



wwPDB X-ray Structure Validation Summary Report ⓘ

May 21, 2020 – 07:01 pm BST

PDB ID : 4J4Y
Title : Triple mutant GraVN
Authors : Jiao, L.; Ouyang, S.; Liang, M.; Niu, F.; Shaw, N.; Wu, W.; Ding, W.; Jin, C.;
Zhu, Y.; Zhang, F.; Wang, T.; Li, C.; Zuo, X.; Luan, C.H.; Li, D.; Liu, Z.J.
Deposited on : 2013-02-07
Resolution : 2.45 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Xtrriage (Phenix) : 1.13
EDS : 2.11
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) : 2.11

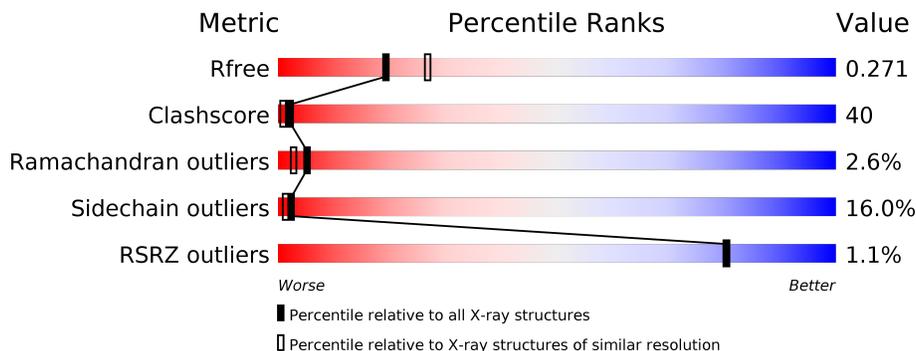
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.45 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	1544 (2.48-2.44)
Clashscore	141614	1613 (2.48-2.44)
Ramachandran outliers	138981	1598 (2.48-2.44)
Sidechain outliers	138945	1598 (2.48-2.44)
RSRZ outliers	127900	1523 (2.48-2.44)

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 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	A	257	
1	B	257	
1	C	257	
1	D	257	

2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 7842 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 NP protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	250	Total 1935	C 1226	N 327	O 368	S 14	0	0	0
1	B	248	Total 1924	C 1218	N 324	O 368	S 14	0	0	0
1	C	251	Total 1944	C 1231	N 328	O 371	S 14	0	0	0
1	D	248	Total 1924	C 1218	N 324	O 368	S 14	0	0	0

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	SER	-	EXPRESSION TAG	UNP E2DQZ4
A	-1	ASN	-	EXPRESSION TAG	UNP E2DQZ4
A	0	ALA	-	EXPRESSION TAG	UNP E2DQZ4
A	69	ASP	ARG	ENGINEERED MUTATION	UNP E2DQZ4
A	72	GLU	LYS	ENGINEERED MUTATION	UNP E2DQZ4
A	79	GLU	LYS	ENGINEERED MUTATION	UNP E2DQZ4
B	-2	SER	-	EXPRESSION TAG	UNP E2DQZ4
B	-1	ASN	-	EXPRESSION TAG	UNP E2DQZ4
B	0	ALA	-	EXPRESSION TAG	UNP E2DQZ4
B	69	ASP	ARG	ENGINEERED MUTATION	UNP E2DQZ4
B	72	GLU	LYS	ENGINEERED MUTATION	UNP E2DQZ4
B	79	GLU	LYS	ENGINEERED MUTATION	UNP E2DQZ4
C	-2	SER	-	EXPRESSION TAG	UNP E2DQZ4
C	-1	ASN	-	EXPRESSION TAG	UNP E2DQZ4
C	0	ALA	-	EXPRESSION TAG	UNP E2DQZ4
C	69	ASP	ARG	ENGINEERED MUTATION	UNP E2DQZ4
C	72	GLU	LYS	ENGINEERED MUTATION	UNP E2DQZ4
C	79	GLU	LYS	ENGINEERED MUTATION	UNP E2DQZ4
D	-2	SER	-	EXPRESSION TAG	UNP E2DQZ4
D	-1	ASN	-	EXPRESSION TAG	UNP E2DQZ4
D	0	ALA	-	EXPRESSION TAG	UNP E2DQZ4

