



# Full wwPDB X-ray Structure Validation Report ⓘ

Jan 24, 2017 – 05:18 PM EST

PDB ID : 5TJ2  
Title : Gasdermin-B C-terminal domain containing the polymorphism residues  
Gly299:Ser306 fused to maltose binding protein  
Authors : Chao, L.K.; Herzberg, O.  
Deposited on : 2016-10-03  
Resolution : 2.80 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-20028442
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-20028442

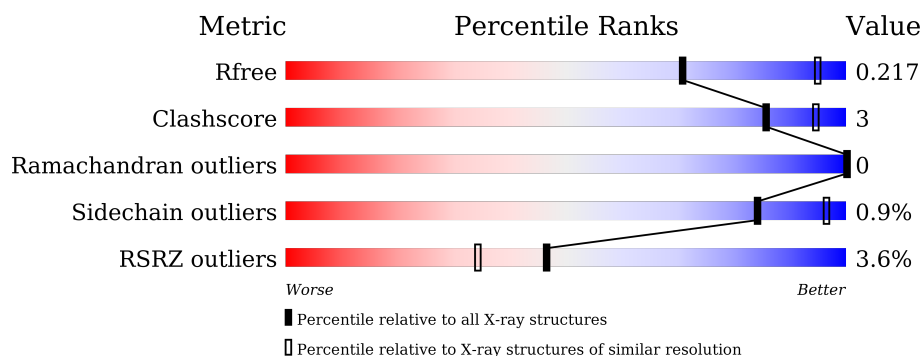
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.80 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	2393 (2.80-2.80)
Clashscore	102246	2827 (2.80-2.80)
Ramachandran outliers	100387	2782 (2.80-2.80)
Sidechain outliers	100360	2784 (2.80-2.80)
RSRZ outliers	91569	2404 (2.80-2.80)

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	553	<div> <div>93%</div> <div> <div></div> <div></div> <div></div> <div></div> </div> <div>• •</div> </div>
1	B	553	<div> <div>91%</div> <div> <div></div> <div></div> <div></div> <div></div> </div> <div>6%</div> <div>•</div> </div>
1	C	553	<div> <div>%</div> <div> <div></div> <div></div> <div></div> <div></div> </div> <div>88%</div> <div>7%</div> <div>•</div> </div>
1	D	553	<div> <div>12%</div> <div> <div></div> <div></div> <div></div> <div></div> </div> <div>91%</div> <div>5%</div> <div>5%</div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 15875 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 Sugar ABC transporter substrate-binding protein, Gasdermin-B fusion protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	534	Total	C	N	O	S	0	0	0
			4029	2579	659	778	13			
1	B	536	Total	C	N	O	S	0	0	0
			4053	2598	663	779	13			
1	C	529	Total	C	N	O	S	0	0	0
			3854	2480	624	738	12			
1	D	528	Total	C	N	O	S	0	0	0
			3810	2447	623	728	12			

There are 64 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	82	ALA	ASP	conflict	UNP A0A178SBV6
A	83	ALA	LYS	conflict	UNP A0A178SBV6
A	172	ALA	GLU	conflict	UNP A0A178SBV6
A	173	ALA	ASN	conflict	UNP A0A178SBV6
A	239	ALA	LYS	conflict	UNP A0A178SBV6
A	362	ALA	LYS	conflict	UNP A0A178SBV6
A	363	ALA	ASP	conflict	UNP A0A178SBV6
A	367	ASN	-	linker	UNP A0A178SBV6
A	368	ALA	-	linker	UNP A0A178SBV6
A	369	ALA	-	linker	UNP A0A178SBV6
A	370	ALA	-	linker	UNP A0A178SBV6
A	1306	SER	PRO	engineered mutation	UNP Q8TAX9
A	?	-	ASP	deletion	UNP Q8TAX9
A	?	-	MET	deletion	UNP Q8TAX9
A	?	-	ASP	deletion	UNP Q8TAX9
A	?	-	TYR	deletion	UNP Q8TAX9
B	82	ALA	ASP	conflict	UNP A0A178SBV6
B	83	ALA	LYS	conflict	UNP A0A178SBV6
B	172	ALA	GLU	conflict	UNP A0A178SBV6
B	173	ALA	ASN	conflict	UNP A0A178SBV6

