



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 5, 2026 – 03:42 PM UTC

PDB ID : 1URJ / pdb_00001urj
Title : Single stranded DNA-binding protein(ICP8) from Herpes simplex virus-1
Authors : Panjikar, S.; Mapelli, M.; Tucker, P.A.
Deposited on : 2003-10-30
Resolution : 3.00 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

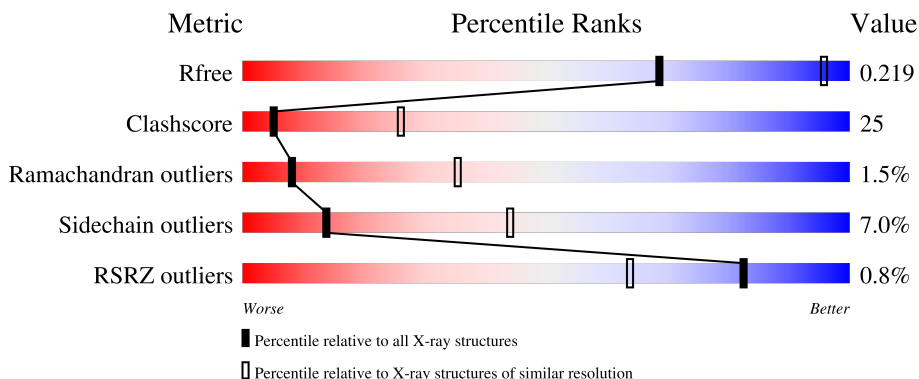
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.00 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	2672 (3.00-3.00)
Clashscore	190562	2977 (3.00-3.00)
Ramachandran outliers	187476	2877 (3.00-3.00)
Sidechain outliers	187428	2880 (3.00-3.00)
RSRZ outliers	180081	2671 (3.00-3.00)

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	1136	
1	B	1136	

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 15798 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 DNA-BINDING PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1029	Total	C	N	O	S	0	0	1
			7840	4968	1384	1444	44			
1	B	1031	Total	C	N	O	S	0	0	1
			7853	4975	1386	1448	44			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	254	SER	CYS	engineered mutation	UNP P04296
A	455	SER	CYS	engineered mutation	UNP P04296
B	254	SER	CYS	engineered mutation	UNP P04296
B	455	SER	CYS	engineered mutation	UNP P04296

- Molecule 2 is ZINC ION (CCD ID: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Zn	0	0
			1	1		
2	B	1	Total	Zn	0	0
			1	1		

- Molecule 3 is MERCURY (II) ION (CCD ID: HG) (formula: Hg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	4	Total	Hg	0	0
			4	4		
3	B	4	Total	Hg	0	0
			4	4		

