



wwPDB X-ray Structure Validation Summary Report i

Dec 16, 2025 – 10:09 AM JST

PDB ID : 9WVF / pdb_00009wvf
Title : Crystal structure of HLA-A*11:01 in complex with KRAS G12S 10-mer peptide (VVVGASGVGK)
Authors : Jiali, Z.; Linlin, Z.
Deposited on : 2025-09-19
Resolution : 2.80 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 i 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](#) i) were used in the production of this report:

MolProbity : 4-5-2 with Phenix2.0
Xtriage (Phenix) : 2.0
EDS : 3.0
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
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) : 2.47

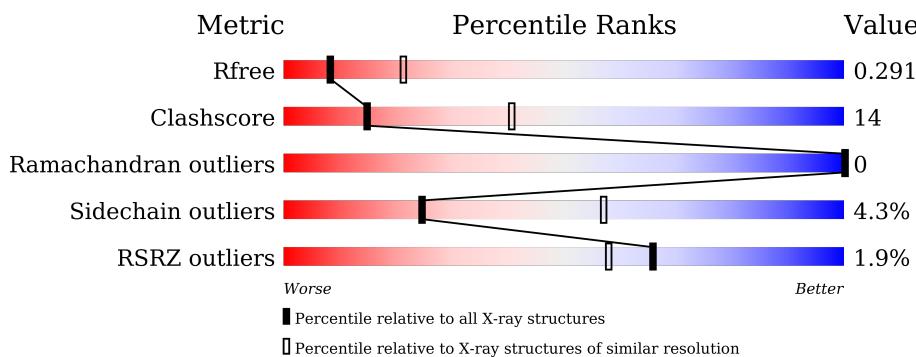
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 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}	164625	3657 (2.80-2.80)
Clashscore	180529	4123 (2.80-2.80)
Ramachandran outliers	177936	4071 (2.80-2.80)
Sidechain outliers	177891	4073 (2.80-2.80)
RSRZ outliers	164620	3659 (2.80-2.80)

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.



Continued on next page...

Continued from previous page...



2 Entry composition [\(i\)](#)

There are 4 unique types of molecules in this entry. The entry contains 12501 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 HLA class I histocompatibility antigen, A alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	275	Total	C	N	O	S	0	0	0
		2241	1393	408	431	9				
1	D	275	Total	C	N	O	S	0	0	0
		2235	1390	405	431	9				
1	G	275	Total	C	N	O	S	0	0	0
		2235	1390	405	431	9				
1	J	275	Total	C	N	O	S	0	0	0
		2229	1387	402	431	9				

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	initiating methionine	UNP P04439
A	9	TYR	PHE	variant	UNP P04439
A	90	ASP	ALA	variant	UNP P04439
A	105	PRO	SER	variant	UNP P04439
D	0	MET	-	initiating methionine	UNP P04439
D	9	TYR	PHE	variant	UNP P04439
D	90	ASP	ALA	variant	UNP P04439
D	105	PRO	SER	variant	UNP P04439
G	0	MET	-	initiating methionine	UNP P04439
G	9	TYR	PHE	variant	UNP P04439
G	90	ASP	ALA	variant	UNP P04439
G	105	PRO	SER	variant	UNP P04439
J	0	MET	-	initiating methionine	UNP P04439
J	9	TYR	PHE	variant	UNP P04439
J	90	ASP	ALA	variant	UNP P04439
J	105	PRO	SER	variant	UNP P04439

- Molecule 2 is a protein called Beta-2-microglobulin.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	B	99	Total	C	N	O	S		
			828	528	140	157	3	0	0
2	E	99	Total	C	N	O	S		
			821	523	140	155	3	0	0
2	H	99	Total	C	N	O	S		
			818	523	137	155	3	0	0
2	K	99	Total	C	N	O	S		
			821	523	140	155	3	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	0	MET	-	initiating methionine	UNP P61769
E	0	MET	-	initiating methionine	UNP P61769
H	0	MET	-	initiating methionine	UNP P61769
K	0	MET	-	initiating methionine	UNP P61769

- Molecule 3 is a protein called VAL-VAL-VAL-GLY-ALA-SER-GLY-VAL-GLY-LYS.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	10	Total	C	N	O			
			61	38	11	12	0	0	0
3	F	10	Total	C	N	O			
			61	38	11	12	0	0	0
3	I	10	Total	C	N	O			
			61	38	11	12	0	0	0
3	L	10	Total	C	N	O			
			61	38	11	12	0	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	8	Total O		
			8 8	0	0
4	B	4	Total O		
			4 4	0	0
4	D	1	Total O		
			1 1	0	0
4	E	2	Total O		
			2 2	0	0
4	G	6	Total O		
			6 6	0	0

Continued on next page...

