



# wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 28, 2016 – 03:36 AM EDT

PDB ID : 1RZO  
Title : Agglutinin from Ricinus communis with galactosa  
Authors : Gabdoulkhakov, A.G.; Savochkina, Y.; Konareva, N.; Krauspenhaar, R.; Stoeva, S.; Nikonov, S.V.; Voelter, W.; Betzel, C.; Mikhailov, A.M.  
Deposited on : 2003-12-26  
Resolution : 2.63 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20027939

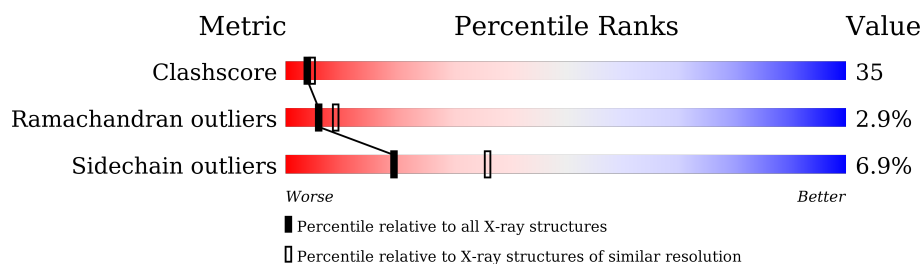
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.63 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))
Clashscore	102246	3781 (2.68-2.60)
Ramachandran outliers	100387	3722 (2.68-2.60)
Sidechain outliers	100360	3722 (2.68-2.60)

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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	262	
1	C	262	
2	B	262	
2	D	262	

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 8759 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.

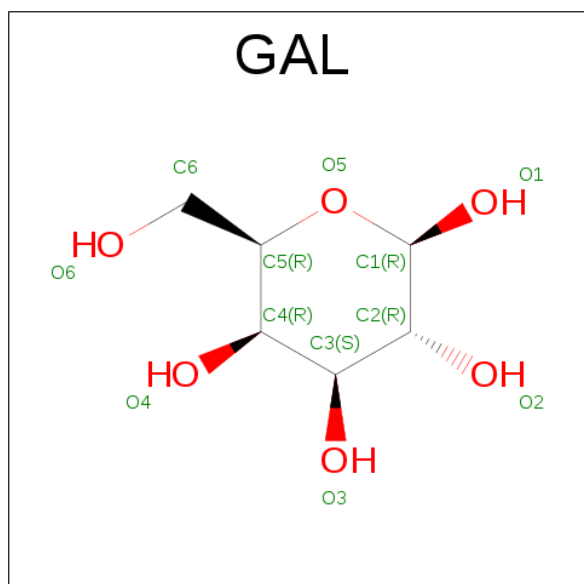
- Molecule 1 is a protein called Agglutinin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	259	Total	C	N	O	S	0	0	0
			2038	1289	356	385	8			
1	C	260	Total	C	N	O	S	0	0	0
			2045	1294	357	386	8			

- Molecule 2 is a protein called Agglutinin.

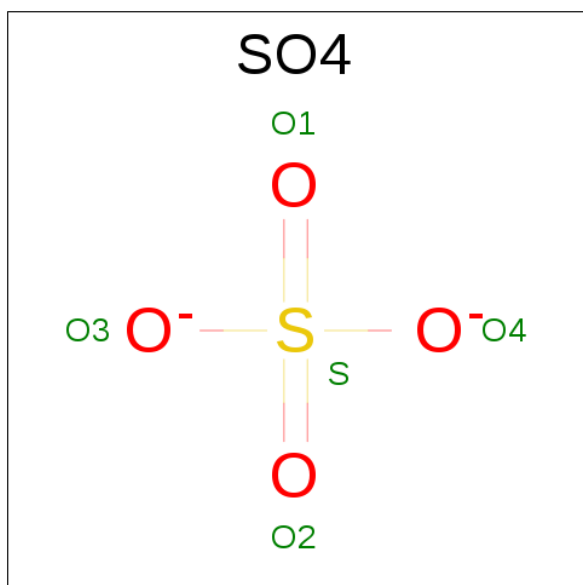
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	262	Total	C	N	O	S	0	0	0
			2044	1284	363	385	12			
2	D	262	Total	C	N	O	S	0	0	0
			2044	1284	363	385	12			

