



Full wwPDB X-ray Structure Validation Report ⓘ

Aug 22, 2017 – 03:57 PM EDT

PDB ID : 4U1K
Title : HLA class I micropolymorphisms determine peptide-HLA landscape and dictate differential HIV-1 escape through identical epitopes
Authors : Rizkallah, P.J.; Cole, D.K.; Fuller, A.; Sewell, A.K.
Deposited on : unknown
Resolution : 2.09 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 ⓘ symbol.

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

MolProbity : 4.02b-467
Mogul : 1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20029824
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
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-20029824

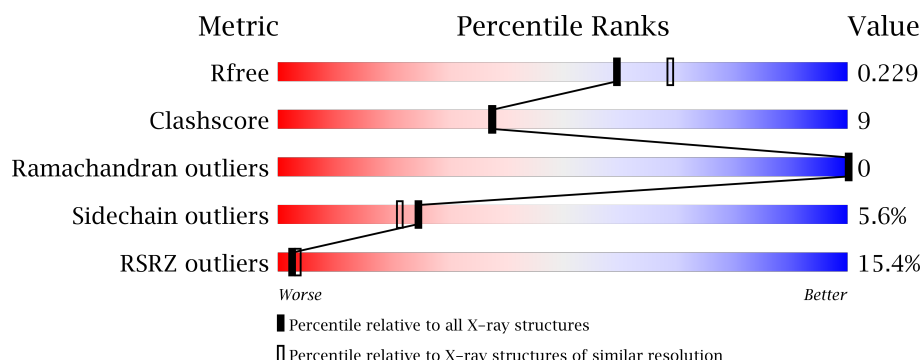
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.09 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}	100719	4243 (2.10-2.10)
Clashscore	112137	4788 (2.10-2.10)
Ramachandran outliers	110173	4740 (2.10-2.10)
Sidechain outliers	110143	4741 (2.10-2.10)
RSRZ outliers	101464	4275 (2.10-2.10)

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 $\leq 5\%$. 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	277	<div> <div>16%</div> <div>79%</div> <div>19%</div> </div>
1	D	277	<div> <div>17%</div> <div>82%</div> <div>17%</div> </div>
2	B	100	<div> <div>18%</div> <div>83%</div> <div>14%</div> </div>
2	E	100	<div> <div>12%</div> <div>82%</div> <div>17%</div> </div>
3	C	9	<div> <div>56%</div> <div>44%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	F	9	 <div>89%11%</div>

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
4	GOL	A	301	-	-	-	X
4	GOL	A	302	-	X	X	X
4	GOL	A	303	-	-	-	X
4	GOL	D	303	-	-	X	X
4	GOL	D	305	-	-	X	X
6	SO4	A	307	-	-	-	X

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 6989 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, B-7 alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	277	Total	C	N	O	S	0	2	0
			2287	1418	417	445	7			
1	D	277	Total	C	N	O	S	0	4	0
			2302	1426	420	449	7			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	initiating methionine	UNP P01889
D	0	MET	-	initiating methionine	UNP P01889

- 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			
2	E	100	Total	C	N	O	S	0	0	0
			837	533	141	159	4			

There are 2 discrepancies between the modelled and reference sequences:

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

- Molecule 3 is a protein called Protein Nef.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	9	Total	C	N	O	S	0	0	0
			76	48	16	11	1			
3	F	9	Total	C	N	O	S	0	0	0
			76	48	16	11	1			

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



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

- Molecule 5 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			4	2	2		
5	A	1	Total	C	O	0	0
			4	2	2		
5	A	1	Total	C	O	0	0
			4	2	2		
5	D	1	Total	C	O	0	0
			4	2	2		
5	D	1	Total	C	O	0	0
			4	2	2		
5	D	1	Total	C	O	0	0
			4	2	2		

