



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 19, 2016 – 09:05 PM GMT

PDB ID : 4Z78  
Title : Weak TCR binding to an unstable insulin epitope drives type 1 diabetes  
Authors : Rizkallah, P.J.; Cole, D.K.  
Deposited on : 2015-04-06  
Resolution : 2.30 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-20026982  
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-20026982

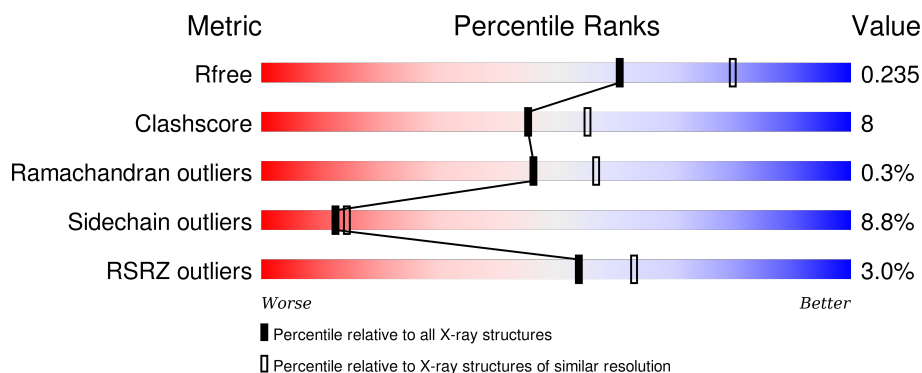
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.30 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.





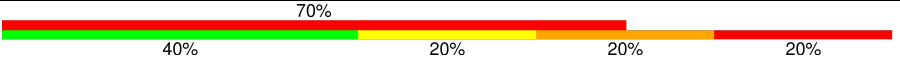
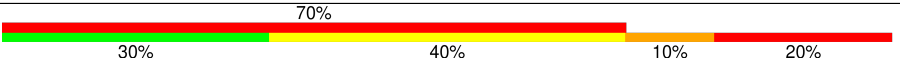
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	91344	3852 (2.30-2.30)
Clashscore	102246	4452 (2.30-2.30)
Ramachandran outliers	100387	4410 (2.30-2.30)
Sidechain outliers	100360	4409 (2.30-2.30)
RSRZ outliers	91569	3857 (2.30-2.30)

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>%</div> <div> <div></div> <div>81%</div> <div>17%</div> <div>.</div> </div> </div>
1	D	277	<div> <div>%</div> <div> <div></div> <div>83%</div> <div>14%</div> <div>.</div> </div> </div>
1	G	277	<div> <div>3%</div> <div> <div></div> <div>83%</div> <div>14%</div> <div>.</div> </div> </div>
2	B	100	<div> <div></div> <div> <div></div> <div>85%</div> <div>13%</div> <div>.</div> </div> </div>
2	E	100	<div> <div>%</div> <div> <div></div> <div>85%</div> <div>14%</div> <div>.</div> </div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	H	100	
3	C	10	
3	F	10	
3	I	10	

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	302	-	-	-	X
4	EDO	H	301	-	-	-	X
5	GOL	A	304	-	-	-	X
5	GOL	A	305	-	-	-	X
5	GOL	D	302	-	-	-	X
5	GOL	H	302	-	-	-	X
5	GOL	H	303	-	-	-	X
6	SO4	A	306	-	-	-	X
6	SO4	A	307	-	-	-	X
6	SO4	A	308	-	-	-	X
6	SO4	D	303	-	-	-	X
6	SO4	D	305	-	-	-	X
6	SO4	D	306	-	-	-	X
6	SO4	G	301	-	-	-	X

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 10098 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 H-2 class I histocompatibility antigen, K-D alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	277	Total	C	N	O	S	0	0	0
			2287	1449	406	424	8			
1	D	277	Total	C	N	O	S	0	0	0
			2287	1449	406	424	8			
1	G	277	Total	C	N	O	S	0	1	0
			2298	1455	410	425	8			

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	initiating methionine	UNP P01902
A	114	HIS	GLN	conflict	UNP P01902
A	276	PRO	-	expression tag	UNP P01902
D	0	MET	-	initiating methionine	UNP P01902
D	114	HIS	GLN	conflict	UNP P01902
D	276	PRO	-	expression tag	UNP P01902
G	0	MET	-	initiating methionine	UNP P01902
G	114	HIS	GLN	conflict	UNP P01902
G	276	PRO	-	expression tag	UNP P01902

- 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	1	0
			844	538	142	160	4			
2	H	100	Total	C	N	O	S	0	0	0
			837	533	141	159	4			

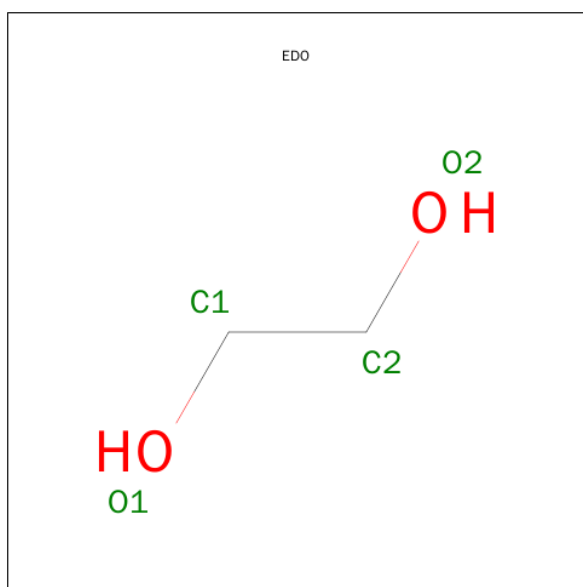
There are 3 discrepancies between the modelled and reference sequences:

