

EXPERIMENTAL  
ARTICLES

## Overproduction of Noncanonical Amino Acids by *Escherichia coli* Cells

E. V. Sycheva<sup>1</sup>, T. A. Yampol'skaya, E. S. Preobrajenskaya, A. E. Novikova,  
N. G. Matrosov, and N. V. Stoyanova

Ajinomoto-Genetika Research Institute, 1st Dorozhnyi proezd, 1, Moscow, 117545 Russia

Received March 15, 2007

**Abstract**—Overproduction of noncanonical amino acids norvaline and norleucine by *Escherichia coli* with inactivated acetohydroxy acid synthases was demonstrated. The cultivation conditions for the overproduction of noncanonical amino acids were studied. The effect of the restoration of acetohydroxy acid synthase activity, increased expression of the *leuABCD* operon, and inactivation of the biosynthetic threonine deaminase on norvaline and norleucine synthesis was studied. When grown under valine limitation, *E. coli* cells with inactivated acetohydroxy acid synthases and an elevated level of expression of the valine operon were shown to accumulate norvaline and norleucine (up to 0.8 and 4 g/l, respectively). These results confirm the existing hypothesis of norvaline and norleucine formation from 2-ketobutyrate by leucine biosynthesis enzymes.

**Key words:** noncanonical amino acids, norvaline, norleucine, 2-ketobutyrate, isopropylmalate synthase.

**DOI:** 10.1134/S0026261707060094

Unusual (or infrequently occurring) amino acids are present in a number of cellular peptides. They are often included in peptide antibiotics produced by nonribosomal synthesis [1]; in other cases, misincorporation of these amino acids into proteins by the ribosomes occurs when heterologous proteins are expressed in bacterial cells [2, 3]. Norvaline and norleucine are two of the noncanonical amino acids. Norleucine is known to be a methionine analogue [4]; it can be substituted for methionine and leucine residues in bacterial proteins [3, 5, 6]. Norvaline can be misincorporated in proteins instead of leucine [2]. Misincorporation of norleucine in the course of expression of heterologous proteins is promoted by cultivation on norleucine-containing media and under leucine or methionine starvation [7–9]. This approach was used to obtain norleucine-containing analogues of biologically active proteins with novel characteristics (interleukine-2, interferon, etc.) [5, 9–11]. For example, the substitution of methionine residues with norleucine ones makes the proteins more resistant to oxidation, methionine-specific proteases, and modifying agents [10].

Moreover, approaches are being developed which enable accurate directional inclusion of noncanonical amino acids into proteins [12]. Thus, in order to increase the structural diversity of amino acid residues in cellular proteins, over 30 unnatural amino acids were used as “building blocks” for both prokaryotic and

eukaryotic proteins [13, 14]. The progress in rational protein design promote interest in the biosynthesis of new optically active polypeptide structural units.

Biosynthesis of noncanonical amino acids by bacteria is therefore of special interest. The scheme for norvaline and norleucine biosynthesis was initially suggested by Kisumi et al. for *Serratia marcescens* [15, 16]. The authors suggested that norvaline and norleucine are formed in the leucine biosynthetic pathway from 2-ketobutyrate and 2-ketovalerate, respectively. Some experimental results indicate that *Escherichia coli*, like *S. marcescens*, uses leucine biosynthesis enzymes to synthesize noncanonical amino acids [3, 7].

The goal of the present work was to study norvaline and norleucine overproduction in an *E. coli* strain with all the isoenzymes of acetohydroxy acid synthases inactivated. Our results confirm the suggested mechanism of the biosynthesis of noncanonical amino acids via the leucine biosynthesis pathway.

### MATERIALS AND METHODS

**Bacterial strains and cultivation conditions.** In the present work, the strains *Escherichia coli* K-12 (VKPM B7), MG1655, leuA55 [17], and their derivatives were used. The Luria–Bertani broth [18] and solid media with 1.2% agarose were used for cultivation. The minimal medium composition was as follows (g/l): Na<sub>2</sub>HPO<sub>4</sub>, 3.5; KH<sub>2</sub>PO<sub>4</sub>, 1.5; NH<sub>4</sub>Cl, 1.0; glucose, 4.0; and thiamine, 0.012.

<sup>1</sup> Corresponding author; e-mail: esycheva@mail.ru

For analysis of norvaline and norleucine accumulation, the strains were grown in a fermentation medium of the following composition (g/l): glucose, 60; (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 18; KH<sub>2</sub>PO<sub>4</sub>, 2; MgSO<sub>4</sub> × 7H<sub>2</sub>O, 1; CaCO<sub>3</sub>, 25; and thiamine, 0.02. The cells were inoculated into test tubes with 2 ml of the medium and grown for 72 h at 32°C.

