Available online at www.sciencedirect.com

SCIENCE DIRECT

Biochimica et Biophysica Acta 1723 (2005) 256–264

BIOCHIMICA ET BIOPHYSICA ACTA
BBA

http://www.elsevier.com/locate/bba

Regular paper

Genes of the thymidine salvage pathway: Thymine-7-hydroxylase from a *Rhodotorula glutinis* cDNA library and *iso*-orotate decarboxylase from *Neurospora crassa*☆

Jeffrey A. Smiley*, Melisa Kundracik[1], Daniel A. Landfried[2], Vincient R. Barnes Sr.[3], Armend A. Axhemi

Department of Chemistry, Youngstown State University, Youngstown, OH 44555, United States

Received 18 November 2004; received in revised form 3 February 2005; accepted 4 February 2005
Available online 24 February 2005

Abstract

Genes for two enzymes in the thymidine salvage pathway, thymine-7-hydroxylase (THase; official name thymine dioxygenase) and *iso*-orotate decarboxylase (IDCase) have been isolated from fungal sources. THase was isolated from a *Rhodotorula glutinis* cDNA library using a degenerate oligonucleotide based on the published amino acid sequence. The coding sequence was transferred to an *Escherichia coli* expression system, from which recombinant THase activity was measured using ^{14}C-labeled thymine. The THase sequence shows an almost complete avoidance of codons ending in A or T: 95.8% GC content is present in the third position of codons. A connection between this codon bias and the role of the thymidine salvage pathway in pyrimidine metabolism is proposed. The THase sequence is similar to Group I Fe^{+2}-dependent, αKG-dependent dioxygenases. The *R. glutinis* THase gene was used to locate the probable THase genes in the sequenced genomes of *Neurospora crassa* and *Aspergillus nidulans*. The genes neighboring THase in these two genomes are similar to each other, and are similar to the mammalian 2-amino-3-carboxymuconate-6-semialdhyde decarboxylase (ACMSD), leading to their identification as IDCase genes. The *N. crassa* version was isolated by PCR of genomic DNA, and IDCase activity was measured in recombinant *E. coli* carrying this gene. A new family of decarboxylases, using similar substrates, is identified by virtue of the protein sequence similarity.

Keywords: Thymine dioxygenase; Uracil-5-carboxylate decarboxylase; Group I Fe^{+2}, αKG-dependent dioxygenase; *Rhodotorula glutinis*; *Neurospora crassa*; 2-Amino-3-carboxymuconate-6-semialdhyde decarboxylase; Codon bias; Enzyme families; Pyrimidine metabolism

1. Introduction

The thymidine salvage[4] pathway [1], apparently present in only a small number of organisms, provides a metabolic twist in the biosynthesis of pyrimidines and pyrimidine nucleotides for the organisms in which it is present (Fig. 1). Normally, pyrimidine nucleotide metabolism begins with the de novo synthesis of UMP [2], from which nucleotides

Abbreviations: THase, thymine-7-hydroxylase (thymine dioxygenase); IDCase, isoorotate decarboxylase (uracil-5-carboxylate decarboxylase); ACMSD, 2-amino-3-carboxymuconate-6-semialdhyde decarboxylase; IPTG, isopropyl-β-D-thiogalactopyranoside; 5-HMU, 5-hydroxymethyluracil

☆ The genes for *R. glutinis* THase and *N. crassa* IDCase have been deposited with GenBank, National Center for Biotechnology Information, Bethesda, MD, and have been assigned GenBank Accession Number AY622311 and GenBank Accession Number AY622310, respectively.

* Corresponding author. Tel.: +1 330 941 7103; fax: +1 330 941 1579.
E-mail address: jasmiley@cc.ysu.edu (J.A. Smiley).

[1] Student in the National Science Foundation Research Experience for Undergraduates in Chemistry Program at Youngstown State University, Summer 2003. Current Address: College of Wooster, Wooster, OH, United States.

[2] Current address: Department of Chemistry, Penn State University, University Park, PA, United States.

[3] Current address: Department of Cell Biology, Lerner Research Institute of the Cleveland Clinic, Cleveland, OH, United States.

[4] The expression "thymidine salvage" has been used elsewhere to describe the conversion of thymidine to nucleotides, for example by thymidine kinase. In this communication, we use the term to describe the conversion of thymidine to uracil (plus the ribose portion of the nucleoside), and the utilization of uracil in the total pyrimidine nucleotide pool by conversion to UMP.

0304-4165/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.bbagen.2005.02.001

