Genomic Structure, Chromosomal Localization, Start of Transcription, and Tissue Expression of the Human p40-phox, A New Component of the Nicotinamide Adenine Dinucleotide Phosphate-Oxidase Complex


p40-phox is a newly isolated cytosolic component of the nicotinamide adenine dinucleotide phosphate (NADPH)-oxidase that copurifies with p67-phox. Although its function is not well defined, preliminary evidence indicates that it is a component of the cytosolic complex. We report the characterization of the human p40-phox gene, which is single copy and spans approximately 18 kb with 10 exons. Based on fluorescent in situ hybridization (FISH) studies and analysis of somatic hybrid cell lines, the chromosomal location of p40-phox is human chromosome 22q13.1. The start of transcription has been mapped to bp -156. The expression of p40-phox is restricted to hematopoietic cells. In addition to identifying the mRNA transcript on Northern blot analysis in cells known to express components of the NADPH-oxidase, polymorphonuclear leukocytes, monocytes, B lymphoblastoid cell lines, and eosinophils, p40-phox is also expressed in two other cell types of white cell lineage, mast cells, and basophils. In addition, the mRNA for p40-phox is expressed in megakaryocytic cells, but not in erythrocytes.

© 1996 by The American Society of Hematology.

From the Pediatric Branch, National Cancer Institute, National Center for Human Genome Research, National Institutes of Health, Bethesda, MD; the Division of Molecular Medicine, University College London, London, England; and the Division of Human Genetics and Molecular Biology, The Children’s Hospital of Philadelphia, Philadelphia, PA.

Submitted December 27, 1995; accepted May 15, 1996.

Supported in part by National Institutes of Health Grant No. HG00425.

Address reprint requests to Stephen J. Chanock, MD, Pediatric Branch, National Cancer Institute, National Institutes of Health, Bldg 10, Room 13N240, Bethesda, MD 20892.

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.

2714-2721

Blood, Vol 88, No 7 (October 1), 1996: pp 2714-2721
Here we report the structure of the gene for p40-phox and demonstrate its location on human chromosome 22q13.1. The immediate 5' upstream region has been characterized and the start of transcription has been mapped to bp -156 (relative to the start of translation). The tissue specificity of the p40-phox gene expression is restricted to hematopoietic cells. We have determined that p40-phox is expressed at the message level in mast cells, basophils, eosinophils, and also in megakaryocytes. Furthermore, p40-phox mRNA is detectable in early undifferentiated myeloid cell lines.

MATERIALS AND METHODS

Isolation of genomic clones for the p40-phox gene. Genomic clones of the human p40-phox gene were isolated by screening a human genomic library (Embl3A) with a full-length cDNA probe labeled by random primer method. First strand cDNA was synthesized from RNA extracted from normal human monocytes using a commercial kit (Boehringer Mannheim, Indianapolis, IN) with avian myeloblastosis leukemia virus reverse transcriptase (RT) (C. Carter, Department of Transfusion Medicine, National Institutes of Health, Bethesda, MD). The cdNA probe for the screening of the genomic library was generated by polymerase chain reaction (PCR) using two unique oligonucleotides ATGGCTGTGCCCCAGACGTG and TCAATGGCATCGTGTTGTAGACCCT. The buffers and concentrations were used according to the manufacturer (Perkin-Elmer, Norwalk, CT). The conditions for the PCR reaction were as follows: 30 cycles of 94° for 1 minute, 55° for 1 minute, 72° for 2 minutes, followed by a 7-minute extension at 72°C. The PCR product was subcloned into the TA-vector, pCRII (Invitrogen, La Jolla, CA). The sequence was confirmed by the dideoxynucleotide chain termination method and compared with two separate clones isolated during the initial screening reported by Wientjes.19 A total of 1.5 X 10⁶ clones were lifted on nitrocellulose filters and hybridized overnight at 42°C in 50% formamide, 6 x SSC (saline sodium citrate solution), 50 mmol/L NaPO₄ pH 7.0, 0.1% sodium pyrophosphate, 6 x Denhard' s solution, 0.1 mg/ml salmon sperm DNA, and 0.1% sodium dodecyl sulfate (SDS). Four consecutive washes were performed at 42°C in 2 x SSC, 1 x SSC, 0.5 x SSC, and 2 x SSC. Two separate genomic phage clones were purified and analyzed by Southern blot analysis using unique, end-labeled oligonucleotide probes corresponding to the published cDNA sequence and by direct sequencing of subcloned fragments.

