Cloning of the cDNA and Chromosome Localization of the Gene for Human Thymidine Kinase 2*

Magnus Johansson and Anna Karlsson‡

From the Medical Nobel Institute, Department of Medical Biochemistry and Biophysics, Karolinska Institute, S-171 77 Stockholm, Sweden

Human thymidine kinase 2 (TK2) is a deoxyribonucleoside kinase that phosphorylates thymidine, deoxycytidine, and deoxyuridine. The enzyme also phosphorylates anti-viral and anti-cancer nucleoside analogs. We have identified an expressed sequence tag cDNA that encoded a 27.5-kDa protein ~30% similar to the human deoxycytidine kinase and deoxyguanosine kinase. The protein was expressed in Escherichia coli and shown to have similar substrate specificity as reported for purified native human TK2. The recombinant TK2 was shown to phosphorylate the anti-cancer nucleoside analog 2'-,2''-difluordeoxycytidine. Northern blot analysis showed two mRNA species at 2.4 and 4.0 kilobases predominately expressed in liver, pancreas, muscle, and brain. We identified a sequence-tagged site designed from the 3' region of the TK2 cDNA. The sequence-tagged site has been mapped to 81–84 centi-morgans from the top linkage group of chromosome 16, which corresponds to the 16q22 region. Our data show that deoxycytidine kinase, deoxyguanosine kinase, and TK2 belong to a family of closely related enzymes. At the time of this report all of the known human deoxyribonucleoside kinases have been cloned. This provides the opportunity to characterize their individual contribution to therapeutic and toxic effects of nucleoside analogs.

Nucleoside analogs used in anti-cancer and anti-viral therapy are administered as pro-drugs, and their pharmacological actions are dependent upon intracellular phosphorylation by deoxyribonucleoside kinases. Biochemical studies show that there are four distinct human deoxyribonucleoside kinases with different subcellular locations (1). Deoxycytidine kinase (dCK) and thymidine kinase 1 (TK1) are cytosolic enzymes whereas deoxyguanosine kinase (dGK) and thymidine kinase 2 (TK2) are considered to be located in the mitochondria. The cDNAs encoding TK1, dCK, and dGK have been cloned (2–4).

Human TK2 purified from leukemic spleen phosphorylates the pyrimidines thymidine, deoxycytidine, and deoxyuridine (5, 6). Purified TK2 also phosphorylates the nucleoside analogs 3'azido-2', 3' -deoxythymidine (AZT) and 1-(2-deoxy-2-fluoro-β-D-arabinofuranosyl)-5-iodouracil (FIAU) that are active against human immunodeficiency virus and hepatitis B virus, respectively (5–8). These nucleoside analogs are also substrates of TK1, but the relative importance of the two thymidine kinases for activation of these nucleoside analogs in different tissues is not known. Both AZT and FIAU cause adverse effects that are due to mitochondrial DNA damage (9). The suggested mitochondrial location of TK2 and the low TK1 expression in the affected tissues have implied that TK2-mediated phosphorylation may be important.

We have recently cloned human mitochondrial dGK by identifying proteins homologous to dCK encoded by expressed sequence tag (EST) cDNAs (4). We now report the cloning of an EST cDNA that encoded an enzyme with ~30% similar primary structure as compared with dCK and dGK. This enzyme had the same substrate specificity and size as described for purified native human TK2. The cloning of all four human deoxyribonucleoside kinases will make it possible to determine the relative contribution of each enzyme to the therapeutic and toxic effects of nucleoside analogs. These data are important for development and evaluation of novel anti-cancer and anti-viral nucleoside analogs to achieve optimal therapeutic index.

EXPERIMENTAL PROCEDURES

Cloning of Human TK2 cDNA—The GenBank sequence data base at the National Center of Biotechnology Information was accessed on the Internet World Wide Web. The Basic Local Alignment Search Tool (BLAST) was used to identify the EST clone-encoded proteins homologous but not identical to predicted amino acid sequences of dCK (3) and dGK (4). The EST cDNA clones were ordered from Research Genetics Inc., and the DNA sequences were determined as described below.

