



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 05:37 AM GMT

PDB ID : 2RG9  
Title : Crystal structure of viscum album mistletoe lectin I in native state at 1.95 Å resolution, comparison of structure active site conformation in ricin and in viscumin  
Authors : Karpechenko, N.U.; Timofeev, V.I.; Gabdoulkhakov, A.G.; Mikhailov, A.M.  
Deposited on : 2007-10-03  
Resolution : 1.95 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 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
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) : trunk26865

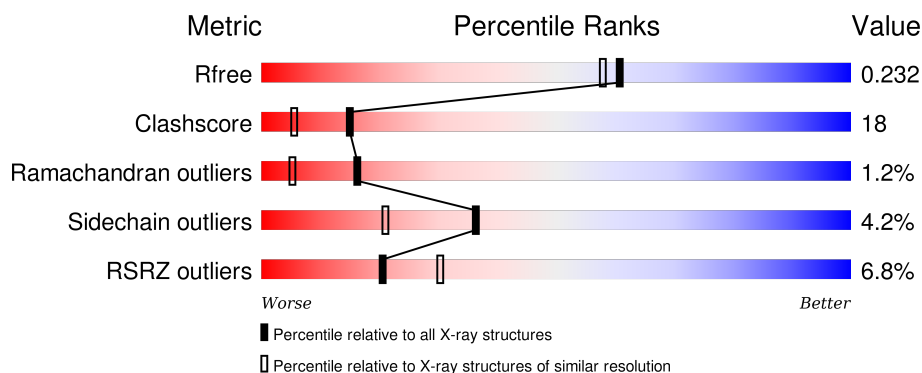
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.95 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	1833 (1.96-1.96)
Clashscore	102246	1953 (1.96-1.96)
Ramachandran outliers	100387	1936 (1.96-1.96)
Sidechain outliers	100360	1936 (1.96-1.96)
RSRZ outliers	91569	1835 (1.96-1.96)

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	249	<div> <div>8%</div> <div>74%</div> <div>21%</div> <div>• •</div> </div>
2	B	263	<div> <div>6%</div> <div>82%</div> <div>13%</div> <div>5%</div> </div>

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	NAG	A	5003	-	-	-	X
4	SO4	A	4001	-	-	X	-
5	AZI	B	4005	-	-	-	X
7	GOL	A	4010	-	-	-	X
7	GOL	A	4020	-	-	X	-
7	GOL	B	4013	-	-	-	X
7	GOL	B	4017	-	-	X	-
7	GOL	B	4018	-	-	-	X

## 2 Entry composition [i](#)

There are 8 unique types of molecules in this entry. The entry contains 4396 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 Beta-galactoside-specific lectin 1 chain A isoform 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	249	Total	C	N	O	S	0	0	0
			1921	1218	328	371	4			

There are 17 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	5	SER	ARG	CONFLICT	UNP P81446
A	8	THR	VAL	CONFLICT	UNP P81446
A	9	VAL	THR	CONFLICT	UNP P81446
A	10	GLN	HIS	CONFLICT	UNP P81446
A	19	SER	ARG	CONFLICT	UNP P81446
A	27	PHE	TYR	CONFLICT	UNP P81446
A	36	ASN	GLU	CONFLICT	UNP P81446
A	49	GLU	ASP	CONFLICT	UNP P81446
A	51	SER	GLN	CONFLICT	UNP P81446
A	84	GLN	ASP	CONFLICT	UNP P81446
A	90	LYS	ARG	CONFLICT	UNP P81446
A	99	GLN	HIS	CONFLICT	UNP P81446
A	100	ASP	LEU	CONFLICT	UNP P81446
A	208	HIS	GLN	CONFLICT	UNP P81446
A	222	LEU	ILE	CONFLICT	UNP P81446
A	227	VAL	PHE	CONFLICT	UNP P81446
A	233	ILE	VAL	CONFLICT	UNP P81446

