Identification of a Novel Exon and Spliced Form of Duffy mRNA That Is the Predominant Transcript in Both Erythroid and Postcapillary Venule Endothelium

By Sadahiko Iwamoto, Jianping Li, Toshinori Omi, Sigenori Ikemoto, and Eiji Kajii

The Duffy gene has been shown not to be split by introns, even in its 5' untranslated region, and to be expressed not only in erythroid but in postcapillary venule endothelium of almost every organ in the body. To further investigate the transcriptional start position in erythroid and postcapillary venule endothelium, we performed 5'-rapid amplification of cDNA ends (5'-RACE). While every positive clone of 5'-RACE encoded the identical sequence of previously identified cDNA downstream from nucleotide 203, the upstream sequences were different. The upstream sequences corresponded to the sequence from nucleotide −279 to −308/−357 in erythroblasts and from −279 to −356/−383 in lung and were regarded as comprising a novel exon. This novel exon encoded seven residues initiated with a methionine, linked to nucleotide 203 in-frame and in agreement with the GT-AG splicing rule. The major erythroid transcriptional start position was identified in human erythroblastic cells

THE Duffy blood group system consists of four major phenotypes: Fya(a+b−), Fyb(a−b+), Fy(a+b+), and Fya(a−b−). Antisera anti-Fya and anti-Fyb define the two principal antigens. The murine monoclonal antibody (MoAb) anti-Fyb6 defines another Duffy antigenic determinant present in all red blood cells except for those of the Fya(a−b−) phenotype. MoAb anti-Fyb6 precipitates a glycoprotein from Duffy-positive red blood cells, named glycoprotein D, of 35 to 45 kD. The Duffy cDNA and its genomic DNA have been cloned, and the one-base substitution resulted in an amino acid substitution at residue 44 that seemed to be responsible for deciding the Fya/Fyb allotetrameric system.

Fya(a−b−) erythrocytes are resistant to invasion by the malarial parasites Plasmodium vivax and P knowlesi. The receptors for parasite invasion of erythrocytes are related to the Duffy blood group antigen. This antigen has also been shown to function as a receptor for chemokines, including interleukin-8 (IL-8), monocyte chemotactic protein-1 (MCP-1), homologous genes of IL-8 receptors, which are the highly homologous genes of FY, are not split by introns within the coding regions but are split in the 5' and 3' untranslated regions (UTRs). The genes of IL-8 receptors, which are the highly homologous genes of FY, are not split by introns within the coding regions but are split in the 5' UTRs; il8ra is formed from two exons and il8rb is formed from 11 exons. From its broad organ distribution and the doubtful genomic structure of FY, we speculated that its 5' region should be complexed for tissue-specific transcription. Then, we planned to reanalyze the transcriptional start position by 5'-rapid amplification of cDNA ends (5'-RACE) for the subsequent reevaluation of the promoter region for organ-specific transcription.

MATERIALS AND METHODS

Cell culture. Normal erythroblasts were obtained from a Duffy-positive individual ([Fya(a+b−)]) by methylcellulose culture in a 35-mm non-tissue culture dish (Falcon, Lincoln Park, NJ) as previously described. On day 14, erythroid burst-forming unit (BFU-E)-derived red colonies were picked up under an inverted microscope and were observed by total RNA extraction using the acid-phenol guanidinium method.

Reticulocyte RNA was isolated from Duffy-positive individuals as previously described. Human erythroblastic cells by primer extension and in bone marrow by ribonuclease protection analysis at 34 bases upstream from the first ATG codon. Distinctively, in lung and kidney, the transcription was started at 62 bases upstream from the ATG. Both Northern blotting and reverse transcription-polymerase chain reaction followed by Southern analysis indicated a predominance of the novel spliced form of mRNA of about 50- to 200-fold comparing with the unspliced form, in every studied organ and erythroid lineage cells. The spliced form of cDNA has been transfected into a human erythroblastic cell line, K562, and the expressed protein reacted with Duffy-specific murine monoclonal antibody Fy6. These studies indicate that the product from the spliced form of mRNA is the major product of the Duffy gene in the erythroid lineage and postcapillary venule endothelium.

© 1996 by The American Society of Hematology.

From the Department of Legal Medicine and Human Genetics, Jichi Medical School, Tochigi, Japan.

Submitted March 10, 1995; accepted August 15, 1995.

Supported by Grant-in-Aid for Scientific Research No. 06404028 from the Ministry of Education, Science, and Culture of Japan.

