Isoleucine binding and regulation of *Escherichia coli* and *Staphylococcus aureus* threonine dehydratase (IlvA)

Mi-Kyung Yun[1], Chitra Subramanian[1,*], Karen Miller[1], Pamela Jackson[1], Christopher D. Radka[1,2], and Charles O. Rock[1,†]

[1]Department of Host-Microbe Interactions, St. Jude Children's Research Hospital, Memphis, Tennessee 38105

[2]Department of Microbiology, Immunology, and Molecular Genetics, University of Kentucky, Lexington, Kentucky 40536

* Correspondence to: Chitra Subramanian, Email: chitra.subramanian@stjude.org

† Deceased

Running Title: Isoleucine regulation of IlvA

Abstract

Threonine deaminase (IlvA) is the first enzyme in the isoleucine branch of the branched-chain amino acid biosynthetic pathway. Kinetic studies of isoleucine feedback inhibition of *Escherichia coli* IlvA (EcIlvA) introduced the concept of allosteric regulation in enzymes. While the crystal structure of EcIlvA shows a tetrameric assembly of protomers with distinct catalytic and regulatory domains, the isoleucine binding site within the regulatory domain (EcIlvA_R) and its regulatory mechanism remained undefined. Here, we present the high-resolution crystal structure of the EcIlvA_R•Ile complex, identifying the isoleucine binding site. The EcIlvA_R•Ile complex reveals that isoleucine binding induces a conformational change in Phe352, propagating a 23 Å shift down the regulatory domain. We propose this conformational shift forms an allosteric pathway that extends to the adjacent protomer's active site, mediating allosteric regulation across the protomer-protomer interface. Supporting this, the EcIlvA(F352A) mutant, though binding isoleucine with high affinity, is not inhibited due to the absence of the Phe352 sidechain. Additionally, *Staphylococcus aureus* IlvA (SaIlvA) is not feedback-regulated by isoleucine. Structural analysis of the SaIlvA regulatory domain reveals a different subdomain organization, preventing isoleucine binding. These findings enhance our understanding of IlvA's allosteric regulation and offer opportunities for engineering feedback-resistant IlvA variants for biotechnological applications.

Introduction

Threonine dehydratase (IlvA; also known as threonine deaminase; EC 4.3.1.19) is the first step in the isoleucine arm of the branched-chain amino acid (BCAA) biosynthetic pathway (Fig. 1) (1, 2). The pyridoxal 5′-phosphate (PLP)-dependent reaction converts threonine to 2-iminobutyrate that is then converted by RidA (3, 4) to 2-ketobutyrate to continue isoleucine biosynthesis (1, 2). The BCAA pathway is regulated by IlvA and was one of the earliest examples of the first enzyme in a biosynthetic pathway being feedback regulated by the end-product, in this case isoleucine (Fig. 1) (5, 6). The kinetics of purified IlvA led Changeux to suggest that threonine and isoleucine bind to different sites on the enzyme and that the binding of isoleucine decreases the binding of threonine (7). The ideas developed to explain IlvA kinetics led to the development of the two-state model that has become the blueprint for understanding the allosteric regulation of enzyme activity throughout biology (8). Further research established that isoleucine is a potent IlvA inhibitor, whereas valine is not (9-12), but there is some disagreement as to whether there are two or four isoleucine binding sites in the *E. coli* IlvA (EcIlvA) (10, 13, 14). The 2.8 Å crystal structure of EcIlvA (PDB ID: 1TDJ) revealed that it is a tetramer, consisting of a dimer of carboxy-terminal regulatory domains projecting outward from a core formed by the tetramerization of the PLP-containing amino-terminal catalytic domains (15, 16). This EcIlvA structure appears to be the active (relaxed or R state) allosteric conformation of the assembly with the PLP cofactor covalently attached to Lys62 in the catalytic domain and isoleucine absent from the regulatory domain.

Each EcIlvA carboxyl-terminal regulatory domain comprises two ACT-like subdomains (15, 16). The ACT domain is a structural motif of 70-80 amino acids that is present in many regulatory enzymes involved in transcriptional control, amino acid and nucleotide

metabolism (17-19). The prototypical ACT domain is composed of four β sheets and two α helices arranged in a βαββαβ fold. The first crystal structure containing an ACT domain was *E. coli* D-3-phosphoglycerate dehydrogenase (PGDH), a tetrameric protein with one ACT domain per subunit (20). The ACT domains form a symmetrical dimer, and two serine regulatory ligands were observed at the dimer interface. It has now been established that ACT domains can be combined in different ways to transmit allosteric regulatory signals through the binding of ligands at ACT domain interfaces (17-19). The isoleucine regulatory site in the EcIlvA regulatory domain has not been structurally characterized but has been inferred from molecular docking simulation studies using the crystal structure of the active EcIlvA coupled with kinetic analyses of site-directed mutations. These studies identified residue changes that result in resistance of IlvA to isoleucine inhibition, and these were interpreted as defining the isoleucine binding site (14, 21, 22).

Defining the structural transitions that switch EcIlvA from an active (R state) to an inactive (T state) conformation are key to understanding the mechanism of this important regulatory protein. Specifically, the identification of the isoleucine binding site and determination of how isoleucine binding triggers a conformational change that alters the EcIlvA active site environment remain key unanswered questions. Here, we report the structure of the isolated EcIlvA regulatory domain with isoleucine bound at a single site at the interface between the two ACT subdomains. The conformation of the isoleucine-bound regulatory domain is distinct from that of the isoleucine-free regulatory domain observed in the intact EcIlvA structure. An allosteric pathway transmits this information to the active site of an adjacent protomer in the tetrameric assembly.

BCAAs are also essential in *Staphylococcus aureus*, not only for protein synthesis but also for generating precursors for fatty acid synthesis (Fig. 1) (24). More specifically,

isoleucine has a key role in *S. aureus* in controlling fatty acid composition and membrane stability/fluidity (24-27). Whether *S. aureus* IlvA (SaIlvA) is regulated by isoleucine like EcIlvA is unknown, and a comparison of the enzymes from the two organisms is important to further understand the regulatory panorama of this key enzyme. Therefore, we also investigated the regulation of SaIlvA and showed that it is refractory to isoleucine regulation. We discovered the structural basis for this observation by comparing the EcIlvA and SaIlvA regulatory domains, which revealed key differences in the ACT domain organization that lead to the disruption of the isoleucine binding site in SaIlvA. These conclusions are further supported by complementation and biochemical studies using an IlvA knockout strain of *S. aureus*.

Results

Crystal structure of E. coli IlvA

To facilitate the analysis of the EcIlvA allosteric regulatory mechanism, we first attempted to obtain a higher resolution crystal structure of the enzyme than the current 2.8 Å structure (PDB ID: 1TDJ) (15). His-tagged EcIlvA (Fig. S1*A*) was purified by affinity chromatography followed by gel filtration chromatography (Fig. S1*B*), and we obtained crystals in space group I222. The hexa-His tag was not removed, and His tagged protein was used for all the assays. The structure was determined at 1.87 Å resolution, and the asymmetric unit contains one EcIlvA monomer that is composed of two globular domains (Fig. 2*A* and Table S1). The amino terminal catalytic domain comprises residues 1-320 (Fig. S2*A*) and contains the enzyme active site with PLP covalently bound to Lys62 (Figs. 2*A* and S2*B*), and the carboxy terminal regulatory domain comprises residues 335-514 (Figs. 2*A* and S2*B*). Residues 482-496 in the regulatory domain are disordered and their

conformation could not be determined (Fig. S2*B*). The independently-folded catalytic and regulatory domains are linked by helix H14 that comprises residues 321-334 (Fig. S2*A*), which is referred to as the 'neck' (15) (Figs. 2*A* and S2*B*). Overall, the higher resolution EcIlvA structure superimposed well with the published structure (PDB ID: 1TDJ) with the root mean square deviation (RMSD) of 1.4 Å over 493 Cα atoms. Comparing this higher resolution EcIlvA structure with the published structure, we found that four regions had different conformations (Fig. 2*A*). Inspection of the published structure with the electron density maps reveals that minor corrections are needed in the misfitted regions (regions 1 to 3) of the published structure (Fig. S3). Region 1 is close to the catalytic site and region 3 (loop L22) is involved in the conformational coupling of the regulatory and catalytic domains (see below). The corrected models of these regions exhibit similar conformations to the higher resolution structure (Fig. S3). The conformational difference in region 4 is related to the binding of glycerol. Consequently, the higher resolution glycerol-bound EcIlvA structure provides higher quality structural information.

The EcIlvA tetramer is organized as a dimer of dimers, and the protomers are identical in our structure because the tetramer is generated through crystallographic symmetry operations (Fig. S2*C*). Gel filtration chromatography (Fig. S1*C*) and prior work (15) are consistent with the observed tetrameric organization of the EcIlvA protomers. The amino terminal catalytic domains create the core tetramer with PLP bound to Lys62 in all four protomers (Fig. S2*C*), and the carboxy terminal regulatory domains form dimers that protrude from either side of the core. To validate the crystal structure organization, we used size-exclusion chromatography coupled with small angle X-ray scattering (SEC-SAXS) analysis to compare the symmetry generated EcIlvA tetramer with the volume occupied by the EcIlvA tetramer in solution (Fig. S4). SAXS experiments showed that the

sample was globular and monodisperse tetramer with a molecular weight of 239.6 kDa (Figs. S4*A*, S4*D*, and Table S2). A comparison of the EcIlvA SAXS experimental data set with the predicted profile of the EcIlvA tetramer based on the crystal structure (Fig. S2*C*) shows that the SAXS data are a very close fit to the theoretical behavior of the EcIlvA tetramer (χ^2 = 1.40) (Fig. S4*G*). The EcIlvA tetramer fits very well into the aligned *ab initio* electron density maps generated from the solution scattering data, with the correlation coefficient of 0.92. This shows a close agreement between the EcIlvA crystal and solution structures (Fig. 2*B*). SAXS data indicates that the crystallographic organization is present in solution.

The regulatory domain contains a bound glycerol in our structure that was not present in PDBID: 1TDJ (Figs. 2*A* and S2*B*). This difference in the two structures likely arises from the use of glycerol as a cryoprotectant in our structure, whereas the prior analysis did not use frozen crystals (15, 16). The bound glycerol appears to elicit a conformational change in the regulatory domain. The glycerol and two structured waters are bound by helix H15, loops L14 and L21 that cluster at the tip of the regulatory domain furthest from the catalytic domain (Fig. 2*C*). The 2Fo-Fc electron density map clearly shows the bound glycerol and the two structured waters (Fig. S5*A*). The hydroxyl groups of glycerol form hydrogen-bonding interactions with the main chain carbonyl oxygen atom of Ile460 and amide nitrogen atoms of Gly350 and Ile460, with the amide nitrogen atom of Leu353 and side chains of Glu347 and Tyr369 via water molecules (Fig. 2*D*). The 3 carbon atoms of glycerol engage in a hydrophobic pocket composed of residues Phe352 and Ile460. This glycerol binding pocket is significant because it corresponds to the isoleucine binding site as discussed below. In PDB ID: 1TDJ, residues Glu347 and Asn459 are pointing toward the solvent, and a water molecule interacts with the main chain carbonyl oxygen atoms of

residue Glu348 on loop L14 and Ile460 on loop L21 (Fig. 2*E*). Comparing the regulatory domains of the glycerol-free and glycerol-bound structures, there are structural differences in Phe352 and Loop L14, but the conformation of Phe367 and its adjacent residues remain unchanged (Fig. 2*C*). The dose-response curve indicates that isoleucine effectively inhibits EcIlvA activity, whereas glycerol, despite binding to the isoleucine binding site of regulatory domain, shows negligible inhibition activity, maintaining an activity level of 93.4% even at a high concentration of 80 mM (Fig. 2*F*). Therefore, both glycerol-free and glycerol-bound EcIlvA were used as the isoleucine-free EcIlvA structure for further discussion.

Crystal structure of the E. coli IlvA regulatory domain (EcIlvA_R)

We failed to obtain full-length EcIlvA crystals containing bound isoleucine. Therefore, we engineered an expression construct that encompassed the proposed EcIlvA regulatory domain starting at residue Glu335 (EcIlvA_R) based on the EcIlvA crystal structure (Figs. S1*A* and S2*A*). The isolated EcIlvA_R was soluble and successfully purified (Fig. S1*B*), and was a 36 kDa dimer based on gel filtration chromatography (Fig. S1*C*). Crystals were obtained in space group $P3_221$, and the structure was determined at 1.60 Å resolution (Table S1). There is one EcIlvA_R in the asymmetric unit (Fig. 3*A*) and the dimeric assembly is generated by crystallographic symmetry (Fig. 3*B*). Each EcIlvA_R comprises six helices (H15-H20) and eight strands (S11-S18) (Fig. 3*A*) folded into two ACT-like subdomains connected by loop L19: ACT1 (helices H15-H17 and strands S11-S14) and ACT2 (helices H18-H20 and strands S15-S18). ACT1 and ACT2 are very similar with the RMSD of 2.3 Å over 69 $C\alpha$ atoms. The prototypical ACT domain has a $\beta\alpha\beta\beta\alpha\beta$ topology but ACT1 and ACT2 each have an extra helix (H17 and H20, respectively) leading to a $\beta\alpha\beta\beta\alpha\beta\alpha$ organization (Fig. 3*A*). H17 and H20 form the dimerization interface (Fig. 3*B*).

The disordered region (482-496) in the regulatory domain of the full-length EcIlvA structure was well-defined in the EcIlvA_R structure and comprises loop L23 and helix H19 (Figs. 3D and S2). To validate crystal structure organization, Wwe used SEC-SAXS analysis to compare the dimeric EcIlvA_R crystal structure with the volume occupied by the EcIlvA_R dimer in solution (Figs. S4). SAXS experiments showed that EcIlvA_R behaves as a dimer with molecular weight of 42.4 kDa (Figs. S4*B*, S4*E and* Table S2). A comparison of the EcIlvA_R experimental data with the predicted profile of the EcIlvA_R dimer based on the crystal structure (Fig. 3*B*) shows that the SAXS data are a very close fit to the theoretical behavior of the EcIlvA_R dimer (χ^2 =3.10) (Fig. S4*H*). The EcIlvA_R crystallographic dimer was fit into the aligned *ab initio* electron density map generated from the solution scattering data, with the correlation coefficient of 0.82. This shows the close alignment of the EcIlvA_R crystal and solution structures (Fig. 3*C*).

The EcIlvA_R structure revealed the presence of one isoleucine molecule (Figs. 3*A* and S5*B*) that binds between sheets S11-S13, loop L14 and helix H15 of the ACT1 subdomain, and loop L21 and sheet S16 of ACT2 subdomain (Fig. 3*D*). This is the same location where glycerol is bound to full-length EcIlvA (Fig. *3D*). The amino group of isoleucine forms hydrogen-bonding interactions with the main chain carbonyl oxygen atom of Ile460 and side chains of Glu347 and Asn459 that close over the isoleucine pocket (Fig. 3*E*). The isoleucine carboxyl group forms hydrogen-bonding interactions with the amide nitrogen atoms of Phe352 and Ile460. Residues Ser351 and Phe352 are at the amino terminus of helix H15, which also allows a favorable dipole interaction with the carboxyl group. In addition, the carboxyl group interacts with the carbonyl oxygen atom of Trp458 and the amide nitrogen atoms of Gly350 and Leu353 via water molecules (Fig. 3*E*). The

side chain of the bound isoleucine engages a hydrophobic pocket composed of residues Ile345, Ser351, F352, Tyr369, Ala377, Ile379, Ile460, and Phe463 (Fig. 3*E*).

