



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 07:39 PM GMT

PDB ID : 4PJX  
Title : Structure of human MR1-Ac-6-FP in complex with human MAIT C-A11 TCR  
Authors : Birkinshaw, R.W.; Rossjohn, J.  
Deposited on : 2014-05-12  
Resolution : 2.25 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 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
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) : trunk26865

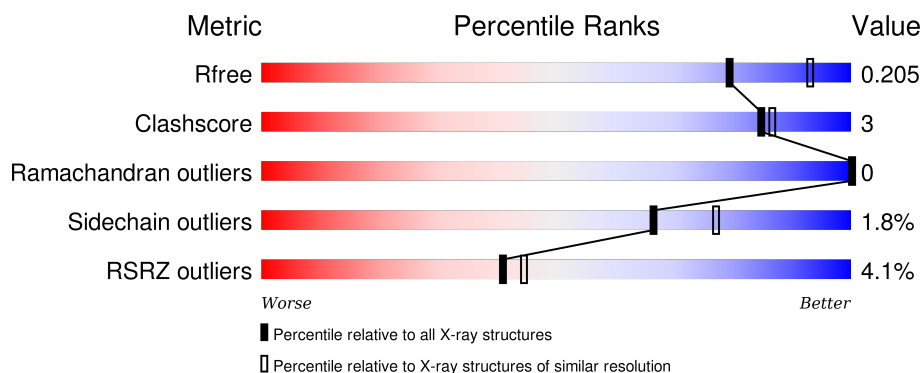
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.25 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	1640 (2.28-2.24)
Clashscore	102246	1095 (2.26-2.26)
Ramachandran outliers	100387	1063 (2.26-2.26)
Sidechain outliers	100360	1063 (2.26-2.26)
RSRZ outliers	91569	1647 (2.28-2.24)

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>88%</div> <div>8%</div> <div>.</div> </div>
1	C	271	<div> <div>3%</div> <div>83%</div> <div>9%</div> <div>8%</div> </div>
2	B	100	<div> <div>%</div> <div>95%</div> <div>.</div> <div>.</div> </div>
2	D	100	<div> <div>13%</div> <div>91%</div> <div>5%</div> <div>.</div> </div>
3	E	205	<div> <div>5%</div> <div>85%</div> <div>6%</div> <div>9%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	G	205	
4	F	246	
4	H	246	

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	CL	A	302	-	-	-	X
7	GOL	B	101	-	-	-	X
7	GOL	B	102[A]	-	-	-	X
7	GOL	B	102[B]	-	-	-	X
7	GOL	C	302	-	-	-	X
7	GOL	C	303	-	-	-	X
7	GOL	F	301	-	-	-	X

## 2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 13290 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	262	Total	C	N	O	S	0	4	0
			2179	1396	379	393	11			
1	C	250	Total	C	N	O	S	0	0	0
			2026	1295	351	369	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	engineered mutation	UNP Q95460
C	0	MET	-	initiating methionine	UNP Q95460
C	261	SER	CYS	engineered mutation	UNP Q95460

- 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
			796	508	136	149	3			
2	D	96	Total	C	N	O	S	0	0	0
			732	471	123	136	2			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	0	MET	-	initiating methionine	UNP P61769
D	0	MET	-	initiating methionine	UNP P61769

- Molecule 3 is a protein called TCR-alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	187	Total	C	N	O	S	0	0	0
			1397	897	225	267	8			

*Continued on next page...*

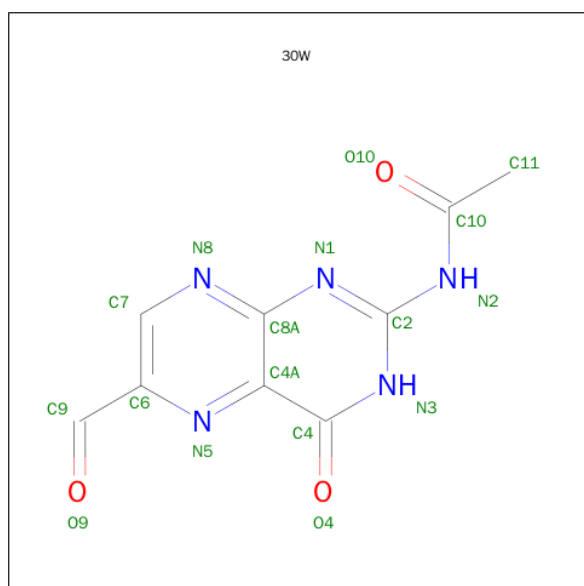
Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	G	198	Total	C	N	O	S	0	0	0
			1508	957	243	299	9			

