Anal. Chem. **2000,** *72,* 858−864

DNA Microsatellite Analysis Using Ion-Pair Reversed-Phase High-Performance Liquid Chromatography

Joseph M. Devaney,*,†,‡ James E. Girard,‡ and Michael A. Marino†

Transgenomic Inc., 11 Firstfield Road, Suite E, Gaithersburg, Maryland 20878, and American University, 4400 Massachusetts Avenue, Washington, D.C. 20016

Genotyping based on short tandem repeat (STR) regions is used in human identification and parentage testing, gene mapping studies, cancer diagnostics, and diagnosis of hereditary diseases. Analysis of STR systems using slab gel electrophoresis requires lengthy and labor-intensive procedures. Therefore, alternative methods such as capillary electrophoresis or ion-pair reversed-phase high-performance liquid chromatography (IPRP HPLC) have been used to analyze DNA. IPRP HPLC offers an attractive substitute to gel electrophoresis for STR analysis because of the reduced analysis time, and there is no need for the waste disposal associated with radioisotopic, enzyme-linked, or fluorescence detection systems. We evaluated the use of IPRP HPLC for the sizing and typing of STR alleles from the HUMTHOl locus. The IPRP HPLC conditions (column temperature, flow rate, percent organic modifier per minute) were optimized for the separation of PCR products. Using the optimized separation conditions, the alleles of the HUMTHO1 system were sized in their native state (double stranded) with the use of internal markers. The typing results correlated 100% to accepted methods of DNA typing. The analysis time for the HUMTHOl locus was less than 14 min, and the alleles could be peak captured for further examination following such as sequencing.

The genetic uniqueness of individuals is a central tenet of human biology. One method to examine the uniqueness of an individual is with microsatellite or short tandem repeat (STR) length polymorphisms. STRs are tandemly repeated units of DNA sequence ranging between 2 and 7 base pairs (bps) in length that have high levels of heterozygosity, distinguishable alleles or different forms of a gene that are polymorphic based on the number of repeating units, and are amplified using the polymerase chain reaction (PCR).

STR systems are useful for disease diagnostics (fragile X, Huntington disease),[1,2] paternity testing,[3,4] physical and genetic mapping,[5−7] cancer detection,[8−10] and forensics.[11,12] The human genome has an abundance of these sequence occurrences averaging one tri- or tetranucleotide repeat system every 15 kb.[13] There are over 2000 STR systems (di-, tri-, tetra-, and pentanucleotide repeats) mapped out in the human genome, and it is estimated that there are thousands more to be discovered.[14]

One example of a STR system is the locus HUMTHO1.[15,16] This particular locus or location on a chromosome is a tetrameric (4-bp repeat) system used extensively in the forensic community because PCR amplification of this system yields fewer artifacts, such as extra band production and repeat slippage or stuttering, that can interfere with analysis.[17−19] The HUMTHO1 locus is located in intron 1 of the human tyrosine gene and has a repeating unit of (AATG).[20] Because this STR system is well characterized, it is used as the model for our paper.

Many different methods are used to establish the number of repeating units of a STR loci. One of the most widely used

* Corresponding author: (phone) 240-631-2001 ext 34; (fax) 240-6312440; (e-mail) jdevaney@transgenomic.com.

† Transgenomic Inc.

‡ American University.

(1) Patsalis, P. C.; Sismani, C.; Stylianou, S.; Ioannou, P.; Joseph, G.; Manoli, P.; Holden, J. J.; Hettinger, J. A. *Am. J. Med. Genet.* **1999**, *84*, 217−220.

(2) Bruland, O.; Almquist, E. W.; Goldberg, Y. P.; Boman, H.; Hayden, M. R.; Knappskog, P. M. *Clin. Genet.* **1999**, *55*, 198−202.

(3) Thomson, J. A.; Pilotti, V.; Stevens, P.; Ayres, K. L.; Debenham, P. G. *Forensic Sci. Int.* **1999**, *100*, 1−16.

(4) Sirchia, S. M.; Garagiola, I.; De Andreis, C. Gazzoli, I. Gramegna, M. *Mol. Cell Probes* **1996**, *10*, 155−158.

(5) Shaw, S. H.; Carrasquillo, M. M.; Kashuk, C.; Puffenberger, E. G.; Chakravarti, A. *Genome Res.* **1998**, *8*, 111−123.

(6) Mansfield, D. C.; Brown, A. F.; Green D. K.; Carothers, A. D.; Morris, S. W.; Evans, H. J. Wright, A. F. *Genomics* **1994**, *24*, 225−233.

(7) Hughes, A. E. *Genomics* **1993**, *15*, 433−434.

(8) Schullerus, D.; von Knobloch, R.; Chudek, J.; Herbes, J. Kovacs, G. *Int. J. Cancer* **1999**, *80*, 22−24.

(9) Schmitt, C.; Cire, K.; Schattenkirchner, S.; Pollok, M.; Baldamus, C. A.; Krug, B.; Kirchner, T.; Wunsch, P. H.; Krieg, T.; Staak, M. *Transplant. Int.* **1998**, *11*, 382−386.

(10) Noviello, C.; Courjal, F.; Theillet, C. *Clin. Cancer Res.* **1996**, *2*, 1601−1606.

(11) Wallin, J. M.; Buoncristiani, M. R.; Lazruk, K. D.; Fildes, N.; Holt, C. L.; Walsch, P. S. *J. Forensic Sci.* **1998**, *43*, 854−870.

(12) Halos, S. C.; Chu, J. Y.; Ferreon, A. C.; Magno, M. M. *Forensic Sci. Int.* **1999**, *101*, 27−32.

