Full wwPDB X-ray Structure Validation Report ⓘ

Jul 26, 2023 – 12:40 AM EDT

PDB ID	:	1AO7
Title	:	COMPLEX BETWEEN HUMAN T-CELL RECEPTOR, VIRAL PEPTIDE (TAX), AND HLA-A 0201
Authors	:	Garboczi, D.N.; Ghosh, P.; Utz, U.; Fan, Q.R.; Biddison, W.E.; Wiley, D.C.
Deposited on	:	1997-07-21
Resolution	:	2.60 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 ⓘ 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 ⓘ) were used in the production of this report:

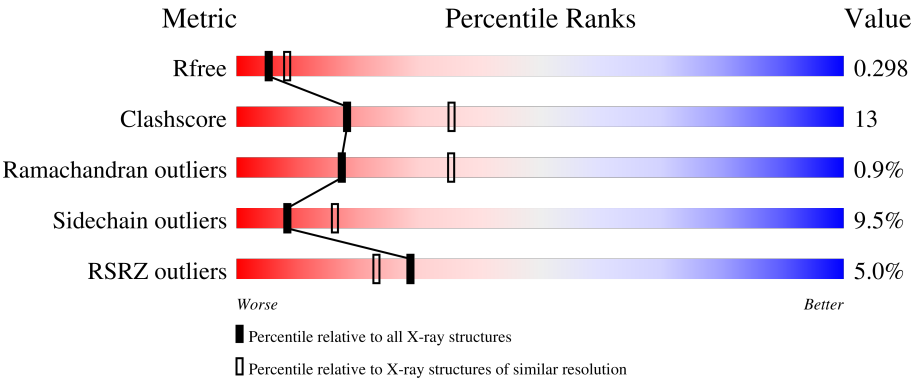
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.34
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.34

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 2.60 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	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

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	275	8% 71% 26%
2	B	100	% 77% 20%
3	C	9	56% 44%
4	D	204	% 38% 16% 44%
5	E	245	4% 52% 29% 15%

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	EMC	B	100	-	-	X	-

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 5711 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-A 0201.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	274	Total	C	N	O	S	122	0	0
			2237	1398	408	422	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	5	0	0
			828	524	140	158	6			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	67	CYS	TYR	conflict	UNP P61769
B	91	CYS	LYS	conflict	UNP P61769

- Molecule 3 is a protein called TAX PEPTIDE.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	9	Total	C	N	O	0	0	0
			77	56	9	12			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	115	Total	C	N	O	S	25	0	0
			886	553	148	182	3			

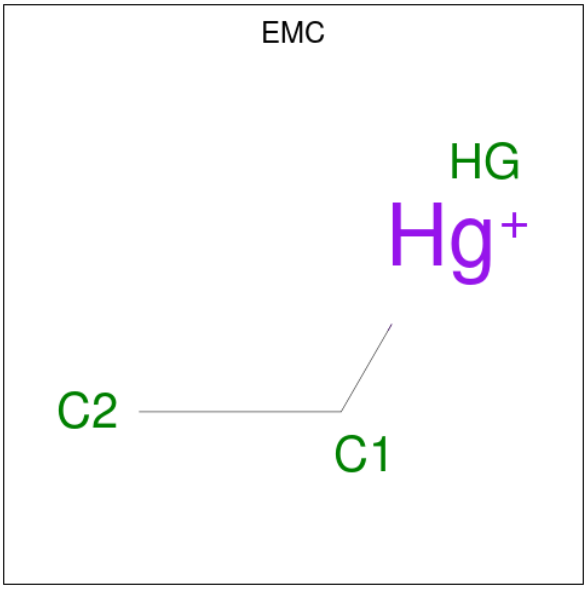
- Molecule 5 is a protein called T CELL RECEPTOR BETA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	209	Total	C	N	O	S	40	0	0
			1640	1041	286	305	8			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	95	ARG	SER	conflict	UNP 3002925
E	96	PRO	PHE	conflict	UNP 3002925
E	97	GLY	PRO	conflict	UNP 3002925
E	98	LEU	ARG	conflict	UNP 3002925
E	99	ALA	GLN	conflict	UNP 3002925
E	100	GLY	PRO	conflict	UNP 3002925
E	101	GLY	SER	conflict	UNP 3002925
E	102	ARG	TYR	conflict	UNP 3002925
E	103	PRO	ASN	conflict	UNP 3002925
E	107	TYR	PHE	conflict	UNP 3002925
E	116A	THR	LEU	conflict	UNP 3002925
E	191	ALA	CYS	conflict	UNP 3002925

- Molecule 6 is ETHYL MERCURY ION (three-letter code: EMC) (formula: C_2H_5Hg).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	B	1	Total C Hg 3 2 1	0	0
6	B	1	Total C Hg 3 2 1	0	0

- Molecule 7 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	13	Total O 13 13	0	0

Continued on next page...

