



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 9, 2020 – 06:55 PM BST

PDB ID : 2ZUW  
Title : Crystal structure of Galacto-N-biose/Lacto-N-biose I phosphorylase in complex with GlcNAc and sulfate  
Authors : Hidaka, M.; Nishimoto, M.; Kitaoka, M.; Wakagi, T.; Shoun, H.; Fushinobu, S.  
Deposited on : 2008-10-28  
Resolution : 2.11 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.13.1  
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) : 2.13.1

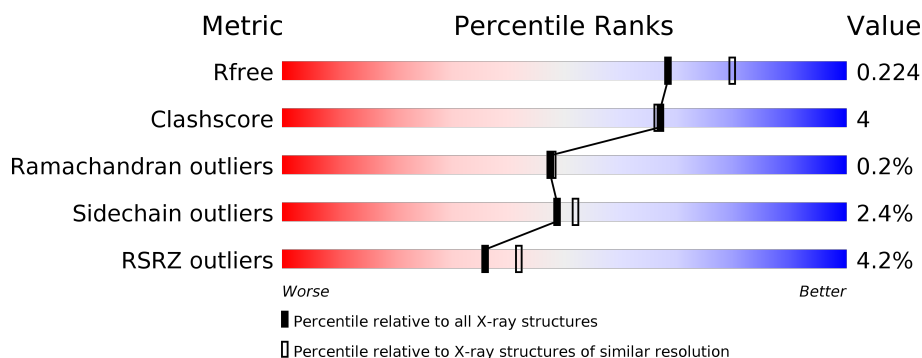
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.11 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	6241 (2.14-2.10)
Clashscore	141614	6778 (2.14-2.10)
Ramachandran outliers	138981	6705 (2.14-2.10)
Sidechain outliers	138945	6706 (2.14-2.10)
RSRZ outliers	127900	6112 (2.14-2.10)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	759	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red, orange, yellow, green, grey);"></div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> <span>%</span> <span>91%</span> <span>6% ..</span> </div> </div>
1	B	759	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red, orange, yellow, green, grey);"></div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> <span>6%</span> <span>87%</span> <span>8% ..</span> </div> </div>
1	C	759	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red, orange, yellow, green, grey);"></div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> <span>2%</span> <span>92%</span> <span>6% .</span> </div> </div>
1	D	759	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red, orange, yellow, green, grey);"></div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> <span>7%</span> <span>81%</span> <span>11% • 6%</span> </div> </div>

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 25782 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 Lacto-N-biose phosphorylase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	745	Total	C	N	O	S	0	0	0
			5926	3782	995	1133	16			
1	B	734	Total	C	N	O	S	0	0	0
			5833	3723	983	1111	16			
1	C	748	Total	C	N	O	S	0	0	0
			5950	3795	1000	1139	16			
1	D	713	Total	C	N	O	S	0	0	0
			5686	3635	957	1078	16			

There are 32 discrepancies between the modelled and reference sequences:

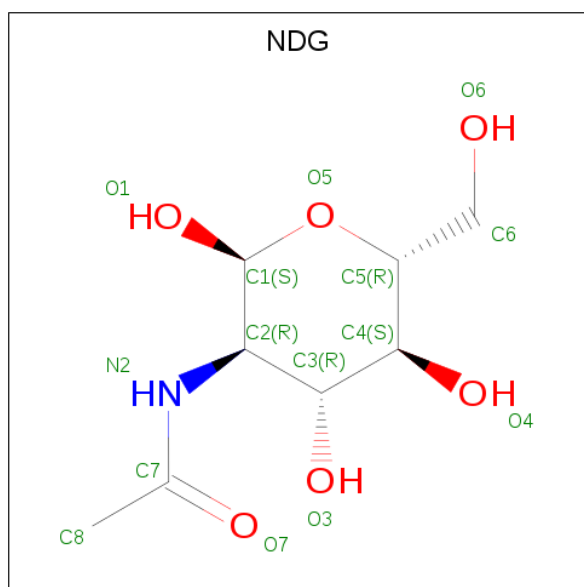
Chain	Residue	Modelled	Actual	Comment	Reference
A	752	LEU	-	expression tag	UNP Q5NU17
A	753	GLU	-	expression tag	UNP Q5NU17
A	754	HIS	-	expression tag	UNP Q5NU17
A	755	HIS	-	expression tag	UNP Q5NU17
A	756	HIS	-	expression tag	UNP Q5NU17
A	757	HIS	-	expression tag	UNP Q5NU17
A	758	HIS	-	expression tag	UNP Q5NU17
A	759	HIS	-	expression tag	UNP Q5NU17
B	752	LEU	-	expression tag	UNP Q5NU17
B	753	GLU	-	expression tag	UNP Q5NU17
B	754	HIS	-	expression tag	UNP Q5NU17
B	755	HIS	-	expression tag	UNP Q5NU17
B	756	HIS	-	expression tag	UNP Q5NU17
B	757	HIS	-	expression tag	UNP Q5NU17
B	758	HIS	-	expression tag	UNP Q5NU17
B	759	HIS	-	expression tag	UNP Q5NU17
C	752	LEU	-	expression tag	UNP Q5NU17
C	753	GLU	-	expression tag	UNP Q5NU17
C	754	HIS	-	expression tag	UNP Q5NU17
C	755	HIS	-	expression tag	UNP Q5NU17
C	756	HIS	-	expression tag	UNP Q5NU17