(13) Beckman, J. S.; Weber, J. L. *Genomics* **1992**, *12*, 627−631.

(14) Wang, Y.; Ju, J.; Carpenter, B. A.; Atherton, J. M.; Sensbaugh, G. F.; Mathies, R. A. *Anal. Chem.* **1995**, *67*, 1197−1203.

(15) Edwards, A.; Civitello, A.; Hammond, H. A.; Caskey, C. T. *Am. J. Hum. Genet.* **1991**, *49*, 746−756.

(16) Polymeropoulos, M. H.; Xiao, H.; Rath, D. S.; Merril, C. R. *Nucleic Acids Res.* **1991**, *19*, 4018.

(17) Levison, G.; Gutman, G. A. *Mol. Biol. Evol.* **1987**, *4*, 397−401.

(18) Schlotterer, C.; Tautz, D. *Nucleic Acids Res.* **1992**, *20*, 211−215.

(19) Sprecher, C. J., Puers, C.; Lins, A. M.; Schumm, J. W. *BioTechniques* **1996**, *20*, 266−276.

(20) Puers, C.; Hammond, H. A.; Jin, L., Caskey, C. T.; Schumm, J. W. *Am. J. Hum. Genet.* **1993**, *53*, 953−958.

10.1021/ac9908896 CCC: $19.00
Published on Web 01/19/2000

methods of STR analysis is polyacrylamide gel electrophoresis[21] with silver staining detection,[22] use of intercalation dye,[23,24] or fluorescent labeling[25] of the PCR product with the use of a fluorescent tag on the 5′-end of the primer. Other techniques accepted for the sizing and typing of STR systems are capillary electrophoresis[26−30] and capillary array electrophoresis.[31−33] In addition, new technologies such as matrix-assisted laser desorption/ionization time-of-flight mass spectrometry[34−36] are being used for STR analysis. All of these techniques represent powerful tools for the analysis of STR systems. However, each of these methods require special preparation such as cleanup and desalting of the PCR product,[31−36] diluting of the PCR product,[26] denaturation of the product,[27−30,32] labeling of the PCR product with a fluorescent tag,[27−32] or labeling the PCR product with biotin.[34] An analytical method that does not require any preparation step before the analysis of PCR product is ion-pair reversed-phase (IPRP) liquid chromatography.

Recently, IPRP liquid chromatography has become an analytical technique for the separation of oligonucleotides,[37] the separation and sizing of restriction enzyme (RE)-cleaved plasmid DNA,[38,39] the detection of mutations and single nucleotide polymorphisms,[40] the quantification of gene expression,[41] and the separation of PCR products.[38,39]

In this paper, we use IPRP as an analytical technique to size and type PCR products from the HUMTHO1 STR system. First, the separation parameters of column temperature, flow rate (mL/min), and linear gradient of percent acetonitrile per minute were optimized using restriction enzyme-cleaved plasmid ladders (RE ladders) and HUMTHOl PCR products. The optimal analysis conditions that gave the highest resolution were a column temperature of 54 °C, a flow rate of 0.8 mL/min, and a percent ACN per minute of 0.25. Using these conditions, we adapted a methodology used by Butler et al.[26] that applies dual standards to calculate a line of regression. From this linear regression analysis, any unknown peaks between the standards can be entered into the equation of the line and a number of base pairs assigned. Using a standard ladder constructed of PCR products, base pair sizes were calculated for alleles 5, 6, 7, 8, 9, 9.3, and 10 and used for the typing of unknown PCR products from the HUMTHO1 STR system. This analysis is rapid (<14 min) and requires no postamplification procedure to occur to the PCR products (DNA is left in its native state), and the PCR product could be peak captured for further analysis (i.e., sequencing).[42]

MATERIALS AND METHODS

Chemicals. HPLC grade acetonitrile (ACN; EM Science, Gibbstown, NJ) and triethylammonium acetate (TEAA; Applied Biosystems, Foster City, CA) were used to constitute the mobile phase.

DNA Extraction. Whole blood samples were collected in 5-mL tubes containing the anticoagulant EDTA. DNA was extracted from 300 μL of whole blood using the Puregene DNA isolation kit (Gentra Systems, Inc., Minneapolis, MN) following the manufacturer's protocol.

STR Amplification. The yields of purified genomic DNA isolated from nucleated cells fell within the range of 5−15 μg for 300 μL of whole blood. The concentrations were determined by UV spectroscopy using a spectrophotometer.

DNA amplification was performed as described in the Perkin-Elmer ABI PRISM STR protocol (Perkin-Elmer, Foster City, CA). The amplified tetrameric STR locus is HUMTHOl. This locus is amplified using primers from the Forensic Service of the British Home Office for use in forensic casework.[42]

Preparation of the Mobile Phase. The mobile phase consists of 0.1 M TEAA (solvent A) and 0.1 M TEAA−25% ACN (solvent B). To keep the concentration of TEAA constant and unaffected by volume contraction during the mixture of organic solvents with water, the mobile phase was prepared as follows: for solvent B, 50 mL of the 2 M TEAA stock solution was added to 250 mL of ACN, and immediately this was diluted to 950 mL with deionized water. The flask was capped and inverted several times to release any gas. Then, the final volume was adjusted to exactly 1000 mL with deionized water.

Instrumentation. The HPLC system used for purification of markers and analysis of PCR products is a Transgenomic (San Jose, CA) nonmetallic PEEK system HPLC with a variable-wavelength detector set at 260 nm, a DNAsep column (Transgenomic), and an autosampler with the capacity to handle 96 samples. The Transgenomic system uses Hitachi model D-7000 chromatography data station software for data analysis.