**Design of strain B7Δ*ilvBN*Δ*ilvGM*Δ*ilvIH* deficient in all the isoenzymes of acetohydroxy acid synthases.** Strain B7Δ*ilvBN*Δ*ilvGM*Δ*ilvIH* was obtained from strain B7 by sequential deletion of the *ilvIH*, *ilvBN*, and *ilvGM* operons by means of λRed-dependent recombination [19]. For inactivation of *ilvIH* genes, a linear DNA fragment was inserted into strain B7 chromosome, carrying the *att* sites of phage λ and a selective marker of chloramphenicol resistance, flanked by short (36 bp) sequences homologous to the 5' and 3' fragments of the *ilvIH* operon. Oligonucleotides *ilvIH*I1 (5'-ctttccaccttctcctgtttattcttattaccctgaagcctgctttttatactaaagttgg-3') and *ilvIH*I2 (5'-acatgttgggctgtaaatgctgcat-tgagatcattccgctcaagttagataaaaaagctgaac-3') were used for PCR amplification of the cassette excised by the λInt/Xis system. The pMW118-λattL-CmR-λattR plasmid [20] was used as template. Excision of the *cat* gene, flanked by the *attL* and *attR* sites of phage λ, from the chromosome of strain B7Δ*ilvIH*::*cat* was carried out according to [21] using the pMW-int/xis-ts plasmid. Deletion of the *ilvBN* genes in strain B7Δ*ilvIH* was carried out in a similar way. For this purpose, the primers *ilvBN*I1 (5'-taaacatcgctggatcgactgattacgctgcactttgagcctgctttttataactaagttgg-3') and *ilvBN*I2 (5'-tcccgaa-agctggcccagaagaaaaggactggagccgctcaagttagataaaaa-gctgaac-3') were used. In order to inactivate the *ilvBN* genes in the chromosome of strain B7Δ*ilvIH*, the primers *ilvGM*I1 (5'-gtttctcaagattcaggacggggaac-taactatgaatgaagcctgctttttataactaagttgg-3') and *ilvGM*I2 (5'-tcagctttctcgtgctcattttatattccttttgcgctcaagttagataa-aaaagctgaac-3') were used. All the chromosomal modifications in strain B7Δ*ilvBN*Δ*ilvGM*Δ*ilvIH* were confirmed by PCR.

**Inactivation of the *ilvA* and *leuA* genes, and modification of the regulatory regions of the *leuABCD* and *leuA*<sup>G479C</sup>*BCD* operons.** Inactivation of the *ilvA* gene in strain B7Δ*ilvBN*Δ*ilvGM*Δ*ilvIH* was carried out by incorporation of the *attL*λ-*cat*-*attR*λ cassette according to [19]. For inactivation of the *ilvA* gene, the *ilvA*I1 (5'-cgcggtgcgcgataaatcgaaactggggggttaatg-aagcctgctttttataactaagttgg-3') and *ilvA*I2 (5'-cctga-taagcgaagcgcctatcagcatttttcctacgctcaagttagataa-aaaagctgaac-3') primers were used. For inactivation of the *leuA* gene in strain B7, the *leuA*I1 (5'-gtgt-gtagctttgcgttgcaactcttttcgacttctgaagcctgctttttataactaagttgg-3') and *leuA*I2 (5'-tcgataccacattgctgacgggtaacag-gcgttaccgctcaagttagataaaaaagctgaac-3') primers were used.

Replacement of the regulatory region of the *leuABCD* operon in strain B7 by the constitutive P<sub>L</sub> promoter of phage lambda was carried out according to

[19]. The chromosome of strain BW25113 cat-P<sub>L</sub>-ydd [22] was used as PCR template, oligonucleotides *Leu*-*attR* (5'-aattagctaatcttacggatgcagaactcagctggcgcctcaag-ttagtataaaaaagctgaac-3') and P<sub>L</sub>-*leu* (5'-aaaataatgact-gctggcctatggttggccttagctgtttcctcttagacggccaatgct-3') were used as primers.

The regulatory region of the *leuA*<sup>G479C</sup>*BCD* operon of strain *leuA*55 was substituted by the P<sub>L</sub> promoter of phage lambda in a similar way. The strain thus obtained was used to transfer the promoter-containing cassette *attL*λ-*cat*-*attR*λ-P<sub>L</sub>-*leuA*<sup>G479C</sup>*BCD* into strain B7 by transduction with phage P1.

**Plasmids.** The plasmid pMIV-P<sub>ilvIH</sub>-*ilvIH* was obtained by cloning the DNA fragment containing the wild type *ilvIH* operon on a low copies vector pMIV-5JS (kindly provided by L. R. Ptitzyn, Ajinomoto-Genetika Research Institute). The *ilvIH* operon was obtained by PCR amplification of the chromosomal DNA of strain MG1655 with the primers *ilvIHL*58 (5'-aggagctcagagttcctgtaactccttag-3') and *ilvIHR*60 (5'-actctagatccaggttccca-3'). The PCR product flanked due to the amplification by the *Sac*I and *Xba*I restriction sites was cloned within the vector pMIV-5JS. Strain B7Δ*ilvBN*Δ*ilvGM*Δ*ilvIH* was used as the recipient for cloning.

