



# wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 27, 2017 – 04:26 PM EDT

PDB ID : 5EWS  
Title : Sugar binding protein - human galectin-2  
Authors : Su, J.Y.; Si, Y.L.  
Deposited on : unknown  
Resolution : 2.00 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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20030345  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
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-20030345

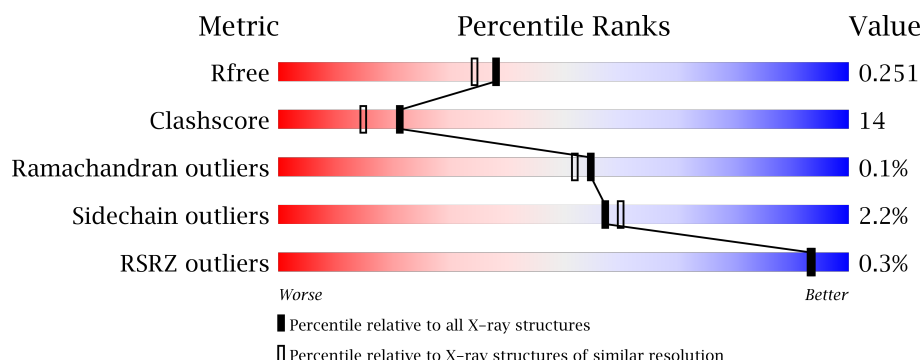
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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.




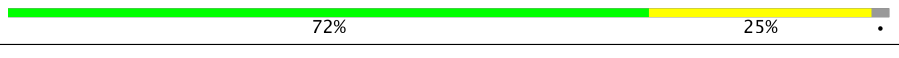

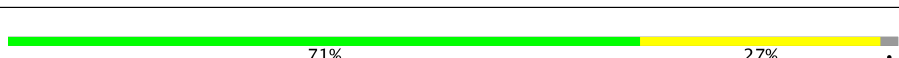


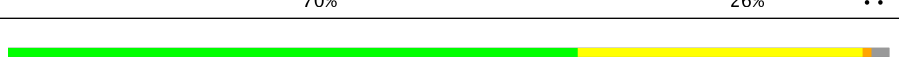

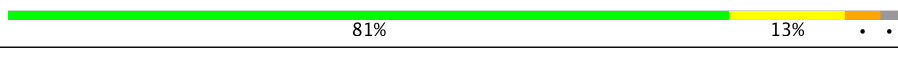
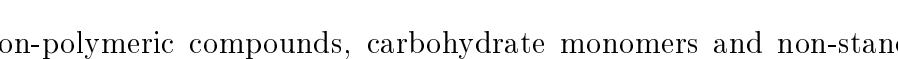
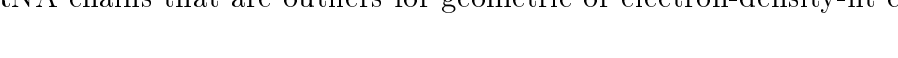
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	100719	6609 (2.00-2.00)
Clashscore	112137	7775 (2.00-2.00)
Ramachandran outliers	110173	7679 (2.00-2.00)
Sidechain outliers	110143	7678 (2.00-2.00)
RSRZ outliers	101464	6696 (2.00-2.00)

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	134	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, orange 1%, yellow 68%, green 28%, grey 3%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>68%</span> <span>28%</span> <span>..</span> </div> </div>
1	B	134	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 0%, orange 0%, yellow 78%, green 19%, grey 3%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>78%</span> <span>19%</span> <span>..</span> </div> </div>
1	C	134	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 0%, orange 0%, yellow 63%, green 33%, grey 4%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>63%</span> <span>33%</span> <span>..</span> </div> </div>
1	D	134	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 0%, orange 0%, yellow 73%, green 23%, grey 4%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>73%</span> <span>23%</span> <span>..</span> </div> </div>
1	E	134	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 0%, orange 0%, yellow 72%, green 25%, grey 3%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>72%</span> <span>25%</span> <span>..</span> </div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	134	
1	G	134	
1	H	134	
1	I	134	
1	J	134	
1	K	134	
1	L	134	
1	M	134	
1	N	134	
1	O	134	
1	P	134	