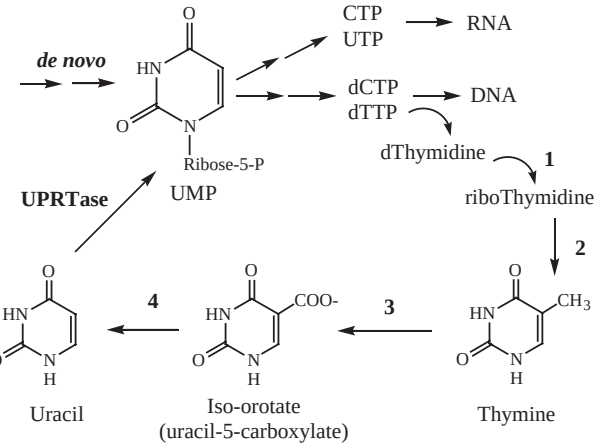

Fig. 1. The thymidine salvage pathway (numbered reactions) and its relationship with other pyrimidine metabolites. Reaction 1: Pyrimidine deoxynucleoside-2′-hydroxylase. Reaction 2: Nucleoside phosphorylase or hydrolase. Reaction 3: Thymine-7-hydroxylase. The reaction of thymine to *iso*-orotate requires three separate oxidation reactions, each requiring O_2 and α-KG and releasing CO_2 and succinate. The intermediates in the conversion of thymine to *iso*-orotate are 5-hydroxymethyluracil (5-HMU) and 5-formyluracil. Reaction 4: *Iso*-orotate decarboxylase (uracil-5-carboxylate decarboxylase). Uracil phosphoribosyltransferase (UPRTase) allows the usage of salvaged uracil for conversion to UMP and further pyrimidine nucleotide biosynthesis.

for both RNA (cytosine and uracil ribonucleotides) and DNA (cytosine and thymine deoxyribonucleotides) are eventually formed. Preformed uracil can be utilized by most organisms via its conversion to UMP in one step by the enzyme uracil phosphoribosyltransferase (UPRTase, Fig. 1). A second, two-step route is possible in many organisms, utilizing the enzymes uridine phosphorylase and uridine kinase; however, *Neurospora crassa*, one of the organisms known to possess the thymidine salvage pathway, cannot utilize uracil in the absence of UPRTase [3]. Thymine and its nucleosides and nucleotides are not usual precursors to UMP, since the pyrimidine 5-methyl group cannot normally be removed enzymatically.

Thymine-7-hydroxylase (THase; official name, thymine dioxygenase; EC 1.14.11.6) oxidizes the pyrimidine 5-methyl group to a carboxylate in three steps in which the oxidation intermediates are separate substrates [4–8]. One oxygen atom from O_2 is transferred to the substrate's 5-substituent, with the remaining oxygen atom from O_2 transferred to α-ketoglutarate (α-KG) to yield succinate and CO_2. Following three oxidation steps by THase, the conversion of thymine to uracil is completed by enzymatic decarboxylation by uracil-5-carboxylate decarboxylase (*iso*-orotate decarboxylase, IDCase; EC 4.1.1.66). Organisms with this pathway can utilize thymine or its nucleosides as a pyrimidine source in the absence of the de novo pathway; in fact, thymine can be utilized as a total nitrogen source by *Rhodotorula glutinis* [7], with much of the resulting uracil apparently not utilized for pyrimidine nucleotide metabolism.

Mechanistically, THase catalyzes the hydroxylation of a chemically unreactive methyl group, and is thus a type of enzyme receiving heightened recent interest [9]. IDCase is similarly intriguing mechanistically: The structure of the substrate suggests comparison with orotidine-5′ -monophosphate decarboxylase (ODCase) [10,11], but the altered positions of the labile carboxylates on the uracil portion of the respective substrates – C5 for IDCase and C6 for ODCase – probably necessitate different catalytic mechanisms. Regardless of any similarity with ODCase, however, IDCase catalyzes a decarboxylation that is mechanistically different from typical enzymatic decarboxylations [12]. In this study, we sought the identification of the THase and IDCase genes from fungal sources as a means of extending the resources available for mechanistic enzymological studies and metabolic studies of different fungi, particularly medically or agriculturally important fungi in which the thymidine salvage pathway may be present.

The N-terminal amino acid sequence of purified THase from *R. glutinis* [4], as well as the amino acid sequence of a proteolytic fragment covalently modified by a mechanism-based inactivator, 5-ethynyluracil [4,5], has been determined previously. Using a cDNA library constructed from mRNA of *R. glutinis* grown with thymine as the sole nitrogen source, and a degenerate oligonucleotide designed from the N-terminal amino acid sequence, we first sought the THase gene as a possible starting point for the isolation of the IDCase gene.

The *R. glutinis* THase gene should allow, by sequence comparison, the identification of the THase genes in the sequenced genomes of *N. crassa* [13] and *Aspergillus nidulans*. The location of the THase genes in these genomes might provide a link to the IDCase genes, since genes for consecutive enzymes in metabolic pathways are often located consecutively in genomes.

An unexpected clue for the identification of the IDCase gene came from the sequence identification of another decarboxylase, 2-amino-3-carboxymuconate-6-semialdhyde decarboxylase (ACMSD), from mammalian [14,15] and

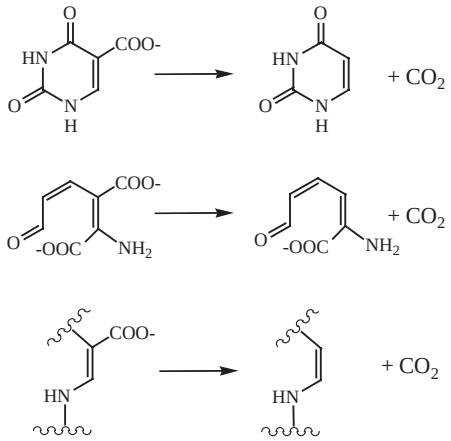

Fig. 2. Comparison of the reactions catalyzed by IDCase (top) and ACMSD (middle), with structural similarity (bottom).