- Molecule 6 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



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

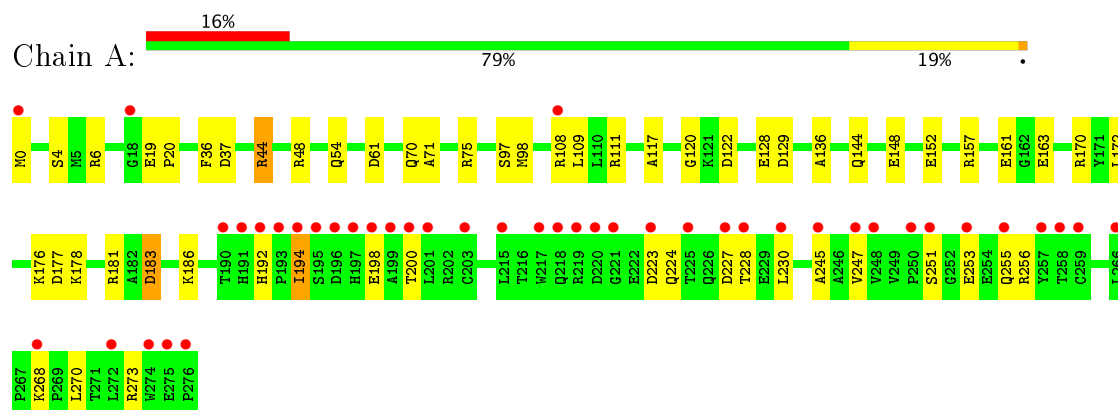
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	206	Total	O	0	0
			206	206		
7	B	31	Total	O	0	0
			31	31		
7	C	9	Total	O	0	0
			9	9		
7	D	205	Total	O	0	0
			205	205		
7	E	32	Total	O	0	0
			32	32		
7	F	14	Total	O	0	0
			14	14		

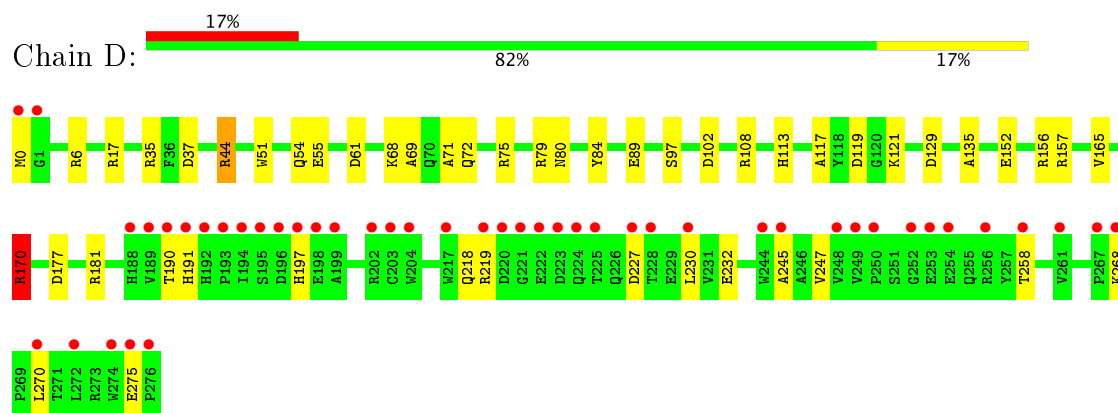
3 Residue-property plots [i](#)

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

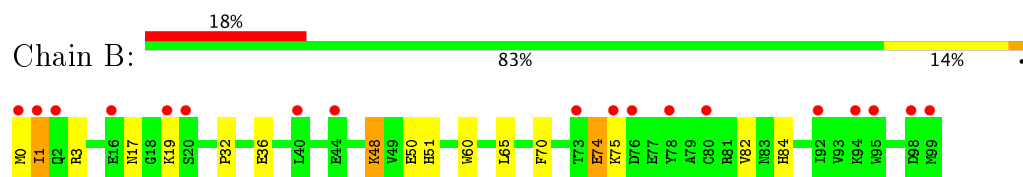
- Molecule 1: HLA class I histocompatibility antigen, B-7 alpha chain



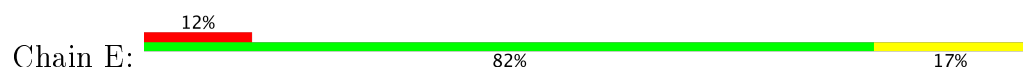
- Molecule 1: HLA class I histocompatibility antigen, B-7 alpha chain