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	757	HIS	-	expression tag	UNP Q5NU17
C	758	HIS	-	expression tag	UNP Q5NU17
C	759	HIS	-	expression tag	UNP Q5NU17
D	752	LEU	-	expression tag	UNP Q5NU17
D	753	GLU	-	expression tag	UNP Q5NU17
D	754	HIS	-	expression tag	UNP Q5NU17
D	755	HIS	-	expression tag	UNP Q5NU17
D	756	HIS	-	expression tag	UNP Q5NU17
D	757	HIS	-	expression tag	UNP Q5NU17
D	758	HIS	-	expression tag	UNP Q5NU17
D	759	HIS	-	expression tag	UNP Q5NU17

- Molecule 2 is 2-acetamido-2-deoxy-alpha-D-glucopyranose (three-letter code: NDG) (formula:  $C_8H_{15}NO_6$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			15	8	1	6		
2	B	1	Total	C	N	O	0	0
			15	8	1	6		
2	C	1	Total	C	N	O	0	0
			15	8	1	6		
2	D	1	Total	C	N	O	0	0
			15	8	1	6		

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



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

- Molecule 4 is NITRATE ION (three-letter code: NO3) (formula: NO<sub>3</sub>).

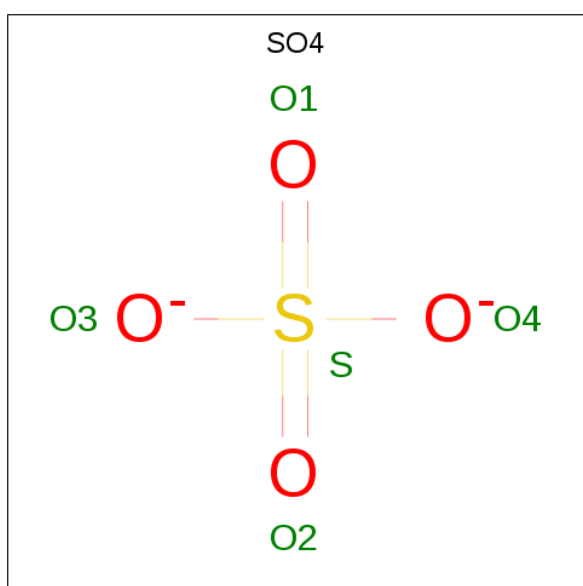


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	N	O	0	0
			4	1	3		

- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	1	Total	Mg	0	0
			1	1		

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



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

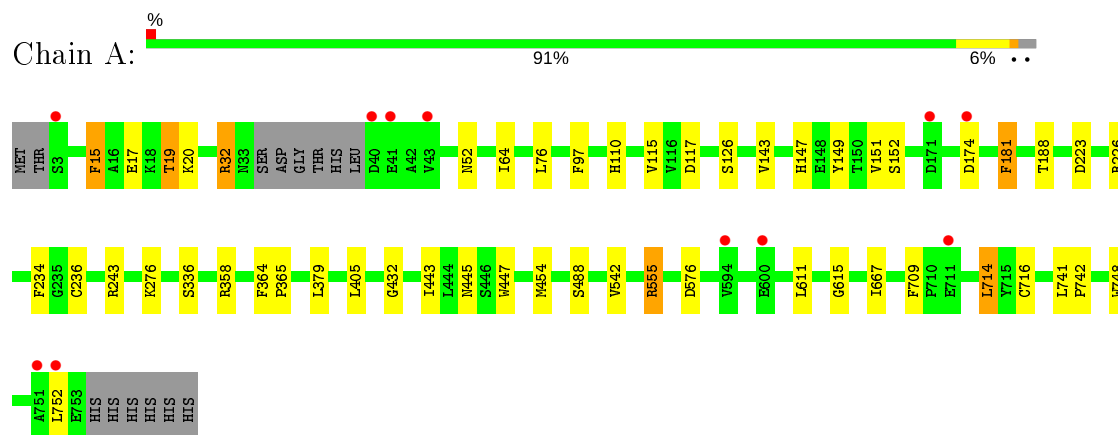
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	627	Total	O	0	0
			627	627		
7	B	533	Total	O	2	0
			533	533		
7	C	601	Total	O	1	0
			601	601		
7	D	520	Total	O	0	0
			520	520		

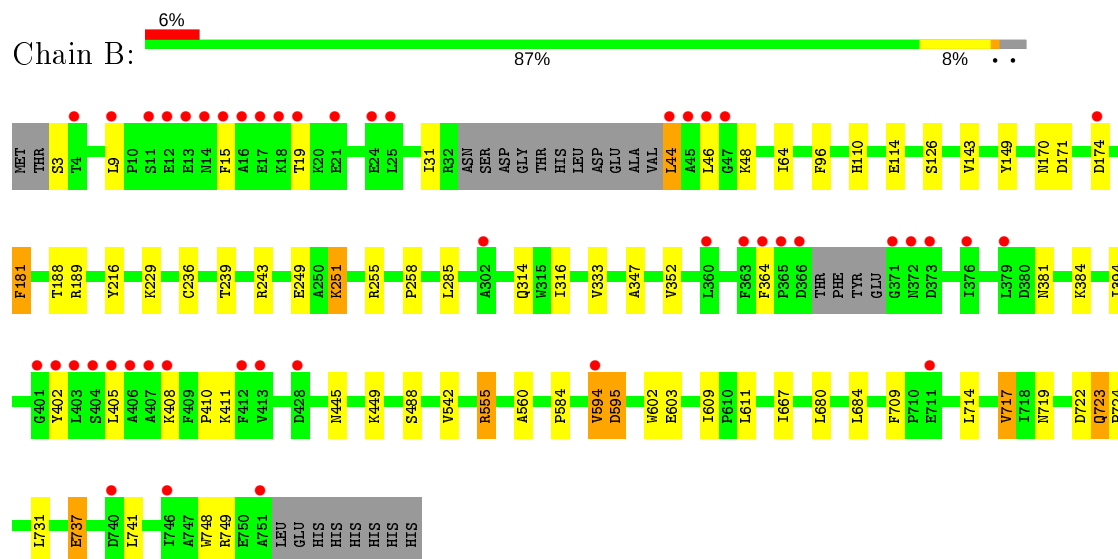
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: Lacto-N-biose phosphorylase



