Computational and Structural Biotechnology Journal 10 (2014) 41–50

COMPUTATIONAL AND STRUCTURAL BIOTECHNOLOGY JOURNAL

journal homepage: www.elsevier.com/locate/csbj

Mini Review

Variations in metabolic pathways create challenges for automated metabolic reconstructions: Examples from the tetrahydrofolate synthesis pathway

Valérie de Crécy-Lagard *

Department of Microbiology and Cell Science and Genetics Institute, University of Florida, Gainesville, FL, United States

ARTICLE INFO

Available online 11 June 2014

Keywords:
Non-orthologous displacements
Paralogs
Metabolic reconstruction

ABSTRACT

The availability of thousands of sequenced genomes has revealed the diversity of biochemical solutions to similar chemical problems. Even for molecules at the heart of metabolism, such as cofactors, the pathway enzymes first discovered in model organisms like *Escherichia coli* or *Saccharomyces cerevisiae* are often not universally conserved. Tetrahydrofolate (THF) (or its close relative tetrahydromethanopterin) is a universal and essential C_1-carrier that most microbes and plants synthesize *de novo*. The THF biosynthesis pathway and enzymes are, however, not universal and alternate solutions are found for most steps, making this pathway a challenge to annotate automatically in many genomes. Comparing THF pathway reconstructions and functional annotations of a chosen set of folate synthesis genes in specific prokaryotes revealed the strengths and weaknesses of different microbial annotation platforms. This analysis revealed that most current platforms fail in metabolic reconstruction of variant pathways. However, all the pieces are in place to quickly correct these deficiencies if the different databases were built on each other's strengths.

© 2014 de Crécy-Lagard. Published by Elsevier B.V. on behalf of the Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Contents

1. Introduction

In order to deal with the flood of data pouring from next-generation sequencing machines [1], robust and automated microbial genome annotation pipelines have become an acute necessity. The steps from gene calling to function prediction have been streamlined in annotation platforms, allowing laboratories with little bioinformatics capacity to annotate microbial genomes in a short amount of time [2–4]. Most of these pipelines base their function prediction calls on sequence similarity; however, this process is still far from perfect and high numbers of erroneous annotations remain [5–7]. Adding other types of information beyond sequence similarity, such as biological contexts by metabolic reconstruction, gene context by physical clustering, or phylogenetic conservation by co-distribution analyses, can greatly improve the quality of functional annotations [7–9]. These methods are slowly becoming part of the annotation pipelines [10,11], improving functional calls and also allowing the identification of gaps ('holes') also called "missing genes" in metabolic pathways [12,13]. Subsequent detailed comparative

* Department of Microbiology and Cell Science, University of Florida, P.O. Box 110700, Gainesville, FL 32611-0700, United States. Tel.: +1 352 392 9416; fax: +1 352 392 5922.
E-mail address: vcrecy@ufl.edu.

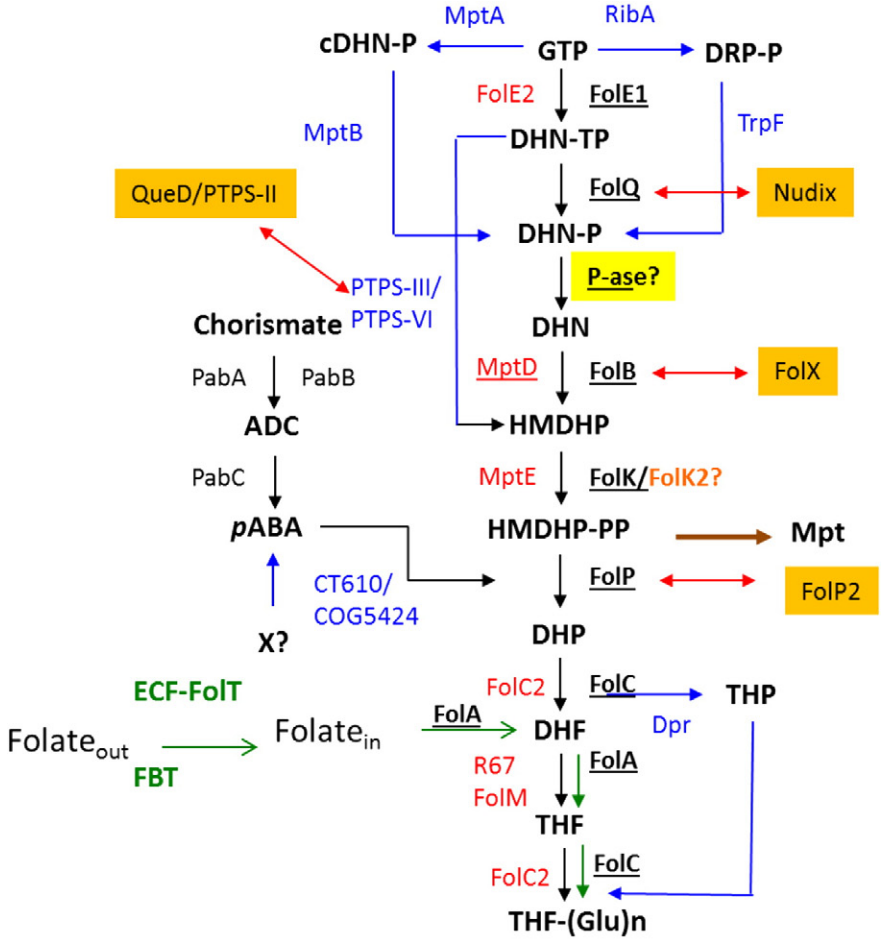
http://dx.doi.org/10.1016/j.csbj.2014.05.008
2001-0370/© 2014 de Crécy-Lagard. Published by Elsevier B.V. on behalf of the Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

