



wwPDB X-ray Structure Validation Summary Report

Feb 28, 2014 – 05:38 AM GMT

PDB ID : 4J4X
Title : Crystal structure of 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 : 2.51 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

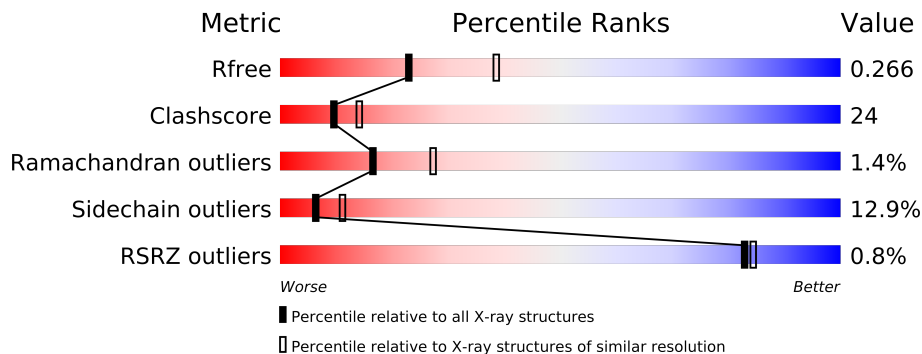
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

1 Overall quality at a glance

The reported resolution of this entry is 2.51 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}	66092	2784 (2.50-2.50)
Clashscore	79885	3562 (2.50-2.50)
Ramachandran outliers	78287	3480 (2.50-2.50)
Sidechain outliers	78261	3482 (2.50-2.50)
RSRZ outliers	66119	2785 (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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	257	
1	B	257	
1	C	257	
1	D	257	
1	E	257	
1	F	257	

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 11737 atoms, of which 0 are hydrogen and 0 are deuterium.

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
1	A	250	Total	C	N	O	S	0	0	0
			1938	1230	332	362	14			
1	B	248	Total	C	N	O	S	0	0	0
			1921	1220	329	358	14			
1	C	246	Total	C	N	O	S	0	0	0
			1910	1213	327	356	14			
1	D	249	Total	C	N	O	S	0	0	0
			1930	1226	331	359	14			
1	E	250	Total	C	N	O	S	0	0	0
			1938	1229	331	364	14			
1	F	248	Total	C	N	O	S	0	0	0
			1921	1220	329	358	14			

There are 18 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
B	-2	SER	-	EXPRESSION TAG	UNP E2DQZ4
B	-1	ASN	-	EXPRESSION TAG	UNP E2DQZ4
B	0	ALA	-	EXPRESSION TAG	UNP E2DQZ4
C	-2	SER	-	EXPRESSION TAG	UNP E2DQZ4
C	-1	ASN	-	EXPRESSION TAG	UNP E2DQZ4
C	0	ALA	-	EXPRESSION TAG	UNP E2DQZ4
D	-2	SER	-	EXPRESSION TAG	UNP E2DQZ4
D	-1	ASN	-	EXPRESSION TAG	UNP E2DQZ4
D	0	ALA	-	EXPRESSION TAG	UNP E2DQZ4
E	-2	SER	-	EXPRESSION TAG	UNP E2DQZ4
E	-1	ASN	-	EXPRESSION TAG	UNP E2DQZ4
E	0	ALA	-	EXPRESSION TAG	UNP E2DQZ4
F	-2	SER	-	EXPRESSION TAG	UNP E2DQZ4
F	-1	ASN	-	EXPRESSION TAG	UNP E2DQZ4

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	0	ALA	-	EXPRESSION TAG	UNP E2DQZ4

- Molecule 2 is water.

