



wwPDB X-ray Structure Validation Summary Report ⓘ

May 16, 2020 – 09:52 pm BST

PDB ID : 3O6F
Title : Crystal structure of a human autoimmune TCR MS2-3C8 bound to MHC class II self-ligand MBP/HLA-DR4
Authors : Yin, Y.; Li, Y.; Martin, R.; Mariuzza, R.A.
Deposited on : 2010-07-29
Resolution : 2.80 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

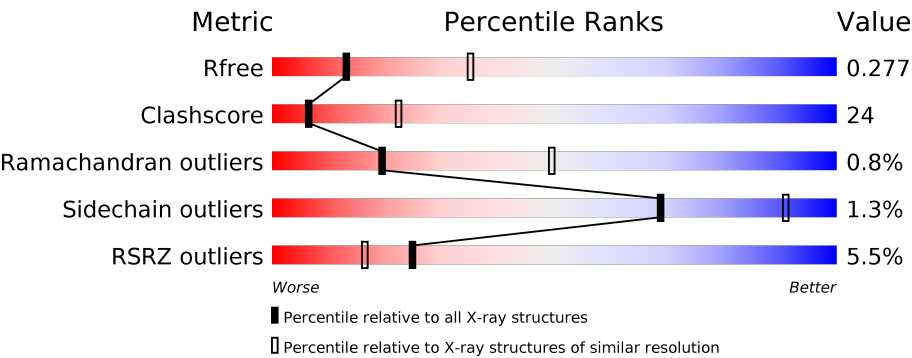
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 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	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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	182	<div><div>4%</div><div><div></div><div>57%</div><div>40%</div><div>••</div></div></div>
1	E	182	<div><div>3%</div><div><div></div><div>70%</div><div>28%</div><div>•</div></div></div>
2	B	221	<div><div>2%</div><div><div></div><div>50%</div><div>37%</div><div>•</div><div>12%</div></div></div>
2	F	221	<div><div>2%</div><div><div></div><div>62%</div><div>29%</div><div>•</div><div>5%</div></div></div>
3	C	206	<div><div>6%</div><div><div></div><div>64%</div><div>29%</div><div>•</div><div>7%</div></div></div>
3	G	206	<div><div>12%</div><div><div></div><div>57%</div><div>33%</div><div>•</div><div>9%</div></div></div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	D	245	 60%40%.
4	H	245	 12%46%52%..

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 12887 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 II histocompatibility antigen, DR alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	178	Total	C	N	O	S	0	0	0
			1446	940	232	269	5			
1	E	179	Total	C	N	O	S	0	0	0
			1456	947	237	267	5			

- Molecule 2 is a protein called HLA class II histocompatibility antigen, DRB1-4 beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	195	Total	C	N	O	S	0	0	0
			1578	1001	271	301	5			
2	F	210	Total	C	N	O	S	0	0	0
			1671	1054	294	318	5			

- Molecule 3 is a protein called T-cell receptor alpha chain C region.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	191	Total	C	N	O	S	4	0	0
			1441	897	241	296	7			
3	G	188	Total	C	N	O	S	0	0	0
			1448	902	247	291	8			

- Molecule 4 is a protein called T-cell receptor beta-1 chain C region.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	245	Total	C	N	O	S	0	0	0
			1919	1210	327	374	8			
4	H	243	Total	C	N	O	S	0	0	0
			1901	1201	324	369	7			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	189	SER	CYS	ENGINEERED MUTATION	UNP A0A5B9
H	189	SER	CYS	ENGINEERED MUTATION	UNP A0A5B9

