Full wwPDB X-ray Structure Validation Report (i)

Jan 23, 2017 – 07:59 PM EST

PDB ID	:	5MEN
Title	:	Human Leukocyte Antigen A02 presenting ILAKFLHWL, in complex with cognate T-Cell Receptor
Authors	:	Rizkallah, P.J.; Lloyd, A.; Crowther, M.; Cole, D.K.; Sewell, A.K.
Deposited on	:	2016-11-16
Resolution	:	2.81 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
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

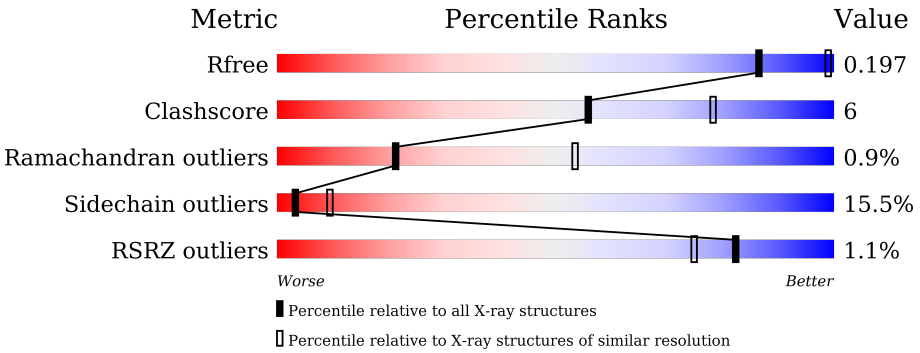
MolProbity	:	4.02b-467
Mogul	:	1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20028442
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 2.81 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	2676 (2.84-2.80)
Clashscore	102246	3124 (2.84-2.80)
Ramachandran outliers	100387	3072 (2.84-2.80)
Sidechain outliers	100360	3074 (2.84-2.80)
RSRZ outliers	91569	2690 (2.84-2.80)

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 >=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 <=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	73% 24% .
2	B	100	82% 17% .
3	C	9	56% 33% 11%
4	D	200	% 67% 27% 7% .
5	E	240	3% 75% 18% 6% .

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	EDO	A	302	-	-	-	X
6	EDO	B	301	-	-	-	X
6	EDO	B	302	-	-	-	X
6	EDO	B	303	-	-	-	X
7	GOL	A	303	-	-	-	X
8	SO4	A	305	-	-	-	X

2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 6739 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
1	A	276	Total	C	N	O	S	0	0	0
			2254	1408	410	427	9			

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

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

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	0	MET	-	initiating methionine	UNP P61769

- Molecule 3 is a protein called ILE-LEU-ALA-LYS-PHE-LEU-HIS-TRP-LEU.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	9	Total	C	N	O	0	0	0
			82	59	13	10			

- Molecule 4 is a protein called Protein TRAV22,Human nkt tcr alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	200	Total	C	N	O	S	0	0	0
			1562	978	259	318	7			

There are 29 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	90	ASP	-	linker	UNP A0A0B4J277
D	91	SER	-	linker	UNP A0A0B4J277
D	92	ALA	-	linker	UNP A0A0B4J277

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	93	THR	-	linker	UNP A0A0B4J277
D	94	SER	-	linker	UNP A0A0B4J277
D	95	GLY	-	linker	UNP A0A0B4J277
D	96	THR	-	linker	UNP A0A0B4J277
D	97	TYR	-	linker	UNP A0A0B4J277
D	98	LYS	-	linker	UNP A0A0B4J277
D	99	TYR	-	linker	UNP A0A0B4J277
D	100	ILE	-	linker	UNP A0A0B4J277
D	101	PHE	-	linker	UNP A0A0B4J277
D	102	GLY	-	linker	UNP A0A0B4J277
D	103	THR	-	linker	UNP A0A0B4J277
D	104	GLY	-	linker	UNP A0A0B4J277
D	105	THR	-	linker	UNP A0A0B4J277
D	106	ARG	-	linker	UNP A0A0B4J277
D	107	LEU	-	linker	UNP A0A0B4J277
D	108	LYS	-	linker	UNP A0A0B4J277
D	109	VAL	-	linker	UNP A0A0B4J277
D	110	LEU	-	linker	UNP A0A0B4J277
D	111	ALA	-	linker	UNP A0A0B4J277
D	112	ASN	-	linker	UNP A0A0B4J277
D	113	ILE	ILE	linker	UNP K7N5M3
D	114	GLN	GLN	linker	UNP K7N5M3
D	115	ASN	ASN	linker	UNP K7N5M3
D	116	PRO	PRO	linker	UNP K7N5M3
D	117	ASP	ASP	linker	UNP K7N5M3
D	118	PRO	PRO	linker	UNP K7N5M3

