut, S. Balasubramanian, Nucleic Acids Res 40 (2012) 4727-4741.

[11]   J.L. Huppert, A. Bugaut, S. Kumari, S. Balasubramanian, Nucleic Acids Res 36 (2008) 6260-6268.

[12]   S. Kumari, A. Bugaut, J.L. Huppert, S. Balasubramanian, Nature chemical biology 3 (2007) 218-221.

[13]   V. Marcel, P.L.T. Tran, C. Sagne, G. Martel-Planche, L. Vaslin, M.-P. Teulade-Fichou, J. Hall, J.-L. Mergny, P. Hainaut, E. Van Dyck, Carcinogenesis 32 (2011) 271-278.

[14]   D. Gomez, T. Lemarteleur, L. Lacroix, P. Mailliet, J.L. Mergny, J.F. Riou, Nucleic Acids Res 32 (2004) 371-379.

[15]   N.K. Gray, M.W. Hentze, Embo J 13 (1994) 3882-3891.

[16]   A.W. van der Velden, A.A.M. Thomas, International Journal of Biochemistry & Cell Biology 31 (1999) 87-106.

[17]   J.Y. Cheung, X.Q. Zhang, J. Song, E. Gao, J.E. Rabinowitz, T.O. Chan, J. Wang, Clinical and translational science 3 (2010) 189-196.

[18]   P. Teriete, C.M. Franzin, J. Choi, F.M. Marassi, Biochemistry-Us 46 (2007) 6774-6783.

[19]   S. Despa, J. Bossuyt, F. Han, K.S. Ginsburg, L.G. Jia, H. Kutchai, A.L. Tucker, D.M. Bers, Circ Res 97 (2005) 252-259.

[20]   W. Fuller, P. Eaton, J.R. Bell, M.J. Shattock, Faseb J 18 (2004) 197-199.

[21]   B. Silverman, W. Fuller, P. Eaton, J. Deng, J.R. Moorman, J.Y. Cheung, A.F. James, M.J. Shattock, Cardiovascular research 65 (2005) 93-103.

[22]   M.J. Shattock, Current opinion in pharmacology 9 (2009) 160-166.

[23]   A.J. Beevers, A. Kukol, Protein Science 15 (2006) 1127-1132.

[24]   J. Bossuyt, S. Despa, J.L. Martin, D.M. Bers, Journal of Biological Chemistry 281 (2006) 32765-32773.

[25]   Q. Song, S. Pallikkuth, J. Bossuyt, D.M. Bers, S.L. Robia, Journal of Biological Chemistry 286 (2011) 9120-9126.

[26]   A.J. Beevers, A. Kukol, Journal of Biological Chemistry 282 (2007) 32742-32748.

[27]   J.P. Morth, B.P. Pedersen, M.S. Toustrup-Jensen, T.L.M. Sorensen, J. Petersen, J.P. Andersen, B. Vilsen, P. Nissen, Nature 450 (2007) 1043-1049.

[28]   T. Shinoda, H. Ogawa, F. Cornelius, C. Toyoshima, Nature 459 (2009) 446-450.

[29]   O. Kikin, L. D'Antonio, P.S. Bagga, Nucleic Acids Res 34 (2006) W676-W682.

[30]   R. Lorenz, S.H. Bernhart, C.H.Z. Siederdissen, H. Tafer, C. Flamm, P.F. Stadler, I.L. Hofacker, Algorithms for Molecular Biology 6 (2011) PMC3319429.

[31]   F. Sievers, A. Wilm, D. Dineen, T.J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Soding, J.D. Thompson, D.G. Higgins, Mol Syst Biol 7 (2011) 539.