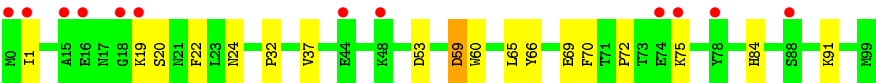


- Molecule 2: Beta-2-microglobulin

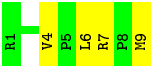


- Molecule 2: Beta-2-microglobulin

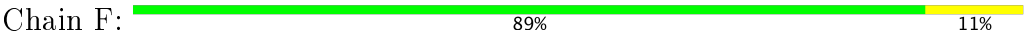




● Molecule 3: Protein Nef



● Molecule 3: Protein Nef



4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	46.27Å 49.13Å 107.46Å 91.37° 93.57° 95.75°	Depositor
Resolution (Å)	48.86 – 2.09 48.86 – 2.09	Depositor EDS
% Data completeness (in resolution range)	97.5 (48.86-2.09) 96.5 (48.86-2.09)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.11 (at 2.08Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, R_{free}	0.176 , 0.226 0.183 , 0.229	Depositor DCC
R_{free} test set	2757 reflections (5.36%)	DCC
Wilson B-factor (Å ²)	28.0	Xtriage
Anisotropy	0.508	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 50.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	6989	wwPDB-VP
Average B, all atoms (Å ²)	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.33% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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	0.94	0/2350	0.97	8/3192 (0.3%)
1	D	0.95	2/2365 (0.1%)	0.93	5/3212 (0.2%)
2	B	0.62	0/860	0.77	0/1162
2	E	0.68	0/860	0.79	1/1162 (0.1%)
3	C	1.21	0/78	1.25	1/104 (1.0%)
3	F	1.10	0/78	1.31	1/104 (1.0%)
All	All	0.88	2/6591 (0.0%)	0.92	16/8936 (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
2	B	0	1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	97	SER	CB-OG	5.75	1.49	1.42
1	D	165	VAL	CB-CG1	-5.04	1.42	1.52

All (16) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	44	ARG	NE-CZ-NH2	7.39	123.99	120.30
1	D	170	ARG	NE-CZ-NH2	-6.89	116.85	120.30
1	A	122	ASP	CB-CG-OD1	6.88	124.50	118.30
1	D	157	ARG	NE-CZ-NH2	6.87	123.73	120.30
2	E	59	ASP	CB-CG-OD1	6.40	124.06	118.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	129	ASP	CB-CG-OD1	6.23	123.90	118.30
1	A	98	MET	CG-SD-CE	-5.77	90.97	100.20
3	F	7	ARG	NE-CZ-NH2	-5.75	117.43	120.30
3	C	9	MET	CA-CB-CG	-5.54	103.89	113.30
1	A	157	ARG	NE-CZ-NH2	5.51	123.05	120.30
1	D	37	ASP	CB-CG-OD1	5.24	123.02	118.30
1	A	48	ARG	NE-CZ-NH1	5.18	122.89	120.30
1	D	6	ARG	NE-CZ-NH1	5.09	122.84	120.30
1	A	98	MET	CA-CB-CG	5.04	121.88	113.30
1	A	170	ARG	NE-CZ-NH1	-5.03	117.78	120.30
1	D	129	ASP	CB-CG-OD1	5.03	122.82	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	0	MET	Peptide

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	2287	0	2122	46	0
1	D	2302	0	2133	37	0
2	B	837	0	803	10	0
2	E	837	0	803	8	0
3	C	76	0	86	11	0
3	F	76	0	86	0	0
4	A	18	0	23	16	0
4	D	30	0	39	18	0
5	A	12	0	18	3	0
5	D	12	0	18	1	0
6	A	5	0	0	0	0
7	A	206	0	0	22	0
7	B	31	0	0	1	0
7	C	9	0	0	0	0
7	D	205	0	0	8	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	E	32	0	0	0	0
7	F	14	0	0	0	0
All	All	6989	0	6131	111	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 9.