Address reprint requests to Sadahiko Iwamoto, MD, Department of Legal Medicine and Human Genetics, Jichi Medical School, 331-1 Minamikawachi-machi, Kawachi-gun, Tochigi 329-04, Japan.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. section 1734 solely to indicate this fact.

© 1996 by The American Society of Hematology. 0006-4971/96/8701-0003$3.00/0

were obtained from the Japanese Cancer Research Resources Bank (Tokyo, Japan) and were maintained in RPMI 1640 medium containing 10% fetal calf serum (FCS). UT7 cells were provided by Dr. N. Komatsu (Department of Hematology, Jichi Medical School, Tochigi, Japan) and were maintained in Iscove’s modified Dulbecco’s medium (IMDM) containing 10% FCS and 1 ng/mL recombinant granulocyte-macrophage colony-stimulating factor (GM-CSF; Sumitomo Pharmaceutical Co, Osaka, Japan).

RNA preparation from human tissues. Lung, kidney, liver, spleen, bone marrow (BM), and brain tissues were prepared from a 54-year-old Japanese male patient under autopsy, after receiving permission for use for research purposes. Total RNA was extracted from 1 g of each tissue by the acid-phenol guanidinium method.

RNA analysis by RACE. To determine the complete sequences of the 5' UTR, we performed 5'RACE from the total RNA of BFU-E-derived cells and lung tissue according to the procedure of Frohman et al.11 with some modifications. An antisense Duffy glycoprotein-specific primer (ATTCCAGTGGACGTTGGAAAC, complimentary to nucleotide 1198 to 1218 of Fy71-81 cDNA) was used to synthesize cDNA. The cDNA was purified by ultrafiltration to remove excess primer and nucleotides, followed by poly(A)-tailed with terminal deoxynucleotidyl transferase (Life Technologies Inc, Gaithersburg, MD). The poly(A)-tailed cDNA was amplified with a nested antisense primer (GTGCAACCGCACTGTCGATAG, complimentary to nucleotide 735 to 716) and an anchor primer (ATCGTGGACATGCATAGCTAATG) in the following polymerase chain reaction (PCR) under stringent conditions (denaturation at 94°C for 1 minute, annealing at 55°C for 1 minute, and extension at 72°C for 3 minutes, for 35 cycles).

To determine the complete sequence of the novel cDNA, we performed 3'RACE. First-strand cDNA was synthesized with the anchor primer, and the 5' end-specific sense primer (TGTCTGCGGATTAGTGAC, complimentary to nucleotide 340 to 359, and ATCCGTCGACATCGATACGT) were used in the following reverse transcription (RT)-PCR (using a random primer labeling kit (Promega, Madison, WI). Poly(A)+ RNA was prepared with the PolyATract mRNA Isolation System II (Promega). The primers were subsequently amplified (RT-PCR) for 20 cycles with common anchor primer (TGTCTGCGGATTAGTGAC, complimentary to nucleotide 1203 to 1226 with Xho I linker) and specific primer for the novel 5' end (326 to 307) or for the 5' end of Fy71-81 (GGCTTCCCCAGGACTGTTCCTG, corresponding to nucleotide 1 to 22). The PCR products were separated on agarose gels and blotted on Hybond-N+ nylon membranes. The blotted membranes were hybridized with common ORF probe at 65°C for 12 hours. A final wash was performed at 65°C for 45 minutes in 0.1× SSPE with 0.1% SDS. The radioactivity of the bands was measured by a radioanalytic imaging apparatus (AMBIS Systems Inc, San Diego, CA). In the preliminary study, the amplification curves of exon 0.1 and Fy71-81 transcripts were linearly exponential and parallel until 27 cycles.

Construction of expression plasmid and transfection. The full-length cDNA with novel 5' end was amplified as above from an Fy(a-b+) individual and subcloned in pCR II vector, and the insert without misincorporation was cloned into pcDNAI vector (Invitrogen). The K562 cells were transfected with 2 μg of plasmid DNA by Lipofectamine (Life Technologies Inc) according to the manufacturer's instructions. Mock-transfected K562 cells were transfected with the same expression plasmid lacking the Duffy cDNA. Expression of the Duffy antigen in K562 cells was assayed by staining the cells with MoAb Fy6, which was provided by Red Cross Bloodbank (Osaka, Japan). The cells incubated with Fy6 were washed three times and were subsequently incubated with fluorescein isothiocyanate-conjugated goat anti-mouse IgG. After incubation, cells were washed three times with phosphate-buffered saline (PBS), and fluorescence was measured on a Spectrum III flow cytometer (Becton Dickinson, Mountain View, CA).