- Molecule 4 is a protein called TCR-beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	F	237	Total	C	N	O	S	0	0	0
			1789	1131	310	339	9			
4	H	241	Total	C	N	O	S	0	0	0
			1840	1161	318	352	9			

- Molecule 5 is N-(6-formyl-4-oxo-3,4-dihydropteridin-2-yl)acetamide (three-letter code: 30W) (formula: C<sub>9</sub>H<sub>7</sub>N<sub>5</sub>O<sub>3</sub>).

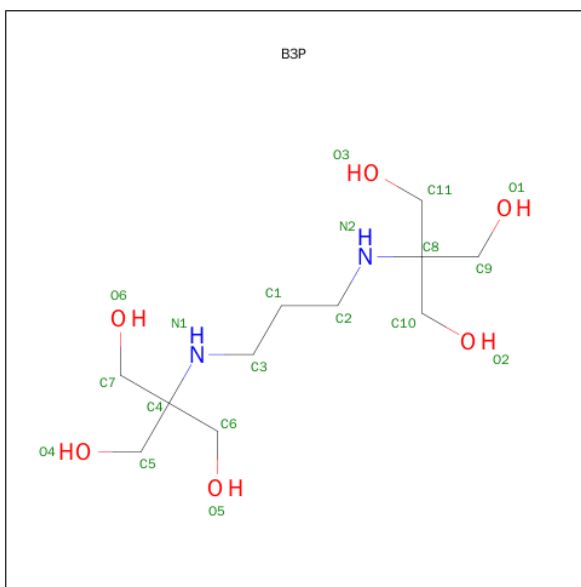


- Molecule 7 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total	C	O	0	0
			6	3	3		
7	B	1	Total	C	O	0	0
			6	3	3		
7	B	1	Total	C	O	0	1
			12	6	6		
7	C	1	Total	C	O	0	0
			6	3	3		
7	C	1	Total	C	O	0	0
			6	3	3		
7	F	1	Total	C	O	0	0
			6	3	3		

- Molecule 8 is 2-[3-(2-HYDROXY-1,1-DIHYDROXYMETHYL-ETHYLAMINO)-PROPYL AMINO]-2-HYDROXYMETHYL-PROPANE-1,3-DIOL (three-letter code: B3P) (formula:  $C_{11}H_{26}N_2O_6$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
8	A	1	Total	C	N	O	0	0
			19	11	2	6		
8	C	1	Total	C	N	O	0	0
			19	11	2	6		

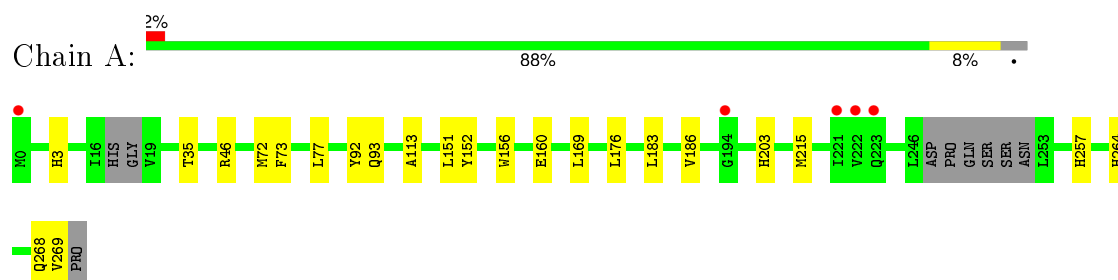
- Molecule 9 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	197	Total	O	0	0
			197	197		
9	B	79	Total	O	0	0
			79	79		
9	C	139	Total	O	0	0
			139	139		
9	D	34	Total	O	0	0
			34	34		
9	E	91	Total	O	0	0
			91	91		
9	F	92	Total	O	0	0
			92	92		
9	G	134	Total	O	0	0
			134	134		
9	H	144	Total	O	0	0
			144	144		

