



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 19, 2020 – 05:40 am BST

PDB ID : 3SJV  
Title : Crystal structure of the RL42 TCR in complex with HLA-B8-FLR  
Authors : Gras, S.; Wilmann, P.G.; Zhenjun, C.; Hanim, H.; Yu Chih, L.; Kjer-Nielsen, L.; Purcell, A.W.; Burrows, S.R.; Mccluskey, J.; Rossjohn, J.  
Deposited on : 2011-06-22  
Resolution : 3.10 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

<https://www.wwpdb.org/validation/2017/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
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

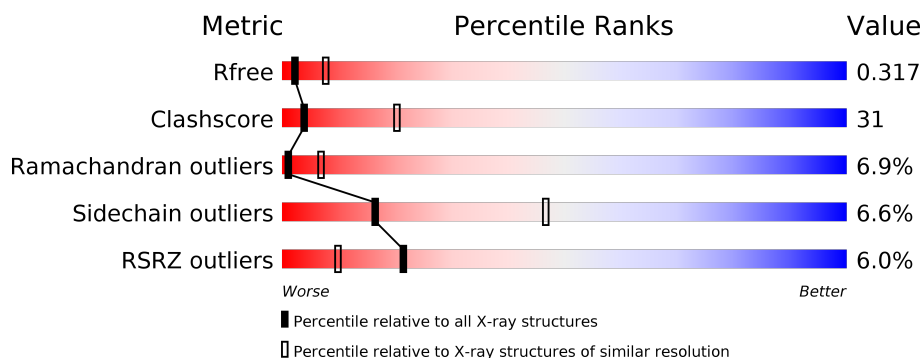
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.10 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))
$R_{free}$	130704	1094 (3.10-3.10)
Clashscore	141614	1184 (3.10-3.10)
Ramachandran outliers	138981	1141 (3.10-3.10)
Sidechain outliers	138945	1141 (3.10-3.10)
RSRZ outliers	127900	1067 (3.10-3.10)

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 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	277	<div> <div>10%</div> <div>43%</div> <div>46%</div> <div>10%</div> </div>
1	F	277	<div> <div>13%</div> <div>45%</div> <div>49%</div> <div>5%</div> </div>
1	K	277	<div> <div>11%</div> <div>49%</div> <div>45%</div> <div>6%</div> </div>
1	P	277	<div> <div>5%</div> <div>48%</div> <div>44%</div> <div>8%</div> </div>
2	B	100	<div> <div>49%</div> <div>41%</div> <div>9%</div> </div>
2	G	100	<div> <div>5%</div> <div>35%</div> <div>56%</div> <div>7%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	L	100	
2	Q	100	
3	C	9	
3	H	9	
3	M	9	
3	R	9	
4	D	203	
4	I	203	
4	N	203	
4	S	203	
5	E	244	
5	J	244	
5	O	244	
5	T	244	

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 26455 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, B-8 alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	277	Total	C	N	O	S	0	0	0
			2258	1398	412	441	7			
1	F	277	Total	C	N	O	S	0	0	0
			2258	1398	412	441	7			
1	K	277	Total	C	N	O	S	0	0	0
			2258	1398	412	441	7			
1	P	277	Total	C	N	O	S	0	0	0
			2258	1398	412	441	7			

- 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	0	0
			829	528	140	158	3			
2	L	100	Total	C	N	O	S	0	0	0
			837	533	141	159	4			
2	Q	100	Total	C	N	O	S	0	0	0
			837	533	141	159	4			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	0	MET	-	initiating methionine	UNP P61769
G	0	MET	-	initiating methionine	UNP P61769
L	0	MET	-	initiating methionine	UNP P61769
Q	0	MET	-	initiating methionine	UNP P61769

- Molecule 3 is a protein called Epstein-Barr nuclear antigen 3.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	9	Total	C	N	O	0	0	0
			75	49	15	11			
3	H	9	Total	C	N	O	0	0	0
			75	49	15	11			
3	M	9	Total	C	N	O	0	0	0
			75	49	15	11			
3	R	9	Total	C	N	O	0	0	0
			75	49	15	11			

