



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 15, 2017 – 01:28 am GMT

PDB ID : 5D7I  
Title : Structure of human MR1-Ac-6-FP in complex with human MAIT M33.64 TCR  
Authors : Keller, A.N.; Woolley, R.E.; Rossjohn, J.  
Deposited on : 2015-08-14  
Resolution : 2.00 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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : trunk28620  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
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) : recalc28949

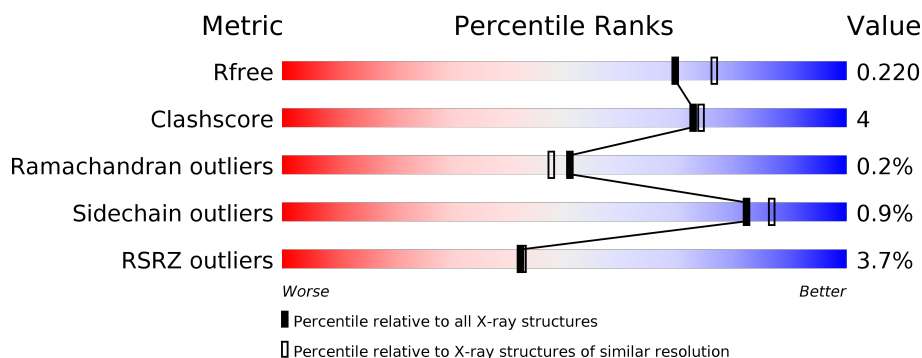
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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.



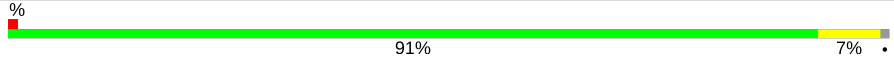
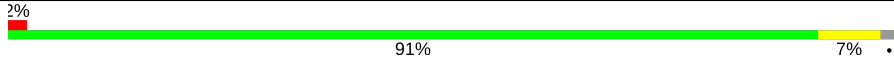
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	100719	6609 (2.00-2.00)
Clashscore	112137	7775 (2.00-2.00)
Ramachandran outliers	110173	7679 (2.00-2.00)
Sidechain outliers	110143	7678 (2.00-2.00)
RSRZ outliers	101464	6696 (2.00-2.00)

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>3%</div> <div> <div></div> <div>86%</div> <div>11%</div> <div></div> </div> <div></div> </div>
1	C	271	<div> <div>4%</div> <div> <div></div> <div>92%</div> <div>7%</div> <div></div> </div> <div></div> </div>
2	B	100	<div> <div>%</div> <div> <div></div> <div>96%</div> <div></div> </div> <div></div> </div>
2	D	100	<div> <div>21%</div> <div> <div></div> <div>88%</div> <div>10%</div> <div></div> </div> <div></div> </div>
3	E	204	<div> <div>2%</div> <div> <div></div> <div>90%</div> <div>8%</div> <div></div> </div> <div></div> </div>
3	G	204	<div> <div>3%</div> <div> <div></div> <div>87%</div> <div>9%</div> <div></div> </div> <div></div> </div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
4	F	245	 % 91% 7% •
4	H	245	 2% 91% 7% •

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
7	GOL	A	303	-	-	-	X
7	GOL	C	303	-	-	-	X
7	GOL	E	301	-	-	-	X
7	GOL	H	301	-	-	-	X
7	GOL	H	303	-	-	-	X

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 14217 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	263	Total	C	N	O	S	0	2	0
			2158	1384	377	387	10			
1	C	269	Total	C	N	O	S	0	5	0
			2221	1418	388	405	10			

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 Beta-2-microglobulin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	99	Total	C	N	O	S	0	0	0
			817	520	138	157	2			
2	D	98	Total	C	N	O	S	0	0	0
			806	514	134	156	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 M33.64 TCR Alpha Chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	G	198	Total	C	N	O	S	0	2	0
			1525	961	247	307	10			