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
2	GAL	D	201	-	-	-	X
2	GAL	H	201	-	-	-	X
2	GAL	J	201	-	-	-	X
2	GAL	M	201	-	-	-	X
2	GAL	N	201	-	-	-	X
3	BGC	D	202	-	-	-	X
3	BGC	J	202	-	-	-	X
3	BGC	M	202	-	-	X	X
3	BGC	N	202	-	-	-	X

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 16873 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 Galectin-2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	131	Total	C	N	O	S	0	0	0
			1012	637	173	196	6			
1	A	131	Total	C	N	O	S	0	0	0
			1017	640	175	196	6			
1	C	131	Total	C	N	O	S	0	0	0
			1017	640	175	196	6			
1	D	131	Total	C	N	O	S	0	0	0
			1014	639	175	194	6			
1	E	131	Total	C	N	O	S	0	0	0
			1017	640	175	196	6			
1	F	131	Total	C	N	O	S	0	0	0
			1017	640	175	196	6			
1	G	131	Total	C	N	O	S	0	0	0
			1017	640	175	196	6			
1	H	131	Total	C	N	O	S	0	0	0
			1017	640	175	196	6			
1	I	131	Total	C	N	O	S	0	0	0
			1017	640	175	196	6			
1	J	131	Total	C	N	O	S	0	0	0
			1017	640	175	196	6			
1	K	131	Total	C	N	O	S	0	0	0
			1017	640	175	196	6			
1	L	131	Total	C	N	O	S	0	0	0
			1017	640	175	196	6			
1	M	131	Total	C	N	O	S	0	0	0
			1017	640	175	196	6			
1	N	131	Total	C	N	O	S	0	0	0
			1017	640	175	196	6			
1	O	131	Total	C	N	O	S	0	0	0
			1017	640	175	196	6			
1	P	131	Total	C	N	O	S	0	0	0
			1017	640	175	196	6			

There are 48 discrepancies between the modelled and reference sequences:

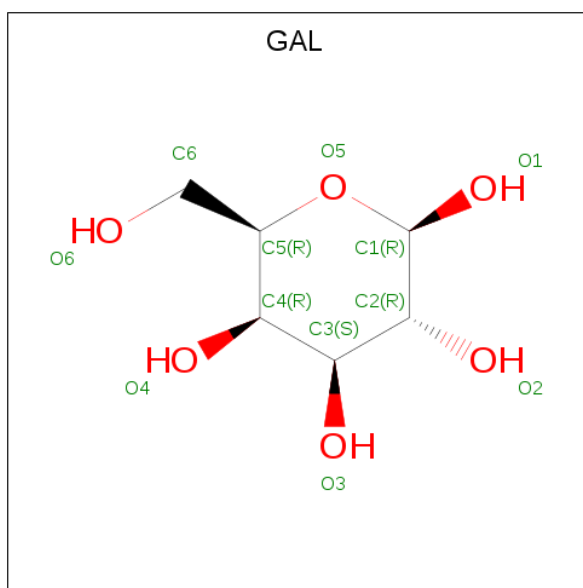
Chain	Residue	Modelled	Actual	Comment	Reference
B	1	GLY	-	expression tag	UNP P05162
B	2	SER	-	expression tag	UNP P05162
B	3	HIS	-	expression tag	UNP P05162
A	1	GLY	-	expression tag	UNP P05162
A	2	SER	-	expression tag	UNP P05162
A	3	HIS	-	expression tag	UNP P05162
C	1	GLY	-	expression tag	UNP P05162
C	2	SER	-	expression tag	UNP P05162
C	3	HIS	-	expression tag	UNP P05162
D	1	GLY	-	expression tag	UNP P05162
D	2	SER	-	expression tag	UNP P05162
D	3	HIS	-	expression tag	UNP P05162
E	1	GLY	-	expression tag	UNP P05162
E	2	SER	-	expression tag	UNP P05162
E	3	HIS	-	expression tag	UNP P05162
F	1	GLY	-	expression tag	UNP P05162
F	2	SER	-	expression tag	UNP P05162
F	3	HIS	-	expression tag	UNP P05162
G	1	GLY	-	expression tag	UNP P05162
G	2	SER	-	expression tag	UNP P05162
G	3	HIS	-	expression tag	UNP P05162
H	1	GLY	-	expression tag	UNP P05162
H	2	SER	-	expression tag	UNP P05162
H	3	HIS	-	expression tag	UNP P05162
I	1	GLY	-	expression tag	UNP P05162
I	2	SER	-	expression tag	UNP P05162
I	3	HIS	-	expression tag	UNP P05162
J	1	GLY	-	expression tag	UNP P05162
J	2	SER	-	expression tag	UNP P05162
J	3	HIS	-	expression tag	UNP P05162
K	1	GLY	-	expression tag	UNP P05162
K	2	SER	-	expression tag	UNP P05162
K	3	HIS	-	expression tag	UNP P05162
L	1	GLY	-	expression tag	UNP P05162
L	2	SER	-	expression tag	UNP P05162
L	3	HIS	-	expression tag	UNP P05162
M	1	GLY	-	expression tag	UNP P05162
M	2	SER	-	expression tag	UNP P05162
M	3	HIS	-	expression tag	UNP P05162
N	1	GLY	-	expression tag	UNP P05162
N	2	SER	-	expression tag	UNP P05162
N	3	HIS	-	expression tag	UNP P05162

