



wwPDB X-ray Structure Validation Summary Report

Aug 15, 2022 – 04:27 pm BST

PDB ID : 7ZAL
Title : FNIP family proteins from Cafeteria roenbergensis virus (CroV): leucine-rich repeats with novel structural features
Authors : Huyton, T.; Goerlich, D.
Deposited on : 2022-03-22
Resolution : 2.73 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  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](#) ) were used in the production of this report:

MolProbity : 4.02b-467
Xtriage (Phenix) : 1.13
EDS : 2.29
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0267
CCP4 : 7.1.010 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.29

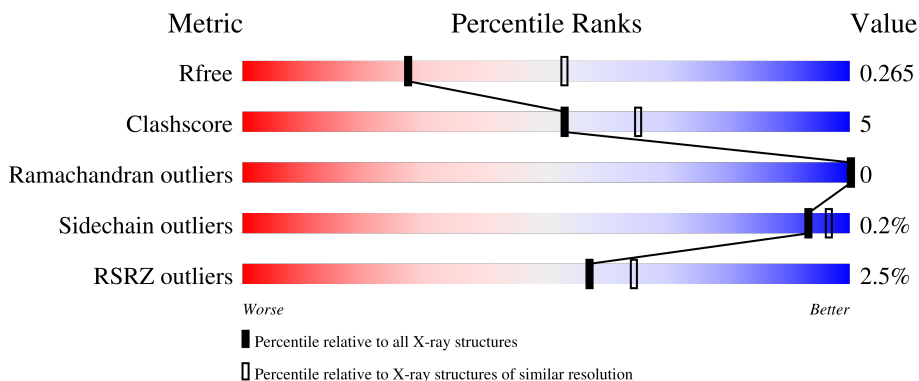
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.73 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.



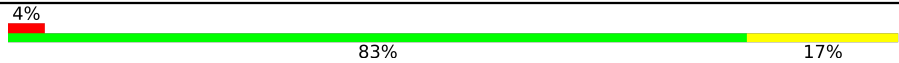
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1271 (2.76-2.72)
Clashscore	141614	1322 (2.76-2.72)
Ramachandran outliers	138981	1297 (2.76-2.72)
Sidechain outliers	138945	1298 (2.76-2.72)
RSRZ outliers	127900	1243 (2.76-2.72)

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 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 $\leq 5\%$. 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	AAA	292	
1	BBB	292	
1	CCC	292	
1	DDD	292	
1	EEE	292	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	FFF	292	 <p>A horizontal bar chart representing the quality of chain. The bar is divided into three segments: a small red segment on the left labeled '4%', a large green segment in the middle labeled '83%', and a yellow segment on the right labeled '17%'.</p>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 14800 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 Crov539.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	AAA	292	2416	1591	383	439	3	0	0	0
1	BBB	292	2416	1591	383	439	3	0	0	0
1	CCC	292	2416	1591	383	439	3	0	0	0
1	DDD	292	2416	1591	383	439	3	0	0	0
1	EEE	292	2416	1591	383	439	3	0	0	0
1	FFF	292	2416	1591	383	439	3	0	0	0

There are 12 discrepancies between the modelled and reference sequences:

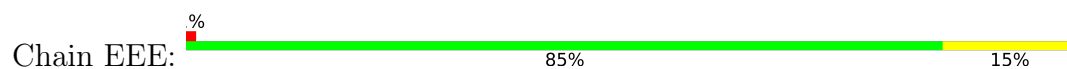
Chain	Residue	Modelled	Actual	Comment	Reference
AAA	-1	GLY	-	expression tag	UNP E3T5W0
AAA	0	SER	-	expression tag	UNP E3T5W0
BBB	-1	GLY	-	expression tag	UNP E3T5W0
BBB	0	SER	-	expression tag	UNP E3T5W0
CCC	-1	GLY	-	expression tag	UNP E3T5W0
CCC	0	SER	-	expression tag	UNP E3T5W0
DDD	-1	GLY	-	expression tag	UNP E3T5W0
DDD	0	SER	-	expression tag	UNP E3T5W0
EEE	-1	GLY	-	expression tag	UNP E3T5W0
EEE	0	SER	-	expression tag	UNP E3T5W0
FFF	-1	GLY	-	expression tag	UNP E3T5W0
FFF	0	SER	-	expression tag	UNP E3T5W0

- Molecule 2 is water.

