ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/49810510

Identification of biochemical and putative biological role of a xenolog from Escherichia coli using structural analysis

ARTICLE *in* PROTEINS STRUCTURE FUNCTION AND BIOINFORMATICS · APRIL 2011

Impact Factor: 2.63 · DOI: 10.1002/prot.22949 · Source: PubMed

READS

54

6 AUTHORS, INCLUDING:

Varun Bhaskar

Max Planck Institute of Biochemistry

7 PUBLICATIONS **22** CITATIONS

SEE PROFILE

Manoj Kumar

NBA SCHOOL OF BUSINESS NEW DELHI

121 PUBLICATIONS **2,077** CITATIONS

SEE PROFILE

Sunil Kumar Tripathi

Regional Centre for Biotechnology

19 PUBLICATIONS **115** CITATIONS

SEE PROFILE

Sankaran Krishnaswamy

Madurai Kamaraj University

91 PUBLICATIONS **1,648** CITATIONS

SEE PROFILE

Available from: Sankaran Krishnaswamy

Retrieved on: 11 January 2016

PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

Identification of biochemical and putative biological role of a xenolog from *Escherichia coli* using structural analysis

Varun Bhaskar, Manoj Kumar, Sankar Manicka, Sunil Tripathi, Aparna Venkatraman, and S. Krishnaswamy*

Department of Genetic Engineering, School of Biotechnology, Madurai Kamaraj University, Madurai 625021, India

ABSTRACT

YagE is a 33 kDa prophage protein encoded by CP4-6 prophage element in *Escherichia coli* K12 genome. Here, we report the structures of YagE complexes with pyruvate (PDB Id 3N2X) and KDGal (2-keto-3-deoxy galactonate) (PDB Id 3NEV) at 2.2Å resolution. Pyruvate depletion assay in presence of glyceraldehyde shows that YagE catalyses the aldol condensation of pyruvate and glyceraldehyde. Our results indicate that the biochemical function of YagE is that of a 2-keto-3-deoxy gluconate (KDG) aldolase. Interestingly, *E. coli* K12 genome lacks an intrinsic KDG aldolase. Moreover, the over-expression of YagE increases cell viability in the presence of certain bactericidal antibiotics, indicating a putative biological role of YagE as a prophage encoded virulence factor enabling the survival of bacteria in the presence of certain antibiotics. The analysis implies a possible mechanism of antibiotic resistance conferred by the over-expression of prophage encoded YagE to *E. coli*.

Proteins 2011; 79:1132–1142.
© 2010 Wiley-Liss, Inc.

Key words: yagE; prophage CP4-6; 2-keto-3-deoxy galaconate; KDG aldolase; pyruvate; Schiff's base; antibiotic resistance; TCA cycle.

INTRODUCTION

Horizontal transfer is one of the reasons for development of drug resistance and pathogenicity among prokaryotes.[1] During infections, phages act as vectors for horizontal gene transfer and thus are one of the causative agents of rapid transmission of drug resistance from one species of bacterium to another. Many of the virulence factors and resistance genes of bacteria are encoded by phage remnants present in the bacterial genome.[1] These genes evolve to acquire novel activities while retaining their original scaffolds.

Divergent evolution is one of the most common mechanisms by which nature introduces and maintains variation. In proteins this is done by introducing subtle variations in the active site while retaining the overall structural framework. A classical example of this can be seen in the TIM barrel proteins, where even a single amino acid substitution in the loops that form the active site can change the activity of the protein.[2] *N*-Acetylneuraminic acid lyase (NAL) superfamily of proteins has been found to share a common scaffold whereas coding for a variety of enzymes.[3] NAL super family of proteins catalyze six different kinds of reactions, which include *N*-Acetylneuraminic acid lyase (NAL), dihydrodipicolinate synthase (DHDPS), 2-keta-3-deoxy gluconate aldolase or 2-keta-3-deoxy galactonate aldolase (KDGA), 5-keto-4-deoxyglucarate dehydratase (KDGDH), trans-*o*-hydroxybenzylidine pyruvate hydratase aldolase (HBPHA), trans-20-carboxybenzalpyruvate hydratase-aldolase (CBPHA), and 4-deoxy-5-oxoglucarate dehydratase (DOGDH).[3] There is an underlying unity among these six families of proteins considering the structural framework and the mechanism they employ to catalyze the reaction. All the enzymes classified under the NAL subfamily bind to a common substrate, that is, pyruvate and undergo a series of changes including a Schiff's base formation between the carbonyl oxygen of the pyruvate moiety and the NZ atom of a lysine residue from β6 which is at the base of the active site.[4,5] This is followed by a proton abstraction from the hydroxyl group of tyrosine and a proton relay, thus resulting in catalysis.

On the basis of the comparison of the active sites of each subfamily, the residues occurring in the active site have been classified into primary and secondary active site residues.[3,6] The primary active site residues include a crucial lysine residue that is involved in pyruvate binding and Schiff's base formation and a tyrosine residue that helps in proton donation to the substrate. The two other residues that can be classified as part of the primary active site residues are a threonine or a

Additional Supporting Information may be found in the online version of this article.

*Correspondence to: S. Krishnaswamy, Department of Genetic Engineering, School of Biotechnology, Madurai Kamaraj University, Madurai 625021, India. E-mail: krishna@mrna.tn.nic.in
Received 12 September 2010; Revised 10 November 2010; Accepted 19 November 2010
Published online 29 November 2010 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/prot.22949

© 2010 WILEY-LISS, INC.

serine and a tyrosine (from the adjacent monomer) that is involved in the proton relay and the final transfer of proton to the solvent, respectively.[7] Mutation of any of these residues has been shown to hamper the activity severely in both NAL and DHDPS.[7] The secondary residues in the active site are generally involved in substrate binding and are not involved directly in catalysis.

As a Structural Genomics Initiative, our group started studying the proteins encoded by phage remnants. YagE is a prophage protein from the *Escherichia coli* K12 genome and is encoded in the CP4-6 prophage gene cluster.[8] Sequence analysis of YagE shows that it is related to the NAL superfamily of enzymes. The structure of the apo form of YagE, reported by our group confirmed its close resemblance to the NAL superfamily of proteins with a typical TIM barrel fold of $(\alpha/\beta)_8$ with three α-helices at the C-terminus.[9] In continuation of this work, the structure of YagE bound to pyruvate is reported here. On the basis of docking experiments and enzyme assays, we have assigned a possible biochemical role to YagE. We have determined the structure of YagE in presence of pyruvate and glyceraldehyde to a resolution of 2.2 Å. The active site shows the presence of KDGal which clearly suggests YagE is indeed a KDG aldolase. *E. coli* K12 genome lacks an inherent KDG aldolase. Surprisingly, over-expression of YagE was shown to result in increased cell viability in presence of bactericidal antibiotics like Norflaxacin, Ampicillin, and Streptomycin. A putative biological role of this protein as a potential virulence factor which on over-expression can enhance the survival of the bacterium in the presence of certain antibiotics is deduced

MATERIALS AND METHODS

Cloning, expression, and purification

The initial clone of YagE gene was obtained from Genobase, Tokyo, in a pCA24N vector. YagE was subcloned into pET28a using *Kpn*I and *Not*I sites and the native YagE was expressed and purified as described earlier.[9] The 6XHis tag was not removed for crystallization. After gel filtration, the protein was dialyzed against a buffer containing 50 m*M* Tris-HCl pH 7.5, 50 m*M* sodium chloride, 10% w/v glycerol, 1 m*M* DTT, and 1 m*M* EDTA. The protein was concentrated to 3 mg/mL.

