



# Full wwPDB X-ray Structure Validation Report ⓘ

Jan 31, 2017 – 10:34 AM EST

PDB ID : 5J50  
Title : Structure of tetrameric jacalin complexed with Gal beta-(1,3) GalNAc-alpha-OPNP  
Authors : Abhinav, K.V.; Sharma, K.; Surolia, A.; Vijayan, M.  
Deposited on : 2016-04-01  
Resolution : 2.05 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-20028442
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-20028442

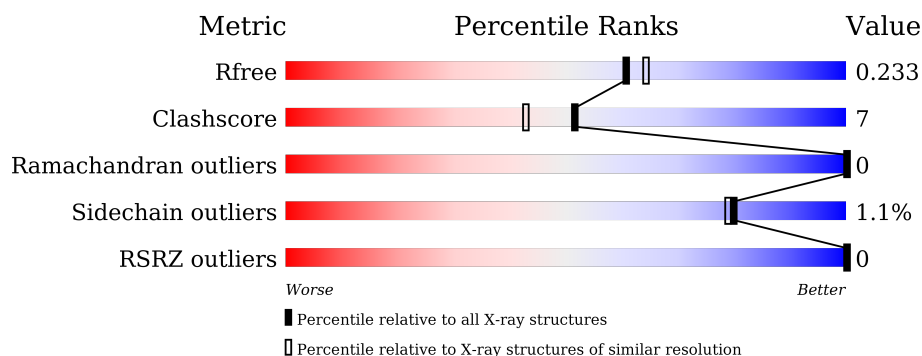
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.05 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	1192 (2.04-2.04)
Clashscore	102246	1269 (2.04-2.04)
Ramachandran outliers	100387	1258 (2.04-2.04)
Sidechain outliers	100360	1258 (2.04-2.04)
RSRZ outliers	91569	1194 (2.04-2.04)

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	133	<div> <div>90%</div> <div>10%</div> </div>
1	C	133	<div> <div>89%</div> <div>11%</div> <div>.</div> </div>
1	E	133	<div> <div>89%</div> <div>10%</div> <div>.</div> </div>
1	G	133	<div> <div>89%</div> <div>11%</div> </div>
2	B	19	<div> <div>79%</div> <div>5%</div> <div>16%</div> </div>
2	D	19	<div> <div>79%</div> <div>5%</div> <div>16%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	F	19	 79%21%
2	H	19	 63%16%21%

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
3	EDO	A	203	-	-	-	X
3	EDO	C	204	-	-	X	X
3	EDO	C	205	-	-	X	-
3	EDO	C	207	-	-	X	-
3	EDO	E	202	-	-	-	X
3	EDO	E	203	-	-	X	X
3	EDO	F	101	-	-	-	X
3	EDO	H	101	-	-	X	X
5	A2G	C	202	-	-	-	X
6	NPO	C	203	-	-	-	X

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 5143 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 Agglutinin alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	133	Total	C	N	O	S	0	0	0
			1028	672	157	197	2			
1	C	133	Total	C	N	O	S	0	0	0
			1022	669	156	195	2			
1	E	133	Total	C	N	O	S	0	0	0
			1033	675	158	198	2			
1	G	133	Total	C	N	O	S	0	0	0
			1029	672	159	196	2			

- Molecule 2 is a protein called Agglutinin beta-3 chain.

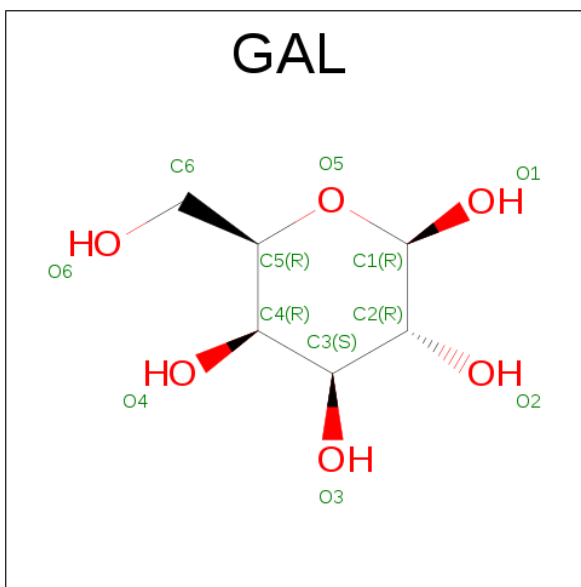
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	B	16	Total	C	N	O	0	0	0
			110	70	19	21			
2	D	16	Total	C	N	O	0	0	0
			103	66	18	19			
2	F	15	Total	C	N	O	0	0	0
			99	64	17	18			
2	H	15	Total	C	N	O	0	0	0
			97	63	16	18			