*Continued on next page...*

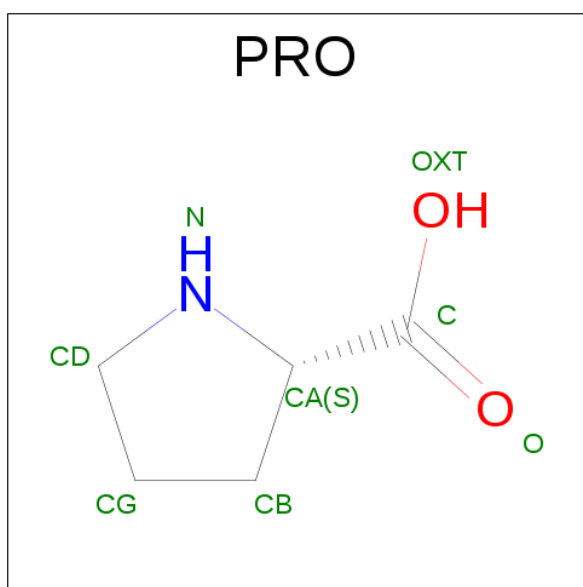
Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	200	Total	C	N	O	S	0	0	0
			1548	976	248	314	10			

- Molecule 4 is a protein called M33.64 TCR Beta Chain.

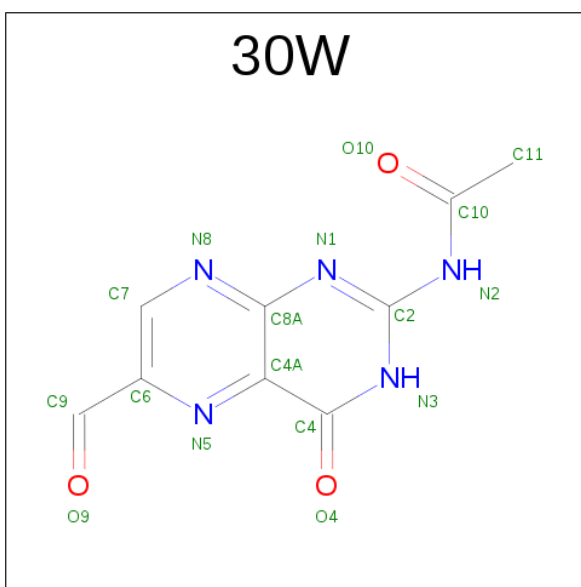
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	H	241	Total	C	N	O	S	0	0	0
			1865	1170	329	358	8			
4	F	243	Total	C	N	O	S	0	0	0
			1880	1179	328	365	8			

- Molecule 5 is PROLINE (three-letter code: PRO) (formula:  $C_5H_9NO_2$ ).



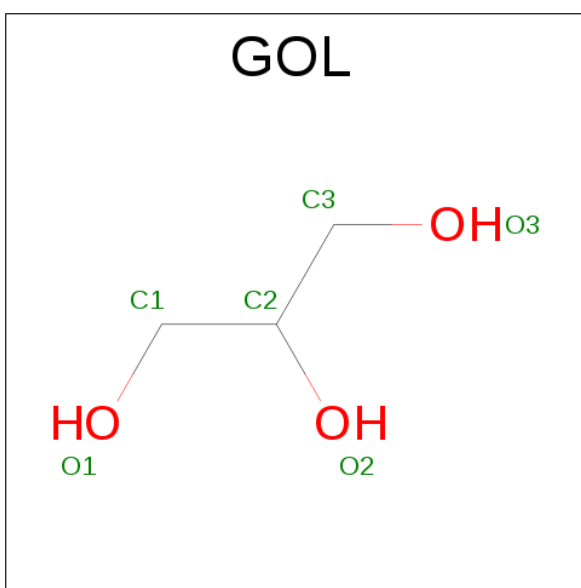
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	N	O	0	0
			7	5	1	1		
5	C	1	Total	C	N	O	0	0
			7	5	1	1		

- Molecule 6 is N-(6-formyl-4-oxo-3,4-dihydropteridin-2-yl)acetamide (three-letter code: 30W) (formula:  $C_9H_7N_5O_3$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	1	Total	C	N	O	0	0
			16	9	5	2		
6	C	1	Total	C	N	O	0	0
			16	9	5	2		

- 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	A	1	Total	C	O	0	0
			6	3	3		

