W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 9, 2023 – 06:22 PM EDT

PDB ID	:	5CFD
Title	:	Crystal Structure of DTT treated Human Cardiovirus SAFV-3
Authors	:	Mullapudi, E.; Plevka, P.
Deposited on	:	2015-07-08
Resolution	:	2.50 Å(reported)

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

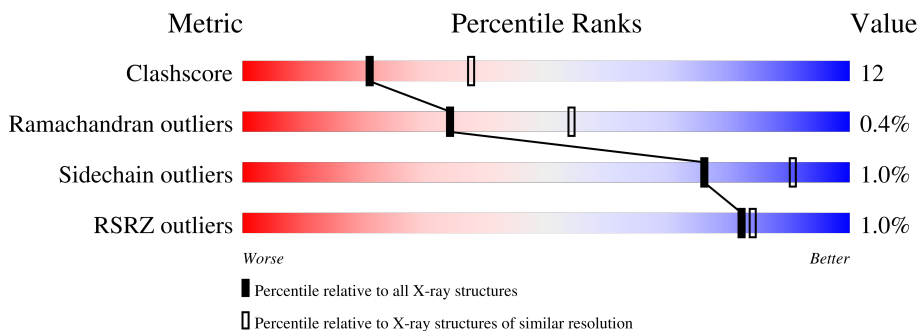|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation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

1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.50 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
Clashscore		12
Ramachandran outliers		0.4%
Sidechain outliers		1.0%
RSRZ outliers		1.0%

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))
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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% 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	252	79% 21%
2	B	232	78% 21% •
3	C	258	3% 74% 25% ••
4	D	24	71% 17% 8% •

WORLDWIDE
wwPDB
PROTEIN DATA BANK

2 Entry composition ⓘ

There are 5 unique types of molecules in this entry. The entry contains 6350 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 VP1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	252	Total 1996	C 1273	N 332	O 384	S 7	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	36	PHE	VAL	conflict	UNP C0MHL9

- Molecule 2 is a protein called VP3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	232	Total 1804	C 1166	N 285	O 343	S 10	0	0	0

- Molecule 3 is a protein called VP2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	258	Total 2023	C 1274	N 364	O 378	S 7	0	0	0

- Molecule 4 is a protein called VP4.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	D	24	Total 189	C 115	N 32	O 42	0	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	124	Total 124	O 124	0	0

Continued on next page...

Continued from previous page...

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	99	Total 99	O 99	0	0
5	C	107	Total 107	O 107	0	0
5	D	8	Total 8	O 8	0	0

3 Residue-property plots ⓘ

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 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.
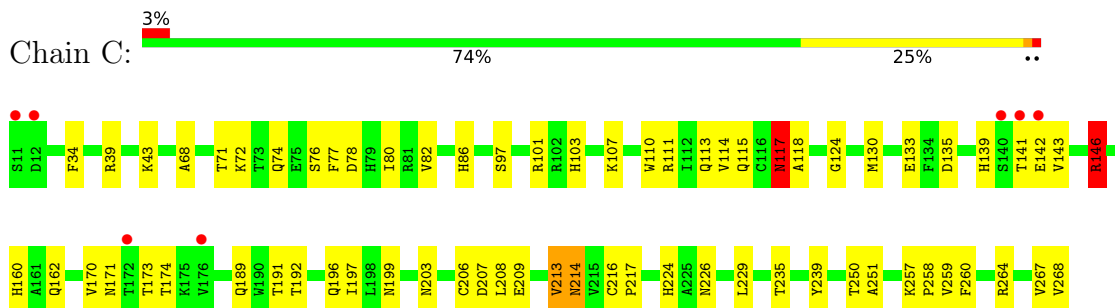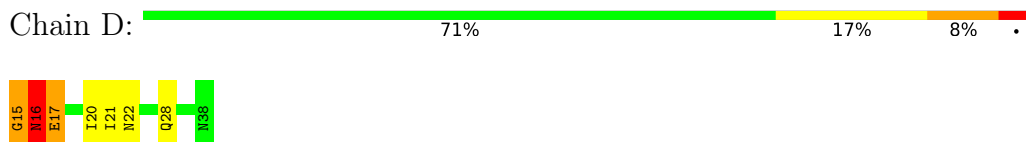
- Molecule 1: VP1

Chain A:

- Molecule 2: VP3

Chain B:

- Molecule 3: VP2

Chain C:

- Molecule 4: VP4

Chain D:
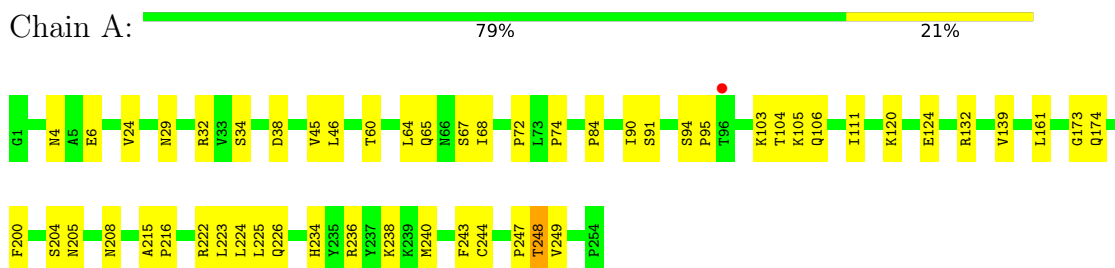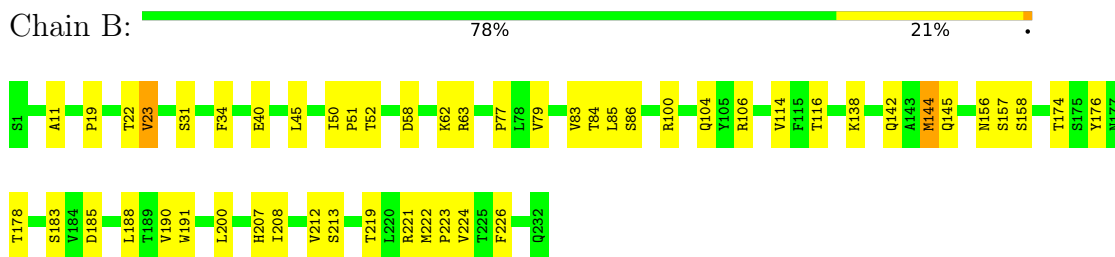
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α, β, γ	299.95Å 299.95Å 723.47Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	70.00 – 2.50 69.03 – 2.50	Depositor EDS
% Data completeness (in resolution range)	67.4 (70.00-2.50) 67.5 (69.03-2.50)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.35 (at 2.51Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.215 , (Not available) 0.212 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å^2)	28.8	Xtriage
Anisotropy	0.097	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.68 , 35.6	EDS
L-test for twinning[2]	$< |L| > = 0.58, < L^2 > = 0.43$	Xtriage
Estimated twinning fraction	0.023 for -h,-k,l	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	6350	wwPDB-VP
Average B, all atoms (Å^2)	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.23% 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	0.76	13/2055 (0.6%)	0.81	8/2808 (0.3%)
2	B	1.03	17/1860 (0.9%)	0.82	5/2553 (0.2%)
3	C	1.17	25/2078 (1.2%)	0.93	10/2843 (0.4%)
4	D	2.26	11/191 (5.8%)	1.32	3/259 (1.2%)
All	All	1.06	66/6184 (1.1%)	0.88	26/8463 (0.3%)

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	C	0	1

The worst 5 of 66 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	259	VAL	C-O	-14.24	0.96	1.23
3	C	203	ASN	C-O	-13.30	0.98	1.23
2	B	23	VAL	C-O	-13.24	0.98	1.23
4	D	28	GLN	C-O	-13.16	0.98	1.23
2	B	212	VAL	C-O	-12.94	0.98	1.23

The worst 5 of 26 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	C	146	ARG	NE-CZ-NH1	16.24	128.42	120.30
3	C	111	ARG	NE-CZ-NH1	11.34	125.97	120.30
1	A	132	ARG	CG-CD-NE	-10.95	88.80	111.80
1	A	132	ARG	NE-CZ-NH1	10.56	125.58	120.30
3	C	259	VAL	O-C-N	8.52	136.33	122.70

There are no chirality outliers.

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	C	117	ASN	Mainchain

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	1996	0	1937	48	0
2	B	1804	0	1760	54	0
3	C	2023	0	1960	63	0
4	D	189	0	166	7	0
5	A	124	0	0	3	0
5	B	99	0	0	2	0
5	C	107	0	0	2	0
5	D	8	0	0	0	0
All	All	6350	0	5823	144	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 12.

The worst 5 of 144 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:A:46:LEU:HD13	1:A:68:ILE:HG23	1.46	0.97
2:B:138:LYS:H	2:B:142:GLN:NE2	1.68	0.90
1:A:104:THR:HG23	1:A:106:GLN:OE1	1.74	0.87
1:A:238:LYS:NZ	4:D:17:GLU:OE2	2.09	0.81
2:B:104:GLN:HE22	2:B:221:ARG:HH21	1.28	0.81

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	250/252 (99%)	238 (95%)	11 (4%)	1 (0%)	34	54
2	B	230/232 (99%)	220 (96%)	9 (4%)	1 (0%)	34	54
3	C	256/258 (99%)	247 (96%)	9 (4%)	0	100	100
4	D	22/24 (92%)	20 (91%)	1 (4%)	1 (4%)	2	3
All	All	758/766 (99%)	725 (96%)	30 (4%)	3 (0%)	34	54

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	D	16	ASN
1	A	244	CYS
2	B	223	PRO

5.3.2 Protein sidechains ⓘ

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	233/233 (100%)	232 (100%)	1 (0%)	91	97
2	B	206/206 (100%)	203 (98%)	3 (2%)	65	85
3	C	222/222 (100%)	219 (99%)	3 (1%)	67	86
4	D	21/21 (100%)	21 (100%)	0	100	100
All	All	682/682 (100%)	675 (99%)	7 (1%)	76	90

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	178	THR
3	C	114	VAL
3	C	146	ARG
3	C	117	ASN
2	B	144	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 28 such sidechains are listed below:

Mol	Chain	Res	Type
3	C	38	GLN
4	D	30	GLN
3	C	113	GLN
3	C	224	HIS
3	C	86	HIS

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
4	D	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	D	17:GLU	C	18:GLY	N	1.19

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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(Å²)	Q<0.9
1	A	252/252 (100%)	-0.56	1 (0%)	92	93	19, 25, 41, 51	0
2	B	232/232 (100%)	-0.67	0	100	100	18, 24, 40, 55	0
3	C	258/258 (100%)	-0.55	7 (2%)	54	58	19, 24, 49, 72	0
4	D	24/24 (100%)	-0.12	0	100	100	23, 40, 62, 64	0
All	All	766/766 (100%)	-0.58	8 (1%)	82	84	18, 24, 46, 72	0

The worst 5 of 8 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	C	141	THR	3.5
3	C	140	SER	3.3
1	A	96	THR	3.3
3	C	172	THR	2.8
3	C	176	VAL	2.6

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 monosaccharides 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.