wwPDB X-ray Structure Validation Summary Report (i)

Jun 17, 2024 – 12:24 PM EDT

PDB ID	:	3UTS
Title	:	1E6-A*0201-ALWGPDPAAA Complex, Monoclinic
Authors	:	Rizkallah, P.J.; Cole, D.K.; Sewell, A.K.; Bulek, A.M.
Deposited on	:	2011-11-26
Resolution	:	2.71 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

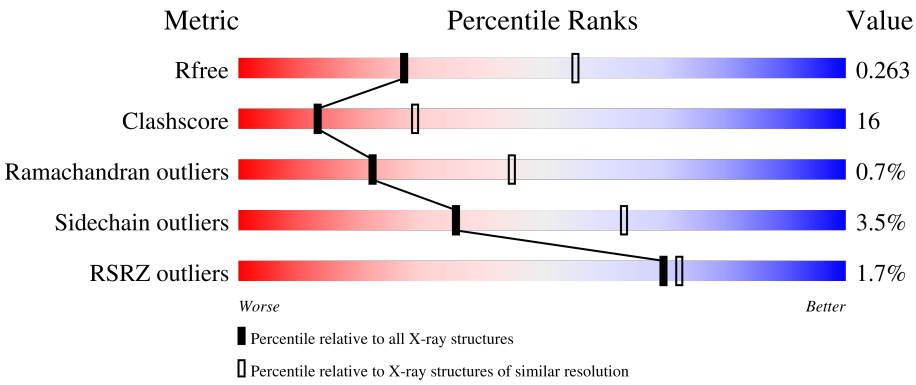
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.71 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	3359 (2.74-2.70)
Clashscore	141614	3686 (2.74-2.70)
Ramachandran outliers	138981	3622 (2.74-2.70)
Sidechain outliers	138945	3623 (2.74-2.70)
RSRZ outliers	127900	3276 (2.74-2.70)

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	276	74% 25% •
1	F	276	2% 76% 22% •
2	B	100	4% 84% 16%
2	G	100	4% 78% 21% •
3	C	10	80% 20%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	H	10	10% 90% 10%
4	D	201	% 69% 27% •
4	I	201	69% 28% •
5	E	246	2% 69% 29% •
5	J	246	2% 76% 21% •

2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 13552 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 HLA class I histocompatibility antigen, A-2 alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	276	2253	1408	410	426	9	0	0	0
1	F	276	2264	1414	414	427	9	0	1	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	100	837	533	141	159	4	0	0	0
2	G	100	837	533	141	159	4	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	0	MET	-	INITIATING METHIONINE	UNP P61769
G	0	MET	-	INITIATING METHIONINE	UNP P61769

- Molecule 3 is a protein called Insulin.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	C	10	69	45	11	13	0	0	0
3	H	10	69	45	11	13	0	0	0

- Molecule 4 is a protein called 1E6 TCR Alpha Chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	D	200	1579	989	260	320	10	0	0	0

Continued on next page...

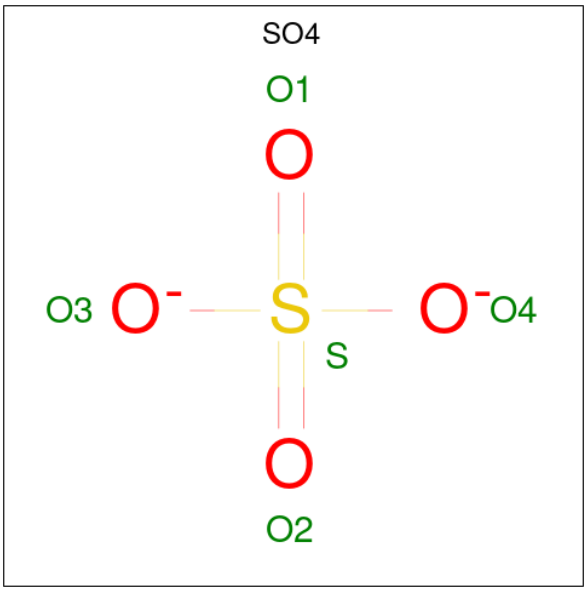
Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	I	201	1586	994	261	321	10	0	0	0

- Molecule 5 is a protein called 1E6 TCR Beta Chain.