*Continued on next page...*

*Continued from previous page...*

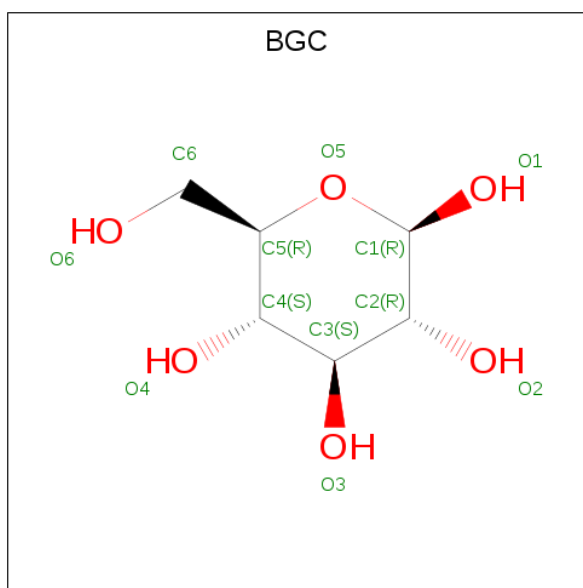
Chain	Residue	Modelled	Actual	Comment	Reference
O	1	GLY	-	expression tag	UNP P05162
O	2	SER	-	expression tag	UNP P05162
O	3	HIS	-	expression tag	UNP P05162
P	1	GLY	-	expression tag	UNP P05162
P	2	SER	-	expression tag	UNP P05162
P	3	HIS	-	expression tag	UNP P05162

- Molecule 2 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
2	C	1	Total C O 11 6 5	0	0
2	D	1	Total C O 11 6 5	0	0
2	G	1	Total C O 11 6 5	0	0
2	H	1	Total C O 11 6 5	0	0
2	I	1	Total C O 11 6 5	0	0
2	J	1	Total C O 11 6 5	0	0
2	M	1	Total C O 11 6 5	0	0
2	N	1	Total C O 11 6 5	0	0

- Molecule 3 is BETA-D-GLUCOSE (three-letter code: BGC) (formula:  $C_6H_{12}O_6$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	C	1	Total	C	O	0	0
			12	6	6		
3	D	1	Total	C	O	0	0
			12	6	6		
3	G	1	Total	C	O	0	0
			12	6	6		
3	H	1	Total	C	O	0	0
			12	6	6		
3	I	1	Total	C	O	0	0
			12	6	6		
3	J	1	Total	C	O	0	0
			12	6	6		
3	M	1	Total	C	O	0	0
			12	6	6		
3	N	1	Total	C	O	0	0
			12	6	6		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	22	Total	O	0	0
			22	22		
4	A	29	Total	O	0	0
			29	29		
4	C	24	Total	O	0	0
			24	24		

*Continued on next page...*

*Continued from previous page...*


Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	D	28	Total 28	O 28	0	0
4	E	29	Total 29	O 29	0	0
4	F	16	Total 16	O 16	0	0
4	G	25	Total 25	O 25	0	0
4	H	27	Total 27	O 27	0	0
4	I	40	Total 40	O 40	0	0
4	J	32	Total 32	O 32	0	0
4	K	24	Total 24	O 24	0	0
4	L	21	Total 21	O 21	0	0
4	M	36	Total 36	O 36	0	0
4	N	34	Total 34	O 34	0	0
4	O	23	Total 23	O 23	0	0
4	P	15	Total 15	O 15	0	0