[32]   D. Altshuler, R.M. Durbin, G.R. Abecasis, D.R. Bentley, A. Chakravarti, A.G. Clark, F.S. Collins, F.M. De la Vega, P. Donnelly, M. Egholm, P. Flicek, S.B. Gabriel, R.A. Gibbs, B.M. Knoppers, E.S. Lander, H. Lehrach, E.R. Mardis, G.A. McVean, D. Nickerson, L. Peltonen, A.J. Schafer, S.T. Sherry, J. Wang, R.K. Wilson, R.A. Gibbs, D. Deiros, M. Metzker, D. Muzny, J. Reid, D. Wheeler, J. Wang, J.X. Li, M. Jian, G. Li, R.Q. Li, H.Q. Liang, G. Tian, B. Wang, J. Wang, W. Wang, H.M. Yang, X.Q. Zhang, H.S. Zheng, E.S. Lander, D.L. Altshuler, L. Ambrogio, T. Bloom, K. Cibulskis, T.J. Fennell, S.B. Gabriel, D.B. Jaffe, E. Shefler, C.L. Sougnez, D.R. Bentley, N. Gormley, S. Humphray, Z. Kingsbury, P. Koko-Gonzales, J. Stone, K.J. McKernan, G.L. Costa, J.K. Ichikawa, C.C. Lee, R. Sudbrak, H. Lehrach, T.A. Borodina, A. Dahl, A.N. Davydov, P. Marquardt, F. Mertes, W. Nietfeld, P. Rosenstiel, S. Schreiber, A.V. Soldatov, B. Timmermann, M. Tolzmann, M. Egholm, J. Affourtit, D. Ashworth, S. Attiya, M. Bachorski, E. Buglione, A. Burke, A. Caprio, C. Celone, S. Clark, D. Conners, B. Desany, L. Gu, L. Guccione, K. Kao, A. Kebbel, J. Knowlton, M. Labrecque, L. McDade, C. Mealmaker, M. Minderman, A. Nawrocki, F. Niazi, K. Pareja, R. Ramenani, D. Riches, W. Song, C. Turcotte, S. Wang, E.R. Mardis, D. Dooling, L. Fulton, R. Fulton, G. Weinstock, R.M. Durbin, J. Burton, D.M. Carter, C. Churcher, A. Coffey, A. Cox, A. Palotie, M. Quail, T. Skelly, J. Stalker, H.P. Swerdlow, D. Turner, A. De Witte, S. Giles, R.A. Gibbs, D. Wheeler, M. Bainbridge, D. Challis, A. Sabo, F. Yu, J. Yu, J. Wang, X.D. Fang, X.S. Guo, R.Q. Li, Y.R. Li, R.B. Luo, S. Tai, H.L. Wu, H.C. Zheng, X.L. Zheng, Y. Zhou, H.M. Yang, G.T. Marth, E.P. Garrison, W. Huang, A. Indap, D. Kural, W.P. Lee, W.F. Leong, W.C. Huang, A. Indap, D. Kural, W.P. Lee, W.F. Leong, A.R. Quinlan, C. Stewart, M.P. Stromberg, A.N. Ward, J.T. Wu, C. Lee, R.E. Mills, X.H. Shi, M.J. Daly, M.A. DePristo, D.L. Altshuler, A.D. Ball, E. Banks, T. Bloom, B.L. Browning, K. Cibulskis, T.J. Fennell, K.V. Garimella, S.R. Grossman, R.E. Handsaker, M. Hanna, C. Hartl, D.B. Jaffe, A.M. Kernytsky, J.M. Korn, H. Li, J.R. Maguire, S.A. McCarroll, A. McKenna, J.C. Nemesh, A.A. Philippakis, R.E. Poplin, A. Price, M.A. Rivas, P.C. Sabeti, S.F. Schaffner, E. Shefler, I.A. Shlyakhter, D.N. Cooper, E.V. Ball, M. Mort, A.D. Phillips, P.D. Stenson, J. Sebat, V. Makarov, K. Ye, S.C. Yoon, C.D. Bustamante, A.G. Clark, A. Boyko, J. Degenhardt, S. Gravel, R.N. Gutenkunst, M. Kaganovich, A. Keinan, P. Lacroute, X. Ma, A. Reynolds, L. Clarke, P. Flicek, F. Cunningham, J. Herrero, S. Keenen, E. Kulesha, R. Leinonen, W. McLaren, R. Radhakrishnan, R.E. Smith, V. Zalunin, X.Q. Zheng-Bradley, J.O. Korbel, A.M. Stutz, S. Humphray, M. Bauer, R.K. Cheetham, T. Cox, M. Eberle, T. James, S. Kahn, L. Murray, K. Ye, F.M. De La Vega, Y.T. Fu, F.C.L. Hyland, J.M. Manning, S.F. McLaughlin, H.E. Peckham, O. Sakarya, Y.A. Sun, E.F. Tsung, M.A. Batzer, M.K. Konkel, J.A. Walker, R. Sudbrak, M.W. Albrecht, V.S. Amstislavskiy, R. Herwig, D.V. Parkhomchuk, S.T. Sherry, R. Agarwala, H. Khouri, A.O. Morgulis, J.E. Paschall, L.D. Phan, K.E. Rotmistrovsky, R.D. Sanders, M.F. Shumway, C.L. Xiao, G.A. McVean, A. Auton, Z. Iqbal, G. Lunter, J.L. Marchini, L. Moutsianas, S. Myers, A. Tumian, B. Desany, J. Knight, R. Winer, D.W. Craig, S.M. Beckstrom-Sternberg, A. Christoforides, A.A. Kurdoglu, J. Pearson, S.A. Sinari, W.D. Tembe, D. Haussler, A.S. Hinrichs, S.J. Katzman, A. Kern, R.M. Kuhn, M. Przeworski, R.D. Hernandez, B. Howie, J.L. Kelley, S.C. Melton, G.R. Abecasis, Y. Li, P. Anderson, T. Blackwell, W. Chen, W.O. Cookson, J. Ding, H.M. Kang, M. Lathrop, L.M. Liang, M.F. Moffatt, P. Scheet, C. Sidore, M.

