



# Full wwPDB X-ray Structure Validation Report ⓘ

Jun 2, 2016 – 05:13 AM EDT

PDB ID : 5C0I  
Title : HAL-A02 carrying RQFGPDFPTI  
Authors : Rizkallah, P.J.; Bulek, A.M.; Cole, D.K.; Sewell, A.K.  
Deposited on : 2015-06-12  
Resolution : 1.53 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](mailto: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 : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20027674  
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-20027674

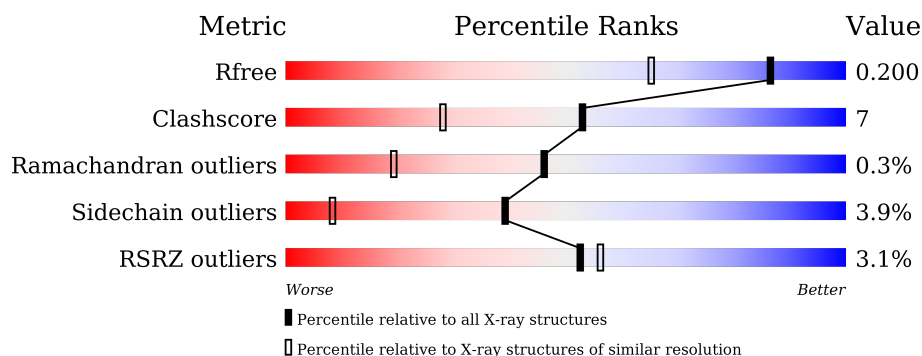
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.53 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	1555 (1.56-1.52)
Clashscore	102246	1627 (1.56-1.52)
Ramachandran outliers	100387	1594 (1.56-1.52)
Sidechain outliers	100360	1592 (1.56-1.52)
RSRZ outliers	91569	1555 (1.56-1.52)

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  $\geq 3$ , 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>2%</div> <div> <div></div> <div>81%</div> <div>15%</div> <div>.</div> </div> </div>
2	B	100	<div> <div>6%</div> <div> <div></div> <div>77%</div> <div>20%</div> <div>..</div> </div> </div>
3	C	10	<div> <div></div> <div> <div></div> <div>80%</div> <div>20%</div> </div> </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	EDO	A	303	-	-	-	X
4	EDO	A	304	-	-	-	X
4	EDO	A	305	-	-	-	X
5	GOL	A	306	-	-	-	X
5	GOL	A	307	-	-	-	X
5	GOL	B	107	-	-	-	X
5	GOL	B	108	-	-	X	X
6	SO4	A	308	-	-	X	-
6	SO4	A	309	-	-	-	X

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 3703 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	277	Total	C	N	O	S	0	17	0
			2411	1495	448	456	12			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	initiating methionine	UNP P01892

- 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	2	0
			852	541	143	164	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 Marker peptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	10	Total	C	N	O	0	1	0
			95	61	18	16			

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



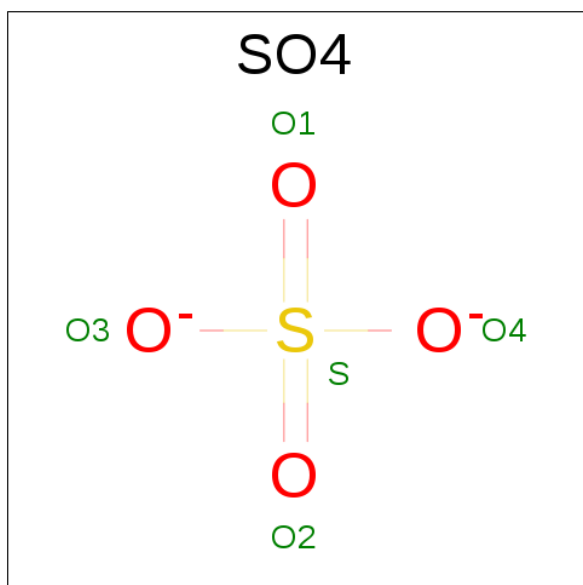
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			4	2	2		
4	A	1	Total	C	O	0	0
			4	2	2		
4	A	1	Total	C	O	0	0
			4	2	2		
4	A	1	Total	C	O	0	0
			4	2	2		
4	A	1	Total	C	O	0	0
			4	2	2		
4	B	1	Total	C	O	0	0
			4	2	2		
4	B	1	Total	C	O	0	0
			4	2	2		
4	B	1	Total	C	O	0	0
			4	2	2		
4	B	1	Total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