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

- Molecule 3 is a protein called Insulin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	10	Total	C	N	O	S	0	0	0
			81	53	13	14	1			
3	F	10	Total	C	N	O	S	0	0	0
			81	53	13	14	1			
3	I	10	Total	C	N	O	S	0	0	0
			81	53	13	14	1			

- 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	D	1	Total	C	O	0	0
			4	2	2		
4	H	1	Total	C	O	0	0
			4	2	2		

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



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	D	1	Total	C	O	0	0
			6	3	3		
5	E	1	Total	C	O	0	0
			6	3	3		
5	H	1	Total	C	O	0	0
			6	3	3		
5	H	1	Total	C	O	0	0
			6	3	3		

- Molecule 6 is SULFATE ION (three-letter code: SO4) (formula:  $O_4S$ ).



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		
6	A	1	Total	O	S	0	0
			5	4	1		
6	A	1	Total	O	S	0	0
			5	4	1		
6	D	1	Total	O	S	0	0
			5	4	1		
6	D	1	Total	O	S	0	0
			5	4	1		
6	D	1	Total	O	S	0	0
			5	4	1		
6	D	1	Total	O	S	0	0
			5	4	1		
6	G	1	Total	O	S	0	0
			5	4	1		

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	100	Total	O	0	0
			100	100		
7	B	46	Total	O	0	0
			46	46		
7	D	87	Total	O	0	0
			87	87		

*Continued on next page...*

*Continued from previous page...*

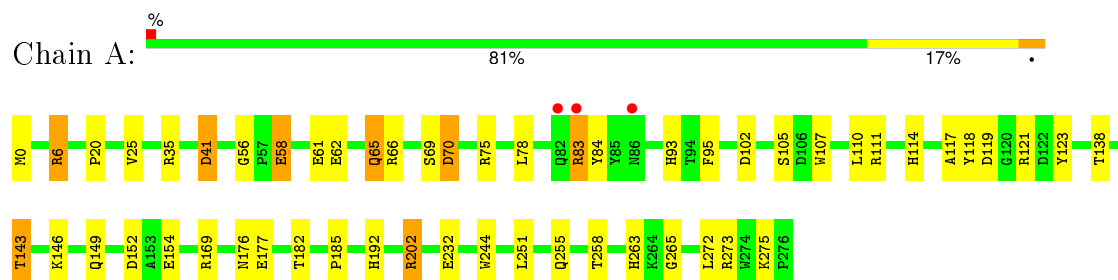
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	E	27	Total 27	O 27	0	0
7	G	63	Total 63	O 63	0	0
7	H	41	Total 41	O 41	0	0



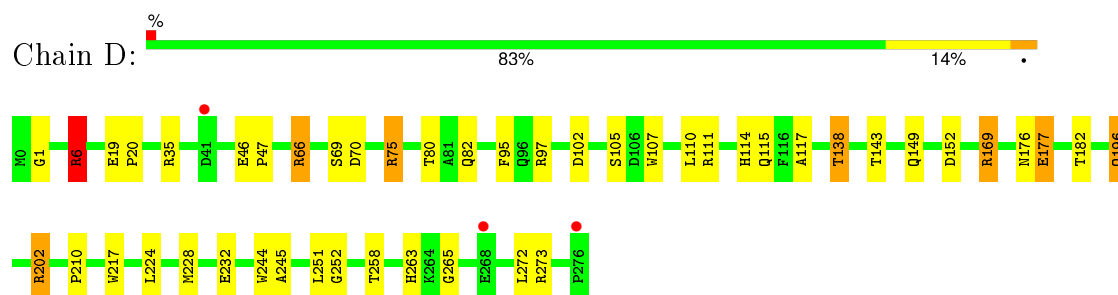
### 3 Residue-property plots

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.

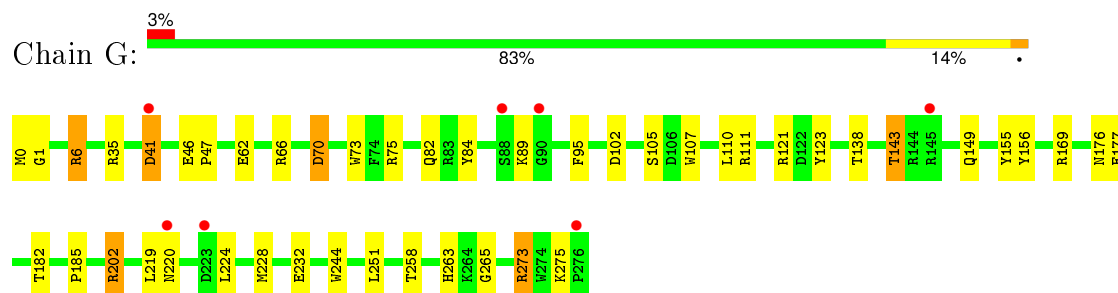
- Molecule 1: H-2 class I histocompatibility antigen, K-D alpha chain