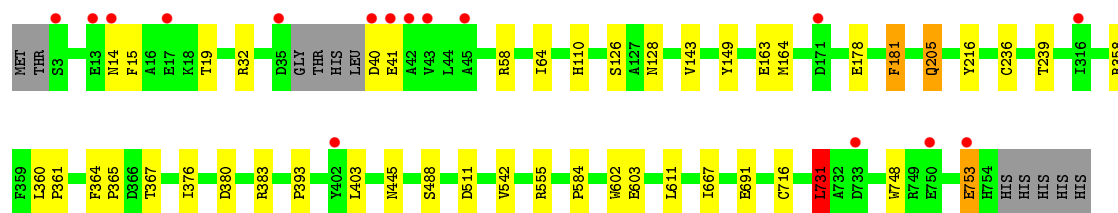
- Molecule 1: Lacto-N-biose phosphorylase



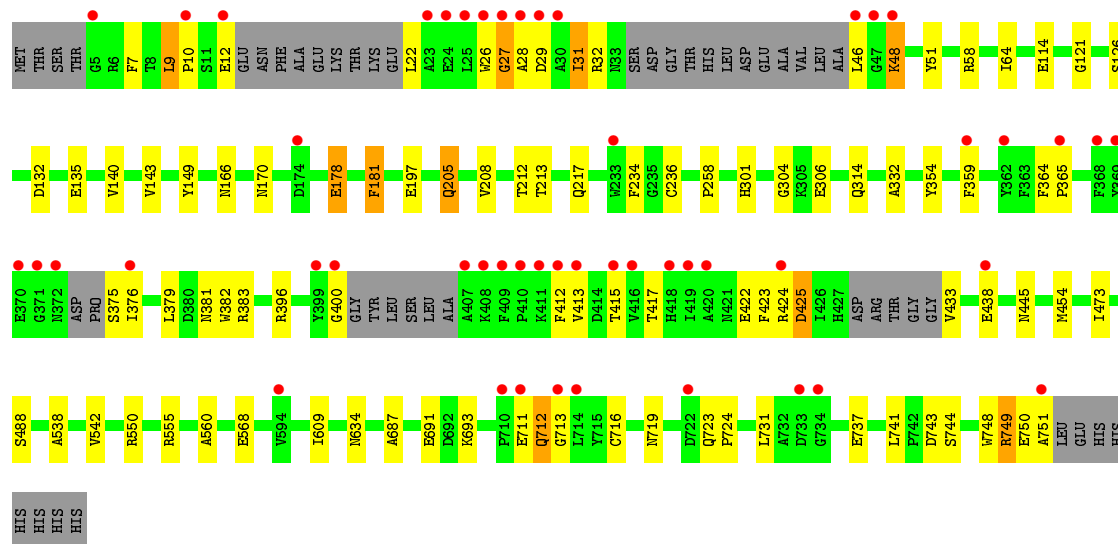
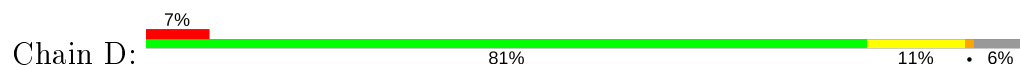
- Molecule 1: Lacto-N-biose phosphorylase







• Molecule 1: Lacto-N-biose phosphorylase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	67.91Å 112.04Å 118.66Å 105.21° 90.43° 107.32°	Depositor
Resolution (Å)	41.07 – 2.11 40.88 – 2.11	Depositor EDS
% Data completeness (in resolution range)	97.4 (41.07-2.11) 97.3 (40.88-2.11)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.66 (at 2.12Å)	Xtriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.169 , 0.225 0.169 , 0.224	Depositor DCC
$R_{free}$ test set	8997 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	22.2	Xtriage
Anisotropy	0.020	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 51.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.025 for h,-h-k,-l	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	25782	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 3.53% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, MG, SO4, NDG, NO3

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	0.65	0/6093	0.66	1/8299 (0.0%)
1	B	0.58	0/5997	0.62	0/8166
1	C	0.63	1/6118 (0.0%)	0.65	1/8333 (0.0%)
1	D	0.63	1/5846 (0.0%)	0.65	0/7958
All	All	0.62	2/24054 (0.0%)	0.64	2/32756 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	236	CYS	CB-SG	-7.78	1.69	1.82
1	C	236	CYS	CB-SG	-6.72	1.70	1.82

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	731	LEU	CA-CB-CG	5.50	127.95	115.30
1	A	223	ASP	CB-CG-OD1	5.10	122.89	118.30

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	5926	0	5615	30	0
1	B	5833	0	5536	52	0
1	C	5950	0	5631	32	0
1	D	5686	0	5386	64	0
2	A	15	0	12	0	0
2	B	15	0	12	0	0
2	C	15	0	12	0	0
2	D	15	0	12	0	0
3	A	6	0	8	1	0
3	B	6	0	8	2	0
3	C	12	0	16	0	0
3	D	12	0	16	1	0
4	A	4	0	0	0	0
5	B	1	0	0	0	0
6	D	5	0	0	0	0
7	A	627	0	0	7	0
7	B	533	0	0	12	0
7	C	601	0	0	7	0
7	D	520	0	0	11	0
All	All	25782	0	22264	177	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 4.

