



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 13, 2017 – 10:56 am GMT

PDB ID : 2OL7
Title : The crystal structure of OspA mutant
Authors : Makabe, K.; Terechko, V.; Koide, S.
Deposited on : 2007-01-18
Resolution : 1.35 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

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	recalc28949

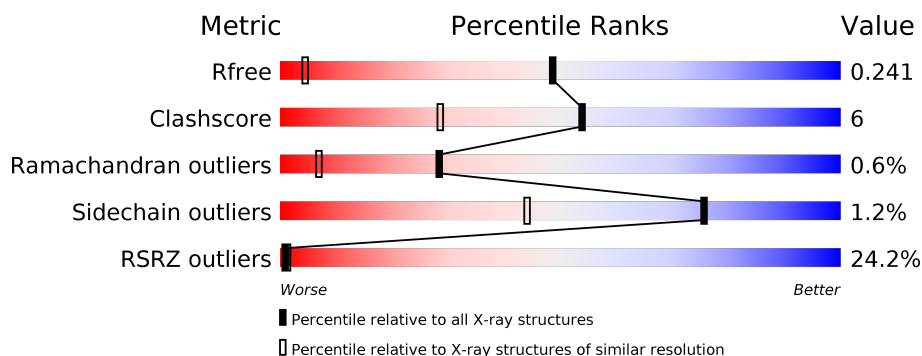
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.35 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}	100719	1024 (1.38-1.34)
Clashscore	112137	1063 (1.38-1.34)
Ramachandran outliers	110173	1048 (1.38-1.34)
Sidechain outliers	110143	1048 (1.38-1.34)
RSRZ outliers	101464	1025 (1.38-1.34)

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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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 $\leq 5\%$. 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	251	<div> <div>8%</div> <div>93%</div> <div>6%</div> </div>
1	B	251	<div> <div>40%</div> <div>88%</div> <div>7%</div> <div>••</div> </div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 4335 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 Outer surface protein A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	250	Total	C	N	O	S	0	15	0
			1921	1197	309	412	3			
1	B	246	Total	C	N	O	S	0	7	0
			1846	1145	300	400	1			

There are 46 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	23	GLY	-	EXPRESSION TAG	UNP Q45040
A	24	SER	-	EXPRESSION TAG	UNP Q45040
A	25	HIS	-	EXPRESSION TAG	UNP Q45040
A	26	MET	-	EXPRESSION TAG	UNP Q45040
A	37	SER	GLU	ENGINEERED	UNP Q45040
A	45	SER	GLU	ENGINEERED	UNP Q45040
A	46	SER	LYS	ENGINEERED	UNP Q45040
A	48	ALA	LYS	ENGINEERED	UNP Q45040
A	60	ALA	LYS	ENGINEERED	UNP Q45040
A	64	SER	LYS	ENGINEERED	UNP Q45040
A	83	ALA	LYS	ENGINEERED	UNP Q45040
A	104	SER	GLU	ENGINEERED	UNP Q45040
A	107	SER	LYS	ENGINEERED	UNP Q45040
A	126	ILE	PHE	ENGINEERED	UNP Q45040
A	127	ILE	ASN	ENGINEERED	UNP Q45040
A	128	ILE	GLU	ENGINEERED	UNP Q45040
A	129	ASP	LYS	ENGINEERED	UNP Q45040
A	131	ILE	GLU	ENGINEERED	UNP Q45040
A	132	ILE	VAL	ENGINEERED	UNP Q45040
A	133	ILE	SER	ENGINEERED	UNP Q45040
A	239	SER	LYS	ENGINEERED	UNP Q45040
A	240	SER	GLU	ENGINEERED	UNP Q45040
A	254	SER	LYS	ENGINEERED	UNP Q45040
B	23	GLY	-	EXPRESSION TAG	UNP Q45040
B	24	SER	-	EXPRESSION TAG	UNP Q45040

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	25	HIS	-	EXPRESSION TAG	UNP Q45040
B	26	MET	-	EXPRESSION TAG	UNP Q45040
B	37	SER	GLU	ENGINEERED	UNP Q45040
B	45	SER	GLU	ENGINEERED	UNP Q45040
B	46	SER	LYS	ENGINEERED	UNP Q45040
B	48	ALA	LYS	ENGINEERED	UNP Q45040
B	60	ALA	LYS	ENGINEERED	UNP Q45040
B	64	SER	LYS	ENGINEERED	UNP Q45040
B	83	ALA	LYS	ENGINEERED	UNP Q45040
B	104	SER	GLU	ENGINEERED	UNP Q45040
B	107	SER	LYS	ENGINEERED	UNP Q45040
B	126	ILE	PHE	ENGINEERED	UNP Q45040
B	127	ILE	ASN	ENGINEERED	UNP Q45040
B	128	ILE	GLU	ENGINEERED	UNP Q45040
B	129	ASP	LYS	ENGINEERED	UNP Q45040
B	131	ILE	GLU	ENGINEERED	UNP Q45040
B	132	ILE	VAL	ENGINEERED	UNP Q45040
B	133	ILE	SER	ENGINEERED	UNP Q45040
B	239	SER	LYS	ENGINEERED	UNP Q45040
B	240	SER	GLU	ENGINEERED	UNP Q45040
B	254	SER	LYS	ENGINEERED	UNP Q45040

