



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 19, 2016 – 07:19 PM GMT

PDB ID : 4RZS  
Title : Lac repressor engineered to bind sucralose, unliganded tetramer  
Authors : Arbing, M.A.; Cascio, D.; Kosuri, S.; Church, G.M.  
Deposited on : 2014-12-24  
Resolution : 2.71 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.  
We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>  
with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Mogul : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026982  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20026982

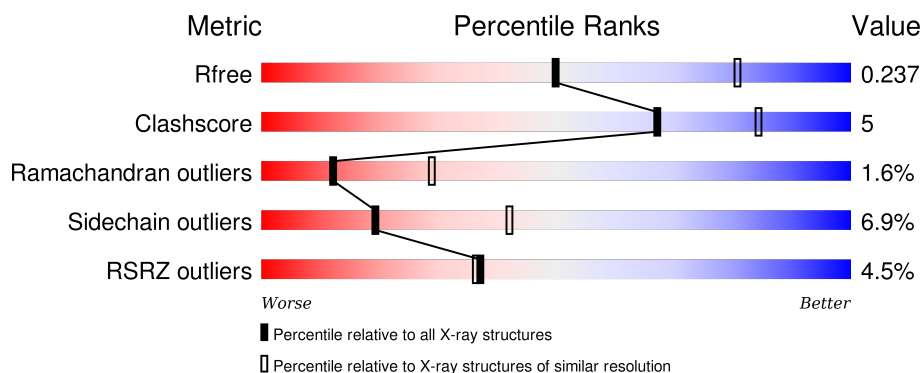
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.71 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	2439 (2.74-2.70)
Clashscore	102246	2771 (2.74-2.70)
Ramachandran outliers	100387	2726 (2.74-2.70)
Sidechain outliers	100360	2727 (2.74-2.70)
RSRZ outliers	91569	2443 (2.74-2.70)

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	381	<div> <div>2%</div> <div>79%</div> <div>10%</div> <div>9%</div> </div>
1	B	381	<div> <div>10%</div> <div>79%</div> <div>10%</div> <div>7%</div> </div>
1	C	381	<div> <div>2%</div> <div>70%</div> <div>7%</div> <div>22%</div> </div>
1	D	381	<div> <div>%</div> <div>68%</div> <div>8%</div> <div>22%</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
2	GOL	C	501	-	-	-	X
2	GOL	D	501	-	-	-	X

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 9530 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 Lac repressor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	347	Total	C	N	O	S	0	1	0
			2547	1587	451	497	12			
1	B	355	Total	C	N	O	S	0	7	0
			2595	1619	459	505	12			
1	C	298	Total	C	N	O	S	0	0	0
			2157	1343	384	419	11			
1	D	296	Total	C	N	O	S	0	1	0
			2184	1363	388	422	11			

There are 100 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-20	MET	-	EXPRESSION TAG	UNP C9QQT3
A	-19	GLY	-	EXPRESSION TAG	UNP C9QQT3
A	-18	SER	-	EXPRESSION TAG	UNP C9QQT3
A	-17	SER	-	EXPRESSION TAG	UNP C9QQT3
A	-16	HIS	-	EXPRESSION TAG	UNP C9QQT3
A	-15	HIS	-	EXPRESSION TAG	UNP C9QQT3
A	-14	HIS	-	EXPRESSION TAG	UNP C9QQT3
A	-13	HIS	-	EXPRESSION TAG	UNP C9QQT3
A	-12	HIS	-	EXPRESSION TAG	UNP C9QQT3
A	-11	HIS	-	EXPRESSION TAG	UNP C9QQT3
A	-10	SER	-	EXPRESSION TAG	UNP C9QQT3
A	-9	SER	-	EXPRESSION TAG	UNP C9QQT3
A	-8	GLY	-	EXPRESSION TAG	UNP C9QQT3
A	-7	LEU	-	EXPRESSION TAG	UNP C9QQT3
A	-6	VAL	-	EXPRESSION TAG	UNP C9QQT3
A	-5	PRO	-	EXPRESSION TAG	UNP C9QQT3
A	-4	ARG	-	EXPRESSION TAG	UNP C9QQT3
A	-3	GLY	-	EXPRESSION TAG	UNP C9QQT3
A	-2	SER	-	EXPRESSION TAG	UNP C9QQT3
A	-1	HIS	-	EXPRESSION TAG	UNP C9QQT3
A	0	MET	-	EXPRESSION TAG	UNP C9QQT3