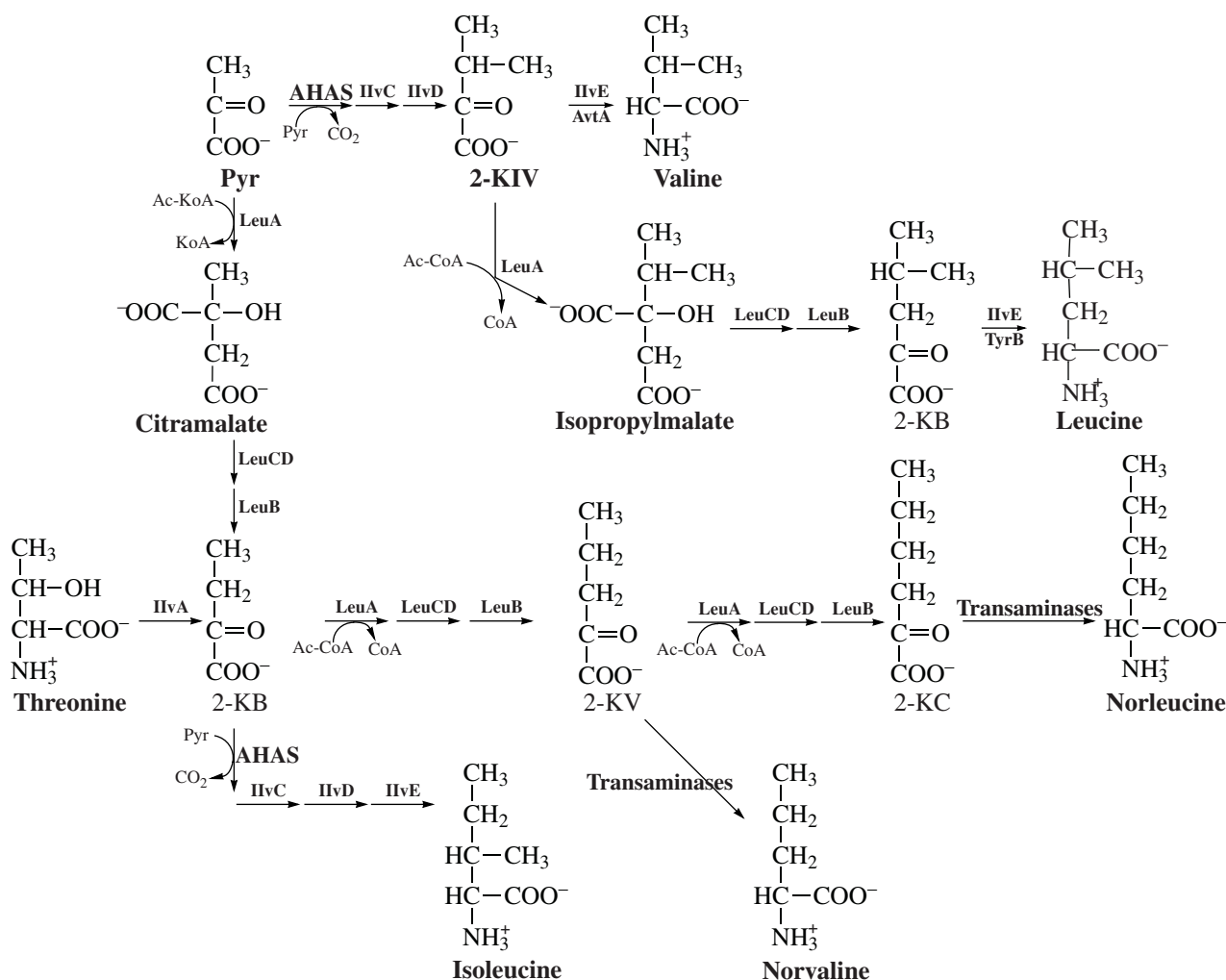
The plasmid pBR-*leuABCD* (kindly provided by E. A. Slivinskaya, Ajinomoto-Genetika Research Institute) contains on the vector the *Mun*I fragment of *E. coli* chromosome including the wild type *leuABCD* and the *leuO* gene encoding the transcription activator.

**Analytical techniques.** Identification of norvaline and norleucine in *E. coli* culture fluid was carried out by HPLC-MS, as described in [23]. The content of norleucine and norvaline in the culture fluid was determined by RF-HPLC with fluorescence detection [23].

The enzymatic activity of isopropylmalate synthase (IPMS) in cell extracts was determined as described in [24]. In order to obtain the extracts, the cells were grown to the mid-exponential phase.

Homology of the isopropylmalate synthases from *E. coli*, *Salmonella. typhimurium*, and *Mycobacterium tuberculosis* was determined using BLAST (<http://www.ncbi.nlm.nih.gov/BLAST/>) or PSI-BLAST software packages. Amino acid sequences of isopropylmalate synthases were aligned using ClustalW software (<http://www.ebi.ac.uk/clustalw/>).

**Analysis of norvaline and norleucine formation in resting *E. coli* cells.** In order to obtain resting *E. coli* cells, the culture was grown in 750-ml flasks with 50 ml of the fermentation medium supplemented with isoleucine (100 mg/l) and valine (100 mg/l). The cultures were grown at 37°C with aeration for 24 h to OD<sub>540</sub> = 10. The cells were collected by centrifugation, washed with 0.9% NaCl (10 ml), and resuspended in a medium of the following composition (g/l): (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 5; MgSO<sub>4</sub> × 7H<sub>2</sub>O, 1; KH<sub>2</sub>PO<sub>4</sub>, 2; thiamine, 0.02; CaCO<sub>3</sub>, 5. The optical density of the cell suspension was 30 U. The suspension of resting cells (3 ml) was supple-



**Fig. 1.** Biosynthesis of branched chain amino acids in *E. coli* and the tentative scheme of norvaline and norleucine in *E. coli* and *S. marcescens*. LeuA, IPMS; LeuCD, IPMI; LeuB, IPMD; IlvA, threonine deaminase; IlvC, ketoacid isomeroreductase; IlvD, dihydroxyacid dehydratase; IlvE, transaminase B; AvtA, transaminase C; TyrA, aromatic transaminase; 2-KIV, 2-ketoisovalerate; Pyr, pyruvate; 2-KB, 2-ketobutyrate; 2-KV, 2-ketovalerate; 2-KIC, 2-ketoisocaproate; 2-KC, 2-ketocaproate.

mented with 2-ketobutyrate to a final concentration of 10 mM; the substrate was not added to control samples. The cells (with and without the substrate) were grown at 37°C with aeration for 24 h; the culture fluid was then centrifuged. Amino acid content in the supernatant was determined by RF-HPLC with fluorescence detection.

