



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 10:09 PM GMT

PDB ID : 1S7X
Title : Crystal structures of the murine class I major histocompatibility complex H-2Db in complex with LCMV-derived gp33 index peptide and three of its escape variants
Authors : Velloso, L.M.; Michaelsson, J.; Ljunggren, H.G.; Schneider, G.; Achour, A.
Deposited on : 2004-01-30
Resolution : 2.41 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
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>
with specific help available everywhere you see the ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

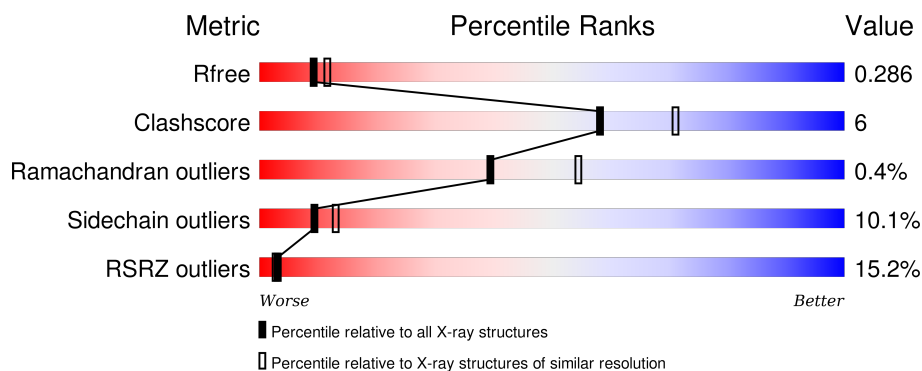
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.41 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}	91344	3386 (2.44-2.40)
Clashscore	102246	3897 (2.44-2.40)
Ramachandran outliers	100387	3837 (2.44-2.40)
Sidechain outliers	100360	3838 (2.44-2.40)
RSRZ outliers	91569	3396 (2.44-2.40)

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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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	338	<div> <div>13%</div> <div>66%</div> <div>13%</div> <div>•</div> <div>18%</div> </div>
1	D	338	<div> <div>11%</div> <div>68%</div> <div>11%</div> <div>••</div> <div>18%</div> </div>
1	G	338	<div> <div>17%</div> <div>64%</div> <div>16%</div> <div>•</div> <div>18%</div> </div>
1	J	338	<div> <div>20%</div> <div>67%</div> <div>13%</div> <div>•</div> <div>18%</div> </div>
2	B	99	<div> <div>2%</div> <div>81%</div> <div>16%</div> <div>•</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	E	99	<div><div></div><div>2%</div><div>77%</div><div>22%</div><div></div></div>
2	H	99	<div><div></div><div>2%</div><div>82%</div><div>18%</div><div></div></div>
2	K	99	<div><div></div><div>13%</div><div>75%</div><div>22%</div><div></div></div>
3	C	9	<div><div></div><div></div><div>89%</div><div>11%</div><div></div></div>
3	F	9	<div><div></div><div>11%</div><div>89%</div><div>11%</div><div></div></div>
3	I	9	<div><div></div><div>22%</div><div>67%</div><div>33%</div><div></div></div>
3	L	9	<div><div></div><div>89%</div><div>89%</div><div>11%</div><div></div></div>

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 13420 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 H-2 class I histocompatibility antigen, D-B alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	276	Total	C	N	O	S	0	0	0
			2264	1430	400	425	9			
1	D	276	Total	C	N	O	S	0	0	0
			2264	1430	400	425	9			
1	G	276	Total	C	N	O	S	0	0	0
			2264	1430	400	425	9			
1	J	276	Total	C	N	O	S	0	0	0
			2264	1430	400	425	9			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	99	Total	C	N	O	S	0	0	0
			818	523	138	151	6			
2	E	99	Total	C	N	O	S	0	0	0
			818	523	138	151	6			
2	H	99	Total	C	N	O	S	0	0	0
			818	523	138	151	6			
2	K	99	Total	C	N	O	S	0	0	0
			818	523	138	151	6			

- Molecule 3 is a protein called Glycoprotein 9-residue peptide.

