



# Full wwPDB X-ray Structure Validation Report ⓘ

Dec 12, 2023 – 04:20 pm GMT

PDB ID : 4AW4  
Title : Engineered variant of *Listeria monocytogenes* InlB internalin domain with an additional leucine rich repeat inserted  
Authors : Niemann, H.H.; Heinz, D.W.  
Deposited on : 2012-05-31  
Resolution : 1.93 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 : 1.8.4, CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.36

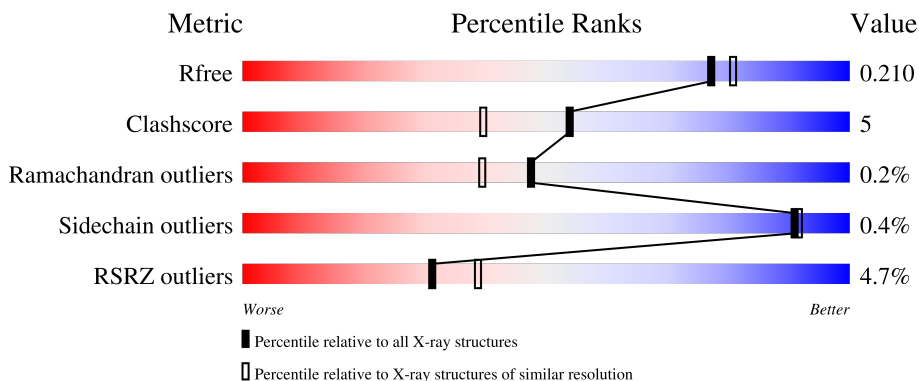
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.93 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))
$R_{free}$	130704	4310 (1.96-1.92)
Clashscore	141614	1023 (1.94-1.94)
Ramachandran outliers	138981	1007 (1.94-1.94)
Sidechain outliers	138945	1007 (1.94-1.94)
RSRZ outliers	127900	4250 (1.96-1.92)

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	311	 6% 86% 11% .
1	B	311	 4% 89% 9% ..
1	C	311	 4% 83% 14% ..

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	SO4	A	1348	-	-	X	-
3	GOL	B	1346	-	-	X	-

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 8080 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 INTERNALIN B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	301	2440	1559	409	471	1	0	10	0
1	B	307	2499	1599	414	485	1	0	12	0
1	C	307	2502	1600	415	486	1	0	12	0

There are 75 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	33	GLY	-	expression tag	UNP P25147
A	34	ALA	-	expression tag	UNP P25147
A	35	MET	-	expression tag	UNP P25147
A	99	ASN	-	insertion	UNP P25147
A	100	LEU	-	insertion	UNP P25147
A	101	THR	-	insertion	UNP P25147
A	102	SER	-	insertion	UNP P25147
A	103	LEU	-	insertion	UNP P25147
A	104	ASN	-	insertion	UNP P25147
A	105	LEU	-	insertion	UNP P25147
A	106	SER	-	insertion	UNP P25147
A	107	ASN	-	insertion	UNP P25147
A	108	ASN	-	insertion	UNP P25147
A	109	GLN	-	insertion	UNP P25147
A	110	ILE	-	insertion	UNP P25147
A	111	THR	-	insertion	UNP P25147
A	112	ASP	-	insertion	UNP P25147
A	113	ILE	-	insertion	UNP P25147
A	114	SER	-	insertion	UNP P25147
A	115	PRO	-	insertion	UNP P25147
A	116	ILE	-	insertion	UNP P25147
A	117	GLN	-	insertion	UNP P25147
A	118	TYR	-	insertion	UNP P25147