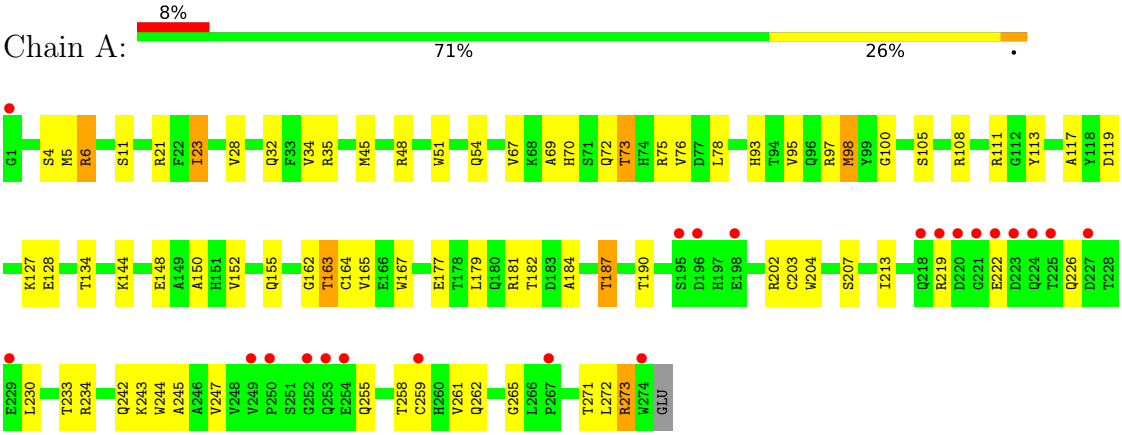
Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	B	13	Total O 13 13	0	0
7	C	1	Total O 1 1	0	0
7	D	2	Total O 2 2	0	0
7	E	8	Total O 8 8	0	0

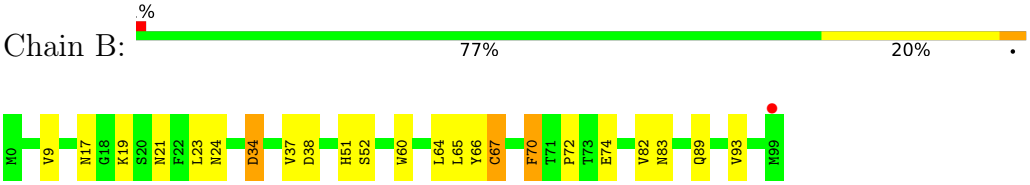
3 Residue-property plots (i)