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			6	3	3		
5	A	1	Total	C	O	0	0
			6	3	3		
5	B	1	Total	C	O	0	0
			6	3	3		
5	B	1	Total	C	O	0	0
			6	3	3		

- Molecule 6 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



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

- Molecule 7 is CALCIUM ION (three-letter code: CA) (formula: Ca).

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

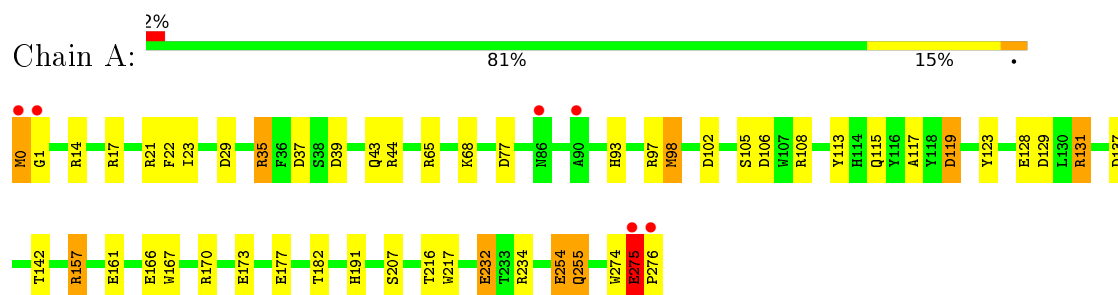
- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	189	Total	O	0	0
			189	189		
8	B	71	Total	O	0	0
			71	71		
8	C	10	Total	O	0	0
			10	10		

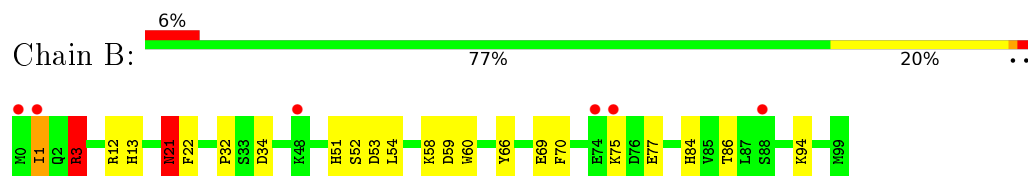
### 3 Residue-property plots [i](#)

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 ( $\text{RSRZ} > 2$ ). 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.

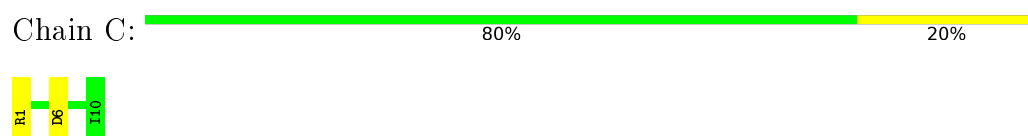
- Molecule 1: HLA class I histocompatibility antigen, A-2 alpha chain



- Molecule 2: Beta-2-microglobulin



- Molecule 3: Marker peptide





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	55.70 Å 79.74 Å 58.08 Å 90.00° 115.67° 90.00°	Depositor
Resolution (Å)	28.30 – 1.53 28.30 – 1.53	Depositor EDS
% Data completeness (in resolution range)	98.8 (28.30-1.53) 98.8 (28.30-1.53)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.33 (at 1.53 Å)	Xtriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.159 , 0.191 0.172 , 0.200	Depositor DCC
$R_{free}$ test set	3448 reflections (5.35%)	DCC
Wilson B-factor (Å <sup>2</sup> )	19.1	Xtriage
Anisotropy	0.725	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 45.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.023 for l,-k,h	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	3703	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	25.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.18% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup> Intensities estimated from amplitudes.

