Full wwPDB X-ray Structure Validation Report (i)

Nov 7, 2023 – 03:23 am GMT

PDB ID	:	7BNG
Title	:	Complex structure of SH3b domain with L-canavanine
Authors	:	Malecki, P.H.; Sabala, I.
Deposited on	:	2021-01-22
Resolution	:	1.47 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 (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:

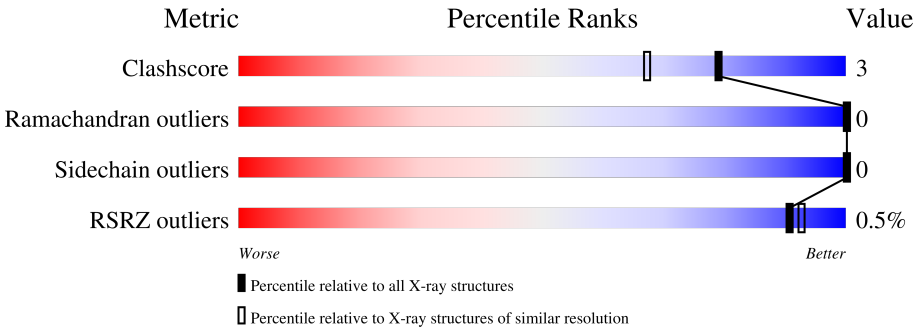
MolProbity	:	4.02b-467
Mogul	:	FAILED
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	FAILED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 1.47 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	141614	4955 (1.50-1.46)
Ramachandran outliers	138981	4846 (1.50-1.46)
Sidechain outliers	138945	4844 (1.50-1.46)
RSRZ outliers	127900	4614 (1.50-1.46)

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 1858 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.

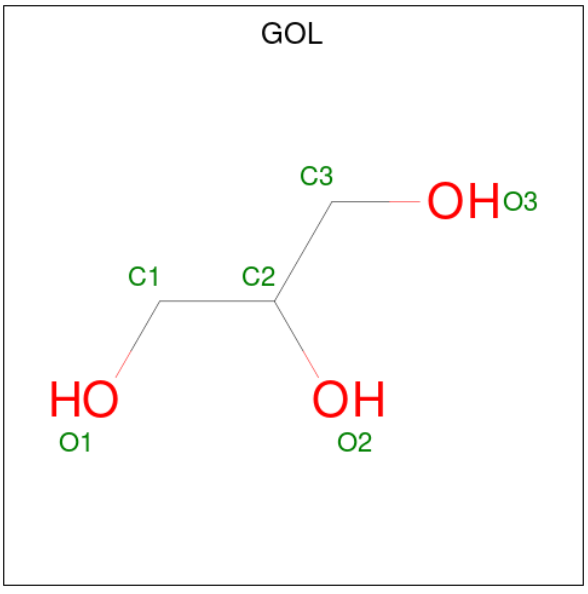
- Molecule 1 is a protein called Lysostaphin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	96	Total	C	N	O	S	0	3	0
			779	500	134	144	1			
1	B	96	Total	C	N	O	S	0	3	0
			782	500	134	147	1			

There are 4 discrepancies between the modelled and reference sequences:

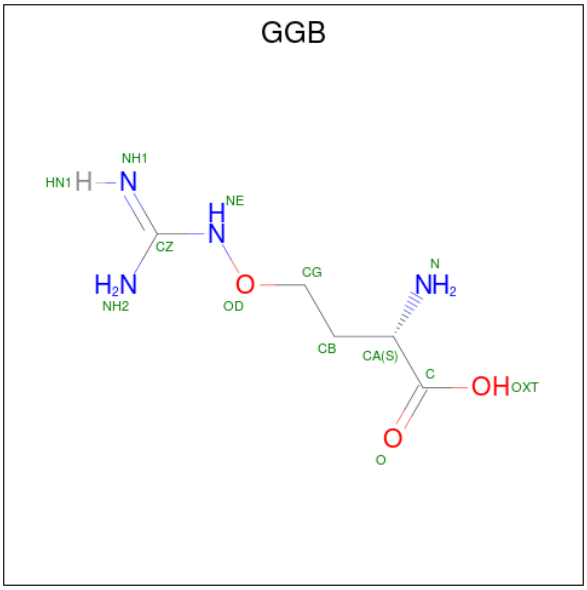
Chain	Residue	Modelled	Actual	Comment	Reference
A	398	SER	THR	conflict	UNP A0A133QQ41
A	400	ALA	SER	conflict	UNP A0A133QQ41
B	398	SER	THR	conflict	UNP A0A133QQ41
B	400	ALA	SER	conflict	UNP A0A133QQ41

