



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 2, 2017 – 12:20 PM EST

PDB ID : 5U17  
Title : Structure of human MR1-DA-6-FP in complex with human MAIT A-F7 TCR  
Authors : Keller, A.N.; Rossjohn, J.  
Deposited on : 2016-11-28  
Resolution : 2.15 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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:

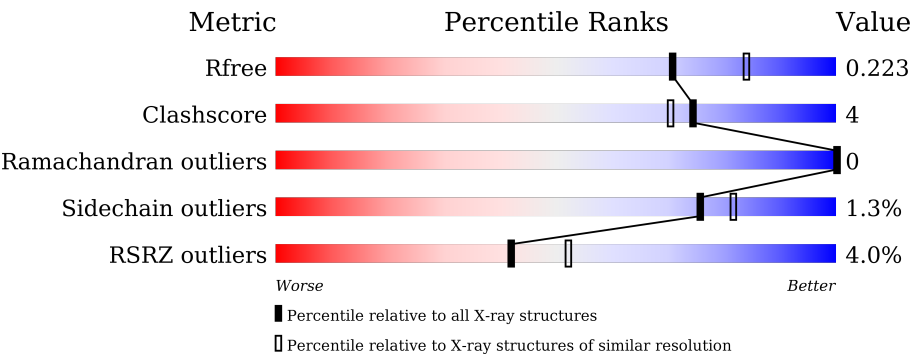
MolProbity : 4.02b-467  
Mogul : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20028442  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20028442

# 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 2.15 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	1045 (2.16-2.16)
Clashscore	102246	1152 (2.16-2.16)
Ramachandran outliers	100387	1131 (2.16-2.16)
Sidechain outliers	100360	1131 (2.16-2.16)
RSRZ outliers	91569	1050 (2.16-2.16)

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	271	<div><div>2%</div><div><div></div><div></div><div></div><div></div></div><div>89%7%..</div></div>
1	C	271	<div><div>5%</div><div><div></div><div></div><div></div><div></div></div><div>86%11%.</div></div>
2	B	203	<div><div>8%</div><div><div></div><div></div><div></div><div></div></div><div>87%9%.</div></div>
2	D	203	<div><div>2%</div><div><div></div><div></div><div></div><div></div></div><div>91%8%.</div></div>
3	E	245	<div><div>%</div><div><div></div><div></div><div></div><div></div></div><div>89%10%</div></div>
3	G	245	<div><div>4%</div><div><div></div><div></div><div></div><div></div></div><div>87%10%..</div></div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
4	F	99	
4	H	99	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
6	GOL	A	302	-	-	-	X
6	GOL	B	301	-	-	-	X
6	GOL	C	302	-	-	-	X
6	GOL	C	303	-	-	-	X
6	GOL	C	304	-	-	-	X
6	GOL	C	305	-	-	-	X
6	GOL	E	301	-	-	-	X
6	GOL	H	102	-	-	-	X
7	ACT	D	301	-	-	-	X
7	ACT	F	101	-	-	-	X

## 2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 13789 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	264	Total	C	N	O	S	0	0	0
			2157	1379	373	395	10			
1	C	262	Total	C	N	O	S	0	1	0
			2152	1375	370	396	11			

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 MAIT T-cell receptor alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	195	Total	C	N	O	S	0	2	0
			1514	958	244	303	9			
2	D	200	Total	C	N	O	S	0	1	0
			1565	986	251	319	9			

