



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 26, 2024 – 11:40 PM EDT

PDB ID : 4LKD  
Title : Crystal Structure of Pseudomonas aeruginosa Lectin LecA Complexed with GalA-QRS at 2.31 Å Resolution  
Authors : Kadam, R.U.; Stocker, A.; Reymond, J.-L.  
Deposited on : 2013-07-07  
Resolution : 2.31 Å (reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

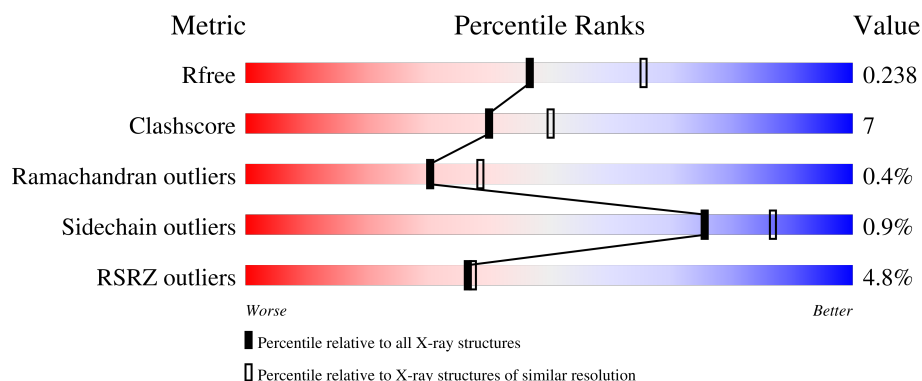
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.31 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))
$R_{free}$	164625	5963 (2.30-2.30)
Clashscore	180529	6698 (2.30-2.30)
Ramachandran outliers	177936	6640 (2.30-2.30)
Sidechain outliers	177891	6640 (2.30-2.30)
RSRZ outliers	164620	5963 (2.30-2.30)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . 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	121	<div> <div>0%</div> <div>88%</div> <div>12%</div> <div>.</div> </div>
1	B	121	<div> <div>0%</div> <div>87%</div> <div>12%</div> <div>.</div> </div>
1	C	121	<div> <div>2%</div> <div>86%</div> <div>14%</div> <div>.</div> </div>
1	D	121	<div> <div>3%</div> <div>87%</div> <div>12%</div> <div>.</div> </div>
1	E	121	<div> <div>6%</div> <div>85%</div> <div>14%</div> <div>.</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	121	<div> <div>2%</div> <div>90%</div> <div>10%</div> </div>
1	G	121	<div> <div>3%</div> <div>87%</div> <div>13%</div> </div>
1	H	121	<div> <div>3%</div> <div>89%</div> <div>11%</div> </div>
2	I	4	<div> <div>100%</div> <div>50%</div> <div>50%</div> </div>
2	J	4	<div> <div>25%</div> <div>50%</div> <div>50%</div> </div>
2	K	4	<div> <div>50%</div> <div>25%</div> <div>25%</div> <div>50%</div> </div>
2	L	4	<div> <div>100%</div> <div>25%</div> <div>75%</div> </div>
2	M	4	<div> <div>75%</div> <div>25%</div> <div>25%</div> <div>25%</div> <div>25%</div> </div>
2	N	4	<div> <div>50%</div> <div>25%</div> <div>25%</div> <div>50%</div> </div>
2	O	4	<div> <div>50%</div> <div>50%</div> <div>50%</div> </div>
2	P	4	<div> <div>75%</div> <div>50%</div> <div>25%</div> <div>25%</div> </div>

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 8681 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 PA-I galactophilic lectin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	121	Total	C	N	O	S	0	0	0
			901	567	156	175	3			
1	B	121	Total	C	N	O	S	0	0	0
			901	567	156	175	3			
1	C	121	Total	C	N	O	S	0	0	0
			901	567	156	175	3			
1	D	121	Total	C	N	O	S	0	0	0
			901	567	156	175	3			
1	E	121	Total	C	N	O	S	0	0	0
			901	567	156	175	3			
1	F	121	Total	C	N	O	S	0	0	0
			901	567	156	175	3			
1	G	121	Total	C	N	O	S	0	0	0
			901	567	156	175	3			
1	H	121	Total	C	N	O	S	0	0	0
			901	567	156	175	3			

- Molecule 2 is a protein called peptide QRSA.