Snyder, X.W. Zhan, S. Zollner, P. Awadalla, F. Casals, Y. Idaghdour, J. Keebler, E.A. Stone, M. Zilversmit, L. Jorde, J.C. Xing, E.E. Eichler, G. Aksay, C. Alkan, I. Hajirasouliha, F. Hormozdiari, J.M. Kidd, S.C. Sahinalp, P.H. Sudmant, E.R. Mardis, K. Chen, A. Chinwalla, L. Ding, D.C. Koboldt, M.D. McLellan, D. Dooling, G. Weinstock, J.W. Wallis, M.C. Wendl, Q.Y. Zhang, R.M. Durbin, C.A. Albers, Q. Ayub, S. Balasubramaniam, J.C. Barrett, D.M. Carter, Y.A. Chen, D.F. Conrad, P. Danecek, E.T. Dermitzakis, M. Hu, N. Huang, M.E. Hurles, H.J. Jin, L. Jostins, T.M. Keane, T.M. Keane, S.Q. Le, S. Lindsay, Q.A. Long, D.G. MacArthur, S.B. Montgomery, L. Parts, J. Stalker, C. Tyler-Smith, K. Walter, Y.J. Zhang, M.B. Gerstein, M. Snyder, A. Abyzov, A. Abyzov, S. Balasubramanian, R. Bjornson, J.A. Du, F. Grubert, L. Habegger, R. Haraksingh, J. Jee, E. Khurana, H.Y.K. Lam, J. Leng, X.J. Mu, A.E. Urban, Z.D. Zhang, Y.R. Li, R.B. Luo, G.T. Marth, E.P. Garrison, D. Kural, A.R. Quinlan, C. Stewart, M.P. Stromberg, A.N. Ward, J.T. Wu, C. Lee, R.E. Mills, X.H. Shi, S.A. McCarroll, E. Banks, M.A. DePristo, R.E. Handsaker, C. Hartl, J.M. Korn, H. Li, J.C. Nemesh, J. Sebat, V. Makarov, K. Ye, S.C. Yoon, J. Degenhardt, M. Kaganovich, L. Clarke, R.E. Smith, X.Q. Zheng-Bradley, J.O. Korbel, S. Humphray, R.K. Cheetham, M. Eberle, S. Kahn, L. Murray, K. Ye, F.M. De la Vega, Y.T. Fu, H.E. Peckham, Y.A. Sun, M.A. Batzer, M.K. Konkel, C.L. Xiao, Z. Iqbal, B. Desany, T. Blackwell, M. Snyder, J.C. Xing, E.E. Eichler, G. Aksay, C. Alkan, I. Hajirasouliha, F. Hormozdiari, J.M. Kidd, K. Chen, A. Chinwalla, L. Ding, M.D. McLellan, J.W. Wallis, M.E. Hurles, D.F. Conrad, K. Walter, Y.J. Zhang, M.B. Gerstein, M. Snyder, A. Abyzov, J.A. Du, F. Grubert, R. Haraksingh, J. Jee, E. Khurana, H.Y.K. Lam, J. Leng, X.J. Mu, A.E. Urban, Z.D. Zhang, R.A. Gibbs, M. Bainbridge, D. Challis, C. Coafra, H. Dinh, C. Kovar, S. Lee, D. Muzny, L. Nazareth, J. Reid, A. Sabo, F.L. Yu, J. Yu, G.T. Marth, E.P. Garrison, A. Indap, W.F. Leong, A.R. Quinlan, C. Stewart, A.N. Ward, J.T. Wu, K. Cibulskis, T.J. Fennell, S.B. Gabriel, K.V. Garimella, C. Hartl, E. Shefler, C.L. Sougnez, J. Wilkinson, A.G. Clark, S. Gravel, F. Grubert, L. Clarke, P. Flicek, R.E. Smith, X.Q. Zheng-Bradley, S.T. Sherry, H.M. Khouri, J.E. Paschall, M.F. Shumway, C.L. Xiao, G.A. McVean, S.J. Katzman, G.R. Abecasis, T. Blackwell, E.R. Mardis, D. Dooling, L. Fulton, R. Fulton, D.C. Koboldt, R.M. Durbin, S. Balasubramaniam, A. Coffey, T.M. Keane, D.G. MacArthur, A. Palotie, C. Scott, J. Stalker, C. Tyler-Smith, M.B. Gerstein, S. Balasubramanian, A. Chakravarti, B.M. Knoppers, L. Peltonen, G.R. Abecasis, C.D. Bustamante, N. Gharani, R.A. Gibbs, L. Jorde, J.S. Kaye, A. Kent, T. Li, A.L. McGuire, G.A. McVean, P.N. Ossorio, C.N. Rotimi, Y.Y. Su, L.H. Toji, C. Tyler-Smith, L.D. Brooks, A.L. Felsenfeld, J.E. McEwen, A. Abdallah, R. Christopher, N.C. Clemm, F.S. Collins, A. Duncanson, E.D. Green, M.S. Guyer, J.L. Peterson, A.J. Schafer, G.R. Abecasis, D.L. Altshuler, A. Auton, L.D. Brooks, R.M. Durbin, R.A. Gibbs, M.E. Hurles, G.A. McVean, G.P. Consortium, Nature 467 (2010) 1061-1073.