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	149	THR	ASP	ENGINEERED MUTATION	UNP C9QQT3
A	150	ALA	VAL	ENGINEERED MUTATION	UNP C9QQT3
A	156	LEU	ILE	ENGINEERED MUTATION	UNP C9QQT3
A	193	ASP	SER	ENGINEERED MUTATION	UNP C9QQT3
B	-20	MET	-	EXPRESSION TAG	UNP C9QQT3
B	-19	GLY	-	EXPRESSION TAG	UNP C9QQT3
B	-18	SER	-	EXPRESSION TAG	UNP C9QQT3
B	-17	SER	-	EXPRESSION TAG	UNP C9QQT3
B	-16	HIS	-	EXPRESSION TAG	UNP C9QQT3
B	-15	HIS	-	EXPRESSION TAG	UNP C9QQT3
B	-14	HIS	-	EXPRESSION TAG	UNP C9QQT3
B	-13	HIS	-	EXPRESSION TAG	UNP C9QQT3
B	-12	HIS	-	EXPRESSION TAG	UNP C9QQT3
B	-11	HIS	-	EXPRESSION TAG	UNP C9QQT3
B	-10	SER	-	EXPRESSION TAG	UNP C9QQT3
B	-9	SER	-	EXPRESSION TAG	UNP C9QQT3
B	-8	GLY	-	EXPRESSION TAG	UNP C9QQT3
B	-7	LEU	-	EXPRESSION TAG	UNP C9QQT3
B	-6	VAL	-	EXPRESSION TAG	UNP C9QQT3
B	-5	PRO	-	EXPRESSION TAG	UNP C9QQT3
B	-4	ARG	-	EXPRESSION TAG	UNP C9QQT3
B	-3	GLY	-	EXPRESSION TAG	UNP C9QQT3
B	-2	SER	-	EXPRESSION TAG	UNP C9QQT3
B	-1	HIS	-	EXPRESSION TAG	UNP C9QQT3
B	0	MET	-	EXPRESSION TAG	UNP C9QQT3
B	149	THR	ASP	ENGINEERED MUTATION	UNP C9QQT3
B	150	ALA	VAL	ENGINEERED MUTATION	UNP C9QQT3
B	156	LEU	ILE	ENGINEERED MUTATION	UNP C9QQT3
B	193	ASP	SER	ENGINEERED MUTATION	UNP C9QQT3
C	-20	MET	-	EXPRESSION TAG	UNP C9QQT3
C	-19	GLY	-	EXPRESSION TAG	UNP C9QQT3
C	-18	SER	-	EXPRESSION TAG	UNP C9QQT3
C	-17	SER	-	EXPRESSION TAG	UNP C9QQT3
C	-16	HIS	-	EXPRESSION TAG	UNP C9QQT3
C	-15	HIS	-	EXPRESSION TAG	UNP C9QQT3
C	-14	HIS	-	EXPRESSION TAG	UNP C9QQT3
C	-13	HIS	-	EXPRESSION TAG	UNP C9QQT3
C	-12	HIS	-	EXPRESSION TAG	UNP C9QQT3
C	-11	HIS	-	EXPRESSION TAG	UNP C9QQT3
C	-10	SER	-	EXPRESSION TAG	UNP C9QQT3
C	-9	SER	-	EXPRESSION TAG	UNP C9QQT3
C	-8	GLY	-	EXPRESSION TAG	UNP C9QQT3

