



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 01:20 PM GMT

PDB ID : 3TBY  
Title : CRYSTAL STRUCTURE OF THE MURINE CLASS I MAJOR HISTO-COMPATIBILITY COMPLEX H-2DB IN COMPLEX WITH THE LCMV-DERIVED GP33 ALTERED PEPTIDE ligand (V3P, Y4F)  
Authors : Duru, A.D.; Allerbring, E.B.; Uchtenhagen, H.; Mazumdar, P.A.; Badia-Martinez, D.; Madhurantakam, C.; Sandalova, T.; Nygren, P.; Achour, A.  
Deposited on : 2011-08-08  
Resolution : 2.50 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](mailto: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

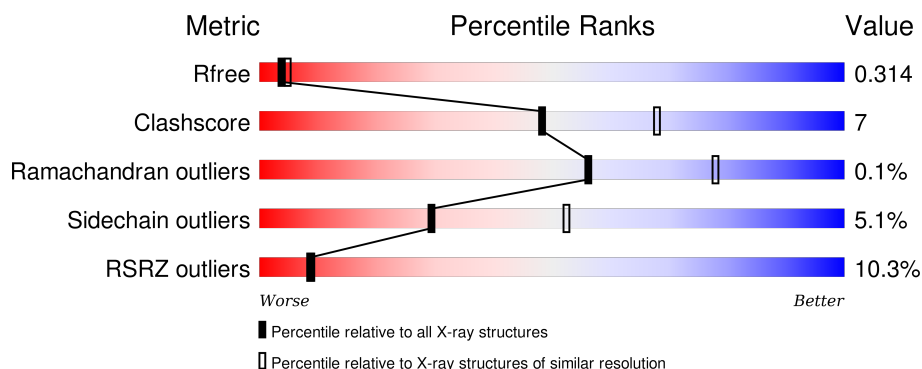
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 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))
$R_{free}$	91344	3553 (2.50-2.50)
Clashscore	102246	4242 (2.50-2.50)
Ramachandran outliers	100387	4156 (2.50-2.50)
Sidechain outliers	100360	4158 (2.50-2.50)
RSRZ outliers	91569	3562 (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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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	 10% 64% 14% • 21%
1	D	338	 9% 67% 11% • 21%
1	G	338	 11% 64% 14% • 22%
1	J	338	 13% 65% 12% • 22%
2	B	99	 3% 76% 22% •

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	E	99	<div><div></div><div>2%69%28%</div><div></div></div>
2	H	99	<div><div></div><div>%86%13%</div><div></div></div>
2	K	99	<div><div></div><div>4%81%17%</div><div></div></div>
3	C	9	<div><div></div><div>67%33%</div><div></div></div>
3	F	9	<div><div></div><div>56%44%</div><div></div></div>
3	I	9	<div><div></div><div>56%44%</div><div></div></div>
3	L	9	<div><div></div><div>22%67%11%</div><div></div></div>

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 12496 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	266	Total	C	N	O	S	0	0	0
			2187	1383	386	409	9			
1	D	266	Total	C	N	O	S	0	0	0
			2187	1383	386	409	9			
1	G	264	Total	C	N	O	S	0	0	0
			2175	1376	384	406	9			
1	J	262	Total	C	N	O	S	0	0	0
			2159	1366	381	403	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
			817	521	137	152	7			
2	E	99	Total	C	N	O	S	0	0	0
			817	521	137	152	7			
2	H	99	Total	C	N	O	S	0	0	0
			817	521	137	152	7			
2	K	99	Total	C	N	O	S	0	0	0
			817	521	137	152	7			

- Molecule 3 is a protein called GLYCOPROTEIN GPC.