RESULTS

Isolation of a spliced form of FY cDNA. For further characterization of the 5' end of Duffy mRNA, we performed 5'-RACE and isolated seven positive clones from erythroblast RNA and nine clones from lung RNA. All clones en-
cated from nucleotide -278 to -355 (from nucleotide 203 to 383). The sequences were regarded as complete, since no obvious banding pattern was obtained with the Fy71-8 specific oligo.

To determine the difference in TSP between hematopoietic organs and nonhematopoietic organs, we attempted more sensitive RNAse protection study. In bone marrow RNA, the exon 0.1 antisense RNA probe made a main signal at nucleotide 212, from which 157 (from nucleotide 203 to 359) and 21 (seven codons in exon 0.1) should be subtracted. The resulting position of 34 bases from the initiating ATG, and the main transcript length is estimated at about 1,120 nucleotides. In contrast, no obvious banding pattern was obtained with the Fy71-81-specific oligo.

Relative quantification of the FY transcripts. To determine the relative level of expression of the two distinct FY mRNAs, three specific probes were used to three strips of Northern blots of HEL cell, BM, lung, and brain RNA. The common ORF probe detected 1.2-kb band in all tissues. For personal use only.

initiated with exon 0.1, we performed 3-RACE using the exon-specific primer and an anchor primer. We isolated four positive clones and showed that the spliced form of Duffy cDNA, which was temporarily designated Fy0.1 clone, shared an identical reading frame and a polyadenylation signal with the Fy71-81 cDNA downstream from nucleotide 203 (Fig 1B).

Characterization of the transcription start points (TSP) of FY mRNA in erythroid and postcapillary venule endothelium. To characterize the TSP for FY gene in erythroid cells without imposing a PCR step bias, poly(A)+ RNA of a human erythroleukemia cell line, HEL, were extended using antisense oligo primers corresponding to each 5' sequence and were analyzed in parallel on the same sequencing gel. The HEL cell has been already shown to express Duffy antigen/chemokine receptor. A reaction primed with an oligo nucleotide specific for exon 0.1 identified one major and two minor products: 54, 52, and 78 nucleotides, respectively (Fig 2). Thus, HEL cells, the major TSP lies 34 bp upstream from the initiating ATG, and the main transcript length is estimated at about 1,120 nucleotides. In contrast, no obvious banding pattern was obtained with the Fy71-81-specific oligo.
81 was not obtained. The resolving power of Northern blotting was insufficient to distinguish the transcriptional size between erythroid and the two organs studied, while from TSP analysis, the erythroid transcript is expected to be 48 nucleotides short compared with lung. The predominance of the Fy0.1 transcript was also revealed by RT-PCR study in lung, kidney, liver, spleen, BM, and brain (not shown).

Relative expression of the FY mRNA. Distribution of FY transcript in hematopoietic cell lines was tested by RT-PCR followed by Southern blot analysis with hybridization to a common ORF probe. The novel spliced-form FY cDNA was expressed strongly in reticulocytes, BFU-E–derived erythroblasts, HEL cells, UT7 cells (megakaryoblastic cell line with

Fig 2. Identification of the TSP in erythroid cells by primer extension. The extended products of HEL cell RNA were primed with antisense oligonucleotides specific for exon 0.1 and Fy71-81 transcripts and were separated on the same gel. In parallel, DNA sequencing reaction is shown for determination of product length. The gel was exposed to X-ray film with intensifying screen at -80°C for 7 days. Arrows with numerals indicate the extended products with the exon 0.1-specific primer. The Fy71-81-specific primer did not show any positive band.

