W O R L D W I D E
wwPDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 10, 2023 – 05:25 PM EDT

PDB ID	:	4JPJ
Title	:	Crystal structure of the germline-targeting HIV-1 gp120 engineered outer domain, eOD-GT6
Authors	:	Julien, J.-P.; Jardine, J.; Schief, W.R.; Wilson, I.A.
Deposited on	:	2013-03-19
Resolution	:	2.50 Å(reported)

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

MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
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

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 2.50 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	Percentile Ranks	Value
Rfree		0.337
Clashscore		5
Ramachandran outliers		2.0%
Sidechain outliers		8.3%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution
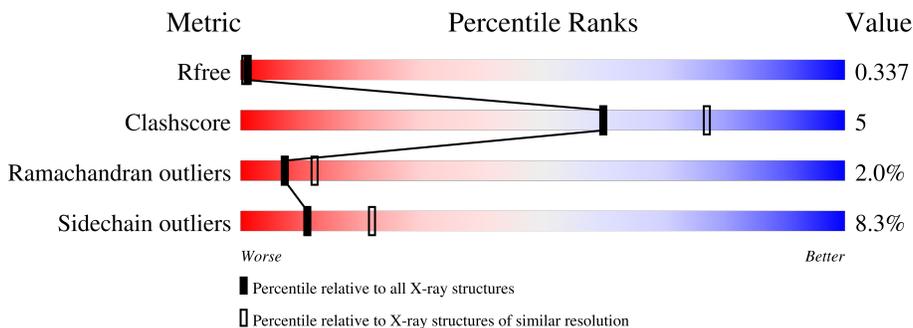
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)

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 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%

| Mol	Chain	Length	Quality of chain
1	A	181	74% 18% •• 6%
1	B	181	80% 10% • 7%
1	C	181	74% 10% • 13%
1	D	181	74% 14% • 10%

2   Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 5228 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 Germline-targeting HIV-1 gp120 engineered outer domain, eOD-GT6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	170	Total	C	N	O	S	0	0	0
			1308	809	226	264	9			
1	B	169	Total	C	N	O	S	0	0	0
			1301	807	224	261	9			
1	C	157	Total	C	N	O	S	0	0	0
			1210	751	207	244	8			
1	D	162	Total	C	N	O	S	0	0	0
			1252	777	211	255	9			

- Molecule 2 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).
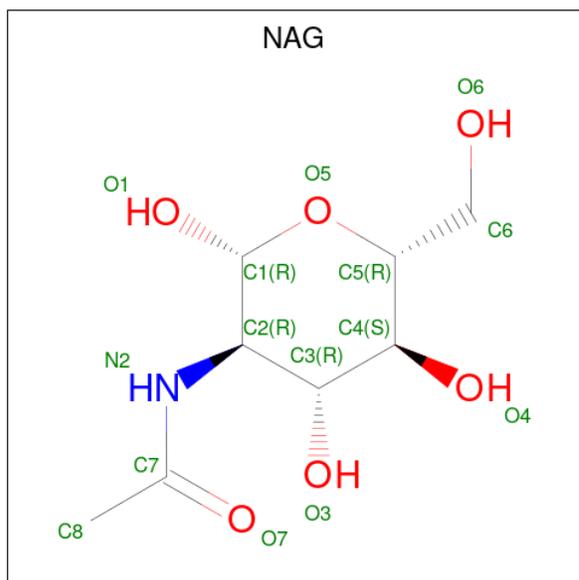
NAG

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			14	8	1	5		

Continued on next page...

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total 14	C 8	N 1	O 5	0	0
2	B	1	Total 14	C 8	N 1	O 5	0	0
2	B	1	Total 14	C 8	N 1	O 5	0	0
2	C	1	Total 14	C 8	N 1	O 5	0	0
2	C	1	Total 14	C 8	N 1	O 5	0	0
2	D	1	Total 14	C 8	N 1	O 5	0	0
2	D	1	Total 14	C 8	N 1	O 5	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	9	Total 9	O 9	0	0
3	B	23	Total 23	O 23	0	0
3	C	9	Total 9	O 9	0	0
3	D	4	Total 4	O 4	0	0

3 Residue-property plots ⓘ

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. 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. 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: Germline-targeting HIV-1 gp120 engineered outer domain, eOD-GT6
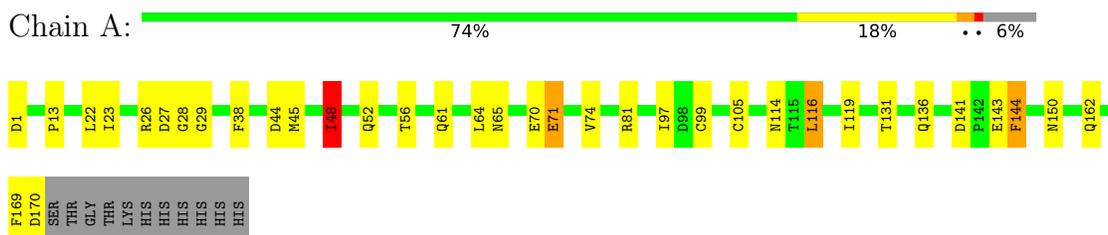
Chain A:

