WORLDWIDE PDB
PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report ⓘ

Feb 26, 2014 – 04:36 PM GMT

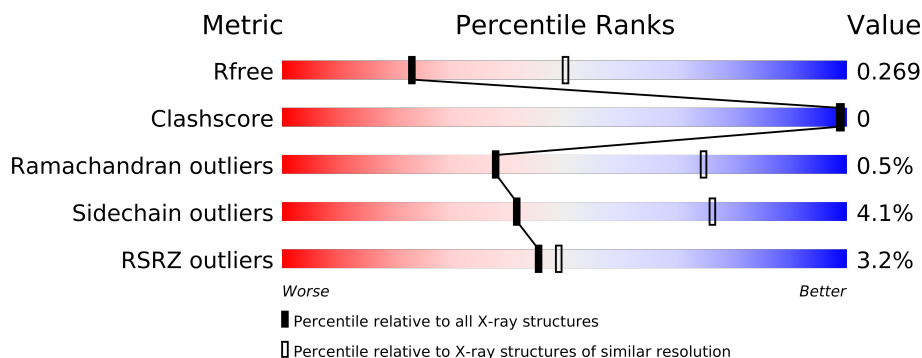|		
PDB ID	:	4H5L
Title	:	Crystal Structure of Toscana Virus Nucleocapsid Protein Hexamer
Authors	:	Raymond, D.D.; Smith, J.L.
Deposited on	:	2012-09-18
Resolution	:	2.75 Å(reported)

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

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

1   Overall quality at a glance ⓘ

The reported resolution of this entry is 2.75 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	Percentile Ranks	Value
Rfree		0.269
Clashscore		0
Ramachandran outliers		0.5%
Sidechain outliers		4.1%
RSRZ outliers		3.2%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	66092	2406 (2.80-2.72)
Clashscore	79885	2995 (2.80-2.72)
Ramachandran outliers	78287	2941 (2.80-2.72)
Sidechain outliers	78261	2944 (2.80-2.72)
RSRZ outliers	66119	2409 (2.80-2.72)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	253	
1	B	253	
1	C	253	
1	D	253	
1	E	253	
1	F	253	

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 11238 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Nucleoprotein.

| Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	243	Total	C	N	O	S	0	0	0
			1869	1186	325	347	11			
1	B	244	Total	C	N	O	S	0	0	0
			1870	1187	325	347	11			
1	C	244	Total	C	N	O	S	0	0	0
			1873	1188	326	348	11			
1	D	244	Total	C	N	O	S	0	0	0
			1873	1188	326	348	11			
1	E	244	Total	C	N	O	S	0	0	0
			1873	1188	326	348	11			
1	F	244	Total	C	N	O	S	0	0	0
			1863	1182	322	348	11			

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	O	0	0
			1	1		
2	B	2	Total	O	0	0
			2	2		
2	C	2	Total	O	0	0
			2	2		
2	D	5	Total	O	0	0
			5	5		
2	E	3	Total	O	0	0
			3	3		
2	F	4	Total	O	0	0
			4	4		

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

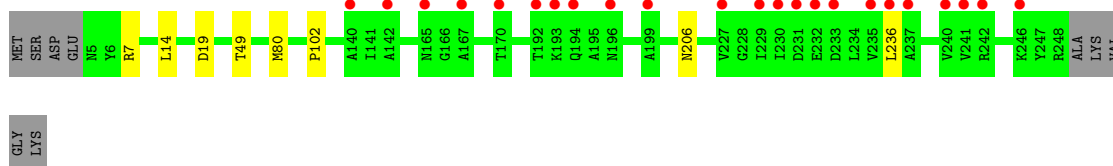- Molecule 1: Nucleoprotein

Chain A:
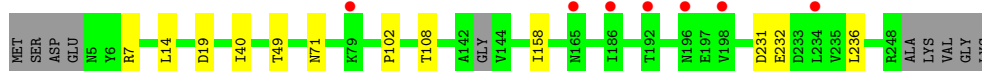
- Molecule 1: Nucleoprotein

Chain B:

