wwPDB X-ray Structure Validation Summary Report (i)

Sep 24, 2023 – 10:01 PM EDT

PDB ID	:	5UKB
Title	:	VSV N PROTEIN IN COMPLEX WITH INHIBITORY NANOBODY 1004
Authors	:	Hanke, L.; Knockenhauer, K.E.; Ploegh, H.L.; Schwartz, T.U.
Deposited on	:	2017-01-20
Resolution	:	5.47 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 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:

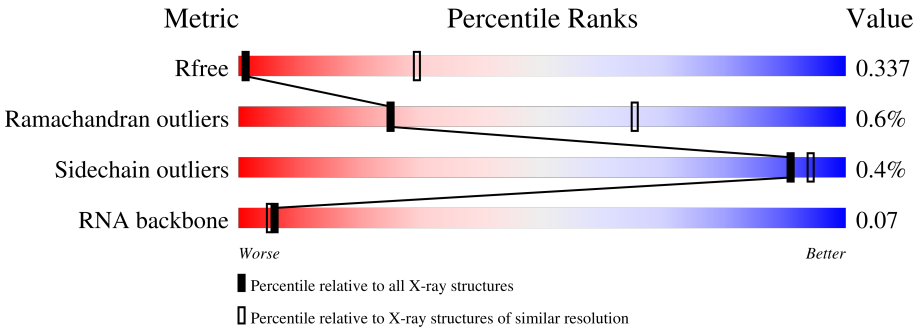
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 5.47 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))
R_{free}	130704	1019 (7.12-3.82)
Ramachandran outliers	138981	1014 (7.12-3.82)
Sidechain outliers	138945	1190 (7.12-3.80)
RNA backbone	3102	1074 (7.80-3.00)

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%

Mol	Chain	Length	Quality of chain
1	a	138	84% · 12%
1	b	138	86% · 12%
1	c	138	84% · 12%
1	d	138	86% · 12%
1	e	138	86% · 12%
2	A	423	96% · ·
2	B	423	97% · ·

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	C	423	97%
2	D	423	97%
2	E	423	97%
3	R	45	16% 64% 20%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 22210 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-vesicular stomatitis virus N VHH.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	d	122	Total	C	N	O	S	0	0	0
			946	596	166	181	3			
1	c	122	Total	C	N	O	S	0	0	0
			946	596	166	181	3			
1	b	122	Total	C	N	O	S	0	0	0
			946	596	166	181	3			
1	a	122	Total	C	N	O	S	0	0	0
			946	596	166	181	3			
1	e	122	Total	C	N	O	S	0	0	0
			946	596	166	181	3			

There are 80 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
d	1	GLN	-	expression tag	UNP A0A192B6J5
d	2	VAL	-	expression tag	UNP A0A192B6J5
d	125	GLY	-	expression tag	UNP A0A192B6J5
d	126	GLY	-	expression tag	UNP A0A192B6J5
d	127	LEU	-	expression tag	UNP A0A192B6J5
d	128	PRO	-	expression tag	UNP A0A192B6J5
d	129	GLU	-	expression tag	UNP A0A192B6J5
d	130	THR	-	expression tag	UNP A0A192B6J5
d	131	GLY	-	expression tag	UNP A0A192B6J5
d	132	GLY	-	expression tag	UNP A0A192B6J5
d	133	HIS	-	expression tag	UNP A0A192B6J5
d	134	HIS	-	expression tag	UNP A0A192B6J5
d	135	HIS	-	expression tag	UNP A0A192B6J5
d	136	HIS	-	expression tag	UNP A0A192B6J5
d	137	HIS	-	expression tag	UNP A0A192B6J5
d	138	HIS	-	expression tag	UNP A0A192B6J5
c	1	GLN	-	expression tag	UNP A0A192B6J5
c	2	VAL	-	expression tag	UNP A0A192B6J5
c	125	GLY	-	expression tag	UNP A0A192B6J5