<sup>2</sup> Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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, CA, 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	1.13	7/2477 (0.3%)	1.30	36/3355 (1.1%)
2	B	0.99	2/875 (0.2%)	1.10	5/1182 (0.4%)
3	C	1.13	0/98	1.63	4/130 (3.1%)
All	All	1.09	9/3450 (0.3%)	1.26	45/4667 (1.0%)

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	1

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	173[A]	GLU	CD-OE2	-7.80	1.17	1.25
1	A	173[B]	GLU	CD-OE2	-7.80	1.17	1.25
1	A	217	TRP	CG-CD1	6.97	1.46	1.36
1	A	274	TRP	CE3-CZ3	6.57	1.49	1.38
1	A	254	GLU	CD-OE2	-5.73	1.19	1.25
2	B	21	ASN	CG-OD1	5.33	1.35	1.24
1	A	113	TYR	CG-CD2	-5.27	1.32	1.39
1	A	97	ARG	CZ-NH2	-5.24	1.26	1.33
2	B	66	TYR	CZ-OH	5.18	1.46	1.37

All (45) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	97	ARG	NE-CZ-NH1	16.45	128.53	120.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	35[A]	ARG	NE-CZ-NH2	11.15	125.88	120.30
1	A	35[B]	ARG	NE-CZ-NH2	11.15	125.88	120.30
2	B	3	ARG	NE-CZ-NH1	10.87	125.73	120.30
1	A	157[A]	ARG	NE-CZ-NH2	-10.56	115.02	120.30
1	A	157[B]	ARG	NE-CZ-NH2	-10.56	115.02	120.30
1	A	157[A]	ARG	NE-CZ-NH1	10.49	125.55	120.30
1	A	157[B]	ARG	NE-CZ-NH1	10.49	125.55	120.30
1	A	97	ARG	NE-CZ-NH2	-10.35	115.13	120.30
3	C	6	ASP	CB-CG-OD2	-8.57	110.59	118.30
1	A	35[A]	ARG	NE-CZ-NH1	-7.80	116.40	120.30
1	A	35[B]	ARG	NE-CZ-NH1	-7.80	116.40	120.30
1	A	234	ARG	NE-CZ-NH1	7.52	124.06	120.30
2	B	34	ASP	CB-CG-OD1	7.22	124.80	118.30
1	A	131	ARG	NE-CZ-NH2	-6.76	116.92	120.30
1	A	21	ARG	NE-CZ-NH2	-6.72	116.94	120.30
3	C	1[A]	ARG	NE-CZ-NH2	-6.34	117.13	120.30
3	C	1[B]	ARG	NE-CZ-NH2	-6.34	117.13	120.30
2	B	12	ARG	NE-CZ-NH2	-6.30	117.15	120.30
1	A	22	PHE	CB-CG-CD2	-6.06	116.56	120.80
1	A	234	ARG	NE-CZ-NH2	-6.02	117.29	120.30
1	A	17	ARG	NE-CZ-NH2	-5.91	117.34	120.30
1	A	232[A]	GLU	OE1-CD-OE2	-5.89	116.23	123.30
1	A	232[B]	GLU	OE1-CD-OE2	-5.89	116.23	123.30
1	A	106	ASP	CB-CG-OD2	-5.88	113.01	118.30
1	A	131	ARG	NE-CZ-NH1	5.87	123.23	120.30
1	A	170	ARG	NE-CZ-NH1	5.84	123.22	120.30
1	A	137	ASP	CB-CG-OD1	5.77	123.49	118.30
2	B	3	ARG	NE-CZ-NH2	-5.75	117.42	120.30
1	A	102	ASP	CB-CG-OD1	5.65	123.39	118.30
1	A	129	ASP	CB-CG-OD1	5.48	123.23	118.30
1	A	35[A]	ARG	CD-NE-CZ	5.47	131.25	123.60
1	A	35[B]	ARG	CD-NE-CZ	5.47	131.25	123.60
1	A	21	ARG	NE-CZ-NH1	5.43	123.02	120.30
3	C	6	ASP	CB-CG-OD1	5.34	123.10	118.30
1	A	119	ASP	CB-CG-OD2	-5.32	113.51	118.30
1	A	65	ARG	NE-CZ-NH2	-5.31	117.65	120.30
1	A	77	ASP	CB-CG-OD2	-5.25	113.57	118.30
1	A	108	ARG	CB-CG-CD	5.22	125.16	111.60
1	A	123	TYR	CZ-CE2-CD2	-5.15	115.17	119.80
1	A	113	TYR	CB-CG-CD1	-5.13	117.92	121.00
1	A	37	ASP	CB-CG-OD1	5.12	122.91	118.30
1	A	137	ASP	CB-CG-OD2	-5.03	113.77	118.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	39	ASP	CB-CG-OD2	-5.01	113.79	118.30
2	B	34	ASP	CB-CG-OD2	-5.00	113.80	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	275	GLU	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	2411	0	2241	27	0
2	B	852	0	812	18	0
3	C	95	0	94	0	0
4	A	20	0	30	2	0
4	B	20	0	30	1	0
5	A	12	0	16	2	0
5	B	12	0	16	6	0
6	A	10	0	0	3	0
7	B	1	0	0	0	0
8	A	189	0	0	5	0
8	B	71	0	0	0	0
8	C	10	0	0	0	0
All	All	3703	0	3239	44	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 7.

