

NIH Public Access

Author Manuscript

Proteins. Author manuscript; available in PMC 2009 December 11

Published in final edited form as: Proteins. 2002 August 1; 48(2): 427–430. doi:10.1002/prot.10160.

Crystal Structure of *Escherichia coli* EC1530, a Glyoxylate Induced Protein YgbM

Y. Kim¹, T. Skarina³, S. Beasley³, R. Laskowski⁴, C. Arrowsmith³, A. Joachimiak^{1,*}, A. Edwards^{2,3,*}, and A. Savchenko²

¹Biosciences Division, Structural Biology Center, Argonne National Laboratory, Argonne, Illinois

²Banting and Best Department of Medical Research, University of Toronto, Toronto, Ontario, Canada

³Clinical Genomics Centre/Proteomics, University Health Network, Toronto, Ontario, Canada

⁴Department of Crystallography, Birkbeck College, London, United Kingdom

Introduction

The crystal structure of YgbM (EC1530) (Fig. 1), a glyoxylate induced protein from *Escherichia coli*, has been determined and refined to 1.63 Å by multiple-wavelength anomalous dispersion (MAD) method. YgbM is encoded by DNA bases 2862259–2863035 and belongs to a protein family of Pfam-B_7694.¹ The gene is clustered with MutS (DNA mismatch repair protein), serine/threonine protein phosphatase, glycerol-3-phosphate regulon repressor, 3-hydroxyisobutyrate dehydrogenase, l-fuculose phosphate aldolase, gluconate permease, Rpos (RNA polymerase sigma factor), Nlpd (lipoprotein Nlpd), Pcm (protein-l-isoaspartate o-methyltransferase), *SurE* (stationary phase survival protein).

The Se-Met derivative of YgbM crystallized in the C2 space group with unit cell dimensions of a = 104.907 Å, b = 74.368 Å, c = 39.376 Å, $\beta = 98.81^{\circ}$. There is one 258-residue protein per asymmetric unit. This structure adopts the common TIM (triosephosphate isomerase) barrel (β/α)₈, in which an eight-membered cylindrical β -sheet is surrounded by eight helices.² Similar to other TIM barrel structures, all of the turns between the α -helices and the subsequent β -strands at the N-terminal end of the barrel are composed of only three or four residues, whereas the corresponding loops at the C-terminal end are longer and form a part of the potential active site. Inside of the TIM barrel, several hydrophilic side-chains from the C-terminal loops as well as two well-ordered water molecules coordinate to a Mg²⁺, presumably forming an active site (Fig. 2). As expected, a Dali search³ found several structures with relatively high similarity, which include 4XIS, 1A0C-A, 1QUM-A, 1DE5, and 1BYB with Z scores higher than 10. Further biochemical and structural analyses are in progress.

Materials and Methods

Protein Cloning Expression and Purification. The ORF of *ygbM* was amplified by PCR from *E. coli* genomic DNA (ATCC). The gene was cloned into the *NdeI* and *Bam*HI sites of a modified pET15b cloning vector (Novagen) in which the TEV protease cleavage site replaced

Atomic coordinates have been deposited into the Protein Data Bank (PDB) as 1k77.

^{© 2002} WILEY-LISS, INC.

^{*}Correspondence to: Dr. Andrzej Joachimiak, Biosciences Division, Structural Biology Center, Argonne National Laboratory, Argonne, IL 60439. andrzejji@anl.gov.

the thrombin cleavage site and a double-stop codon was introduced downstream from the BamHI site. This construct provides for an N-terminal hexa-histidine tag separated from the gene by a TEV protease recognition site (ENLYFQ \downarrow G). The fusion protein was overexpressed in E. coli BL21-Gold (DE3) (Stratagene) harboring an extra plasmid encoding three rare tRNAs (AGG and AGA for Arg, ATA for Ile). The cells were grown in LB at 37° C to an OD₆₀₀ of approximately 0.6 and protein expression induced with 0.4 mM IPTG. After induction, the cells were incubated overnight with shaking at 15°C. The harvested cells were resuspended in binding buffer (500 mM NaCl, 5% Glycerol, 50 mM HEPES pH 7.5, 5 mM imidazole), flashfrozen in liquid N_2 , and stored at -70° C. The thawed cells were lysed by sonication after the addition of 0.5% NP-40 and 1 mM each of PMSF and benzamidine. The lysate was clarified by centrifugation (27000g for 30 min) and passed through a DE52 column preequilibrated in binding buffer. The flow-through fraction was then applied to a metal chelate affinity column charged with Ni²⁺. The hexa-histidine tag was eluted from the column in elution buffer (500 mM NaCl, 5% Glycerol, 50 mM HEPES pH 7.5, 500 mM imidazole), and the tag then cleaved from the protein by treatment with recombinant His-tagged TEV protease. The cleaved protein was then resolved from the cleaved His-tag and the His-tagged protease by flowing the mixture through a second Ni²⁺-column.

The YgbM protein was dialyzed in 10 mM HEPES, pH 7.5, 500 mM NaCl, and concentrated by using a BioMax concentrator (Millipore). Before crystallization, any particulate matter was removed from the sample by passing through a 0.2- μ m Ultrafree-MC centrifugal filter (Millipore). For the preparation of selenomethionine (SeMet) enriched protein, the *E. coli* YgbM was expressed in the methionine auxotroph strain B834(DE3) of *E. coli* (Novagen) and purified under the same conditions as the native protein in supplemented M9 media. The reducing reagent β -mercaptoethanol (5 mM) was added to all purification buffers.

