Full wwPDB X-ray Structure Validation Report ⓘ

Nov 3, 2024 – 05:12 PM EST

PDB ID	:	1RKT
Title	:	Crystal structure of yfiR, a putative transcriptional regulator from Bacillus subtilis
Authors	:	Anderson, W.F.; Rajan, S.S.; Yang, X.; Midwest Center for Structural Genomics (MCSG)
Deposited on	:	2003-11-23
Resolution	:	1.95 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:

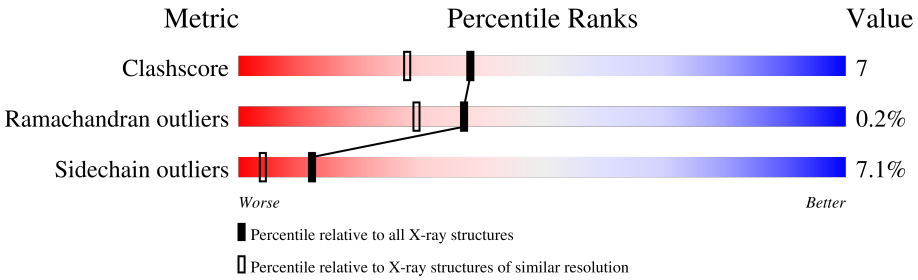
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	FAILED
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation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

1 Overall quality at a glance

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

The reported resolution of this entry is 1.95 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	3412 (1.96-1.96)
Ramachandran outliers	177936	3390 (1.96-1.96)
Sidechain outliers	177891	3390 (1.96-1.96)

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	205	79% 18% •
1	B	205	78% 19% ••

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	UNX	A	431	-	-	X	-

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 3548 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 protein yfiR.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	204	Total	C	N	O	Se	0	7	0
			1708	1083	289	333	3			
1	B	200	Total	C	N	O	Se	0	6	0
			1678	1069	280	326	3			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	MET	modified residue	UNP O31560
A	35	MSE	MET	modified residue	UNP O31560
A	59	MSE	MET	modified residue	UNP O31560
A	167	MSE	MET	modified residue	UNP O31560
B	1	MSE	MET	modified residue	UNP O31560
B	35	MSE	MET	modified residue	UNP O31560
B	59	MSE	MET	modified residue	UNP O31560
B	167	MSE	MET	modified residue	UNP O31560

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	16	Total	X	0	0
			16	16		
2	B	23	Total	X	0	0
			23	23		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	67	Total	O	0	0
			67	67		
3	B	56	Total	O	0	0
			56	56		

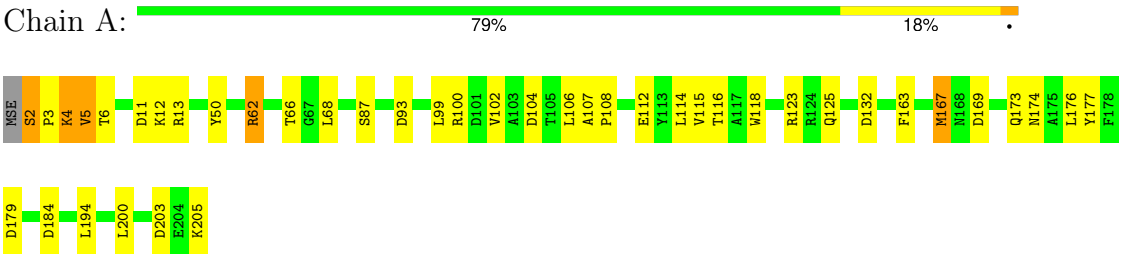
3 Residue-property plots (i)

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.

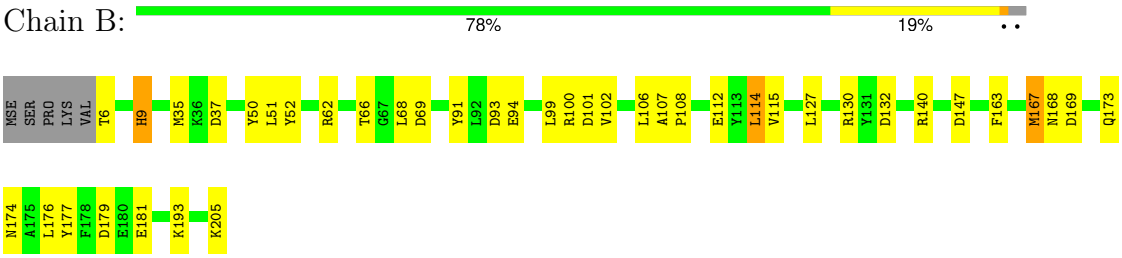
- Molecule 1: protein yfiR

Chain A:



- Molecule 1: protein yfiR

Chain B:



4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	47.98Å 48.77Å 57.23Å 71.96° 70.65° 80.20°	Depositor
Resolution (Å)	30.00 – 1.95	Depositor
% Data completeness (in resolution range)	97.0 (30.00-1.95)	Depositor
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.45 (at 1.91Å)	Xtriage
Refinement program	REFMAC 5.1.24	Depositor
R, R_{free}	0.208 , 0.254	Depositor
Wilson B-factor (Å^2)	35.3	Xtriage
Anisotropy	0.424	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.019 for -k,-h,-l	Xtriage
Total number of atoms	3548	wwPDB-VP
Average B, all atoms (Å^2)	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.44% 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, 0.333 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	0.50	3/1739 (0.2%)	0.71	8/2333 (0.3%)
1	B	0.40	0/1710	0.70	7/2293 (0.3%)
All	All	0.45	3/3449 (0.1%)	0.70	15/4626 (0.3%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	50	TYR	CG-CD2	8.09	1.49	1.39
1	A	50	TYR	CE1-CZ	7.01	1.47	1.38
1	A	167	MSE	SE-CE	-5.33	1.64	1.95

All (15) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	132	ASP	CB-CG-OD2	6.50	124.15	118.30
1	A	132	ASP	CB-CG-OD2	6.30	123.97	118.30
1	A	93	ASP	CB-CG-OD2	5.84	123.55	118.30
1	A	184	ASP	CB-CG-OD2	5.62	123.36	118.30
1	A	104	ASP	CB-CG-OD2	5.49	123.24	118.30
1	A	179	ASP	CB-CG-OD2	5.49	123.24	118.30
1	B	69	ASP	CB-CG-OD2	5.46	123.21	118.30
1	A	203	ASP	CB-CG-OD2	5.38	123.15	118.30
1	B	169	ASP	CB-CG-OD2	5.38	123.14	118.30
1	B	179	ASP	CB-CG-OD2	5.33	123.09	118.30
1	A	11	ASP	CB-CG-OD2	5.21	122.99	118.30
1	B	147	ASP	CB-CG-OD2	5.14	122.92	118.30
1	B	37	ASP	CB-CG-OD2	5.11	122.90	118.30
1	B	101	ASP	CB-CG-OD2	5.08	122.87	118.30
1	A	169	ASP	CB-CG-OD2	5.04	122.84	118.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

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	1708	0	1692	22	0
1	B	1678	0	1654	25	1
2	A	16	0	0	2	0
2	B	23	0	0	1	0
3	A	67	0	0	1	1
3	B	56	0	0	1	0
All	All	3548	0	3346	46	1

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 7.

All (46) 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:163:PHE:CD2	1:A:167:MSE:HE3	1.96	1.00
1:A:163:PHE:CD2	1:A:167:MSE:CE	2.51	0.94
1:A:163:PHE:CE2	1:A:167:MSE:CE	2.52	0.92
1:A:68:LEU:CD2	2:A:431:UNX:UNK	1.82	0.90
1:A:68:LEU:HD21	2:A:431:UNX:UNK	1.40	0.81
1:A:163:PHE:CE2	1:A:167:MSE:HE1	2.14	0.81
1:B:99:LEU:O	1:B:102[A]:VAL:HG23	1.80	0.80
1:A:2:SER:N	1:A:3:PRO:CD	2.48	0.77
1:B:62:ARG:O	1:B:66:THR:HG23	1.90	0.70
1:A:4:LYS:O	1:A:5:VAL:HG22	1.92	0.69
1:B:35:MSE:HG3	1:B:50:TYR:OH	1.92	0.69
2:B:401:UNX:UNK	2:B:409:UNX:UNK	1.37	0.68
1:A:163:PHE:HE2	1:A:167:MSE:HE1	1.58	0.66
1:A:112:GLU:O	1:A:115:VAL:HG12	1.98	0.64
1:A:62:ARG:O	1:A:66:THR:HG23	2.00	0.62
1:A:163:PHE:HD2	1:A:167:MSE:HE3	1.56	0.61
1:B:114:LEU:HD12	1:B:127[B]:LEU:HD21	1.85	0.58

