Evaluation of IP-RP-HPLC for Length Determination of the Trinucleotide Repeat Fragments in Huntington's Disease

Marzena Skrzypczak, Wolfram Heinritz, Anna M. Schulz, Marek Mierzejewski, and **Ursula G. Froster***
Institute of Human Genetics, University of Leipzig, Philipp-Rosenthal-Str. 55, 04103 Leipzig, Germany

Abstract

Expansion of an unstable trinucleotide $(CAG)_n$ repeat region within exon 1 of the gene *IT15* causes autosomal, dominantly inherited Huntington's disease (HD). The number of CAG-repeats varies from 6 to 35 in normal individuals, whereas in affected patients the expanded allele contains 40 or more CAG-repeats. Thus, exact determination of both alleles of the gene (normal and expanded) on the molecular level is of great importance for clinical diagnosis and prognosis regarding the course of the disease. In our study, we optimized and evaluated a highly sensitive, automated, and economical molecular method for length characterization of the trinucleotide fragment expansion such as $(CAG)_n$ repeat region based on ion-pair reversed-phase high-performance liquid chromatography (IP-RP-HPLC). We found that IP-RP-HPLC can be used for exact fragment length measuring between 60–280 bp as a sensitive and advantageous alternative method to conventional techniques.

Introduction

Expansion of the number of trinucleotide repeats is the underlying cause of more than 20 neurological and genetic disorders, such as Huntington's disease (HD; MIM#143100), dystrophia myotonica (DM1; MIM#160900), Fragile-X syndrome (MIM+309550), spinal and bulbar muscular atrophy (SBMA; MIM#313200), or different types of spinocerebellar ataxia (SCA2; MIM#183090) (1–3).

HD is an autosomal, dominantly inherited neuropsychiatric and movement disorder. An unstable $(CAG)_n$ repeat region within the coding exon 1 of the IT15 gene (also known as huntingtin), localized on the short arm of chromosome 4 (4p16.3), is responsible for HD (4,5). The variable number of $(CAG)_n$ repeats is related to different alleles and phenotypes. Normal alleles have up to 35 CAG repeats, but those containing between 28–35 CAG copies (intermediate allele) may expand in the next generation if transmitted by a male to his offspring (6,7). Incomplete penetration is known for alleles from 36 to 39 CAG-repeats; the clinical phenotype is not clearly predictable (8). In affected patients, the CAG-repeat number ranges from 40 to more than 100 in the juvenile form of the disease (9,10).

The identification and exact determination of CAG repeats in the pathologically expanded allele is necessary to confirm the clinical diagnosis and to give a prognosis regarding the course of the disease (11,12). Most of the methods used for this purpose are time-consuming, laborious, and expensive. Especially fragment length analysis by polyacrylamide slab gel electrophoresis (PAGE), which is at present the most popular method for detection and exact determination of trinucleotide repeat expansions, has certain disadvantages. Beside the expense and amount of work, it requires fluorescent-labeled probes (13). It has already been shown that detection of the expanded trinuclotide repeats is possible using a denaturing high-performance liquid chromatography (DHPLC) for pre-implantation genetic diagnosis; however, it was not used for accurate sizing the affected and unaffected alleles (14). The aim of our study was to develop a highly sensitive, automated, and economical molecular approach for the characterization of alleles related to one of the trinucleotide repeat disorders, HD. For these purposes, we used a new methodology combination followed by ion-pair reversed-phase high-performance liquid chromatography (IP-RP-HPLC).

Experimental

DNA samples

To establish this new method, we used DNA samples from 22 patients with clinically diagnosed HD, which were collected between 1998–2005. All DNA samples were isolated from peripherial blood according to standard procedures using the Invisorb Blood Giga Kit 100 (Invitek, Berlin-Buchs, Germany) and stored in TE buffer at 4°C in the DNA bank of the Institute of Human Genetics, University of Leipzig. The informed consent of the patients was given in written form.

Amplification of the CAG-repeat region

A commercially available kit for amplification of GC-rich fragments was used: TaKaRa LA Taq Polymerase (TaKaRa Mirus Bio,

Author to whom correspondence should be sent: email: reichsi@medizin.uni-leipzig.de.

