



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 08:29 AM GMT

PDB ID : 3ETI
Title : Structure of a cubic crystal form of X (ADRP) domain from FCoV
Authors : Wojdyla, J.A.; Manolaridis, I.; Tucker, P.A.
Deposited on : 2008-10-08
Resolution : 2.20 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
Mogul : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
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) : trunk26865

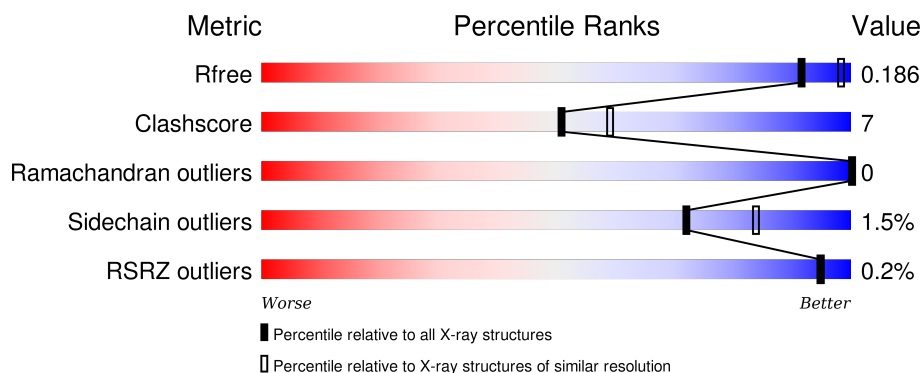
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.20 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}	91344	3774 (2.20-2.20)
Clashscore	102246	4477 (2.20-2.20)
Ramachandran outliers	100387	4404 (2.20-2.20)
Sidechain outliers	100360	4405 (2.20-2.20)
RSRZ outliers	91569	3781 (2.20-2.20)

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	168	<div> <div>89%</div> <div>11%</div> </div>
1	B	168	<div> <div>85%</div> <div>15%</div> </div>
1	C	168	<div> <div>87%</div> <div>13%</div> </div>
1	D	168	<div> <div>86%</div> <div>13%</div> </div>
1	E	168	<div> <div>88%</div> <div>13%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	168	<div><div>%</div><div><div></div><div>88%</div><div>11%</div></div><div></div></div>
1	G	168	<div><div>%</div><div><div></div><div>83%</div><div>17%</div></div><div></div></div>
1	H	168	<div><div></div><div><div></div><div>88%</div><div>11%</div></div><div></div></div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 13011 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 macro domain of Non-structural protein 3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	168	Total	C	N	O	S	0	8	0
			1348	866	231	247	4			
1	B	168	Total	C	N	O	S	0	5	0
			1326	852	224	246	4			
1	C	168	Total	C	N	O	S	0	5	0
			1329	852	229	244	4			
1	D	168	Total	C	N	O	S	0	5	0
			1331	854	228	245	4			
1	E	168	Total	C	N	O	S	0	4	0
			1322	848	225	245	4			
1	F	168	Total	C	N	O	S	0	6	0
			1338	859	229	246	4			
1	G	168	Total	C	N	O	S	0	6	0
			1334	857	228	245	4			
1	H	168	Total	C	N	O	S	0	7	0
			1340	860	227	249	4			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	122	MET	LEU	ENGINEERED	UNP Q98VG9
B	122	MET	LEU	ENGINEERED	UNP Q98VG9
C	122	MET	LEU	ENGINEERED	UNP Q98VG9
D	122	MET	LEU	ENGINEERED	UNP Q98VG9
E	122	MET	LEU	ENGINEERED	UNP Q98VG9
F	122	MET	LEU	ENGINEERED	UNP Q98VG9
G	122	MET	LEU	ENGINEERED	UNP Q98VG9
H	122	MET	LEU	ENGINEERED	UNP Q98VG9


- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	331	Total 331	O 331	0	0
2	B	305	Total 305	O 305	0	0
2	C	293	Total 293	O 293	0	0
2	D	296	Total 296	O 296	0	0
2	E	312	Total 312	O 312	0	0
2	F	231	Total 231	O 231	0	0
2	G	260	Total 260	O 260	0	0
2	H	315	Total 315	O 315	0	0

