Small effect of the androgen receptor gene GGN repeat polymorphism on serum testosterone levels in healthy men

Veerle Bogaert1, Griet Vanbillemont 1, Youri Taes 1, Dirk De Bacquer 2,3, Ellen Deschepper 2,3, Kristel Van Steen 4,5,6 and Jean-Marc Kaufman 1

1Department of Endocrinology, 9 K12 I.E, Ghent University Hospital, De Pintelaan 185, 9000 Gent, Belgium, 2Department of Public Health and 3Biostatistical Unit, Ghent University, Ghent, Belgium, 4Montefiore Institute – Bioinformatics, University of Liège, Liège, Belgium, 5StepGen cvba, Merelbeke, Belgium and 6Department of Medical Genetics, Ghent University, Ghent, Belgium

(Correspondence should be addressed to J-M Kaufman; Email: jean.kaufman@ugent.be)

Abstract

Objective: The human androgen receptor (AR) contains a polyglutamine and a polyglycine stretch which are highly polymorphic and are coded respectively by a CAG and GGN repeat in exon 1 of the AR gene. Although the in vitro studies indicated a possible effect of the GGN repeat polymorphism on the AR gene transcription and clinical observations suggest that it might modulate the androgen action, its functional significance remains unclear. We wanted to assess whether the GGN repeat affects the serum testosterone levels in healthy men, which is the expected outcome through feedback regulation if it influences androgen action as has been shown to be the case for the CAG repeat.

Design and patients: A population based cross-sectional cohort study including 1476 healthy young, middle-aged, and elderly men.

Measurement: Testosterone and LH levels were determined by immunoassay; free testosterone (FT) levels were calculated. Genotyping of the GGN repeat was performed using the sequencing technique.

Results: The GGN repeat number was significantly associated with circulating testosterone and FT levels (P<0.017 and P<0.013 respectively). However, taking into account that age, body mass index, and CAG are already in the regression model, the GGN repeat could explain only a small part of the variation of both testosterone and FT.

Conclusion: To our knowledge, this study is the first to demonstrate a significant positive association between the GGN repeat and androgen levels in a large cohort of healthy men. Although the present study thus adds credence to the view that the polyglycine tract in the AR can modulate AR action, this effect appears to be only small so that its clinical relevance remains questionable.

European Journal of Endocrinology 161 171–177

Introduction

Androgens exert their biological activities by binding and activating the androgen receptor (AR). The AR is a member of the nuclear hormone-receptor superfamily of transcription factors encoded by a gene localized on the long arm (Xq11–12) of the X-chromosome. Within the N-terminal transactivation domain of the AR (exon 1), a well described CAG repeat polymorphism is present which encodes a polyglutamine tract and affects androgen action. Its length is inversely correlated with AR mRNA and protein levels (1, 2) and affects transcriptional efficacy of the AR through conformational changes and association with coactivators (1, 3). Through modulation of androgen action, the CAG repeat can affect the hypothalamic–pituitary feedback regulation with longer CAG repeats expected to be associated with the relatively diminished androgen feedback. This results in a positive association between the CAG repeat and circulating (free) testosterone (4) in healthy men. It has been reported that the CAG repeat modulates, through its influence on AR activity, several clinical androgen-related effects such as prostate cancer risk, spermatogenesis, bone density, hair growth, and cardiovascular risk factors (5).

Downstream of the CAG repeat polymorphism in exon 1 of the AR gene, another repeat polymorphism is present: the GGN repeat that encodes a polymorphic polyglycine (6). While the CAG is a simple repeat and varies in length from 8 to 35, the GGN repeat is complex, represented by (GGT)nGGG(GGT)2(GGC)n with n varying from 10 to 35. The functional effect of the GGN polymorphism remains unclear. Considerable discrepancies exist between the findings of the in vitro studies analyzing the effect of different GGN repeat lengths on AR activity in various cell cultures (7–11). Nevertheless, several epidemiological studies have suggested a significant, albeit modest, impact of the repeat on clinical androgenic effects: the GGN repeat has been linked to the risk of prostate cancer (12),
to alopecia (13), spermatogenesis (14), occurrence of hypospadias (15), and breast cancer (16).