Madison, WI) and the following primers as previously described by Warner et al. (15): HD1- 5'-atgaaggccttcgagtccctcaagtccttc and HD3- 5'-ggcggtggctgttgctgctgctgctgc.

PCR was carried out in a volume of 50 µL containing 0.5 µL TaKaRa LA Taq Polymerase (5 units/µL), 25 µL 2 × GC buffer (5 mM Mg^{2+}), 8 µL dNTP mix (2.5 mM of each dNTP), 1 µL of each primer (10 pmol), and 7 µL DNA (350 ng). PCR procedure started with initial denaturation at 94°C for 1 min, followed by 35 cycles of denaturation (94ºC for 30 s), annealing (70ºC for 30 s) and extension (72ºC for 2 min), and a final extension step at 72ºC for 7 min using the Mastercycler Gradient (Eppendorf, Hamburg, Germany). Amplification effect was controlled by electrophoresis in a 2.5% metaphor agarose gel (Serva, Heidelberg, Germany).

IP-RP-HPLC analysis

HPLC analysis was performed on a semi-automated Wave Nucleic Acid Fragment Analysis System (Model HT 3500) using the DNASep cartridge (Transgenomic, Omaha, NE). The PCR products (15 µL of each sample) were eluted at a flow rate of 0.9 mL/min with a linear increasing acetonitrile gradient using the universal gradient program. The mobile phase consisted of a mixture of two buffers (A + B), each containing 0.1 M of the ion-pair reagent triethylammoniumacetate (TEAA) and additional 25% acetonitrile (buffer B). The initial percentage of buffer B was 39%, and the final percentage was 61%. Run time was 16 min under non-denaturing conditions (50ºC).

DNA fragments were detected by UV-absorption at 260 nm. For fragment length assignment with the WAVE-specific Navigator software, we evaluated commercially available and parallel running 20-bp-DNA molecular weight markers (5 µL for injection, range 20–500 bp) from Lonza, Bio Science Rockland (Rockland, ME).

Multiple linear three-point regression was created by the Navigator™ software using the "size calling" option. The first regression graph was defined for two standard points below and one point above the detected fragment size, and the value was determined. Another graph was created from one standard point below and two standard points above, and a second value was assigned. Both values were averaged to determine the size of the fragment (16).

Standard analysis by PAGE

To confirm the results of IP-RP-HPLC, the fragment lengths for each patient were also resolved by means of polyacrylamide gel electrophoresis on an ABI Prism 377 and ABI Prism 3100-Avant Genetic Analyzer (Applied Biosystems, Foster City, CA). Fluorescent dye-labelled PCR products (primer sequences as described under Amplification of the CAG-repeat region section) were distributed by PAGE and analyzed in comparison to the GeneScan 500 TAMRA standard using the GeneScan and Genotyper analysis software (Applied Biosystems) for polyacrylamide gel electrophoresis and GeneMapper software for capillary electrophoresis. No significant differences between these two methods were observed. Further we use 'PAGE' as the name for both analysis methods.

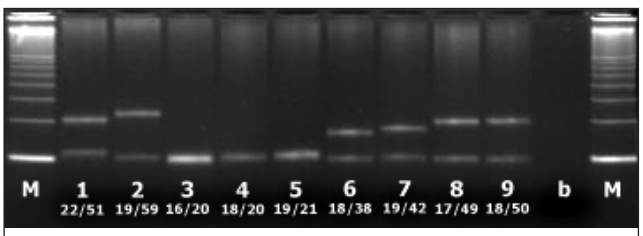

Figure 1. Gel electrophoresis in 2.5% metaphor agarose for the amplification products. Lanes M = DNA size marker (100 bp). Lanes 1–2, 6–9 = positive controls with 51, 59, 38, 42, 49, and 50 CAG-repeats on the affected alleles, respectively. Lanes 3, 4, 5 = normal controls, and Lane b = reagent blank.