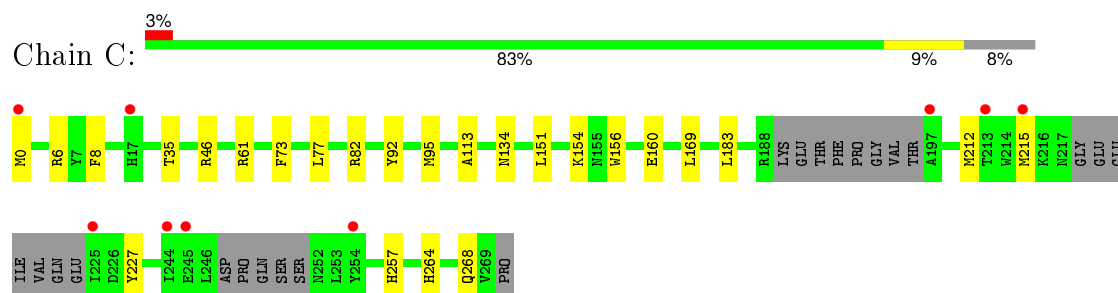
### 3 Residue-property plots [i](#)

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.

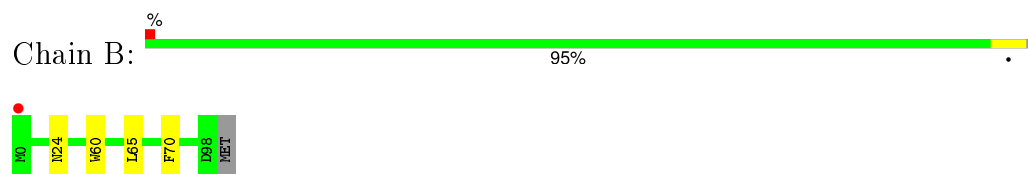
- Molecule 1: Major histocompatibility complex class I-related gene protein



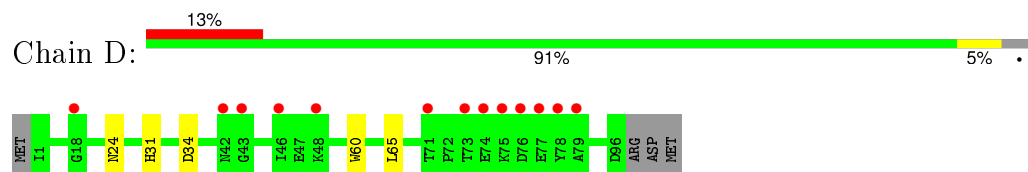
- Molecule 1: Major histocompatibility complex class I-related gene protein