*Continued on next page...*

*Continued from previous page...*

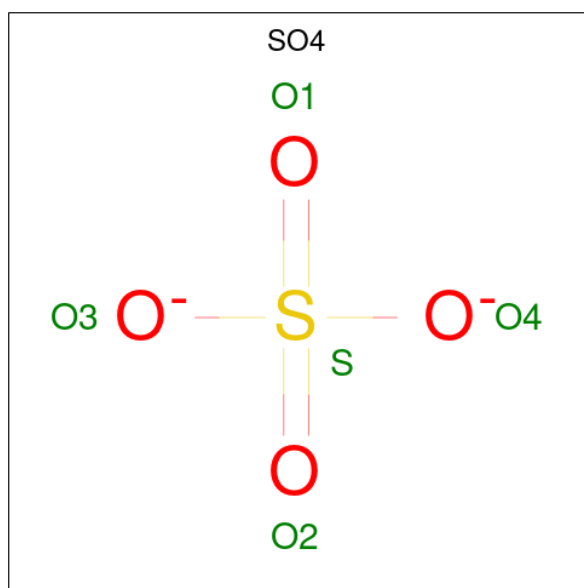
Chain	Residue	Modelled	Actual	Comment	Reference
A	119	LEU	-	insertion	UNP P25147
A	120	PRO	-	insertion	UNP P25147
B	33	GLY	-	expression tag	UNP P25147
B	34	ALA	-	expression tag	UNP P25147
B	35	MET	-	expression tag	UNP P25147
B	99	ASN	-	insertion	UNP P25147
B	100	LEU	-	insertion	UNP P25147
B	101	THR	-	insertion	UNP P25147
B	102	SER	-	insertion	UNP P25147
B	103	LEU	-	insertion	UNP P25147
B	104	ASN	-	insertion	UNP P25147
B	105	LEU	-	insertion	UNP P25147
B	106	SER	-	insertion	UNP P25147
B	107	ASN	-	insertion	UNP P25147
B	108	ASN	-	insertion	UNP P25147
B	109	GLN	-	insertion	UNP P25147
B	110	ILE	-	insertion	UNP P25147
B	111	THR	-	insertion	UNP P25147
B	112	ASP	-	insertion	UNP P25147
B	113	ILE	-	insertion	UNP P25147
B	114	SER	-	insertion	UNP P25147
B	115	PRO	-	insertion	UNP P25147
B	116	ILE	-	insertion	UNP P25147
B	117	GLN	-	insertion	UNP P25147
B	118	TYR	-	insertion	UNP P25147
B	119	LEU	-	insertion	UNP P25147
B	120	PRO	-	insertion	UNP P25147
C	33	GLY	-	expression tag	UNP P25147
C	34	ALA	-	expression tag	UNP P25147
C	35	MET	-	expression tag	UNP P25147
C	99	ASN	-	insertion	UNP P25147
C	100	LEU	-	insertion	UNP P25147
C	101	THR	-	insertion	UNP P25147
C	102	SER	-	insertion	UNP P25147
C	103	LEU	-	insertion	UNP P25147
C	104	ASN	-	insertion	UNP P25147
C	105	LEU	-	insertion	UNP P25147
C	106	SER	-	insertion	UNP P25147
C	107	ASN	-	insertion	UNP P25147
C	108	ASN	-	insertion	UNP P25147
C	109	GLN	-	insertion	UNP P25147
C	110	ILE	-	insertion	UNP P25147

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	111	THR	-	insertion	UNP P25147
C	112	ASP	-	insertion	UNP P25147
C	113	ILE	-	insertion	UNP P25147
C	114	SER	-	insertion	UNP P25147
C	115	PRO	-	insertion	UNP P25147
C	116	ILE	-	insertion	UNP P25147
C	117	GLN	-	insertion	UNP P25147
C	118	TYR	-	insertion	UNP P25147
C	119	LEU	-	insertion	UNP P25147
C	120	PRO	-	insertion	UNP P25147

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



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

Continued on next page...

