Supplemental File pasted in below

Metabolic engineering of L-leucine production in *Escherichia coli* and *Corynebacterium glutamicum*：A review

Ying-Yu Wang [a], Jian-Zhong Xu[a, b], Wei-Guo Zhang [a, *]

[a] The Key Laboratory of Industrial Biotechnology, Ministry of Education, School of Biotechnology, Jiangnan University, 1800# Lihu Road, WuXi 214122, People's Republic of China

[b] The Key Laboratory of Carbohydrate Chemistry and Biotechnology, Ministry of Education, School of Biotechnology, Jiangnan University, 1800# Lihu Road, WuXi 214122, People's Republic of China

[*] Corresponding author: Wei-Guo Zhang; E-mail: zhangwg168@126.com;

Tel: +86-510-85329312; Fax: +86-510-85329312

Enzymes and its regulation involved in L-leucine biosynthesis

In *E. coli* and *C. glutamicum*, L-leucine is synthesized from two molecules of pyruvate in a pathway comprising seven reactions, acetohydroxyacid synthetase (AHAS), acetohydroxyacid isomeroreductase (AHAIR), dihydroxyacid dehydratase (DHAD), branched-chain amino acid transaminase (TA), α-isopropylmalate synthase (IPMS), α-isopropylmalate dehydratase (IPMD), and β-isopropylmalate dehydrogenase (IPMDH), respectively.

Acetohydroxyacid synthetase (AHAS, EC 2.2.1.6), as the first common enzyme of the pathways for the synthesis of BCAAs, catalyzes the condensation of either two molecules of pyruvate to form α-acetolactate in L-leucine and L-valine biosynthetic pathway. Also, the enzyme catalyzes the condensation of one molecule of pyruvate to one molecule of α-ketobutyrate to form α-aceto-α-hydroxybutyrate in L-isoleucine biosynthetic pathway [1]. Only one AHAS was found in *C. glutamicum* [2], which is predominantly similar to *E. coli* AHAS III (encoded by *ilvIH*) in sequence. AHAS consists of two large and two small subunits, which are encoded by *ilvB* and *ilvN*, respectively. The AHAS is responsible for feedback inhibition by all three BCAAs, but only to a maximum of 50% in the presence of 10 mmol of each BCAA. In addition, it is also competitively inhibited by α-ketoisovalerate [2-4]. In contrast with *C. glutamicum*, *E. coli* has three AHAS isoenzymes: AHAS I (encoded by *ilvBN*), AHAS II (encoded by *ilvGM*), and AHAS III [5]. Interestingly, these AHAS isoenzymes have different biochemical properties and regulatory mechanisms [6]. For example, AHAS II is not affected by L-valine, whereas the activity of AHAS I and III is strongly inhibited by L-valine. In addition, AHAS II in the *E. coli* K-12 is not expressed because the large subunit of AHAS II encoded by *ilvG* gene contains frameshift mutation [7]. Although these AHAS isoenzymes are capable of performing the same reactions, great diversity is manifested in the substrate specificities. AHAS I

prefers to use pyruvate as substrate for L-valine biosynthesis, whereas the AHAS II and III act almost specifically on α-ketobutyrate and thus lead to biosynthesize of L-isoleucine.

Acetohydroxyacid isomeroreductase (AHAIR, EC 1.1.1.86), encoded by *ilvC*, catalyzes the reductive isomerisation of α-acetolactate to form α, β-dihydroxyisovalerate in the L-valine and L-leucine biosynthesis, or the conversion of α-aceto-α-hydroxybutyrate to form α, β-dihydroxy-β-methylvalerate in the L-isoleucine biosynthesis [8]. This reaction consumes one mol of NADPH as a cofactor. The coding gene *ilvC* combines with *ilvBN* to form the *ilvBNC* operon, and its expression is negatively controlled by transcriptional attenuation. Noted that this transcriptional attenuation is mediated by all three BCAAs [9]. In addition, the expression of the *ilvBNC* operon is positively controlled by α-ketobutyrate [2]. However, as there is a positive transcriptional activator IlvY, the expression of *E. coli ilvC* is induced by acetolactate [10].