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	-7	LEU	-	EXPRESSION TAG	UNP C9QQT3
C	-6	VAL	-	EXPRESSION TAG	UNP C9QQT3
C	-5	PRO	-	EXPRESSION TAG	UNP C9QQT3
C	-4	ARG	-	EXPRESSION TAG	UNP C9QQT3
C	-3	GLY	-	EXPRESSION TAG	UNP C9QQT3
C	-2	SER	-	EXPRESSION TAG	UNP C9QQT3
C	-1	HIS	-	EXPRESSION TAG	UNP C9QQT3
C	0	MET	-	EXPRESSION TAG	UNP C9QQT3
C	149	THR	ASP	ENGINEERED MUTATION	UNP C9QQT3
C	150	ALA	VAL	ENGINEERED MUTATION	UNP C9QQT3
C	156	LEU	ILE	ENGINEERED MUTATION	UNP C9QQT3
C	193	ASP	SER	ENGINEERED MUTATION	UNP C9QQT3
D	-20	MET	-	EXPRESSION TAG	UNP C9QQT3
D	-19	GLY	-	EXPRESSION TAG	UNP C9QQT3
D	-18	SER	-	EXPRESSION TAG	UNP C9QQT3
D	-17	SER	-	EXPRESSION TAG	UNP C9QQT3
D	-16	HIS	-	EXPRESSION TAG	UNP C9QQT3
D	-15	HIS	-	EXPRESSION TAG	UNP C9QQT3
D	-14	HIS	-	EXPRESSION TAG	UNP C9QQT3
D	-13	HIS	-	EXPRESSION TAG	UNP C9QQT3
D	-12	HIS	-	EXPRESSION TAG	UNP C9QQT3
D	-11	HIS	-	EXPRESSION TAG	UNP C9QQT3
D	-10	SER	-	EXPRESSION TAG	UNP C9QQT3
D	-9	SER	-	EXPRESSION TAG	UNP C9QQT3
D	-8	GLY	-	EXPRESSION TAG	UNP C9QQT3
D	-7	LEU	-	EXPRESSION TAG	UNP C9QQT3
D	-6	VAL	-	EXPRESSION TAG	UNP C9QQT3
D	-5	PRO	-	EXPRESSION TAG	UNP C9QQT3
D	-4	ARG	-	EXPRESSION TAG	UNP C9QQT3
D	-3	GLY	-	EXPRESSION TAG	UNP C9QQT3
D	-2	SER	-	EXPRESSION TAG	UNP C9QQT3
D	-1	HIS	-	EXPRESSION TAG	UNP C9QQT3
D	0	MET	-	EXPRESSION TAG	UNP C9QQT3
D	149	THR	ASP	ENGINEERED MUTATION	UNP C9QQT3
D	150	ALA	VAL	ENGINEERED MUTATION	UNP C9QQT3
D	156	LEU	ILE	ENGINEERED MUTATION	UNP C9QQT3
D	193	ASP	SER	ENGINEERED MUTATION	UNP C9QQT3

- Molecule 2 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
2	A	1	Total	C	O	0	0
			6	3	3		
2	B	1	Total	C	O	0	0
			6	3	3		
2	C	1	Total	C	O	0	0
			6	3	3		
2	D	1	Total	C	O	0	0
			6	3	3		

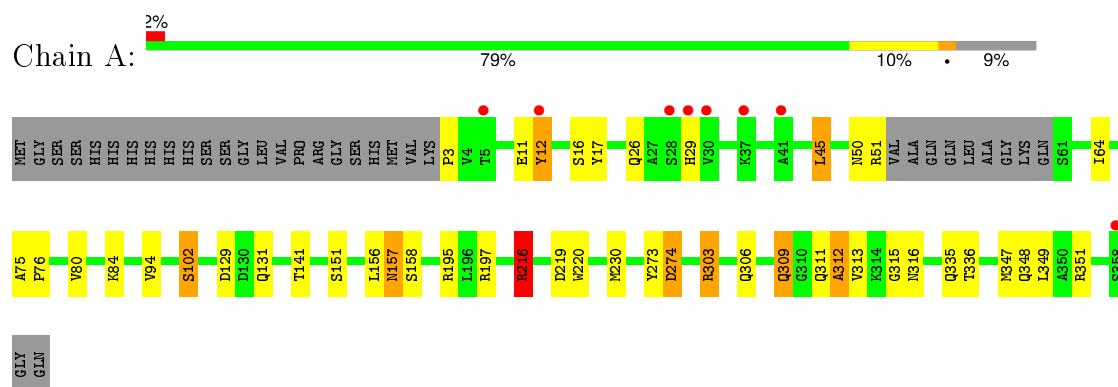
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	12	Total	O	0	0
			12	12		
3	B	4	Total	O	0	0
			4	4		
3	C	2	Total	O	0	0
			2	2		
3	D	5	Total	O	0	0
			5	5		