*Continued from previous page...*

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

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

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	129	Total O 129 129	0	0

*Continued on next page...*



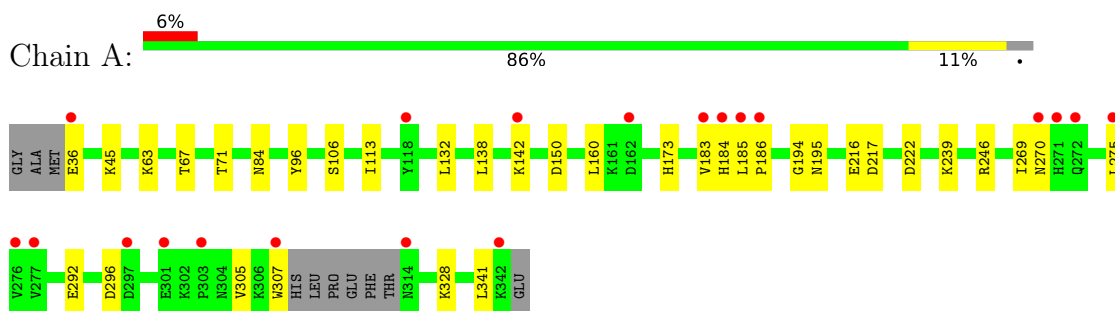
*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
4	B	188	Total 188	O 188	0	0
4	C	146	Total 146	O 146	0	0

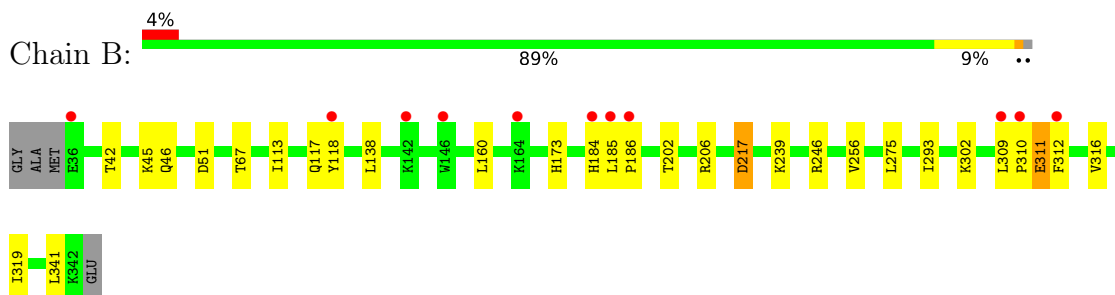
### 3 Residue-property plots [i](#)

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.

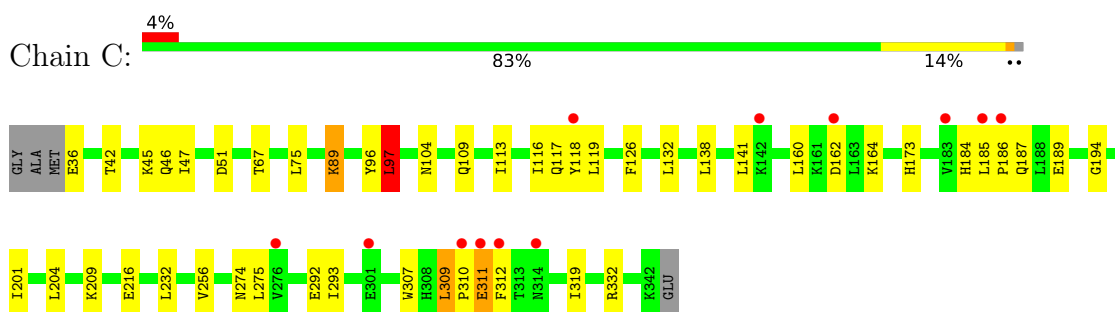
- Molecule 1: INTERNALIN B



- Molecule 1: INTERNALIN B



- Molecule 1: INTERNALIN B



## 4 Data and refinement statistics i

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	179.08Å 102.15Å 69.81Å 90.00° 90.23° 90.00°	Depositor
Resolution (Å)	19.69 – 1.93 19.69 – 1.95	Depositor EDS
% Data completeness (in resolution range)	94.3 (19.69-1.93) 94.4 (19.69-1.95)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.47 (at 1.94Å)	Xtrriage
Refinement program	REFMAC 5.6.0117	Depositor
R, $R_{free}$	0.167 , 0.205 0.173 , 0.210	Depositor DCC
$R_{free}$ test set	2016 reflections (2.26%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	24.7	Xtrriage
Anisotropy	0.591	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 32.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.28$	Xtrriage
Estimated twinning fraction	0.048 for -1/2*h-3/2*k,-1/2*h+1/2*k,-l 0.049 for -1/2*h+3/2*k,1/2*h+1/2*k,-l 0.347 for 1/2*h-3/2*k,-1/2*h-1/2*k,-l 0.348 for 1/2*h+3/2*k,1/2*h-1/2*k,-l 0.052 for -h,-k,l	Xtrriage
Reported twinning fraction	0.664 for H, K, L 0.192 for 1/2H-3/2K, -1/2H-1/2K, -L 0.144 for 1/2H+3/2K, 1/2H-1/2K, -L	Depositor
Outliers	0 of 89100 reflections	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	8080	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	35.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.31% 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 [i](#)