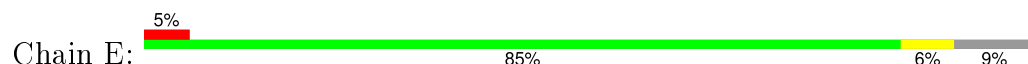
- Molecule 2: Beta-2-microglobulin



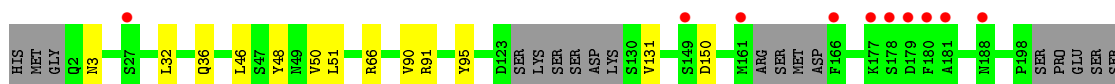
- Molecule 2: Beta-2-microglobulin



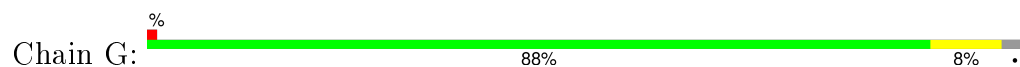
- Molecule 3: TCR-alpha



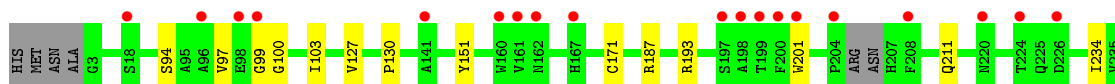
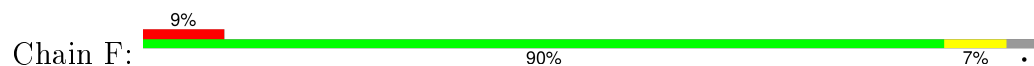




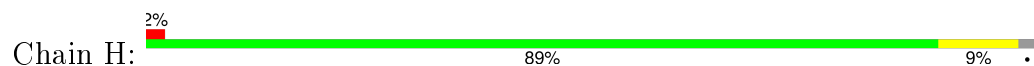
- Molecule 3: TCR-alpha



- Molecule 4: TCR-beta



- Molecule 4: TCR-beta



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	213.92Å 69.58Å 142.50Å 90.00° 103.74° 90.00°	Depositor
Resolution (Å)	32.94 – 2.25 32.65 – 2.25	Depositor EDS
% Data completeness (in resolution range)	99.8 (32.94-2.25) 99.9 (32.65-2.25)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.11 (at 2.24Å)	Xtriage
Refinement program	BUSTER 2.10.0	Depositor
R, $R_{free}$	0.170 , 0.203 0.174 , 0.205	Depositor DCC
$R_{free}$ test set	4840 reflections (5.26%)	DCC
Wilson B-factor (Å <sup>2</sup> )	36.0	Xtriage
Anisotropy	0.486	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 60.3	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	1 of 96819 reflections (0.001%)	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	13290	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	46.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.55% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: GOL, CL, B3P, 30W

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	0.53	0/2243	0.65	0/3047
1	C	0.53	0/2086	0.65	0/2837
2	B	0.52	0/819	0.64	0/1116
2	D	0.45	0/755	0.64	0/1037
3	E	0.54	0/1429	0.67	0/1946
3	G	0.54	0/1542	0.68	0/2097
4	F	0.48	0/1835	0.70	0/2506
4	H	0.50	0/1889	0.69	0/2579
All	All	0.52	0/12598	0.67	0/17165

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	2179	0	2052	18	0
1	C	2026	0	1875	19	0
2	B	796	0	734	2	0
2	D	732	0	633	3	0
3	E	1397	0	1257	6	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	G	1508	0	1386	12	0
4	F	1789	0	1631	7	0
4	H	1840	0	1701	15	0
5	A	16	0	6	0	0
5	C	16	0	6	0	0
6	A	1	0	0	1	0
7	A	6	0	8	0	0
7	B	18	0	24	1	0
7	C	12	0	16	2	0
7	F	6	0	8	0	0
8	A	19	0	26	0	0
8	C	19	0	26	0	0
9	A	197	0	0	5	0
9	B	79	0	0	1	0
9	C	139	0	0	4	0
9	D	34	0	0	1	0
9	E	91	0	0	0	0
9	F	92	0	0	1	0
9	G	134	0	0	2	0
9	H	144	0	0	0	0
All	All	13290	0	11389	69	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 3.

All (69) 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:G:164:MET:HE1	4:H:140:LYS:HE3	1.18	1.16
1:C:257:HIS:HD2	9:C:539:HOH:O	1.30	1.11
1:C:257:HIS:CD2	9:C:539:HOH:O	2.08	1.00
1:A:264:HIS:CE1	9:A:597:HOH:O	2.25	0.89
1:C:0:MET:HE2	1:C:169:LEU:HD13	1.59	0.82
3:G:164:MET:CE	4:H:140:LYS:HE3	2.08	0.81
9:G:434:HOH:O	4:H:99:GLY:HA2	1.79	0.80
3:G:159:LEU:HB3	4:H:171:CYS:HB2	1.78	0.65
1:A:93:GLN:OE1	6:A:302:CL:CL	2.54	0.63
1:C:151:LEU:HD22	3:G:51:LEU:HD12	1.82	0.61
1:C:35:THR:HB	1:C:46:ARG:HD2	1.84	0.60
3:G:28:GLY:HA3	3:G:93:SER:HB3	1.83	0.60
3:G:164:MET:HE1	4:H:140:LYS:CE	2.13	0.60