Crystallization, soaking, and data collection

Diffraction quality crystals were obtained in 100 m*M* HEPES pH 6.5, 200 m*M* magnesium chloride, and 15% PEG 3350[9] using micro batch method after 48 h incubation at 20°C. Each crystallization droplet contains 5 mg/mL of protein. For YagE-pyruvate complex, crystals were soaked in cryo-solution containing 15% PEG 3350, 200 m*M* $MgCl_2$, 100 m*M* HEPES pH 6.5, 25% ethylene glycol, and 50 m*M* of pyruvate for 5 min before freezing on the cryo stream. Complete X-ray diffraction data were collected using the soaked crystal at BM-14 beam line at ESRF, Grenoble, France.

Cocrystallization of YagE with pyruvate and glyceraldehyde was done in 100 m*M* HEPES pH 6.5, 200 m*M* magnesium chloride, 15% PEG 3350, 50 m*M* Pyruvate, and 50 m*M* glyceraldehyde. For these crystals, 15% PEG 3350, 200 m*M* $MgCl_2$, 100 m*M* HEPES pH 6.5, 25% ethylene glycol, and 50 m*M* of pyruvate, and 50 m*M* glyceraldehyde was used as the cryo solvent and the crystals were soaked for 5 min before freezing on the cryo stream. X-ray diffraction data were collected using the soaked crystal at BM-14 beam line at ESRF, Grenoble, France.

Data processing and structure solution

XDS was used for processing the YagE-Pyruvate data.[10] The data were truncated at 2.2 Å and the cell parameters were refined at this stage. The R_{merge} and R_{meas} at this resolution shell were 20.7% and 32%, respectively. For YagE cocrystallized with pyruvate and glyceraldehyde, the data to a resolution of 2.2 Å was used and the cell parameters were refined. The R_{merge} and R_{meas} for the highest resolution shell of this data were 21.8% and 39.8%, respectively. The data were scaled using XSCALE. The resulting structure factor amplitudes were used for molecular replacement and refinement.

Structure of YagE-Pyruvate was determined by molecular replacement using the apo-YagE as the phasing model in the PHASER program. The MR solution from PHASER[11] was later refined using Refmac5[12,13] and Phenix.[14] COOT was used for manual inspection and re-building.[15] NCS restrain was also imposed during the refinement. Water molecules were added during refmac5 refinement using aRP water option in ccp4 suite[16] and manually checked using COOT.

Docking

Docking was done using GLIDE module in Schrodinger,[17,18] GOLD,[19] and AUTODOCK softwares.[20] Protein Preparation, Ligand Preparation, and Structure Minimization were done in the respective modules, except for GOLD. For GOLD, Ligand and Proteins were minimized in Glide and was then used for docking in GOLD. Sialic acid alditol and 2-keto-3-deoxy gluconate (KDG) were docked on apo YagE while L-aspartate semialdehyde (L-ASA) was docked on pyruvate bound YagE structure. HINAL docked with alditol and KDG with *SsKDGA*,[21] were used as controls. The docking results were screened based on the conservation of key interactions that were seen in crystal structure of H1NAL with sialic acid alditol and KDGA with KDG. Since L-ASA binding is seen only after pyruvate binding during

dihydrodipicolinate synthesis, the DHDPS structure from *T.maritima* (PDB Id 1O5K) bound with pyruvate was used as the template for docking L-ASA.[22] Of the many possible conformations generated for the ligands, only those that retained chirality as seen in other crystal structures were used for docking. Structural visualizations and distance measurements were done by Chimera and Pymol.[23,24]

Enzyme assay for NAL, DHDPS, and KDGA activities

NAL activity was monitored using spectrophotometric observation of decrease in absorbance of NADH at 340 nm. The assay was performed in a final volume of 1 mL at both 25°C and 37°C. The substrate *N*-acetyl neuraminic acid is cleaved by NAL into pyruvate and *N*-acetyl mannosamine. The resulting pyruvate is reduced to lactic acid using a hydride from NADH, abstracted by lactate dehydrogenase (LDH). The NAL assay was performed as described in Comb *et al.*[25] The assay cocktail contained 20 m*M* sodium phosphate pH 7.2, 0.15 m*M* NADH, 30 μg of rabbit lactate dehydrogenase, and 2 m*M* Neu5Ac, and varying concentrations of YagE ranging from 0.6 μ*M* to 3 μ*M* while 0.6 μ*M* of NAL was used as a positive control.

The spectrophotometric assay for KDG aldolase activity was designed to estimate the amount of pyruvate consumed in the presence of excess glyceraldehyde after incubating YagE with both pyruvate and glyceraldehyde at 37°C.[26] The NADH-LDH assay was used to estimate the amount of unreacted pyruvate. The reaction mixture contained 100 m*M* Tris-HCl pH 7.5, 0.16 m*M* NADH, 30 μg of rabbit lactate dehydrogenase, 10 m*M* pyruvate, 20 m*M* glyceraldehyde, and 3 μ*M* of YagE.

The DHDPS activity was estimated similar to that of KDGA, the only difference being that 10 m*M* of L-ASA was used instead of 20 m*M* glyceraldehyde. A spectrophotometric assay to monitor the spontaneous formation of dipicolinic acid, which will be reflected in the increase in absorbance at 270 nm when pyruvate and L-ASA are incubated with the enzyme in the presence of imidazole buffer at pH 7.4 was also performed. The 1 mL assay mixture contained final concentrations of 100 m*M* imidazole buffer, pH 7.4, 5 m*M* sodium pyruvate, 2 m*M* aspartate semialdehyde, and various concentrations of YagE ranging from 0.6 to 2 μ*M*. The assay mixture was incubated at 37°C and monitored at fixed time intervals. The assay was done as described by Yugari *et al.*[27,28]

Antibiotic resistance

E. coli BL21 (DE3) containing pET28a: YagE was inoculated in LB media containing 50 μg/mL kanamycin and grown to an OD_{600} of 0.2. IPTG (0.1 m*M*) was used to induce the culture at an OD of 0.2 at 25°C. The culture was grown at 25°C till the OD of the culture reached 0.4. 1 μg/mL Norfloxacin (dissolved in DMSO) or 100 μg/mL Ampicillin or 100 μg/mL was introduced in the respective culture. Then, the growth of the cells was monitored for 2 h after the addition of the antibiotic and 100 μL of samples were collected and the cells were washed in PBS. The Colony forming units (CFU) were estimated by plating the respective dilutions of samples. CFU was used as a parameter to check the survival of the bacterium.