*Continued on next page...*

*Continued from previous page...*

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

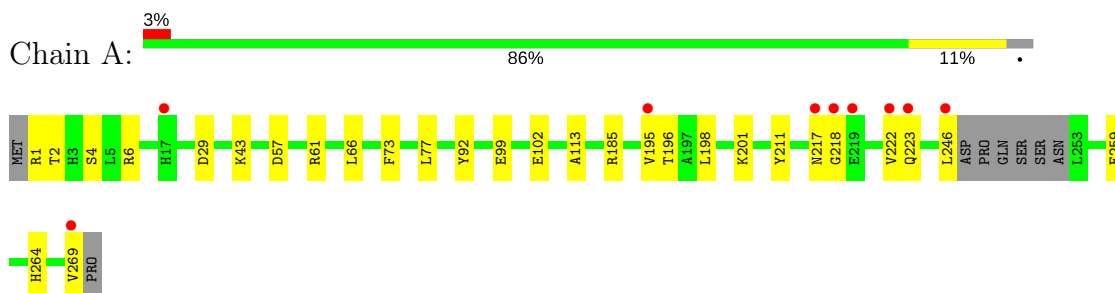
- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	243	Total	O	0	0
			243	243		
8	B	93	Total	O	0	0
			93	93		
8	C	204	Total	O	0	0
			204	204		
8	D	46	Total	O	0	0
			46	46		
8	G	160	Total	O	0	0
			160	160		
8	H	238	Total	O	0	0
			238	238		
8	E	159	Total	O	0	0
			159	159		
8	F	148	Total	O	0	0
			148	148		

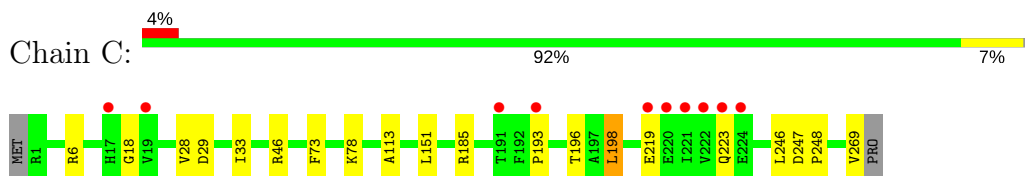
### 3 Residue-property plots

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.

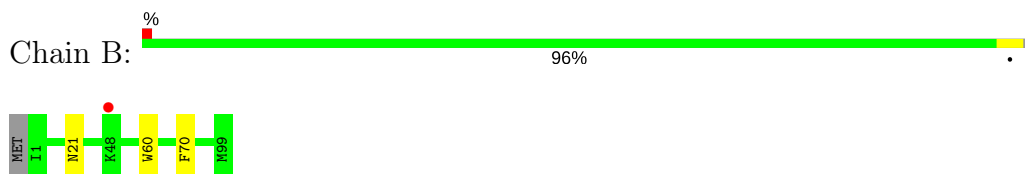
- Molecule 1: Major histocompatibility complex class I-related gene protein



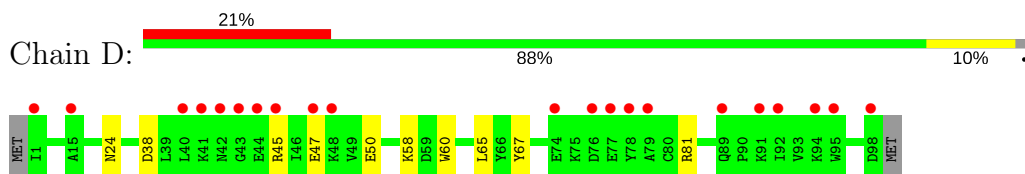
- Molecule 1: Major histocompatibility complex class I-related gene protein



