Horne Christopher (Orcid ID: 0000-0003-1318-514X)
Kind Laura (Orcid ID: 0000-0002-1897-5936)

1

Title:

On the structure and function of *Escherichia coli* YjhC: an oxidoreductase involved in bacterial sialic acid metabolism

Running title:

On the structure and function of *E. coli* YjhC

Authors:

Christopher R. Horne ^a, Laura Kind ^b, James S. Davies ^a, Renwick C.J. Dobson ^{a,c,*}

Author Affiliations:

- a) Biomolecular Interaction Centre and School of Biological Sciences, University of Canterbury, PO Box 4800, Christchurch 8041, New Zealand.
- b) Department of Biomedicine, University of Bergen, Bergen 5020, Norway.
- c) Bio21 Molecular Science and Biotechnology Institute, Department of Biochemistry and Molecular Biology, University of Melbourne, Parkville, Victoria 3010, Australia.

****Corresponding authors:***

Prof. Renwick Dobson, Biomolecular Interaction Centre and School of Biological Sciences, University of Canterbury, Christchurch 8140, New Zealand.

Email: renwick.dobson@canterbury.ac.nz

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: [10.1002/prot.25846](https://doi.org/10.1002/prot.25846)

Acknowledgements:

This work was supported by the following funding to R.C.J.D: **1)** the New Zealand Royal Society Marsden Fund (contract UOC1506); **2)** a Ministry of Business, Innovation and Employment Smart Ideas grant (contract UOCX1706); and **3)** the Biomolecular Interaction Centre (University of Canterbury). **This work was also supported by funding from the Canterbury Medical Research Foundation to C.R.H.** We acknowledge the Australian Synchrotron, the staff at the MX2 beamline and Dr Nigel Kirby and Dr Tim Ryan for provision of synchrotron facilities along with their assistance in using the beamlines. This research was undertaken in part using the MX2 beamline at the Australian Synchrotron, part of ANSTO, and made use of the ACRF detector. We thank Professor Antony Fairbanks and Dr Stewart Alexander for the chemical synthesis of the sialic acid 1,7 lactone. We thank Dr Rachel North for her insight on **bacterial** sialic acid metabolism.

Abstract:

Human pathogenic and commensal bacteria evolved the ability to scavenge host-derived sialic acids and subsequently degrade them as a source of nutrition. Expression of the *Escherichia coli yjhBC* operon is controlled by the repressor protein *nanR*, which regulates the core machinery responsible for the import and catabolic processing of sialic acid. The role of the *yjhBC* encoded proteins is not known—here we demonstrate that the enzyme YjhC is an oxidoreductase involved in bacterial sialic acid degradation. First, we demonstrate *in vivo* using knockout experiments that YjhC is broadly involved in carbohydrate metabolism, including that of *N*-acetyl-D-glucosamine, *N*-acetyl-D-galactosamine and *N*-acetylneuraminic acid. Differential scanning fluorimetry demonstrates that YjhC binds *N*-acetylneuraminic acid and its lactone variant, along with NAD(H), which is consistent with its role as an oxidoreductase. Next, we solved the crystal structure of YjhC in complex with the NAD(H) cofactor to 1.35 Å resolution. The protein fold belongs to the Gfo/Idh/MocA protein family. The dimeric assembly observed in the crystal form is confirmed through solution studies. Ensemble refinement reveals a flexible loop region that may play a key role during catalysis, providing essential contacts to stabilize the substrate—a unique feature to YjhC among closely related structures. Guided by the structure, *in silico* docking experiments support the binding of sialic acid and several common derivatives in the binding pocket with an overall positive charge distribution. Taken together, our results verify the role of YjhC as a *bona fide* oxidoreductase/dehydrogenase and provide the first evidence to support its involvement in sialic acid metabolism.

Keywords: YjhC, Oxidoreductase, Gfo/Idh/MocA family, NAD, Sialic acid, X-ray
Crystallography, Small Angle X-ray Scattering, Analytical Ultracentrifugation

Introduction

Mammalian cell surfaces are decorated with a complex array of glycoconjugates^{1;2}. Located at the termini of many cell surface glycoconjugates are sialic acids—a family of nine-carbon amino monosaccharide units³ that are comprised of more than 50 naturally occurring and structurally distinct variants⁴. The most common form of sialic acid is *N*-acetylneuraminic acid, herein abbreviated to Neu5Ac (**Figure 1A**)^{3;4}, while other common variants include the lactone (between carbon 1 and 7) and the 9-O-acetylated derivatives of Neu5Ac (**Figure 1B**). Given their terminal location on glycoconjugates, these negatively-charged amino sugars mediate a diverse array of cellular interactions (recognition and adhesion processes), through carbohydrate-binding proteins and their associated receptors^{1;2}.

(Insert Figure 1)

Pathogenic and commensal bacteria evolved the ability to scavenge host-derived sialic acids³, which are a source of carbon, nitrogen and energy⁵. In addition, the catabolism of sialic acids yields precursors for peptidoglycan biosynthesis, such as *N*-acetylglucosamine^{6;7}. Lastly, some bacteria can decorate their cell surfaces with host-derived sialic acids to camouflage themselves against the host innate immune system^{1;8}. Given the accessibility of sialic acids in mucus rich environments, their utilization offers pathogenic bacteria and commensals, such as *Escherichia coli*, a competitive advantage to colonize and persist within the human host^{1;9;10}.

In *E. coli*, the *nanATEK-yhcH* operon encodes a lyase (*nanA*), a sugar-proton symporter (*nanT*), an epimerase (*nanE*), and a kinase (*nanK*), which together are essential for the

utilization of Neu5Ac^{5; 6; 8}. The expression of these enzymes from the *nanATEK-yhcH* operon is controlled by a transcriptional regulator, NanR, which is in turn regulated by the presence of sialic acids¹¹. In addition to this core operon, NanR also regulates the *nanCMS* and *yjhBC* operons, collectively termed the sialoregulon (**Figure 1C**)¹¹. The *nanCMS* operon encodes an outer membrane porin (*nanC*), which is responsible for the selective import of sialic acid^{12; 13}, a sialate mutarotase (*nanM*) that catalyses the periplasmic conversion of the α -anomer of Neu5Ac to the thermodynamically favourable β -anomer¹⁴, and a 9-O-acetylated Neu5Ac esterase (*nanS*) that permits growth on this sialic acid derivative¹⁵. Although the biological role of both the *E. coli nanATEK-yhcH* and *nanCMS* operons in sialic acid metabolism is reasonably well understood, the biological role of the *yjhBC* operon has yet to be established—here we focus on the role of *yjhC*.

YjhC is hypothesized to function as an oxidoreductase for less common forms of sialic acid⁵. Our bioinformatic analysis places YjhC in the Gfo/Idh/MocA family of oxidoreductase/dehydrogenase enzymes that use nicotinamide adenine dinucleotide (NAD) or nicotinamide adenine dinucleotide phosphate (NADP) as a cofactor. This analysis was achieved via a sequence homology search (see Materials and Methods). Typical substrates of the Gfo/Idh/MocA family oxidoreductases include pyranose sugars^{16; 17}, although some members have shown catalytic activity towards other substrate types, such as biliverdin; a product of heme catabolism¹⁸ or trans-dihydrodiols; an aromatic hydrocarbon¹⁹.

Here we address two research questions: 1) Is YjhC involved in amino sugar catabolism? and 2) Is YjhC a *bona fide* oxidoreductase/dehydrogenase? To address these questions, we used BiOLOG Phenotype Microarrays²⁰ to report on the biological role of YjhC, along with

thermal shift assays to screen potential substrates for binding. Further, we have solved the crystal structure of *E. coli* YjhC in complex with the cofactor NAD(H) at a resolution of 1.35 Å. Guided by the structure, the substrate-binding pocket was examined, and using *in silico* docking experiments, we support the binding of the acidic sugar, Neu5Ac in its cyclized and open-chain forms, along with the sialic acid 1,7 lactone and the 9-O-acetylated variant.

Materials and Methods

Bioinformatics

To search and compare protein sequences for YjhC, the basic local alignment search tool (*BLAST*) program and *BLASTp*²¹ was used. The *DALI* server²² was used for protein structure comparison. Multiple sequence alignments of YjhC homologues with known atomic structures were generated using *Clustal Omega*²³ and visualized using *ESPrpt 3.0*²⁴. Amino acid sequences and atomic structures of homologues were sourced from the UniProt database and Protein Data Bank (PDB), respectively.

Cloning of the YjhC expression construct

The gene encoding YjhC from *E. coli* (UniProt-P39353) was synthesized commercially by GenScript and supplied in a cloning vector, designated pUC57. In order to use a N-terminal His-tag, the *yjhC* gene was digested from pUC57 using the high-fidelity restriction enzymes *Bam*HI and *Hind*III (New England Biolabs), gel purified and ligated with T4 DNA ligase (Takara) to generate pET30ΔSE/*yjhC*. The pET30ΔSE expression vector, conferring kanamycin resistance was kindly supplied by Dr Hironori Suzuki. The recombinant expression vector was transformed into DH5α competent cells, purified using a DNA-Spin™

Plasmid DNA Purification Kit (iNtRon Biotechnology), and verified by DNA sequencing (Macrogen).