## RESULTS AND DISCUSSION

**Overproduction of norvaline and norleucine by *E. coli* B7ΔilvBNΔilvGMΔilvIH** uses common metabolic pathways for the biosynthesis of branched chain amino acids, leucine, valine, and isoleucine (Fig. 1). Aceto-hydroxy acid synthases (AHAS) catalyze the first common reaction of the biosynthesis of branched chain amino acids. *E. coli* is known to possess only two iso-enzymes of aceto-hydroxy acid synthases, AHAS I (the product of the *ilvBN* genes) and AHAS III (encoded by the *ilvIH* genes). AHAS II (the product of the *ilvGM*

genes) is inactive due to a nonsense mutation in the *ilvG* gene.

In order to investigate the biosynthesis of branched chain amino acids in *E. coli*, we designed strain B7ΔilvBNΔilvGMΔilvIH; in this strain all the genes encoding AHAS isoenzymes were inactivated. Since the synthesis of branched chain amino acids was blocked, it required isoleucine and valine for growth in minimal media. Since leucine can be synthesized from 2-ketoisovalerate (the product of valine deamination) via 2-ketoisocaproate (the keto precursor of leucine) in the so-called “ketoacid chain elongation pathway,” leucine was not required for growth.

The *ilvGM* are known to be a part of the *ilvGMEDA* operon, which encodes the enzymes required for the biosynthesis of branched chain amino acids. Deletion of the *ilvGM* genes was designed in such a way that the polar effect on expression of the distant genes of the

**Table 1.** Accumulation of norvaline and norleucine by *E. coli* strains B7, B7 $\Delta$ ilvBN $\Delta$ ilvGM $\Delta$ ilvIH, and its derivatives\*

Strain	Supplements**	Norvaline, g/l	Norleucine, g/l
B7	–	<0.1	<0.1
B7 $\Delta$ ilvBN $\Delta$ ilvGM $\Delta$ ilvIH	Ile, Val	0.9	1.9
B7 $\Delta$ ilvBN $\Delta$ ilvGM $\Delta$ ilvIH	Ile, Val, Leu	<0.1	<0.1
B7 $\Delta$ ilvBN $\Delta$ ilvGM $\Delta$ ilvIH	Ile, Val***	<0.1	<0.1
B7 $\Delta$ ilvBN $\Delta$ ilvGM $\Delta$ ilvIH/pMIV-P <sub>ilvIH</sub> -ilvIH	–	<0.1	<0.1
B7 $\Delta$ ilvBN $\Delta$ ilvGM $\Delta$ ilvIH/pBR-leuABCD	Ile, Val	0.8	3.9
B7 $\Delta$ ilvBN $\Delta$ ilvGM $\Delta$ ilvIH $\Delta$ ilvA	Ile, Val	0.5	1.6

Notes: \* The cells were grown in test tubes in the fermentation medium.

\*\* Isoleucine was added to a final concentration of 100 mg/l; valine, to 100 mg/l; and leucine, to 100 mg/l.

\*\*\* Valine was added to a final concentration of 500 mg/l.

*ilvGMEDA* operon was minimized. Indeed, the growth of strain B7 $\Delta$ ilvBN $\Delta$ ilvGM $\Delta$ ilvIH on the minimal medium was restored by the introduction of the pMV-P<sub>ilvIH</sub>-ilvIH plasmid carrying the AHAS III genes; this is evidence of the *ilvEDA* expression level sufficient for growth.

We found that strain B7 $\Delta$ ilvBN $\Delta$ ilvGM $\Delta$ ilvIH, grown in a fermentation medium with isoleucine (100 mg/l) and valine (100 mg/l) in the absence of leucine, accumulated amino acids in the culture fluid. Their separation by thin-layer chromatography in the solvent system isopropyl alcohol: ethyl acetate: water: ammonia (80 : 80 : 50 : 25 vol/vol) revealed  $R_f$  values close to those of branched chain amino acids. By the method of HPLC-MS, these amino acids were identified as norvaline and norleucine [23]. The overall production of norvaline and norleucine from glucose by strain B7 $\Delta$ ilvBN $\Delta$ ilvGM $\Delta$ ilvIH was approximately 5% (Table 1).

**Confirmation of the suggested pathway of norvaline and norleucine biosynthesis in *E. coli* strain deficient in AHAS.** The mechanism of norvaline and norleucine formation has been initially proposed for *S. marcescens* [15, 16]. The authors suggested the synthesis of norvaline and norleucine from 2-ketobutyrate and 2-ketovalerate, respectively, by the enzymes performing the synthesis of leucine from 2-ketoisovalerate, isopropylmalate synthase (IMPS, product of the *leuA* gene), isopropylmalate isomerase (IPMI, product of the *leuCD* genes), and isopropylmalate dehydrogenase (IPMD, product of the *leuB* gene) (Fig. 1). In *S. marcescens*, IPMS is known to catalyze acetyl-CoA condensation not only with 2-ketoisovalerate (the precursor of leucine and valine), but also with other ketoacids, 2-ketobutyrate and 2-ketovalerate [15]. The reaction products are metabolized by IPMI and IPMD, which also have broad substrate specificity. In the first cycle of the “ketoacid chain elongation pathway”,

2-ketovalerate is formed from 2-ketobutyrate; in the second cycle of this pathway, it is converted to 2-ketocaproate. Amination of 2-ketovalerate and 2-ketocaproate by aminotransferases IlvE, AvtA, and TyrA results in formation of norvaline and norleucine, respectively.