All (177) 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:359:PHE:HD1	7:D:831:HOH:O	1.37	1.04
1:B:251:LYS:HE2	7:B:961:HOH:O	1.58	1.00
1:C:691:GLU:HG2	7:C:964:HOH:O	1.67	0.93
1:D:412:PHE:HA	1:D:415:THR:HG22	1.52	0.92
1:D:121:GLY:O	3:D:4010:GOL:H12	1.69	0.91
1:B:449:LYS:HB3	3:B:4006:GOL:H32	1.56	0.86
1:A:110:HIS:HD2	1:A:126:SER:HB2	1.41	0.83
1:A:20:LYS:HE2	7:A:2060:HOH:O	1.77	0.83
1:B:110:HIS:HD2	1:B:126:SER:HB2	1.42	0.82
1:D:412:PHE:HA	1:D:415:THR:CG2	2.10	0.81
1:C:15:PHE:O	1:C:19:THR:HG23	1.83	0.77
1:D:7:PHE:HB2	7:D:1375:HOH:O	1.83	0.77
1:C:163:GLU:HG2	7:C:1001:HOH:O	1.86	0.73
1:B:46:LEU:HD12	7:B:1962:HOH:O	1.88	0.73
1:B:737:GLU:HG2	7:B:2233:HOH:O	1.89	0.71

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:381:ASN:HA	1:B:384:LYS:CE	2.20	0.71
1:D:26:TRP:O	1:D:28:ALA:N	2.24	0.70
1:D:32:ARG:HB3	1:D:51:TYR:HB2	1.74	0.70
1:B:44:LEU:N	7:B:1962:HOH:O	2.25	0.69
1:B:722:ASP:OD1	1:B:723:GLN:NE2	2.26	0.69
1:C:110:HIS:HD2	1:C:126:SER:HB2	1.56	0.69
1:D:375:SER:CB	1:D:415:THR:HG21	2.22	0.68
1:A:110:HIS:CD2	1:A:126:SER:HB2	2.27	0.66
1:B:594:VAL:HG23	1:B:595:ASP:H	1.60	0.66
1:D:711:GLU:C	1:D:713:GLY:H	1.99	0.66
1:D:691:GLU:HG2	7:D:1234:HOH:O	1.98	0.64
1:B:249:GLU:OE1	1:B:255:ARG:NH1	2.31	0.63
1:B:64:ILE:HD13	1:B:181:PHE:HB3	1.78	0.63
1:B:236:CYS:HB3	1:B:316:ILE:HG21	1.79	0.63
1:B:555:ARG:HD2	7:B:1757:HOH:O	1.98	0.63
1:B:110:HIS:CD2	1:B:126:SER:HB2	2.30	0.63
1:A:709:PHE:HB2	1:A:714:LEU:HD12	1.82	0.62
1:D:412:PHE:CA	1:D:415:THR:HG22	2.27	0.62
1:D:724:PRO:HA	1:D:741:LEU:O	2.00	0.61
1:B:717:VAL:HG22	1:B:741:LEU:HD11	1.83	0.61
1:A:17:GLU:OE2	1:A:17:GLU:HA	2.01	0.61
1:D:731:LEU:HD11	1:D:737:GLU:HG3	1.82	0.61
1:B:381:ASN:HA	1:B:384:LYS:HE2	1.82	0.60
1:D:140:VAL:HG12	1:D:143:VAL:CG1	2.31	0.60
1:D:375:SER:HB3	1:D:415:THR:HG21	1.82	0.60
1:C:110:HIS:CD2	1:C:126:SER:HB2	2.37	0.60
1:B:381:ASN:HA	1:B:384:LYS:HE3	1.83	0.59
1:D:29:ASP:HB3	1:D:396:ARG:HH22	1.66	0.59
1:C:376:ILE:HD12	7:C:1002:HOH:O	2.02	0.58
1:D:48:LYS:HB2	7:D:811:HOH:O	2.03	0.58
1:D:400:GLY:C	7:D:894:HOH:O	2.40	0.58
1:C:380:ASP:HB3	1:C:383:ARG:NH2	2.19	0.58
1:B:609:ILE:N	1:B:609:ILE:HD12	2.18	0.57
1:A:15:PHE:O	1:A:19:THR:HG23	2.04	0.57
1:A:64:ILE:HD13	1:A:181:PHE:HB3	1.87	0.57
1:D:413:VAL:CG2	7:D:2109:HOH:O	2.53	0.57
1:D:413:VAL:HG23	7:D:2109:HOH:O	2.05	0.57
1:B:449:LYS:CB	3:B:4006:GOL:H32	2.32	0.56
1:C:364:PHE:CD2	1:C:365:PRO:HD2	2.40	0.56
1:B:251:LYS:CE	7:B:961:HOH:O	2.34	0.56
1:D:64:ILE:HD13	1:D:181:PHE:HB3	1.88	0.56

