



wwPDB X-ray Structure Validation Summary Report ⓘ

May 22, 2020 – 10:11 pm BST

PDB ID : 5W1W
Title : Structure of the HLA-E-VMAPRTLVL/GF4 TCR complex
Authors : Gras, S.; Walpole, N.; Farenc, C.; Rossjohn, J.
Deposited on : 2017-06-05
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

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

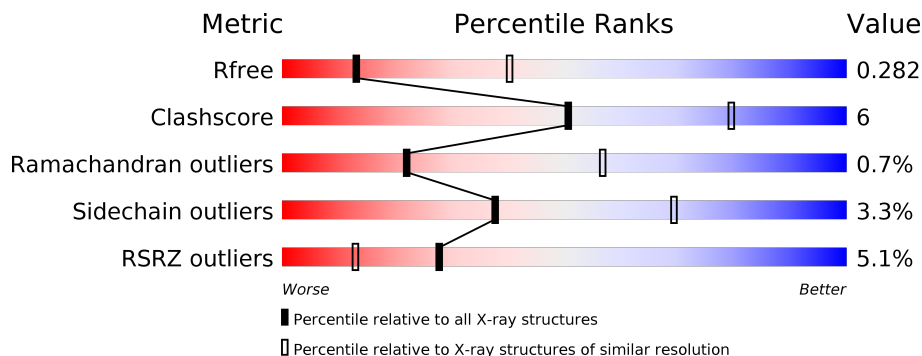
1 Overall quality at a glance

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 ≥ 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	278	
1	F	278	
1	K	278	
1	P	278	
2	B	100	
2	G	100	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	L	100	 16% 80% 20%
2	Q	100	 20% 75% 23%
3	C	9	 67% 33%
3	H	9	 89% 11%
3	M	9	 11% 78% 22%
3	R	9	 78% 22%
4	D	207	 80% 17%
4	I	207	 77% 19%
4	N	207	 4% 78% 17%
4	S	207	 2% 80% 18%
5	E	246	 79% 19%
5	J	246	 85% 13%
5	O	246	 2% 80% 18%
5	T	246	 81% 17%

2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 26478 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, alpha chain E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	272	Total	C	N	O	S	0	0	0
			2221	1392	398	424	7			
1	F	272	Total	C	N	O	S	0	0	0
			2221	1392	398	424	7			
1	K	272	Total	C	N	O	S	0	1	0
			2230	1397	399	427	7			
1	P	272	Total	C	N	O	S	0	1	0
			2230	1397	399	427	7			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	100	Total	C	N	O	S	0	1	0
			848	539	145	160	4			
2	G	100	Total	C	N	O	S	0	1	0
			848	539	145	160	4			
2	L	100	Total	C	N	O	S	0	1	0
			848	539	145	160	4			
2	Q	100	Total	C	N	O	S	0	1	0
			848	539	145	160	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 leader peptide of HLA class I histocompatibility antigen, A alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	9	Total	C	N	O	S	0	0	0
			69	45	12	11	1			
3	H	9	Total	C	N	O	S	0	0	0
			69	45	12	11	1			
3	M	9	Total	C	N	O	S	0	0	0
			69	45	12	11	1			
3	R	9	Total	C	N	O	S	0	0	0
			69	45	12	11	1			

- Molecule 4 is a protein called GF4 T cell receptor alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	201	Total	C	N	O	S	0	0	0
			1551	975	253	314	9			
4	I	202	Total	C	N	O	S	0	0	0
			1560	980	255	316	9			
4	N	199	Total	C	N	O	S	0	0	0
			1537	967	251	311	8			
4	S	202	Total	C	N	O	S	0	0	0
			1556	978	254	315	9			