On the other hand, one might postulate that if the polymorphic polyglycine tract affects the AR activity and thus androgen sensitivity, the GGN repeat would be expected to affect the hypothalamic–pituitary–gonadal axis and thus contribute to the between-subject variation of serum androgens, similarly to what has been observed for the CAG repeat (4). Up till now, the data on the relationship between the number of GGN repeats and androgen levels has been limited and inconclusive. One study in men with prostate cancer reported a positive relationship between the number of GGN repeats and androgen levels (17). In this study, short GGN repeats were associated with lower free testosterone (FT) serum levels, whereas no association was demonstrated in another study involving healthy young Swedish men (14).

The aim of our study was to investigate the role of the AR GGN repeat, alone or in combination with the CAG repeat polymorphism, in the regulation of serum FT levels in healthy men. This study provides new information on the association between GGN repeat of the AR and steroid hormone levels.

**Materials and methods**

**Study population**

This study included 1476 healthy subjects of Caucasian origin, contributed by three cohorts of young, middle-aged, and elderly men that were recruited in the same geographical area. The exclusion criteria were defined as illnesses or medication use that may affect sex steroid levels, body composition, or bone metabolism. The study protocols were approved by the ethical committee of the Ghent University Hospital and written informed consent was obtained from all participants.

Young men (24–45 years) were included from the Siblos study (4, 18), a population-based observational study designed to investigate genetic contribution to the determination of sex steroid levels and bone mass using a sib-pair design. For this study, we included a group of 358 unrelated men, i.e. a single representative of the nuclear families as previously described (4) to avoid interdependence of the data due to the family structure.

Middle-aged men (35–49 years) were recruited from the occupational Belstress study, an epidemiological cohort study focused on job stress, cardiovascular risk, and other health issues. The objectives, design, and methodology of this project have been described elsewhere (19). Our study was performed in a sub-sample of the Belstress study. In the Belstress population, serum testosterone was measured in 2322 men working in the geographical area of Ghent. Included in our study is a group of 857 men representative of the full range of serum FT concentrations in the overall study population as previously described (4). Healthy community-dwelling elderly men (71–86 years; n = 267) were recruited from the population registry of the semi-rural community Merelbeke (Belgium). Other characteristics were extensively described previously (20–22).

**Biochemical determinations**

Venous blood was obtained after overnight fasting and before 1000 h. All samples were stored at −80 °C until batch analysis. Commercial immunoassay kits were used to determine the serum concentrations of testosterone (Orion Diagnostica, Espoo, Finland) and LH (ECLIA, Roche Diagnostics). The free fraction of testosterone was calculated from serum testosterone, SHBG, and albumin concentrations using a previously validated equation (23).

**Genotyping**

Genomic DNA was extracted from EDTA-treated blood using a commercial kit (Puregene kit; Gentra Systems, Minneapolis, MN, USA). The AR exon 1 region encoding the CAG repeat was amplified using PCR with forward primer 5'-GAATCTGGTCCCAGGCTGC-3', fluorescently labeled with NED or FAM and reverse primer 5'-TTCTCTCAGAGACGAGTA-3'. Each PCR was initiated with a 5-min denaturation step at 95 °C and terminated with a 20-min extension step at 72 °C; in-between reaction profiles were as follows: denaturation at 95 °C for 60 s, annealing at 62 °C for 60 s, and extension at 72 °C for 90 s for 35 cycles. The PCR products were mixed with a Genescan 400 HD ROX size standard and deionized formamide, heat denatured, and electrophoresed on an ABI Prism 310 or a 96-capillary 3730 xl genetic analyzer (ABI Prism; Perkin–Elmer Applied Biosystems, Foster City, CA, USA).

The GGN repeat was amplified using PCR with forward primer 5'-CCGCTTCATCCTGGCCAC-3' and reverse primer 5'-GCCGCCAGGTACCACCACAT-3'. Each PCR was performed using a commercial enhancer kit (Invitrogen) and was initiated with a 12-min denaturation step at 96 °C and terminated with a 20-min extension step at 72 °C. Each reaction was as follows: denaturation at 95 °C for 60 s, annealing at 54 °C for 60 s, and extension at 72 °C for 180 s for 40 cycles.