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	350	Total O 350 350	0	0
2	B	218	Total O 218 218	0	0

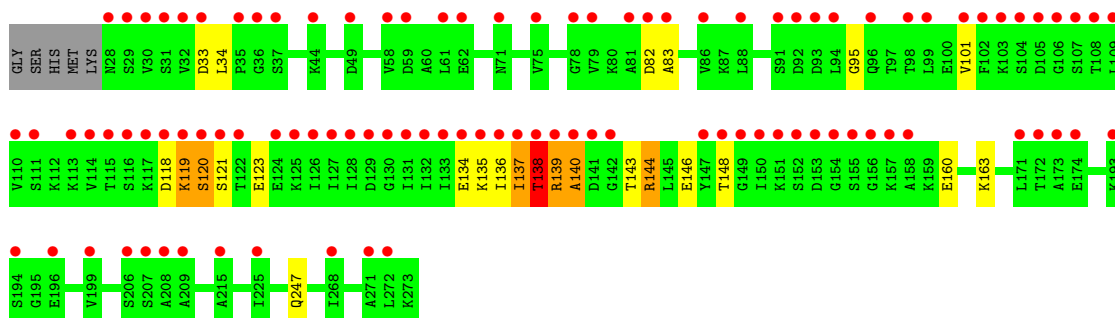
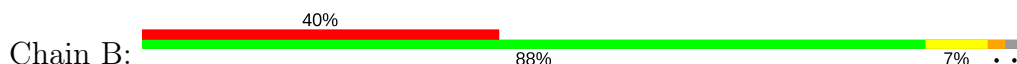
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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: Outer surface protein A



• Molecule 1: Outer surface protein A



4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	37.13Å 55.33Å 64.79Å 85.68° 74.05° 86.45°	Depositor
Resolution (Å)	20.00 – 1.35 29.37 – 1.35	Depositor EDS
% Data completeness (in resolution range)	95.2 (20.00-1.35) 93.5 (29.37-1.35)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.27 (at 1.35Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.184 , 0.218 0.212 , 0.241	Depositor DCC
R_{free} test set	5139 reflections (4.99%)	DCC
Wilson B-factor (Å ²)	13.7	Xtriage
Anisotropy	0.378	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.40 , 54.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	4335	wwPDB-VP
Average B, all atoms (Å ²)	18.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.51% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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	0.64	0/1975	0.78	0/2653
1	B	0.56	0/1875	0.70	0/2523
All	All	0.60	0/3850	0.74	0/5176

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	B	0	3

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	119	LYS	Peptide
1	B	137	ILE	Peptide
1	B	138	THR	Peptide

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	1921	0	2041	16	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	1846	0	1940	31	0
2	A	350	0	0	6	0
2	B	218	0	0	8	0
All	All	4335	0	3981	47	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 6.

The worst 5 of 47 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:119:LYS:HB2	1:B:120:SER:HB3	1.29	1.15
1:A:69:LYS:HE2	2:A:524:HOH:O	1.52	1.07
1:B:119:LYS:HB2	1:B:120:SER:CB	1.95	0.97
1:B:139:ARG:HB2	1:B:143:THR:O	1.72	0.88
1:B:119:LYS:CB	1:B:120:SER:HB3	2.04	0.87

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	263/251 (105%)	259 (98%)	4 (2%)	0	100	100
1	B	251/251 (100%)	243 (97%)	5 (2%)	3 (1%)	15	1
All	All	514/502 (102%)	502 (98%)	9 (2%)	3 (1%)	28	7

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	120	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	140	ALA
1	B	138	THR

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	231/216 (107%)	229 (99%)	2 (1%)	82	55
1	B	219/216 (101%)	216 (99%)	3 (1%)	71	37
All	All	450/432 (104%)	445 (99%)	5 (1%)	75	47

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	119	LYS
1	A	193	LYS
1	B	138	THR
1	B	139	ARG
1	B	144	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	270	ASN
1	B	270	ASN
1	B	96	GLN
1	A	246	GLN
1	B	247	GLN

5.3.3 RNA ⓘ

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 carbohydrates 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	250/251 (99%)	0.74	20 (8%) 13 16	9, 14, 23, 29	0
1	B	246/251 (98%)	2.08	100 (40%) 0 1	11, 22, 29, 34	0
All	All	496/502 (98%)	1.40	120 (24%) 1 1	9, 16, 28, 34	0

The worst 5 of 120 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	119	LYS	10.0
1	B	120	SER	9.7
1	B	28	ASN	9.1
1	B	152	SER	7.5
1	B	127	ILE	7.2

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 carbohydrates 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.