[33]   K.W. Lim, P. Alberti, A. Guedin, L. Lacroix, J.-F. Riou, N.J. Royle, J.-L. Mergny, A.T. Phan, Nucleic Acids Res 37 (2009) 6239-6248.

[34]   R. Lorenz, S.H. Bernhart, J. Qin, C. Honer Zu Siederdissen, A. Tanzer, F. Amman, I.L. Hofacker, P.F. Stadler, IEEE/ACM transactions on computational biology and bioinformatics / IEEE, ACM (2013) 832-844.

[35]   C. Menendez, S. Frees, P.S. Bagga, Nucleic Acids Res 40 (2012) W96-W103.

[36]   P. Flicek, I. Ahmed, M.R. Amode, D. Barrell, K. Beal, S. Brent, D. Carvalho-Silva, P. Clapham, G. Coates, S. Fairley, S. Fitzgerald, L. Gil, C. Garcia-Giron, L. Gordon, T. Hourlier, S. Hunt, T. Juettemann, A.K. Kahari, S. Keenan, M. Komorowska, E. Kulesha, I. Longden, T. Maurel, W.M. McLaren, M. Muffato, R. Nag, B. Overduin, M. Pignatelli, B. Pritchard, E. Pritchard, H.S. Riat, G.R.S. Ritchie, M. Ruffier, M. Schuster, D. Sheppard, D. Sobral, K. Taylor, A. Thormann, S. Trevanion, S. White, S.P. Wilder, B.L. Aken, E. Birney, F. Cunningham, I. Dunham, J. Harrow, J. Herrero, T.J.P. Hubbard, N. Johnson, R. Kinsella, A. Parker, G. Spudich, A. Yates, A. Zadissa, S.M.J. Searle, Nucleic Acids Res 41 (2013) D48-D55.