The sequencing of PCR products was carried out using 1 μl of PCR product and 1 μM of non-fluorescent primer (2 μl; 10% DMSO was added to the forward primer) and 0.5 μl DMSO, 2.5 μl of ABI Prism BigDye Terminator Ready Reaction Kit (2001, Applied Biosystems) and 4 μl of double-distilled water to adjust the volume to 10 μl.

The sequencing products were electrophoresed on a 96-capillary 3730 xl Genetic analyzer (ABI Prism, Perkin–Elmer Applied Biosystems).
**Statistical analysis**

The Spearman’s rho was calculated to investigate the association between CAG and GGN. We further studied the association between hormone parameters and GGN repeat number using linear regression models. Adherence to normal distribution of outcomes of interest was formally tested using Kolmogorov–Smirnov, prior to model building. Free testosterone and LH were LN transformed to enhance normality. By incorporating an interaction term between CAG and GGN, we were able to investigate whether the CAG repeat acted as an effect modifier to the GGN repeat. Age and body mass index (BMI), known to be the major determinants of serum testosterone in healthy men (24), were included as confounding factors in all the regression models. Since mutations in the GGN repeat may alter the relation between the number of GGN repeats and androgen levels, individuals with such a mutation were excluded from this statistical analysis. All analyses were performed using SPSS 15.0 software (SPSS, Chicago, IL, USA).

**Results**

**Population characteristics**

Patients’ characteristics are summarized in Table 1. All hormone values were within the normal reference range. The allele size varied from 9 to 41 for the CAG repeat. The GGN locus is less variable, with the GGN repeat number ranging from 10 to 27. The median number of CAG was 21 and the most common alleles for GGN were 23 (52.4%) and 24 (31.8%). The CAG and GGN length distribution is shown in Fig. 1. Among 1476 study subjects, 39 men showed a deviation from the consensus GGN sequence (Table 2). Their genotypes were as follows: (GGT)4GGG(GGT)2(GGC)16; (GGT)2GGG(GGT)4(GGC)16; (GGT)2GGG(GGT)2(GGC)18 and (GGT)2GGG(GGT)2(GGC)18.

One mutation was found at the 3’ end of the GGN repeat: a C to T conversion resulted in the following sequence (GGT)2GGG(GGT)2(GGC)GAGGTG.

As shown in Table 2, the extra internal GGT does not always coincide with \( n = 20 \) and only 7 men had a repeat number of 20 without the extra GGT.

![Figure 1](https://example.com/fig1.png)

**Figure 1** The CAG and GGN allele distribution in 1476 healthy men. CAG repeat lengths show a near Gaussian distribution ranging from 9 to 41 repeats. For the GGN repeat, 23 and 24 are the two predominant alleles enclosing near to 84% of the population.

We found a weak, statistically significant inverse correlation between CAG and GGN \((r = -0.10, P < 0.001)\).

**GGN and serum FT**

Our data suggested a positive linear correlation between the number of GGN repeats and hormone levels with a bend when the number of GGN repeats were greater than 24 (results not shown). However, no statistically significant evidence was found to describe the GGN-hormone relationship by a quadratic trend. The combined effect of CAG and GGN was associated neither with testosterone and ln(FT) nor with ln(LH). The GGN genotype was a positive predictor of testosterone and ln(FT) levels \((\beta = 0.058, P = 0.029 \text{ and } \beta = 0.057, P = 0.032 \text{ respectively})\) and was found to be an independent positive predictor when included together with age, BMI, and CAG in the regression models (Table 3). However, the effect size was only modest: taken into account that CAG, age, and BMI are already in

<table>
<thead>
<tr>
<th></th>
<th>Total population ((n = 1476))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (year)</td>
<td>45 38–53</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>25.7 23.6–27.9</td>
</tr>
<tr>
<td>Testosterone (ng/dl)</td>
<td>495.7 382.0–632.0</td>
</tr>
<tr>
<td>FT (ng/dl)</td>
<td>11.2 8.3–14.0</td>
</tr>
<tr>
<td>LH (U/l)</td>
<td>4.6 3.4–6.3</td>
</tr>
<tr>
<td>CAG</td>
<td>21 20–24</td>
</tr>
<tr>
<td>GGN</td>
<td>17 17–18</td>
</tr>
</tbody>
</table>
the statistical model, the GGN repeat could explain only 0.4 percent of the variation of both testosterone and FT. An age-stratified analysis of young (24–39 years) and older (71–86 years) subjects did not reveal an age-specific effect (data not shown). However, this analysis did not allow for a definite exclusion of an age-specific effect due to insufficient power.