- Molecule 4 is a protein called RL42 T cell receptor, alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	197	Total	C	N	O	S	1	0	0
			1541	962	258	313	8			
4	I	197	Total	C	N	O	S	1	0	0
			1541	962	258	313	8			
4	N	197	Total	C	N	O	S	1	0	0
			1541	962	258	313	8			
4	S	197	Total	C	N	O	S	1	0	0
			1541	962	258	313	8			

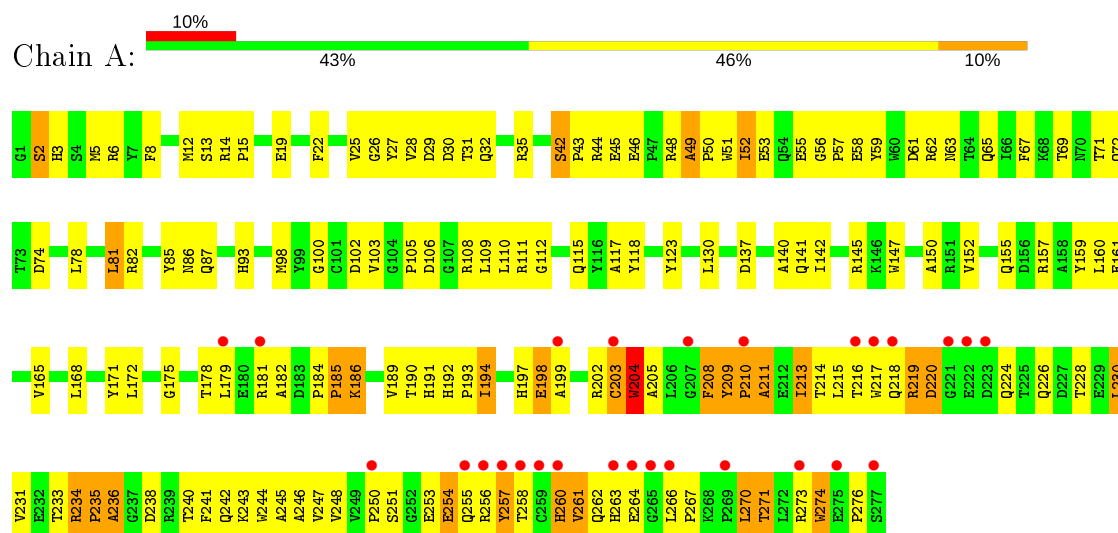
- Molecule 5 is a protein called RL42 T cell receptor, beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	241	Total	C	N	O	S	2	0	0
			1908	1205	331	363	9			
5	J	240	Total	C	N	O	S	2	0	0
			1903	1202	330	362	9			
5	O	241	Total	C	N	O	S	2	0	0
			1908	1205	331	363	9			
5	T	241	Total	C	N	O	S	2	0	0
			1908	1205	331	363	9			

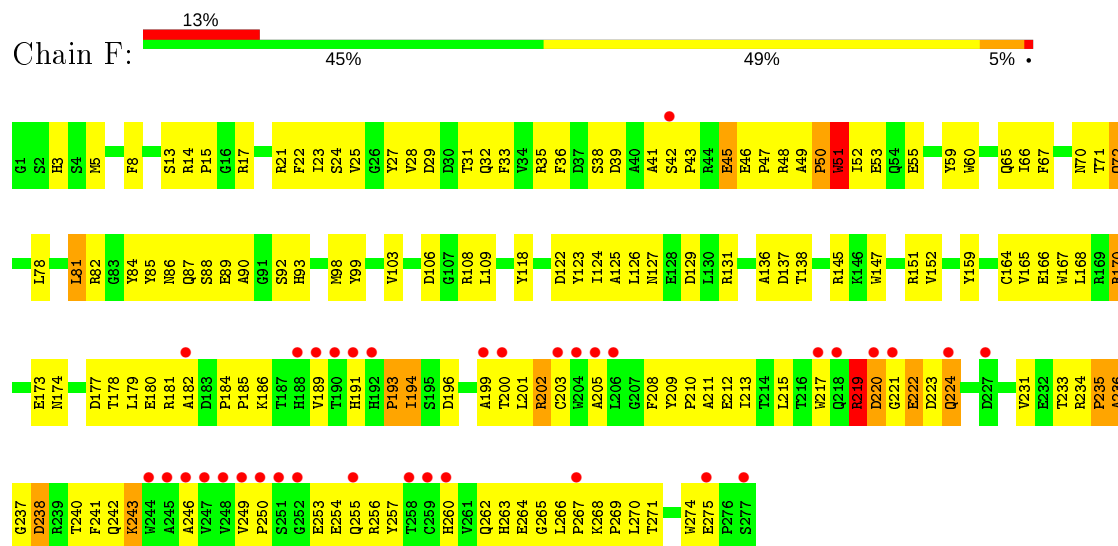
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA 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, B-8 alpha chain

