



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 08:16 AM GMT

PDB ID : 3DXA  
Title : Crystal Structure of the DM1 TCR in complex with HLA-B\*4405 and decamer EBV antigen  
Authors : Archbold, J.K.; Macdonald, W.A.; Gras, S.; Rossjohn, J.  
Deposited on : 2008-07-23  
Resolution : 3.50 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

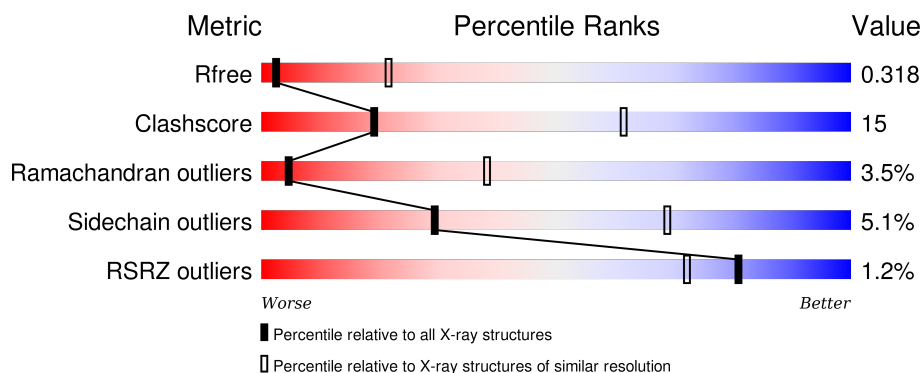
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.50 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	1051 (3.60-3.40)
Clashscore	102246	1157 (3.60-3.40)
Ramachandran outliers	100387	1120 (3.60-3.40)
Sidechain outliers	100360	1121 (3.60-3.40)
RSRZ outliers	91569	1058 (3.60-3.40)

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  $\geq 3$ , 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	276	<div> <div>65%</div> <div>29%</div> <div>5%</div> </div>
1	F	276	<div> <div>67%</div> <div>27%</div> <div>6%</div> </div>
1	K	276	<div> <div>5%</div> <div>69%</div> <div>25%</div> <div>5%</div> </div>
2	B	99	<div> <div>69%</div> <div>26%</div> <div>5%</div> </div>
2	G	99	<div> <div>69%</div> <div>26%</div> <div>5%</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	L	99	<div><div></div><div>3%</div><div>68%</div><div>27%</div><div>5%</div></div>
3	C	10	<div><div></div><div>30%</div><div>40%</div><div>30%</div></div>
3	H	10	<div><div></div><div>50%</div><div>40%</div><div>10%</div></div>
3	M	10	<div><div></div><div>50%</div><div>20%</div><div>20%</div><div>10%</div></div>
4	D	199	<div><div></div><div>56%</div><div>31%</div><div>10%</div><div></div></div>
4	I	199	<div><div></div><div>2%</div><div>63%</div><div>27%</div><div>8%</div><div></div></div>
4	N	199	<div><div></div><div>3%</div><div>65%</div><div>26%</div><div>8%</div><div></div></div>
5	E	244	<div><div></div><div>77%</div><div>21%</div><div></div></div>
5	J	244	<div><div></div><div>76%</div><div>22%</div><div></div></div>
5	O	244	<div><div></div><div>%</div><div>77%</div><div>21%</div><div></div></div>

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 20171 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 complex HLA-B\*4402.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	276	Total	C	N	O	S	0	0	0
			2258	1408	407	436	7			
1	F	276	Total	C	N	O	S	0	0	0
			2258	1408	407	436	7			
1	K	276	Total	C	N	O	S	0	0	0
			2258	1408	407	436	7			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	116	TYR	ASP	ENGINEERED	UNP P30481
F	116	TYR	ASP	ENGINEERED	UNP P30481
K	116	TYR	ASP	ENGINEERED	UNP P30481

