wwPDB X-ray Structure Validation Summary Report (i)

Jan 31, 2016 – 11:15 PM GMT

PDB ID	:	1WUF
Title	:	Crystal structure of protein GI:16801725, member of Enolase superfamily from Listeria innocua Clip11262
Authors	:	Fedorov, A.A.; Fedorov, E.V.; Yew, W.S.; Gerlt, J.A.; Almo, S.C.; Burley, S.K.; New York SGX Research Center for Structural Genomics (NYSGXRC)
Deposited on	:	2004-12-07
Resolution	:	2.90 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

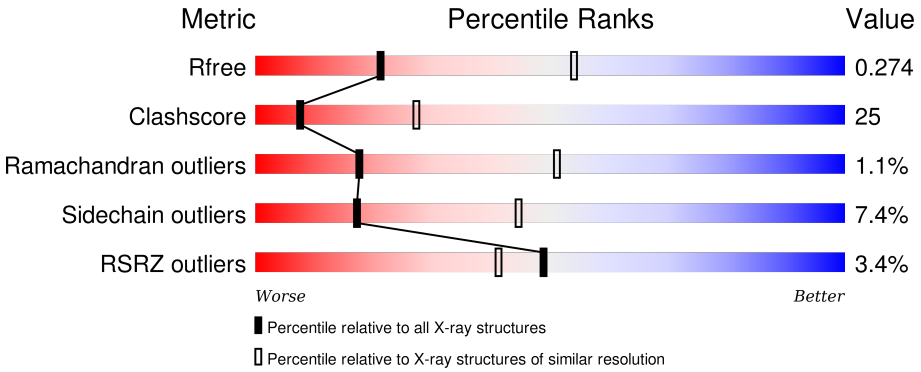
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.90 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	1451 (2.90-2.90)
Clashscore	102246	1668 (2.90-2.90)
Ramachandran outliers	100387	1630 (2.90-2.90)
Sidechain outliers	100360	1632 (2.90-2.90)
RSRZ outliers	91569	1456 (2.90-2.90)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	393	2% 52% 39% 6%
1	B	393	5% 53% 37% 6%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	MG	A	5001	-	-	-	X
2	MG	B	5002	-	-	-	X

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 5894 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called hypothetical protein lin2664.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	371	Total	C	N	O	S	0	0	0
			2946	1891	495	548	12			
1	B	371	Total	C	N	O	S	0	0	0
			2946	1891	495	548	12			

There are 38 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	982	GLY	-	EXPRESSION TAG	UNP Q927X3
A	983	HIS	-	EXPRESSION TAG	UNP Q927X3
A	984	HIS	-	EXPRESSION TAG	UNP Q927X3
A	985	HIS	-	EXPRESSION TAG	UNP Q927X3
A	986	HIS	-	EXPRESSION TAG	UNP Q927X3
A	987	HIS	-	EXPRESSION TAG	UNP Q927X3
A	988	HIS	-	EXPRESSION TAG	UNP Q927X3
A	989	HIS	-	EXPRESSION TAG	UNP Q927X3
A	990	HIS	-	EXPRESSION TAG	UNP Q927X3
A	991	HIS	-	EXPRESSION TAG	UNP Q927X3
A	992	HIS	-	EXPRESSION TAG	UNP Q927X3
A	993	GLY	-	EXPRESSION TAG	UNP Q927X3
A	994	LEU	-	EXPRESSION TAG	UNP Q927X3
A	995	VAL	-	EXPRESSION TAG	UNP Q927X3
A	996	PRO	-	EXPRESSION TAG	UNP Q927X3
A	997	ARG	-	EXPRESSION TAG	UNP Q927X3
A	998	GLY	-	EXPRESSION TAG	UNP Q927X3
A	999	SER	-	EXPRESSION TAG	UNP Q927X3
A	1000	HIS	-	EXPRESSION TAG	UNP Q927X3
B	1982	GLY	-	EXPRESSION TAG	UNP Q927X3
B	1983	HIS	-	EXPRESSION TAG	UNP Q927X3
B	1984	HIS	-	EXPRESSION TAG	UNP Q927X3
B	1985	HIS	-	EXPRESSION TAG	UNP Q927X3
B	1986	HIS	-	EXPRESSION TAG	UNP Q927X3
B	1987	HIS	-	EXPRESSION TAG	UNP Q927X3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	1988	HIS	-	EXPRESSION TAG	UNP Q927X3
B	1989	HIS	-	EXPRESSION TAG	UNP Q927X3
B	1990	HIS	-	EXPRESSION TAG	UNP Q927X3
B	1991	HIS	-	EXPRESSION TAG	UNP Q927X3
B	1992	HIS	-	EXPRESSION TAG	UNP Q927X3
B	1993	GLY	-	EXPRESSION TAG	UNP Q927X3
B	1994	LEU	-	EXPRESSION TAG	UNP Q927X3
B	1995	VAL	-	EXPRESSION TAG	UNP Q927X3
B	1996	PRO	-	EXPRESSION TAG	UNP Q927X3
B	1997	ARG	-	EXPRESSION TAG	UNP Q927X3
B	1998	GLY	-	EXPRESSION TAG	UNP Q927X3
B	1999	SER	-	EXPRESSION TAG	UNP Q927X3
B	2000	HIS	-	EXPRESSION TAG	UNP Q927X3