Chromatographic Conditions. The chromatographic conditions explored were column temperature, flow rate (mL/min), and

(21) Hammond, H. A.; Jin, L.; Zhong, Y.; Caskey, C. T.; Chakraborty, R. *Am. J. Hum. Genet.* **1994**, *55*, 175−189.
(22) Bassam, B. J.; Caetano-Anolles, G.; Gresshoff, P. M. *Anal. Biochem.* **1991**, *196*, 80−83.
(23) Zhu, H.; Clark, S. M.; Benson, S. C.; Rye, H. S.; Glazer, A. N.; Mathies, R. A. *Anal. Chem.* **1994**, *66*, 1941−1948.
(24) Lahiri, D. K.; Zhang, A.; Nurnberger, J. I. *Biochem. Mol. Med.* **1997**, *60*, 70−75.
(25) Lins, A. M.; Micka, K. A.; Sprecher, C. J.; Taylor, J. A.; Backer, J. W.; Rabback, D. R.; Bever, R. A.; Creacy, S. D.; Schumm, J. W. *J. Forensic Sci.* **1998**, *43*, 1168−1180.
(26) Butler, J. M.; McCord, B. R.; Jung, J. M.; Lee, J. A.; Budowle, B.; Allen, R. O. *Electrophoresis* **1995**, *16*, 974−980.
(27) Zhang, N.; Yeung, E. S. *J. Chromatogr., A* **1997**, *768*, 135−141.
(28) Gao, Q.; Yeung, E. S. *Anal. Chem.* **1998**, *70*, 1382−1388.
(29) Wenz, H.-M.; Roberston, J. M.; Menchen, S.; Oaks, F.; Demorest, D. M.; Scheibler, D.; Rosenblum, B. B.; Wike, C.; Gilbert, D. A.; Efcavitch, J. W. *Genome Res.* **1998**, *8*, 69−80.
(30) Lazaruk, K.; Walsh, P. S.; Oaks, F.; Gilbert, D.; Rosenblum, B. B.; Menchen, S.; Scheibler, D.; Wenz, H. M.; Holt, C.; Wallin J. *Electrophoresis* **1998**, *19*, 86−93.
(31) Wang, Y.; Ju, J.; Carpenter, B. A.; Atherton, J. M.; Sensabaugh, G. F.; Mathies, R. A. *Anal. Chem.* **1995**, *67*, 1197−1203.
(32) Mansfield, E. S.; Robertson, J. M.; Vainer, M.; Isenberg, A. R.; Frazier, R. R.; Ferguson, K.; Chow, S.; Harris, D. W.; Barker, D. L.; Gill, P. D.; Budowle, B.; McCord, B. R. *Electrophoresis* **1998**, *19*, 101−107
(33) Gao, Q.; Pang, H.; Yeung, E. S. *Electrophoresis* **1999**, *20*, 1518−1526.
(34) Ross, P. L.; Belgrader, P. *Anal. Chem.* **1997**, *69*, 3966−3972.
(35) Taranenko, N. I.; Golovlev, V. V.; Allman, S. L.; Taranenko, N. V.; Chken, C. H.; Hong, J.; Chang, L. Y. *Rapid Commun. Mass Spectrom.* **1998**, *12*, 413−418.
(36) Butler, J. M.; Li, J.; Shaler, T. A.; Monforte, J. A.; Becker, C. H. *Int. J. Legal Med.* **1999**, *112*, 45−49.
(37) Huber, C. G.; Oefner, P. J.; Bonn, G. K. *Anal. Biochem.* **1993**, *212*, 351−358.
(38) Huber, C. G.; Oefner, P. J.; Bonn, G. K. *Anal. Chem.* **1995**, *67*, 7, 578−585.
(39) Huber, C. G.; Oefner, P. J.; Preuss, E.; Bonn, G. K. *Nucleic Acids Res.* **1993**, *21*, 1061−1066.
(40) Underhill, P. A.; Jin, L.; Lin. A. A.; Mehdi, S. Q.; Jenkins, T.; Vollrath, D.; Davis, R. W.; Cavalli-Sforza, L. L.; Oefner, P. J. *Genome Res.* **1997**, *7*, 996−1005.
(41) Heyward-Lester, A.; Oefner, P. J.; Doris, P. A. *BioTechniques* **1996**, *20*, 250−257.
(42) Marino, M. A.; Devaney, J. M.; Smith, J. K.; Girard, J. E. *Electrophoresis* **1998**, *19*, 108−118.

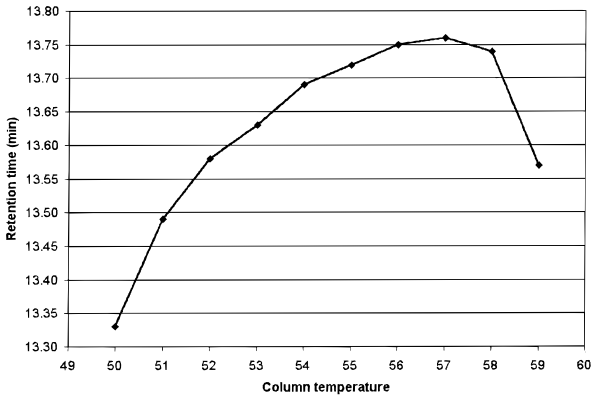

Figure 1. Retention time for a 173-bp HUMTHO1 product with increasing column temperature. The other chromatographic conditions are given in the text.

percent ACN per minute. For the optimization of conditions, the starting percentage of B was 40 and the ending percentage of B was 65. Once the optimization of conditions was completed, the HPLC conditions for the sizing and typing of the HUMTHO1 locus were a column temperature of 54 °C, a flow rate of 0.8 mL/min, and a percent ACN per minute of 0.25. The starting and ending points for the linear gradient were 43% B to 57% B.