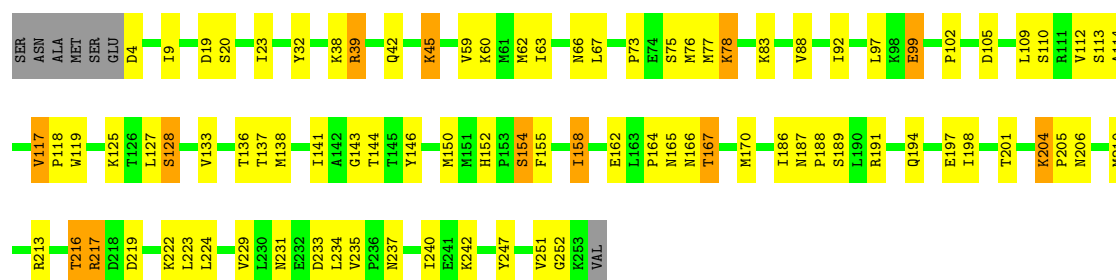
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	47	Total O 47 47	0	0
2	B	40	Total O 40 40	0	0
2	C	19	Total O 19 19	0	0
2	D	27	Total O 27 27	0	0
2	E	23	Total O 23 23	0	0
2	F	23	Total O 23 23	0	0

3 Residue-property plots

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.

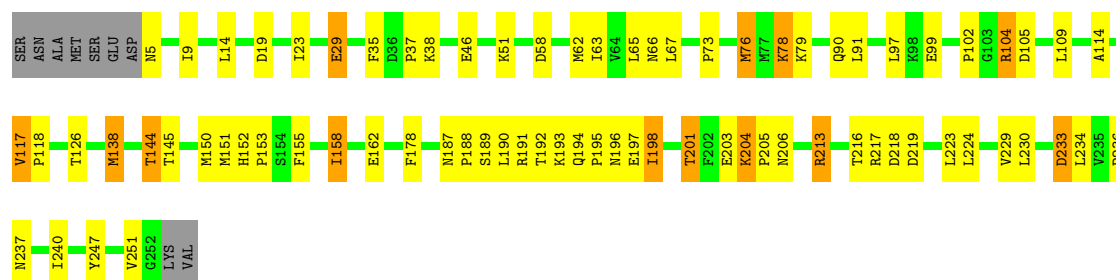
• Molecule 1: NP protein

Chain A: 



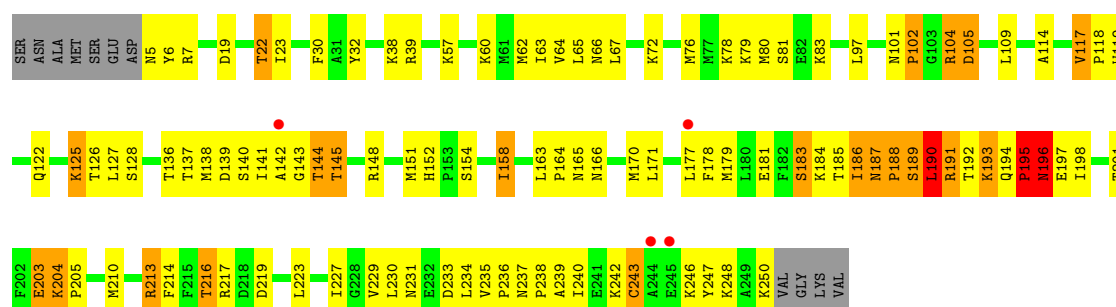
• Molecule 1: NP protein

Chain B: 

