



wwPDB X-ray Structure Validation Summary Report ⓘ

May 15, 2020 – 12:15 am BST

PDB ID : 6PUH
Title : Structure of human MAIT A-F7 TCR in complex with human MR1-Ribityl-less
Authors : Awad, W.; Keller, A.N.; Rossjohn, J.
Deposited on : 2019-07-18
Resolution : 1.88 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
Mogul : 1.8.5 (274361), CSD as541be (2020)
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

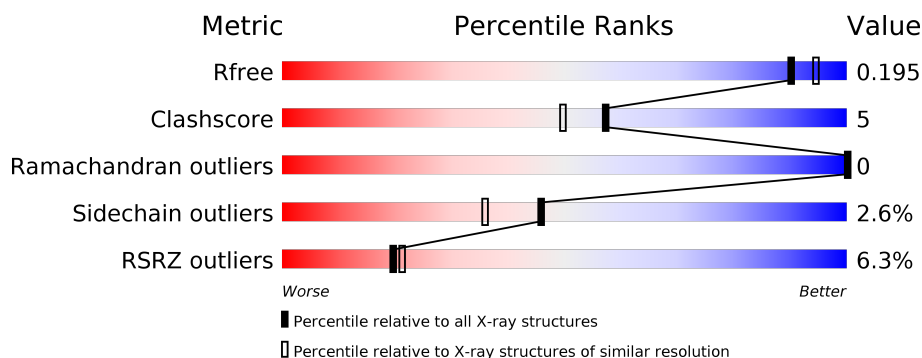
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.88 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	9470 (1.90-1.86)
Clashscore	141614	10282 (1.90-1.86)
Ramachandran outliers	138981	10152 (1.90-1.86)
Sidechain outliers	138945	10152 (1.90-1.86)
RSRZ outliers	127900	9303 (1.90-1.86)

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	271	<div> <div>3%</div> <div> <div></div> <div>85%</div> <div>12%</div> <div>••</div> </div> </div>
1	C	271	<div> <div>2%</div> <div> <div></div> <div>92%</div> <div>7%</div> <div>•</div> </div> </div>
2	B	204	<div> <div></div> <div> <div></div> <div>86%</div> <div>10%</div> <div>••</div> </div> </div>
2	D	204	<div> <div>20%</div> <div> <div></div> <div>83%</div> <div>12%</div> <div>5%</div> </div> </div>
3	E	246	<div> <div>9%</div> <div> <div></div> <div>83%</div> <div>13%</div> <div>••</div> </div> </div>
3	G	246	<div> <div>2%</div> <div> <div></div> <div>89%</div> <div>9%</div> <div>••</div> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	F	100	<div><div></div><div>2%</div><div>96%</div><div></div><div>.</div></div>
4	H	100	<div><div></div><div>19%</div><div>91%</div><div>7%</div><div></div><div>.</div></div>

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 15614 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 Major histocompatibility complex class I-related gene protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	265	Total	C	N	O	S	0	13	0
			2246	1441	388	404	13			
1	C	268	Total	C	N	O	S	0	18	0
			2299	1475	387	422	15			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	initiating methionine	UNP Q95460
A	261	SER	CYS	conflict	UNP Q95460
C	0	MET	-	initiating methionine	UNP Q95460
C	261	SER	CYS	conflict	UNP Q95460

- Molecule 2 is a protein called Human TCR alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	199	Total	C	N	O	S	0	24	0
			1674	1070	260	333	11			
2	D	194	Total	C	N	O	S	0	8	0
			1529	976	240	303	10			

- Molecule 3 is a protein called Human TCR beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	242	Total	C	N	O	S	0	13	0
			1949	1234	331	372	12			
3	G	244	Total	C	N	O	S	0	22	0
			2033	1284	345	391	13			