All (111) 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:D:71:ALA:HB1	4:D:305:GOL:H12	1.40	1.01
4:A:302:GOL:H12	3:C:4:VAL:HG13	1.41	0.99
4:A:302:GOL:C1	3:C:4:VAL:HG13	1.92	0.98
1:A:109:LEU:HD22	1:A:161[B]:GLU:HB3	1.62	0.81
4:A:302:GOL:H11	3:C:4:VAL:N	1.97	0.79
1:D:71:ALA:HB1	4:D:305:GOL:C1	2.15	0.75
1:A:70:GLN:HE22	4:A:301:GOL:H12	1.52	0.74
1:A:163:GLU:OE2	4:A:302:GOL:H2	1.88	0.74
1:A:6:ARG:HG2	7:A:406:HOH:O	1.87	0.74
1:D:75:ARG:CZ	4:D:305:GOL:H11	2.18	0.73
4:A:302:GOL:H11	3:C:4:VAL:H	1.53	0.73
1:A:71:ALA:O	1:A:75:ARG:HD3	1.90	0.71
1:D:170:ARG:NH1	7:D:401:HOH:O	2.23	0.68
1:A:20:PRO:HD2	1:A:75:ARG:HG2	1.76	0.67
1:D:72:GLN:CD	7:D:403:HOH:O	2.31	0.67
1:D:75:ARG:NE	4:D:305:GOL:H11	2.10	0.66
1:A:36:PHE:CD1	7:A:411:HOH:O	2.48	0.65
1:A:163:GLU:OE2	4:A:302:GOL:C2	2.45	0.64
4:A:302:GOL:H11	3:C:4:VAL:HG13	1.77	0.63
1:D:230:LEU:HD23	1:D:245:ALA:HB2	1.80	0.62
5:A:306:EDO:H11	7:A:456:HOH:O	2.00	0.62
1:D:75:ARG:NH2	4:D:305:GOL:H11	2.16	0.60
2:E:22:PHE:CE2	2:E:69:GLU:HG3	2.38	0.59
2:B:48:LYS:HG3	2:B:48:LYS:O	2.02	0.59
1:D:69:ALA:HA	4:D:304:GOL:H2	1.84	0.59
1:D:135:ALA:HB3	5:D:308:EDO:H21	1.85	0.58
1:D:80:ASN:HB3	4:D:303:GOL:H32	1.86	0.58
1:A:4:SER:HB2	7:A:406:HOH:O	2.03	0.58
7:A:566:HOH:O	2:B:1:ILE:HG21	2.04	0.57
4:A:302:GOL:H12	3:C:4:VAL:CG1	2.26	0.57
4:A:302:GOL:C1	3:C:4:VAL:H	2.18	0.56