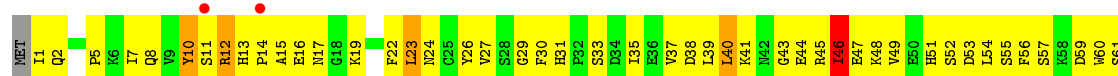


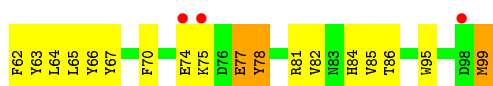
- Molecule 1: HLA class I histocompatibility antigen, B-8 alpha chain



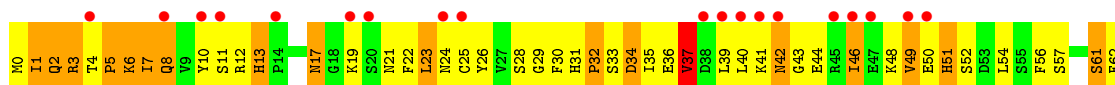
- Molecule 1: HLA class I histocompatibility antigen, B-8 alpha chain



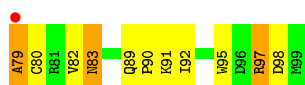
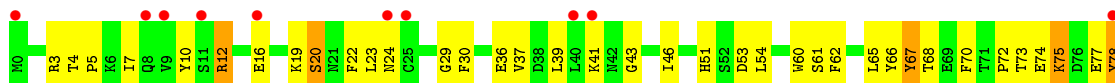




• Molecule 2: Beta-2-microglobulin



• Molecule 2: Beta-2-microglobulin



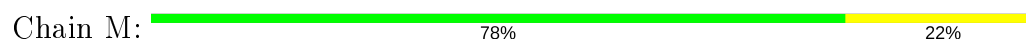
• Molecule 3: Epstein-Barr nuclear antigen 3