- Molecule 3 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



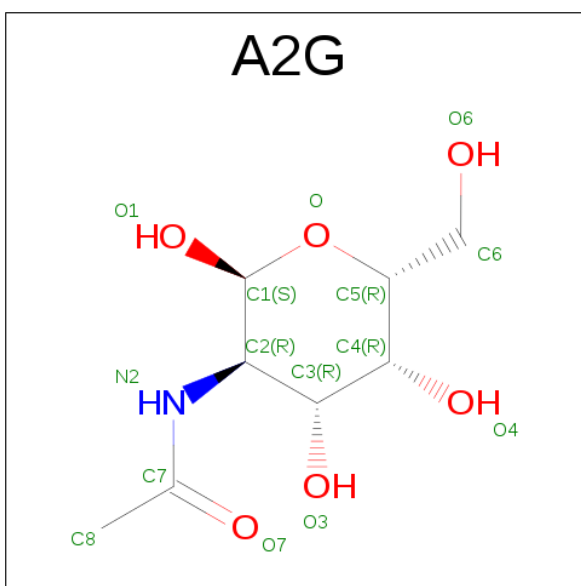
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			4	2	2		
3	A	1	Total	C	O	0	0
			4	2	2		
3	A	1	Total	C	O	0	0
			4	2	2		
3	C	1	Total	C	O	0	0
			4	2	2		
3	C	1	Total	C	O	0	0
			4	2	2		
3	C	1	Total	C	O	0	0
			4	2	2		
3	C	1	Total	C	O	0	0
			4	2	2		
3	E	1	Total	C	O	0	0
			4	2	2		
3	E	1	Total	C	O	0	0
			4	2	2		
3	F	1	Total	C	O	0	0
			4	2	2		
3	G	1	Total	C	O	0	0
			4	2	2		
3	H	1	Total	C	O	0	0
			4	2	2		

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



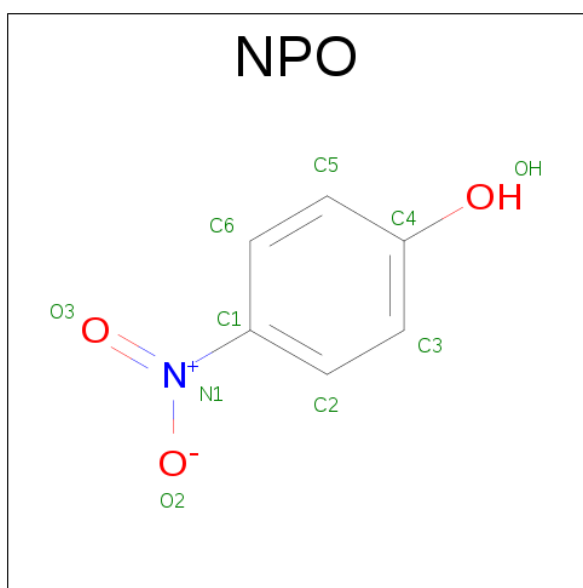
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			11	6	5		
4	C	1	Total	C	O	0	0
			11	6	5		
4	E	1	Total	C	O	0	0
			11	6	5		
4	G	1	Total	C	O	0	0
			11	6	5		

- Molecule 5 is N-ACETYL-2-DEOXY-2-AMINO-GALACTOSE (three-letter code: A2G) (formula:  $C_8H_{15}NO_6$ ).



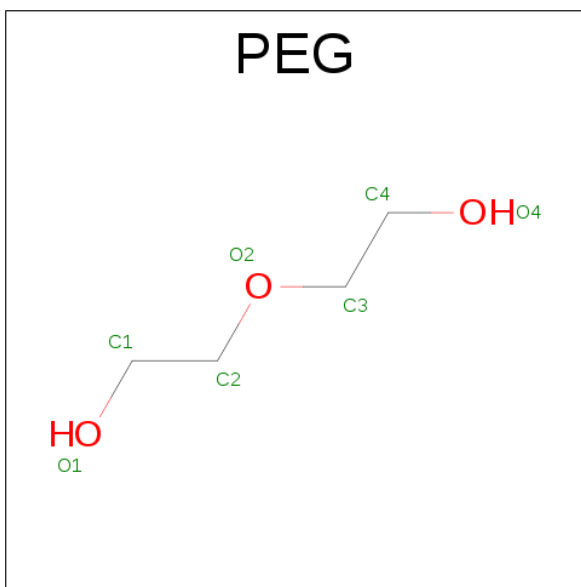
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	N	O	0	0
			14	8	1	5		
5	C	1	Total	C	N	O	0	0
			14	8	1	5		
5	E	1	Total	C	N	O	0	0
			14	8	1	5		
5	G	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 6 is P-NITROPHENOL (three-letter code: NPO) (formula:  $C_6H_5NO_3$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	1	Total	C	N	O	0	0
			10	6	1	3		
6	C	1	Total	C	N	O	0	0
			10	6	1	3		
6	E	1	Total	C	N	O	0	0
			10	6	1	3		
6	G	1	Total	C	N	O	0	0
			10	6	1	3		