Two primers were designed based upon the DNA sequence of the EST clone (5'-CTCTGAATGTTGCCACGTCCACTG and 5'-CTCCTATGGGCAAATGCTCCTCGATTCTC). The primers were used to amplify the full-length cDNA from an adult liver cDNA Marathon library (CLONTECH). A PCR with the Expand PCR Long Template System (Boehringer Mannheim) was used as described in the CLONTECH manual. The PCR products were cloned into pGem-T plasmid vector (Promega). All DNA sequences were determined with the automatic laser fluorescence sequencer (Pharmacia Biotech Inc.).

Expression of Human TK2—Two PCR primers with 5' BamHI and SalI restriction enzyme sites were designed based on the cDNA sequences (5'-ATCGTGATCCATGGATGACTGGATGATGAGC and 5'-ATCGTGATCCATGGATGACTGGATGATGAGC). The PCR-amplified product was cloned into pGEX-4T-1 plasmid vector (Pharmacia) to express the cDNA-encoded protein fused to glutathione S-transferase. The expression plasmid vector was transformed into the Escherichia coli strain BL21 (DE3)pLysS (Stratagene). A transformed colony was inoculated in LB medium supplemented with 100 μg/ml ampicillin and 34 μg/ml chloramphenicol. The bacteria were grown to an OD₆₀₀ = 0.7, and

* This work was supported by grants from the Medical Faculty of the Karolinska Institute, the Knut and Alice Wallenberg Foundation, and the Biomedical Research Program and Human Capital and Mobility Program of the European Community. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed. Tel: 46-8-728 6985; Fax: 46-8-30 51 93.

1 The abbreviations used are: dCK, deoxycytidine kinase; dGK, deoxyguanosine kinase; TK1, thymidine kinase 1; TK2, thymidine kinase 2; AZT, 3’-azido-2’, 3’-deoxythymidine; dFdC, 2’, 2’-difluordeoxycytidine; FIAU, 1-(2-deoxy-2-fluoro-β-D-arabinofuranosyl)-5-iodouracil; EST, expressed sequence tag; PCR, polymerase chain reaction; kb, kilobase(s).

2 Internet address: http://ncbi.nlm.nih.gov/
The cDNA encoded a protein that was homologous but not identical to the protein expressed from dCK and dGK. The encoded protein had no ATG start codon in-frame with the translated part of the cDNA sequence. Isoform A and B differed in the 5' part. Only isoform A had an ATG translation start codon in-frame with the predicted protein sequence. Isoform B had a splice acceptor site located as indicated (12). ORF, open reading frame; Y, C or T; R, A or G.

Protein expression was induced with 0.5 mM isopropyl-1-thio-D-galactopyranoside for 4 h. The cells were harvested by centrifugation at 6000 × g for 10 min and resuspended in phosphate-buffered saline. The bacteria were lysed by sonication for 10 min on ice, and the extract was centrifuged at 10,000 × g for 10 min. The crude extract was loaded onto a glutathione-Sepharose 4B column (Pharmacia), and the recombinant protein was eluted in 100 mM Tris (pH 7.6) supplemented with 10 mM reduced glutathione (Sigma). The size and purity of the recombinant protein were determined by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (Phast system, Pharmacia).

**Enzyme Assay**—The substrate specificity of the recombinant enzyme was determined by a phosphoryl transferase assay with [γ-32P]ATP as described previously (6). The assay was performed in 50 mM Tris (pH 7.6), 10 mM dithiothreitol, 5 mM MgCl₂, 1 mM unlabeled ATP, 25 μCi of [γ-32P]ATP (3000 Ci/mmol, Amersham Life Science, Inc.), and indicated concentrations of deoxyadenosine, deoxyguanosine, thymidine, deoxyuridine, deoxynosine, 9-β-D-arabinofuranosyladenine, 9-β-D-arabinofuranosylcytosine, 9-β-D-arabinofuranosylguanine, 9-β-D-arabinofuranosylthymidine, 2',3'-didehydro-2',3'-dideoxythymidine, 2',3'-dideoxythymidine, 3'-fluoro-2',3'-dideoxythymidine, AZT, and 2-chloro-2'-deoxyadenosine were not phosphorylated by the enzyme. The substrate specificity of the recombinant enzyme is identical to that described for purified native human TK2 with the exception that the tissue-purified native enzyme has been reported to phosphorylate AZT (5, 6). The nucleoside analog dFdC has not, to our knowledge, previously been shown to be a substrate for TK2. Based on the substrate specificity, the predicted molecular size, and the sequence similarities to dCK and dGK, we conclude that the cloned cDNA encodes human TK2.