- Molecule 5 is a protein called Protein TRBV6-5,Human nkt tcr beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	240	Total	C	N	O	S	0	0	0
			1894	1190	327	368	9			

There are 22 discrepancies between the modelled and reference sequences:

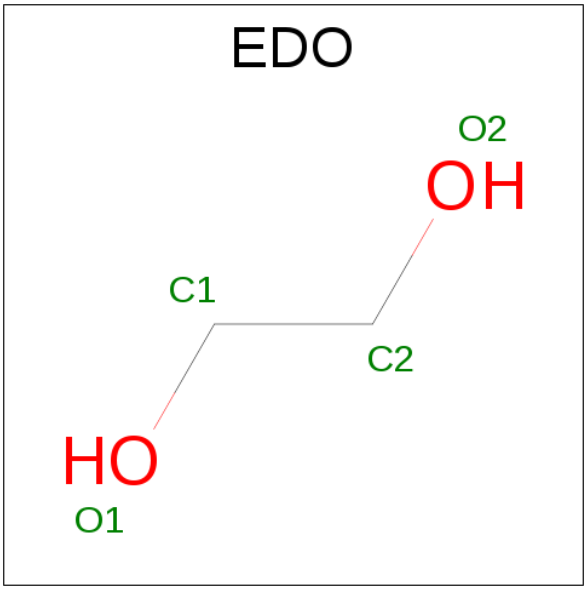
Chain	Residue	Modelled	Actual	Comment	Reference
E	96	GLN	-	linker	UNP A0A0K0K1A5
E	97	GLY	-	linker	UNP A0A0K0K1A5
E	98	THR	-	linker	UNP A0A0K0K1A5
E	99	GLU	-	linker	UNP A0A0K0K1A5
E	100	ALA	-	linker	UNP A0A0K0K1A5
E	101	PHE	-	linker	UNP A0A0K0K1A5

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	102	PHE	-	linker	UNP A0A0K0K1A5
E	103	GLY	-	linker	UNP A0A0K0K1A5
E	104	GLN	-	linker	UNP A0A0K0K1A5
E	105	GLY	-	linker	UNP A0A0K0K1A5
E	106	THR	-	linker	UNP A0A0K0K1A5
E	107	ARG	-	linker	UNP A0A0K0K1A5
E	108	LEU	-	linker	UNP A0A0K0K1A5
E	109	THR	-	linker	UNP A0A0K0K1A5
E	110	VAL	-	linker	UNP A0A0K0K1A5
E	111	VAL	-	linker	UNP A0A0K0K1A5
E	112	GLU	-	linker	UNP A0A0K0K1A5
E	113	ASP	-	linker	UNP A0A0K0K1A5
E	114	LEU	-	linker	UNP A0A0K0K1A5
E	115	ASN	-	linker	UNP A0A0K0K1A5
E	116	LYS	-	linker	UNP A0A0K0K1A5
E	200	ASP	ASN	conflict	UNP K7N5M4

- Molecule 6 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



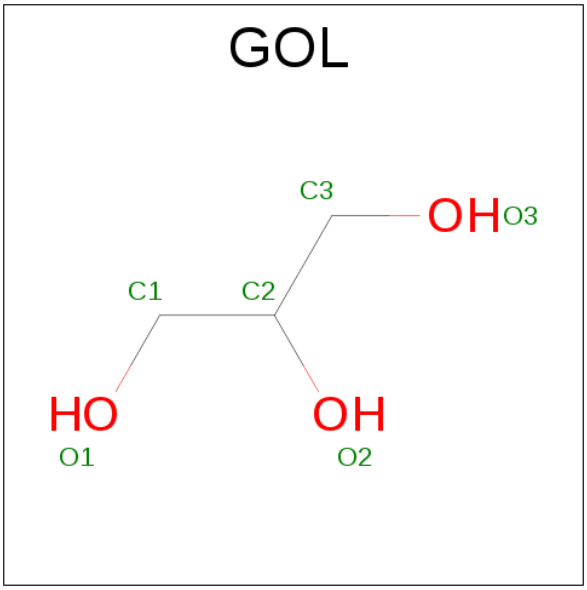
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total C O 4 2 2	0	0
6	A	1	Total C O 4 2 2	0	0
6	B	1	Total C O 4 2 2	0	0