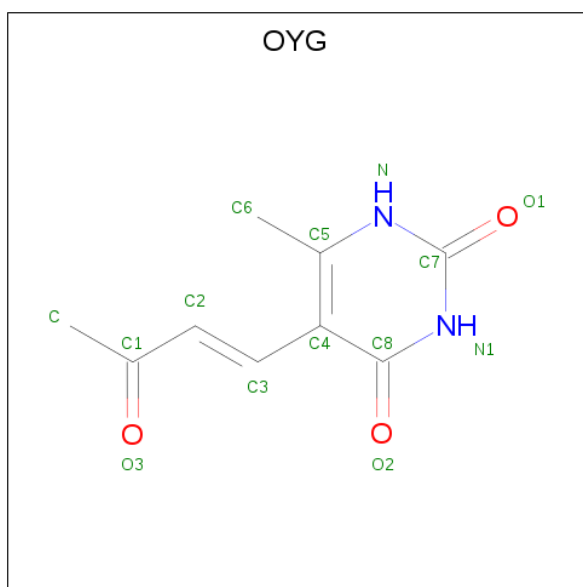
- Molecule 4 is a protein called Beta-2-microglobulin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	F	100	Total	C	N	O	S	0	3	0
			840	540	141	154	5			
4	H	98	Total	C	N	O	S	0	1	0
			793	510	135	145	3			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	0	MET	-	initiating methionine	UNP P61769
H	0	MET	-	initiating methionine	UNP P61769

- Molecule 5 is 6-methyl-5-[(1E)-3-oxobut-1-en-1-yl]pyrimidine-2,4(1H,3H)-dione (three-letter code: OYG) (formula: C₉H₁₀N₂O₃).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	N	O	0	0
			13	9	2	2		
5	C	1	Total	C	N	O	0	0
			13	9	2	2		

- Molecule 6 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	G	1	Total	Na	0	0
			1	1		
6	F	1	Total	Na	0	0
			1	1		

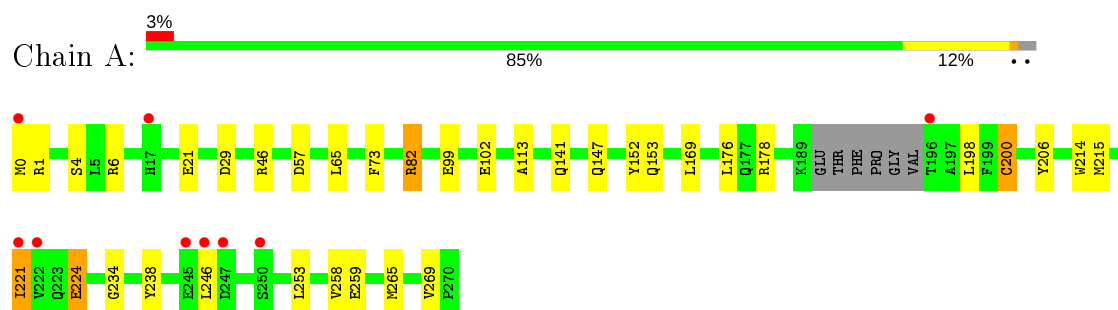
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	378	Total 378	O 378	0	0
7	B	344	Total 344	O 344	0	0
7	C	393	Total 393	O 393	0	0
7	D	176	Total 176	O 176	0	0
7	E	244	Total 244	O 244	0	0
7	F	150	Total 150	O 150	0	0
7	G	417	Total 417	O 417	0	0
7	H	121	Total 121	O 121	0	0

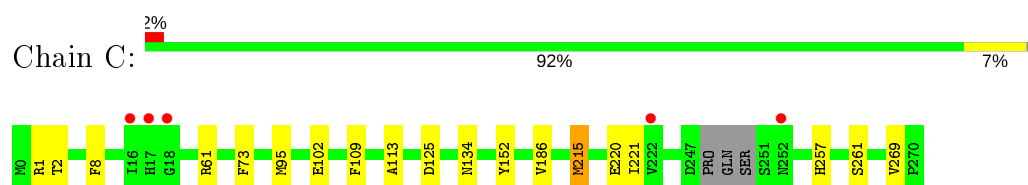
3 Residue-property plots

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.