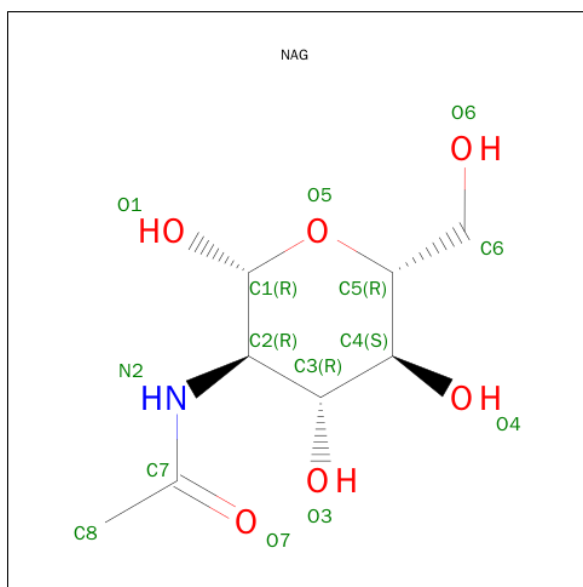
- Molecule 2 is a protein called Beta-galactoside-specific lectin 1 chain B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	263	Total	C	N	O	S	0	0	0
			1996	1238	353	393	12			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	11	ILE	THR	CONFLICT	UNP P81446
B	21	THR	CYS	CONFLICT	UNP P81446
B	30	GLN	HIS	CONFLICT	UNP P81446
B	54	LYS	ARG	CONFLICT	UNP P81446
B	90	ILE	LEU	CONFLICT	UNP P81446
B	92	GLN	GLU	CONFLICT	UNP P81446
B	166	THR	VAL	CONFLICT	UNP P81446
B	168	GLY	SER	CONFLICT	UNP P81446
B	170	GLU	GLN	CONFLICT	UNP P81446
B	189	SER	ASP	CONFLICT	UNP P81446
B	194	ASN	CYS	CONFLICT	UNP P81446
B	254	ASN	LYS	CONFLICT	UNP P81446

- Molecule 3 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



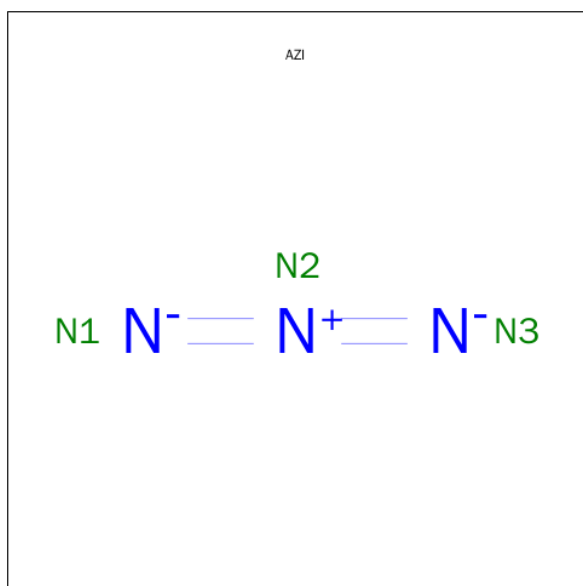
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	B	1	Total	C	N	O	0	0
			14	8	1	5		
3	B	1	Total	C	N	O	0	0
			14	8	1	5		
3	B	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula:  $O_4S$ ).



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

- Molecule 5 is AZIDE ION (three-letter code: AZI) (formula:  $N_3$ ).

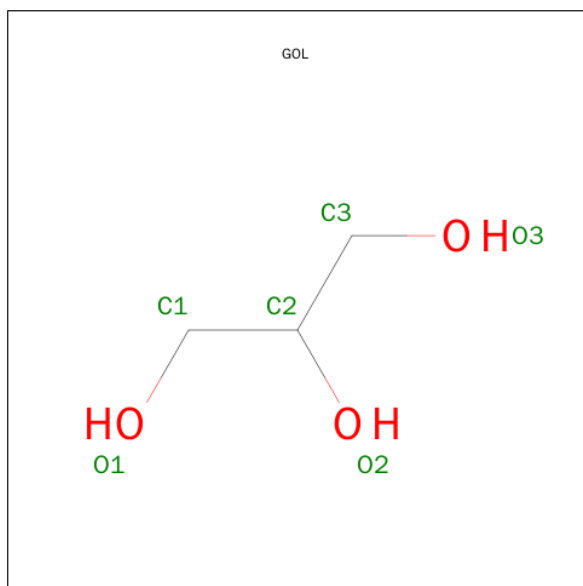


Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1	Total	N	0	0
			3	3		
5	B	1	Total	N	0	0
			3	3		
5	B	1	Total	N	0	0
			3	3		

- Molecule 6 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	2	Total Cl 2 2	0	0

- Molecule 7 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



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

*Continued on next page...*

*Continued from previous page...*

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

- Molecule 8 is water.