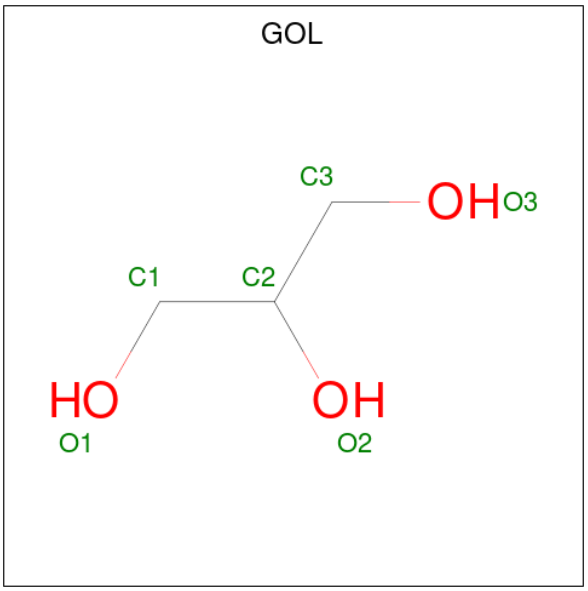
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
5	E	246	1974	1249	341	374	10	0	0	0
5	J	246	1973	1249	341	373	10	0	0	0

- Molecule 6 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	S		
6	A	1	5	4	1	0	0
6	D	1	5	4	1	0	0
6	F	1	5	4	1	0	0

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



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

- Molecule 8 is water.

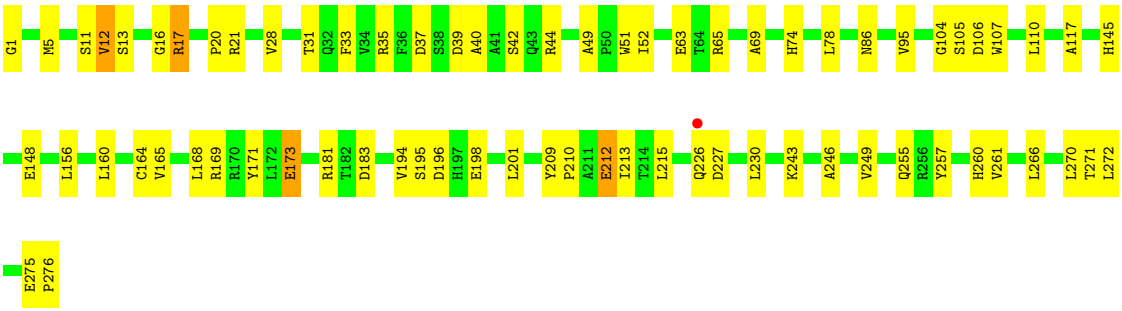
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
8	A	18	18	18	0	0
			Total	O		
8	B	10	10	10	0	0
			Total	O		
8	D	7	7	7	0	0
			Total	O		
8	E	14	14	14	0	0
			Total	O		
8	F	15	15	15	0	0
			Total	O		
8	G	5	5	5	0	0
			Total	O		
8	H	3	3	3	0	0
			Total	O		
8	I	4	4	4	0	0
			Total	O		
8	J	14	14	14	0	0

3 Residue-property plots (i)

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.

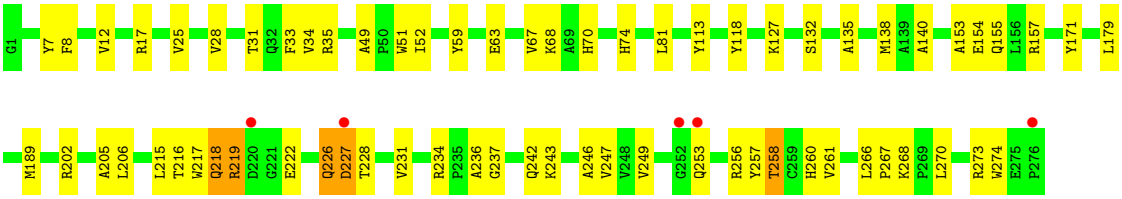
• Molecule 1: HLA class I histocompatibility antigen, A-2 alpha chain

Chain A:



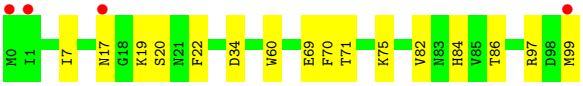
• Molecule 1: HLA class I histocompatibility antigen, A-2 alpha chain

Chain F:



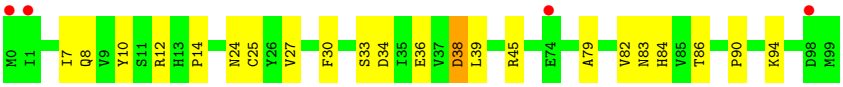
• Molecule 2: Beta-2-microglobulin

Chain B:



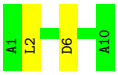
• Molecule 2: Beta-2-microglobulin

Chain G:



- Molecule 3: Insulin

Chain C:

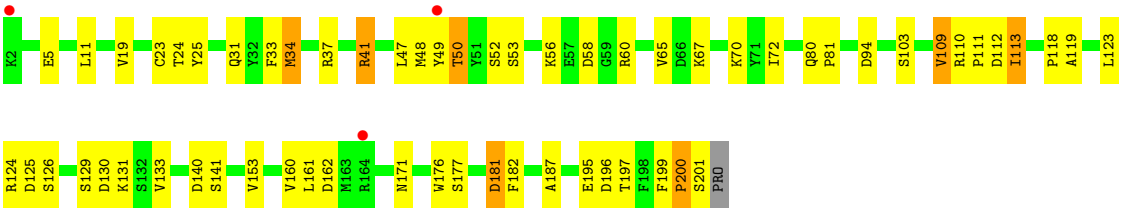


- Molecule 3: Insulin

Chain H:

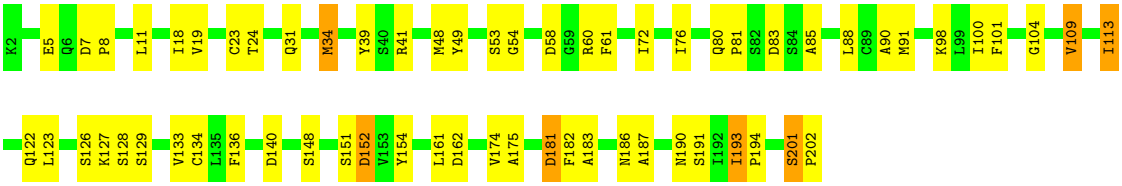
- Molecule 4: 1E6 TCR Alpha Chain

Chain D:



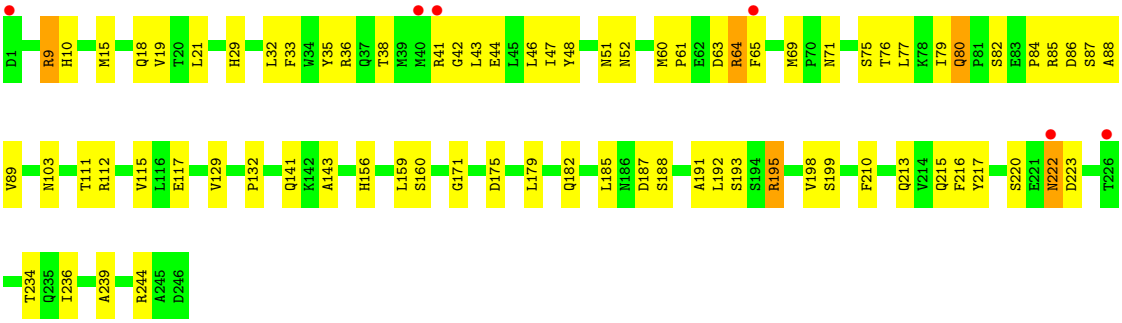
- Molecule 4: 1E6 TCR Alpha Chain

Chain I:



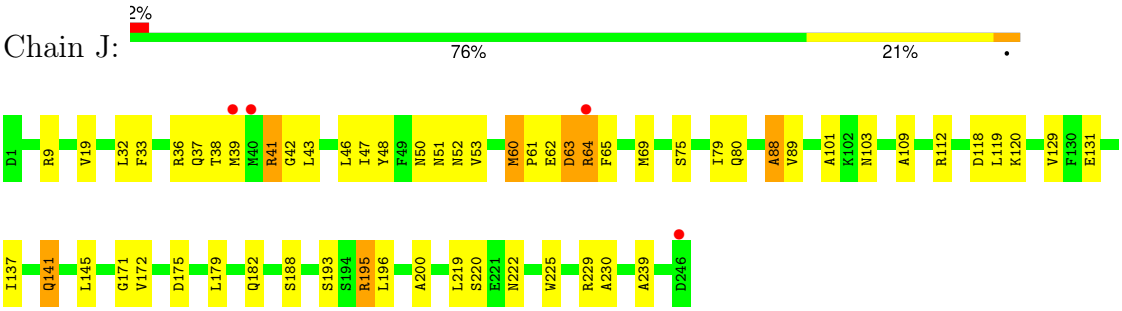
- Molecule 5: 1E6 TCR Beta Chain

Chain E:



- Molecule 5: 1E6 TCR Beta Chain

Chain J:


2%
76%
21%
D1 R9 V19 L32 F33 R36 Q37 T38 M39 M40 R41 G42 L43 L46 I47 Y48 F49 N50 N51 N52 V53 M60 P61 E62 D63 R64 F65 M69 S75 I79 Q80 A88 V89 A101 K102 N103 A109 R112 D118 L119 K120 V129 F130 E131
I137 Q141 L145 G171 V172 D175 L179 Q182 S188 S193 S194 R195 L196 A200 L219 S220 E221 N222 W225 R229 A230 A239 D246

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	93.68Å 84.59Å 126.28Å 90.00° 90.03° 90.00°	Depositor
Resolution (Å)	63.14 – 2.71 63.14 – 2.71	Depositor EDS
% Data completeness (in resolution range)	100.0 (63.14-2.71) 99.2 (63.14-2.71)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.09	Depositor
$< I/\sigma(I) >$ [1]	2.25 (at 2.73Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.201 , 0.269 0.201 , 0.263	Depositor DCC
R_{free} test set	2715 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å^2)	55.5	Xtriage
Anisotropy	0.549	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 18.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.43$, $< L^2 > = 0.26$	Xtriage
Estimated twinning fraction	0.165 for h,-k,-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	13552	wwPDB-VP
Average B, all atoms (Å^2)	60.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.11% 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: SO4, 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	0/2319	0.59	0/3149
1	F	0.59	0/2330	0.57	0/3163
2	B	0.69	0/860	0.59	0/1162
2	G	0.64	0/860	0.55	0/1162
3	C	0.73	0/72	0.63	0/99
3	H	0.70	0/72	0.58	0/99
4	D	0.72	1/1615 (0.1%)	0.65	0/2185
4	I	0.73	1/1623 (0.1%)	0.66	0/2197
5	E	0.64	0/2029	0.60	0/2759
5	J	0.73	0/2028	0.67	1/2759 (0.0%)
All	All	0.67	2/13808 (0.0%)	0.61	1/18734 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	D	23	CYS	CB-SG	-8.22	1.68	1.82
4	I	134	CYS	CB-SG	5.55	1.91	1.82

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	J	63	ASP	N-CA-C	-6.38	93.78	111.00

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	2253	0	2103	64	0
1	F	2264	0	2115	80	0
2	B	837	0	803	12	0
2	G	837	0	803	17	0
3	C	69	0	64	2	0
3	H	69	0	64	3	0
4	D	1579	0	1494	68	0
4	I	1586	0	1501	78	0
5	E	1974	0	1887	72	0
5	J	1973	0	1887	64	0
6	A	5	0	0	1	0
6	D	5	0	0	0	0
6	F	5	0	0	1	0
7	J	6	0	8	0	0
8	A	18	0	0	3	0
8	B	10	0	0	0	0
8	D	7	0	0	2	0
8	E	14	0	0	2	0
8	F	15	0	0	0	0
8	G	5	0	0	0	0
8	H	3	0	0	0	0
8	I	4	0	0	0	0
8	J	14	0	0	0	0
All	All	13552	0	12729	429	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 16.

The worst 5 of 429 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:F:266:LEU:HD22	1:F:270:LEU:CD1	1.62	1.28
1:F:266:LEU:CD2	1:F:270:LEU:HD12	1.66	1.24
1:F:266:LEU:HD23	1:F:270:LEU:HD12	1.31	1.11
4:D:33:PHE:CD1	4:D:72:ILE:HD11	1.90	1.06

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:F:266:LEU:CD2	1:F:270:LEU:CD1	2.27	1.05