• Molecule 3: Epstein-Barr nuclear antigen 3



• Molecule 3: Epstein-Barr nuclear antigen 3

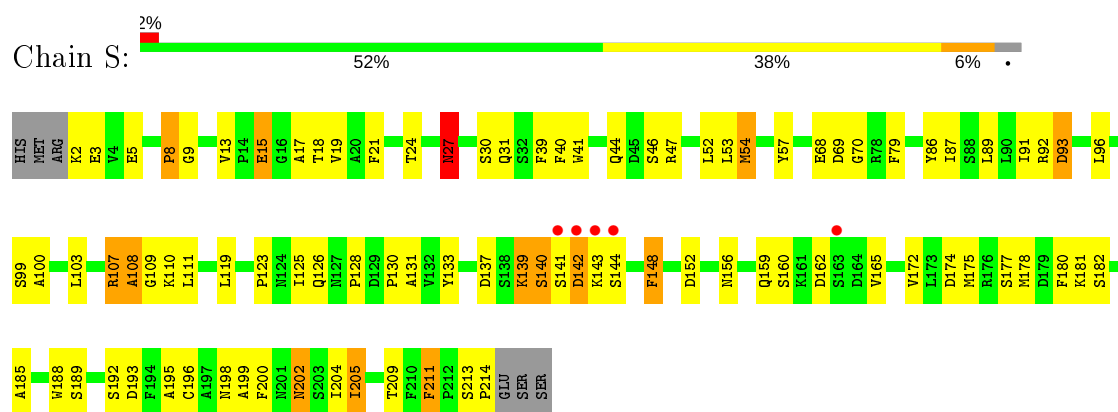


• Molecule 3: Epstein-Barr nuclear antigen 3

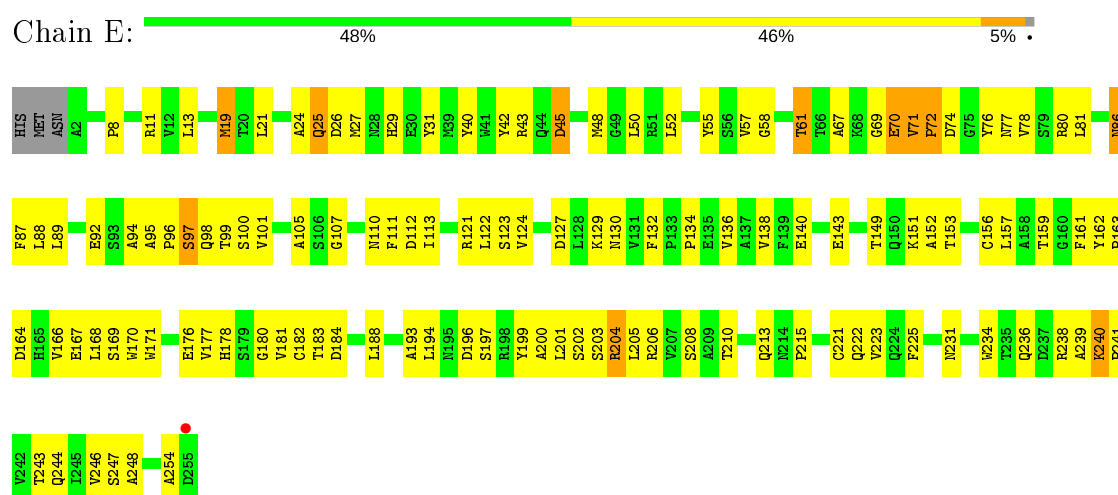








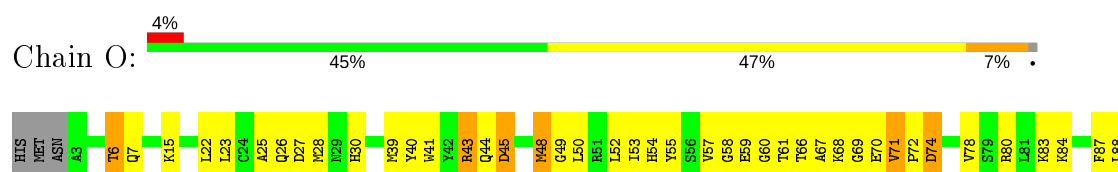
- Molecule 5: RL42 T cell receptor, beta chain