A single P1 clone was isolated from a diploid human P1 library (Human Genome System, Inc, St Louis, MO) by a PCR-based method of screening with two oligonucleotides, F7, GTGAAGATCCTCAAAAGACCTT and R6, GAGATCTCTCCCTCCAGCGA, which generated a 380-bp fragment spanning exon 8/intron 9. Sequencing of this clone was performed to confirm intron/exon boundaries using two different methods, direct sequence analysis and sequence analysis of subcloned fragments of P40-1 into either pBluescript KS+- (Stratagene, La Jolla, CA) or pCR-II (Invitrogen). A preliminary restriction map was constructed and used to confirm the findings of the analysis of the phage clones.

Analysis of genomic map and intron-exon borders. Intron-exon borders were identified by DNA sequence analysis using dideoxynucleotide chain termination method. Intron sequences were confirmed by both Southern blot analysis and PCR-generated fragments, which were subcloned and sequenced using oligonucleotides generated during the analysis.

Mapping the start of transcription. Primer extension was performed as follows. An oligonucleotide, PR4, TCACCTCTCAC-TCTCTCCAGCCAC complementary to the sense cDNA strand beginning at bp -61, upstream of the start site of translation, was labeled with ³²P-adenosine triphosphate (P-ATP) and using polymerase 5'- to 3' activity of approximately 1 X 10⁷ cpm/µg. A total of 5 X 10⁷ cpm of probe was hybridized to total RNA from normal human monocytes and dimethylformamide (DMF) induced PLB-985 cells at 50°C for 1 hour in 12 µL of 100 mmol/L KCl/10 mmol/L MgCl₂/25 mmol/L Tris HCl pH 8.5. The reverse transcription reaction was initiated in a buffer containing 30 mmol/L KCl, 8 mmol/L MgCl₂, 50 mmol/L Tris HCl (pH 8.5), 500 µmol/L dNTP, 25 µg of actinomycin D per mL, 4 U of RNasin and 50 U of avian myeloblastosis virus reverse transcriptase for 90 minutes at 42°C in a shaking bath. After ethanol precipitation, samples were resuspended in formamide buffer, boiled, and analyzed on a 6% acrylamide/8 mol/L urea gel.

Northern blot analysis. Total RNA was isolated using RNAzol (Tel-Test, Inc, Friendswood, TX) from 18 different human tissues (Clontech, Palo Alto, CA). Further characterization of the expression of hematopoietic cells was performed with total RNA extracted from normal human monocytes, neutrophils, and eosinophils. RNA was extracted from the following cell lines, two separate B lymphoblastoid cell lines (one derived from a normal individual and a patient with p47-phox—deficient CGD), three myeloid cell lines, U937, THP-1, and PLB-985, the T-cell line H9, basophil-derived cell line KU 802, mast cell-derived cell line, HMC1, the erythroleukemia cell line, K562 (treated and untreated with hemin), and DAMI (a human megakaryocyte cell line).7-11 In addition, we also analyzed RNA harvested from KU802 and HMC1 cells treated with phorbol myristate acetate (PMA) and the ionophore, A23187, for 4 hours. All RNA samples, except the eosinophils (which were studied by a RT-PCR assay), were evaluated by Northern blot analysis. A total of 20 ug per lane of total RNA was loaded onto a denaturing 1% agarose gel, electrophoresed, and transferred to N+ Hybond (Amer- sham, Arlington Heights, IL) and ultraviolet (UV) cross-linked. The RNA blots were hybridized overnight with random primed p40-phox cDNA probe and washed according to the following conditions, all at 42°C, in 2x SSC, 1x SSC, and 0.2x SSC for 20 minutes each. Afterwards, autoradiography was performed.