Expression and purification of recombinant YjhC

The pET30ΔSE/*yjhC* construct was transformed into *E. coli* BL21(DE3) competent cells (Agilent) and grown in Luria Broth (LB) supplemented with kanamycin (30 mg mL⁻¹) at 37 °C, with shaking at 220 rpm. Expression of YjhC was induced mid-log phase (OD₆₀₀ ≈ 0.6) by the addition of isopropyl β-D-1-thiogalactopyranoside to 1 mM and incubated overnight at 26 °C. Cells were harvested by centrifugation (Thermo Sorvall RC-6-Plus centrifuge) at 8,000 rpm for 10 min and resuspended in buffer A (20 mM Tris-HCl (pH 8.0), 500 mM NaCl, 40 mM imidazole). The cells were lysed by sonication (Hielscher UP200S Ultrasonic Processor) on ice, and cellular debris was pelleted by centrifugation at 16,000 rpm for 30 min.

Purification of YjhC was conducted *via* a two-step procedure: immobilized-metal affinity chromatography (IMAC) and size-exclusion chromatography (SEC). For IMAC the soluble cell lysate was applied onto a His-Trap FF Crude 5 mL column (GE Healthcare) pre-equilibrated with buffer A. The column was washed with buffer A until a steady baseline absorbance was observed. YjhC was eluted using buffer B (20 mM Tris-HCl (pH 8.0), 500 mM NaCl, 400 mM imidazole). Fractions were analyzed by SDS-PAGE and the protein containing fractions were pooled. The sample was spin-concentrated to 500 μL, filtered through a 0.2 μm filter to remove aggregate, and loaded onto a Superdex 200 Increase 10/300 GL SEC column (GE Healthcare) equilibrated with buffer C (20 mM Tris-HCl (pH 8.0), 50 mM NaCl), for further purification and removal of any aggregates. YjhC was purified to near

homogeneity as visualized by SDS-PAGE (**Supplementary Figure 1**). The absorbance at 280 nm (A_{280}) of the purified protein was measured using a Cary 100-Bio UV-Vis spectrophotometer (Agilent Technologies) and the concentration estimated using a molar extinction coefficient of $42,400 \text{ M}^{-1} \text{ cm}^{-1}$ at 280 nm, as calculated by ProtParam²⁵. All purification steps were carried out at 4 °C. Purified YjhC was flash-frozen in liquid nitrogen for storage.

BiOLOG Phenotype MicroArray

Commercially available BiOLOG Phenotype MicroArray plates were used to explore the biological role of YjhC. For this experiment, the BL21(DE3) strain was used as the wild-type, while the chromosomal *yjhC* knockout strain, defined as $\Delta yjhC::kan$ strain was obtained from the *E. coli* Genetic Stock Centre. The 96-well Phenotype MicroArray plates were pre-configured with a variety of compounds and a tetrazolium redox dye, which enabled the phenotypic reaction of growing cells to be monitored in a high-throughput system. Both PM1 and PM2A carbon Phenotype Microarrays were used in this experiment.

For preparation of the inoculum, cells were initially streaked and grown overnight at 37 °C on LB-agar plated media supplemented with 30 mg mL⁻¹ kanamycin. After incubation, cells were resuspended in IF-0a inoculating fluid to an OD₆₀₀ of 0.035 using a Cary 100 Bio UV/Vis spectrophotometer (Agilent Technologies). This inoculum was then combined with the tetrazolium redox dye according to the manufacturer's instructions. Carbon sources were rehydrated by adding 100 µL of the inoculum solution to each plate well (position A1 did not have a carbon source and therefore served as a reference/negative control) and incubated

aerobically without shaking at 37 °C. Metabolism of the various carbon sources was monitored spectrophotometrically at 570 nm using a SpectraMax M5 plate reader (Molecular Devices) over a 48 h period. The reduction of the tetrazolium dye caused a purple color to develop, which is indicative of active respiration and metabolism of the respective compound. For analysis, the OD₅₇₀ for the reference (position A1) was subtracted from each well and then plotted against each strain to generate a kinetic growth curve. Here, a major phenotypic change between strains was defined as a difference in the rate of respiration by more than 50% after 24 h growth (late log phase), on the same carbon source. The entire experiment was performed in duplicate for both wild-type and $\Delta yjhC$:kan bacterial strains. The kinetic growth curves for the PM1 and PM2A plates are presented as supplementary data, along with the contents of each plate (**Supplementary Figure 2 and 3**). The IF-0a inoculating fluid, tetrazolium dye mix, and carbon utilization plates (PM1 and PM2A) were all purchased from BiOLOG Inc. (Harvard, CA, USA). The raw data is included in **Supplementary Figure 4**.

Ligand screening using differential scanning fluorimetry

The QuantStudio 3 real-time PCR system (ThermoFisher Scientific) was used to record the thermal stability of the purified protein with and without potential substrates. Reactions consisted of 10 μ M protein, 5x SYPRO Orange (prepared as a 50x stock) and the desired concentration of potential substrate (Carbon source = 40 mM, NAD(P) = 5 mM) in a final reaction volume of 25 μ L. Samples were heated from 4 °C to a 95 °C at a ramp rate of 0.05 °C, taking fluorescent readings at each time point. Triplicate measurements were performed

for each sample. To validate the experiment, one positive control using a well-characterized melting curve (lysozyme at 10 μ M) , along with two negative controls (protein only and dye only) were conducted. Data analysis was performed using Protein Thermal Shift software (version 1.3–Applied Biosystems), where an apparent melting point (T_m) of each sample in $^{\circ}$ C was obtained from the lowest point of the first derivative plot. A summary of all T_m is shown in **Supplementary Table 1**.

Crystallization, data collection and structure determination of YjhC

Initial crystallization trials for *E. coli* YjhC were performed in-house using the PACT *premier*, Clear Strategy 1, and Clear Strategy 2 commercial screens (Molecular Dimensions). Trials were conducted using the sitting-drop vapor-diffusion method at 20 $^{\circ}$ C with droplets consisting of 400 nL protein solution and 400 nL reservoir solution. Purified YjhC was concentrated to 12 mg mL⁻¹ in buffer C, supplemented with 5 mM of the cofactor NAD, which was identified using differential scanning fluorimetry (**Figure 2 and Supplementary Table 1**). The crystals were cryo-protected by soaking in 85% reservoir solution and 15% glycerol-ethylene glycol (50:50 mix) prior to being flash-frozen in liquid nitrogen. Condition A1 (1.5 M ammonium sulfate, 0.1 M sodium acetate pH 5.5) from Clear Strategy 2 produced large crystals after four weeks growth. The crystals produced in this condition were suitable for diffraction. Diffraction data were collected ($\lambda = 0.9537$ Å) using the MX2 beamline at the Australian Synchrotron with the EIGER X 16 M detector. Data was processed using *XDS*²⁶ and *AIMLESS* from the CCP4 program suite²⁷. Details of the data collection statistics are summarized in **Table 1**.

The structure of YjhC was initially solved using *Auto-Rickshaw*²⁸ to 1.94 Å resolution, in the spacegroup $P2_12_12_1$. Crystallographic phases were determined by molecular replacement using chain A of PDB entry 5A02; an aldose-aldose oxidoreductase from *Caulobacter crescentus*¹⁷ with a sequence identity of 24%. Diffraction data was collected without the use of cryo-protectant, using litho-loops. The refined output model at 1.94 Å from *Auto-Rickshaw*²⁸ was used as a search model for molecular replacement to solve the structure of YjhC at 1.35 Å using *PHASER*²⁹. The model was refined using *REFMAC5*³⁰ from the CCP4 program suite, and *PHENIX REFINE*³¹. To eliminate model bias, simulated annealing using Cartesian restraints was performed. Iterative improvement of the model, including the addition of ligand and water molecules were performed in *COOT*³². Water molecules were manually assessed to fit to the observed density and for hydrogen bonding partners. Final refinement was performed using translation-libration-screw (TLS) refinement to produce a final model of *E. coli* YjhC refined to a R_{work} and R_{free} value of 14.98% and 16.87%, respectively. Model validation was performed using *MOLPROBITY*³³. Details of the structural refinement and model validation are summarized in **Table 1**. To capture possible substrate leads (Neu5Ac and sialic acid 1,7 lactone) in the binding cavity of YjhC, crystals from condition A1 were soaked in cryo-protectant containing 80% reservoir solution, 10% glycerol-ethylene glycol (50:50 mix), and 10% of either Neu5Ac or sialic acid 1,7 lactone, each at a final concentration of 20 mM prior to data collection.

Ensemble refinement

Ensemble refinement was used to evaluate the dynamics and model sources of disorder simultaneously within the final atomic model. Following TLS refinement in *PHENIX REFINE*³¹, all missing residues or truncated side chain atoms (due to insufficient electron density) were added using *COOT*³². Any alternative conformations within the model are automatically removed, while all occupancies are set to one. An ensemble of structures was generated by time-averaged refinement using *PHENIX ENSEMBLE REFINEMENT*³⁴. The averaging time (τ_x) of structure factors used during the refinement was selected based on the resolution of the diffraction data. Harmonic restraints were applied to all three sulfate ions to prevent stochastic displacement—two were placed at the interface and one closer to the active site. To model the contribution of intramolecular disorder an overall TLS model was applied using one group per chain³⁴ with a p_{TLS} of 0.8, which defines the number of atoms included in the TLS refinement. A total of 100 ensemble structures were reported where the R_{work} and R_{free} values decreased by 2.2% and 1.0%, respectively. Refinement statistics are presented in **Supplementary Table 2**.