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-A 0201

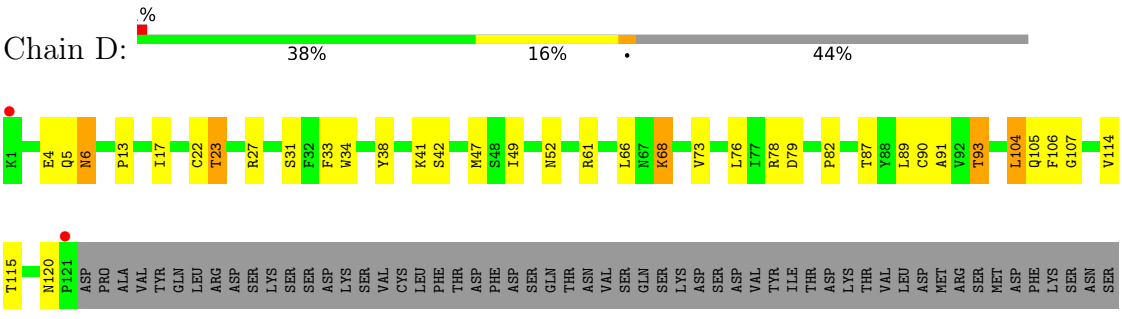


• Molecule 2: BETA-2 MICROGLOBULIN

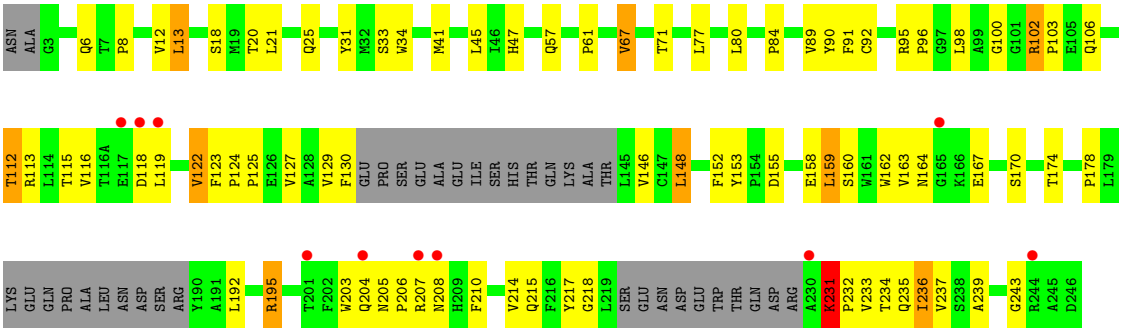


• Molecule 3: TAX PEPTIDE

• Molecule 4: T CELL RECEPTOR ALPHA



• Molecule 5: T CELL RECEPTOR BETA



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	229.30Å 49.50Å 96.00Å 90.00° 89.60° 90.00°	Depositor
Resolution (Å)	6.00 – 2.60 11.95 – 2.34	Depositor EDS
% Data completeness (in resolution range)	94.9 (6.00-2.60) 85.9 (11.95-2.34)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	0.11	Depositor
$< I/\sigma(I) >$ [1]	1.81 (at 2.35Å)	Xtriage
Refinement program	X-PLOR 3.1	Depositor
R, R_{free}	0.245 , 0.320 0.232 , 0.298	Depositor DCC
R_{free} test set	3927 reflections (10.00%)	wwPDB-VP
Wilson B-factor (Å^2)	39.2	Xtriage
Anisotropy	0.189	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 77.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.022 for -h,-k,l	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	5711	wwPDB-VP
Average B, all atoms (Å^2)	43.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.71% 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: EMC

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	0.38	0/2302	0.66	0/3125
2	B	0.40	0/850	0.75	1/1149 (0.1%)
3	C	0.51	0/80	0.82	0/108
4	D	0.41	0/905	0.74	1/1229 (0.1%)
5	E	0.39	0/1685	0.76	2/2292 (0.1%)
All	All	0.39	0/5822	0.72	4/7903 (0.1%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	64	LEU	CA-CB-CG	6.51	130.26	115.30
5	E	231	LYS	N-CA-C	6.18	127.69	111.00
5	E	13	LEU	CA-CB-CG	5.61	128.20	115.30
4	D	120	ASN	N-CA-C	5.51	125.87	111.00

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

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	2237	0	2090	45	0
2	B	828	0	789	15	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	C	77	0	79	4	0
4	D	886	0	845	27	0
5	E	1640	0	1570	59	0
6	B	6	0	0	2	0
7	A	13	0	0	0	0
7	B	13	0	0	0	0
7	C	1	0	0	0	0
7	D	2	0	0	0	0
7	E	8	0	0	0	0
All	All	5711	0	5373	142	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 13.

