W O R L D W I D E
wwPDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 9, 2023 – 02:09 AM EDT

PDB ID	:	6W9V
Title	:	Structure of human MAIT A-F7 TCR in complex with patient MR1-R9H without ligand
Authors	:	Awad, W.; Rossjohn, J.
Deposited on	:	2020-03-23
Resolution	:	1.95 Å(reported)

The following versions of software and data (see references ⓘ) were used in the production of this report:

|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.95 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.
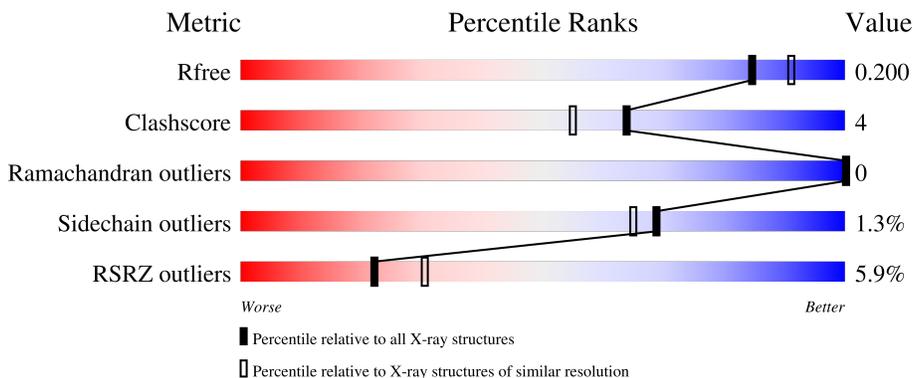
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	2580 (1.96-1.96)
Clashscore	141614	2705 (1.96-1.96)
Ramachandran outliers	138981	2678 (1.96-1.96)
Sidechain outliers	138945	2678 (1.96-1.96)
RSRZ outliers	127900	2539 (1.96-1.96)

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 of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=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	3% / 87% / 11% / •
1	C	271	3% / 87% / 10% / ••
2	B	100	21% / 91% / 6% / ••
2	F	100	95% / 5%
3	D	204	17% / 84% / 8% / 8%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	G	204	88% 10% •
4	E	246	9% 89% 9% •
4	H	246	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
6	GOL	C	301	-	-	X	-

2 Entry composition ⓘ

There are 7 unique types of molecules in this entry. The entry contains 14539 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Major histocompatibility complex class I-related gene protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	265	Total	C	N	O	S	0	20	0
			2276	1467	390	406	13			
1	C	267	Total	C	N	O	S	0	16	0
			2268	1466	385	405	12			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	expression tag	UNP Q95460
A	9	HIS	ARG	engineered mutation	UNP Q95460
A	261	SER	CYS	conflict	UNP Q95460
C	0	MET	-	expression tag	UNP Q95460
C	9	HIS	ARG	engineered mutation	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	98	Total	C	N	O	S	0	2	0
			783	504	131	145	3			
2	F	100	Total	C	N	O	S	0	3	0
			822	529	139	150	4			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	0	MET	-	expression tag	UNP P61769
F	0	MET	-	expression tag	UNP P61769

- Molecule 3 is a protein called TCR-alpha chain.

wwPDB X-ray Structure Validation Summary Report 6W9V

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	188	Total 1460	C 944	N 227	O 280	S 9	0	9	0
3	G	201	Total 1644	C 1048	N 254	O 329	S 13	0	20	0

- Molecule 4 is a protein called TCR-beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	241	Total 1901	C 1205	N 323	O 361	S 12	0	9	0
4	H	244	Total 1967	C 1249	N 334	O 371	S 13	0	16	0