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

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, SO4

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
1	A	0.73	0/2502	0.83	1/3394 (0.0%)
1	B	0.77	0/2571	0.83	2/3493 (0.1%)
1	C	0.78	0/2571	0.86	2/3493 (0.1%)
All	All	0.76	0/7644	0.84	5/10380 (0.0%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	97	LEU	CB-CG-CD1	6.33	121.77	111.00
1	A	150	ASP	CB-CG-OD2	-6.03	112.87	118.30
1	C	232	LEU	CB-CG-CD1	-5.49	101.67	111.00
1	B	217	ASP	CB-CG-OD1	5.48	123.23	118.30
1	B	246	ARG	NE-CZ-NH2	-5.25	117.68	120.30

There are no chirality outliers.

There are no planarity outliers.

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

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2440	0	2538	27	1
1	B	2499	0	2587	24	0
1	C	2502	0	2586	34	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	45	0	0	4	0
2	B	35	0	0	2	0
2	C	30	0	0	2	0
3	A	30	0	40	2	0
3	B	24	0	32	6	0
3	C	12	0	16	1	0
4	A	129	0	0	3	0
4	B	188	0	0	0	0
4	C	146	0	0	2	0
All	All	8080	0	7799	81	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 5.

All (81) 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:C:293[B]:ILE:HG23	1:C:319:ILE:HB	1.66	0.78
1:C:173:HIS:ND1	2:C:1343:SO4:O4	2.19	0.74
1:A:195:ASN:ND2	1:A:217:ASP:OD2	2.23	0.71
1:A:173:HIS:ND1	2:A:1348:SO4:O1	2.21	0.71
1:C:113:ILE:HD12	1:C:138:LEU:HD21	1.72	0.70
1:B:302:LYS:NZ	2:B:1350:SO4:O4	2.22	0.65
1:C:162:ASP:O	1:C:164:LYS:HD3	1.97	0.64
1:B:113:ILE:HD12	1:B:138:LEU:HD21	1.80	0.64
1:A:142[B]:LYS:HG3	4:A:2059:HOH:O	1.96	0.64
1:B:311:GLU:OE2	1:B:312:PHE:O	2.18	0.62
1:B:117:GLN:HG3	1:B:118[B]:TYR:CD2	2.35	0.62
1:A:113:ILE:HD12	1:A:138:LEU:HD21	1.82	0.61
1:B:293[A]:ILE:HG23	1:B:319:ILE:HB	1.83	0.60
1:C:292:GLU:OE2	4:C:2135:HOH:O	2.16	0.60
1:B:256:VAL:HG13	1:C:51[B]:ASP:HB3	1.86	0.58
1:B:275:LEU:HD21	1:B:341:LEU:HD21	1.86	0.57
1:C:46[A]:GLN:HG3	4:C:2010:HOH:O	2.03	0.57
1:B:202:THR:HG22	3:B:1346:GOL:H32	1.88	0.56
1:C:116:ILE:HG22	1:C:141:LEU:HD11	1.88	0.56
1:A:113:ILE:HG23	1:A:132:LEU:HD21	1.88	0.55
1:A:71:THR:HG21	3:A:1351:GOL:H2	1.89	0.54
1:B:206:ARG:HE	3:B:1346:GOL:C2	2.20	0.54
1:B:173:HIS:ND1	2:B:1345:SO4:O1	2.34	0.54
1:C:113:ILE:HD12	1:C:138:LEU:CD2	2.36	0.54

