



# Full wwPDB X-ray Structure Validation Report ⓘ

May 13, 2024 – 04:09 pm BST

PDB ID : 8QRE  
Title : Cholera holotoxin (wildtype)  
Authors : Mojica, N.; Cordara, G.; Krenzel, U.  
Deposited on : 2023-10-06  
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

<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 : 1.8.4, CSD as541be (2020)  
Xtrriage (Phenix) : 1.13  
EDS : 2.36.2  
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) : 2.36.2

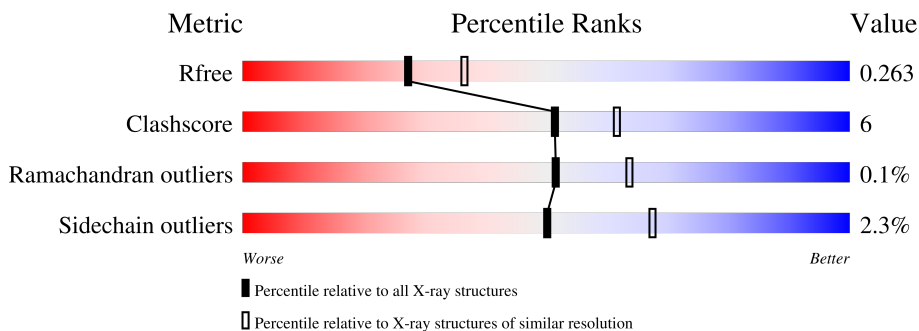
# 1 Overall quality at a glance

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))
$R_{free}$	130704	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (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 of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	A	240	85% 14% .
1	B	240	76% 19% . .
2	C	103	85% 15%
2	D	103	88% 11% .
2	E	103	91% 9%
2	F	103	83% 17%
2	G	103	91% 6% .

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	H	103	 84% 16%
2	I	103	 93% 6%
2	J	103	 90% 10%
2	K	103	 91% 7%
2	L	103	 93% 7%

## 2 Entry composition [i](#)

There are 9 unique types of molecules in this entry. The entry contains 13002 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 Cholera enterotoxin subunit A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	237	Total 1989	C 1236	N 369	O 378	S 6	0	10	0
1	B	231	Total 1947	C 1212	N 355	O 374	S 6	0	11	0

- Molecule 2 is a protein called Cholera enterotoxin subunit B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	D	103	Total 853	C 532	N 150	O 164	S 7	0	5	0
2	E	103	Total 851	C 533	N 149	O 162	S 7	0	5	0
2	F	103	Total 859	C 537	N 150	O 165	S 7	0	6	0
2	G	103	Total 837	C 523	N 148	O 159	S 7	0	3	0
2	H	103	Total 859	C 539	N 150	O 163	S 7	0	6	0
2	C	103	Total 844	C 529	N 148	O 160	S 7	0	4	0
2	I	103	Total 877	C 550	N 156	O 164	S 7	0	7	0
2	J	103	Total 826	C 517	N 144	O 158	S 7	0	2	0
2	K	103	Total 834	C 522	N 145	O 159	S 8	0	3	0
2	L	103	Total 833	C 521	N 145	O 160	S 7	0	3	0

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	18	HIS	TYR	engineered mutation	UNP P01556

*Continued on next page...*

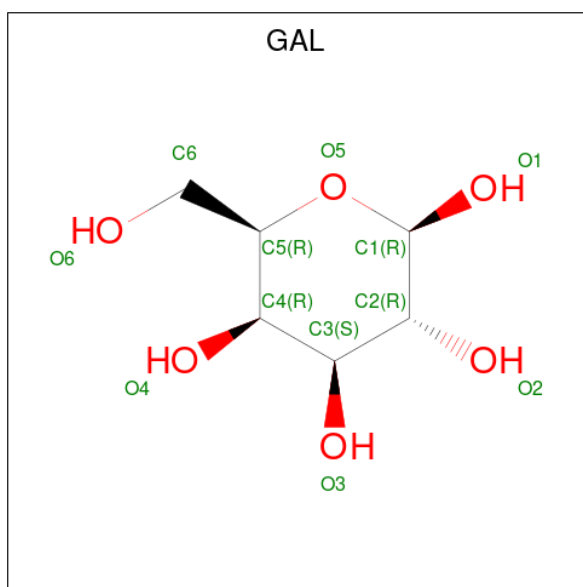
*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	47	THR	ILE	engineered mutation	UNP P01556
E	18	HIS	TYR	engineered mutation	UNP P01556
E	47	THR	ILE	engineered mutation	UNP P01556
F	18	HIS	TYR	engineered mutation	UNP P01556
F	47	THR	ILE	engineered mutation	UNP P01556
G	18	HIS	TYR	engineered mutation	UNP P01556
G	47	THR	ILE	engineered mutation	UNP P01556
H	18	HIS	TYR	engineered mutation	UNP P01556
H	47	THR	ILE	engineered mutation	UNP P01556
C	18	HIS	TYR	engineered mutation	UNP P01556
C	47	THR	ILE	engineered mutation	UNP P01556
I	18	HIS	TYR	engineered mutation	UNP P01556
I	47	THR	ILE	engineered mutation	UNP P01556
J	18	HIS	TYR	engineered mutation	UNP P01556
J	47	THR	ILE	engineered mutation	UNP P01556
K	18	HIS	TYR	engineered mutation	UNP P01556
K	47	THR	ILE	engineered mutation	UNP P01556
L	18	HIS	TYR	engineered mutation	UNP P01556
L	47	THR	ILE	engineered mutation	UNP P01556

- Molecule 3 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total Na 1 1	0	0
3	D	1	Total Na 1 1	0	0
3	B	2	Total Na 2 2	0	0

- Molecule 4 is beta-D-galactopyranose (three-letter code: GAL) (formula: C<sub>6</sub>H<sub>12</sub>O<sub>6</sub>).