Protein crystallization

The protein was crystallized by vapor diffusion in hanging drops by mixing $2 \mu L$ of the protein solution (9 mg/mL) with $2 \mu L$ of 0.1 M HEPES, pH 7, 5% PEG 8000 and 5% glycerol, and equilibrated at 20°C over 100 μL of this solution. Crystals, which appeared after 3 days, were flash-frozen in liquid nitrogen with crystallization buffer plus 20% glycerol as cryoprotectant before data collection.

Data Collection

Diffraction data were collected at 100 K at the 19ID beamline of the Structural Biology Center at the Advanced Photon Source, Argonne National Laboratory. The three-wavelength inversebeam MAD data up to 1.63 Å [peak: 12.6620 KeV (0.97946 Å), inflection point: 12.6603 KeV (0.97957 Å), high-energy remote: 13.1000 KeV (0.93927 Å)] were collected from a Se-Met labeled protein crystal. One crystal ($0.2 \times 0.2 \times 0.2$ mm) was used to collect at 100 K all data MAD sets to 1.63 Å with 3 s exposure/1°/frame using 150 mm crystal to detector distance. The total oscillation range was 170 degrees as predicted with use of strategy module within HKL2000 suite.⁴ The space group was C2 with cell dimension of *a* = 104.907, *b* = 74.368, *c* = 39.376, β = 98.81°. All data were processed and scaled with HKL2000 (Table I) to an Rmerge of 7.2%, 7.0%, and 8.0% for inflection point, peak, and remote, respectively.

Structure Determination and Refinement

The structure was determined by MAD phasing⁵ using CNS⁶ and refined to 1.63 Å by using CNS against the averaged peak data. The initial model was built automatically by using ARP/wARP.⁷ The model was further refined to 1.63 Å. Throughout the model was manually adjusted by using O.⁸ The final R was 0.194 and the free R of 0.214 with all data (Table 2). Electron density calculated at 1.5 σ is well connected for all the main-chains and most of the side-chains

except a few areas on the surface of the molecules. The stereochemistry of the structure was checked with PROCHECK⁹ and the Ramachandran plot. The main-chain torsion angles for all residues are in allowed regions.

Acknowledgments

We thank all members of the Structural Biology Center at Argonne National Laboratory for their help in conducting experiments and Lindy Keller for help in preparation of this manuscript.

Grant sponsor: National Institutes of Health; Grant number: GM62414-01; Grant sponsor: U.S. Department of Energy, Office of Biological and Environmental Research; Grant number: W-31-109-Eng-38; Grant sponsor: Ontario Research and Development Challenge Fund.

REFERENCES

- 1. Bateman A, Birney E, Durbin R, Eddy SR, Howe KL, Sonnhammer EL. The Pfam protein families database. Nucleic Acids Res 2000;28:263–266. [PubMed: 10592242]
- 2. Wierenga RK. The TIM-barrel fold: a versatile framework for efficient enzymes. FEBS Lett 2001;492:193–198. [PubMed: 11257493]
- Holm L, Sander C. Dali: a netwok tool for protein structure comparison. Trends Biochem Sci 1995;20:478–480. [PubMed: 8578593]
- 4. Otwinowski Z, Minor W. Processing of x-ray diffraction data collected in oscillation mode. Methods Enzymol 1997;276:307–326.
- 5. Walsh MA, Dementieva I, Evans G, Sanishvili R, Joachimiak A. Taking MAD to the extreme: ultrafast protein structure determination. Acta Crystallogr 1999;D55:1168–1173.
- 6. Brunger AT, et al. Crystallography and NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr 1998;D54:905–921.
- 7. Perrakis A, Morris RJ, Lamzin VS. Automated protein model building combined with iterative structure refinement. Nat Struct Biol 1999;6:458–463. [PubMed: 10331874]
- 8. Jones TA, Zou JY, Cowan SW, Kjeldgaard M. Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr 1991;A47:110–119.
- 9. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr 1993;26:283–291.

