



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2026 – 09:38 PM UTC

PDB ID : 4AR2 / pdb_00004ar2
Title : Dodecahedron formed of penton base protein from adenovirus Ad3
Authors : Burmeister, W.P.; Szolajska, E.; Zochowska, M.; Nerlo, B.; Andreev, I.; Schoehn, G.; Andrieu, J.-P.; Fender, P.; Naskalska, A.; Zubieta, C.; Cusack, S.; Chroboczek, J.
Deposited on : 2012-04-20
Resolution : 3.80 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

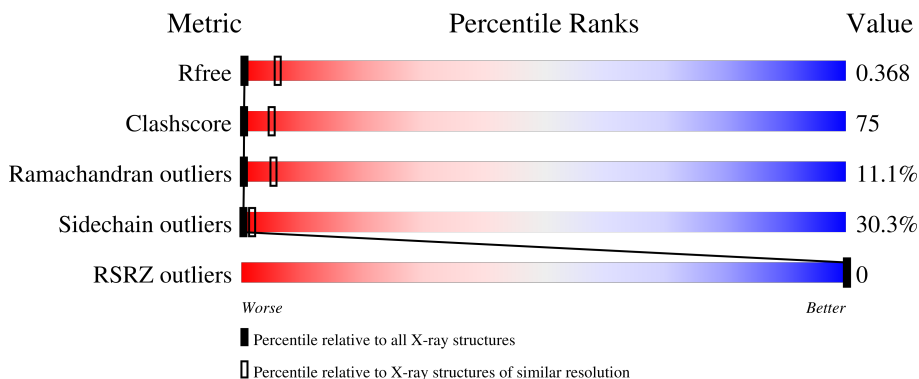
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.80 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	1065 (3.96-3.64)
Clashscore	190562	1012 (3.94-3.66)
Ramachandran outliers	187476	1048 (3.96-3.64)
Sidechain outliers	187428	1043 (3.96-3.64)
RSRZ outliers	180081	1064 (3.96-3.64)

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	509	
2	P	21	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 3757 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 L2 PROTEIN III (PENTON BASE).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	455	3673	2324	629	706	14	0	0	0

- Molecule 2 is a protein called FIBER PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	P	10	83	53	11	18	1	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
P	21	CYS	-	expression tag	UNP P03275

- Molecule 3 is CALCIUM ION (CCD ID: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Ca	0	0
			1	1		

4 Data and refinement statistics i

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	346.04Å 348.55Å 372.87Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	258.20 – 3.80 254.63 – 3.80	Depositor EDS
% Data completeness (in resolution range)	56.8 (258.20-3.80) 47.7 (254.63-3.80)	Depositor EDS
R_{merge}	0.17	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.15 (at 3.75Å)	Xtrriage
Refinement program	REFMAC 5.6.0114	Depositor
R, R_{free}	0.280 , 0.278 0.371 , 0.368	Depositor DCC
R_{free} test set	5040 reflections (2.02%)	wwPDB-VP
Wilson B-factor (Å ²)	101.2	Xtrriage
Anisotropy	0.277	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.26 , 0.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.40$, $\langle L^2 \rangle = 0.22$	Xtrriage
Estimated twinning fraction	0.065 for k,h,-l	Xtrriage
F_o, F_c correlation	0.51	EDS
Total number of atoms	3757	wwPDB-VP
Average B, all atoms (Å ²)	71.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.98% 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](#)

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

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	0.88	3/3758 (0.1%)	1.56	100/5103 (2.0%)
2	P	0.56	0/86	0.84	0/119
All	All	0.88	3/3844 (0.1%)	1.55	100/5222 (1.9%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	79	ILE	CA-CB	10.02	1.68	1.54
1	A	80	ALA	N-CA	-8.70	1.35	1.46
1	A	79	ILE	C-N	-5.95	1.26	1.33

The worst 5 of 100 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	116	ARG	N-CA-C	-15.22	92.75	112.24
1	A	158	THR	N-CA-C	12.46	126.42	108.86
1	A	469	PHE	N-CA-C	-12.37	89.08	109.24
1	A	414	LEU	N-CA-C	-11.96	94.83	110.39
1	A	51	GLY	N-CA-C	11.90	129.11	110.90

There are no chirality outliers.

There are no planarity outliers.

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	3673	0	3600	556	0
2	P	83	0	68	3	0
3	A	1	0	0	0	0
All	All	3757	0	3668	557	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 75.

The worst 5 of 557 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:304:SER:CB	1:A:351:LEU:HB3	1.65	1.24
1:A:114:ASP:CG	1:A:115:GLU:H	1.49	1.13
1:A:131:MET:H	1:A:492:ASN:ND2	1.46	1.12
1:A:431:ASN:O	1:A:521:ARG:HD3	1.52	1.09
1:A:304:SER:HB3	1:A:351:LEU:HB3	1.12	1.08

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	451/509 (89%)	320 (71%)	81 (18%)	50 (11%)	0	5
2	P	8/21 (38%)	5 (62%)	2 (25%)	1 (12%)	0	4
All	All	459/530 (87%)	325 (71%)	83 (18%)	51 (11%)	0	5

5 of 51 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	49	THR
1	A	57	TYR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	59	GLU
1	A	66	THR
1	A	79	ILE

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	412/457 (90%)	284 (69%)	128 (31%)	0 1
2	P	10/20 (50%)	10 (100%)	0	100 100
All	All	422/477 (88%)	294 (70%)	128 (30%)	0 2

5 of 128 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	492	ASN
1	A	514	VAL
1	A	207	GLN
1	A	202	LEU
1	A	518	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 24 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	366	ASN
1	A	432	ASN
1	A	429	GLN
1	A	461	GLN
1	A	112	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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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	455/509 (89%)	-0.68	0 100 100	13, 61, 170, 199	0
2	P	10/21 (47%)	-0.21	0 100 100	121, 128, 130, 130	10 (100%)
All	All	465/530 (87%)	-0.67	0 100 100	13, 61, 169, 199	10 (2%)

There are no RSRZ outliers to report.

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](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
3	CA	A	999	1/1	0.96	0.07	72,72,72,72	1

6.5 Other polymers [i](#)

There are no such residues in this entry.