Notably, there is a conformational change in EcIlvA_R (Fig. 3*F*) that distinguishes it from the isoleucine-free regulatory domain of EcIlvA (Fig. 2*C*). The side chain of Phe352 rotates downward to accommodate the isoleucine terminal methyl group, which triggers a series of structural changes in the residues along helices H15 and H18 (F352-L451-F367-L447-V364) that result in the movement of loops L15 and L20 some 23 Å away from the isoleucine pocket (Fig. *3F*). Loop L20 places the Cα atom of Pro444 3 Å away from its position in the isoleucine-free structure. There are a total of five glycerol molecules in the EcIlvA_R structure, two of them bound in the dimer interface (Fig. 3*A*). Interestingly, a glycerol molecule is bound in a pocket around loop L15, which is formed by a conformational change due to isoleucine binding (Figs. 3*A* and 3*D*). In the isoleucine-free and isoleucine-bound structures, the distances between the Cα atoms of Ser363 and Pro444 are 5.0 Å and 8.6 Å, respectively. Loops L15 and L20 are further apart in the isoleucine-bound structure (Fig. 3*F*). These data show that isoleucine binding requires the movement of the Phe352 sidechain that, in turn, initiates a conformational change that ripples across the EcIlvA regulatory domain.

Biochemical analysis of EcIlvA(F352A)

The structural analysis of EcIlvA_R indicates that Phe352 is the key residue in the isoleucine binding site that is required to initiate the conformational signal when the ligand binds. Therefore, we purified and characterized the EcIlvA(F352A) mutant (Figs. S1*A*-*C*). As reported (9-12, 15), kinetic analysis shows that isoleucine inhibits the binding of threonine to EcIlvA (Fig. 4*A*). In the enzymatic assays used here, the formation of the IlvA

product 2-ketobutyrate was measured (14, 21, 24) and hence the aldimine formation between threonine and PLP was not used for as a readout for the enzymatic assay. Threonine binding is normally cooperative (n_H = 1.9) and at 0.5 mM, isoleucine extinguishes EcIlvA activity regardless of the threonine concentration (Figs. 4*A* and 4*C*). Threonine binding to EcIlvA(F352A) has approximately the same binding kinetics as to wild type EcIlvA, but the binding to EcIlvA(F352A) is not cooperative (n_H = 1.04) and the enzyme remained active in the presence of 5 mM isoleucine (Figs. 4*B* and 4*C*). The ability of isoleucine to inhibit EcIlvA and EcIlvA(F352A) was directly compared, and isoleucine is a potent inhibitor of EcIlvA, but not EcIlvA(F352A) (Fig. 4*D*). Thus, the kinetic analysis of EcIlvA(F352A) supports the proposed role of Phe352 in initiating the conformation change required for isoleucine to inhibit EcIlvA. It was essential to demonstrate that the inability of isoleucine to inhibit the EcIlvA(F352A) mutant was not simply due to the lack of binding to the enzyme. We therefore used a thermal stabilization assay to measure isoleucine binding to EcIlvA and EcIlvA(F352A) (Fig. *4E*). EcIlvA and EcIlvA(F352A) had nearly identical stabilities to thermal denaturation in the absence of isoleucine, and the stability of both proteins was increased ~14 °C in the presence of isoleucine. These data confirm that both proteins bind isoleucine (Fig. *4E*).

Crystal structure of the EcIlvA regulatory domain F352 mutant in complex with isoleucine (EcIlvA_R(F352A)•Ile)

The impact of the Phe352 mutation on the regulatory domain conformation was assessed by purifying EcIlvA_R(F352A) (Figs. S1*A*-*C*) and solving the crystal structure of the EcIlvA_R(F352A)•Ile complex (Table S1 and Fig. S5*C*). The EcIlvA_R(F352A) protein was less stable by 4.9 °C to thermal denaturation compared to EcIlvA_R, but the presence

of isoleucine greatly stabilized both proteins illustrating that isoleucine binds to EcIlvA_R(F352A) (Fig. *4F*). These results using the isolated regulatory domain agree with the data for the full-length protein (Fig. *4E*). Crystals of the EcIlvA_R(F352A)•Ile complex were obtained in space group C2, and the structure was determined at 1.22 Å resolution (Table S1). The asymmetric unit contains a monomer, and the dimeric assembly was generated through crystallographic symmetry. Like the structure of EcIlvA_R•Ile, an isoleucine molecule is bound at the same location in EcIlvA_R(F352A)•Ile (Figs. 5*A* and S5*C*). However, due to the absence of the Phe352 sidechain that initiates the conformational change in the regulatory domain (Fig. 3*F*), isoleucine binding does not trigger a conformational change in the EcIlvA_R(F352A) mutant (Fig. 5*A*). The overlay of EcIlvA_R(F352A) with the regulatory domain of EcIlvA shows that the two structures have the same conformation. A close-up view of the residues between the helices H15 and H18 shows that the sidechains of Phe355, Leu359, Val364, Phe367, Leu447 and Leu451 in EcIlvA_R(F352A) all adopt the same conformation as found in the Isoleucine-free (glycerol bound) regulatory domain of EcIlvA (Fig. 5*B*). Loops L15 and L20 are superimposed, and the Cα distance between Ser363 and Pro444 is 4.4 Å, which is similar to the Isoleucine-free regulatory domain of EcIlvA. A comparison of the overall structure of EcIlvA_R(F352A)•Ile to EcIlvA_R•Ile highlights their differences (Fig. 5*C*), and a close-up view of helices H15 and H18 illustrates how the absence of the Phe352 sidechain prevents the conformational change caused by isoleucine binding to extend down the regulatory domain (Figs. 5*B* and 5*C*).

Threonine binding at the catalytic site

Although the binding mode of PLP is structurally well characterized in the EcIlvA structures reported here, none contain a bound threonine. Therefore, we docked threonine into the catalytic site, and this identified loop L3, loop L10 and helix H5 as key components to accommodating the substrate (Fig. S6*A*). The amino group of threonine is adjacent to the aldehyde carbon atom of PLP and is hydrogen bonded to the carbonyl oxygen atom of Gly241 in loop L10. The carboxyl group is ideally positioned to interact with the positive dipole at the N-terminus of helix H5, and it also forms hydrogen-bonding interactions with the amide nitrogen atom of His90 and the side chain hydroxyl group of Ser86. Finally, the side chain of threonine interacts with the side chains of Glu240 and Val244 in Loop L10. The location of the active site and the binding of threonine are consistent with allosteric regulation operating across the protomer interface. Most notably, loop L10 directly interacts with loop L20' in the regulatory domain of the adjacent protomer (Fig. S6*B*). More specifically, the conformational changes of Pro 441', Ser 443' and Pro444' in loop L20', and D471' and Y472' in loop L22' are well positioned to affect the conformations of Arg234, Leu237, and Phe238 in loop L10 and to impact the binding of the threonine substrate (Figs. S6*B* and S6*C*). We also note that the conformational changes elicited by isoleucine could impact the positioning of the adjacent neck helix H14 and, in turn, the dimer structure (Fig. S6*C*). In terms of the catalytic mechanism, we note that the amino group of the docked threonine is ideally positioned to perform the required initial transamination reaction on the lysine-bound PLP cofactor (2) (Fig. S6*A*).

***Staphylococcus aureus* IlvA - biochemical analysis**

Branched-chain amino acids (BCAAs) are important regulators of virulence and fatty acid synthesis in Gram-positive bacteria including *Staphylococcus aureus* (25). Inspection of the *S. aureus* IlvA (SaIlvA) sequence revealed that, unlike EcIlvA, there is only one ACT domain in the regulatory domain. It was therefore of interest to investigate the regulation of SaIlvA, if any, by isoleucine. We engineered a SaIlvA full-length (residues 1-422) expression construct (Fig. S1*A*), and the expressed protein was soluble and successfully purified (Fig. S1*D*). The molecular weight based on analytical gel filtration chromatography was estimated to be 380 kDa, which suggests that the SaIlvA assembly is an octamer in solution (Fig. S1*E*). Kinetic analyses of SaIlvA revealed that the regulation of the protein's activity is completely different from EcIlvA. Isoleucine does not inhibit the binding kinetics of threonine (Fig. 6*A*). The enzyme kinetics data were analyzed using the Michelis-Menten model and the fitted curve closely followed the observed data points, indicating no cooperativity (Fig. 6A). The enzyme also remained active in the presence of 5 mM isoleucine (Figs. 6*A* and 6*B*). In order to analyze and confirm this fundamental difference between SaIlvA and EcIlvA, we used a different assay. Radioactive [^{14}C]-threonine was used as a substrate for both proteins and the reaction was resolved by TLC. EcIlvA was inhibited by isoleucine while SaIlvA was not (Fig. S7*A*). The same assay was used to determine if the regulation of SaIlvA involved other intermediates in the branched-chain amino acid pathway. The intermediates and the end product of the isoleucine biosynthesis pathway were tested as inhibitors of SaIlvA and none of the metabolites inhibited SaIlvA activity. These included 2-ketobutyrate (IlvA product), α-keto-β-methylvalerate (the substrate for IlvE), 2-methylbutyrate, 2-methylbutyryl-CoA (product of branched chain keto acid dehydrogenase), isoleucine, valine, serine and the end product anteiso15:0 fatty acid (Fig S7*B*). Finally, using the thermal stabilization assay, we confirmed that isoleucine does

not bind to SaIlvA (Fig. 6*C*). Unlike EcIlvA, which was increased ~14 °C in the presence of isoleucine (Fig. *4E*), SaIlvA had nearly identical stabilities to thermal denaturation in both the absence and presence of 5 mM isoleucine (Fig. *6C*). The spectral scans indicate a spectral shift from 410 nm to 430 nm upon threonine binding indicating the external aldimine formation between PLP and threonine for both EcIlvA and SaIlvA (Figs. S7*C* and S7*D*). This shift is similar to other PLP binding enzymes like aspartate aminotransferase that show a shift in the spectra to 430 nm indicating enzyme-substrate interaction (26, 27).

Crystal structure of the S. aureus IlvA regulatory domain (SaIlvA_R)

To understand the differences between EcIlvA and SaIlvA, we performed a crystallographic analysis on the latter. Full length SaIlvA failed to generate crystals, but the regulatory domain (SaIlvA_R) was successfully crystallized. The SaIlvA_R gene (residues 336-422) was cloned with an N-terminal 6xHis-tag (Fig. S1*A*) and the expressed protein was soluble and successfully purified (Fig. S1*D*). The protein was estimated to be a 46 kDa tetramer based on gel filtration chromatography (Fig. S1*E*), and crystals were obtained in space group $I4_1$, and the structure of SaIlvA_R was determined at 2.8 Å resolution (Table S1).. Two tetramers are in the asymmetric unit and are related by a 173° rotation, slightly deviating from perfect 2-fold symmetry, and the RMSD between the two tetramers is 1.2 Å. The tetrameric assembly of SaIlvA_R is consistent with gel filtration analysis and can be considered as a dimer of dimers comprising protomers A/B and C/D (Figs. S1*E* and 7*A*). All eight protomers in the asymmetric unit are structurally very similar with RMSD values between 0.5 to 1.5 Å on Cα atoms in pairwise comparisons.

Unlike EcIlvA_R, which is composed of two ACT domains (ACT1 and ACT2), SaIlvA_R is composed of one ACT domain as predicted from the sequence (Figs. S8*A* and *S8B*)

(28). Protomers A/B and C/D generate dimers that closely resemble the EcIlvA_R structure (Figs. 7*B* and 7*C*). The overall alignment of the central four sheets (S11, S12, S13, and S14) of the SaIlvA_R monomer was quite good with low RMSD value when overlaid with the isolated ACT1 domain of EcIlvA_R, indicating similarity in their core structures (Fig. 7*D*). The RMSD value for the entire ACT domain is 3.4 Å, whereas for the central four sheets, it is significantly lower at 0.6 Å. SaIlvA_R also has an extra helix (H17) in the prototypical ACT domain, like the ACT1 and ACT2 subdomains of EcIlvA_R (Figs. 7*B* and 7*C*). Sheets S12 and S12′ and helices H15 and H15′ from the partner protomers participate in four key interactions along the dimeric interface: Leu352 interacts with Leu352′, Val356′ and Ile366′; Val356 interacts with Leu352′, Arg353′ and Val356′; Ile366 interacts with Leu352′; Phe369 interacts with Phe369′. In addition, Asn357 and Asn357′ on helices H15 and H15’ form hydrogen bonding interactions (Fig. S9*A*). Protomers A/D and B/C also create dimers, and their interface comprises the additional helices H17 and loops L16 that contain a disordered region (Fig. S9*B*).

Despite the structural similarities between EcIlvA_R and the SaIlvA_R A/B and C/D dimers, key differences appear to prevent isoleucine binding to the latter. Helix H15 and loops L14, L15 and L16 exhibit different conformations compared to their EcIlvA_R counterparts, and the H15 and L14 counterparts play crucial roles in the binding of isoleucine in EcIlvA (Figs. 7*B* and 7*C*). Also, the relative orientations of the ACT1 and ACT2 subdomains of EcIlvA_R and the A/B and C/D protomers of SaIlvA_R are significantly different; when ACT1 of EcIlvA_R is aligned with protomer A, protomer B is rotated ~35° from the ACT2 domain of EcIlvA_R (Fig. 7*D*). Isoleucine binds at the interface of ACT1 and ACT2 in EcIlvA_R and the different orientations in the A/B and C/D dimers of SaIlvA_R disrupt the isoleucine binding site (Fig. 7*D*). More specifically, the different

location of sheet S12′ in protomer B corresponding to sheet S16 of the ACT2 domain of EcIlvA disrupts the binding of isoleucine in SaIlvA (Fig. 7*D*). The conformation of helix H15 in SaIlvA_R is different from that in EcIlvA_R, so the isoleucine model obtained by overlaying the entire an ACT domain could not interact with helix H15 (Fig. 7D). Therefore, even when SaIlvA_R and EcIlvA_R•Ile were overlaid only on helix H15 to locate the isoleucine that could interact with helix H15 in SaIlvA_R, the binding of isoleucine on SaIlvA was not favorable (Figs. S9*C* and S9*D*). Specifically, Glu347 interacts with the amino group of the isoleucine in the EcIlvA_R complex, whereas the corresponding residue, Gln347, in the SaIlvA_R is oriented toward the solvent. Two other isoleucine interacting residues, Asn459 and Ile460, are located in sheet S16 of EcIlvA_R, whereas their counterparts, Asp365′ and Ile366′, are located in sheet S12′, which has a different orientation compared to sheet S16 in EcIlvA_R. Leu352 in SaIlvA_R corresponds to Phe352 in EcIlvA, a key residue for the feedback inhibition by isoleucine (Figs. S9*C* and S9*D*). This analysis indicates that isoleucine binding is not favorable in SaIlvA and therefore unregulated by isoleucine.