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:H:127:VAL:HG23	4:H:237:ALA:HB3	1.84	0.58
1:C:264:HIS:CE1	9:C:539:HOH:O	2.56	0.58
4:F:127:VAL:HG23	4:F:237:ALA:HB3	1.89	0.55
3:G:91:ARG:HG2	3:G:95:TYR:HA	1.89	0.54
1:A:151:LEU:HD22	3:E:51:LEU:HD12	1.90	0.54
1:A:72:MET:HG2	9:F:423:HOH:O	2.08	0.54
1:A:46[B]:ARG:NE	9:A:587:HOH:O	2.41	0.54
1:C:215:MET:HG3	1:C:257:HIS:CE1	2.43	0.53
1:A:113:ALA:HB2	2:B:60:TRP:CE2	2.44	0.53
1:C:113:ALA:HB2	2:D:60:TRP:CE2	2.44	0.52
1:C:8:PHE:CE2	1:C:95:MET:HG3	2.44	0.52
1:A:169[A]:LEU:HD23	1:A:176:LEU:HD13	1.93	0.50
2:D:31:HIS:HD2	9:D:125:HOH:O	1.95	0.49
1:C:35:THR:HB	1:C:46:ARG:CD	2.42	0.49
1:A:3:HIS:CD2	1:A:169[B]:LEU:HD21	2.48	0.49
1:A:186:VAL:HG11	1:A:269:VAL:HG22	1.94	0.49
9:G:434:HOH:O	4:H:100:GLY:N	2.36	0.49
1:A:152:TYR:CD1	4:F:99:GLY:HA3	2.47	0.49
1:C:160:GLU:OE2	9:C:533:HOH:O	2.20	0.48
1:A:264:HIS:NE2	9:A:597:HOH:O	2.36	0.47
1:A:203:HIS:HE1	9:B:244:HOH:O	1.97	0.47
3:E:36:GLN:HB2	3:E:46:LEU:HD11	1.97	0.46
2:D:24:ASN:HB3	2:D:65:LEU:HD11	1.98	0.46
1:C:154:LYS:HD3	3:G:51:LEU:HD11	1.96	0.46
4:H:217:LEU:HD22	4:H:230:PRO:HD2	1.97	0.46
4:H:153:ASP:CG	4:H:176:PRO:HG3	2.37	0.46
4:H:130:PRO:HD3	4:H:143:LEU:HG	1.98	0.45
3:E:48:TYR:CE2	4:F:100:GLY:HA2	2.52	0.45
1:A:35:THR:CG2	1:A:46[A]:ARG:HD3	2.47	0.45
4:F:151:TYR:HB2	4:F:187:ARG:HG2	1.99	0.45
1:C:212:MET:HB2	7:C:303:GOL:H12	1.99	0.44
4:F:94:SER:HB3	4:F:103:ILE:HD13	1.99	0.44
3:G:36:GLN:HB2	3:G:46:LEU:HD11	2.00	0.43
4:F:211:GLN:HG3	4:F:234:ILE:HG23	2.00	0.43
1:A:264:HIS:HE1	9:A:597:HOH:O	1.82	0.43
1:A:215:MET:HG3	1:A:257:HIS:CD2	2.52	0.43
2:B:24:ASN:HB3	2:B:65:LEU:HD11	2.01	0.43
1:C:61:ARG:HH22	4:H:97:VAL:HG11	1.84	0.43
4:F:130:PRO:HD2	4:F:201:TRP:CZ2	2.54	0.43
1:C:215:MET:HG3	1:C:257:HIS:ND1	2.34	0.43
3:E:50:VAL:O	3:E:66:ARG:HD3	2.20	0.42