All (44) 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:216[B]:THR:HG22	8:A:513:HOH:O	1.72	0.90
2:B:21:ASN:HD22	2:B:22:PHE:H	1.18	0.88

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:23:ILE:HD13	5:B:108:GOL:H31	1.54	0.87
2:B:54:LEU:H	5:B:108:GOL:H32	1.42	0.84
2:B:77:GLU:OE2	2:B:94:LYS:NZ	2.20	0.74
1:A:68:LYS:HB3	5:A:307:GOL:H11	1.68	0.74
1:A:232[A]:GLU:OE2	8:A:401:HOH:O	2.07	0.71
1:A:131:ARG:HH22	1:A:157[B]:ARG:NH1	1.88	0.71
2:B:53:ASP:HA	5:B:108:GOL:H32	1.80	0.64
2:B:13:HIS:H	2:B:21:ASN:HD21	1.44	0.63
1:A:93:HIS:HE1	8:A:503:HOH:O	1.81	0.63
2:B:21:ASN:ND2	2:B:22:PHE:H	1.95	0.63
2:B:3:ARG:NH2	2:B:59:ASP:O	2.32	0.62
2:B:84:HIS:HD2	2:B:86:THR:OG1	1.84	0.60
2:B:54:LEU:H	5:B:108:GOL:C3	2.14	0.59
2:B:54:LEU:N	5:B:108:GOL:H32	2.16	0.58
1:A:98:MET:HE3	1:A:98:MET:C	2.24	0.57
1:A:0:MET:HB3	1:A:182:THR:CG2	2.34	0.57
1:A:29:ASP:O	5:A:306:GOL:H31	2.08	0.54
1:A:191:HIS:HE1	1:A:254:GLU:OE1	1.92	0.53
1:A:166:GLU:HB3	6:A:308:SO4:O2	2.09	0.53
2:B:21:ASN:HD22	2:B:22:PHE:N	1.97	0.53
1:A:216[B]:THR:CG2	8:A:513:HOH:O	2.44	0.52
1:A:131:ARG:HH22	1:A:157[B]:ARG:CZ	2.22	0.52
1:A:98:MET:HE3	1:A:98:MET:O	2.10	0.51
1:A:167:TRP:CE2	6:A:308:SO4:O4	2.64	0.51
1:A:93:HIS:HD2	1:A:119:ASP:OD2	1.95	0.50
2:B:58:LYS:O	4:B:105:EDO:H22	2.11	0.50
1:A:255:GLN:NE2	1:A:255:GLN:H	2.12	0.48
1:A:131:ARG:NH2	1:A:157[B]:ARG:NH1	2.60	0.48
1:A:142:THR:OG1	4:A:303:EDO:H12	2.14	0.47
1:A:161:GLU:OE1	6:A:309:SO4:O1	2.33	0.47
2:B:1:ILE:HG13	2:B:3:ARG:HH11	1.80	0.47
1:A:131:ARG:NH2	1:A:157[B]:ARG:CZ	2.78	0.46
1:A:44:ARG:HE	4:A:305:EDO:C2	2.30	0.45
1:A:43:GLN:O	1:A:68:LYS:HE3	2.17	0.44
1:A:117:ALA:HB2	2:B:60:TRP:CE2	2.54	0.43
2:B:32:PRO:O	2:B:84:HIS:HE1	2.02	0.43
1:A:14:ARG:NH2	8:A:411:HOH:O	2.52	0.42
1:A:275:GLU:H	1:A:276:PRO:HD2	1.85	0.41
1:A:1:GLY:HA2	1:A:105[A]:SER:OG	2.20	0.41
2:B:51:HIS:HD2	2:B:52:SER:O	2.04	0.40
2:B:22:PHE:CE2	2:B:69[A]:GLU:HG2	2.57	0.40