### 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 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.

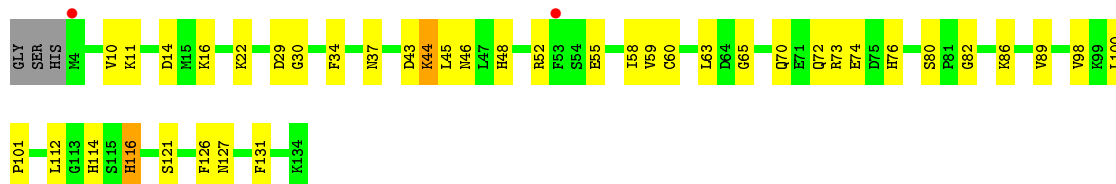
#### • Molecule 1: Galectin-2

Chain B: 



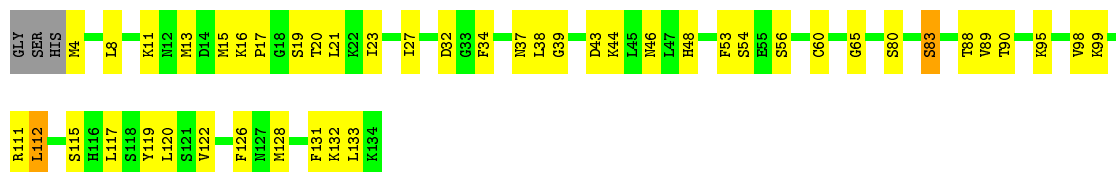
#### • Molecule 1: Galectin-2

Chain A: 



#### • Molecule 1: Galectin-2

Chain C: 



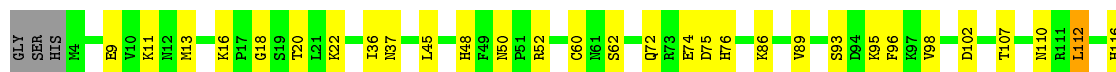
#### • Molecule 1: Galectin-2

Chain D: 



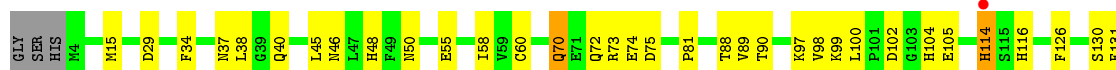
#### • Molecule 1: Galectin-2

Chain E: 