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



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

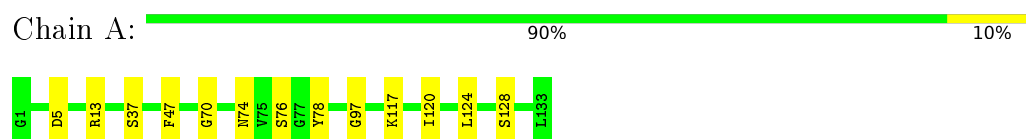
- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	87	Total	O	0	0
			87	87		
8	B	10	Total	O	0	0
			10	10		
8	C	107	Total	O	0	0
			107	107		
8	D	10	Total	O	0	0
			10	10		
8	E	103	Total	O	0	0
			103	103		
8	F	8	Total	O	0	0
			8	8		
8	G	94	Total	O	0	0
			94	94		
8	H	8	Total	O	0	0
			8	8		

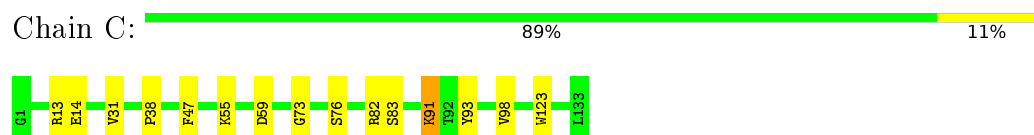
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). 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.

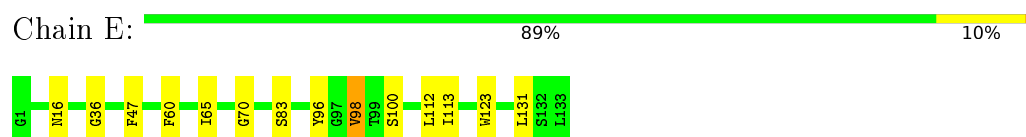
- Molecule 1: Agglutinin alpha chain



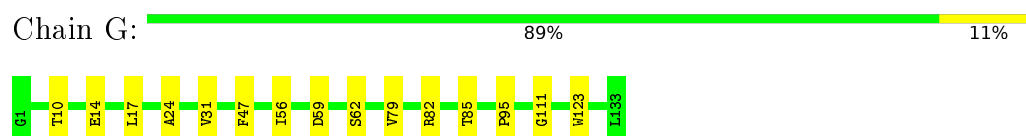
- Molecule 1: Agglutinin alpha chain



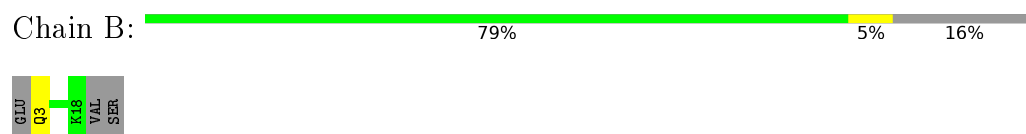
- Molecule 1: Agglutinin alpha chain



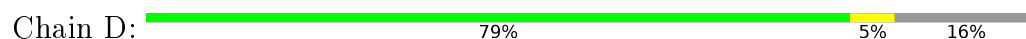
- Molecule 1: Agglutinin alpha chain

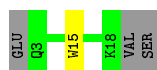


- Molecule 2: Agglutinin beta-3 chain



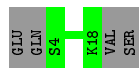
- Molecule 2: Agglutinin beta-3 chain





- Molecule 2: Agglutinin beta-3 chain

Chain F: 79% 21%



- Molecule 2: Agglutinin beta-3 chain

Chain H: 63% 16% 21%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	58.53Å 81.65Å 66.89Å 90.00° 100.00° 90.00°	Depositor
Resolution (Å)	51.27 – 2.05 51.27 – 2.05	Depositor EDS
% Data completeness (in resolution range)	99.7 (51.27-2.05) 99.7 (51.27-2.05)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.91 (at 2.05Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, $R_{free}$	0.182 , 0.233 0.181 , 0.233	Depositor DCC
$R_{free}$ test set	3885 reflections (11.07%)	DCC
Wilson B-factor (Å <sup>2</sup> )	13.1	Xtriage
Anisotropy	0.373	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 44.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.44$ , $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	5143	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	14.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.21% 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: PEG, NPO, GAL, EDO, A2G