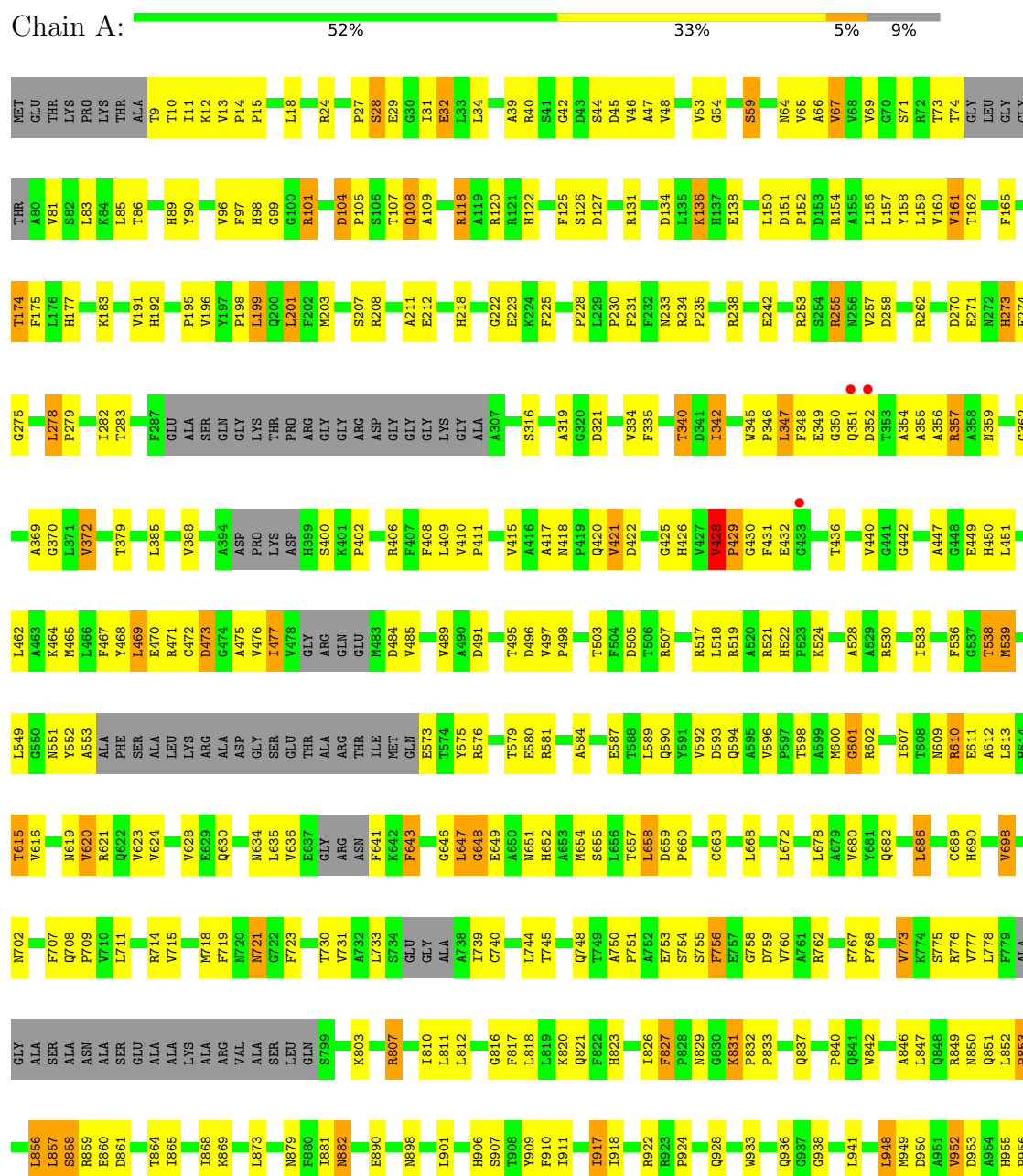
- Molecule 4 is water.

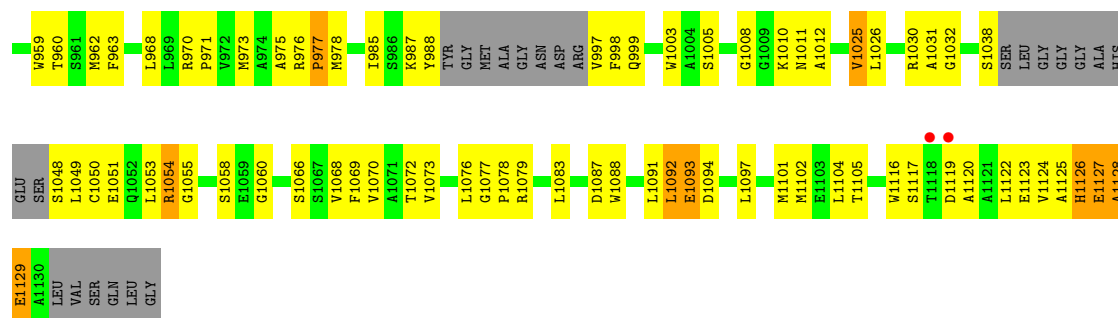
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	50	Total 50	O 50	0	0
4	B	45	Total 45	O 45	0	0