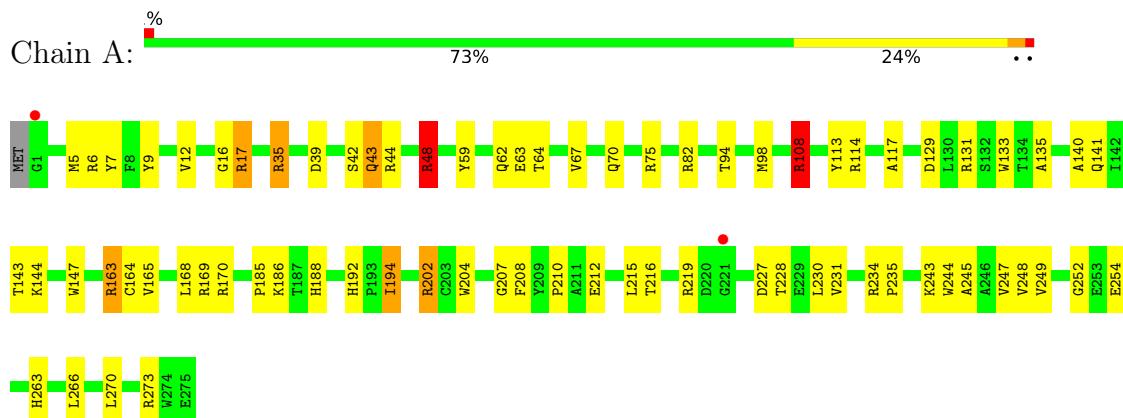
Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	H	4	Total O 4 4	0	0
4	J	3	Total O 3 3	0	0
4	K	1	Total O 1 1	0	0

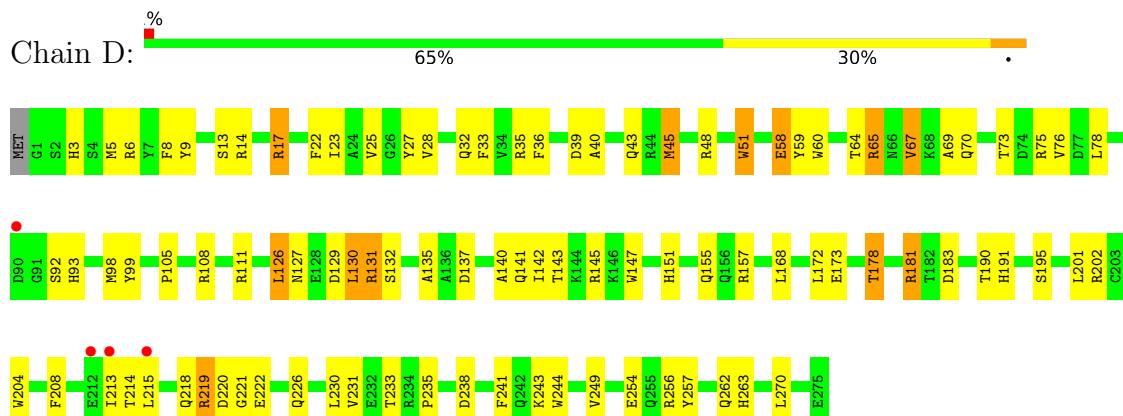
3 Residue-property plots

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.

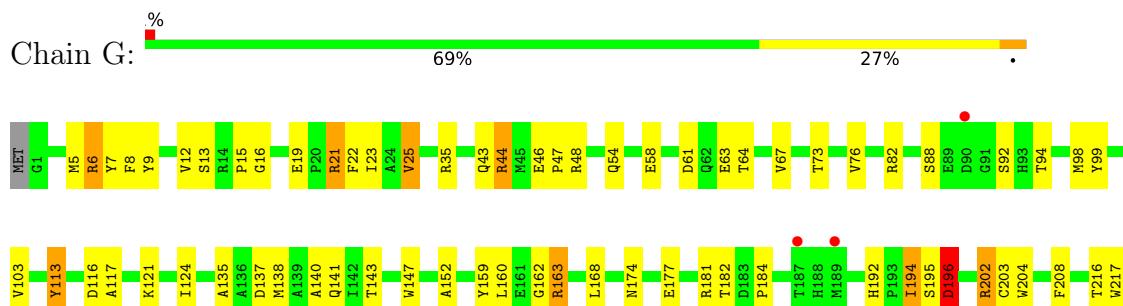
- Molecule 1: HLA class I histocompatibility antigen, A alpha chain