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:E:91:ARG:HG2	3:E:95:TYR:HA	2.01	0.42
4:H:154:HIS:HB3	4:H:215:TYR:HB2	2.02	0.42
1:A:203:HIS:HD2	9:A:592:HOH:O	2.03	0.42
4:H:211:GLN:HG3	4:H:234:ILE:HG23	2.01	0.42
7:B:101:GOL:O2	1:C:134:ASN:ND2	2.53	0.42
1:A:156:TRP:HA	1:A:160:GLU:HB2	2.02	0.41
1:C:77:LEU:HD13	1:C:92:TYR:HB2	2.02	0.41
1:C:227:TYR:HD1	7:C:303:GOL:H11	1.85	0.41
1:C:156:TRP:HA	1:C:160:GLU:HB2	2.02	0.41
4:H:149:GLY:O	4:H:187:ARG:HD2	2.21	0.41
3:E:32:LEU:HD12	3:E:90:VAL:HG22	2.03	0.40
1:A:77:LEU:HD13	1:A:92:TYR:HB2	2.03	0.40
3:G:32:LEU:HD12	3:G:90:VAL:HG22	2.02	0.40
3:G:50:VAL:O	3:G:66:ARG:HD3	2.21	0.40
3:G:122:ARG:HB2	4:H:129:GLU:HB2	2.04	0.40

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	260/271 (96%)	252 (97%)	8 (3%)	0	100	100
1	C	242/271 (89%)	236 (98%)	6 (2%)	0	100	100
2	B	97/100 (97%)	97 (100%)	0	0	100	100
2	D	94/100 (94%)	94 (100%)	0	0	100	100
3	E	181/205 (88%)	179 (99%)	2 (1%)	0	100	100
3	G	196/205 (96%)	194 (99%)	2 (1%)	0	100	100
4	F	233/246 (95%)	229 (98%)	4 (2%)	0	100	100
4	H	239/246 (97%)	236 (99%)	3 (1%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	1542/1644 (94%)	1517 (98%)	25 (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%)	224 (99%)	3 (1%)	76	85
1	C	209/241 (87%)	204 (98%)	5 (2%)	57	67
2	B	86/95 (90%)	85 (99%)	1 (1%)	78	87
2	D	72/95 (76%)	71 (99%)	1 (1%)	74	84
3	E	137/182 (75%)	134 (98%)	3 (2%)	60	70
3	G	159/182 (87%)	156 (98%)	3 (2%)	65	75
4	F	178/211 (84%)	174 (98%)	4 (2%)	60	70
4	H	190/211 (90%)	188 (99%)	2 (1%)	80	88
All	All	1258/1458 (86%)	1236 (98%)	22 (2%)	66	79

All (22) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	73	PHE
1	A	183	LEU
1	A	268	GLN
2	B	70	PHE
1	C	6	ARG
1	C	73	PHE
1	C	82	ARG
1	C	183	LEU
1	C	268	GLN
2	D	34	ASP
3	E	3	ASN
3	E	131	VAL

