



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 14, 2017 – 01:48 pm GMT

PDB ID : 2VF1
Title : X-RAY CRYSTALLOGRAPHIC STRUCTURE OF THE PICOBIR-NAVIRUS CAPSID
Authors : Duquerroy, S.; Da Costa, B.; Vigouroux, A.; Lepault, J.; Navaza, J.; Delmas, B.; Rey, F.A.
Deposited on : 2007-10-29
Resolution : 3.40 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

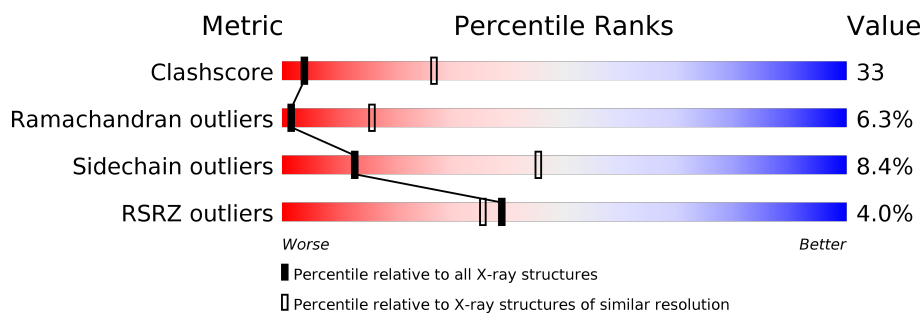
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.40 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	112137	1832 (3.50-3.30)
Ramachandran outliers	110173	1789 (3.50-3.30)
Sidechain outliers	110143	1789 (3.50-3.30)
RSRZ outliers	101464	1709 (3.50-3.30)

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	525	<div> <div>4%</div> <div>50%</div> <div>41%</div> <div>8%</div> </div>
1	B	525	<div> <div>4%</div> <div>50%</div> <div>40%</div> <div>9%</div> </div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 8286 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 CAPSID PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	525	Total	C	N	O	S	0	0	0
			4117	2634	671	794	18			
1	B	525	Total	C	N	O	S	0	0	0
			4117	2634	671	794	18			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	179	TYR	UNK	SEE REMARK 999	UNP Q9Q1V2
B	179	TYR	UNK	SEE REMARK 999	UNP Q9Q1V2

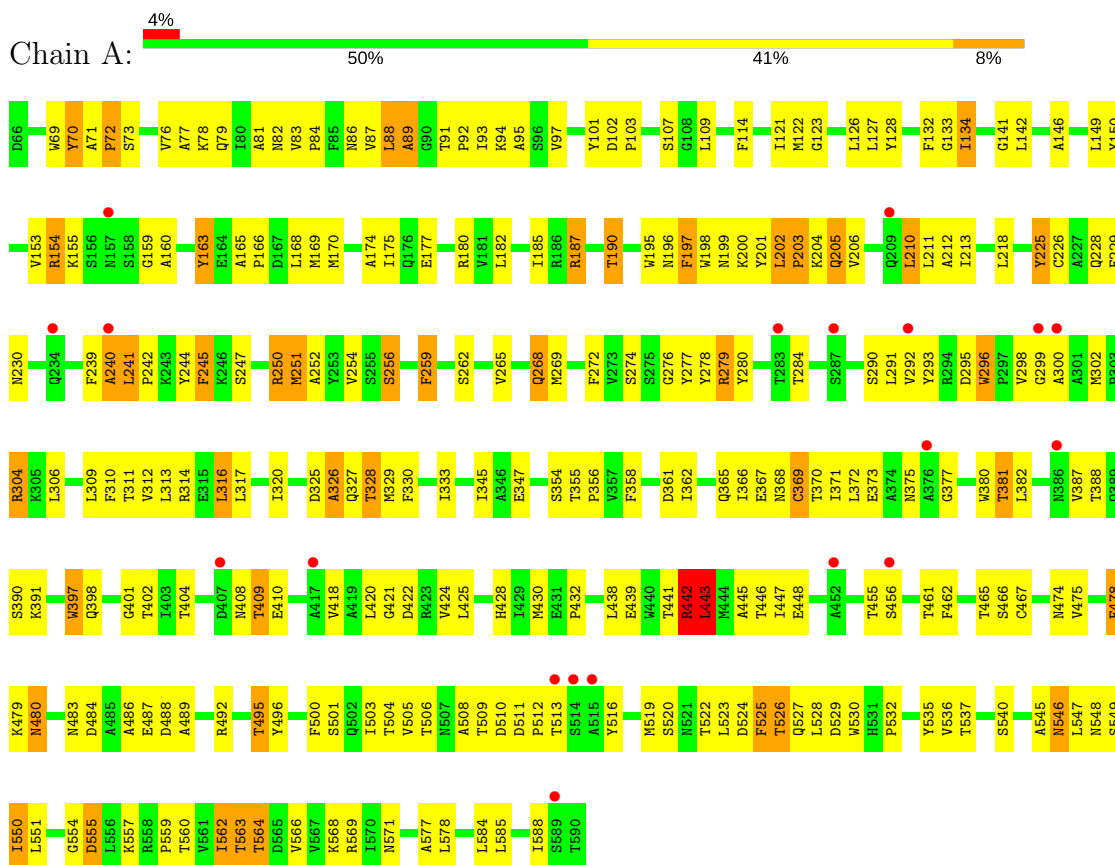
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	22	Total	O	0	0
			22	22		
2	B	30	Total	O	0	0
			30	30		