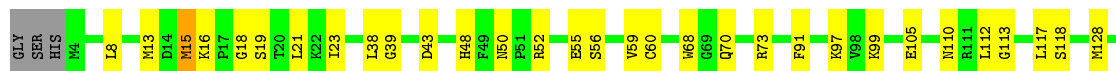
- Molecule 1: Galectin-2



- Molecule 1: Galectin-2



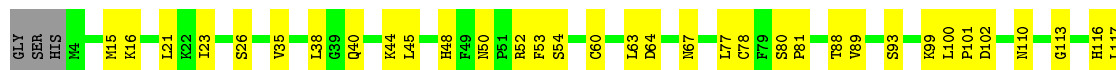
- Molecule 1: Galectin-2



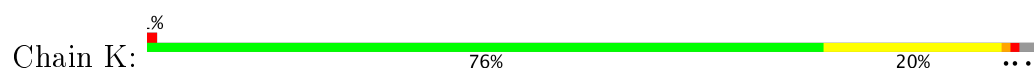
- Molecule 1: Galectin-2



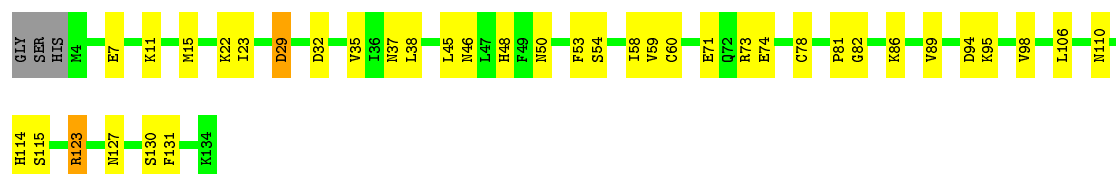
- Molecule 1: Galectin-2



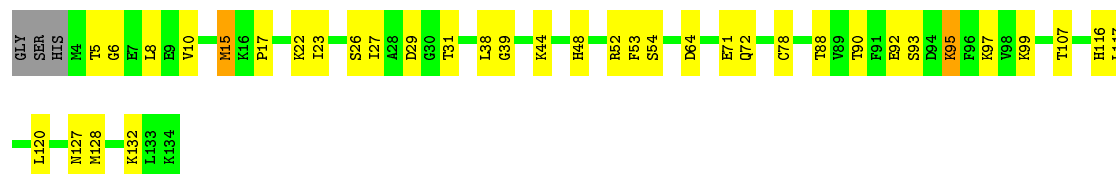
- Molecule 1: Galectin-2



- Molecule 1: Galectin-2



- Molecule 1: Galectin-2



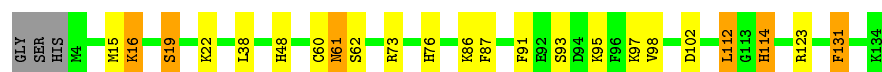
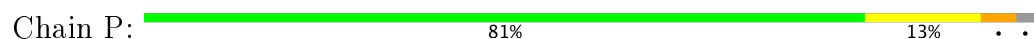
- Molecule 1: Galectin-2



- Molecule 1: Galectin-2



- Molecule 1: Galectin-2



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	38.40 Å 106.94 Å 121.57 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.95 – 2.00 48.95 – 2.00	Depositor EDS
% Data completeness (in resolution range)	97.6 (48.95-2.00) 92.7 (48.95-2.00)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.94 (at 2.00 Å)	Xtriage
Refinement program	PHENIX 1.10.1_2155	Depositor
R, $R_{free}$	0.214 , 0.249 0.217 , 0.251	Depositor DCC
$R_{free}$ test set	1901 reflections (1.57%)	DCC
Wilson B-factor (Å <sup>2</sup> )	25.0	Xtriage
Anisotropy	0.495	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 32.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.44$ , $\langle L^2 \rangle = 0.26$	Xtriage
Estimated twinning fraction	0.469 for h,-k,-l 0.459 for -h,k,-l 0.458 for -h,-k,l	Xtriage
Reported twinning fraction	0.490 for -h,-k,l	Depositor
Outliers	0 of 127533 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	16873	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 34.02 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 7.2946e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<sup>1</sup> Intensities estimated from amplitudes.

<sup>2</sup> Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 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: BGC, GAL

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	0.33	0/1039	0.58	0/1395
1	B	0.26	0/1033	0.50	0/1387
1	C	0.29	0/1039	0.56	0/1395
1	D	0.30	0/1036	0.54	0/1391
1	E	0.30	0/1039	0.60	1/1395 (0.1%)
1	F	0.31	0/1039	0.55	0/1395
1	G	0.27	0/1039	0.52	0/1395
1	H	0.30	0/1039	0.57	2/1395 (0.1%)
1	I	0.28	0/1039	0.54	0/1395
1	J	0.33	0/1039	0.54	0/1395
1	K	0.26	0/1039	0.56	0/1395
1	L	0.28	0/1039	0.54	0/1395
1	M	0.31	0/1039	0.55	0/1395
1	N	0.34	0/1039	0.56	0/1395
1	O	0.27	0/1039	0.53	0/1395
1	P	0.27	0/1039	0.60	1/1395 (0.1%)
All	All	0.30	0/16615	0.55	4/22308 (0.0%)