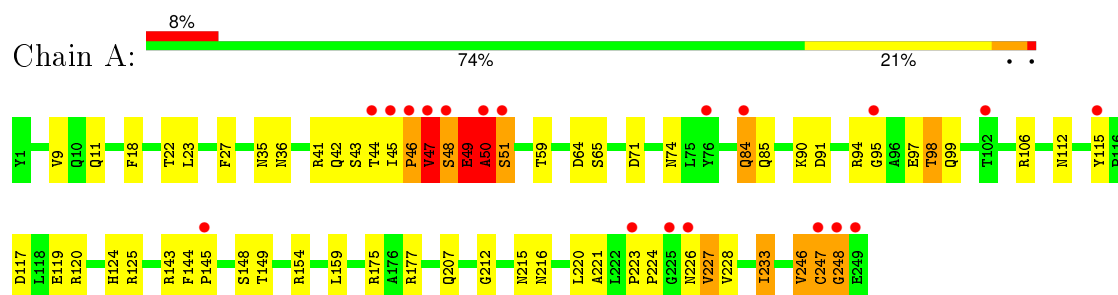
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	113	Total	O	0	0
			113	113		
8	B	204	Total	O	0	0
			204	204		



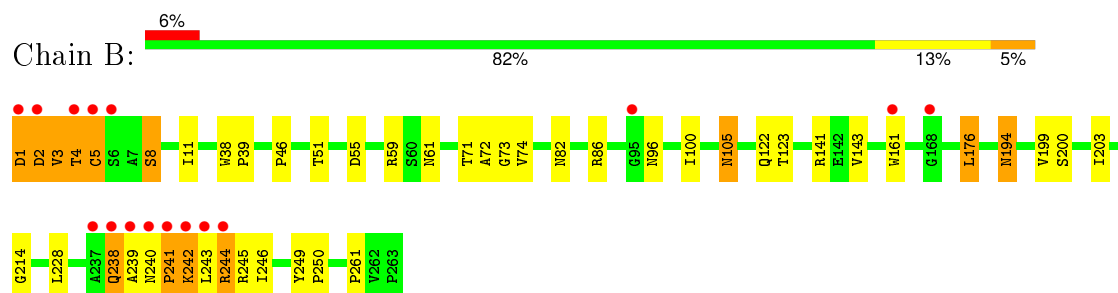
### 3 Residue-property plots [i](#)

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 ( $\text{RSRZ} > 2$ ). 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.

- Molecule 1: Beta-galactoside-specific lectin 1 chain A isoform 1



- Molecule 2: Beta-galactoside-specific lectin 1 chain B



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 65 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	106.95Å 106.95Å 310.95Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	8.00 – 1.95 8.00 – 1.95	Depositor EDS
% Data completeness (in resolution range)	(Not available) (8.00-1.95) 99.9 (8.00-1.95)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.86 (at 1.95Å)	Xtriage
Refinement program	PHENIX	Depositor
R, $R_{free}$	0.222 , 0.239 0.213 , 0.232	Depositor DCC
$R_{free}$ test set	3801 reflections (5.00%)	DCC
Wilson B-factor (Å <sup>2</sup> )	36.8	Xtriage
Anisotropy	0.055	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.44 , 70.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	0 of 76030 reflections	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	4396	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	46.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.54% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: AZI, SO4, CL, NAG, GOL

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	0.40	0/1960	0.51	0/2670
2	B	0.39	0/2035	0.55	0/2777
All	All	0.40	0/3995	0.53	0/5447

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	A	1	8
2	B	1	5
All	All	2	13

There are no bond length outliers.

There are no bond angle outliers.

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	49	GLU	CA
2	B	5	CYS	CA

All (13) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	246	VAL	Peptide
1	A	248	GLY	Peptide
1	A	46	PRO	Peptide
1	A	47	VAL	Peptide

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Group
1	A	48	SER	Peptide
1	A	49	GLU	Peptide
1	A	50	ALA	Peptide
1	A	51	SER	Peptide
2	B	1	ASP	Peptide
2	B	238	GLN	Peptide
2	B	242	LYS	Peptide
2	B	244	ARG	Peptide
2	B	4	THR	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	1921	0	1881	89	0
2	B	1996	0	1920	62	0
3	A	14	0	13	3	0
3	B	42	0	39	4	0
4	A	5	0	0	3	0
5	A	3	0	0	0	0
5	B	6	0	0	0	0
6	A	2	0	0	1	0
7	A	48	0	64	16	0
7	B	42	0	56	12	0
8	A	113	0	0	2	0
8	B	204	0	0	3	0
All	All	4396	0	3973	142	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 18.