Statistical analysis

The average allele size (number of corresponding CAG-repeats) was calculated for each method. The difference and 95% confidence interval of the mean between allele length was established as well. We compared IP-RP-HPLC with PAGE by per-

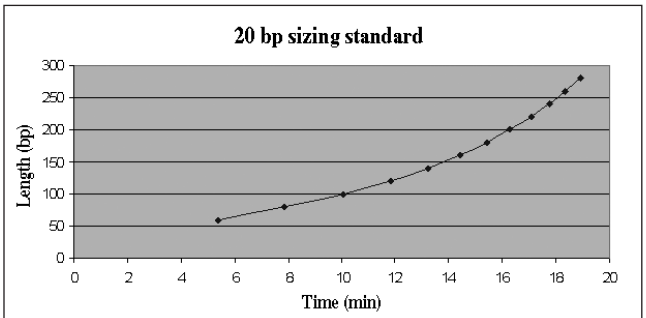

Figure 2. Plot of DNA fragments vs. retention time from 60 to 280 bp for 20-bp-DNA standard. Retention time was automatically determined by Navigator software.

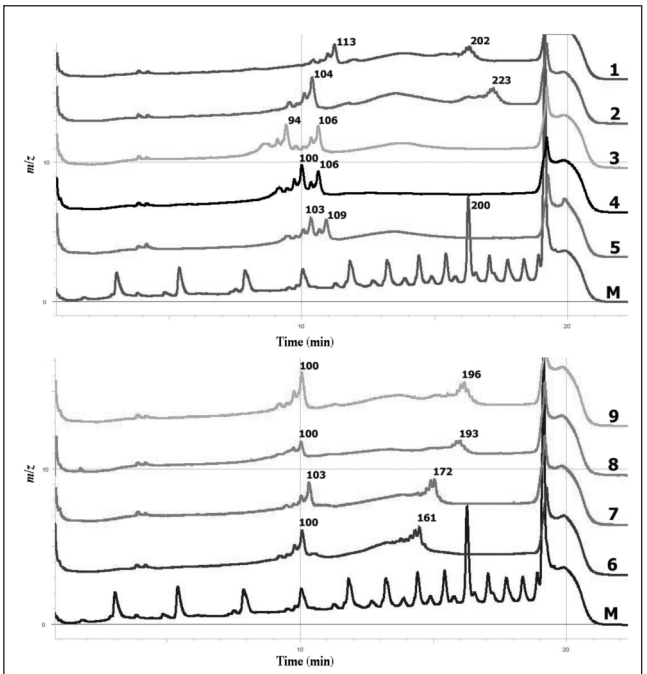

Figure 3. Chromatogram pattern of analysis with the 20-bp-DNA standard: (A) curves 1, 2 = samples with 51 and 59 CAG-rp on the affected allele; curves 3, 4 , 5 = normal samples with 16/20, 19/21, 18/20 CAG-rp, respectively; M = 20-bp-DNA standard. (B) Curves 6, 7, 8, 9 = samples with 38, 42, 49, and 50 CAG-rp on the affected allele, respectively; M = 20-bp-DNA standard.

forming paired t-student test for both alleles for each patient. For the DNA standards, the correlation coefficient to the PAGE method was also calculated.

Results and Discussion

Amplification of the CAG-repeat region

Amplification of sequences containing expanded trinucleotide repeat regions is difficult using conventional methods. Many scientific groups have already worked on a simple and reproducible protocol to amplify the CAG-repeat expansion within the IT15 gene (9,14,17,18). The IT15 gene regions are GC-rich and tend to form certain secondary structures. Therefore, certain PCR modifications are necessary to amplify these sequences e.g., the methylation sensitive conversion of C to U residues by bisulphite treatment of single stranded DNA (18) or the exchange of dGTP with 7-deaza-dGTP (17).

In our study, the commercially available Takara LA Taq Polymerase System, which has been developed to amplify GC-rich regions, was tested. This product contains a mixture of a "simple" Taq Polymerase with a second proofreading polymerase optimized for the amplification of long DNA templates. Moreover, a buffer was specially designed to amplify DNA regions with high GC-content or with a significant amount of secondary structures. The amplified products were qualitatively evaluated by means of gel electrophoresis (Figure 1). This was the first assay to screen for and to distinguish between normal (lane 3, 4, 5) and expanded (lane 1–2, 6–9) alleles in the probands.