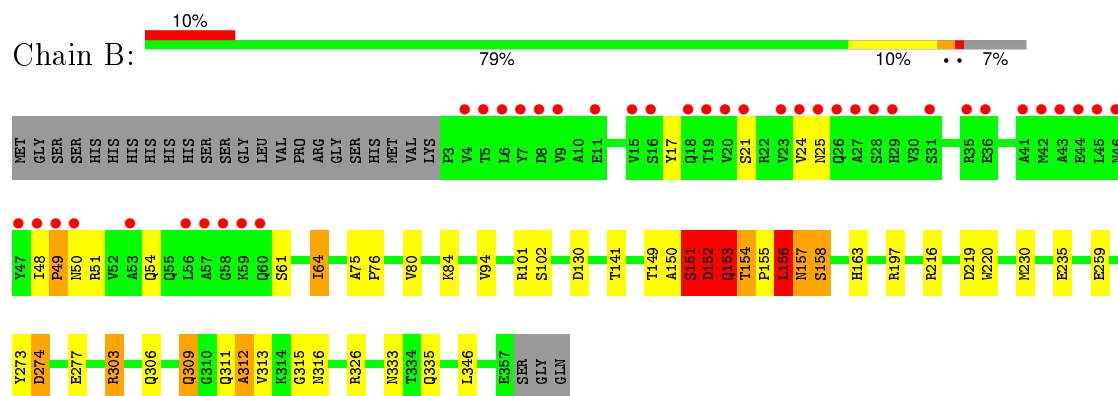
### 3 Residue-property plots

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

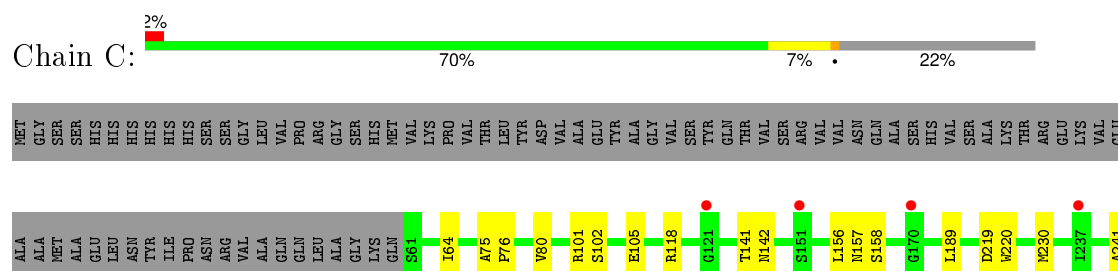
- Molecule 1: Lac repressor



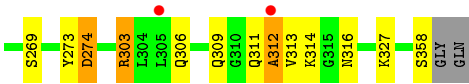
- Molecule 1: Lac repressor



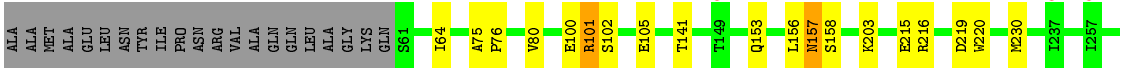
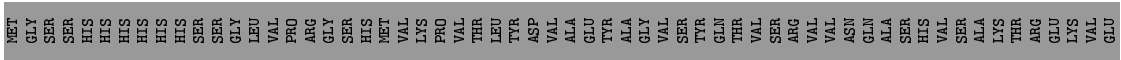
- Molecule 1: Lac repressor







● Molecule 1: Lac repressor



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	91.09Å 111.60Å 189.33Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	94.49 – 2.71 82.08 – 2.71	Depositor EDS
% Data completeness (in resolution range)	68.1 (94.49-2.71) 68.2 (82.08-2.71)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.03 (at 2.69Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, $R_{free}$	0.199 , 0.239 0.200 , 0.237	Depositor DCC
$R_{free}$ test set	3638 reflections (11.11%)	DCC
Wilson B-factor (Å <sup>2</sup> )	73.3	Xtriage
Anisotropy	0.074	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 40.7	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	0 of 36386 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	9530	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	79.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.17% 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: 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	0.78	0/2586	0.81	3/3527 (0.1%)
1	B	10.85	5/2635 (0.2%)	2.24	9/3600 (0.2%)
1	C	0.68	0/2186	0.76	0/2982
1	D	0.73	0/2214	0.80	2/3015 (0.1%)
All	All	5.71	5/9621 (0.1%)	1.35	14/13124 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	B	0	5
1	D	0	1
All	All	0	7

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	152[A]	ASP	CB-CG	357.17	9.01	1.51
1	B	152[B]	ASP	CB-CG	357.17	9.01	1.51
1	B	151[A]	SER	CB-OG	163.70	3.55	1.42
1	B	151[B]	SER	CB-OG	163.70	3.55	1.42
1	B	277	GLU	CG-CD	5.15	1.59	1.51