RESULTS AND DISCUSSION

Optimization of the Chromatographic Conditions. For this study, the chromatographic parameters consisting of column temperature, flow rate, and organic modifier concentration per minute were optimized for the separation of dsDNA in the form of restriction enzyme-cleaved plasmid DNA and PCR products from the HUMTHO1 locus. The RE-cleaved dsDNA serves as markers for the sizing and typing experiments. The results were the same for the RE-cleaved DNA and the PCR-amplified DNA; therefore, only the data from the PCR-amplified DNA are shown.

Column temperature is an important parameter for dsDNA separation using liquid chromatography because the column temperature has been shown to affect the resolution between fragments and have an affect on the physical properties of a DNA fragment.[39,40,43] The effects of column temperature on the separation of dsDNA in the form of PCR-amplified DNA from the HUMTHO1 locus were examined. The chromatographic conditions were held constant throughout the column temperature at a flow rate of 0.9 mL/min and a linear gradient where the percentage of ACN was increased 0.5%/min (2.0% B/min). The column temperature was the only chromatographic condition changed during the analyses.

One of the first effects of raising the temperature of the column is that the dsDNA fragment (173-bp HUMTHO1 homozygote sample) is retained longer in the stationary phase (Figure 1). With an increase in temperature of the column, the dsDNA begins to straighten due to the defolding and derotating of the double helix.[39] The defolding and elimination of any secondary structure in the DNA helix exposes more phosphate groups on the DNA backbone for ion-pairing formation with the TEAA ion. This allows stronger interaction with the alkylated polymer beads, giving a longer retention time because a higher percentage of acetonitrile will be required to move the dsDNA from the stationary phase into the mobile phase at higher temperatures.

Table 1. Change in Column Temperature and the Effect on the Change in Retention Time of a HUMTHO1 (9.3, 9.3)

temp 1 (T_1, °C)	retention time for T_1 (R_1, min)	temp 2 (T_2, °C)	retention time for T_2 ($R_{2,\,\mathrm{min}}$)	Δ ($R_2 - R_1$) (min)
50	13.33	51	13.49	0.16
51	13.49	52	13.58	0.09
52	13.58	53	13.63	0.05
53	13.63	54	13.69	0.06
54	13.69	55	13.72	0.03
55	13.72	56	13.75	0.03
56	13.75	57	13.76	0.01
57	13.76	58	13.74	−0.02
58	13.74	59	13.57	−0.17

The HUMTHO1 PCR product increases in retention time (RT) from 50 to 57 °C, but the RT decreases once the column temperature reaches 57 °C. Table 1 represents the change in retention time for each degree change in column temperature. The temperature change from 50 to 51 °C is the most substantial change in RT. As can be seen in Table 1, the increase in temperature causes the RT to reach a plateau where column temperature increases do not affect the retention time substantially. The temperature change from 53 to 54 °C causes a larger change in retention time than the changes in retention time that occur from 54 to 55, 55 to 56, and 56 to 57 °C. It is possible that the dsDNA helix for the HUMTHO1 (9.3, 9.3) sample has defolded and derotated to reach a maximum linear state at the temperature of 54 °C. Therefore, any increases in temperature are not going to produce any further ion-pairing sites for the TEAA to interact with on the DNA. Once the column temperature reaches 57 °C, the dsDNA begins to denature, thus reducing the RT.

The same temperature gradient was repeated on fragments of different RE-cleaved plasmid DNA (pBR322 cleaved with *Hae*III, pUC18 cleaved with *Hae*III, and with *Msp*I) because this sizing method will be used for all different types of PCR products (different sequences, different base pair lengths, different percentages of adenine and thymine). Therefore, fragments of different base pair sizes and different sequences were analyzed with increasing column temperature. The results were the same as the temperature gradient performed on the PCR product except until the column temperature reached 55 °C. Upon examination of the chromatograms, the 458-bp peak from pBR322 cleaved with *Hae*III showed a decrease in retention time at a temperature of 55 °C (Figure 2). Panel A shows five peaks (434, 458, 504, 540, and 587 bp) from the pBR322 cleaved with *Hae*III at a column temperature of 54 °C. Panel B shows the same fragments separated at 55 °C; only the 458-bp fragment coelutes with the 434-bp fragment. The decrease in retention time of the 458-bp fragment is caused by the partial denaturation of the helix. Because of the partial denaturation of this fragment at a temperature of 55 °C, the sizing integrity while using the RE ladders would be compromised. Therefore, 54 °C is the upper limit for sizing when using the RE ladders. The optimum temperature for analysis of the dsDNA fragments analyzed in this study lies below 55 °C because all possible markers have to be in a nondenatured state because

(43) Huber, C. G.; Berti, N. *Anal. Chem.* **1996**, *68*, 2959−2965.