All (142) 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:A:21:ARG:NE	1:A:23:ILE:HD11	1.91	0.86
1:A:187:THR:HG21	1:A:261:VAL:HG21	1.57	0.84
4:D:93:THR:HG21	4:D:104:LEU:HD22	1.60	0.83
1:A:32:GLN:HE21	1:A:48:ARG:HG3	1.44	0.81
5:E:21:LEU:HD12	5:E:77:LEU:HD23	1.62	0.80
1:A:69:ALA:O	1:A:73:THR:HG22	1.83	0.77
5:E:163:VAL:HG12	5:E:210:PHE:HD1	1.50	0.75
1:A:127:LYS:HE2	1:A:134:THR:OG1	1.87	0.74
5:E:146:VAL:HG12	5:E:195:ARG:HD3	1.74	0.70
4:D:61:ARG:HD2	4:D:78:ARG:O	1.91	0.70
4:D:93:THR:HG21	4:D:104:LEU:CD2	2.21	0.69
4:D:52:ASN:HB3	4:D:68:LYS:HD2	1.75	0.68
1:A:32:GLN:NE2	1:A:48:ARG:HG3	2.09	0.67
5:E:218:GLY:H	5:E:234:THR:HG22	1.59	0.67
5:E:163:VAL:HG12	5:E:210:PHE:CD1	2.30	0.67
4:D:93:THR:HG22	4:D:104:LEU:HD13	1.76	0.66
5:E:130:PHE:HB2	5:E:146:VAL:HG23	1.78	0.66
1:A:111:ARG:HD2	1:A:113:TYR:OH	1.96	0.66
5:E:233:VAL:O	5:E:235:GLN:HG3	1.95	0.65
5:E:236:ILE:HD13	5:E:236:ILE:N	2.10	0.65
5:E:159:LEU:HD12	5:E:160:SER:N	2.11	0.65
1:A:213:ILE:HG13	1:A:262:GLN:O	1.97	0.64
1:A:150:ALA:HA	5:E:102:ARG:NH1	2.13	0.64
5:E:95:ARG:HG2	5:E:106:GLN:HB2	1.81	0.62

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
5:E:18:SER:HB3	5:E:80:LEU:O	1.99	0.62
5:E:100:GLY:HA3	5:E:102:ARG:HE	1.65	0.61
4:D:91:ALA:HB1	4:D:104:LEU:HD12	1.81	0.61
5:E:158:GLU:HB2	5:E:215:GLN:HB3	1.81	0.61
4:D:5:GLN:NE2	4:D:90:CYS:H	1.98	0.60
4:D:90:CYS:O	4:D:106:PHE:HA	2.02	0.60
5:E:217:TYR:HA	5:E:234:THR:HG22	1.83	0.60
5:E:122:VAL:HA	5:E:153:TYR:O	2.00	0.60
5:E:89:VAL:HG22	5:E:113:ARG:HG3	1.84	0.60
5:E:8:PRO:O	5:E:112:THR:HB	2.02	0.59
4:D:27:ARG:HH11	4:D:27:ARG:HG3	1.67	0.59
5:E:206:PRO:HA	5:E:243:GLY:O	2.03	0.59
4:D:93:THR:CG2	4:D:104:LEU:HD22	2.33	0.58
4:D:17:ILE:HD13	4:D:78:ARG:HA	1.85	0.58
4:D:38:TYR:HB2	4:D:41:LYS:HD2	1.86	0.58
1:A:93:HIS:HD2	1:A:119:ASP:OD2	1.86	0.58
1:A:4:SER:O	1:A:28:VAL:O	2.22	0.58
1:A:167:TRP:NE1	3:C:1:LEU:HD22	2.19	0.57
4:D:82:PRO:HA	4:D:114:VAL:HB	1.86	0.57
4:D:61:ARG:HD3	4:D:79:ASP:HB3	1.85	0.57
5:E:203:TRP:HE3	5:E:210:PHE:CE2	2.22	0.57
2:B:17:ASN:ND2	2:B:74:GLU:HG2	2.20	0.56
1:A:202:ARG:HD2	1:A:244:TRP:CD2	2.40	0.56
1:A:163:THR:HG22	1:A:164:CYS:N	2.21	0.55
1:A:162:GLY:O	1:A:165:VAL:HG22	2.07	0.55
5:E:129:VAL:HG23	5:E:239:ALA:HB3	1.89	0.55
1:A:6:ARG:NH2	1:A:113:TYR:CE2	2.75	0.55
1:A:6:ARG:NH2	1:A:113:TYR:CD2	2.75	0.55
2:B:37:VAL:HG22	2:B:82:VAL:HG22	1.89	0.54
5:E:203:TRP:HE3	5:E:210:PHE:HE2	1.55	0.54
4:D:91:ALA:HA	4:D:105:GLN:O	2.08	0.54
5:E:125:PRO:HB3	5:E:152:PHE:HB3	1.89	0.54
5:E:129:VAL:HG23	5:E:239:ALA:CB	2.38	0.54
5:E:102:ARG:N	5:E:102:ARG:HD3	2.23	0.53
2:B:21:ASN:HB3	2:B:70:PHE:CE1	2.43	0.53
2:B:52:SER:N	6:B:100:EMC:C2	2.72	0.53
4:D:5:GLN:HE21	4:D:107:GLY:HA3	1.74	0.52
1:A:73:THR:HG21	3:C:6:PRO:HB2	1.91	0.52
1:A:204:TRP:CH2	1:A:244:TRP:CD1	2.98	0.52
5:E:203:TRP:CE3	5:E:210:PHE:HE2	2.26	0.52
1:A:21:ARG:CZ	1:A:23:ILE:HD11	2.40	0.52

