



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 5, 2026 – 04:32 AM UTC

PDB ID : 8H2N / pdb_00008h2n
Title : gp96 RNA polymerase from P23-45 phage (crystal 2)
Authors : Chaban, A.; Sokolova, M.L.; Tagami, S.
Deposited on : 2022-10-06
Resolution : 4.41 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
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

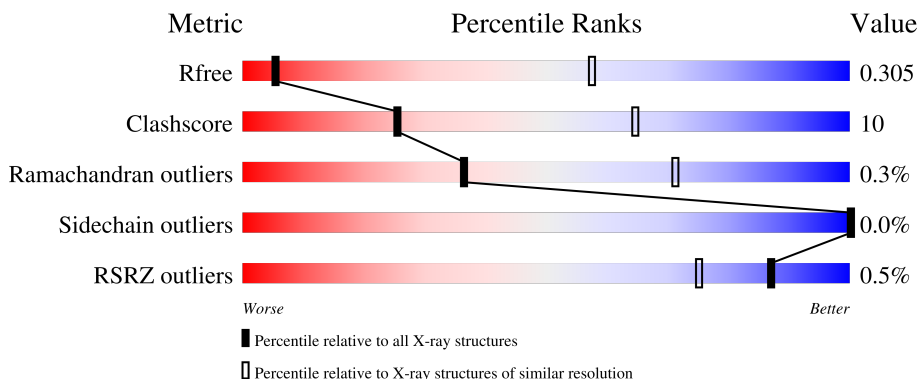
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 4.41 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	1092 (4.92-3.90)
Clashscore	190562	1139 (4.92-3.90)
Ramachandran outliers	187476	1028 (4.92-3.90)
Sidechain outliers	187428	1012 (4.92-3.90)
RSRZ outliers	180081	1088 (4.92-3.90)

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	821	65% 21% 14%
1	B	821	66% 20% 14%
1	C	821	68% 18% 14%
1	D	821	68% 19% 14%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 22384 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 Tape tail measure protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	709	Total	C	N	O	Se	0	0	0
			5595	3533	979	1077	6			
1	B	709	Total	C	N	O	Se	0	0	0
			5595	3533	979	1077	6			
1	C	709	Total	C	N	O	Se	0	0	0
			5595	3533	979	1077	6			
1	D	709	Total	C	N	O	Se	0	0	0
			5595	3533	979	1077	6			

There are 92 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	304	MSE	-	initiating methionine	UNP A7XXD0
A	305	ALA	-	expression tag	UNP A7XXD0
A	306	HIS	-	expression tag	UNP A7XXD0
A	307	HIS	-	expression tag	UNP A7XXD0
A	308	HIS	-	expression tag	UNP A7XXD0
A	309	HIS	-	expression tag	UNP A7XXD0
A	310	HIS	-	expression tag	UNP A7XXD0
A	311	HIS	-	expression tag	UNP A7XXD0
A	312	SER	-	expression tag	UNP A7XXD0
A	313	ALA	-	expression tag	UNP A7XXD0
A	314	ALA	-	expression tag	UNP A7XXD0
A	315	LEU	-	expression tag	UNP A7XXD0
A	316	GLU	-	expression tag	UNP A7XXD0
A	317	VAL	-	expression tag	UNP A7XXD0
A	318	LEU	-	expression tag	UNP A7XXD0
A	319	PHE	-	expression tag	UNP A7XXD0
A	320	GLN	-	expression tag	UNP A7XXD0
A	321	GLY	-	expression tag	UNP A7XXD0
A	322	PRO	-	expression tag	UNP A7XXD0
A	323	GLY	-	expression tag	UNP A7XXD0
A	324	ARG	-	expression tag	UNP A7XXD0