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	0.96	0/1057	0.97	2/1436 (0.1%)
1	C	1.00	1/1051 (0.1%)	0.96	1/1428 (0.1%)
1	E	0.96	1/1062 (0.1%)	0.91	1/1441 (0.1%)
1	G	0.93	1/1058 (0.1%)	0.91	1/1436 (0.1%)
2	B	1.07	0/112	0.94	0/153
2	D	1.08	1/105 (1.0%)	0.95	0/143
2	F	1.03	0/101	0.89	0/138
2	H	0.92	0/99	0.89	0/136
All	All	0.97	4/4645 (0.1%)	0.94	5/6311 (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	C	0	1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	123	TRP	CD2-CE2	5.71	1.48	1.41
1	G	123	TRP	CD2-CE2	5.43	1.47	1.41
1	C	123	TRP	CD2-CE2	5.37	1.47	1.41
2	D	15	TRP	CD2-CE2	5.23	1.47	1.41

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	5	ASP	CB-CG-OD1	6.12	123.81	118.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	59	ASP	CB-CG-OD1	6.11	123.80	118.30
1	A	13	ARG	NE-CZ-NH2	-5.64	117.48	120.30
1	E	98	VAL	CB-CA-C	-5.52	100.91	111.40
1	G	59	ASP	CB-CG-OD1	-5.13	113.68	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	76	SER	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	1028	0	980	8	0
1	C	1022	0	970	21	0
1	E	1033	0	994	10	0
1	G	1029	0	984	10	0
2	B	110	0	105	1	0
2	D	103	0	93	0	0
2	F	99	0	92	0	0
2	H	97	0	91	5	0
3	A	12	0	18	1	0
3	C	16	0	24	19	0
3	E	8	0	12	5	0
3	F	4	0	6	0	0
3	G	4	0	6	2	0
3	H	4	0	6	4	0
4	A	11	0	10	0	0
4	C	11	0	10	0	0
4	E	11	0	10	0	0
4	G	11	0	10	0	0
5	A	14	0	12	2	0
5	C	14	0	12	3	0
5	E	14	0	12	3	0
5	G	14	0	12	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	A	10	0	4	0	0
6	C	10	0	4	2	0
6	E	10	0	4	0	0
6	G	10	0	4	0	0
7	E	7	0	10	1	0
8	A	87	0	0	2	0
8	B	10	0	0	1	0
8	C	107	0	0	4	0
8	D	10	0	0	1	0
8	E	103	0	0	2	0
8	F	8	0	0	0	0
8	G	94	0	0	1	0
8	H	8	0	0	0	0
All	All	5143	0	4495	63	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 7.

All (63) 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)
6:C:203:NPO:O3	6:C:203:NPO:N1	1.66	1.26
1:C:73:GLY:HA3	3:C:204:EDO:H11	1.46	0.94
1:C:55:LYS:HD2	3:C:205:EDO:H22	1.53	0.90
3:C:204:EDO:H21	8:C:477:HOH:O	1.78	0.82
1:C:47:PHE:HE2	5:C:202:A2G:H8B	1.48	0.77
1:C:73:GLY:CA	3:C:204:EDO:H11	2.15	0.76
1:C:55:LYS:HD2	3:C:205:EDO:C2	2.16	0.75
1:E:47:PHE:HE2	5:E:205:A2G:H8B	1.54	0.72
3:C:206:EDO:H11	8:C:480:HOH:O	1.90	0.72
2:B:3:GLN:HG2	8:B:104:HOH:O	1.91	0.71
1:E:36:GLY:HA2	3:E:203:EDO:H12	1.73	0.71
5:E:205:A2G:H8	8:E:384:HOH:O	1.91	0.70
1:A:97:GLY:H	3:A:202:EDO:H12	1.58	0.69
2:H:16:GLY:H	3:H:101:EDO:C1	2.08	0.67
2:H:16:GLY:H	3:H:101:EDO:H11	1.60	0.66
3:C:204:EDO:H12	8:C:442:HOH:O	1.95	0.66
1:C:55:LYS:CD	3:C:205:EDO:H22	2.25	0.65
1:G:47:PHE:HE2	5:G:203:A2G:H8B	1.64	0.63
1:E:36:GLY:CA	3:E:203:EDO:H12	2.32	0.59
1:C:73:GLY:HA3	3:C:204:EDO:C1	2.27	0.59