Fig 3. RNase protection analysis. The antisense RNA probes for Fy0.1 (404 bases, arrow) and Fy71-81 transcript (470 bases) were synthesized from linearized plasmid containing each cDNA; the Fy0.1 clone encoded from nucleotide -383 to -279/203 to 359; the Fy71-81 clone encoded from nucleotide 1 to 359. The RNA probes were hybridized with 50 µg of BM, lung, or kidney total RNA and treated with RNase T1/RNase A; lane 1 is the product hybridized with Fy0.1 probe, lane 2 is with Fy71-81 probe. The major protected fragment with Fy0.1 probe in BM RNA is indicated by a bold arrow, and the fragments in lung and kidney are indicated by the open triangle. The protected fragments with Fy71-81 probe are indicated by an arrowhead. The Fy0.1 probes hybridized with tRNA and treated with (+) or without (-) RNase are shown in the right lanes. The radiolabeled RNA size markers are on the left.
erythroid features\(^1\)), and KU812 cells (premature basophilic cell line with erythroid features\(^2\)); weakly in K562 cells (erythroleukemic cell line), HL60 cells (promyelocytic cell line), and Raji cells (B lymphoid cell line; Fig 5); and very weakly in KG1 (myeloblastic cell line) and MOLT4 (T lymphoid cell line) cells, which expression was detected after prolonged exposure (not shown). The Fy71-81 cDNA was expressed weakly in reticulocytes, BFU-E, HEL, UT7 and KU812, but was not detected in K562, HL60, KG1, and MOLT4. Without reverse-transcription, no positive bands were detected in any PCR samples. The radioactivity of individual bands was measured by a radioanalytic imaging apparatus and is listed in Fig 5. The ratio of the Fy0.1 cDNA to the Fy71-81 cDNA was 1:0.018 in reticulocytes, 1:0.022 in BFU-E, 1:0.006 in HEL cells, 1:0.003 in UT-7 cells, and 1:0.010 in KU812 cells, respectively. In other cells, the Fy71-81 bands were too weak to estimate the accurate relative expression. In general, the results are consistent with the predominance of spliced-form mRNA suggested by primer extension, the RNase protection study, and Northern blotting. The spliced-form mRNA was expressed strongly in erythroid-committed cells, other than K562 cells. In myeloid (KG1, HL60) and lymphoid (Raji, MOLT4) cells, the transcription of FY was very weak.

**Surface expression of Duffy product on transfected K562 cells.** To investigate the influence of amino terminal modification on antigenicity, the FY cDNA with novel 5’-end and Fyb genotype was used to transfet K562 cells. The transfected cells were harvested and stained with Duffy antigen-specific MoAb Fy6. Positive staining was obtained on the cells transfected with pcDNA-Fyb0.1 (Fig 6). Approximately 5% of the transfected cells were stained with the antibody. No staining was observed when cells were transfected with the expression vector alone (mock), or cells were stained only with second antibody (Fig 6). In triplicate experiments, reproducible results were obtained.

**DISCUSSION**

We have previously shown that the Duffy gene was constructed in a single exon, and its predicted promoter sequence was highly pyrimidine-rich (82%), in contrast with most eukaryotic promoters, which are GC-rich or include TATA and CCAAT boxes. To reanalyze the TSP of FY in erythroid cells and to discover the TSP of postcapillary venule endothelium, we performed 5’-RACE and isolated a novel Fy0.1 cDNA that was distinct from the Fy71-81 clone of Chaudhuri et al\(^3\) at its 5’ end. We have not yet been able to isolate an identical clone to Fy71-81 using this procedure. The 5’ end of the novel cDNA was located at -278 bases upstream from the TSP of Fy71-81 and was identified as exon 0.1. The exon 0.1 encodes seven residues that are initiated with a methionine and connect with nucleotide 203 in-frame and in accordance with the GT-AG splicing rule. The Fy0.1 cDNA is regarded as a splicing isoform. A major erythroid TSP was identified by primer extension analysis with HEL cell RNA at 34 bp upstream from the first ATG. The TSP was confirmed by RNase protection assay with BM RNA. In lung and kidney RNA, the transcription started at 82 bp upstream from the ATG. Then, instead of the nine N-terminal residues of the Fy71-81 cDNA product, seven residues of the Fy0.1 cDNA are predicted to be a signal sequence for the initial membrane penetration into endoplasmic reticulum. The upstream sequence of the novel initial ATG is also lacking TATA or CCAAT boxes, but includes short GC stretches.

Northern blotting and RT-PCR followed by Southern blotting clearly showed the predominant expression of the Fy0.1 transcript in erythroid cells and in every organ studied. The predominance of the spliced-form transcript in erythroid cells is consistent with the absence of an obvious band in primer extension analysis primed with Fy71-81-specific oligo and with the unsuccessful cloning of Fy71-81 cDNA by 5’-RACE. Furthermore, dominance was also revealed by RNase protection assay. The Fy71-81 probe showed the splicing position mainly and exactly. Therefore, it is obvious that the Fy0.1 transcript is the main transcript in both erythroid and postcapillary venule endothelium.