- Molecule 3 is SUGAR (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
3	A	1	Total C O 12 6 6	0	0
3	B	1	Total C O 12 6 6	0	0
3	B	1	Total C O 12 6 6	0	0
3	B	1	Total C O 12 6 6	0	0
3	B	1	Total C O 12 6 6	0	0
3	B	1	Total C O 12 6 6	0	0
3	C	1	Total C O 12 6 6	0	0
3	D	1	Total C O 12 6 6	0	0
3	D	1	Total C O 12 6 6	0	0
3	D	1	Total C O 12 6 6	0	0
3	D	1	Total C O 12 6 6	0	0

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



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

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	B	1	Total	O	S	0	0
			5	4	1		
4	B	1	Total	O	S	0	0
			5	4	1		
4	A	1	Total	O	S	0	0
			5	4	1		
4	A	1	Total	O	S	0	0
			5	4	1		
4	B	1	Total	O	S	0	0
			5	4	1		
4	C	1	Total	O	S	0	0
			5	4	1		
4	D	1	Total	O	S	0	0
			5	4	1		
4	D	1	Total	O	S	0	0
			5	4	1		
4	C	1	Total	O	S	0	0
			5	4	1		

- Molecule 5 is water.

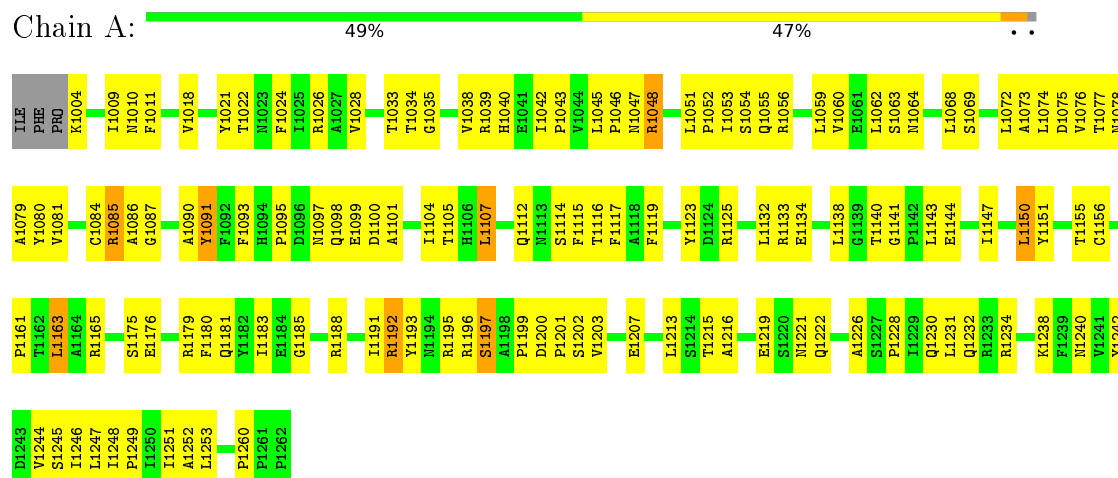
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	131	Total	O	0	0
			131	131		
5	B	117	Total	O	0	0
			117	117		
5	C	82	Total	O	0	0
			82	82		
5	D	76	Total	O	0	0
			76	76		

### 3 Residue-property plots

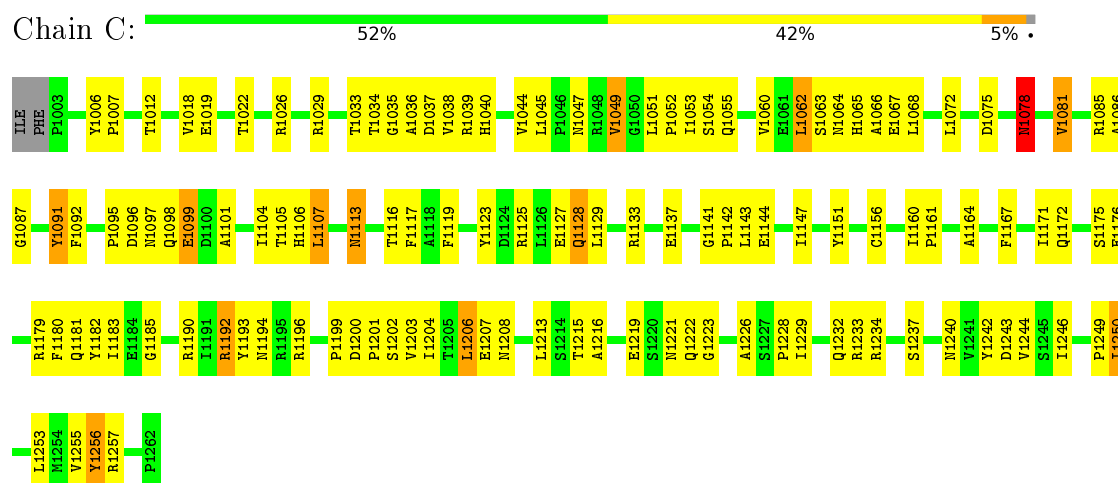
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according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