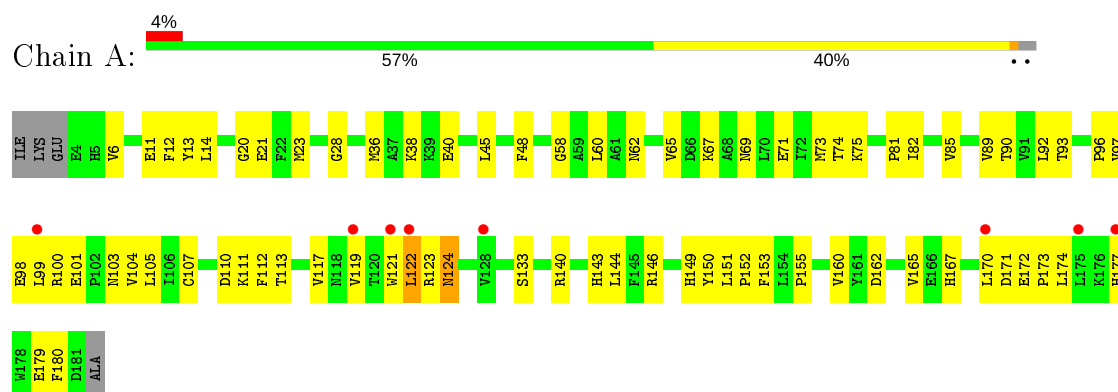
- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	2	Total O 2 2	0	0
5	B	3	Total O 3 3	0	0
5	C	8	Total O 8 8	0	0
5	D	4	Total O 4 4	0	0
5	E	1	Total O 1 1	0	0
5	F	4	Total O 4 4	0	0
5	G	5	Total O 5 5	0	0

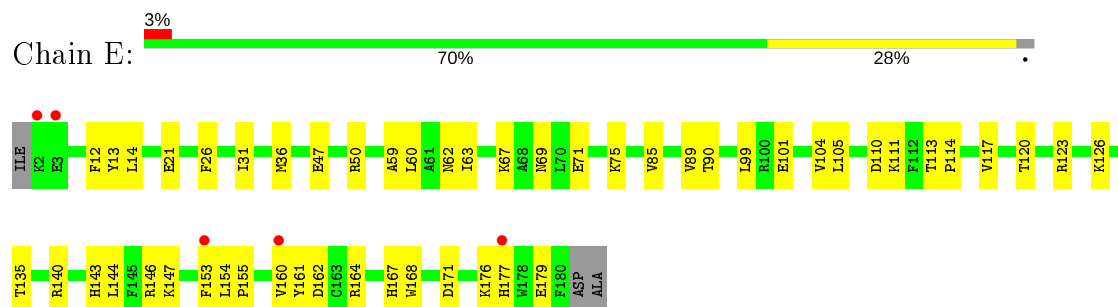
3 Residue-property plots

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.

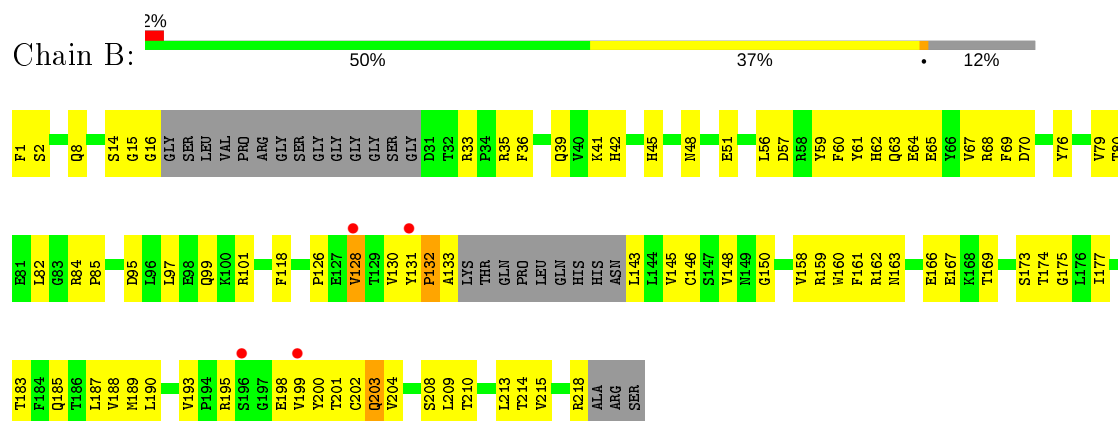
- Molecule 1: HLA class II histocompatibility antigen, DR alpha chain



