Full wwPDB X-ray Structure Validation Report ⓘ

Nov 11, 2024 – 01:27 PM EST

PDB ID	:	1XRG
Title	:	Conserved hypothetical protein from Clostridium thermocellum Cth-2968
Authors	:	Zhao, M.; Chang, J.; Habel, J.; Kataeva, I.; Xu, H.; Chen, L.; Lee, D.; Nguyen, J.; Chang, S.-H.; Horanyi, P.; Florence, Q.; Zhou, W.; Tempel, W.; Lin, D.; Zhang, H.; Arendall III, W.B.; Ljundahl, L.; Liu, Z.-J.; Rose, J.; Richardson, J.S.; Richardson, D.C.; Wang, B.-C.; Southeast Collaboratory for Structural Genomics (SECSG)
Deposited on	:	2004-10-14
Resolution	:	2.20 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

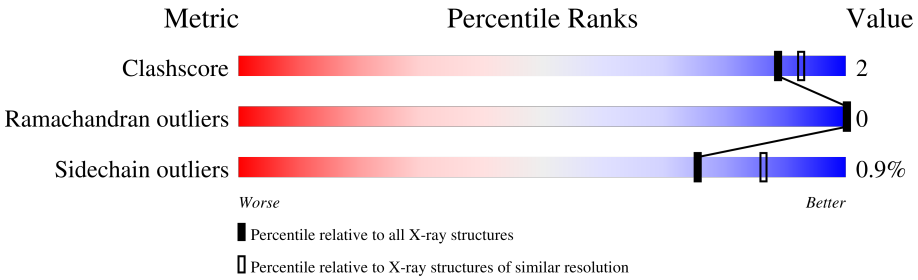
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	FAILED
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

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 2.20 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))
Clashscore	180529	6634 (2.20-2.20)
Ramachandran outliers	177936	6560 (2.20-2.20)
Sidechain outliers	177891	6561 (2.20-2.20)

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 of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. 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.

Note EDS failed to run properly.

Mol	Chain	Length	Quality of chain
1	A	156	76% · 20%
1	B	156	84% · 12%
1	C	156	83% · · 12%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 3105 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 Putative translation initiation inhibitor, yjgF family.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	125	Total	C	N	O	S	Se	0	1	0
			949	608	151	187	2	1			
1	B	138	Total	C	N	O	S	Se	0	1	0
			1043	669	164	206	2	2			
1	C	137	Total	C	N	O	S	Se	0	1	0
			1039	667	165	203	2	2			

There are 96 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-29	MSE	-	cloning artifact	GB 48859893
A	-28	GLY	-	cloning artifact	GB 48859893
A	-27	SER	-	cloning artifact	GB 48859893
A	-26	SER	-	cloning artifact	GB 48859893
A	-25	HIS	-	cloning artifact	GB 48859893
A	-24	HIS	-	cloning artifact	GB 48859893
A	-23	HIS	-	cloning artifact	GB 48859893
A	-22	HIS	-	cloning artifact	GB 48859893
A	-21	HIS	-	cloning artifact	GB 48859893
A	-20	HIS	-	cloning artifact	GB 48859893
A	-19	SER	-	cloning artifact	GB 48859893
A	-18	SER	-	cloning artifact	GB 48859893
A	-17	GLY	-	cloning artifact	GB 48859893
A	-16	LEU	-	cloning artifact	GB 48859893
A	-15	VAL	-	cloning artifact	GB 48859893
A	-14	PRO	-	cloning artifact	GB 48859893
A	-13	ARG	-	cloning artifact	GB 48859893
A	-12	GLY	-	cloning artifact	GB 48859893
A	-11	SER	-	cloning artifact	GB 48859893
A	-10	GLN	-	cloning artifact	GB 48859893
A	-9	SER	-	cloning artifact	GB 48859893
A	-8	THR	-	cloning artifact	GB 48859893
A	-7	SER	-	cloning artifact	GB 48859893