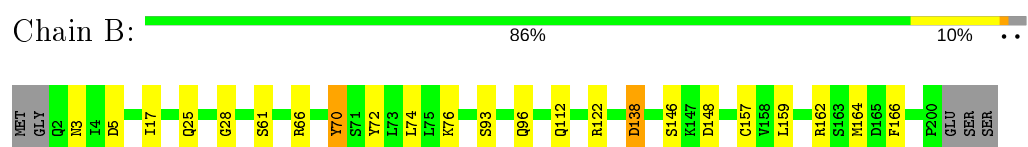
- Molecule 1: Major histocompatibility complex class I-related gene protein



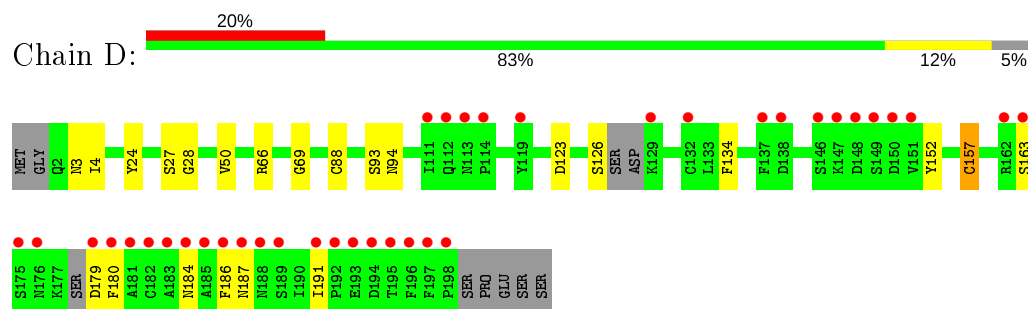
- Molecule 1: Major histocompatibility complex class I-related gene protein