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
3	E	150	ASP
4	F	97	VAL
4	F	171	CYS
4	F	193	ARG
4	F	236	SER
3	G	6	GLN
3	G	125	LYS
3	G	146	SER
4	H	17	GLN
4	H	193	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	93	GLN
2	B	31	HIS
1	C	111	GLN
1	C	134	ASN
1	C	257	HIS
3	E	120	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 12 ligands modelled in this entry, 1 is monoatomic - leaving 11 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	30W	A	301	1	16,17,18	1.93	4 (25%)	19,24,25	3.26	8 (42%)
7	GOL	A	303	-	5,5,5	0.21	0	5,5,5	0.52	0
8	B3P	A	304	-	18,18,18	0.53	0	23,23,23	1.32	4 (17%)
7	GOL	B	101	-	5,5,5	0.25	0	5,5,5	0.47	0
7	GOL	B	102[A]	-	5,5,5	0.19	0	5,5,5	0.35	0
7	GOL	B	102[B]	-	5,5,5	0.19	0	5,5,5	0.31	0
5	30W	C	301	1	16,17,18	1.79	3 (18%)	19,24,25	3.44	7 (36%)
7	GOL	C	302	-	5,5,5	0.15	0	5,5,5	0.21	0
7	GOL	C	303	-	5,5,5	0.13	0	5,5,5	0.23	0
8	B3P	C	304	-	18,18,18	0.50	0	23,23,23	1.72	3 (13%)
7	GOL	F	301	-	5,5,5	0.15	0	5,5,5	0.34	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	30W	A	301	1	-	0/4/4/6	0/2/2/2
7	GOL	A	303	-	-	0/4/4/4	0/0/0/0
8	B3P	A	304	-	-	0/28/28/28	0/0/0/0
7	GOL	B	101	-	-	0/4/4/4	0/0/0/0
7	GOL	B	102[A]	-	-	0/4/4/4	0/0/0/0
7	GOL	B	102[B]	-	-	0/4/4/4	0/0/0/0
5	30W	C	301	1	-	0/4/4/6	0/2/2/2
7	GOL	C	302	-	-	0/4/4/4	0/0/0/0
7	GOL	C	303	-	-	0/4/4/4	0/0/0/0
8	B3P	C	304	-	-	0/28/28/28	0/0/0/0
7	GOL	F	301	-	-	0/4/4/4	0/0/0/0

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	301	30W	C2-N2	-4.81	1.33	1.39
5	C	301	30W	C2-N2	-4.79	1.33	1.39
5	A	301	30W	C2-N1	2.27	1.42	1.34

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	301	30W	C4-N3	2.73	1.38	1.33
5	C	301	30W	C4-N3	2.75	1.38	1.33
5	C	301	30W	C4-C4A	3.17	1.47	1.41
5	A	301	30W	C4-C4A	3.80	1.48	1.41

All (22) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	C	301	30W	C4A-C4-N3	-6.05	115.31	123.59
5	C	301	30W	N1-C2-N3	-5.71	117.44	126.22
5	A	301	30W	N1-C2-N3	-5.67	117.51	126.22
5	A	301	30W	C4A-C4-N3	-5.45	116.13	123.59
5	C	301	30W	C2-N2-C10	-5.17	122.13	130.17
5	A	301	30W	C2-N2-C10	-4.76	122.77	130.17
5	C	301	30W	C4-C4A-C8A	-3.24	117.86	119.94
8	A	304	B3P	C5-C4-N1	-3.05	100.30	109.19
5	A	301	30W	C4-C4A-C8A	-2.33	118.45	119.94
8	A	304	B3P	C2-N2-C8	2.01	119.07	116.07
5	A	301	30W	N8-C8A-N1	2.15	119.22	116.14
8	C	304	B3P	C7-C4-N1	2.41	116.23	109.19
8	A	304	B3P	C7-C4-N1	2.47	116.41	109.19
8	A	304	B3P	C3-N1-C4	3.64	121.49	116.07
8	C	304	B3P	C2-N2-C8	4.01	122.04	116.07
5	C	301	30W	N2-C2-N3	4.24	129.54	116.44
5	A	301	30W	N2-C2-N3	4.38	129.96	116.44
5	C	301	30W	C2-N1-C8A	5.04	121.16	115.09
5	A	301	30W	C2-N1-C8A	5.29	121.46	115.09
8	C	304	B3P	C3-N1-C4	5.94	124.92	116.07
5	A	301	30W	C4-N3-C2	7.37	126.03	115.31
5	C	301	30W	C4-N3-C2	8.13	127.12	115.31

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	B	101	GOL	1	0
7	C	303	GOL	2	0

## 5.7 Other polymers

There are no such residues in this entry.