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	-6	LEU	-	cloning artifact	GB 48859893
A	-5	TYR	-	cloning artifact	GB 48859893
A	-4	LYS	-	cloning artifact	GB 48859893
A	-3	LYS	-	cloning artifact	GB 48859893
A	-2	ALA	-	cloning artifact	GB 48859893
A	-1	GLY	-	cloning artifact	GB 48859893
A	0	LEU	-	cloning artifact	GB 48859893
A	1	MSE	MET	modified residue	GB 48859893
A	83	MSE	MET	modified residue	GB 48859893
B	-29	MSE	-	cloning artifact	GB 48859893
B	-28	GLY	-	cloning artifact	GB 48859893
B	-27	SER	-	cloning artifact	GB 48859893
B	-26	SER	-	cloning artifact	GB 48859893
B	-25	HIS	-	cloning artifact	GB 48859893
B	-24	HIS	-	cloning artifact	GB 48859893
B	-23	HIS	-	cloning artifact	GB 48859893
B	-22	HIS	-	cloning artifact	GB 48859893
B	-21	HIS	-	cloning artifact	GB 48859893
B	-20	HIS	-	cloning artifact	GB 48859893
B	-19	SER	-	cloning artifact	GB 48859893
B	-18	SER	-	cloning artifact	GB 48859893
B	-17	GLY	-	cloning artifact	GB 48859893
B	-16	LEU	-	cloning artifact	GB 48859893
B	-15	VAL	-	cloning artifact	GB 48859893
B	-14	PRO	-	cloning artifact	GB 48859893
B	-13	ARG	-	cloning artifact	GB 48859893
B	-12	GLY	-	cloning artifact	GB 48859893
B	-11	SER	-	cloning artifact	GB 48859893
B	-10	GLN	-	cloning artifact	GB 48859893
B	-9	SER	-	cloning artifact	GB 48859893
B	-8	THR	-	cloning artifact	GB 48859893
B	-7	SER	-	cloning artifact	GB 48859893
B	-6	LEU	-	cloning artifact	GB 48859893
B	-5	TYR	-	cloning artifact	GB 48859893
B	-4	LYS	-	cloning artifact	GB 48859893
B	-3	LYS	-	cloning artifact	GB 48859893
B	-2	ALA	-	cloning artifact	GB 48859893
B	-1	GLY	-	cloning artifact	GB 48859893
B	0	LEU	-	cloning artifact	GB 48859893
B	1	MSE	MET	modified residue	GB 48859893
B	83	MSE	MET	modified residue	GB 48859893
C	-29	MSE	-	cloning artifact	GB 48859893

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	-28	GLY	-	cloning artifact	GB 48859893
C	-27	SER	-	cloning artifact	GB 48859893
C	-26	SER	-	cloning artifact	GB 48859893
C	-25	HIS	-	cloning artifact	GB 48859893
C	-24	HIS	-	cloning artifact	GB 48859893
C	-23	HIS	-	cloning artifact	GB 48859893
C	-22	HIS	-	cloning artifact	GB 48859893
C	-21	HIS	-	cloning artifact	GB 48859893
C	-20	HIS	-	cloning artifact	GB 48859893
C	-19	SER	-	cloning artifact	GB 48859893
C	-18	SER	-	cloning artifact	GB 48859893
C	-17	GLY	-	cloning artifact	GB 48859893
C	-16	LEU	-	cloning artifact	GB 48859893
C	-15	VAL	-	cloning artifact	GB 48859893
C	-14	PRO	-	cloning artifact	GB 48859893
C	-13	ARG	-	cloning artifact	GB 48859893
C	-12	GLY	-	cloning artifact	GB 48859893
C	-11	SER	-	cloning artifact	GB 48859893
C	-10	GLN	-	cloning artifact	GB 48859893
C	-9	SER	-	cloning artifact	GB 48859893
C	-8	THR	-	cloning artifact	GB 48859893
C	-7	SER	-	cloning artifact	GB 48859893
C	-6	LEU	-	cloning artifact	GB 48859893
C	-5	TYR	-	cloning artifact	GB 48859893
C	-4	LYS	-	cloning artifact	GB 48859893
C	-3	LYS	-	cloning artifact	GB 48859893
C	-2	ALA	-	cloning artifact	GB 48859893
C	-1	GLY	-	cloning artifact	GB 48859893
C	0	LEU	-	cloning artifact	GB 48859893
C	1	MSE	MET	modified residue	GB 48859893
C	83	MSE	MET	modified residue	GB 48859893

- Molecule 2 is UNKNOWN ATOM OR ION (three-letter code: UNX) (formula: X).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	13	Total X 13 13	0	0
2	B	2	Total X 2 2	0	0
2	C	2	Total X 2 2	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	17	Total O 17 17	0	0
3	B	20	Total O 20 20	0	0
3	C	20	Total O 20 20	0	0

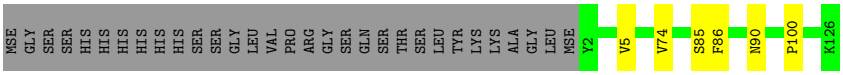
3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes 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.