Chromosomal localization. Aliquots of DNA from human-rodent hybrid cell lines containing different human chromosome(s) (Coriell Institute for Medical Research, Camden, NJ) were screened by PCR, using primers F7 and R6 under the conditions described earlier. The subchromosomal map position was assessed by fluorescence in situ hybridization (FISH) analysis. Metaphase spreads were prepared from methotrexate-synchronized and 5-bromodeoxyuridine (BrdU)-treated cultures of a healthy donor. Colcemid treatment, standard protocols, P1-DNA was labeled with biotin-16-dUTP (Boehringer Mannheim) by nick-translation. A total of 100 ng of labeled DNA were precipitated in the presence of 5 µg of Cot1-DNA and 10 µg of salmon sperm DNA, resuspended in 10 µL hybridization solution (50% formamide, 2x SSC, 10% dextran sulfate), denatured for 5 minutes at 76°C, and preannealed for 30 minutes at 37°C. The probe was then added to previously denatured metaphase chromosome preparations, hybridized, and detected using avidin-FITC (Vector Laboratories, Burlingame, CA). The Brd U-induced R-banding pattern was generated through a fluorescence photolysis step, followed by staining with propidium iodide. Fluorescence signals were acquired using a cooled CCD-camera (Photometrics, Tucson, AZ), mounted on a Leica DMRBE-microscope, and visualized using Gene Join.

Subregional localization of p40-phox was performed with a PCR assay using 15 somatic cell hybrids from a panel of 26 hybrid cell lines. Primers, p40-phoxF, GCCATCGTGTGTAGACCC and p40-phoxR, GCCATCGTGTGTAGACCC, were designed using...
the program PRIMER (M.J. Daly, S. Lincoln, and E.S. Lander, Whitehead Institute, Cambridge, MA, 1992) and were used to amplify a 174-bp fragment surrounding exon 10. Conditions for the PCR reaction were as follows: 95°C for 5 minutes, 35 cycles of 94°C for 15 seconds, 62°C for 15 seconds, and 72°C for 1 minute 22 seconds, and a final 7-minute extension at 72°C. The PCR products were electrophoresed on a 1.5% agarose gel and visualized by ethidium bromide staining.

RESULTS

Isolation of genomic clones and mapping of p40-phox gene. Two overlapping recombinant Emb13A bacteriophage clones, L40-1 and L40-2, were isolated from a human genomic library by screening with a full-length cDNA probe containing the coding region of p40-phox mRNA (Fig 1). A single P1 clone, P40-1, was obtained and its analysis by Southern blot and sequencing of PCR products confirmed the map generated by the overlapping phage clones. P40-1 extends far upstream of the start of transcription. Analysis of these three clones resulted in a composite map of selected restriction sites, confirmed by genomic Southern blots (Fig 1). Exon positions were determined by sequencing of subclones either directly isolated from the phage clones or generated by PCR using unique oligonucleotide primers identified during the analysis. Southern blot analysis of normal genomic DNA digested with the enzymes shown in Fig 1 did not indicate large rearrangements or deletions (data not shown).

One bacteriophage clone, L40-2, contains a 7-kb insert that extends from intron 1 to intron 4 and thus, does not