Structural model of the SaIlvA assembly

In the absence of a full-length SaIlvA crystal structure, we predicted the assembly using a combination of the protomer AlphaFold model (AF-Q2FF63) and our crystal structure of the SaIlvA_R tetramer. The AlphaFold model was first compared with the crystal structure of EcIlvA (PDB ID: 9D2Q) and the catalytic domains are very similar, consistent with their 36.3% sequence identity (Figs. S8*A* and S8*B*). The structure of the regulatory domain of the model is also very similar to the domain observed in the crystal structure of SaIlvA_R (PDB ID: 9D2T), except for loops L14 and L16 (Fig. S8*C*). Thus, the

individual domains of AlphaFold model appear to be excellent. However, when the model is superimposed on the tetrameric SaIlvA_R to generate the full assembly, there is no close contact between the catalytic domains (Fig. S10*A*). Two domains of SaIlvA are connected by a linker, allowing for potential flexibility in their relative orientation. To obtain an improved oligomeric structure, the AlphaFold model of the dimeric SaIlvA was first predicted using ChimeraX and ColabFold (Fig. S10*B*) (29, 30). The catalytic domain assembly of the resulting dimeric SaIlvA model is similar to that of the dimeric EcIlvA with interactions between the catalytic domains (Figs. S10*B* and S10*C*). In the crystal structure of SaIlvA_R, the N-termini of protomers A and C align in the same direction, while those of the protomers B and D face the opposite direction (Fig. 8*A*). Therefore, when the model of the dimeric SaIlvA was superimposed on protomers A/C and B/D of SaIlvA_R (Fig. 8*B*), the resulting tetramer model has a central regulatory domain assembly with paired catalytic domains extending in opposite directions (Fig. 8*C*). This is opposite of the EcIlvA assembly in which the catalytic domains constitute the core of the tetramer (Figs. 2*B* and 8*C*). Due to the agreement between copies in the asymmetric unit, it is unlikely that this organization is due to a crystallographic artifact.

As noted above, gel filtration of the purified SaIlvA suggested an octameric assembly but the modeling suggests a tetramer. To resolve this issue, we used SEC-SAXS analysis to determine the volume occupied by SaIlvA in solution (Figs. 8*D*, S4*C*, S4*F*, S4*I*, and Table S2). Comparison of the SaIlvA SAXS experimental data set with the predicted profile of the generated tetrameric model of SaIlvA is very poor, but the SAXS data fit well to the theoretical behavior of a SaIlvA octamer (χ^2 = 1.37) (Fig. S4*I*). Specifically, the SAXS analysis is consistent with SaIlvA existing as an octamer with a molecular weight of 321 kDa in solution, which agrees well with the gel filtration profile

(Table S2 and Fig. S1*E*). A SaIlvA octamer, consisting of two associated tetramers, was constructed that optimally fits the *ab initio* electron density maps generated from the solution scattering data, with the correlation coefficient of 0.79 (Fig. 8*D*).

Complementation of a SaIlvA knockout strain by EcIlvA and EcIlvA(F352A)

The Δ*SaIlvA* knockout strain PDJ78 was transformed by three plasmids, pCN38 (empty vector) (31), pCN38 SaIlvA or pCN38 EcIlvA. The three strains all grew in 'complete defined medium' (CDM) that contains isoleucine as expected, and the strain complemented with pCN38 SaIlvA also grew in CDM minus isoleucine consistent with the presence of the active enzyme that lacks feedback inhibition. However, the strains complemented with pCN38 or pCN38 EcIlvA could not grow in the absence of isoleucine due to the lack of the enzyme and the presence of feedback regulated EcIlvA, respectively (Fig. 9*A*). To understand the effect of the feedback regulation on fatty acid synthesis in these three strains, the cells at the end of the growth curve shown in panel A were harvested and their phospholipids were extracted. The fatty acids from the phospholipids were transformed into methyl esters and subjected to gas chromatography. In CDM, the fatty acid composition was remarkably indistinguishable in the three strains with a slight increase in straight chain odd fatty acids in the strain with SaIlvA (Fig. 9*B*). However, in the absence of isoleucine, there was an increase in *antieso* odd chain fatty acids in the pCN38 SaIlvA strain compared to the strain with pCN38 EcIlvA. This indicates that isoleucine was copiously synthesized to engage in fatty acid synthesis with SaIlvA present but not with EcIlvA, again confirming feedback inhibition with the latter (Fig. 9*C*). Also, an increase in straight chain even fatty acid was observed with the EcIlvA indicating that the decrease in *antieso* odd chain fatty acids was compensated with an increase in straight

chain even fatty acids in the absence of isoleucine (Fig. 9*C*). This change in the *antieso* to straight chain fatty acids would affect the membrane fluidity preventing the growth of *S. aureus* when complemented with EcIlvA.

To further analyze the regulation of isoleucine synthesis by EcIlvA, we expressed the F352A mutant of EcIlvA (pCN38 EcIlvA (F352A)) in the Δ*SaIlvA* knockout. The growth curves of this strain clearly show that it can grow in both the presence and absence of isoleucine (Fig. 9*D*), which is not the case with the EcIlvA strain (Fig. 9*A*). The fatty acid analysis (Figs. 10*E* and 10*F*) shows that there is a significant increase in *antieso* odd chain fatty acids and a decrease in the straight chain even fatty acids in the EcIlvA (F352A) compared to the EcIlvA even when grown in CDM. A significant decrease in the straight chain even fatty acids was observed when grown in the absence of isoleucine. The branched-chain fatty acids made due to the lack of feedback regulation in the EcIlvA (F352A) mutant allow the Δ*SaIlvA* knockout to grow in the absence of isoleucine in the media. These results confirm the importance of F352 in regulating isoleucine synthesis by EcIlvA.

Discussion

Despite the wealth of knowledge surrounding IlvA biochemistry, neither the isoleucine binding site within the IlvA regulatory domain nor the structural consequences of isoleucine binding had been directly observed before the studies presented here. Our studies initially focused on the IlvA from *E. coli*, and although we failed to structurally characterize full-length EcIlvA in its isoleucine-bound inactive state, we succeeded in determining the

crystal structure of the EcIlvA_R regulatory domain bound to isoleucine. Considering the historical importance of IlvA to the development of allosteric theory, it is gratifying that the binding of isoleucine results in defined conformational changes within the regulatory domain that are propagated across the structure via a clear allosteric pathway. When considered in terms of the full-length dimer-dimer tetrameric assembly of EcIlvA, the direction of the allosteric pathway strongly suggests that the allosteric regulation occurs across the dimer interface with the isoleucine binding to the EcIlvA_R of one protomer negatively regulating the active site of the catalytic module of the adjacent protomer (Fig. 10). The structural data that support this are twofold. First, the active site where threonine binds in the catalytic module is 43 Å away from the isoleucine site on the partner protomer regulatory domain but 54 Å away from its own regulatory domain (Fig. 10). Second, the crystal structures of EcIlvA_R presented here with and without bound isoleucine show how the isoleucine binding signal is transmitted 23 Å from the isoleucine binding site to the L20 loop, which is located at the dimer interface and directly contacts the partner active site (Figs. 10 and S6*B*). This requirement for allostery to operate within a quaternary structure that undergoes small but significant conformational changes upon effector binding is another basic tenet of the fundamental theory. Although we were not able to characterize these presumed changes, they may involve the neck helix H14 that connects the regulatory and catalytic subdomains of EcIlvA (Figs. 10 and S6*C*) and is in a position to be affected by the allosteric pathway. When the previously determined crystal structure of EcIlvA is compared with the higher resolution structure presented here, the catalytic and regulatory domains flanking the neck helix are in slightly different relative positions which suggests a degree of interdomain flexibility within the protomer.

The identification of Phe352 as the key residue at the start of the allosteric pathway directly adjacent to the bound isoleucine is an important finding that we tested directly by biochemically and structurally characterizing the F352A point mutant (Figs. 4, 5 and 10). These analyses confirmed the importance of this residue. EcIlvA(F352A) still binds isoleucine, but the structure of EcIlvA_R(F352A) is very similar to EcIlvA and the allosteric pathway is unaffected. Thus, F352A prevents the initial communication as predicted. Also, EcIlvA(F352A) is no longer inhibited by isoleucine despite clearly binding the effector. A number of studies have identified point mutants that render EcIlvA refractory to inhibition by isoleucine (12, 14, 21, 32). These residues are shown in Fig. S11 and they all occur within the regulatory domain, either directly affecting the isoleucine binding site or the allosteric pathway. It was reported that the ligand binding site mutants E347F, F352A, Y369L, and I460F became remarkably insensitive to isoleucine inhibition, with increased IC50 values by factors 43, 34, 91, and 49, respectively (14). Arg362 in loop L15 makes a salt bridge with Glu394 in the EcIlvA_R structure. The R362F mutant, lacking salt bridge, became strongly feedback-resistant with an increased IC50 value of 104-fold compared to that of the wild-type EcIlvA (14). Furthermore, previous studies with EcIlvA, mutations in residues P441L, G445E, A446E, L447A, L451A, and L454A, which are in the allosteric pathway, showed resistance to isoleucine inhibition (12, 14, 21, 32). It should be noted that the location of the isoleucine binding site is different from the binding site that was inferred by computational modeling and mutagenesis (21).

Unlike *E. coli*, *S. aureus* maintains its membrane with branched-chain fatty acids (33, 34). Isoleucine derived exogenously or synthesized *de novo* is converted to keto acids and diverted into FASII. Branched chain *iso* and *antieso* fatty acids both contribute to membrane fluidity. It has been shown that removal of isoleucine from the media increases

iso fatty acids to maintain membrane fluidity (35). This key role of isoleucine in *S. aureus* prompted us to analyze SaIlvA and its regulation. The sequence of SaIlvA suggests that the regulatory domain SaIlvA_R has a different architecture than EcIlvA_R with a single ACT subdomain, and we demonstrated that SaIlvA is not allosterically controlled by isoleucine and indeed does not bind isoleucine. The essentiality of isoleucine to membrane phospholipids in *S. aureus* may explain why allosteric feedback inhibition is not present in SaIlvA (Fig. 9). The crystal structure of SaIlvA_R reveals that it is a tetrameric assembly with two dimers that each resemble the EcIlvA_R protomer, but the isoleucine binding site is not present due to differing orientations of the paired ACT domains. The tetrameric assembly of the ACT domains in SaIlvA_R reflects the variety of ways in which this domain can generate multimers (20, 36, 37). Modeling suggests that the tetrameric assembly of SaIlvA_R forms the core of the SaIlvA assembly, unlike EcIlvA where the catalytic domain creates the core of the tetramer. This further suggests that it is the catalytic domains that extend outwards from the regulatory domains, which is opposite from the EcIlvA assembly. SEC-SAXS data and gel filtration profiles both agree that SaIlvA exists as an octamer in solution. The SEC-SAXS analyses are consistent with two associated tetrameric assemblies, but we have no clear model to explain this. Although the crystal structure of SaIlvA_R has two tetramers in the crystal asymmetric unit, the optimal SaIlvA octamer based on the SEC-SAXS is not consistent with this asymmetric unit, which is therefore unlikely to structurally determine the SaIlvA octamer.

IlvA plays a crucial role not only in BCAA (branched-chain amino acids) production in bacteria but also in plants, offering commercial opportunities for amino acid overproduction through fermentation (38, 39) and for enhancing crop production (40, 41). Isoleucine, an essential amino acid for humans, is vital for immunity, blood sugar regulation, and

hormone production (42-45). Our analysis of IlvA's feedback inhibition mechanism and the key residues involved in its active and allosteric binding sites opens new possibilities for developing small molecule effectors and engineered variants to enhance isoleucine production. For instance, our experiments with the EcIlvA (F352A) mutant demonstrated that it can bypass isoleucine feedback regulation, leading to increased isoleucine biosynthesis during fermentation. Additionally, we showed in *S. aureus* that combining IlvA sequence analysis from other bacteria with AlphaFold structural predictions provides valuable insights into their allosteric regulation. This approach could potentially be extended beyond bacteria to crops, enhancing their nutritional and therapeutic value (46).

Experimental procedures

Materials

All chemicals and reagents were reagent grade. Pyridoxal phosphate (PLP) monohydrate was purchased from TCI (Product Number C0377). Crystallization screening kits and Sypro Orange dye were purchased from NeXtal Biotechnologies and Invitrogen, respectively. *E. coli* strains were routinely grown in rich broth (10 g/l tryptone, 5 g/l yeast extract, 5 g/l NaCl*).*

*Cloning, Expressio*n and Purification of IlvA

The *EcIlvA* gene (b3772) and the regulatory domain (amino acids 335-514) of *EcIlvA* were amplified by PCR using primers designed for Gibson Assembly cloning into NdeI and EcoRI digested pET28a to obtain pPJ669 and pPJ668, respectively. Site directed mutagenesis was performed with pPJ669 and pPJ668 using quick change lightning mutagenesis (Stratagene) to obtain F352A mutant in the full length and regulatory domain of IlvA, respectively. SaIlvA (SaUSA300_2014) and its regulatory domain were amplified by PCR using primers designed for Gibson Assembly into pET28a. The SaIlvA was cloned into NcoI and XhoI digested vector while the regulatory domain was cloned into NdeI and EcoRI digested vector to obtain pPJ665 and pPJ 666 respectively. The *SaIlvA* knockout (PDJ78) was generated in strain AH1263 by allelic replacement. For this approximately 1000 bp region the upstream and downstream of SAUSA300_2014 were PCR amplified using primers SaIlvA UP-EcoRI-F, SaIlvA UP -R, SaIlvA DN-F, SaIlvA DN-KpnI-R and gel purified. The PCR products were assembled into pJB38 by Gibson Assembly (New England Biolabs) yielding pPJ676. Plasmid pPJ676 was transformed into AH1263 by electroporation and the knockout was generated as previously described (47). The

knockout strain PDJ78 was confirmed by PCR. For testing complementation of Δ*IlvA,* the *SaIlvA* and *EcIlvA* genes were cloned into pCN38 vector (31). Briefly, for cloning *EcIlvA*, two fragments were generated by PCR, a 189 bp fragment containing the SarAP1 promoter, and 1633 bp fragment containing *EcIlvA*. Similarly, for *SaIlvA*, a 191 bp fragment containing the SarAP1 promoter and a 1347 bp fragment containing *SaIlvA* were generated. The two fragments for each EcIlvA and SaIlvA were assembled with SphI and BamHI digested pCN38 by Gibson Assembly. The plasmids were sequenced and transformed into Δ*IlvA* to test complementation. The pCN38 EcIlvA F352A mutant plasmid was generated by site-directed mutagenesis as described above. The list of primers and plasmids generated is given in Table S3. The pET28a plasmids containing the EcIlvA, EcIlvA(F352A), EcIlvA_R, EcIlvA_R(F352A), SaIlvA and SaIlvA_R were transformed into BL21(DE3) cells and grown to an OD_{600} of 0.7 at 37 °C with shaking at 200 rpm. The culture was induced for expression with 1mM IPTG and grown overnight at 16 °C with shaking at 200 rpm. For the full-length proteins EcIlvA, EcIlvA(F352A), and SaIlvA, cells were harvested and resuspended in 20 mM potassium phosphate buffer, pH 7.5, 500 mM NaCl, 0.1 mM TCEP, 1 mM PLP, 10 mM imidazole and were broken by two passages through a cell disruptor. The cell lysate was centrifuged at 20,000 g for 45 minutes and IlvA was purified from the supernatant using a Ni-NTA column by washing with 20 column volumes of each 20 mM potassium phosphate buffer, pH 7.5, 500 mM NaCl, 0.1 mM TCEP, 0.1 mM PLP containing 10 mM imidazole or 20 mM imidazole or 40 mM Imidazole. IlvA was eluted with 20 mM potassium phosphate buffer (pH 7.5), 500 mM NaCl, 0.1 mM TCEP, 0.1 mM PLP and 250 mM imidazole. The regulatory domain proteins IlvA_R and IlvA_R (F352A) pellets were resuspended in 20 mM Tris, pH 8.0, 500 mM NaCl, 10 mM imidazole. The protein was purified as described above except the buffer contained 20 mM