All (14) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	152[A]	ASP	CB-CG-OD1	-68.77	56.41	118.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	152[B]	ASP	CB-CG-OD1	-68.77	56.41	118.30
1	B	152[A]	ASP	CB-CG-OD2	49.96	163.26	118.30
1	B	152[B]	ASP	CB-CG-OD2	49.96	163.26	118.30
1	B	152[A]	ASP	CA-CB-CG	-25.52	57.27	113.40
1	B	152[B]	ASP	CA-CB-CG	-25.52	57.27	113.40
1	A	216	ARG	NE-CZ-NH1	6.11	123.35	120.30
1	B	216	ARG	NE-CZ-NH2	-5.95	117.32	120.30
1	A	216	ARG	NE-CZ-NH2	-5.61	117.50	120.30
1	B	50	ASN	N-CA-C	5.48	125.79	111.00
1	D	101	ARG	N-CA-C	-5.42	96.36	111.00
1	B	326	ARG	NE-CZ-NH1	5.42	123.01	120.30
1	D	216	ARG	NE-CZ-NH2	-5.17	117.72	120.30
1	A	347	MET	CA-CB-CG	5.09	121.96	113.30

There are no chirality outliers.

All (7) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	315	GLY	Peptide
1	B	149	THR	Peptide
1	B	151[B]	SER	Peptide
1	B	315	GLY	Peptide
1	B	49	PRO	Peptide
1	B	61	SER	Peptide
1	D	315	GLY	Peptide

## 5.2 Too-close contacts [\(i\)](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2547	0	2515	19	0
1	B	2595	0	2541	50	0
1	C	2157	0	2131	13	0
1	D	2184	0	2188	17	0
2	A	6	0	8	0	0
2	B	6	0	8	0	0
2	C	6	0	8	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	D	6	0	8	0	0
3	A	12	0	0	0	0
3	B	4	0	0	0	0
3	C	2	0	0	0	0
3	D	5	0	0	0	0
All	All	9530	0	9407	91	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 5.

All (91) 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:B:152[B]:ASP:HB3	1:B:153[B]:GLN:NE2	1.34	1.41
1:B:154[B]:THR:CG2	1:B:156[B]:LEU:HG	1.79	1.11
1:B:154[B]:THR:HG22	1:B:156[B]:LEU:HG	1.15	1.09
1:B:152[B]:ASP:CB	1:B:153[B]:GLN:NE2	2.16	1.08
1:B:130:ASP:CB	1:B:152[B]:ASP:HB2	1.97	0.94
1:B:151[B]:SER:OG	1:B:152[B]:ASP:N	1.99	0.87
1:B:153[B]:GLN:O	1:B:154[B]:THR:C	2.12	0.85
1:B:152[B]:ASP:HB3	1:B:153[B]:GLN:HE22	1.02	0.84
1:B:151[B]:SER:O	1:B:153[B]:GLN:N	2.13	0.82
1:B:154[B]:THR:CG2	1:B:156[B]:LEU:CG	2.58	0.81
1:B:155[B]:PRO:HG2	1:B:156[B]:LEU:N	1.97	0.80
1:B:51:ARG:CB	1:B:54:GLN:CB	2.61	0.79
1:B:154[B]:THR:HG21	1:B:156[B]:LEU:CD2	2.16	0.74
1:B:130:ASP:CB	1:B:152[B]:ASP:CB	2.70	0.70
1:A:11:GLU:HB3	1:D:203:LYS:HD2	1.75	0.69
1:D:75:ALA:HB3	1:D:76:PRO:HD3	1.75	0.68
1:B:75:ALA:HB3	1:B:76:PRO:HD3	1.76	0.68
1:B:154[B]:THR:HG21	1:B:156[B]:LEU:HD21	1.75	0.67
1:A:351:ARG:NH1	1:B:259:GLU:O	2.27	0.67
1:B:150[B]:ALA:CB	1:B:158:SER:HB2	2.25	0.66
1:B:157:ASN:HD22	1:B:157:ASN:N	1.93	0.66
1:B:155[B]:PRO:CG	1:B:156[B]:LEU:N	2.58	0.65
1:A:75:ALA:HB3	1:A:76:PRO:HD3	1.79	0.65
1:B:153[B]:GLN:O	1:B:154[B]:THR:O	2.14	0.64
1:B:154[B]:THR:HG22	1:B:156[B]:LEU:CG	2.10	0.63
1:D:355:ARG:O	1:D:356:LEU:C	2.38	0.63
1:B:152[A]:ASP:O	1:B:153[A]:GLN:C	2.36	0.62
1:A:157:ASN:HD22	1:A:157:ASN:N	1.96	0.62