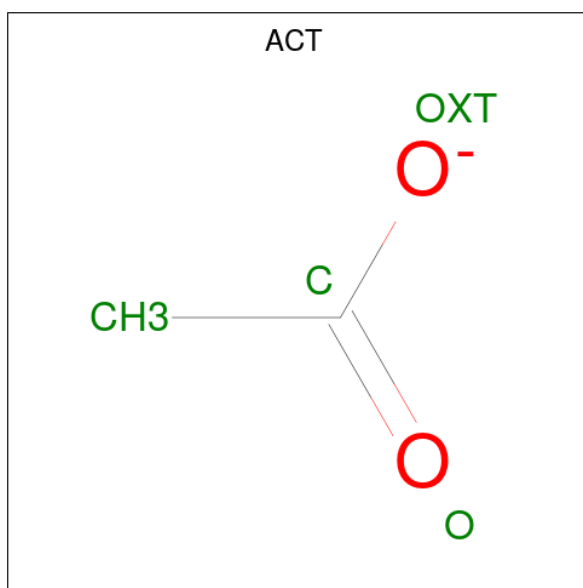
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O 12 6 6	0	0
4	D	1	Total C O 12 6 6	0	0
4	E	1	Total C O 12 6 6	0	0
4	F	1	Total C O 12 6 6	0	0
4	G	1	Total C O 12 6 6	0	0
4	H	1	Total C O 12 6 6	0	0
4	I	1	Total C O 12 6 6	0	0
4	J	1	Total C O 12 6 6	0	0
4	K	1	Total C O 12 6 6	0	0
4	L	1	Total C O 12 6 6	0	0

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



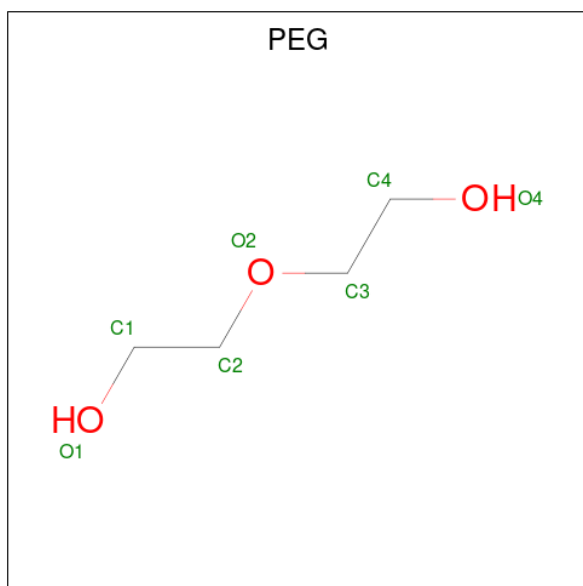
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O 6 3 3	0	0
5	A	1	Total C O 6 3 3	0	0
5	A	1	Total C O 6 3 3	0	0
5	D	1	Total C O 6 3 3	0	0
5	E	1	Total C O 6 3 3	0	0
5	G	1	Total C O 6 3 3	0	0
5	H	1	Total C O 6 3 3	0	0
5	B	1	Total C O 6 3 3	0	0
5	C	1	Total C O 6 3 3	0	0
5	I	1	Total C O 6 3 3	0	0
5	I	1	Total C O 6 3 3	0	0
5	J	1	Total C O 6 3 3	0	0
5	L	1	Total C O 6 3 3	0	0
5	L	1	Total C O 6 3 3	0	0

- Molecule 6 is ACETATE ION (three-letter code: ACT) (formula:  $C_2H_3O_2$ ).



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

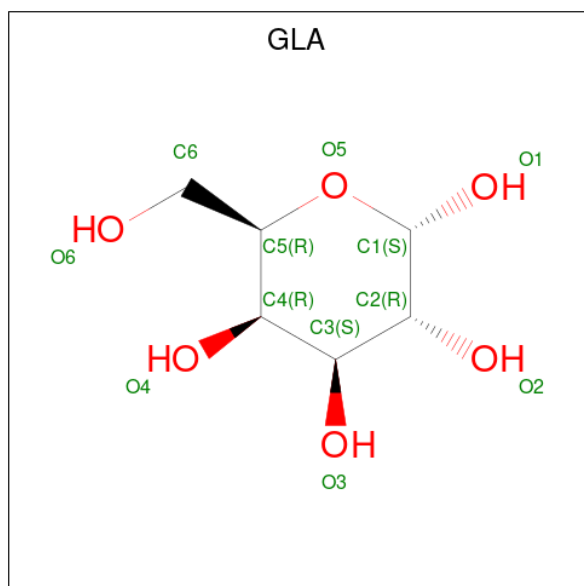
- Molecule 7 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula:  $C_4H_{10}O_3$ ).





Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	G	1	Total	C	O	0	0
			7	4	3		

- Molecule 8 is alpha-D-galactopyranose (three-letter code: GLA) (formula: C<sub>6</sub>H<sub>12</sub>O<sub>6</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	C	1	Total	C	O	0	0
			12	6	6		

- Molecule 9 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	46	Total	O	0	0
			46	46		
9	D	35	Total	O	0	0
			35	35		
9	E	30	Total	O	0	0
			30	30		
9	F	31	Total	O	0	0
			31	31		
9	G	36	Total	O	0	0
			36	36		
9	H	28	Total	O	0	0
			28	28		
9	B	37	Total	O	0	0
			37	37		
9	C	28	Total	O	0	0
			28	28		

*Continued on next page...*

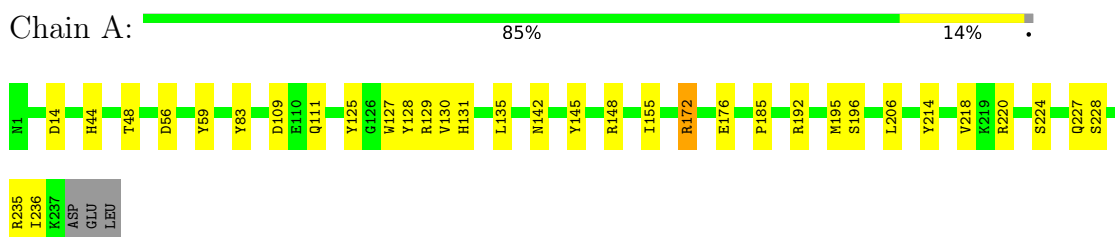
*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
9	I	16	Total 16	O 16	0	0
9	J	23	Total 23	O 23	0	0
9	K	16	Total 16	O 16	0	0
9	L	28	Total 28	O 28	0	0

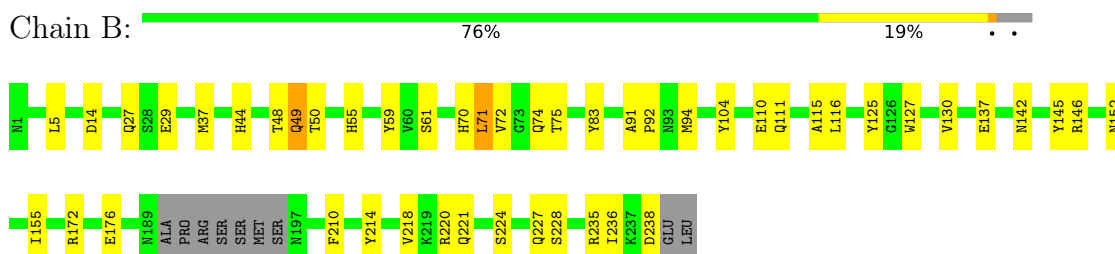
### 3 Residue-property plots [i](#)

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. 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. 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.

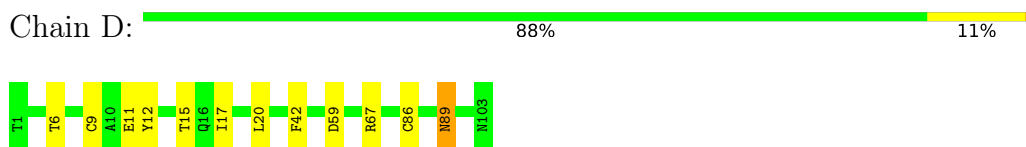
- Molecule 1: Cholera enterotoxin subunit A



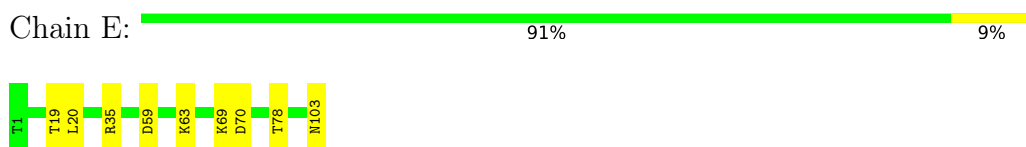
- Molecule 1: Cholera enterotoxin subunit A



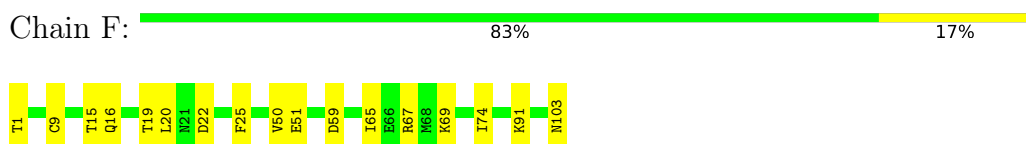
- Molecule 2: Cholera enterotoxin subunit B



- Molecule 2: Cholera enterotoxin subunit B



- Molecule 2: Cholera enterotoxin subunit B




- Molecule 2: Cholera enterotoxin subunit B

Chain G:  91% 6%




- Molecule 2: Cholera enterotoxin subunit B

Chain H:  84% 16%




- Molecule 2: Cholera enterotoxin subunit B

Chain C:  85% 15%




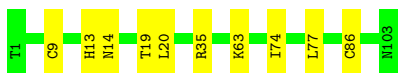
- Molecule 2: Cholera enterotoxin subunit B