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:D:42:SER:HA	5:E:91:PHE:CE1	2.45	0.52
5:E:102:ARG:H	5:E:102:ARG:CD	2.23	0.52
5:E:162:TRP:HA	5:E:167:GLU:HA	1.92	0.52
5:E:127:VAL:HG21	5:E:237:VAL:O	2.11	0.51
5:E:6:GLN:HE22	5:E:92:CYS:H	1.59	0.51
1:A:117:ALA:HB2	2:B:60:TRP:CE2	2.46	0.50
1:A:203:CYS:O	1:A:244:TRP:HB2	2.11	0.50
1:A:219:ARG:HB3	1:A:222:GLU:HB2	1.93	0.50
5:E:100:GLY:HA3	5:E:102:ARG:HH21	1.75	0.50
1:A:187:THR:HA	1:A:204:TRP:O	2.11	0.50
1:A:184:ALA:HB2	1:A:265:GLY:O	2.12	0.50
5:E:13:LEU:O	5:E:116:VAL:HA	2.12	0.50
5:E:204:GLN:O	5:E:206:PRO:HD3	2.12	0.49
5:E:203:TRP:CE3	5:E:210:PHE:CE2	3.00	0.49
1:A:202:ARG:HD2	1:A:244:TRP:CE3	2.48	0.49
5:E:233:VAL:HG12	5:E:234:THR:N	2.27	0.49
5:E:236:ILE:HD13	5:E:236:ILE:H	1.75	0.48
5:E:123:PHE:O	5:E:152:PHE:HA	2.14	0.48
1:A:258:THR:HG22	1:A:273:ARG:HA	1.95	0.48
5:E:159:LEU:HB2	5:E:214:VAL:HG22	1.96	0.48
1:A:230:LEU:HG	1:A:245:ALA:HB2	1.96	0.47
2:B:51:HIS:HB3	2:B:66:TYR:CE2	2.49	0.47
4:D:34:TRP:HB2	4:D:47:MET:HB2	1.96	0.47
2:B:65:LEU:HD12	2:B:66:TYR:H	1.80	0.47
1:A:181:ARG:HG2	1:A:182:THR:N	2.28	0.47
1:A:51:TRP:CZ2	1:A:179:LEU:HD11	2.50	0.46
5:E:146:VAL:HG12	5:E:195:ARG:CD	2.44	0.46
1:A:108:ARG:HH11	1:A:108:ARG:HG2	1.81	0.46
2:B:19:LYS:O	2:B:72:PRO:HD2	2.15	0.46
5:E:34:TRP:O	5:E:45:LEU:HD12	2.17	0.45
5:E:6:GLN:OE1	5:E:112:THR:HG22	2.17	0.45
4:D:5:GLN:HE22	4:D:89:LEU:HA	1.82	0.45
2:B:51:HIS:C	6:B:100:EMC:C2	2.85	0.45
4:D:5:GLN:HG3	4:D:107:GLY:HA3	1.99	0.45
2:B:19:LYS:HE3	2:B:19:LYS:HB3	1.66	0.45
1:A:70:HIS:HA	1:A:73:THR:CG2	2.47	0.44
4:D:22:CYS:HB3	4:D:73:VAL:HB	2.00	0.44
5:E:218:GLY:H	5:E:234:THR:CG2	2.29	0.44
1:A:28:VAL:HG11	1:A:179:LEU:HD13	1.98	0.44
1:A:11:SER:OG	1:A:78:LEU:HD11	2.17	0.44
5:E:100:GLY:CA	5:E:102:ARG:HH21	2.30	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:259:CYS:O	1:A:271:THR:HA	2.18	0.43