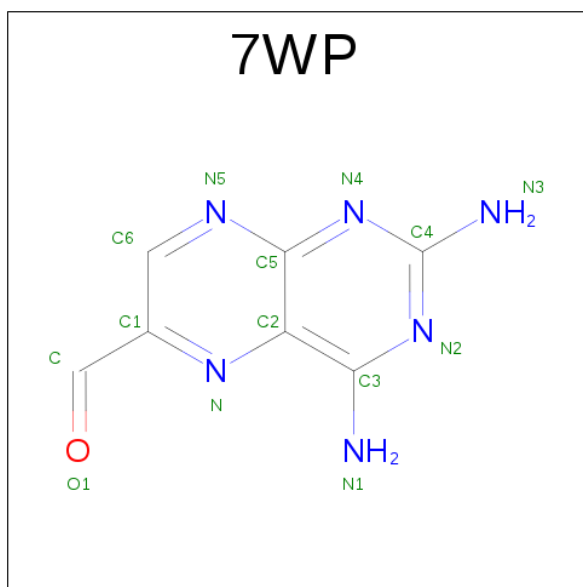
- Molecule 3 is a protein called MAIT T-cell receptor beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	244	Total	C	N	O	S	0	0	0
			1920	1207	334	370	9			
3	G	241	Total	C	N	O	S	0	2	0
			1899	1195	326	369	9			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	F	97	Total	C	N	O	S	0	0	0
			792	507	135	148	2			
4	H	99	Total	C	N	O	S	0	0	0
			817	520	138	157	2			

- Molecule 5 is 2,4-diaminopteridine-6-carbaldehyde (three-letter code: 7WP) (formula:  $C_7H_6N_6O$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	N	0	0
			13	7	6		
5	C	1	Total	C	N	0	0
			13	7	6		

- Molecule 6 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	C	O	0	0
			6	3	3		
6	B	1	Total	C	O	0	0
			6	3	3		
6	C	1	Total	C	O	0	0
			6	3	3		
6	C	1	Total	C	O	0	0
			6	3	3		
6	C	1	Total	C	O	0	0
			6	3	3		
6	C	1	Total	C	O	0	0
			6	3	3		
6	E	1	Total	C	O	0	0
			6	3	3		
6	H	1	Total	C	O	0	0
			6	3	3		
6	H	1	Total	C	O	0	0
			6	3	3		

- Molecule 7 is ACETATE ION (three-letter code: ACT) (formula: C<sub>2</sub>H<sub>3</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	D	1	Total	C	O	0	0
			4	2	2		
7	F	1	Total	C	O	0	0
			4	2	2		

- Molecule 8 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	H	1	Total	Na	0	0
			1	1		
8	G	1	Total	Na	0	0
			1	1		
8	E	1	Total	Na	0	0
			1	1		

- Molecule 9 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	159	Total	O	0	0
			159	159		
9	B	96	Total	O	0	0
			96	96		
9	C	149	Total	O	0	0
			149	149		
9	D	134	Total	O	0	0
			134	134		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	E	141	Total 141	O 141	0	0
9	F	32	Total 32	O 32	0	0
9	G	102	Total 102	O 102	0	0
9	H	69	Total 69	O 69	0	0



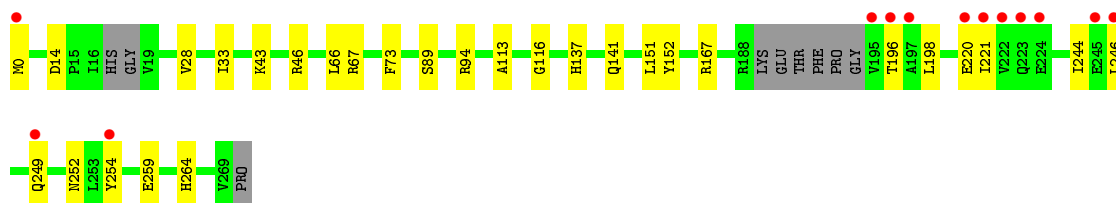
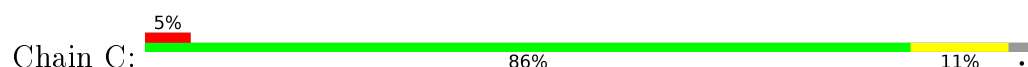
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