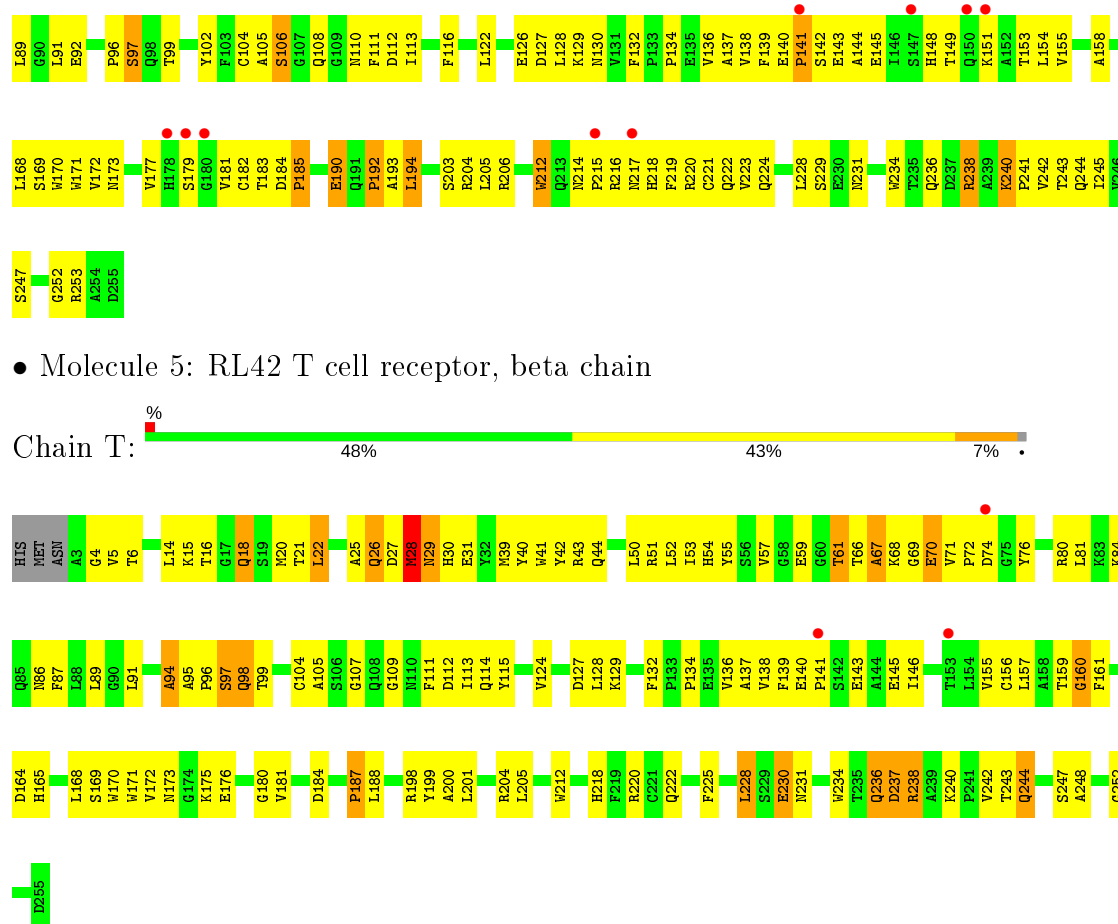


- Molecule 5: RL42 T cell receptor, beta chain



- Molecule 5: RL42 T cell receptor, beta chain





• Molecule 5: RL42 T cell receptor, beta chain

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	100.08Å 185.00Å 217.68Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.99 – 3.10 19.99 – 3.10	Depositor EDS
% Data completeness (in resolution range)	90.7 (19.99-3.10) 99.9 (19.99-3.10)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.38 (at 3.09Å)	Xtriage
Refinement program	PHENIX 1.6.1_357	Depositor
R, $R_{free}$	0.253 , 0.321 0.260 , 0.317	Depositor DCC
$R_{free}$ test set	3716 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	69.3	Xtriage
Anisotropy	0.265	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.23 , 39.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	26455	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	86.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 18.73% 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, 0.333 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	0.24	0/2320	0.40	0/3154
1	F	0.23	0/2320	0.40	0/3154
1	K	0.22	0/2320	0.40	0/3154
1	P	0.23	0/2320	0.40	0/3154
2	B	0.24	0/852	0.38	0/1152
2	G	0.22	0/852	0.41	0/1152
2	L	0.21	0/860	0.39	0/1162
2	Q	0.22	0/860	0.38	0/1162
3	C	0.23	0/76	0.43	0/98
3	H	0.31	0/76	0.35	0/98
3	M	0.28	0/76	0.45	0/98
3	R	0.28	0/76	0.36	0/98
4	D	0.26	0/1574	0.41	0/2132
4	I	0.26	0/1574	0.42	0/2132
4	N	0.24	0/1574	0.42	0/2132
4	S	0.27	0/1574	0.44	0/2132
5	E	0.25	0/1959	0.43	0/2662
5	J	0.28	0/1954	0.46	0/2655
5	O	0.24	0/1959	0.43	0/2662
5	T	0.25	0/1959	0.44	0/2662
All	All	0.24	0/27135	0.42	0/36805