Overproduction of norvaline and norleucine by the cells of strain B7 $\Delta$ ilvBN $\Delta$ ilvGM $\Delta$ ilvIH with inactivated AHAS when grown under valine limitation in the absence of leucine (Table 1) indicates the existence of a similar mechanism for formation of noncanonical amino acids in *E. coli*. Cultivation in the minimal medium without leucine probably results in the derepression of the *leuABCD* operon, while the absence of AHAS changes the intracellular ketoacid pools: the pool of 2-ketobutyrate increases, while the pool of 2-ketoisovalerate decreases. Under such conditions, instead of 2-ketoisovalerate, IPMS utilizes the substrate present in excess, 2-ketobutyrate.

The biochemical properties of *E. coli* IPMS have not been studied to this point; this enzyme, however, exhibits high homology with the *S. thyphimurium* IPMS (E-value <  $3 \cdot 10^{-180}$ ; identity, 92%; BLAST software), which can utilize 2-ketobutyrate and pyruvate in vitro as alternative substrates ( $K_M^{2-KIB} = 0.06$  mM;  $K_M^{2-KIV} = 1.1$  mM; and  $K_M^{pyruvate} = 10$  mM) [24].

The three-dimensional structure has been determined for a single enzyme of the isopropylmalate synthase class, IPMS of *M. tuberculosis* [25]. X-ray structure analysis revealed the amino acid residues participating in the formation of the active and regulatory sites of the enzyme [25]. Alignment of amino acid sequences of IPMS from *M. tuberculosis*, *E. coli*, and *S. thyphimurium* revealed that in spite of their low homology (E-value <  $9e-19$ ; identity, 28%; PSI-BLAST software), the amino acid residues interacting with 2-ketoacids in the substrate binding site (Arg-80, Asp-81, Glu-218, Thr-254, His-285, His-287, His-379,

<i>E. coli</i>	-----	
<i>S. typhimurium</i>	-----	
<i>M. tuberculosis</i>	MTTESPDAYTESFGAHTIVKPGPPRVGQPSWNPQRASSMPVNRYPFAEEVEPIRLRN	(60)
<i>E. coli</i>	-----MSQQVII FDTTL <b>RD</b> GEQALQASLSVKEKLIQIALALERMGVDMVEVGFVSSPG	(53)
<i>S. typhimurium</i>	-----MSQQVII FDTTL <b>RD</b> GEQALQASLSAKEKLIQIALALERMGVDMVEVGFVSSPG	(53)
<i>M. tuberculosis</i>	RTWPDRVIDRAPLWCADVLD <b>RD</b> GNQALIDPMS PARKRRMFDLLVRMGYKEIEVGFPSASQT	(120)
<i>E. coli</i>	DFESVQTIARQ---VKNSRVCALARCVEKIDIVAAESLKVAAEFRIHTFIATSPMHIATK	(110)
<i>S. typhimurium</i>	DFESVQTIART---IKNSRVCALARCVEKIDIVAAQALKVADAFRIHTFIATSPMHIATK	(110)
<i>M. tuberculosis</i>	DFDFVREIEEQGAI PDDVTTIQVLTQCRPELIERTFQACSGAPRAIVHFYNSTSI LQRRVV	(180)
<i>E. coli</i>	LRSTLDEVIERAIY---MVKRARNYTD---DVEF <b>SC</b> EDAGRTPIADLARVVEAAINAGA	(163)
<i>S. typhimurium</i>	LRSTLDEVIERAVY---MVKRARNYTD---DVEF <b>SC</b> EDAGRTPVDDLARVVEAAINAGA	(163)
<i>M. tuberculosis</i>	FRANRAEVQAIATDGARKCQEQAAYPGTQWRFEYS <b>PE</b> SYTGTELEYAKQVCDVAVGEVIA	(240)
<i>E. coli</i>	TT-----INIPD <b>TV</b> GYTMPFEFAGIISGLYERVPNIDKAIISV <b>HT</b> HDDLGLAVGNLAA	(217)
<i>S. typhimurium</i>	RT-----INIPD <b>TV</b> GYTMPFEFAGIISGLYERVPNIDKAIISV <b>HT</b> HDDLGLAVGNLAA	(217)
<i>M. tuberculosis</i>	PTPERPIIFNLPA <b>TV</b> EMTTPNVYADSI EWMSRNLANRESVILSL <b>HP</b> HNDRGTAVAAAEIG	(300)
<i>E. coli</i>	VHAGARQVEGAMNGIGERAGNC SLEEVIMAIKVRKIDILNVHTAINHQEIWRTSQVLSQIC	(277)
<i>S. typhimurium</i>	VHAGARQVEGAMNGIGERAGNCAL EEVIMAIKVRKIDIMNVHTINNHHEIWRTSQVLSQIC	(277)
<i>M. tuberculosis</i>	FAAGADRIEGCLFNGERTGNVCLVTLGLNLF <b>SR</b> ----GVDPQIDFSNIDEIRRTVEYCN	(356)
<i>E. coli</i>	NMPIPANKAIVGSGAFHSSG <b>I</b> HODGV <b>LK</b> -----NRE <b>NY</b> EIMTPE	(317)
<i>S. typhimurium</i>	NMPIPANKAIVGSGAFHSSG <b>I</b> HODGV <b>LK</b> -----NRE <b>NY</b> EIMTPE	(317)
<i>M. tuberculosis</i>	QLPVHERHPYGGDLVYTA <b>FS</b> GS <b>H</b> QDAINKGLDAMKLDADAADCDDMLWQ <b>V</b> YLPIDPR	(416)
<i>E. coli</i>	SIG-LNQIQLNLTSRSGRAAVKHRMDEMGYKESEYNLDNLYDAFLKLDKKGQVFDYDLE	(376)
<i>S. typhimurium</i>	SIG-LNQIQLNLTSRSGRAAVKHRMEEMGYKDTDYNMDHLYDAFLKLDKKGQVFDYDLE	(376)
<i>M. tuberculosis</i>	DVGRTYEAVIRVNSQSGKGGVAYIMKTDHGLSLPRRLQIE <b>F</b> SQV <b>I</b> Q <b>I</b> AE <b>G</b> TAGEGEGEVS	(476)
<i>E. coli</i>	ALAFIGKQ <b>Q</b> EE-----PEHFRLDYFSVQSGSNDIATAAVKLACGEEVKAEAAANG <b>NP</b> VD	(430)
<i>S. typhimurium</i>	ALAFIN <b>K</b> Q <b>Q</b> EE-----PEHFRLDYFSVQSGSSDIATASVKLACGEEIKAEAAANG <b>NP</b> VD	(430)
<i>M. tuberculosis</i>	PKEMWDAFAEEYLAPVRPLERIRQHVDAADDDGGT <b>S</b> ITATV <b>K</b> INGVETE <b>I</b> SGSG <b>NP</b> LA	(536)
<i>E. coli</i>	AVYQAINRIT <b>E</b> YN <b>V</b> ELV <b>K</b> YSLTAKGHGKDAL <b>G</b> QVDIVANYNGRRF <b>H</b> GVG----LATD <b>I</b> VE	(486)
<i>S. typhimurium</i>	AIYQAINRIT <b>G</b> YD <b>V</b> ELV <b>K</b> YDLNAKGQGDAL <b>G</b> QVDIVNVHHGRRF <b>H</b> GVG----LATD <b>I</b> VE	(486)
<i>M. tuberculosis</i>	AFVHALADVG-FDVA <b>V</b> LD <b>Y</b> EH <b>A</b> MSAGDD <b>A</b> Q <b>A</b> AYVEAS <b>V</b> TIASPAQ <b>P</b> GEAGR <b>H</b> ASDP <b>V</b> T	(595)
<i>E. coli</i>	SSAKAMVHVLNNIWRAAEVEKELQRKAQHNNEN <b>K</b> ETV-----	(523)
<i>S. typhimurium</i>	SSAKAMVHVLNNIWRAAEVEKELQRKAQ <b>K</b> EN <b>N</b> KETV-----	(523)
<i>M. tuberculosis</i>	IASPAQ <b>P</b> GEAGR <b>H</b> ASDP <b>V</b> T <b>S</b> KT <b>V</b> WG <b>V</b> GI <b>A</b> PS <b>I</b> TT <b>A</b> SLRA <b>V</b> VS <b>A</b> V <b>N</b> RA <b>A</b> R	(644)