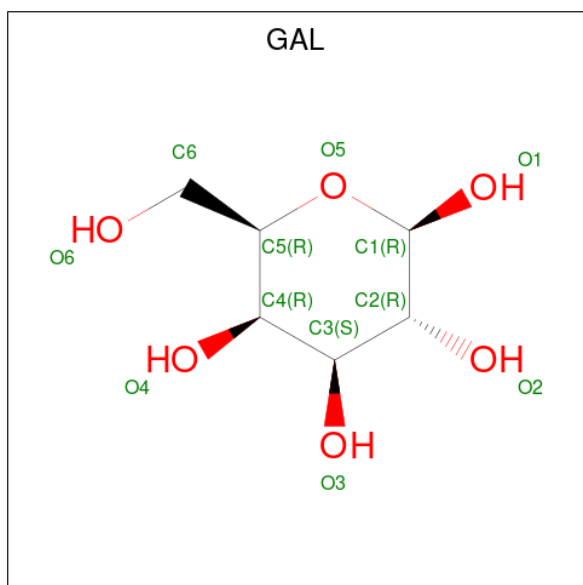
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	I	4	Total	C	N	O	0	0	1
			27	14	8	5			
2	J	2	Total	C	N	O	0	0	0
			20	11	6	3			
2	K	2	Total	C	N	O	0	0	0
			13	7	3	3			
2	L	4	Total	C	N	O	0	0	1
			27	14	8	5			
2	M	3	Total	C	N	O	0	0	0
			25	14	7	4			
2	N	2	Total	C	N	O	0	0	0
			13	8	3	2			

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	O	2	Total	C	N	O	0	0	0
			17	10	4	3			
2	P	4	Total	C	N	O	0	0	1
			27	14	8	5			

- Molecule 3 is beta-D-galactopyranose (three-letter code: GAL) (formula: C<sub>6</sub>H<sub>12</sub>O<sub>6</sub>).

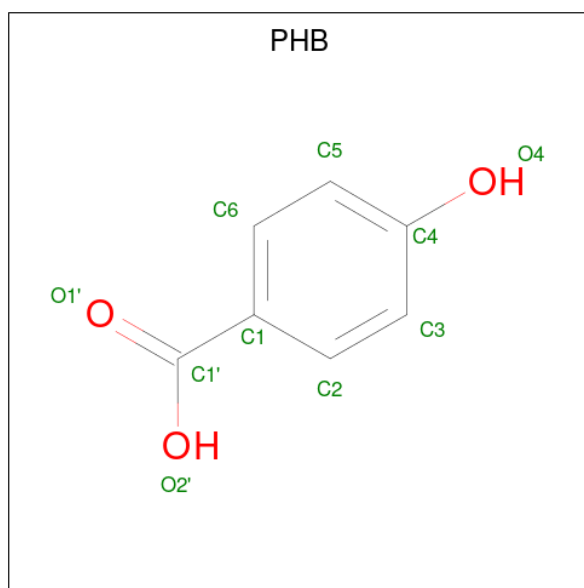


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			11	6	5		
3	B	1	Total	C	O	0	0
			11	6	5		
3	C	1	Total	C	O	0	0
			11	6	5		
3	D	1	Total	C	O	0	0
			11	6	5		
3	E	1	Total	C	O	0	0
			11	6	5		
3	F	1	Total	C	O	0	0
			11	6	5		
3	G	1	Total	C	O	0	0
			11	6	5		
3	H	1	Total	C	O	0	0
			11	6	5		

- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total Ca 1 1	0	0
4	B	1	Total Ca 1 1	0	0
4	C	1	Total Ca 1 1	0	0
4	D	1	Total Ca 1 1	0	0
4	E	1	Total Ca 1 1	0	0
4	F	1	Total Ca 1 1	0	0
4	G	1	Total Ca 1 1	0	0
4	H	1	Total Ca 1 1	0	0

- Molecule 5 is P-HYDROXYBENZOIC ACID (three-letter code: PHB) (formula:  $C_7H_6O_3$ ).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	I	1	Total C O 9 7 2	0	0
5	J	1	Total C O 9 7 2	0	0
5	K	1	Total C O 9 7 2	0	0
5	L	1	Total C O 9 7 2	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	M	1	Total	C	O	0	0
			9	7	2		
5	N	1	Total	C	O	0	0
			9	7	2		
5	O	1	Total	C	O	0	0
			9	7	2		
5	P	1	Total	C	O	0	0
			9	7	2		

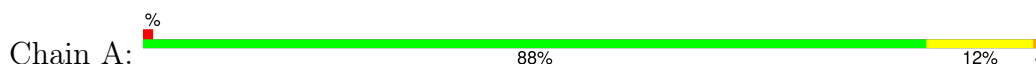
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	185	Total	O	0	0
			185	185		
6	B	162	Total	O	0	0
			162	162		
6	C	129	Total	O	0	0
			129	129		
6	D	136	Total	O	0	0
			136	136		
6	E	112	Total	O	0	0
			112	112		
6	F	156	Total	O	0	0
			156	156		
6	G	134	Total	O	0	0
			134	134		
6	H	96	Total	O	0	0
			96	96		
6	I	2	Total	O	0	0
			2	2		
6	J	5	Total	O	0	0
			5	5		
6	K	7	Total	O	0	0
			7	7		
6	L	5	Total	O	0	0
			5	5		
6	M	1	Total	O	0	0
			1	1		
6	N	3	Total	O	0	0
			3	3		
6	O	3	Total	O	0	0
			3	3		

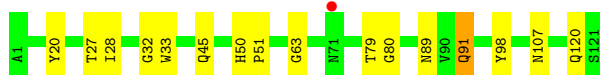
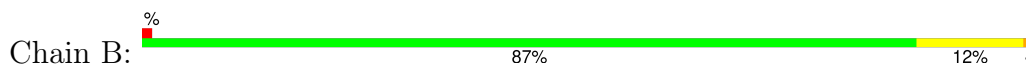
### 3 Residue-property plots [i](#)

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.