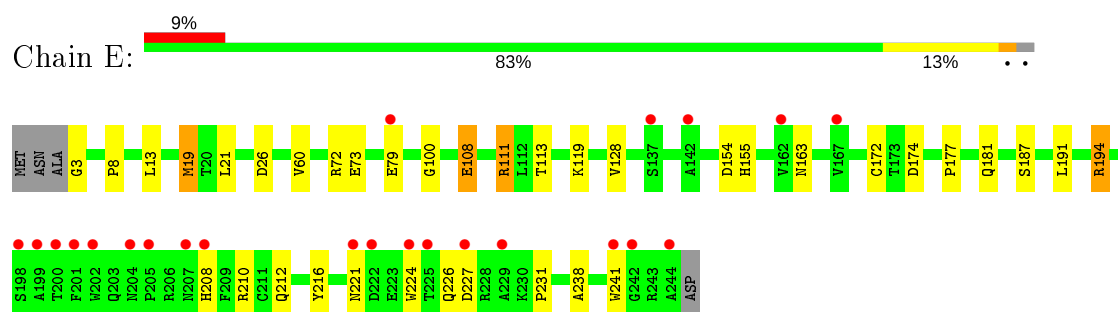
- Molecule 2: Human TCR alpha chain



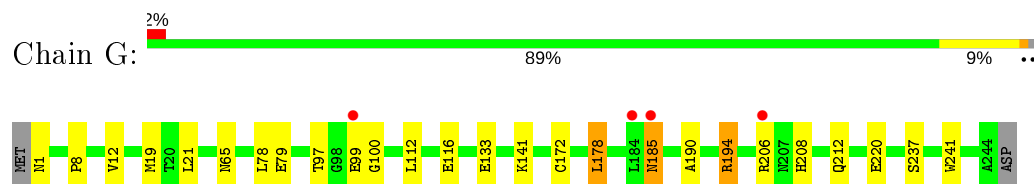
- Molecule 2: Human TCR alpha chain



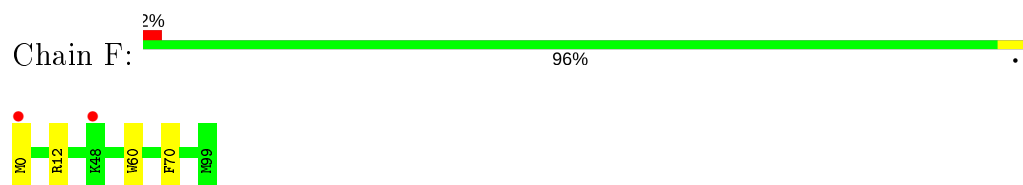
- Molecule 3: Human TCR beta chain



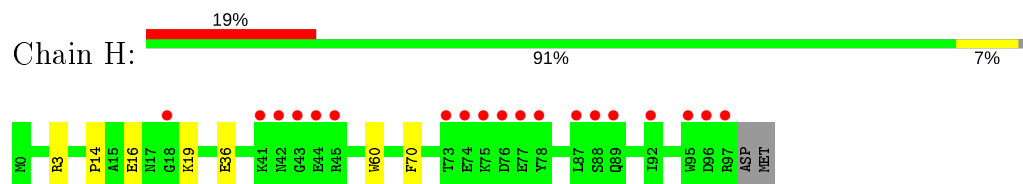
- Molecule 3: Human TCR beta chain



- Molecule 4: Beta-2-microglobulin



- Molecule 4: Beta-2-microglobulin



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	218.48Å 70.56Å 143.98Å 90.00° 104.74° 90.00°	Depositor
Resolution (Å)	48.43 – 1.88 48.43 – 1.88	Depositor EDS
% Data completeness (in resolution range)	99.8 (48.43-1.88) 99.8 (48.43-1.88)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.26 (at 1.88Å)	Xtriage
Refinement program	PHENIX 1.16_3549	Depositor
R, R_{free}	0.160 , 0.195 0.160 , 0.195	Depositor DCC
R_{free} test set	8699 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å ²)	27.3	Xtriage
Anisotropy	0.476	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 50.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	15614	wwPDB-VP
Average B, all atoms (Å ²)	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.61% 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: OYG, NA

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	0.66	2/2347 (0.1%)	0.70	1/3188 (0.0%)
1	C	0.63	0/2411	0.71	0/3276
2	B	0.67	0/1773	0.75	2/2402 (0.1%)
2	D	0.56	0/1582	0.65	0/2146
3	E	0.54	0/2039	0.63	0/2777
3	G	0.69	0/2136	0.72	0/2903
4	F	0.55	0/869	0.68	1/1178 (0.1%)
4	H	0.51	0/819	0.61	0/1114
All	All	0.62	2/13976 (0.0%)	0.69	4/18984 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	200	CYS	CB-SG	-5.58	1.72	1.81
1	A	206	TYR	CD2-CE2	5.05	1.47	1.39

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	178	ARG	NE-CZ-NH1	5.93	123.27	120.30
2	B	66	ARG	NE-CZ-NH1	5.70	123.15	120.30
2	B	66	ARG	NE-CZ-NH2	-5.42	117.59	120.30
4	F	12	ARG	NE-CZ-NH1	-5.20	117.70	120.30

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	2246	0	2141	26	0
1	C	2299	0	2200	16	0
2	B	1674	0	1638	23	0
2	D	1529	0	1430	12	0
3	E	1949	0	1853	28	0
3	G	2033	0	1950	27	0
4	F	840	0	808	2	0
4	H	793	0	743	4	0
5	A	13	0	0	0	0
5	C	13	0	0	0	0
6	F	1	0	0	0	0
6	G	1	0	0	0	0
7	A	378	0	0	15	1
7	B	344	0	0	10	1
7	C	393	0	0	9	1
7	D	176	0	0	1	0
7	E	244	0	0	8	1
7	F	150	0	0	1	0
7	G	417	0	0	12	0
7	H	121	0	0	3	0
All	All	15614	0	12763	128	3

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 5.

The worst 5 of 128 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:B:138[A]:ASP:OD1	7:B:301:HOH:O	1.90	0.86
3:E:73:GLU:OE2	7:E:301:HOH:O	1.97	0.81
3:G:65[B]:ASN:ND2	7:G:405:HOH:O	2.17	0.77
1:A:147:GLN:OE1	7:A:401:HOH:O	2.04	0.76
7:C:933:HOH:O	3:G:206[B]:ARG:HG3	1.86	0.74

All (3) symmetry-related close contacts are listed below. The label for Atom-2 includes the sym-

metry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
7:E:459:HOH:O	7:E:492:HOH:O[4_559]	2.04	0.16
7:B:374:HOH:O	7:B:448:HOH:O[4_548]	2.07	0.13
7:A:576:HOH:O	7:C:909:HOH:O[4_558]	2.15	0.05

