



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 2, 2017 – 01:12 PM EST

PDB ID : 5U6Q  
Title : Structure of human MR1-3-F-SA in complex with human MAIT A-F7 TCR  
Authors : Keller, A.N.; Rossjohn, J.  
Deposited on : 2016-12-08  
Resolution : 1.90 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](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:

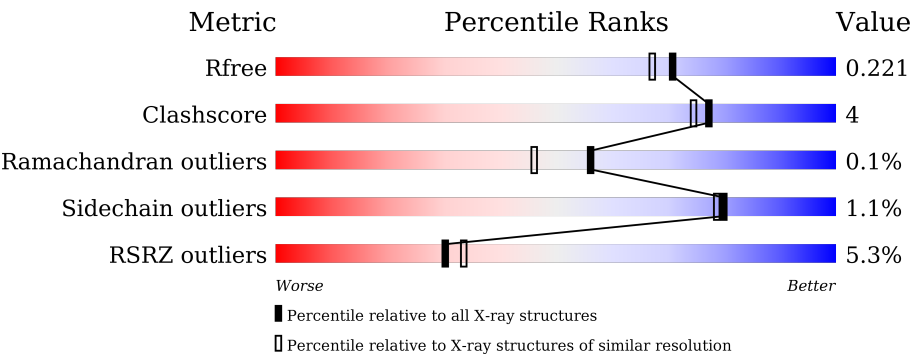
MolProbity : 4.02b-467  
Mogul : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20028442  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal 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 1.90 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.



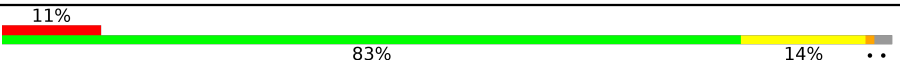
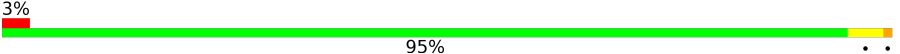
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	91344	4755 (1.90-1.90)
Clashscore	102246	5398 (1.90-1.90)
Ramachandran outliers	100387	5338 (1.90-1.90)
Sidechain outliers	100360	5339 (1.90-1.90)
RSRZ outliers	91569	4766 (1.90-1.90)

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></div><div><div></div><div></div><div></div><div></div></div><div>88%8%</div></div>
1	C	271	<div><div>6%</div><div></div><div></div><div></div><div></div></div> <div>89%7%</div>
2	B	203	<div><div>14%</div><div></div><div></div><div></div><div></div></div> <div>84%12%</div>
2	D	203	<div><div></div><div></div><div></div><div></div><div></div></div> <div>89%9%</div>
3	E	245	<div><div></div><div></div><div></div><div></div><div></div></div> <div>91%8%</div>
3	G	245	<div><div>7%</div><div></div><div></div><div></div><div></div></div> <div>88%9%</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	C	302	-	-	-	X
7	ACT	A	302	-	-	-	X
7	ACT	C	303	-	-	-	X

## 2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 14317 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	C	261	Total	C	N	O	S	0	1	0
			2157	1377	372	397	11			
1	A	263	Total	C	N	O	S	0	2	0
			2184	1401	373	399	11			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	0	MET	-	initiating methionine	UNP Q95460
C	261	SER	CYS	conflict	UNP Q95460
A	0	MET	-	initiating methionine	UNP Q95460
A	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	D	200	Total	C	N	O	S	0	1	0
			1565	986	251	319	9			
2	B	195	Total	C	N	O	S	0	0	0
			1503	954	239	301	9			

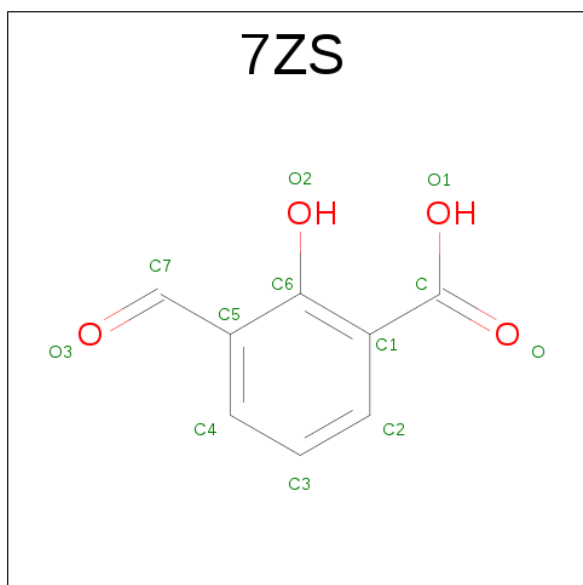
- Molecule 3 is a protein called MAIT T-cell receptor beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	244	Total	C	N	O	S	0	2	0
			1943	1219	339	376	9			
3	G	239	Total	C	N	O	S	0	0	0
			1879	1183	323	364	9			