Chain I:  93% 6%



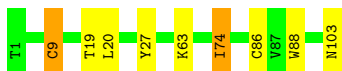
- Molecule 2: Cholera enterotoxin subunit B

Chain J:  90% 10%



- Molecule 2: Cholera enterotoxin subunit B

Chain K:  91% 7%



- Molecule 2: Cholera enterotoxin subunit B

Chain L:  93% 7%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	60.66Å 108.13Å 124.77Å 90.00° 95.96° 90.00°	Depositor
Resolution (Å)	124.09 – 2.30 124.09 – 1.50	Depositor EDS
% Data completeness (in resolution range)	98.6 (124.09-2.30) 59.3 (124.09-1.50)	Depositor EDS
$R_{merge}$	0.26	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	0.98 (at 1.50Å)	Xtrriage
Refinement program	REFMAC 5.8.0419	Depositor
R, $R_{free}$	0.216 , 0.263 0.220 , 0.263	Depositor DCC
$R_{free}$ test set	7803 reflections (5.13%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	16.7	Xtrriage
Anisotropy	0.185	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.40 , 50.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	13002	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	44.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.58% 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, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: ACT, GOL, GAL, GLA, PEG, NA

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	0.43	0/2044	0.48	0/2768
1	B	0.40	0/1999	0.47	0/2709
2	C	0.39	0/858	0.48	0/1156
2	D	0.39	0/867	0.49	0/1170
2	E	0.40	0/865	0.48	0/1166
2	F	0.39	0/873	0.52	0/1177
2	G	0.40	0/851	0.49	0/1148
2	H	0.41	0/873	0.48	0/1177
2	I	0.38	0/891	0.49	0/1199
2	J	0.37	0/840	0.49	0/1134
2	K	0.37	0/848	0.47	0/1144
2	L	0.36	0/847	0.47	0/1144
All	All	0.39	0/12656	0.48	0/17092

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1989	0	1861	24	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	1947	0	1809	40	0
2	C	844	0	849	9	0
2	D	853	0	848	15	0
2	E	851	0	853	11	0
2	F	859	0	856	17	0
2	G	837	0	837	11	0
2	H	859	0	866	13	0
2	I	877	0	890	13	0
2	J	826	0	825	9	0
2	K	834	0	833	10	0
2	L	833	0	831	7	0
3	A	1	0	0	0	0
3	B	2	0	0	0	0
3	D	1	0	0	0	0
4	A	12	0	12	1	0
4	D	12	0	12	0	0
4	E	12	0	12	0	0
4	F	12	0	12	0	0
4	G	12	0	11	0	0
4	H	12	0	12	0	0
4	I	12	0	12	0	0
4	J	12	0	12	0	0
4	K	12	0	12	1	0
4	L	12	0	12	0	0
5	A	18	0	24	0	0
5	B	6	0	8	0	0
5	C	6	0	8	0	0
5	D	6	0	8	0	0
5	E	6	0	8	0	0
5	G	6	0	8	0	0
5	H	6	0	8	1	0
5	I	12	0	16	0	0
5	J	6	0	8	0	0
5	L	12	0	16	0	0
6	C	4	0	3	0	0
6	F	4	0	3	0	0
6	G	4	0	3	0	0
7	G	7	0	10	0	0
8	C	12	0	12	0	0
9	A	46	0	0	2	0
9	B	37	0	0	3	0
9	C	28	0	0	2	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
9	D	35	0	0	0	0
9	E	30	0	0	0	0
9	F	31	0	0	2	0
9	G	36	0	0	0	0
9	H	28	0	0	1	0
9	I	16	0	0	1	0
9	J	23	0	0	1	0
9	K	16	0	0	2	0
9	L	28	0	0	1	0
All	All	13002	0	12420	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 6.

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)
2:D:9[B]:CYS:SG	2:D:15[B]:THR:CG2	2.22	1.28
2:F:9[B]:CYS:SG	2:F:15[B]:THR:CG2	2.26	1.24
1:B:146[B]:ARG:HH11	1:B:146[B]:ARG:HG2	1.01	1.17
2:F:9[B]:CYS:SG	2:F:15[B]:THR:HG23	1.85	1.12
1:B:227:GLN:OE1	2:I:74[A]:ILE:HG22	1.53	1.08
2:F:51:GLU:OE2	2:F:91:LYS:HE2	1.56	1.04
2:D:9[B]:CYS:SG	2:D:15[B]:THR:HG22	1.95	1.03
2:F:9[B]:CYS:SG	2:F:15[B]:THR:HG21	2.03	0.98
1:B:146[B]:ARG:HG2	1:B:146[B]:ARG:NH1	1.78	0.92
2:D:6:THR:HA	2:D:17:ILE:HD11	1.56	0.86
2:D:9[B]:CYS:SG	2:D:15[B]:THR:HG23	2.17	0.83
2:D:9[B]:CYS:SG	2:D:15[B]:THR:HG21	2.19	0.82
2:K:74:ILE:HD13	2:L:77:LEU:HD13	1.62	0.81
2:G:74:ILE:HD13	2:H:77:LEU:HD13	1.64	0.78
2:I:81[B]:LYS:HD2	2:I:103:ASN:ND2	2.00	0.76
1:A:235:ARG:NH2	2:E:70:ASP:OD2	2.17	0.76
2:H:103:ASN:ND2	9:H:301:HOH:O	2.17	0.75
2:F:65:ILE:O	2:F:69[B]:LYS:HD3	1.88	0.74
2:D:6:THR:HA	2:D:17:ILE:CD1	2.17	0.74
1:B:224:SER:O	2:J:74:ILE:HD11	1.89	0.72
1:B:49[B]:GLN:O	1:B:50:THR:OG1	2.05	0.72
2:I:103:ASN:O	9:I:301:HOH:O	2.10	0.68
1:B:49[B]:GLN:C	1:B:50:THR:HG1	1.97	0.68
1:B:228:SER:CB	2:J:74:ILE:HG12	2.24	0.67

Continued on next page...