- Molecule 1: HLA class I histocompatibility antigen, A alpha chain



- Molecule 1: HLA class I histocompatibility antigen, A alpha chain





- Molecule 1: HLA class I histocompatibility antigen, A alpha chain



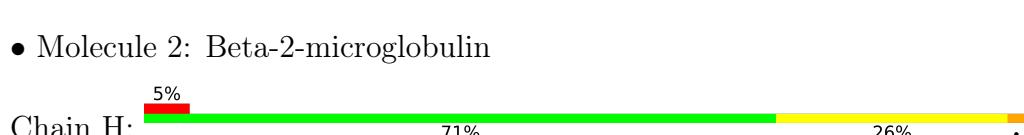
- Molecule 2: Beta-2-microglobulin



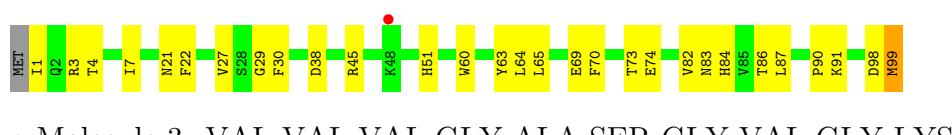
- Molecule 2: Beta-2-microglobulin



- Molecule 2: Beta-2-microglobulin



- Molecule 2: Beta-2-microglobulin

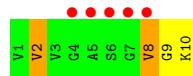




- Molecule 3: VAL-VAL-VAL-GLY-ALA-SER-GLY-VAL-GLY-LYS



- Molecule 3: VAL-VAL-VAL-GLY-ALA-SER-GLY-VAL-GLY-LYS



- Molecule 3: VAL-VAL-VAL-GLY-ALA-SER-GLY-VAL-GLY-LYS



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	118.86 Å 121.53 Å 132.56 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	44.79 – 2.80 44.79 – 2.80	Depositor EDS
% Data completeness (in resolution range)	99.7 (44.79-2.80) 99.7 (44.79-2.80)	Depositor EDS
R_{merge}	0.22	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle^1$	1.22 (at 2.81 Å)	Xtriage
Refinement program	PHENIX (1.20.1_4487: ???)	Depositor
R , R_{free}	0.265, 0.291 0.265, 0.291	Depositor DCC
R_{free} test set	2236 reflections (4.66%)	wwPDB-VP
Wilson B-factor (Å ²)	64.0	Xtriage
Anisotropy	0.599	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32, 45.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.000 for k,h,-l	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	12501	wwPDB-VP
Average B, all atoms (Å ²)	77.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 31.19 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.1536e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

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	0.84	0/2302	1.21	2/3125 (0.1%)
1	D	0.85	0/2296	1.19	0/3118
1	G	0.85	1/2296 (0.0%)	1.17	2/3118 (0.1%)
1	J	0.71	0/2290	1.05	4/3111 (0.1%)
2	B	0.84	0/851	1.15	0/1152
2	E	0.88	0/844	1.17	0/1143
2	H	0.88	0/841	1.18	0/1140
2	K	0.50	0/844	0.76	0/1143
3	C	0.87	0/60	1.32	0/78
3	F	0.90	0/60	1.33	0/78
3	I	0.89	0/60	1.29	0/78
3	L	0.37	0/60	0.63	0/78
All	All	0.81	1/12804 (0.0%)	1.14	8/17362 (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	A	0	11
1	D	0	8
1	G	0	7
1	J	0	5
2	B	0	2
2	E	0	1
2	H	0	1
All	All	0	35

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	G	184	PRO	C-O	-5.33	1.20	1.24

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	J	82	ARG	N-CA-C	-6.12	104.52	111.07
1	A	208	PHE	CA-CB-CG	5.68	119.48	113.80
1	A	194	ILE	N-CA-C	-5.67	106.94	111.81
1	J	105	PRO	N-CA-C	-5.27	106.20	113.53
1	J	173	GLU	N-CA-C	-5.10	107.44	113.97

There are no chirality outliers.

5 of 35 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	17	ARG	Sidechain
1	A	48	ARG	Sidechain
1	A	6	ARG	Sidechain
1	A	75	ARG	Sidechain
1	A	82	ARG	Sidechain

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	2241	0	2089	52	1
1	D	2235	0	2078	76	4
1	G	2235	0	2078	58	5
1	J	2229	0	2067	51	4
2	B	828	0	794	29	0
2	E	821	0	781	44	0
2	H	818	0	779	30	2
2	K	821	0	781	26	0
3	C	61	0	70	8	0
3	F	61	0	70	5	0
3	I	61	0	70	8	0
3	L	61	0	70	6	0
4	A	8	0	0	4	0
4	B	4	0	0	1	0
4	D	1	0	0	0	0
4	E	2	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	G	6	0	0	0	0
4	H	4	0	0	0	0
4	J	3	0	0	0	0
4	K	1	0	0	0	0
All	All	12501	0	11727	328	8

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 14.