- 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
			829	528	140	158	3			
2	G	99	Total	C	N	O	S	0	1	0
			832	530	140	159	3			
2	L	99	Total	C	N	O	S	0	0	0
			829	528	140	158	3			

- Molecule 3 is a protein called EBV decapeptide epitope.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	10	Total	C	N	O	0	0	0
			91	59	14	18			
3	H	10	Total	C	N	O	0	0	0
			90	59	14	17			

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	M	10	Total	C	N	O	0	0	0
			91	59	14	18			

- Molecule 4 is a protein called DM1 T cell receptor alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	199	Total	C	N	O	S	0	0	0
			1560	975	262	313	10			
4	I	199	Total	C	N	O	S	0	0	0
			1560	975	262	313	10			
4	N	199	Total	C	N	O	S	0	0	0
			1560	975	262	313	10			

- Molecule 5 is a protein called DM1 T cell receptor beta chain.

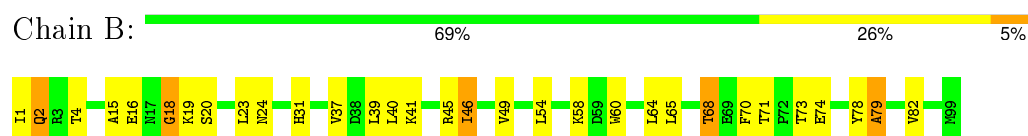
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	244	Total	C	N	O	S	0	0	0
			1985	1244	356	379	6			
5	J	244	Total	C	N	O	S	0	0	0
			1985	1244	356	379	6			
5	O	244	Total	C	N	O	S	0	0	0
			1985	1244	356	379	6			



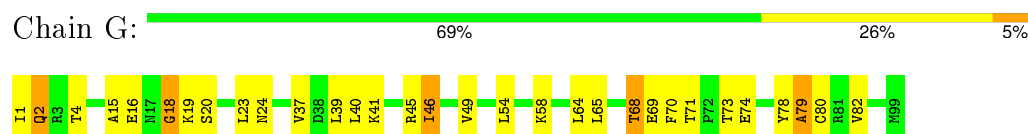
- Molecule 1: HLA class I histocompatibility complex HLA-B\*4402