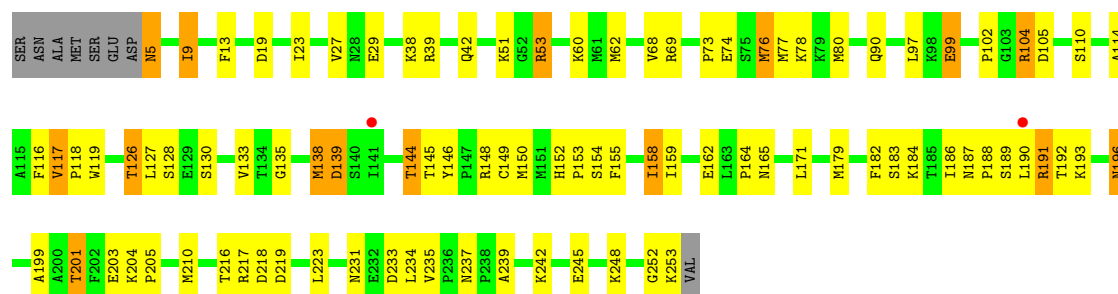


• Molecule 1: NP protein

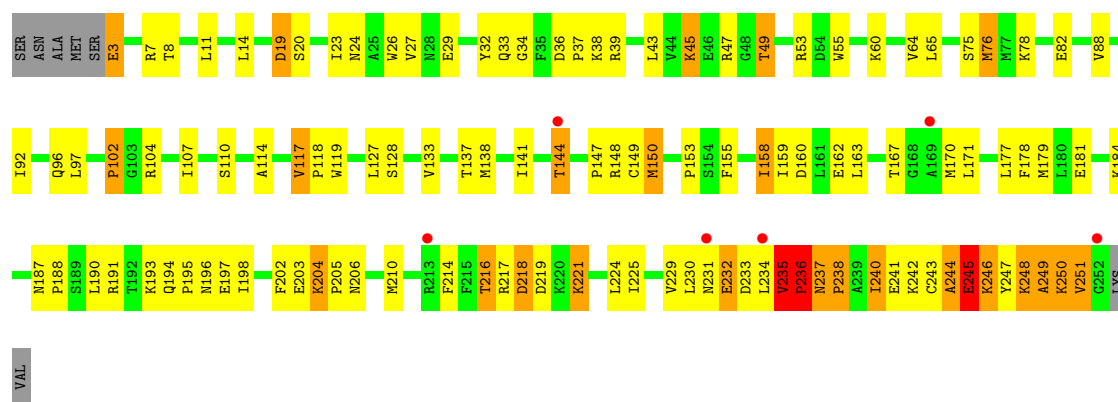
Chain C: 



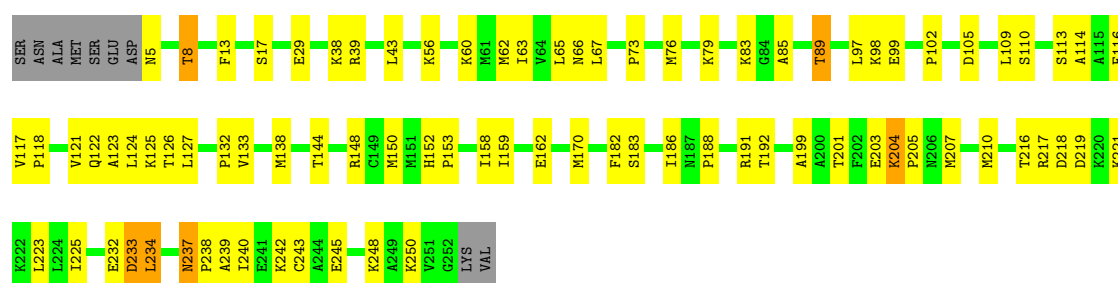
• Molecule 1: NP protein

Chain D: 

- Molecule 1: NP protein

Chain E: 

- Molecule 1: NP protein

Chain F: 

4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	51.29Å 96.34Å 100.80Å 64.57° 81.72° 85.10°	Depositor
Resolution (Å)	49.59 – 2.51 49.59 – 2.51	Depositor EDS
% Data completeness (in resolution range)	96.8 (49.59-2.51) 96.8 (49.59-2.51)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.36 (at 2.51Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, R_{free}	0.195 , 0.266 0.195 , 0.266	Depositor DCC
R_{free} test set	2905 reflections (5.33%)	DCC
Wilson B-factor (Å ²)	48.4	Xtriage
Anisotropy	0.017	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 43.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 57405 reflections	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	11737	wwPDB-VP
Average B, all atoms (Å ²)	58.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.87% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

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	1.22	0/1973	0.73	1/2664 (0.0%)
1	B	1.19	0/1956	0.71	0/2642
1	C	1.12	2/1945 (0.1%)	0.76	2/2627 (0.1%)
1	D	1.11	0/1965	0.69	0/2653
1	E	1.03	2/1973 (0.1%)	0.73	2/2665 (0.1%)
1	F	1.11	0/1956	0.72	0/2642
All	All	1.13	4/11768 (0.0%)	0.72	5/15893 (0.0%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	195	PRO	N-CD	5.49	1.55	1.47
1	C	205	PRO	N-CD	5.37	1.55	1.47
1	E	236	PRO	N-CD	5.36	1.55	1.47
1	E	238	PRO	N-CD	5.24	1.55	1.47

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	187	ASN	C-N-CD	6.10	141.21	128.40
1	C	204	LYS	C-N-CD	5.64	140.25	128.40
1	E	237	ASN	C-N-CD	5.59	140.13	128.40
1	E	235	VAL	C-N-CD	5.52	139.99	128.40
1	A	252	GLY	N-CA-C	-5.32	99.80	113.10

There are no chirality outliers.