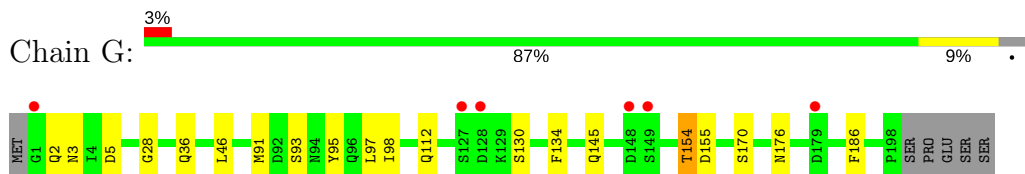
- Molecule 2: Beta-2-microglobulin



- Molecule 2: Beta-2-microglobulin

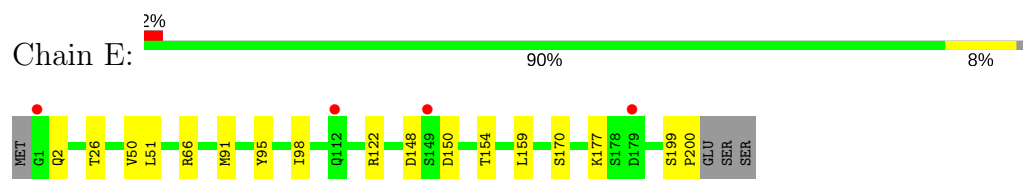


- Molecule 3: M33.64 TCR Alpha Chain

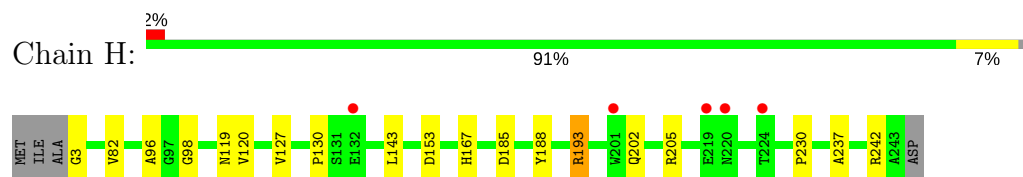




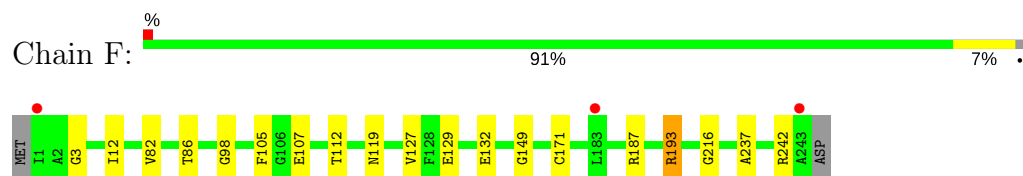
- Molecule 3: M33.64 TCR Alpha Chain



- Molecule 4: M33.64 TCR Beta Chain



- Molecule 4: M33.64 TCR Beta Chain



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	214.30Å 69.66Å 141.19Å 90.00° 103.99° 90.00°	Depositor
Resolution (Å)	51.74 – 2.00 51.74 – 2.00	Depositor EDS
% Data completeness (in resolution range)	97.8 (51.74-2.00) 97.8 (51.74-2.00)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.54 (at 2.00Å)	Xtriage
Refinement program	PHENIX	Depositor
R, $R_{free}$	0.178 , 0.220 0.178 , 0.220	Depositor DCC
$R_{free}$ test set	6634 reflections (4.96%)	DCC
Wilson B-factor (Å <sup>2</sup> )	28.6	Xtriage
Anisotropy	0.495	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 49.0	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	14217	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.31% 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, 3OW

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	0.41	0/2231	0.55	0/3034
1	C	0.34	0/2300	0.51	0/3130
2	B	0.38	0/840	0.52	0/1141
2	D	0.34	0/829	0.51	0/1127
3	E	0.40	0/1583	0.56	0/2149
3	G	0.41	0/1564	0.59	0/2125
4	F	0.36	0/1928	0.55	0/2630
4	H	0.39	0/1913	0.56	0/2608
All	All	0.38	0/13188	0.55	0/17944

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
4	F	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	F	105	PHE	Peptide

## 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	2158	0	2034	23	0
1	C	2221	0	2100	13	0
2	B	817	0	765	2	0
2	D	806	0	752	11	0
3	E	1548	0	1448	17	0
3	G	1525	0	1400	16	0
4	F	1880	0	1778	10	0
4	H	1865	0	1767	15	0
5	A	7	0	7	2	0
5	C	7	0	7	1	0
6	A	16	0	6	0	0
6	C	16	0	6	0	0
7	A	12	0	16	3	0
7	B	12	0	16	0	0
7	C	6	0	8	0	0
7	E	6	0	8	0	0
7	F	6	0	8	2	0
7	H	18	0	24	2	0
8	A	243	0	0	4	0
8	B	93	0	0	0	0
8	C	204	0	0	2	0
8	D	46	0	0	2	0
8	E	159	0	0	0	0
8	F	148	0	0	1	0
8	G	160	0	0	4	0
8	H	238	0	0	3	0
All	All	14217	0	12150	97	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 (97) 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)
1:A:61:ARG:HB2	7:A:303:GOL:H31	1.50	0.91