4:D:13:PRO:HA	4:D:115:THR:HG23	2.00	0.43
5:E:127:VAL:HA	5:E:148:LEU:O	2.18	0.43
1:A:97:ARG:HG2	1:A:98:MET:N	2.34	0.43
2:B:23:LEU:O	2:B:67:CYS:HA	2.18	0.43
1:A:11:SER:HB3	1:A:95:VAL:HB	2.00	0.43
2:B:9:VAL:HG12	2:B:93:VAL:HG12	2.00	0.43
4:D:4:GLU:HB2	4:D:23:THR:HG22	2.00	0.43
5:E:178:PRO:HA	5:E:192:LEU:HD13	2.00	0.43
1:A:5:MET:O	1:A:100:GLY:HA3	2.19	0.42
1:A:73:THR:HA	3:C:8:TYR:CE1	2.54	0.42
2:B:9:VAL:HA	2:B:24:ASN:O	2.19	0.42
5:E:57:GLN:HB2	5:E:61:PRO:HG3	2.01	0.42
1:A:234:ARG:HE	1:A:242:GLN:HE21	1.66	0.42
4:D:93:THR:CG2	4:D:104:LEU:HD13	2.46	0.42
5:E:124:PRO:HD3	5:E:232:PRO:HB3	2.01	0.42
1:A:187:THR:HB	1:A:272:LEU:HD11	2.01	0.42
4:D:52:ASN:HA	4:D:66:LEU:HB3	2.00	0.42
5:E:84:PRO:HA	5:E:116:VAL:HB	2.01	0.42
5:E:205:ASN:HB3	5:E:208:ASN:ND2	2.34	0.42
5:E:47:HIS:CD2	5:E:67:VAL:HB	2.55	0.42
5:E:96:PRO:HG2	5:E:102:ARG:HG2	2.02	0.42
4:D:33:PHE:HZ	5:E:103:PRO:HB2	1.84	0.41
1:A:72:GLN:O	1:A:76:VAL:HG23	2.20	0.41
1:A:152:VAL:HG22	3:C:7:VAL:HG21	2.02	0.41
4:D:52:ASN:HB3	4:D:68:LYS:CD	2.46	0.41
5:E:67:VAL:HG23	5:E:77:LEU:HA	2.01	0.41
1:A:144:LYS:O	1:A:148:GLU:HG3	2.20	0.41
1:A:233:THR:OG1	1:A:243:LYS:HD2	2.20	0.41
2:B:34:ASP:OD1	2:B:34:ASP:N	2.53	0.41
5:E:102:ARG:HD3	5:E:102:ARG:H	1.81	0.41
5:E:90:TYR:HB2	5:E:112:THR:HG23	2.03	0.40
5:E:102:ARG:N	5:E:102:ARG:CD	2.83	0.40
5:E:159:LEU:HD12	5:E:159:LEU:C	2.41	0.40
2:B:51:HIS:HA	2:B:65:LEU:O	2.22	0.40
5:E:205:ASN:HB3	5:E:208:ASN:HD21	1.86	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	272/275 (99%)	260 (96%)	12 (4%)	0	100 100
2	B	98/100 (98%)	95 (97%)	3 (3%)	0	100 100
3	C	7/9 (78%)	7 (100%)	0	0	100 100
4	D	113/204 (55%)	105 (93%)	7 (6%)	1 (1%)	17 35
5	E	201/245 (82%)	182 (90%)	14 (7%)	5 (2%)	5 9
All	All	691/833 (83%)	649 (94%)	36 (5%)	6 (1%)	17 35