Continued on next page...

Continued from previous page...

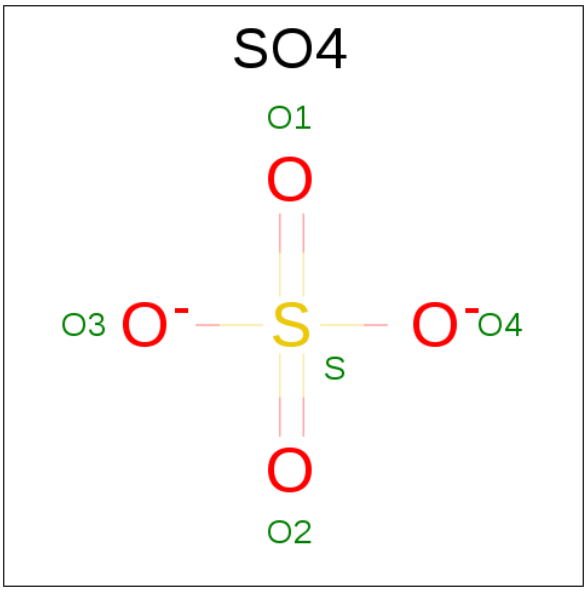
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	B	1	Total C O 4 2 2	0	0
6	B	1	Total C O 4 2 2	0	0

- 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 6 3 3	0	0

- Molecule 8 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	1	Total O S 5 4 1	0	0
8	A	1	Total O S 5 4 1	0	0
8	A	1	Total O S 5 4 1	0	0
8	B	1	Total O S 5 4 1	0	0
8	E	1	Total O S 5 4 1	0	0

- Molecule 9 is water.

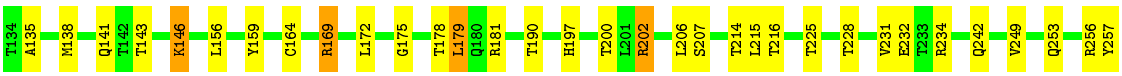
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	A	23	Total O 23 23	0	0
9	B	13	Total O 13 13	0	0
9	C	1	Total O 1 1	0	0
9	D	11	Total O 11 11	0	0
9	E	11	Total O 11 11	0	0

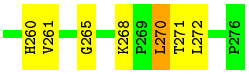
3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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:



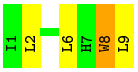


- Molecule 2: Beta-2-microglobulin

Chain B:

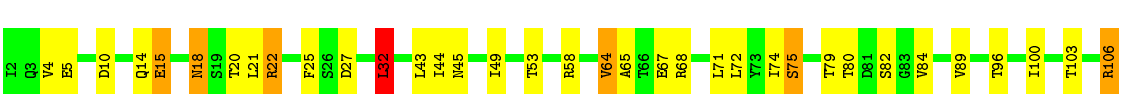
- Molecule 3: ILE-LEU-ALA-LYS-PHE-LEU-HIS-TRP-LEU

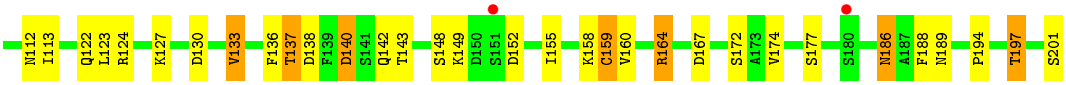
Chain C:



- Molecule 4: Protein TRAV22,Human nkt tcr alpha chain

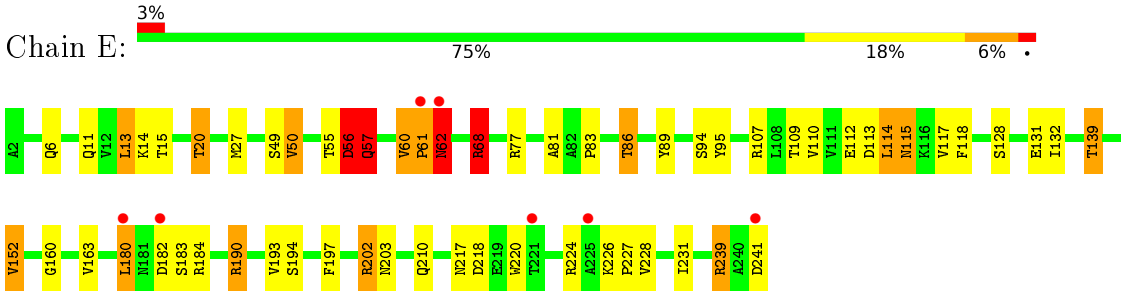
Chain D:





- Molecule 5: Protein TRBV6-5,Human nkt tcr beta chain

Chain E:


3%
75%
18%
6%
A2 Q6 Q11 V12 L13 K14 T15 T20 M27 S49 V50 T55 D56 Q57 V60 P61 M62 R68 R77 A81 A82 P83 T86 Y89 S94 Y95 R107 L108 T109 V110 V111 E112 D113 L114 N115 K116 V117 F118 S128 E131 I132 T139
V152 G160 V163 L180 N181 D182 S183 R184 R190 V193 S194 F197 R202 N203 Q210 N217 D218 E219 W220 T221 R224 A225 K226 P227 V228 I231 R239 A240 D241

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	93.16Å 48.69Å 118.10Å 90.00° 108.15° 90.00°	Depositor
Resolution (Å)	38.03 – 2.81 56.11 – 2.81	Depositor EDS
% Data completeness (in resolution range)	99.6 (38.03-2.81) 99.6 (56.11-2.81)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.10	Depositor
$< I/\sigma(I) >$ [1]	2.07 (at 2.81Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.189 , 0.272 0.199 , 0.197	Depositor DCC
R_{free} test set	1274 reflections (5.37%)	DCC
Wilson B-factor (Å^2)	55.3	Xtriage
Anisotropy	0.296	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 41.0	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	0.94	EDS
Total number of atoms	6739	wwPDB-VP
Average B, all atoms (Å^2)	57.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.24% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, SO4, EDO

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	0.90	0/2320	1.07	12/3149 (0.4%)
2	B	0.99	0/860	1.03	1/1162 (0.1%)
3	C	1.16	1/85 (1.2%)	1.05	0/113
4	D	0.83	1/1596 (0.1%)	0.99	3/2171 (0.1%)
5	E	0.82	0/1945	0.99	5/2649 (0.2%)
All	All	0.88	2/6806 (0.0%)	1.02	21/9244 (0.2%)

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
1	A	0	2
4	D	0	1
5	E	0	3
All	All	0	6

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	8	TRP	CB-CG	-6.00	1.39	1.50
4	D	5	GLU	CG-CD	5.07	1.59	1.51

All (21) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	E	180	LEU	CA-CB-CG	8.37	134.56	115.30
1	A	5	MET	CG-SD-CE	-7.10	88.84	100.20
5	E	239	ARG	NE-CZ-NH1	6.70	123.65	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	E	202	ARG	NE-CZ-NH1	6.45	123.52	120.30
1	A	45	MET	CG-SD-CE	6.32	110.31	100.20
1	A	270	LEU	CB-CG-CD1	6.20	121.54	111.00
1	A	101	CYS	CB-CA-C	-6.20	98.00	110.40
5	E	68	ARG	CG-CD-NE	5.92	124.22	111.80
1	A	270	LEU	CB-CG-CD2	5.71	120.71	111.00
5	E	56	ASP	CA-C-N	-5.53	105.03	117.20
1	A	202	ARG	NE-CZ-NH1	5.52	123.06	120.30
1	A	75	ARG	NE-CZ-NH1	5.37	122.99	120.30
4	D	32	LEU	CA-CB-CG	5.30	127.48	115.30
1	A	82	ARG	NE-CZ-NH1	5.29	122.94	120.30
4	D	164	ARG	NE-CZ-NH1	5.24	122.92	120.30
1	A	35	ARG	NE-CZ-NH2	-5.24	117.68	120.30
1	A	82	ARG	NE-CZ-NH2	-5.20	117.70	120.30
4	D	22	ARG	NE-CZ-NH2	5.14	122.87	120.30
1	A	102	ASP	CB-CG-OD1	5.12	122.91	118.30
2	B	76	ASP	CB-CG-OD1	5.07	122.87	118.30
1	A	108	ARG	NE-CZ-NH1	5.07	122.83	120.30