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	325	ASN	-	expression tag	UNP A7XXD0
A	326	ASN	-	expression tag	UNP A7XXD0
B	304	MSE	-	initiating methionine	UNP A7XXD0
B	305	ALA	-	expression tag	UNP A7XXD0
B	306	HIS	-	expression tag	UNP A7XXD0
B	307	HIS	-	expression tag	UNP A7XXD0
B	308	HIS	-	expression tag	UNP A7XXD0
B	309	HIS	-	expression tag	UNP A7XXD0
B	310	HIS	-	expression tag	UNP A7XXD0
B	311	HIS	-	expression tag	UNP A7XXD0
B	312	SER	-	expression tag	UNP A7XXD0
B	313	ALA	-	expression tag	UNP A7XXD0
B	314	ALA	-	expression tag	UNP A7XXD0
B	315	LEU	-	expression tag	UNP A7XXD0
B	316	GLU	-	expression tag	UNP A7XXD0
B	317	VAL	-	expression tag	UNP A7XXD0
B	318	LEU	-	expression tag	UNP A7XXD0
B	319	PHE	-	expression tag	UNP A7XXD0
B	320	GLN	-	expression tag	UNP A7XXD0
B	321	GLY	-	expression tag	UNP A7XXD0
B	322	PRO	-	expression tag	UNP A7XXD0
B	323	GLY	-	expression tag	UNP A7XXD0
B	324	ARG	-	expression tag	UNP A7XXD0
B	325	ASN	-	expression tag	UNP A7XXD0
B	326	ASN	-	expression tag	UNP A7XXD0
C	304	MSE	-	initiating methionine	UNP A7XXD0
C	305	ALA	-	expression tag	UNP A7XXD0
C	306	HIS	-	expression tag	UNP A7XXD0
C	307	HIS	-	expression tag	UNP A7XXD0
C	308	HIS	-	expression tag	UNP A7XXD0
C	309	HIS	-	expression tag	UNP A7XXD0
C	310	HIS	-	expression tag	UNP A7XXD0
C	311	HIS	-	expression tag	UNP A7XXD0
C	312	SER	-	expression tag	UNP A7XXD0
C	313	ALA	-	expression tag	UNP A7XXD0
C	314	ALA	-	expression tag	UNP A7XXD0
C	315	LEU	-	expression tag	UNP A7XXD0
C	316	GLU	-	expression tag	UNP A7XXD0
C	317	VAL	-	expression tag	UNP A7XXD0
C	318	LEU	-	expression tag	UNP A7XXD0
C	319	PHE	-	expression tag	UNP A7XXD0
C	320	GLN	-	expression tag	UNP A7XXD0