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			

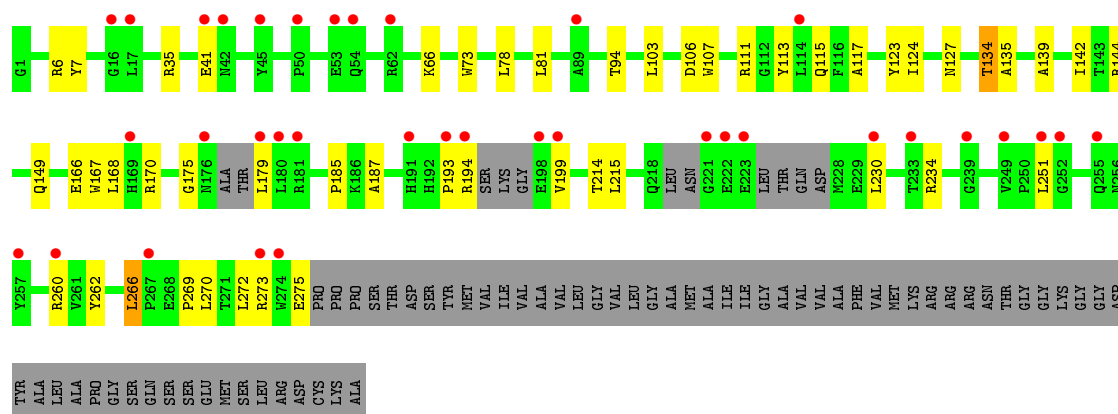
There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	3	PRO	VAL	ENGINEERED MUTATION	UNP P07399
C	4	PHE	TYR	ENGINEERED MUTATION	UNP P07399
C	9	MET	CYS	ENGINEERED MUTATION	UNP P07399
F	3	PRO	VAL	ENGINEERED MUTATION	UNP P07399
F	4	PHE	TYR	ENGINEERED MUTATION	UNP P07399
F	9	MET	CYS	ENGINEERED MUTATION	UNP P07399
I	3	PRO	VAL	ENGINEERED MUTATION	UNP P07399
I	4	PHE	TYR	ENGINEERED MUTATION	UNP P07399
I	9	MET	CYS	ENGINEERED MUTATION	UNP P07399
L	3	PRO	VAL	ENGINEERED MUTATION	UNP P07399
L	4	PHE	TYR	ENGINEERED MUTATION	UNP P07399
L	9	MET	CYS	ENGINEERED MUTATION	UNP P07399

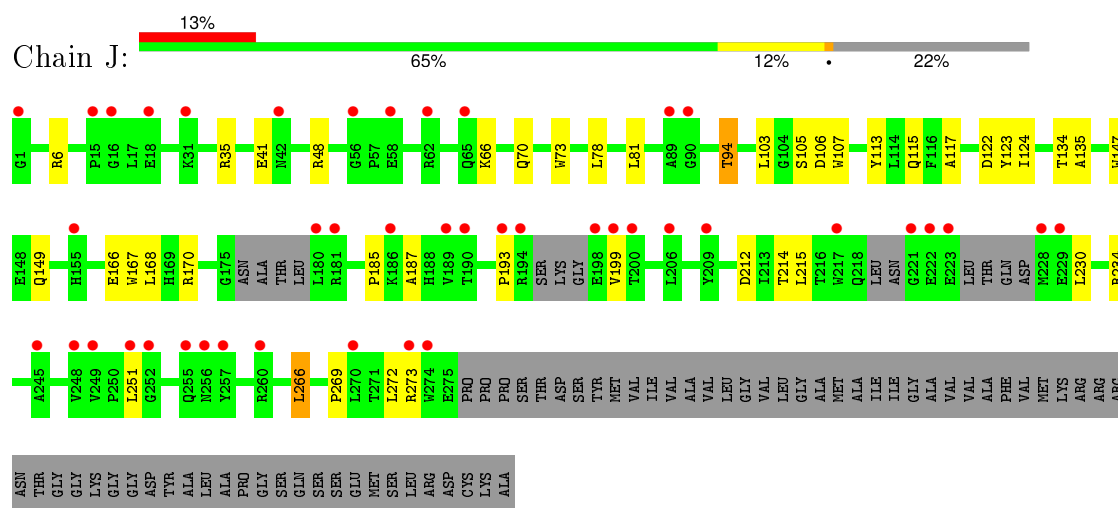
- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	38	Total O 38 38	0	0
4	B	26	Total O 26 26	0	0
4	C	2	Total O 2 2	0	0
4	D	35	Total O 35 35	0	0
4	E	24	Total O 24 24	0	0
4	F	1	Total O 1 1	0	0
4	G	49	Total O 49 49	0	0
4	H	19	Total O 19 19	0	0
4	I	2	Total O 2 2	0	0
4	J	22	Total O 22 22	0	0
4	K	13	Total O 13 13	0	0
4	L	1	Total O 1 1	0	0

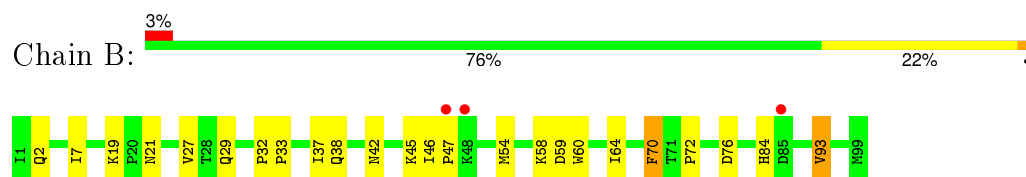




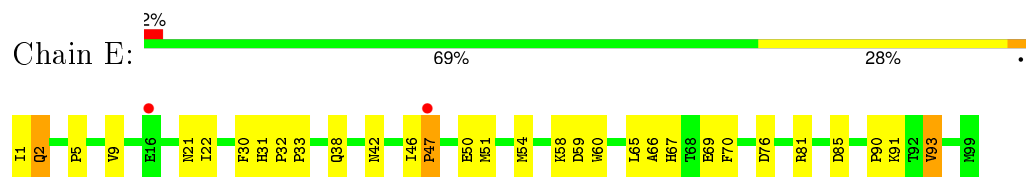
- Molecule 1: H-2 class I histocompatibility antigen, D-B alpha chain