No significant association between ln(LH) and GGN could be found.

In an exploratory analysis for a subpopulation of 1076 men for which estradiol (E$_2$) was available, we did not observe an association between the GGN repeat number and (free)E$_2$ levels. Also, inclusion of E$_2$ in our linear regression models had no major effect on the association between the GGN repeat number and FT.

**Mutations in GGN repeat and androgen levels**

Alterations in the GGN genotype (different from the repeat number itself) had no significant influence on the androgen levels. However, some trend has been observed: men with GGN≤25 appeared to have higher FT levels when compared to men with GGN≤25 and an extra GGT. This in contrast with men having a GGN≥26 and a CAG≥21, where lower FT concentrations were found compared to men who have the same characteristics for the two repeats, but have an extra GGT (data not shown).

**Discussion**

This is the first large study with more than 1000 subjects relating GGN-polymorphism to circulating androgen levels. Previously, clinical and epidemiological studies have reported an association of the GGN repeat length in exon 1 of the AR with androgen-related clinical findings. If this is indeed the case, the GGN repeat would be expected to affect the androgen action and thus feedback regulation of the gonadal axis and serum FT levels in healthy men. Such an effect has been demonstrated for the CAG repeat polymorphism (4). The present study shows that the same holds true for the GGN repeat. However, the impact of this GGN repeat on serum FT is rather modest, explaining only a small part of the variation in serum FT in healthy men, independently of age, BMI, and CAG repeat polymorphism.

An explanation of this small effect may reside in the fact that the hypothalamus–pituitary feedback is not only controlled by testosterone levels but also by E$_2$ levels. Therefore, E$_2$ could theoretically blunt the effect of different GGN repeat lengths (resulting in an altered feedback) onto circulating androgen levels. However, this would be rather difficult to examine since the relationship between testosterone–E$_2$ is bi-directional. Not only has E$_2$ an effect on testosterone levels due to its involvement in the feedback regulation, testosterone itself influences E$_2$ levels via the aromatase pathway.

### Table 2 Hormone levels in healthy men with a deviation from the consensus GGN genotype.

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Total no. of individuals</th>
<th>No. of GGC</th>
<th>Frequency</th>
<th>Testosterone (ng/dl)</th>
<th>FT (ng/dl)</th>
<th>LH (U/l)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(GGT)$_2$GGG(GGT)$_2$(GGC)$_2$GAGGG 1435</td>
<td>/</td>
<td>/</td>
<td></td>
<td>506.9±173.3</td>
<td>11.7±4.3</td>
<td>5.5±4.2</td>
</tr>
<tr>
<td>(GGT)$_2$GGG(GGT)$_2$(GGC)$_2$GAGG   31</td>
<td>17</td>
<td>3</td>
<td>337.3±108.4</td>
<td>7.6±2.1</td>
<td>2.2±0.9</td>
<td></td>
</tr>
<tr>
<td>(GGT)$_2$GGG(GGT)$_2$(GGC)$_2$GAGG   18</td>
<td>18</td>
<td>1</td>
<td>434.0</td>
<td>10.0</td>
<td>3.9</td>
<td></td>
</tr>
<tr>
<td>(GGT)$_2$GGG(GGT)$_2$(GGC)$_2$GAGG   19</td>
<td>19</td>
<td>1</td>
<td>215.0</td>
<td>5.7</td>
<td>2.6</td>
<td></td>
</tr>
<tr>
<td>(GGT)$_2$GGG(GGT)$_2$(GGC)$_2$GAGG   20</td>
<td>20</td>
<td>24</td>
<td>519.5±171.0</td>
<td>12.7±4.5</td>
<td>5.3±1.7</td>
<td></td>
</tr>
<tr>
<td>(GGT)$_2$GGG(GGT)$_2$(GGC)$_2$GAGG   21</td>
<td>21</td>
<td>3</td>
<td>727.9±379.4</td>
<td>17.1±6.4</td>
<td>5.3±2.0</td>
<td></td>
</tr>
<tr>
<td>(GGT)$_2$GGG(GGT)$_2$(GGC)$_2$GAGG   22</td>
<td>22</td>
<td>3</td>
<td>514.0</td>
<td>10.8</td>
<td>2.1</td>
<td></td>
</tr>
<tr>
<td>(GGT)$_2$GGG(GGT)$_2$(GGC)$_2$GAGG   23</td>
<td>23</td>
<td>4</td>
<td>510.0±70.8</td>
<td>12.2±1.4</td>
<td>3.0±0.5</td>
<td></td>
</tr>
<tr>
<td>(GGT)$_2$GGG(GGT)$_2$(GGC)$_2$GAGG   24</td>
<td>24</td>
<td>5</td>
<td>745.4±123.9</td>
<td>19.7±2.4</td>
<td>5.4±2.2</td>
<td></td>
</tr>
<tr>
<td>(GGT)$_2$GGG(GGT)$_2$(GGC)$_2$GAGG   25</td>
<td>25</td>
<td>6</td>
<td>570.0±213.9</td>
<td>17.0±3.4</td>
<td>2.5±0.5</td>
<td></td>
</tr>
</tbody>
</table>