DNA size standard profiles on HPLC

For accuracy of sizing and homogeneity of conditions, it was indispensable to run a new injection of 20-bp-DNA standard for analyses on different days and for each change of buffer (16,19). Kinetics of the outflowing fragments represents hyperbolical dependency of the DNA fragment length in a linear buffer gradient B in time (Figure 2).

Sizing

The size of the CAG-repeat regions (and the corresponding number of CAG-repeats) for 22 German HD patients was calculated with two linear three-point regression curves in order to receive a mean value (16). This procedure was necessary to avoid errors in measurement due to the non-linear, hyperbolic kinetics of standards and the samples. The results were compared to those received with the polyacrylamide electrophoresis as classical method (Figures 3 and 4). PAGE analysis based on an ABI Prism 377 is often replaced now by the application of capillary sequencers. Therefore, the probes were additionally sized by means of ABI Prism 3100-Avant Genetic Analyzer (Applied Biosystems), but there was no significant difference in number of sized CAG-repeats observed.

For each individual, the shorter CAG-repeat fragment is known as the first allele and the longer fragment as the second allele. While comparing IP-RP-HPLC and PAGE, significant differences were found for the second allele, but the differences were not significant for the first allele analyzed against the 20 bp DNA standard (Table I). The mean differences for the determined number of CAG-repeats for the first and second allele were 0.35 and 0.72, respectively (Table I).

Furthermore, the mean of difference for the number of CAG-repeats of each allele does not go over 3 between the two methods (IP-RP-HPLC vs. PAGE). The tolerance of error was recently suggested by the EMQN (20) with ± 1 CAG-repeat for normal alleles and up to ± 3 CAG-repeats for alleles over 40 CAG-repeats. The possible explanation for these differences could be usage of non-denaturing IP-RP-HPLC conditions. Previously, such differences were also observed when comparing denaturing and non-denaturing gel electrophoresis (21). Another plausible reason for discrepancy may be the usage of size standards of random sequence, which differ in GC percentage (22).

However, there are some disadvantages of IP-RP-HPLC. Due to the fixed resolution capacity of IP-RP-HPLC (~ 2% discrimination power), there is an upper limit for exact determination of fragment size and number of trinucleotide repeats. It means that only for fragments up to approximately 300–400 bp, it is possible to distinguish between two adjacent allele sizes. Furthermore, the chromatographic kinetics switch to an exponential function

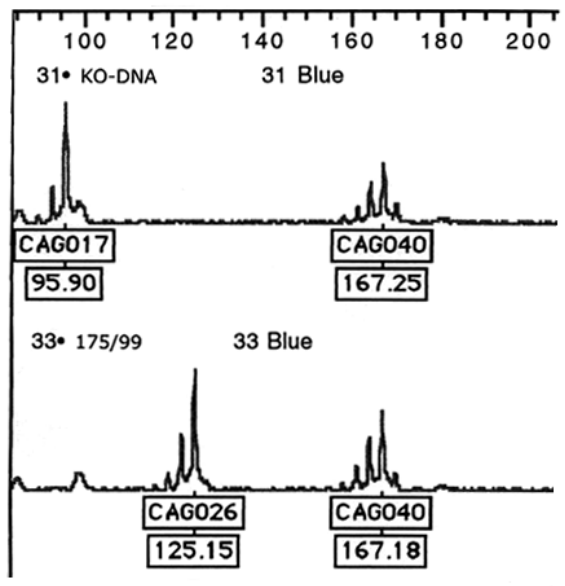

Figure 4. PAGE electropherogram of analysis with the GeneScan 500 TAMRA standard patient 175/99 (bottom) and positive control sample.