Page 4

	٨	i.	
1	١	۱	

10	20											
		30	40	50	60	70	80	90	100	110	120	1
PRFAANLSMMFTEVPI	TERFAAAR	AGFDAVEFLF	TNISTLOI	ROLEONHLTL	ALFNTAPGD	INAGEWGLSA	PGREHEAHADI	DLALEYA	LALNCEQUHV	MAGVVPAGEL	ARRYA	
PRFAANLSMMFTEVPE	FIERFAAAR	AGFDAVEFLF	TNYSTLQI	KOLEONHLTL	ALFNTAPGD	INAGENGLSA	PGREHEAHADI	DLALEYA	LALNCEQUHVI	MAGVVPAGED	AERY A	FID
RFAANLSMMPTEVPI	TERFAAAR	AGFDAVEFLF	TOYSTLOI	ROLEONHLTL	ALFNTAPGD	INAGEWGLSA	PORCHEAHADI	DLALEYA	LALNCEQVHV	MAGVVPAGED	AERY A	VFID
RFAANLSMMPTEVP	FIERFAAAR	AGFDAVEFLFI	TOYSTLOI	KOLEONHLTL	ALFNTAPGD	INAGENGLISA	PGREHEAHADI	DLALEYA	LALNCEOVHVI	MAGVVPAGED	AERTA	VFID
							PGREHEARADI	DLALEYA	LALNCEQUHVI	MAGVVPAGED	AERY G	FIC
							PGREADSHEDI	DLALEYA	LALOCPNVHI	MSAVVPEGAS	REEYKQ	FIL
							PGREQOSHRDI	DLALEYA	LALDCPNVHI	MAAVVPEGAL	RELYKR	FV.
							PGREASFRRAI	AUDYA	OVLONRHIHV	MAGLLPAGAD	RIRHE	TYL
							PGREADFRGT	GRALEYA	GVIGNDRIHV	MAGLIPADAD	RARCA	TL
							PDRVQEFREG	DKALAYA	RVLGNRQVNC	LAGIRPOGHE	CASV K	FV
						MGLGA	PGROAAFREGI	QAVRYA	KALOCPRINL	MAGRVPQGAD	RIAVKAENA	VFL
140	150	160	170	180	190	200	210	220	230	240	250	;
	+	+	+	+	+	+		+	+	+	+	
		PHYLPSSOYOA										
		PHI PSS ICA							EINX PERL			
	ALCOPUTATION AL							and the second second	The state of the s			
A ADDRALL/PUBYLVS	B. C. DESTR.	BUY POONA								DEVGYOG	GCEYKPRGLT	EEG
AADRFALHGERILVE		PHYLPSSOYOA	TVEEVAR	NVFIQLDTFH	AQKVDGNLT	HLIRDYACKY	AHVQIAGLPOR	HEPOD-G	EINY LFRL	DEVGYQG I	GCEYKPRGLT	EEG
AADRFAPHGERILVE	ALSPGVK-	PHYLFSSOYOA	IVEEVAR	NVFIQLDTFH	AQKVDGNLT	HLIRDYAGKY	AHVQIAGLPOR	HEPOD-G	EINYPELFRL EINYPELFRL	DEVGYQG I DEVGYQG I	GCEYKPRGLT GCEYKPRGLT	EEG
AADRFAPHGERILVE ASDEYEPYGIEIQLE	ALSPEVK-	PHYLFSSOYOAI PNYLLESOFDTI	VELVDR	ONVFIQLDTFH	AQKVDGNLT AQKVDGNLT AQNVDGNLA	HLIRDYAGKY HLIRDYAGKY RLIRDYAGKY	AHVQIAGLPOR AHVQIAGLPOR AHVQIAGLPOR	HEPOD-G HEPOD-G HEPOE-G	EINYPWLFRL EINYPWLFRL EINYQYIFDK	DEVGYQG I DEVGYQG I DEVGYQG I DEIGYTG V	GCEYKPRGLT GCEYKPRGLT GCEYKPRGLT	EEGI EEGI VTGI
(AADRFAPHGERILVE (ASDKYKPYGIRIQLE (AAEKFKPHGIRILLE	ALSPGVK- ALSPEVK- ALSPQVK-	PHYLFSSOYOAI PNYLLKSOFDTI PNYLLKSOFDTI	AIVEEVARI	ONVFIQLDTFH ONVFIQLDTFH ONVFIQLDTFH	AQKVDGNLT AQKVDGNLT AQNVDGNLA AQNVDGNLA	-HLIRDYAGKY -HLIRDYAGKY -RLTDKLWGKH -QLTDLLKDKY	CAHVQIAGLPOR CAHVQIAGLPOR CAHVQIASVPOR CAHVQIASVPOR	HEPOD-G HEPOD-G HEPOK-G HEPNE-G	EINYPWLFRL EINYPWLFRL EINYQYIFDK EINYTYLFDK	DEVGYQG I DEVGYQG I DEVGYQG I DEIGYTG V DQNGYRG I	GCEYKPRGLT GCEYKPRGLT GCEYKPRGLT GCEYKPRGET GCEYKPRGET	EEGI EEGI VTGI
AADRFAPHGERILYE ASDEYEPYGIRIQLE AAEEFEPHGIRILLE AAREAAGAGLTVLIE	ALSPGVK- ALSPEVK- ALSPQVK- PINNRDM-	PHYLFSSOYOAI PNYLLKSOFDTI PNYLLKSOFDTI PGFFLTRODDA(IVEEVARI	NVFIQLDIFH NVFIQLDIFH NVFIQLDIFH NVFIQLDIFH NVFIQLDIFH	AQKVDGNLT AQKVDGNLT AQNVDGNLA AQNVDGKLS CQIVEGDLA	HLIRDYAGKY HLIRDYAGKY RLTDKLMGKI QLTDLLKDKY TTLRRDIGNY	CAHVQIAGLPOR CAHVQIAGLPOR CAHVQIASVPOR CHIQIASVPOR CHVQIAGVPOR	HEPOD-G HEPOD-G HEPOK-G HEPNE-G HEPODDN	EINYPHLFRL EINYPHLFRL EINYGYIFDR EINYTYLFDR EVNYPYLFEL	DEVGYQG I DEVGYQG I DEVGYQG I DEIGYTGYV DQNGYRG I DRLGYQG I	GCEYKPRGLT GCEYKPRGLT GCEYKPRGLT GCEYKPRGET GCEYKPKIET	EEGI EEGI VTGI VTGI