RESULTS AND DISCUSSION

Structure of YagE in complex with pyruvate

The structure of YagE bound to pyruvate (PDB Id 3N2X) was determined in P 2_1 2_1 2, orthorhombic space group with cell parameters a = 141.180 Å, b = 155.20 Å, c = 55.72 Å, and $\alpha = \beta = \gamma = 90$. The data collection and refinement summary are shown in Table I. The asymmetric unit consisted of a tetramer formed by a dimer of dimers. The Non Crystallographic Symmetry (NCS) determined using Polarrfn from ccp4 suite showed 222 symmetry which is not uncommon in NAL superfamily of proteins [Fig. 1(A)]. The overall structural fold of this structure was similar to the apoYagE structure with a typical $(\alpha/\beta)_8$ barrel and C-terminal helices. The active sites of apoYagE and YagE-pyruvate complex structures show very similar architecture with minimal change in orientation of residues [Fig. 1(B)]. It has also been noted that the "**b**" axis of YagE-pyruvate complex has a reduction in length by ~ 3 Å when compared to the unit cell dimensions of apo YagE. Analysis of protein dihedral angle by PROCHECK showed that all the residues except tyrosine 119 from all the four monomers were placed in a favorable region of the Ramachandran plot. The YagE monomer bound to pyruvate superposed with Apo YagE shows RMSD of 0.18 Å, while the superposition with NAL, DHDPS, and KDG had RMSDs of 1.08 Å, 1.01 Å, and 1.09 Å, respectively. The two tight dimer interfaces are formed between chains A and B and C and D, respectively. The tight dimer interface buries a total surface area of 1675 $Å^2$ and has eight salt bridges. In contrast to the tight dimer interface, the other interface that is formed by chains A and C and B and D chains is quite loose and buries a total surface area of 960 $Å^2$ and consists of four salt bridges. The interface interaction seen in this structure is very similar to the interactions present in the apo YagE structure.

The electron density map corresponding to lysine 174, which forms the base of the active site showed a continuous density with the ligand's electron density. This led to the conclusion that pyruvate could be covalently linked to the lysine through an imine bond between NZ atom of lysine and carbonyl carbon (C2) of pyruvate forming a Schiff's base, as is already established in NAL super-

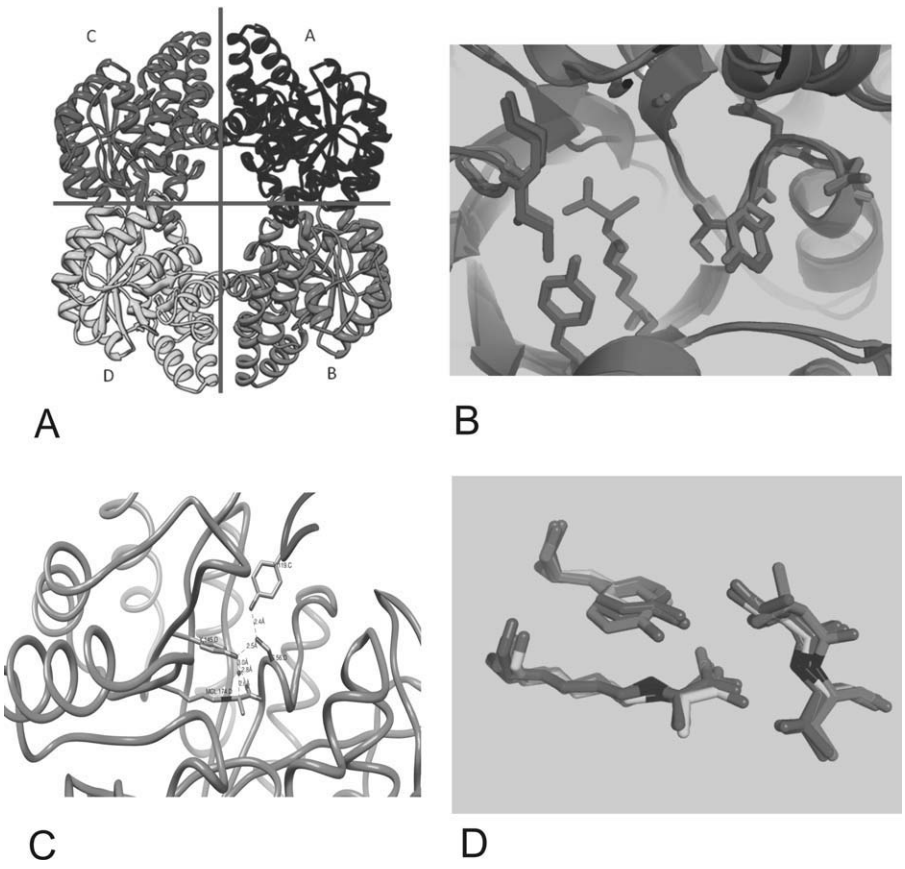

Figure 1

A: YagE pyruvate tetramer with 222 NCS. **B**: Superposition of Apo YagE active site (Dark) on active site of YagE pyruvate complex (Gray). **C**: Catalytic triad of YagE formed by Tyr 145 and Ser 56 of chain A and Tyr 119 of chain B. Figure also shows the Schiff's base formed between Lys 174 and pyruvate moiety and the released water. **D**: Superposition of YagE pyruvate complex monomers active site (dark) with active sites of H1NAL (light gray), Tm DHDPS pyruvate complex (gray), Ss KDGA (white), all these molecules in complex with pyruvate.

family of enzymes.[3,21,22,29] The presence of imine bond was confirmed using $3F_o-2F_c$ weighted difference map and electron density at the imine bond was clearly visible at 1σ. Density for a water molecule was seen at 2.71Å from carbonyl carbon and it is hypothesized that this could be the water that was released during the Schiff's base formation [Fig. 1(C)]. The superposition of YagE-pyruvate complex active site on pyruvate complexes of NAL superfamily enzymes is shown in Figure 1(D).

Previous reports on NAL superfamily of enzymes indicate the presence of a catalytic triad apart from the Schiff's base-forming lysine. This triad is formed by a tyrosine above the lysine residue, serine/threonine from GXXG motif, and a tyrosine residue from the adjacent monomer in the dimer.[7] Such an arrangement is essential for proton abstraction and relay resulting in transport of a proton to (or from) the solvent. The proton abstraction and relay are necessary for both Schiff's base formation and for subsequent aldol condensation/cleavage. The triad in YagE is formed by tyrosine 145, serine 56, and tyrosine 119. The distance between the hydroxyl group of Y145 and Oγ of S56 is 2.39Å, while that between the hydroxyl group of Y119 and Oγ of S56 is 2.53Å [Fig. 1(C)]. Thus, the YagE-pyruvate structure shows that the catalytic triad and the critical lysine residue are in the appropriate configuration. This along with the Schiff's base formation between lysine and pyruvate suggests that YagE can be active.