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 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	2258	0	2098	149	0
1	F	2258	0	2098	137	0
1	K	2258	0	2098	132	0
1	P	2258	0	2098	136	0
2	B	829	0	794	51	0
2	G	829	0	794	86	0
2	L	837	0	803	81	0
2	Q	837	0	803	48	0
3	C	75	0	79	3	0
3	H	75	0	79	8	0
3	M	75	0	79	4	0
3	R	75	0	79	5	0
4	D	1541	0	1466	109	0
4	I	1541	0	1464	83	0
4	N	1541	0	1464	148	0
4	S	1541	0	1464	83	0
5	E	1908	0	1819	109	0
5	J	1903	0	1814	95	0
5	O	1908	0	1819	138	0
5	T	1908	0	1819	109	0
All	All	26455	0	25031	1573	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 31.

The worst 5 of 1573 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:I:12:ASN:HB3	4:I:122:LYS:HE3	1.29	1.14
5:J:96:PRO:HB3	5:J:124:VAL:HB	1.32	1.07
5:O:142:SER:HB3	5:O:145:GLU:HB3	1.42	0.99
4:N:207:GLU:N	4:N:208:ASP:HB2	1.77	0.99
4:D:30:SER:HA	4:D:108:ALA:CB	1.92	0.99

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	275/277 (99%)	193 (70%)	54 (20%)	28 (10%)	0	3
1	F	275/277 (99%)	196 (71%)	57 (21%)	22 (8%)	1	5
1	K	275/277 (99%)	196 (71%)	62 (22%)	17 (6%)	1	9
1	P	275/277 (99%)	210 (76%)	43 (16%)	22 (8%)	1	5
2	B	97/100 (97%)	77 (79%)	15 (16%)	5 (5%)	2	12
2	G	97/100 (97%)	75 (77%)	16 (16%)	6 (6%)	1	9
2	L	98/100 (98%)	70 (71%)	17 (17%)	11 (11%)	0	2
2	Q	98/100 (98%)	72 (74%)	20 (20%)	6 (6%)	1	9
3	C	7/9 (78%)	6 (86%)	1 (14%)	0	100	100
3	H	7/9 (78%)	6 (86%)	1 (14%)	0	100	100
3	M	7/9 (78%)	5 (71%)	2 (29%)	0	100	100
3	R	7/9 (78%)	7 (100%)	0	0	100	100
4	D	195/203 (96%)	139 (71%)	39 (20%)	17 (9%)	1	4
4	I	195/203 (96%)	140 (72%)	42 (22%)	13 (7%)	1	7
4	N	195/203 (96%)	120 (62%)	54 (28%)	21 (11%)	0	2
4	S	195/203 (96%)	146 (75%)	37 (19%)	12 (6%)	1	9
5	E	239/244 (98%)	192 (80%)	40 (17%)	7 (3%)	4	24
5	J	238/244 (98%)	191 (80%)	35 (15%)	12 (5%)	2	13
5	O	239/244 (98%)	172 (72%)	54 (23%)	13 (5%)	2	12
5	T	239/244 (98%)	196 (82%)	31 (13%)	12 (5%)	2	13
All	All	3253/3332 (98%)	2409 (74%)	620 (19%)	224 (7%)	1	7