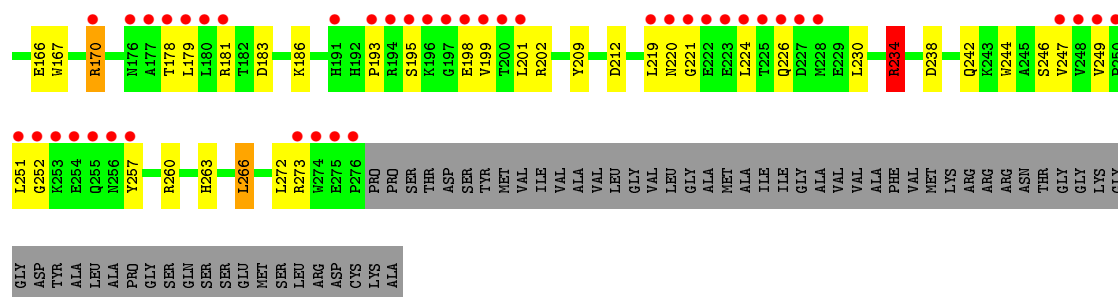
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	9	Total	C	N	O	S	0	0	0
			72	48	11	12	1			
3	F	9	Total	C	N	O	S	0	0	0
			72	48	11	12	1			
3	I	9	Total	C	N	O	S	0	0	0
			72	48	11	12	1			
3	L	9	Total	C	N	O	S	0	0	0
			72	48	11	12	1			

There are 8 discrepancies between the modelled and reference sequences:

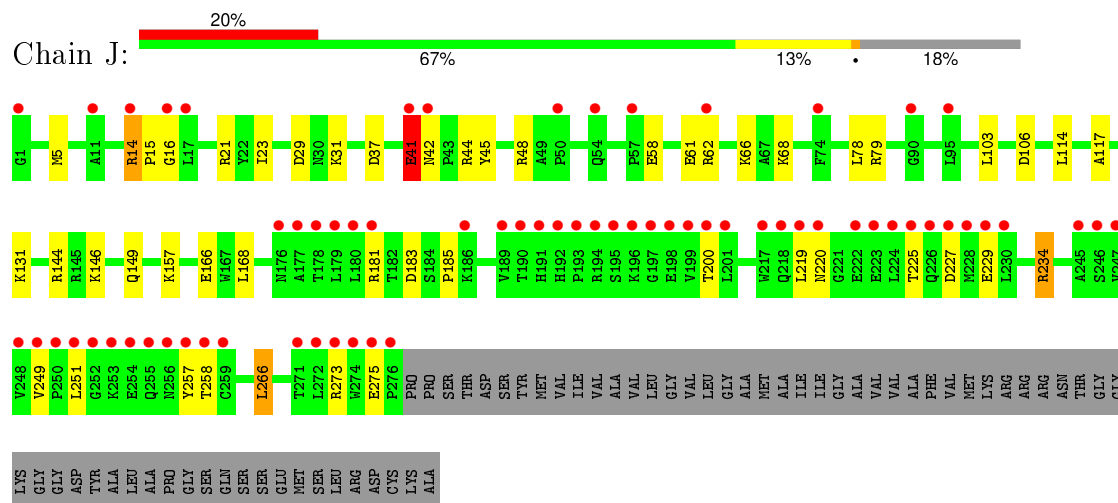
Chain	Residue	Modelled	Actual	Comment	Reference
C	4	PHE	TYR	ENGINEERED	UNP P07399
C	9	MET	CYS	SEE REMARK 999	UNP P07399
F	4	PHE	TYR	ENGINEERED	UNP P07399
F	9	MET	CYS	SEE REMARK 999	UNP P07399
I	4	PHE	TYR	ENGINEERED	UNP P07399
I	9	MET	CYS	SEE REMARK 999	UNP P07399
L	4	PHE	TYR	ENGINEERED	UNP P07399
L	9	MET	CYS	SEE REMARK 999	UNP P07399

- Molecule 4 is water.