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:36:PHE:CE1	7:A:411:HOH:O	2.58	0.56
1:A:172:LEU:O	1:A:176:LYS:HB3	2.06	0.56
1:A:136:ALA:HB2	5:A:305:EDO:H11	1.88	0.56
1:A:4:SER:CB	7:A:406:HOH:O	2.53	0.56
4:A:302:GOL:H32	3:C:4:VAL:HG13	1.89	0.55
5:A:306:EDO:C1	7:A:456:HOH:O	2.54	0.55
1:A:36:PHE:CG	7:A:411:HOH:O	2.60	0.55
2:B:48:LYS:CG	2:B:48:LYS:O	2.54	0.55
4:A:303:GOL:H2	7:D:408:HOH:O	2.05	0.55
1:A:6:ARG:NH1	7:A:406:HOH:O	2.38	0.54
1:D:69:ALA:CA	4:D:304:GOL:H2	2.36	0.54
1:A:37:ASP:N	7:A:411:HOH:O	2.40	0.54
1:D:80:ASN:HA	4:D:303:GOL:H31	1.90	0.54
1:D:89:GLU:CD	7:D:478:HOH:O	2.45	0.54
1:A:71:ALA:O	1:A:75:ARG:CD	2.56	0.54
1:A:97:SER:HB2	7:A:442:HOH:O	2.09	0.53
1:A:54:GLN:NE2	7:A:413:HOH:O	2.41	0.53
4:A:302:GOL:H32	3:C:4:VAL:CG1	2.38	0.53
1:D:68:LYS:NZ	7:D:402:HOH:O	2.25	0.53
1:A:61[B]:ASP:OD1	4:A:303:GOL:H32	2.08	0.53
1:A:253:GLU:OE2	1:A:256:ARG:NH1	2.42	0.53
2:E:32:PRO:O	2:E:84:HIS:HE1	1.92	0.53
1:A:75:ARG:NH2	7:A:412:HOH:O	2.41	0.52
1:D:170:ARG:HH11	1:D:170:ARG:HG3	1.76	0.51
1:A:44:ARG:NH2	7:A:414:HOH:O	2.43	0.51
1:D:51:TRP:O	1:D:54:GLN:HG2	2.11	0.50
4:A:303:GOL:O2	7:A:401:HOH:O	2.19	0.50
2:B:50:GLU:HG3	7:B:125:HOH:O	2.11	0.50
1:A:227:ASP:O	1:A:247:VAL:HG23	2.10	0.50
1:A:192:HIS:HB2	1:A:200:THR:HB	1.93	0.50
1:D:55:GLU:OE2	1:D:170:ARG:NE	2.45	0.49
2:B:36:GLU:O	2:B:82:VAL:HA	2.12	0.49
1:D:108:ARG:HB2	1:D:108:ARG:CZ	2.43	0.49
2:B:32:PRO:O	2:B:84:HIS:HE1	1.97	0.48
1:A:255:GLN:O	1:A:273:ARG:NH1	2.47	0.48
1:A:152:GLU:OE1	3:C:6:LEU:HB3	2.14	0.48
1:A:111:ARG:HD2	1:A:128:GLU:HG3	1.96	0.47
1:A:36:PHE:CZ	7:A:411:HOH:O	2.68	0.47
1:A:111:ARG:NH1	7:A:402:HOH:O	2.23	0.47
2:E:24:ASN:HB3	2:E:65:LEU:HD11	1.97	0.47
4:D:305:GOL:H31	7:D:421:HOH:O	2.14	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:80:ASN:OD1	4:D:303:GOL:H31	2.15	0.46
1:D:71:ALA:CB	4:D:305:GOL:H12	2.29	0.46
1:A:117:ALA:HB2	2:B:60:TRP:CE2	2.50	0.46
1:A:161[B]:GLU:C	1:A:161[B]:GLU:OE1	2.53	0.46
1:D:119:ASP:OD2	4:D:302:GOL:O1	2.28	0.45
1:A:37:ASP:C	7:A:411:HOH:O	2.56	0.45
1:A:44:ARG:NH1	1:D:75:ARG:NH1	2.65	0.45
1:D:152[A]:GLU:O	1:D:156:ARG:HG2	2.18	0.44
1:D:177:ASP:O	1:D:181:ARG:NH1	2.50	0.44
4:D:305:GOL:H32	7:D:462:HOH:O	2.18	0.44
1:D:102:ASP:OD2	1:D:113:HIS:HE1	2.02	0.43
1:A:181:ARG:NH2	1:A:183:ASP:OD2	2.52	0.43
1:A:70:GLN:HE22	4:A:301:GOL:C1	2.27	0.43
1:D:80:ASN:HA	4:D:303:GOL:C3	2.49	0.43
1:D:84:TYR:CE1	4:D:303:GOL:H32	2.53	0.43
2:E:37:VAL:HG21	2:E:66:TYR:CE1	2.52	0.43
1:A:152:GLU:OE2	3:C:7:ARG:NH1	2.46	0.43
1:D:80:ASN:CB	4:D:303:GOL:H32	2.47	0.43
1:D:69:ALA:N	4:D:304:GOL:H2	2.33	0.43
1:A:224:GLN:O	1:A:228:THR:OG1	2.32	0.42
7:A:414:HOH:O	1:D:79:ARG:NH1	2.52	0.42
1:A:194:ILE:HG12	1:A:198:GLU:HB2	2.01	0.42
1:D:230:LEU:CD2	1:D:245:ALA:HB2	2.47	0.42
1:A:108:ARG:NE	7:A:425:HOH:O	2.52	0.42
1:D:35:ARG:HD3	2:E:53:ASP:OD2	2.20	0.42
2:B:17:ASN:HD21	2:B:74:GLU:HA	1.86	0.41
1:D:170:ARG:NH1	1:D:170:ARG:HG3	2.34	0.41
2:B:51:HIS:HA	2:B:65:LEU:O	2.19	0.41
1:A:19:GLU:CB	1:A:75:ARG:NH2	2.84	0.41
1:A:144:GLN:O	1:A:148:GLU:HG3	2.21	0.41
1:A:230:LEU:HD23	1:A:245:ALA:HB2	2.03	0.41
1:A:36:PHE:CD2	7:A:411:HOH:O	2.72	0.41
1:D:113:HIS:HD2	7:D:553:HOH:O	2.03	0.41
2:E:59:ASP:O	2:E:60:TRP:HB2	2.20	0.41
1:A:120:GLY:O	2:B:3:ARG:NH1	2.51	0.41
1:A:108:ARG:HB3	1:A:108:ARG:HH21	1.86	0.40
1:D:44:ARG:NH2	1:D:61:ASP:OD1	2.46	0.40
2:E:19:LYS:O	2:E:72:PRO:HD2	2.22	0.40
1:D:117:ALA:HB2	2:E:60:TRP:CE2	2.56	0.40