Analytical ultracentrifugation

Sedimentation velocity experiments were performed with a XL-I analytical ultracentrifuge (Beckman Coulter) using double sector quartz cells and epon center-pieces in an An-50 Ti 8-hole rotor. Data was obtained at 45,000 rpm using 380 μL protein at three different concentrations (7, 14 and 20 μM) in buffer C as the sample, and 400 μL buffer C as the reference. A total of 100 scans were collected at 20 °C using radial absorbance scans at 280 nm and a step size of 0.003 cm. All data were analyzed using *UltraScan 4.0*, release 2578³⁵. Sedimentation data were evaluated by the two-dimensional spectrum analysis (2DSA)³⁶,

which affords an unbiased hydrodynamic model, where the sedimentation coefficient and frictional ratio (f/f_0) can be floated independently. Furthermore, 50 Monte-Carlo iterations were performed on the 2DSA data set providing 95% confidence intervals for all measured parameters. All fitting procedures were completed using the *UltraScan* LIMS cluster. The buffer density, buffer viscosity and an estimate of the partial specific volume of the protein sample based on the amino acid sequence were determined using *UltraScan*. Data analysis are summarized in **Table 2**.

Small angle X-ray scattering

Small angle X-ray scattering (SAXS) data were collected on the SAXS/WAXS beamline equipped with a Pilatus 1M detector (170 mm $\times$ 170 mm, effective pixel size, 172 μm $\times$ 172 μm) at the Australian Synchrotron. A sample detector distance of 1600 mm was used, which provided a q range of 0.006-0.5 $\text{\AA}^{-1}$. 50 μL of purified YjhC protein (8 mg mL^{-1}) was injected onto an inline Superdex S200 Increase 5/150 GL SEC column (GE Healthcare), equilibrated with 20 mM Tris-HCl (pH 8.0), 150 mM NaCl, supplemented with 5% (v/v) glycerol and 0.1% (w/v) sodium azide, using a flow rate of 0.45 mL min^{-1} . Coflow SAXS was used to minimize sample dilution and maximize signal to noise³⁷. Scattering data was collected in one second exposures ($\lambda = 1.0332 \text{ \AA}$) over a total of 400 frames, using a 1.5 mm glass capillary, at 12 $^{\circ}\text{C}$. 2D intensity plots were radially averaged, normalized to sample transmission, and background subtracted using the *Scatterbrain* software package (Australian Synchrotron).

Analysis of the scattering data was performed using the *ATSAS* software package (version 2.8.4)³⁸. *PrimusQT*³⁹ was used to perform the Guinier analysis and to generate the pairwise distribution function $P(r)$, calculated using an indirect Fourier transform. Molecular mass was estimated using the *SAXS-MoW2* package⁴⁰, while *CRY SOL*⁴¹ was used to evaluate the solution scattering from the known atomic coordinates of YjhC (solved in this study) by fitting to the experimental scattering data. Data collection and analysis are summarized in **Table 2**.

In silico docking experiments

The molecular interactions that may occur between a small molecule and a protein were predicted using the web service *SwissDock*⁴² interfaced with the *CHARMM* package⁴³. To initiate docking, we defined a search space around the center of mass of our suspected binding cavity as a set of x, y and z coordinates (x center/size = -13.76/17.57; y center/size = 1.11/15.70; z center/size = -19.20/17.85). Secondly, the target protein and ligand were processed to remove the solvent, repair any truncated sidechains using a rotamer library, add hydrogens, assign partial charges and convert all files to a Mol2 file format. Binding modes were then scored using their estimated binding free energy and full fitness, which is defined as the total energy of the system along with a solvation term. All docking results were visualized in *UCSF Chimera*⁴⁴ using *ViewDock*. The binding modes are summarized in **Supplementary Table 3**.

NAD kinetic assay

Kinetic analysis of YjhC was performed using a Cary 100 Bio UV/Vis spectrophotometer (Agilent Technologies), where the reaction was observed by measuring the increase in absorbance at 340 nm due to the consumption of NAD⁺. Assays were performed in duplicate, using acrylic cuvettes with a path length of 1 cm. All assay mixtures were equilibrated at 25 °C for 20 min using the heat-controlled Peltier module, prior to measurement. All other stock solutions were kept on ice. The assay mixture consisted of: 20 mM Tris-HCl pH 8.0, 150 mM NaCl, 2 mM NAD⁺ (Sigma-Aldrich) and Neu5Ac (Carbosynth), which varied from 5 mM to 60 mM (980 µL total). The assay was initiated following the additional of 20 µL YjhC at a final concentration of 0.05 µM. The assay was repeated using sialic acid 1,7 lactone as the substrate and 0.2 mM NADH (Sigma-Aldrich). In this experiment, the reaction was observed by measuring the decrease in absorption at 340 nm due to the conversion of NADH to NAD⁺. Controls without enzyme were performed to measure the background turnover of NAD⁺/NADH in solution. All initial rate data were analyzed and fitted with the Michaelis-Menten model using *Prism 8.2.0* (GraphPad Software, La Jolla California USA).

Results and Discussion

YjhC is involved in amino sugar catabolism

A protein's primary sequence is often used to predict its function. However, it is estimated that approximately one-third of all bacterial proteins have such low sequence similarity to do this accurately, meaning many databases annotate protein sequence with an incorrect function^{45; 46}. This is often the case with homologous proteins that have diverse substrate specificities

⁴⁷, such as the Gfo/Idh/MocA family. To investigate the function of the oxidoreductase YjhC and its predicted role in sialic acid catabolism, two approaches were taken.

First, we examined the phenotypic consequence of knocking out the *yjhC* gene of *E. coli*, when grown in a range of different carbon sources by using the BiOLOG Phenotype MicroArray system. In this experiment, the overall metabolic activity of the *yjhC* knockout was significantly reduced compared with the wild-type when grown in 15 different carbon sources (**Figure 2A**). These were L-arabinose, D-arabinose, *N*-acetyl-D-glucosamine, D-trehalose, D-mannose, D-mannitol, D-fructose, α -D-glucose, maltose, α -D-lactose, *N*-acetyl-D-galactosamine, maltotriose, *N*-acetylneuraminic acid, β -D-allose, and D-galactonic acid- γ -lactone. Interestingly, most of these carbon sources are carbohydrates (**Figure 2B**), where the pyranose sugars displayed the most pronounced catabolic divergence to the wild-type, as identified through the kinetic growth curves (**Supplementary Figure 2 and 3**). In contrast, an increase in the metabolic activity of the *yjhC* knockout was observed for four different carbon sources; D-saccharic acid, mucic acid, D-malic acid and L-malic acid, when compared with the wild-type (**Figure 2A**)—the reason for this increase was not clear. Phenotype MicroArrays measure respiration, not growth. As such, where there is reduced metabolism for the *yjhC* knockout compared to the wild-type, the cells may be incapable of converting these sole carbon sources into the downstream metabolites required for cell division, thus leading to potential false negatives ⁴⁸. Overall these data support the annotation of YjhC within the Gfo/Idh/MocA family, with a pyranose sugar as their likely substrate.

(Insert Figure 2)

Next, we used differential scanning fluorimetry to test whether the compounds implicated in the Phenotype MicroArray experiment bound directly to YjhC. Binding is indicated by a positive or negative thermal shift of the unfolding temperature for the protein (T_m)⁴⁹. To verify the role of YjhC as a *bona fide* oxidoreductase, the cofactors NAD⁺ and NADP⁺ were screened. The addition of NAD⁺ produced a positive thermal shift of 5.85 °C (**Figure 2C**), whereas NADP⁺ had no significant change (0.12 °C). NADH produced a similar shift to the oxidized cofactor of 6.94 °C (**Figure 2D**). Within the Gfo/Idh/MocA family, members who modify carbohydrates typically catalyze their respective reactions *via* a ping-pong mechanism, where the cofactor binds first, followed by the substrate^{16; 17; 50}. In the absence of NAD⁺/NADH, the carbon compounds tested gave only small thermal shifts, within a range of 0.06-0.37 °C, which we did not consider significant (**Supplementary Table 1**). However, in the presence of NAD⁺/NADH, both Neu5Ac and the sialic acid 1,7 lactone had a significant effect on the thermal stability of YjhC with an additional thermal shift of 2.86 and 0.98 °C, respectively (**Figure 2C-D**). The remaining carbon compounds did not significantly affect the thermal stability and only produced thermal shifts between 0.01-0.48 °C (**Supplementary Table 1**).

Taken together, our Phenotype MicroArray and differential scanning fluorimetry experiments suggest that Neu5Ac or analogous sialic acid variants could be the biologically relevant substrate of YjhC. Moreover, we show that YjhC can preferentially bind the cofactor NAD(H) over the cofactor NADP(H), providing strong evidence that YjhC is an oxidoreductase.

YjhC adopts a common oxidoreductase architecture

To add further weight to the hypothesis that YjhC is a *bona fide* oxidoreductase, as predicted from the bioinformatic analysis and thermal shift assays, the structure was solved at a resolution of 1.35 Å (**Table 1**). The structure is in complex with NAD(H) and has two monomers in the asymmetric unit that associate closely to form a dimer (**Figure 3A**), which is discussed later in more detail.

(Insert Table 1)

(Insert Figure 3)

The monomer of YjhC shares a typical architecture to other Gfo/Idh/MocA protein family members and consists of two domains (**Figure 3B**). The N-terminal domain (residues 1-117) adopts the distinctive Rossmann fold of six β -strands⁵¹, alongside five neighboring α -helices. The C-terminal α/β domain (residues 118-368) consists of an eight-stranded β -sheet and seven α -helices. This long β -sheet predominately contributes to a flat-faced dimer interface and is a common feature of this protein family⁴⁷. The two monomeric chains of YjhC are nearly identical with a root mean square deviation (RMSD) of 0.128 Å (C α positions).

