Full wwPDB X-ray Structure Validation Report (i)

Dec 19, 2023 – 03:37 AM EST

PDB ID	:	2P5W
Title	:	Crystal structures of high affinity human T-cell receptors bound to pMHC reveal native diagonal binding geometry
Authors	:	Sami, M.; Rizkallah, P.J.; Dunn, S.; Li, Y.; Moysey, R.; Vuidepot, A.; Baston, E.; Todorov, P.; Molloy, P.; Gao, F.; Boulter, J.M.; Jakobsen, B.K.
Deposited on	:	2007-03-16
Resolution	:	2.20 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*
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:

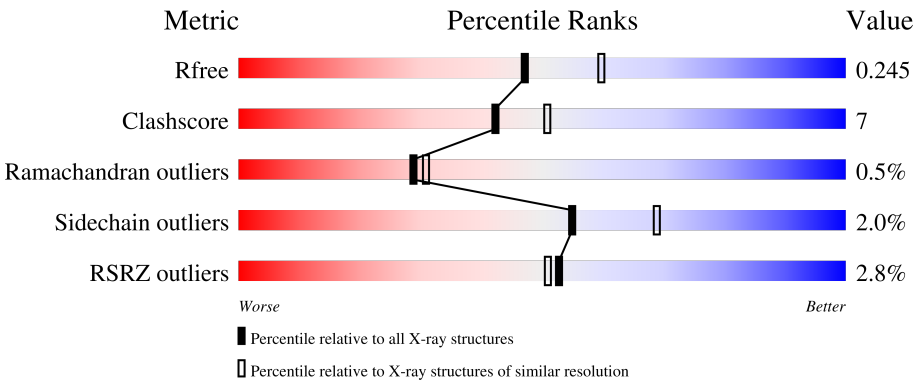
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.20 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	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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	3%, 85%, 13%, •
2	B	100	%, 90%, 8%, ••
3	C	9	67%, 33%
4	D	192	7%, 84%, 14%, •
5	E	243	85%, 13%, •

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	SO4	B	806	-	-	-	X

2 Entry composition (i)

There are 11 unique types of molecules in this entry. The entry contains 7417 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	2267	1417	413	428	9	0	2	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	845	536	144	160	5	0	1	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	0	MET	-	insertion	UNP P61769
B	91	CYS	LYS	conflict	UNP P61769

- Molecule 3 is a protein called Cancer/testis antigen 1B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	9	75	49	11	13	2	0	0	0

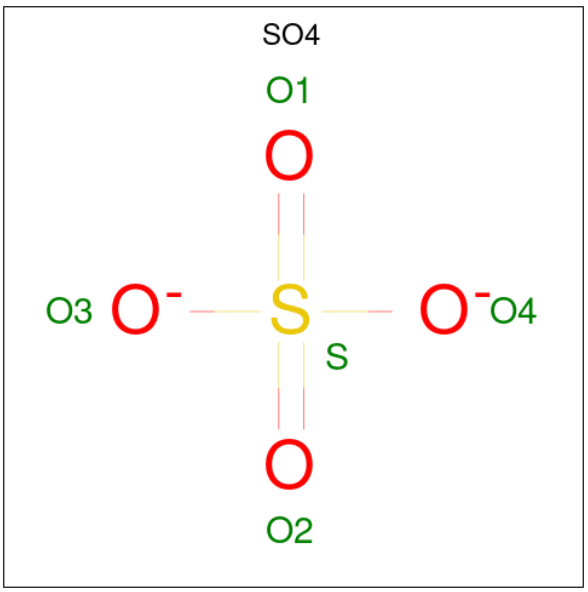
- Molecule 4 is a protein called T-CELL RECEPTOR, ALPHA CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	D	192	1497	934	252	302	9	0	3	0

- Molecule 5 is a protein called Hypothetical protein.