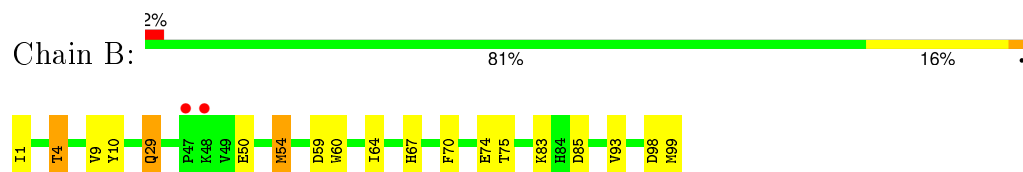
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	107	Total O 107 107	0	0
4	B	91	Total O 91 91	0	0
4	C	2	Total O 2 2	0	0
4	D	126	Total O 126 126	0	0
4	E	74	Total O 74 74	0	0
4	F	5	Total O 5 5	0	0
4	G	145	Total O 145 145	0	0
4	H	78	Total O 78 78	0	0
4	I	4	Total O 4 4	0	0
4	J	112	Total O 112 112	0	0
4	K	53	Total O 53 53	0	0
4	L	7	Total O 7 7	0	0



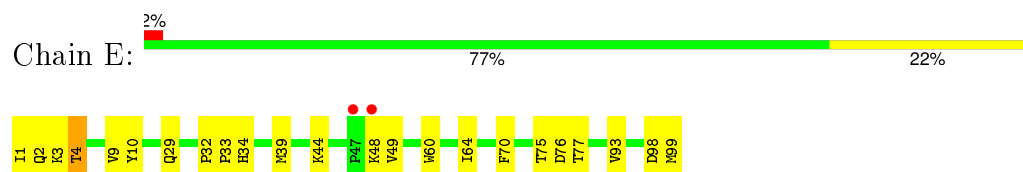
- Molecule 1: H-2 class I histocompatibility antigen, D-B alpha chain



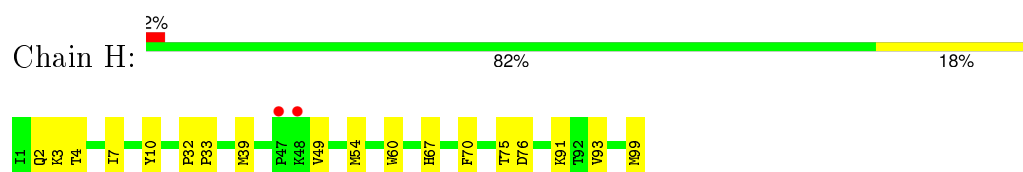
- Molecule 2: Beta-2-microglobulin



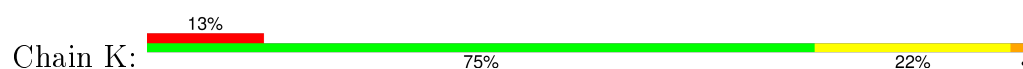
- Molecule 2: Beta-2-microglobulin

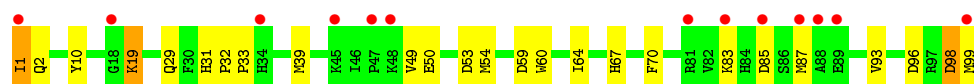


- Molecule 2: Beta-2-microglobulin



- Molecule 2: Beta-2-microglobulin

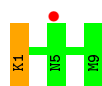
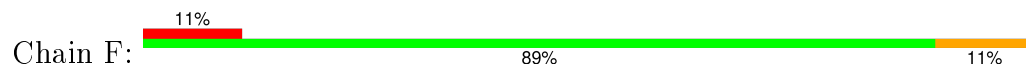




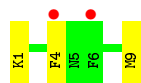
- Molecule 3: Glycoprotein 9-residue peptide



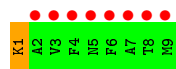
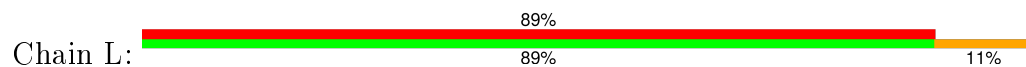
- Molecule 3: Glycoprotein 9-residue peptide