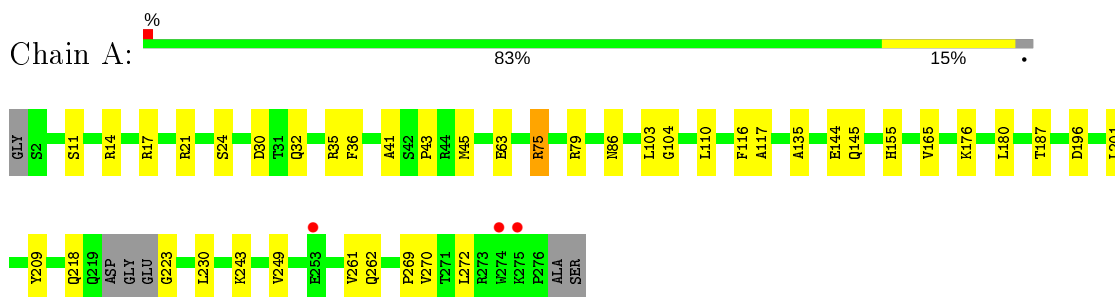
- Molecule 5 is a protein called GF4 T cell receptor beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	243	Total	C	N	O	S	0	0	0
			1926	1213	334	374	5			
5	J	243	Total	C	N	O	S	0	0	0
			1926	1213	334	374	5			
5	O	243	Total	C	N	O	S	0	0	0
			1926	1213	334	374	5			
5	T	243	Total	C	N	O	S	0	0	0
			1926	1213	334	374	5			

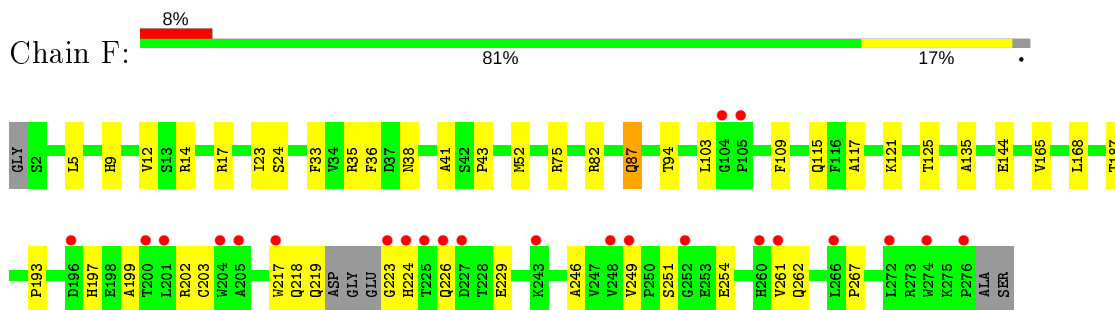
3 Residue-property plots [i](#)

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.

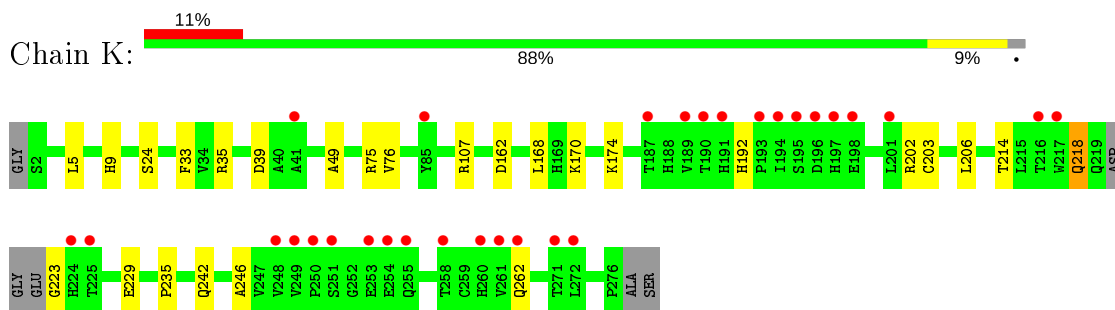
- Molecule 1: HLA class I histocompatibility antigen, alpha chain E



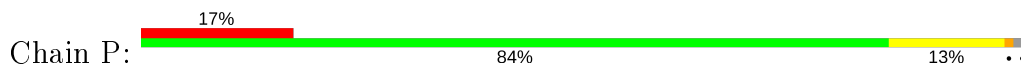
- Molecule 1: HLA class I histocompatibility antigen, alpha chain E