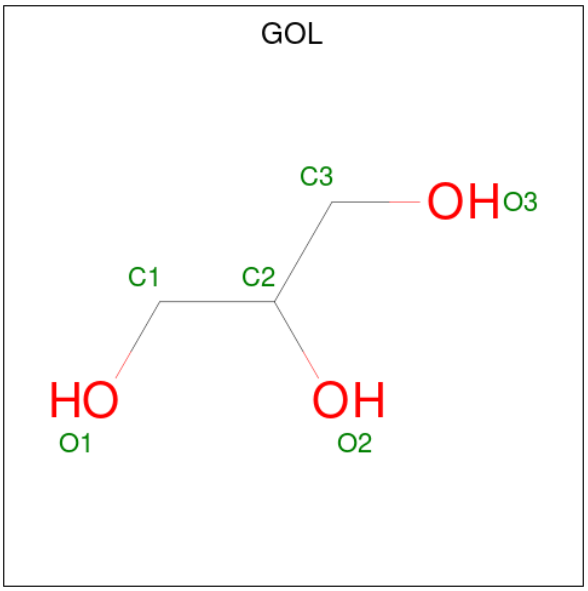
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
5	E	242	1945	1222	334	377	12	0	4	0

- 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
			Total	O	S			
6	A	1	5	4	1		0	0
			Total	O	S			
6	A	1	5	4	1		0	0
			Total	O	S			
6	B	1	5	4	1		0	0
			Total	O	S			
6	D	1	5	4	1		0	0
			Total	O	S			
6	E	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	A	1	6	3	3	0	0
			Total	C	O		
7	A	1	6	3	3	0	0
			Total	C	O		
7	A	1	6	3	3	0	0
			Total	C	O		
7	A	1	6	3	3	0	0
			Total	C	O		
7	B	1	6	3	3	0	0
			Total	C	O		
7	B	1	6	3	3	0	0
			Total	C	O		
7	D	1	6	3	3	0	0
			Total	C	O		
7	D	1	6	3	3	0	0
			Total	C	O		
7	D	1	6	3	3	0	0
			Total	C	O		
7	D	1	6	3	3	0	0
			Total	C	O		
7	E	1	6	3	3	0	0

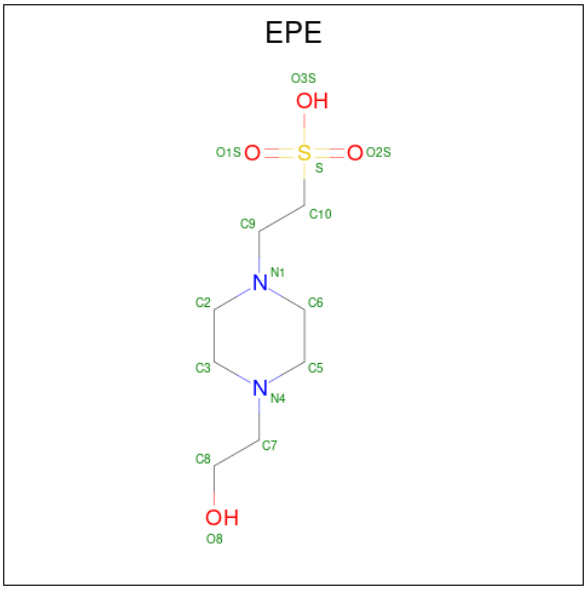
- Molecule 8 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Ca		
8	B	1	1	1	0	0

- Molecule 9 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Mg		
9	B	1	1	1	0	0

- Molecule 10 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID (three-letter code: EPE) (formula: $C_8H_{18}N_2O_4S$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
10	E	1	15	8	2	4	1	0	0

- Molecule 11 is water.