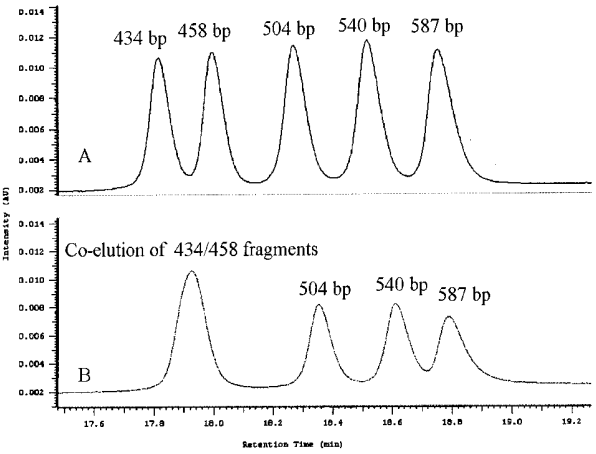

Figure 2. Separation of RE fragments at the column temperatures of 54 (A) and 55 (B) °C.

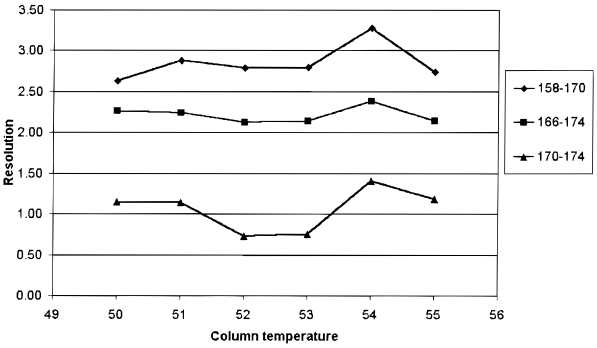

Figure 3. Resolution for HUMTHO1 PCR products (bp) with increasing column temperature (°C).

any unwinding of the dsDNA helix for any marker would compromise sizing accuracy.

To find the optimal temperature for separation of dsDNA, different allele combinations of HUMTHO1 PCR products were analyzed using column temperatures starting at 50 °C and were repeatedly analyzed as the column temperature was raised by 1-deg increments to 55 °C (Figure 3). The resolution was calculated using the equation

$$R_s = 2(t_2 - t_1)/W_1 + W_2 \qquad (1)$$

where R_s is resolution, t_n is retention time of a component, and W_n is peak width of a component. The resolution for all samples increased until reaching a maximum at 54 °C. At 55 °C, the resolution dropped. The increase in retention time and increase in resolution with increasing column temperature counters the fact that as a DNA molecule increases in size, the dsDNA becomes more randomly coiled in free solution.[44] However, in the chromatographic column, the dsDNA is not in free solution. The DNA flows through the narrow channels of a packing material. This causes stretching of the DNA molecules by shear forces.[44] Therefore, the effect of the increased ion-pairing interaction with the exposed phosphate groups and the removal of secondary structure in the dsDNA helix contributes to increased resolution for HUMTHOl allele pairs.

(44) Ringo, M. C.; Evans, C. E. *Anal. Chem.* **1997**, *69*, 643−649.

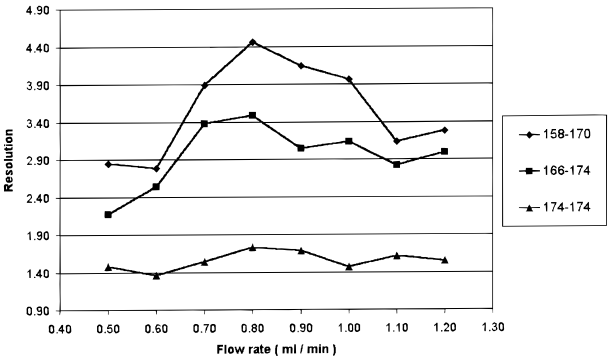

Figure 4. Resolution for HUMTHO1 PCR products (bp) with increasing flow rate.

The results for the HUMTHO1 alleles point to a column temperature of 54 °C as the optimum column temperature for the sizing and typing experiments. At this column temperature, no partial melting of any of the HUMTHO1 alleles occurs. With an optimum column temperature for the analysis of STR products, the next step is to find the proper flow rate for optimum resolution.

The chromatography used in this study has a mobile phase that contains a polar compound. Consequently, flow rate should effect the separation of dsDNA.[45−49] From earlier work dealing with the separation of RE fragments, an increase in flow rate exerted a pronounced increase in resolution between a 162- and 210-bp fragments.[39] However, STR alleles can differ by as little as 1 bp.

For the HUMTHO1 alleles, the flow rate was examined from 0.5 to 1.2 mL/min in increasing steps of 0.1 mL/min. The conditions of separation included a column temperature of 54 °C and a linear increase in the percentage of ACN per minute of 0.5.

Figure 4 is the result of resolution calculations using eq 1 for the HUMTHO1 PCR products with increases in the flow rate. From the figure, the resolution for all alleles reaches a maximum at a flow rate of 0.8 mL/min. For the HUMTHO1 (9, 10) alleles, a resolution of 1.73 was attained. This resolution (≥1.5), according to the literature, allows for a degree overlap of less than 1%, and mixtures with concentrations as high as 10:1 could easily be distinguished.[50]

The optimal flow rate (0.8 mL/min), coupled with the optimal column temperature (54 °C), was used to find the optimal percentage of ACN per minute for analysis of dsDNA using the DNAsep column.