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:143:VAL:HB	1:A:149:TYR:OH	2.06	0.55
1:A:447:TRP:O	3:A:4005:GOL:H31	2.07	0.55
1:A:181:PHE:HB2	1:A:188:THR:HG21	1.89	0.55
1:B:719:ASN:ND2	1:B:723:GLN:O	2.36	0.55
1:C:511:ASP:OD2	1:D:550:ARG:HD3	2.07	0.54
1:D:26:TRP:O	1:D:27:GLY:C	2.45	0.54
1:C:19:THR:HG22	1:C:403:LEU:HD23	1.90	0.53
1:C:64:ILE:HD13	1:C:181:PHE:HB3	1.91	0.53
1:D:750:GLU:O	1:D:751:ALA:C	2.44	0.53
1:D:473:ILE:HD11	1:D:542:VAL:HG11	1.89	0.53
1:C:205:GLN:H	1:C:205:GLN:CD	2.11	0.53
1:C:542:VAL:HA	1:C:667:ILE:O	2.09	0.53
1:D:140:VAL:HG12	1:D:143:VAL:HG11	1.90	0.53
1:D:58:ARG:HD2	1:D:178:GLU:HG2	1.90	0.53
1:D:205:GLN:H	1:D:205:GLN:CD	2.13	0.52
1:D:379:LEU:HG	1:D:383:ARG:HH12	1.74	0.52
1:B:15:PHE:O	1:B:19:THR:HG23	2.08	0.52
1:B:9:LEU:O	1:B:31:ILE:HA	2.09	0.52
1:C:32:ARG:HD3	7:C:1575:HOH:O	2.10	0.52
1:C:364:PHE:CG	1:C:365:PRO:HD2	2.45	0.52
1:D:445:ASN:O	1:D:488:SER:HA	2.10	0.52
1:D:438:GLU:HB3	1:D:693:LYS:NZ	2.24	0.51
1:A:234:PHE:CZ	1:A:454:MET:HB3	2.46	0.51
1:B:555:ARG:HD3	1:B:560:ALA:HB3	1.93	0.51
1:B:609:ILE:HD11	7:B:1404:HOH:O	2.11	0.51
1:A:32:ARG:HD3	7:A:887:HOH:O	2.10	0.50
1:D:422:GLU:HA	1:D:425:ASP:OD2	2.11	0.50
1:D:711:GLU:C	1:D:713:GLY:N	2.65	0.50
1:A:110:HIS:HE1	7:A:807:HOH:O	1.95	0.49
1:B:709:PHE:HB2	1:B:714:LEU:HD12	1.94	0.49
1:A:445:ASN:O	1:A:488:SER:HA	2.11	0.49
1:C:205:GLN:HB3	7:C:919:HOH:O	2.13	0.49
1:D:9:LEU:HA	7:D:1809:HOH:O	2.12	0.49
1:D:31:ILE:HG13	1:D:32:ARG:N	2.27	0.49
1:C:584:PRO:HD2	7:C:901:HOH:O	2.12	0.49
1:B:402:TYR:HB2	1:B:405:LEU:HD13	1.95	0.48
1:A:97:PHE:CE1	1:A:226:ARG:HD2	2.49	0.48
1:B:181:PHE:HB2	1:B:188:THR:HG21	1.96	0.48
1:D:711:GLU:O	1:D:713:GLY:N	2.47	0.48
1:D:383:ARG:NH1	1:D:744:SER:OG	2.47	0.48
1:C:602:TRP:CE3	1:C:603:GLU:HG3	2.49	0.47

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:749:ARG:NH2	7:D:2064:HOH:O	2.47	0.47
1:B:410:PRO:HA	7:B:1449:HOH:O	2.13	0.47
1:D:170:ASN:ND2	1:D:365:PRO:HD3	2.30	0.47
1:D:140:VAL:CG1	1:D:143:VAL:HG11	2.45	0.47
1:D:413:VAL:O	1:D:417:THR:HG23	2.15	0.47
1:C:143:VAL:HB	1:C:149:TYR:CZ	2.50	0.47
1:C:393:PRO:HB3	7:C:992:HOH:O	2.15	0.47
1:D:379:LEU:HG	1:D:383:ARG:NH1	2.30	0.46
1:D:234:PHE:CZ	1:D:454:MET:HB3	2.49	0.46
1:C:380:ASP:HB3	1:C:383:ARG:HH22	1.79	0.46
1:D:359:PHE:CD1	7:D:831:HOH:O	2.27	0.46
1:C:216:TYR:HB3	1:C:239:THR:HG22	1.98	0.46
1:C:555:ARG:HE	1:C:555:ARG:HA	1.80	0.46
1:D:132:ASP:HB3	1:D:135:GLU:O	2.16	0.46
1:B:114:GLU:HB2	1:B:258:PRO:HG2	1.98	0.46
1:D:555:ARG:HD2	1:D:560:ALA:HB3	1.97	0.46
1:B:555:ARG:HE	1:B:555:ARG:HA	1.80	0.46
1:C:445:ASN:O	1:C:488:SER:HA	2.16	0.46
1:D:10:PRO:HA	1:D:32:ARG:O	2.16	0.45
1:D:716:CYS:HB3	1:D:748:TRP:CE3	2.50	0.45
1:D:719:ASN:ND2	1:D:723:GLN:O	2.38	0.45
1:C:731:LEU:HD23	1:C:731:LEU:N	2.31	0.45
1:D:375:SER:HB2	1:D:415:THR:HG21	1.99	0.45
1:B:724:PRO:HA	1:B:741:LEU:O	2.17	0.45
1:A:716:CYS:HB3	1:A:748:TRP:CE3	2.52	0.45
1:B:243:ARG:HA	1:B:243:ARG:HD2	1.44	0.44
1:B:408:LYS:O	1:B:410:PRO:HD3	2.18	0.44
1:B:609:ILE:CD1	7:B:1404:HOH:O	2.64	0.44
1:B:555:ARG:CD	7:B:1757:HOH:O	2.61	0.44
1:A:143:VAL:HB	1:A:149:TYR:CZ	2.52	0.44
1:B:542:VAL:HA	1:B:667:ILE:O	2.17	0.44
1:B:748:TRP:O	1:B:749:ARG:NH1	2.50	0.44
1:C:128:ASN:N	1:C:128:ASN:HD22	2.15	0.44
1:A:143:VAL:HG12	7:A:2021:HOH:O	2.16	0.44
1:A:576:ASP:HB3	1:A:615:GLY:HA2	2.00	0.44
1:C:58:ARG:HD2	1:C:178:GLU:OE1	2.18	0.44
1:A:432:GLY:HA2	7:A:1535:HOH:O	2.17	0.44
1:B:584:PRO:HD2	7:B:815:HOH:O	2.17	0.44
1:C:716:CYS:HB3	1:C:748:TRP:CE3	2.53	0.43
1:D:166:ASN:ND2	1:D:364:PHE:HE2	2.16	0.43
1:D:217:GLN:HA	1:D:314:GLN:HA	2.00	0.43