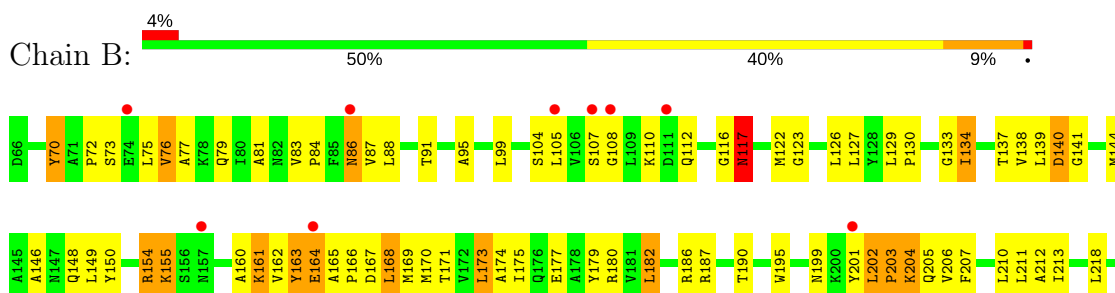
3 Residue-property plots

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



• Molecule 1: CAPSID PROTEIN



N548	N549	N550	G554	D555	L556	K557	P559	T560	V561	L562	K568	N571	A577	L578	S581	A582	N583	L584	L585	S586	N587	L588	S589	T590																												
V475	L476	Y477	F478	K479	W482	N483	A486	E487	D488	Q491	R492	V493	S498	S501	Q502	I503	T504	V505	T506	N507	A508	D511	P512	T513	S514	A515	S520	L523	D524	F525	T526	Q527	L528	D529	W530	I533	I534	Y535	V536	T539	S540	V541	H542	N543	V544	A545	N546	L547				
N386	V387	T388	Q389	S390	K391	C392	Q393	V394	L395	L396	W397	G401	N408	T409	E410	V418	A419	L420	G421	D422	R423	V424	L425	M430	Y434	S435	D436	V437	L438	E439	W440	T441	R442	L443	M444	A445	T446	I447	A452	S457	V460	T461	F462	T465	A374	S466	C467	E470	L471			
G299	A300	A301	N302	P303	R304	K305	L306	V312	N237	L316	I320	A326	Q327	T328	K329	F330	T333	G338	R250	S339	B340	S344	F347	I348	E352	T353	S354	T355	P356	V357	F358	D359	V360	D361	I362	L363	A364	Q365	I366	E367	N368	C369	T370	I371	L372	E373	Y293	R294	D295	W296	P297	V298
N224	Y225	C226	F229	N230	Q234	K235	I236	N237	T238	F239	A240	L241	P242	K243	Y244	F245	K246	S247	L248	L249	R250	M251	S255	F259	S262	V265	T266	G267	Q268	M269	S275	G276	Y277	Y278	R279	Y280	S281	A282	T283	S287	G288	T289	S290	L291	V292	Y293	R294	D295	W296	P297	V298	

