wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 5, 2024 – 01:21 pm BST

PDB ID	:	6EU4
Title	:	Structure of Acinetobacter phage vb_AbaP_AS12 gp42 tailspike
Authors	:	Taylor, N.M.I.; Shneider, M.M.; Leiman, P.G.
Deposited on	:	2017-10-27
Resolution	:	1.79 Å(reported)

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

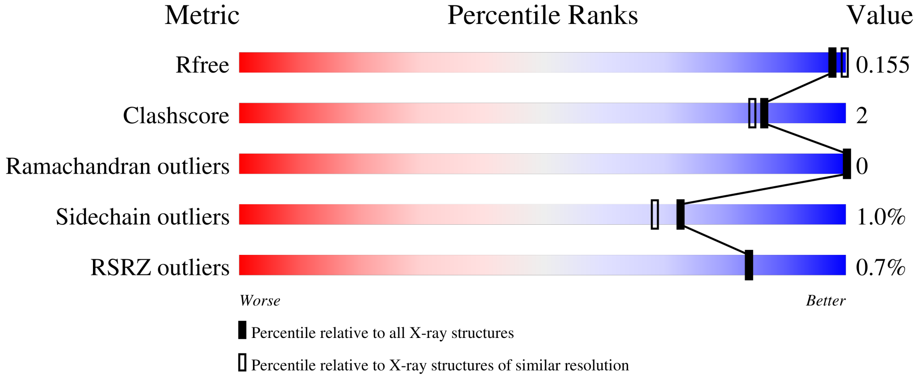
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, 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

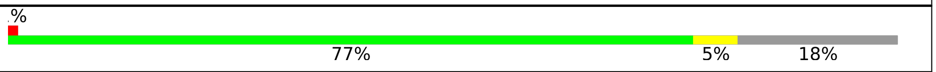
The reported resolution of this entry is 1.79 Å.

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}	164625	7108 (1.80-1.80)
Clashscore	180529	8162 (1.80-1.80)
Ramachandran outliers	177936	8077 (1.80-1.80)
Sidechain outliers	177891	8076 (1.80-1.80)
RSRZ outliers	164620	7108 (1.80-1.80)

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	716	
1	B	716	
1	C	716	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 28742 atoms, of which 13118 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 Tail spike protein.

Mol	Chain	Residues	Atoms							ZeroOcc	AltConf	Trace
1	A	587	Total 8899	C 2860	H 4387	N 746	O 884	S 11	Se 11	0	14	0
1	B	587	Total 8857	C 2848	H 4362	N 744	O 881	S 11	Se 11	0	12	0
1	C	587	Total 8866	C 2850	H 4369	N 743	O 882	S 11	Se 11	0	12	0

There are 15 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	GLY	-	expression tag	UNP A0A218KRF6
A	2	SER	-	expression tag	UNP A0A218KRF6
A	3	GLY	-	expression tag	UNP A0A218KRF6
A	4	SER	-	expression tag	UNP A0A218KRF6
A	5	THR	-	expression tag	UNP A0A218KRF6
B	1	GLY	-	expression tag	UNP A0A218KRF6
B	2	SER	-	expression tag	UNP A0A218KRF6
B	3	GLY	-	expression tag	UNP A0A218KRF6
B	4	SER	-	expression tag	UNP A0A218KRF6
B	5	THR	-	expression tag	UNP A0A218KRF6
C	1	GLY	-	expression tag	UNP A0A218KRF6
C	2	SER	-	expression tag	UNP A0A218KRF6
C	3	GLY	-	expression tag	UNP A0A218KRF6
C	4	SER	-	expression tag	UNP A0A218KRF6
C	5	THR	-	expression tag	UNP A0A218KRF6