*Continued on next page...*

*Continued from previous page...*

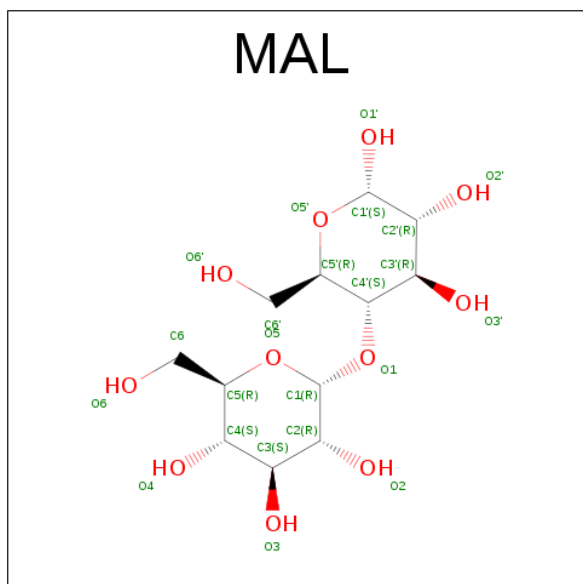
Chain	Residue	Modelled	Actual	Comment	Reference
B	239	ALA	LYS	conflict	UNP A0A178SBV6
B	362	ALA	LYS	conflict	UNP A0A178SBV6
B	363	ALA	ASP	conflict	UNP A0A178SBV6
B	367	ASN	-	linker	UNP A0A178SBV6
B	368	ALA	-	linker	UNP A0A178SBV6
B	369	ALA	-	linker	UNP A0A178SBV6
B	370	ALA	-	linker	UNP A0A178SBV6
B	1306	SER	PRO	engineered mutation	UNP Q8TAX9
B	?	-	ASP	deletion	UNP Q8TAX9
B	?	-	MET	deletion	UNP Q8TAX9
B	?	-	ASP	deletion	UNP Q8TAX9
B	?	-	TYR	deletion	UNP Q8TAX9
C	82	ALA	ASP	conflict	UNP A0A178SBV6
C	83	ALA	LYS	conflict	UNP A0A178SBV6
C	172	ALA	GLU	conflict	UNP A0A178SBV6
C	173	ALA	ASN	conflict	UNP A0A178SBV6
C	239	ALA	LYS	conflict	UNP A0A178SBV6
C	362	ALA	LYS	conflict	UNP A0A178SBV6
C	363	ALA	ASP	conflict	UNP A0A178SBV6
C	367	ASN	-	linker	UNP A0A178SBV6
C	368	ALA	-	linker	UNP A0A178SBV6
C	369	ALA	-	linker	UNP A0A178SBV6
C	370	ALA	-	linker	UNP A0A178SBV6
C	1306	SER	PRO	engineered mutation	UNP Q8TAX9
C	?	-	ASP	deletion	UNP Q8TAX9
C	?	-	MET	deletion	UNP Q8TAX9
C	?	-	ASP	deletion	UNP Q8TAX9
C	?	-	TYR	deletion	UNP Q8TAX9
D	82	ALA	ASP	conflict	UNP A0A178SBV6
D	83	ALA	LYS	conflict	UNP A0A178SBV6
D	172	ALA	GLU	conflict	UNP A0A178SBV6
D	173	ALA	ASN	conflict	UNP A0A178SBV6
D	239	ALA	LYS	conflict	UNP A0A178SBV6
D	362	ALA	LYS	conflict	UNP A0A178SBV6
D	363	ALA	ASP	conflict	UNP A0A178SBV6
D	367	ASN	-	linker	UNP A0A178SBV6
D	368	ALA	-	linker	UNP A0A178SBV6
D	369	ALA	-	linker	UNP A0A178SBV6
D	370	ALA	-	linker	UNP A0A178SBV6
D	1306	SER	PRO	engineered mutation	UNP Q8TAX9
D	?	-	ASP	deletion	UNP Q8TAX9
D	?	-	MET	deletion	UNP Q8TAX9