3 Residue-property plots

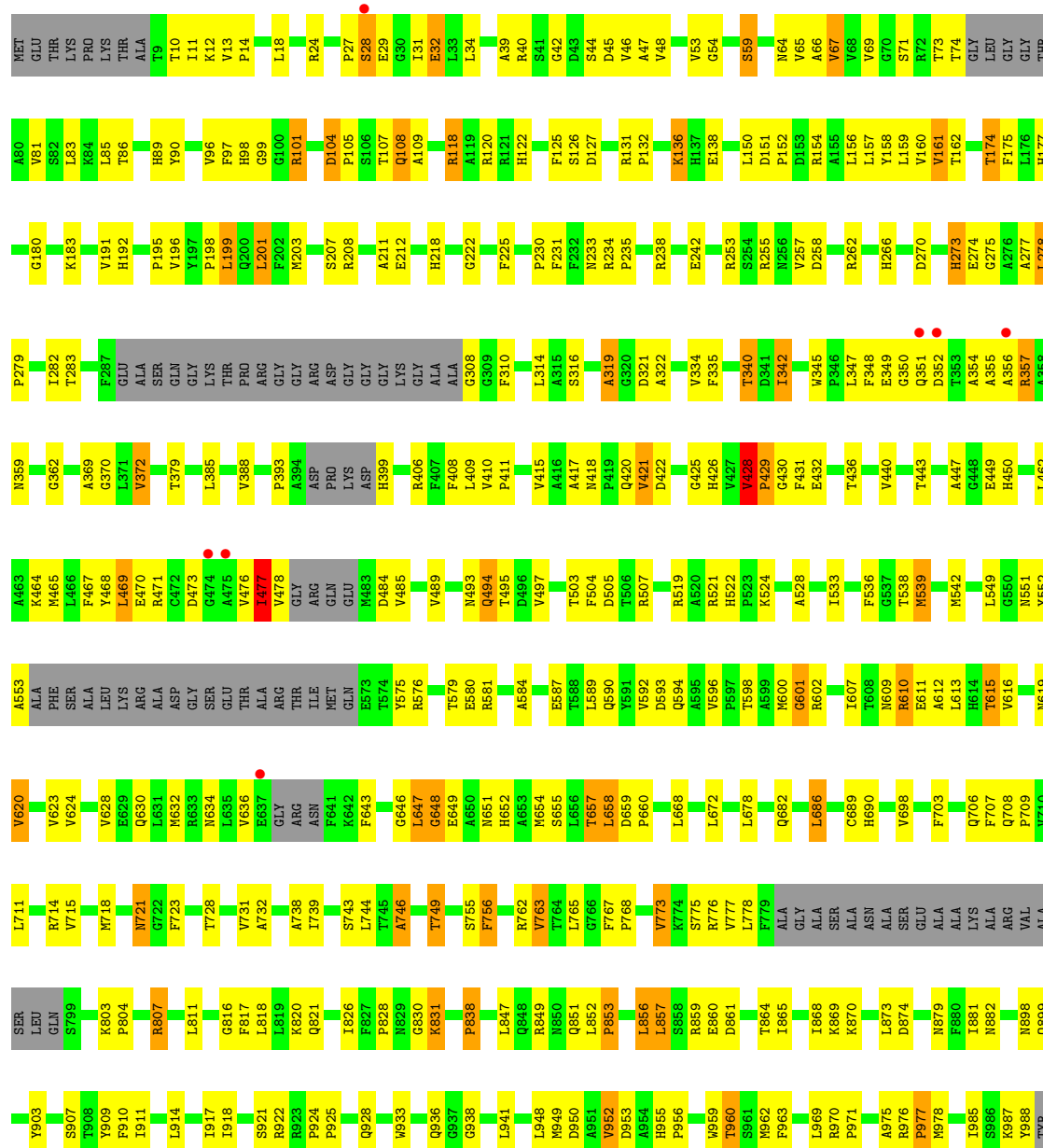
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: MAJOR DNA-BINDING PROTEIN





• Molecule 1: MAJOR DNA-BINDING PROTEIN



GLY		V1073		GLY		S1066	
MET				NET		S1067	
ALA		L1076		ALA		V1068	
GLY		G1077		GLY		F1069	
ASN		P1078		ASN		V1070	
ASP		R1079		ASP		A1071	
ARG				ARG		T1072	
V997		L1083		V997			
F998				F998			
Q999		D1087		Q999			
		W1088					
W1003				W1003			
		L1091					
G1008		L1092		G1008			
G1009		E1093		G1009			
K1010		D1094		K1010			
N1011				N1011			
A1012		L1097		A1012			
C1013				C1013			
P1014		M1101		P1014			
		M1102					
V1025		E1103		V1025			
L1026		L1104		L1026			
		T1105					
R1030				R1030			
A1031		A1108		A1031			
		L1109					
C1035				C1035			
		W1116					
S1038		S1117		S1038			
SER		T1118		SER			
LEU		D1119		LEU			
GLY		A1120		GLY			
GLY		L1122		GLY			
ALA		E1123		ALA			
HIS		V1124		HIS			
GLU		A1125		GLU			
		H1126					
SER		E1127		SER			
S1048		A1128		S1048			
L1049		E1129		L1049			
C1050				C1050			
E1051		A1130		E1051			
Q1052		LEU		Q1052			
L1053		VAL		L1053			
R1054		SER		R1054			
G1055		GLN		G1055			
I1056		LEU		I1056			
I1057		GLY		I1057			
S1058				S1058			
E1059				E1059			
G1060				G1060			
S1066				S1066			
S1067				S1067			
V1068				V1068			
F1069				F1069			
V1070				V1070			
A1071				A1071			
T1072				T1072			