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:53:ASP:CA	5:B:108:GOL:H32	2.50	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	292/277 (105%)	287 (98%)	4 (1%)	1 (0%)	46	19
2	B	100/100 (100%)	99 (99%)	1 (1%)	0	100	100
3	C	8/10 (80%)	8 (100%)	0	0	100	100
All	All	400/387 (103%)	394 (98%)	5 (1%)	1 (0%)	46	19

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	275	GLU

### 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	250/233 (107%)	239 (96%)	11 (4%)	35	6
2	B	97/95 (102%)	92 (95%)	5 (5%)	29	4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	C	10/9 (111%)	10 (100%)	0	100	100
All	All	357/337 (106%)	341 (96%)	16 (4%)	39	5

All (16) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	0	MET
1	A	35[A]	ARG
1	A	35[B]	ARG
1	A	98	MET
1	A	115[A]	GLN
1	A	115[B]	GLN
1	A	128	GLU
1	A	177[A]	GLU
1	A	177[B]	GLU
1	A	207	SER
1	A	255	GLN
2	B	1	ILE
2	B	3	ARG
2	B	21	ASN
2	B	70	PHE
2	B	75	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (11) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	86	ASN
1	A	93	HIS
1	A	180	GLN
1	A	191	HIS
1	A	192	HIS
1	A	255	GLN
2	B	21	ASN
2	B	24	ASN
2	B	51	HIS
2	B	84	HIS
2	B	89	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 ⓘ

Of 17 ligands modelled in this entry, 1 is monoatomic - leaving 16 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	$\# Z  > 2$
4	EDO	A	301	-	3,3,3	0.45	0	2,2,2	1.04	0
4	EDO	A	302	-	3,3,3	0.19	0	2,2,2	1.22	0
4	EDO	A	303	-	3,3,3	0.84	0	2,2,2	0.76	0
4	EDO	A	304	-	3,3,3	0.42	0	2,2,2	0.39	0
4	EDO	A	305	-	3,3,3	0.32	0	2,2,2	1.41	0
5	GOL	A	306	-	5,5,5	0.47	0	5,5,5	1.19	1 (20%)
5	GOL	A	307	-	5,5,5	0.57	0	5,5,5	2.52	2 (40%)
6	SO4	A	308	-	4,4,4	0.67	0	6,6,6	0.70	0
6	SO4	A	309	-	4,4,4	0.69	0	6,6,6	0.33	0
4	EDO	B	102	-	3,3,3	0.62	0	2,2,2	0.85	0
4	EDO	B	103	-	3,3,3	0.47	0	2,2,2	0.24	0
4	EDO	B	104	-	3,3,3	0.46	0	2,2,2	0.64	0
4	EDO	B	105	-	3,3,3	0.59	0	2,2,2	0.49	0
4	EDO	B	106	-	3,3,3	0.55	0	2,2,2	0.32	0
5	GOL	B	107	-	5,5,5	0.27	0	5,5,5	0.66	0
5	GOL	B	108	-	5,5,5	1.12	1 (20%)	5,5,5	0.93	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	EDO	A	301	-	-	0/1/1/1	0/0/0/0
4	EDO	A	302	-	-	0/1/1/1	0/0/0/0
4	EDO	A	303	-	-	0/1/1/1	0/0/0/0
4	EDO	A	304	-	-	0/1/1/1	0/0/0/0
4	EDO	A	305	-	-	0/1/1/1	0/0/0/0
5	GOL	A	306	-	-	0/4/4/4	0/0/0/0
5	GOL	A	307	-	-	0/4/4/4	0/0/0/0
6	SO4	A	308	-	-	0/0/0/0	0/0/0/0
6	SO4	A	309	-	-	0/0/0/0	0/0/0/0
4	EDO	B	102	-	-	0/1/1/1	0/0/0/0
4	EDO	B	103	-	-	0/1/1/1	0/0/0/0
4	EDO	B	104	-	-	0/1/1/1	0/0/0/0
4	EDO	B	105	-	-	0/1/1/1	0/0/0/0
4	EDO	B	106	-	-	0/1/1/1	0/0/0/0
5	GOL	B	107	-	-	0/4/4/4	0/0/0/0
5	GOL	B	108	-	-	0/4/4/4	0/0/0/0