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	K	0	1
1	P	0	1
All	All	0	2

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	P	112	LEU	CB-CG-CD2	-8.56	96.44	111.00
1	E	112	LEU	CB-CG-CD2	-8.01	97.38	111.00
1	H	15	MET	C-N-CA	5.61	135.74	121.70
1	H	43	ASP	CB-CG-OD1	-5.09	113.72	118.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	K	114	HIS	Peptide
1	P	114	HIS	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	1017	0	988	42	0
1	B	1012	0	983	23	0
1	C	1017	0	988	38	0
1	D	1014	0	986	24	0
1	E	1017	0	988	28	0
1	F	1017	0	988	25	0
1	G	1017	0	988	22	0
1	H	1017	0	988	26	0
1	I	1017	0	988	24	0
1	J	1017	0	988	21	0
1	K	1017	0	988	33	0
1	L	1017	0	988	29	0
1	M	1017	0	988	31	0
1	N	1017	0	988	46	0
1	O	1017	0	988	33	0
1	P	1017	0	988	20	0
2	C	11	0	10	1	0
2	D	11	0	10	2	0
2	G	11	0	10	0	0
2	H	11	0	10	4	0
2	I	11	0	10	1	0
2	J	11	0	10	3	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	M	11	0	10	4	0
2	N	11	0	10	3	0
3	C	12	0	10	1	0
3	D	12	0	10	3	0
3	G	12	0	10	0	0
3	H	12	0	10	1	0
3	I	12	0	10	1	0
3	J	12	0	10	3	0
3	M	12	0	10	6	0
3	N	12	0	10	2	0
4	A	29	0	0	15	0
4	B	22	0	0	7	0
4	C	24	0	0	5	0
4	D	28	0	0	4	0
4	E	29	0	0	13	0
4	F	16	0	0	6	0
4	G	25	0	0	4	0
4	H	27	0	0	2	0
4	I	40	0	0	7	0
4	J	32	0	0	5	0
4	K	24	0	0	8	0
4	L	21	0	0	9	1
4	M	36	0	0	8	0
4	N	34	0	0	7	1
4	O	23	0	0	10	0
4	P	15	0	0	4	0
All	All	16873	0	15961	442	1

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 14.

The worst 5 of 442 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:O:93:SER:HB2	1:O:114:HIS:HB2	1.44	0.98
1:A:43:ASP:OD2	1:A:44:LYS:HE3	1.65	0.94
1:I:132:LYS:HG2	1:O:130:SER:HB2	1.54	0.89
1:A:55:GLU:HG2	3:D:202:BGC:H6C2	1.58	0.84
1:M:52:ARG:NH2	2:M:201:GAL:O4	2.10	0.83

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the sym-

metry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:L:212:HOH:O	4:N:311:HOH:O[1_454]	1.76	0.44

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	129/134 (96%)	125 (97%)	4 (3%)	0	100	100
1	B	129/134 (96%)	126 (98%)	3 (2%)	0	100	100
1	C	129/134 (96%)	127 (98%)	2 (2%)	0	100	100
1	D	129/134 (96%)	127 (98%)	2 (2%)	0	100	100
1	E	129/134 (96%)	124 (96%)	4 (3%)	1 (1%)	22	15
1	F	129/134 (96%)	127 (98%)	2 (2%)	0	100	100
1	G	129/134 (96%)	127 (98%)	2 (2%)	0	100	100
1	H	129/134 (96%)	128 (99%)	1 (1%)	0	100	100
1	I	129/134 (96%)	127 (98%)	2 (2%)	0	100	100
1	J	129/134 (96%)	126 (98%)	3 (2%)	0	100	100
1	K	129/134 (96%)	123 (95%)	4 (3%)	2 (2%)	11	5
1	L	129/134 (96%)	127 (98%)	2 (2%)	0	100	100
1	M	129/134 (96%)	126 (98%)	3 (2%)	0	100	100
1	N	129/134 (96%)	127 (98%)	2 (2%)	0	100	100
1	O	129/134 (96%)	128 (99%)	1 (1%)	0	100	100
1	P	129/134 (96%)	126 (98%)	3 (2%)	0	100	100
All	All	2064/2144 (96%)	2021 (98%)	40 (2%)	3 (0%)	55	52

All (3) Ramachandran outliers are listed below:



Mol	Chain	Res	Type
1	K	114	HIS
1	E	116	HIS
1	K	115	SER

### 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	116/118 (98%)	113 (97%)	3 (3%)	51	52
1	B	115/118 (98%)	114 (99%)	1 (1%)	82	87
1	C	116/118 (98%)	112 (97%)	4 (3%)	42	40
1	D	115/118 (98%)	113 (98%)	2 (2%)	66	70
1	E	116/118 (98%)	116 (100%)	0	100	100
1	F	116/118 (98%)	112 (97%)	4 (3%)	42	40
1	G	116/118 (98%)	114 (98%)	2 (2%)	66	70
1	H	116/118 (98%)	114 (98%)	2 (2%)	66	70
1	I	116/118 (98%)	113 (97%)	3 (3%)	51	52
1	J	116/118 (98%)	115 (99%)	1 (1%)	82	87
1	K	116/118 (98%)	116 (100%)	0	100	100
1	L	116/118 (98%)	112 (97%)	4 (3%)	42	40
1	M	116/118 (98%)	113 (97%)	3 (3%)	51	52
1	N	116/118 (98%)	113 (97%)	3 (3%)	51	52
1	O	116/118 (98%)	112 (97%)	4 (3%)	42	40
1	P	116/118 (98%)	112 (97%)	4 (3%)	42	40
All	All	1854/1888 (98%)	1814 (98%)	40 (2%)	57	60