- Molecule 5 is ACETATE ION (three-letter code: ACT) (formula: $C_2H_3O_2$).
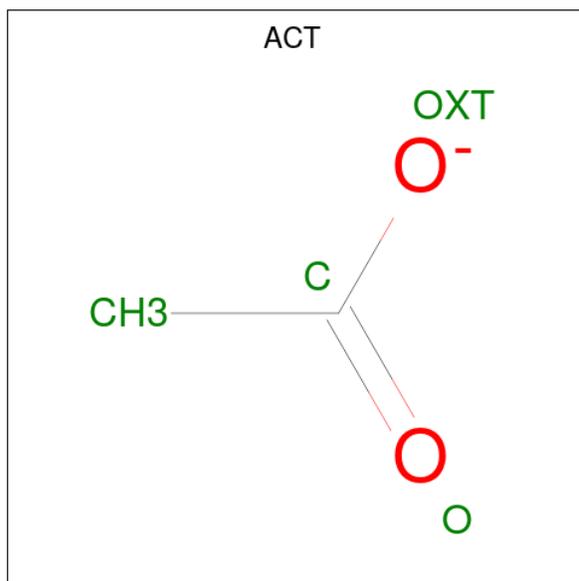
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total 4	C 2	O 2	0	0
5	H	1	Total 4	C 2	O 2	0	0

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

GOL

C3

OH O3

C1 C2

HO

O1

OH

O2

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

- Molecule 7 is water.

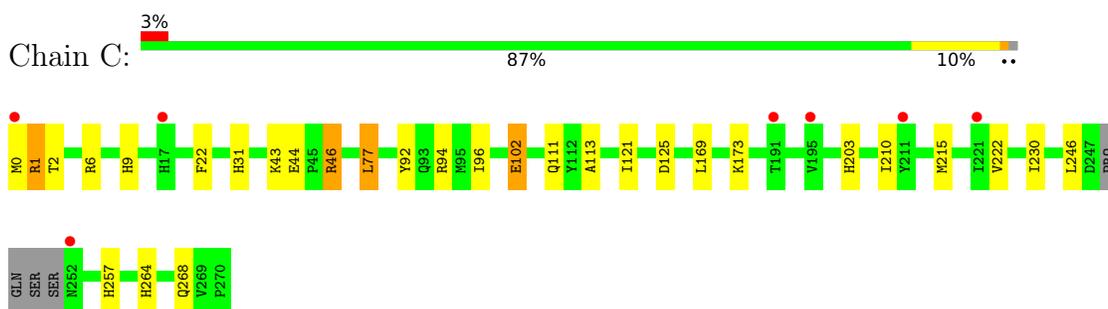| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	247	Total 247	O 247	0	0
7	B	67	Total 67	O 67	0	0
7	C	229	Total 229	O 229	0	0
7	D	121	Total 121	O 121	0	0
7	E	138	Total 138	O 138	0	0
7	F	105	Total 105	O 105	0	0
7	G	202	Total 202	O 202	0	0
7	H	277	Total 277	O 277	0	0

WORLDWIDE
PDB
PROTEIN DATA BANK

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.
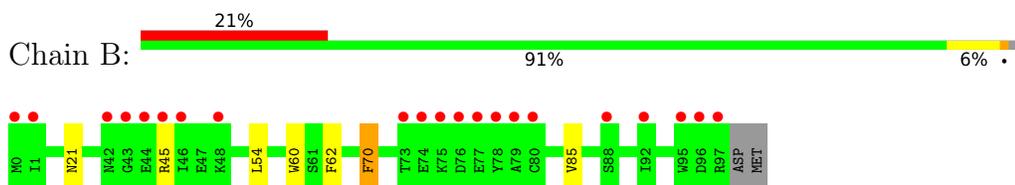
• Molecule 1: Major histocompatibility complex class I-related gene protein
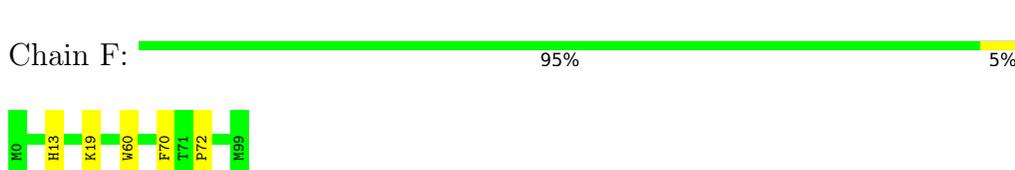
Chain A:

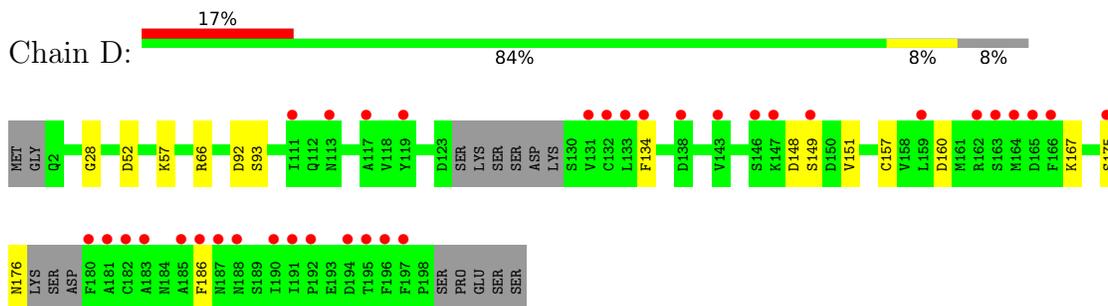• Molecule 1: Major histocompatibility complex class I-related gene protein

Chain C:

• Molecule 2: Beta-2-microglobulin

Chain B:

• Molecule 2: Beta-2-microglobulin

Chain F:

• Molecule 3: TCR-alpha chain

Chain D:

17%

84%   8%   8%

MET GLY Q2 G28 D52 K57 R66 D92 S93 I111 Q112 N113 A117 V118 Y119 D123 SER LYS SER SER ASP LYS S130 V131 C132 L133 F134 D138 V143 S146 K147 D148 S149 D150 V151 C157 V158 L159 D160 M161 R162 S163 M164 D165 F166 K167 S175

N176 LYS SER ASP F180 A181 C182 A183 N184 A185 F186 N187 N188 S189 I190 I191 P192 E193 D194 T195 F196 F197 P198 SER PRO GLU SER SER

• Molecule 3: TCR-alpha chain

Chain G:

88%   10%

MET GLY Q2 E9 S27 G28 S61 Y70 S71 Y72 L73 L74 S93 Q112 K125 C132 D138 Q145 S146 C157 V158 L159 M164 K167 C182 A183 N184 A185 F186 S202 SER

• Molecule 4: TCR-beta chain

Chain E:

9%

89%   9%

MET ASN ALA G3 L13 K14 M19 D26 R72 R77 L78 E79 W96 T97 G98 R111 V128 P131 S132 I136 S137 K141 H155 V162 N163 V167 R194 V197 S198 A199 T200 F201 W202 Q203 N204 P205 R206 N207 H208 F209 Y216 S219

E220 N221 D222 E223 W224 T225 Q226 D227 A238 W241 G242 R243 ALA ASP

• Molecule 4: TCR-beta chain

Chain H:

91%   7%

MET N1 V12 L13 M19 R72 N120 E125 V128 E157 C172 L184 N185 D186 R194 T200 R206 R210 C211 Q212 A238 A244 ASP

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	216.99Å　　70.41Å　　142.90Å 90.00°　　104.15°　　90.00°	Depositor
Resolution (Å)	46.19　–　1.95 47.91　–　1.95	Depositor EDS
% Data completeness (in resolution range)	99.6 (46.19-1.95) 99.6 (47.91-1.95)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.80 (at 1.95Å)	Xtriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, R_{free}	0.160　,　0.200 0.162　,　0.200	Depositor DCC
R_{free} test set	7601 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	35.9	Xtriage
Anisotropy	0.325	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 47.8	EDS
L-test for twinning [2]	$< |L| > = 0.49, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	14539	wwPDB-VP
Average B, all atoms (Å^2)	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.69% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ 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: ACT, GOL