- Molecule 3: Glycoprotein 9-residue peptide



- Molecule 3: Glycoprotein 9-residue peptide



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	91.90Å 123.32Å 100.20Å 90.00° 102.82° 90.00°	Depositor
Resolution (Å)	19.80 – 2.41 31.93 – 2.41	Depositor EDS
% Data completeness (in resolution range)	(Not available) (19.80-2.41) 98.9 (31.93-2.41)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.55 (at 2.42Å)	Xtriage
Refinement program	REFMAC 5.1.24	Depositor
R, R_{free}	0.203 , 0.254 0.243 , 0.286	Depositor DCC
R_{free} test set	2033 reflections (2.45%)	DCC
Wilson B-factor (Å ²)	37.7	Xtriage
Anisotropy	0.082	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.39 , 57.7	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	2 of 82996 reflections (0.002%)	Xtriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	13420	wwPDB-VP
Average B, all atoms (Å ²)	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 13.69% 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, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

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	0.46	0/2331	0.74	6/3166 (0.2%)
1	D	0.51	0/2331	0.78	6/3166 (0.2%)
1	G	0.53	1/2331 (0.0%)	0.78	9/3166 (0.3%)
1	J	0.58	1/2331 (0.0%)	0.76	6/3166 (0.2%)
2	B	0.58	0/844	0.79	2/1146 (0.2%)
2	E	0.58	0/844	0.77	2/1146 (0.2%)
2	H	0.59	0/844	0.78	1/1146 (0.1%)
2	K	0.50	0/844	0.78	4/1146 (0.3%)
3	C	0.76	0/73	0.70	0/95
3	F	0.84	0/73	0.77	0/95
3	I	1.00	1/73 (1.4%)	0.87	0/95
3	L	0.72	0/73	0.67	0/95
All	All	0.54	3/12992 (0.0%)	0.77	36/17628 (0.2%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	J	41	GLU	C-N	17.52	1.74	1.34
1	G	146	LYS	CE-NZ	6.06	1.64	1.49
3	I	9	MET	CG-SD	-5.05	1.68	1.81

The worst 5 of 36 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	J	41	GLU	CA-C-N	-9.36	96.60	117.20
1	J	41	GLU	O-C-N	7.65	134.94	122.70
1	A	227	ASP	CB-CG-OD2	6.43	124.09	118.30
1	A	129	ASP	CB-CG-OD2	6.42	124.07	118.30
1	G	29	ASP	CB-CG-OD2	6.18	123.86	118.30

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	2264	0	2136	19	0
1	D	2264	0	2136	30	0
1	G	2264	0	2136	25	0
1	J	2264	0	2135	24	0
2	B	818	0	791	13	0
2	E	818	0	791	18	0
2	H	818	0	791	15	0
2	K	818	0	791	14	0
3	C	72	0	74	0	0
3	F	72	0	74	1	0
3	I	72	0	74	0	0
3	L	72	0	74	3	0
4	A	107	0	0	2	0
4	B	91	0	0	6	0
4	C	2	0	0	0	0
4	D	126	0	0	3	0
4	E	74	0	0	4	0
4	F	5	0	0	0	0
4	G	145	0	0	6	0
4	H	78	0	0	2	0
4	I	4	0	0	0	0
4	J	112	0	0	4	0
4	K	53	0	0	1	0
4	L	7	0	0	1	0
All	All	13420	0	12003	144	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 6.