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:60:PHE:HE2	3:E:203:EDO:H21	1.70	0.57
1:C:13:ARG:HH12	3:C:207:EDO:C1	2.18	0.57
1:A:47:PHE:HE2	5:A:205:A2G:H8B	1.70	0.56
1:C:91:LYS:HE3	1:C:93:TYR:OH	2.07	0.55
1:C:13:ARG:NH1	3:C:207:EDO:H21	2.22	0.54
2:H:15:TRP:HA	3:H:101:EDO:H12	1.89	0.54
1:A:74:ASN:HA	1:A:78:TYR:O	2.09	0.53
1:C:38:PRO:HB3	3:C:206:EDO:H12	1.90	0.53
1:G:17:LEU:HD12	1:G:17:LEU:C	2.29	0.52
8:D:107:HOH:O	3:G:201:EDO:H11	2.10	0.52
1:E:83:SER:HA	1:E:96:TYR:O	2.09	0.52
5:C:202:A2G:H8A	8:C:438:HOH:O	2.10	0.52
1:G:56:ILE:O	1:G:56:ILE:HG22	2.10	0.51
1:A:47:PHE:CE2	5:A:205:A2G:H8B	2.47	0.49
1:E:70:GLY:HA2	8:E:321:HOH:O	2.12	0.49
1:C:13:ARG:HH12	3:C:207:EDO:H21	1.77	0.49
1:C:47:PHE:CE2	5:C:202:A2G:H8B	2.38	0.49
1:C:13:ARG:HH22	3:C:207:EDO:H22	1.78	0.48
1:C:13:ARG:HH12	3:C:207:EDO:C2	2.26	0.48
6:C:203:NPO:C1	6:C:203:NPO:O3	2.55	0.48
1:C:82:ARG:HH11	3:C:204:EDO:C2	2.25	0.48
1:A:120:ILE:HG12	1:A:124:LEU:HA	1.96	0.47
1:G:79:VAL:O	2:H:17:ALA:HB2	2.14	0.47
7:E:201:PEG:O1	5:E:205:A2G:H8A	2.15	0.47
1:C:55:LYS:CG	3:C:205:EDO:H22	2.46	0.46
1:G:24:ALA:HB1	1:G:82:ARG:HA	1.98	0.45
1:G:111:GLY:H	3:G:201:EDO:C2	2.29	0.45
1:C:82:ARG:HD2	3:C:204:EDO:H22	1.98	0.45
1:E:60:PHE:CE2	3:E:203:EDO:H21	2.51	0.45
1:C:14:GLU:HB3	1:C:31:VAL:HB	1.97	0.45
2:H:15:TRP:HA	3:H:101:EDO:C1	2.47	0.44
1:G:85:THR:OG1	1:G:95:PRO:HA	2.16	0.44
1:A:70:GLY:HA2	8:A:313:HOH:O	2.18	0.43
1:A:76:SER:HB3	8:A:344:HOH:O	2.19	0.43
1:G:24:ALA:CB	1:G:82:ARG:HA	2.49	0.43
1:G:10:THR:HG23	8:G:301:HOH:O	2.18	0.42
1:A:117:LYS:HG3	1:A:128:SER:HB2	2.01	0.41
1:E:16:ASN:HB3	3:E:202:EDO:H21	2.01	0.41
1:G:14:GLU:HB3	1:G:31:VAL:HB	2.00	0.41
1:C:13:ARG:HH22	3:C:207:EDO:C2	2.32	0.41
1:C:83:SER:HB2	1:C:98:VAL:O	2.20	0.41

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:65:ILE:HG21	1:E:113:ILE:HD12	2.04	0.40
1:E:112:LEU:O	1:E:131:LEU:HA	2.21	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	131/133 (98%)	125 (95%)	6 (5%)	0	100	100
1	C	131/133 (98%)	124 (95%)	7 (5%)	0	100	100
1	E	131/133 (98%)	127 (97%)	4 (3%)	0	100	100
1	G	131/133 (98%)	122 (93%)	9 (7%)	0	100	100
2	B	14/19 (74%)	14 (100%)	0	0	100	100
2	D	14/19 (74%)	13 (93%)	1 (7%)	0	100	100
2	F	13/19 (68%)	13 (100%)	0	0	100	100
2	H	13/19 (68%)	13 (100%)	0	0	100	100
All	All	578/608 (95%)	551 (95%)	27 (5%)	0	100	100

There are no Ramachandran outliers to report.