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
c	126	GLY	-	expression tag	UNP A0A192B6J5
c	127	LEU	-	expression tag	UNP A0A192B6J5
c	128	PRO	-	expression tag	UNP A0A192B6J5
c	129	GLU	-	expression tag	UNP A0A192B6J5
c	130	THR	-	expression tag	UNP A0A192B6J5
c	131	GLY	-	expression tag	UNP A0A192B6J5
c	132	GLY	-	expression tag	UNP A0A192B6J5
c	133	HIS	-	expression tag	UNP A0A192B6J5
c	134	HIS	-	expression tag	UNP A0A192B6J5
c	135	HIS	-	expression tag	UNP A0A192B6J5
c	136	HIS	-	expression tag	UNP A0A192B6J5
c	137	HIS	-	expression tag	UNP A0A192B6J5
c	138	HIS	-	expression tag	UNP A0A192B6J5
b	1	GLN	-	expression tag	UNP A0A192B6J5
b	2	VAL	-	expression tag	UNP A0A192B6J5
b	125	GLY	-	expression tag	UNP A0A192B6J5
b	126	GLY	-	expression tag	UNP A0A192B6J5
b	127	LEU	-	expression tag	UNP A0A192B6J5
b	128	PRO	-	expression tag	UNP A0A192B6J5
b	129	GLU	-	expression tag	UNP A0A192B6J5
b	130	THR	-	expression tag	UNP A0A192B6J5
b	131	GLY	-	expression tag	UNP A0A192B6J5
b	132	GLY	-	expression tag	UNP A0A192B6J5
b	133	HIS	-	expression tag	UNP A0A192B6J5
b	134	HIS	-	expression tag	UNP A0A192B6J5
b	135	HIS	-	expression tag	UNP A0A192B6J5
b	136	HIS	-	expression tag	UNP A0A192B6J5
b	137	HIS	-	expression tag	UNP A0A192B6J5
b	138	HIS	-	expression tag	UNP A0A192B6J5
a	1	GLN	-	expression tag	UNP A0A192B6J5
a	2	VAL	-	expression tag	UNP A0A192B6J5
a	125	GLY	-	expression tag	UNP A0A192B6J5
a	126	GLY	-	expression tag	UNP A0A192B6J5
a	127	LEU	-	expression tag	UNP A0A192B6J5
a	128	PRO	-	expression tag	UNP A0A192B6J5
a	129	GLU	-	expression tag	UNP A0A192B6J5
a	130	THR	-	expression tag	UNP A0A192B6J5
a	131	GLY	-	expression tag	UNP A0A192B6J5
a	132	GLY	-	expression tag	UNP A0A192B6J5
a	133	HIS	-	expression tag	UNP A0A192B6J5
a	134	HIS	-	expression tag	UNP A0A192B6J5
a	135	HIS	-	expression tag	UNP A0A192B6J5

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
a	136	HIS	-	expression tag	UNP A0A192B6J5
a	137	HIS	-	expression tag	UNP A0A192B6J5
a	138	HIS	-	expression tag	UNP A0A192B6J5
e	1	GLN	-	expression tag	UNP A0A192B6J5
e	2	VAL	-	expression tag	UNP A0A192B6J5
e	125	GLY	-	expression tag	UNP A0A192B6J5
e	126	GLY	-	expression tag	UNP A0A192B6J5
e	127	LEU	-	expression tag	UNP A0A192B6J5
e	128	PRO	-	expression tag	UNP A0A192B6J5
e	129	GLU	-	expression tag	UNP A0A192B6J5
e	130	THR	-	expression tag	UNP A0A192B6J5
e	131	GLY	-	expression tag	UNP A0A192B6J5
e	132	GLY	-	expression tag	UNP A0A192B6J5
e	133	HIS	-	expression tag	UNP A0A192B6J5
e	134	HIS	-	expression tag	UNP A0A192B6J5
e	135	HIS	-	expression tag	UNP A0A192B6J5
e	136	HIS	-	expression tag	UNP A0A192B6J5
e	137	HIS	-	expression tag	UNP A0A192B6J5
e	138	HIS	-	expression tag	UNP A0A192B6J5

- Molecule 2 is a protein called Nucleocapsid.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	419	Total	C	N	O	S	0	0	0
			3316	2111	556	631	18			
2	C	419	Total	C	N	O	S	0	0	0
			3316	2111	556	631	18			
2	B	419	Total	C	N	O	S	0	0	0
			3316	2111	556	631	18			
2	A	419	Total	C	N	O	S	0	0	0
			3316	2111	556	631	18			
2	E	419	Total	C	N	O	S	0	0	0
			3316	2111	556	631	18			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	0	MET	-	expression tag	UNP A6H4P1
D	1	ALA	-	expression tag	UNP A6H4P1
C	0	MET	-	expression tag	UNP A6H4P1
C	1	ALA	-	expression tag	UNP A6H4P1
B	0	MET	-	expression tag	UNP A6H4P1