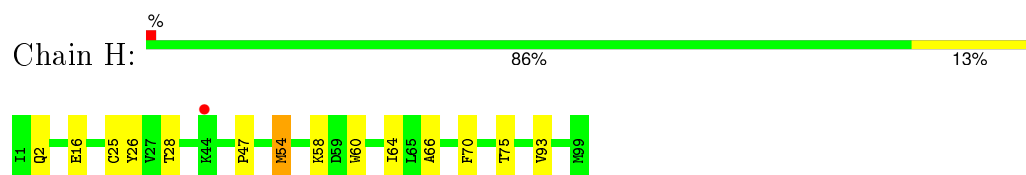
- Molecule 2: Beta-2-microglobulin



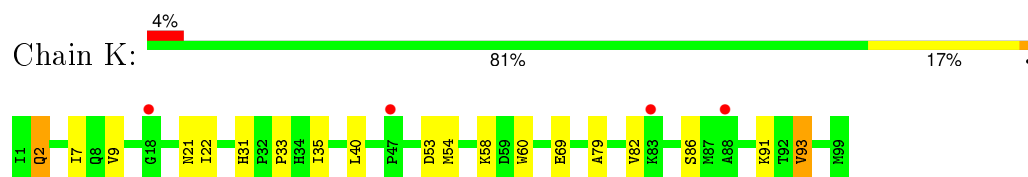
- Molecule 2: Beta-2-microglobulin



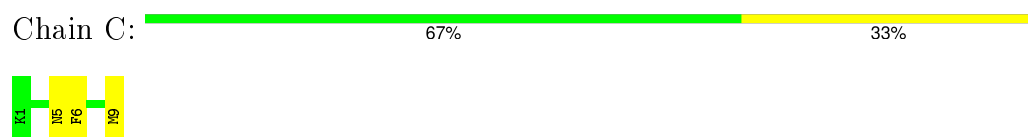
- Molecule 2: Beta-2-microglobulin



## ● Molecule 2: Beta-2-microglobulin



## ● Molecule 3: GLYCOPROTEIN GPC



## ● Molecule 3: GLYCOPROTEIN GPC



## ● Molecule 3: GLYCOPROTEIN GPC



## ● Molecule 3: GLYCOPROTEIN GPC





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	92.82Å 124.27Å 99.65Å 90.00° 103.29° 90.00°	Depositor
Resolution (Å)	20.00 – 2.50 19.97 – 2.50	Depositor EDS
% Data completeness (in resolution range)	97.9 (20.00-2.50) 97.9 (19.97-2.50)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.91 (at 2.50Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, $R_{free}$	0.284 , 0.315 0.284 , 0.314	Depositor DCC
$R_{free}$ test set	3747 reflections (5.31%)	DCC
Wilson B-factor (Å <sup>2</sup> )	40.2	Xtriage
Anisotropy	0.044	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 38.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.44$ , $\langle L^2 \rangle = 0.26$	Xtriage
Outliers	1 of 74349 reflections (0.001%)	Xtriage
$F_o, F_c$ correlation	0.86	EDS
Total number of atoms	12496	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 11.22% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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	0.50	0/2250	0.59	0/3052
1	D	0.48	0/2250	0.60	0/3052
1	G	0.50	0/2237	0.58	0/3032
1	J	0.49	0/2221	0.59	0/3010
2	B	0.54	0/843	0.68	0/1144
2	E	0.58	0/843	0.73	2/1144 (0.2%)
2	H	0.60	0/843	0.72	0/1144
2	K	0.61	0/843	0.72	0/1144
3	C	0.77	0/74	0.85	0/97
3	F	0.67	0/74	0.84	0/97
3	I	0.68	0/74	0.87	0/97
3	L	0.82	0/74	0.69	0/97
All	All	0.53	0/12626	0.63	2/17110 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	E	47	PRO	N-CA-C	7.36	131.25	112.10
2	E	47	PRO	CB-CA-C	-5.54	98.16	112.00

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	2187	0	2049	32	0
1	D	2187	0	2049	32	0
1	G	2175	0	2036	28	0
1	J	2159	0	2019	27	0
2	B	817	0	785	20	0
2	E	817	0	785	20	0
2	H	817	0	785	9	0
2	K	817	0	785	18	0
3	C	72	0	72	5	0
3	F	72	0	72	3	0
3	I	72	0	72	3	0
3	L	72	0	72	7	0
4	A	38	0	0	2	0
4	B	26	0	0	1	0
4	C	2	0	0	0	0
4	D	35	0	0	6	0
4	E	24	0	0	1	0
4	F	1	0	0	0	0
4	G	49	0	0	3	0
4	H	19	0	0	0	0
4	I	2	0	0	0	0
4	J	22	0	0	1	0
4	K	13	0	0	0	0
4	L	1	0	0	0	0
All	All	12496	0	11581	174	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 7.

