wwPDB X-ray Structure Validation Summary Report (i)

Oct 23, 2021 – 11:39 AM EDT

PDB ID	:	1IJS
Title	:	CPV (STRAIN D) mutant A300D, complex (VIRAL COAT/DNA), VP2, PH=7.5, T=4 DEGREES C
Authors	:	Llamas-Saiz, A.L.; Agbandje-McKenna, M.; Parker, J.S.L.; Wahid, A.T.M.; Parrish, C.R.; Rossmann, M.G.
Deposited on	:	1996-09-12
Resolution	:	3.25 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 following versions of software and data (see references (1)) were used in the production of this report:

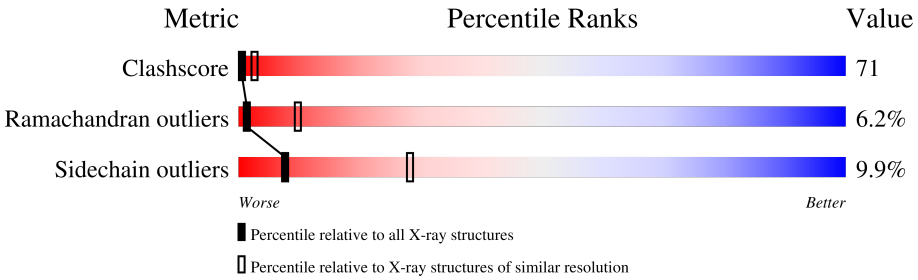
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

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 3.25 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	1251 (3.30-3.22)
Ramachandran outliers	138981	1229 (3.30-3.22)
Sidechain outliers	138945	1228 (3.30-3.22)

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	N	9	33% 67%
2	A	2	100%
3	P	584	29% 51% 11% • 6%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 4571 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 DNA chain called DNA (5'-D(*CP*CP*AP*CP*CP*CP*CP*AP*A)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	N	9	Total	C	N	O	P	0	0	0
			178	84	33	52	9			

- Molecule 2 is a DNA chain called DNA (5'-D(*AP*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	2	Total	C	N	O	P	0	0	0
			40	19	8	11	2			

- Molecule 3 is a protein called PROTEIN (PARVOVIRUS COAT PROTEIN).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	P	548	Total	C	N	O	S	0	0	0
			4353	2765	742	830	16			

There are 2 discrepancies between the modelled and reference sequences:

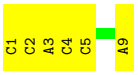
Chain	Residue	Modelled	Actual	Comment	Reference
P	300	ASP	ALA	engineered mutation	UNP P30129
P	386	GLN	LYS	conflict	UNP P30129

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.

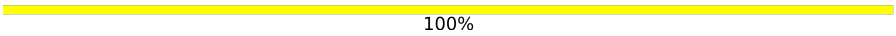
• Molecule 1: DNA (5'-D(*CP*CP*AP*CP*CP*CP*CP*AP*A)-3')

Chain N:



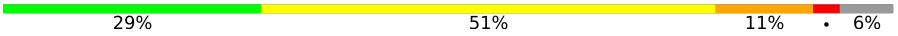
• Molecule 2: DNA (5'-D(*AP*C)-3')

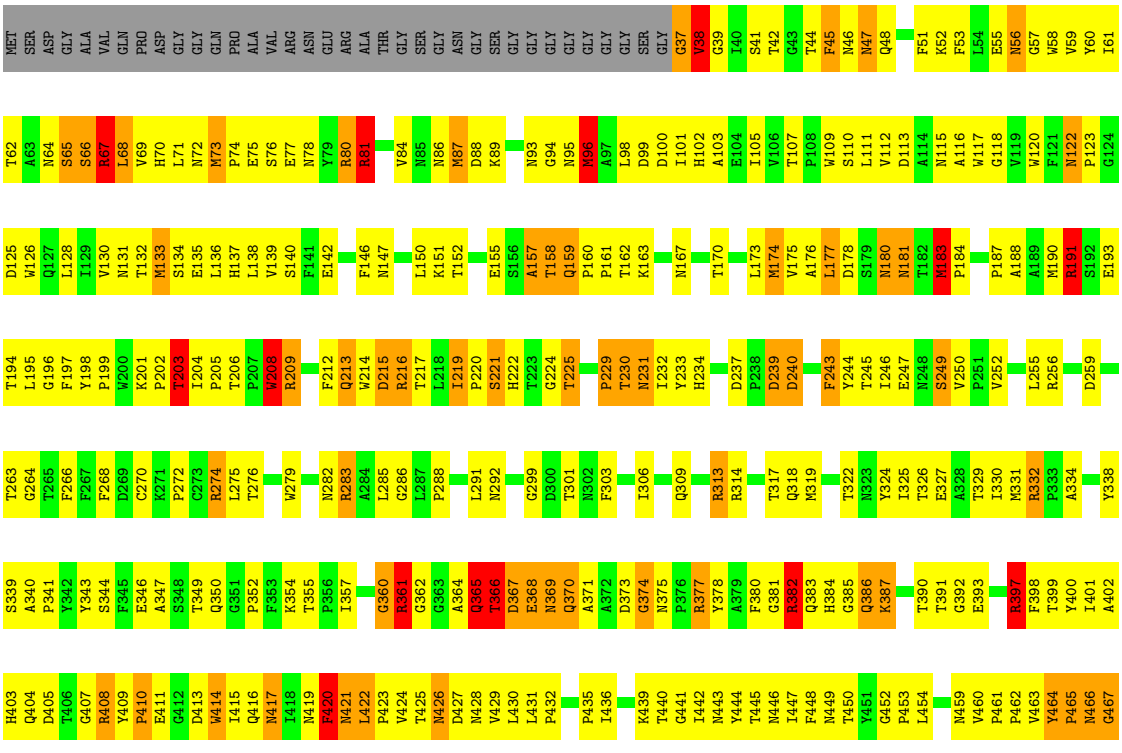
Chain A:

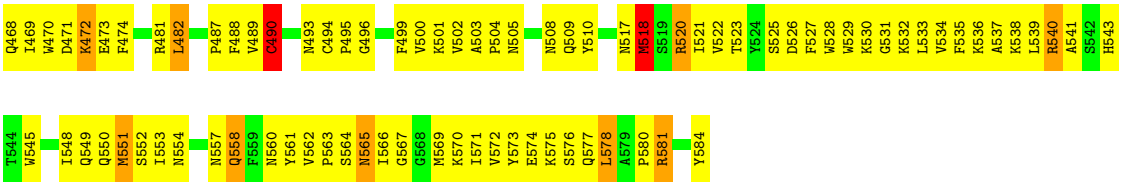


• Molecule 3: PROTEIN (PARVOVIRUS COAT PROTEIN)

Chain P:






Q468
I469
W470
D471
K472
E473
F474
R481
L482
P487
F488
V489
C490
M493
C494
P495
G496
F499
V500
K501
V502
A503
P504
N505
N508
Q509
Y510
N517
M518
S519
R520
I521
V522
T523
Y524
S525
D526
F527
W528
W529
K530
G531
K532
L533
V534
F535
K536
A537
K538
L539
R540
A541
S542
H543
T544
W545
I548
Q549
Q550
M551
S552
I553
N554
N557
Q558
F559
N560
Y561
V562
P563
S564
N565
I566
G567
G568
M569
K570
I571
V572
Y573
E574
K575
S576
Q577
L578
A579
P580
R581
Y584

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	267.60Å 268.50Å 274.30Å 61.90° 62.60° 60.20°	Depositor
Resolution (Å)	(Not available) – 3.25 15.01 – 3.25	Depositor EDS
% Data completeness (in resolution range)	(Not available) ((Not available)-3.25) 32.5 (15.01-3.25)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	11.97 (at 3.25Å)	Xtriage
Refinement program	NONE	Depositor
R, R_{free}	(Not available) , (Not available) 0.529 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å^2)	32.4	Xtriage
Anisotropy	0.329	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.27 , 50.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	0.014 for -h+k,-h,-h+l 0.014 for -k,h-k,-k+l 0.028 for h,h-k,h-l 0.035 for -h+k,k,k-l 0.018 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.06	EDS
Total number of atoms	4571	wwPDB-VP
Average B, all atoms (Å^2)	11.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.88% 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	N	1.29	1/198 (0.5%)	1.10	0/299
2	A	1.07	0/44	1.07	0/65
3	P	1.11	10/4483 (0.2%)	1.25	37/6133 (0.6%)
All	All	1.12	11/4725 (0.2%)	1.25	37/6497 (0.6%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
3	P	0	1

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	P	81	ARG	CD-NE	12.24	1.67	1.46
3	P	490	CYS	N-CA	-10.05	1.26	1.46
3	P	490	CYS	CA-C	9.46	1.77	1.52
3	P	239	ASP	C-N	-8.09	1.15	1.34
1	N	1	DC	C2'-C1'	-7.39	1.44	1.52

The worst 5 of 37 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	P	81	ARG	CG-CD-NE	12.11	137.24	111.80
3	P	490	CYS	N-CA-CB	11.21	130.77	110.60
3	P	81	ARG	NE-CZ-NH2	8.14	124.37	120.30
3	P	81	ARG	CA-CB-CG	7.98	130.96	113.40
3	P	314	ARG	NE-CZ-NH2	7.73	124.17	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	P	81	ARG	Sidechain

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	N	178	0	99	7	0
2	A	40	0	23	12	0
3	P	4353	0	4142	605	0
All	All	4571	0	4264	624	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 71.

The worst 5 of 624 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)
3:P:96:MET:CE	3:P:221:SER:H	1.02	1.57
3:P:424:VAL:HG21	3:P:429:VAL:CG2	1.35	1.51
3:P:490:CYS:C	3:P:490:CYS:CA	1.77	1.49
2:A:1:DA:C3'	2:A:2:DC:H5"	1.48	1.44
3:P:276:THR:CG2	3:P:581:ARG:HB2	1.48	1.44

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
3	P	546/584 (94%)	442 (81%)	70 (13%)	34 (6%)	1 10

5 of 34 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	P	38	VAL
3	P	159	GLN
3	P	369	ASN
3	P	370	GLN
3	P	386	GLN

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
3	P	477/496 (96%)	430 (90%)	47 (10%)	8 28

5 of 47 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	P	249	SER
3	P	382	ARG
3	P	349	THR
3	P	366	THR
3	P	393	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 26 such sidechains are listed below:

Mol	Chain	Res	Type
3	P	370	GLN
3	P	403	HIS
3	P	557	ASN
3	P	386	GLN
3	P	428	ASN

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 monosaccharides 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

The following chains have linkage breaks:

Mol	Chain	Number of breaks
3	P	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	P	239:ASP	C	240:ASP	N	1.15

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.