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

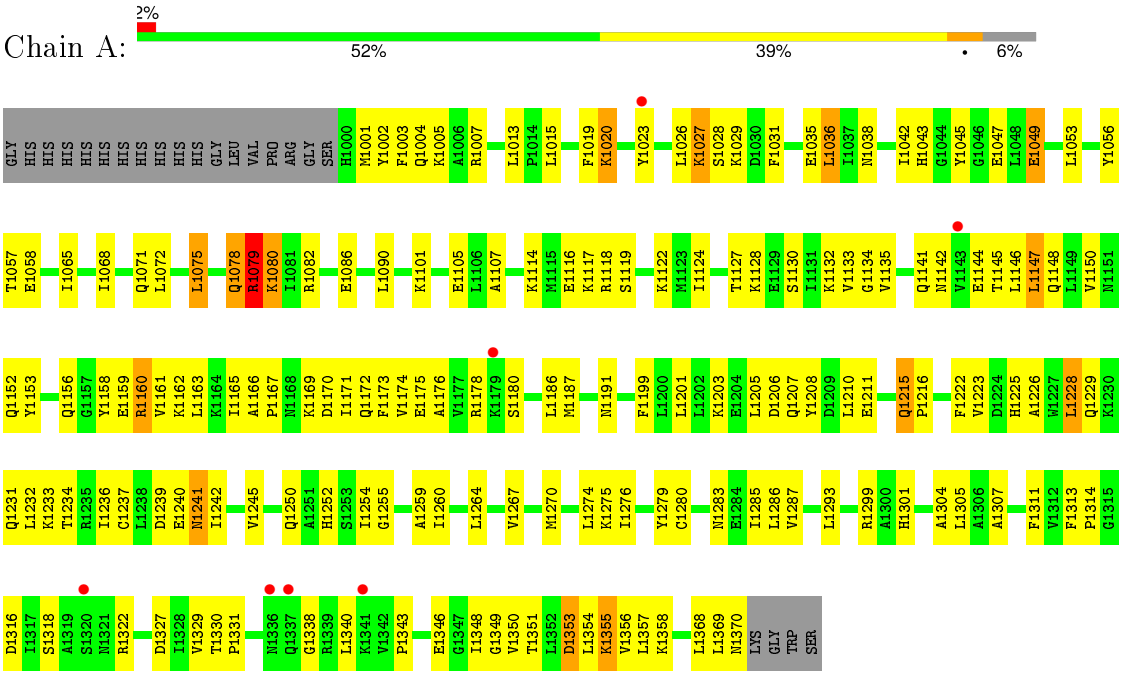
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	1	Total Mg 1 1	0	0
2	A	1	Total Mg 1 1	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

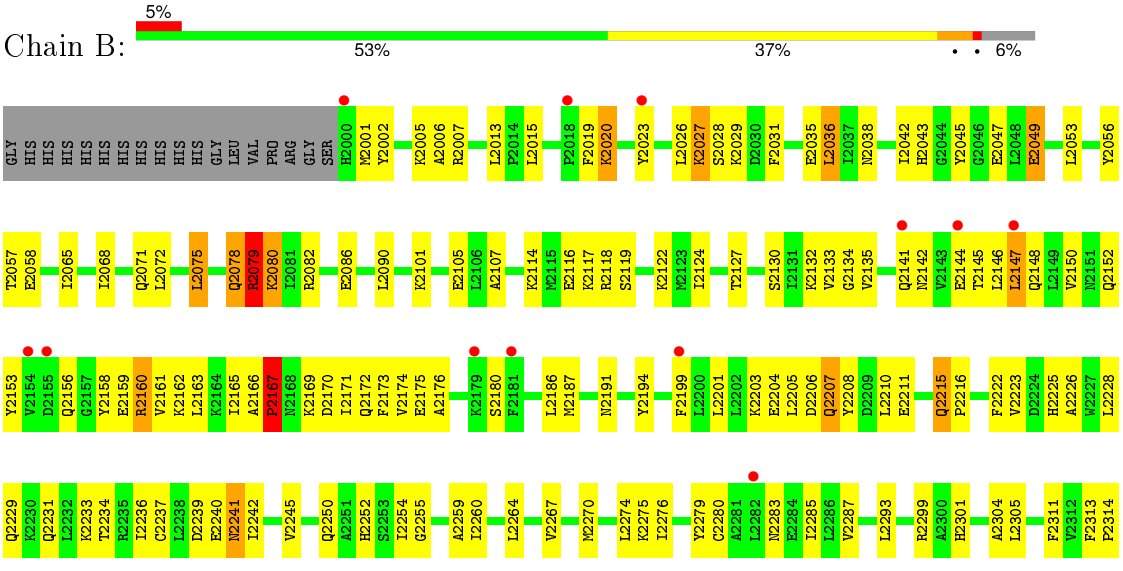
- Molecule 1: hypothetical protein lin2664