The worst 5 of 174 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:E:22:ILE:HD13	2:E:69:GLU:HA	1.68	0.74
1:J:66:LYS:HG2	3:L:4:PHE:CE1	2.26	0.70
2:E:81:ARG:HD3	2:E:90:PRO:HB2	1.74	0.69
3:L:2:ALA:HB1	3:L:3:PRO:HD2	1.74	0.68
1:G:35:ARG:NH2	2:H:54:MET:O	2.25	0.68

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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	258/338 (76%)	242 (94%)	16 (6%)	0	100	100
1	D	258/338 (76%)	242 (94%)	16 (6%)	0	100	100
1	G	254/338 (75%)	237 (93%)	17 (7%)	0	100	100
1	J	252/338 (75%)	236 (94%)	16 (6%)	0	100	100
2	B	97/99 (98%)	94 (97%)	3 (3%)	0	100	100
2	E	97/99 (98%)	93 (96%)	4 (4%)	0	100	100
2	H	97/99 (98%)	93 (96%)	3 (3%)	1 (1%)	19	34
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%)	5 (71%)	2 (29%)	0	100	100
All	All	1438/1784 (81%)	1354 (94%)	83 (6%)	1 (0%)	56	78

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	H	47	PRO

### 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	224/280 (80%)	214 (96%)	10 (4%)	34	59
1	D	224/280 (80%)	216 (96%)	8 (4%)	42	69
1	G	223/280 (80%)	212 (95%)	11 (5%)	31	55
1	J	221/280 (79%)	210 (95%)	11 (5%)	30	53
2	B	93/94 (99%)	88 (95%)	5 (5%)	27	49
2	E	93/94 (99%)	87 (94%)	6 (6%)	21	39
2	H	93/94 (99%)	86 (92%)	7 (8%)	17	31
2	K	93/94 (99%)	89 (96%)	4 (4%)	35	61
3	C	7/7 (100%)	7 (100%)	0	100	100
3	F	7/7 (100%)	6 (86%)	1 (14%)	4	7
3	I	7/7 (100%)	6 (86%)	1 (14%)	4	7
3	L	7/7 (100%)	5 (71%)	2 (29%)	0	0
All	All	1292/1524 (85%)	1226 (95%)	66 (5%)	29	52

5 of 66 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	F	4	PHE
1	G	166	GLU
2	K	2	GLN
1	G	41	GLU
1	G	115	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 32 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	255	GLN
2	E	38	GLN
2	K	2	GLN
2	E	31	HIS
3	F	5	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 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 [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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	266/338 (78%)	0.80	34 (12%) 5 5	18, 38, 73, 109	0
1	D	266/338 (78%)	0.75	31 (11%) 6 6	17, 38, 68, 124	0
1	G	264/338 (78%)	0.85	36 (13%) 4 4	17, 37, 76, 101	0
1	J	262/338 (77%)	0.90	43 (16%) 2 2	16, 41, 77, 110	0
2	B	99/99 (100%)	0.36	3 (3%) 54 59	19, 33, 44, 53	0
2	E	99/99 (100%)	0.39	2 (2%) 68 72	21, 33, 46, 48	0
2	H	99/99 (100%)	0.24	1 (1%) 84 86	17, 32, 43, 51	0
2	K	99/99 (100%)	0.62	4 (4%) 42 47	22, 37, 46, 51	0
3	C	9/9 (100%)	0.29	0 100 100	18, 25, 30, 33	0
3	F	9/9 (100%)	-0.06	0 100 100	18, 21, 26, 30	0
3	I	9/9 (100%)	0.42	0 100 100	21, 23, 30, 34	0
3	L	9/9 (100%)	0.36	0 100 100	22, 24, 26, 26	0
All	All	1490/1784 (83%)	0.70	154 (10%) 9 9	16, 36, 70, 124	0

The worst 5 of 154 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	177	ALA	8.2
1	A	251	LEU	6.8
1	G	180	LEU	6.7
1	D	178	THR	6.4
1	A	180	LEU	5.7

### 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 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.