Common features of NAL superfamily active site in comparison with YagE structures

NAL super family of enzymes possess similar active sites. Fersht *et al.* showed that a single amino acid mutation of conserved leucine 150 to arginine changed the activity of enzyme from NAL to DHDPS.[2] Here, based on sequence alignment and structural superposition, we have compared the active site architecture of YagE with NAL, DHDPS, and KDG family of enzymes. The equivalent residues are summarized in Supporting Information Table I.

NAL, DHDPS, and KDG active sites when compared to YagE at its 57th position, have a threonine instead of a

Table I
Data Collection and Refinement Statistics

Data collection	YagE-pyruvate structure	YagE-KDGal structure
X-Ray source	BM14-4 ESRF	BM14-4 ESRF
Space group	$P2_12_12$	$P2_12_12$
Cell dimensions		
a, b, c (Å)	141.13, 153.55, 55.53	141.80, 155.20, 55.71
Molecules/AU	4	4
Solvent content (%)	41.30	41.86
Wavelength (Å)	0.97	0.95
Resolution (Å)	49−2.2 (2.33−2.20)	49.18−2.2 (2.24−2.20)
R_{merge} (%)	8.8 (20.70)	8.2 (21.80)
I/σ (I)	16.93 (5.90)	14.01 (5.40)
Average *B* value		
Protein atoms	13.6	20.3
Ligand atoms	14.6	31.3
Solvent atoms	18.6	29.0
Reflections	62,116 (9222)	61,463 (8987)
Completeness (%)	96.30 (92.40)	95.80 (87.80)
Refinement		
Resolution (Å)	49−2.2 (2.25−2.20)	104.83−2.2 (2.25−2.20)
Protein atoms	9040	9104
Ligand atoms	28	44
Solvent atoms	511	522
*R_{work}/R_{free} (%)	17.8/22.2 (18.1/24.3)	19.7/24.1 (23.9/28.8)
Deviation from ideality		
Bond length	0.023	0.024
Bond angle	1.866	1.868
Procheck *G* score (overall avg)	0.02	−0.02
PDB Id	3N2X	3NEV

*R_{work} is the measure of agreement between the observed and computed structure factor amplitudes for the set of reflections. $R_{work} = \Sigma_I (||F_{Obs}| - |F_{calc}||)/\Sigma_I |F_{Obs}|$, I = all $h\ k\ l$.
*R_{free} is the measure of agreement between the observed and computed structure factor amplitudes for the set of reflections that is omitted in the refinement process.

glycine. In addition, when compared to NAL, only a single amino acid substitution of Ser 207 to Ala 221 is seen in YagE. YagE lacks the critical arginine residue, which is present in DHDPS that is essential for L-ASA binding. Similarly at sequence level, in contrast to KDGA, YagE possesses a F147 residue in the place of Y132 in KDGA. Thus, it becomes difficult to assign a possible role to YagE just based on the active site comparison, since all these three families of proteins have at most one active site residue difference when compared to YagE, as has also been reported earlier.[9] Docking studies were carried out to try and ascertain the nature of the possible ligand.

Identification of biochemical function

Computational analysis

Since in NAL super family of proteins, the key interactions of the substrate with the primary catalytic residues, which includes the Schiff's base forming lysine, proton donating tyrosine and the GXXG motif are essential, loss of these interactions in the docked ligand implies that this orientation of ligand is probably not appropriate. Apart from this, the interaction of each atom of the ligand with the binding residue was cross checked with the crystal structure.

Pyruvate was docked on to the apo YagE template and was compared with the crystal structure of YagE-pyruvate Complex. The comparison is tabulated in Supporting Information Table II. The docking results show that pyruvate indeed interacts with Lys 174, Tyr 145, and Ser 56, although the main chain interactions are not seen in GLIDE results. Docking in GOLD and AUTODOCK showed that instead of the carbonyl group, carboxyl group of pyruvate interacts with Lys174. The orientation given by the docking results may be hypothesized to indicate the initial binding of pyruvate to the active site in which pyruvate may be drawn into the cavity due to the electrostatic attraction from Lys 174.

Ring opening of a sialic acid is an essential step in case of NAL catalysis.[3] Docking of sialic acid did not produce any biologically relevant conformations on either H1NAL or YagE. Hence, the ring opened sialic acid alditol (SIA alditol), that is, alditol, which is a substrate analogue of sialic acid, was chosen for docking. Ligand preparation using SIA alditol in GLIDE generated 32 possible conformations of the ligand, all of which were used for docking on H1 NAL (PDB ID: 1F7B) and YagE (PDB ID: 2V8Z).[9] Docking of Alditol on H1 NAL was used for validating the docking experiment and was verified with Alditol bound H1NAL structure (PDB ID: 1F73).[3] For AUTODOCK and GOLD also the ligand was docked flexibly but at the same time, chirality of the ligand was retained. The results of the docking trials are tabulated in Supporting Information Table III.

Docking of YagE with SIA alditol using GLIDE, showed that none of the 32 conformations of the substrate docked favorably (Supporting Information Fig. 1). Only 2 out of 32 conformations of the ligand showed interaction of O2 of SIA alditol with Lys 174 NZ as well as the O1A of SIA alditol with Gly 57NH. Even in these two conformations of ligand, the rest of the groups did not show other interactions. Also, the orientation of these two conformations of ligand especially from C3–C9 was almost oppositely placed with respect to the positioning of these groups in the crystal structure. The same was seen with GOLD and AUTODOCK results.

In fact, a closer analysis suggested that the alditol binding pocket itself is not perfectly shaped in case of YagE due to the steric hindrance in the region of Ala 221 and between the residues 260–265 in YagE compared to H1NAL (Supporting Information Fig. 2). Our docking results and structural analysis led us to believe that YagE could not be a potential NAL enzyme since it lacks the proper active site and if it has to be a NAL, it has to undergo a structural rearrangement in presence of its substrate.