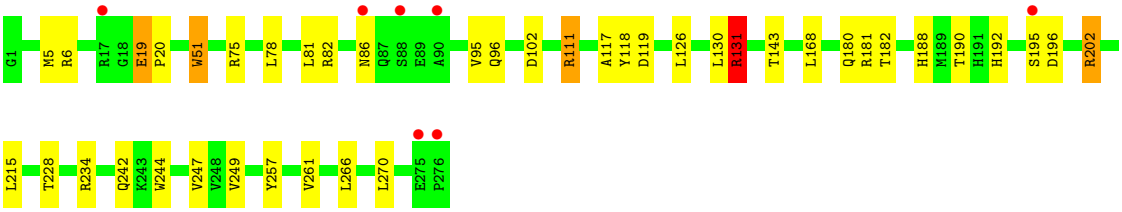
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
11	A	224	224	224	0	0
			Total	O		
11	B	104	104	104	0	0
			Total	O		
11	C	10	10	10	0	0
			Total	O		
11	D	115	115	115	0	0
			Total	O		
11	E	222	222	222	0	0

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.

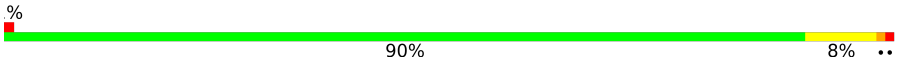
- Molecule 1: HLA class I histocompatibility antigen, A-2 alpha chain

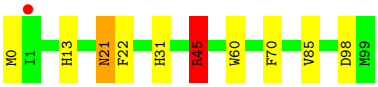
Chain A:



- Molecule 2: Beta-2-microglobulin

Chain B:



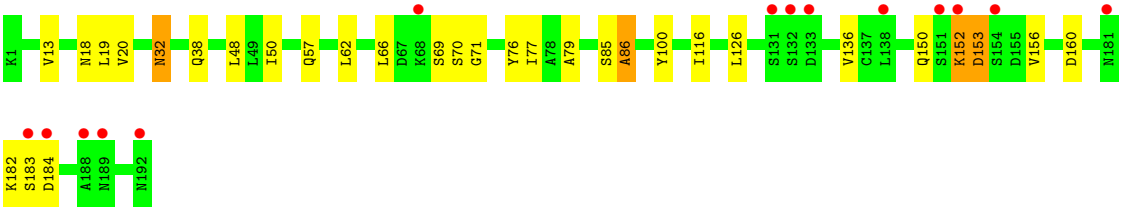


- Molecule 3: Cancer/testis antigen 1B

Chain C:

- Molecule 4: T-CELL RECEPTOR, ALPHA CHAIN

Chain D:



- Molecule 5: Hypothetical protein

Chain E:

85%
13%

M0
F8
Q9
L19
Q23
D24
M25
N26
Q35
R42
L43
I44
H45
V48
S49
V50
G51
M52
T53
D54
E57
V58
Y62
N63
V86
L94
G95
N96
R107
L108
T109
V110
L111
V124
H151
L154
D170
L174
L187
T196
A234
A240
D241
GLN

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	74.59Å 52.33Å 118.93Å 90.00° 98.13° 90.00°	Depositor
Resolution (Å)	24.80 – 2.20 24.81 – 2.20	Depositor EDS
% Data completeness (in resolution range)	99.3 (24.80-2.20) 99.3 (24.81-2.20)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	0.14	Depositor
$< I/\sigma(I) >$ [1]	2.42 (at 2.19Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.165 , 0.245 0.169 , 0.245	Depositor DCC
R_{free} test set	2334 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å^2)	26.5	Xtriage
Anisotropy	0.245	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 46.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	7417	wwPDB-VP
Average B, all atoms (Å^2)	18.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.44% 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, EPE, CA, MG, 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	0.80	1/2339 (0.0%)	0.80	1/3175 (0.0%)
2	B	0.76	0/868	0.84	1/1173 (0.1%)
3	C	0.70	0/76	1.00	0/101
4	D	0.71	0/1528	0.84	0/2079
5	E	0.81	0/1999	0.84	1/2720 (0.0%)
All	All	0.78	1/6810 (0.0%)	0.83	3/9248 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	19	GLU	CB-CG	5.51	1.62	1.52