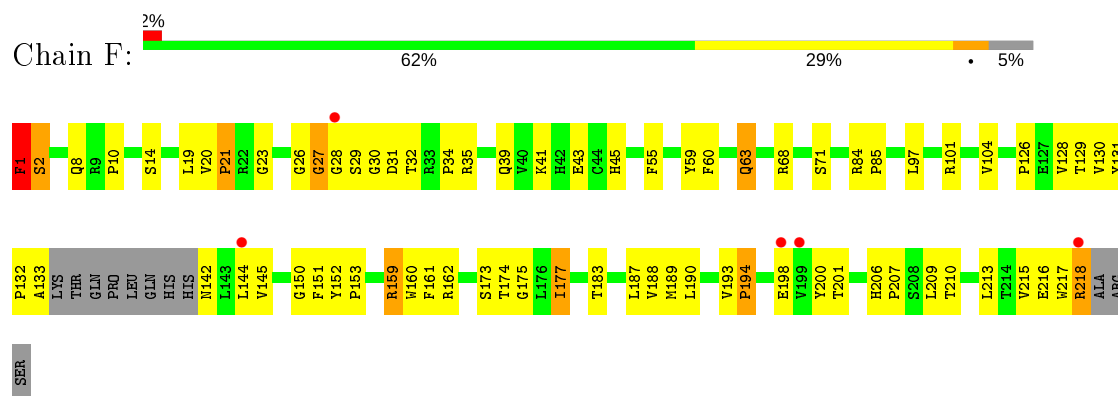
- Molecule 1: HLA class II histocompatibility antigen, DR alpha chain



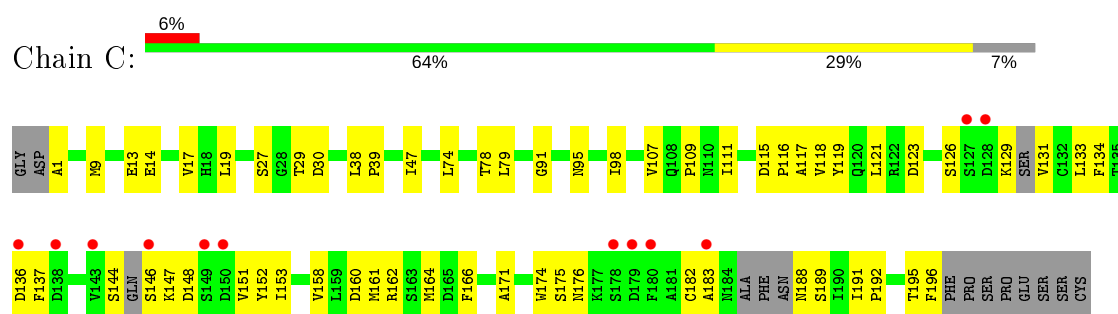
- Molecule 2: HLA class II histocompatibility antigen, DRB1-4 beta chain



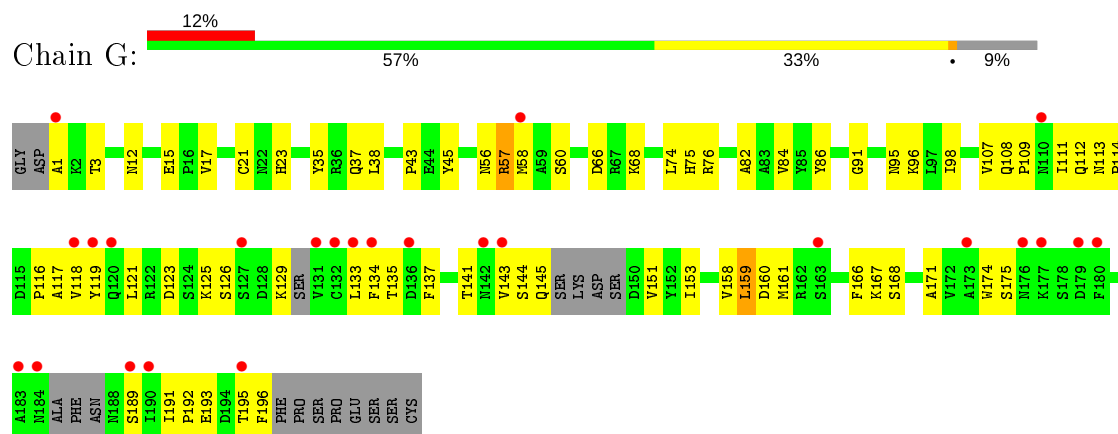
- Molecule 2: HLA class II histocompatibility antigen, DRB1-4 beta chain