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:H:98:GLY:H	7:H:301:GOL:H11	1.36	0.91
2:D:38:ASP:HB2	2:D:81:ARG:HB3	1.54	0.88
3:G:3:ASN:ND2	3:G:5:ASP:OD1	2.06	0.87
3:G:170:SER:OG	4:H:193:ARG:NE	2.15	0.80
1:A:2:THR:OG1	8:A:401:HOH:O	2.03	0.75
3:E:170:SER:OG	4:F:193:ARG:NE	2.19	0.75
2:D:38:ASP:HA	2:D:45:ARG:HH12	1.53	0.73
1:A:195:VAL:HG11	1:C:78:LYS:HE2	1.71	0.71
4:F:129:GLU:OE1	4:F:242:ARG:NH1	2.23	0.71
1:C:198:LEU:HD13	1:C:246:LEU:HD21	1.74	0.70
2:D:38:ASP:HA	2:D:45:ARG:NH1	2.07	0.69
3:G:130:SER:O	8:G:301:HOH:O	2.11	0.68
3:G:112:GLN:OE1	8:G:302:HOH:O	2.12	0.68
1:A:211:TYR:OH	8:A:402:HOH:O	2.13	0.66
1:C:6[A]:ARG:NH2	1:C:29:ASP:O	2.25	0.66
3:E:2:GLN:HE21	3:E:98:ILE:CD1	2.10	0.65
2:D:45:ARG:NH1	8:D:101:HOH:O	2.30	0.65
4:H:3:GLY:N	8:H:403:HOH:O	2.30	0.63
3:E:2:GLN:NE2	3:E:98:ILE:HD11	2.12	0.63
3:G:145:GLN:NE2	8:G:306:HOH:O	2.33	0.60
3:E:2:GLN:HG2	3:E:26:THR:HG22	1.83	0.60
3:G:176:ASN:N	3:G:176:ASN:OD1	2.36	0.59
7:F:301:GOL:O2	8:F:401:HOH:O	2.12	0.57
3:E:2:GLN:HE21	3:E:98:ILE:HD13	1.70	0.56
4:H:96:ALA:O	8:H:401:HOH:O	2.18	0.56
3:E:150:ASP:HB2	3:E:177:LYS:HD3	1.86	0.56
1:A:198:LEU:HD23	1:A:246:LEU:HD11	1.87	0.56
3:G:112:GLN:CD	3:G:112:GLN:H	2.09	0.55
2:B:21:ASN:HB3	2:B:70:PHE:CE2	2.43	0.54
3:G:176:ASN:ND2	8:G:307:HOH:O	2.41	0.54
1:A:1:ARG:HD2	8:A:582:HOH:O	2.07	0.54
3:E:2:GLN:NE2	3:E:98:ILE:CD1	2.71	0.53
2:D:58:LYS:HG3	8:D:126:HOH:O	2.09	0.53
2:D:38:ASP:OD1	2:D:45:ARG:NH1	2.42	0.53
4:H:98:GLY:N	7:H:301:GOL:H11	2.15	0.52
3:E:159:LEU:HB3	4:F:171:CYS:HB2	1.90	0.52
4:F:12:ILE:HD12	4:F:216:GLY:HA2	1.90	0.52
3:G:36:GLN:HB2	3:G:46:LEU:HD11	1.92	0.52
1:C:113:ALA:HB2	2:D:60:TRP:CE2	2.45	0.52
1:A:259:GLU:OE1	1:A:264[B]:HIS:NE2	2.44	0.50
3:E:170:SER:O	4:F:193:ARG:NH1	2.38	0.50

