wwPDB X-ray Structure Validation Summary Report (i)

Oct 12, 2024 – 09:03 PM EDT

PDB ID	:	2DLB
Title	:	X-ray Crystal Structure of Protein yopT from Bacillus subtilis. Northeast Structural Genomics Consortium Target SR412
Authors	:	Kuzin, A.P.; Chen, Y.; Seetharaman, J.; Ho, C.-K.; Cunningham, K.; Janjua, H.; Conover, K.; Ma, L.-C.; Xiao, R.; Acton, T.B.; Montelione, G.T.; Hunt, J.F.; Tong, L.; Northeast Structural Genomics Consortium (NESG)
Deposited on	:	2006-04-18
Resolution	:	1.20 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) 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 (i)) were used in the production of this report:

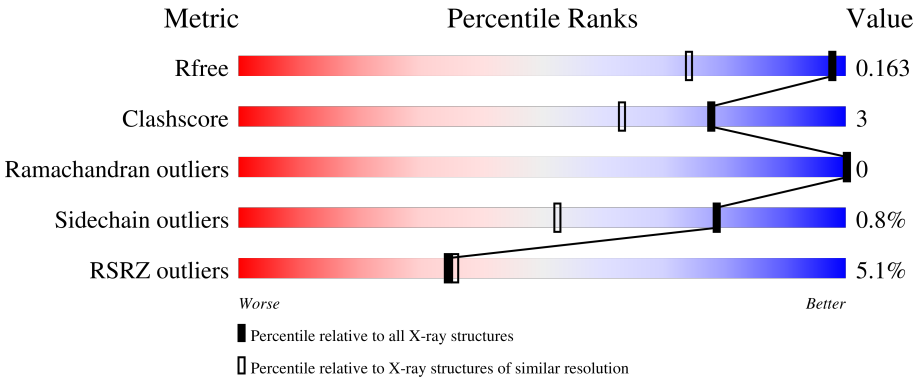
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (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 1.20 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))
R_{free}	164625	1079 (1.20-1.20)
Clashscore	180529	1183 (1.20-1.20)
Ramachandran outliers	177936	1146 (1.20-1.20)
Sidechain outliers	177891	1146 (1.20-1.20)
RSRZ outliers	164620	1078 (1.20-1.20)

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.

Mol	Chain	Length	Quality of chain
1	A	80	5% 79% 9% 12%
1	B	80	4% 76% 10% • 12%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 1294 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 yopT.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	70	Total	C	N	O	S	Se	0	1	0
			553	348	86	117	1	1			
1	B	70	Total	C	N	O	S	Se	2	2	0
			561	352	88	119	1	1			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	2001	MSE	MET	modified residue	UNP O34498
A	2060	MSE	MET	modified residue	UNP O34498
A	2073	LEU	-	expression tag	UNP O34498
A	2074	GLU	-	expression tag	UNP O34498
A	2075	HIS	-	expression tag	UNP O34498
A	2076	HIS	-	expression tag	UNP O34498
A	2077	HIS	-	expression tag	UNP O34498
A	2078	HIS	-	expression tag	UNP O34498
A	2079	HIS	-	expression tag	UNP O34498
A	2080	HIS	-	expression tag	UNP O34498
B	4001	MSE	MET	modified residue	UNP O34498
B	4060	MSE	MET	modified residue	UNP O34498
B	4073	LEU	-	expression tag	UNP O34498
B	4074	GLU	-	expression tag	UNP O34498
B	4075	HIS	-	expression tag	UNP O34498
B	4076	HIS	-	expression tag	UNP O34498
B	4077	HIS	-	expression tag	UNP O34498
B	4078	HIS	-	expression tag	UNP O34498
B	4079	HIS	-	expression tag	UNP O34498
B	4080	HIS	-	expression tag	UNP O34498

- Molecule 2 is water.

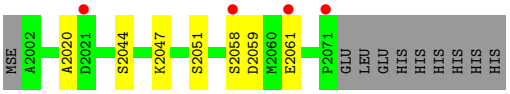
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	97	Total O 97 97	0	0
2	B	83	Total O 83 83	0	0

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.

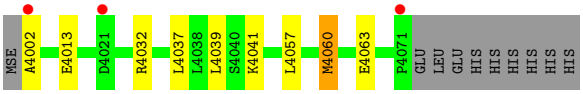
- Molecule 1: yopT

Chain A:



- Molecule 1: yopT

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	32.35Å 76.12Å 78.39Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 1.20 30.00 – 1.20	Depositor EDS
% Data completeness (in resolution range)	83.1 (30.00-1.20) 89.5 (30.00-1.20)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.95 (at 1.20Å)	Xtriage
Refinement program	SHELXL-97	Depositor
R, R_{free}	0.154 , 0.173 0.165 , 0.163	Depositor DCC
R_{free} test set	2797 reflections (4.90%)	wwPDB-VP
Wilson B-factor (Å^2)	9.3	Xtriage
Anisotropy	0.361	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 33.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.011 for -h,l,k	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	1294	wwPDB-VP
Average B, all atoms (Å^2)	17.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 15.07% 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

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	0.74	1/557 (0.2%)	1.14	0/752
1	B	0.74	0/565	1.27	5/763 (0.7%)
All	All	0.74	1/1122 (0.1%)	1.21	5/1515 (0.3%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	2044	SER	CB-OG	-6.41	1.33	1.42

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	4032	ARG	NE-CZ-NH1	-9.80	115.40	120.30
1	B	4032	ARG	CD-NE-CZ	8.39	135.35	123.60
1	B	4032	ARG	NH1-CZ-NH2	6.72	126.79	119.40
1	B	4057	LEU	CA-CB-CG	5.52	128.00	115.30
1	B	4063	GLU	OE1-CD-OE2	5.32	129.68	123.30

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	553	0	553	4	0
1	B	561	0	558	4	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	97	0	0	0	0
2	B	83	0	0	0	0
All	All	1294	0	1111	6	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 3.

The worst 5 of 6 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:2020:ALA:O	1:B:4002:ALA:HB1	1.92	0.69
1:A:2047:LYS:HE3	1:A:2051:SER:OG	2.06	0.55
1:A:2059:ASP:OD2	1:A:2061:GLU:OE1	2.29	0.51
1:A:2058[A]:SER:OG	1:B:4013:GLU:O	2.30	0.49
1:B:4037:LEU:O	1:B:4060:MSE:HE1	2.16	0.46

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	69/80 (86%)	69 (100%)	0	0	100	100
1	B	70/80 (88%)	70 (100%)	0	0	100	100
All	All	139/160 (87%)	139 (100%)	0	0	100	100

There are no Ramachandran outliers to report.

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	65/72 (90%)	65 (100%)	0	100	100
1	B	66/72 (92%)	65 (98%)	1 (2%)	60	25
All	All	131/144 (91%)	130 (99%)	1 (1%)	79	52

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	4060	MSE

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

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

There are no oligosaccharides in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

5.7 Other polymers

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	69/80 (86%)	0.30	4 (5%)	30	30	7, 13, 28, 41	1 (1%)
1	B	69/80 (86%)	0.34	3 (4%)	40	41	6, 13, 28, 41	2 (2%)
All	All	138/160 (86%)	0.32	7 (5%)	34	36	6, 13, 29, 41	3 (2%)

The worst 5 of 7 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	4002	ALA	4.4
1	A	2071	PRO	3.5
1	B	4071	PRO	3.2
1	A	2061	GLU	2.9
1	B	4021	ASP	2.8

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 monosaccharides in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.