[37]   T.M. Bryan, P. Baumann, Mol. Biotechnol. 49 (2011) 198-208.

[38]   J.R. Williamson, M.K. Raghuraman, T.R. Cech, Cell 59 (1989) 871-880.

Tables

Table 1: The highest scoring predicted G-Quadruplex forming Sequences (GQS) from FXYD1 orthologues and their location within the gene. The analysis was performed with QGRS mapper [29].

Organism	Sequence of highest scoring GQS	G-score	Genomic location[1]
Homo sapiens	GGGAGACUGCGGGUAUUCUGGGGAGAGGG	39	Intronic (6:7)
Musmusculus	GGGAGGAAGGAGGGAGAGGGUUUGGAGGG	38	Intronic (7:8)
Canis lupus	GGGGGCGAAGGGUGGGCUGGGAUGGCCGGG	42	3'-UTR
Pan troglodytes	GGGAGACUGCGGGUAUUUGGGGAGAGGG	39	Intronic (6:7)
	GGGUUGAAGGGCGGCGAGGGGUGGGG	39	Intronic (7:8)
Bostaurus	GGGCGCGGGGGGUCGGGGAUCGGG	42	Intronic (6:7)
Rattus norvegicus	GGGCAGGUGAGGCUGGG	21	Intronic (1:2)
	GGAUGGAAGGUAGG	21	Intronic (2:3)
	GGCGGUGGGGG	21	Intronic (5:6)
	GGCACGGGGAGGUAAGG	21	Intronic (5:6)
	GGGAGGAAGGAGGG	21	Intronic (7:8)
	GGCGGGUUGGAGGG	21	Intronic (7:8)
Feliscatus	GGGAGACUUUGGGGGUUUGGGGGUGAGGG	40	Intronic (5:6)
Otolemurgarnettii	GGGCGCAGGGUGGGGUGGGUGAGGCGGG	40	Intronic (4:5)
Tursiopstruncatus	GGGAGUUAGGGGGUGCUGGGCUGGG	38	Intronic (2:3)

Species	Sequence	Value	Genomic location[1]
Equuscaballus	GGGAGUUGGGAGUGGGGUUUGGG	42	Intronic (3:4)
Ailuropodamelanoleuca	GGGAGACUUCGGGUGUUUGGGGGUGAGGG	40	Intronic (5:6)
Pongoabelii	GGGAGACUGCGGGUAUUUUGGGGAGAGGG	39	Intronic (5:6)
	GGGUUGAAGGGCGGCGAGGGGUGGGG	39	Intronic (6:7)
Oryctolaguscuniculus	GGGAGAGUGGGUGGGGGUCCUGGG	40	Intronic (5:6)
Gorilla gorillagorilla	GGUGGCGGUGG	21	Intronic (1:2)
Sus scrofa	GGGGGUGGGGGUGGGGGUGGGGG	83	Intronic (2:3)
Ovisaries	GGGCUG*GGG*CAAAGGGGGAGGG	41	Intronic (1:2)
Monodelphisdomestica	GGGGGUGGGGAGGAGGGAUGGG	40	5'-UTR
	GGGAGAUGGGGGGGGGUAGGUGGG	40	Intronic (2:3)
positive control[2]	AGGGCTAGGGCTAGGGCTAGGG	42	N/A
negative control_A[3]	CGTGGGGAGATTGGGGAGCGCA	0	N/A
negative control_B	GGTGTGCGTGTGCGAGCGAGAGAGAGTGG	0	N/A

[1] The genomic location specifies the intron between the numbered exons

[2] The G-quadruplex structure of this DNA sequence was determined by nuclear magnetic resonance (Protein databank-ID: 2KM3)

[3] All controls were DNA. The negative control_A is a randomised sequence with the same base composition as the positive control. The negative control_B has the same base composition as the *Homo Sapiens* GQS.

Table 2: Most stable RNA structures predicted by RNAfold from the Vienna RNA package and minimum free energy (MFE). In case the G-quadruplex was not the most stable structure the MFE of suboptimal structures was calculated with RNAeval.