- Molecule 1: Germline-targeting HIV-1 gp120 engineered outer domain, eOD-GT6

Chain B:

- Molecule 1: Germline-targeting HIV-1 gp120 engineered outer domain, eOD-GT6

Chain C:

- Molecule 1: Germline-targeting HIV-1 gp120 engineered outer domain, eOD-GT6

Chain D:
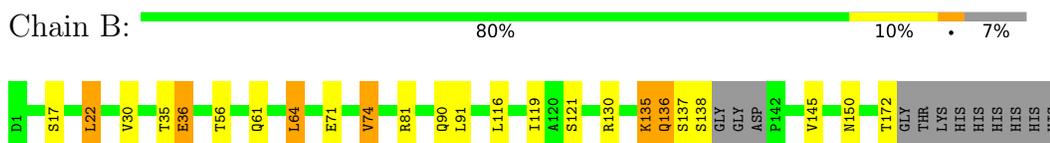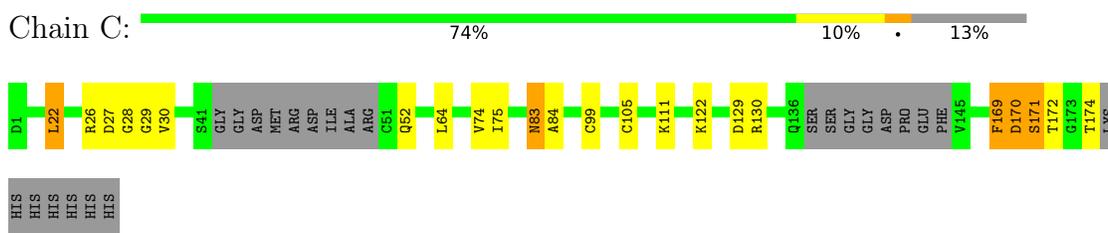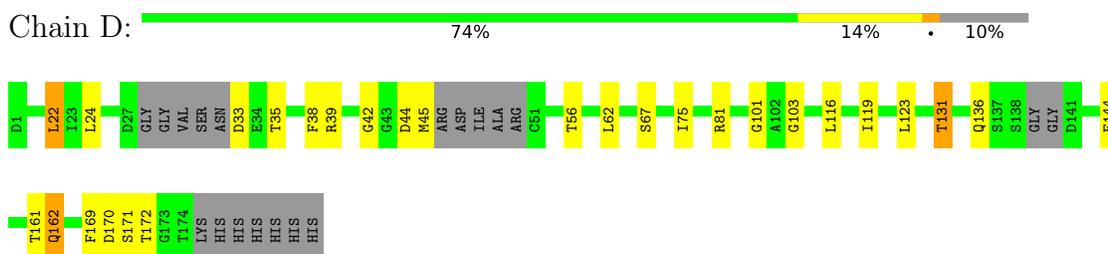
4　Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	44.96Å　217.73Å　44.99Å 90.00°　119.97°　90.00°	Depositor
Resolution (Å)	38.97　–　2.50 38.97　–　2.50	Depositor EDS
% Data completeness (in resolution range)	99.7 (38.97-2.50) 99.6 (38.97-2.50)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.14	Depositor
$< I/\sigma(I) >$ [1]	2.80 (at 2.51Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, R_{free}	0.210　,　0.262 0.298　,　0.337	Depositor DCC
R_{free} test set	1438 reflections (5.59%)	wwPDB-VP
Wilson B-factor (Å^2)	28.0	Xtriage
Anisotropy	0.156	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 4.9	EDS
L-test for twinning[2]	$< |L| > = 0.54, < L^2 > = 0.38$	Xtriage
Estimated twinning fraction	0.000 for -h-l,k,h 0.000 for l,k,-h-l 0.000 for h,-k,-h-l 0.207 for -h-l,-k,l 0.000 for l,-k,h	Xtriage
Reported twinning fraction	0.489 for H, K, L 0.102 for L, -K, H 0.408 for H+L, -K, -L	Depositor
Outliers	0 of 25744 reflections	Xtriage
F_o,F_c correlation	0.83	EDS
Total number of atoms	5228	wwPDB-VP
Average B, all atoms (Å^2)	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.00% 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 ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: NAG

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	0.36	0/1333	0.63	0/1806
1	B	0.34	0/1325	0.58	0/1794
1	C	0.33	0/1231	0.56	0/1668
1	D	0.33	0/1274	0.55	0/1724
All	All	0.34	0/5163	0.58	0/6992

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	A	0	1
1	C	0	1
All	All	0	2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	29	GLY	Peptide
1	C	169	PHE	Peptide

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	1308	0	1238	12	0
1	B	1301	0	1233	9	0
1	C	1210	0	1151	11	0
1	D	1252	0	1179	14	0
2	A	28	0	26	0	0
2	B	28	0	26	0	0
2	C	28	0	26	0	0
2	D	28	0	26	0	0
3	A	9	0	0	0	0
3	B	23	0	0	0	0
3	C	9	0	0	1	0
3	D	4	0	0	0	0
All	All	5228	0	4905	46	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 5.