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	45	ARG	NE-CZ-NH1	7.06	123.83	120.30
1	A	131	ARG	NE-CZ-NH1	6.04	123.32	120.30
5	E	174	LEU	CA-CB-CG	5.06	126.94	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	2267	0	2122	32	0
2	B	845	0	807	10	0
3	C	75	0	79	3	0
4	D	1497	0	1441	22	0
5	E	1945	0	1840	32	0
6	A	15	0	0	0	0
6	B	5	0	0	0	0
6	D	5	0	0	0	0
6	E	5	0	0	0	0
7	A	24	0	32	1	0
7	B	12	0	16	0	0
7	D	24	0	32	3	0
7	E	6	0	8	0	0
8	B	1	0	0	0	0
9	B	1	0	0	0	0
10	E	15	0	17	0	0
11	A	224	0	0	3	1
11	B	104	0	0	1	1
11	C	10	0	0	0	0
11	D	115	0	0	0	0
11	E	222	0	0	4	0
All	All	7417	0	6394	91	1

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 7.

All (91) 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)
5:E:109:THR:HG22	5:E:111:LEU:CD1	1.91	1.01
5:E:42:ARG:NH1	5:E:58:VAL:HG11	1.84	0.92
5:E:109:THR:HG22	5:E:111:LEU:HD11	1.53	0.86
2:B:21:ASN:HD22	2:B:22:PHE:H	1.18	0.86
4:D:32:ASN:C	4:D:32:ASN:HD22	1.83	0.82
5:E:42:ARG:CZ	5:E:58:VAL:HG11	2.11	0.81
1:A:82:ARG:HB2	7:A:1008:GOL:H31	1.65	0.79
5:E:109:THR:CG2	5:E:111:LEU:HD11	2.16	0.76
2:B:13:HIS:H	2:B:21:ASN:HD21	1.33	0.73
4:D:153:ASP:OD1	4:D:182:LYS:NZ	2.21	0.73
3:C:3:LEU:HD23	3:C:6:ILE:HD11	1.72	0.71
5:E:52:MET:HE2	5:E:54[B]:ASP:OD2	1.90	0.71
5:E:109:THR:HG22	5:E:111:LEU:HD12	1.72	0.69