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	?	-	ASP	deletion	UNP Q8TAX9
D	?	-	TYR	deletion	UNP Q8TAX9

- Molecule 2 is MALTOSE (three-letter code: MAL) (formula:  $C_{12}H_{22}O_{11}$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			23	12	11		
2	B	1	Total	C	O	0	0
			23	12	11		
2	C	1	Total	C	O	0	0
			23	12	11		
2	D	1	Total	C	O	0	0
			23	12	11		

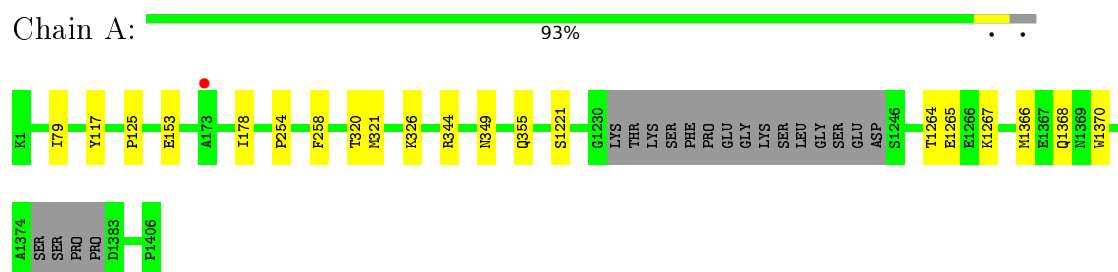
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	21	Total	O	0	0
			21	21		
3	B	12	Total	O	0	0
			12	12		
3	C	3	Total	O	0	0
			3	3		
3	D	1	Total	O	0	0
			1	1		

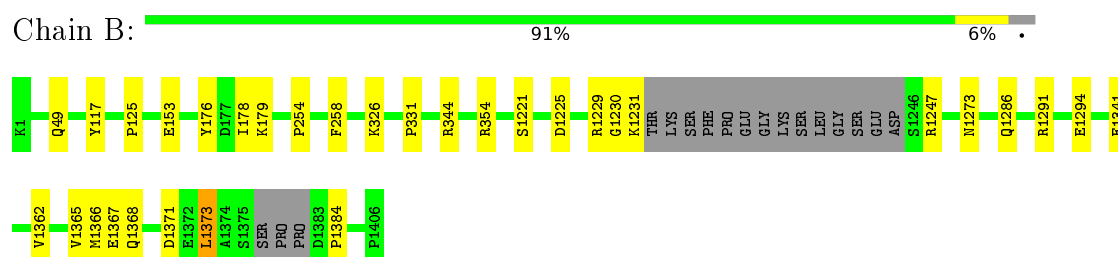
### 3 Residue-property plots

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.