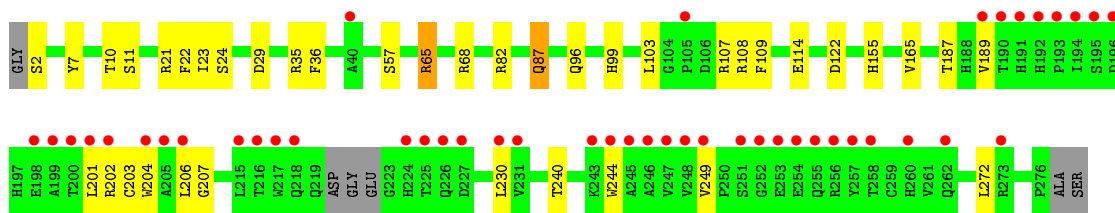


- Molecule 1: HLA class I histocompatibility antigen, alpha chain E

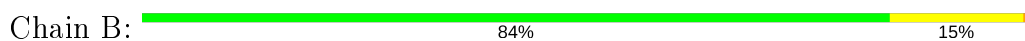


- Molecule 1: HLA class I histocompatibility antigen, alpha chain E

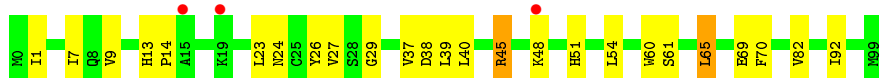
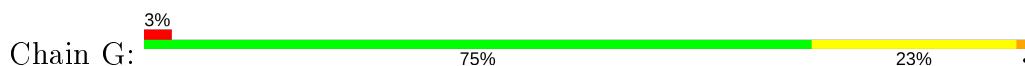




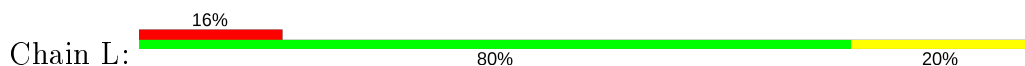
• Molecule 2: Beta-2-microglobulin



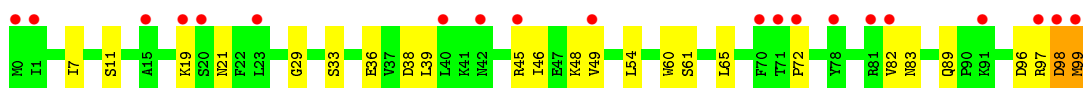
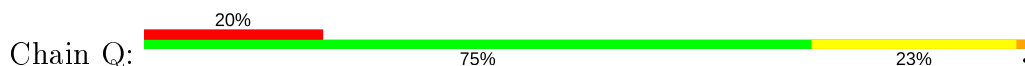
• Molecule 2: Beta-2-microglobulin



• Molecule 2: Beta-2-microglobulin



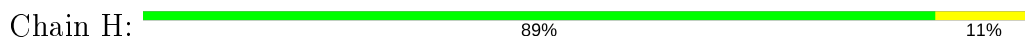
• Molecule 2: Beta-2-microglobulin



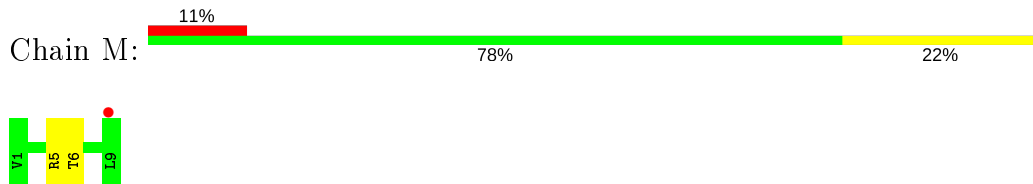
• Molecule 3: leader peptide of HLA class I histocompatibility antigen, A alpha chain



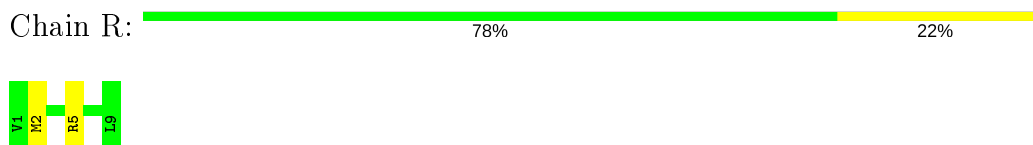
• Molecule 3: leader peptide of HLA class I histocompatibility antigen, A alpha chain