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:102:GLU:HG2	8:A:401:HOH:O	2.10	0.49
3:G:28:GLY:HA3	3:G:93[B]:SER:OG	2.12	0.49
1:A:246:LEU:HD12	1:A:246:LEU:H	1.77	0.49
2:D:45:ARG:NH2	2:D:47:GLU:OE2	2.45	0.49
1:A:113:ALA:HB2	2:B:60:TRP:CE2	2.48	0.49
3:G:170:SER:HG	4:H:193:ARG:HE	1.56	0.48
4:H:127:VAL:HG23	4:H:237:ALA:HB3	1.95	0.48
1:A:196:THR:HB	1:A:246:LEU:HD13	1.96	0.48
1:A:61:ARG:CB	7:A:303:GOL:H31	2.35	0.47
1:C:185:ARG:NH1	8:C:408:HOH:O	2.43	0.47
3:E:148:ASP:OD1	3:E:177:LYS:NZ	2.32	0.47
4:H:202:GLN:HA	4:H:242:ARG:O	2.14	0.47
1:A:246:LEU:HD12	1:A:246:LEU:N	2.30	0.47
3:E:91:MET:HG2	3:E:95:TYR:HA	1.96	0.47
4:H:130:PRO:HD3	4:H:143:LEU:HG	1.96	0.47
4:H:119:ASN:ND2	4:H:185:ASP:HB2	2.30	0.47
1:C:269:VAL:C	5:C:301:PRO:HA	2.35	0.46
3:E:2:GLN:CG	3:E:26:THR:HA	2.44	0.46
1:C:219:GLU:OE1	1:C:219:GLU:N	2.49	0.46
1:C:196:THR:HG22	1:C:246:LEU:HD23	1.97	0.46
4:H:205:ARG:HA	4:H:205:ARG:NE	2.31	0.46
3:G:154:THR:HG22	3:G:155:ASP:O	2.16	0.45
1:A:269:VAL:HG12	5:A:301:PRO:C	2.36	0.45
1:A:57:ASP:OD2	7:A:303:GOL:H32	2.17	0.45
1:A:77:LEU:HD13	1:A:92:TYR:HB2	1.98	0.45
1:A:269:VAL:HG12	5:A:301:PRO:O	2.16	0.44
1:C:193:PRO:HD3	8:C:462:HOH:O	2.17	0.44
4:F:127:VAL:HG23	4:F:237:ALA:HB3	2.00	0.44
1:A:43:LYS:HD2	1:A:66:LEU:HD12	2.00	0.44
2:D:50:GLU:HB3	2:D:67:TYR:CZ	2.52	0.44
1:C:46:ARG:HD3	1:C:46:ARG:HA	1.85	0.43
1:C:151:LEU:HD22	3:E:51:LEU:HD12	2.00	0.43
4:F:98:GLY:H	7:F:301:GOL:H31	1.83	0.43
4:F:86:THR:HG23	4:F:112:THR:HA	2.00	0.43
4:H:153:ASP:HB3	4:H:188:TYR:CE2	2.53	0.42
2:D:24:ASN:HB3	2:D:65:LEU:HD11	2.01	0.42
4:F:149:GLY:O	4:F:187:ARG:HD2	2.18	0.42
1:A:185:ARG:NH2	1:A:201:LYS:HD3	2.34	0.42
1:A:222:VAL:HA	1:A:223:GLN:HA	1.75	0.42
1:C:247:ASP:HA	1:C:248:PRO:HD3	1.91	0.42
3:G:2:GLN:NE2	3:G:98:ILE:HD11	2.35	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:217:ASN:HA	1:A:218:GLY:O	2.20	0.42
4:H:120:VAL:HG12	4:H:230:PRO:HB2	2.02	0.42
4:H:167:HIS:HD2	8:H:613:HOH:O	2.02	0.42
2:D:45:ARG:NE	2:D:47:GLU:OE2	2.54	0.41
3:E:199:SER:HA	3:E:200:PRO:HD3	1.90	0.41
1:C:28:VAL:HG23	1:C:33:ILE:HD13	2.02	0.41
3:G:91:MET:HG2	3:G:95:TYR:HA	2.02	0.41
3:E:2:GLN:HG2	3:E:26:THR:HA	2.01	0.41
3:E:50:VAL:O	3:E:66:ARG:HD3	2.21	0.41
3:E:122:ARG:NH2	4:F:132:GLU:HG2	2.36	0.41
1:A:4:SER:HB3	1:A:99:GLU:HG2	2.03	0.41
1:A:6[A]:ARG:NH2	1:A:29:ASP:O	2.45	0.40
3:G:170:SER:O	4:H:193:ARG:NH1	2.52	0.40
3:G:134:PHE:HB2	3:G:186:PHE:CZ	2.55	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	261/271 (96%)	252 (97%)	9 (3%)	0	100	100
1	C	272/271 (100%)	263 (97%)	7 (3%)	2 (1%)	25	18
2	B	97/100 (97%)	97 (100%)	0	0	100	100
2	D	96/100 (96%)	96 (100%)	0	0	100	100
3	E	198/204 (97%)	195 (98%)	3 (2%)	0	100	100
3	G	198/204 (97%)	192 (97%)	6 (3%)	0	100	100
4	F	241/245 (98%)	233 (97%)	7 (3%)	1 (0%)	38	33
4	H	239/245 (98%)	233 (98%)	6 (2%)	0	100	100
All	All	1602/1640 (98%)	1561 (97%)	38 (2%)	3 (0%)	51	48

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	F	3	GLY
1	C	223	GLN
1	C	18	GLY