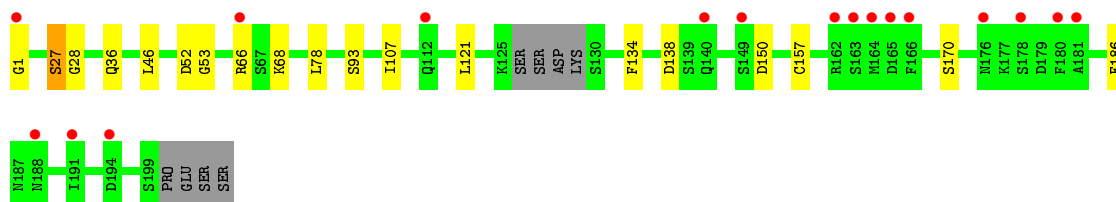
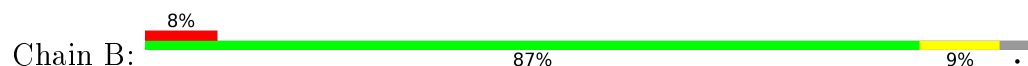
- Molecule 1: Major histocompatibility complex class I-related gene protein



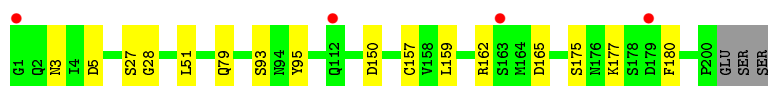
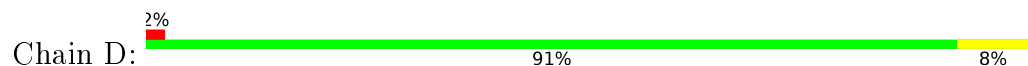
- Molecule 1: Major histocompatibility complex class I-related gene protein



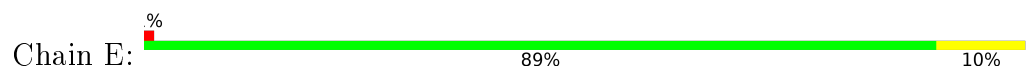
- Molecule 2: MAIT T-cell receptor alpha chain



- Molecule 2: MAIT T-cell receptor alpha chain

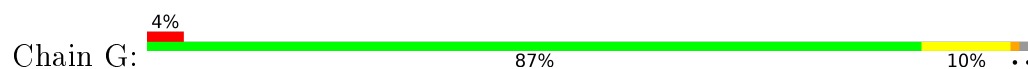


- Molecule 3: MAIT T-cell receptor beta chain

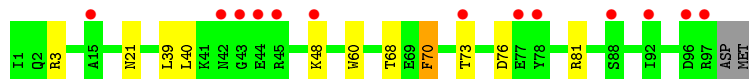
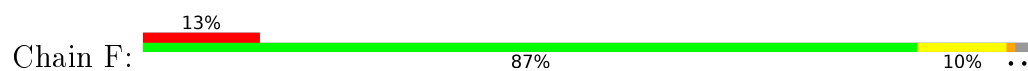