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

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	277/277 (100%)	266 (96%)	11 (4%)	0	100	100
1	D	279/277 (101%)	270 (97%)	9 (3%)	0	100	100
2	B	98/100 (98%)	96 (98%)	2 (2%)	0	100	100
2	E	98/100 (98%)	97 (99%)	1 (1%)	0	100	100
3	C	7/9 (78%)	6 (86%)	1 (14%)	0	100	100
3	F	7/9 (78%)	6 (86%)	1 (14%)	0	100	100
All	All	766/772 (99%)	741 (97%)	25 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	237/235 (101%)	227 (96%)	10 (4%)	34	33
1	D	239/235 (102%)	222 (93%)	17 (7%)	17	13
2	B	95/95 (100%)	89 (94%)	6 (6%)	21	17
2	E	95/95 (100%)	90 (95%)	5 (5%)	26	24
3	C	9/9 (100%)	9 (100%)	0	100	100
3	F	9/9 (100%)	9 (100%)	0	100	100
All	All	684/678 (101%)	646 (94%)	38 (6%)	25	21

All (38) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	0	MET
1	A	177	ASP
1	A	178	LYS
1	A	183	ASP
1	A	186	LYS
1	A	194	ILE
1	A	223	ASP
1	A	251	SER
1	A	268	LYS
1	A	270	LEU
2	B	1	ILE
2	B	19	LYS
2	B	48	LYS
2	B	70	PHE
2	B	74	GLU
2	B	75	LYS
1	D	0	MET
1	D	17	ARG
1	D	44	ARG
1	D	121	LYS
1	D	170	ARG
1	D	190	THR
1	D	191	HIS
1	D	197	HIS
1	D	218	GLN
1	D	219	ARG
1	D	227	ASP
1	D	232	GLU
1	D	247	VAL
1	D	258	THR
1	D	268	LYS
1	D	270	LEU
1	D	275	GLU
2	E	1	ILE
2	E	20	SER
2	E	70	PHE
2	E	75	LYS
2	E	91	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (10) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	70	GLN
1	A	113	HIS
1	A	188	HIS
1	A	260	HIS
2	B	84	HIS
1	D	113	HIS
1	D	155	GLN
1	D	191	HIS
1	D	218	GLN
2	E	84	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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

15 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	$\# Z > 2$
4	GOL	A	301	-	5,5,5	1.18	0	5,5,5	1.63	1 (20%)
4	GOL	A	302	-	5,5,5	1.89	3 (60%)	5,5,5	3.65	4 (80%)
4	GOL	A	303	-	5,5,5	1.06	0	5,5,5	1.62	1 (20%)
5	EDO	A	304	-	3,3,3	0.56	0	2,2,2	0.37	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
5	EDO	A	305	-	3,3,3	0.47	0	2,2,2	0.26	0
5	EDO	A	306	-	3,3,3	0.48	0	2,2,2	0.18	0
6	SO4	A	307	-	4,4,4	0.53	0	6,6,6	0.34	0
4	GOL	D	301	-	5,5,5	0.97	0	5,5,5	1.26	1 (20%)
4	GOL	D	302	-	5,5,5	0.56	0	5,5,5	0.68	0
4	GOL	D	303	-	5,5,5	1.92	1 (20%)	5,5,5	2.28	2 (40%)
4	GOL	D	304	-	5,5,5	0.79	0	5,5,5	1.37	1 (20%)
4	GOL	D	305	-	5,5,5	1.10	0	5,5,5	1.44	1 (20%)
5	EDO	D	306	-	3,3,3	0.83	0	2,2,2	0.48	0
5	EDO	D	307	-	3,3,3	0.62	0	2,2,2	0.26	0
5	EDO	D	308	-	3,3,3	0.49	0	2,2,2	0.74	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
4	GOL	A	301	-	-	0/4/4/4	0/0/0/0
4	GOL	A	302	-	-	0/4/4/4	0/0/0/0
4	GOL	A	303	-	-	0/4/4/4	0/0/0/0
5	EDO	A	304	-	-	0/1/1/1	0/0/0/0
5	EDO	A	305	-	-	0/1/1/1	0/0/0/0
5	EDO	A	306	-	-	0/1/1/1	0/0/0/0
6	SO4	A	307	-	-	0/0/0/0	0/0/0/0
4	GOL	D	301	-	-	0/4/4/4	0/0/0/0
4	GOL	D	302	-	-	0/4/4/4	0/0/0/0
4	GOL	D	303	-	-	0/4/4/4	0/0/0/0
4	GOL	D	304	-	-	0/4/4/4	0/0/0/0
4	GOL	D	305	-	-	0/4/4/4	0/0/0/0
5	EDO	D	306	-	-	0/1/1/1	0/0/0/0
5	EDO	D	307	-	-	0/1/1/1	0/0/0/0
5	EDO	D	308	-	-	0/1/1/1	0/0/0/0