AADRFAPHGKRILVE ASDKYKPYGIRIQLE AABKFKPHGIRILLE AAREAAGAGLTVLIE AARAAAAQGTVLIE	ALSPGVK- ALSPEVK- PINNRDR- PINNRDR- PINTRDN-	PHYLESSOYOAI PNYLLKSOFDTI PNYLLKSOFDTI PGFFLTRODDA(PGYFLNRODOG(A IVEEVARI A IVEEVARI VVELVDRI D IVNLVDRI A ICREVGAI	DNVFIQLDTFH DNVFIQLDTFH DNVFIQLDTFH DNVFIQLDYFH DNVFIQLDYFH DNLQVQFDCYH	AQKVDGNLT AQKVDGNLT AQNVDGNLA AQNVDGKLS CQIVEGDLA CQIVEGDVA	-HLIRDYAGK) -HLIRDYAGK) -RLIDYAGKI -QLIDLLKDKU -TTLRRDIGN) MKLKRDIAG-J	HVQIAGLP R HVQIAGLP R HVQIASVP R HIQIASVP R HIQIAGVP R HIQIAGVP R	HEPOD-G HEPOD-G HEPOE-G HEPNE-G HEPODDN HEPOV-G	EINYPWLFRL EINYPWLFRL EINYPYLFRL EINYTYLFDK EVNYPYLFEL ELNYPYLFEV	DEVGYQG I DEVGYQG I DEVGYQG I DEIGYTGYV DQNGYRG I DRLGYQG I DTLGYDG I	GCEYKPRGLT GCEYKPRGLT GCEYKPRGET GCEYKPRGET GCEYRPRGDT GCEYRPRGDT	EEGI REGI VTGI SAGI
AADRFAPHGERILYE ASDEYEPYGIRIQLE AAEEFEPHGIRILLE AAREAAGAGLTVLIE	CALSPEVK- CALSPEVK- CALSPEVK- CALSPEVK- CONTROM- CONTROM- CONTROM- CONTROM-	PHYLFSSQYQAI PNYLLKSQFDTI PGFFLTRQDDA PGFFLTRQDDA PGYFLNRQDDG PGFYLQNTRQAI	IVEEVARI VEEVARI VELVDRI IVELVDRI ICREVGAI	NVFICLDTFH NVFICLDTFH NVFICLDTFH NVFICLDTFH NLCVCFDCYH NLCVCFDCYH	AQKVDGNLT AQKVDGNLT AQNVDGNLA AQNVDGNLS CQIVEGDLA CQIVEGUNA MQIMEGDLA	-HLIPDYAGK) -HLIPDYAGK) -RLTOKLNGKI -QLTOLLKDKV -TTLRRDIGN) MKLKRDIAG-J -RTLENQLAAJ	AHVQIAGLP R AHVQIAGLP R AHVQIASVP R AHVQIASVP R AHIQIASVP R AHIQIAGVP R CHIQIAGVP R CHIQIAGVP R	HEPOD-G HEPOD-G HEPOE-G HEPNE-G HEPODDM HEPOV-G HEPGT-G	EINYPALFRL EINYPALFRL EINYPYLFDK EINYPYLFBL ENYPYLFBL ELNYPYLFEV EINYAFLFEH	DEVGYQG I DEVGYQG I DEVGYQG I DELGYTG V DQNGYRG I DRLGYQG I DRLGYQG I	GCEYKPRGLT GCEYKPRGLT GCEYKPRGET GCEYKPRGET GCEYRPRGT GCEYRPRGT GCEYRPRACT	EEGI EEGI VTGI SAGI SAGI
	RPAANLSNOPTEVP RFAANLSNOPTEVP RFAANLSNOPTEVP RFAANLSNOPTEVP RFAANLSNOPTEVP RFAANLSNOPTEVD RFAANLSNOPTEVD RFAANLSNOPTEVD 140 140	ВРААНСКИОТТВУР ГИВРАЛВ ВРААНСКОТТВУР ГИВРАЛВ ВРААНСКОТВУР ГИВРАЛВ ВРААНСКИ УРНОГИВРАЛВ ВРААНСКИ, УРНОГИВРАТИВИ ВРААНСКИ, УРНОГИВИТИВИ, УРНОГИВРАТИВИ ВРААНСКИ, УРНОГИВИТИВИ ВРААНСКИ, УРНОГИВИТИВИ, УРНОГИВИТИВИ, УРНОГИВИ ВРААНСКИ, УРНОГИВИСТИВИ, УРНОГИВИТИВИ ВРААНСКИ, УРНОГИВИ ВРАСИНСКИ, УРНОГИВИ ВРААНСКИ, УРНОГИВИТИВИ, УРНОГИВИТИВИ, УРНОГИВИ ВРААНСКИ, УРНОГИВИТИВИ, УРНОГИВИ ВРААНСКИ, УРНОГИВИТИВИ, УРНОГИВИ ВРААНСКИ ВРАСИНСКИ ВРАСИНСКИ, УРНОГИВИ ВРААНСКИ ВРАСИНСКИ ВРАСИНСКИ ВРАСИНСКИ ВРАВИСИ ВРАВИСКИ, УРНОГИВИ ВРАСИНСКИ ВРАВИСИ ВРАСИНСКИ ВРАВИСИ ВРАВИСИ ВРАВИЛИ ВРАСИНСКИ ВРАСИНСКИ ВРАВИСИ ВРАВИСИ ВРАВИЛИ ВРАВИСИ ВРАВИСИ ВРАВИСИ ВРАВИСИ ВРАВИСИ ВРАВИСИ ВРАСИНСКИ ВРАВИСИ ВРАВИСИ ВРАВИСИ ВРАВИСИ ВРАВИСИ ВРАВИСИ ВРАВИСИ ВРАВИСИ ВРАВИСИ ВРАВИСИ ВРАВИСИ ВРАВИЛИ ВРАВИСИ ВРАВИСИ ВРАВИ ВРАВИСИ ВР	RPAALLSKETTEVPFIERPAARFAGTONVERLEF RPAALLSKETTEVPFIERPAARFAGTONVERLEF KPAALTSKETTEVPFIERPAARFAGTONVERLEF RPAALTSKETTEVPFIERPAARFAGTOVVERLEF RPAALLSKETTEVPFIERPAARFAGTOVVERLEF RPAALSKETTEVPFIERPAARFAGTOVVERLEF RPAALSKETTEVPFIERPAARFAGTOVVERLEF RPAALSKETTEVPFIERPAARFAGTOVVERLEF RPAALSKETTEVFIERPAARFAGTOVVERLEF RPAALSKETTEVFIERPAARFAGTOVVERLEF RPAALSKETTEVFIERPAARFAGTOVVERLEF RPAALSKETTEVFIERPAARFAGTOVVERLEF RPAALSKETTEVFIERPAARFAGTOVVERLEF RPAALSKETTEVFIERPAARFAGTOVVERLEF