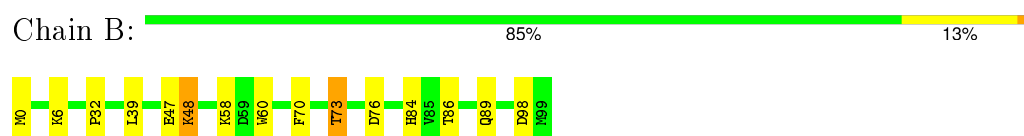
- Molecule 1: H-2 class I histocompatibility antigen, K-D alpha chain



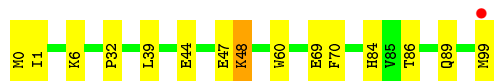
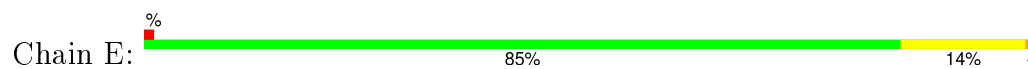
- Molecule 1: H-2 class I histocompatibility antigen, K-D alpha chain



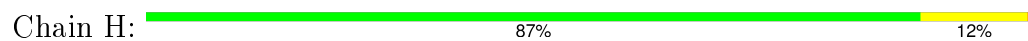
- Molecule 2: Beta-2-microglobulin



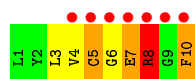
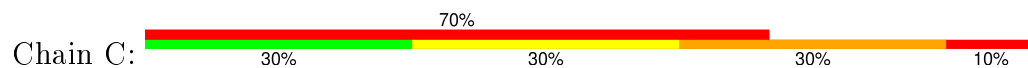
- Molecule 2: Beta-2-microglobulin



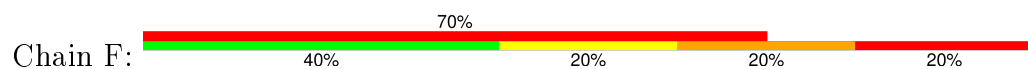
- Molecule 2: Beta-2-microglobulin



- Molecule 3: Insulin



- Molecule 3: Insulin



- Molecule 3: Insulin



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	46.17Å 151.57Å 182.23Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	78.09 – 2.30 78.09 – 2.30	Depositor EDS
% Data completeness (in resolution range)	100.0 (78.09-2.30) 100.0 (78.09-2.30)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.13	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.53 (at 2.29Å)	Xtriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.188 , 0.233 0.193 , 0.235	Depositor DCC
$R_{free}$ test set	2925 reflections (5.35%)	DCC
Wilson B-factor (Å <sup>2</sup> )	34.9	Xtriage
Anisotropy	0.082	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 34.0	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	2 of 57644 reflections (0.003%)	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	10098	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 21.63 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 6.7426e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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, 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	0.85	1/2357 (0.0%)	0.91	4/3205 (0.1%)
1	D	0.83	0/2357	0.90	6/3205 (0.2%)
1	G	0.80	0/2368	0.94	7/3219 (0.2%)
2	B	0.73	0/860	0.80	0/1162
2	E	0.63	0/867	0.77	0/1172
2	H	0.72	0/860	0.79	0/1162
3	C	0.91	0/82	1.29	0/107
3	F	1.03	0/82	1.28	1/107 (0.9%)
3	I	0.73	0/82	1.32	1/107 (0.9%)
All	All	0.80	1/9915 (0.0%)	0.89	19/13446 (0.1%)

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
3	C	0	1
3	I	0	2
All	All	0	4

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	58	GLU	CD-OE1	5.10	1.31	1.25

All (19) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	G	121	ARG	NE-CZ-NH2	-10.28	115.16	120.30
1	G	121	ARG	NE-CZ-NH1	8.94	124.77	120.30
1	G	202	ARG	NE-CZ-NH2	-7.21	116.70	120.30
1	A	121	ARG	NE-CZ-NH1	-7.06	116.77	120.30
3	I	9	GLY	N-CA-C	-6.71	96.32	113.10
1	D	75	ARG	NE-CZ-NH2	6.35	123.47	120.30
1	D	202	ARG	NE-CZ-NH1	6.21	123.40	120.30
1	D	169	ARG	NE-CZ-NH1	-6.20	117.20	120.30
1	A	202	ARG	NE-CZ-NH2	-6.14	117.23	120.30
1	D	1	GLY	N-CA-C	-5.97	98.18	113.10
1	A	121	ARG	NE-CZ-NH2	5.75	123.18	120.30
1	G	1	GLY	N-CA-C	-5.58	99.15	113.10
3	F	4	VAL	CB-CA-C	-5.30	101.33	111.40
1	A	202	ARG	NE-CZ-NH1	5.25	122.92	120.30
1	D	6	ARG	NE-CZ-NH1	5.23	122.92	120.30
1	G	202	ARG	NE-CZ-NH1	5.16	122.88	120.30
1	D	115	GLN	N-CA-CB	-5.15	101.33	110.60
1	G	84	TYR	CB-CG-CD2	-5.08	117.95	121.00
1	G	228	MET	CA-CB-CG	5.08	121.93	113.30