Sample Name	Dot bracket notation[1]	MFE in kcal/mol
positive control	`AGGGCTAGGGCTAGGGCTAGGG` `.+++...+++...+++...+++`	-12.65
negative control_A	`CGTGGGGAGATTGGGGAGCGCA` `.(((..............))).`	-0.30
negative control_B	`GGUGUGCGUGUGCGAGCGAGAGAGAGUGG` `.((.((((....))).)).............`	-2.70

Homo sapiens QGS	`GGGAGACUGCGGGUAUUCUGGGGAGAGGG`
	`......((.(((.....))).))......`
	`.......((.((((...))))).)).....`
	`+++.......+++......+++....+++`
		-4.30
		-4.10
		-3.51

Bostaurus QGS	`GGGCGCGGGGGGUCGGGGAUCGGG` `+++...+++.....+++....+++`	-8.37

[1] The symbols '(' and ')' represent canonical base pairs, '+' represents guanine bases taking part in G-tetrad formation, '.' represents unpaired bases

Table 3: P-values obtained from a Student two-tailed-t-test of R_f values comparing samples with K^+-buffer with samples in K^+ free buffer.

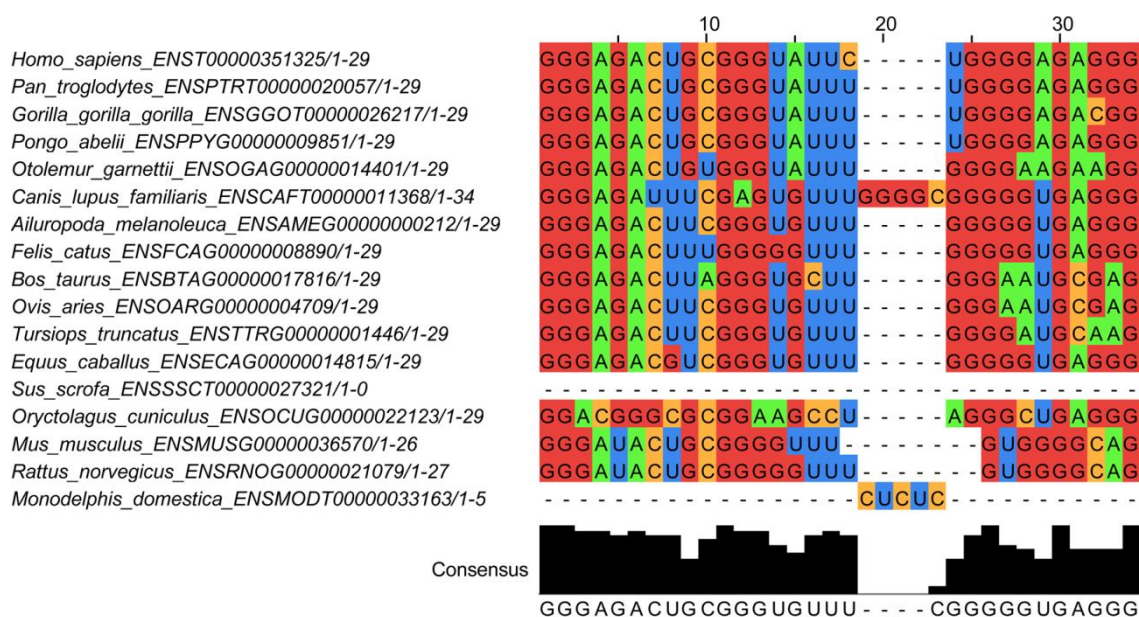| $[K^+]$ in M	Positive control	Negative control_A	Negative control_B	Human GQS	Bovine GQS
0.05	$4.2 \cdot 10^{-5}$	0.74	0.46	$3.3 \cdot 10^{-6}$	$2.0 \cdot 10^{-7}$
0.10	$4.2 \cdot 10^{-7}$	0.72	0.34	$3.1 \cdot 10^{-7}$	$5.3 \cdot 10^{-11}$