Table 1. p40-phox Intron-Exon Boundary Sequences

<table>
<thead>
<tr>
<th>Exon</th>
<th>5' Boundary</th>
<th>Intron Length</th>
<th>3' Boundary</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>GCC GAG AG</td>
<td>gtaagttgccggggtgagccgccccggc-5.8kb-</td>
<td>acctttttccccctctctcggaacag T GAC TTT</td>
</tr>
<tr>
<td>II</td>
<td>ACC GAC TTT</td>
<td>gtaagccagactctatctctttccaaaaccc-0.8kb-</td>
<td>tccccccacactctctctctctcag GCT TCC TCC</td>
</tr>
<tr>
<td>III</td>
<td>ACA GCA CCA G</td>
<td>gttgctggccactcccgtctgtg-2.3kb-</td>
<td>aggagctctttttcggtctctcag AGC TCC TCC</td>
</tr>
<tr>
<td>IV</td>
<td>GCC TAC ATG AAG</td>
<td>gtaaggctggccactcccctgcttgg-2.8kb-</td>
<td>cggagctctttttcctgtctctcag CC AAA AAG</td>
</tr>
<tr>
<td>V</td>
<td>ACC CCG AAA GT</td>
<td>gtaagccagcccccaggtgtctcc-1.2kb-</td>
<td>aacagccccctttttctctcag G AAG AAG</td>
</tr>
<tr>
<td>VI</td>
<td>CCG AGA GCA GAG</td>
<td>gttgctggccactcccctgcttgg-2.2kb-</td>
<td>ctgctggccactcccctgcttggcag GCT CAG TCC</td>
</tr>
<tr>
<td>VII</td>
<td>GAC TGG CTK GAG</td>
<td>gtaagttccagaagttaggagagggg-2.2kb-</td>
<td>tggctgctggccactcccctgcttggcag GGC ACT TCC</td>
</tr>
<tr>
<td>VIII</td>
<td>GAC ATC AA</td>
<td>gctgtgctggccagggaggggag-0.3kb-</td>
<td>ttcctgctggccactcccctgcttggcag GGC ACT TCC</td>
</tr>
<tr>
<td>IX</td>
<td>GAG CTC ACA AG</td>
<td>gtattcggagtcgtaaagggctgtgctg-1.6kb-</td>
<td>atattcggagtcgtaaagggctgctg GGC GGC TCC</td>
</tr>
</tbody>
</table>

Exonic sequence is indicated in upper case letters and intronic sequence is indicated in lower case letters. Numbers above each exon refer to the cDNA sequence beginning with the start site of translation. The sizes of the introns are approximate and have been confirmed by both PCR analysis and Southern blot analysis.
include either the start site of transcription or translation. The 3' clone, L40-1, contains a 13-kb insert, overlapping L40-2 by 4 kb between intron 3 and 4 and extends approximately 3 kb into the 3' untranslated region.

**Exon-intron structure of coding region of p40-phox.** The structure of the coding region of the gene for p40-phox was determined by analysis of the phage clones and the P1 clone. Sizes of exons and introns were confirmed by PCR amplification of genomic DNA and fragments subcloned in the TA-vector, pCR1. The gene spans approximately 18 kb and is divided into 10 exons (Fig 1). The exon-intron boundaries were identified by direct sequencing of either bacteriophage clones or PCR amplified fragments derived from L40-1, L40-2, and P40-1 (Table 1). All splice junction sequences conform to the GT/AG rule. The complete sequencing of all 10 exons confirms a product of 339 amino acids, as predicted by the translation product of the published cDNA sequence. The sizes of polypeptide regions encoded by the individual exons ranged from 11 to 64 amino acids, with the last exon encoding for the largest span of amino acids. Of note, the SH3 region, which may play a critical role in interactions with other components of the NADPH-oxidase, is distributed between three exons, 6, 7, and 8. Introns varied in size from 300 bases (intron 8) to 5.8 kb (intron 1).

During our analysis, we determined an error in the published cDNA sequence for p40-phox in the 5' upstream region.19 The previously published sequence from -60 to -130 is incorrect with respect to the orientation of the 5' upstream sequence. Its position relative to flanking sequences could not be verified in our genomic clone or in PCR amplification of this region from normal human genomic DNA. A second cDNA clone from the original screening was obtained and analyzed. Its sequence matched the genomic P1 clone isolated, as well as material amplified from normal human genomic DNA. The sequence for the open-reading frame was identical to the previously published sequence. On further analysis, the published sequence from -60 to -128 matches in retrograde complement the correct sequence from -28 to -96. We believe this represents a cloning artifact. Figure 2B shows the corrected sequence (underlined) seen in the second cDNA clone, the genomic clone, and material amplified from normal genomic DNA.