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	100.91Å 145.37Å 162.03Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 3.00 20.00 – 3.00	Depositor EDS
% Data completeness (in resolution range)	98.0 (20.00-3.00) 97.6 (20.00-3.00)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.44 (at 3.00Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.235 , 0.286 0.217 , 0.219	Depositor DCC
R_{free} test set	1211 reflections (2.56%)	wwPDB-VP
Wilson B-factor (Å ²)	50.5	Xtriage
Anisotropy	0.403	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 33.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	15798	wwPDB-VP
Average B, all atoms (Å ²)	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.90% 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: HG, ZN

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	0.50	1/8009 (0.0%)	0.99	34/10878 (0.3%)
1	B	0.51	1/8023 (0.0%)	0.99	31/10898 (0.3%)
All	All	0.51	2/16032 (0.0%)	0.99	65/21776 (0.3%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1129	GLU	C-N	-5.36	1.25	1.33
1	B	1129	GLU	C-N	-5.06	1.26	1.33

The worst 5 of 65 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	357	ARG	N-CA-C	-12.39	97.87	113.23
1	B	357	ARG	N-CA-C	-12.37	97.89	113.23
1	A	827	PHE	CA-C-N	9.95	129.58	119.24
1	A	827	PHE	C-N-CA	9.95	129.58	119.24
1	B	428	VAL	CA-C-N	-8.37	109.38	119.84

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	7840	0	7742	392	0
1	B	7853	0	7752	385	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	4	0	0	0	0
3	B	4	0	0	0	0
4	A	50	0	0	6	0
4	B	45	0	0	4	0
All	All	15798	0	15494	769	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 25.

The worst 5 of 769 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:B:494:GLN:HE21	1:B:494:GLN:HA	1.02	1.16
1:B:715:VAL:HA	1:B:718:MET:HE3	1.37	1.07
1:A:283:THR:HG23	1:A:420:GLN:HE21	1.21	1.04
1:A:715:VAL:HA	1:A:718:MET:HE3	1.39	1.03
1:A:1119:ASP:HB2	1:A:1122:LEU:HG	1.38	1.03

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	1007/1136 (89%)	912 (91%)	81 (8%)	14 (1%)	9	36
1	B	1011/1136 (89%)	912 (90%)	82 (8%)	17 (2%)	7	32
All	All	2018/2272 (89%)	1824 (90%)	163 (8%)	31 (2%)	8	35

5 of 31 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	428	VAL
1	A	429	PRO
1	A	858	SER
1	A	1128	ALA
1	B	428	VAL

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	818/887 (92%)	760 (93%)	58 (7%)	13	43
1	B	819/887 (92%)	763 (93%)	56 (7%)	14	45
All	All	1637/1774 (92%)	1523 (93%)	114 (7%)	14	44

5 of 114 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1127	GLU
1	B	1025	VAL
1	B	278	LEU
1	B	977	PRO
1	B	749	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 48 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	218	HIS
1	B	630	GLN
1	B	266	HIS
1	B	450	HIS
1	B	677	ASN

5.3.3 RNA ⓘ

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 10 ligands modelled in this entry, 10 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	1029/1136 (90%)	-0.23	5 (0%) 87 72	13, 34, 68, 79	0
1	B	1031/1136 (90%)	-0.26	11 (1%) 78 57	12, 33, 66, 80	0
All	All	2060/2272 (90%)	-0.24	16 (0%) 82 64	12, 34, 67, 80	0

The worst 5 of 16 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	352	ASP	3.8
1	A	1118	THR	3.0
1	B	351	GLN	2.7
1	B	474	GLY	2.5
1	B	356	ALA	2.4

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
3	HG	A	2135	1/1	0.87	0.17	24,24,24,24	1
3	HG	A	2134	1/1	0.88	0.21	85,85,85,85	1
3	HG	B	2134	1/1	0.96	0.13	58,58,58,58	1
3	HG	B	2132	1/1	0.97	0.04	57,57,57,57	1
3	HG	A	2132	1/1	0.98	0.03	49,49,49,49	1
3	HG	B	2135	1/1	0.98	0.12	25,25,25,25	1
3	HG	B	2133	1/1	0.99	0.04	66,66,66,66	1
3	HG	A	2133	1/1	0.99	0.06	59,59,59,59	1
2	ZN	B	2131	1/1	0.99	0.02	29,29,29,29	0
2	ZN	A	2131	1/1	1.00	0.04	35,35,35,35	0

6.5 Other polymers [i](#)

There are no such residues in this entry.