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	114	HIS	Peptide
3	C	8	ARG	Peptide
3	I	4	VAL	Peptide
3	I	8	ARG	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	2153	38	0
1	D	2287	0	2153	36	0
1	G	2298	0	2165	26	0
2	B	837	0	803	9	0
2	E	844	0	811	6	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	H	837	0	803	8	0
3	C	81	0	81	18	0
3	F	81	0	81	24	0
3	I	81	0	81	14	0
4	A	12	0	18	1	0
4	D	4	0	6	0	0
4	H	4	0	6	0	0
5	A	12	0	16	4	0
5	D	6	0	8	0	0
5	E	6	0	8	1	0
5	H	12	0	16	2	0
6	A	20	0	0	0	0
6	D	20	0	0	0	0
6	G	5	0	0	1	0
7	A	100	0	0	2	0
7	B	46	0	0	0	0
7	D	87	0	0	1	1
7	E	27	0	0	0	0
7	G	63	0	0	1	1
7	H	41	0	0	2	0
All	All	10098	0	9209	147	1

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 8.

All (147) 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)
3:C:5:CYS:HB3	3:C:8:ARG:HH11	1.02	1.14
3:C:5:CYS:HB3	3:C:8:ARG:NH1	1.61	1.13
3:C:5:CYS:CB	3:C:8:ARG:HH11	1.64	1.11
3:F:10:PHE:OXT	3:F:10:PHE:CD1	2.11	1.04
3:F:8:ARG:NH2	3:F:8:ARG:H	1.60	0.99
3:F:10:PHE:OXT	3:F:10:PHE:HD1	1.40	0.98
3:F:8:ARG:N	3:F:8:ARG:HH21	1.67	0.93
1:A:152:ASP:OD2	3:C:8:ARG:NE	2.05	0.87
1:A:70:ASP:OD1	3:C:4:VAL:HG23	1.80	0.82
1:A:143:THR:HG22	3:C:10:PHE:HB3	1.60	0.82
1:A:143:THR:CG2	3:C:10:PHE:HB3	2.09	0.82
1:A:58:GLU:OE2	1:D:138:THR:HG23	1.80	0.82
3:I:10:PHE:C	3:I:10:PHE:CD1	2.53	0.81

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:66:ARG:HG2	3:C:4:VAL:HG12	1.64	0.80
1:G:156:TYR:OH	3:I:4:VAL:O	2.00	0.77
3:I:7:GLU:O	3:I:8:ARG:HB3	1.82	0.77
3:F:8:ARG:HH21	3:F:8:ARG:H	0.82	0.74
1:D:143:THR:HG23	3:F:9:GLY:CA	2.19	0.72
2:B:73:THR:HG22	2:B:76:ASP:H	1.55	0.71
2:H:73:THR:HG22	2:H:76:ASP:H	1.56	0.71
1:D:217:TRP:N	1:D:228:MET:HE3	2.06	0.70
1:G:73:TRP:HE1	3:I:8:ARG:HB2	1.56	0.69
1:A:83:ARG:HG2	1:A:84:TYR:N	2.09	0.68
1:A:93:HIS:HD2	1:A:119:ASP:OD2	1.76	0.68
1:A:93:HIS:HB2	5:A:305:GOL:H12	1.77	0.67
1:A:84:TYR:CD1	1:A:84:TYR:C	2.69	0.66
1:D:6:ARG:NH2	1:D:102:ASP:OD1	2.29	0.66
1:D:143:THR:HG23	3:F:9:GLY:HA3	1.78	0.65
3:C:5:CYS:CB	3:C:8:ARG:NH1	2.40	0.63
1:G:73:TRP:HE1	3:I:8:ARG:CB	2.11	0.62
1:D:176:ASN:ND2	1:D:177:GLU:OE2	2.32	0.62
1:G:75[B]:ARG:HD3	6:G:301:SO4:O3	1.98	0.62
1:D:196:GLN:HE21	1:D:196:GLN:HA	1.64	0.62
1:A:41:ASP:OD1	1:A:41:ASP:N	2.33	0.62
3:F:9:GLY:O	3:F:10:PHE:CD2	2.52	0.62
2:H:84:HIS:HD2	2:H:86:THR:OG1	1.83	0.61
1:A:78:LEU:HB3	5:A:305:GOL:H11	1.82	0.61
1:A:6:ARG:NH2	1:A:102:ASP:OD1	2.33	0.61
1:A:232:GLU:OE2	2:B:6:LYS:NZ	2.31	0.61
1:D:143:THR:HG23	3:F:9:GLY:HA2	1.81	0.61
1:D:217:TRP:H	1:D:228:MET:HE3	1.66	0.61
1:A:118:TYR:HD2	5:A:304:GOL:H31	1.66	0.60
2:E:48:LYS:HE3	2:E:48:LYS:HA	1.83	0.60
2:B:48:LYS:HE3	2:B:48:LYS:HA	1.84	0.59
1:G:263:HIS:CD2	1:G:265:GLY:H	2.20	0.59
2:E:84:HIS:HD2	2:E:86:THR:OG1	1.84	0.59
2:B:84:HIS:HD2	2:B:86:THR:OG1	1.86	0.59
1:A:192:HIS:CE1	2:B:98:ASP:HB3	2.39	0.57
1:G:70:ASP:OD1	3:I:5:CYS:N	2.36	0.57
1:G:107:TRP:CZ3	1:G:169:ARG:HG2	2.39	0.57
3:I:10:PHE:C	3:I:10:PHE:HD1	2.06	0.57
1:D:263:HIS:CD2	1:D:265:GLY:H	2.22	0.56
1:G:41:ASP:OD1	1:G:41:ASP:N	2.31	0.56
1:D:202:ARG:HD2	1:D:244:TRP:CD2	2.40	0.56