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



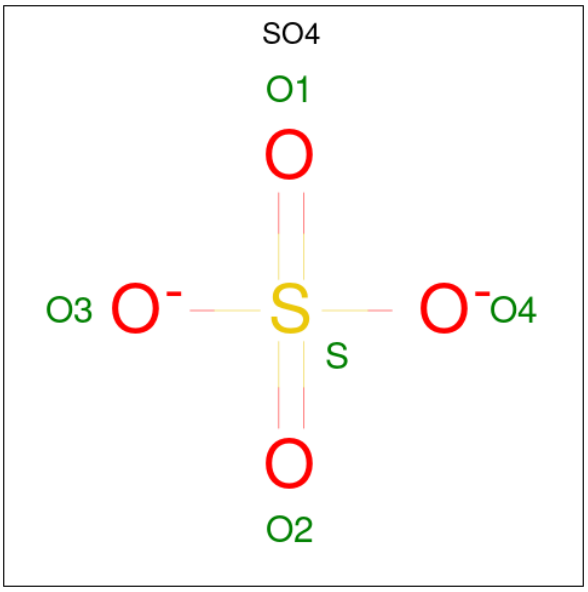
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C O 6 3 3	0	0
2	A	1	Total C O 6 3 3	0	0

- Molecule 3 is L-CANAVANINE (three-letter code: GGB) (formula: $C_5H_{12}N_4O_3$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C N O 12 5 4 3	0	0

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total O S 5 4 1	0	1
4	A	1	Total O S 5 4 1	0	0
4	A	1	Total O S 5 4 1	0	0
4	A	1	Total O S 5 4 1	0	1
4	B	1	Total O S 5 4 1	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	125	Total O 135 135	0	12
5	B	100	Total O 113 113	0	16

SEQUENCE-PLOTS INFOmissingINFO

3 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	36.90Å 55.23Å 48.99Å 90.00° 93.41° 90.00°	Depositor
Resolution (Å)	36.83 – 1.47 36.83 – 1.47	Depositor EDS
% Data completeness (in resolution range)	98.7 (36.83-1.47) 90.3 (36.83-1.47)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.18 (at 1.47Å)	Xtriage
Refinement program	PHENIX 1.18.2	Depositor
R, R_{free}	0.159 , 0.200 (Not available) , (Not available)	Depositor DCC
R_{free} test set	1064 reflections (3.20%)	wwPDB-VP
Wilson B-factor (Å^2)	12.9	Xtriage
Anisotropy	0.244	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.40 , 78.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	1858	wwPDB-VP
Average B, all atoms (Å^2)	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.54% 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.

4 Model quality

4.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GGB, SO4, GOL

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	0/810	0.69	0/1099
1	B	0.57	0/813	0.76	0/1104
All	All	0.54	0/1623	0.73	0/2203

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

4.2 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	779	0	773	6	0
1	B	782	0	769	3	0
2	A	12	0	16	0	0
3	A	12	0	11	0	0
4	A	20	0	0	0	0
4	B	5	0	0	0	0
5	A	135	0	0	1	0
5	B	113	0	0	0	0
All	All	1858	0	1569	9	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.

All (9) 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:414:GLU:HB3	1:A:452[A]:VAL:HG23	1.72	0.72
1:B:452[B]:VAL:HG12	1:B:461:VAL:HG12	1.76	0.65
1:A:452[B]:VAL:HG12	1:A:461:VAL:HG12	1.82	0.61
1:B:414:GLU:HB3	1:B:452[A]:VAL:HG23	1.84	0.59
1:A:488:LEU:HD21	1:A:492:ILE:HG13	1.98	0.44
1:A:412:LYS:HG2	1:A:452[B]:VAL:HG23	1.98	0.44
1:A:449:TYR:CG	1:A:461:VAL:HB	2.55	0.41
1:B:447:ILE:C	1:B:447:ILE:HD12	2.40	0.41
1:A:448[A]:LYS:HE3	5:A:635:HOH:O	2.21	0.41

There are no symmetry-related clashes.

4.3 Torsion angles (i)

4.3.1 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	97/96 (101%)	95 (98%)	2 (2%)	0	100	100
1	B	98/96 (102%)	96 (98%)	2 (2%)	0	100	100
All	All	195/192 (102%)	191 (98%)	4 (2%)	0	100	100

There are no Ramachandran outliers to report.

4.3.2 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	85/82 (104%)	85 (100%)	0	100	100
1	B	86/82 (105%)	86 (100%)	0	100	100
All	All	171/164 (104%)	171 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

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

4.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

4.5 Carbohydrates (i)

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

4.6 Ligand geometry (i)

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

4.7 Other polymers (i)

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

4.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

5 Fit of model and data

5.1 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	96/96 (100%)	-0.37	0 100 100	11, 16, 27, 31	0
1	B	96/96 (100%)	-0.32	1 (1%) 82 84	11, 17, 32, 49	0
All	All	192/192 (100%)	-0.34	1 (0%) 91 93	11, 16, 30, 49	0

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	466	TYR	5.4

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

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

5.3 Carbohydrates

There are no monosaccharides in this entry.

5.4 Ligands

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	GGB	A	502	12/12	0.90	0.13	23,25,30,31	0
2	GOL	A	503	6/6	0.91	0.09	21,23,24,28	0
4	SO4	A	507[A]	5/5	0.94	0.11	41,42,42,42	5
2	GOL	A	501	6/6	0.95	0.10	17,17,21,22	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	SO4	A	504[A]	5/5	0.97	0.13	24,24,26,27	5
4	SO4	A	506	5/5	0.98	0.08	27,28,28,29	0
4	SO4	A	505	5/5	0.98	0.09	25,26,26,27	0
4	SO4	B	501	5/5	0.98	0.10	28,29,30,31	0

5.5 Other polymers (i)

There are no such residues in this entry.