Tris, pH 8.0, 500 mM NaCl and imidazole. SaIlvA_R pellets were resuspended in 20 mM Bis-Tris, pH 6.5, 500 mM NaCl, 10 mM imidazole, and 0.5% CHAPS. The protein was purified as described above except the buffer contained 20 mM Bis-Tris, pH 6.5, 500 mM NaCl, 250 mM imidazole and 0.5% CHAPS. All proteins except EcIlvA_R were further purified by gel filtration chromatography using a HiLoad 16/600 Superdex 200 column (Cytiva Life Sciences). EcIlvA_R was dialyzed against 20 mM Tris, pH 7.5. 200 mM NaCl. EcIlvA, and EcIlvA(F352A) were eluted with 20 mM potassium phosphate buffer, pH 7.5, 200 mM NaCl, 0.1 mM TCEP. EcIlvA_R(F352A) was eluted with 20 mM Tris, pH 7.5, 200 mM NaCl. SaIlvA was eluted with 20 mM potassium phosphate buffer, pH 7.5, 500 mM NaCl, 0.1 mM TCEP. SaIlvA_R was eluted with 20 mM Bis-Tris, pH 6.5, 200 mM NaCl. The molecular weight was calculated by analyzing the peak size on a HiLoad 13/300 Super 200 analytical column (Cytiva Life Sciences). A standard curve was generated using thyroglobulin (669 kDa), ferritin (440 kDa), aldolase (158 kDa), conalbumin (75 kDa), ovalbumin (44 kDa), ribonuclease A (13.7 kDa), and aprotinin (6.5 kDa).

IlvA activity assay

The activity assay was performed by a spectrophotometric method using purified recombinant EcIlvA, EcIlvA(F352A) and SaIlvA proteins as described previously (14, 21, 24). Briefly, the reaction mixture contained 100 mM potassium phosphate buffer, pH 7.5, 20 μM PLP, 50 nM enzyme, and varied concentration of threonine (0 to 100 mM) for EcIlvA and EcIlvA(F352A). Assays in Figs. 4*A* and 4*B* were carried out in the presence and absence of 0.05 mM and 0.5 mM isoleucine for EcIlvA and 5 mM isoleucine for EcIlvA(F352). The effect of isoleucine on the activities of EcIlvA and EcIlvA(F352A) in Fig. 4*D* were assayed at the presence of 7 mM threonine and varied concentration of

isoleucine (0 to 0.5 mM). The reaction mixture was the same except 200 nM of enzyme and varied concentration of threonine (0 to 200 mM) for SaIlvA. Assays in Fig. 6*A* were carried out in the presence and absence of 0.5 mM and 5 mM isoleucine. The effect of isoleucine on the activity of SaIlvA in Fig. 6*B* was assayed at the presence of 30 mM threonine and varied concentration of isoleucine (0 to 5 mM). The reaction was carried at 30 °C and the formation of 2-ketobutyrate was measured continuously at 230 nm on a 96-well clear flat bottom UV-transparent microplate (Corning, Product No. 3635) using a BioTek Synergy H1 Microplate Reader (BioTek Instruments, Inc.). All measurements were repeated three times. The enzyme kinetics were calculated by fitting allosteric sigmoidal and Michaelis-Meneten equations for EcIlvA and SaIlvA, respectively, within GraphPad Prism 10.0.2 (GraphPad Software, Boston, Massachusetts USA, www.graphpad.com).

X-ray crystallography

EcIlvA, EcIlvA_R•Ile, EcIlvA_R(F352A)•Ile, and SaIlvA_R crystals were grown by the sitting and hanging drop vapor diffusion methods at 20 °C using a Mosquito crystallization robot (TTP Labtech). A final 5 mM isoleucine was added to the EcIlvA protein solution (10.2 mg/ml in 20 mM potassium phosphate, pH 7.5, 0.2 M sodium chloride, and 0.1 mM TCEP) and incubated for 30 minutes at room temperature before crystallization. The sitting drops were produced by mixing 0.2 μl of the protein solution with 0.2 μl of reservoir solution (0.1 M HEPES, pH 7.0, 0.1 M magnesium chloride, and 15% PEG 4000). Crystals were transferred into a cryo-protection solution (0.1 M HEPES, pH 7.0, 0.1 M magnesium chloride, 15% PEG 4000, 1 mM isoleucine, and 30 % glycerol) and flash frozen in liquid nitrogen. Diffraction data were collected at the SERCAT beamline 22-ID at the Advanced Photon Source and processed using xia2 (48). The EcIlvA structure was solved by

molecular replacement method with EcIlvA molecule (PDB ID: 1TDJ) as a search model using Phaser Crystallographic Software (49). To obtain the EcIlvA_R•Ile crystals, the sitting drops were produced by mixing 0.2 μl of the purified EcIlvA_R protein solution (10 mg/ml in 20 mM Tris-HCl, pH 7.5, 0.2 M sodium chloride, and 0.1 mM EDTA) with 0.2 μl of reservoir solution (0.2 M ammonium acetate and 20% PEG 3350). For cryoprotection, crystals were transferred into a reservoir solution supplemented with 30% glycerol, and subsequently cryocooled in liquid nitrogen. X-ray diffraction data were collected at the 17-ID-1 AMX beamline (50) at the National Synchrotron Light Source II (NSLS II), Brookhaven National Laboratory. Diffraction data were processed using autoPROC toolbox (51) coupled with XDS (52), POINTLESS (53), and AIMLESS (53). The EcIlvA_R structure was solved by molecular replacement method with the regulatory domain (residues 335-514) of the EcIlvA molecule (PDB ID: 1TDJ) as a search model using Phaser Crystallographic Software. To obtain the EcIlvA_R(F352A)•Ile crystals, a final 5 mM isoleucine was added to the EcIlvA_R(F352A) protein solution (9.4 mg/ml in 20 mM Tris-HCl, pH 7.5, and 0.2 M sodium chloride). Crystals were grown using a reservoir consisting of 0.04 M sodium acetate, pH 4.0, 0.08 M ammonium acetate, and 6% PEG 4000. Crystals were transferred into a cryo-protection solution (0.07 M sodium acetate, pH 4.0, 0.14 M ammonium acetate, 10.5% PEG 4000, 2 mM isoleucine, and 30 % glycerol) and flash frozen in liquid nitrogen. Diffraction data were collected at beamline 5.0.2 of the Advanced Light Source. Diffraction data were processed using XDS and AIMLESS. The EcIlvA_R(F352A)•Ile structure was solved by molecular replacement method with EcIlvA_R•Ile structure (PDB ID: 9D2R) as a search model using Phaser Crystallographic Software. To obtain the SaIlvA_R crystals, the hanging drops were produced by mixing 1 μl of the purified SaIlvA_R protein solution (10 mg/ml in 20 mM Bis-Tris, pH 6.5, and 0.2 M sodium chloride) with 1 μl of reservoir

solution (0.1 M CAPSO, pH 9.0, 1.35 M ammonium sulfate and 0.2 M sodium chloride). For cryoprotection, crystals were transferred into a mixture of mineral oil and Paraton-N, and flash frozen in liquid nitrogen. Diffraction data were collected at the SERCAT beamline 22-ID at the Advanced Photon Source and processed using HKL2000 (54). The SaIlvA_R structure was solved by molecular replacement method with the regulatory domain (residues 336-422) of the SaIlvA AlphaFold model structure (AF-Q2FF63) as a search model using Phaser Crystallographic Software. All structures were refined and optimized using PHENIX (55) and COOT (56), respectively. Data collection and refinement statistics are summarized in Table S1. Figures were rendered with PyMOL (version 2.5.5, Schrödinger, LLC).

SEC-SAXS analysis

SEC-SAXS was performed at BioCAT (beamline 18ID at the Advanced Photon Source, Chicago) with in-line size exclusion chromatography (SEC) to separate sample from aggregates and other contaminants thus ensuring optimal sample quality and multiangle light scattering (MALS), dynamic light scattering (DLS) and refractive index measurement (RI) for additional biophysical characterization (SEC-MALS-SAXS). The samples were loaded on a Superdex 200 Increase 10/300 GL column (Cytiva) run by a 1260 Infinity II HPLC (Agilent Technologies) at 0.6 ml/min. The flow passed through (in order) the Agilent UV detector, a MALS detector and a DLS detector (DAWN Helios II, Wyatt Technologies), and an RI detector (Optilab T-rEX, Wyatt). The SAXS sheath flow cell (57) consists of a 1.0 mm internal diameter quartz capillary with ~20 mm walls (effective path length 0.542 mm). A coflowing buffer sheath is used to separate samples from the capillary walls to minimize radiation damage. Scattering intensity was recorded using an Eiger2 XE 9M

(Dectris) detector which was placed 3.6 m from the sample giving us access to a q-range of 0.003 $Å^{-1}$ to 0.42 $Å^{-1}$. Exposures (0.5 s) were acquired every 1 s during elution. Buffer blanks were created by averaging regions flanking the elution peak and subtracted from exposures selected from the elution peak to create the I(q) vs q curves used for subsequent analyses. SAXS data were processed, and the overall parameters computed following standard procedures of the software package BioXTAS RAW 2.1.4 (58). The Guinier plots, the P(r) function and the SAXS parameters in Table S2 were calculated using described methods (59-63). The Evolving factor analysis (EFA) function (64) in RAW for EcIlvA_R was used to deconvolve overlapping chromatography data. The intensity at zero scattering angle I(0) and radius of gyration R_g were obtained by performing a linear fit to the Guinier plot. The particle dimension Dmax was determined from the pair-distance distribution P(r) with the program GNOM (59). Reconstructions (n = 20) were aligned, averaged, and refined in slow mode using DENSS (65) to calculate the electron density for the final three-dimensional volume. For EcIlvA and EcIlvA_R, the 'Align output to PDB/MRC' option was used with EcIlvA tetramer model (PDBID: 9D2Q) and EcIlvA_R dimer model (PDBID: 9D2R), respectively, during the DENSS run. The EcIlvA tetramer model, the EcIlvA_R dimer model and SaIlvA_R octamer model were fit into their SAXS electron density using UCSF ChimeraX, version 1.8 (30). Theoretical scattering profiles and qualities of fit (χ^2) were obtained from the FoXS webserver (66) using the atomic coordinates of EcIlvA tetramer, EcIlvA_R dimer, SaIlvA octamer and their experimental SAXS profiles. The theoretical dimensionless Kratky plots were computed in RAW using the theoretical scattering profiles obtained from FoXS.

Thermal stability measurements

Purified proteins EcIlvA, EcIlvA(F352A), EcIlvA_R, EcIlvA_R(F352A) and SaIlvA were subjected to thermal denaturation in the presence of Sypro Orange dye as described in (67). Briefly, 10 μM of each of the proteins were added to 30 μl of 20 mM HEPES, pH 7.5, 150 mM NaCl and 10X Sypro Orange dye in a 96 well plate. The stability of the proteins was analyzed in the presence and absence of 5 mM isoleucine. The plate was centrifuged at 1500 rpm for 2 min before being placed in an Applied Biosciences 7500 RealTime PCR instrument. A thermal scan from 25 °C to 95 °C was performed using an increment rate of 1 °C/min. All experiments were performed in triplicate. The data were fitted using Boltzmann sigmoidal equation within GraphPad Prism 10.0.2 (GraphPad Software, Boston, Massachusetts USA, www.graphpad.com).

Molecular docking

Molecular docking of threonine substrate with the EcIlvA protein was performed using Molecular Operating Environment (MOE) software (2018.01; Chemical Computing Group). For docking calculation, the threonine ligand was drawn using ChemDraw Professional and converted to a 3D model (version 17.1, PerkinElmer Informatics). An IlvA-PLP structure for docking calculation was generated from the IlvA structure (PDB ID: 9D2Q). All molecules except the protein and PLP were removed. The ligand and protein structures were imported into MOE and prepared using the “QuickPrep” function. After performing docking using the “Dock” function of MOE, a docking model with the possibility of reacting to the cofactor PLP was selected.

EcIlvA and SaIlvA assay using radioactive threonine

5 mM 14[C] threonine-specific activity (989 μCi/mmol) was incubated with 5 μM EcIlvA or SaIlvA in the presence of 40 μM PLP and 50 mM HEPES buffer or phosphate buffer (pH

7.5) for 30 minutes. 10 μl of the assay was spotted on a Silica gel 60 thin layer chromatography plate. The TLC was resolved using 1-Butanol: acetic acid: water (80:20:4 v/v/v). The TLC was air-dried and exposed to a phosphorimager screen overnight and scanned using Typhoon FLA9500 (GE Healthcare)

Growth curve and fatty acid analysis

For growth curve and fatty acid analysis ΔIlvA strain containing empty vector (pCN38) or pSaIlvA or pEcIlvA grown in Complete Defined Medium (CDM) overnight. This was diluted to OD600 of 0.1 and grown in CDM or CDM minus Isoleucine for 5-6 hours. For fatty acid analysis the cells were harvested after 6 hours and lipids were extracted by Bligh Dyer method (68). Methyl esters of the fatty acids were generated with the lipid fraction and analyzed by gas chromatography (69).

EcIlvA and SaIlvA spectra

The spectral scan of EcIlvA and SaIlvA was performed using nanodrop in the presence of Buffer (20 mM K PO4 Buffer (pH 7.5), 200 mM NaCl, 0.1 mM TCEP) plus 0.1 mM plp, or 30 mM threonine as indicated in the figure. The data was smoothed using Graphpad prism software.

Data availability

The atomic coordinates and structure factors have been deposited in the worldwide Protein Data Bank (EcIlvA, PDB ID: 9D2Q; EcIlvA_R•Ile, PDB ID: 9D2R; EcIlvA_R(F352A)•Ile, PDB ID: 9D2S; and SaIlvA_R, PDB ID: 9D2T). SAXS data and models have been deposited in the Small Angle Scattering Biological Data Bank

(SASDBD, https://www.sasbdb.org/) (70) under the accession codes SASDVC7 (EcIlvA), SASDVD7 (EcIlvA_R), and SASDVE7 (SaIlvA). All other data produced for this work are contained within the article.

Keywords

Escherichia coli, threonine dehydratase, IlvA, isoleucine, branched-chain amino acids, allosteric regulation, conformational change, fatty acid, X-ray crystallography, enzyme catalysis.