- Molecule 1: Nucleoprotein
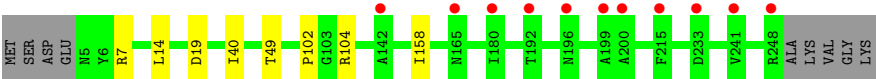
Chain C:

- Molecule 1: Nucleoprotein

Chain D:

- Molecule 1: Nucleoprotein

Chain E:

- Molecule 1: Nucleoprotein

Chain F:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	50.35Å 93.75Å 95.52Å 67.92° 85.94° 87.95°	Depositor
Resolution (Å)	50.22 – 2.75 50.22 – 2.75	Depositor EDS
% Data completeness (in resolution range)	95.2 (50.22-2.75) 95.2 (50.22-2.75)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.09	Depositor
$< I/\sigma(I) >$ [1]	1.89 (at 2.77Å)	Xtriage
Refinement program	BUSTER 2.10	Depositor
R, R_{free}	0.219 , 0.238 0.251 , 0.269	Depositor DCC
R_{free} test set	2020 reflections (5.06%)	DCC
Wilson B-factor (Å^2)	52.5	Xtriage
Anisotropy	0.413	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 41.0	EDS
Estimated twinning fraction	0.020 for -h,-l,-k	Xtriage
L-test for twinning	$< |L| > = 0.46, < L^2 > = 0.28$	Xtriage
Outliers	0 of 39945 reflections	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	11238	wwPDB-VP
Average B, all atoms (Å^2)	82.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.99% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

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	0.45	0/1900	0.60	0/2565
1	B	0.45	0/1902	0.61	0/2569
1	C	0.45	0/1905	0.59	0/2573
1	D	0.45	0/1905	0.60	0/2573
1	E	0.45	0/1905	0.58	0/2573
1	F	0.45	0/1895	0.59	0/2562
All	All	0.45	0/11412	0.60	0/15415

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1869	0	0	0	0
1	B	1870	0	0	0	0
1	C	1873	0	0	1	0
1	D	1873	0	0	3	0
1	E	1873	0	0	2	0
1	F	1863	0	0	1	0
2	A	1	0	0	0	0
2	B	2	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	C	2	0	0	0	0
2	D	5	0	0	0	0
2	E	3	0	0	0	0
2	F	4	0	0	0	0
All	All	11238	0	0	4	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 0.

All (4) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:C:80:MET:O	1:D:104:ARG:NE	2.46	0.48
1:D:173:ASP:OD2	1:D:248:ARG:NH2	2.52	0.42
1:E:80:MET:O	1:F:104:ARG:NE	2.53	0.41
1:D:80:MET:O	1:E:104:ARG:NH2	2.54	0.40

There are no symmetry-related clashes.

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	239/253 (94%)	226 (95%)	11 (5%)	2 (1%)	27	63
1	B	242/253 (96%)	231 (96%)	10 (4%)	1 (0%)	43	79
1	C	242/253 (96%)	230 (95%)	11 (4%)	1 (0%)	43	79
1	D	242/253 (96%)	231 (96%)	10 (4%)	1 (0%)	43	79
1	E	242/253 (96%)	234 (97%)	7 (3%)	1 (0%)	43	79
1	F	242/253 (96%)	233 (96%)	8 (3%)	1 (0%)	43	79
All	All	1449/1518 (96%)	1385 (96%)	57 (4%)	7 (0%)	38	74

All (7) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	232	GLU
1	A	102	PRO
1	B	102	PRO
1	C	102	PRO
1	D	102	PRO
1	E	102	PRO
1	F	102	PRO