There are no chirality outliers.

All (6) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	105	SER	Peptide
1	A	175	GLY	Peptide
4	D	127	LYS	Peptide
5	E	160	GLY	Peptide
5	E	56	ASP	Peptide
5	E	57	GLN	Peptide

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	2254	0	2103	20	0
2	B	837	0	803	7	0
3	C	82	0	90	5	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	D	1562	0	1480	26	0
5	E	1894	0	1787	29	0
6	A	8	0	12	0	0
6	B	12	0	18	0	0
7	A	6	0	8	0	0
8	A	15	0	0	0	0
8	B	5	0	0	0	0
8	E	5	0	0	0	0
9	A	23	0	0	0	0
9	B	13	0	0	0	0
9	C	1	0	0	0	0
9	D	11	0	0	0	0
9	E	11	0	0	0	0
All	All	6739	0	6301	80	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 6.

All (80) 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)
4:D:159:CYS:SG	5:E:190:ARG:NH2	2.33	1.02
2:B:4:THR:HA	2:B:86:THR:HG21	1.53	0.89
1:A:107:TRP:O	1:A:169:ARG:NH1	2.13	0.81
1:A:67:VAL:HG13	3:C:2:LEU:HD11	1.69	0.73
4:D:32:LEU:HD22	4:D:89:VAL:HG22	1.69	0.73
1:A:69:ALA:HA	1:A:72:GLN:HE21	1.57	0.69
5:E:115:ASN:HD22	5:E:115:ASN:N	1.92	0.68
5:E:60:VAL:HG22	5:E:61:PRO:HD2	1.75	0.68
4:D:4:VAL:HG12	4:D:100:ILE:HG22	1.75	0.67
4:D:137:THR:OG1	4:D:138:ASP:N	2.26	0.66
5:E:83:PRO:O	5:E:86:THR:HG22	1.96	0.65
1:A:234:ARG:HE	1:A:242:GLN:HE21	1.49	0.61
4:D:65:ALA:O	4:D:68:ARG:NH1	2.33	0.61
5:E:14:LYS:HD3	5:E:114:LEU:HD11	1.83	0.59
4:D:113:ILE:HG13	4:D:140:ASP:HA	1.83	0.59
2:B:4:THR:CA	2:B:86:THR:HG21	2.30	0.58
5:E:60:VAL:HG22	5:E:61:PRO:CD	2.34	0.57
1:A:249:VAL:HG22	1:A:257:TYR:CZ	2.41	0.55
5:E:11:GLN:HG3	5:E:13:LEU:HD13	1.89	0.55
5:E:55:THR:OG1	5:E:57:GLN:NE2	2.33	0.54

