Full wwPDB X-ray Structure Validation Report (i)

Oct 4, 2023 – 06:14 PM EDT

PDB ID	:	6U51
Title	:	Anti-Sudan ebolavirus Nucleoprotein Single Domain Antibody Sudan B (SB) Complexed with Sudan ebolavirus Nucleoprotein C-terminal Domain 610-738
Authors	:	Taylor, A.B.; Sherwood, L.J.; Hart, P.J.; Hayhurst, A.
Deposited on	:	2019-08-26
Resolution	:	2.52 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:

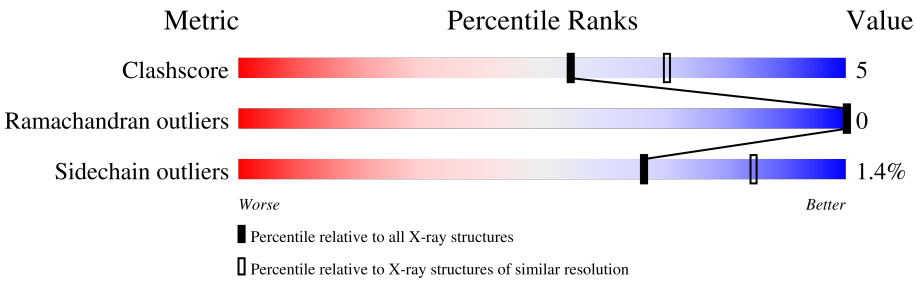
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	FAILED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 2.52 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	6463 (2.54-2.50)
Ramachandran outliers	138981	6335 (2.54-2.50)
Sidechain outliers	138945	6337 (2.54-2.50)

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%

Note EDS failed to run properly.

Mol	Chain	Length	Quality of chain
1	A	120	82% 16% • •
1	C	120	90% 8% •
2	B	141	56% 9% 35%
2	D	141	55% 8% 37%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 3345 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 Anti-Sudan ebolavirus Nucleoprotein Single Domain Antibody Sudan B (SB).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	C	118	901	552	168	177	4	0	0	0
1	A	119	908	557	169	178	4	0	0	0

- Molecule 2 is a protein called Nucleoprotein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	92	766	500	116	148	2	0	0	0
2	D	89	747	487	113	145	2	0	0	0

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	598	LYS	-	expression tag	UNP Q9QP77
B	599	ILE	-	expression tag	UNP Q9QP77
B	600	HIS	-	expression tag	UNP Q9QP77
B	601	HIS	-	expression tag	UNP Q9QP77
B	602	HIS	-	expression tag	UNP Q9QP77
B	603	HIS	-	expression tag	UNP Q9QP77
B	604	HIS	-	expression tag	UNP Q9QP77
B	605	HIS	-	expression tag	UNP Q9QP77
B	606	GLY	-	expression tag	UNP Q9QP77
B	607	GLY	-	expression tag	UNP Q9QP77
B	608	GLY	-	expression tag	UNP Q9QP77
B	609	SER	-	expression tag	UNP Q9QP77
D	598	LYS	-	expression tag	UNP Q9QP77
D	599	ILE	-	expression tag	UNP Q9QP77
D	600	HIS	-	expression tag	UNP Q9QP77

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	601	HIS	-	expression tag	UNP Q9QP77
D	602	HIS	-	expression tag	UNP Q9QP77
D	603	HIS	-	expression tag	UNP Q9QP77
D	604	HIS	-	expression tag	UNP Q9QP77
D	605	HIS	-	expression tag	UNP Q9QP77
D	606	GLY	-	expression tag	UNP Q9QP77
D	607	GLY	-	expression tag	UNP Q9QP77
D	608	GLY	-	expression tag	UNP Q9QP77
D	609	SER	-	expression tag	UNP Q9QP77

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	C	1	1	1	0	0
			Total	O		
3	B	7	7	7	0	0
			Total	O		
3	A	14	14	14	0	0
			Total	O		
3	D	1	1	1	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. 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. 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.

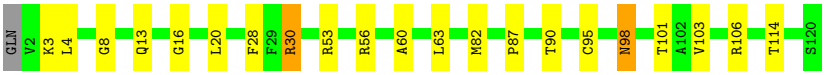
Note EDS failed to run properly.

- Molecule 1: Anti-Sudan ebolavirus Nucleoprotein Single Domain Antibody Sudan B (SB)

Chain C:

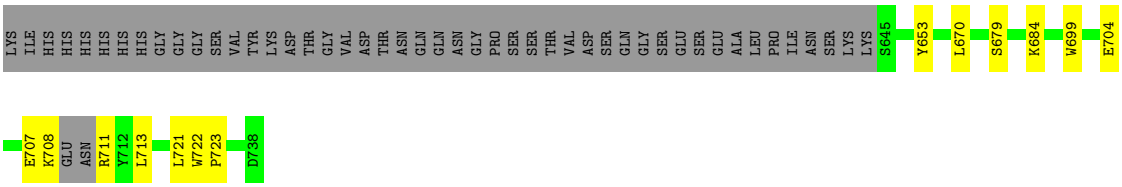
- Molecule 1: Anti-Sudan ebolavirus Nucleoprotein Single Domain Antibody Sudan B (SB)

Chain A:



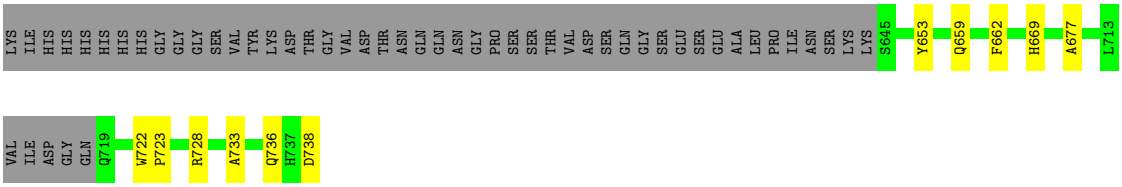
- Molecule 2: Nucleoprotein

Chain B:



- Molecule 2: Nucleoprotein

Chain D:



4 Data and refinement statistics (i)

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	P 61 2 2	Depositor
Cell constants a, b, c, α, β, γ	117.88Å 117.88Å 144.96Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	58.94 – 2.52	Depositor
% Data completeness (in resolution range)	99.5 (58.94-2.52)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.12 (at 2.52Å)	Xtriage
Refinement program	PHENIX 1.15.2_3472	Depositor
R, R_{free}	0.208 , 0.250	Depositor
Wilson B-factor (Å^2)	49.2	Xtriage
Anisotropy	0.305	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	3345	wwPDB-VP
Average B, all atoms (Å^2)	57.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.17% 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	0.31	0/924	0.57	0/1248
1	C	0.30	0/917	0.56	0/1238
2	B	0.30	0/789	0.47	0/1069
2	D	0.31	0/770	0.47	0/1043
All	All	0.30	0/3400	0.52	0/4598

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 (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	908	0	873	13	0
1	C	901	0	864	10	0
2	B	766	0	725	11	0
2	D	747	0	702	7	0
3	A	14	0	0	0	0
3	B	7	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
All	All	3345	0	3164	34	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 5.

All (34) 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)
2:B:708:LYS:HE3	2:B:711:ARG:HD3	1.56	0.85
1:A:101:THR:HG22	2:D:677:ALA:HB3	1.65	0.76
2:B:708:LYS:HG2	2:B:708:LYS:O	1.91	0.70
1:A:98:ASN:OD1	1:A:106:ARG:NH1	2.28	0.67
1:A:30:ARG:O	1:A:53:ARG:NH1	2.28	0.62
1:C:60:ALA:HB3	1:C:63:LEU:HD12	1.81	0.62
2:B:721:LEU:HG	2:B:723:PRO:HD2	1.82	0.62
1:C:48:VAL:HG13	1:C:63:LEU:HD13	1.82	0.61
1:C:30:ARG:NH1	1:C:53:ARG:O	2.39	0.56
2:B:699:TRP:HZ2	2:B:713:LEU:HD11	1.72	0.55
2:B:708:LYS:HE3	2:B:711:ARG:CD	2.34	0.52
1:C:103:VAL:HG11	2:B:653:TYR:CZ	2.46	0.51
1:A:8:GLY:O	1:A:114:THR:HG21	2.10	0.51
1:A:87:PRO:O	1:A:90:THR:HG22	2.13	0.49
2:B:708:LYS:O	2:B:708:LYS:CG	2.57	0.48
2:D:722:TRP:HB3	2:D:723:PRO:HD3	1.94	0.48
1:A:103:VAL:HG11	2:D:653:TYR:CZ	2.49	0.47
1:C:87:PRO:O	1:C:90:THR:HG22	2.14	0.47
1:C:61:PRO:HG3	2:D:659:GLN:HE22	1.79	0.46
1:A:60:ALA:HB3	1:A:63:LEU:HD12	1.98	0.44
2:B:670:LEU:HD22	1:A:56:ARG:HD3	1.99	0.44
2:B:722:TRP:HB3	2:B:723:PRO:HD3	2.00	0.43
2:D:733:ALA:O	2:D:736:GLN:NE2	2.40	0.43
1:C:90:THR:HG23	1:C:90:THR:O	2.18	0.43
1:A:3:LYS:HE2	1:A:28:PHE:CZ	2.54	0.42
1:A:13:GLN:O	1:A:16:GLY:N	2.52	0.42
1:A:20:LEU:HD11	1:A:82:MET:CE	2.50	0.42
1:C:56:ARG:HH22	2:D:669:HIS:CE1	2.38	0.41
2:B:679:SER:HA	2:B:684:LYS:O	2.20	0.41
1:A:4:LEU:HD13	1:A:95:CYS:SG	2.60	0.41
2:D:662:PHE:CD1	2:D:728:ARG:HA	2.55	0.41
1:C:60:ALA:HB3	1:C:63:LEU:CD1	2.49	0.41
1:A:30:ARG:HE	1:A:30:ARG:HB3	1.70	0.41
1:C:103:VAL:HG11	2:B:653:TYR:OH	2.21	0.41

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	117/120 (98%)	115 (98%)	2 (2%)	0	100	100
1	C	116/120 (97%)	111 (96%)	5 (4%)	0	100	100
2	B	88/141 (62%)	87 (99%)	1 (1%)	0	100	100
2	D	85/141 (60%)	84 (99%)	1 (1%)	0	100	100
All	All	406/522 (78%)	397 (98%)	9 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	95/96 (99%)	93 (98%)	2 (2%)	53	76
1	C	94/96 (98%)	94 (100%)	0	100	100
2	B	83/125 (66%)	81 (98%)	2 (2%)	49	73
2	D	81/125 (65%)	80 (99%)	1 (1%)	71	87
All	All	353/442 (80%)	348 (99%)	5 (1%)	67	85

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

Mol	Chain	Res	Type
2	B	704	GLU
2	B	707	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	30	ARG
1	A	98	ASN
2	D	738	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
2	D	659	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)

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

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

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

6.3 Carbohydrates (i)

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

6.4 Ligands (i)

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

6.5 Other polymers (i)

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