- Molecule 3: MAIT T-cell receptor beta chain



- 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, $\alpha$ , $\beta$ , $\gamma$	215.29Å 69.74Å 142.36Å 90.00° 104.02° 90.00°	Depositor
Resolution (Å)	45.35 – 2.15 45.35 – 2.15	Depositor EDS
% Data completeness (in resolution range)	100.0 (45.35-2.15) 100.0 (45.35-2.15)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.97 (at 2.16Å)	Xtriage
Refinement program	PHENIX	Depositor
R, $R_{free}$	0.180 , 0.224 0.178 , 0.223	Depositor DCC
$R_{free}$ test set	5599 reflections (5.01%)	DCC
Wilson B-factor (Å <sup>2</sup> )	31.7	Xtriage
Anisotropy	0.266	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 44.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	13789	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.84% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, ACT, 7WP, 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	0.52	0/2222	0.64	1/3022 (0.0%)
1	C	0.49	0/2216	0.59	0/3014
2	B	0.46	0/1553	0.57	0/2108
2	D	0.52	0/1600	0.60	0/2170
3	E	0.50	0/1971	0.61	1/2682 (0.0%)
3	G	0.43	0/1950	0.58	1/2657 (0.0%)
4	F	0.40	0/815	0.58	0/1108
4	H	0.48	0/840	0.60	0/1141
All	All	0.48	0/13167	0.60	3/17902 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	66	LEU	CA-CB-CG	-7.56	97.92	115.30
3	G	178	LEU	CA-CB-CG	5.45	127.84	115.30
3	E	194	ARG	NE-CZ-NH2	-5.28	117.66	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	2157	0	2028	18	0
1	C	2152	0	2016	21	0
2	B	1514	0	1413	15	0
2	D	1565	0	1474	12	0
3	E	1920	0	1812	17	0
3	G	1899	0	1766	18	0
4	F	792	0	745	6	0
4	H	817	0	765	7	0
5	A	13	0	0	0	0
5	C	13	0	0	0	0
6	A	6	0	8	1	0
6	B	6	0	8	2	0
6	C	24	0	32	2	0
6	E	6	0	8	1	0
6	H	12	0	16	1	0
7	D	4	0	3	1	0
7	F	4	0	3	0	0
8	E	1	0	0	0	0
8	G	1	0	0	0	0
8	H	1	0	0	0	0
9	A	159	0	0	5	0
9	B	96	0	0	4	0
9	C	149	0	0	2	0
9	D	134	0	0	1	0
9	E	141	0	0	3	0
9	F	32	0	0	0	0
9	G	102	0	0	2	0
9	H	69	0	0	0	0
All	All	13789	0	12097	101	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 4.

All (101) 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)
3:E:206:ARG:HH21	6:E:301:GOL:H31	1.26	0.98
1:A:16:ILE:HG22	1:A:17:HIS:H	1.41	0.85
1:A:1:ARG:O	9:A:401:HOH:O	2.03	0.77
1:C:249:GLN:OE1	9:C:401:HOH:O	2.02	0.77
2:B:150[B]:ASP:OD1	9:B:401:HOH:O	2.03	0.75
2:B:66[B]:ARG:NH2	9:B:402:HOH:O	2.07	0.74