Within the N-terminal domain, the first residues of a glycine-rich motif in the turn between the first β -strand and α -helix serve as a fingerprint for cofactor specificity, interacting with the pyrophosphate bridge of the cofactor NAD(P). While NADP(H) binding enzymes normally possess a GxGxxA consensus sequence after the first β -strand^{52; 53}, a GxGxxG motif typically **correlates with** the binding of NAD(H)—the latter is observed in YjhC between residues G8 and G13. Most residues were well defined in the electron density map, except the sequence ₂₈₈SEMDGAIAYGHPGK₃₀₁ within the C-terminal domain, which was disordered in both

chains (yellow dotted line in **Figure 3A-B**). Furthermore, three residues at the C-terminus were disordered and thus not modelled in the final refined structure. However, using electrospray ionization mass spectrometry we verified that the full sequence (residues 1-372) was expressed and purified.

Cofactor binding

The crystal structure was solved in complex with the cofactor NAD(H). As shown in **Figure 3A** (insert), electron density for NAD(H) was unambiguously observed for the entire cofactor. NAD(H) was bound within a crevice formed between the loop regions of β -strands, $\beta 1$ and $\beta 4$ of the Rossmann fold within the N-terminal domain in which there was a change in the direction of the β -strand order (**Figure 3B**). The binding motif at this region, ${}_8\text{GxGxxG}_{13}$ confers cofactor specificity for NAD(H) over NADP(H)⁵⁴. Specifically, the cofactor was recognized by the side chains of residues F12, D33, N36, N70, H73, K91, W160 and K161, through an interaction that was mediated by an ordered water molecule (**Supplementary Figure 5**). Further, the cofactor was recognized by the main chain atoms of residues G10, Y11 and G13, through an ordered water molecule (**Supplementary Figure 5**). The side chain of residue D33 was observed to form a bifurcated hydrogen bond with the ribose of the adenine moiety. In addition to an increased thermal stability of the protein, which was observed in the presence of NAD⁺ over NADP⁺ (**Supplementary Table 1**), the locality of residue D33 provided further evidence for specificity towards the cofactor NAD(H) as it would sterically hinder the access of the 2'-phosphate group of NADP(H). The oxygen of the primary amide group, located within the nicotinamide moiety, was hydrogen-bonded to the side chain of residues E90, E314, and the main chain carbonyl oxygen of residue G117,

through an ordered water molecule (**Supplementary Figure 5**). In addition to this extensive hydrogen-bonding network, the nicotinamide ribose moiety was in an *anti*-C2'-endo conformation, which was stabilized through an aromatic edge-to-face interaction, or a C-H- π hydrogen bond with residue F12⁵⁵. The adenine ribose moiety also existed in an *anti*-C2'-endo conformation, and was predominately stabilized through a hydrophobic patch, which was formed by residues Y32, W68, P69, and L72 (**Supplementary Figure 5**). There was no observed difference between chains A and B with respect to the cofactor binding network. This observation was reinforced in the ensemble structure with residues G10, Y11, F12, G13, D33, N36, N70, H73, K91, and W160, along with NAD(H) itself, displaying little to no movement (**Figure 4C**). The only observed differences in the ensemble model was residue K161, which displayed increased flexibility relative to neighboring residues, and residue D90, which formed a hydrogen bond with the nitrogen of the amide group of NAD(H). Taken together, the well-ordered cofactor binding network highlights how tightly bound NAD(H) is within the crevice formed at the C-terminal edge of the Rossmann fold. Lastly, this binding mode is analogous to Gfo/Idh/MocA family members who bind NAD(H).

Ensemble refinement demonstrates mobile active site loops

To evaluate the disordered regions of YjhC we combined molecular dynamics simulations with crystallographic data, and generated an ensemble of models by time-averaged refinement, using *PHENIX ENSEMBLE REFINEMENT*³⁴. This refinement strategy extracts the local molecular motion and samples the global disorder within the protein, through which it can provide more insight into dynamic regions that are functionally important, rather than just an averaged 'snap-shot' of the atomic structure. In this study, the ensemble refinement

generated a Boltzmann-weighted population of 100 structures, with a significant reduction of the mF_o-DF_c electron density differences and improved R_{work} and R_{free} factors by 2.2% and 1.0%, respectively (**Supplementary Table 2**).

(Insert Figure 4)

Overall, the ensemble model revealed that, while YjhC displayed a well-ordered core, four loop regions (**Figure 4A-B**) along with the C-terminus displayed varied degrees of flexibility. The four flexible loop regions; L1-L4 are identified through the root mean square fluctuation (RMSF) for the backbone of both chain A and B, which is displayed in **Figure 4A**. While these loop regions are situated on the surface of the protein, they center around the substrate- and cofactor-binding cavities (**Figure 4B**). The flexibility in loops L1 and L3 may function to facilitate the diffusion of substrates to the binding cavity, with L3 capping the active site during catalysis—this suggested link between protein flexibility and catalysis has been hypothesized by others^{56; 57}. The fluctuations within the overall structure are visualized in **Supplementary Movie 1**.

E. coli YjhC is a dimer

Within the Gfo/Idh/MocA family the oligomeric assemblies of solved structures are predominately dimeric or tetrameric, although in one case a monomeric structure was reported⁵⁸. The structure of *E. coli* YjhC contained two molecules in the asymmetric unit, which form a dimeric assembly (**Figure 3A**), supported by a long, open flat-faced β -sheet in an antiparallel arrangement and *via* a two-fold rotation axis. Using the *PDBePISA* server⁵⁹, we determined that approximately 1666 Å² of accessible surface area was buried against the

interface of the opposing monomer, while the total estimated ΔG^{int} value of the dimer interface was approximately $-62.3 \text{ kcal mol}^{-1}$. Collectively, this formed a rigid, sandwich-like structure that was stabilized through face-to-face hydrophobic interactions, interdigitation of side chains from the opposing monomers, and a hydrogen-bonding network of nine residues per monomer (**Supplementary Figure 6**). In addition, two sulfate ions were modelled at this interface, where were both coordinated to residue R354 of chain A and B. Given their external location and a feature of the crystallization condition, these sulfate ions are unlikely to be of functional relevance.

To assess whether this dimeric assembly in the crystal structure was biologically relevant in solution, we used analytical ultracentrifugation and small angle X-ray scattering. Using analytical ultracentrifugation, three protein concentrations were analyzed (7, 14 and 20 μM) to assess if the dimeric assembly showed any signs of self-association. Following 2DSA-Monte Carlo analysis, all three concentrations were consistent with a single peak and supported a homogenous species in solution (**Figure 5A**). These data were consistent with the observation in SEC experiments, which showed a single monodisperse species (**Supplementary Figure 1**). The sedimentation coefficient of 5.60 S (95% confidence interval is 5.49- 5.72) was consistent with a dimeric species, while the frictional ratio (f/f_0) of 1.26 (95% confidence interval is 0.79-1.73), suggested that the molecule was slightly elongated. Further analysis using UltraScan predicted a molecular mass of 88.6 kDa, which was consistent with the calculated dimeric mass of 86.1 kDa (**Table 2**).

(Insert Figure 5)

(Insert Table 2)

Small angle X-ray scattering data were also consistent with a dimer in solution (**Figure 5B**, SASBDB accession code: SASDFZ3), when analyzed by the online server *SAXS-MoW2*⁴⁰. Measured on a relative scale, the molecular weight was 84.5 kDa, which was within 2% of the calculated dimeric mass (86.1 kDa). As shown in **Figure 5B**, the theoretical scattering profile, which was generated from the crystal structure of YjhC, fit the data well to high q , with a reduced X^2 value of 0.62. This supported that the dimeric assembly of YjhC, which was observed in the crystal structure was consistent with our solution structure. Guinier analysis produced a linear plot with a radius of gyration (R_g) of 31.14 Å and a forward scattering value (I_0) of 0.049 (**Table 2**). The linearity of the Guinier plot further supported a single monodisperse species in solution, which was free of any significant amount of aggregation or interparticle interference. An indirect Fourier transform of the scattering data gave a maximum interparticle dimension (D_{max}) of 107.1 Å (**Figure 5C**). This value agreed with the maximum interparticle dimension of ~107 Å, which was determined from the crystal structure. Overall, these data propose *E. coli* YjhC is a tight dimer, where the crystal structure closely models the solution structure.

Structural comparison with other Gfo/Idh/MocA family members

A search of the PDB using the *DALI* server²² was used to gain insight into the overall folding, structural features and possible catalytic functions of YjhC. The numerous search hits were reduced to seven by excluding PDB coordinate entries without a publication available and excluding PDB coordinate entries of identical protein sequence (**Table 3**). A multiple

sequence alignment was generated using these seven homologs to investigate consensus regions to and identity characteristic features that may uncover potential catalytic residues (**Supplementary Figure 7**). **Figure 6A** presents the sequence conservation among the search hits mapped onto the monomeric structure of YjhC using *UCSF Chimera*, while **Figure 6B** presents a structural superimposition of the monomeric assembly for each protein. All seven homologs were found to catalyze redox reactions with carbohydrate-based substrates. However, despite sharing a similar overall tertiary fold, the sequence identity between each structure was low compared to YjhC (<25%). This trend was observed across the Gfo/Idh/MocA family, which likely reflects their diverse catalytic function ⁴⁷.