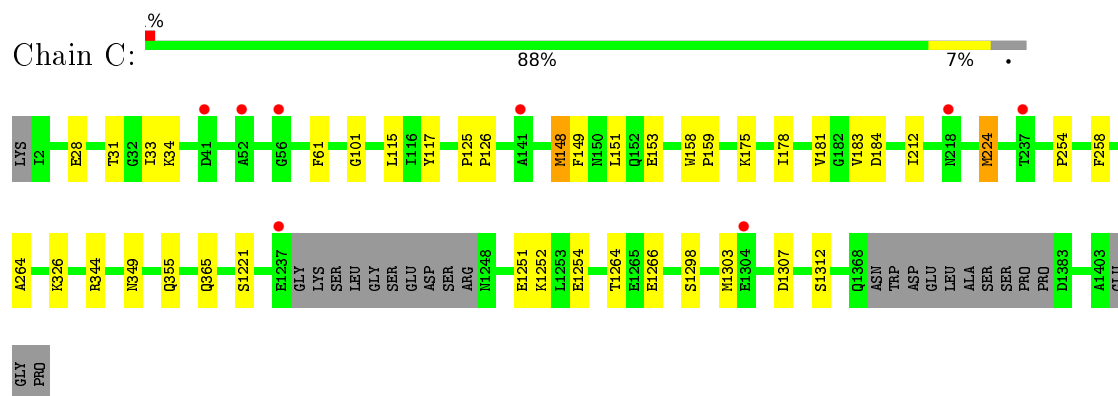
- Molecule 1: Sugar ABC transporter substrate-binding protein,Gasdermin-B fusion protein



- Molecule 1: Sugar ABC transporter substrate-binding protein,Gasdermin-B fusion protein

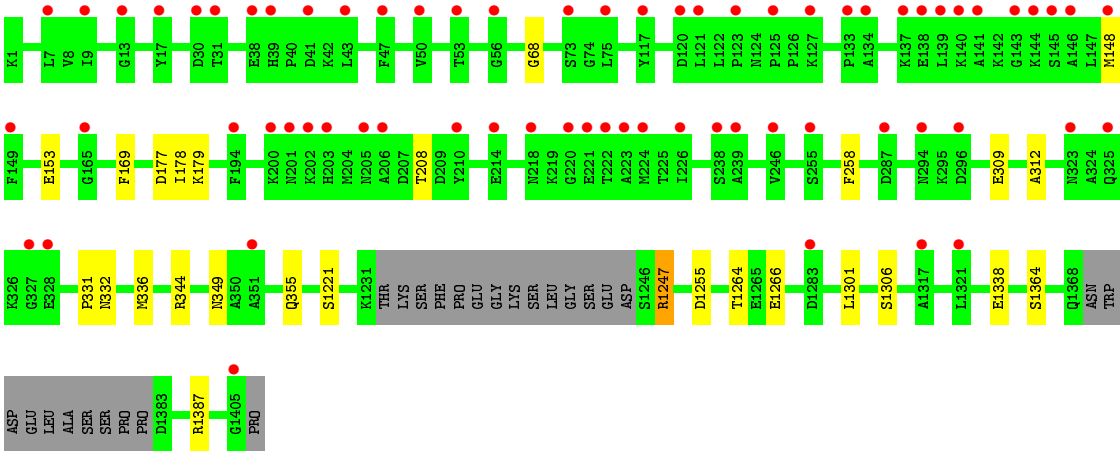


- Molecule 1: Sugar ABC transporter substrate-binding protein,Gasdermin-B fusion protein



- Molecule 1: Sugar ABC transporter substrate-binding protein,Gasdermin-B fusion protein





## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	144.04Å 152.75Å 255.77Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	42.50 – 2.80 42.50 – 2.50	Depositor EDS
% Data completeness (in resolution range)	99.6 (42.50-2.80) 99.4 (42.50-2.50)	Depositor EDS
$R_{merge}$	0.15	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.41 (at 2.51Å)	Xtriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.213 , 0.252 0.216 , 0.217	Depositor DCC
$R_{free}$ test set	3433 reflections (5.21%)	DCC
Wilson B-factor (Å <sup>2</sup> )	40.6	Xtriage
Anisotropy	0.034	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 23.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	15875	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	67.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.23% 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: MAL

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	0.62	0/4109	0.71	0/5588
1	B	0.61	0/4133	0.71	2/5616 (0.0%)
1	C	0.55	0/3933	0.70	3/5372 (0.1%)
1	D	0.49	0/3887	0.68	2/5316 (0.0%)
All	All	0.57	0/16062	0.70	7/21892 (0.0%)

There are no bond length outliers.