There are no planarity outliers.

5.2 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 hydrogens

added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1938	0	1990	63	1
1	B	1921	0	1973	65	0
1	C	1910	0	1959	143	1
1	D	1930	0	1985	85	0
1	E	1938	0	1983	157	0
1	F	1921	0	1970	76	0
2	A	47	0	0	5	0
2	B	40	0	0	2	0
2	C	19	0	0	5	0
2	D	27	0	0	4	0
2	E	23	0	0	12	0
2	F	23	0	0	3	0
All	All	11737	0	11860	572	1

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 24.

The worst 5 of 572 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:E:194:GLN:CB	1:E:197:GLU:OE2	1.71	1.39
1:C:139:ASP:OD2	1:C:148:ARG:NH1	1.60	1.34
1:E:194:GLN:HB2	1:E:197:GLU:OE2	1.15	1.32
1:C:250:LYS:O	2:C:315:HOH:O	1.52	1.26
1:E:20:SER:O	1:E:24:ASN:ND2	1.65	1.25

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	Distance(Å)	Clash(Å)
1:A:19:ASP:OD2	1:C:7:ARG:NH2[1_654]	1.95	0.25

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	248/257 (96%)	245 (99%)	2 (1%)	1 (0%)	43	66
1	B	246/257 (96%)	242 (98%)	3 (1%)	1 (0%)	43	66
1	C	244/257 (95%)	224 (92%)	10 (4%)	10 (4%)	4	5
1	D	247/257 (96%)	242 (98%)	4 (2%)	1 (0%)	43	66
1	E	248/257 (96%)	234 (94%)	8 (3%)	6 (2%)	9	13
1	F	246/257 (96%)	241 (98%)	4 (2%)	1 (0%)	43	66
All	All	1479/1542 (96%)	1428 (97%)	31 (2%)	20 (1%)	16	27

5 of 20 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	145	THR
1	C	190	LEU
1	C	196	ASN
1	C	203	GLU
1	C	143	GLY

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	211/217 (97%)	185 (88%)	26 (12%)	7	12
1	B	209/217 (96%)	185 (88%)	24 (12%)	8	15
1	C	208/217 (96%)	181 (87%)	27 (13%)	6	11
1	D	210/217 (97%)	179 (85%)	31 (15%)	4	8
1	E	211/217 (97%)	180 (85%)	31 (15%)	4	8
1	F	209/217 (96%)	186 (89%)	23 (11%)	9	16
All	All	1258/1302 (97%)	1096 (87%)	162 (13%)	6	11

5 of 162 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	213	ARG
1	D	110	SER
1	F	113	SER
1	C	217	ARG
1	D	51	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	66	ASN
1	F	237	ASN
1	D	90	GLN
1	C	187	ASN
1	D	152	HIS

5.3.3 RNA ⓘ

There are no RNA chains 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 carbohydrates 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

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, 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%)	-0.19	0 100 100	30, 47, 68, 96	0
1	B	248/257 (96%)	-0.12	0 100 100	30, 48, 74, 87	0
1	C	246/257 (95%)	0.09	4 (1%) 68 71	31, 60, 103, 120	0
1	D	249/257 (96%)	-0.07	2 (0%) 83 84	34, 59, 86, 95	0
1	E	250/257 (97%)	0.10	6 (2%) 56 58	37, 67, 104, 123	0
1	F	248/257 (96%)	-0.12	0 100 100	34, 58, 78, 91	0
All	All	1491/1542 (96%)	-0.05	12 (0%) 83 84	30, 56, 91, 123	0

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	245	GLU	3.1
1	E	252	GLY	2.6
1	E	144	THR	2.5
1	E	169	ALA	2.4
1	E	234	LEU	2.3

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 carbohydrates 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.