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:224:SER:O	2:F:74:ILE:HD11	1.94	0.67
2:D:11[B]:GLU:OE1	2:E:35:ARG:CZ	2.42	0.67
1:B:146[B]:ARG:HH11	1:B:146[B]:ARG:CG	1.90	0.66
2:L:22:ASP:OD2	9:L:301:HOH:O	2.14	0.66
1:B:94[A]:MET:SD	1:B:115:ALA:HB2	2.36	0.65
2:I:74[A]:ILE:HD12	2:I:75:ALA:N	2.11	0.65
2:C:2:PRO:HG3	2:C:8:LEU:HD23	1.80	0.64
1:B:228:SER:HB3	2:J:74:ILE:HG12	1.80	0.63
2:K:27:TYR:OH	9:K:301:HOH:O	2.14	0.63
2:C:15:THR:OG1	9:C:301:HOH:O	2.14	0.63
2:J:13:HIS:CD2	2:J:14:ASN:OD1	2.52	0.63
1:B:227:GLN:CD	2:I:74[A]:ILE:HG22	2.19	0.63
2:F:20:LEU:HB3	2:F:22[B]:ASP:OD1	1.99	0.62
2:G:103:ASN:O	2:G:103:ASN:ND2	2.33	0.61
2:E:103:ASN:HD22	2:F:25:PHE:HA	1.65	0.60
1:B:48[A]:THR:O	1:B:49[A]:GLN:HB2	2.02	0.60
2:D:11[B]:GLU:OE1	2:E:35:ARG:NE	2.35	0.59
1:B:49[A]:GLN:O	1:B:50:THR:OG1	2.13	0.59
1:B:236:ILE:HD13	2:I:63:LYS:HB2	1.84	0.59
2:J:63:LYS:HG2	9:J:306:HOH:O	2.02	0.58
1:B:48[B]:THR:HG22	1:B:48[B]:THR:O	2.01	0.58
1:B:70:HIS:O	1:B:74:GLN:HG2	2.03	0.58
1:B:137:GLU:HG3	1:B:137:GLU:O	2.03	0.58
1:B:220:ARG:HA	2:I:78:THR:HB	1.85	0.58
1:A:185:PRO:HB3	2:I:45:GLY:HA3	1.85	0.57
1:A:235:ARG:HD3	4:A:302:GAL:H62	1.84	0.57
1:B:146[B]:ARG:NH1	1:B:221:GLN:OE1	2.37	0.57
1:A:111[A]:GLN:O	1:A:111[A]:GLN:HG3	2.05	0.56
2:I:81[B]:LYS:HD2	2:I:103:ASN:HD21	1.69	0.56
2:I:74[A]:ILE:HD13	2:J:77:LEU:CD1	2.36	0.55
1:A:125:TYR:HD1	1:A:145:TYR:CE2	2.25	0.55
1:A:220:ARG:HA	2:E:78:THR:HB	1.89	0.55
2:K:9[B]:CYS:HB2	2:K:86[B]:CYS:SG	2.47	0.55
2:H:47:THR:H	5:H:202:GOL:H12	1.72	0.54
1:A:228:SER:CB	2:F:74:ILE:HG12	2.37	0.54
1:B:104:TYR:OH	1:B:176:GLU:OE2	2.16	0.53
2:F:51:GLU:OE2	2:F:91:LYS:CE	2.45	0.53
2:G:9[B]:CYS:HB2	2:G:86[B]:CYS:SG	2.49	0.53
2:C:67:ARG:NH1	9:C:302:HOH:O	2.32	0.52
2:H:20:LEU:HD13	2:H:42:PHE:CZ	2.45	0.52
1:A:228:SER:HB3	2:F:74:ILE:HG12	1.92	0.51