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	1229	ARG	NE-CZ-NH1	7.04	123.82	120.30
1	C	1303	MET	CG-SD-CE	6.93	111.29	100.20
1	C	224	MET	CG-SD-CE	-6.70	89.49	100.20
1	D	1247	ARG	NE-CZ-NH1	6.51	123.56	120.30
1	D	1301	LEU	N-CA-C	-6.06	94.63	111.00
1	C	148	MET	CG-SD-CE	5.93	109.69	100.20
1	B	354	ARG	NE-CZ-NH1	5.33	122.96	120.30

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	4029	0	3931	11	0
1	B	4053	0	3977	29	0
1	C	3854	0	3623	31	0
1	D	3810	0	3555	23	0
2	A	23	0	22	0	0
2	B	23	0	22	0	0
2	C	23	0	22	0	0
2	D	23	0	22	0	0
3	A	21	0	0	0	0
3	B	12	0	0	0	0
3	C	3	0	0	0	0
3	D	1	0	0	0	0
All	All	15875	0	15174	88	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 3.

All (88) 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:1231:LYS:HA	1:B:1231:LYS:HE2	1.50	0.91
1:C:184:ASP:HB3	1:C:365:GLN:OE1	1.72	0.89
1:C:126:PRO:HD2	1:C:224:MET:CE	2.06	0.86
1:D:1247:ARG:NH2	1:D:1255:ASP:OD2	2.15	0.78
1:C:126:PRO:HD2	1:C:224:MET:HE1	1.68	0.75
1:B:1291:ARG:NH1	1:B:1294:GLU:OE1	2.22	0.72
1:B:1231:LYS:CA	1:B:1231:LYS:HE2	2.16	0.71
1:B:1230:GLY:O	1:B:1231:LYS:HG2	1.91	0.71
1:A:320:THR:HG22	1:A:321:MET:CE	2.23	0.69
1:D:331:PRO:HG2	1:D:336:MET:HE3	1.75	0.68
1:B:1384:PRO:HB3	1:C:101:GLY:HA3	1.77	0.67
1:B:1231:LYS:CE	1:B:1231:LYS:HA	2.23	0.67
1:C:178:ILE:HG23	1:C:1221:SER:HB3	1.77	0.66
1:A:178:ILE:HG23	1:A:1221:SER:HB3	1.79	0.64
1:D:169:PHE:CE2	1:D:336:MET:HE2	2.32	0.64
1:C:126:PRO:HD2	1:C:224:MET:HE3	1.79	0.64
1:C:1251:GLU:HA	1:C:1254:GLU:HB2	1.80	0.64
1:D:178:ILE:HG23	1:D:1221:SER:HB3	1.80	0.62
1:B:178:ILE:HG23	1:B:1221:SER:HB3	1.81	0.61
1:D:331:PRO:HB2	1:D:336:MET:HE3	1.82	0.61
1:D:169:PHE:HE2	1:D:336:MET:HE2	1.67	0.60
1:A:320:THR:HG22	1:A:321:MET:HE2	1.83	0.59