**Sequence Analysis and Comparison**—We have previously compared the sequences of dCK, dGK, and herpes simplex virus type-1 thymidine kinase (4). The comparison shows that five of the six primary structure motifs that are conserved in most herpesvirus thymidine kinases (13) match the regions that are highly conserved between dCK and dGK. Alignment of the predicted amino acid sequences of dCK, dGK, and TK2 showed that the five regions are also conserved in the TK2 sequence (Fig. 3).

TK2 is considered to be located in the mitochondria (1, 5). Most proteins that are translocated into the mitochondria have an N-terminal signal sequence with characteristic properties such as high content of basic and hydrophobic amino acid residues, few acidic residues, and an amphipathic α-helical structure (14). These features are present in the N-terminal sequence of human mitochondrial dGK (4). The N-terminal sequence of TK2, however, lacks several of these features, and we were not able to identify a mitochondrial translocation signal motif in the TK2 sequence.

**Northern Blot Analysis of TK2 mRNA**—We used multiple tissue Northern blots to determine the length and expression pattern of TK2 mRNA. Northern blot analysis of the same tissues has previously been used to determine the expression patterns of dCK and dGK mRNAs (4). The TK2 probe hybridized with two bands at ∼2.4 and ∼4.0 kb (Fig. 4). Both TK2 mRNA species were expressed in all tissues with the highest levels in liver, pancreas, muscle, and brain. Cross-hybridization with the 1.0–1.3-kb dGK mRNA or the 2.8-kb dCK mRNA was observed.

**Chromosomal Localization of the TK2 Gene**—The PCR primers that were used to map two of the sequence-tagged sites deposited in the GenBank™ database were derived from the 3' region of an EST TK2 cDNA (sequence-tagged sites clone WI-11439 and WI-15863) (15). Both sequence-tagged sites have been mapped between the gene markers D16S400 and D16S421 by PCR screening of a human-rat hybrid panel. The gene markers are located on chromosome 16 at 8455.
81–84 centimorgans from the top linkage group (16). We compared this location with the locations of other genes that have been mapped with both fluorescence in situ hybridization and PCR screening of the same radiation hybrid panel (16). The comparison showed that the position of the human TK2 gene can be assigned to chromosome 16q22 (Fig. 5).

**DISCUSSION**

We cloned and recombinantly expressed human TK2 cDNA. Thereby the cDNA sequences of all four known human deoxyribonucleoside kinases are available. The sequence data showed that human dCK, dGK, and TK2 are closely related. These three enzymes are also sequence-related to both prokaryotic and viral deoxyribonucleoside kinases (17, 18). Human TK1 cannot be included in this enzyme family based on sequence similarity as most of the substrate binding regions are different (2). The cell cycle regulation of TK1 also separates it from the other deoxyribonucleoside kinases. TK1 is strictly cell cycle-regulated and expressed only in the S-phase, whereas dCK, dGK, and TK2 are considered to be constitutively expressed throughout the cell cycle (1). The constitutively expressed enzymes are believed to supply nondividing cells with deoxyribonucleotides for DNA repair. As a result of the broad substrate specificities of dCK, dGK, and TK2, these enzymes are sufficient for the phosphorylation of all the naturally occurring purine and pyrimidine deoxyribonucleosides.

We were not able to identify a mitochondrial translocation signal in the primary structure of TK2. This does not exclude that the TK2 protein is translocated into the mitochondria, as there are other signals in addition to the typical N-terminal motif (14). There are, however, reports that TK2 activity also has been detected in the cytosol (19). One possibility is that both cytosolic and mitochondrial forms of TK2 exist, and the presently cloned TK2 cDNA encodes the cytosolic enzyme. By Northern blot analysis we have found two mRNA species with different lengths; further studies will be required to determine the differences between these TK2 mRNAs. The cloning of TK2 will allow immunocytochemistry analysis to determine the true intracellular location of TK2.