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:F:9:GLY:O	3:F:10:PHE:CG	2.58	0.56
1:A:263:HIS:CD2	1:A:265:GLY:H	2.24	0.55
3:F:8:ARG:NH2	3:F:8:ARG:N	2.38	0.55
1:D:217:TRP:HD1	1:D:228:MET:HE2	1.72	0.55
1:D:19:GLU:HB3	1:D:75:ARG:NH1	2.22	0.54
1:G:6:ARG:NH2	1:G:102:ASP:OD1	2.40	0.54
1:A:202:ARG:HD3	1:A:244:TRP:CE3	2.43	0.53
1:G:232:GLU:OE2	2:H:6:LYS:NZ	2.36	0.53
3:F:9:GLY:O	3:F:10:PHE:CB	2.56	0.53
1:G:219:LEU:O	1:G:220:ASN:HB2	2.09	0.53
2:H:81:ARG:HG2	7:H:427:HOH:O	2.09	0.53
3:I:10:PHE:CD1	3:I:10:PHE:O	2.62	0.53
1:G:202:ARG:HD2	1:G:244:TRP:CD2	2.44	0.52
1:D:80:THR:HG21	3:F:10:PHE:HB2	1.92	0.52
3:C:6:GLY:O	3:C:7:GLU:HG3	2.09	0.52
3:I:3:LEU:HD22	3:I:4:VAL:H	1.75	0.51
1:D:202:ARG:HD3	1:D:244:TRP:CE3	2.45	0.51
1:A:202:ARG:HD2	1:A:244:TRP:CD2	2.45	0.51
2:E:99:MET:OXT	5:E:101:GOL:H32	2.10	0.51
1:A:84:TYR:CD1	1:A:84:TYR:O	2.64	0.51
2:H:96:ASP:H	5:H:302:GOL:H12	1.75	0.50
1:A:118:TYR:HD2	5:A:304:GOL:C3	2.25	0.50
1:A:107:TRP:CZ3	1:A:169:ARG:HG2	2.47	0.50
3:F:7:GLU:H	3:F:8:ARG:NH2	2.10	0.50
1:D:80:THR:HG21	3:F:10:PHE:CG	2.47	0.50
1:D:82:GLN:HG3	7:D:479:HOH:O	2.12	0.50
1:G:66:ARG:CZ	3:I:3:LEU:HD23	2.42	0.49
1:G:123:TYR:OH	1:G:143:THR:CG2	2.60	0.49
3:F:3:LEU:HD12	3:F:4:VAL:H	1.76	0.49
1:D:217:TRP:HD1	1:D:228:MET:CE	2.25	0.49
1:A:123:TYR:OH	1:A:143:THR:CG2	2.60	0.49
3:F:7:GLU:HG3	3:F:8:ARG:N	2.27	0.49
1:D:217:TRP:CD1	1:D:228:MET:CE	2.96	0.49
2:H:96:ASP:H	5:H:302:GOL:C1	2.26	0.49
1:D:252:GLY:O	1:G:89:LYS:HE3	2.14	0.48
1:G:258:THR:HG22	1:G:273:ARG:HG3	1.96	0.48
1:G:107:TRP:CH2	1:G:169:ARG:HG2	2.49	0.48
1:D:107:TRP:CZ3	1:D:169:ARG:HG2	2.48	0.48
1:A:61:GLU:OE2	1:A:65:GLN:NE2	2.47	0.48
1:D:217:TRP:CD1	1:D:228:MET:HE2	2.50	0.47
3:C:5:CYS:CB	3:C:8:ARG:HD3	2.44	0.47