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:21:ASN:ND2	2:B:22:PHE:H	1.92	0.65
4:D:38:GLN:HE22	5:E:35:GLN:HE22	1.43	0.65
2:B:21:ASN:HD22	2:B:22:PHE:N	1.94	0.64
4:D:69:SER:O	4:D:71:GLY:N	2.28	0.64
2:B:85:VAL:HG13	11:B:1032:HOH:O	1.97	0.62
5:E:48:VAL:HG22	5:E:52:MET:HE3	1.82	0.61
5:E:9:GLN:HB3	5:E:108:LEU:HD13	1.83	0.61
4:D:20:VAL:O	4:D:20:VAL:HG23	2.01	0.60
1:A:190:THR:OG1	1:A:202:ARG:HB3	2.02	0.59
1:A:180:GLN:NE2	11:A:1138:HOH:O	2.36	0.58
5:E:19:LEU:HD21	5:E:108:LEU:HD21	1.84	0.58
5:E:42:ARG:NH1	5:E:58:VAL:CG1	2.61	0.58
5:E:52:MET:CE	5:E:54[B]:ASP:OD2	2.51	0.58
1:A:19:GLU:HG2	1:A:20:PRO:HD2	1.86	0.57
1:A:234:ARG:HE	1:A:242:GLN:HE21	1.53	0.56
5:E:240:ALA:O	5:E:241:ASP:HB2	2.05	0.56
1:A:119:ASP:HB3	2:B:0:MET:HG3	1.88	0.56
5:E:151:HIS:CD2	11:E:1074:HOH:O	2.59	0.55
5:E:196:THR:HG23	11:E:1107:HOH:O	2.07	0.55
1:A:81:LEU:HD13	1:A:118:TYR:CD1	2.42	0.54
5:E:151:HIS:HD2	11:E:1074:HOH:O	1.90	0.54
2:B:45:ARG:CG	2:B:45:ARG:HH11	2.21	0.54
4:D:116:ILE:HG23	7:D:1003:GOL:H12	1.90	0.53
1:A:96:GLN:OE1	2:B:31:HIS:HE1	1.89	0.53
4:D:150:GLN:HE21	4:D:160:ASP:HA	1.73	0.53
1:A:202:ARG:HD2	1:A:244:TRP:CD2	2.45	0.52
5:E:26:ASN:ND2	11:E:1040:HOH:O	2.43	0.52
1:A:228:THR:HG22	1:A:247:VAL:HG23	1.92	0.52
5:E:42:ARG:HD2	5:E:57:GLU:OE2	2.10	0.52
5:E:44:ILE:HG22	5:E:45:HIS:CD2	2.45	0.51
4:D:48:LEU:HD22	4:D:57:GLN:HG2	1.92	0.51
1:A:266:LEU:HD13	1:A:270:LEU:HG	1.92	0.51
4:D:32:ASN:C	4:D:32:ASN:ND2	2.57	0.51
3:C:3:LEU:HD23	3:C:6:ILE:CD1	2.40	0.50
1:A:131:ARG:CG	1:A:131:ARG:HH11	2.24	0.50
5:E:124:VAL:HG23	5:E:234:ALA:HB3	1.94	0.49
1:A:215:LEU:CD2	1:A:261:VAL:HG22	2.43	0.49
1:A:19:GLU:HG3	1:A:75:ARG:HD2	1.93	0.49
5:E:48:VAL:CG2	5:E:52:MET:HE3	2.43	0.49
5:E:86:VAL:HG22	5:E:107[A]:ARG:HG3	1.94	0.49
4:D:85:SER:O	4:D:86:ALA:HB2	2.13	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:D:183:SER:O	4:D:184:ASP:HB2	2.13	0.48
1:A:143:THR:HG23	3:C:9:CYS:HA	1.95	0.48
1:A:249:VAL:HG12	1:A:257:TYR:CZ	2.48	0.48
5:E:62:TYR:C	5:E:63:ASN:HD22	2.17	0.48
1:A:117:ALA:HB2	2:B:60:TRP:CE2	2.49	0.48
4:D:152:LYS:NZ	4:D:156:VAL:HG11	2.30	0.47
1:A:192:HIS:NE2	2:B:98:ASP:OD1	2.30	0.47
4:D:18:ASN:ND2	4:D:79:ALA:H	2.13	0.47
4:D:62:LEU:CD2	4:D:77:ILE:HG12	2.45	0.47
1:A:249:VAL:HG12	1:A:257:TYR:CE2	2.51	0.46
1:A:51:TRP:CD1	1:A:51:TRP:C	2.90	0.45
1:A:188:HIS:HD2	11:A:1123:HOH:O	1.99	0.45
1:A:78:LEU:HD23	1:A:95:VAL:HG23	1.97	0.45
4:D:126:LEU:N	4:D:126:LEU:HD12	2.32	0.45
1:A:5:MET:HB2	1:A:168:LEU:HD13	1.99	0.45
4:D:20:VAL:O	4:D:20:VAL:CG2	2.64	0.45
1:A:75:ARG:NH1	5:E:50:VAL:HG21	2.32	0.45
1:A:19:GLU:HG2	1:A:20:PRO:CD	2.46	0.44
5:E:154:LEU:C	5:E:154:LEU:HD23	2.37	0.44
1:A:126:LEU:HG	1:A:130:LEU:HA	2.00	0.44
1:A:102:ASP:OD2	1:A:111:ARG:HD3	2.18	0.43
4:D:50:ILE:HD12	4:D:66:LEU:HB2	1.99	0.43
5:E:23:GLN:NE2	5:E:25:MET:H	2.15	0.43
5:E:86:VAL:HG22	5:E:107[B]:ARG:HG2	2.00	0.43
4:D:13:VAL:HG21	4:D:19:LEU:HD21	2.00	0.42
1:A:195:SER:OG	1:A:196:ASP:N	2.52	0.42
1:A:19:GLU:CG	1:A:20:PRO:HD2	2.47	0.41
1:A:182:THR:HG23	11:A:1182:HOH:O	2.21	0.41
1:A:6:ARG:NH1	1:A:102:ASP:OD1	2.50	0.41
4:D:100:TYR:CD2	5:E:94:LEU:HD22	2.55	0.41
4:D:76:TYR:HB3	7:D:1004:GOL:H31	2.02	0.41
4:D:150:GLN:HE21	4:D:160:ASP:CA	2.34	0.41
1:A:181:ARG:O	1:A:181:ARG:CG	2.69	0.41
4:D:76:TYR:HB3	7:D:1004:GOL:H12	2.03	0.41
5:E:170:ASP:HB2	5:E:187:LEU:HD12	2.02	0.41
5:E:8:PHE:CD1	5:E:151:HIS:HB3	2.55	0.40
5:E:52:MET:C	5:E:52:MET:SD	2.99	0.40