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	0.65	0/2400	0.72	3/3261 (0.1%)
1	C	0.60	0/2383	0.73	3/3239 (0.1%)
2	B	0.50	0/812	0.61	0/1107
2	F	0.60	0/854	0.69	0/1161
3	D	0.59	0/1520	0.73	2/2068 (0.1%)
3	G	0.70	0/1743	0.78	1/2365 (0.0%)
4	E	0.54	0/1976	0.66	1/2692 (0.0%)
4	H	0.67	0/2067	0.72	1/2814 (0.0%)
All	All	0.62	0/13755	0.71	11/18707 (0.1%)

There are no bond length outliers.

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	46[A]	ARG	NE-CZ-NH1	-7.00	116.80	120.30
1	A	46[B]	ARG	NE-CZ-NH1	-7.00	116.80	120.30
1	C	1	ARG	NE-CZ-NH1	6.93	123.77	120.30
3	G	157	CYS	CA-CB-SG	6.70	126.06	114.00
4	H	210	ARG	NE-CZ-NH1	-6.16	117.22	120.30

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.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2276	0	2182	22	0
1	C	2268	0	2177	28	0
2	B	783	0	716	5	0
2	F	822	0	782	3	0
3	D	1460	0	1347	8	0
3	G	1644	0	1594	15	0
4	E	1901	0	1778	14	0
4	H	1967	0	1881	17	0
5	A	4	0	3	0	0
5	H	4	0	3	0	0
6	A	12	0	16	0	0
6	C	6	0	8	5	0
6	F	6	0	8	0	0
7	A	247	0	0	10	2
7	B	67	0	0	2	0
7	C	229	0	0	14	0
7	D	121	0	0	2	2
7	E	138	0	0	4	0
7	F	105	0	0	1	0
7	G	202	0	0	7	0
7	H	277	0	0	8	0
All	All	14539	0	12495	107	2

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 107 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:C:44:GLU:OE1	7:C:401:HOH:O	1.76	1.00
4:E:111:ARG:NH2	7:E:301:HOH:O	1.99	0.95
1:C:94:ARG:HH12	6:C:301:GOL:H12	1.35	0.92
1:A:92:TYR:OH	7:A:401:HOH:O	1.88	0.90
1:C:173:LYS:NZ	7:C:402:HOH:O	2.00	0.89

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
7:A:632:HOH:O	7:D:396:HOH:O[2_658]	2.03	0.17
7:A:493:HOH:O	7:D:321:HOH:O[2_658]	2.17	0.03

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	281/271 (104%)	279 (99%)	2 (1%)	0	100	100
1	C	279/271 (103%)	274 (98%)	5 (2%)	0	100	100
2	B	98/100 (98%)	98 (100%)	0	0	100	100
2	F	101/100 (101%)	101 (100%)	0	0	100	100
3	D	191/204 (94%)	186 (97%)	5 (3%)	0	100	100
3	G	220/204 (108%)	216 (98%)	4 (2%)	0	100	100
4	E	248/246 (101%)	245 (99%)	3 (1%)	0	100	100
4	H	258/246 (105%)	255 (99%)	3 (1%)	0	100	100
All	All	1676/1642 (102%)	1654 (99%)	22 (1%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	245/241 (102%)	242 (99%)	3 (1%)	71	68
1	C	242/241 (100%)	238 (98%)	4 (2%)	60	55
2	B	82/95 (86%)	81 (99%)	1 (1%)	71	68
2	F	90/95 (95%)	89 (99%)	1 (1%)	73	71
3	D	150/181 (83%)	149 (99%)	1 (1%)	84	82
3	G	194/181 (107%)	188 (97%)	6 (3%)	40	28