4 Data and refinement statistics

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α , β , γ	407.42Å 407.42Å 808.63Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	49.98 – 3.40 78.82 – 3.40	Depositor EDS
% Data completeness (in resolution range)	70.0 (49.98-3.40) 43.5 (78.82-3.40)	Depositor EDS
R_{merge}	0.24	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.32 (at 3.41Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, R_{free}	0.273 , 0.272 (Not available) , (Not available)	Depositor DCC
R_{free} test set	NotAvailable	DCC
Wilson B-factor (Å ²)	47.8	Xtriage
Anisotropy	0.386	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.39 , -69.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.39$, $\langle L^2 \rangle = 0.21$	Xtriage
Estimated twinning fraction	0.087 for -h,-k,l	Xtriage
F_o, F_c correlation	0.41	EDS
Total number of atoms	8286	wwPDB-VP
Average B, all atoms (Å ²)	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.69% 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	0.47	0/4209	0.72	1/5737 (0.0%)
1	B	0.47	0/4209	0.73	2/5737 (0.0%)
All	All	0.47	0/8418	0.72	3/11474 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	443	LEU	N-CA-C	-5.40	96.41	111.00
1	B	443	LEU	N-CA-C	-5.32	96.64	111.00
1	B	296	TRP	C-N-CD	5.13	139.17	128.40

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	4117	0	4038	281	0
1	B	4117	0	4038	302	0
2	A	22	0	0	2	0
2	B	30	0	0	0	0
All	All	8286	0	8076	543	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 33.

The worst 5 of 543 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:72:PRO:HG2	1:B:76:VAL:HG11	1.34	1.05
1:B:419:ALA:HB1	1:B:447:ILE:HD13	1.40	1.03
1:A:187:ARG:HH11	1:A:187:ARG:HG2	1.26	0.99
1:A:442:ARG:N	1:A:442:ARG:HH11	1.62	0.97
1:A:381:THR:HG22	1:A:402:THR:H	1.30	0.96

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	523/525 (100%)	408 (78%)	84 (16%)	31 (6%)	2	17
1	B	523/525 (100%)	414 (79%)	74 (14%)	35 (7%)	1	14
All	All	1046/1050 (100%)	822 (79%)	158 (15%)	66 (6%)	1	16

5 of 66 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	197	PHE
1	A	202	LEU
1	A	203	PRO
1	A	205	GLN
1	A	326	ALA

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	449/449 (100%)	410 (91%)	39 (9%)	12	43
1	B	449/449 (100%)	413 (92%)	36 (8%)	14	48
All	All	898/898 (100%)	823 (92%)	75 (8%)	13	45

5 of 75 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	510	ASP
1	B	140	ASP
1	B	478	PHE
1	A	526	THR
1	A	564	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	148	GLN
1	B	176	GLN
1	B	389	GLN
1	A	483	ASN
1	B	398	GLN

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 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	525/525 (100%)	0.40	19 (3%) 43 39	16, 31, 60, 85	0
1	B	525/525 (100%)	0.45	23 (4%) 35 32	14, 31, 62, 103	0
All	All	1050/1050 (100%)	0.43	42 (4%) 39 35	14, 31, 62, 103	0

The worst 5 of 42 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	407	ASP	5.4
1	A	299	GLY	4.4
1	A	287	SER	4.3
1	B	107	SER	3.6
1	B	105	LEU	3.5

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.