Values are given as mean±s.e.m.

---

**Table 3 Independent predictive role of GGN genotypes for testosterone, FT, and LH.**

<table>
<thead>
<tr>
<th></th>
<th>Testosterone (ng/dl)</th>
<th>FT (ng/dl)</th>
<th>LH (U/l)</th>
</tr>
</thead>
<tbody>
<tr>
<td>β</td>
<td>P</td>
<td>β</td>
<td>P</td>
</tr>
<tr>
<td>Age</td>
<td>$-0.184 ± 0.002$</td>
<td>&lt;0.001</td>
<td>$-0.59 ± 0.001$</td>
</tr>
<tr>
<td>BMI</td>
<td>$-0.284 ± 0.007$</td>
<td>&lt;0.001</td>
<td>$-0.92 ± 0.006$</td>
</tr>
<tr>
<td>CAG</td>
<td>$0.114 ± 0.008$</td>
<td>&lt;0.001</td>
<td>$0.082 ± 0.007$</td>
</tr>
<tr>
<td>GGN</td>
<td>$0.059 ± 0.013$</td>
<td>0.017</td>
<td>$0.052 ± 0.011$</td>
</tr>
</tbody>
</table>

Only GGN included as a main effector. Age, BMI, and CAG were included as cofactors. Values are given as $\beta±S.E.M$. P values ≤0.05 are considered significant.
In any case, in an exploratory analysis for a subpopulation of men for which E2 was available (n=1076), we did not observe an association between the GGN repeat number and (free)E2 levels. Also, inclusion of E2 in our linear regression models had no major effect on the association between the GGN repeat number and FT (data not shown).

In epidemiologic investigations, somewhat conflicting results have been reported for the association between the GGN polymorphism and prostate cancer risk. An increased risk of prostate cancer has been associated with a short number (25, 26) or high number (27) of GGN repeats. The most common allele has been reported to be correlated with androgenic baldness, while 24 GGN repeats seemed to have a protective role (13). Lower semen volumes were found in men with GGN lengths shorter than 23 (14). A decreased risk of breast cancer has been described for women with long GGN alleles (28), whereas others (16) found a reverse relationship. Overall, clinical and epidemiological findings on a possible effect of the GGN repeat polymorphism on androgen-related clinical effects are thus far from univocal.

To the best of our knowledge, only two studies have investigated the influence of GGN repeat numbers on serum androgen levels. In one study in men with prostate cancer (n = 134 and median age = 67), Giwercman et al. (17) observed that the free testosterone index, calculated as testosterone/SHBG, was higher in men with GGN > 23 when compared to those with GGN < 23 (P = 0.057) and GGN = 23 (P = 0.027). In a second cohort study of 305 healthy young Swedish men (14), Lundin et al. found no difference in testosterone and FT levels between the three groups of subjects (GGN < 23, GGN = 23 and GGN > 23). The modest amplitude of the association between GGN and F(T) demonstrated in our study, would indeed require a larger cohort than in the latter study to be picked up.