Dihydroxyacid dehydratase (DHAD, EC 4.2.1.9), encoded by *ilvD*, catalyzes the next biosynthetic step in which α-ketoisovalerate and α-keto-β-methylvalerate are formed from α, β-dihydroxyisovalerate and α, β-dihydroxy-β-methylvalerate, respectively. The α-ketoisovalerate, as the direct precursor of L-valine, is a branch point compound and it is also the common substrate for the biosynthesis of L-leucine and D-pantothenate [11]. DHAD is inhibited by either L-valine or L-leucine, whereas it is not subject to cooperative feedback inhibition by all three BCAAs. In *E. coli*, the coding gene *ilvD* combines with the *ilvGMEA* genes to form an operon (*ilvA* encodes L-threonine dehydratase, which converts L-threonine to form α-ketobutyrate), and its expression is controlled by transcriptional attenuation mediated by all three BCAAs.

Isopropylmalate synthase (IPMS, EC 4.1.3.12), encoded *by leuA*, is the real-limited enzyme of the specific pathway in the L-leucine biosynthesis. It catalyzes the formation of α-isopropylmalate from α-

ketoisovalerate and acetyl-CoA [12]. This enzyme is subjected to strong feedback inhibition by L-leucine (K_i=0.4 mM) [13]. In addition, Patek et al [13] pointed out that the addition of L-leucine to the growth medium leads to a decrease of the specific activity of enzymes in the IPM pathway [13], which indicates a negative control of the respective genes at the transcriptional level. IPMS possesses tenfold higher affinity for α-ketoisovalerate than for branched-chain amino acid transaminase (TA). Thus, it is more beneficial for L-leucine biosynthesis than for L-valine biosynthesis.

Isopropylmalate dehydratase (IPMI, EC 4.2.1.33), isomerizes the α-isopropylmalate to form β-isopropylmalate. IPMI consists of large and small subunits, which are encoded by *leuC* and *leuD*, respectively. IPMI from *E. coli* consists of a heterodimer of 50 kDa and 23 kDa subunits, whereas the proteins encoded by the homologous genes in *C. glutamicum* are 52 kDa and 22 kDa, respectively.

Isopropylmalate dehydrogenase (IPMDH, EC 1.1.1.85), encoded by *leuB*, converts β-isopropylmalate to form α-ketoisocaproate. IPMDH is strongly repressed by L-leucine in *C. glutamicum*. In addition, the gene *leuB* together with the genes *leuCD* is repressed by the leucine and tryptophan biosynthesis regulator (LtbR) [14]. The coding gene *leuB* combines with *leuACD* genes to form an operon in *E. coli* and its expression is controlled by the transcriptional attenuation mediated by L-leucine.

Branched-chain amino acid transaminase (TA, EC 2.6.1.6), which is also called transaminase B, is encoded by the *ilvE* gene in both *E. coli* and *C. glutamicum*. This enzyme transfers the amine moiety of L-glutamate to α-ketoisocaproate and thus form L-leucine by using pyridoxal 5'-phosphate (PLP) as a coenzyme [15]. In addition, this reaction is catalyzed by tyrosine-repressible transaminase (TrTA, EC 2.6.1.57), which is encoded by *tyrB* in *E. coli* [11, 16, 17]. In *E. coli*, the *tyrB* is also involved in the synthesis of the aromatic amino acids [15, 18]. As the last common enzyme of the BCAAs

biosynthesis, TA can remove the amine moiety of L-glutamate to form L-valine, L-isoleucine. Although the reaction catalyzed by TA is reversible, the balance of the reaction greatly favors the direction to BCAAs synthesis.

The main abbreviations in the manuscript

Abbreviations in Fig. 1: *Glucose-6-P*, Glucose-6-phosphate; *Glyceraldehyde-3P*, Glyceraldehyde-3-phosphate; *AHAS*, acetohydroxyacid synthetase; *AHAIR*, acetohydroxyacid isomeroreductase; *DHAD*, dihydroxyacid dehydratase; *TA*, branched-chain amino acid transaminase; *IPMS*, isopropylmalate synthase; *IPMD*, isopropylmalate dehydratase; *IPMDH*, isopropylmalate dehydrogenase; PTS, phosphotransferase system; *OAA*, oxaloacetate; *GDH*, glutamate dehydratase; *TCA*, tricarboxylic acid cycle; *PP pathway*, Pentose phosphate pathway; *Lrp*, Leucine response protein.

