wwPDB X-ray Structure Validation Summary Report (i)

Oct 30, 2023 – 06:31 PM JST

PDB ID	:	4XLX
Title	:	Crystal structure of BjKS from Bradyrhizobium japonicum
Authors	:	Hu, Y.; Zheng, Y.; Ko, T.P.; Liu, W.; Guo, R.T.
Deposited on	:	2015-01-14
Resolution	:	2.00 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 (1)) were used in the production of this report:

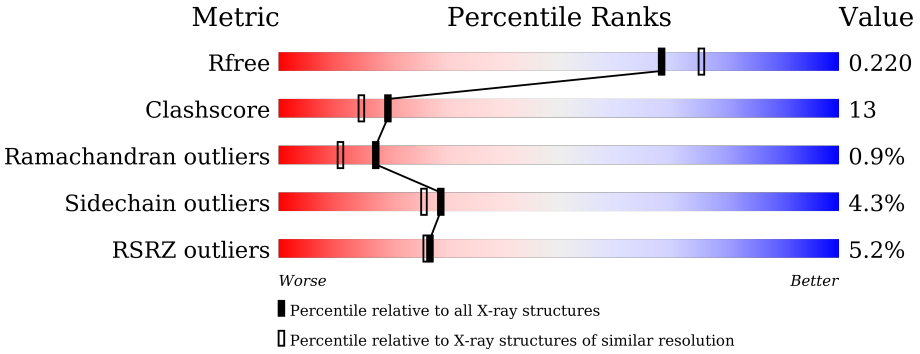
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.00 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}	130704	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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	300	5% 78% 14% • •
1	B	300	6% 70% 21% • 7%
1	C	300	4% 76% 14% • 6%
1	D	300	4% 66% 19% • 11%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 9417 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 Uncharacterized protein blr2150.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	288	2263	1421	413	417	12	0	0	0
1	B	279	2189	1379	398	400	12	0	0	0
1	C	281	2211	1394	402	403	12	0	0	0
1	D	267	2103	1334	379	379	11	0	0	0

- Molecule 2 is water.

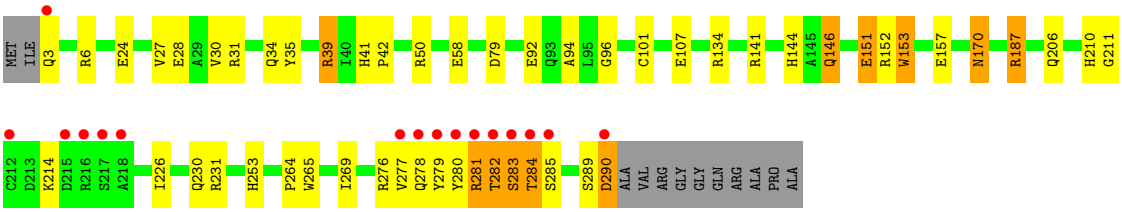
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	188	188	188	0	0
2	B	137	137	137	0	0
2	C	183	183	183	0	0
2	D	143	143	143	0	0

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.

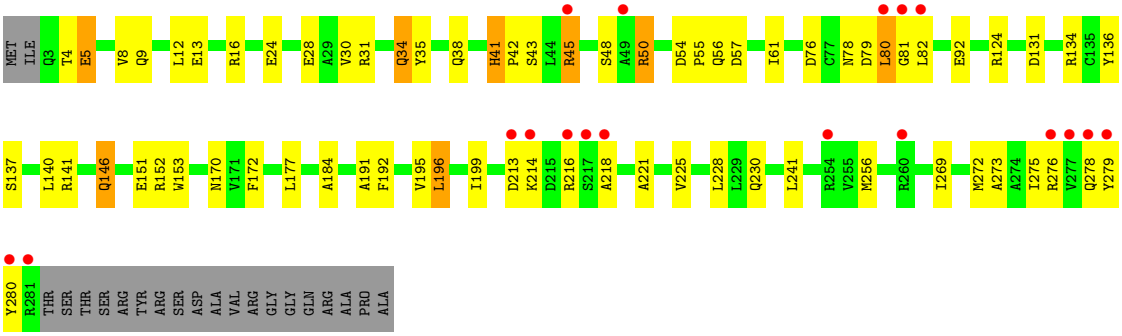
• Molecule 1: Uncharacterized protein blr2150