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:197:GLU:HG3	7:D:1792:HOH:O	2.18	0.43
1:D:301:HIS:O	1:D:304:GLY:N	2.39	0.43
1:A:276:LYS:HE2	7:A:1055:HOH:O	2.18	0.43
1:B:445:ASN:O	1:B:488:SER:HA	2.18	0.43
1:B:189:ARG:HD2	7:B:1674:HOH:O	2.18	0.43
1:B:333:VAL:HG23	1:B:352:VAL:HG11	2.01	0.43
1:A:555:ARG:HD3	7:A:1008:HOH:O	2.19	0.43
1:D:538:ALA:HB1	1:D:687:ALA:HB1	2.00	0.42
1:B:216:TYR:O	1:B:314:GLN:HA	2.19	0.42
1:B:402:TYR:CB	1:B:405:LEU:HD13	2.49	0.42
1:B:143:VAL:HB	1:B:149:TYR:OH	2.20	0.42
1:B:680:LEU:O	1:B:684:LEU:HG	2.19	0.42
1:A:234:PHE:CE1	1:A:454:MET:HB3	2.55	0.42
1:A:117:ASP:OD2	1:A:147:HIS:ND1	2.53	0.42
1:A:741:LEU:HA	1:A:742:PRO:HD3	1.93	0.42
1:C:164:MET:HE2	1:C:164:MET:HB3	1.96	0.42
1:D:114:GLU:HB2	1:D:258:PRO:HG2	2.01	0.42
1:A:542:VAL:HA	1:A:667:ILE:O	2.20	0.41
1:B:170:ASN:HD21	1:B:364:PHE:HZ	1.64	0.41
1:C:602:TRP:CZ3	1:C:603:GLU:HG3	2.55	0.41
1:D:332:ALA:HA	1:D:354:TYR:O	2.19	0.41
1:D:438:GLU:HB3	1:D:693:LYS:HZ3	1.85	0.41
1:A:115:VAL:HG22	1:A:151:VAL:HG22	2.02	0.41
1:C:753:GLU:HG2	1:C:753:GLU:H	1.71	0.41
1:D:212:THR:N	1:D:213:THR:HA	2.35	0.41
1:B:143:VAL:HB	1:B:149:TYR:CZ	2.55	0.41
1:A:364:PHE:CD2	1:A:365:PRO:HD2	2.56	0.41
1:C:360:LEU:HB3	1:C:361:PRO:HA	2.03	0.41
1:B:602:TRP:CE3	1:B:603:GLU:HG3	2.55	0.41
1:D:22:LEU:O	1:D:26:TRP:HB2	2.20	0.41
1:B:96:PHE:CE1	1:B:229:LYS:HE2	2.56	0.41
1:D:143:VAL:HB	1:D:149:TYR:CZ	2.56	0.41
1:D:9:LEU:O	1:D:31:ILE:HD12	2.21	0.41
1:A:97:PHE:CZ	1:A:226:ARG:HD2	2.56	0.40
1:B:239:THR:HG21	1:B:285:LEU:HD21	2.02	0.40
1:B:347:ALA:HB2	1:B:394:ILE:CG2	2.51	0.40
1:D:568:GLU:HB3	1:D:634:ASN:HB3	2.02	0.40
1:A:76:LEU:O	1:A:152:SER:HA	2.21	0.40
1:D:208:VAL:HG22	1:D:306:GLU:HB2	2.04	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	741/759 (98%)	711 (96%)	30 (4%)	0	100	100
1	B	728/759 (96%)	694 (95%)	32 (4%)	2 (0%)	41	40
1	C	744/759 (98%)	718 (96%)	25 (3%)	1 (0%)	51	53
1	D	701/759 (92%)	663 (95%)	34 (5%)	4 (1%)	25	20
All	All	2914/3036 (96%)	2786 (96%)	121 (4%)	7 (0%)	47	48

All (7) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	594	VAL
1	D	27	GLY
1	D	712	GLN
1	B	595	ASP
1	C	41	GLU
1	D	382	TRP
1	D	423	PHE

### 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	621/634 (98%)	604 (97%)	17 (3%)	44	47
1	B	611/634 (96%)	597 (98%)	14 (2%)	50	53
1	C	624/634 (98%)	615 (99%)	9 (1%)	67	72

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	D	596/634 (94%)	578 (97%)	18 (3%)	41	43
All	All	2452/2536 (97%)	2394 (98%)	58 (2%)	49	52