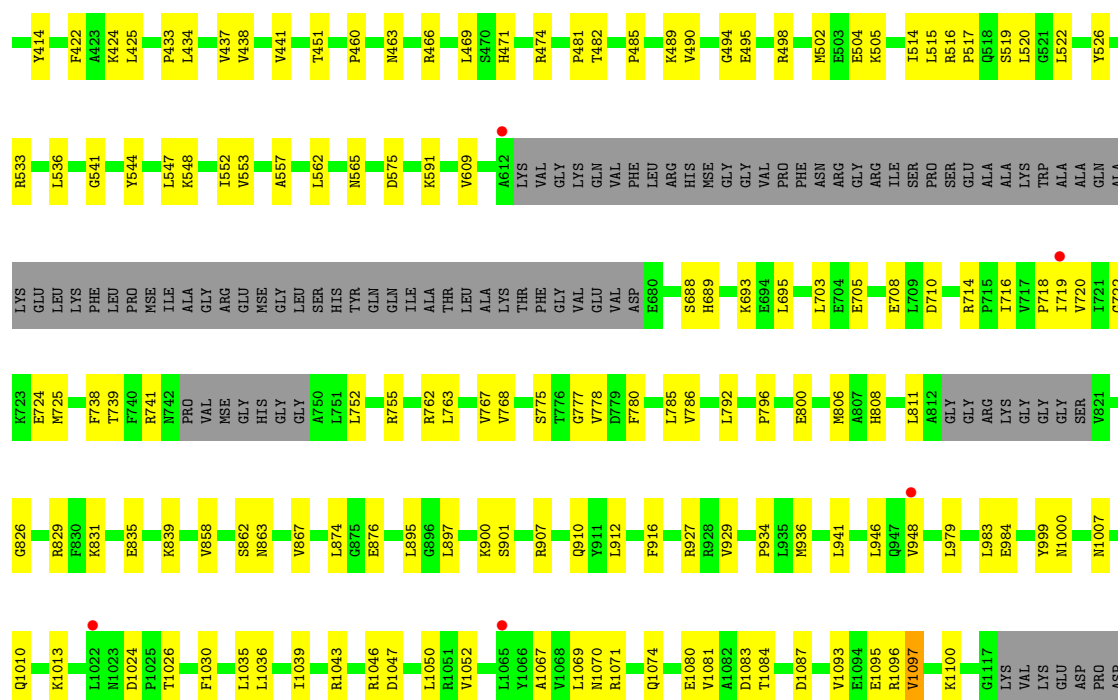
Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	321	GLY	-	expression tag	UNP A7XXD0
C	322	PRO	-	expression tag	UNP A7XXD0
C	323	GLY	-	expression tag	UNP A7XXD0
C	324	ARG	-	expression tag	UNP A7XXD0
C	325	ASN	-	expression tag	UNP A7XXD0
C	326	ASN	-	expression tag	UNP A7XXD0
D	304	MSE	-	initiating methionine	UNP A7XXD0
D	305	ALA	-	expression tag	UNP A7XXD0
D	306	HIS	-	expression tag	UNP A7XXD0
D	307	HIS	-	expression tag	UNP A7XXD0
D	308	HIS	-	expression tag	UNP A7XXD0
D	309	HIS	-	expression tag	UNP A7XXD0
D	310	HIS	-	expression tag	UNP A7XXD0
D	311	HIS	-	expression tag	UNP A7XXD0
D	312	SER	-	expression tag	UNP A7XXD0
D	313	ALA	-	expression tag	UNP A7XXD0
D	314	ALA	-	expression tag	UNP A7XXD0
D	315	LEU	-	expression tag	UNP A7XXD0
D	316	GLU	-	expression tag	UNP A7XXD0
D	317	VAL	-	expression tag	UNP A7XXD0
D	318	LEU	-	expression tag	UNP A7XXD0
D	319	PHE	-	expression tag	UNP A7XXD0
D	320	GLN	-	expression tag	UNP A7XXD0
D	321	GLY	-	expression tag	UNP A7XXD0
D	322	PRO	-	expression tag	UNP A7XXD0
D	323	GLY	-	expression tag	UNP A7XXD0
D	324	ARG	-	expression tag	UNP A7XXD0
D	325	ASN	-	expression tag	UNP A7XXD0
D	326	ASN	-	expression tag	UNP A7XXD0

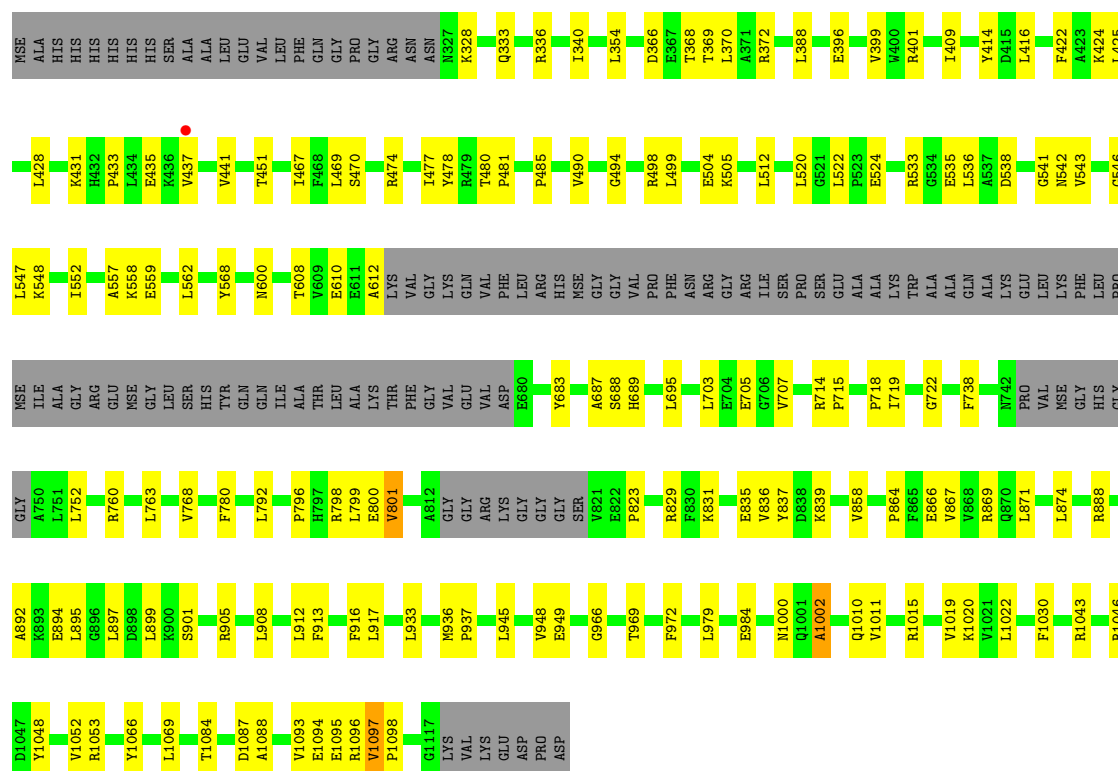
- Molecule 2 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Mg 1 1	0	0
2	B	1	Total Mg 1 1	0	0
2	C	1	Total Mg 1 1	0	0
2	D	1	Total Mg 1 1	0	0