The optimization of IPRP liquid chromatography involves the relationship of the effects of pH, organic modifier type, organic modifier concentration, and the effects of the type of IP agent. For the separation of dsDNA using the DNAsep column, the pH, organic modifier type, concentration of organic modifier, and IP agent have been optimized.[39] However, the effect of the organic modifier concentration per minute has not been examined. The organic modifier concentration can control the separation based on the mechanisms proposed for IPRP separation.[50,51]

(45) Tanaka, N.; Yoshimura, T.; Araki, M. *J. Chromatogr.* **1987**, *406*, 247−256.
(46) McGuffin, V. L.; Evans, C. E. *J. Microcolumn Sep.* **1991**, *3*, 513−520.
(47) McGuffin, V. L.; Evans, C. E.; Chen, S. H. *J. Microcolumn Sep.* **1993**, *5*, 3−10.
(48) Hirabayshi, J.; Kasai, K. *Anal. Biochem.*, **1989**, *178*, 336−341.
(49) Ksai, K. *J. Chromatogr.* **1993**, *618*, 203−221.
(50) Potter, R. L.; Lewis, R. V. *High Performance Liquid Chromatography Advances and Perspectives*, 4th ed.; Academic Press: Orlando, 1986; Chapter 1.

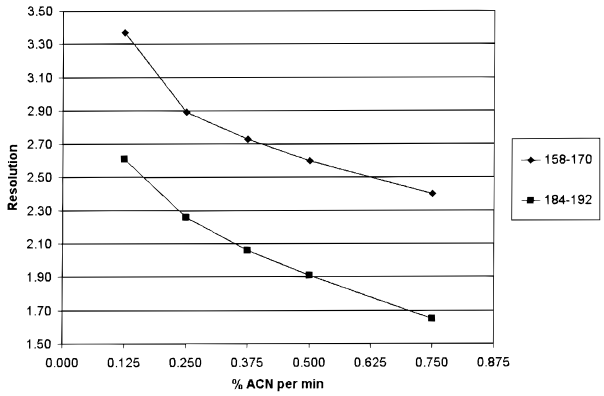

Figure 5. Resolution for HUMTHO1 PCR product and RE fragments (bp) with increasing percent ACN per minute.

For the analysis of dsDNA, the rate of increase in percent ACN per minute was changed, while the same chromatographic conditions were used. The flow rate was 0.8 mL/min, the column temperature was 54 °C, and the gradient conditions were linear. The changes in ACN were 0.125, 0.25, 0.375, 0.5, and 0.75%/min.

To examine the effects of changing the percent ACN per minute on resolution (R_s), the HUMTHO1 6, 9 alleles and the 184−192 bp (Δ bp = 8) RE fragments were analyzed using the conditions discussed above (Figure 5). The R_s increased as the percent ACN per minute decreased. This correlates well with the difference in retention time as the percent ACN per minute changes. At the lower values, the dsDNA fragments have more time to interact with the stationary phase and allow for maximum IP agent interaction. However, time becomes a factor. The analysis at 0.125% ACN/min gave the highest resolution values but the total separation time was over 25 min (data not shown), whereas with 0.25% ACN/min, the total analysis took just over 15 min. With this in mind, 0.25% ACN/min is used for all dsDNA analysis.

The optimal conditions for the analysis of PCR-amplified dsDNA and RE fragments are a column temperature of 54 °C, a flow rate of 0.8 mL/min, and a rate of increase in ACN of 0.25%/min. The actual starting and ending points for a linear gradient are calculated on the basis of the size of dsDNA to be separated.

Sizing and Typing of the HUMTHO1 Locus. The starting and ending points for the linear gradients are calculated on the basis of the size of dsDNA that is to be separated. This section uses the optimized column conditions (column temperature, flow rate, percent ACN per minute) found in the optimization of chromatographic conditions.

To measure the length (bp) of a DNA fragment in gel electrophoresis, a relationship is established between the mobility of an unknown fragment and the mobility of fragments of known length. With gel electrophoresis, problems arise in the fact that gel inhomogeneities exist. In addition, temperature and voltage gradients across the gel affect the electrophoretic mobility of dsDNA fragments, and the sequence of dsDNA affects the electrophoretic mobility. However, the idea of using the mobility of dsDNA fragments of known length to size unknown fragments has become a popular method to size dsDNA.

The above method of using the mobility of DNA fragments of known size to establish the base pair size of unknown fragments was used to size and type alleles PCR-amplified from the HUMTHO1 locus using IPRP HPLC. We used the retention times of RE fragments of known base pair size to draw a line of regression. With this line of regression, a retention time of an unknown peak interpolates with the equation of the line between the standards to find the number of base pairs in the unknown peak. In the case of the HUMTHO1 locus, all alleles for the United States population fall between 154 and 174 bp.[52] Therefore, the markers used for the sizing of the HUMTHO1 locus should be <154 and > -174 bp.

Table 2. Sizing Method Using Two Markers Applied to a HUMTHO1 (9.3, 9.3)

markers (bp)	allele bp size	calcd bp size	Δbp between sizing method and allele size
123, 213	174	183.17	9.17
124, 213	174	182.64	8.64
147, 213	174	180.74	6.74
123, 192	174	178.82	4.82
124, 192	174	178.54	4.54
147, 192	174	177.93	3.93
147, 190	174	175.95	1.95
147, 184	174	175.15	1.15

The sizing of PCR fragments has been accomplished using two internal markers or standards.[28] Butler et al. used 150- and 300-bp fragments as markers to size a STR system by drawing a least-squares fit line between the two markers.[26] The retention times of the alleles of the STR system were fit to the least-squares line and a base pair size was calculated. This same methodology was applied to IPRP HPLC to size HUMTHO1 alleles.

The first method examined for sizing of HUMTHO1 alleles was the use of RE fragments as markers. The size of the marker (fragment) could have an effect on the sizing accuracy. Therefore, different sized markers were tested for the accuracy of sizing the dsDNA HUMTHO1 PCR products of known base pair length. The markers that allowed for the highest accuracy in sizing would be used to size a group of unknown samples.