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total 1	Zn 1	0	0

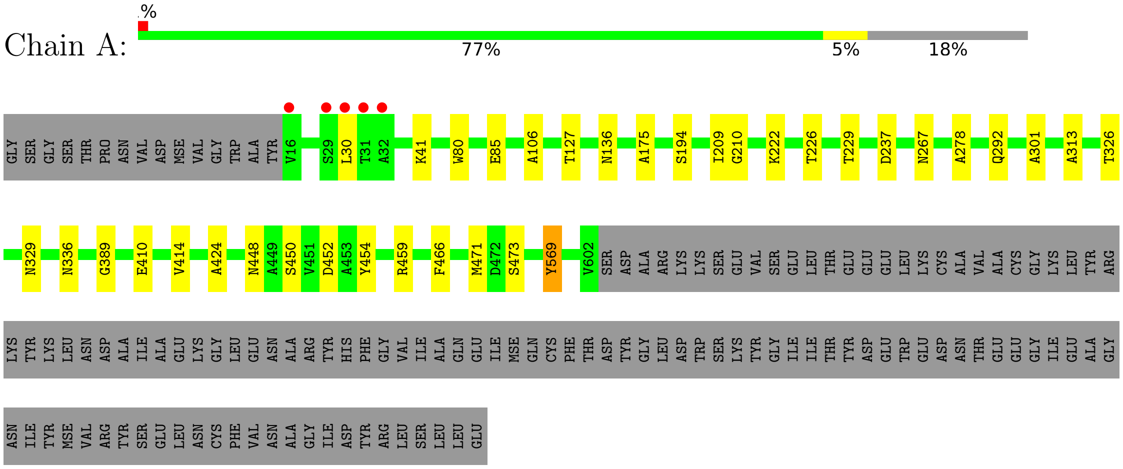
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	710	Total O 710 710	0	0
3	B	723	Total O 723 723	0	0
3	C	686	Total O 686 686	0	0

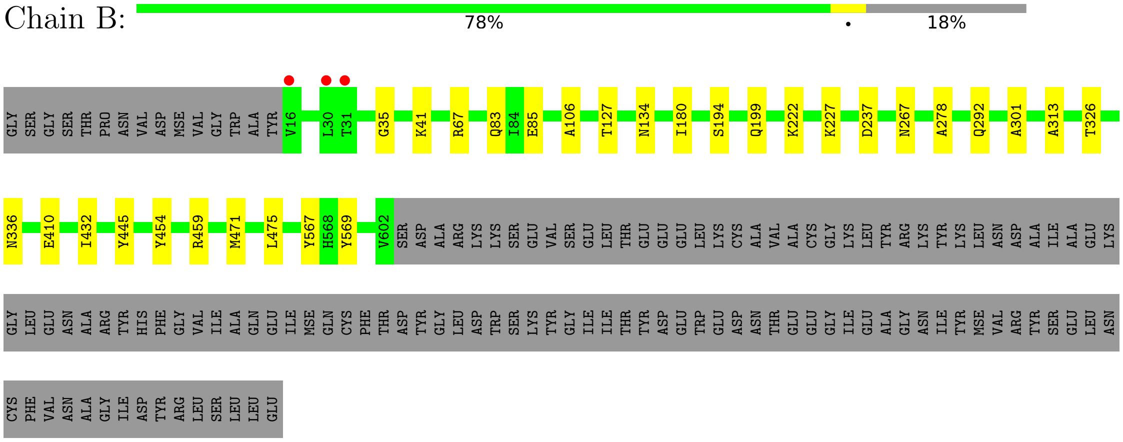
3 Residue-property plots ⓘ

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.