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:17:TYR:CE2	1:D:153:GLN:HG3	2.35	0.62
1:D:157:ASN:HD22	1:D:157:ASN:N	1.98	0.62
1:C:75:ALA:HB3	1:C:76:PRO:HD3	1.81	0.61
1:D:340:ARG:HD3	1:D:340:ARG:O	2.02	0.58
1:B:235:GLU:HA	1:C:327:LYS:NZ	2.19	0.58
1:B:155[B]:PRO:HG2	1:B:156[B]:LEU:H	1.70	0.56
1:B:154[B]:THR:HB	1:B:156[B]:LEU:HD11	1.86	0.56
1:B:154[B]:THR:HG21	1:B:156[B]:LEU:CG	2.35	0.54
1:C:306:GLN:HB3	1:C:312:ALA:HB2	1.90	0.53
1:A:306:GLN:HB3	1:A:312:ALA:HB2	1.90	0.53
1:D:353:VAL:HA	1:D:356:LEU:HD12	1.89	0.53
1:D:306:GLN:HB3	1:D:312:ALA:HB2	1.92	0.51
1:B:306:GLN:HB3	1:B:312:ALA:HB2	1.92	0.51
1:A:273:TYR:O	1:A:274:ASP:CB	2.59	0.51
1:A:273:TYR:O	1:A:274:ASP:HB2	2.11	0.50
1:C:273:TYR:O	1:C:274:ASP:HB2	2.12	0.50
1:C:273:TYR:O	1:C:274:ASP:CB	2.61	0.49
1:B:303:ARG:NH2	1:B:313:VAL:O	2.46	0.49
1:D:273:TYR:O	1:D:274:ASP:HB2	2.13	0.49
1:B:273:TYR:O	1:B:274:ASP:CB	2.60	0.48
1:B:273:TYR:O	1:B:274:ASP:HB2	2.12	0.48
1:C:303:ARG:NH2	1:C:313:VAL:O	2.46	0.48
1:D:303:ARG:NH2	1:D:313:VAL:O	2.47	0.48
1:B:21:SER:O	1:B:25:ASN:ND2	2.46	0.48
1:D:273:TYR:O	1:D:274:ASP:CB	2.62	0.48
1:A:303:ARG:NH2	1:A:313:VAL:O	2.47	0.47
1:B:155[B]:PRO:CG	1:B:156[B]:LEU:H	2.27	0.47
1:D:355:ARG:O	1:D:356:LEU:O	2.33	0.47
1:B:163:HIS:CD2	1:B:163:HIS:H	2.32	0.47
1:A:129:ASP:HB3	1:A:195:ARG:NH1	2.30	0.46
1:B:154[B]:THR:O	1:B:155[B]:PRO:C	2.51	0.46
1:B:76:PRO:O	1:B:80:VAL:HG23	2.16	0.45
1:B:157:ASN:N	1:B:157:ASN:ND2	2.61	0.45
1:A:16:SER:HB2	1:D:153:GLN:HE22	1.81	0.45
1:A:3:PRO:HB2	1:D:215:GLU:OE2	2.16	0.45
1:B:333:ASN:OD1	1:B:335:GLN:HB2	2.17	0.45
1:B:154[B]:THR:HB	1:B:156[B]:LEU:CD1	2.46	0.45
1:B:309:GLN:HB2	1:B:309:GLN:HE21	1.61	0.44
1:B:154[B]:THR:CG2	1:B:156[B]:LEU:CD2	2.89	0.44
1:C:76:PRO:O	1:C:80:VAL:HG23	2.18	0.44
1:B:151[B]:SER:HG	1:B:152[B]:ASP:H	1.54	0.44