(Insert Table 3)

YjhC contains the sequence ₈₃NKKHVFCEKP₉₂, which closely resembles the consensus motif AGKHVxCEKP—a fingerprint of sugar dehydrogenases ⁵⁰. Interestingly, within this motif we identified a conserved *cis*-peptide between the lysine and proline residue for all homologs including YjhC (residues K91 and P92). This *cis* conformation orientates the lysine residue towards the cofactor, where it can form a cation-*pi* interaction with the nicotinamide ring or hydrogen bond to the 2'-hydroxyl group of the nicotinamide ribose. Together, these interactions likely play a crucial role in maintaining the nicotinamide ring in the correct conformation during the catalytic reaction. In addition, this lysine residue is hypothesized to act as a molecular switch, preventing substrate rebinding by moving closer to the nicotinamide ring when the cofactor is reduced ^{16; 60}.

(Insert Figure 6)

Structural comparison of the binding cavity of these homologous proteins demonstrated that the binding cavity was open (**Figure 6C**), whereas in YjhC the binding cavity was found to be occluded (**Figure 6C**), primarily by disordered loop L3 within the C-terminal domain, which is highly flexible (**Figure 4B**). This loop may have a catalytic role and provide essential contacts to stabilize the substrate—a unique feature of YjhC among closely related structures.

Insight into YjhC substrate specificity at the binding cavity

Electrostatics play a key role in substrate binding, so we mapped the electrostatic potential onto the molecular surface of the substrate binding cavity for *E. coli* YjhC (**Supplementary Figure 8A**). Situated between the N- and C-terminal domains, this cavity had a positive charge distribution that would support the binding of acidic amino sugars, such as Neu5Ac and its derivatives. This observation was consistent with our data analysis using differential scanning fluorimetry (**Figure 2C-D**). Typically, within the Gfo/Idh/MocA family the catalyzed reaction precedes *via* a ping-pong mechanism, in which the tightly bound cofactor is required for the substrate to bind. Following binding, the substrate is either oxidized or reduced⁴⁷. To initiate the transfer of the hydride ion, the C4 position of the nicotinamide ring must be positioned near the substrate (~3.5 Å), while surrounding residues, depending upon the specific chemistry, can assist in the catalytic reaction.

Interestingly, we observed ambiguous density in the mF_o-DF_c map of the crystal structure that was close to the C4 position of the nicotinamide ring (**Supplementary Figure 8B**) and close enough to interact with the neighboring K91 and H176 residues. However, no molecule was modelled into this density, as nothing within the crystallization condition appeared to fit the

density when refined. Perhaps this could be an **artifact** from the expression and purification of the protein, with variable occupancy (<1.0).

To help deduce what the unmodelled density was and further identify potential substrates we performed *in silico* docking experiments using *SwissDock*⁴², interfaced with the *CHARMM* package⁴³ on Neu5Ac and several other sialic acid derivatives. These were selected based on their relative distribution in the human gastrointestinal tract, and within the small and large intestine of mice⁵. Neu5Ac and its open chain form, along with the 1,7 lactone and 9-O-acetylated sialic acid variant all presented favorable binding modes in which they were: 1) nestled within the binding cavity with favorable electrostatic interactions; 2) positioned close to the C4 position of the nicotinamide ring, poised for hydride transfer; and 3) orientated in a way that a redox reaction would be chemically plausible (**Figure 7**). These potential ligands were scored based on their estimated ΔG , number of hydrogen bonding partners, and full fitness (**Supplementary Table 3**), which was defined as the total energy of the system, along with a solvation term⁴³. As a control, to illustrate the efficiency of the docking strategy, NAD(H) was docked into YjhC with an RMSD difference to the crystal structure of less than 0.5 Å.

(Insert Figure 7)

Interestingly, the C4 position of Neu5Ac and the 9-O-acetylated variant were poised for hydride transfer and may suggest the position in which the hydride would be removed (oxidation) or added (reduction) during the catalytic reaction. In contrast, the open chain form of Neu5Ac, and the sialic acid 1,7 lactone variant both presented different carbon positions,

relative to the nicotinamide ring. The positioning of the open chain Neu5Ac (C7 position) and the sialic acid 1,7 lactone (C1 position), although not modelled as closely to the C4 position of the nicotinamide could likely be reoriented during conformational changes that inherently may occur during initial binding. In this case, disordered loop L3 (**Figure 5B**) may be involved in stabilizing the lactone during catalysis. Unfortunately, none of the compounds that were modelled in **Supplementary Figure 8**, fit the ambiguous density, which supported our hypothesis that this molecule was likely **an artifact** from the expression and purification of the protein.

Following the identification of several promising leads through docking experiments and differential scanning fluorimetry, attempts were made to co-crystallize the protein with the potential substrates (Neu5Ac and sialic acid 1,7 lactone), or alternatively soak the protein crystals with these molecules in the presence of the cofactor. Unfortunately, following data collection, there was no difference in the electron density within the binding cavity. However, ambiguous density was still observed close to the C4 position of the nicotinamide ring, as illustrated in **Supplementary Figure 8B**. We believe the presence of this unknown molecule(s) potentially hindered the binding of both Neu5Ac and sialic acid 1,7 lactone.

In addition, we have kinetically tested these promising leads using a spectrophotometric NADH assay, by measuring the change in absorbance at 340 nm upon addition of each molecule. Interestingly, this presented a very small rate for Neu5Ac and gave an apparent K_m of 68.8 mM (**Supplementary Figure 9**). No observable rate was measured in the absence of YjhC (control), or when using sialic acid 1,7 lactone as the substrate. It is probable that the

low observable rate reflects the ambiguous density observed in the binding cavity (Supplementary Figure 8B), and therefore may be inhibiting the reaction in YjhC.

Conclusion

Here, we report the characterization of YjhC from *E. coli* in an effort to elucidate its role in amino sugar metabolism. The results from our *in vivo* (BiOLOG Phenotype MicroArray) and *in vitro* (differential scanning fluorimetry assays) experiments argue that YjhC is involved in sialic acid metabolism. The structure of the protein, the presence of NADH in the active site, and binding studies demonstrate that the enzyme binds NAD(H) as a cofactor, consistent with its role as an oxidoreductase/dehydrogenase from the Gfo/Idh/MocA family. Unfortunately, we were unable to identify the substrate(s) for the enzyme, despite several promising leads, although sialic acid and closely related variants are potential substrates for YjhC.

References

1. Vimr ER, Kalivoda KA, Deszo EL, Steenbergen SM. 2004. Diversity of microbial sialic acid metabolism. *Microbiology and Molecular Biology Reviews*. 68(1):132-153.
2. Kelm S, Schauer R. 1997. Sialic acids in molecular and cellular interactions. San Diego: Elsevier Academic Press Inc. p. 137-240.
3. Almagro-Moreno S, Boyd EF. 2009. Insights into the evolution of sialic acid catabolism among bacteria. *BioMed Central Evolutionary Biology*. 9(1):118.

4. Angata T, Varki A. 2002. Chemical diversity in the sialic acids and related alpha-keto acids: An evolutionary perspective. *Chemical Reviews*. 102(2):439-469.
5. Vimr ER. 2013. Unified theory of bacterial sialometabolism: How and why bacteria metabolise host sialic acids. *International Scholarly Research Notices Microbiology*. 2013:816713.
6. Severi E, Hood DW, Thomas GH. 2007. Sialic acid utilisation by bacterial pathogens. *Microbiology*. 153(9):2817-2822.
7. Davies JS, Coombes D, Horne CR, Pearce FG, Friemann R, North RA, Dobson RCJ. 2019. Functional and solution structure studies of amino sugar deacetylase and deaminase enzymes from *staphylococcus aureus*. *FEBS Letters*. 593(1):52-66.
8. Vimr ER, Troy FA. 1985. Identification of an inducible catabolic system for sialic acids (nan) in *escherichia coli*. *Journal of Bacteriology*. 164(2):845-853.
9. Almagro-Moreno S, Boyd EF. 2009. Sialic acid catabolism confers a competitive advantage to pathogenic *vibrio cholerae* in the mouse intestine. *Infection and Immunity*. 77(9):3807-3816.
10. Chang DE, Smalley DJ, Tucker DL, Leatham MP, Norris WE, Stevenson SJ, Anderson AB, Grissom JE, Laux DC, Cohen PS et al. 2004. Carbon nutrition of *escherichia coli* in the mouse intestine. *Proceedings of the National Academy of Sciences of the United States of America*. 101(19):7427-7432.
11. Kalivoda KA, Steenbergen SM, Vimr ER. 2013. Control of the *escherichia coli* sialoregulon by transcriptional repressor nanr. *Journal of Bacteriology*. 195(20):4689-4701.
12. Condemine G, Berrier C, Plumbridge J, Ghazi A. 2005. Function and expression of an *n*-acetylneuraminic acid-inducible outer membrane channel in *escherichia coli*. *Journal of Bacteriology*. 187(6):1959-1965.