Table I. HPLC Compared to PAGE for Measuring Both Alleles by Paired Student t-Test

	HPLC		PAGE		HPLC vs PAGE*		
	Mean in 22 patients	SD	Mean in 22 patients	SD	Mean of differences (95% CI†)	p-value	Correlation coefficient
Allele 1 (Shorter)	18.44	2.755	18.09	2.942	0.349	0.111 (–0.087–0.784)	0.943
Allele 2 (Longer)	38.72	11.84	38.00	11.92	0.724	0.003 (0.274–1.172)	0.996

* By paired student t-test; † CI = confidence interval.

for large fragments and linear regression fails at fragment size determination. Thus, for those rare cases of HD patients with an expanded allele containing > 100 CAG-repeats (juvenile onset HD), IP-RP-HPLC is unsuitable for exact determination of these alleles but may be sufficient to detect them with higher measurement error.

Moreover, HPLC appears absolutely inapplicable for diagnostic purposes regarding the detection of very large trinucleotide repeat expansions as observed in patients with Fragile-X syndrome (> 200 CGG-repeats within the promoter region of the *FMR1* gene) or dystrophia myotonica type 1 (up to several thousand CTG-repeats within the 3'-UTR of the *DMPK* gene). For these analyses, special PCR methods (e.g., long range PCR) or Southern Blot techniques are advisable.

Altogether, the obtained results of IP-RP-HPLC were concordant with the results obtained by the standard method (PAGE) and in accordance with the suggestions of the EMQN. The results of our study imply that IP-RP-HPLC is a suitable alternative method for fragment sizing compared with PAGE. The semi-automated procedure, time, and cost efficiency are obvious advantages of IP-RP-HPLC. Trinucleotide repeat expansions can be detected by chromatographic technique with high resolution. Moreover, unlabeled amplification products can be analyzed without a prior denaturation. It has already been shown that IP-RP-HPLC was successfully used for sizing of short tandem repeats (STR) within HUMTHO1 locus (23). Thus, IP-RP-HPLC is not only a good alternative diagnostic method for HD analysis but also for other diseases caused by trinucleotide repeat expansions, such as SBMA and different types of SCA, (unpublished result) and for sizing of other repeat fragments (23).

Acknowledgments

This study was partly supported by Max-Buchner-Forschungsstiftung (Grant 2498) and Marzena Skrzypczak was supported by KAAD (Bonn/Germany).