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:42:THR:HG23	1:B:46[B]:GLN:OE1	2.08	0.53
1:A:45:LYS:HE2	1:A:67:THR:HG22	1.90	0.53
1:B:202:THR:CG2	3:B:1346:GOL:H12	2.39	0.53
1:C:118[B]:TYR:O	1:C:119:LEU:HD23	2.08	0.53
1:C:307:TRP:HB3	1:C:309:LEU:HD13	1.91	0.53
1:A:296:ASP:N	2:A:1346:SO4:O4	2.34	0.52
1:B:160:LEU:O	1:B:185:LEU:HD21	2.10	0.51
1:A:113:ILE:HD12	1:A:138:LEU:CD2	2.43	0.49
1:C:113:ILE:HG23	1:C:132:LEU:HD21	1.95	0.49
1:A:36:GLU:HB3	1:A:96:TYR:CE1	2.49	0.48
1:C:113:ILE:CD1	1:C:138:LEU:HD21	2.43	0.48
1:A:269:ILE:HG22	1:A:270:ASN:N	2.28	0.48
1:A:195:ASN:HB3	4:A:2079:HOH:O	2.13	0.48
1:C:189:GLU:OE2	1:C:209[B]:LYS:NZ	2.47	0.48
1:C:45:LYS:HE2	1:C:67:THR:HG22	1.96	0.47
1:A:160:LEU:O	1:A:185:LEU:HD21	2.14	0.47
1:C:160:LEU:O	1:C:185:LEU:HD21	2.15	0.47
1:B:309:LEU:HD11	1:B:316:VAL:HG11	1.98	0.46
1:B:45:LYS:HE2	1:B:67:THR:HG22	1.97	0.46
1:B:113:ILE:HD12	1:B:138:LEU:CD2	2.45	0.46
1:C:36:GLU:HB3	1:C:96:TYR:CE1	2.51	0.46
1:A:71:THR:HG21	3:A:1351:GOL:C2	2.45	0.46
1:C:274:ASN:HB2	3:C:1345:GOL:H32	1.97	0.45
1:A:183:VAL:HG13	1:A:184:HIS:N	2.31	0.45
1:C:310:PRO:O	1:C:311:GLU:HB3	2.16	0.45
1:B:310:PRO:O	1:B:311:GLU:HB3	2.16	0.45
1:B:206:ARG:HD2	3:B:1346:GOL:H2	1.98	0.44
1:A:305:VAL:CG1	1:A:307:TRP:CH2	3.00	0.44
1:A:63:LYS:HD2	1:C:293[A]:ILE:CD1	2.48	0.44
1:C:75:LEU:HB3	1:C:97:LEU:HD12	2.00	0.44
1:C:311:GLU:OE2	1:C:312:PHE:O	2.35	0.44
1:A:328:LYS:NZ	4:A:2124:HOH:O	2.41	0.43
1:A:292:GLU:O	2:A:1356:SO4:O1	2.36	0.43
1:A:305:VAL:HG11	1:A:307:TRP:CH2	2.54	0.42
1:B:202:THR:CG2	3:B:1346:GOL:C1	2.97	0.42
1:C:194:GLY:HA2	1:C:216:GLU:O	2.18	0.42
1:B:217:ASP:HA	1:B:239:LYS:O	2.19	0.42
1:B:51:ASP:HB3	1:C:256:VAL:HG13	2.01	0.42
1:C:184:HIS:C	1:C:186:PRO:HD3	2.40	0.42
1:C:104:ASN:HA	1:C:126:PHE:HB2	2.02	0.42
1:C:275:LEU:HD23	1:C:307:TRP:CE3	2.55	0.41