All (142) 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:A:247:CYS:HB2	2:B:5:CYS:SG	1.55	1.45
1:A:49:GLU:HG2	1:A:50:ALA:CA	1.66	1.25

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:247:CYS:CB	2:B:5:CYS:SG	2.33	1.15
1:A:9:VAL:HG23	1:A:11:GLN:H	1.13	1.08
2:B:59:ARG:HH22	7:B:4017:GOL:H32	1.17	1.07
1:A:119:GLU:OE2	4:A:4001:SO4:O2	1.74	1.04
1:A:49:GLU:CG	1:A:50:ALA:CA	2.37	1.02
2:B:194:ASN:HD22	2:B:194:ASN:H	1.01	0.98
2:B:96:ASN:HD22	2:B:228:LEU:HD11	1.28	0.95
1:A:247:CYS:CB	2:B:5:CYS:HG	1.74	0.94
1:A:148:SER:HB2	7:A:4020:GOL:H11	1.50	0.93
1:A:48:SER:HB2	1:A:49:GLU:HA	1.49	0.93
1:A:27:PHE:HZ	1:A:51:SER:HG	1.15	0.91
1:A:247:CYS:HA	2:B:5:CYS:SG	2.13	0.89
2:B:240:ASN:OD1	2:B:241:PRO:HD2	1.73	0.88
1:A:97:GLU:HG3	7:A:4010:GOL:H12	1.56	0.86
1:A:49:GLU:HG2	1:A:50:ALA:N	1.79	0.85
1:A:247:CYS:CA	2:B:5:CYS:SG	2.64	0.84
2:B:59:ARG:HH22	7:B:4017:GOL:C3	1.89	0.84
2:B:61:ASN:HD21	3:B:5004:NAG:C1	1.91	0.84
2:B:194:ASN:N	2:B:194:ASN:HD22	1.77	0.82
1:A:47:VAL:HG23	1:A:47:VAL:O	1.80	0.81
1:A:9:VAL:HG23	1:A:11:GLN:N	1.95	0.80
2:B:96:ASN:HD22	2:B:228:LEU:CD1	1.96	0.79
1:A:47:VAL:CG2	1:A:47:VAL:O	2.30	0.79
2:B:59:ARG:NH2	7:B:4017:GOL:H32	1.96	0.78
1:A:95:GLY:H	7:A:4010:GOL:H2	1.50	0.76
2:B:96:ASN:ND2	2:B:228:LEU:HD11	2.00	0.75
2:B:240:ASN:OD1	2:B:241:PRO:CD	2.39	0.71
1:A:48:SER:HB2	1:A:49:GLU:CA	2.21	0.69
1:A:46:PRO:HA	1:A:47:VAL:HG12	1.73	0.69
2:B:141:ARG:HG3	7:B:4018:GOL:O1	1.94	0.68
1:A:46:PRO:CA	1:A:47:VAL:HG12	2.24	0.67
2:B:11:ILE:HG13	8:B:3199:HOH:O	1.94	0.67
2:B:194:ASN:ND2	2:B:194:ASN:H	1.83	0.67
1:A:246:VAL:HG12	1:A:246:VAL:O	1.94	0.67
1:A:246:VAL:CG1	1:A:246:VAL:O	2.43	0.67
1:A:65:SER:HB2	7:A:4009:GOL:H31	1.76	0.66
1:A:42:GLN:HG3	7:A:4008:GOL:H31	1.77	0.66
1:A:143:ARG:HH22	7:A:4019:GOL:H32	1.62	0.65
2:B:243:LEU:N	2:B:243:LEU:HD12	2.11	0.65
2:B:241:PRO:O	2:B:244:ARG:CB	2.45	0.64
2:B:238:GLN:HA	2:B:238:GLN:NE2	2.13	0.64