13. Wirth C, Condemine G, Boiteux C, Berneche S, Schirmer T, Peneff CM. 2009. Nanc crystal structure, a model for outer-membrane channels of the acidic sugar-specific kdgm porin family. *Journal of Molecular Biology*. 394(4):718-731.
14. Severi E, Müller A, Potts JR, Leech A, Williamson D, Wilson KS, Thomas GH. 2008. Sialic acid mutarotation is catalyzed by the *escherichia coli* β -propeller protein yjht. *Journal of Biological Chemistry*. 283(8):4841-4849.
15. Rangarajan ES, Ruane KM, Proteau A, Schrag JD, Valladares R, Gonzalez CF, Gilbert M, Yakunin AF, Cygler M. 2011. Structural and enzymatic characterization of nans (yjhs), a 9-o-acetyl n-acetylneuraminic acid esterase from *escherichia coli* o157:H7. *Protein Science*. 20(7):1208-1219.
16. Nurizzo D, Halbig D, Sprenger GA, Baker EN. 2001. Crystal structures of the precursor form of glucose-fructose oxidoreductase from *zymomonas mobilis* and its complexes with bound ligands. *Biochemistry*. 40(46):13857-13867.
17. Taberman H, Andberg M, Koivula A, Hakulinen N, Penttila M, Rouvinen J, Parkkinen T. 2015. Structure and function of *caulobacter crescentus* aldose-aldose oxidoreductase. *Biochemical Journal*. 472:297-307.
18. Kikuchi A, Park SY, Miyatake H, Sun DY, Sato M, Yoshida T, Shiro Y. 2001. Crystal structure of rat biliverdin reductase. *Nature Structural Biology*. 8(3):221-225.
19. Carbone V, Sumii R, Ishikura S, Asada Y, Hara A, El-Kabbani O. 2008. Structure of monkey dimeric dihydrodiol dehydrogenase in complex with isoascorbic acid. *Acta Crystallographica Section D-Biological Crystallography*. 64:532-542.
20. Bochner BR, Gadzinski P, Panomitros E. 2001. Phenotype microarrays for high-throughput phenotypic testing and assay of gene function. *Genome Research*. 11(7):1246-1255.

21. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped blast and psi-blast: A new generation of protein database search programs. *Nucleic Acids Research*. 25(17):3389-3402.
22. Holm L, Rosenstrom P. 2010. Dali server: Conservation mapping in 3d. *Nucleic Acids Research*. 38(Web Server issue):W545-549.
23. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Soeding J et al. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using clustal omega. *Molecular Systems Biology*. 7.
24. Robert X, Gouet P. 2014. Deciphering key features in protein structures with the new endsript server. *Nucleic Acids Research*. 42(W1):W320-W324.
25. Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A. 2003. Expasy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Research*. 31(13):3784-3788.
26. Kabsch W. 2010. Xds. *Acta Crystallographica Section D: Biological Crystallography*. 66(2):125-132.
27. Evans PR, Murshudov GN. 2013. How good are my data and what is the resolution? *Acta Crystallographica Section D-Biological Crystallography*. 69:1204-1214.
28. Panjikar S, Parthasarathy V, Lamzin VS, Weiss MS, Tucker PA. 2005. Auto-rickshaw: An automated crystal structure determination platform as an efficient tool for the validation of an x-ray diffraction experiment. *Acta Crystallographica Section D*. 61(4):449 -457.
29. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. 2007. Phaser crystallographic software. *J Appl Crystallogr*. 40(Pt 4):658-674.

30. Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F, Vagin AA. 2011. Refmac5 for the refinement of macromolecular crystal structures. *Acta Crystallographica Section D, Biological Crystallography*. 67(Pt 4):355-367.
31. Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, Terwilliger TC, Urzhumtsev A, Zwart PH, Adams PD. 2012. Towards automated crystallographic structure refinement with phenix.Refine. *Acta Crystallographica Section D-Structural Biology*. 68:352-367.
32. Emsley P, Cowtan K. 2004. Coot: Model building tools for molecular graphics. *Acta Crystallographica Section D, Biological Crystallography*. 60(Pt 12 Pt 1):2126-2132.
33. Chen VB, Arendall WB, III, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC. 2010. Molprobity: All-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D*. 66:12-21.
34. Burnley BT, Afonine PV, Adams PD, Gros P. 2012. Modelling dynamics in protein crystal structures by ensemble refinement. *eLife*. 1:29.
35. Demeler B. 2005. Ultrascan - a comprehensive data analysis software package for analytical ultracentrifugation experiments. In: Scott DJ, Harding SE, Rowe AJ, editors. *Analytical ultracentrifugation: Techniques and methods*. The Royal Society of Chemistry. p. 210-230.
36. Brookes E, Cao WM, Demeler B. 2010. A two-dimensional spectrum analysis for sedimentation velocity experiments of mixtures with heterogeneity in molecular weight and shape. *Eur Biophys J Biophys Lett*. 39(3):405-414.
37. Kirby N, Cowieson N, Hawley AM, Mudie ST, McGillivray DJ, Kusel M, Samardzic-Boban V, Ryan TM. 2016. Improved radiation dose efficiency in solution saxs using a sheath flow sample environment. *Acta Crystallographica Section D-Structural Biology*. 72:1254-1266.

38. Franke D, Petoukhov MV, Konarev PV, Panjkovich A, Tuukkanen A, Mertens HDT, Kikhney AG, Hajizadeh NR, Franklin JM, Jeffries CM et al. 2017. Atsas 2.8: A comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J Appl Crystallogr.* 50:1212-1225.
39. Konarev PV, Volkov VV, Sokolova AV, Koch MHJ, Svergun DI. 2003. Primus: A windows pc-based system for small-angle scattering data analysis. *J Appl Crystallogr.* 36:1277-1282.
40. Fischer H, Neto MD, Napolitano HB, Polikarpov I, Craievich AF. 2010. Determination of the molecular weight of proteins in solution from a single small-angle x-ray scattering measurement on a relative scale. *J Appl Crystallogr.* 43:101-109.
41. Svergun D, Barberato C, Koch MH. 1995. Crysol - a program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates. *J Appl Crystallogr.* 28(6):768-773.
42. Grosdidier A, Zoete V, Michielin O. 2011. Swissdock, a protein-small molecule docking web service based on eadock dss. *Nucleic Acids Research.* 39:W270-W277.
43. Grosdidier A, Zoete V, Michielin O. 2011. Fast docking using the charmm force field with eadock dss. *J Comput Chem.* 32(10):2149-2159.
44. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. Ucsf chimera - a visualization system for exploratory research and analysis. *J Comput Chem.* 25(13):1605-1612.
45. Chang YC, Hu ZJ, Rachlin J, Anton BP, Kasif S, Roberts RJ, Steffen M. 2016. Combrex-db: An experiment centered database of protein function: Knowledge, predictions and knowledge gaps. *Nucleic Acids Research.* 44(D1):D330-D335.
46. Schnoes AM, Brown SD, Dodevski I, Babbitt PC. 2009. Annotation error in public databases: Misannotation of molecular function in enzyme superfamilies. *PLoS Comput Biol.* 5(12):13.

47. Taberman H, Parkkinen T, Rouvinen J. 2016. Structural and functional features of the nad(p) dependent gfo/idh/moca protein family oxidoreductases. *Protein Science*. 25(4):778-786.
48. Bochner BR. 2009. Global phenotypic characterization of bacteria. *FEMS microbiology reviews*. 33(1):191-205.
49. Niesen FH, Berglund H, Vedadi M. 2007. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nature Protocols*. 2(9):2212-2221.
50. Wiegert T, Sahm H, Sprenger GA. 1997. Expression of the zymomonas mobilis gfo gene for nadp-containing glucose:Fructose oxidoreductase (gfor) in escherichia coli - formation of enzymatically active pregfor but lack of processing into a stable periplasmic protein. *European Journal of Biochemistry*. 244(1):107-112.
51. Rao ST, Rossmann MG. 1973. Comparison of super-secondary structures in proteins. *Journal of Molecular Biology*. 76(2):241.
52. Hanukoglu I, Gutfinger T. 1989. Cdna sequence of adrenodoxin reductase - identification of nadp-binding sites in oxidoreductases. *European Journal of Biochemistry*. 180(2):479-484.
53. Hua YH, Wu CY, Sargsyan K, Lim C. 2014. Sequence-motif detection of nad(p)-binding proteins: Discovery of a unique antibacterial drug target. *Scientific Reports*. 4:7.
54. Bellamacina CR. 1996. The nicotinamide dinucleotide binding motif: A comparison of nucleotide binding proteins. *Faseb J*. 10(11):1257-1269.
55. Nishio M, Umezawa Y, Fantini J, Weiss MS, Chakrabarti P. 2014. Ch- π hydrogen bonds in biological macromolecules. *Physical Chemistry Chemical Physics*. 16(25):12648-12683.
56. Garcia-Viloca M, Gao J, Karplus M, Truhlar DG. 2004. How enzymes work: Analysis by modern rate theory and computer simulations. *Science*. 303(5655):186-195.
57. Ramanathan A, Agarwal PK. 2011. Evolutionarily conserved linkage between enzyme fold, flexibility, and catalysis. *PLoS biology*. 9(11):e1001193.