All (1) 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)
11:A:1054:HOH:O	11:B:1085:HOH:O[2_556]	2.14	0.06

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	276/276 (100%)	268 (97%)	8 (3%)	0	100	100
2	B	99/100 (99%)	96 (97%)	3 (3%)	0	100	100
3	C	7/9 (78%)	7 (100%)	0	0	100	100
4	D	193/192 (100%)	179 (93%)	10 (5%)	4 (2%)	7	4
5	E	244/243 (100%)	236 (97%)	8 (3%)	0	100	100
All	All	819/820 (100%)	786 (96%)	29 (4%)	4 (0%)	29	31

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	D	70	SER
4	D	153	ASP
4	D	86	ALA
4	D	152	LYS

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	234/232 (101%)	229 (98%)	5 (2%)	53	67

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	B	96/95 (101%)	93 (97%)	3 (3%)	40	51
3	C	9/9 (100%)	9 (100%)	0	100	100
4	D	172/169 (102%)	170 (99%)	2 (1%)	71	83
5	E	215/212 (101%)	211 (98%)	4 (2%)	57	71
All	All	726/717 (101%)	712 (98%)	14 (2%)	55	71

All (14) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	51	TRP
1	A	86	ASN
1	A	111	ARG
1	A	131	ARG
1	A	202	ARG
2	B	21	ASN
2	B	45	ARG
2	B	70	PHE
4	D	32	ASN
4	D	136	VAL
5	E	23	GLN
5	E	52	MET
5	E	63	ASN
5	E	96	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (41) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	43	GLN
1	A	70	HIS
1	A	86	ASN
1	A	141	GLN
1	A	174	ASN
1	A	180	GLN
1	A	188	HIS
1	A	224	GLN
1	A	226	GLN
1	A	242	GLN
1	A	253	GLN
2	B	2	GLN
2	B	21	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	24	ASN
2	B	31	HIS
2	B	89	GLN
4	D	6	GLN
4	D	18	ASN
4	D	32	ASN
4	D	54	GLN
4	D	63	ASN
4	D	81	GLN
4	D	113	HIS
4	D	117	GLN
4	D	125	GLN
4	D	145	GLN
4	D	147	ASN
4	D	181	ASN
4	D	189	ASN
4	D	192	ASN
5	E	9	GLN
5	E	23	GLN
5	E	26	ASN
5	E	35	GLN
5	E	45	HIS
5	E	55	GLN
5	E	60	ASN
5	E	63	ASN
5	E	96	ASN
5	E	151	HIS
5	E	181	ASN