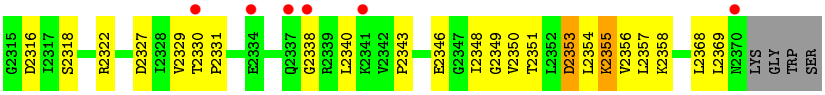
Chain A:



- Molecule 1: hypothetical protein lin2664

Chain B:




G2315
D2316
I2317
S2318
R2322
D2327
I2328
V2329
T2330
P2331
E2334
Q2337
G2338
R2339
L2340
K2341
V2342
P2343
E2346
G2347
I2348
G2349
V2350
T2351
L2352
D2353
L2354
K2355
V2356
L2357
K2358
L2368
L2369
M2370
LYS
GLY
TRP
SER

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	54.72Å 87.35Å 161.49Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	25.00 – 2.90 29.65 – 2.87	Depositor EDS
% Data completeness (in resolution range)	(Not available) (25.00-2.90) 95.0 (29.65-2.87)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.17 (at 2.90Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.236 , 0.274 0.235 , 0.274	Depositor DCC
R_{free} test set	811 reflections (4.73%)	DCC
Wilson B-factor (Å^2)	47.1	Xtriage
Anisotropy	0.118	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 45.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.32$	Xtriage
Outliers	0 of 17929 reflections	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	5894	wwPDB-VP
Average B, all atoms (Å^2)	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.44% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: MG

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$\|Z\|>5$	RMSZ	#$\|Z\|>5$
1	A	0.42	0/2998	0.68	1/4044 (0.0%)
1	B	0.43	0/2998	0.69	1/4044 (0.0%)
All	All	0.42	0/5996	0.69	2/8088 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	1079	ARG	N-CA-C	6.15	127.60	111.00
1	B	2079	ARG	N-CA-C	6.08	127.43	111.00

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2946	0	3006	158	0
1	B	2946	0	3006	158	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
All	All	5894	0	6012	302	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including

hydrogen atoms). The all-atom clashscore for this structure is 25.

The worst 5 of 302 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:2330:THR:OG1	1:B:2331:PRO:HD3	1.71	0.90
1:A:1241:ASN:H	1:A:1241:ASN:ND2	1.69	0.89
1:B:2241:ASN:H	1:B:2241:ASN:HD22	0.90	0.89
1:B:2241:ASN:HD22	1:B:2241:ASN:N	1.66	0.88
1:A:1241:ASN:HD22	1:A:1241:ASN:H	0.91	0.88

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	369/393 (94%)	337 (91%)	28 (8%)	4 (1%)	17	51
1	B	369/393 (94%)	336 (91%)	29 (8%)	4 (1%)	17	51
All	All	738/786 (94%)	673 (91%)	57 (8%)	8 (1%)	17	51

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1180	SER
1	B	2180	SER
1	A	1079	ARG
1	B	2079	ARG
1	A	1191	ASN

5.3.2 Protein sidechains

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	316/334 (95%)	293 (93%)	23 (7%)	17 45
1	B	316/334 (95%)	292 (92%)	24 (8%)	16 43
All	All	632/668 (95%)	585 (93%)	47 (7%)	17 44

5 of 47 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1357	LEU
1	B	2036	LEU
1	B	2353	ASP
1	B	2005	LYS
1	B	2049	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 34 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1283	ASN
1	B	2071	GLN
1	B	2283	ASN
1	B	2004	GLN
1	A	1156	GLN

5.3.3 RNA

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
1	A	371/393 (94%)	0.08	7 (1%) 70 66	14, 39, 70, 87	0
1	B	371/393 (94%)	0.29	18 (4%) 33 27	15, 40, 70, 87	0
All	All	742/786 (94%)	0.18	25 (3%) 49 41	14, 40, 70, 87	0

The worst 5 of 25 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	2155	ASP	3.6
1	A	1023	TYR	3.5
1	B	2337	GLN	3.4
1	B	2147	LEU	3.3
1	B	2000	HIS	3.3

6.2 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no carbohydrates in this entry.

6.4 Ligands

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
2	MG	A	5001	1/1	0.87	0.67	12.00	65,65,65,65	0
2	MG	B	5002	1/1	0.86	0.58	11.74	60,60,60,60	0

6.5 Other polymers (i)

There are no such residues in this entry.