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:16:ILE:HG22	1:A:17:HIS:N	2.03	0.72
1:C:196:THR:HG22	1:C:246:LEU:HD13	1.73	0.71
1:A:269:VAL:O	9:A:402:HOH:O	2.12	0.67
2:B:1:GLY:HA2	2:B:27:SER:H	1.58	0.67
2:D:28:GLY:HA3	2:D:93[B]:SER:OG	1.93	0.67
2:B:52:ASP:OD1	2:B:66[A]:ARG:HD3	1.95	0.66
1:A:155:ASN:OD1	9:A:403:HOH:O	2.14	0.66
3:G:208:HIS:NE2	3:G:239:GLU:OE1	2.27	0.65
1:A:1:ARG:NH2	1:A:178:ARG:O	2.30	0.64
1:A:252:ASN:OD1	1:A:252:ASN:N	2.31	0.64
2:D:157:CYS:HB2	3:E:194:ARG:NH2	2.14	0.62
6:B:301:GOL:O3	9:B:403:HOH:O	2.16	0.61
2:B:53:GLY:HA3	6:B:301:GOL:H11	1.82	0.60
3:E:222:ASP:OD1	9:E:401:HOH:O	2.16	0.60
1:C:116:GLY:O	4:F:3:ARG:NH2	2.36	0.58
2:D:150:ASP:HB2	2:D:177:LYS:HD2	1.85	0.58
2:D:79:GLN:NE2	9:D:402:HOH:O	2.30	0.58
1:C:259:GLU:HG3	1:C:264[B]:HIS:CE1	2.39	0.57
1:C:221:ILE:HD12	1:C:244:ILE:CD1	2.37	0.55
1:C:113:ALA:HB2	4:F:60:TRP:CE2	2.41	0.55
2:B:28:GLY:HA3	2:B:93:SER:HB3	1.87	0.54
1:C:67:ARG:HH12	6:C:304:GOL:H12	1.73	0.54
3:E:30:ASN:O	3:E:51:SER:HA	2.07	0.54
4:F:21:ASN:HB3	4:F:70:PHE:CE2	2.43	0.53
3:G:30:ASN:O	3:G:51:SER:HA	2.09	0.53
3:G:59:GLU:HG2	3:G:60:VAL:HG13	1.91	0.53
1:C:152:TYR:CD1	3:E:100:GLY:HA3	2.44	0.52
3:G:210:ARG:HG3	3:G:239:GLU:HB3	1.91	0.52
2:B:138:ASP:OD1	9:B:404:HOH:O	2.19	0.52
3:G:99[B]:GLU:OE1	3:G:101:SER:OG	2.26	0.52
2:B:170:SER:OG	3:G:194:ARG:NE	2.40	0.52
1:C:252:ASN:HB2	1:C:254:TYR:CE2	2.45	0.52
3:E:8:PRO:HD2	3:E:21:LEU:HD22	1.93	0.51
2:D:159:LEU:HB3	3:E:172:CYS:HB2	1.92	0.51
4:F:40:LEU:HD11	4:F:81:ARG:HB2	1.93	0.51
3:E:131:PRO:HD3	3:E:144:LEU:HG	1.93	0.51
4:H:49:VAL:HG22	4:H:68:THR:CG2	2.41	0.50
3:G:95:VAL:HG12	3:G:96:TRP:CD1	2.46	0.49
2:B:134:PHE:HB2	2:B:186:PHE:CE1	2.47	0.49
2:B:1:GLY:CA	2:B:27:SER:H	2.25	0.49
1:C:43:LYS:HD2	1:C:66:LEU:HD13	1.95	0.49