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)

Of 20 ligands modelled in this entry, 2 are monoatomic - leaving 18 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	$\#\|Z\| > 2$
10	EPE	E	901	-	15,15,15	0.75	1 (6%)	18,20,20	1.11	1 (5%)
6	SO4	D	801	-	4,4,4	0.20	0	6,6,6	0.17	0
7	GOL	B	1001	-	5,5,5	0.37	0	5,5,5	0.37	0
6	SO4	A	803	-	4,4,4	0.20	0	6,6,6	0.13	0
7	GOL	A	1006	-	5,5,5	0.43	0	5,5,5	0.14	0
7	GOL	A	1009	-	5,5,5	0.58	0	5,5,5	0.70	0
6	SO4	E	804	-	4,4,4	0.12	0	6,6,6	0.16	0
7	GOL	A	1008	-	5,5,5	0.66	0	5,5,5	0.84	0
7	GOL	E	1002	-	5,5,5	0.43	0	5,5,5	0.42	0
7	GOL	B	1010	-	5,5,5	0.57	0	5,5,5	0.47	0
7	GOL	A	1005	-	5,5,5	0.44	0	5,5,5	0.35	0
7	GOL	D	1004	-	5,5,5	0.33	0	5,5,5	0.26	0
6	SO4	A	805	-	4,4,4	0.14	0	6,6,6	0.14	0
7	GOL	D	1011	-	5,5,5	0.60	0	5,5,5	0.30	0
6	SO4	B	806	-	4,4,4	0.19	0	6,6,6	0.08	0
7	GOL	D	1003	-	5,5,5	0.58	0	5,5,5	0.90	0
6	SO4	A	802	-	4,4,4	0.15	0	6,6,6	0.06	0
7	GOL	D	1007	-	5,5,5	0.41	0	5,5,5	0.21	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
10	EPE	E	901	-	-	2/9/19/19	0/1/1/1
7	GOL	B	1001	-	-	4/4/4/4	-
7	GOL	A	1006	-	-	4/4/4/4	-
7	GOL	A	1009	-	-	2/4/4/4	-
7	GOL	A	1008	-	-	2/4/4/4	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
7	GOL	E	1002	-	-	3/4/4/4	-
7	GOL	B	1010	-	-	2/4/4/4	-
7	GOL	A	1005	-	-	0/4/4/4	-
7	GOL	D	1004	-	-	4/4/4/4	-
7	GOL	D	1011	-	-	4/4/4/4	-
7	GOL	D	1003	-	-	2/4/4/4	-
7	GOL	D	1007	-	-	2/4/4/4	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
10	E	901	EPE	C10-S	2.56	1.81	1.77

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
10	E	901	EPE	O3S-S-C10	2.72	110.17	105.77

There are no chirality outliers.

All (31) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	A	1006	GOL	O1-C1-C2-C3
7	A	1009	GOL	O1-C1-C2-C3
7	B	1001	GOL	C1-C2-C3-O3
7	D	1007	GOL	O1-C1-C2-C3
7	B	1010	GOL	O1-C1-C2-O2
7	A	1006	GOL	C1-C2-C3-O3
7	A	1008	GOL	C1-C2-C3-O3
7	B	1001	GOL	O1-C1-C2-C3
7	B	1010	GOL	O1-C1-C2-C3
7	D	1003	GOL	O1-C1-C2-C3
7	D	1004	GOL	O1-C1-C2-C3
7	D	1004	GOL	C1-C2-C3-O3
7	D	1011	GOL	O1-C1-C2-C3
7	D	1011	GOL	C1-C2-C3-O3
7	E	1002	GOL	O1-C1-C2-C3
7	A	1006	GOL	O2-C2-C3-O3
7	A	1008	GOL	O2-C2-C3-O3
7	A	1009	GOL	O1-C1-C2-O2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
7	B	1001	GOL	O2-C2-C3-O3
7	D	1007	GOL	O1-C1-C2-O2
7	D	1011	GOL	O2-C2-C3-O3
7	E	1002	GOL	O1-C1-C2-O2
7	A	1006	GOL	O1-C1-C2-O2
7	D	1004	GOL	O1-C1-C2-O2
7	D	1004	GOL	O2-C2-C3-O3
7	B	1001	GOL	O1-C1-C2-O2
7	D	1003	GOL	O1-C1-C2-O2
7	D	1011	GOL	O1-C1-C2-O2
10	E	901	EPE	C10-C9-N1-C2
10	E	901	EPE	C10-C9-N1-C6
7	E	1002	GOL	C1-C2-C3-O3