### 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	109/113 (96%)	108 (99%)	1 (1%)	84	84
1	C	107/113 (95%)	106 (99%)	1 (1%)	84	84
1	E	111/113 (98%)	109 (98%)	2 (2%)	66	62
1	G	109/113 (96%)	108 (99%)	1 (1%)	84	84
2	B	11/15 (73%)	11 (100%)	0	100	100
2	D	9/15 (60%)	9 (100%)	0	100	100
2	F	9/15 (60%)	9 (100%)	0	100	100
2	H	9/15 (60%)	9 (100%)	0	100	100
All	All	474/512 (93%)	469 (99%)	5 (1%)	80	79

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	37	SER
1	C	91	LYS
1	E	98	VAL
1	E	100	SER
1	G	62	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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 ⓘ

25 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	$\# Z  > 2$
3	EDO	A	201	-	3,3,3	0.32	0	2,2,2	1.12	0
3	EDO	A	202	-	3,3,3	0.52	0	2,2,2	0.58	0
3	EDO	A	203	-	3,3,3	0.62	0	2,2,2	0.76	0
4	GAL	A	204	5	11,11,12	1.38	2 (18%)	15,15,17	1.23	1 (6%)
5	A2G	A	205	4,6	14,14,15	1.97	5 (35%)	15,19,21	1.73	4 (26%)
6	NPO	A	206	5	8,10,10	5.11	1 (12%)	11,13,13	1.54	2 (18%)
4	GAL	C	201	5	11,11,12	1.10	1 (9%)	15,15,17	2.28	2 (13%)
5	A2G	C	202	4,6	14,14,15	2.52	2 (14%)	15,19,21	1.57	5 (33%)
6	NPO	C	203	5	8,10,10	7.82	1 (12%)	11,13,13	1.72	2 (18%)
3	EDO	C	204	-	3,3,3	0.30	0	2,2,2	1.19	0
3	EDO	C	205	-	3,3,3	0.54	0	2,2,2	0.70	0
3	EDO	C	206	-	3,3,3	0.55	0	2,2,2	0.65	0
3	EDO	C	207	-	3,3,3	0.61	0	2,2,2	0.49	0
7	PEG	E	201	-	6,6,6	0.96	0	5,5,5	0.73	0
3	EDO	E	202	-	3,3,3	0.39	0	2,2,2	1.08	0
3	EDO	E	203	-	3,3,3	0.29	0	2,2,2	1.15	0
4	GAL	E	204	5	11,11,12	1.18	1 (9%)	15,15,17	1.53	4 (26%)
5	A2G	E	205	4,6	14,14,15	2.04	4 (28%)	15,19,21	1.45	3 (20%)
6	NPO	E	206	5	8,10,10	3.65	2 (25%)	11,13,13	3.52	5 (45%)
3	EDO	F	101	-	3,3,3	0.41	0	2,2,2	1.45	0
3	EDO	G	201	-	3,3,3	0.46	0	2,2,2	1.44	0
4	GAL	G	202	5	11,11,12	1.08	1 (9%)	15,15,17	1.94	7 (46%)
5	A2G	G	203	4,6	14,14,15	2.01	3 (21%)	15,19,21	1.85	5 (33%)
6	NPO	G	204	5	8,10,10	4.07	2 (25%)	11,13,13	1.90	3 (27%)
3	EDO	H	101	-	3,3,3	0.26	0	2,2,2	1.34	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
3	EDO	A	201	-	-	0/1/1/1	0/0/0/0
3	EDO	A	202	-	-	0/1/1/1	0/0/0/0
3	EDO	A	203	-	-	0/1/1/1	0/0/0/0
4	GAL	A	204	5	-	0/2/19/22	0/1/1/1
5	A2G	A	205	4,6	-	0/6/23/26	0/1/1/1
6	NPO	A	206	5	-	0/4/4/4	0/1/1/1
4	GAL	C	201	5	-	0/2/19/22	0/1/1/1
5	A2G	C	202	4,6	-	0/6/23/26	0/1/1/1
6	NPO	C	203	5	-	0/4/4/4	0/1/1/1
3	EDO	C	204	-	-	0/1/1/1	0/0/0/0
3	EDO	C	205	-	-	0/1/1/1	0/0/0/0
3	EDO	C	206	-	-	0/1/1/1	0/0/0/0
3	EDO	C	207	-	-	0/1/1/1	0/0/0/0
7	PEG	E	201	-	-	0/4/4/4	0/0/0/0
3	EDO	E	202	-	-	0/1/1/1	0/0/0/0
3	EDO	E	203	-	-	0/1/1/1	0/0/0/0
4	GAL	E	204	5	-	0/2/19/22	0/1/1/1
5	A2G	E	205	4,6	-	0/6/23/26	0/1/1/1
6	NPO	E	206	5	-	0/4/4/4	0/1/1/1
3	EDO	F	101	-	-	0/1/1/1	0/0/0/0
3	EDO	G	201	-	-	0/1/1/1	0/0/0/0
4	GAL	G	202	5	-	0/2/19/22	0/1/1/1
5	A2G	G	203	4,6	-	0/6/23/26	0/1/1/1
6	NPO	G	204	5	-	0/4/4/4	0/1/1/1
3	EDO	H	101	-	-	0/1/1/1	0/0/0/0