Note EDS was not executed.

#### • Molecule 1: Agglutinin

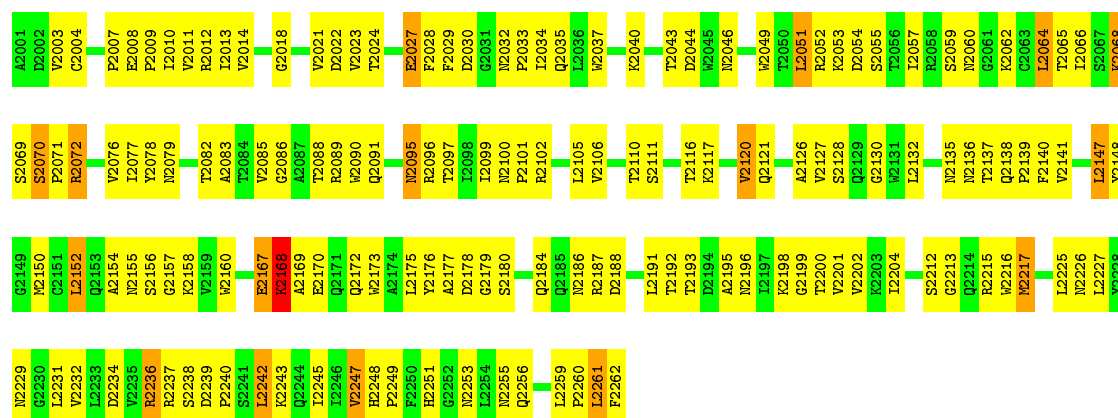


#### • Molecule 1: Agglutinin

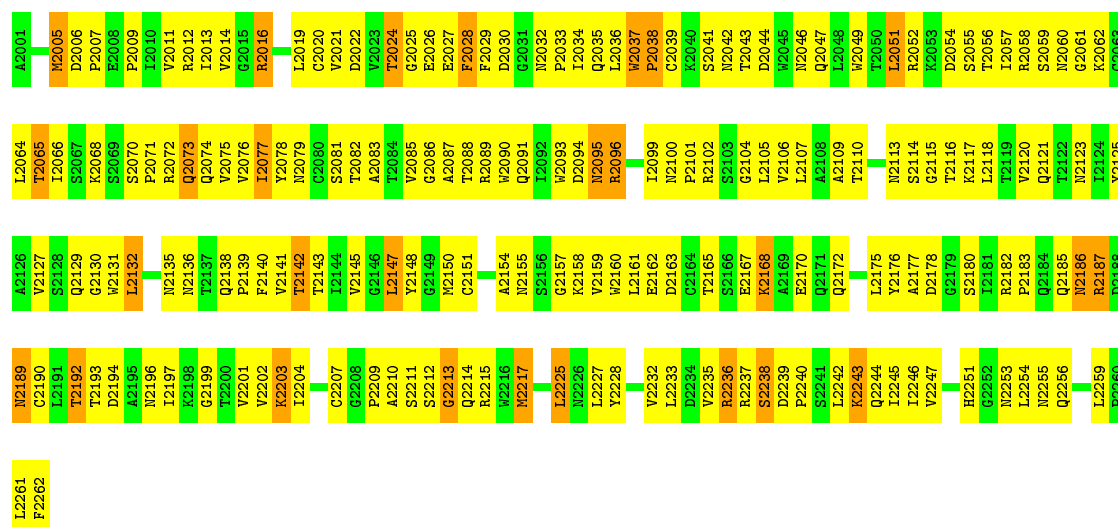


#### • Molecule 2: Agglutinin





## • Molecule 2: Agglutinin



## 4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	97.63 Å 97.63 Å 207.83 Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	26.11 – 2.63	Depositor
% Data completeness (in resolution range)	99.6 (26.11-2.63)	Depositor
$R_{merge}$	0.06	Depositor
$R_{sym}$	0.06	Depositor
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.225 , 0.276	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	8759	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	61.0	wwPDB-VP