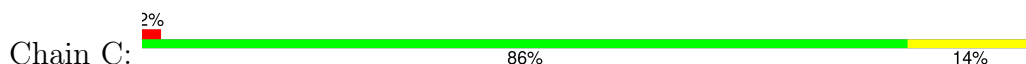
- Molecule 1: PA-I galactophilic lectin



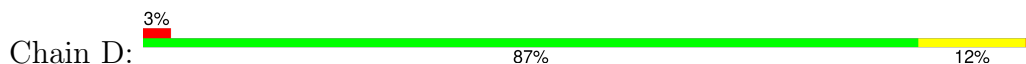
- Molecule 1: PA-I galactophilic lectin



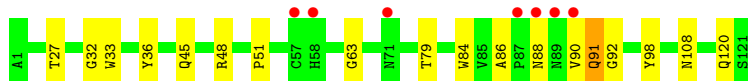
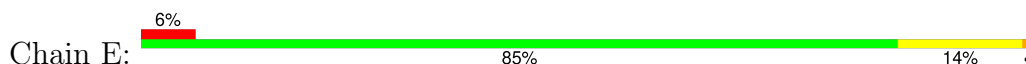
- Molecule 1: PA-I galactophilic lectin



- Molecule 1: PA-I galactophilic lectin

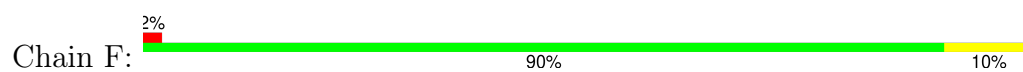


- Molecule 1: PA-I galactophilic lectin

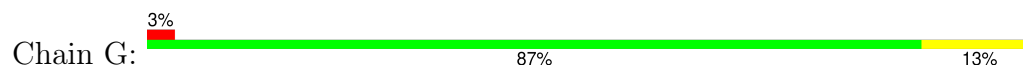


- Molecule 1: PA-I galactophilic lectin

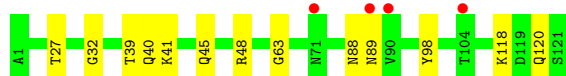
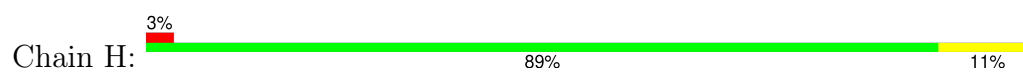




- Molecule 1: PA-I galactophilic lectin



- Molecule 1: PA-I galactophilic lectin



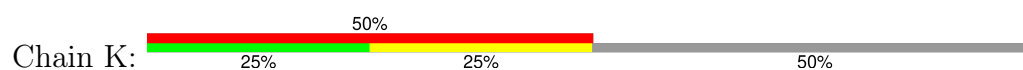
- Molecule 2: peptide QRSA



- Molecule 2: peptide QRSA



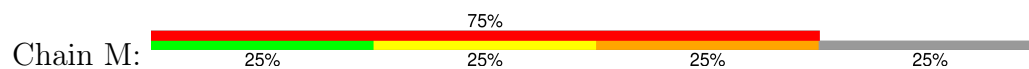
- Molecule 2: peptide QRSA



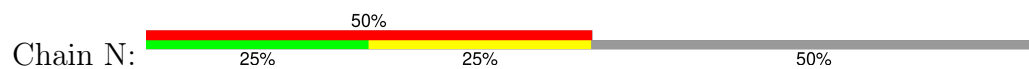
- Molecule 2: peptide QRSA



## ● Molecule 2: peptide QRSA



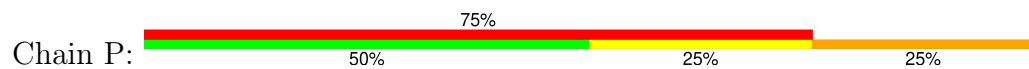
## ● Molecule 2: peptide QRSA



## ● Molecule 2: peptide QRSA



## ● Molecule 2: peptide QRSA



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	159.22Å 149.24Å 87.02Å 90.00° 110.96° 90.00°	Depositor
Resolution (Å)	48.46 – 2.31 48.46 – 2.31	Depositor EDS
% Data completeness (in resolution range)	92.9 (48.46-2.31) 98.2 (48.46-2.31)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.30 (at 2.32Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.6.4_486)	Depositor
R, $R_{free}$	0.215 , 0.247 0.205 , 0.238	Depositor DCC
$R_{free}$ test set	4088 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	25.6	Xtriage
Anisotropy	0.545	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 45.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	8681	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 16.56% 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: CA, GAL, PHB