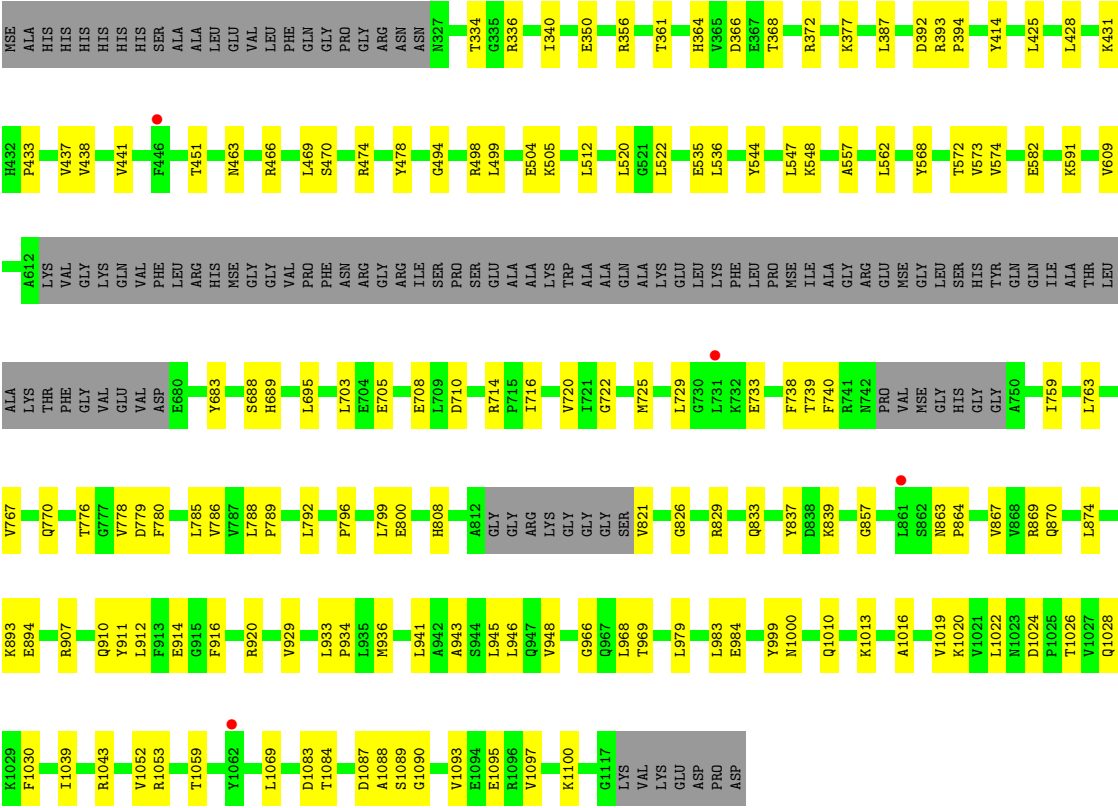
• Molecule 1: Tape tail measure protein

Chain C: 68% 18% 14%



• Molecule 1: Tape tail measure protein

Chain D: 68% 19% 14%



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	118.16Å 128.91Å 136.76Å 90.00° 96.84° 90.00°	Depositor
Resolution (Å)	44.37 – 4.41 44.37 – 4.41	Depositor EDS
% Data completeness (in resolution range)	99.2 (44.37-4.41) 82.0 (44.37-4.41)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	0.58 (at 4.45Å)	Xtriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, R_{free}	0.245 , 0.303 0.250 , 0.305	Depositor DCC
R_{free} test set	2000 reflections (7.71%)	wwPDB-VP
Wilson B-factor (Å ²)	146.8	Xtriage
Anisotropy	0.140	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 169.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.44$, $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.87	EDS
Total number of atoms	22384	wwPDB-VP
Average B, all atoms (Å ²)	203.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.55% 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: MG

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	0.19	1/5672 (0.0%)	0.40	0/7644
1	B	0.23	1/5672 (0.0%)	0.39	0/7644
1	C	0.17	0/5672	0.39	0/7644
1	D	0.16	0/5672	0.38	0/7644
All	All	0.19	2/22688 (0.0%)	0.39	0/30576

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	1097	VAL	CA-CB	10.67	1.60	1.54
1	A	361	THR	CA-C	5.18	1.59	1.52

There are no bond angle outliers.

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	5595	0	5684	121	0
1	B	5595	0	5684	123	0
1	C	5595	0	5684	104	0
1	D	5595	0	5684	95	0
2	A	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
All	All	22384	0	22736	431	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 10.

The worst 5 of 431 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:1010:GLN:HE22	1:D:1013:LYS:HE3	1.47	0.79
1:C:1002:ALA:HB1	1:C:1010:GLN:HG3	1.66	0.78
1:C:520:LEU:HD13	1:C:522:LEU:HD13	1.66	0.77