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
4	E	202/212 (95%)	198 (98%)	4 (2%)	55	48
4	H	214/212 (101%)	208 (97%)	6 (3%)	43	33
All	All	1419/1458 (97%)	1393 (98%)	26 (2%)	69	53

5 of 26 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	G	70[A]	TYR
3	G	146[B]	SER
4	H	206[A]	ARG
3	G	146[A]	SER
3	G	146[C]	SER

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

Mol	Chain	Res	Type
4	E	28	ASN
3	G	188	ASN
4	H	120	ASN

5.3.3   RNA ⓘ

There are no RNA molecules in this entry.

5.4   Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5   Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6   Ligand geometry ⓘ

6 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	ACT	A	301	-	3,3,3	1.51	1 (33%)	3,3,3	1.45	0
6	GOL	A	303	-	5,5,5	0.93	0	5,5,5	1.38	1 (20%)
6	GOL	F	101	-	5,5,5	1.23	0	5,5,5	0.99	0
5	ACT	H	1601	-	3,3,3	1.60	1 (33%)	3,3,3	1.32	0
6	GOL	C	301	-	5,5,5	0.87	0	5,5,5	0.95	0
6	GOL	A	302	-	5,5,5	1.29	1 (20%)	5,5,5	0.93	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
6	GOL	F	101	-	-	0/4/4/4	-
6	GOL	C	301	-	-	2/4/4/4	-
6	GOL	A	303	-	-	2/4/4/4	-
6	GOL	A	302	-	-	0/4/4/4	-

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	H	1601	ACT	CH3-C	2.40	1.59	1.49
6	A	302	GOL	C3-C2	2.34	1.61	1.51
5	A	301	ACT	CH3-C	2.19	1.58	1.49

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	A	303	GOL	C3-C2-C1	-2.45	102.18	111.70

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	A	303	GOL	C1-C2-C3-O3
6	C	301	GOL	O1-C1-C2-C3
6	A	303	GOL	O2-C2-C3-O3
6	C	301	GOL	O1-C1-C2-O2

There are no ring outliers.

1 monomer is involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	C	301	GOL	5	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, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å²)	Q<0.9
1	A	265/271 (97%)	-0.10	8 (3%)	50	59	27, 39, 72, 89	11 (4%)
1	C	267/271 (98%)	-0.03	7 (2%)	56	65	27, 37, 62, 83	15 (5%)
2	B	98/100 (98%)	0.82	21 (21%)	0	1	33, 60, 83, 93	5 (5%)
2	F	100/100 (100%)	-0.19	0	100	100	29, 41, 58, 65	5 (5%)
3	D	188/204 (92%)	0.73	35 (18%)	1	1	29, 49, 89, 114	7 (3%)
3	G	201/204 (98%)	-0.25	0	100	100	26, 36, 56, 71	15 (7%)
4	E	241/246 (97%)	0.17	22 (9%)	9	15	31, 50, 91, 107	10 (4%)
4	H	244/246 (99%)	-0.07	1 (0%)	92	95	26, 36, 54, 78	15 (6%)
All	All	1604/1642 (97%)	0.09	94 (5%)	22	30	26, 41, 80, 114	83 (5%)

The worst 5 of 94 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	D	180	PHE	6.0
4	E	225	THR	5.8
3	D	162	ARG	5.6
3	D	197	PHE	5.1
3	D	163	SER	4.9

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
5	ACT	A	301	4/4	0.82	0.19	72,75,75,75	0
6	GOL	A	302	6/6	0.87	0.19	42,45,45,46	6
6	GOL	C	301	6/6	0.90	0.13	58,60,61,62	0
5	ACT	H	1601	4/4	0.94	0.10	52,57,57,60	0
6	GOL	A	303	6/6	0.96	0.11	35,38,43,44	0
6	GOL	F	101	6/6	0.97	0.08	34,35,37,39	0

6.5 Other polymers ⓘ

There are no such residues in this entry.