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

Mol	Chain	Res	Type
1	A	15	PHE
1	A	19	THR
1	A	32	ARG
1	A	52	ASN
1	A	174	ASP
1	A	181	PHE
1	A	236	CYS
1	A	243	ARG
1	A	336	SER
1	A	358	ARG
1	A	379	LEU
1	A	405	LEU
1	A	443	ILE
1	A	555	ARG
1	A	611	LEU
1	A	714	LEU
1	A	752	LEU
1	B	3	SER
1	B	44	LEU
1	B	48	LYS
1	B	171	ASP
1	B	174	ASP
1	B	181	PHE
1	B	251	LYS
1	B	411	LYS
1	B	555	ARG
1	B	611	LEU
1	B	717	VAL
1	B	723	GLN
1	B	731	LEU
1	B	737	GLU
1	C	14	ASN
1	C	40	ASP
1	C	181	PHE
1	C	205	GLN
1	C	358	ARG

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	C	367	THR
1	C	611	LEU
1	C	731	LEU
1	C	753	GLU
1	D	9	LEU
1	D	12	GLU
1	D	31	ILE
1	D	46	LEU
1	D	48	LYS
1	D	126	SER
1	D	178	GLU
1	D	181	PHE
1	D	205	GLN
1	D	376	ILE
1	D	381	ASN
1	D	424	ARG
1	D	425	ASP
1	D	433	VAL
1	D	609	ILE
1	D	712	GLN
1	D	743	ASP
1	D	749	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (21) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	110	HIS
1	A	128	ASN
1	A	170	ASN
1	A	690	ASN
1	A	720	ASN
1	B	52	ASN
1	B	128	ASN
1	B	690	ASN
1	B	720	ASN
1	C	110	HIS
1	C	128	ASN
1	C	170	ASN
1	C	641	ASN
1	C	690	ASN
1	C	720	ASN
1	D	128	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	D	166	ASN
1	D	170	ASN
1	D	381	ASN
1	D	690	ASN
1	D	720	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 13 ligands modelled in this entry, 1 is monoatomic - leaving 12 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	$\# Z  > 2$
3	GOL	C	4007	-	5,5,5	0.26	0	5,5,5	0.66	0
3	GOL	C	4008	-	5,5,5	0.31	0	5,5,5	0.53	0
2	NDG	C	4003	-	15,15,15	0.62	0	21,21,21	1.04	1 (4%)
3	GOL	D	4010	-	5,5,5	0.38	0	5,5,5	0.54	0
6	SO4	D	4011	-	4,4,4	0.17	0	6,6,6	0.21	0
2	NDG	B	4002	-	15,15,15	0.63	0	21,21,21	1.27	2 (9%)
3	GOL	D	4009	-	5,5,5	0.29	0	5,5,5	0.81	0
3	GOL	A	4005	-	5,5,5	0.33	0	5,5,5	0.75	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	NDG	A	4001	-	15,15,15	0.59	0	21,21,21	1.07	1 (4%)
3	GOL	B	4006	-	5,5,5	0.29	0	5,5,5	0.76	0
4	NO3	A	4012	-	1,3,3	3.78	1 (100%)	0,3,3	0.00	-
2	NDG	D	4004	-	15,15,15	0.64	0	21,21,21	1.27	3 (14%)

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
3	GOL	C	4007	-	-	0/4/4/4	-
3	GOL	C	4008	-	-	4/4/4/4	-
2	NDG	C	4003	-	-	0/6/26/26	0/1/1/1
3	GOL	D	4010	-	-	2/4/4/4	-
2	NDG	B	4002	-	-	0/6/26/26	0/1/1/1
3	GOL	D	4009	-	-	0/4/4/4	-
3	GOL	A	4005	-	-	2/4/4/4	-
2	NDG	A	4001	-	-	0/6/26/26	0/1/1/1
3	GOL	B	4006	-	-	2/4/4/4	-
2	NDG	D	4004	-	-	0/6/26/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	4012	NO3	O1-N	3.78	1.41	1.24

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	4002	NDG	C1-C2-N2	-4.54	105.47	110.73
2	A	4001	NDG	C1-C2-N2	-3.58	106.59	110.73
2	D	4004	NDG	C1-C2-N2	-3.54	106.62	110.73
2	C	4003	NDG	C1-C2-N2	-3.51	106.66	110.73
2	D	4004	NDG	O5-C1-C2	2.45	111.98	109.52
2	D	4004	NDG	C3-C2-N2	2.16	114.69	110.62
2	B	4002	NDG	O5-C1-C2	2.04	111.56	109.52

There are no chirality outliers.

All (10) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	C	4008	GOL	O1-C1-C2-C3
3	A	4005	GOL	O1-C1-C2-C3
3	B	4006	GOL	C1-C2-C3-O3
3	B	4006	GOL	O2-C2-C3-O3
3	C	4008	GOL	C1-C2-C3-O3
3	C	4008	GOL	O1-C1-C2-O2
3	A	4005	GOL	O1-C1-C2-O2
3	C	4008	GOL	O2-C2-C3-O3
3	D	4010	GOL	O1-C1-C2-O2
3	D	4010	GOL	O1-C1-C2-C3

There are no ring outliers.