58. Whitby FG, Phillips JD, Hill CP, McCoubrey W, Maines MD. 2002. Crystal structure of a biliverdin ix alpha reductase enzyme-cofactor complex. *Journal of Molecular Biology*. 319(5):1199-1210.
59. Krissinel E, Henrick K. 2005. Detection of protein assemblies in crystals. In: Berthold MR, Glen R, Diederichs K, Kohlbacher O, Fischer I, editors. *Computational life sciences, proceedings*. p. 163-174.
60. Schu M, Faust A, Stosik B, Kohring GW, Giffhorn F, Scheidig AJ. 2013. The structure of substrate-free 1,5-anhydro-d-fructose reductase from *Sinorhizobium meliloti* 1021 reveals an open enzyme conformation. *Acta Crystallogr Sect F Struct Biol Cryst Commun*. 69(Pt 8):844-849.
61. Thoden JB, Holden HM. 2011. Biochemical and structural characterization of wlba from *Bordetella pertussis* and *Chromobacterium violaceum*: Enzymes required for the biosynthesis of 2,3-diacetamido-2,3-dideoxy-d-mannuronic acid. *Biochemistry*. 50(9):1483-1491.
62. Sugiura M, Nakahara M, Yamada C, Arakawa T, Kitaoka M, Fushinobu S. 2018. Identification, functional characterization, and crystal structure determination of bacterial levoglucosan dehydrogenase. *Journal of Biological Chemistry*. 293(45):17375-17386.
63. Fukano K, Ozawa K, Kokubu M, Shimizu T, Ito S, Sasaki Y, Nakamura A, Yajima S. 2018. Structural basis of l-glucose oxidation by scyllo-inositol dehydrogenase: Implications for a novel enzyme subfamily classification. *PloS one*. 13(5).
64. Dambe TR, Kuhn AM, Brossette T, Giffhorn F, Scheidig AJ. 2006. Crystal structure of nadp(h)-dependent 1,5-anhydro-d-fructose reductase from *Sinorhizobium morelense* at 2.2 Å resolution: Construction of a nadh-accepting mutant and its application in rare sugar synthesis. *Biochemistry*. 45(33):10030-10042.

65. Kubiak RL, Holden HM. 2011. Combined structural and functional investigation of a c-3"-ketoreductase involved in the biosynthesis of dtdp-l-digitoxose. *Biochemistry*. 50(26):5905-5917.
66. North RA, Horne CR, Davies JS, Remus DM, Muscroft-Taylor AC, Goyal P, Wahlgren WY, Ramaswamy S, Friemann R, Dobson RCJ. 2018. "Just a spoonful of sugar...": Import of sialic acid across bacterial cell membranes. *Biophysical reviews*. 10(2):219-227.
67. DeLano WL. 2002. The PyMOL molecular graphics program. San Carlos: Delano Scientific.

prot_25846_figure_2.eps

prot_25846_figure_3.eps

prot_25846_figure_4.eps

prot_25846_figure_5.eps

prot_25846_figure_6.ai

prot_25846_figure_7.eps

Table 1 | Data collection statistics and refinement parameters ^a.

Structure	<i>E. coli</i> YjhC in complex with NAD(H)
Data collection statistics	
X-ray wavelength (Å)	0.9537
Temperature (K)	100
Detector	EIGER X 16 M (Dectris)
Crystal-to-detector distance (mm)	200
Space group	<i>P</i> 2 ₁ 2 ₁ 2 ₁
a, b, c (Å)	51.51, 84.67, 188.12
α, β, γ (°)	90, 90, 90
Resolution range (Å)	45.18-1.35 (1.398-1.35)
Volume fraction of solvent	50.39
V _m (Å ³ /Da)	2.49
CC _{1/2} [†]	0.999 (0.552)
Multiplicity	12.8 (7.8)
Wilson <i>B</i> -factor (Å)	18.70
Total No. of reflections	2,315,217 (154,583)
No. of unique reflections	181,008 (17,796)
<i>R</i> _{merge}	0.08 (2.11)
Data completeness (%)	99.0 (99.3)
Average <i>I</i> /σ (<i>I</i>)	14.2 (1.0)
Refinement statistics	
Resolution (Å)	45.18-1.35 (1.398-1.35)
Reflections used in refinement	180,786 (17,762)
Reflections used for <i>R</i> _{free}	9103 (915)
Total number of residues	710
Total non-H atoms	6629
Protein/ Water	5593/ 933
Ligands	103
Average temperature factors (Å ²)	
Overall	27.48
Protein/ Water	25.76/ 38.49
Ligands	28.18
<i>R</i> _{work} (%)	14.98
<i>R</i> _{free} (%)	16.87
Geometric RMSD	
Bond (Å)	0.008
Angle (°)	1.061
<i>MolProbity</i> statistics [‡]	
Rotamer outliers (%)	0.83
Clashscore	2.13
Ramachandran plot	
Favoured/allowed/disallowed regions (%)	97.58/2.42/0.0
PDB Code	6O15

^a Values in parentheses are for the highest resolution shell[†] CC_{1/2} is the correlation between random half-sets of data[‡] *MolProbity* is a structure-validation web service ³³

RMSD = root mean square deviation

Table 2 | Summary of data collection and analysis from the solution study of YjhC.

Sedimentation velocity analysis	
Sedimentation coefficient ($\times 10^{-13}$ S)	5.61 (5.51, 5.71)
Frictional ratio (f/f_0)	1.24 (0.97, 1.52)
Molecular weight (kDa)	8.67 (5.96, 11.38)
Partial specific volume of YjhC (mL/g)	0.7329
Buffer density (g/cm ³)	1.0009
Buffer viscosity (cp)	1.0086
SAXS data-collection parameters	
Instrument	Australian Synchrotron SAXS/WAXS beamline
Detector	PILATUS 1M (Dectris)
Wavelength (Å)	1.0332
Maximum flux at sample	8×10^{12} photons per second at 12 keV
Camera length (mm)	1600
q range (Å ⁻¹)	0.006-0.5
Exposure time	Continuous 1 second frame measurements
Sample configuration	SEC-SAXS with coflow
Sample temperature (°C)	12
SAXS data analysis	
Guinier analysis	
$I(0)$ (cm ⁻¹)	0.049 ± 0.00081
R_g (Å)	31.14 ± 0.36
$P(r)$ analysis	
$I(0)$ (cm ⁻¹)	0.049 ± 0.00014
R_g (Å)	31.80 ± 0.13
$D_{\max}$ (Å)	107.08
Porod volume (Å ⁻³)	130,000
CRY SOL analysis	
χ^2 value	0.62
SAXS-MoW2 ^a	
Estimated molecular weight (kDa)	84.5

^a <http://saxs.ifsc.usp.br/> ⁴⁰

Table 3 | Result of structural homology search using the *DAL*/server.

Protein (<i>abbreviation</i>)	Cofactor	Quaternary structure	Z score ^a	r.m.s.d (Å)	% sequence I.D.	PDB ID ^{^reference}
UDP-GlucNAcA Dehydrogenase (WlbA)	NAD	Tetramer	37.4	2.3	22	3Q2K ⁶¹
Levogluconan dehydrogenase (LGDH)	NAD	Tetramer	37.4	2.2	21	6A3I ⁶²
Scyllo-inositol dehydrogenase (IDH)	NAD	Tetramer	37.1	2.3	19	5YA8 ⁶³
1,5-anhydro-D-fructose reductase (AFR)	NADP	Dimer	36.4	2.3	23	2GLX ⁶⁴
Aldose-aldose oxidoreductase (AAOR)	NADP	Dimer	34.9	2.5	24	5A02 ¹⁷
Glucose-fructose oxidoreductase (GFOR)	NADP	Tetramer	34.4	2.6	24	1H6D ¹⁶
C-3'-ketoreductase (KijD10)	NADP	Tetramer	31.2	2.6	18	3RC2 ⁶⁵

^aPairwise similarity score measured by the *DAL*/server. Structural homologs are ranked by this score.

Dear Professor Garcia-Moreno

Many thanks to you and the reviewers for your constructive comments, much appreciated. Here we submit a revised manuscript, addressing, point by point, each of the reviewer's comments:

On the structure and function of *Escherichia coli* Yjhc: an oxidoreductase involved in bacterial sialic acid metabolism.

by Christopher R. Horne, Laura Kind, James S. Davies, Renwick C.J. Dobson

Please find here a detailed response to each of the points you highlighted in your letter. Below we have presented the reviewers comments in *italics*, our response in **red text**, and changes we have made in the manuscript to address these comments in **blue text**. All of the changes are included in a revised document that includes tracked changes. In addition, we have made minor editorial changes to improve the writing in the manuscript; these are also highlighted with track changes.

Reviewer: 1

1) *I am sure the authors have tried to co-crystallize the protein with potential substrates or soak the crystals with them. Most likely these trials failed, but they need to be discussed. Negative results are sometimes as instructive as positive results. If the authors have not tried, they should, so that they don't have to resort to the docking experiments in order to describe "substrate" binding.*

Response: **We have addressed this by including a new paragraph on page 19 to explain that both co-crystallization and soaking experiments (with cofactor) were performed but were unsuccessful. We believe the molecule(s) in the binding cavity adjacent to NAD(H), evidenced by the ambiguous density (Supplementary Figure 8B), hinders the binding of Neu5Ac and the sialic acid 1,7 lactone. The "Crystallization, data collection and structure determination of Yjhc" section of the Materials and Methods has been amended to include these trials.**

The text on page 19 (*Results and Discussion*) now reads:

Following the identification of several promising leads through docking experiments and differential scanning fluorimetry, attempts were made to co-crystallize the protein with the potential substrates (Neu5Ac and the sialic acid 1,7 lactone), or alternatively soak the protein crystals with these molecules in the presence of the cofactor. Unfortunately, following data collection, there was no difference in the electron density within the binding cavity. However, ambiguous density was still observed close to the C4 position of the nicotinamide ring, as illustrated in Supplementary Figure 8B. We believe the presence of this unknown molecule(s) potentially hinders the binding of both Neu5Ac and the sialic acid 1,7 lactone.

The text on page 8 (*Materials and Methods*) reads:

In an attempt to capture possible substrate leads (Neu5Ac and sialic acid 1,7 lactone) in the binding cavity of YjhC, crystals from condition A1 were soaked in cryo-protectant containing 80% reservoir solution, 10% glycerol-ethylene glycol (50:50), and 10% of either Neu5Ac or sialic acid 1,7 lactone each at a final concentration of 20 mM prior to data collection.