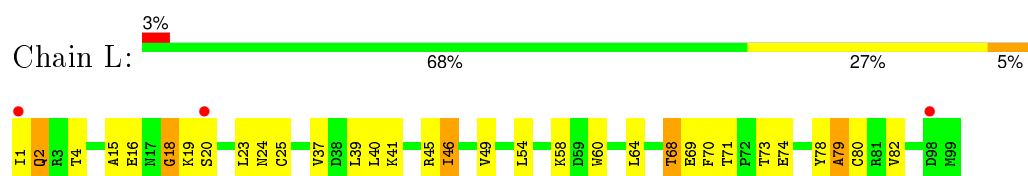
- Molecule 2: Beta-2-microglobulin



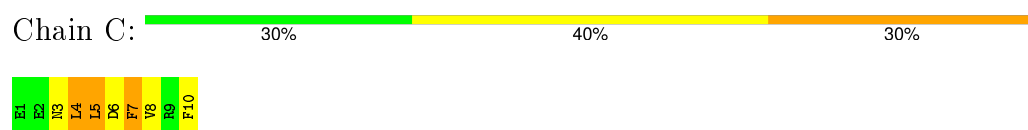
- Molecule 2: Beta-2-microglobulin



- Molecule 2: Beta-2-microglobulin



- Molecule 3: EBV decapeptide epitope



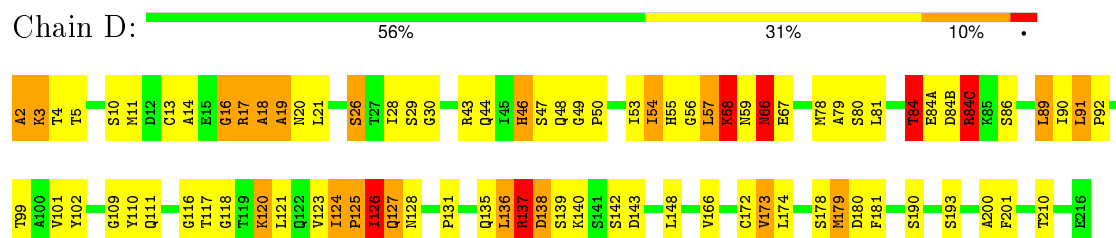
- Molecule 3: EBV decapeptide epitope



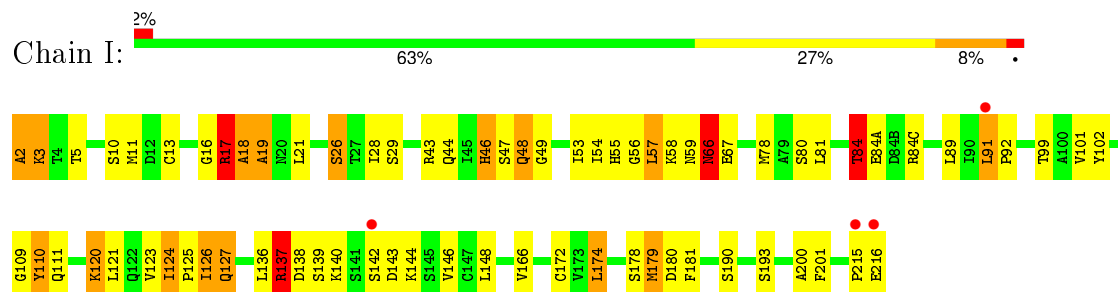
- Molecule 3: EBV decapeptide epitope



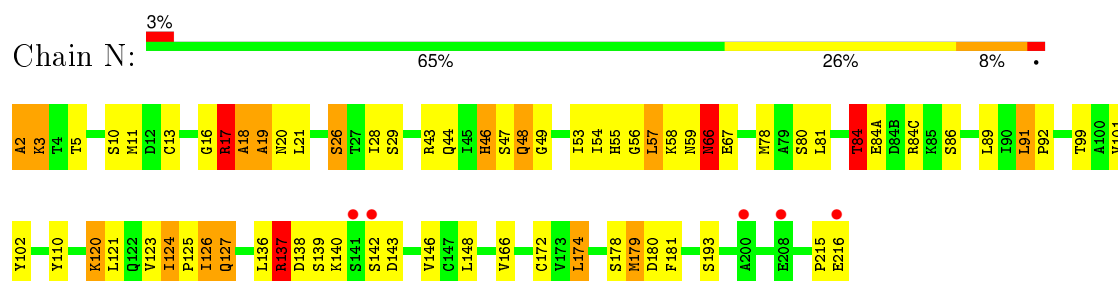
- Molecule 4: DM1 T cell receptor alpha chain



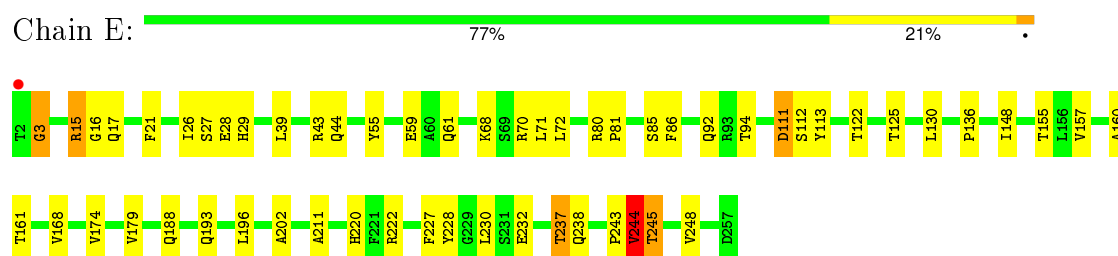
- Molecule 4: DM1 T cell receptor alpha chain