### 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	225/241 (93%)	224 (100%)	1 (0%)	93	95
1	C	235/241 (98%)	233 (99%)	2 (1%)	82	87
2	B	91/95 (96%)	91 (100%)	0	100	100
2	D	90/95 (95%)	90 (100%)	0	100	100
3	E	172/181 (95%)	171 (99%)	1 (1%)	89	92
3	G	164/181 (91%)	162 (99%)	2 (1%)	75	80
4	F	200/205 (98%)	196 (98%)	4 (2%)	60	64
4	H	198/205 (97%)	196 (99%)	2 (1%)	80	84
All	All	1375/1444 (95%)	1363 (99%)	12 (1%)	82	87

All (12) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	73	PHE
1	C	73	PHE
1	C	198	LEU
3	G	97	LEU
3	G	154	THR
4	H	82	VAL
4	H	193	ARG
3	E	154	THR
4	F	82	VAL
4	F	107	GLU
4	F	119	ASN

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
4	F	193	ARG

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

Mol	Chain	Res	Type
3	E	2	GLN

### 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 ⓘ

14 ligands are modelled in this entry.

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	PRO	A	301	-	6,7,8	0.65	0	7,8,10	1.38	1 (14%)
6	30W	A	302	1	16,17,18	1.36	2 (12%)	20,24,25	2.19	4 (20%)
7	GOL	A	303	-	5,5,5	0.29	0	5,5,5	0.38	0
7	GOL	A	304	-	5,5,5	0.26	0	5,5,5	0.29	0
7	GOL	B	101	-	5,5,5	0.37	0	5,5,5	0.75	0
7	GOL	B	102	-	5,5,5	0.30	0	5,5,5	0.34	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	PRO	C	301	-	6,7,8	1.04	1 (16%)	7,8,10	1.67	1 (14%)
6	30W	C	302	1	16,17,18	1.33	2 (12%)	20,24,25	2.22	4 (20%)
7	GOL	C	303	-	5,5,5	0.33	0	5,5,5	0.66	0
7	GOL	E	301	-	5,5,5	0.37	0	5,5,5	0.19	0
7	GOL	F	301	-	5,5,5	0.43	0	5,5,5	0.48	0
7	GOL	H	301	-	5,5,5	0.42	0	5,5,5	1.07	0
7	GOL	H	302	-	5,5,5	0.30	0	5,5,5	0.48	0
7	GOL	H	303	-	5,5,5	0.29	0	5,5,5	0.52	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	PRO	A	301	-	-	0/0/9/11	0/1/1/1
6	30W	A	302	1	-	0/4/4/6	0/2/2/2
7	GOL	A	303	-	-	0/4/4/4	0/0/0/0
7	GOL	A	304	-	-	0/4/4/4	0/0/0/0
7	GOL	B	101	-	-	0/4/4/4	0/0/0/0
7	GOL	B	102	-	-	0/4/4/4	0/0/0/0
5	PRO	C	301	-	-	0/0/9/11	0/1/1/1
6	30W	C	302	1	-	0/4/4/6	0/2/2/2
7	GOL	C	303	-	-	0/4/4/4	0/0/0/0
7	GOL	E	301	-	-	0/4/4/4	0/0/0/0
7	GOL	F	301	-	-	0/4/4/4	0/0/0/0
7	GOL	H	301	-	-	0/4/4/4	0/0/0/0
7	GOL	H	302	-	-	0/4/4/4	0/0/0/0
7	GOL	H	303	-	-	0/4/4/4	0/0/0/0

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	C	301	PRO	CA-C	2.15	1.53	1.50
6	C	302	30W	C4-N3	2.83	1.38	1.33
6	A	302	30W	C4-C4A	2.99	1.47	1.41
6	C	302	30W	C4-C4A	3.20	1.47	1.41
6	A	302	30W	C4-N3	3.25	1.38	1.33

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	C	302	30W	C4A-C4-N3	-7.08	113.41	123.48
6	A	302	30W	C4A-C4-N3	-6.88	113.69	123.48
5	C	301	PRO	O-C-CA	-3.83	116.22	125.15
5	A	301	PRO	CB-CA-C	-2.95	108.64	112.70
6	C	302	30W	C2-N1-C8A	2.01	117.41	115.11
6	A	302	30W	C2-N1-C8A	2.38	117.83	115.11
6	C	302	30W	C4-C4A-N5	2.46	121.38	118.68
6	A	302	30W	C4-C4A-N5	2.87	121.82	118.68
6	A	302	30W	C4-N3-C2	4.80	123.78	115.18
6	C	302	30W	C4-N3-C2	5.18	124.45	115.18