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	0.21	0/924	0.39	0/1262
1	B	0.21	0/924	0.39	0/1262
1	C	0.20	0/924	0.39	0/1262
1	D	0.21	0/924	0.38	0/1262
1	E	0.21	0/924	0.40	0/1262
1	F	0.21	0/924	0.39	0/1262
1	G	0.21	0/924	0.39	0/1262
1	H	0.20	0/924	0.38	0/1262
2	I	0.26	0/26	0.44	0/33
2	J	0.20	0/19	0.33	0/23
2	K	0.17	0/12	0.42	0/14
2	L	0.21	0/26	0.36	0/33
2	M	0.18	0/24	0.42	0/30
2	N	0.15	0/12	0.25	0/15
2	O	0.19	0/16	0.31	0/19
2	P	0.20	0/26	0.32	0/33
All	All	0.21	0/7553	0.39	0/10296

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts

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	901	0	861	13	0
1	B	901	0	861	15	0
1	C	901	0	861	15	0
1	D	901	0	861	15	0
1	E	901	0	861	13	0
1	F	901	0	861	11	0
1	G	901	0	861	15	0
1	H	901	0	861	11	0
2	I	27	0	25	3	0
2	J	20	0	20	0	0
2	K	13	0	8	4	0
2	L	27	0	25	6	0
2	M	25	0	25	4	0
2	N	13	0	9	1	0
2	O	17	0	15	0	0
2	P	27	0	25	1	0
3	A	11	0	10	1	0
3	B	11	0	10	0	0
3	C	11	0	10	0	0
3	D	11	0	10	0	0
3	E	11	0	10	0	0
3	F	11	0	10	0	0
3	G	11	0	10	0	0
3	H	11	0	10	0	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
4	C	1	0	0	0	0
4	D	1	0	0	0	0
4	E	1	0	0	0	0
4	F	1	0	0	0	0
4	G	1	0	0	0	0
4	H	1	0	0	0	0
5	I	9	0	4	1	0
5	J	9	0	4	0	0
5	K	9	0	4	2	0
5	L	9	0	4	3	0
5	M	9	0	4	0	0
5	N	9	0	4	0	0
5	O	9	0	4	0	0
5	P	9	0	4	0	0
6	A	185	0	0	0	0
6	B	162	0	0	3	0
6	C	129	0	0	1	0

Continued on next page...

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	D	136	0	0	4	0
6	E	112	0	0	1	0
6	F	156	0	0	1	0
6	G	134	0	0	4	0
6	H	96	0	0	1	0
6	I	2	0	0	1	0
6	J	5	0	0	0	0
6	K	7	0	0	0	0
6	L	5	0	0	0	0
6	M	1	0	0	0	0
6	N	3	0	0	0	0
6	O	3	0	0	0	0
All	All	8681	0	7152	96	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 (96) 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:K:202:GLN:HG2	5:K:301:PHB:H2	1.36	1.04
1:F:32:GLY:H	1:G:45:GLN:HE22	1.18	0.91
1:B:45:GLN:HE22	1:H:32:GLY:H	1.20	0.89
1:D:32:GLY:H	1:E:45:GLN:HE22	1.15	0.89
1:A:45:GLN:HE22	1:C:32:GLY:H	1.21	0.88
1:A:32:GLY:H	1:C:45:GLN:HE22	1.20	0.86
1:F:45:GLN:HE22	1:G:32:GLY:H	1.23	0.86
1:G:59:ASP:HB2	6:G:415:HOH:O	1.76	0.85
1:D:45:GLN:HE22	1:E:32:GLY:H	1.23	0.83
1:B:32:GLY:H	1:H:45:GLN:HE22	1.29	0.81
1:F:51:PRO:HG3	2:N:203:ARG:HA	1.67	0.75
2:I:204:SER:H	5:I:301:PHB:H2	1.52	0.73
1:G:57:CYS:SG	6:G:415:HOH:O	2.50	0.70
6:D:368:HOH:O	5:L:301:PHB:H3	1.93	0.69
1:B:91:GLN:HE21	1:B:91:GLN:N	1.91	0.69
1:G:76:PRO:HG3	6:G:415:HOH:O	1.97	0.65
1:C:53:GLN:HE22	2:K:202:GLN:CG	2.09	0.65
1:B:89:ASN:HA	6:B:445:HOH:O	1.97	0.64
1:C:53:GLN:HE22	2:K:202:GLN:HG3	1.61	0.64
2:L:203:ARG:NH2	5:L:301:PHB:H2	2.12	0.64
2:K:202:GLN:HG2	5:K:301:PHB:C2	2.20	0.62