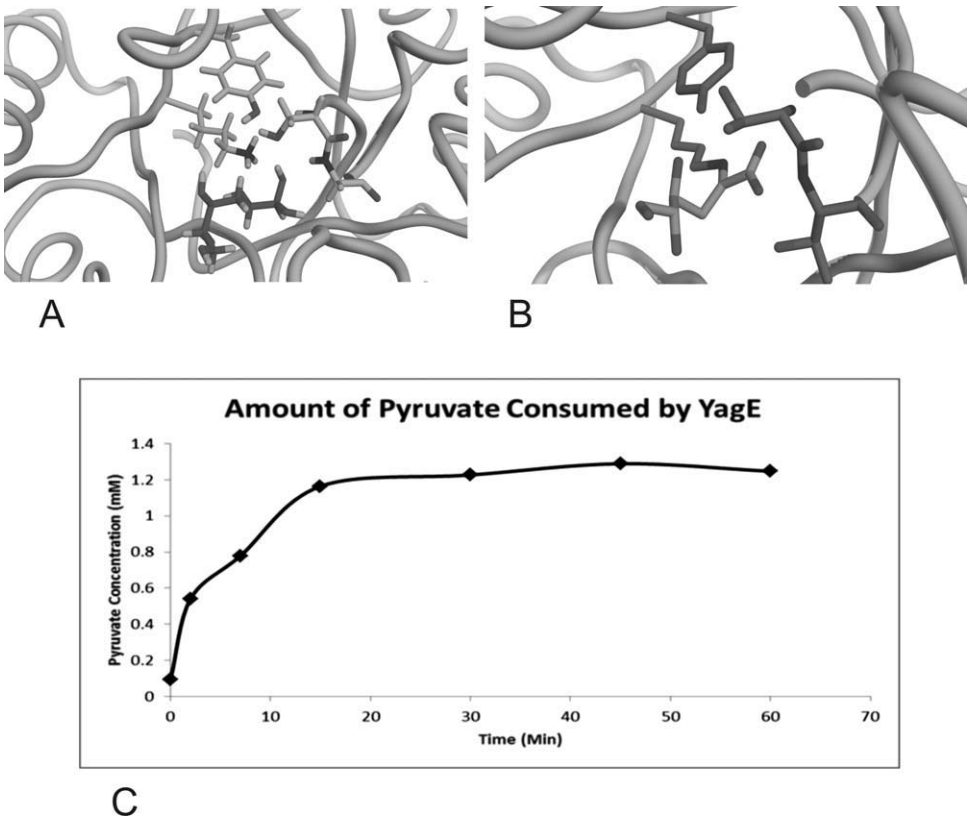

Figure 2

A: Docking of KDG (dark) on Apo YagE structure. The interactions with S56, G57, Y145, and K174 are shown. **B**: Crytal structure of KDG in complex with SsKDGA.[22] Schiffs base formation between K155 and KDG is shown along with interactions with residues T43, T44, and Y130. Ligand orientations in both these active sites are similar. **C**: Graph showing the consumption of pyruvate by YagE at various time intervals. The enzyme reaches a saturation around 15 min.

Since in DHDPS reaction, L-ASA binding is preceded by pyruvate binding and Schiff's base formation,[6] the docking of L-ASA on DHDPS requires a template structure of DHDPS bonded with pyruvate. Such a structure is available from *Thermatoga maritima* (PDB ID: 105K),[22] where the pyruvate is bonded covalently to Lys 161. This experiment was used for validating the docking results and the docking of L-ASA on *Tm*DHDPS did give favorable interactions, as reported in the crystal structure of DHDPS with succinate β-semialdehyde (SAS-analogue of L-ASA, lacking amino group).[6] The chirality of L-ASA was checked after docking and only results having the S-ASA conformation were analyzed.

On the contrary, docking of YagE (bound to pyruvate) with L-ASA showed that in none of the conformations L-ASA was able to enter the cavity [Supporting Information Fig. 2(A)]. All the ligands lacked the key interaction of carbonyl oxygen with Tyr 145 OH. The results are summarized in Supporting Information Table IVA. Since YagE lacks the conserved Arg, we tried to dock the mutated YagE, that is, L150R, with L-ASA. Mutation was done in COOT. Docking of L-ASA on mutated version of YagE imine did not give the correct orientation of ligand. [Supporting Information, Fig. 2(B)]. The docking results on YagE also failed to show the Tyr 145 interaction which is essential for catalysis. These results are summarized in Supporting Information Table IVB.

Docking of KDG was done with apoYagE as the template. Since stereo chemical isomerism will result in change of sugar, the chirality of KDG was retained all through the docking. KDGA without any bound ligand was used as the template for validation of docking. In case of KDGA, a few interactions were water-mediated.[21] Therefore, docking was done in the presence of these ordered water molecules.

The comparison of interactions that were found with docked ligand on YagE and the ones observed in crystal structure of KDG-KDGA complex are shown in Supporting Information Table V.[21] The results show that KDG has been docked favorably in YagE [Fig. 2(A)]. Almost all the critical interactions of KDG with the primary residues are preserved. Lack of interactions of O4 with Tyr 130 (Tyr 145 in case of YagE) and Thr 157 is reasoned for lack of Schiff's base formation in docking, while the

structure has been determined after Schiff's base formation [Fig. 2(B)]. In case of YagE also Tyr 145 and Thr 176 failed to interact with O4 of KDG, which was again reasoned for lack of Schiff's base formation. The structure of YagE pyruvate complex shows the water which is hypothesized as the water released due to Schiff's base formation, occupies the same position as that of O4 of KDG that could possibly interact with Thr 176 OG and Tyr 145 OH. Thus, the argument that the interaction of O4 of KDG with Thr176 OG and Tyr 145 OH will be seen only after Schiff's base formation is further strengthened. Also, presence of a glycine 57 in YagE instead of a threonine (corresponding residue at 44 is a threonine in SsKDGA) justifies the lack of water-mediated interaction with O6 which has been observed with Thr44 Oγ in KDGA. Since Tyr 203 has an approximate distance of 5 Å with O6 before Schiff's base formation, it may show similar interactions as such with Tyr 132 in KDGA after Schiff's base formation. Hence, in the case of YagE, Tyr 203 may play the same role as Tyr 132 residue in KDGA as mentioned in case of SsKDGA.[21] Thus O6 of KDG may solely interact with Tyr 203. The computational analysis outlined suggested that, YagE is probably a KDG aldolase.

Enzymatic assay

The activity of YagE was quantified as 77.63 μ*M* pyruvate being consumed/min under the given conditions using pyruvate standard. The pyruvate consumption activity of YagE at various time intervals shows that the reaction reaches saturation at around 15 min [Fig. 2(C)].

In case of NAL assay, decrease in the optical density of NADH was not observed even after varying the parameters of assay, like temperature and enzyme concentrations indicating that the sialic acid is not being cleaved by YagE. Since at comparable substrate and enzyme concentrations, NAL control showed activity, this shows that YagE does not have a NAL activity. A similar situation was seen for DHDPS assay. Therefore, we conclude that YagE has a KDG Aldolase like activity.

The crystal structure obtained from YagE treated with pyruvate and glyceraldehyde, both in the crystallization solution and in the crystalline state, showed the KDGal moiety. This confirmed the KDG Aldolase activity of YagE.

Structure of YagE in complex with KDGal

The structure of YagE bound to KDGal (PDB Id 3NEV) was determined in P 2(1) 2(1) 2, orthorhombic space group with a cell parameters a = 141.8 Å, b = 155.202 Å, c = 55.721 Å, and α = β = γ = 90. The data collection and refinement summary are shown in Table I. The asymmetric unit consisted of a tetramer formed by dimer of dimers similar to YagE-Pyruvate structure. A 222 NCS was seen in YagE-KDGal structure. When comparing the cell dimension of apo YagE and YagE Pyruvate complex with the YagE KDGal complex, it is found that with respect to apo YagE structure the "**b**" axis has a reduction in length by ∼ 1.1Å, while compared YagE Pyruvate structure the "**b**" axis increases by ∼ 1.7Å. Analysis of protein dihedral angle by PROCHECK showed that all the residues except Tyrosine 119 from all the four monomers were placed in favored region of Ramachandran plot. The overall structural fold of this structure was similar to the previously mentioned YagE structures. Two dimer interfaces are formed between chains A and B and C and D, respectively. The tight dimer interface buries a total surface area of 1685 Å^2 and has eight salt bridges. In contrast to the tight dimer interface, the other interface that is formed by chains A and C and B and D chains is quite loose and buries a total surface area of 939 Å^2 and consists of three salt bridges. The interface interaction seen in this structure is very similar to the interactions present in the apo YagE structure and YagE-Pyruvate structure.