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:G:143:THR:HB	3:I:8:ARG:NH2	2.29	0.47
1:G:62:GLU:OE1	1:G:66:ARG:NE	2.48	0.47
1:A:117:ALA:HB2	2:B:60:TRP:CE2	2.50	0.47
1:G:202:ARG:HD3	1:G:244:TRP:CE3	2.50	0.47
3:I:8:ARG:HG3	3:I:8:ARG:O	2.15	0.47
1:G:185:PRO:HD3	1:G:263:HIS:CD2	2.50	0.46
1:D:232:GLU:OE2	2:E:6:LYS:HE2	2.16	0.46
1:G:46:GLU:HB2	1:G:47:PRO:HD2	1.98	0.46
3:I:8:ARG:O	3:I:8:ARG:CG	2.64	0.46
1:D:228:MET:HE2	1:D:245:ALA:HB1	1.98	0.46
3:C:5:CYS:HB2	3:C:8:ARG:HH11	1.70	0.45
3:C:5:CYS:HA	3:C:8:ARG:HD2	1.98	0.45
1:A:185:PRO:HD3	1:A:263:HIS:CD2	2.52	0.45
3:C:5:CYS:HB2	3:C:8:ARG:HD3	1.99	0.45
2:H:32:PRO:O	2:H:84:HIS:HE1	2.00	0.45
1:A:20:PRO:HG2	1:A:75:ARG:HG3	1.99	0.44
1:G:155:TYR:CE2	3:I:3:LEU:HD11	2.53	0.44
1:D:117:ALA:HB2	2:E:60:TRP:CE2	2.52	0.44
1:A:143:THR:HG22	3:C:10:PHE:CB	2.40	0.44
1:D:66:ARG:CA	3:F:4:VAL:HG11	2.49	0.43
1:A:119:ASP:HB3	2:B:0:MET:HA	2.00	0.43
1:D:152:ASP:OD2	3:F:8:ARG:NH1	2.51	0.43
1:A:123:TYR:OH	1:A:143:THR:HG21	2.17	0.43
2:H:34:ASP:HB2	7:H:441:HOH:O	2.18	0.43
3:F:7:GLU:H	3:F:8:ARG:HH21	1.67	0.43
1:D:97:ARG:HH21	1:D:114:HIS:CD2	2.37	0.42
1:G:123:TYR:OH	1:G:143:THR:HG21	2.18	0.42
1:A:154:GLU:HG3	7:A:497:HOH:O	2.20	0.42
1:A:56:GLY:H	4:A:302:EDO:H21	1.85	0.42
1:A:69:SER:OG	3:C:4:VAL:HG21	2.19	0.42
2:E:32:PRO:O	2:E:84:HIS:HE1	2.02	0.42
1:A:258:THR:HG22	1:A:273:ARG:HG3	2.01	0.42
1:D:258:THR:HG22	1:D:273:ARG:HG3	2.01	0.42
1:D:143:THR:HA	3:F:9:GLY:HA2	2.02	0.42
1:D:46:GLU:HB2	1:D:47:PRO:HD2	2.01	0.42
1:A:62:GLU:OE2	1:A:66:ARG:CD	2.68	0.42
2:B:32:PRO:O	2:B:84:HIS:HE1	2.02	0.41
1:A:202:ARG:NH2	7:A:401:HOH:O	2.34	0.41
3:F:10:PHE:C	3:F:10:PHE:CD1	2.86	0.41
3:F:8:ARG:NH2	3:F:8:ARG:HB2	2.36	0.41
1:A:62:GLU:OE2	1:A:66:ARG:HD2	2.20	0.41

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:217:TRP:CD1	1:D:228:MET:HE3	2.56	0.41
1:A:146:LYS:HE2	3:C:10:PHE:HD2	1.84	0.41
1:D:202:ARG:CD	1:D:244:TRP:CE3	3.04	0.41
2:B:48:LYS:CE	2:B:48:LYS:HA	2.47	0.41
1:D:196:GLN:NE2	1:D:196:GLN:HA	2.35	0.41
1:D:210:PRO:O	1:D:263:HIS:HE1	2.04	0.41
1:G:202:ARG:NH2	7:G:401:HOH:O	2.32	0.41
1:G:258:THR:HG22	1:G:273:ARG:CG	2.51	0.41
3:C:6:GLY:O	3:C:7:GLU:CG	2.69	0.40
1:D:20:PRO:HD2	1:D:75:ARG:HG3	2.02	0.40
3:F:8:ARG:CZ	3:F:8:ARG:HB2	2.52	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
7:D:435:HOH:O	7:G:431:HOH:O[3_745]	2.04	0.16

## 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	275/277 (99%)	271 (98%)	4 (2%)	0	100	100
1	D	275/277 (99%)	270 (98%)	5 (2%)	0	100	100
1	G	276/277 (100%)	270 (98%)	6 (2%)	0	100	100
2	B	98/100 (98%)	97 (99%)	1 (1%)	0	100	100
2	E	99/100 (99%)	96 (97%)	3 (3%)	0	100	100
2	H	98/100 (98%)	97 (99%)	1 (1%)	0	100	100
3	C	8/10 (80%)	6 (75%)	1 (12%)	1 (12%)	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
3	F	8/10 (80%)	4 (50%)	2 (25%)	2 (25%)	0 0
3	I	8/10 (80%)	5 (62%)	2 (25%)	1 (12%)	0 0
All	All	1145/1161 (99%)	1116 (98%)	25 (2%)	4 (0%)	46 57

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	F	8	ARG
3	I	8	ARG
3	C	7	GLU
3	F	9	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	236/236 (100%)	214 (91%)	22 (9%)	11 13
1	D	236/236 (100%)	219 (93%)	17 (7%)	18 22
1	G	237/236 (100%)	217 (92%)	20 (8%)	14 16
2	B	95/95 (100%)	88 (93%)	7 (7%)	17 21
2	E	96/95 (101%)	87 (91%)	9 (9%)	11 12
2	H	95/95 (100%)	90 (95%)	5 (5%)	28 37
3	C	8/8 (100%)	4 (50%)	4 (50%)	0 0
3	F	8/8 (100%)	5 (62%)	3 (38%)	0 0
3	I	8/8 (100%)	6 (75%)	2 (25%)	1 0
All	All	1019/1017 (100%)	930 (91%)	89 (9%)	12 15