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:D:64:VAL:O	4:D:67:GLU:O	2.25	0.54
4:D:148:SER:CB	4:D:155:ILE:HD12	2.38	0.53
4:D:25:PHE:O	4:D:68:ARG:HB3	2.09	0.53
5:E:202:ARG:HB2	5:E:202:ARG:NH1	2.24	0.52
4:D:186:ASN:HB2	4:D:189:ASN:ND2	2.25	0.52
5:E:86:THR:HB	5:E:109:THR:HA	1.92	0.51
1:A:249:VAL:HG22	1:A:257:TYR:CE1	2.47	0.49
5:E:15:THR:O	5:E:81:ALA:O	2.31	0.49
1:A:74:HIS:CE1	1:A:97:ARG:HH21	2.31	0.48
4:D:18:ASN:HB3	4:D:75:SER:HA	1.96	0.48
4:D:4:VAL:CG1	4:D:100:ILE:HG22	2.43	0.48
3:C:8:TRP:CD1	3:C:8:TRP:N	2.81	0.48
2:B:83:ASN:HA	2:B:87:LEU:HD12	1.96	0.48
1:A:143:THR:OG1	3:C:9:LEU:O	2.30	0.47
5:E:152:VAL:HA	5:E:210:GLN:O	2.14	0.47
4:D:84:VAL:HG22	4:D:106:ARG:HD3	1.95	0.47
4:D:152:ASP:O	4:D:177:SER:OG	2.28	0.47
5:E:197:PHE:O	5:E:203:ASN:ND2	2.47	0.47
1:A:67:VAL:CG1	3:C:2:LEU:HD11	2.40	0.46
4:D:123:LEU:HB2	4:D:133:VAL:HG12	1.97	0.46
4:D:15:GLU:OE1	4:D:112:ASN:N	2.49	0.45
5:E:56:ASP:C	5:E:57:GLN:HG3	2.37	0.45
1:A:172:LEU:HD23	1:A:179:LEU:HD12	1.99	0.45
4:D:172:SER:OG	5:E:190:ARG:NH1	2.48	0.45
4:D:45:ASN:N	4:D:45:ASN:HD22	2.15	0.45
4:D:136:PHE:HB2	4:D:188:PHE:CE2	2.52	0.45
1:A:190:THR:OG1	1:A:202:ARG:HB3	2.17	0.44
2:B:40:LEU:HD23	2:B:45:ARG:HA	1.98	0.44
4:D:18:ASN:CB	4:D:74:ILE:O	2.66	0.44
5:E:6:GLN:NE2	5:E:89:TYR:O	2.45	0.44
4:D:21:LEU:HD22	4:D:72:LEU:HD23	1.99	0.43
5:E:118:PHE:CD2	5:E:184:ARG:HD3	2.53	0.43
5:E:62:ASN:HA	5:E:62:ASN:HD22	1.61	0.43
5:E:56:ASP:CA	5:E:57:GLN:HG3	2.49	0.43
1:A:216:THR:HG23	1:A:260:HIS:HB2	2.00	0.43
1:A:159:TYR:CZ	1:A:164:CYS:HB2	2.53	0.42
4:D:122:GLN:C	4:D:123:LEU:HD12	2.40	0.42
4:D:194:PRO:O	4:D:197:THR:HG22	2.20	0.42
5:E:115:ASN:N	5:E:115:ASN:ND2	2.62	0.42
5:E:20:THR:HG22	5:E:77:ARG:HD3	2.01	0.42
1:A:146:LYS:HG3	5:E:95:TYR:OH	2.19	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:55:GLU:HG2	1:A:59:TYR:CD2	2.55	0.42
4:D:20:THR:C	4:D:21:LEU:HD12	2.40	0.42
4:D:32:LEU:CD2	4:D:89:VAL:HG22	2.45	0.42
5:E:220:TRP:CE3	5:E:226:LYS:HB3	2.55	0.42
5:E:49:SER:OG	5:E:68:ARG:HG3	2.20	0.42
3:C:6:LEU:CD1	5:E:50:VAL:HG22	2.50	0.41
1:A:132:SER:OG	1:A:133:TRP:N	2.53	0.41
1:A:52:ILE:HG13	1:A:52:ILE:O	2.20	0.41
4:D:143:THR:HG21	4:D:194:PRO:HG3	2.03	0.41
5:E:128:SER:O	5:E:132:ILE:HG23	2.21	0.41
1:A:253:GLN:HE21	1:A:256:ARG:HD3	1.84	0.41
2:B:4:THR:CB	2:B:86:THR:HG21	2.52	0.40
1:A:135:ALA:HB3	1:A:141:GLN:HE21	1.87	0.40
2:B:17:ASN:OD1	2:B:97:ARG:NH2	2.49	0.40
2:B:84:HIS:ND1	2:B:86:THR:HB	2.36	0.40
5:E:131:GLU:OE2	5:E:139:THR:HG22	2.21	0.40
5:E:193:VAL:HG22	5:E:194:SER:N	2.36	0.40
1:A:64:THR:O	1:A:68:LYS:HG3	2.22	0.40
5:E:112:GLU:OE1	5:E:112:GLU:N	2.53	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	274/276 (99%)	250 (91%)	22 (8%)	2 (1%)	26 60
2	B	98/100 (98%)	94 (96%)	4 (4%)	0	100 100
3	C	7/9 (78%)	6 (86%)	1 (14%)	0	100 100
4	D	198/200 (99%)	179 (90%)	18 (9%)	1 (0%)	34 68
5	E	238/240 (99%)	207 (87%)	27 (11%)	4 (2%)	11 34