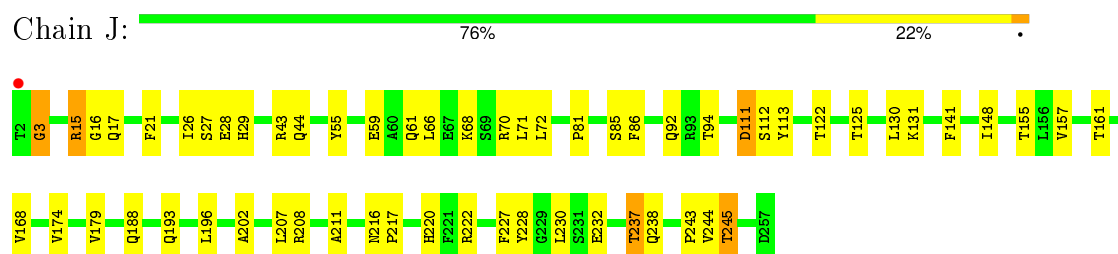
- Molecule 4: DM1 T cell receptor alpha chain



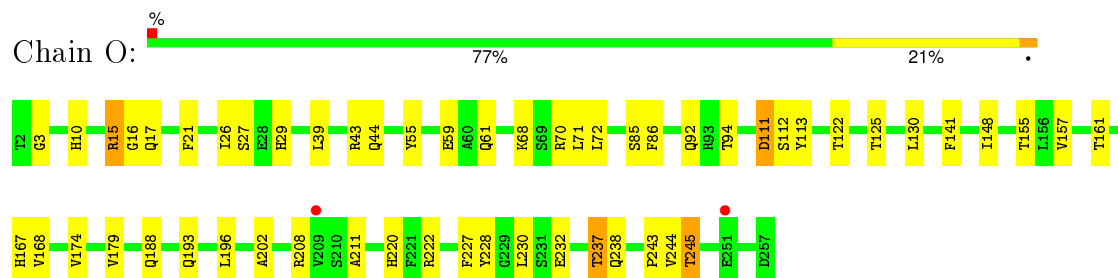
- Molecule 5: DM1 T cell receptor beta chain



- Molecule 5: DM1 T cell receptor beta chain



- Molecule 5: DM1 T cell receptor beta chain





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 65 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	121.95Å 121.95Å 695.91Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	100.00 – 3.50 49.37 – 3.50	Depositor EDS
% Data completeness (in resolution range)	97.4 (100.00-3.50) 97.5 (49.37-3.50)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.21 (at 3.48Å)	Xtriage
Refinement program	REFMAC 5.2	Depositor
R, $R_{free}$	0.286 , 0.330 0.283 , 0.318	Depositor DCC
$R_{free}$ test set	1970 reflections (5.29%)	DCC
Wilson B-factor (Å <sup>2</sup> )	73.8	Xtriage
Anisotropy	0.055	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.27 , 59.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.42$ , $\langle L^2 \rangle = 0.24$	Xtriage
Outliers	2 of 39228 reflections (0.005%)	Xtriage
$F_o, F_c$ correlation	0.81	EDS
Total number of atoms	20171	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	63.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.62% 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	0.42	0/2319	0.75	2/3154 (0.1%)
1	F	0.43	0/2319	0.65	0/3154
1	K	0.39	0/2319	0.65	1/3154 (0.0%)
2	B	0.41	0/852	0.66	0/1152
2	G	0.40	0/858	0.61	0/1160
2	L	0.36	0/852	0.58	0/1152
3	C	0.55	0/92	0.91	0/121
3	H	0.49	0/91	0.97	0/121
3	M	0.46	0/92	1.00	1/121 (0.8%)
4	D	0.50	0/1596	0.93	4/2171 (0.2%)
4	I	0.44	0/1596	0.66	0/2171
4	N	0.41	0/1596	0.68	1/2171 (0.0%)
5	E	0.43	0/2038	0.57	1/2767 (0.0%)
5	J	0.42	0/2038	0.56	0/2767
5	O	0.40	0/2038	0.55	0/2767
All	All	0.42	0/20696	0.67	10/28103 (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	20
1	F	0	18
1	K	0	18
2	B	0	9
2	G	0	9
2	L	0	9
4	D	0	32
4	I	0	21
4	N	0	22
All	All	0	158