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	108	GOL	O3-C3	2.14	1.51	1.42

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	306	GOL	O2-C2-C3	-2.11	98.31	108.47
5	A	307	GOL	C3-C2-C1	3.24	124.72	111.06
5	A	307	GOL	O1-C1-C2	3.92	129.83	109.97

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

8 monomers are involved in 14 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	303	EDO	1	0
4	A	305	EDO	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	306	GOL	1	0
5	A	307	GOL	1	0
6	A	308	SO4	2	0
6	A	309	SO4	1	0
4	B	105	EDO	1	0
5	B	108	GOL	6	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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	277/277 (100%)	0.01	6 (2%) 65 68	13, 21, 37, 72	0
2	B	100/100 (100%)	0.20	6 (6%) 25 26	16, 25, 53, 80	0
3	C	10/10 (100%)	-0.19	0 100 100	17, 22, 28, 28	0
All	All	387/387 (100%)	0.05	12 (3%) 52 56	13, 22, 42, 80	0

All (12) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	0	MET	7.3
2	B	1	ILE	6.0
1	A	276	PRO	5.5
2	B	0	MET	4.5
1	A	275	GLU	4.0
2	B	75	LYS	3.5
2	B	48	LYS	3.3
1	A	1	GLY	2.7
2	B	74	GLU	2.6
2	B	88	SER	2.5
1	A	86	ASN	2.5
1	A	90	ALA	2.1

### 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 ⓘ

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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
4	EDO	A	305	4/4	0.85	0.28	24.20	35,46,48,50	0
4	EDO	A	303	4/4	0.91	0.21	8.35	29,33,33,44	0
5	GOL	B	108	6/6	0.83	0.20	6.73	28,36,44,54	0
6	SO4	A	309	5/5	0.86	0.32	5.88	58,65,82,87	0
5	GOL	A	307	6/6	0.89	0.17	4.68	30,36,39,58	0
4	EDO	A	304	4/4	0.89	0.16	3.56	27,27,30,34	0
5	GOL	B	107	6/6	0.93	0.14	2.91	23,36,40,56	0
5	GOL	A	306	6/6	0.88	0.13	2.42	40,48,54,54	0
4	EDO	B	102	4/4	0.96	0.10	1.76	19,22,27,28	0
4	EDO	A	302	4/4	0.92	0.10	1.71	30,30,31,34	0
4	EDO	A	301	4/4	0.95	0.13	1.49	37,38,40,49	0
4	EDO	B	106	4/4	0.91	0.10	0.92	31,33,35,44	0
4	EDO	B	104	4/4	0.89	0.20	0.87	36,39,42,43	0
4	EDO	B	105	4/4	0.86	0.13	0.44	35,37,38,38	0
4	EDO	B	103	4/4	0.94	0.07	-0.93	28,31,35,36	0
6	SO4	A	308	5/5	0.92	0.27	-	50,56,70,79	0
7	CA	B	101	1/1	0.98	0.20	-	53,53,53,53	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.