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:E:212:GLN:OE1	9:E:402:HOH:O	2.19	0.49
4:H:49:VAL:HA	4:H:68:THR:HG22	1.95	0.49
3:E:59:GLU:HG2	3:E:60:VAL:HG13	1.95	0.48
1:A:3:HIS:CD2	1:A:169:LEU:HD11	2.48	0.48
1:C:167:ARG:HH11	6:C:303:GOL:H11	1.78	0.48
2:D:93[B]:SER:HB2	7:D:301:ACT:H3	1.96	0.48
3:E:128:VAL:HG23	3:E:238:ALA:HB3	1.96	0.48
1:C:0:MET:HB2	9:C:521:HOH:O	2.12	0.48
1:A:116:GLY:O	4:H:3:ARG:NH2	2.47	0.47
1:A:113:ALA:HB2	4:H:60:TRP:CE2	2.49	0.47
1:C:259:GLU:HG3	1:C:264[B]:HIS:ND1	2.30	0.47
1:C:221:ILE:HD12	1:C:244:ILE:HD11	1.95	0.47
1:C:14:ASP:HB3	1:C:89:SER:HB2	1.96	0.47
2:B:121:LEU:HB3	3:G:129:PHE:HB3	1.98	0.46
2:B:157:CYS:HB2	3:G:194:ARG:NH2	2.31	0.46
4:H:39:LEU:HD12	4:H:68:THR:HG23	1.97	0.46
2:D:3:ASN:ND2	2:D:5:ASP:OD1	2.41	0.46
1:A:16:ILE:CG2	1:A:17:HIS:H	2.19	0.46
4:F:39:LEU:HD12	4:F:68:THR:OG1	2.16	0.45
1:A:156:TRP:CZ3	1:A:161:CYS:HB2	2.51	0.45
4:H:94:LYS:O	6:H:101:GOL:H11	2.16	0.45
3:G:131:PRO:HD3	3:G:144:LEU:HG	1.99	0.45
1:C:220:GLU:HG2	1:C:221:ILE:H	1.81	0.45
2:B:78:LEU:HD22	2:B:107:ILE:HD11	1.97	0.44
3:G:204:ASN:HB3	3:G:207:ASN:ND2	2.32	0.44
1:A:16:ILE:CG2	1:A:17:HIS:N	2.73	0.44
1:C:46:ARG:HA	1:C:46:ARG:HD3	1.66	0.44
3:E:95:VAL:HG12	3:E:96:TRP:CD1	2.52	0.44
3:G:155:HIS:HB3	3:G:216:TYR:HB2	1.98	0.44
1:A:245:GLU:O	9:A:404:HOH:O	2.21	0.43
2:B:68:LYS:HE3	3:E:220:GLU:OE2	2.18	0.43
3:E:223:GLU:HG3	9:E:526:HOH:O	2.18	0.43
3:G:210:ARG:NH1	9:G:406:HOH:O	2.39	0.43
2:B:36:GLN:HB2	2:B:46:LEU:HD11	2.01	0.43
3:G:13:LEU:HD22	3:G:17:GLN:HE21	1.82	0.43
2:D:175:SER:HB3	2:D:180:PHE:CG	2.54	0.43
1:A:1:ARG:HB3	9:A:546:HOH:O	2.18	0.42
2:D:162:ARG:NH1	2:D:165:ASP:OD1	2.32	0.42
1:C:28:VAL:HG23	1:C:33:ILE:HD13	2.01	0.42
1:A:160:GLU:OE1	6:A:302:GOL:O3	2.35	0.42
1:A:122:PHE:HB2	1:A:129:TRP:CH2	2.55	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:H:45:ARG:NH1	4:H:47:GLU:OE2	2.53	0.42
3:E:224:TRP:CZ2	3:E:226:GLN:HB2	2.56	0.41
4:F:73:THR:OG1	4:F:76:ASP:OD2	2.31	0.41
1:C:151:LEU:HD22	2:D:51:LEU:HD12	2.02	0.41
1:C:137:HIS:O	1:C:141:GLN:HG2	2.21	0.41
1:C:198:LEU:HG	1:C:246:LEU:HD11	2.03	0.41
2:D:162:ARG:NH1	2:D:165:ASP:HA	2.36	0.41
3:G:154:ASP:OD1	3:G:177:PRO:HG3	2.19	0.41
3:G:26:ASP:HB3	9:G:405:HOH:O	2.20	0.41
1:A:152:TYR:CD1	3:G:100:GLY:HA3	2.56	0.40
2:D:95:TYR:CG	3:E:98:GLY:HA2	2.56	0.40
3:E:154:ASP:CG	3:E:177:PRO:HG3	2.42	0.40
3:G:13:LEU:HD22	3:G:17:GLN:NE2	2.35	0.40

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	260/271 (96%)	251 (96%)	9 (4%)	0	100	100
1	C	257/271 (95%)	252 (98%)	5 (2%)	0	100	100
2	B	193/203 (95%)	188 (97%)	5 (3%)	0	100	100
2	D	199/203 (98%)	194 (98%)	5 (2%)	0	100	100
3	E	242/245 (99%)	235 (97%)	7 (3%)	0	100	100
3	G	241/245 (98%)	236 (98%)	5 (2%)	0	100	100
4	F	95/99 (96%)	94 (99%)	1 (1%)	0	100	100
4	H	97/99 (98%)	97 (100%)	0	0	100	100
All	All	1584/1636 (97%)	1547 (98%)	37 (2%)	0	100	100



There are no Ramachandran outliers to report.