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	274/271 (101%)	268 (98%)	6 (2%)	0	100	100
1	C	282/271 (104%)	280 (99%)	2 (1%)	0	100	100
2	B	221/204 (108%)	219 (99%)	2 (1%)	0	100	100
2	D	196/204 (96%)	191 (97%)	5 (3%)	0	100	100
3	E	253/246 (103%)	248 (98%)	5 (2%)	0	100	100
3	G	264/246 (107%)	259 (98%)	5 (2%)	0	100	100
4	F	100/100 (100%)	100 (100%)	0	0	100	100
4	H	97/100 (97%)	96 (99%)	1 (1%)	0	100	100
All	All	1687/1642 (103%)	1661 (98%)	26 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	240/241 (100%)	231 (96%)	9 (4%)	33	21
1	C	250/241 (104%)	245 (98%)	5 (2%)	55	47
2	B	196/181 (108%)	191 (97%)	5 (3%)	46	36
2	D	166/181 (92%)	159 (96%)	7 (4%)	30	18
3	E	213/212 (100%)	205 (96%)	8 (4%)	33	21
3	G	227/212 (107%)	222 (98%)	5 (2%)	52	43
4	F	93/95 (98%)	92 (99%)	1 (1%)	73	70
4	H	85/95 (90%)	83 (98%)	2 (2%)	49	39
All	All	1470/1458 (101%)	1428 (97%)	42 (3%)	46	32

5 of 42 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	261[B]	SER
2	D	179	ASP
3	G	194	ARG
2	D	27	SER
2	D	163	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
2	D	176	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

Of 4 ligands modelled in this entry, 2 are monoatomic - leaving 2 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
5	OYG	A	301	1	12,13,14	1.91	3 (25%)	9,17,19	7.15	4 (44%)
5	OYG	C	801	1	12,13,14	1.70	2 (16%)	9,17,19	7.17	5 (55%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	OYG	A	301	1	-	0/4/4/5	0/1/1/1
5	OYG	C	801	1	-	1/4/4/5	0/1/1/1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	301	OYG	C8-N1	3.75	1.39	1.33
5	A	301	OYG	C5-N	3.66	1.39	1.34
5	C	801	OYG	C8-N1	3.50	1.39	1.33
5	C	801	OYG	C5-N	3.46	1.39	1.34
5	A	301	OYG	C3-C2	2.66	1.40	1.31

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	C	801	OYG	C4-C8-N1	-14.31	114.42	124.40
5	C	801	OYG	C8-N1-C7	13.86	126.84	115.14
5	A	301	OYG	C8-N1-C7	13.81	126.80	115.14
5	A	301	OYG	C4-C8-N1	-12.74	115.52	124.40
5	A	301	OYG	C1-C2-C3	-9.46	115.53	125.86

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	C	801	OYG	C2-C3-C4-C5

There are no ring outliers.

No monomer is involved in short contacts.

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(Å ²)	Q<0.9
1	A	265/271 (97%)	-0.27	9 (3%) 45 46	21, 32, 61, 72	7 (2%)
1	C	268/271 (98%)	-0.23	5 (1%) 66 68	22, 30, 49, 75	11 (4%)
2	B	199/204 (97%)	-0.51	0 100 100	19, 27, 45, 55	8 (4%)
2	D	194/204 (95%)	0.65	40 (20%) 1 1	24, 42, 76, 98	5 (2%)
3	E	242/246 (98%)	0.09	23 (9%) 8 9	26, 42, 74, 89	6 (2%)
3	G	244/246 (99%)	-0.17	4 (1%) 72 74	19, 27, 44, 61	6 (2%)
4	F	100/100 (100%)	-0.15	2 (2%) 65 67	23, 36, 54, 63	5 (5%)
4	H	98/100 (98%)	0.68	19 (19%) 1 1	27, 48, 63, 70	1 (1%)
All	All	1610/1642 (98%)	-0.05	102 (6%) 20 21	19, 33, 66, 98	49 (3%)

The worst 5 of 102 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	180	PHE	5.8
2	D	191	ILE	5.7
3	E	225	THR	5.2
2	D	197	PHE	4.9
2	D	147	LYS	4.9

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 [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
5	OYG	A	301	13/14	0.94	0.10	24,27,30,31	0
5	OYG	C	801	13/14	0.94	0.10	24,27,29,34	0
6	NA	F	101	1/1	0.95	0.21	38,38,38,38	0
6	NA	G	301	1/1	0.99	0.06	28,28,28,28	0

6.5 Other polymers [i](#)

There are no such residues in this entry.