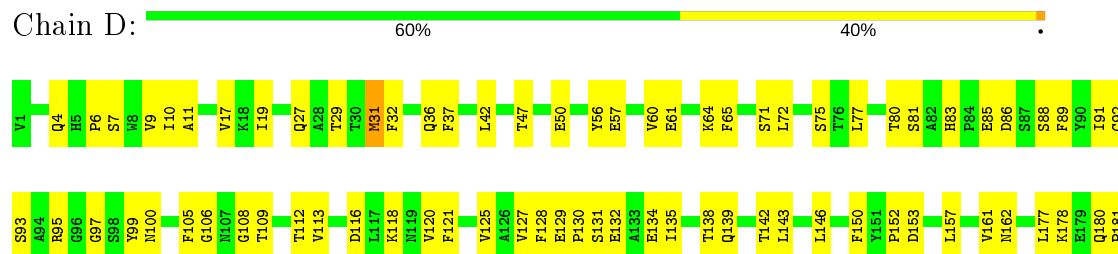
- Molecule 3: T-cell receptor alpha chain C region

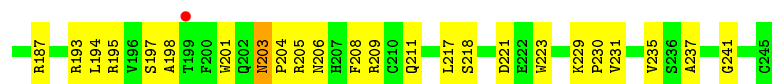


- Molecule 3: T-cell receptor alpha chain C region

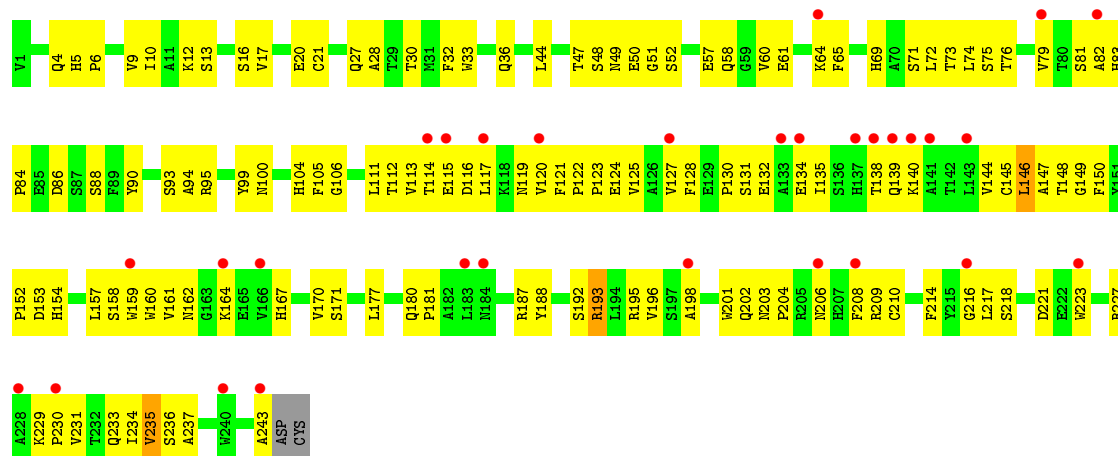


- Molecule 4: T-cell receptor beta-1 chain C region





• Molecule 4: T-cell receptor beta-1 chain C region



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	102.48 Å 218.40 Å 98.41 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.20 – 2.80 49.21 – 2.80	Depositor EDS
% Data completeness (in resolution range)	99.8 (49.20-2.80) 99.8 (49.21-2.80)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.51 (at 2.81 Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.6.2_432)	Depositor
R, R_{free}	0.239 , 0.279 0.237 , 0.277	Depositor DCC
R_{free} test set	2782 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å ²)	58.4	Xtriage
Anisotropy	0.243	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 47.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.011 for l,-k,h	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	12887	wwPDB-VP
Average B, all atoms (Å ²)	66.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.37% 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

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	0.48	0/1491	0.64	0/2036
1	E	0.47	0/1501	0.64	0/2048
2	B	0.51	0/1622	0.69	0/2204
2	F	0.62	1/1717 (0.1%)	0.72	3/2330 (0.1%)
3	C	0.50	0/1469	0.68	0/2005
3	G	0.45	0/1476	0.62	0/2008
4	D	0.48	0/1972	0.65	0/2690
4	H	0.45	0/1954	0.66	0/2666
All	All	0.50	1/13202 (0.0%)	0.66	3/17987 (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
2	F	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	1	PHE	CD2-CE2	5.22	1.49	1.39

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	218	ARG	NE-CZ-NH2	6.10	123.35	120.30
2	F	159	ARG	NE-CZ-NH1	5.78	123.19	120.30
2	F	194	PRO	N-CA-C	5.01	125.13	112.10