### 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	227/241 (94%)	221 (97%)	6 (3%)	54	55
1	C	228/241 (95%)	226 (99%)	2 (1%)	84	89
2	B	165/180 (92%)	164 (99%)	1 (1%)	90	94
2	D	176/180 (98%)	175 (99%)	1 (1%)	90	94
3	E	207/211 (98%)	205 (99%)	2 (1%)	82	87
3	G	204/211 (97%)	201 (98%)	3 (2%)	72	78
4	F	87/94 (93%)	85 (98%)	2 (2%)	58	62
4	H	91/94 (97%)	90 (99%)	1 (1%)	80	85
All	All	1385/1452 (95%)	1367 (99%)	18 (1%)	76	82

All (18) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	73	PHE
1	A	94	ARG
1	A	152	TYR
1	A	169	LEU
1	A	223	GLN
1	A	252	ASN
2	B	27	SER
1	C	73	PHE
1	C	94	ARG
2	D	27	SER
3	E	99	GLU
3	E	194	ARG
4	F	48	LYS
4	F	70	PHE
3	G	112	LEU
3	G	178	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
3	G	194	ARG
4	H	70	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	249	GLN
3	E	212	GLN
3	G	17	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](#)

Of 16 ligands modelled in this entry, 3 are monoatomic - leaving 13 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	7WP	A	301	1	13,14,15	0.74	0	10,20,21	0.64	0
6	GOL	A	302	-	5,5,5	0.30	0	5,5,5	0.48	0
6	GOL	B	301	-	5,5,5	0.35	0	5,5,5	0.82	0
5	7WP	C	301	1	13,14,15	0.54	0	10,20,21	0.58	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
6	GOL	C	302	-	5,5,5	0.51	0	5,5,5	0.87	0
6	GOL	C	303	-	5,5,5	0.28	0	5,5,5	0.53	0
6	GOL	C	304	-	5,5,5	0.31	0	5,5,5	0.31	0
6	GOL	C	305	-	5,5,5	0.36	0	5,5,5	0.25	0
7	ACT	D	301	-	0,3,3	0.00	-	0,3,3	0.00	-
6	GOL	E	301	-	5,5,5	0.52	0	5,5,5	0.49	0
7	ACT	F	101	-	0,3,3	0.00	-	0,3,3	0.00	-
6	GOL	H	101	-	5,5,5	0.30	0	5,5,5	1.01	0
6	GOL	H	102	-	5,5,5	0.34	0	5,5,5	0.49	0

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	7WP	A	301	1	-	0/0/0/2	0/2/2/2
6	GOL	A	302	-	-	0/4/4/4	0/0/0/0
6	GOL	B	301	-	-	0/4/4/4	0/0/0/0
5	7WP	C	301	1	-	0/0/0/2	0/2/2/2
6	GOL	C	302	-	-	0/4/4/4	0/0/0/0
6	GOL	C	303	-	-	0/4/4/4	0/0/0/0
6	GOL	C	304	-	-	0/4/4/4	0/0/0/0
6	GOL	C	305	-	-	0/4/4/4	0/0/0/0
7	ACT	D	301	-	-	0/0/0/0	0/0/0/0
6	GOL	E	301	-	-	0/4/4/4	0/0/0/0
7	ACT	F	101	-	-	0/0/0/0	0/0/0/0
6	GOL	H	101	-	-	0/4/4/4	0/0/0/0
6	GOL	H	102	-	-	0/4/4/4	0/0/0/0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

7 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	302	GOL	1	0
6	B	301	GOL	2	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	C	303	GOL	1	0
6	C	304	GOL	1	0
7	D	301	ACT	1	0
6	E	301	GOL	1	0
6	H	101	GOL	1	0

## 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	264/271 (97%)	0.04	5 (1%) 70 78	23, 32, 64, 106	0
1	C	262/271 (96%)	0.16	13 (4%) 32 43	21, 35, 66, 91	0
2	B	195/203 (96%)	0.45	17 (8%) 13 19	24, 43, 86, 99	0
2	D	200/203 (98%)	0.04	4 (2%) 68 75	24, 34, 61, 79	0
3	E	244/245 (99%)	-0.11	2 (0%) 87 90	25, 36, 59, 87	0
3	G	241/245 (98%)	0.11	10 (4%) 41 51	24, 40, 71, 88	0
4	F	97/99 (97%)	0.69	13 (13%) 4 7	27, 54, 86, 91	0
4	H	99/99 (100%)	0.03	0 100 100	23, 36, 62, 76	0
All	All	1602/1636 (97%)	0.14	64 (3%) 42 52	21, 36, 72, 106	0