RPAALSKETTEVFIERPAARFAGTOV 140 150 160 ADBRANGKETLVERLEPOVX-PHYLFESSOVAL	BPAAMLGNUPTEVPFIERPAAREKAGPDAVERUPTUTSIO BPAAMLGNUPTUPTEERPARKIAGPDAVERUPTUTSIO KPAAMLSNUPTUPTUPTUBERPAARIAGPDAVERUPTUTSIO KPAAMLTNUPTUPENPFLORESKISKIGPTUVTUSIO RPAAMLGNUPTUPENPFLORESKISKIGPTUVTUSIO RPAAMLSNUPTUPENPFLORESKISKIGPTUVTUSIO RPAAMLSNUPTUPENPFLORESKISKIGPTUVTUSIO RPAAMLSNUPTUPENPFLORESKISKIGPTUTUSIO RPAAMLSNUPTUPENPFLORESKISKIGPTUTUSIO RPAAMLSNUPTUPENPFUDIARENAARIGERVERTUPENPFTUSIAARIJ RPAAMLSNUPTUPENPFUDIARENAARIGERVERTUPENPFTUSIAARIJ RPAAMLSNUPTUPENPFUDIARENAARIGERVERTUPENPFTUSIAARIJ RPSAMLSNUPTUPENPFUDIARENARIGERVERTUPENPFTUSIAARIJ RPSAMLSNUPTUPENPFTUPIARENARIGERVERTUPENPFTUSIAARIJ RPSAMLSNUPTUPENPFTUPIARENDIARENTUPENPFTUSIARI RPSAMLSNUPTUPENPFTUPIARENTUPENPFTUSIARI RPSAMLSNUPTUPENPFTUPIARENTUPENPFTUSIARI RPSAMLSNUPTUPENPFTUPIARENTUPENPFTUSIARI RPSAMLSNUPTUPENPFTUPIARENTUPENPFTURI	BPAAMLSGNOPTEVPFIERPAARLAGEDAVEFLEPTYTSTICIOGEGUSUHLT. BPAAMLSGNOPTEVPFIERPAARLAGEDAVEFLEPTYTSTICIOGEGUSUHLT. KPAAMLSGNOPTEVPFIERPAARLAGEVIVVELANDTYTSTISOIGUGUSUHLT. KPAAMLSGNOPTEVPFIERPAARLAGEVIVVELANDTYTSTISOIGUSUHLT. KPAAMLSGNOPTEVPFIERPAARLAGEVIVVELANDTYTSTISOIGUSUHLT. KPAAMLSGNOPTENPFIERPAARLAGEVIVVELANDTYTSTISOIGUSUHLT. REAAMLSGNOPTENPFIERPAARLAGEVIVVELANDTYTSSIGUSUHLENGUSUUSUHLENGU	BPAAMLGSOFTEVP FIERPAAREAGENAVEELPPT015010000E00HLTALEPTANE0 BPAAMLGSOFTEVP FIERPAARAAGENAVEELPPT0150010000V0HLTALEPTANE0 BPAAMLSOFTAPT0P FIERPAARAAGENAVEELPPT0150010000V0HLTALEPTANE0 BPAAMLSOFTAPT0P FIERPAARAAGENAVEELPPT0150010000V0HLTALEPTANE0 BPAAMLSOFTAPT0P FIERPAARAAGENAVEELPPT015001000F0HT0AGENAUEPT0PA00 BPAAMLSOFTAPT0PFLOREFAARAAGENAVEELPPT010000LAREARENOUGUT0LEPTANO BPAAMLSOFTAPT0FIERPT010000LAREARENOUGUT0LEPTANE0 BPAAMLSOFTAPT0FIERPT01000LAREARENOUGUT0LEPTANE0 BPAAMLSOFTAPT0FIERPT0AAAAGENAUEPT01000LAREARENOUGUT0LEPTANACE BPAAMLSOFTAPT0FIERPT0AAAAGENAUEPT01000LAREARENOUGUT0LEPTANACE BPAAMLSOFTAPT0FIERPT0AAAAGENAUEPT01000LAREARENOUGUT0LEPTANACE BPAAMLSOFTAPT0FIERPT0AAAAGENAUEPT01000LAREARENOUGUT0LEPTANE0 BPAAMLSOFTAPT0FIERPT0AAAAGENAUEPT0100LAREARENOUGUT0LEPTANACE BPAAMLSOFTAPT0FIERPT0AAAAGENAUEPT0100LAREARENOUGUT0LEPTANACE BPAAMLSOFTAPT0FIERPT0AAAAGENAUEPT0100LAREARENOUGUT0LEPTANACE BPAAMLSOFTAPT0FIERPT0AAAAGENAUEPT010AAGENAUEPT010AEEPT010AEEPT010AEEPT010AEEPT010AEEPT010AEEPT010AEEPT00AEET0000LAT00EEPT00AEET00AEEPT010AEEPT010AEEPT010AEEPT010AEEPT010AEEPT010AEEPT010AEEPT010AEEPT010AEEPT010AEEPT010AEEPT000AEET000EPT000EET000EPT010AEEPT010EEPT010EEPT010EEPT010EEPT010EEPT0	AADRFAFHGKRILVEALSFGVK-PHYLFSSGYGALAIVERVARDNVFIGLDYFHAGKVDGHLI-HLIIDYAGK AADRFAFHGKRILVEALSFGVK-PHYLFSSGYGALAIVERVARDNVFIGLDYFHAGKVDGHLI-HLIIDYAGK	BPAAMLSGNOPTEVPFIERPAARLAGEDAVEFLEPYVTSTOLOGKOLSGNULTALEPTA-CDINAGENGLAALPORINGHALADD BPAAMLSGNOPTEVPFIERPAARLAGEDAVEFLEPYVTSTOLOGKOLSGNULTALEPTA-CDINAGENGLAALPORINGHALDD BPAAMLSGNOPTEVPFIERPAARLAGENVEVVIERUNTVYTSTOLOGKOLSGNULTALEPTA-CDINAGENGLAALPORINGHALDD BPAAMLSGNOPTEVPFIERPAARLAGENVEVVIERUNTVYTSTSSULGKULTUNGUSULTULEPTA-CDINAGENGLAALPORINGHALDD BPAAMLSGNOPTENPFIERPAARLAGENVIERUNTVEPYVTSASULGKULTUNGUSULTULEPTA-CDINAGENGLAALPORINGHALDD