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:1362:VAL:O	1:B:1365:VAL:CG2	2.52	0.57
1:D:331:PRO:CG	1:D:336:MET:HE3	2.34	0.57
1:C:1307:ASP:HA	1:D:312:ALA:HB1	1.87	0.57
1:B:1362:VAL:O	1:B:1365:VAL:HG22	2.06	0.55
1:D:169:PHE:CD2	1:D:336:MET:CE	2.90	0.55
1:C:31:THR:CB	1:C:33:ILE:HD12	2.36	0.55
1:C:148:MET:HE3	1:C:212:ILE:HG22	1.90	0.54
1:C:148:MET:CE	1:C:212:ILE:HG22	2.38	0.54
1:B:1373:LEU:HD23	1:B:1373:LEU:O	2.08	0.54
1:D:169:PHE:CD2	1:D:336:MET:HE1	2.44	0.53
1:D:331:PRO:CB	1:D:336:MET:HE3	2.39	0.53
1:A:1366:MET:O	1:A:1368:GLN:N	2.43	0.52
1:B:1366:MET:HE3	1:B:1373:LEU:HD13	1.92	0.52
1:D:169:PHE:CE2	1:D:336:MET:CE	2.93	0.51
1:C:115:LEU:HD21	1:C:224:MET:CE	2.41	0.50
1:D:177:ASP:OD1	1:D:179:LYS:HB2	2.11	0.50
1:A:1366:MET:HE1	1:A:1370:TRP:HE3	1.75	0.50
1:C:254:PRO:HB3	1:C:326:LYS:HD3	1.94	0.49
1:B:1225:ASP:OD2	1:C:175:LYS:HE2	2.12	0.49
1:B:1366:MET:O	1:B:1368:GLN:N	2.46	0.48
1:D:148:MET:SD	1:D:208:THR:HG21	2.54	0.48
1:B:179:LYS:NZ	1:B:1341:GLU:OE2	2.46	0.47
1:A:254:PRO:HB3	1:A:326:LYS:HD3	1.96	0.47
1:B:254:PRO:HB3	1:B:326:LYS:HD3	1.98	0.46
1:B:1362:VAL:HA	1:B:1365:VAL:HG22	1.97	0.46
1:B:1362:VAL:HA	1:B:1365:VAL:CG2	2.46	0.46
1:C:1312:SER:CB	1:D:309:GLU:HG3	2.46	0.46
1:C:184:ASP:HB3	1:C:365:GLN:CD	2.34	0.46
1:C:181:VAL:HG12	1:C:183:VAL:HG22	1.98	0.45
1:B:179:LYS:CE	1:B:1341:GLU:OE2	2.65	0.45
1:C:149:PHE:CE1	1:C:151:LEU:HD23	2.51	0.45
1:B:176:TYR:CE1	1:B:331:PRO:HG3	2.51	0.45
1:C:28:GLU:HG3	1:C:34:LYS:HA	1.99	0.45
1:C:126:PRO:CD	1:C:224:MET:HE1	2.44	0.44
1:C:115:LEU:HD21	1:C:224:MET:HE1	1.98	0.44
1:B:1366:MET:CE	1:B:1373:LEU:HD13	2.47	0.44
1:B:153:GLU:OE1	1:B:344:ARG:NH1	2.52	0.43
1:D:1338:GLU:OE1	1:D:1387:ARG:HB3	2.17	0.43
1:C:254:PRO:CB	1:C:326:LYS:HD3	2.47	0.43
1:C:61:PHE:CE2	1:C:264:ALA:HB2	2.54	0.43
1:C:1264:THR:HG22	1:C:1266:GLU:H	1.83	0.43

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:158:TRP:N	1:C:159:PRO:CD	2.80	0.43
1:A:153:GLU:OE1	1:A:344:ARG:NH1	2.50	0.42
1:D:153:GLU:OE1	1:D:344:ARG:NH1	2.53	0.42
1:D:169:PHE:HD2	1:D:336:MET:CE	2.30	0.42
1:A:254:PRO:CB	1:A:326:LYS:HD3	2.50	0.42
1:C:1251:GLU:O	1:C:1252:LYS:C	2.57	0.42
1:D:169:PHE:HD2	1:D:336:MET:HE1	1.82	0.42
1:B:1362:VAL:O	1:B:1365:VAL:HG23	2.18	0.42
1:C:153:GLU:OE1	1:C:344:ARG:NH1	2.52	0.42
1:B:1230:GLY:C	1:B:1231:LYS:HG2	2.40	0.42
1:B:1366:MET:HB3	1:B:1373:LEU:HD13	2.01	0.42
1:C:349:ASN:HB3	1:C:355:GLN:HB2	2.02	0.42
1:A:349:ASN:HB3	1:A:355:GLN:HB2	2.01	0.41
1:B:117:TYR:CZ	1:B:125:PRO:HG3	2.55	0.41
1:B:1384:PRO:CB	1:C:101:GLY:HA3	2.48	0.41
1:B:254:PRO:CB	1:B:326:LYS:HD3	2.51	0.41
1:A:1264:THR:HG22	1:A:1267:LYS:HD2	2.02	0.41
1:D:1264:THR:HG22	1:D:1266:GLU:H	1.86	0.41
1:D:68:GLY:HA3	1:D:332:ASN:O	2.21	0.41
1:C:117:TYR:CZ	1:C:125:PRO:HG3	2.56	0.41
1:B:49:GLN:HG3	1:B:1367:GLU:O	2.21	0.40
1:C:1312:SER:HB2	1:D:309:GLU:HG3	2.03	0.40
1:A:117:TYR:CZ	1:A:125:PRO:HG3	2.57	0.40
1:D:349:ASN:HB3	1:D:355:GLN:HB2	2.03	0.40
1:B:1362:VAL:C	1:B:1365:VAL:HG22	2.41	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	528/553 (96%)	513 (97%)	15 (3%)	0	100	100