Note EDS failed to run properly.

• Molecule 1: Putative translation initiation inhibitor, yjgF family

Chain A:

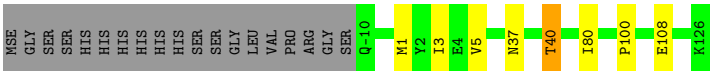


• Molecule 1: Putative translation initiation inhibitor, yjgF family

Chain B:

• Molecule 1: Putative translation initiation inhibitor, yjgF family

Chain C:



4 Data and refinement statistics (i)

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α, β, γ	80.44Å 80.44Å 137.20Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	19.90 – 2.20	Depositor
% Data completeness (in resolution range)	99.6 (19.90-2.20)	Depositor
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	17.22 (at 2.19Å)	Xtriage
Refinement program	REFMAC refmac_5.2.0005	Depositor
R, R_{free}	0.193 , 0.225	Depositor
Wilson B-factor (Å^2)	23.2	Xtriage
Anisotropy	0.019	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.033 for -h,-k,l	Xtriage
Total number of atoms	3105	wwPDB-VP
Average B, all atoms (Å^2)	23.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.82% 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, 0.333 respectively for untwinned datasets, and 0.375, 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: UNX

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	0.79	0/967	0.63	0/1309
1	B	0.77	0/1062	0.64	1/1436 (0.1%)
1	C	0.75	0/1058	0.66	0/1428
All	All	0.77	0/3087	0.64	1/4173 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	45	ASP	CB-CG-OD1	5.17	122.96	118.30

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	949	0	968	6	0
1	B	1043	0	1059	5	0
1	C	1039	0	1063	6	0
2	A	13	0	0	0	0
2	B	2	0	0	0	0
2	C	2	0	0	0	0
3	A	17	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	20	0	0	0	0
3	C	20	0	0	0	0
All	All	3105	0	3090	11	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 2.

All (11) 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:A:86:PHE:CE2	1:A:90:ASN:ND2	2.55	0.75
1:A:86:PHE:CZ	1:A:90:ASN:ND2	2.62	0.67
1:A:100:PRO:HB3	1:B:5:VAL:HB	1.88	0.54
1:C:37:ASN:HB3	1:C:40:THR:CG2	2.42	0.50
1:B:100:PRO:HB3	1:C:5:VAL:HB	1.96	0.47
1:A:5:VAL:CG2	1:C:100:PRO:HB3	2.45	0.47
1:B:74:VAL:HG12	1:C:3:ILE:HG12	1.96	0.47
1:A:74:VAL:HG12	1:B:3:ILE:HG12	1.98	0.45
1:B:86:PHE:CE1	1:B:106[B]:CYS:SG	3.07	0.43
1:C:80:ILE:O	1:C:108:GLU:HA	2.20	0.42
1:A:5:VAL:HG21	1:C:100:PRO:HB3	2.03	0.40

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	124/156 (80%)	120 (97%)	4 (3%)	0	100	100
1	B	137/156 (88%)	135 (98%)	2 (2%)	0	100	100
1	C	136/156 (87%)	133 (98%)	3 (2%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	397/468 (85%)	388 (98%)	9 (2%)	0	100 100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains (i)

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	106/129 (82%)	105 (99%)	1 (1%)	75 86
1	B	115/129 (89%)	115 (100%)	0	100 100
1	C	115/129 (89%)	113 (98%)	2 (2%)	56 71
All	All	336/387 (87%)	333 (99%)	3 (1%)	75 86

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	85	SER
1	C	1	MSE
1	C	40	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

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

5.5 Carbohydrates (i)

There are no oligosaccharides in this entry.

5.6 Ligand geometry (i)

Of 17 ligands modelled in this entry, 17 are unknown - 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

EDS failed to run properly - this section is therefore empty.

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

EDS failed to run properly - this section is therefore empty.

6.3 Carbohydrates

EDS failed to run properly - this section is therefore empty.

6.4 Ligands

EDS failed to run properly - this section is therefore empty.

6.5 Other polymers

EDS failed to run properly - this section is therefore empty.