BPAAMLSGNOPTENPFIERPAARLAGENVIERUNTVEPYVTSASULGKULTUNGUSULTULEPTA-CDINAGENGLAALPORINGHANANANANANANANANANANANANANANANANANANA	RPAAMLSHOTEVPFIERPAARLAGPDAVELEPPTYTSTED OKOLEGIHLITLALPITA-ODINAGENGIAALOGENAL	RPAAMLGBOOTEVPFIERPAARELOGTONVEELPYTOTSJOLOGOLUGU (NELLELLENTATOGDINGGBOLDALPPERHAMMADDICALEXTANA, WITTENDENDERVERUNG (NELLENTATOGDINGGBOLDALPPERHAMMADDICALEXTANA, WITTENDENDERVERUNG) RPAAMLSBOOTEVPFIERPAARLOGDINVEELPYTOTSJOLOGOLUOUVOILT-HUNTATOGDINGGBOLDALPPERHAMMADDICALEXTANA, WITTENDENDERVERUNG (NELLENTATOGDINGGBOLDALPPERHAMMADDICALEXTANA, WITTENDENDERVERUNG) RPAAMLSBOOTEVPFIERPAARLOGDINVEELPYTOTSJOLOGOLUOVOILT-HUNTATOGDINGGBOLDALPPERHAMMADDICALEXTANA, WITTENDENDERVERUNG (NELLENTATOGDINGGBOLDALPPERHAMBEDICALEXTANA, WITTENDENDERVERUNG) RPAAMLSBOOTEVPFIERPAARLOGDINVEELPYTOTSJOLOGOLUOVOILT-HUNTATOGDINGGBOLDINGENDU (NELEXTANA, GORDINGERIAS) RPAAMLSBOOTEVPFIERPAARLOGDINVEELPYTOTSJOLOGOLUOVOILT-HUNTATOGDINGGBOLDINGERIAS) RPAAMLSBOOTENTEPFFIERPAARLOGDINGUUTELPYTOTSJOLOGOLUONUUTELPYTOTSJOLUUTENDINGGBOLLENGTALTANA, GORDINGGBOLLAN, TORANGONUUTELPYTOTSJOLUUTENDINGUUTELPYTOTSJOLUUTENDINGGBOLLONUUTELPYTOTSJOLUUTENDINGGBOLUONUUTENDINGGBOLLAN, TORANGONUUTELPYTOTSJOLUUTENDINGUUTENDI (NELEXTANA), GORDINGGBOLLAN, TORANGONUUTELPYTOTSJOLUUTENDINGGBOLUONUUTENDI (NELEXTANA), GORDINGGBOLLAN, TORANGONUUTENDI (NELEXTANA), GORDINA, GORDI	RPAAMISHED TEVPFIER PAARE LAGFDAVE EVEPYOTSTED (DKOLED "HETALEPYTA"CD: NAGENGLAAL POR HEARADCI LA LEVALAD" INCOMPAGNU VAGE RPAAMISHED TEVPFIER PAARE LAGFDAVE EVEPYOTSTED (DKOLED "HETALEPYTA"CD: NAGENGLAAL POR HEARADCI LA LEVALAD" INCOMPAGNU VAGE RPAAMISHED TEVPFIER PAARA LAGFDAVE EVEPYOTSTED (DKOLED "HETALEPYTA"CD: NAGENGLAAL POR HEARADCI LA LEVALAD" INCOMPAGNU VAGE RPAAMISHED TEVPFIER PAARA LAGFDAVE EVEPYOTSTED (DKOLED "HETALEPYTA"CD: NAGENGLAAL POR HEARADCI LA LEVALAD" INFO RPAAMISHED TEVPFIER PAARA LAGFDAVE EVENTSTED (DKOLED "HETALEPYTA"CD: NAGENGLAAL POR HEARADCI LA LEVALAD" INTO RPAAMISHED TEVPFIER PAARA LAGFDAVE EVENTSTED (DKOLED "HETALEPYTA"CD: NAGENGLAAL POR HEARADCI LA LEVALAD" INFO RPAAMISHED TEVPFIER PAARA LAGFDAVE EVENTSTED (DKOLED "HETALEPYTA" DC) NAGENGLAAL POR HEARADCI LA LEVALAD" INTO RPAAMISHED TEVPFIER PAARA LAGFDAVE EVENTSTED (DKOLED "HETALEPYTA" DC) NAGENGLAAL POR HEARADCI LA LEVALAD" INTO RPAAMISHED TEVPFIER PAARA LAGFDAVE EVENTSTED (DKOLED "HETALEPYTA" DC) NAGENGLAAL POR HEAGTO TAK HEAVEN RPAAMISHED TEVPFIER PAARA LAGFDAVE EVENTSTED (DKOLED "HETALEPYTED "DKOLED "HETALEPYTED	RPAAMLSHOPTEVPFIERPAARELAGTONVELEPTVTSTLOIGKOLSUPHITALEUTIKAIRUS INAGEGISALPORELIHANKDI LAALEVALUUT UVENUVVYADUVYADUUT - AIEUTEK RPAAMLSHOPTUVPFIERPAARELAGTONVELEPTVTSTOIGOOLOUVUPHITALEUTIKAIRUS INAGEGISALPORELIHANKDI LAALEVALUUT USUVVADUVYADUUT - AIEUTEK RPAAMLSHOPTUPFIERPAARELAGTONVELEPTVTSTOIGOOLOUVUPHITALEUTIKAIRUS INAGEGISALPORELIHANKDI LAALEVALUUTUUSUVVADUVYADUUT - AIEUTEK RPAAMLSHOPTUPFIERPAARELAGTONVELEPTVTSTOIGOOLOUVUPHITALEUTIKAIRUS INAEKSANAU REVANUTIONEVPELEREARELAGTONVELEPTVTSTOISOULUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUU

ygbM	DA-WRGS	Ec1530 (apc066) - ygbm	
4738	/DA-WR	sp Q46891 YGBM_ECOLI HYPOTHETICAL 29.2 KDA	{258 ai
97687	DA-WR	gb AAG14970.1 AF242208 3 (AF242208) unknow	(258 aa)
97702	DA-WR	gb AAG14983.1 AF242210 2 (AF242210) unknow	(258 aa)
97694	DA-WR	gb AAG14976.1 AF242209 1 (AF242209) unknow	(258 aa)
4739	PQK-YK	sp Q57151 YGBM HAEIN HYPOTHETICAL PROTEIN H	(258 a)
21734	AP-YKDA	gb AAK03447.1 (AE006174) unknown [Pasteur	(260 aa)
47672	LRR-WRATERDG-	pir D83577 conserved hypothetical protein	(265 aa)
4740	LEP-YLGR	sp Q44015 YGB4 ALCEU HYPOTHETICAL 28.3 KDA	(260 a)
47790	XET-HNVV	pir A83457 conserved hypothetical protein	(260 aa)
70523	RD-LTGOGSAAA	ref NP 102092.1 hypothetical protein [Mes	(265 aa)
54272	LRSYNDRRGHPEAGQ	ref NP 112484.1 hypothetical protein HT03	(204 aa)