Chain A:



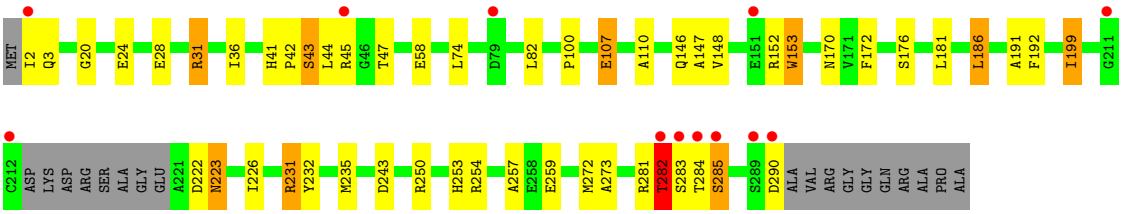
• Molecule 1: Uncharacterized protein blr2150

Chain B:



• Molecule 1: Uncharacterized protein blr2150

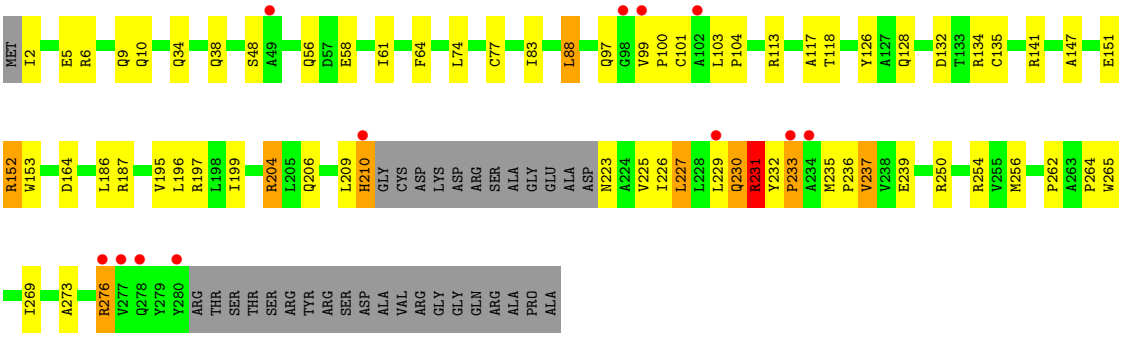
Chain C:



• Molecule 1: Uncharacterized protein blr2150

Chain D:

4%
66%
19%
•
11%


MET I2 E5 R6 Q9 Q10 Q34 Q38 S48 A49 Q56 D57 E58 I61 F64 L74 C77 I83 L88 Q97 G98 V99 P100 C101 A102 L103 P104 R113 A117 T118 Y126 A127 Q128 D132 T133 R134 C135 R141 A147 E151
R152 W153 D164 L186 R187 V195 L196 R197 L198 I199 R204 L205 Q206 L209 H210 GLY CYS ASP LYS ASP ARG SER ALA GLY GLU ALA ASP N223 A224 V225 I226 L227 L228 L229 Q230 R231 Y232 P233 A234 M235 P236 V237 V238 E239 R250 R254 V255 M256 P262 A263 P264 W265
I269 A273 R276 V277 Q278 Y279 Y280 ARG THR SER THR SER ARG TYR ARG SER ASP ALA VAL ARG GLY GLY GLN ARG ALA PRO ALA

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	65.03Å 130.07Å 66.39Å 90.00° 95.62° 90.00°	Depositor
Resolution (Å)	25.00 – 2.00 24.81 – 2.00	Depositor EDS
% Data completeness (in resolution range)	98.0 (25.00-2.00) 95.4 (24.81-2.00)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	8.02 (at 1.99Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, R_{free}	0.176 , 0.220 0.176 , 0.220	Depositor DCC
R_{free} test set	3582 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å^2)	20.7	Xtriage
Anisotropy	0.424	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 57.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.028 for l,-k,h	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	9417	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 4.53% 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 (i)