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:124:HIS:HD2	4:A:4001:SO4:O1	1.81	0.63
1:A:47:VAL:N	1:A:48:SER:O	2.29	0.62
2:B:2:ASP:O	2:B:4:THR:HG23	2.00	0.62
1:A:112:ASN:HD21	3:A:5003:NAG:C1	2.14	0.61
1:A:49:GLU:HG3	1:A:50:ALA:CA	2.29	0.60
1:A:247:CYS:HB2	2:B:5:CYS:HG	0.78	0.60
1:A:84:GLN:HG2	1:A:84:GLN:O	2.02	0.60
2:B:11:ILE:HG12	2:B:51:THR:OG1	2.01	0.60
2:B:243:LEU:HD12	2:B:243:LEU:H	1.67	0.60
1:A:47:VAL:N	1:A:48:SER:CA	2.65	0.59
1:A:154:ARG:HD3	8:A:3288:HOH:O	2.03	0.59
2:B:61:ASN:ND2	3:B:5004:NAG:C1	2.65	0.58
2:B:240:ASN:C	2:B:242:LYS:H	2.08	0.57
1:A:119:GLU:CD	4:A:4001:SO4:O2	2.42	0.57
1:A:212:GLY:HA2	1:A:233:ILE:HG23	1.86	0.56
1:A:18:PHE:O	1:A:22:THR:HG23	2.06	0.56
1:A:47:VAL:CG2	1:A:94:ARG:O	2.54	0.56
1:A:106:ARG:NH1	1:A:106:ARG:HB2	2.22	0.55
2:B:243:LEU:CD1	2:B:243:LEU:H	2.20	0.55
2:B:176:LEU:HD23	7:B:4018:GOL:H11	1.87	0.55
1:A:35:ASN:H	2:B:2:ASP:CB	2.20	0.55
2:B:71:THR:O	2:B:74:VAL:HG22	2.07	0.54
3:B:5002:NAG:O7	7:B:4013:GOL:H2	2.07	0.54
2:B:46:PRO:HD2	8:B:3034:HOH:O	2.06	0.54
1:A:47:VAL:N	1:A:48:SER:C	2.63	0.53
1:A:149:THR:HB	7:A:4020:GOL:H12	1.93	0.51
2:B:82:ASN:HD21	7:B:4017:GOL:H11	1.74	0.51
1:A:49:GLU:HG2	1:A:50:ALA:C	2.28	0.50
1:A:48:SER:CB	1:A:49:GLU:HA	2.23	0.49
1:A:148:SER:HB2	7:A:4020:GOL:C1	2.33	0.49
2:B:240:ASN:O	2:B:242:LYS:N	2.45	0.49
1:A:216:ASN:ND2	6:A:4021:CL:CL	2.82	0.49
7:B:4014:GOL:H32	8:B:3178:HOH:O	2.11	0.49
1:A:220:LEU:HB2	1:A:228:VAL:HG12	1.94	0.49
1:A:115:TYR:O	1:A:119:GLU:HG3	2.13	0.49
1:A:27:PHE:HZ	1:A:51:SER:OG	1.88	0.49
1:A:143:ARG:NH2	7:A:4019:GOL:H32	2.26	0.49
1:A:97:GLU:CG	7:A:4010:GOL:H12	2.37	0.48
2:B:240:ASN:C	2:B:242:LYS:N	2.66	0.48
1:A:84:GLN:NE2	1:A:85:GLN:HE21	2.12	0.48
2:B:73:GLY:HA2	2:B:122:GLN:HE22	1.78	0.48