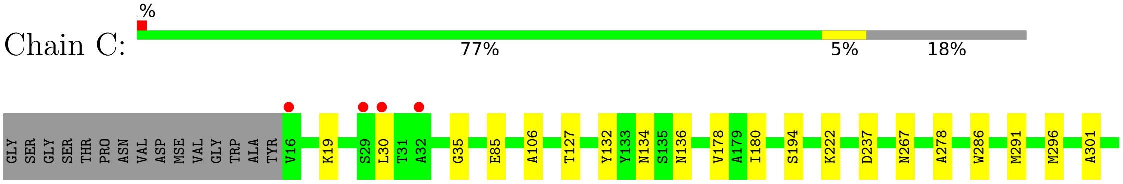
- Molecule 1: Tail spike protein

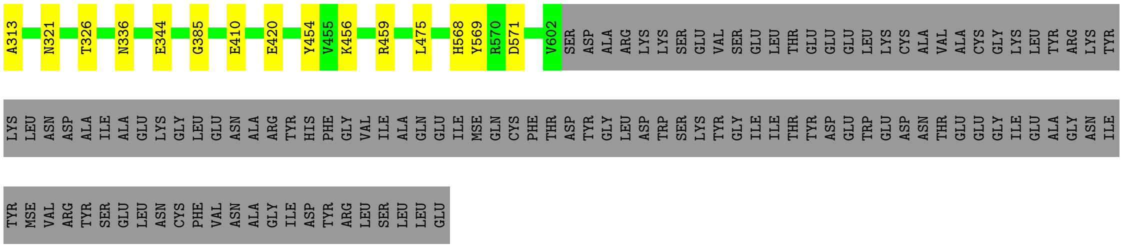


- Molecule 1: Tail spike protein



- Molecule 1: Tail spike protein




A313
N321
T326
N336
E344
G385
E410
E420
Y454
V455
K456
R459
L475
H568
Y569
R570
D571
V602
SER
ASP
ALA
ARG
LYS
LYS
SER
GLU
VAL
SER
GLU
LEU
THR
GLU
GLU
GLU
LEU
LYS
CYS
ALA
VAL
ALA
CYS
GLY
LYS
LEU
TYR
ARG
LYS
TYR
LYS
LEU
ASN
ASP
ALA
ILE
ALA
GLU
LYS
GLY
LEU
GLU
ASN
ALA
ARG
TYR
HIS
PHE
GLY
VAL
ILE
ALA
GLN
GLU
ILE
MSE
GLN
CYS
PHE
THR
ASP
TYR
GLY
LEU
ASP
TRP
SER
LYS
TYR
GLY
ILE
ILE
THR
TYR
ASP
GLU
TRP
GLU
ASP
ASN
THR
GLU
GLU
GLY
ILE
GLU
ALA
GLY
ASN
ILE
TYR
MSE
VAL
ARG
TYR
SER
GLU
LEU
ASN
CYS
PHE
VAL
ASN
ALA
GLY
ILE
ASP
TYR
ARG
LEU
SER
LEU
LEU
GLU

4 Data and refinement statistics ⓘ

Property	**Value**	**Source**
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	93.20Å 143.34Å 176.51Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.77 – 1.79 47.77 – 1.79	Depositor EDS
% Data completeness (in resolution range)	99.4 (47.77-1.79) 99.5 (47.77-1.79)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.99 (at 1.79Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.125 , 0.146 0.136 , 0.155	Depositor DCC
R_{free} test set	10966 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	17.0	Xtriage
Anisotropy	0.255	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 47.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	28742	wwPDB-VP
Average B, all atoms (Å^2)	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.54% 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.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: ZN

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	0.41	0/4651	0.60	0/6323
1	B	0.42	0/4629	0.62	0/6294
1	C	0.41	0/4633	0.61	0/6298
All	All	0.41	0/13913	0.61	0/18915

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 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	4512	4387	4386	22	0
1	B	4495	4362	4352	19	0
1	C	4497	4369	4369	18	0
2	A	1	0	0	0	0
3	A	710	0	0	4	4
3	B	723	0	0	3	1
3	C	686	0	0	2	3
All	All	15624	13118	13107	53	4

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 2.

The worst 5 of 53 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:267[A]:ASN:ND2	3:A:901:HOH:O	1.99	0.95
1:B:292[B]:GLN:NE2	3:B:801:HOH:O	1.99	0.93
1:C:85:GLU:OE2	3:C:801:HOH:O	2.01	0.78
1:A:210:GLY:HA3	1:A:226[B]:THR:HG21	1.67	0.76
1:B:83:GLN:OE1	3:B:802:HOH:O	2.08	0.72

All (4) 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)
3:A:1333:HOH:O	3:B:930:HOH:O[4_466]	1.98	0.22
3:A:1437:HOH:O	3:C:1316:HOH:O[3_546]	2.09	0.11
3:A:1437:HOH:O	3:C:1275:HOH:O[3_546]	2.09	0.11
3:A:1437:HOH:O	3:C:1238:HOH:O[3_546]	2.15	0.05

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	601/716 (84%)	585 (97%)	16 (3%)	0	100	100
1	B	598/716 (84%)	581 (97%)	17 (3%)	0	100	100
1	C	599/716 (84%)	584 (98%)	15 (2%)	0	100	100
All	All	1798/2148 (84%)	1750 (97%)	48 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	493/572 (86%)	487 (99%)	6 (1%)	67	62
1	B	490/572 (86%)	487 (99%)	3 (1%)	84	82
1	C	491/572 (86%)	486 (99%)	5 (1%)	73	68
All	All	1474/1716 (86%)	1460 (99%)	14 (1%)	73	72

5 of 14 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	410	GLU
1	B	569	TYR
1	C	569	TYR
1	C	326	THR
1	C	410	GLU

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

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 1 ligands modelled in this entry, 1 is monoatomic - 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

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	577/716 (80%)	-0.89	5 (0%)	81	80	6, 20, 42, 79	13 (2%)
1	B	577/716 (80%)	-0.86	3 (0%)	87	87	5, 20, 42, 78	10 (1%)
1	C	577/716 (80%)	-0.89	4 (0%)	84	84	6, 21, 42, 78	11 (1%)
All	All	1731/2148 (80%)	-0.88	12 (0%)	84	84	5, 20, 42, 79	34 (1%)

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	16	VAL	4.2
1	B	16	VAL	4.2
1	B	30	LEU	4.0
1	A	31	THR	3.8
1	C	16	VAL	3.5

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

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, 95th 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
2	ZN	A	801	1/1	0.98	0.15	34,34,34,34	0

6.5 Other polymers ⓘ

There are no such residues in this entry.