3 Residue-property plots [i](#)


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: macro domain of Non-structural protein 3

Chain A: 




- Molecule 1: macro domain of Non-structural protein 3

Chain B: 




- Molecule 1: macro domain of Non-structural protein 3

Chain C: 




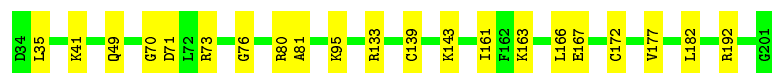
- Molecule 1: macro domain of Non-structural protein 3

Chain D: 




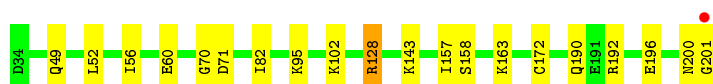
- Molecule 1: macro domain of Non-structural protein 3

Chain E: 

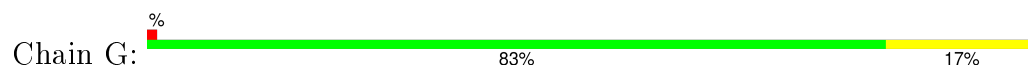


- Molecule 1: macro domain of Non-structural protein 3

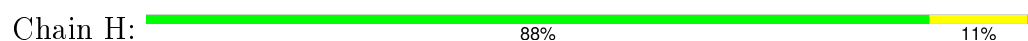
Chain F: 



- Molecule 1: macro domain of Non-structural protein 3



- Molecule 1: macro domain of Non-structural protein 3



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 3	Depositor
Cell constants a, b, c, α , β , γ	218.99Å 218.99Å 218.99Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	24.80 – 2.20 24.80 – 2.20	Depositor EDS
% Data completeness (in resolution range)	100.0 (24.80-2.20) 100.0 (24.80-2.20)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.16 (at 2.19Å)	Xtriage
Refinement program	REFMAC 5.4.0069	Depositor
R, R_{free}	0.147 , 0.181 0.151 , 0.186	Depositor DCC
R_{free} test set	8795 reflections (5.26%)	DCC
Wilson B-factor (Å ²)	25.1	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 70.4	EDS
Estimated twinning fraction	0.095 for l,-k,h	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.43$, $\langle L^2 \rangle = 0.25$	Xtriage
Outliers	1 of 175890 reflections (0.001%)	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	13011	wwPDB-VP
Average B, all atoms (Å ²)	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.64% 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, 0.375 respectively for untwinned datasets, and 0.333, 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	0.89	0/1392	0.81	1/1878 (0.1%)
1	B	0.89	1/1361 (0.1%)	0.81	2/1838 (0.1%)
1	C	0.92	0/1364	0.86	2/1842 (0.1%)
1	D	0.94	0/1363	0.83	1/1840 (0.1%)
1	E	0.91	1/1354 (0.1%)	0.84	1/1829 (0.1%)
1	F	0.84	0/1373	0.80	1/1853 (0.1%)
1	G	0.86	0/1372	0.79	0/1852
1	H	0.93	0/1378	0.79	3/1861 (0.2%)
All	All	0.90	2/10957 (0.0%)	0.82	11/14793 (0.1%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	139	CYS	CB-SG	6.03	1.92	1.82
1	B	139	CYS	CB-SG	5.17	1.91	1.82

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	192	ARG	NE-CZ-NH2	-12.44	114.08	120.30
1	C	192	ARG	NE-CZ-NH1	10.55	125.58	120.30
1	B	128	ARG	NE-CZ-NH1	6.95	123.77	120.30
1	H	80	ARG	NE-CZ-NH1	-6.74	116.93	120.30
1	H	73	ARG	NE-CZ-NH1	-6.14	117.23	120.30

There are no chirality outliers.

There are no planarity outliers.