All (89) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	0	MET
1	A	6	ARG

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	25	VAL
1	A	35	ARG
1	A	41	ASP
1	A	65	GLN
1	A	70	ASP
1	A	83	ARG
1	A	95	PHE
1	A	105	SER
1	A	110	LEU
1	A	111	ARG
1	A	138	THR
1	A	143	THR
1	A	149	GLN
1	A	176	ASN
1	A	177	GLU
1	A	182	THR
1	A	251	LEU
1	A	255	GLN
1	A	272	LEU
1	A	275	LYS
2	B	39	LEU
2	B	47	GLU
2	B	48	LYS
2	B	58	LYS
2	B	70	PHE
2	B	73	THR
2	B	89	GLN
3	C	3	LEU
3	C	5	CYS
3	C	8	ARG
3	C	10	PHE
1	D	6	ARG
1	D	35	ARG
1	D	66	ARG
1	D	69	SER
1	D	70	ASP
1	D	95	PHE
1	D	105	SER
1	D	110	LEU
1	D	111	ARG
1	D	138	THR
1	D	149	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	D	177	GLU
1	D	182	THR
1	D	196	GLN
1	D	224	LEU
1	D	251	LEU
1	D	272	LEU
2	E	0	MET
2	E	1	ILE
2	E	39	LEU
2	E	44	GLU
2	E	47	GLU
2	E	48	LYS
2	E	69	GLU
2	E	70	PHE
2	E	89	GLN
3	F	4	VAL
3	F	8	ARG
3	F	10	PHE
1	G	0	MET
1	G	6	ARG
1	G	35	ARG
1	G	41	ASP
1	G	70	ASP
1	G	82	GLN
1	G	95	PHE
1	G	105	SER
1	G	110	LEU
1	G	111	ARG
1	G	138	THR
1	G	143	THR
1	G	149	GLN
1	G	176	ASN
1	G	177	GLU
1	G	182	THR
1	G	224	LEU
1	G	251	LEU
1	G	273	ARG
1	G	275	LYS
2	H	39	LEU
2	H	58	LYS
2	H	70	PHE
2	H	73	THR

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	H	89	GLN
3	I	4	VAL
3	I	10	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (23) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	82	GLN
1	A	86	ASN
1	A	87	GLN
1	A	93	HIS
1	A	114	HIS
1	A	149	GLN
1	A	196	GLN
1	A	226	GLN
1	A	263	HIS
2	B	84	HIS
1	D	82	GLN
1	D	86	ASN
1	D	114	HIS
1	D	149	GLN
1	D	196	GLN
1	D	226	GLN
1	D	263	HIS
2	E	84	HIS
1	G	65	GLN
1	G	86	ASN
1	G	149	GLN
1	G	263	HIS
2	H	84	HIS

### 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

20 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	EDO	A	301	-	3,3,3	0.47	0	2,2,2	0.57	0
4	EDO	A	302	-	3,3,3	1.21	0	2,2,2	1.16	0
4	EDO	A	303	-	3,3,3	0.33	0	2,2,2	0.49	0
5	GOL	A	304	-	5,5,5	0.70	0	5,5,5	1.98	2 (40%)
5	GOL	A	305	-	5,5,5	0.76	0	5,5,5	1.54	1 (20%)
6	SO4	A	306	-	4,4,4	0.78	0	6,6,6	0.82	0
6	SO4	A	307	-	4,4,4	0.59	0	6,6,6	0.38	0
6	SO4	A	308	-	4,4,4	0.52	0	6,6,6	0.47	0
6	SO4	A	309	-	4,4,4	0.59	0	6,6,6	0.45	0
4	EDO	D	301	-	3,3,3	0.50	0	2,2,2	0.30	0
5	GOL	D	302	-	5,5,5	0.69	0	5,5,5	1.30	1 (20%)
6	SO4	D	303	-	4,4,4	0.64	0	6,6,6	0.93	0
6	SO4	D	304	-	4,4,4	0.71	0	6,6,6	0.54	0
6	SO4	D	305	-	4,4,4	0.51	0	6,6,6	0.45	0
6	SO4	D	306	-	4,4,4	0.76	0	6,6,6	0.50	0
5	GOL	E	101	-	5,5,5	0.93	0	5,5,5	1.27	1 (20%)
6	SO4	G	301	-	4,4,4	0.70	0	6,6,6	0.91	0
4	EDO	H	301	-	3,3,3	0.22	0	2,2,2	1.09	0
5	GOL	H	302	-	5,5,5	0.78	0	5,5,5	0.82	0
5	GOL	H	303	-	5,5,5	1.29	0	5,5,5	1.38	1 (20%)

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
5	GOL	A	304	-	-	0/4/4/4	0/0/0/0
5	GOL	A	305	-	-	0/4/4/4	0/0/0/0
6	SO4	A	306	-	-	0/0/0/0	0/0/0/0
6	SO4	A	307	-	-	0/0/0/0	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	D	301	-	-	0/1/1/1	0/0/0/0
5	GOL	D	302	-	-	0/4/4/4	0/0/0/0
6	SO4	D	303	-	-	0/0/0/0	0/0/0/0
6	SO4	D	304	-	-	0/0/0/0	0/0/0/0
6	SO4	D	305	-	-	0/0/0/0	0/0/0/0
6	SO4	D	306	-	-	0/0/0/0	0/0/0/0
5	GOL	E	101	-	-	0/4/4/4	0/0/0/0
6	SO4	G	301	-	-	0/0/0/0	0/0/0/0
4	EDO	H	301	-	-	0/1/1/1	0/0/0/0
5	GOL	H	302	-	-	0/4/4/4	0/0/0/0
5	GOL	H	303	-	-	0/4/4/4	0/0/0/0

There are no bond length outliers.