The markers to be used in this study are the 123-, 124-, 184-, 192-, and 213-bp fragments from the pBR322 plasmid cleaved with the *Hae*III RE and the 147- and 190-bp fragments from the pUC18 plasmid cleaved with the *Msp*I RE. All markers were IPRP HPLC purified before use as sizing markers.

Markers of different sizes were combined and a HUMTHO1 (9.3, 9.3) homozygote sample was analyzed three times (Table 2). The size of the HUMTHO1 (9.3, 9.3) allele is according to the primer sites and sequence is 173 bp. However, an extra base pair is added to the size of the PCR product with the use of *Termus aquaticus* (Taq) polymerase[53] making the PCR product 174 bp. Taq has the ability to catalyze the nontemplated addition of nucleotides to the 3′-termini of blunt-ended DNA duplexes.[29,53] Therefore, the PCR-amplified DNA has two extra bases attached as overhangs on the 3′-termini of the PCR products.

The markers that gave the size calculations closest to the actual allele size were the 147- and 184-bp markers, as can be seen in

(51) Molnar, I.; Horváth, Cs. *Clin. Chem.* **1976**, *22*, 1497−1502.

(52) Budowle, B.; Smerick, J. B.; Keys, K. M.; Moretti, T. R. *J. Forensic Sci.* **1997**, *42*, 846−849.

(53) Clark, J. M. *Nucleic Acids Res.* **1988**, *16*, 9677−9686.

(54) Magnuson, V. L.; Ally, D. S.; Nylund, S. J.; Karanjawala, Z. E.; Rayman, J. B.; Knapp, J. I.; Lowe, A. L.; Ghosh, S.; Collins, F. S. *BioTechniques* **1996**, 21, 700−709.

Table 3. Sizing Method Applied to HUMTHO1 Alleles Using RE Fragments of 147 and 184 bp ($n = 5$ for Each Allele)

HUMTHO1 allele	allele bp size	av size using RE fragments	std dev of allele size	Δbp between sizing method and allele size
6	159	160.11	0.39	1.11
7	163	164.47	0.38	1.47
8	167	168.56	0.33	1.56
9	171	172.32	0.26	1.32
9.3	174	174.55	0.34	0.55
10	175	176.40	0.34	1.40

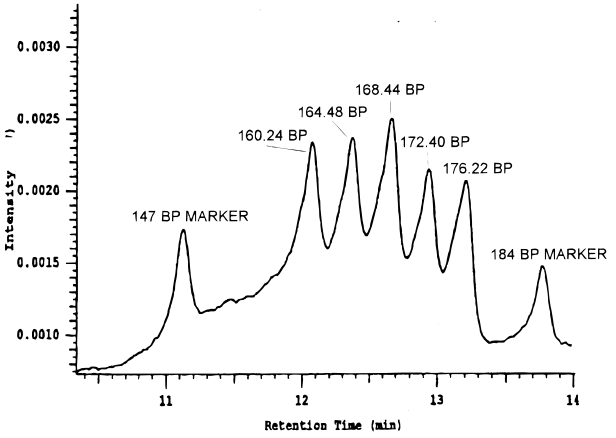

Figure 6. Sizing of HUMTHO1 alleles (6, 7, 8, 9, 10) using 147- and 184-bp markers. The ladder is composed of PCR products from the HUMTHO1 locus (alleles 6, 7, 8, 9, and 10). The calculated base pair sizes are above the alleles. The markers are from the pUC18-Msp I and pBR322−*Hae*III ladders. The separation conditions were as follows: column temperature of 54 °C, flow rate of 0.8 mL/min, and a gradient of acetonitrile of 0.25%/min.

Table 2. Therefore, these markers are used for all the analyses from this point on.

The marker pair consisting of the 147- and 184-bp fragments was used to fit the HUMTHO1 samples using least-squares analysis (Table 3). The table contains the results of the sizing using the two markers, the standard deviation for each from the actual size, and the number of base pairs each calculated size differed from the experimental size from the primers.

The calculated size for each allele was on the average 1.28 bp larger than the size of the allele calculated from the intron sequence with the addition of an extra base pair from the DNA polymerase,[30] and the standard deviation for the sizing using the 147- and 184-bp markers was on the average 0.34 bp. When two markers were used to size the dsDNA (Table 2), it was found that the closer the base pair size of the marker to the base pair size of the target DNA, the more accurate the sizing assay.

Since the base pair size generated for the alleles with the use of the two markers (147 and 184 bp) is larger than the size of the PCR product, a standard for sizing HUMTHO1 PCR products of unknown length is needed. Therefore, the sizes generated in Table 3 are used as a standard for comparing unknown samples; a ladder was made of the HUMTHO1 6, 7, 8, 9, and 10 alleles (an example chromatogram is in Figure 6) and sized using the markers. The sizing of the 9.3 allele follows this (an example chromatogram is in Figure 7). The 9.3 allele coelutes with the 10 allele and therefore

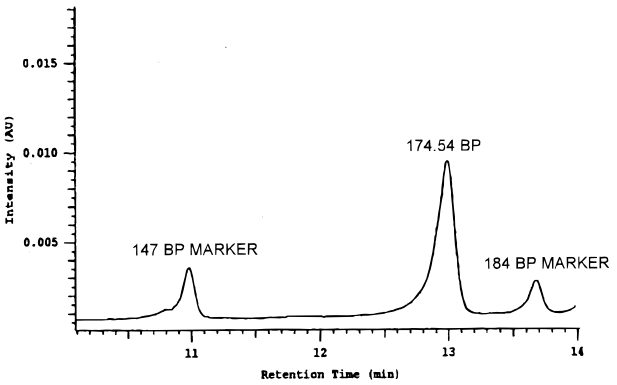

Figure 7. Sizing of HUMTHO1 9.3 allele using 147- and 184-bp markers. The calculated base pair size is above the allele. The markers are from the pUC18−*Msp*I and pBR322−*Hae*III ladders. The separation conditions are the same as in Figure 6.