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:157:ASN:N	1:A:157:ASN:ND2	2.65	0.44
1:B:48:ILE:CB	1:B:49:PRO:HD2	2.48	0.44
1:A:309:GLN:HB2	1:A:309:GLN:HE21	1.63	0.44
1:A:219:ASP:O	1:A:220:TRP:HB2	2.18	0.44
1:B:153[B]:GLN:C	1:B:154[B]:THR:O	2.56	0.43
1:A:76:PRO:O	1:A:80:VAL:HG23	2.18	0.43
1:A:216:ARG:HH11	1:A:216:ARG:HG2	1.84	0.43
1:C:157:ASN:HD21	1:C:314:LYS:HA	1.84	0.43
1:D:76:PRO:O	1:D:80:VAL:HG23	2.19	0.43
1:B:235:GLU:HA	1:C:327:LYS:HZ1	1.84	0.43
1:B:235:GLU:HA	1:C:327:LYS:HZ3	1.85	0.42
1:D:100:GLU:HA	1:D:101:ARG:CB	2.49	0.42
1:D:219:ASP:O	1:D:220:TRP:HB2	2.20	0.42
1:B:219:ASP:O	1:B:220:TRP:HB2	2.19	0.41
1:C:241:ALA:HA	1:C:269:SER:O	2.20	0.41
1:B:152[B]:ASP:C	1:B:153[B]:GLN:CG	2.89	0.41
1:B:84:LYS:HA	1:B:94:VAL:HG21	2.03	0.41
1:A:84:LYS:HA	1:A:94:VAL:HG21	2.03	0.41
1:C:219:ASP:O	1:C:220:TRP:HB2	2.21	0.40
1:B:64:ILE:HD13	1:B:64:ILE:HG21	1.81	0.40
1:C:189:LEU:HA	1:C:189:LEU:HD23	1.95	0.40
1:A:12:TYR:CE1	1:A:45:LEU:HD21	2.57	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	A	344/381 (90%)	320 (93%)	20 (6%)	4 (1%)	16 38
1	B	360/381 (94%)	330 (92%)	17 (5%)	13 (4%)	4 9
1	C	296/381 (78%)	280 (95%)	12 (4%)	4 (1%)	14 34

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	D	295/381 (77%)	282 (96%)	10 (3%)	3 (1%)	19	44
All	All	1295/1524 (85%)	1212 (94%)	59 (5%)	24 (2%)	12	24

All (24) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	152[A]	ASP
1	B	152[B]	ASP
1	B	156[A]	LEU
1	B	156[B]	LEU
1	C	101	ARG
1	D	312	ALA
1	A	274	ASP
1	A	312	ALA
1	B	101	ARG
1	B	153[A]	GLN
1	B	153[B]	GLN
1	B	274	ASP
1	B	312	ALA
1	C	274	ASP
1	C	312	ALA
1	D	274	ASP
1	A	141	THR
1	B	141	THR
1	D	141	THR
1	B	154[A]	THR
1	B	154[B]	THR
1	C	141	THR
1	A	102	SER
1	B	24	VAL

### 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	267/311 (86%)	243 (91%)	24 (9%)	12	26
1	B	265/311 (85%)	247 (93%)	18 (7%)	20	43
1	C	224/311 (72%)	211 (94%)	13 (6%)	25	51
1	D	231/311 (74%)	215 (93%)	16 (7%)	19	42
All	All	987/1244 (79%)	916 (93%)	71 (7%)	19	40

All (71) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	12	TYR
1	A	26	GLN
1	A	29	HIS
1	A	45	LEU
1	A	50	ASN
1	A	51	ARG
1	A	64	ILE
1	A	102	SER
1	A	131	GLN
1	A	151	SER
1	A	156	LEU
1	A	157	ASN
1	A	158	SER
1	A	197	ARG
1	A	216	ARG
1	A	230	MET
1	A	303	ARG
1	A	309	GLN
1	A	311	GLN
1	A	316	ASN
1	A	335	GLN
1	A	336	THR
1	A	348	GLN
1	A	349	LEU
1	B	17	TYR
1	B	64	ILE
1	B	102	SER
1	B	151[A]	SER
1	B	151[B]	SER
1	B	153[A]	GLN
1	B	153[B]	GLN
1	B	156[A]	LEU
1	B	156[B]	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	157	ASN
1	B	158	SER
1	B	197	ARG
1	B	230	MET
1	B	303	ARG
1	B	309	GLN
1	B	311	GLN
1	B	316	ASN
1	B	346	LEU
1	C	64	ILE
1	C	102	SER
1	C	105	GLU
1	C	118	ARG
1	C	142	ASN
1	C	156	LEU
1	C	158	SER
1	C	230	MET
1	C	303	ARG
1	C	309	GLN
1	C	311	GLN
1	C	316	ASN
1	C	358	SER
1	D	64	ILE
1	D	102	SER
1	D	105	GLU
1	D	156	LEU
1	D	157	ASN
1	D	158	SER
1	D	230	MET
1	D	303	ARG
1	D	309	GLN
1	D	311	GLN
1	D	316	ASN
1	D	335	GLN
1	D	340	ARG
1	D	342	LEU
1	D	348	GLN
1	D	355	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (30) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	117	GLN
1	A	131	GLN
1	A	157	ASN
1	A	181	GLN
1	A	234	ASN
1	A	309	GLN
1	B	25	ASN
1	B	60	GLN
1	B	117	GLN
1	B	131	GLN
1	B	157	ASN
1	B	163	HIS
1	B	181	GLN
1	B	234	ASN
1	B	309	GLN
1	B	311	GLN
1	C	117	GLN
1	C	131	GLN
1	C	157	ASN
1	C	234	ASN
1	C	309	GLN
1	C	311	GLN
1	D	117	GLN
1	D	131	GLN
1	D	153	GLN
1	D	157	ASN
1	D	181	GLN
1	D	234	ASN
1	D	309	GLN
1	D	311	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 carbohydrates in this entry.