The worst 5 of 144 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:J:41:GLU:C	1:J:42:ASN:N	1.74	1.40
1:J:41:GLU:OE2	1:J:42:ASN:N	1.80	1.14
2:H:99:MET:C	4:H:102:HOH:O	1.88	1.12

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:99:MET:CG	4:B:180:HOH:O	2.01	1.07
1:D:14:ARG:HG2	1:D:14:ARG:HH11	0.92	1.05

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	274/338 (81%)	258 (94%)	14 (5%)	2 (1%)	26	37
1	D	274/338 (81%)	263 (96%)	9 (3%)	2 (1%)	26	37
1	G	274/338 (81%)	253 (92%)	20 (7%)	1 (0%)	39	54
1	J	274/338 (81%)	260 (95%)	13 (5%)	1 (0%)	39	54
2	B	97/99 (98%)	94 (97%)	3 (3%)	0	100	100
2	E	97/99 (98%)	95 (98%)	2 (2%)	0	100	100
2	H	97/99 (98%)	95 (98%)	2 (2%)	0	100	100
2	K	97/99 (98%)	94 (97%)	3 (3%)	0	100	100
3	C	7/9 (78%)	6 (86%)	1 (14%)	0	100	100
3	F	7/9 (78%)	6 (86%)	1 (14%)	0	100	100
3	I	7/9 (78%)	6 (86%)	1 (14%)	0	100	100
3	L	7/9 (78%)	6 (86%)	1 (14%)	0	100	100
All	All	1512/1784 (85%)	1436 (95%)	70 (5%)	6 (0%)	39	54

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	227	ASP
1	G	221	GLY
1	D	16	GLY

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	195	SER
1	J	41	GLU

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	234/280 (84%)	211 (90%)	23 (10%)	10	14
1	D	234/280 (84%)	211 (90%)	23 (10%)	10	14
1	G	234/280 (84%)	210 (90%)	24 (10%)	9	12
1	J	234/280 (84%)	207 (88%)	27 (12%)	7	9
2	B	93/94 (99%)	84 (90%)	9 (10%)	10	14
2	E	93/94 (99%)	84 (90%)	9 (10%)	10	14
2	H	93/94 (99%)	88 (95%)	5 (5%)	27	42
2	K	93/94 (99%)	83 (89%)	10 (11%)	8	10
3	C	7/7 (100%)	6 (86%)	1 (14%)	4	4
3	F	7/7 (100%)	6 (86%)	1 (14%)	4	4
3	I	7/7 (100%)	5 (71%)	2 (29%)	0	0
3	L	7/7 (100%)	6 (86%)	1 (14%)	4	4
All	All	1336/1524 (88%)	1201 (90%)	135 (10%)	9	13

5 of 135 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	E	64	ILE
1	G	114	LEU
2	K	1	ILE
2	E	70	PHE
1	G	44	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 22 such sidechains are listed below:

Mol	Chain	Res	Type
1	G	54	GLN
2	H	67	HIS
2	K	67	HIS
1	G	97	GLN
1	G	220	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 carbohydrates 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 ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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	276/338 (81%)	0.97	43 (15%) 3 2	13, 30, 53, 61	0
1	D	276/338 (81%)	0.74	37 (13%) 4 4	10, 25, 51, 60	0
1	G	276/338 (81%)	1.02	56 (20%) 1 1	8, 26, 53, 61	0
1	J	276/338 (81%)	1.34	68 (24%) 1 1	10, 29, 62, 65	0
2	B	99/99 (100%)	0.23	2 (2%) 68 67	13, 21, 29, 32	0
2	E	99/99 (100%)	0.21	2 (2%) 68 67	11, 21, 28, 32	0
2	H	99/99 (100%)	0.12	2 (2%) 68 67	11, 21, 29, 34	0
2	K	99/99 (100%)	0.80	13 (13%) 5 4	14, 32, 41, 44	0
3	C	9/9 (100%)	0.92	0 100 100	23, 26, 32, 34	0
3	F	9/9 (100%)	0.95	1 (11%) 7 7	18, 21, 25, 29	0
3	I	9/9 (100%)	1.32	2 (22%) 1 1	13, 20, 28, 31	0
3	L	9/9 (100%)	2.96	8 (88%) 0 0	34, 37, 45, 51	0
All	All	1536/1784 (86%)	0.85	234 (15%) 3 3	8, 26, 56, 65	0

The worst 5 of 234 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	177	ALA	10.5
1	J	177	ALA	10.0
1	A	177	ALA	9.5
1	J	179	LEU	8.3
1	J	249	VAL	7.7

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates [i](#)

There are no carbohydrates 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.