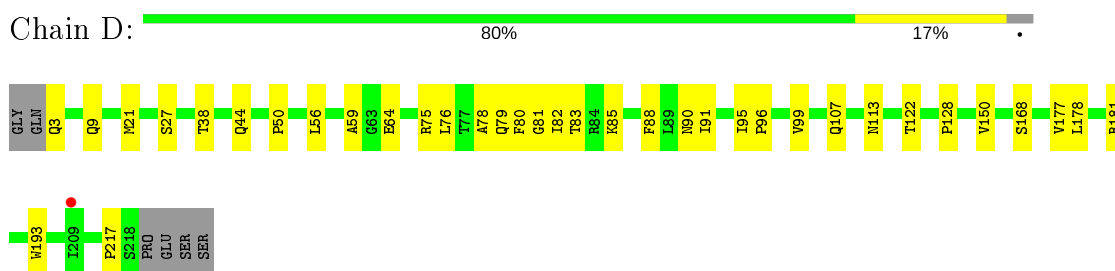
- Molecule 3: leader peptide of HLA class I histocompatibility antigen, A alpha chain



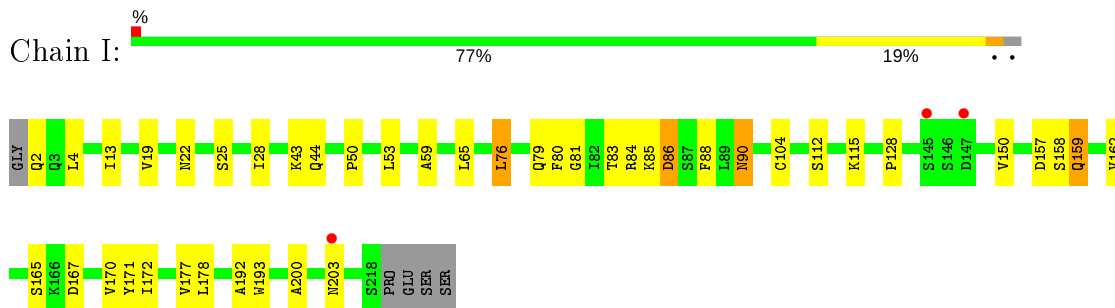
- Molecule 3: leader peptide of HLA class I histocompatibility antigen, A alpha chain



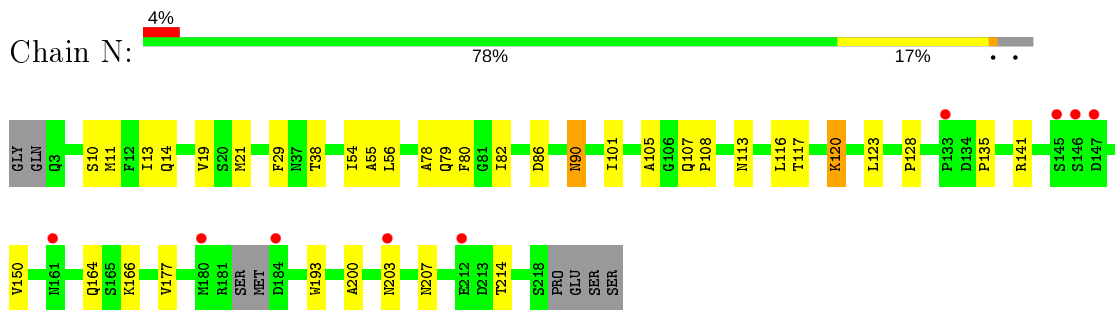
- Molecule 4: GF4 T cell receptor alpha chain



- Molecule 4: GF4 T cell receptor alpha chain

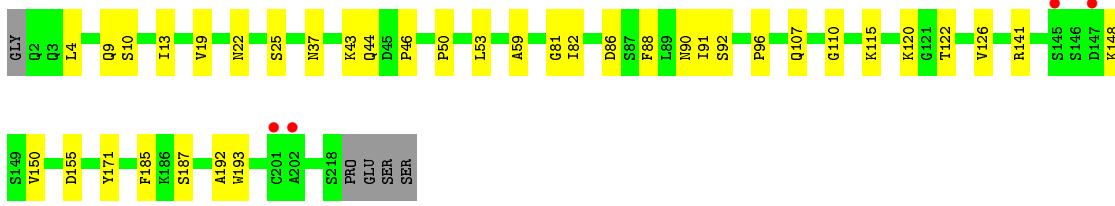


- Molecule 4: GF4 T cell receptor alpha chain

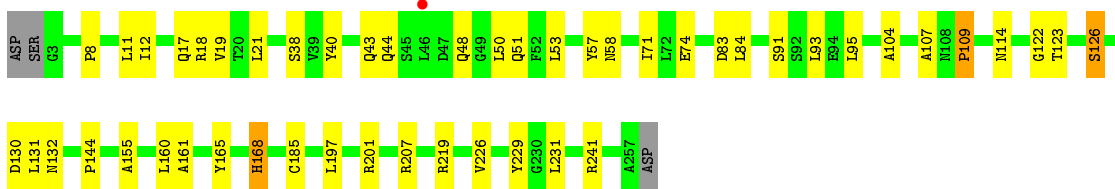
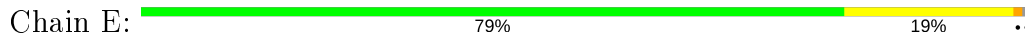