## 5.8 Polymer linkage issues

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	262/271 (96%)	-0.08	5 (1%) 70 74	25, 37, 62, 79	1 (0%)
1	C	250/271 (92%)	-0.06	9 (3%) 46 50	27, 41, 72, 92	1 (0%)
2	B	99/100 (99%)	-0.06	1 (1%) 84 85	27, 41, 66, 74	0
2	D	96/100 (96%)	0.42	13 (13%) 4 4	34, 63, 100, 110	0
3	E	187/205 (91%)	0.14	10 (5%) 30 33	26, 52, 93, 116	0
3	G	198/205 (96%)	-0.26	2 (1%) 84 85	28, 39, 64, 83	0
4	F	237/246 (96%)	0.25	21 (8%) 12 13	30, 49, 96, 125	2 (0%)
4	H	241/246 (97%)	-0.22	4 (1%) 73 77	29, 41, 62, 81	2 (0%)
All	All	1570/1644 (95%)	-0.01	65 (4%) 41 44	25, 42, 83, 125	6 (0%)

All (65) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	F	201	TRP	5.6
4	H	182	ALA	5.4
4	F	200	PHE	5.1
4	F	240	TRP	5.0
1	A	222	VAL	4.5
2	B	0	MET	4.5
4	F	161	VAL	3.9
4	F	198	ALA	3.8
2	D	76	ASP	3.7
4	F	224	THR	3.6
4	F	199	THR	3.5
1	C	254	TYR	3.5
4	F	141	ALA	3.4
2	D	78	TYR	3.4
1	A	0	MET	3.3
2	D	42	ASN	3.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
4	H	184	ASN	3.3
2	D	71	THR	3.3
3	E	166	PHE	3.2
1	C	215	MET	3.2
4	H	97	VAL	3.1
2	D	73	THR	3.1
3	E	178	SER	3.0
3	G	97	LEU	3.0
2	D	48	LYS	2.9
4	F	162	ASN	2.8
3	E	179	ASP	2.8
1	C	0	MET	2.7
3	G	95	TYR	2.7
3	E	177	LYS	2.7
1	C	244	ILE	2.7
2	D	79	ALA	2.7
4	F	220	ASN	2.7
2	D	43	GLY	2.7
2	D	77	GLU	2.6
1	A	221	ILE	2.6
2	D	75	LYS	2.6
3	E	161	MET	2.5
2	D	46	ILE	2.5
1	C	213	THR	2.4
4	F	167	HIS	2.4
1	C	245	GLU	2.4
4	F	98	GLU	2.4
4	F	160	TRP	2.4
3	E	188	ASN	2.3
2	D	18	GLY	2.3
1	C	17	HIS	2.3
4	F	99	GLY	2.3
4	H	99	GLY	2.3
3	E	149	SER	2.3
4	F	197	SER	2.2
4	F	96	ALA	2.2
2	D	74	GLU	2.2
4	F	208	PHE	2.2
3	E	180	PHE	2.2
4	F	226	ASP	2.2
4	F	204	PRO	2.2
3	E	181	ALA	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	225	ILE	2.1
1	A	194	GLY	2.1
4	F	18	SER	2.1
4	F	239	ALA	2.0
3	E	27	SER	2.0
1	A	223	GLN	2.0
1	C	197	ALA	2.0

## 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	GOL	B	102[B]	6/6	0.90	0.65	23.64	10,14,14,15	6
7	GOL	B	102[A]	6/6	0.90	0.65	16.83	31,35,35,36	6
7	GOL	C	302	6/6	0.91	0.20	6.60	65,70,72,72	0
6	CL	A	302	1/1	1.00	0.17	4.30	48,48,48,48	0
7	GOL	B	101	6/6	0.95	0.19	3.94	37,40,42,46	0
7	GOL	F	301	6/6	0.83	0.20	3.30	69,74,74,75	0
7	GOL	C	303	6/6	0.90	0.23	2.86	61,70,73,73	0
8	B3P	A	304	19/19	0.86	0.26	1.34	36,42,69,69	0
7	GOL	A	303	6/6	0.94	0.14	1.13	40,43,48,49	0
8	B3P	C	304	19/19	0.86	0.21	1.00	35,52,82,84	0
5	30W	A	301	16/17	0.98	0.18	-0.06	28,30,34,36	0
5	30W	C	301	16/17	0.98	0.15	-0.24	27,32,36,36	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.