In the large cohort of men investigated in our study, we detected a significant but very weak negative association between the GGN and CAG repeat length.

If the GGN repeat does influence AR function, then its relationship with the androgen levels may be related to its influence on the transactivating activity of the AR. Although the exact mechanism remains unclear, several studies report an influence of the GGN repeat on the transactivation function of the AR. Observations by Ding et al. (8), suggested an association between shorter GGN repeats and higher transactivating activity of the AR, with the number of repeats ranging from 19 to 23, seems compatible with our results. Indeed, shorter GGN repeats would then be related to an increased androgen feedback and ultimately to reduced testosterone levels. However, limited data from the in vitro studies on GGN repeat length in relation to AR action have not been univocal. The observation of Goa et al. (9) indicated that a complete deletion of the GGN repeat resulted in a 30% decrease of the AR transcriptional activity. Also, Werner et al. (11) found that shortening of the GGN repeat leads to a gradual reduction of AR transactivating capacity in CHO-cells. Furthermore, higher repeat numbers have been associated with higher translation efficiency of the AR gene resulting in higher protein levels and AR activity (7). The discrepancies among these in vitro studies are most likely caused by differences in the study design. The range of CAG or GGN repeat examined, cell type, and choice of promoters for vectors all may add to the diversity of observed in vitro effects of GGN repeat length on AR action.

Although it is well established that the (GGC)n part of the repeat is polymorphic, less is known about the variations in other regions of the repeat. Several point mutations in the AR have been found in men with androgen insensitivity syndrome (29) but also in healthy men (30). So far in all reports, the cytosine to thymine (C–T) nucleotide conversion, resulting in (GGT)3GGG(GGT)3(GGC)n or (GGT)3GGG(GGT)4(GGC)n has always been accompanied by a GGN repeat number (n) of 20 (29, 30). This is in contrast to our cohort, where this mutation also coincided with a repeat number (n) of 12, 16, 18, 19 and 21. In our study population, two new point mutations were noticed in the GGN repeat and one new C to T conversion was found just outside the consensus sequence of the repeat. However, the relevancy of these alterations remains unclear. Although these mutations did not seem to give rise to profound changes in a male phenotype, it can not be excluded that they might have an impact on AR action and androgen-related phenotypes. Indeed, even though the alterations do not affect the polypeptide sequence, synonymous codon substitution could alter the rate of mRNA translation, hereby influencing the protein folding leading to increased levels of protein misfolding (31). We were not able to establish an association between these mutations in the GGN repeat and androgen levels because the small number of observations for each mutation would only allow for detection of drastic effects on serum testosterone.

In conclusion, our study in a large cohort of healthy men revealed for the first time a positive association between FT and GGN repeat polymorphism of the AR. These findings support the hypothesis that the AR polymorphism affects androgen action with respect to modulation of androgen feedback action and testosterone secretion. Although these findings seem to be in line with reports of an association of the GGN repeat length with androgen-related clinical features, the latter findings were rather inconsistent and the effect of the GGN repeat polymorphism on serum FT in the present study is of fairly modest amplitude. Thus, the present study adds credence to the view that the polyglycine tract in the AR modulates AR action, but this effect appears to be small and its clinical relevance needs to be established.

www.eje-online.org
Declarations of interest

The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

Funding

This work was supported by the Fund for Scientific Research (FWO Vlaanderen Grant G0662-07). V Bogaert is a PhD student funded by Bijzonder Onderzoeksfonds (BOF) of the Ghent University. G Vanbillomont is a PhD student funded by the Flemish Fund for Scientific Research.

Acknowledgements

The authors are grateful to K Toye, I Bocquaert, and K Mertens, for their excellent technical assistance.

References

1 Albertelli MA, Scheller A, Brogley M & Robins DM. Replacing the mouse androgen receptor with human alleles demonstrates glutamine tract length-dependent effects on physiology and tumorigenesis in mice. Molecular Endocrinology 2006 20 1248–1260.


Received 29 March 2009
Accepted 19 April 2009