wwPDB X-ray Structure Validation Summary Report (i)

Aug 20, 2023 – 12:33 AM JST

PDB ID	:	8IIC
Title	:	Crystal structure of Israeli acute paralysis virus RNA-dependent RNA polymerase delta40 mutant (residues 41-546)
Authors	:	Fang, X.; Lu, G.; Hou, C.; Gong, P.
Deposited on	:	2023-02-24
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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

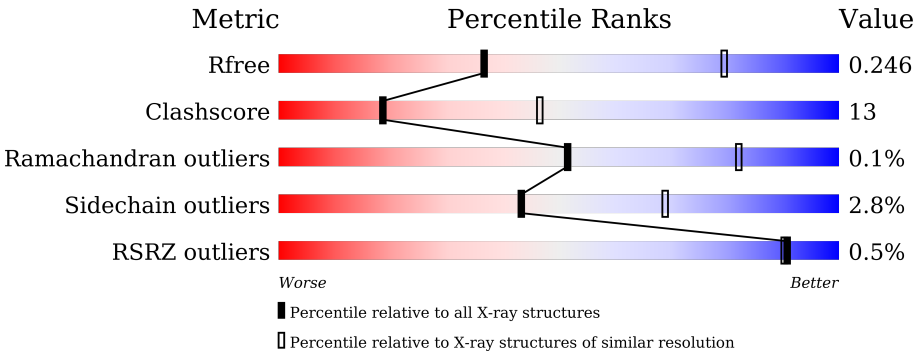
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 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))
R_{free}	130704	1191 (3.30-3.22)
Clashscore	141614	1251 (3.30-3.22)
Ramachandran outliers	138981	1229 (3.30-3.22)
Sidechain outliers	138945	1228 (3.30-3.22)
RSRZ outliers	127900	1154 (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% 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	524	% 55% 18% 27%
1	B	524	59% 21% · 19%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 6070 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 Polymerase polyprotein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	383	Total	C	N	O	S	0	0	0
			2931	1864	492	548	27			
1	B	425	Total	C	N	O	S	0	0	0
			3133	1987	530	586	30			

- Molecule 2 is water.

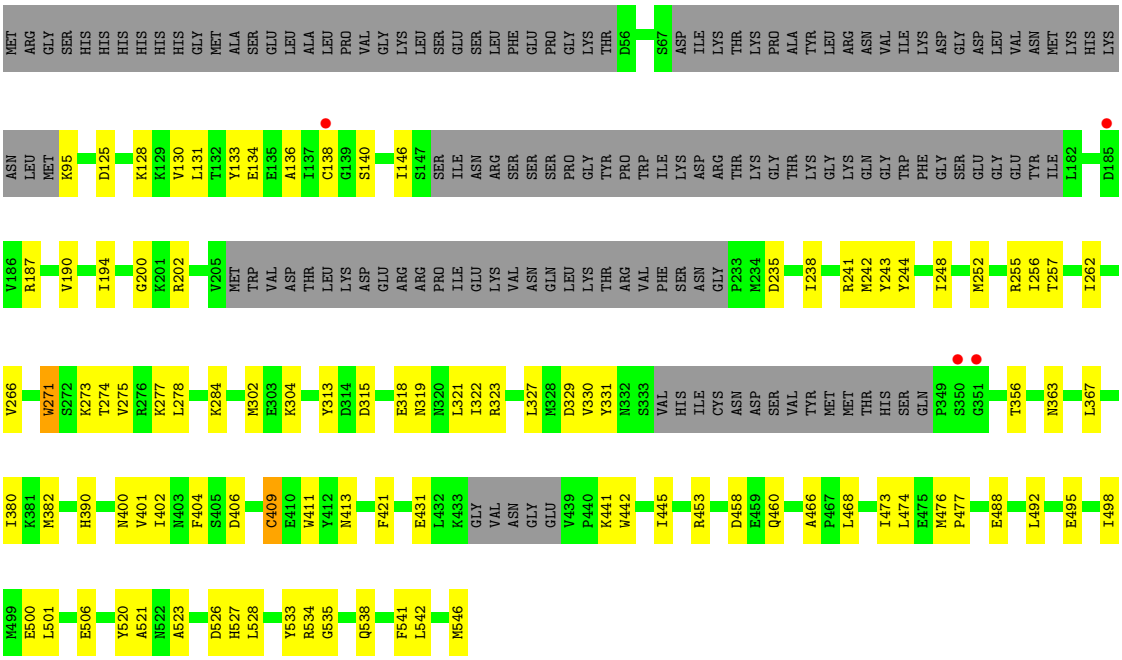
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	4	Total	O	0	0
			4	4		
2	B	2	Total	O	0	0
			2	2		

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 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.