1:A:435:GLU:OE2	1:B:1007:ASN:ND2	2.17	0.77
1:A:547:LEU:HB3	1:A:567:LEU:HG	1.67	0.76

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	701/821 (85%)	645 (92%)	53 (8%)	3 (0%)	30	66
1	B	701/821 (85%)	650 (93%)	51 (7%)	0	100	100
1	C	701/821 (85%)	651 (93%)	46 (7%)	4 (1%)	21	58
1	D	701/821 (85%)	663 (95%)	36 (5%)	2 (0%)	36	71
All	All	2804/3284 (85%)	2609 (93%)	186 (7%)	9 (0%)	36	71

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	435	GLU
1	C	1002	ALA
1	C	1088	ALA
1	C	1097	VAL
1	D	1088	ALA

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	597/670 (89%)	596 (100%)	1 (0%)	87	85
1	B	597/670 (89%)	597 (100%)	0	100	100
1	C	597/670 (89%)	597 (100%)	0	100	100
1	D	597/670 (89%)	597 (100%)	0	100	100
All	All	2388/2680 (89%)	2387 (100%)	1 (0%)	100	100

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	779	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	910	GLN
1	B	1023	ASN
1	D	1023	ASN
1	C	1010	GLN
1	D	333	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](#)

Of 4 ligands modelled in this entry, 4 are 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	703/821 (85%)	-0.20	5 (0%) 84 71	149, 198, 248, 272	0
1	B	703/821 (85%)	-0.16	5 (0%) 84 71	138, 196, 265, 299	0
1	C	703/821 (85%)	-0.22	1 (0%) 92 87	148, 199, 255, 303	0
1	D	703/821 (85%)	-0.16	4 (0%) 85 73	146, 200, 252, 300	0
All	All	2812/3284 (85%)	-0.18	15 (0%) 87 75	138, 198, 256, 303	0

The worst 5 of 15 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	437	VAL	5.9
1	A	384	LEU	3.6
1	B	1065	LEU	2.9
1	A	380	THR	2.6
1	B	1022	LEU	2.5

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](#)

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($\text{\AA}^2$)	Q<0.9
2	MG	B	1201	1/1	0.40	0.17	140,140,140,140	0
2	MG	D	1201	1/1	0.61	0.20	125,125,125,125	0
2	MG	C	1201	1/1	0.70	0.17	177,177,177,177	0
2	MG	A	1201	1/1	0.75	0.14	149,149,149,149	0

6.5 Other polymers [i](#)

There are no such residues in this entry.