



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 10, 2026 – 03:15 PM UTC

PDB ID : 6IGF / pdb_00006igf
Title : Crystal structure of Human Papillomavirus type 52 pentamer
Authors : Li, Z.H.; Song, S.; He, M.Z.; Gu, Y.; Li, S.W.
Deposited on : 2018-09-25
Resolution : 2.75 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity : 4-5-2 with Phenix2.0
Xtriage (Phenix) : 2.0
EDS : 3.0
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4 : 9.0.010 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.49

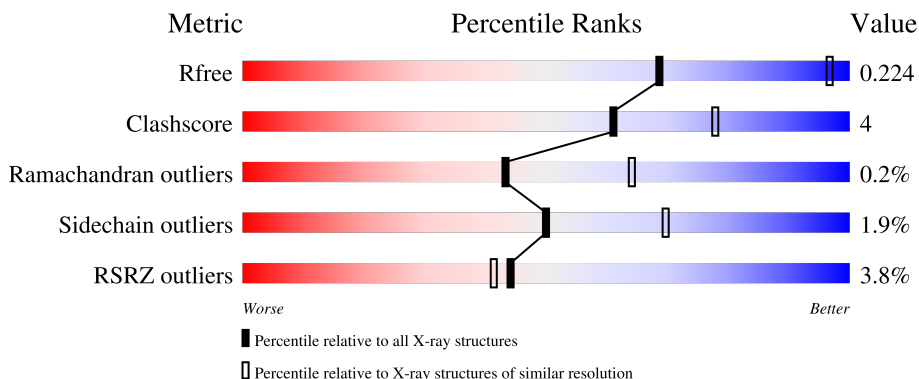
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.75 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}	180053	1009 (2.76-2.76)
Clashscore	190562	1044 (2.76-2.76)
Ramachandran outliers	187476	1024 (2.76-2.76)
Sidechain outliers	187428	1024 (2.76-2.76)
RSRZ outliers	180081	1009 (2.76-2.76)

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 $\leq 5\%$. 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	529	 4% 69% 9% 22%
1	B	529	 4% 72% 7% 20%
1	C	529	 2% 72% 8% 19%
1	D	529	 2% 68% 9% 22%
1	E	529	 2% 74% 8% 19%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	529	 4% 69% 10% 21%
1	G	529	 3% 68% 10% 21%
1	H	529	 3% 68% 10% 21%
1	I	529	 2% 70% 8% 21%
1	J	529	 3% 70% 9% 21%

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 33255 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 Major capsid protein L1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	413	Total 3293	C 2105	N 544	O 621	S 23	0	1	0
1	B	421	Total 3344	C 2139	N 552	O 631	S 22	0	1	0
1	C	427	Total 3377	C 2149	N 562	O 644	S 22	0	0	0
1	D	414	Total 3291	C 2103	N 545	O 621	S 22	0	0	0
1	E	429	Total 3400	C 2168	N 567	O 642	S 23	0	1	0
1	F	417	Total 3316	C 2119	N 548	O 626	S 23	0	1	0
1	G	418	Total 3322	C 2120	N 550	O 630	S 22	0	1	0
1	H	416	Total 3305	C 2111	N 547	O 625	S 22	0	0	0
1	I	416	Total 3300	C 2108	N 547	O 623	S 22	0	0	0
1	J	417	Total 3307	C 2113	N 548	O 625	S 21	0	0	0

4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	306.77Å 105.07Å 196.90Å 90.00° 125.80° 90.00°	Depositor
Resolution (Å)	42.95 – 2.75 42.95 – 2.75	Depositor EDS
% Data completeness (in resolution range)	99.8 (42.95-2.75) 99.8 (42.95-2.75)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.20	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.35 (at 2.77Å)	Xtrriage
Refinement program	PHENIX (1.11.1_2575: ???)	Depositor
R, R_{free}	0.187 , 0.235 (Not available) , 0.224	Depositor DCC
R_{free} test set	6708 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å ²)	43.8	Xtrriage
Anisotropy	0.865	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 42.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	33255	wwPDB-VP
Average B, all atoms (Å ²)	52.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.89% 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, 0.333 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	0.30	1/3381 (0.0%)	0.53	1/4578 (0.0%)
1	B	0.29	0/3434	0.50	0/4652
1	C	0.25	0/3464	0.50	0/4693
1	D	0.30	0/3376	0.53	0/4571
1	E	0.31	0/3491	0.51	0/4730
1	F	0.28	0/3404	0.50	0/4609
1	G	0.32	0/3411	0.56	1/4620 (0.0%)
1	H	0.28	0/3390	0.50	0/4590
1	I	0.30	0/3385	0.54	1/4583 (0.0%)
1	J	0.32	0/3392	0.56	0/4593
All	All	0.30	1/34128 (0.0%)	0.52	3/46219 (0.0%)

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	C	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	166	PRO	CA-C	-7.72	1.47	1.52

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	166	PRO	O-C-N	-8.35	117.22	121.15
1	I	53	LYS	N-CA-C	6.33	113.97	108.78
1	G	71	LEU	CA-CB-CG	-5.18	98.16	116.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	139	ALA	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	3293	0	3191	32	0
1	B	3344	0	3248	27	1
1	C	3377	0	3262	24	0
1	D	3291	0	3188	35	1
1	E	3400	0	3300	23	0
1	F	3316	0	3215	39	1
1	G	3322	0	3216	38	0
1	H	3305	0	3199	42	0
1	I	3300	0	3196	33	0
1	J	3307	0	3205	31	1
All	All	33255	0	32220	275	2

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 4.