Figures

Figure 1: Multiple sequence alignmentof pre-mRNAs from *H. sapiens* and orthologues. The *H. sapiens*GQS sequence from Table 1 was highlighted.Note that the sequences shown are based on the best alignment to the human GQS and the highest scoring GQS of that particular organism. The dashes stand for gaps inserted by the multiple alignment algorithm to achieve the best match between the sequences.
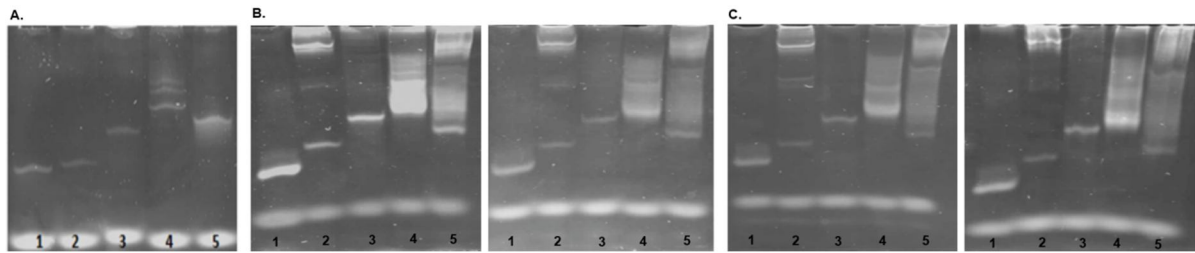
Figure 2: Native 30% PAGE of GQS oligos (table 1) in absence and presence of K$^+$ions. Lanes **1:** positive control**; 2:**negativecontrol_A**; 3:**negative control_B**; 4:**human GQS**; 5:**bovine GQS. **A.** Samples incubated in 0.02 M TrisOAc buffer solution without K$^+$ ions. The positive and negative control_A controls migrated at the same rate on the gel. Negative control_B migrated less than other controls but faster than human and bovine GQS. Human GQS was the slowest migrating sample. **B.