**Fig. 2.** Alignment of amino acid sequences of *E. coli*, *S. typhimurium*, and *M. tuberculosis* IPMS. Amino acid residues participating in the active site are shown in bold and highlighted. Amino acid residues of the regulatory site are highlighted. The numbering of amino acid residues is according to the *M. tuberculosis* sequence.

and Tyr-410) are identical in these organisms (Fig. 2). However, the amino acid residues of the regulatory site responsible for the allosteric inhibition with leucine (Leu-535, Ala-536, Val-551, Tyr-554, Ala-558, Ala-565, and Ala-567) do not exhibit high conservatism (Fig. 2). Moreover, the conservative amino acid residues of the regulatory domain of *M. tuberculosis* IPMS (Tyr-554, Ala-558, and Ala-565) (Fig. 2) interact with the aliphatic groups of leucine [25]. This structure of the regulatory center of IPMS ensures enzyme specificity to leucine inhibition and the inability of the structurally similar amino acids, norvaline and norleucine, to interact with IPMS and inhibit it.

Interestingly, neither the amino acid residues of the substrate binding site nor the metal ion associated with the enzyme interact with the aliphatic groups of 2-ketoisovalerate [25]. This may be the explanation for the broad substrate specificity of IPMS. Thus, considering the conservative structure of the substrate binding site in isopropylmalate synthases, it can be suggested that *E. coli* IPMS, similar to *S. typhimurium* IPMS,

has a broad substrate specificity and can use 2-ketobutyrate instead of 2-ketoisovalerate as a methyl group acceptor.

In order to obtain additional proof of the mechanism of the biosynthesis of noncanonical amino acids in *E. coli*, we have studied the effect of various factors on norleucine and norvaline accumulation by strain B7 $\Delta$ ilvBN $\Delta$ ilvGM $\Delta$ ilvIH (Table 1). Addition of leucine into the medium containing isoleucine and valine was found to prevent formation of the noncanonical amino acids due to the IPMS inhibition and repression of the *leuABCD* operon. An excess of valine in the medium (500 mg/l) had a similar effect; valine can be deaminated with formation of 2-ketoisovalerate, thus preventing the interaction of 2-ketobutyrate with isopropylmalate synthase. Introduction of pMIV-P<sub>ilvIH</sub>-ilvIH (restoring the AHAS III activity) into strain B7 $\Delta$ ilvBN $\Delta$ ilvGM $\Delta$ ilvIH also resulted in suppression of norleucine and norvaline synthesis. Increased expression of the *leuABCD* operon achieved by the introduction of the pBR-*leuABCD* plasmid into strain

**Table 2.** Formation of norvaline (Nva) and norleucine (Nle) by resting *E. coli* cells with different levels of IPMS expression\*

Strain	IPMS, specific activity, nmol/(min mg)	Substrate**	Nva, $\mu$ M	Nle, $\mu$ M
B7	41	–	<10	<10
		2-KB	43	<10
B7 cat-P <sub>L</sub> -leuABCD	680	–	13	24
		2-KB	181	18
B7 cat-P <sub>L</sub> -leuA <sup>G479C</sup> BCD	534	–	80	49
		2-KB	551	34
B7 $\Delta$ leuA	0	–	<10	<10
		2-KB	<10	<10

Notes: \* As determined by RF-HPLC.