All (64) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	195	VAL	7.2
1	C	222	VAL	6.2
1	A	251	SER	5.9
1	C	223	GLN	5.7
1	A	252	ASN	5.6
1	C	246	LEU	4.1
1	C	196	THR	3.9
4	F	77	GLU	3.8
3	G	18	SER	3.8
2	B	149	SER	3.6
2	B	165	ASP	3.5
2	B	1	GLY	3.5
2	B	163	SER	3.5
3	G	221	ASN	3.4
2	B	178	SER	3.4
3	G	225	THR	3.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	B	162	ARG	3.4
4	F	97	ARG	3.2
2	D	1	GLY	3.2
3	E	184	LEU	3.1
3	G	227	ASP	3.1
2	D	179	ASP	3.0
2	B	164	MET	3.0
4	F	15	ALA	3.0
1	C	0	MET	3.0
1	C	249	GLN	2.9
4	F	73	THR	2.8
3	G	223	GLU	2.8
3	G	224	TRP	2.8
1	C	245	GLU	2.8
4	F	45	ARG	2.8
2	B	112	GLN	2.7
1	C	221	ILE	2.7
4	F	42	ASN	2.7
4	F	48	LYS	2.7
2	B	194	ASP	2.7
4	F	43	GLY	2.6
1	C	224	GLU	2.6
2	B	140	GLN	2.5
4	F	92	ILE	2.5
2	B	181	ALA	2.5
2	B	180	PHE	2.5
1	A	191	THR	2.4
3	G	79	GLU	2.4
2	D	112	GLN	2.4
4	F	78	TYR	2.4
1	A	195	VAL	2.4
4	F	44	GLU	2.4
2	B	188	ASN	2.4
3	G	219	SER	2.4
2	B	191	ILE	2.3
2	B	66[A]	ARG	2.3
1	C	220	GLU	2.3
1	C	254	TYR	2.3
3	G	220	GLU	2.3
2	D	163	SER	2.3
1	C	197	ALA	2.2
2	B	166	PHE	2.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
4	F	96	ASP	2.2
2	B	176	ASN	2.1
1	A	222	VAL	2.1
3	G	205	PRO	2.1
3	E	185	ASN	2.1
4	F	88	SER	2.1

## 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](#)

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
7	ACT	F	101	4/4	0.89	0.33	6.28	57,58,58,58	0
6	GOL	H	102	6/6	0.88	0.25	4.73	72,72,72,73	0
6	GOL	C	304	6/6	0.93	0.14	4.46	40,43,45,47	0
7	ACT	D	301	4/4	0.80	0.19	4.22	56,56,60,62	0
6	GOL	C	305	6/6	0.94	0.18	3.66	52,54,56,58	0
6	GOL	A	302	6/6	0.86	0.21	3.55	66,67,68,68	0
6	GOL	B	301	6/6	0.87	0.19	3.26	47,51,52,52	0
6	GOL	C	303	6/6	0.88	0.17	2.82	46,47,48,49	0
6	GOL	E	301	6/6	0.93	0.17	2.61	44,47,47,49	0
6	GOL	C	302	6/6	0.94	0.13	2.55	32,34,43,46	0
6	GOL	H	101	6/6	0.95	0.14	1.92	31,35,36,36	0
8	NA	H	103	1/1	0.87	0.14	0.13	53,53,53,53	0
5	7WP	A	301	13/14	0.96	0.15	-0.09	23,28,30,31	0
8	NA	G	301	1/1	0.93	0.11	-0.51	43,43,43,43	0
5	7WP	C	301	13/14	0.95	0.12	-0.62	20,26,31,31	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
8	NA	E	302	1/1	0.98	0.08	-1.56	40,40,40,40	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.