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	1	ALA	-	expression tag	UNP A6H4P1
A	0	MET	-	expression tag	UNP A6H4P1
A	1	ALA	-	expression tag	UNP A6H4P1
E	0	MET	-	expression tag	UNP A6H4P1
E	1	ALA	-	expression tag	UNP A6H4P1

- Molecule 3 is a RNA chain called RNA (45-MER).

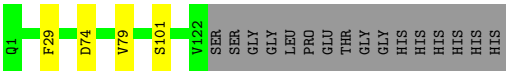
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
3	R	45	900	405	90	360	45	0	0	0

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.

- Molecule 1: Anti-vesicular stomatitis virus N VHH

Chain d:

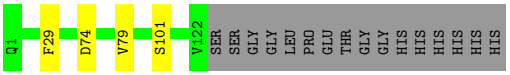


- Molecule 1: Anti-vesicular stomatitis virus N VHH

Chain c:

- Molecule 1: Anti-vesicular stomatitis virus N VHH

Chain b:

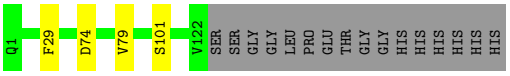


- Molecule 1: Anti-vesicular stomatitis virus N VHH

Chain a:

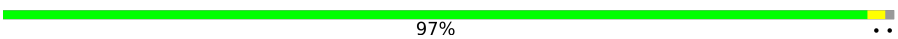
- Molecule 1: Anti-vesicular stomatitis virus N VHH

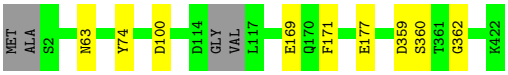
Chain e:



- Molecule 2: Nucleocapsid

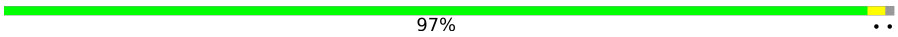
Chain D:

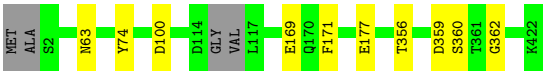




• Molecule 2: Nucleocapsid

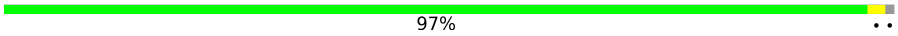
Chain C:

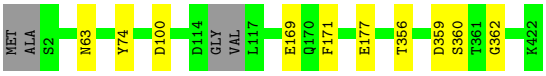




• Molecule 2: Nucleocapsid

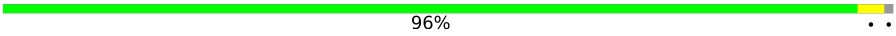
Chain B:

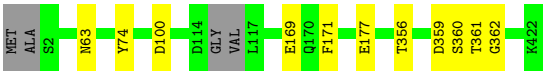




• Molecule 2: Nucleocapsid

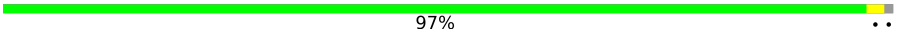
Chain A:





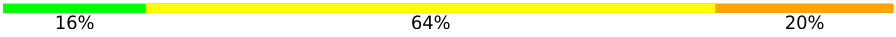
• Molecule 2: Nucleocapsid

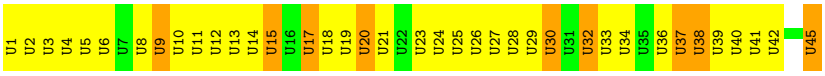
Chain E:



• Molecule 3: RNA (45-MER)

Chain R:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	240.12Å 335.50Å 75.90Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	137.51 – 5.47 137.51 – 5.47	Depositor EDS
% Data completeness (in resolution range)	97.4 (137.51-5.47) 97.4 (137.51-5.47)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.17	Depositor
$< I/\sigma(I) >$ [1]	1.24 (at 5.42Å)	Xtriage
Refinement program	PHENIX (1.11.1_2575: ???)	Depositor
R, R_{free}	0.336 , 0.338 0.336 , 0.337	Depositor DCC
R_{free} test set	1997 reflections (9.70%)	wwPDB-VP
Wilson B-factor (Å^2)	292.7	Xtriage
Anisotropy	0.275	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.40 , 423.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.39$, $< L^2 > = 0.22$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.81	EDS
Total number of atoms	22210	wwPDB-VP
Average B, all atoms (Å^2)	309.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.78% 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