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:I:74[A]:ILE:HD13	2:J:77:LEU:HD11	1.93	0.51
2:C:20:LEU:HD13	2:C:42:PHE:CZ	2.46	0.51
1:A:228:SER:OG	2:G:74:ILE:CG1	2.59	0.50
1:A:155:ILE:O	9:A:401:HOH:O	2.19	0.50
2:L:20:LEU:HD13	2:L:42:PHE:CZ	2.46	0.50
2:K:63:LYS:HG2	9:K:303:HOH:O	2.11	0.50
2:D:9[B]:CYS:HB2	2:D:86[B]:CYS:SG	2.52	0.49
2:H:9[A]:CYS:HB2	2:H:86[A]:CYS:SG	2.52	0.49
2:H:20:LEU:HD22	2:H:42:PHE:CG	2.47	0.49
1:B:125:TYR:HD1	1:B:145:TYR:CE2	2.31	0.49
2:H:22:ASP:O	2:H:82:VAL:HG13	2.13	0.49
2:K:74:ILE:HD13	2:L:77:LEU:CD1	2.39	0.49
2:D:20:LEU:HD13	2:D:42:PHE:CZ	2.48	0.49
1:B:155:ILE:O	9:B:402:HOH:O	2.20	0.49
2:G:67[B]:ARG:NH2	2:H:69:LYS:O	2.46	0.49
1:A:236:ILE:HD13	2:E:63[B]:LYS:HB2	1.95	0.48
2:F:67:ARG:NH1	9:F:304:HOH:O	2.34	0.48
2:J:9[A]:CYS:HB2	2:J:86[A]:CYS:SG	2.53	0.48
2:D:12:TYR:OH	2:E:35:ARG:HD2	2.13	0.48
1:B:72:VAL:O	1:B:75:THR:OG1	2.28	0.48
2:G:67[B]:ARG:HH21	2:H:70:ASP:HA	1.79	0.48
2:C:20:LEU:HD22	2:C:42:PHE:CG	2.49	0.47
2:D:20:LEU:HD22	2:D:42:PHE:CG	2.50	0.47
2:C:9[B]:CYS:HB2	2:C:86[B]:CYS:SG	2.55	0.47
1:B:83:TYR:CZ	1:B:130:VAL:HG11	2.50	0.47
1:A:236:ILE:HD13	2:E:63[A]:LYS:HB2	1.97	0.47
1:B:146[B]:ARG:NH1	1:B:146[B]:ARG:CG	2.59	0.47
2:F:50:VAL:HG11	2:F:69[B]:LYS:HD2	1.98	0.46
2:L:20:LEU:HD22	2:L:42:PHE:CG	2.50	0.46
1:A:148:ARG:HD2	2:F:103:ASN:OXT	2.16	0.46
2:G:74:ILE:HD13	2:H:77:LEU:CD1	2.41	0.46
2:F:15[B]:THR:HG23	2:F:16:GLN:N	2.32	0.45
1:B:27:GLN:HG3	1:B:37:MET:SD	2.57	0.45
2:H:8[B]:LEU:HD12	2:H:8[B]:LEU:O	2.17	0.45
2:D:67[B]:ARG:NH2	2:E:69:LYS:O	2.49	0.45
1:B:235[C]:ARG:NH2	9:B:410:HOH:O	2.50	0.45
1:B:48[B]:THR:O	1:B:48[B]:THR:CG2	2.64	0.45
2:K:88:TRP:CH2	4:K:201:GAL:H62	2.52	0.45
2:G:22:ASP:O	2:G:82:VAL:HG13	2.17	0.44
1:A:227:GLN:NE2	9:A:402:HOH:O	2.27	0.44
1:B:146[B]:ARG:NH2	9:B:404:HOH:O	2.37	0.44

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:K:74:ILE:CD1	2:L:77:LEU:HD13	2.42	0.43
1:B:29:GLU:OE2	2:C:23:LYS:NZ	2.52	0.43
1:B:71[A]:LEU:HD13	1:B:71[A]:LEU:HA	1.87	0.43
1:B:127:TRP:CZ3	1:B:142:ASN:HB2	2.52	0.43
1:A:192:ARG:O	1:A:196:SER:HB2	2.17	0.43
1:A:129:ARG:NH1	1:A:131:HIS:HB2	2.33	0.43
1:B:214:TYR:O	1:B:218:VAL:HG23	2.18	0.43
1:A:214:TYR:O	1:A:218:VAL:HG23	2.18	0.43
1:A:83:TYR:CZ	1:A:130:VAL:HG11	2.54	0.42
2:F:1:THR:N	9:F:311:HOH:O	2.52	0.42
1:B:110:GLU:O	1:B:111[B]:GLN:C	2.58	0.42
2:C:25:PHE:CD2	2:C:43[A]:LYS:HG3	2.54	0.42
1:A:127:TRP:CZ3	1:A:142:ASN:HB2	2.54	0.42
2:K:19:THR:C	2:K:20:LEU:HD22	2.40	0.42
1:B:94[A]:MET:CE	1:B:115:ALA:HB2	2.50	0.42
2:J:19:THR:C	2:J:20:LEU:HD22	2.40	0.41
1:A:128:TYR:CB	1:A:135:LEU:HD11	2.50	0.41
2:D:67[B]:ARG:CZ	2:E:69:LYS:HB3	2.51	0.41
2:I:74[A]:ILE:HD12	2:I:74[A]:ILE:C	2.41	0.41
2:G:19:THR:C	2:G:20:LEU:HD22	2.40	0.41
2:H:33:GLY:O	2:H:34[B]:LYS:HB2	2.21	0.41
2:G:9[B]:CYS:CB	2:G:86[B]:CYS:HG	2.33	0.41
2:E:19:THR:C	2:E:20:LEU:HD22	2.41	0.41
2:D:89:ASN:HD22	2:D:89:ASN:C	2.24	0.41
1:A:172:ARG:HD2	1:A:176:GLU:OE1	2.21	0.41
2:F:19:THR:C	2:F:20:LEU:HD22	2.41	0.41
2:G:67[B]:ARG:NH1	2:H:29:GLU:OE1	2.53	0.40
1:B:5:LEU:HD12	1:B:94[B]:MET:CE	2.51	0.40
1:B:228:SER:OG	2:K:74:ILE:CG1	2.69	0.40
1:A:44:HIS:CG	1:A:59:TYR:HB2	2.57	0.40
1:A:185:PRO:HB3	2:I:45:GLY:CA	2.51	0.40
1:B:116:LEU:HD13	1:B:210:PHE:CD1	2.56	0.40
1:B:44:HIS:CG	1:B:59:TYR:HB2	2.56	0.40
1:B:91:ALA:HB1	1:B:92:PRO:HD2	2.04	0.40
2:C:92:THR:HA	2:C:93:PRO:HA	1.95	0.40
2:K:103:ASN:HB3	2:L:25:PHE:CE1	2.56	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	245/240 (102%)	235 (96%)	10 (4%)	0	100	100
1	B	238/240 (99%)	227 (95%)	9 (4%)	2 (1%)	19	23
2	C	105/103 (102%)	103 (98%)	2 (2%)	0	100	100
2	D	106/103 (103%)	104 (98%)	2 (2%)	0	100	100
2	E	106/103 (103%)	103 (97%)	3 (3%)	0	100	100
2	F	107/103 (104%)	104 (97%)	3 (3%)	0	100	100
2	G	104/103 (101%)	101 (97%)	3 (3%)	0	100	100
2	H	107/103 (104%)	104 (97%)	3 (3%)	0	100	100
2	I	109/103 (106%)	105 (96%)	4 (4%)	0	100	100
2	J	103/103 (100%)	101 (98%)	2 (2%)	0	100	100
2	K	104/103 (101%)	101 (97%)	3 (3%)	0	100	100
2	L	104/103 (101%)	101 (97%)	3 (3%)	0	100	100
All	All	1538/1510 (102%)	1489 (97%)	47 (3%)	2 (0%)	51	64