Table 4. Allele Base Pair Sizes ± 0.45 nt To Be Used for Typing the HUMTHO1 Locus

HUMTHO1 allele (type)	calcd size	calcd size range for each allele
6	160.11	159.66, 160.56
7	164.47	164.02, 164.92
8	168.56	168.11, 169.01
9	172.32	171.87, 172.77
9.3	174.29	173.84, 174.74
10	176.40	175.95, 176.85

to obtain a size for the allele it must be analyzed separately from the 10 allele.

To establish 99.7% confidence in binning with single nucleotide separation, nonoverlapping bins should be established with ±0.45-nucleotide range with average size of the alleles in a ladder.[29] This allows for a series of bins of 0.90 nucleotide to be established. Because 99.7% confidence is derived from 3 standard deviations of sizing precision, the ±0.45-nucleotide range is achieved with 0.15 nucleotide standard deviation in sizing. This method of typing is currently accepted for forensic casework in the United States and is used to analyze samples for both national and state convicted offender DNA databases. With the use of the ladder as a standard, a type can be given to alleles that fall within ±0.45 nt of the allele sizes for the ladder.

The base pair sizes for the HUMTHO1 alleles calculated from analysis with the internal standards were binned with ±0.45 nt (Table 4). This is used to type unknown samples. The table will allow for easy typing of the samples. Once an unknown sample's allele sizes are calculated, the calculated size is compared to Table 4, which contains the bins for each fragment. The data located in Table 4 are gathered before each series of 15 samples are analyzed.

The unknown samples were analyzed with a gradient of 47% eluent B to 58% eluent B, and the results for sizing and typing using Table 4 are shown in Table 5. The samples were analyzed on an ABI 310 genetic analyzer using standard protocols from the Armed Forces DNA Identification Laboratory (Rockville, MD), and the type for each sample was given.

Using the typing method with the 147-bp marker and the 184-bp marker allowed for the typing of all samples correctly. This typing procedure allowed for the correct typing of all 9.3 alleles,

Table 5. Allele Sizes and Types Calculated Using the 147- and 184-bp Markers and Table 4

sample	allele bp sizes	allele type according to Table 4	allele type using ABI 310 genetic analyzer
1	171.95, 173.95	9, 9.3	9, 9.3
2	168.73, 176.30	8, 10	8, 10
3	172.08, 176.05	9, 10	9, 10
4	173.92	9.3, 9.3	9.3, 9.3
5	159.76	6, 6	6, 6
6	164.53, 172.40	7, 9	7, 9
7	174.52	9.3, 9.3	9.3, 9.3
8	160.43, 168.90	6, 8	6, 8
9	164.13, 172.16	7, 9	7, 9
10	160.28, 172.34	6, 9	6, 9
11	164.14, 172.16	7, 9	7, 9
12	160.46, 172.57	6, 9	6, 9
13	160.29, 168.65	6, 8	6, 8
14	172.56, 174.67	9, 9.3	9, 9.3
15	164.40, 172.40	7, 9	7, 9

but when injected with the ladder, the 9.3 allele coelutes with the 10 allele. To allow the sizing and typing method to work, the 9.3 allele is analyzed separately from the ladder to generate a standard size. The separation of the 9.3 and 10 allele calls for single base pair resolution at 175 bp. Currently, single base pair separation can be obtained with fragments with a length up to 124 bp (data not shown). The conditions of separation do not allow for the single base resolution at the base pair size of the HUMTHO1 PCR products.

This sizing method can be used for any STR loci with any markers. In the case of the HUMTHO1 locus, the best markers were the fragments of sizes 147 and 184 bp because these markers gave relatively close base pair approximations. However, any markers could be used for sizing as long as a ladder is analyzed first to get the base pair sizes for each allele. With this method, there was no interference from components of the PCR mix with the size markers. Therefore, the method of using size markers composed of RE-cleaved plasmid DNA accurately sizes PCR products from STR loci using HPLC as a separation method.

CONCLUSION

The separation of dsDNA on a HPLC with a DNAsep column was optimized with the use of RE fragments and PCR products with respect to column temperature, flow rate (mL/min), and percent ACN per minute. The optimal conditions for the highest resolution and fastest separation were a column temperature of 54 °C, flow rate of 0.8 mL/min, and 0.25% ACN/min. With these optimal conditions, the resolution obtained for HUMTHO1 allele pairs ranged from 4.95 for a HUMTHO1 sample consisting of the allele pair 6, 10 (165 and 175 bp) to a resolution of 1.55 for a HUMTHO1 sample consisting of the allele pair 9, 9.3 (171 and 174 bp). These optimal conditions were used to size and type HUMTHO1 alleles in their natural state (double-stranded) with the use of RE-cleaved dsDNA as sizing markers. These conditions could be used with adjustments in the starting and ending points for eluent B to size and type any tetranucleotide STR system with a size range up to 350 bp.

ACKNOWLEDGMENT

We thank the members of the Center for Molecular and Medical Genetics for their support and for the use of the facility. J.M.D. is grateful to the National Science Foundation for the Chemometric Fellowship that supported the research.

Received for review August 5, 1999. Accepted December 3, 1999.

AC9908896