## 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: GAL, SO4

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	0.45	0/2083	0.68	0/2836
1	C	0.43	0/2091	0.63	0/2847
2	B	0.44	0/2089	0.75	1/2850 (0.0%)
2	D	0.36	0/2089	0.71	0/2850
All	All	0.42	0/8352	0.69	1/11383 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	2064	LEU	N-CA-C	-5.09	97.25	111.00

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	2038	0	1998	110	0
1	C	2045	0	2006	122	0
2	B	2044	0	2015	149	0
2	D	2044	0	2015	212	0
3	A	12	0	11	2	0
3	B	60	0	55	9	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	C	12	0	11	0	0
3	D	48	0	45	2	0
4	A	15	0	0	0	0
4	B	15	0	0	0	0
4	C	10	0	0	0	0
4	D	10	0	0	0	0
5	A	131	0	0	5	0
5	B	117	0	0	10	0
5	C	82	0	0	1	0
5	D	76	0	0	5	0
All	All	8759	0	8156	573	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 35.

The worst 5 of 573 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:2033:PRO:HA	2:D:2076:VAL:HG12	1.40	1.02
1:C:1246:ILE:H	1:C:1246:ILE:HD12	1.29	0.97
2:D:2159:VAL:HG22	2:D:2203:LYS:HA	1.48	0.94
2:D:2072:ARG:HG2	2:D:2121:GLN:HE22	1.32	0.93
2:D:2225:LEU:HB2	2:D:2232:VAL:HG12	1.51	0.92

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	257/262 (98%)	243 (95%)	13 (5%)	1 (0%)	39 63
1	C	258/262 (98%)	229 (89%)	25 (10%)	4 (2%)	12 22

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	B	260/262 (99%)	225 (86%)	29 (11%)	6 (2%)	8	13
2	D	260/262 (99%)	199 (76%)	42 (16%)	19 (7%)	1	1
All	All	1035/1048 (99%)	896 (87%)	109 (10%)	30 (3%)	6	9

5 of 30 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	2070	SER
2	B	2167	GLU
2	B	2168	LYS
2	D	2024	THR
2	D	2073	GLN

### 5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	221/224 (99%)	211 (96%)	10 (4%)	34	59
1	C	222/224 (99%)	205 (92%)	17 (8%)	16	30
2	B	230/230 (100%)	215 (94%)	15 (6%)	21	40
2	D	230/230 (100%)	210 (91%)	20 (9%)	13	23
All	All	903/908 (99%)	841 (93%)	62 (7%)	19	36

5 of 62 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	1062	LEU
1	C	1099	GLU
2	D	2189	ASN
1	C	1081	VAL
1	C	1113	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 30 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	1040	HIS
1	C	1078	ASN
2	D	2138	GLN
1	C	1055	GLN
1	C	1113	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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