**TABLE I**

<table>
<thead>
<tr>
<th>Substrate</th>
<th>Relative phosphorylation</th>
</tr>
</thead>
<tbody>
<tr>
<td>dAdo</td>
<td>&lt;0.1</td>
</tr>
<tr>
<td>dCyd</td>
<td>0.4</td>
</tr>
<tr>
<td>dGuo</td>
<td>&gt;0.01</td>
</tr>
<tr>
<td>dThd</td>
<td>1.0</td>
</tr>
<tr>
<td>dIfo</td>
<td>=0.01</td>
</tr>
<tr>
<td>AraT</td>
<td>0.3</td>
</tr>
<tr>
<td>FLT</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>AZT</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>dFdC</td>
<td>0.09</td>
</tr>
</tbody>
</table>

81–84 centimorgans from the top linkage group (16). We compared this location with the locations of other genes that have been mapped with both fluorescence in situ hybridization and PCR screening of the same radiation hybrid panel (16). The comparison showed that the position of the human TK2 gene can be assigned to chromosome 16q22 (Fig. 5).

**FIG. 2.** cDNA and predicted amino acid sequences of isoform A. The region within the box was not present in cDNA isoform B. The translation start codon is underlined, and the possible polyclad莺encylation signal is in lowercase letters.
spectively (20, 21). The TK2 gene has been located to chromosome 16 by electrophoretic separation of the human and mouse TK2 isoenzymes in a human-mouse somatic cell hybrid panel (22). Our data confirmed the location of the TK2 gene to chromosome 16 and showed that it was located at q22. This region is shared with the genes causing Bardet-Biedl syndrome 2, Marner type of cataract, and macular dystrophy type 1 (23, 24). No phenotypic features of these diseases imply, however, that alterations of the TK2 protein or TK2 expression is involved in the pathogenesis.

The nucleoside analog FIAU is a potent inhibitor of hepatitis B virus replication (25). However, long term treatment with this drug causes liver failure, pancreatitis, myopathy, and peripheral neuropathy (26). FIAU phosphorylation is important for both the therapeutic and adverse effects of FIAU (25). Our data showed that the TK2 mRNAs were predominantly expressed in liver, pancreas, muscle, and brain. Since these organs are mainly affected by FIAU toxicity, there may be a correlation between TK2 expression levels and FIAU toxicity. The anti-human immunodeficiency virus nucleoside analog AZT also causes adverse effects due to mitochondrial DNA damage (28). Low levels of AZT phosphorylating activity have been reported for native purified TK2 (5, 6). However, we were not able to detect any AZT phosphorylating activity by the recombinant TK2. Possible explanations of this discrepancy include post-translational modifications of TK2 that are not present in the recombinant enzyme or the possibility that the tissue-purified TK2 was contaminated with a small amount of TK1. Another possibility would be that we have cloned a new, previously not identified, isoenzyme of thymidine kinase. The identical size, specificity of natural substrates, and chromosomal localization of native TK2 and the cloned enzyme, however, strongly suggest that these proteins are identical.

**Fig. 3.** Alignment of the predicted amino acid sequences of human TK2, dCK, and dGK. The black boxes indicate conserved amino acid residues. S1–S5 show five of the six regions conserved in most of the cloned herpesvirus thymidine kinases (13). Numbering of the amino acid residues is shown as TK2, dCK, and dGK.

**Fig. 4.** Multiple tissue Northern blot analysis of human TK2, dCK, and dGK mRNA expression. The TK2 probe hybridized with two mRNA species at 2.4 and 4.0 kb. Actin was used as a control to determine the relative amount of mRNA from each tissue. PBL, peripheral blood leukocytes.
several solid malignant tumors (29). The compound is phosphorylated by dCK (30). Our data showed that dFdC is also a substrate of the recombinant TK2. The pharmacological effects of dFdC phosphorylation by TK2 remain to be elucidated. The clononing of all the known human deoxyribonucleoside kinases will enable us to address the questions of the role of TK2 role in activation of FIAU, AZT, and dFdC. It will also be possible to determine the relative contribution of each deoxyribonucleoside kinase to the pharmacological effects of novel nucleoside analogs. These data are important for the development of new anti-cancer and anti-virus therapies on a rational basis.

REFERENCES

Cloning of the cDNA and Chromosome Localization of the Gene for Human Thymidine Kinase 2
Magnus Johansson and Anna Karlsson

doi: 10.1074/jbc.272.13.8454

Access the most updated version of this article at http://www.jbc.org/content/272/13/8454

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 30 references, 11 of which can be accessed free at http://www.jbc.org/content/272/13/8454.full.html#ref-list-1