\*\* 2-ketobutyrate (2-KB) was added to a final concentration of 10 mM.

B7 $\Delta$ ilvBN $\Delta$ ilvGM $\Delta$ ilvIH resulted in a twofold increase in norleucine production. In this case, the total production of norvaline and norleucine from glucose was approx. 8%. Thus, our results confirm the proposed scheme of the biosynthesis of noncanonical amino acids.

Interestingly, inactivation of the *ilvA* gene, which blocked 2-ketobutyrate formation from threonine, did not prevent norvaline and norleucine synthesis by strain B7 $\Delta$ ilvBN $\Delta$ ilvGM $\Delta$ ilvIH (Table 1). This is an indication that, in the absence of functionally active threonine deaminase, 2-ketobutyrate may be synthesized by leucine biosynthesis enzymes from pyruvate, rather than threonine (Fig. 1). This alternative pathway of 2-ketobutyrate formation from pyruvate is used by some organisms for threonine-independent isoleucine biosynthesis [26, 27].

**Formation of norvaline and norleucine by the resting *E. coli* cells with different levels of expression of leucine biosynthesis genes.** For additional confirmation of norvaline and norleucine formation from 2-ketobutyrate by leucine biosynthesis enzymes, formation of these amino acids from 2-ketobutyrate was studied in the resting cells of strains with different levels of expression of leucine biosynthesis genes: B7 with the wild type *leuABCD* operon; B7 cat-P<sub>L</sub>-leuABCD, containing *leuABCD* genes under control of phage lambda P<sub>L</sub> promoter (expression of *leuABCD* genes is increased); B7 cat-P<sub>L</sub>-leuA<sup>G479C</sup>BCD, in which gene *leuA*, encoding feedback-resistant IPMS, as well as the other genes of the leucine operon are under control of the P<sub>L</sub> promoter; and B7 $\Delta$ leuA, in which *leuA* gene is inactivated. The level of expression of the *leuABCD* operon was determined from isopropylmalate synthase activity (Table 2).

The results of these experiments have demonstrated that accumulation of norvaline and norleucine in resting *E. coli* cells depended on the level of expression of leucine biosynthesis enzymes (Table 2). Strain B7 cat-

P<sub>L</sub>-leuA<sup>G479C</sup>BCD exhibited the highest level of norvaline and norleucine synthesis. Strain B7 $\Delta$ leuA with the deleted *leuA* gene did not synthesize norvaline and norleucine. The cells of strains B7 cat-P<sub>L</sub>-leuABCD and B7 cat-P<sub>L</sub>-leuA<sup>G479C</sup>BCD accumulated norvaline and norleucine even in the absence of exogenous 2-ketobutyrate; this is possibly due to their synthesis from the intracellular 2-ketobutyrate.

It should be mentioned that, unlike strain B7 $\Delta$ ilvBN $\Delta$ ilvGM $\Delta$ ilvIH, *E. coli* cells with functionally active AHASs grown without exogenous branched chain amino acids preferably synthesize norvaline rather than norleucine (Table 2). This is possibly related to the differences in the level of expression of aminotransferases able to aminate the keto-precursors of norvaline and norleucine.

Thus, the results obtained indicate that norvaline and norleucine in *E. coli* are formed from 2-ketobutyrate by the enzymes of the leucine biosynthesis pathway; *E. coli* cells with the blocked activity of acetohydroxyacid synthases are capable of overproduction of these noncanonical amino acids.

## REFERENCES

1. Katz, E. and Demain, A.L., The Peptide Antibiotics of *Bacillus*: Chemistry, Biogenesis and Possible Functions, *Bacteriol. Rev.*, 1977, vol. 41, pp. 449–474.
2. Apostol, I., Levine, J., Lippincott, J., Leach, J., Hess, E., Glascock, C.B., Weickert, M.J., and Blackmore, R., Incorporation of Norvaline at Leucine Positions in Recombinant Human Hemoglobin Expressed in *Escherichia coli*, *J. Biol. Chem.*, 1997, vol. 272, pp. 28980–28988.
3. Bogosian, G., Violand, B.N., Dorward-King, E.J., Workman, W.E., Jung, P.E., and Kane, J.F., Biosynthesis and Incorporation Into Protein of Norleucine by *Escherichia coli*, *J. Biol. Chem.*, 1989, vol. 264, pp. 531–539.