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:207:GLN:O	2:B:5:CYS:HA	2.14	0.47
2:B:243:LEU:HD23	2:B:245:ARG:HH21	1.79	0.47
1:A:42:GLN:H	7:A:4008:GOL:C3	2.26	0.47
1:A:35:ASN:H	2:B:2:ASP:HB2	1.79	0.47
1:A:47:VAL:HG21	1:A:94:ARG:O	2.15	0.47
1:A:221:ALA:HA	1:A:227:VAL:HG12	1.95	0.47
1:A:45:ILE:HA	1:A:46:PRO:HD3	1.62	0.47
1:A:48:SER:H	1:A:99:GLN:HG3	1.79	0.47
2:B:243:LEU:N	2:B:243:LEU:CD1	2.75	0.47
1:A:223:PRO:HB2	1:A:224:PRO:HD3	1.97	0.47
2:B:38:TRP:CG	2:B:39:PRO:HD2	2.49	0.46
2:B:240:ASN:OD1	2:B:241:PRO:N	2.48	0.46
1:A:42:GLN:C	1:A:44:THR:H	2.19	0.46
2:B:123:THR:OG1	7:B:4014:GOL:H2	2.15	0.46
2:B:239:ALA:O	2:B:240:ASN:HB2	2.16	0.46
1:A:143:ARG:HH21	7:A:4009:GOL:H12	1.81	0.46
1:A:46:PRO:O	1:A:49:GLU:HB3	2.15	0.46
2:B:249:TYR:CZ	7:B:4006:GOL:H12	2.51	0.46
2:B:199:VAL:O	2:B:200:SER:HB2	2.16	0.46
1:A:106:ARG:HH11	1:A:106:ARG:HB2	1.79	0.45
1:A:247:CYS:SG	2:B:3:VAL:CG1	3.04	0.45
1:A:112:ASN:ND2	3:A:5003:NAG:C1	2.79	0.45
1:A:90:LYS:O	1:A:112:ASN:HB3	2.17	0.45
1:A:144:PHE:HA	1:A:145:PRO:HD3	1.79	0.45
2:B:249:TYR:CG	2:B:250:PRO:HD2	2.52	0.44
2:B:38:TRP:CD1	2:B:39:PRO:HD2	2.53	0.44
1:A:95:GLY:O	1:A:98:THR:HG23	2.18	0.43
2:B:61:ASN:HD21	3:B:5004:NAG:C2	2.30	0.43
1:A:117:ASP:O	1:A:120:ARG:HB2	2.17	0.43
1:A:47:VAL:H	1:A:48:SER:C	2.16	0.43
2:B:261:PRO:HG2	7:B:4018:GOL:H31	2.01	0.43
1:A:119:GLU:OE2	1:A:125:ARG:HG3	2.19	0.43
1:A:41:ARG:HA	7:A:4008:GOL:H32	2.01	0.43
1:A:35:ASN:O	1:A:36:ASN:HB2	2.19	0.43
2:B:194:ASN:HD21	2:B:214:GLY:HA2	1.84	0.42
1:A:59:THR:HA	1:A:64:ASP:O	2.19	0.42
1:A:149:THR:H	7:A:4020:GOL:H11	1.83	0.42
1:A:90:LYS:O	1:A:91:ASP:HB2	2.20	0.42
2:B:203:ILE:HD12	2:B:246:ILE:HG22	2.01	0.42
1:A:46:PRO:HB3	1:A:47:VAL:HG12	2.02	0.42
2:B:249:TYR:CD2	2:B:250:PRO:HD2	2.55	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:27:PHE:HZ	1:A:51:SER:CB	2.33	0.42
3:A:5003:NAG:H61	8:A:3271:HOH:O	2.18	0.42
2:B:72:ALA:HB1	7:B:4014:GOL:H12	2.01	0.42
2:B:100:ILE:HD11	2:B:105:ASN:ND2	2.35	0.41
1:A:18:PHE:CD1	1:A:177:ARG:NH1	2.89	0.41
1:A:84:GLN:HE21	1:A:85:GLN:HG2	1.86	0.41
1:A:84:GLN:NE2	1:A:85:GLN:HG2	2.35	0.41
1:A:71:ASP:HB3	1:A:74:ASN:OD1	2.21	0.41
1:A:215:ASN:OD1	2:B:8:SER:HB3	2.21	0.41
1:A:45:ILE:HG21	1:A:45:ILE:HD13	1.69	0.41
1:A:46:PRO:CB	1:A:47:VAL:HG12	2.50	0.41
2:B:194:ASN:ND2	2:B:214:GLY:O	2.54	0.41
2:B:55:ASP:HB2	2:B:86:ARG:HH22	1.86	0.41
2:B:100:ILE:HD11	2:B:105:ASN:HD22	1.86	0.40
1:A:149:THR:N	7:A:4020:GOL:H11	2.35	0.40
1:A:175:ARG:NH2	7:A:4011:GOL:O2	2.54	0.40
1:A:35:ASN:H	2:B:2:ASP:HB3	1.87	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	247/249 (99%)	235 (95%)	8 (3%)	4 (2%)	12	3
2	B	261/263 (99%)	249 (95%)	10 (4%)	2 (1%)	24	11
All	All	508/512 (99%)	484 (95%)	18 (4%)	6 (1%)	16	5

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	49	GLU

Continued on next page...



*Continued from previous page...*

Mol	Chain	Res	Type
1	A	50	ALA
1	A	248	GLY
2	B	3	VAL
2	B	241	PRO
1	A	43	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	209/215 (97%)	200 (96%)	9 (4%)	35	20
2	B	217/222 (98%)	208 (96%)	9 (4%)	37	22
All	All	426/437 (98%)	408 (96%)	18 (4%)	36	21

All (18) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	23	LEU
1	A	47	VAL
1	A	84	GLN
1	A	98	THR
1	A	159	LEU
1	A	226	ASN
1	A	227	VAL
1	A	233	ILE
1	A	247	CYS
2	B	1	ASP
2	B	2	ASP
2	B	5	CYS
2	B	8	SER
2	B	105	ASN
2	B	143	VAL
2	B	161	TRP
2	B	176	LEU
2	B	194	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (12) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	36	ASN
1	A	84	GLN
1	A	124	HIS
1	A	127	GLN
2	B	44	ASN
2	B	61	ASN
2	B	82	ASN
2	B	96	ASN
2	B	105	ASN
2	B	122	GLN
2	B	190	GLN
2	B	194	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](#)