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	198/206 (96%)	188 (95%)	10 (5%)	33	67
1	B	197/206 (96%)	189 (96%)	8 (4%)	41	76
1	C	198/206 (96%)	192 (97%)	6 (3%)	53	86
1	D	198/206 (96%)	188 (95%)	10 (5%)	33	67
1	E	198/206 (96%)	189 (96%)	9 (4%)	38	72
1	F	196/206 (95%)	190 (97%)	6 (3%)	52	85
All	All	1185/1236 (96%)	1136 (96%)	49 (4%)	41	76

All (49) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	7	ARG
1	A	14	LEU
1	A	19	ASP
1	A	40	ILE
1	A	49	THR
1	A	71	ASN
1	A	108	THR
1	A	158	ILE
1	A	231	ASP
1	A	236	LEU
1	B	7	ARG
1	B	14	LEU
1	B	19	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	40	ILE
1	B	49	THR
1	B	71	ASN
1	B	158	ILE
1	B	235	VAL
1	C	7	ARG
1	C	14	LEU
1	C	19	ASP
1	C	49	THR
1	C	206	ASN
1	C	236	LEU
1	D	7	ARG
1	D	14	LEU
1	D	19	ASP
1	D	40	ILE
1	D	49	THR
1	D	138	MET
1	D	158	ILE
1	D	223	LEU
1	D	238	SER
1	D	248	ARG
1	E	7	ARG
1	E	14	LEU
1	E	19	ASP
1	E	40	ILE
1	E	49	THR
1	E	120	THR
1	E	158	ILE
1	E	206	ASN
1	E	233	ASP
1	F	7	ARG
1	F	14	LEU
1	F	19	ASP
1	F	40	ILE
1	F	49	THR
1	F	158	ILE

Some 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 chains 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 carbohydrates in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

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 (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	243/253 (96%)	0.21	7 (2%)	49	53	40, 80, 145, 240	0
1	B	244/253 (96%)	0.07	2 (0%)	83	85	34, 78, 131, 205	0
1	C	244/253 (96%)	0.48	23 (9%)	9	8	42, 86, 179, 229	0
1	D	244/253 (96%)	-0.04	1 (0%)	90	92	32, 65, 108, 142	0
1	E	244/253 (96%)	0.04	3 (1%)	75	78	30, 76, 138, 178	0
1	F	244/253 (96%)	0.14	11 (4%)	32	34	28, 70, 152, 233	0
All	All	1463/1518 (96%)	0.15	47 (3%)	45	49	28, 76, 150, 240	0

All (47) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	196	ASN	7.4
1	C	237	ALA	3.9
1	C	142	ALA	3.9
1	F	241	VAL	3.8
1	E	248	ARG	3.6
1	A	165	ASN	3.4
1	F	248	ARG	3.3
1	C	165	ASN	3.2
1	C	246	LYS	3.1
1	A	198	VAL	3.1
1	F	196	ASN	3.0
1	C	193	LYS	2.9
1	C	240	VAL	2.9
1	C	230	ILE	2.9
1	A	192	THR	2.9
1	E	161	LEU	2.8
1	C	241	VAL	2.8
1	A	186	ILE	2.8
1	A	234	LEU	2.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	227	VAL	2.7
1	C	194	GLN	2.7
1	C	199	ALA	2.6
1	F	165	ASN	2.6
1	C	232	GLU	2.5
1	F	199	ALA	2.5
1	B	202	PHE	2.5
1	C	236	LEU	2.5
1	C	167	ALA	2.5
1	F	215	PHE	2.5
1	F	142	ALA	2.4
1	C	235	VAL	2.3
1	C	170	THR	2.3
1	F	200	ALA	2.2
1	C	242	ARG	2.2
1	C	192	THR	2.2
1	F	180	ILE	2.2
1	F	192	THR	2.1
1	E	243	SER	2.1
1	A	196	ASN	2.1
1	C	231	ASP	2.1
1	B	97	LEU	2.1
1	D	70	GLY	2.1
1	C	140	ALA	2.1
1	C	229	ILE	2.1
1	C	233	ASP	2.1
1	A	79	LYS	2.0
1	F	233	ASP	2.0

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 carbohydrates in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.