The worst 5 of 275 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:54:ASN:HB2	1:A:64:LEU:HD12	1.54	0.89
1:J:79:LYS:HE2	1:J:456:LYS:HE3	1.56	0.87
1:B:120:ILE:HD12	1:E:298:PRO:HB3	1.54	0.87
1:A:25:ASP:OD2	1:A:320:ARG:NH1	2.11	0.84
1:G:443:ASP:HB3	1:G:446:LYS:HG3	1.61	0.82

All (2) 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:D:284:ASN:ND2	1:F:94:TYR:O[3_555]	1.95	0.25
1:B:92:SER:OG	1:J:320:ARG:NH2[2_556]	2.16	0.04

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	406/529 (77%)	391 (96%)	15 (4%)	0	100	100
1	B	414/529 (78%)	401 (97%)	12 (3%)	1 (0%)	43	64
1	C	423/529 (80%)	404 (96%)	19 (4%)	0	100	100
1	D	406/529 (77%)	392 (97%)	14 (3%)	0	100	100
1	E	424/529 (80%)	405 (96%)	18 (4%)	1 (0%)	43	64
1	F	410/529 (78%)	395 (96%)	15 (4%)	0	100	100
1	G	413/529 (78%)	393 (95%)	17 (4%)	3 (1%)	18	34
1	H	408/529 (77%)	391 (96%)	15 (4%)	2 (0%)	24	43
1	I	408/529 (77%)	392 (96%)	16 (4%)	0	100	100
1	J	409/529 (77%)	394 (96%)	15 (4%)	0	100	100
All	All	4121/5290 (78%)	3958 (96%)	156 (4%)	7 (0%)	43	64

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	G	287	ASN
1	H	177	PRO
1	B	142	PRO
1	E	177	PRO
1	H	176	THR

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	368/469 (78%)	363 (99%)	5 (1%)	59	75
1	B	374/469 (80%)	369 (99%)	5 (1%)	61	76
1	C	377/469 (80%)	367 (97%)	10 (3%)	39	62
1	D	367/469 (78%)	358 (98%)	9 (2%)	42	64
1	E	380/469 (81%)	372 (98%)	8 (2%)	47	67
1	F	370/469 (79%)	364 (98%)	6 (2%)	55	73
1	G	371/469 (79%)	362 (98%)	9 (2%)	43	65
1	H	368/469 (78%)	363 (99%)	5 (1%)	59	75
1	I	367/469 (78%)	358 (98%)	9 (2%)	42	64
1	J	368/469 (78%)	363 (99%)	5 (1%)	59	75
All	All	3710/4690 (79%)	3639 (98%)	71 (2%)	50	70

5 of 71 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	I	44	THR
1	I	54	ASN
1	I	291	VAL
1	D	62	LYS
1	D	61	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 28 such sidechains are listed below:

Mol	Chain	Res	Type
1	F	257	GLN
1	J	171	HIS
1	H	136	ASN
1	I	199	GLN
1	G	239	GLN

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 oligosaccharides 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	413/529 (78%)	0.17	19 (4%) 37 36	28, 55, 113, 200	1 (0%)
1	B	421/529 (79%)	0.04	21 (4%) 34 34	25, 47, 104, 163	1 (0%)
1	C	427/529 (80%)	0.11	11 (2%) 57 55	33, 53, 112, 170	0
1	D	414/529 (78%)	-0.06	13 (3%) 51 49	27, 47, 96, 195	0
1	E	429/529 (81%)	-0.21	13 (3%) 52 50	20, 39, 98, 154	1 (0%)
1	F	417/529 (78%)	0.17	22 (5%) 32 31	25, 51, 111, 166	1 (0%)
1	G	418/529 (79%)	-0.15	16 (3%) 44 41	17, 38, 97, 154	1 (0%)
1	H	416/529 (78%)	0.14	16 (3%) 44 41	29, 53, 105, 151	0
1	I	416/529 (78%)	-0.11	12 (2%) 53 52	24, 43, 93, 169	0
1	J	417/529 (78%)	-0.25	18 (4%) 40 38	20, 36, 92, 150	0
All	All	4188/5290 (79%)	-0.01	161 (3%) 44 41	17, 47, 103, 200	5 (0%)

The worst 5 of 161 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	478	LEU	6.8
1	B	478	LEU	6.6
1	F	478	LEU	5.9
1	I	408	PHE	5.8
1	G	286	GLY	5.5

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 oligosaccharides 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.