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	815/825 (99%)	736 (90%)	72 (9%)	7 (1%)	21 53

All (7) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
5	E	57	GLN
1	A	105	SER
4	D	15	GLU
5	E	62	ASN
5	E	227	PRO
5	E	61	PRO
1	A	265	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	232/232 (100%)	195 (84%)	37 (16%)	3 9
2	B	95/95 (100%)	86 (90%)	9 (10%)	11 29
3	C	8/8 (100%)	8 (100%)	0	100 100
4	D	180/180 (100%)	144 (80%)	36 (20%)	1 4
5	E	206/206 (100%)	176 (85%)	30 (15%)	4 11
All	All	721/721 (100%)	609 (84%)	112 (16%)	3 9

All (112) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	6	ARG
1	A	15	PRO
1	A	17	ARG
1	A	25	VAL
1	A	28	VAL
1	A	32	GLN
1	A	35	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	43	GLN
1	A	52	ILE
1	A	53	GLU
1	A	74	HIS
1	A	88	SER
1	A	98	MET
1	A	124	ILE
1	A	130	LEU
1	A	138	MET
1	A	146	LYS
1	A	156	LEU
1	A	169	ARG
1	A	178	THR
1	A	179	LEU
1	A	181	ARG
1	A	197	HIS
1	A	200	THR
1	A	206	LEU
1	A	207	SER
1	A	214	THR
1	A	215	LEU
1	A	225	THR
1	A	228	THR
1	A	231	VAL
1	A	232	GLU
1	A	261	VAL
1	A	268	LYS
1	A	270	LEU
1	A	271	THR
1	A	272	LEU
2	B	19	LYS
2	B	32	PRO
2	B	39	LEU
2	B	47	GLU
2	B	70	PHE
2	B	71	THR
2	B	75	LYS
2	B	85	VAL
2	B	86	THR
4	D	10	ASP
4	D	14	GLN
4	D	18	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
4	D	22	ARG
4	D	27	ASP
4	D	32	LEU
4	D	43	LEU
4	D	44	ILE
4	D	49	ILE
4	D	53	THR
4	D	58	ARG
4	D	64	VAL
4	D	71	LEU
4	D	75	SER
4	D	79	THR
4	D	80	THR
4	D	82	SER
4	D	96	THR
4	D	103	THR
4	D	106	ARG
4	D	124	ARG
4	D	130	ASP
4	D	133	VAL
4	D	137	THR
4	D	140	ASP
4	D	142	GLN
4	D	149	LYS
4	D	158	LYS
4	D	159	CYS
4	D	160	VAL
4	D	164	ARG
4	D	167	ASP
4	D	174	VAL
4	D	186	ASN
4	D	197	THR
4	D	201	SER
5	E	13	LEU
5	E	20	THR
5	E	27	MET
5	E	50	VAL
5	E	56	ASP
5	E	60	VAL
5	E	62	ASN
5	E	68	ARG
5	E	86	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
5	E	94	SER
5	E	107	ARG
5	E	110	VAL
5	E	113	ASP
5	E	114	LEU
5	E	115	ASN
5	E	117	VAL
5	E	139	THR
5	E	152	VAL
5	E	163	VAL
5	E	180	LEU
5	E	182	ASP
5	E	183	SER
5	E	190	ARG
5	E	217	ASN
5	E	218	ASP
5	E	224	ARG
5	E	228	VAL
5	E	231	ILE
5	E	239	ARG
5	E	241	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (26) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	3	HIS
1	A	43	GLN
1	A	54	GLN
1	A	72	GLN
1	A	74	HIS
1	A	174	ASN
1	A	218	GLN
1	A	242	GLN
1	A	253	GLN
2	B	8	GLN
4	D	6	GLN
4	D	14	GLN
4	D	30	ASN
4	D	45	ASN
4	D	78	GLN
4	D	112	ASN
4	D	147	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
4	D	186	ASN
5	E	37	GLN
5	E	57	GLN
5	E	62	ASN
5	E	115	ASN
5	E	136	GLN
5	E	199	GLN
5	E	222	GLN
5	E	230	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