The worst 5 of 46 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:D:170:ASP:N	1:D:171:SER:HA	1.92	0.82
1:A:27:ASP:N	1:A:28:GLY:HA2	2.03	0.74
1:A:143:GLU:O	1:A:144:PHE:HB2	1.90	0.71
1:A:97:ILE:HG22	1:A:105:CYS:HB3	1.78	0.66
1:A:116:LEU:HA	1:A:119:ILE:HG22	1.79	0.64

There are no symmetry-related clashes.

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	168/181 (93%)	144 (86%)	18 (11%)	6 (4%)	3	4
1	B	165/181 (91%)	157 (95%)	6 (4%)	2 (1%)	13	24
1	C	151/181 (83%)	136 (90%)	11 (7%)	4 (3%)	5	8
1	D	154/181 (85%)	135 (88%)	18 (12%)	1 (1%)	25	43
All	All	638/724 (88%)	572 (90%)	53 (8%)	13 (2%)	7	12

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	48	ILE
1	A	71	GLU
1	A	144	PHE
1	B	30	VAL
1	B	136	GLN

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	148/158 (94%)	134 (90%)	14 (10%)	8	17
1	B	147/158 (93%)	132 (90%)	15 (10%)	7	14
1	C	139/158 (88%)	128 (92%)	11 (8%)	12	24
1	D	144/158 (91%)	136 (94%)	8 (6%)	21	40
All	All	578/632 (92%)	530 (92%)	48 (8%)	11	22

5 of 48 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	172	THR
1	C	111	LYS
1	C	22	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	74	VAL
1	C	130	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	114	ASN
1	D	118	GLN
1	B	150	ASN
1	C	83	ASN
1	C	90	GLN

5.3.3 RNA ⓘ

There are no RNA molecules 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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

8 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#$|Z| > 2$	Counts	RMSZ	#$|Z| > 2$
2	NAG	D	202	1	14,14,15	0.61	0	17,19,21	1.02	1 (5%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#|Z| > 2	Counts	RMSZ	#|Z| > 2
2	NAG	D	201	1	14,14,15	0.44	0	17,19,21	1.52	1 (5%)
2	NAG	B	201	1	14,14,15	0.53	0	17,19,21	1.91	3 (17%)
2	NAG	B	202	1	14,14,15	0.62	0	17,19,21	1.08	1 (5%)
2	NAG	A	202	1	14,14,15	0.57	0	17,19,21	1.52	2 (11%)
2	NAG	C	202	1	14,14,15	0.48	0	17,19,21	1.49	3 (17%)
2	NAG	C	201	1	14,14,15	0.50	0	17,19,21	1.28	2 (11%)
2	NAG	A	201	1	14,14,15	0.34	0	17,19,21	2.30	4 (23%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	NAG	D	202	1	-	0/6/23/26	0/1/1/1
2	NAG	D	201	1	-	1/6/23/26	0/1/1/1
2	NAG	B	201	1	-	3/6/23/26	0/1/1/1
2	NAG	B	202	1	-	2/6/23/26	0/1/1/1
2	NAG	A	202	1	-	0/6/23/26	0/1/1/1
2	NAG	C	202	1	-	0/6/23/26	0/1/1/1
2	NAG	C	201	1	-	3/6/23/26	0/1/1/1
2	NAG	A	201	1	-	1/6/23/26	0/1/1/1

There are no bond length outliers.

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	201	NAG	C1-O5-C5	7.98	123.00	112.19
2	D	201	NAG	C1-O5-C5	4.92	118.85	112.19
2	B	201	NAG	C1-O5-C5	4.91	118.85	112.19
2	B	201	NAG	O5-C5-C6	4.87	114.83	107.20
2	A	202	NAG	C1-O5-C5	4.44	118.21	112.19

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	201	NAG	O5-C5-C6-O6
2	C	201	NAG	O5-C5-C6-O6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
2	B	202	NAG	O5-C5-C6-O6
2	B	201	NAG	C4-C5-C6-O6
2	B	202	NAG	C4-C5-C6-O6

There are no ring outliers.

No monomer is involved in short contacts.

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 ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.