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	49[A]	GLN
1	B	49[B]	GLN

### 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	211/204 (103%)	204 (97%)	7 (3%)	38	53
1	B	206/204 (101%)	198 (96%)	8 (4%)	32	46
2	C	93/89 (104%)	91 (98%)	2 (2%)	52	69
2	D	94/89 (106%)	92 (98%)	2 (2%)	53	70
2	E	94/89 (106%)	93 (99%)	1 (1%)	73	86
2	F	95/89 (107%)	94 (99%)	1 (1%)	73	86
2	G	92/89 (103%)	88 (96%)	4 (4%)	29	40
2	H	95/89 (107%)	94 (99%)	1 (1%)	73	86
2	I	96/89 (108%)	95 (99%)	1 (1%)	76	87
2	J	91/89 (102%)	90 (99%)	1 (1%)	73	86
2	K	92/89 (103%)	89 (97%)	3 (3%)	38	53
2	L	92/89 (103%)	89 (97%)	3 (3%)	38	53
All	All	1351/1298 (104%)	1317 (98%)	34 (2%)	50	65

All (34) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	14	ASP
1	A	48	THR
1	A	56	ASP
1	A	109	ASP
1	A	172	ARG
1	A	195	MET
1	A	206	LEU
2	D	59	ASP
2	D	89	ASN
2	E	59	ASP
2	F	59	ASP
2	G	9[A]	CYS
2	G	9[B]	CYS
2	G	74	ILE
2	G	103	ASN
2	H	59	ASP
1	B	14	ASP
1	B	55	HIS
1	B	61	SER
1	B	71[A]	LEU
1	B	71[B]	LEU
1	B	152	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	172	ARG
1	B	238	ASP
2	C	59	ASP
2	C	62	LYS
2	I	103	ASN
2	J	35	ARG
2	K	9[A]	CYS
2	K	9[B]	CYS
2	K	74	ILE
2	L	9[A]	CYS
2	L	9[B]	CYS
2	L	59	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
2	D	16	GLN
2	D	89	ASN
2	E	103	ASN
2	C	103	ASN
2	I	103	ASN