Of 25 ligands modelled in this entry, 2 are monoatomic - leaving 23 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
4	SO4	A	4001	-	4,4,4	1.93	2 (50%)	6,6,6	0.55	0
5	AZI	A	4004	-	0,2,2	0.00	-	0,1,1	0.00	-
7	GOL	A	4007	-	5,5,5	0.32	0	5,5,5	0.45	0
7	GOL	A	4008	-	5,5,5	0.37	0	5,5,5	0.41	0
7	GOL	A	4009	-	5,5,5	0.38	0	5,5,5	0.37	0
7	GOL	A	4010	-	5,5,5	0.32	0	5,5,5	0.35	0
7	GOL	A	4011	-	5,5,5	0.31	0	5,5,5	0.40	0
7	GOL	A	4016	-	5,5,5	0.34	0	5,5,5	0.50	0
7	GOL	A	4019	-	5,5,5	0.33	0	5,5,5	0.42	0
7	GOL	A	4020	-	5,5,5	0.37	0	5,5,5	0.47	0
3	NAG	A	5003	-	14,14,15	0.51	0	15,19,21	0.84	1 (6%)
5	AZI	B	4003	-	0,2,2	0.00	-	0,1,1	0.00	-
5	AZI	B	4005	-	0,2,2	0.00	-	0,1,1	0.00	-
7	GOL	B	4006	-	5,5,5	0.38	0	5,5,5	0.28	0
7	GOL	B	4012	-	5,5,5	0.35	0	5,5,5	0.23	0
7	GOL	B	4013	-	5,5,5	0.34	0	5,5,5	0.36	0
7	GOL	B	4014	-	5,5,5	0.32	0	5,5,5	0.17	0
7	GOL	B	4015	-	5,5,5	0.37	0	5,5,5	0.51	0
7	GOL	B	4017	-	5,5,5	0.32	0	5,5,5	0.33	0
7	GOL	B	4018	-	5,5,5	0.35	0	5,5,5	0.47	0
3	NAG	B	5001	2	14,14,15	0.61	0	15,19,21	1.04	1 (6%)
3	NAG	B	5002	-	14,14,15	0.51	0	15,19,21	1.52	2 (13%)
3	NAG	B	5004	-	14,14,15	0.52	0	15,19,21	0.84	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
4	SO4	A	4001	-	-	0/0/0/0	0/0/0/0
5	AZI	A	4004	-	-	0/0/0/0	0/0/0/0
7	GOL	A	4007	-	-	0/4/4/4	0/0/0/0
7	GOL	A	4008	-	-	0/4/4/4	0/0/0/0
7	GOL	A	4009	-	-	0/4/4/4	0/0/0/0
7	GOL	A	4010	-	-	0/4/4/4	0/0/0/0
7	GOL	A	4011	-	-	0/4/4/4	0/0/0/0
7	GOL	A	4016	-	-	0/4/4/4	0/0/0/0
7	GOL	A	4019	-	-	0/4/4/4	0/0/0/0
7	GOL	A	4020	-	-	0/4/4/4	0/0/0/0
3	NAG	A	5003	-	-	0/6/23/26	0/1/1/1
5	AZI	B	4003	-	-	0/0/0/0	0/0/0/0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	AZI	B	4005	-	-	0/0/0/0	0/0/0/0
7	GOL	B	4006	-	-	0/4/4/4	0/0/0/0
7	GOL	B	4012	-	-	0/4/4/4	0/0/0/0
7	GOL	B	4013	-	-	0/4/4/4	0/0/0/0
7	GOL	B	4014	-	-	0/4/4/4	0/0/0/0
7	GOL	B	4015	-	-	0/4/4/4	0/0/0/0
7	GOL	B	4017	-	-	0/4/4/4	0/0/0/0
7	GOL	B	4018	-	-	0/4/4/4	0/0/0/0
3	NAG	B	5001	2	-	0/6/23/26	0/1/1/1
3	NAG	B	5002	-	-	0/6/23/26	0/1/1/1
3	NAG	B	5004	-	-	0/6/23/26	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	4001	SO4	O4-S	-2.21	1.39	1.47
4	A	4001	SO4	O1-S	-2.03	1.40	1.47

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	5002	NAG	C2-N2-C7	-3.31	118.79	123.04
3	A	5003	NAG	C2-N2-C7	-2.23	120.17	123.04
3	B	5001	NAG	C2-N2-C7	2.25	125.93	123.04
3	B	5002	NAG	C1-O5-C5	4.21	117.58	112.25