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	D	303	GOL	O2-C2	-3.89	1.31	1.43
4	A	302	GOL	O2-C2	-2.56	1.35	1.43
4	A	302	GOL	C1-C2	2.16	1.60	1.52
4	A	302	GOL	O1-C1	2.45	1.52	1.42

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	302	GOL	O2-C2-C3	-4.48	87.67	108.84
4	A	303	GOL	O1-C1-C2	2.20	121.15	110.07
4	D	305	GOL	O1-C1-C2	2.21	121.20	110.07
4	A	302	GOL	C3-C2-C1	2.28	120.59	111.52
4	D	303	GOL	O1-C1-C2	2.30	121.65	110.07
4	D	301	GOL	O1-C1-C2	2.32	121.75	110.07
4	A	301	GOL	C3-C2-C1	2.54	121.62	111.52
4	D	304	GOL	O1-C1-C2	2.54	122.89	110.07
4	A	302	GOL	O3-C3-C2	3.79	129.18	110.07
4	D	303	GOL	O3-C3-C2	4.46	132.56	110.07
4	A	302	GOL	O1-C1-C2	5.20	136.27	110.07

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

10 monomers are involved in 38 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	301	GOL	2	0
4	A	302	GOL	11	0
4	A	303	GOL	3	0
5	A	305	EDO	1	0
5	A	306	EDO	2	0
4	D	302	GOL	1	0
4	D	303	GOL	6	0
4	D	304	GOL	3	0
4	D	305	GOL	8	0
5	D	308	EDO	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 ⓘ

6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å ²)	Q<0.9
1	A	277/277 (100%)	0.90	43 (15%)	2	3	16, 36, 124, 139	0
1	D	277/277 (100%)	0.81	46 (16%)	2	2	17, 35, 122, 140	0
2	B	100/100 (100%)	0.85	18 (18%)	2	2	25, 60, 89, 111	0
2	E	100/100 (100%)	0.60	12 (12%)	5	6	25, 56, 86, 111	0
3	C	9/9 (100%)	0.33	0	100	100	21, 25, 33, 34	0
3	F	9/9 (100%)	0.26	0	100	100	20, 25, 33, 34	0
All	All	772/772 (100%)	0.81	119 (15%)	2	3	16, 43, 117, 140	0