**Chromosomal localization of p40-phox.** Screening DNA from human-rodent hybrid cell lines containing different human chromosome(s) with a p40-phox-specific PCR assay indicated that the gene maps to human chromosome 22 (data not shown). The subchromosomal mapping position of the P1-clone, P40-1, was determined by FISH to high-resolution

---

**Fig 2. (A) Identification of the start of transcription by primer extension assay.** Samples were run on 6% acrylamide/8 mol/L urea gel. The first eight lanes show a duplicate sequence analysis of the region above the start of transcription using the primer PR4 in the following orientation T, G, C, A. Two blank lanes separate the sequence analysis from the primer extension assay, which is shown in lanes 1 to 8. Lane 1, total RNA purified from normal monocyte donor no. 1; lane 2, total RNA purified from normal monocyte donor no. 2; lane 3, total RNA from DMF-induced PLB-985 (preparation no. 1); lane 4, total RNA from DMF-induced PLB-985 (preparation no. 2); lane 5, total RNA from DMF-induced PLB-985 (preparation no. 3); lane 6, TrRNA control; lane 7, RNA from normal human genomic DNA. The major site for the start of transcription is indicated by the upper arrow. The lower arrow illustrates a possible secondary site. (B) 616 base pairs of the 5' upstream sequence for p40-phox. The start of transcription is indicated by an arrow. Consensus binding sites for AP-2 and NF-E1 are underlined at bp -457 to -468 and -425 to -431 according to the start of translation. A correction of the published cDNA sequence19 for the 5' upstream region is highlighted with a dotted underline and includes the region corresponding to -60 to -130 from the start of translation. The primer, PR4, used for sequence analysis and primer extension analysis shown in (A) is indicated.
The BrdU-induced G-banding pattern allows for the chromosome and band identification. The probe hybridized to chromosomal map position 22q13. Ideogram of chromosome 22 displays schematically the G-banding pattern with the 550 band resolution. The bar besides the ideogram denotes the chromosomal mapping position of the P1 clone.

Subregional localization of p40-phox on chromosome 22 was accomplished using a subset of 15 somatic cell hybrids from a panel of 26 hybrid cell lines. The extended panel divides chromosome 22 into 25 regions or bins and has been used to map over 300 markers to unique locations on the chromosome.35 The p40-phox primer pair, p40-phoxF, and p40-phoxR (see Materials and Methods), gave a positive signal in the following cell lines: GM10888, GM11220, C1-9/5878, AJ09, RAJ5BE, CI 1-1/TW, CI 15-1/PB and no specific signal for the following cell hybrids: GM11685, CI-4/GB, Rad-37a, 514 AA2, GM11221, D655, CI-3/5878, and CI 21-2/PB (47A).36-38 These results map p40-phox below the hybrid breakpoint in D655 and above the one in CI-3/5878, which corresponds to bin 15 of the extended somatic cell hybrid panel.35 Using this PCR amplification technique, we have independently confirmed the FISH data and provide greater resolution of assignment to the region, 22q 13.1, using the breakpoints of the somatic cell hybrids.

The transcriptional initiation site, primer extension analysis was performed and identified a major start site. The start of transcription maps to purine G, located at base pair, -156, upstream of the ATG initiation of translation (Fig 2A). The band corresponding to bp 156 was seen consistently in six different experiments with RNA prepared from human normal monocytes, as well as RNA from differentiated PLB-985 cells. On two occasions, a second band was observed corresponding to the A at bp -129. All primer extension experiments were performed with the primer, PR4, internally labeled and complimentary to the 5' end of the corrected p40-phox upstream sequence shown in Fig 2B (see below). A total of 616 base pairs sequence upstream of the start of translation are reported (Fig 2B). Sequence analysis of this region shows several notable features. There is no consensus sequence for either a TATA or CAAT motif. Three hundred and one bases upstream from the start of transcription is a consensus sequence for AP2 binding and 269 bases upstream of the start of transcription is an NF-E1 consensus binding site.