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:309:LEU:HD11	1:B:316:VAL:CG1	2.50	0.41
1:C:117:GLN:HG3	1:C:118[B]:TYR:CD2	2.54	0.41
1:B:202:THR:HG22	3:B:1346:GOL:C1	2.49	0.41
1:A:84:ASN:HA	1:A:106:SER:O	2.19	0.41
1:A:194:GLY:HA2	1:A:216:GLU:O	2.20	0.41
1:A:63:LYS:HD2	1:C:293[A]:ILE:HD11	2.02	0.41
1:A:275:LEU:HD21	1:A:341:LEU:HD21	2.03	0.41
1:C:201:ILE:HB	1:C:204:LEU:HD12	2.03	0.41
1:C:332:ARG:NE	2:C:1346:SO4:O1	2.48	0.41
1:A:173:HIS:CB	2:A:1348:SO4:O1	2.69	0.40
1:B:184:HIS:C	1:B:186:PRO:HD3	2.41	0.40
1:C:187[B]:GLN:O	1:C:187[B]:GLN:HG2	2.22	0.40
1:A:184:HIS:C	1:A:186:PRO:HD3	2.42	0.40
1:C:89:LYS:NZ	1:C:109:GLN:OE1	2.47	0.40
1:A:217:ASP:HA	1:A:239:LYS:O	2.20	0.40
1:C:42:THR:HB	1:C:47:ILE:HD11	2.03	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:222:ASP:OD2	1:A:246[B]:ARG:NH2[2_556]	2.19	0.01

## 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	307/311 (99%)	297 (97%)	10 (3%)	0	100 100
1	B	317/311 (102%)	303 (96%)	13 (4%)	1 (0%)	41 32
1	C	317/311 (102%)	303 (96%)	13 (4%)	1 (0%)	41 32
All	All	941/933 (101%)	903 (96%)	36 (4%)	2 (0%)	47 39

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	311	GLU
1	C	311	GLU

### 5.3.2 Protein sidechains [i](#)

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	289/287 (101%)	289 (100%)	0	100	100
1	B	297/287 (104%)	297 (100%)	0	100	100
1	C	297/287 (104%)	294 (99%)	3 (1%)	76	71
All	All	883/861 (103%)	880 (100%)	3 (0%)	91	93

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	89	LYS
1	C	97	LEU
1	C	309	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	195	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 monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

33 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	SO4	B	1344	-	4,4,4	0.35	0	6,6,6	0.34	0
2	SO4	B	1351	-	4,4,4	0.41	0	6,6,6	0.43	0
2	SO4	C	1347	-	4,4,4	0.38	0	6,6,6	0.26	0
3	GOL	A	1350	-	5,5,5	0.43	0	5,5,5	0.27	0
3	GOL	A	1353	-	5,5,5	0.25	0	5,5,5	0.48	0
2	SO4	A	1345	-	4,4,4	0.53	0	6,6,6	0.54	0
2	SO4	B	1353	-	4,4,4	0.29	0	6,6,6	0.16	0
2	SO4	A	1344	-	4,4,4	0.41	0	6,6,6	0.27	0
2	SO4	B	1350	-	4,4,4	0.33	0	6,6,6	0.57	0
2	SO4	C	1349	-	4,4,4	0.33	0	6,6,6	0.25	0
2	SO4	A	1343	-	4,4,4	0.32	0	6,6,6	0.41	0
2	SO4	C	1348	-	4,4,4	0.33	0	6,6,6	0.25	0
3	GOL	A	1352	-	5,5,5	0.38	0	5,5,5	0.37	0
2	SO4	C	1350	-	4,4,4	0.49	0	6,6,6	0.18	0
2	SO4	A	1346	-	4,4,4	0.34	0	6,6,6	0.08	0
2	SO4	A	1356	-	4,4,4	0.42	0	6,6,6	0.52	0
3	GOL	B	1343	-	5,5,5	0.39	0	5,5,5	1.22	1 (20%)
2	SO4	B	1349	-	4,4,4	0.35	0	6,6,6	0.14	0
3	GOL	B	1347	-	5,5,5	0.45	0	5,5,5	0.45	0
3	GOL	A	1349	-	5,5,5	0.43	0	5,5,5	0.76	0
3	GOL	B	1346	-	5,5,5	0.34	0	5,5,5	0.27	0
2	SO4	B	1345	-	4,4,4	0.30	0	6,6,6	0.57	0
2	SO4	A	1355	-	4,4,4	0.32	0	6,6,6	0.17	0
2	SO4	A	1348	-	4,4,4	0.33	0	6,6,6	0.51	0
2	SO4	C	1346	-	4,4,4	0.35	0	6,6,6	0.20	0
2	SO4	A	1347	-	4,4,4	0.32	0	6,6,6	0.25	0
3	GOL	C	1345	-	5,5,5	0.36	0	5,5,5	0.27	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	GOL	A	1351	-	5,5,5	0.55	0	5,5,5	0.27	0
2	SO4	B	1352	-	4,4,4	0.61	0	6,6,6	0.38	0
3	GOL	B	1348	-	5,5,5	0.57	0	5,5,5	0.99	0
2	SO4	A	1354	-	4,4,4	0.39	0	6,6,6	0.59	0
3	GOL	C	1344	-	5,5,5	0.59	0	5,5,5	0.55	0
2	SO4	C	1343	-	4,4,4	0.39	0	6,6,6	0.45	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. '2' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	GOL	A	1351	-	-	2/4/4/4	-
3	GOL	A	1350	-	-	4/4/4/4	-
3	GOL	B	1348	-	-	2/4/4/4	-
3	GOL	A	1353	-	-	2/4/4/4	-
3	GOL	A	1352	-	-	2/4/4/4	-
3	GOL	B	1343	-	-	0/4/4/4	-
3	GOL	B	1346	-	-	2/4/4/4	-
3	GOL	C	1344	-	-	2/4/4/4	-
3	GOL	A	1349	-	-	4/4/4/4	-
3	GOL	B	1347	-	-	4/4/4/4	-
3	GOL	C	1345	-	-	2/4/4/4	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	1343	GOL	O2-C2-C3	-2.38	98.62	109.12