All (119) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	219	ARG	7.5
1	A	194	ILE	7.0
1	D	250	PRO	6.8
1	A	274	TRP	6.4
2	E	1	ILE	6.2
1	A	195	SER	5.8
1	A	199	ALA	5.8
1	D	258	THR	5.7
1	D	225	THR	5.6
2	E	0	MET	5.6
1	D	221	GLY	5.6
2	B	0	MET	5.4
1	A	257	TYR	5.3
1	D	228	THR	5.3
1	A	223	ASP	5.2
1	A	225	THR	5.2
1	D	197	HIS	5.1
1	A	250	PRO	5.1
1	A	198	GLU	4.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	D	199	ALA	4.7
1	D	195	SER	4.6
1	A	221	GLY	4.6
1	D	220	ASP	4.4
1	D	252	GLY	4.4
1	D	223	ASP	4.2
1	A	197	HIS	4.2
1	A	0	MET	4.1
1	A	268	LYS	4.1
1	D	230	LEU	4.1
1	D	194	ILE	4.0
1	D	245	ALA	4.0
1	D	0	MET	3.9
1	A	220	ASP	3.9
1	A	219	ARG	3.9
1	D	275	GLU	3.8
2	B	95	TRP	3.8
1	A	276	PRO	3.8
1	D	254	GLU	3.8
2	B	1	ILE	3.8
1	D	222	GLU	3.7
1	A	258	THR	3.7
1	A	191	HIS	3.6
2	E	75	LYS	3.5
1	D	268	LYS	3.5
1	A	253	GLU	3.5
1	A	230	LEU	3.4
1	D	261	VAL	3.4
1	A	200	THR	3.4
1	A	251	SER	3.4
1	A	245	ALA	3.3
1	D	198	GLU	3.3
1	D	192	HIS	3.3
1	A	201	LEU	3.3
1	D	276	PRO	3.2
1	D	272	LEU	3.2
1	A	227	ASP	3.1
1	D	274	TRP	3.0
1	D	253	GLU	3.0
1	D	1	GLY	3.0
1	D	203	CYS	2.9
2	B	40	LEU	2.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	E	15	ALA	2.9
2	B	99	MET	2.9
1	D	204	TRP	2.8
2	E	48	LYS	2.8
1	D	189	VAL	2.8
1	D	227	ASP	2.7
2	E	16	GLU	2.7
1	A	193	PRO	2.7
1	D	256	ARG	2.7
1	D	191	HIS	2.7
1	A	192	HIS	2.7
2	B	16	GLU	2.7
2	B	76	ASP	2.7
1	A	190	THR	2.7
2	E	88	SER	2.6
1	A	275	GLU	2.6
1	A	247	VAL	2.6
1	A	217	TRP	2.5
2	B	20	SER	2.5
1	A	215	LEU	2.5
1	D	270	LEU	2.5
1	A	203	CYS	2.5
2	E	74	GLU	2.5
1	D	188	HIS	2.4
1	D	249	VAL	2.4
1	D	217	TRP	2.4
1	D	224	GLN	2.4
1	D	196	ASP	2.4
1	A	18	GLY	2.4
2	B	94	LYS	2.3
2	B	73	THR	2.3
2	E	44	GLU	2.3
2	B	78	TYR	2.3
1	A	259	CYS	2.3
1	A	228	THR	2.3
1	A	196	ASP	2.3
2	B	80	CYS	2.3
1	A	272	LEU	2.2
2	E	78	TYR	2.2
2	B	2	GLN	2.2
2	B	92	ILE	2.2
1	A	218	GLN	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	E	18	GLY	2.1
1	D	267	PRO	2.1
1	A	266	LEU	2.1
1	A	248	VAL	2.1
1	D	244	TRP	2.1
2	B	75	LYS	2.1
1	D	190	THR	2.1
1	D	202	ARG	2.1
1	D	248	VAL	2.1
2	B	98	ASP	2.1
2	B	44	GLU	2.0
1	D	193	PRO	2.0
2	B	19	LYS	2.0
1	A	108	ARG	2.0
2	E	19	LYS	2.0
1	A	255	GLN	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 carbohydrates 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. 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, 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	LLDF	B-factors(Å ²)	Q<0.9
4	GOL	D	305	6/6	0.85	0.33	7.62	27,40,44,65	0
4	GOL	A	302	6/6	0.87	0.33	6.50	24,28,33,37	0
4	GOL	A	301	6/6	0.85	0.20	5.19	27,38,40,41	0
4	GOL	D	303	6/6	0.93	0.24	4.16	24,27,32,33	0
4	GOL	A	303	6/6	0.84	0.23	3.45	31,32,36,39	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
6	SO4	A	307	5/5	0.94	0.19	2.76	43,68,71,76	0
5	EDO	A	304	4/4	0.95	0.17	1.09	32,36,39,41	0
4	GOL	D	301	6/6	0.91	0.16	0.54	38,40,47,54	0
5	EDO	D	306	4/4	0.94	0.15	0.51	24,25,25,26	0
5	EDO	D	307	4/4	0.90	0.16	0.42	37,46,46,49	0
4	GOL	D	302	6/6	0.77	0.20	0.23	41,52,61,63	0
5	EDO	A	306	4/4	0.96	0.12	-1.58	39,45,49,52	0
5	EDO	A	305	4/4	0.91	0.11	-1.92	39,41,45,51	0
5	EDO	D	308	4/4	0.85	0.32	-	41,41,42,46	0
4	GOL	D	304	6/6	0.86	0.27	-	24,37,45,54	0

6.5 Other polymers [i](#)

There are no such residues in this entry.