**Discussion**

p40-phox is encoded by a single-copy gene that encompasses 18 kb and consists of 10 exons. Knowledge of the genomic structure will serve as the foundation for studying the regulation of p40-phox gene expression in hematopoietic cells. The localization by FISH of p40-phox to chromosome 22q has been established and the subregional localization has been accomplished using a PCR-based assay on a panel of somatic cell hybrids. Using PCR amplification of primers specific to exon 10, we have regionally mapped human p40-phox to 22q13.1. These results agree with and refine the FISH results. Unlike the p47-phox gene, the p40-phox gene appears to be a single copy.33 Other genes, which have been mapped to this region, include granulocyte-macrophage colony-stimulating factor (GM-CSF) receptor beta chain, histone H1α, hippocampal inward rectifier channel, interleukin-2 receptor beta, myoglobin, and parvalbumin (Budarf et al, submitted).

Expression of p40-phox message is restricted to cells of
hematopoietic origin, neutrophils, monocytes, eosinophils, B lymphocytes, basophils, mast cells, and megakaryocytes. The pattern of expression of p40-phox included cells of the classical white cell lineage not known to generate large quantities of superoxide, mast cells, and basophils. In our study, we were unable to demonstrate the expression of the cytosolic NADPH-oxidase components, p47-phox, and p67-phox in basophils, mast cells, and megakaryocytes (data not shown). Expression of these two cytosolic factors was limited to neutrophils, monocytes, B lymphoblastoid cells, and eosinophils, all of which generate superoxide following activity of the NADPH-oxidase. p22-phox, (data not shown) was present in all cell types tested, as has been previously reported. Like p22-phox, p40-phox is abundantly expressed in an undifferentiated human leukemia cell line, PLB-985, and levels of mRNA do not significantly increase during differentiation. In this regard, we speculate that p40-phox may participate in a function in addition to the NADPH-oxidase.

The major site for the start of transcription has been mapped to a single base pair, G at bp −156, relative to the ATG of the start of translation by primer extension. We cannot exclude the possibility of a second site at bp −129, shown in Fig 2A. However, we observed this band in only two of six experiments. Furthermore, the possible second site was not restricted to a specific source of RNA (either normal human monocytes or differentiated PLB-985 cells). The sequence surrounding the start site of transcription at −156 bp does not contain a consensus sequence for a TATA or CAAT box. Several GC rich regions are apparent. In this regard, the upstream region of p40-phox resembles p47-phox, which also lacks these DNA motifs in the immediate upstream region. Ongoing studies will hopefully characterize the cis-acting elements required for expression of p40-phox.
in hematopoietic cells, particularly in myeloid and monocytic cells.

Knowledge of the gene structure will be useful in determining if rare patients with a chronic granulomatous disease phenotype may have a mutation(s) in p40-phox gene. It will be particularly interesting to investigate patients with p67-phox deficient CGD who are reported to have decreased levels of p40-phox protein. Furthermore, the possible demonstration of a mutation in p40-phox gene would further establish the functional significance of p40-phox in the cytosolic complex of the NADPH-oxidase.

NOTE ADDED IN PROOF

p40-phox has been named neutrophil cytosolic factor 4 by the nomenclature committee of the Human Genome Database. The human symbol is NCF4.

ACKNOWLEDGMENT

The authors acknowledge the advice of Drs Lee Helman and Mary Dinauer. The authors acknowledge the gift of total RNA from the following sources, cell type RNA from Dr Pat Elwood, the HMC1 and KU802 from Dr Dean Metcalfe, eosinophils from Dr Thomas Nutman, K562 and DAM1 from Dr Volker Blank and T-cell lines from Dr Kishor Bhatia. The authors appreciate the technical assistance of Terrence McDonald and Zonglu Lui. Sequences for the intron-exon borders of p40-phox gene have been submitted to Genbank and are available under accession numbers: U50720, U50721, U50722, U50723, U50724, U50725, U50726, U50727, U50728, and U50729.

REFERENCES


Genomic structure, chromosomal localization, start of transcription, and tissue expression of the human p40-phox, a new component of the nicotinamide adenine dinucleotide phosphate-oxidase complex

S Zhan, N Vazquez, S Zhan, FB Wientjes, ML Budarf, E Schrock, T Ried, ED Green and SJ Chanock