Recently, Tournamille et al\(^3\) identified a novel cap site in the Duffy gene at 484 nucleotides upstream from the first ATG identified by Chaudhuri et al\(^3\) using 5’-RACE and the primer extension procedure. Their TSP is similar to our data in erythroid cells. However, a first ATG appears at the same
A SPliced FORM OF DUFFY mRNA

position as the Fy0.1 cDNA in their proposed long 5' non-coding region, and several termination codons emerge in the following intronic sequence. Thus, their cDNA seems to be the prespliced hnRNA. It was difficult to rule out the possibility that our RT-PCR products from Fy71-81 transcript were amplified from the hnRNA. The difference in the transcriptional size (0.4 kb) between Fy1.0 and Fy71-81, which was suggested by Northern blotting, supports this possibility.

The presence of multiple transcription initiation sites and tissue-specific promoters involved in tissue-specific gene expression have been reported for a growing number of other genes, such as band 3,24 band 4.1,11 glycoporin C,12 and glucocorticoid receptor.25 Duffy gene transcripts have been shown to be widely distributed in many organs, including lung, muscle, spleen, heart, pancreas, kidney, and brain.39 Postcapillary venule endothelium is the cell that expresses the Duffy blood group antigen in these organs.10 We showed that the spliced transcript was also the predominant transcript in those organs with a difference in TSP from erythroid cells. The TSP in lung and kidney was 48 bases upstream from that of the erythroid cell. The postcapillary venule endothelium-specific promoter must exist in the upstream. In addition, fetal brain has been shown to transcribe remarkably larger RNA compared with other organs and adult brain.35 The brain, Duffy gene should be transcribed in a time- and organ-specific manner. The manner in which the respective promoter initiates the respective transcription start sites must be determined in the future.

In RT-PCR analysis, the strong signal was observed restrictedly in erythroid-committed cells among eight leukemic cell lines. The Fy(a-b-) individuals do not produce FY mRNA in the BM,4 in accordance with the absence of Duffy glycoprotein on their erythrocytes. However, Chaudhuri et al8 showed that lung, spleen, colon, and kidney of Duffy-negative individuals produce mRNA of the same size but in less quantity than those of Duffy-positive individuals. Moreover, Peiper et al26 showed the expression of Duffy glycoprotein in endothelial cells of Fy(a-b-) individuals. Tournamille et al27 showed that a T-to-C substitution at 46 bases upstream from their proposed cap site, which we also found in three Duffy-negative, black individuals living in Japan, abolished erythroid expression. The mutation disrupted the inverted consensus binding site for the GATA transcription factor (double underscore in Fig 1B). The GATA motif is positioned downstream from the TSP of lung and kidney and 30 bp upstream from the TSP of erythroid cells. This GATA motif should play the core role for the promoter in erythroid cells but will not be essential to transcribe FY gene in postcapillary venule endothelium. If the endothelial cell lines expressing the Duffy product were available, this subject should be resolved.

The recombinant Fy71-81 was transfected and successfully expressed on K562 cells4 and human embryonic kidney cell line 293.7 The expressed proteins reacted with Duffy glycoprotein-specific murine MoAbs Fy6 and Fy3. We also showed that the Fy0.1 cDNA products expressed on K562 cells successfully reacted with MoAb Fy6. This fact indicates that the substitution of nine residues in Fy71-81 with the novel seven residues of Fy0.1 does not influence the epitope for Fy6 and the proper embedding of the glycoprotein into the plasma membrane. Therefore, we conclude that the product from spliced-form FY mRNA is the main product of red blood cells.

From the comparison of the deduced amino acid sequences of humans and simians, Chaudhuri et al8 proposed the recognition site of Fy6 to be the residue 22—40, not the N-terminus. Our expression study supports this speculation. We are now processing the binding study of the Fyb0.1
cloned product and promiscuous chemokines to determine the influence of the N-terminal substitution on the ligand binding. The dissociation constant value for IL-8 binding to K562 cells was slightly higher than those values for erythrocytes and HEL cells.8,9,10

ACKNOWLEDGMENT


REFERENCES

Identification of a novel exon and spliced form of Duffy mRNA that is the predominant transcript in both erythroid and postcapillary venule endothelium

S Iwamoto, J Li, T Omi, S Ikemoto and E Kajii