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	274/276 (99%)	260 (95%)	14 (5%)	0	100	100
1	F	275/276 (100%)	255 (93%)	18 (6%)	2 (1%)	22	45
2	B	98/100 (98%)	94 (96%)	4 (4%)	0	100	100
2	G	98/100 (98%)	91 (93%)	5 (5%)	2 (2%)	7	17
3	C	8/10 (80%)	7 (88%)	1 (12%)	0	100	100
3	H	8/10 (80%)	7 (88%)	1 (12%)	0	100	100
4	D	198/201 (98%)	185 (93%)	11 (6%)	2 (1%)	15	35
4	I	199/201 (99%)	187 (94%)	10 (5%)	2 (1%)	15	35
5	E	244/246 (99%)	228 (93%)	15 (6%)	1 (0%)	34	58
5	J	244/246 (99%)	236 (97%)	6 (2%)	2 (1%)	19	41
All	All	1646/1666 (99%)	1550 (94%)	85 (5%)	11 (1%)	22	45

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	D	130	ASP
1	F	227	ASP
5	J	64	ARG
5	E	85	ARG
4	I	151	SER

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	232/232 (100%)	225 (97%)	7 (3%)	41	69
1	F	233/232 (100%)	226 (97%)	7 (3%)	41	69
2	B	95/95 (100%)	92 (97%)	3 (3%)	39	67
2	G	95/95 (100%)	94 (99%)	1 (1%)	73	89
3	C	5/5 (100%)	5 (100%)	0	100	100
3	H	5/5 (100%)	5 (100%)	0	100	100
4	D	181/182 (100%)	173 (96%)	8 (4%)	28	54
4	I	182/182 (100%)	174 (96%)	8 (4%)	28	54
5	E	216/216 (100%)	206 (95%)	10 (5%)	27	52
5	J	216/216 (100%)	209 (97%)	7 (3%)	39	67
All	All	1460/1460 (100%)	1409 (96%)	51 (4%)	36	63

5 of 51 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	12	VAL
2	G	38	ASP
5	J	179	LEU
1	F	35	ARG
1	F	219	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 44 such sidechains are listed below:

Mol	Chain	Res	Type
2	G	13	HIS
4	I	189	ASN
2	G	24	ASN
4	I	144	ASN
5	J	29	HIS

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

4 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	$\#\|Z\| > 2$
6	SO4	F	277	-	4,4,4	0.28	0	6,6,6	0.05	0
7	GOL	J	247	-	5,5,5	0.71	0	5,5,5	0.36	0
6	SO4	D	203	-	4,4,4	0.30	0	6,6,6	0.11	0
6	SO4	A	277	-	4,4,4	0.28	0	6,6,6	0.13	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
7	GOL	J	247	-	-	4/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	J	247	GOL	O1-C1-C2-C3
7	J	247	GOL	C1-C2-C3-O3
7	J	247	GOL	O1-C1-C2-O2
7	J	247	GOL	O2-C2-C3-O3

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	F	277	SO4	1	0
6	A	277	SO4	1	0

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

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(Å^2)	Q<0.9
1	A	276/276 (100%)	-0.17	1 (0%)	92	93	25, 56, 98, 118	0
1	F	276/276 (100%)	0.02	5 (1%)	68	70	28, 69, 141, 173	0
2	B	100/100 (100%)	-0.05	4 (4%)	38	37	30, 55, 98, 119	0
2	G	100/100 (100%)	0.03	4 (4%)	38	37	35, 62, 99, 128	0
3	C	10/10 (100%)	-0.15	0	100	100	31, 32, 38, 40	0
3	H	10/10 (100%)	0.17	1 (10%)	7	5	30, 35, 44, 47	0
4	D	200/201 (99%)	-0.06	3 (1%)	73	76	28, 53, 97, 131	0
4	I	201/201 (100%)	-0.08	0	100	100	26, 49, 88, 111	0
5	E	246/246 (100%)	-0.05	6 (2%)	59	60	26, 55, 103, 139	0
5	J	246/246 (100%)	-0.20	4 (1%)	72	74	25, 45, 89, 108	0
All	All	1665/1666 (99%)	-0.08	28 (1%)	70	72	25, 55, 106, 173	0

The worst 5 of 28 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	G	0	MET	7.8
2	G	1	ILE	6.9
5	E	40	MET	5.1
1	F	220	ASP	4.6
5	E	222	ASN	4.4

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 monosaccharides 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. 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
7	GOL	J	247	6/6	0.57	0.31	42,69,71,74	0
6	SO4	F	277	5/5	0.83	0.39	100,100,100,100	0
6	SO4	A	277	5/5	0.84	0.41	99,100,100,100	0
6	SO4	D	203	5/5	0.85	0.38	100,100,100,100	0

6.5 Other polymers (i)

There are no such residues in this entry.