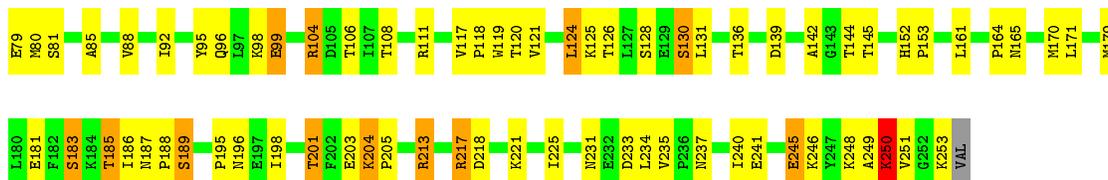
Continued on next page...

Continued from previous page...

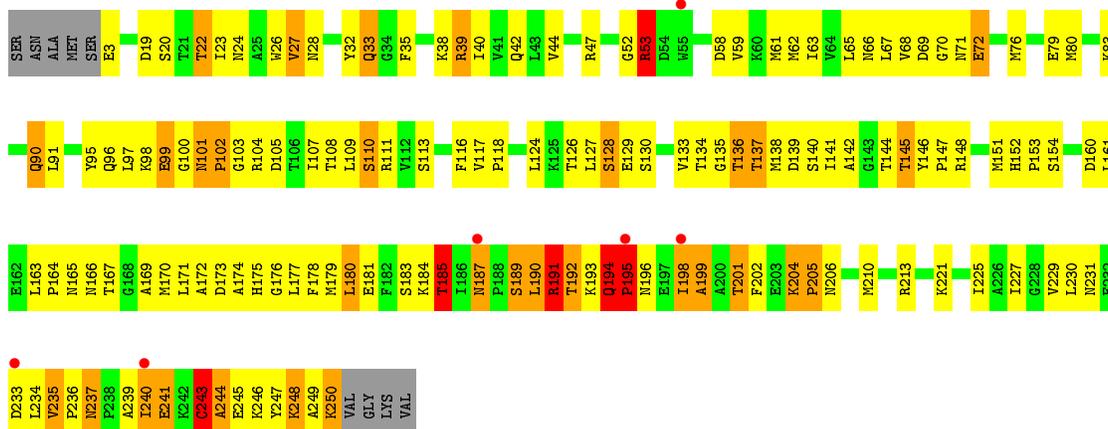
Chain	Residue	Modelled	Actual	Comment	Reference
D	69	ASP	ARG	ENGINEERED MUTATION	UNP E2DQZ4
D	72	GLU	LYS	ENGINEERED MUTATION	UNP E2DQZ4
D	79	GLU	LYS	ENGINEERED MUTATION	UNP E2DQZ4

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	39	Total O 39 39	0	0
2	B	21	Total O 21 21	0	0
2	C	36	Total O 36 36	0	0
2	D	19	Total O 19 19	0	0



• Molecule 1: NP protein



4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	54.16Å 76.80Å 77.16Å 82.79° 69.51° 69.36°	Depositor
Resolution (Å)	28.81 – 2.45 28.82 – 2.45	Depositor EDS
% Data completeness (in resolution range)	95.2 (28.81-2.45) 95.0 (28.82-2.45)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.49 (at 2.45Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, R_{free}	0.210 , 0.291 0.205 , 0.271	Depositor DCC
R_{free} test set	2046 reflections (5.39%)	wwPDB-VP
Wilson B-factor (Å ²)	48.1	Xtriage
Anisotropy	0.075	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 44.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.004 for h,l,h-k 0.004 for h,h-l,k 0.469 for h,h-k,h-l 0.013 for -h,-h+k,-l 0.013 for -h,-k,-h+l 0.015 for -h,-l,-k 0.015 for -h,-h+l,-h+k	Xtriage
Reported twinning fraction	0.536 for H,K,L 0.464 for H,H-K,H-L	Depositor
Outliers	0 of 37978 reflections	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	7842	wwPDB-VP
Average B, all atoms (Å ²)	70.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.34% of the height of the origin peak. No significant pseudotranslation is detected.*

¹ Intensities estimated from amplitudes.