## 5.6 Ligand geometry

4 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	GOL	A	501	-	5,5,5	0.24	0	5,5,5	0.56	0
2	GOL	B	501	-	5,5,5	0.71	0	5,5,5	0.51	0
2	GOL	C	501	-	5,5,5	0.49	0	5,5,5	0.58	0
2	GOL	D	501	-	5,5,5	0.53	0	5,5,5	0.60	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
2	GOL	A	501	-	-	0/4/4/4	0/0/0/0
2	GOL	B	501	-	-	0/4/4/4	0/0/0/0
2	GOL	C	501	-	-	0/4/4/4	0/0/0/0
2	GOL	D	501	-	-	0/4/4/4	0/0/0/0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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	347/381 (91%)	-0.04	8 (2%) 64 64	33, 63, 136, 162	0
1	B	355/381 (93%)	0.32	39 (10%) 7 6	36, 68, 182, 208	1 (0%)
1	C	298/381 (78%)	0.14	6 (2%) 68 69	55, 91, 123, 144	0
1	D	296/381 (77%)	0.03	5 (1%) 73 74	47, 71, 101, 120	0
All	All	1296/1524 (85%)	0.12	58 (4%) 37 36	33, 74, 149, 208	1 (0%)

All (58) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	27	ALA	10.5
1	B	58	GLY	8.8
1	B	18	GLN	6.7
1	B	50	ASN	6.1
1	B	35	ARG	5.8
1	B	24	VAL	5.1
1	B	47	TYR	5.0
1	B	20	VAL	4.9
1	B	29	HIS	4.8
1	B	59	LYS	4.8
1	B	6	LEU	4.5
1	B	43	ALA	4.5
1	B	25	ASN	4.5
1	B	49	PRO	4.4
1	B	19	THR	4.4
1	B	60	GLN	4.3
1	B	5	THR	3.9
1	B	15	VAL	3.9
1	B	57	ALA	3.9
1	A	28	SER	3.8
1	B	23	VAL	3.6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	41	ALA	3.5
1	B	4	VAL	3.4
1	B	21	SER	3.3
1	B	8	ASP	3.3
1	B	42	MET	3.3
1	B	7	TYR	3.2
1	B	28	SER	3.1
1	B	46	ASN	2.9
1	B	36	GLU	2.9
1	B	56	LEU	2.7
1	B	9	VAL	2.7
1	B	31	SER	2.7
1	D	257	ILE	2.7
1	B	16	SER	2.6
1	B	45	LEU	2.6
1	C	312	ALA	2.5
1	B	44	GLU	2.4
1	B	53	ALA	2.4
1	A	12	TYR	2.4
1	B	48	ILE	2.3
1	C	170	GLY	2.3
1	D	310	GLY	2.3
1	C	151	SER	2.3
1	C	121	GLY	2.3
1	D	267	ASP	2.3
1	A	30	VAL	2.2
1	A	358	SER	2.2
1	A	41	ALA	2.2
1	A	37	LYS	2.2
1	D	149	THR	2.2
1	D	237	ILE	2.1
1	A	5	THR	2.1
1	B	26	GLN	2.1
1	C	237	ILE	2.1
1	A	29	HIS	2.1
1	C	305	LEU	2.1
1	B	11	GLU	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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

### 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

### 6.4 Ligands [i](#)

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	GOL	D	501	6/6	0.88	0.29	9.73	56,64,75,78	0
2	GOL	C	501	6/6	0.84	0.26	3.16	59,64,70,71	0
2	GOL	B	501	6/6	0.91	0.17	1.11	51,60,70,70	0
2	GOL	A	501	6/6	0.91	0.17	0.23	48,53,67,68	0

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

There are no such residues in this entry.