5.1 Standard geometry (i)

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	0.97	0/2310	0.86	2/3138 (0.1%)
1	B	0.89	0/2235	0.83	0/3037
1	C	0.95	0/2257	0.89	2/3067 (0.1%)
1	D	0.92	2/2148 (0.1%)	0.85	2/2921 (0.1%)
All	All	0.93	2/8950 (0.0%)	0.86	6/12163 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	256	MET	CG-SD	-6.49	1.64	1.81
1	D	64	PHE	CE2-CZ	5.28	1.47	1.37

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	282	THR	N-CA-C	7.39	130.95	111.00
1	A	39	ARG	NE-CZ-NH1	-6.31	117.14	120.30
1	D	132	ASP	CB-CG-OD1	5.65	123.39	118.30
1	A	187	ARG	NE-CZ-NH2	-5.26	117.67	120.30
1	D	204	ARG	NE-CZ-NH2	-5.26	117.67	120.30

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	2263	0	2220	46	0
1	B	2189	0	2152	67	0
1	C	2211	0	2177	51	0
1	D	2103	0	2079	66	0
2	A	188	0	0	6	0
2	B	137	0	0	8	0
2	C	183	0	0	3	0
2	D	143	0	0	3	0
All	All	9417	0	8628	223	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 13.

The worst 5 of 223 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:199:ILE:HD11	1:B:273:ALA:HA	1.29	1.12
1:D:135:CYS:HB2	2:D:513:HOH:O	1.57	1.05
1:C:235:MET:HE1	1:D:235:MET:HG3	1.41	1.01
1:C:281:ARG:HG3	1:C:282:THR:HG23	1.45	0.95
1:B:199:ILE:HD11	1:B:273:ALA:CA	1.99	0.92

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	286/300 (95%)	281 (98%)	2 (1%)	3 (1%)	15 9
1	B	277/300 (92%)	272 (98%)	4 (1%)	1 (0%)	34 30
1	C	277/300 (92%)	269 (97%)	6 (2%)	2 (1%)	22 16
1	D	263/300 (88%)	252 (96%)	7 (3%)	4 (2%)	10 4

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	1103/1200 (92%)	1074 (97%)	19 (2%)	10 (1%)	17 11

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	282	THR
1	C	282	THR
1	D	230	GLN
1	D	233	PRO
1	C	285	SER

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	230/237 (97%)	221 (96%)	9 (4%)	32 30
1	B	221/237 (93%)	211 (96%)	10 (4%)	27 24
1	C	225/237 (95%)	216 (96%)	9 (4%)	31 29
1	D	213/237 (90%)	203 (95%)	10 (5%)	26 22
All	All	889/948 (94%)	851 (96%)	38 (4%)	29 26

5 of 38 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	48	SER
1	D	231	ARG
1	D	88	LEU
1	D	196	LEU
1	D	276	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 19 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	210	HIS
1	D	34	GLN
1	D	170	ASN
1	D	10	GLN
1	B	93	GLN

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

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 (i)

6.1 Protein, DNA and RNA chains (i)

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, 95th 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	288/300 (96%)	-0.10	16 (5%)	24	23	10, 20, 62, 83	0
1	B	279/300 (93%)	0.05	18 (6%)	18	18	12, 26, 54, 79	0
1	C	281/300 (93%)	-0.11	12 (4%)	35	34	11, 21, 48, 74	0
1	D	267/300 (89%)	-0.11	12 (4%)	33	32	11, 24, 51, 62	0
All	All	1115/1200 (92%)	-0.07	58 (5%)	27	26	10, 23, 54, 83	0

The worst 5 of 58 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	284	THR	10.3
1	A	282	THR	8.8
1	A	283	SER	7.2
1	C	282	THR	6.8
1	C	283	SER	6.6

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

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

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.