- Molecule 4: GF4 T cell receptor alpha chain

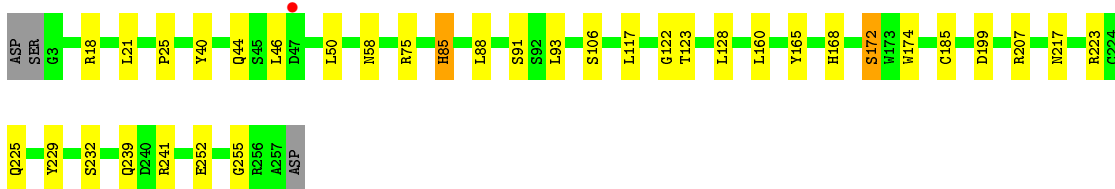
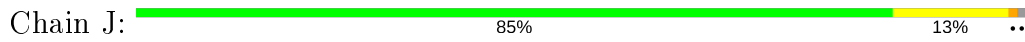




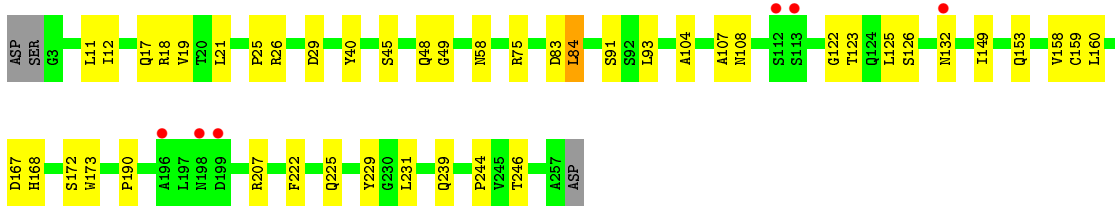
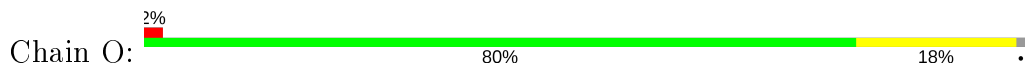
• Molecule 5: GF4 T cell receptor beta chain



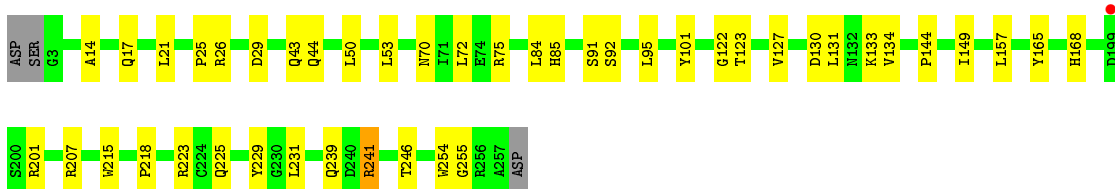
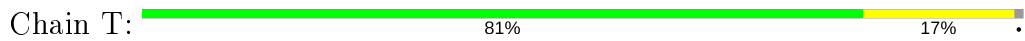
• Molecule 5: GF4 T cell receptor beta chain