Supporting information

This article contains supporting information.

Acknowledgments

We thank the St. Jude Children's Research Hospital (SJCRH) Biomolecular X-Ray Crystallography Center for support with X-ray data collection, and the SJCRH Hartwell Center DNA Sequencing Shared Resource for DNA sequencing. We particularly thank Dr. Stephen W. White for thoughtful suggestions and valuable input in developing this manuscript.

The X-ray data were collected at three shared resources. (1) The Advanced Photon Source (APS) is a U. S. Department of Energy (DOE) Office of Science user facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. Diffraction data were collected at the Southeast Regional Collaborative Access Team (SER-CAT) 22-ID beamline. SER-CAT is supported by its member institutions (https://www3.ser.aps.anl.gov/contact-

us#TITLE_SER_CAT_Memberships), equipment grants (S10_RR25528, S10_RR028976 and S10_OD027000) from the National Institutes of Health, and funding from the Georgia Research Alliance. SAXS data were collected at BioCAT (18-ID beamline), which is supported by grant P30 GM138395 from the National Institute of General Medical Sciences of the National Institutes of Health. (2) Diffraction data were collected at the Center for BioMolecular Structure (CBMS) at the National Synchrotron Light Source II at Brookhaven National Laboratory and is primarily supported by the National Institutes of Health, National Institute of General Medical Sciences (NIGMS) through a Center Core P30 Grant (P30GM133893), and by the DOE Office of Biological and Environmental Research (KP1607011). The National Synchrotron Light Source II is operated under Contract No. DE-SC0012704 for the U.S. Department of Energy (DOE), Office of Science. (3) Diffraction data were collected at beamline 5.0.2 of the Advanced Light Source, a DOE Office of Science User Facility at Lawrence Berkeley Laboratory under Contract No. DE-AC02-05CH11231, and is supported in part by the ALS-ENABLE program funded by the National Institutes of Health, National Institute of General Medical Sciences, grant P30 GM124169-01.

Funding and additional information

This work was supported by National Institutes of Health grants AI166116 (C.D.R.), GM034496 (C.O.R.), Cancer Center Support Grant CA21765, and ALSAC, St. Jude Children's Research Hospital. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflict of interest

The authors declare that they have no conflicts of interest with the contents of this article.

Abbreviations

The abbreviations used are: IlvA, threonine deaminase; BCAA, branched-chain amino acids; PLP, pyridoxal 5′-phosphate; PGDH, phosphoglycerate dehydrogenase; ACT domain, a structural motif of 70-80 amino acids.

References

1. Amorim Franco, T. M., and Blanchard, J. S. (2017) Bacterial Branched-Chain Amino Acid Biosynthesis: Structures, Mechanisms, and Drugability *Biochemistry* **56**, 5849-5865
2. Liang, Y. F., Long, Z. X., Zhang, Y. J., Luo, C. Y., Yan, L. T., Gao, W. Y. *et al.* (2021) The chemical mechanisms of the enzymes in the branched-chain amino acids biosynthetic pathway and their applications *Biochimie* **184**, 72-87
3. Irons, J. L., Hodge-Hanson, K., and Downs, D. M. (2020) RidA Proteins Protect against Metabolic Damage by Reactive Intermediates *Microbiol Mol Biol Rev* **84**,
4. Shen, W., Borchert, A. J., and Downs, D. M. (2022) 2-Aminoacrylate stress damages diverse PLP-dependent enzymes in vivo *J Biol Chem* **298**, 101970
5. Umbarger, H. E. (1956) Evidence for a negative-feedback mechanism in the biosynthesis of isoleucine *Science* **123**, 848
6. Umbarger, H. E., and Brown, B. (1958) Isoleucine and valine metabolism in Escherichia coli. VII. A negative feedback mechanism controlling isoleucine biosynthesis *J Biol Chem* **233**, 415-420
7. Changeux, J. P. (1961) The feedback control mechanisms of biosynthetic L-threonine deaminase by L-isoleucine *Cold Spring Harb Symp Quant Biol* **26**, 313-318
8. Monod, J., Wyman, J., and Changeux, J. P. (1965) On the Nature of Allosteric Transitions: A Plausible Model *J Mol Biol* **12**, 88-118
9. Eisenstein, E. (1991) Cloning, expression, purification, and characterization of biosynthetic threonine deaminase from Escherichia coli *J Biol Chem* **266**, 5801-5807
10. Eisenstein, E., Yu, H. D., and Schwarz, F. P. (1994) Cooperative binding of the feedback modifiers isoleucine and valine to biosynthetic threonine deaminase from Escherichia coli *J Biol Chem* **269**, 29423-29429
11. Eisenstein, E. (1995) Allosteric regulation of biosynthetic threonine deaminase from Escherichia coli: effects of isoleucine and valine on active-site ligand binding and catalysis *Arch Biochem Biophys* **316**, 311-318

12. Eisenstein, E., Yu, H. D., Fisher, K. E., Iacuzio, D. A., Ducote, K. R., and Schwarz, F. P. (1995) An expanded two-state model accounts for homotropic cooperativity in biosynthetic threonine deaminase from Escherichia coli *Biochemistry* **34**, 9403-9412
13. Koerner, K., Rahimi-Laridjani, I., and Grimminger, H. (1975) Purification of biosynthetic threonine deaminase from Escherichia coli *Biochim Biophys Acta* **397**, 220-230
14. Chen, L., Chen, Z., Zheng, P., Sun, J., and Zeng, A. P. (2013) Study and reengineering of the binding sites and allosteric regulation of biosynthetic threonine deaminase by isoleucine and valine in Escherichia coli *Appl Microbiol Biotechnol* **97**, 2939-2949
15. Gallagher, D. T., Gilliland, G. L., Xiao, G., Zondlo, J., Fisher, K. E., Chinchilla, D. *et al.* (1998) Structure and control of pyridoxal phosphate dependent allosteric threonine deaminase *Structure* **6**, 465-475
16. Gallagher, D. T., Eisenstein, E., Fisher, K. E., Zondlo, J., Chinchilla, D., Yu, H. D. *et al.* (1998) Polymorphous crystallization and diffraction of threonine deaminase from Escherichia coli *Acta Crystallogr D Biol Crystallogr* **54**, 467-469
17. Chipman, D. M., and Shaanan, B. (2001) The ACT domain family *Curr Opin Struct Biol* **11**, 694-700
18. Grant, G. A. (2006) The ACT domain: a small molecule binding domain and its role as a common regulatory element *J Biol Chem* **281**, 33825-33829
19. Lang, E. J., Cross, P. J., Mittelstadt, G., Jameson, G. B., and Parker, E. J. (2014) Allosteric ACTion: the varied ACT domains regulating enzymes of amino-acid metabolism *Curr Opin Struct Biol* **29**, 102-111
20. Schuller, D. J., Grant, G. A., and Banaszak, L. J. (1995) The allosteric ligand site in the Vmax-type cooperative enzyme phosphoglycerate dehydrogenase *Nat Struct Biol* **2**, 69-76
21. Wu, M., Sun, Y., Zhu, M., Zhu, L., Lu, J., and Geng, F. (2021) Molecular Dynamics-Based Allosteric Prediction Method to Design Key Residues in Threonine Dehydrogenase for Amino-Acid Production *ACS Omega* **6**, 10975-10983
22. Sander, T., Farke, N., Diehl, C., Kuntz, M., Glatter, T., and Link, H. (2019) Allosteric Feedback Inhibition Enables Robust Amino Acid Biosynthesis in E. coli by Enforcing Enzyme Overabundance *Cell Syst* **8**, 66-75 e68
23. Hammel, M. (2012) Validation of macromolecular flexibility in solution by small-angle X-ray scattering (SAXS) *Eur Biophys J* **41**, 789-799
24. Davis, L. (1965) A spectrophotometric method for the assay of threonine dehydratase *Anal Biochem* **12**, 36-40
25. Kaiser, J. C., and Heinrichs, D. E. (2018) Branching Out: Alterations in Bacterial Physiology and Virulence Due to Branched-Chain Amino Acid Deprivation *mBio* **9**,
26. Hill, M. P., Carroll, E. C., Vang, M. C., Addington, T. A., Toney, M. D., and Larsen, D. S. (2010) Light-enhanced catalysis by pyridoxal phosphate-dependent aspartate aminotransferase *J Am Chem Soc* **132**, 16953-16961
27. Metzler, C. M., Rogers, P. H., Arnone, A., Martin, D. S., and Metzler, D. E. (1979) Investigation of crystalline enzyme-substrate complexes of pyridoxal phosphate-dependent enzymes *Methods Enzymol* **62**, 551-558
28. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. *et al.* (2004) UCSF chimera - A visualization system for exploratory research and analysis *J Comput Chem* **25**, 1605-1612
29. Mirdita, M., Schutze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., and Steinegger, M. (2022) ColabFold: making protein folding accessible to all *Nat Methods* **19**, 679-682
30. Pettersen, E. F., Goddard, T. D., Huang, C. C., Meng, E. C., Couch, G. S., Croll, T. I. *et al.* (2021) UCSF ChimeraX: Structure visualization for researchers, educators, and developers *Protein Sci* **30**, 70-82
31. Charpentier, E., Anton, A. I., Barry, P., Alfonso, B., Fang, Y., and Novick, R. P. (2004) Novel cassette-based shuttle vector system for gram-positive bacteria *Appl Environ Microbiol* **70**, 6076-6085

32. Chinchilla, D., Schwarz, F. P., and Eisenstein, E. (1998) Amino acid substitutions in the C-terminal regulatory domain disrupt allosteric effector binding to biosynthetic threonine deaminase from Escherichia coli *J Biol Chem* **273**, 23219-23224
33. Chen, X., Teoh, W. P., Stock, M. R., Resko, Z. J., and Alonzo, F., 3rd (2021) Branched chain fatty acid synthesis drives tissue-specific innate immune response and infection dynamics of Staphylococcus aureus *PLoS Pathog* **17**, e1009930
34. Parsons, J. B., Frank, M. W., Subramanian, C., Saenkham, P., and Rock, C. O. (2011) Metabolic basis for the differential susceptibility of Gram-positive pathogens to fatty acid synthesis inhibitors *Proc Natl Acad Sci U S A* **108**, 15378-15383
35. Frank, M. W., Whaley, S. G., and Rock, C. O. (2021) Branched-chain amino acid metabolism controls membrane phospholipid structure in Staphylococcus aureus *J Biol Chem* **297**, 101255
36. Kotaka, M., Ren, J., Lockyer, M., Hawkins, A. R., and Stammers, D. K. (2006) Structures of R- and T-state aspartokinase III -: Mechanisms of the allosteric transition and inhibition by lysine *Journal of Biological Chemistry* **281**, 31544-31552
37. Yoshida, A., Tomita, T., Kurihara, T., Fushinobu, S., Kuzuyama, T., and Nishiyama, M. (2007) Structural Insight into concerted inhibition of alpha 2 beta 2-type aspartate kinase from Corynebacterium glutamicum *J Mol Biol* **368**, 521-536
38. Park, J. H., and Lee, S. Y. (2010) Fermentative production of branched chain amino acids: a focus on metabolic engineering *Appl Microbiol Biotechnol* **85**, 491-506
39. Sanchez, S., Rodríguez-Sanoja, R., Ramos, A., and Demain, A. L. (2018) Our microbes not only produce antibiotics, they also overproduce amino acids *J Antibiot* **71**, 26-36
40. Joshi, V., Joung, J. G., Fei, Z. J., and Jander, G. (2010) Interdependence of threonine, methionine and isoleucine metabolism in plants: accumulation and transcriptional regulation under abiotic stress *Amino Acids* **39**, 933-947
41. Ufaz, S., and Galili, G. (2008) Improving the content of essential amino acids in crop plants: Goals and opportunities *Plant Physiol* **147**, 954-961
42. Doering, C. B., and Danner, D. J. (2000) Amino acid deprivation induces translation of branched-chain alpha-ketoacid dehydrogenase kinase *Am J Physiol Cell Physiol* **279**, C1587-1594
43. Yeo, T., Probert, F., Sealey, M., Saldana, L., Geraldes, R., Hoeckner, S. *et al.* (2021) Objective biomarkers for clinical relapse in multiple sclerosis: a metabolomics approach *Brain Commun* **3**, fcab240
44. Yu, D., Richardson, N. E., Green, C. L., Spicer, A. B., Murphy, M. E., Flores, V. *et al.* (2021) The adverse metabolic effects of branched-chain amino acids are mediated by isoleucine and valine *Cell Metab* **33**, 905-922 e906
45. Holecek, M. (2018) Branched-chain amino acids in health and disease: metabolism, alterations in blood plasma, and as supplements *Nutr Metab (Lond)* **15**, 33
46. Holecek, M. (2013) Branched-chain amino acids and ammonia metabolism in liver disease: therapeutic implications *Nutrition* **29**, 1186-1191
47. Bose, J. L., Fey, P. D., and Bayles, K. W. (2013) Genetic tools to enhance the study of gene function and regulation in Staphylococcus aureus *Appl Environ Microbiol* **79**, 2218-2224
48. Winter, G. (2010) xia2: an expert system for macromolecular crystallography data reduction *J Appl Crystallogr* **43**, 186-190
49. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software *J Appl Crystallogr* **40**, 658-674
50. Schneider, D. K., Soares, A. S., Lazo, E. O., Kreitler, D. F., Qian, K., Fuchs, M. R. *et al.* (2022) AMX - the highly automated macromolecular crystallography (17-ID-1) beamline at the NSLS-II *J Synchrotron Radiat* **29**, 1480-1494
51. Vonrhein, C., Flensburg, C., Keller, P., Sharff, A., Smart, O., Paciorek, W. *et al.* (2011) Data processing and analysis with the autoPROC toolbox *Acta Crystallogr D Biol Crystallogr* **67**, 293-302