genomics and experimental studies are then required to fill these pathway holes [14] because these are difficult to fill accurately using current automated gap-filling methods, even if a few success stories have been reported [15,16].

Tetrahydrofolate (THF) is a tripartite cofactor comprised of a pterin core attached to a *p*-aminobenzoate (*p*ABA) moiety and a glutamyl tail (Fig. 1). The THF synthesis pathway is complex and has been biochemically and genetically characterized extensively in *Escherichia coli* (black route in Fig. 1), with only one gene remaining to be identified (yellow highlight in Fig. 1). As the THF synthesis enzymes in yeast and *Arabidopsis thaliana* are very similar to the *E. coli* ones, this pathway was seen as an example of the uniformity of metabolism [17,18]. This view has been shattered with the advent of whole genome sequencing and the availability of thousands of genomes of diverse taxonomic origin has uncovered alternate solutions for nearly every step of the pathway. This diversity makes THF synthesis ideal to evaluate automated microbial functional annotation platforms. In this review, we provide a detailed description of all known pathway variations in THF synthesis in Bacteria and Archaea, then use these to check the annotations of the corresponding genes and the adequate calling of the THF pathway in the most common platforms used by experimentalists for gene functional annotation and pathway predictions (listed in Table 1).

1.1. Examples of non-orthologous displacements in the THF pathways

Several examples of non-orthologous enzymes catalyzing the same catalytic steps are found in the THF pathway. These can be analogous but non-homologous families, where totally different folds have been



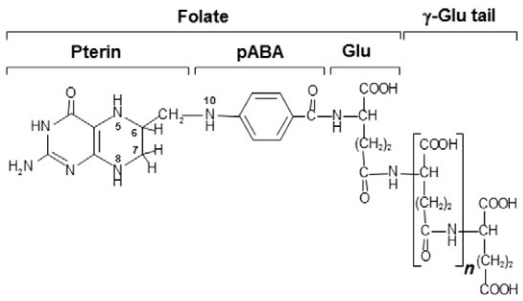


Fig. 1. Known variations and paralogs in the THF pathway. Code: underlined, canonical enzymes; red, non-orthologous displacements; blue, alternate pathways; green, salvage; yellow box, unknown gene; orange box, paralogs not in THF pathway; and orange, paralogs in folate pathway. Enzymes names are given in Table 2. Abbreviations: DHN-TP, dihydroneopterin triphosphate; DHN-MP, dihydroneopterin monophosphate; cDHNP, 7,8-dihydro-D-neopterin 2′,3′-cyclic phosphate; DRP-P, 2,5-diamino-6-ribosylamino-4(3*H*)-pyrimidinone 5′-phosphate; HMDHP, 6-hydroxymethyldihydropterin; HMDHP-PP 6-hydroxymethyldihydropterin diphosphate; *p*ABA, *p*-aminobenzoate; ADC, aminodeoxychorismate; DHP, dihydropteroate; THP, tetrahydropteroate; DHF, dihydrofolate; THF, tetrahydrofolate. THF-$(Glu)_n$, polyglutamylated THF; Mpt, methanopterin.

Download English Version:

https://daneshyari.com/en/article/2079231

Download Persian Version:

https://daneshyari.com/article/2079231

Daneshyari.com