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:93:ASP:OD1	1:B:193:LYS:NZ	2.24	0.57
1:B:91:TYR:OH	1:B:168:ASN:ND2	2.34	0.56
1:A:2:SER:N	1:A:3:PRO:HD2	2.23	0.54
1:A:2:SER:N	1:A:3:PRO:HD3	2.23	0.53
1:B:163:PHE:O	1:B:167:MSE:HG3	2.09	0.53
1:B:99:LEU:HA	1:B:102[B]:VAL:HG13	1.92	0.51
1:B:112:GLU:O	1:B:115:VAL:HG12	2.11	0.50
1:B:99:LEU:HD22	1:B:102[B]:VAL:HG11	1.94	0.49
1:A:118:TRP:NE1	1:B:112:GLU:OE1	2.39	0.48
1:A:12:LYS:HE3	3:A:475:HOH:O	2.13	0.48
1:A:115:VAL:HG11	1:B:115:VAL:HG11	1.96	0.47
1:B:35:MSE:HG3	1:B:50:TYR:CZ	2.50	0.47
1:A:163:PHE:CD2	1:A:167:MSE:HE1	2.36	0.46
1:B:68:LEU:HD23	1:B:130:ARG:HH21	1.80	0.46
1:B:108:PRO:HB3	1:B:176[B]:LEU:HD21	1.98	0.46
1:B:9:HIS:CD2	1:B:9:HIS:C	2.89	0.45
1:A:102:VAL:HG13	1:A:176:LEU:HD23	1.98	0.45
1:A:107:ALA:N	1:A:108:PRO:HD2	2.33	0.43
1:A:116[B]:THR:OG1	1:A:123:ARG:NH1	2.51	0.43
1:B:107:ALA:N	1:B:108:PRO:HD2	2.33	0.43
1:B:114:LEU:HD12	1:B:127[B]:LEU:HD11	1.99	0.43
1:B:99:LEU:CD2	1:B:102[B]:VAL:HG11	2.48	0.43
1:B:173:GLN:O	1:B:177:TYR:HD1	2.01	0.43
1:B:99:LEU:HA	1:B:102[A]:VAL:HG22	2.01	0.42
1:B:35:MSE:CG	1:B:50:TYR:OH	2.65	0.41
1:A:173:GLN:O	1:A:177:TYR:HD1	2.03	0.41
1:B:62:ARG:HD3	3:B:479:HOH:O	2.20	0.41
1:B:114:LEU:HA	1:B:127[B]:LEU:HD21	2.03	0.40
1:B:51:LEU:HD23	1:B:52:TYR:CZ	2.56	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:140:ARG:NH1	3:A:486:HOH:O[1_545]	1.94	0.26

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	209/205 (102%)	205 (98%)	3 (1%)	1 (0%)	25 16
1	B	204/205 (100%)	203 (100%)	1 (0%)	0	100 100
All	All	413/410 (101%)	408 (99%)	4 (1%)	1 (0%)	37 37

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	5	VAL

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	185/175 (106%)	168 (91%)	17 (9%)	7 2
1	B	180/175 (103%)	169 (94%)	11 (6%)	15 6
All	All	365/350 (104%)	337 (92%)	28 (8%)	12 3

All (28) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	2	SER
1	A	4	LYS
1	A	6	THR
1	A	13	ARG
1	A	62	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	87[A]	SER
1	A	87[B]	SER
1	A	99	LEU
1	A	100	ARG
1	A	106	LEU
1	A	114	LEU
1	A	125	GLN
1	A	174	ASN
1	A	194[A]	LEU
1	A	194[B]	LEU
1	A	200	LEU
1	A	205	LYS
1	B	6	THR
1	B	9	HIS
1	B	94[A]	GLU
1	B	94[B]	GLU
1	B	100	ARG
1	B	106	LEU
1	B	114	LEU
1	B	167	MSE
1	B	174	ASN
1	B	181	GLU
1	B	205	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (10) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	110	GLN
1	A	125	GLN
1	A	148	GLN
1	A	166	ASN
1	A	168	ASN
1	A	174	ASN
1	B	14	GLN
1	B	125	GLN
1	B	168	ASN
1	B	174	ASN

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

Of 39 ligands modelled in this entry, 39 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

There are no such residues in this entry.

5.8 Polymer linkage issues

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.