52. Kabsch, W. (2010) Xds *Acta Crystallographica Section D-Biological Crystallography* **66**, 125-132
53. Agirre, J., Atanasova, M., Bagdonas, H., Ballard, C. B., Basle, A., Beilsten-Edmands, J. *et al.* (2023) The CCP4 suite: integrative software for macromolecular crystallography *Acta Crystallogr D Struct Biol* **79**, 449-461
54. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode *Methods Enzymol* **276**, 307-326
55. Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M. *et al.* (2012) Towards automated crystallographic structure refinement with phenix.refine *Acta Crystallogr D Biol Crystallogr* **68**, 352-367
56. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot *Acta Crystallogr D Biol Crystallogr* **66**, 486-501
57. Kirby, N., Cowieson, N., Hawley, A. M., Mudie, S. T., McGillivray, D. J., Kusel, M. *et al.* (2016) Improved radiation dose efficiency in solution SAXS using a sheath flow sample environment *Acta Crystallogr D* **72**, 1254-1266
58. Hopkins, J. B., Gillilan, R. E., and Skou, S. (2017) BioXTAS RAW: improvements to a free open-source program for small-angle X-ray scattering data reduction and analysis *J Appl Crystallogr* **50**, 1545-1553
59. Svergun, D. I. (1992) Determination of the Regularization Parameter in Indirect-Transform Methods Using Perceptual Criteria *J Appl Crystallogr* **25**, 495-503
60. Franke, D., Jeffries, C. M., and Svergun, D. I. (2015) Correlation Map, a goodness-of-fit test for one-dimensional X-ray scattering spectra *Nat Methods* **12**, 419-+
61. Rambo, R. P., and Tainer, J. A. (2013) Accurate assessment of mass, models and resolution by small-angle scattering *Nature* **496**, 477-+
62. Piiadov, V., Ares de Araujo, E., Oliveira Neto, M., Craievich, A. F., and Polikarpov, I. (2019) SAXSMoW 2.0: Online calculator of the molecular weight of proteins in dilute solution from experimental SAXS data measured on a relative scale *Protein Sci* **28**, 454-463
63. Franke, D., Jeffries, C. M., and Svergun, D. I. (2018) Machine Learning Methods for X-Ray Scattering Data Analysis from Biomacromolecular Solutions *Biophys J* **114**, 2485-2492
64. Meisburger, S. P., Taylor, A. B., Khan, C. A., Zhang, S., Fitzpatrick, P. F., and Ando, N. (2016) Domain movements upon activation of phenylalanine hydroxylase characterized by crystallography and chromatography-coupled small-angle X-ray scattering *J Am Chem Soc* **138**, 6506-6516
65. Grant, T. D. (2018) *Ab initio* electron density determination directly from solution scattering data *Nat Methods* **15**, 191-193
66. Schneidman-Duhovny, D., Hammel, M., Tainer, J. A., and Sali, A. (2016) FoXS, FoXSDock and MultiFoXS: Single-state and multi-state structural modeling of proteins and their complexes based on SAXS profiles *Nucleic Acids Res* **44**, W424-429
67. Huynh, K., and Partch, C. L. (2015) Analysis of protein stability and ligand interactions by thermal shift assay *Curr Protoc Protein Sci* **79**, 28 29 21-28 29 14
68. Bligh, E. G., and Dyer, W. J. (1959) A rapid method of total lipid extraction and purification *Can J Biochem Physiol* **37**, 911-917
69. Christie, W. W. (2003) *Lipid Analysis*, 3 Ed., The Oily Press, Bridgwater, UK
70. Kikhney, A. G., Borges, C. R., Molodenskiy, D. S., Jeffries, C. M., and Svergun, D. I. (2020) SASBDB: Towards an automatically curated and validated repository for biological scattering data *Protein Sci* **29**, 66-75

Figure legends

Figure 1. Role of threonine deaminase (IlvA) in the initiation of isoleucine biosynthesis. Threonine dehydratase/deaminase (IlvA) initiates the isoleucine arm of the branched-chain amino acid biosynthesis. RidA accelerates the spontaneous conversion of the IlvA product, 2-iminobutyrate, to 2-ketobuytrate, which is condensed with pyruvate by acetolactate synthase (IlvB/N) followed by ketoacid isomeroreductase (IlvC), dihydroxyacid dehydrase (IlvD) and branched-chain amino acid transaminase (IlvE) to result in isoleucine. Isoleucine is a potent feedback regulator of EcIlvA that maintains the intracellular isoleucine homeostasis by tuning the isoleucine biosynthetic rate to its utilization for protein synthesis. SaIlvA is not feedback regulated by isoleucine but is important for both amino acid metabolism and lipid metabolism in *S. aureus*.

Figure 2. Crystal structures of *E. coli* IlvA. *A*, structure of the EcIlvA monomer at 1.87 Å resolution (PDB ID: 9D2Q). It consists of an amino terminal catalytic domain (purple) connected to a carboxy terminal regulatory domain (yellow) by helix H14 (the neck, blue). This new higher resolution EcIlvA structure clarifies the conformation of 3 regions in the previously reported structure (PDB ID: 1TDJ) (dashed red circles, 1 to 3). The PLP cofactor (green balls) is covalently attached to Lys62 in the catalytic domain and there is a glycerol (salmon balls) bound in the regulatory domain. *B*, the aligned electron density map occupied by EcIlvA in solution was determined by small angle X-ray scattering (SAXS) and matches the crystal structure of the tetramer. *C*, the EcIlvA regulatory domain in the new structure (PDB ID; 9D2Q) (yellow) overlaid with the regulatory domain in (PDB ID: 1TDJ) (gray) illustrating how the conformation of loops L14, L16 and L21 are altered by glycerol

binding. *Inset*, a close-up view of the area where the conformational change occurs upon binding of isoleucine. *D*, the regulatory domain conformational change that occurs within the pocket in the presence of bound glycerol. Glycerol is depicted as salmon balls-sticks, hydrogen bond interactions are shown with dotted red lines, and water molecules are red spheres. *E*, the glycerol binding pocket in the ligand-free EcIlvA regulatory domain (PDB ID: 1TDJ). A water molecule is shown with red sphere. *F*, the dose response curve shows the effect of glycerol on the activity of EcIlvA. The initial velocity was measured in the presence of 7 mM threonine. Isoleucine was used as the inhibitor control compound.

Figure 3. Crystal structure of the EcIlvA_R•Ile complex. *A*, overview of the EcIlvA_R•Ile complex protomer with one isoleucine molecule (magenta balls) bound (PDB ID: 9D2R). The rotated view (below) illustrates the positions of the two helices (H17 and H20) that are added to the core structure consisting of two ACT domains inked by loop L19. Five glycerol molecules are depicted as gray sticks. *B*, the structure of dimeric EcIlvA_R•Ile complex with the protomers shown in cyan and salmon. The rotated view (below) shows how helices H17 and H20 mediate the formation of the dimer. *C*, the EcIlvA_R•Ile dimer fits well into the aligned electron density experimentally determined by SAXS in solution. *D*, superimposition of EcIlvA_R•Ile (cyan) with the glycerol (salmon balls-sticks) bound regulatory domain of EcIlvA (yellow) (PDB ID: 9D2Q). Isoleucine (magenta balls-sticks) is bound between loops L14 and L21 triggering structural changes in the two adjacent helices H15 and H18 and loops L15, L20, and L22. A glycerol molecule, depicted as gray sticks, is bound around Loop L15 of EcIlvA_R•Ile. *E*, close-up views of the isoleucine binding site. Isoleucine is depicted as magenta balls-sticks, water molecules (Wat) are red balls and hydrogen bonds are dashed red lines. *F*, an overlay of the

EcIlvA_R•Ile complex (cyan) with the EcIlvA regulatory domain (yellow; PDB ID: 9D2Q). The change in the sidechain rotamer of Phe352 must occur to accommodate isoleucine and leads to a cascade of structural changes along helices H15 and H18 that result in the movement of loops L15 and L20.

Figure 4. Isoleucine effect on the activity and stability of EcIlvA. A, effect of isoleucine on the steady-state kinetics of EcIlvA. *B*, effect of isoleucine on the steady-state kinetics of EcIlvA(F352A). The theoretical curve was analyzed by the allosteric sigmoidal equation using GraphPad Prism 9.4.1. for both EcIlvA and EcIlvA(F352A). The kcat values for EcIlvA and EcIlvA(F352A) were calculated to be 144.7 ± 1.5 sec^{-1} and 149.2 ± 2.2 sec^{-1}, respectively. Khalf, kcat and Hill coefficient (n_H) values are presented as mean values ± standard error (SE). *C*, close-up plots showing the steady-state kinetics of EcIlvA and EcIlvA(F352A) at the low concentration of threonine (0 – 10 mM). *D*, effect of isoleucine on the activity of EcIlvA and EcIlvA(F352A). The initial velocity was measured in the presence of 7 mM threonine. The data for EcIlvA is the same as the data shown in Fig. 2*F*. *E*, the thermal stability of EcIlvA and EcIlvA(F352A) in the absence and presence of 5 mM isoleucine. *F*, the thermal stability of EcIlvA_R and EcIlvA_R(F352A) in the absence and presence of 5 mM isoleucine. Data represented as mean ± standard deviation (error bars) from triplicate sets.

Figure 5. Crystal structure of EcIlvA_R(F352A). *A*, crystal structure of the EcIlvA_R(F352A)•Ile complex (PDB ID: 9D2S). EcIlvA_R(F352A)•Ile is superimposed on EcIlvA (top, PDB ID: 9D2Q) and EcIlvA_R(bottom, PDB ID:9D2R). *B*, close-up view of the

isoleucine binding site of EcIlvA_R(F352A)•Ile (brown) and EcIlvA (yellow). *C*, close-up view of the isoleucine binding site in EcIlvA_R(F352A)•Ile (brown) and EcIlvA_R•Ile (cyan).

Figure 6. Isoleucine effect on SaIlvA. *A*, effect of isoleucine on the steady-state kinetics of SaIlvA. The theoretical curve was analyzed by the Michaelis-Menten equation using GraphPad Prism 10.1.2. The kcat value was calculated to be 64.4 ± 1.3 sec^{-1}. Km and kcat values are presented as mean values ± standard error (SE). *B*, effect of isoleucine on the activity of SaIlvA. *Inset*, plot at the high concentration of isoleucine (0 – 5 mM). The initial velocity was measured in the presence of 30 mM threonine. *C*, the thermal stability of SaIlvA in the absence and presence of 5 mM isoleucine. Data represented as mean ± standard deviation (error bars) from triplicate sets.

Figure 7. Crystal structure of SaIlvA_R. *A*, tetramer of the SaIlvA_R crystal structure (PDB ID: 9D2T). Each monomer of the tetramer is related by an approximate twofold axis (read symbol and arrows). Two tetramers are in the asymmetric unit. *B*, dimeric structure of SaIlvA_R. Protomers A and B are shown in green and orange, respectively. *C*, monomeric structure of EcIlvA_R•Ile (PDB ID: 9D2R). The two ACT domains are shown in cyan ribbon with highlighted yellow and blue. The bound isoleucine is shown with magenta balls-sticks. *D*, the overlayed structures of dimeric SaIlvA_R (green and orange) and monomeric EcIlvA_R•Ile (cyan). The protomer A of SaIlvA_R is superimposed on the ACT1 domain of EcIlvA_R•Ile. The conformational difference of helix H15 is shown as a black arrow.

Figure 8. Modelled quaternary structure of SaIlvA. *A*, tetramer of SaIlvA_R is shown in green (protomer A), orange (protomer B), blue (protomer C), and yellow (protomer D). The N-terminus of each protomer is shown as a sphere. The N-termini of protomers A and C align in the same direction, while those of the protomers B and D face the opposite way. *B*, dimer of SaIlvA generated using AlphaFold. The regulatory domains of SaIlvA dimer can be overlayed on protomers A/C and B/D. *C*, SaIlvA tetramer model was generated using AlphaFold model of SaIlvA dimer and SaIlvA_R tetramer crystal structure. The four C-terminal regulatory domains form the core of the tetramer. *D*, the SaIlvA octamer model, which consists of two tetramers, fitted into the electron density determined by small angle X-ray scattering (SAXS) in solution (χ^2 = 1.37). A tetramer model is shown in green, orange, blue, and yellow, and the other tetramer is shown in cyan.

Figure 9. Complementation of Δ*IlvA*. *A*, growth curve of Δ*IlvA* containing empty vector (pCN38) or pSaIlvA or pEcIlvA grown in Complete Defined Medium (CDM) or CDM minus Isoleucine. *B*, fatty acid analysis by Gas Chromatography of Δ*IlvA* containing empty vector (pCN38) or pSaIlvA or pEcIlvA grown in CDM. *C*, fatty acid analysis by Gas Chromatography of Δ*IlvA* containing empty vector (pCN38) or pSaIlvA or pEcIlvA grown in CDM minus Isoleucine. *D*, growth curve of Δ*IlvA* containing pEcIlvA (F352A) mutant grown in Complete Defined Medium (CDM) or CDM minus Isoleucine. *E*, fatty acid analysis by Gas Chromatography of Δ*IlvA* pEcIlvA (F352A) mutant grown in CDM. *F*, fatty acid analysis by Gas Chromatography of Δ*IlvA* pEcIlvA (F352A) mutant grown in CDM minus Isoleucine. The data in *A* and *D* are representative of experiments done twice. The data in *B*, *C*, *E*, and *F* are from triplicate sets (mean ± SE). The value for the t-test is given in red.

Figure 10. Allosteric regulation of isoleucine on EcIlvA. Loop interactions between the catalytic and regulatory domain in the active site of EcIlvA. The catalytic domain of protomer 1, regulatory domain of protomer 2 and necks of both protomers of EcIlvA are shown in purple, yellow, and blue, respectively. The EcIlvA_R•Ile structure (cyan cartoon) is superimposed on the EcIlvA structure. The features of protomer 2 are indicated with a prime. The docked threonine molecule is represented by ball-and-sticks with gray carbon atoms, blue nitrogen atom, and red oxygen atoms. PLP and the active-site lysine (K62) are shown as sticks. The red arrow indicates the distance from isoleucine to L20 loop, while the black arrow indicates the distance from isoleucine to PLP of the partner protomer.

Figure 1.

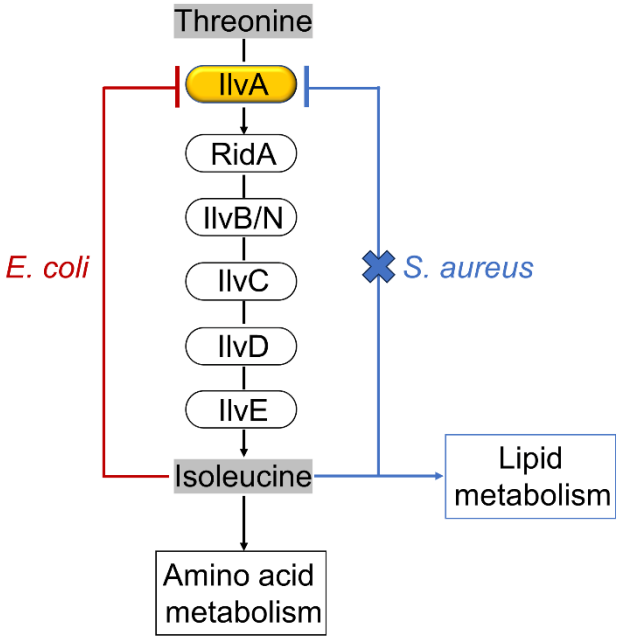

Threonine
IlvA
RidA
IlvB/N
IlvC
IlvD
IlvE
Isoleucine
E. coli
S. aureus
Lipid metabolism
Amino acid metabolism

Figure 2.