There are no chirality outliers.

All (26) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1349	GOL	O1-C1-C2-C3
3	A	1350	GOL	C1-C2-C3-O3
3	A	1351	GOL	O1-C1-C2-C3
3	A	1352	GOL	O1-C1-C2-C3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
3	A	1353	GOL	O1-C1-C2-C3
3	B	1347	GOL	O1-C1-C2-C3
3	B	1347	GOL	C1-C2-C3-O3
3	B	1347	GOL	O2-C2-C3-O3
3	A	1349	GOL	O2-C2-C3-O3
3	A	1349	GOL	C1-C2-C3-O3
3	B	1346	GOL	C1-C2-C3-O3
3	C	1344	GOL	O1-C1-C2-C3
3	A	1349	GOL	O1-C1-C2-O2
3	A	1350	GOL	O2-C2-C3-O3
3	A	1351	GOL	O1-C1-C2-O2
3	B	1347	GOL	O1-C1-C2-O2
3	C	1344	GOL	O1-C1-C2-O2
3	A	1352	GOL	O1-C1-C2-O2
3	A	1353	GOL	O1-C1-C2-O2
3	B	1346	GOL	O2-C2-C3-O3
3	A	1350	GOL	O1-C1-C2-O2
3	C	1345	GOL	O2-C2-C3-O3
3	B	1348	GOL	O1-C1-C2-C3
3	C	1345	GOL	C1-C2-C3-O3
3	B	1348	GOL	O1-C1-C2-O2
3	A	1350	GOL	O1-C1-C2-C3

There are no ring outliers.

10 monomers are involved in 17 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	1350	SO4	1	0
2	A	1346	SO4	1	0
2	A	1356	SO4	1	0
3	B	1346	GOL	6	0
2	B	1345	SO4	1	0
2	A	1348	SO4	2	0
2	C	1346	SO4	1	0
3	C	1345	GOL	1	0
3	A	1351	GOL	2	0
2	C	1343	SO4	1	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	301/311 (96%)	0.41	20 (6%) 18 24	21, 35, 70, 94	0
1	B	307/311 (98%)	0.15	11 (3%) 42 50	17, 29, 61, 84	0
1	C	307/311 (98%)	0.21	12 (3%) 39 47	18, 29, 60, 86	0
All	All	915/933 (98%)	0.25	43 (4%) 31 39	17, 31, 65, 94	0