5 of 40 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	I	14	ASP
1	L	29	ASP
1	P	16	LYS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	I	131	PHE
1	L	95	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 47 such sidechains are listed below:

Mol	Chain	Res	Type
1	H	76	HIS
1	K	67	ASN
1	P	61	ASN
1	J	50	ASN
1	K	72	GLN

### 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](#)

16 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	$\# Z  > 2$
2	GAL	C	201	3	11,11,12	1.71	3 (27%)	13,15,17	3.60	5 (38%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	BGC	C	202	2	12,12,12	2.61	6 (50%)	17,17,17	0.84	1 (5%)
2	GAL	D	201	3	11,11,12	1.66	2 (18%)	13,15,17	2.71	5 (38%)
3	BGC	D	202	2	12,12,12	2.59	5 (41%)	17,17,17	1.00	1 (5%)
2	GAL	G	201	3	11,11,12	1.66	2 (18%)	13,15,17	1.24	1 (7%)
3	BGC	G	202	2	12,12,12	2.53	6 (50%)	17,17,17	0.70	0
2	GAL	H	201	3	11,11,12	1.64	2 (18%)	13,15,17	2.99	5 (38%)
3	BGC	H	202	2	12,12,12	2.48	6 (50%)	17,17,17	1.11	2 (11%)
2	GAL	I	201	3	11,11,12	1.63	2 (18%)	13,15,17	3.44	5 (38%)
3	BGC	I	202	2	12,12,12	2.51	5 (41%)	17,17,17	0.85	0
2	GAL	J	201	3	11,11,12	1.69	3 (27%)	13,15,17	2.62	5 (38%)
3	BGC	J	202	2	12,12,12	2.58	5 (41%)	17,17,17	0.94	1 (5%)
2	GAL	M	201	3	11,11,12	2.08	3 (27%)	13,15,17	3.16	5 (38%)
3	BGC	M	202	2	12,12,12	2.23	5 (41%)	17,17,17	1.59	4 (23%)
2	GAL	N	201	3	11,11,12	1.62	2 (18%)	13,15,17	2.79	5 (38%)
3	BGC	N	202	2	12,12,12	2.56	5 (41%)	17,17,17	0.98	1 (5%)

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
2	GAL	C	201	3	-	0/2/19/22	0/1/1/1
3	BGC	C	202	2	-	0/2/22/22	0/1/1/1
2	GAL	D	201	3	-	0/2/19/22	0/1/1/1
3	BGC	D	202	2	-	0/2/22/22	0/1/1/1
2	GAL	G	201	3	-	0/2/19/22	0/1/1/1
3	BGC	G	202	2	-	0/2/22/22	0/1/1/1
2	GAL	H	201	3	-	0/2/19/22	0/1/1/1
3	BGC	H	202	2	-	0/2/22/22	0/1/1/1
2	GAL	I	201	3	-	0/2/19/22	0/1/1/1
3	BGC	I	202	2	-	0/2/22/22	0/1/1/1
2	GAL	J	201	3	-	0/2/19/22	0/1/1/1
3	BGC	J	202	2	-	0/2/22/22	0/1/1/1
2	GAL	M	201	3	-	0/2/19/22	0/1/1/1
3	BGC	M	202	2	-	0/2/22/22	0/1/1/1
2	GAL	N	201	3	-	0/2/19/22	0/1/1/1
3	BGC	N	202	2	-	0/2/22/22	0/1/1/1

The worst 5 of 62 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	M	201	GAL	C2-C3	-4.98	1.45	1.52
2	H	201	GAL	C2-C3	-3.21	1.48	1.52
2	J	201	GAL	C2-C3	-3.14	1.48	1.52
2	D	201	GAL	C2-C3	-3.09	1.48	1.52
2	C	201	GAL	C2-C3	-3.04	1.48	1.52