Reference

1. Liu Y, Li Y, Wang X: Acetohydroxyacid synthases: evolution, structure, and function. Appl Microbiol Biotechnol. 2016, 100(20):8633-8649.
2. Eggeling I, Cordes C, Eggeling L, Sahm H: Regulation of acetohydroxy acid synthase in *Corynebacterium glutamicum* during fermentation of α-ketobutyrate to L-isoleucine. Appl Microbiol Biotechnol. 1987, 25(4):346-351.
3. Elisakova V, Patek M, Holatko J, Nesvera J, Leyval D, Goergen JL, Delaunay S: Feedback-resistant acetohydroxy acid synthase increases valine production in *Corynebacterium glutamicum*. Appl Environ Microbiol. 2005, 71(1):207-213.
4. Leyval D, Uy D, Delaunay S, Goergen JL, Engasser JM: Characterisation of the enzyme activities involved in the valine biosynthetic pathway in a valine-producing strain of *Corynebacterium glutamicum*. J Biotechnol. 2003, 104(1-3):241-252.
5. Guardiola J, De Felice M, Lamberti A, Iaccarino M: The acetolactate synthase isoenzymes of *Escherichia coli* K-12. Mol Gen GeneT. 1977, 156(1):17-25.
6. Yang CR, Shapiro BE, Hung SP, Mjolsness ED, Hatfield GW: A mathematical model for the branched chain amino acid biosynthetic pathways of *Escherichia coli* K12. J Biol Chem. 2005, 280(12):11224-11232.
7. Rogerson AC, Freundlich M: Control of isoleucine, valine and leucine biosynthesis VIII. Mechanism of growth inhibition by leucine in relaxed and stringent strains of *Escherichia coli* K-12. BBA-Gen Sub. 1970, 208(1):87-98.
8. Zabalza A, Zulet A, Gil-Monreal M, Igal M, Royuela M: Branched-chain amino acid biosynthesis inhibitors: herbicide efficacy is associated with an induced carbon-nitrogen imbalance. J Plant Physiol. 2013, 170(9):814-821.

9. Morbach S, Junger C, Sahm H, Eggeling L: Attenuation control of *ilvBNC* in *Corynebacterium glutamicum*: evidence of leader peptide formation without the presence of a ribosome binding site. J Biosci Bioeng. 2000, 90(5):501-507.
10. Wek RC, Hatfield GW: Transcriptional activation at adjacent operators in the divergent-overlapping *ilvY* and *ilvC* promoters of *Escherichia coli*. J Mol Biol. 1988, 203(3):643-663.
11. Umbarger HE: Amino acid biosynthesis and its regulation. Annu Rev Biochem. 1978, 47(1):533-606.
12. Bückle-Vallant V, Krause FS, Messerschmidt S, Eikmanns BJ: Metabolic engineering of *Corynebacterium glutamicum* for 2-ketoisocaproate production. Appl Microbiol Biotechnol. 2013, 98(1):297-311.
13. Pátek M, Krumbach K, Eggeling L, Sahm H: Leucine synthesis in *Corynebacterium glutamicum*: enzyme activities, structure of *leuA*, and effect of *leuA* inactivation on lysine synthesis. Appl EnvironMicrobiol. 1994, 60(1):133-140.
14. Brune I, Jochmann N, Brinkrolf K, Huser AT, Gerstmeir R, Eikmanns BJ, Kalinowski J, Puhler A, Tauch A: The IclR-type transcriptional repressor LtbR regulates the expression of leucine and tryptophan biosynthesis genes in the amino acid producer *Corynebacterium glutamicum*. J Bacteriol. 2007, 189(7):2720-2733.
15. McHardy AC, Tauch A, Rückert C, Pühler A, Kalinowski J: Genome-based analysis of biosynthetic aminotransferase genes of *Corynebacterium glutamicum*. J Biotechnol. 2003, 104(1-3):229-240.
16. Berg C, Liu L, Vartak N, Whalen W, Wang B: The branched-chain amino acid transaminase genes and their products in *Escherichia coli*. Biosynthesis of branched-chain amino acids-1990 VCH Publishers. Weinheim, Germany 1990:131-162.
17. Yang J, Camakaris H, Pittard J: Molecular analysis of tyrosine□and phenylalanine□mediated repression of the *tyrB* promoter by the TyrR protein of *Escherichia coli*. Mol Microbiol. 2002, 45(5):1407-1419.
18. Haney SA, Platko J, Oxender D, Calvo J: Lrp, a leucine-responsive protein, regulates branched-chain amino acid transport genes in *Escherichia coli*. J Bacteriol. 1992, 174(1):108-115.