### 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](#)

Of 33 ligands modelled in this entry, 4 are monoatomic - leaving 29 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
5	GOL	E	202	-	5,5,5	0.11	0	5,5,5	0.27	0
5	GOL	C	203	-	5,5,5	0.16	0	5,5,5	0.28	0
4	GAL	L	201	-	12,12,12	0.45	0	17,17,17	0.42	0
5	GOL	L	202	-	5,5,5	0.14	0	5,5,5	0.28	0
5	GOL	L	203	-	5,5,5	0.11	0	5,5,5	0.26	0
5	GOL	A	303	-	5,5,5	0.09	0	5,5,5	0.23	0
5	GOL	I	202	-	5,5,5	0.13	0	5,5,5	0.26	0
6	ACT	G	202	-	3,3,3	1.16	0	3,3,3	0.77	0
5	GOL	B	303	-	5,5,5	0.12	0	5,5,5	0.27	0
5	GOL	G	204	-	5,5,5	0.13	0	5,5,5	0.27	0
4	GAL	A	302	-	12,12,12	0.88	1 (8%)	17,17,17	0.47	0
6	ACT	C	201	-	3,3,3	1.03	0	3,3,3	0.80	0
8	GLA	C	202	-	12,12,12	0.79	0	17,17,17	0.47	0
4	GAL	K	201	-	12,12,12	0.21	0	17,17,17	0.28	0
5	GOL	A	304	-	5,5,5	0.10	0	5,5,5	0.22	0
4	GAL	E	201	-	12,12,12	0.56	0	17,17,17	0.39	0
4	GAL	D	202	-	12,12,12	0.68	0	17,17,17	0.53	0
6	ACT	F	201	-	3,3,3	0.99	0	3,3,3	0.80	0
5	GOL	I	203	-	5,5,5	0.11	0	5,5,5	0.24	0
4	GAL	J	201	-	12,12,12	0.66	0	17,17,17	0.53	0
7	PEG	G	201	-	6,6,6	0.19	0	5,5,5	0.09	0
5	GOL	D	203	-	5,5,5	0.11	0	5,5,5	0.25	0
5	GOL	A	305	-	5,5,5	0.11	0	5,5,5	0.25	0
4	GAL	F	202	-	12,12,12	0.62	0	17,17,17	0.33	0
5	GOL	J	202	-	5,5,5	0.10	0	5,5,5	0.27	0
4	GAL	G	203	-	12,12,12	0.58	0	17,17,17	0.31	0
4	GAL	I	201	-	12,12,12	0.49	0	17,17,17	0.27	0
4	GAL	H	201	-	12,12,12	0.54	0	17,17,17	0.42	0
5	GOL	H	202	-	5,5,5	0.11	0	5,5,5	0.26	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
5	GOL	E	202	-	-	2/4/4/4	-
5	GOL	C	203	-	-	2/4/4/4	-
4	GAL	L	201	-	-	0/2/22/22	0/1/1/1
5	GOL	L	202	-	-	4/4/4/4	-
5	GOL	L	203	-	-	2/4/4/4	-
5	GOL	A	303	-	-	2/4/4/4	-
5	GOL	I	202	-	-	4/4/4/4	-
5	GOL	B	303	-	-	4/4/4/4	-
5	GOL	G	204	-	-	0/4/4/4	-
4	GAL	A	302	-	-	2/2/22/22	0/1/1/1
8	GLA	C	202	-	-	1/2/22/22	0/1/1/1
4	GAL	K	201	-	-	2/2/22/22	0/1/1/1
5	GOL	A	304	-	-	2/4/4/4	-
4	GAL	E	201	-	-	0/2/22/22	0/1/1/1
4	GAL	D	202	-	-	0/2/22/22	0/1/1/1
5	GOL	I	203	-	-	2/4/4/4	-
4	GAL	J	201	-	-	0/2/22/22	0/1/1/1
7	PEG	G	201	-	-	1/4/4/4	-
5	GOL	D	203	-	-	4/4/4/4	-
5	GOL	A	305	-	-	0/4/4/4	-
4	GAL	F	202	-	-	0/2/22/22	0/1/1/1
5	GOL	J	202	-	-	2/4/4/4	-
4	GAL	G	203	-	-	0/2/22/22	0/1/1/1
4	GAL	I	201	-	-	0/2/22/22	0/1/1/1
4	GAL	H	201	-	-	1/2/22/22	0/1/1/1
5	GOL	H	202	-	-	4/4/4/4	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	302	GAL	C4-C5	2.72	1.58	1.53

There are no bond angle outliers.

There are no chirality outliers.

All (41) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	303	GOL	O2-C2-C3-O3
5	A	304	GOL	C1-C2-C3-O3

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
5	H	202	GOL	C1-C2-C3-O3
5	B	303	GOL	O1-C1-C2-C3
5	B	303	GOL	C1-C2-C3-O3
5	C	203	GOL	O1-C1-C2-C3
5	I	202	GOL	O1-C1-C2-C3
5	J	202	GOL	C1-C2-C3-O3
5	L	202	GOL	O1-C1-C2-C3
4	K	201	GAL	O5-C5-C6-O6
5	A	303	GOL	C1-C2-C3-O3
5	D	203	GOL	O1-C1-C2-C3
5	D	203	GOL	C1-C2-C3-O3
5	E	202	GOL	C1-C2-C3-O3
5	H	202	GOL	O1-C1-C2-C3
5	I	202	GOL	C1-C2-C3-O3
5	I	203	GOL	O1-C1-C2-C3
5	L	202	GOL	C1-C2-C3-O3
5	L	203	GOL	C1-C2-C3-O3
5	A	304	GOL	O2-C2-C3-O3
5	D	203	GOL	O1-C1-C2-O2
5	D	203	GOL	O2-C2-C3-O3
5	H	202	GOL	O2-C2-C3-O3
5	B	303	GOL	O1-C1-C2-O2
5	I	202	GOL	O2-C2-C3-O3
5	I	203	GOL	O1-C1-C2-O2
5	L	202	GOL	O1-C1-C2-O2
5	H	202	GOL	O1-C1-C2-O2
5	B	303	GOL	O2-C2-C3-O3
5	C	203	GOL	O1-C1-C2-O2
5	J	202	GOL	O2-C2-C3-O3
5	L	202	GOL	O2-C2-C3-O3
4	A	302	GAL	C4-C5-C6-O6
4	H	201	GAL	O5-C5-C6-O6
5	I	202	GOL	O1-C1-C2-O2
4	A	302	GAL	O5-C5-C6-O6
5	E	202	GOL	O2-C2-C3-O3
4	K	201	GAL	C4-C5-C6-O6
7	G	201	PEG	C4-C3-O2-C2
8	C	202	GLA	C4-C5-C6-O6
5	L	203	GOL	O2-C2-C3-O3

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	302	GAL	1	0
4	K	201	GAL	1	0
5	H	202	GOL	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

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.