All (25) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	205	A2G	O-C1	-3.34	1.38	1.43
5	E	205	A2G	C1-C2	-2.98	1.48	1.52
5	G	203	A2G	C1-C2	-2.86	1.48	1.52
5	A	205	A2G	C1-C2	-2.12	1.49	1.52
5	A	205	A2G	C2-N2	-2.10	1.42	1.46
6	E	206	NPO	C3-C4	2.11	1.43	1.38
4	A	204	GAL	O3-C3	2.12	1.47	1.43
5	E	205	A2G	C3-C2	2.16	1.57	1.52
6	G	204	NPO	OH-C4	2.19	1.42	1.37
4	C	201	GAL	O5-C5	2.32	1.48	1.43
4	G	202	GAL	O5-C5	2.36	1.48	1.43
5	A	205	A2G	C8-C7	2.57	1.55	1.50
5	G	203	A2G	O-C1	2.84	1.48	1.43
4	E	204	GAL	O5-C1	2.87	1.48	1.43

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	E	205	A2G	O4-C4	3.18	1.50	1.43
4	A	204	GAL	C4-C5	3.42	1.60	1.53
5	A	205	A2G	O7-C7	4.50	1.33	1.23
5	E	205	A2G	O7-C7	5.32	1.35	1.23
5	G	203	A2G	O7-C7	5.69	1.36	1.23
5	C	202	A2G	O7-C7	5.96	1.37	1.23
5	C	202	A2G	O-C1	6.79	1.54	1.43
6	E	206	NPO	O3-N1	9.78	1.42	1.22
6	G	204	NPO	O3-N1	11.11	1.44	1.22
6	A	206	NPO	O3-N1	14.13	1.50	1.22
6	C	203	NPO	O3-N1	22.04	1.66	1.22

All (43) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	C	201	GAL	O5-C5-C6	-7.42	91.45	107.34
6	E	206	NPO	C6-C1-N1	-6.11	115.13	119.51
6	C	203	NPO	O3-N1-C1	-4.46	115.45	118.67
6	G	204	NPO	C2-C1-N1	-4.12	116.56	119.51
4	G	202	GAL	O2-C2-C3	-3.55	103.02	110.19
5	G	203	A2G	O7-C7-C8	-3.23	116.12	122.07
5	E	205	A2G	O7-C7-C8	-2.92	116.69	122.07
4	C	201	GAL	O2-C2-C1	-2.89	103.45	109.23
5	A	205	A2G	C3-C4-C5	-2.87	105.10	110.23
5	G	203	A2G	O3-C3-C2	-2.84	103.29	109.37
4	G	202	GAL	O5-C5-C6	-2.84	101.26	107.34
4	G	202	GAL	C6-C5-C4	-2.82	105.91	112.99
5	G	203	A2G	O4-C4-C3	-2.76	104.14	110.36
4	E	204	GAL	O3-C3-C4	-2.74	104.19	110.36
6	C	203	NPO	C3-C2-C1	-2.55	116.62	120.13
4	G	202	GAL	O4-C4-C3	-2.49	104.75	110.36
5	C	202	A2G	C4-C3-C2	-2.45	107.53	111.34
5	C	202	A2G	O7-C7-C8	-2.45	117.56	122.07
6	E	206	NPO	C6-C5-C4	-2.39	117.12	119.86
6	E	206	NPO	C3-C2-C1	-2.27	117.01	120.13
5	A	205	A2G	O4-C4-C3	-2.18	105.44	110.36
5	G	203	A2G	C1-O-C5	-2.16	108.96	112.14
4	G	202	GAL	O5-C1-C2	-2.15	107.46	110.89
5	A	205	A2G	O7-C7-C8	-2.11	118.19	122.07
6	G	204	NPO	O3-N1-C1	-2.10	117.15	118.67
4	E	204	GAL	O4-C4-C3	-2.06	105.72	110.36
5	C	202	A2G	O-C5-C4	-2.01	106.80	110.13