11 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	EDO	A	301	-	3,3,3	0.47	0	2,2,2	0.23	0
6	EDO	A	302	-	3,3,3	0.63	0	2,2,2	0.15	0
7	GOL	A	303	-	5,5,5	0.31	0	5,5,5	0.95	0
8	SO4	A	304	-	4,4,4	0.51	0	6,6,6	0.28	0
8	SO4	A	305	-	4,4,4	0.37	0	6,6,6	0.19	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
8	SO4	A	306	-	4,4,4	0.45	0	6,6,6	0.12	0
6	EDO	B	301	-	3,3,3	0.46	0	2,2,2	0.41	0
6	EDO	B	302	-	3,3,3	0.58	0	2,2,2	0.27	0
6	EDO	B	303	-	3,3,3	0.55	0	2,2,2	0.27	0
8	SO4	B	304	-	4,4,4	0.42	0	6,6,6	0.35	0
8	SO4	E	301	-	4,4,4	0.40	0	6,6,6	0.14	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	EDO	A	301	-	-	0/1/1/1	0/0/0/0
6	EDO	A	302	-	-	0/1/1/1	0/0/0/0
7	GOL	A	303	-	-	0/4/4/4	0/0/0/0
8	SO4	A	304	-	-	0/0/0/0	0/0/0/0
8	SO4	A	305	-	-	0/0/0/0	0/0/0/0
8	SO4	A	306	-	-	0/0/0/0	0/0/0/0
6	EDO	B	301	-	-	0/1/1/1	0/0/0/0
6	EDO	B	302	-	-	0/1/1/1	0/0/0/0
6	EDO	B	303	-	-	0/1/1/1	0/0/0/0
8	SO4	B	304	-	-	0/0/0/0	0/0/0/0
8	SO4	E	301	-	-	0/0/0/0	0/0/0/0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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, 95^{th} 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%)	-0.11	0	100	100	25, 50, 88, 110	0
2	B	100/100 (100%)	-0.26	0	100	100	25, 41, 62, 71	0
3	C	9/9 (100%)	-0.44	0	100	100	34, 37, 39, 43	0
4	D	200/200 (100%)	-0.08	2 (1%)	84	77	36, 64, 94, 112	0
5	E	240/240 (100%)	0.07	7 (2%)	55	43	37, 60, 104, 128	0
All	All	825/825 (100%)	-0.07	9 (1%)	82	75	25, 56, 92, 128	0

All (9) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	E	61	PRO	5.8
5	E	62	ASN	5.6
5	E	180	LEU	4.3
4	D	151	SER	4.1
5	E	241	ASP	3.6
5	E	221	THR	2.8
5	E	182	ASP	2.2
5	E	225	ALA	2.1
4	D	180	SER	2.0

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

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^{th} 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^2)	Q<0.9
8	SO4	A	305	5/5	0.86	0.49	9.54	120,125,126,129	0
6	EDO	B	301	4/4	0.89	0.31	3.57	47,51,56,61	0
6	EDO	A	302	4/4	0.83	0.28	3.49	57,58,58,60	0
6	EDO	B	302	4/4	0.94	0.33	3.28	50,52,53,57	0
7	GOL	A	303	6/6	0.94	0.21	3.27	39,45,45,46	0
6	EDO	B	303	4/4	0.94	0.28	2.93	48,53,57,57	0
6	EDO	A	301	4/4	0.96	0.20	0.82	45,49,50,50	0
8	SO4	B	304	5/5	0.95	0.15	-0.88	83,84,88,90	0
8	SO4	A	304	5/5	0.95	0.14	-3.29	71,81,82,84	0
8	SO4	A	306	5/5	0.76	0.33	-	108,112,115,116	0
8	SO4	E	301	5/5	0.91	0.14	-	83,91,94,95	0

6.5 Other polymers

There are no such residues in this entry.