4. Lawrence, D.A., Regulation of the Methionine Feedback-Sensitive Enzyme in Mutants of *Salmonella typhimurium*, *J. Bacteriol.*, 1972, vol. 109, pp. 8–11.
5. Cirino, P.C., Tang, Y., Takahashi, K., Tirrell, D.A., and Arnold, F.H., Global Incorporation of Norleucine in Place of Methionine in Cytochrome P450 BM-3 Heme Domain Increases Peroxygenase Activity, *Biotechnol. Bioeng.*, 2003, vol. 83, pp. 729–734.
6. Yuan, T. and Vogel, H.J., Substitution of the Methionine Residues of Calmodulin with the Unnatural Amino Acid Analogs Ethionine and Norleucine: Biochemical and Spectroscopic Studies, *Protein Sci.*, 1999, vol. 8, pp. 113–121.
7. Tsai, L.B., Lu, H.S., Kenney, W.C., Curless, C.C., Klein, M.L., Lai, P.H., Fenton, D.M., Altrock, B.W., and Man, M.B., Control of Misincorporation of de novo Synthesized Norleucine into Recombinant Interleukin-2 in *E. coli*, *Biochem. Biophys. Res. Commun.*, 1988, vol. 156, pp. 733–739.
8. Miyazawa, T., Yokoyama, S., and Miyake, T., Method for Producing Protein Containing Nonprotein Amino Acids, *US Patent* 4, 879,223, 1989.
9. Fenton, D., Lai, P.-H., Lu, H., Mann, M., and Tsai, L., Control of Norleucine Incorporation Into Recombinant Proteins, *US Patent* 5,599,690, 1997.
10. Gilles, A.M., Marliere, P., Rose, T., Sarfati, R., Longin, R., Meier, A., Femandjian, S., Monnot, M., Cohen, G.N., and Barzu, O., Conservative Replacement of Methionine by Norleucine in *Escherichia coli* Adenylate Kinase, *J. Biol. Chem.*, 1988, vol. 263 P, pp. 8204–8209.
11. Brunner, D.P., Harbour, G.C., Kirschner, R.J., Pinner, J.F., and Garlick, R.L., Fermentation Method for Producing Norleucine, *US Patent* 5,622,845, 1997.
12. Bock, A., Invading the Genetic Code, *Science*, 2001, vol. 292, pp. 453–454.
13. Xie, J. and Schultz, P.G., An Expanding Genetic Code, *Methods*, 2005, vol. 36, pp. 227–238.
14. Zang, Z., Alfonta, L., Tian, F., Bursulaya, B., Uryu, S., King, D.S., and Schultz, P.G., Selective Incorporation of 5-Hydroxytryptophan into Proteins in Mammalian Cells, *Proc. Natl. Acad. Sci. USA*, 2004, vol. 101, pp. 8882–8887.
15. Kisumi, M., Sugiura, M., and Chibata, I., Biosynthesis of Norvaline, Norleucine, and Homoisoleucine in *Serratia marcescens*, *J. Biochem.*, 1976, vol. 80, pp. 333–339.
16. Kisumi, M., Sugiura, M., and Chibata, I., Norleucine Accumulation by a Norleucine-Resistant Mutant of *Serratia marcescens*, *Appl. Environ. Microbiol.*, 1977, vol. 34, pp. 135–138.
17. Gusjatiner, M.M., Lunts, M.G., Kozlov, Ju.I., Ivanovskaya, L.V., and Voroshilova, E.B., *Mutant Alpha-Isopropylmalate Synthase (IPMS), DNA Encoding Mutant IPMS, Method of Preparing Escherichia coli Strain, Method Of L-leucine Preparing*, US Patent 6,403,342. 2002.
18. Sambrook, J., Fritsch, E.F., and Maniatis, T., *Molecular Cloning: a Laboratory Manual. 2nd Ed*, New York: Cold Spring Harbor, 1989.
19. Datsenko, K.A. and Wanner, B.L., One-Step Inactivation of Chromosomal Genes in *Escherichia coli* K-12 Using PCR Products, *Proc. Natl. Acad. Sci. USA*, 2000, vol. 97, pp. 6640–6645.
20. Katashkina, Zh.I., Skorokhodova, A.Yu., Zimenkov, D.V., Gulevich, A.Yu., Minaeva, N.I., Doroshenko, V.G., Biryukova, I.V., and Mashko, S.V., Tuning the Expression Level of a Gene Located on a Bacterial Chromosome, *Mol. Biol.*, 2005, vol. 39, no. 5, pp. 823–831 [*Mol. Biol. (Engl. Transl.)*, vol. 39, no. 5, pp. 719–726].
21. Van Dein, S., Iwatani, S., Usuda, Y., Matsui, K., Nakai, Y., Suzuki, T., and Moriya, M., *Method for Producing L-lysine or L-threonine*, US Patent Application, 2006/0154344.2005.
22. Livshits, V.A., Vitushkiva, M.V., Mashko, S.V., Doroshenko, V.G., Biryukova, I.V., Katashkina, Zh.I., Skorokhodova, A.Yu., and Belareva, A.V., *Process for Producing L-amino acid using Escherichia*, EP1449918A1.2004
23. Novikova, A.E., Stoyanova, N.V., Sycheva, E.V., and Kolokolova, A.V., Identification and Analysis of Non-proteinogeneous Amino Acids Norvaline and Norleucine in the Culture Fluids of *Escherichia coli* Strains, *Sorbtsionnye I Khromatograficheskie Protsessy*, 2006, vol. 6, no. 5, pp. 796–806.
24. Kohlhaw, G., Leary, T.R., and Umbarger, H.E.,  $\alpha$ -Isopropylmalate Synthase from *Salmonella typhimurium*. Purification and Properties, *J. Biol. Chem.*, 1969, vol. 244, pp. 2218–2225.
25. Koon, N., Squire, C.J., and Baker, E., Crystal Structure of LeuA from *Mycobacterium tuberculosis*, a Key Enzyme in Leucine Biosynthesis, *Proc. Natl. Acad. Sci. USA*, 2004, vol. 10, pp. 8295–8300.
26. Xu, H., Zhang, Y., Guo, X., Staempfli, A.A., Chiao, J., Jiang, W., and Zhao, G., Isoleucine Biosynthesis in *Leptospira interrogans* Serotype *Lai* Strain 56601 Proceeds Via a Threonine-Independent Pathway, *J. Bacteriol.*, 2004, vol. 186, pp. 5400–5409.
27. Howell, D.M., Xu, H., and White, R.H., (R)-Citramalate Synthase in Methanogenic Archaea, *J. Bacteriol.*, 1999, vol. 181, pp. 331–333.