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:53:GLN:HE22	2:L:203:ARG:HH22	1.46	0.62
2:I:203:ARG:HH11	2:I:203:ARG:H	1.49	0.60
1:C:5:GLU:OE1	6:C:360:HOH:O	2.17	0.59
1:B:120:GLN:HG2	1:G:119:ASP:HA	1.84	0.58
1:A:120:GLN:HG2	1:D:119:ASP:HA	1.86	0.57
1:D:27:THR:HG21	1:E:27:THR:HG21	1.86	0.57
1:F:63:GLY:HA2	1:F:98:TYR:CZ	2.40	0.57
1:B:107:ASN:ND2	6:B:444:HOH:O	2.35	0.56
1:E:51:PRO:HG3	2:M:202:GLN:HB3	1.88	0.55
2:M:202:GLN:O	2:M:203:ARG:HG2	2.07	0.55
1:E:86:ALA:HB1	1:E:90:VAL:HG22	1.89	0.54
1:A:63:GLY:HA2	1:A:98:TYR:CZ	2.43	0.53
1:A:7:LEU:HD21	6:B:387:HOH:O	2.09	0.53
1:D:63:GLY:HA2	1:D:98:TYR:CZ	2.43	0.53
1:H:40:GLN:HG2	1:H:41:LYS:N	2.24	0.53
1:G:78:ASN:ND2	6:G:307:HOH:O	2.37	0.53
1:C:119:ASP:HA	1:E:120:GLN:HG2	1.90	0.53
1:D:39:THR:O	6:D:356:HOH:O	2.18	0.53
2:L:203:ARG:HH21	5:L:301:PHB:H2	1.73	0.53
1:F:48:ARG:HD2	1:G:33:TRP:CE3	2.44	0.52
1:G:120:GLN:HG2	1:H:120:GLN:CD	2.29	0.52
1:D:45:GLN:HG3	1:D:79:THR:HG23	1.92	0.51
1:C:120:GLN:HG2	1:D:120:GLN:CD	2.31	0.51
1:D:48:ARG:HD2	1:E:33:TRP:CE3	2.46	0.51
2:I:203:ARG:NE	6:I:401:HOH:O	2.43	0.51
1:H:88:ASN:O	1:H:89:ASN:HB2	2.10	0.50
1:C:36:TYR:HA	1:C:108:ASN:OD1	2.11	0.50
1:B:63:GLY:HA2	1:B:98:TYR:CZ	2.47	0.50
1:E:63:GLY:HA2	1:E:98:TYR:CZ	2.47	0.49
1:A:36:TYR:HA	1:A:108:ASN:OD1	2.12	0.49
1:E:91:GLN:HG3	1:E:92:GLY:N	2.27	0.49
1:F:120:GLN:HB2	1:H:118:LYS:O	2.13	0.49
1:H:63:GLY:HA2	1:H:98:TYR:CZ	2.48	0.49
1:C:63:GLY:HA2	1:C:98:TYR:CZ	2.47	0.48
1:A:50:HIS:ND1	1:A:51:PRO:HD2	2.30	0.47
1:C:120:GLN:CD	1:D:120:GLN:HG2	2.34	0.47
1:F:50:HIS:ND1	1:F:51:PRO:HD2	2.29	0.47
1:G:63:GLY:HA2	1:G:98:TYR:CZ	2.50	0.47
1:B:27:THR:HG21	1:H:27:THR:HG21	1.96	0.47
1:E:45:GLN:HG3	1:E:79:THR:HG23	1.96	0.47
1:F:33:TRP:CE3	1:G:48:ARG:HD2	2.49	0.47

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:84:TRP:CH2	1:E:86:ALA:HA	2.51	0.46
1:A:71:ASN:H	1:A:71:ASN:HD22	1.63	0.46
6:H:396:HOH:O	2:P:202:GLN:NE2	2.48	0.46
6:F:327:HOH:O	1:G:40:GLN:HG3	2.16	0.46
2:L:203:ARG:HG2	2:L:203:ARG:NH1	2.31	0.46
1:B:120:GLN:OE1	1:F:120:GLN:HG2	2.17	0.45
1:A:71:ASN:H	1:A:71:ASN:ND2	2.13	0.45
1:A:33:TRP:CE3	1:C:48:ARG:HD2	2.52	0.45
1:A:27:THR:HG21	1:C:27:THR:HG21	1.99	0.44
1:G:50:HIS:ND1	1:G:51:PRO:HD2	2.33	0.44
1:H:88:ASN:HD22	1:H:89:ASN:HD22	1.66	0.44
1:B:33:TRP:CE3	1:H:48:ARG:HD2	2.53	0.44
1:D:88:ASN:HA	6:D:423:HOH:O	2.16	0.44
1:C:45:GLN:HG3	1:C:79:THR:HG23	2.00	0.43
1:C:50:HIS:ND1	1:C:51:PRO:HD2	2.34	0.43
1:F:36:TYR:HA	1:F:108:ASN:OD1	2.18	0.43
2:M:202:GLN:CG	2:M:203:ARG:H	2.32	0.43
2:L:202:GLN:HG2	2:L:202:GLN:O	2.19	0.43
1:D:50:HIS:CG	1:D:51:PRO:HD2	2.54	0.43
1:G:120:GLN:CD	1:H:120:GLN:HG2	2.39	0.43
1:B:45:GLN:HG3	1:B:79:THR:HG23	2.01	0.42
1:E:48:ARG:HD3	6:E:303:HOH:O	2.18	0.42
1:B:50:HIS:ND1	1:B:51:PRO:HD2	2.34	0.42
1:F:27:THR:HG21	1:G:27:THR:HG21	2.01	0.42
1:E:36:TYR:HA	1:E:108:ASN:OD1	2.19	0.41
2:L:203:ARG:N	2:L:203:ARG:NE	2.68	0.41
1:B:28:ILE:O	1:B:80:GLY:HA3	2.20	0.41
1:A:104:THR:HB	3:A:201:GAL:O3	2.20	0.41
1:A:105:TYR:HA	1:A:108:ASN:ND2	2.36	0.41
2:M:202:GLN:HG2	2:M:203:ARG:N	2.36	0.41
1:B:20:TYR:O	1:B:91:GLN:HB2	2.21	0.40
1:D:48:ARG:HG2	6:D:323:HOH:O	2.21	0.40
1:B:89:ASN:HD22	1:D:22:PRO:HB3	1.85	0.40
1:C:100:ASP:HB3	1:C:105:TYR:CZ	2.57	0.40