21 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	GAL	A	5001	-	12,12,12	2.06	4 (33%)	17,17,17	1.60	4 (23%)
4	SO4	A	6001	-	4,4,4	0.16	0	6,6,6	0.22	0
4	SO4	A	6004	-	4,4,4	0.35	0	6,6,6	0.06	0
4	SO4	A	6005	-	4,4,4	0.38	0	6,6,6	0.10	0
3	GAL	B	5002	-	12,12,12	2.21	4 (33%)	17,17,17	1.46	3 (17%)
3	GAL	B	5003	-	12,12,12	2.04	4 (33%)	17,17,17	1.48	4 (23%)
3	GAL	B	5004	-	12,12,12	2.19	4 (33%)	17,17,17	1.54	4 (23%)
3	GAL	B	5005	-	12,12,12	2.19	4 (33%)	17,17,17	1.62	4 (23%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	GAL	B	5006	-	12,12,12	2.25	4 (33%)	17,17,17	1.51	4 (23%)
4	SO4	B	6002	-	4,4,4	0.27	0	6,6,6	0.10	0
4	SO4	B	6003	-	4,4,4	0.19	0	6,6,6	0.12	0
4	SO4	B	6006	-	4,4,4	0.25	0	6,6,6	0.07	0
3	GAL	C	5501	-	12,12,12	2.21	4 (33%)	17,17,17	1.54	4 (23%)
4	SO4	C	6501	-	4,4,4	0.23	0	6,6,6	0.06	0
4	SO4	C	6504	-	4,4,4	0.23	0	6,6,6	0.24	0
3	GAL	D	5502	-	12,12,12	2.30	4 (33%)	17,17,17	1.45	4 (23%)
3	GAL	D	5503	-	12,12,12	2.31	4 (33%)	17,17,17	1.45	4 (23%)
3	GAL	D	5504	-	12,12,12	2.15	4 (33%)	17,17,17	1.51	4 (23%)
3	GAL	D	5505	-	12,12,12	2.23	4 (33%)	17,17,17	1.32	2 (11%)
4	SO4	D	6502	-	4,4,4	0.17	0	6,6,6	0.09	0
4	SO4	D	6503	-	4,4,4	0.18	0	6,6,6	0.09	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	GAL	A	5001	-	-	0/2/22/22	0/1/1/1
4	SO4	A	6001	-	-	0/0/0/0	0/0/0/0
4	SO4	A	6004	-	-	0/0/0/0	0/0/0/0
4	SO4	A	6005	-	-	0/0/0/0	0/0/0/0
3	GAL	B	5002	-	-	0/2/22/22	0/1/1/1
3	GAL	B	5003	-	-	0/2/22/22	0/1/1/1
3	GAL	B	5004	-	-	0/2/22/22	0/1/1/1
3	GAL	B	5005	-	-	0/2/22/22	0/1/1/1
3	GAL	B	5006	-	-	0/2/22/22	0/1/1/1
4	SO4	B	6002	-	-	0/0/0/0	0/0/0/0
4	SO4	B	6003	-	-	0/0/0/0	0/0/0/0
4	SO4	B	6006	-	-	0/0/0/0	0/0/0/0
3	GAL	C	5501	-	-	0/2/22/22	0/1/1/1
4	SO4	C	6501	-	-	0/0/0/0	0/0/0/0
4	SO4	C	6504	-	-	0/0/0/0	0/0/0/0
3	GAL	D	5502	-	-	0/2/22/22	0/1/1/1
3	GAL	D	5503	-	-	0/2/22/22	0/1/1/1
3	GAL	D	5504	-	-	0/2/22/22	0/1/1/1
3	GAL	D	5505	-	-	0/2/22/22	0/1/1/1
4	SO4	D	6502	-	-	0/0/0/0	0/0/0/0
4	SO4	D	6503	-	-	0/0/0/0	0/0/0/0

The worst 5 of 44 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	5502	GAL	C3-C2	-4.97	1.39	1.52
3	D	5503	GAL	C3-C2	-4.89	1.39	1.52
3	D	5505	GAL	C3-C2	-4.89	1.39	1.52
3	B	5003	GAL	C3-C2	-4.89	1.39	1.52
3	B	5005	GAL	C3-C2	-4.82	1.39	1.52

The worst 5 of 41 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	5001	GAL	O3-C3-C4	-3.73	101.94	110.36
3	B	5004	GAL	O3-C3-C4	-3.63	102.18	110.36
3	C	5501	GAL	O3-C3-C4	-3.55	102.36	110.36
3	B	5005	GAL	O3-C3-C4	-3.49	102.50	110.36
3	B	5003	GAL	O3-C3-C4	-3.38	102.73	110.36

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

7 monomers are involved in 13 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	5001	GAL	2	0
3	B	5002	GAL	3	0
3	B	5004	GAL	2	0
3	B	5005	GAL	4	0
3	B	5006	GAL	1	0
3	D	5502	GAL	1	0
3	D	5503	GAL	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 [i](#)

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

EDS was not executed - this section will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

### 6.3 Carbohydrates [i](#)

EDS was not executed - this section will therefore be empty.

### 6.4 Ligands [i](#)

EDS was not executed - this section will therefore be empty.

### 6.5 Other polymers [i](#)

EDS was not executed - this section will therefore be empty.