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	0.25	0/967	0.55	1/1312 (0.1%)
1	b	0.25	0/967	0.55	1/1312 (0.1%)
1	c	0.26	0/967	0.55	1/1312 (0.1%)
1	d	0.25	0/967	0.55	1/1312 (0.1%)
1	e	0.26	0/967	0.55	1/1312 (0.1%)
2	A	0.24	0/3391	0.41	0/4589
2	B	0.24	0/3391	0.41	0/4589
2	C	0.24	0/3391	0.41	0/4589
2	D	0.24	0/3391	0.41	0/4589
2	E	0.24	0/3391	0.41	0/4589
3	R	0.94	0/989	1.56	10/1526 (0.7%)
All	All	0.31	0/22779	0.56	15/31031 (0.0%)

There are no bond length outliers.

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	R	20	U	P-O3'-C3'	8.95	130.44	119.70
3	R	9	U	O4'-C1'-N1	6.60	113.48	108.20
3	R	9	U	O4'-C4'-C3'	-6.11	97.89	104.00
3	R	34	U	O4'-C1'-N1	5.88	112.91	108.20
3	R	15	U	C2-N1-C1'	5.79	124.65	117.70

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

Due to software issues we are unable to calculate clashes - this section is therefore empty.

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	120/138 (87%)	110 (92%)	8 (7%)	2 (2%)	9	41
1	b	120/138 (87%)	110 (92%)	8 (7%)	2 (2%)	9	41
1	c	120/138 (87%)	110 (92%)	8 (7%)	2 (2%)	9	41
1	d	120/138 (87%)	110 (92%)	8 (7%)	2 (2%)	9	41
1	e	120/138 (87%)	110 (92%)	8 (7%)	2 (2%)	9	41
2	A	415/423 (98%)	395 (95%)	19 (5%)	1 (0%)	47	81
2	B	415/423 (98%)	395 (95%)	19 (5%)	1 (0%)	47	81
2	C	415/423 (98%)	395 (95%)	19 (5%)	1 (0%)	47	81
2	D	415/423 (98%)	395 (95%)	19 (5%)	1 (0%)	47	81
2	E	415/423 (98%)	395 (95%)	19 (5%)	1 (0%)	47	81
All	All	2675/2805 (95%)	2525 (94%)	135 (5%)	15 (1%)	25	65

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	d	101	SER
1	c	101	SER
1	b	101	SER
1	a	101	SER
1	e	101	SER

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	98/110 (89%)	97 (99%)	1 (1%)	76	86
1	b	98/110 (89%)	97 (99%)	1 (1%)	76	86
1	c	98/110 (89%)	97 (99%)	1 (1%)	76	86
1	d	98/110 (89%)	97 (99%)	1 (1%)	76	86
1	e	98/110 (89%)	97 (99%)	1 (1%)	76	86
2	A	361/363 (99%)	360 (100%)	1 (0%)	92	95
2	B	361/363 (99%)	360 (100%)	1 (0%)	92	95
2	C	361/363 (99%)	360 (100%)	1 (0%)	92	95
2	D	361/363 (99%)	360 (100%)	1 (0%)	92	95
2	E	361/363 (99%)	360 (100%)	1 (0%)	92	95
All	All	2295/2365 (97%)	2285 (100%)	10 (0%)	91	94

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

Mol	Chain	Res	Type
2	A	100	ASP
1	e	29	PHE
2	E	100	ASP
2	C	100	ASP
1	b	29	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	b	82	GLN
2	B	208	HIS
2	E	208	HIS
2	A	208	HIS
1	e	82	GLN

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
3	R	44/45 (97%)	35 (79%)	8 (18%)

5 of 35 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
3	R	2	U
3	R	3	U
3	R	4	U
3	R	5	U
3	R	6	U

5 of 8 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
3	R	39	U
3	R	32	U
3	R	14	U
3	R	12	U
3	R	20	U

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

6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.