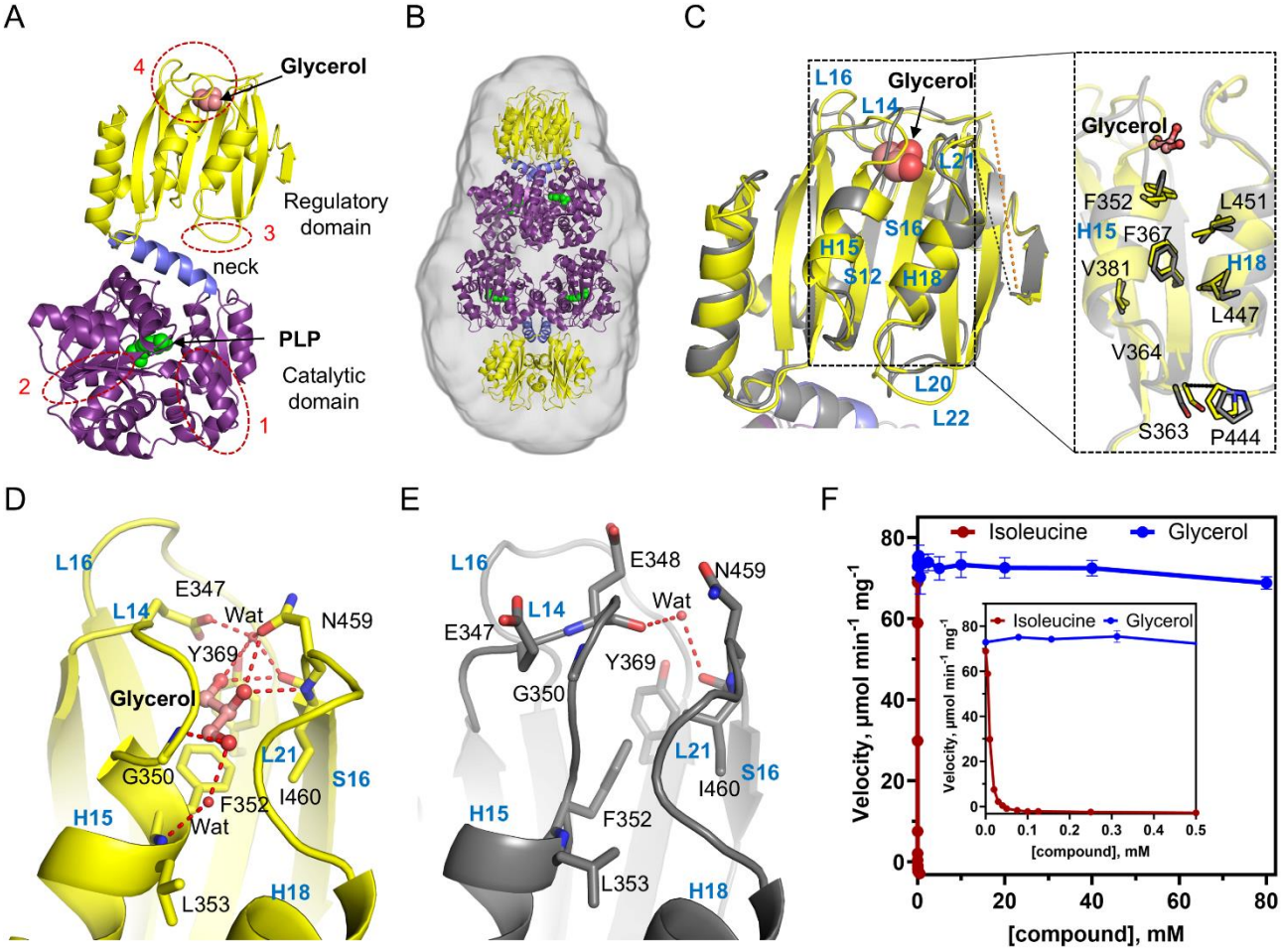

A
Glycerol
4
Regulatory domain
3
neck
PLP
Catalytic domain
2
1
B
C
L16
Glycerol
L14
L21
S16
H15
S12
H18
L20
L22
Glycerol
F352
L451
H15
F367
V381
H18
L447
V364
S363
P444
D
L16
E347
L14
Wat
N459
Y369
Glycerol
G350
L21
S16
I460
F352
H15
Wat
L353
H18
E
L16
E348
N459
L14
Wat
E347
Y369
G350
L21
S16
I460
H15
F352
L353
H18
F
Isoleucine
Glycerol
Velocity, µmol min-1 mg-1
[compound], mM
0
20
40
60
80
0.0
0.1
0.2
0.3
0.4
0.5

Figure 3.

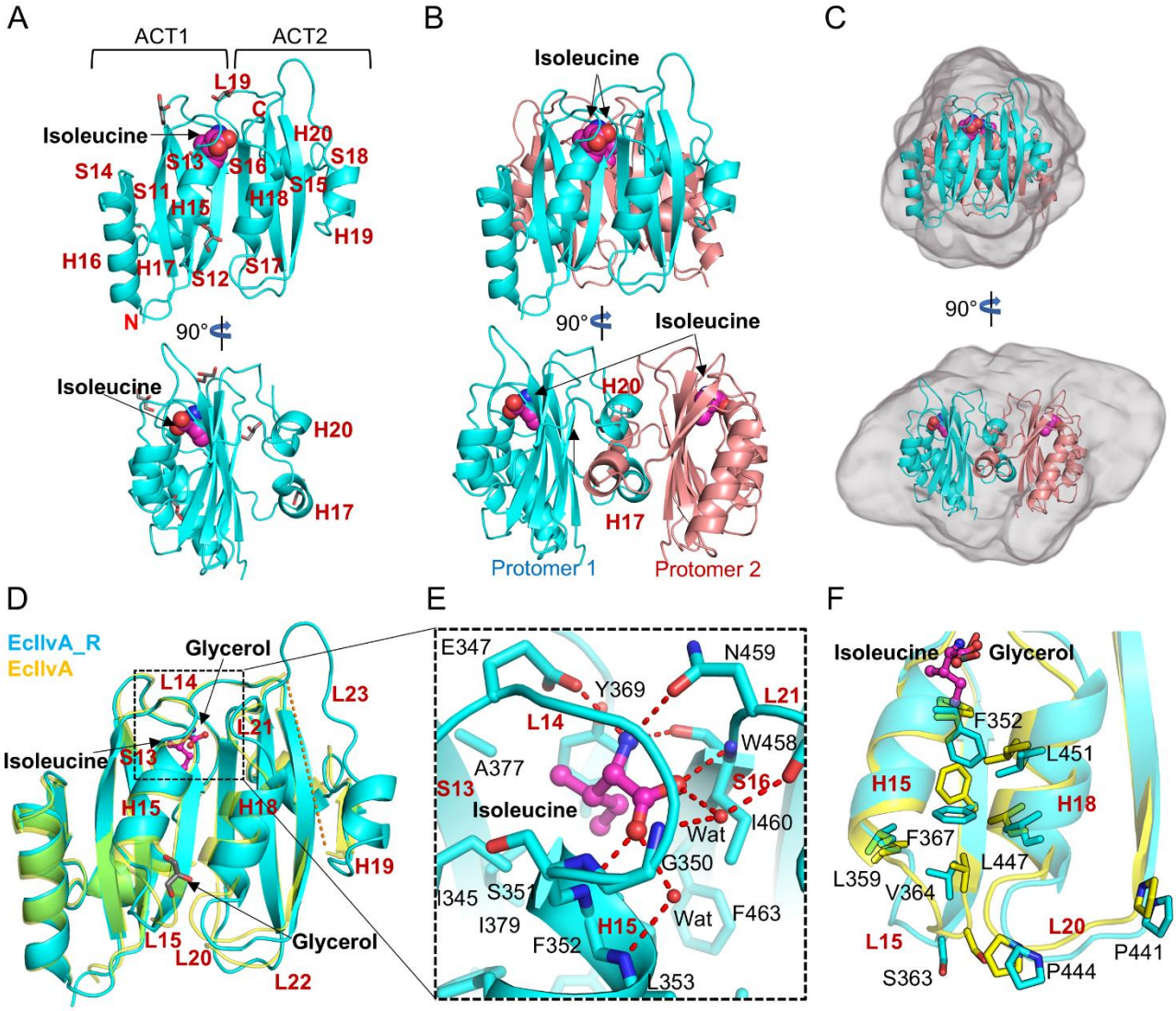

A
ACT1
ACT2
L19
C
Isoleucine
H20
S18
S14
S13
S16
S15
S11
H15
H18
H19
H16
H17
S12
S17
N
90°
Isoleucine
H20
H17
B
Isoleucine
90°
Isoleucine
H20
H17
Protomer 1
Protomer 2
C
90°
D
EcIlvA_R
EcIlvA
Glycerol
L14
L23
L21
Isoleucine
S13
H15
H18
H19
L15
L20
Glycerol
L22
E
E347
N459
Y369
L21
L14
W458
A377
S13
S16
Isoleucine
I460
Wat
G350
I345
S351
Wat
F463
I379
H15
F352
L353
F
Isoleucine
Glycerol
F352
L451
H15
H18
F367
L447
L359
V364
L20
L15
P441
S363
P444

Figure 4.

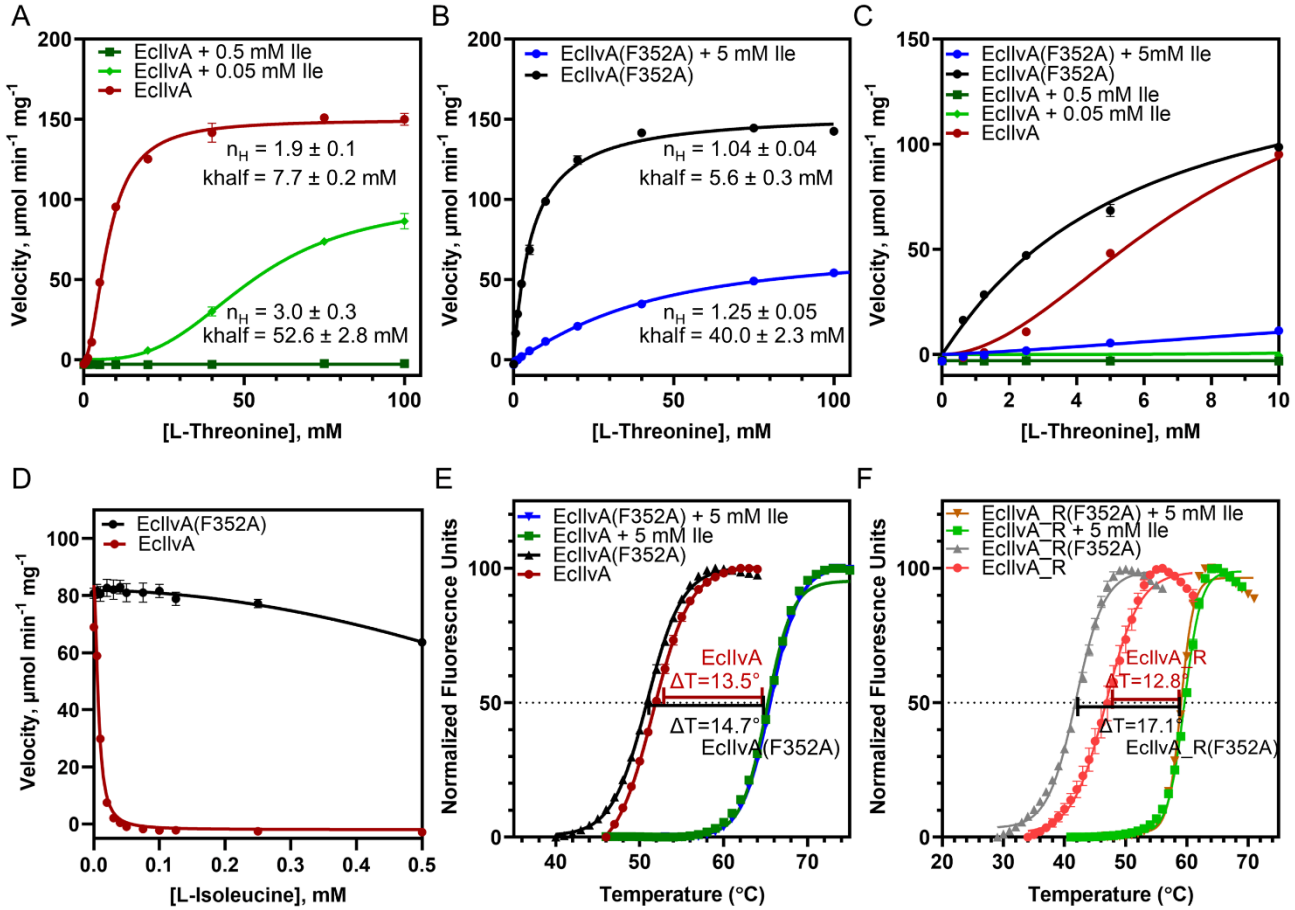

A
EcIlvA + 0.5 mM Ile
EcIlvA + 0.05 mM Ile
EcIlvA
n_H = 1.9 ± 0.1
khalf = 7.7 ± 0.2 mM
n_H = 3.0 ± 0.3
khalf = 52.6 ± 2.8 mM
Velocity, µmol min^{-1} mg^{-1}
[L-Threonine], mM
B
EcIlvA(F352A) + 5 mM Ile
EcIlvA(F352A)
n_H = 1.04 ± 0.04
khalf = 5.6 ± 0.3 mM
n_H = 1.25 ± 0.05
khalf = 40.0 ± 2.3 mM
Velocity, µmol min^{-1} mg^{-1}
[L-Threonine], mM
C
EcIlvA(F352A) + 5mM Ile
EcIlvA(F352A)
EcIlvA + 0.5 mM Ile
EcIlvA + 0.05 mM Ile
EcIlvA
Velocity, µmol min^{-1} mg^{-1}
[L-Threonine], mM
D
EcIlvA(F352A)
EcIlvA
Velocity, µmol min^{-1} mg^{-1}
[L-Isoleucine], mM
E
EcIlvA(F352A) + 5 mM Ile
EcIlvA + 5 mM Ile
EcIlvA(F352A)
EcIlvA
EcIlvA
ΔT=13.5°
ΔT=14.7°
EcIlvA(F352A)
Normalized Fluorescnce Units
Temperature (°C)
F
EcIlvA_R(F352A) + 5 mM Ile
EcIlvA_R + 5 mM Ile
EcIlvA_R(F352A)
EcIlvA_R
EcIlvA_R
ΔT=12.8°
ΔT=17.1°
EcIlvA_R(F352A)
Normalized Fluorescnce Units
Temperature (°C)

Figure 5.

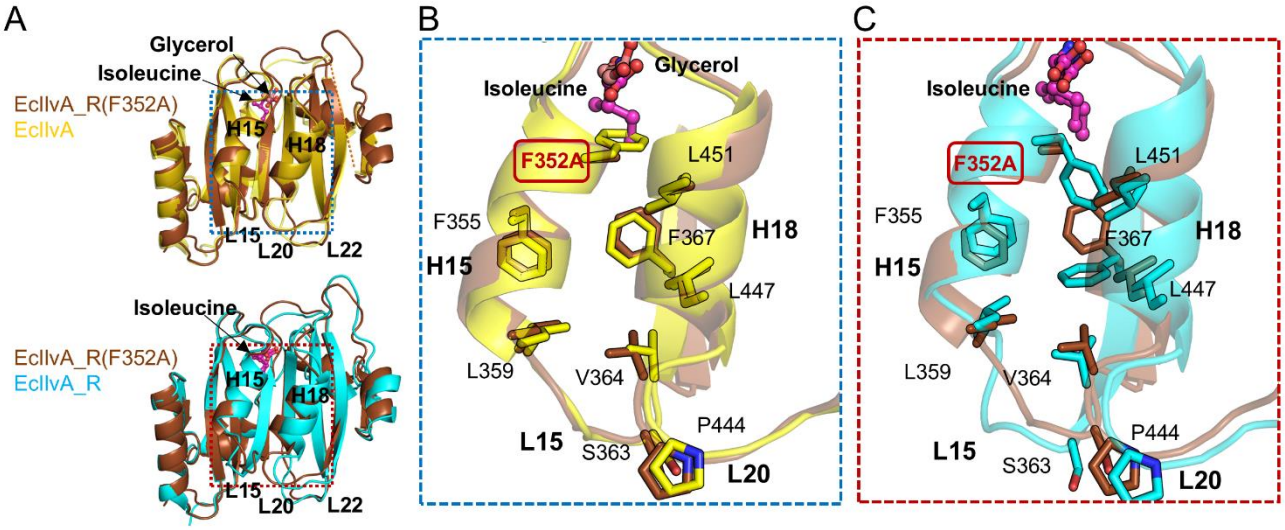

A
Glycerol
Isoleucine
EclIvA_R(F352A)
EclIvA
H15
H18
L15
L20
L22
Isoleucine
EclIvA_R(F352A)
EclIvA_R
H15
H18
L15
L20
L22
B
Isoleucine
Glycerol
F352A
L451
F355
F367
H18
H15
L447
L359
V364
P444
L15
S363
L20
C
Isoleucine
F352A
L451
F355
F367
H18
H15
L447
L359
V364
P444
L15
S363
L20

Figure 6.