3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	D	4010	GOL	1	0
3	A	4005	GOL	1	0
3	B	4006	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	745/759 (98%)	-0.27	11 (1%)	73	77	12, 21, 35, 66	0
1	B	734/759 (96%)	0.08	46 (6%)	20	24	13, 24, 55, 69	0
1	C	748/759 (98%)	-0.25	16 (2%)	63	68	11, 21, 38, 64	0
1	D	713/759 (93%)	0.09	50 (7%)	16	20	12, 23, 53, 64	0
All	All	2940/3036 (96%)	-0.09	123 (4%)	36	42	11, 22, 49, 69	0

All (123) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	406	ALA	7.5
1	D	24	GLU	6.1
1	D	25	LEU	6.0
1	D	371	GLY	5.9
1	B	594	VAL	5.8
1	B	45	ALA	5.4
1	D	710	PRO	5.1
1	D	365	PRO	5.0
1	D	47	GLY	4.9
1	D	376	ILE	4.9
1	D	412	PHE	4.7
1	B	44	LEU	4.7
1	B	407	ALA	4.7
1	D	418	HIS	4.5
1	D	415	THR	4.5
1	D	370	GLU	4.3
1	B	14	ASN	4.3
1	B	46	LEU	4.2
1	B	405	LEU	4.2
1	B	401	GLY	4.2
1	B	12	GLU	4.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	364	PHE	4.0
1	B	13	GLU	3.9
1	A	752	LEU	3.9
1	B	47	GLY	3.9
1	D	362	TYR	3.9
1	C	40	ASP	3.8
1	B	11	SER	3.7
1	D	408	LYS	3.5
1	C	13	GLU	3.5
1	D	5	GLY	3.5
1	D	23	ALA	3.5
1	B	25	LEU	3.5
1	D	413	VAL	3.4
1	D	713	GLY	3.4
1	B	404	SER	3.3
1	C	14	ASN	3.3
1	D	28	ALA	3.3
1	D	424	ARG	3.3
1	B	402	TYR	3.3
1	B	373	ASP	3.3
1	D	416	VAL	3.3
1	D	29	ASP	3.2
1	B	363	PHE	3.1
1	D	10	PRO	3.1
1	C	43	VAL	3.1
1	A	41	GLU	3.1
1	D	410	PRO	3.1
1	B	19	THR	3.1
1	D	46	LEU	3.1
1	B	413	VAL	3.1
1	B	17	GLU	3.1
1	D	407	ALA	3.0
1	D	174	ASP	3.0
1	B	371	GLY	3.0
1	B	379	LEU	3.0
1	D	30	ALA	3.0
1	B	403	LEU	3.0
1	D	411	LYS	2.9
1	D	419	ILE	2.9
1	C	42	ALA	2.9
1	D	372	ASN	2.9
1	C	35	ASP	2.9

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	751	ALA	2.8
1	D	399	TYR	2.8
1	B	408	LYS	2.8
1	B	16	ALA	2.7
1	C	171	ASP	2.7
1	D	369	TYR	2.7
1	B	711	GLU	2.7
1	B	751	ALA	2.7
1	B	4	THR	2.7
1	D	722	ASP	2.7
1	C	41	GLU	2.7
1	B	412	PHE	2.7
1	B	302	ALA	2.7
1	B	372	ASN	2.6
1	B	18	LYS	2.5
1	D	27	GLY	2.5
1	A	40	ASP	2.5
1	B	366	ASP	2.5
1	A	174	ASP	2.5
1	D	26	TRP	2.5
1	D	714	LEU	2.5
1	C	750	GLU	2.5
1	A	3	SER	2.4
1	B	21	GLU	2.4
1	D	409	PHE	2.4
1	B	360	LEU	2.4
1	C	17	GLU	2.4
1	A	600	GLU	2.4
1	B	24	GLU	2.4
1	D	368	PHE	2.4
1	D	12	GLU	2.3
1	C	3	SER	2.3
1	D	438	GLU	2.3
1	A	751	ALA	2.3
1	B	365	PRO	2.3
1	A	171	ASP	2.3
1	B	428	ASP	2.3
1	B	740	ASP	2.3
1	D	733	ASP	2.2
1	D	711	GLU	2.2
1	B	9	LEU	2.2
1	B	746	ILE	2.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	174	ASP	2.2
1	D	420	ALA	2.2
1	A	594	VAL	2.2
1	C	402	TYR	2.2
1	D	359	PHE	2.1
1	D	48	LYS	2.1
1	C	753	GLU	2.1
1	D	400	GLY	2.1
1	A	711	GLU	2.1
1	B	376	ILE	2.1
1	D	734	GLY	2.1
1	C	45	ALA	2.1
1	B	15	PHE	2.1
1	C	733	ASP	2.1
1	D	233	TRP	2.1
1	A	43	VAL	2.0
1	D	594	VAL	2.0
1	C	316	ILE	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

## 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<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	B-factors( $\text{\AA}^2$ )	Q<0.9
4	NO3	A	4012	4/4	0.81	0.16	41,41,42,43	0
3	GOL	D	4010	6/6	0.86	0.18	50,51,52,52	0
3	GOL	D	4009	6/6	0.88	0.18	26,30,31,36	0
3	GOL	B	4006	6/6	0.89	0.17	31,35,37,39	0
3	GOL	A	4005	6/6	0.90	0.13	26,28,30,33	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	GOL	C	4007	6/6	0.93	0.09	19,22,24,26	0
6	SO4	D	4011	5/5	0.93	0.14	70,71,71,71	0
3	GOL	C	4008	6/6	0.93	0.14	35,39,40,42	0
2	NDG	B	4002	15/15	0.95	0.10	17,22,25,25	0
2	NDG	C	4003	15/15	0.97	0.14	18,19,22,25	0
2	NDG	A	4001	15/15	0.97	0.13	13,18,21,23	0
2	NDG	D	4004	15/15	0.97	0.15	18,22,23,25	0
5	MG	B	4013	1/1	0.99	0.03	30,30,30,30	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.