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	A	1348	0	1407	16	1
1	B	1326	0	1373	18	0
1	C	1329	0	1378	22	0
1	D	1331	0	1379	19	0
1	E	1322	0	1365	22	0
1	F	1338	0	1389	17	0
1	G	1334	0	1388	21	0
1	H	1340	0	1386	23	3
2	A	331	0	0	8	4
2	B	305	0	0	7	0
2	C	293	0	0	6	2
2	D	296	0	0	8	1
2	E	312	0	0	10	2
2	F	231	0	0	9	0
2	G	260	0	0	6	1
2	H	315	0	0	7	3
All	All	13011	0	11065	149	9

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 7.

The worst 5 of 149 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:71[A]:ASP:OD1	1:A:73:ARG:NH1	2.03	0.90
1:A:71[A]:ASP:OD1	1:A:73:ARG:CZ	2.23	0.85
1:F:70:GLY:O	2:F:359:HOH:O	1.95	0.84
1:H:71[A]:ASP:OD1	1:H:73:ARG:CZ	2.28	0.82
1:F:95[B]:LYS:NZ	2:F:360:HOH:O	2.11	0.82

The worst 5 of 9 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:H:80:ARG:NH1	2:H:489:HOH:O[7_555]	1.58	0.62
2:E:473:HOH:O	2:G:385:HOH:O[7_555]	1.62	0.58
1:H:201:GLY:C	2:A:437:HOH:O[11_456]	1.85	0.35
2:D:427:HOH:O	2:D:495:HOH:O[9_555]	1.94	0.26

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:A:448:HOH:O	2:C:432:HOH:O[5_555]	1.99	0.21

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	174/168 (104%)	172 (99%)	2 (1%)	0	100	100
1	B	171/168 (102%)	166 (97%)	5 (3%)	0	100	100
1	C	171/168 (102%)	167 (98%)	4 (2%)	0	100	100
1	D	171/168 (102%)	168 (98%)	3 (2%)	0	100	100
1	E	170/168 (101%)	166 (98%)	4 (2%)	0	100	100
1	F	172/168 (102%)	169 (98%)	3 (2%)	0	100	100
1	G	172/168 (102%)	168 (98%)	4 (2%)	0	100	100
1	H	173/168 (103%)	169 (98%)	4 (2%)	0	100	100
All	All	1374/1344 (102%)	1345 (98%)	29 (2%)	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	A	151/143 (106%)	148 (98%)	3 (2%)	63	76

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	148/143 (104%)	144 (97%)	4 (3%)	52	64
1	C	148/143 (104%)	147 (99%)	1 (1%)	88	94
1	D	148/143 (104%)	144 (97%)	4 (3%)	52	64
1	E	147/143 (103%)	146 (99%)	1 (1%)	88	94
1	F	149/143 (104%)	148 (99%)	1 (1%)	88	94
1	G	149/143 (104%)	147 (99%)	2 (1%)	76	87
1	H	150/143 (105%)	148 (99%)	2 (1%)	76	87
All	All	1190/1144 (104%)	1172 (98%)	18 (2%)	72	84

5 of 18 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	36	ILE
1	D	71	ASP
1	G	143	LYS
1	B	172	CYS
1	C	36	ILE

Some 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 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	168/168 (100%)	-0.59	0 100 100	24, 34, 49, 54	0
1	B	168/168 (100%)	-0.60	1 (0%) 90 90	27, 35, 52, 67	0
1	C	168/168 (100%)	-0.63	0 100 100	23, 34, 50, 57	0
1	D	168/168 (100%)	-0.60	0 100 100	26, 34, 52, 61	0
1	E	168/168 (100%)	-0.61	0 100 100	26, 34, 49, 58	0
1	F	168/168 (100%)	-0.64	1 (0%) 90 90	28, 38, 55, 77	0
1	G	168/168 (100%)	-0.60	1 (0%) 90 90	27, 37, 54, 76	0
1	H	168/168 (100%)	-0.56	0 100 100	25, 34, 49, 56	0
All	All	1344/1344 (100%)	-0.61	3 (0%) 95 95	23, 35, 52, 77	0

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	201	GLY	4.7
1	F	201	GLY	3.4
1	B	201	GLY	2.1

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.