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

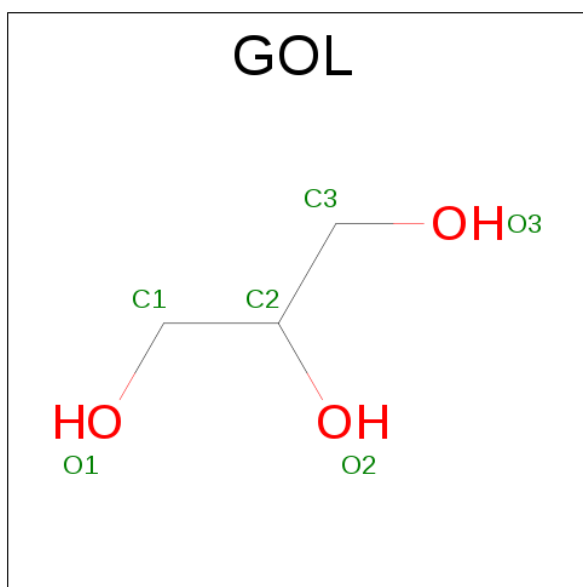
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	F	97	Total	C	N	O	S	0	0	0
			800	511	135	152	2			
4	H	99	Total	C	N	O	S	0	0	0
			820	522	138	157	3			

- Molecule 5 is 3-methanoyl-2-oxidanyl-benzoic acid (three-letter code: 7ZS) (formula:  $C_8H_6O_4$ ).



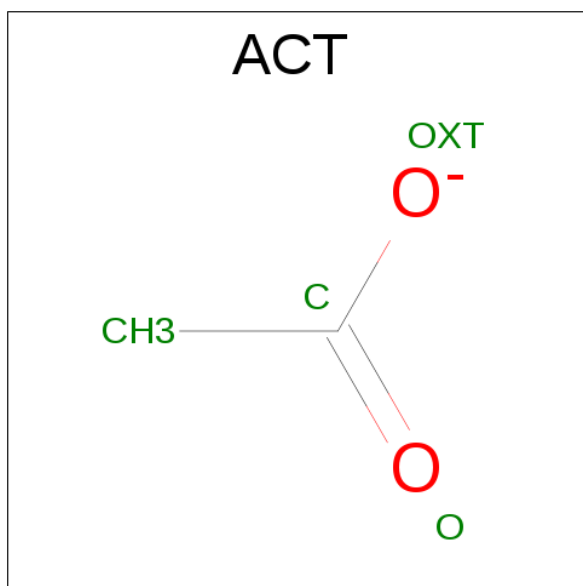
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	C	1	Total	C	O	0	0
			11	8	3		
5	A	1	Total	C	O	0	0
			11	8	3		

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



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	C	1	Total	C	O	0	0
			6	3	3		

- Molecule 7 is ACETATE ION (three-letter code: ACT) (formula:  $C_2H_3O_2$ ).



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

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

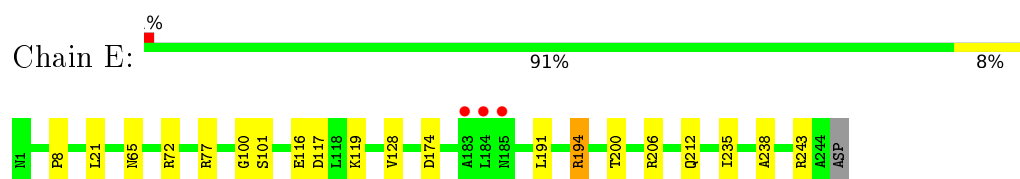
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	G	1	Total Na 1 1	0	0

- Molecule 9 is water.

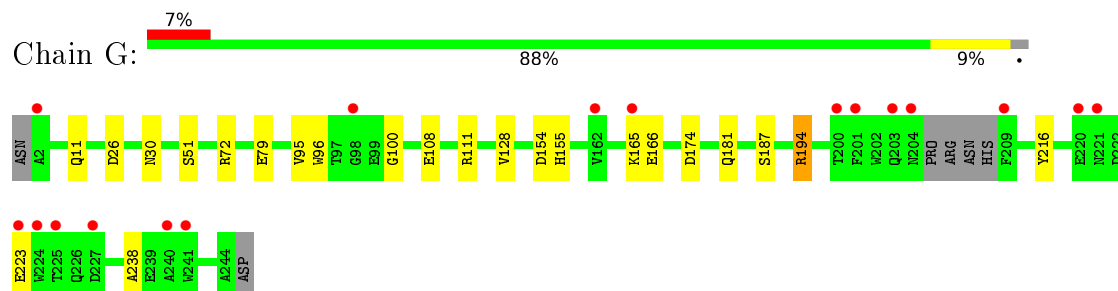
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	C	253	Total O 253 253	0	0
9	D	227	Total O 227 227	0	0
9	E	291	Total O 291 291	0	0
9	F	78	Total O 78 78	0	0
9	A	229	Total O 229 229	0	0
9	B	106	Total O 106 106	0	0
9	G	139	Total O 139 139	0	0
9	H	106	Total O 106 106	0	0