There are no bond length outliers.

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	D	173	VAL	CG1-CB-CG2	7.37	122.69	110.90
4	N	166	VAL	CG1-CB-CG2	7.15	122.34	110.90
5	E	244	VAL	CG1-CB-CG2	6.49	121.28	110.90
1	K	103	VAL	CG1-CB-CG2	6.48	121.27	110.90
1	A	261	VAL	CG1-CB-CG2	6.00	120.49	110.90

There are no chirality outliers.

5 of 158 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	105	PRO	Peptide
1	A	106	ASP	Peptide
1	A	14	ARG	Peptide
1	A	18	GLY	Peptide
1	A	39	ASP	Peptide

## 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	2258	0	2128	85	0
1	F	2258	0	2128	82	0
1	K	2258	0	2128	73	0
2	B	829	0	794	22	0
2	G	832	0	799	21	0
2	L	829	0	794	22	0
3	C	91	0	86	15	0
3	H	90	0	86	17	0
3	M	91	0	86	10	0
4	D	1560	0	1485	63	0
4	I	1560	0	1485	56	0
4	N	1560	0	1485	52	0
5	E	1985	0	1879	46	0
5	J	1985	0	1879	49	0
5	O	1985	0	1879	44	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	20171	0	19121	596	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 15.

The worst 5 of 596 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)
4:N:18:ALA:CB	4:N:19:ALA:HA	1.83	1.08
4:I:18:ALA:CB	4:I:19:ALA:HA	1.84	1.06
1:A:246:ALA:HB3	1:A:247:VAL:HA	1.42	1.02
1:F:246:ALA:HB3	1:F:247:VAL:HA	1.40	1.01
1:K:246:ALA:HB3	1:K:247:VAL:HA	1.40	1.01

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	274/276 (99%)	229 (84%)	35 (13%)	10 (4%)	4	37
1	F	274/276 (99%)	234 (85%)	33 (12%)	7 (3%)	7	45
1	K	274/276 (99%)	232 (85%)	34 (12%)	8 (3%)	6	42
2	B	97/99 (98%)	83 (86%)	13 (13%)	1 (1%)	19	66
2	G	98/99 (99%)	84 (86%)	13 (13%)	1 (1%)	19	66
2	L	97/99 (98%)	84 (87%)	12 (12%)	1 (1%)	19	66
3	C	8/10 (80%)	6 (75%)	1 (12%)	1 (12%)	0	6
3	H	8/10 (80%)	6 (75%)	1 (12%)	1 (12%)	0	6
3	M	8/10 (80%)	6 (75%)	1 (12%)	1 (12%)	0	6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
4	D	197/199 (99%)	149 (76%)	33 (17%)	15 (8%)	1	15
4	I	197/199 (99%)	150 (76%)	34 (17%)	13 (7%)	1	19
4	N	197/199 (99%)	151 (77%)	33 (17%)	13 (7%)	1	19
5	E	242/244 (99%)	217 (90%)	20 (8%)	5 (2%)	9	50
5	J	242/244 (99%)	218 (90%)	19 (8%)	5 (2%)	9	50
5	O	242/244 (99%)	216 (89%)	23 (10%)	3 (1%)	16	61
All	All	2455/2484 (99%)	2065 (84%)	305 (12%)	85 (4%)	4	38