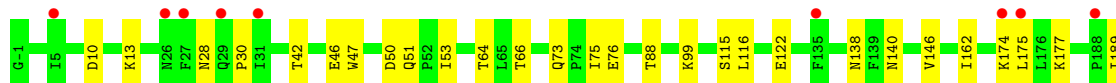
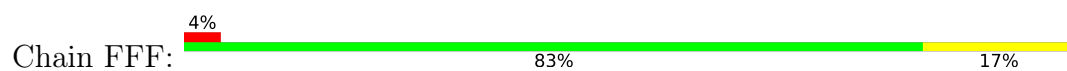
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	AAA	65	Total 65	O 65	0	0
2	BBB	43	Total 43	O 43	0	0
2	CCC	89	Total 89	O 89	0	0
2	DDD	37	Total 37	O 37	0	0
2	EEE	37	Total 37	O 37	0	0
2	FFF	33	Total 33	O 33	0	0



- Molecule 1: Crov539



- Molecule 1: Crov539



4 Data and refinement statistics

Property	Value	Source
Space group	P 2 21 21	Depositor
Cell constants a, b, c, α , β , γ	101.01Å 113.45Å 236.24Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.44 – 2.73 49.39 – 2.73	Depositor EDS
% Data completeness (in resolution range)	99.7 (49.44-2.73) 99.7 (49.39-2.73)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.61 (at 2.73Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.220 , 0.267 0.218 , 0.265	Depositor DCC
R_{free} test set	3577 reflections (4.93%)	wwPDB-VP
Wilson B-factor (Å ²)	72.2	Xtriage
Anisotropy	0.674	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	(Not available) , (Not available)	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	14800	wwPDB-VP
Average B, all atoms (Å ²)	95.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.68% 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	AAA	0.58	0/2489	0.63	0/3397
1	BBB	0.57	0/2489	0.59	0/3397
1	CCC	0.58	0/2489	0.61	0/3397
1	DDD	0.58	0/2489	0.59	0/3397
1	EEE	0.58	0/2489	0.59	0/3397
1	FFF	0.58	0/2489	0.57	0/3397
All	All	0.58	0/14934	0.60	0/20382

There are no bond length outliers.

There are no bond angle outliers.

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	AAA	2416	0	2424	20	0
1	BBB	2416	0	2424	35	0
1	CCC	2416	0	2424	17	0
1	DDD	2416	0	2424	36	1
1	EEE	2416	0	2424	27	0
1	FFF	2416	0	2424	27	0
2	AAA	65	0	0	3	0
2	BBB	43	0	0	1	0
2	CCC	89	0	0	1	0
2	DDD	37	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	EEE	37	0	0	3	0
2	FFF	33	0	0	2	0
All	All	14800	0	14544	157	1

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 5.

The worst 5 of 157 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:EEE:249:TYR:HE2	1:EEE:253:LEU:HD21	1.42	0.84
1:AAA:158:ASN:ND2	2:AAA:301:HOH:O	2.11	0.81
1:AAA:205:ASP:HA	1:AAA:227:PRO:HD2	1.66	0.76
1:EEE:205:ASP:HA	1:EEE:227:PRO:HD2	1.66	0.75
1:FFF:205:ASP:HA	1:FFF:227:PRO:HD2	1.67	0.75

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:DDD:26:ASN:OD1	1:DDD:234:ASN:ND2[3_645]	1.85	0.35

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	AAA	290/292 (99%)	274 (94%)	16 (6%)	0	100	100
1	BBB	290/292 (99%)	278 (96%)	12 (4%)	0	100	100
1	CCC	290/292 (99%)	274 (94%)	16 (6%)	0	100	100
1	DDD	290/292 (99%)	269 (93%)	21 (7%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	EEE	290/292 (99%)	273 (94%)	17 (6%)	0	100	100
1	FFF	290/292 (99%)	280 (97%)	10 (3%)	0	100	100
All	All	1740/1752 (99%)	1648 (95%)	92 (5%)	0	100	100

There are no Ramachandran outliers to report.

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	AAA	281/281 (100%)	280 (100%)	1 (0%)	91	94
1	BBB	281/281 (100%)	281 (100%)	0	100	100
1	CCC	281/281 (100%)	280 (100%)	1 (0%)	91	94
1	DDD	281/281 (100%)	280 (100%)	1 (0%)	91	94
1	EEE	281/281 (100%)	281 (100%)	0	100	100
1	FFF	281/281 (100%)	281 (100%)	0	100	100
All	All	1686/1686 (100%)	1683 (100%)	3 (0%)	93	96

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	86	PHE
1	CCC	86	PHE
1	DDD	86	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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 [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	AAA	292/292 (100%)	0.40	2 (0%) 87 90	51, 71, 120, 176	0
1	BBB	292/292 (100%)	0.33	8 (2%) 54 61	71, 98, 127, 139	0
1	CCC	292/292 (100%)	0.30	1 (0%) 94 96	49, 71, 102, 141	0
1	DDD	292/292 (100%)	0.51	16 (5%) 25 28	68, 97, 147, 183	0
1	EEE	292/292 (100%)	0.28	4 (1%) 75 80	71, 100, 133, 151	0
1	FFF	292/292 (100%)	0.49	13 (4%) 33 36	78, 120, 146, 162	0
All	All	1752/1752 (100%)	0.38	44 (2%) 57 64	49, 92, 137, 183	0

The worst 5 of 44 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	BBB	-1	GLY	4.4
1	FFF	27	PHE	3.9
1	DDD	265	ILE	3.5
1	DDD	218	LYS	3.5
1	BBB	19	HIS	3.3

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 monosaccharides 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.