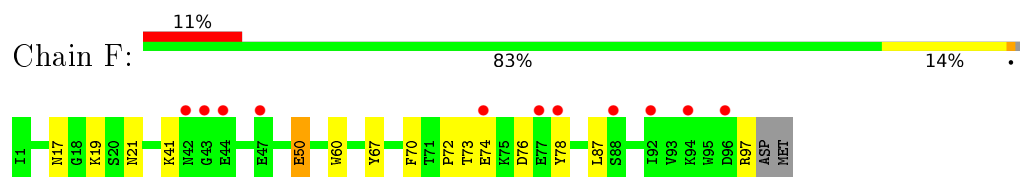




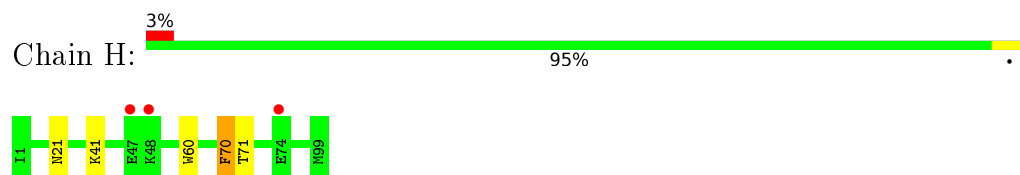
- Molecule 3: MAIT T-cell receptor beta chain



- 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, $\alpha$ , $\beta$ , $\gamma$	217.67Å 70.51Å 143.37Å 90.00° 104.77° 90.00°	Depositor
Resolution (Å)	44.39 – 1.90 44.39 – 1.90	Depositor EDS
% Data completeness (in resolution range)	97.4 (44.39-1.90) 97.4 (44.39-1.90)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.57 (at 1.89Å)	Xtriage
Refinement program	PHENIX	Depositor
R, $R_{free}$	0.182 , 0.222 0.180 , 0.221	Depositor DCC
$R_{free}$ test set	8116 reflections (5.03%)	DCC
Wilson B-factor (Å <sup>2</sup> )	23.4	Xtriage
Anisotropy	0.657	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 46.0	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	14317	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.66% 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, ACT, 7ZS, 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	0.42	0/2250	0.57	1/3059 (0.0%)
1	C	0.44	0/2223	0.56	0/3023
2	B	0.41	0/1537	0.56	0/2088
2	D	0.46	0/1600	0.60	0/2170
3	E	0.49	0/1994	0.59	0/2712
3	G	0.38	0/1927	0.55	0/2620
4	F	0.34	0/823	0.52	0/1118
4	H	0.38	0/843	0.55	0/1144
All	All	0.43	0/13197	0.57	1/17934 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	66	LEU	CA-CB-CG	-7.51	98.03	115.30

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	2184	0	2062	16	0
1	C	2157	0	2026	11	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	1503	0	1399	15	0
2	D	1565	0	1474	17	0
3	E	1943	0	1836	18	0
3	G	1879	0	1773	19	0
4	F	800	0	753	9	0
4	H	820	0	772	4	0
5	A	11	0	0	0	0
5	C	11	0	0	0	0
6	C	6	0	8	2	0
7	A	4	0	3	0	0
7	C	4	0	3	0	0
8	G	1	0	0	0	0
9	A	229	0	0	3	0
9	B	106	0	0	1	0
9	C	253	0	0	1	0
9	D	227	0	0	2	0
9	E	291	0	0	5	0
9	F	78	0	0	1	0
9	G	139	0	0	2	0
9	H	106	0	0	2	0
All	All	14317	0	12109	93	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.

