



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 5, 2026 – 07:51 PM UTC

PDB ID : 2WC9 / pdb_00002wc9
Title : Crystal structure of the g2p (large terminase) nuclease domain from the bacteriophage SPP1 with bound Mn
Authors : Smits, C.; Chechik, M.; Kovalevskiy, O.V.; Shevtsov, M.B.; Foster, A.W.; Alonso, J.C.; Antson, A.A.
Deposited on : 2009-03-10
Resolution : 2.50 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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
Mogul	:	2022.3.0, CSD as543be (2022)
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

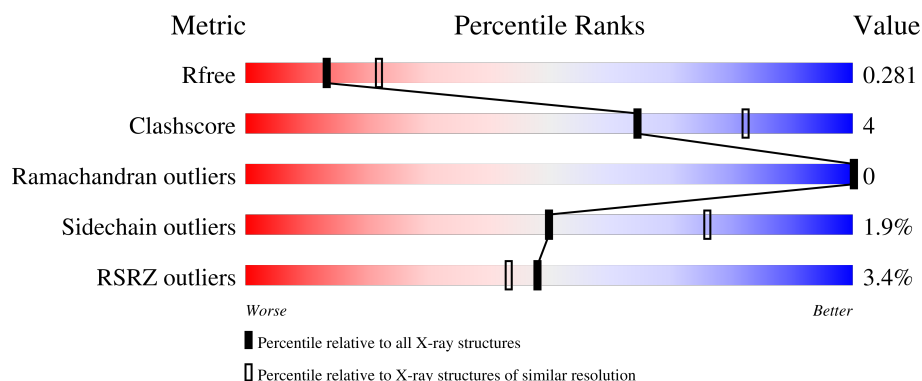
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.50 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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	180053	5829 (2.50-2.50)
Clashscore	190562	6492 (2.50-2.50)
Ramachandran outliers	187476	6378 (2.50-2.50)
Sidechain outliers	187428	6380 (2.50-2.50)
RSRZ outliers	180081	5833 (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 $\leq 5\%$. 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	212	

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 1481 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 TERMINASE LARGE SUBUNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	177	Total	C	N	O	Se	0	0	0
			1450	908	260	281	1			

There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	211	MSE	-	expression tag	UNP P54308
A	212	GLY	-	expression tag	UNP P54308
A	213	SER	-	expression tag	UNP P54308
A	214	SER	-	expression tag	UNP P54308
A	215	HIS	-	expression tag	UNP P54308
A	216	HIS	-	expression tag	UNP P54308
A	217	HIS	-	expression tag	UNP P54308
A	218	HIS	-	expression tag	UNP P54308
A	219	HIS	-	expression tag	UNP P54308
A	220	HIS	-	expression tag	UNP P54308
A	221	SER	-	expression tag	UNP P54308
A	222	SER	-	expression tag	UNP P54308
A	223	GLY	-	expression tag	UNP P54308
A	224	LEU	-	expression tag	UNP P54308
A	225	VAL	-	expression tag	UNP P54308
A	226	PRO	-	expression tag	UNP P54308
A	227	ARG	-	expression tag	UNP P54308
A	228	GLY	-	expression tag	UNP P54308
A	229	SER	-	expression tag	UNP P54308
A	230	HIS	-	expression tag	UNP P54308
A	231	MSE	-	expression tag	UNP P54308

- Molecule 2 is MANGANESE (II) ION (CCD ID: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	2	Total	Mn	0	0
			2	2		

- Molecule 3 is SULFATE ION (CCD ID: SO4) (formula: O₄S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	S	0	0
			5	4	1		