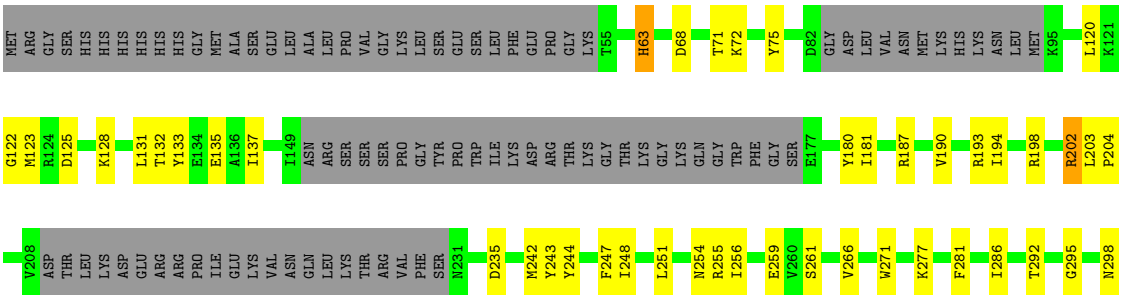
• Molecule 1: Polymerase polyprotein

Chain A:



• Molecule 1: Polymerase polyprotein

Chain B:




A306
Y313
D314
D315
E318
R323
L327
M328
D329
V330
N352
P353
A354
T355
T356
P357
L358
I362
L367
R368
F371
C374
I380
K381
M382
T383
D386
F387
V391
S392
M393
V394
V401
I402
N403
F404
S405
C409
E410
W411
Y412
N413
T416
I417
F421
F426
L432
LYS
GLY
VAL
ASN
GLY
GLU
V439
P440
K441
W442
R443
Y450
R453
D458
K462
E465
A466
P467
L468
C469
M470
D471
T472
I473
L474
E475
N478
E488
L492
E495
I498
M499
E500
L501
S502
T511
I515
Y520
D526
H527
L528
D529
I530
Y536
R540
F541
L542
M546

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 62 2 2	Depositor
Cell constants a, b, c, α, β, γ	167.93Å 167.93Å 218.32Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	23.78 – 3.25 23.78 – 3.25	Depositor EDS
% Data completeness (in resolution range)	95.1 (23.78-3.25) 95.1 (23.78-3.25)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	5.16 (at 3.23Å)	Xtriage
Refinement program	PHENIX 1.17.1	Depositor
R, R_{free}	0.211 , 0.246 0.211 , 0.246	Depositor DCC
R_{free} test set	1401 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å^2)	62.6	Xtriage
Anisotropy	0.085	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 63.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	6070	wwPDB-VP
Average B, all atoms (Å^2)	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.80% 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	0.57	0/2989	0.71	0/4049
1	B	0.46	0/3194	0.67	0/4339
All	All	0.52	0/6183	0.69	0/8388

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

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	2931	0	2666	66	0
1	B	3133	0	2717	85	0
2	A	4	0	0	0	0
2	B	2	0	0	0	0
All	All	6070	0	5383	151	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 13.

The worst 5 of 151 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:440:PRO:HG2	1:B:443:ARG:HH21	1.51	0.74

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:262:ILE:HD11	1:A:356:THR:HG22	1.71	0.73
1:B:520:TYR:HE1	1:B:526:ASP:HB3	1.55	0.72
1:B:382:MET:HE3	1:B:411:TRP:HZ2	1.56	0.70
1:A:244:TYR:CZ	1:A:323:ARG:HG3	2.26	0.70

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	371/524 (71%)	347 (94%)	24 (6%)	0	100 100
1	B	415/524 (79%)	387 (93%)	27 (6%)	1 (0%)	47 77
All	All	786/1048 (75%)	734 (93%)	51 (6%)	1 (0%)	51 82

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	63	HIS

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	283/463 (61%)	274 (97%)	9 (3%)	39 66

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	B	279/463 (60%)	272 (98%)	7 (2%)	47 71
All	All	562/926 (61%)	546 (97%)	16 (3%)	43 69

5 of 16 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	526	ASP
1	B	421	PHE
1	A	541	PHE
1	B	412	TYR
1	A	527	HIS

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

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, 95^{th} 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	383/524 (73%)	-0.39	4 (1%) 82 82	17, 40, 79, 96	0
1	B	425/524 (81%)	-0.30	0 100 100	25, 60, 85, 105	1 (0%)
All	All	808/1048 (77%)	-0.34	4 (0%) 91 90	17, 53, 84, 105	1 (0%)

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	138	CYS	2.6
1	A	350	SER	2.6
1	A	351	GLY	2.3
1	A	185	ASP	2.0

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

There are no such residues in this entry.