**Samples incubated in buffer with 0.05 M K$^+$concentration.The positive control was the fastest migrating sample. The bovine QGS now migrated faster than negative control_B. In the order of fastest to slowest migrating sample: positive control > negative control_A> bovine GQS > negative control_B> human GQS. **C.**Samples incubated in buffer with 0.10 M K$^+$concentration.Samples migrated with a similar trend as seen in B.
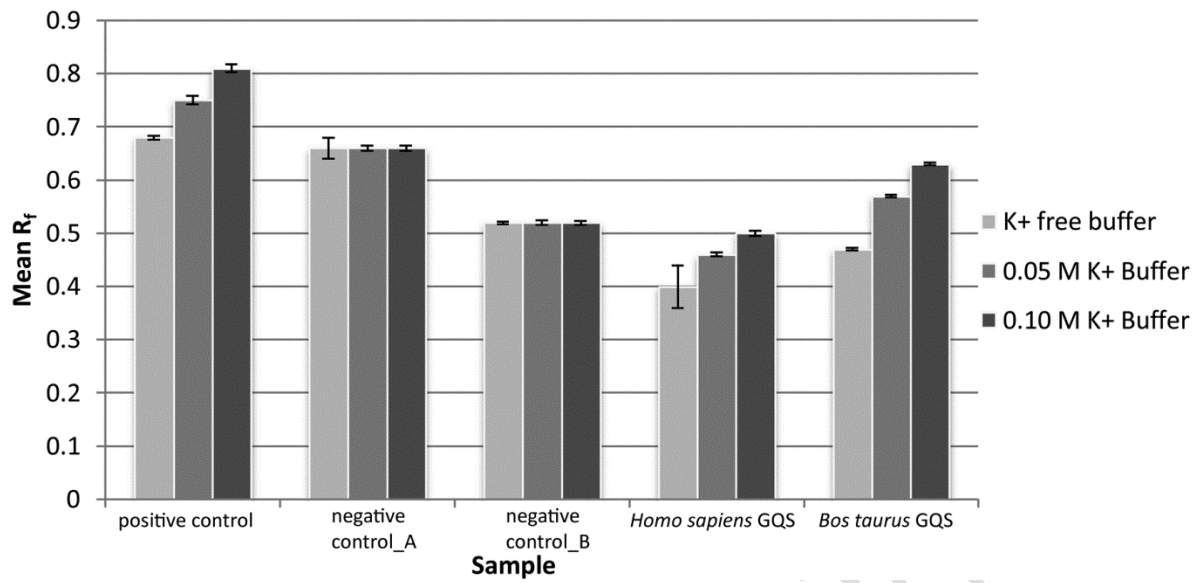
Figure 3: Average relative migration distances R_f from n=5 experiments obtained from native PAGE experiments for samples treated in K^+ free and K^+ containing buffer.
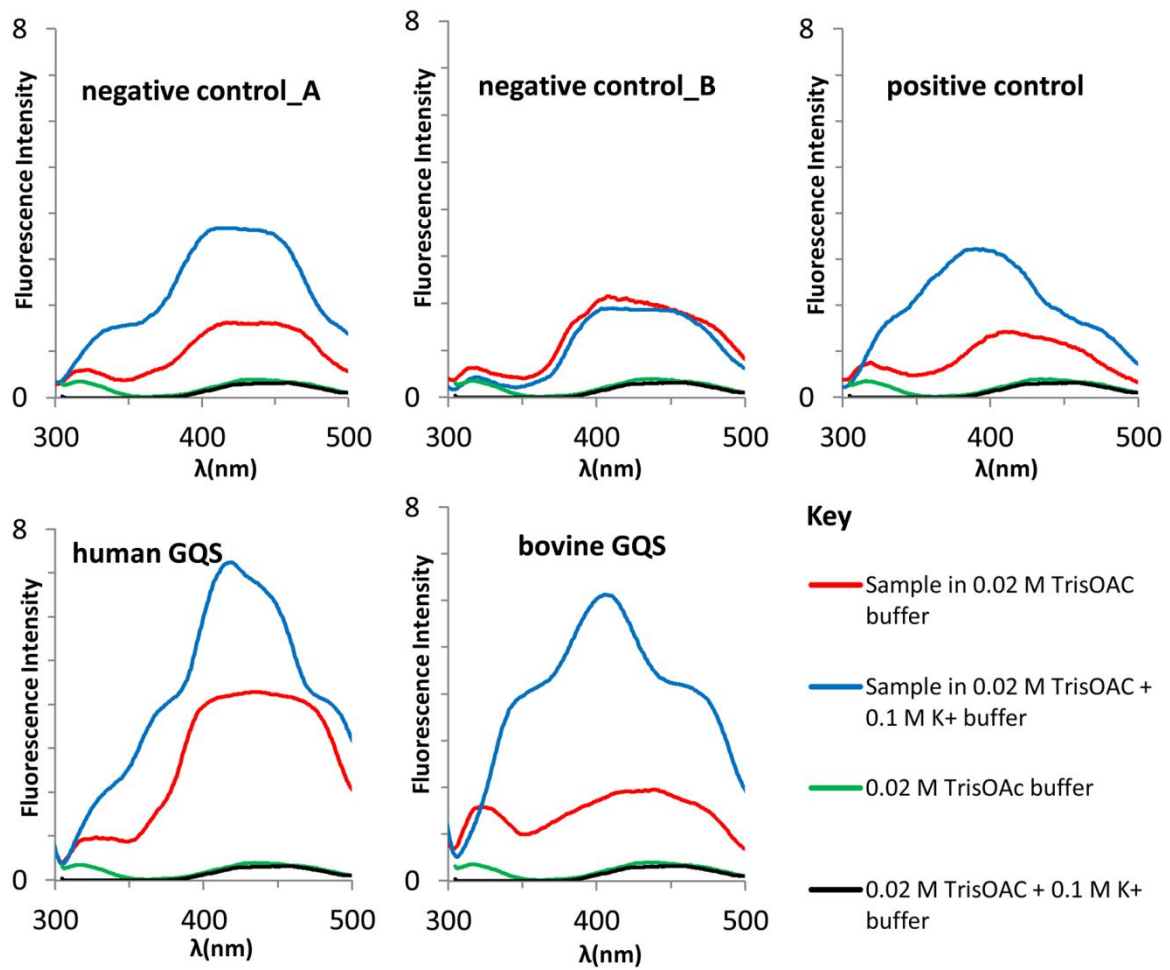
Figure 4: Fluorescence emission spectra of samples excited at 260 nm in the presence of K$^+$ and absence of K$^+$ (blue and red curves). In addition the spectra of buffers were recorded (green and black curves). The oligonucleotide concentration was 5 µM for the DNA controls and 1.5µM for RNA samples.