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

15 monomers are involved in 37 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	4001	SO4	3	0
7	A	4008	GOL	3	0
7	A	4009	GOL	2	0
7	A	4010	GOL	3	0
7	A	4011	GOL	1	0
7	A	4019	GOL	2	0
7	A	4020	GOL	5	0
3	A	5003	NAG	3	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	B	4006	GOL	1	0
7	B	4013	GOL	1	0
7	B	4014	GOL	3	0
7	B	4017	GOL	4	0
7	B	4018	GOL	3	0
3	B	5002	NAG	1	0
3	B	5004	NAG	3	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	249/249 (100%)	0.16	19 (7%)	17 26	32, 46, 76, 105	0
2	B	263/263 (100%)	-0.10	16 (6%)	25 34	25, 37, 67, 107	0
All	All	512/512 (100%)	0.03	35 (6%)	20 30	25, 42, 74, 107	0

All (35) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	240	ASN	10.0
2	B	239	ALA	9.8
2	B	243	LEU	9.4
1	A	223	PRO	8.2
1	A	247	CYS	7.6
1	A	225	GLY	5.9
2	B	241	PRO	5.7
1	A	50	ALA	5.4
2	B	244	ARG	5.1
2	B	242	LYS	4.8
2	B	2	ASP	4.6
1	A	248	GLY	4.6
2	B	238	GLN	4.4
2	B	1	ASP	4.4
1	A	51	SER	4.2
1	A	115	TYR	3.8
1	A	145	PRO	3.6
2	B	5	CYS	3.6
1	A	47	VAL	3.3
1	A	249	GLU	3.0
1	A	102	THR	2.9
2	B	161	TRP	2.9
2	B	6	SER	2.8
1	A	44	THR	2.8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	46	PRO	2.8
1	A	226	ASN	2.7
2	B	95	GLY	2.6
1	A	84	GLN	2.4
1	A	95	GLY	2.4
2	B	168	GLY	2.3
1	A	45	ILE	2.2
1	A	48	SER	2.2
2	B	237	ALA	2.1
2	B	4	THR	2.1
1	A	76	TYR	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 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
3	NAG	A	5003	14/15	0.72	0.29	3.82	77,91,96,106	0
7	GOL	B	4018	6/6	0.88	0.18	3.80	48,58,61,69	0
7	GOL	A	4010	6/6	0.68	0.27	3.05	70,80,84,88	0
5	AZI	B	4005	3/3	0.70	0.27	2.86	41,41,51,58	0
7	GOL	B	4013	6/6	0.79	0.15	2.05	60,63,85,88	0
7	GOL	A	4016	6/6	0.79	0.16	1.80	49,51,58,61	0
5	AZI	A	4004	3/3	0.96	0.13	1.75	42,42,44,55	0
7	GOL	A	4019	6/6	0.76	0.22	0.47	69,74,76,80	0
7	GOL	A	4011	6/6	0.81	0.20	0.25	61,68,80,81	0
3	NAG	B	5001	14/15	0.89	0.11	0.22	32,43,57,57	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
7	GOL	B	4015	6/6	0.95	0.08	-0.09	31,44,47,52	0
7	GOL	A	4008	6/6	0.91	0.12	-0.12	62,70,73,78	0
7	GOL	B	4012	6/6	0.97	0.08	-0.20	34,35,44,45	0
7	GOL	B	4006	6/6	0.91	0.12	-0.45	61,67,68,74	0
6	CL	A	4021	1/1	0.98	0.03	-	71,71,71,71	0
3	NAG	B	5002	14/15	0.60	0.33	-	51,81,90,93	0
7	GOL	B	4014	6/6	0.87	0.26	-	31,46,64,66	0
7	GOL	A	4009	6/6	0.86	0.25	-	56,66,72,80	0
6	CL	A	4022	1/1	0.89	0.15	-	58,58,58,58	1
4	SO4	A	4001	5/5	0.96	0.17	-	53,56,60,63	0
5	AZI	B	4003	3/3	0.96	0.08	-	47,47,56,57	0
3	NAG	B	5004	14/15	0.73	0.31	-	51,78,85,90	0
7	GOL	A	4007	6/6	0.84	0.27	-	77,82,88,88	0
7	GOL	B	4017	6/6	0.78	0.29	-	67,70,80,87	0
7	GOL	A	4020	6/6	0.82	0.18	-	54,60,63,80	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.