• Molecule 5: GF4 T cell receptor beta chain



• Molecule 5: GF4 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, α , β , γ	73.37Å 225.93Å 276.27Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.29 – 3.10 49.63 – 3.10	Depositor EDS
% Data completeness (in resolution range)	99.8 (50.29-3.10) 99.8 (49.63-3.10)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.19 (at 3.12Å)	Xtrriage
Refinement program	BUSTER 2.10.2	Depositor
R, R_{free}	0.215 , 0.263 0.230 , 0.282	Depositor DCC
R_{free} test set	4202 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å ²)	66.5	Xtrriage
Anisotropy	0.552	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 70.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	26478	wwPDB-VP
Average B, all atoms (Å ²)	80.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.44% 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, 0.333 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	0.37	0/2287	0.64	0/3110
1	F	0.39	0/2287	0.65	1/3110 (0.0%)
1	K	0.37	0/2296	0.61	0/3122
1	P	0.39	0/2296	0.59	0/3122
2	B	0.37	0/871	0.67	0/1176
2	G	0.38	0/871	0.65	0/1176
2	L	0.38	0/871	0.63	0/1176
2	Q	0.40	0/871	0.65	0/1176
3	C	0.45	0/69	0.69	0/92
3	H	0.45	0/69	0.66	0/92
3	M	0.43	0/69	0.72	0/92
3	R	0.39	0/69	0.66	0/92
4	D	0.40	0/1585	0.73	0/2151
4	I	0.40	0/1594	0.71	0/2163
4	N	0.39	0/1570	0.69	0/2130
4	S	0.41	0/1590	0.69	0/2158
5	E	0.37	0/1976	0.70	1/2689 (0.0%)
5	J	0.37	0/1976	0.71	0/2689
5	O	0.38	0/1976	0.68	1/2689 (0.0%)
5	T	0.37	0/1976	0.68	0/2689
All	All	0.38	0/27169	0.67	3/36894 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	E	83	ASP	N-CA-C	-5.83	95.26	111.00
5	O	83	ASP	N-CA-C	-5.67	95.68	111.00
1	F	197	HIS	CB-CA-C	-5.59	99.22	110.40

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	2221	0	2066	22	0
1	F	2221	0	2066	26	0
1	K	2230	0	2071	15	0
1	P	2230	0	2071	27	0
2	B	848	0	815	8	0
2	G	848	0	815	13	0
2	L	848	0	815	12	0
2	Q	848	0	814	35	0
3	C	69	0	83	2	0
3	H	69	0	83	0	0
3	M	69	0	83	2	0
3	R	69	0	83	1	0
4	D	1551	0	1481	22	0
4	I	1560	0	1489	26	0
4	N	1537	0	1466	23	0
4	S	1556	0	1483	20	0
5	E	1926	0	1828	25	0
5	J	1926	0	1828	16	0
5	O	1926	0	1828	25	0
5	T	1926	0	1828	23	0
All	All	26478	0	25096	294	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 294 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:Q:96:ASP:OD1	2:Q:99:MET:N	1.59	1.34
2:Q:96:ASP:CG	2:Q:99:MET:HA	1.61	1.20
2:Q:96:ASP:OD1	2:Q:99:MET:CA	1.98	1.11
2:Q:96:ASP:OD1	2:Q:99:MET:HA	1.58	0.99
2:L:29:GLY:HA2	2:L:61:SER:HB2	1.50	0.94

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	268/278 (96%)	248 (92%)	20 (8%)	0	100	100
1	F	268/278 (96%)	243 (91%)	23 (9%)	2 (1%)	22	57
1	K	269/278 (97%)	244 (91%)	24 (9%)	1 (0%)	34	69
1	P	269/278 (97%)	252 (94%)	16 (6%)	1 (0%)	34	69
2	B	99/100 (99%)	97 (98%)	2 (2%)	0	100	100
2	G	99/100 (99%)	95 (96%)	4 (4%)	0	100	100
2	L	99/100 (99%)	93 (94%)	6 (6%)	0	100	100
2	Q	99/100 (99%)	96 (97%)	3 (3%)	0	100	100
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%)	6 (86%)	1 (14%)	0	100	100
3	R	7/9 (78%)	7 (100%)	0	0	100	100
4	D	199/207 (96%)	184 (92%)	13 (6%)	2 (1%)	15	49
4	I	200/207 (97%)	185 (92%)	14 (7%)	1 (0%)	29	64
4	N	195/207 (94%)	176 (90%)	18 (9%)	1 (0%)	29	64
4	S	200/207 (97%)	182 (91%)	15 (8%)	3 (2%)	10	39
5	E	241/246 (98%)	221 (92%)	18 (8%)	2 (1%)	19	54
5	J	241/246 (98%)	220 (91%)	16 (7%)	5 (2%)	7	30
5	O	241/246 (98%)	218 (90%)	20 (8%)	3 (1%)	13	44
5	T	241/246 (98%)	222 (92%)	16 (7%)	3 (1%)	13	44
All	All	3256/3360 (97%)	3001 (92%)	231 (7%)	24 (1%)	22	57