The worst 5 of 46 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	M	201	GAL	C2-C3-C4	-3.01	105.63	110.88
3	M	202	BGC	O5-C1-C2	-2.65	105.64	110.04
3	H	202	BGC	O5-C1-C2	-2.05	106.64	110.04
2	J	201	GAL	C2-C3-C4	2.12	114.57	110.88
2	N	201	GAL	C1-O5-C5	2.15	115.13	112.17

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

14 monomers are involved in 27 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	201	GAL	1	0
3	C	202	BGC	1	0
2	D	201	GAL	2	0
3	D	202	BGC	3	0
2	H	201	GAL	4	0
3	H	202	BGC	1	0
2	I	201	GAL	1	0
3	I	202	BGC	1	0
2	J	201	GAL	3	0
3	J	202	BGC	3	0
2	M	201	GAL	4	0
3	M	202	BGC	6	0
2	N	201	GAL	3	0
3	N	202	BGC	2	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 ⓘ

### 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	131/134 (97%)	-0.20	2 (1%) 74 73	19, 34, 48, 60	0
1	B	131/134 (97%)	-0.28	0 100 100	16, 31, 41, 47	0
1	C	131/134 (97%)	-0.39	0 100 100	17, 30, 43, 49	0
1	D	131/134 (97%)	-0.27	0 100 100	20, 32, 43, 54	0
1	E	131/134 (97%)	-0.35	0 100 100	16, 28, 41, 60	0
1	F	131/134 (97%)	-0.13	1 (0%) 86 85	17, 35, 48, 52	0
1	G	131/134 (97%)	-0.26	0 100 100	20, 31, 44, 59	0
1	H	131/134 (97%)	-0.38	0 100 100	19, 30, 42, 50	0
1	I	131/134 (97%)	-0.35	0 100 100	18, 29, 39, 47	0
1	J	131/134 (97%)	-0.32	0 100 100	20, 32, 45, 52	0
1	K	131/134 (97%)	-0.36	1 (0%) 86 85	14, 28, 40, 60	0
1	L	131/134 (97%)	-0.08	0 100 100	18, 36, 50, 53	0
1	M	131/134 (97%)	-0.38	0 100 100	18, 29, 40, 45	0
1	N	131/134 (97%)	-0.27	0 100 100	21, 32, 45, 53	0
1	O	131/134 (97%)	-0.18	3 (2%) 61 60	19, 34, 49, 59	0
1	P	131/134 (97%)	-0.28	0 100 100	16, 31, 43, 51	0
All	All	2096/2144 (97%)	-0.28	7 (0%) 93 93	14, 31, 46, 60	0

The worst 5 of 7 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	O	53	PHE	2.8
1	O	114	HIS	2.7
1	A	53	PHE	2.7
1	F	114	HIS	2.5
1	K	5	THR	2.4

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

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

## 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
2	GAL	H	201	11/12	0.81	0.24	8.26	38,42,46,50	0
2	GAL	N	201	11/12	0.89	0.21	8.02	33,40,47,48	0
3	BGC	M	202	12/12	0.82	0.19	7.61	37,44,51,54	0
2	GAL	J	201	11/12	0.87	0.18	6.08	33,37,44,47	0
3	BGC	J	202	12/12	0.81	0.21	3.87	40,47,49,52	0
2	GAL	D	201	11/12	0.86	0.16	3.74	29,36,43,45	0
3	BGC	N	202	12/12	0.81	0.18	3.23	42,50,55,59	0
3	BGC	D	202	12/12	0.86	0.16	2.66	34,40,47,49	0
2	GAL	M	201	11/12	0.85	0.14	2.15	21,35,40,40	0
3	BGC	H	202	12/12	0.79	0.14	1.60	38,43,46,47	0
2	GAL	I	201	11/12	0.93	0.11	1.37	25,28,31,32	0
3	BGC	I	202	12/12	0.93	0.12	0.17	29,34,37,38	0
2	GAL	C	201	11/12	0.93	0.09	-0.19	23,27,32,33	0
3	BGC	G	202	12/12	0.95	0.11	-0.30	19,29,34,41	0
3	BGC	C	202	12/12	0.96	0.09	-0.88	24,32,35,36	0
2	GAL	G	201	11/12	0.93	0.10	-0.97	14,17,21,25	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.