There are no symmetry-related clashes.



## 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	119/121 (98%)	116 (98%)	3 (2%)	0	100	100
1	B	119/121 (98%)	114 (96%)	5 (4%)	0	100	100
1	C	119/121 (98%)	117 (98%)	2 (2%)	0	100	100
1	D	119/121 (98%)	115 (97%)	4 (3%)	0	100	100
1	E	119/121 (98%)	112 (94%)	6 (5%)	1 (1%)	16	20
1	F	119/121 (98%)	116 (98%)	3 (2%)	0	100	100
1	G	119/121 (98%)	116 (98%)	3 (2%)	0	100	100
1	H	119/121 (98%)	114 (96%)	5 (4%)	0	100	100
2	I	2/4 (50%)	1 (50%)	1 (50%)	0	100	100
2	L	2/4 (50%)	1 (50%)	0	1 (50%)	0	0
2	M	1/4 (25%)	0	0	1 (100%)	0	0
2	P	2/4 (50%)	1 (50%)	0	1 (50%)	0	0
All	All	959/984 (98%)	923 (96%)	32 (3%)	4 (0%)	30	39

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	L	204	SER
2	P	203	ARG
2	M	203	ARG
1	E	88	ASN

### 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	94/94 (100%)	93 (99%)	1 (1%)	70	83
1	B	94/94 (100%)	93 (99%)	1 (1%)	70	83
1	C	94/94 (100%)	94 (100%)	0	100	100
1	D	94/94 (100%)	92 (98%)	2 (2%)	48	66
1	E	94/94 (100%)	93 (99%)	1 (1%)	70	83
1	F	94/94 (100%)	94 (100%)	0	100	100
1	G	94/94 (100%)	94 (100%)	0	100	100
1	H	94/94 (100%)	93 (99%)	1 (1%)	70	83
2	I	3/3 (100%)	3 (100%)	0	100	100
2	J	2/3 (67%)	2 (100%)	0	100	100
2	K	1/3 (33%)	1 (100%)	0	100	100
2	L	3/3 (100%)	3 (100%)	0	100	100
2	M	3/3 (100%)	3 (100%)	0	100	100
2	N	1/3 (33%)	1 (100%)	0	100	100
2	O	1/3 (33%)	1 (100%)	0	100	100
2	P	3/3 (100%)	2 (67%)	1 (33%)	0	0
All	All	769/776 (99%)	762 (99%)	7 (1%)	75	87

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	71	ASN
1	B	91	GLN
1	D	39	THR
1	D	40	GLN
1	E	91	GLN
1	H	39	THR
2	P	203	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (26) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	45	GLN
1	A	71	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	78	ASN
1	B	45	GLN
1	B	91	GLN
1	C	45	GLN
1	C	53	GLN
1	C	89	ASN
1	C	120	GLN
1	D	45	GLN
1	D	120	GLN
1	E	45	GLN
1	E	91	GLN
1	F	45	GLN
1	F	78	ASN
1	F	91	GLN
1	G	45	GLN
1	G	120	GLN
1	H	40	GLN
1	H	45	GLN
1	H	78	ASN
1	H	88	ASN
1	H	89	ASN
1	H	120	GLN
2	O	202	GLN
2	P	202	GLN