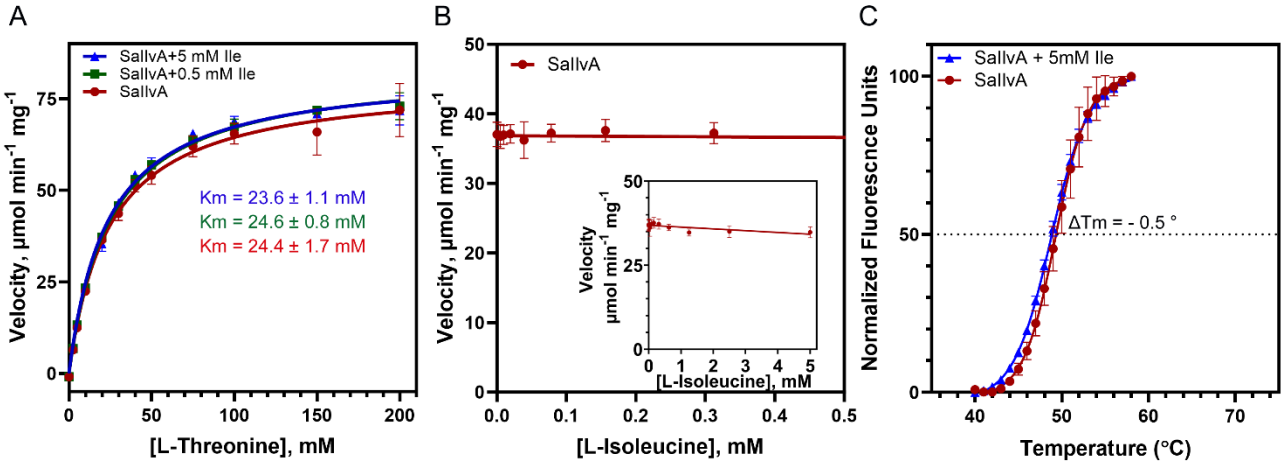

A
SalIvA+5 mM Ile
SalIvA+0.5 mM Ile
SalIvA
Km = 23.6 ± 1.1 mM
Km = 24.6 ± 0.8 mM
Km = 24.4 ± 1.7 mM
Velocity, µmol min-1 mg-1
[L-Threonine], mM
B
SalIvA
Velocity, µmol min-1 mg-1
[L-Isoleucine], mM
Velocity µmol min-1 mg-1
[L-Isoleucine], mM
C
SalIvA + 5mM Ile
SalIvA
ΔTm = - 0.5 °
Normalized Fluorescnce Units
Temperature (°C)

Figure 7.

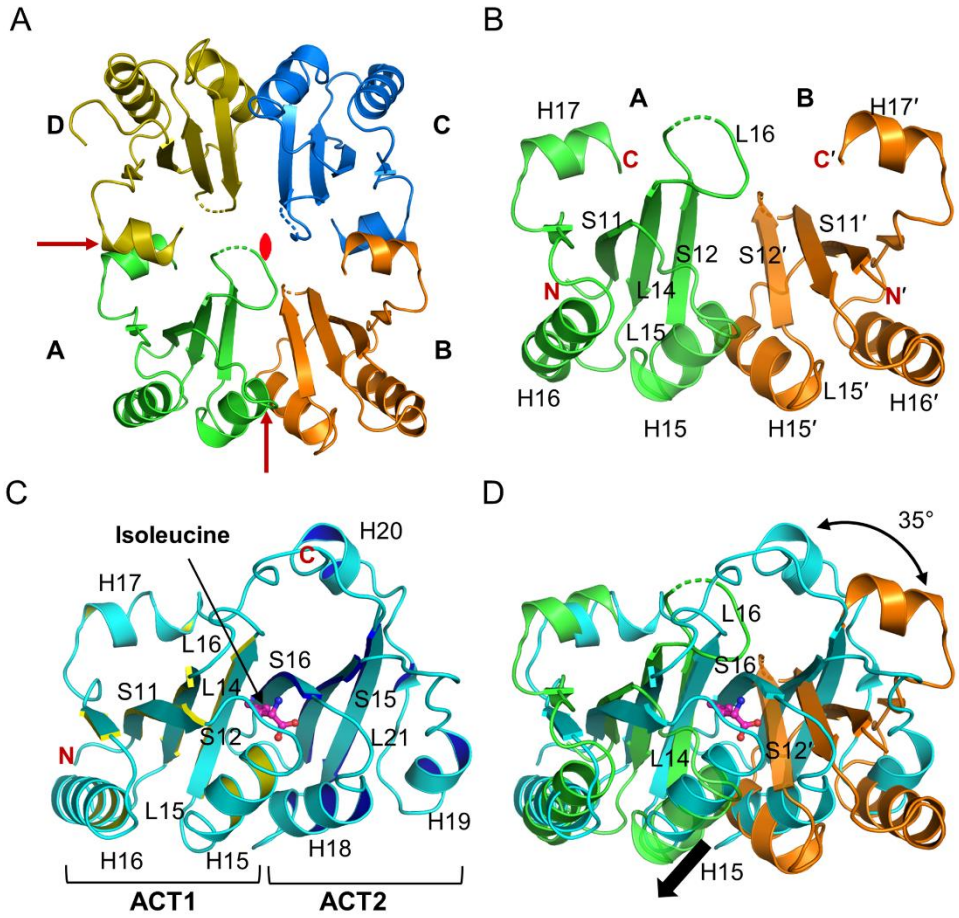

A
D
C
A
B
B
H17
A
B
H17′
L16
C
C′
S11
S12
S12′
S11′
N
L14
N′
L15
H16
L15′
H16′
H15
H15′
C
Isoleucine
H20
C
H17
L16
S16
S11
L14
S15
S12
L21
N
L15
H19
H16
H15
H18
ACT1
ACT2
D
35°
L16
S16
L14
S12′
H15

Figure 8.

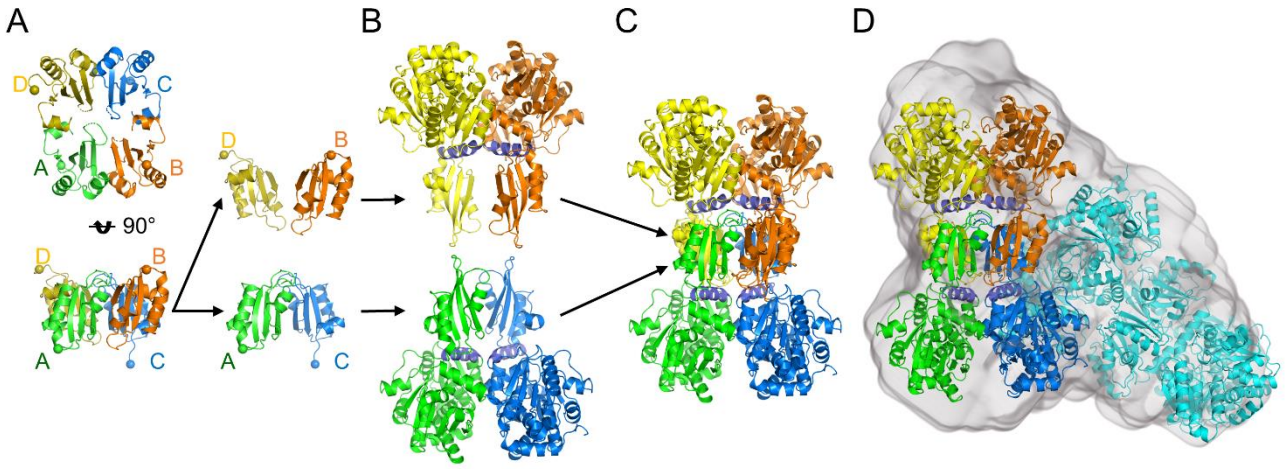


Figure 9.

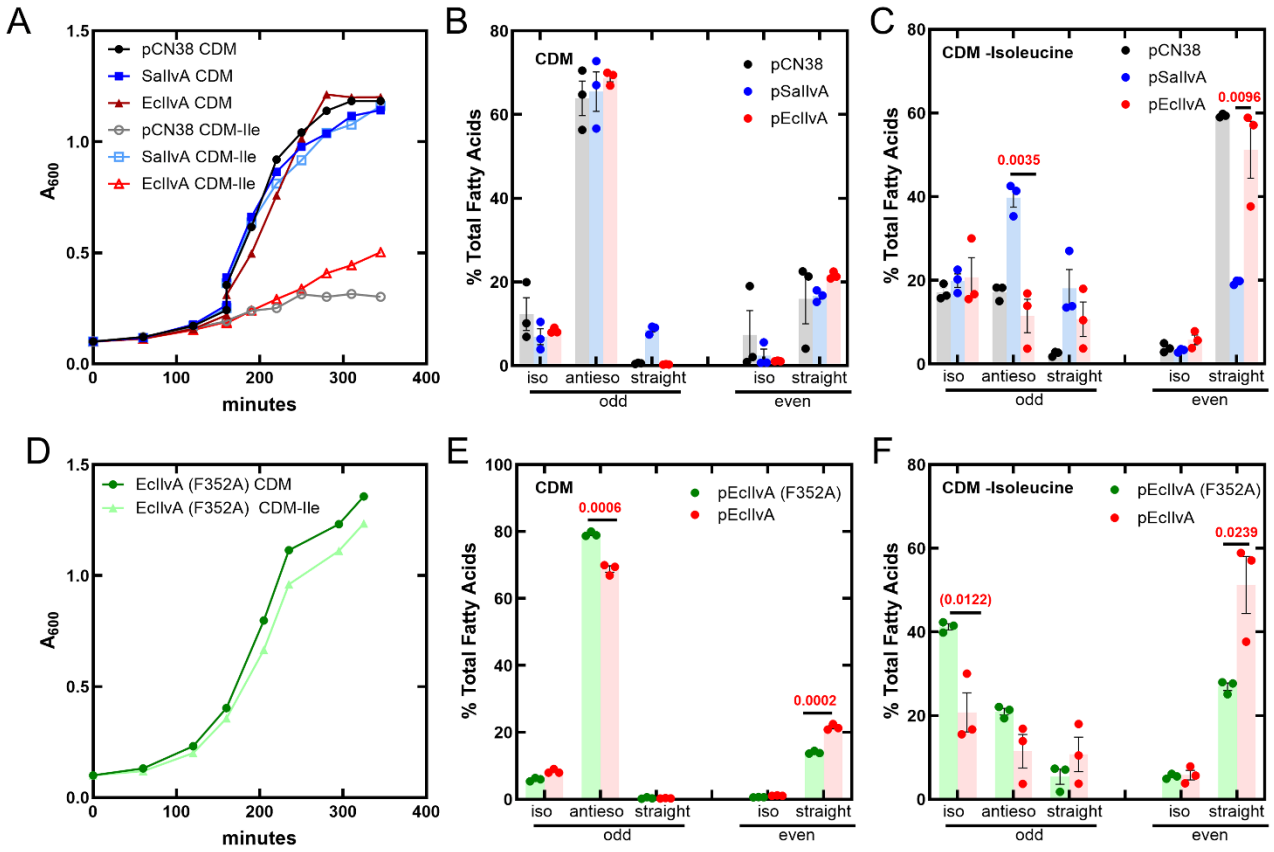
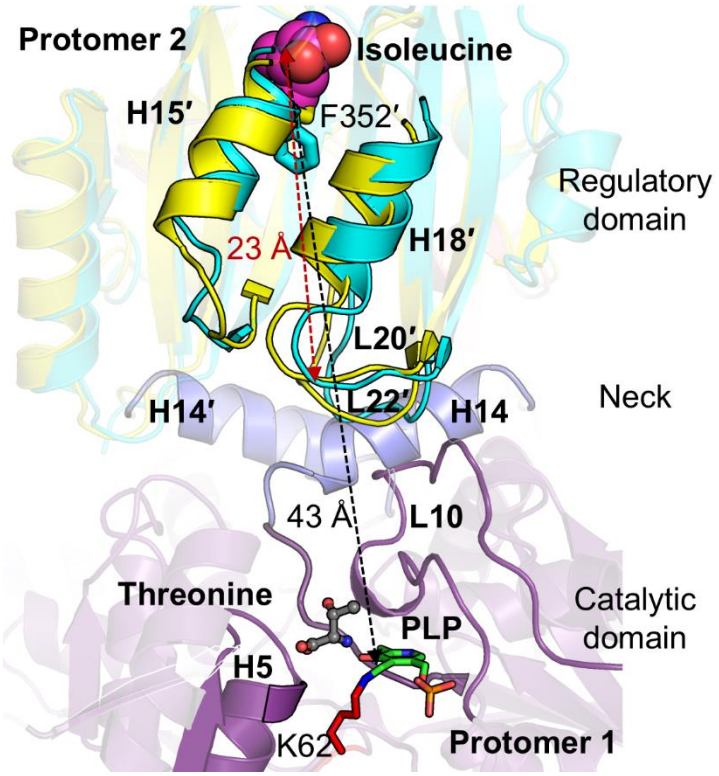

A
pCN38 CDM
SalIvA CDM
EclIvA CDM
pCN38 CDM-Ile
SalIvA CDM-Ile
EclIvA CDM-Ile
A600
minutes
B
CDM
pCN38
pSalIvA
pEclIvA
% Total Fatty Acids
iso
antieso
straight
odd
iso
straight
even
C
CDM -Isoleucine
pCN38
pSalIvA
pEclIvA
0.0035
0.0096
D
EclIvA (F352A) CDM
EclIvA (F352A) CDM-Ile
E
CDM
pEclIvA (F352A)
pEclIvA
0.0006
0.0002
F
CDM -Isoleucine
(0.0122)
0.0239

Figure 10.


Protomer 2
Isoleucine
H15′
F352′
Regulatory domain
23 Å
H18′
L20′
H14′
L22′
H14
Neck
43 Å
L10
Threonine
PLP
Catalytic domain
H5
K62
Protomer 1