5 of 224 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	30	ASP

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	211	ALA
1	A	235	PRO
1	A	267	PRO
2	B	14	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	236/236 (100%)	221 (94%)	15 (6%)	17	48
1	F	236/236 (100%)	220 (93%)	16 (7%)	16	45
1	K	236/236 (100%)	224 (95%)	12 (5%)	24	56
1	P	236/236 (100%)	220 (93%)	16 (7%)	16	45
2	B	94/95 (99%)	86 (92%)	8 (8%)	10	37
2	G	94/95 (99%)	86 (92%)	8 (8%)	10	37
2	L	95/95 (100%)	84 (88%)	11 (12%)	5	22
2	Q	95/95 (100%)	89 (94%)	6 (6%)	18	48
3	C	6/6 (100%)	4 (67%)	2 (33%)	0	0
3	H	6/6 (100%)	5 (83%)	1 (17%)	2	9
3	M	6/6 (100%)	6 (100%)	0	100	100
3	R	6/6 (100%)	6 (100%)	0	100	100
4	D	177/183 (97%)	166 (94%)	11 (6%)	18	49
4	I	177/183 (97%)	166 (94%)	11 (6%)	18	49
4	N	177/183 (97%)	162 (92%)	15 (8%)	10	37
4	S	177/183 (97%)	169 (96%)	8 (4%)	27	60
5	E	205/208 (99%)	190 (93%)	15 (7%)	14	43
5	J	205/208 (99%)	195 (95%)	10 (5%)	25	57
5	O	205/208 (99%)	193 (94%)	12 (6%)	19	50
5	T	205/208 (99%)	191 (93%)	14 (7%)	16	45

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
All	All	2874/2912 (99%)	2683 (93%)	191 (7%)	16	47

5 of 191 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	I	210	PHE
1	K	244	TRP
4	S	211	PHE
5	J	48	MET
5	J	253	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 72 such sidechains are listed below:

Mol	Chain	Res	Type
1	K	3	HIS
2	L	13	HIS
4	S	202	ASN
1	K	65	GLN
1	K	174	ASN

### 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 [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, 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	277/277 (100%)	0.28	27 (9%) 7 2	21, 78, 193, 223	0
1	F	277/277 (100%)	0.37	35 (12%) 3 1	30, 82, 214, 249	0
1	K	277/277 (100%)	0.37	31 (11%) 5 2	31, 93, 215, 245	0
1	P	277/277 (100%)	0.12	14 (5%) 28 13	25, 87, 200, 220	0
2	B	99/100 (99%)	-0.16	0 100 100	34, 67, 116, 130	0
2	G	99/100 (99%)	0.32	5 (5%) 28 13	49, 100, 159, 188	0
2	L	100/100 (100%)	1.51	32 (32%) 0 0	66, 193, 228, 236	0
2	Q	100/100 (100%)	0.65	11 (11%) 5 2	45, 136, 184, 191	0
3	C	9/9 (100%)	-0.87	0 100 100	27, 31, 41, 48	0
3	H	9/9 (100%)	-0.84	0 100 100	22, 34, 43, 45	0
3	M	9/9 (100%)	-0.63	0 100 100	29, 34, 52, 62	0
3	R	9/9 (100%)	-0.83	0 100 100	22, 27, 46, 54	0
4	D	197/203 (97%)	0.06	4 (2%) 65 44	31, 78, 133, 162	0
4	I	197/203 (97%)	-0.33	1 (0%) 91 81	20, 46, 116, 143	0
4	N	197/203 (97%)	0.61	20 (10%) 6 2	35, 97, 178, 201	0
4	S	197/203 (97%)	-0.15	5 (2%) 57 34	21, 51, 132, 166	0
5	E	241/244 (98%)	-0.27	1 (0%) 92 84	28, 64, 113, 133	2 (0%)
5	J	240/244 (98%)	-0.46	1 (0%) 92 84	20, 41, 90, 137	2 (0%)
5	O	241/244 (98%)	-0.01	9 (3%) 41 21	27, 61, 177, 189	2 (0%)
5	T	241/244 (98%)	-0.31	3 (1%) 79 61	20, 52, 114, 145	2 (0%)
All	All	3293/3332 (98%)	0.09	199 (6%) 21 10	20, 72, 193, 249	8 (0%)

The worst 5 of 199 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	218	GLN	10.5
1	K	237	GLY	8.2
1	A	217	TRP	7.4
1	F	249	VAL	6.9
1	A	221	GLY	6.6

## 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.