### 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 oligosaccharides in this entry.

### 5.6 Ligand geometry ⓘ

Of 24 ligands modelled in this entry, 8 are monoatomic - leaving 16 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
5	PHB	L	301	3,2	9,9,10	1.25	0	11,11,13	1.09	1 (9%)
5	PHB	I	301	3,2	9,9,10	1.21	0	11,11,13	0.96	0
3	GAL	H	201	4,5	11,11,12	0.90	0	15,15,17	1.15	1 (6%)
5	PHB	J	301	3,2	9,9,10	1.21	0	11,11,13	1.11	1 (9%)
5	PHB	O	301	3,2	9,9,10	1.21	0	11,11,13	0.95	0
5	PHB	P	301	3,2	9,9,10	1.22	0	11,11,13	1.05	1 (9%)
3	GAL	B	201	4,5	11,11,12	0.85	0	15,15,17	1.07	2 (13%)
3	GAL	D	201	4,5	11,11,12	0.95	0	15,15,17	1.14	2 (13%)
3	GAL	A	201	4,5	11,11,12	0.91	0	15,15,17	1.13	2 (13%)
3	GAL	E	201	4,5	11,11,12	0.90	0	15,15,17	1.24	2 (13%)
3	GAL	C	201	4,5	11,11,12	0.95	0	15,15,17	1.12	1 (6%)
5	PHB	M	301	3,2	9,9,10	1.25	0	11,11,13	0.97	0
5	PHB	K	301	3,2	9,9,10	1.20	0	11,11,13	1.10	1 (9%)
3	GAL	G	201	4,5	11,11,12	0.91	0	15,15,17	1.15	2 (13%)
5	PHB	N	301	3,2	9,9,10	1.31	0	11,11,13	1.02	1 (9%)
3	GAL	F	201	4,5	11,11,12	0.90	0	15,15,17	1.20	2 (13%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	PHB	L	301	3,2	-	0/2/2/4	0/1/1/1
5	PHB	I	301	3,2	-	2/2/2/4	0/1/1/1
3	GAL	H	201	4,5	-	0/2/19/22	0/1/1/1
5	PHB	J	301	3,2	-	0/2/2/4	0/1/1/1
5	PHB	O	301	3,2	-	0/2/2/4	0/1/1/1
5	PHB	P	301	3,2	-	0/2/2/4	0/1/1/1
3	GAL	B	201	4,5	-	0/2/19/22	0/1/1/1
3	GAL	D	201	4,5	-	0/2/19/22	0/1/1/1
3	GAL	A	201	4,5	-	0/2/19/22	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	GAL	E	201	4,5	-	0/2/19/22	0/1/1/1
3	GAL	C	201	4,5	-	0/2/19/22	0/1/1/1
5	PHB	M	301	3,2	-	0/2/2/4	0/1/1/1
5	PHB	K	301	3,2	-	2/2/2/4	0/1/1/1
3	GAL	G	201	4,5	-	0/2/19/22	0/1/1/1
5	PHB	N	301	3,2	-	2/2/2/4	0/1/1/1
3	GAL	F	201	4,5	-	0/2/19/22	0/1/1/1

There are no bond length outliers.

All (19) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	E	201	GAL	C1-O5-C5	3.21	116.49	112.19
3	H	201	GAL	C1-O5-C5	2.88	116.04	112.19
3	F	201	GAL	C1-O5-C5	2.85	116.01	112.19
3	A	201	GAL	C1-O5-C5	2.64	115.73	112.19
3	D	201	GAL	C1-O5-C5	2.61	115.68	112.19
3	B	201	GAL	C1-O5-C5	2.55	115.60	112.19
3	G	201	GAL	C1-O5-C5	2.49	115.52	112.19
3	C	201	GAL	C1-O5-C5	2.46	115.48	112.19
3	F	201	GAL	C1-C2-C3	-2.40	106.16	109.64
3	A	201	GAL	C1-C2-C3	-2.33	106.25	109.64
5	K	301	PHB	O4-C4-C5	2.30	126.42	120.00
3	E	201	GAL	C1-C2-C3	-2.27	106.34	109.64
5	L	301	PHB	O4-C4-C5	2.24	126.25	120.00
5	J	301	PHB	O4-C4-C5	2.23	126.21	120.00
3	B	201	GAL	C1-C2-C3	-2.13	106.55	109.64
5	N	301	PHB	O4-C4-C5	2.11	125.88	120.00
3	G	201	GAL	C1-C2-C3	-2.09	106.60	109.64
5	P	301	PHB	O4-C4-C5	2.09	125.81	120.00
3	D	201	GAL	O2-C2-C3	-2.04	105.92	110.15

There are no chirality outliers.

All (6) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	I	301	PHB	C6-C1-C1'-O1'
5	I	301	PHB	C2-C1-C1'-O1'
5	N	301	PHB	C6-C1-C1'-O1'
5	N	301	PHB	C2-C1-C1'-O1'
5	K	301	PHB	C6-C1-C1'-O1'

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
5	K	301	PHB	C2-C1-C1'-O1'

There are no ring outliers.