All (43) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	275	LEU	6.8
1	A	162	ASP	4.9
1	C	310	PRO	4.7
1	B	310	PRO	4.6
1	C	118[A]	TYR	4.5
1	B	118[A]	TYR	4.4
1	A	272	GLN	4.2
1	A	118[A]	TYR	4.2
1	A	270	ASN	3.7
1	A	342	LYS	3.7
1	C	312	PHE	3.6
1	A	307	TRP	3.5
1	B	185	LEU	3.5
1	C	185	LEU	3.4
1	A	301	GLU	3.3
1	A	185	LEU	3.3
1	A	183	VAL	3.3
1	A	297	ASP	3.2
1	A	186	PRO	3.2
1	C	162	ASP	3.2
1	B	36	GLU	3.1
1	A	277	VAL	3.0
1	B	312	PHE	2.9

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	311	GLU	2.8
1	C	142[A]	LYS	2.8
1	A	271	HIS	2.8
1	C	183	VAL	2.8
1	A	142[A]	LYS	2.6
1	B	184	HIS	2.6
1	A	276	VAL	2.5
1	B	309	LEU	2.5
1	C	301	GLU	2.5
1	B	142[A]	LYS	2.5
1	C	186	PRO	2.3
1	B	186	PRO	2.3
1	A	184	HIS	2.3
1	C	276	VAL	2.2
1	A	314	ASN	2.2
1	A	303	PRO	2.1
1	B	146[A]	TRP	2.1
1	C	314	ASN	2.1
1	A	36	GLU	2.1
1	B	164	LYS	2.1

## 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( $\text{\AA}^2$ )	Q<0.9
3	GOL	C	1344	6/6	0.67	0.21	55,56,59,60	0
2	SO4	A	1345	5/5	0.70	0.33	41,42,45,47	5
3	GOL	B	1347	6/6	0.71	0.19	49,57,60,61	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	SO4	B	1351	5/5	0.74	0.25	35,40,45,45	5
2	SO4	C	1350	5/5	0.75	0.24	44,45,45,51	5
2	SO4	A	1343	5/5	0.77	0.19	45,49,51,53	5
3	GOL	A	1350	6/6	0.78	0.18	55,67,73,73	0
2	SO4	A	1356	5/5	0.79	0.36	39,39,45,47	5
2	SO4	B	1353	5/5	0.79	0.22	55,63,63,64	5
3	GOL	A	1352	6/6	0.80	0.18	58,61,65,65	0
3	GOL	A	1353	6/6	0.81	0.20	60,63,64,64	0
3	GOL	B	1348	6/6	0.83	0.20	36,43,50,54	0
2	SO4	A	1344	5/5	0.84	0.19	43,46,50,52	5
2	SO4	A	1347	5/5	0.85	0.23	40,46,49,51	5
3	GOL	A	1351	6/6	0.86	0.13	49,54,55,60	0
2	SO4	B	1352	5/5	0.86	0.20	40,42,56,57	5
3	GOL	B	1346	6/6	0.87	0.39	52,53,53,62	0
2	SO4	A	1355	5/5	0.87	0.15	48,50,56,58	5
2	SO4	B	1344	5/5	0.90	0.17	45,46,48,49	5
3	GOL	C	1345	6/6	0.90	0.15	38,41,41,44	6
2	SO4	A	1346	5/5	0.91	0.36	50,51,54,56	5
2	SO4	C	1347	5/5	0.91	0.14	64,66,75,76	0
2	SO4	B	1350	5/5	0.91	0.19	54,66,68,73	0
3	GOL	A	1349	6/6	0.92	0.34	26,27,28,31	6
2	SO4	C	1349	5/5	0.92	0.19	71,78,79,82	0
2	SO4	A	1348	5/5	0.93	0.15	38,40,42,48	5
2	SO4	C	1346	5/5	0.93	0.24	62,66,72,74	0
2	SO4	A	1354	5/5	0.93	0.13	62,63,68,68	0
2	SO4	C	1348	5/5	0.93	0.18	57,73,80,86	0
3	GOL	B	1343	6/6	0.94	0.22	34,42,48,51	0
2	SO4	B	1345	5/5	0.94	0.15	31,38,39,43	5
2	SO4	C	1343	5/5	0.94	0.14	40,42,42,47	5
2	SO4	B	1349	5/5	0.95	0.20	59,64,71,74	0

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

There are no such residues in this entry.