5 of 24 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
5	E	84	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
5	J	122	GLY
1	K	107	ARG
5	T	122	GLY
4	D	181	ARG

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	235/239 (98%)	229 (97%)	6 (3%)	46 74
1	F	235/239 (98%)	231 (98%)	4 (2%)	60 83
1	K	236/239 (99%)	230 (98%)	6 (2%)	47 75
1	P	236/239 (99%)	227 (96%)	9 (4%)	33 66
2	B	96/95 (101%)	92 (96%)	4 (4%)	30 62
2	G	96/95 (101%)	92 (96%)	4 (4%)	30 62
2	L	96/95 (101%)	93 (97%)	3 (3%)	40 70
2	Q	96/95 (101%)	90 (94%)	6 (6%)	18 48
3	C	8/8 (100%)	7 (88%)	1 (12%)	4 18
3	H	8/8 (100%)	7 (88%)	1 (12%)	4 18
3	M	8/8 (100%)	8 (100%)	0	100 100
3	R	8/8 (100%)	7 (88%)	1 (12%)	4 18
4	D	178/183 (97%)	173 (97%)	5 (3%)	43 73
4	I	179/183 (98%)	172 (96%)	7 (4%)	32 65
4	N	176/183 (96%)	170 (97%)	6 (3%)	37 69
4	S	178/183 (97%)	172 (97%)	6 (3%)	37 69
5	E	211/214 (99%)	200 (95%)	11 (5%)	23 55
5	J	211/214 (99%)	204 (97%)	7 (3%)	38 69
5	O	211/214 (99%)	206 (98%)	5 (2%)	49 76
5	T	211/214 (99%)	206 (98%)	5 (2%)	49 76

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	2913/2956 (98%)	2816 (97%)	97 (3%)	38 69

5 of 97 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	J	18	ARG
1	K	218	GLN
4	S	141	ARG
5	J	172	SER
5	J	252	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 43 such sidechains are listed below:

Mol	Chain	Res	Type
5	J	168	HIS
2	L	13	HIS
5	T	132	ASN
5	J	194	GLN
5	J	225	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 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, 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	272/278 (97%)	-0.06	3 (1%) 80 64	28, 57, 115, 136	0
1	F	272/278 (97%)	0.46	23 (8%) 10 4	41, 82, 147, 163	0
1	K	272/278 (97%)	0.56	30 (11%) 5 2	40, 95, 163, 176	0
1	P	272/278 (97%)	0.74	46 (16%) 1 0	60, 99, 170, 185	0
2	B	100/100 (100%)	-0.19	0 100 100	43, 58, 84, 99	0
2	G	100/100 (100%)	0.31	3 (3%) 50 27	64, 92, 122, 134	0
2	L	100/100 (100%)	0.96	16 (16%) 1 1	69, 125, 147, 155	0
2	Q	100/100 (100%)	1.03	20 (20%) 1 0	87, 132, 154, 158	0
3	C	9/9 (100%)	-0.10	0 100 100	30, 36, 39, 48	0
3	H	9/9 (100%)	-0.09	0 100 100	41, 47, 54, 56	0
3	M	9/9 (100%)	0.22	1 (11%) 5 2	59, 61, 66, 83	0
3	R	9/9 (100%)	-0.10	0 100 100	67, 69, 77, 90	0
4	D	201/207 (97%)	0.10	1 (0%) 91 81	26, 52, 119, 131	0
4	I	202/207 (97%)	0.03	3 (1%) 73 54	27, 51, 114, 131	0
4	N	199/207 (96%)	0.27	9 (4%) 33 16	45, 74, 152, 158	0
4	S	202/207 (97%)	0.07	4 (1%) 65 44	42, 59, 118, 127	0
5	E	243/246 (98%)	-0.09	1 (0%) 92 84	29, 57, 92, 108	0
5	J	243/246 (98%)	-0.06	1 (0%) 92 84	35, 53, 86, 109	0
5	O	243/246 (98%)	0.16	6 (2%) 57 34	42, 81, 118, 129	0
5	T	243/246 (98%)	-0.08	1 (0%) 92 84	28, 61, 93, 111	0
All	All	3300/3360 (98%)	0.23	168 (5%) 28 13	26, 72, 148, 185	0

The worst 5 of 168 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	225	THR	8.8
1	P	245	ALA	7.5
4	N	145	SER	6.8
1	P	201	LEU	5.9
1	P	200	THR	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.