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	F	68	ARG	Sidechain

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	1446	0	1370	83	0
1	E	1456	0	1393	50	0
2	B	1578	0	1451	83	0
2	F	1671	0	1552	70	0
3	C	1441	0	1333	67	0
3	G	1448	0	1377	71	0
4	D	1919	0	1809	114	0
4	H	1901	0	1794	138	0
5	A	2	0	0	0	0
5	B	3	0	0	0	0
5	C	8	0	0	0	0
5	D	4	0	0	0	0
5	E	1	0	0	0	0
5	F	4	0	0	3	0
5	G	5	0	0	1	0
All	All	12887	0	12079	597	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 24.

The worst 5 of 597 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:D:135:ILE:HD11	4:D:198:ALA:CB	1.58	1.30
3:C:121:LEU:CD2	4:D:131:SER:N	2.08	1.17
4:D:128:PHE:CE2	4:D:146:LEU:HD12	1.80	1.16
4:D:135:ILE:HD11	4:D:198:ALA:HB2	1.17	1.15
3:C:121:LEU:HD21	4:D:131:SER:H	1.16	1.09

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	176/182 (97%)	160 (91%)	15 (8%)	1 (1%)	25	56
1	E	177/182 (97%)	164 (93%)	13 (7%)	0	100	100
2	B	189/221 (86%)	176 (93%)	11 (6%)	2 (1%)	14	41
2	F	206/221 (93%)	185 (90%)	17 (8%)	4 (2%)	8	26
3	C	183/206 (89%)	162 (88%)	20 (11%)	1 (0%)	29	61
3	G	180/206 (87%)	154 (86%)	25 (14%)	1 (1%)	25	56
4	D	243/245 (99%)	225 (93%)	18 (7%)	0	100	100
4	H	241/245 (98%)	210 (87%)	28 (12%)	3 (1%)	13	39
All	All	1595/1708 (93%)	1436 (90%)	147 (9%)	12 (1%)	19	49

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	F	2	SER
4	H	235	VAL
2	B	132	PRO
2	B	2	SER
4	H	146	LEU

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	159/166 (96%)	157 (99%)	2 (1%)	69	91

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	E	160/166 (96%)	160 (100%)	0	100	100
2	B	167/189 (88%)	164 (98%)	3 (2%)	59	86
2	F	176/189 (93%)	172 (98%)	4 (2%)	50	82
3	C	159/183 (87%)	159 (100%)	0	100	100
3	G	163/183 (89%)	160 (98%)	3 (2%)	59	86
4	D	212/216 (98%)	207 (98%)	5 (2%)	49	81
4	H	209/216 (97%)	208 (100%)	1 (0%)	88	96
All	All	1405/1508 (93%)	1387 (99%)	18 (1%)	69	91

5 of 18 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	D	193	ARG
4	D	203	ASN
3	G	57	ARG
4	D	81	SER
4	D	161	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 24 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	62	ASN
2	F	45	HIS
4	H	40	GLN
1	E	84	ASN
1	E	143	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 [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(Å²)	Q<0.9
1	A	178/182 (97%)	0.14	8 (4%)	33	23	32, 66, 117, 124	0
1	E	179/182 (98%)	0.12	5 (2%)	53	43	27, 51, 98, 108	0
2	B	195/221 (88%)	-0.05	4 (2%)	63	54	29, 54, 93, 133	0
2	F	210/221 (95%)	0.01	5 (2%)	59	49	26, 48, 93, 116	0
3	C	191/206 (92%)	0.28	12 (6%)	20	12	24, 52, 128, 139	1 (0%)
3	G	188/206 (91%)	0.60	25 (13%)	3	2	28, 68, 137, 147	0
4	D	245/245 (100%)	0.02	1 (0%)	92	91	34, 57, 95, 125	0
4	H	243/245 (99%)	0.71	30 (12%)	4	2	30, 92, 128, 143	0
All	All	1629/1708 (95%)	0.23	90 (5%)	25	16	24, 60, 124, 147	1 (0%)

The worst 5 of 90 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	H	243	ALA	7.7
3	G	190	ILE	7.1
4	H	140	LYS	5.7
3	C	143	VAL	5.4
3	G	119	TYR	5.1

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 ⓘ

There are no carbohydrates in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.