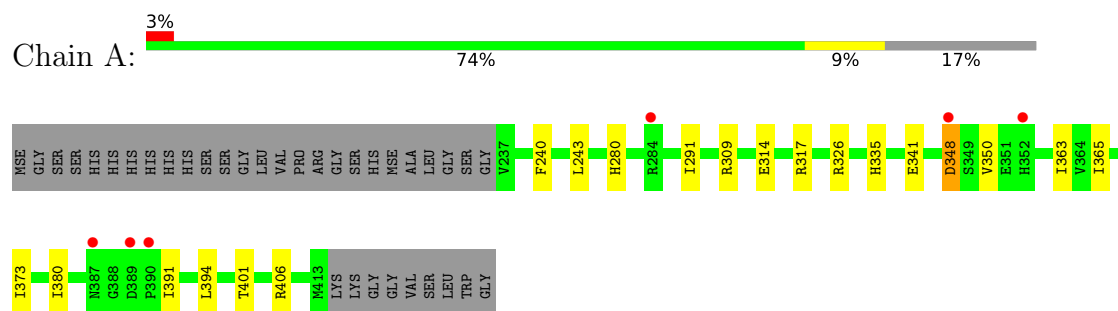
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	24	Total	O	0	0
			24	24		

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 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: TERMINASE LARGE SUBUNIT



4 Data and refinement statistics

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α , β , γ	69.71Å 69.71Å 73.80Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	24.60 – 2.50 24.60 – 2.50	Depositor EDS
% Data completeness (in resolution range)	100.0 (24.60-2.50) 99.9 (24.60-2.50)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.09 (at 2.50Å)	Xtriage
Refinement program	REFMAC 5.5.0082	Depositor
R, R_{free}	0.195 , 0.259 0.208 , 0.281	Depositor DCC
R_{free} test set	349 reflections (4.66%)	wwPDB-VP
Wilson B-factor (Å ²)	59.1	Xtriage
Anisotropy	0.003	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 25.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.035 for -h,-k,l	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	1481	wwPDB-VP
Average B, all atoms (Å ²)	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.95% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: SO4, MN

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	0.65	0/1479	0.82	0/1998

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	1450	0	1405	11	1
2	A	2	0	0	0	0
3	A	5	0	0	0	0
4	A	24	0	0	0	0
All	All	1481	0	1405	11	1

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.

All (11) 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:317:ARG:HD2	1:A:341:GLU:OE2	2.03	0.57
1:A:365:ILE:HD13	1:A:373:ILE:HG22	1.85	0.57
1:A:348:ASP:OD2	1:A:348:ASP:N	2.41	0.54
1:A:243:LEU:HD23	1:A:363:ILE:HB	1.94	0.50
1:A:240:PHE:HB3	1:A:243:LEU:HG	1.94	0.48
1:A:314:GLU:OE1	1:A:314:GLU:N	2.44	0.43
1:A:380:ILE:O	1:A:380:ILE:HG23	2.18	0.43
1:A:373:ILE:HG12	1:A:401:THR:HB	2.00	0.43
1:A:309:ARG:NH2	1:A:335:HIS:O	2.52	0.42
1:A:280:HIS:HB2	1:A:291:ILE:HD11	2.03	0.40
1:A:350:VAL:HG22	1:A:406:ARG:NH2	2.36	0.40

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	Interatomic distance (Å)	Clash overlap (Å)
1:A:326:ARG:NH2	1:A:380:ILE:O[6_655]	2.17	0.03

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	175/212 (82%)	169 (97%)	6 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	155/178 (87%)	152 (98%)	3 (2%)	50 76

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	348	ASP
1	A	391	ILE
1	A	394	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	335	HIS

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 3 ligands modelled in this entry, 2 are monoatomic - leaving 1 for Mogul analysis.

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	$\# Z > 2$
3	SO4	A	1415	-	4,4,4	0.33	0	6,6,6	0.42	0

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	176/212 (83%)	0.30	6 (3%) 48 43	19, 37, 42, 43	1 (0%)

All (6) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	284	ARG	9.3
1	A	389	ASP	3.8
1	A	387	ASN	3.0
1	A	390	PRO	2.7
1	A	348	ASP	2.5
1	A	352	HIS	2.3

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
2	MN	A	1414	1/1	0.44	0.31	43,43,43,43	1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($\text{\AA}^2$)	Q<0.9
3	SO4	A	1415	5/5	0.85	0.09	55,56,58,59	1
2	MN	A	1413	1/1	0.88	0.25	44,44,44,44	1

6.5 Other polymers [i](#)

There are no such residues in this entry.