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	530/553 (96%)	516 (97%)	14 (3%)	0	100	100
1	C	523/553 (95%)	509 (97%)	14 (3%)	0	100	100
1	D	522/553 (94%)	508 (97%)	14 (3%)	0	100	100
All	All	2103/2212 (95%)	2046 (97%)	57 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 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	404/446 (91%)	401 (99%)	3 (1%)	88	97
1	B	407/446 (91%)	401 (98%)	6 (2%)	72	93
1	C	359/446 (80%)	357 (99%)	2 (1%)	90	98
1	D	349/446 (78%)	346 (99%)	3 (1%)	84	96
All	All	1519/1784 (85%)	1505 (99%)	14 (1%)	84	96

All (14) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	79	ILE
1	A	258	PHE
1	A	1265	GLU
1	B	258	PHE
1	B	1247	ARG
1	B	1273	ASN
1	B	1286	GLN
1	B	1371	ASP
1	B	1373	LEU
1	C	258	PHE
1	C	1298	SER
1	D	258	PHE
1	D	1306	SER

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	D	1364	SER

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

Mol	Chain	Res	Type
1	C	201	ASN

### 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	MAL	A	1501	-	24,24,24	0.64	0	35,35,35	1.35	6 (17%)
2	MAL	B	1501	-	24,24,24	0.61	0	35,35,35	1.08	1 (2%)
2	MAL	C	1501	-	24,24,24	0.69	0	35,35,35	1.00	2 (5%)
2	MAL	D	1501	-	24,24,24	0.72	0	35,35,35	1.36	5 (14%)

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	MAL	A	1501	-	-	0/8/48/48	0/2/2/2
2	MAL	B	1501	-	-	0/8/48/48	0/2/2/2
2	MAL	C	1501	-	-	0/8/48/48	0/2/2/2
2	MAL	D	1501	-	-	0/8/48/48	0/2/2/2

There are no bond length outliers.

All (14) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1501	MAL	O1-C1-O5	-2.81	103.37	110.69
2	A	1501	MAL	O2'-C2'-C3'	-2.65	104.38	110.36
2	D	1501	MAL	O1-C4'-C5'	-2.57	102.48	109.33
2	A	1501	MAL	O3'-C3'-C2'	-2.14	105.54	110.36
2	C	1501	MAL	C2'-C3'-C4'	2.16	114.40	109.63
2	D	1501	MAL	C3-C4-C5	2.17	114.10	110.23
2	A	1501	MAL	O4-C4-C3	2.23	115.38	110.36
2	D	1501	MAL	C1'-C2'-C3'	2.27	114.41	110.68
2	A	1501	MAL	C1-O5-C5	2.44	118.53	113.74
2	C	1501	MAL	C1-O5-C5	2.53	118.70	113.74
2	A	1501	MAL	C1'-C2'-C3'	2.76	115.21	110.68
2	B	1501	MAL	C1-O5-C5	2.77	119.17	113.74
2	D	1501	MAL	C3'-C4'-C5'	2.79	117.22	110.85
2	D	1501	MAL	C2'-C3'-C4'	3.33	116.99	109.63