2) *It could also be that the mysterious electron density described in Supplementary material hinders the binding of the substrate. This also needs to be discussed.*

Response: We have addressed this in the new paragraph on page 18, as illustrated in response to the first comment by Reviewer 1.

3) *Concerning the “mysterious density”:*

Is this also present in the second crystal, which was measured to 1.94 Å resolution, and which is not reported in detail here? (Note aside: since the other crystal was measured in the presence of glycerol/ethylene glycol it may be worth adding this as a second structure to the manuscript).

And,

It would be rather interesting to have a look in this crystal form and also to have a look at an $F(\text{crystal 1})-F(\text{crystal 2})$ -map, provided the isomorphism between the crystals is high enough. It is possible that the glycerol/ethylene glycol combination washes out the stuff that leads to this mysterious density.

Response: The “mysterious density” is also present in the lower resolution dataset in which it looks virtually identical to the high-resolution dataset. Essentially, both structures are the same, as all crystals grew in the same crystallization condition, belong to the same space group, and have virtually the same unit cell dimensions. The only difference between the two datasets was the use of cryo-protection (glycerol/ethylene glycol) during collection of the lower resolution dataset—we identified just one glycerol on the surface, which was located away from the binding cavity.

We generated an isomorphous difference ($F(\text{crystal 1})-F(\text{crystal 2})$) map, using both the low- and high-resolution datasets and found no notable difference in the electron density within the binding cavity. Thus, the cryo-protectant does not seem to wash away the “mysterious density”. As such, we have chosen not to add this structure to the manuscript as it does not add any new knowledge. For the reviewer’s reference, we have included a figure illustrating the electron density of the high- and low-resolution datasets ($2F_o-DF_c$ is contoured to 1.0σ and mF_o-DF_c is contoured to $-3.0/3.0\sigma$, respectively), as well as the isomorphous difference map between each model (blue mesh = 1.0σ , green mesh = 3.0σ).

4) Further, what does the “mysterious density” look like after ensemble refinement?

Response: We have looked at the density after the ensemble refinement and can conclude that there is no notable difference in the appearance of the “mysterious density”. We note that some of the lower occupancy waters surrounding this region have moved inside and around the “mysterious density”, while all surrounding residues display very little dynamic fluctuations as shown in Figure 4.

5) The authors mention an EIGAR detector. The correct spelling is EIGER.

Response: We have corrected the spelling to EIGER in both the *Materials and Methods* section (“Crystallization, data collection and structure determination of YjhC”) and Table I.

6) The authors mention *pi*-stacking between ribose and F12. Ribose does not possess a *pi*-system, hence *pi*-stacking cannot occur. What the authors most likely want to describe is a hydrophobic stacking of the ribose unit against the aromatic *pi*-system of F12. Sometimes this is also described as C-H...Pi hydrogen bonds. The authors should have a look at some papers by M. Nishio et al. concerning such interactions.

Response: We have addressed this by removing all references to *pi*-stacking and now describe this interaction as an aromatic edge-to-face interaction or C-H- π hydrogen bond. We have also referenced this interaction (M. Nishio et al. 2014)(Reference 55).

The text on page 14 now reads:

In addition to this extensive hydrogen-bonding network, the nicotinamide ribose moiety is in an *anti*-C2'-endo conformation and is stabilized through an aromatic edge-to-face interaction, or C-H- π hydrogen bond with residue F12⁵⁵.

7) With an $I/\sigma(I)$ value of 2.7 in the highest resolution shell, the X-ray data could have been collected to higher resolution. Nowadays it is customary to include everything until $I/\sigma(I)$ of 1.0 or until a $CC(1/2)$ of 0.3 provided that the values are statistically significant (look for * in the XDS output). Any cutoffs based on R-values should be avoided, since they are statistically meaningless. A higher resolution of the structure may also be able to resolve the mysterious density issue (see above).

Response: We have reprocessed the data to 1.35 Å ($I/\sigma(I)$ is 1.0 and $CC_{1/2}$ is 0.552 in the outer shell), re-refined the structure and updated the coordinates within the PDB (ID 6O15). With respect to the XDS output, these cut-off values are statistically significant. We thank you for the suggestion. The necessary figures (3, 4 and supplementary figure 8), along with Table I have all been updated.

8) In Table III, it is not clear according to which criteria the entries of the table are ordered. Please specify.

Response: We have addressed this by rearranging Table III to rank entries by Z score and have included this ranking criterion in the table footer.

Reviewer: 2

1) There is one critical experiment that is missing. If you know the substrate is Neu5Ac or its lactone derivative, it is an oxidoreductase and needs NAD^+ , why was the reaction not carried out and the product identified. This is a simple experiment that will exactly identify the reaction catalyzed by the enzyme. This is an important experiment that will dramatically improve the value of the manuscript.

Response: We have attempted a kinetic assay on these promising leads using a spectrophotometric NAD^+ assay and indeed found that with Neu5Ac there is a very slow rate. For this experiment we collected initial rate data for Neu5Ac, which gave an apparent K_m of 68.8 mM. We note there was no observable rate for the sialic acid lactone, or in the absence of enzyme.

We think this very slow rate for Neu5Ac reflects the “mysterious density” observed in the binding cavity, which may be inhibiting the reaction in YjhC. We have included an additional figure to present this data (**Supplementary Figure 9**), added a new paragraph to page 19, and provided a brief description of the kinetic assay in the *Materials and Methods* section (page 10-11).

The text on page 19 (*Results and Discussion*) reads:

In addition, we have kinetically tested these promising leads using a spectrophotometric NADH assay, by measuring the change in absorbance at 340 nm upon addition of each molecule. Interestingly, this presented a very small rate for Neu5Ac and gave an apparent K_m of 65.1 mM (**Supplementary Figure 9**). No observable rate was measured in the absence of YjhC (control), or when using sialic acid 1,7 lactone as the substrate. It is probable that the low observable rate reflects the ambiguous density observed in the binding cavity (**Supplementary Figure 8B**), and therefore may be inhibiting the reaction in YjhC.

The text on page 10-11 (*Materials and Methods*) reads:

NAD kinetic assay

Kinetic analysis of YjhC was performed using a Cary 100 Bio UV/Vis spectrophotometer (Agilent Technologies), where the reaction was observed by measuring the increase in absorbance at 340 nm due to the consumption of NAD⁺. Assays were performed in duplicate, using acrylic cuvettes with a path length of 1 cm. All assay mixtures were equilibrated at 25 °C for 20 mins in the heat-controlled Peltier module prior to measurement. All other stock solutions were kept on ice. The assay mixture consisted of: 20 mM Tris–HCl pH 8.0, 150 mM NaCl, 2 mM NAD⁺ (Sigma-Aldrich) and Neu5Ac (Carbosynth), which varied from 5 mM to 60 mM (980 µL total). The assay was initiated following the additional of 20 µL YjhC at a final concentration of 0.05 µM. The assay was repeated using sialic acid 1,7 lactone as the substrate and 0.2 mM NADH (Sigma-Aldrich). In this experiment, the reaction was observed by measuring the decrease in absorption at 340 nm due to the conversion of NADH to NAD⁺. Controls without enzyme were performed to measure the background turnover of NAD⁺/NADH in solution. All Initial rate data were analysed and fitted with the Michaelis-Menten model using *GraphPad Prism 8.2.0* (GraphPad Software, La Jolla California USA).

2) A less critical experiment is a control of the *yjhC* knockout growth with a three-carbon sugar.

Response: We are uncertain what the rationale for this query is, but we do note that the BiOLOG Phenotype MicroArray experiment did include glycerol (PM1, B03), which presented no significant difference in respiration between the control and the *yjhC* gene knockout.

3) Minor comment: I am not sure what maximum resolution in crystallography for a refined structure is. Should we cut out the word maximum and just say, the structure was refined to a resolution of.....

Response: We have removed the word maximum and rewritten the sentence (page 4 and page 12). The re-processed resolution of 1.35 Å has been amended.

The sentence on page 4 now reads:

Further, we have solved the crystal structure of *E. coli* YjhC in complex with the cofactor NAD(H) at a resolution of 1.35 Å.

The sentence on page 12 now reads:

To add further weight to the hypothesis that YjhC is a *bona fide* oxidoreductase, as predicted from the bioinformatic analysis and thermal shift assays, the structure was solved at a resolution of 1.35 Å (Table I).

Many thanks for your consideration.

Yours sincerely,

Prof. Ren Dobson

Co-Director, Biomolecular Interaction Centre (BIC) <http://www.bic.canterbury.ac.nz/>

School of Biological Sciences | Te Kura Pūtaiao Koiora

University of Canterbury | Te Whare Wānanga o Waitaha

Private Bag 4800

Christchurch | Ōtautahi

New Zealand | Aotearoa

E: renwick.dobson@canterbury.ac.nz

T: ++64 (03) 369-5145 x 95145

F: ++64 (03) 364-2590

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Horne, CR;Kind, L;Davies, JS;Dobson, RCJ

Title:

On the structure and function of Escherichia coli YjhC: An oxidoreductase involved in bacterial sialic acid metabolism

Date:

2019-11-21

Citation:

Horne, C. R., Kind, L., Davies, J. S. & Dobson, R. C. J. (2019). On the structure and function of Escherichia coli YjhC: An oxidoreductase involved in bacterial sialic acid metabolism. PROTEINS-STRUCTURE FUNCTION AND BIOINFORMATICS, 88 (5), pp.654-668. <https://doi.org/10.1002/prot.25846>.

Persistent Link:

<http://hdl.handle.net/11343/286647>