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	D	6	ASN
5	E	119	LEU
5	E	231	LYS
5	E	164	ASN
5	E	155	ASP
5	E	122	VAL

5.3.2 Protein sidechains (i)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	230/231 (100%)	208 (90%)	22 (10%)	8 16
2	B	95/95 (100%)	89 (94%)	6 (6%)	18 36
3	C	8/8 (100%)	8 (100%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
4	D	99/182 (54%)	90 (91%)	9 (9%)	9 18
5	E	177/209 (85%)	156 (88%)	21 (12%)	5 9
All	All	609/725 (84%)	551 (90%)	58 (10%)	8 16

All (58) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	6	ARG
1	A	23	ILE
1	A	34	VAL
1	A	35	ARG
1	A	45	MET
1	A	54	GLN
1	A	67	VAL
1	A	73	THR
1	A	75	ARG
1	A	98	MET
1	A	105	SER
1	A	128	GLU
1	A	155	GLN
1	A	163	THR
1	A	177	GLU
1	A	187	THR
1	A	190	THR
1	A	207	SER
1	A	226	GLN
1	A	247	VAL
1	A	255	GLN
1	A	273	ARG
2	B	34	ASP
2	B	38	ASP
2	B	67	CYS
2	B	70	PHE
2	B	83	ASN
2	B	89	GLN
4	D	6	ASN
4	D	23	THR
4	D	31	SER
4	D	49	ILE
4	D	68	LYS
4	D	76	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
4	D	87	THR
4	D	93	THR
4	D	104	LEU
5	E	12	VAL
5	E	20	THR
5	E	25	GLN
5	E	31	TYR
5	E	33	SER
5	E	41	MET
5	E	67	VAL
5	E	71	THR
5	E	98	LEU
5	E	102	ARG
5	E	112	THR
5	E	115	THR
5	E	118	ASP
5	E	148	LEU
5	E	159	LEU
5	E	170	SER
5	E	174	THR
5	E	195	ARG
5	E	207	ARG
5	E	231	LYS
5	E	236	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (16) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	32	GLN
1	A	74	HIS
1	A	86	ASN
1	A	93	HIS
1	A	141	GLN
1	A	242	GLN
2	B	31	HIS
2	B	83	ASN
4	D	5	GLN
4	D	37	GLN
4	D	52	ASN
5	E	6	GLN
5	E	11	GLN
5	E	37	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
5	E	208	ASN
5	E	235	GLN

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)

2 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	EMC	B	101	2	1,2,2	0.70	0	-		
6	EMC	B	100	2	1,2,2	1.21	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.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	B	100	EMC	2	0

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{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	274/275 (99%)	-0.14	22 (8%) 12 9	6, 32, 119, 136	30 (10%)
2	B	100/100 (100%)	-0.67	1 (1%) 82 80	6, 25, 57, 81	2 (2%)
3	C	9/9 (100%)	-1.24	0 100 100	13, 19, 23, 24	0
4	D	115/204 (56%)	-0.45	2 (1%) 70 66	11, 33, 65, 76	7 (6%)
5	E	209/245 (85%)	-0.01	10 (4%) 30 24	16, 48, 90, 101	10 (4%)
All	All	707/833 (84%)	-0.24	35 (4%) 28 23	6, 35, 98, 136	49 (6%)

All (35) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	196	ASP	7.8
5	E	118	ASP	7.7
1	A	1	GLY	6.4
1	A	249	VAL	5.7
1	A	221	GLY	5.5
1	A	195	SER	5.0
1	A	223	ASP	5.0
1	A	220	ASP	5.0
1	A	219	ARG	4.5
1	A	250	PRO	4.0
4	D	121	PRO	3.7
1	A	253	GLN	3.5
5	E	119	LEU	3.2
2	B	99	MET	3.1
5	E	201	THR	3.0
1	A	274	TRP	3.0
1	A	252	GLY	3.0
5	E	230	ALA	2.8
1	A	225	THR	2.8
5	E	117	GLU	2.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	227	ASP	2.6
1	A	254	GLU	2.6
1	A	259	CYS	2.6
1	A	222	GLU	2.4
5	E	204	GLN	2.4
4	D	1	LYS	2.4
1	A	198	GLU	2.4
5	E	207	ARG	2.3
1	A	224	GLN	2.3
1	A	218	GLN	2.2
5	E	165	GLY	2.2
1	A	267	PRO	2.2
5	E	244	ARG	2.1
1	A	229	GLU	2.0
5	E	208	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, 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	B-factors(Å^2)	Q<0.9
6	EMC	B	100	3/3	0.98	0.07	44,44,57,69	0
6	EMC	B	101	3/3	0.99	0.05	50,50,51,53	3

6.5 Other polymers (i)

There are no such residues in this entry.