A continuous electron density from NZ atom of Lysine and carbonyl carbon (C2) of the KDGal forming a Schiff's base similar to YagE Pyruvate structure is seen [Fig. 3(A)]. The presence of NZ-C2 bond was confirmed using $3F_o-2F_c$ weighted difference map and was clearly visible at 1σ. The superposition of YagE KDGal complex active site on other KDG aldolase structures also shows the presence of a similar bond.

The active site of ApoYagE, YagE-Pyruvate complex structure show almost similar architecture with minimal change in orientation of residues [Fig. 3(B)]. The comparison of SsKDGA in complex with KDGal and YagE-KDGal structure shows that the KDGal is oriented similarly in both structures [Fig. 3(C)]. As previously mentioned, the catalytic triad formed by Y145, S56, and Y119 is well oriented and similar to the YagE-pyruvate structure. The distance between the hydroxyl group of Y145 and Oγ of S56 is 2.49Å, while that between the hydroxyl group of Y119 and Oγ of S56 is 2.49Å [Fig. 3(D)]. The Y 145 hydroxyl group interacts with O4 atom of KDGal which also corresponds to aldehyde oxygen of glyceraldehyde. This interaction also is seen in SsKDGA[21] structure and is essential for aldol condensation between pyruvate moiety and glyceraldehyde. In YagE-KDGal structure O6 interacts directly with backbone O of Gly202 and NH of Ala221, while in case of SsKDGA in complex with KDG, O5 of KDG shows a water mediated interaction with structurally equivalent Gly179 and Ala198. The YagE KDGal structure shows that both lysine 174 and the catalytic triad are intact. The formation of Schiff's base between KDGal and lysine shows that YagE is a KDG Aldolase. Since only pyruvate and glyceraldehyde were provided in the crystallization condition, both KDG and KDGal could be bound at the active site. However, since this structure has only KDGal

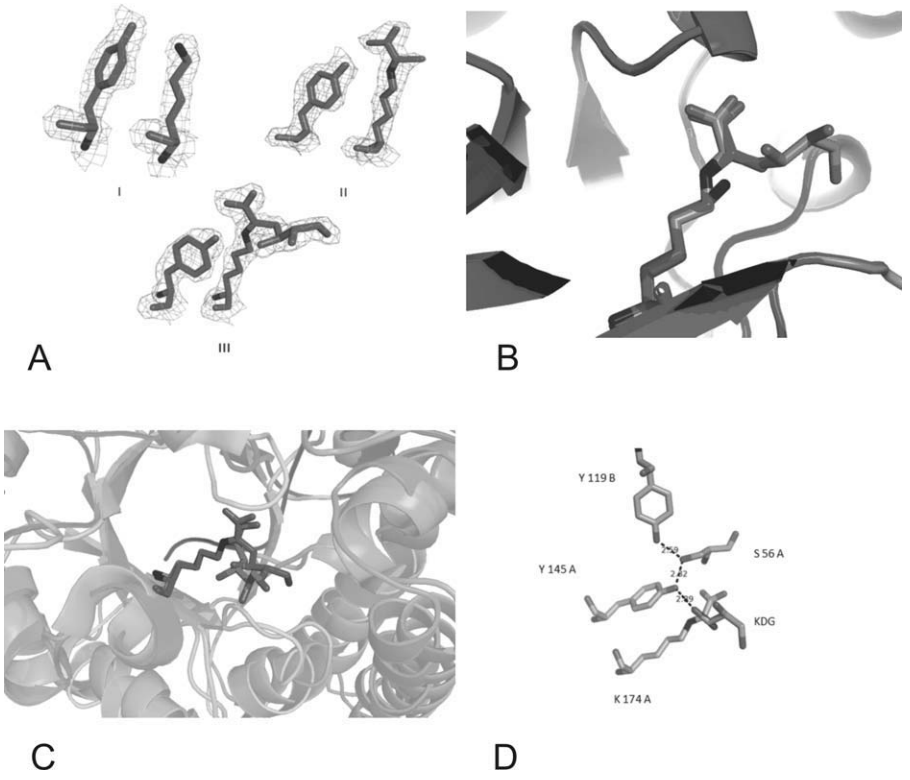

Figure 3

A: $2F_o - F_c$ map around the active site for apoYagE (I), YagE-Pyruvate (II), and YagE-KDGal (III) complex structures. **B**: Orientation of Lys 174 and ligand in apo YagE (dark), YagE-Pyruvate (gray), and YagE-KDGal structures (light gray). **C**: Comparison of orientation of covalently bonded KDGal in YagE (gray) and SsKDGA (light gray) structures. **D**: Covalently bonded KDG to Lys 174NZ and an intact catalytic triad seen in YagE-KDGal structure.

in all four active sites, probable stereo selectivity remains a putative question for further analysis.

Antibiotic resistance

Bacterial cultures expressing YagE showed better growth and survival when compared to the uninduced cultures as evident from the growth curve and estimated CFU/mL (Figs. 4 and 5). The uninduced culture showed a decrease in OD_{600} after 1 h of antibiotic addition while the induced culture containing norfloxacin, ampicillin and streptomycin showed growth similar to the culture without antibiotics. The estimated CFU/mL in the induced culture was up to 10-fold in excess to that in uninduced culture in case of Norfloxacin while it was almost 200–300-fold excess for ampicillin and streptomycin. This clearly suggests that overexpression of YagE increases the viability of the bacteria in the presence of antibiotics.

To confirm that this activity is unique to YagE, we designed a grid screen in which we tried three concentrations of norfloxacin, that is, 100 ng/mL (less than the bactericidal conc.), 500 ng/mL, 1 μg/mL, and three concentrations of IPTG, that is, 0.01 m*M*, 0.1 m*M*, and 1 m*M*. The results show that in 100 ng/mL of norfloxacin, cultures induced with all the three concentrations of IPTG showed a similar growth profile. When the norfloxacin concentration of 500 ng/mL and 1 μg/mL was used, the culture induced with 0.01 m*M* IPTG showed drastic decrease in bacterial survival, as estimated by CFU/mL. On the other hand, the cultures induced with 0.1 m*M* and 1 m*M* IPTG showed increased survival suggesting YagE indeed promotes cell survival (Supporting Information Fig. 4). Since YagE is not known to be a direct target of any of these drugs, the possibility of YagE binding to these drugs is rather low. Spectroscopic studies show there is no quenching of antibiotic fluorescence or absorbance in presence of YagE (Supporting Information Figs. 5, 6, and Table VI).