5 of 85 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	198	GLU
1	A	246	ALA
4	D	11	MET
4	D	17	ARG
4	D	18	ALA

### 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	237/237 (100%)	229 (97%)	8 (3%)	44	79
1	F	237/237 (100%)	229 (97%)	8 (3%)	44	79
1	K	237/237 (100%)	228 (96%)	9 (4%)	40	76
2	B	94/94 (100%)	89 (95%)	5 (5%)	28	67
2	G	95/94 (101%)	89 (94%)	6 (6%)	22	63
2	L	94/94 (100%)	88 (94%)	6 (6%)	22	62
3	C	10/10 (100%)	8 (80%)	2 (20%)	1	9
3	H	10/10 (100%)	9 (90%)	1 (10%)	9	41
3	M	10/10 (100%)	8 (80%)	2 (20%)	1	9
4	D	178/178 (100%)	166 (93%)	12 (7%)	20	61

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
4	I	178/178 (100%)	166 (93%)	12 (7%)	20	61
4	N	178/178 (100%)	165 (93%)	13 (7%)	17	57
5	E	216/216 (100%)	206 (95%)	10 (5%)	33	72
5	J	216/216 (100%)	207 (96%)	9 (4%)	36	74
5	O	216/216 (100%)	207 (96%)	9 (4%)	36	74
All	All	2206/2205 (100%)	2094 (95%)	112 (5%)	29	69

5 of 112 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	G	80	CYS
4	I	174	LEU
5	O	55	TYR
4	I	5	THR
4	I	81	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 39 such sidechains are listed below:

Mol	Chain	Res	Type
1	F	144	GLN
4	I	44	GLN
5	O	92	GLN
1	F	188	HIS
1	F	263	HIS

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry

There are no ligands in this entry.

## 5.7 Other polymers

There are no such residues in this entry.

## 5.8 Polymer linkage issues

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, 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	276/276 (100%)	-0.10	1 (0%) 93 90	64, 64, 64, 64	13 (4%)
1	F	276/276 (100%)	-0.13	0 100 100	64, 64, 64, 64	12 (4%)
1	K	276/276 (100%)	0.16	14 (5%) 32 24	64, 64, 64, 64	16 (5%)
2	B	99/99 (100%)	0.08	0 100 100	64, 64, 64, 64	6 (6%)
2	G	99/99 (100%)	-0.08	0 100 100	64, 64, 64, 64	5 (5%)
2	L	99/99 (100%)	0.33	3 (3%) 54 43	64, 64, 64, 64	9 (9%)
3	C	10/10 (100%)	-0.06	0 100 100	64, 64, 64, 64	0
3	H	10/10 (100%)	0.11	0 100 100	64, 64, 64, 64	3 (30%)
3	M	10/10 (100%)	-0.10	0 100 100	64, 64, 64, 64	0
4	D	199/199 (100%)	-0.19	0 100 100	64, 64, 64, 65	13 (6%)
4	I	199/199 (100%)	0.06	4 (2%) 68 59	64, 64, 64, 65	17 (8%)
4	N	199/199 (100%)	0.14	5 (2%) 61 50	64, 64, 64, 65	21 (10%)
5	E	244/244 (100%)	-0.22	1 (0%) 93 90	25, 64, 64, 65	21 (8%)
5	J	244/244 (100%)	-0.18	1 (0%) 93 90	25, 64, 64, 65	20 (8%)
5	O	244/244 (100%)	0.00	2 (0%) 87 80	25, 64, 64, 65	20 (8%)
All	All	2484/2484 (100%)	-0.03	31 (1%) 81 72	25, 64, 64, 65	176 (7%)

The worst 5 of 31 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	O	209	VAL	4.3
1	K	228	THR	3.9
5	E	2	THR	3.8
1	K	201	LEU	3.8
4	I	215	PRO	3.7



## 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.