B

gi 249 gi 101 gi 101 gi 249 gi 127 gi 127 gi 113 gi 249 gi 113 gi 134 gi 136

С

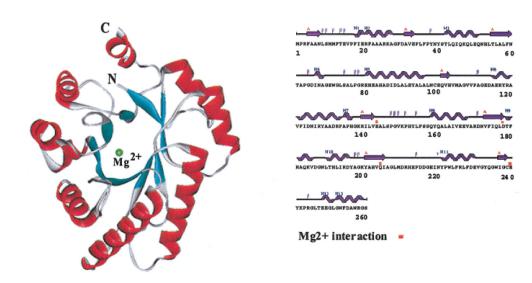


Fig. 1.

A: Protein sequences are compared between different species. B: The ribbon drawing shows the TIM structure of EC1530. α -helices, outside of the barrel, are shown in red, β -strands, inside of the barrel in cyan, and one of two Mg²⁺ (the major site) in the barrel is shown in green. The Mg²⁺ in the minor site located on the outside surface is not shown. C: The secondary elements were indicated above the one-letter amino acid codes of *E. coli* ygbM. Residues interacting with Mg²⁺ are also indicated in red dots.

Kim et al.

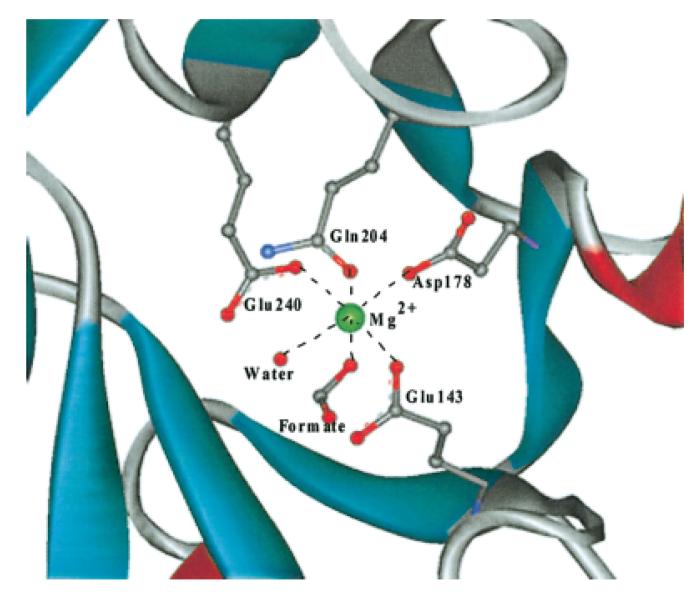


Fig. 2. Putative catalytic site including Mg^{2+} (major site) is shown. The Mg^{2+} is coordinated to an ordered water molecule, a formate, two glutamate, a glutamine, aspartate residues forming a square-bi-pyramid conformation.

Kim et al.

Summary of Crystal and MAD Data

Unit cell Space group MW Da (residues) Mol (AU) SeMet (AU)	a = 104.907 Å, b = C2 29200(258) 1	$ \begin{array}{l} a = \ 104.907 \ \bar{A}, \ b = 74.368 \ \bar{A}, \ c = 39.376 \ \bar{A}, \ \beta = 98.81^\circ \\ C2 \\ 29200(258) \\ 1 \\ 4 \end{array} $	٥.			
	MAD	MAD data collection				
	Edge	Peak	Remote	ď	Parent	
Wavelength (Å)	0.97957	0.97946	0.93927	0.0	0.97946	
Resolution range (Å)	1.63	1.63	1.63			(1.69 - 1.63)
No. of unique reflections	35666	35621	36010	35621		(2656)
Completeness (%)	95.3	95.7	97.1	95.7		(70.8)
R merge (%)	6.6	7.2	6.4	7.2		(26.6)

_
\leq
_
_U
-
~
_
1
÷÷-
~
uthor
-
~
\geq
Aar
~
<u> </u>
~~
0
ISC
-
0
+

II 318F1 NIH-PA Author Manuscript

Kim et al.

\mathbf{x}
<u>S</u>
· 🗕
<u> </u>
co.
tatis
<u> </u>
6
3
Ś
. <u> </u>
Ξ
~
<u> </u>
Б
-
bh
Š
0
<u> </u>
-
6
ت.
ŝ
~
5

Crystallographic Statistics						
Phasing Resolution range 39–2.4 Å Phasing power FOM Density modification, FOM (1.8 Å)	Inf Fried 3.05 0.45	Iso 1.91 0.38	Peak Fried 2.71 0.43	Iso 1.33 0.30	h-Remote Fried 1.74 0.32	All 0.75 0.96
Refinement Resolution range (Å) No. of treflections of cutoff R-value Free R-value Rund length (1–2) (Å) Angle (°) Dihedral (°) Improper (°) Dihedral (°) Improper (°) Mana B-factor (Å2) Angle (°) Dihedral (°) Improper (°) Mana B-factor (Å2) Angle (°) Manator side chain Frotein atoms (2149) Protein atoms (2149) Protein atoms (2149) Protein atoms (2149) Remater (3) Mg (2) Mg (2) Kanater (272) Ramater plot statistics (%) Residues in most favored regions Residues in additional allowed regions					$\begin{array}{c} 39-1.63\\ 35198\\ 0.0\\ 0\\ 0.194\\ 0.214\ (4024)\\ 0.065\\ 1.20\\ 2.1.9\\ 0.84\\ 18.70\\ 1.20\\ 2.1.9\\ 0.84\\ 18.70\\ 112.52\\ 15.70\\ 12.52\\ 8.69\\ 30.76\\ 30.76\\ 0.0\\ 0.0\\ 0.0\end{array}$	