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	534/553 (96%)	-0.63	1 (0%) 95 94	23, 39, 71, 120	0
1	B	536/553 (96%)	-0.63	0 100 100	22, 40, 72, 124	0
1	C	529/553 (95%)	-0.29	8 (1%) 76 68	33, 74, 112, 133	0
1	D	528/553 (95%)	0.51	68 (12%) 5 2	41, 107, 158, 215	0
All	All	2127/2212 (96%)	-0.26	77 (3%) 46 34	22, 58, 132, 215	0

All (77) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	220	GLY	10.5
1	D	41	ASP	8.9
1	D	221	GLU	8.3
1	D	141	ALA	8.1
1	D	121	LEU	7.4
1	D	351	ALA	7.3
1	D	38	GLU	6.8
1	D	214	GLU	6.7
1	D	30	ASP	5.5
1	D	145	SER	5.1
1	D	75	LEU	5.1
1	D	120	ASP	4.7
1	D	201	ASN	4.7
1	D	203	HIS	4.6
1	D	218	ASN	4.3
1	C	52	ALA	4.1
1	D	223	ALA	4.1
1	C	141	ALA	4.0
1	D	13	GLY	3.9
1	D	200	LYS	3.8
1	D	222	THR	3.7

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	123	PRO	3.7
1	D	17	TYR	3.6
1	D	125	PRO	3.5
1	D	134	ALA	3.5
1	D	56	GLY	3.4
1	D	255	SER	3.3
1	D	206	ALA	3.2
1	D	137	LYS	3.2
1	D	138	GLU	3.2
1	D	246	VAL	3.1
1	D	140	LYS	3.0
1	D	127	LYS	3.0
1	D	224	MET	3.0
1	D	1317	ALA	3.0
1	A	173	ALA	2.9
1	D	202	LYS	2.9
1	D	39	HIS	2.9
1	D	323	ASN	2.8
1	C	41	ASP	2.8
1	D	205	ASN	2.8
1	D	43	LEU	2.8
1	D	146	ALA	2.8
1	D	210	TYR	2.7
1	D	239	ALA	2.7
1	D	149	PHE	2.7
1	D	238	SER	2.7
1	D	165	GLY	2.7
1	D	143	GLY	2.7
1	D	327	GLY	2.6
1	D	287	ASP	2.6
1	D	7	LEU	2.6
1	D	296	ASP	2.6
1	D	139	LEU	2.6
1	D	1321	LEU	2.5
1	D	148	MET	2.5
1	D	50	VAL	2.5
1	D	117	TYR	2.5
1	D	144	LYS	2.4
1	D	294	ASN	2.4
1	D	1283	ASP	2.4
1	D	133	PRO	2.4
1	C	1237	GLU	2.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	325	GLN	2.3
1	C	1304	GLU	2.3
1	D	73	SER	2.2
1	D	194	PHE	2.2
1	D	328	GLU	2.2
1	D	53	THR	2.2
1	C	218	ASN	2.2
1	D	31	THR	2.1
1	D	47	PHE	2.1
1	D	9	ILE	2.1
1	C	56	GLY	2.1
1	D	1405	GLY	2.1
1	C	237	THR	2.1
1	D	226	ILE	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( $\text{\AA}^2$ )	Q<0.9
2	MAL	C	1501	23/23	0.97	0.13	-0.52	43,53,74,79	0
2	MAL	A	1501	23/23	0.97	0.12	-0.98	28,32,38,43	0
2	MAL	B	1501	23/23	0.97	0.11	-1.21	24,29,35,36	0
2	MAL	D	1501	23/23	0.86	0.13	-1.22	84,104,120,128	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.