Previous studies show cell death by bactericidal antibiotics can be mediated by an increase in TCA cycle rate at 0.5 h of antibiotic addition.[30] Since YagE is a KDG Aldolase, it can use pyruvate and glyceraldehyde as a substrate. Hence we hypothesize that YagE overexpression leads to increased cell survival by reducing the cellular pool of pyruvate and thus the rate of the TCA cycle. Our

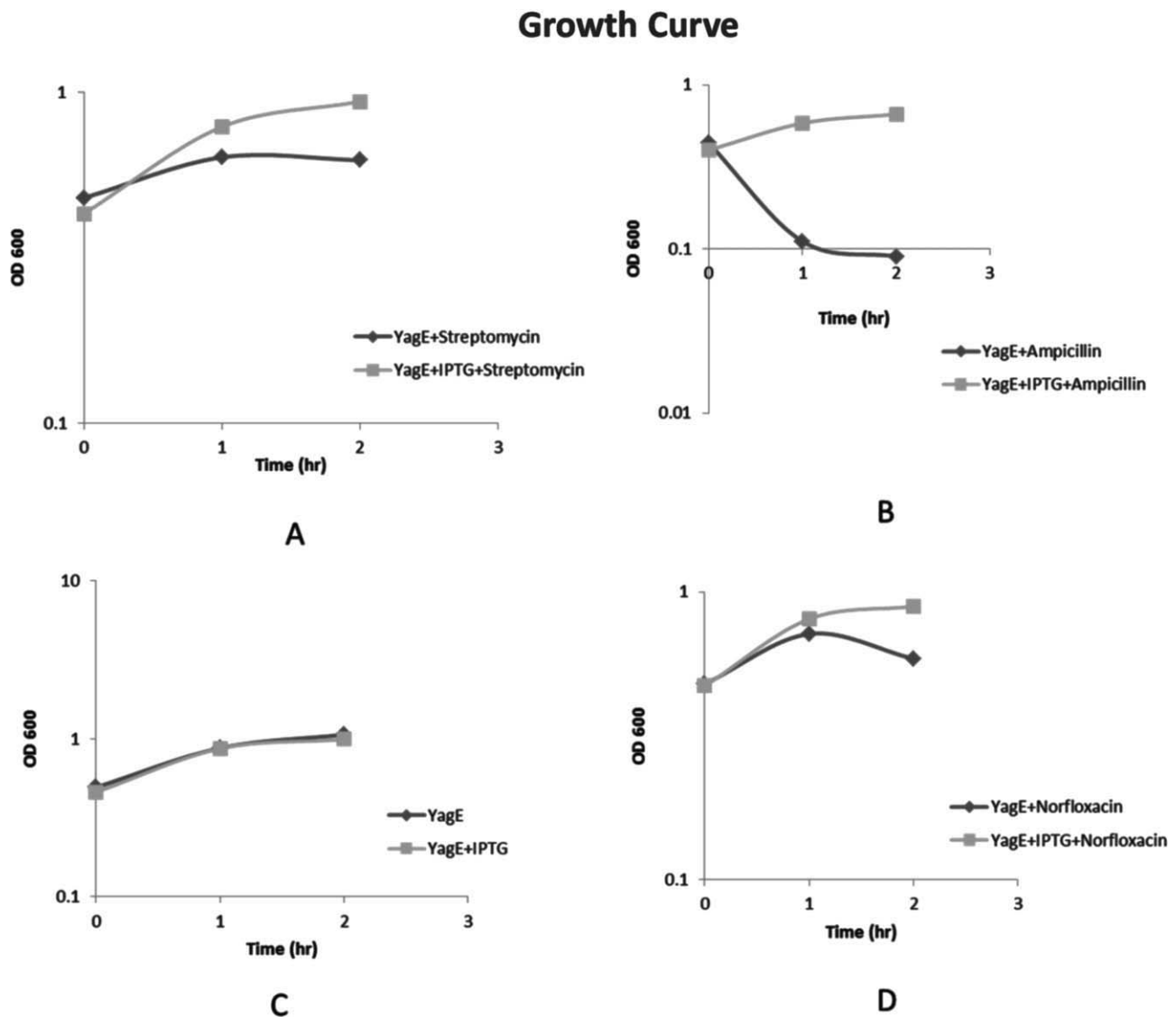

Figure 4

Growth curve for YagE under various conditions.

initial studies indicate that after 1 h the cellular pyruvate concentration was always higher in uninduced culture compared to induced culture (both with norfloxacin) by almost 8–10 μ*M* (data not shown). This indicates that overexpression of YagE reduces the effective concentrations of cellular pool of pyruvate utilized for TCA cycle.

CONCLUSION

We have determined the structure of YagE, a CP4-6 prophage-encoded protein from *E. coli* K12 in complex with pyruvate. The formation of a Schiff's base between YagE and pyruvate and the presence of intact catalytic triad indicate that it is a fully functional protein. Our docking and assay results indicate that YagE may be a KDG aldolase and negate the possibility of YagE being either a NAL or a DHDPS-like enzyme. The high-resolution structure of YagE-KDGal complex shows that YagE has KDGA activity. The interactions of KDGal with YagE are very similar to other KDG aldolases except for the absence of some water mediated interactions. Although the precise *in vivo* role of YagE is yet to be elucidated, this study suggests that YagE is a potential virulence factor and may be important as no homolog of KDGA is found in *E. coli* genome.[31] It is interesting to note that since the archeal KDGA are involved in alternative ED Pathway, the presence of YagE as a replacement for KDGA might suggest that alternative ED pathway could be active in *E. coli* also. Further experiments are needed to explore this possibility indicated by our study.

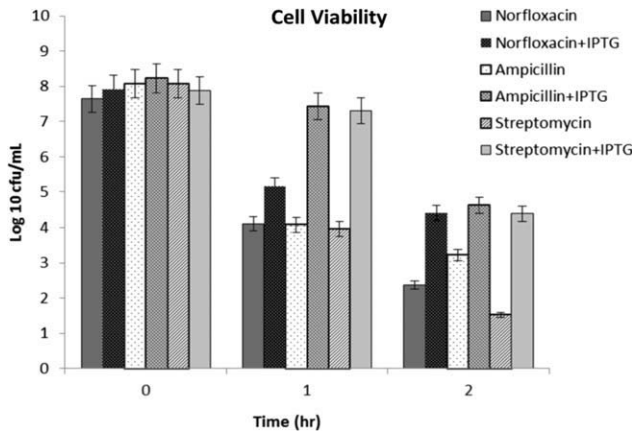

Figure 5

Comparison of Log_{10}CFU/mL of induced (0.1 m*M* IPTG) and uninduced YagE clones in the presence of 1 μg/mL norfloxacin, 100 μg/mL Ampicillin, and 100 μg/mL streptomycin.