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	G	202	GAL	O4-C4-C5	-2.00	103.94	109.23
4	E	204	GAL	C1-C2-C3	2.13	112.13	109.55
5	C	202	A2G	C8-C7-N2	2.13	120.19	116.10
5	E	205	A2G	O4-C4-C5	2.20	115.03	109.23
5	C	202	A2G	C2-N2-C7	2.22	126.00	123.11
6	A	206	NPO	C6-C1-N1	2.25	121.13	119.51
4	G	202	GAL	C3-C4-C5	2.42	114.54	110.23
5	G	203	A2G	C8-C7-N2	2.72	121.31	116.10
5	E	205	A2G	C8-C7-N2	2.88	121.62	116.10
4	A	204	GAL	C1-C2-C3	3.04	113.24	109.55
6	E	206	NPO	C5-C6-C1	3.06	124.35	120.13
6	G	204	NPO	C6-C1-N1	3.06	121.71	119.51
4	E	204	GAL	C1-O5-C5	3.20	116.85	112.14
5	A	205	A2G	C8-C7-N2	3.56	122.92	116.10
6	A	206	NPO	O3-N1-C1	3.82	121.42	118.67
6	E	206	NPO	C2-C1-N1	8.53	125.63	119.51

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

15 monomers are involved in 42 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	202	EDO	1	0
5	A	205	A2G	2	0
5	C	202	A2G	3	0
6	C	203	NPO	2	0
3	C	204	EDO	7	0
3	C	205	EDO	4	0
3	C	206	EDO	2	0
3	C	207	EDO	6	0
7	E	201	PEG	1	0
3	E	202	EDO	1	0
3	E	203	EDO	4	0
5	E	205	A2G	3	0
3	G	201	EDO	2	0
5	G	203	A2G	1	0
3	H	101	EDO	4	0

## 5.7 Other polymers

There are no such residues in this entry.

## 5.8 Polymer linkage issues

There are no chain breaks in this entry.

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	133/133 (100%)	-0.50	0 100 100	6, 12, 20, 32	0
1	C	133/133 (100%)	-0.51	0 100 100	6, 12, 17, 18	0
1	E	133/133 (100%)	-0.49	0 100 100	6, 12, 23, 27	0
1	G	133/133 (100%)	-0.53	0 100 100	6, 13, 20, 28	0
2	B	16/19 (84%)	-0.26	0 100 100	9, 14, 29, 32	0
2	D	16/19 (84%)	-0.35	0 100 100	9, 13, 26, 30	0
2	F	15/19 (78%)	-0.16	0 100 100	12, 16, 25, 29	0
2	H	15/19 (78%)	-0.44	0 100 100	8, 12, 25, 35	0
All	All	594/608 (97%)	-0.49	0 100 100	6, 12, 22, 35	0

There are no RSRZ outliers to report.

### 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( $\text{\AA}^2$ )	Q<0.9
3	EDO	H	101	4/4	0.95	0.27	22.76	20,23,23,24	0
3	EDO	E	202	4/4	0.91	0.30	19.88	22,26,27,27	0
3	EDO	C	204	4/4	0.96	0.22	10.10	21,22,23,28	0
3	EDO	A	203	4/4	0.92	0.26	9.83	27,29,30,34	0
3	EDO	F	101	4/4	0.95	0.15	5.02	18,19,19,22	0
5	A2G	C	202	14/15	0.96	0.13	3.98	11,15,16,18	0
3	EDO	E	203	4/4	0.96	0.12	3.23	20,21,21,23	0
6	NPO	C	203	10/10	0.90	0.14	2.10	17,19,23,26	0
5	A2G	G	203	14/15	0.95	0.13	1.20	14,17,20,21	0
3	EDO	G	201	4/4	0.95	0.11	0.61	16,17,18,20	0
5	A2G	A	205	14/15	0.95	0.10	-0.15	11,16,18,18	0
4	GAL	A	204	11/12	0.94	0.10	-0.39	15,18,20,20	0
5	A2G	E	205	14/15	0.95	0.09	-0.88	15,16,19,22	0
3	EDO	C	207	4/4	0.93	0.30	-	30,32,32,32	0
4	GAL	G	202	11/12	0.95	0.16	-	18,20,22,23	0
4	GAL	E	204	11/12	0.96	0.07	-	18,19,21,21	0
4	GAL	C	201	11/12	0.95	0.09	-	15,16,17,19	0
3	EDO	C	206	4/4	0.93	0.27	-	25,28,28,28	0
3	EDO	A	202	4/4	0.96	0.30	-	19,21,22,22	0
6	NPO	E	206	10/10	0.93	0.11	-	19,21,25,26	0
6	NPO	A	206	10/10	0.90	0.14	-	17,21,26,26	0
3	EDO	A	201	4/4	0.98	0.26	-	20,21,21,25	0
6	NPO	G	204	10/10	0.92	0.14	-	20,22,25,25	0
3	EDO	C	205	4/4	0.90	0.28	-	25,25,28,33	0
7	PEG	E	201	7/7	0.84	0.19	-	25,27,29,30	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.