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	301	PRO	2	0
7	A	303	GOL	3	0
5	C	301	PRO	1	0
7	F	301	GOL	2	0
7	H	301	GOL	2	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	263/271 (97%)	-0.17	9 (3%) 46 46	18, 29, 62, 110	0
1	C	269/271 (99%)	-0.23	10 (3%) 42 43	25, 35, 71, 121	0
2	B	99/100 (99%)	-0.11	1 (1%) 82 82	23, 38, 71, 78	0
2	D	98/100 (98%)	0.79	21 (21%) 1 1	30, 58, 92, 117	0
3	E	200/204 (98%)	-0.31	4 (2%) 65 65	24, 35, 64, 87	0
3	G	198/204 (97%)	0.16	6 (3%) 51 51	17, 35, 84, 115	0
4	F	243/245 (99%)	-0.16	3 (1%) 79 78	24, 40, 63, 100	0
4	H	241/245 (98%)	-0.26	5 (2%) 64 63	19, 30, 62, 96	0
All	All	1611/1640 (98%)	-0.11	59 (3%) 42 43	17, 35, 74, 121	0

All (59) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	222	VAL	7.9
1	A	17	HIS	6.0
3	E	1	GLY	6.0
3	G	149	SER	5.1
2	D	98	ASP	5.1
4	H	220	ASN	4.4
3	G	127	SER	4.2
4	F	183	LEU	4.2
2	D	92	ILE	3.8
4	F	1	ILE	3.7
2	D	48	LYS	3.5
2	D	89	GLN	3.3
2	D	79	ALA	3.3
2	D	78	TYR	3.2
1	A	219	GLU	3.2
2	D	45	ARG	3.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
4	F	243	ALA	3.1
1	C	219	GLU	3.0
1	A	217	ASN	3.0
2	D	74	GLU	3.0
1	C	193	PRO	2.9
1	A	218	GLY	2.9
1	A	246	LEU	2.9
2	D	44	GLU	2.9
4	H	132	GLU	2.7
4	H	219	GLU	2.7
3	G	1	GLY	2.7
2	B	48	LYS	2.7
1	C	220	GLU	2.7
3	G	128	ASP	2.7
2	D	15	ALA	2.6
2	D	77	GLU	2.5
3	E	179	ASP	2.5
1	C	223	GLN	2.5
1	A	222	VAL	2.5
2	D	47	GLU	2.4
1	C	19	VAL	2.4
1	A	195	VAL	2.4
3	E	149	SER	2.4
1	C	191	THR	2.3
3	G	179	ASP	2.3
2	D	42	ASN	2.3
2	D	1	ILE	2.3
2	D	43	GLY	2.3
2	D	41	LYS	2.2
1	C	224	GLU	2.2
4	H	224	THR	2.2
2	D	76	ASP	2.2
1	A	223	GLN	2.2
3	E	112	GLN	2.1
1	C	221	ILE	2.1
1	C	17	HIS	2.1
4	H	201	TRP	2.1
2	D	91	LYS	2.1
2	D	40	LEU	2.1
1	A	269	VAL	2.1
3	G	148	ASP	2.0
2	D	95	TRP	2.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	D	94	LYS	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	H	303	6/6	0.60	0.25	10.11	58,62,63,63	0
7	GOL	E	301	6/6	0.75	0.28	9.64	57,62,64,67	0
7	GOL	A	303	6/6	0.78	0.26	6.88	56,57,59,63	0
7	GOL	H	301	6/6	0.96	0.16	2.37	26,33,35,38	0
7	GOL	C	303	6/6	0.84	0.19	2.26	43,46,47,47	0
7	GOL	F	301	6/6	0.95	0.12	1.91	28,33,34,41	0
7	GOL	B	102	6/6	0.93	0.11	1.62	31,35,39,41	0
7	GOL	B	101	6/6	0.92	0.13	1.40	33,34,40,42	0
7	GOL	A	304	6/6	0.94	0.14	1.01	47,48,50,52	0
7	GOL	H	302	6/6	0.74	0.14	0.89	51,54,58,61	0
6	30W	A	302	16/17	0.98	0.11	-0.46	20,23,28,28	0
6	30W	C	302	16/17	0.96	0.08	-0.78	22,27,30,31	0
5	PRO	C	301	7/8	0.84	0.33	-	77,95,97,104	0
5	PRO	A	301	7/8	0.69	0.19	-	62,65,71,74	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.