Antibiotic stress is known to reduce the sugar intake by bacteria through down-regulation of the symporters which involves reduced transcription of the entire operon involved in sugar metabolism including sugar anabolism genes.[30] As the bacteria over-expressing YagE showed increased viability, it is possible that if genes involved in sugar anabolism are controlled by monocistronic transcripts, their up-regulation during stress may lead to increased bacterial viability. If these factors are phage-encoded, this could be an example of a symbiotic association between prophage and bacteria. This could also imply that a novel mechanism of antibiotic resistance may develop in bacteria that does not directly inhibit drug binding but may instead alter the sugar metabolism to prolong survival.

ACKNOWLEDGMENTS

The authors thank DBT for the Centre for Excellence in Bioinformatics facilities at Madurai Kamaraj University, DBT Programme Support for Macromolecular X-ray Diffraction Facility, and Indo-Israel DBT project. They thank the anonymous referee for detailed comments and suggestions.

REFERENCES

1. Casjens S. Prophages and bacterial genomics: what have we learned so far? Mol Microbiol 2003;49:277–300.
2. Joerger AC, Mayer S, Fersht AR. Mimicking natural evolution in vitro: an *N*-acetylneuraminate lyase mutant with an increased dihydrodipicolinate synthase activity. Proc Natl Acad Sci USA 2003;100:5694–5699.
3. Barbosa JA, Smith BJ, DeGori R, Ooi HC, Marcuccio SM, Campi EM, Jackson WR, Brossmer R, Sommer M, Lawrence MC. Active site modulation in the *N*-acetylneuraminate lyase sub-family as revealed by the structure of the inhibitor-complexed *Haemophilus influenzae* enzyme. J Mol Biol 2000;303:405–421.
4. Laber B, Gomis-Ruth FX, Romao MJ, Huber R. *Escherichia coli* dihydrodipicolinate synthase. Identification of the active site and crystallization. Biochem J 1992;288 (Part 2):691–695.
5. Lawrence MC, Barbosa JA, Smith BJ, Hall NE, Pilling PA, Ooi HC, Marcuccio SM. Structure and mechanism of a sub-family of enzymes related to *N*-acetylneuraminate lyase. J Mol Biol 1997;266:381–399.
6. Blickling S, Renner C, Laber B, Pohlenz HD, Holak TA, Huber R. Reaction mechanism of *Escherichia coli* dihydrodipicolinate synthase investigated by X-ray crystallography and NMR spectroscopy. Biochemistry 1997;36:24–33.
7. Dobson RC, Valegard K, Gerrard JA. The crystal structure of three site-directed mutants of *Escherichia coli* dihydrodipicolinate synthase: further evidence for a catalytic triad. J Mol Biol 2004;338:329–339.
8. Blattner FR, Plunkett G, III, Bloch CA, Perna NT, Burland V, Riley M, Collado-Vides J, Glasner JD, Rode CK, Mayhew GF, Gregor J, Davis NW, Kirkpatrick HA, Goeden MA, Rose DJ, Mau B, Shao Y. The complete genome sequence of *Escherichia coli* K-12. Science 1997;277:1453–1462.
9. Manicka S, Peleg Y, Unger T, Albeck S, Dym O, Greenblatt HM, Bourenkov G, Lamzin V, Krishnaswamy S, Sussman JL. Crystal structure of YagE, a putative DHDPS-like protein from *Escherichia coli* K12. Proteins 2008;71:2102–2108.
10. Kabsch W. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. J Appl crystallography 1993;26:795–800.
11. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. J Appl Crystallogr 2007;40:658–674.
12. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr 1997;53:240–255.
13. Vagin AA, Steiner RA, Lebedev AA, Potterton L, McNicholas S, Long F, Murshudov GN. REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. Acta Crystallogr D Biol Crystallogr 2004;60:2184–2195.
14. Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Cryst 2010;D66:213–221.
15. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 2004;60:2126–2132.
16. CCP4Team. The CCP4 suite: programs for protein crystallography. Acta Crystallogr D Biol Crystallogr 1994;50:760–763.
17. Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, Repasky MP, Knoll EH, Shelley M, Perry JK, Shaw DE, Francis P, Shenkin PS. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. J Med Chem 2004;47:1739–1749.
18. Halgren TA, Murphy RB, Friesner RA, Beard HS, Frye LL, Pollard WT, Banks JL. Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. J Med Chem 2004;47:1750–1759.
19. Verdonk ML, Cole JC, Hartshorn MJ, Murray CW, Taylor RD. Improved protein-ligand docking using GOLD. Proteins 2003;52:609–623.
20. Garrett MM, Ruth H, William L, Michel FS, Richard KB, David SG, Arthur JO. AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. J Comput Chem 2002;30:2785–2791.
21. Theodossis A, Walden H, Westwick EJ, Connaris H, Lamble HJ, Hough DW, Danson MJ, Taylor GL. The structural basis for substrate promiscuity in 2-keto-3-deoxygluconate aldolase from the Entner-Doudoroff pathway in *Sulfolobus solfataricus*. J Biol Chem 2004;279:43886–43892.

22. Pearce FG, Perugini MA, McKerchar HJ, Gerrard JA. Dihydrodipicolinate synthase from *Thermotoga maritima*. Biochem J 2006;400:359–366.
23. Delano WL. The PyMOL molecular graphics system. DeLano Scientific, San Carlos, CA, USA, 2002.
24. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera—a visualization system for exploratory research and analysis. J Comput Chem 2004;25:1605–1612.
25. Comb DG, Roseman S. The sialic acids. I. The structure and enzymatic synthesis of *N*-acetylneuraminic acid. J Biol Chem 1960;235: 2529–2537.
26. Wolterink-Van Loo S, Andr'e Van E, Marco AJS, Jasper A, Bauke WD, Van Der Oost J. Biochemical and structural exploration of the catalytic capacity of *Sulfolobus* KDG aldolases. Biochem J 2007;403:421–430.
27. Yugari Y, Gilvarg C. The condensation step in diaminopimelate synthesis. J Biol Chem 1965;240:4710–4716.
28. Dobson RC, Devenish SR, Turner LA, Clifford VR, Pearce FG, Jameson GB, Gerrard JA. Role of arginine 138 in the catalysis and regulation of *Escherichia coli* dihydrodipicolinate synthase. Biochemistry 2005;44:13007–13013.
29. Borthwick EB, Connell SJ, Tudor DW, Robins DJ, Shneier A, Abell C, Coggins JR. *Escherichia coli* dihydrodipicolinate synthase: characterization of the imine intermediate and the product of bromopyruvate treatment by electrospray mass spectrometry. Biochem J 1995;305 (Part 2):521–524.
30. Michael AK, Daniel JD, Boris H, Carolyn AL, James JC. A common mechanism of cellular death induced by bactericidal antibiotics. Cell 2007;130:797–810.
31. Kenneth ER. EcoGene: a genome sequence database for *Escherichia coli* K-12. NAR 2000;28:60–64.
32. Harry GB. Solid-state fluorescence of the trihydrate phases of ampicillin and amoxicillin. AAPS PharmSciTech 2005;6:444–448.
33. Gongwu S, Yu H, Zhaoxia C. A fluorescence spectroscopic study of the interaction between norfloxacin and DNA. Can J Anal Sci Spectroscopy 2004;49:203–209.