The worst 5 of 93 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:D:150:ASP:HB2	2:D:177:LYS:HD2	1.63	0.78
2:D:170:SER:OG	3:E:194:ARG:NE	2.17	0.77
3:G:111:ARG:NH1	3:G:154:ASP:O	2.17	0.77
3:E:65[A]:ASN:OD1	3:E:77:ARG:HB3	1.88	0.74
2:B:112:GLN:NE2	9:B:301:HOH:O	2.20	0.73

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	259/271 (96%)	256 (99%)	3 (1%)	0	100	100
1	C	258/271 (95%)	253 (98%)	4 (2%)	1 (0%)	39	27
2	B	191/203 (94%)	183 (96%)	7 (4%)	1 (0%)	34	21
2	D	199/203 (98%)	196 (98%)	3 (2%)	0	100	100
3	E	244/245 (100%)	239 (98%)	5 (2%)	0	100	100
3	G	235/245 (96%)	227 (97%)	8 (3%)	0	100	100
4	F	95/99 (96%)	94 (99%)	1 (1%)	0	100	100
4	H	97/99 (98%)	96 (99%)	1 (1%)	0	100	100
All	All	1578/1636 (96%)	1544 (98%)	32 (2%)	2 (0%)	56	46

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	112	GLN
1	C	222	VAL

### 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	232/241 (96%)	229 (99%)	3 (1%)	76	73
1	C	230/241 (95%)	227 (99%)	3 (1%)	76	73
2	B	164/180 (91%)	162 (99%)	2 (1%)	78	76

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	D	176/180 (98%)	174 (99%)	2 (1%)	80	79
3	E	211/211 (100%)	210 (100%)	1 (0%)	92	92
3	G	203/211 (96%)	201 (99%)	2 (1%)	82	81
4	F	89/94 (95%)	88 (99%)	1 (1%)	80	79
4	H	92/94 (98%)	91 (99%)	1 (1%)	80	79
All	All	1397/1452 (96%)	1382 (99%)	15 (1%)	80	79

5 of 15 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	F	50	GLU
1	A	66	LEU
3	G	79	GLU
3	E	194	ARG
2	B	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	96	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 ⓘ

Of 6 ligands modelled in this entry, 1 is monoatomic - leaving 5 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	7ZS	A	301	1	8,11,12	0.89	1 (12%)	11,15,16	0.88	0
7	ACT	A	302	-	0,3,3	0.00	-	0,3,3	0.00	-
5	7ZS	C	301	1	8,11,12	0.81	0	11,15,16	1.06	1 (9%)
6	GOL	C	302	-	5,5,5	0.43	0	5,5,5	0.85	0
7	ACT	C	303	-	0,3,3	0.00	-	0,3,3	0.00	-

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	7ZS	A	301	1	-	0/0/4/6	0/1/1/1
7	ACT	A	302	-	-	0/0/0/0	0/0/0/0
5	7ZS	C	301	1	-	0/0/4/6	0/1/1/1
6	GOL	C	302	-	-	0/4/4/4	0/0/0/0
7	ACT	C	303	-	-	0/0/0/0	0/0/0/0

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	301	7ZS	C7-C5	-2.32	1.46	1.51

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	C	301	7ZS	C1-C6-C5	-2.35	119.74	120.95

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	C	302	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%)	-0.19	3 (1%) 82 84	19, 27, 52, 71	1 (0%)
1	C	261/271 (96%)	-0.05	16 (6%) 25 27	17, 28, 62, 92	1 (0%)
2	B	195/203 (96%)	0.58	29 (14%) 3 3	22, 42, 95, 106	0
2	D	200/203 (98%)	-0.29	2 (1%) 84 86	16, 25, 54, 74	0
3	E	244/245 (99%)	-0.38	3 (1%) 81 83	17, 26, 47, 73	1 (0%)
3	G	239/245 (97%)	0.13	17 (7%) 19 21	23, 39, 88, 109	0
4	F	97/99 (97%)	0.45	11 (11%) 7 7	22, 47, 75, 94	0
4	H	99/99 (100%)	-0.11	3 (3%) 54 57	19, 33, 60, 77	0
All	All	1598/1636 (97%)	-0.02	84 (5%) 30 33	16, 30, 73, 109	3 (0%)

The worst 5 of 84 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	222	VAL	6.2
2	B	199	SER	5.9
2	B	112	GLN	5.8
2	B	178	SER	5.5
2	B	200	PRO	5.5

### 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. 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	A	302	4/4	0.87	0.22	3.48	39,42,42,44	0
7	ACT	C	303	4/4	0.93	0.20	3.34	35,38,39,43	0
6	GOL	C	302	6/6	0.83	0.17	2.34	40,45,47,51	0
5	7ZS	A	301	11/12	0.96	0.15	1.23	24,28,32,33	0
8	NA	G	301	1/1	0.91	0.16	1.15	30,30,30,30	0
5	7ZS	C	301	11/12	0.94	0.11	0.20	19,22,24,26	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.