² Theoretical 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	A	1.22	0/1970	0.77	1/2663 (0.0%)
1	B	1.02	0/1959	0.80	1/2649 (0.0%)
1	C	1.25	0/1979	0.80	3/2675 (0.1%)
1	D	1.06	2/1959 (0.1%)	0.86	4/2649 (0.2%)
All	All	1.14	2/7867 (0.0%)	0.81	9/10636 (0.1%)

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
1	D	0	1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	205	PRO	N-CD	5.38	1.55	1.47
1	D	195	PRO	N-CD	5.12	1.55	1.47

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	72	GLU	C-N-CD	6.23	141.49	128.40
1	D	187	ASN	C-N-CD	6.22	141.45	128.40
1	D	101	ASN	C-N-CD	6.04	141.09	128.40
1	B	101	ASN	C-N-CD	5.92	140.84	128.40
1	A	187	ASN	C-N-CD	5.84	140.67	128.40

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	D	194	GLN	Peptide

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	1935	0	1964	107	0
1	B	1924	0	1945	214	0
1	C	1944	0	1972	106	0
1	D	1924	0	1947	218	0
2	A	39	0	0	4	0
2	B	21	0	0	6	0
2	C	36	0	0	5	0
2	D	19	0	0	3	0
All	All	7842	0	7828	615	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 40.

The worst 5 of 615 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:123:ALA:O	1:B:127:LEU:CD2	1.78	1.30
1:B:103:GLY:N	1:B:106:THR:OG1	1.70	1.25
1:A:7:ARG:O	1:A:10:ALA:O	1.55	1.23
1:B:174:ALA:CA	1:B:244:ALA:HB2	1.68	1.23
1:B:162:GLU:OE1	2:B:308:HOH:O	1.54	1.20

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	248/257 (96%)	237 (96%)	8 (3%)	3 (1%)	13	12
1	B	246/257 (96%)	221 (90%)	15 (6%)	10 (4%)	3	1
1	C	249/257 (97%)	238 (96%)	8 (3%)	3 (1%)	13	12
1	D	246/257 (96%)	217 (88%)	19 (8%)	10 (4%)	3	1
All	All	989/1028 (96%)	913 (92%)	50 (5%)	26 (3%)	5	3

5 of 26 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	7	ARG
1	B	104	ARG
1	B	237	ASN
1	C	4	ASP
1	C	7	ARG

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	211/217 (97%)	182 (86%)	29 (14%)	3	3
1	B	210/217 (97%)	171 (81%)	39 (19%)	1	1
1	C	212/217 (98%)	183 (86%)	29 (14%)	3	3
1	D	210/217 (97%)	172 (82%)	38 (18%)	1	1
All	All	843/868 (97%)	708 (84%)	135 (16%)	2	1

5 of 135 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	204	LYS
1	C	53	ARG
1	D	198	ILE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	207	MET
1	C	4	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	166	ASN
1	A	175	HIS
1	B	66	ASN
1	B	237	ASN
1	C	28	ASN

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	250/257 (97%)	-0.28	1 (0%) 92 93	34, 59, 85, 102	0
1	B	248/257 (96%)	-0.07	4 (1%) 72 69	46, 82, 111, 128	0
1	C	251/257 (97%)	-0.31	0 100 100	34, 59, 83, 97	0
1	D	248/257 (96%)	-0.01	6 (2%) 59 54	46, 79, 116, 161	0
All	All	997/1028 (96%)	-0.17	11 (1%) 80 80	34, 70, 105, 161	0

The worst 5 of 11 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	198	ILE	3.6
1	D	195	PRO	3.3
1	B	146	TYR	3.0
1	B	202	PHE	2.9
1	D	233	ASP	2.6

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

There are no such residues in this entry.