4 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	L	301	PHB	3	0
5	I	301	PHB	1	0
3	A	201	GAL	1	0
5	K	301	PHB	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	121/121 (100%)	-0.25	1 (0%) 82 83	16, 22, 32, 53	0
1	B	121/121 (100%)	-0.12	1 (0%) 82 83	16, 24, 36, 49	0
1	C	121/121 (100%)	0.04	3 (2%) 58 59	18, 27, 37, 54	0
1	D	121/121 (100%)	0.04	4 (3%) 49 51	18, 27, 37, 57	0
1	E	121/121 (100%)	0.11	7 (5%) 30 32	17, 26, 42, 64	0
1	F	121/121 (100%)	-0.03	3 (2%) 58 59	18, 26, 36, 53	0
1	G	121/121 (100%)	0.05	4 (3%) 49 51	19, 27, 37, 50	0
1	H	121/121 (100%)	0.21	4 (3%) 49 51	19, 30, 45, 65	0
2	I	4/4 (100%)	2.84	4 (100%) 0 0	57, 58, 60, 60	0
2	J	2/4 (50%)	1.87	1 (50%) 0 0	40, 40, 40, 49	0
2	K	2/4 (50%)	2.92	2 (100%) 0 0	50, 50, 50, 57	0
2	L	4/4 (100%)	3.40	4 (100%) 0 0	42, 45, 47, 47	0
2	M	3/4 (75%)	4.03	3 (100%) 0 0	51, 51, 57, 65	0
2	N	2/4 (50%)	5.51	2 (100%) 0 0	48, 48, 48, 51	0
2	O	2/4 (50%)	2.62	2 (100%) 0 0	56, 56, 56, 61	0
2	P	4/4 (100%)	3.01	3 (75%) 0 0	49, 58, 63, 64	0
All	All	991/1000 (99%)	0.08	48 (4%) 36 37	16, 26, 45, 65	0

All (48) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	M	204	SER	6.5
2	N	203	ARG	5.9
2	N	202	GLN	5.2
2	I	204	SER	4.7
1	E	88	ASN	4.5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	E	90	VAL	4.1
2	L	203	ARG	3.7
2	L	205	ALA	3.7
2	P	203	ARG	3.7
2	P	204	SER	3.6
1	E	87	PRO	3.6
2	K	203	ARG	3.5
1	C	71	ASN	3.3
2	L	204	SER	3.2
2	O	203	ARG	3.1
2	P	205	ALA	3.1
1	D	89	ASN	3.0
1	H	71	ASN	3.0
2	L	202	GLN	3.0
2	M	203	ARG	2.8
1	F	89	ASN	2.8
2	M	202	GLN	2.7
1	A	88	ASN	2.7
1	C	88	ASN	2.6
1	B	71	ASN	2.5
1	D	71	ASN	2.5
1	E	89	ASN	2.5
2	J	203	ARG	2.5
2	K	202	GLN	2.4
1	G	71	ASN	2.4
1	G	88	ASN	2.3
1	F	71	ASN	2.3
2	I	205	ALA	2.3
2	I	202	GLN	2.2
1	D	39	THR	2.2
1	G	89	ASN	2.2
2	I	203	ARG	2.1
1	E	57	CYS	2.1
2	O	202	GLN	2.1
1	C	59	ASP	2.1
1	E	58	HIS	2.1
1	D	40	GLN	2.1
1	G	1	ALA	2.1
1	H	90	VAL	2.1
1	E	71	ASN	2.1
1	F	88	ASN	2.1
1	H	104	THR	2.0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	H	89	ASN	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(Å <sup>2</sup> )	Q<0.9
3	GAL	B	201	11/12	0.87	0.13	14,23,27,28	0
3	GAL	C	201	11/12	0.88	0.12	26,30,34,35	0
3	GAL	G	201	11/12	0.88	0.13	27,31,34,36	0
5	PHB	I	301	9/10	0.89	0.15	23,30,46,51	0
3	GAL	H	201	11/12	0.90	0.12	23,28,33,33	0
5	PHB	N	301	9/10	0.90	0.14	25,33,45,48	0
5	PHB	O	301	9/10	0.90	0.14	34,37,52,59	0
5	PHB	L	301	9/10	0.91	0.14	29,31,35,39	0
4	CA	H	202	1/1	0.92	0.15	33,33,33,33	0
5	PHB	M	301	9/10	0.92	0.11	22,26,36,38	0
3	GAL	D	201	11/12	0.93	0.10	22,26,38,40	0
3	GAL	E	201	11/12	0.93	0.10	19,23,28,30	0
4	CA	D	202	1/1	0.93	0.12	32,32,32,32	0
5	PHB	P	301	9/10	0.93	0.11	26,31,36,37	0
3	GAL	A	201	11/12	0.94	0.09	20,25,27,28	0
5	PHB	J	301	9/10	0.94	0.09	19,22,34,36	0
5	PHB	K	301	9/10	0.94	0.10	30,32,45,49	0
3	GAL	F	201	11/12	0.94	0.07	22,27,29,30	0
4	CA	F	202	1/1	0.96	0.05	26,26,26,26	0
4	CA	G	202	1/1	0.97	0.09	31,31,31,31	0
4	CA	E	202	1/1	0.98	0.08	26,26,26,26	0
4	CA	B	202	1/1	0.98	0.09	25,25,25,25	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	CA	C	202	1/1	0.99	0.10	31,31,31,31	0
4	CA	A	202	1/1	0.99	0.03	22,22,22,22	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.