wwPDB X-ray Structure Validation Summary Report (i)

Jan 30, 2021 – 02:09 PM EST

PDB ID	:	3DXI
Title	:	Crystal structure of the N-terminal domain of a putative aldolase (BVU_2661) from Bacteroides vulgatus
Authors	:	Eswaramoorthy, S.; Pabalan, A.A.; Burley, S.K.; Swaminathan, S.; New York SGX Research Center for Structural Genomics (NYSGXRC)
Deposited on	:	2008-07-24
Resolution	:	2.04 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 following versions of software and data (see references (i)) were used in the production of this report:

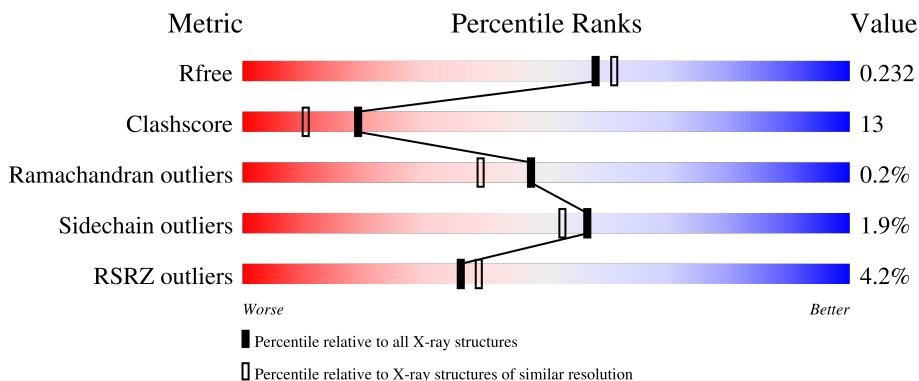|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.16
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.16

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 2.04 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	Percentile Ranks	Value
Rfree		0.232
Clashscore		13
Ramachandran outliers		0.2%
Sidechain outliers		1.9%
RSRZ outliers		4.2%

Worse .... *Better*

▌ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1692 (2.04-2.04)
Clashscore	141614	1773 (2.04-2.04)
Ramachandran outliers	138981	1752 (2.04-2.04)
Sidechain outliers	138945	1752 (2.04-2.04)
RSRZ outliers	127900	1672 (2.04-2.04)

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	320	4% / 69% / 24% / 7%
1	B	320	4% / 70% / 22% / · / 7%

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 5201 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 aldolase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	299	Total	C	N	O	S	Se	0	0	0
			2372	1536	391	430	2	13			
1	B	299	Total	C	N	O	S	Se	0	0	0
			2372	1536	391	430	2	13			

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	MSE	-	expression tag	UNP A6L3P9
A	0	SER	-	expression tag	UNP A6L3P9
A	1	LEU	-	expression tag	UNP A6L3P9
A	311	GLU	-	expression tag	UNP A6L3P9
A	312	GLY	-	expression tag	UNP A6L3P9
A	313	HIS	-	expression tag	UNP A6L3P9
A	314	HIS	-	expression tag	UNP A6L3P9
A	315	HIS	-	expression tag	UNP A6L3P9
A	316	HIS	-	expression tag	UNP A6L3P9
A	317	HIS	-	expression tag	UNP A6L3P9
A	318	HIS	-	expression tag	UNP A6L3P9
B	-1	MSE	-	expression tag	UNP A6L3P9
B	0	SER	-	expression tag	UNP A6L3P9
B	1	LEU	-	expression tag	UNP A6L3P9
B	311	GLU	-	expression tag	UNP A6L3P9
B	312	GLY	-	expression tag	UNP A6L3P9
B	313	HIS	-	expression tag	UNP A6L3P9
B	314	HIS	-	expression tag	UNP A6L3P9
B	315	HIS	-	expression tag	UNP A6L3P9
B	316	HIS	-	expression tag	UNP A6L3P9
B	317	HIS	-	expression tag	UNP A6L3P9
B	318	HIS	-	expression tag	UNP A6L3P9

- Molecule 2 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	249	Total 249	O 249	0	0
2	B	208	Total 208	O 208	0	0

3 Residue-property plots ⓘ

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: Putative aldolase

Chain A:

• Molecule 1: Putative aldolase

Chain B:
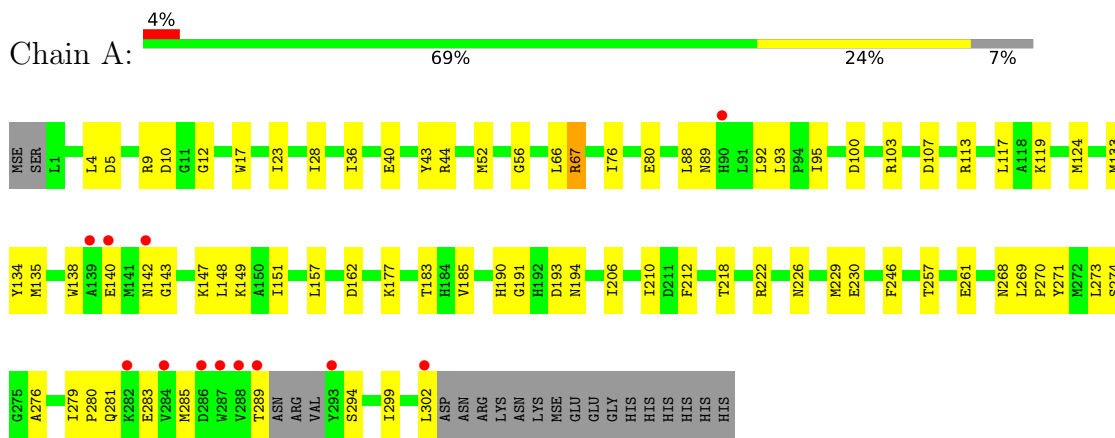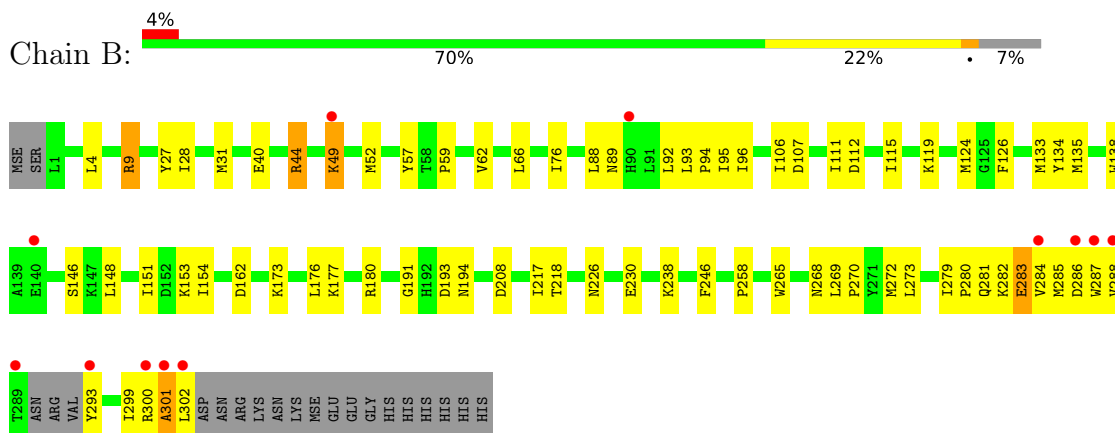
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	57.26Å 65.64Å 177.84Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	36.81 – 2.04 36.81 – 2.04	Depositor EDS
% Data completeness (in resolution range)	92.7 (36.81-2.04) 92.9 (36.81-2.04)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.93 (at 2.05Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.197 , 0.232 0.197 , 0.232	Depositor DCC
R_{free} test set	1641 reflections (3.85%)	wwPDB-VP
Wilson B-factor (Å^2)	12.9	Xtriage
Anisotropy	0.839	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 47.5	EDS
L-test for twinning[2]	$< |L| > = 0.48, < L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	5201	wwPDB-VP
Average B, all atoms (Å^2)	15.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.31% 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	0.33	0/2408	0.63	2/3237 (0.1%)
1	B	0.32	0/2408	0.63	0/3237
All	All	0.33	0/4816	0.63	2/6474 (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	294	SER	N-CA-C	-5.37	96.52	111.00
1	A	44	ARG	N-CA-C	-5.12	97.16	111.00

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	2372	0	2422	61	0
1	B	2372	0	2422	73	0
2	A	249	0	0	10	0
2	B	208	0	0	9	0
All	All	5201	0	4844	125	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 125 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:119:LYS:HD2	2:A:539:HOH:O	1.49	1.09
1:A:273:LEU:HD22	1:A:299:ILE:HD11	1.52	0.91
1:A:149:LYS:HD2	2:A:525:HOH:O	1.72	0.88
1:B:153:LYS:HE2	2:B:500:HOH:O	1.78	0.83
1:A:230:GLU:HG3	1:B:272:MSE:HE1	1.60	0.83

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	295/320 (92%)	290 (98%)	5 (2%)	0	100	100
1	B	295/320 (92%)	291 (99%)	3 (1%)	1 (0%)	41	31
All	All	590/640 (92%)	581 (98%)	8 (1%)	1 (0%)	47	39

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	301	ALA

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	261/266 (98%)	257 (98%)	4 (2%)	65	63
1	B	261/266 (98%)	255 (98%)	6 (2%)	50	44
All	All	522/532 (98%)	512 (98%)	10 (2%)	57	53

5 of 10 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	9	ARG
1	B	44	ARG
1	B	146	SER
1	A	193	ASP
1	B	49	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	226	ASN
1	B	226	ASN
1	B	194	ASN
1	A	194	ASN
1	B	89	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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

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 ⓘ

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, 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(Å²)	Q<0.9
1	A	286/320 (89%)	0.06	12 (4%)	36	39	6, 13, 26, 37	0
1	B	286/320 (89%)	0.09	12 (4%)	36	39	6, 13, 29, 38	0
All	All	572/640 (89%)	0.07	24 (4%)	36	39	6, 13, 29, 38	0

The worst 5 of 24 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	288	VAL	5.8
1	A	289	THR	5.0
1	A	293	TYR	4.7
1	B	302	LEU	4.7
1	B	288	VAL	4.5

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.