References

1. G. Bates and H. Lehrach. Trinucleotide repeat expansions and human genetic disease. *Bioessays* **16(4):** 277–284 (1994).
2. R.R. Sinden, V.N. Potaman, E.A. Oussatcheva, C.E. Pearson, Y.L. Lyubchenko, and L.S. Shlyakhtenko. Triplet repeat DNA structures and human genetic disease: dynamic mutations from dynamic DNA. *J. Biosci.* **27(1 Suppl 1):** 53–65 (2002).
3. S. Mohmood, A. Sherwani, F. Khan, R.H. Khanand, and A. Azfer. DNA trinucleotide repeat expansion in neuropsychiatric patients. *Med Sci. Monit.* **9(9):** RA237–245 (2003).
4. H.S. Wang, C.R. Greenberg, J. Hewitt, D. Kalousek, and M.R. Hayden. Subregional assignment of the linked marker G8 (D4S10) for Huntington disease to chromosome 4p16.1-16.3. *Am. J. Hum. Genet.* **39:** 392–396 (1986).
5. The Huntington's Disease Collaborative Research Group. Novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell* **72:** 971–983 (1993).
6. F.O. Walker. Huntington's disease. *Lancet* **369:** 218–228 (2007).
7. Y.P. Goldberg, B. Kremer, S.E. Andrew, J. Theilmann, R.K. Graham, F. Squitieri, H. Telenius, S. Adam, A. Sajoo, E. Starr, A. Heiberg, G. Wolff, and M.R. Hayden. Molecular analysis of new mutations for Huntington's disease: intermediate alleles and sex of origin effects. *Nat. Genet.* **5(2):** 174–179 (1993).
8. D.C. Rubinsztein, J. Leggo, R. Coles, E. Almqvist, V. Biancalana, J.J. Cassiman, K. Chotai, M. Connarty, et al. Phenotypic characterization of individuals with 30-40 CAG repeats in the Huntington disease (HD) gene reveals HD cases with 36 repeats and apparently normal elderly individuals with 36-39 repeats. *Am. J. Hum. Genet.* **59(1):** 16–22 (1996).
9. J.M. Milunsky, T.A. Maher, B.A. Loose, B.T. Darras, and M. Ito. XL PCR for the detection of large trinucleotide expansions in juvenile Huntington's disease. *Clin. Genet.* **64:** 70–73 (2003).
10. F.A. Nahhas, J Garbern, K.M. Krajewski, B.B. Roa, and G.L. Feldman. Juvenile onset Huntington disease resulting from a very large maternal expansion. *Am. J. Med. Genet.* **137A(3):** 328–331 (2005).
11. M. Duyao, C. Ambrose, R. Myers, A. Novelletto, F. Persichetti, M. Frontali, S. Folstein, C. Ross, et al. Trinucleotide repeat length instability and age of onset in Huntington's Disease. *Nat. Genet.* **4:** 387–392 (1993).
12. L. Djousse, B. Knowlton, M. Hayden, E.W. Almqvist, R. Brinkman, C. Ross, R. Margolis, A. Rosenblatt, et al. Interaction of normal and expanded CAG repeat sizes influences age at onset of Huntington disease. *Am. J. Med. Genet.* **119A:** 279–283 (2003).
13. T. Tóth, I. Findlay, B. Nagy, and Z. Papp. Accurate sizing of (CAG)n repeats causing Huntington Disease by fluorescent PCR. *Clin. Chem.* **43:** 2422–2423 (1997).
14. K.C. Drury, M.C. Liu, S. Lilleberg, S. Kipersztok, and R.S. Williams. Results on single cell PCR for Huntington's gene and WAVE product analysis for preimplantation genetic diagnosis. *Mol. Cell Endocrinol.* **183:** S1–S4 (2001).
15. J.P. Warner, L.H. Barron, and D.J. Brock. A new polymerase chain reaction (PCR) assay for the trinucleotide repeat that is unstable and expanded on Huntington's disease chromosomes. *Mol. Cell Probes* **7(3):** 235–239 (1993).
16. C.G. Huber, P.J. Oefner, and G.K. Bonn. Rapid and accurate sizing of DNA fragments by ion-pair chromatography on alkylated nonporous poly(styrene-divinylbenzene) particles. *Anal. Chem.* **67:** 578–585 (1995).
17. S. Hecimovic, I. Barisic, A. Muller, I. Petkovic, I. Baric, I. Ligutic, and K. Pavelic. Expand Long PCR for fragile X mutation detection. *Clin. Genet.* **52(3):** 147–154 (1997).
18. I. Panagopoulos, C. Lassen, U. Kristoffersson, and P. Aman. A novel PCR-based approach for the detection of the Huntington disease associated trinucleotide repeat expansion. *Hum. Mutat.* **13(3):** 232–236 (1999).
19. K. Munson, R. Haefele, A. Kuklin, D. Gjerde, and P. Taylor. Sizing of DNA fragments with the WAVE® Nucleic Acid Fragment Analysis System. Transgenomic Inc.: Application Note 107 (2000).
20. EMQN (European Molecular Genetics Quality Network; c/o Regional Molecular Genetics Laboratory, St Mary's Hospital, Hathersage Road, Manchester M13 0JH, UK). Final report EQA scheme for Huntington disease (2004).
21. P.D. 2nd Chastain, E.E. Eichler, S. Kang, D.L. Nelson, S.D. Levene, and R.R. Sinden. Anomalous rapid electrophoretic mobility of DNA containing triplet repeats associated with human disease genes. *Biochemistry* **34(49):** 16125–16131 (1995).
22. O. Bruland, E.W. Almqvist, Y.P. Goldberg, H. Boman, M.R. Hayden, and P.M. Knappskog. Accurate determination of the number of CAG repeats in the Huntington disease gene using sequence-specific internal DNA standard. *Clin. Genet.* **55(3):** 198–202 (1999).
23. J.M. Devaney, J.E. Girard, and M.A. Marino. DNA microsatellite analysis using ion-pair reversed-phase high-performance liquid chromatography. *Anal. Chem.* **72(4):** 858–864 (2000).

Manuscript received June 18, 2008;
Revision received September 16, 2008.