The worst 5 of 328 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)
2:H:4:THR:HA	2:H:86:THR:HG21	1.48	0.94
1:A:5:MET:HE2	1:A:7:TYR:HE1	1.37	0.88
1:D:45:MET:H	1:D:64:THR:HG22	1.41	0.85
2:E:4:THR:HA	2:E:86:THR:HG21	1.59	0.85
3:L:6:SER:HB2	3:L:8:VAL:HG23	1.59	0.84

The worst 5 of 8 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:108:ARG:NH2	1:J:48:ARG:O[3_554]	1.34	0.86
1:G:256:ARG:NH1	1:J:256:ARG:NE[1_565]	1.52	0.68
1:G:256:ARG:CD	1:J:256:ARG:NH2[1_565]	1.61	0.59
1:G:256:ARG:CG	1:J:256:ARG:NH2[1_565]	1.87	0.33
1:D:173:GLU:OE2	1:G:177:GLU:OE2[3_444]	1.90	0.30

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	273/276 (99%)	262 (96%)	11 (4%)	0	100 100
1	D	273/276 (99%)	265 (97%)	8 (3%)	0	100 100
1	G	273/276 (99%)	263 (96%)	10 (4%)	0	100 100
1	J	273/276 (99%)	261 (96%)	12 (4%)	0	100 100
2	B	97/100 (97%)	92 (95%)	5 (5%)	0	100 100
2	E	97/100 (97%)	94 (97%)	3 (3%)	0	100 100
2	H	97/100 (97%)	95 (98%)	2 (2%)	0	100 100
2	K	97/100 (97%)	94 (97%)	3 (3%)	0	100 100
3	C	8/10 (80%)	7 (88%)	1 (12%)	0	100 100
3	F	8/10 (80%)	7 (88%)	1 (12%)	0	100 100
3	I	8/10 (80%)	5 (62%)	3 (38%)	0	100 100
3	L	8/10 (80%)	6 (75%)	2 (25%)	0	100 100
All	All	1512/1544 (98%)	1451 (96%)	61 (4%)	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	231/233 (99%)	221 (96%)	10 (4%)	25 57
1	D	230/233 (99%)	217 (94%)	13 (6%)	17 46
1	G	230/233 (99%)	218 (95%)	12 (5%)	19 50
1	J	229/233 (98%)	219 (96%)	10 (4%)	24 56
2	B	94/95 (99%)	91 (97%)	3 (3%)	34 68
2	E	92/95 (97%)	89 (97%)	3 (3%)	33 67
2	H	92/95 (97%)	90 (98%)	2 (2%)	47 79
2	K	92/95 (97%)	91 (99%)	1 (1%)	70 90
3	C	6/6 (100%)	6 (100%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
3	F	6/6 (100%)	6 (100%)	0	100 100
3	I	6/6 (100%)	4 (67%)	2 (33%)	0 0
3	L	6/6 (100%)	6 (100%)	0	100 100
All	All	1314/1336 (98%)	1258 (96%)	56 (4%)	25 57

5 of 56 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	E	74	GLU
2	K	99	MET
1	G	103	VAL
1	J	262	GLN
1	J	173	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 27 such sidechains are listed below:

Mol	Chain	Res	Type
1	G	62	GLN
1	G	262	GLN
1	J	180	GLN
1	G	86	ASN
2	H	8	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	275/276 (99%)	0.27	2 (0%)	84	79	45, 75, 103, 144
1	D	275/276 (99%)	0.40	4 (1%)	71	64	62, 86, 116, 153
1	G	275/276 (99%)	0.24	3 (1%)	77	71	56, 80, 104, 146
1	J	275/276 (99%)	0.25	2 (0%)	84	79	47, 77, 107, 140
2	B	99/100 (99%)	0.06	0	100	100	50, 61, 83, 111
2	E	99/100 (99%)	0.21	1 (1%)	79	73	49, 73, 103, 132
2	H	99/100 (99%)	0.33	5 (5%)	34	27	52, 75, 111, 119
2	K	99/100 (99%)	-0.30	1 (1%)	79	73	43, 51, 80, 106
3	C	10/10 (100%)	1.04	1 (10%)	14	11	75, 92, 120, 133
3	F	10/10 (100%)	1.16	2 (20%)	3	3	81, 95, 124, 130
3	I	10/10 (100%)	1.79	5 (50%)	0	1	68, 91, 124, 157
3	L	10/10 (100%)	1.29	3 (30%)	1	1	74, 95, 137, 141
All	All	1536/1544 (99%)	0.26	29 (1%)	66	58	43, 77, 111, 157

The worst 5 of 29 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	I	7	GLY	4.2
3	F	8	VAL	4.0
1	D	215	LEU	3.5
3	I	6	SER	3.4
1	G	187	THR	3.0

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.