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	D	302	GOL	O3-C3-C2	2.02	120.21	109.97
5	A	305	GOL	O1-C1-C2	2.23	121.27	109.97
5	H	303	GOL	O1-C1-C2	2.28	121.51	109.97
5	E	101	GOL	O3-C3-C2	2.44	122.34	109.97
5	A	304	GOL	C3-C2-C1	2.51	121.64	111.06
5	A	304	GOL	O3-C3-C2	2.86	124.48	109.97

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	302	EDO	1	0
5	A	304	GOL	2	0
5	A	305	GOL	2	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	E	101	GOL	1	0
6	G	301	SO4	1	0
5	H	302	GOL	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<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%)	-0.21	3 (1%) 82 86	17, 32, 55, 83	0
1	D	277/277 (100%)	-0.27	3 (1%) 82 86	18, 36, 62, 93	0
1	G	277/277 (100%)	-0.13	7 (2%) 61 70	23, 41, 72, 97	0
2	B	100/100 (100%)	-0.36	0 100 100	19, 35, 58, 68	0
2	E	100/100 (100%)	-0.16	1 (1%) 84 88	21, 44, 78, 95	0
2	H	100/100 (100%)	-0.38	0 100 100	23, 37, 56, 70	0
3	C	10/10 (100%)	4.90	7 (70%) 0 0	32, 45, 62, 64	7 (70%)
3	F	10/10 (100%)	5.93	7 (70%) 0 0	33, 50, 63, 69	7 (70%)
3	I	10/10 (100%)	4.41	7 (70%) 0 0	44, 58, 66, 70	7 (70%)
All	All	1161/1161 (100%)	-0.09	35 (3%) 54 63	17, 38, 68, 97	21 (1%)

All (35) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	F	6	GLY	14.3
3	F	5	CYS	11.6
3	F	4	VAL	9.7
3	C	6	GLY	9.7
3	F	10	PHE	8.7
3	C	5	CYS	8.6
3	C	10	PHE	8.4
3	I	9	GLY	8.4
3	I	5	CYS	7.6
3	I	4	VAL	7.0
3	I	10	PHE	7.0
3	C	4	VAL	6.7
3	C	9	GLY	6.4
3	F	7	GLU	6.3
3	F	9	GLY	5.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
3	C	8	ARG	5.1
3	C	7	GLU	4.9
3	F	8	ARG	4.6
3	I	8	ARG	4.3
3	I	6	GLY	4.3
1	D	41	ASP	4.1
3	I	7	GLU	4.0
1	G	88	SER	3.5
2	E	99	MET	2.9
1	A	82	GLN	2.9
1	A	86	ASN	2.8
1	D	268	GLU	2.6
1	G	220	ASN	2.4
1	G	223	ASP	2.3
1	G	145	ARG	2.3
1	A	83	ARG	2.3
1	G	90	GLY	2.2
1	G	41	ASP	2.2
1	D	276	PRO	2.2
1	G	276	PRO	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 [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, 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
6	SO4	A	308	5/5	0.71	0.63	49.37	84,94,108,113	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
6	SO4	D	306	5/5	0.76	0.38	15.63	71,88,103,110	0
6	SO4	A	306	5/5	0.91	0.29	15.57	51,57,79,83	0
5	GOL	D	302	6/6	0.81	0.25	11.82	46,51,57,59	0
6	SO4	A	307	5/5	0.87	0.42	10.80	78,90,101,104	0
5	GOL	H	302	6/6	0.81	0.40	10.05	32,43,48,60	0
6	SO4	D	305	5/5	0.93	0.32	8.13	75,82,93,101	0
5	GOL	A	304	6/6	0.88	0.19	6.99	34,38,40,47	0
6	SO4	D	303	5/5	0.96	0.19	6.20	48,57,72,75	0
5	GOL	H	303	6/6	0.26	0.32	4.98	53,58,69,69	0
6	SO4	G	301	5/5	0.85	0.37	4.98	65,74,99,112	0
4	EDO	A	302	4/4	0.76	0.23	4.69	39,51,52,53	0
4	EDO	H	301	4/4	0.95	0.18	3.00	35,36,38,47	0
5	GOL	A	305	6/6	0.87	0.34	2.55	38,47,50,50	0
4	EDO	A	303	4/4	0.96	0.16	1.06	39,39,41,42	0
4	EDO	A	301	4/4	0.95	0.14	0.77	46,46,48,48	0
4	EDO	D	301	4/4	0.89	0.15	0.73	52,55,55,58	0
5	GOL	E	101	6/6	0.82	0.16	0.59	44,52,56,59	0
6	SO4	D	304	5/5	0.79	0.42	-	70,85,91,94	0
6	SO4	A	309	5/5	0.86	0.39	-	87,90,97,107	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.