There are no ring outliers.

3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	A	1008	GOL	1	0
7	D	1004	GOL	2	0
7	D	1003	GOL	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%)	-0.28	7 (2%)	57	55	5, 15, 34, 70	0
2	B	100/100 (100%)	-0.37	1 (1%)	82	81	6, 14, 32, 61	0
3	C	9/9 (100%)	0.37	0	100	100	3, 5, 9, 11	0
4	D	192/192 (100%)	0.09	14 (7%)	15	14	7, 18, 55, 68	0
5	E	242/243 (99%)	-0.41	1 (0%)	92	91	5, 12, 27, 45	0
All	All	819/820 (99%)	-0.23	23 (2%)	53	51	3, 15, 37, 70	0

All (23) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	1	ILE	5.8
4	D	183	SER	4.3
4	D	132	SER	4.3
4	D	184	ASP	3.8
5	E	0	MET	3.7
1	A	17	ARG	3.6
4	D	154	SER	3.6
1	A	90	ALA	3.3
4	D	68	LYS	3.3
1	A	88	SER	3.1
4	D	152	LYS	3.1
4	D	189	ASN	3.0
1	A	275	GLU	2.6
4	D	188	ALA	2.6
4	D	192	ASN	2.6
4	D	151	SER	2.4
1	A	276	PRO	2.3
1	A	195	SER	2.3
4	D	138	LEU	2.2
4	D	133	ASP	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
4	D	181	ASN	2.1
4	D	131[A]	SER	2.1
1	A	86	ASN	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 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
6	SO4	B	806	5/5	0.71	0.47	100,100,100,100	0
7	GOL	B	1010	6/6	0.71	0.24	38,44,48,50	0
7	GOL	D	1003	6/6	0.79	0.30	21,25,31,39	0
7	GOL	A	1006	6/6	0.82	0.24	30,40,51,54	0
7	GOL	A	1009	6/6	0.86	0.30	22,27,35,39	0
7	GOL	D	1004	6/6	0.86	0.30	26,28,37,59	0
7	GOL	A	1008	6/6	0.87	0.34	9,27,35,51	0
7	GOL	A	1005	6/6	0.88	0.15	28,37,40,48	0
6	SO4	A	805	5/5	0.89	0.29	65,77,80,90	0
7	GOL	B	1001	6/6	0.91	0.19	22,26,38,49	0
6	SO4	A	802	5/5	0.91	0.18	90,91,95,97	0
7	GOL	D	1007	6/6	0.92	0.18	42,44,54,56	0
6	SO4	D	801	5/5	0.93	0.30	52,55,64,72	0
6	SO4	A	803	5/5	0.93	0.35	51,60,72,78	0
7	GOL	E	1002	6/6	0.93	0.12	22,27,30,50	0
6	SO4	E	804	5/5	0.95	0.25	38,50,61,63	0
7	GOL	D	1011	6/6	0.96	0.23	17,32,47,49	0
8	CA	B	701	1/1	0.97	0.12	44,44,44,44	0
9	MG	B	702	1/1	0.97	0.07	23,23,23,23	0
10	EPE	E	901	15/15	0.97	0.22	35,54,75,76	0

6.5 Other polymers (i)

There are no such residues in this entry.